Giulietta

The mathematics book which wants to be collaborative

http://laurent.claessens-donadello.eu/pdf/giulietta.pdf

Laurent Claessens

October 9, 2024

6c499d4c9c4d292b7cf7b39351b1a4e701b0162f

Copyright 2011-2024
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the chapter entitled "GNU Free Documentation License".

This file contains everything I know in math and some more by other contributors.
Subdivided in two main parts

Le Frido in French contains the mathematics for the first few years in mathematics: from the construction of the rationals to the Markov chains trough Lebesgue measure, representation theory and fields extensions.

Giulietta in English contains whatever mathematics at higher level.

Every contributions are most welcome.

Rolling release The last version can be downloaded at

http://laurent.claessens-donadello.eu/pdf/giulietta.pdf

The French part The French part is a book in itself which can be downloaded here:

http://laurent.claessens-donadello.eu/pdf/lefrido.pdf

If you want to recompile the book at home, see the instructions

https://github.com/LaurentClaessens/mazhe/blob/master/COMPILATION_giulietta.md
https://github.com/LaurentClaessens/mazhe

Index thématique

1 : cardinalité
2 : morphismes et compagnie
3 : arithmétique modulo, théorème de Bézout
4 : divisibilité
5 : équations diophantiennes
6 : types d'anneaux
7 : sous-groupes
8 : groupe symétrique
9 : action de groupe
10 : classification de groupes
11 : fonction indicatrice d'Euler
12 : produit semi-direct de groupes
13 : théorie des représentations
14 : espaces vectoriels
15 : rang
16 : formes bilinéaires et quadratiques
17 : endomorphismes cycliques
18 : extension de corps et polynômes
19 : polynômes
20 : polynôme d'endomorphismes
21 : dualité
22 : injections
23 : tribu, algèbre de parties, λ-systèmes et co.
24 : théorie de la mesure
25 : normes
26 : inégalités
27 : constructions topologiques
28 : espaces métriques, normés
29 : limite et continuité
30 : intégration
31 : connexité
32 : séparation
33 : compacts
34 : densité
35 : application réciproque
36 : applications continues et bornées
37 : suite de Cauchy, espace complet
38 : caractérisations séquentielles
39 : valeurs propres, définie positive
40 : norme matricielle, norme opérateur et rayon spectral
41 : série de matrices
42 : décomposition de matrices
43 : systèmes d'équations linéaires
44 : réduction, diagonalisation
45 : déterminant
46 : espaces de fonctions
47 : fonctions Lipschitz
48 : suites et séries
49 : suite de fonctions
50 : exponentielle
51 : logarithme
52 : fonction puissance
53 : sommation finie et infinie
54 : polynôme de Taylor
55 : formule des accroissements finis
56 : dérivation
57 : différentielle
58 : équations différentielles
59 : convexité
60 : espaces de Hilbert, base hilbertienne
61 : analyse complexe, fonctions holomorphes
62 : permuter des limites
63 : déduire la nullité d'une fonction depuis son intégrale
64 : inversion locale, fonction implicite
65 : points fixes
66 : changement de variables
67 : techniques de calcul
68 : méthodes de calcul
69 : méthode de Newton
70 : prolongement d'applications
71 : opérations sur les distributions
72 : convolution
73 : séries de Fourier
74 : transformée de Fourier
75 : gaussienne
76 : lemme de transfert
77 : invariants de similitude
78 : isométries
79 : enveloppes
80 : intégration sur des variétés
81 : dénombrements
82 : caractérisation de distributions en probabilités
83 : théorème central limite
84 : indépendance d'événements et de variables aléatoires
85 : probabilités et espérances conditionnelles
86 : chaine de Markov
87 : manifolds
88 : Immersions and Embeddings
89 : Lie groups
90 : $\mathrm{GL}(n, \mathbb{C})$
91 : $\mathrm{SU}(n)$
92 : flow of a vector field
93 : exponential

THTOC1

Thème 1 : cardinalité Le Frido [1] ne définit pas la notion de nombre cardinal; ça nous mènerait trop loin. Au lieu de cela, nous allons nous contenter des notions d'équipotence, surpotence et subpotence, et démontrer un certain nombre de résultats en utilisant sans retenue le lemme de Zorn 1.23.

(1) Ensemble infini, définition 1.117.

(2) Ensemble dénombrable, définition 1.129.

(3) Cardinal d'un ensemble fini, définition 1.126.

(4) Définition d'équipotence, surpotence et subpotence, notations $A \succ B$ et $A \approx B$, définition 1.115.

(5) Toute partie d'un ensemble fini est finie, lemme 1.119.

(6) Si A est un ensemble fini ou dénombrable, alors il existe une surjection $\mathbb{N} \to A$, lemme 1.131.

(7) Si A est un ensemble infini et si $f\colon A \to B$ est une application injective, alors $f(A)$ est infini, proposition 1.124.

(8) Toute partie infinie de $\mathbb{N}$ est dénombrable, proposition 1.132

(9) Une bijection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$, proposition 1.135.

(10) Une décomposition de $\mathbb{N}$ en une infinité de parties équipotentes à $\mathbb{N}$, corolaire 1.136.

(11) Si il existe une surjection $\mathbb{N} \to A$, alors A est fini ou dénombrable, lemme 1.137.

(12) Une union dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable, proposition 1.138

(13) Tout ensemble infini contient une partie en bijection avec $\mathbb{N}$, proposition 1.142.

(14) Toute partie d'un ensemble fini est finie, et toute partie d'un ensemble dénombrable est finie ou dénombrable, proposition 1.143.

(15) Si $A \succeq B$ et $B \succeq A$, alors $A \approx B$, théorème de Cantor-Schröder-Bernstein 1.146

(16) Le théorème de Cantor 1.149 dit qu'il n'existe pas de surjection d'un ensemble vers son ensemble des parties. On en déduit qu'il n'existe pas d'ensemble contenant tous les ensembles (corolaire 1.151).

(17) Si A est infini et si $A \succeq B$, alors $A \approx A \cup B$ par le lemme 1.153.

(18) Si S est un ensemble infini alors il existe une bijection $\varphi\colon \{0,1\} \times S \to S$, proposition 1.154.

(19) Si A est infini, alors $A \times \mathbb{N} \approx A$, proposition 1.156.

(20) Si A est infini et si $B \prec A$, alors $A \backslash B \approx A$, lemme 1.158.

(21) Si A est infini, alors $A \approx A \times A$, théorème 1.162.

Il y a aussi des résultats de cardinalité autour des extensions de corps.

(1) Si $\mathbb{K}$ est un corps infini, alors $\mathbb{K}[X] \approx \mathbb{K}$.

(2) Le théorème de Steinitz 6.135 affirme que tout corps admet une unique clôture algébrique. La preuve utilise pas mal de cardinalité ainsi que le lemme de Zorn 1.23.

THTOC2

Thème 2 : morphismes et compagnie

(1) Un morphisme est un concept algébrique. Il s'agit d'une application (pas spécialement inversible) qui préserve la structure. Quand on parle de morphisme, il faut donc préciser la structure. On dit «morphisme de groupe», «morphisme d'espace vectoriel», «morphisme d'anneaux», etc.

(2) Morphisme de module, définition 1.330.

1. Ici je mets la référence [1]; pas parce qu'elle est utile ici, mais parce que je veux être sûr qu'elle soit numéro 1 de façon à être facilement reconnaissable. Elle indique les affirmations à propos desquelles la lectrice doit être doublement attentive.

(3) Morphisme de groupes, définition 1.38.

(4) Un isomorphisme d'espaces topologiques est une application continue, inversible, dont l'inverse est continue, 7.41. On dit aussi un homéomorphisme.

(5) Un difféomorphisme est différentiable d'inverse différentiable, définition 10.265.

(6) Un C^k-difféomorphisme est une application C^k d'inverse C^k. Définition 10.265.

(7) Un morphisme de groupes de Lie est un morphisme de groupe C^∞. Nous ne demandons pas que l'inverse ait une régularité particulière.

Le mot «homomorphisme» signifie exactement «morphisme», et, sauf incohérence de ma part, il n'est pas utilisé dans le Frido.

THEMEooNRZHoo

Thème 3 : arithmétique modulo, théorème de Bézout

(1) Pour $\mathbb{Z}^*$ c'est le théorème 1.236.

(2) Théorème de Bézout dans un anneau principal : corolaire 3.73.

(3) Théorème de Bézout dans un anneau de polynômes : théorème 6.47.

(4) En parlant des racines de l'unité et des générateurs du groupe unitaire dans le lemme 19.6. Au passage nous y parlerons de solfège.

(5) La proposition 1.262 qui donne tout entier assez grand comme combinaisons de a et b à coefficients positifs est utilisée en chaines de Markov, voir la définition 38.33 et ce qui suit.

(6) PGCD et PPCM sont dans la définition 1.186.

(7) Calcul effectif du PGCD puis des coefficients de Bézout : sous-sections 3.2.1.1 et 3.2.1.2.

THTOC4

Thème 4 : divisibilité

(1) Si a divise bc et si a est premier avec c, alors a divise b. Lemme de Gauss 3.12.

(2) Si a divise b et b divise a, alors $a = b$, lemme 3.14.

THTOC5

Thème 5 : équations diophantiennes

(1) Équation $ax + by = c$ dans $\mathbb{N}$, équation (3.58).

(2) Dans 3.2.6, nous résolvons $ax + by = c$ en utilisant Bézout (théorème 1.236).

(3) L'exemple 3.91 donne une application de la pure notion de modulo pour $x^2 = 3y^2 + 8$. Pas de solutions.

(4) L'exemple 3.92 résout l'équation $x^2 + 2 = y^3$ en parlant de l'extension $\mathbb{Z}[i\sqrt{2}]$ et de stathme.

(5) Les propositions 3.96 et 3.98 parlent de triplets pythagoriciens.

(6) Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

(7) La proposition 1.135 donne une bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ en résolvant dans $\mathbb{N}$ (entre autres) l'équation $k = y^2 + x$ pour k fixé.

THEMEooVIQIoo

Thème 6 : types d'anneaux

(1) Définition d'anneau : définition 1.42.

(2) La définition d'anneau principal est 1.228; pour un idéal principal, c'est 1.227.

(3) $\mathbb{Z}$ est intègre, lemme 1.252, principal et euclidien (proposition 1.255).

(4) $\mathbb{Z}[X]$ n'est pas principal (voir (5)).

ITEMooNQQMooSnuKvW

(5) Si A est un anneau intègre [2] qui n'est pas un corps, alors $A[X]$ n'est pas principal, lemme 1.256.

(6) L'anneau des fonctions holomorphes sur un compact donné est principal, proposition 26.68.

2. Définition 1.197.

(7) L'anneau $\mathbb{Z}[i\sqrt{3}]$ n'est pas factoriel, exemple 3.69.

(8) L'anneau $\mathbb{Z}[i\sqrt{5}]$ n'est ni factoriel ni principal, exemple 3.82.

(9) Tous les idéaux de $\mathbb{Z}/6\mathbb{Z}$ sont principaux, mais $\mathbb{Z}/6\mathbb{Z}$ n'est pas principal. Exemple 1.248.

anneau principal

— Définition 1.228.

— Il est intègre, définition 1.228.

— Il est noetherien, lemme 3.78.

— il est factoriel, théorème 3.80.

anneau euclidien

— Définition 1.251.

— Il est principal par 1.254

anneau intègre

— Définition 1.197.

anneau noetherien

— Définition 3.77.

anneau factoriel

— Définition 3.64.

— Il est intègre, définition 3.64.

corps

— Corps, définition 1.209.

— Il est un anneau intègre, lemme 1.198.

— Il est un anneau principal, proposition 3.81.

— Il est un anneau factoriel, proposition 3.81.

Dans un anneau intègre

— Un anneau fini intègre est un corps, proposition 3.62.

— A est intègre si et seulement si $A[X]$ est intègre, théorème 3.104.

— Si un pgcd[3] est inversible, tous les pgcd sont inversibles, lemme 1.200.

Dans un anneau principal

— p est premier si et seulement si il est irréductible si et seulement si l'idéal pA est maximal, proposition 1.234.

— Tous les idéaux sont principaux, définition 1.228.

— Si $\text{pgcd}(a,b)=1$, $\text{pgcd}(a,c)=1$, alors $\text{pgcd}(a,bc)=1$, lemme 3.75.

Dans un anneau factoriel

— Si p est irréductible et si $p \mid ab$, alors p divise a ou b, lemme 3.65.

— Tout élément irréductible est premier, proposition 3.66.

— Si p est irréductible, pA est un idéal premier, lemme 3.67.

— Si p est irréductible, A/pA est un anneau intègre, lemme 3.67 et proposition 1.231.

— P et Q sont primitifs si et seulement si PQ est primitif, lemme 3.119.

— Formule $c(PQ)=c(P)c(Q)$. Lemme 3.119.

— L'anneau des polynômes $\mathcal{P}(A)=A[X]$ est factoriel, théorème 6.35.

Dans un corps commutatif

— L'anneau $\mathbb{K}[X]$ est factoriel, proposition 6.36, euclidien et principal, lemme 3.110.

3. pgcd, définition 1.186.

— Les seuls idéaux de $\mathbb{K}$ sont $\{0\}$ et $\mathbb{K}$, lemme 1.183.

— L'anneau $\mathcal{P}(\mathbb{K})$ est euclidien, lemme 3.110.

élément irréductible

— Définition 1.189

élément premier

— Définition 1.188

— Si A est intègre et unitaire, l'élément p est premier si et seulement si l'idéal pA est premier, proposition 1.199.

éléments associés

— Définition 3.54.

Dans un anneau, un idéal est 1.182.

idéal premier

— Définition 1.229.

— Idéal premier si et seulement si A/I est anneau intègre, proposition 1.231.

idéal maximal

— Idéal maximal si et seulement si A/I est un corps, proposition 1.231.

idéal principal

— Définition 1.227.

— Dans un anneau unitaire intègre, p est irréductible si et seulement si il n'y a pas d'idéal principal I tel que $pA \subsetneq I \subsetneq A$, proposition 1.199.

THTOC7

Thème 7 : sous-groupes

(1) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 8.313.

(2) Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 13.40 ou proposition 13.41.

THEMEooQEEWo6XDO6s

Thème 8 : groupe symétrique

(1) Définition 1.274.

(2) La signature $\epsilon : S_n \to \{-1, 1\}$ est l'unique morphisme surjectif de S_n sur $\{-1, 1\}$, proposition 1.300(1).

(3) La table des caractères du groupe symétrique S_4 est donné dans la section 16.5.

(4) Le groupe symétrique S_4 est le groupe des symétries affines du tétraèdre régulier [4], proposition 18.195.

(5) Le groupe alterné A_5 est l'unique groupe simple d'ordre [5] 60, proposition 5.55.

(6) La proposition 5.45 donne la position du groupe alterné dans le groupe symétrique : A_n est un sous-groupe caractéristique de S_n, et l'unique sous-groupe d'indice 2.

THEMEooKZHBo6RKTOC9

Thème 9 : action de groupe

(1) Définition d'une action de groupe sur un ensemble : 1.367.

(2) Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96.

(3) La formule de Burnside (théorème 2.40) parle du nombre d'orbites pour l'action d'un groupe fini sur un ensemble fini.

(4) Des applications de la formule de Burnside : le jeu de la roulette et l'affaire du collier, 18.10.15.1 et 18.10.15.2.

(5) Le groupe symétrique S_n agit sur l'anneau $\mathbb{K}[T_1, \ldots, T_n]$, lemme 1.368.

4. Définition 12.147.
5. Ordre d'un groupe, définition 1.267.

THTOC10

Thème 10 : classification de groupes

(1) Ordre d'un groupe, définition 1.267.

(2) Structure des groupes d'ordre pq, théorème 5.25.

(3) Le groupe alterné est simple, théorème 5.51.

(4) Définition 5.6 d'un p-groupe.

(5) Théorème de Sylow 5.11.

(6) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.

(7) $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.32.

(8) Le groupe $(\mathbb{Z}/p\mathbb{Z})^*$ est cyclique, proposition 3.137.

(9) Le premier théorème d'isomorphisme 2.6 dit que si θ est un morphisme de groupes, alors $G/\ker(\theta) = \mathrm{Image}(\theta)$.

(10) Le deuxième théorème d'isomorphisme 2.7 dit que si N est normal dans G, alors $NH/N = H/(H \cap N)$.

(11) Le troisième théorème d'isomorphisme 2.9 dit que, avec des hypothèses de normalité, $(G/M)/(N/M) = G/N$.

À propos d'ordre dans un groupe.

(1) L'ordre d'un groupe est son cardinal. L'ordre d'un élément est le plus petit n tel que $g^n = e$, définition 1.268.

(2) L'ordre d'un élément divise l'ordre du groupe, corolaire 2.14.

(3) Si H est normal dans G, alors $|G/H| = |G|/|G|$, théorème de Lagrange 2.13.

(4) Si G est un groupe cyclique, pour tout diviseur p de $|G|$, le groupe G contient au moins un élément d'ordre p. Théorème de Cauchy 3.26.

THMooUDYMo

Thème 11 : fonction indicatrice d'Euler

(1) Définition de $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*$ dans 5.29.

(2) Propriétés genre $\varphi(pq) = \varphi(p)\varphi(q)$, corolaire 5.43.

(3) $\varphi(n) = \mathrm{Card}(\Delta_n)$, proposition 19.12.

(4) $n = \sum_{d\,|\,n} \varphi(d)$, lemme 5.38.

(5) Le théorème de Euler-Fermat 5.34 donne $a^{\varphi(n)} \in [1]_n$ dès que a et n sont premiers entre eux.

(6) Si A et B sont premiers entre eux, il existe p, m tels que $A^p = mB + 1$, proposition 5.34.

THTOC12

Thème 12 : produit semi-direct de groupes

(1) Définition 2.47.

(2) Le corolaire 2.49 donne un critère pour prouver qu'un produit NH est un produit semi-direct.

(3) L'exemple 18.177 donne le groupe des isométries du carré comme un produit semi-direct.

(4) Le théorème 5.25 classifie les groupes d'ordre pq (p, q premiers distincts) à grands coups de produit semi-directs.

(5) Le théorème 18.79 donne les isométries de $\mathbb{R}^n$ par $\mathrm{Isom}(\mathbb{R}^n) = T(n) \times_\rho O(n)$ où $T(n)$ est le groupe des translations.

(6) La proposition 18.81 donne une décomposition du groupe orthogonal $\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2$ où $C_2 = \{\mathrm{Id}, R\}$ où R est de déterminant -1.

(7) La proposition 11.63 donne $\mathrm{Aff}(\mathbb{R}^n) = T(n) \times_\rho \mathrm{GL}(n,\mathbb{R})$ où $\mathrm{Aff}(\mathbb{R}^n)$ est le groupe des applications affines bijectives de $\mathbb{R}^n$.

THTOC13

Thème 13 : théorie des représentations

(1) Définition 4.135.

(2) Table des caractères du groupe diédral, section 18.17.

(3) Table des caractères du groupe symétrique S_4, section 16.5.

THTOC14

Thème 14 : espaces vectoriels

(1) Les notations $\mathcal{L}(V,W)$, $\mathcal{L}^0(V,W)$, $\mathrm{GL}(V,W)$ ainsi que les notions de morphismes et d'isomorphismes d'espaces vectoriels normés sont dans la définition 10.272.

(2) Existence d'une base. Pour un espace vectoriel quelconque, proposition 4.24.

(3) Théorème de la base incomplète. Pour un espace vectoriel quelconque, théorème 4.25.

THTOC15

Thème 15 : rang

(1) Définition pour une application linéaire : 4.46, pour une matrice : 4.106. L'équivalence est la proposition 4.110.

(2) Le théorème du rang, théorème 4.47

(3) Pour une applicaton linéaire entre deux espaces vectoriels de même dimension finie, il est équivalent d'être injectif, surjectif ou bijectif, c'est le corolaire 4.49.

(4) Pour prouver que des matrices sont équivalentes et pour les mettre sous des formes canoniques, nous avons le lemme 4.112 et son corolaire 4.113.

(5) Tout hyperplan de $\mathbb{M}(n,\mathbb{K})$ coupe $\mathrm{GL}(n,\mathbb{K})$, corolaire 4.113. Cela utilise la forme canonique sus-mentionnée.

(6) Le lien entre application duale et orthogonal de la proposition 8.195 utilise la notion de rang.

(7) Le lemme 8.303 parle de commutant et utilise la notion de rang. Ce lemme sert à prouver diverses conditions équivalentes à être un endomorphisme cyclique dans le théorème 8.304.

THEMEooOAJKTHTOC16

Thème 16 : formes bilinéaires et quadratiques

(1) Les formes bilinéaires sont définies en 8.124.

(2) Forme quadratique, définition 8.125.

(3) Équivalence de forme quadratiques, définition 8.252. Deux formes quadratiques sont équivalentes si et seulement si elles ont même signature, proposition 8.253.

(4) Une isométrie d'une forme bilinéaire est affine ou linéaire, théorème 8.157.

(5) Forme bilinéaire dégénérée, définition 8.129.

(6) Une forme bilinéaire est non-dégénérée si et seulement si sa matrice associée est inversible, c'est la proposition 8.229.

(7) Une isométrie d'une forme bilinéaire est linéaire ou affine par le théorème 8.157.

(8) Toute forme quadratique admet des bases orthogonales, théorème 8.245 pour le cas général; proposition 8.256 pour le cas de $\mathbb{R}^n$, en se basant sur le théorème spectral.

(9) Base q-orthogonale pour une forme quadratique, théorème 8.245.

(10) Le concept de projection orthogonale est la définition 12.142 en dimension finie et la définition 25.7 dans le cas des espaces de Hilbert.

(11) Produit hermitien, définition 8.176. Opérateur hermitien ($A^\dagger = A$), opérateur unitaire ($A^\dagger A = 1$), définition 8.179

matrice

(1) Matrice d'une forme quadratique, définition 8.148.

(2) Théorème de Sylvester à propos de signature (définition 8.164) de forme quadratique réelle : 8.251.

orthogonalité

(1) Il existe une base q-orhogonale, proposition 8.243 et théorème 8.245.

(2) Définition de l'orthogonal $A^\perp$ et du noyau $\ker(b)$, définitions 8.131, 8.132.

(3) Diverses propriétés comme $A^\perp = \text{Span}(A)^\perp$ et $V \cap V^\perp = \ker(b_V)$, propositions 8.133 et 8.134.

(4) En dimension finie, $(V^\perp)^\perp = V$, lemme 8.137.

(5) Nous avons $\dim(V) + \dim(V^\perp) = \dim(E)$, lemme 8.136.

(6) Pour une forme quadratique strictement définie positive ou négative, $E = F \oplus E^\perp$, lemme 8.139.

THTOC17

Thème 17 : endomorphismes cycliques

(1) Définition 8.105.

(2) Groupe cyclique, définition 1.326.

(3) Son lien avec le commutant donné dans la proposition 8.301 et le théorème 8.304.

(4) Utilisation dans le théorème de Frobenius (invariants de similitude), théorème 8.292.

THEMEooZYKF

Thème 18 : extension de corps et polynômes

(1) Définition d'une extension de corps 6.59.

(2) Définition de polynôme minimal: 6.64.

(3) Pour l'extension du corps de base d'un espace vectoriel et les propriétés d'extension des applications linéaires, voir la section 8.15.

(4) Extension de corps de base et similitude d'application linéaire (ou de matrices, c'est la même chose), théorème 8.306.

(5) Extension de corps de base et cyclicité des applications linéaires, corolaire 8.305.

(6) À propos d'extensions de $\mathbb{Q}$, le lemme 6.183.

(7) Corps de rupture : définition 6.114 existence par la proposition 6.120. Il n'y a pas unicité.

(8) Corps de décomposition : définition 6.142. Attention : le plus souvent nous parlons de corps de décomposition d'un seul polynôme. Cette définition est un peu surfaite. Existence par la proposition 6.143 qui le donne même comme extension par toutes les racines, et unicité à isomorphisme près par le théorème 6.145, énoncé de façon plus simple (mais pas indépendante !) en la proposition 6.146.

(9) Si $\mathbb{K}$ est algébrique clos et si $\alpha\colon \mathbb{K} \to \mathbb{L}$ est une extension algébrique, alors $\alpha(\mathbb{K}) = \mathbb{L}$ par le lemme 6.80.

Un trio de résultats d'enfer est :

(1) Dans un anneau principal qui n'est pas un corps, un idéal est maximal si et seulement si il est engendré par un irréductible [6] (proposition 1.244).

(2) Dans un anneau, un idéal I est maximal si et seulement si A/I est un corps (proposition 1.220)

(3) Si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ est principal (lemme 3.110).

6. Irréductibilité, y compris polynôme irréductible, définition 1.189.

THTOC19

Thème 19 : polynômes

Définitions Soient un anneau A, un corps $\mathbb{K}$, une extension $\mathbb{L}$ de $\mathbb{K}$ et un élément $\alpha \in \mathbb{L}$.

(1) La définition la plus formelle est en tant que module produit $A^{(\mathbb{N})}$, définition 1.359. Le produit et l'évaluation sont définis en 1.362 et la formule $(PQ)(x) = P(x)Q(x)$ dans 1.363.

(2) Si $\mathbb{A}$ est commutatif, alors $\mathbb{A}[X]$ est également commutatif, lemme 1.364.

(3) En ce qui concerne la notation $A[X]$, elle ne devrait pas être utilisée, voir 1.19.2. L'ensemble des polynômes sera noté $\mathcal{P}(A)$ et ceux de degré n (définition 1.361), $\mathcal{P}_n(A)$.

(4) $A[X]$, définition 1.359; l'anneau $\mathbb{K}[X]$ a même définition parce que c'est un cas particulier. L'évaluation d'un polynôme en un élément de l'anneau, $P(\alpha)$ est définie en 1.362.

(5) Liens entre $\mathbb{K}[\alpha]$, $\mathbb{K}[X]$, $\mathbb{K}(\alpha)$ et $\mathbb{K}(X)$ lorsque α est transcendant, proposition 6.100. Et la proposition 6.103 pour le cas où α est algébrique[7].

(6) Si A est un anneau et si α est un élément d'une extension de A (comme anneau), nous écrivons $A[\alpha]$ pour le plus petit sous-anneau de B contenant A et α. C'est la définition 3.102.

(7) $\mathbb{K}(X)$, le corps des fractions de $\mathbb{K}[X]$, définition 6.84. Si $R = P/Q$ dans $\mathbb{K}(X)$, l'évaluation est $R(\alpha) = P(\alpha)Q(\alpha)^{-1}$, définition 6.85.

(8) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est le plus petit corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, définition 6.86.

(9) À propos de polynômes à plusieurs variables.

— Anneau de polynômes : $A[X_1, \dots, X_n]$ est la définition 3.50.

— Corps engendré : $\mathbb{K}(\alpha_1, \dots, \alpha_n)$ est la définition 6.139.

— Corps des fractions rationnelles : $\mathbb{K}(X_1, \dots, X_n)$ est la définition 6.140.

contenu À propos de $c(P)$.

(1) Définition du contenu $c(P)$, 3.113.

(2) $c(PQ) = c(P)c(Q)$, lemme de Gauss 19.46, et lemme 3.119 pour un anneau factoriel.

(3) Si $c(P) = 1$, on dit que P est primitif au sens des pgcd, définition 3.115.

(4) Si $a \in A$ nous avons $c(aP) = ac(P)$ par le lemme 1.201.

(5) Un polynôme primitif est la définition 3.115.

Dérivation Plusieurs propriétés du polynôme dérivé P'.

(1) La définition 6.155 donne $P' = \sum_{k=1}^{n} ka_iX^{k-1}$.

(2) Règle de Leibniz, lemme 6.156.

(3) Caractérisation de $P' = 0$ en fonction de la caractéristique de $\mathbb{K}$, lemme 6.157.

Coefficients dans un anneau commutatif (1) Les polynômes à coefficients dans un anneau commutatif sont à la section 3.7.

(2) Un anneau A est intègre si et seulement si $A[X]$ est intègre; théorème 3.104.

Coefficients dans un corps Les polynômes à coefficients dans un corps sont à la section 6.3.

(1) Si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ est euclidien et principal, lemme 3.110.

(2) Nous parlons de l'idéal des polynômes annulateurs dans le théorème 6.43.

(3) Le théorème 6.43 dit que $\mathbb{K}[X]$ est un anneau principal et que tous ses idéaux sont engendrés par un unique polynôme unitaire.

(4) Le polynôme minimal est irréductible, proposition 6.67.

(5) Quelques formules sur le pgcd, lemme 6.57.

7. Définition 6.71.

Polynôme primitif (1) Un polynôme est irréductible sur A si et seulement si il est irréductible et primitif sur le corps des fractions, corolaire 3.134.

Polynôme d'endomorphisme C'est la section 8.7.

Racines et factorisation (1) Si $\mathbb{A}$ est un anneau, la proposition 3.123 factorise une racine.

(2) Si $\mathbb{A}$ est un anneau, la proposition 3.128 factorise une racine avec sa multiplicité.

(3) Si $\mathbb{A}$ est un anneau, le théorème 3.130 factorise plusieurs racines avec leurs multiplicités.

(4) Le théorème 3.130 nous indique que lorsqu'on a autant de racines (multiplicité comprise) que le degré, alors nous avons toutes les racines.

(5) Si $\mathbb{K}$ est un corps et α une racine dans une extension, le polynôme minimal de α divise tout polynôme annulateur par la proposition 6.101.

(6) Le théorème 6.111 annule un polynôme de degré n ayant $n+1$ racines distinctes.

(7) La proposition 6.186 nous annule un polynôme à plusieurs variables lorsqu'il a trop de racines.

(8) En analyse complexe, le principe des zéros isolés 17.136 annule en gros toute série entière possédant un zéro non isolé.

(9) Polynômes irréductibles sur $\mathbb{F}_q$.

THTOC20

Thème 20 : polynôme d'endomorphismes

(1) Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 8.301.

(2) Racine carrée d'une matrice hermitienne positive, proposition 13.27.

(3) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.

(4) Décomposition de Dunford, théorème 8.265.

(5) Algorithme des facteurs invariants 4.114.

THEMEooULGFTHTOC21

Thème 21 : dualité Ne pas confondre dual algébrique et dual topologique d'un espace vectoriel.

(1) Définition du dual, 4.126.

(2) Le dual d'un espace de Banach est de Banach, proposition 7.267.

(3) Définition de la base duale 4.127.

(4) $\dim(E) = \dim(E^*)$, lemme 4.127.

(5) Base préduale (existence, unicité) : proposition 4.131.

(6) Théorème de représentation de Riez 27.186 : $L^p = (L^q)'$, et en particulier $L^\infty = (L^1)'$.

(7) Il n'est pas vrai que $(L^\infty)' = L^1$, voir la proposition 27.232.

(8) Dans un espace de Banach [8], $\|x\| = \max_{\substack{\varphi\in E'\\ \|\varphi\|=1}} |\varphi(x)|$, proposition 27.174.

Dual topologique et algébrique Ils sont définis par 4.126. Le dual algébrique est l'ensemble des formes linéaires, et le dual topologique ne considère que les formes linéaires continues (en dimension infinie, les applications linéaires ne sont pas toutes continues).

Topologie Une topologie possible sur le dual d'un espace vectoriel topologique est celle $*$-faible de la définition 7.358.

Nous comparons les topologies faibles et de la norme en la section 10.15.

Théorèmes de dualité Quelques théorèmes établissent des dualités entre des espaces courants.

(1) En dimension finie, $x \mapsto b(x,.)$ est isomorphisme $E \to E^*$ lorsque b est une forme bilinéaire, proposition 8.135.

(2) Le théorème de représentation de Riesz 25.18 pour les espaces de Hilbert.

(3) La proposition 27.184 pour les espaces $L^p\big([0,1]\big)$ avec $1 < p < 2$.

(4) Le théorème de représentation de Riesz 27.186 pour les espaces L^p en général.

Tous ces théorèmes donnent la dualité par l'application $\Phi_x = \langle x, . \rangle$.

8. Espace de Banach, définition 7.265.

THTOC22

Thème 22 : injections

(1) L'espace de Sobolev $H^1(I)$ s'injecte de façon compacte dans $C^0(\bar{I})$, proposition 31.6.

(2) L'espace de Sobolev $H^1(I)$ s'injecte de façon continue dans $L^2(I)$, proposition 31.6.

(3) L'espace $L^2(\Omega)$ s'injecte continument dans $\mathscr{D}'(\Omega)$ (les distributions), proposition 30.35.

INTooVDSCoo THTOC23

Thème 23 : tribu, algèbre de parties, λ-systèmes et co. Il existe des centaines de notions de mesures et de classes de parties.

(1) Le plus souvent lorsque nous parlons de mesure est que nous parlons de mesure positive, définition 14.18 sur un espace mesuré avec une tribu, définition 14.1.

(2) Une mesure extérieure est la définition 14.10

(3) Une algèbre de partie : définition 14.13. Une mesure sur une algèbre de parties : définition 14.11. L'intérêt est que si on connait une mesure sur une algèbre de parties, elle se prolonge en une mesure sur la tribu engendrée par le théorème de prolongement de Hahn 14.80.

(4) Un λ-système : définition 14.31.

(5) Une mesure complexe : définition 14.239.

En théorie de l'intégration, si X est une partie de $\mathbb{R}^n$, la convention est de considérer des fonctions

$$f\colon \big(X, \mathcal{L}\mathrm{eb}(X)\big) \to \big(\mathbb{R}, \mathcal{B}\mathrm{or}(\mathbb{R})\big).$$

Voir les points 14.120 et 14.166 pour les conventions à ce propos.

À propos d'applications mesurables :

(1) Définition d'une application mesurable, définition 14.44.

(2) Une fonction continue est borélienne, théorème 14.56.

(3) Si les f_n sont mesutables (au sens des boréliens), alors $\sup_n f_n$ est mesurable, lemme 14.100.

À propos de tribu induite:

(1) Définition 14.8.

(2) Les boréliens induits sont bien les boréliens de la topologie induite[9] : $\mathcal{B}\mathrm{or}(Y) = \mathcal{B}\mathrm{or}(X)_Y$, théorème 14.57.

Tribu engendrée.

(1) Tribu engendrée par des parties, définition 14.4.

THEMEooKLVRTHTOC24

Thème 24 : théorie de la mesure

Mesure À propos de mesure.

(1) Tribu des boréliens, définition 14.50.

(2) Mesure positive, mesure finie et σ-finie, c'est la définition 14.18.

(3) Le produit de tribus est donné par la définition 14.127,

(4) Produit d'une mesure par une fonction, définition 14.212.

(5) le produit d'espaces mesurés est donné par la définition 14.247.

(6) Mesure de Lebesgue sur $\mathbb{R}$, définition 14.143.

(7) Une partie de $\mathbb{R}$ non mesurable au sens de Lebesgue : l'exemple 14.157.

(8) Mesure de Lebesgue sur $\mathbb{R}^N$, définition 14.249.

(9) Mesure à densité, définition 14.212.

Opérations On peut faire quelques opérations sur les mesures.

9. Topologie induite, définition 7.26.

(1) Si w est une fonction et μ une mesure, on définit $w \cdot \mu$ par 14.212 : $(w \cdot \mu)(A) = \int_A w d\mu$).

Théorèmes d'approximation Il est important de pouvoir approcher des fonctions continues ou L^p par des fonctions étagées, sinon on ne parvient pas à faire tourner la machine de l'intégration de Lebesgue.

(1) Si $(S, \mathcal{A}, \mu)$ est un espace mesuré et si $f : S \to [0, +\infty]$ est une fonction mesurable, le théorème fondamental d'approximation 14.118 dit qu'il existe une suite croissante de fonctions étagées qui converge vers f.

(2) Les fonctions simples sont denses dans L^p, proposition 27.48.

(3) Encadrement d'un borélien A par un fermé F et un ouvert V par le lemme 14.86 : $F \subset A \subset V$ avec $\mu(V \backslash F) < \epsilon$.

(4) Approximation L^p de la fonction caractéristique d'un borélien par une fonction continue par le théorème 14.242.

Produit (1) La tribu produit, définition 14.127.

(2) La mesure produit, définition 14.246.

THEMEooUJVXBHZ0625

Thème 25 : normes

Définitions (1) Espace vectoriel normé : définition 7.158.

(2) Produit scalaire, définition 8.168.

(3) Norme associée à un produit scalaire (cas réel), théorème 10.1.

(4) Norme associée à une forme sesquilinéaire (cas complexe), proposition 9.113.

(5) Produit scalaire sur $\mathbb{R}^n$, définition 8.173. Norme sur $\mathbb{R}^n$, définition 10.3.

(6) Forme hermitienne sur $\mathbb{C}^n$, définition 8.183. Norme sur $\mathbb{C}^n$, définition 9.114.

(7) La proposition 10.44 donne les normes $\|x\|_1$, $\|x\|_2$ et $\|x\|_\infty$ sur $\mathbb{R}^n$.

(8) Sur $\mathbb{R}^n$, la proposition 17.105 dit que $\|x\|_p$ est une norme.

Topologie (1) Métrique associée à une norme $d(x, y) = \|x - y\|$, définition 7.163.

(2) La topologie d'un espace vectoriel normé est la topologie métrique du théorème 7.117.

(3) Les boules sont les boules métrique définies en (7.111).

Inégalités (1) En général pour les normes $\|.\|_p$, il y a des inégalités dans 17.106 et 17.98.

(2) La proposition 17.111 donne l'inégalité $\|x\|_q \leqslant n^{\frac{1}{q}-\frac{1}{p}} \|x\|_p$ dès que $0 < q < p$.

Équivalence de norme (1) Définition de l'équivalence de norme 10.46.

(2) La proposition 10.47 sur l'équivalence des normes $\|.\|_2$, $\|.\|_1$ et $\|.\|_\infty$ dans $\mathbb{R}^n$.

(3) Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes par le théorème 10.49.

Autres (1) Montrer que le problème $a - b$ est stable dans l'exemple 34.27.

(2) La proposition 12.114 donnant $\rho(A) \leqslant \|A\|$ utilise l'équivalence de toutes les normes sur un espace vectoriel de dimension finie (théorème 10.49.).

(3) La norme $x \mapsto \|x\|$ est une application continue, proposition 7.166.

Norme opérateur et d'algèbre voir le thème 40.

THTOC26

Thème 26 : inégalités Dans $\mathbb{C}$ nous avons $|a + b| \leqslant |a| + |b|$ par la proposition 9.102(6).

Inégalité de Clarkson Pour $f, g \in L^p$ nous avons l'inégalité de Clarkson 27.149 :

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \leqslant \frac{1}{2}\Big(\|f\|_p^p + \|g\|_p^p\Big). \tag{-1.1}$$

Inégalité de Young Nous avons

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} \tag{-1.2}$$

par la proposition 27.30.

Inégalité de Jensen (1) Une version discrète pour $f\big(\sum_i \lambda_i x_i\big)$, la proposition 17.102.

(2) Une version intégrale pour $f\big(\int \alpha d\mu\big)$, la proposition 27.31.

(3) Une version pour l'espérance conditionnelle, la proposition 36.86.

Inégalité de Hölder Il en existe de nombreuses versions et variations.

(1) Hölder pour L^p: $\|fg\|_1 \leqslant \|f\|_p \|g\|_q$, proposition 27.33.

(2) Hölder pour ℓ^p: $\|x\|_q \leqslant n^{\frac{1}{q}-\frac{1}{p}} \|x\|_p$, proposition 17.111.

(3) $\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p}\|x\|_\infty$, théorème 17.106

(4) $\|x\|_p \leqslant n^{1/p}\|x\|_q$, corolaire 17.107.

Inégalité de Minkowsky (1) Pour une forme quadratique[10] q sur $\mathbb{R}^n$ nous avons $\sqrt{q(x+y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$. Proposition 10.12.

(2) Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$. Proposition 27.39.

(3) L'inégalité de Minkowsky sous forme intégrale s'écrit sous forme déballée

$$\left[\int_X \Big(\int_Y |f(x,y)| d\nu(y)\Big)^p d\mu(x)\right]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y).$$

ou sous forme compacte

$$\left\| x \mapsto \int_Y f(x,y) d\nu(y) \right\|_p \leqslant \int_Y \|f_y\|_p d\nu(y)$$

C'est la proposition 27.41.

Transformée de Fourier Pour tout $f \in L^1(\mathbb{R}^n)$ nous avons $\|\hat{f}\|_\infty \leqslant \|f\|_1$, lemme 29.12.

Inégalité des normes Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 27.70.

c_r inégalité La proposition 27.139 donne $\big|f(\omega) + g(\omega)\big|^r \leqslant c_r |f(\omega)|^r + c_r |g(\omega)|^r$.

THEMEooYRIWoTHXOo27

Thème 27 : constructions topologiques

topologie produit Si X et Y sont des espaces topologiques, nous pouvons construire une topologie sur $X \times Y$.

(1) La définition de la topologie produit est 7.15.

(2) Pour les espaces vectoriels normés, le produit est donné par la définition 7.226.

(3) L'équivalence entre la topologie de la norme produit et la topologie produit est le lemme 7.226.

(4) Quand V et W sont des espaces métriques, la topologie considérée sur $V \times W$ est celle de la définition 7.226. C'est à la fois la topologie de la norme produit et la topologie produit.

(5) La convergence dans un espace vectoriel est si et seulement si il y a convergence composante par composante, proposition 7.65.

(6) Dans le cas d'espaces normés, la topologie produit est la même que celle de la norme produit, lemme 7.227.

topologie induite Si X est un espace topologique et si A est une partie de X, nous mettons une topologie sur A.

10. Définition 8.125.

(1) Le topologie induite, définition 7.26.

(2) Si X est un espace vectoriel normé, la topologie induite est celle de la norme restreinte : lemme 7.122.

topologie quotient Si X est topologique et si $\sim$ est une relation d'équivalence, nous définissons une topologie sur $X/\sim$.

(1) La topologie quotient est définie en 7.49.

(2) Si X est vectoriel normé, la topologie sur $X/\sim$ est aussi donnée par une norme quotient de la définition 7.355. La proposition 7.357 donne l'équivalence entre la topologie quotient et la norme quotient.

topologie rendant continues des applications Si nous avons une application $f\colon X \to Y$ et si Y est topologique, nous mettons une topologie sur X rendant f continue.

(1) Définition de la plus petite topologie rendant des applications continues, proposition 7.44.

THTOC28

Thème 28 : espaces métriques, normés

(1) Un espace métrique est un ensemble muni d'une distance, définition 7.115.

(2) La distance entre un point et un ensemble est la définition 7.148.

(3) Le théorème-définition 7.117 donne la topologie sur un espace métrique en disant que les boules ouvertes sont une base de la topologie (définition 7.2).

(4) La définition de la convergence d'une suite est la définition 7.13.

(5) Dans un espace vectoriel normé, une application est continue si et seulement si elle est bornée, proposition 10.70.

(6) Un espace vectoriel topologique [11] qui possède en tout point une base dénombrable de topologie accepte une distance, théorème 7.282.

THEMEooGVCCoBTHTOC29

Thème 29 : limite et continuité

(1) Limite d'une fonction en un point : définition 7.110. Il n'y a pas unicité en général comme le montre l'exemple 7.59 dans un espace non séparé.

(2) Définition de la continuité d'une fonction en un point et sur une partie de l'espace de départ : définition 7.36.

(3) Caractérisation de la limite dans $\mathbb{R}$, proposition 12.1.

(4) Unicité de la limite d'une suite dans un espace séparé : proposition 7.62. Unicité de la limite d'une fonction, toujours dans le cas d'un espace séparé : proposition 7.113.

(5) La proposition 7.361 donne l'unicité de la limite dans le cas des espaces duaux pour la topologie $*$-faible. La proposition 7.113 nous dira qu'il y a unicité dès que l'espace d'arrivée est séparé.

(6) Continuité sur une partie si et seulement si continue en chaque point, c'est le théorème 7.205.

(7) Voir l'exemple 12.54 traité en détail pour la (non) continuité d'une fonction qui fait un saut en un point.

(8) La fonction $f(x,y) = x + y$ est continue, lemme 9.29.

THEMEooHINHoTHTOC30

Thème 30 : intégration À propos d'intégration.

Vocabulaire La différence entre une intégrale qui existe fonction intégrable est dans 14.187.

L'ordre dans lequel les choses sont faites — Nous commençons par considérer des fonctions $f\colon \Omega \to [0,+\infty]$ dans la définition 14.167.

11. Définition 7.170.

— Nous donnerons ensuite quelques propriétés restreintes aux fonctions à valeurs positives, par exemple
  (1) La convergence monotone 14.178,
  (2) Lemme de Fatou 14.182.
  (3) (presque) linéarité pour les fonctions positives, théorème 14.183.
— La définition pour les fonctions à valeurs dans $\mathbb{R}$ puis $\mathbb{C}$ est 14.186.
— Pour les fonctions à valeurs dans un espace vectoriel, c'est la définition 14.197.

primitive et intégrale (1) La définition 14.285 donne $\int_a^b f = \int_{]a,b[} f$ lorsque f est intégrable sur $]a,b[$.

(2) Lorsque f n'est pas intégrable sur $]a,b[$ nous pouvons poser $\lim_{x\to b} \int_a^b f$ et dire que c'est une intégrale impropre, définition 14.297.

Quelque résultats (1) Intégrale associée à une mesure, définition 14.167
(2) L'existence d'une primitive pour toute fonction continue est le théorème 12.424.
(3) La définition d'une primitive est la définition 12.198.
(4) Primitive et intégrale, proposition 14.286.
(5) Intégrale impropre, définition 14.297.

Intégrale et mesure (1) L'intégrale de la fonction 1 donne la mesure : $\int_B 1 d\mu = \mu(B)$, c'est le lemme 14.175.
(2) Le théorème de Radon-Nikodym 14.235 donne une densité pour certaines mesures.
(3) Le produit d'une mesure par une fonction donnée par la définition 14.212 introduit aussi une densité : $(w \cdot \mu)(A) = \int_A w d\mu$.

Limite dans les bornes (1) Si $f\colon \mathbb{R} \to \mathbb{R}$ est intégrable alors $\lim_{b\to\infty} \int_{[a,b]} f = \int_{[0,\infty[} f$.
(2) Nous avons $\int_{\mathbb{R}} f = \lim_{R\to\infty} \int_{-R}^{R} f$ par la proposition 14.262.

Autre résultats (1) Si $A, B \subset \Omega$ sont des parties disjointes, alors $\int_{A\cup B} f = \int_A f + \int_B f$, proposition 14.194.
(2) La σ-additivité dénombrable, $\int_{\bigcup_i A_i} f d\mu = \sum_{i=0}^{\infty} \int_{A_i} f d\mu$ est dans les propositions 14.210 et 14.211.

THTOC31

Thème 31 : connexité

(1) Définition 7.70
(2) L'image d'un connexe par une fonction continue est connexe, lemme 7.218.
(3) Un espace topologique est connexe si et seulement si il vérifie les valeurs intermédiaires, proposition 9.57.
(4) Dans un connexe, une partie ouverte et fermée est soit vide soit l'espace entier, proposition 7.74(4).
(5) Connexité par arcs, définition 9.64.
(6) Si U est connexe et si $U \subset S \subset \bar{U}$, alors S est connexe, proposition 7.75.
(7) Une partie de $\mathbb{R}^2$ qui est connexe, mais pas connexe par arcs, proposition 21.56.
(8) Une partie de $\mathbb{R}$ est connexe si et seulement si elle est un intervalle, proposition 9.53.
(9) Le groupe $\mathrm{SL}(n,\mathbb{K})$ est connexe par arcs, proposition 13.18.
(10) Le groupe $\mathrm{GL}(n,\mathbb{C})$ est connexe par arcs, proposition 13.19.
(11) Le groupe $\mathrm{GL}(n,\mathbb{R})$ a exactement deux composantes connexes par arcs, proposition 13.20.
(12) Le groupe $\mathrm{O}(n,\mathbb{R})$ n'est pas connexe, lemme 13.15.
(13) Les groupes $\mathrm{U}(n)$ et $\mathrm{SU}(n)$ sont connexes par arcs, lemme 13.16.

(14) Pour tout $n \geqslant 2$, le groupe $\mathrm{SO}(n)$ est connexe, le groupe $\mathrm{O}(n)$ a deux composantes connexes, proposition 13.4.

(15) Connexité des formes quadratiques de signature donnée, proposition 17.117.

(16) Dans un espace vectoriel normé, les connexes par arcs sont connexes par arcs C^1, proposition 15.6.

THTOC32

Thème 32 : séparation

(1) Espace topologique normal, définition 21.108.

(2) Le dual d'un espace vectoriel normé sépare les points, corolaire 27.175.

(3) Les fermés et les compacts sont séparés dans un espace vectoriel topologique, proposition 7.271.

(4) Le théorème de Hahn-Banach 27.169 parle d'hyperplan qui sépare des parties convexes disjointes (et autres hypothèses).

(5) Les ouverts séparent les points et les compacts, lemme 7.105.

THEMEooQQBHTHLF033

Thème 33 : compacts

Propriétés générales Quelques propriétés de compacts.

(1) La définition d'un ensemble compact est la définition 7.80.

(2) Ne pas confondre le compactifié d'Alexandrov 7.106 avec la droite réelle achevée 12.24.

(3) Une partie est relativement compacte si sa fermeture est compacte, définition 7.103. Une union finie de relativement compacts est relativement compacte, lemme 7.104.

(4) Si M est un compact de $A \times B$, alors $M \subset K \times L$ où K est compact de A et L de B, proposition 7.101.

(5) Un fermé dans un compact est compact, lemme 7.96

(6) Dans un espace Hausdorff[12], les compacts sont fermés, 7.96(2).

(7) Un espace est compact si et seulement si toute intersection finie de fermé est non vide, théorème 7.109.

(8) Tout compact d'un espace topologique séparé est fermé, lemme 7.96(2).

(9) Dans un espace vectoriel réel de dimension finie, les compacts sont les fermés bornés par le théorème 9.24.

(10) Le théorème de Borel-Lebesgue 9.19 dit qu'un intervalle[13] de $\mathbb{R}$ est compact si et seulement si il est de la forme $[a, b]$.

(11) Théorème des fermés emboîtés dans le cas compact, corolaire 7.93. À ne pas confondre avec celui dans le cas des espaces métrique, théorème 7.299.

(12) L'image d'un compact par une fonction continue est un compact, théorème 7.220.

(13) Si $f : K \to X$ est une bijection continue, sa réciproque est continue, lemme 7.222.

(14) Suites dans un compact

(14a) Toute suite dans un compact admet une sous-suite convergente, théorème 7.290.

(14b) Dans $\mathbb{R}^n$, toute suite dans un compact admet une sous-suite convergente, théorème 9.58. La démonstration de ce théorème est non seulement plus compliquée que le cas général, mais utilise en plus le cas dans $\mathbb{R}$; lequel cas n'est pas démontré de façon directe dans le Frido.

(14c) Un espace métrique est compact si et seulement si toute suite contient une sous-suite convergente. C'est le théorème de Bolzano-Weierstrass 7.144. La démonstration de ce théorème est indépendante.

12. Hausdorff, définition 7.60.
13. Définition 1.21.

(15) Une fonction continue sur un compact est bornée et atteint ses bornes, théorème 7.146.

(16) Une fonction continue sur un compact y est uniformément continue, théorème de Heine 12.78.

(17) Une bijection continue $f : K \to X$ entre un compact et un séparé est un isomorphisme, 7.221.

Produits de compacts À propos de produits de compacts. C'est un compact dans tous les cas métriques[14].

(1) Les produits d'espaces métriques compacts sont compacts. Il s'agit du théorème de Tykhonov que nous verrons ce résultat dans les cas suivants.

— $\mathbb{R}$, lemme 9.21.

— Produit fini d'espaces métriques compacts, théorème 7.314.

— Produit dénombrable d'espaces métriques compacts, théorème 7.316.

Composante connexe À propos de composantes connexes et de compacts.

(1) Si K est compact dans $\mathbb{C}$, alors la partie $\mathbb{C}\backslash K$ possède exactement une composante connexe non bornée. Lemme 26.49.

THEooPUII5BHDDC34

Thème 34 : densité

(1) Densité de $\mathbb{Q}$ dans $\mathbb{R}$, proposition 9.16.

(2) Densité des polynômes dans $\left(C^0([0,1]), \|.\|_\infty\right)$, théorème de Bernstein 36.185, ou une conséquence de Stone-Weierstrass 12.418.

(3) Densité des polynômes dans $\left(C^0(I), \|.\|_\infty\right)$ lorsque $I = [a,b]$, corolaire 36.186.

(4) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$ pour $1 \leqslant p < \infty$, théorème 27.51.

(5) Densité de $\mathscr{S}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 31.15.

(6) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 31.17.
Cela est utilisé pour le théorème de trace 31.19.

(7) Les applications étagées dans les applications mesurables (qui plus est avec limite croissante), théorème fondamental d'approximation 27.54.

(8) Les fonctions continues à support compact dans $L^2(I)$, théorème 27.55.

(9) Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$. Deux démonstrations indépendantes par le théorème 28.8 et le théorème 27.89.

Les densités sont bien entendu utilisées pour prouver des formules sur un espace en sachant qu'elles sont vraies sur une partie dense. Mais également pour étendre une application définie seulement sur une partie dense. C'est par exemple ce qui est fait pour définir la trace γ_0 sur les espaces de Sobolev $H^s(\mathbb{R}^d)$ en utilisant le théorème d'extension 17.127.

Comme presque tous les théorèmes importants, le théorème de Stone-Weierstrass possède de nombreuses formulations à divers degrés de généralité.

— Le lemme 12.414 le donne pour la racine carré.

— Le théorème 12.420 donne la densité des polynômes dans les fonctions continues sur un compact.

— Le théorème 12.417 est une généralisation qui donne la densité uniforme d'une sous-algèbre de $C(X,\mathbb{R})$ dès que X sépare les points.

— Le théorème 12.418 donne le même résultat pour la densité dans $C(X,\mathbb{C})$.

— Le lemme 28.1 est une version pour les polynômes trigonométriques.

— Le lemme 12.414 est un cas particulier du théorème 12.420, mais nous en donnons une démonstration indépendante afin d'isoler la preuve de la généralisation 12.418. Une version pour les polynômes trigonométriques sera donnée dans le lemme 28.1.

14. Si vous connaissez des exemples non métriques de produits de compacts qui ne sont pas compacts, écrivez-moi.

Le théorème de Stone-Weierstrass est utilisé, entre autres nombreuses choses, pour prouver la densité des polynômes trigonométriques dans les fonctions continues sur S^1, voir la proposition 27.106.

THEMooOXNQoo

Thème 35 : application réciproque

(1) Définition 7.212.

(2) Dans le cas des réels, des exemples sont donnés en 9.9.

(3) Continuité, théorème de la bijection 12.51.

(4) Continuité de la réciproque pour $f : K \to X$, lemme 7.222.

(5) Si f est strictement monotone sur un intervalle, son inverse est continue, théorème de la bijection 12.51.

(6) Théorème de la bijection 12.51 (qui contient aussi de la continuité).

(7) Dérivabilité, proposition 12.178.

(8) Une application continue est injective si et seulement si elle est strictement monotone, lemme 12.48.

THEMooYCBU

Thème 36 : applications continues et bornées

(1) Application continue, définition 7.36.

(2) Une application linéaire non continue : exemple 10.71 de $e_k \mapsto ke_k$. Les dérivées partielles sont calculées en (25.156).

(3) La dérivation sur les polynômes (exemple 10.72) donne un autre exemple d'application linéaire non continue.

(4) Une application linéaire est bornée si et seulement si elle est continue, proposition 10.70.

(5) Une forme sesquilinéaire est bornée si et seulement si elle est continue, proposition 25.2.

THMooOCXT

Thème 37 : suite de Cauchy, espace complet Nous parlons d'espaces topologiques complets. À ne pas confondre avec un espace mesuré complet, définition 14.68.

(1) Corps complet : définition 1.379(5), espace métrique complet : définition 7.263.

(2) L'image d'une suite de Cauchy est de Cauchy. Si (x_n) est de Cauchy, alors $\big(f(x_n)\big)$ est de Cauchy quand f est une isométrie, lemme 17.131.

(3) La définition 7.262 donne la notion de suite de Cauchy dans un espace métrique.

(4) La définition 7.260 donne la notion de suite de τ-Cauchy dans un espace vectoriel topologique.

(5) Deux espaces métriques (avec une distance) peuvent être isomorphes en tant qu'espaces topologiques, mais ne pas avoir les mêmes suites de Cauchy, exemple 7.266.

(6) La proposition 7.269 donne l'équivalence entre les suites de Cauchy et les suites τ-Cauchy dans le cas des espaces vectoriels topologiques *normés*.

(7) L'exemple 7.266 est un exemple pire que simplement une suite de Cauchy qui ne converge pas. Le problème de convergence de cette suite n'est pas simplement que la limite n'est pas dans l'espace; c'est que la suite de Cauchy donnée ne convergerait même pas dans $\mathbb{R}$.

(8) Le théorème 17.135 est un théorème de complétion d'un espace métrique.

(9) Dans $\mathbb{R}$, une suite est convergente si et seulement si elle est de Cauchy, théorème 7.287(2).

(10) Toute suite convergente dans un espace métrique est de Cauchy, proposition 7.270.

Quelques espaces qui sont complets sont listés ci-dessous. Attention : la complétude est bien une propriété de la métrique; le même ensemble peut être complet pour une distance et pas pour une autre. Souvent, cependant la distance à considérer est donnée par le contexte.

(1) Les réels $\mathbb{R}$, théorème 7.287.

(2) Les complexes $\mathbb{C}$, proposition 7.289.

(3) Un espace vectoriel normé sur un corps complet est complet, proposition 7.293.

(4) La proposition 12.359 donne quelques espaces complets. Soit X un espace topologique métrique, (Y,d) un espace métrique complet. Alors les espaces

(4a) $\left(C_b^0(X,Y),\|.\|_\infty\right)$

(4b) $\left(C_0^0(X,Y),\|.\|_\infty\right)$

(4c) $\left(C_0^k(X,Y),\|.\|_\infty\right)$

sont complets.

(5) Le lemme 12.360 dit que $\left(C^0(A,B),\|.\|_\infty\right)$ est complet dès que A est compact et B est complet.

(6) L'espace $\mathscr{D}(K)$ est complet tant pour la topologie des seminormes[15] que pour la topologie métrique (qui sont les mêmes). C'est la proposition 30.28.

(7) L'espace $\mathscr{S}(\Omega)$ est complet et métrisable par la proposition 30.80.

(8) L'espace $L^p(\Omega,\mathcal{A},\mu)$ par le théorème 27.44.

La limite uniforme d'une suite de fonctions dérivables n'est pas spécialement dérivable. Même si les fonctions sont de classe C^∞, la limite n'est pas spécialement mieux que continue. En effet, le théorème de Stone-Weierstrass 12.420 nous dit que les polynômes (qui sont C^∞) sont denses dans les fonctions continues sur un compact pour la norme uniforme. Vous ne pouvez donc pas espérer que $\left(C^p(X,Y),\|.\|_\infty\right)$ soit complet en général.

THTOC38

Thème 38 : caractérisations séquentielles Diverses caractérisations séquentielles.

continuité Des résultats de continuité séquentielle.

(1) Fonction séquentiellement continue, définition 7.208. Dans un espace métrisable séparé, la continuité séquentielle est équivalente à la continuité, proposition 7.255.

(2) Quand une topologie est donnée par un ensemble dénombrable de seminormes, il est métrisable, proposition 7.350.

(3) La continuité implique la continuité séquentielle, proposition 7.210 et corolaire 7.137.

(4) Une version spéciale pour $\mathbb{R}^m$ est donnée par le théorème 12.219.

fermeture Fermeture séquentielle, proposition 7.252.

compacité Un espace topologique séquentiellement compact : toute suite possède une sous-suite convergente, définition 7.89.

THEMEooYEVLooTHTOC39

Thème 39 : valeurs propres, définie positive

À propos de valeurs propres (1) Définition des valeurs propres d'une forme quadratique : définition 8.254.

(2) Définition de valeur propre et vecteur propre pour un endomorphisme $f\colon V\to V$, définition 8.86.

À propos de choses définies positives (1) Une application bilinéaire est définie positive lorsque $g(u,u)\geqslant 0$ et $g(u,u)=0$ si et seulement si $u=0$ est la définition 8.126.

(2) Un opérateur ou une matrice est défini positif si toutes ses valeurs propres sont positives, c'est la définition 8.230.

(3) Pour une matrice symétrique, définie positive si et seulement si $\langle Ax,x\rangle>0$ pour tout x. C'est le lemme 8.234.

(4) Une application linéaire est définie positive si et seulement si sa matrice associée l'est. C'est la proposition 8.235.

Remarque : nous ne définissons pas la notion de matrice définie positive dans le cas d'une matrice non symétrique.

15. Seminorme, définition 7.326, et la topologie des seminormes est la définition 7.342.

THEMEooOJJFoTWMOAtD

Thème 40 : norme matricielle, norme opérateur et rayon spectral Quelques définitions

(1) Définition de la norme opérateur : définition 10.55.

(2) Définition du rayon spectral 10.64.

La norme matricielle n'est rien d'autre que la norme opérateur de l'application linéaire donnée par la matrice.

(1) Lien entre norme matricielle et rayon spectral, le théorème 12.120 assure que $\|A\|_2 = \sqrt{\rho(A^tA)}$.

(2) Lien entre valeurs propres et norme opérateur : le lemme 12.121 pour les matrices symétriques strictement définies positives donne $\|A\|_2 = \lambda_{max}$.

(3) Pour une matrice diagonale, $\|D\|_2 = \max\{|\lambda_i|\}$, lemme 12.122.

(4) Pour toute norme algébrique nous avons $\rho(A) \leqslant \|A\|$, proposition 12.114.

(5) Dans le cadre du conditionnement de matrice. Voir en particulier la proposition 34.110 qui utilise le théorème 12.120.

(6) Rayon spectral et convergence de méthode itérative, proposition 34.148.

Pour la norme opérateur nous avons les résultats suivants.

(1) La majoration $\|Au\| \leqslant \|A\|\|u\|$ est le lemme 10.66.

(2) Définition d'une algèbre : 1.347 et pour une norme d'algèbre : 10.63.

(3) La norme opérateur est une norme d'algèbre : $\|AB\| \leqslant \|A\|\|B\|$, lemme 10.69.

(4) Pour des espaces vectoriels normés, être borné est équivalent à être continu : proposition 10.70.

(5) Le lemme à propos d'exponentielle de matrice 15.154 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}.$$

La norme opérateur est utilisée pour donner une norme sur les produits tensoriels, définition 10.233.

Une norme matricielle donne une topologie. Il y a donc également des liens entre rayon spectral et convergence de série. Dans cette optique, pour les séries de matrices, voir le thème 41.

Notations :

(1) $\mathcal{L}(E,F)$ est l'ensemble des applications linéaires $E \to F$.

(2) $\mathcal{L}^0(E,F)$ est l'ensemble des applications linéaires continues $E \to F$ (par analogie avec C^0). Pour définir $\mathcal{L}^0(E,F)$ il faut que E et F soient des espaces vectoriels topologiques; le plus souvent ce seront des espaces vectoriels normés.

(3) $\mathrm{End}(E)$ est l'ensemble des applications linéaires $E \to E$.

(4) $\mathrm{End}^0(E)$ est l'ensemble des applications linéaires continues $E \to E$.

THEMEooPQKDoTHMOFH

Thème 41 : série de matrices

(1) Rayon spectral et norme opérateur : thème 40.

(2) Exponentielle de matrices : thème 50.

(3) Série entière de matrices : section 15.13.

(4) Pour la série $\sum_k A^k = (1-A)^{-1}$.

— Pour un espace de Banach : proposition 10.299.

— Pour les matrices nilpotentes : proposition 8.214.

— En lien avec le rayon spectral (si et seulement si $\rho(A) < 1$) dans la proposition 15.150.

— Le lemme 15.47 parle de la série entière $\sum_{n\in\mathbb{N}} z^{nk} = (1-z^k)^{-1}$.

Cette série est utilisée entre autres dans la proposition 34.168 pour prouver qu'une M-matrice irréductible vérifie $A^{-1} > 0$.

DECooWTAI THTOC42

Thème 42 : décomposition de matrices

(1) Décomposition de Bruhat, théorème 13.39.

(2) Décomposition de Dunford, théorème 8.265.

(3) Décomposition polaire 13.32 des matrices symétriques et la proposition 17.57 pour la régularité.

THTOC43

Thème 43 : systèmes d'équations linéaires

— Algorithme des facteurs invariants 4.114.

— La méthode du gradient à pas optimal permet de résoudre par itérations $Ax = b$ lorsque A est symétrique strictement définie positive. Il s'agit de minimiser une fonction bien choisie. Propositions 17.113 pour l'existence et 17.114 pour la méthode.

THTOC44

Thème 44 : réduction, diagonalisation Des résultats qui parlent diagonalisation

(1) Définition d'un endomorphisme diagonalisable : 8.215.

(2) Conditions équivalentes au fait d'être diagonalisable en termes de polynôme minimal, y compris la décomposition en espaces propres : théorème 8.219.

(3) Diagonalisation simultanée 8.222, pseudo-diagonalisation simultanée 10.40.

(4) Diagonalisation d'exponentielle 15.132 utilisant la décomposition de Dunford.

(5) Décomposition polaire théorème 13.32. $M = SQ$, S est symétrique, réelle, définie positive, Q est orthogonale.

(6) Décomposition de Dunford 8.265. $u = s + n$ où s est diagonalisable et n est nilpotent, $[s, n] = 0$.

(7) Réduction de Jordan (bloc-diagonale) 8.295.

(8) L'algorithme des facteurs invariants 4.114 donne $U = PDQ$ avec P et Q inversibles, D diagonale, sans hypothèse sur U. De plus les éléments de D forment une chaine d'éléments qui se divisent l'un l'autre.

Le théorème spectral et ses variantes :

(1) Théorème spectral, matrice symétrique, théorème 8.227. Via le lemme de Schur complexe 12.100.

(2) Théorème spectral autoadjoint (c'est le même, mais vu sans matrices), théorème 10.8

(3) Théorème spectral hermitien, lemme 10.20.

(4) Théorème spectral, matrice normales, théorème 12.102.

Pour les résultats de décomposition dont une partie est diagonale, voir le thème 42 sur les décompositions. Réduction de quadriques :

(1) Réduction de Gauss, théorème 8.241.

Trigonalisation.

(1) Le lemme de Schur complexe 12.100 dit que toute matrice est unitairement équivalente à une matrice triangulaire suppérieure.

THMooUXJM THTOC45

Thème 45 : déterminant

(1) Déterminant d'une matrice : définition 4.79.

(2) Déterminant d'un endomorphisme 8.15.

(3) Déterminant $\det_\varphi$ d'une application linéaire entre espaces différents, proposition 10.75.

(4) Le lemme 10.5 donne la formule $\det(f) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \langle e_{\sigma(i)}, f(e_i)\rangle$.

(5) Principales propriétés algébriques du déterminant : la proposition 8.16.

(6) La formule $\det(AB) = \det(A)\det(B)$ est la proposition 8.249 pour des matrices et la proposition 8.16(1) pour les applications linéaires.

(7) Déterminant et manipulations de lignes et colonnes, section 4.3.11 et les propositions qui précèdent à partir du lemme 4.81 qui dit que $\det(A) = \det(A^t)$.

(8) Les n-formes alternées forment un espace de dimension 1, proposition 8.9.

(9) Déterminant d'une famille de vecteurs 8.10.

(10) Calcul d'un déterminant de taille 2×2 : équation (4.104).

(11) Interprétations géométriques

(11a) À propos d'orthogonalité, le déterminant est très lié au produit vectoriel en dimension 3. Et il le généralise en dimension supérieure.

(11a.i) Liaison au produit vectoriel (orthogonalité) dans la proposition 10.37.

(11a.i) En particulier le lemme 10.38 nous dit comment un déterminant donne un vecteur orthogonal à une famille donnée de vecteurs.

(11b) Déterminant et aires, volumes

(11b.ii) Déterminant et mesure de Lebesgue : théorème 14.271.

(11b.ii) Aire du parallélogramme : il y a la formule avec le produit vectoriel dans la proposition 18.55, mais l'aire proprement dite, avec une intégrale est dans la proposition 20.26.

(11b.ii) Volume du parallélépipède avec le produit mixte et le déterminant 3×3, 18.56.

Tant que nous en sommes dans les interprétations géométriques, il faut lier déterminant, produit vectoriel, orthogonalité et mesure en notant que l'élément de volume lors de l'intégration en dimension 3 est donné par (20.210) : $dS = \|T_u \times T_v\|$ qui est la norme du produit vectoriel des vecteurs tangents au paramétrage.

Nous voyons dans l'équation (20.207) que l'élément de volume pour une partie de dimension n dans $\mathbb{R}^m$ (à l'occasion d'y intégrer une fonction) est donné par un déterminant mettant en jeu les vecteurs tangents du paramétrage.

(12) Le déterminant de Vandermonde est à la proposition 8.18. Il est utilisé à divers endroits :

(12a) Pour prouver que u est nilpotente si et seulement si $\mathrm{Tr}(u^p) = 0$ pour tout p (lemme 8.211)

(12b) Pour prouver qu'un endomorphisme possédant $\dim(E)$ valeurs propres distinctes est cyclique (proposition 8.301).

THEMooNMYKoTHTeOT46

Thème 46 : espaces de fonctions En ce qui concerne les densités, voir le thème 34.

(1) $C^\infty(A)$ est l'ensemble des fonctions de classe C^∞ sur A. Les éléments de $C^\infty(A)$ peuvent être à valeurs dans $\mathbb{R}$ ou dans $\mathbb{C}$ selon le contexte.

(2) $\mathscr{D}(A)$ est l'ensemble des fonctions de classe C^∞ dont le support est un compact dans A.

Pour les ensembles $\mathscr{S}(A)$, $\mathscr{D}(A)$ et $L^p(A)$, les fonctions sont à valeurs dans $\mathbb{C}$. La raison est que, de toutes façons, le passage à la transformée de Fourier produit en général des fonctions à valeurs dans $\mathbb{C}$ même si les fonctions de départ sont à valeurs dans $\mathbb{R}$.

Topologie Les espaces de fonctions sont souvent munis de topologies définies par des seminormes [16].

(1) La topologie des seminormes est la définition 7.326.

(2) La définition 30.15 donne les topologies sur $C^\infty(\Omega)$, $\mathscr{D}(K)$ et $\mathscr{D}(\Omega)$.

(3) La topologie $*$-faible sur $\mathscr{D}'(\Omega)$ est donnée par la définition 30.31.

(4) $p_{K,m}(f) = \sum_{|\mu|\leqslant m} \|\partial^\mu f\|_K$ est une seminorme sur $C^\infty(\Omega)$, proposition 30.10.

16. Définition 7.326.

L'espace $L^2\big([0,2\pi]\big)$ C'est un espace très important, entre autres parce qu'il est de Hilbert et est bien adapté à la transformée de Fourier.

(1) Un rappel de la construction en 27.91.
(2) Le produit scalaire [17] $\langle f,g\rangle$ est donné en (27.477) et la base trigonométrique est (27.478).
(3) La densité des polynômes trigonométriques dans $L^p(S^2)$ est le théorème 27.89 ou le théorème 28.8, au choix.
(4) Une conséquence de cette densité est que le système trigonométrique est une base hilbertienne [18] de L^2 par le lemme 27.132.

L'espace L^2 est discuté en analyse fonctionnelle, dans la section 27.7 et les suivantes parce que l'étude de L^2 utilise entre autres l'inégalité de Hölder 27.33.

Le fait que L^2 soit un espace de Hilbert est utilisé dans la preuve du théorème de représentation de Riesz 27.184.

Si $(\Omega,\mathcal{A},\mu)$ est un espace mesuré, alors $L^p(\Omega,\mathcal{A},\mu)$ est un espace de Banach; c'est le théorème de Riesz-Fischer 27.45.

THTOC47

Thème 47 : fonctions Lipschitz

(1) Définition : 12.316.
(2) La notion de Lipschitz est utilisée pour définir la stabilité d'un problème, définition 34.26.
(3) Toute fonction Lipschitz est uniformément continue, proposition 12.321.

THTOC48

Thème 48 : suites et séries

Opérations sur les suites (1) Les limites $x_k+y_k\to x+y$ et $\lambda x_k\to\lambda x$ sont la proposition 7.202.
(2) La limite $x_ky_k\to xy$ est la proposition 9.27.

Suites La proposition 9.26 donne une caractérisation de limite de suite dans un espace vectoriel normé.

(1) Les suites adjacentes, c'est la définition 9.39.
(2) Les séries alternées, théorème 10.140. Il s'agit de dire que $\sum_{k=0}^{\infty}(-1)^k a_k$ converge quand a_k est décroissante et tend vers zéro.
(3) Le concept de suite adjacente sert à étudier la série de Taylor de $\ln(x+1)$, voir le lemme 15.101 et ce qui l'entoure.
(4) La définition de la convergence absolue est la définition 10.92.
(5) Une suite réelle croissante et majorée converge, proposition 9.35.
(6) Toute suite dans un compact admet une sous-suite convergente, théorème 7.290.
(7) Pour tout réel, il existe une suite croissante de rationnels qui y converge, proposition 9.17.

Produit de Cauchy (1) Dans une algèbre normée, proposition 10.107,
(2) Dans $\mathbb{C}$, théorème 15.32.

Calcul de suites (1) Somme : $x_n+y_n\to x+y$ est la proposition 9.28.

Série Les séries sont en général dans la section 10.7.

(1) Quelques séries usuelles en 10.9.3 : série harmonique, géométrique, de Riemann, et la mythique arithmético-géométrique.
(1a) La série est associative : $\sum_k(a_k+b_k)=\sum_k a_k+\sum_k b_k$. C'est la proposition 10.102.

17. Produit scalaire, définition 8.168.
18. Définition 25.27.

(1b) La série harmonique diverge : $\sum_k \frac{1}{k} = \infty$, exemple 10.133.

(1c) La série géométrique : $\sum_{k=0}^{N} q^k = \frac{1-q^{N+1}}{1-q}$, proposition 10.135.

(1d) Une autre cool série : $\sum_{k=1}^{N} \frac{1}{k(k+1)} = \frac{N}{N+1}$, lemme 10.139.

(2) Critère des séries alternées, théorème 10.140.

(3) Convergence d'une série implique convergence vers zéro du terme général, proposition 10.100.

(4) Dans une algèbre normée : $(\sum_{k=0}^{\infty} a_k)b = \sum_{k=0}^{\infty}(a_k b)$, proposition 10.106.

(5) Produit de Cauchy : théorème 15.32 et proposition 10.107.

Sommes infinies En ce qui concerne les sommes finies, la notation $\sum_{i=1}^{N}$ est définie en 1.85. Pour permuter les termes d'une somme avec un élément du groupe symétrique, nous avons la proposition 1.310.

Voici quelques résultats à propos de sommes infinies.

(1) Une somme indexée par un ensemble quelconque est la définition 10.112.

(2) La proposition 10.116 donne une caractérisation pour le sommes de réels positifs.

(3) La définition de la somme d'une infinité de termes est donnée par la définition 10.89.

(4) Une somme de termes positifs indexée par un ensemble indénombrable est toujours infinie par le lemme 10.122.

(5) si la série converge, on peut regrouper ses termes sans modifier la convergence ni la somme (associativité); pour les sommes infinies l'associativité et la commutativité dans une série sont perdues. Néanmoins, il subsiste que

(5a) si la série converge absolument, on peut modifier l'ordre des termes sans modifier la convergence ni la somme (commutativité, proposition 10.109).

(6) Permuter une somme infinie avec une application linéaire : $f(\sum_{i\in I} v_i) = \sum_{i\in I} f(v_i)$, c'est la proposition 10.127.

Série entières (1) Rayon de convergence, définition 15.12.

(2) Convervence absolue à l'intérieur du rayon de convergence, lemme d'Abel 15.18.

(3) La fonction définie par la série entière $z \mapsto \sum_{k=0}^{\infty} a_n z^n$ est holomorphe dans son disque de convergence par la proposition 15.42.

(4) La série entière pour $\frac{1}{1-z^k}$, pour $\frac{1}{\omega - z}$ et pour $\frac{1}{(\omega-z)^k}$ sont dans le lemme 15.47.

THTOC49

Thème 49 : suite de fonctions

(1) Une limite uniforme de fonctions continues est continue, proposition 12.358.

(2) Sous certaines hypothèses, si $f_i \to f$, alors $f_i' \to f'$, théorème 12.374.

(3) Si les f_k sont continues et si la convergence $\sum_k f_k$ est uniforme, alors la somme est continue, théorème 12.370.

THEMEooKXSGoBETQNo50

Thème 50 : exponentielle Toutes les exponentielles sont définies par la série

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

tant que la somme a un sens.

Réels Voici le plan que nous suivons dans le Frido :

— L'exponentielle est définie par sa série en 15.60.

— Nous démontrons qu'elle vérifie l'équation différentielle $y' = y$, $y(0) = 1$ (théorème 15.76).

— L'équation différentielle $y' = \lambda y$ est dans la proposition 15.80.
— Nous démontrons l'unicité de la solution à cette équation différentielle.
— Nous démontrons qu'elle est égale à $x \mapsto y(1)^x$. Cela donne la définition du nombre e comme valant $y(1)$.
— Nous définissons le logarithme comme l'application réciproque de l'exponentielle (définition 15.83).
— Les fonctions trigonométriques (sinus et cosinus) sont définies par leurs séries. Il est alors montré que $e^{ix} = \cos(x) + i\sin(x)$ (lemme 18.11).

Propriétés

— La formule $a^{-x} = 1/a^x$ est la proposition 12.398(3).
— $\exp(x) = e^x$, proposition 15.79.
— Nous avons l'encadrement $2.5 < e < 3$, lemme 15.81.
— Le nombre e est irrationnel, proposition 15.82.

Complexes (1) La définition de $\exp(a + ib)$ est la définition 15.60.

(2) Les principales propriétés, dont $e^{z+w} = e^z e^w$, sont dans la proposition 18.9.
(3) Nous avons $e^{ix} = e^{iy}$ si et seulement si il existe $k \in \mathbb{Z}$ tel que $y = x + 2k\pi$, corolaire 18.24.
(4) Le fait que $e^{i\theta}$ donne tous les nombres complexes de norme 1 est la proposition 18.61(1).
(5) Le groupe des racines de l'unité est donné par l'équation (19.1).

Algèbre normée commutative Pour la définition c'est la proposition 15.60 et pour la régularité C^∞ c'est la proposition 15.65.

Idem non commutatif Il y a une tentative de théorème 15.66, mais c'est principalement pour les matrices qu'il y a des résultats.

Matrices De nombreux résultats sont disponibles pour les exponentielles de matrices.

(1) $e^{sA}e^{tA} = e^{(s+t)A}$, proposition 15.70.
(2) Si A est une matrice, alors $(e^tA)'(u) = Ae^{uA}$, proposition 15.72.
(3) Les sections 10.14 et 15.5.3 parlent d'exponentielle de matrices.
(4) L'exponentielle donne lieu à une fonction de classe C^∞, proposition 15.151.
(5) Le lemme à propos d'exponentielle de matrice 15.154 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}.$$

(6) La proposition 15.132 : si $A \in \mathbb{M}(n, \mathbb{R})$ a un polynôme caractéristique scindé, alors A est diagonalisable si et seulement si e^A est diagonalisable.
(7) La section 15.13.3 parle des fonctions exponentielle et logarithme pour les matrices. Entre autres la dérivation et les séries.
(8) Pour résoudre des équations différentielles linéaires : sous-section 32.6.1.
(9) La proposition 15.131 dit que l'exponentielle est surjective sur $\mathrm{GL}(n, \mathbb{C})$.
(10) La proposition 10.302 : si u est un endomorphisme, alors $\exp(u)$ est un polynôme en u.
(11) Calcul effectif : sous-section 15.13.4.
(12) Proposition 13.23 : si $A \in \mathbb{M}(n, \mathbb{C})$ alors $e^{\mathrm{Tr}(A)} = \det(e^A)$.
(13) Les séries entières de matrices sont traitées autour de la proposition 15.148.

Paramétrisation du cercle (1) La proposition 18.61 parle des propriétés de

$$\begin{aligned} \varphi\colon]0, 2\pi[&\to S^1 \\ t &\mapsto e^{it}. \end{aligned} \tag{-1.3}$$

THTOC51

Thème 51 : logarithme

(1) Le logarithme pour les réels strictement positifs $\ln\colon]0,\infty[\to \mathbb{R}$ est donné en la définition 15.83; c'est l'application réciproque de exp.

(2) Les principales propriétés sont dans la proposition 15.86 : $\ln(xy) = \ln(x) + \ln(y)$ etc.

(3) Dérivée : $\ln'(x) = \frac{1}{x}$, proposition 15.85.

(4) La proposition 15.102 donne la série

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k. \tag{-1.4}$$

(5) L'exemple 20.152 donne l'encadrement $0.644 \leqslant \ln(2) \leqslant 0.846$.

(6) La proposition 15.131 dit que toute matrice complexe admet un logarithme. En particulier une série explicite est donnée pour le logarithme d'un bloc de Jordan [19].

(7) Sur les complexes, le logarithme $\ln\colon \mathbb{C}^* \to \mathbb{C}$ est la définition 26.29. Attention : ce n'est pas la seule définition possible.

(8) La série harmonique diverge à vitesse logarithmique, et la série des inverses des nombres premiers, c'est encore plus lent : théorème 15.120.

(9) Détermination du logarithme le long d'un chemin dans $\mathbb{C}$, définition 26.45.

(10) Un logarithme continu d'une fonction, définition 26.44.

(11) Théorème de Borsuk 26.47 : il y a toujours un logarithme continu le long d'un chemin homotope à une constante (et il y a même équivalence avec admettre une extension sur $\mathbb{C}^*$).

(12) Si $z \in \mathbb{C}$, nous avons $\int \frac{1}{x+z} dx = \ln(x+z)$, proposition 26.42.

(13) La proposition 26.58 dit qu'une application f a un logarithme continu si et seulement si $\mathrm{Ind}(f \circ \gamma, 0) = 0$.

THEMEooBSBLoTWTAQ52

Thème 52 : fonction puissance Il y a beaucoup de choses à dire...

Définition Nous considérons, pour $a > 0$, la fonction $g_a\colon \mathbb{R} \to \mathbb{R}$ donnée par $g_a(x) = a^x$. La définition de cette fonction se fait en de nombreuses étapes.

(1) a^n pour $n \in \mathbb{N}$ en la définition 1.206.

(2) a^n pour $n \in \mathbb{Z}$ en la définition 12.378.

(3) $a^{1/n}$ pour $n \in \mathbb{Z}$ en la définition 12.381.

(4) a^q pour $q \in \mathbb{Q}$ en la définition 12.381.

(5) $\sqrt[n]{x}$ en la définition 12.383.

(6) La fonction g_a est Cauchy-continue sur $\mathbb{Q}$, c'est la proposition 12.395.

(7) a^x pour $a > 0$ et $x \in \mathbb{R}$ en la définition 12.397.

(8) a^z pour $a > 0$ et $z \in \mathbb{C}$ en la définition 18.5.

Quelques propriétés Quelques propriétés.

(1) Si b, m sont des entiers, il existe $n \in \mathbb{N}$ tel que $b^n > m$, proposition 1.208.

(2) Pour tout $q \in \mathbb{Q}$, il y a un $\sqrt{q}$ dans $\mathbb{R}$, proposition 1.474.

(3) Pour $a > 0$ et $x, y \in \mathbb{R}$ nous avons $a^x a^y = a^{x+y}$, proposition 12.398.

(4) Si $a > 0$ et $x, y \in \mathbb{R}$ nous avons $(a^x)^y = (a^y)^x = a^{xy}$ par la proposition 12.408.

(5) Si $\alpha > 0$, alors $x \mapsto x^\alpha$ est strictement croissante pour $x \geqslant 0$, proposition 12.403.

(6) La formule $a^{-x} = 1/a^x$ est la proposition 12.398(3).

19. Jordan, théorème 8.295.

(7) La fonction puissance $g_a(x) = a^x$ est continue, proposition 12.398.

Croissance (1) La fonction puissance $f_\alpha(x) = x^\alpha$ est strictement croissante pour les $x > 0$, proposition 12.403

(2) La fonction puissance $g_a(x) = a^x$ est strictement croissante, proposition 12.398.

(3) $\lim_{x\to\infty} x^\alpha = \infty$, proposition 12.409.

Continuité, Dérivation Comme toutes les choses sur la fonction puissance, les preuves sont assez différentes selon que l'on parle de a^x ou de x^a.

(1) La fonction $f_\alpha(x) = x^\alpha$ est continue, proposition 12.407.

(2) La fonction a^x est dérivable et sa dérivée vérifie $g'_a(x) = g_a(x)g'_a(0)$, proposition 12.425.

(3) La formule de dérivation pour $x \mapsto x^q$ avec $q \in \mathbb{Q}$ est la proposition 12.435.

(4) La dérivation de $x \mapsto x^\alpha$ avec $\alpha \in \mathbb{R}$ est la proposition 14.294. Si elle est tellement loin, c'est parce qu'elle nécessite de permuter une limite de fonctions avec une dérivée.

(5) Pour la formule générale de dérivation de $x \mapsto a^x$ demande de savoir les logarithmes (proposition 15.95).

L'équation fonctionnelle L'exponentielle et plus généralement la fonction puissance $g_a(x) = a^x$ peut être introduite au moyen d'une équation fonctionnelle au lieu de l'équation différentielle usuelle. Cette fameuse équation fonctionnelle est

$$f(x + y) = f(x)f(y) \tag{-1.5}$$

en la définition 12.427.

(1) L'équivalence entre l'équation fonctionnelle et l'équation différentielle est donnée par la proposition 12.433.

(2) La fonction $g_a(x) = a^x$ vérifie l'équation fonctionnelle $g_a(x + y) = g_a(x)g_a(y)$ et les conséquences. C'est la définition 12.427 et les choses qui suivent.

(3) L'équation fonctionnelle pour une fonction continue $f\colon \mathbb{R} \to S^1$ est traitée dans la proposition 12.437.

Une définition alternative de la fonction puissance serait de poser directement

$$a^x = e^{x\ln(a)}.$$

De là les propriétés se déduisent facilement. Dans cette approche, les choses se mettent dans l'ordre suivant :

— Définir $\exp(x)$ par sa série pour tout x.
— Démontrer que $\exp(q) = \exp(1)^q$ pour tout rationnel q (première partie de la proposition 15.79).
— Définir $e = \exp(1)$.
— Définir, pour x irrationnel, $a^x = \exp(x\ln(a))$.
— Prouver que $e^x = \exp(x)$ pour tout x.

THEMEooMKLBooTOCxx

Thème 53 : sommation finie et infinie La définition du symbole $\sum_{i\in I}$ se fait en trois étapes et deux demi, chacune se basant sur la précédente.

(1) Si $(A, +)$ est un ensemble avec une loi de composition interne, $\sum_{i=0}^n a_i$ est en 1.85.

(2) Si I est fini et si est à valeur dans un groupe commutatif, $\sum_{i\in I} f(i)$ est 1.309.

(3) Enfin si I est un ensemble quelconque, la définition 10.112 introduit la notion de famille sommable dans un espace vectoriel normé.

(4) Si (a_k) est une suite dans un espace vectoriel normé, la somme $\sum_{k=0}^\infty a_k$ est avec les sommes partielles dans la définition 10.89.

(5) La somme au sens de Cesàro est la somme des moyennes partielles, définition 10.141.

Notez que $\sum_{k=0}^{\infty} a_k$ n'est pas un cas particulier de $\sum_{k\in\mathbb{N}} a_k$. Une différence de taille entre les deux est que pour que $\sum_{k=0}^{\infty} a_k$ existe, il suffit que les a_k puissent être sommés dans cet ordre. À l'inverse pour que $\sum_{k\in\mathbb{N}} a_k$ existe, il faut que l'ordre de sommation puisse être arbitraire.

Si vous voulez sommer des séries encore moins convergentes, vous pouvez avoir envie d'utiliser la supersomme[2], ou la η-régularisation[3].

THTOC54

Thème 54 : polynôme de Taylor

Énoncés Il existe de nombreux énoncés du théorème de Taylor, et en particulier beaucoup de formules pour le reste.

(1) Énoncé : théorème 12.461.

(2) Une majoration du reste est dans le théorème 15.53

(3) De classe C^2 sur $\mathbb{R}^n$, proposition 12.468.

(4) Avec un reste donné par un point dans $]x, a[$, proposition 12.472.

(5) Avec reste intégral, théorème 20.149, et théorème 20.147 pour le cas simple $\mathbb{R} \to \mathbb{R}$.

(6) Le polynôme de Taylor généralise à l'utilisation de toutes les dérivées disponibles le résultat de développement limité donné par la proposition 10.257.

(7) Pour les fonctions holomorphes, il y a le théorème 26.76 qui donne une série de Taylor sur un disque de convergence.

Utilisation Des polynômes de Taylor sont utilisés pour démontrer des théorèmes par-ci par-là.

(1) Il est utilisé pour justifier la méthode de Newton autour de l'équation (34.95).

(2) On utilise pas mal de Taylor dans les résultats liant extrémum et différentielle/hessienne. Par exemple la proposition 17.70.

Quelques développements Voici quelques développements limités à savoir. Ils sont calculables en utilisant la formule de Taylor-Young (proposition 12.477).

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + x^n\alpha(x) \qquad \text{ordre } n, \text{proposition 15.97}$$

$$\cos(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k}}{(2k)!} + x^{2p+1}\alpha(x) \qquad \text{ordre } 2p+1, \text{proposition 18.75}$$

$$\sin(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + x^{2p+2}\alpha(x) \qquad \text{ordre } 2p+1, \text{proposition 18.75}$$

$$\ln(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + \alpha(x)x^n \qquad \text{ordre } n, \text{proposition 15.100}$$

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \qquad \text{exact proposition 15.102}$$

$$\ln(2) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \qquad \text{exact proposition 15.102}$$

$$(1+x)^l = \sum_{k=0}^{l} \binom{l}{k} x^k \qquad \text{exact si } l \text{ est entier.}$$

$$(1+x)^l = 1 + \sum_{k=1}^{n} \frac{l(l-1)\dots(l-k+1)}{k!} x^k + x^n\alpha(x) \qquad \text{ordre } n.$$

Dans toutes ces formules, la fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ vérifie $\lim_{t\to 0} \alpha(t) = 0$.

Le développement limité en 0 d'une fonction paire ne contient que les puissances de x d'exposant pair. Voir comme exemple le développement de la fonction cosinus.

INTERNooXFNT5BNTOC55

Thème 55 : formule des accroissements finis Il en existe plusieurs formes :

(1) Une version adaptée aux espaces normés de dimension finie, avec hypothèse de différentiabilité, est le théorème 12.314. La formule $\|f(b)-f(a)\|_n \leqslant \sup_{x\in[a,b]} \|df_x\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|b-a\|_m$.

(2) Une version pour les dérivées partielles est dans le lemme 12.237. Pour rappel, la définition de la dérivation partielle est 10.244.

(3) La formule $f(a+\epsilon e_i) = f(a) + \epsilon(\partial_i f)(a) + \epsilon\alpha(\epsilon)$, proposition 12.238.

(4) L'existence de $c \in]a,b[$ tel que
$$f'(c) = \frac{f(b)-f(a)}{b-a} \tag{-1.7}$$
est le théorème des accroissements finis proprement dit. C'est le théorème 12.192.

(5) Il existe un c entre a et b tel que
$$f(b) = f(a) + (b-a)(\partial_\beta f)(c) \tag{-1.8}$$
où $\beta = b-a$ est la proposition 12.236.

(6) La formule $f(a+h) = f(a)+hf'(a)+\alpha(h)$ pour une fonction $\mathbb{R}\to\mathbb{R}$ en le théorème 10.257.

(7) Une généralisation pour les intervalles non bornés : théorème 12.193.

(8) Espaces vectoriels normés, théorème 10.292

(9) La formule
$$f(a+h) = f(a) = df_a(h) + \alpha(h) \tag{-1.9}$$
est le lemme 10.271.

THTOC56

Thème 56 : dérivation

(1) Définition de la dérivée, définition 10.245.

(2) Dérivée de fonction composée, proposition 12.181 dans le cas réel.

(3) Dérivée partielle de fonction composée, théorème 12.301.

(4) La formule $(f\circ\gamma)'(a) = (\nabla f)\big(\gamma(a)\big)\cdot\gamma'(a)$, corolaire 12.302.

(5) $f(\lambda x)' = \lambda f'(x)$, lemme 12.176.

THTOC57

Thème 57 : différentielle

Généralités (1) La différentielle est définie en général pour des espaces vectoriels normés par la proposition 10.243

(2) Différentielle d'une application linéaire, lemme 10.269.

(3) Nous parlons de différentielle en dimension finie et donnons une interprétation géométrique en 12.22.1.

(4) La recherche d'extrémums d'une fonction sur $\mathbb{R}^n$ passe par la seconde différentielle, proposition 17.70.

(5) Lien entre différentielle seconde (hessienne) et convexité en la proposition 17.97 et le corolaire 17.98.

(6) La différentielle est liée aux dérivées partielles par les formules données au lemme 12.255
$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a)\cdot u. \tag{-1.10}$$

(7) La matrice de df en a est $(df_a)_{ij} = df_a(e_j)\cdot e_i = (\partial_j f_i)(a)$, proposition 12.256.

Différentielle et dérivées partielles À propos de fonctions de classe C^k, définition 10.261.

(1) Une fonction est de classe C^1 si et seulement si ses dérivées partielles sont continues, théorème 12.294.

(2) Une fonction est C^n si et seulement si ses dérivées partielles sont C^{n-1}, c'est le théorème 12.330.

(3) Différentiabilité en un seul point si les dérivées partielles sont continues en ce point : proposition 12.292.

Fonctions composées À propos de la formule $d(f\circ g)_a = dg_{f(a)}\circ df_a$, il y a deux théorèmes très semblables.

(1) Le théorème 10.276 insiste sur des hypothèses locales.

(2) Le théorème 10.277 fait des hypothèses plus globales pour s'alléger l'esprit, mais fait une récurrence pour dire que $f\circ g$ est de classe C^r si f et g le sont.

(3) Dans le cas d'espaces vectoriels normés de dimension finie, proposition 10.280.

(4) Composée avec une application linéaire, propositions 10.279 et 10.278.

Applications à valeurs dans les endomorphismes (1) $A\colon U \to \mathrm{End}(V,W)$ est C^p si ses composantes le sont, proposition 12.334.

(2) $A\colon U \to \mathrm{End}(V,W)$ est C^p si $s\mapsto A(s)v$ est C^p pour tout v, proposition 12.334

THTOC58

Thème 58 : équations différentielles L'utilisation des théorèmes de point fixe pour l'existence de solutions à des équations différentielles est fait dans le chapitre sur les points fixes.

(1) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 20.34 pour les équations différentielles.

(2) Le théorème de Schauder 20.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 17.39) sans la condition Lipschitz

(3) Le théorème de Cauchy-Lipschitz 17.39 est utilisé à plusieurs endroits :

— Pour calculer la transformée de Fourier de $e^{-x^2/2}$ dans le lemme 29.22.

(4) Théorème de stabilité de Lyapunov 32.43.

(5) Le système proie-prédateur de Lotka-Volterra 32.44

(6) Équation de Schrödinger, théorème 32.51.

(7) L'équation $(x-x_0)^\alpha u = 0$ pour $u\in\mathscr{D}'(\mathbb{R})$, théorème 30.51.

(8) La proposition 32.46 donne un résultat sur $y''+qy=0$ à partir d'une hypothèse de croissance.

(9) Équation de Hill $y''+qy=0$, proposition 32.48.

THTOC59

Thème 59 : convexité

Fonctions convexes L'essentiel des résultats sur les fonctions convexes sont dans la section 17.12. On a surtout :

(1) Définition des fonctions convexes : 17.76 et 17.94 en dimension supérieure.

(2) En termes de différentielles, 17.95 pour la différentielle première et 17.98 pour la hessienne.

(3) Une courbe paramétrée convexe est la définition 21.90.

(4) L'enveloppe convexe d'une courbe fermée simple et convexe : 21.92.

(5) Courbure et convexité d'une courbe paramétrée : section 21.13.4.

(6) Une courbe paramétrée convexe est localement le graphe d'une fonction convexe par le lemme 21.91.

(7) La convexité est utilisée dans la méthode du gradient à pas optimal de la proposition 17.114.

(8) La fonction $t\mapsto t^p$ est strictement convexe sur les positifs dans le lemme 17.87.

(9) Une inégalité sympa : $a^r + b^r \leqslant (a+b)^r \leqslant 2^{r-1}(a^r + b^r)$ pour $a, b > 0$ et $r > 1$, lemme 27.136.

En termes de parties convexes, on a :

Parties convexes (1) Définition 7.154 d'une partie convexe d'un espace vectoriel.

(2) Une boule est convexe, proposition 11.28.

(3) Un polygone convexe est défini en 18.168, et les racines de l'unité forment un polygone convexe par la proposition 18.169.

THTOC60

Thème 60 : espaces de Hilbert, base hilbertienne Beacoup de choses concerant les espaces L^2, et en particulier la base trigonométrique sont dans le thème 46.

(1) Toute partie orthonormale d'un espace de Hilbert est libre, proposition 25.24.

(2) Tout espace de Hilbert possède une base hilbertienne, proposition 25.37.

(3) Unicité de la décomposition dans une base hilbertienne, proposition 25.38.

THTOC61

Thème 61 : analyse complexe, fonctions holomorphes

(1) Le lien entre différentielle et dérivée complexe est donné par les équations

$$df_{z_0}(z) = f'(z_0)z = (\partial_z f)(z_0)z \tag{-1.11}$$

par (12.838) et la proposition 26.4. Cela se résume par la formulation lapidaire $f' = \partial_z f$.

(2) Série de Laurent $f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$, théorème 28.26.

THEMEooJGEHooTHZQOC62

Thème 62 : permuter des limites

Permuter des dérivées partielles Pour permuter les dérivées partielles, la clef est d'être C^k.

(1) Si une fonction est de classe C^2, le théorème de Schwarz 12.343 dit que

$$\partial_k \partial_l f = \partial_l \partial_k f. \tag{-1.12}$$

(2) Pour une fonction de classe C^m, on peut permuter m dérivées partielles avec la proposition 12.352.

Fonctions définies par une intégrale Les théorèmes sur les fonctions définies par une intégrale, section 17.4. Nous avons entre autres

(1) $\partial_i \int_B f = \int_B \partial_i f$, avec B compact, proposition 17.24.

(2) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 17.15 pour la continuité de $x \mapsto \int_\Omega f(x,\omega)d\mu(\omega)$.

(3) Pour la fonction $F(x) = \int_\Omega f(x,\omega)d\mu(\omega)$, nous avons la dérivation sous l'intégrale par la formule de Leibniz

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\mu(\omega) \tag{-1.13}$$

démontrée en le théorème 17.19 ou 17.23.

Des variations avec des dérivées partielles et des différentielles sont dans 17.24 et dans 17.25.

(4) Si $f\colon \mathbb{C}\times\Omega \to \mathbb{C}$ est holomorphe (pour $\mathbb{C}$), alors F est holomorphe et

$$F'(z) = \int_\Omega \frac{\partial f}{\partial z}(z,\omega)d\mu(\omega). \tag{-1.14}$$

(5) Pour des dérivées partielles multiples, nous avons la formule

$$(\partial^\alpha F)(x) = \int_\Omega (\partial^\alpha f_\omega)(a)d\mu(\omega) \tag{-1.15}$$

dans la proposition 17.21.

(6) Si l'intégrale est uniformément convergente, nous avons le théorème 17.16 qui donne la continuité de $F(x) = \int_\Omega f(x,\omega)d\mu(\omega)$.

(7) Pour dériver $\int_B g(t,z)dt$ avec B compact dans $\mathbb{R}$ et $g\colon \mathbb{R}\times\mathbb{C}\to\mathbb{C}$, il faut aller voir la proposition 26.64.

(8) En ce qui concerne le x dans la borne, le théorème 14.286 lie primitive et intégrale en montrant que $F(x) = \int_a^x f(t)dt$ est une primitive de f (sous certaines conditions). Le théorème fondamental de l'analyse 14.287 en est une conséquence.

(9) Si T est une **distribution**, alors nous avons

$$T\big(x\mapsto(\partial_y^\alpha\phi)(x,y_0)\big) = \partial_y^\alpha\Big(T\big(x\mapsto\phi(x,y)\big)\Big)_{y=y_0}. \tag{-1.16}$$

C'est la proposition 30.55.

Dérivée et intégrale (1) Dériver $\int_\Omega f(x,\omega)d\omega$, théorème 17.23.

Limite et intégrale (1) Théorème de la convergence monotone, théorème 14.178.

(2) $\int\sum_n f_n = \sum_n\int f_n$ dans le corolaire 14.181.

Fubini Le théorème de Fubini permet non seulement de permuter des intégrales, mais également des sommes parce que ces dernières peuvent être vues comme des intégrales sur $\mathbb{N}$ muni de la tribu des parties et de la mesure de comptage[20]. Nous utilisons cette technique pour permuter une somme et une intégrale dans l'équation (26.297).

L'utilisation de Fubini pour permuter des intégrales (sur deux variables différentes) ou deux sommes est expliquée dans 14.305.

C'est par exemple utilisé pour permuter deux sommes dans le cadre des chaines de Markov en 38.8.

— le théorème de Fubini-Tonelli 14.301 demande que la fonction soit mesurable et positive;

— le théorème de Fubini 14.304 demande que la fonction soit intégrable (mais pas spécialement positive);

— le corolaire 14.303 demande l'intégrabilité de la valeur absolue des intégrales partielles pour déduire que la fonction elle-même est intégrable.

Limite et dérivées, différentielle (1) Permuter limite et dérivée dans le cas $\mathbb{R}\to\mathbb{R}$, théorème 12.374.

(2) Permuter limite et dérivées partielles, théorème 12.377.

(3) Permuter limite et différentielle, théorème 15.9.

Quelques remarques sur les techniques de démonstration.

(1) Le résultat fondamental 12.374 est démontré sans recourir à des intégrales. Une preuve alternative, plus courte, avec des intégrales est donnée en 14.329.

(2) Permuter limite et dérivée partielle, théorème 12.377.

(3) Permuter série et différentielle, théorème 15.9.

Somme et dérivée Permuter somme et différentielle, théorème 15.9.

Limite et mesure Une mesure n'est pas toujours une limite, mais la définition d'une mesure positive sur un espace mesurable parle de permuter limite et mesure : définition 14.18(2).

THEMEooDASVoTWOOfr

Thème 63 : déduire la nullité d'une fonction depuis son intégrale Des résultats qui disent que si $\int f = 0$ c'est que $f = 0$ dans un sens ou dans un autre.

(1) Il y a le lemme 14.199 qui dit ça.

(2) Un lemme du genre dans L^2 existe aussi pour $\int f\varphi = 0$ pour tout φ. C'est le lemme 27.76.

(3) Et encore un pour L^p dans la proposition 27.189.

20. Mesure de comptage, définition 14.263.

(4) Si $\int f\chi = 0$ pour tout χ à support compact alors $f = 0$ presque partout, proposition 30.1.

(5) Si $\int_a^c f = 0$ pour tout c, alors $f = 0$, proposition 14.325.

(6) En utilisant le théorème de représentation de Riesz, on peut prouver que $\int_\Omega f\chi = 0$ implique $f = 0$ pour tout $f \in L^p$, proposition 27.189.

(7) La proposition 27.21 donne $f = 0$ dans L^p lorsque $\int fg = 0$ pour tout $g \in L^q$.

(8) Une fonction $h \in C_c^\infty(I)$ admet une primitive dans $C_c^\infty(I)$ si et seulement si $\int_I h = 0$. Théorème 17.2.

Dans le même ordre d'idées, si $f > 0$ et si $\mu(X) > 0$ alors $\int_X f > 0$ par le lemme 14.201.

THTOC64

Thème 64 : inversion locale, fonction implicite

(1) Théorème d'inversion locale dans un Banach : théorème 17.47.

(2) Fonction implicite dans un Banach : théorème 17.48.

(3) Utilisé pour montrer que le flot d'une équation différentielle est un C^p-difféomorphisme local, voir 32.36.

(4) Pour le théorème de Von Neumann 17.61.

THEMEooWAYJoTHTOC65

Thème 65 : points fixes

(1) Il y a plusieurs théorèmes de points fixes.

Théorème de Picard 17.34 donne un point fixe comme limite d'itérés d'une fonction Lipschitz. Il aura pour conséquence le théorème de Cauchy-Lipschitz 17.39, l'équation de Fredholm, théorème 17.38 et le théorème d'inversion locale dans le cas des espaces de Banach 17.47.

Théorème de Brouwer qui donne un point fixe pour une application d'une boule vers elle-même. Nous allons donner plusieurs versions et preuves.

(1a) Dans $\mathbb{R}^n$ en version C^∞ via le théorème de Stokes, proposition 20.30.

(1b) Dans $\mathbb{R}^n$ en version continue, en s'appuyant sur le cas C^∞ et en faisant un passage à la limite, théorème 20.32.

(1c) Dans $\mathbb{R}^2$ via l'homotopie, théorème 26.66. Oui, c'est très loin. Et c'est normal parce que ça va utiliser la formule de l'indice qui est de l'analyse complexe [21].

Théorème de Markov-Kakutani 20.36 qui donne un point fixe à une application continue d'un convexe fermé borné dans lui-même.

Théorème de Schauder C'est une version valable en dimension infinie du théorème de Brouwer. Théorème 20.33

(2) Pour les équations différentielles

(2a) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 20.34 pour les équations différentielles.

(2b) Le théorème de Schauder 20.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 17.39) sans la condition Lipschitz, mais alors sans unicité de la solution. Notons que de ce point de vue nous sommes dans la même situation que la différence entre le théorème de Brouwer et celui de Picard : hors hypothèse de type «contraction», point d'unicité.

(3) En calcul numérique

— La convergence d'une méthode de point fixe est donnée par la proposition 34.48.

— La convergence quadratique de la méthode de Newton est donnée par le théorème 34.54.

— En calcul numérique, section 34.5

21. On aime bien parce que ça ne demande pas Stokes, mais quand même hein, c'est pas gratos non plus.

— Méthode de Newton comme méthode de point fixe, sous-section 34.6.2.

(4) D'autres utilisations de points fixes.

— Processus de Galton-Watson, théorème 38.59.

— Dans le théorème de Max-Milgram 25.61, le théorème de Picard est utilisé.

THTOC66

Thème 66 : changement de variables Il n'existe rien en mathémaitque qui s'appelle «changement de variables». Il n'existe que des compositions de fonctions. Ce snobisme terminologique étant, voici un certain nombre de résultats de changement de variables.

(1) Dans des intégrales, théorème 14.281.

(2) Dans des limites, le lemme 7.177 donne $\lim_{x\to a} f(x) = \lim_{x\to b} f(x+a-b)$ si la limite existe.

(3) Dans une équation aux dérivées partielles, exemple 33.132.

(4) Limite de fonction composée $\lim_{x\to a}(f\circ g)$, propositions 7.178 et 7.179.

THEMEooLTCI66THTOC67

Thème 67 : techniques de calcul Il y en a pour tous les gouts.

Primitives et intégrales Toute la section 18.2 donne des trucs et astuces pour trouver des primitives et des intégrales.

Limite à deux variables Les exemples de limites à plusieurs variables font souvent intervenir des coordonnées polaires (du théorème 18.227) ou autres fonctions trigonométriques. Ils sont donc placés beaucoup plus bas que la théorie.

— Méthode du développement asymptotique, sous-section 18.14.2.

— Méthode des coordonnées polaires, sous-section 18.14.1.

— Utilisation du théorème de la fonction implicite, dans l'exemple 18.229.

THTOC68

Thème 68 : méthodes de calcul

(1) Théorème de Rothstein-Trager 20.97.

(2) Algorithme des facteurs invariants 4.114.

(3) Méthode de Newton, théorème 34.63

(4) Calcul d'intégrale par suite équirépartie 28.10.

THTOC69

Thème 69 : méthode de Newton

(1) Nous parlons un petit peu de méthode de Newton en dimension 1 dans 34.6.

(2) La méthode de Newton fonctionne bien avec les fonctions convexes par la proposition 34.56.

(3) La méthode de Newton en dimension n est le théorème 34.63.

(4) Un intervalle de convergence autour de α s'obtient par majoration de $|g'|$, proposition 34.48.

(5) Un intervalle de convergence quadratique s'obtient par majoration de $|g''|$, théorème 34.54.

(6) En calcul numérique, section 34.6.

(7) Méthode de Newton pour calculer $\sqrt{A}$, exemple 34.57.

THTOC70

Thème 70 : prolongement d'applications

(1) Prolongement de fonction définie sur une partie dense, théorème 17.128

(2) Lemme de Borel 15.168.

(3) Prolongement méromorphe de la fonction Γ d'Euler.

(4) Le théorème de Tietze prolonge des fonction continues définies sur un fermé. Prolongement continu dans le cas métrique, théorème 27.178. Dans le cas d'un espace normal, théorème 26.21.

THTOC71

Thème 71 : opérations sur les distributions

(1) Convolution d'une distribution par une fonction, définition par l'équation (30.227).

(2) Dérivation d'une distribution, proposition-définition 30.37.

(3) Produit d'une distribution par une fonction, définition 30.36.

THTOC72

Thème 72 : convolution

(1) Définition 27.65, et principales propriétés sur $L^1(\mathbb{R})$ dans le théorème 27.66.

(2) Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 27.70.

(3) $\varphi \in L^1(\mathbb{R})$ et $\psi \in \mathscr{S}(\mathbb{R})$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R})$, proposition 27.245.

(4) Les suites régularisantes : $\lim_{n\to\infty} \rho_n * f = f$ dans la proposition 29.18.

(5) Convolution d'une distribution par une fonction, définition par l'équation (30.227).

(6) Nous avons $(f * g)' = f * g'$, proposition 27.71.

(7) Si $f \in L^1(\mathbb{R}^d)$ et $g \in C^\infty(\mathbb{R}^d)$, alors $f * g \in C^\infty$, corolaire 27.72.

(8) Somme de variables aléatoires : $f_{x+Y} = f_X * f_Y$, proposition 36.42.

THMooHWEB36HMD673

Thème 73 : séries de Fourier

— Le système trigonométrique est donné en la définition 27.84.

— Les coefficients de Fourier de $c_n(f)$ sont donnés par 27.86.

— Formule sommatoire de Poisson, proposition 29.15.

— Inégalité isopérimétrique, théorème 28.23.

— Fonction continue et périodique dont la série de Fourier ne converge pas, proposition 28.21.

— Nous allons montrer la convergence de $\sum_{k\in\mathbb{Z}} c_k(f)e^{inx}$ vers $f(x)$ dans divers cas :

  (1) Si f est continue et périodique, convergence au sens de Cesàro, théorème de Fejér 28.7.

  (2) Convergence au sens $L^2\Big([0,2\pi]\Big)$ dans le théorème 27.133.

  (3) Si f est continue, périodique et si sa série de Fourier converge uniformément, théorème 28.15.

  (4) Si f est périodique et la série des coefficients converge absolument pour tout x, proposition 28.16.

  (5) Si f est périodique et de classe C^1, théorème 28.17.

  (6) Unicité des coefficients de Fourier, corolaire 28.18.

  Il est cependant faux de croire que la continuité et la périodicité suffisent à obtenir une convergence, comme le montre la proposition 28.21.

THTOC74

Thème 74 : transformée de Fourier

(1) Définition sur L^1, définition 29.1.

(2) La transformée de Fourier d'une fonction $L^1(\mathbb{R}^d)$ est continue, proposition 29.11.

(3) L'espace de Schwartz [22] est stable par transformée de Fourier. L'application $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ est continue. Proposition 29.20

(4) L'application $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ est une bijection. Formule d'inversion, proposition 29.26.

22. Définition 27.233.

THTOC75

Thème 75 : gaussienne

(1) Le calcul de l'intégrale

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

est fait de deux façons dans l'exemple 14.307. Dans les deux cas, le théorème de Fubini 14.304 est utilisé.

(2) Le lemme 29.22 calcule la transformée de Fourier de $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ qui donne $\hat{g}_\epsilon(\xi) = \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-\|\xi\|^2/4\epsilon}$.

(3) Le lemme 29.23 donne une suite régularisante à base de gaussienne.

(4) Elle est utilisée pour régulariser une intégrale dans la preuve de la formule d'inversion de Fourier 29.26

THEMEooJREIooTHTOC76

Thème 76 : lemme de transfert Il y a deux résultats qui portent ce nom. Le premier est dans la théorie de Fourier, le résultat $\hat{f}' = i\xi\hat{f}$.

(1) Lemme 29.19 sur $\mathscr{S}(\mathbb{R}^d)$

(2) Lemme 31.10 pour L^2.

L'autre lemme de transfert est en théorie des tribus, le résultat $\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big)$ du lemme 14.48. Celui-ci est d'ailleurs plutôt nommé «lemme de transport».

Il existe aussi un théorème de transfert 36.96 qui parle de variables aléatoires :

$$E(f \circ X) = \int_\Omega f\big(X(\omega)\big) dP(\omega) = \int_{\mathbb{R}^d} f(x) dP_X(x) \tag{-1.17}$$

THTOC77

Thème 77 : invariants de similitude

(1) Théorème 8.292.

(2) Pour prouver que la similitude d'applications linéaires résiste à l'extension du corps de base, théorème 8.306.

(3) Pour prouver que la dimension du commutant d'un endomorphisme de E est de dimension au moins $\dim(E)$, lemme 8.303.

(4) Nous verrons dans la remarque 8.293 à propos des invariants de similitude que toute matrice est semblable [23] à la matrice bloc-diagonale constituées des matrices compagnon (définition 8.287) de la suite des polynômes minimaux.

THMooVUCLooTHTOC78

Thème 78 : isométries Il y a $(\mathbb{R}^n, \|.\|)$ et $(\mathbb{R}^n, d)$.

Les isométries de $\|.\|$ sont linéaires, tandis que les isométries de la distance contiennent aussi les translations et les rotations de centre différent de l'origine.

Ne pas confondre une isométrie d'un espace affine avec une isométrie d'un espace euclidien [24]. Les isométries d'un espace euclidien préservent le produit scalaire et fixent donc l'origine (lemme 10.17). Les isométries des espaces affines par contre conservent les distances (définition 11.64) et peuvent donc déplacer l'origine de l'espace vectoriel sur lequel il est modelé; typiquement les translations sont des isométries de l'espace affine mais pas de l'espace euclidien.

Parfois, lorsqu'on coupe les cheveux en quatre, il faut faire attention en parlant de $\mathbb{R}^n$: soit on en parle comme d'un espace métrique (muni de la distance), soit on en parle comme d'un espace normé (muni de la norme ou du produit scalaire).

Général Quelques résultats généraux et en vrac à propos d'isométries.

23. Définition 4.111.
24. Définition 8.172.

(1) Définition d'une isométrie pour une forme bilinéaire, 8.155. Pour une forme quadratique : définition 8.154.

(2) Définition du groupe orthogonal 8.44, et le spécial orthogonal $\mathrm{SO}(n)$ en la définition 8.47. Le groupe $\mathrm{SO}(2)$ est le groupe des rotations, par corolaire 18.138.

(3) La rotation $R_A(\theta)$ d'un angle θ autour du point $A \in \mathbb{R}^2$ est donnée par la définition 18.127.

(4) La proposition 18.139 donne à toute rotation $R_0(\theta)$ une matrice de la forme connue. C'est autour de cela que nous définissons les angles, définition 18.157.

(5) Le groupe orthogonal est le groupe des isométries de $\mathbb{R}^n$, proposition 8.46.

(6) Les isométries de l'espace euclidien sont affines, 8.157.

(7) Les isométries de l'espace euclidien comme produit semi-direct : $\mathrm{Isom}(\mathbb{R}^n) \simeq T(n) \times_\rho \mathrm{O}(n)$, théorème 18.79.

(8) Isométries du cube, section 5.7.

(9) Nous parlons des isométries affines du tétraèdre régulier dans la proposition 18.195.

Groupe diédral Le groupe diédral est un peu central dans la théorie des isométries de $(\mathbb{R}^2, d)$ parce que beaucoup de sous-groupes finis des isométries de $(\mathbb{R}^2, d)$ sont en fait isomorphes au groupe diédral.

(1) Générateurs du groupe diédral, proposition 18.175.

(2) Un sous-groupe fini des isométries de $(\mathbb{R}^2, d)$ contenant au moins une réflexion est isomorphe au groupe diédral par le théorème 18.200.

(3) Le théorème 18.202 dit que le groupe des isométries propres d'une partie quelconque de $(\mathbb{R}^2, d)$ est soit cyclique soit isomorphe au groupe diédral.

Isométries et réflexions Dans un espace euclidien, toute isométrie peut être décomposée en réflexions autour d'hyperplans. Voici quelques énoncés à ce propos.

(1) Définition d'une réflexion dans $\mathbb{R}^2$ 18.121.

(2) La caractérisation en termes de projection orthogonale est le lemme 18.106; en termes de médiatrice c'est le lemme 18.112.

(3) Définition d'un hyperplan 8.307.

(4) En dimension 2, une rotation est définie comme composée de deux réflexions en la définition 18.118.

(5) En dimension 2, les réflexions ont un déterminant -1 par le lemme 18.131.

(6) Les isométries du plan $(\mathbb{R}^2, d)$ sont données dans le théorème 18.191, et sont au plus 3 réflexions par le théorème 18.189.

(7) Décomposition des isométries de $\mathbb{R}^n$ en réflexions par le lemme 18.87.

(8) En particulier, les éléments de $\mathrm{SO}(3)$ sont des compositions de deux réflexions par le corolaire 18.89.

(9) Une isométrie de $\mathbb{R}^n$ préserve l'orientation si et seulement si elle est la composition d'un nombre pair de réflexions. C'est le théorème 18.95.

Sous-groupe fini (1) Les sous-groupes finis des isométries de $(\mathbb{R}^2, d)$ sont cycliques, théorème 18.200.

(2) Les sous-groupes finis de $\mathrm{SO}(3)$ sont listés dans 18.224.

(3) Les sous-groupes finis de $\mathrm{SO}(2)$ sont cycliques, lemme 18.142.

THTOC79

Thème 79 : enveloppes

(1) L'ellipse de John-Loewner donne un ellipsoïde de volume minimum autour d'un compact dans $\mathbb{R}^n$, théorème 17.122.

(2) Le cercle circonscrit à une courbe donne un cercle de rayon minimal contenant une courbe fermée simple, proposition 21.89.

(3) Enveloppe convexe du groupe orthogonal 13.38.

(4) Enveloppe convexe d'une courbe fermée plane comme intersection des demi-plans tangents, proposition 21.95.

THTOC80

Thème 80 : intégration sur des variétés Définition d'une variété : 20.1.

orientation La notion d'orientation commence avec l'orientation des bases d'un espace vectoriel et continue jusqu'à orienter des variétés à partir de ses cartes.

(1) Classe d'orientation sur les bases d'un espace vectoriel, définition 8.29.

(2) Orientation sur une surface, définition 20.60.

(3) Variété orientable, définition 20.17.

Intégrale Nous définissons d'abord l'intégrale sur une carte, et ensuite sur une variété. En ce qui concerne les formes différentielles, nous n'intégrons que des n-formes sur des variétés de dimension n.

(1) Intégrale d'une fonction sur une carte, définition 20.9.

théorème de Stokes, théorème de Green et compagnie Tous ces théorèmes sont des conséquences plus ou moins directes de celui de Stokes, et des généralisations du théorème fondamental de l'analyse.

(1) Forme générale, théorème 20.71.

(2) Rotationnel et circulation, théorème 24.8.

Le théorème de Stokes peut être utilisé pour montrer le théorème de Brouwer, proposition 20.30.

THTOC81

Thème 81 : dénombrements

— Coloriage de roulette (18.10.15.1) et composition de colliers (18.10.15.2).

— Nombres de Bell, théorème 15.170.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples [25], théorème 26.99.

THTOC82

Thème 82 : caractérisation de distributions en probabilités

(1) La probabilité conjointe est la définition 36.24.

(2) La fonction de répartition est la définition 36.87.

(3) La fonction caractéristique est la définition 36.90.

THTOC83

Thème 83 : théorème central limite

(1) Pour les processus de Poisson [26], théorème 40.22.

25. Éléments simples, lemme 19.13.
26. Définition 40.5.

THTOC84

Thème 84 : indépendance d'événements et de variables aléatoires

(1) Événements indépendants, définition 36.9.
(2) Variables aléatoires indépendantes, définition 36.13.
(3) Espérance de variables aléatoires indépendantes : $E(X_1 \cdots X_n) = E(X_1) \dots E(X_n)$, proposition 36.37.
(4) Densité de variables aléatoires indépendantes : $f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$, proposition 36.27.

THTOC85

Thème 85 : probabilités et espérances conditionnelles Les deux définitions de base, sur lesquelles se basent toutes les choses conditionnelles sont :

— L'espérance conditionnelle d'une variable aléatoire sachant une tribu : $E(X|\mathcal{F})$ de la définition 36.63.

Les autres sont listées ci-dessous.

Probabilité conditionnelle .

Plusieurs probabilités conditionnelles.

— D'un événement en sachant un autre : la définition 36.55 donne

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Cela est la définition de base. L'autre est une définition dérivée.

— D'un événement vis-à-vis d'une variable aléatoire discrète. C'est par la définition 36.81 qui définit la variable aléatoire

$$P(A|X)(\omega) = P(A|X = X(\omega)).$$

Dans le cas continu, c'est la définition 36.82 :

$$P(A|X) = P(A|\sigma(X)) = E(\mathbb{1}_A|\sigma(X)).$$

— D'un événement par rapport à une tribu. C'est la variable aléatoire

$$P(A|\mathcal{F}) = E(\mathbb{1}_A|\mathcal{F}).$$

Espérances conditionnelles Plusieurs espérances conditionnelles.

— D'une variable aléatoire par rapport à un événement, définition 36.75 :

$$E(X|A) = \frac{E(X\mathbb{1}_A)}{P(A)}. \tag{-1.18}$$

— d'une variable aléatoire par rapport à une tribu. La variable aléatoire $E(X|\mathcal{F})$ est la variable aléatoire $\mathcal{F}$-mesurable telle que

$$\int_B E(X|\mathcal{F}) = \int_B X$$

pour tout $X \in \mathcal{F}$. Si $X \in L^2(\Omega, \mathcal{A}, P)$ alors $E(X|\mathcal{F}) = \texttt{proj}_K(X)$ où K est le sous-ensemble de $L^2(\Omega, \mathcal{A}, P)$ des fonctions $\mathcal{F}$-mesurables (théorème 36.63). Cela au sens des projections orthogonales.

— d'une variable aléatoire par rapport à une autre. La définition 36.65 est une variation sur le même thème :

$$E(X|Y) = E(X|\sigma(Y)),$$

Notons que partout, si X est une variable aléatoire, la notation «sachant X» est un raccourci pour dire «sachant la tribu engendrée par X».

Quelque formules.

(1) Pour l'espérance conditionnelle d'une variable aléatoire prenant seulement une quantité dénombrable de valeurs : $E(X|A) = \sum_{k=0}^{\infty} y_k P(X = y_k|A)$ par le lemme 36.76.
(2) La probabilité conditionnelle se factorise par rapport à l'union disjointe par le lemme 36.59 : $P\big(\bigcup_{i=0}^{\infty} A_i|B\big) = \sum_{i=0}^{\infty} P(A_i|B)$.

THTOC86

Thème 86 : chaine de Markov

(1) Construction d'une chaine de Markov dont la matrice de transition est donnée, lemme 38.42.

(2) Principales propriétés d'une chaine de Markov homogène, proposition 38.39.

(3) Si la chaine de Markov est irréductible, alors il y a unicité de l'état stationnaire, proposition 38.16. Mais attention : cela ne veut pas encore dire que la chaine converge effectivement vers cet état.

(4) Si la chaine est irréductible et apériodique, alors il y a convergence en loi vers l'unique loi invariante, théorème 38.37.

(5) État absorbant, définition 38.44.

THTOC87

Thème 87 : manifolds

(1) Inverse function theorem: if the differential is bijective, the map is a local diffeomorphism, proposition 49.142.
(2) There exist a unique v such that $X = \frac{d}{dt}\left[\varphi_\alpha(s_0 + tv)\right]_{t=0}$, proposition 49.64 and lemma 49.65.
(3) By lemma 49.63, a vector $v \in T_aM$ can be written as

$$v(f) = \sum_{k=1}^{n} v_k \partial_k (f \circ \varphi)(s). \tag{-1.19}$$

(4) We can also write

$$X_x = \sum_{i=1}^{n} X_i(x)\partial_i \tag{-1.20}$$

for $X_i \in C^k(M, \mathbb{R})$ by proposition 49.85.
(5) The tangent space T_pM is a vector space of the same dimension as M, proposition 49.53.
(6) The vectors ∂_i are a basis of T_pM, proposition 49.56.
(7) If X is a vector field, it can be written as

$$X_x(f) = \sum_{k=1}^{n} v_k(x)\partial_k(f \circ \varphi)\big(\varphi^{-1}(x)\big), \tag{-1.21}$$

with $v_k \in C^k(M, \mathbb{R})$, lemma 49.86.

THTOC88

Thème 88 : Immersions and Embeddings En français, «embedding» est «plongement».

(1) Immersion, definition 49.110.
(2) Embedding, definition 49.126.
(3) Immersions are locally embeddings, theorem 49.127.

THTOC89

Thème 89 : Lie groups

(1) Topological group, definition 48.49.
(2) Lie group, definition 52.1.
(3) Lie subgroup, definition 52.24.
(4) Lie group structure on a closed subgroup of a Lie group, Cartan theomrem 53.64.

THTOC90

Thème 90 : $\mathbf{GL}(n, \mathbb{C})$

(1) $\mathrm{GL}(n, \mathbb{C})$ is a Lie group, proposition 55.17

THTOC91

Thème 91 : $\mathrm{SU}(n)$

(1) The part $\mathrm{SU}(n)$ is closed in $\mathrm{GL}(n, \mathbb{C})$, lemma 55.20.
(2) The analytic Lie group structure on $\mathrm{SU}(n)$ is given by proposition 55.27.
(3) The group $\mathrm{SU}(2)$ is compact, proposition 55.34.

THTOC92

Thème 92 : flow of a vector field

(1) Definition, existence, unicity, proposition 49.116.
(2) Complete vector field, definition 49.118.
(3) Specification to invariant vector fields, equations (53.64).
(4) On a Lie group, $\alpha_X(s + t) = \alpha_X(s)\alpha_X(t)$, proposition 53.26.
(5) $\Phi(t, g) = g\Phi(t, e)$, proposition 53.28.

THTOC93

Thème 93 : exponential

(1) Definition 53.30.

(2) exp is continuous, proposition 53.32.

(3) exp is smooth, proposition 53.35.

(4) $d\exp = \mathrm{Id}$ (sort of), lemma 53.37.

(5) $\phi(\exp(X)) = \exp\big(d\phi_e(X)\big)$, lemma 53.38.

(6) The exponential is a smooth local diffeomorphism, proposition 53.47.

(7) The exponential is an analytic local diffeomorhism, theorem 53.48.

(8) The map $\exp\colon U \to V$ is an analytic diffeomorphism, theorem 53.48.

(9) $\exp\big(t(X+Y)\big) = \exp(tX)\exp(tY)\exp(t\alpha(t))$ and other like that, lemma 53.55.

(10) $[X,Y] = \frac{d}{dt}\Big[\exp\big(\sqrt{t}X\big)\exp\big(\sqrt{t}Y\big)\exp\big(-\sqrt{t}X\big)\exp\big(-\sqrt{t}Y\big)\Big]_{t=0}$, lemma 53.54.

-1.1 Conventions sur les matrices et changement de bases

SECooBTTTooZZABWA

-1.1.1 Matrices et applications linéaires

SUBSECooAFPDooOzXdGz

Le lien entre matrice et application linéaire est donné par la définition 4.68. L'application d'une matrice à un vecteur est (4.84). Le lien le plus simple entre l'application linéaire et les éléments de matrice est donné par la proposition 4.71. Voici les relations :

$$T_{\alpha i} = T(e_i)_\alpha \tag{-1.22a}$$

$$T(e_i) = \sum_\alpha T_{\alpha i} f_\alpha \tag{-1.22b}$$

$$T(x) = \sum_{i\alpha} T_{\alpha i} x_i f_\alpha \tag{-1.22c}$$

$$T(x)_\alpha = \sum_i T_{\alpha i} x_i. \tag{-1.22d}$$

De la même manière nous utiliserons (rarement) la notation suivante (définition 9.109) si $x \in \mathbb{R}^n$ et $T \in \mathbb{M}(n \times n)$:

$$xT = \sum_{ij} x_i T_{ij} e_j. \tag{-1.23}$$

À partir de là, il est possible de parler de vecteur propre à gauche lorsque $xT = \lambda x$.

Cela définit une application $\psi\colon \mathbb{M}(n \times m, \mathbb{K}) \to \mathcal{L}(E,F)$ qui a plein de propriétés.

(1) C'est une bijection, proposition 4.71(4).

(2) C'est un isomorphisme d'algèbre, proposition 4.74.

(3) C'est un isomorphisme d'espaces vectoriels, proposition 4.72.

(4) Isomorphisme d'algèbres et d'anneaux, proposition 4.74.

(5) Isomorphisme d'espaces topologiques, proposition 4.74.

Lorsque nous avons une base orthonormée[27] nous avons aussi les propositions 8.184 et 8.184 qui donnent des formules avec produit scalaire :

(1) $T_{\alpha i} = e_\alpha \cdot T(e_i)$

(2) $x \cdot Ay = \sum_{kl} A_{kl} x_k y_l$.

où le point est le produit scalaire usuel de $\mathbb{R}^n$.

27. Définition 8.170.

-1.1.2 Le changement de base

Soit un espace vectoriel V muni de deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$. Le lemme 4.115 donne le lien entre les vecteurs de base :

(1) $f_\alpha = \sum_i Q_{i\alpha} e_i$
(2) $e_i = \sum_\alpha Q^{-1}_{\alpha i} f_\alpha$

La proposition 4.116 donne un certain nombre de formules pour les coordonnées des vecteurs :

(1) $y_\alpha = \sum_i Q^{-1}_{\alpha i} x_i$
(2) $x_i = \sum_\alpha Q_{i\alpha} y_\alpha$.
(3) $x_i = (Qy)_i$
(4) $x = Qy$

La transformation de la matrice d'une application linéaire lors d'un changement de base est la proposition 4.119. Soit une application linéaire $T : V \to V$ de matrices A et B dans les bases $\{e_i\}$ et $\{f_\alpha\}$. Si les bases sont liées par $f_\alpha = \sum_i Q_{i\alpha} e_i$, alors les matrices A et B sont liées par

$$B = Q^{-1} A Q. \tag{-1.24}$$

En ce qui concerne la base duale, si $e'_j = \sum_i A_{ij} e_i$, alors

$$\alpha'_j = \sum_j A^{-1}_{jk} \alpha_k, \tag{-1.25}$$

proposition 4.132.

-1.1.3 Changement de base : matrice d'une forme bilinéaire

La proposition 8.151 fait le changement de matrice d'une forme bilinéaire lors d'un changement de base. Si la matrice de q dans la base $\{e_i\}$ est A et celle dans la base $\{f_\alpha\}$ est B, alors

$$B = Q^t A Q. \tag{-1.26}$$

Pour comparaison avec la loi de transformation des matrices des applications linéaires, voir la remarque 8.152.

Plus généralement, si ϕ est une application linéaire, la matrice de $q \circ \phi$ est $\phi^t q \phi$, proposition 8.150.

-1.2 Multiindice et liste d'indices

NORMooRRZCooMOKAzY

-1.1.

Je crois qu'il y a quelques incohérences de notations/dénominations dans le texte. En principe quand on parle de $\mathbb{R}^n$, un **multiindice** [4] est une vecteur d'entiers positifs à n composantes. Si $\alpha = (2, 1)$ alors nous avons la notation

$$\partial^\alpha f = \frac{\partial^2}{\partial x_1} \frac{\partial}{\partial x_2} f \tag{-1.27}$$

Cette notation pose problème lorsque, par exemple, $\partial_1^2 \partial_2 f \neq \partial_1 \partial_2 \partial_1 f$.

Elle pose également problème lorsque l'on veut faire une récurrence sur l'ordre de dérivation en ajoutant une seule dérivation à la fois.

C'est pourquoi nous introduisons le concept de **liste d'indices**. En parlant de $\mathbb{R}^n$, une liste d'indices est un vecteur arbitrairement long (mais fini) d'entiers dans $\{1, \dots, n\}$. Si, dans $\mathbb{R}^7$, $\alpha = (1, 3, 1, 5)$, alors

$$\partial^\alpha f = \partial_1 \partial_3 \partial_1 \partial_5 f. \tag{-1.28}$$

Si α est une liste d'indices de longueur p, une **queue de** α est une liste d'indices de longueur $0 < k \leqslant p$ de la forme $(\alpha_{p-k+1}, \alpha_{p-k+2}, \dots, \alpha_p)$.

Voir aussi la définition 3.47 pour $\alpha!$ et <++>

-1.3 Anglicismes

SECooPBZVooCVInFT

Voici quelques anglicismes dont je ne me souviens jamais.

(1) Une σ-algèbre est une tribu, définition 14.1.

(2) Le «uniform boundedness principle» est le théorème de Banach-Steinhaus 10.151.

(3) Un anneau est ce qu'on appelle «*ring*» en anglais. Un corps est en anglais «*field*». De plus le mot «*field*» comprend la commutativité. Donc certains utilisent le mot «corps» pour dire «corps commutatif» et parlent alors d'anneau *à division* pour parler de corps non commutatifs.

Contents

Thématique **3**
-1.1 Conventions sur les matrices et changement de bases . . . 44
-1.1.1 Matrices et applications linéaires . . . 44
-1.1.2 Le changement de base . . . 45
-1.1.3 Changement de base : matrice d'une forme bilinéaire . . . 45
-1.2 Multiindice et liste d'indices . . . 45
-1.3 Anglicismes . . . 46

Table des matières **47**

Index **105**

Liste des notations **145**

I Le Frido **157**

0 Introduction **159**
0.1 Auteurs, contributeurs, sources et remerciements . . . 159
0.1.1 Ceux qui ont travaillé sur le Frido . . . 159
0.1.2 Aide directe, mais pas volontairement sur le Frido . . . 160
0.1.3 Des gens qui ont fait un travail qui m'a bien servi . . . 161
0.2 Originalité . . . 161
0.3 Les choses qui doivent vous faire tiquer . . . 162
0.4 Quelques choix qui peuvent provoquer des quiproquos . . . 162
0.5 Autres choix pas spécialement standards . . . 163
0.5.1 Mathématique intéressante . . . 163
0.6 Sage est là pour vous aider . . . 163
0.6.1 Lancez-vous dans Sage . . . 164
0.6.2 Exemples de ce que Sage peut faire pour vous . . . 164
0.7 Comment contribuer et aider ? . . . 165
0.7.1 Des preuves qui manquent . . . 165
0.7.2 Du texte qui manque . . . 165
0.7.3 Des exemples qui manquent . . . 165
0.7.4 Trucs de programmation et de LaTeX . . . 165
0.8 Les politiques éditoriales . . . 166
0.8.1 Licence libre . . . 166
0.8.2 pdflatex . . . 166
0.8.3 utf8 . . . 166
0.8.4 Notations . . . 166
0.8.5 De la bibliographie . . . 166
0.8.6 Faire des références à tout . . . 166
0.8.7 Des listes de liens internes . . . 166
0.8.8 Pas de références vers le futur . . . 166

1 Construction des ensembles de nombres **167**
1.1 Quelques éléments sur les ensembles . . . 167
1.1.1 Petit mot d'introduction . . . 167
1.1.2 Injection, surjection, bijection . . . 168
1.1.3 Ensemble ordonné . . . 168
1.1.4 Lemme de Zorn . . . 170
1.1.5 Complémentaire . . . 171
1.1.6 Quelques relations ensemblistes . . . 171
1.1.7 Relations d'équivalence . . . 174
1.2 Quelques structures algébriques . . . 175
1.3 Les naturels . . . 176
1.3.1 Applications définies par récurrence . . . 177
1.3.2 Addition sur les naturels . . . 180
1.3.3 Ordre sur les naturels . . . 183
1.3.4 Multiplication dans les naturels . . . 188
1.3.5 Presque unicité des triplets naturels . . . 191
1.3.6 Écriture d'un naturel dans une base . . . 195
1.4 Les entiers . . . 200
1.4.1 Opposé . . . 201
1.4.2 Ordre sur $\mathbb{Z}$ . . . 202
1.5 Quelques résultats de cardinalité . . . 203
1.5.1 Équipotence, surpotence, subpotence . . . 203
1.5.2 Un peu d'infinité . . . 204
1.5.3 Dénombrabilité et ensemble des naturels . . . 210
1.5.4 Théorème de Cantor-Schröder-Bernstein . . . 215
1.5.5 Comparabilité cardinale . . . 218
1.5.6 Théorème de Cantor, ensemble des ensembles . . . 219
1.5.7 Ajouter ou soustraire des cardinalités . . . 220
1.6 Groupes . . . 225
1.6.1 Définition, unicité du neutre . . . 225
1.6.2 Groupe ordonné . . . 225
1.6.3 Classes de conjugaison . . . 226
1.7 Anneaux . . . 227
1.7.1 Diviseurs . . . 229
1.7.2 Élément irréductible et premier . . . 230
1.7.3 Anneau intègre . . . 231
1.7.4 Positivité . . . 235
1.7.5 Fonction puissance . . . 235
1.8 Idéal dans un anneau . . . 236
1.8.1 Division euclidienne . . . 238
1.9 Anneau principal et idéal premier . . . 240
1.10 Anneau intègre . . . 242
1.10.1 Sous-groupes de $(\mathbb{Z}, +)$ . . . 242
1.10.2 Théorème de Bézout . . . 242
1.11 Anneau euclidien . . . 246
1.12 Le groupe et anneau des entiers . . . 248
1.12.1 PGCD, PPCM et Bézout . . . 248
1.13 Sous-groupe normal . . . 250
1.13.1 Permutations, groupe symétrique . . . 254
1.13.2 Décomposition en cycles . . . 255
1.13.3 Permutation un peu ordonnées . . . 261
1.14 Corps . . . 263
1.14.1 Définitions, morphismes . . . 263

1.15 Symbole de sommation . . . 264
1.15.1 Somme à valeurs dans un groupe commutatif . . . 264
1.16 Symbole de produit . . . 267
1.16.1 Sous-groupe engendré . . . 268
1.17 Module sur un anneau . . . 270
1.17.1 Module produit . . . 271
1.17.2 Sous-module . . . 273
1.18 Caractéristique d'un anneau . . . 275
1.18.1 Caractéristique deux . . . 276
1.19 Polynômes . . . 277
1.19.1 Polynômes d'une variable . . . 277
1.19.2 La notation $A[X]$ . . . 279
1.19.3 Action du groupe symétrique . . . 280
1.19.4 Corps des fractions . . . 281
1.19.5 Corps totalement ordonné . . . 285
1.20 Les rationnels . . . 287
1.20.1 Relation d'ordre . . . 288
1.20.2 Caractéristique . . . 290
1.21 Suite de Cauchy dans un corps totalement ordonné . . . 290
1.21.1 Suites de Cauchy dans les rationnels . . . 292
1.22 Insuffisance des rationnels . . . 294
1.23 Les réels . . . 298
1.23.1 L'ensemble . . . 298
1.23.2 Relation d'ordre . . . 300
1.23.3 Complétude . . . 307
1.23.4 Intervalles . . . 310
1.23.5 Maximum, supremum et compagnie . . . 311
1.23.6 Racines . . . 317
1.23.7 Corps valué . . . 318
1.23.8 Partie entière, partie fractionnaire . . . 318
1.24 Les complexes . . . 318

2 Théorie des groupes **321**
2.1 Groupe dérivé . . . 321
2.2 Théorèmes d'isomorphismes . . . 322
2.3 Indice d'un sous-groupe et ordre des éléments . . . 324
2.4 Suite de composition . . . 325
2.5 Groupes résolubles . . . 329
2.6 Action de groupes . . . 330
2.7 Produit semi-direct de groupes . . . 336
2.8 Groupe de torsion . . . 338
2.9 Famille presque nulle . . . 338

3 Anneaux **339**
3.1 Inversibles et nilpotents . . . 339
3.2 PGCD, PPCM et éléments inversibles . . . 340
3.2.1 Calcul effectif du PGCD et théorème de Bézout . . . 340
3.2.2 Générateurs . . . 342
3.2.3 Décomposition en facteurs premiers . . . 344
3.2.4 Ordre d'un élément dans un groupe fini . . . 348
3.2.5 Écriture des fractions . . . 350
3.2.6 Équation diophantienne linéaire à deux inconnues . . . 351
3.2.7 Quotients . . . 353
3.3 Binôme de Newton et morphisme de Frobenius . . . 353

3.3.1 Théorème multinomial . . . 356
3.4 Polynômes de plusieurs variables . . . 358
3.4.1 Divisibilité et classes d'association . . . 358
3.4.2 PGCD et PPCM . . . 359
3.4.3 Anneaux intègres et corps . . . 360
3.5 Anneau factoriel . . . 360
3.5.1 Autour du théorème de Bézout . . . 362
3.5.2 Idéal premier . . . 364
3.5.3 Anneau noethérien . . . 364
3.6 Anneau $\mathbb{Z}/6\mathbb{Z}$ . . . 367
3.6.1 Équations diophantiennes . . . 370
3.6.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$ . . . 371
3.7 Polynômes à coefficients dans un anneau commutatif . . . 373
3.7.1 Monômes . . . 374
3.7.2 Évaluation . . . 374
3.7.3 Polynômes sur un anneau intègre . . . 375
3.7.4 Division euclidienne . . . 375
3.7.5 Polynôme primitif . . . 376
3.7.6 Racines des polynômes . . . 379
3.7.7 Quelques identités . . . 382
3.7.8 Générateurs pour le groupe multiplicatif . . . 382

4 Espaces vectoriels (début) 385
4.1 Parties libres, génératrices, bases et dimension . . . 385
4.1.1 Et en dimension infinie . . . 391
4.1.2 Espace librement engendré . . . 393
4.2 Applications linéaires . . . 394
4.2.1 Définition . . . 394
4.2.2 Linéarité et bases . . . 397
4.2.3 Rang . . . 398
4.3 Matrices . . . 402
4.3.1 Définitions . . . 403
4.3.2 Identifier matrices et applications linéaires . . . 405
4.3.3 Ensemble des applications linéaires . . . 408
4.3.4 Déterminant . . . 409
4.3.5 Déterminant en petite dimension . . . 409
4.3.6 Manipulations de lignes et de colonnes . . . 410
4.3.7 Réduction de Gauss . . . 415
4.3.8 Matrices inversibles . . . 417
4.3.9 Inversibilité et déterminant . . . 419
4.3.10 Quelques ensembles de matrices particuliers . . . 419
4.3.11 Déterminant et combinaisons de lignes et colonnes . . . 420
4.3.12 Transvections . . . 420
4.3.13 Mineur, rang . . . 421
4.3.14 Matrices équivalentes et semblables . . . 423
4.3.15 Algorithme des facteurs invariants . . . 424
4.4 Changement de base . . . 426
4.4.1 Changement de base : vecteurs de base . . . 426
4.4.2 Changement de base : coordonnées . . . 426
4.4.3 Changement de base : matrice d'une application linéaire . . . 427
4.5 Espaces de polynômes . . . 428
4.6 Projection et orthogonalité . . . 429
4.7 Dualité . . . 430
4.7.1 Changement de base duale . . . 432

4.7.2 Orthogonal . . . 432
4.8 Représentation de groupe . . . 433
4.9 Somme directe d'espaces vectoriels . . . 434
4.9.1 Structure réelle . . . 435

5 Classification de certains groupes 439
5.1 Théorèmes de Sylow . . . 439
5.2 Groupe monogène . . . 443
5.3 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$ . . . 444
5.4 Groupes abéliens finis . . . 446
5.5 Groupes d'ordre pq . . . 448
5.5.1 Fonction indicatrice d'Euler . . . 451
5.5.2 Générateurs . . . 454
5.5.3 Fonction indicatrice d'Euler (propriétés) . . . 455
5.6 Groupe symétrique, groupe alterné . . . 455
5.6.1 Le groupe alterné . . . 455
5.6.2 Sous-groupes normaux . . . 461
5.6.3 Indice . . . 462
5.7 Isométries du cube . . . 464

6 Corps 467
6.1 Généralités . . . 467
6.1.1 Corps ordonnés . . . 467
6.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$ . . . 467
6.1.3 Corps premier . . . 469
6.1.4 Petit théorème de Fermat . . . 471
6.1.5 Nombres de Sophie Germain . . . 471
6.2 Théorème des deux carrés . . . 474
6.2.1 Un peu de structure dans $\mathbb{Z}[i]$ . . . 474
6.2.2 Résultats chinois . . . 478
6.3 Polynômes à coefficients dans un corps . . . 480
6.3.1 Irréductibilité . . . 485
6.3.2 Idéaux . . . 486
6.3.3 Identité de Bézout . . . 488
6.3.4 Lemme et théorème de Gauss . . . 489
6.3.5 Polynômes sur un corps et pgcd . . . 489
6.4 Extension de corps . . . 492
6.4.1 Extension et polynôme minimal . . . 493
6.4.2 Extensions algébriques et éléments transcendants . . . 495
6.4.3 Extension algébrique et polynôme minimal . . . 498
6.4.4 Extensions et polynômes . . . 499
6.4.5 Racines de polynômes . . . 509
6.4.6 Corps de rupture . . . 510
6.4.7 Pile d'extensions . . . 512
6.4.8 Clôture algébrique . . . 513
6.4.9 Polynômes à plusieurs variables . . . 519
6.4.10 Corps de décomposition . . . 520
6.4.11 Non irréductible ou pas corps ? . . . 524
6.4.12 Clôture algébrique . . . 524
6.4.13 Dérivée de polynômes . . . 525
6.4.14 Extensions séparables . . . 526
6.5 Idéal maximum . . . 533
6.5.1 Idéal maximum . . . 533
6.6 Polynômes symétriques et alternés . . . 534

6.6.1 Polynômes symétriques, alternés ou semi-symétriques . . . 534
6.6.2 Polynôme symétrique élémentaire . . . 534
6.6.3 Relations coefficients racines . . . 537
6.7 Minuscule morceau sur la théorie de Galois . . . 538

7 Topologie générale **539**
7.1 Éléments généraux de topologie . . . 539
7.1.1 Définitions et propriétés de base . . . 539
7.1.2 Base de topologie . . . 539
7.1.3 Fermés . . . 540
7.1.4 Quelques exemples . . . 541
7.1.5 Topologie produit . . . 543
7.1.6 Adhérence, fermeture, intérieur, point d'accumulation et point isolé . . . 543
7.1.7 Topologie sur $\mathbb{Q}$ . . . 547
7.1.8 Applications continues . . . 547
7.2 Topologie rendant continue . . . 549
7.2.1 Topologie quotient . . . 551
7.3 Suites et convergence . . . 552
7.3.1 Convergence dans un fermé . . . 552
7.3.2 Pour des limites uniques : séparabilité . . . 553
7.3.3 Topologie induite . . . 554
7.3.4 Fonctions équivalentes . . . 555
7.4 Connexité . . . 556
7.5 Compacité . . . 559
7.5.1 Définition et notions connexes . . . 559
7.5.2 Espace localement compact . . . 560
7.5.3 Autres compacité . . . 561
7.5.4 Quelques propriétés . . . 562
7.5.5 Espace relativement compact . . . 564
7.5.6 Compactifié d'Alexandrov . . . 565
7.5.7 Propriété d'intersection finie . . . 567
7.6 Limite de fonction . . . 567
7.7 Topologie, distances et normes . . . 569
7.7.1 Distance et topologie métrique . . . 569
7.7.2 Topologie métrique et induite . . . 570
7.7.3 Intérieur, adhérence et frontière . . . 570
7.7.4 Boules ouvertes, fermées, sphères . . . 572
7.7.5 Continuité séquentielle . . . 572
7.7.6 Continuité et compacité . . . 573
7.7.7 Distance à un ensemble . . . 576
7.7.8 Convexité . . . 576
7.7.9 Norme . . . 577
7.8 Espaces vectoriels topologiques . . . 580
7.8.1 Base de topologie . . . 581
7.8.2 Corps topologique . . . 583
7.8.3 Voisinage symétrique et équilibré . . . 583
7.8.4 Limite de suites . . . 586
7.8.5 Espace topologique localement convexe . . . 587
7.8.6 Limite de suites . . . 588
7.9 Applications continues . . . 589
7.9.1 Continuité . . . 589
7.9.2 Continuité et topologie induite . . . 591
7.9.3 Continuité et connexité . . . 593
7.9.4 Continuité et compacité . . . 594

7.9.5 Continuité de la réciproque sur un compact . . . 595
7.9.6 Espace vectoriel topologique . . . 595
7.10 Produit fini d'espaces vectoriels normés . . . 595
7.10.1 Distance et norme produit . . . 595
7.11 Topologie réelle en dimension n . . . 598
7.11.1 Ouverts et fermés . . . 598
7.11.2 Point d'accumulation, point isolé . . . 599
7.11.3 Limite de suite . . . 599
7.12 Topologie et distance . . . 600
7.12.1 Suites et espaces métriques . . . 602
7.12.2 Espace métrisable . . . 603
7.13 Suites de Cauchy, métrique et espaces complets . . . 604
7.13.1 Généralités . . . 604
7.13.2 Espace de Banach . . . 605
7.13.3 Espace topologique métrique . . . 607
7.13.4 Compacts, fermés . . . 608
7.13.5 Équivalence entre Cauchy et $\tau-$Cauchy . . . 616
7.14 Norme, espace vectoriel normé . . . 617
7.15 Espaces métriques . . . 620
7.15.1 Espaces métrisables . . . 620
7.15.2 Fonctions continues . . . 621
7.15.3 Ensembles enchainés . . . 627
7.15.4 Produit fini d'espaces métriques . . . 628
7.15.5 Équicontinuité . . . 629
7.15.6 Continuité uniforme . . . 630
7.16 Ensembles nulle part denses . . . 630
7.17 Seminormes . . . 631
7.17.1 Seminorme . . . 631
7.17.2 Fonctionnelle de Minkowski . . . 632
7.18 Topologie des seminormes . . . 636
7.18.1 Topologie initiale des seminormes . . . 637
7.18.2 Topologie avec des boules . . . 638
7.18.3 Seminorme et métrisable . . . 644
7.18.4 Boules et topologie . . . 645
7.18.5 Espace dual . . . 648
7.18.6 Espace $C^k(\mathbb{R}, E')$ . . . 649
7.19 Espaces de Baire . . . 650

8 Espaces vectoriels (encore) 651
8.1 Déterminants . . . 651
8.2 Applications multilinéaires . . . 651
8.2.1 Continuité . . . 652
8.2.2 Formes multilinéaires alternées . . . 653
8.2.3 Déterminant d'une famille de vecteurs . . . 655
8.2.4 Déterminant d'un endomorphisme . . . 657
8.2.5 Déterminant de Vandermonde . . . 660
8.2.6 Déterminant de Gram . . . 662
8.2.7 Déterminant de Cauchy . . . 663
8.2.8 Matrice de Sylvester . . . 663
8.2.9 Théorème de Kronecker . . . 666
8.3 Orientation . . . 667
8.3.1 Cas vectoriel . . . 667
8.3.2 Cas affine . . . 669
8.4 Hermitien, orthogonal, adjoint . . . 669

8.4.1 Opérateur orthogonal, matrice orthogonale . . . 671
8.5 Topologie . . . 672
8.5.1 Boules et sphères . . . 672
8.5.2 Ouverts, fermés, intérieur et adhérence . . . 674
8.5.3 Point isolé, point d'accumulation . . . 680
8.5.4 Des exemples . . . 681
8.6 Valeur propre et vecteur propre . . . 684
8.6.1 Généralités . . . 684
8.6.2 Dans le vif du sujet . . . 685
8.7 Polynômes d'endomorphismes . . . 686
8.7.1 Polynômes d'endomorphismes . . . 686
8.7.2 Polynôme minimal et minimal ponctuel . . . 689
8.7.3 Polynôme caractéristique . . . 696
8.8 Formes bilinéaires et quadratiques . . . 699
8.8.1 Dégénérescence d'une forme bilinéaire . . . 700
8.8.2 Orthogonal pour une forme bilinéaire . . . 700
8.8.3 Formes quadratiques . . . 703
8.8.4 Isotropie . . . 705
8.9 Formes bilinéaires et quadratiques . . . 706
8.9.1 Matrice associée à une forme bilinéaire . . . 706
8.9.2 Changement de base : matrice d'une forme bilinéaire . . . 707
8.9.3 Isométrie, forme quadratique et bilinéaire . . . 708
8.9.4 Isométries . . . 709
8.10 Signature, théorème de Sylvester . . . 711
8.11 Produit scalaire, produit hermitien . . . 712
8.11.1 Hermitien, unitaire, etc. . . . 714
8.11.2 Éléments de matrice . . . 715
8.11.3 Transposée : pas d'approche naïve . . . 716
8.11.4 Transposée : la bonne approche . . . 718
8.11.5 Polynômes de Lagrange . . . 721
8.11.6 Dual de $\mathbb{M}(n, \mathbb{K})$ . . . 721
8.12 Diagonalisation et trigonalisation . . . 722
8.12.1 Matrices semblables . . . 723
8.12.2 Trace de matrices semblables . . . 723
8.12.3 Endomorphismes nilpotents . . . 724
8.12.4 Endomorphismes diagonalisables . . . 726
8.12.5 Diagonalisation : cas complexe, pas toujours . . . 730
8.12.6 Diagonalisation : cas réel . . . 730
8.12.7 Matrice définie positive . . . 732
8.12.8 Réduction de Gauss . . . 736
8.12.9 Orthogonalité pour une forme bilinéaire . . . 739
8.12.10 Équivalence de formes quadratiques . . . 743
8.12.11 Diagonalisation . . . 743
8.13 Fonctions . . . 745
8.14 Sous espaces caractéristiques . . . 746
8.14.1 Théorèmes de décomposition . . . 748
8.14.2 Valeurs singulières . . . 750
8.15 Extension du corps de base . . . 751
8.15.1 Extension des applications linéaires . . . 751
8.15.2 Projections . . . 753
8.15.3 Rang, polynôme minimal, polynôme caractéristique . . . 756
8.16 Frobenius et Jordan . . . 758
8.16.1 Matrice compagnon . . . 758

8.16.2 Réduction de Frobenius . . . 759
8.16.3 Forme normale de Jordan . . . 761
8.17 Commutant et endomorphismes cycliques . . . 763
8.17.1 Endomorphisme cyclique . . . 763
8.17.2 Commutant : cas diagonalisable . . . 763
8.17.3 Commutant : cas général . . . 766
8.18 Hyperplans et formes linéaires . . . 768
8.18.1 Trouver la matrice d'une symétrie donnée . . . 770
8.19 Théorème de Burnside . . . 772
8.20 Ellipsoïde . . . 774
8.21 Système d'équations linéaires : méthode de Gauss . . . 775

9 Analyse réelle : topologie et continuité 777
9.1 Intervalles . . . 777
9.2 Application réciproque . . . 778
9.2.1 Définitions . . . 778
9.2.2 Graphe de la fonction réciproque . . . 779
9.3 Topologie sur l'ensemble des réels . . . 779
9.3.1 Compacité pour les réels . . . 780
9.3.2 Conséquence : les fermés bornés sont compacts . . . 783
9.3.3 Suites et limites dans les réels . . . 784
9.3.4 Opérations sur les limites . . . 784
9.3.5 Exemples . . . 785
9.3.6 Limites infinies . . . 785
9.3.7 Topologie et matrices . . . 786
9.3.8 Suites croissantes et bornées . . . 787
9.3.9 Suites adjacentes . . . 788
9.3.10 Limite supérieure et inférieure . . . 789
9.3.11 Ouverts, voisinage, topologie . . . 791
9.3.12 Intervalles et connexité . . . 792
9.3.13 Théorème de Bolzano-Weierstrass . . . 795
9.3.14 Recouvrement par des intervalles ouverts . . . 797
9.4 Connexité par arcs . . . 798
9.4.1 Des exemples . . . 800
9.4.2 Quelques mots à propos de la droite réelle achevée . . . 802
9.5 Continuité . . . 803
9.5.1 Opération sur la continuité . . . 804
9.5.2 La fonction la moins continue du monde . . . 805
9.5.3 Approche topologique . . . 806
9.5.4 Limite sur des chemins . . . 807
9.5.5 Racine carrée . . . 807
9.5.6 Module sur les nombres complexes . . . 808
9.5.7 Théorème de Perron-Frobenius . . . 811
9.5.8 Limite à gauche et à droite . . . 814
9.6 Norme à partir d'un produit scalaire . . . 815
9.6.1 Continuité de la racine carrée, invitation à la topologie induite . . . 815
9.6.2 Second degré . . . 817

10 Espaces vectoriels normés 819
10.0.1 Norme, produit scalaire et Cauchy-Schwarz (cas réel) . . . 819
10.0.2 Topologie . . . 822
10.1 Théorème spectral autoadjoint . . . 823
10.1.1 Inégalité de Minkowski . . . 825
10.1.2 Cauchy-Schwarz etc. cas complexe . . . 827

10.1.3 Diagonalisation : cas complexe, ce qu'on a . . . 828
10.1.4 Projection et orthogonalité . . . 828
10.1.5 Théorème de Pythagore . . . 830
10.1.6 Produit vectoriel . . . 830
10.1.7 Produit mixte . . . 834
10.1.8 Procédé de Gram-Schmidt . . . 835
10.1.9 Pseudo-réduction simultanée . . . 836
10.2 Approximations . . . 837
10.2.1 Quelques exemples de normes sur $\mathbb{R}^n$ . . . 838
10.3 Équivalence des normes . . . 839
10.3.1 En dimension finie . . . 840
10.3.2 Contre-exemple en dimension infinie . . . 843
10.4 Norme opérateur . . . 844
10.4.1 Norme d'algèbre . . . 846
10.5 Application linéaire continue et bornée . . . 847
10.5.1 Déterminant entre espaces différents . . . 849
10.5.2 Suites . . . 851
10.5.3 Continuité du produit de matrices . . . 854
10.5.4 Continuité de la composition . . . 855
10.6 Applications multilinéaires . . . 855
10.7 Séries . . . 856
10.7.1 Les trois types de convergence . . . 857
10.7.2 Somme téléscopique . . . 861
10.7.3 Distributivité . . . 861
10.7.4 Produit de Cauchy . . . 861
10.8 Sommes de familles infinies . . . 862
10.8.1 Convergence commutative . . . 862
10.8.2 Somme non dénombrables . . . 869
10.8.3 Sommes dénombrables . . . 870
10.9 Série réelle . . . 873
10.9.1 Critères de convergence absolue . . . 873
10.9.2 Critères de convergence simple . . . 875
10.9.3 Quelques séries usuelles . . . 875
10.9.4 Séries alternées . . . 878
10.9.5 Moyenne de Cesàro . . . 879
10.9.6 Écriture décimale d'un réel . . . 880
10.9.7 Théorème de Banach-Steinhaus . . . 883
10.9.8 Convergence forte . . . 886
10.10 Application ouverte . . . 887
10.11 Produit tensoriel d'espaces vectoriels . . . 890
10.12 Produit tensoriel . . . 891
10.12.1 Définition générale . . . 891
10.12.2 Constructions du produit tensoriel . . . 893
10.12.3 Construction abstraite . . . 893
10.12.4 Bases . . . 897
10.12.5 Construction par les formes multilinéaires . . . 899
10.12.6 Décomposabilité . . . 902
10.12.7 Produit tensoriel d'applications linéaires . . . 903
10.12.8 Contraction . . . 905
10.12.9 Tenseurs symétriques et alternés . . . 906
10.12.10 Décomposition . . . 911
10.12.11 Règle de Leibniz pour le produit extérieur . . . 913
10.12.12 Produit intérieur . . . 920

10.12.13 Pull back . . . 924
10.12.14 Norme . . . 924
10.12.15 Applications bilinéaires, matrices et produit tensoriel . . . 925
10.12.16 Application d'opérateurs . . . 926
10.12.17 Convergence en norme et par composante . . . 926
10.13 Calcul différentiel dans un espace vectoriel normé . . . 927
10.13.1 Limite de fonctions . . . 927
10.13.2 Définition de la différentielle . . . 927
10.13.3 Pas en dimension infinie . . . 928
10.13.4 Dérivée directionnelle . . . 928
10.13.5 Théorème des accroissements finis . . . 932
10.13.6 Différentielle d'ordre 2 . . . 934
10.13.7 Quelques mots à propos des différentielles d'ordre supérieur . . . 934
10.13.8 Différentielle d'applications linéaires . . . 936
10.13.9 Accroissements finis . . . 937
10.13.10 Notations pour les applications linéaires . . . 937
10.13.11 (non ?) Différentiabilité des applications linéaires . . . 937
10.13.12 Dérivation en chaine et formule de Leibniz . . . 938
10.13.13 Composition . . . 940
10.13.14 Règle de Leibniz . . . 944
10.13.15 Inverse . . . 944
10.13.16 Différentiation de produit . . . 944
10.13.17 Formule des accroissements finis . . . 948
10.13.18 Applications multilinéaires . . . 951
10.13.19 Différentielle partielle . . . 952
10.13.20 L'inverse, sa différentielle . . . 953
10.14 Exponentielle de matrice . . . 955
10.15 Espace dual . . . 956
10.15.1 Topologies . . . 956
10.15.2 Module de continuité . . . 958
10.16 Mini introduction aux nombres p-adiques . . . 959
10.16.1 La flèche d'Achille . . . 959
10.16.2 La tortue et Achille . . . 960
10.16.3 Dans les nombres p-adiques, c'est vrai . . . 961

11 Espaces affines **963**
11.1 Vecteurs agissant sur un espace . . . 963
11.2 Repères cartésiens affines . . . 964
11.3 Classification affine des coniques . . . 964
11.4 Applications affines . . . 967
11.4.1 Autres propriétés . . . 968
11.5 Isomorphismes . . . 969
11.6 Sous espaces affines . . . 970
11.7 Barycentre . . . 972
11.7.1 Sous-espaces affines . . . 972
11.7.2 Enveloppe convexe . . . 974
11.7.3 Applications affines et barycentre . . . 977
11.8 Repères, coordonnées cartésiennes et barycentriques . . . 978
11.8.1 Équation de droite . . . 981
11.8.2 Associativité, coordonnées barycentriques dans un triangle . . . 982
11.9 Applications affines sur $\mathbb{R}^n$ . . . 983
11.9.1 Structure de groupe pour les applications affines . . . 984
11.10 Isométries . . . 986

12 Analyse réelle : limites et dérivation **987**
12.1 Limite de fonctions . . . 987
12.1.1 Définition . . . 987
12.1.2 Quelques règles de calcul . . . 989
12.2 Limites pointées et épointées . . . 993
12.2.1 Théorèmes de composition de limites . . . 995
12.2.2 Discussion pointée Vs épointée . . . 998
12.3 Limites en l'infini . . . 1001
12.3.1 Limite en des nombres . . . 1003
12.3.2 Limites quand tout va bien . . . 1003
12.3.3 Limites de fonctions . . . 1005
12.3.4 Limite à gauche et à droite . . . 1005
12.4 Limite en compactifié d'Alexandrov . . . 1006
12.4.1 Prolongement par continuité . . . 1007
12.4.2 Prolongement par continuité . . . 1009
12.4.3 Théorème de la bijection . . . 1009
12.5 Limite et continuité . . . 1013
12.5.1 Prolongement des rationnels vers les réels . . . 1015
12.6 Espace des fonctions continues . . . 1018
12.7 Uniforme continuité . . . 1022
12.8 Fonctions sur un compact . . . 1025
12.9 Polynômes, théorème de d'Alembert . . . 1025
12.9.1 Polynômes sur les réels . . . 1025
12.9.2 Polynômes sur les complexes . . . 1026
12.10 Trigonalisation . . . 1032
12.10.1 Trigonalisation : généralités . . . 1032
12.10.2 Trigonalisation : cas complexe . . . 1033
12.11 Matrices, spectre et norme . . . 1040
12.11.1 Rayon spectral . . . 1040
12.11.2 Normes de matrices et d'applications linéaires . . . 1045
12.12 Géométrie dans l'espace . . . 1049
12.12.1 Droites dans l'espace . . . 1050
12.12.2 Projection orthogonale . . . 1052
12.12.3 Plan médiateur . . . 1054
12.12.4 Tétraèdre . . . 1056
12.13 Géométrie dans le plan . . . 1056
12.14 Dérivée : exemples introductifs . . . 1061
12.14.1 La vitesse . . . 1061
12.14.2 La tangente à une courbe . . . 1061
12.14.3 L'aire en dessous d'une courbe . . . 1063
12.15 Dérivation de fonctions réelles . . . 1063
12.15.1 Exemples . . . 1065
12.15.2 Interprétation géométrique : tangente . . . 1066
12.15.3 Interprétation géométrique : approximation affine . . . 1066
12.16 Règles de calcul . . . 1067
12.16.1 Dérivée de la réciproque . . . 1071
12.16.2 Dérivée de fonction composée . . . 1074
12.16.3 Dérivée de fonction périodique . . . 1075
12.17 Dérivation et croissance . . . 1075
12.17.1 Théorèmes de Rolle et des accroissements finis . . . 1077
12.17.2 Règle de l'Hospital . . . 1080
12.17.3 Dérivée et primitive . . . 1081
12.18 Fonctions de plusieurs variables . . . 1082

12.18.1 Graphes de fonctions à plusieurs variables . . . 1085
12.18.2 Courbes de niveau . . . 1086
12.19 Limites à plusieurs variables . . . 1091
12.19.1 Caractérisation de la limite par les suites . . . 1093
12.19.2 Règle de l'étau . . . 1094
12.19.3 Méthode des chemins . . . 1095
12.20 Dérivée directionnelle . . . 1097
12.20.1 Dérivée partielle et directionnelles . . . 1098
12.20.2 Gradient : direction de plus grande pente . . . 1098
12.20.3 Gradient : orthogonal au plan tangent . . . 1099
12.20.4 Mise en bouche en dimension 2 . . . 1099
12.20.5 Accroissements finis et dérivées partielles . . . 1101
12.21 Formes différentielles . . . 1102
12.21.1 Décomposition dans la base duale . . . 1102
12.21.2 L'isomorphisme musical . . . 1103
12.22 Différentielle . . . 1104
12.22.1 Exemples introductifs . . . 1104
12.22.2 Différentielle . . . 1105
12.22.3 Matrice de la différentielle . . . 1106
12.22.4 Différentielle, dual et forme différentielle . . . 1107
12.22.5 Ce n'est pas la différentielle extérieure . . . 1108
12.22.6 Continuité, dérivabilité et différentiabilité . . . 1109
12.22.7 Calcul de valeurs approchées . . . 1112
12.22.8 Différentielle et tangente . . . 1113
12.22.9 Prouver qu'une fonction n'est pas différentiable . . . 1115
12.22.10 Gradient . . . 1119
12.22.11 Linéarité . . . 1119
12.23 Produit . . . 1119
12.23.1 Difficulté d'ordre supérieur . . . 1121
12.23.2 Solution : produit tensoriel . . . 1122
12.23.3 Formes bilinéaires . . . 1122
12.24 Différentielle de fonction composée . . . 1123
12.24.1 Fonctions composées . . . 1125
12.25 Autres trucs sur la différentielle . . . 1126
12.25.1 Différentielle et dérivées partielles . . . 1126
12.25.2 Plan tangent . . . 1126
12.25.3 Notes idéologiques quant au concept de plan tangent . . . 1126
12.25.4 Gradient et recherche du plan tangent . . . 1127
12.26 Jacobienne . . . 1131
12.26.1 Rappels et définitions . . . 1131
12.27 Fonctions de classe C^1 . . . 1132
12.28 Différentielle et dérivée complexe . . . 1141
12.28.1 Quelques règles de calcul . . . 1144
12.29 Théorèmes des accroissements finis . . . 1144
12.30 Fonctions Lipschitziennes . . . 1145
12.31 Différentielles d'ordre supérieur . . . 1147
12.31.1 Différentielle et dérivées partielles . . . 1147
12.31.2 Application à valeurs dans les opérateurs . . . 1151
12.31.3 Espaces d'applications multilinéaires et identifications . . . 1152
12.31.4 Fonctions différentiables plusieurs fois . . . 1156
12.31.5 Différentielle seconde, fonction de classe C^2 . . . 1157
12.31.6 Ordre supérieur . . . 1160
12.32 Suites et séries : généralités . . . 1163

12.32.1 Norme suprémum . . . 1163
12.32.2 Convergence uniforme . . . 1163
12.32.3 Série de fonctions . . . 1167
12.33 Permuter limite et dérivée . . . 1169
12.34 La fonction puissance . . . 1175
12.34.1 Sur les naturels . . . 1176
12.34.2 Sur les rationnels, racines . . . 1177
12.35 Densité des polynômes . . . 1192
12.35.1 Théorème de Stone-Weierstrass . . . 1192
12.36 Primitive de fonction continue . . . 1197
12.36.1 Dérivation de la fonction puissance (première) . . . 1198
12.36.2 Équation fonctionnelle . . . 1200
12.36.3 Dérivation de la fonction puissance (seconde) . . . 1203
12.36.4 Vers les complexes . . . 1204
12.37 Applications analytique entre espaces de Banach . . . 1205
12.37.1 En général . . . 1205
12.37.2 Série de puissance $\mathbb{C}$ . . . 1207
12.37.3 L'opérateur d'inversion est analytique . . . 1207
12.38 Polynômes de Taylor . . . 1211
12.38.1 Fonctions « petit o » . . . 1215
12.38.2 Autres formulations . . . 1217
12.38.3 Formule et reste . . . 1217
12.38.4 Reste intégral . . . 1218
12.39 Développement limité autour de zéro . . . 1218
12.39.1 Généralités . . . 1218
12.39.2 Formule de Taylor-Young . . . 1219
12.39.3 Règles de calcul . . . 1221
12.40 Développement ailleurs qu'à l'origine . . . 1223
12.41 Développement au voisinage de l'infini . . . 1223
12.41.1 La fonction puissance : remarques pour la suite . . . 1223
12.42 Fonctions réelles de deux variables réelles . . . 1223
12.42.1 Limites de fonctions à deux variables . . . 1224
12.42.2 Dérivées partielles . . . 1225
12.42.3 Différentielle et accroissement . . . 1225
12.43 Les fonctions à valeurs vectorielles . . . 1226
12.44 Fonctions vectorielles de plusieurs variables . . . 1226
12.45 Limites à plusieurs variables . . . 1227
12.46 Champs de vecteurs . . . 1230
12.46.1 Matrice jacobienne . . . 1231
12.47 Divergence, rotationnel et l'opérateur nabla . . . 1231
12.48 Interprétation de la divergence . . . 1233
12.49 Quelques formules de Leibniz . . . 1236

13 Analyse sur des groupes **1237**
13.1 Action de groupe et connexité . . . 1237
13.2 Espaces de matrices . . . 1239
13.2.1 Dilatations et transvections . . . 1239
13.2.2 Connexité de certains groupes . . . 1245
13.2.3 Densité . . . 1247
13.2.4 Racine carrée d'une matrice hermitienne positive . . . 1249
13.2.5 Racine carrée d'une matrice symétrique positive . . . 1250
13.2.6 Décomposition polaires : cas réel . . . 1251
13.2.7 Enveloppe convexe . . . 1254
13.2.8 Décomposition de Bruhat . . . 1256

13.3 Sous-groupes du groupe linéaire . . . 1258

14 Tribus, théorie de la mesure, intégration 1263
14.1 Tribus . . . 1263
14.1.1 Généralités . . . 1263
14.1.2 Tribu engendrée . . . 1264
14.1.3 Tribu induite . . . 1265
14.2 Théorie de la mesure . . . 1265
14.2.1 Mesure sur un ensemble de parties . . . 1265
14.2.2 Mesure sur une algèbre de parties . . . 1266
14.2.3 Mesure sur une tribu, espace mesuré . . . 1268
14.2.4 Mesures sigma-finies . . . 1268
14.2.5 Suite du texte . . . 1269
14.2.6 Mesure extérieure . . . 1275
14.3 Applications mesurables . . . 1278
14.3.1 Propriétés . . . 1278
14.3.2 D'une tribu à l'autre . . . 1278
14.4 Tribu borélienne . . . 1279
14.4.1 Applications continues et boréliennes . . . 1281
14.4.2 Tribu de Baire . . . 1283
14.5 Espace mesuré complet . . . 1285
14.5.1 Partie négligeable . . . 1285
14.5.2 Prolongement . . . 1294
14.5.3 Mesure image . . . 1297
14.5.4 Régularité d'une mesure . . . 1298
14.5.5 Théorème de récurrence . . . 1305
14.6 Mesurabilité des fonctions à valeurs réelles . . . 1305
14.6.1 Fonctions à valeurs réelles sur un espace mesurable . . . 1306
14.6.2 Fonction étagée . . . 1312
14.6.3 Fonctions réelles à variables réelles . . . 1316
14.7 Tribu produit . . . 1317
14.7.1 Produit d'espaces mesurables . . . 1317
14.7.2 Le cas des boréliens . . . 1319
14.8 Mesure de Lebesgue sur $\mathbb{R}$ . . . 1320
14.8.1 Mesure et tribu de Lebesgue . . . 1325
14.8.2 Propriétés de la mesure de Lebesgue . . . 1326
14.8.3 Fonctions mesurables . . . 1330
14.8.4 Ensemble de Vitali (non mesurable) . . . 1331
14.8.5 Ensemble de Cantor . . . 1331
14.8.6 Mesure positive sans intervalle . . . 1334
14.9 Intégrale par rapport à une mesure . . . 1335
14.9.1 Définition pour les fonctions à valeurs positives . . . 1335
14.9.2 Premières propriétés . . . 1337
14.9.3 Propriétés plus avancées . . . 1340
14.9.4 Fonctions à valeurs réelles . . . 1344
14.9.5 Additivité de l'intégrale . . . 1346
14.9.6 Fonctions à valeurs vectorielles (dimension finie) . . . 1346
14.9.7 Quelques propriétés . . . 1349
14.9.8 Permuter limite et intégrale . . . 1350
14.9.9 Additivité de l'intégrale de Lebesgque . . . 1353
14.9.10 Produit d'une mesure par une fonction (mesure à densité) . . . 1355
14.9.11 Mesure et topologie . . . 1356
14.10 Mesure à densité . . . 1359
14.10.1 Théorème de Radon-Nikodym . . . 1359

14.10.2 Mesure complexe . . . 1374
14.10.3 Théorème d'approximation . . . 1375
14.11 Produit de mesures . . . 1376
14.12 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$ . . . 1381
14.12.1 Propriétés d'unicité . . . 1382
14.12.2 Régularité . . . 1384
14.13 Propriétés de l'intégrale de Lebesgue . . . 1384
14.13.1 Quelques limites dans les bornes . . . 1385
14.13.2 Mesure de comptage et série . . . 1388
14.13.3 Théorème de la moyenne . . . 1390
14.14 Changement de variables dans une intégrale multiple . . . 1390
14.14.1 Des lemmes . . . 1390
14.14.2 Déterminant et mesure de Lebesgue . . . 1391
14.14.3 Ensembles négligeables . . . 1393
14.14.4 Parties et fonctions mesurables . . . 1394
14.14.5 Le théorème et sa démonstration . . . 1395
14.14.6 Exemples . . . 1400
14.15 Changement d'espace mesuré . . . 1400
14.16 Primitives et intégrales . . . 1401
14.16.1 Théorème fondamental du calcul intégral . . . 1402
14.16.2 Exemples et applications . . . 1403
14.16.3 Intégrales impropres . . . 1405
14.17 Théorème de Fubini-Tonelli et de Fubini . . . 1407
14.17.1 Utiliser Fubini sur $\mathbb{R}^n$ . . . 1415
14.17.2 Absolue continuité et variation bornée . . . 1416
14.17.3 Dérivation dans le bornes . . . 1418
14.17.4 Permuter limite et dérivée . . . 1421

15 Suites et séries de fonctions **1425**
15.1 Séries de fonctions . . . 1425
15.1.1 Intégration de séries de fonctions . . . 1425
15.1.2 Différentiabilité . . . 1426
15.2 Séries entières . . . 1433
15.2.1 Disque de convergence . . . 1433
15.2.2 Somme et produit de séries . . . 1438
15.2.3 Convergence normale . . . 1444
15.2.4 Dérivation . . . 1447
15.2.5 Intégration . . . 1450
15.3 Séries de Taylor . . . 1450
15.3.1 Polynôme de Taylor d'une série entière . . . 1450
15.3.2 Une majoration pour le reste . . . 1451
15.3.3 Fonctions analytiques . . . 1452
15.4 Algèbre engendrée par une matrice . . . 1454
15.5 Exponentielle sur une algèbre normée . . . 1454
15.5.1 Définition . . . 1454
15.5.2 Différentielles . . . 1457
15.5.3 Exponentielle de matrice . . . 1459
15.6 Exponentielle et logarithme dans les réels . . . 1462
15.6.1 L'équation différentielle . . . 1462
15.6.2 Existence . . . 1463
15.6.3 Le nombre de Neper e . . . 1464
15.6.4 Application réciproque : logarithme . . . 1466
15.6.5 Approximations numériques de e . . . 1469
15.6.6 Résumé des propriétés de l'exponentielle . . . 1470

15.6.7 Dérivée de la fonction puissance . . . 1471
15.6.8 Dérivée du logarithme . . . 1472
15.6.9 Taylor pour l'exponentielle . . . 1472
15.6.10 Analycité . . . 1473
15.6.11 Autres propriétés et petits calculs . . . 1473
15.6.12 Taylor pour le logarithme . . . 1473
15.6.13 Développements et calcul de limites . . . 1477
15.6.14 Une petite intégrale . . . 1477
15.7 Vitesses des puissances, de l'exponentielle et du logarithme . . . 1478
15.7.1 Un peu de théorie . . . 1478
15.7.2 Nombres premiers . . . 1484
15.7.3 Quelques limites . . . 1485
15.8 Trigonométrie hyperbolique . . . 1486
15.9 Séries entières de matrices . . . 1488
15.9.1 Différentiabilité . . . 1488
15.10 Exponentielle de matrices . . . 1490
15.10.1 Diagonalisabilité d'exponentielle . . . 1491
15.11 Étude d'asymptote . . . 1493
15.12 Développement en série . . . 1494
15.12.1 Série génératrice d'une suite . . . 1494
15.12.2 Développement en série et Taylor . . . 1494
15.12.3 Resommer une série . . . 1496
15.13 Séries entières de matrices . . . 1501
15.13.1 Rayon de convergence . . . 1501
15.13.2 Convergence et rayon spectral . . . 1501
15.13.3 Exponentielle et logarithme de matrice . . . 1503
15.13.4 Calcul effectif de l'exponentielle d'une matrice . . . 1507
15.13.5 Opérateur d'inversion . . . 1508
15.14 Lemme de Borel . . . 1508
15.14.1 Fonctions plateaux, Urysohn, partition de l'unité . . . 1508
15.14.2 Lemme de Urysohn . . . 1512
15.14.3 Le lemme de Borel . . . 1513
15.15 Nombres de Bell . . . 1514

16 Représentations et caractères **1519**
16.1 Représentations et caractères . . . 1519
16.1.1 Crochet de dualité et transformée de Fourier . . . 1521
16.1.2 Groupes non abéliens . . . 1522
16.1.3 Représentations linéaires des groupes finis . . . 1522
16.1.4 Module . . . 1524
16.1.5 Structure hermitienne . . . 1525
16.1.6 Caractères . . . 1526
16.2 Équivalence de représentations et caractères . . . 1526
16.2.1 Représentation régulière . . . 1529
16.2.2 Caractères et représentations : suite et fin . . . 1530
16.3 Représentation produit tensoriel . . . 1533
16.4 Exemple sur le groupe symétrique . . . 1533
16.5 Table des caractères du groupe symétrique S_4 . . . 1534
16.5.1 Calculs à partir de rien ou presque . . . 1534
16.5.2 À propos de la représentation ρ_s . . . 1536
16.5.3 À propos de la représentation ρ_u . . . 1536

17 Encore de l'analyse (et c'est pas fini) **1539**
17.1 Densité des polynômes . . . 1539

17.2 Primitive et intégrale . . . 1539
17.2.1 Théorème taubérien de Hardy-Littlewood . . . 1540
17.2.2 Théorème de Müntz . . . 1544
17.3 Intégrales convergeant uniformément . . . 1546
17.3.1 Définition et propriété . . . 1546
17.3.2 Critères de convergence uniforme . . . 1547
17.4 Fonctions définies par une intégrale . . . 1548
17.4.1 Continuité sous l'intégrale . . . 1548
17.4.2 Le coup du compact . . . 1549
17.4.3 Dérivabilité sous l'intégrale . . . 1550
17.4.4 Dérivation et intégrale . . . 1552
17.4.5 Différentiabilité sous l'intégrale . . . 1553
17.5 Théorème d'Hadamard . . . 1556
17.6 Deux théorèmes de point fixe . . . 1557
17.6.1 Points fixes attractifs et répulsifs . . . 1557
17.6.2 Picard . . . 1558
17.7 Théorèmes de point fixes et équations différentielles . . . 1560
17.7.1 Théorème de Cauchy-Lipschitz . . . 1560
17.8 Théorèmes d'inversion locale et de la fonction implicite . . . 1568
17.8.1 Mise en situation . . . 1568
17.8.2 Théorème d'inversion locale . . . 1569
17.8.3 Théorème de la fonction implicite . . . 1572
17.8.4 Exemple . . . 1574
17.9 Décomposition polaire (régularité) . . . 1575
17.10 Théorème de Von Neumann . . . 1577
17.11 Recherche d'extrémums . . . 1580
17.11.1 Extrema à une variable . . . 1580
17.11.2 Extrema libre . . . 1581
17.11.3 Extremums et Hessienne . . . 1582
17.11.4 Un peu de recettes de cuisine . . . 1584
17.11.5 Extrema liés . . . 1585
17.12 Fonctions convexes . . . 1587
17.12.1 Inégalité des pentes . . . 1588
17.12.2 Convexité et régularité . . . 1589
17.12.3 Dérivées d'une fonction convexe . . . 1589
17.12.4 Graphe d'une fonction convexe . . . 1591
17.12.5 Convexité et hessienne . . . 1597
17.12.6 Quelques inégalités . . . 1602
17.12.7 Norme l^p . . . 1604
17.12.8 Hölder . . . 1605
17.13 Algorithme du gradient à pas optimal . . . 1610
17.14 Formes quadratiques, signature, et lemme de Morse . . . 1614
17.15 Ellipsoïde de John-Loewner . . . 1618
17.16 Prolongement de fonctions . . . 1624
17.16.1 Encore du prolongement . . . 1626
17.17 Complétion d'un espace métrique . . . 1628
17.17.1 Principe des zéros isolés . . . 1633
17.18 Un petit extra . . . 1634

18 Trigonométrie, isométries **1637**
18.1 Trigonométrie . . . 1637
18.1.1 Définitions, périodicité et quelques valeurs remarquables . . . 1637
18.1.2 Fonction puissance (pour les complexes) . . . 1638
18.1.3 Formules de trigonométrie . . . 1640

18.2 Trucs et astuces de calcul d'intégrales . . . 1650
18.2.1 Quelques intégrales « usuelles » . . . 1651
18.2.2 Reformer un carré au dénominateur . . . 1653
18.2.3 Décomposition en fractions simples . . . 1653
18.3 Très modeste approximation de π . . . 1654
18.3.1 Les fonctions tangente et arc tangente . . . 1655
18.3.2 La fonction arc sinus . . . 1657
18.3.3 La fonction arc cosinus . . . 1659
18.3.4 Une meilleure approximation de π . . . 1661
18.3.5 Angle entre deux vecteurs . . . 1661
18.3.6 Aire du parallélogramme . . . 1662
18.4 Paramétrisation du cercle . . . 1664
18.5 Cercle trigonométriques . . . 1664
18.5.1 Bijection continue . . . 1665
18.5.2 Inverse . . . 1668
18.5.3 Cercle trigonométrique . . . 1668
18.5.4 Du point de vue de la tribu, mesure et co. . . . 1669
18.6 Exemples trigonométriques . . . 1672
18.6.1 Quelques équations trigonométriques . . . 1673
18.6.2 Développements en série . . . 1673
18.7 Isométries de l'espace euclidien . . . 1674
18.7.1 Structure du groupe $\mathrm{Isom}(\mathbb{R}^n)$ . . . 1675
18.8 Isométries dans $\mathbb{R}^n$ . . . 1677
18.8.1 Préserver l'orientation . . . 1681
18.9 Groupes finis d'isométries . . . 1683
18.9.1 Points fixés par une affinité . . . 1684
18.10 Classification des isométries dans $\mathbb{R}^2$ . . . 1686
18.10.1 Projection orthogonale . . . 1686
18.10.2 Réflexions . . . 1686
18.10.3 Segment, plan médiateur et équidistance . . . 1688
18.10.4 Translations et réflexions . . . 1689
18.10.5 Rotations . . . 1690
18.10.6 Rotation d'un angle donné . . . 1693
18.10.7 Rotations vectorielles . . . 1695
18.10.8 Matrice des transformations orthogonales . . . 1697
18.10.9 Rotations, SO(2) et matrice de rotation . . . 1699
18.10.10 Rotation et application affine . . . 1700
18.10.11 Angle orienté . . . 1701
18.10.12 Angles et nombres complexes . . . 1705
18.10.13 Polygone convexe . . . 1709
18.10.14 Groupe diédral . . . 1711
18.10.15 Applications : du dénombrement . . . 1718
18.10.16 Classification . . . 1719
18.10.17 Classification des isométries de $\mathbb{R}$ . . . 1722
18.10.18 Isométries du tétraèdre régulier . . . 1722
18.10.19 Représentation de S_4 via les isométries du tétraèdre . . . 1724
18.11 Transformations de Lorentz . . . 1725
18.11.1 Sous-groupe fini d'isométries du plan . . . 1728
18.11.2 Relations trigonométriques dans un triangle rectangle . . . 1732
18.11.3 Pavages du plan . . . 1734
18.12 Un peu de structure de $\mathrm{O}(n)$ . . . 1749
18.12.1 Valeurs propres dans $\mathrm{O}(n)$ . . . 1749
18.12.2 Sous-groupes finis de SO(3) . . . 1752

18.13 Systèmes de coordonnées . . . 1761
18.13.1 Coordonnées polaires . . . 1761
18.13.2 Coordonnées cylindriques . . . 1769
18.13.3 Coordonnées sphériques . . . 1770
18.14 Calcul de limites . . . 1771
18.14.1 Méthode des coordonnées polaires . . . 1772
18.14.2 Méthode du développement asymptotique . . . 1775
18.15 Quelques intégrales avec de la trigonométrie . . . 1776
18.15.1 Changement de variables . . . 1776
18.15.2 Coordonnées polaires . . . 1777
18.15.3 Coordonnées cylindriques . . . 1778
18.15.4 Un autre système utile . . . 1780
18.16 Aire d'une surface de révolution . . . 1781
18.17 Table de caractères du groupe diédral . . . 1783
18.17.1 Représentations de dimension un . . . 1783
18.17.2 Représentations de dimension deux . . . 1783
18.17.3 Le compte pour n pair . . . 1785
18.17.4 Le compte pour n impair . . . 1786

19 Corps finis, racines de l'unité **1787**
19.1 Le groupe des racines de l'unité dans les nombres complexes . . . 1787
19.1.1 Le groupe . . . 1787
19.1.2 Fonction indicatrice d'Euler . . . 1789
19.1.3 Décomposition en éléments simples . . . 1791
19.2 Chiffrement RSA . . . 1792
19.2.1 Mise en place par Bob . . . 1792
19.2.2 Chiffrement . . . 1792
19.2.3 Déchiffrement . . . 1793
19.2.4 Une imprudence à ne pas commettre . . . 1793
19.2.5 Problèmes calculatoires . . . 1794
19.2.6 La solidité de RSA . . . 1794
19.2.7 Note non mathématique pour doucher l'enthousiasme . . . 1794
19.3 Polynômes cyclotomiques . . . 1795
19.3.1 Définitions et propriétés . . . 1795
19.3.2 Nombres premiers . . . 1799
19.4 Corps finis . . . 1800
19.4.1 Théorème de Wedderburn . . . 1800
19.4.2 Existence, unicité . . . 1802
19.4.3 Symboles de Legendre et carrés . . . 1805
19.4.4 Théorème de Chevalley-Warning . . . 1810
19.4.5 Contenu d'un polynôme . . . 1813
19.4.6 Théorème de l'élément primitif . . . 1813
19.4.7 Construction de $\mathbb{F}_{p^n}$ . . . 1818
19.4.8 Exemple : étude de $\mathbb{F}_{16}$ . . . 1820
19.4.9 Polynômes irréductibles sur $\mathbb{F}_q$ . . . 1822
19.4.10 Matrices . . . 1825
19.5 Constructions à la règle et au compas . . . 1826
19.5.1 Quelques constructions . . . 1826
19.5.2 Nombres constructibles . . . 1828
19.5.3 Polygones constructibles . . . 1830

20 Intégration sur des variétés **1837**
20.1 Variétés . . . 1837
20.1.1 Introduction . . . 1837

20.1.2 Définition, carte . . . 1837
20.1.3 Ancienne définition . . . 1838
20.1.4 Espace tangent . . . 1840
20.2 Intégration . . . 1840
20.2.1 Le problème pour une intégration globale . . . 1840
20.2.2 Intégrale sur une carte . . . 1840
20.2.3 Quelques expressions pour l'élément de volume . . . 1843
20.3 Intégrale sur une variété . . . 1844
20.3.1 Mesure sur une carte . . . 1844
20.3.2 Intégrale sur une carte . . . 1845
20.3.3 Exemples . . . 1846
20.3.4 Orientation . . . 1846
20.4 Longueur, aire, volumes etc. . . . 1848
20.4.1 Quelques aires faciles . . . 1848
20.5 Autres théorèmes de points fixes . . . 1850
20.5.1 Brouwer . . . 1850
20.5.2 Théorème de Schauder . . . 1855
20.5.3 Théorème de Cauchy-Arzella . . . 1856
20.5.4 Théorème de Markov-Kakutani . . . 1857
20.6 Intégrales curvilignes . . . 1858
20.6.1 Chemins de classe C^1 . . . 1858
20.6.2 Intégrer une fonction . . . 1859
20.6.3 Intégrer un champ de vecteurs . . . 1860
20.6.4 Intégrer une forme différentielle sur un chemin . . . 1860
20.6.5 Intégration d'une forme différentielle sur un chemin . . . 1861
20.6.6 Interprétation physique : travail . . . 1863
20.6.7 Intégrer un champ de vecteurs sur un bord en $2D$ . . . 1863
20.6.8 Intégrer une forme différentielle sur un bord en $2D$ . . . 1864
20.6.9 Intégrer une forme différentielle sur un bord en $3D$ . . . 1864
20.6.10 Intégrer un champ de vecteurs sur un bord en $3D$ . . . 1864
20.6.11 Dérivées croisées et forme différentielle exacte . . . 1864
20.7 Surfaces paramétrées . . . 1866
20.7.1 Graphe d'une fonction . . . 1867
20.7.2 Intégrale sur une partie de $\mathbb{R}^m$ . . . 1868
20.8 Intégrales de surface . . . 1868
20.8.1 Intégrale d'un champ de vecteurs . . . 1868
20.9 Aires et intégrales . . . 1869
20.9.1 Aire d'une surface paramétrée . . . 1869
20.9.2 Intégrale d'une fonction sur une surface . . . 1871
20.9.3 Intégrale d'une 2-forme . . . 1871
20.10 Flux d'un champ de vecteurs à travers une surface . . . 1872
20.11 Divergence, Green, Stokes . . . 1874
20.11.1 Théorème de la divergence . . . 1875
20.11.2 Lacets et homotopie . . . 1876
20.11.3 Formule de Green . . . 1876
20.11.4 Formule de Stokes . . . 1878
20.12 Résumé des intégrales vues . . . 1879
20.12.1 L'intégrale d'une fonction sur les réels . . . 1879
20.12.2 Intégrale d'une fonction sur un chemin . . . 1880
20.12.3 Intégrale d'une fonction sur une surface . . . 1880
20.12.4 Intégrale d'une fonction sur un volume . . . 1880
20.12.5 Conclusion pour les fonctions . . . 1881
20.12.6 Circulation d'un champ de vecteurs . . . 1881

20.12.7 Flux d'un champ de vecteurs . . . 1881
20.12.8 Conclusion pour les champs de vecteurs . . . 1882
20.12.9 Attention pour les surfaces fermées ! . . . 1882
20.13 Formes différentielles exactes et fermées . . . 1884
20.14 Théorème d'Abel angulaire . . . 1886
20.15 Passage à la limite sous le signe intégral . . . 1889
20.15.1 Intégrale en dimension un . . . 1889
20.15.2 Intégrales convergentes . . . 1890
20.15.3 La méthode de Rothstein-Trager . . . 1890
20.16 Rappel sur les intégrales usuelles . . . 1896
20.17 Intégrales le long de chemins . . . 1896
20.17.1 Circulation d'un champ de vecteur . . . 1896
20.18 Circulation d'un champ conservatif . . . 1898
20.19 Intégration de fonction à deux variables . . . 1900
20.19.1 Intégration sur un domaine rectangulaire . . . 1900
20.19.2 Intégration sur un domaine non rectangulaire . . . 1901
20.19.3 Changement de variables . . . 1903
20.20 Les intégrales triples . . . 1903
20.20.1 Volume . . . 1905
20.21 Un petit peu plus formel . . . 1906
20.21.1 Intégration sur un domaine non rectangulaire . . . 1906
20.22 Aire et primitive . . . 1908
20.22.1 Longueur d'arc de courbe . . . 1908
20.22.2 Aire de révolution . . . 1909
20.23 L'aire en dessous d'une courbe . . . 1909
20.24 Propriétés des intégrales . . . 1910
20.25 Techniques d'intégration . . . 1911
20.25.1 Intégration par parties . . . 1912
20.25.2 Changement de variables – pour trouver des primitives . . . 1914
20.25.3 Changement de variables – pour calculer des intégrales . . . 1915
20.25.4 Intégrations des fractions rationnelles réduites . . . 1917
20.25.5 Quelques formules à connaitre . . . 1918
20.25.6 Taylor avec reste intégral . . . 1918
20.25.7 Approximation de ln(2) . . . 1922
20.26 Constructions plus naïves de l'intégrale dans le cas réel . . . 1924
20.26.1 Mesure de Lebesgue, version rapide . . . 1924
20.26.2 Pavés et subdivisions . . . 1924
20.26.3 Intégrale d'une fonction en escalier . . . 1927
20.26.4 Intégrales partielles . . . 1928
20.26.5 Réduction d'une intégrale multiple . . . 1928
20.26.6 Propriétés de l'intégrale . . . 1930
20.26.7 Intégrales multiples, cas général . . . 1931
20.26.8 Réduction d'une intégrale multiple . . . 1931
20.26.9 Intégrales sur des parties de $\mathbb{R}^2$ . . . 1933
20.26.10 Intégrales sur des parties de $\mathbb{R}^3$ . . . 1937
20.26.11 Fonctions et ensembles non bornés . . . 1938
20.26.12 Lemme de Morse . . . 1939
20.27 Autres intégrales sympathiques . . . 1941
20.27.1 Intégrale de Wallis . . . 1941
20.27.2 Formule de Stirling . . . 1944
20.27.3 La fonction sinus cardinal, intégrale de Dirichlet . . . 1948

21 Arcs paramétrés **1959**
21.1 Définitions . . . 1959

21.2 Longueur d'arc . . . 1959
21.3 Abscisse curviligne . . . 1962
21.3.1 Formule intégrale de la longueur . . . 1963
21.4 Suite du chapitre . . . 1969
21.5 Autres exemples . . . 1970
21.6 Élément de longueur . . . 1971
21.6.1 Élément de longueur : cartésiennes . . . 1971
21.6.2 Élément de longueur : polaires (1) . . . 1971
21.6.3 Élément de longueur : polaires (2) . . . 1972
21.6.4 Approximation de la longueur par des cordes . . . 1974
21.7 Arc géométrique . . . 1975
21.7.1 Abscisse curviligne et paramétrage normal . . . 1977
21.7.2 Tangente à une courbe paramétrée . . . 1982
21.8 Un peu de topologie . . . 1983
21.9 Repère de Frenet . . . 1986
21.9.1 Torsion . . . 1988
21.10 Hors des coordonnées normales . . . 1989
21.11 Tracer des courbes paramétriques dans $\mathbb{R}^2$ . . . 1992
21.12 Courbes planes . . . 1993
21.12.1 Angle . . . 1993
21.12.2 Courbure signée . . . 1994
21.12.3 Degré, indice et homotopie . . . 1997
21.13 Courbes fermées planes . . . 2004
21.13.1 Cercle circonscrit . . . 2004
21.13.2 Description locale . . . 2006
21.13.3 Enveloppe convexe . . . 2007
21.13.4 Courbure et convexité . . . 2011
21.13.5 Théorème des quatre sommets . . . 2012
21.13.6 Espace topologique normal . . . 2014
21.13.7 Théorème d'Urysohn . . . 2014

22 Géométrie hyperbolique **2019**
22.1 Inversion . . . 2019
22.1.1 Cercles perpendiculaires . . . 2019
22.1.2 Inversion . . . 2020

23 Espaces projectifs **2025**
23.1 Sous espaces projectifs . . . 2025
23.2 Espace projectifs comme « complétés » d'espaces affines . . . 2028
23.3 Théorème de Pappus . . . 2030
23.4 Homographies . . . 2031
23.4.1 Homographies . . . 2031
23.4.2 Le groupe projectif . . . 2032
23.4.3 Repères projectifs . . . 2033
23.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$ . . . 2037
23.4.5 Birapport . . . 2039
23.5 Coordonnées homogènes . . . 2044
23.5.1 Curiosité : matrice de translation . . . 2045
23.5.2 Dualité . . . 2045
23.5.3 Polynômes . . . 2047
23.6 La sphère de Riemann $P_1(\mathbb{C})$ . . . 2048
23.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$ . . . 2049
23.6.2 Homographies . . . 2052
23.6.3 Birapport . . . 2058

23.6.4 Division harmonique . . . 2061
23.6.5 Groupe circulaire . . . 2065
23.6.6 Action du groupe modulaire . . . 2068

24 Analyse vectorielle **2073**
24.1 Le théorème de Green . . . 2073
24.2 Théorème de la divergence dans le plan . . . 2077
24.2.1 La convention de sens de parcours . . . 2077
24.2.2 Théorème de la divergence . . . 2078
24.3 Théorème de Stokes . . . 2078
24.4 Théorème de Gauss . . . 2080
24.5 Coordonnées curvilignes . . . 2082
24.5.1 Base locale . . . 2082
24.5.2 Importance de l'orthogonalité . . . 2082
24.5.3 Coordonnées polaires . . . 2084
24.5.4 Coordonnées cylindriques . . . 2084
24.5.5 Coordonnées sphériques . . . 2085
24.5.6 Gradient en coordonnées curvilignes . . . 2085
24.5.7 Divergence en coordonnées curvilignes . . . 2086
24.5.8 Laplacien en coordonnées curvilignes orthogonales . . . 2089
24.5.9 Rotationnel en coordonnées curvilignes orthogonales . . . 2089
24.6 Les formules . . . 2090
24.6.1 Coordonnées polaires . . . 2090
24.6.2 Coordonnées cylindriques . . . 2090
24.6.3 Coordonnées sphériques . . . 2091

25 Espaces de Hilbert **2093**
25.1 Espaces de Hilbert . . . 2093
25.1.1 Sous-espace vectoriel fermé ? ? ? . . . 2096
25.2 Théorème de la projection . . . 2097
25.3 Systèmes orthogonaux et bases . . . 2100
25.3.1 Orthogonal d'une partie . . . 2100
25.3.2 Dual, théorème de représentation de Riesz . . . 2101
25.3.3 Séparabilité . . . 2102
25.3.4 Base hilbertienne . . . 2104
25.3.5 Décomposition dans une base hilbertienne . . . 2110
25.3.6 Digression sur les normes opérateurs . . . 2114
25.3.7 Applications linéaires et continuité . . . 2115
25.4 Théorème de Kochen-Specker . . . 2117
25.5 Théorème de Lax-Milgram . . . 2118

26 Analyse complexe **2123**
26.1 Fonctions holomorphes . . . 2123
26.1.1 Équations de Cauchy-Riemann . . . 2123
26.1.2 Intégrale sur un chemin dans $\mathbb{C}$ . . . 2126
26.1.3 Intégrales sur des chemins fermés . . . 2126
26.1.4 Homotopie entre applications . . . 2128
26.1.5 Intégrale et homotopie . . . 2130
26.1.6 Théorème de Tietze (espace normal) . . . 2132
26.2 Logarithme complexe . . . 2134
26.2.1 La fonction argument . . . 2134
26.2.2 Une définition possible du logarithme . . . 2137
26.2.3 Pas plus de continuité . . . 2139
26.2.4 Pas d'unicité : autres déterminations de l'argument . . . 2140

26.2.5 Pas d'unicité : développement en série . . . 2142
26.2.6 Pas d'unicité : laquelle choisir ? . . . 2142
26.2.7 Logarithme comme primitive . . . 2142
26.2.8 Logarithme sur un chemin . . . 2144
26.2.9 Lacets, indice et homotopie . . . 2146
26.2.10 Théorème de Cauchy et analycité . . . 2152
26.2.11 Théorème de Brouwer en dimension 2 . . . 2156
26.2.12 Principe des zéros isolés . . . 2157
26.2.13 Prolongement de fonctions holomorphes . . . 2159
26.2.14 Théorème de Runge . . . 2159
26.3 Intégrales de fonctions holomorphes . . . 2161
26.3.1 Holomorphie sous l'intégrale . . . 2162
26.3.2 Mesure de Radon . . . 2166
26.4 Conditions équivalentes à l'holomorphie . . . 2168
26.5 Singularités, pôles et méromorphe . . . 2169
26.6 Dénombrement des solutions d'une équation diophantienne . . . 2173
26.7 Fonctions d'Euler . . . 2179
26.7.1 Euler et factorielle . . . 2182
26.8 Exponentielle et logarithme complexe . . . 2182
26.8.1 Propriétés de l'exponentielle . . . 2182
26.8.2 Intégrale de Fresnel . . . 2183
26.9 Théorème de Weierstrass . . . 2185
26.10 Analyse complexe en plusieurs variables . . . 2187
26.10.1 Inverse de fonctions analytiques . . . 2187

27 Analyse fonctionnelle **2189**
27.1 Théorème d'isomorphisme de Banach . . . 2189
27.2 Théorème d'Ascoli . . . 2189
27.3 Espaces de Lebesgue L^p . . . 2194
27.3.1 Généralités . . . 2194
27.3.2 Un peu de convergence de suites . . . 2198
27.3.3 L'espace L^∞ . . . 2199
27.3.4 Quelques identifications . . . 2201
27.3.5 Inégalité de Young, Jensen, Hölder et de Minkowski . . . 2202
27.3.6 Ni inclusions ni inégalités . . . 2213
27.3.7 Complétude . . . 2214
27.3.8 Théorèmes d'approximation . . . 2220
27.3.9 Densité des fonctions infiniment dérivables à support compact . . . 2221
27.3.10 Approximation . . . 2224
27.3.11 Partie uniformément intégrable . . . 2225
27.4 Autres résultats de convergence . . . 2226
27.5 Convolution . . . 2227
27.5.1 Fonction plateau, lemme d'Urysohn . . . 2231
27.5.2 Partition de l'unité . . . 2233
27.5.3 Approximation de l'unité . . . 2234
27.6 Intégrale sur des variétés . . . 2236
27.6.1 Formes différentielles . . . 2236
27.6.2 Intégrale d'une fonction sur une sous-variété . . . 2237
27.6.3 Densité des polynômes trigonométriques . . . 2238
27.7 Espaces L^2, généralités . . . 2239
27.8 L'espace $L^2(\mathbb{R}^d)$ . . . 2241
27.9 L'espace $L^2(S^1)$ . . . 2241
27.9.1 Espace mesuré . . . 2242
27.9.2 Topologie . . . 2243

27.9.3 Système trigonométrique . . . 2249
27.9.4 Convolution . . . 2250
27.9.5 Approximation de l'unité . . . 2252
27.9.6 Base hilbertienne (suite des polynômes trigonométriques) . . . 2256
27.9.7 Convolution, bis . . . 2258
27.10 L'espace de lebesgue $L^2([a,b])$ . . . 2258
27.11 Sur $[0,2\pi[$ . . . 2261
27.12 Sur $[-T,T[$ . . . 2262
27.12.1 Le cas dans $[0,2\pi]$ . . . 2262
27.13 Des inégalités . . . 2264
27.13.1 c-r inégalité . . . 2264
27.13.2 Inégalité de Hanner . . . 2267
27.13.3 Inégalités de Clarkson . . . 2272
27.14 Théorème de la projection normale . . . 2276
27.14.1 Espace uniformément convexe . . . 2276
27.14.2 Uniforme convexité des espaces de Lebesgue . . . 2279
27.14.3 Théorème de la projection normale . . . 2279
27.15 Jauge . . . 2281
27.16 Théorèmes de Hahn-Banach . . . 2283
27.16.1 Applications $\mathbb{R}$-linéaires et $\mathbb{C}$-linéaires . . . 2283
27.16.2 Hahn-Banach, théorème d'extension dominée . . . 2285
27.16.3 Hyperplan séparateur . . . 2288
27.16.4 Prolongement de fonctionnelles (dimension finie) . . . 2290
27.16.5 Prolongement de fonctionnelles (dimension infinie) . . . 2293
27.17 Théorème de Tietze . . . 2295
27.17.1 Pas de bicontinues entre dimensions différentes . . . 2298
27.18 Dualité, réflexivité et théorème de représentation de Riesz . . . 2301
27.18.1 Le théorème de Jordan . . . 2315
27.18.2 Théorème de Jordan . . . 2328
27.19 Topologie faible . . . 2329
27.19.1 Espace de Banach réflexif . . . 2333
27.19.2 Espaces L^∞ . . . 2333
27.20 Espace de Schwartz . . . 2334
27.20.1 Topologie . . . 2336
27.20.2 Produit de convolution . . . 2338
27.21 Théorème de Montel . . . 2338
27.22 Espaces de Bergman . . . 2340

28 Séries de Fourier **2345**
28.1 Densité des polynômes trigonométriques . . . 2345
28.1.1 Convergence pour les fonctions continues (via Weierstrass) . . . 2345
28.1.2 Convergence pour les fonctions continues (via Fejér) . . . 2345
28.1.3 Densité dans L^p . . . 2349
28.1.4 Suite équirépartie, critère de Weyl . . . 2349
28.2 Fonctions de Dirichlet . . . 2352
28.3 Coefficients et série de Fourier . . . 2353
28.3.1 Le contre-exemple que nous attendions tous . . . 2356
28.3.2 Inégalité isopérimétrique . . . 2359
28.3.3 À propos des coefficients . . . 2361
28.4 Série de Laurent . . . 2362

29 Transformation de Fourier **2369**
29.1 Transformée de Fourier sur $L^1(\mathbb{R}^d)$ . . . 2370
29.1.1 Formule sommatoire de Poisson . . . 2373

29.2 Suite régularisante . . . 2376
29.3 Transformée de Fourier dans l'espace de Schwartz . . . 2378
29.3.1 Quelques transformées de Fourier . . . 2380
29.3.2 Formule d'inversion . . . 2383
29.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$ . . . 2386
29.4.1 Le problème . . . 2387
29.4.2 Extension de $L^1 \cap L^2$ vers L^2 . . . 2387
29.4.3 Une formule de Leibniz . . . 2389

30 Distributions **2391**
30.1 Dérivée faible . . . 2392
30.1.1 Dérivée partielle au sens faible . . . 2392
30.1.2 Dérivée faible partielle . . . 2394
30.1.3 Formules de Leibniz . . . 2395
30.2 Topologie et convergence sur des espaces de fonctions . . . 2396
30.2.1 Limite inductive . . . 2396
30.2.2 Les espaces classiques . . . 2396
30.2.3 Espaces de fonctions et topologie inductive . . . 2403
30.2.4 Convergence dans les fonctions test . . . 2403
30.3 Distributions . . . 2405
30.3.1 Multiplication d'une distribution par une fonction . . . 2408
30.3.2 Dérivée de distribution . . . 2408
30.3.3 Ordre et support d'une distribution . . . 2410
30.4 L'espace $C^\infty(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d))$ . . . 2412
30.4.1 Dérivation . . . 2415
30.5 Une équation de distribution . . . 2415
30.6 Localisation, principe de recollement . . . 2416
30.7 Permuter distributions, dérivées et intégrales . . . 2421
30.8 Distributions tempérées . . . 2425
30.8.1 Topologie . . . 2427
30.8.2 Distributions associées à des fonctions . . . 2427
30.8.3 Composition avec une fonction . . . 2428
30.8.4 Transformée de Fourier d'une distribution tempérée . . . 2428
30.8.5 Convolution d'une distribution par une fonction . . . 2428
30.8.6 Approximation de la distribution de Dirac . . . 2429
30.8.7 Peigne de Dirac . . . 2432
30.9 L'espace $C^\infty(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d))$ . . . 2433
30.9.1 Propriétés générales . . . 2433
30.9.2 Dérivation . . . 2436

31 Espaces de Sobolev, équations elliptiques **2439**
31.1 Espaces de Sobolev . . . 2439
31.1.1 Sur un intervalle de $\mathbb{R}$ . . . 2439
31.1.2 Sur un ouvert de $\mathbb{R}^n$ . . . 2444
31.1.3 Espace de Sobolev fractionnaire . . . 2446
31.2 Trace . . . 2448
31.3 Théorème de plongement . . . 2451

32 Équations différentielles ordinaires **2455**
32.1 Équation homogène, solution particulière . . . 2456
32.2 Que faire avec $f(z)dz = g(t)dt$? . . . 2457
32.3 Équations linéaires du premier ordre . . . 2458
32.3.1 Pourquoi la variation des constantes fonctionne toujours ? . . . 2459
32.4 Équations à variables séparées . . . 2460

32.4.1 La méthode rapide . . . 2460
32.4.2 La méthode plus propre . . . 2461
32.4.3 Les théorèmes . . . 2461
32.5 Équations linéaires d'ordre supérieur . . . 2463
32.5.1 Équations et systèmes linéaires à coefficients constants . . . 2463
32.5.2 Si les coefficients ne sont pas constants ? . . . 2464
32.6 Système d'équations linéaires . . . 2464
32.6.1 La magie de l'exponentielle. 2464
32.6.2 . . .mais la difficulté . . . 2465
32.6.3 La recette . . . 2465
32.6.4 Système d'équations linéaires avec matrice constante . . . 2466
32.6.5 Système d'équations linéaires avec matrice non constante . . . 2466
32.7 Réduction de l'ordre . . . 2466
32.8 Autour de Cauchy-Lipschitz . . . 2469
32.8.1 Fuite des compacts et explosion en temps fini . . . 2469
32.8.2 Écart entre deux conditions initiales . . . 2471
32.8.3 Flot d'un champ de vecteurs . . . 2472
32.8.4 Stabilité de Lyapunov . . . 2486
32.8.5 Système proies-prédateurs de Lotka-Volterra . . . 2490
32.9 Équation du second ordre . . . 2493
32.9.1 Wronskien . . . 2493
32.9.2 Avec second membre . . . 2494
32.9.3 Équation $y'' + q(t)y = 0$ . . . 2494
32.9.4 Équation de Hill . . . 2496
32.10 Différents types d'équations différentielles . . . 2499
32.10.1 Équation homogène . . . 2499
32.10.2 Équation de Bernoulli . . . 2499
32.10.3 Équation de Riccati . . . 2499
32.10.4 Équation différentielle exacte . . . 2500
32.11 Distributions pour les équations différentielles . . . 2501
32.11.1 Équation de Schrödinger . . . 2501
32.12 Équations différentielles du premier ordre . . . 2504
32.13 Premier ordre, variables séparables . . . 2507
32.14 Équations différentielles linéaires du premier ordre . . . 2509
32.14.1 Méthode de variation de la constante . . . 2511
32.15 Équations différentielles linéaires du second ordre . . . 2512
32.15.1 Équations différentielles linéaires du second ordre homogènes à coefficients constants . . . 2512
32.15.2 Linéaires du second ordre à coefficients constants, non homogènes . . . 2514
32.16 Fonction de Green . . . 2515

33 Équations aux dérivées partielles 2519
33.1 Symbole principal, équation des caractéristiques . . . 2519
33.2 Méthode des caractéristiques pour l'ordre 1 . . . 2519
33.2.1 Un exemple complet un peu minimal . . . 2520
33.2.2 Un théorème d'existence et d'unicité . . . 2522
33.3 Méthode des caractéristiques pour l'ordre 2 . . . 2526
33.3.1 Principe général . . . 2526
33.3.2 Exemple : l'équation d'onde . . . 2527
33.4 Classification des équations du second ordre . . . 2528
33.4.1 Problème au limite . . . 2529
33.5 Principe du maximum . . . 2530
33.6 Quelques exemples . . . 2534
33.6.1 Un changement de variables . . . 2534

34 Numérique **2537**
34.1 Introduction . . . 2537
34.2 Représentations numériques . . . 2537
34.2.1 Entier relatif en complément à deux (binaire) . . . 2537
34.2.2 Représentation en virgule flottante . . . 2539
34.2.3 Simple précision, IEEE-754 . . . 2539
34.3 Problèmes pour écrire des nombres . . . 2542
34.3.1 Troncature : la base . . . 2542
34.3.2 Troncature : le drift . . . 2543
34.3.3 Quelques bonnes règles . . . 2544
34.3.4 Erreur de “cancellation” . . . 2544
34.3.5 Calcul d’une dérivée . . . 2546
34.3.6 Erreur d’absorption . . . 2546
34.4 Conditionnement et stabilité . . . 2547
34.4.1 Comment choisir et penser le K ? . . . 2550
34.5 Un peu de points fixes . . . 2551
34.5.1 Choix de la fonction à point fixe . . . 2551
34.5.2 Convergence quadratique . . . 2552
34.5.3 Convergence . . . 2554
34.6 Méthode de Newton . . . 2555
34.6.1 « Justification » par la formule de Taylor . . . 2555
34.6.2 « Justification » par points fixes . . . 2556
34.6.3 Convergence de la méthode de Newton . . . 2557
34.6.4 Racine carrée par la méthode de Newton . . . 2558
34.6.5 Formalisation de l’algorithme . . . 2559
34.6.6 Caractéristiques . . . 2560
34.6.7 Exemple de la racine carrée . . . 2560
34.6.8 Si multiplicité . . . 2561
34.6.9 Et la dérivée ? . . . 2561
34.6.10 Méthode de Newton : le cas général . . . 2561
34.7 Estimation de l’ordre de convergence . . . 2563
34.8 Autres méthodes . . . 2564
34.8.1 Méthode de Schröder . . . 2564
34.8.2 Halley . . . 2564
34.9 Méthode des sécantes variables . . . 2564
34.9.1 Aitken . . . 2565
34.10 Équations algébrique . . . 2566
34.10.1 Résoudre un système linéaire . . . 2566
34.10.2 Caractéristiques . . . 2566
34.10.3 Définitions . . . 2567
34.11 Équations non linéaires . . . 2567
34.11.1 Méthode de bisection . . . 2569
34.12 Efficacité . . . 2571
34.13 Exemples sous forme d’exercices . . . 2571
34.14 Approximations de fonctions . . . 2575
34.14.1 Critère d’interpolation . . . 2575
34.14.2 Base de Newton . . . 2576
34.14.3 Méthode des minimums quadratiques . . . 2577
34.14.4 Notre espace de Hilbert . . . 2578
34.14.5 Droite de régression . . . 2579
34.15 Conditionnement d’une matrice . . . 2580
34.15.1 Perturbation du vecteur . . . 2582
34.15.2 Perturbation de la matrice . . . 2584

34.16 Système linéaires (généralités) . . . 2585
34.16.1 Les méthodes directes . . . 2585
34.16.2 Méthodes itératives . . . 2586
34.17 Système linéaires (méthodes directes) . . . 2586
34.17.1 Inversion de matrice triangulaire . . . 2586
34.17.2 Transformation gaussienne . . . 2587
34.17.3 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 2588
34.17.4 Méthode de Gauss sans pivot (décomposition LU) . . . 2589
34.17.5 Matrice de permutation élémentaire . . . 2593
34.18 Méthode de Gauss avec pivot partiel (décomposition PLU) . . . 2594
34.18.1 L'idée . . . 2594
34.18.2 Le théorème . . . 2595
34.18.3 D'un point de vue algorithmique . . . 2598
34.18.4 Exemples . . . 2600
34.19 Résolution de systèmes linéaires (suite) . . . 2603
34.19.1 Déterminant . . . 2603
34.19.2 Plusieurs termes indépendants . . . 2603
34.19.3 Cholesky . . . 2604
34.20 Système linéaire (méthodes itératives) . . . 2607
34.20.1 La méthode générale . . . 2608
34.20.2 Jacobi . . . 2608
34.20.3 Gauss-Seidel . . . 2608
34.20.4 Autres . . . 2608
34.21 Indices connectés, matrice irréductible . . . 2608
34.22 Localisation des valeurs propres . . . 2610
34.22.1 Matrices à diagonale dominante . . . 2613
34.22.2 M-matrice . . . 2615

35 Méthode des différences finies 2619
35.1 Problèmes de dimension un . . . 2619
35.1.1 Un schéma à cinq points . . . 2620
35.1.2 Exemple . . . 2625
35.2 Problèmes de dimension deux . . . 2625
35.2.1 Discrétisation en croix . . . 2626
35.2.2 Discrétisation en carré . . . 2627
35.2.3 Résolution de la discrétisation en croix . . . 2628
35.3 Consistance, convergence . . . 2630
35.3.1 Définitions, mise en place . . . 2630
35.3.2 Exemple . . . 2632
35.3.3 Consistance, stabilité et convergence . . . 2633
35.3.4 Exemple : schéma à cinq points, laplacien en croix . . . 2634
35.4 Autres laplaciens . . . 2635
35.4.1 Travail avec le laplacien à 9 points . . . 2638

36 Variables aléatoires et théorie des probabilités 2639
36.1 Espace de probabilité . . . 2639
36.2 Variables aléatoires . . . 2640
36.2.1 Loi d'une variable aléatoire . . . 2640
36.2.2 Indépendance . . . 2641
36.2.3 Lois conjointes et indépendance . . . 2644
36.2.4 Densité jointe . . . 2647
36.2.5 Espérance . . . 2650
36.2.6 Moment . . . 2652
36.2.7 Somme et produit de variables aléatoires indépendantes . . . 2652

36.2.8 Variance . . . 2655
36.2.9 Covariance . . . 2656
36.2.10 Probabilité conditionnelle : événements . . . 2656
36.2.11 Espérance conditionnelle . . . 2659
36.2.12 Probabilité conditionnelle : tribu . . . 2667
36.2.13 Variables de Rademacher indépendantes . . . 2669
36.2.14 Un petit paradoxe . . . 2671
36.2.15 Inégalité de Jensen . . . 2677
36.2.16 Fonction de répartition . . . 2678
36.2.17 Fonction caractéristique . . . 2678
36.2.18 Fonction génératrice des moments, transformée de Laplace . . . 2679
36.2.19 Théorème de transfert . . . 2680
36.2.20 Changement de variables . . . 2683
36.3 Convergence . . . 2683
36.3.1 Convergence et fonction caractéristique . . . 2688
36.4 Loi des grands nombres, théorème central limite . . . 2693
36.4.1 Loi des grands nombres . . . 2693
36.4.2 Théorème central limite . . . 2695
36.4.3 Marche aléatoire . . . 2698
36.5 Les lois usuelles . . . 2699
36.5.1 Loi de Bernoulli . . . 2699
36.5.2 Loi binomiale . . . 2699
36.5.3 Loi multinomiale . . . 2700
36.5.4 Loi géométrique . . . 2701
36.5.5 Loi de Poisson . . . 2701
36.5.6 Loi exponentielle . . . 2702
36.5.7 Approximation de la binomiale par une Poisson . . . 2706
36.5.8 Loi de Poisson et loi exponentielle . . . 2706
36.5.9 Loi normale . . . 2708
36.5.10 Vecteurs gaussiens . . . 2709
36.5.11 Variable aléatoire de Rademacher . . . 2714
36.5.12 Loi de Student . . . 2715
36.5.13 Indépendance, covariance et variance de somme . . . 2716
36.6 Estimation des grands écarts . . . 2716
36.7 Simulations de réalisations de variables aléatoires . . . 2720
36.7.1 Générateur uniforme . . . 2720
36.7.2 Simulation par inversion . . . 2721
36.7.3 Algorithme de Box-Muller . . . 2722
36.7.4 Méthode du rejet . . . 2722
36.7.5 Simuler une loi géométrique à l'ordinateur . . . 2724
36.7.6 Simuler une loi exponentielle à l'ordinateur . . . 2724
36.7.7 Simuler une loi de Poisson à l'ordinateur . . . 2724
36.8 Sage . . . 2725
36.8.1 Loi exponentielle . . . 2725
36.8.2 Inverser des lois . . . 2725
36.9 Monte-Carlo . . . 2726
36.9.1 Intervalle de confiance . . . 2727
36.10 Résultats qui se démontrent avec des variables aléatoires . . . 2729
36.10.1 Nombres normaux . . . 2729
36.10.2 Théorème de Bernstein . . . 2731

37 Statistiques **2737**
37.1 Notations et hypothèses . . . 2737
37.2 Modèle statistique . . . 2737

37.3 Modèles d'échantillonnages . . . 2740
37.4 Estimation ponctuelle . . . 2743
37.5 Statistiques et estimateurs . . . 2745
37.5.1 Qualité des estimateurs . . . 2745
37.5.2 Méthode des moments . . . 2746
37.5.3 Méthode de substitution . . . 2747
37.5.4 Méthode du maximum de vraisemblance . . . 2748
37.5.5 Exemples sous forme d'exercices . . . 2748
37.5.6 Estimation d'une fonction de répartition . . . 2750
37.5.7 Exemples sous forme d'exercices . . . 2751
37.5.8 Espérance et variance d'un estimateur . . . 2752
37.6 Estimation par intervalle de confiance . . . 2753
37.6.1 Région de confiance . . . 2756
37.6.2 Fonction pivotale . . . 2756
37.6.3 Sondage de proportion . . . 2759
37.7 Estimer une densité lorsqu'on ne sait rien . . . 2760
37.7.1 Distance entre des mesures . . . 2761
37.7.2 Estimateur par fenêtres glissantes . . . 2762
37.8 Test d'hypothèses, prise de décision . . . 2764
37.8.1 Exemple : qualité des pièces d'usine . . . 2764
37.8.2 Exemple : la résistance d'un fil . . . 2764
37.8.3 Vocabulaire et théorie . . . 2765
37.8.4 Risque de première et seconde espèce . . . 2766
37.8.5 Modèle paramétrique de loi gaussienne . . . 2767
37.9 Tests paramétriques . . . 2768
37.10 Tests d'adéquation . . . 2769

38 Chaînes de Markov à temps discret **2775**
38.1 Généralités . . . 2775
38.1.1 Matrice stochastique . . . 2777
38.2 Chaînes de Markov sur un ensemble fini . . . 2778
38.2.1 Graphe de transition . . . 2780
38.2.2 Nombre de visites . . . 2780
38.2.3 Récurrent et transient . . . 2782
38.2.4 Chaînes irréductibles . . . 2783
38.2.5 Périodique et irréductible . . . 2788
38.3 Marche aléatoire sur $\mathbb{Z}$ . . . 2790
38.3.1 Chaînes de Markov homogènes . . . 2792
38.3.2 Chaîne de Markov définie par récurrence . . . 2794
38.4 Classification des états . . . 2798
38.5 Mesure invariante . . . 2800
38.6 Convergence vers l'équilibre . . . 2802
38.7 Processus de Galton-Watson . . . 2803

39 Martingales **2807**
39.1 Convergence de martingales . . . 2807
39.2 Temps d'arrêt et martingale terminée . . . 2810
39.2.1 Nombre de montées . . . 2810
39.2.2 Théorèmes d'arrêt . . . 2814
39.3 Décomposition de martingales . . . 2817
39.4 Problème de la ruine du joueur . . . 2818
39.4.1 Le cas où la pièce est truquée . . . 2819
39.4.2 Le cas où la pièce est non truquée . . . 2823
39.4.3 Un petit complément . . . 2825

40 Processus de Poisson **2827**
40.1 Processus de Poisson . . . 2827
40.1.1 Processus a-Poisson . . . 2827
40.1.2 Processus c-Poisson . . . 2827
40.1.3 Processus n-Poisson . . . 2828
40.1.4 Processus o-Poisson . . . 2828
40.1.5 Équivalence de définitions . . . 2843
40.1.6 Estimateurs . . . 2843
40.2 Quelques trucs sur la simulation . . . 2845
40.2.1 Le théorème central limite pour Markov . . . 2845
40.2.2 Feuille 5 . . . 2845
40.2.3 Feuille 6 . . . 2845
40.2.4 Feuille 7 . . . 2846
40.2.5 Simuler des lois conditionnelles . . . 2846

41 Langages **2847**
41.1 Alphabets et mots . . . 2847
41.2 Langages . . . 2848

42 Utilisation dans les autres sciences **2851**
42.1 Démystification du MRUA . . . 2851
42.1.1 Preuve de la formule . . . 2851
42.1.2 Interprétation graphique . . . 2852
42.2 Relativité en mécanique newtonienne . . . 2852
42.2.1 Relativité du mouvement . . . 2852
42.2.2 Bob et Alice . . . 2852
42.3 Invariance de la vitesse de la lumière . . . 2853
42.3.1 Champ de gravitation et électrique . . . 2853
42.3.2 Support du champ : pas d'éther . . . 2853
42.3.3 Le problème . . . 2854
42.4 Conséquences . . . 2854
42.4.1 Ligne d'univers . . . 2854
42.4.2 Transformations de Lorentz . . . 2855
42.4.3 Conditions d'existence . . . 2858
42.4.4 La notion d'intervalle . . . 2859
42.4.5 Le cône de lumière d'un point . . . 2859
42.4.6 Contraction des longueurs . . . 2860
42.4.7 Dilatation des intervalles de temps . . . 2860
42.4.8 Invariance de l'intervalle . . . 2861
42.4.9 Vitesse limite . . . 2865
42.5 Applications . . . 2865
42.5.1 Le GPS . . . 2866
42.5.2 Les ondes électromagnétiques . . . 2866
42.6 Mécanique relativiste . . . 2866
42.6.1 Des problèmes, toujours des problèmes . . . 2866
42.6.2 Loi d'addition des vitesses . . . 2867
42.6.3 L'action d'une force . . . 2867
42.6.4 Équivalence entre la masse et l'énergie . . . 2869
42.7 Principe de correspondance . . . 2869

43 Exemples avec Sage **2871**
43.1 Graphiques . . . 2871
43.1.1 Autres . . . 2871

44 Épilogue : la constante de Weiner **2891**
44.0.1 Résultats prouvés dans le Frido . . . 2891
44.0.2 Autres résultats . . . 2892
44.0.3 Conjectures . . . 2892

45 Développements possibles **2893**
45.1 Algèbre et géométrie . . . 2893
45.2 Analyse . . . 2897
45.3 Anciennes leçons . . . 2900

II Giulietta **2909**

46 Some old results **2911**
46.1 Some old questions . . . 2911
46.2 Still some questions and problems . . . 2911
46.3 Some statements waiting for proofs . . . 2911

47 Categories **2915**
47.1 Functor . . . 2915
47.1.1 Functor and equivalence . . . 2915
47.2 Direct limit . . . 2915
47.2.1 Direct limit of sets . . . 2915
47.2.2 Direct limit of vector spaces . . . 2916
47.2.3 Direct limit in categories . . . 2916
47.3 Categories with tensor product . . . 2917

48 Topology **2921**
48.1 Main definitions . . . 2921
48.1.1 Separability axioms . . . 2921
48.1.2 Axioms of countability . . . 2921
48.2 Topology and convergence . . . 2925
48.3 Locally convex . . . 2926
48.3.1 Topological group . . . 2928
48.4 Metrisable groups . . . 2928
48.5 One point compactification . . . 2929
48.5.1 Example: the Riemann sphere . . . 2929
48.6 Topological approximation . . . 2930
48.6.1 Introduction . . . 2930
48.6.2 Generalization . . . 2931
48.6.3 Order and topology . . . 2931
48.7 Partition of unity . . . 2932
48.8 Directed sets and net . . . 2933
48.9 Homotopy group . . . 2934
48.10 Covering spaces . . . 2934
48.10.1 Universal covering . . . 2934
48.10.2 Monodromy action . . . 2935
48.11 Haar measure . . . 2935
48.11.1 Modular function . . . 2943
48.11.2 Riesz-Markov representation theorem . . . 2944

49 Manifolds **2961**
49.1 Topological manifold . . . 2961
49.1.1 Isomorphisms . . . 2962
49.1.2 Manifold from maps . . . 2963

49.1.3 Topology . . . 2965
49.1.4 Covering with compact closure . . . 2965
49.2 Differentiable manifolds . . . 2969
49.2.1 Definition, charts . . . 2969
49.2.2 Topology . . . 2971
49.2.3 Regularity . . . 2972
49.2.4 Partition of unity . . . 2973
49.2.5 Product of manifolds . . . 2976
49.3 Tangent vector . . . 2977
49.3.1 Tangent vector . . . 2977
49.3.2 Vector space structure on the tangent space . . . 2978
49.3.3 Basis vectors . . . 2979
49.3.4 Tangent vector on a product manifold . . . 2982
49.4 Differential of a map . . . 2982
49.4.1 Chain rule and inverse . . . 2988
49.4.2 Topology on a tangent space . . . 2989
49.4.3 Smooth invariance of the domain . . . 2990
49.5 Vector field . . . 2990
49.5.1 Vector field . . . 2993
49.6 Tangent and cotangent bundle . . . 2993
49.6.1 Dual basis . . . 2995
49.6.2 Field of basis . . . 2996
49.6.3 Commutator of vector fields . . . 2998
49.7 Submanifold . . . 3000
49.8 Flow and integral curve . . . 3003
49.8.1 Next stuff . . . 3006
49.8.2 Differential of a map . . . 3007
49.8.3 Examples . . . 3008
49.8.4 Some Leibniz formulas . . . 3008
49.8.5 Cotangent bundle . . . 3009
49.8.6 Immersion, embedding . . . 3009
49.9 Distribution and Frobenius theorem . . . 3011
49.9.1 Analytic . . . 3015
49.9.2 Tensor algebra . . . 3015
49.10 Rank theorem . . . 3015
49.11 Exterior calculus . . . 3019
49.11.1 The exterior algebra . . . 3019
49.11.2 Differential of k-forms . . . 3020
49.11.3 Musical isomorphism . . . 3021
49.11.4 Pull-back and push-forward . . . 3022
49.12 Lie derivative . . . 3023
49.12.1 Notions of topology . . . 3024
49.12.2 Differential, path and extensions . . . 3025
49.12.3 Vector fields . . . 3025
49.12.4 Cohomology . . . 3025
49.12.5 Integral curve . . . 3028
49.12.6 Christoffel symbols . . . 3029
49.12.7 Commutator of vector field . . . 3030
49.13 Exercises . . . 3030

50 Representations **3043**
50.1 Definitions . . . 3043
50.1.1 Complex conjugate representation . . . 3043
50.1.2 Faithful representation . . . 3044

50.1.3 Schur lemma . . . 3044
50.1.4 Irreducible representations . . . 3045
50.1.5 Direct sum of representations . . . 3045
50.1.6 Tensor product of representations . . . 3046

51 Lie algebras **3049**
51.1 Adjoint representation . . . 3050
51.2 Representation of the complex algebra . . . 3051
51.3 Jordan decomposition . . . 3052
51.4 Killing form . . . 3053
51.5 Solvable and nilpotent algebras . . . 3054
51.6 Flags and nilpotent Lie algebras . . . 3060
51.7 Semisimple Lie algebras . . . 3061
51.7.1 Cartan criterion . . . 3063
51.7.2 More about radical . . . 3064
51.7.3 Lorentz algebra . . . 3066
51.8 Clifford algebra . . . 3067
51.8.1 Definition and universal problem . . . 3067
51.8.2 Representations of the algebra $\mathfrak{su}(2) = \mathfrak{so}(3)$ . . . 3068
51.9 Cartan subalgebras in complex Lie algebras . . . 3070
51.10 Root spaces in semisimple complex Lie algebras . . . 3073
51.10.1 Introduction and notations . . . 3073
51.10.2 Root spaces . . . 3074
51.10.3 Generators . . . 3079
51.10.4 Subalgebra $\mathfrak{sl}(2)_i$ . . . 3080
51.10.5 Chevalley basis . . . 3081
51.10.6 Coefficients in the Cartan matrix . . . 3084
51.10.7 Simple roots . . . 3085
51.10.8 Weyl group . . . 3085
51.10.9 Abstract root system . . . 3088
51.10.10 Abstract Cartan matrix . . . 3095
51.10.11 Dynkin diagrams . . . 3096
51.10.12 Example of reconstruction by hand . . . 3101
51.10.13 Reconstruction . . . 3103
51.10.14 Cartan-Weyl basis . . . 3106
51.10.15 Cartan matrix . . . 3107
51.11 Other results . . . 3109
51.11.1 Abstract Cartan matrix . . . 3109
51.11.2 About group representations . . . 3110
51.11.3 Modules and reducibility . . . 3111
51.11.4 Weight and dual spaces . . . 3111
51.11.5 Dynkin diagram . . . 3113
51.11.6 Weyl: other results . . . 3125
51.11.7 Longest element . . . 3125
51.11.8 Weyl group and representations . . . 3126
51.11.9 Chevalley basis (deprecated) . . . 3127
51.12 Iwasawa decomposition of Lie groups . . . 3128
51.12.1 Cartan decomposition . . . 3128
51.12.2 Root space decomposition . . . 3129
51.12.3 Positivity, convex cone and partial ordering . . . 3130
51.12.4 Iwasawa decomposition . . . 3130
51.13 Other results about Cartan algebras . . . 3131
51.14 Representations . . . 3136
51.14.1 List of the weights of a representation . . . 3137

51.14.2 Tensor product of representations . . . 3139
51.15 Semi-direct product . . . 3142
51.15.1 From Lie algebra point of view . . . 3142
51.15.2 Group algebra . . . 3144

52 Lie groups **3145**
52.1 Lie groups . . . 3145
52.1.1 Connected component of Lie groups . . . 3146
52.1.2 What is $g^{-1}dg$? . . . 3147
52.1.3 Product of Lie groupe . . . 3147
52.1.4 Left and right translations . . . 3147
52.2 Two words about Lie algebra . . . 3147
52.2.1 Adjoint map . . . 3147
52.3 Universal enveloping algebra . . . 3149
52.3.1 Adjoint map in $\mathcal{U}(\mathcal{G})$ . . . 3150
52.3.2 Invariant fields . . . 3151
52.3.3 Representation of Lie groups . . . 3151
52.4 Lie subgroup . . . 3152
52.5 Semi-direct product of Lie groups . . . 3153
52.5.1 Introduction by exact short sequence . . . 3155

53 Lie group and Lie algebra **3157**
53.1 Lie algebra of a Lie group . . . 3157
53.2 Matrix Lie group and its algebra . . . 3161
53.3 Fundamental vector field . . . 3167
53.4 Invariant vector fields . . . 3168
53.5 Exponential map . . . 3169
53.5.1 Integral curve . . . 3169
53.5.2 Integral curve and exponential . . . 3171
53.5.3 Flow and exponential . . . 3171
53.5.4 Invariant vector and derivation . . . 3175
53.6 Exponential as bijection on a neighborhood . . . 3176
53.6.1 Analytic Lie group, Taylor formula . . . 3176
53.6.2 Exponential and subspace . . . 3181
53.6.3 Campbell-Baker-Hausdorff formula . . . 3185
53.6.4 Smooth and analytic Lie group . . . 3185
53.6.5 Topological Lie subgroup . . . 3186
53.6.6 Action from Lie algebra to Lie group . . . 3186
53.6.7 Example: determining the smooth isometries of the flat vector space . . . 3187
53.6.8 Other stuff . . . 3187
53.6.9 Lie algebra of a Lie subgroup . . . 3189
53.7 Lifting a Lie subalgebra . . . 3189
53.8 Covering . . . 3190
53.9 Representations from the Lie algebra to the Lie group . . . 3198
53.10 Old stuff . . . 3200
53.11 Cosets . . . 3201
53.12 Connected components . . . 3205
53.13 Adjoint group, inner automorphisms . . . 3207
53.14 Compact Lie algebra . . . 3211
53.15 Real Lie algebras . . . 3212
53.15.1 Real and complex vector spaces . . . 3212
53.15.2 Real and complex Lie algebras . . . 3212
53.15.3 Split real form . . . 3213
53.15.4 Compact real form . . . 3213

53.15.5 Involutions . . . 3214
53.15.6 Cartan decomposition . . . 3217
53.16 Root spaces in the real case . . . 3218
53.16.1 Iwasawa decomposition . . . 3219
53.17 The group $SL(2,\mathbb{R})$ and its algebra . . . 3221
53.17.1 Iwasawa decomposition . . . 3221
53.17.2 A companion: $A\bar{N}$ . . . 3222
53.17.3 Killing form . . . 3222
53.17.4 Abstract root space setting . . . 3223
53.17.5 Isomorphism . . . 3223
53.18 The complex algebra $\mathfrak{sl}(2,\mathbb{C})$ and its representations . . . 3223
53.19 The group SO(3) and its Lie algebra . . . 3225
53.19.1 Rotations of functions . . . 3225
53.20 Verma module . . . 3226
53.21 Cyclic modules and representations . . . 3226
53.21.1 Choice of basis . . . 3227
53.21.2 Roots and highest weight vectors . . . 3228
53.21.3 Dominant weight . . . 3228
53.21.4 Verma modules . . . 3229

54 A lot of algebra **3231**
54.1 Algebraic structures . . . 3231
54.1.1 Algebraic structures . . . 3231
54.1.2 Morphisms and such . . . 3231
54.1.3 Group ring . . . 3232
54.2 Tensor products . . . 3232
54.2.1 Tensor product of vector spaces . . . 3232
54.2.2 Some conversions . . . 3233
54.2.3 Contraction of tensors . . . 3233
54.2.4 Tensor product of groups . . . 3234
54.3 Modules over algebras . . . 3234
54.3.1 Endomorphism . . . 3235
54.3.2 Dimension of a module . . . 3236
54.3.3 Tensor product of modules (first) . . . 3237
54.3.4 Tensor product of module (second) . . . 3237
54.3.5 Tensor product of modules (third) . . . 3238
54.3.6 Explicit building of tensor product . . . 3239
54.3.7 Unitary group . . . 3240
54.4 Module over unital ring . . . 3240
54.4.1 Projective module of finite type on unital algebra . . . 3241
54.4.2 Projective module of finite type on unital ring . . . 3241
54.5 Bialgebras and co-properties . . . 3244
54.5.1 Bialgebras . . . 3244
54.5.2 Other co-properties . . . 3246
54.6 Hopf algebras . . . 3247
54.6.1 Convolution product . . . 3247
54.6.2 Opposite algebra . . . 3250
54.6.3 Dual of a Hopf algebra . . . 3251
54.6.4 Involution, $*$-Hopf algebra . . . 3252
54.6.5 Example: universal enveloping algebra . . . 3253
54.6.6 Modules and representations . . . 3255
54.6.7 Matrix elements of modules . . . 3256
54.6.8 Quasitriangular Hopf algebra . . . 3256
54.7 Convolution semigroup . . . 3257

54.8 Modules bialgebras . . . 3258
54.8.1 Module (bi)algebras . . . 3258
54.8.2 Tensor product of representations . . . 3259
54.9 Poisson structure . . . 3259
54.9.1 Homomorphism . . . 3260
54.9.2 Tensor product . . . 3260
54.10 Lie bialgebra . . . 3260
54.11 Poisson-Lie group . . . 3261
54.12 Cyclic cohomology . . . 3263
54.12.1 Example: de Rham homology . . . 3263
54.12.2 Hochschild cohomology . . . 3263
54.12.3 Hochschild groups of cohomology . . . 3265
54.12.4 Homomorphisms . . . 3266
54.12.5 The cup product . . . 3266

55 Examples of groups and representations 3269
55.1 Connectedness of some groups . . . 3269
55.2 Young tableau and diagrams . . . 3269
55.2.1 Representations of $\mathrm{GL}(V)$ . . . 3270
55.2.2 Representations of $O(p,q)$ . . . 3270
55.3 Two words about $\mathfrak{su}(3)$ . . . 3271
55.4 Representations of $\mathfrak{so}(n)$ . . . 3271
55.4.1 Ladder operators . . . 3271
55.4.2 Root spaces and Dynkin diagram for $\mathfrak{so}(2l)$ . . . 3273
55.5 Orthogonal algebra in the odd dimensional case . . . 3274
55.5.1 An odd dimensional example: $\mathfrak{so}(5)$ . . . 3274
55.5.2 Cartan matrix and Dynkin diagram of $\mathfrak{so}(2l+1)$ . . . 3276
55.5.3 Irreducible representations of $\mathfrak{so}(2,l-1)$ . . . 3277
55.5.4 Elementary and basic representations of $\mathfrak{so}(2l+1)$ . . . 3278
55.6 Explicit choices in $\mathfrak{so}(2,l-1)$ . . . 3280
55.6.1 Decompositions and commutators for $\mathcal{Q}$ . . . 3281
55.6.2 Commutators between root spaces and $\mathcal{Q}$ . . . 3281
55.6.3 Commutators in the root spaces . . . 3282
55.6.4 Killing form . . . 3282
55.7 Iwasawa decomposition for $\mathfrak{sl}(2,\mathbb{C})$ . . . 3283
55.8 Symplectic group . . . 3284
55.8.1 Iwasawa decomposition . . . 3284
55.8.2 Isomorphism . . . 3286
55.8.3 Reductive structure on the symplectic group . . . 3287
55.9 Heisenberg group and algebra . . . 3287
55.9.1 The exponential mapping . . . 3288
55.10 Hermitian conjugate, unitary operators . . . 3289
55.11 The group $\mathrm{GL}(V,W)$ of invertible operators . . . 3291
55.12 The group $\mathrm{SU}(n)$ of special unitary operators . . . 3292
55.12.1 Some settings . . . 3293
55.12.2 The center of $\mathrm{SU}(n)$ . . . 3295
55.12.3 Lie group . . . 3295
55.13 Representations of $U(1)$ . . . 3296
55.14 The group $SU(2)$ . . . 3298
55.14.1 SU(2) as compact group . . . 3299
55.14.2 Pauli matrices . . . 3299
55.14.3 Link with SO(3) . . . 3304
55.14.4 The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ . . . 3311
55.14.5 Irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$ . . . 3314

55.14.6 Representations of $\mathfrak{su}(2)$ . . . 3319
55.14.7 Haar measure on SU(2) . . . 3319
55.14.8 Building some representations for SU(2) . . . 3320
55.14.9 Special case: $j = \frac{1}{2}$ . . . 3321
55.14.10 Clebsch-Gordan . . . 3322
55.15 Representations of SO(3) . . . 3322
55.16 Lorentz group . . . 3323
55.16.1 Adjoint map . . . 3324
55.16.2 Structure . . . 3325
55.16.3 Isomorphism $\mathrm{SO}(3,1)^{\uparrow} = \mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$ . . . 3329
55.16.4 Lie group structure . . . 3336
55.16.5 Lie algebra of SO(3, 1) . . . 3336
55.17 Representations of $\mathrm{SL}(2,\mathbb{R})$ and SU(2) . . . 3342
55.18 Representations of $\mathfrak{so}(2, d-1)$ . . . 3343
55.18.1 Verma module . . . 3343
55.18.2 Singular module . . . 3344
55.18.3 The quotient for the scalar singleton . . . 3346

56 Fiber bundle **3349**
56.1 Vector bundle . . . 3349
56.1.1 Transition functions . . . 3350
56.1.2 Inverse construction . . . 3351
56.1.3 Equivalence of vector bundle . . . 3351
56.1.4 Sections of vector bundle . . . 3352
56.1.5 Tensor product . . . 3353
56.2 Tensor bundle . . . 3360
56.2.1 Dual basis . . . 3361
56.3 Differential form . . . 3363
56.3.1 Interior product . . . 3367
56.3.2 Exterior derivative . . . 3368
56.3.3 Pull back . . . 3374
56.4 Vector valued differential forms . . . 3376
56.4.1 A digression: $T_Y\mathcal{G}$ and $\mathcal{G}$ . . . 3376
56.5 Lie algebra valued differential forms . . . 3379
56.6 Principal bundle . . . 3379
56.6.1 Transition functions . . . 3380
56.6.2 Morphisms and such... . . . 3382
56.6.3 Frame bundle: first . . . 3383
56.6.4 Frame bundle: second . . . 3383
56.6.5 Sections of principal bundle . . . 3385
56.6.6 Equivalence of principal bundle . . . 3386
56.6.7 Reduction of the structural group . . . 3387
56.6.8 Density . . . 3388
56.7 Associated bundle . . . 3388
56.7.1 Transition functions . . . 3388
56.7.2 Sections on associated bundle . . . 3389
56.7.3 Associated and vector bundle . . . 3390
56.7.4 Equivariant functions for a vector field . . . 3391
56.7.5 Gauge transformations . . . 3391
56.8 Adjoint bundle . . . 3392
56.9 Connection on vector bundle: local description . . . 3392
56.9.1 Connection and transition functions . . . 3393
56.9.2 Torsion and curvature . . . 3394
56.10 Connexion on vector bundle: algebraic view . . . 3395

56.10.1 Exterior derivative . . . 3396
56.10.2 Divergence, gradient and Laplacian (general) . . . 3397
56.10.3 Divergence, gradient and Laplacian (Riemannian case) . . . 3397
56.10.4 Divergence, gradient and Laplacian (coordinatewise) . . . 3397
56.10.5 Soldering form and torsion . . . 3398
56.10.6 Example: Levi-Civita . . . 3399
56.11 Connection on principal bundle . . . 3400
56.11.1 First definition: 1-form . . . 3400
56.11.2 Horizontal space . . . 3400
56.11.3 Curvature . . . 3402
56.12 Exterior covariant derivative and Bianchi identity . . . 3402
56.13 Covariant derivative on associated bundle . . . 3403
56.13.1 Curvature on associated bundle . . . 3405
56.13.2 Connection on frame bundle . . . 3405
56.13.3 Holonomy . . . 3408
56.13.4 Connection and gauge transformation . . . 3408
56.14 Product of principal bundle . . . 3410
56.14.1 Putting together principal bundle . . . 3410
56.14.2 Connections . . . 3411
56.14.3 Representations . . . 3411
56.15 Connections . . . 3412
56.15.1 Gauge potentials . . . 3412
56.15.2 Covariant derivative . . . 3413
56.15.3 Gauge transformation . . . 3413
56.15.4 Curvature . . . 3415
56.16 Hodge decomposition theorem and harmonic forms . . . 3418

57 Integration on manifolds 3421
57.1 Orientation and volume forme . . . 3421
57.1.1 Orientation . . . 3421
57.1.2 Connected atlas . . . 3422
57.1.3 Volume form . . . 3424
57.2 Integral of differential forms . . . 3427
57.2.1 Integration over a chart . . . 3428
57.2.2 Integration over a compact part . . . 3428
57.2.3 Non compact . . . 3430
57.2.4 Non compact (another try) . . . 3431
57.3 Manifold with boundary . . . 3435
57.3.1 Some facts about $\mathbb{R}^n_+$ . . . 3435
57.3.2 Manifold with boundary . . . 3437
57.3.3 Tangent space . . . 3441
57.4 Stokes theorem . . . 3442
57.5 Integration of a differential form . . . 3445
57.5.1 Open set in $\mathbb{R}^n$ . . . 3445
57.5.2 One chart on a manifold . . . 3446
57.5.3 On manifold that require a finite atlas . . . 3447
57.5.4 Integrating by part . . . 3449

58 Lie groups of transformations 3451
58.1 Groups of transformations . . . 3451
58.2 Lie groups of transformations . . . 3451
58.2.1 Action of the Lie algebra . . . 3453
58.3 Cosets and homogeneous spaces . . . 3454
58.4 Isotropy group . . . 3454

59 Classical mechanics **3457**
59.1 Some symplectic and Poisson geometry . . . 3457
59.1.1 Symplectic manifold . . . 3457
59.1.2 Poisson manifold . . . 3457
59.1.3 Hamiltonian action . . . 3458
59.1.4 Coadjoint orbits . . . 3459
59.1.5 Example of adjoint and coadjoint orbit . . . 3459
59.1.6 A second example . . . 3460
59.1.7 Kostant theorem . . . 3461
59.1.8 Central extension . . . 3462
59.2 Constrained system . . . 3462
59.2.1 Holonomic constraints . . . 3462
59.2.2 Primary constraints . . . 3463
59.2.3 Passage to Hamiltonian formalism . . . 3464
59.2.4 Secondary constraints . . . 3465
59.3 Lie groupoids and algebroids . . . 3465
59.3.1 Example: tangent bundle TM . . . 3465
59.3.2 Example: a Lie algebra . . . 3465
59.3.3 Example: gauge algebroid . . . 3466
59.3.4 Poisson structure . . . 3467
59.4 Lagrangian formalism . . . 3468
59.5 Groupoids . . . 3468
59.5.1 Example: when $\Gamma_0 = \{e\}$ . . . 3470
59.5.2 Example: the null groupoid . . . 3470
59.5.3 The case $\alpha = \beta$ . . . 3470
59.5.4 An example on a vector bundle . . . 3471
59.5.5 Orbits . . . 3471
59.5.6 Morphism . . . 3471
59.6 Lie groupoid . . . 3471

60 Hilbert spaces **3473**
60.1 Basis and orthonormal systems . . . 3473
60.2 Operators on Hilbert spaces . . . 3475
60.2.1 Adjoint, unitary and projection operator . . . 3475
60.2.2 Topology on space of continuous endomorphism . . . 3476
60.2.3 Compact operators . . . 3476
60.2.4 Hilbert-Schmidt operators . . . 3479
60.2.5 The Schatten-von Neumann ideal . . . 3480
60.2.6 Normal operators on Hilbert space . . . 3481
60.3 Spectral theory on Banach algebras . . . 3482
60.4 Spectral theorem and some consequences . . . 3483
60.4.1 Spectrum . . . 3483
60.4.2 Note about operator algebras . . . 3484
60.4.3 Spectrum: next steps . . . 3484
60.5 Operators with compact resolvent . . . 3491
60.6 Strong, weak and other topologies . . . 3492
60.6.1 Ultraweak topology . . . 3492
60.7 Polar cone . . . 3494
60.8 Representations of Lie groups on Hilbert spaces . . . 3494
60.8.1 Induced representation . . . 3495
60.8.2 Unitary induced representation . . . 3497

61 Sobolev spaces **3499**
61.1 Distributions . . . 3499

61.1.1 Dirac distribution . . . 3500
61.1.2 Distribution defined from functions . . . 3500
61.1.3 Fourier transform and Schwartz functions . . . 3501
61.1.4 Support of a distribution . . . 3501
61.1.5 Duality . . . 3502
61.2 Measure, distribution and integral . . . 3503
61.2.1 Example: the Lebesgue measure . . . 3504
61.2.2 Integration on more general spaces . . . 3505
61.2.3 Integration of vector valued functions . . . 3506
61.2.4 Weak integral . . . 3506
61.3 Distribution on groups . . . 3506
61.3.1 Convolution product . . . 3507
61.3.2 Representations . . . 3508
61.3.3 Representation on a Hilbert space . . . 3508
61.3.4 Slightly more general . . . 3509
61.3.5 Representation of $M_0(G)$ . . . 3509
61.4 A boundary aware Sobolev space . . . 3510
61.5 Smooth diffeomorphisms . . . 3512
61.6 Restriction operator . . . 3517
61.7 Boundaries . . . 3521
61.8 Older work . . . 3523

62 Analysis **3527**
62.1 Fréchet spaces and algebras . . . 3527
62.2 Gårding spaces . . . 3528
62.2.1 Strongly continuous action and smooth vector . . . 3528
62.2.2 Action on Hilbert space . . . 3530
62.3 Operator symbol . . . 3534
62.3.1 A case without problem . . . 3534
62.3.2 A problem . . . 3534
62.3.3 Basic definitions . . . 3534
62.3.4 Topology on $\mathcal{S}^m(X \times \mathbb{R}^N)$ . . . 3537
62.3.5 Asymptotic expansions . . . 3541
62.3.6 Tempered functions . . . 3542
62.4 First construction of oscillatory integrals . . . 3543
62.4.1 Construction . . . 3543
62.4.2 Parametric oscillatory integral . . . 3546
62.5 Other constructions of oscillatory integral . . . 3548
62.5.1 Derivation definition . . . 3548
62.5.2 Lattice way to define oscillatory integrals . . . 3550
62.6 Decomposition into direct sum . . . 3551
62.6.1 Integral over $V \times V$ . . . 3553
62.6.2 Normal operator . . . 3554
62.7 Pseudo-differentiable operators . . . 3555
62.7.1 Differential operator . . . 3555
62.7.2 Laplace operator . . . 3555
62.8 Invariant differential operators on Lie groups . . . 3556
62.9 Pseudo differential operators . . . 3559
62.9.1 Composition . . . 3559
62.9.2 Fourier transform and pseudo-differential operator . . . 3559
62.9.3 Example . . . 3560
62.9.4 Trace operators . . . 3560
62.9.5 Asymptotic expansions . . . 3560
62.10 Dirac and Laplace type operators . . . 3561

62.11 Wodzicki residue . . . 3561
62.12 Interpolation theory . . . 3562
62.13 Trace class operators . . . 3563
62.13.1 Trace . . . 3563
62.14 Dixmier traces . . . 3564
62.14.1 Banach limit . . . 3564
62.14.2 Infinitesimal operator . . . 3565
62.14.3 Dixmier trace . . . 3565
62.14.4 Dixmier: second . . . 3567
62.14.5 Noncommutative integral . . . 3568
62.14.6 Residues . . . 3568

63 Partial derivative equation **3569**
63.1 Basic definitions . . . 3569
63.2 Principal symbol . . . 3570
63.2.1 Classification with respect to the boundary conditions . . . 3572
63.2.2 Characteristic speed . . . 3574
63.2.3 Decoupling a system . . . 3574
63.3 Some examples . . . 3575
63.4 Principle of superposition . . . 3575
63.4.1 Separation of variables . . . 3577

64 Finite elements **3581**
64.1 Lax-Milgram theorem . . . 3581
64.2 Variational formulation (not too rigorous) . . . 3581
64.3 Galerkin's approximation . . . 3583
64.4 Gradient on a boundary . . . 3586
64.5 Locally integrable functions . . . 3587
64.6 An approximation result . . . 3587
64.7 Lax-Milgram with a boundary condition . . . 3588

65 Chain complexes **3591**
65.1 Chain complexes . . . 3591
65.1.1 Homotopy groups . . . 3591
65.1.2 Exact sequences . . . 3592
65.1.3 Euler-Poincaré . . . 3593
65.2 Cochain complexes . . . 3593
65.2.1 Singular chains . . . 3594
65.3 Singular homology . . . 3594

66 Homogeneous and symmetric spaces **3597**
66.1 Action of groups on sets . . . 3597
66.1.1 Fundamental and invariant fields . . . 3597
66.2 Rough introduction to homogeneous spaces . . . 3598
66.2.1 Killing induced product . . . 3599
66.2.2 Homogeneous space . . . 3600
66.2.3 Invariant metric on homogeneous space (first) . . . 3601
66.2.4 Homogeneous metric on homogeneous spaces . . . 3604
66.3 Symmetric spaces . . . 3606
66.3.1 Basic facts . . . 3606
66.3.2 Choice of a Cartan involution . . . 3608
66.3.3 Affine Symmetric spaces . . . 3608
66.3.4 Symmetric pair . . . 3610
66.3.5 Symmetric spaces as quotient . . . 3612

66.3.6 Symmetric Lie algebras . . . 3613
66.3.7 Connection on symmetric spaces . . . 3615
66.3.8 Canonical connection and covariant derivative . . . 3616
66.4 The group SO(2, 1) and its algebra . . . 3617
66.4.1 Root spaces for $\mathfrak{so}(2,1)$ . . . 3618
66.4.2 Position of SO(1, n) inside SO(2, n) . . . 3618
66.4.3 Iwasawa decomposition for SO(1, n) . . . 3620
66.5 Iwasawa decomposition for SO(2, n) . . . 3621
66.5.1 Cartan decomposition and compatible reductive decomposition . . . 3621
66.5.2 Maximal abelian subalgebra . . . 3621
66.5.3 Compatibility conditions between $\mathcal{H}$ and $\mathcal{G}$ . . . 3622
66.5.4 First choice: AN . . . 3625
66.5.5 A companion: $A\bar{N}$ . . . 3627
66.5.6 Second choice . . . 3627
66.5.7 Application: an isomorphism of Lie algebra . . . 3631
66.6 Openness of orbits in homogeneous spaces . . . 3631
66.7 Symplectic symmetric spaces . . . 3633
66.7.1 Example . . . 3633
66.7.2 Algebraic setting . . . 3633
66.7.3 Symplectic triple . . . 3634
66.7.4 Example on the Heisenberg algebra . . . 3635
66.7.5 Realization as coadjoint orbit . . . 3636
66.8 Hermitian and symplectic spaces . . . 3640
66.8.1 The Chevalley cohomology . . . 3643
66.8.2 Involutive symmetric Lie algebras . . . 3647
66.8.3 Symplectic symmetric spaces and involutive Lie algebra . . . 3648
66.8.4 Symmetric spaces and coadjoint orbits . . . 3649
66.8.5 Elementary solvable symmetric spaces . . . 3650
66.8.6 Mid-point map . . . 3650
66.8.7 Kähler structures . . . 3650
66.8.8 Symplectic structure on the Iwasawa component . . . 3651
66.8.9 Iwasawa coordinates . . . 3652
66.8.10 Summary of the construction . . . 3654
66.8.11 Continuation . . . 3654
66.9 Elementary normal symplectic spaces . . . 3656
66.10 Pyatetskii-Shapiro structure theorem . . . 3657

67 Heat kernel expansions **3661**
67.1 Mellin transform . . . 3661
67.2 General setting . . . 3661
67.2.1 Mellin . . . 3662
67.2.2 Residues and zeta function . . . 3662
67.2.3 Boundary conditions for heat kernel expansions . . . 3663
67.2.4 Boundary and Dirac . . . 3663

68 From Clifford algebras to Dirac operator **3665**
68.1 Invitation: Clifford algebra in quantum field theory . . . 3665
68.1.1 Schrödinger, Klein-Gordon and Dirac . . . 3665
68.1.2 First representation . . . 3666
68.1.3 Some consequences of the universal property . . . 3667
68.1.4 Trace . . . 3667
68.2 Spinor representation . . . 3669
68.2.1 Explicit representation . . . 3672
68.2.2 A remark . . . 3673

68.2.3 General two dimensional Clifford algebra . . . 3673
68.3 Spin group . . . 3673
68.3.1 Studying the group structure . . . 3675
68.3.2 Redefinition of Spin(V) . . . 3680
68.3.3 A few about Lie algebra . . . 3681
68.3.4 Grading ΛW . . . 3682
68.3.5 Clifford algebra for $V = \mathbb{R}^2$ . . . 3682
68.4 Clifford modules . . . 3685
68.5 Spin structure . . . 3687
68.5.1 Example: spin structure on the sphere S^2 . . . 3687
68.5.2 Spinor bundle . . . 3688
68.6 Dirac operator . . . 3690
68.6.1 Preliminary definition . . . 3690
68.6.2 Definition of Dirac . . . 3692
68.7 Dirac operator on $\mathbb{R}^2$ . . . 3694
68.7.1 Connection on SO($\mathbb{R}^2$) . . . 3695
68.7.2 Construction of γ . . . 3696
68.7.3 Covariant derivative on $\Gamma(\mathcal{S})$ . . . 3697
68.7.4 Dirac operator on the euclidian $\mathbb{R}^2$ . . . 3697
68.7.5 Dirac operator as elliptic pseudo-differential operator . . . 3698
68.8 Clifford algebras and Morita equivalence . . . 3698
68.8.1 Example: quantum field theory . . . 3699
68.8.2 An other definition of the Dirac operator . . . 3699

69 Relativistic fields and group theory **3701**
69.1 Mathematical framework of field theory . . . 3701
69.1.1 Axioms of the (quantum) relativistic field theory . . . 3701
69.1.2 Symmetries and Wigner's theorem . . . 3701
69.1.3 Projective representations . . . 3706
69.1.4 Representations and power expansions . . . 3707
69.2 The symmetry group of nature . . . 3708
69.2.1 Spin and double covering . . . 3708
69.2.2 How to implement the Poincaré group . . . 3710
69.2.3 Momentum operator . . . 3710
69.2.4 Pure Lorentz transformation . . . 3711
69.2.5 Rebuilding of a basis for $\mathscr{H}$ . . . 3712
69.2.6 Little group . . . 3712
69.2.7 Positive mass . . . 3714
69.2.8 Null mass . . . 3714

70 Relativistic fields and fiber bundle formalism **3715**
70.1 Example: electromagnetism . . . 3715
70.2 Connections for the gauge invariance . . . 3716
70.2.1 Little more general, slightly more formal . . . 3717
70.2.2 A "final" formalism . . . 3717
70.2.3 The electromagnetic field F . . . 3718
70.3 Spin manifold for Lorentz invariance . . . 3719
70.4 Interactions . . . 3720
70.4.1 Spin zero . . . 3720
70.4.2 Non zero spin formalism . . . 3722
70.5 Singletons as sections of bundle . . . 3723

71 Conformal fields theory **3725**
71.1 The conformal group . . . 3725

71.1.1 Preliminary discussion . . . 3725
71.1.2 Generators . . . 3726
71.1.3 Exponentiation . . . 3730
71.1.4 What about really preserving the angles? . . . 3733
71.2 Two dimensional conformal group . . . 3733

72 Banach and C^*-algebras **3737**
72.1 Commutative Banach algebra . . . 3737
72.1.1 Structure space . . . 3737
72.1.2 Topology on $\Delta(\mathfrak{A})$ . . . 3739
72.1.3 An example . . . 3740
72.2 Commutative C^*-algebras . . . 3742
72.3 Functional calculus in unital C^*-algebras . . . 3745
72.3.1 The isomorphism $C^*(A, \mathbb{1}) \leftrightarrow C(\sigma(A))$ . . . 3749
72.4 Positivity . . . 3750
72.5 Ordering relation . . . 3753
72.6 Approximate unit . . . 3756
72.7 Ideal in Banach algebras . . . 3760
72.8 States . . . 3761
72.9 States on unital C^*-algebras . . . 3761
72.9.1 States on non unital C^*-algebras . . . 3764
72.10 Uniqueness of the norm . . . 3768
72.10.1 Convexity . . . 3768
72.11 Representation . . . 3769
72.11.1 Representation of involutive algebra . . . 3769
72.11.2 Cyclic representations of C^*-algebra . . . 3772
72.11.3 Primitive spectrum . . . 3773
72.11.4 GNS construction . . . 3774
72.11.5 Universal representation . . . 3776
72.12 Spaces of matrices . . . 3777
72.13 Stinespring theorem . . . 3778
72.14 Representations . . . 3782
72.15 Pure states . . . 3783
72.15.1 Existence of pure states . . . 3789
72.16 Generalization of matrix C^*-algebra . . . 3790
72.16.1 Example . . . 3791
72.17 Tensor product . . . 3791
72.17.1 Examples: Hopf algebra of functions . . . 3792
72.18 Traces over C^*-algebra . . . 3792

73 Compact quantum groups **3795**
73.1 Definitions . . . 3795
73.1.1 Example: representation of groups . . . 3795
73.1.2 Matrix quantum group . . . 3796
73.1.3 Haar measure on compact quantum groups . . . 3797
73.2 Construction of $\mathrm{SU}_q(n)$ . . . 3799
73.2.1 Determinant . . . 3800
73.2.2 Norm . . . 3802
73.3 Representations of $\mathrm{SU}_q(3)$ . . . 3803
73.4 Quantized universal algebras . . . 3810
73.4.1 Example with $\mathfrak{sl}(2, \mathbb{C})$ . . . 3813
73.5 Quantum universal enveloping algebra . . . 3813
73.5.1 Admissible modules . . . 3814
73.5.2 Example on $U_q\,\mathfrak{sl}(2, \mathbb{C})$ . . . 3815

73.6 Quantized function algebra . . . 3820

74 von Neumann algebras **3823**
74.1 Functional, representation and automorphism . . . 3823
74.2 Commutant . . . 3823
74.3 Examples of von Neumann algebras . . . 3825
74.3.1 Algebra $L^\infty(X)$ . . . 3825
74.3.2 Countable direct sum of Hilbert spaces . . . 3826
74.3.3 Direct limit . . . 3826
74.4 Continuous dimensions . . . 3828
74.5 Cantor . . . 3829
74.6 More general state . . . 3829
74.7 Group measure space construction . . . 3829
74.7.1 First attempt . . . 3830
74.7.2 Second attempt . . . 3831
74.7.3 First generalisation . . . 3832
74.7.4 Second generalisation . . . 3833
74.7.5 One particular case . . . 3833
74.8 More about projections . . . 3835
74.8.1 Comparison of projections . . . 3837
74.9 Type I factor and factorization . . . 3838
74.9.1 Tensor product of von Neumann algebras . . . 3842
74.10 Dimensions . . . 3842
74.10.1 Finite and infinite projections . . . 3842
74.10.2 Finite factor . . . 3846
74.10.3 Rational and real dimensions . . . 3847
74.10.4 Summary . . . 3847
74.10.5 Purely infinite factor . . . 3848
74.11 Tracial functional . . . 3849
74.12 Modules over von Neumann algebras . . . 3852
74.12.1 Modular conjugation and factor of type II_1 . . . 3852
74.12.2 Dimension . . . 3856
74.13 Position of submodules . . . 3858
74.13.1 Summary . . . 3862
74.13.2 Properties of the dimension function . . . 3862
74.14 Decomposition of operators and representations . . . 3864
74.14.1 Motivation . . . 3864
74.14.2 Conventions and definitions . . . 3864
74.14.3 Examples . . . 3865
74.14.4 Decompositions of operators . . . 3866
74.14.5 Diagonal and decomposable operators . . . 3868
74.14.6 Decompositions of representations . . . 3872
74.15 Index theory . . . 3874
74.15.1 Introduction . . . 3874
74.15.2 Example . . . 3875
74.15.3 Isomorphisms of abstract von Neumann algebras . . . 3876
74.15.4 Index of finite subfactors . . . 3877
74.15.5 Subfactors . . . 3879
74.15.6 Example of index bigger than 4 . . . 3883
74.15.7 Properties of the sequence of M_i, P_i . . . 3883
74.15.8 Some interesting functions . . . 3885
74.16 Modular theory . . . 3886
74.16.1 Modular isometry . . . 3886
74.16.2 Example . . . 3888

74.16.3 Other example . . . 3888
74.16.4 Still an other example . . . 3890
74.16.5 Tomita's theorem . . . 3890
74.16.6 Modular group . . . 3892
74.16.7 Standard positive cone . . . 3893

75 Dirichlet forms **3895**
75.1 KMS states and CAR algebras . . . 3895
75.1.1 KMS states . . . 3895
75.1.2 CAR algebras . . . 3895
75.1.3 KMS symmetric maps . . . 3896
75.2 Dirichlet and traces . . . 3898
75.3 Modules over C^*-algebra . . . 3901
75.3.1 Example: gradient of a function . . . 3901
75.3.2 Left and right representation . . . 3901

76 K-theory **3905**
76.1 The group K_0 . . . 3905
76.1.1 Example . . . 3907
76.1.2 Universality . . . 3907
76.2 Vector bundle and K-theory . . . 3908
76.2.1 Reduced group . . . 3910
76.2.2 Functorial description . . . 3910
76.2.3 External product . . . 3910
76.2.4 Clutching function . . . 3911

77 BF theory **3913**
77.1 BF theory and relativity . . . 3913
77.2 Palatini formalism . . . 3914
77.2.1 Other point of view . . . 3915

78 BTZ black hole from identifications **3917**
78.1 Low dimensional anti de Sitter spaces . . . 3917
78.1.1 Two dimensional anti de Sitter . . . 3917
78.1.2 Three dimensional anti de Sitter . . . 3917
78.2 Simple example on AdS_2 . . . 3918
78.3 Still AdS_2 . . . 3919
78.4 Causal structure of AdS_2 . . . 3920
78.4.1 Second point of view . . . 3922
78.4.2 Light cone . . . 3922
78.4.3 The graphical way . . . 3923
78.4.4 The black hole . . . 3924
78.4.5 Closed orbits . . . 3924
78.4.6 Time orientation . . . 3924
78.4.7 Singularity and physical space . . . 3925
78.4.8 Light cone . . . 3926
78.5 Causal structure of AdS_3 . . . 3927
78.5.1 Isometries of AdS_3 . . . 3927
78.5.2 Orbits in AdS_3 . . . 3927
78.5.3 The metric . . . 3929
78.6 Identifications in AdS_3 . . . 3930
78.6.1 Global description of the black hole . . . 3932
78.6.2 Derivation of the horizons . . . 3933
78.6.3 General considerations . . . 3936

78.6.4 The metric . . . 3937

79 BTZ black holes in anti de Sitter spaces 3939
79.1 Introduction . . . 3939
79.2 Introduction . . . 3941
79.2.1 Physics and mathematics of black holes . . . 3941
79.2.2 BTZ black hole . . . 3943
79.2.3 Generalization and group setting . . . 3943
79.2.4 Some notations . . . 3944
79.3 Visite guidée . . . 3944
79.3.1 En termes de BTZ . . . 3944
79.3.2 En termes de généralisations . . . 3945
79.4 Connectedness of groups and anti de Sitter spaces . . . 3946
79.4.1 The quotient for anti de Sitter . . . 3946
79.5 Symmetric space structure on anti de Sitter . . . 3947
79.5.1 Anti de Sitter as symmetric space . . . 3948
79.6 Open and closed orbits . . . 3948
79.6.1 Open orbits in anti de Sitter spaces . . . 3948
79.7 Some complements . . . 3950
79.7.1 A first brute force computation . . . 3950
79.7.2 Search for $Z(K)$ . . . 3951
79.7.3 The same with the "old" Iwasawa decomposition . . . 3951
79.7.4 A non-open orbit and a precision . . . 3952
79.7.5 The same with the "old" Iwasawa decomposition . . . 3953
79.7.6 Orbits as homogeneous spaces . . . 3954
79.8 Causality, light cone and related topics on anti de Sitter . . . 3955
79.8.1 Time orientation . . . 3959
79.8.2 Action of H and $\mathrm{Ad}(\mathcal{Q})$ . . . 3959
79.8.3 Two other characterizations of the singularity . . . 3960
79.8.4 A criterion with the tangent spaces . . . 3962
79.8.5 Orbits and topology . . . 3963
79.8.6 The volume form method . . . 3964
79.9 Existence of a non trivial horizon . . . 3969
79.10 Characterization by angles in $SO(l-1)$ . . . 3971
79.10.1 Another (useless) characterisation . . . 3972
79.11 Characterisation as orbit of group (by the equation) . . . 3972
79.11.1 The old three dimensional case . . . 3972
79.11.2 Characterization by induction on the dimension . . . 3973
79.12 Organization of the next few pages . . . 3977
79.13 Some old results . . . 3978
79.13.1 Horizon in the three dimensional case . . . 3978
79.13.2 Topology and horizon . . . 3978
79.14 The horizon of the BTZ black hole . . . 3979
79.15 Conclusion . . . 3986
79.16 The algebras without matrices . . . 3986
79.16.1 The structure theorem by Pyatetskii-Shapiro . . . 3986
79.17 Characterisation of the horizon (vanishing norm) . . . 3987
79.18 Conclusions and perspectives . . . 3987
79.19 BTZ from the structure of $\mathfrak{so}(2,n)$ . . . 3987
79.20 Introduction . . . 3987
79.20.1 Anti de Sitter space and the BTZ black hole . . . 3987
79.20.2 The way we describe the BTZ black hole . . . 3988
79.20.3 Organization of the paper . . . 3989
79.21 Structure of the algebra . . . 3990

79.21.1 The Iwasawa component . . . 3990
79.21.2 The compact part . . . 3991
79.21.3 Dimensional slices . . . 3992
79.21.4 Reductive decomposition . . . 3993
79.21.5 Properties of the reductive decompositions . . . 3995
79.21.6 A convenient basis for the root spaces and computations . . . 4005
79.21.7 Properties of the basis . . . 4012
79.21.8 Some exponentials . . . 4013
79.21.9 Classification of the basis by the spaces . . . 4014
79.22 The causally singular structure . . . 4015
79.22.1 Closed orbits . . . 4015
79.22.2 Vanishing norm criterion . . . 4018
79.22.3 Existence of the black hole . . . 4021
79.23 Some more computations . . . 4023
79.24 Description of the horizon . . . 4027
79.24.1 Induction on the dimension . . . 4027
79.24.2 Examples of surjective groups . . . 4030
79.24.3 Backward induction . . . 4031
79.25 Dirac operator on AdS_2 . . . 4032
79.25.1 Clifford algebra and spin group . . . 4032
79.25.2 Relation between $SU(1,1)$ and $\mathrm{SO}(2,1)$ . . . 4033
79.25.3 Spin structure on AdS_2 . . . 4034
79.25.4 Spinor bundle and connection . . . 4035
79.25.5 Clifford algebra $(1,1)$ . . . 4036
79.25.6 Parallel transport . . . 4038
79.25.7 Covariant derivative . . . 4039
79.25.8 Another way to write a section (wrong way to do) . . . 4041
79.25.9 Once again . . . 4041
79.26 Dirac operator on AdS_3 . . . 4045
79.26.1 Spin structure on AdS_3 . . . 4046
79.26.2 Connection on the spinor bundle . . . 4048
79.26.3 Horizontal lift . . . 4049
79.26.4 Spin structure on AdS_3 . . . 4050
79.26.5 Left invariance of Dirac . . . 4054
79.27 Spin structure and Dirac operator on AdS_l . . . 4056
79.27.1 Frame bundle and spin structure . . . 4056
79.27.2 Connection . . . 4058
79.27.3 Dirac operator . . . 4059
79.27.4 Frame bundle . . . 4060
79.27.5 Spin structure . . . 4061
79.27.6 Reduction of the structural group . . . 4062
79.28 Conclusion . . . 4062

80 General non commutative geometry **4065**
80.1 Non commutative differential forms . . . 4065
80.1.1 Universal differential forms . . . 4065
80.1.2 Connes differential forms . . . 4069
80.1.3 Example: two points space . . . 4074
80.1.4 Example: manifold . . . 4076
80.2 Fredholm modules . . . 4076
80.2.1 Introductory example . . . 4076
80.2.2 Definition . . . 4077
80.2.3 Cycle associated with Fredholm module . . . 4077
80.3 Hochschild cohomology . . . 4079

80.4 Fredholm module and conformal structure . . . 4079
80.5 Fredholm modules and K-cycles . . . 4080
80.6 Spectral triple . . . 4080
80.6.1 General spectral triple . . . 4080
80.6.2 Commutative real triple . . . 4081
80.6.3 Analysis on a spectral triple . . . 4082
80.6.4 Spectral triple over a manifold . . . 4082
80.6.5 Distance over general triple . . . 4084
80.6.6 Real triple . . . 4084
80.6.7 Example: compact manifold . . . 4085
80.7 Non commutative vector bundle . . . 4085
80.7.1 The category of complex vector bundles . . . 4085
80.8 Hermitian structure and compatible connection . . . 4086
80.8.1 Hermitian structures . . . 4086
80.8.2 Morita equivalence . . . 4089
80.8.3 Differential over a module . . . 4090
80.8.4 Connections over Hermitian modules . . . 4090
80.8.5 Universal compatible connection . . . 4091
80.8.6 Curvature . . . 4093
80.8.7 Transformation law under the unitary group . . . 4093
80.8.8 Connection on bimodule . . . 4094
80.8.9 Connection on dual module . . . 4094
80.8.10 Connections and Connes' calculus . . . 4094
80.8.11 Scalar product on $\pi(\Omega^p\mathfrak{A})$ and Hilbert structure . . . 4095
80.9 Connections on modules and Yang-Mills functional . . . 4097
80.9.1 Potential vector and field strength . . . 4097
80.9.2 Action of the gauge group . . . 4097
80.9.3 Yang-Mills . . . 4098
80.9.4 Fermionic action . . . 4099
80.10 Forms and physics formalism . . . 4101
80.10.1 Unitary group . . . 4101
80.11 Yang-Mills on the two point space . . . 4102
80.11.1 The simplest module . . . 4103

81 Deformations: formal aspects **4105**
81.1 Twists of module (co)algebras . . . 4105
81.1.1 Twisting the module algebras . . . 4105
81.1.2 Twisting the bialgebra itself . . . 4106
81.1.3 Twisting the left module algebra . . . 4107
81.2 Twist and star product . . . 4107
81.3 Formal Extension lemma . . . 4110

82 Deformations: non-formal aspects **4117**
82.1 Introduction to Moyal star product . . . 4117
82.1.1 General definition of a star product . . . 4117
82.1.2 Definition of the Moyal star product . . . 4117
82.2 Rieffel's deformation by action of $\mathbb{R}^d$ . . . 4119
82.3 Weyl product . . . 4119
82.4 The twisting map . . . 4122
82.5 Wigner function . . . 4122
82.6 Deformation by action of group . . . 4123
82.7 A first analysis question . . . 4125
82.7.1 A simple toy model . . . 4125
82.7.2 The case of interest . . . 4126

82.7.3 Smooth vectors of the action . . . 4126
82.7.4 Oscillatory definition on $SL(2, R)$ . . . 4127
82.8 Method improvement . . . 4129
82.8.1 Change of point of view for the integral . . . 4129
82.8.2 Adapted vector field and by part integration . . . 4129
82.8.3 New way to express the problem . . . 4130
82.8.4 Topology choice . . . 4131
82.8.5 Convergence condition . . . 4132
82.8.6 Some inequalities in $C^\infty(R)$ . . . 4133
82.8.7 Compactly supported function . . . 4134
82.9 The construction . . . 4135
82.9.1 The problem with Rieffel . . . 4137
82.10 Klein-Gordon calculus . . . 4138
82.10.1 Affine space primer . . . 4138
82.10.2 Space-time structure . . . 4138
82.10.3 Observer . . . 4139
82.10.4 Action of $\mathscr{P}_0$ on Ω . . . 4141
82.10.5 The spacial symmetry P_ω . . . 4142
82.10.6 Momentum . . . 4143
82.10.7 Invariant measure . . . 4145
82.10.8 Sobolev setting, Bargmann and Wigner representation . . . 4145
82.10.9 Back to operator P_ω . . . 4147
82.10.10 Passive symbol . . . 4148
82.10.11 The spacial bundle . . . 4149
82.10.12 Measures on spacial bundle . . . 4150
82.10.13 Active symbol . . . 4151
82.10.14 The operator ∇^λ . . . 4152

83 WKB quantization **4153**
83.1 WKB quantization . . . 4153
83.1.1 Definitions and general setting . . . 4154
83.1.2 Deformation of Iwasawa subgroups . . . 4157
83.2 Deformation of $SL(2, \mathbb{R})$ . . . 4160
83.2.1 Actions and Symplectic structure . . . 4160
83.2.2 Guessing the star product . . . 4165
83.2.3 Formula for the product . . . 4168
83.3 Non formal Extension lemma . . . 4171
83.4 Deformation of $SU(1, n)$ . . . 4173
83.5 One dimensional split extensions of Heisenberg algebras . . . 4174
83.5.1 Introduction . . . 4174
83.5.2 General definitions . . . 4174
83.5.3 Symplectic structure . . . 4175
83.5.4 Isomorphisms . . . 4177
83.5.5 Extensions with non trivial $\overline{X}$ . . . 4177
83.5.6 Get the μ back . . . 4181
83.5.7 Jump from one kernel to another . . . 4181

84 Deformation of anti de Sitter spaces **4183**
84.1 Group structure on the open orbits . . . 4183
84.1.1 Global structure . . . 4183
84.1.2 Alternative more intrinsic proofs . . . 4189
84.1.3 Local group structure . . . 4192
84.2 Deformation of AdS_4 . . . 4196
84.2.1 Isomorphism with $\mathcal{R}$ . . . 4200

84.2.2 The symplectic issue . . . 4202
84.3 Isospectral deformations of M . . . 4203
84.3.1 Complement: importance of left invariance . . . 4204
84.4 Perspectives . . . 4204
84.5 Spin structure on the black hole . . . 4205
84.5.1 Frame bundle over R_1 . . . 4205
84.5.2 Frame bundle over $\mathcal{U}$ . . . 4205

85 Two notes for further developments **4207**
85.1 Formula of Unterberger on $\mathrm{SL}(2,\mathbb{R})$ . . . 4207
85.1.1 Action on the dual of its Lie algebra . . . 4210
85.1.2 First action on the two dimensional anti de Sitter space . . . 4210
85.1.3 Second action on the two dimensional anti de Sitter space . . . 4211
85.2 Deformation of $\mathrm{SO}(2,n)$ . . . 4212
85.2.1 Applying the extension lemma with old Iwasawa . . . 4212
85.2.2 Decomposition as split extension . . . 4215
85.2.3 Conclusion and perspectives . . . 4216

86 Gravitation and noncommutative geometry **4217**
86.1 Inner and outer automorphisms . . . 4217
86.2 Spectral action . . . 4220
86.3 Gauge theory . . . 4220
86.4 Scalar curvature . . . 4221
86.5 Boundary conditions and heat kernel expansion . . . 4221
86.5.1 Example . . . 4221
86.6 Boundary . . . 4224

87 Levy Processes and such **4225**
87.1 Lévy process . . . 4225
87.2 Quantum probability space and stochastic processes . . . 4225
87.3 Lévy process . . . 4226
87.4 Schürmann triple . . . 4227
87.5 A representation . . . 4230
87.5.1 Ideals . . . 4231
87.5.2 Decomposition . . . 4233
87.6 Gaussian process . . . 4238
87.6.1 Poisson process . . . 4239
87.6.2 Gaussian cocycle . . . 4239
87.6.3 Gaussian generator . . . 4242

88 Complements **4247**
88.1 Alternative formalism for the quantum mechanics . . . 4247
88.2 Description of Maxima computations . . . 4248
88.2.1 Projections and decompositions . . . 4248
88.2.2 Symplectic computations . . . 4249
88.2.3 $\mathrm{Ad}(g)^*(\delta\xi^*) = E^*$ . . . 4250
88.3 Some of my favourite mistakes . . . 4250
88.4 Frequently asked questions . . . 4251
88.4.1 Are open orbits simply connected? . . . 4251
88.4.2 Why not de Sitter? . . . 4251

III Matlab **4253**

89 Matlab **4255**

89.1 Exercices des séances . . . 4255
89.1.1 Bases et calcul matriciel . . . 4255
89.1.2 Vecteurs à éléments équidistants . . . 4259
89.1.3 Polynômes et approximations au sens des moindres carrés . . . 4262
89.1.4 Intégration numérique et résolution d'équations différentielles . . . 4264
89.1.5 Exercices variés . . . 4267
89.2 Anciens tests et examens . . . 4270
89.3 Conseils généraux . . . 4290
89.3.1 Écriture d'une fonction . . . 4290
89.3.2 Conception des fonctions . . . 4290
89.3.3 Autres . . . 4290

IV Exercices 4291

90 Mathématique générale pour le starter SVT (Besançon) 4293
90.1 Théorie . . . 4293
90.1.1 Suites . . . 4293
90.1.2 Techniques pour majorer et minorer . . . 4294
90.2 Rappels : exponentielles et logarithmes . . . 4294
90.3 Exponentielles et logarithmes . . . 4295
90.4 Fonctions et graphes . . . 4296
90.5 Limites du côté de l'infini . . . 4297
90.6 Limite de suites . . . 4299
90.7 Étude de fonctions, première partie . . . 4317
90.8 Étude de fonctions, suite . . . 4320
90.9 Intégration . . . 4325
90.10 Équations différentielles . . . 4337
90.11 Révisions . . . 4350
90.12 Anciennes interrogations . . . 4351
90.12.1 Septembre 2010 . . . 4351
90.12.2 DS octobre 2010, un . . . 4357
90.12.3 DS octobre 2010, deux . . . 4362
90.12.4 Épreuve complémentaire décembre 2010 . . . 4369
90.12.5 Examen décembre 2010 . . . 4371
90.12.6 DS décembre 2011 . . . 4375
90.12.7 Épreuve complémentaire, 15 décembre 2010 . . . 4380

91 Géométrie analytique (Besançon) 4381
91.1 Espaces vectoriels normés . . . 4381
91.1.1 Normes . . . 4381
91.1.2 Topologie . . . 4386
91.2 Courbes et surfaces . . . 4390
91.3 Limite et continuité . . . 4404
91.4 Calcul différentiel . . . 4412
91.4.1 Dérivées partielles . . . 4412
91.4.2 Différentielles . . . 4420
91.4.3 Dérivée d'applications composées . . . 4423
91.5 Intégrales multiples . . . 4435
91.6 Autres exercices . . . 4448
91.7 Exercices pour aller plus loin . . . 4477

92 Exercices d'analyse numérique (Louvain-la-Neuve) 4483

93 Pour des ingénieurs (Louvain-la-Neuve) 4507
93.1 Déterminants et systèmes d'équations . . . 4507
93.2 Opérations sur les matrices . . . 4516
93.3 Espaces vectoriels . . . 4519
93.3.1 Orthogonalité . . . 4525
93.4 Valeurs propres et vecteurs propres . . . 4528
93.5 Triangularisation . . . 4538
93.6 Formes quadratiques . . . 4539
93.7 Interrogation de mars 2010 . . . 4547

94 Pour les géographes (Bruxelles) 4551

95 Pour les pharmaciens (Bruxelles) 4557
95.1 Outils math . . . 4558
95.2 Autres retrouvés . . . 4627

96 Exercices de calcul différentiel et intégral 1 4635
96.1 Prérequis . . . 4635
96.2 Limites et continuité . . . 4643
96.3 Suites numériques . . . 4648
96.4 Limites . . . 4650
96.5 Dérivées et optimisation . . . 4655
96.6 Primitives et intégration . . . 4657
96.6.1 Longueur d'un arc de courbe . . . 4667
96.6.2 Aire d'une surface de révolution . . . 4668
96.7 Équations différentielles . . . 4670
96.7.1 Équations à variables séparées . . . 4670
96.7.2 Équations homogènes . . . 4671
96.7.3 Équations linéaires . . . 4672
96.7.4 Problèmes divers . . . 4673
96.7.5 Équations différentielles du second ordre . . . 4676
96.8 Fonctions de deux variables réelles . . . 4682
96.8.1 Tracer . . . 4682
96.8.2 Limites à deux variables . . . 4682
96.8.3 Dérivées partielles, différentielles totales . . . 4690
96.8.4 Différentiabilité, accroissements finis . . . 4691
96.8.5 Plan tangent . . . 4696
96.8.6 Dérivées de fonctions composées . . . 4699
96.8.7 Dérivées de fonctions implicites . . . 4702
96.8.8 Extrema . . . 4703
96.8.9 Quelques fautes usuelles . . . 4725
96.9 Intégrales de surface, Stokes et Green . . . 4726
96.10 Continuité de fonctions réelles . . . 4729
96.11 Intégrales, longueur de courbes, EDO's linéaires . . . 4730
96.12 Calcul de limites . . . 4732
96.13 Dérivabilité . . . 4733
96.14 Intégration . . . 4734
96.15 Quelques corrections . . . 4735
96.16 Intégration . . . 4739
96.16.1 Série A . . . 4739
96.17 Équations différentielles du premier ordre . . . 4740
96.17.1 Exercices . . . 4740
96.18 Théorème de la fonction implicite . . . 4742
96.18.1 Exercices . . . 4742

96.19 Intégrales curvilignes . . . 4746
96.19.1 Exercices . . . 4746
96.20 Intégrales de surface . . . 4749

97 Exercices de calcul différentiel et intégral 2 4753
97.1 Supremum, maximum . . . 4753
97.2 Suites . . . 4756
97.2.1 Suites définies par récurrence . . . 4764
97.3 Calcul de limites . . . 4767
97.3.1 Limites à deux variables . . . 4772
97.4 Limite et continuité . . . 4776
97.5 Dérivées partielles et différentiabilité . . . 4786
97.6 Séries et séries de puissances . . . 4806
97.7 Exercices de topologie . . . 4814
97.7.1 Exercices ultra basiques . . . 4814
97.7.2 Exercices simplement basiques . . . 4821
97.8 Fonctions d'une variable réelle (suite) . . . 4824
97.9 Développements de Taylor et Maclaurin . . . 4828
97.10 Optimisation sans contraintes . . . 4833
97.11 Équations différentielles . . . 4837
97.11.1 Équations différentielles du premier ordre . . . 4837
97.11.2 Équations différentielles du second ordre . . . 4843
97.11.3 Équations différentielles : modélisation . . . 4849
97.12 Intégrales multiples . . . 4850
97.13 Théorème de la fonction implicite . . . 4857
97.14 Variétés et extrémums liés . . . 4863
97.15 Intégrales curvilignes . . . 4868
97.16 Intégrales de surface, Stokes et Green . . . 4872
97.17 Autres . . . 4881
97.18 Suites de fonctions . . . 4887
97.19 Séries de fonctions . . . 4892
97.20 Existence d'intégrales . . . 4897
97.21 Fonctions définies par des intégrales . . . 4901
97.22 Convergence, continuité et dérivation sous le signe intégral . . . 4902
97.23 Quelques propriétés des espaces fonctionnels . . . 4911
97.24 Équations différentielles . . . 4912
97.24.1 Équations différentielles résolubles . . . 4912
97.24.2 Équation de Bernoulli . . . 4917
97.24.3 Équations de Ricatti . . . 4921
97.24.4 Équations homogènes . . . 4921
97.24.5 Équations différentielles exactes. Facteurs intégrants . . . 4923
97.24.6 L'équation $y' = f\left(\frac{at+by+c}{a't+b'y+c'}\right)$ . . . 4926
97.24.7 Équation d'Euler . . . 4928
97.24.8 Équation dont on peut réduire l'ordre . . . 4929
97.24.9 Quelques exemples d'équations résolubles . . . 4933
97.24.10 L'équation différentielle linéaire du second ordre . . . 4936
97.24.11 Équation différentielle implicite du premier ordre . . . 4938
97.25 Systèmes différentiels linéaires à coefficients constants . . . 4944

98 Exercices pour analyse CTU 4953
98.1 TD 1 : Fonctions monotones, bijections . . . 4953
98.2 TD 2 : Dérivation, fonction composées . . . 4957
98.3 TD 3 : Rappels de trigonométrie . . . 4960

98.4 TD 4 : Fonction trigonométriques réciproques 4962
98.5 TD 5 : Trigonométrie hyperbolique 4965
98.6 TD 6 : Calcul intégral, introduction 4966
98.7 TD 7 : Recherche de primitives 4968
98.8 TD 8 : Intégrale définie 4970
98.9 TD 9 : Équations différentielles : généralités 4973
98.10 TD 10 : Équations différentielles : résolution 4975
98.11 TD 11 : Développements limités 4980
98.12 Interrogations des années précédentes 4984
98.13 Autres exercices 4988
98.14 Exercices en réserve 4992
98.15 Autres 5007

99 GNU Free Documentation License **5011**

Bibliographie **5019**

Index

ad-nilpotent, 3059
α-series of weight, 3138
ϵ-trace, 3849
$\mathfrak{g}$-module, 3043
C^*-algebra, 3742
G-module, 3043
G-structure, 3387
H-invariant subspace, 3599
p-summable Fredholm module, 4077
T_0 topological space, 2921
T_1 topological space, 2921
T_2 topological space, 2921
w^*-limit, 3768
w^*-weak topology, 3739
λ-système, 1272
k-forme différentielle, 1103
p-groupe, 439
p-Sylow, 440
q-orthogonal, 706

Abel
  angulaire, 1886
  convergence radiale, 1445
Abel
  angulaire, 1886
  convergence radiale, 1445
abélianisé, 322
abscisse
  curviligne, 1977
absolument continue, 1416
absolute value of an operator, 3488
absolutely continuous, 3830
absorbant, 2798
abstract
  Cartan matrix, 3095, 3109
  root system, 3088
acceleration, 3615
accélération d'un chemin, 1959
accroissement, 1225
accroissements finie
  dérivée partielle, 933
action, 280
  adjointe, 331
  associated with Lagrangian, 3468
  de groupe
    Wedderburn, 1800
  domaine fondamental, 333
  fidèle, 330
  libre, 336
  of Spin(2) on $\mathbb{R}^2$, 3684
  of Spin(p,q) on End S, 3691
  of a group on a manifold, 3451
  transitive, 336
action de groupe, 661
  sur des matrices, 1939
action fidèle, 330
action transitive, 336
active symbol, 4151
adjoint, 669
  action
    Hopf algebra, 3253
  bundle, 3392
  group, 3207
  in Clifford algebra, 3667
  map, 3147
  operator, 3476
  representation
    Lie group on its Lie algebra, 3050, 3148
    on Hilbert space, 4219
adjoint in Minkowsky space, 3324
admissible
  $U_q\mathfrak{g}$-module, 3814
  function, 4148
affine
  application, 967
  espace, 963
  map, 3609
  sous-espace, 970
  space, 4138
  symmetric space, 3609
aire, 1935
aire dans $\mathbb{R}^2$, 1848
algebra, 3231
  B-module, 3258
  Clifford, 3666, 3683
  of classical ΨDO, 3560
  quaternion, 3683
algèbre, 274

de parties, 1266
engendrée, 275
algèbre de Banach, 605
algèbre engendrée, 1454
algébrique, 495
algébriquement
indépendant, 538
algorithme, 2567
consistant, 2567
facteurs invariants, 424
fortement consistant, 2567
stable, 2567
algorithme convergent, 2567
alignement
dans un espace projectif, 2026
allowed
Young diagram, 3270
almost
complex structure, 3641
normal subgroup, 3888
alphabet, 2847
alterné
groupe, 455
polynôme, 534
analytic Lie group, 3145
analytique entre espaces de Banach, 1206
anchor, 3465
angle
d'une courbe, 1996
entre deux droites, 1707
orienté de vecteurs, 1702
angle between roots, 3108
angle entre deux vecteurs, 1661
anneau, 176
$\mathbb{Z}/n\mathbb{Z}$, 449, 1792, 1800, 1804
à division, 467
de séries formelles, 1515
euclidien
facteurs invariants, 424
factoriel, 360
noethérien, 364
principal, 363, 475, 694
utilisation, 477
quotient par un idéal, 229
Anneau
$\mathbb{Z}/n\mathbb{Z}$
polynôme cyclotomique, 1797
anneau commutatif, 176
anneau intègre, 231
anneau topologique, 583
antipode, 3247
apériodique
état d'une chaine de Markov, 2789
application
de classe C^k, 934
définie positive, 699
différentiable, 928, 934, 1109, 1145, 1570, 1939
extrémums lié, 1585
en escalier, 1926
linéaire
théorème de Banach-Steinhaus, 885
mesurable, 1271
ouverte, 598
semi-définie positive, 699
tangente, 1114
application k-linéaire, 651
application k-multilinéaire, 651
application affine, 710
application bilinéaire, 699
application multilinéaire
décomposable, 902
application ouverte, 887
application quotient, 551
application réciproque, 591
applications homotopes, 2128
applications linéaires semblables, 723
approximate unit, 3756
approximation
de fonctions
par des polynômes, 2732
de l'unité, 2234
par polynômes, 1541
polynomiale, 2159
arc
géométriques, 1976
paramétré, 1959
arc cosinus, 1660
arc sinus, 1658
arc tangente, 1656
archimédien, 188
associated form with a representation, 3770
associativity
constraint, 2917
of a WKB quantization, 4155
associée
subdivision, 1927
associés
éléments d'un anneau, 358
asymptotic expansion, 3542
asymptotiquement pivotale, 2756
asymptotiquement stable, 4731, 4850
atlas, 2961, 2969, 2970
attractif
point fixe, 1557
automorphism

of affine symplectic manifold, 4154
automorphisme, 548
d'espace vectoriel, 395
automorphisme de groupes, 175
axiom
countability, 2921
separability, 2921
axiome
du choix, 168
axis
of rotation in SO(3), 3225

Baire
espace, 650
théorème, 650
tribu, 1283
Baire
espace, 650
théorème, 650
tribu, 1283
balanced product, 3238
balancing, 2919
Banach
algebra, 3483
limit, 3564
space, 3473, 3483
Banach-Alaoglu theorem, 3492
barycentre
cas affine, 972
cas vectoriel, 1683
enveloppe convexe, 976
base, 386
d'un module, 273
de Newton, 2576
de topologie
dénombrable, 572
espace métrique, 572
duale, 430
espace préhilbertien, 2104
hilbertienne, 2104
utilisation, 2359
locale, 2082
base associée à un repère cartésien, 964
base canonique de $\mathbb{R}^m$, 390
base de topologie, 539
base de topologie et continuité, 548
base de voisinages, 540
base orthonormée, 713
base préduale, 431
basic
representation, 3140
basis, 4060
of T_xM, 3383
of a module, 3234
of an abstract root system, 3090
of Hilbert space, 3474
of topology, 2921
Bergman (espace), 2340
Bernoulli, 2699
somme, 2717
Berry-Esséen (borne), 2697
Bessel
inégalité, 2104
betty number, 3593
Bézout
anneau principal, 363
calcul effectif, 342
nombres entiers, 242
polynômes, 488
BHTZ
action, 3931
subgroup, 3931
biais
d'estimateur, 2746
bialgebra, 3245
Lie, 3260
Poisson, 3260
Bianchi identities, 3394, 3402, 4094
bidifferential operator, 3556
bien
conditionné, 2548
enchainé, 627
biholomorphe, 2187
bijection, 168, 591
bilinéaire, 651
bilinear form
coercive, 3581
elliptic, 3581
bimodule, 3234
over a C^*-algebra, 3901
binormale, 1987
birapport, 2039
birapport dans $\mathbb{C} \cup \{\infty\}$, 2058
birégulier
point sur une courbe, 1974
Black hole, 3989
black hole, 3937, 3959
singularity, 3931
Bogoliubov transform, 3896
Bolzano-Weierstrass
espaces métriques, 574
bon
ordre, 170
bonne seminorme, 2397
boost, 3327, 4140
in direction of ∂_1, 4141
bord, 570

Borel measures, 3505
borélienne
fonction, 1281
tribu, 1279
boréliens, 1279
borné, 573
partie de V, 672
temps d'arrêt, 2810
bornée
différentielle, 1132
partie de $\mathbb{R}^m$, 1022
suite, 787
Bott periodicity, 3911
boule
ouverte, 569
boule dans un corps, 285
boule fermée, 572
boundary, 2922, 3437
extreme, 3783
morphism, 3591
of a chain complex, 3592
of a singular p-chain, 3594
bounded
element of $L^2(M)$, 3877
measure, 3505
operator, 3482
positive measure, 3505
subset of $\mathscr{E}(\mathcal{U})$, 2927
braid Hopf algebra, 3256
braiding, 2919
branch
in a Dynkin diagram, 3142
Bruhat (décomposition), 1256
BTZ black hole, 3943
bundle
associated, 3388
principal, 3379
Burnisde
formule, 334

Campbell-Baker-Hausdorff formula, 3653
Campbell-Baker-Hausdorff formula, 3653
canonic
decomposition
of a symmetric Lie algebra, 3614
extension of a representation, 3770
spectral triple, 4082
canonical
basis of a representation, 3136
canonique
base, 390
décomposition, 175
espace affine, 964
Cantor
ensemble, 1332
CAR algebra, 3895
caractère, 1526
abélien, 1519
de S_4, 1534
groupe diédral, 1783
irréductible, 1526
caractéristique
d'un anneau, 275
d'une équation différentielle, 2519
polynôme, 696
sous-groupe, 226
cardinal, 208
cardioïde, 1973, 4397
carré
dans un corps fini, 1805
carrée
matrice, 403
Cartan
abstract matrix, 3109
decomposition, 3128, 3217
involution, 3128, 3214, 3608
matrix, 3081, 3108
abstract, 3095
of $\mathfrak{so}(2l+1)$, 3277
subalgebra, 3071
subgroup, 3110
Cartan-Dieudonné theorem, 3676
Cartan-Weyl basis, 3107
carte, 1838
catégorie
ensemble de première, 630
Cauchy
critère
uniforme, 1163
déterminant, 663
formule, 2153
suite, 616
théorème, 439
Cauchy sequences, 3538
Cauchy-continue, 1015
Cauchy-Riemann, 2123
Cauchy-Schwarz, 819, 2095
causal
effect, 3956
influence, 3942
solvable symmetric black hole, 3943
structure, 3956
Cayley
théorème, 439
Cayley-Hamilton theorem, 2912
cellule d'un pavage, 1926
center

of a bimodule, 3235
of a Lie algebra, 3644
central
decreasing sequence, 3054
extension, 3462
centrale (application), 1526
centralisateur, 226, 228
centralizer, 3070
centre
d'un anneau, 228
d'un groupe, 226
d'une rotation, 1691
cercle
circonscrit à une courbe, 2004
dans la sphère de Riemann, 2050
cercle-droite, 2050
cercles
perpendiculaires, 2019
cesàro mean, 3567
Cesàro
moyenne, 879
chain
complex, 3591
condition, 3055
chaine, 627
de Markov, 2775
convergence, 2803
finie, 2778
homogène, 2775
irréductible, 2778
récurrente positive, 2786
chaine de Markov apériodique, 2789
chaine de Markov régulière, 2778
champ
conservatif, 1898
de vecteurs, 1860
champ de vecteur conservatif, 1863
champ dérivant d'un potentiel, 1898
changement de variable, 1976
character
of a cycle, 4079
of a homomorphism, 3263
of a representation, 3110, 3864
of an abelian group, 3110
of an algebra, 3737
characteristic
Euler-Poincaré, 3593
value, 3476, 3477
chart, 2970
chart on topological manifold, 2961
chart with boundary, 3437
Chasles, 963
chemin, 1876
dans $\mathbb{R}^p$, 1959
chemin C^1 par morceaux, 1858
chemin de Jordan, 1876
chemin régulier, 1876
chevalley
coboundary, 3462
cohomology, 3643
Chevalley
basis, 3081, 3128
cohomology, 3649
Christoffel symbol, 3395, 3399, 3615
circulation, 1896
classe
d'association, 358
de conjugaison, 226
classe d'association, 358
classe de conjugaison
dans S_4 , 260
classical
pseudo-differential operator, 3560
Clifford
algebra, 3666
grading of, 3674
Universal property of, 3067
group, 3674
module, 3685
multiplication, 3685
closed
operator, 3475
sesquilinear form, 3480
trace, 3263
closure, 3475
of a submodule, 3243
clôture algébrique, 496
clutching function, 3911
co-Jacobi, 3246
co-Poisson bialgebra, 3260
co-skew symmetric, 3246
co-weight, 3127
coaction, 3497
coadjoint
action, 3459, 4126
orbit, 4182
coalgebra, 3244
$*$-coalgebra, 3252
coassociativity, 3245
coboundary, 4239
morphism, 3593
cochain complex, 3593
cocycle
of a Schürmann triple, 4228
codifferential, 3419
codimension, 391

coefficient binomial, 353
coefficients
de Fourier, 2238
coefficients binomiaux, 353
coefficients de Fourier, 2250
coercion, 2118
coercive, 1610
colinéarité, 2025
combinaison
convexe, 972
combinatoire, 1718
commutant, 763, 3784, 3823
commutateur
dans un groupe, 321
commutator of vector fields, 2998
compacité, 574, 623, 627, 2163, 2185
sous-groupes du groupe linéaire, 1258
théorème de Dini, 1165
utilisation, 1620
théorème de Montel, 2338
compact, 559, 675
arc paramétré, 1959
boule unité, 783
et fonction continue, 574, 795
fermé et borné, 783
intervalle $[a, b]$, 782
le coup du, 1550
Lie algebra, 3211, 3213
localement, 560
locally, 2923
opérateur, 2189
operator, 3476
produit dénombrable, 629
produit fini, 628
quantum group, 3795
quasi, 559
real form, 3213
relatively, 2923
resolvent, 3491
séquentiellement, 561
suite exhaustive, 625
Compactifié, 4823
compactly embedded, 3211
compacts et fermés, 562
compatible
connection, 4090, 4095
isomorphism and connection, 4094
complément
à deux, 2538
complémentaire, 171
complet, 4385
$\mathbb{R}$
corps, 310
espace métrique, 617
corps, 285
espace mesuré, 1285
espace topologique, 605
métrique, 605
complète
famille de projecteurs, 274
complete
metric space, 2923
symbol, 3555
topological group, 2928
complété
d'un espace métrique, 1630
complete vector field, 3004
completely
positive, 3778
completely reducible representation, 3046
completion, 3473
of a measure, 3864
complétion
projective, 2028
complétion d'espace métrique, 1630
complétude, 1626, 1630, 2214
espaces L^p, 2216
complex conjugate representation, 3044
complex conjugate vector space, 3043
complex structure, 3641
complexe conjugué, 320
complexification
of a vector space, 3212
of Clifford algebra, 3667
composante, 1226
composante connexe, 557
composition
suite de, 325
concaténation
de langages, 2848
de mots, 2847
concave, 1587
log-concave, 1594
condition initiale, 2505
conditionally positive, 4227
conditionnement
absolu, 2547
d'une matrice inversible, 2581
relatif asymptotique, 2567
Cône de lumière, 2859
conformal map, 3725
conformal structure, 4079
conjugate
module, 4089
conjugation, 3212
conjugué hermitien, 715

conjugués
éléments d'une extension, 495
connected part in topological space, 2922
connectés
indices d'une matrice, 2608, 2609
connection, 3412
Levi-Civita, 3395, 3689
on $\mathrm{SO}(\mathbb{R}^2)$, 3695
on an hermitian module, 4090
on Connes' calculus, 4094
on principal bundle, 3400
on symmetric spaces, 3615
on vector bundle, 3392
Connes differential form, 4070
connexe par arc, 798
connexité, 627
définition, 556
et intervalles, 792
fonction holomorphe, 1633
indice d'une courbe, 2147
le groupe $\mathrm{GL}^+(n,\mathbb{R})$, 1247
par arc
fonction différentiable, 1145
points d'accumulation, 623
prolongement analytique, 1634
signature d'une forme quadratique, 1615
théorème de Runge, 2159
théorème des valeurs intermédiaires, 807
utilisation
Brouwer, 2156
consistance
estimateur, 2745
ordre, 2631
constructible
angle, 1830
point, 1826
réel, 1826
construction
des réels, 299
contenu, 376
continue
fonction entre espaces métriques, 621
fonction entre espaces topologiques, 547
forme différentielle, 1103
continuité
fonction définie par une intégrale, 1548
séquentielle, 590
continuous, 2925
functional calculus
selfadjoint in C^*-algebra, 3747
selfadjoint operators, 3488
functional on Banach space, 3484
spectrum, 3481
strongly, 3528
uniformly, 2922
contraction, 905, 1558, 3739
of a tensor, 3233
contragredient representation, 3459
contravariant
functor, 2915
convergence
commutative, 862
dans un espace vectoriel normé, 784
de martingales, 2816
en loi, 2684
en moyenne quadratique, 4911
en probabilité, 2683
in topological spaces, 2925
of a net, 2933
ordre, 2563
presque surement, 2683
quadratique, 2552
rapidité, 1484, 2374, 2375, 2562, 2717
strong in $\mathcal{L}(\mathscr{H})$, 3476
suite
dans un corps, 285
suite dans $\mathbb{R}^m$, 599
suite numérique, 784, 1541, 2349
Abel angulaire, 1886
uniforme, 1163
intégrale, 1546
théorème de Dini, 1165
convergence absolue, 858
convergence de suite, 542
convergence forte, 886
convergence normale, 858
convergence uniforme
série de fonctions, 858
convex, 3783
cone, 3130
subset, 3768
convexe
courbe plane, 2006
fonction sur $\mathbb{R}^n$, 1597
norme, 4381
convexité
barycentre, 974
enveloppe de $O(n)$, 1255
fonction, 1587
inégalité de Jensen, 1602
locale, 587
méthode de Newton, 2558
utilisation, 1620
convolution, 2653, 2805, 4226
on coalgebra, 3257
on compact quantum groups, 3798

over a bialgebra, 3247
product of distribution, 3507
semigroup, 3257
twisted, 4119, 4120
convolution sur S^1, 2250
coordinate
local, 3522
coordonnées
cartésiennes
dans un espace affine, 979
curvilignes, 2082
cylindrique, 1769
dans un espace affine, 964
homogène, 2044
sphériques, 1770
coordonnées barycentrique, 980
coordonnées polaires, 1763
coproduct, 3245
corepresentation, 3797
corps, 236
complet, 285
de décomposition, 520
de rupture, 510
polynôme cyclotomique, 1797
des fractions, 281, 284
extension, 527, 536
fini, 1804, 1808, 1811
Wedderburn, 1800
formellement réel, 467
ordonné, 285
premier, 469
corps algébriquement clos, 496
corps valué, 318
cosinus, 1637
hyperbolique, 1486
cosphere, 3555, 3568
cotangent bundle, 3009
counit, 3245
countability axioms, 2921
courbe, 1959
efficacité, 2766
étude métrique, 2359
fermée, 1993
simple, 1993
courbe de Jordan, 1876
courbe de niveau, 1084, 1086
courbe simple, 1876
courbure, 1987
signée, 1994
totale, 1996
covariance, 2656
covariant
derivative, 3413
exterior , 3396
on associated bundle, 3403
on vector bundle, 3392
star product, 4157
covering
universal, 2934
covering map, 2934
critère
Abel, 1434
Abel pour intégrales, 1547
Cauchy
uniforme, 1163
de Cauchy, 310, 617
de la racine, 4552
Weierstrass, 1547
série de fonctions, 1168
critère du quotient, 874
critique
Galton-Watson, 2805
point, 1581
point d'un arc, 1974
région, 2765
valeur, 2766
cup product, 3266
current, 3503
1-form, 3718
curvature, 3394, 3415, 4093
(scalar) of a spectral triple, 4221
extrinsic, 4222
form, 3402
on Connes' calculus, 4094
on principal bundle, 3402
cycle, 255, 3263, 3592
cyclic
module, 3227
representation, 3773
vector, 3773, 3892
cyclique
endomorphisme, 692
matrice, 692
cycloïde, 4395
coordonnées normales, 1979
longueur, 1973

d'Alembert operator, 4146
d'Alembert operator, 4146
décalage, 2539
decomposable
in an abstract root system, 3090
décomposition
Bruhat, 1256
canonique, 175
corps, 520
Dunford, 749

application, 1491
exponentielle de matrice, 1191
Jordan
et exponentielle de matrice, 1490
polaire, 1252
primaire, 748
sous-espaces caractéristiques, 748
spectrale, 748
decomposition
Cartan, 3128
Iwasawa, 3130, 3219
Jordan, 3052
of operator, 3866
polar in von Neumann algebra, 3825
root space, 3129
décomposition de Hahn, 1364
décomposition de Jordan, 1366
décomposition décimale, 882
definition representation, 3044
deformation, 4153
degenerated representation of C^*-algebra, 3773
degré
application $S^1 \to S^1$, 1997
d'une représentation, 433
extension de corps, 493
degré d'un polynôme, 277
demi-plan, 1057
dénombrable, 210
à l'infini, 561
dénombrement, 1718
partitions de $\{1, \dots, n\}$, 1515
dense, 545
nulle part, 630
densité, 1630
d'une variable aléatoire, 2640
dans un espace de fonction
critère de Weyl, 2349
de $\mathbb{Q}$ dans $\mathbb{R}$, 780
utilisation, 1594
de $\mathrm{GL}(n, \mathbb{R})$ dans $\mathbb{M}(n, \mathbb{R})$, 1247
de $\mathscr{D}(\mathbb{R}^n)$ dans $L^1(\mathbb{R}^n)$, 2337
de $C_c^\infty(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$, 2221
de $L^2\big([0,1]\big)$ dans $L^p\big([0,1]\big)$, 2223
de $S^+(n, \mathbb{R})$ dans $S^{++}(n, \mathbb{R})$, 1251
des fonctions étagées dans L^p, 2221
des polynômes
dans $C_c^0[0,1]$, 2732
matrices diagonalisables dans $\mathbb{M}(n, \mathbb{C})$, 1247
mesure, 1355
points extrémaux dans $\mathcal{L}$, 1254
prolongement, 1626
densité conjointe, 2645
densité d'une mesure, 1355
densité jointe, 2647
density, 3388
déplacement, 1734
dérivabilité
fonction définie par une intégrale, 1550
lemme de Borel, 1513
dérivable
au sens complexe, 1141
fonction, 2455
derivation
inner, 3208
of a distribution, 3500
of a Lie algebra, 3049, 3142
over a C^*-algebra, 3901
dérivation
au sens des distribution
Sobolev, 2442
dérivé
groupe, 321
derived
Lie algebra, 3054
series, 3054
dérivée, 929
dans Sobolev $H^1(I)$, 2439
distributionnelle, 2408
faible, 2394
fonction à valeurs dans E', 650
partielle, 1097, 1225
seconde, 1064
dérivée directionnelle, 928, 929
dérivée faible, 2392
dérivée partielle, 929
déterminant, 654
Cauchy, 1544
d'un endomorphisme, 658
d'une famille de vecteurs, 655
de Cauchy, 663
et inversibilité, 658
forme linéaire alternée, 653
Gram, 662, 1544
interprétation géométrique, 1391
matrice, 409
résultant, 664, 1891
utilisation, 1620
Vandermonde, 660
détermination
logarithme, 2141
principale, 2141
détermination du logarithme, 2144
développable
en série entière, 1494
développement
asymptotique, 1223

limité
en zéro, 1218
fonction holomorphe, 2123
premier ordre, 932
Taylor, 1939
diagonal operator, 3868
diagonale
dominante, 2613
fortement dominante, 2613
strictement dominante, 2613
diagonalisable, 726
et polynôme minimum scindé, 727
exponentielle, 1491
diagonalisation
cas complexe, 1034
cas réel, 731
endomorphisme autoadjoint, 823
simultanée, 729
diamètre, 1022
diédral, 1718
diffeomorphism, 3232
difféomorphisme, 935
de classe C^k, 1156
différence
centrée, 2619
divisée, 2576
progressive, 2619
régressive, 2619
différentiabilité, 1145
differentiable
manifold, 2969
map, 3232
vector, 4156
differential, 3530
form, 3009
vector-valued, 3376
operator, 3555
left invariant, 3556
differential form, 3363
différentielle, 928
de $u \mapsto u^{-1}$, 953
partielle, 952
totale, 1225
dilatation, 1240
dilatation (matrice), 421
dimension, 390
n-formes multilinéaires alternées, 654
définition, 390
function, 3243
of a free module, 3236
of a spectral triple, 4081
simple, 4081
spectrum, 4081
of a symplectic triple, 3648
of a von Neumann algebra, 3837
sous espace affine, 970
utilisation, 976
Dirac
distribution, 3500
dirac
matrices, 3671, 3673
operator, 3666, 3699
on (M, g), a spin manifold, 3692
on $\mathbb{R}^{1,3}$, 3673
on $\mathbb{R}^2$, 3694, 3697
on functions, 3694
dirac type operator, 3561
direct
integral decomposition, 3864
limit, 2916
sum of Lie algebras, 3142
system, 2916
direct sum representation, 3045
directed set, 2933
direction, 929, 1982
sous-espace affine, 970
Direction, 4021
Dirichlet
algebra, 3904
completely, 3899
form, 3899
C^*, 3899
noyau, 2346
théorème, 2346
théorème (sur les nombres premiers), 1800
displacement group, 3606
disque de convergence, 1433
distance, 569
associée à une norme, 579
compatible, 609
entre deux mesures de probabilités, 2761
entre un ensemble et un point, 4822
invariante, 609
point et ensemble, 601, 4388
distance discrète, 601
distance on spectral triple, 4084
distance produit, 595
distingué
sous-groupe, 226
distribution, 2405, 3499
de Dirac, 2426
équation de Schrödinger, 2501
marginal, 4225
produit par une fonction, 2408
tempérée, 2425
distribution on a manifold, 3011

divergence, 1231, 3397
diviseur
  dans un anneau, 229
  de zéro, 231
  de zéro à droite, 231
  polynôme, 375
diviseur de zéro, 231
divisible, 194
division
  euclidienne, 238, 375
  harmonique, 2061
Dixmier trace, 3568
domaine, 777
  fondamental d'une action, 333
dominant weight, 3113
dominé
  modèle statistique, 2744
dominée
  convergence (Lebesgue), 1352
  mesure, 1359
dominée par le dessus, 632
dominée par une seminorme, 632
double covering
  of SO(2), 3684
  of $\mathrm{SO}_0(p,q)$, 3678
  of $\mathrm{SO}(V)$, 3681
drift
  process, 4238
Drinfel'd twist, 4108
droite
  dans la sphère de Riemann, 2049
  projective, 2025
droite affine, 981, 1050
droite réelle achevée, 802
droite vectorielle, 1050
droites parallèles, 1050
droites perpendiculaires, 1050
dual
  connection, 4094
  d'un espace de Hilbert, 2101
  de $\mathbb{M}(n,\mathbb{K})$, 721
  de $L^p(\Omega)$, 2311
  de L^p, 2306
  Hopf algebra, 3256
  in strict monoidal categories, 2918
  module, 3502
  of a module, 3594, 4087
  of a vector space, 3009
  of Hopf algebra, 3251
  pair of Hopf algebra, 3251
  space, 3502
  topological, 3502
dual algébrique, 430
dual module, 3685
dual topologique, 956, 2093
Dunford
  décomposition, 749
dyadique, 1382
  développement, 306
Dynkin
  coefficient, 3113
  diagram, 3113
    of $\mathfrak{so}(2l+1)$, 3277
Dynkin diagram, 3096

ecart, 2924
ecart, 2924
écart-type, 2655
échantillon, 2738, 2740
effectif
  empirique, 2770
effective
  action, 3451
  symmetric space, 3612
efficacité
  courbe, 2766
  d'une méthode itérative, 2571
eigenvalue, 3484
electromagnetism, 3715
élément
  inversible
    dans un anneau, 231
élément
  de surface, 1871
  de torsion, 338
élément de surface, 1844
élément maximal, 169
élément minimal, 169
élément premier, 230
élément régulier, 180
élément régulier à gauche, 180
élémentaire
  polynôme symétrique, 534
elementary
  normal symplectic algebra, 3656
  representation, 3140
  solvable
    exact triple, 3650
  solvable exact triple, 3648
ellipsoïde, 774
elliptic
  boundary condition, 4224
  pde, 3571
elliptic operator, 3555
elliptique
  équation aux dérivées partielles, 2529
embedding, 3010

endomorphism
of ΛW, 3670
of module, 3240
endomorphisme, 395
cyclique, 692
décomposition
polaire, 1252
diagonalisable, 772, 1249, 2497
Dunford, 749
diagonalisation, 731
nilpotent
Dunford, 749
préservant une forme quadratique, 1261
sous-espace stable, 749, 2497
endomorphisme direct, 668
endomorphisme préserve l'orientation, 668
energy
associated with a Lagrangian, 3468
in the representations of $\mathfrak{so}(2, d-1)$, 3343
energy-momentum, 4143
Engel theorem, 3059
engendré, 236
λ-système, 1273
corps, extension, 502
idéal dans un anneau, 236
sous-espace affine, 970
sous-groupe, 268
ensemble
archimédien, 188
de Cantor, 1332
différence symétrique, 173
infini, 204
ensemble connexe, 556
ensemble des mots, 2847
ensemble étoilé, 4881
ensemble fini, 204
ensemble ordonné, 169
ensemble quotient, 174
ensembles
disjoints, 167
entier, 200
entire function, 3897
entrelacement, 1526
enveloping algebra, 3234
enveloppe
convexe, 974
epimorphism, 3231
équation
différentielle
linéaire, 2458
ordinaire d'ordre 1, 2455
variables séparées, 2460
générale de degré n, 538
equation
Klein-Gordon, 3665
Schrödinger, 3665
Yang-Baxter, 3256
équation
aux variations, 2478
de Riccati, 2499
des classes, 333
des orbites, 333
différentielle
étude qualitative, 2497
Hill, 2496
homogène, 2499
système, 2497
diophantienne, 351, 370, 372
Fredholm, 1560
orbite-stabilisateur, 332
équation de droite, 1056
équation différentielle
linéaire du premier ordre, 2509
linéaire du premier ordre, homogène, 2509
linéaire du second ordre, 2512
linéaire du second ordre, homogène, 2512
premier ordre, 2504
second ordre, 2505
variables séparables, 2507
équation exponentielle, 1199
équation fonctionnelle, 1199
équation homogène associée, 2509
équicontinu, 629
équipotent, 203
equivalence
of categories, 2915
of Lévy process, 4226
of observer, 3702
of principal bundle, 3386
of representation of C^*-algebra , 3773
of spectral triple, 4081
of states, 3790
of vector bundle, 3351
representation of Clifford, 3685
équivalence
arcs paramétrés, 1975
chemin, 1968
classe de fonctions, 2194
de représentations, 1526
de suites, 786
homotopie, 2149
suite de composition, 327
équivalence de forme quadratiques, 743
équivalence de normes, 840
equivalence of orientation, 3421
equivalence of projectors, 3837

equivalent
Murray-von Neumann, 3905
equivalent representations, 3044
equivariant, 3389, 3460, 3495
function, 3413
momentum map, 3461
vector field on principal bundle, 3386, 3466
ergodic, 3830
erreur, 2577, 2607
assignation, 2542
de consistance, 2630
discrétisation, 2631
quadratique, 2580
troncature, 2542
erreur relative, 2542
escalier, 1313
ESET, 3648
espace
L^2
Sobolev, 2442
L^p, 2198
de Baire, 650
de Bergman, 2340
de fonctions
L^p, 2216
Sobolev H^1, 2442
de Hilbert
espace de Sobolev H^1, 2442
de probabilité, 2639
de Schwartz, 2334, 2425
de Sobolev, 2439, 2446
euclidien, 835
mesurable, 1263
mesuré, 1268
complété, 1288
métrique, 569
base de topologie, 572
projectif, 2025
propre, 685
séparé, 553
topologique
métrisable, 620
vectoriel, 271
dimension, 654
vectoriel topologique
métrisable, 603
espace de Banach, 605
espace de Sobolev, 2444
espace tangent, 1840
espace topologique, 539
espace topologique normal, 2014
espace vectoriel
topologique, 580
espace vectoriel normé, 577
espérance, 2650
conditionnelle, 2660, 2661
événement, 2661
variable aléatoire, 2666
essential subspace of a representation, 3770
estimateur, 2745
biais, 2746
consistant, 2745
de fonction de répartition, 2750
maximum de vraisemblance, 2748
estimateur convergeant, 2745
estimation
des grands écarts, 2717
étagée
fonction, 1313
état
apériodique, 2789
récurrent, 2782
récurrent positif, 2782
transitoire, 2782
Éther, 2854
étoile de Kleene, 2849
étranger
dans leur ensemble, 376
étrangers
polynômes, 376
Euclide
algorithme étendu, 340
lemme, 344
euclidien
anneau, 246
espace, 713
Euler
fonction, 4908
fonction β, 4910
euler-Poincaré characteristic, 3593
évaluation
polynôme plusieurs variables, 358
even spectral triple, 4081
Événement, 2859
événement, 2639
event horizon, 3959
exact
intervalle de confiance, 2754
exact sequence
short, 3155
exact triple, 3648, 3649
excès
intervalle de confiance, 2754
exhaustive (suite de compacts), 625
exponential from a Lie algebra, 3171
exponentielle, 1470

convergence, 877
de matrice, 955, 1490, 1491, 1503
utilisation, 1579
existence, 1463
rapide, 1794
unicité, 1200
exposant, 446, 772
d'un groupe, 253
extension
corps de base, 751
de corps, 492, 536
algébrique, 505
finie, 1813
simple, 502
utilisation, 1831
isométrie, 1627
lemma, 4172
of group, 3153
of Heisenberg algebra, 4174
séparable, 530
extension algébrique, 496
extension algébriquement clos, 496
extension of Lie groups, 3155
split, 3155
exterior
covariant derivative, 3402
derivative, 3396
point, 3959
extrapolation, 2576
extreme point, 3783
extrémité
d'un intervalle, 792
extrémum, 1939
lié, 1585
local
relatif, 1585
extrémum local, 1580
extrémums, 1583
volume d'un ellipsoïde, 1620
extrinsic curvature, 4222

face
of a singular p-simplex, 3594
face
of a singular p-simplex, 3594
facteur
intégrant, 2500
factor, 3824
finite, 3852
of type II, 3840
of type I, 3840
of Type I_n, 3840
of type II_1, 3829, 3840, 3852, 3853
of type II_∞, 3840
of type III, 3840
purely infinite, 3846
semifinite, 3846
factoriel
anneau, 360
factorisation
de polynôme, 381, 509
faisceau de droites, 2047
faithful
functional, 3849
functor, 2915
state, 3761
faithful representation, 3044
famille
sommable, 866
famille trigonométrique
sur S^1, 2249
Fatou, 1341
Fejér
noyau, 2346
fermé, 540, 675
fermeture séquentielle, 602
fermionic action, 4099
fibration, 4185
field
algebraic structure, 3231
of Hilbert space, 3864
strength, 3722, 4097
filtration, 2807
filtration naturelle, 2807
final projection, 3838
fine
subdivision, 1959
finite
factor, 3842
projection, 3842
finite dimensional
module over Hopf algebra, 3255
finite type module, 3234
first countable, 2921
fixateur, 330
flag, 3060
flat, 3021
flot, 2472, 2520
flow, 3003
flux
d'un champ de vecteur, 1875
flux d'un champ de vecteurs, 1873
fonction, 777
Γ d'Euler, 2179
à décroissance rapide, 2335
borélienne, 1281
caractéristique, 1929

d'une variable aléatoire, 2678
continue
égales, 603
convexe, 1587, 1594
croissante, 778
de Dirichlet, 2352
de Möbius, 1822
de répartition, 2678
décroissante, 778
définie par une intégrale, 1421, 1509, 1548, 1553, 2185
Γ d'Euler, 2179
utilisation, 2705
en escalier intégrable, 1927
étagée, 2221
génératrice, 2679
holomorphe, 1141, 2185
théorème de Montel, 2338
image, 777
méromorphe
Γ d'Euler, 2179
monotone, 778
valeurs vectorielles, 1226
fonction analytique, 1211
fonction continue en un point, 547
fonction dérivée, 929
fonction intégrable, 1344
fonction périodique, 1075
fonctionnelle
énergie, 2122
fonctionnelle de Minkowski, 633
fonctions équivalentes, 555
fondamental
domaine d'une action, 333
form core, 3481
form of degree k, 4101
formal
extension lemma, 4115
forme
bilinéaire, 699
non dégénérée, 700
différentielle, 1102
exacte, 1884
fermée, 1884
linéaire
différentielle, 1585
quadratique, 1615, 1617, 1939
groupe orthogonal, 1261
forme bilinéaire symétrique, 699
forme canonique
fonction simple, 1313
matrice de transition, 2799
forme linéaire, 430, 653
forme multilinéaire alternée, 653
forme quadratique, 699
formle multilinéaire antisymétrique, 651
formule
Bayes, 2657
Burnside, 334
d'expulsion (produit vectoriel), 833
de Cauchy, 2153
Hadamard, 1437
inversion Möbius, 1823
probabilité totales, 2657
sommatoire de Poisson, 2374
Taylor
utilisation, 2562
formule binomiale, 354
Formule de Leibniz, 2395
formule de Stirling, 1945
formule de Taylor
reste intégral, 1920
formule des classes, 332
Fourier, 2374
série
utilisation, 2359
transformée
groupe abélien fini, 1521
fourier transform, 3553
Fourier transform, 3501, 3523, 3525
twisted, 4118
fraction
rationnelle
intégration, 1891
fraction dyadique, 306, 2017
fraction rationnelle, 1917
fractions
rationnelles, 281
fractions (corps), 284
frame, 3384
bundle, 3405
on AdS_l, 4060
moving, 3384
Fréchet
algebra, 3527
space, 3527
Fredholm
équation, 1560
module
associated with a K-cycle, 4080
even, 4077
odd, 4077
operator, 4076
free
action, 3597
family in a module, 3234

part of a black hole, 3944
vector, 4138
Frenet
formules, 1989
fréquence
empirique, 2770, 2780
Fresnel
intégrale, 2183
Frobenius
morphisme, 276
frontière, 545, 570, 679
Fubini
théorème
dans $\mathbb{R}^n$, 1415
full functor, 2915
functional, 3484
functional calculus
continuous, 3488
functor
contravariant, 2915
functor (covariant), 2915
fundamental
1-form, 3376
representation, 3812
sequence of subset, 2926
vector field, 3167, 3400, 3452, 3597, 3928, 3962
on $\mathbb{R}^2$, 4162
Future, 3989
future
directed vector, 3719, 3942, 3959
horizon, 3931
interior region, 3931

Gårding space, 3533
Gårding space, 3533
Galerkin approximation, 3584
Galton-Watson
sous-critique, 2805
sur-critique, 2805
gauge
invariance, 3716
of electromagnetism, 3716
potential, 4097
(universal connection), 4092
principal potential, 3412
transformation, 3413
local description, 3409
of principal bundle, 3391
of section of associated bundle, 3391
gauged
Dirac action, 4100
Dirac operator, 4100
Gauss
lemme
polynômes, 489
polynomial, 3810
somme de, 1806
Gaussian
part of a representation, 4239
Schürmann triple, 4238
Gelfand
theorem, 3743
topology, 3739
transform, 3739
générateur, 269
generating family of a module, 3234
generator
of infinitesimal rotations, 3225
génératrice
partie d'un module, 273
géométrie
avec des groupes, 1712, 2068
avec nombres complexes, 1712, 2068
géométrique
avec des nombres complexes, 2359
Gershgorin
disque, 2610
global symmetry, 3716
globally group type, 4183
GNS representation, 3898
graded differential, 4078
gradient, 931, 1111, 1119, 1131, 3397
Gram (déterminant), 662
Gram-Schmidt, 835
graphe, 778, 1066
de transition (chaine de Markov), 2780
fonction, 1083
fonction de deux variables, 1085
grassmann connection, 4091
Grönwall (lemme), 2455, 2456
grothendiek group, 3906
group
adjoint, 3207
homotopy, 2934
isotropy, 3454, 3471
Lorentz, 3710
Weyl, 3086
group ring, 3232
groupe, 175
$\mathrm{GL}(n, \mathbb{R})$, 1617
p-groupe, 439
action, 2068
utilisation, 1261
agissant sur un ensemble
diédral, 1711
alterné, 455

circulaire, 2065
de Galois, 538
de permutation, 1783
de permutations, 1718
caractères de S_4, 1534
de torsion, 338
dérivé, 321
de $\mathrm{GL}(n, \mathbb{K})$, 1245
de $\mathrm{SL}(n, \mathbb{K})$, 1245
du groupe alterné, 458
du groupe symétrique, 457
des isométries
espace métrique, 601
des symétries, 1730
diédral, 1711, 1718
générateurs (preuve), 1712
générateurs (utilisation), 1783
en géométrie, 1711
et géométrie, 654, 1718, 2068
isométries du cube, 464
fini, 441, 449, 1718, 1792, 1800, 1804
alterné, 457
diédral, 1711
Wedderburn, 1800
linéaire, 1256
décomposition polaire, 1252
enveloppe convexe de $\Omega(n)$, 1255
hyperplan, 722
sous-groupes compacts, 1258
modulaire, 2068
orthogonal, 671
d'une forme quadratique, 1261
partie génératrice, 457, 1792, 2068
permutation, 654, 661, 1256, 1792
diédral, 1711
projectif, 2032
quotient, 325
simple, 226
spécial orthogonal, 672
symétrique, 254
action sur un triangle, 1523
groupe abélien, 175
groupe commutatif, 175
groupe cyclique, 269
groupe de pavage, 1734
groupe dérivé
de $\mathrm{GL}(n, \mathbb{C})$, 1239
groupe ordonné, 225
groupe résoluble, 329
groupe simple, 226
groupoid, 3468

Haar
measure, 3797
state, 3798
Haar
measure, 3797
state, 3798
Haar functional, 2938
Haar measure, 2938, 2958
Hadamard
conditions, 2530
formule, 1437
Hahn-Banach theorem, 3483, 3494
Hamiltonian
action, 3458, 3649
weak, 3458
field, 3458
Hardy-Littlewood (théorème), 1541
harmonic form, 3420
harmonique
fonction, 4480, 4591
hasse diagram, 2931
Hausdorff, 553
Hausdorff space, 2923
heat kernel, 3661
height of a representation, 3137
Heine (théorème), 1023
heisenberg
algebra, 3656
hermitian, 3476
metric, 3650
structure on a module, 4087
symmetric space, 3641
Hermitian
functional, 4227
hermitian conjugate, 3289
hermitian operator, 3290
hermitienne, 715
hessienne, 1158
hexagonal identities, 2919
highest weigh vector
for $U_q\,\mathfrak{sl}(2, \mathbb{C})$, 3817
highest weight, 3111
for group representation, 3110
in $U_h\mathfrak{g}$-modules, 3813
module, 3817
Hilbert
espace, 2095
Hilbert space, 3473, 3508
Hilbert-Schmidt
class operator, 3492
operator, 3478, 3479
Hochschild
coboundary, 3263
cohomology, 3265
Hodge operator, 3021

hölder inequality, 3480
holomorphe, 1141, 2187
    sur un compact, 1141
holonomic constraint, 3462
holonomy group, 3408
homeomorphism, 3232
homéomorphisme, 548
homogène
    chaine de Markov, 2775
homogeneous
    linear PDE, 3569
    Riemannian metric, 3604
    space, 3454, 3598, 3612, 3640, 3954
homographie, 2031, 2068
    sur $\mathbb{C} \cup \{\infty\}$, 2054
homology
    group, 3592
homomorphism, 3231
    of Lie group, 3153
    of Poisson algebra, 3260
    of principal bundle, 3382
    of symmetric space, 3606
homotopie, 1876
homotopie à extrémité fixées, 1876
homotopy
    of morphism of chain complexes, 3592
homotopy group, 2934
homset, 2915
Hopf algebra, 3247
    $*$-Hopf algebra, 3252
    dual, 3256
    quasitriangular, 3256
    representation, 3255
horizon, 3937
Horizon, 3989
horizontal, 3400
    lift, 3400, 3401
    space, 3400
hull-kernel topology, 3773
hyperbolic
    pde, 3571
hyperbolique
    équation aux dérivées partielles, 2529
hyperplan, 768, 1677
    de $\mathbb{M}(n, \mathbb{K})$, 722
    sépare
        au sens strict, 2288
    séparer
        au sens large, 2288
hypothèse
    alternative, 2765
    composite, 2765
    multiple, 2765
    nulle, 2765
    simple, 2765

idéal
    bilatère, 228
    dans un anneau, 228
    principal
        à droite, 240
        à gauche, 240
idéal
    bilatère, 228
    dans un anneau, 228
    principal
        à droite, 240
        à gauche, 240
idéal à droite, 228
ideal in Banach algebra, 3760
idéal maximal, 237
identifiable, 2744
identification vector, 3931
identité
    parallélogramme, 4381
identité de polarisation, 704
identity
    in monoidal category, 2917
image, 777
immersion, 3002, 3009
incompressible
    champ de vecteur, 1233
increasing
    net of projections, 3876
indécomposable
    module, 274
indecomposable module, 3111
indépendance
    affine, 978
    algébrique, 538
    événements, 2641
        utilisation, 2730, 2818
    projective, 2033
    sous tribus, 2641
    variables aléatoires, 2641
independent
    with respect to a state, 4226
index, 3875, 3877
indicatrice d'Euler, 451
indice, 324
    d'une courbe dans $\mathbb{C}$, 2146
    de rotation, 1998
indice d'inertie, 711
induced
    Clifford module, 3685
    distribution, 3502
    representation, 3495

topology, 2921
induced orientation, 3440
inductif, 171
induite
topologie, 676
tribu, 1265
inégalité
arithmético-géométrique, 1603
Bessel, 2104
Cauchy-Schwarz, 819, 2095
de Khintchine, 2714
de la moyenne, 1145
des pentes, 1588
Hölder, 2203
utilisation, 2654, 2733
isopérimétrique, 2359
Jensen, 1602
espérance conditionnelle, 2677
pour une somme, 1602
version intégrale, 2202
Kantorovitch, 1603
Markov, 2693
Minkowski, 2206
triangulaire, 569, 577
inférence statistique, 2737
infimum, 311
infinite
projection, 3842
infinitesimal
of order 1, 3567
infinitesimal operator, 3565
initial projection, 3838
injection, 168, 591
inner
automorphism, 3072, 3208, 3215, 3266, 4217
fluctuation of the metric, 4219
intégrable
fonction à valeurs vectorielles, 1348
fonction non en escalier, 1933
fonction positive, 1938
integrable distribution, 3011
integrable scalar, 3506
integral
curve, 3023
lower, 3505
on a spectral triple, 4084
upper, 3504
integral curve, 3003
integral manifold, 3011
intégrale
calcul, 2349
convergente, 1405, 1890
d'une fonction sur une carte, 1845
d'une fonction sur une variété, 2238
d'une forme différentielle, 1861
fonction en escalier, 1927
Fresnel, 2183
impropre, 1405
sur un chemin, 1859
intégrale d'une fonction, 1335
intégrale d'une forme différentielle, 1861
intégrale de Dirichlet, 1956
intégrale existe, 1344
intégrale sur une carte, 1841
intégration
fraction rationnelle, 1891
intéressante, 163
intérieur, 543
d'un ensemble, 674
point, 674
interior
product, 3019
region, 3959
internal
metric, 3914
space, 3914
interpolation, 2576
Intervalle, 2859
intervalle, 170, 777, 798
longueur, 1321
intervalle de confiance
asymptotique, 2758
invariance cyclique
trace, 404
invariant
de similitude, 759
form, 3053
function on a group, 4170
Lie subalgebra, 3062
vector subspace, 3043, 3052
invariant subspace, 3044
invariant vector field, 3157
invariante
mesure
pour une chaine de Markov, 2800
inverse
dans un groupe, 225
root, 3079, 3103
inverse généralisé, 2721
inversion, 2020
dans le groupe symétrique, 260
inversion dans $\mathbb{C} \cup \{\infty\}$, 2052
involution, 729, 3252
automorphism, 3612
Cartan, 3128, 3608
on a $\mathbb{K}$-algebra, 3252

on algebra, 3482
on complex vector space, 3252
on Hopf algebra, 3252
involutive
automorphism, 3214, 3608, 3610, 3611
Lie algebra, 3647
involutive distribution, 3012
irrationalité
$\sqrt{2}$, 294
irreducible
abstract root system, 3088
algebraically, 3783
representation, 3043
topologically, 3783
vector space, 3052
irreducible representation, 3044
irréductible
chaine de Markov, 2778
dans un anneau, 230
module, 274
représentation, 1524
isobarycentre, 972
isométrie
de forme quadratique, 708
de l'espace euclidien $\mathbb{R}^2$, 1712
espace euclidien
isométries du cube, 464
groupe, 601
isométrie (forme bilinéaire), 708
isométrie d'espaces métriques, 600
isométrie positive, 1681
isometry
in Hilbert space, 3476
isomorphism
of abstract root system, 3089
of categories, 2915
of Lie groups, 3150
of modules, 3853
of principal bundle, 3382
of symplectic
symmetric spaces, 3649
triple, 3648
isomorphism of topological manifold, 2962
isomorphisme, 891
$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, 1804
d'anneaux, 227
de corps, 263
espace affine, 969
espace vectoriel normé, 937
espaces vectoriels, 396
isomorphisme d'espace topologique, 548
isospectral
manifold, 4221
isotrope
cône, 705
sous espace, 705
totalement, 705
vecteur, 705
isotropic
representation, 3604
subspace
completely, 3669
in $\mathbb{R}^2$, 3682
isotropy
group, 3451, 3454, 3471
Linear group, 3454
Iwasawa
coordinates, 3652
decomposition, 3130, 3219
of $\mathrm{SL}(2,\mathbb{C})$, 3283
of $\mathrm{SO}(1,n)$, 3620
of $\mathrm{SP}(2,\mathbb{R})$, 3284
of $SL(2,\mathbb{R})$, 3221
group, 3130, 3947
Iwasawa decomposition
of $\mathrm{SO}(2,1)$, 3617
of $\mathrm{SO}(2,n)$, 3621

Jacobi, 3246
identity, 3049
Jacobi, 3246
identity, 3049
jacobien, 1131, 1231
Jacobson topology, 3773
jauge, 2281
Jordan
forme, 4947
réduction, 762
Jordan decomposition, 3052
in Hilbert space, 3494
Jordan-Hölder, 326
junk form, 4069, 4070

K-cycle, 4080
K-cycle, 4080
kähler
2-form, 3651
manifold, 3650
Kaplansky density theorem, 3868
kernel
for a WKB quantization, 4155
for extension $(\overline{X}, 0, 2)$ of Heisenberg, 4180
ket-bra operator, 4089
Killing
form, 3053, 3114
bi-invariance, 3148
killing

form, 4160
vector field, 3023
KMS
condition, 3895
symmetric map, 3896
Kronecker
symbole, 390
Kronecker product, 3139

lacet, 1876
lacet, 1876
lacet de Jordan, 1876
ladder operators, 3068
Lagrange
multiplicateur, 1585
polynôme, 721
théorème, 325
Lagrange-Poisson structure, 3468
Lagrangian, 3468
subspace, 3457
vector field, 3468
lagrangien, 1585
langage, 2848
itéré, 2849
itéré strict, 2849
unité, 2848
vide, 2848
Laplace
transformée, 2679
Laplace operator, 3524
laplace operator, 3397
laplace type operator, 3561
Laplace-Beltrami operator, 3420, 3661
laplacien, 2089, 4423
layer, 3137
Lebesgue measure, 3504
left
bounded vector, 3853
ideal, 3760
invariant
kernel, 4155
support, 3838
left invariant
differential operator, 3556
function on a group, 2928
linear form on a Hopf algebra, 3255
operator, 3556
vector field, 3157, 3597
left invariant vector field, 3157
left translation, 3147, 3157
left-invariant vector field, 3158
Legendre
symbole, 1805
legendre mapping, 3468

Leibniz, 1068
applications entre espaces vectoriels normés, 951
Leibniz formula, 3009
lemme
Borel, 1513
d'Euclide, 344
de Borel-Cantelli, 2691
de Gauss
contenu de polynôme, 1813
pour des entiers, 344
de Morse, 1939
de Schreider, 327
de Slutsky, 2690
de transport, 1279
de Zorn, 171
des noyaux, 687
Fatou, 1341
Gauss
dans un anneau principal, 363
polynômes, 489
Grönwall, 2455, 2456
Hadamard, 1556
Schur complexe, 1033
Schur réel, 730
lemme de regroupement, 2643
length
in Weyl group, 3125
length of a root, 3108
lettres, 2847
levi subalgebra, 3064
Levi-Civita, 2084
Levi-Civita connection, 3407
Lévy
process, 4225, 4226
lexicographic ordering, 3109, 3119, 3122
libre
action, 336
partie, 385
partie d'un module, 273
librement engendré, 393
Lie
algebra
compact, 3211
derived, 3054
representation, 3043
solvable, 3055
algebroid, 3465
bialgebra, 3260
coalgebra, 3260
derivative
of a p-form, 3023
of a vector field, 3023

group, 3145
of transformation, 3451
solvable, 3055
groupoid, 3471
theorem, 3055, 3061
Lie algebra, 3049
Lie algebra of a Lie group, 3158
Lie subgroup, 3152
lift
horizontal, 3401
lifting property
covering space, 2934
projective module, 3241
light
cone, 3957
ray, 3956
light-like, 3941
Ligne d'univers, 2854
liminal C^*-algebra , 3873
limite, 745
d'ensembles, 1271
d'une fonction, 567
de fonctions holomorphes, 2185
de suite
espace topologique, 542
inférieure, 789
inversion, 1515, 1541, 2185, 2198
permutation
utilisation, 2680
suite dans $\mathbb{R}^m$, 599
suite numérique, 784
supérieure, 789
unicité, 568
limite à droite, 814, 1005
limite à gauche, 814
limite de suite, 542
limite épointée, 993
limite inductive, 2396
limite inférieure, 789
limite pointée, 993
linéaire
application, 394
linéairement indépendant
module, 271
linear
PDE, 3569
linear map between modules, 3852
Liouville measure
for $\mathrm{SL}(2,\mathbb{R})$, 4169
Lipschitz, 1558, 4082
localement, 1146
lipschitz
function, 4085
Lipschitzienne, 1145
liste d'indices, 45
little group, 3713
local symmetry (physics), 3716
localement
intégrable, 1405
locally
bounded representation, 3528
group type, 4183
integrable, 3587
isomorphic
Lie groups, 3150
trivial
principal bundle, 3382
locally convex, 2926
locally finite, 2923
log-concave, 1594
logarithme
complexe, 2137
dans $\mathbb{C}$, 2134
de matrice, 1504
sur les réels positifs, 1466
logarithme d'une application, 2144
loi
χ^2, 2715
binomiale
comportement asymptotique, 2717
conjointe, 2644
de Poisson, 2701
des grands nombres
forte, 2693
pour les chaines de Markov, 2802
processus de Poisson, 2842
utilisation, 2717, 2730
marginale, 2644
normale
vecteur gaussien, 2709
parente, 2737, 2738
parente d'un échantillon, 2740
réciprocité quadratique, 1808
sans mémoire, 2704
Student, 2715
loi d'une variable aléatoire, 2640
loi exponentielle, 2702
loi stationnaire, 2778
longueur
arc géométrique, 1976
d'un arc paramétré compact, 1960
d'un intervalle, 1321
d'une arrête, 1925
élément de, 1971
longueur d'arc, 1962
Lorentz

group, 3710, 4138
lorentz
algebra, 3066
Lorentz group, 3324
Lotka-Volterra, 2490
lowest weight
in $U_h\mathfrak{g}$-module, 3813
lucky number, 3139

M-matrice, 2615
M-matrice, 2615
maigre, 1283
maigre (ensemble), 630
majorant, 169, 311
essentiel, 2200
manifold, 2969
marginal
distribution, 4225
Markov
inégalité, 2693, 2743
Markov property, 3884
Markovian, 3895
completely
map on a C^*-algebra, 3895
semigroup on $L^2(A,\tau)$, 3900
quadratic functional, 3899
semigroup on $L^2(A,\tau)$, 3900
martingale, 2807
bornée dans $L^2(\Omega)$, 2808
mass hyperboloid, 4143
matrice, 403, 1256, 2068
compagnon, 758
creuse, 2585
cyclique, 692
d'une application linéaire, 406
de dilatation, 421
de permutation, 421
de similitude, 1141
de Sylvester, 663
de transvection, 420
dense, 2585
équivalence, 423
dans le groupe linéaire, 1261
hermitienne
racine carrée, 1249
jacobienne, 1119, 1131
normale, 1034
orthogonale, 419, 671
permutation
élémentaire, 2593
racine carrée, 1249
réductible, 2609
semblable, 1249
semblables, 1617
symétrique, 1617
réelle, 1615
trigonalisable, 1032
matrice d'une forme bilinéaire, 706
matrice de transition, 2776
matrice positive, 811
matrice primitive, 811
matrice stochastique, 2777
matrice-colonne, 403
matrice-ligne, 403
matrices
similitude, 423
matrices équivalentes, 723
matrices semblables, 723
matrix
element of a module, 3256
quantum group, 3796
maximal
norm, 3791
tensor product, 3791
maximale
partie orthonormale, 2105
maximum, 169, 186
global, 1580
local, 1580
maximum ideal, 3760
maximum local, 1580
Maxwell, 3715, 3719
measurable operator, 3568
measure, 3503
Lebesgue, 3504
positive, 3504
real, 3504
Mellin transform, 3661, 4224
méromorphe, 2171
mesurable
application, 1271
au sens de m^*, 1277
ensemble, 1269
fonction, 1278
Lebesgue, 1924
mesure
σ-finie, 1268
angle entre vecteurs, 1703
complexe, 1374
dans une carte, 1844
de Borel, 1303
de comptage, 1388, 1414
de Lebesgue, 1325
de Radon, 1303, 2166
extérieure, 1265
externe, 1924
finie, 1268

image, 1297
positive, 1268
probabilité, 2639
produit, 1379
régulière, 1303
extérieure, 1303
intérieure, 1303
sur un ensemble de parties, 1265
mesure σ-finie, 1266
mesure à densité, 1355
mesure absolument continue, 1355, 1359
mesure finie, 1266
mesure signée, 1268
mesures mutuellement singulières, 1360
méthode
des chemins, 1095
Newton, 2562
cas convexe, 2558
Méthode
de Newton, 2556
métrisable
espace vectoriel topologique, 603
mid-point map, 3650
minimal
norm, 3791
projection, 3839
minimum, 186
ensemble ordonné, 169
minimum local, 1580
Minkowski
space, 3942, 4138
Minkowsky product, 3323
minorant, 169, 311
mixed
boundary condition, 3664
state, 3783
modèle
échantillonnage, 2738, 2740
paramétrique, 2738
modèle statistique, 2737
modulaire (groupe), 2068
modular
condition, 3892
conjugation, 3853, 3892
function, 4130
group of automorphisms, 3892
isometry, 3887
operator, 3892
modular function, 2943
module, 3111
à gauche, 270
de continuité, 958
dual, 3594
finite type, 3234
highest weight, 3228
indecomposable, 3111
indécomposable, 274
irréductible, 274
of coefficients, 3594
over $U_h\mathfrak{g}$
finite dimensional, 3811
over a C^*-algebra, 3901
over a von Neumann algebra, 3852
over Hopf algebra, 3255
projective, 3241
right, 3234
semisimple, 3111
simple, 274, 3111
module d'un nombre complexe, 808
module produit, 271
moment, 2652
d'une force, 4587
fonction génératrice, 2679
momentum map, 3458
dual, 3458
monodromy action, 2935
monogène, 269
extension de corps, 502
monoidal
category, 2917
monoïde, 475
monôme, 374
monomorphism, 3231
monotone par morceaux, 1317
monotonie, 2495
Morita equivalence, 3698
morphism
of balanced product, 3238
of chain complexes, 3591
of cochain complex, 3593
of groupoid, 3471
of right module, 3234
of short exact sequences, 3592
morphisme
d'algèbres, 274
d'anneaux, 176
de corps, 263
espace vectoriel normé, 937
Frobenius, 276
morphisme de groupes, 175
morphisme de modules, 270
morphisme de produits tensoriels, 891
mot, 2847
longueur d'un mot, 2847
mot vide, 2847
nombre d'occurrences, 2847

mot non vide, 2847
moving frame, 3384
Moyal star product
on $\mathbb{R}^n$, 4117
on general symplectic manifold, 4119
moyenne
de Cesàro, 879
empirique, 2655
empirique d'un échantillon, 2740
quadratique, 2655
multiindice, 45, 356
multiplicateur
de Lagrange, 1585
multiplication, 188
multiplicative, 3737
multiplicité
racine d'un polynôme, 379
racine de $f(x) = 0$, 2555
valeur propre
algébrique, 1031
géométrique, 1031
multiplicity of a weight, 3136
musical isomorphism, 3021

nabla, 1127
nabla, 1127
natural topology, 3201, 3599
négatif, 302
négligeable
partie d'un espace mesuré, 1285
net, 2933
neutre
dans un groupe, 225
Newton
méthode, 2562
nilpotent, 231
Lie algebra, 3055
niveau de confiance, 2754
nombre
complexe
norme 1, 1800
de Fermat, 1831
dénormalisé, 2540
normal, 2729
normalisé, 2540
premier, 449, 475, 1484, 1792, 1800, 1804
dans leur ensemble, 248
théorème des deux carrés, 477
tours d'une courbe plane, 1996
nombre de montées, 2810
nombre premier
décomposition, 344
polynôme cyclotomique, 1797
nombres complexes, 318
non dénombrable, 210
non-rotating massive BTZ black hole, 3931
noncommutative
integral, 3568
vector bundle, 4086
nondegenerate
hermitian structure, 4087
representation, 3770
norm
maximal, 3791
norm of an operator, 3482
normal
j-subalgebra, 3657
(almost) subgroup, 3888
arc paramétré, 1977
element of an involutive algebra, 3745
elementary j-algebra, 3657
endomorphisme, 1034
family of functions, 2933
functional, 3849
map between von Neumann algebras, 3876
neighbourhood, 3178, 3707
nombre, 2729
operator, 3481, 3554
sous-groupe, 226
space, 2923
subgroup, 3050, 3153
normal j-algebra, 3657
normal extérieur
vecteur, 1875
normale
loi réduite, 2708
principale, 1987
normalisateur, 226
normalizer, 3070
norme, 577, 4488
d'algèbre, 846
d'application linéaire, 844
d'une application linéaire, 844
euclidienne
dans $\mathbb{R}^m$, 824
supremum, 838
sur $\mathbb{Z}[i\sqrt{5}]$, 366
normé
espace vectoriel, 579
norme maximum, 786
norme quotient, 647
norme suprémum, 1163
normes équivalentes, 840
noyau
application vers un groupe, 322
d'une forme bilinéaire, 706
Dirichlet, 2346

Fejér, 2346
vers un espace vectoriel, 396
noyau d'une forme bilinéaire, 700
null
groupoid, 3470
null vector, 3345
nulle part dense, 630

observation, 2639
observation, 2639
observer, 4139
odd spectral triple, 4081
one point compactification, 2929
opérateur
défini positif, 732
opérateur hermitien, 715
opérateur unitaire, 715
opérateurs
compatibles, 2117
operator
d'Alembert, 4146
differential, 3555
Dirac, 3700
elliptic, 3555
Laplace-Beltrami, 3420
process, 4226
pseudo-differential, 3559
symbol, 3534
opposés
chemins, 1968
opposite algebra, 3231, 3250
orbit in a groupoid, 3471
orbite
d'un point sous une action, 330
orbite d'un élément sous une permutation, 254
order
of a distribution, 3499
of a pole, 2930
ordering
linear partial relation, 3130
ordre, 169
bon ordre, 170
d'un élément, 251
d'un polynôme, 379
d'une matrice carrée, 403
dans un corps, 285
de convergence d'un schéma, 2631
distribution, 2410
partiel, 169
sur un anneau factoriel, 362
total, 169
ordre d'un groupe, 251
orientable, 3421
orientation, 667, 1846, 1869
orientation of a volume form, 3424
orientation on manifold, 3421
origine
abscisse curviligne, 1977
repère affine, 979
orthochronous, 3326
orthogonal, 713, 2100
coordonnées curviligne, 2082
famille de projecteurs, 274
matrice, 671
opérateur, 671
sous-espace, 432
orthogonal pour une forme bilinéaire, 700
orthonormé, 824
système, 2102
oscillation
d'une fonction, 1018
d'une fonction en un point, 1018
oscillatory integral, 3546
osculateur (cercle), 1992
outer
equivalent automorphism, 3891
outer automorphism, 4217
ouvert, 539, 675

Palatini action, 3915
Palatini action, 3915
parabolic
pde, 3571
parabolique
équation aux dérivées partielles, 2529
paracompact, 2923
parallèle
sous-espaces affines, 970
parallélogramme, 1849
paramétrage, 1968
normale, 1977
paramétrages
admissible, 1976
parametrix, 3559
parametrization, 3030
parity
of a representation, 3138
Parseval, 2110
partial
derivative, 3530
Fourier transform, 4167
isometry, 3476
order, 2931
partie
génératrice, 269, 385
régulière, 1218
totale, 2102
partie absolument convexe, 584

partie absorbante, 584
partie convexe, 576
partie discrète, 545
partie entière, 318
partie équilibrée, 584
partie fractionnaire, 318
partie inductive, 782
partie libre
module, 271
partie linéaire, 967
partie négative, 1361
partie nulle, 1361
partie positive, 1361
partie symétrique, 583
partition
dénombrable mesurable, 1308
of unity, 2932
subordinate, 2932
partition de l'unité, 2233
partition of unity, 2933, 2973, 3352
partition of unity subordinate, 2973
passage aux classes, 551
passive symbol, 4148
Past, 3989
past
horizon, 3931
interior region, 3931
Pauli matrices, 3300
pauli matrices, 3070, 3683
pavable, 1925
pavage du plan, 1734
pavé, 1924
Pearson
théorème, 2770
peigne de Dirac, 2432
pentagonal axiom, 2917
Perceval, 3501
période, 1075
permutation, 254
matrice, 421
permutation impaire, 260
permutation paire, 260
permuter
dérivée et intégrale
$\mathbb{R}^n$, 1553
dans $\mathbb{R}$, 1550
dans $\mathbb{R}$ avec les bornes, 1418
dérivée et limite, 1172
différentielle et intégrale
$\mathbb{R}^n$, 1554
intégrale
et série, 1414
limite et intégrale
convergence dominée, 1352
convergence monotone, 1340
espace mesuré, 1548, 1549
série entière et dérivation, 1447
série entière et intégration, 1450
somme et intégrale, 1341, 1425
permuter limite et intégrale, 1548
Peter-Weyl theorem, 3864
petit théorème de Fermat, 471
pgcd
calcul effectif, 341
dans un anneau intègre, 230
polynômes, 491
phase
function, 3544
space, 4144
physical space, 3926
pivotale, 2756
plan
projectif, 2025
tangent, 1131, 1225
plan affine, 1050
plan médiateur, 1054
plan vectoriel, 1050
Plancherel, 2110, 3501
plancherel measure, 3874
plongement, 1627
Poincaré
group, 3710, 4138
orthochroneous, 4145
Poincaré (demi-plan), 2068
point
d'équilibre
stable, 2486
pondéré, 972
spectrum, 3481
point adhérent, 544
point critique
définition, 1940
point d'accumulation, 547
point d'équilibre, 2486
point extrémal dans un convexe, 1254
point fixe, 2805
attractif, 1557
Brouwer, 1853
Picard, 1558
Schauder, 1855
point intérieur, 543
point isolé, 547
Poisson
formule sommatoire, 2374
generator, 4239
processus, 2828

Poisson structure, 3457
on an algebra, 3259
and Moyal product, 4165
on algebroid, 3468
on dual algebroid, 3467
on symplectic manifold, 3458
Poisson-Lie group, 3261
polaire
nombre complexes, 1667
polar
decomposition
in von Neumann algebra, 3825
operator on Hilbert space, 3489
polarization, 3793
good, 3636
identity, 3835
pole, 2930
at infinity, 2930
simple, 2930
pôle, 1753, 2169
polydisque, 2187
polygone, 1709
polygone convexe, 1710
polynôme
à plusieurs indéterminées, 666, 1811
alterné, 534
annulateur, 686, 690
caractéristique, 696, 1039
contenu, 376
cyclotomique, 1795
irréductibilité, 1797
propriétés, 1796
d'endomorphisme, 1249
décomposition de Dunford, 749
de Bernstein, 2732
irréductible
sur $\mathbb{F}_q$, 1823
minimal, 498
d'un élément d'une extension, 494
ponctuel, 690
racines, 536
semi-symétrique, 534
symétrique, 534, 536, 661, 666, 1811
élémentaire, 534, 537
trigonométrique, 2238
polynôme de Lagrange, 721
polynôme de plusieurs variables, 358
polynôme de Taylor, 1211
polynôme dérivé, 525
polynôme irréductible, 230
polynôme primitif, 376
polynôme scindé, 485
polynôme séparable, 529
polynôme sur un espace de Banach, 1205
polynôme trigonométrique, 2249
polynômes, 277
portée
mesure, 1361
poset, 2931
positif, 302
positive
completely
map between C^*-algebra, 3778
map in a C^*-algebra, 3895
semigroup on $L^2(A, \tau)$, 3900
conditionally, 4227
cone
standard, 3893
defined matrix, 2912
element
in a vector space, 3130
of a C^*-algebra, 3750
form on involutive algebra, 3769
map between C^*-algebra, 3764, 3895
measure, 3504
operator, 3484
sesquilinear form, 3480
vector, 3130
weight, 3136
positive root, 3079
potential, 3715
vector, 4097
potentiel, 1866, 1898
ppcm
dans un anneau intègre, 230
pre-C^*-module, 4087
pre-Hilbert, 3473
pre-inner product, 3762
prébase, 541
précision
simple, 2539, 2540
préhilbertien, 2095
premier
corps, 469
deux éléments d'un anneau principal, 248
deux polynômes entre eux, 376
idéal, 240
premier type
région solide, 1937
préserve l'orientation, 669, 1681
presque
nulle, 338
partout, 1271
surjective, 2295
primary constraints, 3464
primitif

élément d'un corps, 1804
élément d'une extension de corps, 502
racine, 1817
triplet pythagoricien, 371
primitive, 1063, 1896
de fonction continue, 1197
et intégrale, 1406
fonction, 1082
primitive et intégrale, 1402
primitive spectrum, 3773
principal
anneau, 240
bundle, 3379
idéal, 240
symbol, 3555, 3559, 3560
principe
prolongement analytique, 1634
zéros isolés, 1633
Principe de correspondance, 2869
probabilité
conditionnelle, 2656
probability
measure, 3832
problem
Dirichlet boundary condition, 3572
Neumann boundary condition, 3572
stationary boundary condition, 3572
well posed, 3572
problème
aux limites d'évolution, 2530
aux limites stationnaires, 2529
bien posé, 2530
limite de Dirichlet, 2529
limite de Von Neumann, 2529
problème de Cauchy, 2505
process
operator, 4226
processus
adapté à une filtration, 2807
arrêté, 2816
croissant prévisible, 2817
Galton-Watson, 2803
Poisson, 2828
sans mémoire, 2704
processus arrêté, 2817
processus de comptage, 2827
processus de Poisson, 2827
product
exterior, 3019
produit, 188
d'une mesure par une fonction, 1355
de convolution, 2227
et Fourier, 2371
de langages, 2848
de mots, 2847
distribution et fonction, 2408
espaces mesurés, 1381
mixte, 831, 834
scalaire
en général, 712
sur $\mathbb{M}(n,\mathbb{R})$, 1046
semi-direct, 337
tensoriel
de représentations, 1533
vectoriel, 830
produit de Cauchy, 1441, 1443
produit extérieur, 909
produit hermitien, 715
produit intérieur, 920
produit pseudo-scalaire, 713
produit remarquable, 2161
produit scalaire sur $\mathbb{R}^n$, 713
produit tensoriel, 891, 894
projecteur
dans un module, 274
projectif
complétion, 2028
droite, 2025, 2026
espace, 2025
groupe, 2032
hyperplan, 2026
plan, 2025
repère, 2033
sous-espace, 2025
projection
finite or infinite, 3842
in a C^*-algebra , 3752
in Hilbert space, 3476
initial and final, 3838
minimal, 3839
orthogonale, 2099
projection normale, 2277
projection orthogonale, 1686
projective
module, 3241
real space, 3008
representation, 3706
prolongement
analytique, 1634
utilisation, 2339
de fonctions, 1626
lemme de Borel, 1513
méromorphe de la fonction Γ, 2179
par continuité, 1008, 1009
dans $H^1(I)$, 2442
par densité, 1626

théorème de Hahn, 1294
prolongement de fonctionnelle linéaire, 2293
propriété d'intersection finie, 567
propriété d'intersection non vide, 561
propriété des valeurs intermédiaires, 795
propriété universelle, 891
pseudo-differential operator, 3559
pseudo-dimension, 2327
pseudo-Riemannian, 3719
norm, 3941
puissance
d'un langage, 2849
d'un mot, 2848
d'un point, 2019
d'un test, 2766
d'une inversion, 2052
pull back, 924
pull-back
of a k-form, 3022
of a vector field, 3022
pure state, 3783
purely infinite, 3846
push-forward
of a k-form, 3022
of a vector field, 3022

QED, 3716
QED, 3716
quadratic representation, 3606
quantization, 4153
quantized
algebra of regular functions, 3820
quantum
derivative, 3901
determinant, 3800
field theory, 3665
group
compact, 3795
matrix, 3796
minor, 3800
probability space, 4225
random variable, 4225
stochastic process, 4225
universal enveloping algebra, 3813
quasi
linear PDE, 3569
quasi-compact, 559
quasitriangular Hopf algebra, 3256
quaternion, 532, 3319
algebra, 3683
queue de liste d'indices, 45
quotient, 238, 375
dans une suite de composition, 326
de groupe, 325
de groupes, 448
quotient d'un chemin, 2315
quotient d'un espace vectoriel, 890
quotient topology, 2925

racine
carré
de matrice hermitienne, 1249
carré de matrice
hermitienne positive, 1249
d'un polynôme, 379
de l'unité, 772, 1712, 1797, 1800, 2068
primitive, 1789
utilisation, 1796
de polynôme, 509
multiple, 2555
primitive, 1817
simple, 2555
racine
carré
de matrice hermitienne, 1249
carré de matrice
hermitienne positive, 1249
d'un polynôme, 379
de l'unité, 772, 1712, 1797, 1800, 2068
primitive, 1789
utilisation, 1796
de polynôme, 509
multiple, 2555
primitive, 1817
simple, 2555
racine carrée, 807
racine de l'unité, 1787
racine multiple, 379
racine simple, 379
radical
of a Lie algebra, 3058
of a quadratic form, 3058
Radon-Nikodỳm theorem, 3830
raffinement, 1959
subdivision d'un pavé, 1926
rang, 398, 654, 742
classe d'équivalence, 423
diagonalisation, 731
différentielle, 1585
utilisation, 2101
rang d'une forme quadratique, 711
rang d'une matrice, 421
rank, 3015
of a complex Lie algebra, 3073
rank of a Lie algebra, 3106, 3128, 3134
rare, 1283
rationnels, 287
ray, 3701

rayon
de convergence, 1433
de courbure, 1987
de torsion, 1988
spectral, 846, 1191
real
J-real quadratic functional, 3899
form
of a vector space, 3119
of complex vector space, 3212
structure
on a spectral triple, 4081, 4084
réciproque
dérivabilité, 1077
recouvrement, 559
rectangle
produit de tribus, 1317
rectifiable, 1960
arc géométrique, 1976
récurrent
état, 2782
nul, 2782
point d'un système dynamique, 1305
positif, 2782
reduced abstract root system, 3088
reduced K group, 3910
reducible
principal bundle, 3387
representation of Clifford, 3685
réduction
d'endomorphisme, 749
Jordan, 762
réduction de Frobenius, 759
reduction of a principal bundle, 3387
reductive
AdS_n, 3947
homogeneous space, 3599
réel, 299
refinement of $\mathcal{W}$, 2967
réflexif, 2301
réflexion, 1722
dans $\mathbb{R}^2$, 1686
glissée, 1719
par rapport à un hyperplan, 1677
région
critique, 2765
de confiance exacte, 2756
de rejet, 2765
règle du produit nul, 231
regular, 2972, 3071, 3134, 3135
Lagrangian, 3468
left representation of a Hopf algebra, 3256
measure, 3505
quadratic form, 3899
representation, 4153
régularité
d'une mesure, 1303
extérieure de la mesure de Lebesgue, 1328
intérieure de la mesure de Lebesgue, 1330
régulier
arc, 1974
point d'un arc, 1974
régulier à droite, 230
régulière
surface, 1866
rejet
région dans une prise de décision, 2765
relation anormale, 2572
relation binaire, 169
relation d'équivalence, 174
relations
coefficient-racines, 537
de Chasles, 963, 1910
relativement compact, 564
relativistic invariance, 3720
relèvement, 1997
repère
affine, 979
cartésien
espace affine, 964
de Frenet, 1988
projectif, 2033
représentation, 433
de groupe fini
caractères de S_4, 1534
fidèle, 433
groupe diédral, 1783
irréductible, 1524
produit tensoriel, 1533
régulière gauche, 1529
virgule fixe, 2539
representation
of $U(\mathcal{A})$, 3149
adjoint, 3050, 3148
fundamental, 3812
isotropic, 3604
linear continuous, 3508
nondegenerate, 3770
of $\Gamma(p, q)$, 3675
of $\mathfrak{sl}(2, \mathbb{C})$, 3224
of $\mathfrak{su}(2)$, 3068
of $U_h\mathfrak{g}$, 3811
of a C^*-algebra, 3772
of a Lie group on a Hilbert, 3530
of a symmetric space, 3640
of Clifford algebra, 3670, 3672, 3685

of Hopf algebra, 3255
of involutive algebra, 3769
of Lie algebra, 3043
of Lie group, 3043
on Hilbert space, 3494
quadratic, 3606
regular
left on Hopf algebra, 3256
regular left, 3495, 4153
regular right, 3256
unitarisable, 3043
unitary, 3043, 3508
Wigner-Bargmann, 4145
Représentation
virgule flottante normalisée, 2539
representation of a Lie algebra, 3049
répulsif
point fixe, 1557
résidu
méthode itérative, 2607
residual spectrum, 3481
residue, 3568
résolvante, 2466
resolvent, 3484
reste, 238, 375
d'un développement limité, 1218
restricted root (real case), 3218
restriction d'une distribution, 2416
résultant, 663
utilisation, 666, 1891
ribbon category, 2920
Rieffel, 4119
Riemann
fonction, 1481, 4892
sphere, 2929
Riemannian symmetric pair, 3610
Riesz-Fisher theorem, 3474, 3508
right
bounded vector, 3853
invariant
vector field, 3157, 3597
support, 3838
right invariant
linear form on a Hopf algebra, 3255
right translation, 3157
right-invariant vector field, 3158
rigid
object in monoidal category, 2918
ring, 3231
risque
première espèce, 2766
quadratique, 2745
seconde espèce, 2766
root, 3074
abstract, 3088
inverse, 3103
reflexion, 3125
restricted, 3129
restricted (real case), 3218
simple, 3127
space, 3074
decomposition, 3129
in $\mathfrak{so}(2,1)$, 3617
vectors, 3123
rotation
en dimension 2, 1690
on functions, 3225
rotation d'angle θ, 1694
rotation-homothétie, 2051
rupture
corps, 510

safe region, 3931
safe region, 3931
scalar
curvature of a spectral triple, 4221
representation, 3344
scale invariance, 3566
Schatten-von Neumann ideal, 3480
schéma
consistant, 2630
schéma discret convergent, 2631
schéma numérique, 2630
Schrödinger, 2501
schur
module, 3270
Schur (théorème), 1527
Schürmann triple, 4227
Schwartz functions, 3501
second countable, 2921
secondary constraint, 3465
section, 3412
de graphe, 1084
local description, 3409
of an associated bundle, 3389
of associated bundle, 3413
of field of Hilbert space, 3864
of principal bundle, 3385
of vector bundle, 3352
propriété des, 1376
segment
dans $\mathbb{R}^p$, 792
dans un espace affine, 972
self polar cone, 3494
self-adjoint operator, 3476
semi direct product, 3156
semi-défini positif, 733

semi-direct product
of Lie algebras, 3143
of Lie groups, 3153, 3155
semi-linear
PDE, 3569
semi-simple
endomorphisme, 693
semi-symétrique
polynôme, 534
semicontinuous
lower, 3504
upper, 3504
semifinite
factor, 3846
functional on von Neumann algebra, 3849
seminorm, 2924
seminorme, 631
seminorme quotient, 647
seminormes séparantes, 644
semisimple, 3052
endomorphism, 3053, 3111
Lie algebra, 3062
module, 3111
separability
axiom, 2921
separable, 2921, 2925
Hilbert space, 3474
topological space, 2922
séparable, 2102
élément d'une extension, 530
espace topologique, 553
extension de corps, 530
polynôme non constant, 529
separate, 2924
separating vector, 3892
separation, 3473
sépare
les points, 1193
séquentiellement fermé, 602
série
dans un espace vectoriel normé, 856
de Fourier, 2353, 2374
utilisation, 2359
de puissance, 1433
donnant $(1-A)^{-1}$, 953
entière, 1433, 1515, 2374
Abel angulaire, 1886
processus de Markov, 2805
utilisation, 2680
fonctions, 1541, 2374
génératrice d'une suite, 1494
géométrique, 876
nombres, 1541
numérique, 1484, 1515
Riemann, 877
Taylor, 1495
série convergente, 856
série de Laurent, 2363
série de puissance, 1205
série de Taylor, 1211
série divergente, 856
série entière
fonctions holomorphes, 2153
série harmonique, 875
series of weight, 3138
Serre-Swan theorem, 4086
sesquilinéaire, 714
sesquilinear, 3473
form, 3480
sharp, 3021
shift operator, 3564
short exact sequence, 3592
sign of an operator, 3489
signature, 260
forme quadratique, 711
similitude, 1141
simple
abstract root, 3109
dimension of a spectral triple, 4081
extension de corps, 502
fonction, 1312
Lie algebra, 3062
module, 274, 3111
root, 3079, 3085, 3127
system, 3109
weight, 3136
simplex
singular, 3594
standard, 3594
simplification property, 3906
simply transitive action, 3597
singular
n-cycles, 3594
p-chain, 3594
p-simplex, 3594
chain
complex, 3594
homology with coefficients in $\mathbb{A}$, 3594
point in a black hole, 3959
submodule, 3345
value of an operator, 3478
Singular point, 3988
singularité, 2169
singularité effaçable, 2169
singularité isolée, 2169
singularity, 3959

sinus, 1637
hyperbolique, 1486
sinus cardinal, 1652, 1948
smooth
diffeomorphism, 3521
in spectral triple, 4082
map, 3530
open in $\mathbb{R}^d$, 3522
vector, 3529
smooth distribution, 3011
smooth manifold with boundary, 3438
smoothing operator, 3559
Sobolev space, 3524
soldering form, 3398
solfège, 1788
solution
générale, 2504
particulière, 2504
weak, 3582
solvable
Lie algebra, 3055
Lie group, 3055
somme
inférieure, 1931
partielle, 856
supérieure, 1931
somme directe, 434
somme directe (de représentations), 1523
somme directe topologique, 581
somme partielles
Abel angulaire, 1886
sommet, 2012
Sophie Germain, 471
sous arc, 1959
sous-additivité
sur algèbre de parties, 1266
sous-anneau, 236
sous-corps premier, 469
sous-espace
affine engendré par une partie, 970
caractéristique, 746
sous-espace topologique, 546
sous-groupe
caractéristique, 226
distingué, 449
dans le groupe alterné, 457
engendré, 268
normal, 325, 448
sous-groupe distingué, 226
sous-martingale, 2807
sous-module, 273
sous-module engendré, 271
space-like, 3941
space-time, 3719, 4138
spacial
bundle, 4149
symmetry, 4139
spatial rotation, 3327
spectral
action principle, 4219, 4220
radius, 3484
theorem
compact operators, 3478
selfadjoint operators, 3487
triple, 4080
spectre
matrice hermitienne, 828
matrice symétrique réelle, 731
spectre d'un endomorphisme, 685
spectrum, 3484
continuous, 3481
point, 3481
primitive, 3773
pure point, 3484
residual, 3481
sphère, 572
de Riemann, 2048
sphere, 3008
spin
group
on $\mathbb{R}^2$, 3684
on $\mathbb{R}^{1,3}$, 3674
manifold, 3687
norm, 3674
of representation of $\mathfrak{so}(3)$, 3070
representation
of $\mathfrak{so}(2l+1)$, 3278
of SU(2), 3322
structure, 3687, 3720
on $\mathbb{R}^2$, 3694
spinor, 3390
bundle, 3689
connection, 3690
field, 3689
gauged, 4099
representation, 3670
split
extension, 3143, 3153
real form, 3213
square integrable, 3898
square root of an operator, 3488
stabilisateur, 330
stabiliser of a flag, 3060
stabilité
d'un point d'équilibre, 2486
Lyapunov, 2486

stable, 2547
schéma numérique, 2633
stably isomorphic, 3908
standard
p-simplex, 3594
bialgebra structure, 3261
form
of a von Neumann algebra, 3893
form of a von Neumann algebra, 3877
star product
differential, 4117
formal, 4117
state
equivalent, 3790
mixed, 3783
on non-unital C^*-algebra, 3766
on unital C^*-algebra, 3761
one particle, 3711
pure, 3783
vector, 3774
stathme
sur $\mathbb{Z}[i]$, 474
stathme euclidien, 246
stationnaire
chaine de Markov, 2800
statistique, 2745
statistiques
descriptives, 2737
Stinespring theorem, 3778
strict
monoidal category, 2917
strictement
convexe
sur $\mathbb{R}^n$, 1597
strictement convexe, 1587
string
of roots, 3084
strong topology, 3492
strongly continuous
action, 3528
representation, 3530
structure
complexe, 1994
equations, 3394
group, 3379
space, 3737
structure d'anneau canonique, 228
structure réelle, 435
Student, 2715, 2758
subdivision, 1926
associée à une fonction, 1927
d'un intervalle, 1959
submanifold, 3000
submersion, 3002
subordonnée
norme, 844
subpotent, 203
subrepresentation, 3530
suite, 338
arithmético-géométrique, 877
de Cauchy, 604, 605
dans un corps, 285
de fonctions, 2198
théorème de Montel, 2338
de fonctions intégrables, 1548, 2185
de Jordan-Hölder, 326
définie par itération, 2562
équirépartie, 2349
critère de Weyl, 2349
exacte, 336
régularisante, 2376
suite de composition, 325
suite de variables aléatoires de Bernoulli, 2803
suites adjacentes, 788
support, 1927
d'une permutation, 254
distribution, 2410
famille d'éléments, 338
support of a distribution, 3502
supremum, 311
d'une suite d'ensembles, 1263
limit, 3490
norm, 2924
sur-martingale, 2807
surface paramétrée, 1866
surjection, 168, 591
surpotent, 203
sweedler notation, 3244
Sylow
p-Sylow, 440
Sylvester (matrice), 663
symbol
active, 4151
complete, 3555
Klein-Gordon
passive, 4148
of a pseudo-differential operator, 3559
of weight m, 4123
principal, 3555, 3570
symbole
de Legendre, 1805
symbole principal, 2519
symbolic vector, 4137
symétrique
polynôme, 534
symmetric

bimodule, 3901
derivation on a C^*-algebra, 3901
elements, 3612
Lie algebra, 3613
pair, 3610
product, 3615
sesquilinear form, 3480
space, 3600, 3606, 3612, 3640, 3948
Lie group, 3611
symplectic space, 3649
symmetry, 3649
in an affine space, 3608
of a root, 3085
of quantum system, 3702, 4247
symplectic
complement, 3457
gradient, 3458
group, 3457
Lie algebra, 3457, 4214
manifold, 3457
on R, 3651
symmetric space, 3633
triple, 3634
vector space, 3457
symplectomorphism, 3458
système
fondamental, 2463
orthonormé, 2102
trigonométrique, 2103, 2238
système trigonométrique, 2260

tame spectral triple, 4095
tame spectral triple, 4095
tangent
bundle, 2995
space, 2993
vecteur unitaire, 1987
tangent space, 2977
tangente, 1655, 1982
tangente hyperbolique, 1488
Tangente hyperbolique, 2863
taubérien, 1540
taux d'accroissement, 1588
Taylor
série entière, 1494
tempered
distribution, 3502
function, 3542
temps d'arrêt, 2810
temps de première atteinte, 2781
temps de retour, 2781
tenseur, 899
tenseur alterné, 906
tenseur symétrique, 906
tensor algebra, 3015
tensor bundle, 3360
tensor field, 3360
tensor product
of C^*-algebra, 3791
of algebra representations, 3259
of group representations, 3139
of Lie algebra representations, 3139
of matrices, 3139
of modules, 3237, 3238
of modules over Hopf algebra, 3255
of vector spaces, 3232
of von Neumann algebras, 3842
terminée
martingale, 2815
test, 2765
bilatéral, 2766
unilatéral, 2766
tétraèdre régulier, 1056
theorem
Cartan-Dieudonné, 3676
Engel, 3059
Gelfand, 3743
Hodge decomposition, 3420
Lie, 3061
spectral
compact operators, 3478
Stinespring, 3778
théorème
accroissements finis, 949
dans $\mathbb{R}$, 1078
forme générale, 1145
Ascoli, 2190
Banach-Steinhaus, 885
avec seminormes, 2190
base incomplète, 390
Beppo-Levi, 1340
Bézout
polynômes, 488
utilisation, 664
Bolzano-Weierstrass, 574
Borel-Cantelli, 2639
Borel-Lebesgue, 783
Brouwer
dimension 2, 2156
Carathéodory, 976
Cauchy
groupe, 348
Cauchy-Arzela, 1856
Cauchy-Lipschitz, 1561
Cayley-Hamilton, 698, 1248
central limite, 2695
processus de Poisson, 2843

Chevalley-Warning, 1811
chinois, 480
  anneau principal, 246
Cochran, 2742
Cochrane, 2743
convergence
  dominée de Lebesgue, 1352
  monotone, 1340
de Baire, 650
de Jordan, 2328
de représentation de Riesz, 2101
décomposition des noyaux
  et exponentielle de matrice, 1507
des deux carrés, 477
  version faible, 475
Dini, 1165
Dirichlet, 2346
  forme faible, 1800
Doob, 2816
du rang, 399
élément primitif, 1813
élément primitif, 531, 1814
extension d'isométrie, 1627
Fejér, 2348
fonction implicite dans $\mathbb{R}^n$, 1573
fonction implicite dans Banach, 1572
Fubini
  dans $\mathbb{R}^n$, 1415
  espace mesuré, 1410
  version compacte dans $\mathbb{R}^2$, 1900
Fubini-Tonelli, 1407
fuite des compacts, 2469
Gauss
  polynômes, 489
Gauss-Wantzel, 1831
Glivenko-Cantelli, 2750
Hahn-Banach, 2286
Hardy-Littlewood, 1541
Heine, 1023
incidence, 2026
inversion locale, 1570
  utilisation, 1585, 1939
isomorphisme
  second, 322
  troisième, 324
isomorphisme de Banach, 2189
Kronecker, 666
Lagrange, 325
Lie-Kolchin, 1036
Lotka-Volterra, 2490
Markov-Takutani, 1857
Montel, 2338
Pappus
  affine, 2030
  projectif, 2031
Pearson, 2770
petit de Fermat, 471
Picard, 1558
point fixe
  Brouwer, 2156
projection
  cas vectoriel, 2099
  partie fermée convexe, 2097
prolongement de Hahn, 1294
prolongement de Riemann, 2170
Radon-Nikodym, 1367
  complexe, 1375
Rolle, 1077
Rothstein-Trager, 1891
Runge, 2159
Schauder, 1855
Schur, 1527
spectral, 748
  autoadjoint, 823
  matrice symétrique, 731
  matrices normales, 1034
stabilité de Lyapunov, 2486
Stone-Weierstrass, 1194, 1196, 1197
Sylvester, 742
taubérien, 1540
taubérien faible, 1888
transfert, 2681
Tykhonov, 626
  dénombrable, 629
  fini, 628
valeurs intermédiaires, 807
Von Neumann, 1579
Wedderburn, 1800
Weierstrass, 574
théorème de Baire, 630
théorème de Brouwer, 1854
théorème de Cartan-Dieudonné, 1679
théorème des extrémums liés, 1585
théorème fondamental du calcul intégral, 1402
time
  orientable, 3719
  orientation, 3719, 3959
    on AdS_2, 3925
time orientation, 3941
time-like, 3941
topological
  conjugate, 3502
  dual, 3502
  group, 3159
  group of transformations, 3451
  Lie subgroup, 3186

subspace, 2921
topological embedding, 3010
topological group, 2928
topological manifold, 2961
with boundary, 3437
topologie, 539, 676
∗-faible, 648, 2407
p-adique, 961
discrète, 541
engendrée par une famille, 541
forte, 846
grossière, 541
métrique, 569
sur $\mathscr{D}(\Omega)$, 2397
sur $\mathscr{D}(K)$, 2397
sur $C^\infty(\Omega)$, 2397
sur dual topologique, 648
usuelle sur $\mathbb{R}^n$, 676
topologie d'une norme, 822
topologie des seminormes, 638
topologie faible, 2329
topologie induite, 545
topologie initiale, 549
topologie produit, 543
topologie quotient, 551
topology
hull-kernel, 3773
induced, 2921
Jacobson, 3773
natural on G/H, 3201
on submanifold, 3002
quotient, 2925, 3382
strong on $\mathcal{L}(\mathscr{H})$, 3476
ultraweak, 3492
uniform on $\mathcal{L}(\mathscr{H})$, 3476
weak on $\mathcal{L}(\mathscr{H})$, 3476
with seminorms, 2924
topology on manifold, 2971
tore, 4445
torsion, 1988
d'un groupe, 338
free, connection, 3394
of a connection, 3394
of exterior derivative, 3398
submodule, 3859
total subset in a pre-Hilbert, 3473
totale, 2102
totalizing vector, 3782
trace, 2448, 3565
ϵ-trace, 3849
class operator, 3479, 3492, 3563
dual de $\mathbb{M}(n,\mathbb{K})$, 721
endomorphisme, 724
in a balanced category, 2919
matrice, 724
of an operator, 3479, 3480, 3493, 3560, 3563
over C^*-algebra, 3792
over an algebra, 3231
over finite factor, 3852
produit scalaire sur $\mathbb{M}(n,\mathbb{R})$, 1046
unicité pour la propriété de trace, 722
vector, 3475
tracial functional
on von Neumann algebra, 3849
transcendant, 495
transformation
Fourier, 2374
gaussienne, 2587
group (topological), 3451
Lie group, 3451
transformée
de Cauchy, 2166
de Fourier, 2369, 2678
continuité, 2372
groupe abélien fini, 1521
Fourier
distribution tempérée, 2428
Laplace, 2679, 2680
transient
état, 2782
transition
probabilité, 2775
transition function, 3349
transitive, 3597
transitoire
état, 2782
translation
in affine space, 4138
invariant function on a group, 2928
operator, 3517
transposé, 670
transposée, 718
transposition, 259
transvection, 1240
transvection (matrice), 420
transvection group, 3613
transversale, 333
tribu, 1263
borélienne, 1279
de Baire, 1283
de Lebesgue, 1325
induite, 1265
produit, 1317
tribu de Lebesgue sur S^1, 1671
tribu engendrée, 1264
tribu engendrée par une application, 1264

trigonalisation
et polynôme caractéristique, 1032
simultanée, 1036
triple
elementary solvable, 3647
exact, 3647
symplectic, 3647
triplet
pythagoricien, 371
triplet naturel, 176
trivial
principal bundle, 3382, 3387, 3388
twist, 4105
based on $\mathcal{U}(\mathfrak{b})$, 4107
twisted
convolution, 4119, 4120
twisting map, 4122
type
fini
en algèbre, 533
espace vectoriel, 388
type I factor, 3840
type II factor, 3840
type III factor, 3840

ultraweak topology, 3492
ultraweak topology, 3492
unconditional convergence, 3474
unicité
des mesures, 1274
uniform continuity, 2928
uniformément continue, 630
uniformément convexe, 2276
uniformément intégrable, 2225
uniformisable topology, 2924
uniformly
continuous, 2922
unipotent, 231
unit
in a Banach algebra, 3483
unitaire
normale principale, 1987
unitarizable, 3254
unitary
endomorphism of module, 3240
group
of a module, 4091
of an algebra, 4096, 4101
induced representation, 3497
operator, 3476
representation, 3043, 3495, 3508
topological space, 2931
unitary operator, 3290
universal
R-matrix, 3256
algebra, 4065
compatible connection, 4091
connection, 4091
covering, 2934
deformation formula, 4108
enveloping algebra, 3149
problem, 3068
representation, 3776
upper limit, 2913
usinage, 2925

vacuum
of a GNS representation, 3771
vector, 3344
vacuum
of a GNS representation, 3771
vector, 3344
valeur
principale (distribution), 2426
propre, 685
valeur absolue, 318
p-adique, 961
valeur principale, 2135
valeur propre
d'une forme quadratique, 744
forme quadratique, 744
valeur singulière, 750
valuation
p-adique, 961
valuation d'un polynôme, 277
Vandermonde (déterminant), 660
vanishing cycle, 3266
variable
de décision, 2766
variable aléatoire, 2640
absolument continue, 2640
Bernoulli
marche aléatoire, 2790
utilisation, 2732
binomiale
utilisation, 2818
centrée, 2654
de Bernoulli
utilisation, 2818
de Rademacher, 2714
variance, 2655
empirique, 2655, 2741
empirique corrigée, 2741
vecteur gaussien, 2709
variation bornée, 1417
variation des constantes, 2459, 2464
variation totale, 1417
variational

formulation, 3582
variété, 1585, 1837
variété
orientée, 1847
variété orientable, 1847
vecteur
cyclique, 692
gaussien, 2709
propre, 685
unitaire normal, 1987
unitaire tangent, 1982
vecteur propre à gauche, 811
vector
form, 3770
potential, 4101
state, 3774
vector bundle, 3349
vector bundle isomorphism, 3349
vector field, 2993
vector-valued differential form, 3376
Verma module, 3226, 3229
vertical space, 3400
vielbein, 3699
Vitali (ensemble), 1331
vitesse d'un chemin, 1959
vitesse de convergence de suites, 1483
voisinage, 540, 676
volume
d'une région solide, 1937
element, 3685
région bornée dans $\mathbb{R}^3$, 1935
volume dans $\mathbb{R}^3$, 1848
volume form, 3424
von Neumann algebra, 3824
vraisemblance, 2748

weak integral, 3506
weak integral, 3506
weak topology, 3492
weight
dominant, 3113
for endomorphism, 3111
for representation, 3111
function, 4123
in $U_q\,\mathfrak{sl}(2,\mathbb{C})$, 3817
in a Dynkin diagram, 3096
space, 3111
weight of a vector, 3136
Weiner
constante, 2891
Weyl
group, 3085, 3086
product, 4119, 4170
Weyl group
abstract setting, 3088
width
of a representation, 3137
Wigner function, 4123
Wigner-Bargmann representation, 4145
WKB quantization, 4154
wodzicki residue, 3562
Wronskien, 2493

Yang-Baxter, 3800
Yang-Baxter, 3800
Yang-Baxter equation, 3256
Yang-Mills, 4098
field strength, 3415, 3718
Young
diagram, 3269
allowed, 3270
tableau, 3269
standard, 3269

Liste des notations

Algèbre

$V_1 \hat{\otimes} V_2$ product of $\mathbb{K}[\![h]\!]$-modules , page 3811

(α, β) Inner product on the dual $\mathfrak{h}^*$ of a Cartan algebra, page 3076

$(A, \Delta)^{cop}$ coopposite algebra, page 3250

$(A, \Delta)^{op}$ opposite algebra, page 3250

$[E, f]$ vector bundle over $X \times S^1$, page 3911

$[M : N]$ The index of two modules, page 3875

$[M_1 : M_0]$ Index of a subfactor., page 3877

$\mathrm{Ad}(u)\xi$ Adjoint representation of $U(\mathfrak{A})$ over $\mathscr{H}$, page 4219

bi-Diff(G) Bidifferential operators on G, page 3556

$\mathfrak{A}^\times$ The set of invertible elements of the algebra $\mathfrak{A}$; for example $\mathbb{R}^\times = \mathbb{R}\backslash\{0\}$, page 3676

$\mathfrak{A}^0$ Opposite algebra, page 3231

$\mathfrak{A}^e$ Enveloping algebra, page 3234

$\mathfrak{A}^N = \mathbb{C}^N \otimes_{\mathbb{C}} \mathfrak{A}$ A complex module over $\mathfrak{A}$, page 3234

$\mathscr{D}'$ distribution space, page 3499

χ A representation of $\Gamma(p, q)$, page 3674

$\mathrm{Cl}_F(E)$ Closure of the submodule E in F, page 3243

Δf Laplace operator, page 3524

Δ Basis of the roots, page 3127

Δ_p The standard p-simplex, page 3594

$\mathrm{Diff}^G(G)$ space of G-left invariant differential operators on G, page 3556

$\tilde{\rho}\colon (\mathbb{R}^2)^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ Spinor representation, page 3682

$\tilde{\rho}\colon (\mathbb{R}^{1+3})^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ Spinor representation, page 3670

$N\colon \Gamma(p, q) \to \Gamma(p, q)$ Spin norm, page 3674

$\mathbb{E}$ Spacial bundle, page 4149

$\mathbb{H}$ quaternionic algebra, page 3683

$\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ Algebra of endomorphism of the module $\mathcal{E}$, page 3240

$\mathrm{Fun}(M)$ Functions on the manifold M, page 4153

$\Gamma(E)$ Space of sections of the vector bundle E, page 3352

γ_i Abstract definition of Dirac matrices, page 3671

γ_i Explicit form of gamma matrices, page 3673

$\hat{G}$ Set of irreducible representations of the group G, page 3874

$\mathrm{Hom}(L, M)$ Space of linear maps from L to M, page 3502

$\hom_{\mathscr{C}}(A, B)$ The homset of arrows from A to B in the category $\mathscr{C}$, page 2915

$\mathrm{Ind}_{\mathfrak{A}}^{\mathrm{Cl}(V)}(E_1)$ Induced Clifford module, page 3685

$\mathrm{Inn}(\mathfrak{A})$ Group of inner automorphisms of an algebra, page 4217

$\mathrm{Irr}_{\mathfrak{g}}(\mathfrak{g})$ the unique cyclic highest weight $\mathfrak{g}$-module with highest weight λ., page 3230

ΛW Space of spinor representation, page 3670

$\Lambda W^{\pm}$ Spinor space, page 3682

$\langle M_0, P \rangle$ The von Neumann algebra generated by M_0 and P, page 3879

$\mathfrak{g}^x$ An algebra derived from $\mathfrak{g}$, page 3132

$\lim_{\rightarrow} A_i$ the direct limit of $(A_i)_{i \in I}$, page 2916

$\mathfrak{n}$ Restricted roots, page 3219

$\mathscr{P}$ Poincaré group, page 4138

$\mathfrak{M}^n(\mathfrak{A})$ C^*-algebra of matrices, page 3777

$\mathfrak{sp}$ Symplectic algebra, page 3457

$\mathcal{L}(H)$ Space of continuous endomorphisms of H, page 3782

$\mathcal{P}$ Set of pseudo-differential operator, page 3559

$\mathcal{Z}(\mathcal{G})$ Center of a Lie algebra, page 3644

$\mathscr{L}_M(\mathcal{E}, \mathcal{F})$ Space of linear maps between the modules $\mathcal{E}$ and $\mathcal{F}$, page 3852

$\Omega(M, V)$ V valued 1-forms, page 3009

$\Omega\mathfrak{A}$ Universal algebra of $\mathfrak{A}$, page 4065

$\Omega_D\mathfrak{A}$ Connes differential forms, page 4070

ω_i^j Connection form, page 3396

Op Weyl quantization, page 4122

$\mathrm{Out}(\mathfrak{A})$ Group of outer automorphisms of an algebra, page 4217

$\partial_i \sigma$ The face i of the singular p-simplex σ, page 3594

Φ, Φ^+ Root system, page 3127

Prim Primitive spectrum, page 3773

$\mathrm{Ray}\,\mathscr{H}$ Rays in a Hilbert space, page 3701

σ_ω Bargmann-Winger representation, page 4145

$\mathrm{SP}(V, \Omega)$ Symplectic group, page 3457

$\mathrm{Spec}(M, D)$ The spectrum of the Dirac operator over M, page 4221

$\mathcal{Q}$ A subgroup of $\mathcal{G}$, page 3677

$\mathcal{R}_z$ Trick to compute open orbits, page 3962

$\star^G$ Star product on G, page 4123

$\star^X$ Star product on X, page 4124

$\mathscr{S}$ Schwartz space, page 3501

$\mathscr{S}'$ Tempered distributions, page 3502

θ_ξ Coadjoint orbit of ξ, page 3459

$\tilde{K}(X)$., page 3908

$\mathrm{Tr}(A)$ Trace of the operator A, page 3563

A_α Gauge potentials, page 3412

AdS_n Anti de Sitter space, page 3947

$B_n(C)$ n-boundaries of a chain complex, page 3592

$C(X)$ Space of continuous functions, page 2924

$C\big(\mathrm{SU}_q(2)\big)$ the quantum matrix group, page 3797

C^r Functions of class C^r, page 3530

C^∞ Smooth functions, page 3530

$C^\infty(T)$ Space of smooth vectors of T, page 3530

$C_0(X)$ Space of continuous functions which decrease to zero at infinity in a certain sense, page 2924

$C_p(X, \mathbb{A})$ Set of singular p-chains in X with coefficients in $\mathbb{A}$, page 3594

$C_\flat(V, A)$ Set of continuous bounded map, page 3528

d_ω Exterior covariant derivative associated with the connection form ω, page 3402

$E_1 \sim E_2$ Vector bundle E_1 and E_2 are equivalent, page 3908

$E_1 \simeq_s E_1$ Vector bundles E_1 and E_2 are stably isomorphic, page 3908

E_{ij} Matrix full of zero's and 1 at position ij, page 3053

f_v $g \mapsto T(g)v$, page 3530

$F_{\mu\nu}$ Electromagnetic field strength, page 3716

FL Legendre mapping, fibre derivative of the Lagrangian L, page 3468

$G(\mathfrak{A})$ Set of invertible elements in $\mathfrak{A}$, page 3485

G/H as analytic manifold, page 3202

$H_2^k(\mathbb{R}^N)$ Sobolev space, page 3524

$H_n(C)$ the nth homology group of C, page 3592

$H_n(X, \mathbb{A})$ The singular homology of X with coefficients in $\mathbb{A}$, page 3594

J_μ Electromagnetic 4-current, page 3716

$K(X)$, page 3909

L' Topological dual of L, page 3502

L^* Linear maps from L to $\mathbb{C}$, page 3502

$L^2(M)$ Hilbert space of the standard form for a finite factor, page 3877

$L^2_\varphi(M)$ A completion of the von Neumann algebra M, page 3886

$M(X)$ Set of Borel measures on X, page 3505

$M_0(X)$ Set of compact supported Borel measures, page 3505

$Op(f)$ Active symbol of f, page 4151

$R(G)$ Dual of $C^\infty(G)$, page 3506

$T(f)$ Generators of Gårding space, page 3533

T_f Distribution defined by a function, page 3500

$U(\mathfrak{A})$ Unitary group of the algebra $\mathfrak{A}$, page 4096

$U(\mathcal{A})$ Universal enveloping algebra, page 3149

$U(\mathcal{E})$ Space of unitary endomorphism of the module $\mathcal{E}$, page 3240

$U(n,\theta)$ Rotation operator on functions, page 3225

$U_q\mathfrak{b}_\pm q$ Hopf subalgebra of $U_q\mathfrak{g}$, page 3814

$U_q\mathfrak{h}$ Hopf subalgebra of $U_q\mathfrak{g}$, page 3814

$V(\mathfrak{A})$ A class of projectors defined in order to introduce K-theory., page 3905

$V^{\mathbb{C}}$ Complexification of V, page 3212

$V_\lambda^{GL(D)}$ The Schur module, representation of $GL(D)$, page 3270

$v_{\tilde{P}}$ Velocity of observer $P \cdot \omega_0$, page 4140

$W, \underline{W}$ Totally isotropic subspace, page 3669

$W^{\mathbb{R}}$ Restriction of a complex vector spaces to $\mathbb{R}$, page 3212

X^* Image of a tensor in the universal enveloping algebra, page 3149

$x^\perp$ Space orthogonal to x, page 3675

$Z_n(C)$ n-cycles of a chain complex, page 3592

$[\mathbb{L} : \mathbb{K}]$ degré d'une extension de corps, page 493

(p) idéal engendré par p, page 236

ad adjoint action on Hopf algebra, page 3253

$\mathcal{L}(E,F)$ Ensemble des applications linéaires de E dans F, page 395

$\mathbb{C}[G]_q$ quantized algebra of regular functions, page 3820

$\mathbb{F}_p$ lorsque p est premier, page 469

$\mathbb{F}_{p^n}$ corps fini à p^n éléments, page 1802

$\mathbb{K}(A)$ corps contenant $\mathbb{K}$ et A, page 502

$\mathbb{K}[A]$ anneau contenant $\mathbb{K}$ et A, page 502

$\mathrm{Frac}(A)$ Le corps des fractions de l'anneau A, page 284

$\mathrm{Fun}(X,Y)$ les applications de X vers Y, page 228

$\mathrm{S}(E)$ Les opérateurs autoadjoints de E, page 715

∇f gradient de la fonction f, page 1119

$\pi_n(X,b)$ Homotopy group of X on the base point b, page 2934

$\mathtt{proj}_V$ projection de $V \times W$ sur V, page 598

$\mathrm{res}(P,Q)$ résultant des polynômes P et Q, page 663

$\mathrm{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A, page 385

$\sqrt{A}$ racine d'une matrice hermitienne positive, page 1249

$\theta_\alpha(P)$ la multiplicité de α par rapport à P, page 379

$A[X]$ tous les polynômes de degré fini à coefficients dans A, page 279

$A_n[X]$ les polynômes à coefficients dans A et de degré inférieur à n, page 374

$C(P)$ matrice compagnon, page 758

$c^{\pi}_{l,v}$ matrix element of a module, page 3256

$c^{V}_{l,v}$ matrix element of a module, page 3256

$D \mid P$ D divise P, page 375

$df_a(u)$ Application de la différentielle de f sur le vecteur u, page 1105

$E_\lambda(T)$ Espace propre de T, page 685

$f^{(n)}$ La n-ième dérivée de la fonction f, page 1213

$mat_{\mathcal{B}}(q)$ matrice de q dans la base $\mathcal{B}$, page 1618

$o(x)$ fonction tendant rapidement vers zéro, page 1215

$U(A)$ ensemble des inversibles, page 231

$U_h\mathfrak{g}$ a Hopf algebra, page 3810

U_n Le groupe des racines n^{e} de l'unité, page 1787

$U_q\mathfrak{g}$ quantum universal enveloping algebra, page 3813

Ensembles de matrices

$S^+(n,\mathbb{R})$ matrices symétriques définies positives, page 733

$S^{++}(n,\mathbb{R})$ matrices symétriques strictement définies positives, page 733

$\mathrm{Aut}(E)$ automorphisme de l'espace vectoriel E, page 395

$\mathcal{L}^0(E,F)$ applications linéaires bornées (continues), page 937

$\mathrm{End}(E)$ les endomorphismes de E, page 395

$\mathrm{O}(n,\mathbb{R})$ le groupe des matrices orthogonales, page 672

$\Omega(E)$ formes quadratiques non dégénérées, page 1614

$Q(E)$ formes quadratiques réelles sur E, page 699

$Q^+(E)$ formes quadratiques positives, page 1614

$Q^{++}(E)$ formes quadratiques strictement définies positives, page 1614

$S_n^{p,q}(\mathbb{R})$ matrices symétriques réelles de signature (p,q), page 1614

$\Delta(\mathfrak{A})$ structure space if the C^*-algebra $\mathfrak{A}$, page 3737

$\dim_M(E)$ The dimension of E as module over M, page 3857

$\mathscr{H}_1 = L^2(M)$ A completion of a von Neumann algebra, page 3854

$\mathscr{H}_\infty$ $\mathscr{H}_1 \otimes l^2$, page 3854

$\leqslant$ Ordering relation in C^*-algebra, page 3753

$L^2(A,\tau)$ Space associated to the trace τ in a C^*-algebra, page 3899

$L^\infty(A,\tau)$ von Neumann algebra associated to the trace τ on the C^*-algebra A, page 3899

$M_1 \bar{\otimes} M_2$ tensor product of von Neumann algebras, page 3842

$p \preceq q$ A partial ordering notion over the projections of a von Neumann algebra, page 3835

$p \sim q$ Equivalence of projectors, page 3837

$\beta(s)$ Vecteur unitaire de la binormale, page 1987

$\gamma \sim g$ Équivalence d'arcs paramétrés, page 1975

$\nu(s)$ Vecteur unitaire de la normale principale, page 1987

$c(s)$ rayon de courbure, page 1987

$t(s)$ Torsion, page 1988

$\mathfrak{A}_{\mathbb{R}}$ The set of selfadjoint elements in $\mathfrak{A}$, page 3750

$(\theta_\alpha)_i^j$ Matrix associated with a connection, page 3392

$[\omega \wedge \eta]$ Combination of the wedge product and the Lie bracket in the case of Lie algebra-valued forms, page 3379

Ad Adjoint representation, page 3050

$\mathrm{Ad}(P)$ Adjoint bundle of the principal bundle P, page 3392

Δ Laplace-Beltrami operator, page 3420

$\mathcal{D}$ Dirac operator, page 3692

$\gamma\colon \mathfrak{X}(M) \to \operatorname{End}\Gamma(\mathcal{S})$ A key ingredient for Dirac operator, page 3691

$\tilde{\nabla}\colon \mathfrak{X}(M) \times \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$ Covariant derivative for the spinor connection, page 3690

$\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ The algebra of ket-bras on the module $\mathfrak{A}$, page 4089

$\mathcal{C}(\mathcal{E})$ The space of compatible connections on the module $\mathcal{E}$, page 4090

$\mathcal{CC}(\mathcal{E})$ space of universal compatible connections, page 4091

$\Omega(M, E)$ the set of E-valued differential forms, page 3376

$C(\mathcal{E})$ the space of compatible connections on the hermitian module $\mathcal{E}$, page 4091

$F(\nabla)$ Curvature of a connection, page 4094

$\mathcal{B}$ Space of bounded operators on a Banach space, page 3482

$⨍ T$ The noncommutative integral, page 3568

Δf Laplace operator, page 3397

$\mathcal{F}$ Fourier transform, page 4167

∇f Gradient of the function f, page 3397

$\nabla \cdot X$ divergence of the vector field X, page 3397

$\mathcal{K}(\mathscr{H})$ The space of compact operators over the Hilbert space $\mathscr{H}$, page 3563

$\mathscr{L}(\mathscr{H})$ Space of trace class operators over $\mathscr{H}$, page 3493

$\mathscr{L}^p$ A functional space around the Dixmier trace, page 3567

$\Psi^n(M)$ Space of classical pseudo-differential operators on M, page 3560

$\rho(A)$ Resolvent of A, page 3484

$\sigma(A)$ Spectrum of A, page 3484

$\sigma_n(T)$ The sum of the n first characteristic values of the operator T, page 3476

$\mathrm{Spec}(A)$ Spectrum of A, page 3484

$\operatorname{Tr} P$ Trace of the operator P, page 3560

$C^n(G, \mathcal{G})$ Space of functions from $G \times \cdots \times G$ into $\mathcal{G} \otimes \mathcal{G}$, page 3262

$L_x^* P$ Left translated differential operator, page 3556

l_∞ Space of complex-valued bounded sequences, page 3564

$r(A)$ Spectral radius of A, page 3484

Géométrie

$P(E)$ l'espace projectif de E, page 2025

$\mathrm{Aut}(M, \omega, \nabla)$ Automorphism group of an affine symplectic manifold, page 4154

$\mathrm{Aut}\,\mathfrak{a}$ Group of automorphisms of $\mathfrak{a}$, page 3207

$\mathrm{Cl}(2)$ Clifford algebra of $\mathbb{R}^2$, page 3683

$\mathrm{Cl}(p, q)$ Clifford algebra of $\mathbb{R}^{1,3}$, page 3666

$\mathrm{Cl}(p, q)^{\pm}$ Grading of Clifford algebra, page 3674

$\Gamma(p, q)$ Clifford group, page 3674

$\mathrm{Int}(\mathfrak{a})$ Adjoint group of $\mathfrak{a}$, page 3207

$\mathfrak{Lie}(\Gamma(p, q)^+)$ Algèbre de $\Gamma(p, q)^+$, page 3681

$\mathcal{U}(\mathfrak{b}_1)^{\mathbb{B}_2}$ The elements of $\mathcal{U}(\mathfrak{b}_1)$ that are fixed by the action of $\mathbb{B}_2$., page 4110

$\mathcal{Z}(\mathfrak{h})$ the centralizer of $\mathfrak{h}$, page 3070

$\partial\mathfrak{a}$ The Lie algebra of $\mathrm{Aut}(\mathfrak{a})$, page 3207

$\mathfrak{gl}(\mathfrak{a})$ space of endomorphisms with usual bracket, page 3207

$\mathrm{SO}(p, q)$ Isometry group of $\mathbb{R}^{p,q}$, page 3619

$\mathrm{Spin}(p, q)$ Spin group of $\mathbb{R}^{1,3}$, page 3674

$\mathfrak{spin}(p, q)$ Lie algebra of the group $\mathrm{Spin}(p, q)$, page 3681

$\mathrm{Spin}(V)$ The spin group, page 3680

$\mathrm{Spin}^c(V)$ A group related to Spin, page 3680

$A \lhd B$ A is a normal subgroup of B, page 3155

$GL(\mathfrak{a})$ The group of nonsingular endomorphisms of $\mathfrak{a}$, page 3207

$W_\pm$ The set of elements of the form $(w, \pm w)$ in $W \oplus W$, page 3635

(α, β) inner product on the dual $\mathfrak{h}^*$., page 3076

$(x_0 : \ldots : x_n)$ coordonnées homogènes dans un espace projectif, page 2044

$\mathcal{C}\mathrm{onv}(A)$ enveloppe convexe, page 974

$\mathrm{PGL}(E)$ groupe projectif, page 2032

B^o orthogonal dans le dual, page 433

$G[\mathbb{C}]$ combinaisons d'éléments de G à coefficients dans $\mathbb{C}$, page 1524

$l(w)$ length in the Weyl group, page 3125

$P_1(\mathbb{C})$ sphère de Riemann, page 2048

t_α a basis of $\mathfrak{h}$, page 3076

Chaînes de Markov

$\pi(x)$ lié au temps de retour, page 2784

M' The commutant of M, page 3823

$p \vee q$ The smallest projection bigger than p and q, page 3835

$p \wedge q$ The biggest projection smaller than p and q, page 3835

Probabilités et statistique

$\sigma(f)$ La tribu engendrée par une variable aléatoire ou une application, page 1264

$a \wedge b$ $\min(a, b)$, page 2814

K_X matrice de covariance d'un vecteur gaussien, page 2709

$m(\mathcal{A})$ Ensemble des fonctions $\mathcal{A}$-mesurables, page 1281

K_m Ideal in $\mathrm{SU}_q(n)$, page 4231

$L(\Lambda)$ Module of highest weight Λ on $U_h\mathfrak{g}$, page 3812

Théorie des groupes

$(G/H)_g$ classes à gauche, page 331

$[a]_p$ ensemble des $a + kp$, page 353

$[G, G]$ groupe dérivé, page 321

$[g, h]$ commutateur dans un groupe, page 321

$\mathrm{Aff}(\mathbb{R}^n)$ Le groupe des applications affines bijectives de $\mathbb{R}^n$., page 985

gr_G groupe engendré, page 268

$\hat{G}$ groupe des caractères de G, page 1519

σ_x réflexion par rapport à x, page 1722

A_n groupe alterné, page 455

$D(G)$ groupe dérivé, page 321

D_n groupe diédral, page 1711

G^{ab} groupe abélianisé de G, page 322

$N \times_\phi H$ produit semi-direct, page 337

S_n le groupe symétrique, page 254

$N \lhd G$ Le sous-groupe N est normal dans G, page 226

Topologie et théorie des ensembles

$\mathrm{Adh}(A)$ adhérence de A, page 544

$\complement A$ Le complémentaire de l'ensemble A, page 171

$\mathrm{Diam}(A)$ Diamètre de la partie A, page 1022

∂A La frontière de l'ensemble A, page 679

$A \Delta B$ différence symétrique, page 173

A^c complémentaire de A, page 171

Analyse

$C^\infty(I, \mathscr{D}'(\mathbb{R}^d))$ fonctions à valeurs dans les distributions, page 2413

$(S, \hat{\mathcal{F}}, \hat{\mu})$ complété de l'espace mesuré $(S, \hat{\mathcal{F}}, \hat{\mu})$, page 1288

$\mathcal{L}(E, F)$ Les applications linéaires de E vers F, page 937

$\arg(z)$ La valeur principale de l'argument de $z \in \mathbb{C}$, page 2135

$\mathbb{D}_b$ l'ensemble des écritures décimales en base b, page 880

$\mathbb{R}$ l'ensemble des réels, page 299

$\mathbb{R}^+$ les réels positifs ou nuls, page 302

exp exponentielle, page 1470

$\mathscr{H}'$ dual, page 2101

$\mathrm{Isom}(X)$ Le groupe des isométries de X, page 601

$\liminf a_n$ limite inférieure, page 789

$\limsup a_n$ limite supérieure, page 789

$\mathcal{L}^p$ espace de Lebesgue, sans les classes, page 2195

$\mu \ll \nu$ La mesure μ est absolument continue par rapport à la mesure ν, page 1359

$\mu \perp \nu$ mesures mutuellement singulières, page 1360

μ^* La mesure extérieure associée à la mesure μ, page 1276

∂_z,$\partial_{\bar{z}}$ dérivées partielles d'une fonction complexe, page 2123

$\mathtt{proj}_K(x)$ projection orthogonale de x sur y, page 2099

$\sigma(\mathcal{D})$ tribu engendrée par $\mathcal{D}$, page 1264

$\mathscr{D}(\Omega)$ Les fonctions C^∞ à support compact sur Ω, page 2405

$\mathscr{S}'(\mathbb{R}^d)$ espace des distributions tempérées, page 2425

$A^2(\Omega)$ espace de Bergman, page 2340

$A^\perp$ orthogonal d'une partie., page 2100

$C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ Fonctions à valeurs dans les distributions., page 2501

$C_c(I)$ fonctions continues à support compact dans I, page 2224

$f \sim g$ fonctions ayant des limites équivalentes, page 1095

$H^1(\Omega)$ espace de Sobolev sur Ω, page 2444

$H^1(I)$ espace de Sobolev, page 2439

$H^m(M)$ espace de Sobolev, page 2446

$L^1_{loc}(\Omega)$ locally integrable functions on Ω, page 3587

$L^1_{loc}(I)$ fonctions intégrables sur les compacts de I, page 2439

L^p espace de Lebesgue avec les classes, page 2196

$M_i\varphi$ La fonction $x \mapsto x_i\varphi(x)$, page 2334

$S_n f$ somme partielle de série de Fourier, page 2263

Introduction in English

This document is a compilation of many stuff I wrote during my mathematical live. There are mainly three parts :

— Exercises for undergrad (in French).

— Theory for undergrad (in French, mainly devoted to l'agrégation)

— Theory I studied during my thesis (in English)

— Some research material (the BTZ part, in English)

Some parts of this text were written in 2003 wile others were written yesterday; don't expect a high quality everywhere. This document thus takes the point of view of the learner with some consequences. As far as I can judge my own work:

(1) There are *much* more details in the proofs in this text that what you can find in other textbooks.

(2) This is not a text in which you can get a deep understanding of what you are reading.

(3) There are (many) holes in the text.

There are still open questions in the sense that there are points I didn't understand when I wrote. I think that these points are clearly indicated with footnotes or special environment "Problem and misunderstanding". Let me know if you know some answers.

By the way, let me know any comment or suggestions.

Première partie

Le Frido

Chapitre 0

Introduction

0.1 Auteurs, contributeurs, sources et remerciements

Les remerciements, dans chaque catégorie, sont mis dans l'ordre chronologique approximatif.

0.1.1 Ceux qui ont travaillé sur le Frido

(1) Carlotta Donadello pour l'ensemble du cours de CTU de géométrie analytique 2010-2011. Une grosse partie de « analyse réelle » vient de là.

(2) Les étudiants de géométrie analytique en CTU 2010-2011 ont détecté d'innombrables coquilles. Les étudiants du cours présentiel de géométrie analytique 2011-2012 ont signalé un certain nombre d'incorrections dans les exercices et les corrigés. Les agrégatifs de Besançon 2011-2012 pour leurs plans et leurs développements.

(3) Lilian Besson pour m'avoir signalé un paquet de fautes, et quelques points pas clairs en statistiques.

(4) Plouf qui m'a signalé une coquille dans le fil la-selection-scientifique-de-la-semaine-numero-106.

(5) Benjamin de Block pour des coquilles et une mise au point sur les conventions à propos de $\mathbb{R}^+$ et $(\mathbb{R}^+)^*$.

(6) Olivier Garet pour avoir répondu à plein de questions de probabilités.

(7) François Gannaz pour de la relecture et une version plus claire de la preuve (et de l'énoncé) de la proposition 19.8.

(8) Danarmk pour des réponses à des questions dans les commentaires (allongement pour éviter un Overfull hbox) `http://linuxfr.org/nodes/110155/comments/1675589`. Et aussi pour une discussion à propos de la topologie sur $\mathscr{D}(\Omega)$.

(9) Cédric Boutilier pour des réponses à des questions de probabilité statistique. `https://github.com/LaurentClaessens/mazhe/issues/16`

(10) Remsirems pour des réponses à des questions d'analyse [1]

(11) Bertrand Desmons pour plusieurs patchs rendant plus clairs de nombreux passages sur les suites de Cauchy dans $\mathbb{Q}$.

(12) Anthony Ollivier pour m'avoir fait remarquer qu'il n'est pas vrai que $A[X]$ est euclidien lorsque A est intègre (contre-exemple : $A = \mathbb{Z}$). Ça fait une faute de moins dans le Frido.

(13) ybailly pour avoir détecté un bon nombre de coquilles dans la partie sur les ensembles de nombres.

(14) Éric Guirbal pour le remplacement de `frenchb` par `french`.

1. `http://linuxfr.org/nodes/110155/comments/1675813`

(15) cdrcprds pour une réponse à une question d'algèbre, démonstration à l'appui à propos de pgcd. Également pour sa relecture sans pitié de la partie sur la cardinalité (en particulier $A \approx A\backslash B$) et pour avoir pointé l'utilité du théorème de comparabilité cardinale.

(16) Antoine Bensalah pour avoir répondu à une question sur Lax-Milgram tout en même temps que pointé une erreur dans la démonstration et fourni l'exemple 25.62 sur l'optimalité de l'inégalité.

(17) Guillaume Deschamps pour ses remarques à propos du fait que le chapitre « constructions des ensembles » est très ardu.

(18) Guillaume Barriere pour sa relecture attentive jusqu'aux corps.

(19) Samy Clementz pour avoir découvert une faute dans la définition de mesure positive sur un espace mesurable.

(20) Sylvain Rousseau pour avoir clarifié une construction dans le théorème de Bower version C^∞.

(21) Maxmax pour des typos dans l'index thématique.

(22) Laurent Choulette pour une typo dans les propriétés du neutre d'un groupe.

(23) Pierre Lairez pour la démonstration du théorème d'inversion de limite et de dérivée 12.374 sans passer par les intégrales (et les lemmes correspondants à propos du module de continuité).

(24) Gregory Berhuy pour des réponses d'algèbres dans les catégories facile, moyen et difficile.

(25) Benoît Tran pour avoir signalé un paquet de typos dans la démonstration de l'ellipsoïde de John-Loewner et ses dépendances.

(26) Provatiscus pour avoir signalé un paquet de choses pas claires, et surtout pour avoir trouvé une faute dans la démonstration du fait qu'une fonction continue sur $\mathbb{Q}$ se prolonge en une fonction continue sur $\mathbb{R}$. Et pour cause : cet énoncé est faux. https://github.com/LaurentClaessens/mazhe/issues/124

(27) William pour l'environnement `example` qui gère correctement le triangle.

(28) Colin Pitrat pour de nombreuses remarques, typos et relecture de théorèmes.

(29) Bruno Turgeon pour une très belle moisson de fautes de frappe (euphémisme pour dire « mon ignorance crasse de l'orthographe »).

(30) Sacha Dhénin pour une belle quantité de fautes de frappes et pour avoir soulevé quelques points pas clairs (par exemple la définition du sous-groupe engendré dans le cas de la partie vide).

(31) Patrice Goyer pour m'avoir signalé quelques fautes et pas mal de points pas clairs dans les polynômes et dans la théorie généraliste des ensembles. Et pour m'avoir fait remarquer (deux fois) que mon script de déploiement ne marchait pas.

(32) Alain Vigne pour une quantité (presque) indénombrable de fautes d'orthographe et de mauvaises tournures de phrase dans les espaces vectoriels, la construction de nombres, la théorie des groupes et les anneaux. Il m'a également pointé quelques fautes et points vraiment pas clairs dans des démonstrations dont la correction a permis de bien améliorer la qualité du texte.

(33) Quentin Guyot pour une quantité de typos proche du nombre de Graham dont beaucoup me semblent impossible à détecter sans une lecture attentive ; (au moins) huit entrées dans l'erratum lui sont dues (au sens où c'est lui qui les a trouvées, pas qu'il les a causées).

(34) Jean Abou Samra pour la démonstration de la connextité par arcs C^1.

(35) jperon pour le classement de l'index en comptant les é comme des é. Voir [5].

0.1.2 Aide directe, mais pas volontairement sur le Frido

(1) Plein de monde pour diverses contributions à des énoncés d'exercices. Pierre Bieliavsky pour les énoncés d'analyse numérique (MAT1151 à Louvain la Neuve 2009-2010). Jonathan Di

Cosmo pour certaines corrections de MAT1151. François Lemeux, exercices sur les normes de matrices et correction de coquilles. Martin Meyer et Mustapha Mokhtar-Kharroubi pour certains exercices du cours *Outils mathématiques* (surtout ceux des DS et examens).

(2) Nicolas Richard et Ivik Swan pour les parties des exercices et rappels de calcul différentiel et intégral (Université libre de Bruxelles, 2003-2004) qui leurs reviennent.

(3) Carlotta Donadello pour la partie géométrie analytique : topologie dans $\mathbb{R}^n$, courbes, intégrales, limites. (Université de Franche-Comté 2010-2012)

(4) Le forum usenet de math, en particulier pour la construction des corps fini dans le fil « Vérifier qu'un polynôme est primitif » initié le 20 décembre 2011.

(5) Mihai Bostan nous a donné ses notes manuscrites de son cours présentiel de géométrie analytique 2009-2010. (Presque) Toute la structure du cours de géométrie analytique lui est due (qui est maintenant fondue un peu partout dans les chapitres d'analyse).

0.1.3 Des gens qui ont fait un travail qui m'a bien servi

(1) Arnaud Girand pour avoir mis ses développements bien faits en ligne. Une bonne vingtaine de résultats un peu partout dans ces notes viennent de lui.

(2) Le site `http://www.les-mathematiques.net` m'a donné les preuves de nombreux résultats.

(3) Pierre Monmarché pour son document en ligne tout plein de développements, et des réponses à des questions.

(4) Tous les contributeurs du Wikipédia francophone (et aussi un peu l'anglophone) doivent être remerciés. J'en ai pompé des quantités astronomiques ; des articles utilisés sont cités à divers endroits du texte, mais ce n'est absolument pas exhaustif.

(5) Les intervenants du fil « Antisymétrisation, alterné, déterminant et caractéristique » sur `les-mathematiques.net` m'ont bien aidé pour la section sur les déterminants 8.1 (bien qu'ils ne le savent pas).

(6) Xavier Mauquoy pour l'énoncé et la preuve du théorème 3.36.

(7) David Revoy pour les dessins de Pepper&Carrot de la couverture.

J'ai souvent donné entre parenthèses à côté des mots « théorème », « lemme » ou « proposition » une ou plusieurs références vers les sources de la preuve que je donne. Ce sont parfois des liens vers la bibliographie ; parfois aussi des liens hypertextes vers des sites, des blogs, etc. Tous ces gens ont fait du bon boulot. Sans toute cette « communauté », l'internet serait mort [2].

0.2 Originalité

Ces notes ne sont pas originales par leur contenu : ce sont toutes des choses qu'on trouve facilement sur internet ; je crois que la bibliographie est éloquente à ce sujet. Ce cours se distingue des autres sur les points suivants.

La longueur J'ai décidé de ne pas me soucier de la taille du fichier. Il fera cinq mille pages s'il le faut, mais il restera en un bloc. Étant donné qu'il n'existe qu'une seule mathématique, il ne m'a pas semblé intéressant de produire une division artificielle entre l'analyse, la géométrie ou l'algèbre. Tous les résultats d'une branche peuvent (et sont) être utilisés dans toutes les autres branches.

Dans cette optique, je me suis évertué à ne créer que des références « vers le haut ». À moins d'oubli de ma part [3], il n'y a aucun endroit du texte qui dépend d'un lemme démontré plus bas. Le fait qu'un théorème B soit plus bas qu'un théorème A signifie qu'on peut démontrer A sans savoir B.

2. Cette dernière phrase doit être comprise comme un appel à ne pas utiliser Moodle et autres iCampus pour diffuser vos cours de math, ou en tout cas pas comme moyen exclusif.

3. Par exemple pour les théorèmes pour lesquels je n'ai pas lu ni a fortiori écrit de preuves.

La licence Ce document est publié sous une licence libre. Elle vous donne explicitement le droit de copier, modifier et redistribuer.

Les mises à jour Ce document est régulièrement mis à jour. Des fautes d'orthographe sont corrigées (presque) chaque jour. Si vous me signalez une faute de mathématique, elle sera corrigée.

Transparence Je ne fais pas semblant que ces notes soient parfaites. Les points sur lesquels je ne suis pas sûr, les preuves que j'ai inventées moi-même sont clairement indiqués pour inciter le lecteur à redoubler de prudence. Une liste de questions à résoudre est incluse en la section 0.7. Voir 0.3 pour plus de détails.

0.3 Les choses qui doivent vous faire tiquer

SECooWVHBooCaYoXP

Un cours de math doit toujours être lu attentivement, surtout si vous avez l'intention de resservir à un jury le fruit de vos lectures. Dans ce livre, trois éléments doivent vous faire redoubler de prudence.

La référence [1] D'abord les références à [1] indiquent qu'une bonne partie de ce qui suit est de l'invention personnelle de l'auteur. Cela ne veut évidemment pas dire que c'est moi qui ai découvert le résultat. Ça veut dire que je n'ai pas trouvé le résultat ou certaines parties de la preuve.

Les notes en bas de page Certaines notes en bas de page sont écrites dans une fonte spéciale [4]. Elles indiquent des points sur lesquels je doute ou des étapes de calculs que je ne parviens pas à reproduire en suivant mes sources. Lorsque vous voyez une telle note, redoublez de prudence, allez voir la source, et écrivez-moi si vous pouvez résoudre le problème.

Les environnements dédiés Et enfin certains problèmes sont indiqués plus longuement dans un environnement dédié en petits caractères comme ceci :

¡¡ Avertissement/question au lecteur !! 0.1
Les choses écrites comme ceci sont des questions ou des éléments sur lesquels j'ai un doute. Lisez-les attentivement. Ces notes mentionnent des points que personnellement je n'oserais pas affirmer plein d'aplomb à un jury d'agrégation.

0.4 Quelques choix qui peuvent provoquer des quiproquos

Comme tout cours de mathématique, ce cours fait des choix qui sont parfois discutables. Voici quelques points sur lesquels les choix faits ici ne sont peut-être pas ceux fait par tout le monde. Ce sont donc des points sur lesquels vous devez faire attention pour éviter les quiproquos lors par exemple d'un oral dans un concours.

(1) Nous utilisons la définition « épointée » de limite d'une fonction en un point. Elle diffère de celle donnée par le ministère de l'enseignement en France. Si votre but est de passer un concours d'enseignement en France, vous devriez lire 12.2 ; dans tous les autres cas, la définition prise ici est celle qu'il vous faut.

(2) Un compact est une partie d'un espace topologique pour lequel tout recouvrement par des ouverts admet un sous-recouvrement fini. Le fait d'être séparable n'est pas inclus dans la définition de compact. De nombreux textes français incluent la séparabilité dans la compacité.

(3) Le logarithme sur $\mathbb{C}$ est une application $\ln : \mathbb{C}^* \to \mathbb{C}$ définie partout sauf en zéro. Elle n'est donc pas continue sur la fameuse demi-droite. À ne pas confondre avec une *détermination* du logarithme qui est par définition continue et donc non définie sur la demi-droite.
Cela est un choix très discutable. La raison de donner à la notation « ln » cette signification est simplement de suivre l'usage de Sage. Pour Sage, $\ln(-1)$ existe et vaut $i\pi$.
Voir les remarques 26.41.

4. Les notes comme celle-ci signifient qu'il y a certaines choses dont je ne suis pas sûr.

(4) Le mot « corps » n'implique pas la commutativité bien que tous les corps du Frido soient commutatifs, et nous n'utilisons pas la terminologie « anneau à division ». Voir la section -1.3 et la discussion 6.1.

0.5 Autres choix pas spécialement standards

Nous listons ici quelques choix qui n'induisent pas de différences ou d'incompatibilité avec les autres cours, mais qui doivent être compris et justifiés.

(1) Nous n'utilisons pas les notations $o(x)$ ou autres $O(N^2)$. D'abord parce que je n'ai jamais très bien compris comment elles fonctionnent, et ensuite (surtout) parce que ces notations induisent en erreur. Ce sont des notations qui cachent, sous des notations a peu près intuitive, l'utilisation de théorèmes pas simples.
Écrire
$$f(x) = P(x) + o(x^2), \tag{0.1}$$
c'est un peu comme quand on écrit (horreur !)
$$F(x) = \int f(x)dx + C. \tag{0.2}$$
Où est le x à droite ? Quel est le statut de C ?
Même chose pour la notation $f(x) = P(x) + o(x^2)$. Le x de $o(x^2)$ est-il le x qu'on a à gauche ? Si $g(x) = Q(x) + o(x^2)$, est-ce le même o que celui de f ?

(2) Nous allons être plus calme avec la notation $A[X]$ pour les polynômes sur l'anneau A, et encore moins $A[X_1, \dots, X_n]$ pour les polynômes de n variables. Au lieu de cela nous utilisons $\mathcal{P}(A)$ et $\mathcal{P}_n(A)$.
Est-ce que vous diriez que $A[X] = A[Y]$? Quelle est exactement la nature de X dans $P = X^2 + 1$ ou dans $P(X) = X^2 + 1$? Si $P \in A[X]$ vaut $P(X) = X^2$ et si $Q \in A[Y]$ vaut $Q(Y) = Y^2$, est-ce que vous oseriez écrire $P = Q$?

0.5.1 Mathématique intéressante

DEFooDABVooKdDyBw

Définition 0.2.
Une notion mathématique est ***intéressante*** *si elle permet de répondre à une question que l'on peut se poser sans connaitre la notion.*

Exemple 0.3.
Étant donné un segment dans le plan, quels sont les triangles rectangles dont ce segment est l'hypoténuse ?

Nous n'avons pas besoin de cercles pour poser cette question. Mais nous avons besoin de connaissances sur les cercles pour y répondre. Donc l'étude des cercles est intéressante. △

Exemple 0.4.
Comment fonctionne la gravitation ?

Cette question peut être posée sans connaitre de calcul tensoriel, d'équations différentielles ou d'intégrales. Et pourtant, tous ces concepts sont utiles pour y répondre. △

0.6 Sage est là pour vous aider

Il existe de nombreux logiciels de mathématique. Notre préféré est Sage pour une raison très précise : Sage est (en simplifiant) un module pour python. Donc quand on travaille en Sage, on dispose de tout Python. La syntaxe et la structure de Sage ne sont pas *ad hoc* pour faire des maths, et ce qu'on apprend en Sage peut être recyclé pour faire n'importe quoi : navigateur web, script de manipulation de texte, traitement d'image, réseau neuronaux, ...

Sage est un logiciel disponible pour l'épreuve de modélisation de l'agrégation de mathématique ; il y a donc de bonnes chances que vous en ayez l'usage.

0.6.1 Lancez-vous dans Sage

(1) Aller sur `http://www.sagemath.org`,

(2) créer un compte,

(3) créer des feuilles de calcul et s'amuser!!

Il y a beaucoup de documentation sur le site officiel [5], et nous vous conseillons particulièrement le livre [6].

Si vous comptez utiliser régulièrement ce logiciel, je vous recommande *chaudement* de l'installer sur votre ordinateur.

0.6.2 Exemples de ce que Sage peut faire pour vous

Ce livre est émaillé de petits bouts de code en Sage montrant ses différentes fonctionnalités là où nous en avons besoin [6]. Voici une liste (non exhaustive) de ce que Sage peut faire pour vous.

(1) Calculer des limites de fonctions, exemples 43.1 et 43.2.

(2) Tracer des graphes de fonctions, exemple 43.2.

(3) Tracer des courbes en trois dimensions, voir exemple 12.207. Notez que pour cela vous devez installer aussi le logiciel Jmol. Pour Ubuntu, c'est dans le paquet `icedtea6-plugin`.

(4) Calculer des dérivées partielles de fonctions à plusieurs variables, voir exemple 43.3.

(5) Résoudre des systèmes d'équations linéaires. Voir les exemples 43.4 et 43.5. Lire aussi la documentation.

(6) Tout savoir d'une forme quadratique, voir exemple 43.6.

(7) Calculer la matrice hessienne de fonctions de deux variables, déterminer les points critiques, déterminer le genre de la matrice hessienne aux points critiques et écrire les extrémums de la fonctions (sous réserve d'être capable de résoudre certaines équations), voir les exemples 43.7 et 43.8.

(8) Indiquer une infinité de solutions à une équation en utilisant des paramètres. L'exemple 43.9 montre ça avec une équation algébrique. Un exemple concernant des fonctions trigonométriques :

```
sage: solve(sin(x)/cos(x)==1,x,to_poly_solve=True)
[x == 1/4*pi + pi*z1]
sage: solve(sin(x)**2==cos(x)**2,x,to_poly_solve=True)
[sin(x) == cos(x), x == -1/4*pi + 2*pi*z86, x == 3/4*pi + 2*pi*z84]
```

Notez l'option `to_poly_solve=true` dans `solve`.

(9) Calculer des dérivées symboliquement, voir exemple 43.10.

(10) Calculer des approximations numériques comme dans l'exemple 43.11.

(11) Calculer dans un corps de polynômes modulo comme $\mathbb{F}_p[X]/P$ où P est un polynôme à coefficients dans $\mathbb{F}_p$. Voir l'exemple 19.62.

Sage peut en général faire tout ce que vous êtes capable de faire à l'entrée en master et probablement bien plus, à la notable exception des limites à deux variables.

Remarque 0.5.
Sage peut toutefois vous induire en erreur si vous n'y prenez pas garde parce qu'il sait des choses en mathématique que vous ne savez pas. Par conséquent il peut parfois vous donner des réponses (mathématiquement exactes) auxquelles vous ne vous attendez pas. Voir par exemple 15.146 pour le logarithme de nombres négatifs. Et aussi ceci :

5. http://www.sagemath.org
6. Soit un vrai besoin comme tracer un graphique en 3D, soit de la paresse comme calculer une grosse dérivée.

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: limit(1/x,x=0)
7 Infinity
8 sage: limit(1/x**2,x=0)
9 +Infinity
```

tex/sage/sageSnip017.sage

Sage fait une différence entre `Infinity` et `+Infinity` et donne

$$\lim_{x\to 0} \frac{1}{x} = \infty \tag{0.3}$$

ainsi que

$$\lim_{x\to 0} \frac{1}{x^2} = +\infty. \tag{0.4}$$

Voir aussi la compactification en un point d'Alexandroff 7.106.

0.7 Comment contribuer et aider ?

SecooCKWWooBFgnea

0.7.1 Des preuves qui manquent

Vous trouverez un peu partout des énoncés sans preuves. Certaines sont surement très faciles, et d'autres probablement assez compliquées. N'hésitez pas à rédiger une preuve et me l'envoyer.

Vous pouvez m'envoyer vos preuves sous forme de « c'est bien fait dans tel cours », avec une URL.

Ne me dites juste pas « c'est bien fait dans tel *livre* ». Je ne travaille pas à l'université, et je n'ai pas accès à une bibliothèque universitaire ; je n'ai donc pas réellement accès à ces fameux « livres » dont tout le monde parle.

0.7.2 Du texte qui manque

Vous remarquerez que de nombreuses pages du Frido sont des enchainements de théorèmes et démonstrations sans articulations. Autrement dit, il manque ce qu'à l'agrégation on dirait à l'oral quand on présente le plan. Si vous avez des idées de choses à ajouter ici où là, faites le moi savoir.

0.7.3 Des exemples qui manquent

Si vous connaissez de bons exemples, faites-le moi savoir.

0.7.4 Trucs de programmation et de LaTeX

(1) Comment faire en sorte que les mots commençant par « é » soient avec les « e » dans l'index, et non avant les « a » ? Il me faudrait un mécanisme plus automatique que faire `machin@truc`.
Une fonction en python qui prend en entrée le fichier bbl sous forme de string et qui ressort sous forme de string le fichier bbl modifié me convient.

(2) Il y a des problèmes dans la table des matières. « Table des matières », « Index », et « Liste des notations » ne pointent pas vers la bonne page.

(3) Écrire un script (en python ou autre) qui prend en argument deux numéros ou noms de chapitres et qui retourne l'ensemble des lignes du premier qui contient des `ref` ou `eqref` dont le label correspondant est dans le second.

Attention : il faut tenir compte de `input` de façon récursive.

Bonus : calculer le hash sha1 de chaque ligne du résultat et ne pas l'afficher si il se trouve dans la liste du fichier `commons.py`.

0.8 Les politiques éditoriales

Certaines parties de ce texte ne respectent pas les politiques éditoriales. Ce sont des erreurs de jeunesse, et j'en suis le premier triste.

0.8.1 Licence libre

Je crois que c'est clair.

0.8.2 pdflatex

Tout est compilable avec pdfLaTeX. Pas de pstricks, de psfrag ou de ps<quoiquecesoit>.

0.8.3 utf8

Je crois que c'est clair.

0.8.4 Notations

On essaie d'être cohérent dans les notations et les conventions. Pour la transformée de Fourier par exemple, je crois que la définition du produit scalaire dans L^2, des coefficients de Fourier, de la transformation et de la transformation inverse sont cohérents. Cela demande, lorsqu'on suit un livre qui ne suit pas les conventions utilisées ici, de convertir parfois massivement.

0.8.5 De la bibliographie

On évite d'écrire en haut de chapitre « les références pour ce chapitre sont ... ». Il est mieux d'écrire au niveau des théorèmes, entre parenthèses, les références.

Lorsqu'on écrit l'énoncé d'un théorème sans retranscrire la démonstration, il faut mettre une référence vers un document *en ligne* qui en contient la preuve. Il est vraiment fastidieux de chercher une preuve sur internet et de tomber sur des dizaines de documents qui donnent l'énoncé mais pas la preuve.

0.8.6 Faire des références à tout

Lorsqu'un utilise le théorème des accroissements finis, il ne faut pas écrire « d'après le théorème des accroissements finis, blablabla ». Il faut écrire un `\ref` explicite vers le résultat. Cela alourdit un peu le texte, mais lorsqu'on joue avec un texte de plus de 2000 pages, il est parfois laborieux de trouver le résultat qu'on cherche (surtout si il existe plusieurs versions d'un résultat et que l'on veut faire référence à une version particulière).

0.8.7 Des listes de liens internes

Le début du Frido contient une espèce d'index thématique. Il serait bon de l'étoffer.

0.8.8 Pas de références vers le futur

Dans le Frido, *aucune* preuve ne peut faire une référence vers un résultat prouvé plus bas. On n'utilise pas le théorème 10 dans la démonstration du théorème 7. Cela est une contrainte forte sur le découpage en chapitres et sur l'ordre de présentation des matières.

Il est bien entendu accepté et même encouragé de mettre des notes du type « Nous verrons plus loin un théorème qui ... ». Tant que ce théorème n'est pas *utilisé*, ça va.

Chapitre 1

Construction des ensembles de nombres

1.1 Quelques éléments sur les ensembles

1.1.1 Petit mot d'introduction

Le Frido n'est pas supposé être lu dans l'ordre de la première à la dernière page ; les matières y sont présentées dans l'ordre logique mathématique, et non dans l'ordre logique pédagogique, et encore moins par ordre de difficulté croissante.

1.1 (On saute la théorie des ensembles)**.**
En mathématique, si on lit une démonstration et que l'on veut vraiment tout justifier, et justifier toutes les étapes de tous les résultats utilisés, on tombe forcément un jour sur les axiomes.

Or l'axiomatique est un sujet particulièrement difficile. Nous n'allons donc pas « tout justifier » jusque là. Nous n'allons même pas préciser quel système d'axiome est utilisé. En particulier nous n'allons pas donner l'axiomatique des ensembles : nous allons supposer connus les ensembles et leurs principales propriétés.

Bref. Nous supposons avoir une théorie des ensembles qui tient la route. En particulier nous supposons connues les notions suivantes :

(1) ensemble vide,

(2) ensemble, appartenance, intersection, union,

(3) application entre deux ensembles, notation $f(x)$ pour désigner l'image de x par f,

(4) produit cartésien de plusieurs ensembles.

Ce sont toutes des choses dont la construction à partir des axiomes n'est en aucun cas évidente. En particulier, des « définitions » comme « l'intersection de deux ensembles est l'ensemble contenant exactement les éléments communs aux deux ensembles » ne sont pas correctes parce qu'elles passent à côté de l'existence et de l'unicité d'un tel ensemble.

1.2 (On saute la grammaire)**.**
Nous n'allons pas non plus formaliser la grammaire des expressions mathématiques [1]. Nous supposons que vous êtes capables de lire des expressions comme

$$x \in \mathbb{N} \Rightarrow \{a \geqslant x\} \text{ est infini.} \tag{1.1}$$

Tout cela pour dire que le Frido ne traitera que de la partie facile de la mathématique.

DefEnsemblesDisjoints

Définition 1.3.
*Deux ensembles A et B sont **disjoints** si leur intersection est vide [2] ; en d'autres termes, si il n'existe aucun élément commun à A et B.*

1. J'utilise ici des mots que je ne comprends pas, juste pour me donner l'air malin.
2. Remarquez que les mots « intersection » et « vide » sont de ceux que nous avons décidé de ne pas définir.

1.4.
Remarquez par exemple que la première phrase de l'article de Wikipédia sur la construction de $\mathbb{N}$ est « Partant de la théorie des ensembles, on identifie 0 à l'ensemble vide, puis on construit ... ». Il est bien précisé que l'on part d'une théorie des ensembles.

1.5.
La suite de ce chapitre sera essentiellement sans exemple parce qu'avant d'avoir construit les ensembles de nombres, je ne sais pas très bien quels exemples on peut donner de quoi que ce soit.

1.1.2 Injection, surjection, bijection

Définition 1.6.
Soient deux ensembles E et F. Une application $f : E \to F$ est

(1) ***surjective*** *si pour tout $y \in F$, il existe $x \in E$ tel que $y = f(x)$;*

(2) ***injective*** *si pour tout $y \in F$, il existe au plus un $x \in E$ tel que $y = f(x)$;*

(3) ***bijective*** *si elle est à la fois injective et surjective.*

La méthode la plus courante pour démontrer qu'une application $f : E \to F$ est injective est de considérer $x, y \in E$ tels que $f(x) = f(y)$, et de prouver à partir de là que $x = y$. Ou alors de supposer $x \neq y$ et d'obtenir une contradiction.

La technique de la contradiction est évidemment la plus courante lorsque l'égalité $f(x) = g(x)$ implique une équation faisant intervenir $1/(x - y)$.

LEMooWBYSooFqyqQP

Lemme 1.7.
Soient deux ensembles A et B ainsi qu'une application $f : A \to B$. Nous supposons qu'il existe une application $g : B \to A$ telle que $f \circ g = \mathrm{Id}_B$ et $g \circ f = \mathrm{Id}_A$.

Alors f est une bijection.

Démonstration. En deux parties.

(1) **Injection** Supposons que $f(a) = f(b)$. Alors en appliquant g des deux côtés, et en utilisant le fait que $g \circ f = \mathrm{Id}_A$, nous trouvons $a = b$.

(2) **Surjection** Soit $x \in B$. Posons $a = g(x)$. Alors, en utilisant le fait que $f \circ g = \mathrm{Id}_B$ nous avons
$$f(a) = (f \circ g)(x) = x. \tag{1.2}$$
Donc x est dans l'image de f et f est surjective.

□

PROPooBXVSooZXmwKC

Proposition 1.8 ([1]).
Si $f : A \to B$ est injective, et si X et Y sont des parties de A, alors $f(X \cap Y) = f(X) \cap f(Y)$.

Démonstration. En deux inclusions.

(1) **Première inclusion** Si $\xi \in f(X \cap Y)$, alors il existe $s \in X \cap Y$ tel que $\xi = f(s)$. Étant donné que $s \in X$, nous avons $\xi = f(s) \in f(X)$. Et comme $s \in Y$ nous avons aussi $\xi = f(s) \in f(Y)$. Au final $\xi = f(x) \in f(X) \cap f(Y)$.

(2) **Deuxième inclusion** Si $\xi \in f(X) \cap f(Y)$, il existe $x \in X$ et $y \in Y$ tels que $\xi = f(x)$ et $\xi = f(y)$. Par injectivité de f, nous avons $x = y$ et donc $\xi \in f(X \cap Y)$.

□

1.1.3 Ensemble ordonné

NORooLMBYooYjUoju

1.9.
L'**axiome du choix** que nous acceptons peut s'énoncer comme ceci[7] : Étant donné un ensemble X d'ensembles non vides, il existe une fonction définie sur X, appelée fonction de choix, qui à chacun d'entre eux associe un de ses éléments.

DEFooRFVTooUUuFuE

Définition 1.10 ([8])**.**
Une ***relation binaire*** *d'un ensemble E vers un ensemble F est une partie de $E \times F$.*

Si G est une relation binaire entre E et F, nous notons $x\mathcal{R}_G y$ et nous disons que x est en relation avec y.

DefooFLYOooRaGYRk

Définition 1.11.
Une ***relation d'ordre*** *sur un ensemble E est une relation binaire*[3] *(notée $\leqslant$) sur E telle que pour tous $x, y, z \in E$,*

réflexivité *: $x \leqslant x$*

antisymétrie *: $x \leqslant y$ et $y \leqslant x$ implique $x = y$*

transitivité *: $x \leqslant y$ et $y \leqslant z$ implique $x \leqslant z$.*

Pour suivre les notations de la définition 1.10, la partie G de $E \times E$ est une relation d'ordre lorsque

(1) $(x, x) \in G$ pour tout $x \in G$,

(2) Si $(x, y) \in G$ et $(y, x) \in G$, alors $x = y$,

(3) Si $(x, y) \in G$ et $(y, z) \in G$, alors $(x, z) \in G$.

Dans la suite nous n'allons plus écrire de relations binaires en détaillant l'ensemble sous-jacent.

Lorsque nous avons un ensemble E et une relation d'ordre $\leqslant$ sur E, nous disons que le couple $(E, \leqslant)$ est un ***ensemble ordonné****.*

DEFooVGYQooUhUZGr

Définition 1.12.
Un ensemble ordonné est ***totalement ordonné*** *si deux éléments sont toujours comparables : si $x, y \in E$ alors nous avons soit $x \leqslant y$ soit $y \leqslant x$. Si les éléments ne sont pas tous comparables, nous disons que l'ordre est* ***partiel****.*

DEFooBZNRooYRPGme

Définition 1.13 ([9])**.**
Soit un ensemble ordonné $(E, \leqslant)$. Un élément $M \in E$ est un ***élément maximal*** *de E si pour tout $x \in E$ tel que $M \leqslant x$, nous avons $M \leqslant x \Rightarrow x = M$.*

Nous disons que $m \in E$ est un ***élément minimal*** *si pour tout $x \in E$, nous avons $x \leqslant m \Rightarrow x = m$.*

DEFooDNWRooTiMAzK

Définition 1.14.
Soit un ensemble ordonné $(E, \leqslant)$ et une partie A de E. Nous disons que $m \in A$ est un ***minimum*** *de A si pour tout $x \in A$, l'élément m est comparable à x et $m \leqslant x$.*

Un élément $p \in E$ est un ***minorant*** *de A si pour tout $a \in A$, les éléments p et a sont comparables et $p \leqslant a$.*

Les notions de ***maximum*** *et de* ***majorant*** *sont définies de façon analogue.*

LEMooPCRFooXRGrUr

Lemme 1.15.
Si E est un ensemble ordonné et si A est une partie finie totalement ordonnée de E, alors A possède un unique minimum et un unique maximum.

NORMooVHIBooJAOsou

1.16.
Notons qu'il n'est pas demandé à un élément maximal[4] d'être comparable à tous les autres éléments. Si $(E, \leqslant)$ n'est pas totalement ordonné, un élément maximal peut ne majorer qu'une partie de E.

Un élément maximal est plus grand que tous les éléments avec lesquels il est comparable.

Dans un ensemble totalement ordonné, les notions d'élément maximal de E et de maximum[5] de E coïncident ; dans le cas d'un ordre partiel, ces notions sont distinctes.

3. Définition 1.10.
4. Définition 1.13
5. Définition 1.14.

Il se peut que nous parlions d'un « maximum » ou « élément maximum » au lieu d'un « élément maximal », en particulier en utilisant le lemme de Zorn. Si vous voyez de telles choses, n'hésitez pas à me le dire.

Par exemple, quand on applique le lemme de Zorn pour démontrer l'existence d'une base dans un espace vectoriel de dimension infinie, on obtient une famille libre qui est maximale, c'est-à-dire qui est un élément maximal dans l'ensemble, ordonné par inclusion, des familles libres de l'espace vectoriel. C'est un élément maximal, et non un maximum, car l'ensemble des familles libres d'un espace vectoriel non réduit à $\{0\}$ n'admet pas de maximum (pour l'inclusion). Voir 4.25 et 4.26.

Lorsqu'une partie possède un minimum, ce dernier est nommé le « plus petit élément » de la partie. Attention : il n'en existe pas toujours. D'innombrables exemples pourront être vus lorsque nous aurons construit $\mathbb{Q}$ et $\mathbb{R}$. Typiquement les intervalles du type $]a, b[$.

Exemple 1.17 ([1])**.**
Soit un ensemble E ainsi que $a \neq b$ dans E. Nous considérons les parties

$$A = \{P \subset E \text{ tel que } a \in P, b \notin P\} \tag{1.3a}$$
$$B = \{P \subset E \text{ tel que } b \in P, a \notin P\}. \tag{1.3b}$$

Et enfin nous considérons l'ensemble $F = A \cup B$, c'est-à-dire l'ensemble des parties de E qui contiennent soit a soit b mais pas les deux. Nous ordonnons partiellement F par l'inclusion.

Dans $(F, \subset)$, l'élément $E\backslash\{b\}$ est un élément maximal, mais pas un majorant. △

DEFooLJEAooBLGsiS

Définition 1.18.
Un ensemble ordonné est ***bien ordonné*** *si toute partie non vide possède un plus petit élément.*

Autrement dit, l'ensemble ordonné E est bien ordonné si pour toute partie non vide A, il existe $x \in A$ tel que $x \leqslant y$ pour tout $y \in A$.

1.19.
Quelques remarques.

(1) L'inégalité stricte (définie par : $x < y$ si et seulement si $x \leqslant y$ et $x \neq y$) n'est pas une relation d'ordre parce qu'elle n'est pas réflexive.

(2) Nous verrons dans la remarque 1.442 que l'intervalle $[-1, 1]$ dans $\mathbb{R}$ n'est pas bien ordonné.

(3) Un ensemble bien ordonné est forcément totalement ordonné parce que toutes les parties de la forme $\{x, y\}$ possèdent un minimum. Par conséquent x et y doivent être comparables : $x \leqslant y$ ou $y \leqslant x$.

Exemple 1.20.
Si E est un ensemble, l'inclusion est un ordre sur l'ensemble des parties de E, mais pas un ordre total parce que si X, Y sont des parties de E, alors nous n'avons pas automatiquement soit $X \subset Y$ soit $Y \subset X$. △

La notion d'ordre permet d'introduire la notion d'intervalle.

DefEYAooMYYTz

Définition 1.21.
Soit un ensemble totalement ordonné $(E, \leqslant)$. Un ***intervalle*** *de E est une partie I telle que tout élément compris entre deux éléments de I soit dans I. En language mathématique, la partie I de E est un intervalle si*

$$\forall a, b \in I, (a \leqslant x \leqslant b) \Rightarrow x \in I.$$

1.1.4 Lemme de Zorn

Nous admettons l'axiome du choix qui s'énonce de la façon suivante[10] :

Pour tout ensemble X d'ensembles non vides, il existe une fonction définie sur X, appelée fonction de choix, qui à chaque ensemble A appartenant à X associe un élément de cet ensemble A.

DefGHDfyyz

Définition 1.22 (Ensemble inductif[11])**.**
Un ensemble est ***inductif*** *si toute partie totalement ordonnée admet un majorant.*

LemUEGjJBc

Lemme 1.23 (Lemme de Zorn[12])**.**
Tout ensemble ordonné inductif non vide admet au moins un élément maximal.

PROPooFOETooWYLOeq

Proposition 1.24 ([13])**.**
Soient un ensemble inductif $(E, \leqslant)$ et $b \in E$. Il existe un élément maximal[6] *$m \in E$ tel que $b \leqslant m$.*

Démonstration. En plusieurs parties.

(1) **Un ensemble** Nous considérons

$$E_b = \{x \in E \text{ tel que } b \leqslant x\}. \tag{1.4}$$

(2) **E_b est inductif** Soit une partie non vide totalement ordonnée A de E_b. Puisque E est inductif, la partie A admet un majorant $m \in E$.
Comme A est non vide, nous pouvons considérer $x \in A$. Nous avons $b \leqslant x$ parce que $x \in A \subset E_b$. Mais comme m est un majorant de A, $x \leqslant m$. Bref, nous avons les inégalités

$$b \leqslant x \leqslant m. \tag{1.5}$$

Donc $m \in E_b$. Nous avons prouvé que m est un majorant de A contenu dans E_b. Donc E_b est inductif.

(3) **Zorn** Puisque $(E_b, \leqslant)$ est inductif, le lemme de Zorn 1.23 nous indique que E_b a un élément maximal. Nous le notons m.

(4) **m est maximal dans E** Supposons avoir un élément $a \in E$ tel que $m \leqslant a$. Nous avons

$$b \leqslant m \leqslant a, \tag{1.6}$$

et donc $a \in E_b$. Mais m est maximal dans E_b, donc $a = m$.

□

1.1.5 Complémentaire

AppComplement

Définition 1.25.
Soit E, un ensemble et A, une partie de E (c'est-à-dire un sous-ensemble de E). Le ***complémentaire*** *de l'ensemble A, dans E, noté $\complement A$ est l'ensemble des éléments de E qui ne font pas partie de A :*

$$\complement A = E \backslash A = \{x \in E \text{ tel que } x \notin A\}. \tag{1.7}$$

Nous allons aussi régulièrement noter le complémentaire de A par A^c.

1.1.6 Quelques relations ensemblistes

LEMooHRKAooRskzQL

Lemme 1.26 (Quelques relations ensemblistes)**.**
Soient $A, B, C \subset X$. Nous avons

(1) $X \backslash (A \cap B) = (X \backslash A) \cup (X \backslash B)$.

ITEMooQCGUooKnWfBo

(2) $X \backslash (A \cup B) = (X \backslash A) \cap (X \backslash B)$.

6. Définition 1.13.

ITEMooXWKCooUASxlh

(3) $A \cap (B\backslash C) = (A \cap B)\backslash C$.

LEMooBGQRooCVLGsh

Lemme 1.27 ([1])**.**
Si $A \subset B$ et si $f : B \to B$ est injective, alors

EQooHLEHooADUbJQ

$$f(A) = f(B)\backslash f(B\backslash A). \tag{1.8}$$

Démonstration. Les deux inclusions.

(1) $\subseteq$ Soit $a \in A$. Alors $a \in A \subset B$, de telle sorte que $f(a) \in f(B)$. Mais a n'est pas dans $B\backslash A$, et comme f est injective, $f(A) \cap f(B\backslash A) = \varnothing$. Donc, vu que $f(a) \in f(A)$, nous avons $f(a) \notin f(B\backslash A)$.

(2) **Dans l'autre sens** Suppose $y \in f(B)\backslash f(B\backslash A)$. Vu que $y \in f(B)$, il existe $b \in B$ tel que $y = f(b)$. Mais comme $y \notin f(B\backslash A)$, nous avons $b \notin B\backslash A$. Autrement dit, $b \in A$ et donc $f(b) \in f(A)$.

□

LEMooHPUNooPmViwi

Lemme 1.28 ([14])**.**
Pour tout ensembles A, B, C nous avons

(1) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

ITEMooSIDWooNrDjoU

(2) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

LemPropsComplement

Lemme 1.29.
Quelques propriétés à propos des complémentaires. Si E est un ensemble et si A et B sont des sous-ensembles de E, nous avons

ITEMooWIYBooKMLVnZ

(1) $\complement\complement A = A$, *en d'autres termes,* $E\backslash(E\backslash A) = A$,

ItemLemPropComplementiii

(2) $A\backslash B = A \cap \complement B$.

ITEMooNHDUooWtURqQ

(3) $(A\backslash B)^c = A^c \cup B$.

ITEMooTBWKooTChOmC

(4) $A^c\backslash B^c = B\backslash A$.

Démonstration. Plusieurs points.

(1) **Pour (1)**

(2) **Pour (2)**

(3) **Pour (3)** Il faut le faire en deux inclusions.

(3a) $(A\backslash B)^c \subset A^c \cup B$ Supposons que $x \in (A\backslash B)^c$. Si $x \in A^c$, c'est bon. Supposons que x n'est pas dans A^c, et montrons que $x \in B$. Le fait que x ne soit pas dans A^c signifie que $x \in A$. Si x n'était pas dans B, alors x serait dans $A\backslash B$, ce qui est contraire à l'hypothèse. Donc $x \in B$.

(3b) $A^c \cup B \subset (A\backslash B)^c$ Supposons d'abord que $x \in A^c$. Comme $A\backslash B \subset A$, si $x \in A^c$, alors $x \in (A\backslash B)^c$.
Si $x \in B$, alors x n'est pas dans $A\backslash B$ et donc x est dans $(A\backslash B)^c$.

(4) **Pour (4)** Pour cette égalité, nous séparons 4 cas suivant que x est dans A ou B ou non. Bref, nous écrivons la table de vérité :

A	1	1	0	0
B	1	0	1	0
$A^c\backslash B^c$	0	0	1	0
$B\backslash A$	0	0	1	0

(1.9)

Les deux dernières lignes étant égales, nous avons l'égalité d'ensembles annoncée.

□

DefBMLooVjlSG

Définition 1.30 (différence symétrique)**.**
Si A et B sont des ensembles, l'ensemble $A\Delta B$ est la ***différence symétrique*** *d'ensembles :*

$$A\Delta B = (A \cup B)\backslash(A \cap B). \tag{1.10}$$

C'est l'ensemble des éléments étant soit dans A soit dans B mais pas dans les deux, ni dans aucun des deux. La table de vérité de $A\Delta B$ est intéressante :

A	1	1	0	0
B	1	0	1	0
$A\Delta B$	0	1	1	0

(1.11)

EQooOJBOooKkKbYp

La deuxième colonne signifie que si $x \in A$ et $x \in B^c$, alors $x \in A\Delta B$.

LemCUVoohKpWB

Lemme 1.31 ([15])**.**
Si A et B sont des parties d'un ensemble, nous avons

$$A\Delta B = (A\backslash B) \cup (B\backslash A), \tag{1.12}$$

EQooMWTNooKXfFyU

et aussi

ItemVUCooHAztC

(1) $A^c\Delta B^c = A\Delta B$.

ItemVUCooHAztCii

(2) $(A\Delta B)\Delta B = A$.

ITEMooSPZXooPTgisP

(3) $(A\Delta B)^c = (A^c \cap B^c) \cup (A \cap B)$.

ITEMooSMXWooYcWsRC

(4) Associativité : $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

Démonstration. Pour (1.12), il suffit de voir que la table de vérité de $(A\backslash B) \cup (B\backslash A)$ est la même que (1.11).

Et pour le reste, c'est parti.

(1) **Pour (1)** Nous rappelons l'égalité $X^c\backslash Y^c = Y\backslash X$ du lemme 1.29(4). De là nous écrivons

$$A^c\Delta B^c = (A^c \cup B^c)\backslash(A^c \cap B^c) = (A^c \cap B^c)^c\backslash(A^c \cup B^c)^c = (A \cup B)\backslash(A \cap B) = A\Delta B. \tag{1.13}$$

(2) **Pour (2)** Ici, il faut remarquer que $(A\Delta B) \cup B = A \cup B$ et que $(A\Delta B) \cap B = B\backslash A$, donc

$$(A\Delta B)\Delta B = (A \cup B)\backslash(B\backslash A) = A. \tag{1.14}$$

(3) **Pour (3)** Il s'agit d'utiliser le lemme 1.29(3) :

$$\begin{aligned}(A\Delta B)^c &= \Big((A \cup B)\backslash(A \cap B)\Big)^c && \text{(1.15a)}\\ &= (A \cup B)^c \cup (A \cap B) && \text{(1.15b)}\\ &= (A^c \cap B^c) \cup (A \cap B). && \text{(1.15c)}\end{aligned}$$

(4) **Pour l'associativité (4)** Nous écrivons les tables de vérités selon que x est dans A, B, C ou non. D'abord

A	1	1	1	1	0	0	0	0
B	1	1	0	0	1	1	0	0
C	1	0	1	0	1	0	1	0
$B\Delta C$	0	1	1	0	0	1	1	0
$A\Delta(B\Delta C)$	1	0	0	1	0	1	1	0

(1.16)

La quatrième ligne s'écrit sur le modèle de (1.11) en regardant les deuxièmes et troisièmes lignes. La dernière ligne se fait avec la première et la quatrième.

L'autre table de vérité se fait de la même manière :

A	1	1	1	1	0	0	0	0
B	1	1	0	0	1	1	0	0
C	1	0	1	0	1	0	1	0
$A\Delta B$	0	0	1	1	1	1	0	0
$(A\Delta B)\Delta C$	1	0	0	1	0	1	1	0

(1.17)

Puisque les lignes pour $A\Delta(B\Delta C)$ et pour $(A\Delta B)\Delta C$ sont identiques, nous avons égalité.

□

1.1.7 Relations d'équivalence

appDefinitionMp

Définition 1.32.
*Si E est un ensemble, une **relation d'équivalence** sur E est une relation binaire*[7] *$\sim$ qui est à la fois*

réflexive *$x \sim x$ pour tout $x \in E$,*

symétrique *$x \sim y$ si et seulement si $y \sim x$;*

transitive *si $x \sim y$ et $y \sim z$, alors $x \sim z$.*

DEFooRHPSooHKBZX1

Définition 1.33.
*Si E est un ensemble et si $\sim$ est une relation d'équivalence sur E, alors nous notons $E/\sim$ l'**ensemble quotient**, c'est-à-dire l'ensemble des classes d'équivalence dans E. Un élément de $E/\sim$ est de la forme*

$$[a] = \{x \in E \text{ tel que } x \sim a\}. \tag{1.18}$$

Lemme 1.34.
Soit un ensemble E et une relation d'équivalence $\sim$. Pour $a, b \in E$, nous avons $[a] = [b]$ si et seulement si $a \sim b$.

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que $[a] = [b]$. Par réflexivité, $a \sim a$ et nous avons $a \in [a] = [b]$. Mais $a \in [b]$ signifie $a \sim b$, ce qu'il fallait.

(2) $\Leftarrow$ Nous supposons que $a \sim b$, et nous démontrons que $[a] \subset [b]$ (pour l'inclusion inverse, vous devriez vous en sortir tout seul). Si $x \in [a]$, alors $x \sim a$. Mais $a \sim b$. Donc $x \sim a \sim b$, ce qui implique $x \sim b$ par transitivité. Or dire $x \sim b$ implique $x \in [b]$.

□

EXooYDRVooHsANlC

Exemple 1.35.
Sur l'ensemble de tous les polygones du plan, la relation « a le même nombre de côtés » est une relation d'équivalence. Plus précisément, si P et Q sont deux polygones, nous disons que $P \sim Q$ si et seulement si P et Q ont le même nombre de côtés. C'est une relation d'équivalence :

— un polygone P a toujours le même nombre de côtés que lui-même : $P \sim P$;

— si P a le même nombre de côtés que Q ($P \sim Q$), alors Q a le même nombre de côtés que P ($Q \sim P$) ;

— si P a le même nombre de côtés que Q ($P \sim Q$) et que Q a le même nombre de côtés que R ($Q \sim R$), alors P a le même nombre de côtés que R ($P \sim R$).

△

7. Définition 1.10.

Exemple 1.36.
Soit f une application entre deux ensembles E et F. Nous définissons une relation d'équivalence sur E par

$$x \sim y \Leftrightarrow f(x) = f(y). \tag{1.19}$$

Nous notons par $\pi\colon E \to E/\sim$ la projection canonique. L'application

$$\begin{aligned} g\colon E/\sim &\to F \\ [x] &\mapsto f(x) \end{aligned} \tag{1.20}$$

est bien définie et injective. Elle n'est pas surjective tant que f ne l'est pas. La **décomposition canonique** de f est

$$f = g \circ \pi. \tag{1.21}$$

△

1.2 Quelques structures algébriques

Nous collectons ici les définitions des principales structures algébriques.

DEFooBMUZooLAfbeM

Définition 1.37 (Groupe)**.**
*Un **groupe** est un ensemble G muni d'une opération interne $\cdot : G \times G \to G$ telle que*

(1) pour tous $g, h, k \in G$, $g \cdot (h \cdot k) = (g \cdot h) \cdot k$,

(2) il existe un élément $e \in G$ tel que $e \cdot g = g \cdot e = g$ pour tout $g \in G$,

(3) pour tout $g \in G$, il existe un élément $h \in G$ tel que $g \cdot h = h \cdot g = e$.

*Un groupe est **commutatif** ou **abélien** si $g \cdot h = h \cdot g$ pour tout $g, h \in G$.*

Notons que nous avons écrit $g \cdot h$ et non $\cdot(g, h)$ comme une notation purement fonctionnelle nous l'aurait suggéré. Dans les exemples concrets, selon les cas, la loi de groupe appliquée à g et h sera notée tantôt $g + h$, tantôt $g \cdot h$ ou, le plus souvent pour un groupe générique, simplement gh.

DEFooBEHTooMeCOTX

Définition 1.38 (morphisme, automorphisme)**.**
*Soient deux groupes G et H. Un **morphisme** entre G et H est une application $\alpha\colon G \to H$ telle que pour tout $g, h \in G$ nous ayons $\alpha(gh) = \alpha(g)\alpha(h)$.*

Comme d'habitude, un isomorphisme est un morphisme bijectif.

*Un **automorphisme** de G est un isomorphisme de G vers G lui-même. L'ensemble des automorphismes de G est noté* $\mathrm{Aut}(G)$.

PROPooJHHPooIpciPA

Proposition 1.39.
Si $\alpha\colon G \to H$ est un morphisme de groupes, alors $\alpha(G)$ est un sous-groupe de H.

LEMooKJMWooQeIwgF

Lemme 1.40.
Si G est un groupe, alors $\mathrm{Aut}(G)$ *est un groupe pour la composition.*

Démonstration. Soient α et β des automorphismes de G. Alors nous prouvons que $\alpha \circ \beta$ est un automorphisme de G :

$$(\alpha \circ \beta)(gh) = \alpha\big(\beta(g)\beta(h)\big) = \alpha\big(\beta(g)\big)\alpha\big(\beta(h)\big) = (\alpha \circ \beta)(g)(\alpha \circ \beta)(h). \tag{1.22}$$

□

LEMooDPISooRoAFmt

Lemme 1.41.
Si G et H sont des groupes isomorphes, alors les groupes $\mathrm{Aut}(G)$ *et* $\mathrm{Aut}(H)$ *sont isomorphes.*

Démonstration. Soit un isomorphisme $f\colon G \to H$. D'abord pour tout $\alpha \in \mathrm{Aut}(G)$, l'application $f \circ \alpha \circ f^{-1}$ est un automorphisme de H. Cela est rapidement vérifié parce que f, α et f^{-1} sont des bijections et des morphismes.

Nous pouvons donc considérer l'application

$$\begin{aligned}\psi\colon \mathrm{Aut}(G) &\to \mathrm{Aut}(H)\\ \alpha &\mapsto f \circ \alpha \circ f^{-1}.\end{aligned} \tag{1.23}$$

(1) ψ **est un morphisme** Soient $\alpha, \beta \in \mathrm{Aut}(G)$. Vu que f est une bijection, nous pouvons introduire $f^{-1} \circ f$ partout où ça nous plaît :

$$\psi(\alpha\beta) = f \circ \alpha \circ \beta \circ f^{-1} = f \circ \alpha \circ f^{-1} \circ f \circ \beta \circ f^{-1} = \psi(\alpha) \circ \psi(\beta). \tag{1.24}$$

(2) ψ **est injective** Si $\alpha\beta \in \mathrm{Aut}(G)$ sont tels que $\psi(\alpha) = \psi(\beta)$, alors

$$f \circ \alpha \circ f^{-1} = f \circ \beta \circ f^{-1}. \tag{1.25}$$

Comme f est une bijection, cela implique que $\alpha = \beta$.

(3) ψ **est surjective** Si $\alpha\colon H \to H$ est un automorphisme, alors $\sigma = \psi(f^{-1} \circ \sigma \circ f)$ où il est facile de vérifier que $f^{-1} \circ \sigma \circ f \in \mathrm{Aut}(G)$.

□

DefHXJUooKoovob

Définition 1.42 (Anneau[16])**.**
Un ***anneau***[8] *est un triplet* $(A, +, \cdot)$ *avec les conditions*

(1) $(A, +)$ *est un groupe*[9] *commutatif. Nous notons* 0 *le neutre.*

(2) La multiplication est associative et nous notons 1 *le neutre.*

ITEMooGMNOooSTGiXw

(3) La multiplication est distributive par rapport à l'addition.

L'anneau $(A, +, \cdot)$ *est* ***commutatif*** *si pour tout* $a, b \in A$ *nous avons* $a \cdot b = b \cdot a$.

DEFooSPHPooCwjzuz

Définition 1.43 (Morphisme d'anneaux[17])**.**
Si $(A, +, \cdot)$ *et* $(B, +, \cdot)$ *sont des anneaux*[10]*, un* ***morphisme d'anneaux*** *est une application* $f\colon A \to B$ *telle que*

(1) $f(a + b) = f(a) + f(b)$

(2) $f(a \cdot b) = f(a) \cdot f(b)$

(3) $f(1) = 1$.

Étant bien entendu que les significations de 1, $+$ *et* $\cdot$ *sont différentes à gauche et à droite.*

1.3 Les naturels

DEFooHJBOooWVallq

Définition 1.44 ([18])**.**
Un ***triplet naturel*** *est un triplet* $(\mathcal{N}, o, s)$ *où* $\mathcal{N}$ *est un ensemble,* o *est un élément de* $\mathcal{N}$ *et* s *est une application* $s\colon \mathcal{N} \to \mathcal{N}$ *satisfaisant les propriétés suivantes :*

(1) s *est injective,*

ITEMooQAKJooGKdJsM

(2) $s(\mathcal{N}) = \mathcal{N} \backslash \{o\}$

ITEMooXPYEooFajywh

(3) Si $A \subset \mathcal{N}$ *est tel que* $o \in A$ *et* $s(A) \subset A$*, alors* $A = \mathcal{N}$.

8. Nous faisons le choix qu'un anneau admet toujours un neutre pour la multiplication. Certains ouvrages parlent dans ce cas d'anneau unitaire.
9. Groupe, définition 1.37.
10. Définition 1.42.

Le théorème suivant est typiquement de ceux qui vont demander de gratter la théorie axiomatique des ensembles avec une certaine précision [11].

THOooOXMHooXYgMqb

Théorème 1.45.
Il existe un[12] *triplet naturel.*

1.46 (Définition de $\mathbb{N}$)**.**
Pour la suite, nous considérons un triplet naturel $(\mathcal{N}, o, s)$ et nous notons $\mathbb{N} = \mathcal{N}$. Donc la nature de tous les objets que nous allons considérer à partir de maintenant dépend du choix de triplet naturel que nous faisons à présent. Le théorème 1.83 nous assurera que peu de choses devraient réellement dépendre de ce choix.

Nous notons 0 l'élément o et 1 l'élément $s(0)$. C'est tout ce dont nous avons besoin dans l'immédiat.

1.3.1 Applications définies par récurrence

PROPooXTRCooKwrWkq

Proposition 1.47 (Récurrence[18])**.**
Soit un triplet naturel $(\mathcal{N}, o, s)$ et une application $P\colon \mathcal{N} \to \{0,1\}$ vérifiant[13]

(1) $P(o) = 1$,

(2) pour tout $a \in \mathcal{N}$, si $P(a) = 1$, alors $P\big(s(a)\big) = 1$.

Alors $P(x) = 1$ pour tout $x \in \mathcal{N}$.

Démonstration. Nous posons

$$A = \{x \in \mathcal{N} \text{ tel que } P(x) = 1\}. \tag{1.26}$$

Cet ensemble vérifie la propriété 1.44(3). Donc $A = \mathcal{N}$. □

THOooEJPYooZFVnez

Théorème 1.48 ([13, 19])**.**
Soient E un ensemble, g une application de E dans E et b un élément de E. Alors il existe une unique application $f\colon \mathbb{N} \to E$ telle que :

(1) $f(0) = b$

(2) $f\big(s(n)\big) = g\big(f(n)\big)$ *pour tout* $n \in \mathbb{N}\backslash\{0\}$.

Démonstration. Nous commençons par l'unicité. Soient f_1 et f_2 deux telles applications. Nous posons

$$A = \{n \in \mathbb{N} \text{ tel que } f_1(n) = f_2(n)\}. \tag{1.27}$$

Nous avons $0 \in A$ parce que $f_1(0) = f_2(0) = b$.

Supposons que $f_1(k) = f_2(k)$. Alors nous avons

$$f_1\big(s(k)\big) = g\big(f_1(k)\big) = g\big(f_2(k)\big) = f_2\big(s(k)\big). \tag{1.28}$$

Nous en déduisons que $s(k) \in A$. Autrement dit $s(A) \subset A$. La définition 1.44(3) nous indique alors que $A = \mathbb{N}$, c'est-à-dire que $f_1 = f_2$.

Nous montrons à présent l'existence en plusieurs étapes.

(1) **L'ensemble est assez grand** Nous considérons l'ensemble $\mathcal{A}$ des parties $A \subset \mathbb{N} \times E$ telles que

11. Ou alors il y a quelque chose qui m'échappe. Écrivez-moi si vous connaissez une construction « simple ».

12. Nous verrons plus tard que toute partie infinie d'un triplet naturel fournit un nouveau triplet naturel ; il en existe donc plusieurs.

13. Les plus pointilleuses diront que 1 n'est pas encore défini. Bon j'avoue. Ce qui est important est que P prenne ses valeurs dans un ensemble contenant deux éléments distincts. Si maintenant vous râlez parce que « deux » est encore moins défini, prenez un ensemble quelconque A et dites que P prend ses valeurs dans $\{A, \mathcal{P}(A)\}$. Mais êtes-vous bien certaine que $\mathcal{P}(A) \neq A$?

(1a) $(0,b) \in A$

(1b) $(n,x) \in A \Rightarrow \big(s(n),g(x)\big) \in A$.

L'ensemble $\mathcal{A}$ est non vide parce que $\mathbb{N} \times E \in \mathcal{A}$.

(2) **Le plus petit** Nous posons

$$G = \bigcap_{A \in \mathcal{A}} A, \tag{1.29}$$

et nous prouvons que $G \in \mathcal{A}$. D'abord $(0,b) \in G$ parce que cet élément est dans chacun des $A \in G$. Ensuite si $(n,x) \in G$, alors pour tout $A \in \mathcal{A}$ nous avons $(n,x) \in A$ et donc $\big(s(n),g(x)\big) \in A$. Par conséquent $\big(s(n),g(x)\big) \in \bigcup_{A\in\mathcal{A}} A = G$.
Pour $n \in \mathbb{N}$ nous posons

$$G_n = \{x \in E \text{ tel que } (n,x) \in G\}. \tag{1.30}$$

Nous avons en particulier que $b \in G_0$ parce que $(0,b) \in G$.

(3) G **contient un** (n,x) **pour tout** n Nous prouvons que pour tout $n \in \mathbb{N}$, il existe $x \in E$ tel que $(n,x) \in G$. Nous faisons ça avec la proposition 1.47 en posant

$$\begin{aligned} P\colon \mathbb{N} &\to \{0,1\} \\ n &\mapsto \begin{cases} 1 & \text{si } G_n \neq \varnothing \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.31}$$

Puisque $(0,b) \in G$ nous avons $P(0) = 1$. Supposons que $P(k) = 1$ et montrons que $P\big(s(k)\big) = 1$. Comme $P(k) = 1$, il existe $x \in E$ tel que $(k,x) \in G$. De ce fait, $\big(s(k),g(x)\big) \in G$, ce qui donne $G_{s(k)} \neq \varnothing$ et $P\big(s(k)\big) = 1$.

(4) G_n **est un singleton** Nous avons vu que G_n n'est jamais vide. Nous allons montrer que G_n est un singleton pour tout n. Pour cela nous posons

$$\begin{aligned} P\colon \mathbb{N} &\to \{0,1\} \\ n &\mapsto \begin{cases} 1 & \text{si } G_n \text{ est un singleton} \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.32}$$

Nous prouvons par récurrence que $P(n) = 1$ pour tout n.

(4a) $P(0) = 1$ Nous commençons par prouver que $P(0) = 1$. Nous savons que $(0,b) \in G_0$. Supposons $a \neq b$ tel que $(0,a) \in G_0$. Alors en posant $G' = G\backslash\{(0,a)\}$ nous avons $G' \in \mathcal{A}$.
En effet $(0,b) \in G'$ parce que $(0,b) \in G$ et $(0,b) \neq (0,a)$. De plus si $(n,x) \in G'$, alors $\big(s(n),g(x)\big) \in G$. Mais comme $s(n) \neq 0$ nous avons $\big(s(n),g(x)\big) \neq (0,a)$ et donc $\big(s(n),g(x)\big) \in G'$.
L'ensemble G' serait un élément de $\mathcal{A}$ strictement inclus dans G. Impossible. Donc G_0 est un singleton.

(4b) **Récurrence** Supposons que $P(k) = 1$, c'est-à-dire que G_k est un singleton. Soit e l'unique élément de G_k : $(k,e) \in G$. Nous avons alors aussi que $\big(s(k),g(e)\big) \in G$. Nous devons prouver que si $y \in G_{s(k)}$, alors $y = g(e)$.
Supposons donc $y \neq g(e)$ soit dans $G_{s(k)}$. Nous posons

$$G' = G\backslash\{\big(s(k),y\big)\}. \tag{1.33}$$

Nous prouvons que $G' \in \mathcal{A}$. D'abord $(0,b) \in G'$ parce que $s(k) \neq 0$. Soit ensuite $(m,z) \in G'$. Si $m = k$, alors $z = e$ (parce que par hypothèse G_k est un singleton) et nous savons que $\big(s(m),g(e)\big) \in G'$. Si par contre $m \neq k$, comme s est injective, nous avons aussi $s(m) \neq s(k)$. Donc $\big(s(m),g(z)\big) \neq \big(s(k),y\big)$ et $\big(s(m),g(z)\big) \in G'$. Donc $G' \in \mathcal{A}$ et est strictement plus petit que G. Contradiction.
Nous concluons que $G_{s(k)}$ est un singleton, c'est-à-dire que $P\big(s(k)\big) = 1$.

(4c) **Conclusion** Nous avons prouvé que G_n est un singleton pour tout n.

(5) **Et enfin**

Nous définissons $f(n)$ comme étant l'unique élément de G_n. Puisque $(0,b)\in G$ nous avons $G_0=\{b\}$ et donc $g(0)=b$.

Par définition de f, nous avons $\big(n,f(n)\big)\in G$. Parce que $G\in\mathcal{A}$ nous avons alors

$$\big(s(n),g\big(f(n)\big)\big)\in G. \tag{1.34}$$

Autrement dit, $G_{s(n)}=\{g\big(f(n)\big)\}$. Cela montre que

$$f\big(s(n)\big)=g\big(f(n)\big), \tag{1.35}$$

et donc que f vérifie les propriétés demandées.

□

NORMooLNXMooWIblPf

1.49 ([1]).
Le théorème 1.48 nous permet de définir une suite $(x_n)_{n\in\mathbb{N}}$ par récurrence en posant par exemple

$$\begin{cases} x_1=b & \text{(1.36a)}\\ x_{n+1}=f(x_n). & \text{(1.36b)}\end{cases}$$

Parfois nous pouvons avoir envie de définir une suite par récurrence de telle sorte que chaque nouveau terme dépende de tous les précédents :

$$\begin{cases} x_1=b & \text{(1.37a)}\\ x_{n+1}=g(x_1,\dots,x_{n-1}). & \text{(1.37b)}\end{cases}$$

Pour ce faire, il faut un peu adapter.

Posons $E'=E\cup(E\times E)\cup(E\times E\times E)\cup\dots$. Nous supposons avoir une application $g\colon E'\to E'$ qui ne prend ses valeurs que dans E : $g(E')\subset E$. De plus nous considérons l'application

$$\begin{aligned} a\colon\mathbb{N}&\to E'\\ n&\mapsto\big(f(0),\dots,f(n)\big).\end{aligned} \tag{1.38}$$

Le théorème 1.48 appliqué à $g\circ a$ nous donne l'existence et l'unicité d'une application $f\colon\mathbb{N}\to E'$ telle que

$$\begin{cases} f(0)=b & \text{(1.39a)}\\ f(n+1)=g\big(f(0),\dots,f(n)\big). & \text{(1.39b)}\end{cases}$$

CORooVNHKooRkKtXf

Corolaire 1.50 ([13]).
Soient deux ensembles X,Y, une application $\alpha\colon X\to Y$ et une application $\beta\colon Y\to Y$. Alors il existe une unique application $H\colon X\times\mathbb{N}\to Y$ telle que

(1) $H(x,0)=\alpha(x)$ pour tout élément $x\in X$;

(2) $H(x,n+1)=\beta(H(x,n))$ pour tout élément $x\in X$ et pour tout $n\in\mathbb{N}$.

Démonstration. Pour faire le lien avec les notations du théorème 1.48, nous notons $E=\mathrm{Fun}(X,Y)$, $b=\alpha\in E$ et

$$\begin{aligned} g\colon E&\to E\\ s&\mapsto\beta\circ s.\end{aligned} \tag{1.40}$$

Le théorème 1.48 donne alors l'existence d'une application $f\colon\mathbb{N}\to E$ telle que

(1) $f(0)=b$

(2) $f(n+1)=g\big(f(n)\big)$.

Nous définissons alors

$$\begin{aligned} H\colon X\times\mathbb{N}&\to Y\\ (x,n)&\mapsto f(n)x, \end{aligned} \tag{1.41}$$

et nous vérifions qu'elle satisfait aux exigences.

(1) D'abord nous avons

$$H(x,0)=f(0)x=b(x)=\alpha(x). \tag{1.42}$$

(2) Ensuite, pour $x\in X$ et $n\in\mathbb{N}$ nous avons :

$$\begin{aligned} H(x,n+1)&=f(x+1)x && \text{(1.43a)}\\ &=g\big(f(n)\big)x && \text{(1.43b)}\\ &=g\big(f(n)\big)x && \text{(1.43c)}\\ &=\big(\beta\circ f(n)\big)x && \text{(1.43d)}\\ &=\beta\big(f(n)x\big) && \text{(1.43e)}\\ &=\beta\big(H(x,n)\big). && \text{(1.43f)} \end{aligned}$$

Et voilà. □

1.3.2 Addition sur les naturels

DEFooIJIEooZaAdSs

Définition 1.51 (élément régulier[20])**.**
*Soit un ensemble E muni d'une opération $*\colon E\times E\to E$. Un élément $s\in E$ est **régulier à gauche** si pour tout $x,y\in E$ nous avons*

$$s*x=s*y\Rightarrow x=y. \tag{1.44}$$

L'élément s est régulier à droite si pour tout $x,y\in E$ nous avons

$$x*s=y*s\Rightarrow x=y. \tag{1.45}$$

*Il est **régulier** si il est régulier à gauche et à droite.*

PROPooVFOXooXmwpFh

Proposition-Définition 1.52 ([18, 1])**.**
Il existe une unique fonction $f\colon\mathbb{N}\times\mathbb{N}\to\mathbb{N}$ vérifiant

ITEMooILZSooNYIkYR

(1) $f(a,0)=a$ pour tout $a\in\mathbb{N}$

ITEMooZWHQooBAjZyE

(2) $f\big(a,s(b)\big)=s\big(f(a,b)\big)$ pour tout $a,b\in\mathbb{N}$.

Pour $a,b\in\mathbb{N}$ nous notons $f(a,b)=a+b$.

LEMooMJMTooOtUuJT

Lemme 1.53 ([18])**.**
Pour tout $a\in\mathbb{N}$ nous avons $s(a)=a+1$.

Démonstration. Nous avons :

$$\begin{aligned} s(a)&=s(a+0) && \text{(1.46a)}\\ &=a+s(0) && \text{(1.46b)}\\ &=a+1. && \text{(1.46c)} \end{aligned}$$

SUBEQooMNBLooTQruhE
SUBEQooAGUSooGjiYGj
SUBEQooUZQDooWtNBHQ

Justifications.

(1) Pour (1.46a) c'est dans la définition 1.52(1) de la somme.

(2) Pour (1.46b), c'est dans la définition 1.52(2) de la somme.

(3) Pour (1.46c). Le symbole « 1 » désigne l'élément $s(0)$ dans $\mathbb{N}$.

□

PROPooTLTSooGNMTmV

Proposition 1.54 ([18]).
En ce qui concerne la somme dans $\mathbb{N}$.

ITEMooIFFPooXfftfG

(1) *La somme est associative et commutative.*

ITEMooSGRVooPAVFYK

(2) *L'élément* 0 *est neutre.*

ITEMooNUTHooJWWzGv

(3) *Tous les éléments de* $\mathbb{N}$ *sont réguliers*[14] *par rapport à l'addition.*

Démonstration. En plusieurs parties.

(1) **Associative** Nous devons prouver que $(a+b)+c=a+(b+c)$ pour tout $a,b,c\in\mathbb{N}$. Pour ce faire, nous fixons $a,b\in\mathbb{N}$ et nous prouvons l'égalité demandée par récurrence sur c.
Pour $c=0$, nous avons $(a+b)+c=a+b$ et $a+(b+c)=a+b$. Donc nous sommes d'accord[15].
Nous vérifions avec $s(c)$:

$$\begin{aligned}(a+b)+s(c)&=s\big((a+b)+c\big) &(1.47a)\\ &=s\big(a+(b+c)\big) &(1.47b)\\ &=a+s(b+c) &(1.47c)\\ &=a+\big(b+s(c)\big). &(1.47d)\end{aligned}$$

SUBEQooDDUBooAjuuZq

Justifications.

— Pour (1.47b). C'est l'hypothèse de récurrence. À ce stade, je vous conseille d'être capable de rédiger complètement la récurrence et l'appel au théorème 1.48.

(2) **Neutre** La définition de l'addition contient déjà $a+0=a$. Nous prouvons par récurrence que $0+a=a$ pour tout $a\in\mathbb{N}$.
Pour $a=0$, l'égalité demandé est correcte : $0+0=0$ parce que pour tout x dans $\mathbb{N}$, $0+x=x$. Pour $s(a)$ nous avons

$$0+s(a)=s(0+a)=s(a). \tag{1.48}$$

La dernière égalité est l'hypothèse de récurrence.
Nous posons

$$A=\{a\in\mathbb{N}\text{ tel que }0+a=a\}. \tag{1.49}$$

Nous avons prouvé que $0\in A$ et que $s(A)\subset A$. Le théorème 1.48 nous assure alors que $A=\mathbb{N}$.

(3) **Commutativité** Nous fixons $a\in\mathbb{N}$ et nous prouvons par récurrence sur b que $a+b=b+a$ pour tout $b\in\mathbb{N}$. Cela va être décomposé en plusieurs étapes.

(4) $a+0=0+a$ Pour $b=0$ c'est correct, car $b+0=0+b=b$ parce que 0 est neutre.

(5) $a+1=1+a$ Nous démontrons par récurrence sur a que $a+1=1+a$. Avec $a=0$ c'est déjà fait. Pour les autres,

$$\begin{aligned}s(a)+1&=(a+1)+1 &\text{lemme 1.53} &(1.50a)\\ &=(1+a)+1 &\text{hypothèse récurrence} &(1.50b)\\ &=1+(a+1) &\text{associativité} &(1.50c)\\ &=1+s(a). & &(1.50d)\end{aligned}$$

14. Élément régulier pour une opération, définition 1.51.
15. Notez que nous n'avons pas utilisé le fait que 0 était neutre des deux côtés – chose que nous n'avons pas encore démontré. Nous avons seulement utilisé $a+0=a$, qui est dans la définition de la somme.

(6) $a+b=b+a$ Nous y voici. Nous fixons a et nous prouvons par récurrence que $a+b=b+a$. Pour $b=0$ c'est déjà fait. Pour les autres,

$$\begin{aligned}
a+s(b)&=a+(b+1) && && (1.51a)\\
&=(a+b)+1 && && (1.51b)\\
&=(b+a)+1 && \text{hypothèse récurrence} && (1.51c)\\
&=b+(a+1) && && (1.51d)\\
&=b+(1+a) && \text{commutativité avec 1} && (1.51e)\\
&=(b+1)+a && \text{associativité} && (1.51f)\\
&=s(b)+a. && && (1.51g)
\end{aligned}$$

Récurrence terminée.

(7) **Régularité** Nous devons prouver que, pour tout $a,x,y\in\mathbb{N}$, si $a+x=a+y$ alors $x=y$. Nous allons procéder par récurrence en posant

$$A=\{a\in\mathbb{N}\text{ tel que }\forall x,y\in\mathbb{N},a+x=a+y\Rightarrow x=y\}. \tag{1.52}$$

Puisque $0+x=x$ et $0+y=y$, nous avons $0\in A$. Supposons à présent que $a\in A$ et montrons que $s(a)\in A$. Soient $x,y\in\mathbb{N}$ tels que $a+x=a+y$. Nous avons :

$$\begin{aligned}
&s(a)+x=s(a)+y && && (1.53a)\\
\Rightarrow\quad &s(a+x)=s(a+y) && && (1.53b)\\
\Rightarrow\quad &a+x=a+y && && (1.53c)\\
\Rightarrow\quad &x=y && && (1.53d)
\end{aligned}$$

SUBEQooNJHZooSIKPxN SUBEQooWDJLooJbzIFe SUBEQooTLYZooFsMaJD

Justifications.

— Pour (1.53b). En utilisant la définition de l'addition et la commutativité, nous avons $s(a)+x=s(a+x)$.

— Pour (1.53c). Parce que s est injective ; c'est dans la définition 1.44 d'un triplet naturel.

— Pour (1.53d). Parce que $a\in A$.

Nous avons prouvé que $s(A)\subset A$, et donc que $A=\mathbb{N}$.

□

LEMooCOMSooEWrumL

Lemme 1.55.
Nous avons $0\neq 1$.

Démonstration. Par définition $1=s(0)$. Comme s est à valeurs dans $\mathbb{N}\backslash\{0\}$, nous ne pouvons pas avoir $s(0)=0$. □

LEMooQBHFooCuCusQ

Lemme 1.56 ([1])**.**
Si $a+b=0$*, alors* $a=b=0$.

Démonstration. Soient $a,b\in\mathbb{N}$ tels que $a+b=0$, et supposons que $b\neq 0$. Par la définition 1.44(2), nous avons $b=s(c)$ pour un certain $c\in\mathbb{N}$.

Dans ce cas nous avons $a+b=a+s(c)=s(a+c)\neq 0$ parce que l'image de s ne contient pas 0. Hélas, par hypothèse nous avons $a+b=0$. Nous avons obtenu une contradiction, et nous déduisons que $b=0$.

Maintenant que nous savons que $b=0$, il reste $0=a+b=a+0=a$. □

1.3.3 Ordre sur les naturels

DEFooAXZSooTEMjlV

Définition 1.57 ([18]).
Pour $a, b \in \mathbb{N}$, nous notons $a \leqslant b$ si il existe $x \in \mathbb{N}$ tel que $a + x = b$.
Nous notons également $a < b$ si $a \leqslant b$ et $a \neq b$.

LEMooWMYPooLTMyWb

Lemme 1.58.
Nous avons $a \leqslant s(a)$ pour tout $a \in \mathbb{N}$.

Démonstration. Cela est une conséquence du lemme 1.53 : $s(a) = a + 1$. □

PROPooVXBBooZcghrA

Proposition 1.59.
La relation $\leqslant$ est une relation d'ordre compatible avec l'addition.

Démonstration. Plusieurs choses à vérifier.

(1) **Réflexive** Nous avons $a \leqslant a$ parce que $a + 0 = a$.

(2) **Antisymétrique** Soient $a, b \in \mathbb{N}$ tels que $a \leqslant b$ et $b \leqslant a$. Il existe $x, y \in \mathbb{N}$ tels que

$$b = a + x \tag{1.54a}$$
$$a = b + y. \tag{1.54b}$$

En substituant la seconde équation dans la première, $b = (b + y) + x$ que nous récrivons, en utilisant l'associativité [16],

$$0 + b = b + (x + y). \tag{1.55}$$

En utilisant la régularité, $0 = x + y$ et donc $x = y = 0$ par le lemme 1.56. Cela donne alors $a = b$.

(3) **Transitive** Si $a \leqslant b$ et $b \leqslant c$, nous avons $n, p \in \mathbb{N}$ tels que $b = a + n$ et $c = b + p$. Donc

$$c = (a + n) + p = a + (n + p), \tag{1.56}$$

ce qui signifie que $a \leqslant c$. Notez l'utilisation de l'associativité de la somme, démontrée en la proposition 1.54(1).

(4) **Compatibilité** Soient $a, b, n \in \mathbb{N}$ tels que $a \leqslant b$. Nous devons montrer que $a + n \leqslant b + n$. Puisque $a \leqslant b$, il existe $x \in \mathbb{N}$ tel que $b = a + x$. Par conséquent,

$$b + n = a + x + n = (a + n) + x, \tag{1.57}$$

qui signifie bien que $a + n \leqslant b + n$.

□

LEMooPVRQooXPMKTt

Lemme 1.60.
À propos d'ordre et de stricte inégalité.

ITEMooGWWFooYGPCZw

(1) Si $x \leqslant a$ et $b \neq 0$, alors $x < a + b$.

ITEMooRWGWooAfkrri

(2) Si $x \leqslant a$, alors $x < s(a)$.

ITEMooWCOIooMWrCag

(3) Si $x < a$, alors $s(x) \leqslant a$.

Démonstration. En plusieurs parties.

(1) **Pour (1)** Si $x \leqslant a$, il existe $d \in \mathbb{N}$ tel que $x + d = a$. Nous avons alors aussi

$$x + d + b = a + b, \tag{1.58}$$

ce qui signifie que $x \leqslant a + b$. Mais si x était égal à $a + b$, nous aurions $d + b = 0$, ce qui impliquerait [17] $d = b = 0$, alors que l'hypothèse stipule que $b \neq 0$. Donc $x \neq a + b$.

16. Proposition 1.54(1).
17. Par le lemme 1.56.

(2) **Pour (2)** Il s'agit seulement d'utiliser le point (1) avec $b = 1$ et le fait que $s(a) = a + 1$ par le lemme 1.53.

(3) **Pour (3)** Par hypothèse, il existe $b \neq 0$ tel que $x + b = a$. Puisque $b \neq 0$, il existe $c \in \mathbb{N}$ tel que $b = s(c)$ et donc, tel que

$$x + s(c) = a. \tag{1.59}$$

En utilisant le fait que $s(c) = c + 1$ ainsi que l'associativité et la commutativité de l'addition (proposition 1.54(1)) nous avons

$$a = x + s(c) = s(x) + c, \tag{1.60}$$

ce qui prouve que $s(x) \leqslant a$.

□

LEMooCSIXooHeuWEd

Lemme 1.61.
L'élément 0 *est l'unique plus petit élément de* $\mathbb{N}$.

Démonstration. Puisque 0 est neutre pour l'addition[18], nous avons $a + 0 = a$ pour tout $a \in \mathbb{N}$ et donc $0 \leqslant a$ pour tout a. Cela veut dire que 0 est plus petit que tout élément de $\mathbb{N}$.

En ce qui concerne l'unicité, soit $z \in \mathbb{N}$ tel que $z \leqslant a$ pour tout $a \in \mathbb{N}$. Si $z \neq 0$, il existe $x \in \mathbb{N}$ tel que $z = s(x)$. Nous avons donc $z \geqslant x$ en même temps que $x \leqslant z$. Cela implique $z = x$ (parce qu'une relation d'ordre est symétrique) et donc $z = z + 1$. En utilisant la régularité de z pour l'addition nous en déduisons que $0 = 1$, ce qui est impossible par le lemme 1.55. □

LEMooJRZKooOMhOkH

Lemme 1.62.
Si $a \leqslant b$ *et* $b \leqslant a$, *alors* $a = b$.

Démonstration. L'inégalité $a \leqslant b$ dit qu'il existe $x \in \mathbb{N}$ tel que $a + x = b$. En mettant cela dans l'inégalité $b \leqslant a$ nous trouvons $a + x \leqslant a$ qui donne, via la proposition 1.59 : $x \leqslant 0$. Nous en déduisons que $x = 0$ parce que zéro est l'unique minimum de $\mathbb{N}$ par le lemme 1.61. □

PROPooGCCRooFBYrlo

Proposition 1.63.
Le couple $(\mathbb{N}, \leqslant)$ *est totalement ordonné*[19].

Démonstration. Soit $a \in \mathbb{N}$. Nous devons prouver que pour tout $x \in \mathbb{N}$ nous avons $x \leqslant a$ ou $a \leqslant x$ (non exclusifs). Nous posons

$$A = \{x \in \mathbb{N} \text{ tel que } x \leqslant a\} \tag{1.61a}$$
$$B = \{x \in \mathbb{N} \text{ tel que } a \leqslant x\}, \tag{1.61b}$$

et nous prouvons que $A \cup B = \mathbb{N}$ en montrant que $0 \in A \cup B$ et que $s(A \cup B) \subset A \cup B$.

Nous avons $0 \in A \subset A \cup B$ par le lemme 1.61.

Pour étudier $s(A \cup B)$, nous considérons $x \in A \cup B$ et nous subdivisons en deux cas selon que $x \in A$ ou $x \in B$.

(1) **Si $x \in B$** Si $x \in B$, alors $a \leqslant x \leqslant s(x)$ parce que $x \leqslant s(x)$ par le lemme 1.58. Donc $s(x) \in B \subset A \cup B$.

(2) **Si $x \in A$** Si $x \in A$, il y a deux possibilités : $x = a$ et $x \neq a$. Si $x = a$, alors $a \leqslant s(x)$ et donc $s(x) \in A \subset A \cup B$.

Si $x \neq a$, alors le lemme 1.60(3) nous indique que $s(x) \leqslant a$ et donc $s(x) \in A \subset A \cup B$.

Nous avons donc prouvé que $s(A \cup B) \subset A \cup B$, et donc que $A \cup B = \mathbb{N}$. □

PROPooMZOWooHmsXzI

Proposition 1.64 ([18, 1])**.**
L'ensemble ordonné $(\mathbb{N}, \leqslant)$ *vérifie les propriétés suivantes.*

18. Proposition 1.54(2).
19. Définition 1.12.

ITEMooJLAHooDKukfH
ITEMooYAJIooEFmOpB
ITEMooSRGOooNYJJHY
ITEMooKIHZooDRTCdx

(1) L'élément 0 est l'unique minimum de $\mathbb{N}$.

(2) Toute partie non vide a un unique plus petit élément.

(3) L'ensemble $\mathbb{N}$ *n'a pas de plus grand élément.*

(4) Toute partie non vide majorée a un unique plus grand élément.

Démonstration. Point par point.

(1) **Pour (1)** C'est le lemme 1.61.

(2) **Pour (2)** Soit une partie A non vide dans $\mathbb{N}$. Si $0 \in A$, nous avons fini.

(2a) **L'ensemble B** Nous supposons donc que A ne contient pas zéro et nous définissons

$$B = \{n \in \mathbb{N} \backslash A \text{ tel que } n \leqslant a, \forall a \in A\}. \tag{1.62}$$

(2b) **Un élément particulier dans B** L'ensemble B vérifie :

— $0 \in B$

— $B \neq \mathbb{N}$ parce que A est non vide.

La contraposée de la condition (3) de la définition 1.44 d'un triplet naturel implique que $s(B) \not\subset B$. Autrement dit, il existe $b \in B$ tel que $s(b) \notin B$.

(2c) **Deux fonctions sur A** Puisque $b \in B$, nous avons une application $c\colon A \to \mathbb{N}$ telle que $b + c(a) = a$. Nous avons $c(a) \neq 0$ parce que $c(a) = 0$ signifierait $b = a$, ce qui est impossible parce que $a \in A$ et $b \in B$.

Comme pour tout $a \in A$, l'élément $c(a)$ est non nul, il existe une fonction $d\colon A \to \mathbb{N}$ telle que $c(a) = d(a) + 1$.

(2d) **$s(b) \in A$** Supposons que $s(b) \notin A$. Alors il existe $a \in A$ tel que $s(b) \leqslant a$ est faux. Puisque l'ordre est total (proposition 1.63), nous avons

EQooYQSFooPSPJMt
$$a \leqslant s(b). \tag{1.63}$$

Comme $b \in B$ nous avons aussi

EQooIPAWooDBSJEa
$$b \leqslant a. \tag{1.64}$$

Et enfin nous avons

EQooWFHRooGDSBFD
$$b \neq a \tag{1.65}$$

parce que $a \in A$ et $b \in B$.

Les conditions (1.64) et (1.65) se résument en $b < a$. Le lemme 1.60(3) nous indique alors que $s(b) \leqslant a$. Cela mis à côté de (1.63) conclut que $a = s(b)$, et donc que $s(b)$ est dans A. Contradiction. Nous en concluons que $s(b) \in A$.

(2e) **$s(b)$ est un minimum de A** En utilisant la commutativité et l'associativité de la somme nous avons, pour tout $a \in A$:

$$a = b + c(a) = b + \big(d(a) + 1\big) = (b + 1) + d(a) = s(b) + d(a). \tag{1.66}$$

Donc $s(b) \leqslant a$ pour tout $a \in A$. Mais comme $s(b) \in A$, l'élément $s(b)$ est bien un minimum de A.

(2f) **Unicité** Si a et a' sont des minimums de A, alors $a \leqslant a'$ et $a' \leqslant a$. Nous en déduisons que $a = a'$.

(3) **Pour (3)** Si $M \in \mathbb{N}$ majore tous les éléments de $\mathbb{N}$, alors en particulier $M \geqslant s(M)$. Mais le lemme 1.58 nous indique que $M \leqslant s(M)$. Nous avons donc $s(M) = M$, c'est-à-dire $M = M + 1$. En utilisant la régularité de M [20], nous trouvons $0 = 1$, ce qui est impossible par le lemme 1.55.

20. Dit plus simplement : en simplifiant par M.

(4) **Pour (4)** Soit A une partie non vide et majorée de $\mathbb{N}$.

(4a) **L'ensemble B** Nous posons

$$B = \{n \in \mathbb{N} \backslash A \text{ tel que } a \leqslant n, \forall a \in A\}. \tag{1.67}$$

(4b) **B est non vide** Soit un majorant M de A : pour tout $a \in A$ nous avons $a \leqslant M$. Nous avons $s(M) \notin A$, parce que si $s(M)$ était dans A, ce serait un élément de A strictement plus grand que tout $a \in A$. Donc B est non vide parce qu'il contient $s(M)$.

(5) **Minimum** Puisque B est non vide, il possède un plus petit élément que nous notons b. Nous savons que $b \neq 0$ parce que sinon A serait vide. Il existe donc $c \in \mathbb{N}$ tel que $b = s(c)$.

(6) **$a \leqslant c$ pour tout $a \in A$** Comme $s(c) \in B$ nous avons $a < s(c)$ pour tout $a \in A$. Donc, par le lemme 1.60(3) nous avons $s(a) \leqslant s(c)$, c'est-à-dire $a + 1 \leqslant c + 1$. Par régularité nous avons $a \leqslant c$.

Nous avons prouvé que $a \leqslant c$ pour tout $a \in A$.

(7) **$c \in A$** Si c n'est pas dans A, alors il est dans B et il contredit la minimalité de b. Donc c est dans A.

(8) **Conclusion** L'élément c est dans A tout en étant plus petit que tout élément de A.

(9) **Unicité** Si x est un élément minium de A, alors nous avons $x \leqslant c$ parce que x est minimum et $c \leqslant x$ parce que c est minimum, et donc $x = c$.

□

LEMooKUWUooPLWelf

Lemme 1.65.
Toute partie finie non vide de $\mathbb{N}$ est majorée et minorée.

LEMooOEJOooOgaxzi

Lemme-Définition 1.66.
Si A est une partie de $\mathbb{N}$, il existe un unique élément $m \in \mathbb{N}$ tel que

$$\begin{cases} m \in A & \text{(1.68a)} \\ m \leqslant a \, \forall a \in A. & \text{(1.68b)} \end{cases}$$

Cet élément est noté $\min(A)$ *et nommé* ***minimum de*** *A.*

Si A est majoré, il existe un unique élément $M \in \mathbb{N}$ tel que

$$\begin{cases} M \in A & \text{(1.69a)} \\ M \geqslant a \, \forall a \in A. & \text{(1.69b)} \end{cases}$$

Cet élément est noté $\max(A)$ *et nommé* ***maximum de*** *A.*

Nous verrons dans le lemme 1.74 qu'une partie de $\mathbb{N}$ admet un maximum si et seulement si elle est finie.

LEMooYMRJooYIAhBb

Lemme 1.67 ([1])**.**
Quelques affirmations sur l'ordre dans $\mathbb{N}$.

ITEMooTLOIooTWNtod

(1) Il n'existe pas de $n \in \mathbb{N}$ tel que $n < 0$.

ITEMooPJKQooGfLCUM

(2) Si $a, b \in \mathbb{N}$ vérifient $a > b$, alors il n'existe pas de x dans $\mathbb{N}$ tel que $a + x = b$.

Démonstration. En deux parties.

(1) **Pour (1)** Nous savons par la proposition 1.64(1) que 0 est l'unique minimum de $\mathbb{N}$. Nous avons donc forcément $0 \leqslant n$. Si n vérifie de plus $n \leqslant 0$ alors nous avons $n = 0$ par symétrie de la relation d'ordre $\leqslant$. Il n'est donc pas possible d'avoir $n \neq 0$.

(2) **Pour (2)** Si $b \leqslant a$ il existe $c \in \mathbb{N}$ tel que $b + c = a$. Et comme $a \neq b$, c n'est pas nul et il existe $y \in \mathbb{N}$ tel que $c = s(y)$. Bref, nous avons

$$b + s(y) = a. \tag{1.70}$$

Si de plus il existe $x \in \mathbb{N}$ tel que $a + x = b$ nous aurions

$$a + x + s(y) = a. \tag{1.71}$$

Comme a est régulier pour l'addition [21], nous avons

$$x + s(y) = 0, \tag{1.72}$$

ce qui signifie, par le lemme 1.56 que $x = s(y) = 0$. Puisque s prend ses valeurs dans $\mathbb{N}\backslash\{0\}$, cela est impossible.

□

DEFooKBUFooLvMHrf

Définition 1.68.
Soient $a, b \in \mathbb{N}$ tels que $a \leqslant b$. Nous notons par $\{a, \dots, b\}$ l'ensemble

$$\{x \in \mathbb{N} \text{ tel que } a \leqslant x \leqslant b\}. \tag{1.73}$$

PROPooFYMJooWihvhk

Proposition 1.69.
Toute application $\mathbb{N} \to \mathbb{N}$ strictement croissante est injective.

Démonstration. Soit une application strictement croissante $f : \mathbb{N} \to \mathbb{N}$. Soient $a, b \in \mathbb{N}$ tels que $f(a) = f(b)$. Puisque l'ordre est total [22], nous supposons que $a \leqslant b$. Si $a = b$ nous avons terminé. Nous supposons donc que $a \neq b$, c'est-à-dire que $a < b$. Par stricte croissance nous avons alors $f(a) < f(b)$ qui signifie $f(a) \leqslant f(b)$ et $f(a) \neq f(b)$. Contradiction. Il n'existe donc pas de $a \neq b$ tels que $f(a) = f(b)$. L'application f est donc injective. □

LEMooFKLPooPrmeUU

Lemme 1.70 ([1])**.**
Si S n'est pas majoré dans $\mathbb{N}$, alors il existe une bijection $\mathbb{N} \to S$.

Démonstration. Nous considérons l'application suivante :

$$\begin{aligned} g : S &\to S \\ n &\mapsto \min\{x \in S \text{ tel que } x > n\}. \end{aligned} \tag{1.74}$$

Cette application est bien définie parce que tout partie non vide de $\mathbb{N}$ a un plus petit élément [23]. Maintenant nous définissons $f : \mathbb{N} \to S$ par

$$\begin{cases} f(0) = \min(S) & \text{(1.75a)} \\ f(n+1) = g\big(f(n)\big). & \text{(1.75b)} \end{cases}$$

C'est le théorème 1.48 qui nous permet de le faire. Nous montrons que f est bijective.

(1) **Injective** Nous avons

$$f(n+1) \in \{x \in S \text{ tel que } x > f(n)\}. \tag{1.76}$$

Donc f est strictement croissante. Elle est donc injective.

(2) **Surjective** Soit $a \in S$. Nous allons voir que a est dans l'image de f. Pour cela nous posons

$$A = \{x \in \mathbb{N} \text{ tel que } f(x) < a\}. \tag{1.77}$$

Cet ensemble est majoré par a. En effet si $x \in A$ nous avons $x \leqslant f(x) < a$. La partie A de $\mathbb{N}$ possède un maximum. Nous notons $M = \max(A)$. Ce M a deux propriétés intéressantes.

21. Proposition 1.54(3).
22. Proposition 1.63.
23. Proposition 1.64(2).

(2a) **D'abord** Puisque $M \in A$, nous avons $f(M) < a$. Une autre façon de dire cela est de dire que

$$a \in \{x \in S \text{ tel que } x > f(M)\}. \tag{1.78}$$

Or $f(M+1) = \min\{x \in S \text{ tel que } x > f(M)\}$. Donc $f(M+1) \leqslant a$.

(2b) **Ensuite** Puisque M est le maximum de A, $M+1$ majore A, c'est-à-dire que $f(M+1) \geqslant a$.

(2c) **Les deux ensemble** Nous avons prouvé que $f(M+1) \leqslant a$ et $f(M+1) \geqslant a$. Nous en déduisons, par le lemme 1.62, que $f(M+1) = a$.

□

NORMooQXASooMXqhjI

1.71.
Durant la preuve du lemme 1.70, nous n'avons pas été loin de prouver que

$$\big(\min(S), S, g\big) \tag{1.79}$$

est un triplet naturel.

Toute partie non bornée de $\mathbb{N}$ donne lieu à un triplet naturel.

DEFooAZAYooVjNzmy

Définition 1.72 ([1]).
Soit un ensemble muni d'une loi de composition interne $(A, +)$. Soit $n \in \mathbb{N}$ et $a \in A$. Nous définissons $n \times A$ par

$$\begin{cases} 0 \times a = 0 & \text{(1.80a)} \\ (n+1) \times a = n \times a + a. & \text{(1.80b)} \end{cases}$$

DEFooLCWLooYrToFv

Définition 1.73.
*Un ensemble totalement ordonné muni d'une loi de composition interne $(A, +, \leqslant)$ est **archimédien** si pour tout $x, y \in A$ avec $x > 0$, il existe $n \in \mathbb{N}$ tel que $n \times x \geqslant y$ (voir la définition 1.72).*

LEMooGQUWooYJQfJB

Lemme 1.74.
Une partie de $\mathbb{N}$ admet un maximum si et seulement si elle est finie.

1.3.4 Multiplication dans les naturels

PROPooBBQPooRgPOjf

Proposition-Définition 1.75.
Il existe une unique fonction $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ telle que

ITEMooNTUUooDAUVsV

(1) $f(a, 0) = 0$ pour tout $a \in \mathbb{N}$

ITEMooPPZZooQQabwn

(2) $f\big(a, s(b)\big) = f(a, b) + a$ pour tout $a, b \in \mathbb{N}$.

*Cette fonction est la **multiplication** et nous notons $f(a, b) = a \times b$, voire ab quand il n'y a pas d'ambigüité. Le nombre $a \times b$ est nommé le **produit** de a par b.*

Démonstration. En deux parties.

(1) **Fonctions définies par récurrence** Soit $a \in \mathbb{N}$. Par le théorème 1.48, il existe une unique application $f_a : \mathbb{N} \to \mathbb{N}$ telle que

EQSooWVCTooNTVjKU

$$\begin{cases} f_a(0) = 0 & \text{(1.81a)} \\ f_a\big(s(b)\big) = f_a(b) + a & \text{(1.81b)} \end{cases}$$

(2) **Existence** Nous considérons, pour chaque $a \in \mathbb{N}$ la fonction f_a définie par les conditions (1.81). En posant $f(a, b) = f_a(b)$, nous avons une application qui vérifie toutes les conditions.

(3) **Unicité** Soient des applications f et g vérifiant les propriétés demandées. Soit $a \in \mathbb{N}$. Nous pouvons définir $f_a : \mathbb{N} \to \mathbb{N}$ et $g_a : \mathbb{N} \to \mathbb{N}$ par $f_a(n) = f(a, b)$ et $g_a(b) = g(a, b)$.

Les applications f_a et g_a vérifient toutes deux les conditions (1.81), et sont donc égales : $f_a = g_a$ pour tout a. Donc $f = g$.

□

1.76.
Nous supposons que la lectrice connait déjà la priorité des opérations. Elle saura donc interpréter des expressions comme $a \times b + c$ comme voulant dire $(a \times b) + c$ sans que nous ayons à ajouter des parenthèses.

PROPooGHDOooFYRmon

Proposition 1.77 ([18, 1])**.**
La multiplication a les propriétés suivantes.

ITEMooHFWRooDCEpjj
(1) $n \times 1 = n$ *pour tout* $n \in \mathbb{N}$.

ITEMooRSYMooSUrRsl
(2) $1 \times n = n$ *pour tout* $n \in \mathbb{N}$.

ITEMooWJPOooRUYjwQ
(3) *La multiplication est commutative.*

ITEMooNBYKooXnGRrf
(4) $0 \times n = 0$ *pour tout* $n \in \mathbb{N}$

ITEMooLJQBooVpUxUv
(5) *L'élément* 1 *est neutre pour la multiplication.*

ITEMooDYLIooETIBEL
(6) *La multiplication est distributive par rapport à l'addition.*

ITEMooQBFSooWGDQYX
(7) *La multiplication est associative.*

Démonstration. Point par point.

(1) **Pour (1)** Nous avons $n \times 1 = n \times s(0) = (n \times 0) + n = 0 + n = n$. Donc $n \times 1 = n$.

(2) **Pour (2)** Nous le faisons par récurrence. Par définition c'est vrai pour $n = 0$. En ce qui concerne la récurrence, nous supposons que $1 \times n = n$, et nous prouvons que $1 \times s(n) = s(n)$:

$$1 \times s(n) = (1 \times n) + 1 = n + 1 = s(n). \tag{1.82}$$

(3) **Pour (3)** Soit $a \in \mathbb{N}$. Nous prouvons par récurrence sur $b \in \mathbb{N}$ que $a \times b = b \times a$. Pour $b = 0$ c'est bon. Pour la récurrence, nous supposons que $a \times b = b \times a$ et nous prouvons que $a \times s(b) = s(b) \times a$:

$$\begin{aligned}
a \times s(b) &= a \times b + a && && \text{(1.83a)}\\
&= b \times a + a && \text{récurrence} && \text{(1.83b)}\\
&= (b \times a) + (1 \times a) && \text{par (2)} && \text{(1.83c)}\\
&= (b + 1) \times a && \text{distributivité} && \text{(1.83d)}\\
&= s(b) \times a. && && \text{(1.83e)}
\end{aligned}$$

(4) **Pour (4)** C'est vrai pour $n = 0$ par la définition 1.75(1). En ce qui concerne $s(n)$, nous avons

$$0 \times s(n) = (0 \times n) + 0 = 0. \tag{1.84}$$

(5) **Pour (5)** C'est la combinaison de (1) et (2).

(6) **Pour (6)** Soient $a, b \in \mathbb{N}$. Nous prouvons par récurrence sur $c \in \mathbb{N}$ que[24]

$$(a + b) \times c = a \times c + b \times c. \tag{1.85}$$

Pour $c = 0$, nous avons $(a + b) \times 0 = 0$ ainsi que $a \times 0 = b \times 0 = 0$ en vertu des points précédents sur la multiplication par zéro. Pour la récurrence nous utilisons associativité et commutativité de la somme :

$$\begin{aligned}
(a + b) \times s(c) &= (a + b) \times c + (a + b) && \text{(1.86a)}\\
&= a \times c + b \times c + a + b && \text{(1.86b)}\\
&= (a \times c + a) + (b \times c + b) && \text{(1.86c)}\\
&= \big(a \times s(c)\big) + \big(b \times s(c)\big). && \text{(1.86d)}
\end{aligned}$$

24. Nous n'écrivons pas toutes les parenthèses parce que les règles de priorité des opérations sont supposées connues. J'invite cependant la lectrice à remarquer qu'une formalisation de ces règles n'est probablement pas facile. Pour que tout soit rigoureux, il faudrait un algorithme qui parcourt une suite de caractères et l'interprète en ajoutant correctement les parenthèses.

(7) **Pour (7)** Soient $a, b \in \mathbb{N}$. Nous démontrons par récurrence sur $c \in \mathbb{N}$ que $(a \times b) \times c = a \times (b \times c)$. Pour $c = 0$ l'égalité est triviale. Nous supposons que l'égalité est correcte pour c, et nous la prouvons pour $s(c)$:

$$\begin{aligned}
(a \times b) \times s(c) &= \big((a \times b) \times c\big) + a \times b && && (1.87a)\\
&= \big(a \times (b \times c)\big) + a \times b && \text{récurrence} && (1.87b)\\
&= \big((b \times c) \times a\big) + b \times a && \text{commutativité} && (1.87c)\\
&= (b \times c + b) \times a && \text{distributivité} && (1.87d)\\
&= \big(b \times s(c)\big) \times a && \text{définition 1.75(2)} && (1.87e)\\
&= a \times \big(b \times s(c)\big) && \text{commutativité.} && (1.87f)
\end{aligned}$$

□

Lemme 1.78 ([1]).
La multiplication est compatible avec l'ordre :

$$a \leqslant b \Rightarrow a \times n \leqslant b \times n \tag{1.88}$$

pour tout $n \in \mathbb{N}$.

Démonstration. Par la définition 1.57 de l'ordre, si $a \leqslant b$, il existe $c \in \mathbb{N}$ tel que $b = a + c$. En utilisant la distributivité [25], nous avons

$$b \times n = (a + c) \times n = a \times n + c \times n. \tag{1.89}$$

Nous en déduisons que $a \times n \leqslant b \times n$ parce que $c \times n \in \mathbb{N}$. □

LEMooEHYEooLDudfn

Lemme 1.79.
Si $a \times b = 0$, alors $a = 0$ ou $b = 0$ (ou les deux).

Démonstration. Supposons que $a \neq 0$. Alors il existe $c \in \mathbb{N}$ tel que $a = s(c)$. Nous avons

$$0 = a \times b = s(c) \times b = c \times b + b. \tag{1.90}$$

Le lemme 1.56 nous dit alors que $c \times b = b = 0$. □

LEMooGUXGooBcKJdS

Lemme 1.80 ([1]).
Si $a < b$ et si $n \neq 0$, alors

$$a \times n < b \times n. \tag{1.91}$$

Démonstration. L'hypothèse $a \leqslant b$ implique qu'il existe $c \in \mathbb{N}$ tel que $a + c = b$. De plus $c \neq 0$ parce que $a \neq b$. En utilisant la distributivité [26], nous avons

$$b \times n = (a + c) \times n = (a \times n) + (c \times n). \tag{1.92}$$

Cela prouve que $a \times n \leqslant b \times n$. Et comme c et n ne sont pas nuls, nous avons même [27] $c \times n \neq 0$ et donc $a \times n < b \times n$. □

Une version dans $\mathbb{Z}$ sera le lemme 1.110.

LEMooSFUKooBNAple

Lemme 1.81 ([1]).
Soient $a \neq 0$ et $b > 1$ dans $\mathbb{N}$. Alors

$$ab > a. \tag{1.93}$$

25. Proposition 1.77(6).
26. Proposition 1.77(6).
27. Lemme 1.79.

Démonstration. Il s'agit d'une application du lemme 1.80 en partant de l'inégalité $1 < b$ et en la « multipliant » par a. □

Proposition 1.82 ([1]).
Tous les naturels non nuls sont réguliers par rapport à la multiplication. Autrement dit, si $a \neq 0$, alors nous avons

$$a \times x = a \times y \Rightarrow x = y. \tag{1.94}$$

Démonstration. Soit $a \neq 0$ dans $\mathbb{N}$. Nous supposons que $a \times x = a \times y$. Puisque l'ordre sur $\mathbb{N}$ est total (proposition 1.63), nous pouvons supposer que $y \geqslant x$; sinon il suffit de permuter les rôles de x et y dans tout ce qui suit.

Il existe $d \in \mathbb{N}$ tel que $y = x + d$. En utilisant l'hypothèse $a \times y = a \times x$ et la distributivité [28],

$$a \times x = a \times y = a \times (x + d) = (a \times x) + (a \times d). \tag{1.95}$$

Puisque $(a \times x)$ est régulier pour la somme [29] nous en déduisons que

$$0 = a \times d. \tag{1.96}$$

Le lemme 1.79 dit alors que $a = 0$ ou que $d = 0$. Étant donné que $a \neq 0$ par hypothèse, nous déduisons que $d = 0$, c'est-à-dire que $x = y$. □

1.3.5 Presque unicité des triplets naturels

Il existe de nombreux triplets naturels ; l'existence d'un triplet naturel est un théorème de la théorie des ensembles que nous avons accepté. Nous avons déjà à peu près montré que toute partie non bornée de $\mathbb{N}$ donne lieu à un nouveau triplet naturel. Voir 1.71.

Nous voyons maintenant que tous les triplets naturels sont équivalents au moins pour l'ordre.

THOooFUXMooJuigHK

Théorème 1.83 ([1]).
Soient des triplets naturels $(\mathcal{N}_1, o_1, s_1)$ et $(\mathcal{N}_2, o_2, s_2)$. Alors

(1) il existe une unique application $f : \mathcal{N}_1 \to \mathcal{N}_2$ telle que
   (1a) $f(o_1) = o_2$
   (1b) $f \circ s_1 = s_2 \circ f$.
(2) Une telle application est une bijection croissante.

Démonstration. En plusieurs points.

(1) **Existence** Nous voyons $(\mathcal{N}_1, o_1, s_1)$ comme un triplet naturel, et $\mathcal{N}_2$ comme un simple ensemble. Nous pouvons appliquer le théorème 1.48 à $(\mathcal{N}_1, o_1, s_1)$. L'élément o_1 va jouer le rôle de 0 alors que o_2 va jouer le rôle de b. L'application g est s_2. Bref, il existe une unique application $f : \mathcal{N}_1 \to \mathcal{N}_2$ telle que
   (1a) $f(o_1) = o_2$
   (1b) $f\big(s_1(n)\big) = s_2\big(f(n)\big)$
   pour tout $n \in \mathcal{N}_1$.

(2) **Unicité** Le théorème 1.48 donne déjà l'unicité. Nous la faisons quand même, juste pour vous faire plaisir. Soit g, une autre application vérifiant les mêmes conditions. Pour faire la récurrence de façon très explicite, nous posons

$$\begin{aligned} P : \mathcal{N}_1 &\to \{0, 1\} \\ x &\mapsto \begin{cases} 1 & \text{si } g(x) = f(x) \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.97}$$

28. Proposition 1.77(6).
29. Proposition 1.54(3).

Notre but est de prouver que $P(x) = 1$ pour tout $x \in \mathcal{N}_1$, en utilisant la récurrence telle que décrite dans la proposition 1.47.

Nous avons $f(o_1) = o_2 = g(o_1)$. Donc $P(o_1) = 1$. Nous supposons que, pour un certain $a \in \mathcal{N}_1$, nous ayons $P(a) = 1$, et nous prouvons que $P\big(s_1(a)\big) = 1$.

Nous avons $g(a) = f(a)$, et nous prenons s_2 des deux côtés, nous avons succéssivement

$$(s_2 \circ g)(a) = (s_2 \circ f)(a) \tag{1.98a}$$
$$(g \circ s_1)(a) = (f \circ s_1)(a) \tag{1.98b}$$
$$g\big(s_1(a)\big) = f\big(s_1(a)\big). \tag{1.98c}$$

La dernière égalité signifie que $P\big(s_1(a)\big) = 1$. La proposition 1.47 implique que $P(x) = 1$ pour tout $x \in \mathcal{N}_1$.

(3) **Bijection, définir l'inverse** Nous allons trouver un inverse et le lemme 1.7 nous dit que c'est suffisant. La partie « existence », en inversant les rôles de $\mathcal{N}_1$ et $\mathcal{N}_2$ nous donne une application $g\colon \mathcal{N}_2 \to \mathcal{N}_1$ telle que

(3a) $g(o_2) = o_1$

(3b) $g \circ s_2 = s_1 \circ g$.

Nous allons prouver que g est un inverse de f.

(4) $f \circ g = \mathrm{Id}$ Nous posons $A = \{x \in \mathcal{N}_2 \text{ tel que } (f \circ g)(x) = x\}$. Nous avons

$$f\big(g(o_2)\big) = f(o_1) = o_2, \tag{1.99}$$

et donc $o_2 \in A$.

Supposons que $x \in A$. Alors

$$(f \circ g)\big(s_2(x)\big) = (f \circ \underbrace{g \circ s_2}_{s_1 \circ g})(x) \tag{1.100a}$$
$$= (\underbrace{f \circ s_1}_{= s_2 \circ f} \circ g)(x) \tag{1.100b}$$
$$= (s_2 \circ f \circ g)(x) \tag{1.100c}$$
$$= s_2\big((f \circ g)(x)\big) \tag{1.100d}$$
$$= s_2(x) \tag{1.100e}$$

Donc $s_2(x) \in A$. Nous en déduisons que $A = \mathcal{N}_2$ par le point (3) de la définition 1.44 d'un triplet naturel.

(5) $g \circ f = \mathrm{Id}$ J'imagine que c'est la même chose que dans l'autre sens (ci-dessus) [30].

□

PROPooCCVNooYUYcqG

Proposition 1.84.
L'ensemble structuré $(\mathbb{N}, +, \times, \leqslant)$ est archimédien[31]*. En d'autres termes, pour tout $a, b \in \mathbb{N}\backslash\{0\}$, il existe $n \in \mathbb{N}$ tel que*

$$b < n \times a. \tag{1.101}$$

Démonstration. Soient $a, b \in \mathbb{N}$ avec $a \neq 0$.

Si $a > b$, nous avons le résultat avec $n = 1$.

Si $a = b$, en prenant $n = s(1)$ nous avons le résultat. En effet $s(1) \times a = a + a$. Puisque $a \neq 0$, nous avons $a + a \geqslant a$ et $a + a \neq a$, donc $s(1) \times a > a$.

La vraie vie est avec $a < b$. Nous posons

$$X = \{x \in \mathbb{N} \text{ tel que } 1 \leqslant x \times a \leqslant b\} \tag{1.102}$$

30. Je n'ai pas fait les calculs ; écrivez-moi si ça pose un problème.
31. Définition 1.73.

et

$$B = \{x \times a \text{ tel que } x \in X\}. \tag{1.103}$$

L'ensemble X est non vide parce que $1 \in X$. L'ensemble B est alors également non vide, et majoré par b. La proposition 1.64(4) nous indique alors que B possède un plus grand élément que nous allons noter $x_0 \times a$ ($x_0 \in X$).

Nous posons $n = s(x_0)$, et nous avons

$$x_0 \times a < x_0 \times a + a = s(x_0) \times a = n \times a. \tag{1.104}$$

Nous en déduisons que $n \times a$ n'est pas dans B parce que $x_0 \times a$ est le plus grand élément de B. Donc x_0 n'est pas dans X ; nous n'avons donc pas les inégalités

$$1 \leqslant n \times a \leqslant b. \tag{1.105}$$

Laquelle des deux inégalités est fausse ? Puisque $n = s(x_0) \geqslant 1$ et que $a \geqslant 1$, nous avons $1 \leqslant n \times a$. Donc c'est la seconde inégalité qui est fausse. Nous avons donc $n \times a > b$. □

DEFooNEVNooJlmJOC

Proposition-Définition 1.85.
Soit A un ensemble muni d'une loi de composition interne[32] *notée $+$. Si nous avons une application $\alpha : \mathbb{N} \to A$, alors il existe une unique application $f : \mathbb{N} \to A$ telle que* SUBEQSooKJOFooOOxDKH

$$\begin{cases} f(0) = \alpha(0) & \text{(1.106a)} \\ f(k) = f(k-1) + \alpha(k). & \text{(1.106b)} \end{cases}$$

La valeur de $f(k)$ est notée $\sum_{i=0}^{k} \alpha(i)$.

Démonstration. Application du théorème 1.48. □

1.86.
Les équations de définition (1.106) signifient que nous définissons la notation $\sum_{i=0}^{N} \alpha(i)$ par récurrence par les conditions suivantes : ITEMooIPDTooEhOxea

(1) $\sum_{i=0}^{0} \alpha(i) = \alpha(0)$,
(2) $\sum_{i=0}^{k} \alpha(i) = \sum_{i=0}^{k-1} \alpha(i) + \alpha(k)$.

PROPooXXGHooLafGsI

Proposition 1.87 (La multiplication est une somme itérée[18])**.**
Pour tout $a, b \in \mathbb{N}$, nous avons

$$\sum_{i=1}^{n} a = a \times n. \tag{1.107}$$

Démonstration. Nous le faisons par récurrence en partant de $n = 1$. Avec $n = 1$ nous avons $\sum_{i=1}^{1} a = a$, et $a \times 1 = a$. Donc c'est bon.

Pour la récurrence nous avons :

$$a \times s(n) = a \times n + a = \sum_{i=1}^{n} a + a = \sum_{i=1}^{n+1} a = \sum_{i=1}^{s(n)} a. \tag{1.108}$$

□

LEMooIETGooMyrilW

Lemme 1.88.
Soit $a > 1$ dans $\mathbb{N}$. Pour tout $n \geqslant 1$ nous avons $na \leqslant a^n$.

32. Peut-être un anneau, mais comme nous avons l'intention, dans les propositions 1.87 et suivantes, de faire des sommes vers $(\mathbb{N}, +)$, plutôt un monoïde.

Démonstration. Par récurrence. Avec $n = 1$ nous avons bien $a \leqslant a$; pas de problème. Supposons que $na \leqslant a^n$, et montrons le pas de récurrence. Nous avons :

$$\begin{aligned}
(n+1)a &= na + a && && (1.109a)\\
&\leqslant na + na && \text{parce que } a \leqslant na && (1.109b)\\
&= 2na && && (1.109c)\\
&\leqslant 2a^n && \text{récurrence} && (1.109d)\\
&\leqslant aa^n && \text{parce que } a \geqslant 2 && (1.109e)\\
&= a^{n+1}. && && (1.109f)
\end{aligned}$$

□

Proposition 1.89 ([18]).
Soit $a > 1$. Alors

(1) l'application $n \mapsto a^n$ est strictement croissante ;

(2) l'ensemble $\{a^n \text{ tel que } n \in \mathbb{N}\}$ n'est pas majoré.

Démonstration. Nous avons

$$\begin{aligned}
a^{n+1} &= a^n \times a && && (1.110a)\\
&> a^n \times 1 && \text{lemme 1.81} && (1.110b)\\
&= a^n. && && (1.110c)
\end{aligned}$$

Cela prouve le premier point.

Pour le second point, soit $m \in \mathbb{N}$. Nous devons trouver $N \in \mathbb{N}$ tel que $a^N \geqslant m$. Puisque $\mathbb{N}$ est archimédien [33], nous pouvons considérer N tel que $Na > m$. Le lemme 1.56 nous assure alors que

$$m < Na \leqslant a^N. \tag{1.111}$$

□

THOooKDJVooRIJRHP

Théorème 1.90 (division euclidienne [18]).
Pour tout $a \in \mathbb{N}$, pour tout $b \in \mathbb{N}\backslash\{0\}$, il existe un unique couple $(q, r) \in \mathbb{N}^2$ tel que $a = bq + r$ avec $0 \leqslant r < b$.

*Si $r = 0$, nous disons que a est **divisible** par b.*

Démonstration. Existence puis unicité.

(1) **Existence**

Nous posons

$$A = \{bx \text{ tel que } x \in \mathbb{N}, bx \leqslant a\}. \tag{1.112}$$

L'ensemble A contient 0 (avec $x = 0$) et est majoré par a. Donc il possède un plus grand élément que nous notons bq. Puisque $bq \in A$, nous avons $bq \leqslant a$ et donc il existe $r \in \mathbb{N}$ tel que

$$bq + r = a. \tag{1.113}$$

EQooIUICooLenNBP

Il reste à montrer que $r < b$. Supposons que $r \geqslant b$. Il existerait alors un x tel que $b + x = r$. En mettant ça dans (1.113),

$$bq + b + x = a, \tag{1.114}$$

c'est-à-dire $b(q + 1) + x = a$, qui signifierait $b(q + 1) \leqslant a$, ce qui est faux parce que bq est le plus grand élément de A.

33. Proposition 1.84.

(2) **Unicité** Supposons que nous ayons

$$a = bq + r = bq' + r' \tag{1.115}$$

avec $0 \leqslant r < b$ et $0 \leqslant r' < b$. Il y a trois possibilités : $q' < q$, $q' = q$ et $q' > q$.

(2a) **Si $q' < q$** Alors il existe $x \in \mathbb{N}$ tel que $q' + x = q$, et nous avons

$$b(q' + x) + r = bq' + bx + r, \tag{1.116}$$

ce qui, après distribution et simplification, donne $r' = bx + r$. Puisque nous avons $x \geqslant 1$, il vient

$$r' = bx + r \geqslant b + r \geqslant b. \tag{1.117}$$

Cela n'est pas possible parce que $r' < b$. Le cas $q' < q$ n'est pas possible.

(2b) **Si $q' = q$** Nous avons alors immédiatement $bq + r = bq + r'$ et donc $r = r'$. Unicité.

(2c) **Si $q' > q$** En posant $q + x = q'$ nous trouvons la même impossibilité que dans le cas $q' < q$.

□

1.3.6 Écriture d'un naturel dans une base

1.91.
Nous avons déjà donné la notation $1 = s(0)$. Nous continuons avec $2 = s(1)$, $3 = s(2)$, $4 = s(3)$, $5 = s(4)$, $6 = s(5)$, $7 = s(6)$, $8 = s(7)$ et $9 = s(8)$.

Nous allons maintenant voir comment écrire des nombres plus grands.

Si $b > 1$ et $N \in \mathbb{N}$ sont donnés, nous notons

$$C_{b,N} = \big\{u \in \{0, \dots, b-1\}^{N+1} \text{ tel que } u_N \neq 0\big\}. \tag{1.118}$$

où les u_i sont numérotés à partir de 0 ; donc dire $u_N \neq 0$ revient à dire que le *dernier* est non nul, et non l'avant dernier. Nous définissons [34]

$$\begin{aligned} \varphi_{b,N} \colon C_{b,N} &\to \mathbb{N} \\ u &\mapsto \sum_{i=0}^{N} u_i b^i. \end{aligned} \tag{1.119}$$

EQooWWTUooHAnSEv

Cette application $\varphi_{b,N}$ sera encore bien étudiée pour la partie décimale d'un réel. Voir la définition 10.370.

LEMooJUGKooGsbrhi

Lemme 1.92 ([18]).
Soient $b > 1$, $N \geqslant 0$ ainsi que $u \in C_{b,N}$. Alors

$$b^N \leqslant \varphi_{b,N}(u) < b^{N+1}. \tag{1.120}$$

EQooYHTLooNwqIIq

Démonstration. En séparant la somme nous avons

$$\varphi_{b,N}(u) = u_N b^N + \sum_{i=0}^{N-1} u_i b^i. \tag{1.121}$$

Puisque $u_N \geqslant 1$ nous avons $b^N \leqslant u_N b^N$, et donc

$$b^N \leqslant u_N b^N \leqslant \varphi_{b,N}(u). \tag{1.122}$$

Voilà qui prouve la première inégalité de (1.120).

Pour prouver que $\varphi_{b,N}(u) < b^{N+1}$, nous faisons une récurrence sur N.

34. Le symbole de sommation est défini par 1.85.

(1) **Pour $N = 0$** Nous devons prouver que $\varphi_{b,0}(u) < b$. Par définition $\varphi_{b,N}(u) = u_0 b^0$. Puisque $u \in \{0, \dots, b-1\}^{N+1}$, nous avons $u_0 \leqslant b - 1 < b$.

(2) **Récurrence** Nous supposons que pour tout $u \in C_{b,N}$ nous avons $\varphi_{b,N}(u) < b^{N+1}$. Et nous devons montrer que pour tout $v \in C_{b,N+1}$ nous avons $\varphi_{b,N+1}(v) < b^{N+2}$.
Nous posons $u = (v_0, \dots, v_N)$; nous avons alors

$$\begin{aligned}
\varphi_{b,N+1}(v) &= v_{N+1}b^{N+1} + \sum_{i=0}^{N} v_i b^i && && (1.123a)\\
&= v_{N+1}b^{N+1} + \varphi_{b,N}(u) && && (1.123b)\\
&< v_{N+1}b^{N+1} + b^{N+1} && \text{récurrence} && (1.123c)\\
&= (v_{N+1}+1)b^{N+1} && && (1.123d)\\
&\leqslant bb^{N+1} && \text{parce que } v_{N+1} \leqslant b-1 && (1.123e)\\
&= b^{N+2}. && && (1.123f)
\end{aligned}$$

□

LEMooKDKJooSkhcJS

Lemme 1.93 ([1]).
Soient $x \in \mathbb{N}$ ainsi que $b \geqslant 2$. Nous posons

$$N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.124}$$

Alors

(1) Si $n > N$ alors $\varphi_{b,n}(u) > x$ pour tout $u \in C_{b,n}$.

(2) Si $n < N$ alors $\varphi_{b,n}(u) < x$ pour tout $u \in C_{b,n}$.

Démonstration. En deux parties.

(1) **Si $n > N$** Nous avons, par définition de $C_{b,n}$ que $u_n \neq 0$, de telle sorte que

$$\varphi_{b,n}(u) \geqslant u_n b^n \geqslant b^n > x. \tag{1.125}$$

La dernière inégalité est due au fait que $n \notin \{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}$.

(2) **Si $n < N$** Nous avons

$$x \geqslant b^N > \varphi_{b,n}(u). \tag{1.126}$$

Le seconde inégalité est une conséquence du lemme 1.92.

□

Théorème 1.94 ([18]).
Soit $b \geqslant 2$. Si $x \in \mathbb{N}$, alors il existe un unique $N \in \mathbb{N}$ et un unique $u \in C_{b,N}$ tels que

$$x = \varphi_{b,N}(u). \tag{1.127}$$

Démonstration. Nous commençons par $x < b$. Dans ce cas, $N = 0$ parce que si $u_k \neq 0$ avec $k \neq 0$, nous avons

$$\sum_{i=0}^{N} u_i b^i \geqslant u_k b^k \geqslant b > x. \tag{1.128}$$

Donc $x = x_0 b^0 = u_0$. Bref, dans le cas $x < b$ nous avons obligatoirement $N = 0$ et $u_0 = x$.

Nous étudions à présent le cas $x \geqslant b$ que nous subdivisons en plusieurs étapes.

(1) $N \geqslant 1$ Si $N = 0$, alors $\varphi_{b,0}(u) = u_0 < b \leqslant x$. Donc $N \geqslant 1$.
Notons incidemment que nous pouvons parler de $N - 1$ à partir de maintenant.

(2) **Unicité, préambule** Le lemme 1.93 nous indique que si $x = \varphi_{b,N}(u)$ pour un certain $N \in \mathbb{N}$ et un certain $u \in C_{b,N}$, alors

$$N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.129}$$

Nous posons

$$X_k = \sum_{i=k}^{N} u_i b^{i-k}, \tag{1.130}$$

et nous allons montrer que le couple (X_{k+1}, u_k) est le résultat de la division euclidienne[35] de X_k par b.

D'abord, $u_k < b$, donc ça a bien la tête d'un reste. Ensuite, pour le quotient,

$$\begin{aligned} bX_{k+1} + u_k &= b \sum_{i=k+1}^{N} u_i b^{i-(k+1)} + u_k && (1.131a)\\ &= \sum_{i=k+1}^{N} u_i b^{i-k} + u_k && (1.131b)\\ &= \sum_{i=k}^{N} u_i b^{i-k} && (1.131c)\\ &= X_k. && (1.131d) \end{aligned}$$

(3) **Unicité** En quoi cela fait-il avancer la choucroute ? Supposons que $\varphi_{b,N}(u) = \varphi_{b,M}(v)$. Alors nous avons déjà prouvé que

$$M = N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.132}$$

Ensuite nous devons montrer que $u = v$. Nous posons $X_k = \sum_{i=k}^{N} u_i b^{i-k}$ et $Y_k = \sum_{i=k}^{N} v_i b^{i-k}$. Notez que

$$X_0 = Y_0 = x. \tag{1.133}$$

Si $X_k = Y_k$, alors par unicité de la division euclidienne nous avons $X_{k+1} = Y_{k+1}$ et $u_k = v_k$. Par récurrence nous avons $X_k = Y_k$ et $u_k = v_k$ pour tout k.

(4) **Existence** Soit $x \in \mathbb{N}$. Nous posons $y_0 = x$ et

$$y_k = by_{k+1} + u_k \tag{1.134}$$

avec $u_k < b$. Vus l'unicité dans la division euclidienne et le théorème[36] 1.48 permettant la définition par récurrence, ces conditions définissent deux suites (u_k) et (y_k) dans $\mathbb{N}$.

Montrons qu'il existe un $N \in \mathbb{N}$ tel que $y_n = 0$ pour tout $n \geqslant N + 1$. Nous avons :

$$\begin{aligned} 2y_{k+1} &\leqslant by_{k+1} && \text{parce que } b \geqslant 2 && (1.135a)\\ &\leqslant y_k && \text{pcq } by_{k+1} + u_k = y_k. && (1.135b) \end{aligned}$$

Bref : $2y_{k+1} \leqslant y_k$. Par récurrence[37] nous trouvons que

$$2^k y_k \leqslant x \tag{1.136}$$

parce que $y_0 = x$. Par le lemme 1.88, si k est assez grand,

$$2ky_k \leqslant 2^k y_k \leqslant x. \tag{1.137}$$

35. Théorème 1.90.
36. Nous ne citerons pas toujours ce théorème à chaque fois que nous définissons quelque chose par récurrence.
37. Faut-il citer la proposition 1.47 et donner explicitement la fonction P ?

Puisque $\mathbb{N}$ est archimédien[38], nous pouvons considérer $s \in \mathbb{N}$ tel que $2s > x$. À ce moment nous avons

$$y_n = 0 \tag{1.138}$$

pour tout $n \geqslant s$. Nous posons

$$N = \max\{k \text{ tel que } y_k \neq 0\}. \tag{1.139}$$

Prouvons par récurrence sur l que

$$y_{N-l} = \sum_{i=N-l}^{N} u_i b^{(i+l)-N}. \tag{1.140}$$

EQooZBKQooFqcckr

Notez que $i + l \geqslant N - l + l = N$, donc $(i + l) - N$ a un sens.

(4a) **Pour $l = 0$** Avec $l = 0$ nous avons $\sum_{i=N-l}^{N} u_i b^{(i+l)-N} = u_N$. Il faut donc voir que $y_N = u_N$. Nous avons

$$y_N = by_{N+1} + u_N. \tag{1.141}$$

En se rappelant que $y_{N+1} = 0$, nous avons le résultat.

(4b) **Pour $l + 1$** Pour la récurrence nous avons le calcul suivant :

$$\begin{aligned} y_{N-l-1} &= by_{N-l} + u_{N-l-1} && \text{(1.142a)}\\ &= b \sum_{i=N-l}^{N} u_i b^{(i+l)-N} + u_{N-l-1} && \text{(1.142b)}\\ &= \sum_{i=N-l}^{N} u_i b^{(i+l)-N+1} + u_{N-l-1} && \text{(1.142c)}\\ &= \sum_{i=N-l-1}^{N} u_i b^{(i+l+1)-N}. && \text{(1.142d)} \end{aligned}$$

La récurrence est prouvée. L'égalité (1.140) est validée pour tout l.

En posant $l = N$ dans (1.140) nous trouvons

$$y_0 = \sum_{i=0}^{N} u_i b^i. \tag{1.143}$$

Mais la définition de la suite (y_k) contient $y_0 = x$. Donc nous avons prouvé que

$$x = \sum_{i=0}^{N} u_i b^i = \varphi_{b,N}(u). \tag{1.144}$$

□

Exemple 1.95.
Comment écrire le nombre b en base b ? Nous devons trouver un N et une suite (u_i) tels que

$$b = \sum_{i=0}^{N} u_i b^i. \tag{1.145}$$

Il est facile de voir que le choix $N = 1$ et $u = (0, 1)$ fonctionne bien : $b = 1 \times b^1 + 0$. Nous avons donc

$$b = \varphi_{b,1}(1, 0). \tag{1.146}$$

Nous écrivons cela plus sobrement $b = 10$. △

38. Proposition 1.84.

1.96.
À part des cas très exceptionnels, nous utilisons toujours la base $b = s(9) = s^9(0)$. Nous nous permettons donc d'écrire « 64 » le nombre $\varphi_{s(9),2}(6,4)$. Vous saviez que tout groupe simple d'ordre $\varphi_{s(9),2}(6,0)$ est isomorphe au groupe alterné $A_{\varphi_{s(9),0}(5)}$? C'est la proposition 5.55.

La proposition suivante dit que le nombre qui a le plus de chiffres est le plus grand.

Proposition 1.97 ([18])**.**
Si $u \in C_{b,N}$ et $v \in C_{b,M}$ avec $M > N$ alors $\varphi_{b,N}(u) < \varphi_{b,M}(v)$.

Démonstration. Le lemme 1.92 nous dit que

$$b^N \leqslant \varphi_{b,N}(u) < b^{N+1} \tag{1.147}$$

et

$$b^M \leqslant \varphi_{b,M}(v) < b^{M+1}. \tag{1.148}$$

Puisque $M > N$ nous avons $b^{N+1} \leqslant b^M$ et donc

$$\varphi_{b,N}(u) < b^{N+1} \leqslant b^M \leqslant \varphi_{b,M}(v). \tag{1.149}$$

□

La proposition suivante dit que si deux nombres s'écrivent avec le même nombre de chiffres, le plus grand est celui dont le premier chiffre différent est le plus grand. Autrement dit, les nombres en écriture de position se classent par ordre lexicographique.

Proposition 1.98.
Soient $u, v \in C_{b,N}$ tels que $u_i = v_i$ pour $i = r+1, \dots, N$. Si $u_r > v_r$ alors $\varphi_{b,N}(u) > \varphi_{b,N}(v)$.

Démonstration. En découpant les sommes nous avons

$$\varphi_{b,N}(u) = \sum_{i=r+1}^{N} u_i b^i + u_r b^r + \sum_{i=0}^{r-1} u_i b^i \tag{1.150}$$

et

$$\varphi_{b,N}(v) = \sum_{i=r+1}^{N} u_i b^i + v_r b^r + \sum_{i=0}^{r-1} v_i b^i. \tag{1.151}$$

Puisque $b^r > \sum_{i=0}^{r-1} u_i b^i$ (lemme 1.92), nous avons aussi

$$b^r + \sum_{i=0}^{r-1} u_i b^i > \sum_{i=0}^{r-1} v_i b^i. \tag{1.152}$$

EQooTZPBooTeauhX

Et le calcul final :

$$\begin{aligned}
\varphi_{b,N}(v) &< \sum_{i=r+1}^{N} v_i b^i + v_r b^r + b^r + \sum_{i=0}^{r-1} u_i b^i && \text{pcq (1.152)} && \text{(1.153a)}\\
&= \sum_{i=r+1}^{N} u_i b^i + (v_r+1) b^r + \sum_{i=0}^{r-1} u_i b^i && && \text{(1.153b)}\\
&\leqslant \sum_{i=r+1}^{N} u_i b^i + u_r b^r + \sum_{i=0}^{r-1} u_i b^i && \text{pcq } u_r \geqslant v_r + 1 && \text{(1.153c)}\\
&= \sum_{i=0}^{N} u_i b^i && && \text{(1.153d)}\\
&= \varphi_{b,N}(u). && && \text{(1.153e)}
\end{aligned}$$

Et voilà. □

1.4 Les entiers

PROPooFIKUooVHlvTt

Proposition-Définition 1.99 ([18]).
Soient $a, b, a', b' \in \mathbb{N}$. Nous disons que $(a, b) \sim (a', b')$ si et seulement si

$$a + b' = b + a' \tag{1.154}$$

ITEMooZQSHooSDfdvK

(1) $\sim$ est une relation d'équivalence sur $\mathbb{N}^2$.

(2) Si $(a, b) \sim (a', b')$ et $(x, y) \sim (x', y')$ alors

$$(a + x, b + y) \sim (a' + x', b' + y'). \tag{1.155}$$

*L'ensemble des **entiers** est*

$$\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim, \tag{1.156}$$

et nous notons $\overline{a, b} \in \mathbb{Z}$ la classe de $(a, b) \in \mathbb{N} \times \mathbb{N}$.

Démonstration. En plusieurs points.

(1) **Symétrie** C'est la commutativité de la somme dans $\mathbb{N}$, proposition 1.54(1).

(2) **Réflexive** Immédiat.

(3) **Transitive** Nous supposons que $(a, b) \sim (u, v)$ et que $(u, v) \sim (x, y)$. Alors nous avons

$$a + v = u + b \tag{1.157a}$$
$$u + y = v + x. \tag{1.157b}$$

En additionnant membre à membre,

$$a + v + u + y = u + b + v + x. \tag{1.158}$$

La commutativité nous permet de mettre u et v à droite dans chacun des deux membres. Ensuite la proposition 1.54(3) nous permet de simplifier par $u + v$. Il reste $a + y = b + x$, qui signifie $(a, b) \sim (x, y)$.

(4) **Pour (2)** L'hypothèse donne les égalités

$$a + b' = b + a' \tag{1.159a}$$
$$x + y' = y + x' \tag{1.159b}$$

En sommant, et en utilisant l'associativité,

$$(a + x) + (b' + y') = (b + y) + (a' + x'). \tag{1.160}$$

Cela signifie bien que $(a + x, b + y) \sim (a' + x', b' + y')$.

□

Lemme 1.100.
Soient $a, b \in \mathbb{N}$. Nous avons $(a, b) \sim (0, 0)$ si et seulement si $a = b$.

Démonstration. Dire que $(a, b) \sim (0, 0)$ est équivalent à dire que $a + 0 = b + 0$, ou encore que $a = b$. □

Proposition-Définition 1.101 ([18]).
Soient $a, b, x, y \in \mathbb{N}$. L'application

$$\begin{aligned} f\colon \overline{(a, b)} \times \overline{(x, y)} &\to \mathbb{Z} \\ \big((a', b'), (x', y')\big) &\mapsto (a' + x', b' + y') \end{aligned} \tag{1.161}$$

est constante.

Nous nommons sa valeur $\overline{(a, b)} + \overline{(x, y)}$.

Démonstration. Cela est une conséquence de la proposition 1.99(2). □

Proposition 1.102.
La paire $(\mathbb{Z}, +)$ est un groupe commutatif.

Démonstration. En plusieurs points.

(1) **Neutre** Le neutre est $e = \overline{(0,0)}$. En effet,

$$\overline{(a,b)} + \overline{(0,0)} = \overline{(a+0, b+0)} = \overline{(a,b)}. \tag{1.162}$$

De même $e + \overline{(a,b)} = \overline{(a,b)}$ par commutativité de la somme dans $\mathbb{N}$.

(2) **Inverse** Il est facile de vérifier que $\overline{(b,a)}$ est l'inverse de $\overline{(a,b)}$.

(3) **Associativité** Calcul direct en utilisant l'associativité dans $\mathbb{N}$.

□

PROPooVMGYooUWsbSr

Proposition 1.103.
L'application

$$\begin{aligned} \iota\colon \mathbb{N} &\to \mathbb{Z} \\ n &\mapsto \overline{(n,0)} \end{aligned} \tag{1.163}$$

est un morphisme[39] *injectif.*

Démonstration. Le fait que ce soit un morphisme est le calcul

$$\iota(a+b) = \overline{(a+b,0)} = \overline{(a,0)} + \overline{(b,0)} = \iota(a) + \iota(b). \tag{1.164}$$

Pour l'injectivité, supposons que $\iota(a) = \iota(b)$. Alors $\overline{(a,0)} = \overline{(b,0)}$, c'est-à-dire $a + 0 = b + 0$. Donc $a = b$. □

1.4.1 Opposé

LEMooSABNooZZDIes

Lemme 1.104.
Tout élément de $\mathbb{Z}$ a un représentant de la forme $(a,0)$ ou $(0,b)$.

Démonstration. Soient $a, b \in \mathbb{N}$. Si $b \leqslant a$, alors nous avons

$$(a,b) \sim (a-b, 0) \tag{1.165}$$

où la différence est calculée dans $\mathbb{N}$ et a un sens parce que nous avons supposé $b \leqslant a$. Si par contre $a \leqslant b$ alors

$$(a,b) \sim (0, b-a). \tag{1.166}$$

Puisque l'ordre sur $\mathbb{N}$ est total[40], tous les cas sont couverts. □

Lemme-Définition 1.105.
Soit $z \in \mathbb{Z}$. L'application[41]

$$\begin{aligned} f\colon z &\to \mathbb{Z} \\ (a,b) &\mapsto \overline{(b,a)} \end{aligned} \tag{1.167}$$

est constante.

Nous nommons $-z$ sa valeur.

Démonstration. Soient (a,b) et (x,y) dans z. Nous avons successivement :

39. Certes $\mathbb{N}$ n'est pas un groupe, donc le mot « morphisme » est un peu abusé, mais vous voyez ce que je veux dire.
40. Proposition 1.63.
41. Pour rappel, z est une classe d'équivalence dans $\mathbb{N} \times \mathbb{N}$, c'est-à-dire une partie de $\mathbb{N} \times \mathbb{N}$. Ça a un sens de prendre z comme ensemble sur lequel on définit une fonction.

— $(a,b)\sim(x,y)$.
— $a+y=b+x$.
— $(b,a)\sim(y,x)$
— $\overline{(b,a)}=\overline{(y,x)}$
— $f(a,b)=f(x,y)$.

D'où la constance de f. □

PROPooJMETooHQGwnv

Proposition 1.106 ([1]).
Soient $a,b\in\mathbb{Z}$. Nous avons

ITEMooWNYGooMopPcq

(1) $(-a)b=-(ab)$

Lemme 1.107.
Nous avons

ITEMooSQFGooQPgIMu

(1) $\mathbb{Z}=\iota(\mathbb{N})\cup-\iota(\mathbb{N})$

ITEMooHQUQooJeqULl

(2) $\iota(\mathbb{N})\cap-\iota(\mathbb{N})=\{0\}$.

Démonstration. Nous avons

SUBEQooVJGVooCUxtvk

$$\iota(\mathbb{N})=\{\overline{(n,0)}\text{ tel que }n\in\mathbb{N}\}\tag{1.168a}$$

SUBEQooAPGRooQCkYRr

$$-\iota(\mathbb{N})=\{\overline{(0,n)}\text{ tel que }n\in\mathbb{N}\}.\tag{1.168b}$$

Montrons à présent les deux points.

(1) **Pour (1)** Nous savons par le lemme 1.104 que tous les éléments de $\mathbb{Z}$ sont de la forme (1.168a) ou (1.168b).

(2) **Pour (2)** Si $z\in\iota(\mathbb{N})\cap-\iota(\mathbb{N})$, il existe $n,m\in\mathbb{N}$ tels que $\overline{(n,0)}=\overline{(0,m)}$, ce qui signifie en particulier que $(n,0)\sim(0,m)$ ou encore que $n+m=0$. Le lemme 1.56 dit alors que $n=m=0$.
Nous avons donc $z=\overline{(0,0)}=0$.

□

PROPooCHGRooRksGGO

Proposition 1.108.
L'application

$$\begin{aligned}\iota\colon\mathbb{N}&\to\mathbb{Z}\\ n&\mapsto\overline{(n,0)}\end{aligned}\tag{1.169}$$

est un morphisme injectif qui respecte l'ordre.

Démonstration. Le fait que ce soit un morphisme injectif est la proposition 1.103.
Il faut encore prouver que ça respecte l'ordre [42] □

1.4.2 Ordre sur $\mathbb{Z}$

Si $z\in\mathbb{Z}$, nous disons que $z\in\mathbb{N}$ lorsque $z\in\iota(\mathbb{N})$. C'est un abus de notation qu'il est difficile de ne pas faire.

PROPooMYYDooOABOdB

Proposition-Définition 1.109 (Relation d'ordre [18]).
Nous disons que $x\leqslant y$ si et seulement si $y-x\in\mathbb{N}$.
L'ensemble $(\mathbb{Z},\leqslant)$ est totalement ordonné.

Une version dans $\mathbb{R}$ sera le lemme 1.434.

42. Faites-le en envoyez-moi la preuve.

LEMooSVDDooWsyxNP

Lemme 1.110.
Soient $a > 0$ et $b > 1$ dans $\mathbb{Z}$. Nous avons

$$ab > a. \tag{1.170}$$

PROPooYFUBooJUZgwH

Proposition 1.111 ([1])**.**
Quelques propriétés de l'ordre dans $\mathbb{Z}$.

(1) *Si $a, b \geqslant 0$, alors $ab \geqslant 0$.*
(2) *Si $a > 0$, alors $-a < 0$.*
(3) *Pour tout $x \in \mathbb{Z}$ nous avons $x^2 \geqslant 0$ et $x^2 = 0$ si et seulement si $x = 0$.*

ITEMooBUHIooLJOSLQ

(4) *Si $ab \geqslant 0$ et si $a \geqslant 0$ alors $b \geqslant 0$.*

ITEMooGDFPooBbpBgn

(5) *Si $ab > 0$ alors soit $a, b > 0$, soit $a, b < 0$.*

ITEMooSAJPooEDSgXJ

(6) *Si $a, b > 0$, alors $ab > 0$.*

LEMooMYEIooNFwNVI

Lemme 1.112.
Toute partie bornée de $\mathbb{Z}$ possède un plus grand élément.

PROPooYJBMooZrzkNX

Proposition 1.113.
Soit $a, b \in \mathbb{Z}$ tels que a divise b. Alors $|a| \leqslant |b|$.

LEMooJNXIooBmdOVi

Lemme 1.114.
L'ensemble $\mathbb{Z}$ est infini dénombrable.

1.5 Quelques résultats de cardinalité

1.5.1 Équipotence, surpotence, subpotence

Les notions d'équipotence, surpotence et de subpotence permettent de comparer les « tailles » des ensembles sans avoir besoin de la théorie des ordinaux. Tout ceci ne sera pas très souvent utile par la suite. Un exemple d'utilisation de ces notions est le théorème de Steinitz 6.135 qui démontre l'existence de clôture algébrique pour tout corps.

DEFooXGXZooIgcBCg

Définition 1.115 ([21, 22])**.**
Soient deux ensembles A et B.

(1) *Les ensembles A et B sont **équipotents** si il existe une bijection entre A et B. Nous notons $A \approx B$.*
(2) *L'ensemble A est **surpotent** à B si il existe une surjection de A vers B. Nous notons $A \succeq B$.*
(3) *L'ensemble A est **subpotent** à B si il existe une injection de A vers B. Nous notons $A \preceq B$.*

Nous disons également « strictement » surpotent quand il y a surpotence mais pas équipotence, et de même pour la subpotence. Les symboles $\succ$ et $\prec$ sont alors utilisés.

PROPooWSXTooMQPcNG

Proposition 1.116 ([1, 13])**.**
L'ensemble A est subpotent à B si et seulement si B est surpotent à A.

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que A est subpotent à B. Il existe une injection $\varphi: A \to B$. Nous définissons $f: B \to A$ par

$$f(x) = \begin{cases} \varphi^{-1}(x) & \text{si } x \in \varphi(A) \\ a & \text{sinon} \end{cases} \tag{1.171}$$

où a est un élément quelconque de A. Cette application est bien définie parce que φ est injective, de telle sorte que φ^{-1} est bien définie. Puisque φ est définie sur tout A, l'application f est une surjection.

(2) $\Leftarrow$ Nous supposons que B est surpotent à A. Il existe donc une surjection $\varphi\colon B\to A$. Pour chaque $x\in A$ nous considérons un élément $b_x\in\varphi^{-1}(x)$, qui existe parce que φ est surjective. Nous considérons ensuite l'application

$$\begin{aligned} f\colon A&\to B\\ x&\mapsto b_x. \end{aligned} \tag{1.172}$$

Nous prouvons que f est une injection. Supposons que $x,y\in A$ soient tels que $f(x)=f(y)$. Nous avons $b_x=b_y$. Donc

$$x=\varphi(b_x)=\varphi(b_y)=y. \tag{1.173}$$

Nous avons prouvé que $x=y$, et donc que f est injective.

□

Vu que l'ensemble des ensembles n'existe pas [43], nous n'allons pas énoncer le fait que ces notions donnent une relation d'ordre sur les ensembles ; il faudrait parler de classes et nous ne nous en sortirions pas. Nous allons toutefois énoncer quelques résultats qui vont dans ce sens. Pour en savoir plus, vous pouvez lire les différentes pages de Wikipédia sur les nombres cardinaux.

1.5.2 Un peu d'infinité

DefEOZLooUMCzZR

Définition 1.117 (ensemble Dedekind infini)**.**
*Un ensemble est **infini** si il peut être mis en bijection avec un de ses sous-ensembles propres (c'est-à-dire différent de lui-même).*

*Un ensemble est **fini** si il n'est pas infini.*

1.118.
Nous adoptons les notions d'ensembles finis et infinis au sens de Dedekind. De nombreuses sources (dont wikipédia [23, 24]) définissent un ensemble fini comme étant un ensemble en bijection avec une partie de $\mathbb{N}$ de la forme $\{0,\dots,N\}$. Alors un ensemble est infini si il n'est pas fini.

Cependant, d'une part *les deux définitions d'ensembles infinis ne sont pas équivalentes*, mais d'autre part, elles sont équivalentes si on accepte l'axiome du choix [44]. Or le Frido accepte l'axiome du choix sans vergogne et sous toutes ses formes. Nous démontrerons donc, en utilisant le lemme de Zorn, qu'un ensemble A est fini (définition 1.117) si et seulement si il existe une bijection $\{0,\dots,N\}\to A$ pour un certain $N\in\mathbb{N}$. Ce sera le théorème 1.126.

LEMooTUIRooEXjfDY

Lemme 1.119.
Toute partie d'un ensemble fini est finie.

Démonstration. Nous allons prouver la contraposée : si un ensemble contient une partie infinie, alors il est infini. Soit $A\subset B$ où A est infini. Nous allons prouver que B est infini. En vertu de la définition 1.117, il existe une partie $A'\subsetneq A$ et une bijection $\sigma\colon A'\to A$.

Nous considérons la partie $B'=A'\cup(B\backslash A)$, qui est une partie stricte de B. Puisque $A'\cap(B\backslash A)=\emptyset$, nous pouvons définir

$$\begin{aligned} \varphi\colon B'&\to B\\ x&\mapsto \begin{cases} \sigma(x) & \text{si } x\in A'\\ x & \text{si } x\in B\backslash A. \end{cases} \end{aligned} \tag{1.174}$$

Montrons que φ est une bijection.

(1) **Surjectif** Nous avons $\varphi(A')=A$ et $\varphi(B\backslash A)=B\backslash A$. Donc

$$\varphi(B')=\varphi(A')\cup\varphi(B\backslash A)=A\cup(B\backslash A)=B. \tag{1.175}$$

43. Voir le corolaire 1.151.
44. Et même seulement l'axiome du choix dénombrable ; si vous voulez en savoir plus, lisez la page wikipédia [25].

(2) **Injectif** Si $\varphi(x) = \varphi(y)$, nous avons 4 possibilités suivant que x et y sont dans A' ou $B\backslash A$.
Si $x, y \in A'$, alors $\varphi(x) = \varphi(y)$ implique $\sigma(x) = \sigma(y)$ et donc $x = y$ parce que σ est injective.
Si $x \in A'$ et $y \in B\backslash A$ alors $\varphi(x) = \sigma(x) \in A$ et $\varphi(y) = y \in B\backslash A$. Ce cas n'est pas possible. Le cas $x \in B\backslash A$ et $y \in A'$ n'est pas possible non plus.
Si $x, y \in B\backslash A$, alors $\varphi(x) = \varphi(y)$ implique immédiatement $x = y$.

Nous avons une bijection entre B' et B alors que B' est un sous-ensemble strict de B. Donc B est infini. □

PROPooVOKDooOStPzU

Proposition 1.120 ([1, 13]).
Si A est fini et si $\omega \notin A$, alors $A \cup \{\omega\}$ est fini.

Démonstration. Supposons que $A \cup \{\omega\}$ est infini. Il existe un sous-ensemble strict de $A \cup \{\omega\}$ en bijection avec $A \cup \{\omega\}$. Soient donc $B \subsetneq A \cup \{\omega\}$ et $\sigma\colon B \to A \cup \{\omega\}$ une bijection.

Il y a deux possibilités : soit ω est dans B, soit non.

(1) $\omega \notin B$ Alors $B \subset A$, et il existe $x \in B$ tel que $\sigma(x) = \omega$. Considérons $B' = B\backslash\{x\}$; cela est une partie propre de A. Ensuite nous définissons

$$\begin{aligned} \varphi\colon B\backslash\{x\} &\to A \\ a &\mapsto \sigma(a). \end{aligned} \tag{1.176}$$

C'est injectif parce que σ est injective, et c'est surjectif parce que

$$\varphi\big(B\backslash\{x\}\big) = \sigma(B)\backslash\sigma(\{\omega\}) = \sigma(B)\backslash\{\omega\} = (A \cup \{\omega\})\backslash\{\omega\} = A. \tag{1.177}$$

Pour la dernière égalité nous avons utilisé le fait que ω n'est pas dans A.

(2) **Si $\omega \in B$** Plusieurs choses à dire.

(2a) $\sigma(\omega) \neq \omega$
Si $\omega \in B$, nous commençons par prouver que $\sigma(\omega) \neq \omega$. En effet si $\sigma(\omega) = \omega$, alors nous pouvons poser

$$\begin{aligned} \sigma'\colon B\backslash\{\omega\} &\to A \\ x &\mapsto \sigma(x). \end{aligned} \tag{1.178}$$

Ce σ' est une bijection entre A et la partie propre $B\backslash\{A\}$. Cela est impossible parce que nous avons supposé que A est fini.

(2b) **Definition de φ** Puisque B est une partie propre de $A \cup \{\omega\}$, il existe $x \in A\backslash B$. Nous considérons $B' = \big(B\backslash\{\omega\}\big) \cup \{x\}$ et nous définissons

$$\begin{aligned} \varphi\colon B' &\to A \\ b &\mapsto \begin{cases} \sigma(b) & \text{si } b \neq x \\ \sigma(\omega) & \text{si } b = x. \end{cases} \end{aligned} \tag{1.179}$$

Nous montrons à présent que φ est une bijection.

(2c) **Injectif** Soient $u, v \in B'$ tels que $\varphi(u) = \varphi(v)$. Il y a 4 possibilités suivant que u ou v est égal à x.

— Si $u = v = x$ on est bon.

— Si $u = x$ et $v \neq x$. D'une part $\varphi(u) = \varphi(v) = \sigma(v)$ parce que $v \neq x$. Mais d'autre part $\varphi(u) = \varphi(x) = \sigma(\omega)$. Donc $\sigma(v) = \sigma(\omega)$ et $v = \omega$ par injectivité de σ. Or cela est impossible parce que $v \in B'$ alors que $\omega \notin B'$. Ici nous utilisons le fait que ω ne peut pas être égal à x parce que $x \in A$ alors que ω est hors de A.
Bref, ce cas est impossible.

— Si $u \neq x$ et $v = x$, alors c'est le même cas que le précédent avec quelque adaptations.

— Si $u \neq x$ et $v \neq x$, alors $\varphi(u) = \sigma(u)$ et $\varphi(v) = \sigma(v)$, de telle sorte que $\sigma(u) = \sigma(v)$ et donc $u = v$ par injectivité de σ.

(2d) **Surjective** Soit $y \in A$. Vu que $\sigma : B \to A \cup \{\omega\}$ est surjective, il existe $b \in B$ tel que $\sigma(b) = y$. Notez que $b \neq x$ parce que $x \in A \backslash B$.
Si $b \neq \omega$ alors $b \in B\backslash\{\omega\} \subset B'$. Nous pouvons calculer $\varphi(b)$. Étant donné que $b \neq x$, nous avons $\varphi(b) = \sigma(b) = y$.
Si au contraire $b = \omega$, alors $\varphi(x) = \sigma(\omega) = y$. Dans les deux cas, y est dans l'image de φ.

Dans tous les cas nous avons construit une bijection entre une partie propre $B' \subsetneq A$ et A, ce qui est absurde parce que A est un ensemble fini. □

La proposition suivante est à peu près prise comme définition d'un ensemble fini dans [26] qui donne également une preuve de l'équivalence avec notre définition.

PROPooBYKCooGDkfWy

Proposition 1.121.
L'ensemble $\mathbb{N}$ *est infini*[45].

Démonstration. Nous considérons la partie propre[46]

$$A = \{n \in \mathbb{N} \text{ tel que } n \geqslant 1\}. \tag{1.180}$$

Ensuite nous posons

$$\begin{aligned} \sigma : \mathbb{N} &\to A \\ x &\mapsto s(x). \end{aligned} \tag{1.181}$$

Le fait que σ prenne ses valeurs dans A est parce que s prend ses valeurs dans $\mathbb{N}\backslash\{0\} = A$.

(1) σ **est injective** Parce que s l'est.

(2) σ **est surjective** C'est dans la définition 1.44(2)

Nous avons donc une bijection entre $\mathbb{N}$ et un sous-ensemble strict. □

LEMooYIWDooIQFSad

Lemme 1.122.
Soit $N \in \mathbb{N}$. *La partie* $\{0, \dots, N\}$ *est finie*[47].

Démonstration. Par récurrence sur N. Avec $N = 0$, la partie $\{0\}$ est finie parce que son seul sous-ensemble propre est $\varnothing$ qui n'est pas en bijection avec $\{0\}$.

Supposons que $\{0, \dots, N\}$ est fini. Alors $\{0, \dots, N\} \cup \{N + 1\}$ est fini par le lemme 1.120. □

LEMooJDGOooHdyJnu

Lemme 1.123.
Si $p \neq q$, *alors il n'existe pas de bijection entre* $\{0, \dots, p\}$ *et* $\{0, \dots, q\}$.

Démonstration. Supposons pour fixer les idées que $p \leqslant q$. Dans ce cas $\{0, \dots, p\}$ est une partie stricte de $\{0, \dots, q\}$. Vu que $\{0, \dots, q\}$ est fini (lemme 1.122), il n'y a pas de bijection avec ses parties strictes. □

PROPooWKSIooHcfYPN

Proposition 1.124 ([1])**.**
Si A *est infini et si* $\sigma : A \to B$ *est injective, alors* B *est infini.*

Démonstration. Nous allons prouver que $\sigma(A)$ est une partie infinie de B. Puisque A est infini, nous pouvons considérer une partie $A' \subsetneq A$ et une bijection $\varphi_A : A' \to A$. Nous définissons

$$\begin{aligned} \varphi_B : \sigma(A') &\to \sigma(A) \\ y &\mapsto \sigma\Big(\varphi_A\big(\sigma^{-1}(y)\big)\Big). \end{aligned} \tag{1.182}$$

45. Définition 1.117.
46. Le fait que ce soit une partie propre est dû au fait que 0 n'est pas dedans d'une part parce que le lemme 1.58 dit que $0 \leqslant 1$, et d'autre part parce que le lemme 1.56 donne $0 \neq 1$.
47. Pour rappel, la définition de $\{0, \dots, N\}$ est 1.68.

Cette définition a un sens parce que si $y \in \sigma(A')$, alors il existe un unique $x \in A'$ tel que $\sigma(x) = y$ parce que σ est injective. De là, $\varphi_A(x) \in A$ et nous pouvons lui appliquer σ.

Nous montrons que φ_B est une bijection.

(1) **Injective** Supposons que $\varphi_B(a) = \varphi_B(b)$, c'est-à-dire que

$$(\sigma \circ \varphi_A \circ \sigma^{-1})(a) = (\sigma \circ \varphi_A \circ \sigma^{-1})(b). \tag{1.183}$$

Étant donné que σ et φ_A sont injectives, nous avons $\sigma^{-1}(a) = \sigma^{-1}(b)$. En appliquant σ des deux côtés, nous trouvons $a = b$.

(2) **surjective** Soit $y \in \sigma(A)$. En prenant $x \in (\sigma \circ \varphi_A^{-1} \circ \sigma^{-1})(y)$ nous avons $\varphi_B(x) = y$.

Donc B contient une partie infinie ($\sigma(A)$). Le lemme 1.119 conclut que B est infini. □

LEMooPGPVooZzlFvf

Lemme 1.125.
À propos d'applications entre ensembles finis.

ITEMooNCCUooBGrtdn

(1) Si $\sigma\colon A \to B$ est une application quelconque et si A est fini, alors $\sigma(A)$ est une partie finie de B.

ITEMooKQMFooSzmXrd

(2) Si $\sigma\colon A \to B$ est surjective et si A est fini, alors B est fini.

Démonstration. Nous allons utiliser le lemme de Zorn. Nous considérons l'ensemble

$$\mathcal{A} = \{X \subset A \text{ tel que } \sigma\colon X \to \sigma(A) \text{ est injective}\} \tag{1.184}$$

que nous ordonnons (partiellement) par l'inclusion.

(1) **$\mathcal{A}$ est inductif** Soit une partie totalement ordonnée $\mathcal{F}$ de $\mathcal{A}$. Nous considérons $Y = \bigcup_{X\in\mathcal{F}} X$, et nous prouvons que Y est un majorant de $\mathcal{F}$.
Pour cela nous commençons par prouver que $Y \in \mathcal{A}$. Soient $a, b \in Y$ tels que $\sigma(a) = \sigma(b)$. Il existe $X_1, X_2 \in \mathcal{F}$ tels que $a \in X_1$ et $b \in X_2$. Supposons pour fixer les idées que $X_1 \leqslant X_2$ ($\mathcal{F}$ étant totalement ordonné nous avons toujours $X_1 \leqslant X_2$ ou $X_2 \leqslant X_1$). Puisque l'ordre est l'inclusion, cela signifie que $X_1 \subset X_2$. Nous avons donc $a, b \in X_2$, alors que σ est injective sur X_2. Donc $\sigma(a) = \sigma(b)$ implique $a = b$, et σ est injective sur Y. Nous avons donc prouvé que $Y \in \mathcal{A}$.
Puisque pour tout $X \in \mathcal{F}$ nous avons $X \subset Y$, nous avons $X \leqslant Y$ (dans $\mathcal{A}$) pour tout $X \in \mathcal{F}$. Bref, Y est un majorant de $\mathcal{F}$ dans $\mathcal{A}$.
Toute partie totalement ordonnée de $\mathcal{A}$ est majorée. Cela signifie que $\mathcal{A}$ est inductif [48].

(2) **Zorn** L'ensemble $\mathcal{A}$ étant inductif et non vide (les singletons dans A sont dans $\mathcal{A}$), il possède un élément maximal [49] par le lemme de Zorn 1.23. Nous nommons A' un élément maximal dans $\mathcal{A}$.

(3) **Bijective** L'application $\sigma\colon A' \to \sigma(A)$ est injective parce que $A' \in \mathcal{A}$. Nous devons prouver qu'elle est surjective.
Supposons que $y \in \sigma(A)\backslash\sigma(A')$. Alors il existe $a \in A\backslash A'$ tel que $\sigma(a) = y$. Dans ce cas, la partie $A' \cup \{a\}$ est un majorant de A' dans $\mathcal{A}$, ce qui est impossible.
Donc $\sigma\colon A' \to \sigma(A)$ est bijective.

(4) **Conclusion** L'ensemble A' est fini en tant que partie de l'ensemble fini A (lemme 1.119). L'application σ étant injective, la proposition 1.124 conclut que $\sigma(A')$ est fini. Et comme $\sigma(A')$ n'est autre que $\sigma(A)$ nous avons fini.

La partie (1) est prouvée. La partie (2) est maintenant facile. La partie (1) dit que $\sigma(A)$ est une partie finie de B, mais si σ est surjective, alors $\sigma(A) = B$. □

PROPooJLGKooDCcnWi

Proposition-Définition 1.126 (Cardinal d'un ensemble fini[1])**.**
Soit un ensemble non vide I.

48. Plus précisément c'est l'ensemble ordonné $(\mathcal{A}, \subset)$ qui est inductif.
49. Définition 1.13 ; voir aussi 1.16.

ITEMooMNMTooEOIjdo

(1) L'ensemble I est fini[50] *si et seulement si il existe une bijection entre I et $\{0,\dots,N\}$ pour un certain $N\in\mathbb{N}$.*

ITEMooZJFUooSNUSIk

(2) Si I est fini, il existe un unique $N\in\mathbb{N}$ tel que I soit en bijection avec $\{0,\dots,N\}$.

Dans ce cas, le nombre $N+1$ est le ***cardinal*** *de I, et est noté* $\text{Card}(I)$*. Pour l'ensemble vide, nous définissons* $\text{Card}(\varnothing)=0$.

Démonstration. En plusieurs parties.

(1) **(1)⇐**
Soit un ensemble A en bijection avec $\{0,\dots,N\}$. Nous avons vu que $\{0,\dots,N\}$ est fini dans le lemme 1.122. Le lemme 1.125(2) conclut que A est fini.

(2) **(1)⇒** Le vrai sport est de faire l'implication inverse. Nous supposons que A est un ensemble fini, et nous allons prouver qu'il est en bijection avec $\{0,\dots,N\}$ pour un N bien choisi. Nous allons utiliser le lemme de Zorn. Soit

$$\mathcal{A}=\{(N,\varphi)\text{ tel que }\varphi\colon\{0,\dots,N\}\to A\text{ est injective}\}. \tag{1.185}$$

Nous mettons sur $\mathcal{A}$ la relation d'ordre donnée par $(N_1,\varphi_1)\leqslant(N_2,\varphi_2)$ lorsque

(2a) $N_1\leqslant N_2$

(2b) φ_2 étend φ_1, c'est-à-dire que $\varphi_2=\varphi_1$ sur $\{0,\dots,N\}$.

(2a) **$\mathcal{A}$ est inductif** Soit une partie $\mathcal{F}$ totalement ordonnée de $\mathcal{A}$. Nous considérons la partie suivante de $\mathbb{N}$:

$$S=\{n\in\mathbb{N}\text{ tel que }\exists\varphi\text{ tel que }(n,\varphi)\in\mathcal{F}\}. \tag{1.186}$$

Si S est majoré, alors il a un maximum (proposition 1.64(4)). Si M est le maximum de S, le (M,φ) de $\mathcal{F}$ qui correspond à ce maximum est un majorant de $\mathcal{F}$.
Supposons –pour l'absurde– que $\mathcal{F}$ n'est pas majoré dans $\mathcal{A}$; en particulier S n'est pas majoré dans n. Pour chaque $n\in S$, il existe une application φ_n telle que $(n,\varphi_n)\in\mathcal{F}$. Cela nous permet de définir

$$\begin{aligned}\phi\colon S&\to A\\ n&\mapsto\varphi_n(n).\end{aligned} \tag{1.187}$$

Montrons que ϕ est injective. Si $\phi(m)=\phi(n)$, alors $\varphi_m(m)=\varphi_n(n)$. Supposons pour fixer les idées que $n\leqslant m$. Vu que (n,φ_n) et (m,φ_m) sont dans $\mathcal{F}$ qui est ordonné, φ_m prolonge φ_n ; en particulier $\varphi_n(n)=\varphi_m(n)$. Mais comme φ_m est injective, $m=n$.
Nous avons donc une injection $\phi\colon S\to A$. Mais S est non borné et donc infini[51]. La proposition 1.124 conclut que A est infini, ce qui est contraire aux hypothèses.
Donc $\mathcal{F}$ a un majorant et $\mathcal{A}$ est inductif.

(2b) **Lemme de Zorn** Le lemme de Zorn dit que $\mathcal{A}$ a un élément maximal.

(2c) **Conclusion** Soit (N,φ) un élément maximal de $\mathcal{A}$. Nous allons prouver que $\varphi\colon\{0,\dots,N\}\to A$ est une bijection. Que φ soit injective est une conséquence du fait que (N,φ) est dans $\mathcal{A}$.
Pour prouver que φ est surjective, nous supposons qu'elle ne l'est pas. Soit $a\in A$ qui n'est pas dans l'image de φ. En posant

$$\begin{aligned}\phi\colon\{0,\dots,N+1\}&\to A\\ x&\mapsto\begin{cases}\varphi(x)&\text{si }x\neq N+1\\ a&\text{si }x=N+1,\end{cases}\end{aligned} \tag{1.188}$$

le couple $(N+1,\phi)$ majore strictement (N,φ). Ce qui est une contradiction.

50. Ensemble fini, définition 1.117.

51. Le lemme 1.70 dit qu'il existe une bijection entre S et $\mathbb{N}$. De là nous concluons que S est infini parce qu'un ensemble en bijection avec un ensemble infini est infini par la proposition 1.124.

(3) **(2) existence** L'existence est ce que nous venons de montrer ci-dessus.

(4) **(2) unicité** Supposons que I soit en bijection avec $\{0,\dots,M\}$ et avec $\{0,\dots,N\}$. Il existe donc une bijection entre $\{0,\dots,M\}$ et $\{0,\dots,N\}$. Par le lemme 1.123, cela implique que $M=N$.

□

Nous ne définissons pas ce qu'est le cardinal d'un ensemble infini ; c'est très compliqué et ça ne nous servira pas.

LEMooRWUDooTgoRXH

Lemme 1.127.
Si A est une partie infinie de $\mathbb{N}$, alors pour tout n, la partie $A\backslash\{0,\dots,n\}$ est non vide.

Démonstration. Si $A\backslash\{0,\dots,n\}$ était vide, cela signifierait que A est une partie de $\{0,\dots,n\}$. Or nous savons que $\{0,\dots,n\}$ est fini (lemme 1.122), et que toute partie d'un ensemble fini est finie (lemme 1.119). Donc nous aurions que A est fini, ce qui est contraire à l'hypothèse. □

LEMooVFPNooVmdUXY

Lemme 1.128 ([1]).
Union d'ensembles finis.

ITEMooBUCZooYLCuIe

(1) Si A et B sont des ensembles finis disjoints, alors $A\cup B$ est fini et

$$\operatorname{Card}(A\cup B)=\operatorname{Card}(A)+\operatorname{Card}(B). \tag{1.189}$$

ITEMooCCWOooYwgGBp

(2) Si A et B sont des ensembles finis, alors $A\cup B$ est fini.

ITEMooYJSZooXQXkOX

(3) Si A est fini et si $B\subset A$ alors

$$\operatorname{Card}(A\backslash B)=\operatorname{Card}(A)-\operatorname{Card}(B). \tag{1.190}$$

ITEMooSWJCooEpBVkG

(4) Si A et B sont des ensembles quelconques, alors

$$\operatorname{Card}(A\cup B)=\operatorname{Card}(A)+\operatorname{Card}(B)-\operatorname{Card}(A\cap B). \tag{1.191}$$

ITEMooJDUUooVMvAOn

(5) Si les $\{A_i\}_{i=1,\dots,n}$ sont des ensembles deux à deux disjoints, alors

$$\operatorname{Card}\big(\bigcup_{i=1}^{n}A_i\big)=\sum_{i=1}^{n}\operatorname{Card}(A_i). \tag{1.192}$$

ITEMooNMFSooBvsNyq

(6) Si I ou J est infini, alors $I\cup J$ est infini.

Démonstration. Point par point.

(1) **Pour (1)** Puisque A et B sont finis, la proposition 1.126 nous dit qu'il existe des naturels N et M ainsi que des bijections $\varphi_A\colon\{0,\dots,N\}\to A$ et $\varphi_B\colon\{0,\dots,M\}\to B$. Maintenant l'application

$$\begin{aligned}\varphi\colon\{0,\dots,M+N+1\}&\to A\cup B\\ n&\mapsto\begin{cases}\varphi_A(n) & \text{si } n\leqslant N\\ \varphi_B(n-N-1) & \text{si } n>N\end{cases}\end{aligned} \tag{1.193}$$

est une bijection. Le fait que A et B soient disjoints est important pour l'injectivité. La proposition 1.126 nous dit qu'alors $A\cup B$ est fini. De plus, par définition le cardinal de $A\cup B$ est $N+M$.

(2) **Pour (2)** Nous ne supposons plus que A et B sont disjoints. Nous posons $I=A$ et $J=B\backslash A$. Avec ça, I et J sont disjoints et finis (comme parties des ensembles finis, lemme 1.119), et vérifient $I\cup J=A\cup B$. Le point (1) indique que $I\cup J$ est fini.

(3) **Pour (3)**
L'ensemble A peut être écrit sous la forme d'une union disjointe : $A = B \cup (A\backslash B)$. Les ensembles B et $A\backslash B$ étant disjoints, nous avons

$$\text{Card}(A) = \text{Card}(B) + \text{Card}(A\backslash B). \tag{1.194}$$

(4) **Pour (4)** Nous utilisons quelques égalités d'ensembles pour ramener $A \cup B$ à des cas déjà traités :

$$A \cup B = A \cup (B\backslash A) = A \cup \big(B\backslash(A \cap B)\big). \tag{1.195}$$ EQooSZDYooPBOdYv

Nous avons en particulier utilisé $B\backslash A = B\backslash(A \cap B)$. La chose intéressante dans (1.195) est que l'union est disjointe et que $A \cap B \subset B$. Nous pouvons donc écrire

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}\big(B\backslash(A \cap B)\big) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B). \tag{1.196}$$

(5) **Pour (5)** Récurrence en utilisant le point (4).

(6) **Pour (6)** Toute partie d'un ensemble fini est finie (lemme 1.119). Donc si $I \cup J$ était fini, I et J devraient l'être.

□

DefEnsembleDenombrable

Définition 1.129.
Un ensemble est ***dénombrable*** *si il peut être mis en bijection avec* $\mathbb{N}$*. Il est* ***non dénombrable*** *si il est infini et ne peut pas être mis en bijection avec* $\mathbb{N}$*.*

Une chose vraiment amusante avec cette définition que l'on met en rapport avec la définition 1.117, c'est qu'un ensemble fini n'est ni dénombrable ni non dénombrable [52].

LEMooDTAEooIBdHyo

Lemme 1.130.
Si A est dénombrable et si il existe une surjection $f\colon A \to B$, alors B est fini ou dénombrable.

LEMooSRZWooASgEfy

Lemme 1.131.
Si A est un ensemble fini ou dénombrable, alors il existe une surjection $\mathbb{N} \to A$.

1.5.3 Dénombrabilité et ensemble des naturels

PROPooOBKMooWEGCvM

Proposition 1.132 ([1, 13])**.**
Toute partie infinie de $\mathbb{N}$ *est dénombrable.*

Démonstration. Soit A, une partie infinie de $\mathbb{N}$.

(1) **Définition de σ** Nous voulons construire une application $\sigma\colon \mathbb{N} \to A$ telle que

$$\begin{cases} \sigma(0) = \min(A) & \text{(1.197a)} \quad \text{SUBEQooEIEMooZcTOWT} \\ \sigma(k+1) = \min\Big(A\backslash\sigma\big(\{0,\dots,k\}\big)\Big) & \text{(1.197b)} \quad \text{SUBEQooWWOAooAEfrPx} \end{cases}$$

Les lâches, par prudence, diront juste que c'est défini par récurrence et n'insisteront pas. Nous, nous insistons.
Nous allons définir $\sigma(n)$ à l'aide du théorème 1.48. Pour cela nous posons $E = \mathcal{P}(A)$, $b = \varnothing$ et

$$\begin{aligned} g\colon E &\to E \\ Z &\mapsto \begin{cases} A & \text{si } Z = A \\ Z \cup \{\min(A\backslash Z)\} & \text{sinon.} \end{cases} \end{aligned} \tag{1.198}$$

Notons que la proposition 1.64(2) nous indique que toute partie non vide de $\mathbb{N}$ possède un minimum ; la définition de g a donc un sens. Le théorème 1.48 donne alors une application $f\colon \mathbb{N} \to E$ telle que

52. Beaucoup de sources disent qu'un ensemble est dénombrable lorsqu'il est en bijection avec une partie de $\mathbb{N}$. Cela laisse la porte ouverte aux ensembles finis. Par exemple Wikipédia[27].

(1a) $f(0) = b = \varnothing$

(1b) $f(n+1) = g\big(f(n)\big)$ pour tout $n \geqslant 0$.

Prouvons par récurrence que $f(n)$ est un ensemble fini pour tout n. D'abord $f(0) = \varnothing$. Ensuite, si $n \geqslant 0$ est tel que $f(n)$ est fini, alors en particulier $f(n) \neq A$ et nous avons

$$f(n+1) = g\big(f(n)\big) = f(n) \cup \{\min\big(A\backslash f(n)\big)\}. \tag{1.199}$$

Dans ce cas, $f(n+1)$ est également fini comme union de deux ensembles finis.

Nous posons

$$\sigma(n) = \min\big(A\backslash f(n)\big). \tag{1.200}$$

EQooGHQHooRnXDdo

Avec $n = 0$, nous avons $\sigma(0) = \min\big(A\backslash\varnothing\big) = \min(A)$. La condition (1.197a) est donc déjà satisfaite.

Nous devons encore prouver (1.197b). Pour tout n, la relation entre $f(n)$ et $\sigma(n)$ est donnée par

$$\begin{cases} f(0) = \varnothing & \text{(1.201a)} \\ f(n+1) = f(n) \cup \sigma(n). & \text{(1.201b)} \end{cases}$$

Par récurrence nous avons alors

$$f(n) = \bigcup_{k=0}^{n-1} \{\sigma(k)\} = \sigma\big(\{0, \dots, n-1\}\big) \tag{1.202}$$

EQooPXFEooYzhtBe

pour tout $n \geqslant 1$. Nous avons alors la condition 1.197b en substituant (1.202) dans la définition (1.200) écrite avec $n+1$:

$$\sigma(n+1) = \min\big(A\backslash f(n+1)\big) = \min\Big(A\backslash\sigma\big(\{0, \dots, n\}\big)\Big). \tag{1.203}$$

(2) **σ est strictement croissante** Vu que $A\backslash\sigma\{0, \dots, k\} \subset A\backslash\sigma\{0, \dots, k-1\}$, le minimum est plus grand ou égal : $\sigma(k+1) \geqslant \sigma(k)$. Mais $\sigma(k+1)$ est sélectionné dans l'ensemble $A\backslash\sigma\{0, \dots, k\}$, qui ne contient justement pas $\sigma(k)$. Donc $\sigma(k+1) \neq \sigma(k)$.

(3) **σ est définie sur $\mathbb{N}$** Il faut montrer que pour tout k, l'ensemble $A\backslash\sigma\{0, \dots, k\}$ est non vide. Si il l'était, cela signifierait que $A \subset \sigma\{0, \dots, k\}$. Par le lemme 1.125(1), la partie $\sigma\{0, \dots, k\}$ est finie dans $\mathbb{N}$. Le lemme 1.119 dit alors qu'en tant que partie de $\sigma\{0, \dots, k\}$, l'ensemble A est fini. Mais comme les hypothèses disent que A est infini, nous avons une contradiction et nous concluons que σ est bien définie sur tout $\mathbb{N}$.

(4) **σ est injective** Une application $\mathbb{N} \to \mathbb{N}$ strictement croissante est injective par la proposition 1.69.

(5) **σ est surjective** Soit $a \in A$. Vu que σ est strictement croissante et que $\sigma(0) \geqslant 0$, nous avons $\sigma(a) \geqslant a$. Si $\sigma(a) = a$ nous avons terminé. Supposons $\sigma(a) > a$. Alors

$$\min\big(A\backslash\sigma\{0, \dots, a\}\big) > a. \tag{1.204}$$

EQooNHTBooQexzwV

Si $\sigma\big(\{0, \dots, a\}\big)$ ne contenait pas a, alors $A\backslash\sigma(\{0, \dots, a\})$ le contiendrait et nous n'aurions pas l'inégalité (1.204). Donc $a \in \sigma\big(\{0, \dots, a\}\big)$ et a est bien dans l'image de σ.

□

1.133.

La proposition 1.132 pourrait être prouvée plus facilement en acceptant le théorème de Cantor-Schröder-Bernstein 1.146. Il existe une injection $A \to \mathbb{N}$ parce que A est une partie de $\mathbb{N}$. Mais puisque A est infini, il possède une partie dénombrable. Cela donne une surjection $A \to \mathbb{N}$ et donc une injection $\mathbb{N} \to A$. Le théorème de Cantor-Schröder-Bernstein conclut.

Cela dit, une telle preuve demanderait des outils plus complexes.

1.134.
La proposition suivante donne une bijection explicite entre $\mathbb{N}$ et $\mathbb{N} \times \mathbb{N}$. Elle n'a rien de transcendante, mais je ne résiste pas à la donner ici parce qu'elle est utilisée dans l'article *Un peu de programmation transfinie* de David Madore[53]. Son utilité est de pouvoir créer un langage de programmation pouvant traiter des paires d'entiers rien qu'en traitant des entiers.

PROPooLPKUooAlsYJg

Proposition 1.135 (Une bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$)**.**
La fonction

$$\begin{aligned} f\colon \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (x,y) &\mapsto \begin{cases} y^2 + x & si\ x < y \\ x^2 + x + y & si\ y \leqslant x. \end{cases} \end{aligned} \tag{1.205}$$

est une bijection.

Démonstration. Il s'agit de prouver qu'elle est injective et surjective. Dans la suite, tous les nombres sont des entiers positifs.

(1) **f est injective**
Pour $k \in \mathbb{N}$ donné, nous allons prouver que
(1a) l'équation $f(x,y) = k$ possède au maximum une solution avec $x < y$,
(1b) l'équation $f(x,y) = k$ possède au maximum une solution avec $y \leqslant x$,
(1c) si $k = y^2 + x$ avec $x < y$ alors il est impossible que $k = x'^2 + x' + y'$ avec $y' \leqslant x'$.
On y va.
(1a) Nous supposons $y^2 + x = t^2 + z$ avec $x < y$ et $z < t$. Pour fixer les idées, nous supposons $t > y$ et nous posons $t = y + s$ ($s \geqslant 1$). En substituant, et en isolant z,

$$\begin{aligned} z &= x - 2sy - s^2 && (1.206a)\\ &< x - 2sy && (1.206b)\\ &< x - 2sx && (1.206c)\\ &= x(1 - 2s) && (1.206d)\\ &< 0. && (1.206e) \end{aligned}$$

Impossible parce que $z \geqslant 0$.
(1b) De même nous supposons $x^2 + x + y = z^2 + z + t$ avec $y \leqslant x$ et $t \leqslant z$. Nous posons $z = x + s$, et nous déballons le même genre de calculs en isolant t.
(1c) Enfin nous supposons $y^2 + x = z^2 + z + t$ avec $x < y$ et $t \leqslant z$. Les plus courageux diviseront en trois cas : $y < z$, $y = z$ et $y > z$ et feront les calculs. Par exemple, pour le cas $y > z$ nous posons $y = z + s$ et nous substituons :

$$(y+s)^2 + x = z^2 + z + t \tag{1.207}$$

qui donne

$$x = z + t - 2zs - s^2 < 2z - 2zs - s^2 = 2z(1-s) - s^2 \leqslant -s < 0 \tag{1.208}$$

parce que $s \geqslant 1$, donc $1 - s \leqslant 0$.

(2) **f est surjective**
Nous devons prouver que tous les éléments de $\mathbb{N}$ sont dans l'image de $\mathbb{N} \times \mathbb{N}$ par f. En premier lieu, $0 = f(0,0)$. C'est un bon début. Soit $a \in \mathbb{N}$ non nul ; nous montrons que tous les nombres de a^2 à $(a+1)^2$ sont des images de f. D'abord $a^2 = f(0,a)$, ensuite les nombres

$$f(1,a), f(2,a), \dots, f(a-1,a) \tag{1.209}$$

prennent les valeurs $a^2 + 1$, …, $a^2 + a - 1$. Enfin nous avons $f(a,0) = a^2 + a$ et les nombres $f(a,1), \dots, f(a,a)$ prennent les valeurs de $a^2 + a + 1$ à $a^2 + 2a = (a+1)^2 - 1$.

53. Et comme j'aime beaucoup cet article, il me fallait une excuse pour le placer ici.
http://www.madore.org/~david/weblog/d.2017-08-18.2460.html.

□

Sachez que cette fonction s'étend aux ordinaux (mais là ce n'est plus pour rigoler).

CORooNRPIooZPSmqa

Corolaire 1.136.
Il existe des parties $\{\mathbb{N}_i\}_{i\in\mathbb{N}}$ *telles que* $\bigcup_{i\in\mathbb{N}}\mathbb{N}_i=\mathbb{N}$ *et que chaque* $\mathbb{N}_i$ *soit en bijection avec* $\mathbb{N}$

Démonstration. Nous considérons la bijection $f\colon\mathbb{N}\to\mathbb{N}\times\mathbb{N}$ donnée par (l'inverse de celle donnée) par la proposition 1.135, et nous posons

$$\mathbb{N}_i=f^{-1}(i,\mathbb{N}). \tag{1.210}$$

L'application

$$f\colon\mathbb{N}_i\to\{(i,k)\}_{k\in\mathbb{N}} \tag{1.211}$$

est une bijection. Or l'ensemble $\{(i,k)\}_{k\in\mathbb{N}}$ est évidemment en bijection avec $\mathbb{N}$. Par composition nous avons le résultat. □

LEMooDLWFooNAJbbq

Lemme 1.137 ([1]).
Si il existe une surjection $\mathbb{N}\to A$*, alors* A *est fini ou dénombrable.*

Démonstration. Pour chaque $a\in A$, l'ensemble $f^{-1}(a)$ est une partie de $\mathbb{N}$.

(1) **Une application** La proposition 1.64(2) nous permet de poser

$$\begin{aligned}\sigma\colon A&\to\mathbb{N}\\ a&\mapsto\min\big(f^{-1}(a)\big).\end{aligned} \tag{1.212}$$

(2) **σ est injective** Supposons que $\sigma(a)=\sigma(b)$. Nous appelons x ce nombre :

$$x=\min\big(f^{-1}(a)\big)=\min\big(f^{-1}(b)\big). \tag{1.213}$$

Nous avons $x\in f^{-1}(a)\cap f^{-1}(b)$, ce qui implique que $f(x)=a$ et que $f(x)=b$; donc $a=b$. Donc σ est une injection.

(3) **A est infini** Si A est fini, le lemme est prouvé. Donc à partir de maintenant nous supposons que A est infini. Le but est de prouver qu'il est dénombrable, c'est-à-dire de construire une bijection $A\to\mathbb{N}$.

(4) **$\sigma(A)$ est dénombrable** Puisque $\sigma\colon A\to\mathbb{N}$ est injective et que A est infini, la proposition 1.124 dit que $\sigma(A)$ est infini dans $\mathbb{N}$. La proposition 1.132 nous dit alors que $\sigma(A)$ est dénombrable.
Soit une bijection $\varphi\colon\sigma(A)\to\mathbb{N}$.

(5) **La candidate bijection** Nous posons

$$f=\varphi\circ\sigma\colon A\to\mathbb{N} \tag{1.214}$$

et nous allons prouver que c'est une bijection.

(6) **Injective** Puisque φ et σ sont injectives, l'égalité $(\varphi\sigma)(a)=(\varphi\sigma)(b)$ implique immédiatement $a=b$.

(7) **Surjective** Soit $k\in\mathbb{N}$. Puisque φ et σ sont des injections, nous pouvons poser $a=(\sigma^{-1}\varphi^{-1})(k)$. Il est alors immédiat que $f(a)=k$.

□

PROPooENTPooSPpmhY

Proposition 1.138 ([1, 27]).
Une union dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable.

Démonstration. Soient A_i des ensembles finis ou dénombrables. Nous posons $A = \bigcup_{i\in\mathbb{N}} A_i$, et nous considérons les parties $\mathbb{N}_i$ du corolaire 1.136. Puisque A_i est dénombrable ou fini et que $\mathbb{N}_i$ est dénombrable, il existe une surjection $\varphi_i\colon \mathbb{N}_i \to A_i$.

Nous définissons $s\colon \mathbb{N} \to \mathbb{N}$ par $n \in \mathbb{N}_{s(n)}$, et nous posons enfin

$$\begin{aligned}\varphi\colon \mathbb{N} &\to A\\ n &\mapsto \varphi_{s(n)}(n).\end{aligned} \tag{1.215}$$

Nous prouvons que φ est surjective.

Soit $a \in A_i$. Il existe $n \in \mathbb{N}_i$ tel que $a = \varphi_i(n)$. Mais comme $n \in \mathbb{N}_i$, nous avons $s(n) = i$. Donc

$$a = \varphi_i(n) = \varphi_{s(n)}(n) = \varphi(n). \tag{1.216}$$

Donc $\varphi\colon \mathbb{N} \to A$ est surjective.

Le lemme 1.137 conclut que A est fini ou dénombrable. □

LEMooKFBAooHxgOsg

Lemme 1.139 ([1])**.**
Si A est dénombrable, alors l'ensemble des parties finies de A est dénombrable.

LEMooRXSRooBUWOyb

Lemme 1.140 ([1])**.**
Si N est un ensemble dénombrable, alors il existe une bijection $g\colon \{1,2\} \times N \to N$.

Démonstration. D'abord nous définissons une bijection $\varphi\colon \{0,1\} \times \mathbb{N} \to \mathbb{N}$ par

$$\begin{aligned}\varphi\colon \{0,1\} \times \mathbb{N} &\to \mathbb{N}\\ (n,k) &\mapsto 2k + n.\end{aligned} \tag{1.217}$$

Ensuite si $f\colon \mathbb{N} \to N$ est une bijection, il suffit de poser $g(n,k) = f\big(\varphi(n,k)\big)$. □

PROPooDMZHooXouDrQ

Proposition 1.141 (Produit d'ensembles dénombrables[27])**.**
Deux résultats.

(1) Si N est un ensemble dénombrable, alors pour tout $n \in \mathbb{N}$, l'ensemble N^n est dénombrable.

ITEMooWNBGooZjoIjZ

(2) Si A et B sont dénombrables, alors $A \times B$ est dénombrable.

Les ensembles dénombrables sont les plus petits ensembles infinis possibles, comme en témoigne la proposition suivante.

PROPooUIPAooCUEFme

Proposition 1.142.
Tout ensemble infini contient une partie en bijection avec $\mathbb{N}$.

Démonstration. Soient un ensemble infini E_0 et une partie propre E_1 en bijection avec E_0. Nous notons $\varphi\colon E_0 \to E_1$ une bijection.

Soit $x_0 \in E_0 \backslash E_1$ (axiome du choix et tout ça). Nous définissons

$$\begin{aligned}\psi\colon \mathbb{N} &\to \{\varphi^k(x_0)\}\\ n &\mapsto \varphi^n(x_0)\end{aligned} \tag{1.218}$$

et nous allons prouver que c'est une bijection. Que ce soit surjectif est immédiat. Pour l'injectivité, soit $\varphi^k(x_0) = \varphi^l(x_0)$ avec $k \neq l$. Supposons pour fixer les notations que $k > l$. Alors, vu que φ est inversible nous pouvons écrire

$$x_0 = \varphi^{k-l}(x_0) = \varphi\big(\varphi^{k-l-1}(x_0)\big) \tag{1.219}$$

où il est entendu que $\varphi^0(x_0) = x_0$. Cela signifie que x_0 est dans l'image de φ, c'est-à-dire dans E_1, ce que nous avons exclu par choix de x_0 dans $E_0 \backslash E_1$. Donc en réalité $\varphi^k(x_0) \neq \varphi^l(x_0)$ dès que $k \neq l$. □

PropQEPoozLqOQ

Proposition 1.143.
Toute partie d'un ensemble fini est finie, et toute partie d'un ensemble dénombrable est finie ou dénombrable.

Démonstration. Soient un ensemble E ainsi qu'une partie infinie $A \subset E$. Nous notons $\varphi: A \to A'$ une bijection entre A et une partie propre A' de A. Dans ce cas, l'application

$$\begin{aligned} \phi: E &\to (E\backslash A) \cup A' \\ x &\mapsto \begin{cases} x & \text{si } x \in E\backslash A \\ \varphi(x) & \text{si } x \in A \end{cases} \end{aligned} \tag{1.220}$$

est une bijection entre E et une partie propre de E. Donc E est infini.

Par contraposée nous déduisons que toute partie d'un ensemble fini est finie.

En ce qui concerne les parties d'ensembles dénombrables, soit une partie A d'un ensemble dénombrable E. Nous avons une bijection $t\varphi: E \to \mathbb{N}$. La restriction $\varphi: A \to \varphi(A)$ est une bijection entre A est une partie de $\mathbb{N}$.

— Si $\varphi(E)$ est infinie, elle est dénombrable (proposition 1.132). Dans ce cas A est en bijection avec un ensemble dénombrable. Il est donc dénombrable.

— Si $\varphi(E)$ est fini, alors le lemme 1.125(2) nous dit que A est fini.

□

LEMooGTOTooFbpvzU

Lemme 1.144.
Soit un ensemble E non dénombrable ainsi qu'une application $f: E \to F$ où F est un ensemble quelconque. Si $f(E)$ est dénombrable (ou fini), alors il existe $y \in f(E)$ tel que $f^{-1}(y)$ est indénombrable.

Démonstration. Nous avons

$$E = \bigcup_{y \in F} f^{-1}(y). \tag{1.221}$$

Si tous les $f^{-1}(y)$ sont dénombrables, alors E est une union dénombrable (F est dénombrable) d'ensembles dénombrables. Il serait donc dénombrable (proposition 1.138), ce qui est contraire à l'hypothèse. □

1.5.4 Théorème de Cantor-Schröder-Bernstein

LEMooTNMHooBpdzab

Lemme 1.145 ([28])**.**
Soient un ensemble A et une partie B de A. Si il existe une injection $f: A \to B$ alors il existe une bijection $\alpha: A \to B$.

Nous donnons deux preuves de ce lemme.

Première preuve de 1.145. Nous posons $Y = A\backslash B$ et nous décomposons la preuve en étapes.

(1) **Les $f^k(Y)$ sont disjoints** Vu que f prend ses valeurs dans B, nous avons $f^k(Y) \subset B$ pour tout k. Et vu que $Y = A\backslash B$, nous avons

$$f^k(Y) \cap Y = \varnothing \tag{1.222}$$

EQooDNHJooFJBrDq

pour tout k. L'application f étant injective, elle vérifie $f(C \cap D) = f(C) \cap f(D)$. Nous appliquons f^m des deux côtés de (1.222) :

$$f^{k+m}(Y) \cap f^m(Y) = \varnothing \tag{1.223}$$

pour tout $k, m \in \mathbb{N}$.

(2) **Une décomposition** Nous posons

$$X = \bigcup_{k\in\mathbb{N}} f^k(Y) = Y \cup \bigcup_{k=1}^{\infty} f^k(Y). \tag{1.224}$$

Vu que $f(X) \subset B$ nous avons l'égalité

$$B = f(X) \cup \big(B\backslash f(X)\big). \tag{1.225}$$

(3) $A\backslash X = B\backslash f(X)$ Supposons $x \in A\backslash X$. Vu que $Y = A\backslash B$ est dans X, l'élément x n'est pas dans $A\backslash B$ et donc est dans B parce qu'il est dans A. Mais x n'est pas dans X et en particulier pas dans $f(X)$ parce que $f(X) \subset X$. Donc x est dans $B\backslash f(X)$.
Dans l'autre sens, nous supposons que $x \in B\backslash f(X)$. Vu que $B \subset A$ nous avons $x \in A$. Comme x est hors de $f(X)$, il est hors des $f^k(Y)$ pour $k \geqslant 1$. Mais $x \in B$, donc x est hors de $A\backslash B = f^0(Y)$. Donc x est hors de $f^k(Y)$ pour tout $k \geqslant 0$. Donc x est hors de X.

(4) **La bijection** Nous considérons l'application

$$\begin{aligned} \alpha\colon A &\to B \\ x &\mapsto \begin{cases} f(x) & \text{si } x \in X \\ x & \text{si } x \in A\backslash X. \end{cases} \end{aligned} \tag{1.226}$$

Nous démontrons dans les points suivants que α est bijective.

(5) **Injective** Nous supposons $\alpha(x) = \alpha(y)$. Il y a 4 possibilités suivant que x et y soient dans X ou $A\backslash X$.
Si $x, y \in X$ alors $f(x) = f(y)$ et donc $x = y$ parce que f est injective.
Si $x \in X$ et $y \in A\backslash X$, alors $f(x) = y$. Mais $f(x) \in f(X)$ et $y \in A\backslash X = B\backslash f(X)$. Donc l'élément $f(x) = y$ est dans $f(X) \cap \big(B\backslash f(X)\big) = \varnothing$. Il n'est donc pas possible d'avoir $\alpha(x) = \alpha(y)$ avec $x \in X$ et $y \in A\backslash X$.
Si $x \in A\backslash X$ et $y \in X$, c'est la même chose.
Si $x, y \in A\backslash X$, alors $x = \alpha(x) = \alpha(y) = y$.

(6) **Surjective** Soit $y \in B$. Il y a deux possibilités : $y \in X$ et $y \in A\backslash X$. La première se divise en deux : $y \in Y$ et $y \in \bigcup_{k=1}^{\infty} f^k(Y)$. On y va.

(6a) $y \in Y$ Ce cas n'est pas possible parce que $y \in B$ alors que $Y = A\backslash B$.

(6b) $y \in f^k(Y)$ **avec** $k \geqslant 1$ Nous avons

$$y \in f\big(f^{k-1}(Y)\big) \subset f(X) \subset \alpha(A). \tag{1.227}$$

(6c) $y \in A\backslash X$ Alors $y = \alpha(y)$.

□

Deuxième preuve de 1.145[13]. Nous posons $Y = A\backslash B$ et

$$\mathcal{M} = \{M \subset A \text{ tel que } Y \cup f(M) \subset M\}. \tag{1.228}$$

Nous allons dire de nombreuses choses à propos de ce $\mathcal{M}$.

(1) **$\mathcal{M}$ est non vide** Nous avons $A \in \mathcal{M}$ parce que Y et $f(A)$ sont dans A.

(2) **Si $M \in \mathcal{M}$ alors $Y \subset M$** C'est dans la définition de $\mathcal{M}$.

(3) **Encore un ensemble** Vu que $\mathcal{M}$ est non vide, nous pouvons poser

$$X = \bigcap_{M\in\mathcal{M}} M \tag{1.229}$$

sans nous poser trop de questions. Cela étant fait, nous pouvons passer aux choses sérieuses.

(4) $f(X) \subset M$ **pour tout** $M \in \mathcal{M}$ Soit $M \in \mathcal{M}$. Nous avons

$$\begin{aligned} f(X) &\subset f(M) && \text{(1.230a)}\\ &\subset Y \cup f(M) && \text{(1.230b)}\\ &\subset M && \text{(1.230c)} \end{aligned}$$

Justifications.

— Pour (1.230a). Parce que $X \subset \mathcal{M}$.

— Pour (1.230c). Parce que $M \in \mathcal{M}$.

(5) $X \in \mathcal{M}$ Vu que $Y \subset M$ pour tout M dans $\mathcal{M}$, nous avons $Y \subset \bigcap_{M\in\mathcal{M}} M = X$. Nous devons prouver $f(X) \subset X$. Nous venons de prouver que $f(X) \subset M$ pour tout $M \in \mathcal{M}$, donc

$$f(X) \subset \bigcap_{M\in\mathcal{M}} M = X. \tag{1.231}$$

L'ensemble X est le plus petit élément de $\mathcal{M}$ pour l'inclusion.

(6) $Y \cup f(X) \subset X$ Ça fait partie de $X \in \mathcal{M}$. Mais c'est bien de le dire explicitement parce que nous allons l'utiliser quelques fois dans la suite.

(7) $Y \cup f(X) \in \mathcal{M}$ Vu que $X \in \mathcal{M}$, nous savons déjà que $Y \cup f(X) \subset X$. En appliquant f des deux côtés,

$$f\big(Y \cup f(X)\big) \subset f(X). \tag{1.232}$$

En ajoutant Y des deux côtés,

$$Y \cup f\big(Y \cup f(X)\big) \subset Y \cup f(X), \tag{1.233}$$

et donc $Y \cup f(X) \in \mathcal{M}$.

(8) $Y \cup f(X) = X$ Nous savons déjà que $Y \cup f(X) \in \mathcal{M}$. Vu que X est le plus petit élément de $\mathcal{M}$, nous avons

$$X \subset Y \cup f(X). \tag{1.234}$$

L'inclusion inverse étant déjà faite, nous avons l'égalité.

(9) $B\backslash f(X) = A\backslash X$ Nous avons :

$$\begin{aligned} A\backslash X &= A\backslash\big(Y \cup f(X)\big) && \text{(1.235a)}\\ &= (A\backslash Y) \cap \big(A\backslash f(X)\big) && \text{(1.235b)}\\ &= B \cap \big(A\backslash f(X)\big) && \text{(1.235c)}\\ &= (B \cap A)\backslash f(X) && \text{(1.235d)}\\ &= B\backslash f(X). && \text{(1.235e)} \end{aligned}$$

Justifications :

— Pour (1.235b). Complémentaire de réunion, lemme 1.26(2).

— Pour (1.235c). Parce que $A\backslash Y = B$ du fait que $Y = A\backslash B$.

— Pour (1.235d). Lemme 1.26(3).

— Pour (1.235e). Parce que $B \subset A$.

(10) **Notre bijection** Nous voulons définir

$$\begin{aligned} \alpha\colon A &\to B\\ x &\mapsto \begin{cases} f(x) & \text{si } x \in X\\ x & \text{si } x \in A\backslash X. \end{cases} \end{aligned} \tag{1.236}$$

Pour y parvenir, nous devons prouver que α prend effectivement ses valeurs dans B. Ensuite nous prouverons que α est une bijection.

(11) α **prend ses valeurs dans** B Si $x \in X$ nous avons $\alpha(x) = f(x) \in B$. Si au contraire $x \in A\backslash X$ nous avons
$$\alpha(x) = x \in A\backslash X = B\backslash f(X) \subset B. \tag{1.237}$$

(12) α **est injective** Soient $x, y \in A$ tels que $\alpha(x) = \alpha(y)$. Il y a quatre possibilités suivant que x et y sont dans X ou dans $A\backslash X$.

(12a) $x \in X$, $y \in X$ Alors $f(x) = \alpha(x) = \alpha(y) = f(y)$. Vu que f est injective, nous trouvons que $x = y$.

(12b) $x \in X$, $y \in A\backslash X$ Nous avons $\alpha(x) = f(x) \in f(X)$ et $\alpha(y) = y \in A\backslash X = B\backslash f(X)$. Il n'est donc pas possible d'avoir $\alpha(x) = \alpha(y)$ dans ce cas.

(12c) $x \in A\backslash X$, $y \in X$ Idem.

(12d) $x \in A\backslash X$, $y \in A\backslash X$ Dans ce cas nous avons $\alpha(x) = x$ et $\alpha(y) = y$. Donc $x = y$.

(13) α **est surjective** Soit $b \in B$. Il y a deux possibilités : $b \in f(X)$ ou $b \notin f(X)$.

(13a) **Si** $b \in f(X)$ Soit $x \in X$ tel que $f(x) = b$. Alors $\alpha(x) = f(x) = b$.

(13b) **Si** $b \notin f(X)$ Alors $b \in B\backslash f(X) = A\backslash X$, et donc $\alpha(b) = b$.

□

THOooRYZJooQcjlcl

Théorème 1.146 (Cantor-Schröder-Bernstein)**.**
Soient deux ensembles A et B pour lesquels il existe des injections $f\colon A \to B$ et $g\colon B \to A$. Alors il existe une bijection entre A et B.

Démonstration. La composée $g \circ f\colon A \to A$ est injective et prend ses valeurs dans $g\big(f(A)\big) \subset g(B) \subset A$. Bref, l'application $g \circ f\colon A \to g(B)$ est injective. Le lemme 1.145 donne alors une bijection $\varphi\colon A \to g(B)$.

Nous montrons que $g^{-1} \circ \varphi\colon A \to B$ est une bijection.

(1) **Injective** Nous supposons $x, y \in A$ tels que
$$g^{-1}\big(\varphi(x)\big) = g^{-1}\big(\varphi(y)\big). \tag{1.238}$$
Nous appliquons g des deux côtés : $\varphi(x) = \varphi(y)$. Puisque φ est une bijection, cela entraîne $x = y$.

(2) **Surjective** Soit $b \in B$. En posant $a = \varphi^{-1}\big(g(b)\big)$ nous avons bien $(g^{-1} \circ \varphi)(a) = b$.

□

1.5.5 Comparabilité cardinale

Le théorème de comparabilité cardinale énonce que si A et B sont des ensemble, alors nous avons toujours $A \geq B$ ou $A \leq B$ (ou les deux ; dans ce cas $A \approx B$ par Cantor-Schröder-Bernstein).

THOooCBSKooCmzfUf

Théorème 1.147 (Théorème de comparabilité cardinale[1, 29, 30])**.**
Entre deux ensembles, il existe forcément une injection de l'un dans l'autre.

Démonstration. Nous allons montrer que le graphe d'une injection de A dans B ou de B dans A est donné par un élément maximal (au sens de l'inclusion) de l'ensemble (inductif) des graphes d'injections d'une partie de A dans une partie de B.

Nous allons utiliser le lemme de Zorn 1.23 à l'ensemble [54]
$$\mathcal{A} = \left\{ (X, Y, \varphi) \text{ tel que } \begin{cases} X \subset A \\ Y \subset B \\ \varphi\colon X \to Y \text{ est injective.} \end{cases} \right\} \tag{1.239}$$

54. Attention : dans le Frido, la notation $f\colon A \to B$, signifie que f est définie sur tout l'ensemble A, mais pas qu'elle est surjective sur B.

que nous ordonnons par l'inclusion, c'est-à-dire par $(X_1, Y_1, \varphi_1) < (X_2, Y_2, \varphi_2)$ lorsque $X_1 \subset X_2$, $Y_1 \subset Y_2$ et $\varphi_2|_{X_1} = \varphi_1$.

Nous passons rapidement sur le fait que cet ensemble est inductif, et nous considérons tout de suite un élément maximal (X, Y, φ). Il y a deux possibilités : soit $\varphi(X) = B$, soit $\varphi(X) \neq B$.

(1) **Si $\varphi(X) = B$** Dans ce cas, $\varphi: X \to B$ est une surjection. L'ensemble A est donc surpotent à B. La proposition 1.116 conclut que B est subpotent à A.

(2) **Si $\varphi(X) \neq B$** Nous subdivisons en deux nouveaux cas : soit $X = A$, soit $X \neq A$.

(2a) **Si $X = A$** Alors nous avons une injection $\varphi: A \to B$, et c'est bon.

(2b) **Si $X \neq A$** Nous sommes dans le cas $X \neq A$ et $\varphi(X) \neq B$. Soient $a \in A \backslash X$ et $b \in B \backslash \varphi(X)$. Nous considérons l'application

$$\begin{aligned} \psi: X \cup \{a\} &\to Y \cup \{b\} \\ x &\mapsto \begin{cases} \varphi(x) & \text{si } x \in X \\ b & \text{si } x = a. \end{cases} \end{aligned} \tag{1.240}$$

C'est une application injective. Donc le triplet $(X \cup \{a\}, Y \cup \{b\}, \psi)$ majore (X, Y, φ). Nous avons une contradiction et ce cas n'est pas possible.

□

1.148.
Le théorème de comparabilité cardinale couplé au théorème de Cantor-Schröder-Bernstein nous indique que pour tout ensembles A et B, nous avons soit $A \preceq B$, soit $B \preceq A$. Et si $A \preceq B \preceq A$, alors $A \approx B$.

Nous ne sommes pas loin de dire que la relation $\preceq$ donne un ordre total sur l'ensemble des ensembles. C'est très beau sauf que l'ensemble des ensembles n'existe pas [55]. Il faudrait parler de *classe* des ensembles, mais ça nous mènerait trop loin. Toujours est-il que ces deux théorèmes montrent qu'on n'est pas loin d'avoir un ordre sur les ensembles, et que cela est une des bases possibles pour développer les nombres cardinaux.

1.5.6 Théorème de Cantor, ensemble des ensembles

THOooJPNFooWSxUhd

Théorème 1.149 (Cantor[31])**.**
Un ensemble est toujours strictement subpotent à son ensemble des parties.

Démonstration. Soit un ensemble E et son ensemble des parties $\mathcal{P}(E)$. Nous commençons par prouver qu'il n'existe pas de surjection $E \to \mathcal{P}(E)$. Soit en effet une application $f: E \to \mathcal{P}(E)$. Nous posons

$$D = \{x \in E \text{ tel que } x \notin f(x)\}. \tag{1.241}$$

Nous prouvons que D ne peut pas être dans l'image de f. Supposons que $y \in E$ soit tel que $f(y) = D$.

(1) **Si $y \in D$** Alors par définition de D, nous avons $y \notin f(y) = D$. Contradiction.

(2) **Si $y \notin D$** Alors $y \in f(y) = D$, contradiction.

Donc aucune surjection $f: E \to \mathcal{P}(E)$ n'existe. En particulier pas de bijections.

Par ailleurs, l'application $g: \mathcal{P}(E) \to E$ qui fait $g(\{a\}) = a$ (et n'importe quoi d'autre sur les autres éléments de $\mathcal{P}(E)$) est une surjection $\mathcal{P}(E) \to E$.

Donc $\mathcal{P}(E)$ est toujours strictement surpotent à E. □

1.150.
Le théorème de Cantor implique en particulier qu'il existe (au moins) une infinité dénombrable d'ensembles infinis de cardinalité différentes (plus évidemment une infinité dénombrable d'ensembles finis de cardinalité différentes).

55. Corolaire 1.151.

Pour tout ensemble A, il est donc possible de dire « soit E, un ensemble strictement surpotent à A ».

CORooZMAOooPfJosM

Corolaire 1.151.
Il n'existe pas d'ensemble contenant tous les ensembles.

Démonstration. Si E était un tel ensemble, nous aurions $\mathcal{P}(E) \subset E$ parce que les éléments de $\mathcal{P}(E)$ sont des ensembles. Or cela donnerait une surjection $E \to \mathcal{P}(E)$ alors que cela est impossible par le théorème de Cantor 1.149. □

1.5.7 Ajouter ou soustraire des cardinalités

Nous allons prouver une série de résultats que nous pourrions résumer en « ajouter ou retrancher des parties de cardinalité plus petite ne change pas la cardinalité ».

LEMooUFCAooSyZtZj

Lemme 1.152 ([1])**.**
Si A est infini et B est fini, alors $A \cup B \approx A$.

Démonstration. Nous supposons que A et B sont disjoints[56]. La proposition 1.126 nous permet de considérer une bijection $\psi\colon \{1, \dots, n\} \to B$.

Puisque A est infini, la proposition 1.142 nous permet de considérer $N \subset A$ et une bijection $\varphi\colon \mathbb{N} \to N$.

Maintenant, il s'agit seulement d'insérer B dans A en le mettant « au début » de N et en décalant les autres éléments. La bijection est

$$\begin{aligned} f\colon A \cup B &\to A \\ x &\mapsto \begin{cases} x & \text{si } x \in A\backslash N \\ \varphi\big(\varphi^{-1}(x) + n\big) & \text{si } x \in N \\ \varphi\big(\psi^{-1}(x)\big) & \text{si } x \in B. \end{cases} \end{aligned} \tag{1.242}$$

□

LEMooXMVDooIWLWis

Lemme 1.153.
Si A est infini et si A est surpotent à B, alors $A \approx A \cup B$.

Démonstration. Il existe évidemment une injection $A \to A \cup B$. Donc le théorème de Cantor-Schröder-Bernstein 1.146 nous indique que trouver une injection $A \cup B \to A$ suffira pour la peine.

Nous allons utiliser le lemme de Zorn 1.23 avec l'ensemble

$$\mathcal{A} = \left\{ (X, \varphi_X) \text{ tel que } \begin{cases} X \subset B \\ \varphi_X\colon A \cup X \to A \text{ est injective.} \end{cases} \right\} \tag{1.243}$$

muni de l'ordre de l'inclusion : $(X, \varphi_X) < (Y, \varphi_Y)$ si $X \subset Y$ et $\varphi_Y(x) = \varphi_X(x)$ pour tout $x \in A \cup X$.

(1) <u>$\mathcal{A}$ est inductif</u> Soit une famille $\mathcal{F} = \{(X_i, \varphi_i)\}_{i \in I}$ complètement ordonnée indexée par l'ensemble I. En posant $X = \bigcup_{i \in I} X_i$ et $\varphi(x) = \varphi_i(x)$ dès que $x \in A \cup X_i$, l'élément (X, φ) majore $\mathcal{F}$.

(2) <u>Un maximum</u> Le lemme de Zorn nous assure que $\mathcal{A}$ possède (au moins) un élément maximal. Soit un tel élément maximum (X, φ).

(3) <u>$X \approx B$</u> Ah oui, vous auriez aimé avoir $X = B$. Mais non ; il n'y a pas de garantie. Nous allons montrer que $X \approx B$, et ça suffira.
Vu que $X \subset B$, si X n'est pas équipotent à B, il est strictement inclus dans B. Nous pouvons donc considérer

$$b \in B\backslash X \tag{1.244a}$$
$$a \in A\backslash\varphi(X). \tag{1.244b}$$

56. Adaptez la démonstration au cas où l'intersection n'est pas vide.

Nous considérons alors l'élément $(Y,\psi)\in\mathcal{A}$ défini par

$$Y=X\cup\{b\} \tag{1.245a}$$

$$\psi(x)=\begin{cases}a & \text{si } x=b\\ \varphi(x) & \text{sinon .}\end{cases} \tag{1.245b}$$

Cet élément majore (X,φ).
Donc $X\approx B$.

(4) **Résumé de la situation** Nous avons $A\approx A\cup X$ ainsi qu'une injection $\varphi\colon A\cup X\to A$ et une bijection $\psi\colon B\to X$.

(5) **Conclusion si A est disjoint de B** Si A et B sont disjoints, nous avons une bijection

$$\begin{aligned}l\colon A\cup B&\to A\\ x&\mapsto\begin{cases}\varphi(x) & \text{si } x\in A\\ \varphi\big(\psi(x)\big) & \text{si } x\in B.\end{cases}\end{aligned} \tag{1.246}$$

(6) **Conclusion si A n'est pas disjoint de B** Il suffit de poser $C=B\backslash A$ et nous avons

$$A\cup B=[A\cup(A\cap B)]\cup C. \tag{1.247}$$

Cette union est disjointe, $A\cup(A\cap B)$ est surpotent à A et C est subpotent à B. La conclusion est donc encore valable.

□

La proposition suivante sera utilisée en théorie de la mesure, dans l'exemple 14.73.

PropVCSooMzmIX

Proposition 1.154 ([32, 13])**.**
Si S est un ensemble infini alors il existe une bijection $\varphi\colon\{0,1\}\times S\to S$.

Démonstration. Nous posons $A=\{0\}\times S$ et $B=\{1\}\times S$. L'ensemble A est infini et surpotent à B (pas strictement, mais quand même).

Donc A est idempotent à $A\cup B$ par le lemme 1.153. Mais A est idempotent à S, donc

$$S\approx A\approx A\cup B=\{0,1\}\times S. \tag{1.248}$$

□

CORooJCSIooOeOICJ

Corolaire 1.155.
Si A est un ensemble infini, alors A possède deux sous-ensembles disjoints A_1 et A_2 qui sont tous deux en bijection avec A.

Démonstration. La proposition 1.154 donne une bijection $\varphi\colon\{1,2\}\times A\to A$. Il suffit de poser $A_1=\varphi(1,A)$ et $A_2=\varphi(2,A)$. □

Maintenant que nous pouvons mettre dans A deux copies disjointes de A, il n'est pas très étonnant que nous puissions en mettre une infinité dénombrable. C'est en substance ce que signifie la proposition suivante.

PROPooFKBEooKXqujV

Proposition 1.156.
Si A est infini, alors $A\times\mathbb{N}\approx A$.

Démonstration. La démonstration se base sur le fait qu'à l'intérieur de A, nous pouvons construire autant de copies de A deux à deux disjointes que nous le voulons. La k^{e} « copie » sera naturellement l'image de $k\times A$.

Voyons tout cela en détail.

(1) **Ce que nous allons faire** Nous allons construire, pour tout $i \geqslant 1$ des parties $A_i, B_i \subset A$ telles que
   — $A_i \cap B_i = \varnothing$,
   — $A_i, B_i \subset B_{i-1}$,
   — $A_i \approx B_i \approx A$,
   — $A_i \cap A_j = \varnothing$ si $i \neq j$

(2) **La construction** Nous commençons à zéro en utilisant le corolaire 1.155 pour construire des parties disjointes A_0 et B_0 de A telles que $A_0 \approx B_0 \approx A$.
   Ensuite, puisque $B_0 \approx A$, il existe A_1 et B_1 dans B_0 tels que $A_1 \cap B_1 = \varnothing$ et $A_1 \approx B_1 \approx B_0 \approx A$. Cela est notre construction pour $i = 1$.
   Pour la récurrence, puisque $A_i \approx B_i \approx A$, nous considérons A_{i+1} et B_{i+1} dans B_i tels que $A_{i+1} \cap B_{i+1} = \varnothing$ et $A_{i+1} \approx B_{i+1} \approx B_i \approx A$. C'est encore le corolaire 1.155 qui fait le travail.

(3) **Les propriétés** Nous avons $A_i \cap A_{i+1} = \varnothing$ parce que $A_i \cap A_{i+1} \subset A_i \cap B_i = \varnothing$.
   Nous devons encore montrer que $A_i \cap A_j = \varnothing$ dès que $i \neq j$. Supposons que $j > i$. Nous avons les inclusions
   $$A_j \subset B_{j-1} \subset B_{j-2} \subset \ldots \subset B_i. \tag{1.249}$$
   Donc $A_j \cap A_i \subset B_i \cap A_i = \varnothing$.

(4) **Une injection** Nous pouvons à présent écrire une injection qui termine presque la preuve. Pour cela nous considérons pour tout i, une bijection $\psi_i : A \to A_i$. Ensuite nous posons
   $$\begin{aligned} \varphi : A \times \mathbb{N} &\to A \\ (a, k) &\mapsto \psi_k(a). \end{aligned} \tag{1.250}$$
   Si $\varphi(a, k) = \varphi(b, l)$, alors $\psi_k(a) = \psi_l(b)$. L'élément $\psi_k(a)$ est donc dans $A_k \cap A_l$; ce n'est possible que si $k = l$. Donc $\psi_l(a) = \psi_l(b)$. Cette dernière égalité n'est possible que si $a = b$ parce que ψ_l est une bijection.
   Donc φ est une injection, et nous avons prouvé que $A \times \mathbb{N} \preceq A$.

(5) **La bijection** Nous venons de prouver que $A \times \mathbb{N} \preceq A$. La surpotence $A \times \mathbb{N} \succeq A$ étant évidente, le théorème de Cantor-Schröder-Bernstein 1.146 conclut que $A \times \mathbb{N} \approx A$.

□
LEMooDHWSooFqhano

Lemme 1.157 ([1]).
Soit un ensemble A muni de deux sous-ensembles B et B' équipotents et disjoints. Alors $A\backslash B$ est équipotent à $A\backslash B'$.

Démonstration. Soit une bijection $\psi : B' \to B$. L'application
$$\begin{aligned} \varphi : A\backslash B &\to A\backslash B' \\ x &\mapsto \begin{cases} x & \text{si } x \notin B' \\ \psi(x) & \text{si } x \in B'. \end{cases} \end{aligned} \tag{1.251}$$
est la bijection cherchée. □
LEMooIVCBooHWQiZB

Lemme 1.158 ([33]).
Si A est un ensemble infini et si $B \prec A$, alors $A \approx A\backslash B$.

Démonstration. Nous pouvons écrire
$$A = (A\backslash B) \cup B. \tag{1.252}$$
Le théorème de comparabilité cardinale 1.147 nous indique que soit $A\backslash B \preceq B$, soit $A\backslash B \succeq B$. Nous allons étudier les deux cas.

(1) **Si $A\backslash B \geq B$** Dans ce cas, $(A\backslash B) \cup B \approx A\backslash B$ par le lemme 1.153. Alors, notre résultat est prouvé parce que $A = (A\backslash B) \cup B \approx A\backslash B$.

(2) **Si $A\backslash B \leq B$** Dans ce cas, le lemme 1.153 nous indique que $A = (A\backslash B) \cup B \approx B$. Mais $A \approx B$ est exclu par l'hypothèse. Ce cas est donc impossible.

□

LEMooMRVQooUZSSyL

Lemme 1.159 ([1]).
Si A est infini et si B est une partie strictement subpotente de A, alors il existe $U \subset A$ disjoint de B et équipotent à B.

Démonstration. Le lemme 1.158 nous donne une bijection $\varphi\colon A \to A\backslash B$. Il suffit alors de poser $U = \varphi(B)$. Cette partie est disjointe de B parce que φ prend ses valeurs dans $A\backslash B$. □

Lemme 1.160 ([34]).
Soit un ensemble infini A ainsi qu'un sous-ensemble $B \subset A$. Nous supposons l'existence d'une fonction surjective $f\colon B \to B \times B$.

Alors $B \leq B \times B \leq B \leq A$.

Démonstration. La première est l'hypothèse sur f. La seconde est l'existence (évidente) d'une surjection $B \times B \to B$. La troisième est le fait que B soit inclus dans A. □

LEMooPOEFooXaifhT

Lemme 1.161 ([34]).
Soit un ensemble infini A ainsi qu'un sous-ensemble strictement subpotent $B \subset A$. Nous supposons l'existence d'une fonction surjective $f\colon B \to B \times B$.

Alors f peut être étendue en une injection $f\colon D \to D \times D$ où $D \subset A$ contient strictement B.

Démonstration. Par le lemme 1.159, nous considérons une partie $U \subset A$ disjointe de B et équipotente à B. Nous pouvons écrire le développement

$$(B \cup U) \times (B \cup U) = (B \times B) \cup (B \times U) \cup (U \times B) \cup (U \times U). \tag{1.253}$$

Nous savons que B est surpotent à U (il est même équipotent) ; donc le lemme 1.153 nous dit que $B \cup U \approx B$. De plus il existe une bijection $B \to U$, donc

$$U \approx B \approx B \times B \approx B \times U \approx U \times B \approx U \times U. \tag{1.254}$$

Chacun des ensembles $U \times B$, $B \times U$ et $U \times U$ est équipotent à U. Leur union est donc équipotente [57] à U et nous avons une bijection

$$\varphi\colon U \to U \times B \cup (B \times U) \cup (U \times U). \tag{1.255}$$

Et enfin nous définissons

$$\begin{aligned} g\colon B \cup U &\to (B \times U) \times (B \cup U) \\ x &\mapsto \begin{cases} f(x) & \text{si } x \in B \\ \varphi(x) & \text{si } x \in U. \end{cases} \end{aligned} \tag{1.256}$$

Cette définition est bonne parce que U et B sont disjoints, et g est injective. □

Le théorème suivant est une généralisation de la proposition 1.154. Elle implique, entre autres choses, qu'il existe une bijection entre $\mathbb{R}$ et $\mathbb{R} \times \mathbb{R}$. Pour le cas de $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$, il y a la proposition 1.135 qui donne une bijection explicite et donc sans axiome du choix et sans lemme de Zorn.

THOooDGOVooRdURVi

Théorème 1.162.
Si A est infini, alors $A \approx A \times A$.

57. Lemme 1.153.

Démonstration. Une fois de plus, ce sera le lemme de Zorn qui va s'y coller. Soit l'ensemble

$$\mathcal{A} = \Big\{ (X,\varphi) \text{ tel que } \begin{cases} X \subset A \\ \varphi\colon X \to X \times X \text{ est surjective.} \end{cases} \Big\} \tag{1.257}$$

Cet ensemble est non vide parce que A est infini ; il contient donc une partie dénombrable N, et nous connaissons la surjection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ du lemme 1.135.

Nous ordonnons $\mathcal{A}$ par l'inclusion : $(X,\varphi) \leqslant (Y,\phi)$ lorsque $X \subset Y$ et $\phi|_X = \varphi$. La tambouille usuelle montre que $\mathcal{A}$ est un ensemble inductif et le lemme de Zorn 1.23 donne l'existence d'un élément maximal que nous notons (B,φ).

Puisque B est subpotent à A (parce qu'il est inclus), soit il est strictement subpotent, soit il est équipotent. Nous commençons par montrer que B ne peut pas être strictement subpotent à A.

En effet, si nous avions une surjection $B \to B \times B$, alors que B est strictement subpotent à A. Le lemme 1.161 nous dit alors que φ peut être étendue, ce qui contredirait la maximalité de (B,φ).

Donc la partie B est équipotente à A : il existe une bijection $g\colon A \to B$. Mais nous avons une surjection $B \to B \times B$ et donc aussi une injection $B \times B \to B$. Vu que nous avons par ailleurs une injection $B \to B \times B$, le théorème de Cantor-Schröder-Bernstein 1.146 nous donne une bijection $\phi\colon B \times B \to B$. Avec ça, l'application

$$\begin{aligned} f\colon A \times A &\to A \\ (a,b) &\mapsto \phi\big(g(a),g(b)\big) \end{aligned} \tag{1.258}$$

est une bijection. Donc les ensembles A et $A \times A$ sont équipotents. □

LEMooNKKDooUvSYPO

Lemme 1.163.
Si A est infini, et si pour tout $i \in \mathbb{N}$ nous avons $A_i \approx A$, alors

$$\bigcup_{i\in\mathbb{N}} A_i \approx A. \tag{1.259}$$

Démonstration. Pour chaque $i \in \mathbb{N}$ nous avons une bijection $\varphi_i\colon A_i \to A$. Nous posons

$$\begin{aligned} \varphi\colon A \times \mathbb{N} &\to \bigcup_{i=0}^{\infty} A_i \\ (a,i) &\mapsto \varphi_i(a). \end{aligned} \tag{1.260}$$

EQooCHJAooRpHvnV

Cette application est surjective mais peut-être pas injective parce que les A_i peuvent avoir des intersections non vides. Nous avons alors le calcul

$$\begin{aligned} A &\approx A \times \mathbb{N} && \text{(1.261a)} \\ &\geq \bigcup_{i=0}^{\infty} A_i && \text{(1.261b)} \\ &\geq A && \text{(1.261c)} \end{aligned}$$

SUBEQooICFEooTLuFHZ
SUBEQooRVPRooJJcvkv
SUBEQooFJRGooJncvvv

Justifications :

— Pour (1.261a), c'est la proposition 1.156.

— Pour (1.261b), c'est la surjection (1.260).

— Pour (1.261c), c'est le fait que seulement A_0 possède déjà une surjection vers A.

Donc $\bigcup_i A_i$ est à la fois surpotent et subpotent à A. Il est donc équipotent par le théorème 1.146. □

1.6 Groupes

1.6.1 Définition, unicité du neutre

La définition d'un groupe est la définition 1.37.

LEMooECDMooCkWxXf

Lemme-Définition 1.164 (Unicités)**.**
Dans un groupe, l'inverse et le neutre sont uniques. Plus précisément, si G est un groupe nous avons :

(1) il existe un unique élément $e \in G$ tel que $eg = ge = g$ pour tout $g \in G$, ITEMooOIWTooYqmMPP

(2) pour tout $g \in G$, il existe un unique élément $h \in G$ tel que $gh = hg = e$.

Le e ainsi défini est nommé ***neutre*** *de G. Le h tel que $gh = hg = e$ est nommé l'****inverse*** *de g et est noté g^{-1}.*

Démonstration. Chaque point séparément.

(1) Supposons que e_1 et e_2 vérifient la propriété. Nous avons pour tout $g \in G$: $e_1 g = g e_1 = g$. En particulier pour $g = e_2$ nous écrivons $e_1 e_2 = e_2 e_1 = e_2$. Mais en partant dans l'autre sens : $e_2 g = g e_2 = g$ avec $g = e_1$ nous avons $e_2 e_1 = e_1 e_2 = e_1$. En égalant ces deux valeurs de $e_2 e_1$ nous avons $e_1 = e_2$.
Pour la suite de la preuve nous écrivons e l'unique neutre de G.

(2) Supposons que k_1 et k_2 soient deux inverses de g. On considère alors le produit $k_1 g k_2$. Puisque $k_1 g = e$, on a $k_1 g k_2 = e k_2 = k_2$; mais, comme $g k_2 = e$, on a aussi $k_1 g k_2 = k_1 e = k_1$. Le produit est donc à la fois égal à k_1 et à k_2, et donc $k_1 = k_2$.

□

LEMooWYLRooNOdZnp

Lemme 1.165.
Soient deux groupes G et H. Si $\alpha : G \to H$ est un morphisme de groupes[58]*, alors*

(1) $\alpha(e_G) = e_H$.
(2) $\alpha(g^{-1}) = \alpha(g)^{-1}$.

LEMooBIBFooBHxFYC

Lemme 1.166.
Si G est un groupe et si $h \in G$, alors les applications

$$\begin{aligned} L_h : G &\to G \\ g &\mapsto hg \end{aligned} \tag{1.262}$$

et

$$\begin{aligned} R_h : G &\to G \\ g &\mapsto gh \end{aligned} \tag{1.263}$$

sont des bijections.

Démonstration. D'abord si $L_h(g_1) = L_h(g_2)$, alors $hg_1 = hg_2$ et en multipliant à gauche par h^{-1} nous avons $g_1 = g_2$; donc L_h est injective. Ensuite L_h est surjective parce que si $g \in G$, alors $g = L_h(h^{-1}g)$.
Pour l'application R_h, la preuve est une simple adaptation. □

1.6.2 Groupe ordonné

DEFooEUHFooYvhnLQ

Définition 1.167 ([35])**.**
Soient un groupe $(G, +)$, ainsi qu'une relation d'ordre $\leqslant$ sur G. Nous disons que la relation d'ordre est ***compatible*** *avec la structure de groupe si pour tout $x, y, z \in G$, si $x \leqslant y$ alors $x + z \leqslant y + z$ et $z + x \leqslant z + y$. Dans ce cas, le triple $(G, +, \leqslant)$ est un* ***groupe ordonné****.*
Si $(G, \leqslant)$ est totalement ordonné, nous disons que le groupe est totalement ordonné.

58. Définition 1.38.

1.6.3 Classes de conjugaison

DEFooOLXPooWelsZV

Définition 1.168 (classe de conjugaison)**.**
*Soit un groupe G et un élément $g \in G$. La **classe de conjugaison** de g est la partie*

$$C_g = \{kgk^{-1} \text{ tel que } k \in G\}. \tag{1.264}$$

LEMooQYBJooYwMwGM

Lemme 1.169.
Un groupe est commutatif si et seulement si ses classes de conjugaison sont des singletons.

Démonstration. Supposons que G soit commutatif. Alors

$$C_g = \{kgk^{-1} \text{ tel que } k \in G\} = \{g\}. \tag{1.265}$$

Donc les classes de conjugaison sont des singletons.

Dans l'autre sens, si les classes sont des singletons, on a $kgk^{-1} = g$ pour tous $k, g \in G$. Cela signifie immédiatement que G est commutatif. □

defGroupeCentre

Définition 1.170 (centralisateur[36])**.**
*Soient un groupe G, un sous-groupe H et un élément $h \in H$. Le **centralisateur** de h dans G est l'ensemble des éléments de G qui commutent avec h :*

$$Z_G(h) = \{z \in G \text{ tel que } hz = zh\}. \tag{1.266}$$

Le centralisateur de H dans G est l'ensemble des éléments de G qui commutent avec tous les éléments de H :

$$Z_G(H) = \bigcap_{h \in H} Z_G(h). \tag{1.267}$$

*Le **centre** d'un groupe G est l'ensemble des éléments de G qui commutent avec tous les autres :*

$$Z_G = Z_G(G) = \{z \in G \text{ tel que } gz = zg, \forall g \in G\}. \tag{1.268}$$

DEFooZTSMooBislIy

Définition 1.171 (normalisateur[36])**.**
*Soient un groupe G et un sous-groupe H. Le **normalisateur** de H dans G est*

$$\mathcal{N}_G(H) = \{g \in G \text{ tel que } gH = Hg\}. \tag{1.269}$$

DEFooNIIMooFkZgvX

Définition 1.172 (Sous-groupe normal)**.**
*Un sous-groupe N de G est **normal** ou **distingué** si pour tout $g \in G$ et pour tout $n \in N$, $gng^{-1} \in N$. Autrement dit lorsque $gNg^{-1} \subset N$.*

Lorsque N est normal dans G il est parfois noté $N \lhd G$.

DEFooUXXTooCCLmQe

Définition 1.173.
*Un sous-groupe H de G est un sous-groupe **caractéristique** si $\alpha(H) \subset H$ pour tout automorphisme[59] α de G.*

Lemme 1.174 ([37])**.**
Si H est un sous-groupe caractéristique de G, alors $\alpha(H) = H$ pour tout automorphisme α de G.

Démonstration. Si α est un automorphisme de G, alors α^{-1} est encore un automorphisme de G. En particulier $\alpha^{-1}(H) \subset H$.

Soit $h \in H$. Nous devons prouver que $h \in \alpha(H)$. Pour cela :

$$h = \alpha\big(\alpha^{-1}(h)\big) \in \alpha\big(\alpha^{-1}(H)\big) \subset \alpha(H). \tag{1.270}$$

□

DefGroupeSimple

Définition 1.175 (Groupe simple)**.**
*Un groupe est dit **simple** si il est non trivial et si les seuls sous-groupes normaux qu'il admet sont lui-même et le sous-groupe réduit à l'élément neutre.*

59. Automorphisme de groupe, définition 1.38.

1.7 Anneaux

DEFooKWKGooIOwGTA

Définition 1.176.
*Un **isomorphisme d'anneaux** est un morphisme d'anneaux*[60]*, bijectif.*

La distributivité de la partie (3) de la définition 1.42 ne traite que de l'addition ; pas de la soustraction. Voici une lemme qui dit que ça fonctionne quand même.

LEMooVPYUooRzexke

Lemme 1.177 ([13])**.**
Soient un anneau A ainsi que $a, b, c \in A$. Alors

$$a(b-c) = ab - ac. \tag{1.271}$$

Démonstration. Nous avons le calcul suivant :

SUBEQooKCOWooFeQHUM

$$a(b-c) + ac = a\big((b-c)+c\big) \tag{1.272a}$$

SUBEQooMLLOooNRmIYM

$$= ab. \tag{1.272b}$$

Justifications :

— Pour 1.272a. Distributivité.

— Pour 1.272b. Parce que $(b-c)+c = b$.

Nous avons donc $a(b-c)+ac = ab$ et donc l'égalité demandée en ajoutant $-ac$ des deux côtés. □

LEMooVUSMooWisQpD

Lemme 1.178.
Pour tout élément a d'un anneau nous avons $a \cdot 0 = 0$.

Démonstration. L'élément 0 est le neutre de l'addition. Il peut être écrit $1-1$, et en utilisant la distributivité sous la forme du lemme 1.177,

$$a \cdot 0 = a \cdot (1-1) = a - a = 0. \tag{1.273}$$

Notons que la dernière égalité s'écrit en détail $a-a = a+(-a)$ qui donne le neutre de l'addition. □

PROPooNCCGooXjVyVt

Proposition 1.179.
Dans un anneau[61] *non nul, le neutre pour l'addition est distinct du neutre pour la multiplication.*

Démonstration. Supposons par contraposée que dans un anneau A, nous ayons $1 = 0$. Alors, pour tout $a \in A$, on a $a = 1a = 0a = (1-1)a = a - a = 0$, $a = 0$. Autrement dit $a = 0$ pour tout $a \in A$ et l'anneau est nul. □

LEMooLTERooVKgqjn

Lemme 1.180 ([1])**.**
Un peu d'arithmétique. Soit un anneau A et un élément $a \in A$.

ITEMooUGHCooOPgoeR

(1) $1 \times 1 = 1$.

ITEMooJMBSooVgvVwg

(2) $(-1) \times a = -a$.

ITEMooXJGMooKNLlHU

(3) $-(-a) = a$.

ITEMooYMRKooHVYYKU

(4) $(-1) \times (-1) = 1$.

Démonstration. En plusieurs parties.

(1) **Pour (1)** La définition de 1 est que $1 \times a = a$ pour tout a. En particulier pour $a = 1$ nous avons le résultat.

60. Définition 1.43.
61. Définition 1.42.

(2) **Pour (2)** Nous avons

$$(-1) \times a + a = a \times \big((-1) + 1\big) = a \times 0 = 0. \tag{1.274}$$

Nous avons utilisé le fait que la multiplication était distributive et que le zéro était absorbant (lemme 1.178).

(3) **Pour (3)** Nous avons $-a+a = 0$ par définition de la notation $-a$. Donc a est bien l'inverse de $-a$ pour l'addition.

(4) **Pour (4)** En utilisant les points (2) et (3) nous avons

$$(-1) \times (-1) = -(-1) = 1. \tag{1.275}$$

□

Soit X un ensemble et un anneau $(A, +, \times)$. Nous considérons $\mathrm{Fun}(X, A)$ l'ensemble des applications $X \to A$. Cet ensemble devient un anneau avec les définitions

$$(f + g)(x) = f(x) + g(x) \tag{1.276a}$$
$$(fg)(x) = f(x)g(x). \tag{1.276b}$$

C'est la **structure canonique** d'anneau sur $\mathrm{Fun}(X, A)$.

Définition 1.181.
*Le **centralisateur** de $x \in A$ dans A est l'ensemble*

$$\{y \in A \text{ tel que } xy = yx\}, \tag{1.277}$$

*le **centre** de A est*

$$\{y \in A \text{ tel que } xy = yx, \forall x \in A\}. \tag{1.278}$$

DefooQULAooREUIU

Définition 1.182 (Idéal dans un anneau)**.**
*Un sous-ensemble $I \subset A$ est un **idéal à gauche** si*

(1) I est un sous-groupe pour l'addition,

(2) pour tout $a \in A$, $aI \subset I$.

*De même nous disons que $I \subset A$ est une **idéal à droite** lorsque I est un sous-groupe pour l'addition et $Ia \subset I$ pour tout $a \in A$.*

*Lorsqu'un ensemble est idéal à gauche et à droite, nous disons que c'est un **idéal bilatère**. Lorsque nous parlons d'idéal sans précision, nous parlons d'idéal bilatère.*

LEMooMAHXooXSowdn

Lemme 1.183.
Les seuls idéaux d'un corps sont $\{0\}$ et le corps lui-même.

Démonstration. Soient un corps $\mathbb{K}$ et un idéal I dans A. Si $I = \{0\}$, c'est un idéal, pas de problèmes. Si $I \neq \{0\}$, alors $0 \in I$ parce qu'un idéal doit contenir le neutre de l'addition.

Soit $x \neq 0$ dans I. Alors pour tout $a \in \mathbb{K}$ nous avons $ax \in I$. En particulier avec $a = x^{-1}$ nous voyons que $1 \in I$. De là, $I = \mathbb{K}$ parce que si $x \in \mathbb{K}$, nous avons $x = x \cdot 1 \in xI \subset I$. □

PROPooGXMRooTcUGbi

Proposition-Définition 1.184.
Soit A, un anneau, I un idéal bilatère[62] de A. Nous considérons la relation d'équivalence $x \sim y$ si et seulement si $x - y \in I$. Sur le quotient[63]

$$A/\sim = A/I, \tag{1.279}$$

nous mettons les opérations

62. Définition 1.182.
63. Définition 1.33.

(1) $[x]+[y]=[x+y]$

(2) $[x][y]=[xy]$.

Nous avons alors les résultats suivants :

ITEMooEJPEooRKAqmS

(1) Les opérations sont bien définies,

ITEMooYBEGooTlHgNz

(2) l'ensemble A/I, muni de ces opérations, est un anneau. Le neutre pour l'addition est $[0]$, l'inverse de $[a]$ est $[-a]$ que nous noterons $-[a]$.

ITEMooLNRLooMkoWXZ

(3) la surjection canonique $\pi\colon A\to A/I$ est un morphisme.

Cet anneau est appelé ***anneau quotient****.*

Démonstration. En plusieurs parties.

(1) **Pour (1)** Nous savons que, par définition,

$$\bar{x}=\{x+i \text{ tel que } i\in I\}. \tag{1.280}$$

Calculons le produit de représentants génériques de $\bar{x}$ et de $\bar{y}$:

$$(x+i_1)(y+i_2)=xy+xi_2+yi_1+i_1i_2. \tag{1.281}$$

Puisque I est un idéal, nous avons $xi_2+yi_1+i_1i_2\in I$ et donc bien

$$(x+i_1)(y+i_2)\in\overline{xy}. \tag{1.282}$$

(2) **Pour (2)** Il s'agit de vérifier les conditions de la définition 1.42.
D'abord A/I est un groupe de neutre $[0]$. En effet, vu que $(A,+)$ est un groupe commutatif de neutre 0, nous avons

(2a) Neutre : $[a]+[0]=[a+0]=[a]$.

(2b) Associativité : $[a]+([b]+[c])=[a]+[b+c]=[a+b+c]=[a+b]+[c]=\big([a]+[b]\big)+[c]$.

(2c) Inversibilité : l'inverse de $[a]$ est $[-a]$ parce que $[a]+[-a]=[a-a]=[0]$.

Nous pouvons noter $-[a]$ l'élément $[-a]$. Le groupe A/I est commutatif :

$$[a]+[b]=[a+b]=[b+a]=[b]+[a]. \tag{1.283}$$

Donc $(A/I,+)$ est un groupe commutatif de neutre $[0]$.
L'associativité de A donne l'associativité dans A/I :

$$\big([a][b]\big)[c]=[ab][c]=[abc]=[a][bc]=[a]\big([b][c]\big). \tag{1.284}$$

Et enfin pour la distributivité,

$$[a]\big([b]+[c]\big)=[a][b+c]=[a(b+c)]=[ab+ac]=[ab]+[ac]=[a][b]+[a][c]. \tag{1.285}$$

Nous avons prouvé que A/I est un anneau de neutre $[0]$ et d'unité $[1]$.

(3) **Pour (3)** Nous devons vérifier les trois conditions de la définition 1.43. Cela est immédiat parce que $\pi(x)=[x]$.

□

1.7.1 Diviseurs

DiviseursAnneau

Définition 1.185 (Diviseurs dans un anneau)**.**
Soient $a,b\in A$. On dit que a divise b, ou que a est un ***diviseur (à gauche)*** *de b si il existe $c\in A$ tel que $ac=b$. On dit que c'est un diviseur de b à droite si $ca=b$ pour un certain $c\in A$.*

Un cas particulier est le cas des diviseurs de zéro. L'absence de tels diviseurs dans un anneau est une propriété intéressante : on dit dans ce cas que l'anneau est intègre. Définition 1.197.

Un élément $a \in A$ est **régulier à droite** si $ba = 0$ implique $b = 0$. Il est régulier à gauche si $ab = 0$ implique $b = 0$.

DefrYwbct

Définition 1.186.
*Soient A un anneau commutatif et $S \subset A$. Nous disons que $\delta \in A$ est un **PGCD** de S si*

(1) δ divise[64] *tous les éléments de S.*

ITEMooVCKGooWDXZOj

(2) si d divise également tous les éléments de S, alors d divise δ.

*Nous disons que $\mu \in A$ est un **PPCM** de S si*

(1) $S \mid \mu$,

(2) si $S \mid m$, alors $\mu \mid m$.

Si P et Q sont des polynômes, ce que nous notons $\operatorname{pgcd}(P,Q)$ *est l'unique polynôme unitaire dans* $\operatorname{pgcd}\big(\{P,Q\}\big)$. *Voir 6.53.*

Remarque 1.187.
Au sens de la définition 1.186, le pgcd n'est pas unique. Dans $\mathbb{Z}$ par exemple les nombres 4 et -4 sont tous deux pgcd de $\{4, 16\}$.

Dans $\mathbb{Z}$ cependant, nous modifions implicitement la définition et nous n'acceptons que les positifs, de telle sorte que l'unique pgcd soit effectivement le plus grand pour l'ordre usuel sur $\mathbb{Z}$.

Pour l'unicité dans $\mathbb{Z}$, voir 3.19.

1.7.2 Élément irréductible et premier

DEFooZCRQooWXRalw

Définition 1.188 ([38])**.**
*Soit un anneau commutatif A. Un élément $p \in A$ est **premier** si il est*

(1) non nul,

(2) non inversible,

ITEMooPMTTooCVHPIm

(3) si p divise un produit ab, alors il divise soit a soit b (ou les deux).

DeirredBDhQfA

Définition 1.189 (Élément irréductible[39])**.**
*Un élément d'un anneau commutatif est **irréductible** si il n'est ni inversible, ni le produit de deux éléments non inversibles.*

1.190.
Nous allons voir dans la section 3.6 que le concept d'élément irréductible n'est vraiment utile que dans le cas des anneaux intègres.

Exemple 1.191.
Un corps n'a pas d'élément irréductible parce qu'à part zéro, tous les éléments sont inversibles. Mais 0 n'est pas irréductible parce qu'il peut être écrit comme produit d'éléments non inversibles : $0 = 0 \cdot 0$. △

LEMooJBUJooScsiGc

Lemme 1.192 ([1])**.**
Si p est irréductible et si u est inversible, alors pu est irréductible.

Démonstration. D'abord pu n'est pas inversible parce que p ne l'est pas.

Ensuite supposons que $pu = ab$. Vu que u est inversible, nous avons $p = a(bu^{-1})$. Comme p est irréductible, soit a, soit bu^{-1} est inversible.

Si c'est a, c'est gagnée. Sinon, soit k un inverse de bu^{-1} : $bu^{-1}k = 1$. Nous voyons que $u^{-1}k$ est un inverse de b. Donc b est inversible. □

64. Définition 1.185.

PROPooKDWQooTtScrN

Proposition 1.193.
Les éléments irréductibles de l'anneau $\mathbb{Z}$ sont les nombres premiers[65].

Démonstration. Les seuls inversibles de $\mathbb{Z}$ sont ± 1.

Si p est premier et $p = ab$ avec $a, b \in \mathbb{Z}$, alors nous avons soit $a = \pm 1$ soit $b = \pm 1$. Donc p n'est pas le produit de deux éléments non inversibles.

Dans le sens inverse, supposons que p soit irréductible dans $\mathbb{Z}$. D'abord p ne peut pas être ± 1 parce que ± 1 sont inversibles. Ensuite supposons que $p = ab$. Vu que p est irréductible, nous avons $a = \pm 1$ ou $b = \pm 1$. Autrement dit, dans $p = ab$, soit a soit b est un inversible. □

1.7.3 Anneau intègre

DEFooHRRYooTmbUTH

Définition 1.194 (Éléments nilpotents, unipotents)**.**
*On dit que $a \in A$ est **nilpotent** si il existe $n \in \mathbb{N}$ tel que $a^n = 0$. Il est dit **unipotent** si $a - 1$ est nilpotent, c'est-à-dire si $(a-1)^n = 0$ pour un certain $n \in \mathbb{N}$.*

DEFooCIHVooAhpJxy

Définition 1.195 (Éléments inversibles)**.**
*Un élément $a \in A$ est dit **inversible** si il existe $b \in A$ tel que $ab = 1$.*

L'ensemble $U(A)$ des éléments inversibles de A est un groupe pour la multiplication. Nous notons $A^* = A \backslash \{0\}$.

Conformément à la définition 1.185 de diviseur, nous posons la définition suivante pour les diviseurs de zéro.

DEFooCIYLooFkhVOc

Définition 1.196 (diviseur de zéro[40])**.**
*Un élément $a \neq 0$ est un **diviseur de zéro à gauche** si il existe $x \neq 0$ tel que $ax = 0$. L'élément a est un diviseur de zéro **à droite** si il existe $y \neq 0$ tel que $ya = 0$.*

*Nous disons que a est un **diviseur de zéro** si il est un diviseur de zéro à gauche ou à droite.*

DEFooTAOPooWDPYmd

Proposition-Définition 1.197 (Anneau intègre[1])**.**
Soit A un anneau non réduit à $\{0\}$. Les assertions suivantes sont équivalentes :

ITEMooMXMKooXMYpkN

(1) A ne possède pas de diviseurs de zéro[66].

ITEMooLAJCooFwxXrV

(2) La règle du produit nul s'applique dans A : pour tous $a, b \in A$, si $ab = 0$, alors $a = 0$ ou $b = 0$.

ITEMooQNTFooSRrVPK

(3) On peut simplifier par un même élément non-nul, deux expressions produit dans A qui sont égales : pour tous $a, b, c \in A$ avec $a \neq 0$, si $ab = ac$, alors $b = c$.

*Un anneau non réduit à $\{0\}$ qui vérifie ces propriétés est dit **intègre**.*

Démonstration. En trois implications.

(1) **(1) implique (2)**
Si $ab = 0$ avec $b \neq 0$ alors a est un diviseur de zéro. Vu que nous supposons que A n'a pas de diviseurs de zéros, a est nul. De même, si $a \neq 0$ b devrait être nul.

(2) **(2) implique (3)**
Si $ab = ac$, alors $a(b-c) = 0$ et l'hypothèse dit que soit $a = 0$, soit $b - c = 0$. Donc si $a \neq 0$, alors $b - c = 0$.

(3) **(3) implique (1)** Si $A = \{0\}$, le point (3) n'est pas applicable.
Si $a \neq 0$ et $ax = 0$, alors nous avons aussi $ax = a \times 0$. Par propriété de simplification, $x = 0$. Donc a n'est pas un diviseur de zéro à gauche. Nous prouvons de la même façon qu'il n'y a pas de diviseurs de zéro à droite.

65. Nombre premier, définition 1.188.
66. Définition 1.196.

□

LEMooIKNMooMfvQnu

Lemme 1.198.
Un corps est un anneau intègre[67].

Démonstration. Nous vérifions la définition 1.197, et nous nommons $\mathbb{K}$ le corps considéré.

(1) **Pour (1)** Soit $a \in \mathbb{K}$. Si $ax = 0$ avec $x \neq 0$ alors en multipliant par x^{-1} nous trouvons $a = 0$. Donc a n'est pas un diviseur de zéro non nul.

(2) **Pour (2)** Idem à ce que nous venons de faire. Si dans $ab = 0$ l'un des deux est non nul, en multipliant par son inverse, nous trouvons que l'autre est nul.

(3) **Pour (3)** Si $ab = ac$ avec $a \neq 0$, alors il suffit de multiplier à gauche par a^{-1} (qui existe parce que nous sommes dans un corps) pour obtenir $b = c$.

□

Conséquence : dans un corps nous avons toujours la règle du produit nul, et l'élément nul n'est jamais inversible.

PROPooZBTIooRhAhvg

Proposition 1.199 ([41])**.**
Soit un anneau unitaire et intègre A. Soit $p \in A$.

ITEMooJWRYooHndNpV

(1) p est premier si et seulement si l'idéal pA est premier.

ITEMooGHGCooRkJilg

(2) p est irréductible si et seulement si il n'existe pas d'idéal principal I tel que $pA \subsetneq I \subsetneq A$.

Démonstration. En plusieurs parties.

(1) **(1)⇒**
Supposons que p est premier. Soient $a, b \in A$ tels que $ab \in pA$. En particulier $p \mid ab$, et p étant premier[68], nous avons soit $p \mid a$ soit $p \mid b$. Supposons que $p \mid a$. Il existe $k \in A$ tel que $pk = a$, et donc

$$a = pk \in pA. \tag{1.286}$$

De même si $p \mid b$ nous avons $b \in pA$.

(2) **(1)⇐**
Nous supposons que l'idéal pA est premier, et nous prouvons que p est premier. Soient $a, b \in A$ tels que $p \mid ab$. Il existe $k \in A$ tel que $pk = ab$. Donc $ab \in pA$. L'idéal pA étant premier, nus avons soit $a \in pA$ soit $b \in pA$. Donc soit $a \mid a$ soit $b \mid p$.

(3) **(2)⇒**
Supposons que p est irréductible. Soit un idéal principal I vérifiant $pA \subset I$. Nous allons montrer que soit $I = pA$ soit $I = A$. Vu que I est principal, il existe $a \in A$ tel que $I = aA$. Nous avons $pA \subset aA$, et en particulier $p \in aA$. Notons $b \in A$ un élément tel que $ab = p$.
Vu que p est irréductible, il n'est pas le produit de deux non inversibles. Donc soit a soit b est inversible.
Si a est inversible, alors $I = A$.
Si b est inversible, alors $a = pb^{-1}$, de telle sorte que $a \in pA$. De ce fait $I = pA$.

(4) **(2)⇐**
Enfin, nous supposons qu'il n'existe pas d'idéal principal I tel que $pA \subsetneq I \subsetneq A$, et nous montrons que p est irréductible. Soient $a, b \in A$ tels que $p = ab$.
Nous avons $p \in aA$ et donc $pA \subset aA$. Donc soit $aA = pA$ soit $aA = A$.
Si $aA = pA$, alors il existe $k \in A$ tel que $pk = a$. En multipliant l'égalité $ab = p$ par k nous trouvons $abk = pk = a$. Vu que A est intègre, nous pouvons simplifier par a et trouver $bk = 1$, de telle sorte que b soit inversible.
Si $aA = A$, alors a est inversible parce que $1 \in A = aA$.

67. Définition 1.197.
68. Définition 1.188.

□
LEMooSFHMooQoKsPV

Lemme 1.200 ([1]).
Soient un anneau intègre A, et une partie $S \subset A$. Si un des pgcd *de S est inversible*[69]*, alors ils le sont tous.*

Démonstration. Pour rappel, les pgcd d'une partie de A sont définis dans 1.186. Soit un pgcd inversible de S, ainsi qu'un autre pgcd que nous nommons δ'. Vu que δ' divise tous les éléments de S, il est divisé par δ : $\delta \mid \delta'$. Réciproquement, $\delta' \mid \delta$.

Soient x et y définis par $\delta = x\delta'$ et $\delta' = y\delta$. Nous avons

$$\delta = x\delta' = xy\delta. \tag{1.287}$$

Comme l'anneau A est intègre, nous pouvons simplifier par δ et voir $xy = 1$, ce qui signifie que x et y sont inversibles. Donc si δ est inversible, alors $\delta' = y\delta$ est inversible. □
LEMooMHZQooIcSNSf

Lemme 1.201 ([1, 42]).
Soient un anneau intègre, A, une partie $S \subset A$ et un élément $a \in A$. Nous avons[70]

$$\text{pgcd}(aS) = a\,\text{pgcd}(S). \tag{1.288}$$

Démonstration. Deux inclusions à prouver.

(1) $\text{pgcd}(aS) \subset a\,\text{pgcd}(S)$ Soit un pgcd δ de aS. Nous devons trouver un $\delta' \in \text{pgcd}(S)$ tel que $\delta = a\delta'$. En termes de notations, nous notons $S = \{s_i\}_{i\in I}$. Pour chaque i nous avons $\delta \mid as_i$: il existe $x_i \in A$ tel que

$$\delta x_i = as_i. \tag{1.289}$$

Vu que a divise tous les éléments de aS, il divise n'importe quel pgcd de aS, et en particulier $a \mid \delta$: il existe $\delta' \in A$ tel que $\delta = a\delta'$. Nous montrons que $\delta' \in \text{pgcd}(S)$.

Nous savons que $\delta x_i = as_i$. En remplaçant δ par $a\delta'$, $a\delta' x_i = as_i$. Vu que nous sommes dans un anneau intègre, nous pouvons simplifier par a (définition 1.197(3)) :

$$\delta' x_i = s_i. \tag{1.290}$$

Donc δ' divise tous les éléments de S, et vérifie la première condition pour être un pgcd de S. Pour la seconde condition, nous supposons que d divise tous les éléments de S. Nous avons $d \mid S$, donc $ad \mid aS$. Et comme δ est un pgcd de aS, nous déduisons que ad divise δ. Il existe $y \in A$ tel que

$$ady = \delta. \tag{1.291}$$

Nous remplaçons δ par sa valeur $a\delta'$: $ady = a\delta'$. Encore une fois nous simplifions par a et nous trouvons $dy = \delta'$, c'est-à-dire que δ divise δ'.

(2) $a\,\text{pgcd}(S) \subset \text{pgcd}(aS)$
Soit un pgcd δ de S. Nous voulons que $a\delta \in \text{pgcd}(aS)$. Vu que $\delta \in \text{pgcd}(S)$ nous avons $\delta x_i = s_i$ pour tout i, et donc aussi $a\delta x_i = as_i$, de telle sorte que $a\delta$ divise tous les éléments de aS.

Soit maintenant $d \in A$ divisant tous les éléments de aS. Nous devons prouver que $d \mid a\delta$.

(2a) **Travail préliminaire**
Nous considérons $\delta' \in \text{pgcd}(aS)$. Vu que $\delta \mid s$ pour tout $s \in S$, nous avons aussi $a\delta \mid as$ pour tout $s \in S$. Comme δ' est un est un pgcd de aS, nous avons donc

$$a\delta \mid \delta'. \tag{1.292}$$

Soit $u \in A$ tel que $\delta' = a\delta u$.

69. Définition 1.195.
70. Définition du pgcd : 1.186.

En utilisant la première partie de la preuve, nous avons

$$\delta' \in \operatorname{pgcd}(aS) \subset a \operatorname{pgcd}(S). \tag{1.293}$$

Donc il existe $\delta_1 \in \operatorname{pgcd}(S)$ tel que $\delta' = a\delta_1$. En écrivant l'égalité $\delta' = a\delta u$ avec cette valeur de δ', nous trouvons

EQooVHSSooDdVUeW
$$a\delta_1 = a\delta u, \tag{1.294}$$

et donc $\delta_1 = \delta u$ parce que A est intègre. Vu que $\delta_1 \in \operatorname{pgcd}(S)$, nous avons aussi $\delta u \in \operatorname{pgcd}(S)$. En particulier δu divise tous les éléments de S, et donc divise δ qui est un pgcd de S : $\delta u \mid \delta$. En multipliant par a,

$$\delta' = a\delta u \mid a\delta. \tag{1.295}$$

(2b) **Résumé** Nous avons considéré $\delta \in \operatorname{pgcd}(S)$ et nous sommes en train de prouver que $a\delta \in \operatorname{pgcd}(aS)$. Nous avons déjà prouvé que si $\delta' \in \operatorname{pgcd}(aS)$, alors nous avons $\delta' \mid a\delta$. Nous posons $y \in A$ tel que $\delta' y = a\delta$.

(2c) **Et enfin** Soit d divisant tous les éléments de $a\delta$. Donc $d \mid \delta'$: il existe $x \in A$ tel que $dx = \delta'$. En multipliant par y,

$$dxy = \delta' y = a\delta. \tag{1.296}$$

Nous avons montré que d divise $a\delta$, ce qu'il nous fallait.

□

LEMooZSUNooUmYmgt
Lemme 1.202 ([43]).
Soient un anneau intègre A et une partie $S \subset A$. Si $\delta \in \operatorname{pgcd}(S)$, alors $\operatorname{pgcd}(S/\delta)$ ne contient que des inversibles.

Démonstration. En utilisant le lemme 1.201, nous avons

$$\operatorname{pgcd}(S) = \operatorname{pgcd}\big(\delta(S/\delta)\big) = \delta \operatorname{pgcd}(S/\delta). \tag{1.297}$$

Soit $d \in \operatorname{pgcd}(S/\delta)$. Notre but est de montrer que d est inversible. D'abord

$$\delta d \in \delta \operatorname{pgcd}(S/\delta) = \operatorname{pgcd}(S). \tag{1.298}$$

Donc δd divise tous les éléments de S. Étant donné que δ est un pgcd de S, la définition 1.186(2) nous dit que δd divise δ. Il existe donc $d' \in A$ tel que $\delta dd' = \delta$, c'est à dire

$$\delta(dd' - 1) = 0. \tag{1.299}$$

Étant donné que $\delta \neq 0$ et que A est intègre nous avons $dd' - 1 = 0$ (définition 1.197(2)), c'est-à-dire $dd' = 1$. Cela montre que d' est un inverse de d, et donc que d est inversible. □

LEMooZKASooKstTuK
Lemme 1.203 ([1]).
Soit un anneau intègre A. Si $\delta \in \operatorname{pgcd}(S)$ et si $u \in A$ est inversible, alors $\delta u \in \operatorname{pgcd}(S)$.

Démonstration. Soit $s \in S$. Si $\delta x = s$, alors $u\delta(u^{-1}x) = s$. Donc $u\delta$ divise tous les éléments de S. De plus si $d \mid S$, alors $d \mid \delta$. Dans ce cas il existe y tel que $dy = \delta$. Nous avons alors aussi

$$dxu = \delta u, \tag{1.300}$$

de telle sorte que d divise δu. □

1.7.4 Positivité

DEFooZRMFooCtzMov

Définition 1.204 ([1]).
Soit un anneau A. Une positivité sur A est une partie $A^+ \subset A$ telle que

(1) Pour tout $a \neq 0$ dans A, nous avons[71] $a \in A^+$ *xor* $-a \in A^+$.

(2) Pour tout $a, b \in A^+$, nous avons $a + b \in A^+$.

Oui, il faudrait dire une stricte *positivité.*

PROPooKLOPooBgQqhM

Proposition 1.205 ([1]).
Si A^+ est une positivité sur A, alors en posant $a \leqslant b$ si et seulement si $b - a \in A^+ \cup \{0\}$, nous avons une relation d'ordre[72].

Démonstration. Nous vérifions les conditions de la définition 1.11.

(1) **Réflexivité** Vu que $a - a = 0$ et que $0 \in A \cup \{0\}$, nous avons $a \leqslant a$.

(2) **Antisymétrie** Si $a \leqslant b$ et $b \leqslant a$, nous avons

$$\begin{cases} b - a \in A^+ \cup \{0\} & (1.301a) \\ a - b \in A^+ \cup \{0\}. & (1.301b) \end{cases}$$

Entre $a - b$ et $b - a$, au maximum un des deux peut être dans A^+. Donc soit $a - b = 0$, soit $b - a = 0$ (ou le deux en même temps, ce qui est le cas).

(3) **Transitive** Supposons que $a \leqslant b$ et $b \leqslant c$. Nous avons alors

$$\begin{cases} b - a \in A^+ \cup \{0\} & (1.302a) \\ c - b \in A^+ \cup \{0\}. & (1.302b) \end{cases}$$

Si l'un des deux est nul, c'est facile. Nous supposons qu'ils sont tous les deux dans A^+. Alors leur somme est dans A^+ : $(b-a)+(c-b) \in A^+$. Cela donne tout de suite $c - a \in A^+$ et donc $a \leqslant c$ comme demandé.

□

1.7.5 Fonction puissance

Voici une première définition de la fonction puissance. Il y en aura d'autres, de plus en plus générales. Voir le thème 52.

DEFooGVSFooFVLtNo

Définition 1.206.
Si A est un anneau, si $a \in A$ et si $n \in \mathbb{N}$, nous définissons a^n par récurrence :

(1) $a^0 = 1$ (l'unité pour la multiplication dans A),

ITEMooOUIPooGjAgQb

(2) $a^{k+1} = a \cdot a^k$.

Si vous n'êtes pas sûr de vous, ne citez pas le théorème 1.48. Il est indispensable pour faire fonctionner cette définition, mais vous pouvez faire comme si vous n'avez rien vu.

Le lemme suivant dit que le point (2) de la définition 1.206 aurait pu être écrit $a^k \cdot a$ au lieu de $a \cdot a^k$.

LEMooWPARooYLZlzr

Lemme 1.207 ([1]).
Si A est un anneau, si $a \in A$ et si $n \in \mathbb{N}$, alors

$$a^n = a \cdot a^{n-1} = a^{n-1} \cdot a. \tag{1.303}$$

71. *xor* est le ou exclusif. L'un ou l'autre est vrai, mais les deux en même temps.
72. Définition 1.11.

Démonstration. Cela se prouve par récurrence. Pour $n = 1$ c'est l'égalité $a = a^0 a$ qui est correcte parce que par définition $a^0 = 1$.

Supposons que le résultat soit bon pour n et voyons ce que ça donne pour $n + 1$:

$$\begin{aligned} a^{n+1} &= aa^n && \text{Définition de } a^{n+1} && (1.304a)\\ &= a(a^{n-1}a) && \text{hypothèse de récurrence pour } a^n && (1.304b)\\ &= (aa^{n-1})a && \text{associativité} && (1.304c)\\ &= a^n a && \text{Définition de } a^n. && (1.304d) \end{aligned}$$

□

PROPooPLLSooUpiLKa

Proposition 1.208 ([1])**.**
Soit un naturel $b \in \mathbb{N}\backslash\{0\}$. Soit un naturel $m > 0$. Il existe $n \in \mathbb{N}$ tel que $b^n > m$.

1.8 Idéal dans un anneau

La définition d'un idéal dans un anneau est la définition 1.182.

DefTMNooKXHUd

Définition 1.209 ([44])**.**
*Un **corps** est un anneau*[73] *$(A, +, \times)$ dans lequel tout élément non nul est inversible pour l'opération $\times$ (pour l'opération $+$, tous les éléments sont inversibles parce que $(A, +)$ est un groupe).*

1.210.
Dans le Frido, nous ne parlons que de corps commutatifs ; nous ne le répéterons pas toujours.

1.211.
Pour savoir ce qu'est un « *ring* » ou « *field* » en anglais, voir -1.3.

DefSKTooOTauAR

Définition 1.212 (Idéal engendré par un élément)**.**
*Si p est un élément d'un anneau A alors nous notons (p) l'idéal dans A **engendré** par p, c'est-à-dire pA.*

DefAJVTPxb

Définition 1.213.
*Un sous-ensemble $B \subset A$ d'un anneau est un **sous anneau** si*

(1) $1 \in B$
(2) *B est un sous-groupe pour l'addition*
(3) *B est stable pour la multiplication.*

Remarque 1.214.
Un idéal n'est pas toujours un anneau parce que l'identité pourrait manquer. Un idéal qui contient l'identité est l'anneau complet.

LEMooQAYSooCYJXkC

Lemme 1.215.
L'ensemble $2\mathbb{Z}$ est un idéal[74] *de $\mathbb{Z}$. On peut aussi le noter (2).*

PROPooJALPooHFIObB

Proposition 1.216 (Premier théorème d'isomorphisme pour les anneaux)**.**
Soient A et B des anneaux et un homomorphisme $f : A \to B$. Nous considérons l'injection canonique $j : f(A) \to B$ et la surjection canonique $\phi : A \to A/\ker f$. Alors il existe un unique isomorphisme

$$\tilde{f} : A/\ker f \to f(A) \tag{1.305}$$

tel que $f = j \circ \tilde{f} \circ \phi$.

73. Définition 1.42.
74. Idéal, définition 1.182.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\phi}\downarrow & & \uparrow{\scriptstyle j} \\ A/\ker f & \xrightarrow[\tilde{f}]{} & f(A)\subset B \end{array} \tag{1.306}$$

PropIJJIdsousphi

Proposition 1.217.
Soient I un idéal de A et la projection canonique

$$\phi\colon A\to A/I. \tag{1.307}$$

C'est une bijection entre les idéaux de A contenant I et les idéaux de A/I.
Dit de façon imagée :

$$\{\textit{idéaux de } A \textit{ contenant } I\}\simeq\{\textit{idéaux de } A/I\}. \tag{1.308}$$

EqKbrizu

Démonstration. Si $I\subset J$ et si J est un idéal de A, alors $\phi(J)$ est un idéal dans A/I. En effet un élément de $\phi(J)$ est de la forme $\phi(j)$ et un élément de A/I est de la forme $\phi(i)$. Leur produit vaut

$$\phi(i)\phi(j)=\phi(ij)\in\phi(J). \tag{1.309}$$

Soit maintenant K un idéal dans A/I et soit $J=\phi^{-1}(K)$. Étant donné qu'un idéal doit contenir 0 (parce qu'un idéal est un groupe pour l'addition), $[0]\in K$ et par conséquent $I\subset\phi^{-1}(K)$. □

PROPooUOAVoGZpMVd

Proposition 1.218 ([1])**.**
Si A est un anneau, nous avons les équivalences

ITEMooLAAVooXhTcMe

(1) A est un corps[75]*.*

ITEMooDGZIooRopYGx

(2) A est non nul et ses seuls[76] *idéaux à gauche sont $\{0\}$ et A.*

ITEMooLPWHooDJpTbR

(3) A est non nul et ses seuls idéaux à droite sont $\{0\}$ et A.

Démonstration. Nous allons montrer que le point (1) est équivalent aux deux autres.

(1) **(1) implique (2)** Si I est un idéal à gauche différent de $\{0\}$, alors il contient un certain $a\neq 0$. Puisque A est un corps, il contient un inverse a^{-1}, et comme I est un idéal, $a^{-1}I\subset I$. En particulier $a^{-1}a\in I$. Donc $1\in I$ et $I=A$.

(2) **(2) implique (1)** Supposons que les seuls idéaux de A soient $\{0\}$ et A. Soit $a\in A$. Si a est non nul, alors aA est un idéal de a. Vu qu'il contient $a\neq 0$, nous avons $aA=A$ (par hypothèse, un idéal qui n'est pas $\{1\}$ est A). En particulier, $1\in aA$, c'est-à-dire qu'il existe $b\in A$ tel que $ab=1$. L'élément a est donc inversible.

(3) **(1) implique (3)** Comme pour (1) implique (2).

(4) **(3) implique (1)** Comme pour (2) implique (1).

Notez que je n'ai pas vérifié les deux derniers points. Donc vous devriez le vérifier et m'écrire si il y a un problème. □

DEFIdealMax

Définition 1.219 ([45])**.**
Soit un anneau A. Un idéal $I\neq A$ est dit ***idéal maximal*** *si il n'existe pas d'idéal $J\neq A$ contenant strictement I.*

PROPooSHHWooCyZPPw

Proposition 1.220 (Thème 18)**.**
Un idéal I dans un anneau A est maximum si et seulement si A/I est un corps.

75. Définition 1.209.
76. Je vous laisse vous poser de grandes questions sur le fait que le vide est un idéal ou non.

Démonstration. Soit un idéal maximum $I \subset A$. Alors les idéaux contenant I sont A et I. L'application ϕ de la proposition 1.217 est une bijection, donc l'ensemble des idéaux de A/I ne contient que deux éléments. Les seuls idéaux de A/I sont donc $\{0\}$ et A/I ; donc A/I est un corps par la proposition 1.218.

Dans l'autre sens, c'est la même chose : si A/I est un corps, il possède exactement deux idéaux, donc A ne contient que deux idéaux contenant I. Donc I est un idéal maximum. □

THOooFWYLooOofaPa

Théorème 1.221 (Théorème de Krull[46])**.**
Pour tout idéal propre I d'un anneau commutatif A, il existe au moins un idéal maximal de A contenant I.

1.8.1 Division euclidienne

ThoDivisEuclide

Théorème-Définition 1.222 (Division euclidienne[47])**.**
Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$, avec $0 \leqslant r < b$, tel que

$$a = bq + r. \tag{1.310}$$

*L'opération $(a, b) \mapsto (q, r)$ ainsi définie est la **division euclidienne**. Le nombre q est le **quotient** et r est le **reste** de la division de a par b.*

Démonstration. Remarquons que $r = a - bq$, et donc, une fois l'existence et l'unicité de q établie, celle de r suivra.

(1) **Unicité** Nous supposons avoir $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tels que

$$\begin{cases} 0 \leqslant r < b & \text{(1.311a)} \\ a = qb + r. & \text{(1.311b)} \end{cases}$$

Ce système implique que

$$0 \leqslant a - qb < b. \tag{1.312}$$

En ajoutant qb dans les trois membres de cette inégalité,

$$qb \leqslant a < (q + 1)b. \tag{1.313}$$

Cela implique que

$$q = \max\{k \in \mathbb{Z} \text{ tel que } kb \leqslant a\}. \tag{1.314}$$

Donc q est unique et la relation $a = bq + r$ implique que r est également unique.
Soit

$$E = \{q \in \mathbb{Z} | bq \leqslant a\}.$$

La partie E est non vide (parce qu'elle contient $-|a|$) et admet un majorant : l'élément $|a|$. Elle admet donc un maximum q par le lemme 1.112. Ce maximum vérifie

$$bq \leqslant a < b(q + 1). \tag{1.315}$$

Cela donne $0 \leqslant a - bq < b$ et le résultat, en posant $r = a - qb$.

□

Le lemme suivant est souvent pris pour la définition d'un nombre premier lorsqu'on parle de $\mathbb{N}$ ou $\mathbb{Z}$.

Lemme 1.223 ([48, 1])**.**
Dans $\mathbb{N}$, un nombre est premier si et seulement si il admet exactement deux diviseurs entiers distincts.

Démonstration. En deux parties.

(1) $\Rightarrow$ Soit un élément premier $p \in \mathbb{N}$. Il y a trois possibilités : $p = 0$, $p = 1$ et $p > 1$.
Le nombre $p = 0$ n'est pas premier parce qu'il est nul. Le nombre $p = 1$ n'est pas premier parce qu'il est inversible. Donc nous savons que si p est premier, alors $p > 1$.
Un élément $p > 1$ dans $\mathbb{N}$ a toujours au moins deux diviseurs distincts : 1 et p. Soit un diviseur k de p. Il existe $l \in \mathbb{N}$ tel que $p = kl$. Vu que p est premier et divise le produit kl, il divise k ou l. Disons que p divise k. De cette façon p divise k et k divise p.
Il existe donc $n \in \mathbb{N}$ tel que $k = np$. En y substituant $p = kl$, on trouve $k = np = nkl$. En simplifiant par k, il vient

$$1 = nl, \tag{1.316}$$

ce qui prouve que $n = l = 1$ et donc que $k = p$ et donc que p n'a pas d'autres diviseurs que 1 et p.

(2) $\Leftarrow$ Nous supposons que $p \in \mathbb{N}$ ait exactement deux diviseurs entiers distincts. Nous vérifions que p vérifie les trois conditions de la définition 1.188.

(2a) $p \neq 0$ parce que 0 a nettement plus que deux diviseurs distincts.

(2b) $p \neq 1$ parce que 1 a exactement un diviseur. Donc p n'est pas inversible dans $\mathbb{N}$.

(2c) Soit p admettant exactement deux diviseurs distincts. Soit p divisant le produit ab' pour certains a et b' dans $\mathbb{N}$. Nous supposons que p ne divise pas a, et nous allons prouver que p divise b' en supposant d'abord que p ne divise pas b'.

(2c.iii) **Un ensemble** Pour cela nous posons

$$E = \{x \in \mathbb{N} \text{ tel que } p \mid ax, p \nmid x\}. \tag{1.317}$$

Nous posons $b = \min(E)$. Nous avons pour hypothèse que E est non vide ; en particulier $0 < b$.

(2c.iii) $b < p$ On vérifie que si $p + k \in E$ alors $k \in E$. Donc b ne peut pas être plus grand que p. Vu que p lui-même n'est pas dans E, nous avons $b < p$.

(2c.iii) **Division euclidienne** Nous effectuons la division euclidienne du théorème 1.222 :

$$p = mb + r. \tag{1.318}$$

En multipliant par a, $ar = ap - mab$. Vu que ab est un multiple de p $ap - mab$ est un multiple de p. En particulier ar est divisible en p.

(2c.iii) **La contradiction** Nous avons donc $r \in E$, alors que $r < b$. Impossible.

□
PROPooWMNPooZdvOBt

Proposition 1.224 ([49]).
Dans un anneau intègre[77] *tout élément premier est irréductible*[78].

Démonstration. Soit p, un élément premier dans un anneau intègre A.

(1) p **n'est pas inversible** Cela fait partie de la définition d'un élément premier.

(2) p **n'est pas un produit d'inversibles** Soient $a, b \in A$ tels que $p = ab$. Par le point (3) de la définition 1.188, p divise soit a soit b. Supposons que p divise a. Alors il existe $x \in A$ tel que $a = px$. En remettant dans $p = ab$ nous avons :

$$p = pxb. \tag{1.319}$$
EQooPYBGooLFHMJZ

Mais l'anneau est intègre et permet donc des simplifications par tout élément non nul. La relation 1.319 donne donc

$$1 = xb, \tag{1.320}$$

ce qui signifie que b est inversible.
Un travail similaire montre que a est inversible si p divise b.

77. Si pas intègre, voir l'exemple 1.226.
78. Toutes les définitions dans le thème 6.

□

Exemple 1.225.
Si nous avons l'égalité $7 = ab$ dans $\mathbb{Z}$, alors soit a soit b vaut 1. Mettons $a = 1$. Dans ce cas, $b = 7$ et n'est donc pas inversible. △

Sur un anneau non intègre, la notion d'élément premier n'est pas aussi intéressante que sur un anneau intègre. Par exemple la proposition 1.224 devient fausse.

EXooEIUEooCZCPMC

Exemple 1.226.
Soit l'anneau $\mathbb{Z}^2$. L'élément $(1,0)$ est premier mais pas irréductible.

(1) **$(1,0)$ est premier** L'élément $(1,0)$ est non nul ; ça c'est pas cher. Pour qu'il soit inversible, il faudrait $(1,0)(x,y) = (1,1)$. Entre autres, $0 \times y = 1$, ce qui est impossible. Donc il n'est pas inversible.

Supposons que $(1,0)$ divise le produit $(a,b)(c,d) = (ac,bd)$. Alors il existe (x,y) tel que $(1,0)(x,y) = (ac,bd)$. Cela signifie que $x = ac$ et $0 \times y = bd$. En particulier, soit $b = 0$ soit $d = 0$. Si $b = 0$, nous avons $(a,b) = (a,0)$ et effectivement, $(1,0)$ le divise.

(2) **$(1,0)$ n'est pas irréductible** Nous avons $(1,0) = (1,0)(1,0)$. Donc l'élément $(1,0)$ est le produit de deux éléments non inversibles.

△

1.9 Anneau principal et idéal premier

DEFooMZRKooBPLAWH

Définition 1.227.
Un idéal[79] *I dans A est **principal à gauche** si il existe $a \in I$ tel que $I = Aa$. Il est **principal à droite** si il existe $a \in I$ tel que $I = aA$. Nous disons qu'il est **principal** si il est principal à gauche et à droite.*

DEFooGWOZooXzUlhK

Définition 1.228.
*Un anneau est **principal** si*

(1) il est commutatif et intègre

(2) tous ses idéaux sont principaux[80].

Souvent pour prouver qu'un anneau est principal, nous prouvons qu'il est euclidien (définition 1.251) et nous utilisons la proposition 1.254 qui dit qu'un anneau euclidien est principal.

Une manière de prouver qu'un anneau n'est pas principal est de prouver qu'il n'est pas factoriel, théorème 3.80.

DEFooAQSZooVhvQWv

Définition 1.229.
*Nous disons qu'un idéal I dans A est **premier** si I est strictement inclus dans A et si pour tout $a,b \in A$ tels que $ab \in I$ nous avons $a \in I$ ou $b \in I$.*

LEMooYRPBooYxXXsi

Lemme 1.230.
L'idéal nul (réduit à $\{0\}$) est premier si et seulement si A est intègre.

Démonstration. En deux sens.

(1) **Si $\{0\}$ est premier** Alors $A \neq \{0\}$ parce que $I = \{0\}$ est propre (définition d'idéal premier). De plus, si $ab = 0$, alors $ab \in I$ qui est un idéal premier. Donc soit a soit b est dans I, c'est-à-dire que soit a soit b est nul. Donc A est intègre.

(2) **Si A est intègre** Alors $A \neq \{0\}$ et l'idéal $I = \{0\}$ est strictement inclus dans A. Si $ab \in I$, alors $ab = 0$ et comme A est intègre, soit a soit b est nul, c'est-à-dire appartient à I.

79. Idéal, définition 1.182.
80. Définition 1.227.

□

PROPooRUQKooIfbnQX

Proposition 1.231 ([45]).
Soit un anneau commutatif[81] *et un idéal I dans A.*

ITEMooUGBTooOGrnWl

(1) I est un idéal premier[82] *si et seulement si A/I est un anneau intègre.*

ITEMooGLXSooUjINqR

(2) I est un idéal maximal[83] *si et seulement si A/I est un corps.*

ITEMooTFFQooOUajFw

(3) Tout idéal maximal propre est premier.

Démonstration. En plein d'étapes.

(1) **(1), ⇒** Évacuons le cas trivial pour être sûr. Si $I = \{0\}$ alors A est intègre par le lemme 1.230. Donc $A/I = A/\{0\} = A$ est intègre également.
Soient $a, b \in A$ tels que $[a][b] = [0]$. Donc $[ab] = [0]$, c'est-à-dire $ab \in I$. Puisque I est un idéal premier nous avons $a \in I$ ou $b \in I$, c'est-à-dire $[a] = 0$ ou $[b] = 0$; nous en déduisons que A/I est un anneau intègre.

(2) **(1), ⇐**
Soit $ab \in I$. Alors $[ab] = 0$, ce qui signifie que $[a][b] = 0$ donc que $[a] = 0$ ou que $[b] = 0$ parce que A/I est intègre. Mais la condition $[a] = 0$ signifie $a \in I$, et $[b] = 0$ signifie $b \in I$. Nous avons donc prouvé que soit a soit b est dans I, c'est-à-dire que I est premier.

(3) **(2), ⇒**
Nous devons montrer que tout élément non nul de A/I est inversible. Un élément non nul de A/I est $[x]$ avec $x \in A\backslash I$.
Nous considérons $J = Ax + I$, qui est un idéal parce que pour tout $a \in A$, $aAx + aI \in Ax + I$. Mais comme I est maximal, $J = I$ ou $J = A$.
Si $J = I$, nous aurions que pour tout $a \in A$ et pour tout $i \in I$, $ax + i \in I$. En particulier pour $a = 1$ et $i = 0$ nous aurions $x \in I$, ce qui est contraire à l'hypothèse faite sur x.
Donc $J = A$. En particulier, $1 \in J$, c'est-à-dire qu'il existe $a \in A$ et $i \in I$ tels que $ax + i = 1$. En passant aux classes, $[ax] = 1$, c'est-à-dire $[a][x] = 1$ qui signifie que $[a]$ est un inverse de $[x]$ dans A/I.
Nous avons prouvé que A/I est un corps.

(4) **(2), ⇐** Si $x \in A\backslash I$, il faut prouver que tout idéal contenant I et x est A.
Un idéal contenant I et x doit contenir l'idéal $J = Ax + I$. Comme $x \notin I$, nous avons $[x] \neq 0$ dans A/I. Donc $[x]$ est inversible et il existe $a \in A$ tel que $[ax] = [1]$. C'est-à-dire que $ax - 1 \in I$. Nous avons alors

$$1 = ax + \underbrace{(1 - ax)}_{\in I}. \tag{1.321}$$

C'est-à-dire que $1 \in Ax + I$ et donc $Ax + I = A$.

Enfin nous prouvons que tout idéal maximal propre est premier.

Si I est maximal, A/I est un corps par le point (2), et vu que I est propre, le corps A/I n'est pas réduit à $\{0\}$. Donc le lemme 1.198 dit que A/I est un anneau intègre. Le point (1) dit alors que I est un idéal premier. □

Remarque 1.232.
Puisqu'un corps peut être réduit à $\{0\}$, dans (2), l'idéal peut être A. Mais pas dans (3), parce qu'un idéal premier est propre, ça fait partie de la définition 1.229.

PROPooHABIooBZZQMj

Proposition 1.233 ([50]).
Si A est un anneau commutatif intègre, alors un idéal I dans A est premier si et seulement si A/I est intègre.

81. Tous les anneaux du Frido sont commutatifs
82. Idéal premier, définition 1.229.
83. Idéal maximal, définition 1.219.

Démonstration. Supposons que I soit un idéal premier. Si $[a], [b] \in A/I$ sont tels que $[a][b] = 0$, alors $[ab] = 0$, ce qui signifie que $ab \in I$. Mais alors, puisque I est premier, soit a soit b est dans I. Cela signifie que soit $[a]$ soit $[b]$ est nul dans A/I. Cela prouve que A/I est un anneau intègre.

Dans l'autre sens, nous supposons que A/I est intègre. Cela implique immédiatement que $I \neq A$ parce que A/A n'est pas un anneau intègre (tout le monde est évidemment diviseur de zéro).

Soient donc $a, b \in A$ tels que $ab \in I$. Alors $[a][b] = [ab] = 0$ dans A/I, mais comme A/I est intègre, cela implique que soit $[a]$ soit $[b]$ est nul. Autrement dit, soit a soit b est dans I. □

PROPooSGLNooBYKNyo

Proposition 1.234.
Dans un idéal principal, les conditions suivantes sont équivalentes :

(1) L'élément p est premier.

(2) L'élément p est irréductible.

(3) L'idéal pA est un idéal maximal.

1.10 Anneau intègre

SECAnneauxIntegres

1.10.1 Sous-groupes de $(\mathbb{Z}, +)$

PropSsgpZestnZ

Proposition 1.235 (liste des sous groupes de $\mathbb{Z}$)**.**
À propos de sous-groupes de $\mathbb{Z}$.

(1) Une partie H du groupe $(\mathbb{Z}, +)$ est un sous-groupe si et seulement si il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.

ITEMooOWNZooUsYRok

(2) Si H est un sous-groupe de $(\mathbb{Z}, +)$, il existe un unique n tel que $H = n\mathbb{Z}$.

Démonstration. Soit $H \neq \{0\}$ un sous-groupe de $\mathbb{Z}$. L'ensemble $H \cap \mathbb{N}^*$ contient un élément minimum que nous notons n. Nous avons certainement $n\mathbb{Z} \subset H$ parce que H est un groupe (donc $n + n$ et $-n$ sont dans H dès que n est dans H). Nous devons prouver que $H \subset n\mathbb{Z}$.

Si $x \in H$, par le théorème de division euclidienne 1.222, il existe $q \in \mathbb{Z}$ et $r \in \mathbb{N}$, uniques, tels que $x = nq + r$ et $0 \leqslant r < n$. Nous savons déjà que $nq \in H$, donc $r = x - nq \in H$. Le nombre r est donc un élément de H strictement plus petit que n. Mais nous avions décidé que n serait le plus petit élément de $H \cap \mathbb{N}^*$. Par conséquent $r = 0$ et $x = nq \in n\mathbb{Z}$.

En ce qui concerne l'unicité, supposons que $n\mathbb{Z} = m\mathbb{Z}$. Le nombre n divise m (parce que $m \in m\mathbb{Z} \subset n\mathbb{Z}$) et le nombre m divise n parce que $n \in m\mathbb{Z}$. Par conséquent $n = m$. □

1.10.2 Théorème de Bézout

ThoBuNjam

Théorème 1.236 (Théorème de Bézout[84][51], thème 3)**.**
Deux entiers non nuls $a, b \in \mathbb{Z}^*$ sont premiers entre eux si et seulement si il existe $u, v \in \mathbb{Z}$ tels que

$$au + bv = 1 \tag{1.322}$$

Démonstration. Soit $d = \text{pgcd}(a, b)$ et des nombres u, v tels que $au + bv = 1$. Le PGCD d divise à la fois a et b, et donc divise $au + bv$. Nous en déduisons que d divise 1 et est par conséquent égal à 1.

Nous supposons maintenant que $\text{pgcd}(a, b) = 1$ et nous considérons l'ensemble

$$E = \{au + bv \text{ tel que } u, v \in \mathbb{Z}\} \cap \mathbb{N}^*. \tag{1.323}$$

C'est-à-dire l'ensemble des nombres strictement positifs pouvant s'écrire sous la forme $au + bv$. Cet ensemble est non vide parce qu'il contient par exemple soit a soit $-a$. Soit m le plus petit élément de E et écrivons

$$m = au_1 + bv_1. \tag{1.324}$$

EqMBsfsP

84. Il y a une super application ici : https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/mauvais_prix.pdf.

Par le théorème de division euclidienne[85] (avec a et m), il existe des entiers uniques q et r tels que

$$a = mq + r \tag{1.325}$$

avec $0 \leqslant r < m$. En remplaçant m par sa valeur (1.324), $a = (au_1 + bv_1)q + r$ et

$$r = a(1 - u_1 q) - bv_1 q, \tag{1.326}$$

c'est-à-dire que $r \in \mathbb{Z}a + \mathbb{Z}b$ en même temps que $0 \leqslant r < m$. Si r était strictement positif, il serait dans E. Mais cela est impossible par minimalité de m. Donc $r = 0$ et a est divisible par m.

De la même façon nous prouvons que b est divisible par m. Puisque m divise à la fois a et b nous avons $m = 1$. □

Une généralisation de Bézout 1.236 à plus de 2 variables.

PROPooWSMTooMdfqse

Proposition 1.237.
Si $\{a_i\}_{i=1,\dots,N}$ sont des entiers tels que $\text{pgcd}(a_1, \dots, a_N) = 1$, alors il existe des entiers $\{u_i\}_{i=1,\dots,N}$ tels que

$$\sum_i a_i u_i = 1. \tag{1.327}$$

CorgEMtLj

Corolaire 1.238.
Soient p et q deux entiers premiers entre eux. Alors

$$p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}; \tag{1.328}$$

en particulier, pour tout $x \in \mathbb{Z}$, il existe u_x, v_x entiers tels que $u_x p + v_x q = x$.

Démonstration. Soit $x \in \mathbb{Z}$. Le théorème de Bézout nous donne k et l tels que $kp + lq = 1$. Alors, $(xk)p + (xl)q = x$. □

Notons que l'application $p\mathbb{Z} + q\mathbb{Z}$ vers $\mathbb{Z}$ n'est évidemment pas injective : les u_x et v_x ne sont pas uniques à x fixé, car si $\alpha p + \beta q = x$, alors $(\alpha + kq)p + (\beta - kp)q = x$ pour tout $k \in \mathbb{Z}$.

CORooLINXooBlUKPG

Corolaire 1.239.
Les quotients de $\mathbb{Z}$ sont $\mathbb{Z}/n\mathbb{Z}$.

Démonstration. Tous les idéaux de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$. En effet en vertu de la proposition 1.235, les seuls sous-groupes de $\mathbb{Z}$ (en tant que groupe additif) sont les $n\mathbb{Z}$. Tous les idéaux sont donc de cette forme. De plus les $n\mathbb{Z}$ sont effectivement tous des idéaux[86] : si $a \in n\mathbb{Z}$ et si $k \in \mathbb{Z}$ alors $ak \in n\mathbb{Z}$. □

PropZpintssiprempUzn

Proposition 1.240.
Soient $n \geqslant 2$ un entier et $\phi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ la surjection canonique. Nous noterons $\overline{a} = \phi(a)$. Alors l'ensemble des inversibles de $\mathbb{Z}/n\mathbb{Z}$ est donné par

$$U(\mathbb{Z}/n\mathbb{Z}) = \phi(P_n) = \{\overline{x} \text{ tel que } 0 \leqslant x \leqslant n \text{ tel que } \text{pgcd}(x, n) = 1\}. \tag{1.329}$$

où P_n est l'ensemble $P_n = \{x \in \{0, \dots, n-1\} \text{ tel que } \text{pgcd}(x, n) = 1\}$.

De plus,

$$\text{Card}\left(U(\mathbb{Z}/n\mathbb{Z})\right) = \phi(n). \tag{1.330}$$

Démonstration. Soit $0 \leqslant x \leqslant n$ tel que $\text{pgcd}(x, n) = 1$. Il existe donc[87] $u, v \in \mathbb{Z}$ tels que $ux + vn = 1$. En passant aux classes,

$$\overline{ux} = \overline{1}, \tag{1.331}$$

85. Théorème 1.222.
86. Définition 1.182.
87. Théorème de Bézout 1.236

donc $\overline{u}$ est l'inverse de $\overline{x}$. Cela prouve que $\phi(P_n) \subset U(\mathbb{Z}/n\mathbb{Z})$.

Nous prouvons maintenant l'inclusion inverse. Soient $\overline{x}$ et $\overline{y}$ inverses l'un de l'autre : $\overline{xy} = \overline{1}$. Il existe donc $q \in \mathbb{Z}$ tel que $xy - qn = 1$, ce qui prouve[88] que $\text{pgcd}(x, n) = 1$. □

LEMooZSMEooUmSXWZ

Lemme 1.241.
Un corps[89] *est un anneau intègre.*

Démonstration. En effet, soient un corps $\mathbb{K}$ et deux éléments $x, y \in \mathbb{K}$ tels que $xy = 0$. Si y est inversible, alors nous pouvons multiplier par y^{-1} pour trouver $x = 0$. Cela prouve que $\mathbb{K}$ est un anneau intègre. □

Exemple 1.242.
L'anneau $\mathbb{Z}/6\mathbb{Z}$ n'est pas intègre parce que $3 \cdot 2 = 0$ alors que ni 3 ni 2 ne sont nuls. △

Nous verrons au théorème 3.104 que l'anneau A est intègre si et seulement si $A[X]$ est intègre.

CorZnInternprem

Corolaire 1.243.
L'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre si et seulement si n est premier.

Démonstration. Supposons que n soit premier. La proposition 1.240 donne les inversibles de $\mathbb{Z}/n\mathbb{Z}$ par

$$U(\mathbb{Z}/n\mathbb{Z}) = \{\overline{x} \text{ tel que } 0 \leqslant x \leqslant n \text{ tel que } \text{pgcd}(x, n) = 1\}. \tag{1.332}$$

Mais comme n est premier, $\text{pgcd}(x, n) = 1$ pour tout x, et donc tous les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont inversibles. Donc $\mathbb{Z}/n\mathbb{Z}$ est intègre.

Si n n'est pas premier, alors $n = pq$ avec $1 < p \leqslant q < n$. Alors

$$[p]_n[q]_n = [pq]_n = [0]_n. \tag{1.333}$$

Donc lorsque n n'est pas premier, l'anneau $\mathbb{Z}/n\mathbb{Z}$ possède des diviseurs de zéro et n'est alors pas intègre. □

PropomqcGe

Proposition 1.244 (Thème 18, [1])**.**
Soit A un anneau principal[90] *qui n'est pas un corps. Pour un idéal propre I de A, les conditions suivantes sont équivalentes :*

ITEMooNOVFooEHtcwE

(1) I est un idéal maximal[91] *;*

ITEMooMQWVooNocVEU

(2) I est un idéal premier non nul[92] *;*

ITEMooJBXGooEISNuW

(3) il existe p irréductible[93] *dans A tel que $I = (p)$.*

Démonstration. En plusieurs implications.

(1) **(1) implique (2)** Par hypothèse, I est un idéal propre, de plus il n'est pas égal à $\{0\}$, parce que lorsque A et $\{0\}$ sont les seuls idéaux, nous avons un corps (proposition 1.218). Étant donné que I est un idéal maximal, le quotient A/I est un corps par la proposition 1.220.

Soient maintenant, pour entrer dans le vif du sujet, des éléments $a, b \in A$ tels que $ab \in I$. Dans le corps A/I nous avons $[ab] = 0$, et par définition du produit dans le quotient, $[a][b] = 0$. Par intégrité de l'anneau A/I (un corps est un anneau intègre, lemme 1.241) nous avons soit $[a] = 0$, soit $[b] = 0$, soit les deux en même temps. Dans tous les cas, soit a soit b est dans I.

88. À nouveau avec le Théorème de Bézout.
89. Définition 1.209.
90. Définition 1.228.
91. Définition 1.219.
92. Définition 1.229.
93. Définition 1.189.

(2) **(2) implique (3)** Maintenant I est un idéal premier non réduit à $\{0\}$. Puisque A est un anneau principal, il existe $x \in A$ tel que $I = (x)$. Nous devons prouver que x peut être choisi irréductible ; et nous allons faire plus : nous allons prouver que x ne peut être que irréductible [94].

Supposons que x ne soit pas irréductible. Alors il existe $a, b \in A$ non inversibles tels que $x = ab$. Si $a \in (x)$ alors il existe $k \in A$ tel que $a = xk$, et en particulier, $a = abk$, c'est-à-dire $1 = bk$ (parce que A est principal et donc intègre). Cela signifie que b est inversible alors que nous avions dit qu'il ne l'était pas. Nous en déduisons que a n'est pas dans (x). On montre de manière similaire que b n'est pas dans (x) non plus.

Nous nous retrouvons donc avec $a, b \in A$ tel que $ab \in I$ sans que ni a ni b ne soient dans I. Cela contredit le fait que I soit un idéal premier. En conclusion, x est irréductible.

(3) **(3) implique (1)** Nous avons $I = (p)$ avec p irréductible dans A. Supposons que J est un idéal différent de A contenant I. Comme A est principal, il existe $y \in A$ tel que $J = (y)$. En particulier $p \in J$, donc $p = ay$ pour un certain $a \in A$. Mais p est irréductible, donc soit a est inversible, soit y est inversible. Si y est inversible, alors $J = A$, ce qui est exclu. Si a est inversible, alors $(y) = (p)$, et $I = J$.

□

1.245.
Dans la proposition 1.244, l'hypothèse d'idéal propre est importante. En effet dans le cas $I = A$, nous avons évidemment que I est un idéal maximum. Mais A n'est d'abord pas un idéal premier parce qu'un idéal premier doit être strictement inclus dans l'anneau. Et ensuite, A est en général loin d'être garanti d'être égal à (p) pour un de ses éléments p.

PropoTMMXCx

Proposition 1.246.
Soit A un anneau principal, et soit $p \in A$ un élément irréductible. Alors

(1) (p) est un idéal maximum.

ITEMooKPJQooWuPZXS

(2) $A/(p)$ est un corps.

Démonstration. Nous notons $I = (p)$. Soit un idéal J contenant I. Comme A est principal, J est monogène : $J = (q)$. Mais comme p est dans I qui est dans J, il existe $a \in A$ tel que $p = qa$.

Puisque p est irréductible, soit q, soit a est inversible. Si q est inversible, alors $J = A$. Si a est inversible, alors nous avons $p = qa$, donc $q = pa^{-1}$, ce qui signifie que $q \in (p)$ et donc que $J = I$.

Cela prouve que (p) est un idéal maximum.

Le fait que $A/(p)$ soit un corps est maintenant la proposition 1.220. □

Exemple 1.247.
L'anneau $\mathbb{Z}$ est principal parce qu'il est intègre et que ses seuls idéaux sont les $n\mathbb{Z}$ qui sont principaux : $n\mathbb{Z}$ est engendré par n. △

EXooCJRPooYkWdyr

Exemple 1.248 (Les idéaux de $\mathbb{Z}/n\mathbb{Z}$)**.**
Les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont principaux, mais l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'est pas principal lorsque n n'est pas premier. Nous allons voir ça.

(1) **Les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont principaux** Soit un idéal S dans $\mathbb{Z}/n\mathbb{Z}$. Nous considérons la projection canonique $\phi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. La proposition 1.217 dit que $S = \phi(J)$ où J est un idéal de $\mathbb{Z}$ contenant $n\mathbb{Z}$. Mais le corolaire 1.239 nous dit qu'alors $J = m\mathbb{Z}$ pour un certain m. Pour que $m\mathbb{Z}$ contienne $n\mathbb{Z}$, il faut que m divise n.

Bref, $S = \phi(m\mathbb{Z})$ avec $m \mid n$. Nous montrons maintenant que S est engendré par $[m]_n$. D'abord, l'élément $[m]_n$ est bien dans $\phi(m\mathbb{Z})$. Ensuite un élément de $\phi(m\mathbb{Z})$ est de la forme

$$[km]_n = k[m]_n \in ([m]_n). \tag{1.334}$$

94. ça me semble un peu trop facile. Lisez attentivement, et écrivez-moi pour dire si vous êtes d'accord ou pas.

Donc $S \subset ([m]_n)$. Et l'inclusion dans l'autre sens est tout aussi immédiate : un élément de $([m]_n)$ est de la forme

$$k[m]_n = [km]_n = \phi(km) \in \phi(m\mathbb{Z}). \tag{1.335}$$

(2) **Si n n'est pas premier, $\mathbb{Z}/n\mathbb{Z}$ n'est pas principal** Le fait est que lorsque n n'est pas premier, $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre (corolaire 1.243).

(3) **Moralité** Un anneau comme $\mathbb{Z}/6\mathbb{Z}$ est un anneau dont tous les idéaux sont principaux, mais qui n'est pas principal.

△

Exemple 1.249.
Nous verrons dans la proposition 26.68 que l'anneau des fonctions holomorphes sur un compact de $\mathbb{C}$ est principal. △

ThofPXwiM

Théorème 1.250.
Si A est un anneau principal et si p et q sont premiers entre eux dans A, alors on a l'isomorphisme d'anneaux

$$A/pqA \simeq A/pA \times A/qA. \tag{1.336}$$

1.11 Anneau euclidien

DefAXitWRL

Définition 1.251 (Wikipédia)**.**
Soit A un anneau intègre[95]*. Un **stathme euclidien** sur A est une application $\alpha : A\backslash\{0\} \to \mathbb{N}$ tel que*

ITEMooLVJAooLpjgEz

(1) $\forall a, b \in A\backslash\{0\}$, il existe $q, r \in A$ tel que

$$a = qb + r \tag{1.337}$$

et $\alpha(r) < \alpha(b)$.

(2) Pour tout $a, b \in A\backslash\{0\}$, $\alpha(b) \leqslant \alpha(ab)$.

*Un anneau est **euclidien** si il accepte un stathme euclidien.*

Le stathme est la fonction qui donne le « degré » à utiliser dans la division euclidienne. La contrainte est que le degré du reste soit plus petit que le degré du dividende.

LEMooFUSTooDCcBDb

Lemme 1.252.
L'ensemble $\mathbb{Z}$ avec les opérations usuelles est un anneau intègre[96]*.*

ExwqlCwvV

Exemple 1.253.
Le stathme de $\mathbb{N}$ pour la division euclidienne usuelle est $\alpha(n) = n$. Si $a, b \in \mathbb{N}$ nous écrivons

$$a = qb + r \tag{1.338}$$

où q est l'entier le plus proche *inférieur* à a/b (on veut que le reste soit positif) et $r = a - qb$. Nous avons donc

$$r - b = a - b(q + 1) < a - b\frac{a}{b} = 0, \tag{1.339}$$

ce qui montre que $r < b$. △

Cet exemple ne fonctionne pas avec $\mathbb{Z}$ au lieu de $\mathbb{N}$ parce que le stathme doit avoir des valeurs dans $\mathbb{N}$. Cela ne veut cependant pas dire qu'il n'existe pas de stathme sur $\mathbb{Z}$; cela veut seulement dire que $\alpha(x) = x$ ne fonctionne pas.

95. Défnition 1.197.
96. Anneau intègre, définition 1.197.

Propkllxnv

Proposition 1.254 ([52]).
Tout anneau euclidien[97] *est principal*[98].

Démonstration. Soit A un anneau euclidien et α un stathme sur A. Nous considérons un idéal I non nul de A. Nous devons montrer que I est généré par un élément. En l'occurrence nous allons montrer qu'un élément $a \in I\backslash\{0\}$ qui minimise $\alpha(a)$ va générer[99] I.

Soit $x \in I$. Par construction, il existe $q, r \in A$ tels que $x = aq + r$ avec $r = 0$ ou $\alpha(r) < \alpha(a)$. Étant donné que $x, a \in I$, $r \in I$. Si $r \neq 0$, alors r contredirait la minimalité de $\alpha(a)$. Donc $r = 0$ et $x = aq$, ce qui signifie que I est principal. □

PROPooPJGLooQSrJTU

Proposition 1.255.
L'anneau $\mathbb{Z}$ est principal et euclidien.

Démonstration. Nous allons seulement montrer que $\alpha(x) = |x|$ est un stathme euclidien. Ainsi $\mathbb{Z}$ sera euclidien et donc principal par la proposition 1.254.

D'abord $\mathbb{Z}$ est intègre, c'est le lemme 1.252.

La condition $\alpha(b) \leqslant \alpha(ab)$ est immédiate : $|a| \leqslant |ab|$ pour tout $a, b \in \mathbb{Z}$.

Soient maintenant $a, b \in \mathbb{Z}$. Nous définissons $q_0, r_0 \in \mathbb{N}$ tels que

$$|a| = q_0|b| + r_0 \tag{1.340}$$

avec $r_0 < |b|$. Cela existe parce que $\alpha(x) = x$ est un stathme sur $\mathbb{N}$ par l'exemple 1.253.

(1) **Si $a > 0$ et $b > 0$**
Alors $a = q_0 b + r_0$ et le couple (q_0, r_0) vérifie les conditions de la définition 1.251(1).

(2) **Si $a > 0$ et $b < 0$**
Alors $a = -q_0 b + r_0$, et le couple $(-q_0, r_0)$ vérifie les conditions de la définition 1.251(1).

(3) **Si $a < 0$ et $b > 0$** Alors $a = -q_0 b - r_0$, et le couple $(-q_0, -r_0)$ vérifie les conditions de la définition 1.251(1) parce que

$$\alpha(-r_0) = r_0 < |b| = \alpha(b). \tag{1.341}$$

(4) **Si $a < 0$ et $b < 0$** Alors $a = q_0 b - r_0$, et le couple $(q_0, -r_0)$ vérifie les conditions de la définition 1.251(1).

□

Nous venons de voir que $\mathbb{Z}$ est principal ; le lemme suivant nous dit que $\mathbb{Z}[X]$ n'est lui, pas principal.

LEMooDJSUooJWyxCL

Lemme 1.256 ([53]).
Si A est un anneau intègre[100] *qui n'est pas un corps, alors $A[X]$ n'est pas principal.*

Démonstration. Soit un élément non nul $a \in A$.

(1) **Un idéal principal contenant a et X est $\mathbf{A[X]}$**
Soit (P) un idéal principal contenant a et X. Puisque $a \in (P)$, il existe Q tel que $a = QP$. Donc P divise a dans $\mathbb{Z}[X]$. L'égalité des degrés indique que P est un polynôme constant, c'est-à-dire en réalité un élément de A. Soit $P = k \in A$.
Comme P divise X, nous avons aussi $X = kQ$ pour un certain $Q \in \mathbb{Z}[X]$. L'égalité des degrés dit qu'il existe $k' \in A$ tel que $Q = k'X$ et donc $Q = k'X = k'kQ$, ce qui implique que $kk' = 1$. L'idéal engendré par k contient donc en particulier $kk' = 1$ et donc contient $A[X]$ en entier :

$$1 = k'k \in k'(P) = (P). \tag{1.342}$$

97. Euclidien, définition 1.251.
98. Principal, définition 1.227
99. Un tel élément existe...
100. Définition 1.197.

(2) **Si $(a, X) = A[X]$ alors a est inversible**
Si $(a, X) = A[X]$, en particulier, $1 \in (a, X)$, ce qui signifie qu'il existe des polynômes $U, V \in A[X]$ tels que
$$1 = UX + Va. \tag{1.343}$$
Nous évaluons cette égalité en 0 : comme $(UX)(0) = 0$, nous avons $1 = V(0)a$, ce qui signifie que $V(0)$ est un inverse de A. Donc a est inversible.

(3) **Si a n'est pas inversible alors (a, X) n'est pas principal**
Si (a, X) était principal, alors nous aurions, par ce qui est dit plus haut, $(a, X) = A[X]$. Mais cette dernière égalité impliquerait que a soit inversible.

En conclusion, si A n'est pas un corps, il possède un élément ni nul ni inversible. Dans ce cas, l'idéal (a, X) n'est pas principal dans $A[X]$ et nous en déduisons que $A[X]$ n'est pas un anneau principal. □

Nous verrons dans le lemme 3.110 que si $\mathbb{K}$ est un corps, alors $\mathbb{K}[X]$ est principal.

1.12 Le groupe et anneau des entiers

Certes $(\mathbb{Z}, +)$ est un groupe mais en ajoutant la multiplication, $(\mathbb{Z}, +, \times)$ devient un anneau [101].

1.12.1 PGCD, PPCM et Bézout

Puisque $\mathbb{Z}$ est un anneau intègre, nous avons la définition 1.186 de pgcd et de ppcm.

PROPooAVRGooUfhjwF

Proposition 1.257 (PPCM et PGCD)**.**
Soient $p, q \in \mathbb{Z}^*$.

(1) Le pgcd de p et q est le plus grand diviseur commun de p et q.

(2) Le ppcm de p et q est leur plus petit multiple commun.

Démonstration. Démontrons le premier point. Notons δ le pgcd de p et q. Si d est un diviseur commun de p et q, alors d divise δ. Dans $\mathbb{Z}$, $d \mid \delta$ implique $d \leqslant \delta$ (proposition 1.113). □

Lemme 1.258.
Soient $p, q \in \mathbb{Z}^*$. Les entiers $\text{ppcm}(p, q)$ *et* $\text{pgcd}(p, q)$ *fournissent les isomorphismes de groupes suivants :*

$$p\mathbb{Z} \cap q\mathbb{Z} = \text{ppcm}(p, q)\mathbb{Z} \tag{1.344a}$$
$$p\mathbb{Z} + q\mathbb{Z} = \text{pgcd}(p, q)\mathbb{Z}. \tag{1.344b}$$

DEFooXSPFooPumQSy

Définition 1.259.
Nous disons que deux éléments d'un anneau principal [102] sont ***premiers entre eux*** *si leurs diviseurs communs sont inversibles.*

Vu que $\mathbb{Z}$ est un anneau principal (proposition 1.255), la définition 1.259 d'éléments premiers entre eux s'applique.

LEMooLKLTooXUdsgB

Lemme 1.260.
Dans $\mathbb{Z}$, les nombres p et q sont premiers entre eux si et seulement si $\text{pgcd}(p, q) = 1$.

DefZHRXooNeWIcB

Définition 1.261.
Si nous avons un ensemble d'entiers a_i, nous disons qu'ils sont premiers ***dans leur ensemble*** *si 1 est le PGCD de tous les a_i ensemble.*

101. Définition 1.42.
102. Anneau principal, définition 1.228.

Les nombres 2, 4 et 7 ne sont pas premiers deux à deux (à cause de 2 et 4), mais ils sont premiers dans leur ensemble parce qu'il n'y a pas de diviseurs communs plus grand que 1, au triplet $(2,4,7)$.

La proposition suivante établit que si x est assez grand, alors il peut même être écrit comme une combinaison de p et q à coefficients positifs. Elle sera utilisée pour démontrer que les états apériodiques d'une chaine de Markov peuvent être atteints à tout moment (assez grand), voir la définition 38.33 et ce qui suit.

PropLAbRSE

Proposition 1.262.
Soient a et b deux éléments de $\mathbb{N}$ premiers entre eux. Il existe $N>0$ tel que tout $x>N$ appartient à $a\mathbb{N}+b\mathbb{N}$.

Démonstration. Soient a et b, premiers entre eux, et $x\in\mathbb{N}$. Disons tout de suite, pour éviter les cas triviaux et pénibles, que x, a et b sont strictement positifs.

(1) **Une décomposition pour x**
On applique le théorème 1.222 de division euclidienne à x et $a+b$: il existe des entiers p_x, r_x, uniques, tels que

$$\begin{cases} x=(p_x-1)(a+b)+r_x & \text{(1.345a)}\\ 0\leqslant r_x<a+b. & \text{(1.345b)}\end{cases}$$

En d'autres termes, $p_x(a+b)$ est le premier multiple de $a+b$ supérieur ou égal à x. De plus, p_x est strictement positif car x l'est. Il existe alors des entiers u et v tels que

$$ua+vb=p_x(a+b)-x \tag{1.346}$$ EQooXYSZooJqxPui

par le corolaire 1.238. Ainsi, x peut s'écrire

$$x=(p_x-u)a+(p_x-v)b. \tag{1.347}$$

(2) **Des maximums**
Il s'agit maintenant de savoir si nous pouvons être assuré d'avoir $p_x>u$ et $p_x>v$ dès que x est assez grand. Pour cela, grâce au corolaire 1.238, nous considérons les nombres u_i et v_i définis par

$$u_ia+v_ib=i \tag{1.348}$$

pour $i=1,\ldots,a+b$. Nous posons $u^*=\max\{u_i\}$, $v^*=\max\{v_i\}$, et $p^*=\max\{u^*,v^*\}$. Nous posons alors $N=p^*(a+b)$, et considérons $x>N$.

(3) **Nouvelle décomposition pour x**
Nous voulons écrire

$$x=(p_x-u_k)a+(p_x-v_k)b \tag{1.349}$$ EQooIKNWooBKItYz

pour un certain k. Cela demande $u_ka+v_kb=ua+vb=p_x(a+b)-x$ par l'équation (1.346). Vu que $p_x(a+b)-x>0$, les nombres u_k et v_k existent : il suffit de prendre $k=p_x(a+b)-x$.

(4) **Conclusion**
Avec tous ces choix, nous avons d'abord $x>p^*(a+b)$ et donc

$$x=(p_x-1)(a+b)+r_x>p^*(a+b), \tag{1.350}$$

ce qui donne

$$(p_x-1)(a+b)>p^*(a+b)-r_x>(p^*-1)(a+b). \tag{1.351}$$

ou encore $p_x>p^*$. Nous avons finalement

$$p_x\geqslant p^*\geqslant u^*\geqslant u_k \tag{1.352}$$

et

$$p_x\geqslant p^*\geqslant v^*\geqslant v_k. \tag{1.353}$$

De ce fait, la décomposition (1.349) est celle que nous voulions.

□

1.263.
Une méthode pour obtenir les entiers naturels u et v qui permettent la décomposition $x = au + bv$ est d'abord de choisir u_0 et v_0 tels que au_0 et bv_0 soient les plus proches possibles de $x/2$, puis de décomposer le nombre (relativement petit) $x - au_0 - bv_0$ en $au_1 + bv_1$. Deux nombres u et v qui fonctionnent sont alors $u = u_0 + u_1$ et $v = v_0 + v_1$.

Exemple 1.264.
Écrivons $1000 = u \cdot 7 + v \cdot 5$ avec $u, v \in \mathbb{N}$. D'abord $72 \cdot 7 = 504$ et $100 \cdot 5 = 500$. Nous avons donc

$$1004 = 72 \cdot 7 + 100 \cdot 5. \tag{1.354}$$

Ensuite $4 = 25 - 21 = -3 \cdot 7 + 5 \cdot 5$. Au final,

$$1000 = 75 \cdot 7 + 95 \cdot 5. \tag{1.355}$$

△

1.13 Sous-groupe normal

propGroupeNormal

Proposition 1.265.
Soit N un sous-groupe de G. Les propriétés suivantes sont équivalentes :

ITEMooDYEUooOuKEqQ

(1) N est normal[103] *dans G.*

ITEMooPYTEooZhvrUa

(2) N est une union de classes de conjugaison[104] *de G,*

ITEMooJWTLooBRmriQ

(3) $gNg^{-1} \subseteq N$ pour tout $g \in G$,

ITEMooVRZIooAorhRY

(4) $gNg^{-1} = N$ pour tout $g \in G$,

ITEMooJGUOooYshOZa

(5) $gN = Ng$ pour tout $g \in G$,

ITEMooMRYRooZifCCe

(6) Le normalisateur[105] *de N est G : $\mathcal{N}_G(N) = G$.*

Démonstration. En plusieurs parties.

(1) **(1) implique (3)** C'est la définition de sous-groupe normal.

(2) **(3) implique (4)** Soit $g \in G$. Nous avons $gNg^{-1} \subset N$, mais aussi (en appliquant l'hypothèse à g^{-1}) $g^{-1}Ng \subset N$. En combinant nous avons

$$N \subset g(g^{-1}Ng)g^{-1} \subset gNg^{-1}. \tag{1.356}$$

Nous avons l'inclusion dans les deux sens. Donc l'égalité.

(3) **(4) implique (5)** Soit $g \in G$. Un élément général de gN est de la forme gn avec $n \in N$. Nous devons trouver un $n' \in N$ tel que $gn = n'g$. En posant $n' = gng^{-1}$ nous avons

$$n' = gng^{-1} \in gNg^{-1} \subset N. \tag{1.357}$$

Il est immédiat de prouver que $gn = n'g$. Cela prouve que $gN \subset Ng$.
Le même raisonnement donne $Ng \subset gN$.

(4) **(5) implique (3)** Un élément de gNg^{-1} est $a = gng^{-1}$ avec $n \in N$. Nous devons prouver que $a \in N$. Puisque $gn \in gN$, par hypothèse il existe n' tel que $gn = n'g$. En remplaçant dans la définition de a,

$$a = gng^{-1} = n'gg^{-1} = n' \in N. \tag{1.358}$$

103. Définition 1.172.
104. Définition 1.168.
105. Définition 1.171.

(5) **(5) implique (2)** Pour chaque $a \in G$ nous notons C_a la classe de conjugaison de a dans G :

$$C_a = \{gag^{-1} \text{ tel que } g \in G\}. \tag{1.359}$$

EQooJIEVooCAshfe

Comme $a \in C_a$ (prendre $g = e$ dans (1.359).) nous avons forcément

$$N \subset \bigcup_{n \in N} C_n. \tag{1.360}$$

Prouvons maintenant l'inclusion inverse. Nous avons déjà prouvé que (5) implique (3). Donc si $n \in N$, alors $gng^{-1} \in N$. Nous avons alors

$$C_n = \{gng^{-1} \text{ tel que } g \in G\} \subset N. \tag{1.361}$$

Donc il est vrai que $N = \bigcup_{n \in N} C_n$.

(6) **(2) implique (1)** Nous supposons que $N \subset G$ est un sous-groupe de la forme

$$N = \bigcup_{a \in I} C_a \tag{1.362}$$

où I est une partie de G. Nous devons montrer que pour tout $g \in G$ et pour tout $n \in N$ nous avons $gng^{-1} \in N$. Puisque $n \in N$, il existe $a \in I$ tel que $n \in C_a$ et donc il existe $k \in G$ tel que $n = kak^{-1}$. Nous avons donc

$$gng^{-1} = g(kak^{-1})g^{-1} = (gk)a(gk)^{-1} \in C_a \subset N. \tag{1.363}$$

La lectrice attentive aura remarqué l'utilisation de l'axiome du choix. La prudence l'incitera à ne pas le faire remarquer au jury.

(7) **(6) si et seulement si (5)** C'est la définition du normalisateur.

□

Définition 1.266.
Soit $g \in G$ et $n \in \mathbb{Z}$. Nous définissons g^n par

(1) $g^0 = e$ et $g^n = gg^{n-1}$ si n est positif.

(2) si $n < 0$, nous posons $g^n = (g^{-1})^{-n}$.

L'ordre d'un groupe et l'ordre d'un élément d'un groupe sont deux choses différentes.

DEFooKWBCooMlmpCP

Définition 1.267 (Ordre d'un groupe)**.**
Soit un groupe G.

*(1) Si G est un ensemble fini, l'**ordre** de G est son cardinal*[106]*, et nous le notons $|G|$.*

(2) Si l'ensemble G est infini, nous disons que $|G| = \infty$ et qu'il est d'ordre infini.

Oui : nous pourrions simplement toujours dire « cardinalité » et écrire $\text{Card}(G)$. Au lieu de ça, dans le cas particulier des groupes, il y a une tradition de dire « ordre » et d'écrire $|G|$.

DEFooKSTVooOObpgC

Définition 1.268 (Ordre d'un élément)**.**
*L'**ordre** d'un élément g de G est le naturel*

$$\min\{n \in \mathbb{N}\backslash\{0\} \text{ tel que } g^n = e\}, \tag{1.364}$$

si il existe ; dans le cas contraire, nous disons que l'ordre de g est infini.

1.269.
Nous verrons que le corolaire 2.14 au théorème de Lagrange dira que l'ordre d'un élément divise l'ordre du groupe.

106. Définition 1.126.

LemHUkMxp

Lemme 1.270 ([54, 13]).
Soient un groupe G et deux sous-groupes normaux[107] *H et K tels que $H \cap K = \{e\}$. Alors :*

ITEMooDFVBooSnnlgR

(1) Tout élément de H commute avec tout élément de K.

ITEMooVVBGooZSJqjp

(2) HK est un sous-groupe de G.

IMTEooPCBZooQoZFOD

(3) L'application

$$\begin{aligned} \varphi\colon H \times K &\to HK \\ (h,k) &\mapsto hk \end{aligned} \tag{1.365}$$

est un isomorphisme de groupes.

Démonstration. Point par point.

(1) **(1)** Soient $h \in H$ et $k \in K$. Nous voulons montrer que $hk = kh$. Pour cela nous considérons l'élément $a = hkh^{-1}k^{-1}$. Comme H est normal dans G, nous avons

$$kh^{-1}k^{-1} \in H \tag{1.366}$$

et donc $a \in H$. De même K étant normal dans G, nous avons $hkh^{-1} \in K$ et donc $a \in K$. Au final $a \in H \cap K = \{e\}$. Nous avons prouvé que

$$hkh^{-1}k^{-1} = e, \tag{1.367}$$

et donc que $hk = kh$.

(2) **(2)** Puisque H et K sont des sous-groupes, $\{e\}$ est dans les deux, de telle sorte que $e \in HK$. De plus si $h_i \in H$ et $k_i \in K$, la commutativité du point (1) donne

$$(h_1k_1)(h_2k_2) = h_1h_2k_1k_2 \in HK. \tag{1.368}$$

Donc le produit de deux éléments de HK est dans HK.

(3) **(3)** En trois sous-parties.

(3a) **Morphisme** Soient $h_i \in H$ et $k_i \in K$. En utilisant la commutativité du point (1) nous avons

$$\begin{aligned} \varphi\big((h_1,k_1)(h_2,k_2)\big) &= \varphi(h_1h_2, k_1k_2) && \text{(1.369a)}\\ &= (h_1h_2)(k_1k_2) && \text{(1.369b)}\\ &= (h_1k_1)(h_2k_2) && \text{(1.369c)}\\ &= \varphi(h_1,k_1)\varphi(h_2,k_2). && \text{(1.369d)} \end{aligned}$$

(3b) **Injectif** Si $\varphi(h_1,k_1) = \varphi(h_2,k_2)$ nous avons successivement

$$\begin{aligned} h_1k_1 &= h_2k_2 && \text{(1.370a)}\\ h_1k_1h_2^{-1} &= k_2 && \text{(1.370b)}\\ h_1k_1h_2^{-1}k_1^{-1} &= k_2k_1^{-1} && \text{(1.370c)}\\ h_1h_2^{-1} &= k_2k_1^{-1}. && \text{(1.370d)} \end{aligned}$$

Le membre de gauche est un élément de H et le membre de droite un élément de K. Comme $H \cap K = \{e\}$ nous avons $h_1h_2^{-1} = e$ et $k_2k_1^{-1} = e$, c'est-à-dire $h_1 = h_2$ et $k_1 = k_2$.

(3c) **Surjectif** Un élément général de HK est hk avec $h \in H$ et $k \in K$, c'est-à-dire $\varphi(h,k)$.

□

107. Sous-groupe normal, définition 1.172.

DefvtSAyb

Définition 1.271.
*L'**exposant** du groupe G est le plus petit entier non nul n tel que $g^n = e$ pour tout $g \in G$. S'il n'existe pas un tel n, nous disons que l'exposant du groupe est infini.*

PROPooSWHHooOzqWkw

Proposition 1.272.
À propos d'exposant de groupe et de ppcm.

ITEMooFIPBooFOaPkU

(1) *Si l'ensemble des ordres de tous les éléments d'un groupe est majoré, alors l'exposant du groupe est le plus petit commun multiple des ordres des éléments du groupe.*
ITEMooLIREooYEtcdL

(2) *Pour un groupe fini, l'exposant est le* ppcm *des ordres des éléments du groupe.*

Démonstration. En deux parties.

(1) **Pour (1)** Soit p le ppcm des ordres de tous les éléments du groupe. Si g est d'ordre a, il existe $k \in \mathbb{N}$ tel que $p = ak$. Avec ça nous avons $g^p = (g^a)^k = e^k = e$. Donc p a bien la propriété $g^p = e$ pour tout $g \in G$.
Si $q < p$, nous montrons que q ne peut pas avoir cette propriété. Il existe un élément g dont l'ordre a ne divise pas q. Par la division euclidienne 1.222, nous avons des entiers u et v tels que $q = ua + v$ avec $v < a$. Nous avons alors

$$g^q = g^{ua}g^v = g^v \neq e \tag{1.371}$$

parce que $v < a$ et que a est le plus petit n tel que $g^n = e$.

(2) **Pour (2)** Même chose. Pour un groupe fini, le ppcm existe toujours.

□

<++>
Le théorème de Burnside 8.313 nous donnera un bon paquet d'exemples de groupes d'exposant fini dans $\mathrm{GL}(n, \mathbb{C})$.

PropSRMJooIDPBoW

Proposition 1.273.
Soit un groupe G. Nous considérons un sous-groupe normal H de G ainsi qu'un morphisme $\psi\colon G \to H$. Alors

(1) *$\psi(H)$ est normal dans $\psi(G)$*

(2) *Si G/H est abélien alors $\psi(G)/\psi(H)$ est abélien.*

Démonstration. Soient $h \in H$ et $g \in G$. Alors $\psi(g)\psi(h)\psi(g)^{-1} = \psi(ghg^{-1}) \in \psi(H)$. Donc $\psi(H)$ est normal dans $\psi(G)$.

Pour la seconde partie nous notons $[\dots]$ les classes par rapport à $\psi(H)$ et $\overline{\dots}$ celles par rapport à H. Nous avons

$$\begin{aligned}
[\psi(g_1)][\psi(g_2)] &= \big[\psi(g_1)\psi(g_2)\big] && \text{(1.372a)}\\
&= \big[\psi(g_1g_2)\big] && \text{(1.372b)}\\
&= \{\psi(g_1g_2)\psi(h) \text{ tel que } h \in H\} && \text{(1.372c)}\\
&= \{\psi(g_1g_2h) \text{ tel que } h \in H\} && \text{(1.372d)}\\
&= \psi\Big(\{g_1g_2h \text{ tel que } h \in H\}\Big) && \text{(1.372e)}\\
&= \psi\big(\overline{g_1g_2}\big) && \text{(1.372f)}\\
&= \psi\big(\overline{g_2g_1}\big) && \text{(1.372g)}\\
&= \text{refaire à l'envers} && \text{(1.372h)}\\
&= [\psi(g_2)][\psi(g_1)]. && \text{(1.372i)}
\end{aligned}$$

Par conséquent $\psi(G)/\psi(H)$ est abélien. □

1.13.1 Permutations, groupe symétrique

Nous donnons ici quelques éléments à propos du groupe symétrique. Beaucoup de choses supplémentaires sont reportées à la section 5.6. Voir aussi le thème 8.

DEFooJNPIooMuzIXd

Définition 1.274.
*Soit un ensemble E. Une **permutation** de l'ensemble E est une bijection $E \to E$. Le **groupe symétrique** de E est le groupe des bijections $E \to E$; il est noté S_E.*

*Le **groupe symétrique** S_n est le groupe des permutations de l'ensemble $\{1, \dots, n\}$. C'est donc l'ensemble des bijections $\{1, \dots, n\} \to \{1, \dots, n\}$.*

DEFooSupportPermutation

Définition 1.275.
*Le **support** d'une permutation σ est l'ensemble constitué des éléments modifiés par σ :*

$$\operatorname{supp}\sigma = \{i \in \{1, \dots, n\} \text{ tel que } \sigma(i) \neq i\}.$$

DEFooMVFKooMpXMQy

Définition 1.276 ([54]).
*Soient une permutation $\sigma \in E$ ainsi que $a \in E$. La σ-**orbite** de a est l'ensemble*

$$\Omega_\sigma(a) = \{\sigma^i(a)\}_{i \in \mathbb{N}}. \tag{1.373}$$

LEMooSGWKooKFIDyT

Lemme 1.277 ([55]).
Le groupe symétrique S_n est un ensemble fini contenant $n!$ éléments :

$$\operatorname{Card}(S_n) = n!. \tag{1.374}$$

LEMooUPBOooWbwMTx

Lemme 1.278 ([56]).
Deux résultats.

1. *Tout groupe est isomorphe à un sous-groupe d'un groupe symétrique.*
2. *Tout groupe fini d'ordre n est isomorphe à un sous-groupe de S_n.*

Démonstration. Soit, pour $g \in G$ donné, l'application

$$\begin{aligned} \tau_g \colon G &\to G \\ x &\mapsto gx. \end{aligned} \tag{1.375}$$

Commençons par prouver que cela est une bijection. D'une part, $\tau_g(x) = y$ pour $x = g^{-1}y$ (surjection) et, d'autre part, $\tau_g(x) = \tau_g(y)$ implique $gx = gy$ et donc $x = y$ (injection).

Nous avons donc $\tau_g \in S_G$. L'application

$$\begin{aligned} \varphi \colon G &\to S_G \\ g &\mapsto \tau_g \end{aligned} \tag{1.376}$$

est un morphisme de groupe [108] :

$$\varphi(gh)x = ghx = \varphi(g)(hx) = \varphi(g)\Big(\varphi(h)x\Big) = \Big(\varphi(g) \circ \varphi(h)\Big)x. \tag{1.377}$$

Elle est injective parce que si $\tau_g = \tau_h$ alors $gx = hx$ pour tout x. En particulier $g = h$. Donc $\varphi \colon G \to \operatorname{Image}(\varphi)$ est un isomorphisme entre G et un sous-groupe de S_G (proposition 1.39).

Un groupe fini de cardinal n est isomorphe à un sous-groupe de S_G ; or S_G est isomorphe à un des S_n. □

108. Définition 1.38.

1.13.2 Décomposition en cycles

Définition 1.279 (cycle[54]).
Soit E un ensemble de cardinal[109] *n. Soit un entier $1 \leqslant k \leqslant n$. Un élément $\sigma \in S_E$ est un k-**cycle** si il ne possède qu'une seule orbite*[110] *non réduite à un élément et qu'elle est de cardinal k.*

LEMooADNGooDZpdTb

Lemme 1.280 ([1]).
Soient un k-cycle σ et $a \in \Omega$. Alors

$$\Omega_\sigma(a) = \{a, \sigma(a), \dots, \sigma^{k-1}(a)\} \tag{1.378}$$

et $\sigma^k(a) = a$.

En particulier, les éléments $\sigma^q(a)$ avec $q = 0, \dots, k-1$ sont tous distincts.

Démonstration. Soit l le plus grand entier tel que les $\sigma^i(a)$ avec $0 \leqslant i \leqslant l$ soient tous distincts, et notons $A = \{\sigma^i(a)\}_{i=0,\dots,l}$. Cet ensemble satisfait

— $\text{Card}(A) = l + 1$

— $A \subset \Omega_\sigma(a)$, et donc $\text{Card}(A) \leqslant \text{Card}(\Omega_\sigma(a)) = k$ par le lemme 1.128(3).

Que vaut $\sigma^{l+1}(a)$? Par maximalité de l, $\sigma^{l+1}(a)$ est un des $\sigma^i(a)$ avec $i \leqslant l$. Par injectivité de σ, nous avons donc forcément $\sigma^{l+1}(a) = a$.

Donc pour tout $i > l$ il existe $j \leqslant l$ tel que $\sigma^i(a) = \sigma^j(a)$ (parce que $\sigma^i(a) = \sigma^{i-l-1}(a)$). Nous en déduisons que

$$\Omega_\sigma(a) = \{\sigma^i(a)\}_{0 \leqslant i \leqslant l} = A. \tag{1.379}$$

Le cardinal de $\Omega_\sigma(a)$ étant k par hypothèse nous avons $k = l + 1$, et donc $l = k - 1$. □

LEMooANVHooOQiTwY

Lemme 1.281.
Si σ est une cycle de longueur k, et si $b \in \Omega_\sigma(a)$, alors

$$\Omega_\sigma(a) = \Omega_\sigma(b) = \{\sigma^i(b)\}_{i=0,\dots,k-1}. \tag{1.380}$$

Démonstration. Comme $b \in \Omega_\sigma(a)$, il existe $l \leqslant k-1$ tel que $b = \sigma^l(a)$. Pour tout i nous avons $\sigma^i(a) = \sigma^{k-l+1}(b)$, et donc $\Omega_\sigma(a) \subset \Omega_\sigma(b)$.

Mais pour tout i nous avons aussi $\sigma^i(b) = \sigma^{l+1}(a)$ et donc $\Omega_\sigma(b) \subset \Omega_\sigma(a)$.

Nous avons donc montré que $\Omega_\sigma(a) = \Omega_\sigma(b)$. La seconde égalité est le lemme 1.280 appliqué à b. □

LEMooMIHGooQfALbc

Lemme 1.282 ([1]).
Soient un ensemble fini E, une permutation $\sigma \in S_E$ ainsi que $a \in E$. Si $b \in \Omega_\sigma(a)$, alors $\Omega_\sigma(b) = \Omega_\sigma(a)$.

Lemme 1.283 ([54]).
Tout k-cycle est d'ordre[111] *k.*

Démonstration. Soit le cycle $\{a, \sigma(a), \dots, \sigma^{k-1}(a)\}$. Tous les $\sigma^i(a)$ avec $i \leqslant k-1$ sont distincts et $\sigma^k(a) = a$. Donc σ^k est l'identité, et l'ordre de σ est plus petit ou égal à k.

Si $i \leqslant k-1$, alors $\sigma^i(a) \neq a$ parce que les éléments du cycle sont distincts. Donc $\sigma^i \neq \text{Id}$ pour $i \leqslant k-1$. Nous en déduisons que l'ordre de σ est k. □

LEMooQLSAooBrXDXw

Lemme 1.284 ([54, 57]).
Tout élément du groupe symétrique S_n peut être décomposé en un nombre fini de cycles de supports disjoints.

Cette décomposition est unique à l'ordre près de l'écriture des cycles.

Plus précisément, si σ est une permutation, alors il existe un unique ensemble fini $\{\omega_i\}_{i \in I}$ de cycles de supports disjoints tels que[112] *$\sigma = \prod_{i \in I} \omega_i$.*

109. Définition 1.126.
110. Définition 1.276.
111. Définition 1.268.
112. Ici I est un ensemble fini et vu que les supports sont disjoints, le produit est commutatif.

Démonstration. Soit $\sigma \in S_E$. Si les éléments $\{a, \sigma(a), \dots \sigma^k(a)\}$ sont distincts, alors soit $\sigma^{k+1}(a)$ est distincts des autres, soit $\sigma^{k+1}(a) = a$. Il n'est en effet pas possible d'avoir $\sigma^{k+1}(a) = \sigma^l(a)$ avec $l < k$ parce que ça contredirait l'injectivité de σ.

Soit donc $a \in E$. Nous considérons le cycle $\big(a, \sigma(a), \dots, \sigma^k(a)\big)$ où k est maximum tel que tous les éléments sont distincts.

Soit ce cycle contient tous les éléments de E, soit il existe un élément b hors de ce cycle. Dans le second cas, nous considérons le cycle commençant par b.

Et ça continue... □

LEMooWXXLooIzrwJT

Lemme 1.285 ([54]).
Deux cycles de support disjoint commutent.

LEMooVVPWooMkRjyR

Lemme 1.286 ([58]).
Tout cycle de longueur r est le produit de $r - 1$ transpositions.

Démonstration. Il suffit de vérifier que

$$(a_1, \dots, a_r) = (a_1, a_r)(a_1, a_{r-1}) \dots (a_1, a_2). \tag{1.381}$$

□

LEMooGGLUooUSzuAx

Lemme 1.287 ([1]).
Soit une permutation $\sigma \in S_E$. Soit un cycle c et une permutation s de supports disjoints telles que $\sigma = s \circ c$. Alors

LEMooUHWTooFptoZU

(1) Si $a \in \text{supp}(\sigma)$, alors pour tout $q \in \mathbb{N}$ nous avons $c^q(a) = \sigma^q(a)$.

ITEMooHSDLooIAKYZA

(2) Si $a \in \text{supp}(c)$, alors

$$\Omega_\sigma(a) = \Omega_c(a) = \text{supp}(c). \tag{1.382}$$

(3) Si $a \in \text{supp}(c)$, alors

$$c(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon.} \end{cases} \tag{1.383}$$

Démonstration. En plusieurs parties.

(1) **Si $a \in \text{supp}(c)$, alors $c(a) = \sigma(a)$** Soit $a \in \text{supp}(c)$. Nous savons que $c(a) \neq a$, et vu que c est injective, nous devons aussi avoir $c\big(c(a)\big) \neq c(a)$. Donc a et $c(a)$ sont dans $\text{supp}(c)$. Étant donné que les supports de c et de s sont disjoints, nous déduisons que $c(a)$ n'est pas dans le support de s, et donc que

$$\sigma(a) = (s \circ c)(a) = s\big(c(a)\big) = c(a). \tag{1.384}$$

(2) $\sigma^q(a) = c^q(a)$ Juste une récurrence sur le point précédent : si $b \in \text{supp}(a)$, alors $\sigma(b) = c(b) \in \text{supp}(a)$.

(3) **Si $a \in \text{supp}(c)$, alors $\Omega_\sigma(a) = \Omega_c(a)$** Utilisant le point précédent, ainsi que la définition 1.276 d'une orbite,

$$\Omega_\sigma(a) = \{\sigma^q(a)\} = \{c^q(a)\} = \Omega_c(a). \tag{1.385}$$

(4) $\text{supp}(c) \subset \Omega_c(a)$ Comme toujours, a est un élément de $\text{supp}(c)$. Nous considérons $b \in \text{supp}(c)$ et nous montrons que $b \in \Omega_c(a)$. Étant donné que $b \in \text{supp}(c)$, nous avons $c(b) \neq b$, de telle sorte que $\Omega_c(b)$ contienne au moins deux éléments distincts.

Même chose pour a : l'ensemble $\Omega_c(a)$ contient au moins a et $c(a)$. Vu que c est un cycle, il n'existe qu'une seule orbite non triviale. Donc $\Omega_c(a) = \Omega_c(b)$. En particulier $b \in \Omega_c(b) = \Omega_c(a)$.

(5) $\Omega_c(a) \subset \text{supp}(c)$ Soit $b \in \Omega_c(a)$. Le lemme 1.281 nous permet de dire que $\Omega_c(a) = \Omega_c(b)$. Comme $\Omega_c(b)$ contient au moins deux éléments (parce qu'il est égal à $\Omega_c(a)$ et que a est dans le support de c), nous savons que $c(b) \neq b$ et donc que $b \in \text{supp}(c)$.

(6) **La formule pour $c(x)$** Si $x \in \Omega_\sigma)(a)$, alors $c(x) = \sigma(x)$ par le point (1). Si x n'est pas dans $\Omega_c(a) = \text{supp}(c)$, alors x n'est pas dans le support de c et donc $c(x) = x$.

□

Le lemme suivant permet d'extraire le cycle de σ associé à un élément de E.

LEMooFFTBooCZsaFu

Lemme 1.288 ([1]).
Soient un ensemble fini E ainsi que $\sigma \in S_E$, et $a \in E$ tel que $\sigma(a) \neq a$. Nous posons

$$\begin{aligned} c\colon E &\to E \\ x &\mapsto \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon} \end{cases} \end{aligned} \tag{1.386}$$

Alors

(1) Si $b \in \Omega_\sigma(a)$, nous avons $\Omega_c(b) = \Omega_\sigma(a)$.

(2) Si $b \notin \Omega_\sigma(a)$, nous avons $\Omega_c(b) = \{b\}$.

(3) c est un cycle.

Démonstration. Si $b \in \Omega_\sigma(a)$, alors $\sigma^q(b) \in \Omega_\sigma(a)$ pour tout $q \in \mathbb{N}$, et donc

$$c^q(b) = \sigma^q(b) \in \Omega_\sigma(a) \tag{1.387}$$

pour tout q. Donc nous avons

$$\Omega_c(b) = \{c^q(b) \text{ tel que } q \in \mathbb{N}\} = \{\sigma^q(b) \text{ tel que } q \in \mathbb{N}\} = \Omega_\sigma(b) = \Omega_\sigma(a). \tag{1.388}$$

La dernière égalité est le lemme 1.282.

Si $b \notin \Omega_\sigma(a)$, alors $c(b) = b$ et $\Omega_c(b) = \{b\}$.

Nous avons prouvé que c a une seule orbite de taille plus grands ou égale à 2. Donc c est un cycle. □

Théorème 1.289 ([57]).
Soit un ensemble fini E de cardinal au moins deux. Soit une permutation $\sigma \in S_E$.

(1) Il existe des cycles $c_1, \dots, c_m$ à support disjoints tels que $\sigma = c_1 \circ \dots \circ c_m$.

(2) Cette décomposition est unique à l'ordre près.

Démonstration. Plusieurs points.

(1) **Existence** Nous choisissons des éléments $\{a_i\}_{i=1,\dots,p}$ tels que les $\Omega_\sigma(a_i)$ forment une partition de E en sous-ensembles disjoints. En posant $l_k = \min\{r \text{ tel que } \sigma^r(a_k) = a_k\}$, nous avons

$$\Omega_\sigma(a_k) = \{\sigma^q(a_k)\}_{q=1,\dots,l_k-1} \tag{1.389}$$

et tous les $\sigma^q(a_k)$ sont distincts pour $q = 1, \dots, l_k - 1$.

Posons

$$\begin{aligned} c_k\colon E &\to E \\ x &\mapsto \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a_k) \\ x & \text{sinon.} \end{cases} \end{aligned} \tag{1.390}$$

Le lemme 1.287 dit que c_k est un cycle. Vu que $c(a_k) = \sigma(a_k)$, le cycle c est un l_k-cycle.

Nous montrons à présent que $\sigma = c_1 \circ \dots c_p$. Soit $x \in E$. Il existe un $k \in \{1, \dots, p\}$ tel que

$$x \in \Omega_\sigma(a_k) = \Omega_{c_k}(a_k) = \Omega_{c_k}(x), \tag{1.391}$$

la dernière égalité est parce que $x \in \Omega_{c_k}(a_k)$. Nous en déduisons que $c_k(x) \neq x$. D'autre part si $l \neq k$, alors x n'est pas dans $\Omega_\sigma(a_l)$, et donc $c_l(x) = x$. Au final,

$$(c_1 \circ \dots c_p)(x) = c_k(x) = \sigma(x). \tag{1.392}$$

(2) **Unicité** Nous supposons avoir $\sigma = c_1 \circ \ldots \circ c_p = \gamma_1 \circ \gamma_q$ où les c_i et les γ_j sont deux ensembles de cycles de supports disjoints. Nous avons

$$\operatorname{supp}(\sigma) = \bigcup_{i=1}^{p} \operatorname{supp}(c_i) = \bigcup_{j=1}^{q} \operatorname{supp}(\gamma_j). \tag{1.393}$$

Montrons que si $\operatorname{supp}(c_i) \cap \operatorname{supp}(\gamma_j) \neq \varnothing$, alors $\operatorname{supp}(c_i) = \operatorname{supp}(\gamma_j)$. En effet si $a \in \operatorname{supp}(c_i) \cap \operatorname{supp}(\gamma_j)$, alors

$$\operatorname{supp}(c_i) = \Omega_{c_i}(a) = \Omega_\sigma(a) = \Omega_{\gamma_j}(a) = \operatorname{supp}(\gamma_j). \tag{1.394}$$

Et comme les $\operatorname{supp}(\gamma_j)$ sont disjoints, l'ensemble $\operatorname{supp}(c_i)$ n'a d'intersection qu'avec un et un seul des $\operatorname{supp}(\gamma_j)$. Cela définit donc une application

$$\begin{aligned} u\colon \{1,\ldots,p\} &\to \{1,\ldots,q\} \\ i &\mapsto \text{l'unique } j \text{ tel que } \operatorname{supp}(c_i) = \operatorname{supp}(\gamma_j). \end{aligned} \tag{1.395}$$

Autrement dit, l'application u permet d'écrire

$$\operatorname{supp}(c_i) = \operatorname{supp}\big(\gamma_{u(i)}\big). \tag{1.396}$$

L'application u est injective. En effet si $u(i) = u(l)$, nous avons

$$\operatorname{supp}(c_i) = \operatorname{supp}\big(\gamma_{u(i)}\big) \tag{1.397a}$$
$$\operatorname{supp}(c_l) = \operatorname{supp}\big(\gamma_{u(l)}\big) \tag{1.397b}$$
$$u(i) = u(l). \tag{1.397c}$$

Donc $\operatorname{supp}(c_i) = \operatorname{supp}(c_l)$. Et comme les supports sont disjoints, $i = l$.
L'application u est surjective. En effet, soit $j \in \{1,\ldots,q\}$. Soit $a \in \operatorname{supp}(\gamma_j)$. Il existe un i tel que $a \in \operatorname{supp}(c_i)$. Nous avons alors $a \in \operatorname{supp}(\gamma_j) \cap \operatorname{supp}(c_i)$, autrement dit $u(i) = j$.
Maintenant l'application $u\colon \{1,\ldots,p\} \to \{1,\ldots,q\}$ est bijective. Nous en déduisons que $p = q$. Concluons en montrant que $c_i = \gamma_{u(i)}$. Soit $a \in \operatorname{supp}(c_i) = \operatorname{supp}\big(\gamma_{u(i)}\big)$.
Nous avons

$$c_i(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon} \end{cases} \tag{1.398}$$

et

$$\gamma_j(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon,} \end{cases} \tag{1.399}$$

et donc $c_i = \gamma_j$.

□

LemmvZFWP

Lemme 1.290 ([59]).
Soit $\sigma = (i_1,\ldots,i_k) \in S_n$, un cycle de longueur k et $\theta \in S_n$. Alors

$$\theta\sigma\theta^{-1} = \big(\theta(i_1),\ldots,\theta(i_k)\big). \tag{1.400}$$

Tous les cycles de longueur k sont conjugués entre eux.

PropEAHWXwe

Proposition 1.291 (Classes de conjugaison et structure en cycles[60]).
Une classe de conjugaison[113] *dans S_n est formée des permutations ayant une décomposition en cycles disjoints de même structure. Autrement dit, deux permutations σ et σ' sont conjuguées si et seulement si le nombre k_i de cycles de longueur i dans σ est le même que le nombre k'_i de cycles de longueur i dans σ'.*

113. Définition 1.168.

Démonstration. Soit $\sigma = c_1 \dots c_m$ la décomposition de σ en cycles c_i de supports disjoints. Si τ est une permutation, alors

$$\sigma' = \tau\sigma\tau^{-1} = (\tau c_1 \tau^{-1}) \dots (\tau c_m \tau^{-1}), \tag{1.401}$$

mais $\tau c_i \tau^{-1}$ est un cycle de même longueur que c_i, puisque le lemme 1.290 nous dit que si $\sigma = (a_1, \dots, a_k)$, alors $\tau c \tau^{-1} = \big(\tau(a_1), \dots, \tau(a_k)\big)$. Notons encore que les cycles $\tau c_i \tau^{-1}$ restent à support disjoints.

Donc tous les éléments de la classe de conjugaison de σ sont des permutations de même structure que σ.

Réciproquement, si $\sigma' = c'_1 \dots c'_m$ est une décomposition de σ' en cycles disjoints tels que la longueur des c_i est la même que la longueur des c'_i, alors il suffit de construire des permutations τ_i telles que $\tau_i c_i \tau_i^{-1} = c'_i$, à travers le lemme 1.290. Comme les supports des c_i et des c'_i sont disjoints, la permutation $\tau_1 \dots \tau_m$ conjugue σ et σ'. □

EXooQAXRooBsPURs

Exemple 1.292.
Voyons les classes de conjugaison de S_3. Étant donné que ce groupe agit par définition sur un ensemble à 3 éléments, aucun élément de S_3 ne possède un cycle de plus de 3 éléments. Il y a donc seulement des cycles de longueur deux ou trois (à part les triviaux). Aucun élément de S_3 n'a une décomposition en cycles disjoints contenant deux cycles de deux ou un cycle de deux et un de trois.

En résumé il y a trois classes de conjugaison dans S_3. La première est celle contenant seulement l'identité. La seconde est celle contenant les cycles de longueur deux et la troisième contient les cycles de longueur 3.

Ce sont donc

$$C_1 = \{\mathrm{Id}\} \tag{1.402a}$$
$$C_2 = \{(1,2), (1,3), (2,3)\} \tag{1.402b}$$
$$C_3 = \{(1,2,3), (2,1,3)\}. \tag{1.402c}$$

△

DEFooXNAFooGTbTTJ

Définition 1.293 (transposition)**.**
Une ***transposition*** *est une permutation*[114] *qui échange deux éléments de E. Plus précisément, une bijection $\sigma : E \to E$ est une transposition si il existe $a, b \in E$ tels que*

$$\sigma(x) = \begin{cases} a & \text{si } x = b \\ b & \text{si } x = a \\ x & \text{sinon.} \end{cases} \tag{1.403}$$

ExVYZPzub

Exemple 1.294.
Les classes de conjugaison de S_4. Nous savons que les classes de conjugaison dans S_4 sont caractérisées par la structure des décompositions en cycles (proposition 1.291). Le groupe symétrique S_4 possède donc les classes de conjugaison suivantes.

(1) Le cycle vide qui représente la classe constituée de l'identité seule.
(2) Les transpositions (de type (a, b)) qui sont au nombre de 6.
(3) Les 3-cycles. Pour savoir quel est leur nombre nous commençons par remarquer qu'il y a 4 façons de prendre 3 nombres parmi 4 et ensuite 2 façons de les arranger. Il y a donc 8 éléments dans cette classe de conjugaison.
(4) Les 4-cycles. Le premier est arbitraire (parce que c'est cyclique). Pour le second il y a 3 possibilités, et deux possibilités pour le troisième ; le quatrième est alors automatique. Cette classe de conjugaison contient donc 6 éléments.

114. Une permutation est une bijection, définition 1.274.

ITEMooGCMYooKZgFHX

(5) Les doubles transpositions, du type $(a,b)(c,d)$. Dans ce cas, tous les nombres sont permutés, et l'image de 1 détermine la double transposition. Il y a 3 images possibles, et donc 3 éléments dans cette classe.

△

PropPWIJbu

Proposition 1.295.
Tout élément de S_n peut être écrit sous la forme d'un produit fini de transpositions.

Si E est un ensemble fini, tout élément de S_E pour être écrit sous forme d'un produit fini de transposition de E.

Démonstration. Un élément de S_n se décompose en un nombre fini de cycles par le lemme 1.284 et chacun des cycles peut être décomposé en un nombre fini de transpositions par le lemme 1.286. □

Cette décomposition n'est pas à confondre avec celle en cycles de support disjoints. Par exemple $(1,2,3) = (1,3)(1,2)$.

Le théorème suivant, qui donne la notion de parité d'une permutation, est la clef pour savoir quelles positions du jeu de taquin sont possibles ou impossibles[61, 62].

PROPooKRHEooAxtmRv

Proposition-Définition 1.296 (parité d'une permutation)**.**
À propos de décomposition ne permutations.

(1) Si une permutation peut être écrite sous forme d'un produit d'un nombre pair de transpositions, alors toute décomposition en transpositions sera en quantité paire.

(2) Si une permutation peut être écrite sous forme d'un produit d'un nombre impair de transpositions, alors toute décomposition en transpositions sera en quantité impaire.

*Une permutation qui se décompose en une quantité paire de transpositions est une **permutation paire** (et **impaire** sinon).*

DEFooNHXSooQzCPzD

Définition 1.297.
*La **signature** est l'application*

$$\begin{aligned}\epsilon\colon S_E &\to \{-1,1\}\\ \sigma &\mapsto \begin{cases}1 & \text{si } \sigma \text{ est paire}\\ -1 & \text{si } \sigma \text{ est impaire.}\end{cases}\end{aligned} \tag{1.404}$$

LEMooWGRXooHWyzLC

Lemme 1.298.
*Nous disons qu'un élément $\sigma \in S_n$ est une **inversion** pour les nombres $i<j$ si $\sigma(i)>\sigma(j)$. Soit N_σ le nombre d'inversions que $\sigma \in S_n$ possède (c'est le nombre de couples (i,j) avec $i<j$ tels que $\sigma(i)>\sigma(j)$). Nous avons*

$$\epsilon(\sigma) = (-1)^{N_\sigma} \tag{1.405}$$

où ϵ est la signature[115] *dans S_n.*

LemhxnkMf

Lemme 1.299 ([54])**.**
Un k-cycle est une permutation impaire si k est pair et paire si k est impair.

ProphIuJrC

Proposition 1.300 ([59])**.**
Soit S_n le groupe symétrique.

ITEMooBQKUooFTkvSu

(1) L'application $\epsilon\colon S_n \to \{1,-1\}$ est l'unique homomorphisme surjectif de S_n sur $\{-1,1\}$.

(2) Si $s = t_1\cdots t_k$ est le produit de k transpositions, alors $\epsilon(s) = (-1)^k$.

115. Définition 1.297.

Démonstration. Soit $\sigma, \theta \in S_n$. Afin de montrer que $\epsilon(\sigma\theta) = \epsilon(\sigma)\epsilon(\theta)$, nous divisons les couples (i, j) tels que $i \leqslant j$ en 4 groupes suivant que $\theta(i) \gtrless \theta(j)$ et $\sigma\big(\theta(i)\big) \gtrless \sigma\big(\theta(j)\big)$. Nous notons N_1, N_2, N_3 et N_4 le nombre de couples dans chacun des quatre groupes :

(i,j)	$\sigma\big(\theta(i)\big) < \sigma\big(\theta(j)\big)$	$\sigma\big(\theta(i)\big) > \sigma\big(\theta(j)\big)$
$\theta(i) < \theta(j)$	N_1	N_2
$\theta(i) > \theta(j)$	N_3	N_4

Nous avons immédiatement $N_\theta = N_3 + N_4$ et $N_{\sigma\theta} = N_2 + N_4$. Les éléments qui participent à N_σ sont ceux où $\theta(i)$ et $\theta(j)$ sont dans l'ordre inverse de $\sigma\big(\theta(i)\big)$ et $\sigma\big(\theta(j)\big)$ (parce que θ est une bijection). Donc $N_\sigma = N_2 + N_3$. Par conséquent nous avons

$$\epsilon(\sigma)\epsilon(\theta) = (-1)^{N_2+N_3}(-1)^{N_3+N_4} = (-1)^{N_2+N_4} = (-1)^{N_{\sigma\theta}} = \epsilon(\sigma\theta). \tag{1.406}$$

Nous avons prouvé que ϵ est un homomorphisme. Pour montrer que ϵ est surjectif sur $\{-1, 1\}$ nous devons trouver un élément $\tau \in S_n$ tel que $\epsilon(\tau) = -1$. Si τ est la transposition $1 \leftrightarrow 2$ alors le couple $(1, 2)$ est le seul à être inversé par τ et nous avons $\epsilon(\tau) = -1$.

Avant de montrer l'unicité, nous montrons que si $\sigma = t_1 \dots t_k$ alors $\epsilon(\sigma) = (-1)^k$. Pour cela il faut montrer que $\epsilon(\tau) = -1$ dès que τ est une transposition. Soit τ_{ij}, la transposition (i, j) et $\theta = (i, i + 1, \dots, j - 1)$ alors le lemme 1.290 dit que

$$\tau_{ij} = \theta\tau_{j-1,j}\theta^{-1}. \tag{1.407}$$

La signature étant un homomorphisme,

$$\epsilon(\tau_{ij}) = \epsilon(\theta)\epsilon(\tau_{j-1,j})\epsilon(\theta)^{-1} = \epsilon(\tau_{j-1,j}) = -1. \tag{1.408}$$

Nous passons maintenant à la partie unicité de la proposition. Soit un homomorphisme surjectif $\varphi\colon S_n \to \{-1, 1\}$ et τ, une transposition telle que $\varphi(\tau) = -1$ (qui existe parce que sinon φ ne serait pas surjectif[116]). Si τ' est une autre transposition, il existe $\sigma \in S_n$ tel que $\tau' = \sigma\tau\sigma^{-1}$ (lemme 1.290). Dans ce cas, $\varphi(\tau') = \varphi(\tau) = -1$, et si $\sigma = (\tau_1 \dots \tau_k)$,

$$\varphi(\sigma) = (-1)^k = \epsilon(\sigma). \tag{1.409}$$

□

CORooZLUKooBOhUPG

Corolaire 1.301.
Si $\sigma \in S_n$, alors

$$\epsilon(\sigma) = \epsilon(\sigma^{-1}). \tag{1.410}$$

Démonstration. Comme énoncé par la proposition 1.300, ϵ est un homomorphisme, donc

$$\epsilon(\sigma)\epsilon(\sigma^{-1}) = \epsilon(\sigma\sigma^{-1}) = \epsilon(\mathrm{Id}) = 1. \tag{1.411}$$

Puisque $\epsilon(\sigma)$ et $\epsilon(\sigma^{-1})$ ne peuvent valoir que ± 1, ils doivent être tous les deux égaux à 1 ou tous les deux à -1 pour que le produit soit 1. □

1.13.3 Permutation un peu ordonnées

LEMooEOTGooPslULz

Lemme 1.302 ([1]).
Deux énoncés.

ITEMooQZDRooUYjcgX

(1) Soit $\pi \in S_k$ satisfaisant $\pi(1) = 1$. Nous définissons $\tau \in S_{k-1}$ par $\tau(i) = \pi(i + 1) - 1$. Alors $\epsilon(\tau) = \epsilon(\pi)$.

ITEMooBUGZooVCVhKz

(2) L'application

$$\begin{aligned}\varphi\colon \{\pi \in S_k \text{ tel que } \pi(1) = 1\} &\to S_{k-1}\\ \pi &\mapsto \Big[\tau(i) = \pi(i + 1) - 1\Big]\end{aligned} \tag{1.412}$$

est une bijection.

116. Nous utilisons ici le fait que tous les éléments de S_n sont des produits de transpositions, proposition 1.295.

Démonstration. Nous nommons π' la restriction de π à $\{2,\dots,k\}$. Cela est encore une bijection, de telle sorte que π' puisse être écrite sous forme d'un produit de n transpositions de $\{2,\dots,k\}$ (proposition 1.295) :

$$\pi = \sigma_1 \circ \dots \circ \sigma_n. \tag{1.413}$$

Les σ_i étant des transpositions de $\{2,\dots,k\}$, elles sont des transpositions de $\{1,\dots,k\}$. Tout cela pour dire que π peut être écrite comme produit de transpositions ne faisant pas intervenir 1.

Nous introduisons la bijection

$$\begin{aligned}\psi\colon \{2,\dots,k\} &\to \{1,\dots,k-1\}\\ i &\mapsto i-1.\end{aligned} \tag{1.414}$$

En termes de ψ, la définition de τ s'écrit $\psi(\pi) = \psi\circ\pi\circ\psi^{-1}$, et nous avons

$$\tau = \psi\circ\pi\circ\psi^{-1} \tag{1.415a}$$
$$= \psi\circ\sigma_1\circ\dots\circ\sigma_n\circ\psi^{-1} \tag{1.415b}$$
$$= \psi\circ\sigma_1\circ\psi^{-1}\circ\psi\circ\sigma_2\circ\dots\circ\psi\circ\sigma_n\circ\psi^{-1}. \tag{1.415c}$$

Bref, en notant $\sigma_i' = \psi\sigma_i\psi^{-1}$ nous avons $\tau = \sigma_1'\circ\dots\circ\sigma_n$.

Il suffit maintenant de remarquer que σ_i' est une transposition dans $\{1,\dots,k-1\}$. On vérifie pour cela que si $\sigma = (a,b)$ alors $\sigma' = \big(\psi(a),\psi(b)\big)$. En effet, pour $i = 1,\dots,k-1$, il y a trois possibilités : soit $i = \psi(a)$, soit $i = \psi(b)$ soit i n'est ni l'un ni l'autre.

Si $i = \psi(a)$, alors $\sigma'(i) = (\psi\sigma\psi^{-1})\big(\psi(a)\big) = \psi\sigma(a) = \psi(b)$. Même vérification pour montrer que $\sigma'\big(\psi(b)\big) = \psi(a)$. Si i n'est ni $\psi(a)$ ni $\psi(b)$, alors $\psi^{-1}(i)$ n'est ni a ni b et dans ce cas

$$(\psi\sigma\psi^{-1})(i) = \psi\sigma\big(\psi^{-1}(i)\big) = \psi\big(\psi^{-1}(i)\big) = i. \tag{1.416}$$

Donc la décomposition $\tau = \sigma_1'\circ\dots\circ\sigma_n'$ est une décomposition de τ en n transpositions. Autrement dit le nombre de transpositions est le même pour τ que pour π.

En particulier les signatures de τ et de π sont les mêmes. Cela finit la preuve du point (1).

Prouvons le point (2).

(1) **Injectif** Supposons que $\varphi(\pi) = \varphi(\sigma)$. Alors pour tout $i = 1,\dots,k-1$ nous avons

$$\pi(i+1) - 1 = \sigma(i+1) - 1, \tag{1.417}$$

c'est-à-dire $\pi(j) = \sigma(j)$ pour tout $j = 2,\dots,k$. Vu que $\pi(1) = \sigma(1)$ nous avons bien $\pi = \sigma$.

(2) **Surjectif** Soit $\tau \in S_{k-1}$. En posant

$$\pi(i) = \begin{cases} 1 & \text{si } i = 1\\ \tau(i-1)+1 & \text{sinon,}\end{cases} \tag{1.418}$$

nous avons bien $\tau = \varphi(\pi)$.

□

DEFooDFBEooFElghU

Définition 1.303 ([1, 63]).
Nous notons

$$S_{(k,l)} = \{\pi \in S_{k+l} \text{ tel que } \pi(1) < \dots \pi(k),\ \pi(k+1) < \dots < \pi(k+l)\}, \tag{1.419}$$

et

$$A_{(k,l)} = \{\pi \in S_{(k,l)} \text{ tel que } \pi(1) = 1\}. \tag{1.420}$$

LEMooCKJAooBIAyVs

Lemme 1.304 ([1]).
Deux énoncés.

ITEMooCNWMooIbnHnz

(1) Soit $\pi \in A_{(k,l)}$. En posant $\tau(i) = \pi(i+1) - 1$ nous avons $\tau \in S_{(k-1,l)}$.

ITEMooKNKRooWbLYrF

(2) *L'application*

$$\begin{aligned}\varphi\colon A_{(k,l)} &\to S_{(k-1,l)}\\ \pi &\mapsto \Big[\tau(i) = \pi(i+1) - 1\Big]\end{aligned} \tag{1.421}$$

est une bijection.

(3) *En ce qui concerne la signature,* $\epsilon\big(\varphi(\pi)\big) = \epsilon(\pi)$.

Démonstration. Pour prouver (1), soit d'abord $1 \leqslant i < j \leqslant k-1$. Nous avons

$$\tau(i) = \pi(i+1) - 1 < \pi(i+j) - 1 = \tau(j) \tag{1.422}$$

parce que $i+1$ et $j+1$ sont entre 1 et k. Le même raisonnement tient pour $k+1 \leqslant i < j \leqslant k+l$.

Pour (2). Même preuve que la partie correspondante du lemme 1.302. □

1.14 Corps

La définition d'un corps est 1.209.

1.14.1 Définitions, morphismes

La proposition suivante donne une caractérisation d'un corps, en disant un tout petit peu plus que la définition 1.209.

Proposition 1.305.
L'anneau A est un corps si et seulement si $U(A) = A^*$.

Démonstration. En deux parties.

(1) **Sens direct** Nous supposons que A est un corps. D'une part tous les éléments non nuls sont inversibles, c'est-à-dire $A^* \subset U(A)$.
Pour l'inclusion inverse, nous montrons qu'une élément inversible ne peut pas être nul. Cela n'est autre que le lemme 1.178 couplé à la proposition 1.179 : $a \cdot 0 = 0 \neq 1$ pour tout a.

(2) **Sens inverse** Si $U(A) = A^*$, nous avons immédiatement que tous les éléments non nuls sont inversibles et donc que A est un corps.

□

LEMooJNIBooAURhrt

Lemme 1.306.
Si $\mathbb{K}$ est un corps et si $a \in \mathbb{K}$ vérifie $a^2 = 1$, alors $a = \pm 1$.

Définition 1.307 (Morphisme de corps)**.**
Un corps étant un anneau sans plus de structure, un ***morphisme de corps*** *n'est qu'un morphisme des anneaux*[117].

Le lemme suivant montre que définir un morphisme de corps comme étant simplement un morphisme des anneaux est une bonne idée.

LEMooWBOPooZnsZgQ

Lemme 1.308.
Si $\varphi\colon \mathbb{K} \to \mathbb{K}'$ est un morphisme de corps, alors

(1) pour tout $a \in \mathbb{K}$ nous avons $\varphi(a^{-1}) = \varphi(a)^{-1}$;

(2) le morphisme φ est injectif.

Démonstration. Vu que $\varphi(1) = 1$, nous avons aussi

$$1 = \varphi(aa^{-1}) = \varphi(a)\varphi(a^{-1}). \tag{1.423}$$

Donc, par unicité de l'inverse[118], $\varphi(a^{-1}) = \varphi(a)^{-1}$.

117. Définition 1.43.
118. Lemme 1.164 (2).

Pour l'injectivité nous supposons $\varphi(a) = \varphi(b)$. Étant donné que $\mathbb{K}'$ est un corps, nous pouvons multiplier par $\varphi(b)^{-1}$:

$$\varphi(a)\varphi(b)^{-1} = 1. \tag{1.424}$$

En utilisant le premier point nous avons $1 = \varphi(a)\varphi(b^{-1})$, puis le morphisme d'anneaux : $1 = \varphi(ab^{-1})$, et encore le morphisme d'anneaux nous permet de déduire $ab^{-1} = 1$ et donc $a = b$. □

1.15 Symbole de sommation

1.15.1 Somme à valeurs dans un groupe commutatif

Si S est un ensemble fini, nous savons de la proposition 1.126 qu'il existe un unique $N \in \mathbb{N}$ pour lequel il existe une bijection $\varphi\colon \{0, \dots, N\} \to S$. Cette bijection n'est à priori pas unique.

DEFooLNEXooYMQjRo

Lemme-Définition 1.309 ([1]).
Soient un groupe commutatif $(G, +)$ ainsi qu'un ensemble fini I contenant n éléments. Soit une application $f\colon I \to G$. Si $\sigma_1, \sigma_2\colon \{1, \dots, n\} \to I$ sont deux bijections, alors[119]

$$\sum_{i=1}^{n} f\big(\sigma_1(i)\big) = \sum_{i=1}^{n} f\big(\sigma_2(i)\big). \tag{1.425}$$

La valeur commune est notée

$$\sum_{i\in I} f(i) \tag{1.426}$$

Démonstration. Nous commençons par considérer une transposition σ (qui permute k et l avec $k < l$). Nous avons

$$\sum_{i=1}^{n} f(i) = \sum_{i=1}^{k-1} f(i) + f(k) + \sum_{i=k+1}^{l-1} f(i) + f(l) + \sum_{i=l+1}^{n} f(i) \tag{1.427a}$$

$$= \sum_{i=1}^{k-1} f(i) + f(l) + \sum_{i=k+1}^{l-1} f(i) + f(k) + \sum_{i=l+1}^{n} f(i) \tag{1.427b}$$

$$= \sum_{i=1}^{n} f\big(\sigma(i)\big). \tag{1.427c}$$

Pour cela nous avons utilisé le fait que G est commutatif pour permuter $f(l) \in G$ et $f(k) \in G$ avec $\sum_{i=k+1}^{l-1} f(i) \in G$.

Une permutation quelconque est un produit de telles transpositions (proposition 1.295). Donc pour toute permutation σ nous avons

$$\sum_{i=1}^{n} f\big(\sigma(i)\big) = \sum_{i=1}^{n} f(i). \tag{1.428}$$

□

La définition 1.309 donne lieu à un certain nombre de remarques.

(1) Elle donne la somme sur un ensemble fini. Un problème avec les ensembles infinis (outre la convergence) est l'ordre de sommation. Si vous voulez sommer sur $\mathbb{Z}$, dans quel ordre le faire ?

(2) Pour aller plus loin, et sommer sur des ensembles infinis, rendez-vous dans le thème 53.

119. Pour rappel, le symbole $\sum_{i=1}^{n}$ est défini par 1.85.

PROPooJBQVooNqWErk

Proposition 1.310.
Soient un groupe commutatif $(G,+)$, un ensemble fini I, une application $f\colon I\to G$ et une bijection $\sigma\colon I\to I$. Alors

$$\sum_{i\in I} f(i)=\sum_{i\in I} f\big(\sigma(i)\big). \tag{1.429}$$

Démonstration. Par définition 1.309, si $\varphi\colon\{1,\dots,n\}\to G$ est une bijection, alors

$$\sum_{i\in I} f(i)=\sum_{i=1}^{n}(f\circ\varphi)(i) \tag{1.430}$$

et

$$\sum_{i\in I}(f\circ\sigma)(i)=\sum_{i=1}^{n}(f\circ\varphi\circ\sigma)(i). \tag{1.431}$$

Vu que φ et $\varphi\circ\sigma$ sont des bijections, les deux sont égaux par le lemme 1.309. □

Si nous avons une application $L\colon S\to S$, nous notons

$$\sum_{s\in S} f\big(L(s)\big)=\sum_{s\in S}(f\circ L)(s). \tag{1.432}$$

Cette façon d'écrire donne une interprétation pour la notation $\sum_{g\in G} f(hg)$ qui arrive dans la proposition 1.313. Il s'agit de considérer l'application L_h du lemme 1.166, de considérer [120]

EQooQQBEooFDOBVG

$$\sum_{g\in G} f(hg)=\sum_{g\in G}(f\circ L_h)(g) \tag{1.433}$$

et de faire tourner la définition 1.309. La même chose tient pour définir $\sum_{g\in G}(gh)$ à l'aide de R_h.

LEMooGAMAooOAFhrc

Lemme 1.311 (Changement de variables dans une somme[1])**.**
Soient deux ensembles finis I,J ainsi qu'une bijection $\varphi\colon I\to J$. Soient un groupe abélien G et une application $f\colon I\to G$. Alors

$$\sum_{i\in I} f(i)=\sum_{j\in J} f\big(\varphi^{-1}(j)\big). \tag{1.434}$$

LEMooNDBYooAGEkmw

Lemme 1.312.
Soit un ensemble A fini pouvant être écrit comme une union disjointe $A=\bigcup_{k=1}^{n}A_k$; nous supposons que les A_i sont non vides. Soient un groupe commutatif $(G,+)$ et une application $f\colon A\to G$. Alors

$$\sum_{a\in A} f(a)=\sum_{k=1}^{n}\sum_{a\in A_k} f(a). \tag{1.435}$$

Démonstration. Le lemme 1.119 nous indique que les parties A_k sont des ensembles finis. Nous notons

(1) $N_0=0$, et $N_k=\mathrm{Card}(A_k)$,
(2) $S_k=\sum_{i=1}^{k}N_k$.
(3) $\varphi_k\colon\{1,\dots,N_k\}\to A_k$, une bijection (l'existence est dans la proposition 1.126).

Nous avons $\mathrm{Card}(A)=S_n$ par le lemme 1.128(4). Nous définissons une belle bijection comme il faut :

$$\begin{aligned}\alpha\colon\{1,\dots,S_n\}&\to A\\ i&\mapsto\varphi_{k+1}(i-S_k)\end{aligned} \tag{1.436}$$

pour $i\in\,]S_k,S_{k+1}]$.

120. Le fait que L_h soit une bijection n'a pas d'importance ici.

(1) α **est bien définie** Puisque $i > S_k$ et $i \leqslant S_{k+1}$ nous avons $i - S_k \in \{1, \dots, N_{k+1}\}$, et donc φ_{k+1} s'applique bien à $i - S_k$.

(2) α **est injective** Supposons que $\alpha(i) = \alpha(j)$. Si $i \in]S_k, S_{k+1}]$ et $j \in]S_l, S_{l+1}]$, alors $\alpha(i) = \varphi_{k+1}(i - S_k) \in A_{k+1}$ et $\alpha(j) = \varphi_{l+1}(j - S_l) \in A_{l+1}$. Vu que les A_i sont disjoints, nous avons $k = l$, et donc

$$\varphi_{k+1}(i - S_k) = \varphi_{k+1}(j - S_k). \tag{1.437}$$

Étant donné que φ_{k+1} est injective, nous avons $i - S_k = j - S_k$, ce qui montre que $i = j$.

(3) α **est surjective** Soit $a \in A$. Il existe k tel que $a \in A_k$. Nous avons donc un $s \in \{1, \dots, N_k\}$ tel que $a = \varphi_k(s)$. En posant $i = s + S_k$, nous avons bien $a = \alpha(s + S_k)$ parce que $s + S_k \in]S_{k-1}, S_k]$.

Vu que α est une bijection, nous avons l'égalité

$$\sum_{a \in A} f(a) = \sum_{i=1}^{S_n} (f \circ \alpha)(i). \tag{1.438}$$

Nous avons encore besoin d'introduire une bijection. Nous posons

$$\begin{aligned} \beta_k \colon]S_{k-1}, S_k] &\to A_k \\ i &\mapsto \varphi_k(i - S_{k-1}). \end{aligned} \tag{1.439}$$

C'est une bijection parce que φ_k en est une, et que $i \mapsto i - S_{k-1}$ est une bijection de $]S_{k-1}, S_k]$.

Nous pouvons maintenant terminer :

$$\sum_{a \in A} f(a) = \sum_{i=1}^{S_n} (f \circ \alpha)(i) \tag{1.440a}$$

$$= \sum_{k=1}^{n} \left(\sum_{i=S_{k-1}-1}^{S_k} (f \circ \alpha)(i) \right) \tag{1.440b}$$

SUBEQooNVKWooZqBAau

$$= \sum_{k=1}^{n} \left(\sum_{i \in]S_{k-1}, S_k]} f\big(\varphi_k(i - S_{k-1})\big) \right) \tag{1.440c}$$

$$= \sum_{k=1}^{n} \left(\sum_{i \in]S_{k-1}, S_k]} f\big(\beta_k(i)\big) \right) \tag{1.440d}$$

$$= \sum_{i=1}^{n} \left(\sum_{a \in A_k} f(a) \right). \tag{1.440e}$$

Justifications :

— Pour (1.440b). Associativité de la somme.

□

PROPooWJQQooFINSEc

Proposition 1.313 ([1])**.**
Soient un groupe fini G et une fonction $f \colon G \to A$ où A est un anneau commutatif. Alors

$$\sum_{g \in G} f(g) = \sum_{g \in G} f(gh) = \sum_{g \in G} f(hg) \tag{1.441}$$

pour tout $h \in G$.

Démonstration. Nous avons une bijection $\varphi \colon \{0, \dots, N\} \to G$ garantie par la proposition 1.126. Sa définition est

$$\sum_{g \in G} f(g) = \sum_{i=0}^{N} f\big(\varphi(i)\big). \tag{1.442}$$

Par ailleurs, le lemme 1.166 donne une bijection $L_h\colon G\to G$ et permet de considérer la composée

$$\begin{aligned}\varphi'\colon \{0,\dots,N\}&\to G\\ \varphi'&=L_h\circ\varphi.\end{aligned}\tag{1.443}$$

La proposition 1.309 nous permet d'utiliser la bijection φ' au lieu de φ pour exprimer la somme $\sum_{g\in G}$. Ensuite un jeu de notation utilisant (1.433) donne

$$\begin{aligned}\sum_{g\in G}f(g)&=\sum_{i=0}^{N}f\big(\varphi(i)\big)=\sum_{i=0}^{N}f\big(\varphi'(i)\big)=\sum_{i=0}^{N}(f\circ L_h\circ\varphi)(i)\\ &=\sum_{i=0}^{N}(f\circ L_h)\big(\varphi(i)\big)=\sum_{g\in G}(f\circ L_h)(g)=\sum_{g\in G}f(hg).\end{aligned}\tag{1.444}$$

En ce qui concerne $\sum_{g\in G}f(gh)$, c'est la même chose, en utilisant R_h au lieu de L_h. □

LEMooKSVWooIFsfwm

Lemme 1.314.
Soit un groupe totalement ordonné[121] $(A,+,\leqslant)$. *Soient deux suites* (a_i) *et* (b_i) *dans* G *telles que* $a_i\leqslant b_i$ *pour tout* i. *Alors pour tout* n *nous avons*

$$\sum_{i=0}^{n}a_i\leqslant\sum_{i=0}^{n}b_i.\tag{1.445}$$

Tout cela nous permet de définir une somme sympathique et bien connue.

Lemme 1.315.
Soit $n\in\mathbb{N}$. *Nous avons*

$$\sum_{k=0}^{n}k=\frac{n(n+1)}{2}.\tag{1.446}$$

Démonstration. La preuve est pratiquement immédiate par récurrence. Nous allons donner une preuve plus « constructive », qui formalise l'idée classique d'écrire la somme à l'endroit et à l'envers.

Nous notons S la somme $\sum_{k=0}^{n}k$. Le lemme 1.309 dit que si les $\sigma_i\colon\{0,\dots,n\}\to\{0,\dots,n\}$ sont des bijections, alors $\sum_{k=0}^{n}f\big(\sigma_1(k)\big)=\sum_{k=0}^{n}f\big(\sigma_2(k)\big)$. Nous sommes intéressé au cas $f(i)=i$.

En prenant $\sigma_1(k)=k$ et $\sigma_2(k)=n-k$, nous avons

$$S=\sum_{k=0}^{n}k=\sum_{k=0}^{n}(n-k).\tag{1.447}$$

Donc

$$2S=\sum_{k=0}^{n}\big(k+(n-k)\big)=\sum_{k=0}^{n}n=n\sum_{k=0}^{n}1=n(n+1).\tag{1.448}$$

En divisant par deux, nous obtenons le résultat annoncé. □

1.16 Symbole de produit

NORMooDBOFooQCwbOY

1.316.
Si $(G,\ \cdot\ )$ est un groupe et si $H\subset G$, nous notons le produit des éléments de H par

$$\prod_{g\in H}g=\sum_{g\in H}g\tag{1.449}$$

où à droite, c'est la somme déjà définie. La différence entre $\prod$ et $\sum$ est que nous utilisons $\prod$ pour les groupes notés « multiplicativement » comme $(G,\ \cdot\ )$ alors que nous utilisons $\sum$ lorsque le groupe est noté « additivement » comme $(G,+)$.

121. Définition 1.167.

Dans le cas d'un anneau $(A, +, \cdot)$, la distinction est importante pour savoir quelle opération est sous-entendue.

La définition 1.85(1) signifie qu'une somme vide vaut zéro : $\sum_{x\in\emptyset} x = 0$. Vu que zéro est la façon usuelle de noter le neutre pour une opération notée « + », lorsque l'opération est notée $\cdot$ nous avons

$$\prod_{x\in\emptyset} x = 1 \tag{1.450}$$

EQooCSDSooTxdfzQ

parce que 1 est la façon usuelle de noter le neutre d'une opération notée « $\cdot$ ».

Notez que (1.450) n'est pas une nouvelle définition ou une nouvelle convention. C'est seulement l'égalité $\sum_{x\in\emptyset x} x = 0$, avec des notations adaptées à un groupe dont l'opération est notée multiplicativement.

PROPooQMUDooQQVRIe

Proposition 1.317.
Si E est un ensemble fini et si G est un groupe commutatif, alors pour toute fonction $f : E \to G$ et pour toute permutation[122] *σ de E,*

$$\prod_{i\in E} f(i) = \prod_{i\in E} f\big(\sigma(i)\big) \tag{1.451}$$

Démonstration. C'est exactement la proposition 1.309, sauf qu'ici la loi de groupe est notée multiplicativement au lieu d'additivement. □

1.16.1 Sous-groupe engendré

DefooRDRXooEhVxxu

Définition 1.318 (Sous-groupe engendré)**.**
*Soit A une partie du groupe G. Le sous-groupe **engendré** par A est l'intersection de tous les sous-groupes de G contenant A. Nous notons ce groupe $\mathrm{gr}_G(A)$.*

Lorsque A est fini (disons $A = \{a_1, \dots, a_n\}$), on note aussi le sous-groupe engendré $\langle a_1, \dots, a_n\rangle$.

1.319.
Un sous-groupe engendré n'est jamais vide parce qu'il contient toujours au moins le neutre (parce que c'est un sous-groupe). Si G est un groupe, le sous-groupe $\mathrm{gr}_G(\emptyset)$ lui-même contient e[123].

1.320.
Dans de nombreux cas, le groupe « ambiant » G est entendu par le contexte et nous noterons $\mathrm{gr}(A)$ au lieu de $\mathrm{gr}_G(A)$.

Si par exemple A est la matrice $\begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$, le groupe $\mathrm{gr}(A)$ est à comprendre dans $\mathrm{GL}(2, \mathbb{R})$. Il faudrait être fou pour avoir en tête un autre groupe que $\mathrm{GL}(2, \mathbb{R})$ sans le préciser.

D'ailleurs, connaissez-vous un groupe contenant la matrice A et n'étant pas un sous-groupe de $\mathrm{GL}(2, \mathbb{C})$?

Lemme 1.321.
Si G est un groupe et A une partie de G, alors $\mathrm{gr}(A)$ est un sous-groupe de G.

Le sous-groupe engendré par A est le plus petit (pour l'inclusion) groupe de G contenant A. Plus formellement, nous avons le résultat suivant :

Lemme 1.322.
Tout sous-groupe de G contenant A contient $\mathrm{gr}(A)$.

Démonstration. Si H est un sous-groupe de G contenant A, alors $\mathrm{gr}(A)$ est l'intersection de H avec tous les autres sous-groupes de G contenant A. Il contient donc $\mathrm{gr}(A)$. □

122. Une permutation est une bijection, définition 1.274.
123. Demandez-vous si il est possible que $\mathrm{gr}(\emptyset)$ contienne d'autres éléments que e.

LemFUIZooBZTCiy

Lemme 1.323 ([64]).
Si A est une partie du groupe G, alors le sous-groupe $\mathrm{gr}(A)$ engendré[124] *par A est l'ensemble de tous les produits finis d'éléments de A et de A^{-1} (l'identité est le produit à zéro éléments).*

C'est-à-dire que tout élément de $\mathrm{gr}(A)$ peut être écrit sous la forme[125]

$$\prod_{i=1}^{n} g_i^{a_i} \tag{1.452}$$

où $a_i \in \mathbb{Z}$ et $g\colon \mathbb{N} \to A$ n'est pas spécialement injective : il peut arriver que $g_i = g_j$.

Démonstration. Puisqu'un produit vide est égal à l'identité[126], le lemme est vrai (un peu trivialement) dans le cas où $A = \varnothing$. À partir de maintenant, nous supposons que A est non vide.

Nous nommons $\mathrm{gr}(A)$ le groupe engendré par A et H, l'ensemble

$$H = \{g_1 \dots g_n \text{ tel que } g_i \in A \cup A^{-1}\}. \tag{1.453}$$

Nous commençons par prouver que H est un groupe.

— Puisque A est non vide, nous considérons $a \in A$. Dans ce cas, $e = aa^{-1} \in H$. Donc $e \in H$.

— L'inverse de $g_1 \dots g_n$ est $g_n^{-1} \dots g_1^{-1}$ qui est également dans H.

— Le produit de $g_1 \dots g_n$ par $h_1 \dots h_n$, tous éléments de H, est également dans H[127].

Comme H est un groupe contenant A, nous avons $\mathrm{gr}(A) \subset H$ parce que $\mathrm{gr}(A)$ est une intersection dont un des éléments est H.

Par ailleurs tout groupe contenant A doit contenir les inverses et les produits finis, donc $H \subset \mathrm{gr}(A)$.

Au final, $H = \mathrm{gr}(A)$, ce qu'il fallait. □

LEMooCFTVooKvmyKN

Lemme 1.324.
Soit un groupe G et un sous-groupe $H = \mathrm{gr}(h_1, \dots, h_n)$. Si $\alpha \in G$, alors

$$\alpha H \alpha^{-1} = \mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1}). \tag{1.454}$$

Démonstration. Il s'agit d'une conséquence du lemme 1.323. Un élément de $\mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1})$ est un produit d'éléments de G de la forme $\alpha h_i \alpha^{-1}$ ou $(\alpha h_j \alpha^{-1})^{-1} = \alpha h_j^{-1} \alpha^{-1}$. Or nous avons

$$\alpha h_i \alpha^{-1} \alpha h_j \alpha^{-1} = \alpha h_i h_j \alpha^{-1} \in \alpha H \alpha^{-1}. \tag{1.455}$$

Donc

$$\mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1}) \subset \alpha H \alpha^{-1}. \tag{1.456}$$

L'inclusion dans l'autre sens est du même tonneau. □

DEFooWMFVooLDqVxR

Définition 1.325 (Partie génératrice, groupe monogène).
*Soient un groupe G, et une partie $A \subset G$. Si $\mathrm{gr}(A) = G$, alors nous disons que A est une **partie génératrice** du groupe G.*

*Un groupe est **monogène** si il a une partie génératrice réduite à un seul élément.*

DefHFJWooFxkzCF

Définition 1.326 (Groupe cyclique).
*Un élément $a \in G$ est un **générateur** de G si tous les éléments de G s'écrivent sous la forme a^n pour un certain $n \in \mathbb{Z}$. Un groupe fini et monogène est dit **cyclique**.*

124. Définition 1.318.
125. Les a_i négatifs correspondent aux inverses. Notons que si $g \in A$, il n'y a pas de garanties que g^{-1} soit également dans A.
126. Voir 1.316.
127. Et c'est ici qu'on se rend compte que la décomposition n'est probablement que rarement unique.

1.327.
La différence entre un groupe monogène et un groupe cyclique est qu'un groupe cyclique est fini. Dans un groupe cyclique, à force d'itérer le générateur, nous finissons par tourner en rond – d'où le nom.

Exemple 1.328.
Soit le groupe $\big(\mathbb{Z}/10\mathbb{Z},+\big)$. L'élément $[2]_{10}$ n'est pas générateur parce que ses puissances [128] sont

$$\mathrm{gr}([2]_{10}) = \{[2]_{10}, [4]_{10}, [6]_{10}, [8]_{10}, [0]_{10}\}. \tag{1.457}$$

Par contre l'élément $[3]_{10}$ est générateur : ses puissances sont dans l'ordre

$$[3]_{10}, [6]_{10}, [9]_{10}, [2]_{10}, [5]_{10}, [8]_{10}, [1]_{10}, [4]_{10}, [7]_{10}, [0]_{10}. \tag{1.458}$$

△

Un exemple presque identique, mais un peu masqué sera l'exemple 18.156.

LEMooTHBIooJPjBlp

Lemme 1.329 ([1])**.**
Si $n = dr$, alors

(1) $r\mathbb{Z}/n\mathbb{Z}$ est un groupe.

LEMooSAAXooMezYui

(2) Le groupe $r\mathbb{Z}/n\mathbb{Z}$ est cyclique [129]*.*

ITEMooGCGQooIfZULt

(3) $\mathrm{Card}(r\mathbb{Z}/n\mathbb{Z}) = d = n/r$.

1.17 Module sur un anneau

DEFooHXITooBFvzrR

Définition 1.330 (module sur un anneau[65])**.**
*Soit un anneau A. Un **module à gauche** sur A est la donnée d'un triplet $(M, +, \cdot)$ où*

(1) $+$ est une loi de composition interne à M, c'est-à-dire $+ : M \times M \to M$,

(2) $\cdot$ est une loi de composition externe, c'est-à-dire $\cdot : A \times M \to M$

telles que

(1) $(M, +)$ est un groupe [130]*.*

(2) $a \cdot (x + y) = a \cdot x + a \cdot y$,

(3) $(a + b) \cdot x = a \cdot x + b \cdot x$,

(4) $(ab) \cdot x = a \cdot (b \cdot x)$

(5) $1 \cdot x = x$.

pour tout $a, b \in A$ et $x, y \in M$.

*Si M et N sont des A-modules, un **morphisme** de M vers N est une application $f : M \to N$ qui*

(1) est un morphisme de groupes entre $(M, +)$ et $(N, +)$

(2) vérifie $f(a \cdot x) = a \cdot f(x)$ pour tout $a \in A$, $x \in M$.

L'ensemble des morphismes entre M et N est noté $\mathrm{Hom}_A(M, N)$. Si B est un sous-anneau de A, nous parlons de $\mathrm{Hom}_B(M, N)$ pour parler des morphismes de groupes qui ne vérifient $f(a \cdot x) = a \cdot f(x)$ que pour $a \in B$.

PROPooGARGooDiMqtN

Proposition 1.331.
Si M est un module sur un anneau, alors $(M, +)$ est un groupe commutatif.

128. Attention aux notations ; en général on écrit la loi de groupe de façon multiplicative et on parle des puissances d'un élément, mais ici on écrit la loi de groupe additivement, donc les « puissances » sont en réalité les multiples.
129. Définition 1.326.
130. Nous verrons dans la proposition 1.331 qu'il est forcément commutatif.

Démonstration. Il suffit de calculer $(1+1)\cdot(x+y)$ de deux façons différentes :

$$(1+1)\cdot(x+y) = 1\cdot(x+y)+1\cdot(x+y) = x+y+x+y \tag{1.459}$$

d'une part et

$$(1+1)\cdot(x+y) = (1+1)\cdot x+(1+1)\cdot y = x+x+y+y, \tag{1.460}$$

d'autre part. En égalant les deux expressions, il vient

$$x+y+x+y = x+x+y+y, \tag{1.461}$$

qui se simplifie (nous sommes dans un groupe) en $y+x=x+y$. □

DEFooKHWZooIfxdNc

Définition 1.332.
*Un **espace vectoriel** est un module*[131] *sur un corps commutatif*[132].

DEFooRUKVooLnXxdS

Définition 1.333 ([66])**.**
*Soient un A-module M et un ensemble I. Une famille $\{m_i\}_{i\in I}$ est **libre** si les m_i sont **linéairement indépendants**, c'est-à-dire si pour tout choix d'une partie finie J dans I et d'éléments $(a_j)_{j\in J}$ dans A, si nous avons*

$$\sum_{j\in J} a_j m_j = 0, \tag{1.462}$$

alors $a_j=0$ pour tout j.

DEFooWBOBooJNyyBF

Définition 1.334 ([67])**.**
*Soit S, une partie du A-module M. Le **sous-module engendré** par S est l'ensemble des éléments de M qui sont des combinaisons linéaires finies d'éléments de S, c'est-à-dire de sommes de la forme*

$$\sum_{t\in T} a_t t \tag{1.463}$$

où T est fini dans S et $a_t\in A$.

1.17.1 Module produit

DEFooLCJEooBvVmkV

Lemme-Définition 1.335 ([66])**.**
*Soient un anneau A et un ensemble I. Le A-**module produit** A^I est l'ensemble des applications $I\to A$.*

En termes de notations, nous écrivons ceci :

$$A^I = \{(a_i)_{i\in I}, a_i\in A\}. \tag{1.464}$$

L'ensemble A^I devient un module par les définitions, pour $x,y\in A^I$ et $a\in A$:

EQooPLBSooFuQXrj

$$ax = (ax_i)_{i\in I} \tag{1.465a}$$

EQooODBMooQKLUgd

$$x+y = (x_i+y_i)_{i\in I}. \tag{1.465b}$$

En d'autres termes, $A^I = \mathrm{Fun}(I,A)$.

Démonstration. Il faut vérifier toutes les conditions de la définition 1.330. En guise d'exemple, nous vérifions la distributivité. Soient $a\in A$ et $x,y\in A^I$. Nous avons

$$\begin{aligned}
\big[a\cdot(x+y)\big]_i &= a(x+y)_i && && \text{(1.466a)}\\
&= a(x_i+y_i) && \text{(1.465b)} && \text{(1.466b)}\\
&= ax_i+ay_i && \text{distrib. dans } A && \text{(1.466c)}\\
&= (ax+ay)_i && \text{(1.465b)}. && \text{(1.466d)}
\end{aligned}$$

□

131. Définition 1.330.
132. La condition de commutativité n'est pas indispensable, mais comme nous ne parlerons que de corps commutatifs...

Lemme 1.336.
Pour chaque $i \in I$ nous considérons l'élément $e_i \in A^I$ donné par

$$\begin{aligned} e_i \colon I &\to A \\ j &\mapsto \begin{cases} 1 & \text{si } j = i \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.467}$$

La famille $\{e_i\}_{i\in I}$ est libre[133] *dans A^I.*

Démonstration. Soient J fini dans I ainsi que des éléments $a_j \in A$ ($j \in J$). Nous supposons que[134] $\sum_{j\in J} a_j e_j = 0$. Calculons un peu :

$$\sum_{j\in J} a_j e_j = \sum_{j\in J} (a_j \delta_{ji})_{i\in I} = \left(\sum_{j\in J} a_j \delta_{ji} \right)_{i\in I}. \tag{1.468}$$

Pour que le tout soit nul dans A^I, il faut que

$$\sum_{j\in J} a_j \delta_{ji} \tag{1.469}$$

soit nul pour tout $i \in I$. Si nous fixons $i \in I$, la somme sur j possède un seul terme non annulé par δ_{ji}, et c'est le terme $j = i$. Nous avons donc $a_i = 0$. □

DEFooBMEPooFsCHgb

Définition 1.337.
Nous notons $A^{(I)}$ le sous-module de A^I engendré[135] *par les e_i.*

Lemme 1.338 ([1])**.**
L'ensemble $A^{(I)}$ est l'ensemble des applications $I \to A$ de support fini.

Démonstration. En deux sens.

(1) **Si $x \in A^{(I)}$** Pour rappel, la définition 1.335 nous dit que x est une application $I \to A$. Vu que x est dans le sous-module engendré par les e_i, il existe une partie finie $J \subset I$ telle que

$$x = \sum_{j\in J} x_j e_j. \tag{1.470}$$

Pour $i \in I$ nous avons

$$x(i) = \sum_{j\in J} x_j \delta_{ij} = \begin{cases} x_i & \text{si } i \in J \\ 0 & \text{sinon.} \end{cases} \tag{1.471}$$

Donc le support de x est dans J qui est fini. Vu que toute partie d'un ensemble fini est fini (lemme 1.119), le support de x est fini.

(2) **Si x est de support fini** Supposons que le support de $x \colon I \to A$ soit la partie finie $J \subset I$. En notant $x_j = x(j)$ pour tout $j \in J$, nous avons

$$x = \sum_{j\in J} x_j e_j. \tag{1.472}$$

□

THOooPDZCooJnHbOd

Théorème 1.339 (Propriété universelle de $A^{(I)}$[66])**.**
Soient un anneau A ainsi qu'un A-module P. Pour $\phi \in \mathrm{Hom}_A(A^{(I)}, P)$, nous considérons

$$\begin{aligned} \phi|_I \colon I &\to P \\ i &\mapsto \phi(e_i). \end{aligned} \tag{1.473}$$

133. Définition 1.333.
134. Pour rappel, les sommes finies sont définies par 1.309.
135. Définition 1.334.

(1) *L'application*

$$\begin{aligned} f\colon \mathrm{Hom}_A(A^{(I)},P) &\to \mathrm{Fun}(I,P)\\ \phi &\mapsto \phi|_I \end{aligned} \tag{1.474}$$

est une bijection.

(2) *L'application inverse est* $g\colon \mathrm{Fun}(I,P)\to \mathrm{Hom}_A(A^{(I)},P)$ *donnée par*

$$g(\psi)\big(\sum_{j\in J} a_j e_j\big) = \sum_{j\in J} a_j\psi(j) \tag{1.475}$$

pour tout J *fini dans* I *et choix de* $a_j\in A$.

Démonstration. Nous allons montrer que $g\big(f(\phi)\big)=\phi$ et que $f\big(g(\psi)\big)=\psi$ pour tout $\phi\in \mathrm{Hom}_A(A^{(I)},P)$ et pour tout $\psi\in\mathrm{Fun}(I,P)$.

Dans un premier sens nous avons :

$$g\big(f(\phi)\big)\big(\sum_j a_j e_j\big) = \sum_j a_j f(\phi)(j) \tag{1.476a}$$

$$= \sum_j a_j \phi(e_j) \tag{1.476b}$$

SUBALIGNooBWPLooHclaQg

$$= \phi\big(\sum_j a_j e_j\big). \tag{1.476c}$$

SUBALIGNooUOQPooCwIgZo

Justifications :

— Pour (1.476b), nous avons utilisé le fait que $f(\phi)(i)=\phi|_I(i)=\phi(e_i)$.

— Pour (1.476c), nous utilisons le fait que ϕ est un morphisme de modules.

Et pour l'autre sens,

$$f\big(g(\psi)\big)(i) = g(\psi)(e_i) = \psi(i). \tag{1.477}$$

Vérifions que cela est suffisant pour que f soit une bijection.

(1) **Surjectif** Soit $\psi\in\mathrm{Fun}(I,P)$. Nous avons $f\big(g(\psi)\big)=\psi$, ce qui prouve que ψ est dans l'image de f.

(2) **Injectif** Supposons que $f(\phi_1)=f(\phi_2)$. Alors en appliquant g des deux côtés, il vient $\phi_1=\phi_2$.

□

1.17.2 Sous-module

Soient M un A-module et $x=(x_i)_{i\in I}$ une famille d'éléments de M paramétrée par l'ensemble I. Nous considérons l'application

$$\begin{aligned} \mu_x\colon A^{(I)} &\to M\\ (a_i)_{i\in I} &\mapsto \sum_{i\in I} a_i x_i. \end{aligned} \tag{1.478}$$

Ici $A^{(I)}$ désigne l'ensemble de toutes les applications $I\to A$ de support fini (définition 1.337).

DefBasePouyKj

Définition 1.340.
À l'instar des espaces vectoriels, les modules ont une notion de partie libre, génératrice et de bases :

(1) *Si* μ_x *est surjective, nous disons que* x *est une partie* ***génératrice****.*

(2) *Si* μ_x *est injective, nous disons que la partie* x *est* ***libre****.*

(3) *Si* μ_x *est bijective, nous disons que la partie* x *est une* ***base****.*

Définition 1.341.
Un sous-ensemble $N\subset M$ *est un* ***sous-module*** *si* $(N,+)$ *est un sous-groupe de* $(M,+)$ *et si* $a\cdot x\in N$ *pour tout* $x\in N$ *et pour tout* $a\in A$.

Exemple 1.342.
Un anneau A est lui-même un A-module et ses sous-modules sont les idéaux. △

Définition 1.343.
*Soit M un module sur un anneau commutatif A. Un **projecteur** est une application linéaire $p\colon M \to M$ telle que $p^2 = p$.*

*Une famille $(p_i)_{i\in I}$ sur M est **orthogonale** si $p_i \circ p_j = 0$ pour tout $i \neq j$. La famille est **complète** si $\sum_{i\in I} p_i = \mathbb{1}$.*

ThoProjModpAlsUR

Théorème 1.344.
Soient des sous-modules $M_1, \dots, M_n$ du module M tels que $M = M_1 \oplus \dots \oplus M_n$. Les applications p_i définies par

$$p_i(x_1 + \dots + x_n) = x_i \tag{1.479}$$

forment une famille orthogonale de projecteurs et $p_1 + \dots + p_n = \mathrm{Id}$.

Inversement, si $(p_1, \dots, p_n)$ est une famille orthogonale de projecteurs dans un module $\mathcal{E}$ tel que $\sum_{i=1}^n p_i = \mathrm{Id}$, alors

$$M = \bigoplus_{i=1}^{n} p_i(M). \tag{1.480}$$

Définition 1.345.
*Un module est **simple** ou **irréductible** si il n'a pas d'autres sous-modules que $\{0\}$ et lui-même. Un module est **indécomposable** si il ne peut pas être écrit comme somme directe de sous-modules.*

Un module simple est a fortiori indécomposable. L'inverse n'est pas vrai comme le montre l'exemple suivant.

Exemple 1.346.
Soit $\mathcal{E} = \mathbb{C}[X]/(X^2)$ vu comme $\mathbb{C}[X]$-module. C'est le $\mathbb{C}[X]$-module des polynômes de la forme $aX + b$ avec $a, b \in \mathbb{C}$. L'ensemble des polynômes de la forme aX est un sous-module. Le module $\mathcal{E}$ n'est donc pas simple. Il est cependant indécomposable parce que $\{aX\}$ est le seul sous-module non trivial. En effet si $\mathcal{F}$ est un sous-module de $\mathcal{E}$ contenant $aX + b$ avec $b \neq 0$, alors $\mathcal{F}$ contient $X(aX + b) = bX$ et donc contient tout $\mathcal{E}$. △

DefAEbnJqI

Définition 1.347 (Algèbre[68])**.**
Si $\mathbb{K}$ est un corps commutatif[136]*, une $\mathbb{K}$-**algèbre** A est un espace vectoriel*[137] *muni d'une opération bilinéaire $\times\colon A \times A \to A$, c'est-à-dire telle que pour tout $x, y, z \in A$ et pour tout $\alpha, \beta \in \mathbb{K}$,*

(1) $(x + y) \times z = x \times z + y \times z$
(2) $x \times (y + z) = x \times y + x \times z$
(3) $(\alpha x) \times (\beta y) = (\alpha\beta)(x \times y)$.

*Si A et B sont deux $\mathbb{K}$-algèbres, une application $f\colon A \to B$ est un **morphisme d'algèbres** entre A et B si pour tout $x, y \in A$ et pour tout $\alpha \in \mathbb{K}$,*

(1) $f(xy) = f(x)f(y)$
(2) $f(x + \alpha y) = f(x) + \alpha f(y)$

où nous avons noté xy pour $x \times y$.

LEMooVKLKooSAHmpZ

Lemme 1.348 ([1])**.**
Soient une algèbre A et une famille $(X_i)_{i\in I}$ de sous-algèbres de A (ici I est un ensemble quelconque). Alors la partie $X = \bigcap_{i\in I} X_i$ est une sous-algèbre de A.

Démonstration. Nous devons prouver que si x et y sont dans X et $\lambda \in \mathbb{K}$, alors xy, $x + y$ et λx sont dans X. Pour tout $i \in I$ nous avons $x, y \in X_i$ et donc $xy \in X_i$, $x + y \in X_i$ et $\lambda x \in X_i$ (parce que X_i est une algèbre). Donc xy, $x + y$ et λx sont dans X_i pour tout I, et donc dans X. □

136. Définition 1.209
137. Définition 1.332.

DefkAXaWY

Définition 1.349.
*L'**algèbre engendrée** par X est l'intersection de toutes les sous-algèbres de A contenant X (qui est une algèbre par le lemme 1.348).*

1.18 Caractéristique d'un anneau

LEMDEFooVEWZooUrPaDw

Lemme-Définition 1.350.
Soit l'application

$$\begin{aligned}\mu\colon \mathbb{Z} &\to A\\ n &\mapsto n\cdot 1_A\end{aligned} \tag{1.481}$$

où $n\cdot 1_A$ signifie $\sum_{k=1}^{n} 1_A$.

(1) C'est un morphisme d'anneaux.

(2) Le noyau est un sous-groupe de $\mathbb{Z}$

(3) Il existe un unique $p\in\mathbb{Z}$ tel que $\ker(\mu)=p\mathbb{Z}$.

*Ce p est la **caractéristique** de A.*

Par exemple la caractéristique de $\mathbb{Q}$ est zéro parce qu'aucun multiple de l'unité n'est nul.
À propos de diagonalisation en caractéristique 2, voir l'exemple 8.223.

Lemme 1.351.
Si A est de caractéristique nulle, alors A est infini.

Démonstration. En effet, $\ker\mu=\{0\}$ implique que $n1_A\neq m1_A$ dès que $n\neq m$ et par conséquent A contient $\mathbb{Z}1_A$, et est infini. □

LemHmDaYH

Lemme 1.352.
Soit un anneau A de caractéristique p.

ITEMooPKSEooPKChGM

(1) Si $p>0$, alors nous avons l'isomorphisme d'anneaux

$$\mathbb{Z}1_A\simeq\mathbb{Z}/p\mathbb{Z}. \tag{1.482}$$

ITEMooBTUIooYzOycc

(2) Si $p=0$, alors nous avons l'isomorphisme d'anneaux

$$\mathbb{Z}1_A\simeq\mathbb{Z} \tag{1.483}$$

Démonstration. Pour (1), l'isomorphisme est donné par l'application $n1_A\mapsto\phi(n)$ si ϕ est la projection canonique $\mathbb{Z}\to\mathbb{Z}/p\mathbb{Z}$.
Pour (2), la preuve est encore à faire. Écrivez-moi. □

PropGExaUK

Proposition 1.353.
La caractéristique d'un anneau fini divise son cardinal.

Démonstration. Si A est un anneau, le groupe $\mathbb{Z}$ agit sur A par

$$n\cdot a=a+n1_A. \tag{1.484}$$

Chaque orbite de cette action est de la forme

$$\mathcal{O}_a=\{a+n1_A \text{ tel que } n=0,\dots,p-1\} \tag{1.485}$$

où p est la caractéristique de A. Les orbites ont p éléments et forment une partition de A, donc le cardinal de A est un multiple de p. □

LEMooJQIKooQgukqn

Lemme 1.354 ([69]).
Un anneau totalement ordonné est de caractéristique nulle.

Démonstration. Le morphisme $\mu\colon \mathbb{Z} \to A$, $n \mapsto n1_A$ est strictement croissant, en particulier $\mu(x) \neq \mu(y)$ dès que $x \neq y$. Donc $\ker(\mu) = \{0\}$. □

L'ensemble typique de caractéristique p est $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

PropFrobHAMkTY

Proposition 1.355.
Soit A un anneau commutatif unitaire de caractéristique p. L'application

$$\begin{aligned} \mathrm{Frob}_A\colon A &\to A \\ x &\mapsto x^p \end{aligned} \tag{1.486}$$

est un automorphisme d'anneau unitaire.

Nous le nommons le **morphisme de Frobenius**. Nous utiliserons aussi les itérés du morphisme de Frobenius : $\mathrm{Frob}^k\colon x \mapsto x^{p^k}$.

Exemple 1.356.
Soit à factoriser $X^p - 1$ dans $\mathbb{F}_p$. Grâce au morphisme de Frobenius, nous avons immédiatement

$$X^p - 1 = (X - 1)^p. \tag{1.487}$$

△

LemCaractIntergernbrcartpre

Lemme 1.357.
La caractéristique[138] *d'un anneau intègre est zéro ou un élément premier*[139].

Démonstration. Si A est intègre, alors $\mathbb{Z}1_A$ est a fortiori intègre. Notons p la caractéristique de A. Si $p = 0$, la preuve est finie ; supposons donc que $p \neq 0$. Alors, l'anneau $\mathbb{Z}/p\mathbb{Z}$ est isomorphe à $\mathbb{Z}1_A$, et est donc intègre. Or, la proposition 1.243 dit que $\mathbb{Z}/p\mathbb{Z}$ est intègre si et seulement si p est premier, ce qui conclut la preuve. □

Exemple 1.358.
Il existe des corps dont la caractéristique n'est pas égale au cardinal (contrairement à ce que laisserait penser l'exemple des $\mathbb{Z}/p\mathbb{Z}$). En effet les matrices $n \times n$ inversibles sur $\mathbb{F}_3$ forment un corps qui n'est pas de cardinal trois alors que la caractéristique est 3 :

$$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = 0. \tag{1.488}$$

△

1.18.1 Caractéristique deux

Beaucoup de résultats demandent une caractéristique différente de deux. Qu'a donc de particulier la caractéristique deux ?

Si $\mathbb{K}$ est un corps de caractéristique 2, alors l'égalité $x = -x$ n'implique pas $x = 0$, puisque $2x = 0$ est vérifiée pour tout x. Cela se répercute sur un certain nombre de résultats. Par exemple, en caractéristique deux, une forme antisymétrique n'est pas toujours alternée : voir le lemme 8.8.

138. Définition 1.350.
139. Définition 1.188.

1.19 Polynômes

1.19.1 Polynômes d'une variable

Et voilà la définition que tout le monde attendait ; la définition des anneaux de polynômes. Pour ne pas taper trop fort du premier coup, nous commençons par les polynômes d'une seule variable[140].

L'ensemble des polynômes sur A sera simplement $A^{(\mathbb{N})}$ (notation 1.337). Puisque $\mathbb{N}$ est un ensemble bien particulier possédant plein de structure, nous allons pouvoir installer sur $A^{(\mathbb{N})}$ une structure non seulement de A-module (ça c'est déjà fait), mais en plus d'anneau, ainsi qu'une évaluation.

DEFooFYZRooMikwEL

Définition 1.359.
L'ensemble des ***polynômes*** *en une indéterminée sur l'anneau A est l'anneau*

$$\mathcal{P}(A) = A^{(\mathbb{N})} \tag{1.489}$$

défini en 1.337.

1.360.
En ce qui concerne la notation $A[X]$, voir 1.19.2. Pour $\mathbb{K}(X)$ lorsque $\mathbb{K}$ est un corps, voir 6.84.

DefDegrePoly

Proposition-Définition 1.361 ([70])**.**
Soit P non nul dans $\mathcal{P}(A)$. Nous notons a_n la valeur[141] *de P en $n \in \mathbb{N}$: $P = (a_n)_{n\in\mathbb{N}}$.*

(1) L'ensemble $\{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est fini dans $\mathbb{N}$.

(2) Cet ensemble possède un minimum et un maximum.

Le ***degré*** *de P est*

$$\deg(P) = \max\{n \in \mathbb{N} \text{ tel que } a_n \neq 0\}, \tag{1.490}$$

et la ***valuation*** *de P est*

$$\mathrm{val}(P) = \min\{n \text{ tel que } a_n \neq 0\}. \tag{1.491}$$

Nous notons

$$\mathcal{P}_n(A) = \{P \in \mathcal{P}(A) \text{ tel que } \deg(P) \leqslant n\}, \tag{1.492}$$

EQooAULWooDgogWT

Dans le cas du polynôme nul, l'ensemble $\{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est vide, et les définitions ne s'appliquent pas. Nous convenons que

$$\mathrm{val}(0) = +\infty \tag{1.493a}$$
$$\deg(0) = -\infty. \tag{1.493b}$$

Démonstration. Le fait que P soit non nul implique que $A = \{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est non vide. De plus cet ensemble est fini parce que $P \in A^{(\mathbb{N})}$. Toute partie finie non vide de $\mathbb{N}$ étant majorée et minorée (lemme 1.65), le lemme 1.66 définit correctement le minimum et le maximum de A. □

Vu que $A^{(\mathbb{N})}$ est engendré par les e_i, tout polynôme sur A s'écrit $P = \sum_{i=1}^{n} a_i e_i$.

DEFooNXKUooLrGeuh

Définition 1.362.
Nous ajoutons deux structures à $A^{(\mathbb{N})}$.

L'évaluation *Si $\alpha \in A$ et si $P \in A^{(\mathbb{N})}$, nous définissons $P(\alpha)$ par*

$$P(\alpha) = (\sum_{i=0}^{n} a_i e_i)(\alpha) = \sum_{i=0}^{n} a_i \alpha^i, \tag{1.494}$$

EQooDJISooTEkMOw

140. Pour les polynômes à plusieurs variables, voir la définition 3.50.

141. Ici il y a une énorme subtilité de terminologie. Formellement, P est une application $\mathbb{N} \to A$. Cela n'a rien à voir avec le fait que P puisse être évalué sur A avec des formule du type $P(x) = \sum_n a_n x^n$. D'ailleurs nous n'avons pas encore vu cette évaluation.

étant entendu que $\alpha^0 = 1$ dans A.

Cette définition s'étend immédiatement au cas où B est un anneau qui étend A. Dans ce cas nous pouvons définir $P(b)$ pour tout $P \in A^{(\mathbb{N})}$ et $b \in B$ avec la même formule (1.494).

Le produit *C'est ici que la structure particulière de $\mathbb{N}$ est utilisée. Nous définissons le produit $A^{\mathbb{N}} \times A^{(\mathbb{N})} \to A^{(\mathbb{N})}$ de la façon suivante. Si $(P_k)_{k\in\mathbb{N}}$ est la suite (presque partout nulle) d'éléments de A qui définit P et si $(Q_k)_{k\in\mathbb{N}}$ est celle de Q, nous notons*

$$(PQ)_n = \sum_{k=0}^{n} P_k Q_{n-k}, \tag{1.495}$$

EQooTNCSooKklish

et donc $PQ = \sum_i (PQ)_i e_i$. Plus explicitement,

$$(\sum_{i=0}^{n} a_i e_i)(\sum_{j=0}^{m} b_j e_j) = \sum_{k=0}^{\infty} \Big(\sum_{\substack{(i,j)\in\mathbb{N}^2 \\ i+j=k}} a_i b_j \Big) e_k. \tag{1.496}$$

EQooCIBUooAzpxiL

Notons qu'à droite, la somme sur k est une somme finie.

PROPooGDQCooHziCPH

Proposition 1.363.
Soit un anneau A. À propos de structure sur $A^{(\mathbb{N})}$.

(1) Avec le produit, l'ensemble $A^{(\mathbb{N})}$ devient un anneau.

(2) L'application

$$\begin{aligned} g\colon A^{(\mathbb{N})} &\to A \\ P &\mapsto P(\alpha) \end{aligned} \tag{1.497}$$

est un morphisme d'anneaux[142]*. En particulier, $(PQ)(\alpha) = P(\alpha)Q(\alpha)$.*

Démonstration. En plusieurs points

(1) **Anneau** L'identité pour le produit dans $A^{(\mathbb{N})}$ est le polynôme donné par $a_0 = 1$ et $a_i = 0$ pour $i \neq 0$. Cela se vérifie en utilisant directement la définition (1.496). La distributivité aussi[143].

(2) **Le morphisme** Nous notons P_k les éléments de la suite définissant P et Q_k ceux de Q. Alors nous avons

$$(P+Q)(\alpha) = \sum_k (P_k + Q_k)\alpha^k = \sum_k P_k\alpha^k + \sum_k Q_k\alpha^k = P(\alpha) + Q(\alpha). \tag{1.498}$$

Vous aurez noté que la première égalité était la définition (1.465b). De même,

$$P(\alpha)Q(\alpha) = \big(\sum_n P_n\alpha^n\big)\big(\sum_k Q_k\alpha^k\big) = \sum_k Q_k\big(\sum_n P_n\alpha^n\big)\alpha^k = \sum_k\sum_n Q_k P_n \alpha^{n+k} \tag{1.499a}$$

$$= \sum_m \big(\sum_{l=0}^{m} P_l Q_{m-l}\big)\alpha^m = \sum_m (PQ)_m\alpha^m = (PQ)(\alpha). \tag{1.499b}$$

□

LEMooWVUXooQlaepO

Lemme 1.364.
Si A est commutatif, alors $A^{(\mathbb{N})}$ est commutatif.

Démonstration. Soient $P, Q \in A^{(\mathbb{N})}$; pour rappel, le produit est donné par la définition 1.495. L'application

$$\begin{aligned} \varphi\colon \{0,\dots,n\} &\to \{0,\dots,n\} \\ k &\mapsto n-k \end{aligned} \tag{1.500}$$

142. Définition 1.43.
143. Je n'ai pas fait les calculs, écrivez-moi pour me dire si ça va facilement.

est une bijection. Voici maintenant le calcul :

$$(PQ)_n = \sum_{k=0}^{n} P_k Q_{n-k} \tag{1.501a}$$
$$= \sum_{k=0}^{n} P_{\varphi(k)} Q_{n-\varphi(k)} \tag{1.501b}$$
SUBEQooISTNooLPySIy
$$= \sum_{k=0}^{n} P_{n-k} Q_k \tag{1.501c}$$
$$= \sum_{k=0}^{n} Q_k P_{n-k} \tag{1.501d}$$
SUBEQooCUMAooFjqqHW
$$= (QP)_n. \tag{1.501e}$$

Justifications

— Pour (1.501b). Lemme 1.309 et le fait que φ soit une bijection.

— Pour (1.501d). Commutativité de A.

□

1.19.2 La notation $A[X]$

SUBSECooLEKVooFBPSJz

Si A est un anneau, nous avons déjà défini les polynômes en une indéterminée sur A comme étant le module $A^{(\mathbb{N})}$ qui est devenu un anneau par la proposition 1.363.

Le polynôme donné par la suite $(a_n)_{n\in\mathbb{N}}$ est souvent notée

$$\sum_k a_k X^k. \tag{1.502}$$

Par exemple avec $a = (4, 2, 8)$ nous avons $a = 8X^2 + 2X + 4$. Nous utiliserons souvent cette notation, qui est très pratique parce qu'elle s'adapte bien aux règles de multiplication et d'addition, en particulier la distributivité.

Il y a (au moins) deux façons de comprendre ce que signifie réellement « X » dans cette notation.

1.19.2.1 Première façon (qui botte en touche)

La première est de dire qu'il n'a pas de significations, et que X^2 est un simple abus de notations pour écrire $(0, 0, 1, 0, \cdots)$. Avec cette façon de voir, nous notons l'anneau des polynômes sur A par « $A[X]$ » où le X n'a pas d'autres raisons d'être que d'avertir le lecteur que nous réservons la lettre « X » pour utiliser la notation pratique des polynômes.

1.19.2.2 Seconde façon (la bonne)

SUBSUBSECooKOLoMPNBMdWvXfBHxg

1.365.

La seconde façon de voir le « X » est de nous rappeler que $A^{(\mathbb{N})}$ a une base en tant que module : les e_k dont nous avons parlé plus haut. Nous posons $X = e_1$, et nous prenons la convention $X^0 = 1$. Alors nous avons $e_k = X^k$ et nous notons $A[X]$ l'anneau $A^{(\mathbb{N})}$ exprimé avec X.

Dans les deux cas, il n'est pas vraiment légitime d'écrire des égalités comme « $P(X) = X^2 + 2X - 3$ », et encore moins de dire « Le polynôme P, *évalué* en X vaut $X^2 + 2X - 3$ » : il est plus correct d'écrire « $P = X^2 + 2X - 3$ ».

Le lemme suivant montre que ces notations tombent vraiment à point. La véritable difficulté de l'énoncé est de comprendre qu'il n'est pas trivial.

Nous avons vu dans la définition 1.362 que si B est un anneau qui étend A, et si $P \in A[X]$, alors nous avons une définition de $P(b)$ pour tout $b \in B$. Nous appliquons cela à $B = A[X]$, qui est un anneau qui étend A. Autrement dit, si P et Q sont des polynômes, ça a un sens d'écrire $P(Q)$ et le résultat sera un élément de $A[X]$.

Dans le cas particulier $Q = X$, nous avons une chouette formule.

LEMooGKWQooVOyeDX

Lemme 1.366.
Nous avons

$$P(X) = P \tag{1.503}$$

pour tout $P \in A[X]$.

Démonstration. Si $P = (a_k)_{k\in\mathbb{N}}$ alors par définition $P(\alpha) = \sum_k a_k\alpha^k$ dès que α est dans un anneau B qui étend A. Nous considérons le cas particulier $B = A[X]$ et $\alpha = X$, c'est-à-dire $Q = (0, 1, 0, \ldots)$, l'élément $P(X)$ de $A[X]$ vaut

$$\sum_k a_k X^k, \tag{1.504}$$

EQooABULooFCEasf

qui est exactement P lui-même. □

Mais il faut bien comprendre que si P est le polynôme $(-3, 2, 1, 0, \ldots)$, noté X^2+2X-3, écrire $P(X) = X^2 + 2X - 3$ est une pirouette de notations que rien ne justifie par rapport à simplement écrire $P = X^2 + 2X - 3$.

1.19.3 Action du groupe symétrique

DefActionGroupe

Définition 1.367 (Thème 9)**.**
*Une **action de groupe** G sur un ensemble E est la donnée, pour chaque élément $g \in G$, d'une fonction $\phi_g : E \to E$, de telle sorte que :*

$$\begin{aligned} \phi_e(x) &= x, & \forall x \in E;\\ \phi_{gh}(x) &= \phi_g(\phi_h(x)), & \forall g, h \in G, \forall x \in E. \end{aligned}$$

*On dit dans ce cas que G **agit** sur E.*

Par souci de notations, nous notons $\mathcal{P}_n(A)$ l'anneau des polynômes de n variables sur A. La propriété universelle de $\mathcal{P}_n(A) = A^{(\mathbb{N}^n)}$ du théorème 1.339 nous donne une application

$$g\colon \operatorname{Fun}\big(\mathbb{N}^n, \mathcal{P}_n(A)\big) \to \operatorname{Hom}_A\big(\mathcal{P}_n(A), \mathcal{P}_n(A)\big) \tag{1.505}$$

Avec cela nous pouvons énoncer et démontrer le lemme qui donne l'action de S_n [144] sur $\mathcal{P}_n(A)$.

LEMooIRVQooHvoNBq

Lemme 1.368 ([71])**.**
Pour $\sigma \in S_n$ nous définissons

$$\begin{aligned} \phi_\sigma\colon \mathbb{N}^n &\to \mathcal{P}_n(A)\\ m &\mapsto e_{\sigma(m)}. \end{aligned} \tag{1.506}$$

Alors l'application

$$\begin{aligned} \rho\colon S_n &\to \operatorname{Hom}_A\big(\mathcal{P}_n(A), \mathcal{P}_n(A)\big)\\ \sigma &\mapsto g(\phi\sigma) \end{aligned} \tag{1.507}$$

est une action [145].

Démonstration. Nous commençons par donner une expression à notre ρ. Un élément de $\mathcal{P}_n(A)$ est de la forme $\sum_{m\in\mathbb{N}^n} a_m e_m$, et nous avons [146]

$$\rho(\sigma)\Big(\sum_{m\in\mathbb{N}^n} a_m e_m\Big) = \sum_m a_m \phi_\sigma(m) = \sum_m a_m e_{\sigma(m)}. \tag{1.508}$$

144. Définition du groupe symétrique S_n en 1.274.
145. Définition 1.367.
146. La somme est définie par 1.309, et ça va être important. Ah oui, en réalité partout, les sommes sont finies parce que les a_m ($m \in \mathbb{N}^n$) sont presque tous nuls. Il faudrait écrire sur la somme sur $\{m \in \mathbb{N}^2 \text{ tel que } a_m \neq 0\}$, mais vous vous imaginez la complication dans la notation.

Nous avons tout de suite $\rho(\mathrm{Id}) = \mathrm{Id}$.

En ce qui concerne la composition, nous avons d'une part

$$\rho(\sigma_1\sigma_2)\big(\sum_m a_m e_m\big) = g(\phi_{\sigma_1\sigma_2})\big(\sum_m a_m e_m\big) = \sum_m a_m e_{\sigma_1\sigma_2(m)}, \tag{1.509}$$

et d'autre part,

$$\begin{aligned}
\rho(\sigma_1)\rho(\sigma_2)\big(\sum_m a_m e_m\big) &= \rho(\sigma_1)\big(\sum_m a_m e_{\sigma_2(m)}\big) && \text{(1.510a)}\\
&= \rho(\sigma_1)\big(\sum_m a_{\sigma_2^{-1}(m)} e_m\big) && \text{(1.510b)}\\
&= \sum_m a_{\sigma_2^{-1}(m)} e_{\sigma_1(m)} && \text{(1.510c)}\\
&= \sum_m a_m e_{\sigma_1\sigma_2(m)} && \text{(1.510d)}
\end{aligned}$$

SUBEQooTSCYooCUWiRz

SUBEQooQPGPooVvqJdT

La proposition 1.310 est utilisée pour (1.510b) et pour (1.510d). □

1.19.4 Corps des fractions

DEFooGJYXooOiJQvP

Proposition-Définition 1.369 ([72]).
Soit un anneau commutatif et intègre[147] *A. Nous posons $E = A \times A\backslash\{0\}$, et nous définissons les deux opérations suivantes sur E :*

ITEMooWBWHooYsXFkO

(1) $(a,b) + (c,d) = (ad + cb, bd)$;

ITEMooGOOIooCHqLRl

(2) $(a,b)(c,d) = (ac, bd)$.

Nous notons a/b la classe de (a,b). Alors

(1) La relation $(a,b) \sim (c,d)$ si et seulement si $ad = bc$ est une relation d'équivalence.

(2) Le quotient

$$\mathrm{Frac}(A) = \big(A \times A\backslash\{0\}\big)/\sim . \tag{1.511}$$

est un corps.

ITEMooYUNEooAKZJEh

(3) L'unité dans $\mathrm{Frac}(A)$ *est la classe* $1/1$.

ITEMooOGPCooAZGWoF

(4) L'opposé de a/b est $(-a)/b$.

ITEMooECDBooJdJEza

(5) Si $a \neq 0$, alors l'inverse de a/b est b/a.

Le corps $\mathrm{Frac}(A)$ *est nommé le* ***corps des fractions*** *de A.*

Lorsque A est un anneau de polynômes[148]*, alors les éléments de* $\mathrm{Frac}(A)$ *sont des* ***fractions rationnelles****.*

Le fait que A soit intègre est important pour être certain que $bd \neq 0$ sous l'hypothèse que $b, d \neq 0$.

PROPooUULNooKbwuEw

Proposition 1.370.
Soit un anneau commutatif A. Nous considérons l'application

$$\begin{aligned} i\colon A &\to \mathrm{Frac}(A)\\ a &\mapsto a/1. \end{aligned} \tag{1.512}$$

(1) Elle est injective.

(2) Elle est un morphisme d'anneaux[149].

147. Définition 1.197.
148. Définition 1.359.
149. Définition 1.43.

Démonstration. En plusieurs parties.

(1) **Injective** Supposons que $f(a) = f(b)$, c'est-à-dire que $a/1 = b/1$. En vertu de la relation d'équivalence donnée en 1.369, nous avons $a1 = b1$, c'est-à-dire $a = b$.

(2) **Morphisme d'anneaux** D'abord nous avons

$$\begin{aligned} i(a) + i(b) &= (a/1) + (b + 1) && \text{(1.513a)}\\ &= (a1 + 1b)/1 && \text{(1.513b)}\\ &= i(a + b). && \text{(1.513c)} \end{aligned}$$

Ensuite,

$$i(a)i(b) = (a/1)(b/1) = ab/1 = i(ab). \tag{1.514}$$

Et enfin $i(1) = 1/1$ est bien l'unité de $\mathrm{Frac}(A)$ par 1.369(3).

□

NORMooFBKEooBQuATg

1.371.
Si $a \in A$, nous nous permettrons de noter également a l'élément $a/1$ de $\mathrm{Frac}(A)$. Ce sera toujours clair dans le contexte.

NORMooVYEMooUiUfnE

1.372.
En vertu de la proposition 1.370, à partir de maintenant, nous allons identifier la partie $i(\mathbb{Z})$ à $\mathbb{Z}$. Nous nous autorisons donc à dire que $4 \in \mathbb{Q}$ ou que $-7 \in \mathbb{Q}$, et même que $0 \in \mathbb{Q}$.

PROPooTXGUooRvViwT

Proposition 1.373 (Règle de simplification de fraction[1])**.**
Soit un anneau commutatif A. Si $a, x, y \in A$ avec $a, y \neq 0$, nous avons

(1)

$$ax/a = x/1. \tag{1.515}$$

Si vous avez suivi 1.371, cela signifie que $ax/a = x$ dans $\mathrm{Frac}(A)$.

(2)

$$ax/ay = x/y \tag{1.516}$$

dans $\mathrm{Frac}(A)$.

LEMooBJRCooIZnaid

Lemme 1.374.
Soient un anneau A ainsi que $a, b, x, y \in A$. Si $b, y \neq 0$, nous avons $ay = bx$ dans A si et seulement si nous avons $ab^{-1} = xy^{-1}$ dans $\mathrm{Frac}(A)$ [150].

Démonstration. Deux parties.

(1) $\Rightarrow$ Nous supposons avoir $ay = bx$ dans A. En passant aux classes, cela donne l'égalité $ay/1 = bx/1$ dans $\mathrm{Frac}(A)$. Nous multiplions les deux membres par $(b/1)^{-1}(y/1)^{-1}$. En utilisant les différents points de la définition 1.369, et en particulier (5), nous avons$(b/1)^{-1}(y/1)^{-1} = 1/by$. Nous avons donc

$$(ay/1)(1/by) = (bx/1)(1/by). \tag{1.517}$$

En utilisant les relations de simplification de la proposition 1.373, nous avons $ay/by = a/b$ à gauche et $bx/by = x/y$ à droite. Bref, nous avons

$$a/b = x/y. \tag{1.518}$$

Le membre de gauche peut être récrit $a/b = (a/1)(b/1)^{-1}$, ce qui peut s'écrire ab^{-1}.

150. Rappel de 1.372 : si $a \in A$, nous nommons $a \in \mathrm{Frac}(A)$ l'élément $a/1$, c'est-à-dire la classe de $(a, 1)$

$\Leftarrow$ L'hypothèse s'écrit, sans abus de notations sous la forme

$$(a/1)(b/1)^{-1} = (x/1)(y/1)^{-1}. \tag{1.519}$$

En multipliant les deux côtés par $(b/1)(y/1)$ nous trouvons

$$ay/1 = xb/1. \tag{1.520}$$

Nous pouvons écrire cela explicitement avec la relation d'équivalence qui définit $\text{Frac}(A)$ (définition 1.369) :

$$(ay, 1) \sim (xb, 1), \tag{1.521}$$

c'est-à-dire $ay1 = 1xb$, ce que nous voulions. □

La proposition suivante montre encore que le corps des fractions est le plus petit corps que l'on puisse imaginer à partir d'un anneau.

PROPooIJBEooDjsoHr

Proposition 1.375 ([13, 1])**.**
Soit un anneau commutatif A. Tout corps commutatif contenant un sous-anneau isomorphe[151] *à A contient un sous-corps isomorphe à* $\text{Frac}(A)$.

Démonstration. Soit un corps $\mathbb{K}$ contenant un sous-anneau A' isomorphe à A. Nous notons $\sigma : A' \to A$ un isomorphisme d'anneaux entre A' et A.

(1) **Une partie bien choisie**
Nous considérons la partie suivante de $\mathbb{K}$:

$$S = \{ab^{-1} \text{ tel que } a, b \in A'\}. \tag{1.522}$$

(2) **S est un corps** Deux éléments arbitraires de S sont ab^{-1} et xy^{-1}. Nous devons prouver plusieurs choses.

(2a) **Neutres** En prenant $a = b = 1$ nous avons $ab^{-1} = 1 \in S$. En prenant $a = 0$ et $b = 1$ nous avons $ab^{-1} = 0 \in S$.

(2b) **Somme** Il faut remarquer que $ab^{-1} + xy^{-1} = (ay + xb)(by)^{-1}$. En effet,

$$\begin{aligned}(ay + xb)(by)^{-1} &= (ay + xb)y^{-1}b^{-1} && \text{(1.523a)}\\ &= ayy^{-1}b^{-1} + xby^{-1}b^{-1} && \text{(1.523b)}\\ &= ab^{-1} + xy^{-1} && \text{(1.523c)}\end{aligned}$$

SUBEQooRGPSooXaBGyx
SUBEQooOHJGooWrfPow

Justifications :

— Pour (1.523b). Distributivité.

— Pour (1.523c). Commutativité dans A.

(2c) **Produit** Il s'agit du même genre de calculs en utilisant les mêmes propriétés. Nous avons

$$(ab^{-1})(xy^{-1}) = (ax)(by)^{-1}. \tag{1.524}$$

(3) **Ce qui va être notre isomorphisme**
Ensuite nous montrons que l'application

$$\begin{aligned}\varphi : S &\to \text{Frac}(A)\\ ab^{-1} &\mapsto \sigma(a)/\sigma(b)\end{aligned} \tag{1.525}$$

est bien définie et est un isomorphisme de corps.

151. Morphisme d'anneaux, définition 1.43.

(4) **Bien définie**
Si $ab^{-1} = xy^{-1}$ alors $ay = xb$. Puisque σ est un isomorphisme nous avons aussi $\sigma(a)\sigma(y) = \sigma(x)\sigma(b)$ et donc $\sigma(a)/\sigma(b) = \sigma(x)/\sigma(y)$ par définition des classes de $\text{Frac}(A)$.

(5) **Morphisme** Deux éléments arbitraires de S sont ab^{-1} et xy^{-1}. Calculons un peu :

$$\begin{aligned}
\varphi\big((ab^{-1})(xy^{-1})\big) &= \varphi(axy^{-1}b^{-1}) && \text{(1.526a)}\\
&= \varphi\big((ax)(by)^{-1}\big) && \text{(1.526b)}\\
&= \sigma(ax)/\sigma(by) && \text{(1.526c)}\\
&= \big(\sigma(a)/\sigma(b)\big)\big(\sigma(x)/\sigma(y)\big) && \text{(1.526d)}\\
&= \varphi(ab^{-1})\varphi(xy^{-1}). && \text{(1.526e)}
\end{aligned}$$

Justifications :
— Pour (1.526a). Commutativité dans A.
— Pour (1.526b). Associativité dans A.
— Pour (1.526d). Définition 1.369(2) de la multiplication de fractions.

(6) **Surjectif**
Tout élément de $\text{Frac}(A)$ est de la forme a'/b' avec $a', b' \in A$, et donc de la forme $\sigma(a)/\sigma(b)$ avec $a, b \in A'$. Un tel élément est l'image par φ de $ab^{-1} \in S$.

(7) **Injectif**
Si $\varphi(ab^{-1}) = \varphi(xy^{-1})$ alors $\sigma(a)/\sigma(b) = \sigma(x)/\sigma(y)$, et par définition des classes nous avons $\sigma(a)\sigma(y) = \sigma(b)\sigma(x)$. De là nous avons $\sigma(ay) = \sigma(bx)$ et donc $ay = bx$ (parce que σ est un isomorphisme). Nous en déduisons que $ab^{-1} = xy^{-1}$.

□

1.376.
Soit un anneau A et son anneau des polynômes $\mathcal{P}(A)$. Si $\alpha \in A$, nous avons la définition 1.362 qui donne l'évaluation $P(\alpha)$.

Si par contre P et Q sont des polynômes sur A, nous n'avons pas encore défini ce que serait l'évaluation de la fraction rationnelle P/Q en α. Nous comblons à présent ce manque.

Définition 1.377 (Évaluation d'une fraction rationnelle)**.**
Soit un corps $\mathbb{K}$ contenant l'anneau A. Si $R = P/Q \in \text{Frac}(A)$ et si $\alpha \in \mathbb{K}$ nous définissons[152]

$$R(\alpha) = (P/Q)(\alpha) = P(\alpha)Q^{-1}(\alpha). \tag{1.527}$$

Dans cette formule, les polynômes, l'inverse et le produit sont calculés dans $\mathbb{K}$ et non dans A.

Théorème-Définition 1.378.
Soit A un anneau commutatif intègre.

(1) Il existe un couple $(\mathbb{K}, \epsilon)$ où $\mathbb{K}$ est un corps commutatif et $\epsilon\colon A \to \mathbb{K}$ est un morphisme injectif d'anneaux tels que pour tout $\lambda \in \mathbb{K}$, il existe $(a, b) \in A \times A^*$ tels que

$$\lambda = \epsilon(a)\big(\epsilon(b)\big)^{-1} \tag{1.528}$$

*(2) Si $(\mathbb{K}', \epsilon')$ est un autre couple qui vérifie la propriété, les corps $\mathbb{K}$ et $\mathbb{K}'$ sont isomorphes. Le corps $\mathbb{K}$ associé à l'anneau A est le **corps des fractions** de A, et sera noté* $\text{Frac}(A)$.

(3) Nous posons

$$\begin{aligned}
\sigma\colon A \times A^* &\to \mathbb{K}\\
(a, b) &\mapsto \epsilon(a)\big(\epsilon(b)\big)^{-1}.
\end{aligned} \tag{1.529}$$

Nous avons

$$\sigma(xa, xb) = \sigma(a, b) \tag{1.530}$$

pour tout $a, b, x \in A$.

152. Les fractions rationnelles, définition 1.369.

1.19.5 Corps totalement ordonné

DefKCGBooLRNdJf

Définition 1.379.
Ordre et choses reliées dans un corps.

ITEMooOOOVooJWwIQr

(1) *Un corps* $\mathbb{K}$ *est* ***totalement ordonné*** *si il existe une relation d'ordre total*[153] *tel que*

ITEMooZI3JooWNxnBj

(1a) $x \leqslant y$ *implique* $x + z \leqslant y + z$ *pour tout* $x, y, z \in \mathbb{K}$

CONDooBYYDooElXgPO

(1b) $x \geqslant 0$ *et* $y \geqslant 0$ *implique* $xy \geqslant 0$.

ItemooWUGSooRSRvYC

(2) *Si* $\mathbb{K}$ *est un corps totalement ordonné, nous y définissons la valeur absolue par*

EQooNONAooHLSERO

$$|x| = \begin{cases} x & si\ x \geqslant 0 \\ -x & si\ x \leqslant 0. \end{cases} \tag{1.531}$$

ItemVXOZooTYpcYN

(3) *La suite* (x_n) *dans le corps totalement ordonné* $\mathbb{K}$ *est* ***de Cauchy*** *si pour tout* $\epsilon \in \mathbb{K}^+$, *il existe* $N \in \mathbb{N}$ *tel que si* $p, q \geqslant N$ *alors* $|x_p - x_q| \leqslant \epsilon$.

ITEMooDERQooLmJwFR

(4) *La suite* (x_n) *dans le corps totalement ordonné* $\mathbb{K}$ *est* ***convergente*** *si il existe* $q \in \mathbb{K}$ *tel que pour tout* $\epsilon \in \mathbb{K}^+$, *il existe* N *tel que si* $k \geqslant N$ *alors* $|x_k - q| \leqslant \epsilon$.

ITEMooKZZYooDaidGU

(5) *Un corps totalement ordonné est* ***complet*** *si toute suite de Cauchy y est convergente.*

ITEMooMWASooEzhVyh

(6) *Si* $a, \epsilon \in \mathbb{K}$ *avec* $\epsilon > 0$ *alors nous définissons la* ***boule ouverte*** *de centre* a *et de rayon* ϵ *par*

$$B(a, \epsilon) = \{x \in \mathbb{K} \text{ tel que } |a - x| < \epsilon\}, \tag{1.532}$$

et la ***boule fermée*** *par*

$$\overline{B(a, \epsilon)} = \{x \in \mathbb{K} \text{ tel que } |a - x| \leqslant \epsilon\}. \tag{1.533}$$

Lemme 1.380.
Une suite (x_k) *converge vers* q *si et seulement si pour tout* $\epsilon > 0$, *il existe* $N > 0$ *tel que* $x_k \in B(q, \epsilon)$ *pour tout* $k \geqslant N$.

Démonstration. Il s'agit de mettre côte à côte les points (4) et (6) de la définition 1.379. □

1.381.
Ces boules prendront une nouvelle force avec le super-théorème 7.117.

Parmi ces définitions, celles de suite convergente, de Cauchy et de corps complet seront utilisées dans le cas de $\mathbb{Q}$ (et de $\mathbb{R}$ pour la complétude). Elles seront prouvées être équivalentes aux définitions topologiques dans le cas particulier de $\mathbb{R}$ et $\mathbb{Q}$ lorsque la topologie métrique sera définie. Dans cet état d'esprit nous n'allons pas démontrer tout de suite que $\mathbb{R}$ est un corps complet. Nous allons directement démontrer que c'est un espace topologique complet.

LEMooXJTAooZauchx

Lemme 1.382 (Règle des signes[73])**.**
Soit un corps totalement ordonné $\mathbb{K}$ *ainsi que* $x, y \in \mathbb{K}$. *Nous avons :*

(1) *Si* $x \leqslant 0$ *et* $y \leqslant 0$ *alors* $xy \geqslant 0$.

(2) *Si* $x \leqslant 0$ *et* $y \geqslant 0$ *alors* $xy \leqslant 0$.

(3) *Si* $x \geqslant 0$ *et* $y \leqslant 0$ *alors* $xy \leqslant 0$.

ITEMooRGYAooCUIfss

(4) $0 \leqslant 1$.

ITEMooMRNHooLglPKn

(5) *Si* $x \geqslant 0$ *alors* $x^{-1} \geqslant 0$.

LemooANTJooYxQZDw

Lemme 1.383 (Propriétés de la valeur absolue)**.**
Soit $\mathbb{K}$ *un corps totalement ordonné. Si* $x, y \in \mathbb{K}$ *alors*[154]

153. Définition 1.12.
154. La « valeur absolue » est définie en (1.379)(2).

ItemooNVDIooSuiSoB

(1) *Si* $x \geqslant 0$ *alors* $-x \leqslant 0$.

ITEMooVNAZooSxmtuH

(2) *Si* $x \leqslant 0$ *alors* $-x \geqslant 0$.

ITEMooSDNHooDnjScE

(3) $|x| \geqslant 0$

ITEMooLQLTooTJTPVM

(4) $|x| = 0$ *si et seulement si* $x = 0$

ITEMooVJAEooOEatzY

(5) $|-x| = |x|$.

ItemooOMKNooRlanvk

(6) $|x + y| \leqslant |x| + |y|$.

ITEMooEFMLooYVCuHD

(7) $|xy| = |x||y|$

Démonstration. Point par point

(1) **(1)** Nous partons de $x \geqslant 0$ et nous ajoutons $-x$ des deux côtés en profitant de la définition d'un corps totalement ordonné : $x - x \geqslant -x$ et donc $0 \geqslant -x$, c'est-à-dire $-x \leqslant 0$.

(2) **(2)** Nous partons de $x \leqslant 0$ et nous ajoutons $-x$ des deux côtés.

(3) **(3)** Si $x \geqslant 0$ alors c'est vrai. Sinon, $x \leqslant 0$ et $|x| = -x \geqslant 0$ par le point (1).

(4) **(4)** Si $x = 0$ alors $x = -x$ et $|x| = 0$. Au contraire si $x \neq 0$ alors $-x \neq 0$ et que x soit positif ou négatif, nous aurons toujours $\pm x \neq 0$.

(5) **(5)** Il faut décomposer en deux cas selon que $x \geqslant 0$ et $x \leqslant 0$. Supposons $x \geqslant 0$. Alors d'une part $|x| = x$. D'autre part $-x \leqslant 0$ par le point (1), de telle sorte que

$$|-x| = -(-x) = x. \tag{1.534}$$

Nous avons donc $|x| = |-x| = x$.

Le même raisonnement tient avec $x \leqslant 0$.

(6) **(6)** Nous supposons que $x \leqslant y$ et nous distinguons divers cas suivant la positivité de x et y.

(6a) Si $x, y \geqslant 0$. Dans ce cas, $x + y \geqslant y \geqslant 0$, donc $|x + y| = x + y = |x| + |y|$.

(6b) Si $x, y \leqslant 0$. Dans ce cas, $x + y \leqslant 0$ et nous avons $|x + y| = -x - y = |x| + |y|$.

(6c) Si $x \leqslant 0$ et $y \geqslant 0$. Nous subdivisons encore en deux cas suivant que $x + y$ est positif ou négatif. Si $x + y \geqslant 0$, alors nous écrivons successivement

$$x \leqslant 0 \tag{1.535a}$$

$$x + y \leqslant y \leqslant y + |x| = |x| + |y| \tag{1.535b}$$

et donc $|x + y| = x + y \leqslant |x| + |y|$.

Nous supposons à présent que $x \leqslant 0$, $y \geqslant 0$ et $x + y \leqslant 0$. Dans ce cas il suffit d'écrire $|x + y| = |(-x) + (-y)|$ pour retomber dans le cas précédent à inversion près de x et y.

(7) **Pour (7)** Il suffit de prendre les 4 cas suivant les signes de x et y, et d'utiliser les règles de signes du lemme 1.382 dans la définition 1.531.

□

RemooJCAUooKkuglX

Remarque 1.384.
La partie (6) est très importante parce que c'est elle qui fera presque toutes les majorations dont nous aurons besoin en analyse. En effet elle donne l'inégalité triangulaire de la façon suivante : si $x, y, z \in \mathbb{K}$ nous avons

$$|x - y| = |(x - z) + (z - y)| \leqslant |x - z| + |z - y|. \tag{1.536}$$

LEMooQXDCooPEABBm

Lemme 1.385 (À propos de boules)**.**
Soient un corps totalement ordonné $\mathbb{K}$ *et des éléments* $x, y \in \mathbb{K}$. *Soit aussi* $\epsilon > 0$ *dans* $\mathbb{K}$. *Nous avons :*

ITEMooXJGVooSebiip

(1) $y \in B(x,\epsilon)$ *si et seulement si* $x-\epsilon < y < x+\epsilon$.

ITEMooRUBBooRayiMs

(2) *Si* $y \in \overline{B(x,\epsilon)}$ *alors* $y \in B(x,\epsilon')$ *pour tout* $\epsilon' > \epsilon$.

Démonstration. Pour rappel,

$$|x-y| = \begin{cases} x-y & \text{si } x-y \geqslant 0 \\ y-x & \text{si } x-y \leqslant 0. \end{cases} \tag{1.537}$$

Nous pouvons maintenant démontrer nos assertions.

(1) **(1)** En deux parties.

(1a) $\Rightarrow$ Nous supposons que $|x-y| < \epsilon$.
Si $x-y \geqslant 0$ alors l'hypothèse signifie $x-y < \epsilon$, ce qui donne $y > x-\epsilon$. Mais l'inégalité $x-y \geqslant 0$ donne également $x \geqslant y$ et donc $x+\epsilon \geqslant y+\epsilon > y$. Notez le jeu de l'inégalité non stricte qui se change en inégalité stricte.
Si $x-y \leqslant 0$ nous pouvons faire le même raisonnement.

(1b) $\Leftarrow$ Des inégalités $x-\epsilon < y$ et $y < x+\epsilon$ nous tirons $x-y < \epsilon$ et $y-x < \epsilon$. Donc quel que soit le signe de $x-y$ nous avons toujours $|x-y| < \epsilon$.

(2) **(2)**
C'est immédiat parce que

$$|x-y| \leqslant \epsilon < \epsilon'. \tag{1.538}$$

□

LEMooVZNCooRJatKK

Lemme 1.386.
Tout corps totalement ordonné est de caractéristique nulle.

1.20 Les rationnels

Note : pour l'existence et l'unicité de l'écriture $q = k/d$, il faut voir le théorème 3.33.
Une construction très explicite est faite dans [18]. Ici nous allons prendre plus court :

Définition 1.387.
Le corps des fractions de $\mathbb{Z}$ *(définition 1.369) est noté* $\mathbb{Q}$ *et ses éléments sont les* ***rationnels****.*

La proposition suivante provient de 1.370.

PROPooKUIYooCxMHTX

Proposition 1.388 ([1])**.**
L'application

$$\begin{aligned} i\colon \mathbb{Z} &\to \mathbb{Q} \\ a &\mapsto a/1 \end{aligned} \tag{1.539}$$

est un morphisme injectif.

PROPooDHIAooZysvNs

Proposition 1.389.
L'ensemble des rationnels est infini dénombrable[155]*.*

Démonstration. L'ensemble $\mathbb{Z}$ est infini[156] et la proposition 1.370 donne une injection $f\colon \mathbb{Z} \to \mathbb{Q}$. Donc $f(\mathbb{Z})$ est infini.

L'ensemble $\mathbb{Q}$ contient une partie infinie. Il est donc infini par le lemme 1.119.

155. Ouais, je sais, dans les définitions prises ici, un ensemble dénombrable est toujours infini. Mais l'excès de précision ne tue pas, loin s'en faut.
156. Lemme 1.114.

L'application

$$\begin{aligned} g\colon \mathbb{Z}^2 &\to \mathbb{Q} \\ a,b &\mapsto a/b \end{aligned} \tag{1.540}$$

est surjective alors que $\mathbb{Z}^2$ est dénombrable. Le lemme 1.130 dit alors que $\mathbb{Q}$ est fini ou dénombrable. Vu que nous avons déjà prouvé que $\mathbb{Q}$ était infini, nous déduisons que $\mathbb{Q}$ est infini dénombrable. □

1.20.1 Relation d'ordre

PROPooEBCHooYEQkIO

Proposition 1.390 ([1]).
Soient $a,b\in\mathbb{Z}$ tels que $ab\geqslant 0$[157]*. Alors si $(x,y)\sim(a,b)$, nous avons $xy\geqslant 0$.*

Démonstration. L'hypothèse $(x,y)\sim(a,b)$ dit que $ab=ya$. L'astuce est de calculer $(xy)(bb)$ en nous servant de $xb=ya$:

$$(xy)b^2=yayb=aby^2. \tag{1.541}$$

Vu que $ab\geqslant 0$ et $y^2\geqslant 0$, le membre de droite est positif (proposition 1.111). Nous avons donc aussi $xyb^2\geqslant 0$. La proposition 1.111(4) dit encore que dans ce cas $xy\geqslant 0$. □

PROPooEOZWooLxRPXa

Proposition 1.391 ([1]).
La partie

$$\mathbb{Q}^+=\{a/b \text{ tel que } ab>0\}\subset\mathbb{Q} \tag{1.542}$$

est bien définie et est une positivité[158]

Démonstration. Le fait que ce soit bien défini est la proposition 1.390.

(1) Supposons pour le plaisir de la contradiction que $a/b\in\mathbb{Q}^+$ et $-(a/b)\in\mathbb{Q}^+$. En utilisant la propriété 1.369(4) nous avons $ab>0$ et $(-a)b>0$. En utilisant 1.106(1) nous avons

$$\begin{cases} (ab)>0 & \text{(1.543a)}\\ -(ab)>0. & \text{(1.543b)} \end{cases}$$

En utilisant 1.111(6) nous trouvons $ab(-ab)>0$. En utilisant encore une des nombreuses propriétés déjà citées, nous trouvons $-a^2b^2>0$, c'est à dire

$$a^2b^2<0, \tag{1.544}$$

ce qui est faux parce que $a^2b^2\geqslant 0$.

(2) Supposons que $a/b\in\mathbb{Q}^+$ et $x/y\in\mathbb{Q}^+$ (b et y sont non nuls). Cette hypothèse signifie que $ab>0$ et $xy>0$. Nous devons prouver que $(a/b)+(x/y)\in\mathbb{Q}^+$. Nous avons

$$\frac{a}{b}+\frac{x}{y}=\frac{ay+bx}{by}. \tag{1.545}$$

Nous avons un petit calcul à faire :

EQooLGUJooYAUXPJ

$$by(ay+bx)=aby^2+b^2xy. \tag{1.546}$$

Par hypothèse, $ab>0$. Nous savons aussi que $y^2\geqslant 0$. Mais dans notre cas nous avons $y^2\neq 0$. Donc $aby^2>0$. Même raisonnement pour b^2xy. Bref ce que nous avons dans (1.546) est strictement positif.

□

157. Définition 1.109 pour l'ordre dans $\mathbb{Z}$.
158. Définition 1.204.

DEFooZEXXooUtOhqB

Proposition-Définition 1.392 (Ordre sur $\mathbb{Q}$[18]).
Nous considérons l'ensemble

$$\mathbb{Q}^+ = \{a/b \text{ tel que } ab > 0\} \subset \mathbb{Q} \tag{1.547}$$

qui est une positivité sur $\mathbb{Q}$ par la proposition 1.391. Pour $\alpha, \beta \in \mathbb{Q}$ nous disons $\alpha \leqslant \beta$ si et seulement si

$$\beta - \alpha \in \mathbb{Q}^+ \cup \{0\}. \tag{1.548}$$

Cela définit un ordre total par la proposition 1.205.

PROPooXNBVooIvBmkW

Proposition 1.393 ([1]).
L'application

$$\begin{aligned} i\colon \mathbb{Z} &\to \mathbb{Q} \\ a &\mapsto a/1 \end{aligned} \tag{1.549}$$

est un morphisme injectif qui respecte l'ordre.

Démonstration. Le fait que ce soit un morphisme injectif est la proposition 1.388.

Le respect de l'ordre doit encore être fait [159]. □

1.394.
Le lemme suivant 1.397 dit que tout rationnel est majoré par un naturel. Vu que les naturels et les rationnels ne sont pas du tout les mêmes ensembles (les entiers sont des classes de naturels et les rationnels des classes d'entiers), l'énoncé n'a à priori pas de sens.

La réalité est que nous identifions $\mathbb{N}$ à une partie de $\mathbb{Q}$ via les propositions 1.108 et 1.393. L'énoncé exact serait que si $q \in \mathbb{Q}$, il existe $n \in \mathbb{N}$ tel que $i\big(\iota(n)\big) > q$.

PROPooTEREooZgblWu

Proposition 1.395 ([1]).
Si $q > 0$ dans $\mathbb{Q}$, alors il existe $a, b \in \mathbb{N}$ tels que $q = a/b$.

Démonstration. Posons $q = a/b$ avec $a, b \in \mathbb{Z}$. Si $a, b > 0$, c'est bon. Si $a = 0$ alors $q = 0 = 0/1$ c'est bon aussi. Supposons que ce ne soit pas le cas. L'hypothèse $q > 0$ dit que $q = a/b$ avec $ab > 0$. La proposition 1.111(5) dit qu'il n'y a que deux possibilités : soit $a, b > 0$, soit $a, b < 0$.

Supposons que $a, b < 0$. En vertu des classes d'équivalence, nous avons $a = a/b = (-a)/(-b)$. Vu que nous avons $-a > 0$ et $-b > 0$, c'est bon. □

PROPooPSGBooLuMXoI

Proposition 1.396 ([1]).
Soit $a > 0$ et $b > 1$. Alors $a > a/b$.

Démonstration. Pour prouver que $a > a/b$, nous devons calculer

$$a - \frac{a}{b} = \frac{a}{1} - \frac{a}{b} = \frac{ab - a}{b}. \tag{1.550}$$

Nous devons donc évaluer la quantité $b(ab-a)$. Vu que $b > 1$, le lemme 1.110 dit que $ab > a$. Donc $ab - a > 0$. Le produit $b(ab - a)$ est donc le produit de deux entiers strictement positifs. En vertu de la proposition 1.111(6) nous avons donc $b(ab - a) > 0$, et donc $a > a/b$. □

LEMooEBTIooGMoHsj

Lemme 1.397.
Tout rationnel est strictement majoré par un naturel.

Démonstration. Si $q < 0$, alors $n = 0$ fait l'affaire. Si $q = 0$, alors $n = 1$ fonctionne. Supposons donc que $q = a/b > 0$ avec $ab > 0$ ainsi que $a, b \in \mathbb{N}$ [160]. Si $b = 1$, alors $n = a + 1$ fonctionne.

Si $b > 1$, alors la proposition 1.396 dit que $a > a/b$, et donc $n = a$ fonctionne. □

159. Rédigez une preuve, faites un scan de votre feuille et envoyez-moi.
160. Existe pas la proposition 1.395.

PROPooBTCCooVVvaeL

Proposition 1.398.
Si $q < 1$ dans $\mathbb{Q}$, alors $ql < l$ pour tout $l \in \mathbb{Q}^+$.

Démonstration. Si $q < 0$ c'est facile. Supposons que $0 < q < 1$. Vu que $q - 1 < 0$ et $l > 0$, nous avons

$$ql - l = (q - 1)l < 0. \tag{1.551}$$

Mais dire $ql - l < 0$ signifie $ql < l$. □

PROPooMXGPooDUkOuv

Proposition 1.399.
Le corps $\mathbb{Q}$ est archimédien[161].

LEMooMUYAooDLgDcf

Lemme 1.400.
Si $q \in \mathbb{Q}$, alors il existe $k \in \mathbb{N}$ tel que $kq \in \mathbb{Z}$.

1.20.2 Caractéristique

LEMooYCPUooNxEPhB

Lemme 1.401.
Le corps $\mathbb{Q}$ est de caractéristique[162] *nulle.*

1.21 Suite de Cauchy dans un corps totalement ordonné

LEMooLTBIooSZnvsQ

Lemme 1.402 ([69, 1])**.**
Tout corps commutatif de caractéristique nulle contient un sous-corps isomorphe à $\mathbb{Q}$.

Démonstration. Soit un corps $\mathbb{K}$ de caractéristique nulle. Nous savons du lemme 1.350 que

$$\begin{aligned} \mu\colon \mathbb{Z} &\to \mathbb{K} \\ n &\mapsto n1_{\mathbb{K}} \end{aligned} \tag{1.552}$$

est un morphisme d'anneaux vérifiant $\ker(\mu) = \{0\}$. Nous posons $Z = \mu(\mathbb{Z})$. L'application $\mu\colon \mathbb{Z} \to Z$ est un isomorphisme d'anneaux. Prouvons cela :

(1) **Morphisme** L'application μ est un morphisme par le lemme 1.350.

(2) **Surjectif** Par définition les éléments de Z sont dans l'image de $\mathbb{Z}$.

(3) **Injectif** Si $x, y \in \mathbb{Z}$ vérifient $\mu(x) = \mu(y)$, alors $\mu(x - y) = 0$ parce que μ est un morphisme. Mais $\mathbb{K}$ est de caractéristique nulle, c'est-à-dire $\ker(\mu) = \{0\}$. Donc $x - y = 0$.

Le corps $\mathbb{K}$ contient donc un sous-anneau isomorphe à $\mathbb{Z}$. Puisque $\mathbb{Z}$ et $\mathbb{K}$ sont commutatifs, la proposition 1.375 s'applique et $\mathbb{K}$ contient un sous-corps isomorphe à $\text{Frac}(\mathbb{Z}) = \mathbb{Q}$. □

La proposition suivante donne des précisions à propos du lemme 1.402.

PROPooKNROooFdgIeQ

Proposition 1.403 ([1])**.**
Soit un corps totalement ordonné $\mathbb{K}$. Nous considérons l'application

$$\begin{aligned} \mu\colon \mathbb{Z} &\to \mathbb{K} \\ n &\mapsto n \cdot 1_{\mathbb{K}} \end{aligned} \tag{1.553}$$

et ensuite

$$\begin{aligned} \sigma\colon \mathbb{Q} &\to \mathbb{K} \\ a/b &\mapsto \mu(a)\mu(b)^{-1}. \end{aligned} \tag{1.554}$$

Alors

161. Définition 1.73.
162. Définition 1.350.

(1) L'application σ est bien définie.

(2) L'application σ est un morphisme de corps.

(3) Si $q \leqslant q'$ dans $\mathbb{Q}$, alors $\sigma(q) \leqslant \sigma(q')$.

Démonstration. En plusieurs morceaux.

(1) **σ est bien définie** Montrons que σ est bien définie. Pour cela nous considérons $a, b, x, y \in \mathbb{Z}$ tels que $a/b = x/y$ dans $\mathbb{Q}$. Par définition des classes (définition 1.369 du corps des fractions), nous avons $ay = bx$ dans $\mathbb{Q}$. Vu que μ est un morphisme nous avons alors

$$\mu(a)\mu(y) = \mu(b)\mu(x) \tag{1.555}$$

et donc $\mu(a)\mu(b)^{-1} = \mu(x)\mu(y)^{-1}$, c'est-à-dire $\sigma(a/b) = \sigma(x/y)$. L'application σ est donc bien définie.

(2) **Morphisme pour la somme**
L'application μ est un morphisme d'anneaux, comme déjà dit depuis le lemme 1.350. Notons aussi que, parce que $\mathbb{K}$ est commutatif,

$$\mu(qy)^{-1} = \mu(q)^{-1}\mu(y)^{-1}. \tag{1.556}$$

En utilisant la définition 1.369(1) de la somme nous avons

$$\begin{aligned}
\sigma(p/q + x/y) &= \sigma\big((py + qx)/qy\big) && \text{(1.557a)}\\
&= \big[\mu(py) + \mu(qx)\big]\mu(qy)^{-1} && \text{(1.557b)}\\
&= \mu(py)\mu(qy)^{-1} + \mu(qx)\mu(qy)^{-1} && \text{(1.557c)}\\
&= \mu(p)\mu(q)^{-1} + \mu(x)\mu(y)^{-1} && \text{(1.557d)}\\
&= \sigma(p/q) + \sigma(x/y). && \text{(1.557e)}
\end{aligned}$$

(3) **Morphisme pour le produit** Même genre de calculs que pour la somme.

(4) **Croissante**
Nous savons aussi par le lemme 1.382(4) que $1 \geqslant 0$. Puisque μ est un morphisme d'anneaux,

$$\mu(n+1) = \mu(n) + \mu(1) = \mu(n) + 1 \tag{1.558}$$

La définition 1.379(1a) dit alors que $\mu(n) \geqslant 0$ pour tout $n \in \mathbb{N}$. Nous avons pour la même raison que si $m \geqslant n$ dans $\mathbb{N}$, alors $\mu(m) \geqslant \mu(n)$ dans $\mathbb{K}$.
Soient maintenant $p, q \in \mathbb{N}$, et prouvons que $\sigma(p/q) \geqslant 0$. D'abord

$$\sigma(p/q) = \mu(p)\mu(q)^{-1} \tag{1.559}$$

où $\mu(p) \geqslant 0$ et $\mu(q) \geqslant 0$. Ensuite le lemme 1.382(5) nous indique que $\mu(q)^{-1} \geqslant 0$. Enfin la condition 1.379(1b) nous permet de conclure que $\sigma(p/q) \geqslant 0$.
Finalement, si $q_1 \geqslant q_2$ dans $\mathbb{Q}$, alors $q_1 - q_2 \geqslant 0$, et nous avons

$$\sigma(q_1) = \sigma(q_2 + q_1 - q_2) = \sigma(q_2) + \sigma(q_1 - q_2) \geqslant \sigma(q_2) \tag{1.560}$$

par la condition 1.379(1a).

□

NORMooJRRZooTwTVYG

1.404.
Si $\mathbb{K}$ est un corps totalement ordonné, la proposition 1.403 nous donne un morphisme de corps $\sigma\colon \mathbb{Q} \to \mathbb{K}$ qui respecte l'ordre. Pour $q \in \mathbb{Q}$ et $k \in \mathbb{K}$ nous notons

$$qk = \sigma(q)k. \tag{1.561}$$

EQooERFIooMpZVEs

Nous pourrons donc écrire $\frac{k}{2}$ pour $\sigma(1/2)k$.

Le lemme suivant explique que la notation (1.561) n'est pas complètement idiote.

LEMooWIONooGTKfcJ

Lemme 1.405.
Soit un corps commutatif totalement ordonné $\mathbb{K}$. Soit $k \in \mathbb{K}$. Nous avons

$$k + k = 2k. \tag{1.562}$$

Démonstration. Vu que $\sigma : \mathbb{Q} \to \mathbb{K}$ est un morphisme, il vérifie $\sigma(1) = 1$, donc

$$\begin{aligned} k + k &= \sigma(1)k + \sigma(1)k && \text{(1.563a)}\\ &= \big(\sigma(1) + \sigma(1)\big)k && \text{(1.563b)}\\ &= \sigma(2)k && \text{(1.563c)}\\ &= 2k. && \text{(1.563d)} \end{aligned}$$

□

PROPooTFVOooFoSHPg

Proposition 1.406.
Toute suite convergente dans un corps totalement ordonné est de Cauchy.

Démonstration. Soit un corps totalement ordonné $\mathbb{K}$ et une suite $x_n \xrightarrow{\mathbb{K}} x$. Soit $\epsilon > 0$. Il est important de se rendre compte que $\epsilon \in \mathbb{K}$ et que l'inégalité est au sens de l'ordre dans $\mathbb{K}$; en particulier ce n'est pas $\epsilon \in \mathbb{R}$ ni $\epsilon \in \mathbb{Q}$. D'ailleurs nous n'avons pas encore défini $\mathbb{R}$.

Vu que (x_n) converge vers x, il existe $N \in \mathbb{N}$ tel que pour tout $k > N$,

$$|x_k - x| < \epsilon. \tag{1.564}$$

Soient $p, q > N$. Alors en utilisant la majoration du lemme 1.383(6),

$$|x_p - x_q| = \big|(x_p - x) + (x - x_q)\big| \leqslant |x_p - x| + |x - x_q| \leqslant 2\epsilon. \tag{1.565}$$

EQooMQYGooLpgEQQ

En analyse en général, on s'arrête là et on dit que (x_n) est de Cauchy parce qu'il n'y a pas vraiment de différence entre réaliser une majoration avec ϵ ou avec 2ϵ. Détaillons toutefois comment ça se passe dans le cas où ϵ est un élément d'un corps totalement ordonné.

Le 2ϵ arrivant à la fin de (1.565) est en réalité $\epsilon + \epsilon = \sigma(2)\epsilon$ en vertu de ce qui est raconté en 1.404 et en vertu du lemme 1.405.

Considérons $\epsilon' = \sigma(1/2)\epsilon$, que nous pouvons noter $\epsilon' = \epsilon/2$. Vu que $\epsilon' > 0$, il existe un N' tel que pour tout $p, q > N'$ nous ayons

$$|x_p - x_q| \leqslant 2\epsilon' = \sigma(2)\sigma(1/2)\epsilon = \sigma(1)\epsilon = \epsilon. \tag{1.566}$$

Ce dernier ϵ étant bien celui fixé au début de la preuve, nous en déduisons que (x_n) est de Cauchy. □

1.21.1 Suites de Cauchy dans les rationnels

PropFFDJooAapQlP

Proposition 1.407 ([18]).
Principales propriétés des suites de Cauchy dans $\mathbb{Q}$.

ItemRKCIooJguHdji

(1) Toute suite convergente est de Cauchy[163]*.*

ItemRKCIooJguHdjii

(2) Toute suite de Cauchy est bornée.

ItemRKCIooJguHdjiii

(3) Si $x_n \to 0$ et si (y_n) est bornée, alors $x_n y_n \to 0$

(4) Si (x_n) et (y_n) sont de Cauchy alors $(x_n + y_n)$, $(x_n - y_n)$ et $(x_n y_n)$ sont également de Cauchy.

ITEMooIAFSooAIUpAN

(5) Si il existe $a, b \in \mathbb{Q}$ tels que $x_n \to a$ et $y_n \to b$ alors $x_n + y_n \to a + b$, $x_n - y_n \to a - b$ et $x_n y_n \to ab$.

163. Et non la réciproque, qui sera justement la grande innovation des nombres réels.

ItemRKCIooJguHdjvi

(6) *Soit (x_n) une suite de Cauchy qui ne converge pas vers zéro. Alors il existe n_0 tel que la suite $\left(\frac{1}{x_n}\right)_{n\geqslant n_0}$ soit de Cauchy.*

Démonstration. Point par point.

(1) C'est la proposition 1.406.

(2) Soit (x_n) une suite de Cauchy dans $\mathbb{Q}$. Avec $\epsilon = 1$ dans la définition, si $q > N$, nous avons

$$|x_q - x_N| \leqslant 1. \tag{1.567}$$

Et donc pour tout q plus grand que N, $x_N - 1 \leqslant x_q \leqslant x_N + 1$, ou encore, pour tout n :

$$|x_n| \leqslant \max\{|x_1|, |x_2|, \dots, |x_N|, |x_N + 1|\}. \tag{1.568}$$

La suite est donc bornée.

(3) Soit $\epsilon > 0$. Les hypothèses disent qu'il existe un N tel que $|x_n| \leqslant \epsilon$ dès que $n \geqslant N$. Et il existe aussi $M \geqslant 0$ tel que $|y_n| \leqslant M$ pour tout n. Du coup, lorsque $n \geqslant N$ nous avons $|x_n y_n| \leqslant M\epsilon$.

(4) En ce qui concerne la somme,

$$|x_p + y_p - x_q - y_q| \leqslant |x_p - x_q| + |y_p - y_q|. \tag{1.569}$$

EqDCNBooAzmBi

Soit N_1 tel que si $p, q \geqslant N_1$ alors $|x_p - x_q| \leqslant \epsilon$ et N_2 de même pour la suite (y_n). En prenant $N = \max\{N_1, N_2\}$, la somme (1.569) est plus petite que 2ϵ dès que $p, q \geqslant N$.

Passons à la démonstration du fait que le produit de deux suites de Cauchy est de Cauchy. Les suites (x_n) et (y_n) sont bornées et quitte à prendre le maximum, nous disons qu'elles sont toutes les deux bornées par le nombre M : pour tout n nous avons $|x_n| \leqslant M$ et $|y_n| \leqslant M$. Nous avons :

$$|x_p y_p - x_q y_q| \leqslant |x_p y_p - x_q y_p| + |x_q y_p - x_q y_q| \leqslant |y_p||x_p - x_q| + |x_q||y_p - y_q|. \tag{1.570}$$

Puisque (x_n) et (y_n) sont de Cauchy, si p et q sont assez grands, les deux différences sont majorées par ϵ et nous avons

$$|x_p y_p - x_q y_q| \leqslant M\epsilon + M\epsilon = 2M\epsilon, \tag{1.571}$$

ce qui prouve que $(x_n y_n)$ est de Cauchy.

(5) En ce qui concerne la somme, nous pouvons tout de suite calculer

$$|x_n + y_n - (a + b)| \leqslant |x_n - a| + |y_n - b|. \tag{1.572}$$

Il existe une valeur de n à partir de laquelle le premier terme est plus petit que ϵ et une à partir de laquelle le second terme est plus petit que ϵ. En prenant le maximum des deux, la somme est plus petite que 2ϵ.

En ce qui concerne le produit,

$$|x_n y_n - ab| \leqslant |x_n y_n - a y_n| + |a y_n - ab| \leqslant |y_n||x_n - a| + |a||y_n - b|. \tag{1.573}$$

Les suites $|x_n - a|$ et $|y_n - b|$ convergent vers zéro ; la suite (y_n) est bornée parce que convergente (combinaison des points (1) et (2)) et a (la suite constante) est également bornée. Donc par le point (3), nous avons

$$y_n|x_n - a| + a|y_n - b| \to 0. \tag{1.574}$$

Au passage nous avons également utilisé la propriété de la somme que nous venons de démontrer.

(6) Soit (x_n) une suite de Cauchy dans $\mathbb{Q}$ ne convergeant pas vers zéro : il existe $\alpha > 0$ tel que pour tout $N \in \mathbb{N}$, il existe $n \geqslant N$ tel que $|x_n| > \alpha$. Mais notre suite est de Cauchy, donc il existe $n_0 \in \mathbb{N}$ tel que si $p, q \geqslant n_0$ alors

$$|x_p - x_q| \leqslant \frac{\alpha}{2}. \tag{1.575}$$

En fixant $N = n_0$, on obtient un naturel $n \geqslant n_0$ tel que $|x_n| \geqslant \alpha$. De plus, comme la suite est de Cauchy, si $p > n$ nous avons aussi $|x_n - x_p| \leqslant \frac{\alpha}{2}$. Cela implique $|x_p| \geqslant \frac{\alpha}{2}$ et en particulier $x_p \neq 0$.
Nous venons de prouver que la suite ne s'annule plus à partir de l'indice n, et même que $|x_k| \geqslant \alpha/2$ pour tout $k \geqslant n$. La suite $(1/x_k)_{k \geqslant n}$ est donc bien définie.
Soit $\epsilon > 0$. Soit n_0 tel que $|x_p - x_q| < \epsilon$ pour tout $p, q > n_0$. Soit K plus grand que n_0 et que n. En prenant $p, q \geqslant K$, nous avons $|x_p| > \frac{\alpha}{2}$ et $|x_q| > \frac{\alpha}{2}$. Nous en déduisons que

$$\left| \frac{1}{x_p} - \frac{1}{x_q} \right| \leqslant \frac{|x_q - x_p|}{|x_p x_q|} \leqslant \frac{4}{\alpha^2} |x_q - x_p| \leqslant \frac{4}{\alpha^2} \epsilon. \tag{1.576}$$

Donc $\left(\frac{1}{x_n} \right)$ est de Cauchy.

□

1.22 Insuffisance des rationnels

Nous allons voir qu'il n'existe pas de nombres rationnels x tels que $x^2 = 2$, mais que pourtant il existe une infinité de suites de rationnels (x_n) tels que $x_n^2 \to 2$.

LemJPIUooWFHaFM

Lemme 1.408.
Un entier x est pair si et seulement si l'entier x^2 est pair.

Démonstration. Si x est un nombre pair, alors il existe un entier a tel que $x = 2a$ alors $x^2 = 4a^2$ est pair.

Inversement, si x est impair alors il existe un entier a tel que $x = 2a + 1$ et alors $x^2 = 4a^2 + 4a + 1 = 2(2a^2 + 2a) + 1$ est impair. □

Le théorème 3.36 nous dira que tous les $\sqrt{n}$ sont irrationnels dès que n n'est pas un carré parfait. Voici déjà le résultat pour $n = 2$. Le fait que $\sqrt{2}$ existe dans $\mathbb{R}$ sera la proposition 1.474.

PropooRJMSooPrdeJb

Proposition 1.409 (Irrationalité de $\sqrt{2}$)**.**
Il n'existe pas de fractions d'entiers dont le carré soit égal à 2.

Démonstration. Nous supposons que la fraction d'entiers a/b est telle que $a^2/b^2 = 2$, et nous allons construire une suite d'entiers strictement décroissante et strictement positive, ce qui est impossible.

Grâce au lemme 1.408 nous avons successivement les affirmations suivantes :

— $\frac{a^2}{b^2} = 2$ avec $a \neq 0$ et $b \neq 0$.
— $a^2 = 2b^2$, donc a^2 est pair.
— a est alors pair et a^2 est divisible par 4. Soit $a^2 = 4k$.
— $4k/b^2 = 2$, donc $4k = 2b^2$, donc $b^2 = 2k$ et b^2 est pair.
— Nous déduisons que b est pair.

La fraction $\frac{a/2}{b/2}$ est alors une nouvelle fraction d'entiers dont le carré vaut 2. En procédant de la même façon, en remplaçant a par $a/2$ et b par $b/2$, on obtient que la fraction d'entiers $\frac{a/4}{b/4}$ a la même propriété.

En particulier, tous les nombres de la forme $a/2^n$ sont des entiers. Ils forment une suite strictement décroissante d'entiers strictement positifs. Impossible, me diriez-vous ? Et vous auriez bien

raison : toute partie non vide de $\mathbb{N}$ admet un plus petit élément [164]. Il n'y a donc pas de fractions d'entiers dont le carré vaut 2. □

LEMooOTVUooImvusn

Lemme 1.410 (Série géométrique)**.**
Si $q \neq 1$ dans $\mathbb{Q}$ et $p \in \mathbb{N}$ nous avons

$$\sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{1.577}$$

Démonstration. En posant $S_p = 1 + q + q^2 + \cdots + q^p$, nous avons $S_p - qS_p = 1 - q^{p+1}$ et donc

$$S_p = \sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{1.578}$$

□

Proposition 1.411.
La suite donnée par

$$x_n = 1 + \frac{1}{1!} + \cdots + \frac{1}{n!} \tag{1.579}$$

est de Cauchy et ne converge pas dans $\mathbb{Q}$.

Démonstration. Si $p > q > 0$ nous avons

$$\begin{aligned}
x_p - x_q &= \sum_{k=q+1}^{p} \frac{1}{k!} && \text{(1.580a)}\\
&\leqslant \sum_{k=q+1}^{p} \frac{1}{(q+1)!}\frac{1}{(q+1)^{k-q-1}} && \text{(1.580b)}\\
&\leqslant \frac{1}{(q+1)!} \lim_{p\to\infty} \sum_{k=0}^{p} \frac{1}{(q+1)^k} && \text{(1.580c)}\\
&\leqslant \frac{1}{(q+1)!}\frac{1}{1-\frac{1}{q+1}} && \text{(1.580d)}\\
&\leqslant \frac{1}{(q+1)!}\frac{q+1}{q} && \text{(1.580e)}\\
&\leqslant \frac{1}{q!q}. && \text{(1.580f)}
\end{aligned}$$

SUBEQooAXILooEAcnVB SUBEQooNDPTooDSEYEJ SUBEQooEMHJooSnCUiK

Justifications :

— Pour (1.580b), il s'agit de remplacer dans $k!$ tous les facteurs plus grands que $(q+1)$ par $q+1$. Cela rend le dénominateur plus petit.

— Pour (1.580c), il y a une inégalité parce que la suite $p \mapsto \sum_{k=0}^{p} 1/(q+1)^k$ est une suite strictement croissante.

— Pour (1.580d), le lemme 1.410 donne la valeur de la somme finie. En ce qui concerne la limite, nous avons demandé $p > q > 0$ et donc $q + 1 > 1$. Dans ce cas la limite fonctionne.

Cette inégalité une fois établie nous permet de prouver les assertions. La suite (x_n) est de Cauchy car, pour tout $\epsilon \in \mathbb{Q}$ s'écrivant $\epsilon = \frac{a}{b}$ avec $a, b \in \mathbb{N}$, en prenant $p, q > b$, nous avons

$$x_p - x_q \leqslant \frac{1}{b!b} < \frac{1}{b} < \frac{a}{b} = \epsilon. \tag{1.581}$$

164. Voir [18], et attention : ce n'est pas tout à fait évident.

Montrons par l'absurde que cette suite ne converge pas dans $\mathbb{Q}$. Pour cela, nous supposons que $\lim_{n\to\infty} x_n = \frac{a}{b} \in \mathbb{Q}$. Pour tout $p > q$ nous avons établi

$$0 < x_p - x_q < \frac{1}{qq!}. \tag{1.582}$$

Prenons la limite $p \to \infty$; par stricte croissance de la suite, les inégalités restent strictes :

$$0 < \frac{a}{b} - x_q < \frac{1}{qq!}. \tag{1.583}$$

EqQLCTooQgQQdh

Si $n > b$ alors nous pouvons écrire

$$\frac{a}{b} - x_n = \frac{\alpha}{n!} \tag{1.584}$$

avec $\alpha \in \mathbb{Z}$ parce que le dénominateur commun entre $\frac{a}{b}$ et x_n est dans $n!$. En prenant donc $q > n$ dans (1.583) nous pouvons écrire

$$0 < \frac{\alpha}{q!} < \frac{1}{qq!}, \tag{1.585}$$

c'est-à-dire $0 < \alpha < \frac{1}{q}$, ce qui est impossible pour $\alpha \in \mathbb{Z}$. □

LEMooDTXYooKwmlZh

Lemme 1.412.
Soit $A > 0$ dans $\mathbb{Q}$. Il existe un rationnel $q > 0$ tel que $q^2 < A$.

Démonstration. Vu que $\mathbb{Q}$ est archimédien (proposition 1.399), il existe $n \in \mathbb{N}$ tel que $1 < nA$. Pour ce n, nous avons

$$\left(\frac{1}{n}\right)^2 < \frac{1}{n} < A. \tag{1.586}$$

□

La proposition suivante donne une suite de rationnels qui convergerait dans $\mathbb{R}$ vers $\sqrt{A}$ (non encore défini à ce stade). Il est expliqué dans [74] que la suite peut être vue comme une forme de méthode de Newton 34.54 ; voir l'exemple 34.57. Si vous aimez les dessins et les approches géométriques, il y a une explication sur Wikipédia[75].

PROPooSTQXooHlIGVf

Proposition 1.413 ([74]).
Soient $A > 0$ dans $\mathbb{Q}$ et $x_0 \in \mathbb{Q}$. La suite (x_k) définie par

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{A}{x_k}\right) \tag{1.587}$$

EQooOUIVooUqWhXe

a les propriétés suivantes :

(1) *La suite $y_k = x_k^2$ converge dans $\mathbb{Q}$ vers A.*
(2) *La suite (x_k) est de Cauchy dans $\mathbb{Q}$.*
(3) *La suite (x_k) ne converge pas dans $\mathbb{Q}$ dans le cas de $A = 2$.*

Démonstration. En plusieurs points.

(1) **La suite s_k** En posant $y_k = x_k^2$ nous calculons que

$$y_{k+1} - A = \frac{(y_k - A)^2}{4y_k}. \tag{1.588}$$

Autrement dit, la suite $s_k = y_k - A$ vérifie

$$s_{k+1} = \frac{s_k^2}{4(A + s_k)}. \tag{1.589}$$

Quelle que soit la valeur de $s_0 = x_0^2 - A$, nous avons

$$s_1 = \frac{s_0^2}{4(A+s_0)} = \frac{(x_0^2-A)^2}{4(A+x_0^2-A)} = \frac{(x_0^2-A)^2}{4x_0^2} > 0. \tag{1.590}$$

Donc à partir de s_1, tous les éléments sont positifs. Vu que $A > 0$ nous avons aussi

$$s_{k+1} < \frac{s_k^2}{4s_k} = \frac{s_k}{4} \tag{1.591}$$

et donc $s_k < s_0/4^k$. Donc $s_k \to 0$.

(2) **La suite (y_k)** Nous venons de prouver que si $y_k = A + s_k$, alors $s_k \to 0$. Autrement dit, la suite y_k converge vers A dans $\mathbb{Q}$.
La suite (y_k) est donc de Cauchy par la proposition 1.407(1).

(3) **La suite (x_k) est de Cauchy** Soit $\epsilon > 0$ dans $\mathbb{Q}$. Puisque (y_k) est de Cauchy, il existe $n_0 \in \mathbb{N}$ tel que

$$|x_r^2 - x_s^2| < \epsilon \tag{1.592}$$

pour tout $r, s \geqslant n_0$.
Soit par ailleurs $q \neq 0$ dans $\mathbb{Q}$ tel que $q^2 < A$, assuré par le lemme 1.412. Quitte à augmenter la valeur de n_0, nous supposons que $x_r, x_s > q$, et en particulier que $x_r + x_s \neq 0$. Cela permet d'écrire d'abord

$$x_r^2 - x_s^2 = (x_r + x_s)(x_r - x_s) \tag{1.593}$$

et ensuite de prendre la valeur absolue et de diviser par $|x_r + x_s|$:

$$|x_r - x_s| = \frac{|x_r^2 - x_s^2|}{|x_r + x_s|} < \frac{\epsilon}{2q}. \tag{1.594}$$

Donc (x_k) est une suite de Cauchy.

(4) **Pas de convergence pour $A = 2$** Supposons que $x_k \to a \in \mathbb{Q}$. Dans ce cas nous aurions $x_k^2 \to a^2 = A = 2$ (proposition 1.407(5)). Mais nous savons par la proposition 1.409 que $a^2 = 2$ est impossible dans $\mathbb{Q}$.

□

Notons que cette proposition ne présume en rien de l'existence ou de la non-existence dans $\mathbb{Q}$ d'un élément qui pourrait décemment être nommé $\sqrt{A}$. Il se fait que le théorème 3.36 dira que $\sqrt{n}$ est soit entier, soit irrationnel.

1.414.
Un petit programme en python pour explorer la suite de la proposition 1.413.

```python
#!/usr/bin/python3

def rec(A,x):
    return ((x**2+A)/x)/2

A = 3        # Compute square root of 3
x = 1000     # Initial guess: 1000

for i in range(1,100):
    print(i, x, x**2, x**2-A)
    x = rec(A, x)
```

tex/frido/codeSnip_4.py

1.23 Les réels

Une construction des réels via les coupures de Dedekind est donnée dans [76].

NormooHRDZooRGGtCd

1.415.
La construction des réels va nécessiter un petit « bootstrap » au niveau de la topologie. En effet la notion de suite de Cauchy est une notion topologique (définition 7.260) alors que la topologie métrique (celle entre autres de $\mathbb{Q}$) ne sera définie que par le théorème 7.117. Nous avons donc dû définir en la définition 1.379 *ex nihilo* les notions de

— suite de Cauchy

— suite convergente

— complétude

Nous allons ensuite construire $\mathbb{R}$ comme ensemble de classes d'équivalence de suites de Cauchy dans $\mathbb{Q}$. Ce ne sera que plus tard, après avoir défini la notion d'espace métrique que nous allons voir que sur $\mathbb{R}$, ces trois notions coïncident avec celles topologiques [165]. Et par conséquent que $\mathbb{R}$ sera un espace métrique complet [166].

Dans cette optique, il est intéressant de lire ce que dit Wikipédia à propos des suites de Cauchy dans l'article consacré à la construction des nombres réels[77].

1.23.1 L'ensemble

Soit $\mathcal{E}$ l'ensemble des suites de Cauchy [167] dans $\mathbb{Q}$. Soit aussi l'ensemble $\mathcal{E}_0$ constituée des suites qui convergent vers zéro [168].

En posant

$$x + y = (x_n + y_n) \tag{1.595}$$

et

$$xy = (x_n y_n), \tag{1.596}$$

l'ensemble $\mathcal{E}$ devient un anneau [169] commutatif dont le neutre de l'addition est la suite constante $x_n = 0$ et le neutre pour la multiplication est la suite constante $x_n = 1$.

PROPooNUQVooAAkicK

Proposition 1.416.
La partie $\mathcal{E}_0$ est un idéal[170] *de l'anneau $\mathcal{E}$.*

Démonstration. Nous savons par la proposition 1.407(1) que les suites convergentes sont de Cauchy ; par conséquent $\mathcal{E}_0 \subset \mathcal{E}$.

L'ensemble structuré $(\mathcal{E}_0, +)$ est un sous-groupe de $\mathcal{E}$ par les propriétés de la proposition 1.407 (il s'agit du fait que la somme de deux suites convergeant vers zéro est une suite convergente vers zéro).

En ce qui concerne la propriété fondamentale des idéaux, si $x \in \mathcal{E}_0$ et $y \in \mathcal{E}$ nous devons prouver que $xy \in \mathcal{E}_0$. Puisque $(\mathcal{E}_0, \cdot)$ est commutatif, cela suffira pour être un idéal bilatère. Vu que y est une suite de Cauchy, elle est bornée ; et étant donné que $x \to 0$ nous avons alors $xy \to 0$ (par la proposition 1.407(3)). □

DefooFKYKooOngSCB

Théorème-Définition 1.417 (L'anneau des réels[18])**.**
Sur l'ensemble quotient $\mathcal{E}/\mathcal{E}_0$, les opérations

(1) $\bar{u} + \bar{v} = \overline{u + v}$

165. Proposition 7.269.
166. Théorème 1.456 pour la complétude en tant que corps et théorème 1.406 pour la complétude en tant que espace métrique.
167. Définition 1.379(3)
168. Nous rappelons qu'à ce niveau nous n'avons pas encore prouvé que toutes les suites de Cauchy convergent.
169. Définition 1.42.
170. Définition 1.182.

(2) $\bar{u} \cdot \bar{v} = \overline{uv}$

sont bien définies et donnent à $\mathcal{E}/\mathcal{E}_0$ *une structure de corps commutatif appelé* ***corps des réels*** *et noté* $\mathbb{R}$

Démonstration. Nous divisons la preuve en plusieurs parties.

(1) **Les opérations sont bien définies** La partie $\mathcal{E}_0$ est un idéal par la proposition 1.416. Le quotient est donc bien défini et est un anneau par la proposition 1.184(2).

(2) **Caractérisation des classes** Soit $q \in \mathbb{Q}$ et une suite x convergente vers q. Cette suite est de Cauchy comme toute suite convergente. Montrons que

$$\bar{x} = \{\text{suites qui convergent vers } q\}. \tag{1.597}$$

Si $y \in \bar{x}$ alors $y = x + h$ avec $h \in \mathcal{E}_0$, et comme $h_n \to 0$, on a $y_n \to q$. Réciproquement, si $y_n \to q$ alors pour chaque n nous avons

$$y_n = x_n + (y_n - x_n), \tag{1.598}$$

mais $y_n - x_n \to 0$. Donc la suite $y - x \in \mathcal{E}_0$ ce qui signifie que $y \in \bar{x}$.

(3) **Neutre et unité** Il est vite vérifié que $\bar{0}$, la classe de la suite constante égale à zéro est neutre pour l'addition. De même, $\bar{1}$, est un neutre pour la multiplication.

(4) **Corps** Nous devons prouver que tout élément non nul est inversible. C'est-à-dire que si $x \in \mathcal{E}$ ne converge pas vers zéro [171] alors il existe $y \in \mathcal{E}$ tel que $xy \in \bar{1}$.

Nous savons par la proposition 1.407(6) que x étant une suite de Cauchy dans $\mathbb{Q}$, il existe $n_0 \in \mathbb{N}$ tel que $\left(\frac{1}{x_n}\right)_{n \geqslant n_0}$ est une suite de Cauchy. Nous posons alors

$$y_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ \frac{1}{x_n} & \text{si } n > n_0. \end{cases} \tag{1.599}$$

Nous avons alors

$$(xy)_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ 1 & \text{si } n > n_0 \end{cases} \tag{1.600}$$

et donc $xy \in \bar{1}$.

□

NORMooWBYNooBQaPPk

1.418 (Quelques notations entre $\mathbb{Q}$ et $\mathbb{R}$)**.**
Si $k \mapsto x_k$ est une suite, nous notons (x_k) avec des parenthèses la suite elle-même. Le k dans (x_k) est un indice muet, et dans les cas où il peut y avoir une ambiguïté, nous pouvons noter $(x_k)_{k\in\mathbb{N}}$. Cette dernière notation est plus lourde, mais plus exacte.

Le mieux est d'écrire simplement x la suite, mais alors il faut être prudent et ne pas noter x la limite. Nous éviterons donc d'écrire $x_k \to x$.

Si (x_k) est une suite de Cauchy dans $\mathbb{Q}$, nous notons $\bar{x}$ l'élément de $\mathbb{R}$ qui lui correspond. En fait $\bar{x} = (x_k)$: $\bar{x}$ est la suite-elle même, mais pour nous souvenir de l'origine nous allons adopter cette notation.

D'autre part nous définissons

$$\begin{aligned} \varphi\colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto \overline{[k \mapsto q]}, \end{aligned} \tag{1.601}$$

c'est-à-dire que $\varphi(q)$ est la classe de la suite constante $x_k = q$.

171. $x \in \mathcal{E}$ peut soit ne pas converger du tout, soit converger vers autre chose que zéro.

PropooEPFCooMtDOfP

Proposition 1.419.
Soit l'application

$$\begin{aligned}\varphi\colon \mathbb{Q} &\to \mathbb{R}\\ q &\mapsto \bar{q}.\end{aligned} \tag{1.602}$$

où par $\bar{q}$ nous entendons la classe de la suite constante égale à q (qui est de Cauchy).

(1) C'est un homomorphisme de corps injectif.

(2) $\text{Image}(\varphi)$ *est un sous-corps de* $\mathbb{R}$

(3) $\varphi\colon \mathbb{Q} \to \text{Image}(\varphi)$ *est un isomorphisme de corps.*

Démonstration. Le fait que ce soit un homomorphisme est simplement

— $\varphi(q+q') = \overline{q+q'} = \bar{q} + \overline{q'}$

— $\varphi(qq') = \overline{qq'} = \bar{q}\overline{q'}$.

En ce qui concerne l'injectivité, si q est tel que $\varphi(q) = \bar{0} = \mathcal{E}_0$, c'est que

$$\varphi(q) = \{\text{suites de Cauchy qui convergent vers zéro}\} \tag{1.603}$$

Mais nous savons aussi que [172]

$$\varphi(q) = \bar{q} = \{\text{suites de Cauchy qui convergent vers } q\} \tag{1.604}$$

Nous en déduisons que $q = 0$. □

Lorsque dans la suite nous parlerons d'un élément de $\mathbb{Q}$ comme étant un réel, nous aurons en tête l'image de cet élément par φ.

LEMooYLQBooFistHs

Lemme 1.420.
Soient $q, l \in \mathbb{Q}$ tels que $\bar{q} = \bar{l}$. Alors $q = l$ dans $\mathbb{Q}$.

Démonstration. La suite constante $x_n = q$ est un représentant de $\bar{q}$, et la suite constante $y_n = l$ est représentant de $\bar{l}$. Dire que $\bar{l} = \bar{q}$ signifie qu'il existe une suite $z \in \mathcal{E}_0$ tel que

$$x = y + z. \tag{1.605}$$

Pour tout n nous avons donc $x_n = y_n + z_n$, ou encore

$$z_n = q - l \tag{1.606}$$

pour tout n. Puisque z est une suite constante, elle ne peut appartenir à $\mathcal{E}_0$ que si elle est la suite constante nulle, c'est-à-dire si $q = l$. □

1.23.2 Relation d'ordre

Définition 1.421.
Nous définissons les parties $\mathcal{E}^+$ et $\mathcal{E}^-$ de $\mathcal{E}$ par

(1) $x \in \mathcal{E}^+$ si et seulement si pour tout $\epsilon > 0$ (ϵ est dans $\mathbb{Q}$), il existe N_ϵ tel que $n > N_\epsilon$ implique $x_n > -\epsilon$.

(2) $x \in \mathcal{E}^-$ si et seulement si pour tout $\epsilon > 0$, il existe N_ϵ tel que $n > N_\epsilon$ implique $x_n < \epsilon$.

Nous notons aussi $\mathcal{E}^{++} = \mathcal{E}^+ \backslash \mathcal{E}_0$.

Dans le lemme suivant, le point (2) peut sembler perturbant. Il s'agit de dire que si x est la classe de la suite constante 0, alors il est le neutre pour l'addition dans $\mathbb{R}$.

172. Voir dans la démonstration du théorème 1.417.

LEMooJOYQooCDlhHW

Lemme 1.422 (À propos du zéro)**.**
Nous avons

ITEMooKRBYooZGhhch

(1) $\mathcal{E}^+ \cap \mathcal{E}^- = \{\bar{0}\}$.

ITEMooRCIZooHUIymE

(2) Si $x = \bar{0}$ *alors* $x = 0$.

LEMooSWYXooMKMLYI

Lemme 1.423.
Les parties $\mathcal{E}^+$ *et* $\mathcal{E}^-$ *partitionnent* $\mathcal{E}$ *de la façon suivante :*

ITEMooZRVXooANHspZ

(1) $\mathcal{E}^+ \cap \mathcal{E}^- = \mathcal{E}_0$

(2) $\mathcal{E}^+ \cup \mathcal{E}^- = \mathcal{E}$

Démonstration. On prouve d'abord que $\mathcal{E}^+ \cap \mathcal{E}^- \subset \mathcal{E}_0$, l'inclusion inverse est évidente. Soit $\epsilon > 0$ et $x \in \mathcal{E}^+ \cap \mathcal{E}^-$. Il existe un $N \in \mathbb{N}$ tel que $x_n > -\epsilon$ et $x_n < \epsilon$ pour tout $n \geqslant N$. Par conséquent, $|x_n| \leqslant \epsilon$ pour tout $n \geqslant N$ et la suite x converge vers zéro, c'est-à-dire $x \in \mathcal{E}_0$.

Pour prouver le second point, soit $x \in \mathcal{E} \backslash \mathcal{E}^-$, et prouvons que $x \in \mathcal{E}^+$. La condition $x \notin \mathcal{E}^-$ donne qu'il existe un $\alpha > 0$ (dans $\mathbb{Q}$) tel que pour tout n, il existe $p > n$ avec $x_p > \alpha$. Mais x est une suite de Cauchy, donc nous avons un n_0 tel que si $n, p \geqslant n_0$ alors $|x_n - x_p| \leqslant \frac{\alpha}{2}$. En particulier, si $n \geqslant n_0$, et si $p > n$ est tel que $x_p > \alpha$, on obtient

$$x_n > \frac{\alpha}{2} > 0 \tag{1.607}$$

Par conséquent $x \in \mathcal{E}^+$ parce que $x \in \mathcal{E}$ et les x_n sont tous positifs à partir d'un certain rang. □

LEMooXNWSooHbNcAV

Lemme 1.424 ([1])**.**
Si $x \in \mathcal{E}^{++}$*, alors il existe* N *tel que* $x_n > 0$ *pour tout* $n > N$.

Démonstration. La suite x ne tend pas vers zéro. Donc il existe $\delta > 0$ tel que pour tout $N > 0$ il existe $n > N$ vérifiant $x_n > \delta$.

Mais la suite x est également de Cauchy. Écrivons cette condition pour $\delta/2$. Il existe $N_2 > 0$ tel que $p, q > N_2$ implique $|x_p - x_q| < \delta/2$.

Nous fixons $n > N_2$ tel que $x_n > \delta$. Alors pour tout $p > N_2$ nous avons aussi

$$|x_p - x_n| < \frac{\delta}{2}. \tag{1.608}$$

Cela implique que $x_p > \delta/2 > 0$ pour tout $p > N_2$. □

La proposition suivante est une version plus précise du lemme 1.424.

Proposition 1.425 ([13])**.**
Soit $x \in \mathcal{E}^{++}$*. Il existe* $r \in \mathbb{Q}^+ \backslash \{0\}$ *et* $N \in \mathbb{N}$ *tel que* $x_n \geqslant r$ *pour tout* $n \geqslant N$.

Démonstration. Fixons $\epsilon \in \mathbb{Q}^+ \backslash \{0\}$, et procédons par l'absurde. Du coup nous savons trois choses sur la suite (x_n).

ITEMooTNZDooHKZdBe

(1) Hypothèse absurde : pour tout $q \in \mathbb{Q}^+ \backslash \{0\}$ et pour tout $N \in \mathbb{N}$, il existe $p \geqslant N$ vérifiant $x_p < q$.

(2) Vu que $x \in \mathcal{E}^{++} \subset \mathcal{E}^+$, il existe $P_1 \in \mathbb{N}$ tel que $x_n > -\epsilon$ pour tout $n \geqslant P_1$.

(3) La suite x est de Cauchy. Donc il existe $P_2 \in \mathbb{N}$ tel que si $m, n \geqslant P_2$, nous avons $|x_m - x_n| < \epsilon/2$.

Nous considérons $P \geqslant \max\{P_1, P_2\}$ et nous prenons $q = \epsilon/2$, $N = P$ dans la propriété (1). Il existe donc $p > P$ tel que $x_p < \epsilon/2$.

Prenons aussi $n > p$ et écrivons les deux autres propriétés :

(1) $x_n > -\epsilon$ parce que $n > p > P > P_1$.

(2) $|x_n - x_p| < \epsilon/2$ parce que $n, p > P > P_2$.

Du coup nous avons

$$x_n < x_p + \frac{\epsilon}{2} < \epsilon, \tag{1.609}$$

et donc $-\epsilon < x_n < \epsilon$.

Au final, nous avons prouvé que pour tout $\epsilon \in \mathbb{Q}^+\backslash\{0\}$, il existe P tel que $n > P$ implique $|x_n| < \epsilon$. Cela signifie que

$$x_n \xrightarrow{\mathbb{Q}} 0, \tag{1.610}$$

c'est-à-dire $x \in \mathcal{E}_0$, ce qui est contraire à l'hypothèse. □

LEMooRKSXooFsIohe

Lemme 1.426 ([18]).
Quelques propriétés du partitionnement.

ITEMooRQVKooCnwWOY

(1) $x \in \mathcal{E}^-$ *si et seulement si* $(-x) \in \mathcal{E}^+$

ITEMooJUPOooOBubqA

(2) $x \in \mathcal{E}^+$ *et* $y \in \mathcal{E}^+$ *implique* $x + y \in \mathcal{E}^+$

ITEMooDQLJooPViuVC

(3) $x \in \mathcal{E}^+$ *et* $y \in \mathcal{E}^+$ *implique* $xy \in \mathcal{E}^+$

(4) Si $x, y \in \mathcal{E}$ *sont tels que* $x - y \in \mathcal{E}_0$ *alors soit* $x, y \in \mathcal{E}^+$ *soit* $x, y \in \mathcal{E}^-$.

Démonstration. Point par point.

(1) Définition de $\mathcal{E}^+$ et $\mathcal{E}^-$.

(2) Pour $n \geqslant N_{\epsilon/2}$ nous avons $x_n > -\epsilon/2$ et $y_n > -\epsilon/2$. Donc $x_n + y_n > -\epsilon$.

(3) Si x ou y est dans $\mathcal{E}_0$ alors $xy \in \mathcal{E}_0$ et c'est bon. Si par contre $x, y \in \mathcal{E}^{++}$ alors le lemme 1.424 nous indique que pour n suffisamment grand, $x_n > 0$ et $y_n > 0$. Et dans ce cas, $(xy)_n > 0$, c'est-à-dire $xy \in \mathcal{E}^+$.

(4) Supposons que $x - y \in \mathcal{E}_0$ avec $x \in \mathcal{E}^+$ et prouvons qu'alors $y \in \mathcal{E}^+$. Soit donc $\epsilon > 0$; il existe n_1 tel que $x_n > -\frac{\epsilon}{2}$ dès que $n \geqslant n_1$. Mais $x - y \in \mathcal{E}_0$, donc il existe n_2 tel que $|x_n - y_n| < \frac{\epsilon}{2}$ dès que $n \geqslant n_2$. En prenant n plus grand que n_1 et n_2, nous avons en même temps

$$\begin{cases} x_n > -\frac{\epsilon}{2} & \text{(1.611a)} \\ |x_n - y_n| < \frac{\epsilon}{2}. & \text{(1.611b)} \end{cases}$$

Cela implique que $y_n > -\epsilon$ et donc que $y \in \mathcal{E}^+$.

Nous pouvons de même prouver que si $x \in \mathcal{E}^-$ alors $y \in \mathcal{E}^-$.

□

DefooLMQIooTgzZXd

Définition 1.427 (Positivité dans $\mathbb{R}$).
Vocabulaire et notations.

(1) Nous notons $\mathbb{R} = \mathcal{E}/\mathcal{E}_0$.

(2) Nous notons $\mathbb{R}^+ = \bar{\mathcal{E}}^+$.

(3) Nous notons $\mathbb{R}^- = \bar{\mathcal{E}}^-$.

(4) Un élément de $\mathbb{R}$ *est* ***positif*** *si il est la classe d'une suite de Cauchy appartenant à* $\mathcal{E}^+$.

(5) Un élément de $\mathbb{R}$ *est* ***négatif*** *si il est la classe d'une suite de Cauchy appartenant à* $\mathcal{E}^-$.

(6) Lorsque nous parlons de nombres réels, le symbole « 0 » signifie $\mathcal{E}_0$ *ou plus précisément la classe d'un élément de* $\mathcal{E}_0$ *modulo* $\mathcal{E}_0$.

REMooOCXLooKQrDoq

1.428.
Avec les conventions de la définition 1.427, et en anticipant sur nos connaissances à propos des réels,

(1) zéro est positif et négatif.

(2) L'intersection entre $\mathbb{R}^+$ et $\mathbb{R}^-$ est le singleton $\{0\}$.
(3) L'ensemble des nombres *strictement* positifs est noté $(\mathbb{R}^+)^*$ ou $\mathbb{R}^+\backslash\{0\}$.
(4) Le mot « positif » signifie « positif ou nul » ; le mot « négatif » signifie « négatif ou nul ». Ce sont des conventions qui sont également celles de Wikipédia[78].

DEFooBXHJooOEYPRI

Définition 1.429 (Ordre sur $\mathbb{R}$)**.**
Si $x, y \in \mathbb{R}$ nous notons $x \leqslant y$ si et seulement si $y - x \in \overline{\mathcal{E}^+}$.

PROPooYMJVooNAsXae

Proposition 1.430.
Le couple $(\mathbb{R}, \leqslant)$ est un corps totalement ordonné[173]

Démonstration. Il s'agit de vérifier, dans l'ordre, les définitions 1.11, 1.12 et 1.379(1). Pour la suite nous considérons $x, y, z \in \mathbb{R}$ et des suites de Cauchy a, b, c représentant x, y, z, c'est-à-dire telles que $x = \bar{a}$, $y = \bar{b}$ et $z = \bar{c}$.

(1) **Réflexivité** Pour savoir si $x \geqslant x$, nous devons nous demander si $x - x \in \overline{\mathcal{E}^+}$. Nous avons $x - x = \bar{a} - \bar{a} = \bar{0} = 0$.

(2) **antisymétrie** Nous supposons que $x \geqslant y$ et $y \geqslant x$. Du côté des suites de Cauchy, cela signifie que $a - b \in \mathcal{E}^+$ et $b - a \in \mathcal{E}^+$. Le lemme 1.426(1) nous indique alors que $a - b = -(b - a) \in \mathcal{E}^-$. Donc
$$a - b \in \mathcal{E}^+ \cap \mathcal{E}^- = \{\bar{0}\}. \tag{1.612}$$
Donc le réel $x - y$ est la classe de la suite constante 0. Le lemme 1.422(2) dit alors que $x - y = 0$ ou encore que $x = y$.

(3) **transitivité** Nous supposons que $x \leqslant y$ et $y \leqslant z$. L'hypothèse $x \leqslant y$ signifie $y - x \in \bar{\mathcal{E}}^+$ ou encore que $b - a \in \mathcal{E}^+$. Même chose pour $y \leqslant z$ qui signifie que $c - b \in \mathcal{E}^+$. Nous avons alors
$$\begin{aligned} c - a &= c - b + b - a && \text{(1.613a)}\\ &= (c - b) + (b - a) && \text{(1.613b)}\\ &\in \mathcal{E}^+ + \mathcal{E}^+ && \text{(1.613c)}\\ &\subset \mathcal{E}^+ && \text{(1.613d)} \end{aligned}$$
où nous avons utilisé le lemme 1.426(2). Puisque $c - a \in \mathcal{E}^+$, nous avons $z - x = \overline{c - a} \geqslant 0$.

(4) **Ordre total** Nous devons prouver que pour $x, y \in \mathbb{R}$ nous avons toujours $x \leqslant y$ ou $y \leqslant x$. Supposons que nous n'ayons pas $x \leqslant y$, c'est-à-dire $\overline{b - a} \notin \mathcal{E}^+$. Vu le lemme 1.423(2) nous avons $\overline{b - a} \in \mathcal{E}^-$, ce qui donne, par le lemme 1.426(1) que $\overline{a - b} \in \mathcal{E}^+$, c'est-à-dire $y \leqslant x$.

(5) **Corps ordonné** Enfin nous devons vérifier les deux conditions de la définition 1.379(1). Pour la première condition, nous supposons $x \leqslant y$, c'est-à-dire $b - a \in \mathcal{E}^+$. Nous avons donc
$$(b + c) - (a + c) = b + x - a - c = b - a \in \mathcal{E}^+, \tag{1.614}$$
donc $\overline{a + c} \leqslant \overline{b + c}$, c'est-à-dire $x + z \leqslant y + z$.
Pour la seconde condition, c'est le lemme 1.426(3).

□

Définition 1.431.
Puisque $\mathbb{R}$ est un corps totalement ordonné (proposition 1.430), si $x \in \mathbb{R}$, nous définissons $|x|$ conformément à 1.379(2).

LEMooTJAXooKEqPCG

Lemme 1.432.
L'application
$$\begin{aligned} \varphi\colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto \bar{q} \end{aligned} \tag{1.615}$$
dont nous avons déjà parlé dans la proposition 1.419 est strictement croissante.

173. Corps totalement ordonné, définition 1.379.

Démonstration. Nous supposons $q < l$ dans $\mathbb{Q}$. Nous devons montrer que $\bar{q} \leqslant \bar{l}$ dans $\mathbb{R}$, c'est-à-dire que $\bar{q} \leqslant \bar{l}$ et $\bar{q} \neq \bar{l}$.

Considérons la suite constante $x_n = l - q \in \mathbb{Q}$. Pour tout $\epsilon > 0$ dans $\mathbb{Q}$ nous avons

$$x_n = l - q > 0 > -\epsilon, \tag{1.616}$$

et donc $x_n \in \mathcal{E}^+$. Donc $\overline{l - q} \geqslant 0$. Cela signifie $\bar{q} \leqslant \bar{l}$.

D'autre part le lemme 1.420 dit que $\bar{q} = \bar{l}$ uniquement si $q = l$, ce qui est exclu parce que $q < l$. Donc $\bar{q} \neq \bar{l}$. □

LEMooNHMTooEdtBnQ

Lemme 1.433.
Si $a, b \in \mathbb{R}^+$ satisfont $a + b = 0$, alors $a = b = 0$.

Voici une version dans $\mathbb{R}$ du lemme 1.110.

LEMooKAXFooIPyzJC

Lemme 1.434.
Soient $a > 0$ et $b > 1$ dans $\mathbb{R}$. Nous avons

$$ab > a. \tag{1.617}$$

Remarque 1.435.
Comme déjà mentionné plus haut, à chaque fois que nous parlerons d'un élément de $\mathbb{Q}$ comme étant un élément de $\mathbb{R}$, nous considérons la classe de la suite constante.

PROPooWYPNooCcBERG

Proposition 1.436 ([1])**.**
Soient $a \leqslant b$ dans $\mathbb{R}$.

(1) Si $\lambda > 0$ alors $\lambda a \leqslant \lambda b$.

(2) Si $\lambda < 0$ alors $\lambda a \geqslant \lambda b$.

LemooYNOVooOwoRwD

Lemme 1.437.
Si $x, y, z \in \mathbb{R}$ avec $x > 0$ sont tels que $z > y/x$ alors $zx > y$.

Démonstration. Nous savons que

$$z - \frac{y}{x} \in \mathcal{E}^+ \backslash \{0\} = \mathcal{E}^{++}. \tag{1.618}$$

Puisque $x \in \mathcal{E}^{++}$, multiplier par x fait rester dans $\mathcal{E}^{++}$:

$$zx - x\frac{y}{x} \in \mathcal{E}^{++}. \tag{1.619}$$

Un représentant de $x\frac{y}{x}$ est la suite $n \mapsto x_n \frac{y_n}{x_n} = y_n$. Donc $x\frac{y}{x} = y$. Cela signifie que $zx - y \in \mathcal{E}^{++}$ et donc que $zx > y$. □

LemooMWOUooVWgaEi

Lemme 1.438.
Pour tout $a \in \mathbb{R}$, il existe $p \in \mathbb{N}$ tel que $p > a$.

Démonstration. Nous allons donner deux preuves différentes de ce lemme.

(1) **Première façon**
L'élément a de $\mathbb{R}$ admet un représentant (a_n) qui est une suite de Cauchy dans $\mathbb{Q}$. C'est donc une suite bornée, c'est-à-dire qu'il existe $m, q \in \mathbb{N}$ tels que $|a_n| \leqslant m/q$ pour tout n (proposition 1.407(2)). Soit M un naturel strictement plus grand que m/q [174].

La suite de Cauchy $(M - a_n)_{n\in\mathbb{N}}$ est constituée de rationnels positifs et est donc dans $\mathcal{E}^+$. La classe de $M - a$ est donc un réel positif [175]. Par définition de la relation d'ordre, $M \geqslant a$.

174. Lemme 1.397.
175. Et nous allons d'ailleurs arrêter de toujours préciser « la classe de » lorsque ce n'est pas nécessaire.

(2) **Seconde façon**
La suite (a_n) est majorée par $\frac{m}{q}$, donc on a dans $\mathbb{Q}$ et pour tout n :

$$a_n \leqslant \frac{m}{q} = M \leqslant qM. \tag{1.620}$$

L'application $\varphi\colon \mathbb{Q} \to \mathbb{R}$ est croissante, donc

$$\varphi\big((a_n)\big) \leqslant \varphi(qM). \tag{1.621}$$

□

En corolaire, nous avons

LEMooMWOUooVWgbFi

Lemme 1.439.
Pour tout $x \in \mathbb{R}$, il existe $q \in \mathbb{Z}$ tel que $q < x$.

Démonstration. Utilisation du lemme précédent avec $a = -x$: on prend $q = -p$. □

ThoooKJTTooCaxEny

Théorème 1.440 ([18]).
Le corps $\mathbb{R}$ est archimédien[176].

Démonstration. La proposition 1.430 dit que $\mathbb{R}$ est totalement ordonné. Soient $x, y \in \mathbb{R}$ avec $x > 0$; posons $a = \frac{y}{x}$. Le lemme 1.438 nous donne un $p \in \mathbb{N}$ tel que $p > a$. Nous concluons alors avec le lemme 1.437 :

$$px > ax = \frac{y}{x}x = y. \tag{1.622}$$

□

Le lemme suivant n'est pas loin de dire que $\mathbb{Q}$ est dense dans $\mathbb{R}$, à part que nous n'avons pas encore donné de topologie sur $\mathbb{R}$.

LemooHLHTooTyCZYL

Lemme 1.441.
À propos de rationnels entre des réels.

ITEMooGVTMooQsoTCi

(1) Si $x, y \in \mathbb{R}$ sont tels que $x < y$, alors il existe $s \in \mathbb{Q}$ tel que $x < s < y$.

ITEMooCVDSooAjimCL

(2) Si $\epsilon > 0$ dans $\mathbb{R}$, il existe $n \in \mathbb{N}$ tel que $\frac{1}{n} < \epsilon$.

ITEMooASUXooImjqzK

(3) Si $a \in \mathbb{R}$ vérifie $|a| < \epsilon$ pour tout $\epsilon > 0$, alors $a = 0$.

Démonstration. Nous avons par hypothèse que $y - x > 0$ et donc le fait que $\mathbb{R}$ soit archimédien (théorème 1.440) nous donne $q \in \mathbb{N}$ tel que $q(y - x) > 1$. Soit

$$E = \{n \in \mathbb{Z} \text{ tel que } \frac{n}{q} \leqslant x\}. \tag{1.623}$$

Cet ensemble n'est pas vide à cause du lemme 1.439 ; de plus, comme $|x|q \leqslant n_0$ pour un certain n_0 (à cause du lemme 1.438), l'ensemble E est majoré par n_0. Donc E possède un plus grand élément[177] p qui vérifie

$$\frac{p}{q} \leqslant x < \frac{p+1}{q}. \tag{1.624}$$

De plus $(p+1)/q < y$ parce que

$$\frac{p+1}{q} = \frac{p}{q} + \frac{1}{q} \leqslant x + \frac{1}{q} < x + y - x = y \tag{1.625}$$

où nous avons utilisé l'inégalité stricte $y - x > \frac{1}{q}$.

176. Définition 1.73.
177. Lemme 1.112.

Nous avons donc

$$x < \frac{p+1}{q} < y, \tag{1.626}$$

et le nombre $(p+1)/q$ convient comme s. Le point (1) est prouvé.

Pour le point (2), par le point (1) nous considérons $s \in \mathbb{Q}$ tel que $0 < s < \epsilon$. Si $s = p/q$ avec $p, q \in \mathbb{N}$ nous avons

$$0 < \frac{1}{q} \leqslant \frac{p}{q} < \epsilon. \tag{1.627}$$

Pour (3), supposons $a \neq 0$. Alors $|a| > 0$. Par le point (1), il existe $s \in \mathbb{R}$ tel que $0 < m < |a|$, ce qui est contraire à l'hypothèse. □

REMooXOIOooHjwMvA

Remarque 1.442.
Le lemme 1.441 a également pour conséquence que des ensembles comme $[-1, 1]$ ne sont pas bien ordonnés (définition 1.12). En effet la partie $]0, 1[$ ne possède pas de minimum parce que si $x \in]0, 1[$ alors $0 < x$ et il existe $s \in \mathbb{Q}$ (a fortiori $s \in \mathbb{R}$) tel que $0 < s < x$, c'est-à-dire que x n'est pas un minimum de $]0, 1[$.

Tant que nous y sommes dans les encadrements de réels...

1.443.
Soit $q_0 \in \mathbb{Q}$ tel que $0 \leqslant q_0 < 1$. On définit alors $d_1 \in \{0, 1\}$ comme valant 1 si $2q_0 \geqslant 1$ et 0 sinon. Puis on pose $q_1 = 2q_0 - d_1$.

Poursuivant de la sorte, on crée une suite $(d_n)_{n\geqslant 1}$: c'est le **développement dyadique** de q_0.

PROPooSZZBooUmsPeM

Proposition 1.444 ([1])**.**
Si $q \neq q'$, alors leurs développement dyadiques sont différents.

LEMooRSLIooVrZMxM

Lemme 1.445 ([1])**.**
Soit q, q' deux rationnels tels que $0 \leqslant q < q' < 1$. Il existe deux entiers naturels a et N tels que $q < \frac{a}{2^N} < q'$.

Démonstration. On crée les développements dyadiques de q et q', que l'on note respectivement $(d_n)_{n\geqslant 1}$ et $(d'_n)_{n\geqslant 1}$. Notons

$$E = \{n \in \mathbb{N} \text{ tel que } d_n \neq d'_n\}. \tag{1.628}$$

Comme $q \neq q'$, les développements dyadiques sont différents [178], l'ensemble E est non-vide, et il admet un plus petit élément N. Or, $q < q'$, et donc nécessairement $d_N < d'_N$. On construit alors $a = \sum_{i=1}^{N} d_i 2^i$. □

CorDensiteDyadiques

Corolaire 1.446.
Deux choses.

(1) Pour tous réels x, y tels que $0 \leqslant x < y \leqslant 1$, il existe un nombre de la forme $d = a/2^n$, avec $n \in \mathbb{N}$ et $a \in \mathbb{N}$, $a \leqslant 2^n$, tel que $x < d < y$.

(2) La partie

$$D = \{\frac{a}{2^b} \text{ tel que } a \in \mathbb{Z}, b \in \mathbb{N}\} \tag{1.629}$$

est dense dans $\mathbb{R}$.

*Ce D est nommé l'ensemble des **fractions dyadiques**.*

Démonstration. Soit $\epsilon > 0$. Le lemme 1.441(1) nous permet de considérer un rationnel q tel que $x < q < y$. Une utilisation combinée de 1.441(2) et de 1.208 permet de choisir n tel que $\frac{1}{2^n} < \epsilon$. Nous posons

$$a = \min\{s \in \mathbb{N} \text{ tel que } \frac{s}{2^n} > q\}. \tag{1.630}$$

178. Proposition 1.444.

Nous avons $a/2^n < q + \epsilon$, sinon prendre $a - 1$ aurait convenu. Nous avons donc

$$x < q \leqslant \frac{a}{2^n} < a + \epsilon < y. \tag{1.631}$$

En ce qui concerne la densité, soient $r \in \mathbb{R}$ et $\epsilon > 0$. Nous considérons $n \in \mathbb{N}$ tel que $1/2^n < \epsilon$. Ensuite nous posons

$$a = \min\{a \in \mathbb{N} \text{ tel que } \frac{a}{2^n} \geqslant r\}. \tag{1.632}$$

Nous avons alors $\frac{a}{2^n} - r < \epsilon$. Cela prouve que $B(r, \epsilon) \cap D \neq \varnothing$. □

LEMooEGYLooCGrDrl

Lemme 1.447 ([1])**.**
Soient des réels a, b, x, y tels que

$$a \leqslant x \leqslant b \tag{1.633}$$

et

$$a \leqslant y \leqslant b, \tag{1.634}$$

alors $|x - y| \leqslant |b - a|$.

LEMooTPLUooXiCZHJ

Lemme 1.448.
Soient deux réels a, b tels que

(1) $a \geqslant 0$,

(2) $b \geqslant 0$

(3) $a + b = 0$.

Alors $a = 0$ et $b = 0$.

LEMooNLGSooSGdvAo

Lemme 1.449.
Si $a \in \mathbb{R}$, alors $a^2 \geqslant 0$ et $a^2 = 0$ si et seulement si $a = 0$.

LEMooFQMVooDNaTDT

Lemme 1.450.
Soient $a, b, x \in \mathbb{R}$ tels que $a + x = b + x$. Alors $a = b$.

PROPooVIDJooSzXPLP

Proposition 1.451.
Si $a > b$ dans $\mathbb{R}$, alors

$$\frac{1}{a} < \frac{1}{b}. \tag{1.635}$$

1.23.3 Complétude

Le théorème 17.135 donne une complétion de tout espace métrique en un espace complet. Il serait tentant de l'utiliser ici pour définir $\mathbb{R}$ à partir de $\mathbb{Q}$. Cette méthode ne fonctionne cependant pas parce que la démonstration de 17.135 utilise le fait que $\mathbb{R}$ est complet.

LEMooXCVRooOSZYWv

Lemme 1.452.
Nous avons

$$|\varphi(q)| = \varphi(|q|) \tag{1.636}$$

pour tout $q \in \mathbb{Q}$.

Démonstration. Soit $q \in \mathbb{Q}$. Si $q \geqslant 0$ alors nous avons d'une part $|q| = q$ dans $\mathbb{Q}$, et d'autre part $\varphi(q) \geqslant 0$ dans $\mathbb{R}$. Donc au final

$$|\varphi(q)| = \varphi(q) = \varphi(|q|). \tag{1.637}$$

Supposons au contraire que $q < 0$. Alors $|q| = -q$ dans $\mathbb{Q}$, mais aussi $\varphi(q) \leqslant 0$. Donc

$$|\varphi(q)| = -\varphi(q) = \varphi(-q) = \varphi(|q|). \tag{1.638}$$

□

LemooRTGFooYVstwS

Lemme 1.453 ([18, 1])**.**
Toute suite de Cauchy dans $\mathbb{Q}$ converge dans $\mathbb{R}$ vers le réel qu'elle représente.

Plus précisément, en suivant les notations de 1.418, si (x_k) est une suite de Cauchy dans $\mathbb{Q}$, alors

(1) $\varphi(x_k)$ est une suite de Cauchy dans $\mathbb{R}$.

(2) $\varphi(x_k) \xrightarrow{\mathbb{R}} \bar{x}$.

Ici $\bar{x}$ est la classe de la suite x. C'est donc un élément de $\mathbb{R}$.

Démonstration. Soit (x_n) une suite de Cauchy de $\mathbb{Q}$, c'est-à-dire que $x_k \in \mathbb{Q}$ pour tout k et qu'elle est de Cauchy. Elle représente un réel $\bar{x} \in \mathbb{R}$, et nous voulons prouver que pour la topologie de $\mathbb{R}$ nous avons $\lim_{n\to\infty} \varphi(x_n) = \bar{x}$.

(1) **$\varphi(x_k)$ est de Cauchy**
Soit $\epsilon > 0$ dans $\mathbb{R}$. Nous considérons $\epsilon' \in \mathbb{Q}$ tel que $0 < \epsilon' < \epsilon$. Plus précisément tel que

$$0 < \varphi(\epsilon') < \epsilon. \tag{1.639}$$

Soit $N > 0$ tel que $|x_p - x_q| < \epsilon'$ pour tout $p, q \geqslant N$. Cela existe parce que (x_k) est dans Cauchy dans $\mathbb{Q}$. Nous avons alors

$$\begin{aligned} |\varphi(x_p) - \varphi(x_q)| &= |\varphi(x_p - x_q)| && \text{(1.640a)}\\ &= \varphi\big(|x_p - x_q|\big) && \text{(1.640b)}\\ &\leqslant \varphi(\epsilon') && \text{(1.640c)}\\ &\leqslant \epsilon. && \text{(1.640d)} \end{aligned}$$

SUBEQooZPTIooYsiEqk
SUBEQooDOFUooUmCZpR

Justifications :

— Pour (1.640b). C'est le lemme 1.452.

— Pour (1.640c). L'application φ est croissante, lemme 1.432.

Donc la suite $\varphi(x_k)$ est de Cauchy.

(2) $\varphi(x_k) \xrightarrow{\mathbb{R}} \bar{x}$ Nous devons prouver que pour tout $\epsilon \in \mathbb{R}$, il existe N tel que $n > N$ implique $\varphi(x_n) \in B\big(\bar{x}, \epsilon\big)$. Nous allons faire ça en deux parties. D'abord $\epsilon \in \mathbb{Q}$ et ensuite $\epsilon \in \mathbb{R}$.

(2a) **Avec $\epsilon \in \mathbb{Q}$** Soit $\epsilon \in \mathbb{Q}$. Puisque x est une suite de Cauchy dans $\mathbb{Q}$, il existe N tel que si $p, n > N$ nous avons

$$x_p - \epsilon < x_n < x_p + \epsilon. \tag{1.641}$$

EQooJURNooOoSzDZ

Ces inégalités sont dans $\mathbb{Q}$. Nous fixons p et nous commençons par écrire plus en détail la première inéquation :

$$x_p - \epsilon - x_n < 0. \tag{1.642}$$

Autrement dit, pour tout n nous avons

$$(\overline{x_p - \epsilon})_n < x_n. \tag{1.643}$$

Pour rappel, la suite $\overline{x_p - \epsilon}$ est la suite constante dans $\mathbb{Q}$. La suite

$$n \mapsto x_n - (\overline{x_p - \epsilon})_n \tag{1.644}$$

est dans $\mathcal{E}^+$. Donc, en vertu de la définition 1.429 nous avons

$$\bar{x} - \overline{x_p - \epsilon} \geqslant 0. \tag{1.645}$$

Nous pouvons aussi bien écrire

$$\bar{x} \geqslant \varphi(x_p) - \varphi(\epsilon). \tag{1.646}$$

En prenant l'autre inégalité de (1.641) nous trouvons de la même manière que

$$\bar{x} \leqslant \varphi(x_p) + \varphi(\epsilon). \tag{1.647}$$

Ces deux inégalités ensemble montrent que

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon)\big). \tag{1.648}$$

(2b) **Avec $\epsilon \in \mathbb{R}$** Nous considérons $\epsilon \in \mathbb{R}$ et $\epsilon' \in \mathbb{Q}$ tel que $\varphi(\epsilon') < \epsilon$. Par le point précédent, il existe N tel que $p > N$ implique

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon')\big). \tag{1.649}$$

Étant donné que $\varphi(\epsilon') < \epsilon$ nous avons

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon')\big) \subset B\big(\bar{x}, \epsilon\big). \tag{1.650}$$

□

PROPooZSQYooWRKNGY

Proposition 1.454.
Soit une suite convergente $x_k \xrightarrow{\mathbb{Q}} q$. Alors

$$\varphi(x_k) \xrightarrow{\mathbb{R}} \varphi(q) \tag{1.651}$$

où φ est la fonction qui à un rationnel fait correspondre la classe de la suite constante correspondante[179].

Démonstration. Le fait d'avoir une convergence $x_k \to q$ dans $\mathbb{Q}$ implique que la suite (x_k) est de Cauchy, par la proposition 1.407(1).

Le lemme 1.453 nous indique que $\varphi(x_k)$ est une suite dans $\mathbb{R}$ qui converge vers $\bar{q}$, la classe de la suite (x_k).

À prouver : $\varphi(x) = \bar{q}$. Autrement dit, nous devons prouver que la classe de la suite constante $a_k = q$ et la classe de la suite x sont les mêmes.

La suite $(x_k - q)$ est de Cauchy dans $\mathbb{Q}$ et converge vers zéro par hypothèse. Donc les suites x et (q) sont dans la même classe. □

PROPooFGBOooHiZqbs

Proposition 1.455 ([1])**.**
Deux choses à propos de suites de rationnels convergeant vers un réel.

ITEMooMAVYooKFtqlx

(1) Soit un réel x. Il existe une suite de rationnels strictement croissante qui converge vers x.

ITEMooVOVYooFUbccG

(2) Si de plus $x > 0$, alors la suite (toujours strictement croissante) peut être choisie parmi les rationnels strictement positifs.

Démonstration. Le lemme 1.441 nous sera d'une grande aide. Soit $x \in \mathbb{R}$. Il existe $q_0 \in \mathbb{Q}$ tel que $x - 1 < q_0 < x$. Ensuite nous construisons la suite par récurrence : q_k est choisi tel que $q_{k-1} < q_k < x$. Cela règle le point (1).

Pour (2). Il suffit de faire la même chose, en partant de $0 < q_0 < x$. □

THOooUFVJooYJlieh

Théorème 1.456 (Complétude de $\mathbb{R}$, critère de Cauchy[18])**.**
Nous avons :

(1) Le corps $\mathbb{R}$ est un corps complet (définition 1.379(5))

(2) Une suite dans $\mathbb{R}$ est convergente (définition 1.379(4)) si et seulement si elle est de Cauchy (définition 1.379(3)).

179. Voir les notations en 1.418.

Notez la grande similitude entre ce théorème et le théorème 7.287. Ils ne sont pas équivalents, ne parlent pas exactement du même objet « $\mathbb{R}$ », ni des mêmes notions de suites de Cauchy et de complétude.

Démonstration. Soit (x_n) une suite de Cauchy dans $\mathbb{R}$. Pour chaque n, il existe par le lemme 1.441 un $y_n \in \mathbb{Q}$ tel que

$$x_n - \frac{1}{n} < y_n < x_n + \frac{1}{n}. \tag{1.652}$$

(1) **(y_n) est une suite de Cauchy dans $\mathbb{Q}$** Nous prouvons que (y_n) est une suite de Cauchy dans $\mathbb{Q}$ (définition 1.379(3)). Vu que (x_n) est de Cauchy pour le corps $\mathbb{R}$, si $\epsilon > 0$ dans $\mathbb{R}$ est donné, il existe n_ϵ tel que si $p, q \geqslant n_\epsilon$, alors $|x_p - x_q| < \epsilon$.

Nous avons :

$$|y_p - y_q| \leqslant |y_p - x_p| + |x_p - x_q| + |x_q - y_q| < \frac{1}{p} + \epsilon + \frac{1}{q}. \tag{1.653}$$

En choisissant $N_\epsilon > \max\{n_\epsilon, \frac{1}{\epsilon}\}$ (ce qui est possible par le lemme 1.438), et en prenant $p, q > N_\epsilon$, nous avons

$$|y_p - y_q| \leqslant 3\epsilon, \tag{1.654}$$

ce qui prouve que (y_p) est une suite de Cauchy dans $\mathbb{Q}$, pour la notion de suite de Cauchy dans $\mathbb{Q}$.

(2) **Le réel représenté**

Puisque (y_p) est de Cauchy dans $\mathbb{Q}$, elle représente un réel que nous notons $\bar{y}$.

(3) **Convergence de (x_n)**

Nous prouvons que $x_n \xrightarrow{\mathbb{R}} \bar{y}$.

Nous savons qu'une suite de Cauchy de rationnels converge dans $\mathbb{R}$ vers le réel qu'elle représente, c'est-à-dire : $y_n \xrightarrow{\mathbb{R}} \bar{y}$ où chaque $y_n \in \mathbb{Q}$ est vu comme la suite constante (cela est le lemme 1.453). Autrement dit, pour $\epsilon > 0$, il existe un $N_\epsilon \in \mathbb{N}$ tel que si $p > N_\epsilon$ alors $|\bar{y} - y_p| < \epsilon$. Pour un tel p nous avons

$$|\bar{y} - x_p| \leqslant |\bar{y} - y_p| + |y_p - x_p| \leqslant \epsilon + \frac{1}{p}. \tag{1.655}$$

Donc dès que p est plus grand que $\max\{N_\epsilon, \frac{1}{\epsilon}\}$, nous avons $|\bar{y} - x_p| < 2\epsilon$, ce qui signifie que la suite (x_n) converge vers $\bar{y}$ dans $\mathbb{R}$.

Ceci achève de prouver que $\mathbb{R}$ est un corps complet.

En ce qui concerne l'équivalence entre les suites convergentes et de Cauchy, nous venons de prouver que toute suite de Cauchy dans $\mathbb{R}$ est convergente. La réciproque est la proposition 1.406. □

Nous avons terminé avec la construction des réels. Les propriétés topologiques arrivent en la section 9.3. En particulier le théorème 7.287 pour la complétude de $\mathbb{R}$ en tant qu'espace métrique.

1.23.4 Intervalles

Nous avons déjà défini la notion d'intervalle pour un espace totalement ordonné en 1.21. Nous posons quelques notations dans $\mathbb{R}$.

DEFooAQBUooKLChOW

Définition 1.457.
Soient $a \neq b$ dans $\mathbb{R}$. Nous définissons les parties suivantes de $\mathbb{R}$:

(1) $]a, b[= \{x \in \mathbb{R} \text{ tel que } a < x < b\}$

(2) $[a, b[= \{x \in \mathbb{R} \text{ tel que } a \leqslant x < b\}$

(3) $]a, b] = \{x \in \mathbb{R} \text{ tel que } a < x \leqslant b\}$

(4) $[a, b] = \{x \in \mathbb{R} \text{ tel que } a \leqslant x \leqslant b\}$

(5) $]-\infty, a] = \{x \in \mathbb{R} \text{ tel que } x \leqslant a\}$

(6) $]-\infty, a[= \{x \in \mathbb{R} \text{ tel que } x < a\}$

(7) $]a, \infty[= \{x \in \mathbb{R} \text{ tel que } x > a\}$

(8) $[a, \infty[= \{x \in \mathbb{R} \text{ tel que } x \geqslant a\}$.

(9) $]-\infty, \infty[= \mathbb{R}$.

La proposition 1.466 nous dira que tous les intervalles de $\mathbb{R}$ sont d'une de ces formes.

1.23.5 Maximum, supremum et compagnie

Ce n'est un secret pour personne que $\mathbb{R}$ est un ensemble totalement ordonné[180] : il y a des éléments plus grands que d'autres, et mieux : à chaque fois que je prends deux éléments différents dans $\mathbb{R}$, il y en a un des deux qui est plus grand que l'autre. Il n'y a pas d'*ex æquo* dans $\mathbb{R}$.

La définition du maximum d'une partie est la définition 1.14.

Proposition 1.458.
Soient $x_1, \dots, x_k \in \mathbb{R}$. Alors nous avons

$$\frac{1}{\min(x_i)} = \max\left(\frac{1}{x_i}\right). \tag{1.656}$$

Démonstration. En effet notons $y = \min(x_i)$. Nous avons alors $1/x_i \leqslant 1/y$ pour tout i. □

Définition 1.459.
Soit A, une partie de $\mathbb{R}$.

(1) *Un nombre M est un **majorant** de A si M est plus grand que tous les éléments de A : pour tout $x \in A$, $M \geqslant x$.*

(2) *Un nombre m est un **minorant** de A si m est plus petit que tous les éléments de A : pour tout $x \in A$, $m \leqslant x$.*

Nous parlons de majorant ou de minorants stricts *lorsque les inégalités sont strictes.*

Nous insistons sur le fait que l'inégalité n'est pas stricte. Ainsi, 1 est un majorant de $[0, 1]$. Dès qu'un ensemble a un majorant, il en a plein. Si s majore l'ensemble A, alors $s + 1$, $s + 4$, et $s + \frac{3}{7}$ majorent également A.

Exemple 1.460.
Une petite galerie d'exemples de majorants.

— L'intervalle fermé $[4, 8]$ admet entre autres 8 et 130 comme majorants,

— l'intervalle ouvert $]4, 8[$ admet également 8 et 130 comme majorants,

— 7 n'est pas un majorant de $[1, 5] \cup]8, 32]$,

— 10/10 majore les notes qu'on peut obtenir à un devoir.

— l'intervalle $[4, \infty[$ n'a pas de majorant.

△

DefSupeA

Proposition-Définition 1.461 (Least-upper-bound property[79])**.**
Soit A une partie majorée de $\mathbb{R}$. Il existe un unique élément $M \in \mathbb{R}$ tel que

(1) *$M \geqslant x$ pour tout $x \in A$,*

(2) *pour tout ε, le nombre $M - \varepsilon$ n'est pas un majorant de A, c'est-à-dire qu'il existe un élément $x \in A$ tel que $x > M - \varepsilon$.*

*Cet élément est nommé **supremum** de A et est noté* $\sup(A)$*. De la même façon, **l'infimum** de A, noté* $\inf A$ *est l'unique réel $m \in \mathbb{R}$ vérifiant*

180. Proposition 1.430.

(1) $m \leqslant a$ *pour tout* $a \in A$,

(2) pour tout $\epsilon > 0$, *le nombre* $m + \epsilon$ *n'est pas un minorant : il existe* $a \in A$ *tel que* $a \leqslant m + \epsilon$.

Par convention, si la partie n'est pas bornée vers le haut, nous dirons que son supremum n'existe pas, ou bien qu'il est égal à $+\infty$, suivant les contextes. Pour votre culture générale, sachez toutefois que $\infty \notin \mathbb{R}$.

Démonstration. Nous faisons la preuve pour l'infimum.

(1) **Unicité**
En ce qui concerne l'unicité, soient m_1 et m_2, deux infimums de A. Supposons $m_1 > m_2$. Alors il existe $\epsilon > 0$ tel que $m_2 < m_2 + \epsilon < m_1$ (c'est le lemme 1.441). Cela prouve que $m_2 + \epsilon$ est un minorant de A et donc que m_2 n'est pas un infimum.

(2) **Existence**
Soit A, une partie de $\mathbb{R}$. Nous allons trouver son infimum en suivant une méthode de dichotomie. Pour cela nous allons construire trois suites en même temps de la façon suivante. D'abord nous choisissons un point x_0 de A et un point x_1 qui minore A (qui existe par hypothèse) :

$$\begin{aligned} &x_0 \text{ est un élément de } A,\\ &x_1 \text{ est un minorant de } A,\\ &a_0 = x_0\\ &b_0 = x_1\\ &b_1 = x_1. \end{aligned} \tag{1.657}$$

Ensuite, nous faisons la récurrence suivante :

$$\begin{aligned} x_{n+1} &= \frac{a_n + b_n}{2},\\ a_{n+1} &= \begin{cases} a_n & \text{si } x_{n+1} \text{ minore } A\\ x_{n+1} & \text{sinon,} \end{cases}\\ b_{n+1} &= \begin{cases} x_{n+1} & \text{si } x_{n+1} \text{ minore } A\\ b_n & \text{sinon.} \end{cases} \end{aligned} \tag{1.658}$$

Nous allons montrer que (a_n) et (b_n) sont des suites convergentes de même limite et que cette limite est l'infimum de A.
Soit $n \in \mathbb{N}$; il y a deux possibilités. Soit $a_n = a_{n-1}$ et $b_n = x_n$, soit $a_n = x_n$ et $b_n = b_{n-1}$. Supposons que nous soyons dans le premier cas (le second se traite de façon similaire). Alors nous avons

$$\begin{aligned} |a_n - b_n| &= |a_{n-1} - x_n|\\ &= \left| a_{n-1} - \frac{a_{n-1} + b_{n-1}}{2} \right|\\ &= \frac{1}{2}|a_{n-1} - b_{n-1}|, \end{aligned} \tag{1.659}$$

ce qui prouve que $|a_n - b_n| \to 0$. Nous montrons maintenant que la suite (a_n) est de Cauchy. En effet nous avons

$$|a_n - a_{n-1}| = \begin{cases} 0\\ \left|\frac{a_n - b_n}{2}\right| \end{cases} \leqslant \frac{1}{2n}. \tag{1.660}$$

Il en est de même pour la suite (b_n). Ce sont deux suites de Cauchy (donc convergentes par la proposition 1.406) qui convergent vers la même limite. Soit ℓ cette limite.
Le nombre ℓ minore A. En effet si $a \in A$ est plus petit que ℓ, les éléments b_n tels que $|b_n - \ell| < |a - \ell|$ ne peuvent pas minorer A. D'autre part, pour tout ϵ, le nombre $\ell + \epsilon$ ne peut pas minorer A. En effet, ℓ est la limite de la suite décroissante (a_n), donc il existe a_n entre ℓ et $\ell + \epsilon$. Mais a_n ne minore pas A, donc $\ell + \epsilon$ ne minore pas non plus A.
Nous avons prouvé que toute partie minorée de $\mathbb{R}$ possède un infimum.

La preuve que toute partie majorée possède un supremum se fait de la même façon. □

LEMooSSVKooDPhSkq

Lemme 1.462.
Soit une partie A de $\mathbb{R}$. Si M est un majorant de A, alors $M \geqslant \sup(A)$.

Démonstration. Si $M < \sup(A)$, alors en posant $\epsilon = \sup(A) - M$, le nombre $\sup(A) - \epsilon$ est encore un majorant de A, ce qui est impossible par définition d'un supremum. □

PROPooGOPOooYyZSuU

Proposition 1.463 ([1]).
Soit $A \subset \mathbb{R}$. Alors

ITEMooAQNRooXFnsrA

(1) Si $\lambda \geqslant 0$, nous avons $\inf(\lambda A) = \lambda \inf(A)$.

ITEMooPCDXooRJraYV

(2) Si $\lambda \geqslant 0$, nous avons $\sup(\lambda A) = \lambda \sup(A)$.

ITEMooHPBEooGaKRDf

(3) Si $\lambda \leqslant 0$, nous avons $\inf(\lambda A) = \lambda \sup(A)$.

(1.661)

Démonstration. En plusieurs parties. Vérification des conditions de la définition 1.461.

(1) **Pour (1)** Nous posons $m = \inf(A)$, et nous prouvons que $\lambda m = \inf(\lambda A)$. Si $a \in A$, alors $m \leqslant a$ et donc $\lambda m \leqslant \lambda a$ par la proposition 1.436. Donc λm minore λA.
Vu que $m = \inf(A)$, si $\epsilon > 0$, il existe $a \in A$ tel que $m + \epsilon/\lambda \geqslant a$. Donc, en multipliant par λ, $\lambda m + \epsilon \geqslant \lambda a$. Et donc $\lambda m + \epsilon$ ne minore pas λA.

(2) **Pour (2)** Même raisonnement que pour (1) [181].

(3) **Pour (3) avec $\lambda = -1$** Nous posons $s = \sup(A)$, et nous pouvons que $-s = \inf(-A)$. Soit $a \in A$. Nous avons $s \geqslant a$, et donc $-s \leqslant -a$, de telle sorte que $-s$ minore $-A$.
Soit $\epsilon > 0$. Nous voulons un $a \in A$ tel que $-s + \epsilon \geqslant a$. Vu que $s = \sup(A)$, il existe $a \in A$ tel que $a > s - \epsilon$, et donc, en multipliant par -1 : $-a < -s + \epsilon$.

(4) **Pour (3), cas général** Nous posons $\lambda = -\lambda'$ avec $\lambda' > 0$. Nous avons alors

$$\begin{aligned} \inf(\lambda A) &= \inf\big(-(\lambda' A)\big) && && (1.662a)\\ &= -\sup(\lambda' A) && \text{cas } \lambda = -1 && (1.662b)\\ &= -\lambda' \sup(A) && \text{par (2)} && (1.662c)\\ &= \lambda \sup(A). && && (1.662d) \end{aligned}$$

□

LEMooTRJZooSTuHWs

Lemme 1.464.
Soit une partie bornée A dans $\mathbb{R}$. Si (x_n) est une suite dans A convergente vers $\sup(A)$, alors il existe une sous-suite croissante qui converge également vers $\sup(A)$.

1.23.5.1 Intervalles

LEMooRMUCooMKiTGr

Lemme 1.465 ([1]).
Soit une partie A de $\mathbb{R}$.

ITEMooIQECooFjJFKz

(1) Nous supposons que A admette un supremum qui n'est pas dans A. Si x est un élément de A strictement plus petit que $\sup(A)$, alors il existe $y \in A$ tel que $x < y < \sup(A)$.

(2) Nous supposons que A admette un infimum qui n'est pas dans A. Si x est un élément de A strictement plus grand que $\inf(A)$, alors il existe $y \in A$ tel que $\inf(A) < y < x$.

181. J'ai pas vraiment vérifié. Écrivez-moi si ça pose problème.

Démonstration. Soit $\epsilon > 0$. Par définition 1.461 d'un supremum, le nombre $\sup(A) - \epsilon$ n'est pas un majorant de A. Autrement dit, il existe $y \in A$ tel que $y > \sup(A) - \epsilon$. Vu que $\sup(A)$ est plus grand que tous les éléments de A et qu'il n'est pas lui-même dans A, nous avons aussi $\sup(A) > y$. En mettant bout à bout :

$$x < \sup(A) - \epsilon < y < \sup(A). \tag{1.663}$$

□

PROPooHPMWooQJXCAS

Proposition 1.466 ([1]).
Tous les intervalles[182] *de* $\mathbb{R}$ *sont d'une des formes listées dans la définition 1.457.*

Démonstration. Il y a beaucoup de cas, et nous ne les feront pas tous[183].

(1) **Si I est borné vers le haut et vers le bas** Il y a 4 possibilités suivant que $\inf(I)$ et $\sup(I)$ soient ou non dans I.

(2) **Si $\inf(I) \in I$ et $\sup(I) \in I$** Nous prouvons que $I = [\inf(I), \sup(I)]$.

(2a) **Dans un sens** Si $x \in I$ nous avons $\inf(I) \leqslant x \leqslant \sup(I)$ parce que $\inf(I)$ est un minorant de toute élément de I alors que $\sup(I)$ est un majorant de tout élément de I. Donc $x \in [\inf(I), \sup(I)]$.

(2b) **Dans l'autre sens** Soit $\inf(I) \leqslant x \leqslant \sup(I)$. Vu que I est un intervalle et que $\sup(I)$ et $\inf(I)$ sont deux éléments de I, nous avons $x \in I$.

(3) $\inf(I) \in I$ **et** $\sup(I) \notin I$ Nous allons démontrer que $I = [\inf(I), \sup(I)[$.

(3a) **Dans un sens** Soit $x \in I$. Nous savons (par définition de l'infimum et du supremum) que $\inf(I) \leqslant x \leqslant \sup(x)$. Mais nous sommes dans un cas où $\sup(I) \neq x$ parce que $\sup(I)$ n'est pas dans I. Donc

$$\inf(I) \leqslant x < \sup(I), \tag{1.664}$$

ce qui montre que $I \subset [\inf(I), \sup(I)[$.

(3b) **Dans l'autre sens** Soit $x \in [\inf(I), \sup(I)[$. Nous utilisons le lemme 1.465(1) : il existe $y \in I$ tel que $\inf(I) \leqslant x < y < \sup(I)$. Vu que $\inf(I) \in I$, que $y \in I$ et que I est un intervalle, nous avons aussi $x \in I$ et donc $I \subset [\inf(I), \sup(I)[$.

(4) $\inf(I) \notin I$ **et** $\sup(I) \in I$ C'est le même raisonnement, mais en utilisant l'autre partie du lemme.

(5) $\inf(I) \notin I$ **et** $\sup(I) \notin I$ Nous prouvons que $I =]\inf(I), \sup(I)[$.

(5a) **Dans un sens** Nous avons

$$I \subset \{x \in \mathbb{R} \text{ tel que } \inf(I) \leqslant x \leqslant \sup(I)\}. \tag{1.665}$$

Mais comme $\inf(I)$ et $\sup(I)$ ne sont pas dans I, nous pouvons les enlever dans le membre de droite :

$$I \subset \{x \in \mathbb{R} \text{ tel que } \inf(I) < x < \sup(I)\} =]\inf(I), \sup(I)[. \tag{1.666}$$

(5b) **Dans l'autre sens** Si $x \in I$, nous avons $\inf(I) < x < \sup(I)$. En utilisant les deux parties du lemme, nous avons y_1 et y_2 dans I tels que

$$\inf(I) < y_1 < x < y_2 < \sup(I). \tag{1.667}$$

Vu que I est un intervalle, nous en déduisons que $x \in I$.

(6) $\inf(I) = -\infty$, $\sup(I) \in I$ Nous prouvons que $I =]-\infty, \sup(I)]$.

(6a) **Dans un sens** L'inclusion $I \subset \{x \in \mathbb{R} \text{ tel que } x \leqslant \sup(I)\}$ est automatique : $\sup(I)$ majore tous les éléments de I.

(6b) **Dans l'autre sens** Soit $x \leqslant \sup(I)$. Vu que I n'a pas d'infimum, il existe $y \in I$ tel que $y < x < \sup(I)$. Comme I est un intervalle nous déduisons que $x \in I$.

Je vous laisse voir les autres cas. □

182. Définition 1.21.
183. Si vous avez un doute, écrivez-moi.

1.23.5.2 Quelques exemples

En matière de notations, le maximum de l'ensemble A est noté $\max A$, le supremum est noté $\sup A$. Le minimum et l'infimum sont notés $\min A$ et $\inf A$.

Exemple 1.467.
Exemples de différence entre majorant, supremum et maximum.

— Le nombre 10 est un supremum, majorant et maximum de l'intervalle fermé $[0, 10]$,

— Le nombre 10 est un majorant et un supremum, mais pas un maximum de l'intervalle ouvert $]0, 10[$,

— Le nombre 136 est un majorant, mais ni un maximum ni un supremum de l'intervalle $[0, 10]$.

△

En utilisant les notations concises, ces différents cas s'écrivent ainsi :

$$10 = \max[0, 10] = \sup[0, 10] \quad 10 = \sup[0, 10[\tag{1.668}$$

Exemple 1.468.
Si on dit qu'un pont s'effondre à partir d'une charge de 10 tonnes, alors 10 tonnes est un *supremum* des charges que le pont peut supporter : si on met $9,999999$ tonnes dessus, il tient encore le coup, mais si on ajoute un gramme, alors il s'effondre (on sort de l'ensemble des charges acceptables). △

Exemple 1.469.
Si on dit qu'un pont résiste jusqu'à 10 tonnes, alors 10 tonnes est un *maximum* de la charge acceptable. Sur ce pont-ci, on peut ajouter le dernier gramme. Mais à partir de là, le moindre truc qu'on ajoute, il s'effondre. △

LEMooWCUXooFqTwDK

Lemme 1.470.
À propos de bornes d'un intervalle dans $\mathbb{R}$.

(1) La borne inférieure[184] *d'un intervalle est son infimum,*

(2) la borne supérieure est le supremum.

(3) Si de plus l'intervalle est fermé, l'infimum est un minimum et le supremum est un maximum.

Démonstration. Soit I un des intervalles $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$. Nous allons montrer que dans tous ces cas, a est l'infimum de I.

Nous allons prouver que dans tous ces cas, $a = \bigwedge(I)$ en vérifiant les deux conditions de la définition 1.461. D'abord par définition d'un intervalle, pour tout $t \in I$ nous avons $t \geqslant a$. Première condition vérifiée. Ensuite, nous prenons $\epsilon > 0$. Si $a + \epsilon \geqslant b$, alors n'importe quel élément de I est plus petit que $a + \epsilon$. Si $a + \epsilon < b$, alors nous avons

$$a < a + \epsilon < b. \tag{1.669}$$

Dans ce cas, le lemme 1.441 donne un réel t tel que $a < t < a + \epsilon < b$. Donc $t \in I$ et $a + \epsilon$ n'est pas un minorant de I. □

Exemple 1.471.
Quelques exemples dans les intervalles.

(1) $A = [1, 2]$. Tous les nombres plus petits ou égaux à 1 sont minorants, 1 est infimum et minimum. Le nombre 2 est un majorant, le maximum et le supremum.

(2) $B =]3, \pi[$. Le nombre π est le supremum et est un majorant, mais n'est pas le maximum (parce que $\pi \notin B$). L'ensemble B n'a pas de maximum. Bien entendu, -1000 est un minorant.

Dans les deux cas, le nombre 53 est un majorant. △

184. Ici par « borne inférieure » nous entendons le a dans les intervalles du type $]a, b]$, $[a, b]$, etc.

Il existe évidemment de nombreux exemples plus vicieux.

Exemple 1.472.
Prenons $E = \{\frac{1}{n} \text{ tel que } n \in \mathbb{N}_0\}$, dont les premiers points sont indiqués sur la figure 1.1. Cet ensemble est constitué des nombres 1, $\frac{1}{2}$, $\frac{1}{3}$, ...Le plus grand d'entre eux est 1 parce que tous les nombres de la forme $\frac{1}{n}$ avec $n \geqslant 1$ sont plus petits ou égaux à 1. Le nombre 1 est donc maximum de E.

L'ensemble E n'a par contre pas de minimum parce que tout élément de E s'écrit $\frac{1}{n}$ pour un certain n et est plus grand que $\frac{1}{n+1}$ qui est également dans E.

Prouvons que zéro est l'infimum de E. D'abord, tous les éléments de E sont strictement positifs, donc zéro est certainement un minorant de E. Ensuite, nous savons que pour tout $\varepsilon > 0$, il existe un n tel que $\frac{1}{n}$ est plus petit que ε. L'ensemble E possède donc un élément plus petit que $0 + \varepsilon$, et zéro est bien l'infimum. △

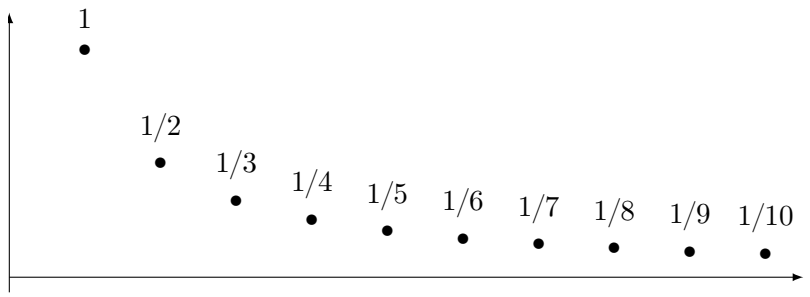


FIGURE 1.1: Les premiers points du type $x_n = 1/n$. LabelFigSuiteUnSurn

L'exemple suivant est une source classique d'erreurs en ce qui concerne l'infimum.

Exemple 1.473.
Les premiers points de l'ensemble $F = \{\frac{(-1)^n}{n} \text{ tel que } n \in \mathbb{N}_0\}$ sont représentés à la figure 1.2. Bien que (comme nous le verrons plus tard) la limite de la suite $x_n = (-1)^n/n$ soit zéro, il n'est pas correct de dire que zéro est l'infimum de l'ensemble F. Le dessin, au contraire, montre bien que -1 est le minium (aucun point est plus bas que -1), tandis que le maximum est $1/2$.

Nous reviendrons avec cet exemple dans la suite. Pour l'instant, ayez bien en tête que zéro n'est rien de spécial pour l'ensemble F en ce qui concerne les notions de maximum, minimum et compagnie. △

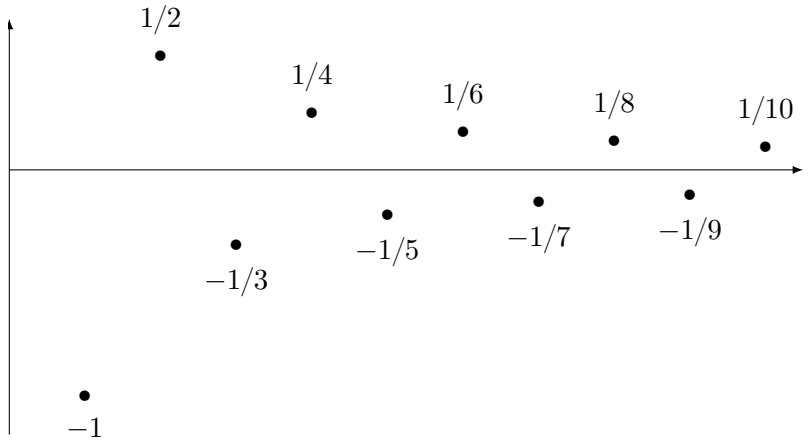


FIGURE 1.2: Les quelques premiers points du type $(-1)^n/n$. LabelFigSuiteInverseAlterne

1.23.6 Racines

SUBSECooMBCNooEqjjTY

Dans cette section, nous définissons $\sqrt{x}$ pour $x \in \mathbb{Q}^+$. Vous notez que c'est fait de façon assez algébrique [185], ou en tout cas, en restant proche des définitions. Des définitions plus technologiques utilisant la continuité de $x \mapsto x^n$ et qui prouvent que l'application est bijective sur un domaine choisi avec prudence existent (voir la définition 12.381). Il est même expliqué dans [74] que la méthode décrite ici permet de définir $\sqrt[n]{x}$ pour tout n entier, et pas seulement pour $n = 2$.

PROPooUHKFooVKmpte

Proposition 1.474.
Soit $q \in \mathbb{Q}^+$. Il existe un unique $r \in \mathbb{R}$ tel que $r^2 = q$.

Plus précisément, en termes des notations de 1.418, pour tout $q \in \mathbb{Q}^+$, il existe un unique $r \in \mathbb{R}^+$ tel que $r^2 = \varphi(q)$.

Démonstration. En deux parties : d'abord l'existence et ensuite l'unicité.

(1) **Existence** Si $q = 0$, c'est $r = 0$. Nous supposons $q > 0$. La suite (x_k) de la proposition 1.413 a la propriété d'être de Cauchy dans $\mathbb{Q}$. Donc il existe un réel r qui est la classe de cette suite. Nous posons donc

$$r = \bar{x}. \tag{1.670}$$

En ce qui concerne r^2, nous avons, par définition du produit dans $\mathbb{R}$,

$$r^2 = \bar{x}^2 = \overline{(x_k^2)}, \tag{1.671}$$

EQooPHLFooAZhebM

c'est la classe de la suite de Cauchy donnée par les x_k^2. Posons $y_k = x_k^2$; la relation (1.671) s'écrit

$$r^2 = \bar{y}. \tag{1.672}$$

La proposition 1.413 nous dit également que y est une suite de Cauchy et que

$$y_k \xrightarrow{\mathbb{Q}} q \tag{1.673}$$

La proposition 1.454 donne alors $\bar{y} = \bar{q}$, et finalement

$$r^2 = \bar{q} = \varphi(q). \tag{1.674}$$

Ici tout n'est pas encore terminé avec l'existence parce qu'il faut nous assurer que $r \geqslant 0$. Ce n'est pas très compliqué : si $r < 0$, alors nous pouvons faire le choix $-r$ qui convient tout aussi bien : $(-r)^2 = r^2$.

(2) **Unicité** Supposons $r_1, r_2 \in \mathbb{R}$ tels que $r_1^2 = r_2^2$. La proposition 1.430 dit que $\mathbb{R}$ est totalement ordonné ; disons pour fixer les idées que $r_1 \leqslant r_2$. Cela signifie, par définition de l'ordre sur $\mathbb{R}$, que $r_2 - r_1 \geqslant 0$. En posant $s = r_2 - r_1$ nous avons $r_2 = r_1 + s$. Passons au carré ; la distribution dans le calcul suivant provient du fait que $\mathbb{R}$ est un corps :

$$r_2^2 = (r_1 + s)^2 = r_1^2 + 2r_1 s + s^2. \tag{1.675}$$

Vu que $r_1^2 = q = r_2^2$, nous avons $2r_1 s + s^2 = 0$ ou encore

$$s(2r_1 + s) = 0. \tag{1.676}$$

Puisque $\mathbb{R}$ est un corps, c'est un anneau intègre [186] et la règle du produit nul s'applique : soit $s = 0$, soit $2r_2 + s = 0$. Puisque $r_2 > 0$ et que $s \geqslant 0$, nous avons $2r_2 + s > 0$ et donc $s = 0$. Nous en déduisons que $r_1 = r_2$.

□

185. Discutable parce que des limites sont utilisées.
186. Lemme 1.198.

1.23.7 Corps valué

DEFooBWXXooAkBBRS

Définition 1.475 (Valeur absolue, corps valué[80, 81])**.**
Soit un corps $\mathbb{K}$*. Une* ***valeur absolue*** *sur* $\mathbb{K}$ *est une application* $|.|: \mathbb{K} \to \mathbb{R}^+$ *telle que*

(1) $|x| = 0$ *si et seulement si* $x = 0$*,*

(2) $|x + y| \leqslant |x| + |y|$

(3) $|xy| \leqslant |x||y|$.

Un corps muni d'une valeur absolue est un ***corps valué****.*

Un corps valué sera un espace topologique métrique dans la définition 7.182. Dans le cas d'un corps totalement ordonné, nous avons une valeur absolue donnée par 1.379(2) et les principales propriétés dans le lemme 1.383.

1.23.8 Partie entière, partie fractionnaire

LEMooLEXTooGAQxGB

Lemme-Définition 1.476 ([82])**.**
Pour tout réel x*, il existe un unique entier* n *vérifiant*

$$n \leqslant x < n + 1. \tag{1.677}$$

Dans ce cas nous avons $x - n \in [0, 1[$.

Le nombre n *ainsi défini est la* ***partie entière*** *de* x*, et il sera noté* $\text{int}(x)$*. Le nombre* $x - \text{int}(x)$ *est la* ***partie fractionnaire*** *de* x *que nous notons* $\text{frac}(x)$.

Il est toutefois à noter que la partie fractionnaire de x *n'est pas garantie d'être une fraction : elle pourrait être irrationnelle. Cette dénomination est donc un peu trompeuse.*

LEMooMMSAooFiRkQd

Lemme 1.477.
Si $\lambda \in \mathbb{R}$ *n'est pas un entier, alors* $\text{int}(\lambda) + 2 > \lambda$.

1.24 Les complexes

La notion de module d'un nombre complexe $|z|$ sera donnée beaucoup plus tard, dans le lemme 9.100. La raison est que le module demande la racine carrée.

Définition 1.478 (Nombres complexes[83])**.**
L'ensemble des ***nombres complexes*** $\mathbb{C}$ *est l'ensemble* $\mathbb{R}^2$ *muni des opérations suivantes :*

(1)

$$\begin{aligned} \times_{\mathbb{C}}: \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ \big((x, y), (x', y')\big) &\mapsto (xx' - yy', xy' + yx') \end{aligned} \tag{1.678}$$

EQooIJWOooZBiKEW

(2)

$$\begin{aligned} +_{\mathbb{C}}: \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ \big((x, y), (x', y')\big) &\mapsto (x + x', y + y') \end{aligned} \tag{1.679}$$

(3)

$$\begin{aligned} \cdot : \mathbb{R} \times \mathbb{C} &\to \mathbb{C} \\ \big(\lambda, (x, y)\big) &\mapsto (\lambda x, \lambda y). \end{aligned} \tag{1.680}$$

Lemme 1.479.
Le triplet $(\mathbb{C}, +_{\mathbb{C}}, \times_{\mathbb{C}})$ *est un anneau*[187] *commutatif dont le neutre pour l'addition est* $(0, 0)$ *et le neutre pour la multiplication est* $(1, 0)$.

187. Définition 1.42.

Démonstration. Ce ne sont que des calculs. Juste pour vous montrer, voici la première partie pour l'associativité :

$$\big((a,b)(x,y)\big)(s,t) = (ax-by, ay+bx)(s,t) \tag{1.681a}$$
$$= (axs-bys-ayt-bxt, axt-byt+ays+bxs). \tag{1.681b}$$

Nous avons utilisé la distributivité sur $\mathbb{R}$, provenant du fait que $\mathbb{R}$ est un corps par le théorème 1.417. □

Lemme 1.480.
L'anneau $\mathbb{C}$ est un corps.

Démonstration. Il suffit de trouver un inverse pour chaque élément non nul. Soit un élément non nul $(a,b) \in \mathbb{C}$. En combinant les lemmes 1.448 et 1.449 nous savons que $a^2+b^2>0$. En particulier, cet élément est inversible dans $\mathbb{R}$, et nous pouvons considérer l'élément suivant de $\mathbb{C}$:

$$z = \big(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\big). \tag{1.682}$$

Prouver que $z(a,b)=(1,0)$ est maintenant juste un calcul. □

Lemme 1.481.
L'application

$$\begin{aligned}\varphi\colon \mathbb{R} &\to \mathbb{C}\\ x &\mapsto (x,0)\end{aligned} \tag{1.683}$$

est un morphisme d'anneaux[188].

Démonstration. Simples calculs. Par exemple

$$\varphi(xx') = (xx',0) = (x,0)(x',0) = \varphi(x)\varphi(x'). \tag{1.684}$$

□

1.482.
Admirez ...

— Un nombre complexe est un couple de réels.
— Un réel est une classe d'équivalence de suites de Cauchy de rationnels.
— Une suite de Cauchy de rationnels est une application $\mathbb{N} \to \mathbb{Q}$ vérifiant certaines propriétés.
— Un rationnel est une classe d'équivalences d'éléments de $\mathbb{Z}$.
— Un élément de $\mathbb{Z}$ est une classe d'équivalence de couples de naturels.
— Un naturel sera ...là c'est plus compliqué. Une construction vraiment rigoureuse des naturels risque d'être en dehors du cadre du Frido.

Bref, les objets que nous manipulons sont d'une effroyable complexité.

1.483.
À partir de maintenant, lorsque nous parlons de $\mathbb{R}$, nous parlons en réalité de $\varphi(\mathbb{R}) \subset \mathbb{C}$.

Lemme 1.484.
Nous avons $(0,1)^2 = (-1,0)$. Nous notons $i=(0,1)$.

Démonstration. Calcul direct à partir de la définition 1.678. □

La proposition suivante donne la forme cartésienne des nombres complexes. Pour la forme trigonométrique, il faudra attendre la proposition 18.62.

188. Définition 1.43.

PROPooKQHLooMFxNLe

Proposition 1.485 ([1])**.**
L'application

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{C} \\ (a,b) &\mapsto a\varphi(1)+bi \end{aligned} \tag{1.685}$$

est un isomorphisme de $\mathbb{R}$-module[189].

Cette proposition permet d'écrire tout nombre complexe sous la forme $a+bi$ pour des réels a et b.

DEFooQDDVooRYDsAJ

Définition 1.486.
*Si $z=a+bi$ est un nombre complexe (avec $a,b\in\mathbb{R}$), son **complexe conjugué** est le nombre $a-bi$.*

L'étude de la série géométrique est reportée à (beaucoup) plus tard, à la proposition 10.135. Dans l'immédiat il nous est possible de calculer la somme partielle.

LEMooAFSCooWEVlvp

Lemme 1.487 (Somme partielle de la série géométrique)**.**
Soit $q\in\mathbb{C}$. Nous avons

$$\sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}. \tag{1.686}$$

Démonstration. Posons $S_N=1+q+\ldots+q^N$. Nous avons évidemment $S_N-qS_N=1-q^{N+1}$ et donc

$$S_N=\sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}. \tag{1.687}$$

EqASYTiCK

□

189. Module, définition 1.330.

Chapitre 2

Théorie des groupes

Pour rappel, la notion de groupe est définie en 1.37.

2.1 Groupe dérivé

DEFooVHZAooUgmesE

Définition 2.1.
*Si G est un groupe et si $g, h \in G$, nous notons $[g,h] = ghg^{-1}h^{-1}$ le **commutateur** de g et h.*

L'élément neutre est toujours un commutateur : pour $g = h$, $[g,g] = ggg^{-1}g^{-1} = e$.

DEFooBNLPooShKYXa

Définition 2.2.
*Le **groupe dérivé** de G est le sous-groupe noté $D(G)$ ou $[G,G]$ engendré*[1] *par les commutateurs.*

Autrement dit, $D(G)$ est l'intersection de tous les sous-groupes de G contenant tous les commutateurs. Le groupe $D(G)$ contient toujours au moins le neutre parce que c'est un groupe.

En vertu du lemme 1.323, le groupe dérivé de G est l'ensemble des produits finis de commutateurs. C'est-à-dire que si S_m est l'ensemble des produits de m commutateurs, alors

$$D(G) = \bigcup_{m=1}^{\infty} S_m. \tag{2.1}$$

LemMMOCooDJJJhy

Lemme 2.3.
Le groupe dérivé est un sous-groupe caractéristique[2]*, et un sous-groupe normal*[3]*.*

Démonstration. Il est évident que si $\alpha \in \mathrm{Aut}(G)$ alors

$$\alpha\big([g,h]\big) = \big[\alpha(g), \alpha(h)\big], \tag{2.2}$$

c'est-à-dire que $D(G)$ est un sous-groupe caractéristique. En particulier si c est un commutateur, alors xcx^{-1} en est encore un, ce qui montre que $D(G)$ est normal dans G. Plus spécifiquement,

$$x(ghg^{-1}h^{-1})x^{-1} = (xgx^{-1})(xhx^{-1})(xg^{-1}x^{-1})(xh^{-1}x^{-1}) \tag{2.3a}$$
$$= (xgx^{-1})(xhx^{-1})(xgx^{-1})^{-1}(xhx^{-1})^{-1}. \tag{2.3b}$$

□

PropAPRGooHBkELf

Proposition 2.4.
Le groupe quotient $G/D(G)$ est abélien.

1. Définition 1.318.
2. Définition 1.173.
3. Définition 1.172.

Démonstration. En ce qui concerne le fait que $G/D(G)$ soit abélien, nous savons que pour tout $g, h \in G$ nous avons $h^{-1}g^{-1}hg \in D(G)$ et donc

$$[g][h] = [gh] = [ghh^{-1}g^{-1}hg] = [hg] = [h][g]. \tag{2.4}$$

□

Le groupe quotient $G/D(G)$ est appelé l'**abélianisé** de G et est parfois noté G^{ab}.

Si $f : G \to A$ est un morphisme entre le groupe G et un groupe abélien A, alors $f\big(D(G)\big) = \{0\}$. Du coup f passe au quotient de G par $D(G)$, et il existe une unique application $\bar{f} : G/D(G) \to A$ telle que $f = \bar{f} \circ \pi$ où $\pi : G \to G/D(G)$ est la projection canonique.

2.2 Théorèmes d'isomorphismes

DEFooWBIYooGNRYOp

Définition 2.5.
*Soient un groupe G, un ensemble X et une application $f : X \to G$. Le **noyau** de f est la partie*

$$\ker(f) = \{x \in X \text{ tel que } f(x) = e\} \tag{2.5}$$

où e est l'élément neutre de G.

Si G est un groupe et si N est un sous-groupe normal, alors l'ensemble G/N a une structure de groupe et la projection canonique $\pi : G \to G/N$ est un morphisme surjectif de noyau N.

ThoPremierthoisomo

Théorème 2.6 (Premier théorème d'isomorphisme)**.**
Soit $\theta : G \to H$ un morphisme de groupe. Alors

(1) $\ker(\theta)$ *est normal dans G,*

(2) $\operatorname{Image}\theta$ *est un sous-groupe de H*

ItemWLCLdk

(3) nous avons un isomorphisme

$$\frac{G}{\ker(\theta)} \simeq \operatorname{Image}\theta \tag{2.6}$$

Démonstration. Point par point.

(1) Le fait que $\ker(\theta)$ soit un sous-groupe de G est clair ; montrons qu'il est normal. Si $g \in G$ et $u \in \ker(\theta)$, alors $\theta(g^{-1}ug) = \theta(g^{-1})\theta(u)\theta(g) = \big(\theta(g)\big)^{-1}\theta(g) = 1_H$, et donc $g^{-1}ug \in \ker(\theta)$.

(2) Il suffit de remarquer que si $h = \theta(g)$ et $h' = \theta(g')$, alors $h^{-1}h' = \theta(g^{-1}g')$.

(3) Si $[g]$ représente la classe de g dans $G/\ker(\theta)$, l'isomorphisme est donné par $\varphi([g]) = \theta(g)$.

□

Ce premier théorème d'isomorphismes permet entre autres de prouver que $SO(3) = \mathrm{SU}(2)/\mathbb{Z}_2$, voir la proposition 55.47.

THOooURXUooQJvkjx

Théorème 2.7 (Deuxième théorème d'isomorphisme)**.**
Soient H et N deux sous-groupes de G et supposons que N soit normal[4]*. Alors*

(1) $NH = HN$ est un sous-groupe.

(2) Le groupe N est normal dans NH.

(3) Le groupe $N \cap H$ est normal dans H.

ItemjRPajc

(4) Nous avons l'isomorphisme

$$\frac{HN}{N} \simeq \frac{H}{H \cap N}. \tag{2.7}$$

4. Si N n'est pas normal, il y aura la proposition 2.8.

Démonstration. Point par point.

(1) Il est clair que $1_G \in NH$. Soient nh et $n'h'$ deux éléments de NH ; alors en tenant compte du fait que N est normal,

$$nhn'h' = n\underbrace{hn'h^{-1}}_{\in N}hh' \in NH. \tag{2.8}$$

Cela prouve que NH est un groupe.

De la même façon, nous prouvons que HN est un groupe par

$$hnh'n' = hh'\underbrace{h'^{-1}nh'}_{\in N}n' \in HN \tag{2.9}$$

Nous devons encore prouver que $HN = NH$. Pour cela, $nh \in HN$, car $nh = hh^{-1}nh$, les trois derniers facteurs formant un élément de N par normalité ; de même $hn \in NH$, montrant que $NH = HN$. Enfin, comme $(nh)^{-1} = h^{-1}n^{-1}$, les inverses de NH sont dans $HN = NH$.

(2) N est normal dans G, a fortiori dans l'un de ses sous-groupes.

(3) Il suffit de voir que, si $h \in H$ et $n \in N \cap H$, alors $hnh^{-1} \in N \cap H$. Or, $hnh^{-1} \in H$ puisque H est un sous-groupe ; et $hnh^{-1} \in N$ car N est un sous-groupe normal de G.

(4) Il faut d'abord remarquer que H et N étant des groupes et le produit NH étant un groupe, nous avons $NH = HN$. Soit le morphisme injectif

$$\begin{aligned} j\colon H &\to HN \\ h &\mapsto h \end{aligned} \tag{2.10}$$

et la surjection canonique

$$\sigma\colon HN \to HN/N \tag{2.11}$$

Nous considérons ensuite l'application composée

$$\begin{aligned} f\colon H &\to HN/N \\ h &\mapsto hN. \end{aligned} \tag{2.12}$$

(4a) **f est surjective** L'application f est surjective parce que l'élément $hnN \in HN/N$ est l'image de h, étant donné que $hnN = hN$.

(4b) $\ker(f) = H \cap N$ Si $a \in H \cap N$, nous avons $f(a) = aN = N$, et donc $H \cap N \subset \ker(f)$. D'autre part, si $h \in H$ vérifie $h \in \ker(f)$, alors $f(h) = hN = N$, ce qui est uniquement possible lorsque $h \in N$.

Le premier théorème d'isomorphisme[5] implique alors que $H/\ker(f) \simeq \text{Image}\, f$, c'est-à-dire

$$H/N \cap H \simeq HN/N. \tag{2.13}$$

□

PROPooVBGMooPTlyLF

Proposition 2.8 (Deuxième théorème d'isomorphisme (suite))**.**
Soient N et H des sous-groupes de G. Si H normalise N, c'est-à-dire si $hNh^{-1} \in N$ pour tout $h \in H$, alors nous avons l'isomorphisme

$$\frac{HN}{N} \simeq \frac{H}{H \cap N}. \tag{2.14}$$

ThoezgBep

Théorème 2.9 (Troisième théorème d'isomorphisme)**.**
Soient N et M deux sous-groupes normaux de G avec $M \subset N$. Alors N/M est normal dans G/M et

$$(G/M)/(N/M) \simeq G/N. \tag{2.15}$$

5. Théorème 2.6.

Démonstration. Afin de montrer que N/M est normal dans G/M, nous considérons $g \in G$, $nM \in N/M$ et nous calculons

$$gnMg^{-1} = gng^{-1}\underbrace{gMg^{-1}}_{=M} = \underbrace{gng^{-1}}_{\in N}M \in N/M. \tag{2.16}$$

Pour prouver l'isomorphisme nous considérons le morphisme

$$\begin{aligned}\varphi\colon G/M &\to G/N\\ gM &\mapsto gN.\end{aligned} \tag{2.17}$$

Ce morphisme est surjectif et son noyau est N/M, parce que $\varphi(gM) = N$ uniquement si $g \in N$. Nous pouvons appliquer le premier théorème d'isomorphisme à φ en écrivant

$$(G/M)/\ker(\varphi) \simeq \text{Image}\,\varphi, \tag{2.18}$$

c'est-à-dire

$$(G/M)/(N/M) \simeq G/N. \tag{2.19}$$

□

2.3 Indice d'un sous-groupe et ordre des éléments

LEMooFNVRooRCkjLc

Lemme 2.10.
Lorsque H est normal dans G, alors la définition

EQooEUESooSeUWHK

$$[a] \cdot [b] = [ab] \tag{2.20}$$

définit une loi de groupe sur l'ensemble G/H.

Démonstration. Le neutre est $[e]$ et l'associativité ne pose pas plus de problème que l'existence d'un inverse. Le point à vérifier est que la formule (2.20) est une bonne définition : $[ah] \cdot [bh'] = [ab]$ pour tout $h, h' \in H$. Nous avons :

$$[ah] \cdot [bh'] = [ahbh'] = [ahb]. \tag{2.21}$$

Pour montrer que c'est $[ab]$, l'astuce est d'introduire bb^{-1} à côté du a :

$$[ahb] = [abb^{-1}hb] = [ab] \tag{2.22}$$

parce que $b^{-1}hb \in H$ du fait que H soit normal dans G. □

EXooFNIKooHxePSs

Exemple 2.11 ([84]).
Il ne faudrait pas croire que le groupe quotient G/H est forcément un sous-groupe de G. Par exemple le quotient $\mathbb{Z}/2\mathbb{Z}$ est l'ensemble $\{0, 1\}$ muni de l'addition. En particulier $1 + 1 = 0$, ce qui est évidemment faux dans $\mathbb{Z}$. Le groupe $(\mathbb{Z}, +)$ ne possède aucun élément d'ordre 2.

Il n'en est pas moins vrai que l'application

$$\begin{aligned}f\colon G &\to G/H\\ g &\mapsto [g]\end{aligned} \tag{2.23}$$

est un morphisme de groupes. △

DEFooMPIAooIeZNaR

Définition 2.12.
*Si H est un sous-groupe d'un groupe fini, l'**indice** de H dans G est le nombre $|G|/|H|$, souvent noté $|G : H|$.*

Le théorème de Lagrange dira en particulier que l'indice est toujours un nombre entier. C'est à ne pas confondre avec le degré d'une extension de corps (définition 6.61).

ThoLagrange

Théorème 2.13 (Théorème de Lagrange)**.**
Soit H un sous-groupe du groupe fini G. Alors ITEMooDPKSooNpOusd

(1) L'ordre de H divise l'ordre de G.

(2) Les trois nombres suivants sont égaux :

- *le nombre de classes de H à gauche,*
- *le nombre de classes de H à droite,*
- *l'indice de H dans G.*

En particulier si H est distingué dans G, nous avons

$$|G/H| = \frac{|G|}{|H|}. \tag{2.24}$$

Démonstration. Nous commençons par montrer que les classes de H ont toutes le même nombre d'éléments que H. En effet pour chaque $g \in G$ nous avons la bijection

$$\begin{aligned} \varphi\colon H &\to gH \\ h &\mapsto gh. \end{aligned} \tag{2.25}$$

L'injectivité de φ est le fait que $gh = gh'$ implique $h = h'$. La surjectivité est par définition de la classe.

Les classes à gauche formant une partition de G, le cardinal de G est le produit de la taille des classes par le nombre de classes :

$$|G| = |H| \cdot \text{ nombre de classes.} \tag{2.26}$$

En particulier nous voyons que $|H|$ divise $|G|$.

La dernière formule exprime simplement que G/H est par définition le nombre de classes de H à gauche (ou à droite) dans G. □

CorpZItFX

Corolaire 2.14.
L'ordre d'un élément[6] *d'un groupe fini divise l'ordre du groupe. En particulier dans un groupe d'ordre n tous les éléments vérifient $g^n = e$.*

Démonstration. Soit G un groupe fini et considérons, à $g \in G$ fixé, le sous-groupe

$$H = \{g^k \text{ tel que } k \in \mathbb{N}\}. \tag{2.27}$$

Par le théorème de Lagrange 2.13, l'ordre de H divise $|G|$, mais l'ordre de H est le plus petit k tel que $g^k = e$, c'est-à-dire l'ordre de g. □

D'autres résultats à propos d'ordres et d'indices de groupes finis dans la proposition 3.30 et le lemme 3.32. En particulier le théorème de Cauchy 3.26 qui dit : si le groupe est cyclique et si p divise l'ordre du groupe G, alors G contient au moins un élément d'ordre p.

2.4 Suite de composition

DefJWZSooNcntfK

Définition 2.15 (Suite de composition)**.**
Soit un groupe G.

*(1) Une **suite de composition** dans G est une suite finie de sous-groupes $(G_i)_{i=0,\dots,n}$ telle que*

$$\{e\} = G_n \subseteq G_{n-1} \subseteq \dots \subseteq G_1 \subseteq G_0 = G \tag{2.28}$$

et telle que G_{i+1} est normal[7] *dans G_i.*

6. Ordre d'un élément, définition 1.268.
7. Nous rappelons au cas où, que « normal » signifie « distingué ».

(2) *Les groupes G_i/G_{i+1} sont les* ***quotients*** *de la suite de composition.*

(3) *Une suite de* ***Jordan-Hölder*** *est une suite de composition dont tous les quotients sont simples.*

L'objet de nos prochaines pérégrinations mathématiques est de montrer que tout groupe fini admet une suite de Jordan-Hölder (théorème 2.21).

LemsKpXCG

Lemme 2.16 (du papillon ou de Zassenhaus[85])**.**
Soient G un groupe et des sous-groupes A et B. Soient A' normal dans A et B' normal dans B. Alors

(1) *$A'(A\cap B')$ est normal dans $A'(A\cap B)$*

(2) *$(A'\cap B)B'$ est normal dans $(A\cap B)B'$*

(3) *Nous avons les isomorphismes de groupes*

$$\frac{A'(A\cap B)}{A'(A\cap B')}\simeq\frac{(A\cap B)B'}{(A'\cap B)B'}\simeq\frac{B'(A\cap B)}{B'(A'\cap B)}.\tag{2.29}$$

Démonstration. Nous n'allons pas démontrer chacun des points ; pour plus de détails, nous dirons simplement que « la preuve est très similaire dans les autres cas ».

Commençons par montrer que $A'(A\cap B')$ est un groupe. Si $a,b\in A'$ et $x,y\in A\cap B'$,

$$axby=xx^{-1}axbx^{-1}xy\tag{2.30}$$

En utilisant la normalité, $x^{-1}ax\in A'$, donc $xx^{-1}axbx^{-1}\in A'$ et donc le tout est dans $A'(A\cap B')$. L'ensemble $A'(A\cap B')$ est également stable pour l'inverse parce que

$$x^{-1}a^{-1}=\underbrace{x^{-1}a^{-1}x}_{\in A'}x^{-1}.\tag{2.31}$$

Nous montrons maintenant que $A'(A\cap B')$ est normal dans $A'(A\cap B)$. Soient $a,b\in A'$, $x\in A\cap B'$ et $f\in A\cap B$. Alors

$$\begin{aligned}(bf)^{-1}(ax)(bf)&=(bf)^{-1}(a\underbrace{xbx^{-1}}_{=c\in A'}xf)&(2.32a)\\&=f^{-1}b^{-1}acxf&(2.32b)\\&=f^{-1}b^{-1}acf\underbrace{f^{-1}xf}_{=y\in A\cap B'}&(2.32c)\\&=\underbrace{f^{-1}b^{-1}acf}_{\in A'}y&(2.32d)\\&\in A'(A\cap B').&(2.32e)\end{aligned}$$

Pour prouver l'isomorphisme

$$\frac{A'(A\cap B)}{A'(A\cap B')}\simeq\frac{(A\cap B)B'}{(A'\cap B)B'},\tag{2.33}$$

nous allons utiliser le deuxième théorème d'isomorphisme (2.8) que nous appliquons à $H=A\cap B$ et $N=A'(A\cap B')$. La vérification que H normalise N est usuelle. Nous commençons par écrire

$$\frac{A'(A\cap B')(A\cap B)}{A'(A\cap B')}\simeq\frac{A\cap B}{A\cap B\cap A'(A\cap B')}.\tag{2.34}$$

EqkphNsE

Pour simplifier un peu cette expression nous prouvons d'abord que

$$(A\cap B)\cap A'(A\cap B')=(A'\cap B)(A\cap B').\tag{2.35}$$

EqkhsyNh

L'inclusion $\supset$ est facile. Pour l'autre sens, étant donné que $A'(A\cap B')\subset A$ nous avons

$$A\cap B\cap A'(A\cap B)=B\cap A'(A\cap B). \tag{2.36}$$

EQooEEVIooHCbasF

Un élément de $B\cap A'(A\cap B)$ est un élément de B qui s'écrit sous la forme $s=ax$ avec $a\in A'$ et $x\in A\cap B$. Nous avons alors $a=sx^{-1}$ avec $s\in B$ et $x^{-1}\in A\cap B$. Par conséquent $a\in B$ et donc $a\in A'\cap B$. Donc un élément de $B\cap A'(A\cap B)$ s'écrit sous la forme ax avec $a\in A'\cap B$ et $x\in A\cap B$. Autrement dit

$$B\cap A'(A\cap B)\subset(A'\cap B)(A\cap B) \tag{2.37}$$

et nous avons

$$(A\cap B)\cap A'(A\cap B)=B\cap A'(A\cap B)\subset(A'\cap B)(A\cap B'), \tag{2.38}$$

et donc l'égalité (2.35). Toujours dans l'idée de simplifier (2.34) nous remarquons que $A\cap B'$ est un sous-ensemble de $A\cap B$, donc $A'(A\cap B')(A\cap B)=A'(A\cap B)$. Il reste donc

$$\frac{A'(A\cap B)}{A'(A\cap B')}=\frac{A\cap B}{(A'\cap B)(A\cap B')}. \tag{2.39}$$

Étant donné que les hypothèses sur A et B sont symétriques, le membre de droite peut aussi s'écrire en inversant A et B. Nous en sommes à

$$\frac{B'(A\cap B)}{B'(A'\cap B)}=\frac{A'(A\cap B)}{A'(A\cap B')}. \tag{2.40}$$

Nous devons encore justifier $B'(A\cap B)=(A\cap B)B'$ et $B'(A'\cap B)=(A'\cap B)B'$. Vérifions la première égalité, et laissons la seconde au lecteur. Si $b\in B'$ et $x\in A\cap B$, alors

$$bx=x\underbrace{x^{-1}bx}_{\in B'}\in(A\cap B)B'. \tag{2.41}$$

□

PROPooUIXNooJdKuAs

Proposition 2.17.
Si G est un groupe fini et si (G_i) est une suite de composition pour G, alors l'ordre de G est le produit des ordres de ses quotients.

Démonstration. Étant donné que G_{i+1} est toujours normal dans G_i, le théorème de Lagrange 2.13 s'applique et, à chaque pas de la suite de composition, nous avons :

$$\left|\frac{G_i}{G_{i+1}}\right|=\frac{|G_i|}{|G_{i+1}|}. \tag{2.42}$$

Il suffit maintenant d'écrire $|G|$ de façon télescopique :

$$|G|=\prod_{0\leqslant i\leqslant n-1}\frac{|G_i|}{|G_{i+1}|} \tag{2.43}$$

□

DEFooSZZSooCrgDyo

Définition 2.18.
*Nous disons que les deux suites de composition $(G_i)_{0\leqslant i\leqslant r}$ et $(H_j)_{0\leqslant j\leqslant s}$ sont **équivalentes** si $r=s$ et si il existe une permutation $\sigma\in S_{r-1}$ telle que*

$$\frac{G_i}{G_{i+1}}\simeq\frac{H_{\sigma(i)}}{H_{\sigma(i)+1}}. \tag{2.44}$$

Proposition 2.19 (Schreider)**.**
Deux suites de composition d'un même groupe admettent des raffinements équivalents.

Démonstration. Soient les suites de composition

$$\{e\} = G_m \subseteq \ldots \subseteq G_1 \subseteq G_0 = G \tag{2.45a}$$
$$\{e\} = H_n \subseteq \ldots \subseteq H_1 \subseteq H_0 = G \tag{2.45b}$$

Nous raffinons la suite (G_i) en ajoutant, entre G_{i+1} et G_i, les sous-groupes

$$G_{i+1} = G_{i+1}(G_i \cap H_n) \subset G_{i+1}(G_i \cap H_{n-1}) \subseteq \ldots \subseteq G_{i+1}(G_i \cap H_0) = G_i, \tag{2.46}$$

et de même pour (H_j). Le groupe $G_{i+1}(G_i \cap H_k)$ est normal dans $G_{i+1}(G_i \cap H_{k-1})$ parce que G_{i+1} étant normal dans G_i, et H_k dans H_{k-1}, le lemme 2.16 s'applique. Nous avons donc bien défini un raffinement.

Nous devons maintenant prouver que les deux raffinements ainsi construits sont des suites de composition équivalentes. D'abord elles ont la même longueur mn parce que chacun des m éléments de la suite (G_i) a été remplacé par n éléments et inversement, chacun des n éléments de la suite (H_j) a été remplacé par m éléments.

Par ailleurs, les quotients du raffinement de (G_i) sont de la forme

$$\frac{G_{i+1}(G_i \cap H_k)}{G_{i+1}(G_i \cap H_{k+1})} \simeq \frac{H_{k+1}(H_k \cap G_i)}{H_{k+1}(H_k \cap G_{i+1})} \tag{2.47}$$ EqPAYTCR

en vertu du lemme du papillon (2.16). Le membre de droite de (2.47) est un des quotients du raffinement de (H_j). □

LemBSicRJ

Lemme 2.20 (Schreider strictement décroissant)**.**
Soient Σ_1 et Σ_2, deux suites de composition strictement décroissantes du groupe G. Alors elles admettent des raffinements équivalents strictement décroissants.

Démonstration. Par hypothèse, Σ_1 et Σ_2 n'ont pas de répétitions. Soient Σ_1'' et Σ_2'', des raffinements équivalents donnés par le lemme de Schreider. Étant donné que ce sont des suites de composition équivalentes, elles ont le même nombre de quotients réduits à $\{e\}$, c'est-à-dire le même nombre de répétitions.

Les suites Σ_1' et Σ_2' obtenues en retirant les répétitions de Σ_1'' et Σ_2'' sont des raffinements équivalents de Σ_1 et Σ_2 et strictement décroissants. □

ThoLgxWIC

Théorème 2.21 (Jordan-Hölder)**.**
À propos de suites de Jordan-Hölder dans un groupe fini.

ITEMooRSDDooNHkFYO

(1) Tout groupe fini admet une suite de Jordan-Hölder.

ITEMooGBOCooBAgnyt

(2) Toutes les suites de Jordan-Hölder dans un groupe fini sont équivalentes.

Démonstration. En deux parties.

(1) **Pour (1)**

(2) **Pour (2)**
Par définition, une suite de Jordan-Hölder n'a pas de raffinement strictement décroissant (à part elle-même) parce que G_{i+1} est normal maximum dans G_i. Si Σ_1 et Σ_2 sont des suites de Jordan-Hölder nous pouvons considérer les raffinements équivalents strictement décroissants Σ_1' et Σ_2' du lemme de Schreider 2.20. Nous avons $\Sigma_1' \sim \Sigma_2'$, mais par ce que nous venons de dire à propos de la maximalité, $\Sigma_1' = \Sigma_1$ et $\Sigma_2' = \Sigma_2$. D'où le résultat.

□

2.5 Groupes résolubles

DefOSYNooTROIKs

Définition 2.22 (Groupe résoluble[86, 87])**.**
*Le groupe G est **résoluble** si il existe une suite finie de sous-groupes G_i*

$$\{e\} = G_0 \subset G_1 \subset \ldots \subset G_{n-1} \subset G_n = G \tag{2.48}$$

avec G_i normal dans G_{i+1} et G_{i+1}/G_i abélien.

Il s'agit d'un groupe qui admet une suite de composition[8] dont les quotients sont abéliens.

LemOARMooYhYmbH

Lemme 2.23 ([88])**.**
Soit G un groupe et H un sous-groupe normal. Le groupe G/H est abélien si et seulement si[9] *$D(G) \subset H$.*

Démonstration. Les propositions suivantes sont équivalentes :

— Le groupe G/H est abélien
— pour tout $x, y \in G$, $[x][y] = [y][x]$
— $[x][y][x^{-1}][y^{-1}] = [e]$
— $[xyx^{-1}y^{-1}] = [e]$
— $[x, y] \in H$, voir la définition 2.1.
— $D(G) \subset H$.

□

PropRWYZooTarnmm

Proposition 2.24 ([88])**.**
Un groupe est résoluble si et seulement si sa suite dérivée termine sur $\{e\}$.

Démonstration. Grâce au lemme 2.3 et à la proposition 2.4, si la suite dérivée termine sur $\{e\}$ alors la suite dérivée est une suite qui répond aux conditions de la définition 2.22 de groupe résoluble.

Il faut donc encore montrer le sens direct. Nous supposons que G est un groupe résoluble et nous étudions sa suite dérivée. Nous avons une suite

$$\{e\} = G_0 \subset G_1 \subset \ldots \subset G_{n-1} \subset G_n = G \tag{2.49}$$

avec G_i normal dans G_{i+1} et G_{i+1}/G_i abélien. Nous allons prouver par récurrence que $D^i(G) \subset G_{n-i}$.

Pour $i = 0$ nous avons bien $G \subset G_n$.

Notre hypothèse de récurrence est :

EqEAQEooEaeJEo
$$D^i(G) \subset G_{n-i}. \tag{2.50}$$

Vu que G_{n-i}/G_{n-i-1} est abélien, le lemme 2.23 dit que

EqEDJXooLQLQcr
$$D(G_{n-i}) \subset G_{n-i-1}. \tag{2.51}$$

En dérivant (2.50) et en tenant compte de (2.51),

$$D^{i+1}(G) \subset D(G_{n-i}) \subset G_{n-i-1}. \tag{2.52}$$

Nous avons donc bien $D^i(G) \subset G_{n-i}$ pour tout i. En particulier pour $i = n$, nous trouvons $D^n(G) \subset G_0 = \{e\}$. □

8. Voir définition 2.15.
9. Ici $D(G)$ est le groupe dérivé de G, voir 2.2.

PropBNEZooJMDFIB

Proposition 2.25.
Soient des groupes G et H. Nous supposons que G est résoluble[10] *et nous considérons un morphisme $\psi\colon G\to H$. Alors $\psi(G)$ est résoluble.*

Démonstration. Puisque G est résoluble, il existe une suite de sous-groupes G_i tels que

$$\{e\}=G_0\subset G_1\subset\ldots\subset G_{n-1}\subset G_n=G \tag{2.53}$$

avec G_i normal dans G_{i+1} et G_{i+1}/G_i abélien. Nous posons $\psi(G)_i=\psi(G_i)$ et nous avons $\psi(G)_0=\psi\big(\{e\}\big)=\{e\}$ ainsi que $\psi(G)_0=\psi(G)$; donc

$$\{e\}=\psi(G)_0\subset\psi(G)_1\subset\ldots\subset\psi(G)_{n-1}\subset\psi(G)_n=\psi(G). \tag{2.54}$$

La proposition 1.273 nous indique que $\psi(G)_i$ est normal dans $\psi(G)_{i+1}$, et que $\psi(G)_{i+1}/\psi(G)_i$ est abélien. □

2.6 Action de groupes

Le concept d'action d'un groupe est donné par la définition 1.367

Lemme 2.26.
Pour tout $g\in G$,

(1) L'application $\phi_g\colon E\to E$ est injective,

(2) Pour l'inverse : $(\phi_g)^{-1}=\phi_{g^{-1}}$.

Démonstration. Si $x,y\in E$ sont tels que $\phi_g(x)=\phi_g(y)$ alors en appliquant $\phi_{g^{-1}}$ aux deux membres nous trouvons

$$(\phi_{g^{-1}}\phi_g)(x)=(\phi_{g^{-1}}\phi_g)(y), \tag{2.55}$$

ce qui donne $x=y$ parce que $\phi_{g^{-1}}\phi_g=\phi_{g^{-1}g}=\phi_e=\mathrm{Id}$.

Les trois dernières égalités écrites disent que $\phi_{g^{-1}}$ est l'inverse[11] de ϕ_g. □

Pour alléger les notations, on convient d'écrire $g\cdot x$, voire plus simplement gx au lieu de $\phi_g(x)$. Le deuxième axiome d'action de groupe dit que la notation ghx ne souffre d'aucune ambiguïté.

DEFooMZXFooXbwGjj

Définition 2.27 (Orbite)**.**
*Si G agit sur un ensemble E, nous notons $G\cdot x$ l'**orbite** de $x\in E$ sous l'action de G :*

$$G\cdot x=\{gx\text{ tel que }g\in G\}.$$

DEFooMDYGooLrOERP

Définition 2.28 (Stabilisateur)**.**
*Si G est un groupe agissant sur l'ensemble E, et si $x\in E$, nous notons G_x ou $\mathrm{Stab}(x)$ le **stabilisateur** de x :*

$$\mathrm{Stab}(x)=G_x=\{g\in G\text{ tel que }g\cdot x=x\}. \tag{2.56}$$

Définition 2.29 (Fixateur)**.**
*Si G est un groupe agissant sur l'ensemble E, et si $g\in G$, nous notons enfin $\mathrm{Fix}(g)$ le **fixateur** de g :*

$$\mathrm{Fix}(g)=\{x\in E\text{ tel que }g\cdot x=x\}. \tag{2.57}$$

DefuyYJRh

Définition 2.30.
*L'action de G sur E est **fidèle** si l'élément neutre est le seul élément de G à fixer tous les points de E, c'est-à-dire si*

$$(gx=x\quad\forall x\in E)\Rightarrow g=e. \tag{2.58}$$

10. Définition 2.22.
11. Si vous décidez de dire ça à un jury dans un concours, soyez prêts à préciser les domaines.

Un exemple d'action fidèle tout à fait non trivial sera donné avec l'action du groupe modulaire sur le plan de Poincaré dans le théorème 23.96.

Le groupe G agit toujours sur lui même à gauche et à droite. L'action à gauche est $g \cdot h = gh$; celle à droite est $g \cdot h = hg^{-1}$.

DEFooCORTooEeOLPT

Définition 2.31.
*L'action **adjointe** définie par $g \cdot h = ghg^{-1}$ est une manière pour un groupe d'agir sur lui-même par automorphismes. Cela est souvent noté $\mathbf{Ad}(g)h = ghg^{-1}$.*

En effet pour tout $g \in G$, l'application $\mathbf{Ad}(g)\colon G \to G$ est un automorphisme de G.

Si H est un sous-groupe de G, nous notons G/H le quotient de G par la relation $g \sim gh$ pour tout $h \in H$. Lorsque la distinction est importante, nous noterons $(G/H)_g$ pour les classes à gauche et $(G/H)_d$ pour les classes à droite.

Nous avons une relation d'équivalence à gauche et une à droite. D'abord

$$x \sim_g y \Leftrightarrow xh = y \tag{2.59}$$

pour un certain $h \in H$. Ensuite

$$x \sim_d y \Leftrightarrow hx = y \tag{2.60}$$

pour un certain $h \in H$.

Le lemme suivant est une généralisation du théorème de Lagrange 2.13.

Lemme 2.32.
L'ensemble $(G/H)_g$ est fini si et seulement si l'ensemble $(G/H)_d$ est fini. Si il en est ainsi, alors $(G/H)_g$ et $(G/H)_d$ ont même cardinal, qui vaut l'indice de H dans G.

Démonstration. L'application

$$\begin{aligned} f\colon (G/H)_g &\to (G/H)_d \\ [x]_g &\mapsto [x^{-1}]_d \end{aligned} \tag{2.61}$$

est une bijection bien définie. En effet si $x \sim_g y$, nous avons $h \in H$ tel que $y^{-1}h = x^{-1}$, c'est-à-dire que $x^{-1} \sim_d y^{-1}$ et f est bien définie. Le fait que f soit surjective est évident. Pour l'injectivité, soient $x, y \in G$ tels que

$$f([x]_g) = f([y]_g). \tag{2.62}$$

Alors $x^{-1} \sim_d y^{-1}$, ce qui implique l'existence de $h \in H$ tel que $hx^{-1} = y^{-1}$, ou encore que $xh^{-1} = y$, ce qui signifie que $x \sim_g y$.

Pour l'énoncé à propos de l'indice, nous procédons en plusieurs étapes simples.

(1) Les classes (les éléments de $(G/H)_g$) forment une partition de G.

(2) Toutes les classes ont le même nombre d'éléments par la bijection

$$\begin{aligned} f\colon [x]_g &\to [y]_g \\ xh &\mapsto yh. \end{aligned} \tag{2.63}$$

(3) Le nombre d'éléments dans une classe est égal à $|H|$ par la bijection

$$\begin{aligned} g\colon [x]_g &\to H \\ xh &\mapsto h. \end{aligned} \tag{2.64}$$

Par conséquent

$$|G| = |H| \cdot \text{nombre de classes} = |H| \cdot \text{Card}\big((G/H)_g\big), \tag{2.65}$$

et nous avons bien

$$\text{Card}\big((G/H)_g\big) = \frac{|G|}{|H|} = |G : H|. \tag{2.66}$$

□

Propszymlr

Proposition 2.33 (Orbite-stabilisateur[59]).
Soient G un groupe agissant sur un ensemble E et $x \in E$.

(1) Les ensembles $G \cdot x$ et G/G_x sont équipotents.

ITEMooCWUGooCOFHYk

(2) L'orbite de x est finie si et seulement si $\mathrm{Stab}(x)$ est d'indice fini dans G. Dans ce cas nous avons

$$\mathrm{Card}(G \cdot x) = |G : \mathrm{Stab}(x)|, \tag{2.67}$$

EqnLCHCE

et

$$|G| = |\mathrm{Stab}(x)||\mathcal{O}_x|. \tag{2.68}$$

EqCewSXT

La formule (2.67) est nommée ***formule des classes****[89].*

Démonstration. En deux points.

(1) Soit l'application

$$\begin{aligned} \psi\colon G \cdot x &\to G/G_x \\ a \cdot x &\mapsto [a]. \end{aligned} \tag{2.69}$$

Cette application est bien définie parce que si $a \cdot x = b \cdot x$, alors il existe $h \in G_x$ tel que $b = ah$, et par conséquent $[a] = [b]$. Cette application est une bijection et par conséquent $G \cdot x$ est équipotent à G/G_x.

(2) Soit $y \in \mathcal{O}_x$ et $A_y = \{g \in G \text{ tel que } g \cdot x = y\}$. Nous avons

$$A_y = s\,\mathrm{Stab}(x). \tag{2.70}$$

En effet, si $a \in A_y$, alors $a \cdot x = y$ et $s^{-1}a \cdot x = s^{-1} \cdot y = x$. Par conséquent $s^{-1}a \in \mathrm{Stab}(x)$. Dans l'autre sens, supposons $a \in s\,\mathrm{Stab}(x)$. Soit $g \in \mathrm{Stab}(x)$ tel que $a = sg$. Alors nous avons $a \cdot x = sg \cdot x = s \cdot x = y$. Donc $a \in A_y$.
Nous venons de prouver que la partie A_y est une classe à gauche de $\mathrm{Stab}(x)$, par conséquent $|A_y| = |\mathrm{Stab}(x)|$ pour tout $y \in \mathcal{O}_x$. Les A_y pour différents y sont disjoints et nous avons de plus

$$\bigcup_{y \in \mathcal{O}_x} A_y = G. \tag{2.71}$$

Les ensembles A_y divisent donc G en $|\mathcal{O}_x|$ paquets de $|\mathrm{Stab}(x)|$ éléments. D'où la formule (2.68).

□

CORooRRVHooTyCjZZ

Corolaire 2.34.
Soit C_g la classe de conjugaison d'un élément g du groupe fini G. Alors

$$\mathrm{Card}(C_g) = |G : Z_G(g)| \tag{2.72}$$

où $Z_G(g)$ est le centralisateur de g dans G[12] *de G.*

Démonstration. C'est une application de la proposition 2.33 (formule (2.67)) dans le cas de l'action adjointe de G sur lui-même.

En effet, si nous considérons l'action adjointe, l'orbite est la classe de conjugaison : $C_g = G \cdot g$. Et le stabilisateur de g pour l'action adjointe n'est autre que le centralisateur de g :

$$\begin{aligned} \mathrm{Fix}(g) &= \{h \in G \text{ tel que } h \cdot g = g\} && \text{(2.73a)} \\ &= \{h \in G \text{ tel que } hgh^{-1} = g\} && \text{(2.73b)} \\ &= \{h \in G \text{ tel que } gh = hg\} && \text{(2.73c)} \\ &= Z_G(g). && \text{(2.73d)} \end{aligned}$$

Donc la formule $\mathrm{Card}(G \cdot g) = |G : G_g|$ devient, dans le cas de l'action adjointe de G sur lui-même : $\mathrm{Card}(C_g) = |G : Z_G(g)|$. □

12. Définition 1.170.

Lemme 2.35.
Soit G un groupe agissant sur l'ensemble E. On définit $x \sim x'$ si et seulement si il existe $g \in G$ tel que $g \cdot x = x'$. Alors

(1) la relation $\sim$ est une relation d'équivalence.

(2) la classe $[x]$ est l'orbite $\mathcal{O}_x$ de x sous G.

CorARFVMP

Corolaire 2.36 (Équation des orbites)**.**
Soient G un groupe agissant sur l'ensemble E et $\mathcal{O}_1, \dots, \mathcal{O}_k$ la liste des orbites (distinctes). Alors

(1) $E = \bigcup_i \mathcal{O}_i$, l'union est disjointe,

(2) $\text{Card}(E) = \sum_i \text{Card}(\mathcal{O}_i)$.

DefcSuYxz

Définition 2.37.
Soit G un groupe agissant sur l'ensemble E. Un ***domaine fondamental*** *ou une* ***transversale*** *est une partie de E contenant un et un seul élément de chaque orbite.*

Autrement dit, les images des éléments d'un domaine fondamental F forment une partition de l'ensemble :

$$E = \bigsqcup_{g \in G} g(F) \tag{2.74}$$

où $g(F) = \phi_g(F) = \{\phi_g(x) \text{ tel que } x \in F\}$. L'union est disjointe, c'est-à-dire que si $g \neq g'$, alors $g(F) \cap g'(F) = \varnothing$.

PropUyLPdp

Proposition 2.38 (Équation des classes[90])**.**
Soit G, un groupe fini opérant sur un ensemble E. Si E'' est un ensemble contenant exactement un élément de chaque orbite dans $E \backslash \text{Fix}_G(E)$, alors

$$|G| = |\text{Fix}_G(E)| + \sum_{x \in E''} \frac{|G|}{|\text{Fix}_G(x)|}. \tag{2.75}$$

EqobuzIK

Si de plus G est un p-groupe, alors

$$|E| = |\text{Fix}_G(E)| \mod p. \tag{2.76}$$

EqbzLEVJ

Démonstration. Par le corolaire 2.36, nous avons $|G| = \sum_{x \in E'} |\mathcal{O}_x|$ où E' est une transversale. En séparant la somme entre les orbites à un élément et les autres,

$$|G| = \text{Card}(\text{Fix}_G(E)) + \sum_{x \in E''} \frac{|G|}{|\text{Fix}_G(x)|} \tag{2.77}$$

EqaggkBs

EqDgYbhm où nous avons utilisé le fait que $|G| = |\text{Fix}_G(x)||\mathcal{O}_x|$.

Si G est un p-groupe alors si $x \in E''$, $\text{Fix}_G(x)$ est un sous-groupe propre de G et donc $|\text{Fix}_G(x)|$ est un diviseur propre de $|G|$. Du coup la fraction $|G|/|\text{Fix}_G(x)|$ est une puissance non nulle de p et l'équation (2.75) devient immédiatement (2.76). □

Corolaire 2.39 (Équation des classes)**.**
Soient G un groupe, et $C_1, \dots, C_l$ la liste de ses classes de conjugaison contenant plus d'un élément. Alors

$$\text{Card}(G) = \text{Card}\big(Z(G)\big) + \sum_i |G : Z_{g_i}| = \text{Card}\big(Z(G)\big) + \sum_i \frac{\text{Card}(G)}{\text{Card}\big(\text{Fix}(g_i)\big)} \tag{2.78}$$

EqkgGmoq

si $g_i \in C_i$.

Démonstration. Étant donné que les classes de conjugaison sont disjointes, le cardinal de G est bien la somme des cardinaux de ses classes. Les classes ne contenant qu'un seul élément sont celles des éléments de $Z(G)$. En ce qui concerne les autres orbites, $\text{Card}(C_{g_i}) = |G : Z_{g_i}|$ par le théorème orbite-stabilisateur (proposition 2.33). □

THOooEFDMooDfosOw

Théorème 2.40 (Formule de Burnside)**.**
Si G est un groupe fini agissant sur l'ensemble fini E et si Ω est l'ensemble des orbites, alors

$$\operatorname{Card}(\Omega) = \frac{1}{|G|}\sum_{g\in G}\operatorname{Card}\big(\operatorname{Fix}(g)\big). \tag{2.79}$$

EqTUsblv

Démonstration. Nous considérons l'ensemble

$$A = \{(g,x) \in G\times E \text{ tel que } gx = x\}, \tag{2.80}$$

et nous en calculons le cardinal de deux façons. D'abord

$$\begin{aligned}
\operatorname{Card}(A) &= \sum_{x\in E}\operatorname{Card}\{g\in G \text{ tel que } gx = x\} && \text{(2.81a)}\\
&= \sum_{x\in E}\operatorname{Card}(\operatorname{Fix}(x)) && \text{(2.81b)}\\
&= \sum_{\omega\in\Omega}\sum_{x\in\omega}\operatorname{Card}(\operatorname{Fix}(x)) && \text{(2.81c)}\\
&= \sum_{\omega\in\Omega}\frac{|G|}{\operatorname{Card}(\omega)} && \text{(2.81d)}\\
&= |G|. && \text{(2.81e)}
\end{aligned}$$

EqyVtkxf

Pour obtenir (2.81d) nous avons utilisé l'équation des classes (2.68). L'autre façon de calculer $\operatorname{Card}(A)$ est de regrouper ainsi :

$$\operatorname{Card}(A) = \sum_{g\in G}\operatorname{Card}\{x\in E \text{ tel que } gx = x\} = \sum_{g\in G}\operatorname{Card}(\operatorname{Fix}(g)). \tag{2.82}$$

En égalisant les deux expressions de $\operatorname{Card}(A)$ nous trouvons

$$|G|\operatorname{Card}(\Omega) = \sum_{g\in G}\operatorname{Card}(\operatorname{Fix}(g)). \tag{2.83}$$

□

PROPooMYKKooLetZWi

Proposition 2.41.
Soient G un groupe, et H, un sous-groupe du centre de G.

(1) H est normal dans G.

(2) Si G/H est monogène, alors G est abélien.

PROPooWPQTooECUPLK

Proposition 2.42 ([91])**.**
L'indice du centre d'un groupe n'est jamais un nombre premier.

LEMooZJGVooRYrujB

Lemme 2.43 ([92])**.**
Soient un groupe cyclique G et un sous groupe H.

ITEMooUNOEooZoElYL

(1) Il existe $r, n \in \mathbb{N}$ tels que $r \mid n$ et $H \simeq r\mathbb{Z}/n\mathbb{Z}$.

ITEMooMUSSooBmNXEd

(2) $\operatorname{Card}(H) = n/r$

ITEMooVPEKooKdPibk

(3) H est cyclique.

En particulier, tout sous-groupe d'un groupe cyclique est cyclique.

Démonstration. En plusieurs parties.

(1) **Pour (1)** Soit un générateur a de G. L'application

$$\begin{aligned} f\colon \mathbb{Z} &\to G \\ k &\mapsto a^k \end{aligned} \tag{2.84}$$

est un morphisme de groupe surjectif.

La partie $f^{-1}(H)$ est un sous-groupe de $\mathbb{Z}$. La proposition 1.235(2) dit qu'il existe $r \in \mathbb{N}$ tel que $f^{-1}(H) = r\mathbb{Z}$. De même $\ker(f)$ est un sous-groupe de $\mathbb{Z}$, de telle sorte qu'il existe $n \in \mathbb{N}$ tel que $\ker(f) = n\mathbb{Z}$.

Nous avons

$$n\mathbb{Z} = \ker(f) = f^{-1}(e) \subset f^{-1}(H) = r\mathbb{Z}. \tag{2.85}$$

Étant donné que $n\mathbb{Z} \subset r\mathbb{Z}$, il existe $d \in \mathbb{N}$ tel que $n = dr$.

Nous considérons maintenant

$$\begin{aligned} g\colon r\mathbb{Z} &\to H \\ k &\mapsto a^k, \end{aligned} \tag{2.86}$$

qui est encore un morphisme de groupes. En utilisant le premier théorème d'isomorphismes 2.6(3), nous avons

$$\frac{r\mathbb{Z}}{\ker(g)} \simeq \text{Image}(g). \tag{2.87}$$

Mais $\ker(g) = n\mathbb{Z}$ et $\text{Image}(g) = H$. Donc nous avons bien $H \simeq r\mathbb{Z}/n\mathbb{Z}$.

(2) **Pour (2)** En ce qui concerne le cardinal, vu que nous avons un isompshisme,

$$\text{Card}(H) = \text{Card}(r\mathbb{Z}/n\mathbb{Z}) = n/r \tag{2.88}$$

par le lemme 1.329(3).

(3) **Pour (3)** Le fait que H soit cyclique est le lemme 1.329(2).

□

THOooRGSTooIWyhqt

Théorème 2.44 (Sous-groupe d'un groupe cyclique[92])**.**
Soit G un groupe cyclique[13] *d'ordre n. Pour chaque $d \in \mathbb{N}$ nous posons*

$$H_d = \{x \in G \text{ tel que } x^d = e\}. \tag{2.89}$$

(1) Ces H_d sont des sous-groupes de G.

(2) $\text{Card}(H_d) = d$.

ITEMooMJCXooCIWSmK

(3) L'ensemble des sous-groupes de G est $\{H_d \text{ tel que } d \mid n\}$.

ITEMooIEGUooPQWEep

(4) Si g est un générateur de G, alors H_d peut être décrit des façons suivantes :

$$H_d = \{x \in G \text{ tel que } x^d = e\} = \text{gr}(g^{n/d}). \tag{2.90}$$

Démonstration. Point par point.

(1) **H_d est un sous-groupe** ITEMooYAMOooVtbCrS Si $g, h \in H_d$, alors $(gh)^d = g^d h^d = ee = e$. De plus $(g^{-1})^d = (g^d)^{-1} = e^{-1} = e$.

(2) **Tout sous-groupe est un H_d avec $d \mid n$** ITEMooXEXFooIrFOTi Soit un sous-groupe H de G. Nous notons $\text{Card}(H) = d$. Le lemme 2.43(3) nous indique que H est cyclique. Soit un générateur $a \in H$; son ordre est le cardinal de H : $\text{ord}(a) = \text{Card}(H) = d$.

Tout élément de H vérifie $x^{\text{ord}(a)} = (a^k)^{\text{ord}(a)} = (a^{\text{ord}(a)})^k = e$. Donc $H \subset H_d$. En particulier $\text{Card}(H) \leqslant \text{Card}(H_d)$.

13. Définition 1.326.

Tout sous-groupe de G est cyclique. En particulier H_d est cyclique, et nous pouvons considérer un générateur : il existe $a \in H_d$ tel que $\text{ord}(a) = \text{Card}(H_d)$. Étant donné que $a \in H_d$, il vérifie $a^d = e$, et donc

$$\text{ord}(a) \mid d. \tag{2.91}$$

Vu que $\text{ord}(a) = \text{Card}(H_d)$ et que $d = \text{Card}(H)$, nous avons $\text{Card}(H_d) \mid \text{Card}(H)$ et en particulier $\text{Card}(H_d) \leqslant \text{Card}(H)$. Donc $\text{Card}(H) = \text{Card}(H_d)$ et au final $H = H_d$.

Nous avons prouvé au passage que $\text{Card}(H_d) = d$. Vu que le cardinal d'un sous-groupe divise le cardinal du groupe [14], nous avons $d \mid n$.

(3) $\text{Card}(H_d) = d$ C'est contenu dans ce que nous avons fait dans la partie (2).

(4) **Si $d \mid n$ alors H_d est un sous-groupe de G** C'est un cas particulier de la partie (1).

(5) **Note** Le point (3) est prouvé. Cela ne contredit pas le fait que H_d soit un sous-groupe de G pour tout d, et pas seulement pour les d qui divisent n. En effet si d ne divise pas n, il existe d' divisant n tel que $H_d = H_{d'}$. Pas besoin de chercher une preuve de ce fait : la preuve est déjà faite. Il s'agit de dire que H_d est un sous-groupe de G pour tout d (partie (1)), et que tous les sous-groupes de G sont de la forme H_d avec $d \mid n$.

(6) **Pour (4)** Soit un générateur g de G. Nous avons d'une part $g^{n/d} \in H_d$ parce que $(g^{n/d})^d = g^n = e$. D'autre part, les éléments

$$g^{n/d}, g^{2n/d}, \dots, g^{(d-1)n/d}, g^n = e \tag{2.92}$$

sont distincts. En effet si $g^{kn/d} = g^{ln/d}$ avec $0 < l < k < d$, alors $g^{(k-l)n/d} = e$. Comme $k - l < d$, nous avons $(k-l)n/d < n$ et donc $g^{(k-l)n/d} \neq e$ parce que g est générateur.

Bref, les $g^{kn/d}$ avec $k = 1, \dots, d$ sont d éléments distincts, tous dans H_d dont le cardinal est d. Donc

$$H_d = \{g^{kn/d}\}_{1,\dots,d}. \tag{2.93}$$

□

DEFooQDHPooCfDEuL

Définition 2.45.
*Soit G un groupe agissant sur un ensemble E. Nous disons que l'action est **transitive** si elle possède une seule orbite. L'action est **libre** si $g \cdot x = g' \cdot x$ implique $g = g'$.*

2.7 Produit semi-direct de groupes

Définition 2.46.
*Une **suite exacte** est une suite d'applications comme suit :*

$$\dots \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \xrightarrow{f_{i+2}} \dots \tag{2.94}$$

où pour chaque i, les applications f_i et f_{i+1} vérifient $\ker(f_{i+1}) = \text{Image}(f_i)$. Lorsque les ensembles A_i sont des groupes, alors nous demandons de plus que les f_i soient des morphismes.

Très souvent nous sommes confrontés à des suites exactes de la forme

$$1 \longrightarrow A \xrightarrow{f} G \xrightarrow{g} B \longrightarrow 1 \tag{2.95}$$

où G, A et B sont des groupes, 1 est l'identité. La première flèche est l'application $\{1\} \to A$ qui à 1 fait correspondre 1. La dernière est l'application $B \to 1$ qui à tous les éléments de B fait correspondre 1. Le noyau de f étant l'image de la première flèche (c'est-à-dire $\{1\}$), l'application f est injective. L'image de g étant le noyau de la dernière flèche (c'est-à-dire B en entier), l'application g est surjective.

14. Théorème de Lagrange 2.13(1).

DEFooKWEHooISNQzi

Définition 2.47.
*Soient N et H deux groupes et un morphisme de groupes $\phi\colon H \to \mathrm{Aut}(N)$. Le **produit semi-direct** de N et H relativement à ϕ, noté $N \times_\phi H$ est l'ensemble $N \times H$ muni de la loi (que l'on vérifiera être de groupe)*

$$(n,h)\cdot(n',h') = (n\phi_h(n'), hh'). \tag{2.96}$$

EqDRgbBI

Attention à l'ordre quelque peu contre-intuitif. Lorsque nous notons $N\times_\phi H$, c'est bien $\phi\colon H \to \mathrm{Aut}(N)$, c'est-à-dire H qui agit sur N et non le contraire.

Lorsque N et H sont des sous-groupes d'un même groupe, le plus souvent ϕ est l'action adjointe définie en 2.31.

Le théorème suivant permet de reconnaitre un produit semi-direct lorsqu'on en voit un.

THOooZNYTooPhnIdE

Théorème 2.48 ([11])**.**
Soit une suite exacte de groupes

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{s} H \longrightarrow 1 \tag{2.97}$$

Si il existe un sous-groupe $\tilde{H}$ de G à partir duquel s est un isomorphisme, alors

$$G \simeq i(N) \times_\sigma \tilde{H} \tag{2.98}$$

où σ est l'action adjointe[15] *de $\tilde{H}$ sur $i(N)$.*

Démonstration. Nous posons $\tilde{N} = i(N)$ et nous allons subdiviser la preuve en petits pas.

(1) $\tilde{N}$ est normal dans G. En effet étant donné que la suite est exacte nous avons $\tilde{N} = \ker(s)$. Le noyau d'un morphisme est toujours un sous-groupe normal.

(2) $\tilde{N} \cap \tilde{H} = \{e\}$. L'application s étant un isomorphisme depuis $\tilde{H}$, il n'y a pas d'éléments de $\tilde{H}$ dans $\ker(s)$ autre que e.

ItemzIaXGM

(3) $G = \tilde{N}\tilde{H}$. Nous considérons $g \in G$ et $h \in \tilde{H}$ tel que $s(g) = s(h)$. L'existence d'un tel h est assurée par le fait que s est surjective depuis $\tilde{H}$. Du coup nous avons $e = s(gh^{-1})$, c'est-à-dire $gh^{-1} \in \ker(s) = \tilde{N}$. Nous avons donc bien la décomposition $g = (gh^{-1})h$, et donc $G = \tilde{N}\tilde{H}$.

ItemUGFjle

(4) L'écriture $g = nh$ avec $n \in \tilde{N}$ et $h \in \tilde{H}$ est unique. Si $nh = n'h'$, alors $n = n'h'h^{-1}$, ce qui signifierait que $h'h^{-1} \in \tilde{N}$. Mais étant donné que $\tilde{H} \cap \tilde{N} = \{e\}$, nous obtenons $h = h'$ et par suite $n = n'$.

ItemUZlrKo

(5) L'application

$$\begin{aligned}\phi\colon G &\to \tilde{N} \times \tilde{H}\\ nh &\mapsto (n,h)\end{aligned} \tag{2.99}$$

est une bijection. C'est une conséquence des points (3) et (4).

(6) Si sur $\tilde{N} \times \tilde{H}$ nous mettons le produit

$$(n,h)\cdot(n',h') = (n\sigma_h n', hh') \tag{2.100}$$

où σ est l'action adjointe du groupe sur lui-même, c'est-à-dire $\sigma_x(y) = xyx^{-1}$, alors ϕ est un isomorphisme. Si $g, g' \in G$ s'écrivent (de façon unique par le point (5)) $g = nh$ et $g' = n'h'$ alors

$$\begin{aligned}\phi(nhn'h') &= \phi(n\underbrace{hn'h^{-1}}_{\in\tilde{N}}hh') && \text{(2.101a)}\\ &= \phi\big((nhn'h^{-1})(hh')\big) && \text{(2.101b)}\\ &= (nhn'h^{-1}, hh') && \text{(2.101c)}\\ &= (n,h)\cdot(n',h') && \text{(2.101d)}\\ &= \phi(nh)\phi(n'h'). && \text{(2.101e)}\end{aligned}$$

15. Le fait que H agisse sur $i(N)$ fait partie du théorème.

□
CoroGohOZ

Corolaire 2.49.
Soit G un groupe, et N, H des sous-groupes de G tels que

(1) H normalise N (c'est-à-dire que $\sigma_h(n) = hnh^{-1} \in N$ pour tout $h \in H$ et $n \in N$ [16]),

(2) $H \cap N = \{e\}$,

(3) $HN = G$.

Alors l'application

$$\begin{aligned} \psi\colon N \times_\sigma H &\to G \\ (n,h) &\mapsto nh \end{aligned} \tag{2.102}$$

est un isomorphisme de groupes.

Démonstration. En plusieurs parties.

(1) ψ **est un morphisme** Simple calcul en utilisant la formule (2.96) du produit dans $N \times_\sigma H$:

$$\psi\big((n,h)(n',h')\big) = \psi\big(n\sigma_h(n'), hh'\big) \tag{2.103a}$$
$$= nhn'h^{-1}hh' \tag{2.103b}$$
$$= nhn'h' \tag{2.103c}$$
$$= \psi(n,h)\psi(n',h'). \tag{2.103d}$$

(2) ψ **est injective** Supposons que $\psi(n,h) = \psi(n',h')$. Alors $nh = n'h'$ et

$$n = n'h'h^{-1}. \tag{2.104}$$

En multipliant par $(n')^{-1}$ nous trouvons

$$h'h^{-1} = (n')^{-1}n \in N. \tag{2.105}$$

Donc $h'h^{-1} \in H \cap N = \{e\}$. Nous en déduisons que $h = h'$ et donc que $n = n'$.

(3) ψ **est surjective** Nous savons que $HN = G$. Soit $g \in G$. Il existe $h \in H$ et $n \in N$ tel que $g^{-1} = h^{-1}n^{-1}$. Avec ces éléments nous avons

$$g = nh = \psi(n,h). \tag{2.106}$$

□

Dans les hypothèses, l'ordre entre N et H est important lorsqu'on dit que c'est N qui agit sur H ; mais l'hypothèse $NH = G$ est équivalente à $HN = G$ (passer à l'inverse pour s'en assurer).

Insistons encore un peu sur la notation : dans $N \times_\sigma H$, c'est H qui agit sur N par σ.

2.8 Groupe de torsion

Soit G un groupe. Un élément $g \in G$ est un **élément de torsion** si il est d'ordre fini. La **torsion** de G est l'ensemble de ses éléments de torsion. Nous disons qu'un groupe est un **groupe de torsion** si tous ses éléments sont de torsion.

Exemple 2.50.
Le groupe additif $\mathbb{Q}/\mathbb{Z}$ est un groupe de torsion parce que si $[x] = [p/q]$, alors $q[x] = [p] = [0]$. △

2.9 Famille presque nulle

Soit $(G,+)$ un groupe abélien et $\mathcal{F} = \{g_i\}_{i\in I}$ une famille d'éléments de G indicés par un ensemble I. Le **support** de $\mathcal{F}$ est l'ensemble $\{i \in I \text{ tel que } g_i \neq 0\}$. La famille est dite **presque nulle** si le support est fini.

Nous disons que $\mathcal{F}$ est une **suite** si $I = \mathbb{N}$.

16. Ou encore que H agit sur N par automorphismes internes.

Chapitre 3

Anneaux

Attention aux conventions. Dans le Frido, un corps peut être réduit à $\{0\}$ et un idéal premier ne peut pas être $\{0\}$. Ces conventions ont une série de conséquences un peu partout, par exemple dans la proposition 1.231 où nous parlons d'idéal maximum propre. Comparez par exemple avec [45]. Soyez attentif.

En cas de doutes, nous suivons les conventions de Wikipédia.

3.1 Inversibles et nilpotents

Le concept d'anneau est la définition 1.42.

LEMooOYZEooLivKWI

Lemme 3.1.
Si a et b commutent, nous avons, pour tout $r \in \mathbb{N}$ et $r > 0$, la formule

$$a^{r+1} - b^{r+1} = (a-b)\left(\sum_{k=0}^{r} a^{r-k}b^k\right). \tag{3.1}$$

Eqarpurmkbk

Démonstration. Démontrons cela par récurrence. Le cas $r = 0$ est évident. Pour un r donné, si (3.1) est vraie, alors

$$\begin{aligned}
a^{r+2} - b^{r+2} &= a^{r+1}a - a^{r+1}b + a^{r+1}b - b^{r+1}b \\
&= a^{r+1}(a-b) + (a^{r+1} - b^{r+1})b \\
&= a^{r+1}(a-b) + (a-b)\left(\sum_{k=0}^{r} a^{r-k}b^k\right)b \\
&= (a-b)\left(a^{r+1} + \left(\sum_{k=0}^{r} a^{r-k}b^k\right)b\right) \\
&= (a-b)\left(a^{r+1} + \sum_{k=0}^{r} a^{r-k}b^{k+1}\right) \\
&= (a-b)\left(a^{r+1} + \sum_{k'=1}^{r+1} a^{(r+1)-k'}b^{k'}\right) \\
&= (a-b)\left(\sum_{k'=0}^{r+1} a^{(r+1)-k'}b^{k'}\right).
\end{aligned}$$

□

Proposition 3.2.
Si a est un élément nilpotent de l'anneau A, alors $1-a$ est inversible. Si a est nilpotent non nul, alors il est diviseur de zéro.

Démonstration. Soit n le minimum tel que $a^n = 0$. En vertu de la formule (3.1) nous avons

$$1 = 1 - a^n = (1-a)(1 + a + \cdots + a^{n-1}) = (1 + a + \cdots + a^{n-1})(1-a). \tag{3.2}$$

La somme $1 + a + \cdots + a^{n-1}$ est donc un inverse de $(1-a)$. □

3.2 PGCD, PPCM et éléments inversibles

La définition de pgcd et ppcm dans un anneau commutatif est la définition 1.186. Dans la plus grande tradition, elle a été introduite sans motivation, et utilisée par-ci par-là. Nous revenons maintenant dessus.

Commençons par donner une autre vision de la divisibilité dans les anneaux intègres.

PropDiviseurIdeaux

Proposition 3.3.
Dans un anneau intègre[1] *A, on a l'équivalence suivante concernant deux éléments $a, b \in A$:*

$$a \mid b \Leftrightarrow (b) \subset (a). \tag{3.3}$$

Donc la divisibilité devient en réalité une relation d'ordre dont nous pouvons chercher un maximum et un minimum. Si S est une partie de A, nous notons $a \mid S$ pour exprimer que $a \mid x$ pour tout $x \in S$; de la même façon, $S \mid b$ signifie que $x \mid b$ pour tout $x \in S$.

Nous rappelons également la définition 1.43 de morphisme d'anneaux. Remarquons que si f est un morphisme, nous avons $f(0) = 0$ et $f(x)^{-1} = f(x^{-1})$.

Lemme 3.4 ([93])**.**
Les éléments inversibles d'un anneau sont diviseurs de tous les éléments.

Démonstration. Soit k inversible d'inverse k' : $kk' = 1$; soit aussi $a \in A$. Alors $a = k(k'a)$, ce qui montre que k divise a. □

Lemme 3.5 ([93])**.**
Dans un anneau, 1 est un pgcd de a et b si et seulement si les seuls diviseurs communs sont les inversibles.

Démonstration. Supposons pour commencer que 1 est un pgcd de a et b. Un diviseur commun de a et b doit donc diviser 1. Or un diviseur de 1 est forcément inversible.

Dans l'autre sens, les diviseurs communs de a et b sont tous inversibles et donc diviseurs de 1. Donc 1 est un pgcd de a et b. □

3.2.1 Calcul effectif du PGCD et théorème de Bézout

subSecIpmnhO

Soient a et b, deux entiers que nous allons prendre positifs. Nous allons voir maintenant l'algorithme de **Euclide étendu** qui est capable, pour a et b donnés, de calculer le PGCD de a et b, et un couple de Bézout (u, v) tel que $ua + vb = \text{pgcd}(a, b)$. Ce calcul est indispensable si on veut implémenter RSA (19.2).

Cela se base sur le lemme suivant.

LemiVqita

Lemme 3.6.
Soient $a, b \in \mathbb{N}$ et des nombres q et r tels que $a = qb + r$. Alors $\text{pgcd}(a, b) = \text{pgcd}(r, b)$.

Démonstration. Il suffit de voir que les diviseurs communs de a et b sont diviseurs de r et que les diviseurs communs de r et b divisent a.

Si s divise a et b, alors dans l'équation

$$\frac{a}{s} = \frac{qb}{s} + \frac{r}{s}$$

1. Définition 1.197.

les termes a/s et qb/s sont entiers, donc r/s est aussi entier, et s divise r.

Inversement, si s divise r et b, alors il divise $qb+r$ et donc a. □

Remarque 3.7.
Ce lemme est surtout intéressant lorsque $a = qb+r$ est la division euclidienne de a par b : en effet, dans ce cas $r < b$, et le calcul du PGCD de deux nombres (a et b) est ramené à un calcul de PGCD de deux nombres plus petits (b et r).

L'algorithme pour calculer $\text{pgcd}(a,b)$ consiste à écrire des divisions euclidiennes successives de la manière suivante :

$$a = q_2 b + r_2 \qquad r_2 < b \tag{3.4a}$$
$$b = q_3 r_2 + r_3 \qquad r_3 < r_2 \tag{3.4b}$$
$$\vdots \tag{3.4c}$$

en remarquant que $\text{pgcd}(a,b) = \text{pgcd}(b,r_2) = \text{pgcd}(r_2,r_3)$. Étant donné que les inégalités $r_2 < b$ et $r_3 < r_2$ sont strictes, en continuant ainsi nous finissons sur zéro, c'est-à-dire qu'il existera un n pour lequel $r_{n+1} = 0$; et donc

$$r_{n-1} = q_{n+1} r_n,$$

et à ce moment nous avons $\text{pgcd}(a,b) = \text{pgcd}(r_{n-1},r_n) = r_n$.

3.2.1.1 Algorithme d'Euclide pour le PGCD

SUBSECooAEBLooFGJRkg

Écrivons l'algorithme[94] en détail (parce que Bézout, ça va être la même chose en cinq fois plus compliqué). On pose

$$r_0 = a \tag{3.5a}$$
$$r_1 = b \tag{3.5b}$$

(ce qui explique que nous n'ayons pas utilisé r_0 et r_1 précédemment). Ensuite on écrit la division euclidienne $a = q_2 b + r_2$, c'est-à-dire $r_0 = q_2 r_1 + r_2$. Cela donne r_2 et q_2 en termes de r_0 et r_1 :

$$r_2 = r_0 - q_2 r_1. \tag{3.6}$$

Ensuite, sachant r_2 nous pouvons continuer :

$$r_1 = q_3 r_2 + r_3 \tag{3.7}$$

donne q_3 et $r_3 = r_1 - q_3 r_2$. On continue avec $r_2 = q_4 r_3 + r_4$. Tout cela pour poser la suite

$$\begin{aligned} r_0 &= a \\ r_1 &= b \\ r_k &= q_{k+2} r_{k+1} + r_{k+2} \end{aligned} \tag{3.8}$$

SUBEQooDYUEooYNQnII

où la troisième équation définit r_{k+2} et q_{k+2} en fonction de r_k et r_{k+1}, à l'aide du théorème de la division euclidienne. La suite (r_k) ainsi construite est strictement décroissante et à chaque étape le lemme 3.6 et le principe de l'algorithme d'Euclide nous donnent

$$\begin{cases} \text{pgcd}(r_k,r_{k+1}) = \text{pgcd}(r_{k+1},r_{k+2}) = \text{pgcd}(a,b) & \text{(3.9a)} \\ 0 \leqslant r_{k+1} < r_k. & \text{(3.9b)} \end{cases}$$

La suite étant strictement décroissante, nous prenons r_n, le dernier non nul : $r_{n+1} = 0$. Dans ce cas, en prenant $k = n-1$ dans la la dernière équation (3.8) devient :

$$r_{n-1} = q_{n+1} r_n \tag{3.10}$$

avec $\text{pgcd}(a,b) = \text{pgcd}(r_n,r_{n-1}) = r_n$.

Exemple 3.8.
Calculons le PGCD de 18 et 231. Pour cela nous écrivons les divisions euclidiennes en chaine :

$$231 = 18 \cdot 12 + 15 \tag{3.11a}$$
$$18 = 1 \cdot 15 + 3 \tag{3.11b}$$
$$15 = 5 \cdot 5 + 0. \tag{3.11c}$$

Donc le PGCD est 3 (le dernier reste non nul). △

3.2.1.2 Algorithme étendu : calcul effectif des coefficients de Bézout

SUBSECooRHSQooEuBWbd

La difficulté est de construire la suite qui donne des coefficients de Bézout. Elle va être construite à l'envers. Nous supposons déjà connaitre la liste complète des r_k jusqu'à $r_n = \text{pgcd}(a, b)$, ainsi que la liste complète des divisions euclidiennes

$$r_k = q_{k+2}r_{k+1} + r_{k+2}. \tag{3.12}$$

Nous voulons trouver les couples (u_k, v_k) de telle façon à avoir à chaque étape

$$r_n = u_k r_k + v_k r_{k-1}. \tag{3.13}$$

Notons que c'est à chaque fois r_n que nous construisons. La première équation de type Bézout à résoudre est

$$r_n = u_n r_n + v_n r_{n-1}, \tag{3.14}$$

sachant que $r_{n-1} = q_{n+1}r_n$. On pose $v_n = 0$ et $u_n = 1$ et c'est bon. Pour la récurrence, supposons les coefficients u_k et v_k connus, et déterminons les coefficients u_{k-1} et v_{k-1}. Pour ce faire, nous égalons les deux expressions pour r_n :

$$r_n = u_k r_k + v_k r_{k-1} = u_{k-1}r_{k-1} + v_{k-1}r_{k-2}; \tag{3.15}$$

dans laquelle nous substituons $r_{k-2} = q_k r_{k-1} + r_k$:

$$u_k r_k + v_k r_{k-1} = u_{k-1}r_{k-1} + v_{k-1}(q_k r_{k-1} + r_k) \tag{3.16}$$
$$= (u_{k-1} + q_k v_{k-1})r_{k-1} + v_{k-1}r_k \tag{3.17}$$

et nous égalons les coefficients de r_k et r_{k-1} pour obtenir

$$\begin{cases} v_{k-1} = u_k & \text{(3.18a)} \\ u_{k-1} = v_k - v_{k-1}q_k. & \text{(3.18b)} \end{cases}$$

Dès que u_k et v_k ainsi que q_k sont connus, on peut calculer u_{k-1} et v_{k-1}.

La dernière équation, celle avec $k = 1$, est

$$r_n = u_1 r_1 + v_1 r_0, \tag{3.19}$$

c'est-à-dire

$$\text{pgcd}(a, b) = u_1 b + v_1 a. \tag{3.20}$$

EqNDMLooDvajAc

Nous avons ainsi trouvé des coefficients de Bézout pour a et b.

3.2.2 Générateurs

3.9.
Les éléments de $\mathbb{Z}/n\mathbb{Z}$ ne sont pas des éléments de $\mathbb{Z}$; ce sont des parties de $\mathbb{Z}$. Pour rappel :

$$[a]_n = \{a + kn \text{ tel que } k \in \mathbb{Z}\}. \tag{3.21}$$

Pour écrire les éléments de $\mathbb{Z}/n\mathbb{Z}$, nous pouvons écrire

$$\{[i]_n\}_{i=1,\dots,n} \tag{3.22}$$

ou

$$\{[i]_n \text{ tel que } 1 \leqslant i \leqslant n\}. \tag{3.23}$$

Mais attention : l'ensemble

$$\bigcup_{i=1}^{n}[i]_n \tag{3.24}$$

est très différent. Ce dernier ensemble est $\mathbb{Z}$.

PROPooEWREooUOSMsE

Proposition 3.10 ([95]).
Soit $n \geqslant 2$.

ITEMooHTRHooDWgXlk

(1) L'ensemble quotient $\mathbb{Z}/n\mathbb{Z}$ contient n éléments

(2) Nous avons $\mathbb{Z}/n\mathbb{Z} = \{[k]_n\}_{k=0,\dots,n-1}$.

Démonstration. Nous montrons que

$$\begin{aligned}\varphi\colon \{1,\dots,n\} &\to \mathbb{Z}/n\mathbb{Z}\\ k &\mapsto [k]_n\end{aligned} \tag{3.25}$$

est une bijection.

(1) **Injective** Supposons que $\varphi(k) = \varphi(l)$ pour $k, l \in \{1,\dots,n\}$. Il existe $t \in \mathbb{Z}$ tel que $l = l + tn$. Si $t > 0$, alors nous avons

$$l + tn \geqslant 1 + n > n, \tag{3.26}$$

ce qui est impossible parce que $k \in \{0,\dots,n\}$. Nous montrons de même que $t < 0$ est impossible. Nous en déduisons que $t = 0$ et donc que $k = l$.

(2) **Surjective** Soit $l \in \mathbb{Z}$. Nous allons montrer que $[l]_n$ est dans l'image de φ. Par la division euclidienne 1.222, il existe $q \in \mathbb{Z}$ et $r < n$ (dans $\mathbb{N}$) tels que $l = qn + r$. Nous venons de prouver que $[l]_n = [r]_n = \varphi(r)$.

Pour le point (2), c'est juste l'image de φ. □

PROPooMTWGooEMbvDi

Proposition 3.11 ([95]).
Soit $n \geqslant 2$ et $m \in \mathbb{Z}$. Nous avons équivalence entre

ITEMooMCIJooItgSHc

(1) $\operatorname{pgcd}(n, m) = 1$,

ITEMooMNKCooPnpyye

(2) $[m]_n$ engendre le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$

ITEMooBPWLooFLdJfr

(3) $[m]_n$ est inversible dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)$.

Démonstration. En trois parties.

(1) **(1) implique (2)**
Si $\operatorname{pgcd}(m, n) = 1$, le théorème de Bézout 1.236 donne des $u, v \in \mathbb{Z}$ tels que $um + vn = 1$, autrement dit $u[m]_n = [1]_n$. Si k est quelconque, nous avons

$$[k]_n = uk[m]_n, \tag{3.27}$$

ce qui signifie que $[k]_n$ est bien un multiple de $[m]_n$ qui est donc générateur.

(2) **(2) implique (3)**
Si $[m]_n$ engendre $\big(\mathbb{Z}/n\mathbb{Z}, +\big)$, il existe en particulier un multiple de $[m]_n$ qui vaut $[1]_n$: il existe $k \in \mathbb{Z}$ tel que $k[m]_n = [1]_n$.
Nous voyons que $[k]_n$ est un inverse de $[m]_n$ dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)$.

(3) **(3) implique (1)**
Nous supposons que $[m]_n$ est inversible dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)$. C'est à dire qu'il existe $k \in \mathbb{Z}$ tel que $[k]_n[m]_n = [1]_n$. Cela signifie qu'il existe $l \in \mathbb{Z}$ tel que

$$km = 1 + ln. \tag{3.28}$$

Par le théorème de Bézout (pris dans l'autre sens), cela signifie que $\text{pgcd}(m, n) = 1$.

□

3.2.3 Décomposition en facteurs premiers

LemPRuUrsD

Lemme 3.12 (Lemme de Gauss)**.**
Soient $a, b, c \in \mathbb{Z}$ tels que a divise bc. Si a est premier avec c, alors a divise b.

Démonstration. Puisque a est premier avec c, nous avons $\text{pgcd}(a, c) = 1$ et le théorème de Bézout 1.236 nous donne donc $s, t \in \mathbb{Z}$ tels que $sa + tc = 1$. En multipliant par b, nous avons $sab + tbc = b$. Mais les deux termes du membre de gauche sont multiples de a parce que a divise bc. Par conséquent b est somme de deux multiples de a et donc est multiple de a. □

LemAXINooOeuMJZ

Lemme 3.13 (Lemme d'Euclide[96])**.**
Soient $a, b \in \mathbb{Z}$. Si le nombre premier p divise le produit ab, alors p divise a ou b.

Démonstration. Comme p est premier, si il ne divise pas a c'est que $\text{pgcd}(a, p) = 1$. Dans ce cas le lemme de Gauss 3.12 implique que p divise b. □

LEMooSRFMooHgEMwj

Lemme 3.14.
Soient $a, b \in \mathbb{N}$ tels que a divise b et b divise a. Alors $a = b$.

Le théorème fondamental de l'arithmétique permet de décomposer des nombres en facteurs premiers.

ThoAJFJooAveRvY

Théorème 3.15 (Décomposition en facteurs premiers[97])**.**
Tout entier strictement positif peut être écrit comme un produit de nombres premiers d'une unique façon, à l'ordre près des facteurs.

En d'autres termes, pour tout entier $n > 1$, il existe une unique suite finie unique $(p_1, k_1), \ldots (p_r, k_r)$ telle que :

(1) les p_i sont des nombres premiers tels que, si $i < j$, alors $p_i < p_j$;

(2) les k_i sont des entiers naturels non nuls ;

(3) $n = \prod_{i=1}^{r} p_i^{k_i}$.

Démonstration. Soit n un entier positif. Nous prouvons l'existence d'une décomposition en facteurs premiers par récurrence. Le nombre $n = 1$ est le produit d'une famille finie de nombres premiers : la famille vide [2].

Supposons que tout entier strictement inférieur à un certain entier $n > 1$ est produit de nombres premiers. Deux possibilités apparaissent pour n : il est premier ou non. Si n est premier, et donc produit d'un unique entier premier, à savoir lui-même, le résultat est vrai. Si n n'est pas premier, il se décompose sous la forme kl avec k et l strictement inférieurs à n. Dans ce cas, l'hypothèse de récurrence implique que les entiers k et l peuvent s'écrire comme produits de nombres premiers. Leur produit aussi, ce qui fournit une décomposition de n en produit de nombres premiers. Par application du principe de récurrence, tous les entiers naturels peuvent s'écrire comme produit de nombres premiers.

2. Voir 1.316.

Nous prouvons maintenant l'unicité. Prenons deux produits de nombres premiers qui sont égaux. Prenons n'importe quel nombre premier p du premier produit. Il divise le premier produit, et, de là, aussi le second. Par le lemme d'Euclide 3.13, p doit alors diviser au moins un facteur dans le second produit. Mais les facteurs sont tous des nombres premiers eux-mêmes, donc p doit être égal à un des facteurs du second produit. Nous pouvons donc simplifier par p les deux produits. En continuant de cette manière, nous voyons que les facteurs premiers des deux produits coïncident précisément. □

LEMooDTQQooYoJABt

Lemme 3.16 ([1])**.**
Nous notons $\mathcal{P}$ l'ensemble des nombres premiers dans $\mathbb{N}$. Soient des suites finies $(a_p)_{p\in\mathcal{P}}$ et $(b_p)_{p\in\mathcal{P}}$. Nous posons

$$a = \prod_{p\in\mathcal{P}} p^{a_p} \quad et \quad b = \prod_{p\in\mathcal{P}} p^{b_p}. \tag{3.29}$$

Alors $a \mid b$ si et seulement si $a_p \leqslant b_p$ pour tout p.

Démonstration. Dire que $a \mid b$ signifie qu'il existe $s \in \mathbb{N}$ tel que $as = b$; le théorème 3.15 nous permet de décomposer s en $s = \prod_{p\in\mathcal{P}} p^{s_p}$. Puisque le produit dans $\mathbb{N}$ est commutatif et associatif,

$$b = as = \prod_{p\in\mathcal{P}} p^{s_p+a_p}. \tag{3.30}$$

Par unicité de la décomposition de b (toujours le théorème 3.15), nous en déduisons que $b_p = s_p + a_p \geqslant a_p$.

Dans l'autre sens, l'hypothèse $a_p \leqslant b_p$ implique l'existence de $s_p \geqslant 0$ tels que $b_p = a_p + s_p$. En posant $s = \prod_{p\in\mathcal{P}} p^{s_p}$, nous avons

$$as = \prod_{p\in\mathcal{P}} p^{s_p+a_p} = \prod_{p\in\mathcal{P}} p^{b_p} = b. \tag{3.31}$$

Donc $a \mid b$. □

LEMooGLZHooUcRNgu

Lemme 3.17.
Soient un nombre premier q ainsi que $a \in \mathbb{Z}$. Soit un entier $n \geqslant 1$. Le nombre q divise a si et seulement si il divise a^n.

Démonstration. Nous numérotons les nombres premiers p_i pour que p_1 soit q. La décomposition en nombre premiers du théorème 3.15 nous dit que

$$a = q^{a_1} \prod_{i\neq 1} p_i^{a_i} \tag{3.32}$$

et

$$a^n = q^{na_1} \prod_{i\neq 1} p_i^{na_i} \tag{3.33}$$

Nous avons équivalence entre les énoncés suivants :

— q divise a
— $a_1 \neq 0$
— $na_1 \neq 0$ (parce que $n \neq 0$)
— q divise a^n.

□

CORooWBSQooQOEmaC

Corolaire 3.18.
Si $n \in \mathbb{N}$ n'est pas une puissance d'un nombre premier, alors il existe $a, b \in \mathbb{N}$ tels que $\gcd(a, b) = 1$ *et $n = ab$.*

Démonstration. Si n n'est pas une puissance d'un nombre premier, alors le théorème 3.15 de décomposition en nombres premiers dit que

$$n = \prod_{i=1}^{r} p_i^{k_i} \tag{3.34}$$

avec au moins $r = 2$, et les k_i non nuls. En prenant $a = p_1^{k_1}$ et $b = \prod_{i=2}^{r} p_i^{k_i}$ nous avons bien $n = ab$. Montrons que a et b n'ont pas de diviseurs communs. Soit, par l'absurde, q un diviseur premier commun à a et b.

Vu que q divise $p_1^{k_1}$, le nombre q divise p_1 par le lemme 3.17. Étant donné que q divise $\prod_{i=2}^{r} p_i^{k_i}$, il divise au moins un des facteurs (lemme d'Euclide 3.13). Disons que q divise $p_i^{k_i}$. Dans ce cas q divise p_i.

Donc q divise p_i et p_1 et donc $q = 1$ parce que p_i et p_1 sont des nombres premiers distincts. □

LEMooBJVJooFyuFeN

Lemme 3.19 ([1])**.**
Dans $\mathbb{N}$*, le pgcd*[3] *et le ppcm sont uniques.*

Démonstration. Supposons que δ_1 et δ_2 soient des pgcd de la partie S. Puisque $\delta_1 \mid S$, nous avons $\delta_1 \mid \delta_2$ parce que δ_2 est un pgcd. Le même raisonnement, inversant δ_1 et δ_2 montre que $\delta_2 \mid \delta_1$. Si (a_p) sont les éléments de la décomposition de δ_1 et (b_p) ceux de δ_2, alors le lemme 3.16 nous indique que $a_p \leqslant b_p$ et $b_p \leqslant a_p$, ce qui implique que $a_p = b_p$.

La démonstration pour le ppcm s'effectue selon le même principe. □

LEMooJIGRooARiIPC

Lemme 3.20.
Soit une partie S *de* $\mathbb{N}$*.*

(1) Le pgcd de S *est le plus grand élément de* $\mathbb{N}$ *divisant tous les éléments de* S*.*

(2) Le ppcm de S *est le plus petit élément de* $\mathbb{N}$ *que tous les éléments de* S *divisent.*

LEMooEVIZooPAkQZW

Lemme 3.21 ([1])**.**
Soient $a, b \in \mathbb{Z}$ *et* $k \in \mathbb{N}$ *tels que* $ab = q^k$ *et* $\gcd(a, b) = 1$*. Alors il existe* $\alpha, \beta \in \mathbb{Z}$ *tels que* $a = \alpha^k$ *et* $b = \beta^k$*.*

Démonstration. Nous décomposons a, b et q en facteurs premiers suivant le théorème 3.15 :

SUBEQooBJEQooDiWbYg

$$a = \prod_i p_i^{a_i} \tag{3.35a}$$

$$b = \prod_i p_i^{b_i} \tag{3.35b}$$

$$q = \prod_i p_i^{q_i}. \tag{3.35c}$$

D'un part, en utilisant la commutativité et l'associativité du produit,

$$ab = \prod_i p_i^{a_i + b_i}. \tag{3.36}$$

D'autre part, puisque $ab = q^k$, nous avons

$$ab = \left(\prod_i p_i^{q_i}\right)^k = \prod_i p_i^{kq_i}. \tag{3.37}$$

En vertu de l'unicité de la décomposition en facteurs premiers, pour chaque i nous avons

$$a_i + b_i = kq_i. \tag{3.38}$$

3. Le pgcd et ppcm sont définis dans 1.186.

Comme a et b sont premiers entre eux, si $a_i \neq 0$ alors $b_i = 0$ et inversement. Prenons un i tel que $a_i \neq 0$. Alors $b_i = 0$ et nous avons $a_i = kq_i$. Idem pour les b_i.

Donc tous les a_i et les b_i qui sont non nuls sont des multiples de k. Nous posons $a_i = ks_i$ et nous reportons dans (3.35a) :

$$a = \prod_i p_i^{ks_i} = (\prod_i p_i^{s_i})^k, \tag{3.39}$$

de telle sorte que a soit une puissance k^{e}. La même chose tient pour b. □

PROPooNQBOooHWqTvs

Proposition 3.22.
Soient $a, b \in \mathbb{Z}\backslash\{0\}$ décomposés en $a = \prod_{p\in\mathcal{P}} p^{a_p}$ et $b = \prod_{p\in\mathcal{P}} p^{b_p}$. En posant

$$m_p = \min\{a_p, b_p\} \tag{3.40a}$$
$$M_p = \max\{a_p, b_p\}, \tag{3.40b}$$

nous avons

$$\text{pgcd}(a, b) = \prod_{p\in\mathcal{P}} p^{m_p} \tag{3.41a}$$
$$\text{ppcm}(a, b) = \prod_{p\in\mathcal{P}} p^{M_p}. \tag{3.41b}$$

Pour rappel, la définition de pgcd et ppcm sont dans 1.186.

Démonstration. Nous commençons par le pgcd. Nous notons $\delta = \prod_{p\in\mathcal{P}} p^{m_p}$ et nous prouvons que δ est un pgcd de $\{a, b\}$. Il y a deux propriétés à vérifier.

(1) **δ divise a et b** Puisque $m_p = \min\{a_p, b_p\}$, nous avons $m_p \leqslant a_p$ et $m_p \leqslant b_p$. Le lemme 3.16 nous dit alors que $\delta \mid a$ et $\delta \mid b$.

(2) **Si s divise a et b** De même, si $s \mid a$ et $s \mid b$, nous avons $s_p \leqslant a_p$ et $s_p \leqslant b_p$, ce qui montre que $s_p \leqslant m_p$ et donc que $s \mid \delta$.

Pour le ppcm, nous posons $\mu = \prod_{p\in\mathcal{P}} p^{M_p}$, et nous prouvons que μ est un ppcm de $\{a, b\}$.

(1) **a et b divisent μ** Pour tout p, nous avons $M_p \geqslant a_p$. Le lemme 3.16 implique que $a \mid \mu$. Idem pour b, donc tous les éléments de $\{a, b\}$ divisent μ.

(2) **Si a et b divisent r** Supposons que a et b divisent un certain nombre r. Alors $a_p \leqslant r_p$ et $b_p \leqslant r_p$. Donc $r_p \geqslant \max\{a_p, b_p\} = M_p$. Nous en déduisons que $\mu \mid r$.

Puisque les pgcd et ppcm sont uniques (lemme 3.19), nous avons prouvé que δ et μ sont les nombres recherchés. □

CORooQIMHooUzLUJY

Corolaire 3.23.
Soit un nombre premier p. Un élément $m \in \mathbb{Z}^*$ vérifie $\text{pgcd}(m, p^n) \neq 1$ si et seulement si $m = qp$ pour un certain q.

Démonstration. Nous considérons les décompositions en facteurs premiers $m = \prod_{s\in\mathcal{P}} s^{a_s}$ et $p^n = \prod_{s\in\mathcal{P}} s^{b_s}$. Par la partie unicité du théorème 3.15, nous savons que $b_s = 0$ pour $s \neq p$ et $b_p = n$.

Nous prenons ensuite l'expression du pgcd donné par la proposition 3.22 :

$$\text{pgcd}(m, p^n) = \prod_{s\in\mathcal{P}} s^{\min\{a_s, b_s\}}. \tag{3.42}$$

Le minimum est toujours zéro lorsque $s \neq p$, donc $\text{pgcd}(m, p^n) = p^{\min\{a_s, n\}}$.

Nous avons donc $\text{pgcd}(m, p^n) \neq 1$ si et seulement si $a_p \neq 0$. Mais $a_p \neq 0$ si et seulement si m est un multiple de p. □

LemheKdsa

Lemme 3.24.
Un entier $n \geqslant 1$ se décompose de façon unique en produit de la forme $n = qm^2$ où q est un entier sans facteurs carrés et m, un entier.

Démonstration. Pour $n = 1$, c'est évident. Nous supposons $n \geqslant 2$.

En ce qui concerne l'existence, nous décomposons n en facteurs premiers [4] et nous séparons les puissances paires des puissances impaires :

$$n = \prod_{i=1}^{r} p_i^{2\alpha_i} \prod_{j=1}^{s} q_j^{2\beta_j+1} \tag{3.43a}$$

$$= \underbrace{\left(\prod_{i=1}^{r} p_i^{2\alpha_i} \prod_{j=1}^{s} q^{2\beta_j} \right)}_{m^2} \underbrace{\prod_{j=1}^{s} q_j}_{q}. \tag{3.43b}$$

Nous passons à l'unicité. Supposons que $n = q_1 m_1^2 = q_2 m_2^2$ avec q_1 et q_2 sans facteurs carrés (dans leur décomposition en facteurs premiers). Soit $d = \text{pgcd}(m_1, m_2)$ et k_1, k_2 définis par $m_1 = dk_1$, $m_2 = dk_2$. Par construction, $\text{pgcd}(k_1, k_2) = 1$. Étant donné que

$$n = q_1 d^2 k_1^2 = q_2 d^2 k_2^2, \tag{3.44}$$

EqWPQtto

nous avons $q_1 k_1^2 = q_2 k_2^2$ et donc k_1^2 divise $q_2 k_2^2$. Mais k_1 et k_2 n'ont pas de facteurs premiers en commun, donc k_1^2 divise q_2, ce qui n'est possible que si $k_1 = 1$ (parce que k_1^2 n'a que des facteurs premiers alors que q_2 n'en a pas). Dans ce cas, $d = m_1$ et m_1 divise m_2. Si $m_2 = lm_1$ alors l'équation (3.44) se réduit à $n = q_1 m_1^2 = q_2 l^2 m_1^2$ et donc

$$q_1 = q_2 l^2, \tag{3.45}$$

ce qui signifie $l = 1$ et donc $m_1 = m_2$. □

Les nombres premiers ne sont pas si rares que ça dans $\mathbb{N}$. Nous allons voir dans 15.120 que la somme des inverses des nombres premiers diverge. Pour comparaison, la somme des inverses des carrés converge par la proposition 10.136. Il y a donc, dans un certains sens, plus de nombres premiers que de carrés ; dans un autre sens, il y en a autant : une infinité dénombrable.

3.2.4 Ordre d'un élément dans un groupe fini

Voir plus d'informations dans la partie 5.2 sur les groupes monogènes.

LEMooRFKQooWTdYcr

Lemme 3.25 ([1]).
Si g est un élément d'ordre s, et si $g^r = e$, alors $s \mid r$.

THOooSUWKooICbzqM

Théorème 3.26 (Théorème de Cauchy[98, 99]).
Soit G, un groupe cyclique d'ordre n. Si d divise n, alors G possède un unique sous-groupe d'ordre d.

En particulier, G possède des éléments de tous les ordres divisant n.

Démonstration. En plusieurs parties.

(1) **Existence** L'hypothèse est que G est un groupe cyclique. Donc nous pouvons considérer une générateur g de G. Nous considérons $H_d = \text{gr}(g^{n/d})$, et nous prouvons que H_d est un sous-groupe d'ordre d.

D'abord nous avons $(g^{n/d})^d = g^n = e$. Donc $|H_d| \leqslant d$. Pour prouver que $|H_d| \geqslant d$, nous supposons par l'absurde qu'il existe $k < d$ tel que $(g^{n/d})^k = e$. Nous aurions alors $g^{nk/d} = e$ avec $nk/d < n$, ce qui contredirait que g est générateur d'un groupe d'ordre n.

(2) **Unicité** Soit H, un sous-groupe de G d'ordre d. Nous allons prouver que $H = \text{gr}(g^{n/d})$. Vu que H et H_d ont le même cardinal, il suffira de montrer que $H \subset H_d$ pour avoir $H = H_d$.

4. Théorème 3.15.

Vu que $\{e\}$ est l'unique sous-groupe d'ordre 1, nous supposons que $d > 1$. Soit $h \neq e$ dans H. Vu que g est générateur de G, il existe $k \in \mathbb{N}^*$ tel que $g^k = h$. Vu que H est d'ordre d, nous avons $h^d = e$ (corolaire 2.14). Donc

$$g^{kd} = e. \tag{3.46}$$

Le lemme 3.25 dit alors que $n \mid kd$. Soit $t \in \mathbb{N}^*$ tel que $nt = kd$; nous pouvons écrire $k = tn/d$ et récrire $g^k = h$ avec cette valeur de k :

$$h = g^{tn/d} = (g^{n/d})^t \in \text{gr}(g^{n/d}). \tag{3.47}$$

Vu que h a été pris arbitrairement dans H, nous en déduisons que $H \subset \text{gr}(g^{n/d})$.

□

Le lemme suivant indique que sous hypothèse de commutativité, l'ordre d'un élément est une notion multiplicative.

LemyETtdy

Lemme 3.27 ([100]).
Soit G un groupe et $a, b \in G$ tels que $ab = ba$ d'ordres respectivement r et s, deux nombres premiers entre eux. Alors l'élément ab est d'ordre rs.

Démonstration. Étant donné que $(ab)^{rs} = a^{rs}b^{rs} = 1$, l'ordre de ab divise rs. Et comme r et s sont premiers entre eux, l'ordre de ab s'écrit sous la forme r_1s_1 avec $r_1 \mid r$ et $s_1 \mid s$. Nous avons

$$a^{r_1s_1}b^{r_1s_1} = (ab)^{r_1s_1} = 1, \tag{3.48}$$

que nous élevons à la puissance r_2 où r_2 est défini en posant $r = r_1r_2$:

$$a^{rs_1}b^{rs_1} = 1. \tag{3.49}$$

Et comme $a^{rs_1} = 1$, nous concluons que $b^{rs_1} = 1$. Donc $s \mid rs_1$. Par le lemme de Gauss 3.12, nous avons en fait $s \mid s_1$. Puisqu'on a aussi $s_1 \mid s$, nous avons $s = s_1$.

Le même type d'argument donne $r = r_1$, et finalement l'ordre de ab est $r_1s_1 = rs$. □

LemSkIOOG

Lemme 3.28 ([59]).
Un sous-groupe d'indice 2 est un sous-groupe normal.

Démonstration. Si H est un tel sous-groupe d'un groupe G, alors G possède exactement deux classes à gauche par rapport à H (théorème de Lagrange 2.13) et se partitionne donc en deux parties : $G = H \cup xH$ avec $x \notin H$. De même pour les classes à droite : $G = H \cup Hx$. Puisque la classe à droite Hx n'est pas H, on a $xH = Hx$, et H est normal dans G par la proposition 1.265. □

PropubeiGX

Lemme 3.29 ([101]).
Soit H, un sous-groupe normal d'indice m de G. Alors pour tout $a \in G$ nous avons $a^m \in H$.

Démonstration. Par définition de l'indice, le groupe G/H est d'ordre m. Donc si $[a] \in G/H$, nous avons $[a]^m = [e]$, ce qui signifie $[a^m] = [e]$, ou encore $a^m \in H$. □

PROPooVWVIooQzuAlA

Proposition 3.30 ([101]).
Soit un groupe fini G et H, un sous-groupe normal d'ordre n et d'indice m avec m et n premiers entre eux. Alors H est l'unique sous-groupe de G à être d'ordre n.

Démonstration. Soit H' un sous-groupe d'ordre n. Si $h \in H'$ alors $h^n = 1$ et $h^m \in H$ par le lemme 3.29. Étant donné que m et n sont premiers entre eux, par le théorème de Bézout 1.236, il existe $a, b \in \mathbb{Z}$ tels que

$$am + bn = 1. \tag{3.50}$$

Et donc, $h = h^1 = (h^m)^a(h^n)^b$. En tenant compte du fait que $h^n = 1$ et $h^m \in H$, nous avons $h \in H$. Ce que nous venons de prouver est que $H' \subset H$ et donc que $H = H'$ parce que $|H'| = |H| = |G|/m$. □

3.31.
Notons que cette proposition ne dit pas qu'il existe un sous-groupe d'ordre n et d'indice m. Il dit juste que si il y en a un et si il est normal, alors il n'y en a pas d'autre.

LemqAUBYn

Lemme 3.32.
L'ensemble des ordres d'un groupe commutatif est stable par PPCM[5].

Autrement dit, si $x \in G$ est d'ordre r et si $y \in G$ est d'ordre s, alors il existe un élément d'ordre $\operatorname{ppcm}(r,s)$.

Démonstration. Soit $m = \operatorname{ppcm}(r,s)$. Afin d'écrire m sous une forme pratique, nous considérons les décompositions en facteurs premiers de r et s :

$$r = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{3.51a}$$

$$s = \prod_{i=1}^{k} p_i^{\beta_i} \tag{3.51b}$$

où $\{p_i\}_{i=1,\dots,k}$ est l'ensemble des nombres premiers arrivant dans les décompositions de r et de s. Si nous posons

$$r' = \prod_{\substack{i=1\\ \alpha_i > \beta_i}}^{k} p_i^{\alpha_i} \tag{3.52a}$$

$$s' = \prod_{\substack{i=1\\ \alpha_i \leqslant \beta_i}}^{k} p_i^{\beta_i}, \tag{3.52b}$$

alors $\operatorname{ppcm}(r,s) = r's'$ et r' et s' sont premiers entre eux. L'élément $x^{r/r'}$ est d'ordre r' et l'élément $y^{s/s'}$ est d'ordre s'. Maintenant nous utilisons le fait que G soit commutatif et le lemme 3.27 pour conclure que l'ordre de $x^{r/r'}y^{s/s'}$ est $r's' = m$. □

3.2.5 Écriture des fractions

THOooWYQVooRBaAAM

Théorème 3.33 ([102]).
Tout élément de $\mathbb{Q}^+$ s'écrit de façon unique comme quotient de deux entiers premiers entre eux.

Démonstration. En deux parties[6]

(1) **Unicité** Supposons avoir $\frac{a}{b} = \frac{c}{d}$ avec $\operatorname{pgcd}(a,b) = \operatorname{pgcd}(c,d) = 1$. Nous avons

$$ad = bc \tag{3.53}$$

donc

(1a) a divise bc mais est premier avec b donc a divise c par le lemme de Gauss 3.12.

(1b) c divise ad mais est premier avec d donc c divise a par le lemme de Gauss 3.12.

En conclusion a divise c et c divise a, ergo[7] $a = c$. L'égalité $b = d$ est alors immédiate.

(2) **Existence** Soit le quotient $\frac{a}{b}$. Nous avons

$$\frac{a}{b} = \frac{a/\operatorname{pgcd}(a,b)}{b/\operatorname{pgcd}(a,b)}, \tag{3.54}$$

qui est encore un quotient d'entiers parce que $\operatorname{pgcd}(a,b)$ divise aussi bien a que b. Il faut montrer que les nombres $a/\operatorname{pgcd}(a,b)$ et $b/\operatorname{pgcd}(a,b)$ sont premiers entre eux. Pour cela

5. Définition 1.186.
6. Définitions des pgcd et ppcm en 1.186.
7. Lemme 3.14.

nous supposons que k est un diviseur commun. En particulier, les nombres $a/k\operatorname{pgcd}(a,b)$ et $b/k\operatorname{pgcd}(a,b)$ sont des entiers, ce qui fait que $k\operatorname{pgcd}(a,b)$ est un diviseur commun de a et b. Étant donné que $\operatorname{pgcd}(a,b)$ est le plus grand tel diviseur, nous devons avoir $k\operatorname{pgcd}(a,b)=\operatorname{pgcd}(a,b)$ c'est-à-dire que $k=1$. Donc les nombres $a/\operatorname{pgcd}(a,b)$ et $b/\operatorname{pgcd}(a,b)$ sont premiers entre eux.

□

PROPooWBFJooWisSZX

Proposition 3.34 ([1])**.**
Soient $l,l,d_1,d_2\in\mathbb{N}$ tels que

$$\frac{k}{d_1}=\frac{l}{d_2} \tag{3.55}$$

avec $\operatorname{pgcd}(k,d_1)=1$.
Alors

(1) $d_1\leqslant d_2$

(2) $d_1=\frac{d_2}{\operatorname{pgcd}(l,d_2)}$.

PROPooRZDDooLJabov

Proposition 3.35.
Les entiers p et q sont premiers entre eux[8] *si et seulement si p^2 et q^2 sont premiers entre eux.*

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que p^2 et q^2 ne sont pas premiers entre eux. Donc il existe δ divisant p^2 et q^2. Si δ' est un facteur premier de δ, alors δ' divise δ et donc aussi p^2 et q^2. Le lemme 3.17 implique que δ divise p et q. Donc p et q ne sont pas premiers entre eux.

(2) $\Leftarrow$ Si p^2 et q^2 sont premiers entre eux, par le théorème de Bézout 1.236 il existe $a,b\in\mathbb{Z}$ tels que

$$ap^2+bq^2=1 \tag{3.56}$$

Dans ce cas, $(ap)p+(bq)q=1$, ce qui montre (par encore Bézout, mais l'autre sens) que p et q sont premiers entre eux.

□

Une des conséquences de ces résultats sera le fait que $\sqrt{n}$ est irrationnelle dès que n n'est pas un carré parfait, théorème 3.36.

Nous avons déjà vu dans la proposition 1.409 que $\sqrt{2}$ était irrationnel. En fait le théorème suivant va nous montrer que le nombre $\sqrt{n}$ est soit entier, soit irrationnel.

THOooYXJIooWcbnbm

Théorème 3.36.
Soit $n\in\mathbb{N}$. Le nombre $\sqrt{n}$ est rationnel si et seulement si n est un carré parfait.

Démonstration. Supposons que $\sqrt{n}$ soit rationnel. Le théorème 3.33 nous donne $p,q\in\mathbb{N}$ premiers entre eux tels que $\sqrt{n}=p/q$. La proposition 3.35 nous enseigne de plus que p^2 et q^2 sont premiers entre eux. Nous avons

$$p^2=nq^2. \tag{3.57}$$

Le nombre q est alors un diviseur commun de q^2 et de p. Donc $q=1$ et $n=p^2$ est un carré parfait. □

3.2.6 Équation diophantienne linéaire à deux inconnues

subsecZVKNooXNjPSf

Soient a, b et c des entiers naturels donnés. Nous considérons l'équation

$$ax+by=c \tag{3.58}$$

EqTOVSooJhxIIq

à résoudre[103] pour $(x,y)\in\mathbb{N}^2$.

8. Premiers entre eux, définition 1.259.

Si a ou b est nul, c'est facile ; nous supposons donc que a et b sont tous deux non nuls. Nous commençons par simplifier l'équation en cherchant les diviseurs communs. Soit $d = \text{pgcd}(a, b)$ et notons $a = da'$, $b = db'$. Nous avons déjà l'équation

$$d(a'x + b'y) = c, \tag{3.59}$$

et donc si c n'est pas un multiple de d, il n'y a pas de solution [9]. Si par contre c est un multiple de d alors nous écrivons $c = c'd$ et l'équation devient

$$a'x + b'y = c' \tag{3.60}$$

C'est maintenant que nous utilisons le théorème de Bézout 1.236 : puisque a' et b' sont premiers entre eux, nous avons la relation $a'u + b'v = 1$ pour certains [10] nombres entiers u et v. Nous récrivons notre équation sous la forme $a'x + b'y = c'(a'u + b'v)$ et rassemblons les termes :

$$a'(x - c'u) = b'(c'v - y). \tag{3.61}$$

C'est-à-dire que si (x, y) est une solution, alors a' divise $b'(c'v - y)$. Mais comme a' et b' sont premiers entre eux, le nombre a' doit forcément [11] diviser $c'v - y$. Disons $c'v - y = ka'$. Alors $a'(x - c'u) = b'ka'$ et donc

$$x = b'k + c'u. \tag{3.62}$$

Nous trouvons alors une expression pour y en injectant ce résultat dans $a'x+b'y = c'$ et en utilisant le théorème de Bézout : $a'c'u = (1 - b'v)c'$. Au final nous avons prouvé que toutes les solutions sont de la forme [EqYCQVooZzHuRq]

$$\begin{cases} x = b'k + c'u & \text{(3.63a)} \\ y = vc' - a'k & \text{(3.63b)} \end{cases}$$

avec $k \in \mathbb{Z}$. Si nous ne voulons réellement que les solutions dans $\mathbb{N}$ et non dans $\mathbb{Z}$, il faut seulement un peu restreindre les valeurs de k. Il en reste un nombre fini parce que l'équation pour x borne k vers le bas tandis que celle pour y borne k vers le haut.

Par ailleurs, il est très vite vérifié que tous les couples (x, y) de la forme (3.63) sont solutions.

Exemple 3.37.
Résoudre l'équation $2x + 6y = 52$.

Nous pouvons factoriser 2 dans le membre de gauche et simplifier alors toute l'équation par 2 :

$$x + 3y = 26. \tag{3.64}$$

Nous cherchons une relation de Bézout pour $u+3v = 1$. Ce n'est heureusement pas très compliqué : $u = -5$, $v = 2$. Nous pouvons alors écrire

$$x + 3y = 26 \times (-5 + 3 \times 2), \tag{3.65}$$

et donc $x + 5 \times 26 = -3(y - 26 \times 2)$, et en posant $k = y - 26 \times 2$ nous avons

$$x = -3k - 130. \tag{3.66}$$

En injectant nous trouvons l'équation $3y - 3k - 130 = 26$ et donc

$$y = 52 + k. \tag{3.67}$$

△

9. Exemple : $8x + 2y = 9$. Le membre de gauche est certainement un nombre pair et il n'y a donc pas de solution.
10. Nous avons décrit un algorithme pour les trouver en démontrant l'équation 3.20.
11. C'est Gauss 3.12 qui le dit, et vous savez que lorsque Gauss dit, c'est *forcément* vrai.

3.2.7 Quotients

Nous écrivons $a = b \mod p$ essentiellement si il existe $k \in \mathbb{Z}$ tel que $b + kp = a$. Plus généralement nous notons $[a]_p = \{a + kp | k \in \mathbb{Z}\}$ et l'écriture « $a = n \mod p$ » peut tout autant signifier $a = [b]_p$ que $a \in [b]_p$. La différence entre les deux est subtile mais peut de temps en temps valoir son pesant d'or.

Proposition 3.38.
Soit $n \in \mathbb{N}$. Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est monogène. Si $n \neq 0$, il est cyclique d'ordre n.

Démonstration. Nous considérons la surjection canonique $\mu: \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Si $a \in \mathbb{Z}$, alors $\mu(a) = a\mu(1)$. Par conséquent $\mathbb{Z}/n\mathbb{Z} = \text{gr}\big(\mu(1)\big)$ parce que tout groupe contenant $\mu(1)$ contient tous les multiples de $\mu(1)$, et par conséquent contient $\mu(\mathbb{Z}) = \mathbb{Z}/n\mathbb{Z}$.

Soit $x \in \mathbb{Z}/n\mathbb{Z}$ et considérons x_0, le plus petit naturel représentant x. Nous notons $x = [x_0]$. Le théorème de la division euclidienne 1.222 donne l'existence de q et r avec $0 \leqslant r < n$ et $q \geqslant 0$ tels que

$$x_0 = nq + r. \tag{3.68}$$

Nous avons $[x_0] = [r] = \mu(r)$ parce que $x_0 - r$ est un multiple de n. Nous avons donc $[x_0] \in \mu(\mathbb{N}_{n-1})$. Par conséquent

$$\mathbb{Z}/n\mathbb{Z} = \mu(\mathbb{Z}) = \mu(\mathbb{N}_{n-1}). \tag{3.69}$$

La restriction $\mu: \mathbb{N}_{n-1} \to \mathbb{Z}/n\mathbb{Z}$ est donc surjective. Montrons qu'elle est également injective. Si $\mu(x_0) = \mu(x_1)$, alors $x_1 = x_0 + kn$. Si nous supposons que $x_1 > x_0$, alors $k > 0$ et si $x_0 \in \mathbb{N}_{n-1}$, alors $x_1 > n - 1$.

L'ordre de $\mathbb{Z}/n\mathbb{Z}$ est donc le même que le cardinal de $\mathbb{N}_{n-1}$, c'est-à-dire n. Le groupe $\mathbb{Z}/n\mathbb{Z}$ est donc fini, d'ordre n et monogène parce que $\mathbb{Z}/n\mathbb{Z} = \text{gr}(\mu(1))$. Il est donc cyclique. □

Lemme 3.39 ([104]).
Soit $q \in \mathbb{N}$ avec $q \geqslant 2$. Soient $n, d \in \mathbb{N}$ tels que $q^d - 1 \mid q^n - 1$. Alors $d \mid n$.

Démonstration. Par le théorème de division euclidienne 1.222, il existe $a, b \in \mathbb{Z}$ tels que $n = ad + b$ avec $0 \leqslant b < d$. En remarquant que $q^d \in [1]_{q^d-1}$ nous avons

$$q^n = (q^d)^a q^b \in [1]_{q^d-1} q^b = [q^b]_{q^d-1}. \tag{3.70}$$

Pour cela nous avons utilisé d'abord le fait que si $a \in [z]_k$, alors $a^n \in [z^n]_k$ et ensuite le fait que $[1]_k x = [x]_k$. D'autre part l'hypothèse $q^d - 1 \mid q^n - 1$ implique

$$q^n \in [1]_{q^d-1}. \tag{3.71}$$

Par conséquent le nombre q^n est à la fois dans $[q^b]_{q^d-1}$ et dans $[1]_{q^d-1}$. Cela implique que

$$[1]_{q^d-1} = [q^b]_{q^d-1}, \tag{3.72}$$

ou encore que $q^b \in [1]_{q^d-1}$ ou encore que $q^d - 1 \mid q^b - 1$.

Étant donné que $b < d$ et que $q \geqslant 2$, nous avons que $q^b - 1 < q^d - 1$; donc pour que $q^d - 1$ divise $q^b - 1$, il faut que $q^b - 1$ soit zéro, c'est-à-dire $b = 0$.

Mais dire $b = 0$ revient à dire que $d \mid n$, ce qu'il fallait démontrer. □

3.3 Binôme de Newton et morphisme de Frobenius

DEFooHXNFooOBKgqD

Définition 3.40 (Coefficient binomial).
*Soient des entiers n et k avec $k \leqslant n$. Nous définissons **coefficients binomiaux** par*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{3.73}$$

EQooJJYGooTymSAt

Lorsque α est un multiindice, la factorielle $\alpha!$ et les coefficients binomiaux sont définis en 3.47.

LEMooHWMNooIINsxu

Lemme 3.41.
Un formule pour les coefficients binomiaux :

$$\binom{n}{k}+\binom{n}{k-1}=\binom{n+1}{k} \tag{3.74}$$

Démonstration. Simple mise au même dénominateur :

$$\frac{n!}{k!(n-k)!}+\frac{n!}{(k-1)!(n-k+1)!}=\frac{n!(n-k+1)+n!k}{k!(n-k+1)!}=\frac{n!(n+1)}{k!(n+1-k)!}. \tag{3.75}$$

□

LEMooUTDTooXAmvdF

Lemme 3.42 ([1])**.**
Nous considérons ces deux ensembles :

$$A=\{I=(i_1,\dots,i_k)\text{ tel que }1\leqslant i_1<\dots<i_k\leqslant n\} \tag{3.76a}$$
$$C=\{\text{parties de cardinal } k \text{ dans } \{1,\dots,n\}\}. \tag{3.76b}$$

Il existe une bijection entre A et C.

Nous avons de plus[12]

EQooSPVNooHoPuOe

$$\operatorname{Card}(A)=\operatorname{Card}(C)=\binom{n}{k} \tag{3.77}$$

PropBinomFExOiL

Proposition 3.43 (Formule binomiale[105])**.**
Pour tout x, $y\in\mathbb{R}$ et $n\in\mathbb{N}$, nous avons

EqNewtonB

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k \tag{3.78}$$

où les coefficients binomiaux sont donnés dans le lemme 3.42.

Démonstration. La preuve se fait par récurrence. La vérification pour $n=0$ et $n=1$ se fait aisément pour peu que l'on se rappelle que $x^0=1$ et que $0!=1$, ce qui donne entre autres $\binom{0}{0}=1$.

EqBinTrav

Supposons que la formule (3.78) soit vraie pour $n\geqslant 1$, et prouvons la pour $n+1$. Nous avons

SUBEQooBEUCooHSRtug

$$(x+y)^{n+1}=(x+y)\cdot\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k \tag{3.79a}$$

SUBEQooRURMooGVIbbL

$$=\sum_{k=0}^{n}\binom{n}{k}x^{n-k+1}y^k+\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^{k+1} \tag{3.79b}$$
$$=x^{n+1}+\sum_{k=1}^{n}\binom{n}{k}x^{n-k+1}y^k+\sum_{k=0}^{n-1}\binom{n}{k}x^{n-k}y^{k+1}+y^{n+1}. \tag{3.79c}$$

Notez le passage de (3.79a) à (3.79b) qui demande la commutativité de la multiplication. La formule binomiale ne se généralise pas de façon directe à un anneau non commutatif.

La seconde grande somme peut être transformée en posant $i=k+1$:

$$\sum_{k=0}^{n-1}\binom{n}{k}x^{n-k}y^{k+1}=\sum_{i=1}^{n}\binom{n}{i-1}x^{n-(i-1)}y^{i-1+1}, \tag{3.80}$$

dans lequel nous pouvons immédiatement renommer i par k. En remplaçant dans la dernière expression de (3.79), nous trouvons

$$(x+y)^{n+1}=x^{n+1}+y^{n+1}+\sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]x^{n-k+1}y^k. \tag{3.81}$$

La thèse découle maintenant de la formule du lemme 3.41. □

12. Pour le coefficient binomial, définition 3.40.

LEMooLPCXooYIzJsD

Lemme 3.44 ([1]).
Si $n \geqslant k$ nous avons

$$\frac{n!}{k!(n-k)!} \leqslant \frac{n^{k-1}}{k!}. \tag{3.82}$$

Démonstration. Nous décomposons le produit définissant $n!$ en les facteurs entre 1 et $n-k$ et ceux entre $n-k+1$ et n :

$$n! = (n-k)! \prod_{i=n-k+1}^{n} i \leqslant n^{k-1}(n-k)!. \tag{3.83}$$

Donc

$$\frac{n!}{k!(n-k)!} \leqslant \frac{n^{k-1}}{k!}. \tag{3.84}$$

□

Tant que nous sommes à démontrer des égalités, en voici une.

LEMooLPOQooICJYdV

Lemme 3.45 ([106]).
Pour $a, b \in \mathbb{R}$ et $n \in \mathbb{N}$ nous avons

$$a^n + (-1)^{n-1}b^n = (a+b)\sum_{k=0}^{n-1}(-1)^k a^{n-1-k}b^k. \tag{3.85}$$

Démonstration. C'est un simple calcul :

$$(a+b)\sum_{k=0}^{n-1}(-1)^k a^{n-1-k}b^k = \sum_{k=0}^{n-1}(-1)^k a^{n-k}b^k + \sum_{k=0}^{n-1}(-1)^k a^{n-k-1}b^{k+1} \tag{3.86a}$$

$$= a^n + \sum_{k=1}^{n-1}(-1)^k a^{n-k}b^k + \sum_{k=0}^{n-2}(-1)^k a^{n-k-1}b^{k+1} + (-1)^{n-1}b^n \tag{3.86b}$$

SUBEQooUCIBooKauEbb

$$= a^n + (-1)^{n-1}b^n \tag{3.86c}$$

SUBEQooLTIHooZPMwVF

Justifications.

— Pour (3.86b). Dans la première somme, nous avons séparé le terme $k = 0$ et dans la seconde nous avons séparé le terme $k = n-1$

— Pour (3.86c). Dans la seconde somme, décaler les termes pour sommer de 1 à $n-1$ et remarquer que ce qu'on obtient annule la première somme.

□

Propqrrdem

Proposition 3.46.
Soit A un anneau commutatif de caractéristique première p. Alors $\sigma(x) = x^p$ est un automorphisme de l'anneau A. Nous avons la formule

$$(a+b)^p = a^p + b^p \tag{3.87}$$

pour tout $a, b \in A$.

Démonstration. Nous utilisons la formule du binôme de la proposition 3.43 et le fait que les coefficients binomiaux non extrêmes sont divisibles par p et donc nuls. □

3.3.1 Théorème multinomial

DEFooCQPRooFeWeOS

Définition 3.47 (multiindices[107]).
*Un **multiindice** pour $\mathbb{R}^n$ est un élément de $\mathbb{N}^n$. Sa taille est définie par*

$$|\alpha| = \sum_i \alpha_i. \tag{3.88}$$

Nous disons que $\beta \leqslant \alpha$ lorsque $\beta_i \leqslant \alpha_i$ pour tout i.

Nous définissons sa factorielle par

$$\alpha! = \prod_{k=1}^{n} (\alpha_k!) = \alpha_1! \times \ldots \times \alpha_n!. \tag{3.89}$$

En ce qui concerne les coefficients binomiaux, nous posons

$$\binom{\alpha}{\beta} = \frac{\alpha!}{(\alpha-\beta)!\beta!} = \prod_{i=1}^{n} \binom{\alpha_i}{\beta_i}. \tag{3.90}$$

EQooKBURooKFBxnt

Si $x \in \mathbb{R}^n$ et si $\alpha \in \mathbb{N}^n$, nous notons

$$x^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i}. \tag{3.91}$$

THOooNHAUooQvuytn

Théorème 3.48 (Théorème multinomial[107]).
Soit $x \in \mathbb{R}^n$, et $\alpha \in \mathbb{N}^n$. Nous avons

$$(x_1 + \ldots + x_n)^k = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=k}} \frac{k!}{\alpha!} x^\alpha. \tag{3.92}$$

Démonstration. Le cas $n = 2$ est la formule binomiale 3.43 :

$$(x_1 + x_2)^k = \sum_{l=0}^{k} \binom{k}{l} x_1^l x_2^{k-l} \tag{3.93a}$$

$$= \sum_{l=0}^{k} \frac{k!}{l!(k-l)!} x_1^l x_2^{k-l} \tag{3.93b}$$

$$= \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2} \tag{3.93c}$$

$$= \sum_{\substack{\alpha \in \mathbb{N}^2 \\ |\alpha|=k}} \frac{k!}{\alpha!} x^\alpha. \tag{3.93d}$$

Et maintenant, nous faisons une récurrence sur n, en tenant k fixé. Nous supposons que la formule est démontrée pour tout $n \leqslant N$, et nous allons avec $N + 1$. Soit $x = (x_1, \ldots, x_{N+1})$, et posons

$y = (x_1, \dots, x_N)$. Nous avons

$$(x_1 + \dots + x_{N+1})^k = \big((x_1 + \dots + x_N) + x_{N+1}\big)^k \tag{3.94a}$$
$$= \sum_{i+j=k} \frac{k!}{i!j!} (x_1 + \dots + x_N)^i x_{N+1}^j \tag{3.94b}$$
$$= \sum_{i+j=k} \frac{k!}{i!j!} \sum_{\substack{\alpha \in \mathbb{N}^N \\ |\alpha|=i}} x_{N+1}^j \frac{i!}{\alpha!} y^\alpha \tag{3.94c}$$
$$= \sum_{j=0}^{k} \sum_{\substack{\alpha \in \mathbb{N}^N \\ |\alpha|=k-j}} x_{N+1} \frac{k!}{j!} \frac{1}{\alpha!} y^\alpha \tag{3.94d}$$
$$= \sum_{\substack{\beta \in \mathbb{N}^{N+1} \\ |\beta|=k}} \frac{k!}{\beta_{N+1}! \alpha!} x_{N+1}^{\beta_{N+1}} x_1^{\alpha_1} \dots x_N^{\alpha_N} \qquad \text{cf. justif.} \tag{3.94e}$$
SUBEQooTQFGooHyUhCi
$$= \sum_{\substack{\beta \in \mathbb{N}^{N+1} \\ |\beta|=k}} x^\beta \frac{k!}{\beta!}. \tag{3.94f}$$

Justification pour (3.94e). Les deux sommes ensemble font une somme sur tous $\beta = (\alpha, j) \in \mathbb{N}^{N+1}$ avec $|\beta| = |\alpha| + j = (k - j) + j = k$. □

LEMooBSYRooDboTor

Lemme 3.49 ([1]).
Soit une application $f : \mathbb{N}^n \to \mathbb{R}$. Nous avons

$$\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=2}} \frac{1}{\alpha!} f(\alpha) = \frac{1}{2} \sum_{k,l} f(e_k + e_l). \tag{3.95}$$

Démonstration. Nous avons

$$\{\alpha \in \mathbb{N}^n \text{ tel que } |\alpha| = 2\} = \{e_k + e_l \text{ tel que } k, l = 1, \dots, n\}, \tag{3.96}$$

mais nous ne pouvons pas sommer en faisant $\sum_{\substack{\alpha \in \mathbb{N}^2 \\ |\alpha|=2}} = \sum_{k,l}$ parce que à droite, les termes (k, l) et (l, k) sont identiques.

Au lieu de ça, remarquons que nous avons une bijection

$$\begin{aligned} \varphi : \{(k, l) \in \{1, \dots, n\}^2 \text{ tel que } k \leqslant n\} &\to \{e_k + e_l\}_{k,l=1,\dots,n} \\ (k, l) &\mapsto e_k + e_l. \end{aligned} \tag{3.97}$$

Prouvons que cette application est injective. Si $\varphi(k, l) = \varphi(s, t)$, alors $e_k + e_l = e_s + e_t$. Le produit scalaire avec e_k donne

$$1 + \delta_{kl} = \delta_{st} + \delta_{tk}. \tag{3.98}$$

(1) **Si $k < l$** Alors $1 = \delta_{sk} + \delta_{tk}$. Nous avons donc forcément $k = s$ ou $k = t$, mais pas les deux en même temps.

(1a) **Si $s = k$** Dans ce csa $t \neq k$, et nous avons $e_k + e_l = e_k + e_t$, donc $e_l = e_t$ et donc $l = t$. Nous avons donc $(k, l) = (s, t)$, ce qu'il fallait.

(1b) **Si $t = k$** Alors $e_l = e_s$ et donc $l = s$, ce qui donnerait $(k, l) = (l, k)$, ce qui est impossible parce que $k < l$.

(2) **Si $k = l$** Alors $2e_k = e_s + t_t$. Le produit scalaire avec e_k donne $2 = \delta_{sk} + \delta_{tk}$ et donc $s = t = k$.

Nous pouvons donc sommer de la façon suivante :

$$\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=2}}\frac{1}{\alpha!}f(\alpha)=\sum_{k\leqslant l}\frac{1}{(e_k+e_l)!}f(e_k+e_l) \tag{3.99a}$$

$$=\sum_{k<l}f(e_k+e_l)+\sum_k\frac{1}{2}f(2e_k) \tag{3.99b}$$

SUBEQooZERNooQeAAgb

$$=\frac{1}{2}\sum_{k\neq l}f(e_k+e_l)+\sum_k\frac{1}{2}f(2e_k) \qquad \text{cf. justif.} \tag{3.99c}$$

$$=\frac{1}{2}\sum_{k,l}f(e_k+e_l). \tag{3.99d}$$

Justification pour (3.99c). Nous avons $f(e_l+e_l)=f(e_l+e_k)$. Donc $\sum_{k<l}f(e_k+e_l)=\sum_{k>l}f(e_k+e_l)$. □

3.4 Polynômes de plusieurs variables

DEFooZNHOooCruuwI

Définition 3.50.
*L'ensemble des **polynôme de** n **variables** sur l'anneau A est $A^{(\mathbb{N}^n)}$, c'est-à-dire l'ensemble des suites indexées par $\mathbb{N}^n$ et dont seulement une quantité finie de coefficients sont non nuls.*

Le produit sur $A[X_1,\dots,X_n]$ est défini par

$$(PQ)(k_1,\dots,k_n)=\sum_{\substack{(l_1,\dots,l_n),(m_1,\dots,m_n)\in\mathbb{N}^n\times\mathbb{N}^n\\ l_i+m_i=k_i}}P_{l_1,\dots,l_n}Q_{m_1,\dots,m_n}. \tag{3.100}$$

3.51.
Dans $A[X_1,\dots,X_n]$, la multiplication n'est pas la multiplication de fonctions $\mathbb{N}^n\to\mathbb{K}$, parce que le but est d'obtenir la multiplication usuelle au niveau des évaluations.

Définition 3.52.
*Si P est un polynôme de n variables sur A, et si $(x_1,\dots,x_n)\in A^n$, **l'évaluation** de P sur $(x_1,\dots,x_n)$ est*

$$P(x_1,\dots,x_n)=\sum_{(k_1,\dots,k_n)\in\mathbb{N}^n}P_{k_1,\dots,k_n}x_1^{k_1}\dots x_n^{k_n}. \tag{3.101}$$

Notez que la somme, bien que sur $\mathbb{N}^n$, est une somme finie.

3.53.
Comme dans le cas des polynômes d'une seule variable, les X_i dans la notation $A[X_1,\dots,X_n]$ sont à prendre à la légère. L'anneau des polynômes de n variables sur A aurait mieux fait d'être noté par exemple par $\mathcal{P}_n(A)$.

Le fait est que nous avons les polynômes élémentaires définis par

$$X_1(k_1,\dots,k_n)=\begin{cases}1 & \text{si } (k_1,\dots,k_n)=(1,0\dots,0)\\ 0 & \text{sinon.}\end{cases} \tag{3.102}$$

et que l'anneau des polynômes peut être vu comme A (les polynômes constants) étendus par les X_i.

Quoi qu'il en soit, les X_i dans la notation $A[X_1,\dots,X_n]$ sont des indices muets. L'anneau $A[X_1,\dots,X_n]$ est exactement le même que $A[T_1,\dots,T_n]$.

3.4.1 Divisibilité et classes d'association

DivisibiliteAnneauxInvEgires

Définition 3.54.
*On dit de deux éléments $a,b\in A$ qu'ils sont **associés** si il existe un inversible $u\in A$ tel que $a=ub$.*

*La **classe d'association** d'un élément $a\in A$ est l'ensemble des éléments qui lui sont associés.*

LemRmVTRq

Lemme 3.55.
Si A est un anneau intègre et si $a, b \in A$ sont deux éléments vérifiant $a \mid b$ et $b \mid a$, alors ils sont associés, c'est-à-dire qu'il existe un inversible $u \in A$ tel que $a = ub$.

Démonstration. Les hypothèses à propos de la divisibilité nous indiquent que $a = xb$ et $b = ya$ pour certains $x, y \in A$. Alors,

$$b(1 - yx) = 0. \tag{3.103}$$

Étant donné que A est intègre, cela montre que $b = 0$ ou $1 - yx = 0$. Si $b = 0$ nous avons immédiatement $a = 0$ et le lemme est prouvé. Si au contraire $yx = 1$, c'est que y et x sont inversibles et inverses l'un de l'autre. □

LEMooJBOXooYkMRrz

Lemme 3.56.
Si A est un anneau et si $a \in A$, la classe d'association de a est $aU(A)$ où $U(A)$ est l'ensemble des éléments inversibles de A.

Exemple 3.57.
Dans $\mathbb{Z}[i]$, les inversibles sont ± 1 et $\pm i$. Donc les éléments associés à z sont z, $-z$, iz et $-iz$.

Notons au passage que la notion de divisibilité dans $\mathbb{Z}[i]$ n'est pas immédiatement intuitive. En effet bien que 7 ne soit pas divisible par 2 (ni dans $\mathbb{Z}$ ni dans $\mathbb{Z}[i]$), le nombre $7 + 6i$ est divisible par $2 + i$ dans $\mathbb{Z}[i]$. En effet :

$$(2 + i)(4 + i) = 7 + 6i. \tag{3.104}$$

△

Exemple 3.58.
Si $\mathbb{K}$ est un corps, l'élément XY de $\mathbb{K}[X, Y]$ n'est pas premier parce que $XY \mid X^2Y^2$ alors que XY ne divise ni X^2 ni Y^2. △

¡¡ Avertissement/question à la lectrice !! 3.59
Est-ce que quelqu'un connaît un anneau contenant $\mathbb{Z}$ dans lequel 7 est divisible en 2 ?

Peut-être $\mathbb{Z}$ étendu par tous les $1/2^n$?

3.4.2 PGCD et PPCM

Pour le théorème de Bézout et autres opérations avec des modulo, voir le thème 3. Le pgcd et le ppcm sont définis en 1.186.

LEMooGWKMooLEepxz

Lemme 3.60.
Soient A un anneau intègre et $S \subset A$. Si δ est un PGCD de S, alors l'ensemble des PGCD de S est la classe d'association de δ.

De la même façon si μ est un PPCM de S, alors l'ensemble des PPCM de S est la classe d'association de μ.

Démonstration. Soient δ un PGCD de S et u un inversible dans A. Si $x \in S$ nous avons $\delta \mid x$ et donc $x = a\delta$. Par conséquent $x = au^{-1}u\delta$ et donc $u\delta$ divise x. De la même manière, si d divise x pour tout $x \in S$, alors d divise δ et donc $\delta = ad$ et $u\delta = aud$, ce qui signifie que d divise $u\delta$.

Dans l'autre sens nous devons prouver que si δ' est un autre PGCD de S, alors il existe un inversible $u \in A$ tel que $\delta' = u\delta$. Comme δ' divise x pour tout $x \in S$, nous avons $\delta' \mid \delta$, et symétriquement nous trouvons $\delta \mid \delta'$. Par conséquent (lemme 3.55), il existe un inversible u tel que $\delta = u\delta'$.

Le même type de raisonnement tient pour le PPCM. □

Si δ est un PGCD de S, nous dirons *par abus de langage* que δ est *le* PGCD de S, gardant en tête qu'en réalité toute sa classe d'association est PGCD. Nous noterons aussi, toujours par abus que $\delta = \text{pgcd}(S)$.

Remarque 3.61.
La classe d'association d'un élément n'est pas toujours très grande. Les inversibles dans $\mathbb{Z}$ étant seulement ± 1, nous pouvons obtenir l'unicité du PGCD et du PPCM en imposant qu'ils soient positifs.

Pour les polynômes, nous obtenons l'unicité en demandant que le PGCD soit unitaire.

Dans les cas pratiques, il y a donc en réalité peu d'ambiguïté à parler du PGCD ou du PPCM d'un ensemble.

3.4.3 Anneaux intègres et corps

Le fait d'être intègre pour un anneau n'assure pas le fait d'être un corps. Nous avons cependant ce résultat pour les anneaux finis.

PropanfinintimpCorp

Proposition 3.62.
Un anneau fini intègre est un corps.

Démonstration. Soit A un tel anneau. Soit $a \neq 0$. Les applications

$$l_a : x \mapsto ax \tag{3.105a}$$
$$r_a : x \mapsto xa \tag{3.105b}$$

sont injectives. En tant que applications injectives entre ensembles finis, elles sont surjectives. Il existe donc b et c tels que $1 = ba = ac$. Nous trouvons que b et c sont égaux parce que [13]

$$b = b(ac) = (ba)c = c. \tag{3.106}$$

Par conséquent b est un inverse de a. □

PropzhFgNJ

Proposition 3.63.
Soit $n \in \mathbb{N}^*$. Les conditions suivantes sont équivalentes :

(1) n est premier.

(2) $\mathbb{Z}/n\mathbb{Z}$ est un anneau intègre.

(3) $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Démonstration. L'équivalence entre les deux premiers points est le contenu du corolaire 1.243. Le fait que $\mathbb{Z}/n\mathbb{Z}$ soit un corps lorsque $\mathbb{Z}/n\mathbb{Z}$ est intègre est la proposition 3.62. Le fait que $\mathbb{Z}/n\mathbb{Z}$ soit intègre lorsque $\mathbb{Z}/n\mathbb{Z}$ est un corps est une propriété générale des corps : ce sont en particulier des anneaux intègres (lemme 1.198). □

3.5 Anneau factoriel

DEFooVCATooPJGWKq

Définition 3.64 (Anneau factoriel)**.**
*Un anneau commutatif A est **factoriel** si il vérifie les propriétés suivantes.*

(1) L'anneau A est intègre [14].

(2) Si $a \in A$ est non nul et non inversible, alors il admet une décomposition en facteurs irréductibles : $a = p_1 \dots p_k$ où les p_i sont irréductibles [15].

ITEMooKVDBooTASwVO

(3) Si $a = q_1 \dots q_m$ est une autre décomposition de a en irréductibles, alors $m = k$ et il existe une permutation [16] $\sigma \in S_k$ telle que p_i et $q_{\sigma(i)}$ soient associés [17].

13. Il faut être un peu souple avec les notations communément employées dans les ouvrages mathématiques, et que nous reprenons telles quelles. Dans l'équation qui suit, $b(ac)$ est le produit de b par l'élément ac, et non quelque chose comme le produit de b avec l'idéal (ac) par exemple.
14. Anneau intègre, définition 1.197.
15. Élément irréductible, définition 1.189.
16. Définition 1.274.
17. Définition 3.54.

Un anneau factoriel permet de caractériser le pgcd et le ppcm de nombres.

LEMooLVKMooSLuzao

Lemme 3.65 ([108]).
Soit un anneau factoriel A et un élément irréductible $p \in A$. Si $p|xy$, alors p divise x ou y ou les deux.

Démonstration. Comme nous sommes dans un anneau factoriel, nous pouvons écrire q, x et y comme produits d'irréductibles, et profiter de la plus ou moins unicité de ces décompositions (la propriété 3.64(3)). Nous notons $q = q_1 \dots q_k$, $x = x_1 \dots x_m$ et $y = y_1 \dots y_l$. L'égalité $pq = xy$ devient :

$$pq_1 \dots q_k = x_1 \dots x_m y_1 \dots y_l. \tag{3.107}$$

Il existe un des x_j ou y_j qui est associé à p. Fixons un i et disons que c'est x_i (si c'est un des y_j, adaptez) : il existe un inversible u tel que $x_i = pu$. Nous avons alors [18]

$$x = pux_1 \dots x_{i-1}x_{i+1} \dots x_m. \tag{3.108}$$

Donc p divise x et fin de l'histoire. □

PROPooOQSXooYidPQv

Proposition 3.66 ([108]).
Dans un anneau factoriel, tout élément irréductible est premier [19].

Démonstration. Soit un anneau factoriel A et un élément irréductible p dans A. Nous devons prouver qu'il est premier.

(1) **p est non nul** Si $p = 0$, nous avons $p = 0 \times 0$. Comme 0 n'est pas inversible (un anneau factoriel est par définition intègre), p serait un produit de deux non inversibles.

(2) **Non inversible** L'élément p est non inversible parce que c'est dans la définition d'un élément irréductible.

(3) **Si $p \mid xy$** Si p divise xy, alors il divise x ou y ; c'est le lemme 3.65.

□

LEMooGKOSooRKtfDJ

Lemme 3.67 ([1]).
Si A est un anneau factoriel, et si p est irréductible dans A, alors :

(1) L'idéal pA est premier.

(2) L'anneau A/pA est intègre.

Démonstration. En deux parties.

(1) **pA est premier** D'abord pA est strictement inclus dans A parce que p n'étant pas inversible, l'élément 1 n'est pas dans pA.
Soient $a, b \in A$ tels que $ab \in pA$. Cela signifie qu'il existe $x \in A$ tel que $px = ab$, ou encore que p divise ab. Le lemme 3.65 dit alors que p divise a ou b. Supposons pour fixer les idées que $p \mid a$. Il existe y tel que $a = py \in pA$.

(2) **L'anneau A/pA est intègre** C'est la proposition 1.231.

□

PROPooOXQMooVEzlyG

Proposition 3.68.
Soit une famille $\{a_n\}$ d'éléments de A qui se décomposent en irréductibles comme

$$a_i = \prod_k p_k^{\alpha_{k,i}}. \tag{3.109}$$

Alors

$$\text{pgcd}\{a_n\} = \prod_k p_k^{\min_i\{\alpha_{k,i}\}}. \tag{3.110}$$

De plus le PGCD est :

18. Le fait que A soit commutatif est utilisé partout.
19. Élément premier, définition 1.188.

(1) *Un multiple de tous les diviseurs communs des a_i.*
(2) *Unique pour cette propriété à multiple près par un inversible*[20].

De la même manière,

$$\text{ppcm}\{a_n\} = \prod_k p_k^{\max_i\{\alpha_{k,i}\}}. \tag{3.111}$$

Un anneau factoriel a une relation de préordre partiel donnée par $a < b$ si a divise b. En termes d'idéaux, cela donne l'ordre inverse de celui de l'inclusion[21] : $a < b$ si et seulement si $(b) \subset (a)$.

EXooCWJUooCDJqkr

Exemple 3.69.
L'anneau $\mathbb{Z}[i\sqrt{3}]$ n'est pas factoriel parce que 4 a au moins deux décompositions distinctes en irréductibles :

$$4 = 2 \cdot 2, \tag{3.112}$$

et

$$4 = (1 + i\sqrt{3})(1 - i\sqrt{3}). \tag{3.113}$$

△

Nous allons voir dans l'exemple 3.88 que $\mathbb{Z}[i\sqrt{2}]$ est factoriel parce qu'il sera euclidien.

3.5.1 Autour du théorème de Bézout

Rappel de notations : si A est un anneau et si $p \in A$, nous notons (p) l'idéal engendré par p, c'est-à-dire pA. C'est la définition 1.212.

LEMooARNUooXqrJGa

Lemme 3.70 ([109]).
Soient un anneau principal[22] *A ainsi que $a, b \in A$.*
(1) *d est PGCD de $\{a, b\}$ si et seulement si $(d) = (a) + (b)$.*
(2) *m est PPCM de $\{a, b\}$ si et seulement si $(m) = (a) \cap (b)$.*

Théorème 3.71 ([110, 109]).
Soient un anneau principal et $S \subset A$ non vide.
(1) *d est un PGCD de S si et seulement si il génère l'idéal engendré par S.*
(2) *m est PPCM de S si et seulement si $(m) = \bigcap_{s\in S}(s)$.*
Toute partie d'un anneau principal admet un PGCD et un PPCM[23].

¡¡ Avertissement/question à la lectrice !! 3.72
La preuve est à revoir.

Démonstration. Puisque l'anneau A est principal, tous ses idéaux sont principaux et donc engendrés par un seul élément. En particulier il existe $\delta, \mu \in A$ tels que

$$(\delta) = \sum_{s\in S}(s) \tag{3.114a}$$

$$(\mu) = \bigcap_{s\in S}(s) \tag{3.114b}$$

(1) **PGCD** Montrons ce que δ est un PGCD de S. Pour tout $x \in S$, nous avons $(x) \subset (\delta)$, et donc $\delta \mid x$. Par ailleurs si $d \mid x$ pour tout $x \in S$, nous avons $(x) \subset (d)$ et donc

$$\sum_{x\in S}(x) \subset (d), \tag{3.115}$$

20. Soyez prudent avec cette affirmation : je n'en n'ai pas de démonstrations sous la main et ne suis pas certain que ce soit vrai.
21. Voir proposition 3.3.
22. Définition 1.228.
23. Ce n'est pas aussi trivial que ça parce qu'il faut encore qu'il existe des éléments vérifiant les formules données.

puis $(\delta) \subset (d)$ et finalement $d \mid \delta$.

(2) **PPCM** Si $x \in S$ nous avons $(\mu) \subset (x)$ et donc $x \mid \mu$. D'autre part si $x \mid m$ pour tout $x \in S$, alors $(m) \subset (x)$ et donc $(m) \subset (\mu)$, finalement $\mu \mid m$.

□

CorimHyXy

Corolaire 3.73 (Théorème de Bézout[110]).
Soit un anneau principal A. Deux éléments $a, b \in A$ sont premiers entre eux si et seulement si il existe un couple $(u, v) \in A^2$ tel que

$$ua + vb = 1. \tag{3.116}$$

À la place de 1 on aurait pu écrire n'importe quel inversible.

Démonstration. Pour cette preuve, nous allons écrire $\operatorname{pgcd}(a, b)$ l'ensemble de PGCD de a et b, c'est-à-dire la classe d'association d'un PGCD.

Si a et b sont premiers entre eux, alors

$$1 \in \operatorname{pgcd}(a, b) = \sum_{x=a,b} (x) = (a) + (b). \tag{3.117}$$

À l'inverse, si nous avons $ua + vb = 1$, alors $1 \in (a) + (b)$, mais puisque $(a) + (b)$ est un idéal principal, $(1) = (a) + (b)$ et donc $1 \in \operatorname{pgcd}(a, b)$. □

Le lemme de Gauss est une application immédiate du théorème de Bézout. Il y aura aussi un lemme de Gauss à propos de polynômes (lemme 6.51), et une généralisation directe au théorème de Gauss, théorème 6.50.

LemSdnZNX

Lemme 3.74 (Lemme de Gauss[111]).
Soit A un anneau principal et $a, b, c \in A$ tels que a divise bc. Si a est premier avec c, alors a divise b.

Démonstration. Comme a est premier avec c, nous avons $\operatorname{pgcd}(a, c) = 1$ et Bézout (1.236) nous donne donc $s, t \in A$ tels que $sa + tc = 1$. En multipliant par b,

$$sab + tbc = b. \tag{3.118}$$

Mais les deux termes du membre de gauche sont multiples de a parce que a divise bc. Par conséquent b est somme de deux multiples de a et donc est multiple de a. □

LEMooQJGIooEtVnyj

Lemme 3.75 ([41]).
Soit un anneau principal A. Soient $a, b, c \in A$ tels que

(1) $\operatorname{pgcd}(a, b) = 1$

(2) $\operatorname{pgcd}(a, c) = 1$.

Alors $\operatorname{pgcd}(a, bc) = 1$.

Démonstration. Le théorème de Bézout 3.73 donne des éléments $u, v, x, y \in A$ tels que $ua + vb = 1$ et $xa + yc = 1$. En multipliant ces équations l'une avec l'autre,

$$(ua + vb)(xa + yc) = 1. \tag{3.119}$$

En développant, nous trouvons

$$(uax + uc + vbx)a + (vy)bc = 1. \tag{3.120}$$

Donc le théorème de Bézout (dans l'autre sens) nous indique que $\operatorname{pgcd}(a, bc) = 1$. □

3.5.2 Idéal premier

PROPooZICGooNmblhl

Proposition 3.76 ([112, 1], thème 6)**.**
Soit un anneau principal A et un élément $p \neq 0$ dans A. Nous avons équivalence de :

ITEMooBTEAooWlFUTX

(1) (p) est un idéal premier,

ITEMooKQRMooBNPDMX

(2) p est un élément premier,

ITEMooZYYJooCWiBhL

(3) p est un élément irréductible,

ITEMooHPAIooYoQzqD

(4) (p) est un idéal maximum propre[24].

Démonstration. En plusieurs implications.

(1) **(1) implique (2)** En plusieurs points.
- La condition $p \neq 0$ est dans les hypothèses de la proposition.
- Si p était inversible, nous aurions $(p) = A$ et donc pas que (p) est un idéal premier.
- Soient $a, b \in A$ tels que $p \mid ab$. En particulier, $(ab) \in (p)$. Mais comme (p) est un idéal premier, cela implique soit $a \in (p)$ soit $b \in (p)$. Donc soit p divise a soit p divise b.

(2) **(2) implique (3)** Un anneau principal est intègre ; c'est dans la définition 1.228. Dans un anneau intègre, tout élément premier est irréductible, c'est la proposition 1.224.

(3) **(3) implique (4)** Soit un idéal I contenant (p). Puisque A est principal, I est engendré par un seul élément ; soit $I = (a)$. Vu que $p \in I$, l'élément a divise p. Mais comme p est un élément premier, $a \mid p$ implique $a = p$ ou $a = 1$. Dans le premier cas, $I = (a) = (p)$, et dans le second cas, $I = (a) = (1) = A$. Donc (p) est bien un idéal maximum.
De plus l'idéal (p) est propre. En effet avoir $(p) = A$ dirait en particulier que $1 \in (p)$, c'est-à-dire qu'il existe $x \in A$ tel que $xp = 1$. Or p étant irréductible, il est non inversible.

(4) **(4) implique (1)** C'est la proposition 1.231(3).

□

Un exemple d'élément premier non irréductible est $[4]_6$ dans l'anneau non principal $\mathbb{Z}/6\mathbb{Z}$. Voir 3.86 et le lemme 3.87.

3.5.3 Anneau noethérien

DEFooPWMHooCnrQuJ

Définition 3.77.
Un anneau est dit ***noethérien*** *si toute suite croissante d'idéaux est stationnaire (à partir d'un certain rang).*

Montrer que tout anneau principal est noethérien est le premier pas pour montrer que tout anneau principal est factoriel.

LEMooHQPVooTfkhRV

Lemme 3.78.
Tout anneau principal[25] *est noethérien.*

Démonstration. Soit (J_n) une suite croissante d'idéaux et J la réunion. L'ensemble J est encore un idéal parce que les J_i sont emboités. Étant donné que l'idéal est principal nous pouvons prendre $a \in J$ tel que $J = (a)$. Il existe N tel que $a \in J_N$. Alors pour tout $n \geqslant N$ nous avons

$$J \subset J_N \subset J_n \subset J. \tag{3.121}$$

La première inclusion est le fait que $J = (a)$ et $a \in J_N$. La seconde est la croissance des idéaux et la troisième est le fait que J est une union. Par conséquent pour tout $n \geqslant N$ nous avons $J_N = J_n = J$. La suite est par conséquent stationnaire. □

24. Ce « propre » n'est pas dans l'énoncé sur Wikipédia. Je ne comprends pas pourquoi, et j'ai posé la question sur la page de discussion.
https://fr.wikipedia.org/wiki/Discussion:Idéal_premier

25. Définition 1.228.

Exemple 3.79.
Il y a moyen d'avoir un anneau noetherien non principal. C'est le cas de $\mathbb{Z}/6\mathbb{Z}$ dont nous parlerons dans 3.87. △

THOooANCAooBChmwp

Théorème 3.80 ([113, 41])**.**
Tout anneau principal est factoriel.

Démonstration. Soit un anneau principal A. Nous devons prouver les trois points de la définition 3.64. D'abord A est intègre parce que ça fait partie de la définition 1.228 d'un anneau principal. Le gros morceau est l'existence et l'unicité d'une décomposition en irréductibles.

mini lemme

Nous prouvons que si a est ni nul ni inversible, il est divisible par un irréductible.

Soit $a \in A$ que nous supposons être ni nul ni inversible. L'idéal aA ne contient pas tout A parce que a n'est pas inversible ; par le théorème de Krull 1.221 il existe un idéal maximal M contenant aA. Tous les idéaux étant principaux dans A, l'idéal M est principal.

Par définition d'idéal principal, il existe $p \in A$ tel que $M = pA$. Résumé :

$$a \in aA \subsetneq M = pA. \tag{3.122}$$

Comme l'idéal pA est maximal, la proposition 1.199 dit que p est irréductible. Et donc $p \mid a$ avec p est irréductible.

Existence

Nous définissons des suites (a_n) et (p_n) par récurrence. D'abord $a_0 = a$, et en suite, si a_n est défini nous définissons a_{n+1} et p_{n+1} de la façon suivante.

(1) Si a_n est inversible, nous nous arrêtons et la suite est finie.

(2) Si a_n n'est pas inversible, nous définissons a_{n+1} et p_{n+1} par la relation donnée par le mini-lemme :

$$a_n = p_{n+1}a_{n+1} \tag{3.123}$$

avec p_{n+1} irréductible.

Pour tout n assez petit pour que la suite ne soit pas finie, nous avons

$$a = a_0 = p_1a_1 = p_1p_2a_2 = \ldots = p_1 \ldots p_na_n \tag{3.124}$$

où les p_i sont irréductibles.

Nous montrons à présent que la suite des (a_n) est finie. Considérons les idéaux $I_n = a_nA$, et montrons qu'ils sont croissants en montrant que a_{n+1} n'est pas dans I_n.

En effet supposons que $a_{n+1} \in I_n = a_nA$. Il existerait $k \in A$ tel que $a_nk = a_{n+1}$, c'est-à-dire $p_{n+1}a_{n+1}k = a_{n+1}$. Vu que l'anneau est intègre et que les a_i sont non nums, nous simplifions par a_{n+1} : $p_{n+1}k = 1$. Cela n'est pas possible parce que p_{n+1} n'est pas inversible. Bref, nous avons uns suite strictement croissante d'idéaux.

L'anneau A est principal et donc noetherien (lemme 3.78). Donc toute suite croissante d'idéaux est stationnaire. Oh mais ça c'est pas possible si la suite des (a_n) est infinie parce que nous venons de prouver qu'elle est toujours strictement croissante.

Bon. Ben c'est que la suite des (a_n) est finie. Il existe donc un N tel que a_N est inversible. Pour ce N nous avons

$$a = a_0 = p_1a_1 = p_1p_2a_2 = \ldots = p_1 \ldots p_Na_N. \tag{3.125}$$

Il nous reste à prouver que p_Na_N est irréductible.

D'abord p_Na_N est non ivnersible parce que p_N n'est pas inversible et a_N est inversible. Ensuite supposons que $p_Na_N = st$ avec $s, t \in A$. Nous allons prouver qu'au moins s ou t est inversible. Nous avons

$$p_N = s(ta_N^{-1}). \tag{3.126}$$

Vu que p_N est irréductible, il n'est pas le produit de deux non-inversibles. Soit s est inversible (alors on a gagné), soit ta_N^{-1} est inversible, et alors t est inversible parce que a_N^{-1} est inversible.

Unicité

Soient des irréductibles $p_1, \dots, p_k$, $q_1, \dots, q_m$ tels que $a = p_1 \dots p_k = q_1 \dots q_m$. Nous utilisons le lemme 3.75 : si p_1 était premier avec tous chacun des q_j, il serait premier avec le produit, ce qui serait absurde parce que le produit est a et p_1 divise a. Bref, il y a un des q_j qui n'est pas premier avec p_1.

Nous considérons $\sigma_1 \in S_1$ tel que p_1 n'est pas premier avec $q_{\sigma_1(1)}$. Au passage, nous notons $q_k^{(1)} = q_{\sigma_1(k)}$. Soit un diviseur commun d non inversible entre p_1 et $q_1^{(1)}$. Nous avons des éléments u, v tels que

$$p_1 = du \quad q_1^{(1)} = dv. \tag{3.127}$$

Vu que p_1 et $q_1^{(1)}$ sont irréductibles, ils ne peuvent pas être produit de deux non inversibles. Donc u et v sont inversibles. Nous pouvons écrire $d = q_1^{(1)}v^{-1}$, et donc

$$p_1 = q_1^{(1)}uv^{-1}. \tag{3.128}$$

Nous récrivons $p_1 \dots p_k = q_1 \dots q_m$ avec ces valeurs :

$$dup_2 \dots p_k = dvq_2^{(1)} \dots q_m^{(1)}. \tag{3.129}$$

Nous simplifions par d et nous avons

$$(up_2)p_3 \dots p_k = (vq_2^{(1)})q_3^{(1)} \dots q_m^{(1)}. \tag{3.130}$$

Par récurrence nous construisons les σ_i et les $q_j^{(i)}$. □

PROPooPVFOooJvWRIh

Proposition 3.81.
Un corps est un anneau principal et un anneau factoriel.

Démonstration. Un anneau est principal quant :

- Commutatif
- intègre
- tous les idéaux sont principaux.

Un corps est toujours commutatif. Un corps est un anneau intègre par le lemme 1.198. Donc un corps, les seuls idéaux sont $\{0\}$ et $\mathbb{K}$. Le premier est principal parce que $0A = \{0\}$. Et le second est également principal parce que $\mathbb{K} = 1\mathbb{K}$.

Donc un corps est un anneau principal.

Le fait qu'un corps soit un anneau factoriel est maintenant le théorème 3.80. □

EXooYCTDooGXAjGC

Exemple 3.82 ($\mathbb{Z}[i\sqrt{5}]$ n'est ni factoriel ni principal)**.**
Puisque $(i\sqrt{5})^2 = -5$, les éléments de $\mathbb{Z}[i\sqrt{5}]$ sont les éléments de $\mathbb{C}$ de la forme $a + bi\sqrt{5}$ avec $a, b \in \mathbb{Z}$. Nous définissons la **norme** sur $\mathbb{Z}[i\sqrt{5}]$ par [26]

$$\begin{aligned} N\colon \mathbb{Z}[i\sqrt{5}] &\to \mathbb{N} \\ z &\mapsto z\bar{z}. \end{aligned} \tag{3.131}$$

Le fait que ce soit à valeurs dans $\mathbb{N}$ est un simple calcul :

$$N(x + iy\sqrt{5}) = (x + iy\sqrt{5})(x - iy\sqrt{5}) = x^2 + 5y^2. \tag{3.132}$$

De plus N est multiplicative : $N(z_1z_2) = N(z_1)N(z_2)$.

26. C'est le carré de la norme usuelle, mais c'est l'usage dans le milieu.

Nous pouvons maintenant déterminer les inversibles de $\mathbb{Z}[i\sqrt{5}]$. Si α est inversible, alors il existe β tel que $\alpha\beta = 1$. Au niveau de la norme,

$$N(\alpha)N(\beta) = 1, \tag{3.133}$$

ce qui implique que $N(\alpha) = 1$. Or l'équation $x^2 + 5y^2 = 1$ dans $\mathbb{N}$ donne $y = 0$, $x = \pm 1$.

Au final, les inversibles de $\mathbb{Z}[i\sqrt{5}]$ sont ± 1.

L'anneau $\mathbb{Z}[i\sqrt{5}]$ n'est alors pas factoriel (définition 3.64) parce que

$$2 \times 3 = (1 + i\sqrt{5})(1 - i\sqrt{5}). \tag{3.134}$$

Cela donne deux décompositions du nombre 6 en produit d'éléments non associés [27] (2 n'est associé qu'à 2 et -2) parce que les inversibles sont 1 et -1.

Le fait que $\mathbb{Z}[i\sqrt{5}]$ ne soit pas factoriel implique qu'il ne soit pas principal, théorème 3.80. △

3.6 Anneau $\mathbb{Z}/6\mathbb{Z}$

SECooSWGKooEeOZTO

Nous allons donner quelques propriétés de cet anneau, et en particulier voir que dans cet anneau non intègre, la notion d'élément irréductible n'est pas très intéressante.

Voici pour commencer un calcul de la table de multiplication de $A = \mathbb{Z}/6\mathbb{Z}$. Pour les multiples de (par exemple) $[4]_6$ nous écrivons

$$1 \times [4]_6 = [4_6] \tag{3.135}$$

et ensuite

$$2 \times [4]_6 = [8]_6 = [2]_6, \tag{3.136}$$

puis

$$3 \times [4]_6 = [2 + 4]_6 = [6]_6 = [0]_6, \tag{3.137}$$

et caetera. Le résultat est :

$\times$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[0]_6$	0	0	0	0	0	0
$[1]_6$	0	1	2	3	4	5
$[2]_6$	0	2	4	0	2	4
$[3]_6$	0	3	0	3	0	3
$[4]_6$	0	4	2	0	4	2
$[5]_6$	0	5	4	3	2	1

(3.138)

Pour ne pas alourdir, nous n'avons pas écrit $[x]_6$ partout au lieu de x.

3.83 (Inversibles)**.**
Les éléments inversibles de $\mathbb{Z}/6\mathbb{Z}$ sont ceux qui ont un $[1]_6$ dans leur table de multiplication. Ce sont donc

$$U(\mathbb{Z}/6\mathbb{Z}) = \{[1]_6, [5]_6\}. \tag{3.139}$$

3.84 (Diviseurs de zéro)**.**
Les diviseurs de zéro sont ceux qui ont un $[0]_6$ dans leur table de multiplication, c'est-à-dire

$$\{[2]_6, [3]_6, [4]_6\}. \tag{3.140}$$

3.85 (Irréductibles)**.**
Les irréductibles sont ceux qui ne sont ni inversibles ni produit de deux éléments non inversibles. Les non inversibles sont :

$$\{[0]_6, [2]_6, [3]_6, [4]_6\}. \tag{3.141}$$

27. Définition 3.54.

Ils sont candidats à être irréductibles. Les produits de ces éléments (on oublie les crochets) sont :

$$2 \times 2 = 4 \tag{3.142a}$$
$$2 \times 3 = 0 \tag{3.142b}$$
$$2 \times 4 = 2 \tag{3.142c}$$
$$3 \times 3 = 3 \tag{3.142d}$$
$$3 \times 4 = 0 \tag{3.142e}$$
$$4 \times 4 = 4. \tag{3.142f}$$

Donc $[0]_6$, $[2]_6$, $[3]_6$ et $[4]_6$ ne sont plus candidats à être irréductible. Bref, il ne reste aucun candidats.

L'anneau $\mathbb{Z}/6\mathbb{Z}$ n'a aucun élément irréductible.

NORMooAXOKooAQMXoB

3.86 (Éléments premiers)**.**
Les éléments non nuls et non inversibles sont 2, 3 et 4.

(1) **Pour 2** L'élément $[2]_6$ divise 2, 4 et 0.
   - Les (a, b) tels que $ab = 2$ sont : $(1, 2)$, $(2, 4)$ et $(5, 4)$. L'élément 2 divise donc toujours a ou b.
   - Les (a, b) tels que $ab = 4$ sont : $(1, 4)$, $(2, 5)$ et $(4, 4)$. L'élément 2 divise donc toujours a ou b.
   - Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 2 divise donc toujours a ou b. En particulier, $[2]_6$ divise $[0]_6$; c'est important.

   Donc $[2]_6$ est un élément premier.

(2) **Pour 3** L'élément $[3]_6$ divise 3 et 0.
   - Les (a, b) tels que $ab = 3$ sont : $(1, 3)$ et $(3, 5)$. L'élément 3 divise donc toujours a ou b.
   - Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 3 divise donc toujours a ou b.

   Donc $[3]_6$ est un élément premier. L'élément $[4]_6$ divise 4, 2 et 0.
   - Les (a, b) tels que $ab = 4$ sont : $(1, 4)$, $(2, 5)$ et $(4, 4)$. L'élément 4 divise donc toujours a ou b.
   - Les (a, b) tels que $ab = 2$ sont : $(1, 2)$, $(2, 4)$ et $(5, 4)$. L'élément 4 divise donc toujours a ou b.
   - Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 4 divise donc toujours a ou b.

   Donc $[4]_6$ est un élément premier.

Au final, les éléments premiers dans $\mathbb{Z}/6\mathbb{Z}$ sont

$$\{[2]_6, [3]_6, [4]_6\}. \tag{3.143}$$

Vous noterez que $\mathbb{Z}/6\mathbb{Z}$ a des éléments premiers non irréductibles. Cela est un contre-exemple à la proposition 3.76 dans le cas d'un anneau non-intègre.

LEMooZSELooGOFEIz

Lemme 3.87 ([1])**.**
L'anneau $\mathbb{Z}/6\mathbb{Z}$ est noetherien, mais ni intègre ni principal[28].

Démonstration. Comme c'est un anneau fini, toute suite croissante de quoi que ce soit devient stationnaire ; donc $\mathbb{Z}/6\mathbb{Z}$ est noetherien.

Puisque $\mathbb{Z}/6\mathbb{Z}$ a des diviseurs de zéro, il n'est pas intègre. Et s'il n'est pas intègre, il n'est pas factoriel non plus. □

28. Toutes les définitions dans le thème 6.

ExeDufyZI

Exemple 3.88.
Prouvons que $\mathbb{Z}[i\sqrt{2}]$ est un anneau euclidien. Pour cela nous démontrons que

$$\begin{aligned} N\colon \mathbb{Z}[i\sqrt{2}] &\to \mathbb{N} \\ a+bi\sqrt{2} &\mapsto a^2+2b^2 \end{aligned} \tag{3.144}$$

EqOZUIooZGmHWJ

est un stathme euclidien.

Soient $z = a + bi\sqrt{2}$, $t = a' + b'i\sqrt{2}$. Nous cherchons q et r tels que la division euclidienne s'écrive $z = qt + r$. Soient $\alpha, \beta \in \mathbb{Q}$ tels que

$$\frac{z}{t} = \alpha + \beta i\sqrt{2}. \tag{3.145}$$

Nous désignons par $\alpha+\epsilon_1$ et $\beta+\epsilon_2$ les entiers les plus proches de α et β. Nous avons $|\epsilon_1|, |\epsilon_2| \leqslant \frac{1}{2}$. Nous posons alors naturellement

$$q = (\alpha+\epsilon_1) + (\beta+\epsilon_2)i\sqrt{2} \tag{3.146}$$

et nous calculons $r = z - qt$:

$$2b'\epsilon_2 - a'\epsilon_1 + i\sqrt{2}\big(\epsilon_1 b' - a'\epsilon_2\big). \tag{3.147}$$

Nous trouvons

$$N(r) = a'^2\epsilon_1^2 + 4b'^2\epsilon_2^2 + 2a'^2\epsilon_1^2 + 2b'^2\epsilon_2^2 \leqslant \frac{3}{4}a'^2 + \frac{3}{2}b'^2. \tag{3.148}$$

D'autre part $N(t) = a'^2 + 2b'^2$, et nous avons donc bien $N(r) < N(t)$.

En ce qui concerne la seconde propriété du stathme, un petit calcul montre que

$$N(zt) = (a^2+2b^2)(a'^2+2b'^2), \tag{3.149}$$

et tant que $t \neq 0$ nous avons bien $N(zt) > N(z)$. △

Notons en particulier que $\mathbb{Z}[i\sqrt{2}]$ est factoriel et principal.

ExluqIkE

Exemple 3.89 (Décomposition en facteurs irréductibles dans $\mathbb{Z}[i\sqrt{2}]$)**.**
Les éléments inversibles de $\mathbb{Z}[i\sqrt{2}]$ sont ± 1, donc deux éléments a et b sont associés (définition 3.54) si et seulement si $a = \pm b$.

De plus si p est irréductible [29], alors $-p$ est irréductible. Les éléments irréductibles de $\mathbb{Z}[i\sqrt{2}]$ arrivent donc par pairs d'éléments associés. Soit $\{p_i\}$ une sélection de un élément irréductible parmi chaque paire. Tout élément x de $\mathbb{Z}[i\sqrt{2}]$ peut alors être écrit $x = \pm p_1^{\alpha_1}\dots p_n^{\alpha_n}$. Cette écriture va être pratique pour comparer des décompositions en facteurs irréductibles d'éléments. △

Le lemme suivant fait en pratique partie de l'exemple 3.92, mais nous l'isolons pour plus de clarté [30].

LemTScCIv

Lemme 3.90.
Si a et b sont deux éléments premiers entre eux de $\mathbb{Z}[i\sqrt{2}]$, et si il existe $y \in \mathbb{Z}[i\sqrt{2}]$ tel que $ab = y^3$, alors a et b sont des cubes (dans $\mathbb{Z}[i\sqrt{2}]$).

Démonstration. D'après l'exemple 3.89 nous pouvons écrire

$$y = \pm p_1^{\sigma_1}\dots p_n^{\sigma_n} \tag{3.150a}$$
$$a = \pm p_1^{\alpha_1}\dots p_n^{\alpha_n} \tag{3.150b}$$
$$b = \pm p_1^{\beta_1}\dots p_n^{\beta_n} \tag{3.150c}$$

29. Définition 1.189
30. Merci à Marvoir pour m'avoir souligné le manque.

où les p_i sont les irréductibles de $\mathbb{Z}[i\sqrt{2}]$ « modulo ± 1 » au sens où la liste des irréductibles est $\{p_i\} \cup \{-p_i\}$ (union disjointe). Étant donné que a et b sont premiers entre eux, α_i et β_i ne peuvent pas être non nuls en même temps alors que leur somme doit faire $3\sigma_i$. Nous avons donc pour chaque i soit $\alpha_i = 3\sigma_i$ soit $\beta_i = 3\sigma_i$ (et bien entendu si $\sigma_i = 0$ alors $\alpha_i = \beta_i = 0$).

Étant donné que ± 1 sont également deux cubes, a et b sont bien des cubes.

Notons que nous avons utilisé de façon capitale le fait que $\mathbb{Z}[i\sqrt{2}]$ était factoriel. □

3.6.1 Équations diophantiennes

ExZPVFooPpdKJc

Exemple 3.91.
L'équation diophantienne

$$x^2 = 3y^2 + 8 \tag{3.151}$$

n'a pas de solution. En effet si nous prenons l'équation modulo 3 nous obtenons

$$[x^2]_3 = [3y^2 + 8]_3 = [8]_3 = [2]_3. \tag{3.152}$$

Or dans $\mathbb{Z}/3\mathbb{Z}$, aucun carré n'est égal à deux : $0^2 = 0 \neq 2$, $1^2 = 1 \neq 2$ et $2^2 = 4 = 1 \neq 2$. △

ExmuQisZU

Exemple 3.92.
Résolvons l'équation diophantienne

$$x^2 + 2 = y^3. \tag{3.153}$$

Une première remarque est que x doit être impair. En effet si $x = 2k$, nous devons avoir y^3 pair. Mais un cube pair est divisible par 8. Donc $y^3 = 8l$ pour un certain l. L'équation devient $4k^2 + 2 = 8l$, c'est-à-dire $2k^2 + 1 = 4l$. Le membre de gauche est impair tandis que celui de droite est pair. Impossible.

Nous pouvons écrire l'équation sous la forme $x^2 + 2 = (x + i\sqrt{2})(x - i\sqrt{2})$. Et nous considérons $\mathbb{Z}[i\sqrt{2}]$ muni de son stathme N donné par (3.144).

L'élément $i\sqrt{2}$ est irréductible parce que $N(i\sqrt{2}) = 2$, et si nous avions $i\sqrt{2} = pq$, alors nous aurions $N(p)N(q) = 2$, ce qui n'est possible que si $N(p)$ ou $N(q)$ vaut 1.

Nous prouvons maintenant que les éléments $x + i\sqrt{2}$ et $x - i\sqrt{2}$ sont premiers entre eux. Supposons que d soit un diviseur commun ; alors il divise aussi la somme et la différence. Donc d divise à la fois $2x$ et $2i\sqrt{2}$.

Étant donné que $i\sqrt{2}$ est irréductible et que $2i\sqrt{2} = (-i\sqrt{2})^3$, les diviseurs de $2i\sqrt{2}$ sont les puissances de $(-i\sqrt{2})$. Alors nous devrions avoir $d = (i\sqrt{2})^\beta$ et donc

$$x = (i\sqrt{2})^\beta q \tag{3.154}$$

pour un certain $q \in \mathbb{Z}[i\sqrt{2}]$. Dans ce cas nous avons $N(x) = 2^\beta N(q)$, mais nous avons déjà précisé que x ne pouvait pas être pair, donc $\beta = 0$ et nous avons $d = 1$.

Comme les nombres $x \pm i\sqrt{2}$ sont premiers entre eux et que leur produit doit être un cube, ils doivent être séparément des cubes (lemme 3.90). Nous devons donc résoudre séparément $x \pm i\sqrt{2} = y^3$.

Cherchons les x et y entiers tels que $x + i\sqrt{2} = y^3$. Si nous posons $z = a + bi\sqrt{2}$, il suffit de calculer z^3 :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('a,b')
(a, b)
sage: z=a+I*sqrt(2)*b
sage: (z**3).expand()
3*I*sqrt(2)*a^2*b - 2*I*sqrt(2)*b^3 + a^3 - 6*a*b^2
```

En identifiant cela à $x+i\sqrt{2}$ nous trouvons le système

$$\begin{cases} x = a^3 - 6ab^2 & (3.155a)\\ 1 = 3a^2b - 2b^3 & (3.155b) \end{cases}$$

où, nous le rappelons, x, a et b sont des entiers. La seconde équation montre que b doit être inversible : $b(3a^2-2b^2)=1$. Il y a donc les possibilités $b=\pm1$. Pour $b=1$ l'équation devient $3a^2-2=1$, c'est-à-dire $a=\pm1$. Pour $b=-1$ l'équation devient $3a^2-2=-1$, impossible. En conclusion les possibilités sont

$$(x,z) = (-5, 1+i\sqrt{2}) \quad (3.156a)$$
$$(x,z) = (5, -1+i\sqrt{2}) \quad (3.156b)$$
$$(3.156c)$$

Le travail avec $x-i\sqrt{2}$ donne les mêmes résultats.

Les deux solutions de l'équation $x^2+2=y^3$ sont alors $(5,3)$ et $(-5,3)$. △

3.6.2 Triplets pythagoriciens et équation de Fermat pour $n=4$

Définition 3.93.
*Les solutions entières (positives) de l'équation $x^2+y^2=z^2$ sont appelés **triplets pythagoriciens**.*

Ils donnent toutes les possibilités de triangles rectangles dont les côtés ont des longueurs entières.

Définition 3.94.
*On dit qu'un triplet pythagoricien est **primitif** si les trois nombres sont premiers dans leur ensemble*[31].

Remarquons que c'est équivalent à montrer que les trois nombres sont premiers deux à deux : en effet, si deux parmi x, y et z sont divisibles par un nombre, alors tous les trois sont divisibles par ce nombre[32], donc les nombres x, y et z sont premiers deux à deux.

LemTripletsPythagoriciensPrimitifs

Lemme 3.95.
Dans un triplet pythagoricien primitif (x,y,z), on a toujours z impair et :

- *soit x impair et y pair ;*
- *soit x pair et y impair.*

Démonstration. Remarquons que le fait d'imposer que le triplet soit primitif, interdit aux nombres x et y d'être pairs en même temps. En effet, si c'était le cas, alors x^2 et y^2 seraient aussi pairs, donc leur somme z^2 aussi, d'ou z serait pair et les trois nombres ne seraient pas premiers entre eux.

Nous montrons à présent que les nombres x et y ne sont pas tous les deux impairs. Par l'absurde, si $x=2a+1$, nous avons $x^2=4a^2+4a+1\in[1]_4$; de la même manière, $y^2\in[1]_4$. On en déduit alors que $z^2=x^2+y^2\in[2]_4$. Le nombre z^2 est donc pair, donc z est pair : disons $z=2c$. Alors, $z^2=4c^2\in[0]_4$. Comme les classes modulo 4 sont disjointes, nous aboutissons à une contradiction. □

PropXHMLooRnJKRi

Proposition 3.96 (Triplets pythagoriciens[114, 115])**.**
Un triplet $(x,y,z)\in(\mathbb{N}^*)^3$ est solution de $x^2+y^2=z^2$ si et seulement si il existe $d\in\mathbb{N}$ et $u,v\in\mathbb{N}^*$ premiers entre eux tels que subeqLVHFooVgWsFx

$$\begin{cases} x = d(u^2-v^2) & (3.157a)\\ y = 2duv & (3.157b)\\ z = d(u^2+v^2) & (3.157c) \end{cases}$$

31. Définition 1.261.
32. Parce que k et k^2 ont les mêmes facteurs premiers.

ou SUBEQSooRQKBooCBsqYA

$$\begin{cases} x = 2duv & (3.158a)\\ y = d(u^2 - v^2) & (3.158b)\\ z = d(u^2 + v^2) & (3.158c) \end{cases}$$

La différence entre les deux est seulement d'inverser les rôles de x et y.

Démonstration. Montrons d'abord que les formules proposées sont bien des solutions ; nous vérifions (3.157) :

$$x^2 + y^2 = d^2(u^2 - v^2)^2 + 4d^2u^2v^2 = d^2(u^2 + v^2)^2, \tag{3.159}$$

qui correspond bien au z^2 proposé.

Une vérification du même style fonctionne pour (3.158).

Nous allons maintenant prouver la réciproque : toute solution est d'une des deux formes proposées. Déterminer les triplets primitifs suffira parce que si (x, y, z) n'est pas une solution primitive, alors en posant $k = \text{pgcd}(x, y, z)$, le triplet $\left(\frac{x}{k}, \frac{y}{k}, \frac{z}{k}\right)$ est primitif. Connaissant les triplets primitifs, nous obtenons tous les autres par simple multiplication.

Soit donc (x, y, z) un triplet pythagoricien primitif. Sans perte de généralité [33], grâce au lemme 3.95, nous pouvons supposer x pair tandis que y et z seront impairs. Comme $x^2 = (z + y)(z - y)$, nous avons

$$\left(\frac{x}{2}\right)^2 = \frac{1}{4}(z + y)(z - y). \tag{3.160}$$

Puisque z et y sont premiers entre eux, les nombres $z + y$ et $z - y$ sont également premiers entre eux [34]. Vu que $(z+y)(z-y)$ est divisible par 4, soit $z+y$ soit $(z-y)$ est divisible par 4. Pour fixer les idées nous supposons que c'est $z + y$, et nous écrivons

$$\left(\frac{x}{2}\right)^2 = \left(\frac{z + y}{4}\right)(z - y). \tag{3.161}$$

Les facteurs premiers (qui arrivent au moins au carré) de $(x/2)$ sont chacun soit dans $(z + y)/4$ soit dans $(z - y)$. Tout deux sont donc des carrés parfaits. Nous posons

$$\frac{z + y}{4} = u^2, \quad z - y = v^2. \tag{3.162}$$

Bien entendu u et v sont non nuls parce que nous avons exclu la possibilité de triplets dont un élément serait nul. Avec tout cela nous avons $(x/2)^2 = u^2v^2$ et donc $x = 2uv$ puis par somme et différence :

$$\begin{cases} x = 2uv & (3.163a)\\ y = v^2 - u^2 & (3.163b)\\ z = u^2 + v^2, & (3.163c) \end{cases}$$

ce qu'il fallait. □

Remarque 3.97.
Les solutions dans lesquelles x, y ou z sont nuls sont faciles à classer. La solution $(1, 0, 1)$ n'est pas dans les formes proposées. En effet elle ne peut pas être de la première forme : avoir $y = 0$ demanderait qu'un nombre parmi d, u et v soit nul, ce qui est interdit. La solution $(1, 0, 1)$ ne peut pas non plus être de la seconde forme parce que x y est pair.

propFKKKooFYQcxE

Proposition 3.98 ([114])**.**
Les équations $x^4 + y^4 = z^2$ et $x^2 + y^4 = z^4$ n'ont pas de solution dans $(\mathbb{N}^*)^3$.

33. En échangeant les rôles de x et y ici, nous obtenons à la fin la seconde forme des solutions.
34. Si $z - y = kn$ et $z + y = km$, faisant la somme et la différence on voit que y et z sont divisibles par k.

Démonstration. Si la première équation n'a pas de solution, alors la seconde n'en a pas non plus parce que z^4 est un carré. Nous nous concentrons donc sur l'équation $x^4+y^4=z^2$ et nous supposons qu'il existe au moins une solution dans $(\mathbb{N}^*)^3$. Nous en choisissons une (x,y,z) avec z minimum (les z dans différentes solutions étant dans $\mathbb{N}$, il en existe forcément un minimum [35]). Du coup, les trois nombres x, y et z sont premiers dans leur ensemble parce qu'une division par leur pgcd donnerait une nouvelle solution qui contredirait la minimalité de z.

Nous posons $x^4=\bar{x}^2$ et $y^4=\bar{y}^2$. Ils vérifient l'équation $\bar{x}^2+\bar{y}^2=z^2$ et par la proposition 3.96, il existe $u,v\in\mathbb{N}^*$ premiers entre eux tels que, sans perte de généralité [36], on ait

$$\begin{cases} \bar{x}=2uv & (3.164a)\\ \bar{y}=u^2-v^2 & (3.164b)\\ z=u^2+v^2. & (3.164c)\end{cases}$$

Si u est pair, alors v est impair (et inversement) parce que $\operatorname{pgcd}(u,v)=1$ Si u est pair, alors $u=2a$ et $v=2b+1$, ce qui donne $\bar{y}=4a^2-4b^2-4b-1\in[-1]_4$. Or nous avons déjà vu qu'un carré est dans $[0]_4$ ou dans $[1]_4$. Il faut donc que u soit impair. Le lemme 3.95 implique alors que v soit pair.

De l'équation 3.164b, nous en déduisons que $v^2+\bar{y}=u^2$; de plus u^2, v^2 et $\bar{y}$ sont premiers dans leur ensemble : en effet, u et v sont premiers entre eux, et si l'un parmi u^2 et v^2 a un facteur commun avec $\bar{y}$, alors l'autre doit l'avoir aussi (dans une égalité $a+b=c$, si deux des nombres ont un diviseur commun, le troisième l'a aussi). Comme $\bar{y}=y^2$, le triplet (v,y,u) est un triplet pythagoricien primitif. Nous appliquons de nouveau la proposition 3.96, en se souvenant que v est pair : il existe donc deux nombres r et s premiers entre eux tels que

$$\begin{cases} v=2rs & (3.165a)\\ y=r^2-s^2 & (3.165b)\\ u=r^2+s^2. & (3.165c)\end{cases}$$

Avec cela, $\bar{x}=2uv=4rs(r^2+s^2)$. Remarquons que les trois nombres r, s et r^2+s^2 sont premiers entre eux dans leur ensemble ; or, comme $\bar{x}$ est un carré ces nombres doivent séparément être des carrés :

$$\begin{cases} r=\alpha^2 & (3.166a)\\ s=\beta^2 & (3.166b)\\ r^2+s^2=\gamma^2. & (3.166c)\end{cases}$$

En mettant les deux premiers dans la troisième équation, nous avons $\alpha^4+\beta^4=\gamma^2$. Donc $(\alpha^2,\beta^2,\gamma)$ est une solution. Nous allons montrer que $\gamma<z$, ce qui terminera la preuve, puisque z était supposé minimal. Nous avons :

$$\begin{aligned} z&=u^2+v^2 && \text{par 3.164c}\\ &=r^2+s^2+4r^2s^2 && \text{par 3.165}\\ &=\gamma^2+4r^2s^2\\ &>\gamma^2,\end{aligned}$$

et a fortiori $\gamma<z$. □

3.7 Polynômes à coefficients dans un anneau commutatif

Lemme 3.99.
Nous considérons un polynôme $P\in A[X]$, et le quotient $A[X]/(P)$. Pour tout polynôme $Q\in A[X]$ nous avons les égalités

$$Q(\bar{X})=\overline{Q(X)}=\bar{Q}. \tag{3.167}$$

35. Voir quelque chose comme le lemme 1.143.
36. En inversant les rôles de x et y au besoin.

Démonstration. Si $Q = \sum_k a_k X^k$, alors par la linéarité de la prise de classes,

$$\bar{Q} = \sum_k a_k \overline{X^k}. \tag{3.168}$$

EQooXQRMooIPGFVM

Nous insistons sur le fait que cette égalité n'est rien d'autre que l'itération de la définition de la somme dans l'espace quotient : $\bar{x} + \bar{y} = \overline{x+y}$ ainsi que du produit $k\bar{x} = \overline{kx}$ (définition 1.184). Toujours par définition du produit appliqué à l'élément $\bar{X}$ nous avons $(\bar{X})^2 = \overline{X^2}$; par récurrence $\overline{X^k} = \bar{X}^k$, et

$$\bar{Q} = \sum_k a_k \bar{X}^k = Q(\bar{X}). \tag{3.169}$$

Le fait que $\bar{Q} = \overline{Q(X)}$ n'est rien d'autre que le fait que dans $A[X]$ nous avons $Q = Q(X)$, comme expliqué dans le lemme 1.366. □

3.7.1 Monômes

3.100.
Les éléments de la forme λX^k avec $\lambda \in A$ et $k \in \mathbb{N}$ sont des **monômes**.

Nous allons aussi considérer

$$A_n[X] = \{P \in A[X] \text{ tel que } \deg(P) \leqslant n\}. \tag{3.170}$$

Cet ensemble est un sous-module libre.

3.7.2 Évaluation

Soit $P \in A[X]$. À priori, P n'est qu'une suite dans A indexée par $\mathbb{N}$.

Nous avons déjà défini son évaluation sur un élément $\alpha \in A$ dans la définition 1.362 :

$$P(\alpha) = \sum_k a_k \alpha^k. \tag{3.171}$$

Cette somme est toujours finie.

NORMooQFTJooLBcPxl

3.101.
L'ensemble $A[X]$ est une algèbre et donc un espace vectoriel. Il possède un unique élément nul qui est celui dont tous les coefficients sont nuls ; cela est immédiat par la construction en tant que suites presque nulles.

Il n'y a à priori pas équivalence entre le fait d'être un polynôme nul et le fait de s'évaluer à zéro sur tous les éléments de A. Cela sera discuté dans le théorème 6.111 et l'exemple 19.34.

DEFooRFBFooKCXQsv

Définition 3.102.
Soient un anneau A et un anneau B qui contient A (comme sous-anneau). Pour $\alpha \in B$ nous définissons $A[\alpha]_B$ comme étant l'intersection de tous les sous-anneaux de B contenant A.

Comme dit plus haut, nous nous permettons d'écrire $A[\alpha]$ sans préciser B lorsque ce dernier sera clair dans le contexte.

PROPooPMNSooOkHOxJ

Proposition 3.103.
Soient un anneau A et un anneau B qui contient A (comme sous-anneau). Pour tout $\alpha \in B$ nous avons

$$A[\alpha] = \{P(\alpha) \text{ tel que } P \in A[X]\} \tag{3.172}$$

où encore une fois, $P(\alpha)$ est calculé dans B ; le contexte est clair là-dessus.

Démonstration. Si A' est un sous-anneau de B contenant A et α, alors A' contient tous les $P(\alpha)$ avec $P \in A[X]$. Nous avons donc

$$\{P(\alpha) \text{ tel que } P \in A[X]\} \subset A[\alpha]. \tag{3.173}$$

Par ailleurs, $\{P(\alpha)$ tel que $P \in A[X]\}$ est un sous-anneau de B contenant A et α. Donc $A[\alpha]$ y est inclus. □

3.7.3 Polynômes sur un anneau intègre

ThoBUEDrJ

Théorème 3.104.
L'anneau A est intègre si et seulement si $A[X]$ est intègre.

Démonstration. Soient P et Q des éléments non nuls de $A[X]$. Puisque l'anneau A est intègre, nous avons

$$\deg(PQ) = \deg(P) + \deg(Q) \tag{3.174}$$

et le produit ne peut pas être nul. L'anneau $A[X]$ est donc intègre.

Si $A[X]$ est intègre, A est intègre parce qu'il peut être considéré comme sous-anneau de $A[X]$. □

3.105.
Si A n'est pas intègre, soient $\alpha, \beta \in A$ non nuls tels que $\alpha\beta = 0$. Le produit des polynômes $X \mapsto \alpha X$ et $X \mapsto \beta$ est $(\alpha X) \cdot (\beta) = 0$; le degré du produit n'est pas la somme des degrés.

Les personnes qui ont tout compris jusqu'ici remarqueront que la notation « $X \mapsto P(X)$ » n'est pas correcte parce que du point de vue que nous adoptons ici, un polynôme n'est pas une application.

Corolaire 3.106.
Si A est intègre, les inversibles de $A[X]$ sont les éléments inversibles de A.

Démonstration. Pour que Q soit inversible, il faut un P tel que $PQ = 1$. Mais l'anneau A étant intègre, les degrés s'additionnent. Par conséquent ils doivent être de degré zéro et il faut que $P, Q \in A$. Donc Q est un inversible de A. □

3.7.4 Division euclidienne

Le théorème suivant établit la **division euclidienne** dans $A[X]$ du polynôme P par un polynôme D.

ThodivEuclPsFexf

Théorème 3.107.
Soit $D \neq 0$ dans $A[X]$ de coefficient dominant inversible dans A. Pour tout $P \in A[X]$, il existe $Q, R \in A[X]$ tels que

$$P = QD + R \tag{3.175}$$

avec $\deg(R) < \deg(D)$.

Les polynômes Q et R sont déterminés de façon univoque par cette condition.

DefMPZooMmMymG

Définition 3.108.
*Le polynôme Q est le **quotient** et R est le **reste** de la division euclidienne de P par D. Si le reste de la division de P par D est nul on dit que D **divise** P et on note $D \mid P$. Autrement dit D divise P si il existe Q tel que $P = QD$.* [37]

3.109.
Le théorème 3.107 nous incite à utiliser le degré comme stathme euclidien sur $A[X]$ dès que A est un anneau intègre. Or cela ne fonctionne pas en général, parce que très peu de polynômes ont à priori un coefficient dominant inversible.

37. Ceci se rapproche tout naturellement des notions générales de divisibilité dans un anneau intègre, vues en sous-section 3.4.1.

LEMooIDSKooQfkeKp

Lemme 3.110 (Thème 18)**.**
Si $\mathbb{K}$ est un corps[38]*, alors l'anneau $\mathbb{K}[X]$ est euclidien et principal.*

Démonstration. Puisque $\mathbb{K}$ est un corps, tous les éléments sont inversibles et le degré donne un stathme par le théorème 3.107. L'anneau $\mathbb{K}[X]$ est donc euclidien et par conséquent principal (proposition 1.254). □

Dans le théorème 6.43 nous donnerons une preuve directe du fait que $\mathbb{K}[X]$ est principal en montrant que tous ses idéaux sont principaux. Nous y démontrerons donc un peu moins pour un peu plus cher, mais avec le plaisir de ne pas devoir passer par un stathme.

DefDSFooZVbNAX

Définition 3.111 ([116])**.**
Soit un anneau A. Deux polynômes P et Q dans $A[X]$ sont dits ***étrangers*** *entre eux si* 1 *est un pgcd*[39] *de P et Q. Un ensemble de polynômes $(P_i)_{i\in I}$ est étranger* ***dans leur ensemble*** *si* 1 *est un* pgcd *des P_i.*

Les polynômes P et Q sont ***premiers entre eux*** *si les seuls diviseurs communs de P et Q sont les inversibles.*

Les notions de polynômes étrangers entre eux ou de polynômes premiers entre eux ne sont pas identiques, comme le montre l'exemple suivant.

Exemple 3.112 ([1])**.**
Soient dans $\mathbb{Z}[X]$ les polynômes $P(X) = 2X + 2$ et $Q(X) = 2X^2 + 2$. Le nombre 2 est diviseur commun et n'est pas un diviseur de 1. Donc 1 n'est pas un pgcd de P et Q. Ils ne sont pas étrangers.

Mais ils sont premiers entre eux parce qu'ils n'ont pas d'autres diviseurs communs que les inversibles (1 et -1). △

3.7.5 Polynôme primitif

DefContenuPolynome

Définition 3.113.
Un ***contenu*** *du polynôme $P = \sum_i a_i X^i \in \mathbb{K}[X]$ est un pgcd de ses coefficients : $c(P) = \text{pgcd}\left(\{a_i\}\right)$.*

NORooQITDooKsTMKm

3.114.
Quand il y a unicité du pgcd d'un ensemble, on peut parler du contenu au singulier.

Le contenu d'un polynôme sera surtout utilisé juste pour savoir si il est inversible ou non. Dans un anneau intègre, ça demande un peu d'abus de langage, mais comme le lemme 1.200 dit que si un pcgd est inversible, ils le sont tous, au moins discuter de l'inversibilité du contenu dans un anneau intègre a un sens et est bien défini.

DEFooDVOOooKaPZQC

Définition 3.115 (Polynôme primitif[117])**.**
Il existe deux notions de polynômes primitifs. Vu qu'il existe des anneaux unitaires qui sont des corps finis[40]*, il faut faire attention au contexte.*

ITEMooCNHAooVflYEW

(1) Si A est un anneau unitaire, P est un ***polynôme primitif*** *$c(P)$ est inversible*[41]*.*

(2) Si $\mathbb{K}$ est un corps fini, un polynôme dans $\mathbb{K}[X]$ est primitif si il est le polynôme minimal d'un générateur du groupe commutatif du corps.

3.116.
Pour rappel, il y a plusieurs façons de périphraser le fait que les coefficients soient premiers entre eux. Nous pouvons dire ...

38. Définition 1.209.
39. Définition 1.186.
40. Je pense à $\mathbb{Z}/n\mathbb{Z}$, mais faites attention, j'ai pas vérifié. Écrivez-moi pour me dire si c'est un bon exemple.
41. Pas mal d'auteurs disent que $c(P)$ soit être égal à 1. Mais en général ces auteurs se rétractent bien vite en disant que $c(P)$ n'est définit qu'à inversible près. Voir aussi 3.114.

(1) Le pgcd de ses coefficients est 1 parce que c'est la définition 1.259 pour avoir des nombres premiers entre eux.

(2) Le contenu de ses coefficients est 1. Parce que le contenu est précisément le pgcd, définition 3.113.

La notion de polynôme primitif au sens du pgcd est particulière aux polynômes à coefficients dans un anneau comme le montre le lemme suivant.

Lemme 3.117.
Si $\mathbb{K}$ est un corps, tout polynôme unitaire dans $\mathbb{K}[X]$ non nul est primitif au sens du pgcd.

Démonstration. Un polynôme unitaire a un 1 parmi ses coefficients, donc le pgcd est forcément 1. □

Lemme 3.118 ([118]).
Soit un anneau factoriel[42] *A. Soient des polynômes primitifs P_1 et P_2 dans $A[X]$. Soient $a_1, a_2 \in A^*$ tels que*

$$a_1P_1 = a_2P_2. \tag{3.176}$$

Alors

(1) Les éléments a_1 et a_2 sont associés dans A.

(2) Les polynômes P_1 et P_2 sont associés dans $A[X]$.

Démonstration. Vu que les P_i sont primitifs, nous avons $c(a_iP_i) = u_ia_i$ où u_i est un inversible ($u_i = c(P_i)$). Les éléments u_1a_1 et u_2a_2 sont donc deux pgcd du même ensemble : les coefficients de a_1P_1 (qui sont les mêmes que ceux de a_2P_2). Le lemme 3.60 entre en jeu pour dire que u_1a_1 et u_2a_2 sont associés : il existe un inversible v tel que $u_1a_1 = vu_2a_2$. Au final, $a_1 = u_1^{-1}vu_2a_2$ et nous voyons que a_1 et a_2 sont associés.

Nous écrivons l'égalité $a_1P1 = a_2P_2$ avec la valeur trouvée de a_1 : $u_1^{-1}vu_2a_2P_1 = a_2P_2$. On voudrait simplifier par a_2 (surtout que A est commutatif), mais comme a_2 n'est pas spécialement inversible, il faut le justifier. L'anneau A est intègre ; donc l'anneau $A[X]$ l'est aussi (théorème 3.104). Et comme $A[X]$ est intègre, on peut faire la simplification.

Au final nous avons $u_1^{-1}vu_2P_1 = P_2$, et comme $u_1^{-1}vuy$ est inversible, P_1 et P_2 sont associés. □

LEMooHJECooTeELgN

Lemme 3.119 (de Gauss[118]).
Soit un anneau factoriel A. Soient $P, Q \in A[X]$.

ITEMooISPDooXdRywE

(1) Les polynômes P et Q sont primitifs si et seulement si le polynôme PQ est primitif.

ITEMooLWEEooLcDlfp

(2) Il existe un inversible $u \in A$ tel que $c(PQ) = uc(P)c(Q)$.

Démonstration. En plusieurs parties.

(1) (1)⇒ Nous supposons que P et Q sont primitifs. Et, par l'absurde, nous supposons que PQ n'est pas primitif. Donc $c(PQ)$ n'est pas inversible dans A. Vu que A est factoriel, tout élément non inversible est un produit d'irréductibles ; c'est le cas de $c(PQ)$. Soit p un irréductible entrant dans la décomposition de $c(pQ)$.

Le lemme 3.67 dit que l'anneau $B = A/pA$ est intègre[43].

Considérons la surjection canonique $\pi\colon A \to B$ que nous prolongeons en un morphisme d'anneaux

$$\begin{aligned}\phi\colon A[X] &\to B[X]\\ \sum_i a_iX^i &\mapsto \sum_i \pi(a_i)X^i.\end{aligned} \tag{3.177}$$

42. Définition 3.64. Je rappelle qu'un anneau factoriel est toujours commutatif.

43. Ici, ma source [118] précise que p est premier. C'est vrai par la proposition 3.66, mais ça me semble inutile. Écrivez-moi pour me dire si le fait que p soit premier est important.

Notons que p ne divise pas tous les coefficients de P. En effet, sinon $c(P)$ serait un multiple de p (proposition 3.68) et ne pourrait pas être inversible. Donc $\phi(P) \neq 0$ dans $B[X]$. De même nous savons que $\phi(Q) \neq 0$.

L'anneau B est intègre parce que A est intègre (proposition 3.104). Donc $\phi(P)\phi(A) \neq 0$, et comme ϕ est un morphisme, $\phi(PQ) \neq 0$, ce qui signifie que p ne divise pas tous les coefficients de PQ, ce qui est contraire à l'hypothèse de l'absurde.

Nous en déduisons que $c(PQ)$ n'est pas inversible, et donc que PQ est primitif.

(2) **(1)⇐**

Nous supposons que PQ est primitif, et nous montrons que P et Q le sont. En factorisant un pgcd des coefficients de P nous pouvons écrire $P = c(P)P_1$ avec P_1 primitif (lemme 1.202). Par la partie précédente nous savons que P_1Q_1 est primitif. Nous avons donc l'égalité

$$PQ = c(P)c(Q)P_1Q_1 \tag{3.178}$$

où à la fois PQ et P_1P_2 sont primitifs. Vu que PQ est primitif, $c(P)c(Q)P_1Q_1$ est primitif, de telle sorte que [44]

$$c\Big(c(P)c(Q)P_1Q_1\Big) = c(P)c(Q)c(P_1Q_1) \tag{3.179}$$

est inversible. Vu que $c(P_1Q_1)$ est également inversible, nous en déduisons que $c(P)c(Q)$ est inversible. Si $u \in A$ vérifie $c(P)c(Q)u = 1$, alors $c(Q)u$ est un inverse de $c(P)$. De même $uc(P)$ est un inverse de $c(Q)$. Bref, ils sont inversibles et les polynômes P et Q sont primitifs.

(3) **Pour (2)**

Soient $P, Q \in A[X]$. Nous considérons les polynômes primitifs P_1, Q_1 et R_1 définis par $P = c(P)P_1$, $Q = c(Q)Q_1$ et $PQ = c(PQ)R_1$. Nous avons

$$c(P)c(Q)P_1Q_1 = PQ = c(PQ)R_1. \tag{3.180}$$

En prenant le contenu des deux côtés nous avons

$$c\Big(c(P)c(Q)P_1Q_1\Big) = c\Big(c(PQ)R_1\Big), \tag{3.181}$$

et en tenant compte du lemme 1.201, cela devient

$$c(P)c(Q)c(P_1Q_1) = c(PQ)c(R_1) \tag{3.182}$$

où les éléments $c(P_1Q_1)$ et $c(R_1)$ sont inversibles. En posant $u = c(P_1Q_1)^{-1}c(R_1)$, nous avons bien

$$c(P)c(Q) = uc(PQ). \tag{3.183}$$

□

3.120.

Dans toute la démonstration, il y a une ambiguïté entre $c(P)$ qui serait le pgcd des coefficients de P, c'est-à-dire un ensemble et $c(P)$ qui serait un représentant de cet ensemble.

Grâce au lemme 1.203, nous savons que si δ est un pgcd et si u est inversible, alors $u\delta$ est un pgcd. Donc à partir de l'équation

$$c(P)c(Q) = uc(PQ), \tag{3.184}$$

en fait nous savons qu'il existe des représentants de $c(P)$, $c(Q)$ et $c(PQ)$ tels que

$$c(P)c(Q) = c(PQ). \tag{3.185}$$

44. Utilisation du lemme 1.201.

3.7.6 Racines des polynômes

Définition 3.121 (Ordre d'un polynôme)**.**
Soit P un polynôme irréductible[45] *de degré n sur $\mathbb{F}_p[X]$. L'**ordre** de P est*

$$\min\{k \text{ tel que } P \mid X^k - 1\}. \tag{3.186}$$

Définition 3.122.
*Soient A un anneau et $P \in A[X]$. On appelle **racine** un élément $\alpha \in A$ tel que $P(\alpha) = 0$; c'est-à-dire que, en remplaçant toutes les occurrences de X par α dans l'expression de P, on obtient 0.*

PropHSQooASRbeA

Proposition 3.123.
Soient A un anneau et P un polynôme non nul dans $A[X]$. Si $\alpha \in A$ est une racine de P alors $X - \alpha$ divise P, et réciproquement.

Démonstration. Nous notons le polynôme $\mu = X - \alpha$ par analogie avec le polynôme minimal dont il sera question dans la très semblable proposition 6.101. Le sens réciproque est clair : si μ divise P, alors α est racine de P.

Pour le sens direct, remarquons que si α est racine de P, alors P est de degré au moins égal à 1, et nous pouvons donc effectuer la division euclidienne[46] de P par μ : il existe des polynômes Q et R tels que

$$P = Q\mu + R \tag{3.187}$$

PropHSQooASRbeA1

avec $\deg(R) < \deg(\mu)$. Donc R est une constante, élément de A : appelons-le a. En évaluant (3.187) en α, il vient

$$0 = P(\alpha) = Q(\alpha)\mu(\alpha) + a, \tag{3.188}$$

et nous en déduisons que $a = 0$, ce qui montre que $P = Q\mu$ et que μ divise P. □

DEFooTGZYooCYiKQa

Définition 3.124 (Racine simple et multiple d'un polynôme)**.**
*Soit A un anneau ainsi qu'un polynôme $P \in A[X]$ et $\alpha \in A$ racine de P. La **multiplicité** de α par rapport à P est l'entier h tel que P est divisible par $(X - \alpha)^h$ mais pas divisible par $(X - \alpha)^{h+1}$. Nous noterons $\theta_\alpha(P)$ la multiplicité de α par rapport à P.*

*Si la multiplicité d'une racine est égale à 1, nous disons que c'est une **racine simple**. Sinon, c'est une **racine multiple**.*

3.125.
Pour une définition générale d'une racine simple de l'équation $f(x) = 0$, voir la définition 34.50. La proposition 3.123 nous indique que toute racine est de multiplicité au moins égale à 1.

LEMooZUUVooQgsdXs

Lemme 3.126.
Soient un anneau A ainsi que $x, r \in A$. Nous avons

$$x^n - r^n = (x - r) \sum_{k=0}^{n-1} x^k r^{n-k-1}. \tag{3.189}$$

45. Définition 1.189.
46. Théorème 3.107.

Démonstration. Nous effectuons la distribution, et nous calculons un peu :

$$(x-r)\sum_{k=0}^{n-1}x^kr^{n-k}=\sum_{k=0}^{n-1}x^{k+1}r^{n-k-1}-\sum_{k=0}^{n-1}x^kr^{n-k} \tag{3.190a}$$
$$=\sum_{k=1}^{n}x^kr^{n-k}-\sum_{k=0}^{n-1}x^kr^{n-k} \tag{3.190b}$$
$$=-r^n+\sum_{k=1}^{n-1}(x^kr^{n-k}-x^kr^{n-k})+x^n \tag{3.190c}$$
$$=x^n-r^n. \tag{3.190d}$$

□

PROPooQCZSooVokxXQ

Proposition 3.127 ([119]).
Soient un anneau A, un polynôme $P\in A[X]$ de degré n, ainsi qu'une racine $\alpha\in A$ de P. Alors il existe un polynôme $Q\in A[X]$ de degré $n-1$ tel que $P=(X-\alpha)Q$.

Démonstration. D'après le lemme 3.126, pour chaque k, nous avons $x^k-r^k=(x-r)S_k(x)$ où S_k est un polynôme de degré $k-1$ (ou moins).

Soit un polynôme P tel que $P(\alpha)=0$. Nous pouvons écrire

$$P(x)=P(x)-P(\alpha) \tag{3.191a}$$
$$=\sum_{k=0}^{n}a_kx^k-\sum_{k=0}^{n}a_k\alpha^k \tag{3.191b}$$
$$=\sum_{k=1}^{n}a_k(x^k-\alpha^k) \tag{3.191c}$$
$$=\sum_{k=1}^{n}a_k(x-\alpha)S_k(x) \tag{3.191d}$$
$$=(x-\alpha)\sum_{k=1}^{n}a_lS_k(x) \tag{3.191e}$$
$$=(x-\alpha)Q(x) \tag{3.191f}$$

où Q est le polynômde de degré $n-1$ donné par $Q=\sum_{k=1}^{n}a_kS_k$. □

PropahQQpA

Proposition 3.128.
L'élément $\alpha\in A$ est une racine de multiplicité[47] h du polynôme P si et seulement si il existe $Q\in A[X]$ tel que $P=(X-\alpha)^hQ$ avec $Q(\alpha)\neq 0$.

Démonstration. Par définition de la multiplicité de α, le polynôme P est divisible par $(X-\alpha)^h$ mais pas par $(X-\alpha)^{h+1}$. Il existe donc un polynôme Q tel que

$$P=(X-\alpha)^hQ. \tag{3.192}$$

Si $Q(\alpha)$ était nul, la proposition 3.127 nous dirait que $Q=(X-\alpha)S$ pour un cetain polynôme S. Cela ferait $P=(X-\alpha)^{h+1}S$, ce qui est contraire à l'hypothèse sur la multiplicité. □

LemIeLhpc

Lemme 3.129.
Soient P et Q des polynômes non nuls de $A[X]$ et $\alpha\in A$. Alors

(1) $\theta_\alpha(P+Q)\leqslant\min\{\theta_\alpha(P),\theta_\alpha(Q)\}$, et l'égalité a lieu si $\theta_\alpha(P)\neq\theta_\alpha(Q)$;

ItemIeLhpciv

(2) $\theta_\alpha(PQ)\geqslant\theta_{\alpha(P)}+\theta_\alpha(Q)$, et l'égalité a lieu si A est intègre.

47. Multiplicité d'une racine, définition 3.124.

Dans le théorème suivant, la partie importante en pratique est le point (2) parce qu'il dit que, lorsque nous cherchons les racines d'un polynôme, nous pouvons nous arrêter lorsque nous en avons trouvé autant que le degré, multiplicité comprise.

ThoSVZooMpNANi

Théorème 3.130.
Soit A un anneau intègre et $P \in A[X]\backslash\{0\}$, un polynôme de degré n.

(1) Si $\alpha_1, \dots, \alpha_p \in A$ sont des racines deux à deux distinctes de multiplicités $k_1, \dots, k_p$, alors il existe $Q \in A[X]$, de degré $n - \sum_{i=1}^{p} k_i$, tel que

$$P = Q \prod_{i=1}^{p} (X - \alpha_i)^{k_i} \tag{3.193}$$

et $Q(\alpha_i) \neq 0$ pour tout i.

ITEMooWGOBooGApPOo

(2) La somme des multiplicités des racines de P est au plus $\deg(P)$.

Démonstration. Si $p = 1$, soit α une racine de multiplicité k de P. La définition de la multiplicité d'une racine nous dit que P est divisible par $(X - \alpha)^k$ mais pas par $(X - \alpha)^{k+1}$. Donc il existe $Q \in A[X]$ tel que $P = Q(X - \alpha)^k$. Il reste à voir que $Q(\alpha) \neq 0$. Cela est une conséquence de la proposition 3.123 : si $Q(\alpha)$ était nul, on pourrait lui factoriser $(X - \alpha)$ et donc avoir $(X - \alpha)^{k+1}$ qui se factorise dans P, ce qui n'est pas possible.

Nous supposons que $p \geqslant 2$ et nous effectuons une récurrence sur p. Nous considérons donc les $p-1$ premières racines $\alpha_1, \dots, \alpha_{p-1}$ et un polynôme $R \in A[X]$ tel que $R(\alpha_i) \neq 0$ pour $i = 1, \dots, p-1$ et

$$P = \underbrace{(X - \alpha_1)^{k_1} \dots (X - \alpha_{p-1})^{k_{p-1}}}_{S} R. \tag{3.194}$$

Par hypothèse $P(\alpha_p) = S(\alpha_p)R(\alpha_p) = 0$. L'anneau A étant intègre, $S(\alpha_p) \neq 0$ parce que $\alpha_i \neq \alpha_p$ pour $i \neq p$. Par conséquent, $R(\alpha_p) = 0$.

Nous devons encore vérifier que la multiplicité α_p est k_p par rapport à R. Pour cela nous utilisons le point (2) du lemme 3.129 afin de dire que le degré de α_p pour $P = SR$ est k_p. Par conséquent

$$R = (X - \alpha_p)^{k_p} T \tag{3.195}$$

avec $T(\alpha_p) \neq 0$ et enfin

$$P = \prod_{i=1}^{p} (X - \alpha_i) T. \tag{3.196}$$

De plus $T(\alpha_i) \neq 0$, sinon $R(\alpha_i)$ serait nul. □

CORooUGJGooBofWLr

Corolaire 3.131.
Un polynôme de degré n sur un anneau intègre possède au maximum n racines distinctes.

Démonstration. Le théorème 3.130(2) dit que la somme des multiplicités des racines de P est au maximum n. Mais la proposition 3.123 dit que toutes les racines ont une multiplicité au moins égale à un. Donc il ne peut pas y en avoir plus de n. □

Proposition 3.132 ([120, 121])**.**
Soit un anneau intègre et $n > 0$. Les racines n^e de l'unité forment un groupe cyclique dont l'ordre divise n.

LEMooHDJIooPTpUCJ

Lemme 3.133.
Soit le polynôme $P = X^2 - 1$ sur l'anneau intègre[48] *A. Si $1 \neq -1$*[49]*, alors les racines de P sont ± 1.*

48. Définition 1.197.

49. À mon avis cette hypothèse n'est pas nécessaire. Et d'ailleurs j'ai un peu du mal à voir des exemples exotiques d'anneaux intègres dans lesquels $1 = -1$. Il y aurait $[0, 1]$ modulo 2, certaines de ses parties, comme $\mathbb{Z}/2\mathbb{Z}$. Et quoi d'autre ?

Démonstration. Le fait que ± 1 sont racines est le lemme 1.180(1)(4). Puisque par hypothèse $1 \neq -1$, le corolaire 3.131 termine la preuve. □

CORooZCSOooHQVAOV

Corolaire 3.134 (Conséquence du lemme de Gauss[122])**.**
Soient A un anneau factoriel et Frac(A) *son corps des fractions. Un polynôme non constant $P \in A[X]$ est irréductible (sur A) si et seulement si il est irréductible et primitif au sens du pgcd*[50] *sur* Frac$(A)[X]$.

Exemple 3.135.
Il ne faudrait pas croire qu'être irréductible dans un anneau A implique d'être irréductible dans le corps des fractions. En effet soit $A = \mathbb{Z}[\sqrt{5}]$ et $P = X^2 - X - 1$. Nous savons que sa factorisation est

$$P = \left(X - \frac{1+\sqrt{5}}{2}\right)\left(X - \frac{1-\sqrt{5}}{2}\right). \tag{3.197}$$

Si vous ne le saviez pas, faites juste le calcul pour vous en assurer.

Ce polynôme est irréductible sur $\mathbb{Z}[\sqrt{5}]$ mais pas irréductible sur Frac$\big(\mathbb{Z}[\sqrt{5}]\big)$. △

3.7.7 Quelques identités

LemISPooHIKJBU

Lemme 3.136 ([1, 123])**.**
Quelques identités de polynômes.

ItemLTBooAcyMtN

(1) Si n est impair, alors $1 + X$ divise $1 + X^n$.

ItemLTBooAcyMtNii

(2) Pour tout n nous avons $X^n - 1 = (X - 1)(1 + X + \cdots + X^{n-1})$.

ITEMooILIVooHtmWLm

(3) $X^n - a^n = (X - a)\sum_{i=0}^{n-1} a^i X^{n-1-i}$.

Démonstration. Plusieurs points.

(1) **Pour (1)** Nous allons utiliser le lemme 3.1 avec $a = 1$, $b = -X$ et $r = n - 1$. Notez que n étant impair, $r = n - 1$ est encore positif. En ce qui concerne le membre de gauche de (3.1), nous remarquons que $(-X)^n = (-1)^n X^n = -X^n$. Nous avons donc :

$$1 + X^n = (a - b)\sum_{k=0}^{n-1} a^{n-1-kb^k} = (1 + X)(\cdots). \tag{3.198}$$

Cela prouve que $1 + X$ divise $1 + X^n$.

(2) **Pour (2)** Posons $P = X + \ldots + X^{n-1}$. Nous avons $XP = P - X + X^n$ et donc

$$(X - 1)(1 + P) = X + XP - 1 - P = X + (P - X + X^n) - 1 - P = X^n - 1. \tag{3.199}$$

(3) **Pour (3)** Déjà fait dans (3.1).

□

3.7.8 Générateurs pour le groupe multiplicatif

PROPooKSCRooPyInSv

Proposition 3.137 ([124])**.**
Si p est un nombre premier, tout sous-groupe de $\big((\mathbb{Z}/p\mathbb{Z})^*, \cdot\big)$ est cyclique[51].

Démonstration. Soit un sous-groupe H d'ordre $|H| = n$ de $(\mathbb{Z}/p\mathbb{Z})^*$. Nous prouvons par récurrence sur n que H est cyclique.

(1) $n = 1$ Si H est un sous-groupe d'ordre 1, alors $H = \{[1]_p\}$ et c'est cyclique, pas de problèmes.

50. Définition 3.115.
51. Définition 1.326.

(2) **Récurrence** Nous supposons que tous les sous-groupes de $\big((\mathbb{Z}/p\mathbb{Z})^*$ d'ordre plus petit que n sont cycliques, et nous prouvons que H est également cyclique.

Nous allons faire deux cas suivant que n est la puissance d'un nombre premier ou non.

(2a) **Si $n = q^k$** Nous supposons que $n = q^k$ pour un certain nombre premier q. Nous supposons aussi, par l'absurde que H n'est pas cyclique. Les éléments de H ont un ordre qui divise q^k (corolaire 2.14), mais pas d'ordre q^k. Donc pour tout x dans H nous avons $x^{q^{k-1}} = 1$. En particulier le polynôme $P = X^{X^{q-1}} - 1$ a au moins q^k racines dans $\mathbb{Z}/p\mathbb{Z}$: les éléments de H.

Hélas le degré de P étant q^{k-1}, il ne peut pas avoir plus de q^{k-1} racines distinctes par 3.130. Contradiction. Donc H est cylcique.

(2b) **Cas général**

Nous supposons que $|H|$ n'est pas la puissance d'un nombre premier.

(2b.ii) **Un morphisme** Le corolaire 3.18 nous indique que $n = ab$ pour deux nombres a, b tels que $\text{pgcd}(a, b) = 1$. Nous considérons l'application

$$\begin{aligned} f\colon H &\to H \\ x &\mapsto x^a. \end{aligned} \tag{3.200}$$

Étant donné que H est abélien, f est un morphisme de groupes. Tout élément $x \in H$ vérifie $f(x)^b = x^{ab} = 1$. Donc

$$\text{Image}(f) \subset \{y \in H \text{ tel que } y^b = 1\}. \tag{3.201}$$

(2b.ii) **Ordre du noyau et de l'image** Par le coup des racines distinctes de polynôme $X^b - 1$, nous avons $|\text{Image}(f)| \leqslant b$. De même nous avons

$$\ker(f) = \{x \in H \text{ tel que } x^a = 1\}, \tag{3.202}$$

et donc $|\ker(f)| \leqslant a$.

Le premier théorème d'isomorphisme 2.6 nous indique que

$$\frac{H}{\ker(f)} \simeq \text{Image}(f), \tag{3.203}$$

et comme tout est abélien, tous les sous groupes sont normaux[52], et nous avons, par le théorème de Lagrange 2.13 que

$$ab = |H| = \underbrace{|\ker(f)|}_{\leqslant a} \underbrace{|\text{Image}(f)|}_{\leqslant b} \leqslant ab. \tag{3.204}$$

Le premier membre est égal au dernier. Donc toutes les inégalités sont des égalités. Nous avons donc $|\ker(f)| = a$ et $|\text{Image}(f)| = b$.

(2b.ii) **Hypothèse de récurrence** C'est le moment d'utiliser l'hypothèse de récurrence : les groupes $\ker(f)$ et $\text{Image}(f)$ sont cycliques. Soient des générateurs h et h'. Quel est l'ordre de hh' ? Le lemme 3.27 nous indique que hh' est d'ordre $ab = n$. Donc hh' est générateur de H qui est donc cyclique.

□

52. « distingué » et « normal » sont synonymes.

Chapitre 4

Espaces vectoriels (début)

4.1 Parties libres, génératrices, bases et dimension

Nous avons déjà défini (dans 1.332) un espace vectoriel comme étant un module sur un corps commutatif. En explicitant un peu, cela donne ceci[125].

Un espace vectoriel sur le corps $\mathbb{K}$ est un ensemble E muni de deux opérations :

— une loi de composition interne $+ : E \times E \to E$,

— une loi de composition externe $\cdot : \mathbb{K} \times E \to E$

telles que

(1) $(E, +)$ soit un groupe abélien,

(2) pour tout $u, v \in E$ et pour tout $k, k' \in \mathbb{K}$,

(2a) $k(u + v) = (ku) + (kv)$

(2b) $(kk')u = k(k'u)$

(2c) $(k + k')u = (ku) + (k'u)$

(2d) $1u = u$

où 1 est le neutre de $\mathbb{K}$ et où nous avons directement adopté la notation ku pour $k \cdot u$.

Si $u \in E$, nous notons $-u$ l'inverse de u dans le groupe $(E, +)$.

DEFooDLHWooAvfhgc

Définition 4.1 (Partie libre)**.**
*Si E est un espace vectoriel, une partie A de E est **libre** si pour tout choix d'un nombre fini d'éléments $\{u_i\}_{i=1,\dots,n}$ de A, l'égalité*

$$a_1u_1 + \cdots + a_nu_n = 0 \tag{4.1}$$

implique $a_i = 0$ pour tout i (ici les a_i sont dans le corps de base).

Remarque 4.2.
Notons que le vecteur nul n'est dans aucune partie libre, ne fût-ce que parce que $a0 = 0$ n'implique pas $a = 0$.

Si A est une partie de l'espace vectoriel E, nous notons $\mathrm{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A. Les coefficients de ces combinaisons linéaires sont dans le corps de base $\mathbb{K}$.

Définition 4.3 (Partie génératrice)**.**
*Une partie B d'un espace vectoriel E est **génératrice** si* $\mathrm{Span}(B) = E$.

Remarque 4.4.
Ces définitions demandent des commentaires en dimension infinie [1].

1. Nous n'avons pas encore défini le concept de dimension, mais nous nous adressons au lecteur trop pressé.

(1) Tout élément peut être écrit comme combinaison linéaire finie d'une partie génératrice. Cela ne signifie pas que nous pouvons extraire une partie finie qui convient pour tous les éléments à la fois. Lorsque l'espace est de dimension infinie, ceci est particulièrement important.

(2) La définition séparée de liberté dans le cas des parties infinies a son importance lorsqu'on parle d'espaces vectoriels de dimension infinie (en dimension finie, aucune partie infinie n'est libre) parce que cela fera une différence entre une base algébrique et une base hilbertienne par exemple.

DEFooNGDSooEDAwTh

Définition 4.5 (Base)**.**
Une ***base*** *de l'espace vectoriel E est une partie à la fois génératrice et libre.*

NORMooJGRNooAKtvWt

4.6.
Supposons que la partie $\{u_1, u_2, u_2\}$ soit une base d'un espace vectoriel E. Un ensemble étant quelque chose de non ordonné, la partie $\{u_3, u_2, u_1\}$ est la même, et $\{u_1, u_1, u_2, u_2, u_3, u_4\}$ est encore la même.

Lorsque l'ordre dans lequel nous numérotons les vecteurs d'une base est important (par exemple pour écrire explicitement la matrice d'une application linéaire), nous devrions écrire (u_1, u_2, u_3). Nous pourrions définir le concept de *base ordonnée* en disant qu'une base ordonnée de E est une application $u\colon I \to E$ où I est un ensemble totalement ordonné et telle que $u(I)$ soit une base de E.

À partir de là, il ne serait plus autorisé de dire « la matrice de f dans une base ». Il faudrait dire « la matrice de f dans une base ordonnée ». Dans ce contexte, une matrice pour une application linéaire d'un espace réel serait une application $I \times I \to \mathbb{R}$.

Bref, tout ça pour dire qu'il y a clairement des incohérences de notations à ce niveau dans le Frido. Il sera souvent écrit $\{u_1, u_2\}$ au lieu de (u_1, u_2) ou, mieux, $u\colon I \to \mathbb{R}$, et il sera souvent dit « base » au lieu de « base ordonnée ».

Nous prouvons à présent que tout élément non nul d'un espace vectoriel possédant une base [2] se décompose de façon unique en combinaison linéaire finie d'éléments d'une base.

PROPooEIQIooXfWDDV

Proposition 4.7 ([1])**.**
Soient un espace vectoriel E sur $\mathbb{K}$ muni d'une base $\{e_i\}_{i\in I}$. Si $v \in E$, alors il existe un unique couple (J, c) où

(1) J est une partie finie de I ;

(2) $c\colon J \to \mathbb{K}$ est une application

(3) $v = \sum_{j\in J} c(j)e_j$.

Les éléments $c(i)$ seront notés v_i et sont les composantes de v dans la base. Sous-entendu, on prolonge c de J vers I par zéro sur $I\backslash J$.

Démonstration. Soit un espace vectoriel E et une base $\{e_i\}_{i\in I}$ où I est un ensemble à priori quelconque. Soit $v \in E$. Puisque $E = \text{Span}\{e_i\}_{i\in I}$, il existe une partie finie J de I et des coefficients $\{v_j\}_{j\in J}$ dans $\mathbb{K}$ tels que

$$v = \sum_{j\in J} v_j e_j. \tag{4.2}$$

Cela donne l'existence.

En ce qui concerne l'unicité, soient J et K des parties finies de I et des coefficients $\{v_j\}_{j\in J}$ et $\{w_k\}_{k\in K}$ tels que

EQooFKNEooICnVPP

$$v = \sum_{j\in J} v_j e_j = \sum_{k\in K} w_k e_k. \tag{4.3}$$

2. Nous n'avons pas démontré que tout espace vectoriel possède une base. Donc à notre niveau, il est possible que ce théorème soit sans objet pour beaucoup d'espaces.

Nous posons $L = J \cup K$. Remarquez que les unions suivantes sont des unions disjointes :

$$L = (J\backslash K) \cup (J \cap K) \cup (K\backslash J) \tag{4.4a}$$
$$J = (J\backslash K) \cup (J \cap K) \tag{4.4b}$$
$$K = (K \cap J) \cup (K\backslash J). \tag{4.4c}$$

Écrivons $0 = v - v$ en utilisant les expressions de v de (4.3) et en décomposant les sommes :

$$0 = \sum_{j\in J} v_j e_j - \sum_{k\in K} w_k e_k \tag{4.5a}$$
$$= \sum_{j\in J\backslash K} v_j e_j + \sum_{j\in J\cap K} v_j e_j - \sum_{k\in K\cap J} w_k e_k - \sum_{k\in K\backslash J} w_k e_k \tag{4.5b}$$
$$= \sum_{j\in J\backslash K} v_j e_j + \sum_{j\in J\cap K} (v_j - w_j) e_j - \sum_{k\in K\backslash J} w_k e_k. \tag{4.5c}$$

Là-dessus, nous posons

$$\alpha_l = \begin{cases} v_l & \text{si } l \in J\backslash K \\ v_l - w_l & \text{si } l \in J \cap K \\ -w_l & \text{si } l \in K\backslash J. \end{cases} \tag{4.6}$$

Nous avons alors

$$\sum_{l\in L} \alpha_l e_l = 0. \tag{4.7}$$

EQooCTWRooZKyaup

Vu que $\{e_i\}_{i\in I}$ est libre, la partie $\{e_l\}_{l\in L}$ est également libre. Donc l'équation (4.7) implique que $\alpha_l = 0$ pour tout $l \in L$.

Nous devons prouver que

$$\{j \in J \text{ tel que } v_j \neq 0\} = \{k \in K \text{ tel que } w_k \neq 0\} \tag{4.8}$$

et que pour l dans cet ensemble, $v_l = w_l$.

Soit $j \in J$; il y a deux possibilités : soit $j \in J\backslash K$, soit $j \in J \cap K$. Dans le premier cas nous avons déjà vu que $\alpha_j = v_j = 0$. Dans le second cas, $\alpha_j = v_j - w_j = 0$, c'est-à-dire $v_j = w_j$.

Donc $j \in J$ vérifiant $v_j \neq 0$ implique $j \in J \cap K$ et l'égalité des coefficients. Idem avec $k \in K$ tel que $w_k \neq 0$ implique $k \in J \cap K$. □

LEMooDJSIooYcsvhO

Lemme 4.8 ([1])**.**
Soit un espace vectoriel admettant des bases. Un endomorphisme est une bijection si et seulement si il change toute base en une base.

Démonstration. En deux parties. Soit un espace vectoriel E possédant des bases et un endomorphisme $f : E \to E$.

(1) **Si f est bijective** Soit une base $\{v_i\}_{i\in I}$; nous devons voir que $\{f(v_i)\}_{i\in I}$ est une base.

(1a) **Libre** Si J est une partie finie de I et si les λ_j sont des scalaires tels que $\sum_{j\in J} \lambda_j f(v_j) = 0$, alors

$$0 = \sum_{j\in J} \lambda_j f(v_j) = f\big(\sum_{j\in J} \lambda_j v_j\big). \tag{4.9}$$

Mais comme f est bijective, cela implique que $\sum_{j\in J} \lambda_j v_j = 0$. En retour, parce que $\{v_i\}$ est une base, cela implique que $\lambda_j = 0$ pour tout j.

(1b) **Générateur** Soit $x \in E$. Puisque f est bijective, il existe un unique $y \in E$ tel que $x = f(y)$. Comme $\{v_i\}_{i\in I}$ est une base, il existe une partie finie $J \subset I$ et des scalaires $\{\lambda_j\}_{j\in J}$ tels que

$$y = \sum_{j\in J} \lambda_j v_j. \tag{4.10}$$

Nous avons alors

$$x = f(y) = \sum_{j\in J} \lambda_j f(v_j), \tag{4.11}$$

qui montre que $\{f(v_i)\}_{i\in I}$ est bien génératrice de E

(2) **Si f change les bases en bases** Soit un endomorphisme changeant toute base en une base. Nous devons prouver qu'il est bijectif.

(2a) **Injective** Nous considérons une base $\{v_i\}_{i\in I}$. La partie $\{f(v_i)\}_{i\in I}$ est par hypothèse également une base.

Soient $x, y \in E$ tels que $f(x) = f(y)$. Il existe J et K finis dans I qui permettent de décomposer x et y respectivement dans la base $\{f(v_i)\}_{i\in I}$. Quitte à poser $J' = J \cup K$, nous supposons que J suffit [3]. Il existe donc des scalaires $\{\lambda_j\}_{j\in J}$ et $\{\mu_j\}_{j\in J}$ tels que $x = \sum_{j\in J} \lambda_j f(v_j)$ et $y = \sum_{j\in J} \mu_j f(v_j)$.

La relation $f(x) = f(y)$ donne immédiatement, par la linéarité de f,

$$\sum_{j\in J} (\lambda_j - \mu_j) f(v_j) = 0. \tag{4.12}$$

Du fait que $\{f(v_i)\}_{i\in I}$ soit une base, nous déduisons que $\lambda_j - \mu_j = 0$ pour tout j. Donc $x = y$, et f est injective.

(2b) **Surjective** Soit $x \in E$. Puisque $\{f(v_i)\}_{i\in I}$ est une base, il existe des scalaires λ_j tels que

$$x = \sum_{j\in J} \lambda_j f(v_j) = f\Big(\sum_{j\in J} \lambda_j v_j\Big). \tag{4.13}$$

Donc f est surjective.

□

PROPooVJFNooSbkzTF

Proposition 4.9 ([1]).
Soit un espace vectoriel E. Soit une partie $B \subset E$ telle que tout élément de E puisse être écrit de façon unique comme combinaison linéaire d'un nombre fini d'éléments de B. Alors B est une base.

LEMooRWQHooIxrQek

Lemme 4.10 ([1]).
Bases dans $\mathbb{R}^n$.

ITEMooKWULooCTmOqM

(1) Si v et w ne sont pas colinéaire dans $\mathbb{R}^2$, alors $\{v, w\}$ est une base de $\mathbb{R}^2$.

(2) Toute partie libre de $\mathbb{R}^n$ contenant n élément est une base.

Définition 4.11.
*Un espace vectoriel est **de type fini** si il contient une partie génératrice finie.*

Nous verrons dans les résultats qui suivent que cette définition est en réalité inutile parce qu'un espace vectoriel sera de type fini si et seulement si il est de dimension finie.

LemytHnlD

Lemme 4.12.
Si E a une famille génératrice de cardinal n, alors toute famille de $n+1$ éléments est liée.

Démonstration. Nous procédons par récurrence sur n. Pour $n = 1$, nous avons $E = \text{Span}(e)$ et donc si $v_1, v_2 \in E$ nous avons $v_1 = \lambda_1 e$, $v_2 = \lambda_2 e$ pour certains éléments non nuls λ_1, λ_2 du corps de base. Nous avons donc $\lambda_2 v_1 - \lambda_1 v_2 = 0$. Cela prouve que $\{v_1, v_2\}$ est liée.

Supposons maintenant que le résultat soit vrai pour $k < n$, c'est-à-dire que pour tout espace vectoriel contenant une partie génératrice de cardinal $k < n$, les parties de $k+1$ éléments sont liées. Soit maintenant un espace vectoriel muni d'une partie génératrice $G = \{e_1, \dots, e_n\}$ de n éléments, et montrons que toute partie $V = \{v_1, \dots, v_{n+1}\}$ contenant $n+1$ éléments est liée. Dans nos notations nous supposons que les e_i sont des vecteurs distincts et les v_i également. Nous les

3. Nous utilisons le fait que l'union de deux parties finies d'un ensemble est finie (lemme 1.128(2)).

supposons également tous non nuls. Étant donné que $\{e_i\}$ est génératrice nous pouvons définir les nombres $\lambda_{i,k}$ par

$$v_i = \sum_{k=1}^{n} \lambda_{i,k} e_k \tag{4.14}$$

Puisque

$$v_{n+1} = \sum_{k=1}^{n} \lambda_{n+1,k} e_k \neq 0, \tag{4.15}$$

quitte à changer la numérotation des e_i, nous pouvons supposer que $\lambda_{n+1,n} \neq 0$. Nous considérons les vecteurs

$$w_i = \lambda_{n+1,n} v_i - \lambda_{i,n} v_{n+1}. \tag{4.16}$$

En calculant un peu,

$$w_i = \lambda_{n+1,n} \sum_{k=1}^{n} \lambda_{i,k} e_k - \lambda_{i,n} \sum_{k=1}^{n} \lambda_{n+1,k} e_k \tag{4.17a}$$

$$= \sum_{k=1}^{n-1} \big(\lambda_{n+1,n}\lambda_{i,k} - \lambda_{i,n}\lambda_{n+1,k}\big) e_k \tag{4.17b}$$

parce que les termes en e_n se sont simplifiés. Donc la famille $\{w_1, \dots, w_n\}$ est une famille de n vecteurs dans l'espace vectoriel $\mathrm{Span}\{e_1, \dots, e_{n-1}\}$; elle est donc liée par l'hypothèse de récurrence. Il existe donc des nombres $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tous nuls, tels que

$$0 = \sum_{i=1}^{n} \alpha_i w_i = \sum_{i=1}^{n} \alpha_i \lambda_{n+1,n} v_i - \sum_{i=1}^{n} \alpha_i \lambda_{i,n} v_{n+1}. \tag{4.18}$$ EqOQGGoU

Vu que $\lambda_{n+1,n} \neq 0$ et que, parmi les α_i, au moins un est non nul, nous avons au moins un des produits $\alpha_i \lambda_{n+1,n}$ qui est non nul. Par conséquent (4.18) est une combinaison linéaire nulle non triviale des vecteurs de $\{v_1, \dots, v_{n+1}\}$. Cette partie est donc liée. □ LemkUfzHl

Lemme 4.13.
Soient L une partie libre et G une partie génératrice d'un espace vectoriel E. Si l'ensemble des parties libres L' telles que $L \subset L' \subset G$ possède un élément maximum[4]*, alors cet élément est une base.*

Qu'entend-on par « maximale » ? La partie B candidate, doit être libre, contenir L, être contenue dans G et de plus avoir la propriété que $\forall x \in G \backslash B$, la partie $B \cup \{x\}$ est liée.

Démonstration. D'abord si G est une base, alors toutes les parties de G sont libres et le maximum est $B = G$. Dans ce cas le résultat est évident. Nous supposons donc que G est liée.

La partie $B = \{b_1, \dots, b_l\}$ est libre parce qu'on l'a prise parmi les libres. Montrons que B est génératrice. Soit $x \in G \backslash B$; par hypothèse de maximalité, $B \cup \{x\}$ est liée, c'est-à-dire qu'il existe des nombres λ_i, λ_x non tous nuls tels que

$$\sum_{i=1}^{l} \lambda_i b_i + \lambda_x x = 0. \tag{4.19}$$ EqxfkevM

Si $\lambda_x = 0$ alors un de λ_i doit être non nul et l'équation (4.19) devient une combinaison linéaire nulle non triviale des b_i, ce qui est impossible parce que B est libre. Donc $\lambda_x \neq 0$ et

$$x = -\frac{1}{\lambda_x} \sum_{i=1}^{l} \lambda_i b_i. \tag{4.20}$$

Donc tous les éléments de $G \backslash B$ sont des combinaisons linéaires des éléments de B, et par conséquent, G étant génératrice, tous les éléments de E sont combinaisons linéaires d'éléments de B. □

4. Encore une fois, à part quelques cas triviaux, il n'est pas clair à ce point que ce maximum existe.

ThonmnWKs

Théorème 4.14 (Théorème de la base incomplète)**.**
Soit E un espace vectoriel de type fini sur le corps $\mathbb{K}$.

ItemBazxTZ

(1) Si L est une partie libre et si G est une partie génératrice contenant L, alors il existe une base B telle que $L \subset B \subset G$.

ITEMooFVJXooGzzpOu

(2) Toute partie libre peut être étendue en une base.

ITEMooFBUAooSSZxgx

(3) Toutes les bases sont finies et ont même cardinal.

ITEMooJIJSooGuJMdt

(4) Si V est un sous-espace vectoriel de E, et si L est une base de V, alors il existe une base de E qui contient L.

Démonstration. Point par point.

(1) Comme E est de type fini, il admet une partie génératrice G de cardinal fini n. Donc une partie libre est de cardinal au plus n par le lemme 4.12. Soit L, une partie libre contenue dans G (ça existe : par exemple $L = \varnothing$). La partie B maximalement libre contenue dans G et contenant L est une base par le lemme 4.13.

(2) Notons que puisque E lui-même est générateur, le point (1) implique que toute partie libre peut être étendue en une base.

(3) Soient B et B', deux bases. En particulier B est génératrice et B' est libre, donc le lemme 4.12 indique que $\mathrm{Card}(B') \leqslant \mathrm{Card}(B)$. Par symétrie on a l'inégalité inverse. Donc $\mathrm{Card}(B) = \mathrm{Card}(B')$.

(4) La partie L étant une base de V, elle est en particulier libre dans E. Par le point (2), L peut être étendue en une base.

□

REMooYGJEooEcZQKa

Remarque 4.15.
Le théorème de la base incomplète 4.14(2) est ce qui permet de construire une base d'un espace vectoriel en « commençant par » une base d'un sous-espace. En effet si H est un sous-espace de E, alors une base de H est une partie libre de E et donc, peut être étendue en une base de E.

DEFooWRLKooArTpgh

Définition 4.16.
*La **dimension** d'un espace vectoriel de type fini est le cardinal*[5] *d'une*[6] *de ses bases.*

Il existe une infinité de bases de $\mathbb{R}^m$. On peut démontrer que le cardinal de toute base de $\mathbb{R}^m$ est m, c'est-à-dire que toute base de $\mathbb{R}^m$ possède exactement m éléments.

Exemple 4.17.
La base de **canonique** de $\mathbb{R}^m$ est la partie $\{e_1, \dots, e_m\}$, où le vecteur e_j est

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{j-ème} \quad .$$

La composante numéro j de e_i est 1 si $i = j$ et 0 si $i \neq j$. Cela s'écrit $(e_i)_j = \delta_{i,j}$ où δ est le **symbole de Kronecker** défini par

$$\delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{4.21}$$

5. Définition 1.126.
6. Le théorème de la base incomplète 4.14(3) montre que cette définition ne souffre d'aucune ambiguïté.

Les éléments de la base canonique de $\mathbb{R}^m$ peuvent donc être écrits $e_i = \sum_{k=1}^{m} \delta_{i,k} e_k$. △

Le théorème suivant est essentiellement une reformulation du théorème 4.14.

ThoMGQZooIgrXjy

Théorème 4.18.
Soit E un espace vectoriel de dimension finie et $\{e_i\}_{i\in I}$ une partie génératrice de E.

ITEMooTZUDooFEgymQ

(1) Il existe $J \subset I$ tel que $\{e_i\}_{i\in J}$ est une base. Autrement dit : de toute partie génératrice nous pouvons extraire une base.

ITEMooCJQGooXwjsfm

(2) Soit $\{f_1, \dots, f_l\}$ une partie libre. Alors nous pouvons la compléter en utilisant des éléments e_i. C'est-à-dire qu'il existe $J \subset I$ tel que $\{f_k\} \cup \{e_i\}_{i\in J}$ soit une base.

Démonstration. En deux parties.

(1) **Pour (1)** C'est le théorème 4.14(1). La partie $\{e_{i_0}\}$ est libre et la partie $\{e_i\}_{i\in I}$ est génératrice. Donc il existe une base B telle que $\{e_{i_0}\} \subset B \subset \{e_i\}_{i\in I}$. Si $B = \{e_i\}_{i\in J}$, alors J est fini[7] parce que toutes les bases d'un espace vectoriel de dimension finie ont le même nombre d'éléments.

(2) **Pour (2)** Utiliser le théorème 4.14(4) avec $V = \text{Span}\{f_1, \dots, f_l\}$.

□

PROPooVEVCooHkrldw

Proposition 4.19.
Si E est un espace vectoriel de dimension finie n, alors

ITEMooZNLDooBISkJyBS

(1) toute partie contenant $n+1$ éléments est liée.

ITEMooSGGCooOUsuBs

(2) toute partie libre contenant n éléments est une base,

(3) toute partie génératrice contenant n éléments est une base.

Démonstration. Soit une partie M contenant $n+1$ éléments. L'espace E possède une partie génératrice contenant n éléments (n'importe quelle base). Donc M est liée par le lemme 4.12.

Une partie libre contenant n éléments peut être étendue en une base ; si ladite extension est non triviale (c'est-à-dire qu'on ajoute vraiment au moins un élément) une telle base contiendra une partie de $n+1$ éléments qui serait liée par le lemme 4.12.

Pour la dernière assertion, soit une partie génératrice $\{v_i\}_{i\in I}$ où I contient n éléments. Par le théorème 4.18(1) nous pouvons en extraire une base : il existe $J \subset I$ tel que $\{v_j\}_{j\in J}$ soit une base. Si l'inclusion $J \subset I$ était stricte, alors la base $\{v_j\}_{j\in J}$ contiendrait moins de n éléments, ce qui serait en contradiction avec le théorème 4.14(3). □

DefCodimension

Définition 4.20.
*Soit F un sous-espace vectoriel de l'espace vectoriel E. La **codimension** de F dans E est*

$$\text{codim}_E(F) = \dim(E/F). \tag{4.22}$$

LEMooGNDPooAhvIek

Lemme 4.21.
L'ensemble $\mathbb{Q}^n$ est un $\mathbb{Q}$-espace vectoriel de dimension n.

4.1.1 Et en dimension infinie

Dans ZFC, en dimension infinie, il existe aussi une base pour tout espace vectoriel ainsi qu'un théorème de la base incomplète. Nous ne parlerons pas de ce qu'il se passe lorsque nous ne considérons que ZF[8].

7. Sauf si I est choisi de façon traitresse. Il se peut que $e_{i_1} = e_{i_2}$ avec $i_1 \neq i_2$.

8. Si vous ne savez pas ce que signifient les sigles « ZF » et « ZFC » vous ne devriez pas être en train de lire ceci, et encore moins penser à le resservir à un jury d'agrégation.

LEMooSSRXooIyfgNz

Lemme 4.22 ([126]).
Soient un $\mathbb{K}$-espace vectoriel E et un sous-espace vectoriel V de E. Soient encore deux sous-espaces vectoriels W_1 et W_2 tels que

(1) $V \cap W_1 = \{0\}$;

(2) $V + W_2 = E$.

Alors il existe un supplémentaire W de V tel que $W_1 \subset W \subset W_2$.

Juste une remarque : dans le Frido le symbole « $\subset$ » ne signifie pas une inclusion stricte.

Démonstration. Nous utilisons le lemme de Zorn.

(1) **Un gros ensemble** Soit

$$\mathcal{A} = \left\{ S \subset E \text{ tel que } \begin{cases} S \text{ est un sous-espace vectoriel de } E \\ W_1 \subset S \subset W_2 \\ S \cap V = \{0\} \end{cases} \right\} \tag{4.23}$$

(2) **Non vide** Puisque $W_1 \in \mathcal{A}$, cet ensemble n'est pas vide.

(3) **Ordre** L'ensemble $\mathcal{A}$ est partiellement ordonné pour l'inclusion.

(4) **$\mathcal{A}$ est inductif** Nous prouvons maintenant que $\mathcal{A}$ est inductif[9]. Pour cela, soit une partie $\mathcal{A}'$ totalement ordonnée et $U = \bigcup_{A \in \mathcal{A}'} A$.
Alors, la partie U est un sous-espace vectoriel de E. En effet si $x, y \in U$, alors il existe $A_1, A_2 \in \mathcal{A}'$ tels que $x \in A_1$ et $y \in A_2$. Comme $\mathcal{A}'$ est totalement ordonné, l'un des ensembles parmi A_1 et A_2 est inclus dans l'autre. Sans perdre de généralité, disons $A_1 \subset A_2$. Alors les opérations s'effectuent dans A_2 : nous avons $x, y \in A_2$, et donc $\lambda x \in A_2 \subset U$ ainsi que $x + y \in A_2 \subset U$.
De plus, U contient W_1, et est contenu dans W_2. Ainsi, $U \in \mathcal{A}$ et majore $\mathcal{A}'$ pour l'inclusion. En bref, $\mathcal{A}$ est bien inductif.

(5) **Utilisation de Zorn** Le lemme de Zorn 1.23 nous donne alors un élément maximal W de $\mathcal{A}$. Cet élément vérifie

(5a) $W \cap V = \{0\}$,

(5b) $W_1 \subset W \subset W_2$,

(5c) pour tout $W' \in \mathcal{A}$, nous avons $W' \subset W$ par maximalité de W.

(6) **Supplémentaire** Montrons que ce W est un supplémentaire de V. Soit $x \in E$. Le but est de trouver une décomposition de x en somme d'un élément de W et un de V. Comme $V + W_2 = E$, nous avons $v \in V$ et $w_2 \in W_2$ tels que

$$x = v + w_2. \tag{4.24}$$

Si $w_2 \in W$ alors c'est fini. Sinon ...
Soit $X = \text{Span}\{W, w_2\}$. Vu que X contient strictement W et que W est maximum dans $\mathcal{A}$, la partie X n'est pas un élément de $\mathcal{A}$. Comme X est un sous-espace vectoriel de E tel que $W_1 \subset X \subset W_2$, la seule possibilité pour que X ne soit pas dans $\mathcal{A}$ est que $X \cap V \neq \{0\}$. Soit donc $y \neq 0$ dans $X \cap V$. Par définition de X,

$$y = w' + \lambda w_2 \tag{4.25}$$

EqDecompo55:296

pour $w' \in W$, $w_2 \in W_2$ et $\lambda \in \mathbb{K}$. Nous avons $\lambda \neq 0$, sinon nous aurions $y \in W \cap V$ et donc $y = 0$ puisque W est dans $\mathcal{A}$. La décomposition (4.25) permet alors d'écrire $w_2 = (y - w')/\lambda$ et finalement

$$x = v + \frac{1}{\lambda}(y - w') = \underbrace{v + \frac{1}{\lambda}y}_{\in V} - \underbrace{\frac{1}{\lambda}w'}_{\in W}. \tag{4.26}$$

La somme d'espaces vectoriels $E = V + W$ est donc établie.

9. Définition 1.22.

□

Corolaire 4.23.
Tout sous-espace vectoriel d'un espace vectoriel possède un supplémentaire.

Démonstration. Soit un espace vectoriel E ainsi qu'un sous-espace vectoriel V. Si $V = E$ nous sommes O.K. Sinon nous considérons $v \in E \backslash V$ et nous posons $W_1 = \mathbb{K}v$ et $W_2 = E$.

Vu que V et W_1 sont des espaces vectoriels, nous avons $V \cap W_1 = \{0\}$, et puisque $W_2 = E$, nous avons $V + W_2 = E$. Le lemme 4.22 nous donne alors un supplémentaire de V. □

PROPooHDCEooMhDjPi

Proposition 4.24 (Base incomplète)**.**
Tout espace vectoriel (non réduit à $\{0\}$) possède une base.

Démonstration. Soit $\mathcal{A}$ l'ensemble des familles libres de E. Il n'est pas vide parce que $\{v\}$ en est une dès que v est non nul dans E. Rapidement :

— l'ensemble $\mathcal{A}$ est ordonné pour l'inclusion,

— si $\mathcal{A}'$ est une partie totalement ordonnée, l'union est un majorant,

— donc $\mathcal{A}$ est inductif,

— soit un maximum F de $\mathcal{A}$.

La partie F est libre parce qu'elle est dans $\mathcal{A}$. Elle est génératrice parce que si v n'est pas dans $\mathrm{Span}(F)$ alors la partie $F \cup \{v\}$ est encore libre, et majore strictement F pour l'inclusion, ce qui n'est pas possible.

Donc F est une base de E. □

THOooOQLQooHqEeDK

Théorème 4.25 (Base incomplète, dimension quelconque)**.**
Soit une partie $\{e_i\}_{i\in I}$ génératrice de l'espace vectoriel E (ici, I est un ensemble quelconque[10]). Soit $I_0 \subset I$ tel que $\{e_i\}_{i\in I_0}$ soit libre.

Alors il existe I_1 tel que $I_0 \subset I_1 \subset I$ tel que $\{e_i\}_{i\in I_1}$ soit une base de E.

Note : une telle partie I_0 existe en prenant un singleton. Mais l'existence n'est pas le sujet ici.

Démonstration. Soit $\mathcal{A}$ l'ensemble des parties J de I telles que $I_0 \subset J \subset I$ et telles que $\{e_i\}_{i\in J}$ soit libre.

Encore une fois, $\mathcal{A}$ est inductif pour l'ordre partiel donné par l'inclusion. Soit J un élément maximum par le lemme de Zorn 1.23. Puisque $J \in \mathcal{A}$, la partie $\{e_i\}_{i\in J}$ est libre. Mais elle est également génératrice parce que si e_k n'est pas dedans, J ne serait pas maximum, étant majorée par $J \cup \{k\}$.

Donc $\{e_i\}_{i\in J}$ engendre tous les e_i avec $i \in I$ et donc, tous les éléments de E. □

NORMooREVQooEFJWta

4.26 ([13])**.**
Une preuve alternative du théorème de la base incomplète serait de prouver que l'ensemble des parties libres est inductif. De ce fait, la proposition 1.24 permet de dire que toute partie libre peut être complétée en une base.

4.1.2 Espace librement engendré

DEFooCPNIooNxsYMY

Définition 4.27 ([127])**.**
*Soient un ensemble S et un corps $\mathbb{K}$. L'espace vectoriel **librement engendré** sur S, noté $F_{\mathbb{K}}(S)$ est l'ensemble des applications $S \to \mathbb{K}$ qui sont non-nulles en un nombre fini de points de S.*

Autrement dit, $\sigma : S \to \mathbb{K}$ est dans $F_{\mathbb{K}}(S)$ si $\{x \in S \text{ tel que } \sigma(x) \neq 0\}$ est fini[11].

Le lemme suivant donne tout son sens à l'expression « librement » engendré. Il dit que $F_{\mathbb{K}}(S)$ possède une base indexée par S lui-même.

10. Un cas d'utilisation intéressant est de poser $I = E$ et $e_i = i$. Pensez-y.
11. Parce que nous l'aimons bien, nous ne résistons pas à faire un renvoi vers la définition 1.117.

LEMooLOPAooUNQVku

Lemme 4.28.
L'ensemble des applications δ_s données par

$$\begin{aligned}\delta_s\colon S &\to \mathbb{K}\\ t &\mapsto \begin{cases}1 & si\ t=s\\ 0 & sinon\end{cases}\end{aligned} \tag{4.27}$$

avec $s\in S$ forment une base[12] *de $F_{\mathbb{K}}(S)$.*

Démonstration. Pour prouver que les δ_s sont générateurs, nous considérons $g\colon S\to\mathbb{K}$ non nul sur la partie finie $\{s_i\}_{i\in I}$ de S. Alors nous avons

$$g=\sum_{i\in I} g(s_i)\delta_{s_i}. \tag{4.28}$$

Pour prouver que les δ_s forment une partie libre, nous supposons avoir $\lambda_i\in\mathbb{K}$ tels que

$$g=\sum_{i\in I}\lambda_i\delta_{s_i}=0 \tag{4.29}$$

Soit $j\in I$. Nous avons

$$0=f(s_j)=\sum_{i\in I}\lambda_i\underbrace{\delta_{s_i}(s_j)}_{=\delta_{i,j}}=\lambda_j. \tag{4.30}$$

Donc les coefficients λ_i sont tous nuls, et nous avons prouvé que la partie est libre. □

Il est parfois pratique d'écrire les éléments de $F_{\mathbb{K}}(S)$ comme sommes « formelles » d'éléments de S. Cela va encore lorsque S est un ensemble n'ayant aucune somme bien définie.

Mais attention : si $S=\mathbb{R}$, l'élément $4+7$ de $F_{\mathbb{R}}(\mathbb{R})$ n'est pas 11. L'élément 11 de $F_{\mathbb{R}}(\mathbb{R})$ est un élément complètement différent. Bref, il n'est pas judicieux d'écrire les éléments de $F_{\mathbb{K}}(S)$ comme des combinaisons linéaires d'éléments de S. Pour $x\in S$ il vaut mieux écrire explicitement δ_x que x. La somme $\delta_x+\delta_y$ est parfaitement bien définie dans l'ensemble des applications de S vers $\mathbb{K}$.

4.2 Applications linéaires

Le lien entre matrice et applications linéaires est la définition 4.68 et toutes les propriétés qui s'en suivent.

4.2.1 Définition

DEFooULVAooXJuRmr

Définition 4.29.
*Soient des espaces vectoriels E et F sur le corps $\mathbb{K}$. Soit un sous-corps $\mathbb{L}$ de $\mathbb{K}$. Une application $T\colon E\to F$ est dite $\mathbb{L}$-**linéaire** si*

— *$T(x+y)=T(x)+T(y)$ pour tout x et y dans E,*

— *$T(\lambda x)=\lambda T(x)$ pour tout λ dans $\mathbb{K}$ et x dans E.*

Nous noterons $\mathcal{L}_{\mathbb{L}}(E,F)$ l'espace des applications $\mathbb{L}$-linéaires de E vers F.

Si vous avez bien suivi, les égalités dans la définition 4.29 sont des égalités dans F.

4.30.
Le plus souvent, si E est un espace vectoriel sur $\mathbb{K}$, alors nous ne considérerons que les applications $\mathbb{K}$-linéaires. Autrement dit, nous écrirons le plus souvent simplement $\mathcal{L}(E,F)$ sans préciser le corps.

Il pourra pourtant arriver que, pour un espace vectoriel sur $\mathbb{C}$, nous considérions les applications seulement $\mathbb{R}$-linéaires. Ce sera le cas dans le lemme 27.161.

12. Définition 4.5.

DefDQRooVGbzSm

Lemme-Définition 4.31.
L'ensemble de toutes les applications linéaires de E vers F est noté $\mathcal{L}(E,F)$ et devient un espace vectoriel sur $\mathbb{K}$ avec les définitions suivantes :

(1) $(T_1+T_2)(x)=T_1(x)+T_2(x)$,

(2) $(\lambda T)(x)=\lambda T(x)$.

EXooMAWMooEaNWpl

Exemple 4.32.
Pour tout b dans $\mathbb{R}$ la fonction $T_b(x)=bx$ est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$. En effet,

— $T_b(x+y)=b(x+y)=bx+by=T_b(x)+T_b(y)$,

— $T_b(ax)=b(ax)=abx=aT_b(x)$.

De la même façon on peut montrer que la fonction T_λ définie par $T_\lambda(x)=\lambda x$ est une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^m$ pour tout λ dans $\mathbb{R}$ et m dans $\mathbb{N}$. △

ex_affine

Exemple 4.33.
Soit $m\in\mathbb{N}$. On fixe λ dans $\mathbb{R}$ et v dans $\mathbb{R}^m$. L'application U_λ de $\mathbb{R}^m$ dans $\mathbb{R}^m$ définie par $U_\lambda(x)=\lambda x+v$ n'est pas une application linéaire lorsque $v\neq 0$, parce que si a est un réel différent de 0 et 1, alors $av\neq v$, d'où

$$U_\lambda(ax)=\lambda(ax)+v\neq a(\lambda x+v)=aU_\lambda(x).$$

△

exampleT_A

Exemple 4.34.
Soit A une matrice fixée de $\mathbb{M}(m\times n,\mathbb{R})$. La fonction $T_A\colon\mathbb{R}^m\to\mathbb{R}^n$ définie par $T_A(x)=Ax$ est une application linéaire. En effet,

— $T_A(x+y)=A(x+y)=Ax+Ay=T_A(x)+T_A(y)$,

— $T_A(ax)=A(ax)=a(Ax)=aT_A(x)$.

△

LEMooLGEHooVEEoiU

Lemme 4.35.
Si une application linéaire est inversible, alors son inverse est linéaire.

Démonstration. Soient une application linéaire inversible $f\colon E\to F$, ainsi que $x,y\in E$. Nous avons

$$f\big(f^{-1}(x)+f^{-1}(y)\big)=f\big(f^{-1}(x)\big)+f\big(f^{-1}(y)\big)=x+y=f\big(f^{-1}(x+y)\big).\tag{4.31}$$

En prenant f^{-1} des deux côtés, nous trouvons

$$f^{-1}(x)+f^{-1}(y)=f^{-1}(x+y).\tag{4.32}$$

De même :

$$f\big(\lambda f^{-1}(x)\big)=\lambda f\big(f^{-1}(x)\big)=\lambda x=f\big(f^{-1}(\lambda x)\big),\tag{4.33}$$

ce qui nous donne $\lambda f^{-1}(x)=f^{-1}(\lambda x)$. □

DEFooOAOGooKuJSup

Définition 4.36 (Quelques ensembles d'applications linéaires)**.**
Soient E et F des espaces vectoriels.

— *L'ensemble des applications linéaires de E vers F est noté $\mathcal{L}(E,F)$, comme déjà dit en 4.31.*

— *Une application linéaire $E\to E$ est un* ***endomorphisme*** *de E. L'ensemble des endomorphismes de E est noté* $\mathrm{End}(E)$.

— *Un endomorphisme bijectif est un* ***automorphisme****. L'ensemble des automorphismes de E est noté* $\mathrm{Aut}(E)$.

— *Une application linéaire bijective $E \to F$ est un **isomorphisme** d'espace vectoriel. L'ensemble des isomorphismes de $E \to F$ est noté*[13] $\mathrm{GL}(E,F)$. *C'est un groupe par le lemme 4.35.*

Remarque 4.37.
Les ensembles définis en 4.36 concernent la structure d'espace vectoriel seulement. Lorsque nous verrons la notion d'espace vectoriel normé, nous demanderons de plus, la continuité, laquelle n'est pas automatique en dimension infinie. Voir aussi les définitions 10.272.

Définition 4.38.
*Si E est un espace vectoriel, si X est un espace vectoriel, et si $f\colon X \to E$ est une application, le **noyau** de f est le noyau de f lorsque E est vu comme un groupe pour l'addition*[14]*, c'est-à-dire la partie*

$$\ker(f) = \{x \in X \text{ tel que } f(x) = 0\}. \tag{4.34}$$

PROPooRLLPooKYzsJp

Proposition 4.39.
Le noyau d'une application linéaire est un sous-espace vectoriel.

Démonstration. Soit une application linéaire $f\colon E \to F$. Si $x, y \in \ker(f)$ et si $\lambda \in \mathbb{K}$ alors

$$f(x+y) = f(x) + f(y) = 0 + 0 = 0, \tag{4.35}$$

donc $x + y \in \ker(f)$, et

$$f(\lambda x) = \lambda f(x) = 0, \tag{4.36}$$

donc $\lambda x \in \ker(f)$. □

LEMooIQJMooJWRbub

Lemme 4.40 ([1])**.**
Soit une application linéaire $f\colon V \to W$. Si $b \in W$ et si $m \in f^{-1}(b)$, alors

$$f^{-1}(b) - m = \ker(f). \tag{4.37}$$

Démonstration. Si $x \in f^{-1}(b) - m$, il existe $\alpha \in f^{-1}(b)$ tel que $x = \alpha - m$. Alors

$$f(x) = f(\alpha) - f(m) = b - b = 0. \tag{4.38}$$

Pour l'inclusion dans l'autre sens, si $z \in \ker(f)$, alors $z = (z+m) - m$ avec $z + m \in f^{-1}(b)$ parce que

$$f(z+m) = f(z) + f(m) = 0 + b = b = . \tag{4.39}$$

□

Proposition 4.41.
Si E et F sont des espaces vectoriels de dimension n et si $\{e_i\}_{i=1,\dots,n}$ et $\{f_i\}_{i=1,\dots,n}$ sont des bases respectivement de E et F, alors il existe une unique application linéaire $T\colon E \to F$ telle que $T(e_i) = f_i$ pour tout i.

Démonstration. En deux parties.

(1) **Existence** Soit $v \in E$. Vu que $\{e_i\}$ est une base, v se décompose de façon unique en $v = \sum_i v_i e_i$. Alors la définition

$$T(v) = \sum_i v_i f_i \tag{4.40}$$

est une bonne définition et satisfait aux exigences.

13. Le fait d'utiliser une notation similaire à celle des matrices inversibles n'est pas anodine : le lecteur en est sans doute conscient.
14. Définition 2.5.

(2) **Unicité** Soient T et U satisfaisant aux exigences. Alors pour tout i nous avons $T(e_i) = U(e_i)$. Si $v \in E$ s'écrit de la forme $v = \sum_i v_i e_i$ alors la linéarité impose $T(v) = \sum_i v_i T(e_i) = \sum_i v_i U(e_i) = U(v)$. Donc $T = U$.

□

LEMooJXFIooKDzRWR

Lemme 4.42 ([1]).
Soient des espaces vectoriels V et W de dimension finie. Soient des bases $\{e_i\}$ de V et $\{f_\alpha\}$ de W. Nous posons

$$\begin{aligned} \varphi_{i\alpha} \colon V &\to W \\ v &\mapsto v_i f_\alpha \end{aligned} \tag{4.41}$$

où v_i est défini par la décomposition (unique) $v = \sum_i v_i e_i$.
Alors :

(1) La partie $\{\varphi_{i\alpha}\}$ est une base de $\mathcal{L}(V, W)$.

ITEMooPMLWooNbTyJI

(2) Au niveau des dimensions : $\dim\big(\mathcal{L}(V,W)\big) = \dim(V)\dim(W)$.

Démonstration. Il faut prouver que $\{\varphi_{i\alpha}\}$ est libre et générateur.

(1) **Générateur** Soit une application linéaire $b \colon V \to W$. En décomposant $b(v)$ dans la base $\{f_\alpha\}$, nous définissons $b_\alpha \colon V \to \mathbb{K}$ par

$$b(v) = \sum_\alpha b_\alpha(v) f_\alpha. \tag{4.42}$$

Nous posons $b_{\alpha i} = b_\alpha(e_i)$. Ainsi,

$$b(v) = \sum_\alpha v_i b_{\alpha i} f_\alpha = \sum_{\alpha i} b_{\alpha i} \varphi_{i\alpha}(v). \tag{4.43}$$

Donc b peut être écrit comme combinaison linéaire des $\varphi_{i\alpha}$.

(2) **Libre** Supposons que $\sum_{i\alpha} a_{i\alpha}\varphi_{i\alpha} = 0$ pour certains coefficients $a_{i\alpha} \in \mathbb{K}$. Nous avons, pour tout $v \in V$:

$$0 = \sum_{i\alpha} a_{i\alpha}\varphi_{i\alpha}(v) = \sum_{i\alpha} a_{i\alpha} v_i f_\alpha, \tag{4.44}$$

mais comme les f_α forment une base, chaque terme de la somme sur α est nul :

$$\sum_i a_{i\alpha} v_i = 0. \tag{4.45}$$

Et comme cela est valable pour tout v et donc, pour tout choix de v_i, nous avons $a_{i\alpha} = 0$ pour tout i et pour tout α.

La formule de dimension est simplement la cardinalité de la base trouvée ; c'est la définition 4.16.
□

4.2.2 Linéarité et bases

Proposition 4.43 ([128]).
Soient deux espaces vectoriels E et F. Une application linéaire[15] *$f \colon E \to F$ est injective si et seulement si* $\ker(f) = \{0\}$.

Démonstration. Nous supposons que f est injective. Si $x \in \ker(f)$, alors $f(x) = 0$. Or f est linéaire, donc $f(0) = 0$. Nous avons donc $f(x) = f(0)$ et donc $x = 0$ parce que f est injective.

Dans l'autre sens, soient x, y tels que $f(x) = f(y)$. Par linéarité de f nous avons $f(x - y) = 0$, et donc $x - y = 0$ parce que $\ker(f) = \{0\}$. Donc $x = y$ et f est injective. □

15. Définition 4.29.

PROPooZFKZooBGLSex

Proposition 4.44 ([128]).
Soit $f \in \mathcal{L}(E,F)$ où E et F sont deux espaces vectoriels.

ITEMooPPMEooIaZqtm
(1) Si f est injective et si $\{v_i\}_{i\in I}$ est libre, alors $\{f(v_i)\}_{i\in I}$ est libre.

ITEMooOZSPooQBrDGi
(2) Si f est surjective et si $\{v_i\}_{i\in I}$ est génératrice, alors $\{f(v_i)\}_{i\in I}$ est génératrice.

ITEMooOIEYooIfdFnv
(3) Si f est une bijection, alors l'image d'une base par f est une base.

Démonstration. En trois parties.

(1) **(1)** Nous devons montrer que $\{f(v_j)\}_{j\in J}$ est libre pour tout J fini dans I. Soit donc une partie finie $J \subset I$ et des scalaires[16] tels que $\sum_{j\in J}\lambda_j f(v_j) = 0$. La linéarité de f donne[17]

$$f\big(\sum_{j\in J}\lambda_j v_j\big) = 0. \tag{4.46}$$

Par injectivité de f nous avons alors $\sum_j \lambda_j v_j = 0$. Comme les v_j eux-même forment une partie libre, nous avons $\lambda_j = 0$ pour tout $j \in J$.

(2) **(2)** Soit $y \in F$. Puisque f est surjective, il existe $x \in E$ tel que $f(x) = y$. Étant donné que $\{v_i\}_{i\in I}$ est générateur, il existe une partie finie $J \subset I$ et des scalaires $\lambda_j \in \mathbb{K}$ tels que

$$x = \sum_{j\in J}\lambda_j v_j. \tag{4.47}$$

En appliquant f aux deux côtés, et en tenant compte de la linéarité de f,

$$y = f(x) = \sum_{j\in J}\lambda_j f(v_j), \tag{4.48}$$

ce qui prouve que y est une combinaison linéaire des $f(v_j)$.

(3) **(3)** Une base est à la fois libre et génératrice et une bijection est à la fois injective et surjective. Les deux premiers points permettent de conclure.

□

CORooXIPKooWThOsr

Corolaire 4.45 ([1]).
Si E et F sont des espaces vectoriels isomorphes de dimensions finies. Alors leurs dimensions sont égales.

Démonstration. Puisque E et F sont isomorphes, il existe une bijection $f\colon E \to F$. Par la proposition 4.44(3), l'image d'une base de E est une base de F. Donc les espaces E et F ont des bases contenant le même nombre d'éléments. □

4.2.3 Rang

La proposition 4.46 et le théorème 4.47 sont valables également en dimension infinie ; ce sera une des rares incursions en dimension infinie de ce chapitre.

DefALUAooSPcmyK

Proposition-Définition 4.46.
L'image d'une application linéaire est un espace vectoriel. La dimension de cet espace est le ***rang*** *de ladite application linéaire.*

Démonstration. Soit une application linéaire $f\colon E \to F$. Nous considérons v, w dans l'image de f ainsi que λ dans le corps de base commun à E et F.

Soient $v_0 \in E$ et $w_0 \in E$ tels que $v = f(v_0)$ et $w = f(w_0)$. Alors $v + w = f(v_0 + w_0)$ et $\lambda v = f(\lambda v_0)$. Donc l'image est bien un espace vectoriel. □

16. Des éléments du corps de base $\mathbb{K}$.
17. Voir les propriétés de la définition 4.29.

ThoGkkffA

Théorème 4.47 (Théorème du rang)**.**
Soient E et F deux espaces vectoriels (de dimensions finies ou non) et soit $f\colon E \to F$ une application linéaire.

Si $(u_s)_{s\in S}$ est une base de $\ker(f)$ et si $\big(f(v_t)\big)_{t\in T}$ est une base de $\text{Image}(f)$ alors

$$(u_s)_{s\in s} \cup (v_t)_{t\in T} \tag{4.49}$$

est une base de E.

En dimension finie, nous avons en plus la formule suivante :

EQooUEOQooLySRiE

$$\text{rk}(f) + \dim(\ker f) = \dim E, \tag{4.50}$$

c'est-à-dire que le rang[18] *de f est égal à la codimension*[19] *du noyau.*

Démonstration. Nous devons montrer que

$$(u_s)_{s\in S} \cup (v_t)_{t\in T} \tag{4.51}$$

est libre et générateur.

Soit $x \in E$. Nous définissons les nombres x_t par la décomposition de $f(x)$ dans la base $\big(f(v_t)\big)$:

$$f(x) = \sum_{t\in T} x_t f(v_t). \tag{4.52}$$

Ensuite le vecteur $x - \sum_t x_t v_t$ est dans le noyau de f, par conséquent nous le décomposons dans la base (u_s) :

$$x - \sum_t x_t v_t = \sum_{s\in S} x_s u_s. \tag{4.53}$$

Par conséquent

$$x = \sum_s x_s u_s + \sum_t x_t v_t. \tag{4.54}$$

En ce qui concerne la liberté nous écrivons

$$\sum_t x_t v_t + \sum_s x_s u_s = 0. \tag{4.55}$$

En appliquant f nous trouvons que

$$\sum_t x_t f(v_t) = 0 \tag{4.56}$$

et donc que les x_t doivent être nuls. Nous restons avec $\sum_s x_s u_s = 0$ qui à son tour implique que $x_s = 0$. □

Un exemple d'utilisation de ce théorème en dimension infinie sera donné dans le cadre du théorème de Fréchet-Riesz, théorème 25.18. Il existe une généralisation du théorème du rang pour les variétés différentiables en le théorème 49.141.

PROPooQCIXooHIyPPq

Proposition 4.48 ([129])**.**
Soit E, un espace vectoriel de dimension finie sur le corps $\mathbb{K}$. Soient V et W des sous-espaces vectoriels de E. Alors

$$\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W). \tag{4.57}$$

18. Définition 4.46.
19. Définition 4.20.

Démonstration. Nous considérons l'application

$$\begin{aligned}\varphi\colon V\times W&\to E\\ (x,y)&\mapsto x+y.\end{aligned}\tag{4.58}$$

C'est une application linéaire dont l'image est $V+W$. Nous avons donc, pour commencer

$$\dim(V+W)=\dim\big(\text{Image}(\varphi)\big).\tag{4.59}$$

Nous appliquons à présent le théorème du rang 4.47 à l'application φ :

$$\begin{aligned}\dim(V+W)&=\dim\big(\text{Image}(\varphi)\big)&\text{(4.60a)}\\ &=\dim(V\times W)-\dim\big(\ker(\varphi)\big)&\text{(4.60b)}\\ &=\dim(V)+\dim(W)-\dim\big(\ker(\varphi)\big).&\text{(4.60c)}\end{aligned}$$

Nous devons maintenant étudier $\ker(\varphi)$. D'abord, $(v,w)\in V\times W$ appartient à $\ker(\varphi)$ si et seulement si $v+w=0$. Nous avons donc

$$\ker(\varphi)=\{(x,-x)\text{ tel que }x\in V\cap W\}.\tag{4.61}$$

Nous montrons à partir de cela que $\dim\big(\ker(\varphi)\big)=\dim(V\cap W)$ en montrant que l'application

$$\begin{aligned}\psi\colon V\cap W&\to\ker(\varphi)\\ x&\mapsto(x,-x)\end{aligned}\tag{4.62}$$

est un isomorphisme d'espaces vectoriels. D'abord ψ est injective parce que si $\psi(x)=\psi(y)$, alors $(x,-x)=(y,-y)$ et donc $x=y$. Ensuite, ψ est surjective parce qu'un élément générique de $\ker(\varphi)$ est $(x,-x)=\psi(x)$ avec $x\in V\cap W$. L'application ψ étant un isomorphisme d'espaces vectoriels, nous avons bien $\dim\big(\ker(\varphi)\big)=\dim(V\cap W)$. □

CORooCCXHooALmxKk

Corolaire 4.49.
Soient deux espaces vectoriels E et F de même dimension finie[20]*. Pour une application linéaire $f\colon E\to F$, les trois conditions suivantes sont équivalentes :*

(1) f est injective ;

(2) f est surjective ;

(3) f est bijective.

Démonstration. Si $f\colon E\to E$ est surjective, alors $\text{rk}(f)=\dim(E)$, ce qui donne, par le théorème du rang 4.47, $\dim\big(\ker(f)\big)=0$, c'est-à-dire que f est injectif.

De la même façon, si f est injective, alors $\dim\big(\ker(f)\big)=0$, ce qui donne $\text{rk}(f)=\dim(E)$ ou encore que f est surjective. □

Exemple 4.50.
Le corolaire 4.49 n'est pas correct en dimension infinie. Par exemple en prenant $f(e_1)=f(e_2)=e_1$ et ensuite $f(e_k)=e_{k-1}$ pour tout $k\geqslant 2$. Cette application est surjective mais pas injective. △

Une conséquence du théorème du rang est que les endomorphismes ont un inverse à gauche et à droite égaux (lorsqu'ils existent). En résumé, ce que le corolaire 4.51 dit est que si $AB=\mathbb{1}$, alors $BA=\mathbb{1}$.

CORooNFJLooJtzFwN

Corolaire 4.51.
Soit un endomorphisme f d'un espace vectoriel de dimension finie. Si f admet un inverse à gauche, alors

(1) f est bijective,

20. Les deux mots sont importants : les dimensions doivent être égales et finies.

(2) f admet également un inverse à droite,

(3) les inverses à gauche et à droite sont égaux.

Tout cela tient également en remplaçant « gauche » par « droite ».

Démonstration. Soit g, un inverse à gauche de f : $gf = \mathrm{Id}$. Cela implique que f est injective et que g est surjective, et donc qu'elles sont toutes deux bijectives par le corolaire 4.49. Puisque f est bijective, elle admet également un inverse à droite, soit h. Nous avons : $gf = \mathrm{Id}$ et $fh = \mathrm{Id}$.

Alors $gfh = h$ parce que $gf = \mathrm{Id}$, mais également $gfh = g$ parce que $fh = \mathrm{Id}$. Donc $g = h$.[21] □

C'est ce corolaire qui nous permet d'écrire f^{-1} sans plus de précisions dès que f est une bijection.

Exemple 4.52 (Pas en dimension infinie)**.**
Tout cela ne fonctionne pas en dimension infinie. Par exemple avec une base $\{e_k\}_{k\in\mathbb{N}}$ nous pouvons considérer l'opérateur

$$f(e_k) = e_{k+1}. \tag{4.63}$$

Il est injectif, mais pas surjectif. Si on pose

$$g(e_k) = \begin{cases} e_{k-1} & \text{si } k \geqslant 1 \\ 0 & \text{si } k = 0 \end{cases} \tag{4.64}$$

alors nous avons $gf = \mathrm{Id}$, mais pas $fg = \mathrm{Id}$ parce que ce $(fg)(e_0) = 0$. △

LEMooRZDTooEuLTrO

Lemme 4.53.
Si E et F sont des espaces vectoriels et si $f\colon E \to F$ est une application linéaire inversible, alors son inverse est également linéaire.

Démonstration. Nous avons $f^{-1}(x + y) = f^{-1}(x) + f^{-1}(y)$. En effet,

$$f\big(f^{-1}(x) + f^{-1}(y)\big) = f\big(f^{-1}(x)\big) + f\big(f^{-1}(y)\big) = x + y. \tag{4.65}$$

De la même façon,

$$f\big(\lambda f^{-1}(x)\big) = \lambda x, \tag{4.66}$$

donc $f^{-1}(\lambda x) = \lambda f^{-1}(x)$. □

PROPooHLUYooNsDgbn

Proposition 4.54.
Soient un espace vectoriel E de dimension finie, un endomorphisme $f\colon E \to E$ et une partie $\{v_i\}_{i\in I}$ tel que $\{f(v_i)\}_{i\in I}$ soit une base.

Alors $\{v_i\}_{i\in I}$ est une base.

Démonstration. Soit $x \in E$. Il existe une partie finie $J \subset I$ et des scalaires λ_j tels que

$$x = \sum_j \lambda_j f(v_j) = f\Big(\sum_j \lambda_j v_j\Big), \tag{4.67}$$

ce qui prouve que f est surjective. Le corolaire 4.49 nous dit alors que f est une bijection. L'application inverse est également linéaire par le lemme 4.53.

Une application linéaire bijective (comme f^{-1}) transforme une base en une base par la proposition 4.44. Donc

$$f^{-1}\big(\{f(v_i)\}\big) \tag{4.68}$$

est une base. □

21. C'est le même argument que celui employé pour la preuve du lemme 1.164 (2), à ceci près que nous devions montrer l'existence de l'inverse à droite.

PROPooADESooATJSrH

Proposition 4.55.
Soit un espace vectoriel E de dimension finie et deux applications linéaires $f, g\colon E \to E$ telles que $g \circ f = \mathrm{Id}$. Alors f et g sont bijectives.

Démonstration. En plusieurs étapes

(1) **f est injective** Si $f(x) = f(y)$, alors en appliquant g nous avons
$$g\big(f(x)\big) = g\big(f(y)\big), \tag{4.69}$$
ce qui donne $x = y$.

(2) **f est surjective** C'est maintenant le corolaire 4.49.

(3) **g est surjective** Pour tout $x \in E$ nous avons $g\big(f(x)\big) = x$. Donc l'image de $f(E)$ par g est E.

(4) **g est injective** C'est maintenant le corolaire 4.49.

□

LEMooDAACooElDsYb

Lemme 4.56 ([130])**.**
Soit une application linéaire $f\colon E \to F$.

ITEMooEZEWooZGoqsZ

(1) L'application f est injective si et seulement si il existe $g\colon F \to E$ telle que $g \circ f = \mathrm{Id}\,|_E$.

(2) L'application f est surjective si et seulement si il existe $g\colon F \to E$ telle que $f \circ g = \mathrm{Id}\,|_F$.

Démonstration. Nous démontrons séparément les deux affirmations.

(1) Si f est injective, alors $f\colon E \to \mathrm{Image}(f)$ est un isomorphisme. Si V est un supplémentaire de $\mathrm{Image}(f)$ dans F (c'est-à-dire $F = \mathrm{Image}(f) \oplus V$) alors nous pouvons poser $g(x+v) = f^{-1}(x)$ où $x + v$ est la décomposition (unique) d'un élément de F en $x \in \mathrm{Image}(f)$ et $v \in V$. Avec cela nous avons bien $g \circ f = \mathrm{Id}$.
Réciproquement, si il existe $g\colon F \to E$ telle que $g \circ f = \mathrm{Id}$ alors $f\colon E \to E$ doit être injective. Parce que si $f(x) = 0$ avec $x \neq 0$ alors $(g \circ f)(x) = 0 \neq x$.

(2) Si f est surjective nous pouvons choisir des éléments $x_1, \dots, x_p$ dans E tels que $\{f(x_i)\}$ soit une base de F. Ensuite nous définissons
$$\begin{aligned} g\colon F &\to E \\ \sum_k a_k f(x_k) &\mapsto \sum_k a_k x_k. \end{aligned} \tag{4.70}$$
Cela donne $f \circ g = \mathrm{Id}\,|_F$ parce que si $v \in F$ alors $v = \sum_k v_k f(x_k)$ avec $v_k \in \mathbb{K}$, et nous avons
$$(f \circ g)(v) = \sum_k v_k (f \circ g)\big(f(x_k)\big) = f\left(\sum_k v_k x_k\right) = \sum_k v_k f(x_k) = v. \tag{4.71}$$
Réciproquement, si il existe $g\colon F \to E$ tel que $f \circ g = \mathrm{Id}$ alors f doit être surjective, parce que
$$F = \mathrm{Image}(f \circ g) = f\big(\mathrm{Image}(g)\big) \subset \mathrm{Image}(f). \tag{4.72}$$

□

4.3 Matrices

Les matrices et les applications linéaires sont deux choses différentes. Une application linéaire [22] est une application d'un espace vectoriel vers un autre, et une matrice est un simple tableau de nombres sur lesquels nous définissons des opérations, de telle sorte à fournir une structure d'espace vectoriel. Le lien entre ces opérations et les opérations correspondantes sur les applications linéaires sera fait plus tard. Voir la définition 4.68 et ce qui s'en suit.

22. Définition 4.29.

4.3.1 Définitions

Les notions topologiques sur les espaces de matrices sont pour plus tard, à commencer par la définition 9.33.

Définition 4.57.
Soit un anneau $\mathbb{A}$ ainsi que des entiers m, n strictement positifs. L'ensemble $\mathbb{M}(n\times m,\mathbb{A})$ est l'ensemble des applications

$$\{1,\dots,n\}\times\{1,\dots,m\}\to\mathbb{A}, \tag{4.73}$$

*et est appelé ensemble des **matrices** $n\times m$ sur $\mathbb{A}$.*

Si A est une matrice, nous notons $A_{i,j}$ au lieu de $A(i,j)$ l'image de (i,j) par l'application A.

Définition 4.58.
Quelques ensembles de matrices particuliers.

(1) *Si $n=m$, alors :*
   — *nous disons que la matrice est **carrée**,*
   — *nous notons $\mathbb{M}(n,\mathbb{A})$ pour $\mathbb{M}(n\times n,\mathbb{A})$,*
   — *n est appelée **ordre** de la matrice.*

(2) *Si $n=1$, alors la matrice est appelée **matrice-ligne**.*

(3) *Si $m=1$, alors la matrice est appelée **matrice-colonne**.*

4.59.
On note les isomorphismes naturels $\mathbb{M}(1\times m,\mathbb{A})\simeq\mathbb{A}^m$ et $\mathbb{M}(n\times 1,\mathbb{A})\simeq\mathbb{A}^n$.

LEMooYWTEooQyLxKv

Lemme-Définition 4.60.
Nous considérons les opérations suivantes sur $\mathbb{M}(n\times m,\mathbb{A})$:

Somme $(A+B)_{i,j}=A_{i,j}+B_{i,j}$,

Produit par un scalaire $(\lambda\cdot A)_{i,j}=\lambda A_{i,j}$ *pour tout $A,B\in\mathbb{M}(n\times m,\mathbb{A})$ et $\lambda\in\mathbb{A}$.*

Alors $\big(\mathbb{M}(n\times m,\mathbb{A}),+,\cdot\big)$ est un $\mathbb{A}$-module[23].

LEMooMBZTooKdGvON

Lemme-Définition 4.61.
Avec la multiplication

$$\begin{aligned}\mathbb{M}(n\times p,\mathbb{A})\times\mathbb{M}(p\times m,\mathbb{A})&\to\mathbb{M}(n\times m,\mathbb{A})\\ (A,B)&\mapsto(AB)_{i,j}=\sum_{k=1}^{p}A_{i,k}B_{k,j},\end{aligned} \tag{4.74}$$

l'espace $\mathbb{M}(n,\mathbb{K})$ est une $\mathbb{K}$-algèbre[24].

Démonstration. Il faut vérifier les trois formules de la définition 1.347.

(1) **Distribution (1)** Nous avons

$$\big((x+y)\times z\big)_{ij}=\sum_k(x+y)_{ik}z_k=\sum_k\big(x_{ik}z_kj+y_{ik}z_{kj}\big)=(xz)_{ij}+(yz)_{ij}. \tag{4.75}$$

(2) **Distribution (2)** Même calcul.

(3) **Multiplication par un scalaire** Encore le même genre de calculs, en partant de

$$\big((\alpha x)\times(\beta y)\big)_{ij}=\sum_k(\alpha x)_{ik}(\beta y)_{kj}. \tag{4.76}$$

23. Définition 1.330
24. Définition 1.347.

□

Définition 4.62.
Pour un élément $A \in \mathbb{M}(n \times m, \mathbb{A})$ nous définissons encore
La transposée $A^t_{i,j} = A_{j,i}$,
La trace $\mathrm{Tr}(A) = \sum_i A_{i,i}$.

Remarque 4.63.
Quelques remarques directes sur les définitions.

(1) La motivation de cette définition pour le produit apparaîtra plus loin, mais le Frido n'étant pas un livre d'introduction, j'imagine que le lecteur a déjà une idée.
(2) Nous verrons plus loin en 8.11.3 que la définition de transposée d'une application linéaire n'est pas tout à fait évidente ; elle sera la définition 8.189.
Ici nous avons bien défini la transposée d'une matrice, pas d'une application linéaire.

Remarque 4.64.
Quelques remarques à propos de structures supplémentaires.

(1) Nous utiliserons (presque) tout le temps des matrices à coefficients dans un corps. Il est clair que, si $\mathbb{K}$ est un corps (commutatif), alors $\mathbb{M}(n \times m, \mathbb{K})$ a une structure d'espace vectoriel sur $\mathbb{K}$.
(2) Par ailleurs, sur les matrices carrées d'ordre n fixé, le produit de deux matrices est bien défini. Ainsi, $\mathbb{M}(n, \mathbb{A})$ se voit conférer une structure d'anneau, dont le neutre pour la multiplication est la matrice carrée $\mathbb{1}_n$ (notée aussi $\mathbb{1}$ lorsqu'il n'y a pas d'ambiguïté sur la taille), donnée par
$$\mathbb{1}_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases} \tag{4.77}$$
Il est vite vu que si A est une matrice carrée d'ordre n, alors $A\mathbb{1} = \mathbb{1}A = A$.

LEMooUXDRooWZbMVN

Lemme 4.65 ([1])**.**
Si A, B et C sont des matrices, nous avons
(1) $(AB)^t = B^t A^t$,
(2) $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB)$.

ITEMooXDYQooAlnArd

Démonstration. La première est un simple calcul :
$$(AB)^t_{i,j} = (AB)_{j,i} = \sum_k A_{j,k}B_{k,i} = \sum_k A^t_{k,j}B^t_{i,k} = (B^tA^t)_{i,j}. \tag{4.78}$$
Pour la seconde :
$$\mathrm{Tr}(ABC) = \sum_{ikl} A_{i,k}B_{k,l}C_{l,i} = \sum_{ikl} C_{l,i}A_{i,k}B_{k,l} = \sum_l (CAB)_{l,l} = \mathrm{Tr}(CAB). \tag{4.79}$$

□

4.66.
La seconde égalité est importante et est nommée **invariance cyclique** de la trace. Elle sert, entre autres nombreuses choses, à prouver que la trace d'une matrice d'une application linéaire ne dépend pas de la base choisie. Ce sera la proposition 8.207.

LEMooLXAHooPRyHaF

Lemme 4.67.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$. Si pour tout $x, y \in \mathbb{K}^n$ nous avons
$$\sum_{ij} A_{i,j}x_iy_j = \sum_{ij} B_{i,j}x_iy_j \tag{4.80}$$
alors $A = B$.

Démonstration. Il suffit de choisir $x_i = \delta_{i,k}$ et $y_j = \delta_{j,l}$ et d'effectuer les sommes ; par exemple

$$\sum_{ij} A_{i,j}\delta_{i,k}y_j = \sum_j A_{k,j}y_j. \tag{4.81}$$

Après avoir effectué toutes les sommes, nous nous retrouvons avec $A_{k,l} = B_{k,l}$, ce qui signifie $A = B$. □

4.3.2 Identifier matrices et applications linéaires

Voir dans l'index thématique, -1.1.1.

Soient deux espaces vectoriels de dimension finie E, F sur le corps $\mathbb{K}$. Nous considérons les bases[25] $\{e_i\}$ pour E et $\{f_\alpha\}$ pour F.

DEFooJVOAooUgGKme

Définition 4.68.
Nous considérons l'application

$$\begin{aligned}\psi\colon \mathbb{M}(n\times m,\mathbb{K}) &\to \mathcal{L}(E,F)\\ A &\mapsto f_A\end{aligned} \tag{4.82}$$

EQooVZQWooMyFFsQ

où f_A est définie par

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i}x_i f_\alpha \tag{4.83}$$

EQooBVGHooJhFbMs

si x_i sont les coordonnées de $x \in E$ dans la base $\{e_i\}$.

4.69.
Nous allons prouver un certain nombre de résultats montrant que cette application a toutes les propriétés imaginables permettant d'identifier les matrices aux applications linéaires : elle est un isomorphisme pour toutes les structures que vous pouvez raisonnablement imaginer.

À cette application ψ il manque cependant une propriété importante : elle n'est pas canonique. Elle dépend des bases choisies. Autrement dit : nous avons à priori autant d'applications ψ différentes qu'il y a de choix de bases sur E et F[26].

Nous allons prouver maintenant quelques résultats montrant que les matrices et les applications linéaires, dans le cas des espaces vectoriels $\mathbb{K}^n$ sont deux présentations de la même chose.

Le fait que ψ est continue sera la proposition 12.117.

4.70.
Lorsque $A \in \mathbb{M}(n,\mathbb{K})$ est une matrice et $x \in \mathbb{K}^n$ un vecteur, nous notons Ax l'élément de $\mathbb{K}^n$ donné par

$$(Ax)_i = \sum_j A_{i,j}x_j. \tag{4.84}$$

EQooQFVTooMFfzol

Autrement dit, $Ax = f_A(x)$.

Cette convention et de nombreuses autres à propos de matrice sera rappelée dans -1.1.

PROPooGXDBooHfKRrv

Proposition-Définition 4.71.
Soient deux espaces vectoriels de dimension finie E, F sur le corps $\mathbb{K}$. Nous considérons les bases $\{e_i\}$ pour E et $\{f_\alpha\}$ pour F.

Nous considérons l'application

$$\begin{aligned}\psi\colon \mathbb{M}(n\times m,\mathbb{K}) &\to \mathcal{L}(E,F)\\ A &\mapsto f_A\end{aligned} \tag{4.85}$$

25. C'est le théorème 4.14 qui nous permet de considérer des bases. Et ce théorème ne fonctionne que parce que nous avons supposé une dimension finie.

26. Bonne question. Est-ce qu'il y a moyen de construire deux choix de bases donnant la même application ψ ? Écrivez-moi si vous savez la réponse.

où f_A est définie par

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i} x_i f_\alpha \tag{4.86}$$

EQooZKEKooNYjvbP

si x_i sont les coordonnées de $x \in E$ dans la base $\{e_i\}$.

Alors

ITEMooKZYYooZPTkpq

(1) Nous avons

$$f_A(e_i)_\alpha = A_{\alpha,i}. \tag{4.87}$$

ITEMooANXFooGIuxUR

(2) Nous avons

$$f_A(e_i) = \sum_\alpha A_{\alpha,i} f_\alpha. \tag{4.88}$$

EQooOKOJooYgteNP

ITEMooXLLLooKfigfB

(3) Nous avons

$$\big(f_A(x)\big)_\alpha = \sum_i A_{\alpha,i} x_i. \tag{4.89}$$

EQooAXRJooUwHbiR

ITEMooHSMLooRJZref

(4) L'application ψ est une bijection.

Si f est une application linéaire, alors la matrice $\psi^{-1}(f)$ est la ***matrice associée*** *à f dans les bases choisies.*

Remarque : les bases ne sont supposées être canoniques en aucun sens du terme. Les dimensions de E et F ne sont pas non plus supposées identiques.

Démonstration. En nous rappelant que $(e_j)_i = \delta_{i,j}$ nous avons

$$f_A(e_j) = \sum_{i\alpha} A_{\alpha,i}(e_j)_i f_\alpha = \sum_\alpha A_{\alpha,j} f_\alpha, \tag{4.90}$$

EQooWGZHooIPovgB

donc $f_A(e_i)_\alpha = A_{\alpha,i}$. Cela prouve la formule du point (1).

Le point (2) est une simple somme sur α de (1).

La formule du point (3) est simplement la composante f_α de la définition 4.86.

Prouvons que ψ est injective. Si $f_A = f_B$, nous avons en particulier $f_A(e_i)_\alpha = f_B(e_i)_\alpha$ et donc $A_{\alpha,i} = B_{\alpha,i}$.

Prouvons que ψ est surjective. Pour cela nous considérons $f \in \mathcal{L}(E,F)$ et nous posons $A_{\alpha,i} = f(e_i)_\alpha$. Nous avons alors $f = f_A$ parce que

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i} x_i f_\alpha = \sum_{i\alpha} f(e_i)_\alpha x_i f_\alpha = \sum_\alpha f(\sum_i x_i e_i)_\alpha f_\alpha = \sum_\alpha f(x)_\alpha f_\alpha = f(x). \tag{4.91}$$

□

La proposition suivante montre que le produit matriciel correspond à la composition d'applications linéaires, pourvu que l'on travaille avec les bases canoniques sur $\mathbb{K}^n$.

PROPooIYVQooOiuRhX

Proposition 4.72 ([1])**.**
Soit un corps commutatif $\mathbb{K}$. Nous considérons des espaces vectoriels E et F munis de bases $\{e_i\}_{i=1,\dots,n}$ et $\{f_\alpha\}_{\alpha=1,\dots,m}$.

L'application déjà définie[27]

$$\psi\colon \mathbb{M}(m\times n,\mathbb{K}) \to \mathcal{L}(E,F) \tag{4.92}$$

est un isomorphisme d'espaces vectoriels.

27. Notez la position du n et du m. Sachez noter les bornes des sommes écrites dans la démonstration.

Démonstration. Le fait que ψ soit une bijection est la proposition 4.71. Nous devons montrer que c'est une application linéaire.

Pour $\lambda \in \mathbb{K}$ nous avons le calcul

$$\psi(\lambda A)(e_k) = f_{\lambda A}(e_k) = \sum_{\alpha i} (\lambda A)_{\alpha,i} \underbrace{(e_k)_i}_{=\delta_{k,i}} f_\alpha = \lambda \sum_\alpha A_{\alpha,k} f_\alpha = \lambda f_A(e_k). \tag{4.93}$$

Donc $\psi(\lambda A) = \lambda\psi(A)$.

Si $A, B \in \mathbb{M}(n, \mathbb{K})$ nous avons de la même façon, $f_{A+B} = f_A + f_B$. □

PROPooCSJNooEqcmFm

Proposition 4.73.
Soient des espaces vectoriels E, F et G de dimensions n, m et p munis de bases[28] *$\{e_i\}$, $\{f_i\}$ et $\{g_i\}$. Nous considérons les applications*

$$\psi\colon \mathbb{M}(m \times n, \mathbb{K}) \to \mathcal{L}(E, F) \tag{4.94a}$$
$$\psi\colon \mathbb{M}(p \times m, \mathbb{K}) \to \mathcal{L}(F, G) \tag{4.94b}$$
$$\psi\colon \mathbb{M}(p \times n, \mathbb{K}) \to \mathcal{L}(E, G). \tag{4.94c}$$

Nous avons

$$\psi(A) \circ \psi(B) = \psi(AB) \tag{4.95}$$

pour toutes matrices $A \in \mathbb{M}(p \times m, \mathbb{K})$ et $B \in \mathbb{M}(m \times n, \mathbb{K})$.

Démonstration. Nous considérons les applications linéaires associées à A et B : $f_A\colon F \to G$ et $f_B\colon E \to F$ et la composée $f_A \circ f_B\colon E \to G$. Et puis c'est le calcul :

$$(f_A \circ f_B)(e_k) = f_A\Big(\sum_{ij} B_{i,j}(e_k)_j f_i\Big) \tag{4.96a}$$
$$= \sum_i B_{i,k} f_A(f_i) \tag{4.96b}$$
$$= \sum_i B_{i,k} \sum_{rs} A_{r,s}(f_i)_s g_r \tag{4.96c}$$
$$= \sum_{ir} B_{i,k} A_{r,i} g_r \tag{4.96d}$$
$$= \sum_r (AB)_{r,k} g_r \tag{4.96e}$$
$$= f_{AB}(e_k). \tag{4.96f}$$

Donc $f_A \circ f_B = f_{AB}$ comme il se doit. □

Nous pouvons particulariser au cas où $E = F = G$.

PROPooFMBFooEVCLKA

Proposition 4.74.
Si E est un espace vectoriel muni d'une base $\{e_i\}$, alors l'application

$$\psi\colon \mathbb{M}(n, \mathbb{K}) \to \mathrm{End}(E) \tag{4.97}$$

est un isomorphisme d'algèbre[29] *et d'anneaux*[30].

Démonstration. Le fait que ψ soit un isomorphisme d'algèbre est juste la combinaison entre les propositions 4.72 et 4.73.

28. Avec trois ensembles, nous renonçons à utiliser des alphabets différents pour numéroter les éléments des bases.
29. Définition 1.347.
30. Définition 1.43

En ce qui concerne l'isomorphisme d'anneaux, il faut en plus identifier les neutres. Le neutre pour la composition d'applications linéaires est l'application identité et le neutre pour la multiplication de matrices est la matrice identité. Nous devons donc montrer que $\psi(\delta) = f_\delta = \mathrm{Id}$. Juste un calcul :

$$f_\delta(x) = \sum_{ij} \delta_{i,j} x_j e_i = \sum_i x_i e_i = x. \tag{4.98}$$

Donc oui, f_δ est l'identité. □

Le fait que ψ est continue sera la proposition 12.117.

Voilà. Soyez bien conscient que l'application ψ dont nous avons beaucoup parlé est surtout intéressante dans le cas des espaces de la forme $\mathbb{K}^n$. Dans ce cas, nous avons une identification canonique entre $\mathbb{M}(n, \mathbb{K})$ et $\mathrm{End}(\mathbb{K}^n)$ qui est un isomorphisme d'anneaux et d'algèbres.

Nous verrons que ce ψ respecte encore les inverses [31] et les déterminants [32].

4.75.
Il convient de ne pas confondre matrice et application linéaire (bien que nous le ferons sans vergogne). Une matrice est un bête tableau de nombres, tandis qu'une application linéaire est une application entre deux espaces vectoriels vérifiant certaines propriétés.

Cependant si les espaces vectoriels E et F sont munis de bases, alors il y a une application

$$\psi\colon \mathbb{M}(m \times n, \mathbb{K}) \to \mathcal{L}(E, F) \tag{4.99}$$

qui a toutes les propriétés imaginables [33].

Cette application dépend des bases choisies. Il n'y a donc pas de trucs comme « la matrice de telle application linéaire » ou comme « voici une matrice, nous considérons l'application linéaire associée ».

Cependant, sur des espaces comme $\mathbb{R}^n$ ou plus généralement sur $\mathbb{K}^n$, nous avons une base canonique et toute personne raisonnable utilise toujours la base canonique (sauf mention du contraire). Dans ces cas il est sans danger de dire « la matrice associée à telle application linéaire » sans préciser les bases.

Mais si un jour vous utilisez une base autre que la base canonique sur $\mathbb{R}^n$, précisez-le et plutôt deux fois qu'une [34].

4.76.
Tant que nous sommes à parler de matrice et d'applications linéaires, les plus acharnés anti-abus de language [35] remarqueront qu'il n'est pas vrai que « étant donné une base, une application linéaire a une matrice ».

En effet, une base est une partie libre et génératrice (définition 4.5). Or une partie d'un ensemble n'est pas muni d'un ordre. Toutes les permutations de colonnes de la matrice sont encore possible d'après l'ordre que l'on met sur les vecteurs de la base.

Encore une fois, la base canonique n'a pas de problème parce que les $\{e_i\}$ de $\mathbb{R}^n$ viennent avec un ordre indiscutable. Plus généralement, très souvent, lorsqu'on construit une base, la construction suggère un ordre.

4.3.3 Ensemble des applications linéaires

PROPooRIFBooGOvsfb

Proposition 4.77 ([1]).
Soient deux espaces vectoriels de dimension finie V et W. Soit des bases $\{e_i\}_{i\in I}$ de V et $\{e'_j\}_{j\in J}$

31. Proposition 4.96.
32. Proposition 8.17.
33. Et elle en aura encore plus lorsque nous aurons vu les déterminants.
34. Au passage, non, les coordonnées polaires ne sont pas une base de $\mathbb{R}^2$. C'est un système de coordonnées, et ce n'est pas la même chose.
35. Dont l'auteur de ces lignes fait partie.

de W. Nous considérons les applications

$$\begin{aligned} f_{ij}\colon V &\to W \\ \sum_k x_k e_k &\mapsto x_i e'_j. \end{aligned} \tag{4.100}$$

EQooBMKQooCCypzP

Nous avons :

ITEMooSEMLooXuxrpk

(1) Ces applications sont linéaires.

(2) Elles forment une base de $\mathcal{L}(V,W)$.

(3) $\dim(V,W) = \dim(V)\dim(W)$.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

PROPooPNQNooAZpojm

Proposition 4.78 ([1]).
Soient des bases $\{e_i\}$ et $\{e'_j)\}$ des espaces vectoriels V et W. Soit une application linéaire $A \in \mathrm{End}(V,W)$. Alors

$$A = \sum_{ij} A_{ji} f_{ij} \tag{4.101}$$

où les f_{ij} sont les éléments de bases définis en (4.100) et A_{ji} sont les éléments de matrices définis par la proposition 4.71.

Démonstration. Il s'agit d'un calcul. En appliquant à e_k, nous avons

$$\sum_{ij} A_{ji} f_{ij}(e_k) = \sum_{ij} A_{ji}\delta_{ik} e'_j = \sum_j A_{jk} e'_j. \tag{4.102}$$

En comparant avec la formule de la proposition 4.71(1), nous avons le résultat. □

4.3.4 Déterminant

DEFooYCKRooTrajdP

Définition 4.79.
*Si $A \in \mathbb{M}(n,\mathbb{K})$ nous définissons le **déterminant** de A par la formule*

$$\det(A) = \sum_{\sigma\in S_n} (-1)^\sigma \prod_{i=1}^{n} A_{i,\sigma(i)} \tag{4.103}$$

où la somme est effectuée sur tous les éléments du groupe symétrique[36] *S_n et où $(-1)^\sigma$ représente la parité de la permutation σ.*

En se souvenant que $|S_n| = n!$, nous sommes frappés de stupeur devant le fait que le nombre de termes dans la somme croît de façon factorielle (c'est plus qu'exponentiel, pour info) en la taille de la matrice. Cette formule est donc sans espoir pour une matrice plus grande que 3×3 ou à la rigueur 4×4 à la main. À l'ordinateur, il est possible de monter plus haut, mais pas tellement.

4.3.5 Déterminant en petite dimension

En dimension deux, le déterminant de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est le nombre

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb. \tag{4.104}$$

EQooQRGVooChwRMd

Ce nombre détermine entre autres le nombre de solutions que va avoir le système d'équations linéaires associé à la matrice.

36. Pour le groupe symétrique, c'est la définition 1.274, le fait que ce soit un groupe fini est le lemme 1.277, et pour la somme sur un groupe fini c'est la définition 1.309.

Pour une matrice 3×3, nous avons le même concept, mais un peu plus compliqué ; nous avons la formule

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33} \end{pmatrix}=\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33} \end{vmatrix}=a_{11}\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33} \end{vmatrix}-a_{12}\begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33} \end{vmatrix}+a_{13}\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32} \end{vmatrix}. \tag{4.105}$$

4.3.6 Manipulations de lignes et de colonnes

Nous voudrions savoir ce qu'il se passe avec le déterminant d'une matrice lorsque nous substituons à une ligne ou une colonne, une combinaison des autres lignes et colonnes. Lorsqu'une matrice est donnée, nous notons C_j sa j^{e} colonne.

LEMooRSJTooQEoOtN

Lemme 4.80 ([1]).
Si A et B sont des matrices, alors

$$(AB)^t=B^tA^t. \tag{4.106}$$

Démonstration. Il suffit de calculer les éléments de matrice :

$$(AB)^t_{i,j}=(AB)_{j,i}=\sum_k A_{j,k}B_{k,i}=\sum_k B^t_{i,k}A^t_{k,j}=(B^tA^t)_{i,j}. \tag{4.107}$$

□

LEMooCEQYooYAbctZ

Lemme 4.81 ([1, 131]).
Si A est une matrice, alors $\det(A)=\det(A^t)$.

Démonstration. Nous commençons par écrire la définition du déterminant :

$$\det(A^t)=\sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n}(A^t)_{i,\sigma(i)}=\sum_{\sigma}\epsilon(\sigma)\prod_i A_{\sigma(i),i}. \tag{4.108}$$

Pour chaque σ séparément, en utilisant la proposition 1.317 pour ré-indexer le produit :

$$\prod_i A_{\sigma(i),i}=\prod_i A_{i,\sigma^{-1}(i)}. \tag{4.109}$$

Nous profitons du fait que l'application $\varphi\colon S_n\to S_n$ donnée par $\varphi(\sigma)=\sigma^{-1}$ soit une permutation de S_n pour appliquer la définition 1.309 et faire la somme sur σ^{-1} :

$$\det(A^t)=\sum_{\sigma}\epsilon(\sigma)\prod_i A_{i,\sigma^{-1}(i)}=\sum_{\sigma}\epsilon(\sigma^{-1})\prod_i A_{i,\sigma(i)}=\det(A) \tag{4.110}$$

où nous avons utilisé le fait que $\epsilon(\sigma^{-1})=\epsilon(\sigma)$ (corolaire 1.301). □

Le fait que $\det(A)=\det(A^t)$ permet, dans toutes les propositions du type « ce qui arrive au déterminant si on change telle ligne ou colonne » de ne donner qu'une preuve pour la partie « ligne » et déduire automatiquement le cas « colonne ». Le lemme suivant donne un exemple d'utilisation.

LEMooWMQWooGWFlmC

Lemme 4.82 ([1]).
Soit une matrice A. Nous considérons la matrice B obtenue à partir de A par la permutation de lignes $L_k\leftrightarrow L_l$ ainsi que la matrice C obtenue à partir de A^t par la permutation de colonnes $C_k\leftrightarrow C_l$.

Alors $C^t=B$.

Démonstration. Calculons les éléments de matrice de C :

$$C_{i,j}=\begin{cases}(A^t)_{i,j} & \text{si } j\neq k, j\neq l\\ (A^t)_{i,k} & \text{si } j=l\\ (A^t)_{i,l} & \text{si } j=k\end{cases}=\begin{cases}A_{j,i} & \text{si } j\neq k, j\neq l\\ A_{k,i} & \text{si } j=l\\ A_{l,i} & \text{si } j=k.\end{cases} \tag{4.111}$$

Ensuite nous prouvons que $C^t = B$ en écrivant les éléments de C^t :

$$(C^t)_{i,j} = C_{j,i} = \begin{cases} A_{i,j} & \text{si } i \neq k, i \neq l \\ A_{k,j} & \text{si } i = l \\ A_{l,j} & \text{si } i = k. \end{cases} \tag{4.112}$$

Cette dernière expression est la matrice A après permutation des lignes $L_k \leftrightarrow L_l$, c'est-à-dire la matrice B. □

Pour la suite nous écrivons δ la matrice « identité », c'est-à-dire celle dont les entrées sont précisément les $\delta_{i,k}$. Nous écrivons également $E_{i,j}$ la matrice contenant des zéros partout sauf en (i,j) où elle a un 1, c'est-à-dire

$$(E_{i,j})_{k,l} = \delta_{i,k}\delta_{j,l}. \tag{4.113}$$

PROPooFQRDooRPfuxk

Proposition 4.83 (Permuter des lignes ou des colonnes $L_k \leftrightarrow L_l$[132, 1])**.**
Soient une matrice $A \in \mathbb{M}(n, \mathbb{K})$, deux entiers $k \neq l$ inférieurs ou égaux à n.

ITEMooAIHWooHXzeys

(1) Si B est la matrice obtenue à partir de A en permutant deux lignes ou deux colonnes, alors

$$\det(A) = -\det(B). \tag{4.114}$$

ITEMooDNHWooOMgmxa

(2) Si B est la matrice obtenue à partir de A par la permutation de lignes $L_k \leftrightarrow L_l$. Alors

$$B = SA \tag{4.115}$$

avec $S = \delta + E_{k,l} + E_{l,k} - E_{k,k} - E_{l,l}$.
Autrement dit : la matrice S est une matrice de permutations de lignes.

ITEMooSHRQooQrqVdO

(3) La matrice S vérifie $\det(S) = -1$

ITEMooQXSEooMWiKbL

(4) Nous avons

$$\det(SA) = \det(S)\det(A). \tag{4.116}$$

Démonstration. Point par point

(1) **(1) pour les colonnes**
Soient k et l fixés, et considérons la permutation des colonnes C_k et C_l. Nous notons α la permutation (k,l) dans S_n (groupe symétrique, définition 1.274). Nous avons

$$B_{i,j} = A_{i,\alpha(j)}, \tag{4.117}$$

ou encore : $A_{i,j} = B_{i,\alpha(j)}$. Par définition,

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)} \tag{4.118}$$

C'est le moment d'utiliser la proposition 1.313 à propos de somme sur des groupes avec $G = S_n$, $h = \alpha$ et

$$f(\sigma) = \epsilon(\sigma) \prod_i A_{i,\sigma(i)}. \tag{4.119}$$

Nous savons que $\epsilon(\alpha) = -1$ et que ϵ est un morphisme par la proposition 1.300(1), donc

$$f(\alpha\sigma) = \epsilon(\alpha\sigma) \prod_i A_{i,(\alpha\sigma)(i)} = -\epsilon(\sigma) \prod_i B_{i,\sigma(i)}. \tag{4.120}$$

Avec ça, nous concluons :

$$\det(A) = \sum_{\sigma \in S_n} f(\sigma) = \sum_{\sigma} f(\alpha\sigma) = -\sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} B_{i,\sigma(i)} = -\det(B). \tag{4.121}$$

(2) **(1) pour les lignes** Que se passe-t-il si nous permutons les lignes L_k et L_l ? Si nous notons B' la matrice obtenue à partir de A par la permutation de lignes $L_k \leftrightarrow L_l$, et C celle obtenue de A^t après permutation de colonnes $C_k \leftrightarrow C_l$ alors nous avons $C^t = B'$. Le lemme 4.82 nous dit que $C^t = B'$. En utilisant le lemme 4.81 sur le déterminant de la transposée,

$$\det(B') = \det(C^t) = \det(C) = -\det(A^t) = -\det(A). \tag{4.122}$$

Voilà qui prouve le résultat pour les permutation de lignes.

(3) **(2)** Si $k = l$, il n'y a pas de permutation, et il est vite vu que la matrice S est l'identité parce qu'il y a quatre fois le terme $E_{k,k}$. Nous supposons donc que $k \neq l$; en particulier $\delta_{k,l} = 0$.

Il s'agit surtout d'un beau calcul :

$$(SA)_{i,j} = \sum_m S_{i,m}A_{m,j} = A_{i,j} + \sum_m (\delta_{k,i}\delta_{l,m} + \delta_{l,i}\delta_{l,m} - \delta_{k,i}\delta_{k,m} - \delta_{l,i}\delta_{l,m})A_{m,j} \tag{4.123a}$$

$$= A_{i,j} + \delta_{k,i}A_{l,j} + \delta_{l,i}A_{k,j} - \delta_{k,i}A_{k,j} - \delta_{l,i}A_{l,j}. \tag{4.123b}$$

Si $i \neq j$ et $i \neq l$, alors $(SA)_{i,j} = A_{i,j}$. Si $i = k$, alors

$$(SA)_{k,j} = A_{k,j} + A_{l,j} - A_{k,j} = A_{l,j}, \tag{4.124}$$

c'est-à-dire que la k^{e} ligne de SA est la l^{e} ligne de A.

Avec $i = l$ nous obtenons la k^{e} ligne de A.

Tout cela montre que SA est la matrice A dans laquelle les lignes k et l ont été échangées, c'est-à-dire $SA = B$.

(4) **(3)** En utilisant la définition du déterminant,

$$\det(S) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} S_{i,\sigma(i)} \tag{4.125a}$$

$$= \sum_{\sigma} \epsilon(\sigma) \prod_{i} \big(\delta_{i,\sigma(i)} + \delta_{k,i}\delta_{l,\sigma(i)} + \delta_{l,i}\delta_{k,\sigma(i)} - \delta_{k,i}\delta_{k,\sigma(i)} - \delta_{l,i}\delta_{l,\sigma(i)}\big). \tag{4.125b}$$

Nous utilisons l'associativité et la commutativité du produit pour séparer les facteurs $i = k$ et $i = l$ des autres :

$$\det(S) = \sum_{\sigma} \epsilon(\sigma) \prod_{\substack{i\neq k\\ i\neq l}} \delta_{i,\sigma(i)}(\delta_{k,\sigma(k)} + \delta_{l,\sigma(k)} - \delta_{k,\sigma(k)})(\delta_{l,\sigma(l)} + \delta_{k,\sigma(l)} - \delta_{l,\sigma(l)}). \tag{4.126}$$

À cause des facteurs $i \neq k$ et $i \neq l$, les σ pour lesquels le tout n'est pas nul doivent vérifier $\delta_{i,\sigma(i)} = 1$ pour tout i différent de k et l. Les deux seuls sont donc $\sigma = \text{Id}$ et la permutation $\sigma = (k, l)$. Pour $\sigma = \text{Id}$, nous avons

$$\prod_{\substack{i\neq k\\ i\neq l}} \delta_{i,i}(\delta_{k,k} + \delta_{l,k} - \delta_{k,k})(\delta_{l,l} + \delta_{k,l} - \delta_{l,l}) = 0. \tag{4.127}$$

Dernier espoir : $\sigma = (k, l)$. Pour ce terme nous avons $\epsilon(\sigma) = -1$ et

$$\prod_{\substack{i\neq k\\ i\neq l}} \delta_{i,i}(\delta_{k,l} + \delta_{l,l} - \delta_{k,l})(\delta_{l,k} + \delta_{k,k} - \delta_{l,k}) = 1. \tag{4.128}$$

Au final dans $\det(S)$, il n'y a que le terme $\sigma = (k, l)$ qui est non nul, et il vaut -1. Donc

$$\det(S) = -1. \tag{4.129}$$

(5) **(4)** Il s'agit de mettre bout à bout les points déjà prouvés :

$$\det(SA) = -\det(A) = \det(S)\det(A). \tag{4.130}$$

□

COROooAZFCooSYINvBl

Corolaire 4.84 ([132]).
Soit une matrice $A \in \mathbb{M}(n, \mathbb{K})$. Si deux lignes ou deux colonnes de A sont égales, alors $\det(A) = 0$.

Démonstration. Si deux colonnes sont égales, la matrice ne change pas lorsqu'on les permute, alors que le déterminant change de signe. La seule possibilité est que $\det(A) = -\det(A)$, ce qui signifie que $\det(A) = 0$. □

Notons que si pour $k \neq l$ nous avons $C_k = \lambda C_l$, alors nous avons aussi $\det(A) = 0$.

La réciproque n'est pas vraie : il existe des matrices dont le déterminant est nul et dont aucune entrée n'est nulle. Par exemple

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}. \tag{4.131}$$

PROPooNGZJooHjtMyn

Proposition 4.85 ([132]).
Soient $A \in \mathbb{M}(n, \mathbb{K})$, et $v \in \mathbb{K}^n$. Si B est la matrice A avec la substitution $L_j \to L_j + v$ et C est la matrice A avec la substitution $L_j \to v$, alors

$$\det(B) = \det(A) + \det(C). \tag{4.132}$$

Démonstration. En utilisant l'associativité de la multiplication,

$$\begin{aligned}
\det(B) &= \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} B_{i,\sigma(i)} && \text{(4.133a)}\\
&= \sum_{\sigma} \epsilon(\sigma) \big(\prod_{i \neq j} B_{i,\sigma(i)} \big) B_{j,\sigma(j)} && \text{(4.133b)}\\
&= \sum_{\sigma} \epsilon(\sigma) \big(\prod_{i \neq j} A_{i,\sigma(i)} \big) (A_{j,\sigma(j)} + v_{\sigma(j)}) && \text{(4.133c)}\\
&= \sum_{\sigma} \epsilon(\sigma) \prod_{i} A_{i,\sigma(i)} + \sum_{\sigma} \epsilon(\sigma) \prod_{i \neq j} C_{i,\sigma(i)} v_{\sigma(j)} && \text{(4.133d)}\\
&= \det(A) + \sum_{\sigma} \epsilon(\sigma) \prod_{i \neq j} C_{i,\sigma(i)} C_{j,\sigma(j)} && \text{(4.133e)}\\
&= \det(A) + \det(C). && \text{(4.133f)}
\end{aligned}$$

SUBEQooKATCooVIbEpv

SUBEQooCOTDooPPrEYJ

Justifications :

— 4.133d parce que pour $i \neq j$ nous avons $A_{i,\sigma(i)} = C_{i,\sigma(i)}$

— 4.133e parce que $v_{\sigma(j)} = C_{j,\sigma(j)}$.

□

PROPooPYNHooLbeVhj

Proposition 4.86 (Combinaison de lignes ou colonnes $L_k \to L_k + \lambda L_l$[132]).
Soient une matrice $A \in \mathbb{M}(n, \mathbb{K})$, deux entiers $k \neq l$ inférieurs ou égaux à n.

ITEMooJSRDooTggEyO

(1) Si B est la matrice obtenue à partir de A par la substitution $L_k \to L_k + \lambda L_l$ ou $C_k \to C_k + \lambda C_l$, alors

$$\det(A) = \det(B). \tag{4.134}$$

ITEMooHKZWooVZDgnf

(2) Si B est la matrice A dans laquelle nous avons opéré la substitution $L_k \to L_k + \lambda L_l$, alors

$$B = UA \tag{4.135}$$

avec $U = \delta + \lambda E_{k,l}$, c'est-à-dire que U est une matrice de combinaison de lignes.

ITEMooPGYJooWTTghT

(3) La matrice U vérifie $\det(U) = 1$.

ITEMooBBEAooZJVGNV

(4) Nous avons

$$\det(UA) = \det(U) \det(A). \tag{4.136}$$

Démonstration. Point par point.

(1) **(1)** Soit la matrice C obtenue à partir de A par $L_k \to \lambda L_l$. En considérant le vecteur $v = \lambda L_l$, nous sommes dans la situation de la proposition 4.85. Donc

$$\det(B) = \det(A) + \det(C). \tag{4.137}$$

Mais dans la matrice C, nous avons $L_k = \lambda L_l$, ce qui implique $\det(C) = 0$ par le corolaire 4.84. Donc $\det(A) = \det(B)$ comme il se devait.

(2) **(2)** Encore un calcul :

$$(UA)_{i,j} = \sum_m \big(\delta_{i,m} + \lambda(E_{k,l})_{i,m}\big)A_{m,j} = A_{i,j} + \lambda\sum_m \delta_{k,i}\delta_{l,m}A_{m,j} = A_{i,j} + \lambda\delta_{l,i}A_{k,j}. \tag{4.138}$$

Cela donne, pour $i = k$ la ligne

$$(UA)_{k,j} = A_{k,j} + \lambda A_{l,j}, \tag{4.139}$$

ce qui correspond bien à $L_k \to L_k + \lambda L_l$.

(3) **(3)** Nous calculons le déterminant de $U = \delta + \lambda E_{k,l}$ avec $k \neq l$. Nous avons dans un premier temps :

$$\det(U) = \sum_{\sigma\in S_n} \epsilon(\sigma)\prod_{i=1}^{n}(\delta_{i,\sigma(i)} + \lambda\delta_{k,i}\delta_{l,\sigma(i)}). \tag{4.140}$$

Puisque nous avons toujours $\delta_{k,i}\delta_{l,i} = 0$, le terme $\sigma = \mathrm{Id}$ donne 1.

Pour les $\sigma \neq \mathrm{Id}$, le facteur $\lambda\delta_{k,i}\delta_{l,\sigma(i)}$ ne s'annule pas, uniquement si $i = k$ et $\sigma(i) = l$. Donc le seul terme non nul autre que $\sigma = \mathrm{Id}$ peut provenir de $\sigma = (k, l)$. Pour ce terme, nous isolons les termes $i = l$ et $i = k$:

$$(\delta_{k,\sigma(k)} + \lambda\delta_{k,k}\delta_{k,\sigma(k)})(\delta_{l,\sigma(l)} + \lambda\delta_{k,l}\delta_{k,\sigma(l)}). \tag{4.141}$$

Le dernier facteur est nul.

(4) **(4)** En mettant bout à bout les résultats prouvés,

$$\det(UA) = \det(A) = \det(U)\det(A). \tag{4.142}$$

□

PROPooXUFKooOaPnna

Proposition 4.87 (Multiplication par un scalaire d'une ligne ou colonne $L_k \to \lambda L_k$[132])**.**
Soient une matrice $A \in \mathbb{M}(n, \mathbb{K})$, un entier $k \leqslant n$. Soit la matrice B obtenue à partir de A en multipliant la ligne L_k par $\lambda \in \mathbb{K}$.

ITEMooBKIGooCDQEDt

(1) $\det(B) = \lambda\det(A)$

ITEMooWRRCooFXkRNW

(2) En considérant la matrice $T = \delta + (\lambda - 1)E_{k,k}$, nous avons

$$B = TA, \tag{4.143}$$

c'est-à-dire que la matrice T est une matrice de multiplication de ligne par un scalaire.

ITEMooOGGDooPVVRzk

(3) Nous avons $\det(T) = \lambda$.

ITEMooIFRVooWQYgkK

(4) Et aussi : $\det(TA) = \det(T)\det(A)$

Démonstration. Point par point.

(1) **(1)** La matrice B est donnée par les éléments

$$B_{i,j} = \begin{cases} A_{i,j} & \text{si } i \neq k \\ \lambda A_{i,j} & \text{si } i = k \end{cases} \tag{4.144}$$

c'est-à-dire $B_{i,j} = \big(1 + (\lambda - 1)\delta_{i,k}\big)A_{i,j}$. Nous mettons cela dans la définition du déterminant de B :

$$\det(B) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} B_{i,\sigma(i)} = \sum_{\sigma} \prod_{i} \big(1 + (\lambda - 1)\delta_{\sigma(i),k}A_{i,\sigma(i)}\big). \tag{4.145}$$

EQooGVMTooPntKew

L'associativité du produit dans $\mathbb{K}$ nous permet de séparer le produit de la façon suivante :

$$\prod_{i=1}^{n} \big(1 + (\lambda - 1)\delta_{\sigma(i),k}\big)A_{i,\sigma(i)} = \prod_{i} \big(1 + (\lambda - 1)\delta_{\sigma(i),k}\big) \prod_{i} A_{i,\sigma(i)} = \lambda \prod_{i} A_{i,\sigma(i)}. \tag{4.146}$$

En remettant dans (4.145), nous trouvons $\det(B) = \det(A)$.

(2) **(2)** C'est un cas particulier de la proposition 4.86(2) en prenant $k = l$ et en adaptant le λ.

(3) **(3)** Nous calculons le déterminant de la matrice $T = \delta + (\lambda - 1)E_{k,k}$. La formule du déterminant donne

$$\det(T) = \sum_{\sigma} \epsilon(\sigma) \prod_{i=1}^{n} \big(\delta_{i,\sigma(i)} + (\lambda - 1)\delta_{k,i}\delta_{k,\sigma(i)}\big). \tag{4.147}$$

Si $i \neq \sigma(i)$, alors non seulement $\delta_{i,\sigma(i)} = 0$, mais en plus $\delta_{k,i}\delta_{k,\sigma(i)} = 0$. Donc seul $\sigma = \mathrm{Id}$ reste dans la somme sur $\sigma \in S_n$. Il reste donc

$$\det(T) = \prod_{i=1}^{n} \big(1 + (\lambda - 1)\delta_{k,i}\big) = \left(\prod_{i\neq k} 1\right) \big(1 + (\lambda - 1)\big) = \lambda \tag{4.148}$$

où nous avons utilisé encore l'associativité pour isoler le facteur $i = k$.

(4) **(4)** Il faut mettre bout à bout les résultats déjà établis :

$$\det(TA) = \lambda \det(A) = \det(T)\det(A). \tag{4.149}$$

□

4.3.7 Réduction de Gauss

NORMooDXYZooEdnztA

4.88.
Nous avons vu les matrices d'opérations élémentaire sur les lignes :

— Permutation de lignes $L_k \leftrightarrow L_l$: $S(n;k,l) = \delta + E_{k,l} + E_{l,k} - E_{k,k} - E_{l,l}$, proposition 4.83.

— Combinaisons de lignes $L_k \to L_k + \lambda L_l$: $U(n;k,l,\lambda) = \delta + \lambda E_{k,l}$, proposition 4.86.

— Multiplication d'une ligne par un scalaire $L_k \to \lambda L_k$: $T(n;k,\lambda) = \delta + (\lambda - 1)E_{k,k}$, proposition 4.87.

Il existe également des matrices pour faire les mêmes opérations, mais sur les colonnes. Ces matrices doivent toutefois multiplier à droite et non à gauche. Par exemple il existe une matrice $C(n,k,l)$ telle que $AC(n,k,l)$ soit la matrice A, avec les colonnes k et l inversées.

Ces matrices seront dans la suite, notées G.

LEMooDXSZooVpbCRj

Lemme 4.89.
Si G est une matrice de manipulation de lignes, alors

$$\det(GA) = \det(G)\det(A) \tag{4.150}$$

EQooLQTVooBYiVVl

pour toute matrice A.

Si H est une matrice de manipulation de colonnes, alors

$$\det(AH) = \det(A)\det(H) \tag{4.151}$$

pour toute matrice A.

PROPooJBTZooNLobpf

Proposition 4.90 (Réduction de Gauss[1])**.**
Soit une matrice $A \in \mathbb{M}(n,\mathbb{K})$ de déterminant non nul : $\det(A) \neq 0$. *Alors il existe des matrices $G_1,\dots,G_N$ toutes de type S, U ou T telles que*

$$G_1 \dots G_N A = \delta. \tag{4.152}$$

Démonstration. Nous procédons par récurrence sur n. D'abord pour $n = 1$, la matrice A contient un seul élément $A_{1,1}$ qui est non nul par hypothèse. Nous pouvons multiplier sa ligne par $1/A_{1,1}$ pour obtenir le résultat. Plus précisément, nous avons l'égalité

$$T(1;1,\frac{1}{A_{1,1}})A = \delta \tag{4.153}$$

dans $\mathbb{M}(1,\mathbb{K})$. Notons que $\mathbb{K}$ est un corps (donc $A_{1,1}$ est inversible) commutatif, ce qui permet d'écrire $1/A_{1,1}$ sans ambiguïté.

Supposons le résultat prouvé pour n, et voyons ce qu'il se passe pour $n+1$. Puisque $\det(A) \neq 0$, aucune de ses colonnes n'est nulle (corolaire 4.84). Il existe donc un k tel que $A_{k,1} \neq 0$.

Par la proposition 4.83, la matrice

$$B^{(1)} = S(n+1;k,1)A \tag{4.154}$$

est une matrice telle que $B^{(1)}_{1,1} = A_{k,1} \neq 0$. Ensuite, par la proposition 4.87 la matrice

$$B^{(2)} = T(n+1;1,\frac{1}{A_{k,1}})B^{(1)} \tag{4.155}$$

vérifie $B^{(2)}_{1,1} = 1$.

Puisque la multiplication par la matrice $U(n+1;k;l;\lambda)$ réalise, par la proposition 4.86, la substitution $L_k \to L_k + \lambda L_l$, la matrice

$$B^{(3)} = \prod_{k=2}^{n+1} U(n+1;k,1,-B^{(1)}_{k,1})B^{(1)} \tag{4.156}$$

a toute sa première colonne nulle à l'exception de $B^{(3)}_{1,1} = 1$.

Nous n'avons pas donné de nom ni démontré de théorèmes à propos de la substitution $C_k \to C_k + \lambda C_l$. En passant éventuellement par les transposées et en utilisant les lemmes 4.80 et 4.81 nous obtenons une matrice $U'(n+1;k,l,\lambda)$ ayant la propriété que la matrice

$$B^{(4)} = \prod_{k=2}^{n+1} U'(n+1;k,1,-B^{(3)}_{1,k})B^{(3)} \tag{4.157}$$

vérifie $B^{(4)}_{1,j} = B^{(4)}_{j,1} = 0$ pour tout j sauf $j = 1$. En d'autres termes, la matrice $B^{(4)}$ est de la forme

$$B^{(4)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix} \tag{4.158}$$

où A' est une matrice de taille n.

Voyons quelques propriétés de A'. Nous savons que

$$B^{(4)} = \prod_i G_i A \tag{4.159}$$

où les G_i sont de type S, T ou U. Puisque $\det(SA) = \det(S)\det(A)$ (et idem pour T et U), nous avons

$$\det(B^{(4)}) = \prod_i \det(G_i)\det(A), \tag{4.160}$$

et comme aucun des $\det(G_i)$ n'est nul, nous avons encore $\det(B^{(4)}) \neq 0$, ce qui implique $\det(A') \neq 0$.

La récurrence peut avoir lieu. Il existe des matrices G'_i telles que

$$G'_1 \dots G'_M A' = \delta \tag{4.161}$$

où les G'_i sont de taille n, ainsi que le δ. En remarquant que

$$S(n+1;k,l) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & S(n;k-1,l-1) & \\ 0 & & & \end{pmatrix}, \tag{4.162}$$

et pareillement pour les matrices T et U, nous voyons qu'en prenant

$$G_i = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & G'_i & \\ 0 & & & \end{pmatrix}, \tag{4.163}$$

nous avons

$$\prod_{i=1}^{M} G_i B^{(3)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \prod_{i=1}^{M} G'_i A' & \\ 0 & & & \end{pmatrix} = \delta_{n+1} \tag{4.164}$$

où nous avons mis un indice sur le dernier δ pour être plus explicite. □

4.3.8 Matrices inversibles

PROPooMLWRooRWfZXE

Proposition-Définition 4.91.
Soit une matrice $A \in \mathbb{M}(n,\mathbb{K})$. Si les matrices B_1 et B_2 de $\mathbb{M}(n,\mathbb{K})$ vérifient

$$AB_1 = B_1A = \delta \tag{4.165}$$

et

$$AB_2 = B_2A = \delta, \tag{4.166}$$

alors $B_1 = B_2$. Dans ce cas, nous disons que A est inversible et nous notons A^{-1} l'unique matrice telle que $AA^{-1} = A^{-1}A = \delta$.

Démonstration. La preuve est réalisée dans le cas général par le lemme 1.164. Mais si vous en voulez une preuve avec les notations d'ici, en voici une.

Nous avons $AB_1 = AB_2$. En multipliant à gauche par B_1, nous trouvons $B_1AB_1 = B_1AB_2$. En remplaçant B_1A par δ des deux côtés, il reste $B_1 = B_2$. □

LEMooGZCTooQigDvC

Lemme 4.92 ([132]).
Si $A \in \mathbb{M}(n,\mathbb{K})$, alors il existe au plus une matrice $B \in \mathbb{M}(n,\mathbb{K})$ telle que $AB = \delta$.

Démonstration. Soient des matrices $B, C \in \mathbb{M}(n,\mathbb{K})$ telles que $AB = AC = \delta$. Nous allons montrer que $B = C$.

Pour cela nous considérons les applications linéaires $f_A, f_B, f_C \in \mathrm{End}(\mathbb{K}^n)$ associées par la proposition 4.71. Puisque $AB = \delta$, par la proposition 4.73, nous avons $f_A \circ f_B = f_{AB} = \mathrm{Id}$. La proposition 4.55 nous dit alors que f_A et f_B sont bijectives.

En particulier, comme $\{e_i\}$ est une base, son image par f_B est une base par la proposition 4.44. La proposition 4.54 dit alors que $\{f_B(e_i)\}$ est une base. Nous décomposons $f_B(e_k) - f_C(e_k)$ dans cette base :

$$f_B(e_k) - f_C(e_k) = \sum_j \alpha_j f_B(e_j) \tag{4.167}$$

où les α_j dépendent à priori de k. Puisque $f_A \circ (f_B - f_C) = 0$, nous avons

$$0 = f_A\big(f_B(e_k) - f_C(e_k)\big) = \sum_j (f_A \circ f_B)(e_j) = \sum_j \alpha_j e_j. \tag{4.168}$$

Donc les α_j sont tous nuls.

Nous en déduisons que $f_B(e_k) = f_C(e_k)$, et donc $f_B = f_C$. Cela implique que $B = C$ par la proposition 4.71(4). □

PROPooECIIooVMCIwz

Proposition 4.93 ([132]).
Si $A, B \in \mathbb{M}(n, \mathbb{K})$ vérifient $AB = \delta$, alors $BA = \delta$.

Démonstration. L'astuce est de poser $C = BA - \delta + B$ et de montrer que $C = B$. Pour cela, un rapide calcul commence par montrer que

$$AC = ABA - A + AB = AB = \delta. \tag{4.169}$$

Donc C est également un inverse à droite de A. Le lemme 4.92 donne alors $C = B$. □

CORooBQLXooTeVfgb

Corolaire 4.94.
Soit $A \in \mathbb{M}(n, \mathbb{K})$. Si il existe $B \in \mathbb{M}(n, \mathbb{K})$ tel que $AB = \delta$, alors A est inversible et son inverse est B.

Démonstration. Il s'agit d'une paraphrase de la proposition 4.93 et de la définition 4.91. □

LEMooZDNVooArIXzC

Lemme 4.95.
Si une matrice A n'est pas inversible, alors le produit AB n'est inversible pour aucune matrice B.

Démonstration. Soient une matrice non inversible A, ainsi qu'une matrice quelconque B. Supposons que AB soit inversible. Alors

$$AB(AB)^{-1} = \delta. \tag{4.170}$$

Donc la matrice $B(AB)^{-1}$ est un inverse de A. Contradiction. □

PROPooNPMCooPmaCwu

Proposition 4.96.
Une matrice est inversible si et seulement si son application linéaire associée est inversible. Dans ce cas, nous avons

$$f_A^{-1} = f_{A^{-1}}. \tag{4.171}$$

Démonstration. Dans le sens direct, si A est inversible nous avons $AA^{-1} = \delta$. Donc

$$f_A \circ f_{A^{-1}} = f_{AA^{-1}} = f_\delta = \mathrm{Id} \tag{4.172}$$

EQooQQOSooBKVqXb

où nous avons utilisé la proposition 4.73 pour la composition et la proposition 4.74 pour l'identité. L'égalité (4.172) indique que f_A est inversible et que son inverse est $f_{A^{-1}}$.

Dans l'autre sens, l'application f_A^{-1} existe. Soit $B \in \mathbb{M}(n, \mathbb{K})$ sa matrice. Alors nous avons

$$f_\delta = \mathrm{Id} = f_A \circ f_B = f_{AB}. \tag{4.173}$$

Le fait que l'application $\psi\colon A \to f_A$ soit une bijection [37] implique que $AB = \delta$, c'est-à-dire que A est inversible et que $B = A^{-1}$. □

37. Proposition 4.71(4).

LEMooMCIDooYBHrbq

Lemme 4.97 ([1]).
Soient une matrice inversible $A \in \mathbb{M}(n, \mathbb{R})$ et $r < n$. Il existe une permutation $\sigma \in S_n$ telle que la matrice $a \in \mathbb{M}(r, \mathbb{R})$ donnée par

$$a_{i,j} = A_{i,\sigma(j)} \tag{4.174}$$

soit inversible.

4.3.9 Inversibilité et déterminant

PROPooAVIXooMtVCet

Proposition 4.98.
Une matrice au déterminant non nul est inversible.

Démonstration. Si A est une matrice telle que $\det(A) \neq 0$, alors la proposition 4.90 nous donne des matrices $G_1, \dots, G_N$ telles que

$$G_1 \dots G_N A = \delta. \tag{4.175}$$

Donc la matrice $G_1 \dots G_N$ est un inverse de A par le corolaire 4.94. □

PROPooEOKBooKUROFg

Proposition 4.99.
Si une matrice A a une ligne ou une colonne de zéros, alors

(1) $\det(A) = 0$,

(2) A n'est pas inversible.

Démonstration. Par définition, nous avons

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}. \tag{4.176}$$

Si la k^{e} ligne est nulle, alors $A_{k,\sigma(k)} = 0$ pour tout σ. Donc tous les produits contiennent un facteur nul. Donc $\det(A) = 0$.

Pour toute matrice B nous avons

$$(AB)_{k,k} = \sum_{l} A_{k,l} B_{l,k}. \tag{4.177}$$

Si la k^{e} ligne de A est nulle nous avons $(AB)_{k,k} = 0$ et donc pas $AB = \delta$. Donc A n'est pas inversible. □

4.3.10 Quelques ensembles de matrices particuliers

Certains ensembles de matrices ont une importance particulière, que nous développerons plus tard.

Définition 4.100 (Groupe linéaire de matrices).
On note $\mathrm{GL}(n, \mathbb{A})$ *l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{A}$, qui sont inversibles. En d'autres termes,* $\mathrm{GL}(n, \mathbb{A}) = U(\mathbb{M}(n, \mathbb{A}))$.

DefMatriceOrthogonale

Définition 4.101 (Groupe orthogonal de matrices).
On dit qu'une matrice A est ***orthogonale*** *si son inverse est sa transposée, c'est-à-dire si $A^{-1} = A^t$. On note* $\mathrm{O}(n, \mathbb{A})$ *l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{A}$, qui sont orthogonales.*

4.3.11 Déterminant et combinaisons de lignes et colonnes

SUBSECOOKMSZVoBPkwQp

Proposition 4.102.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$ *telles que* $\det(A) \neq 0$ *et* $\det(B) \neq 0$. *Alors*

$$\det(AB) = \det(A)\det(B). \tag{4.178}$$

Démonstration. La proposition 4.90 nous donne des matrices de permutations de lignes et de colonnes $G_1, \dots, G_N$ et $G'_1, \dots, G'_N$ telles que[38] EQooDNZUooHBhcZj

$$G_1 \dots G_N A = \delta \tag{4.179a}$$
$$G'_1 \dots G'_N B = \delta. \tag{4.179b}$$

Nous avons

$$(G'_1 \dots G'_N)\underbrace{(G_1 \dots G_N)A}_{=\delta} B = \delta. \tag{4.180}$$

En prenant le déterminant des deux côtés et en tenant compte de (4.150),

$$1 = \det(\delta) = \det\big(G'_1 \dots G'_N G_1 \dots G_N AB\big) = \det(G'_1 \dots G'_N)\det(G_1 \dots G_N)\det(AB). \tag{4.181}$$

Mais en même temps, les équations 4.179 donnent

$$\det(G_1 \dots G_N) = \det(A)^{-1} \tag{4.182a}$$
$$\det(G'_1 \dots G'_N) = \det(B)^{-1}. \tag{4.182b}$$

Cela pour dire que

$$1 = \det(A)^{-1}\det(B)^{-1}\det(AB), \tag{4.183}$$

et donc ce qu'il nous fallait. □

PROPooWVJFooTmqoec

Proposition 4.103.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$ *telles que* $\det(A) = 0$ *et* $\det(B) \neq 0$. *Alors*

$$\det(AB) = \det(BA) = \det(A)\det(B) = 0. \tag{4.184}$$

Démonstration. Il existe des matrices de manipulations de lignes et de colonnes $G_1, \dots, G_N$ telles que $G_1 \dots G_N B = \delta$. Donc

$$0 = \det(A) = \det(G_1 \dots G_N BA) = \det(G_1 \dots G_N)\det(BA). \tag{4.185}$$

Donc $\det(BA) = 0$. □

4.3.12 Transvections

Nous nommons $E_{i,j}$ la matrice remplie de zéros sauf à la case i, j qui vaut 1. Autrement dit

$$(E_{i,j})_{k,l} = \delta_{i,k}\delta_{j,l}. \tag{4.186}$$

Définition 4.104.
Une ***matrice de transvection*** *est une matrice de la forme*

$$T_{i,j}(\lambda) = \mathrm{Id} + \lambda E_{i,j} \tag{4.187}$$

avec $i \neq j$.

38. Les plus acharnés préciseront que pour avoir le même N des deux côtés, il a fallu compléter avec des matrices δ là où il y en avait le moins.

*Une **matrice de dilatation** est une matrice de la forme*

$$D_i(\lambda) = \mathrm{Id} + (\lambda - 1)E_{i,i}. \tag{4.188}$$

Ici le $(\lambda - 1)$ sert à avoir λ et non $1 + \lambda$. C'est donc une matrice qui dilate d'un facteur λ la direction i, tout en laissant le reste inchangé.

*Si σ est une permutation (un élément du groupe symétrique S_n) alors la **matrice de permutation** associée est la matrice d'entrées*

$$(P_\sigma)_{i,j} = \delta_{i,\sigma(j)}. \tag{4.189}$$

LemyrAXQs

Lemme 4.105.
La matrice $T_{i,j}(\lambda)A = (\mathbb{1} + \lambda E_{i,j})A$ est la matrice A à qui on a effectué la substitution

$$L_i \to L_i + \lambda L_j. \tag{4.190}$$

La matrice $AT_{i,j}(\lambda)$ est la substitution

$$C_j \to C_j + \lambda C_i. \tag{4.191}$$

La matrice AP_σ est la matrice A dans laquelle nous avons permuté les colonnes avec σ.
La matrice $P_\sigma A$ est la matrice A dans laquelle nous avons permuté les lignes avec σ^{-1}.

Démonstration. Calculons la composante k, l de la matrice $E_{i,j}A$:

$$(E_{i,j}A)_{k,l} = \sum_m (E_{i,j})_{k,m}A_{m,l} \tag{4.192a}$$
$$= \sum_m \delta_{i,k}\delta_{j,m}A_{m,l} \tag{4.192b}$$
$$= \delta_{i,k}A_{j,l}. \tag{4.192c}$$

C'est donc la matrice pleine de zéros, sauf la ligne i qui est donnée par la ligne j de A. Donc effectivement la matrice

$$A + \lambda E_{i,j}A \tag{4.193}$$

est la matrice A à laquelle on a substitué la ligne i par la ligne i plus λ fois la ligne j.

En ce qui concerne l'autre assertion sur les transvections, le calcul est le même et nous obtenons

$$(AE_{i,j}) = A_{k,i}\delta_{j,l}. \tag{4.194}$$

Pour les matrices de permutation, nous avons

$$(AP_\sigma)_{k,l} = A_{k,\sigma(l)} \tag{4.195}$$

et

$$(P_\sigma A)_{k,l} = \sum_m \delta_{k,\sigma(m)}A_{m,l} = \sum_m \delta_{\sigma^{-1}(k),m}A_{m,l} = A_{\sigma^{-1}(k),l}. \tag{4.196}$$

□

4.3.13 Mineur, rang

Pour la définition du rang d'une matrice, nous en donnons une qui est clairement inspirée de l'application linéaire associée.

DEFooCSGXooFRzLRj

Définition 4.106 ([132]).
*Le **rang** d'une matrice de $\mathbb{M}(n, \mathbb{K})$ est la dimension de la partie de $\mathbb{K}^n$ engendrée par ses colonnes.*

Il est possible d'exprimer le rang d'une matrice de façon plus « intrinsèque » via le concept de mineur.

Définition 4.107 ([133]).
Les mineurs d'une matrice sont les déterminants de ses sous-matrices carrées.

Dans la suite nous désignerons souvent par le mot « mineur » la sous-matrice carrée elle-même au lieu de son déterminant.

Lorsque A est une matrice, nous notons f_A l'application linéaire associée à la matrice A par l'application (4.82).

LEMVecsaRgFixe

Lemme 4.108.
Soit $\mathbb{K}$ un corps commutatif[39]*. Si A est une matrice carrée d'ordre n et de rang r à coefficients dans $\mathbb{K}$, alors il existe des vecteurs $(x_i)_{i=1,\dots,n}$ formant une base de $\mathbb{K}^n$ tels que*

$$f_A(x_i) \neq 0 \tag{4.197}$$

pour $i \leqslant r$ et

$$f_A(x_i) = 0 \tag{4.198}$$

pour $i > r$.

Démonstration. Soit V le sous-espace de $\mathbb{K}^n$ engendré par les colonnes de A. Nous considérons la base canonique $\{e_i\}$ de $\mathbb{K}^n$, ainsi que v_i le vecteur créé par la i^{e} colonne de A. Nous avons

$$v_i = f_A(e_i). \tag{4.199}$$

Les vecteurs v_i engendrent V, donc nous pouvons en extraire une base par le théorème 4.18(1). Soit donc $\{v_j\}_{i\in J}$ une base de V avec $J \subset \{1,\dots,n\}$.

La base de $\mathbb{K}^n$ que nous cherchons commence par les vecteurs $\{e_j\}_{j\in J}$. Ces vecteurs vérifient $f_A(e_j) = v_j \neq 0$ parce que des vecteurs d'une base ne sont jamais nuls.

Pour la suite de la base, nous pourrions penser au théorème de la base incomplète[40], mais les vecteurs ainsi complétant la base ne sont pas garantis de s'annuler par f_A. Voir l'exemple 4.109.

L'idée est d'utiliser le noyau de f_A qui est un sous-espace vectoriel par la proposition 4.39. Soit une base[41] $\{z_k\}$ de $\ker(f)$. Les vecteurs $\{e_j\}_{j\in J}$ forment une base de $\text{Image}(f_A)$. Puisque les z_i forment une base de $\ker(f_A)$, le théorème du rang 4.47 dit alors que $\{e_j\}_{j\in J} \cup \{z_k\}$ est une base de $\mathbb{K}^n$.

Il y a r éléments dans J parce que l'espace engendré par les colonnes de A est de dimension r par hypothèse. Donc il y a $n-r$ éléments dans les $\{z_k\}$ pour que le tout ait le bon nombre d'éléments. □

EXooRKVQooZOGDEf

Exemple 4.109.
Soit la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}. \tag{4.200}$$

Elle est de rang 1. En suivant l'idée de la démonstration, nous commençons la base de $\mathbb{R}^2$ par le vecteur e_1 qui vérifie

$$f_A(e_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{4.201}$$

L'utilisation du théorème de la base incomplète ne permet pas de trouver un second vecteur de base v tel que $f_A(v) = 0$. En effet ce théorème donne juste l'existence d'une completion de la base, mais pas de propriétés particulières de la base obtenue. Elle pourrait donner $v = e_2$ comme second vecteur de base. Mais alors

$$f_A(v) = f_A(e_2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \neq 0. \tag{4.202}$$

39. Comme toujours.
40. Théorème 4.14(2).
41. Cette base contient $n-r$ éléments, mais ce n'est pas très important pour la suite.

Au contraire, le noyau de f_A est donné par le sous-espace engendré par $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Une base convenable est donc $\{e_1, e_1 - e_2\}$. △

PROPooEGNBooIffJXc

Proposition 4.110.
Le rang d'une application linéaire[42] *est égal au rang de sa matrice*[43] *dans n'importe quelle base.*

4.3.14 Matrices équivalentes et semblables

DefBLELooTvlHoB

Définition 4.111.
Deux relations d'équivalence entre les matrices.

ItemPFXCooOUbSCt

(1) *Deux matrices A et B sont **équivalentes** dans $\mathbb{M}(n, \mathbb{K})$ si il existe $P, Q \in \mathrm{GL}(n, \mathbb{K})$ telles que $A = PBQ^{-1}$.*

(2) *Deux matrices sont **semblables** si il existe une matrice $P \in \mathrm{GL}(n, \mathbb{K})$ telle que $A = PBP^{-1}$.*

LemZMxxnfM

Lemme 4.112.
Une matrice de rang[44] *r dans $\mathbb{M}(n, \mathbb{K})$ est équivalente à la matrice par blocs*

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.203}$$

Démonstration. Nous devons prouver que pour toute matrice $A \in \mathbb{M}(n, \mathbb{K})$ de rang r, il existe $P, Q \in \mathrm{GL}(n, \mathbb{K})$ telles que $QAP = J_r$. Soit $\{e_i\}$ la base canonique de $\mathbb{K}^n$, puis $\{f_i\}$ une base telle que $Af_i = 0$ dès que $i > r$, qui existe par le lemme 4.108.

Nous considérons la matrice inversible P telle que $Pe_i = f_i$; ses colonnes sont donc précisément les f_i, si bien que

$$APe_i = Af_i = \begin{cases} 0 & \text{si } i > r \\ \neq 0 & \text{sinon.} \end{cases} \tag{4.204}$$

La matrice AP se présente donc sous la forme

$$AP = \begin{pmatrix} M & 0 \\ * & 0 \end{pmatrix} \tag{4.205}$$

où M est une matrice $r \times r$. Nous considérons maintenant une base $\{g_i\}_{i=1,\dots,n}$ dont les r premiers éléments sont les r premières colonnes de AP et une matrice inversible Q telle que $Qg_i = e_i$. Alors

$$QAPe_i = \begin{cases} e_i & \text{si } i < r \\ 0 & \text{sinon.} \end{cases}. \tag{4.206}$$

Cela signifie que QAP est la matrice J_r. □

CorGOUYooErfOIe

Corolaire 4.113 (Équivalence et rang)**.**
Deux matrices sont équivalentes[45] *si et seulement si elles sont de même rang.*

Démonstration. D'abord il y a des implicites dans l'énoncé. Puisque nous voulons, soit par hypothèse, soit par conclusion, que les matrices A et B soient équivalentes, nous supposons qu'elles ont même dimension. Soient donc A et B deux matrices carrées d'ordre n.

Par le lemme 4.112, deux matrices de même rang r sont équivalentes à J_r. Elles sont donc équivalentes entre elles.

42. Définition 4.46.
43. Définition 4.106.
44. Définition 4.106.
45. Définition 4.111(1).

Inversement, supposons que A et B soient deux matrices équivalentes[46] : $A = PBQ^{-1}$ avec P et Q inversibles. Alors

$$\begin{aligned}\text{Image}(PBQ^{-1}) &= \{PBQ^{-1}v \text{ tel que } v \in \mathbb{K}^n\} && (4.207a)\\ &= PB\underbrace{\{Q^{-1}v \text{ tel que } v \in \mathbb{K}^n\}}_{=\mathbb{K}^n} && (4.207b)\\ &= P\big(B(\mathbb{K}^n)\big). && (4.207c)\end{aligned}$$

L'ensemble $B(\mathbb{K}^n)$ est un sous-espace vectoriel de $\mathbb{K}^n$. Comme le rang de P est maximum, la dimension de $P\big(B(\mathbb{K}^n)\big)$ est la même que celle de $B(\mathbb{K}^n)$. Par conséquent

$$\dim\Big(\text{Image}(PBQ^{-1})\Big) = \dim\big(B(\mathbb{K}^n)\big) = \text{rk}(B). \tag{4.208}$$

Le membre de gauche de cela n'est autre que $\text{rk}(A) = \dim\big(\text{Image}(PBQ^{-1})\big)$. □

4.3.15 Algorithme des facteurs invariants

PropPDfCqee

Proposition 4.114 (Algorithme des facteurs invariants[104])**.**
Soit $(\mathbb{A}, \delta)$ un anneau euclidien muni de son stathme et $U \in \mathbb{M}(n\times m, \mathbb{A})$. Alors il existe $d_1, \dots, d_s \in \mathbb{A}^*$ et des matrices $P \in \text{GL}(m, \mathbb{A})$, $Q \in \text{GL}(n, \mathbb{A})$ tels que nous ayons

$$U = P\begin{pmatrix} d_1 & & & \\ & \ddots & & 0 \\ & & d_s & \\ & 0 & & 0 \end{pmatrix}Q \tag{4.209}$$

avec $d_i \mid d_{i+1}$ pour tout i.

Démonstration. Nous allons donner la preuve plus ou moins sous forme d'algorithme.

D'abord si $U = 0$ c'est bon, on a la réponse. Sinon, nous prenons l'élément (i_0, j_0) dont le stathme est le plus petit et nous l'amenons en $(1, 1)$ par les permutations

$$\begin{aligned}C_1 &\leftrightarrow C_{j_0}\\ L_1 &\leftrightarrow L_{i_0}\end{aligned} \tag{4.210}$$

Ensuite nous traitons la première colonne jusqu'à amener des zéros partout en dessous de $u_{1,1}$ de la façon suivante : pour chaque ligne successivement nous calculons la division euclidienne

$$u_{i,1} = qu_{1,1} + r_i, \tag{4.211}$$

et nous faisons

$$L_i \to L_i - qL_1, \tag{4.212}$$

c'est-à-dire que nous enlevons le maximum possible et il reste seulement r_i en u_{i1}. Vu que le but est de ne laisser que des zéros dans la première colonne, si le reste n'est pas zéro, nous ne sommes pas contents[47]. Dans ce cas nous permutons $L_1 \leftrightarrow L_i$, ce qui aura pour effet de strictement diminuer le stathme de $u_{1,1}$ parce qu'on va mettre en $u_{1,1}$ le nombre r_i dont le stathme est strictement plus petit que celui de $u_{1,1}$.

En faisant ce jeu de division euclidienne puis échange, on diminue toujours le stathme de $u_{1,1}$, donc ça finit par s'arrêter, c'est-à-dire qu'à un certain moment, la division euclidienne de $u_{i,1}$ par $u_{1,1}$ va donner un reste nul et nous serons contents.

46. Définition 4.111.
47. Si il est zéro, nous passons à la ligne suivante

Une fois la première colonne ramenée à la forme

$$C_1 = \begin{pmatrix} u_{1,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.213}$$

nous faisons tout le même jeu avec la première ligne, en faisant maintenant des sommes divisions et permutations de colonnes. Notons que ce faisant, nous ne changeons plus la première colonne.

En fin de compte, nous trouvons une matrice [48]

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \tag{4.214}$$

Si l'élément $u_{1,1}$ ne divise pas un des éléments de A, disons $a_{i,j}$, alors nous opérons

$$C_1 \to C_1 - C_j. \tag{4.215}$$

Cela nous détruit un peu la première colonne, mais ne change pas $u_{1,1}$. Nous avons maintenant

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ * & & & \\ u_{i,j} & & A & \\ * & & & \\ 0 & & & \end{pmatrix} \tag{4.216}$$

Et nous refaisons tout le jeu depuis le début. Cependant, lorsque nous allons nous attaquer à la ligne i, $u_{1,1}$ ne divisera pas $u_{i,j}$, ce qui donnera lieu à une division euclidienne et un échange $L_1 \leftrightarrow L_i$. L'échange consistant à mettre r_i à la place de $u_{1,1}$ et réciproquement, diminuera encore strictement le stathme. Encore une fois, nous allons travailler jusqu'à avoir la matrice sous la forme

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}, \tag{4.217}$$

EqADcNVgI

sauf que cette fois, le stathme de $u_{1,1}$ est strictement plus petit que la fois précédente. Si $u_{1,1}$ ne divise toujours pas tous les éléments de A, nous recommençons encore et encore. En fin de compte, nous finissons par avoir une matrice de la forme (4.217) avec $u_{1,1}$ qui divise tous les éléments de A.

Une fois que cela est fait, il faut continuer en recommençant tout sur la matrice A. Nous avons maintenant

$$U = \begin{pmatrix} u_{1,1} & & 0 \\ & u_{2,2} & \\ 0 & & B \end{pmatrix}. \tag{4.218}$$

Sous cette forme nous avons $u_{1,1} \mid u_{2,2}$ et $u_{1,1}$ divise tous les éléments de B. En effet $u_{1,1}$ divisant tous les éléments de A, il divise toutes les combinaisons de ces éléments. Or tout l'algorithme ne consiste qu'à prendre des combinaisons d'éléments.

Nous finissons donc bien sur une matrice comme annoncée. De plus, n'ayant effectué que des combinaisons de lignes, nous avons seulement multiplié par des matrices inversibles (lemme 4.105). □

48. Nous nommons toujours par la même lettre U la matrice originale et la matrice modifiée, comme il est d'usage en informatique.

4.4 Changement de base

Soit un espace vectoriel E muni de deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$. Les deux bases sont liées entre elles par

EQooFRQRooSMsQQB
$$f_\alpha = \sum_i Q_{i,\alpha} e_i. \tag{4.219}$$

Ici Q n'est pas une application linéaire $E \to E$: Q est seulement un tableau de nombres, donnant les coordonnées des vecteurs f_α dans la base de e_i. Éventuellement Q peut être vu comme une application linéaire $\mathbb{K}^n \to \mathbb{K}^n$.

Dans la suite nous nommerons Q^{-1} la matrice inverse de Q. Inverse au sens des bêtes tableaux de nombres, sans interprétation en tant qu'application linéaire. De même pour Q^t qui est la transposée de Q.

4.4.1 Changement de base : vecteurs de base

LEMooIHZGooOZoYZd

Lemme 4.115.
Soit un espace vectoriel E sur $\mathbb{K}$ ainsi que deux bases $(e_i)_{i=1,\dots,n}$, $(f_\alpha)_{\alpha=1,\dots,n}$ de E liées par $f_\alpha = \sum_i Q_{i,\alpha} e_i$. Alors

EQooZQPAooAbKAdg
$$e_i = \sum_\alpha Q^{-1}_{\alpha,i} f_\alpha. \tag{4.220}$$

Démonstration. Nous multiplions l'égalité $f_\alpha = \sum_i Q_{i,\alpha} e_i$ par le nombre[49] $Q^{-1}_{\alpha,j} \in \mathbb{K}$ et nous sommons sur α :

$$\sum_\alpha Q^{-1}_{\alpha,j} f_\alpha = \sum_{i\alpha} (A_{i,\alpha} Q^{-1}_{\alpha,j}) e_i = e_j. \tag{4.221}$$

□

4.4.2 Changement de base : coordonnées

PROPooNYYOooHqHryX

Proposition 4.116.
Soit un espace vectoriel E sur $\mathbb{K}$. Soient deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$ liées par $f_\alpha = \sum_i Q_{i,\alpha} e_i$. Nous considérons un même vecteur dans les deux bases : $\sum_i x_i e_i = \sum_\alpha y_\alpha f_\alpha$. Alors

ITEMooIBAEooNaUnPD
(1) $y_\alpha = \sum_i Q^{-1}_{\alpha,i} x_i$

ITEMooKPWTooMwdbPu
(2) $x_i = \sum_\alpha Q_{i,\alpha} y_\alpha$.

Démonstration. Soit un vecteur $x \in E$. Il peut être écrit dans les deux bases :

EQooGSJMooGQstMx
$$x = \sum_i x_i e_i = \sum_\alpha y_\alpha f_\alpha. \tag{4.222}$$

En remplaçant e_i par sa valeur (4.220) nous avons l'égalité

$$\sum_{i\alpha} x_i Q^{-1}_{\alpha,i} f_\alpha = \sum_\alpha y_\alpha f_\alpha. \tag{4.223}$$

Puisque les f_α sont linéairement indépendants, l'égalité des sommes donne l'égalité de chacun des termes :

$$y_\alpha = \sum_i x_i Q^{-1}_{\alpha,i}. \tag{4.224}$$

En identifiant $x \in E$ au vecteur dans $\mathbb{K}^n$ de ses coordonnées dans la base $\{e_i\}$ nous pouvons écrire

$$y_\alpha = (Q^{-1} x)_\alpha, \tag{4.225}$$

Le point (1) est prouvé.

49. Attention à la bonne interprétation de ce nombre : on fait bien référence à l'élément situé en (α, j) de la matrice Q^{-1}, et pas autre chose.

En ce qui concerne le point (2), nous repartons encore de (4.222), mais nous y substituons la définition des f_α :

$$\sum_i x_i e_i = \sum_{\alpha i} y_\alpha Q_{i,\alpha} e_i. \tag{4.226}$$

Vous voulez des détails ? Allez, une étape de plus que le strict nécessaire : nous écrivons

$$\sum_i \big(x_i - \sum_\alpha y_\alpha Q_{i,\alpha}\big) e_i = 0. \tag{4.227}$$

Par linéaire indépendance des e_i, nous avons annulation de tous les coefficients, c'est-à-dire

$$x_i = \sum_\alpha Q_{i,\alpha} y_\alpha, \tag{4.228}$$

comme annoncé. □

4.117.
Attention à l'ordre des indices dans la dernière égalité : la matrice Q vient avec les indices dans l'ordre i, α, tandis que la matrice Q^{-1} vient avec les indices dans l'ordre opposé : α, i. C'est pour cela qu'il est intéressant de noter avec des lettres latines les indices se rapportant à la première base, et avec des lettres grecques ceux se rapportant à la seconde base.

NORMooNWKZooPMwYTO

4.118.
Les formules de changement de coordonnées de la proposition 4.116 s'écrivent souvent de la façon suivante :

ITEMooLHQCooBRvSlp

(1) $y_\alpha = (Q^{-1}x)_\alpha$

ITEMooNXUGooJIeoBf

(2) $y = Q^{-1}x$.

ITEMooEFILooNENamW

(3) $x_i = (Qy)_i$

ITEMooMOKHooFEJvIW

(4) $x = Qy$

Ces égalités reposent sur un petit paquet d'abus de notations qu'il convient de bien comprendre. Ici, x et y sont les éléments de $\mathbb{K}^n$ donnés par les composantes de x dans les bases $\{e_i\}$ et $\{f_\alpha\}$, et Q est vu comme une matrice, un opérateur linéaire sur $\mathbb{K}^n$. Autrement dit, le choix des bases permet d'identifier E avec $\mathbb{K}^n$ et la matrice Q avec l'application linéaire f_Q de la proposition 4.71.

4.4.3 Changement de base : matrice d'une application linéaire

PROPooNZBEooWyCXTw

Proposition 4.119.
Soit une application linéaire $t\colon E \to E$ de matrices A et B dans les bases $\{e_i\}$ et $\{f_\alpha\}$. Si les bases sont liées par

$$f_\alpha = \sum_i Q_{i,\alpha} e_i, \tag{4.229}$$

alors les matrices A et B sont liées par

$$B = Q^{-1}AQ. \tag{4.230}$$

Démonstration. L'hypothèse sur le fait que A et B sont les matrices de t signifie que pour tout $x \in E$,

$$t(x) = \sum_{ij} A_{j,i} x_i e_j = \sum_{\alpha\beta} B_{\alpha,\beta} y_\beta f_\alpha. \tag{4.231}$$

En remplaçant e_j par son expression (4.220) en termes des f_α et x_i par son expression $x_i = (Qy)_i$ (proposition 4.116), nous avons

$$(By)_\alpha = \sum_{ij\alpha} A_{j,i}(Qy)_i Q^{-1}_{\alpha,j} f_\alpha \tag{4.232a}$$

$$= \sum_{i\alpha} (Q^{-1}A)_{\alpha,i}(Qy)_i f_\alpha \tag{4.232b}$$

$$= \sum_\alpha (Q^{-1}AQy)_\alpha f_\alpha. \tag{4.232c}$$

Puisque les f_α forment une base, nous en déduisons $Q^{-1}AQy = By$. Et comme y est un élément quelconque de $\mathbb{K}^n$, nous en déduisons l'égalité de matrices

$$B = Q^{-1}AQ. \tag{4.233}$$

ooWKTYooQJfclT

□

Il s'agit bien d'une égalité de matrices ou, à la limite, d'applications linéaires sur $\mathbb{K}^n$, et non d'une égalité d'application linéaire sur E.

4.5 Espaces de polynômes

SecEspacePolynomes

Attention : les polynômes en-soi, font l'objet de la définition 1.359.

Pour chaque $k > 0$ donné, nous définissons

$$\mathcal{P}^k_{\mathbb{R}} = \{p : \mathbb{R} \to \mathbb{R} \,|\, p : x \mapsto a_0 + a_1x + a_2x^2 + \cdots + a_kx^k,\ a_i \in \mathbb{R},\ \forall i = 0, \ldots, k\}. \tag{4.234}$$

Il est facile de se convaincre que la somme de deux polynômes de degré inférieur ou égal à k est encore un polynôme de degré inférieur ou égal à k. En outre il est clair que la multiplication par un scalaire ne peut pas augmenter le degré d'un polynôme. L'ensemble $\mathcal{P}^k_{\mathbb{R}}$ est donc un espace vectoriel muni des opérations héritées de $\mathcal{P}_{\mathbb{R}}$.

La base canonique de l'espace $\mathcal{P}^k_{\mathbb{R}}$ est donnée par les monômes $\mathcal{B} = \{x \mapsto x^j \,|\, j = 0, \ldots, k\}$. Le fait que cela soit une base est vraiment facile à démontrer, et est un exercice très utile si vous ne l'avez pas encore vu dans un cours précédent.

Nous allons maintenant étudier trois applications linéaires de $\mathcal{P}^k_{\mathbb{R}}$ vers d'autres espaces vectoriels.

L'isomorphisme canonique $\phi : \mathcal{P}^k_{\mathbb{R}} \to \mathbb{R}^{k+1}$ Nous définissons ϕ par les relations suivantes

$$\phi(x^j) = e_{j+1}, \qquad \forall j \in \{0, \ldots, k\}.$$

Cela veut dire que pour tout p dans $\mathcal{P}^k_{\mathbb{R}}$, avec $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_kx^k$, l'image de p par ϕ est

$$\phi(p) = \phi\left(\sum_{j=0}^{k} a_jx^j\right) = \sum_{j=0}^{k} a_je_{j+1}.$$

Exemple 4.120.
Soit $k = 5$ on a

$$\phi(-8 - 7x - 4x^2 + 4x^3 + 2x^5) = \begin{pmatrix} -8 \\ -7 \\ -4 \\ 4 \\ 0 \\ 2 \end{pmatrix}. \tag{4.235}$$

△

Cette application est clairement bijective et respecte les opérations d'espace vectoriel, donc c'est un isomorphisme d'espaces vectoriels. L'existence d'un isomorphisme entre $\mathcal{P}^k_{\mathbb{R}}$ et $\mathbb{R}^{k+1}$ est un cas particulier du théorème qui dit que pour chaque m dans $\mathbb{N}_0$ fixée, tous les espaces vectoriels sur $\mathbb{R}$ de dimension m sont isomorphes à $\mathbb{R}^m$. Vous connaissez peut être déjà ce théorème depuis votre cours d'algèbre linéaire.

La dérivation $d : \mathcal{P}^k_{\mathbb{R}} \to \mathcal{P}^{k-1}_{\mathbb{R}}$ L'application de dérivation d fait exactement ce qu'on attend d'elle

$$d(x^0) = d(1) = 0, \qquad d(x^j) = jx^{j-1}, \quad \forall j \in \{1, \ldots, k\}.$$

Cette application n'est pas injective, parce que l'image de p ne dépend pas de la valeur de a_0, donc si deux polynômes sont les mêmes à une constante près ils auront la même image par d.

Exemple 4.121.
Soit $k = 3$ on a

$$d(-8 - 12x + 4x^3) = -12(1) + 4(3x^2) = -12 + 12x^2. \tag{4.236}$$

Noter que $d(-30 - 12x + 4x^3) = d(-8 - 12x + 4x^3)$. Cela confirme, comme mentionné plus haut, que la dérivée n'est pas injective. △

L'intégration $I : \mathcal{P}^k_{\mathbb{R}} \to \mathcal{P}^{k+1}_{\mathbb{R}}$ Nous pouvons définir une application qui est « à une constante près » l'application réciproque de la dérivation. Cette application est définie sur les éléments de base par

$$I(x^j) = \frac{x^{j+1}}{j+1}. \tag{4.237}$$

Bien entendu, la raison d'être et la motivation de cette définition apparaîtront lorsque nous développerons une théorie générale de l'intégration.

Exemple 4.122.
Soit $k = 4$ on a

$$I(6 + 2x + x^2 + x^4) = 6x + x^2 + \frac{x^3}{3} + \frac{x^5}{5}. \tag{4.238}$$

△

Remarque : étant donné que dans la définition de I nous avons décidé d'intégrer entre zéro et x, tous les polynômes dans $\mathcal{P}^{k+1}_{\mathbb{R}}$ qui sont l'image par I d'un polynôme de $\mathcal{P}^k_{\mathbb{R}}$ ont $a_0 = 0$. Cela veut dire que nous pouvons générer toute l'image de I en utilisant un sous-ensemble de la base canonique de $\mathcal{P}^{k+1}_{\mathbb{R}}$, en particulier $\mathcal{B}_1 = \{x \mapsto x^j \,|\, j = 1, \ldots, k\} \subset \mathcal{B}$ nous suffira. Cela n'est guère surprenant, parce que l'image par une application linéaire d'un espace vectoriel de dimension finie ne peut pas être un espace de dimension supérieure.

Les applications de dérivation et intégration correspondent évidemment à des applications linéaires de $\mathcal{P}_{\mathbb{R}}$ dans lui-même.

L'espace de tous les polynômes étant de dimension infinie, il peut servir de contre-exemple assez simple. Dans la sous-section 10.3.2, nous verrons que toutes les normes ne sont pas équivalentes sur l'espace des polynômes.

4.6 Projection et orthogonalité

PropProjScal

Proposition 4.123.
Si nous écrivons $\texttt{proj}_Y$ *l'opération de projection sur la droite qui sous-tend* Y*, alors nous avons*

$$\| \texttt{proj}_Y X\| = \frac{X \cdot Y}{\|Y\|}. \tag{4.239}$$

Démonstration. Les vecteurs X et Y sont des flèches dans l'espace. Nous pouvons choisir un système d'axe orthogonal tel que les coordonnées de X et Y soient

$$X = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} l \\ 0 \\ 0 \end{pmatrix} \tag{4.240}$$

où l est la longueur du vecteur Y. Pour ce faire, il suffit de mettre le premier axe le long de Y, le second dans le plan qui contient X et Y, et enfin le troisième axe dans le plan perpendiculaire aux deux premiers.

Un simple calcul montre que $X \cdot Y = xl + y \cdot 0 + 0 \cdot 0 = xl$. Par ailleurs, nous avons $\| \texttt{proj}_Y X\| = x$. Par conséquent,

$$\| \texttt{proj}_Y X\| = \frac{X \cdot Y}{l} = \frac{X \cdot Y}{\|Y\|}. \tag{4.241}$$

□

Corolaire 4.124.
Si la norme de Y est 1, alors le nombre $X \cdot Y$ est la longueur de la projection de X sur Y.

Démonstration. Poser $\|Y\| = 1$ dans la proposition 4.123. □

4.7 Dualité

PropEJBZooTNFPRj

Proposition 4.125.
Si A est la matrice d'une application linéaire, alors le rang de cette application linéaire est égal au rang de A, c'est-à-dire à la taille de la plus grande matrice carrée de déterminant non nul contenue dans A.

DefJPGSHpn

Définition 4.126.
Soit E un espace vectoriel sur $\mathbb{K}$.

*Une **forme linéaire** sur E est une application linéaire de E sur son corps de base $\mathbb{K}$.*

*Le **dual algébrique** de E, noté E^*, est l'ensemble des formes linéaires sur E. Autrement dit : $E^* = \mathcal{L}(E, \mathbb{K})$.*

Nous verrons les formes multilinéaires en la définition 8.1.

Nous verrons plus tard qu'en dimension infinie, les applications linéaires ne sont pas toujours continues. Nous définirons donc aussi un concept de dual topologique. Voir la proposition 10.70, la remarque 10.73 et la définition 25.1.

DEFooTMSEooZFtsqa

Lemme-Définition 4.127.
Si E est un espace vectoriel et si $\{b_i\}$ est une base de E, alors nous définissons

$$\begin{aligned} b_i^* \colon E &\to \mathbb{K} \\ b_j &\mapsto \delta_{ij}, \end{aligned} \tag{4.242}$$

*et sa prolongation par linéarité. Ces éléments du dual E^* forment une base appelée **base duale**.*

En particulier nous avons

$$\dim(E) = \dim(E^*). \tag{4.243}$$

Notons que si $v \in E$ est un vecteur, ça n'a aucun sens à priori de parler de v^*. Il s'agit bien de définir *toute* la base $\{e_i^*\}$ à partir de toute la base $\{e_i\}$.

LEMooSVRIooFbxfue

Lemme 4.128 (Changement de base duale[1])**.**
Soient un espace vectoriel V de dimension finie, une base $\{e_i\}_{i=1,\dots,n}$ et sa base duale $\{\alpha_i\}_{i=1,\dots,n}$. Soit une bijection linéaire $A \colon V \to V$. Nous considérons $\{\beta_i\}_{i=1,\dots,n}$ la base duale de la base $\{Ae_i\}$.

Alors

$$\beta_i = \sum_j A_{ij}^{-1} \alpha_j. \tag{4.244}$$

Démonstration. Par définition d'une base duale, il faut $\beta_i(Ae_k) = \delta_{ik}$. Vu que $\{\alpha_i\}$ est une base, nous pouvons décomposer β_i de la façon suivante : $\beta_i = \sum_j B_{ij}\alpha_j$. Il nous reste à prouver que $B = A^{-1}$.

Pour cela nous calculons un peu :

$$\delta_{ik} = \beta_i(Ae_k) = \sum_j B_{ij}\alpha_j(Ae_k) = \sum_{lj} B_{ij}A_{lk} \underbrace{\alpha_j(e_l)}_{=\delta_{jl}} = \sum_l B_{il}A_{lk} = (BA)_{ik}. \tag{4.245}$$

Nous avons donc $\delta_{ik} = (BA)_{ik}$, c'est-à-dire $BA = \mathbb{1}$ ou encore $B = A^{-1}$. □

LEMooQLWNooYUpGdo

Lemme 4.129 ([134])**.**
Soit un espace vectoriel E de dimension finie.

ITEMooHHTLooNCjgfn

(1) Si α est une forme linéaire non nulle sur E, alors il existe $x \in E$ tel que $\alpha(x) = 1$.

ITEMooBYAAooUWBKDk

(2) Si $x \neq 0$ dans E, alors il existe une forme linéaire α sur E telle que $\alpha(x) = 1$.

Démonstration. En deux parties.

(1) **Pour (1)** Puisque α est non nulle, nous pouvons considérer $v \in E$ tel que $\alpha(v) \neq 0$. Alors en posant $x = \alpha(v)^{-1}v$, nous avons le résultat.

(2) **Pour (2)** Soit un vecteur non nul que nous écrivons sous la forme $x = \sum_{i=1}^{n} x_i e_i$ (pour une certaine base $\{e_i\}_{i=1,\dots,n}$ de E). Supposons que $x_k \neq 0$. Alors la forme

$$\begin{aligned} \alpha\colon E &\to \mathbb{K} \\ y &\mapsto y_k/x_k \end{aligned} \tag{4.246}$$

fait l'affaire.

□

LEMooKTREooBrnWVz

Lemme 4.130.
Soit un espace vectoriel de dimension finie E sur le corps $\mathbb{K}$. Si $\{\alpha_1, \dots, \alpha_n\}$ est une base de E^*, alors l'application

$$\begin{aligned} \Phi\colon E &\to \mathbb{K}^n \\ x &\mapsto \big(\alpha_1(x), \dots, \alpha_n(x)\big) \end{aligned} \tag{4.247}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. En deux parties.

(1) **Φ est injective** Soit $z \in \ker(\Phi)$. Nous avons $\alpha_i(z) = 0$ pour tout i. Si $z \neq 0$, alors le lemme 4.129 dit qu'il existe $\beta \in E^*$ tel que $\beta(z) \neq 0$.
Décomposons un tel β dans la base de $\{\alpha_i\}$:

$$\beta = \sum_{i=1}^{n} \beta_i \alpha_i. \tag{4.248}$$

Alors nous avons

$$0 \neq \beta(z) = \sum_{i=1}^{n} \beta_i \underbrace{\alpha_i(z)}_{=0} = 0. \tag{4.249}$$

Contradiction. Donc $\ker(\Phi) = \{0\}$ et Φ est injective.

(2) **Φ est surjective** Les espaces vectoriels E, E^* et $\mathbb{K}^n$ ont tout trois, une dimension n. Donc Φ est une application linéaire injective entre deux espaces de même dimension. Elle est donc surjective par le corolaire 4.49.

□

PROPooDBPGooPagbEB

Proposition-Définition 4.131 ([134])**.**
Soit un espace vectoriel E de dimension finie sur le corps $\mathbb{K}$. Toute base du dual E^* est duale d'une unique base de E. Cette base est dite ***préduale****.*

Démonstration. Nous considérons une base $\mathcal{F} = \{\alpha_1, \dots, \alpha_n\}$ de E^*. Nous devons prouver qu'il existe une unique base de E dont la base duale est $\mathcal{F}$.

(1) **Existence** Le lemme 4.130 nous indique que l'application

$$\begin{aligned} \Phi\colon E &\to \mathbb{K}^n \\ x &\mapsto \big(\alpha_1(x), \dots, \alpha_n(x)\big) \end{aligned} \tag{4.250}$$

est un isomorphisme d'espaces vectoriels.

Soit la base canonique de $\mathbb{K}^n$: $(\epsilon_1, \dots, \epsilon_n)$. Puisque Φ est un isomorphisme, $\big(\Phi^{-1}(\epsilon_i)\big)_{i=1,\dots,n}$ est une base de E. Nous allons montrer qu'elle est préduale de (α_i). Nous posons $e_i = \Phi^{-1}(\epsilon_i)$ et nous calculons :

$$\begin{aligned}
\alpha_i(e_j) &= \alpha_i\big(\Phi^{-1}(\epsilon_j)\big) && \text{(4.251a)}\\
&= \Phi\big(\Phi^{-1}(\epsilon_j)\big)_i && \text{(4.251b)}\\
&= (\epsilon_i)_j && \text{(4.251c)}\\
&= \delta_{i,j} && \text{(4.251d)}
\end{aligned}$$

SUBEQooACVAooVNwzMq

SUBEQooOYNHooSVOOyz

Justifications :

— Pour 4.251b, nous remarquons que $\alpha_i(x) = \Phi(x)_i$.

— Pour 4.251d, nous utilisons le fait que les ϵ_j forment la base canonique de $\mathbb{K}^n$.

(2) **Unicité** Soit une base préduale (e_i) de (α_i). Nous avons, par définition, que $\alpha_i(e_j) = \delta_{i,j}$. Donc

$$\big(\alpha_1(e_j), \dots, \alpha_n(e_j)\big) = \epsilon_j. \tag{4.252}$$

Nous appliquons Φ^{-1} à cette dernière équation pour obtenir $e_j = \Phi^{-1}(\epsilon_j)$. Donc les e_j sont déterminés de façon unique à partir des α_i.

□

4.7.1 Changement de base duale

PROPooDOEYooDpUbnc

Proposition 4.132 ([1]).
Soient deux bases $\{e_i\}$ et $\{e'j\}$ de l'espace vectoriel V. Nous notons $\{\alpha_i\}$ et $\{\alpha'_i\}$ les bases duales. Si $e'_j = \sum_i A_{ij}e_i$, alors

$$\alpha'_j = \sum_k A^{-1}_{jk}\alpha_k. \tag{4.253}$$

Démonstration. Nous posons $\alpha'_j = \sum_k B_{kj}\alpha_k$ et nous appliquons ça à e'_i. À gauche nous avons $\alpha'_j(e'_i) = \delta_{ij}$, et à droite,

$$\begin{aligned}
\sum_k B_{kj}\alpha_k(e'_i) &= \sum_k B_{kj}\alpha_k\big(\sum_l A_{li}e_l\big) && \text{(4.254a)}\\
&= \sum_{kl} B_{kj}A_{li}\delta_{kl} && \text{(4.254b)}\\
&= \sum_l B_{lj}A_{li} && \text{(4.254c)}\\
&= (B^tA)_{ji}. && \text{(4.254d)}
\end{aligned}$$

Nous avons donc montré que $B^tA = \mathbb{1}$, ou encore que $B^t = A^{-1}$, ce qu'il fallait. □

4.7.2 Orthogonal

DEFooEQSMooHVzbfz

Définition 4.133.
*Soit E, un espace vectoriel, et F un sous-espace de E. L'**orthogonal** de F est la partie $F^\perp \subset E^*$ donnée par*

$$F^\perp = \{\alpha \in E^* \text{ tel que } \forall x \in F, \alpha(x) = 0\}. \tag{4.255}$$

Equivple

Cette définition d'orthogonal via le dual n'est pas du pur snobisme. En effet, la définition « usuelle » qui ne parle pas de dual,

$$F^\perp = \{y \in E \text{ tel que } \forall x \in F, y \cdot x = 0\}, \tag{4.256}$$

demande la donnée d'un produit scalaire. Évidemment dans le cas de $\mathbb{R}^n$ muni du produit scalaire usuel et de l'identification usuelle entre $\mathbb{R}^n$ et $(\mathbb{R}^n)^*$ via une base, les deux notions d'orthogonal coïncident.

La définition 4.133, au contraire, est intrinsèque : elle ne dépend que de la structure d'espace vectoriel.

Si $B \subset E^*$, on note B^o son orthogonal :

$$B^o = \{x \in E \text{ tel que } \omega(x) = 0, \forall \omega \in B\}. \tag{4.257}$$

Notons qu'on le note B^o et non $B^\perp$ parce qu'on veut un peu s'abstraire du fait que $(E^*)^* = E$. Du coup on impose que B soit dans un dual, et on prend une notation précise pour dire qu'on remonte au pré-dual, et non qu'on va au dual du dual.

PropXrTDIi

Proposition 4.134.
Soient un espace vectoriel E et F, un sous-espace vectoriel de E. Nous avons

$$\dim F + \dim F^\perp = \dim E. \tag{4.258}$$

Démonstration. Soit $\{e_1, \dots, e_p\}$ une base de F que nous complétons en une base $\{e_1, \dots, e_n\}$ de E par le théorème 4.14. Soit $\{e_1^*, \dots, e_n^*\}$ la base duale. Alors nous prouvons que $\{e_{p+1}^*, \dots, e_n^*\}$ est une base de $F^\perp$.

D'abord, ce sont des éléments de $F^\perp$, parce que si $i \leqslant p$ et si $k \geqslant 1$, nous avons $e_{p+k}^*(e_i) = 0$; donc oui, $e_{p+k}^* \in F^\perp$.

Ensuite, en tant que partie d'une base de F^*, c'est une partie libre. Il reste à montrer que c'est générateur.

Enfin $F^\perp \subset \text{Span}\{e_k^*, k \in \{p+1, \dots, n\}\}$ parce que si $\omega = \sum_{k=1}^n \omega_k e_k^*$, alors $\omega(e_i) = \omega_i$, mais nous savons que si $\omega \in F^\perp$, alors $\omega(e_i) = 0$ pour $i \leqslant p$. Donc $\omega = \sum_{k=p+1}^n \omega_k e_k^*$. □

La proposition 8.185 donnera une version plus terre à terre de la proposition 4.134 en disant que si nous avons un produit scalaire, alors $E = F \oplus F^\perp$ où $F^\perp$ est cette fois défini comme l'orthogonal pour le produit scalaire.

4.8 Représentation de groupe

DEFooXVMSooXDIfZV

Définition 4.135 (Représentation)**.**
*Soit un groupe G. Une **représentation** de G est un couple (E, ρ) où E est un espace vectoriel et ρ est une application $\rho\colon G \to \text{GL}(E)$ vérifiant*

$$\rho(g) \circ \rho(h) = \rho(gh). \tag{4.259}$$

pour tout $g, h \in G$.

Définition 4.136.
Une représentation[50] *$\rho\colon G \to \text{GL}(E)$ est **fidèle** si elle est injective*
*La dimension de E est le **degré** de la représentation (E, ρ).*

Le fait que la représentation $\rho\colon G \to \text{GL}(E)$ soit fidèle ne dit pas que chacun des $\rho(g)$ est injectif.

PROPooHNQOooSzeEFG

Proposition 4.137.
Soit un corps $\mathbb{K}$. Si G est un groupe dans $\mathbb{M}(n, \mathbb{K})$ (c'est-à-dire un groupe de matrices $n \times n$ à coefficients dans $\mathbb{K}$), alors l'application

$$\begin{aligned} \rho\colon G &\to \text{GL}(\mathbb{K}^n) \\ A &\mapsto f_A \end{aligned} \tag{4.260}$$

où f_A est l'application linéaire associée à A, est une représentation de G.

50. Définition 4.135.

Démonstration. La représentation dont nous parons n'est autre que l'application ψ de la définition 4.68, dont nous connaissons beaucoup de propriétés. La proposition 4.74 dit, entre autres, que $\psi(AB) = \psi(A)\psi(B)$, c'est-à-dire que $\rho(AB) = \rho(A)\rho(B)$, et donc que ρ est une représentation. □

4.9 Somme directe d'espaces vectoriels

Si V et W sont des espaces vectoriels, ce que nous notons $V \oplus W$ n'est rien d'autre que l'espace vectoriel de l'ensemble $V \times W$.

DEFooJKAWooKkkkwm

Proposition-Définition 4.138 ([135, 136])**.**
Si V et W sont des espaces vectoriels sur le même corps $\mathbb{K}$, alors les définitions

(1) $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$

(2) $\lambda(v, w) = (\lambda v, \lambda w)$

donnent une structure d'espace vectoriel sur $V \times W$.

Cet espace sera noté $V \oplus W$ et est appelé ***somme directe*** *de V et W.*

DEFooIJDNooRUDUYF

Définition 4.139 (Sous-espaces en somme directe[137])**.**
Soient un espace vectoriel E sur $\mathbb{K}$ ainsi que des sous-espaces vectoriels $\{F_i\}_{i\in I}$ (I est un ensemble fini ou infini). Nous disons que les F_i sont ***en somme directe*** *si pour tout élément $u \in \sum_{i\in I} F_i$, il existe un unique ensemble $\{u_i\}_{i\in I}$ de vecteurs tel que*

(1) $u = \sum_{i\in I} u_i$

(2) $u_i \in F_i$ *pour tout i,*

(3) $\{j \in I \text{ tel que } u_j \neq 0\}$ *est fini.*

Si l'espace vectoriel E est un espace vectoriel topologique, nous avons la définition 7.175 qui donne des conditions de compatibilité entre les topologies.

LEMooHWRVooLedAmF

Lemme 4.140.
Soit un espace vectoriel E et deux sous-espaces F_1, F_2 en somme directe : $E = F_1 \oplus F_2$. Alors l'application

$$\begin{aligned}\psi\colon F_1 \times F_2 &\to E\\ (x, y) &\mapsto x + y\end{aligned} \tag{4.261}$$

est une bijection.

Démonstration. L'application est injective parce que si $\psi(x_1, x_2) = \psi(y_1, y_2)$, alors $x_1 + x_2 = y_1 + y_2$. Nommons u ce vecteur. Par unicité de l'ensemble $\{u_1, u_2\}$ tel que $u = u_1 + u_2$, nous avons automatiquement $\{x_1, x_2\} = \{y_1, y_2\}$ et donc $x_1 = y_1$ et $x_2 = y_2$.

En ce qui concerne la surjectivité, si $u \in E$, il existe un ensemble $\{u_1, u_2\}$ avec $u_i \in F_i$ tel que $u = u_1 + u_2 = \psi(u_1, u_2)$. □

LEMooDQMQooInVVDY

Lemme 4.141 ([1, 137])**.**
Soient un espace vectoriel E ainsi que des sous-espaces vectoriels F_i. Nous avons équivalence entre les assertions suivantes.

(1) *Les F_i sont en somme directe*[51]*.*

(2) *Si $\sum_{i\in I} u_i = 0$ avec $u_i \in F_i$ et si $\{j \in I \text{ tel que } u_j \neq 0\}$ est fini, alors tous les u_i sont nuls.*

(3) *Chaque espace F_k est en somme directe avec la somme des précédents, c'est-à-dire que pour tout k,*

$$\left(\sum_{i=1}^{k-1} F_i\right) \cap F_k = \{0\}. \tag{4.262}$$

51. Définition 4.139.

ITEMooPLXGooCOQgen

(4) *Pour tout k,*

$$F_k \cap \left(\sum_{i \neq k} F_i \right) = \{0\}. \tag{4.263}$$

PROPooCASNooEqisqa

Proposition 4.142 ([138])**.**
Soient E un espace vectoriel de dimension finie, et deux sous-espaces F_1 et F_2 satisfaisant

(1) $F_1 \cap F_2 = \{0\}$,

(2) $\dim(F_1) + \dim(F_2) \geqslant \dim(E)$.

Alors $E = F_1 \oplus F_2$.

Démonstration. Soient une base $\{e_i\}_{i\in I}$ de F_1 et $\{f_\alpha\}$ de F_2. Nous commençons par prouver que la partie $B = \{e_i\} \cup \{f_\alpha\}$ est libre.

Supposons en effet, avoir des coefficients a_i et b_α tels que

$$\sum_i a_i e_i + \sum_\alpha b_\alpha f_\alpha = 0. \tag{4.264}$$

Cela implique que $\sum_i a_i e_i = -\sum_\alpha b_\alpha f_\alpha$. Or $\sum_i a_i e_i \in F_1$ et $-\sum_\alpha b_\alpha f_\alpha \in F_2$. Donc les éléments $\sum_i a_i e_i$ et $\sum_\alpha b_\alpha f_\alpha$ sont dans $F_1 \cap F_2 = \{0\}$. Nous avons alors les égalités

$$\sum_i a_i e_i = 0 \tag{4.265}$$

et

$$\sum_\alpha b_\alpha f_\alpha = 0. \tag{4.266}$$

La première implique $a_i = 0$ pour tout i et la seconde implique $b_\alpha = 0$ pour tout α.

Donc B est une partie libre de E contenant $\dim(F_1) + \dim(F_2) \geqslant \dim(E)$ éléments. La proposition 4.19(1) nous indique alors qu'en réalité $\dim(F_1) + \dim(F_2) = \dim(E)$. Comme B est une partie libre contenant $\dim(E)$ éléments, c'est une base par la proposition 4.19(2). □

4.9.1 Structure réelle

DEFooCIFSooVmcNtE

Proposition-Définition 4.143 ([139, 136, 1])**.**
Soit un espace vectoriel E sur $\mathbb{C}$. Il existe une application $\sigma : E \to E$ telle que

(1) $\sigma^2 = \mathrm{Id}_E$

(2) *Pour tout $\alpha, \beta \in \mathbb{R}$, $f(\alpha x + \beta y) = \bar{\alpha}\sigma(x) + \bar{\beta}\sigma(y)$.*

Une telle application est une ***structure réelle*** *sur E.*

Démonstration. La proposition 4.24 nous permet de considérer une base $\{e_i\}_{i\in I}$ de E. Alors, nous définissons

$$\sigma\Big(\sum_{i\in I} \lambda_i e_i\Big) = \sum_{i\in I} \bar{\lambda}_i e_i. \tag{4.267}$$

Notez que la somme est toujours finie. □

PROPooPZHPooNdarzg

Proposition 4.144.
Soit un espace vectoriel E sur $\mathbb{C}$ et une structure réelle[52] *σ sur E. Nous posons*

$$E_{\mathbb{R}} = \{v \in E \text{ tel que } \sigma(v) = v\}. \tag{4.268}$$

Alors

(1) *La partie $E_{\mathbb{R}}$ est un espace vectoriel réel.*

52. Définition 4.143.

(2) Nous avons la décomposition en somme directe[53]

$$E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}. \tag{4.269}$$

Démonstration. En plusieurs parties.

(1) **Espace vectoriel réel** Si $v, w \in E_{\mathbb{R}}$, alors

$$\sigma(v + w) = \sigma(v) + \sigma(w) = v + w, \tag{4.270}$$

et si $\lambda \in \mathbb{R}$,

$$\sigma(\lambda x) = \lambda\sigma(x) = \lambda x. \tag{4.271}$$

Donc $E_{\mathbb{R}}$ est un espace vectoriel réel.

(2) **Première somme directe** Nous définissons

$$E^+ = \{v \in E \text{ tel que } \sigma(v) = v\}, \tag{4.272a}$$
$$E^- = \{v \in E \text{ tel que } \sigma(v) = -v\} \tag{4.272b}$$

Nous prouvons que

$$\begin{aligned} \psi\colon E^+ \times E^- &\to E \\ (a, b) &\mapsto a + b \end{aligned} \tag{4.273}$$

est un isomorphisme d'espaces vectoriels.

Puisque ψ est linéaire, il suffit de prouver qu'elle est bijective.

(2a) **Surjectif** Si $v \in E$, alors en posant $v_+ = \frac{1}{2}\big(v + \sigma(v)\big)$, $v_- = \frac{1}{2}\big(v - \sigma(v)\big)$, nous avons

$$v = v_+ + v_-, \quad v_+ \in E^+, \quad v_- \in E^-, \tag{4.274}$$

et donc $v = \psi(v_+, v_-)$.

(2b) **Injectif** Supposons $\psi(a, b) = \psi(\alpha, \beta)$. Alors $a + b = \alpha + \beta$ et donc $a - \alpha = \beta - b$. Comme $a - \alpha \in E^+$ et $\beta - b \in E^-$, nous savons que $a - \alpha = \beta - b \in E^+ \cap E^-$. Étant donné que $E^+ \cap E^- = \{0\}$, nous avons $a - \alpha = \beta - b = 0$.

Nous avons donc la somme directe $E = E^+ \oplus E^-$.

(3) **Conclusion** Par définition, $E^+ = E_{\mathbb{R}}$. Il nous reste à voir que $E^- = iE^+$. Nous prouvons les inclusions dans les deux sens.

(3a) $E^- \subset iE^+$ Soit $v \in E^-$. Nous avons $iv \in E^+$; en effet

$$\sigma(iv) = \bar{i}\sigma(v) = -i\sigma(v) = iv. \tag{4.275}$$

Donc $iv \in E^+$ pour $v \in E^-$.

(3b) $iE^+ \subset E^-$ Soit $v \in E^+$, et voyons que $iv \in E^-$. En effet,

$$\sigma(iv) = -i\sigma(v) = -iv. \tag{4.276}$$

□

4.145.
Lorsque nous avons une structure réelle σ sur un espace vectoriel complexe E, nous écrivons $E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ sans préciser dans la notation « $E_{\mathbb{R}}$ » que cet ensemble dépend du choix de σ. En particulier si F est un sous-espace vectoriel de E, nous utiliserons la notation $F_{\mathbb{R}}$ relativement à la même involution que celle utilisée pour E.

Lemme 4.146.
Soit un espace vectoriel complexe E muni d'une structure réelle σ. Si F est un sous-espace de E alors $F_{\mathbb{R}} = E_{\mathbb{R}} \cap F$.

53. Définition 4.138.

Démonstration. Par définition,

$$F_{\mathbb{R}} = \{v \in F \text{ tel que } \sigma(v) = v\}. \tag{4.277}$$

(1) $\underline{F_{\mathbb{R}} \subset F}$ C'est dans la définition de $F_{\mathbb{R}}$ (sous-ensemble de F).

(2) $\underline{F_{\mathbb{R}} \subset E_{\mathbb{R}}}$ Si $v \in F_{\mathbb{R}}$, alors $\sigma(v) = v$. Mais cette égalité est précisément celle qui permet d'être dans $E_{\mathbb{R}}$.

□

Vous remarquerez que ce lemme ne fonctionne que parce que nous avons choisi la même structure réelle sur F que sur E.

Chapitre 5

Classification de certains groupes

5.1 Théorèmes de Sylow

Lemme 5.1.
Soient H et K des sous-groupes finis de G. Alors

$$\mathrm{Card}(HK) = \frac{|H| \cdot |K|}{|H \cap K|}. \tag{5.1}$$

Attention : dans ce lemme, l'ensemble HK n'est pas spécialement un groupe. Ce serait le cas si H normalisait K, c'est-à-dire si nous avions $hkh^{-1} \in K$, $\forall (h,k) \in H \times K$.

ThoCauchyGpFini

Théorème 5.2 (Théorème de Cauchy[140])**.**
Soit G un groupe fini et p un nombre premier divisant $|G|$. Alors

(1) G contient un élément d'ordre p.

(2) Si G est un p-groupe, il existe un élément central d'ordre p dans G.

ThoIfdlEB

Lemme 5.3 (Théorème de Cayley)**.**
Si G est un groupe d'ordre n alors il est isomorphe à un sous-groupe du groupe symétrique S_n.

Démonstration. L'action à gauche de G sur lui-même

$$\begin{aligned} \varphi\colon G &\to S_n \\ \varphi(x)g &\mapsto xg \end{aligned} \tag{5.2}$$

est une permutation des éléments de G. Cela donne un morphisme injectif parce que si $\varphi(x) = \varphi(y)$ nous avons $xg = yg$ pour tout g et en particulier pour $g = e$ nous trouvons $x = y$. □

Pour rappel, lorsque p est premier, nous notons $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

LemaQxjcm

Lemme 5.4.
Soit p un diviseur premier de n. Alors il existe un morphisme injectif du groupe symétrique S_n dans $\mathrm{GL}(n, \mathbb{F}_p)$.

Démonstration. Soit $\{e_i\}$ la base canonique de $\mathbb{F}_p^n$. Par exemple $e_1 = \big([1]_p, [0]_p, \dots, [0]_p\big)$. Nous avons le morphisme injectif $\varphi\colon S_n \to \mathrm{GL}(n, \mathbb{F}_p)$ donné par $\varphi(\sigma)e_i = e_{\sigma(i)}$. □

RemFzxxst

Remarque 5.5.
En mettant bout à bout les lemmes 5.3 et 5.4, nous trouvons que si p est un diviseur premier de $|G|$, alors G peut être vu comme un sous-groupe de $\mathrm{GL}(n, \mathbb{F}_p)$.

DEFooPRCHooVZdwST

Définition 5.6.
*Soit p un nombre premier. Un p-**groupe** est un groupe dont tous les éléments sont d'ordre p^m pour un certain m (dépendant de l'élément).*

*Soit G un groupe fini et p, un diviseur premier de $|G|$. Un p-**Sylow** dans G est un p-sous-groupe d'ordre p^n où p^n est la plus grande puissance de p divisant $|G|$.*

Notons que si p est un nombre premier, alors tout groupe d'ordre p^m est un p-groupe.

Lemme 5.7.
Soit G un groupe fini et P, Q des p-sous-groupes. Nous supposons que Q normalise P. Alors PQ est un p-sous-groupe de G.

Si S est un p-Sylow, alors p ne divise pas le nombre $|G : S| = |G|/|S|$.

Propvocmon

Proposition 5.8.
Soit le corps fini $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ (p premier). Soit T le sous-ensemble de $\mathrm{GL}_n(\mathbb{F}_p)$ formé des matrices triangulaires supérieures de rang[1] n et dont les éléments diagonaux sont 1. Alors T est un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$.

Démonstration. Nous commençons par étudier le cardinal de $\mathrm{GL}_n(\mathbb{F}_p)$. Pour la première colonne, la seule contrainte à vérifier est qu'elle ne soit pas nulle. Il y a donc $p^n - 1$ possibilités. Pour la seconde, il faut ne pas être multiple de la première. Il y a donc $p^n - p$ possibilités (parce qu'il y a p multiples possibles de la premières colonne). Pour la k-ième colonne, il faut éviter toutes les combinaisons linéaires des $(k-1)$ premières colonnes. Il y a p^{k-1} telles combinaisons et donc $p^n - p^{k-1}$ possibilités pour la k-ième colonne. Nous avons donc

$$\begin{aligned}\mathrm{Card}\big(\mathrm{GL}(n,\mathbb{F}_p)\big) &= (p^n-1)(p^n-p)\dots(p^n-p^{n-1}) && (5.3a)\\ &= p\cdot p^2\cdots p^{n-1}(p^n-1)(p^{n-1}-1)\dots(p-1) && (5.3b)\\ &= p^{\frac{n(n-1)}{2}}m && (5.3c)\end{aligned}$$

où m est un entier qui ne divise pas p.

En ce qui concerne le cardinal de T, le calcul est plus simple : pour la première ligne nous avons p^{n-1} choix (parce qu'il y a un 1 qui est imposé sur la diagonale), pour la seconde p^{n-2}, etc. En tout nous avons alors

$$|T| = p^{\frac{n(n-1)}{2}}, \tag{5.4}$$

et T est un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$. □

Proposition 5.9.
Soit p un nombre premier. Un groupe fini G est un p-groupe si et seulement l'ordre de G est p^n pour un certain n.

Démonstration. Supposons que G est un p-groupe. Soit q un nombre premier divisant $|G|$. Par le théorème de Cauchy (5.2), le groupe G contient un élément d'ordre q, soit g un tel élément. Étant donné que G est un p-groupe, $g^{p^n} = g^q = e$ pour un certain n. Donc $q = p^n$ et $q = p$ parce que q est premier. Nous venons de prouver que p est le seul nombre premier qui divise $|G|$. L'ordre de G est par conséquent une puissance de p.

Nous nous intéressons maintenant à l'implication inverse. Nous supposons que $|G| = p^n$ pour un certain entier $n \geqslant 0$. Soit $g \in G$; nous notons r l'ordre de G. Le sous-groupe $\mathrm{gr}(g)$ est d'ordre r, donc r divise $|G|$ (par le théorème 2.13 de Lagrange). Le nombre r est alors une puissance de p. □

LemwDYQMg

Lemme 5.10.
Soit G, un groupe fini de cardinal $|G| = n$ et p, un diviseur premier de n. Nous notons $n = p^m \cdot r$ où p ne divise pas r. Soit H un sous-groupe de G et S, un p-Sylow de G. Alors il existe $g \in G$ tel que

$$gSg^{-1} \cap H \tag{5.5}$$

soit un p-Sylow de H.

1. Définition 4.46.

Démonstration. Nous considérons l'ensemble G/S sur lequel H agit. Si $a \in G$, le stabilisateur de $[a]$ dans G/S est

$$\begin{aligned}\mathrm{Fix}\big([a]\big) &= \{h \in H \text{ tel que } [ha] = [a]\} && (5.6a)\\ &= \{h \in H \text{ tel que } a^{-1}ha \in S\} && (5.6b)\\ &= aSa^{-1} \cap H. && (5.6c)\end{aligned}$$

Nous cherchons $a \in G$ tel que l'entier

$$\frac{\mathrm{Card}(H)}{\mathrm{Card}\big(aSa^{-1} \cap H\big)} \tag{5.7}$$

EqZpIUbWx

soit premier avec p. En effet, dans ce cas le groupe $\mathrm{Fix}([a])$ est un p-Sylow de H parce que $|H : aSa^{-1} \cap H|$ ne divise pas p. La formule des orbites (équation (2.68)) nous dit que

$$\frac{|H|}{|aSa^{-1} \cap H|} = \mathrm{Card}\big(\mathcal{O}_{[a]}\big). \tag{5.8}$$

Supposons que toutes les orbites aient un cardinal divisible par p. Étant donné que G/S est une réunion disjointe de ses orbites, nous aurions

$$p \mid \mathrm{Card}(G/S) = \frac{|G|}{|S|} \tag{5.9}$$

alors que S étant un p-Sylow, p ne peut pas diviser $|G|/|S|$. Toutes les orbites n'ont donc pas un cardinal divisible par p, et il existe un $a \in G$ tel que (5.7) soit vérifiée. □

ThoUkPDXf

Théorème 5.11 (Théorème de Sylow)**.**
Soit G un groupe fini et p, un diviseur premier de $|G|$. Alors

ITEMooETYHooXlUMQZ

(1) G possède au moins un p-Sylow[2]*.*

(2) Tout p-sous-groupe de G est contenu dans un p-Sylow.

ItemMzNRVf

(3) Les p-Sylow de G sont conjugués.

ItemkYbdzZ

(4) Si n_p est le nombre de p-Sylow de G, alors n_p divise $|G|$ et $n_p \in [1]_p$.

Démonstration. En plusieurs points.

(1) Nous savons de la remarque 5.5 que G est un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_p)$ et que ce dernier a un p-Sylow par la proposition 5.8. Par conséquent G possède un p-Sylow par le lemme 5.10.

(2) Soit H un p-sous-groupe de G et S, un p-Sylow de G (qui existe par le point précédent). Par le lemme 5.10 il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H. Mais H est un p-groupe et un p-Sylow dans un p-groupe est automatiquement le groupe entier. Par conséquent,

$$H = aSa^{-1} \cap H \tag{5.10}$$

et $H \subset aSa^{-1}$, ce qui signifie que H est inclus dans un p-Sylow.

(3) Soit H un p-Sylow. Nous venons de voir que si S est un p-Sylow quelconque, alors H est inclus au p-Sylow aSa^{-1} pour un certain $a \in G$. Donc H est un p-Sylow inclus dans le p-Sylow aSa^{-1}, donc $H = aSa^{-1}$.

(4) Le fait que n_p divise n vient du fait que tous les p-Sylow ont le même nombre d'éléments (ils sont conjugués) et sont deux à deux disjoints. Donc ils forment une partition de G et $|G| = n_p|S|$ si S est un p-Sylow quelconque.

2. Définition 5.6.

Montrons maintenant que n_p est congru à un modulo p. Soit E l'ensemble des p-Sylow de G. Le groupe G agit sur E par conjugaison. Soit S un p-Sylow et considérons l'ensemble

$$E_S = \{T \in E \text{ tel que } s \cdot T = T, \forall s \in S\}. \tag{5.11}$$

où l'action est celle par conjugaison. C'est l'ensemble des points fixes de E sous l'action de S. L'ensemble E est la réunion des orbites sous S et chacune de ces orbites a un cardinal qui divise $|S| = p^m$. Par conséquent $|\mathcal{O}_T|$ vaut 1 lorsque $T \in E_S$ et est un multiple de p sinon. Nous avons donc

$$|E| \equiv |E_S| \mod p. \tag{5.12}$$

Nous voulons obtenir $|E_S| = 1$. Évidemment $S \in E_S$ parce que si $s \in S$ alors $sSs^{-1} = S$. Nous voudrions montrer que S est le seul élément de E_S. Soit $T \in E_S$, c'est-à-dire que T est un p-Sylow de G tel que

$$sTs^{-1} = T \tag{5.13}$$

pour tout $s \in S$. Soit N le groupe engendré par S et T. Montrons que T est normal dans N. Un élément g dans N s'écrit

$$g = s_1 t_1 \cdots s_r t_r \tag{5.14}$$

avec $s_i \in S$ et $t_i \in T$. Si $t \in T$, en utilisant le fait que T est un groupe et le fait que S le normalise, nous avons

$$gtg^{-1} = s_1 t_1 \dots s_r t_r t t_r^{-1} s_r^{-1} \dots t_1^{-1} s_1^{-1} \in T. \tag{5.15}$$

Donc T est un sous-groupe normal de N. Mais S et T sont conjugués dans N (parce que ils sont des p-Sylow de N), donc il existe un élément $a \in N$ tel que $aTa^{-1} = S$. Mais étant donné que T est normal,

$$S = aTa^{-1} = T. \tag{5.16}$$

Ceci achève la démonstration des théorèmes de Sylow.

□

Proposition 5.12.
Si S est un p-Sylow dans le groupe G alors pour tout $g \in G$, l'ensemble gSg^{-1} est encore un p-groupe.

Démonstration. Si les éléments de S sont d'ordre p^n, alors nous avons

$$(gsg^{-1})^q = gs^q g^{-1} = e. \tag{5.17}$$

Pour avoir $gs^q g^{-1} = e$, il faut et suffit que $gs^q = g$, alors $s^q = e$, c'est-à-dire $q = p^n$. Donc gSg^{-1} est encore un p-Sylow. □

Lemcmbzum

Lemme 5.13 ([141])**.**
Soit G, un groupe fini et p, un nombre premier. Si H et K sont des groupes distincts d'ordre p, alors $H \cap K = \{e\}$.

Démonstration. L'ensemble $H \cap K$ est un sous-groupe de H. Par conséquent son ordre divise celui de H qui est un nombre premier. Par conséquent soit $|H \cap K| = 1$, soit $|H \cap K| = |H|$. Dans le second cas nous aurions $H = K$, alors que nous avons supposé que H et K étaient distincts. □

PropyfhTmf

Proposition 5.14 ([141])**.**
Soit G un groupe fini et n le nombre de sous-groupes d'ordre p dans G. Alors le nombre d'éléments d'ordre p dans G vaut $n(p-1)$.

Démonstration. Si g est un élément d'ordre p dans G, le groupe H engendré par g est d'ordre p. Réciproquement si H est un groupe d'ordre p, tous les éléments de $H\backslash\{e\}$ sont d'ordre p (parce que l'ordre d'un élément divise l'ordre du groupe). Donc l'ensemble des éléments d'ordre p dans G est la réunion des ensembles $H\backslash\{e\}$ où H parcourt les sous-groupes d'ordre p dans G. Chacun de ces ensembles possède $p-1$ éléments et le lemme 5.13 nous assure qu'ils sont disjoints. Par conséquent nous avons $n(p-1)$ éléments d'ordre p dans G. □

Corolaire 5.15.
Un groupe d'ordre premier est cyclique.

Démonstration. Soit p l'ordre de G. Le nombre de sous-groupes d'ordre p est $n = 1$ (et c'est G lui-même). La proposition 5.14 nous dit alors que le nombre d'éléments d'ordre p dans G est $p-1$. Donc tout élément est générateur. □

5.2 Groupe monogène

SECooXIHPooWVSjhT

Groupe monogène : définition 1.325 ; groupe cyclique : définition 1.326.

Le théorème suivant donne quelques informations à propos des groupes monogènes. Il impliquera dans le corolaire 5.42 qu'un groupe monogène d'ordre n possède $\varphi(n)$ générateurs où φ est la fonction indicatrice d'Euler définie en 5.29.

THOooDOMZooOEYHAe

Théorème 5.16.
Un groupe monogène est abélien. Plus précisément,

(1) un groupe monogène infini est isomorphe à $\mathbb{Z}$,

(2) un groupe monogène fini est isomorphe à $\mathbb{Z}/n\mathbb{Z}$ pour un certain n.

Démonstration. Le groupe est abélien parce que $g = a^n$, $g' = a^{n'}$ implique $gg' = q^{n+n'} = g'g$. Nous considérons un générateur a de G (qui existe parce que G est monogène) et le morphisme surjectif

$$\begin{aligned} f\colon \mathbb{Z} &\to G \\ p &\mapsto a^p. \end{aligned} \tag{5.18}$$

Si G est infini, alors f est injective parce que si $a^n = a^{n'}$, alors $a^{n-n'} = e$, ce qui rendrait G cyclique et par conséquent non infini. Nous concluons que si G est infini, alors f est une bijection et donc un isomorphisme $\mathbb{Z} \simeq G$.

Si G est fini, alors f n'est pas injective et a un noyau $\ker f$. Étant donné que $\ker f$ est un sous-groupe de G, il existe un (unique) n tel que $\ker f = n\mathbb{Z}$ et le premier théorème d'isomorphisme (théorème 2.6) nous indique que

$$\mathbb{Z}/\ker f = \mathbb{Z}/n\mathbb{Z} = \text{Image } f = G. \tag{5.19}$$

□

Le lemme suivant donne une démonstration alternative, avec une construction plus explicite de l'isomorphisme.

LemZhxMit

Lemme 5.17 ([1])**.**
À propos de groupes monogènes[3] *et cycliques.*

(1) Soit un groupe cyclique G de cardinal n dont g est un générateur. Alors il existe un isomorphisme

$$\phi\colon G \to (\mathbb{Z}/n\mathbb{Z}, +) \tag{5.20}$$

tel que $\phi(g) = 1$.

3. Définition 1.325.

(2) Si G est un groupe monogène d'ordre infini et si g est un générateur, alors il existe un isomorphisme

$$\phi\colon G \to (\mathbb{Z}, +) \tag{5.21}$$

tel que $\phi(g) = 1$.

(3) Soient G et H deux groupes monogènes de même ordre. Soient g un générateur de G et h, un générateur de H. Il existe un isomorphisme de G sur H qui envoie g sur h.

Démonstration. Commençons par enfoncer une porte ouverte : comme le groupe est monogène, l'ordre du groupe est égal à l'ordre de son générateur. Nous séparons les cas selon que l'ordre soit fini ou non.

(1) **L'ordre de G est fini et vaut n** Si $k \in \mathbb{Z}$, nous notons $[k]_n$ la classe de k modulo n, c'est-à-dire l'ensemble $\{k + pn \text{ tel que } p \in \mathbb{Z}\}$.
Nous construisons l'isomorphisme $\phi\colon G \to \mathbb{Z}/n\mathbb{Z}$ de la façon suivante :

$$\phi(g^m) = [m]_n. \tag{5.22}$$

Cela est une bonne définition parce qu'une égalité du type $g^m = g^{m'}$ implique que m et m' soient dans la même classe modulo n. Nous vérifions que cela est un isomorphisme entre G et $\mathbb{Z}/n\mathbb{Z}$.

(1a) **Morphisme** Pour l'identité, si $x = e$ alors $m = 0$ et $\phi(e) = [0]_n$. Et si $x = g^k$, $y = g^l$ alors $\phi(xy) = \phi(g^{k+l}) = [k + l]_n = [k]_n + [l]_n = \phi(x) + \phi(y)$.

(1b) **Injectif** Supposons $\phi(g^k) = \phi(g^l)$ avec $k \geqslant l$. Nous avons $h^k = h^l$, donc $h^{k-l} = e$, ce qui donne $k - l \in [0]_n$ ou encore $[k]_n = [l]_n$. En particulier $g^k = g^l$.

(1c) **Surjectif** La classe $[k]_n$ est l'image de g^k.

(2) **L'ordre de G est infini**
Si l'ordre de G est infini alors un élément $x \in G$ s'écrit de façon unique sous la forme $x = g^m$ avec $m \in \mathbb{Z}$. Dans ce cas nous définissons directement $\phi(g^m) = m$.
Le reste de la preuve est alors identique au cas d'ordre fini, mais sans les complications liées au modulo.

La dernière assertion s'obtient des précédentes par composition d'isomorphismes.

□

5.3 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$

Notons que $\mathbb{Z}/n\mathbb{Z} = \mathbb{F}_n$ est un groupe pour l'addition tandis que $(\mathbb{Z}/n\mathbb{Z})^*$ est un groupe pour la multiplication. Il ne peut donc pas y avoir d'équivoque.

ThoozyeSn

Théorème 5.18 ([142]).
Pour chaque $x \in (\mathbb{Z}/n\mathbb{Z})^*$ nous considérons l'application

$$\begin{aligned} \sigma_x\colon \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ y &\mapsto xy. \end{aligned} \tag{5.23}$$

L'application

$$\sigma\colon \big((\mathbb{Z}/n\mathbb{Z})^*,\ \cdot\ \big) \to \mathrm{Aut}\left(\mathbb{Z}/n\mathbb{Z}, +\right) \tag{5.24}$$

ainsi définie est un isomorphisme de groupes.

L'énoncé de ce théorème s'écrit souvent rapidement par

$$\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z}) = (\mathbb{Z}/n\mathbb{Z})^*, \tag{5.25}$$

mais il faut bien garder à l'esprit qu'à gauche on considère le groupe additif et à droite celui multiplicatif.

Démonstration. Nous notons $[x]$ la classe de x dans $\mathbb{Z}/n\mathbb{Z}$. Nous avons $\mathbb{Z}/n\mathbb{Z} = [1]$. Soit f un automorphisme de $(\mathbb{Z}/n\mathbb{Z}, +)$; pour tout $r \in \mathbb{Z}$ nous avons

$$f([r]) = f(r[1]) = rf([1]) = [r]f([1]). \tag{5.26}$$

En particulier, puisque f est surjective, il existe un r tel que $f([r]) = [1]$. Pour un tel r nous avons $[1] = [r]f([1])$, c'est-à-dire que nous avons montré que $f([1])$ est inversible dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)$. Nous montrons à présent que [4]

$$\begin{aligned}\sigma\colon \operatorname{Aut}((\mathbb{Z}/n\mathbb{Z}, +)) &\to \big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)\\ f &\mapsto f([1])\end{aligned} \tag{5.27}$$

est un isomorphisme.

Nous commençons par la surjectivité. Soit $[a] \in (\mathbb{Z}/n\mathbb{Z})^*$. Les élément $[a]$ et $[1]$ étant tous deux des générateurs de $(\mathbb{Z}/n\mathbb{Z}, +)$, il existe un automorphisme de $\mathbb{Z}/n\mathbb{Z}$ qui envoie $[1]$ sur $[a]$ par le lemme 5.17. Cela prouve la surjectivité de σ.

En ce qui concerne l'injectivité, considérons des automorphismes f_1 et f_2 de $(\mathbb{Z}/n\mathbb{Z}, +)$ tels que $f_1([1]) = f_2([1])$. Les automorphismes f_1 et f_2 prennent la même valeur sur un générateur et donc sur tout le groupe. Donc $f_1 = f_2$.

Enfin nous prouvons que σ est un morphisme, c'est-à-dire que $\sigma(f \circ g) = \sigma(f)\sigma(g)$. Nous avons

$$f\big(g([1])\big) = f\big(g([1])[1]\big) = g([1])f([1]) = \sigma(f)\sigma(g). \tag{5.28a}$$

□

Ce dernier résultat s'étend aux groupes cycliques.

PROPooBZOMooVOHoYf

Proposition 5.19.
Si G est un groupe cyclique [5] *d'ordre n, alors*

$$\operatorname{Aut}(G) = \big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big). \tag{5.29}$$

Démonstration. Vu que G est cyclique, le lemme 5.17 nous dit que G est isomorphe à $(\mathbb{Z}/n\mathbb{Z}), +$. Maintenant le théorème 5.18 nous indique que

Les égalités suivantes sont en réalité des isomorphismes de groupes :

$$\operatorname{Aut}(G) = \operatorname{Aut}\Big(\mathbb{Z}/n\mathbb{Z}, +\Big) \tag{5.30a}$$

SUBEQooBNGBooEKRZUn

$$= \Big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\Big) \tag{5.30b}$$

SUBEQooZBMEooBSGPUB

Jusitifications.

— Pour (5.30a). Le lemme 5.17 nous dit que G est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$, et le lemme 1.41 dit que des groupes isomorphes ont des groupes d'isomorphismes isomorphes.

— Pour (5.30b). C'est le théorème 5.18.

□

CorwgmoTK

Corolaire 5.20.
Si p divise $q-1$ alors $\operatorname{Aut}(\mathbb{F}_q)$ *possède un unique sous-groupe d'ordre p.*

Démonstration. Si a est un générateur de $\mathbb{F}_q^*$ alors le groupe

$$\operatorname{gr}\Big(a^{\frac{q-1}{p}}\Big) \tag{5.31}$$

EqAdGiiJ

est un sous-groupe d'ordre p. En ce qui concerne l'unicité, soit S un sous-groupe d'ordre p. Il est donc d'indice $(q-1)/p$ dans $\mathbb{F}_q^*$ et le lemme 3.29 nous enseigne que le groupe donné en (5.31) est contenu dans S. Il est donc égal à S parce qu'il a l'ordre de S. Le fait que S soit normal est dû au fait que $\mathbb{F}_q^*$ est abélien. □

4. Le σ donné ici est l'inverse de celui donné dans l'énoncé. Cela ne change évidemment rien à la validité de l'énoncé et de la preuve.

5. Définition 1.326.

5.4 Groupes abéliens finis

Nous rappelons que l'exposant d'un groupe fini est le ppcm des ordres de ses éléments. Dans le cas des groupes abéliens finis, l'exposant joue un rôle important du fait qu'il existe un élément dont l'ordre est l'exposant. C'est le théorème suivant.

Théorème 5.21 (Exposant dans un groupe abélien fini)**.**
Un groupe abélien fini contient un élément dont l'ordre est l'exposant du groupe.

Démonstration. Soit G un groupe abélien fini et $x \in G$, un élément d'ordre maximum m. Nous montrons par l'absurde que l'ordre de tous les éléments de G divise m. Soit donc $y \in G$, un élément dont l'ordre ne divise pas m ; nous notons q son ordre. Vu que q ne divise pas m, le nombre q possède au moins un facteur premier plus de fois que m : soit p premier tel que la décomposition de q contienne p^β et celle de m contienne p^α avec $\beta > \alpha$. Autrement dit,

$$m = p^\alpha m' \tag{5.32a}$$
$$q = p^\beta q' \tag{5.32b}$$

où m' et q' ne contiennent plus le facteur p. L'élément x étant d'ordre m, l'élément x^{p^α} est d'ordre m'. De la même manière, l'élément $y^{q'}$ est d'ordre p^β. Étant donné que p^β et m' sont premiers entre eux, l'élément $x^{p^\alpha}y^{q'}$ est d'ordre $p^\alpha m' > m$. D'où une contradiction avec le fait que x était d'ordre maximal.

Par conséquent l'ordre de tous les éléments de G divise celui de x qui est alors le ppcm des ordres de tous les éléments de G, c'est-à-dire l'exposant de G. □

PropfPRVxi

Proposition 5.22.
Soit G un groupe abélien fini et $x \in G$, un élément d'ordre maximum. Alors

(1) Il existe un morphisme $\varphi\colon G \to \mathrm{gr}(x)$ tel que $\varphi(x) = x$. ItemKRYwjU

(2) Il existe un sous-groupe K de G tel que $G = \mathrm{gr}(x) \oplus K$.

Démonstration. Nous notons a l'ordre de x qui est également l'exposant du groupe G.

Nous allons prouver la première partie par récurrence sur l'ordre du groupe. Si $G = \mathrm{gr}(x)$, alors c'est évident. Soit H un sous-groupe propre de G contenant x et tel que le problème soit déjà résolu pour H : il existe un morphisme $\varphi\colon H \to \mathrm{gr}(x)$ tel que $\varphi(x) = x$. Soit $y \in G\backslash H$, d'ordre b. Nous allons trouver un morphisme $\hat{\varphi}\colon \mathrm{gr}(H, y) \to \mathrm{gr}(x)$ telle que $\hat{\varphi}(x) = x$.

Pour cela nous commençons par construire les applications suivantes :

$$\begin{aligned} \tilde{\varphi}\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(x) \\ (\bar{k}, h) &\mapsto x^{kl}\varphi(h) \end{aligned} \tag{5.33}$$

où l est encore à déterminer, et

$$\begin{aligned} p\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(y, H) \\ (\bar{k}, h) &\mapsto y^k h. \end{aligned} \tag{5.34}$$

Pour que $\tilde{\varphi}$ soit bien définie, il faut que a divise bl. L'application p est bien définie parce que $\bar{k}$ est pris dans $\mathbb{Z}/b\mathbb{Z}$ et que b est l'ordre de y.

Nous allons construire le morphisme $\hat{\varphi}$ en considérant le diagramme

$$\begin{array}{ccccc} \ker(p) & \hookrightarrow & \mathbb{Z}/b\mathbb{Z} \times H & \xrightarrow{p} & \mathrm{gr}(y, H) \\ & & \downarrow{\scriptstyle \tilde{\varphi}} & \swarrow{\scriptstyle \hat{\varphi}} & \\ & & \mathrm{gr}(x) & & \end{array} \tag{5.35}$$

que l'on voudra être commutatif. Puisque p est surjective, les théorèmes d'isomorphismes nous disent que

$$\mathrm{gr}(y, H) \simeq \frac{\mathbb{Z}/b\mathbb{Z} \times H}{\ker p}. \tag{5.36}$$

Si $[\bar{k}, h]$ est la classe de $(\bar{k}, h)$ modulo $\ker(p)$ alors nous voudrions définir $\hat{\varphi}$ par

$$\hat{\varphi}\big([\bar{k}, h]\big) = \tilde{\varphi}(\bar{k}, h). \tag{5.37}$$

EqgesVxc

Pour que cela soit bien défini, il faut que si $(\bar{r}, z) \in \ker p$, alors,

$$\hat{\varphi}\big([\bar{k}\bar{r}, hz]\big) = \hat{\varphi}\big([\bar{k}, h]\big), \tag{5.38}$$

c'est-à-dire que $\tilde{\varphi}(\bar{r}, z) = e$. Du coup la définition (5.37) n'est bonne que si et seulement si

$$\ker(p) \subset \ker(\tilde{\varphi}). \tag{5.39}$$

Nous pouvons obtenir cela en choisissant bien l.

Déterminons d'abord le noyau de p. Pour cela nous considérons un nombre β divisant b tel que $\mathrm{gr}(y) \cap H = \mathrm{gr}(y^\beta)$. Nous aurons $p(\bar{k}, h) = e$ si et seulement si $y^h = e$. En particulier $h = y^{-k} \in \mathrm{gr}(y) \cap H = \mathrm{gr}(y^\beta)$. Si $h = (y^\beta)^m = y^{m\beta}$, alors $k = -m\beta$ et nous avons

$$\ker(p) = \{(-m\beta, y^{m\beta}) \text{ tel que } m \in \mathbb{Z}\}. \tag{5.40}$$

En plus court : $\ker(p) = \mathrm{gr}(\beta, y^{-\beta})$. Nous devons donc fixer l de telle sorte que $\tilde{\varphi}(\beta, y^{-\beta}) = e$. Étant donné que φ prend ses valeurs dans $\mathrm{gr}(x)$, il existe un entier α tel que $\varphi(y^{-\beta}) = x^\alpha$; en utilisant cet α, nous écrivons

$$\tilde{\varphi}(\beta, y^{-\beta}) = x^{\beta l}\varphi(y^{-\beta}) = x^{\beta l + \alpha}. \tag{5.41}$$

Par conséquent nous choisissons $l = -\alpha/\beta$. Nous devons maintenant vérifier que ce choix est légitime, c'est-à-dire que a divise bl et que α/β est un entier.

Étant donné que y est d'ordre b,

$$e = \varphi(y^b) = \varphi(y^{-\beta b/\beta}) = \varphi(y^{-\beta})^{b/\beta} = x^{b\beta/\alpha}. \tag{5.42}$$

Par conséquent a divise $\frac{b\alpha}{\beta} = -bl$.

Pour voir que l est entier, nous nous rappelons que a est l'exposant de G (parce que x est d'ordre maximum) et que par conséquent b divise a. Mais a divise $\alpha\frac{b}{\beta}$. Donc α/β est entier.

Nous passons maintenant à la seconde partie de la preuve. Nous considérons un morphisme $\varphi\colon G \to \mathrm{gr}(x)$ tel que $\varphi(x) = x$. La première partie nous en assure l'existence. Nous montrons que

$$\begin{aligned} \psi\colon G &\to \mathrm{gr}(x) \oplus \ker(\varphi) \\ g &\mapsto \big(\varphi(g), g\varphi(g)^{-1}\big) \end{aligned} \tag{5.43}$$

est un isomorphisme. D'abord $g\varphi(g)^{-1}$ est dans le noyau de φ parce que $\varphi(g)^{-1}$ étant dans $\mathrm{gr}(x)$, et φ étant un morphisme,

$$\varphi\big(g\varphi(g)^{-1}\big) = \varphi(g)\varphi(g)^{-1} = e. \tag{5.44}$$

L'application ψ est un morphisme parce que, en utilisant le fait que G est abélien,

$$\begin{aligned} \psi(g_1g_2) &= \big(\varphi(g_1g_2), g_1g_2\varphi(g_1g_2)^{-1}\big) && \text{(5.45a)} \\ &= \big(\varphi(g_1)\varphi(g_2), g_1\varphi(g_1)^{-1}g_2\varphi(g_2)^{-1}\big) && \text{(5.45b)} \\ &= \psi(g_1)\psi(g_2). && \text{(5.45c)} \end{aligned}$$

L'application ψ est injective parce que si $\psi(g) = (e, e)$ alors $\varphi(g) = e$ et $g\varphi(g)^{-1} = e$, ce qui implique $g = e$.

Enfin ψ est surjective parce qu'elle est injective et que les ensembles de départ et d'arrivée ont même cardinal. En effet par le premier théorème d'isomorphisme (théorème 2.6) appliqué à φ nous avons

$$|G| = |\,\mathrm{gr}(x)| \cdot |\ker(\varphi)|. \tag{5.46}$$

□

ThoRJWVJd

Théorème 5.23.
Si G est un groupe abélien fini non trivial, il exise un unique $r > 0$ et une unique liste de naturels $(d_1, \dots, d_r)$ tels que

(1) $G \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/d_r\mathbb{Z}$

(2) $d_1 \geqslant 1$

(3) d_i *divise* d_{i+1} *pour tout* $i = 1, \dots, r-1$.

Démonstration. Soit x_1 un élément d'ordre maximal dans G. Soit n_1 son ordre et

$$H_1 = \mathrm{gr}(x_1) = \mathbb{F}_{n_1}. \tag{5.47}$$

D'après la proposition 5.22(2), il existe un supplémentaire K_1 tel que $G = \mathbb{F}_{n_1} \oplus K_1$. Si $K_1 = \{e\}$ on s'arrête et on garde $G = \mathbb{F}_{n_1}$. Sinon on continue de la sorte en prenant x_2 d'ordre maximal dans K_1 etc.

Nous devons maintenant prouver l'unicité de cette décomposition. Supposons deux décompositions avec les nombres $(d_1, \dots, d_r)$ et $(s_1, \dots, s_q)$:

$$G = \mathbb{F}_{d_1} \oplus \dots \oplus \mathbb{F}_{d_r} = \mathbb{F}_{s_1} \oplus \dots \oplus \mathbb{F}_{s_q}. \tag{5.48}$$

L'exposant de G est d_r et s_q. Donc $d_r = s_q$. Les complémentaires étant égaux nous avons

$$\mathbb{F}_{d_1} \oplus \dots \oplus \mathbb{F}_{d_{r-1}} = \mathbb{F}_{s_1} \oplus \dots \oplus \mathbb{F}_{s_{q-1}}. \tag{5.49}$$

En continuant nous trouvons $r = q$ et $d_i = s_i$. □

5.5 Groupes d'ordre pq

Lemme 5.24.
Soit G un groupe d'ordre pq où p et q sont des nombres premiers distincts. Nous supposons que $p < q$.

(1) Le groupe G possède un unique q-Sylow.

(2) Cet unique q-Sylow est normal dans G.

(3) Il n'est ni $\{e\}$ ni G.

(4) Le groupe G n'est pas un groupe simple[6].

Démonstration. Soit n_q le nombre de q-Sylow ; par le théorème de Sylow 5.11(1) le groupe G possède des q-Sylow et par 5.11(4),

$$n_q \in [1]_q. \tag{5.50}$$

De plus le nombre n_q divise $|G| = pq$. Donc n_q vaut p, q ou 1. Avoir $n_q = p$ n'est pas possible parce que $n_q \in [1]_q$ et $p < q$. Avoir $n_q = q$ n'est pas possible non plus, pour la même raison. Donc $n_q = 1$. Notons H l'unique q-Sylow de G.

Le fait que H soit normal est une conséquence de 5.11(3) parce que le conjugué de H est encore un q-Sylow alors que H est l'unique q-Sylow.

Vu que

$$1 < p = |H| < pq = |G|, \tag{5.51}$$

le sous-groupe H n'est ni réduit à l'identité ni le groupe entier.

Par conséquent G n'est pas simple parce qu'il contient un sous-groupe normal non trivial. □

Avant de lire le théorème suivant, n'oubliez pas de lire la définition d'un produit semi-direct 2.47.

6. Pas de sous-groupes normaux non triviaux, 1.175.

ThoLnTMBy

Théorème 5.25 ([143]).
Soient deux nombres premiers distincts[7] *p et q avec $q > p$.*

(1) Si p ne divise pas $q-1$ alors tout groupe d'ordre pq est cyclique et plus précisément le seul groupe (à isomorphisme près) d'ordre pq est $\mathbb{Z}/pq\mathbb{Z}$. ITEMooFQXIooFLAiUD

(2) Si $p \mid q-1$, alors il n'existe que deux groupes d'ordre pq :
- *Le groupe abélien et cyclique $\mathbb{Z}/pq\mathbb{Z}$.*
- *Le produit semi-direct non abélien*

$$G = \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} \tag{5.52}$$

EqNuuTRE

où $\varphi(\bar{1})$ est d'ordre p dans $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$.

(3) Si p et q sont premiers entre eux, le produit est direct[8].

Démonstration. Division de la preuve en plusieurs parties.

(1) **Préliminaires avec Sylow**

Soit un groupe G d'ordre pq. Soient H, un q-Sylow et K, un p-Sylow de G. Ils existent parce que p et q sont des diviseurs premiers de $|G|$ (théorème de Sylow 5.11). Si n_q est le nombre de q-Sylow dans G alors n_q divise $|G|$ et $n_q = 1 \mod q$. Donc d'abord n_q vaut 1, p ou q. Ensuite $n_q = q$ est exclu par la condition $n_q = 1 \mod q$; la possibilité $n_q = p$ est également impossible parce que $p = 1 \mod q$ est impossible avec $p < q$. Donc $n_q = 1$ et H est normal dans G.

L'ensemble $H \cap K$ est un sous-groupe à la fois de H et de K, ce qui entraine que (théorème de Lagrange 2.13) $|H \cap K|$ divise à la fois p et q. Nous en déduisons que $|H \cap K| = 1$ et donc que $H \cap K = \{e\}$.

Étant donné que H est normal, l'ensemble HK est un sous-groupe de G. De plus l'application

$$\begin{aligned} \psi\colon H \times K &\to HK \\ (h,k) &\mapsto hk \end{aligned} \tag{5.53}$$

est un bijection. Nous ne devons vérifier seulement l'injectivité. Supposons que $hk = h'k'$. Alors $e = h^{-1}h'k'k^{-1}$, et donc

$$h^{-1}h' = (k'k^{-1})^{-1} \in H \cap K = \{e\}. \tag{5.54}$$

Par conséquent $|pq| = |H \times K| = |HK|$, et $HK = G$. Le corolaire 2.49 nous indique que

$$G = H \times_\varphi K \tag{5.55}$$

EqGiQiEN

où φ est l'action adjointe. Nous devons maintenant identifier cette action. En d'autres termes, nous savons que $H = \mathbb{Z}/q\mathbb{Z}$ et $K = \mathbb{Z}/p\mathbb{Z}$ et que $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ est un morphisme. Nous devons déterminer les possibilités pour φ.

Soit n_p le nombre de p-Sylow de G. Comme précédemment, n_p vaut 1, p ou q et la possibilité $n_p = p$ est exclue. Donc n_p est 1 ou q.

(2) **Si p ne divise pas $q-1$**

Si p ne divise pas $q-1$ alors il n'est pas possible d'avoir $n_p = q$ parce que $n_p \in [1]_p$. Or dire $n_p = q$ demanderait $q \in [1]_p$, c'est-à-dire $q = kp + 1$, qui impliquerait que p divise $q-1$.

La seule possibilité est que $n_p = 1$. Dans ce cas, K est également normal dans G. Du coup le produit semi-direct (5.55) est en réalité un produit direct (φ est triviale) et nous avons

$$G = \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/pq\mathbb{Z}. \tag{5.56}$$

7. Le cas $p = q$ sera traité par la proposition 5.28.

8. Cette affirmation me semble très bizarre. Comment deux nombres premiers distincts pourraient ne pas être premiers entre eux ? ? ?

(3) **Si p divise $q-1$**
Cette fois $n_p = 1$ et $n_p = q$ sont tous deux possibles. Ce que nous savons est que $\varphi(\mathbb{Z}/p\mathbb{Z})$ est un sous-groupe de $\text{Aut}(\mathbb{Z}/q\mathbb{Z})$. Par le premier théorème d'isomorphisme 2.6, nous avons

$$|\varphi(\mathbb{Z}/p\mathbb{Z})| = \frac{|\mathbb{Z}/p\mathbb{Z}|}{|\ker\varphi|}, \tag{5.57}$$

ce qui signifie que $|\varphi(\mathbb{Z}/p\mathbb{Z})|$ divise $|\mathbb{Z}/p\mathbb{Z}| = p$. Par conséquent, $|\varphi(\mathbb{Z}/p\mathbb{Z})|$ est égal à 1 ou p. Si c'est 1, alors l'action est triviale et le produit est direct.

Nous supposons que $|\varphi(\mathbb{Z}/p\mathbb{Z})| = p$. Le corolaire 5.20 nous indique que $\text{Aut}(\mathbb{Z}/q\mathbb{Z})$ possède un unique sous-groupe d'ordre p que nous notons Γ ; c'est-à-dire que $\Gamma = \text{Image}(\varphi)$. Vu que $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \text{Aut}(\mathbb{Z}/q\mathbb{Z})$ est un morphisme, Γ est généré par $\varphi(\bar{1})$ qui est alors un élément d'ordre p, comme annoncé.

(4) **Unicité** Nous nous attaquons maintenant à l'unicité. Soient φ et φ' deux morphismes non triviaux $\mathbb{Z}/p\mathbb{Z} \to \text{Aut}(\mathbb{Z}/q\mathbb{Z})$. Étant donné que $\text{Aut}(\mathbb{Z}/q\mathbb{Z})$ ne possède qu'un seul sous-groupe d'ordre p, nous savons que $\text{Image}(\varphi) = \text{Image}(\varphi') = \Gamma$. Nous pouvons donc parler de φ'^{-1} en tant qu'application de $\mathbb{Z}/p\mathbb{Z}$ dans Γ. Nous montrons que

$$\begin{aligned} f\colon \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} &\to \mathbb{Z}/q\mathbb{Z} \times_{\varphi'} \mathbb{Z}/p\mathbb{Z} \\ (h,k) &\mapsto (h, \alpha(k)) \end{aligned} \tag{5.58}$$

où $\alpha = \varphi'^{-1} \circ \varphi$ est un isomorphisme de groupes. Le calcul est immédiat :

$$\begin{aligned} f(h_1,k_1)f(h_2 m k_2) &= \big(h_1, \alpha(k_1)\big)(h_2, \alpha(k_2)) && \text{(5.59a)}\\ &= \big(h_1\varphi'(\alpha(k_1))h_2 m \alpha(k_1k_2)\big) && \text{(5.59b)}\\ &= f\big(h_1\varphi(k_1)h_2, k_1k_2\big) && \text{(5.59c)}\\ &= f\big((h_1,k_1),(h_2,k_2)\big). && \text{(5.59d)} \end{aligned}$$

Par conséquent $\mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} \simeq \mathbb{Z}/q\mathbb{Z} \times_{\varphi'} \mathbb{Z}/p\mathbb{Z}$.

□

Note : il existe des nombres premiers p et q tels que $q = 1 \mod p$. Par exemple $7 = 1 \mod 3$.
PROPooNSRYooEodtUl

Proposition 5.26 ([54]).
Soit G un groupe fini d'ordre pq où p et q sont deux nombres premiers distincts vérifiant

$$\begin{cases} p \neq 1 \mod q & \text{(5.60a)}\\ q \neq 1 \mod p. & \text{(5.60b)} \end{cases}$$

Alors G est cyclique, abélien et

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{5.61}$$

Démonstration. Soient n_p et n_q les nombres de p-Sylow et q-Sylow. Par le théorème de Sylow 5.11, n_p divise pq et $n_p = 1 \mod p$. Le second point empêche n_p de diviser p. Par conséquent n_p divise q et donc n_p vaut 1 ou q. La possibilité $n_p = q$ est exclue par l'hypothèse $q \neq 1 \mod p$. Donc $n_p = 1$, et de la même façon nous obtenons $n_q = 1$.

Soient S l'unique p-Sylow et T, l'unique q-Sylow. Pour les mêmes raisons que celles exposées plus haut, ce sont deux sous-groupes normaux dans G. Étant donné que S est d'ordre p^n pour un certain n et que l'ordre de S doit diviser celui de G, nous avons $|S| = p$. De la même façon, $|T| = q$. Par conséquent S est un groupe cyclique d'ordre p et nous considérons x, un de ses générateurs. De la même façon soit y, un générateur de T.

Nous montrons maintenant que x et y commutent, puis que xy engendre G. Nous savons que $S \cap T$ est un sous-groupe à la fois de S et de T, de telle façon que $|S \cap T|$ divise à la fois $|S| = p$

et $|T| = q$. Nous avons donc $|S \cap T| = 1$ et donc $S \cap T$ se réduit au neutre. Par ailleurs, S et T sont normaux, donc

$$(xyx^{-1})y^{-1} \in T \tag{5.62a}$$
$$x(yx^{-1}y^{-1}) \in S, \tag{5.62b}$$

donc $xyx^{-1}y^{-1} = e$, ce qui montre que $xy = yx$.

Montrons que xy engendre G. Soit $m > 0$ tel que $(xy)^m = e$. Pour ce m nous avons $x^m = y^{-m}$ et $y^{-m} = x^m$, ce qui signifie que x^m et y^m appartiennent à $S \cap T$ et donc $x^m = y^m = e$. Les nombres p et q divisent donc tous deux m ; par conséquent $\text{ppcm}(p, q) = pq$ divise m. Nous en concluons que xy est d'ordre pq (il ne peut pas être plus) et qu'il est alors générateur.

Pour la suite nous allons d'abord prouver que $G = ST$ puis que $G \simeq S \times T$. Nous savons déjà que $|S \cap T| = 1$, ce qui nous amène à dire que $|ST| = |S||T|$. En effet si $s, s' \in S$ et $t, t' \in t$ et si $st = s't'$, alors $t = s^{-1}s't'$, ce qui voudrait dire que $s^{-1}s' \in T$ et donc que $s^{-1}s' = e$. Au final nous avons

$$|ST| = |S||T| = pq = |G|. \tag{5.63}$$

Par conséquent $G = ST$. En nous rappelant que $S \cap T = \{e\}$ et que S et T sont normaux, le lemme 1.270 nous dit que $G \simeq S \times T$. Le groupe S étant cyclique d'ordre p nous avons $S = \mathbb{Z}/p\mathbb{Z}$ et pour T, nous avons la même chose : $T = \mathbb{Z}/q\mathbb{Z}$. Nous concluons que

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{5.64}$$

□

ThoImkljy

Théorème 5.27 (Théorème de Burnside[90])**.**
Le centre d'un p-groupe non trivial est non trivial.

Démonstration. Soit G un p-groupe non trivial. Nous considérons l'action adjointe G sur lui-même. Les points fixes de cette action sont les éléments du centre :

$$\mathcal{Z}_G = \{z \in G \text{ tel que } \sigma_x(z) = z, \forall x \in G\} = \text{Fix}_G(G). \tag{5.65}$$

Nous utilisons l'équation aux classes (2.38) pour dire que $|G| = |\mathcal{Z}_G| \mod p$. Mais $|\mathcal{Z}_G|$ n'est pas vide parce qu'il contient l'identité. Donc $|\mathcal{Z}_G|$ est au moins d'ordre p. □

PropssttFK

Proposition 5.28.
Si p est un nombre premier, tout groupe d'ordre p ou p^2 est abélien.

Rappel : un groupe d'ordre p ou p^2 est automatiquement un p-groupe.

Démonstration. Si $|G| = p$, alors le théorème de Cauchy 5.2 nous donne l'existence d'un élément d'ordre p. Cet élément est alors automatiquement générateur, G est cyclique et donc abélien.

Si par contre G est d'ordre p^2, alors les choses se compliquent (un peu). D'après le théorème de Burnside 5.27, le centre $\mathcal{Z}$ n'est pas trivial ; il est alors d'ordre p ou p^2. Supposons qu'il soit d'ordre p et prenons $x \in G \backslash \mathcal{Z}$. Alors le stabilisateur de x pour l'action adjointe contient au moins $\mathcal{Z}$ et x, c'est-à-dire que $|\text{Fix}_G(x)| \geqslant p + 1$. Étant donné que $\text{Fix}_G(x)$ est un sous-groupe, son ordre est automatiquement 1, p ou p^2. En l'occurrence, il doit être p^2 (parce que plus grand que p), et donc x doit être central, ce qui est une contradiction. □

5.5.1 Fonction indicatrice d'Euler

DEFooZRYMooZCozga

Définition 5.29.
*La **fonction indicatrice d'Euler** est l'application*

$$\begin{aligned} \varphi\colon \mathbb{N}^* &\to \mathbb{N}^* \\ n &\mapsto \text{Card}\Big(\{m \in \mathbb{N}^* \text{ tel que } 1 \leqslant m \leqslant n, \text{pgcd}(m, n) = 1\}\Big). \end{aligned} \tag{5.66}$$

5.30. (1) Voir le thème 11 pour des formules concernant l'indicatrice d'Euler.

(2) Une vidéo de 3Blue1Brown : https://www.youtube.com/watch?v=EK32jo7i5LQ.

LEMooVGDHooStUaKH

Lemme 5.31 ([144]).
L'élément $[m]_n$ *est inversible dans le groupe* $\big((\mathbb{Z}/n\mathbb{Z})^*,\ \cdot\ \big)$ *si et seulement si* $\operatorname{pgcd}(m,n)=1$.

Démonstration. Dans les deux sens.

(1) $\Rightarrow$ Si $[m_n]$ est inversible, il existe $u\in\mathbb{Z}$ tel que $[u]_n[m]_n=[1]_n$. Cela donne $[um]_n=[1]_n$ ou encore $um\in[1]_n$, c'est-à-dire $um=1+vn$. Le théorème de Bézout 1.236 conclut que $\operatorname{pgcd}(m,n)=1$.

(2) $\Leftarrow$ Si $\operatorname{pgcd}(m,n)=1$, alors le théorème de Bézout 1.236 nous dit qu'il existe $u,v\in\mathbb{Z}$ tels que $um+vn=1$. Cela donne directement $um=1-vn\in[1]_n$ et donc $[u]_n$ est un inverse de $[m]_n$ dans le groupe multiplicatif $\mathbb{Z}/n\mathbb{Z}$.

□

LEMooCLYEooONhWKs

Lemme 5.32.
Nous avons

$$\varphi(n)=\operatorname{Card}\big((\mathbb{Z}/n\mathbb{Z})^\times\big) \tag{5.67}$$

où $A^\times$ *est le groupe des inversibles (pour la multiplication) dans l'anneau* A.

Démonstration. En utilisant le lemme 5.31,

$$\begin{aligned}\big(\mathbb{Z}/n\mathbb{Z}\big)^\times&=\{[m]_n\text{ tel que }\operatorname{pgcd}(m,n)=1\}&(5.68a)\\&=\{[m]_n\text{ tel que }1\leqslant m\leqslant n,\operatorname{pgcd}(m,n)=1\}&(5.68b)\end{aligned}$$

Comme deux entiers différents entre 1 et n ne peuvent pas être dans la même classe modulo n, il y a bijection entre le dernier ensemble et $\{1\leqslant m\leqslant n\text{ tel que }\operatorname{pgcd}(m,n)=1\}$. Donc

$$\varphi(n)=\operatorname{Card}\big(\{[m]_n\text{ tel que }1\leqslant m\leqslant n,\operatorname{pgcd}(m,n)=1\}\big)=\operatorname{Card}\big((\mathbb{Z}/n\mathbb{Z})^\times\big). \tag{5.69}$$

□

LEMooRGIYooRxgyCO

Lemme 5.33 ([145]).
Soit $n,d\in\mathbb{N}$ *tels que* $d\mid n$. *Nous notons* $q=n/d$ *et* $G=\{k[q]_n\text{ tel que }0\leqslant k\leqslant d-1\}$. *Alors, pour* $r\in\mathbb{Z}$ *nous avons* $d[r]_n=[0]_n$ *si et seulement si* $[r]_n\in G$.

Démonstration. Dans les deux sens.

(1) $\Rightarrow$ Nous supposons que $d[r]_n=[0]_n$. Étant donné que $n=dq$ nous avons les implications suivantes :

$$d[r]_n=[0]_n\Rightarrow d\mid dr\Rightarrow dq\mid dr\Rightarrow q\mid r. \tag{5.70}$$

Nous avons donc $r=kq$ pour un certain $k\in\mathbb{Z}$. Par la division euclidienne 1.222, il existe $s,t\in\mathbb{N}$ tels que $k=sd+t$ avec $t<d$. Avec ça, nous avons le calcul

$$[r]_n=k[q]_n=sd[q]_n+t[q]_n=\underbrace{s[dq]_n}_{=[0]_n}+t[q]_n\in G. \tag{5.71}$$

(2) $\Leftarrow$ Si $[r]_n\in G$, il existe $0\leqslant k\leqslant d-1$ tel que $[r]_n=k[q]_n$. De ce fait,

$$d[r]_n=dk[q]_n=k[dq]_n=[0]_n. \tag{5.72}$$

Et voila.

□

La proposition suivante est un pas important dans l'algorithme de Shor permettant aux ordinateurs quantiques de factoriser rapidement des grands nombres[146].

THOooXMBSooXrrfOP

Théorème 5.34 (Euler-Fermat[147]).
Deux énoncés très similaires.

(1) Soient a, n premiers entre eux dans $\mathbb{N}$. Alors Nous avons la formule

$$a^{\varphi(n)} \in [1]_n. \tag{5.73}$$

ITEMooJHZBooXVKMlT

(2) Soient $A, B \in \mathbb{N}$ premiers entre eux. Alors il existe $p, m \in \mathbb{N}$ tels que

$$A^p = mB + 1. \tag{5.74}$$

Démonstration. Vu que $\operatorname{pgcd}(a, n) = 1$, le théorème de Bézout 1.236 donne $ua+vn = 1$, c'est-à-dire $u[a]_n \in [1]_n$. Autrement dit, $[a]_n$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$. De ce fait l'application

$$\begin{aligned} f\colon (\mathbb{Z}/n\mathbb{Z})^\times &\to (\mathbb{Z}/n\mathbb{Z})^\times \\ x &\mapsto [a]_n x \end{aligned} \tag{5.75}$$

est une bijection.

Un produit peut être réindexé par une bijection [9]. En posant $P = \prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} x$, nous avons

$$\begin{aligned} P &= \prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} x && \text{(5.76a)} \\ &= \prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} f(x) && \text{(5.76b)} \\ &= \prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} [a]_n x && \text{(5.76c)} \\ &= \Big(\prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} [a]_n\Big)\Big(\prod_{x\in(\mathbb{Z}/n\mathbb{Z})^\times} x\Big) && \text{(5.76d)} \\ &= [a]_n^{\operatorname{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right)} P && \text{(5.76e)} \end{aligned}$$

En simplifiant par P dans $\mathbb{Z}/n\mathbb{Z}$ nous trouvons

$$[a]_n^{\operatorname{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right)} = [1]_n. \tag{5.77}$$

Et comme le lemme 5.32 donne $\operatorname{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right) = \varphi(n)$, nous avons trouvé $[a]_n^{\varphi(n)} = [1]_n$. Autrement dit,

$$a^{\varphi(n)} \in [1]_n. \tag{5.78}$$

Le point (2) n'est qu'une reformulation. Vu que A et B sont premiers entre eux, nous venons de voir que $A^{\varphi(B)} \in [1]_B$. Il existe donc $m \in \mathbb{Z}$ tel que $A^{\varphi(B)} = 1 + mB$. Vu que A et B sont dans $\mathbb{N}$, le nombre m est contraint d'être positif, c'est-à-dire $m \in \mathbb{N}$. □

LEMooKPKBooPbrHkI

Lemme 5.35 ([145]).
Soient $d \mid n$ dans $\mathbb{N}^*$. Nous considérons le groupe additif $G_d = \{k[q]_n$ tel que $0 \leqslant k \leqslant d-1\}$. Les éléments d'ordre [10] d dans $(\mathbb{Z}/n\mathbb{Z}, +)$ sont les générateurs de G_d.

Démonstration. Les générateurs de G_d sont d'ordre d parce que $|G_d| = d$. Ça, c'était le sens facile. Dans l'autre sens, si $[r]_n$ est d'ordre d, alors $d[r]_n = [0]_n$. D'après le lemme 5.33, cela prouve que $[r]_n \in G_d$.

Comme le groupe engendré par $[r]_n$ est d'ordre d, il est tout G_d. Donc $[r]_n$ est générateur de G_d. □

9. Proposition 1.309. Pour rappel, le produit n'est rien d'autre qu'une somme pour un groupe dont la loi est notée multiplicativement.

10. Ordre d'un élément, définition 1.268.

LEMooQGGLooDkkmcF

Lemme 5.36 ([148]).
L'ensemble des générateurs de $(\mathbb{Z}/n\mathbb{Z},+)$ est $(\mathbb{Z}/n\mathbb{Z})^\times$.

Démonstration. Si $[r]_n$ est générateur de $\mathbb{Z}/n\mathbb{Z}$, il existe k tel que $k[r]_n=[1]_n$. Dans ce cas, $[k]_n$ est l'inverse de $[r]_n$ pour la multiplication. Donc $[r]_n\in(\mathbb{Z}/n\mathbb{Z})^\times$.

À l'inverse, si $[r]_n$ est inversible, alors il existe k tel que $k[r]_n=[1]_n$. Dans ce cas, $[t]_n=kt[r]_n$, ce qui montre que $[r]_n$ est générateur (pour l'addition). □

LEMooRMWRooRSjGPL

Lemme 5.37 ([148]).
Si G est un groupe cyclique d'ordre n, alors G possède $\varphi(n)$ générateurs.

Démonstration. Vu que tous les groupes cycliques d'ordre n sont isomorphes à $(\mathbb{Z}/n\mathbb{Z}),+$, nous nous contentons de prouver le résultat pour ce groupe. Le lemme 5.36 montre que $(\mathbb{Z}/n\mathbb{Z},+)$ possède $\text{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right)$ générateurs.

Mais le lemme 5.32 assure que $\text{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right)=\varphi(n)$. □

PROPooYHUDooUROTiN

Proposition 5.38.
Nous avons la formule

$$n=\sum_{d|n}\varphi(d). \tag{5.79}$$

EqTPHogJ

Démonstration. Nous notons H_d la partie de $(\mathbb{Z}/n\mathbb{Z},+)$ composée des éléments d'ordre d. Nous avons vu dans le lemme 5.35 que H_d sont justement les générateurs de G_d – voir le lemme pour la notation. Mais comme G_d est un groupe cyclique d'ordre d, il contient $\varphi(d)$ générateurs (lemme 5.37) : $\text{Card}(H_d)=\varphi(d)$.

Vu que tous les éléments de $\mathbb{Z}/n\mathbb{Z}$ ont un ordre qui divise n (corolaire 2.14), nous avons l'union disjointe

$$\mathbb{Z}/n\mathbb{Z}=\bigcup_{d|n}H_d, \tag{5.80}$$

et donc au niveau des cardinals,

$$n=\text{Card}(\mathbb{Z}/n\mathbb{Z})=\sum_{d|n}\text{Card}(H_d)=\sum_{d|n}\varphi(d). \tag{5.81}$$

□

LEMooBEJOooDqTirj

Lemme 5.39.
Si p est un nombre premier, alors $\varphi(p^n)=p^n-p^{n-1}$.

Démonstration. Les éléments de $\{1,\dots,p^n\}$ qui ont un pgcd différent de 1 avec p^n sont des nombres qui s'écrivent sous la forme qp avec $q\leqslant p^{n-1}$ [11]. Il y a évidemment p^{n-1} tels nombres.

Par conséquent le cardinal de P_{p^n} est $\varphi(p^n)=p^n-p^{n-1}$. □

¡¡ Avertissement/question à la lectrice !! 5.40
P_n n'a pas été défini.

Définition proposée (et vue par après) : $P_n=\{m\in\mathbb{N}$ tel que $\text{pgcd}(m,n)=1\}$. À mettre donc en lien avec Δ_d.

5.5.2 Générateurs

PropZnmuphiGensn

Proposition 5.41.
Soit $n\in\mathbb{N}\backslash\{0\}$ et le groupe (additif) $\mathbb{Z}/n\mathbb{Z}$. L'élément $[x]_n$ est un générateur de $\mathbb{Z}/n\mathbb{Z}$ si et seulement si $x\in P_n$. En particulier $\mathbb{Z}/n\mathbb{Z}$ est un groupe contenant $\varphi(n)$ générateurs.

11. Corolaire 3.23.

Démonstration. Nous avons $\text{gr}\big([1]_n\big) = \mathbb{Z}/n\mathbb{Z}$. L'élément $[x]_n$ sera générateur si et seulement si il génère $[1]_n$, c'est-à-dire si il existe u tel que $u[x]_n = [1]_n$. Cette dernière égalité étant une égalité de classes dans $\mathbb{Z}/n\mathbb{Z}$, elle sera vraie si et seulement si il existe v tel que

$$ux + vn = 1. \tag{5.82}$$

Cela signifie entre autres que[12] $x\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$, et aussi que $\text{pgcd}(x, n) = 1$ par le théorème de Bézout 1.236, et donc que $x \in P_n$. □

CORooMBLSooMHKmAq

Corolaire 5.42.
Un groupe monogène d'ordre n possède $\varphi(n)$ générateurs, où φ est la fonction indicatrice d'Euler définie en 5.29.

Démonstration. Le théorème 5.16 nous dit qu'un groupe monogène d'ordre n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. La proposition 5.41 nous indique que $\mathbb{Z}/n\mathbb{Z}$ possède $\varphi(n)$ générateurs. □

5.5.3 Fonction indicatrice d'Euler (propriétés)

subSecKGDFooAvEfjf

Corolaire 5.43.
Deux propriétés.

(1) L'indicatrice d'Euler est multiplicative : si p est premier avec q, alors

$$\varphi(pq) = \varphi(p)\varphi(q). \tag{5.83}$$

(2) Si p est un nombre premier,

$$\varphi(p) = (p - 1). \tag{5.84}$$

Démonstration. Nous savons que si p et q sont premiers entre eux, alors le théorème 5.25 nous donne l'isomorphisme de groupe

$$(\mathbb{Z}/pq\mathbb{Z}, +) \simeq (\mathbb{Z}/p\mathbb{Z}, +) \times (\mathbb{Z}/q\mathbb{Z}, +). \tag{5.85}$$

Un élément (x, y) est générateur du produit si et seulement si x est générateur de $\mathbb{Z}/p\mathbb{Z}$ et y est générateur de $\mathbb{Z}/q\mathbb{Z}$. Par la proposition 5.41, il y a $\varphi(p)\varphi(q)$ tels éléments. Par ailleurs le nombre de générateurs de $\mathbb{Z}/pq\mathbb{Z}$ est $\varphi(pq)$, d'où l'égalité.

Si p est premier, nous avons $\varphi(p) = p - 1$ parce que tous les entiers de $\{1, \dots, p - 1\}$ sont premiers avec p. □

5.6 Groupe symétrique, groupe alterné

SECooZFYQooFfopMa

La définition des permutations et du groupe symétrique sont 1.274. Voir aussi le thème 8.

5.6.1 Le groupe alterné

DEFooEIVIooFvVkHH

Définition 5.44.
Le groupe A_n des permutations paires[13] *dans S_n est le* ***groupe alterné****.*

PROPooCPXOooVxPAij

Proposition 5.45.
À propos du groupe alterné dans le groupe symétrique.

(1) Le groupe alterné A_n est un sous-groupe caractéristique[14] *de S_n*

ITEMooWXXUooOWvFgE

(2) Le sous-groupe A_n est d'indice 2 dans S_n.

12. Corolaire 1.238
13. Définition 1.296.
14. Définition 1.173.

ITEMooGGAHooRYgNqq

(3) Le sous-groupe A_n est l'unique sous-groupe d'indice[15] *2 de S_n.*

Démonstration. Soit $\alpha \in \mathrm{Aut}(S_n)$. Étant donné que $\epsilon \circ \alpha$ est un homomorphisme surjectif sur $\{-1,1\}$, par unicité de cet homomorphisme, nous avons $\epsilon \circ \alpha = \epsilon$, et donc $\alpha(A_n) = A_n$. Par le premier théorème d'isomorphisme 2.6, il existe un isomorphisme

$$f\colon S_n/\ker(\epsilon) \to \mathrm{Image}(\epsilon). \tag{5.86}$$

En égalant le nombre d'éléments nous avons $|S_n : \ker\epsilon| = |S_n : A_n| = 2$.

Nous prouvons maintenant l'unicité. Soit H un sous-groupe d'indice 2 dans S_n. Par le lemme 3.28, H est distingué et nous pouvons considérer le groupe S_n/H. Ce dernier ayant 2 éléments, il est isomorphe à $\{-1,1\}$. Soit θ l'isomorphisme. On note φ le morphisme canonique $\varphi\colon S_n \to S_n/H$:

$$S_n \xrightarrow{\varphi} S_n/H \xrightarrow{\theta} \{-1,1\}. \tag{5.87}$$

EqSZBETH

La composition $\theta\circ\varphi$ est alors un homomorphisme surjectif de S_n sur $\{-1,1\}$ et nous avons $\theta\circ\varphi = \epsilon$ par la proposition 1.300. L'enchainement (5.87) nous montre que $H = \ker(\theta\circ\varphi) = \ker(\epsilon) = A_n$. □

PROPooPSZVooSmAgPA

Proposition 5.46 ([149])**.**
Le groupe symétrique S_n peut être écrit comme un produit semi-direct[16] *du groupe alterné :*

$$S_n = A_n \times_\varphi \mathbb{Z}/2\mathbb{Z} \tag{5.88}$$

où l'action de $\mathbb{Z}/2\mathbb{Z}$ sur A_n est la conjugaison par $\sigma = (12)$, c'est-à-dire $\rho(-1)\tau = \sigma\tau\sigma^{-1}$.

Démonstration. Nous avons la suite exacte

$$1 \xrightarrow{i} A_n \xrightarrow{i} S_n \xrightarrow{\epsilon} \{\pm 1\} \longrightarrow 1 \tag{5.89}$$

où les i représentent des inclusions et ϵ est la signature définie en 1.297. Grâce à cette suite et au fait que la signature soit un isomorphisme à partir de la partie $\{\mathrm{Id}, \sigma\}$ (pour σ d'ordre 2, par exemple $\sigma = (12)$), le théorème 2.48 nous dit que

$$S_n \simeq A_n \times_\varphi \{\mathrm{Id}, \sigma\} \tag{5.90}$$

où φ est l'action adjointe de $\{\mathrm{Id}, \sigma\}$ sur A_n. □

PROPooZOWBooIMxxlj

Proposition 5.47.
Si $\beta \in S_n$ est une transposition, nous avons les égalités suivantes d'ensembles :

$$S_n = A_n \cup A_n\beta = A_n \cup \beta A_n. \tag{5.91}$$

Démonstration. Les parties A_n et βA_n ont le même nombre d'éléments. En effet, l'application

$$\begin{aligned} \varphi\colon A_n &\to A_n\beta \\ \sigma &\mapsto \sigma\beta \end{aligned} \tag{5.92}$$

est une bijection.

De plus ces deux ensembles sont disjoints à cause de la proposition 1.300. En effet si $\sigma \in A_n$, alors $\epsilon(\sigma) = 1$. Mais un élément de $A_n\beta$ est de la forme $\sigma\beta$ avec $\sigma \in A_n$. Or ϵ est un homomorphisme, donc $\epsilon(\sigma\beta) = \epsilon(\sigma)\epsilon(\beta) = -1$.

Enfin, la proposition 5.45(2) dit que A_n est d'indice deux dans S_n. Donc la partie

$$A_n \cup A_n\beta \tag{5.93}$$

contient $|S_n|/2 + |S_n|/2 = |S_n|$ éléments. C'est donc S_n. □

15. Définition 2.12.
16. Définition 2.47.

LemiApyfp

Lemme 5.48.
Le groupe dérivé du groupe symétrique est le groupe alterné : $D(S_n) = A_n$.

Démonstration. Tout élément de $D(S_n)$ s'écrit sous la forme $ghg^{-1}h^{-1}$. Quel que soit le nombre de transpositions dans g et h, le nombre de transpositions dans $[g, h]$ est pair. □

PropsHlmvv

Proposition 5.49 ([150])**.**
Soit $n \geqslant 3$. Les 3-cycles $c_i = (1, 2, i)$ avec $i = 3, \dots, n$ engendrent le groupe alterné A_n.

Démonstration. Soit H, le groupe engendré par les c_i. D'abord nous avons

$$c_i = (1, 2, i) = (1, 2)(2, i), \tag{5.94}$$

de telle sorte que $\epsilon(c_i) = 1$. Par conséquent nous avons $H \subset A_n$. Nous montrons par récurrence que $A_n \subset H$.

Pour $n = 3$ il suffit de vérifier que $A_3 = \{\mathrm{Id}, c_3, c_3^2\}$. Supposons avoir obtenu le résultat pour A_{n-1}, et prouvons le pour A_n. Soit $s \in A_n$.

Si $s(n) = n$, alors s se décompose de la même manière que sa restriction s' à $\{1, \dots, n-1\}$. Par l'hypothèse de récurrence, cette restriction, appartenant à A_{n-1}, se décompose en produit des $c_3, \dots, c_{n-1}$ et de leurs inverses.

Si $s(n) = k$ alors nous considérons l'élément $c_n^2 c_k s$. Cet élément envoie n sur n et peut donc être décomposé avec les c_i $(i = 1, \dots, n-1)$ en vertu du point précédent. □

PropiodtBG

Proposition 5.50.
Lorsque $n \geqslant 5$, tous les 3-cycles de A_n sont conjugués. Autrement dit, la classe de conjugaison d'un 3-cycle est l'ensemble des 3-cycles.

Démonstration. Soient les 3-cycles $\sigma = (i_1, i_2, i_3)$ et $\varphi = (j_1, j_2, j_3)$. Nous considérons une bijection α de $\{1, \dots, n\}$ telle que $\alpha(i_s) = j_s$. Nous avons immédiatement que $\alpha \in S_n$ et que $\alpha\sigma\alpha^{-1} = \varphi$. Donc les 3-cycles sont conjugués dans S_n. Il reste à prouver qu'ils le sont dans A_n.

Si α est une permutation paire, la preuve est terminée. Si α est impaire, alors nous devons un peu la modifier. Comme $n \geqslant 5$, nous pouvons prendre s et t, des éléments distincts dans $\{1, \dots, n\}\backslash\{j_1, j_2, j_3\}$ et poser $\tau = (st)$. Puisque la signature est un homomorphisme et que τ et α sont impairs, l'élément $\tau\alpha$ est pair (lemme et proposition 1.299 et 1.295) et est donc dans A_n. Les supports de τ et φ étant disjoints, ces derniers commutent et nous avons

$$(\tau\alpha)\sigma(\tau\alpha)^{-1} = \tau(\alpha\sigma\alpha^{-1})\tau^{-1} = \tau\varphi\tau^{-1} = \varphi. \tag{5.95}$$

Donc σ et φ sont conjugués par $\tau\alpha$ qui est dans A_n. □

ThoURfSUXP

Théorème 5.51 ([54])**.**
Le groupe alterné A_n est simple[17] *pour $n \geqslant 5$.*

Démonstration. Soit N, un sous-groupe normal de A_n non réduit à l'identité. Étant donné que les 3-cycles engendrent A_n (proposition 5.49) et que tous les 3-cycles sont conjugués dans A_n (proposition 5.50), il suffit de montrer que N contient un 3-cycle. En effet si N contient un 3-cycle, le fait qu'il soit normal implique (par conjugaison) qu'il les contienne tous et donc qu'il contient une partie génératrice de A_n.

Soit donc $\sigma \in N$ différent de l'identité. Nous prenons i dans le support de σ et $j = \sigma(i)$. Nous choisissons ensuite $k \in \{1, \dots, n\}\backslash\{i, j, \sigma^{-1}(i)\}$ et $m = \sigma(k)$. Nous considérons la permutation $\alpha = (ijk)$. Étant donné que N est normal, l'élément

$$\theta = (\alpha^{-1}\sigma\alpha)\sigma^{-1} \tag{5.96}$$

17. Pas de sous-groupes normaux non triviaux, définition 1.175.

est dans N. De plus en utilisant le lemme 1.290 et le fait que $\alpha^{-1} = (ikj)$ nous avons

$$\theta = (ikj)(j\sigma(j)m). \tag{5.97}$$

Cela n'est pas spécialement un 3-cycle, mais nous allons en construire un. Nous allons déterminer que θ est soit un 5-cycle, soit un 3-cycle , soit un 2×2-cycle suivant les valeurs de $\sigma(j)$ et m.

Souvenons-nous que nous avons :

— $i \neq j = \sigma(i)$, puisque i est dans le support de σ ;

— $k \neq i$ et $k \neq j$, par définition de k (rappelons aussi que $k \neq \sigma^{-1}(i)$) ;

— $m \neq i$, $m \neq j$ et $m \neq \sigma(j)$ puisque $m = \sigma(k)$.

Il ne nous reste alors seulement les deux possibilités suivantes :

(1) soit $m = k$, soit $m \neq k$, d'une part ;

(2) soit $\sigma(j) = i$, soit $\sigma(j) = k$, soit $\sigma(j)$ n'est ni i, ni k, ni m, d'autre part.

Supposons dans un premier temps que $m = k$; alors

$$\theta = (ik)(j\sigma(j)). \tag{5.98}$$

C'est à priori un 2×2-cycle. Mais si de plus $\sigma(j) = i$, alors

$$\theta = (ijk) \tag{5.99}$$

qui est un 3-cycle ; et si $\sigma(j) = k$, alors

$$\theta = (ikj) \tag{5.100}$$

qui est un autre 3-cycle.

Supposons à présent que $m \neq k$. Si $\sigma(j)$ n'est ni i, ni k, ni m, alors i, j, k, $\sigma(j)$ et m sont cinq nombres différents, et

$$\theta = (i, j, \sigma(j), m, k) \tag{5.101}$$

est un 5-cycle. Si $\sigma(j) = i$, alors

$$\theta = (ikj)(jim) = (imk) \tag{5.102}$$

qui est un 3-cycle. Si $\sigma(j) = k$, alors

$$\theta = (ikj)(jkm) = (ikm) \tag{5.103}$$

qui est encore un 3-cycle.

Bref nous avons montré que θ est soit un 3-cycle, soit un 5-cycle, soit un 2×2-cycle. Si θ est un 3-cycle, la preuve est terminée.

Si $\theta = (ab)(cd)$, alors on considère $e \in \{1, \ldots, n\} \backslash \{a, b, c, d\}$ et nous avons

$$\underbrace{(abe)^{-1}\theta(abe)}_{\in N}\theta^{-1} = (aeb)(ab)(cd)(abe)(an)(cd) = (abe) \in N. \tag{5.104}$$

Si θ est le 5-cycle $(abcde)$, alors l'élément suivant est dans N :

$$(abc)^{-1}\theta(abc)\theta^{-1} = (acb)(abcde)(abc)(aedcb) = (acd). \tag{5.105}$$

Dans tous les cas nous avons trouvé un 3-cycle dans N et nous avons par conséquent $N = A_n$, ce qui fait que A_n ne contient pas de sous-groupes normaux non triviaux. Le groupe alterné A_n est donc simple. □

Nous en déduisons immédiatement que si $n \geqslant 5$, le groupe dérivé de A_n est A_n parce que A_n ne contient pas d'autres sous-groupes non triviaux.

LEMooICEHooGSSpkq

Lemme 5.52.
Le groupe alterné[18] A_6 *n'accepte pas de sous-groupes normaux d'ordre* 60.

18. Définition 5.44.

Démonstration. Soit G normal dans A_6, et a, un élément d'ordre 5 dans G (qui existe parce que 5 divise 60). Soit aussi un élément b d'ordre 5 dans A_6. Les groupes $\text{gr}(a)$ et $\text{gr}(b)$ sont deux 5-Sylow dans A_6. En effet, 5 est un nombre premier, et est la plus grande puissance de 5 dans la décomposition de 60 ; donc $\text{gr}(a)$ est un 5-Sylow dans G. D'autre part, l'ordre de A_6 (qui est $\frac{1}{2}\cdot 6!$) ne possède également que 5 à la puissance 1 dans sa décomposition.

En vertu du théorème de Sylow 5.11(3), les 5-Sylow $\text{gr}(a)$ et $\text{gr}(b)$ sont conjugués et il existe $\tau\in A_6$ tel que $b=\tau a\tau^{-1}$. Mais G étant normal dans A_6, l'élément $\tau a\tau^{-1}$ est encore dans G, de telle sorte que $b\in G$. Du coup G doit contenir tous les éléments d'ordre 5 de A_6.

Les éléments d'ordre 5 de A_6 doivent fixer un des points de $\{1,2,3,4,5,6\}$ puis permuter les autres de façon à n'avoir qu'un seul cycle. Un cycle correspond à écrire les nombres $1,2,3,4,5$ dans un certain ordre. Ce faisant, le premier n'a pas d'importance parce qu'on considère la permutation cyclique, par exemple $(3,5,2,1,4)$ est la même chose que $(5,2,1,4,3)$. Le nombre de cycles sur $\{1,2,3,4,5\}$ est donc de 4!, et par conséquent le nombre d'éléments d'ordre 5 dans A_6 est $6\cdot 4!=144$.

Le groupe G doit contenir au moins 144 éléments alors que par hypothèse il en contient 60 ; contradiction. □

Le théorème suivant montre que tout groupe peut être vu, en agissant sur lui-même, comme une partie du groupe symétrique.

Théorème 5.53.
Un groupe G est isomorphe à un sous-groupe de son groupe symétrique $S(G)$.

Démonstration. Nous considérons φ, la translation à gauche :

$$\begin{aligned}\varphi\colon G&\to S(G)\\ g&\mapsto t_g\end{aligned}\tag{5.106}$$

où $f_g(h)=gh$. Étant donné que

$$\varphi(gh)=ghx=g(t_hx)=t_g\circ t_h(x),\tag{5.107}$$

l'application φ est un morphisme de groupes. Il est injectif parce que si $gx=hx$ pour tout x, en particulier pour $x=e$ nous trouvons $g=h$.

De la même manière, $\varphi(g)x=\varphi(g)y$ implique $x=y$. Cela montre que l'image est bien dans le groupe symétrique.

L'ensemble $\text{Image}(\varphi)$ est donc un sous-groupe de $S(G)$, et φ est un isomorphisme vers ce groupe. □

LEMooMVUGooRiDaDz

Lemme 5.54.
Si $n\geqslant 3$, alors

(1) Le centre de S_n est trivial.

(2) Le groupe S_n est non abélien.

Démonstration. Soit $s\in Z(S_n)$ et trois éléments distincts a, b et c de $\{1,\dots,n\}$. Nous posons $\tau=(ab)$ et nous avons $s\tau=\tau s$. En notant $a'=s(a)$ et $b'=s(b)$ nous avons

$$a'=s(a)=(\tau s\tau^{-1})(a)=(\tau s)(b)=\tau(b')\tag{5.108a}$$
$$b'=s(b)=(\tau s\tau^{-1})(b)=(\tau s)(a)=\tau(a').\tag{5.108b}$$

Donc τ permute a' et b'. Mais comme τ ne permute que a et b, en tant qu'ensembles, $\{a,b\}=\{s(a),s(b)\}$. Le même raisonnement sur $\{b,c\}$ donne $\{b,c\}=\{s(b),s(c)\}$. Et puisque a, b et c sont distincts,

$$\{b\}=\{b,c\}\cap\{a,b\}=\{s(b)\}.\tag{5.109}$$

Cela montre que $s(b)=b$, et donc que le centre de S_n est réduit à la permutation identité.

En ce qui concerne le fait que S_n est non abélien, si nous avions $st = ts$ pour tout $s, t \in S_n$ alors $s = tst^{-1}$ pour tout t. Alors s serait dans le centre de S_n. En bref, si S_n était abélien, son centre serait S_n et non $\{\mathrm{Id}\}$.

□

PROPooUBIWooTrfCat

Proposition 5.55 ([151, 141]).
Tout groupe simple[19] *d'ordre* 60 *est isomorphe au groupe alterné* A_5.

Démonstration. Nous avons la décomposition en nombres premiers $60 = 2^2 \cdot 3 \cdot 5$. Déterminons pour commencer le nombre n_5 de 5-Sylow dans G. Le théorème de Sylow 5.11(4) nous renseigne que n_5 doit diviser 60 et doit être égal à $1 \mod 5$. Les deux seules possibilités sont $n_5 = 1$ et $n_5 = 6$. Étant donné que tous les p-Sylow sont conjugués, si $n_5 = 1$ alors le 5-Sylow serait un sous-groupe invariant à l'intérieur de G, ce qui est impossible vu que G est simple. Donc $n_5 = 6$.

Par le point (3) du théorème de Sylow, le groupe G agit transitivement sur l'ensemble des 5-Sylow par l'action adjointe :

$$g \cdot S = gSg^{-1}. \tag{5.110}$$

Cela donne donc un morphisme $\theta\colon G \to S_6$. Le noyau de θ est un sous-groupe normal. En effet si $k \in \ker\theta$ et si $g \in G$ nous avons

$$\begin{aligned}(gkg^{-1}) \cdot S &= gkg^{-1}Ggk^{-1}g^{-1} && \text{(5.111a)}\\ &= gkTk^{-1}g^{-1} && \text{(5.111b)}\\ &= gTg^{-1} && \text{(5.111c)}\\ &= S && \text{(5.111d)}\end{aligned}$$

où T est le Sylow $T = g^{-1}Sg$. Étant donné que $k \in \ker\theta$ nous avons utilisé $kTk^{-1} = aT$. Au final $gkg^{-1} \cdot S = S$, ce qui prouve que $gkg^{-1} \in \ker\theta$.

Étant donné que $\ker\theta$ est normal dans G, soit il est réduit à $\{e\}$ soit il vaut G. La seconde possibilité est exclue parce qu'elle reviendrait à dire que G agit trivialement, ce qui n'est pas correct étant donné qu'il agit transitivement. Nous en déduisons que $\ker\theta = \{e\}$, que θ est injective et que G est isomorphe à un sous-groupe de S_6.

Par ailleurs le groupe dérivé de G est un sous-groupe normal (et non réduit à l'identité parce que G est non commutatif). Donc $D(G) = G$. Étant donné que $G \subset S_6$, nous avons

$$G = D(G) \subset D(S_6) = A_6 \tag{5.112}$$

parce que le groupe dérivé du groupe symétrique est le groupe alterné (lemme 5.48).

L'ensemble $\theta^{-1}(A_6)$ est distingué dans G. En effet si $\sigma \in A_6$ et si $g \in G$ nous avons

$$\theta\big(g\theta^{-1}(\sigma)g^{-1}\big) = \theta(g)\sigma\theta(g)^{-1} \in A_6. \tag{5.113}$$

Nous en déduisons que $\theta^{-1}(A_6)$ est soit G entier soit réduit à $\{e\}$. Si $\theta^{-1}(A_6) = \{e\}$, alors pour tout $g \in G$ nous aurions $g^2 = e$ parce que $\theta(g^2) \in A_6$. L'ordre de G étant 60, il n'est pas possible que tous ses éléments soient d'ordre 2. Nous en déduisons que $\theta(G) \subset A_6$.

Nous nommons $H = \theta(G)$ et nous considérons l'ensemble $X = A_6/H$ où les classes sont prises à gauche, c'est-à-dire

$$[\sigma] = \{h\sigma \text{ tel que } h \in H\}. \tag{5.114}$$

Évidemment A_6 agit sur X de façon naturelle. Au niveau de la cardinalité,

$$\mathrm{Card}(X) = \frac{|A_6|}{|G|} = \frac{360}{60} = 6. \tag{5.115}$$

Le groupe A_6 agit sur X qui a 6 éléments. Nous avons donc une application $\varphi\colon A_6 \to A_6$. Encore une fois, la simplicité de A_6 montre que $\varphi(A_6) = A_6$.

19. Définition 1.175.

Nous étudions maintenant $\varphi(H)$ agissant sur X. Un élément $x \in A_6$ fixe la classe de l'unité $[e]$ si et seulement si $x \in H$ et par conséquent $\varphi(H)$ est le fixateur de $[e]$ dans X. À la renumérotation près, nous pouvons identifier $\varphi(H)$ au sous-groupe de A_6 agissant sur $\{1,\dots,6\}$ et fixant 6. Nous avons alors $\varphi(H) = S_5 \cap A_6 = A_5$. Nous venons de prouver que φ fournit un isomorphisme entre A_5 et H. Étant donné que H était isomorphe à G, nous concluons que G est isomorphe à A_6. □

5.6.2 Sous-groupes normaux

NORMooQAZTooBQLqDn

5.56 ([152])**.**
Soit le groupe V_4 engendré par les doubles transpositions de S_4. Nous savons de l'exemple 1.294(5) que ce groupe contient exactement 3 éléments non triviaux et l'identité. De plus, comme c'est une classe de conjugaison, V_4 est normal dans S_4.

Lemme 5.57.
Les sous-groupes $\mathrm{Fix}_{S_n}(a)$ *(avec* $a \in \{1,\dots,n\}$*) sont conjugués entre eux.*

Démonstration. Soit $\sigma \in \mathrm{Fix}(a)$ et $s \in S_n$ nous devons prouver que $s\sigma s^{-1}$ est le fixateur d'un élément de $\{1,\dots,n\}$. Nous notons $s(a) = b$. Alors

$$(s\sigma s^{-1})(b) = (s\sigma)(a) = s(a) = b. \tag{5.116}$$

Donc $s\,\mathrm{Fix}(a)s^{-1} \subset \mathrm{Fix}(b)$.

Dans l'autre sens, si $\sigma \in \mathrm{Fix}(b)$ alors $s^{-1}\sigma s \in \mathrm{Fix}(a)$. Mais $\sigma = s(s^{-1}\sigma s)s^{-1}$, donc $\sigma \in s\,\mathrm{Fix}(a)s^{-1}$. □

PROPooOTJAooUbzGZm

Proposition 5.58 (Sous-groupes normaux de S_n [152])**.**
Les sous-groupes normaux de S_n *ne sont pas légions.*

(1) Pour $n = 4$*, les sous-groupes normaux de* S_4 *sont* $\{\mathrm{Id}\}$*,* V_4*,* A_4 *et* S_4*.*

(2) Pour $n \neq 4$*, les sous-groupes normaux de* S_n *sont* $\{\mathrm{Id}\}$*,* A_n *et* S_n*.*

Démonstration. Les cas $n \leqslant 2$ sont un peu triviaux, donc nous faisons $n \geqslant 3$. Soit H normal dans S_n et $s \neq \mathrm{Id}$ dans H ; par le lemme 5.54, s n'est pas dans le centre de S_n et il existe $u \in S_n$ tel que $us \neq su$. Comme u est un produit de transpositions (proposition 1.295), il existe une transposition t telle que $st \neq ts$. Le sous-groupe H est normal et puisque $s \in H$ nous avons aussi $ts^{-1}t^{-1} \in H$. Mais en même temps, la combinaison sts^{-1} est le conjugué d'une transposition et est donc également une transposition (classe de conjugaison de S_4 dans 1.294). Nous en concluons que $sts^{-1}t^{-1}$ est un produit de deux transpositions appartenant à H.

Nous venons de prouver que H contient au moins un produit de deux transpositions. Et ce produit est différent de Id parce que $sts^{-1}t^{-1} = \mathrm{Id}$ impliquerait $st = ts$.

Soient donc deux transpositions $t_1, t_2 \in H$ telles que $t_1t_2 \neq \mathrm{Id}$. Les supports de t_1 et t_2 ont soit 1 soit aucun élément communs.

(1) **Premier cas**
Supposons $t_1 = (a,b)$, $t_2 = (b,c)$ avec a,b,c distincts dans $\{1,\dots,n\}$. Dans ce cas $t_1t_2 = (a,b,c)$ et H contient un cycle de longueur 3. Puisque H est normal et que les cycles de longueur trois sont une classe de conjugaison (exemple 1.294) et que A_n est engendré par ceux-ci (proposition 5.49), $A_n \subset H$. Mais A_n est d'indice deux dans S_n (proposition (2)(2)). Quel nombre plus grand que $n!/2$ divise $n!$? Seulement n lui-même. Donc H est soit A_n soit S_n.

(2) **Second cas**
Le groupe H contient un élément de la forme $(ab)(cd)$ avec a,b,c,d distincts dans $\{1,\dots,n\}$.

(2a) **Si $n = 3$**
Impossible parce que avec $n = 3$ nous n'avons pas quatre éléments distincts.

(2b) **Si $n = 4$**

Le sous-groupe H de S_4 contient un élément de V_4 qui n'est pas l'identité. Par normalité et classes de conjugaison, H contient V_4. Nous devons maintenant prouver que si H n'est pas V_4 alors H est A_4 ou S_4. Nous avons les inclusions $V_4 \subset H \subset S_4$ et donc les inégalités

$$4 \leqslant |H| \leqslant 24. \tag{5.117}$$

Donc le nombre $|H|$ est un multiple de 4 qui divise 24. Les possibilités sont $|H| = 4, 8, 12, 24$. La possibilité $|H| = 4$ donne $H = V_4$; si $|H| = 24$ alors $H = S_4$; si $|H| = 12$ alors H est d'indice 2 dans S_4 et $H = A_n$ (proposition 5.45(3)). Quid de $|H| = 8$?

D'après le corolaire 2.14 au théorème de Lagrange, l'ordre d'un élément divise l'ordre du groupe. Soit x dans H mais pas dans V_4. L'ordre de x peut être 1, 2, 4 ou 8. Ordre 1 serait $x = \text{Id}$. Ordre 8, pas possible parce que S_4 n'a pas d'éléments d'ordre 8.

(2b.ii) **x d'ordre 2**

Prenons la décomposition de x en cycles disjoints. Puisqu'on est dans S_4, ces cycles ne peuvent être que des transpositions. Soit il y en a un (alors H contient une transposition et donc $H = S_4$), soit il y en a deux et alors x est dans V_4.

(2b.ii) **x d'ordre 4**

L'élément x est alors un cycle de longueur 4, et H contient tous les cycles de longueur 4 ; par exemple, le produit $(abcd)(bacd) = (adc)$. Le sous-groupe H contient alors A_4 (parce qu'il contient tous les 3-cycles).

(2c) **Si $n \geqslant 5$**

Soit un élément e [20] distinct de a, b, c et d. Par notre liste préférée des classes de conjugaison (exemple 1.294(5)), le 2-cycle $(c, e)(a, b)$ est conjugué à $(a, b)(c, d)$ et appartient donc à H. Mais alors le produit suivant est également dans H :

$$(ce)(ab)(ab)(cd) = (ce)(cd) = (ecd). \tag{5.118}$$

Donc H contient un 3-cycle, et par conséquent tous les 3-cycles. Encore une fois, cela prouve que H est soit A_n soit S_n.

(2d) **Pourquoi $n = 4$ est spécial ?**

Dans le premier cas, nous montrons tout de suite que $H = V_4$ n'est pas possible. Dans le deuxième cas, nous montrons que, grâce à un élément différent de a, b, c et d, la possibilité $H = V_4$ est exclue. La possibilité $H = V_4$ n'existe que pour $n = 4$.

□

5.6.3 Indice

THOooXDRNooIyaGlv

Théorème 5.59.
Tout sous-groupe d'indice n dans S_n est isomorphe à S_{n-1}.

Démonstration. Pour $n = 1$, il n'y a pas de sous-groupe. Pour $n = 2$, un sous-groupe d'indice 2 ne peut contenir que 1 élément, qui est donc l'identité. Ok pour que {Id} soit égal à S_1 ?

Pour les autres, il y a un peu plus de travail.

(1) **Pour $n = 3$**

Nous avons $|S_3| = 6$. Donc un sous-groupe d'indice 3 dans S_3 contient exactement 2 éléments. Il contient Id et un autre élément $\sigma \in S_3$ qui doit vérifier $\sigma^2 = \text{Id}$ ou $\sigma^2 = \sigma$. Aucun élément de S_3 ne vérifie $\sigma^2 = \sigma$ (à part l'identité). Donc $\sigma^2 = \text{Id}$, ce qui implique que σ est une transposition. Donc

$$H = \{\text{Id}, (12)\} \tag{5.119}$$

ou l'identité avec (23), ou avec (13). Dans tous les cas c'est isomorphe à S_2.

20. e n'est pas l'élément neutre ici

(2) **Pour $n = 4$**

Nous avons $|S_4 : H| = 4$, donc $|H| = 6$. Mais $6 = 2 \times 3$ et $2 \mid 3 - 1$, donc le théorème 5.25 nous dit que H est soit cyclique[21] (et donc abélien), soit un produit semi-direct. Vu que S_4 n'a pas d'éléments d'ordre 6, aucun sous-groupe d'ordre 6 ne peut être cyclique. Nous sommes donc dans le cas du produit semi-direct

$$H = \mathbb{Z}_3 \times_\varphi \mathbb{Z}_2 \tag{5.120}$$

EQooSHWGopEMNRxf

où $\varphi\colon \mathbb{Z}_2 \to \text{Aut}(\mathbb{Z}_3)$ et $\varphi(1)$ est d'ordre 2 dans $\text{Aut}(\mathbb{Z}_3)$. Il convient de nous attarder un peu pour être sûr d'avoir bien compris tout ce qui se trouve dans l'identification (5.120). D'abord un point de notations : ici nous considérons les groupes $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ munis de l'addition. Donc 1 n'est pas le neutre. Ensuite nous savons du théorème 5.18 que $\text{Aut}(\mathbb{Z}/3\mathbb{Z}) = (\mathbb{Z}/3\mathbb{Z})^*$, et que via cette identification, $\varphi(1) = 2 \in (\mathbb{Z}/3\mathbb{Z})^*$ au sens où $\varphi(1)x = 2x$. Nous avons alors $\varphi(1)^2 x = 4x = x$ dans $\mathbb{Z}/3\mathbb{Z}$. Cela montre bien que $\varphi(1)$ est d'ordre 2.

Par rapport à la proposition 5.46, ici nous écrivons $\mathbb{Z}_2 = \big(\{0,1\}, +\big)$ alors que là nous écrivons $\mathbb{Z}_2 = \big(\{-1,1\}, \cdot\big)$. Ce sont les mêmes groupes, mais il convient de remarquer que le 1 ici est le -1 là.

Nous savons par la proposition 5.46 que $S_n = A_n \times_\varphi \mathbb{Z}_2$; en comparant avec (5.120) nous voyons qu'il suffit de prouver que $A_3 = \mathbb{Z}/3\mathbb{Z}$ pour avoir $H = S_3$.

Le groupe A_3 possède $|S_3|/2 = 3$ éléments. Il est vite vu que $A_3 = \{\text{Id}, (12)(31), (12)(32)\}$: ce sont trois éléments de signature paire dans S_3 ; donc c'est S_3. La correspondance $\text{Id} \mapsto 0$, $(12)(13) \mapsto 1$, $(13)(12) \mapsto 2$ donne un isomorphisme avec $(\mathbb{Z}_3, +)$.

(3) **Pour $n \geqslant 5$**

Soit un sous-groupe H d'indice n dans S_n et l'action à gauche de S_n sur $E = S_n/H$ (qui n'est à priori pas un groupe) donnée par $g \cdot [s] = [gs]$.

(3a) **Morphisme $\varphi\colon S_n \to S_E$**

Le φ défini par l'action est un morphisme parce que

$$\varphi(g_1 g_2)[s] = [g_1 g_2 s] = \varphi(g_1)[g_2 s] = \varphi(g_1)\varphi(g_2)[s]. \tag{5.121}$$

Mais il faut également vérifier que pour chaque $g \in G$, l'application $\varphi(g)\colon E \to E$ est bien une permutation. Pour l'injectivité, si $\varphi(g)[s_1] = \varphi(g)[s_2]$ alors $[gs_1] = [gs_2]$, donc il existe $h \in H$ tel que $gs_1 = gs_2 h$, ce qui prouve que $s_1 = s_2 h$ et donc que $[s_1] = [s_2]$. Pour la surjectivité, soit $[t] \in S_n/H$ et résolvons $\varphi(g)[s] = [t]$ par rapport à s. L'élément $s = g^{-1}t$ convient.

(3b) **$\ker(\varphi)$ est normal**

Soit $z \in \ker(\varphi)$, c'est-à-dire que $\varphi(z) = \text{Id}_E$. Alors pour $\sigma \in S_n$ nous avons $\varphi(\sigma z \sigma^{-1}) = \varphi(\sigma)\varphi(z)\varphi(\sigma^{-1}) = \text{Id}_E$.

(3c) **$\ker(\varphi) = \bigcap_{g \in S_n} gHg^{-1}$**

Supposons que $z \in gHg^{-1}$ pour tout g, et calculons $\varphi(z)[s]$. D'abord par hypothèse il existe $h \in H$ tel que $z = shs^{-1}$, donc

$$\varphi(z)[s] = [zs] = [shs^{-1}s] = [sh] = [s], \tag{5.122}$$

ce qui prouve que $\varphi(z) = \text{Id}$.

Dans l'autre sens, soit $z \in \ker(\varphi)$. Donc $\varphi(z)[s] = [s]$. Il existe donc $h \in H$ tel que $zs = sh$, c'est-à-dire tel que $z = shs^{-1}$. La formule demandée est donc prouvée.

(3d) **Questions d'ordre**

Nous savons que $|H| = (n-1)!$ alors que $|A_n| = \frac{n!}{2}$. Donc $|H| < |A_n|$ avec une inégalité stricte. En même temps nous avons $|\ker(\varphi)| \leqslant |H|$ parce que $\ker(\varphi)$ est une intersection dont un des termes est H lui-même. Nous avons alors les inégalités

$$|\ker(\varphi)| \leqslant |H| = (n-1)! < |A_n|. \tag{5.123}$$

21. Définition 1.326.

Mais les seuls sous-groupes normaux de S_n sont A_n, S_n et $\{\mathrm{Id}\}$ (proposition 5.58). Donc $\ker(\varphi) = \mathrm{Id}$ et φ est une injection entre deux ensembles finis de même cardinalité. Cela fait de φ une bijection et donc un isomorphisme de groupes

$$\varphi\colon S_n \to S_E. \tag{5.124}$$

Soit une fonction de numérotation $\psi\colon E \to \{1,\dots,n\}$. Avec cela nous définissons un isomorphisme de groupes

$$\begin{aligned}\tilde{\psi}\colon S_E &\to S_n\\ \sigma &\mapsto \psi\sigma\psi^{-1}.\end{aligned} \tag{5.125}$$

(3e) **Fixateur**

Nous montrons à présent que $(\tilde{\psi}\circ\varphi)(H) = \mathrm{Fix}\big(\psi[\mathrm{Id}]\big)$ où le stabilisateur est pris dans S_n. Pour la première inclusion, soit $h \in H$. Nous avons $(\tilde{\psi}\circ\varphi)(h) = \psi\circ\varphi(h)\psi^{-1}$, qui nous appliquons à $\psi[\mathrm{Id}]$:

$$(\tilde{\psi}\circ\varphi)(h)\psi[\mathrm{Id}] = \psi\circ\varphi(h)[\mathrm{Id}] = \psi[h] = \psi[\mathrm{Id}]. \tag{5.126}$$

Donc $(\tilde{\psi}\circ\varphi)(H) \subset \mathrm{Fix}\big(\psi[\mathrm{Id}]\big)$.

Pour l'autre inclusion, soit $\sigma \in S_n$ tel que $\sigma\psi[\mathrm{Id}] = \psi[\mathrm{Id}]$. Puisque $\sigma \in S_n$ nous avons $s \in S_E$ tel que $\sigma = \tilde{\psi}(s)$. Pour ce s nous avons donc

$$\big(\tilde{\psi}(s)\circ\psi\big)[\mathrm{Id}] = \psi[\mathrm{Id}], \tag{5.127}$$

d'où nous déduisons $s[\mathrm{Id}] = [\mathrm{Id}]$. Cela prouve que s stabilise $[\mathrm{Id}]$ dans S_E. Donc $s = \varphi(h)$ pour un certain $h \in H$, et au final $\sigma = \tilde{\psi}\big(\varphi(h)\big)$.

(3f) **Conclusion**

L'application $\tilde{\psi}\circ\varphi\colon H \to S_n$ est une application dont l'image est le fixateur d'un point. Plus précisément,

$$\tilde{\psi}\circ\varphi\colon H \to \mathrm{Fix}\big(\psi[\mathrm{Id}]\big) \tag{5.128}$$

est un isomorphisme de groupe. Mais le stabilisateur d'un point dans S_n est S_{n-1}.

□

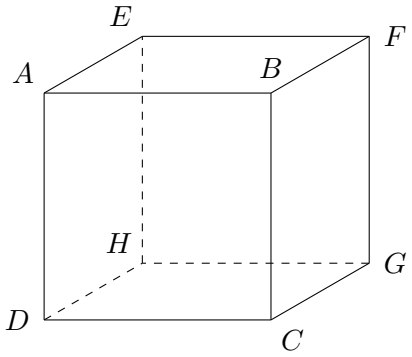
5.7 Isométries du cube

SecPVCmkxM

Les isométries du cube proviennent de [104].

Nous considérons le cube centré en l'origine de $\mathbb{R}^3$ et G, le groupe des isométries de $\mathbb{R}^3$ préservant ce cube. Nous notons aussi G^+ le sous-groupe de G constitué des éléments de déterminant positif. Nous notons

$$\mathcal{D} = \{D_1,\dots,D_4\} \tag{5.129}$$



l'ensemble des grandes diagonales, c'est-à-dire les segments $[AG]$, $[EC]$, $[FD]$, et $[BH]$. Nous savons que G préserve les longueurs et que ces segments sont les plus longs possibles à l'intérieur du cube. Donc G agit sur $\mathcal{D}$ parce qu'il ne peut transformer une grande diagonale qu'en une autre grande diagonale. Nous avons donc un morphisme de groupes

$$\rho\colon G \to S_4. \tag{5.130}$$

Nous montrons que ce morphisme est surjectif en montrant qu'il contient les transpositions. Le groupe G contient la symétrie axiale passant par le milieu M de $[A,E]$ et le milieu N de $[C,G]$. Il est facile de voir que cette symétrie permute $[AG]$ avec $[EC]$. De plus elle laisse $[FD]$ inchangée.

En effet, aussi incroyable que cela paraisse en regardant le dessin, nous avons $FD \perp MN$, parce qu'en termes de vecteurs directeurs,

$$\overrightarrow{ON} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \overrightarrow{OF} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}. \tag{5.131}$$

Étudions à présent le noyau $\ker(\rho)$. Si $f \in \ker(\rho)$ n'est pas l'identité, alors $f(D_i) = D_i$ pour tout i, mais au moins pour une des diagonales les sommets sont inversés. Quitte à renommer les sommets du cube nous supposons que la diagonale $[AG]$ est retournée : $f(A) = G$ et $f(G) = A$. Regardons où peut partir B sous l'effet de f. Étant donné que f préserve les diagonales, $f(B) \in \{B, C\}$, mais étant donné que f est une isométrie, $d\big(f(B), f(G)\big) = d(B, G)$, et nous concluons que $f(B) = H$. Donc la diagonale $[BH]$ est retournée sous l'effet de f. En raisonnant de même, nous voyons que f retourne toutes les diagonales. Donc les éléments non triviaux de $\ker(\rho)$ retournent toutes les diagonales ; il n'y en a donc qu'un seul et c'est la symétrie centrale :

$$\ker(\rho) = \{\mathrm{Id}, s_0\}. \tag{5.132}$$

Le premier théorème d'isomorphisme 2.6 nous permet d'écrire le quotient de groupes :

$$\frac{G}{\{\mathrm{Id}, s_0\}} \simeq S_4. \tag{5.133}$$

Une classe d'équivalence modulo $\ker(\rho)$ dans G est donc toujours de la forme $\{f, f \circ s_0\}$. Et comme s_0 est de déterminant -1, une classe contient toujours exactement un élément de déterminant 1 et un de déterminant -1.

D'autre part, $\ker(\rho)$ est normal dans G parce qu'en tant que matrice, $s_0 = -\mathbb{1}$, donc les problèmes de commutativité ne se posent pas. L'application

$$\begin{aligned} \varphi\colon \frac{G}{\{\mathrm{Id}, s_0\}} &\to G^+ \\ [g] &\mapsto \begin{cases} g & \text{si } \det(g) > 0 \\ g \circ s_0 & \text{sinon} \end{cases} \end{aligned} \tag{5.134}$$

est un isomorphisme de groupes. Et enfin nous pouvons écrire

$$G^+ \simeq S_4. \tag{5.135}$$

Nous allons maintenant utiliser le corolaire 2.49 pour montrer que $G = G^+ \times_\sigma \ker(\rho)$. Les conditions sont remplies :

— $\ker(\rho)$ normalise G^+ parce que $\ker(\rho)$ ne contient que $\pm\mathbb{1}$.

— $\ker(\rho) \cap G^+ = \{\mathrm{Id}\}$.

— $\ker(\rho)G^+ = G$ parce que les classes d'équivalence de G modulo $\ker(\rho)$ sont composées de $\{f, f \circ s_0\}$.

Puisque $G^+ \simeq S_4$ et $\ker(\rho) \simeq \mathbb{Z}/2\mathbb{Z}$ nous pouvons écrire de façon plus brillante que

$$G \simeq S_4 \times_\sigma \mathbb{Z}/2\mathbb{Z}. \tag{5.136}$$

Mais étant donné que la conjugaison par s_0 est triviale, le produit semi-direct est un produit direct :

$$G \simeq S_4 \times \mathbb{Z}/2\mathbb{Z}. \tag{5.137}$$

Il est maintenant du meilleur gout de pouvoir identifier géométriquement ces éléments. Les éléments de $\mathbb{Z}/2\mathbb{Z} = \{\mathrm{Id}, s_0\}$ ne font pas de mystère. Dans S_4 nous avons les classes de conjugaison des éléments Id, (12), (123), (1234) et (12)(34) déterminées durant l'exemple 1.294.

(1) L'élément (12) consiste à permuter deux diagonales et laisser les autres en place. Nous avons déjà vu que c'était une symétrie axiale passant par les milieux de deux côtés opposés. Cela fait 6 axes d'ordre 2.

(2) L'élément (123) fixe une des diagonales. C'est donc la symétrie axiale le long de la diagonale fixée. Par exemple la symétrie d'axe (AG) fait bouger le point B de la façon suivante :

$$B \to D \to E \to B. \tag{5.138}$$

C'est une rotation d'angle $\frac{2\pi}{3}$. En tout, nous avons donc 8 rotations d'ordre 3.

Notons à ce propos que la différence entre (234) et (243) est que la première réalise une rotation d'angle $2\pi/3$ tandis que la seconde, une rotation d'angle $-2\pi/3$.

(3) L'élément (1234) ne maintient aucune des diagonales et est d'ordre 4. C'est donc la rotation d'angle $\pi/2$ ou $-\pi/2$ autour de l'axe passant par les milieux de deux faces opposées. Il y en a 6 comme ça (3 paires de faces puis pour chaque il y a $\pi/2$ et $-\pi/2$), et ça tombe bien 6 est justement la taille de la classe de conjugaison de (1234) dans S_4.

(4) L'élément (12)(34) est le carré de la précédente[22], c'est-à-dire les rotations d'angle π autour des mêmes axes. Cela fait 3 éléments d'ordre 2.

22. En fait c'est (13)(24), le carré de la précédente, mais c'est la même classe de conjugaison.

Chapitre 6

Corps

6.1 Généralités

NORMooGPWRooIKJqqw

6.1.
Nous trouvons parfois le terme **anneau à division**. Cela provient du fait que dans beaucoup de cas on considère uniquement des corps commutatifs ; donc on voudrait une façon de parler d'un anneau dont tous les éléments non nuls sont inversibles. Dans ce cadre on dit :

— Un anneau à division est un anneau dont tous les éléments non nuls sont inversibles,

— Un corps est un anneau à division commutatif.

Pour prendre un exemple de cette différence, le théorème de Wedderburn 19.29 est énoncé ici sous les termes « Tout corps fini est commutatif ». Sous-entendu : la commutativité ne fait pas partie de la définition d'un corps. Par contre dans [104] il est énoncé sous les termes « Tout anneau à division fini est un corps ». Chez lui, un corps est toujours commutatif et un anneau à division est ce que nous appelons ici un corps.

6.1.1 Corps ordonnés

Nous avons vu la définition de corps totalement ordonné en 1.379.

Définition 6.2 ([73]).
Un corps est ***formellement réel*** *si -1 n'est pas une somme de carrés.*

Proposition 6.3.
Un corps totalement ordonné est formellement réel.

Démonstration. Soit un corps totalement ordonné $(\mathbb{K}, \leqslant)$ et $a \in \mathbb{K}$ alors $a^2 \geqslant 0$. En effet si $a \geqslant 0$ alors $a^2 = a \times a \geqslant 0$ directement par la définition 1.379(1b). Si $a \leqslant 0$ alors $-a \geqslant 0$ et

$$a^2 = (-a)^2 \geqslant 0. \tag{6.1}$$

Comme $-1 < 0$, il ne peut donc pas être écrit comme un carré. A fortiori comme somme de carrés. □

6.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$

PROPooLLPMooIVanaO

Proposition 6.4 ([153, 1]).
L'identité est l'unique automorphisme du corps $\mathbb{R}$.

Démonstration. Soit un automorphisme $\sigma : \mathbb{R} \to \mathbb{R}$. Comme pour tout automorphisme,

$$\sigma(a) = \sigma(1a) = \sigma(1)\sigma(a). \tag{6.2}$$

Donc $\sigma(1) = 1$.

(1) **Identité sur les rationnels** De plus

$$\sigma(n) = \sigma(1 + \ldots + 1) = \sigma(1) + \ldots + \sigma(1) = n, \tag{6.3}$$

et

$$\sigma\left(\frac{1}{n}\right) + \ldots + \sigma\left(\frac{1}{n}\right) = \sigma\left(\frac{1}{n} + \ldots + \frac{1}{n}\right) = \sigma(1) = 1. \tag{6.4}$$

Donc $\sigma(1/n) = 1/n$.

Nous en déduisons que pour tout $q \in \mathbb{Q}$, $\sigma(q) = q$. Cela ne suffit pas pour déduire $\sigma(x) = x$ pour tout $x \in \mathbb{R}$ parce que rien n'indique que σ soit continue.

(2) **Positive sur les positifs**
Si $x > 0$ alors $\sigma(x) = \sigma(\sqrt{x})^2 > 0$.

(3) **Croissance**
Si $x > y$ alors $x - y > 0$ et $\sigma(x - y) > 0$. Cela donne $\sigma(x) > \sigma(y)$.

(4) **Identité sur les réels**
Soit un irrationnel $x \in \mathbb{R}$ et une suite (q_i) dans $\mathbb{Q}$ qui converge de façon croissante vers x. Soit $\epsilon > 0$ dans $\mathbb{Q}$. Il existe N tel que si $i > N$ alors $q_i + \epsilon > x$; en appliquant σ à cette inégalité et en se souvenant que σ est l'identité sur $\mathbb{Q}$,

$$q_i + \epsilon > \sigma(x). \tag{6.5}$$

Mais de plus, $q_i < x$ donne $\sigma(q_i) < \sigma(x)$, c'est-à-dire $q_i < \sigma(x)$. En regroupant ces deux inégalités,

$$q_i < \sigma(x) < q_i + \epsilon \tag{6.6}$$

EQooLZOUooPhUNTI

pour tout $\epsilon > 0$ dans $\mathbb{Q}$ et $i > N$. Ce ϵ étant fixé nous pouvons prendre la limite des inégalités (6.6) :

$$x \leqslant \sigma(x) \leqslant x + \epsilon. \tag{6.7}$$

Cela étant valable pour tout $\epsilon > 0$ dans $\mathbb{Q}$, nous avons bien $x = \sigma(x)$.

□

REMooGHEDooOYYUPk

Remarque 6.5.
Certains[153] pensent que l'énoncé de cette proposition, ne parlant que de *corps* $\mathbb{R}$ n'autorise pas l'utilisation d'autre structure réelle que celle de corps. Du coup il faut reconstruire la notion d'ordre à partir seulement du langage des corps. Par exemple en disant que $a > b$ si et seulement si il existe k tel que $a = b + k^2$.

On peut s'en sortir en donnant l'énoncé suivant : « Si $\mathbb{K}$ est un corps isomorphe (en tant que corps) à $\mathbb{R}$ alors son unique automorphisme est l'identité ». Cela se démontre immédiatement en disant que si f est un automorphisme de $\mathbb{K}$ et si ϕ est un isomorphisme $\mathbb{K} \to \mathbb{R}$ alors $\phi \circ f \circ \phi^{-1}$ est un automorphisme de $\mathbb{R}$. Donc il est l'identité et f l'est également.

Attention cependant à prouver que ϕ^{-1} est un morphisme. En effet en posant $\phi^{-1}(x) = a$ et $\phi^{-1}(y) = b$ nous avons

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = x + y \tag{6.8}$$

parce que ϕ est un morphisme. D'autre part,

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = \phi(a + b). \tag{6.9}$$

Donc

$$\phi^{-1}(x + y) = \phi^{-1}\big(\phi(a) + \phi(b)\big) = \phi^{-1}\big(\phi(a + b)\big) = a + b = \phi^{-1}(x) + \phi^{-1}(y). \tag{6.10}$$

PROPooEATMooIPPrRV

Proposition 6.6.
Un automorphisme du corps $\mathbb{C}$ qui fixe $\mathbb{R}$ est soit l'identité soit la conjugaison complexe[1].

1. Par « fixer $\mathbb{R}$ » nous entendons que $\sigma(\mathbb{R}) = \mathbb{R}$, pas spécialement que $\sigma(x) = x$ pour tout $x \in \mathbb{R}$

Démonstration. Soit un automorphisme σ vérifiant la condition de fixer $\mathbb{R}$. Alors la restriction de σ à $\mathbb{R}$ est un automorphisme de $\mathbb{R}$ et y est donc l'identité par la proposition 6.4.

En ce qui concerne les imaginaires purs,

$$-1 = \sigma(-1) = \sigma(ii) = \sigma(i)^2. \tag{6.11}$$

Donc $\sigma(i)$ est un élément de $\mathbb{C}$ vérifiant $\sigma(i)^2 = -1$. C'est-à-dire $\sigma(i) = \pm i$.

Si $\sigma(i) = i$ alors $\sigma = \mathrm{Id}$. Si $\sigma(i) = -i$ alors σ est la conjugaison complexe. □

6.1.3 Corps premier

DEFooQSOFenIBKlq

Définition 6.7.
*Un corps est **premier** si il est son seul sous-corps. Le **sous-corps premier** d'un corps est l'intersection de tous ses sous-corps.*

LEMooPOGUooATyqcf

Lemme 6.8.
Le sous corps premier d'un corps est un corps.

Démonstration. Soit un corps $\mathbb{K}$ et son sous-corps premier $\mathbb{P}$. Si $x, y \in \mathbb{P}$, alors pour tout sous-corps $\mathbb{L}$ de $\mathbb{K}$ nous avons $x, y \in \mathbb{L}$ et donc $xy \in \mathbb{L} \subset \mathbb{P}$. Même type de vérification pour les autres propriétés d'un corps. □

Lemme 6.9.
Un corps premier est commutatif.

Démonstration. Le centre d'un corps est certainement un sous-corps. Par conséquent un corps premier doit être contenu dans son propre centre, c'est-à-dire être commutatif. □

DefXIHLooBAcqYH

Définition 6.10.
Soit p un nombre premier. Nous notons $\mathbb{F}_p = \mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$.

Nous verrons plus loin (section 19.4) comment nous pouvons définir $\mathbb{F}_{p^n}$ lorsque p est premier, ainsi que l'unicité d'un tel corps.

Nous avons par exemple

$$\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\} \tag{6.12}$$

avec la loi $2 = 0$.

Notons que $\mathbb{F}_p$ est un corps [2] possédant p éléments. L'ensemble $\mathbb{F}_p^*$ est un groupe d'ordre $p-1$.

Lemme 6.11.
Les corps $\mathbb{Q}$ et $\mathbb{Z}/p\mathbb{Z}$ (avec p premier) sont premiers.

Démonstration. Tout sous-corps de $\mathbb{Q}$ doit contenir 1, et par conséquent $\mathbb{Z}$. Devant également contenir tous les inverses, il contient $\mathbb{Q}$.

Tout sous-corps de $\mathbb{F}_p$ doit contenir 1 et donc $\mathbb{F}_p$ en entier. Par ailleurs nous savons de la proposition 3.63 que $\mathbb{F}_p$ est un corps lorsque p est premier. □

LEMooCGIYooSwOmHG

Lemme 6.12.
Si $k \in \mathbb{Z}$, nous notons $k_{\mathbb{K}}$ l'élément $k \cdot 1_{\mathbb{K}} = \sum_{i=1}^{k} 1_{\mathbb{K}}$. L'application

$$\begin{aligned} s\colon \mathbb{Z} &\to \mathbb{K} \\ k &\mapsto k_{\mathbb{K}} \end{aligned} \tag{6.13}$$

est un morphisme d'anneaux.

Si la caractéristique de $\mathbb{K}$ est 0, alors s est injective.

2. Quand p est premier, $\mathbb{Z}/p\mathbb{Z}$ est un corps, proposition 3.63.

PROPooXLLOooIAhEue

Proposition 6.13 (Sous-corps premier[154]).
Soit $\mathbb{K}$, un corps[3] *de caractéristique p.*

(1) Si $p > 0$ alors le sous-corps premier de $\mathbb{K}$ est isomorphe à $\mathbb{Z}$.

(2) Si $p = 0$ alors le sous-corps premier de $\mathbb{K}$ est isomorphe à $\mathbb{Q}$.

Démonstration. Pour les besoins de notations, dans la suite si $k \in \mathbb{Z}$, nous notons $k_{\mathbb{K}}$ l'élément $k \cdot 1_{\mathbb{K}} = \sum_{i=1}^{k} 1_{\mathbb{K}}$. Notons aussi $\mathbb{P}$ le sous-corps premier de $\mathbb{K}$.

Étant donné que $1_{\mathbb{K}}$ est dans tout sous-corps, nous devons avoir $\mathbb{Z}1_{\mathbb{K}} \subseteq \mathbb{P}$. Nous avons donc toujours

$$\mathbb{Z}1_{\mathbb{K}} \subset \mathbb{P} \subset \mathbb{K}. \tag{6.14}$$

Maintenant, le lemme 1.352 nous dit la nature de $\mathbb{Z}1_{\mathbb{K}}$ en tant qu'anneaux, et nous pouvons séparer les cas en fonction de la caractéristique.

(1) **Si $p > 0$** Nous avons $\mathbb{Z}1_{\mathbb{K}} \simeq \mathbb{Z}/p\mathbb{Z}$ par le lemme 1.352. L'ensemble $\mathbb{Z}/p\mathbb{Z}$ étant un corps, $\mathbb{Z}1_{\mathbb{K}}$ est un corps et c'est donc le corps premier de $\mathbb{K}$.

(2) **Si $p = 0$** Si $p = 0$, alors $\mathbb{Z}1 \simeq \mathbb{Z}$ en tant qu'anneaux. Dommage que $\mathbb{Z}$ ne soit pas un corps. Soit $n \in \mathbb{Z}1_{\mathbb{K}}$ et $m \in (\mathbb{Z}1_{\mathbb{K}})^*$. L'élément $nm^{-1} \in \mathbb{K}$ est dans $\mathbb{P}$ parce que $\mathbb{P}$ est un corps. Donc l'application

$$\begin{aligned} \psi\colon \mathbb{Q} &\to \mathbb{P} \\ ab^{-1} &\mapsto a_{\mathbb{K}} b_{\mathbb{K}}^{-1} \end{aligned} \tag{6.15}$$

prend bien ses valeurs dans $\mathbb{P}$ et est un morphisme de corps.

La surjectivité de ψ est ce dont nous venons de parler. En ce qui concerne l'injectivité, si $\psi(ab^{-1}) = \psi(xy^{-1})$, alors $a_{\mathbb{K}} b_{\mathbb{K}}^{-1} = x_{\mathbb{K}} y_{\mathbb{K}}^{-1}$ et la commutativité de $\mathbb{K}$ nous permet d'écrire

$$a_{\mathbb{K}} y_{\mathbb{K}} = x_{\mathbb{K}} b_{\mathbb{K}}. \tag{6.16}$$

Étant donné le lemme 6.12, nous déduisons que $(ay - xb)_{\mathbb{K}} = 0$. Et comme $z \mapsto z_{\mathbb{K}}$ est injective, nous en déduisons que $ay - xb = 0$ dans $\mathbb{Z}$, et donc que $ab^{-1} = xy^{-1}$ dans $\mathbb{Q}$ par le lemme 1.374.

□

PropqPPrgJ

Proposition 6.14.
Soit $\mathbb{K}$ un corps et $\mathbb{P}$ son sous-corps premier[4]*. Si $\sigma \in \mathrm{Aut}(\mathbb{K})$ alors $\sigma|_{\mathbb{P}} = \mathrm{Id}$, c'est-à-dire que $\sigma(x) = x$ pour tout $x \in \mathbb{P}$.*

Démonstration. Soit $\mathbb{P}$ le sous-corps premier de $\mathbb{K}$. La proposition 6.13 nous dit que $\mathbb{P}$ est soit $\mathbb{Z}$ soit $\mathbb{Q}$ en fonction de la caractéristique de $\mathbb{K}$. Nous allons faire le cas de caractéristique nulle, c'est-à-dire $\mathbb{P} \simeq \mathbb{Q}$.

Tout élément de $\mathbb{P}$ peut être écrit sous la forme $k_{\mathbb{K}} l_{\mathbb{K}}^{-1}$ pour des $k, l \in \mathbb{Z}$. Nous calculons la valeur de σ sur ces éléments. D'abord σ satisfait $\sigma(1_{\mathbb{K}})$ et préserve la somme, dont $\sigma(k_{\mathbb{K}}) = k_{\mathbb{K}}$ pour tout $k \in \mathbb{Z}$[5]. Étant donné que σ préserve également l'inverse et le produit,

$$\sigma(k_{\mathbb{K}} l_{\mathbb{K}}^{-1}) = \sigma(k_{\mathbb{K}})\sigma(l_{\mathbb{K}})^{-1} = k_{\mathbb{K}} l_{\mathbb{K}}^{-1}. \tag{6.17}$$

Et voila. □

3. Je rappelle que tous les corps du Frido sont commutatifs, sauf mention explicite du contraire.
4. Définition 6.7.
5. Dans le cas de caractéristique non nulle, $\mathbb{P} = \mathbb{Z}$ et vous pouvez vous arrêter ici.

6.1.4 Petit théorème de Fermat

ThoOPQOiO

Théorème 6.15 (Petit théorème de Fermat)**.**
Soit p un nombre premier.

(1) Si $x \in \mathbb{Z}/p\mathbb{Z}$ alors $x^p = x$.

(2) Si $x \in (\mathbb{Z}/p\mathbb{Z})^*$, alors $x^{p-1} = 1$.

(3) Si $x \in (\mathbb{Z}/p\mathbb{Z})^*$ alors $x^{-1} = x^{p-2}$.

ITEMooRNIVooOIzqgc

(4) Si x est premier avec p, alors $x^{p-1} \in [1]_p$.

Démonstration. Étant donné que $\mathbb{Z}/p\mathbb{Z}$ est un corps commutatif et que p est premier, la proposition 3.46 nous indique que $\sigma(x) = x^p$ est un automorphisme. La proposition 6.14 nous indique alors que

$$a^p = a. \tag{6.18}$$

Si a est inversible alors $a^{p-1} = a^p a^{-1} = 1$.

Pour (4). Le nombre x n'est pas premier avec p uniquement lorsque x est multiple de p. Dans ce cas c'est $[a]_p = 0$ dans $\mathbb{Z}/p\mathbb{Z}$ et donc $a^{p-1} = 0$. □

Exemple 6.16.
Soit $\mathbb{K} = \mathbb{F}_{29}$. Le nombre 29 étant premier, $\mathbb{K}$ est un corps. C'est le corps des entiers modulo 29. Nous avons donc

$$-142 = -113 = -84 = -55 = -26 = 3 = 32 = 61 = 90 = 119. \tag{6.19}$$

Le petit théorème de Fermat nous permet aussi de calculer des exposants et des inverses. En effet, puisque $1 = x^{28}$ pour tout $x \in \mathbb{F}_{29}^*$, nous avons $x^{-1} = x^{27}$, et par suite, pour tout entier a,

$$x^{-a} = (x^a)^{27} = x^{27a}. \tag{6.20}$$

Le nombre $27a$ peut être grand par rapport à 29. Mais en réutilisant le fait que $1 = x^{28}$, on obtient

$$x^{-a} = x^{[27a]_{28}}. \tag{6.21}$$

Cette expression doit être comprise comme disant que pour tout $k \in [27a]_{28}$ nous avons $x^{-a} = x^k$.

Chose à retenir : dans les exposants nous calculons modulo 28. △

6.1.5 Nombres de Sophie Germain

DEFooCVFJooKCdVVD

Définition 6.17.
*Un nombre premier p est de **Sophie Germain** si $p > 2$ et si le nombre $q = 2p + 1$ est également premier.*

Notons qu'il existe des nombres premiers de Sophie Germain. Par exemple $p = 3$ donne $q = 2 \times 3 + 1 = 7$. Comme 7 est un nombre premier, le nombre 3 est de Sophie Germain. D'après Wikipédia[155], l'existence d'une infinité de nombres premiers de Sophie Germain est encore une question ouverte[6].

LEMooIYIKooGkoqRJ

Lemme 6.18 ([156, 157])**.**
Soit p, un nombre premier de Sophie Germain. Si $m \in \mathbb{Z}$ n'est pas divisible par q, alors $m^p \in [\pm 1]_q$ où $q = 2p + 1$.

6. Si vous lisez ce paragraphe dans un futur où la question est tranchée, n'hésitez pas à m'écrire pour mettre à jour.

Démonstration. Le petit théorème de Fermat 6.15 nous dit que $m^{q-1} \in [1]_q$. Mais $m^{q-1} = m^{2p} = (m^p)^2$. Donc

$$(m^p)^2 \in [1]_q. \tag{6.22}$$

Puisque q est premier, l'anneau $\mathbb{Z}/q\mathbb{Z}$ est intègre par le corolaire 1.243. Comme $1 \neq -1$, la proposition 3.133 s'applique et nous avons $m^p \in [\pm 1]_q$. □

THOooSZXWooVeHdrh

Théorème 6.19 ([156, 157, 158])**.**
Soit un nombre premier de Sophie Germain (définition 6.17). Il n'existe pas de solution (x, y, z) dans $\mathbb{Z}^3$ au système

$$\begin{cases} x^p + y^p + z^p = 0 \\ [xyz]_p \neq 0. \end{cases} \tag{6.23}$$

Démonstration. Nous supposons que $(x, y, z) \in \mathbb{Z}^3$ est une solution telle que $\text{pgcd}(x, y, z) = 1$.

(1) **x, y et z sont premiers deux à deux** Supposons que k soit un diviseur premier commun à x et y. Alors il existe $\alpha, \beta \in \mathbb{Z}$ tels que $x = k\alpha$ et $y = k\beta$. Du coup on aurait $z^p = -k^p(\alpha^p + \beta^p)$. Donc k divise z^p. Le lemme 3.17 nous indique qu'alors k divise z. Donc $k = 1$ parce que k diviserait x, y et z.
Ce qui est vrai pour le couple (x, y) est encore vrai pour (x, z) et pour (y, z).

(2) **Le fameux u** Nous introduisons à présent une nouvelle variable intermédiaire qui va beaucoup nous aider. En utilisant le lemme 3.45,

$$-x^p = y^p + z^p = (y + z) \sum_{k=0}^{p-1} y^k (-z)^{p-1-k} = (y + z)u \tag{6.24}$$

où nous avons posé $u = \sum_{k=0}^{p-1} y^k(-z)^{p-1-k}$.

(3) **$\text{pgcd}(u, y + z) = 1$** Soit α un nombre premier qui divise u et $y + z$.
D'abord α divise x^p et donc x (lemme 3.17). Autrement dit : $[x]_\alpha = 0$.
Ensuite, $0 = [x + z]_\alpha$, de telle sorte que

$$[y]_\alpha = -[z]_\alpha \tag{6.25}$$

EQooPTICooBNDwpx

Enfin, en prenant la classe de u modulo α, et en substituant (6.25) :

$$0 = [u]_\alpha = \Big[\sum_{k=0}^{p-1} y^k (-z)^{p-1-k} \Big]_\alpha \tag{6.26a}$$

$$= \sum_{k=0}^{p-1} [y^k]_\alpha [y^{p-1-k}]_\alpha \tag{6.26b}$$

$$= \sum_{k=0}^{p-1} [y]_\alpha^{p-1} \tag{6.26c}$$

$$= p[y]_\alpha^{p-1}. \tag{6.26d}$$

Nous avons donc $p[y]_\alpha^{p-1} = 0$. Mais $\mathbb{Z}/\alpha\mathbb{Z}$ est un anneau intègre[7]. Donc la règle du produit nul s'applique[8] et nous avons deux possibilités : $[p]_\alpha = 0$ ou $[y]_\alpha^{p-1} = 0$.
D'abord α divise déjà x. Or $\text{pgcd}(x, y) = 1$. Donc α ne peut pas diviser y (à part si $\alpha = 1$ évidemment), et à fortiori pas y^{p-1} non plus. Nous en déduisons que $[p]_\alpha = 0$. Puisque p est premier, nous avons $p = \alpha$.
Mais nous avons déjà vu que α divisait x. Donc $0 = [x]_\alpha = [x]_p$. Cela n'est pas possible parce que $[xyz]_p \neq 0$.
Nous avons donc prouvé que $\text{pgcd}(u, y + z) = 1$.

7. Corolaire 1.243.
8. Proposition 1.197.

(4) **Utilisation d'un lemme**
Pour rappel, p est impair. Nous avons $(-x)^p = (y+z)u$. Le lemme 3.21 nous dit qu'il existe $a, \alpha \in \mathbb{Z}$ tels que

$$\begin{cases} y+z = a^p \\ u = \alpha^p. \end{cases} \tag{6.27}$$

EQooKMPZooCShIIT

Ce que nous venons de faire pour $y+z$ peut être fait pour les autres couples. Donc il existe $a, b, c \in \mathbb{Z}$ tels que

$$y + z = a^p \tag{6.28a}$$
$$x + z = b^p \tag{6.28b}$$
$$x + y = c^p. \tag{6.28c}$$

(5) **q divise un et un seul des x, y ou z** Supposons que q ne divise ni x, ni y ni z. Alors le lemme 6.18 dit que $x^p \in [\pm 1]_q$, $y^p \in [\pm 1]_q$ et $z^p \in [\pm 1]_q$. Donc les valeurs possibles pour $[x^p + y^p + z^p]_q$ sont $[\pm 3]_q$ et $[\pm 1]_q$.
Mais nous avons supposé que $x^p + y^p + z^p = 0$ et que $q \geqslant 5$ de telle sorte que $[\pm 3]_q \neq 0$. Contradiction. Donc q divise au moins un des trois x, y ou z.
Nous avons déjà montré que x, y et z étaient deux à deux premiers entre eux. Donc q ne peut diviser qu'un seul des trois.

(6) **q divise x** Jusqu'à présent x, y et z avaient des rôles symétriques. Maintenant nous supposons que q divise x.

(7) **$[y^p]_q = [\pm 1]_q$** Nous savons que $c^p = x + y$. En passant aux classes modulo q, nous avons

$$[c^p]_q = [x]_q + [y]_q = [y]_q \tag{6.29}$$

parce que nous avons choisi que q divise x. Nous avons maintenant les implications suivantes :
q ne divise pas y.
$\Rightarrow$ q ne divise pas c^p parce que $[y]_q = [c^p]_q$.
$\Rightarrow$ q ne divise pas c par le lemme 3.17.
$\Rightarrow$ $c^p \in [\pm 1]_q$ par le lemme 6.18.
$\Rightarrow$ $y \in [\pm 1]_q$ toujours parce que $[y]_q = [c^p]_q$.

(8) **$[z^p]_q = [\pm 1]_q$** Parce que les rôles de y et z sont symétriques. Donc ce que nous avons fait pour obtenir $[y^p]_q = [\pm 1]_q$ tient pour avoir $[z^p]_q = [\pm 1]_q$.

(9) **q divise $y+z$** Nous avons dit que q divisait x. Il divise donc aussi $2x$. Nous avons :

$$2x = (x+y) + (x+z) - (y+z) = c^p + b^p - a^p. \tag{6.30}$$

En passant au modulo q, à gauche nous avons $[2x]_q = 0$. Donc

$$[a^p]_q = [c^p]_q + [b^p]_q \tag{6.31a}$$
$$= [y]_q + [z]_q \tag{6.31b}$$
$$= [\pm 1]_q + [\pm 1]_q \tag{6.31c}$$
$$\in \{[2]_q, [-2]_q, [0]_q\}. \tag{6.31d}$$

En tout état de cause, $[a^p]_q$ n'est pas $[\pm 1]_q$. La contraposée du lemme 6.18 nous dit alors que q divise $a^p = y + z$.

(10) **$[y]_q = [-z]_q$** Il s'agit seulement du fait que q divise $y+z$.

(11) **$[u]_q = [p]_q$** Par définition, $u = \sum_{k=0}^{p-1} y^k (-z)^{p-1-k}$. On passe au modulo q, en substituant $[-z]_q$ par $[y]_q$:

$$[u]_q = \sum_{k=0}^{p-1} [y]_q^k [y]_q^{p-1-k} = \sum_{k=0}^{p-1} [y]_q^{p-1} = p[y]_q^{p-1} = [p]_q \tag{6.32}$$

La dernière égalité est le fait que $[y]_q = [\pm 1]_q$ et que $p-1$ est pair.

(12) **La contradiction** Vous vous souvenez de (6.27) qui disait que $u = \alpha^p$ pour un certain p ? Nous avons les implications suivantes :
$[\alpha^p]_q = [p]_q \neq 0 \Rightarrow q$ ne divise pas α^p.
$\Rightarrow q$ ne divise pas α.
$\Rightarrow \alpha^p \in [\pm 1]_q$.
$\Rightarrow [p]_q \in [\pm 1]_q$.
La dernière ligne est clairement fausse parce que $q > p$ et p n'est certainement pas égal à 1 ou -1.

Nous avons prouvé le théorème pour les triples (x, y, z) sans facteur commun. Si k était un diviseur commun de x, y et z, alors $(x/k, y/k, z/k)$ serait encore une solution, ce qui n'est pas le cas. □

6.2 Théorème des deux carrés

PropleGdaT

Proposition 6.20.
Soit p un nombre premier et P un élément de $\mathbb{F}_p[X]$. L'anneau $\mathbb{F}_p[X]/(P)$ est intègre si et seulement si P est irréductible dans $\mathbb{F}_p[X]$.

Démonstration. Supposons que P soit réductible dans $\mathbb{F}_p[X]$, c'est-à-dire qu'il existe $Q, R \in \mathbb{F}_p[X]$ tels que $P = QR$. Dans ce cas, $\bar{Q}$ est diviseur de zéro dans $\mathbb{F}_p[X]/(P)$ parce que $\bar{Q}\bar{R} = 0$.

Nous supposons maintenant que $\mathbb{F}_p[X]/(P)$ ne soit pas intègre : il existe des polynômes $R, Q \in \mathbb{F}_p[X]$ tels que $\bar{Q}\bar{R} = 0$. Dans ce cas le polynôme QR est le produit de P par un polynôme : $QR = PA$. Tous les facteurs irréductibles de A étant soit dans Q soit dans R, il est possible de modifier un peu Q et R pour obtenir $QR = P$, ce qui signifie que P n'est pas irréductible. □

6.2.1 Un peu de structure dans $\mathbb{Z}[i]$

LemSCAlICY

Lemme 6.21.
L'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a + bi &\mapsto a^2 + b^2 \end{aligned} \tag{6.33}$$

est un stathme euclidien pour $\mathbb{Z}[i]$.

Démonstration. Soient $t, z \in \mathbb{Z}[i]\backslash\{0\}$ dont le quotient s'écrit

$$\frac{z}{t} = x + iy \tag{6.34}$$

dans $\mathbb{C}$. Nous considérons $q = a + bi$ où a et b sont les entiers les plus proches de x et y. Si il y a *ex aequo*, on prend au hasard [9]. Alors nous avons

$$|\frac{z}{t} - q| \leqslant \frac{|1 + i|}{2} = \frac{\sqrt{2}}{2} < 1. \tag{6.35}$$

On pose $r = z - qt$ qui est bien un élément de $\mathbb{Z}[i]$. De plus

$$|r| = |z - qt| = |t||\frac{z}{t} - q| < |t|, \tag{6.36}$$

c'est-à-dire que $|r|^2 < |t|^2$ et donc $N(r) < N(t)$. □

Étant donné que $\mathbb{Z}[i]$ est euclidien, il est principal (proposition 1.254).

LemBMEIiiV

Lemme 6.22.
Les éléments inversibles de $\mathbb{Z}[i]$ sont $\{\pm 1, \pm i\}$.

9. Dans l'exemple 1.253, nous prenions toujours l'inférieur parce que le stathme tenait compte de la positivité.

Démonstration. Déterminons les éléments inversibles de $\mathbb{Z}[i]$. Si $z \in \mathbb{Z}[i]^*$, alors il existe $z' \in \mathbb{Z}[i]^*$ tel que $zz' = 1$. Dans ce cas nous aurions

$$1 = N(zz') = N(z)N(z'), \tag{6.37}$$

ce qui est uniquement possible avec $N(z) = N(z') = 1$, c'est-à-dire $z = \pm 1$ ou $z = \pm i$. Nous avons donc

$$\mathbb{Z}[i]^* = \{\pm 1, \pm i\}. \tag{6.38}$$

□

DEFooUCSHooJqGuVB

Définition 6.23 ([159])**.**
*Un **monoïde** est un triplet $(E, *, e)$ où E est un ensemble, e est un élément de E et $*\colon E \times E \to E$ est une loi de composition telle que pour tout $x, y \in E$,*

(1) $x * (y * z) = (x * y) * z$ *(associativité)*

(2) $e * x = x * e = x$ *(e est un neutre).*

Nous notons $\Sigma = \{a^2 + b^2 \text{ tel que } a, b \in \mathbb{N}\}$.

LemIBDPzMB

Lemme 6.24.
L'ensemble $\Sigma = \{a^2 + b^2$ tel que $a, b \in \mathbb{N}\}$ est un sous-monoïde[10] *de $\mathbb{N}$.*

Démonstration. Il suffit de prouver que si $m, n \in \Sigma$, alors le produit mn est également dans Σ. Soit N, le stathme euclidien sur $\mathbb{Z}[i]$ (celui donné par le lemme 6.21). Vu que $n \in \Sigma$, il existe $z \in \mathbb{Z}[i]$ tel que $N(z) = n$. Idem pour m : il existe $z' \in \mathbb{Z}[i]$ tel que $N(z') = m$. Nous avons évidemment $zz' \in \mathbb{Z}[i]$, et

$$N(zz') = N(z)N(z') = nm. \tag{6.39}$$

Donc nm est l'image de zz' par N, ce qui prouve que $nm \in \Sigma$. □

ThospaAEI

Théorème 6.25 (Théorème des deux carrés, version faible)**.**
Un nombre premier est somme de deux carrés si et seulement si $p = 2$ ou $p \in [1]_4$.

Démonstration. Soit p un nombre premier dans Σ. Si $a = 2k$, alors $a^2 = 4k^2$ et $a^2 = 0 \mod 4$. Si au contraire a est impair, $a = 2k + 1$ et $a^2 = 4k^2 + 1 + 4k = 1 \mod 4$. La même chose est valable pour b. Par conséquent, $a^2 + b^2$ est automatiquement $[0]_4$, $[1]_4$ ou $[2]_4$. Évidemment les nombres de la forme $0 \mod 4$ ne sont pas premiers ; parmi les $2 \mod 4$, seul $p = 2$ est premier (et vaut $1^2 + 1^2$).

Nous avons démontré que les seuls premiers de la forme $a^2 + b^2$ sont $p = 2$ et les $p = 1 \mod 4$. Il reste à faire le contraire : démontrer que si un nombre premier p vaut $1 \mod 4$, alors il est premier. Nous considérons l'anneau

$$\mathbb{Z}[i] = \{a + bi \text{ tel que } a, b \in \mathbb{Z}\}. \tag{6.40}$$

puis l'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a + bi &\mapsto a^2 + b^2. \end{aligned} \tag{6.41}$$

Un peu de calcul dans $\mathbb{C}$ montre que pour tout $z, z' \in \mathbb{Z}[i]$, $N(zz') = N(z)N(z')$.

Nous savons que les éléments inversibles de $\mathbb{Z}[i]$ sont ± 1 et $\pm i$ (lemme 6.22).

Le lemme 6.21 montre que $\mathbb{Z}[i]$ est un anneau euclidien parce que N est un stathme. L'anneau $\mathbb{Z}[i]$ étant euclidien, il est principal (proposition 1.254).

Pour la suite, nous allons d'abord montrer que $p \in \Sigma$ si et seulement si p n'est pas irréductible dans $\mathbb{Z}[i]$, puis nous allons voir quels sont les irréductibles de $\mathbb{Z}[i]$.

10. Définition 6.23.

Soit p, un nombre premier dans Σ. Si $p = a^2 + b^2$, alors nous avons $p = (a + ib)(a - bi)$, mais étant donné que p est premier, nous avons $a \neq 0$ et $b \neq 0$. Du coup p n'est pas inversible dans $\mathbb{Z}[i]$, mais il peut être écrit comme le produit de deux non inversibles. Le nombre p est donc non irréductible dans $\mathbb{Z}[i]$.

Dans l'autre sens, nous supposons que p est un nombre premier non irréductible dans $\mathbb{Z}[i]$. Nous avons alors $p = zz'$ avec ni z ni z' dans $\{\pm 1, \pm i\}$. En appliquant N nous avons

$$p^2 = N(p) = N(z)N(z'). \tag{6.42}$$

Comme p est premier par hypothèse, c'est uniquement possible avec $N(z) = N(z') = p$ (avoir $N(z) = 1$ est impossible parce que cela signifirait que z est inversible). Si $z = a + ib$, alors $p = N(z) = a^2 + b^2$, et donc $p \in \Sigma$.

Nous savons déjà que $\mathbb{Z}[i]$ est un anneau principal et n'est pas un corps ; la proposition 1.244 s'applique donc et p sera non irréductible si et seulement si l'idéal (p) est non premier. Le fait que (p) soit un idéal non premier implique que le quotient $\mathbb{Z}[i]/(p)$ est non intègre (c'est la définition d'un idéal premier). Nous cherchons donc les nombres premiers pour lesquels le quotient $\mathbb{Z}[i]/(p)$ n'est pas intègre.

Nous commençons par écrire le quotient $\mathbb{Z}[i]/(p)$ sous d'autres formes. D'abord en remarquant que si I et J sont deux idéaux, on a $(\mathbb{A}/I)/J \simeq (\mathbb{A}/J)/I$. Par conséquent, en tenant compte du fait que $\mathbb{Z}[i] = \mathbb{Z}[X]/(X^2 + 1)$, nous avons

$$\mathbb{Z}[i]/(p) = (\mathbb{Z}[X]/(p))/(X^2 + 1) = \mathbb{F}_p[X]/(X^2 + 1). \tag{6.43}$$

Nous avons donc équivalence des propositions suivantes :

$$p \in \Sigma \tag{6.44a}$$
$$\mathbb{F}_p[X]/(X^2 + 1) \text{ n'est pas intègre} \tag{6.44b}$$
$$X^2 + 1 \text{ n'est pas irréductible dans } \mathbb{F}_p \tag{6.44c}$$
$$X^2 + 1 \text{ admet une racine dans } \mathbb{F}_p \tag{6.44d}$$
$$-1 \in (\mathbb{F}_p^*)^2 \tag{6.44e}$$
$$\exists y \in \mathbb{F}_p^* \text{ tel que } y^2 = -1. \tag{6.44f}$$

Le point (6.44c) vient de la proposition 6.20. Maintenant nous utilisons le fait que p soit un premier impair (parce que le cas de $p = 2$ est déjà complètement traité), donc $(p-1)/2 \in \mathbb{N}$ et nous avons, pour le y de la dernière ligne,

$$(-1)^{(p-1)/2} = (y^2)^{(p-1)/2} = y^{p-1} = 1 \tag{6.45}$$

parce que dans $\mathbb{F}_p$ nous avons $y^{(p-1)} = 1$ par le petit théorème de Fermat (théorème 6.15). Ainsi p doit vérifier

$$1 = (-1)^{(p-1)/2}, \tag{6.46}$$

c'est-à-dire $\frac{p-1}{2} = 0 \mod 2$ ou encore $p = 1 \mod 4$. □

Remarque 6.26.
Il n'est pas dit que les nombres dans $[1]_4$ sont premiers ($9 = 8 + 1$ ne l'est pas par exemple). Le théorème signifie que (à part 2), si un nombre premier est dans $[1]_4$ alors il est somme de deux carrés, et inversement, si un nombre premier est somme de deux carrés, il est dans $[1]_4$.

Théorème 6.27 (Théorème des deux carrés[104])**.**
Soit $n \geqslant 2$ un nombre dont nous notons

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)} \tag{6.47}$$

où $\mathcal{P}$ est l'ensemble des nombres premiers. Alors $n \in \Sigma$ si et seulement si pour tout $p \in \mathcal{P} \cap [3]_4$, nous avons $v_p(n) \in [0]_2$ (c'est-à-dire $v_p(n)$ est pair).

Démonstration. (1) **Condition suffisante.**

Le produit (6.47) est évidemment un produit fini que nous pouvons alors regrouper en quatre parties : $\mathcal{P} \cap [0]_4$, $\mathcal{P} \cap [1]_4$, $\mathcal{P} \cap [2]_4$ et $\mathcal{P} \cap [3]_4$.

— Il n'y a pas de nombres premiers dans $[0]_4$.

— Les nombres premiers de $[1]_4$ sont dans Σ. Le produit d'éléments de Σ étant dans Σ, nous avons

$$\prod_{p\in\mathcal{P}\cap[1]_4} p^{v_p(n)} \in \Sigma. \tag{6.48}$$

— Le seul nombre premier dans $[2]_4$ est 2. C'est un élément de Σ.

— Le produit

$$\prod_{p\in\mathcal{P}\cap[3]_4} p^{v_p(n)} \tag{6.49}$$

est par hypothèse un produit de carrés ($v_p(n)$ est pair), et est donc un carré.

Au final le produit $\prod_{p\in\mathcal{P}} p^{v_p(n)}$ est un produit d'un carré par un élément de Σ, ce qui est encore un élément de Σ.

Pour cette partie, nous avons utilisé et réutilisé le lemme 6.24.

(2) **Condition nécessaire.**

Soit p, un nombre premier. Nous voulons montrer que

$$\{v_p(n) \text{ tel que } n \in \Sigma\} \subset [2]_2. \tag{6.50}$$

Pour montrer cela nous allons procéder par récurrence sur les ensembles

$$E_k = \{v_p(n) \text{ tel que } n \in \Sigma\} \cap \{0, \dots, k\}. \tag{6.51}$$

Il est évident que les éléments de E_0 sont pairs, puisqu'il n'y a que zéro, qui est pair.

Supposons que $E_k \subset [0]_2$, et montrons que $E_{k+1} \subset [0]_2$. Soit un élément de E_{k+1}, c'est-à-dire $v_p(n) \leqslant k+1$ avec $n = a^2 + b^2$. Si $v_p(n) = 0$ alors l'affaire est réglée ; sinon c'est que p divise n. Mais dans $\mathbb{Z}[i]$ nous avons

$$n = a^2 + b^2 = (a + bi)(a - bi) \tag{6.52}$$

Comme $\mathbb{Z}[i]$ est principal, le lemme de Gauss 3.74 nous dit que si p divise n, alors il doit diviser soit $a + bi$, soit $a - bi$ (et en fait, les deux). Nous avons alors $p \mid a$ et $p \mid b$ en même temps. Et donc

$$p^2 \mid a^2 + b^2 = n. \tag{6.53}$$

Posons $a = pa'$ et $b = pb'$ avec $a', b' \in \mathbb{N}$. Nous avons

$$\frac{n}{p^2} = \frac{p^2a'^2 + p^2b'^2}{p^2} = a'^2 + b'^2 \in \Sigma. \tag{6.54}$$

Mais par construction,

$$v_p\left(\frac{n}{p^2}\right) = v_p(n) - 2 < k. \tag{6.55}$$

Donc $v_p(\frac{n}{p^2})$ est pair et par conséquent, $v_p(n)$ doit également être pair.

□

6.2.2 Résultats chinois

Nous allons maintenant parler du système d'équations SUBEQooHVCTooGXvbAH

$$\begin{cases} x = a_1 \mod p & (6.56a)\\ x = a_2 \mod q & (6.56b) \end{cases}$$

avec a_1, a_2 donnés dans $\mathbb{Z}$ et p, q des entiers premiers entre eux. Le lemme chinois nous donne la liste des solutions ainsi qu'une manière de les construire. Le théorème chinois en sera une espèce de corolaire qui établira l'isomorphisme d'anneaux $\mathbb{Z}/pq\mathbb{Z} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$. Voir [160].

Notons que le système (6.56), en suivant les notations usuelles du Frido s'écrit plutôt

$$x \in [a_1]_p \cap [a_2]_q. \tag{6.57}$$

LemCtUeGA

Lemme 6.28 (Lemme chinois [161])**.**
Soient n_1, n_2 deux entiers premiers entre eux[11]*. Soient $a_1, a_2 \in \mathbb{Z}$.* ITEMooZVUSooGoplQW

(1) Il existe $u_1, u_2 \in \mathbb{Z}$ tels que

$$u_1 n_1 + u_2 n_2 = 1. \tag{6.58}$$ EQooIJXOooNEfCBo ITEMooFMFEooQHNLFc

(2) Nous posons

$$a = a_1 u_2 n_2 + a_2 n_1 u_1. \tag{6.59}$$

Nous avons

$$[a_1]_{n_1} \cap [a_2]_{n_2} = [a]_{n_1 n_2}. \tag{6.60}$$

Démonstration. En plusieurs points.

(1) **Pour (1)** Il suffit d'utiliser le théorème de Bézout 1.236

(2) **Pour (2) première inclusion** Soit $x \in [a]_{n_1 n_2}$, et vérifions que $x \in [a_1]_{n_1}$. Il existe $k \in \mathbb{Z}$ tel que

$$x = a_1 u_2 n_2 + a_2 n_1 u_1 + k n_1 n_2. \tag{6.61}$$

En remplaçant $u_2 n_2$ par $1 - u_1 n_1$, nous avons

$$x = a_1(1 - u_1 n_1) + a_2 n_1 u_1 + k n_1 n_2 \in [a_1]_{n_1} \tag{6.62}$$

parce qu'il n'y a que le tout premier terme qui ne contient pas de facteur n_1.
Le fait que $x \in [a_2]_{n_2}$ se vérifie de même[12].

(3) **Pour (2) seconde inclusion** Soit $x \in [a_1]_{n_1} \cap [a_2]_{n_2}$. Il existe $k \in \mathbb{Z}$ tel que $x = a_1 + k n_2$. Donc

$$\begin{aligned} x - a &= a_1 + k n_1 - (a_1 u_2 n_2 + a_2 n_1 u_1) && (6.63a)\\ &= a_1 - a_1 u_2 n_2 + k' n_1 && (6.63b)\\ &= a_1 \underbrace{(1 - u_2 n_2)}_{=u_1 n_1} + k' n_1 && (6.63c)\\ &= a_1 u_1 n_1 + k' n_1, && (6.63d) \end{aligned}$$

ce qui signifie que n_1 divise $x - a$..
De même nous prouvons que n_2 divise $x - a$. Il existe $k, l \in \mathbb{Z}$ tels que

$$x - a = k n_1 = l n_2. \tag{6.64}$$ EQooFMWBooTpqpoj

En particulier n_1 divise $l n_2$ alors que n_1 et n_2 sont premiers entre eux. Le lemme de Gauss 3.74 dit alors que n_1 divise l. Il existe donc $s \in \mathbb{Z}$ tel que $l = s n_1$. En remettant ce l dans (6.64), nous trouvons

$$x - a = l n_1 = s n_1 n_2, \tag{6.65}$$

11. Définition 1.259.
12. Je dis ça au hasard ; je n'ai pas vérifié. Faites-moi signe si c'est bon.

et donc $x-a$ est divisible en n_1n_2, c'est-à-dire

$$x\in[a]_{n_1n_2},\tag{6.66}$$

ce qu'il fallait prouver.

□

THOooQHYLooVMBAfe

Théorème 6.29 (Théorème chinois[162]).
Soient p,q deux naturels premiers entre eux. Si $p,q\geqslant 2$ alors l'application

$$\begin{aligned}\phi\colon\mathbb{Z}/pq\mathbb{Z}&\to\mathbb{Z}/p\mathbb{Z}\times\mathbb{Z}/q\mathbb{Z}\\ [x]_{pq}&\mapsto\big([x]_p,[x]_q\big)\end{aligned}\tag{6.67}$$

est un isomorphisme d'anneaux.

Démonstration. Nous devons prouver que l'application ϕ respecte la somme, le produit et qu'elle est bijective. En ce qui concerne la somme,

$$\begin{aligned}\phi([q]_{pq}+[y]_{pq})&=\phi\big((x+y)\mod pq\big)&(6.68a)\\ &=\big([x+y]_p,[x+y]_q\big)&(6.68b)\\ &=\big([x]_p+[y]_p,[x]_q+[y]_q\big)&(6.68c)\\ &=\big([x]_p,[x]_q\big)+\big([y]_p,[y]_q\big)&(6.68d)\\ &=\phi(x)+\phi(y).&(6.68e)\end{aligned}$$

En ce qui concerne le produit, c'est le même jeu : nous obtenons

$$\phi\big([xy]_{pq}\big)=\phi([x]_{pq}])\phi([y]_{pq})\tag{6.69}$$

en utilisant le fait que $[xy]_p=[x]_p[y]_p$.

Montrons maintenant que ϕ est surjective. Soient $y_1,y_2\in\mathbb{Z}$ et $x\in\mathbb{Z}$. Demander

$$\phi([x]_{pq})=\big([y_1]_p,[y_2]_q\big)\tag{6.70}$$

revient à demander de résoudre le système

$$\begin{cases}[x]_p=[y_1]_p&(6.71a)\\ [x]_q=[y_2]_q&(6.71b)\end{cases}$$

pour l'inconnue $x\in\mathbb{Z}$. Le lemme chinois 6.28 nous assure qu'une solution existe.

En ce qui concerne l'injectivité, nous supposons que x et y soient deux entiers tels que

$$\phi([x]_{pq})=\phi([y]_{pq}).\tag{6.72}$$

Nous en déduisons le système

$$\begin{cases}[x]_p=[y]_p&(6.73a)\\ [x]_q=[y]_q&(6.73b)\end{cases}$$

c'est-à-dire qu'il existe des entiers k et l tels que $x=y+kp$ et $x=y+lq$ ou encore tels que

$$kp+lq=0.\tag{6.74}$$

Étant donné que p et q sont premiers entre eux, le lemme de Gauss 3.12 implique que p divise l : il existe $l'\in\mathbb{Z}$ tel que $lp'=l$. En injectant cela dans $x=y+lq$, nous trouvons $x=y+l'pq$. Donc $[x]_{pq}=[y]_{pq}$, ce qu'il fallait. □

THOooVIGQooUhwBLS

Théorème 6.30 (Théorème chinois[1]).
Soit A un anneau commutatif. Soient $n\geqslant 2$, des éléments $x_1,\dots,x_n$ dans A et des idéaux $I_1,\dots,I_n$ tels que $I_i+I_j=A$ pour tout $i\neq j$.

Alors il existe un $x\in A$ tel que $x-x_i\in I_i$ pour tout $1\leqslant i\leqslant n$.

Démonstration. En plusieurs points.

(1) **Définition de J_k**
Pour $k \in \{1, \dots, n\}$ nous notons J_k le produit $J_k = \prod_{l \neq k} I_l$.

(2) **Si $k \neq i$ alors $J_k \subset I_i$** Un élément de J_k est de la forme

$$\prod_{l \neq k} a_j = a_i \prod_{\substack{l \neq k \\ l \neq i}} a_l \tag{6.75}$$

EQooLVTXooIrdWbH

avec $a_l \in I_l$. Vu que a_i est dans I_i qui est un idéal, tout le produit (6.75) est dans I_i.

(3) $I_i + J_i = A$ En effet, soit $k \neq i$. Nous avons $J_i \subset I_k$ et donc

$$I_i + J_i \subset I_i + I_k = A. \tag{6.76}$$

(4) **Des α et des β**
Soit $i \in \{1, \dots, n\}$. Puisque $I_i + J_i = A$, nous pouvons donc prendre $\alpha_i \in I_i$ et $\beta_i \in J_i$ tels que

$$\alpha_i + \beta_i = 1. \tag{6.77}$$

(5) **Et enfin** Nous posons $x = \sum_{i=1}^{n} \beta_i x_i$. Pour chaque k nous avons

$$\begin{aligned} x - x_k &= \sum_{i=1}^{n} \beta_i x_i - x_k && \text{(6.78a)} \\ &= \sum_{i \neq k} \beta_i x_i + (\beta_k - 1) x_k && \text{(6.78b)} \\ &= \sum_{i \neq k} \beta_i x_i - \alpha_k x_k. && \text{(6.78c)} \end{aligned}$$

Les deux termes sont dans I_k. En effet, d'une part $\beta_i \in J_i \subset I_k$ et I_k est un idéal, donc $\beta_i x_i \in I_k$, et d'autre part $\alpha_k \in I_k$, donc $\alpha_k x_k \in I_k$.
Un idéal étant stable par somme, la somme du tout est dans I_k.

□

Remarque 6.31.
Ce théorème chinois est bien une généralisation du lemme chinois 6.28. En effet, l'élément x dont il est question est solution du problème $x = x_i \mod I_i$. L'hypothèse $I_i + I_j = A$ n'est pas nouvelle non plus étant donné que si p et q sont des entiers premiers entre eux nous avons $p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}$ par le corolaire 1.238.

6.3 Polynômes à coefficients dans un corps

SECooFYOGooQHitgE

Nous supposons que $\mathbb{K}$ est un corps commutatif, et nous étudions l'anneau $\mathbb{K}[X]$, défini en 1.359.

LEMooKSTIooSTxAll

Lemme 6.32 ([118]).
Soit un anneau factoriel A, ainsi que son corps des fraction[13] *$\mathbb{K}$. Si $P \in \mathbb{K}[X]$ est de degré $\geqslant 1$, nous pouvons écrire $P = qP_1$ où*

(1) $q \in \mathbb{K}^\times$

(2) $P_1 \in A[X]$ *est primitif*[14].

13. Définition 1.369.
14. Définition 3.115(1).

Démonstration. Nous considérons le polynôme $P = \sum_{j=0}^{n} \frac{a_j}{s_j} X^j$ avec $n \geqslant 1$, $a_j \in A$, $s_j \in A^*$ et $a_n \neq 0$. Si $s \in \text{ppcm}(\{s_j\}_{j=1,\ldots,n})$, nous avons

$$P = \frac{1}{s} \sum_{j=1}^{n} a'_j X^j \tag{6.79}$$

avec $a'_j \in A$.

Si $d \in \text{pgcd}(\{a'_j\})$, nous pouvons écrire $a'_j = db_j$ où les $b_j \in A$ sont premiers entre eux. Avec tout ça,

$$P = \frac{d}{s} \sum_{j} b_j X^j. \tag{6.80}$$

Maintenant en posant $q = d/s$ et $P_1 = \sum_j b_j X^j$ nous avons ce qu'il faut parce que $d/s \in \mathbb{K}^\times$ et P_1 est primitif. □

LEMooGTPJooPWkffg

Lemme 6.33 ([41])**.**
Soit un polynôme primitif $P \in A[X]$ sur l'anneau factoriel A. Nous notons $\mathbb{K}$ le corps des fractions de A Soient $Q, R \in \mathbb{K}[X]$ tels que $P = QR$. Il existe $q, r \in \mathbb{K}$ tels que

$$P = (qQ)(rR) \tag{6.81}$$

avec qQ, rR dans $A[X]$.

Démonstration. En utilisant le lemme 6.32, nous considérons $qQ = Q_1$ et $rR = R_1$ avec Q_1 et R_1 primitifs dans $A[X]$. Avec ces éléments nous avons

$$R_1 Q_1 = rqPQ = qrP. \tag{6.82}$$

EQooMGXYooIAFBgv

À priori q et p sont dans $\mathbb{K}$. Nous allons prouver que $qr = 1$ …pas tout à fait. Nous allons prouver que qr est inversible. Ensuite il faudra choisir les q et r un peu autrement.

(1) $\underline{qr \in A}$ Posons $P = \sum_i a_i X^i$ ainsi que $qr = s/t$ avec $s, t \in A$. Nous supposons que s/t est une fraction irréductible. L'équation (6.82) nous dit que le polynôme qrP est $R_1 Q_1$ et donc à coefficients dans A. Pour chaque i, il existe donc $b_i \in A$ tel que

$$\frac{sa_i}{t} = b_i, \tag{6.83}$$

ou encore $sa_i = tb_i$. Donc nous avons $t \mid sa_i$ pour tout i. Vu que la fraction $\text{pgcd}(t, s) = 1$, le lemme de Gauss 3.74 nous indique que t divise a_i. Donc si $\delta \in \text{pgcd}(\{a_i\}_{i=1,\ldots,n})$, alors $t \mid \delta$. Vu que P est primitif, δ est inversible, et donc t l'est aussi. Nous avons alors le calcul

$$\frac{s}{t} = \frac{st^{-1}}{tt^{-1}} = st^{-1} \in A. \tag{6.84}$$

Nous avons prouvé que $qr \in A$.

(2) $\underline{qr \textbf{ est inversible}}$
Comme qr est dans A nous pouvons l'utiliser dans le lemme 1.201. Nous savons que $c(P)$ est inversible, disons $c(P) = u$. Nous avons donc, en utilisant le lemme 3.119,

$$qru = qr\, c(P) = u\, c(qrP) = u\, c(Q_1 R_1) = uv\, c(Q_1) c(R_1) \in A^\times \tag{6.85}$$

pour un certain inversible $v \in A^\times$. Donc qru est inversible et comme u est inversible, qr est inversible.

(3) **Conclusion**
Nous avions $P = QR$, et nous avions posé $qQ = Q_1$, $rR = R_1$. Maintenant, en posant $u = (qr)^{-1}$ nous avons

$$(uqQ)(rR) = uqrQR = P. \tag{6.86}$$

En même temps, $uqQ = uQ_1 \in A[X]$ et $rR = R_1 \in A[X]$. Donc les polynômes qui satisfont la demande sont uqQ et rR.

□

THOooJESZooNqScLF

Théorème 6.34 ([41]).
Soient un anneau factoriel A et son corps des fractions $\mathbb{K}$. Les irréductibles[15] *dans $A[X]$ sont :*

(1) les polynômes de degré zéro dont l'unique coefficient est un élément irréductible de A (polynômes constants),

ITEMooVKOOooDeGYYz

(2) les polynômes de $A[X]$ qui sont de degré $\geqslant 1$, primitifs dans $A[X]$ et irréductibles dans $\mathbb{K}[X]$.

Démonstration. Un polynôme est soit de degré zéro soit de degré $\geqslant 1$.

(1) **Degré zéro, $\Rightarrow$**
Soit un polynôme irréductible[16] $P \in \mathcal{P}_0(A)$.
Si nous notons $a \in A$ l'unique coefficient de P, nous devons prouver que a est irréductible dans A.

(1a) **a n'est pas inversible** Si a était inversible dans A, alors nous pourrions considérer le polynôme $Q \in \mathcal{P}_0(A)$ dont le coefficient est a^{-1}. Nous aurions alors $PQ = U$ où $U \in \mathcal{P}(A)$ est le polynôme dont le coefficient est $1 \in A$.
Ce U étant le neutre pour la multiplication dans $\mathcal{P}(A)$, le polynôme P est inversible, ce qui est contraire à l'hypothèse.

(1b) **a n'est pas produit de deux non inversibles** Même jeu : si $a = xy$ avec x et y non inversibles, alors en prenant les polynômes correspondants, P est produit de deux non inversibles, le polynôme P serait produit de deux non inversibles, et donc pas irréductible dans $\mathcal{P}(A)$.

(2) **Degré zéro, $\Leftarrow$**
Si a est irréductible dans A, le polynôme P de coefficient a est irréductible dans $\mathcal{P}(A)$. Même principe de preuve que le point précédent.

(3) **Degré $n \geqslant 1$, $\Rightarrow$**
Soit $P \in \mathcal{P}_n(A)$ irréductible dans $\mathcal{P}(A)$. Nous devons démontrer que P est dans le cas (2). Nous écrivons $P = c(P)P_1$ avec P_1 primitif. Vu que P n'est pas produit de deux non inversibles et que P_1 n'est pas inversible, le polynôme $c(P) \in \mathcal{P}_0(A)$ est inversible[17], et donc $c(P)$ est inversible dans A. Cela démontre déjà que P est primitif.
Pour voir que P est irréductible dans $\mathcal{P}(\mathbb{K})$, nous supposons avoir une égalité $P = QR$ avec $Q, R \in \mathcal{P}(\mathbb{K})$. Par le lemme 6.33, il existe $q, r \in \mathbb{K}$ tels que

$$P = (qQ)(rR) \tag{6.87}$$

avec $qQ, rR \in \mathcal{P}(A)$. Vu que P est irréductible dans $\mathcal{P}(A)$, soit qQ soit rR doit être inversible dans $\mathcal{P}(A)$. Supposons que ce soit qQ. Nous avons donc $qQ \in A^\times$ (qui serait mieux écrit $qQ \in \mathcal{P}_0(A^\times)$), et donc $Q \in \mathcal{P}(\mathbb{K})^\times$.
En résumé nous avons prouvé que si $P = QR$ (dans $\mathcal{P}(\mathbb{K})$), alors soit Q soit R était dans $\mathcal{P}(\mathbb{K})^\times$.

(4) **Degré $n \geqslant 1$, $\Leftarrow$**
Nous supposons que $P \in \mathcal{P}_n(A)$ vérifie la condition (2), et nous montrons qu'il est irréductible dans $\mathcal{P}(A)$. Les polynômes non constants ne sont jamais inversibles, donc pour cette partie-là de la définition d'irréductibilité on est bon.

15. Définition 1.189.
16. Cette notation $\mathcal{P}$ de (1.492) devrait être utilisée plus souvent ; dans le meilleur des mondes, on n'utiliserait jamais $A[X]$.
17. On note de la même manière le polynôme constant $c(P)$ et l'élément $c(P)$ dans A, parce qu'on a beau détester les abus de notations, il y a des limites.

Supposons que $P = QR$ avec $Q, R \in \mathcal{P}(A) \subset \mathcal{P}(\mathbb{K})$. Comme P est irréductible dans $\mathcal{P}(\mathbb{K})$, un des deux Q ou R est inversible dans $\mathcal{P}(\mathbb{K})$. Disons que Q le soit. Donc $Q \in \mathcal{P}(\mathbb{K})^\times$, et nous notons $q \in \mathbb{K}^\times$ l'unique coefficient de Q. Nous avons donc

EQooQUMSooFGIGxE
$$P = qR. \tag{6.88}$$

Comme, dès le début, on avait dit $Q \in \mathcal{P}(A)$, nous avons en réalité $q \in A$. Nous pouvons donc prendre le contenu de (6.88) en appliquant le lemme 1.201 : $c(P) = q\,c(R)$. Nous savons que $c(P)$ est inversible dans A ; nous multiplions par $c(P)^{-1}$:

$$1 = qc(R)c(P)^{-1}. \tag{6.89}$$

Nous voyons que $c(R)c(P)^{-1}$ est un inverse de q dans A. Donc P est irréductible dans $\mathcal{P}(A)$.

□

THOooVZUBooNSvcBr

Théorème 6.35 ([41]).
Si A est un anneau factoriel, l'anneau des polynômes $\mathcal{P}(A)$ est factoriel.

Démonstration. Si $P \in \mathcal{P}(A)$, nous devons prouver l'existence et l'unicité d'un décomposition de P en irréductibles.

Existence

Nous y allons par récurrence sur le degré de P.

(1) **Si** $\deg(P) = 0$ L'unique coefficient de P est un élément de A, et se décompose en irréductibles parce que A est factoriel.

(2) **Si** $\deg(P) = n + 1$ Nous supposons que l'existence est prouvée pour tout $k \leqslant n$. Nous écrivons $P = c(P)P_1^{(0)}$ avec P primitif dans $\mathcal{P}(A)$. Il y a deux possibilités : soit $c(P)$ est inversible, soit non. Si $c(P)$ est inversible, alors P est primitif et nous posons $P_1 = P$ et $c_1 = 1$. Si $c(P)$ n'est pas inversible, alors nous posons $c_1 = c(P)$ et $P_1 = P_1^{(0)}$. Dans les deux cas nous avons

$$P = c_1P_1 \tag{6.90}$$

où P_1 est primitif et c_1 est soit 1 soit non inversible. Si $c_1 = 1$, il n'y a rien à faire avec lui. Sinon, il est non inversible et non nul dans A qui est factoriel. Nous avons donc une décomposition en irréductibles $c_1 = p_1 \dots p_k$. Dans le cas $c_1 = 1$, nous avons $p_1 = 1$ et c'est tout.

En ce qui concerne P_1, il y a deux possibilités : soit il est irréductible dans $\mathcal{P}(\mathbb{K})$ soit pas. Si il est irréductible dans $\mathcal{P}(\mathbb{K})$ alors il est à fortiori irréductible dans $\mathcal{P}(A)$ et

$$P = p_1 \dots p_kP_1 \tag{6.91}$$

est une décomposition de P en irréductibles dans $\mathcal{P}(A)$.

Et si P_1 n'est pas irréductible ? C'est qu'il existe deux non inversibles $Q, R \in \mathcal{P}(\mathbb{K})$ tels que $P = QR$. Le lemme 6.33 permet d'écrire alors

$$P_1 = QR = Q_1R_1 \tag{6.92}$$

avec $Q_1, R_1 \in \mathcal{P}(A)$. Notons que ni P_1 ni R_1 ne sont des constantes : les constantes sont inversibles dans $\mathcal{P}(\mathbb{K})$. Et comme la construction du lemme montre que $\deg(Q_1) = \deg(Q)$ et $\deg(R_1) = \deg(R)$, les polynômes Q_1 et R_1 sont non constants dans $\mathcal{P}(A)$. Nous avons donc $\deg(Q_1) < n + 1$ et $\deg(R_1) < n + 1$.

L'hypothèse de récurrence fonctionne pour les polynômes Q_1 et R_1 dans $\mathcal{P}(A)$. Nous avons donc des polynômes irréductibles $S_1, \dots S_l$ tels que

$$P_1 = Q_1R_1 = S_1 \dots S_l. \tag{6.93}$$

Au final, nous avons la décomposition [18]

$$P = p_1 \dots p_k S_1 \dots S_l. \tag{6.94}$$

Unicité

Soient $P = p_1 \dots p_k S_1 \dots S_m = q_1 \dots q_l T_1 \dots T_n$ deux décompositions de P en irréductibles de A. Les p_i et les q_i sont les polynômes de degré zéro (donc dans A) et les S_i, T_j sont ceux de degré strictement positifs.

(1) **Pour les p_i et q_j**
Comme les polynômes S_i et T_j sont irréductibles leurs contenus sont inversibles, et en utilisant les lemmes 3.119 et 1.201, l'égalité

$$c(p_1 \dots p_k S_1 \dots S_m) = c(q_1 \dots q_l T_1 \dots T_n) \tag{6.95}$$

devient

$$p_1 \dots p_k c(S_1 \dots S_m) = q_1 \dots q_l c(T_1 \dots T_n). \tag{6.96}$$

Il existe donc un inversible $u \in A$ tel que $p_1 \dots p_k = u q_1 \dots q_l$. L'élément (uq_1) est irréductible (lemme 1.192) et donc

$$p_1 \dots p_k = (uq_1) q_2 \dots q_l \tag{6.97}$$

sont deux décompositions en irréductibles du même élément dans A. Donc, parce que A est factoriel, nous avons $k = l$ et il existe une permutation σ qui fait que p_i est associé à $q_{\sigma(i)}$.

(2) **Pour les polynômes**
Nous avons $S_1 \dots S_m = T_1 \dots T_m$. Le polynôme S_1 divise le produit $T_1 \dots T_m$. Comme S_1 est irréductible dans $\mathcal{P}(A)$, il est irréductible dans $\mathcal{P}(\mathbb{K})$ par le théorème 6.34.
L'anneau $\mathcal{P}(\mathbb{K})$ est euclidien par 3.110 et donc principal par 1.254. Or dans un anneau principal, les irréductibles sont premiers par 1.234. Tout ça pour dire que S_1 est premier dans $\mathcal{P}(\mathbb{K})$. Donc, comme il divise un produit, il doit diviser un des facteurs. Disons que $S_1 \mid T_i$ dans $\mathcal{P}(\mathbb{K})$.
Il existe $R \in \mathcal{P}(\mathbb{K})$ tel que $S_1 R = T_i$. Comme S_1 est non inversible (irréductible), et T_i est irréductible, R doit être inversible dans $\mathcal{P}(\mathbb{K})$. Il existe $q \in \mathbb{K}^\times$ tel que $T_i = qS_1$.
Comme T_i et S_1 sont irréductibles, ils sont primitifs dans $\mathcal{P}(A)$ (théorème 6.34). Supposons que q s'écrive sous forme irréductible $q = s/t$ avec $s, t \in A$. On note $T_i = \sum_j a_j X^j$ et $S_1 = \sum_j b_j X^j$. Pour tout i nous avons

$$\frac{sb_j}{t} = a_j, \tag{6.98}$$

ou encore $sb_j = ta_j$. Donc t divise sb_j pour tout j. Mais comme t est premier avec s, l'élément t divise b_j pour tout j. Il divise dont le pgcd des a_j qui est inversible parce que T_i est primitif. Bref, t est inversible dans A. Nous avons donc

$$q = \frac{s}{t} = \frac{st^{-1}}{tt^{-1}} = st^{-1} \in A. \tag{6.99}$$

Nous avons donc $T_i = qS_1$ avec $q \in A$. Comme T_i et S_1 sont irréductibles dans $\mathcal{P}(A)$, l'élément q est inversible dans A.
Nous avons fini de montrer que S_1 était associé à un des T_i. L'égalité

$$S_1 \dots S_m = T_1 \dots T_n \tag{6.100}$$

peut être alors récrite

$$S_1 \dots S_m = qT_1 \dots T_{i-1} S_1 T_{i+1} \dots T_n \tag{6.101}$$

18. Ici [118] utilise le théorème 6.34. Certes, il présente sa récurrence de façon un peu différente, mais j'espère n'avoir pas loupé une étape. Soyez prudente et écrivez-moi si vous voyez une erreur.

Comme A est factoriel, il est intègre, et donc $\mathcal{P}(A)$ est intègre (théorème 3.104). Nous pouvons donc simplifier par S_1 et il reste

$$S_2 \dots S_m = uT_1 \dots T_{i-1}T_{i+1} \dots T_n. \tag{6.102}$$

Par récurrence, nous avons que chacun des S_i est associé à un des T_j.

□

PropqGZXvr

Proposition 6.36 ([118, 163])**.**
Si $\mathbb{K}$ est un corps commutatif, alors l'anneau des polynômes $\mathcal{P}(\mathbb{K})$ est est factoriel[19]*.*

Démonstration. Un corps est un anneau factoriel (proposition 3.81), donc le théorème 6.35 conclut que l'anneau des polynômes sur un corps est factoriel. □

6.3.1 Irréductibilité

Pour rappel, un élément irréductible dans un anneau est la définition 1.189 : non inversible et pas les produit de deux non inversibles.

LEMooTMSUooCWoVHG

Lemme 6.37.
Soit un corps $\mathbb{K}$. Les éléments inversibles dans $\mathbb{K}[X]$ sont les polynômes constants non nuls.

Exemple 6.38.
Si un polynôme $P \in \mathbb{Z}[X]$ n'a que des racines complexes, ça ne l'empêche pas d'être réductible sur $\mathbb{Z}$. La réductibilité ne signifie pas qu'on peut mettre des racines en évidence. Par exemple le polynôme $P = X^4 + 3X^2 + 2$ est réductible sur $\mathbb{Z}$ en

$$P = (X^2 + 1)(X^2 + 2), \tag{6.103}$$

mais n'a pas de racines dans $\mathbb{Z}$. Si on veut réduire plus, il faut sortir de $\mathbb{Z}$.

△

DefCPLSooQaHJKQ

Définition 6.39 (Polynôme scindé)**.**
*Nous disons que $P \in \mathbb{K}[X]$ est **scindé** sur $\mathbb{K}$ si il est produit dans $\mathbb{K}[X]$ de polynômes de degré 1.*

Note : les constantes ne sont donc pas des polynômes scindés.

Proposition 6.40 (Critère d'Eisenstein)**.**
Soit le polynôme $P = \sum_{k=0}^{n} a_n X^n$ dans $\mathbb{Z}[X]$. Nous supposons avoir un nombre premier p tel que

(1) p divise tous les $a_0, \dots, a_{n-1}$,

(2) p ne divise pas a_n,

(3) p^2 ne divise pas a_0.

Alors P est irréductible dans $\mathbb{Q}[X]$.

Si de plus P est primitif au sens du pgcd *(définition 3.115) alors P est irréductible dans $\mathbb{Z}[X]$.*

Démonstration. Nous considérons $\bar{P}$ le polynôme réduit modulo p, c'est-à-dire $\bar{P} \in \mathbb{F}_p[X]$. Étant donné que par hypothèse tous les coefficients sont multiples de p sauf a_n, nous avons $\bar{P} = cX^n$. Supposons par l'absurde que $P = QR$ avec $Q, R \in \mathbb{Q}[X]$. Alors le lemme de Gauss (3.74) impose $P, Q \in \mathbb{Z}[X]$.

Nous avons aussi, au niveau des réductions modulo p que $\bar{Q}\bar{R} = \bar{P}$. Or $\bar{P}$ est un monôme, donc $\bar{Q}$ et $\bar{R}$ doivent également l'être. Donc $\bar{Q} = dX^k$ et $\bar{R} = eX^{n-k}$ et en particulier $\bar{Q}(0) = \bar{R}(0) = 0$, c'est-à-dire que $Q(0)$ et $R(0)$ sont divisibles par p. Cela impliquerait que $a_0 = Q(0)R(0)$ soit divisible par p^2, ce qui est exclu par les hypothèses. Donc P est irréductible.

Supposons de surcroît que P soit primitif au sens du pgcd. Il est donc irréductible et primitif sur $\mathbb{Q}[X]$ et le corolaire 3.134 nous dit alors que P est irréductible sur $\mathbb{Z}[X]$. □

19. Définition 3.64.

Exemple 6.41.
Soit le polynôme $P = 3X^4 + 15X^2 + 10$. Pour appliquer le critère d'Eisenstein il nous faut un nombre premier p divisant 15 et 10, mais pas 3, et dont le carré ne divise pas 10. On aura vite vu que $p = 5$ fait l'affaire. Le polynôme P est donc irréductible sur $\mathbb{Q}[X]$. △

6.3.2 Idéaux

Soit $P \in \mathbb{K}[X]$ un polynôme. Nous notons (P) l'idéal engendré par P :

$$(P) = \{PR \text{ tel que } R \in \mathbb{K}[X]\}. \tag{6.104}$$ EqDefxMkDtW

Lemme 6.42.
Nous avons

(1) $(P) \subset (Q)$ *si et seulement si* Q *divise* P,

(2) $(P) = (Q)$ *si et seulement si* P *et* Q *sont multiples (non nuls) l'un de l'autre.*

Démonstration. Si $(P) \subset (Q)$, en particulier $P \in (Q)$ et il existe $R \in \mathbb{K}[X]$ tel que $P = QR$, ce qui signifie que Q divise P.

Si les idéaux de P et de Q sont identiques, l'un divise l'autre et l'autre divise l'un. Ils sont donc multiples l'un de l'autre. □

ThoCCHkoU

Théorème 6.43.
Soit $\mathbb{K}$ *un corps commutatif.*

ITEMooLZWMooDRsRwW

(1) *L'anneau* $\mathbb{K}[X]$ *est euclidien et principal.*

(2) *Si* I *est un idéal dans* $\mathbb{K}[X]$ *et si* $P \in I$ *est de degré minimal, alors* $(P) = I$.

ITEMooASHKooZqkiCH

(3) *De plus si* $I \neq \{0\}$, *il existe un unique polynôme unitaire* μ *tel que* $I = (\mu)$.

Démonstration. Le point (1) a déjà été démontré dans le lemme 3.110 via le fait que $\mathbb{K}[X]$ est euclidien. Nous allons cependant donner ici une preuve directe que tous les idéaux de $\mathbb{K}[X]$ sont principaux. Si $I = \{0\}$, le résultat est évident. Nous supposons donc I non nul. Soit P de degré minimum parmi les éléments de I. Évidemment $(P) \subset I$. Nous allons démontrer qu'en réalité $(P) = I$.

Soit $P' \in I$. Par le théorème 3.107 de la division euclidienne [20], il existe Q et R dans $\mathbb{K}[X]$ tels que $P' = PQ + R$ avec $\deg(R) < \deg(P)$. Étant donné que $R = P' - PQ$ nous avons $R \in I$ et par conséquent $R = 0$ parce que P a été choisi de degré minimum dans I. Nous avons donc $P' = PQ$ et $I \subset (P)$.

L'existence d'un polynôme unitaire qui génère I est obtenu en choisissant $\mu = P/a_n$ où a_n est le coefficient du terme de plus haut degré. L'unicité d'un tel polynôme est obtenu par le fait que si μ et μ' génèrent le même idéal, alors ils sont multiples l'un de l'autre, or puisqu'ils sont unitaires, ils sont égaux. □

Nous voyons que n'importe quel polynôme de degré minimum dans un idéal génère l'idéal. Une importante conséquence du théorème 6.43 que nous verrons plus bas est que tout polynôme annulateur d'un endomorphisme est divisé par le polynôme minimal (proposition 8.102).

CorsLGiEN

Corolaire 6.44.
Si $\mathbb{K}$ *est un corps et si* P *est un polynôme irréductible, alors l'ensemble* $\mathbb{L} = \mathbb{K}[X]/(P)$ *est un corps. De plus* $\mathbb{L}$ *est un espace vectoriel de dimension* $\deg(P)$.

Démonstration. En effet $\mathbb{K}[X]$ est un anneau principal par le théorème 6.43, par conséquent la proposition 1.246(2) déduit que $\mathbb{K}[X]/(P)$ est un corps.

20. Ici $\mathbb{K}$ est un corps et donc l'hypothèse d'inversibilité est automatiquement vérifiée.

Une base de $\mathbb{L}$ est donnée par les projections de $1, X, X^2, \dots, X^{n-1}$. En effet ces éléments forment une famille libre parce que si $\sum_{k=0}^{n-1} a_k \bar{X}^n = 0$ alors un représentant de cette classe doit être de la forme SP dans $\mathbb{K}[X]$, c'est-à-dire

$$\sum_{k=0}^{n-1} a_k X^k = SP, \tag{6.105}$$

ce qui n'est possible que si $S = 0$ et $a_k = 0$. D'autre part c'est un système générateur. En effet si $P = X^n + Q$ avec $\deg(Q) = n - 1$ alors

$$\bar{X}^{n+l} = \bar{X}^n \bar{X}^l = (\bar{P} - \bar{Q})\bar{X}^l = \bar{Q}\bar{X}^l. \tag{6.106}$$

Nous avons donc exprimé $\bar{X}^{n+l}$ comme une somme de termes de degré $n + l - 1$. Par récurrence nous pouvons exprimer tout $\bar{X}^{n+l}$ comme combinaison d'éléments de degré plus petit que n. □

6.45.
Ce corolaire prendra une nouvelle jeunesse lorsque nous parlerons de polynômes d'endomorphismes, en particulier la proposition 8.112 va donner des précisions.

LEMooGRIMooPxCXAZ

Lemme 6.46 ([164]).
Soit un isomorphisme de corps $\tau\colon \mathbb{K} \to \mathbb{K}'$. Alors

(1) L'application étendue

$$\begin{aligned} \tau\colon \mathbb{K}[X] &\to \mathbb{K}'[X] \\ \sum_i a_i X^i &\mapsto \sum_i \tau(a_i) X^i \end{aligned} \tag{6.107}$$

est un isomorphisme d'anneaux ;

(2) pour tout $P \in \mathbb{K}[X]$, le passage au quotient

$$\begin{aligned} \phi_\tau\colon \mathbb{K}[X]/(P) &\to \mathbb{K}'[X]/\big(\tau(P)\big) \\ \bar{Q} &\mapsto \overline{\tau(Q)} \end{aligned} \tag{6.108}$$

est un isomorphisme d'anneaux (et d'abord est bien définie).

Démonstration. Nous n'allons pas faire explicitement toutes les vérifications, mais tout de même les principales. Montrons que τ respecte le produit entre $\mathbb{K}[X]$ et $\mathbb{K}'[X]$. Nous rappelons que ce produit est défini par la formule (1.495). En notant P_i les coefficients de P et Q_i ceux de Q et en remarquant que la définition de τ est essentiellement que $\tau(P)_i = \tau(P_i)$, nous avons :

$$\tau(PQ) = \tau\Big(\sum_k \big(\sum_{l=0}^{k} P_l Q_{k-l}\big) X^k\Big) \tag{6.109a}$$

$$= \sum_k X^k \sum_{l=0}^{k} \tau(P_l Q_{k-l}) \tag{6.109b}$$

$$= \sum_k X^k \sum_{l=0}^{k} \tau(P_l)\tau(Q_{k-l}) \tag{6.109c}$$

$$= \sum_k X^k \sum_{l=0}^{k} \tau(P)_l \tau(Q)_{k-l} \tag{6.109d}$$

$$= \sum_i \big(\tau(P)_i X^i\big) \sum_j \big(\tau(Q)_j X^j\big) \tag{6.109e}$$

$$= \tau(P)\tau(Q). \tag{6.109f}$$

Passons à l'isomorphisme d'anneaux donné par ϕ_τ.

(1) **Bien définie**
Si $\bar{Q}_1 = \bar{Q}_2$ alors $Q_2 = Q_1 + RP$ pour un certain $R \in \mathbb{K}[X]$. Dans ce cas,

$$\phi_\tau(Q_2) = \overline{\tau(Q_2)} = \overline{\tau(Q_1) + \tau(RP)} \tag{6.110a}$$
$$= \overline{\tau(Q_1) + \tau(R)\tau(P)} \tag{6.110b}$$
$$= \overline{\tau(Q_1)}. \tag{6.110c}$$

L'application ϕ_τ est donc bien définie.

(2) **Injection**
Si $\phi_\tau(\bar{Q}_1) = \phi_\tau(\bar{Q}_2)$ alors $\overline{\tau(Q_1)} = \overline{\tau(Q_2)}$, ce qui signifie que

$$\tau(Q_1) = \tau(Q_2) + R\ \tau(P) \tag{6.111}$$

pour un certain $R \in \mathbb{K}'[X]$. Puisque $\tau\colon \mathbb{K}[X] \to \mathbb{K}'[X]$ est un isomorphisme, nous pouvons y appliquer τ^{-1} pour trouver :

$$Q_1 = Q_2 + \tau^{-1}(R)P, \tag{6.112}$$

ce qui signifie que $\bar{Q}_1 = \bar{Q}_2$.

(3) **Surjection**
Un élément de $\mathbb{K}'[X]/\big(\tau(P)\big)$ est de la forme $\bar{Q}$ avec $Q \in \mathbb{K}'[X]$. C'est l'image par ϕ_τ de l'élément $\overline{\tau^{-1}(Q)} \in \mathbb{K}[X]/(P)$.

(4) **Morphisme**
Nous vous laissons vérifier que l'application ϕ_τ est un morphisme d'anneaux.

□

6.3.3 Identité de Bézout

ThoBezoutOuGmLB

Théorème 6.47 (Bézout)**.**
Les polynômes $P_1, \dots, P_n$ dans $\mathbb{K}[X]$ sont étrangers entre eux si et seulement si il existe des polynômes $Q_1, \dots, Q_n \in \mathbb{K}[X]$ tels que

$$P_1Q_1 + \cdots + P_nQ_n = 1. \tag{6.113}$$

Deux polynômes P et Q ne sont donc pas premiers entre eux si il existe des polynômes x et y tels que l'identité de Bézout soit vérifiée :

$$xP + yQ = 0; \tag{6.114}$$

EqkbbzAi

cette dernière pourra être écrite en termes de la matrice de Sylvester, voir sous-section 8.2.8.

LemuALZHn

Lemme 6.48.
Soient $(P_i)_{i=1,\dots,n} \in \mathbb{K}[X]$ des polynômes étrangers deux à deux. Alors les polynômes

$$Q_i = \prod_{j \neq i} P_j \tag{6.115}$$

sont étrangers entre eux[21].

LemzwkYdn

Lemme 6.49 ([165])**.**
Soit $\mathbb{K}$ un corps commutatif et $\mathbb{A} \subset \mathbb{K}$ un sous anneau de $\mathbb{K}$. Alors $\mathbb{A}[X]$, vu comme idéal de $\mathbb{K}[X]$, est un idéal premier.

En d'autres termes, si $\phi \in \mathbb{K}[X]$, et si il existe $Q \in \mathbb{K}[X]$ unitaire tel que $\phi Q \in \mathbb{A}[X]$, alors $\phi \in \mathbb{A}[X]$.

21. Et non seulement deux à deux.

6.3.4 Lemme et théorème de Gauss

ThoLLgIsig

Théorème 6.50 (Théorème de Gauss)**.**
Soient $P, Q, R \in \mathbb{K}[X]$ tels que P soit premier avec Q et divise QR. Alors P divise R.

Démonstration. Étant donné que P est premier avec Q, le théorème de Bézout[22] nous donne $U, V \in \mathbb{K}[X]$ tels que $PU + QV = 1$. De plus il existe un polynôme S tel que $PS = QR$. En multipliant l'identité de Bézout par R, nous obtenons

$$R = PUR + QVR = PUR + VPS = P(UR + VS), \tag{6.116}$$

ce qui signifie que P divise R. □

Le lemme suivant est une généralisation du lemme de Gauss dans $\mathbb{Z}$ (lemme 3.74).

LemEfdkZw

Lemme 6.51 (Lemme de Gauss[114])**.**
Soient les polynômes unitaires $P, Q \in \mathbb{Q}[X]$. Si $PQ \in \mathbb{Z}[X]$, alors P et Q sont tous deux dans $\mathbb{Z}[X]$.

Démonstration. Soit $a > 0$ le plus petit entier tel que $aP \in \mathbb{Z}[X]$ (c'est le PPCM des dénominateurs) et de la même façon $b > 0$ le plus petit entier tel que $bQ \in \mathbb{Z}[X]$. On pose $P_1 = aP$ et $Q_1 = bQ$.

Si $ab = 1$, alors $a = b = 1$ et nous avons tout de suite $P, Q \in \mathbb{Z}[X]$. Nous supposons donc $ab > 1$ et nous considérons p, un diviseur premier de ab. Ensuite nous considérons la projection

$$\pi_p \colon \mathbb{Z}[X] \to (\mathbb{Z}/p\mathbb{Z})[X]. \tag{6.117}$$

Par définition $abPQ = P_1Q_1 \in \mathbb{Z}[X]$; en prenant la projection,

$$\pi_p(P_1)\pi_p(Q_1) = \pi_p(P_1Q_1) = \pi_P(ab)\pi_p(PQ) = 0 \tag{6.118}$$

parce que $\pi_p(ab) = 0$. Étant donné que $(\mathbb{Z}/p\mathbb{Z})[X]$ est intègre (théorème 3.104), nous avons soit $\pi_p(P_1) = 0$ soit $\pi_p(Q_1) = 0$. Supposons pour fixer les idées que $\pi_p(P_1) = 0$. Alors $P_1 = pP_2$ pour un certain $P_2 \in \mathbb{Z}[X]$. Par ailleurs P est unitaire et $P_1 = aP$, donc le coefficient de plus haut degré de P_1 est a, et nous concluons que p divise a.

Mettons $a = pa'$. Dans ce cas, $pa'P = P_1 = pP_2$, et donc $a'P = P_2 \in \mathbb{Z}[X]$. Cela contredit la minimalité de a. □

6.3.5 Polynômes sur un corps et pgcd

Nous savons qu'un corps est un anneau intègre (lemme 1.198). De plus l'ensemble des polynômes sur un anneau intègre est lui-même un anneau intègre (théorème 3.104). Donc la notion de pgcd à utiliser dans le cas de $\mathbb{K}[X]$ est celle de la définition 1.186.

LEMooXISOooNAMeVX

Lemme 6.52 (Unicité du pgcd à inversibles près)**.**
Soit un corps commutatif $\mathbb{K}$ et $S \subset \mathbb{K}[X]$.

(1) Si δ_1 et δ_2 sont des pgcd[23] de S, alors $\delta_1 = k\delta_2$ avec $k \in \mathbb{K}$.

ITEMooARLJooOQvBcw

(2) Il existe un unique polynôme unitaire dans $\gcd(S)$.

Démonstration. Nous savons que δ_1 est un pgcd de S, mais que δ_2 divise S. Donc $\delta_2 \mid \delta_1$. De la même manière, $\delta_1 \mid \delta_2$. Il existe donc $A, B \in \mathbb{K}[X]$ tels que $\delta_1 = A\delta_2$ et $\delta_2 = B\delta_1$. En substituant,

$$\delta_1 = AB\delta_1. \tag{6.119}$$

22. théorème 6.47.
23. Définition 1.186.

Mais $\mathbb{K}[X]$ possède la propriété de simplification par la proposition 1.197(3). Donc $AB = 1$. Cela signifie entre autres que A et B sont des inversibles de $\mathbb{K}[X]$.

Or les seuls inversibles dans $\mathbb{K}[X]$ sont les éléments de $\mathbb{K}$; si vous en doutez, pensez que le degré de AB est supérieur ou égal à celui de A.

En ce qui concerne l'existence et l'unicité d'un polynôme unitaire dans $\text{pgcd}(S)$. Existence : si $P = \sum_{k=0}^{n} a_k X^k \in \text{pgcd}(S)$, alors le polynôme P/a_n est unitaire et également dans $\text{pgcd}(S)$. Unicité : supposons que P est unitaire dans $\text{pgcd}(S)$ et que Q est dans $\text{pgcd}(S)$. Il existe $k \in \mathbb{K}$ tel que $Q = kP$. Vu que P est unitaire, Q ne l'est pas. □

NORMooUJDJooWfijxT

6.53.
En général, lorsque nous dirons « le » pgcd d'un ensemble de polynômes, nous parlerons du pgcd unitaire, qui existe et est bien défini par le lemme 6.52. En particulier, si P et Q sont des polynômes, nous notons $\text{pgcd}(P, Q)$ l'unique polynôme unitaire parmi les PGCD de P et Q.

LEMooIAGMooHUQtUs

Lemme 6.54 ([166])**.**
Soit un corps commutatif $\mathbb{K}$, deux polynômes quelconques $A, B \in \mathbb{K}[X]$ et un polynôme unitaire G.

Nous avons $G = \text{pgcd}(A, B)$ si et seulement si les deux conditions suivantes sont satisfaites :

(1) Il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = G$,

(2) G divise A et B.

Démonstration. Une implication dans chaque sens.

(1) $\Rightarrow$
Si G est le pgcd de A et B, il est clair que $G|A$ et $G|B$. Il reste donc à montrer l'existence des polynômes U et V vérifiant $AU + BV = G$. Puisque G divise A et B, il existe des polynômes A_1, B_1 tels que $A = GA_1$ et $B = GB_1$.
Nous montrons que les polynômes A_1 et B_1 sont premiers entre eux. S'ils ont un diviseur commun D, alors GD est un diviseur commun à A et B. Or, G est le pgcd de A et B donc $GD|G$; D ne peut être qu'un polynôme constant (c'est-à-dire un élément de $\mathbb{K}$). Mais comme G est unitaire, le coefficient du terme de plus haut degré de GD doit être 1. Donc $D = 1$. L'élément 1 est l'unique diviseur commun de A_1 et B_1 ; donc A_1 et B_1 sont bien premiers entre eux.
D'après le théorème de Bézout 6.47, il existe donc U et V tels que $A_1U + B_1V = 1$. En multipliant par G, nous obtenons l'égalité voulue : $AU + BV = G$.

(2) $\Leftarrow$
Si G vérifie les deux conditions, montrons que G est le pgcd de A et B. Nous savons déjà (par hypothèse) que G divise A et B, il reste à montrer que tous les diviseurs commun à A et B divisent aussi G. Soit donc D un diviseur commun à A et B : il existe A_1 et B_1 tels que $A = DA_1$ et $B = DB_1$. Nous savons que $G = AU + BV$ donc $G = D(A_1U + B_1V)$, et $D|G$.
Par définition, G est bien le pgcd de A et B.

□

Notons qu'en supprimant la condition d'unitarité de G, le résultat tient presque : il suffit de remplacer partout « le pgcd » par « un pgcd ».

LEMooVFBUooLoaOGr

Lemme 6.55 ([167])**.**
Soit un corps $\mathbb{K}$. Soient des polynôme P et Q dans $\mathbb{K}[X]$. Si S est un diviseur commun de P et Q, alors

ITEMooZTCCooPBqzNE

(1) S divise $\text{pgcd}(P, Q)$

ITEMooIIINooUZVtAb

(2) $\deg(S) \leqslant \deg\big(\text{pgcd}(P, Q)\big)$.

Démonstration. Il faut bien relire la définition 1.186 du pgcd ; en particulier le point (2) dit exactement notre point (1).

Étant donné que S divise $\text{pgcd}(P,Q)$ nous avons automatiquement que le degré de S est plus petit que celui de $\text{pgcd}(P,Q)$. □

LEMooGNAMooXRpgBn

Lemme 6.56 ([166])**.**
Soient deux polynômes A, B premiers entre eux. Si le polynôme P est divisible par A et par B alors P est divisible par AB.

Démonstration. Comme $A \mid P$, il existe $Q_1 \in \mathbb{K}[X]$ tel que $P = AQ_1$. Mais B divise $P = AQ_1$ alors que B est premier avec A ; donc d'après le théorème de Gauss 6.50 : $B|Q_1$.

Il existe donc $Q_2 \in \mathbb{K}[X]$ tel que $Q_1 = BQ_2$. On a donc $P = ABQ_2$: P est bien divisible par AB. □

LemUELTuwK

Lemme 6.57 ([166])**.**
Quelques propriétés du PGCD[24] *dans les polynômes. Soient des polynômes $P, Q, R \in \mathbb{K}[X]$.*

ITEMooBPOZooYeFGjl

(1) Nous avons l'égalité[25]

$$\text{pgcd}(P, PQ + R) = \text{pgcd}(P, R). \tag{6.120}$$

ITEMooUVGRooNSGDZn

(2) Si Q et R sont premiers entre eux,

$$\text{pgcd}(P, QR) = \text{pgcd}(P, Q)\,\text{pgcd}(P, R) \tag{6.121}$$

ITEMooYXAHooXibkgV

(3) Si P et Q sont premiers entre eux,

$$\text{pgcd}(P, QR) = \text{pgcd}(P, R) \tag{6.122}$$

Démonstration. Dans la suite si A et B sont des polynômes, nous dirons « les diviseurs de $\{A, B\}$ » pour parler des diviseurs communs de A et B.

(1) **(1)**
Nous montrons que $\{P, PQ + R\}$ a les mêmes diviseurs que $\{P, R\}$.
D'une part, si $A \mid \{P, PQ + R\}$, alors il existe des polynômes B_1 et B_2 tels que $P = AB_1$ et $PQ + R = AB_2$. Donc

$$R = AB_2 - PQ = AB_2 - AB_1Q = A(B_2 - B_1Q), \tag{6.123}$$

et nous concluons que A divise R.
D'autre part, si $A \mid \{P, R\}$ alors il existe des polynômes B_1 et B_2 tels que $P = AB_1$ et $R = AB_2$. Donc

$$PQ + R = AB_1Q + AB_2 = A(B_1Q + B_2), \tag{6.124}$$

et A divise $PQ + R$.
Conclusion : les paires $\{P, PQ + R\}$ et $\{P, R\}$ ont même ensemble de diviseurs, et donc même pgcd.

(2) **(2)**
Nous avons trois polynômes P, Q, R et nous savons que Q et R sont premiers entre eux. Nous notons : $G_1 = \text{pgcd}(P, Q)$ et $G_2 = \text{pgcd}(P, R)$. Il faut montrer que G_1G_2 est le pgcd de P et QR ; pour cela nous allons utiliser le lemme 6.54.

24. Définition 1.186.
25. Notez l'analogie avec le lemme 3.6.

(2a) $\exists U, V$ **tels que** $G_1G_2 = PU + QRV$
Puisque $G_1 = \text{pgcd}(P, Q)$, il existe U_1 et V_1 tels que $G_1 = PU_1 + QV_1$ (lemme 6.54). On a de même : $G_2 = PU_2 + RV_2$. En prenant le produit :

$$G_1G_2 = (PU_1 + QV_1)(PU_2 + RV_2) = P(PU_1U_2 + RU_1V_2 + QV_1V_2) + QR(V_1V_2). \quad (6.125)$$

Donc c'est bon pour ce point.

(2b) G_1 **et** G_2 **sont premiers entre eux**
Si D est un diviseur commun à G_1 et G_2, alors D divise Q et R qui sont premiers entre eux ; D ne peut être qu'un polynôme constant. Tous les diviseurs communs de G_1 et G_2 sont dans $\mathbb{K}$. Mais le pgcd est par définition un diviseur commun unitaire, donc $\text{pgcd}(G_1, G_2) = 1$. Cela signifie que G_1 et G_2 sont premiers entre eux (définition 1.259).

(2c) $G_1G_2 \mid QR$ En effet : $G_1|Q$ et $G_2|R$ donc $G_1G_2|QR$.

(2d) $G_1G_2 \mid P$ Le polynôme P est divisible par G_1 et par G_2, et de plus G_1 et G_2 sont premiers entre eux. Donc le lemme 6.56 conclut que P est divisible par G_1G_2.

(3) **(3)**
Supposons d'abord que $A \in \mathbb{K}[X]$ divise P et QR. Le théorème de Bézout 6.47 assure l'existence de polynômes U et V tels que $PU + QV = 1$. Ensuite l'hypothèse de division nous donne des polynômes B_1 et B_2 tels que $P = AB_1$ et $QR = AB_2$. Nous avons :

$$1 = PU + QV = AB_1U + QV. \quad (6.126)$$

Cela prouve que A est premier avec Q grâce encore à Bézout, mais dans l'autre sens. Donc A est premier avec Q et $A \mid QR$. Donc $A|R$ par le théorème de Gauss 6.50.
Dans l'autre sens, si $A|R$ alors on a évidemment : $A|QR$.
Les diviseurs de $\{P, QR\}$ sont exactement les diviseurs de $\{P, R\}$. En conséquence, nous concluons que les paires $\{P, QR\}$ et $\{P, R\}$ ont le même pgcd.

□

6.4 Extension de corps

SECooLQVdmoBCddR

Lemme 6.58.
Soit $\mathbb{L}$ un corps[26] *fini et $\mathbb{K}$ un sous-corps de $\mathbb{L}$. Alors il existe $s \in \mathbb{N}$ tel que*

$$\text{Card}(\mathbb{L}) = \text{Card}(\mathbb{K})^s. \quad (6.127)$$

EqUzqlIQ

Démonstration. Le corps $\mathbb{L}$ est un $\mathbb{K}$-espace vectoriel de dimension finie. Si s est la dimension alors nous avons la formule (6.127) parce que chaque élément de $\mathbb{L}$ est un s-uple d'éléments de $\mathbb{K}$. □

DEFooFLJJooGJYDOe

Définition 6.59 ([168]).
Soit $\mathbb{K}$ un corps commutatif. Une ***extension*** *de $\mathbb{K}$ est un couple $(\mathbb{L}, j)$ où $\mathbb{L}$ est un corps et $j : \mathbb{K} \to \mathbb{L}$ est un morphisme de corps.*

Nous identifions le plus souvent $\mathbb{K}$ avec $j(\mathbb{K}) \subset \mathbb{L}$, mais il faut savoir que le corps $\mathbb{L}$ étendant $\mathbb{K}$ n'est pas toujours un sur-corps de $\mathbb{K}$. En particulier, l'ensemble $\mathbb{L}$ peut ne pas être une extension de l'ensemble $\mathbb{K}$.

LemooOLIIooXzdppM

Lemme-Définition 6.60.
Si $(\mathbb{L}, i)$ est une extension de $\mathbb{K}$, alors $\mathbb{L}$ devient un espace vectoriel sur $\mathbb{K}$ si nous posons

$$\lambda \cdot x = i(\lambda)x \quad (6.128)$$

pour tout $\lambda \in \mathbb{K}$ et $x \in \mathbb{L}$. La multiplication du membre de droite est celle du corps $\mathbb{L}$.

26. Définition 1.209.

DefUYiyieu

Définition 6.61.
*Le **degré** de $\mathbb{L}$ est la dimension de cet espace vectoriel. Il est noté $[\mathbb{L} : \mathbb{K}]$; notons qu'il peut être infini.*

PROPooCFGUooCQqWZd

Proposition 6.62.
Le corps $\mathbb{C}$ est une extension de $\mathbb{R}$ et son degré[27] *est $[\mathbb{C} : \mathbb{R}] = 2$.*

PROPooPGSCWBSMpM

Proposition 6.63 (Composition des degrés[169])**.**
Si $\mathbb{L}_2$ est une extension de $\mathbb{L}_1$ qui est elle-même une extension de $\mathbb{K}$, alors $\mathbb{L}_2$ est une extension de $\mathbb{K}$ et on a :

EQooOLLQooFdYtnh

$$[\mathbb{L}_2 : \mathbb{K}] = [\mathbb{L}_2 : \mathbb{L}_1][\mathbb{L}_1 : \mathbb{K}]. \tag{6.129}$$

Dans ce cas, si $\{v_i\}_{i\in I}$ est une $\mathbb{K}$-base de $\mathbb{L}_1$ et si $\{w_\alpha\}_{\alpha\in A}$ est une $\mathbb{L}_1$-base de $\mathbb{L}_2$ alors $\{v_i w_\alpha\}_{\substack{i\in I\\ \alpha\in A}}$ est une $\mathbb{K}$-base de $\mathbb{L}_2$.

Notons que la formule (6.129) n'est pas très instructive dans le cas des extensions non finies. La seconde partie, sur les bases, est en réalité nettement plus intéressante.

Démonstration. Soit $a \in \mathbb{L}_2$. Puisque les w_α forment une $\mathbb{L}_1$-base de $\mathbb{L}_2$, il existe des $a_\alpha \in \mathbb{L}_1$ tels que

$$a = \sum_\alpha a_\alpha w_\alpha. \tag{6.130}$$

Mais les v_i forment une $\mathbb{K}$-base de $\mathbb{L}_1$, donc chacun des a_α peut être décomposé comme $a_\alpha = \sum_i a_{\alpha i} v_i w_\alpha$. Donc :

$$a = \sum_{\alpha i} a_{\alpha i} v_i w_\alpha, \tag{6.131}$$

qui donne une décomposition de a en éléments de $\{v_i w_\alpha\}$ à coefficients dans $\mathbb{K}$. La partie proposée est donc génératrice.

Pour prouver qu'elle est également libre, nous supposons avoir des éléments $a_{\alpha i} \in \mathbb{K}$ tels que

$$\sum_{\alpha i} a_{\alpha i} v_i w_\alpha = 0. \tag{6.132}$$

En récrivant sous la forme

$$\sum_\alpha \Big(\sum_i a_{\alpha i} v_i\Big) w_\alpha = 0, \tag{6.133}$$

nous reconnaissons une combinaison linéaire nulle des w_α à coefficients dans $\mathbb{L}_1$. Les coefficients sont donc nuls : $\sum_i a_{\alpha i} v_i = 0$. C'est une combinaison linéaire nulle des v_i à coefficients dans $\mathbb{K}$. Comme les v_i forment une base, les coefficients sont nuls : $a_{\alpha i} = 0$. □

6.4.1 Extension et polynôme minimal

DefCVMooFGSAgL

Lemme-Définition 6.64 (Polynôme minimal)**.**
Soit $\mathbb{L}$ une extension de $\mathbb{K}$ et $a \in \mathbb{L}$. Nous considérons la partie

$$I_a = \{P \in \mathbb{K}[X] \text{ tel que } P(a) = 0\} \tag{6.134}$$

que nous supposons non réduite à $\{0\}$[28]

ITEMooUNLCooIfYZry

(1) La partie I_a est un idéal dans $\mathbb{K}[X]$,

ITEMooDCDRooPDnnbu

(2) la partie I_a est un idéal principal dans $\mathbb{K}[X]$,

ITEMooXFYQooTuMzIu

(3) l'idéal I_a possède un unique générateur unitaire.

27. Définition 6.61.
28. La non trivialité de I_a est une vraie hypothèse. En effet si nous prenons $\mathbb{K} = \mathbb{Q}$ et l'extension $\mathbb{L} = \mathbb{R}$, alors il suffit de prendre un réel a non algébrique sur $\mathbb{Q}$ pour que I_a soit réduit au seul polynôme identiquement nul.

Cet unique générateur unitaire est le ***polynôme minimal*** *de a sur $\mathbb{K}$.*

Démonstration. En plusieurs parties.

(1) **Pour (1)** Soit $P \in I_a$: $P(a) = 0$. Si $Q \in \mathbb{K}[X]$ alors la proposition 1.363 nous indique que

$$(PQ)(a) = P(a)Q(a) = 0. \tag{6.135}$$

Donc $PQ \in I_a$. Comme de plus I_a est clairement vectoriel, I_a est un idéal.
Notez que nous avons utilisé la règle du produit nul justifiée par le fait que $\mathbb{K}$ soit un corps [29] et donc soumis au point (3) de la proposition 1.197.

(2) **Pour (2)** Nous savons par le théorème 6.43 que $\mathbb{K}[X]$ est un anneau principal. En particulier, tous ses idéaux sont principaux, c'est dans la définition 1.228 d'un anneau principal.

(3) **Pour (3)** Le théorème 6.43(3) nous informe alors que I_a possède un unique générateur unitaire.

□

Si nous avons un corps et un élément dans une extension du corps, il n'est pas autorisé de dire « soit le polynôme minimal de cet élément dans le premier corps » parce qu'il n'existe peut-être pas de polynôme annulateur.

6.65.
Dans le cas des opérateurs sur un espace de dimension finie (par exemple les matrices), il existe toujours un polynôme minimal, comme nous le verrons dans le lemme 8.97.

Exemple 6.66.
Le polynôme minimal dépend du corps sur lequel on le considère. Par exemple le nombre imaginaire pur i accepte $X - i$ comme polynôme minimal sur $\mathbb{C}$ et $X^2 + 1$ sur $\mathbb{Q}[X]$. △

PropRARooKavaIT

Proposition 6.67 ([1])**.**
Soit $\mathbb{L}$ une extension de $\mathbb{K}$ et $a \in \mathbb{L}$ dont le polynôme minimal sur $\mathbb{K}$ est $\mu_a \in \mathbb{K}[X]$. Alors itemDOQooYpLvXri

(1) le polynôme μ_a est irréductible [30] sur $\mathbb{K}$;
(2) Le polynôme μ_a est premier [31] avec tout polynôme de $\mathbb{K}[X]$ non annulateur de a.

Démonstration. Une chose à la fois.

(1) D'abord le polynôme μ_a n'est pas inversible parce que seuls les éléments de $\mathbb{K}$ (ceux de degré zéro) peuvent être inversibles [32]. Mais ces polynômes sont constants et ne peuvent donc pas être des polynômes annulateurs de quoi que ce soit.
Ensuite, supposons la décomposition $\mu_a = PQ$ avec $P, Q \in \mathbb{K}[X]$. En évaluant cette égalité en a nous avons

$$0 = P(a)Q(a). \tag{6.136}$$

Puisque nous sommes sur un corps, nous avons la règle du produit nul [33] et nous déduisons que soit $P(a)$ soit $Q(a)$ est nul, ou les deux. Pour fixer les idées, nous supposons $P(a) = 0$.
Dans ce cas, P fait partie de l'idéal annulateur de a, lequel idéal est engendré par μ_a. Donc il existe $S \in \mathbb{K}[X]$ tel que $P = S\mu_a$. En récrivant $\mu_a = PQ$ avec cela nous avons :

$$\mu_a = S\mu_a Q \tag{6.137}$$

ou encore : $SQ = 1$, ce qui signifie que S et Q sont dans $\mathbb{K}$ et inversibles.
Nous concluons que μ_a ne peut pas être écrit sous forme de produit de deux non inversibles.

29. Si vous connaissez un contre-exemple à cette proposition dans le cas où $\mathbb{K}$ serait remplacé par un anneau, écrivez-moi.
30. Définition 1.189.
31. Définition 3.111.
32. Et d'ailleurs, le sont, mais ce n'est pas important ici.
33. Parce qu'un corps est un anneau intègre par le lemme 1.198 et qu'un anneau intègre est justement un anneau sur lequel nous avons la règle du produit nul, voir la définition 1.197.

(2) Soit Q un polynôme non annulateur de a. Soit aussi un diviseur commun P de Q et μ_a dans $\mathbb{K}[X]$. Nous devons prouver que P est un inversible, c'est-à-dire un élément de $\mathbb{K}$ (le fait que P ne soit pas le polynôme nul est évident). Nous avons $\mu_a = PR_1$ et $Q = PR_2$ pour certains polynômes $R_1, R_2 \in \mathbb{K}[X]$. Puisque μ_a est irréductible par (1), il n'est pas produit de deux non inversibles. En d'autres termes, soit P soit R_1 est inversible. Si P n'est pas inversible, alors R_1 est inversible ; disons $R_1 = k \in \mathbb{K}$. Alors

$$0 = \mu_a(a) = P(a)k, \tag{6.138}$$

donc $P(a) = 0$. Mais alors

$$Q(a) = P(a)R_2(a) = 0, \tag{6.139}$$

ce qui est contraire à l'hypothèse selon laquelle Q n'était pas annulateur de a.

Nous retenons donc que P est inversible, ce qu'il fallait montrer.

□

DEFooJQQYooMKqWov

Définition 6.68.
*Deux éléments α et β dans $\mathbb{L}$ sont dit **conjugués** s'ils ont même polynôme minimal. Par exemple i et $-i$ sont conjugués dans $\mathbb{C}$ vu comme extension de $\mathbb{Q}$.*

6.4.2 Extensions algébriques et éléments transcendants

6.4.2.1 Éléments algébriques et transcendants

6.69.
Mise en garde pour les gens qui aiment avoir des notations précises. Dans les lemmes 6.71 et 6.73 ainsi que dans la définition 6.75 il faut garder en tête qu'une extension $\mathbb{L}$ de $\mathbb{K}$ est une application $\iota : \mathbb{K} \to \mathbb{L}$. Il n'est pas spécialement vrai que $\mathbb{K}$ est un sous-sensemble de $\mathbb{L}$.

Pour être rigoureux, beaucoup d'énoncés devraient contenir plus d'applications ι un peu partout.

DEFooBBYGooWoOloR

Définition 6.70.
L'ensemble $A[X]$ devient un $\mathbb{K}$-espace vectoriel avec la définition

$$(\lambda P)_k = \lambda P_k. \tag{6.140}$$

Voici une définition d'un élément algébrique sur un corps. Une caractérisation plus « pratique » sera donnée dans le lemme 6.73.

LEMooLVPLooEkWYDN

Lemme-Définition 6.71 (Élément algébrique et transcendant[170])**.**
Soit une extension $\mathbb{L}$ de $\mathbb{K}$ et $\alpha \in \mathbb{L}$. Nous considérons l'application

$$\begin{aligned} \varphi : \mathbb{K}[X] &\to \mathbb{L} \\ P &\mapsto P(\alpha). \end{aligned} \tag{6.141}$$

Alors

(1) L'application φ est un morphisme d'anneaux[34]*.*

(2) L'application φ est un morphisme de $\mathbb{K}$-espace vectoriel.

*Si φ est injective, nous disons que α est **transcendant**. Sinon, nous disons qu'il est **algébrique**.*

Démonstration. Le fait que φ soit un morphisme d'anneaux est le lemme 1.363 déjà prouvé.

Pour le morphisme de $\mathbb{K}$-espace vectoriel, il faut seulement ajouter le calcul

$$\varphi(\lambda P) = (\lambda P)(\alpha) = \lambda P(\alpha) = \lambda\varphi(P). \tag{6.142}$$

34. Définition 1.43.

Notons la justification suivante qui n'est pas tout à fait triviale :

$$(\lambda P)(\alpha) = \sum_k (\lambda P)_k \alpha^k = \sum_k \lambda P_k \alpha^k = \lambda P(\alpha) \tag{6.143}$$

qui utilise la définition 6.70. □

Exemple 6.72.
L'injectivité de φ n'est pas automatique. Prenons par exemple $\mathbb{L} = \mathbb{Q}[\sqrt{2}]$ dans $\mathbb{R}$. Les polynômes dans $\mathbb{Q}[X]$ ont des degrés arbitrairement élevés en X, tandis que les éléments de $\mathbb{L}$ n'ont pas de degré très élevés en $\sqrt{2}$ parce que $\sqrt{2}\sqrt{2} = 2$. L'ensemble $\mathbb{Q}[\sqrt{2}]$ ne contient donc que des éléments de la forme $a + b\sqrt{2}$ avec $a, b \in \mathbb{Q}$.

Si par contre $x_0 \in \mathbb{R}$ n'est racine d'aucun polynôme (cela existe parce que $\mathbb{R}$ n'est pas dénombrable), alors $\mathbb{Q}[x_0]$ contient tous les $\sum_{k=0}^{N} a_k x_0^k$ avec N arbitrairement grand. Et tous ces nombres sont différents. △

Le lemme suivant donne une caractérisation d'élément algébrique moins abstraite que la définition 6.71.

LEMooTZSSooZmwYji

Lemme 6.73.
Soit $\mathbb{K}$, un corps et $\mathbb{L}$, une extension de $\mathbb{K}$. Un élément $\alpha \in \mathbb{L}$ est algébrique sur $\mathbb{K}$ si et seulement si existe un polynôme non nul $P \in \mathbb{K}[X]$ tel que $P(\alpha) = 0$.

Démonstration. Nous considérons l'application φ de la définition 6.71. Si φ n'est pas injective, c'est qu'il existe un polynôme P dans $\mathbb{K}[X]$ tel que $\varphi(P) = 0$. Dans ce cas, $P(\alpha) = 0$.

À l'inverse si il existe P non nul dans $\mathbb{K}[X]$ tel que $P(\alpha) = 0$, alors $\varphi(P) = 0$ et φ n'est pas injective. □

DEFooYZOYooAesmnP

Définition 6.74.
*Un corps $\mathbb{K}$ est **algébriquement clos** si tout polynôme non constant à coefficients dans $\mathbb{K}$ contient au moins une racine dans $\mathbb{K}$.*

Nous verrons dans le théorème de d'Alembert 12.90 que $\mathbb{C}$ est un corps algébriquement clos.

DEFooREUHooLVwRuw

Définition 6.75 (Extension algébrique, clôture algébrique)**.**
Soient un corps $\mathbb{K}$ et une extension $\alpha\colon \mathbb{K} \to \mathbb{L}$.

*(1) L'extension $\mathbb{L}$ est une extension **algébrique** de $\mathbb{K}$ si tous ses éléments sont algébriques*[35] *sur $\mathbb{K}$, c'est-à-dire sont racines de polynômes à coefficients dans $\alpha(\mathbb{K})$, voir le lemme 6.73.*
ITEMooEIWVooVjJRoR

*(2) L'extension $\mathbb{L}$ est **algébriquement close** si le corps $\mathbb{L}$ est algébriquement clos (définition 6.74).*

*(3) L'extension $\mathbb{L}$ est une **clôture algébrique** du corps $\mathbb{K}$ si elle est une extension algébrique qui est algébriquement close.*

6.76.
Donc une extension est algébrique si elle contient seulement des racines de polynômes ; elle est close si elle contient au moins une racine de chaque polynôme. L'extension est une clôture algébrique si elle est les deux en même temps.

6.77 (Unicité)**.**
Le théorème de Steinitz 6.135 donnera une unicité de la clôture algébrique.

Exemple 6.78.
Le corps $\mathbb{R}$ n'est pas une extension algébrique de $\mathbb{Q}$. En effet il existe seulement une infinité *dénombrable* de polynômes dans $\mathbb{Q}[X]$ et donc une infinité dénombrable de racines de tels polynômes. Toute extension algébrique de $\mathbb{Q}$ est donc dénombrable. Voir aussi la proposition 6.131. △

35. Définition 6.71.

LEMooEYRSooUREeDl

Lemme 6.79.
Un corps est algébriquement clos si et seulement si tous ses polynômes sont scindés[36].

Démonstration. Si tout polynôme est scindé, tout polynôme possède des racines ; c'est l'autre sens qui est plus consistant.

Soit un corps algébriquement clos $\mathbb{K}$. Nous allons effectuer une récurrence sur le degré des polynômes. Si P est un polynôme de degré 1, alors il est scindé.

Nous supposons que tous les polynômes de degré $n-1$ sont scindés. Soit un polynôme P de degré n. Le corps étant algébriquement clos, le polynôme P a une racine que nous notons $a_n \in \mathbb{K}$. La proposition 3.127 nous explique qu'il existe un polynôme Q de degré $n-1$ tel que $P = (X-a_n)Q$.

Par hypothèse de récurrence, le polynôme Q est scindé : il existe $\{a_i\}_{i=1,\dots,n-1}$ dans $\mathbb{K}$ tels que $Q = \prod_{i=1}^{n-1}(X-a_i)$. Au final,

$$P = (X-a_n)Q = \prod_{k=1}^{n}(X-a_k) \tag{6.144}$$

et P est scindé. □

LEMooYVHKooWhewKp

Lemme 6.80.
Soient un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$. Nous supposons que P est scindé :

$$P = \prod_{k=1}^{n}(X-a_k). \tag{6.145}$$

Si α est une racine de P, alors α est l'un des a_k.

Démonstration. Dire que α est une racine de P revient à dire que

$$\prod_{k=1}^{n}(\alpha-a_k) = 0 \tag{6.146}$$

EQooXHSCooVdMiva

Un corps est toujours un anneau intègre (lemme 1.198), c'est-à-dire que la règle du produit nul est utilisable. Dans notre cas, le produit nul (6.146) nous indique que $\alpha - a_k = 0$ pour (au moins) un des k. Donc effectivement α est l'un des a_k. □

Lemme 6.81.
Soient un corps algébriquement clos $\mathbb{K}$ ainsi qu'une extension algébrique $\alpha\colon \mathbb{K} \to \mathbb{L}$. Alors $\alpha(\mathbb{K}) = \mathbb{L}$.

Démonstration. Nous allons montrer que tous les éléments de $\mathbb{L}$ sont dans l'image de α. Soit donc $l \in \mathbb{L}$. Puisque l'extension $\alpha\colon \mathbb{K} \to \mathbb{L}$ est une extension algébrique, il existe un polynôme $P \in \alpha(\mathbb{K})[X]$ tel que $P(l) = 0$.

Étant donné que α est injective, il est possible de considérer le polynôme $Q = \alpha^{-1}(P)$, c'est-à-dire que, si $P = \sum_k a_k X^k$, nous posons $Q = \sum_k \alpha^{-1}(a_k)X^k$.

Le corps $\mathbb{K}$ étant algébriquement clos, le polynôme Q est scindé (proposition 6.79) :

$$Q = \prod_{k=1}^{n}(X-b_k) \tag{6.147}$$

avec $b_k \in \mathbb{K}$. Nous avons alors aussi la factorisation

$$P = \prod_{k=1}^{n}(X-\alpha(b_k)) \tag{6.148}$$

dans $\mathbb{L}[X]$. Nous avons vu que l était une racine de P. Donc l est un des $\alpha(b_k)$ (lemme 6.80). Cela prouve que $l \in \alpha(\mathbb{K})$. □

36. Définition 6.39

6.4.3 Extension algébrique et polynôme minimal

PROPooALFJooDjmIcb

Proposition 6.82 ([93]).
Soit une extension algébrique[37] $\mathbb{L}$ *du corps* $\mathbb{K}$.

(1) *Pour tout* $a \in \mathbb{L}$, *il existe un polynôme* $P \in \mathbb{K}[X]$ *tel que* $P(a) = 0$.

ITEMooEFNFooKYqXDk

(2) *Le polynôme minimal de* a *dans* $\mathbb{K}[X]$ *est l'unique polynôme unitaire irréductible annulant* a.

Démonstration. Le premier point est seulement la définition 6.75 d'une extension algébrique.

L'idéal annulateur $I_a = \{P \in \mathbb{K}[X] \text{ tel que } P(a) = 0\}$ n'est pas réduit à $\{0\}$ parce que $\mathbb{L}$ est une extension algébrique. L'existence du polynôme minimal est le lemme 6.64 et le fait qu'il soit irréductible est la proposition 6.67(1).

Ce qui nous intéresse ici est l'unicité. Soit $\mu_1 \in \mathbb{K}[X]$, un polynôme annulateur de a irréductible et unitaire. Puisque $\mu_1 \in I_a$ et que par définition, $I_a = (\mu)$, il existe $P \in \mathbb{K}[X]$ tel que $\mu_1 = P\mu$. Comme μ n'est pas inversible et que μ_1 est irréductible, P doit être inversible : $\mu_1 = k\mu$ pour un certain $k \in \mathbb{K}$.

Puisque μ et μ_1 sont unitaires, $k = 1$. Donc $\mu_1 = \mu$. □

LEMooHKTMooKEoOuK

Lemme 6.83.
Soient un corps $\mathbb{K}$, *une extension* $\mathbb{L}$ *de* $\mathbb{K}$ *et* $\alpha \in \mathbb{L}$, *un élément algébrique*[38] *sur* $\mathbb{K}$. *Si* μ *est le polynôme minimal de* α *sur* $\mathbb{K}$ *alors*

$$\begin{aligned} \varphi\colon \mathbb{K}[\alpha] &\to \mathbb{K}[X]/(\mu) \\ Q(\alpha) &\mapsto \bar{Q} \end{aligned} \tag{6.149}$$

avec $Q \in \mathbb{K}[X]$ *est un isomorphisme de corps et de* $\mathbb{K}$*-espaces vectoriels.*

Démonstration. D'abord, α est algébrique, donc l'idéal annulateur I_α n'est pas réduit à $\{0\}$, et l'existence d'un polynôme minimal est assurée par le lemme 6.64.

Ensuite, le fait que $\mathbb{K}[X]/(\mu)$ soit un corps est le corolaire 6.44. Nous montrons à présent que φ est un isomorphisme (d'anneaux) ; cela suffit pour en déduire que $\mathbb{K}[\alpha]$ est également un corps.

Ces préliminaires étant dits, nous commençons.

(1) **Bien définie** Nous devons prouver que φ est bien définie, c'est-à-dire que tout élément de $\mathbb{K}[\alpha]$ peut être écrit sous la forme $Q(\alpha)$ pour un $Q \in \mathbb{K}[X]$, et que si $Q_1(\alpha) = Q_2(\alpha)$ alors $\bar{Q}_1 = \bar{Q}_2$.

Le fait que tous les éléments de $\mathbb{K}[\alpha]$ peuvent être écrits sous la forme $Q(\alpha)$ est la proposition 3.103. Supposons que $Q_1(\alpha) = Q_2(\alpha)$. Alors nous définissons $R \in \mathbb{K}[X]$ par $Q_1 = Q_2 + R$, et en évaluant cette égalité en α nous avons

$$Q_1(\alpha) = Q_2(\alpha) + R(\alpha), \tag{6.150}$$

autrement dit $R(\alpha) = 0$. Donc R est dans l'idéal annulateur de α et est donc dans (μ), c'est-à-dire que dans le quotient $\mathbb{K}[X]/(\mu)$ nous avons $\bar{R} = 0$ et donc $\bar{Q}_1 = \bar{Q}_2$.

(2) **Surjective**

Tout élément de $\mathbb{K}[X]/(\mu)$ est de la forme $\bar{Q}$ pour un $Q \in \mathbb{K}[X]$. Or ces éléments sont ceux de l'ensemble d'arrivée de φ.

(3) **Injective**

Si $\bar{Q}_1 = \bar{Q}_2$, alors $Q_1 = Q_2 + R$ avec R dans l'idéal engendré par μ, c'est-à-dire entre autres $R(\alpha) = 0$. Donc $Q_1(\alpha) = Q_2(\alpha)$.

Nous devons encore montrer que nous avons là un morphisme de $\mathbb{K}$-espaces vectoriels.

37. Définition 6.75.
38. Définition 6.71.

(1) Si $k \in \mathbb{K}$ alors $\varphi\big(kQ(\alpha)\big) = \overline{kQ}$. Mais par définition de la structure d'espace vectoriel sur $\mathbb{K}[X]/(\mu)$, $\overline{kQ} = k\bar{Q}$ (vérifier que cette définition de la multiplication par un scalaire sur $\mathbb{K}[X]/(\mu)$ est correcte).
(2) Nous avons aussi $\varphi\big(Q_1(\alpha) + Q_2(\alpha)\big) = \varphi\big((Q_1 + Q_2)(\alpha)\big) = \overline{Q_1 + Q_2} = \bar{Q}_1 + \bar{Q}_2$.

□

6.4.4 Extensions et polynômes

Nous savons déjà depuis la définition 1.359 ce qu'est $A[X]$ pour tout anneau A et donc, à fortiori, pour un corps.

DEFooQPZIooQYiNVh

Définition 6.84.
Soit un corps commutatif[39]*. Nous notons $\mathbb{K}(X)$ le corps des fractions*[40] *de $\mathbb{K}[X]$.*

DEFooZHBZooKlNfGZ

Lemme-Définition 6.85.
Si $R \in \mathbb{K}(X)$, avec $R = P/Q$ et si $\mathbb{L}$ est une extension[41] *de $\mathbb{K}$ contenant l'élément α, alors nous définissons*

$$R(\alpha) = P(\alpha)Q(\alpha)^{-1}. \tag{6.151}$$

Cela est une bonne définition au sens où elle ne dépend pas du choix du représentant (P, Q) pris dans la classe P/Q.

Démonstration. Supposons $R = P_1/Q_1 = P_2/Q_2$. Par définition des classes (définition 1.369) nous avons

EQooKHVNooABuHaQ
$$P_1Q_2 = Q_1P_2. \tag{6.152}$$

Puisque l'évaluation est un morphisme $\mathbb{K}[X] \to \mathbb{K}$[42] nous pouvons évaluer l'équation (6.152) en α :

EQooJAIGooRADgiD
$$P_1(\alpha)Q_2(\alpha) = Q_1(\alpha)P_2(\alpha). \tag{6.153}$$

Cette dernière est une égalité dans le corps $\mathbb{K}$. Nous pouvons donc la multiplier par $Q_2(\alpha)^{-1}Q_1(\alpha)^{-1}$ (et utiliser toutes les hypothèses de commutativité des anneaux et corps) pour obtenir

$$P_1(\alpha)Q_1(\alpha)^{-1} = P_2(\alpha)Q_2(\alpha)^{-1}, \tag{6.154}$$

c'est-à-dire

$$(P_1/Q_1)(\alpha) = (P_2/Q_2)(\alpha). \tag{6.155}$$

□

DEFooVSKGooMyeGel

Proposition-Définition 6.86 ([1])**.**
Soient un corps $\mathbb{K}$, une extension $(\mathbb{L}, j_{\mathbb{L}})$ de $\mathbb{K}$ et un élément $\alpha \in \mathbb{L}$. Nous définissons $\mathbb{K}(\alpha)_{\mathbb{L}}$ comme étant l'intersection de tous les sous-corps de $\mathbb{L}$ contenant $j_{\mathbb{L}}(\mathbb{K})$ et α.

Alors

(1) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est un sous-corps de $\mathbb{L}$,
(2) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est une extension[43] *de $\mathbb{K}$.*

Démonstration. Nous commençons par prouver que $\mathbb{K}(\alpha)_{\mathbb{L}}$ est bien un corps. Si $a, b \in \mathbb{K}(\alpha)_{\mathbb{L}}$ alors il suffit de calculer ab, $a + b$ et a^{-1} dans n'importe quel sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α ; nous avons une garantie que a, b, ab, $a + b$ et a^{-1} sont dans tous les tels sous-corps.

Pour prouver que $\mathbb{K}(\alpha)_{\mathbb{L}}$ est bien une extension, nous devons trouver un morphisme de corps $j : \mathbb{K} \to \mathbb{K}(\alpha)_{\mathbb{L}}$. On constate que prendre $j = j_{\mathbb{L}}$ fonctionne parce que par définition, $\mathbb{K}(\alpha)_{\mathbb{L}}$ est une partie de $\mathbb{L}$ contenant l'image de $j_{\mathbb{L}}$. □

39. Sauf mention du contraire, tous les corps du Frido sont commutatifs.
40. Définition 1.369.
41. Définition 6.59.
42. Lemme 1.363. Certes ce lemme ne parle que d'anneaux, mais à y bien penser, dans le passage de (6.152) à (6.153), nous ne considérons que les structures d'anneaux sur $\mathbb{K}[X]$ et $\mathbb{K}$.
43. Définition 6.59.

LEMooHZLCooPLHkLS

Lemme 6.87.
Soit n tel que $\sqrt{n}$ ne soit pas un rationnel. Si $\alpha \in \{a+b\sqrt{n}\}_{a,b\in\mathbb{Q}}$, alors il existe un unique choix $(x,y) \in \mathbb{Q}^2$ tel que

$$\alpha = x + y\sqrt{n}. \tag{6.156}$$

Démonstration. L'existence est dans la définition de α. Il s'agit de voir l'unicité. Supposons $x + y\sqrt{n} = a + b\sqrt{n}$ avec $x, y, a, b \in \mathbb{Q}$. Si $b \neq y$ nous pouvons écrire

$$\sqrt{n} = \frac{x-a}{b-y}. \tag{6.157}$$

Comme $\sqrt{n}$ n'est pas un rationnel, une telle écriture est impossible. Donc $b = y$. Nous avons alors $x + y\sqrt{n} = a + y\sqrt{n}$ et du coup aussi $x = a$. □

Exemple 6.88.
Nous avons

$$\mathbb{Q}(\sqrt{2})_{\mathbb{R}} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}} \tag{6.158}$$

où à droite nous calculons les sommes et les produits dans $\mathbb{R}$. Le tout est un sous-ensemble de $\mathbb{R}$ qui se révèle être un corps contenant $\mathbb{Q}$ et $\sqrt{2}$.

En particulier, dans $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ nous avons $\sqrt{2}\sqrt{2} = 2$. △

LEMooKVPZooPqPrce

Lemme 6.89.
Les corps $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ et $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ ne sont pas isomorphes.

Démonstration. Supposons l'existence d'un morphisme de corps [44]

$$\psi\colon \mathbb{Q}(\sqrt{2})_{\mathbb{R}} \to \mathbb{Q}(\sqrt{3})_{\mathbb{R}}. \tag{6.159}$$

Nous notons « 1 » à la fois le neutre de la multiplication dans $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ et $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ (qui s'avèrent être les mêmes en tant qu'élément de $\mathbb{R}$, mais ça n'a pas d'importance ici).

Soit $\alpha \in \mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ tel que $\alpha^2 - 2 = 0$. Alors nous avons aussi

$$\psi(\alpha)^2 - 2 = \psi(\alpha^2) - \psi(2) = \psi(\alpha^2 - 2) = \psi(0) = 0. \tag{6.160}$$

Donc $\psi(\alpha)$ est un élément de $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ qui est une racine de $X^2 - 1$.

Or un tel élément n'existe pas dans $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ parce que nous savons que dans $\mathbb{R}$ entier, il n'y a que deux racines : $\pm\sqrt{2}$, et aucune des deux n'est dans $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$. □

EXooJRSUooYhAZkR

Exemple 6.90.
Est-ce que $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend réellement de $\mathbb{L}$? Si $\mathbb{L}_2$ est une extension de $\mathbb{L}$ alors nous avons évidemment [45] $\mathbb{K}(\alpha)_{\mathbb{L}_2} = \mathbb{K}(\alpha)_{\mathbb{L}}$.

Nous commençons par construire un corps $\mathbb{K}$ un peu idiot qui, comme ensemble, est comme $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$, c'est-à-dire la partie

$$\{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}, \tag{6.161}$$

de $\mathbb{R}$.

Mais cette fois nous définissons la multiplication suivante :

$$(a + b\sqrt{2})(c + d\sqrt{2}) = ac + 3bd + (ad + bc)\sqrt{2}. \tag{6.162}$$

C'est un corps parce que tout élément non nul est inversible. En effet, l'équation

EQooIZLEooLPOBcC

$$(a + b\sqrt{2})(x + y\sqrt{2}) = 1 \tag{6.163}$$

44. Définition 1.43. Oui, c'est un bête morphisme d'anneaux. Il n'y a pas plus de structure dans un corps que dans un anneau.

45. Vérifiez-le tout de même.

donne

$$\begin{pmatrix} a & 3b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{6.164}$$

Ce système a une unique solution si et seulement si $\det \begin{pmatrix} a & 3b \\ b & a \end{pmatrix} = 0$. Cela survient si et seulement si

$$a^2 - 3b^2 = 0. \tag{6.165}$$

Les solutions de cette équation dans $\mathbb{R}$ sont $a = \pm\sqrt{3}|b|$. Dès que a ou b est non nul, cela ne peut pas satisfaire $a, b \in \mathbb{Q}$. Donc le déterminant est toujours non nul et il existe $x, y \in \mathbb{Q}$ tels que (6.163) soit satisfaite.

Tout cela nous a donné un corps $\mathbb{K}$ dont $\mathbb{Q}$ est un sous-corps et qui contient l'élément $\sqrt{2}$ de $\mathbb{R}$. Il n'est cependant pas un sous-corps de $\mathbb{R}$.

Ce corps est isomorphe à $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$. En effet, nous montrons que

$$\begin{aligned} \psi\colon \mathbb{K} &\to \mathbb{Q}(\sqrt{3})_{\mathbb{R}} \\ a + b\sqrt{2} &\mapsto a + b\sqrt{3} \end{aligned} \tag{6.166}$$

est un isomorphisme de corps. Pour le produit, nous avons

$$\begin{aligned} \psi\big((a + b\sqrt{2})(c + d\sqrt{2})\big) &= \psi\big(ac + 3bd + (ad + bc)\sqrt{2}\big) && \text{SUBEQooQSZBooHZDTKo (6.167a)} \\ &= ac + 3bd + (ad + bc)\sqrt{3} && \text{SUBEQooPEKHooNPcIjE (6.167b)} \\ &= (a + b\sqrt{3})(c + d\sqrt{3}) && \text{SUBEQooIGBZooMwrmFe (6.167c)} \\ &= \psi(a + b\sqrt{2})\psi(c + d\sqrt{2}). && (6.167d) \end{aligned}$$

Remarques :

— L'application ψ est bien définie grâce au lemme 6.87 couplé au théorème 3.36 appliqué à $n = 2$ et $n = 3$.

— Dans le membre de gauche de (6.167a), $b\sqrt{2}$ est un produit dans $\mathbb{R}$ (d'où l'importance du lemme 6.87 qui permet de re-séparer les éléments de $\mathbb{R}$ partie rationnelle et partie multiple de $\sqrt{2}$), et le produit entre $(a + b\sqrt{2})$ et $(c + d\sqrt{2})$ est un produit dans $\mathbb{K}$.

— Dans (6.167b) et (6.167c), tous les produits sont dans $\mathbb{R}$.

En comparant avec le lemme 6.89, nous avons alors

$$\mathbb{Q}(\sqrt{2})_{\mathbb{K}} = \mathbb{Q}(\sqrt{3})_{\mathbb{R}} \neq \mathbb{Q}(\sqrt{2})_{\mathbb{R}} \tag{6.168}$$

△

6.91.
Nous allons encore enfoncer le clou sur le fait que $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend de $\mathbb{L}$.

Le fait est que si on y pense, l'objet $\sqrt{2}$ n'a aucun rapport avec $\mathbb{Q}$. En effet les objets de $\mathbb{Q}$ sont des classes d'équivalence de couples d'éléments de $\mathbb{Z}$, alors que l'élément $\sqrt{2}$ est une classe d'équivalence de suites de Cauchy dans $\mathbb{Q}$.

Lorsque nous écrivons $\mathbb{Q}(\sqrt{2})$, nous associons des objets de nature complètement différentes, et il n'y a aucune raison à priori de définir la multiplication entre eux d'une façon plutôt qu'une autre.

Plus généralement, dans ZF (nous faisons semblant de suivre ZF tout en sachant que nous ne savons pas ce que c'est réellement [46]), tout est ensemble. Peut-on dire ce que serait $\mathbb{Q}(I)$ si I est un ensemble quelconque ? Attention : en écrivant $\mathbb{Q}(I)$, nous entendons un corps dont I est un élément, pas un corps qui contiendrait comme éléments tous les éléments de I.

Si I est juste un ensemble, quelle définition donner de I^2 ? Il y a plein de choix et rien ne se dégage clairement comme étant pertinent. Si par contre, en guise de I nous considérons l'ensemble

46. En lisant quelques pages de Wikipédia, vous pourrez briller en société, mais ne tentez pas le coup à l'agrégation.

$\sqrt{2}$ (oui, c'est un ensemble : un ensemble de suites de Cauchy dans $\mathbb{Q}$), alors tout de suite nous nous disons que la bonne façon de faire est $\sqrt{2}^2 = 2$. Ce réflexe est juste conditionné par le fait que nous connaissons déjà par ailleurs le corps $\mathbb{R}$. Rien de plus.

Donc oui, $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend de $\mathbb{L}$, mais dans les cas particuliers où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$, il y a une égalité implicite $\mathbb{L} = \mathbb{C}$. Cela étant dit, il n'y a plus d'ambiguïté en écrivant $\mathbb{Q}(\sqrt{2})$.

EXooKLRBooUswzpV

Exemple 6.92.
Soit $\mathbb{K} = \mathbb{Q}$ et $\mathbb{L} = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Afin de montrer que $\mathbb{L} = \mathbb{Q}(\alpha)$ avec $\alpha = \sqrt{2} + \sqrt{3}$ nous devons montrer que $\sqrt{2}$ et $\sqrt{3}$ sont des polynômes en α. △

DefZCYIbve

Définition 6.93.
Soit une extension[47] *de corps $j : \mathbb{K} \to \mathbb{L}$. Soit $A \subset \mathbb{L}$.*

ITEMooJEGUooMsDBhF

(1) Nous notons $\mathbb{K}(A)_{\mathbb{L}}$ le plus petit sous-corps de $\mathbb{L}$ contenant $j(\mathbb{K})$ et A. C'est l'intersection de tous les sous-corps de $\mathbb{L}$ contenant A et $j(\mathbb{K})$.

(2) Nous notons $\mathbb{K}[A]_{\mathbb{L}}$ le plus petit sous-anneau de $\mathbb{L}$ contenant $j(\mathbb{K})$ et A. C'est l'intersection de tous les sous-anneaux de $\mathbb{L}$ contenant A et $j(\mathbb{K})$.

*Nous disons que l'extension $\mathbb{L}$ de $\mathbb{K}$ est **monogène** ou **simple** si il existe $\theta \in \mathbb{L}$ tel que $\mathbb{L} = \mathbb{K}(\theta)$. Un tel élément θ est dit **élément primitif** de $\mathbb{L}$. Il n'est pas nécessairement unique.*

Le plus souvent, l'indice $\mathbb{L}$ dans $\mathbb{K}(A)_{\mathbb{L}}$ et $\mathbb{K}[A]_{\mathbb{L}}$ est omis parce que le contexte est clair[48], et nous avons même très souvent $\mathbb{K} \subset \mathbb{L}$ en tant qu'ensembles. Dans ce cas, l'application j est l'identité et elle sera omise.

Remarque 6.94.
Les ensembles $\mathbb{K}(A)$ et $\mathbb{K}[A]$ sont aussi appelés respectivement corps **engendré** et anneau engendré par A. Cependant il faut bien remarquer que ce sont les parties de $\mathbb{L}$ engendrées par A. Il n'est pas question à priori de parler de corps engendré par A sans dire dans quel corps plus grand nous nous plaçons.

Exemple 6.95.
Nous savons que $\mathbb{R}$ est une extension de $\mathbb{Q}$. Si $a \in \mathbb{R}$ alors $\mathbb{Q}(a)$ est le plus petit corps contenant $\mathbb{Q}$ et a. △

Exemple 6.96.
Nous avons déjà vu à l'occasion de la définition 1.359 que $A[X]$ est l'anneau de tous les polynômes de degré fini en X. Cela rentre dans le cadre de la définition 6.93 parce qu'un anneau contenant X doit contenir tous les X^n.

Notons que même si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ reste un anneau parce qu'un éventuel inverse de X n'est pas dedans[49]. Par contre, $\mathbb{K}(X)$ est un corps parce qu'il contient également les fractions rationnelles. △

ExLQhLhJ

Exemple 6.97.
Si nous prenons $\mathbb{F}_5$ et que nous l'étendons par i, nous obtenons le corps $\mathbb{K} = \mathbb{F}_5(i)$. Nous savons que tous les éléments $a \in \mathbb{F}_5$ sont racines de $X^5 - X$. Mais étant donné que $i^5 = i$, nous avons aussi $x^5 = x$ pour tout $x \in \mathbb{F}_5(i)$. Pour le prouver, utiliser le morphisme de Frobenius. Le polynôme $X^5 - X$ est donc le polynôme nul dans $\mathbb{K}$.

Ceci est un cas très particulier parce que nous avons étendu $\mathbb{F}_p$ par un élément α tel que $\alpha^p = \alpha$. En général sur $\mathbb{F}_p(\alpha)$, le polynôme $X^p - X$ n'est pas identiquement nul, et possède donc au maximum p racines. Pour $x \in \mathbb{F}_p(\alpha)$, nous avons $x^p = x$ si et seulement si $x \in \mathbb{F}_p$. △

47. Définition 6.59.
48. Et je me demande si il est possible de trouver un cas tordu où $\mathbb{K}(A)_{\mathbb{L}} \neq \mathbb{K}(A)_{\mathbb{M}}$. Par exemple lorsque A est dans $\mathbb{L}$ et $\mathbb{M}$, mais que $\mathbb{L}$ n'est pas inclus dans $\mathbb{M}$, ni $\mathbb{M}$ dans $\mathbb{L}$.
49. Lorsqu'on multiplie, les degrés montent toujours.

Dans l'énoncé suivant, la notation $R(\alpha)_{\mathbb{L}}$ signifie que l'évaluation de R sur α se fait en calculant dans le sur-corps $\mathbb{L}$ de $\mathbb{K}$. Cette proposition semble indiquer que $\mathbb{K}(\alpha)$ est donné en termes de $\mathbb{K}(X)$, lequel est défini de façon très intrinsèque sans faire appel implicitement à un sur-corps de $\mathbb{K}$.

PROPooYSFNooFGbbCi

Proposition 6.98 ([1])**.**
Soit une extension $\mathbb{L}$ du corps $\mathbb{K}$ et $\alpha \in \mathbb{L}$. Alors nous avons les isomorphismes de corps suivants :

(1) $\mathbb{K}(\alpha)_{\mathbb{L}} = \text{Frac}\big(\mathbb{K}[\alpha]_{\mathbb{L}}\big)$,

ITEMooATPTooVXKdlK

(2) $\mathbb{K}(\alpha)_{\mathbb{L}} = \{R(\alpha)_{\mathbb{L}} \text{ tel que } R \in \mathbb{K}(X)\}$.

Démonstration. Le corps $\mathbb{K}(\alpha)$ est un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}[\alpha]$ comme sous-anneau. La proposition 1.375 nous dit alors que l'application suivante est un morphisme injectif de corps :

$$\begin{aligned}\epsilon\colon \text{Frac}\big(\mathbb{K}[\alpha]\big) &\to \mathbb{K}(\alpha)\\ P/Q &\mapsto PQ^{-1}.\end{aligned} \tag{6.169}$$

Pour rappel, la notation P/Q est bien une notation pour la classe d'équivalence du couple (P,Q) pour la relation définie en 1.369.

Par ailleurs, la partie $\epsilon\Big(\text{Frac}\big(\mathbb{K}[\alpha]\big)\Big)$ de $\mathbb{L}$ et est un corps contenant $\mathbb{K}$ et α. Donc ce corps fait partie des corps sur lesquels on prend l'intersection pour définir $\mathbb{K}(\alpha)$ [50]. Cela prouve que

$$\mathbb{K}(\alpha) \subset \epsilon\Big(\text{Frac}\big(\mathbb{K}[\alpha]\big)\Big). \tag{6.170}$$

L'application ϵ est donc surjective sur $\mathbb{K}(\alpha)$. Comme elle était déjà injective, elle est bijective.

Pour la seconde partie, veuillez lire la définition 1.377 de l'évaluation d'une fraction rationnelle sur un élément de l'anneau. Si $R = P/Q \in \mathbb{K}(X)$ et si $\alpha \in \mathbb{L}$, nous avons

$$R(\alpha) = P(\alpha)Q(\alpha)^{-1}. \tag{6.171}$$

Tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α doit contenir en particulier $\{P(\alpha) \text{ tel que } P \in \mathbb{K}[X]\}$, les inverses $\{P(\alpha)^{-1} \text{ tel que } P \in \mathbb{K}[X],\ P(\alpha) \neq 0\}$ et les produits de ceux-ci. Donc tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α contient $\{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\}$.

Nous avons donc

$$\{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\} \subset \mathbb{K}(\alpha). \tag{6.172}$$

Mais puisque $\mathbb{K}(\alpha)$ est lui-même un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, il est contenu dans $\{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\}$. D'où l'égalité. □

Pourquoi cela ne contredit pas l'exemple 6.90 ? Lorsque nous écrivons

$$\mathbb{K}(\alpha) = \{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\}, \tag{6.173}$$

certes $\mathbb{K}(X)$ est défini sans faire appel à un corps contenant $\mathbb{K}$. Mais l'évaluation $R(\alpha)$, oui. Pour calculer $R(\alpha)$, il faut écrire $R = P/Q$ et calculer $P(\alpha)Q(\alpha)^{-1}$. Tous les calculs de cette dernière expression doivent se faire dans un sur-corps de $\mathbb{K}$. Il suffit que le sur-corps en question soit un monceau de mauvaise foi comme celui de l'exemple 6.90, et en réalité $\mathbb{K}(\alpha)$ peut ne pas être ce que l'on croit.

Le corolaire suivant montre que les choses s'arrangent.

Corolaire 6.99.
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}_1$ de $\mathbb{K}$, un élément $\alpha \in \mathbb{L}_1$ et une extension $\mathbb{L}_2$ de $\mathbb{L}_1$. Alors

$$\mathbb{K}(\alpha)_{\mathbb{L}_1} = \mathbb{K}(\alpha)_{\mathbb{L}_2}. \tag{6.174}$$

50. Pour rappel, la définition 6.93(1) donne $\mathbb{K}(\alpha)$ comme une intersection.

Démonstration. La proposition 6.98 nous dit que

$$\mathbb{K}(\alpha)_{\mathbb{L}_1} = \{R(\alpha)_{\mathbb{L}_1} \text{ tel que } R \in \mathbb{K}(X)\} \tag{6.175a}$$
$$\mathbb{K}(\alpha)_{\mathbb{L}_2} = \{R(\alpha)_{\mathbb{L}_2} \text{ tel que } R \in \mathbb{K}(X)\}. \tag{6.175b}$$

Mais lorsque $R \in \mathbb{K}(X)$, le calcul de $R(\alpha)$ est exactement le même dans $\mathbb{L}_1$ et dans $\mathbb{L}_2$ parce que $\mathbb{L}_2$ est un sur-corps de $\mathbb{L}_1$ et que les calculs effectifs de $R(\alpha) = P(\alpha)Q(\alpha)^{-1}$ ne font intervenir que des quantités de $\mathbb{K}$ et des puissances de α. □

Ce que ce corolaire nous dit est que si le contexte fixe une extension de $\mathbb{K}$, nous pouvons faire tous les calculs dans cette extension, même si il y a des piles d'extensions à côté.

Typiquement, à chaque fois que nous considérons des sous-corps de $\mathbb{C}$, les extensions se feront dans $\mathbb{C}$: pour tout $\alpha \in \mathbb{C}$, les corps $\mathbb{Q}(\alpha)$, $\mathbb{R}(\alpha)$ se calculent dans $\mathbb{C}$.

PROPooSYQWooFbfQtm

Proposition 6.100.
Soit un corps $\mathbb{K}$, une extension $\mathbb{L}$ et un élément $\alpha \in \mathbb{L}$. Nous considérons l'application

$$\begin{aligned}\varphi\colon \mathbb{K}[X] &\to \mathbb{L}\\ P &\mapsto P(\alpha).\end{aligned} \tag{6.176}$$

ITEMooUZDQooOasiRQ

(1) Si α est transcendant, alors $\mathbb{K}[\alpha] = \mathbb{K}[X]$ (isomorphisme d'anneaux).

(2) Si α est transcendant, alors $\mathbb{K}(\alpha)_{\mathbb{L}} = \mathbb{K}(X)$ (isomorphisme de corps),

(3) Si α est algébrique, alors $\ker(\varphi)$ est un idéal possédant un unique générateur unitaire, lequel est le polynôme minimal[51] de α sur $\mathbb{K}$.

Démonstration. Point par point.

(1) Nous savons que $\mathbb{K}[\alpha] = \{Q(\alpha) \text{ tel que } Q \in \mathbb{K}[X]\}$ (c'est la proposition 3.103). Donc φ est surjective sur $\mathbb{K}[\alpha]$, et est donc bijective. Elle est un isomorphisme[52] parce que le lemme 6.71 dit déjà que c'est un morphisme.

(2) Nous supposons encore que α est transcendant et nous considérons l'application

$$\begin{aligned}\psi\colon \mathbb{K}(X) &\to \mathbb{K}(\alpha)\\ P &\mapsto R(\alpha).\end{aligned} \tag{6.177}$$

Note : cette application n'est pas φ. En effet φ n'est définie que sur $\mathbb{K}[X]$; le corps des fractions $\mathbb{K}(X)$ est nettement plus grand (classes d'équivalence de couples).

Le fait que cette application soit surjective est la proposition 6.98(2). Pour l'injectivité nous supposons que $\psi(R) = 0$, c'est-à-dire que $R(\alpha) = 0$. Nous considérons un représentant (P, Q) de R ; c'est-à-dire $R = P/Q$. L'égalité $R(\alpha) = 0$ signifie $P(\alpha)Q(\alpha)^{-1} = 0$ (égalité dans $\mathbb{L}$). Puisque $\mathbb{L}$ est un corps, c'est un anneau intègre et nous avons la règle du produit nul ; soit $P(\alpha) = 0$, soit $Q(\alpha)^{-1} = 0$. La seconde possibilité est impossible parce que zéro n'est pas inversible. Donc $P(\alpha) = 0$. Donc $\varphi(P) = 0$ et φ étant injective, $P = 0$.

Lorsque $P = 0$, la classe P/Q est nulle dans $\mathbb{K}(X) = \text{Frac}\big(\mathbb{K}[X]\big)$.

(3) C'est le lemme-définition 6.64.

□

PropXULooPCusvE

Proposition 6.101.
Soit un corps $\mathbb{K}$ et une extension $\mathbb{L}$. Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{L}$, une racine de P. Alors le polynôme minimal d'une racine divise[53] tout polynôme annulateur.

Autrement dit, l'idéal engendré par le polynôme minimal est l'idéal des polynômes annulateurs.

51. Définition 6.64.

52. Les amateurs d'écriture inclusive ne seront, je l'espère, pas choqué par « *elle* est *un* isomorphisme » ; c'est une tournure que je propose ici sur le modèle de l'immonde « *elle* est *un* ministre » ou, à peine moins grave, « *il* est *une* sommité ».

53. Définition 3.108.

Démonstration. Nous considérons l'idéal

$$I = \{Q \in \mathbb{K}[X] \text{ tel que } Q(a) = 0\}. \tag{6.178}$$

Le fait que cela soit un idéal est simplement dû à la définition du produit : $(PQ)(a) = P(a)Q(a)$. Par le théorème 6.43, le polynôme minimal μ_a de a est dans I et, qui plus est, le génère : $I = (\mu_a)$. Par conséquent tout polynôme annulateur de a est divisé par μ_a. □

6.4.4.1 Extension algébrique, degré

Proposition 6.102.
Toute extension finie est algébrique.

Démonstration. Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ de degré[54] n de $\mathbb{K}$ et $a \in \mathbb{L}$. Nous devons montrer qu'il existe un polynôme annulateur de a à coefficients dans $\mathbb{K}$.

Soit la partie $S = \{1, a, a^2, \dots, a^n\}$ de $\mathbb{L}$. Si cette partie contient des éléments non distincts, alors c'est plié. En effet, si $a^k = a^l$, alors le polynôme X^{k-l} est un polynôme annulateur de a.

Nous supposons donc que S contienne exactement $n+1$ éléments distincts. Le lemme 4.12 nous assure que S est une partie liée : il existe des éléments $k_i \in \mathbb{K}$ tels que $\sum_{i=0}^{n} k_i a^i = 0$.

Donc le polynôme $\sum_i a_i X^i$ est un polynôme annulateur de a. □

PropURZooVtwNXE

Proposition 6.103 (Propriétés d'extensions algébriques[1])**.**
Soit $\mathbb{K}$ un corps commutatif[55] *et a un élément algébrique*[56] *sur $\mathbb{K}$, de polynôme minimal μ_a de degré n. Alors*

ItemJCMooDgEHajmi

(1) En considérant l'application d'évaluation

$$\begin{aligned} \varphi_a \colon \mathbb{K}[X] &\to \mathbb{L} \\ Q &\mapsto Q(a), \end{aligned} \tag{6.179}$$

nous avons $\mathbb{K}[a] = \text{Image}(\varphi_a)$.

ItemJCMooDgEHajiv

(2) Une base de $\mathbb{K}[a]$ comme espace vectoriel sur $\mathbb{K}$ est donnée par $\{1, a, a^2, \dots, a^{n-1}\}$.

ItemJCMooDgEHajiii

(3) Le degré de l'extension $\mathbb{K}[a]$ est égal au degré du polynôme minimal :

$$\big[\mathbb{K}[a] : \mathbb{K}\big] = n. \tag{6.180}$$

(4) L'anneau $\mathbb{K}[a]$ est l'ensemble des polynômes en a de degré jusqu'à $n-1$ à coefficient dans $\mathbb{K}$.

ItemJCMooDgEHaji

(5) $\mathbb{K}(a) = \mathbb{K}[a]$.

ItemJCMooDgEHajii

(6) Il existe un isomorphisme d'anneaux $\varphi \colon \mathbb{K}[a] \to \mathbb{K}[X]/(\mu_a)$ tel que $\varphi(k) = \bar{k}$ pour tout $k \in \mathbb{K}$. $\mathbb{K}[a] \simeq \mathbb{K}[X]/(\mu_a)$ (isomorphisme d'anneau).

L'intérêt de (6) est qu'il permet de caractériser $\mathbb{K}[a]$ sans avoir recours à un sur-corps de $\mathbb{K}$. Le point (3) indique que le degré d'une extension algébrique est égal au degré du polynôme minimal.

Démonstration. (1) Nous avons $\mathbb{K}[a] \subset \text{Image}(\varphi_a)$ parce que $\text{Image}(\varphi_a)$ est lui-même un sous-anneau de $\mathbb{L}$ contenant $\mathbb{K}$ et a. Pour rappel, $\mathbb{K}[a]$ est l'intersection de tous les tels sous-anneaux.

L'inclusion inverse est le fait que si $Q \in \mathbb{K}[X]$ alors $Q(a) \in \mathbb{K}[a]$ parce que $\mathbb{K}[a]$ est un anneau et contient donc tous les a^n.

54. Définition 6.61.
55. Juste en passant nous rappelons que tous les corps considérés ici sont commutatifs
56. Définition 6.71.

(2) La partie $\{1, a, a^2, \dots, a^{n-1}\}$ est libre parce qu'une combinaison linéaire de ces éléments est un polynôme de degré au plus $n-1$ en a. Un tel polynôme ne peut pas être nul parce que nous avons mis comme hypothèse que le polynôme minimal de a est de degré n.

Rappelons qu'en vertu de la définition 6.64, le polynôme minimal μ_a est unitaire ; donc $\deg\big(\mu_a(X) - X^n\big) \leqslant n-1$. Par conséquent en posant $S(X) = X^n - \mu_a(X)$, nous avons $\deg(S) \leqslant n-1$ et $S(a) = a^n$.

En vertu du point (1), un élément de $\mathbb{K}[a]$ s'écrit $Q(a)$ pour un certain $Q \in \mathbb{K}[X]$. Supposons que Q soit de degré $p > n-1$; alors nous le décomposons en une partie contenant les termes de degré jusqu'à $n-1$ et une partie contenant les autres :

$$Q(X) = Q_1(X) + X^n Q_2(X) \tag{6.181}$$

où $\deg(Q_1) \leqslant n-1$ et $\deg(Q_2) = p-n$. Nous évaluons cette égalité en a :

$$Q(a) = Q_1(a) + S(a)Q_2(a). \tag{6.182}$$

Donc $Q(a)$ est l'image de a par le polynôme $Q_1 + SQ_2$ qui est de degré $p-1$. Par récurrence, $Q(a)$ est l'image de a par un polynôme de degré $n-1$.

Notons que l'idée est très simple : il s'agit de remplacer récursivement tous les a^n par $S(a)$.

(3) Conséquence immédiate de (2).

(4) Conséquence immédiate de (2).

(5) Un élément général non nul de $\mathbb{K}[a]$ est de la forme $Q(a)$ avec $Q \in \mathbb{K}[X]$; il s'agit de lui trouver un inverse. Pour cela nous remarquons que les polynômes $\mu_a(X)$ et $Q(x)$ sont premiers entre eux, sinon μ_a ne serait pas un polynôme minimal (voir la proposition 6.67). Donc le théorème de Bézout 6.47 affirme l'existence d'éléments $U, V \in \mathbb{K}[X]$ tels que

$$U\mu_a + VQ = 1 \tag{6.183}$$

dans $\mathbb{K}[X]$. Nous évaluons cette égalité en a en tenant compte de $\mu_a(a) = 0$ dans $\mathbb{K}[a]$:

$$U(a)\mu_a(a) + V(a)Q(a) = 1 \tag{6.184}$$

dans $\mathbb{K}[a]$. Par conséquent $V(a)Q(a) = 1$, ce qui signifie que $V(a)$ est l'inverse de $Q(a)$.

(6) Nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(\mu_a) &\to \mathbb{K}[a] \\ \bar{R} &\mapsto R(a) \end{aligned} \tag{6.185}$$

et nous montrons qu'elle convient. Pour cela, nous nous souvenons que la proposition 6.101 nous enseigne que (μ_a), l'idéal engendré par μ_a, est égal à l'idéal des polynômes annulateurs de a dans $\mathbb{K}[X]$. Le polynôme μ_a divise tous les éléments de cet idéal ; voir aussi la définition 1.212 de l'idéal (μ_a). Cela étant mis au point, nous passons à la preuve.

(6a) **ψ est bien définie**

Si $\bar{R} = \bar{S}$ alors $R = S + Q$ avec $Q \in (\mu_a)$, et par conséquent $R(a) = S(a) + Q(a)$ avec $Q(a) = 0$.

(6b) **Surjective**

Nous savons que $\mathbb{K}[a] = \text{Image}(\varphi_a)$. Si $x \in \mathbb{K}[a]$ alors il existe $Q \in \mathbb{K}[X]$ tel que $x = Q(a)$. Dans ce cas nous avons aussi $x = \psi(\bar{Q})$.

(6c) **Injective**

Si $\psi(\bar{R}) = 0$ alors $R(a) = 0$, mais comme mentionné plus haut, μ_a engendre l'idéal des polynômes annulateurs de a. Donc $R \in (\mu_a)$ et nous avons $\bar{R} = 0$ dans $\mathbb{K}[X]/(\mu_a)$.

□

Exemple 6.104.
Un fait connu est que $\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. Donc l'inverse de $\sqrt{2}$ s'exprime bien comme un polynôme en $\sqrt{2}$ à coefficients dans $\mathbb{Q}$, ce qui confirme le point (5) de la proposition 6.103. Du point de vue de Bézout, $\mu_{\sqrt{2}}(X) = X^2 - 2$, et nous cherchons des polynômes U et V tels que

$$U(X^2 - 2) + VX = 1. \tag{6.186}$$

cette égalité est réalisée par $U = -\frac{1}{2}$ et $V = \frac{1}{2}X$. Et effectivement $V(\sqrt{2})$ est bien l'inverse de $\sqrt{2}$:

$$V(\sqrt{(2)}) = \frac{1}{2}\sqrt{2}. \tag{6.187}$$

△

PROPooNGJWooYSpwVn

Proposition 6.105 ([170]).
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ de $\mathbb{K}$ et un élément α de $\mathbb{L}$. Il y a équivalence entre les trois points suivants :

ITEMooYTEBooUuEfBz

(1) α *est algébrique sur* $\mathbb{K}$,

ITEMooWMQTooLnepQl

(2) $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$,

ITEMooAQIUooMVZojp

(3) $\mathbb{K}[\alpha]$ *est un* $\mathbb{K}$*-espace vectoriel de dimension finie.*

Si ces affirmation sont vraies, alors $[\mathbb{K}(\alpha) : \mathbb{K}]$ *est le degré du polynôme minimal de* α *sur* $\mathbb{K}$.

Démonstration. Démonstration décomposée en plusieurs implications.

(1) **(1) implique (2)**
Soit α algébrique sur $\mathbb{K}$. Nous considérons le polynôme minimal de α sur $\mathbb{K}$ (définition 6.64). Nous savons par le lemme 6.83 (qui fonctionne parce que α est algébrique) que $\mathbb{K}[\alpha] = \mathbb{K}[X]/(\mu)$ en tant qu'anneaux.
Mais $\mathbb{K}[X]$ est un anneau principal et μ en est un élément irréductible. Donc la proposition 1.244 dit que (μ) est un idéal maximum ; la proposition 1.246 avance encore un peu en disant que $\mathbb{K}[X]/(\mu)$ est un corps.
Donc $\mathbb{K}[X]/(\mu)$ est un corps isomorphe à $\mathbb{K}[\alpha]$ en tant qu'anneaux. En conséquence de quoi $\mathbb{K}[\alpha]$ est un corps.
Le corps $\mathbb{K}[\alpha]$ est un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α ; par définition nous avons donc $\mathbb{K}(\alpha) \subset \mathbb{K}[\alpha]$.
Mais d'autre part, $\mathbb{K}[\alpha]$ est contenu dans tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, donc il est inclus dans l'intersection de tout ces corps, donc $\mathbb{K}[\alpha] \subset \mathbb{K}(\alpha)$.
Nous avons donc l'égalité $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$.

(2) **(2) implique (1)**
Nous montrons que non-(1) implique non-(2). Nous disons donc que α est transcendant sur $\mathbb{K}$; cela implique par la proposition 6.100(1) que $\mathbb{K}[\alpha] = \mathbb{K}[X]$ en tant qu'anneaux. Donc $\mathbb{K}[\alpha]$ n'est pas un corps parce que $\mathbb{K}[X]$ ne l'est pas.
N'étant pas un corps, $\mathbb{K}[\alpha]$ ne peut pas être égal à $\mathbb{K}(\alpha)$ qui, lui, est un corps.

(3) **(1) implique (3)**
L'élément α est maintenant algébrique et nous considérons son polynôme minimal μ. Nous savons par le lemme 6.83 que $\mathbb{K}[\alpha] = \mathbb{K}[X]/(\mu)$ en tant qu'espaces vectoriels. Or $\mathbb{K}[X]/(\mu)$ est de dimension finie $\deg(\mu)$. Donc $\mathbb{K}[\alpha]$ est également de dimension finie.

(4) **(3) implique (1)**
Nous démontrons la contraposée. En supposant que α est transcendant nous avons $\mathbb{K}[\alpha] = \mathbb{K}[X]$ par la proposition 6.100. Or $\mathbb{K}[X]$ n'est pas de dimension finie sur $\mathbb{K}$, donc $\mathbb{K}[\alpha]$ non plus.

□

LEMooIPAXooCNGMQT

Lemme 6.106 ([171]).
Soit $\mathbb{L}$ un corps commutatif et $(\mathbb{K}_i)_{i\in I}$ une famille de sous-corps de $\mathbb{L}$. Alors $\bigcup_{i\in I}\mathbb{K}_i$ est un sous-corps de $\mathbb{L}$.

Lemme 6.107.
Soit $P\in\mathbb{K}[X]$ un polynôme unitaire irréductible de degré n. Il existe une extension $\mathbb{L}$ de $\mathbb{K}$ et $a\in\mathbb{L}$ telle que $\mathbb{L}=\mathbb{K}(a)$ et P est le polynôme minimal de a dans $\mathbb{L}$.

Démonstration. Nous prenons $\mathbb{L}=\mathbb{K}[X]/(P)$ où (P) est l'idéal dans $\mathbb{K}[X]$ généré par P. C'est un corps par le corolaire 6.44. Nous identifions $\mathbb{K}$ avec $\phi(\mathbb{K})$ où

$$\phi\colon \mathbb{K}[X]\to\mathbb{L} \tag{6.188}$$

est la projection canonique. Nous considérons également $a=\phi(X)$.

Nous avons alors $P(a)=0$ dans $\mathbb{L}$. En effet $P(a)=P\big(\phi(X)\big)$ est à voir comme l'application du polynôme P au polynôme X, le résultat étant encore un élément de $\mathbb{L}$. En l'occurrence le résultat est P qui vaut 0 dans $\mathbb{L}$.

Le polynôme P étant unitaire et irréductible, il est minimum dans $\mathbb{L}$.

Nous devons encore montrer que $\mathbb{L}=\mathbb{K}(a)$. Le fait que $\mathbb{K}(a)\subset\mathbb{L}$ est une tautologie parce qu'on calcule $\mathbb{K}(a)$ dans $\mathbb{L}$. Pour l'inclusion inverse soit $Q(X)=\sum_i Q_iX^i$ dans $\mathbb{K}[X]$. Dans $\mathbb{L}$ nous avons évidemment $Q=\sum_i Q_ia^i$. □

PropyMTEbH

Proposition 6.108 ([93]).
Soit $\mathbb{K}$, un corps et $P\in\mathbb{K}[X]$ un polynôme. Soient a et b, deux racines de P dans (éventuellement) une extension $\mathbb{L}$ de $\mathbb{K}$. Si μ_a et μ_b sont les polynômes minimaux de a et b (dans $\mathbb{K}[X]$) et si $\mu_a\neq\mu_b$, alors $\mu_a\mu_b$ divise P dans $\mathbb{K}[X]$.

Démonstration. Nous considérons les idéaux

$$I_a=\{Q\in\mathbb{K}[X]\text{ tel que }Q(a)=0\}; \tag{6.189a}$$
$$I_b=\{Q\in\mathbb{K}[X]\text{ tel que }Q(b)=0\}. \tag{6.189b}$$

Même si $Q(a)$ et $Q(b)$ sont calculés dans $\mathbb{L}$, I_a,I_b sont des idéaux de $\mathbb{K}[X]$. Le polynôme μ_a est par définition le générateur unitaire de I_a, et comme a est une racine de P, nous avons $P\in I_a$ et il existe un polynôme $Q\in\mathbb{K}[X]$ tel que

$$P=\mu_aQ. \tag{6.190}$$ EqvTPoSq

Montrons que $\mu_a(b)\neq0$. Pour cela, nous supposons que $\mu_a(b)=0$, c'est-à-dire que $\mu_a\in I_b$. Il existe alors $R\in\mathbb{K}[X]$ tel que $\mu_a=\mu_bR$. Mais par la proposition 6.67, le polynôme μ_a est irréductible, donc soit μ_b, soit R, est inversible. Comme les inversibles sont les éléments de $\mathbb{K}$ (polynômes de degré zéro), μ_b n'est pas inversible (sinon il serait constant et ne pourait pas être annulateur de b). Donc R est inversible. Disons $R=k$.

Donc $\mu_a=k\mu_b$. Mais puisque μ_a et μ_b sont unitaires, nous avons obligatoirement $k=1$. Cela donnerait $\mu_a=\mu_b$, ce qui est contraire aux hypothèses. Nous en déduisons que $\mu_a(b)\neq0$.

Étant donné que $\mu_a(b)\neq0$, l'évaluation de (6.190) en b montre que $Q(b)=0$, de telle sorte que $Q\in I_b$ et il existe un polynôme S tel que $Q=\mu_bS$, c'est-à-dire tel que $P=\mu_a\mu_bS$, ce qui signifie que $\mu_a\mu_b$ divise P. □

Exemple 6.109.
Soit $P=(X^2+1)(X^2+2)$ dans $\mathbb{R}[X]$. Dans $\mathbb{C}$ nous avons les racines $a=i$ et $b=\sqrt{2}i$ dont les polynômes minimaux sont $\mu_a=X^2+1$ et $\mu_b=X^2+2$. Nous avons effectivement $\mu_a\mu_b$ divise P dans $\mathbb{R}[X]$.

Si par contre nous considérions les racines $a=i$ et $b=-i$, nous aurions $\mu_a=\mu_b=X^2+1$, tandis que le polynôme μ_a^2 ne divise pas P. △

6.4.5 Racines de polynômes

CorDIYooEtmztc

Corolaire 6.110 (Factorisation d'une racine)**.**
Soit $P \in \mathbb{K}[X]$, un polynôme de degré n et $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$. Alors il existe un polynôme Q de degré $n-1$ tel que $P(x) = (X-\alpha)Q$.

Démonstration. Il s'agit d'un cas particulier de la proposition 6.101 : si $\alpha \in \mathbb{K}$ alors son polynôme minimal dans $\mathbb{K}$ est $X - \alpha$; donc $X - \alpha$ divise P. Il existe un polynôme Q tel que $P = (X-\alpha)Q$. Le degré est alors immédiat. □

Avant de lire l'énoncé suivant, allez relire la définition 3.101 pour savoir ce qu'est un polynôme nul.

ThoLXTooNaUAKR

Théorème 6.111 (Polynôme qui a tellement de racines qu'il s'annule)**.**
Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$ un polynôme de degré n possédant $n+1$ racines distinctes $\alpha_1, \ldots, \alpha_{n+1}$, alors $P = 0$.

Démonstration. Si P est de degré 1, il s'écrit $P = aX + b$; si il a comme racines α et β, nous avons le système

$$\begin{cases} a\alpha + b = 0 & \text{(6.191a)} \\ a\beta + b = 0. & \text{(6.191b)} \end{cases}$$

La différence entre les deux donne $a(\alpha - \beta) = 0$. Puisque $\alpha \neq \beta$, la règle du produit nul (lemme 1.198) nous donne $a = 0$. Maintenant que $a = 0$, l'annulation de b est alors immédiate.

Nous faisons maintenant la récurrence en supposant le théorème vrai pour le degré n et en considérant un polynôme P de degré $n+1$ possédant $n+2$ racines distinctes. Puisque $P(\alpha_1) = 0$, le corolaire 6.110 nous donne un polynôme Q de degré n tel que

$$P = (X - \alpha_1)Q. \tag{6.192}$$

EqQGSooNdTWfz

Étant donne que pour tout $i \neq 1$ nous avons $\alpha_i \neq \alpha_1$,

$$0 = P(\alpha_i) = \underbrace{(\alpha_i - \alpha_1)}_{\neq 0} Q(\alpha_i), \tag{6.193}$$

et la règle du produit nul donne $Q(\alpha_i) = 0$. Par conséquent le polynôme Q est de degré n et possède $n+1$ racines distinctes ; tous ses coefficients sont alors nuls par hypothèse de récurrence. Tous les coefficients du produit (6.192) sont alors également nuls. □

ExGRHooBNpjSP

Exemple 6.112.
Un polynôme à plusieurs variables peut s'annuler en une infinité de points sans être nul. Par exemple le polynôme $X^2 + Y^2 - 1 \in \mathbb{R}[X, Y]$ s'annule sur tout un cercle de $\mathbb{R}^2$ mais n'est pas nul, loin s'en faut.

Nous verrons dans la proposition 6.186 une condition pour qu'un polynôme à plusieurs variables s'annule du fait qu'il ait « trop » de racines. △

Remarque 6.113.
L'intérêt du théorème 6.111 est que si l'on prouve qu'un polynôme s'annule sur un corps infini, alors il s'annulera sur n'importe quel autre corps. Nous aurons un exemple d'utilisation de cela dans le théorème de Cayley-Hamilton 13.25.

6.4.6 Corps de rupture

DEFooVALTooDJJmJv

Définition 6.114.
*Soit $P \in \mathbb{K}[X]$ un polynôme irréductible. Une extension $\mathbb{L}$ de $\mathbb{K}$ est un **corps de rupture** pour P si il existe $a \in \mathbb{L}$ tel que $P(a) = 0$ et $\mathbb{L} = \mathbb{K}(a)$.*

NORMALooTPOIooVZAfUo

6.115.
Nous insistons sur le fait que nous ne définissons le concept de corps de rupture que pour un polynôme irréductible à coefficients dans un corps. Les deux points sont importants : irréductible et à coefficient dans un corps.

Nous discuterons brièvement le pourquoi de cela dans la section 6.4.11.

ExemGVxJUC

Exemple 6.116.
Soit $\mathbb{K} = \mathbb{Q}$ et $P = X^2 - 2$. On pose $a = \sqrt{2}$ et $\mathbb{L} = \mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$. De cette façon P est scindé dans $\mathbb{L}$:

$$P = (X - \sqrt{2})(X + \sqrt{2}). \tag{6.194}$$

Le corps $\mathbb{Q}(\sqrt{2})$ est donc un corps de rupture pour P. △

Exemple 6.117.
Dans l'exemple 6.116, nous avions un corps de rupture dans lequel le polynôme P était scindé. Il n'en est pas toujours ainsi. Prenons

$$P = X^3 - 2 \tag{6.195}$$

et $a = \sqrt[3]{2}$. Nous avons, certes, $P(a) = 0$ dans $\mathbb{Q}(\sqrt[3]{2})$, mais P n'est pas scindé parce qu'il y a deux racines complexes. △

Exemple 6.118.
Nous considérons le corps $\mathbb{Z}/p\mathbb{Z}$ où p est un nombre premier. Si $s \in \mathbb{Z}/p\mathbb{Z}$ n'est pas un carré, alors le polynôme $P = X^2 + s$ est irréductible et un corps de rupture de P sur $\mathbb{Z}/p\mathbb{Z}$ est donné par $(\mathbb{Z}/p\mathbb{Z})[X]/(X^2 + s)$, c'est-à-dire l'ensemble des polynômes de degré 1 en $\sqrt{s}$. Le cardinal en est p^2. △

Comme nous allons abondamment parler du quotient $\mathbb{K}[X]/(P)$, nous nous permettons un petit lemme.

LEMooWYYFooXYacdF

Lemme 6.119.
Soit un corps $\mathbb{K}$ et $P \in \mathbb{K}[X]$ non constant. Alors $\mathbb{K}[X]/(P)$ est un corps si et seulement si P est irréductible.

Démonstration. Nous utilisons le trio d'enfer dont il est question dans le thème 18. D'abord $\mathbb{K}[X]$ est un anneau principal par le lemme 3.110. Donc $\mathbb{K}[X]/(P)$ sera un corps si et seulement si (P) est un idéal maximum (proposition 1.220), et cela sera le cas si et seulement si (P) est engendré par un polynôme irréductible (proposition 1.244).

Il ne nous reste qu'à montrer que (P) est engendré par un irréductible si et seulement si P est irréductible. Il y a un sens dans lequel c'est évident.

Soit un irréductible μ tel que $(P) = (\mu)$. En particulier $\mu \in (P)$, c'est-à-dire qu'il existe Q tel que $\mu = PQ$. Puisque μ est irréductible, soit P, soit Q, est inversible. Si P est inversible, c'est-à-dire constant, c'est ce que nous avons exclu par hypothèse. Si par contre Q est inversible, alors $P = k\mu$ pour un certain $k \in \mathbb{K}$, ce qui montre que P est irréductible autant que μ. □

PROPooUBIIooGZQyeE

Proposition 6.120 (Existence d'un corps de rupture)**.**
Soit un corps $\mathbb{K}$ et un polynôme irréductible non constant P. Alors

(1) Le corps $\mathbb{L} = \mathbb{K}[X]/(P)$ est un corps de rupture pour P.

(2) L'élément $\bar{X}$ de $\mathbb{L}$ est une racine de P.

(3) $\mathbb{L} = \mathbb{K}(\bar{X})_{\mathbb{L}}$

Démonstration. Commençons par nous convaincre que $\mathbb{K}[X]/(P)$ est une extension de $\mathbb{K}$ (définition 6.59). Le fait que ce soit un corps est le lemme 6.119. Le morphisme $j\colon \mathbb{K} \to \mathbb{K}[X]/(P)$ est simplement $k \mapsto \bar{k}$ où à droite, $\bar{k}$ voit k dans $\mathbb{K}[X]$ comme étant le polynôme constant. Notez qu'il est automatiquement injectif (lemme 1.308).

Il faut maintenant voir que $\mathbb{K}[X]/(P) = \mathbb{K}(\alpha)$ pour un certain $\alpha \in \mathbb{K}[X]/(P)$. Grâce à notre compréhension des notations acquise dans 1.19.2.2, nous savons que $X \in \mathbb{K}[X]$ et qu'il est donc parfaitement légitime de poser $\alpha = \bar{X}$ dans $\mathbb{K}[X]/(P)$. Il s'agit simplement de l'ensemble $\bar{X} = \{X + QP \text{ tel que } Q \in \mathbb{K}[X]\}$ où X est une notation pour la suite $(0, 1, 0, 0, \ldots)$.

Bref, nous notons $\alpha = \bar{X}$ et nous démontrons que $P(\alpha) = 0$ et que $\mathbb{K}[X]/(P) = \mathbb{K}(\alpha)$ (isomorphisme de corps).

(1) $P(\bar{X}) = 0$

C'est le moment de nous souvenir comment la notation des X fonctionne, et en particulier la pirouette autour de (1.504). D'abord la définition du produit sur $\mathbb{K}[X]/(P)$ est $\bar{P}\bar{Q} = \overline{PQ}$; en particulier si $P = \sum_k a_k X^k$, alors $P(\bar{X}) = \sum_k a_k \bar{X}^k = \sum_k a_k \overline{X^k}$, et

$$P(\bar{X}) = \overline{P(X)} = \bar{P} = 0. \tag{6.196}$$

(2) **L'égalité**

Nous montrons à présent que $\mathbb{K}(\bar{X})_{\mathbb{L}} = \mathbb{L}$. C'est-à-dire que $\mathbb{L}$ est bien engendrée par $\mathbb{K}$ et un seul élément. D'abord, $\mathbb{L} = \mathbb{K}[X]/(P)$ contient bien évidemment $\mathbb{K}$ et $\bar{X}$. Ensuite nous devons prouver que tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et $\bar{X}$ est en réalité $\mathbb{L}$ entier.

Soit $Q \in \mathbb{K}[X]$, et montrons que $\bar{Q}$ est dans tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et $\bar{X}$.

Par le lemme 3.99 nous avons $\bar{Q} = Q(\bar{X})$. Et si un corps contient $\mathbb{K}$ et $\bar{X}$, il doit contenir tous les polynômes en $\bar{X}$ à coefficients dans $\mathbb{K}$. Donc un tel corps doit contenir $Q(\bar{X})$ et donc $\bar{Q}$.

□

Exemple 6.121.
Soit le polynôme $P = X^2 + 1 \in \mathbb{Z}[X]$. Dans le quotient $\mathbb{Z}[X]/(P)$ nous avons $\bar{X}^2 + 1 = 0$ et donc $\bar{X}^2 = -1$. C'est-à-dire que $\mathbb{Z}[X]/(P)$ contient un élément dont le carré est -1. Avouez que c'est bien ce à quoi nous nous attendions.

Notons que $-\bar{X}$ est également une racine de P dans $\mathbb{Z}[X]/(P)$.

En calculant dans les polynômes à coefficients dans $\mathbb{Z}(\bar{X})$ nous avons :

$$(X + \bar{X})(X - \bar{X}) = X^2 - \bar{X}^2 = X^2 + 1, \tag{6.197}$$

c'est-à-dire que P est bien factorisé, et que nous avons retrouvé la multiplication $x^2 + 1 = (x + i)(x - i)$. △

6.122.
Il n'y a évidemment pas unicité d'un corps de rupture pour un polynôme donné. Une raison est qu'un polynôme peut accepter plusieurs racines complètement indépendantes. Le corps étendu par l'une ou l'autre racine donne deux corps de rupture différents. Par exemple dans $\mathbb{Q}[X]$, le polynôme

$$P = X^4 - X^2 - 2 \tag{6.198}$$

a pour racines (dans $\mathbb{C}$) les nombres $\sqrt{2}$ et i. Donc on a deux corps de rupture complètement différents : $\mathbb{Q}(\sqrt{2})$ et $\mathbb{Q}(i)$.

6.123.
La proposition suivante donne une unicité du corps de rupture dans le cas d'un polynôme irréductible. Et nous comprenons pourquoi : un polynôme irréductible n'a fondamentalement qu'une seule racine « indépendante ». Par exemple $X^2 - 2$ a pour racines $\pm\sqrt{2}$. Autre exemple, le polynôme $X^2 + 6X + 13$ a pour racines, dans $\mathbb{C}$, les nombres complexes conjugués $z = -3 + 2i$ et $\bar{z} = -3 - 2i$.

PROPooVJACooNDmlfb

Proposition 6.124 ([164]).
Soient un corps $\mathbb{K}$ et un polynôme irréductible $P \in \mathbb{K}[X]$. Alors toute extension $\mathbb{L}$ contenant une racine α de P admet un unique morphisme de corps

$$\psi\colon \mathbb{K}[X]/(P) \to \mathbb{L} \tag{6.199}$$

tel que $\psi(\bar{X}) = \alpha$.

Dans un tel cas,

ITEMooHRFHooWLIdWU

(1) l'image de ψ est $\mathbb{K}(\alpha)_{\mathbb{L}}$,

(2) si $\mathbb{L} = \mathbb{K}(\alpha)_{\mathbb{L}}$ alors ψ est un isomorphisme.

Démonstration. L'idéal annulateur de α parmi les polynômes de $\mathbb{K}[X]$ n'est pas réduit à $\{0\}$ parce qu'il contient P. Le lemme 6.64 s'applique donc et nous avons μ, le polynôme minimal de α dans $\mathbb{K}[X]$. Il divise P qui est irréductible, donc

$$P = \lambda\mu \tag{6.200}$$

pour un certain $\lambda \in \mathbb{K}$.

Nous posons

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(P) &\to \mathbb{L} \\ \bar{Q} &\mapsto Q(\alpha). \end{aligned} \tag{6.201}$$

(1) **Bien définie** Si $\bar{Q}_1 = \bar{Q}_2$ alors il existe un $R \in \mathbb{K}[X]$ tel que $Q_1 = Q_2 + RP$. Mais alors $\psi(\bar{Q}_1) = Q_1(\alpha) = Q_2(\alpha) + R(\alpha)P(\alpha) = Q_2(\alpha)$.

(2) **Injective**
Si $\psi(\bar{Q}_1) = \psi(\bar{Q}_2)$ alors $Q_1 - Q_2 = R$ pour un certain $R \in \mathbb{K}[X]$ vérifiant $R(\alpha) = 0$. Nous avons alors un polynôme S tel que $R = S\mu = \lambda^{-1}SP$. Donc $\bar{R} = 0$ et donc $\bar{Q}_1 = \bar{Q}_2$.

(3) **Morphisme**
Laissé comme exercice ; la paresse de l'auteur de ces lignes attend vos contributions.

(4) **La condition**
Le morphisme ψ respecte de plus la condition

$$\psi(\bar{X}) = X(\alpha) = \alpha. \tag{6.202}$$

En ce qui concerne l'unicité, fixer $\psi(\bar{X})$ est suffisant pour fixer un morphisme. En effet si $\psi(\bar{X}) = \alpha$, alors

$$\psi(\bar{Q}) = \psi\Big(\sum_k a_k \bar{X}^k\Big) = \sum_k a_k \psi(\bar{X})^k = \sum_k a_k \alpha^k. \tag{6.203}$$

Pour le second point de l'énoncé, il faut remarquer que α est algébrique et non transcendant. Donc en utilisant les propositions 3.103 et 6.103(5) nous trouvons

$$\text{Image}(\psi) = \{Q(\alpha) \text{ tel que } Q \in \mathbb{K}[X]\} = \mathbb{K}[\alpha] = \mathbb{K}(\alpha). \tag{6.204}$$

Et finalement pour le dernier point, un morphisme de corps est toujours injectif. Si il est également surjectif, il sera bijectif. □

6.4.7 Pile d'extensions

LEMooTURZooXnjmjT

Lemme 6.125 ([1]).
Soient un corps $\mathbb{K}$, des extensions $\mathbb{L}_1, \ldots, \mathbb{L}_n$ et des éléments $\alpha_i \in \mathbb{L}_i$ tels que

EQooOCQSooFMkzTc

$$\mathbb{L}_1 = \mathbb{K}(\alpha_1)_{\mathbb{L}_1} \tag{6.205}$$

et

$$\mathbb{L}_k = \mathbb{L}_{k-1}(\alpha_k)_{\mathbb{L}_k}. \tag{6.206}$$

Alors

$$\mathbb{L}_n = \mathbb{K}(\alpha_1, \ldots, \alpha_n)_{\mathbb{L}_n}. \tag{6.207}$$

Démonstration. Nous démontrons par récurrence sur n. Le cas $n = 1$ est simplement l'hypothèse (6.205).

Supposons donc que le lemme soit correct pour n, et étudions le cas $n + 1$. Nous avons, par définition et par hypothèse de récurrence :

$$\mathbb{L}_{n+1} = \mathbb{L}_n(\alpha_{n+1})_{\mathbb{L}_{n+1}} = \Big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\Big)(\alpha_{n+1})_{\mathbb{L}_{n+1}}. \tag{6.208}$$

Notre tâche sera donc de montrer que

$$\Big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\Big)(\alpha_{n+1}) = \mathbb{K}(\alpha_1, \dots, \alpha_{n+1}) \tag{6.209}$$

EQooIHMGooTlPcsd

où nous n'écrivons plus les indices $\mathbb{L}_{n+1}$ partout.

Le membre de gauche est un sous-corps de $\mathbb{L}_{n+1}$ contenant à la fois $\mathbb{K}$ et tous les α_i, si bien que

$$\mathbb{K}(\alpha_1, \dots, \alpha_{n+1}) \subset \big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})_{\mathbb{L}_{n+1}}. \tag{6.210}$$

EQooLLRHooHOiLfk

Il faut donc prouver l'inclusion inverse ; c'est-à-dire montrer que tout élément x du corps $\big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})$ est forcément dans tout sous-corps de $\mathbb{L}_{n+1}$ contenant $\mathbb{K}$ et les α_i. Un tel élément x est, par la proposition 6.98(2), de la forme $r(\alpha_{n+1})$ avec $r \in \mathbb{K}(\alpha_1, \dots, \alpha_n)(X)$, c'est-à-dire

$$P(\alpha_{n+1})Q(\alpha_{n+1})^{-1} \tag{6.211}$$

avec $P, Q \in \mathbb{K}(\alpha_1, \dots, \alpha_n)[X]$.

Prouvons d'abord que si $P \in \mathbb{K}(\alpha_1, \dots, \alpha_n)[X]$, alors $P(\alpha_{n+1})$ est dans tout sous-corps de $\mathbb{L}_{n+1}$ contenant $\mathbb{K}$ et les α_i. Nous pouvons écrire $P = \sum_i a_i X^i$ avec $a_i \in \mathbb{K}(\alpha_1, \dots, \alpha_n)$, et donc

$$P(\alpha_{n+1}) = \sum_i a_i \alpha_{n+1}^i. \tag{6.212}$$

Tout corps contenant $\mathbb{K}$ et les α_1,..., α_n contient les a_i. Par produit, tout corps contenant $\mathbb{K}$, α_1,..., α_{n+1} contient les termes $a_i\alpha_{n+1}^i$, et donc $P(\alpha_{n+1})$ par somme.

De la même façon, si un corps contient $\mathbb{K}$ et les α_i, $(i = 1, \dots, n+1)$, alors il contient $Q(\alpha_{n+1})$. Comme c'est un corps, il contient aussi son inverse $Q(\alpha_{n+1})^{-1}$, et il contient aussi le produit

$$r(\alpha_{n+1}) = P(\alpha_{n+1})Q(\alpha_{n+1})^{-1}. \tag{6.213}$$

On vient ainsi de montrer que tout élément $x \in \big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})$ était dans tout sous-corps de $\mathbb{L}_{n+1}$ qui contient $\mathbb{K}$ et les α_i, $(i = 1, \dots, n+1)$; en d'autres termes :

$$\big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})_{\mathbb{L}_{n+1}} \subset \mathbb{K}(\alpha_1, \dots, \alpha_{n+1}). \tag{6.214}$$

EQooUFSMooJozpqL

Les inclusions (6.210) et (6.214) prouvent l'égalité d'ensembles (6.209) que nous voulions montrer. □

6.4.8 Clôture algébrique

Le concept de clôture algébrique a été défini dans 6.75. Voici un lemme qui dit qu'une clôture algébrique est en quelque sorte une extension algébrique maximale.

LEMooQSCGooMyCktA

Lemme 6.126 ([1]).
Soient un corps $\mathbb{K}$ et une extension algébrique $\mathbb{F}$ de $\mathbb{K}$. Nous supposons que pour toute extension algébrique de $\mathbb{L}$ nous avons $\mathbb{L} = \mathbb{F}$

Alors $\mathbb{F}$ est algébriquement clos[57].

Démonstration. Soit un polynôme $P \in \mathbb{K}[X]$. Nous voudrions prouver que P a des racines dans $\mathbb{F}$. Pour cela, nous voyons P comme un polynôme sur $\mathbb{F}[X]$ et, grace à la proposition 6.120 nous considérons un corps de rupture $\mathbb{L}$ pour P. Puisque $\mathbb{L}$ est une extension de $\mathbb{F}$, nous avons $\mathbb{L} = \mathbb{F}$. Donc $\mathbb{F}$ contient des racines de P. □

57. Définition 6.75(2).

6.127.
Nous avons défini le concept d'extension algébrique en 6.75. Nous allons en construire un petit exemple très piéton.

Souvenez vous que la proposition 1.474 nous donne l'existence et l'unicité d'un réel $\sqrt{2}$ strictement positif dont le carré est 2. Ce réel est irrationnel par la proposition 1.409.

Proposition 6.128 ([172]).
Soit $\mathbb{L} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}$.

(1) *C'est un sous-corps de* $\mathbb{R}$.
ITEMooUSOAooZoBhla

(2) *Tout sous-corps de* $\mathbb{R}$ *contenant* $\mathbb{Q}$ *et* $\sqrt{2}$ *contient* $\mathbb{L}$.

Démonstration. Nous devons d'abord prouver que $\mathbb{L}$ est un corps en vérifiant d'une part que c'est un anneau (définition 1.42) et d'autre part le fait que tous les éléments non nuls sont inversibles.

— La partie $\mathbb{L}$ de $\mathbb{R}$ est stable pour l'addition : dès que $a, b, a', b' \in \mathbb{Q}$,

$$(a + b\sqrt{2}) + (a' + b'\sqrt{2}) = (a + a') + (b + b')\sqrt{2} \in \mathbb{L}. \tag{6.215}$$

— Les neutres 0 et 1 sont dans $\mathbb{L}$.

— Si $\alpha \in \mathbb{L}$, alors $-\alpha \in \mathbb{L}$:

$$-(a + b\sqrt{2}) = -a - b\sqrt{2}. \tag{6.216}$$

— La partie $\mathbb{L}$ est stable pour le produit parce que

$$(a + b\sqrt{2})(a' + b'\sqrt{2}) = (aa' + 2bb') + (ab' + ba')\sqrt{2}. \tag{6.217}$$

— L'inverse d'un élément de $\mathbb{L}$ est dans $\mathbb{L}$. C'est le seul point pas tout à fait évident. D'abord, l'ensemble $\mathbb{R}$ est un corps par le théorème 1.417. Donc pour tout $a, b \in \mathbb{R}$, le nombre

$$\frac{1}{a + b\sqrt{2}} \tag{6.218}$$

existe dans $\mathbb{R}$.
D'abord $a - b\sqrt{2}$ n'est pas nul, parce que si il l'était, nous aurions $\sqrt{2} = a/b \in \mathbb{Q}$ alors que $\sqrt{2}$ n'est pas rationnel par la proposition 1.409. Nous pouvons donc faire le coup de multiplier le numérateur et le dénominateur par le binôme conjugué :

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}. \tag{6.219}$$

Cela est un rationnel. Donc le éléments non nuls de $\mathbb{L}$ ont un inverse qui appartient également à $\mathbb{L}$.

Nous passons à la preuve du point (2). Si $\mathbb{L}'$ est un corps qui contient $\mathbb{Q}$ et $\sqrt{2}$, il doit contenir $b\sqrt{2}$ pour tout $b \in \mathbb{Q}$ et donc aussi tous les $a + b\sqrt{2}$ avec $a, b \in \mathbb{Q}$. Par conséquent, $\mathbb{L}'$ doit contenir au moins tout $\mathbb{L}$. □

Proposition 6.129.
Soit $\mathbb{L} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}$.
ITEMooOMDMooLNhlyh

(1) *C'est un espace vectoriel de dimension* 2 *sur* $\mathbb{Q}$.
ITEMooWGGDooSbsesf

(2) *Si* $\alpha \in \mathbb{L}$, *alors il existe un polynôme* $P \in \mathbb{L}[X]$ *de degré* 2 *ou moins tel que* $P(\alpha) = 0$.
ITEMooPNNYooPtKYwQ

(3) *Le corps* $\mathbb{L}$ *est une extension algébrique de* $\mathbb{Q}$.

Démonstration. En plusieurs parties.

(1) **(1)** Pour la dimension, notez que $\{1, \sqrt{2}\}$ est une partie libre et génératrice de $\mathbb{L}$.

(2) **(2)**
Soit $\alpha \in \mathbb{L}$. La partie $\{1, \alpha, \alpha^2\}$ est de cardinal 1, 2 ou 3. Si c'est 1 ou 2, c'est que $1 = \alpha$ ou $1 = \alpha^2$ ou $\alpha = \alpha^2$. Si par exemple $1 = \alpha$ alors avec $P = X - 1$ nous avons $P(\alpha) = 0$.
Si au contraire $\{1, \alpha, \alpha^2\}$ est de cardinal 3, alors c'est une partie liée par la proposition 4.7. Il existe donc des rationnels a, b, c tels que $a + b\alpha + c\alpha^2 = 0$, c'est-à-dire $P(\alpha) = 0$ avec $P = cX^2 + bX + a$.

(3) **(3)** Nous venons de voir que tous les éléments de $\mathbb{L}$ sont des racines de polynômes de $\mathbb{Q}[X]$.

□
LEMooHWPHooZeWqns

Lemme 6.130.
Si $\mathbb{K}$ est un corps infini, alors $\mathbb{K}[X]$ est équipotent[58] *à $\mathbb{K}$.*

Démonstration. Notons provisoirement $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré n. Nous avons une surjection

$$\begin{aligned}\varphi\colon \mathbb{K}^{n+1} &\to \mathbb{K}_n[X]\\ (k_0, \dots, k_n) &\mapsto \sum_{i=0}^{n} k_i X^i.\end{aligned} \tag{6.220}$$
EQooFGZVooKIMKRA

Par récurrence sur le théorème[59] 1.162, nous avons $\mathbb{K}^{n+1} \approx \mathbb{K}$. La surjection (6.220) dit alors que

$$\mathbb{K}_n[X] \preceq \mathbb{K}^{n+1} \approx \mathbb{K}. \tag{6.221}$$

Mais puisqu'il y a une surjection $\mathbb{K} \to \mathbb{K}_n[X]$, nous avons aussi $\mathbb{K}_n[X] \succeq \mathbb{K}$. Le théorème 1.146 dit alors que $\mathbb{K}_n[X] \approx \mathbb{K}$.

Le lemme 1.163 nous permet alors de conclure que

$$\mathbb{K}[X] = \bigcup_{n=0}^{\infty} \mathbb{K}_n[X] \approx \mathbb{K}. \tag{6.222}$$

□
PROPooVPQFooScWvkS

Proposition 6.131 ([1])**.**
Soit un corps $\mathbb{K}$. Une extension algébrique de $\mathbb{K}$ est

(1) au plus dénombrable si $\mathbb{K}$ est fini,

(2) équipotente à $\mathbb{K}$ si $\mathbb{K}$ est infini.

Démonstration. En deux parties.

(1) **Si $\mathbb{K}$ est fini** Un polynôme non nul possède toujours au maximum un nombre fini de racines (éventuellement zéro) par la proposition 6.111. Par ailleurs, chaque degré de polynôme ayant seulement un nombre fini de possibilités, l'ensemble $\mathbb{K}[X]$ est au plus dénombrable (proposition 1.138).
Pour $P \in \mathbb{K}[X]$ nous avons une surjection de $\mathbb{N}$ vers l'ensemble des racines de P. Nous la notons $\varphi_P\colon \mathbb{N} \to \mathbb{L}$, en posant par exemple $\varphi_P(n) = 1$ si P n'a pas de racines. Enfin nous posons

$$\begin{aligned}\varphi\colon \mathbb{K}[X] \times \mathbb{N} &\to \mathbb{L}\\ (P, n) &\mapsto \varphi_P(n).\end{aligned} \tag{6.223}$$

C'est la fonction qui à un polynôme P et un nombre n fait correspondre la n^{e} racine de P.
Comme $\mathbb{L}$ est une extension algébrique, φ est surjective.
En termes de cardinalité, que $\mathbb{K}[X]$ soit fini ou dénombrable, dans les deux cas, $\mathbb{K}[X] \times \mathbb{N}$ est dénombrable (proposition 1.135). Il existe donc une surjection d'un ensemble dénombrable vers $\mathbb{L}$. Le lemme 1.137 conclut que $\mathbb{L}$ est fini ou dénombrable.

58. Définition 1.115.
59. J'ai quand même du mal à croire qu'il faille vraiment le lemme de Zorn pour prouver que $\mathbb{K}[X]$ est équipotent à $\mathbb{K}$. Si vous connaissez un moyen plus direct, écrivez-moi.

(2) **Si $\mathbb{K}$ est infini** Nous procédons de la même manière, mais nous devons faire appel à des résultats plus technologiques pour maitriser la cardinalité. Nous considérons à nouveau l'application

$$\begin{aligned}\varphi\colon \mathbb{K}[X]\times\mathbb{N}&\to\mathbb{L}\\(P,n)&\mapsto\varphi_P(n).\end{aligned}\tag{6.224}$$

Cette application est encore surjective : $\mathbb{L}\preceq\mathbb{K}[X]\times\mathbb{N}$. Le lemme 6.130 nous assure que $\mathbb{K}[X]\approx\mathbb{K}$ parce que $\mathbb{K}$ est infini. Ensuite la proposition 1.156 nous dit que $\mathbb{K}[X]\times\mathbb{N}\approx\mathbb{K}[X]$. Donc

$$\mathbb{K}\approx\mathbb{K}[X]\approx\mathbb{K}[X]\times\mathbb{N}\geq\mathbb{F}.\tag{6.225}$$

Mais $\mathbb{F}$ est une extension de $\mathbb{K}$. Donc il existe une injection $\mathbb{K}\to\mathbb{F}$, c'est-à-dire $\mathbb{K}\preceq\mathbb{F}$.

Ayant $\mathbb{K}\preceq\mathbb{F}\preceq\mathbb{K}$, le théorème 1.146 implique que $\mathbb{K}\approx\mathbb{F}$.

□

Lemme 6.132 ([1]).
Soient des corps $\mathbb{K}$ et $\mathbb{L}$ ainsi qu'un morphisme de corps $\rho\colon\mathbb{K}\to\mathbb{L}$. Si $P\in\mathbb{K}[X]$ a une racine dans $\mathbb{K}$, alors le polynôme $\rho(P)$ a une racine dans $\mathbb{L}$.

Démonstration. Nous notons $P=\sum_k P_kX^k$. Si $a\in\mathbb{K}$ est une racine de P, alors $\sum_k P_ka^k=0$. Nous appliquons ρ à cette égalité : $\sum_k\rho(P_k)\rho(a)^k=0$, c'est-à-dire $\rho(P)\big(\rho(a)\big)=0$. Donc $\rho(a)\in\mathbb{L}$ est une racine de $\rho(P)$. □

LEMooIIKYooHMNqYn

Lemme 6.133 ([173]).
Nous considérons un triplet $(\mathbb{K},\mathbb{L},\mathbb{F})$ où

(1) $\mathbb{K}$, $\mathbb{L}$ et $\mathbb{F}$ sont des corps

(2) il existe $a\in\mathbb{L}$ algébrique sur $\mathbb{K}$ tel que $\mathbb{L}=\mathbb{K}(a)$ et un morphisme de corps $\alpha\colon\mathbb{K}\to\mathbb{L}$.

(3) $\mathbb{F}$ est une extension algébriquement close de $\mathbb{K}$: il existe un morphisme $\beta\colon\mathbb{K}\to\mathbb{F}$.

Alors il existe un morphisme de corps $\sigma\colon\mathbb{L}\to\mathbb{F}$ tel que $\sigma|_{\alpha(\mathbb{K})}=\beta$.

Note : en pratique, les corps $\mathbb{L}$ et $\mathbb{F}$ sont le plus souvent des sur-corps de $\mathbb{K}$, de telle sorte que les applications α et β sont l'identité. En particulier, la conclusion de ce lemme s'écrit le plus souvent $\sigma|_{\mathbb{K}}=\mathrm{Id}$. Il faut juste savoir que le Frido est un névrosé des notations précises.

Démonstration. Comme $\mathbb{L}$ est monogène, si $\mu_a\in\mathbb{K}[X]$ est le polynôme minimal de $a\in\mathbb{L}$, alors les points (5) et (6) de la proposition 6.103 disent que $\mathbb{L}\simeq\mathbb{K}[a]\simeq\mathbb{K}[X]/(\mu_a)$. Pour référence ultérieure, nous considérons un isomorphisme

$$\varphi\colon\mathbb{L}\to\mathbb{K}[X]/(\mu_a).\tag{6.226}$$

Les coefficients de μ_a sont dans $\mathbb{K}$, donc nous pouvons voir $\mu_a\in\mathbb{F}[X]$. Plus précisément, si $\mu_a=\sum_k a_kX^k$, nous définissons

$$\mu'_a=\sum_k\beta(a_k)X^k\in\mathbb{F}[X].\tag{6.227}$$

Comme $\mathbb{F}$ est algébriquement clos, le polynôme μ'_a possède une racine (au moins) $b\in\mathbb{F}$: $\mu'_a(b)=0$.

Nous posons

$$\begin{aligned}\sigma'\colon\mathbb{K}[X]/(\mu_a)&\to\mathbb{F}\\\overline{\sum_k s_kX^k}&\mapsto\sum_k\beta(s_k)b^k.\end{aligned}\tag{6.228}$$

(1) **σ' est bien définie** Si $P=\sum_k s_kX^k$ et $\mu_a=\sum_k a_kX^k$ $(a_k,s_k\in\mathbb{K})$, alors

$$\sigma'(\overline{P+\mu_a})=\sigma'\big(\overline{\sum_k(a_k+s_k)X^k}\big)=\sum_k\beta(a_k)b^k+\sum_k s_kb^k=\mu'_a(b)+\sigma'(\overline{P})=\sigma'(\overline{P}).\tag{6.229}$$

(2) **σ' est un morphisme de corps** À justifier.

(3) $\varphi\big(\alpha(k)\big) = \bar{k}$ Quand nous parlons de $\bar{k}$, nous parlons de la classe du polynôme de degré zéro donné par $k \in \mathbb{K}$.

(4) **La réponse** Nous posons

$$\sigma = \sigma' \circ \varphi. \tag{6.230}$$

Pour tout $k \in \mathbb{K}$,

$$(\sigma'\varphi\alpha)(k) = \sigma'(\bar{k}) = \beta(k), \tag{6.231}$$

c'est ce qu'il fallait.

□

LEMooUULTooYcytat

Lemme 6.134 ([173]).
Soit un corps $\mathbb{K}$ muni de deux extensions $\alpha\colon \mathbb{K} \to \mathbb{L}$ et $\beta\colon \mathbb{K} \to \mathbb{F}$. Nous supposons que

(1) $\mathbb{L}$ est algébrique sur $\mathbb{K}$;

(2) $\mathbb{F}$ est algébriquement clos.

Alors il existe un morphisme de corps $\sigma\colon \mathbb{L} \to \mathbb{F}$ tel que $\sigma \circ \alpha = \beta$.

Démonstration. Nous allons utiliser le lemme de Zorn 1.23 sur l'ensemble

$$\mathcal{A} = \left\{ (\mathbb{M}, \varphi) \text{ tel que } \begin{cases} \mathbb{M} \text{ est un sous-corps de } \mathbb{L} \\ \alpha(\mathbb{K}) \subset \mathbb{M} \\ \varphi\colon \mathbb{M} \to \mathbb{F} \text{ est une extension de corps} \\ \varphi \circ \alpha = \beta \end{cases} \right\}. \tag{6.232}$$

Nous ordonnons (partiellement) cet ensemble en disant que $(\mathbb{M}_1, \varphi_1) < (\mathbb{M}_2, \varphi_2)$ si $\mathbb{M}_1 \subset \mathbb{M}_2$ et $\varphi_2|_{\mathbb{M}_1} = \varphi_1$. Il se fait que $\mathcal{A}$ est un ensemble inductif et que nous pouvons donc lui appliquer le lemme de Zorn.

Soit un élément maximal $(\mathbb{M}, \varphi)$. Nous allons montrer que $\mathbb{M} = \mathbb{L}$.

Soit $l \in \mathbb{L}$. Puisque $\mathbb{L}$ est une extension algébrique de $\mathbb{K}$, il existe un polynôme P à coefficients dans $\alpha(\mathbb{K})$ tel que $P(l) = 0$. Mais comme $\alpha(\mathbb{K}) \subset \mathbb{M}$, ce polynôme est également à coefficients dans $\mathbb{M}$. Donc l est un élément algébrique sur $\mathbb{M}$.

Nous pouvons donc considérer le triplet $\big(\mathbb{M}, \mathbb{M}(l), \mathbb{F}\big)$ qui vérifie les hypothèses du lemme 6.133. Il existe donc un morphisme de corps $\sigma\colon \mathbb{M}(l) \to \mathbb{F}$ tel que $\sigma|_{\mathbb{M}} = \varphi$.

Nous avons

$$\sigma \circ \alpha = \sigma|_{\sigma(\mathbb{K})} \circ \alpha = \sigma|_{\mathbb{M}} \circ \alpha = \varphi \circ \alpha = \beta. \tag{6.233}$$

Donc l'élément $\big(\mathbb{M}(l), \sigma\big)$ majore $(\mathbb{M}, \varphi)$ dans $\mathcal{A}$.

Par maximalité, nous déduisons que $\mathbb{M} = \mathbb{L}$. Donc le morphisme $\varphi\colon \mathbb{L} \to \mathbb{F}$ vérifie $\varphi \circ \alpha = \beta$, ce qu'il nous fallait. □

THOooEDQKooLEGlDv

Théorème 6.135 (Steinitz[174, 175]).
À propos de clôture algébrique.

(1) Tout corps possède une clôture algébrique[60].

(2) Si $\alpha_1\colon \mathbb{K} \to \mathbb{F}_1$ et $\alpha_2\colon \mathbb{K} \to \mathbb{F}_2$ sont deux clôtures algébriques du même corps $\mathbb{K}$, alors il existe un isomorphisme de corps $\varphi\colon \mathbb{F}_1 \to \mathbb{F}_2$ tel que $\varphi \circ \alpha_1 = \alpha_2$.

Démonstration. Nous commençons par l'existence, en plusieurs points.

(1) **Un ensemble** Nous considérons un ensemble Ω qui contient $\mathbb{K}$, qui est strictement surpotent[61] à $\mathbb{K}$ et qui est infini non dénombrable si $\mathbb{K}$ est fini. Par exemple $\mathcal{P}(\mathbb{K}) \cup \mathbb{K}$ si $\mathbb{K}$ est infini et $\mathbb{R} \cup \mathbb{K}$ si $\mathbb{K}$ est fini (voir le théorème de Cantor 1.149).

60. Définition 6.75.
61. Définition 1.115.

(2) **L'ensemble pour Zorn** Nous considérons l'ensemble des extensions algébriques de $\mathbb{K}$ contenues dans Ω, c'est-à-dire

$$\mathcal{A} = \left\{ (\mathbb{L}, +_{\mathbb{L}}, \times_{\mathbb{L}}) \text{ tel que } \begin{cases} \mathbb{K} \subset \mathbb{L} \subset \Omega \\ (\mathbb{L}, +_{\mathbb{L}}, \times_{\mathbb{L}}) \text{ est une extension algébrique de } (\mathbb{K}, +, \times). \end{cases} \right\} \tag{6.234}$$

Nous ordonnons $\mathcal{A}$ par l'inclusion : nous disons que

$$(\mathbb{L}_1, +_1, \times_1) < (\mathbb{L}_2, +_2, \times_2) \tag{6.235}$$

lorsque $(\mathbb{L}_2, +_2, \times_2)$ est un sur-corps de $(\mathbb{L}_1, +_1, \times_1)$ (en particulier $\mathbb{L}_1 \subset \mathbb{L}_2$).

(3) **$\mathcal{A}$ est inductif** Soit une partie $\mathcal{F} = \{(\mathbb{L}_i, +_i, \times_i)\}_{i\in I}$ de $\mathcal{A}$ que nous supposons être totalement ordonnée. Nous allons lui trouver un majorant dans $\mathcal{A}$. Nous posons $\mathbb{L} = \bigcup_{i\in I} \mathbb{L}_i$, et si $a \in \mathbb{L}_i$, $b \in \mathbb{L}_j$, alors nous définissons

$$\begin{cases} a +_{\mathbb{L}} b = a +_k b & \text{(6.236a)} \\ a \times_{\mathbb{L}} b = a \times_k b & \text{(6.236b)} \end{cases}$$

où $k \in I$ est sélectionné de telle façon à avoir $(\mathbb{L}_i, +_i, \times_i) < (\mathbb{L}_k, +_k, \times_k)$ et $(\mathbb{L}_j, +_j, \times_j) < (\mathbb{L}_k, +_k, \times_k)$. Comme tous les corps L_i sont des sous-corps les uns des autres, c'est une bonne définition.

(4) **Lemme de Zorn** Nous utilisons le lemme de Zorn 1.23. Nous notons $(\mathbb{F}, +, \times)$ un élément maximal de $\mathcal{A}$. Puisque $\mathbb{K}$ en est un sous-corps, il n'y a pas d'ambiguïté de noter $+$ et $\times$ ses opérations.

(5) **Stratégie pour la suite** Nous allons montrer que si $\mathbb{E}$ est une extension algébrique de $\mathbb{F}$, alors $\mathbb{E} = \mathbb{F}$ (le but est d'utiliser le lemme 6.126).

(6) **Un peu de cardinalité** D'abord, comme $\mathbb{F}$ est algébrique sur $\mathbb{K}$, l'ensemble $\mathbb{F}$ est équipotent à $\mathbb{K}$ si $\mathbb{K}$ est infini, et au plus dénombrable, si $\mathbb{K}$ est fini ; c'est la proposition 6.131. En bref :

— Si $\mathbb{K}$ est infini, $\mathbb{K} \approx \mathbb{F} \approx \mathbb{E} \prec \Omega$.

— Si $\mathbb{K}$ est fini, $\mathbb{K} \preceq \mathbb{F} \preceq \mathbb{E} \prec \Omega$ où $\mathbb{E}$ est au maximum dénombrable et Ω est indénombrable.

Dans tous les cas, Ω est strictement surpotent à $\mathbb{F}$, et le lemme 1.158 permet de dire

$$\mathbb{E}\backslash\mathbb{F} \preceq \mathbb{E} \prec \Omega \approx \Omega\backslash\mathbb{F}. \tag{6.237}$$

(7) **Quelques injections** Il existe donc une injection $\varphi\colon \mathbb{E}\backslash\mathbb{F} \to \Omega\backslash\mathbb{F}$. Nous posons

$$\begin{aligned} f\colon \mathbb{E} &\to \Omega \\ x &\mapsto \begin{cases} x & \text{si } x \in \mathbb{F} \\ \varphi(x) & \text{si } x \notin \mathbb{F}. \end{cases} \end{aligned} \tag{6.238}$$

Nous montrons que f est injective. Soient $x, y \in \mathbb{E}$ tels que $f(x) = f(y)$. Si $x, y \in \mathbb{F}$, alors $x = f(x) = f(y) = y$. Si $x \in \mathbb{F}$ et $y \notin \mathbb{F}$, alors $x = \varphi(y)$ alors que $x \in \mathbb{F}$ et $\varphi(y) \in \Omega\backslash F$; ce cas est impossible. Enfin si x et y sont hors de $\mathbb{F}$, alors $f(x) = \varphi(x)$ et $f(y) = \varphi(y)$; donc $\varphi(x) = \varphi(y)$ et $x = y$ par injectivité de φ.

Nous avons donc bien une injection $f\colon \mathbb{E} \to \Omega$.

(8) **La maximalité**

Nous pouvons mettre sur $f(\mathbb{E}) \subset \Omega$ la structure de corps venant de $\mathbb{E}$. Comme $f(\mathbb{F}) = \mathbb{F}$, le corps $f(\mathbb{E})$ est une extension algébrique de $\mathbb{F}$. Par maximalité, $f(\mathbb{E}) = \mathbb{F}$.

Mais si $x \in \mathbb{E}\backslash\mathbb{F}$, alors $f(x) \in \Omega\backslash\mathbb{F}$. Donc en réalité nous avons aussi $\mathbb{E} \subset \mathbb{F}$.

(9) **Conclusion** En conclusion $\mathbb{E} = \mathbb{F}$ et le lemme 6.126 termine en disant que $\mathbb{F}$ est une clôture algébrique de $\mathbb{K}$.

Nous passons à la partie « unicité » de la clôture algébrique. Étant donné que $\mathbb{F}_1$ est une extension algébrique de $\mathbb{K}$ et que $\mathbb{F}_2$ est algébriquement clos, le lemme 6.134 nous donne un morphisme de corps $\sigma : \mathbb{F}_1 \to \mathbb{F}_2$ tel que $\sigma \circ \alpha_1 = \alpha_2$.

Nous sommes donc dans la situation où $\sigma : \mathbb{F}_1 \to \mathbb{F}_2$ est une extension de corps où $\mathbb{F}_1$ est algébriquement clos et $\mathbb{F}_2$ est algébrique. Le lemme 6.80 conclut que $\sigma(\mathbb{F}_1) = \mathbb{F}_2$, c'est-à-dire que σ est surjectif. En tant que morphisme de corps, σ était déjà injective ; elle est donc bijective.

Donc $\sigma : \mathbb{F}_1 \to \mathbb{F}_2$ est un isomorphisme de corps vérifiant $\sigma \circ \alpha_1 = \alpha_2$. □

6.136.
Bien que $\mathbb{C}$ soit une extension algébriquement close de $\mathbb{Q}$, l'ensemble $\mathbb{C}$ n'est pas une clôture algébrique de $\mathbb{Q}$. C'est ce que nous montrons maintenant.

LEMooRDIZooRjWNMa

Lemme 6.137.
Le corps $\mathbb{C}$ n'est pas une clôture algébrique[62] *de $\mathbb{Q}$.*

Démonstration. Nous montrons qu'il existe des éléments de $\mathbb{C}$ qui ne sont pas des racines de polynômes à coefficients rationnels. L'ensemble $\mathbb{Q}$ est dénombrable par la proposition 1.389. L'ensemble des polynômes de degré n à coefficients dans $\mathbb{Q}$ est en bijection avec les n-uples de rationnels, c'est-à-dire avec $\mathbb{Q}^n$ qui est également dénombrable par la proposition 1.141. Enfin l'ensemble des polynômes à coefficients sur $\mathbb{Q}$ est l'union des polynômes de degré fixés, donc dénombrable par la proposition 1.138.

Jusqu'ici nous avons prouvé que l'ensemble des polynômes à coefficients rationnels était dénombrable. Or chaque polynôme possède une quantité finie de racines par le corolaire 3.131. Donc la partie de $\mathbb{C}$ constituée des nombres qui sont des racines de polynômes à coefficients dans $\mathbb{Q}$ est dénombrable. Mais $\mathbb{C}$ n'est pas dénombrable, donc possède des éléments qui ne sont pas des racines de polynômes. □

La proposition suivante est une conséquence de Steinitz 6.135, de 6.62.

PROPooZWVPooStuTAO

Proposition 6.138 (Clôture algébrique de $\mathbb{R}$[176])**.**
Le corps $\mathbb{C}$ est l'unique (à isomorphisme près) clôture algébrique de $\mathbb{R}$. Son degré est 2.

6.4.9 Polynômes à plusieurs variables

Nous avons déjà vu $A[X, Y]$ lorsque A est un anneau en la définition 3.50.

DEFooRHRKooPqLNOp

Définition 6.139.
Soit un corps $\mathbb{K}$. Le corps $\mathbb{K}(X_1, \ldots, X_n)$ est le corps des fractions de l'anneau $\mathbb{K}[X_1, \ldots, X_n]$.

DEFooOCPHooXneutp

Définition 6.140.
Soient un corps $\mathbb{K}$ et une extension $\mathbb{L}$ de $\mathbb{K}$ contenant les éléments $\alpha_1, \ldots, \alpha_n$ de $\mathbb{L}$. Nous définissons $\mathbb{K}(\alpha_1, \ldots, \alpha_n)$ comme étant l'intersection de tous les sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et les α_i.

La proposition suivante est analogue à 6.98(2).

LEMooQEJHooAmSNxU

Lemme 6.141 ([1])**.**
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ et des éléments $\alpha_1, \ldots, \alpha_n$ dans $\mathbb{L}$. Alors

$$\mathbb{K}(\alpha_1, \ldots, \alpha_n) = \{r(\alpha_1, \ldots, \alpha_n) \text{ tel que } r \in \mathbb{K}(X_1, \ldots, X_n)\}. \tag{6.239}$$

62. Clôture algébrique, définition 6.75.

Démonstration. Ce que nous avons à droite est un corps : par exemple pour l'inverse, si $r = P/Q$ alors $r(\alpha_1, \ldots, \alpha_n) = P(\alpha_1, \ldots, \alpha_n)Q(\alpha_1, \ldots, \alpha_n)^{-1}$. Cet élément a un inverse en la fraction $(Q/P)(\alpha_1, \ldots, \alpha_n)$.

Donc à droite nous avons un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ ainsi que les α_i. Donc

$$\mathbb{K}(\alpha_1, \ldots, \alpha_n) \subset \big\{r(\alpha_1, \ldots, \alpha_n) \text{ tel que } r \in \mathbb{K}(X_1, \ldots, X_n)\big\}. \tag{6.240}$$

D'autre part, tout corps contenant $\mathbb{K}$ et les α_i doit contenir tous les $P(\alpha_1, \ldots, \alpha_n)$ ($P \in \mathbb{K}[X_1, \ldots, X_n]$), leurs inverses ainsi que leurs produits ; bref doit contenir tous les $r(\alpha_1, \ldots, \alpha_n)$ avec $r \in \mathbb{K}(X_1, \ldots, X_n]$. □

6.4.10 Corps de décomposition

DEFooEKGZooSkvbum

Définition 6.142.
Soit $\mathbb{K}$ un corps commutatif et $F = (P_i)_{i\in I}$ une famille d'éléments non constants de $\mathbb{K}[X]$. Un ***corps de décomposition*** *de F est une extension $\mathbb{L}$ de $\mathbb{K}$ telle que*

(1) les P_i sont scindés sur $\mathbb{L}$,

(2) $\mathbb{L} = \mathbb{K}(R)$ où $R = \bigcup_{i\in I}\{x \in \mathbb{L} \text{ tel que } P_i(x) = 0\}$.

C'est-à-dire que $\mathbb{L}$ étend $\mathbb{K}$ par toutes les racines de tous les polynômes de F.

PROPooDPOYooFHcqkU

Proposition 6.143 ([177])**.**
Tout polynôme admet un corps de décomposition. Plus précisément, soit un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$ de degré n. Il existe un corps de décomposition $\mathbb{D}$ de la forme $\mathbb{D} = \mathbb{K}(\alpha_1, \ldots, \alpha_n)$ où les α_i sont des racines de P dans $\mathbb{D}$.

Notons que rien dans l'énoncé ne prétend que les α_i soient tous distincts, ni même que certains (voire tous) ne seraient pas dans $\mathbb{K}$.

Démonstration. Soient un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$. Si le degré de P est 0 ou 1, alors $\mathbb{K}$ est un corps de décomposition pour P. Pour le reste nous faisons une récurrence sur le degré de P.

Il y a deux possibilités, soit il existe $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$, soit non.

(1) **Si racine dans $\mathbb{K}$** Alors le corolaire 6.110 nous permet de factoriser $X - \alpha$:

$$P = (X - \alpha)Q \tag{6.241}$$

avec $\deg(Q) = \deg(P) - 1$. Dans ce cas, $\mathbb{K}$ est un corps de rupture de P.

(2) **Si pas de racines dans $\mathbb{K}$**
Nous prenons alors un corps de rupture $\mathbb{L} = \mathbb{K}(\alpha)$ avec $P(\alpha) = 0$ (c'est la proposition 6.120 qui donne l'existence d'un corps de rupture). Dans $\mathbb{L}_1$ nous avons

$$P = (X - \alpha)Q \tag{6.242}$$

avec $Q \in \mathbb{L}_1[X]$ et $\deg(Q) = \deg(P) - 1$.

(3) **Dans les deux cas**
Dans les deux cas, par hypothèse de récurrence nous avons un corps de décomposition pour Q qui se présente sous la forme

$$\mathbb{L} = \mathbb{K}(\alpha_1, \ldots, \alpha_{n-1}). \tag{6.243}$$

De plus, $\mathbb{L}$ est une extension de $\mathbb{L}_1$ parce que c'est une extension du corps sur lequel est Q.

Pour terminer la preuve nous prouvons que

$$\mathbb{D} = \mathbb{K}(\alpha_1, \ldots, \alpha_{n-1}, \alpha) \tag{6.244}$$

est un corps de décomposition de P. Vu que $\mathbb{D}$ contient $\mathbb{K}(\alpha)$ (comme cas particulier du lemme 6.141), dans $\mathbb{D}$ nous avons l'égalité $P = (X-\alpha)Q$. Et vu que $\mathbb{D}$ contient également $\mathbb{K}(\alpha_1, \dots, \alpha_{n-1})$, toujours dans $\mathbb{D}$ nous avons aussi

$$Q = (X - \alpha_1) \dots (X - \alpha_{n-1}). \tag{6.245}$$

Donc nous avons dans $\mathbb{D}$ l'égalité

$$P = (X - \alpha)(X - \alpha_1) \dots (X - \alpha_{n-1}). \tag{6.246}$$

□

LEMooJNGWooTXdGre

Lemme 6.144 ([1]).
Soit un polynôme $P \in \mathbb{K}[X]$, et un corps $\mathbb{L}$ dans lequel P est scindé. Si $P = P_1 \dots P_r$ est la décomposition de P en irréductibles dans $\mathbb{K}$, alors chacun des P_i est scindé dans $\mathbb{L}$.

Démonstration. Juste pour le mentionner, le fait que P ait une décomposition en irréductibles est le fait que $\mathbb{K}[X]$ soit factoriel, c'est-à-dire la proposition 6.36.

Le polynôme P est scindé dans $\mathbb{L}$, c'est-à-dire que, en notant n le degré de P,

$$P = \prod_{i=1}^{n} (X - \lambda_i) \tag{6.247}$$

EQooNFXSooUkWeQu

avec $\lambda_i \in \mathbb{L}$.

Soit $\mathbb{L}_1$, une extension de $\mathbb{L}$ dans laquelle P_1 est scindé. Ensuite, $\mathbb{L}_2$ une extension de $\mathbb{L}_1$ dans laquelle P_2 est scindé et ainsi de suite. Nous avons construit $\mathbb{L}_r$, une extension de $\mathbb{L}$ dans laquelle tous les P_i sont scindés ainsi que P lui-même. Dans ce corps nous avons l'égalité

$$P = \prod_{k=1}^{n} (X - \mu_k) \tag{6.248}$$

EQooBEFUooBnqSUS

où les μ_k sont des éléments des diverses extensions $\mathbb{L}_i$, et sont les racines des P_i. En tout cas, tous sont dans $\mathbb{L}_r$.

Les deux décompositions (6.247) et (6.248) sont des décompositions dans $\mathbb{L}_r[X]$ du polynôme P. Vu que ce dernier est factoriel, en réalité les deux décompositions sont identiques (se souvenir de la définition 3.64), et nous avons $\mu_k \in \mathbb{L}$ pour tout k. Toutes les extensions $\mathbb{L}_i$ sont en réalité triviales, et nous avons $\mathbb{L}_r = \mathbb{L}$.

Bref, tout cela pour dire que les P_i ont toutes leurs racines dans $\mathbb{L}$. □

THOooQVKWooZAAYxK

Théorème 6.145 ([164]).
Soient :

(1) un isomorphisme de corps $\tau : \mathbb{K} \to \mathbb{K}'$;

(2) un polynôme non constant $P \in \mathbb{K}[X]$ de degré n ;

(3) un corps de décomposition $\mathbb{L}$ de P sur $\mathbb{K}$;

(4) un corps de décomposition $\mathbb{L}'$ de $\tau(P)$ sur $\mathbb{K}'$;

Alors τ se prolonge en un isomorphisme $\sigma : \mathbb{L} \to \mathbb{L}'$.

Démonstration. Soit m le nombre de racines de P dans $\mathbb{L}\backslash\mathbb{K}$. Nous faisons une récurrence sur m.

Si $m = 0$ alors $\mathbb{K}$ est un corps de rupture de P ; nous avons

$$P = (X - \lambda_1) \dots (X - \lambda_n) \tag{6.249}$$

avec $\lambda_i \in \mathbb{K}$. Dans ce cas nous avons aussi

$$\tau(P) = \big(X - \tau(\lambda_1)\big) \dots \big(X - \tau(\lambda_n)\big) \tag{6.250}$$

avec $\tau(\lambda_i) \in \mathbb{K}'$. Nous avons donc $\mathbb{L}' = \mathbb{K}'$ et prendre $\sigma = \tau$ fonctionne.

Nous supposons à présent que $m > 0$. Plus précisément nous considérons un polynôme possédant exactement m racines dans $\mathbb{L}\backslash\mathbb{K}$. Soit

$$P = P_1 \dots P_r \tag{6.251}$$

sa décomposition en irréductibles dans $\mathbb{K}[X]$ (notons que $r \leqslant n$ parce que chacun des facteurs est de degré au moins 1). Au moins un des P_i est de degré plus grand ou égal à 2. Nous savons de la proposition 6.36 que $\mathbb{K}[X]$ est un anneau factoriel. Le lemme 6.144 nous assure que les polynômes P_i sont également scindés dans $\mathbb{L}$. Et l'unicité de la décomposition fait en sorte que les racines des P_i sont celles de P.

Soit $\alpha \in \mathbb{L}$, une racine de P_1. Vu que P_1 est irréductible sur $\mathbb{K}$, l'application suivante est un isomorphisme de corps par le lemme 6.83 :

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(P_1) &\to \mathbb{K}[\alpha] \\ \bar{Q} &\mapsto Q(\alpha). \end{aligned} \tag{6.252}$$

Notons que le lemme parle du quotient par le polynôme minimal, mais ici nous avons un irréductible. Un polynôme annulateur irréductible est multiple du polynôme minimal, et l'idéal engendré est le même.

Nous avons aussi la décomposition

$$\tau(P) = \tau(P_1) \dots \tau(P_r), \tag{6.253}$$

et chacun des $\tau(P_i)$ a ses racines dans $\mathbb{L}'$. Soit β, une racine de $\tau(P_1)$ dans $\mathbb{L}'$. Alors nous avons l'isomorphisme

$$\psi'\colon \mathbb{K}'[X]/\big(\tau(P_1)\big) \to \mathbb{K}'[\beta]. \tag{6.254}$$

De plus, par le lemme 6.46, nous savons que τ passe aux classes :

$$\phi_\tau\colon \mathbb{K}[X]/(P_1) \to \mathbb{K}'[X]/\big(\tau(P_1)\big) \tag{6.255}$$

est un isomorphisme d'anneaux. Et enfin, dernier résultat externe à invoquer, la proposition 6.103 nous assure que $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$ et $\mathbb{K}'[\beta] = \mathbb{K}'(\beta)$. Posons pour l'occasion $\mathbb{K}_1 = \mathbb{K}(\alpha)$ et $\mathbb{K}'_1 = \mathbb{K}'(\beta)$.

Nous avons l'enchainement suivant d'isomorphismes de corps [63] :

$$\tau_1 = \psi' \circ \phi_\tau \circ \psi^{-1}\colon \mathbb{K}_1 \to \mathbb{K}[X]/(P_1) \to \mathbb{K}'[X]/\big(\tau(P_1)\big) \to \mathbb{K}'_1. \tag{6.256}$$

Cet isomorphisme $\tau_1\colon \mathbb{K}_1 \to \mathbb{K}'_1$ prolonge τ. Si vous aimez les diagrammes, en voici un sur lequel les i représentent des inclusions et où τ et τ_1 sont des isomorphismes

$$\begin{array}{ccccc} \mathbb{K} & \xrightarrow{i} & \mathbb{K}_1 & \xrightarrow{i} & \mathbb{L} \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau_1} & & \\ \mathbb{K}' & \xrightarrow[i]{} & \mathbb{K}'_1 & \xrightarrow{i} & \mathbb{L}' \end{array} \tag{6.257}$$

Le corps $\mathbb{L}$ est un corps de décomposition de P sur $\mathbb{K}_1$, et le nombre de racines de P dans $\mathbb{L}\backslash\mathbb{K}_1$ est strictement plus petit que m parce qu'il y en a exactement m dans $\mathbb{L}\backslash\mathbb{K}$ et que $\mathbb{K}_1$ en a au moins une de plus que $\mathbb{K}$. Même raisonnement pour $\mathbb{K}'$, $\mathbb{K}'_1$ et $\mathbb{L}'$.

Résumons la situation :

— $\tau_1\colon \mathbb{K}_1 \to \mathbb{K}'_1$ est un isomorphisme de corps ;

— $P \in \mathbb{K}_1[X]$ est un polynôme non constant ;

— $\mathbb{L}$ est un corps de décomposition de P sur $\mathbb{K}_1$;

63. En réalité il est plus exact de dire « isomorphisme d'anneaux », parce que la structure de corps n'est en réalité aucune nouvelle structure par rapport à l'anneau.

— $\mathbb{L}'$ est un corps de décomposition de P sur $\mathbb{K}'_1$;

— le nombre de racines de P dans $\mathbb{L}\backslash\mathbb{K}_1$ est strictement inférieur à m.

Donc, par hypothèse de récurrence sur m, il existe un isomorphisme $\sigma\colon \mathbb{L} \to \mathbb{L}'$ qui prolonge τ_1. Vu que τ_1 prolonge τ, nous avons également σ qui prolonge τ. □

L'énoncé le plus compact pour l'unicité du corps de décomposition (à isomorphisme près) est le suivant :

PropTMkfyM

Proposition 6.146.
Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$. Soient $\mathbb{L}$ et $\mathbb{F}$ deux corps de décomposition de P. Alors il existe un isomorphisme $f\colon \mathbb{L} \to \mathbb{F}$ tel que $f|_{\mathbb{K}} = \mathrm{Id}$.

Démonstration. C'est un cas particulier du théorème 6.145, où nous considérons $\mathbb{K} = \mathbb{K}'$ muni de l'isomorphisme identité. □

Cependant le passage par l'énoncé plus compliqué 6.145 est nécessaire pour les besoins de la récurrence.

6.147.
À propos de terminologie. Lorsque nous disons « *un* corps de décomposition » nous référons à la définition 6.142 et il n'y a pas vraiment unicité. Si nous disons « *le* corps de décomposition » nous référons en général à celui construit dans la proposition 6.143 qui n'est en réalité même pas très explicite.

Quoi qu'il en soit, nous nous permettons de dire « le » corps de décomposition lorsque nous parlons de propriétés invariantes par isomorphisme.

6.148.
La construction du corps de décomposition d'un polynôme fonctionne en prenant successivement le corps de rupture des facteurs irréductibles. Nous insistons sur le fait que cette opération se fait sur chaque facteur irréductible séparément.

L'exemple suivant montre dans quel ordre se passent les choses.

Exemple 6.149.
Soit le polynôme $P = X^4 - 5X^2 + 6$. Sa factorisation en irréductibles est :

$$P = (X^2 - 2)(X^2 - 3). \tag{6.258}$$

Ce polynôme n'est pas irréductible sur $\mathbb{Q}$ et il ne s'agit donc pas de prendre brutalement un corps de rupture pour P. Il s'agit de poser $P = P_1P_2$ avec

$$P_1 = X^2 - 2 \tag{6.259a}$$
$$P_2 = X^2 - 3, \tag{6.259b}$$

de remarquer que P_1 et P_2 sont irréductibles sur $\mathbb{Q}$ et de chercher des corps de rupture pour eux. On commence par P_1. Nous avons le corps de rupture $\mathbb{L}_1 = \mathbb{Q}(\sqrt{2})$ et la factorisation

$$P_1 = (X + \sqrt{2})(X - \sqrt{2}). \tag{6.260}$$

Ensuite nous considérons P_2 dans $\mathbb{L}_1[X]$. Ce P_2 est encore irréductible. Nous lui cherchons un corps de rupture, et c'est $\mathbb{L}_2 = \mathbb{L}_1(\sqrt{3})$ dans lequel nous avons

$$P_2 = (X - \sqrt{3})(X + \sqrt{3}). \tag{6.261}$$

Nous savons (par le lemme 6.125) que

$$\mathbb{L}_2 = \mathbb{L}_1(\sqrt{3}) = \big(\mathbb{Q}(\sqrt{2})\big)(\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}). \tag{6.262}$$

Nous pouvons donc écrire en toute confiance, dans $\mathbb{Q}(\sqrt{2},\sqrt{3})$ la factorisation

$$P=(X+\sqrt{2})(X-\sqrt{2})(X+\sqrt{3})(X-\sqrt{3}). \tag{6.263}$$

Et nous notons que si nous avions commencé par P_2 au lieu de P_1, nous aurions eu le même résultat final. △

CORooELAUooPQGLkR

Corolaire 6.150 ([1]).
Le corps de décomposition d'un polynôme est une extension finie.

Démonstration. Soient un corps $\mathbb{K}$, un polynôme $P\in\mathbb{K}[X]$ et un corps de décomposition $\mathbb{D}$ de P de la forme $\mathbb{D}=\mathbb{K}(\alpha_1,\dots,\alpha_n)$ où les α_i sont les racines de P dans $\mathbb{D}$. Cela existe par la proposition 6.143.

Vu que le lemme 6.125 donne

$$\mathbb{K}(\alpha_1,\dots,\alpha_n)=\big(\mathbb{K}(\alpha_1,\dots,\alpha_{n-1})\big)(\alpha_n), \tag{6.264}$$

le corps $\mathbb{D}$ se construit comme une pile d'extensions finies. Les degrés se composant par le lemme 6.63, au final ce corps de décomposition est une extension finie.

Soit maintenant un corps de décomposition quelconque $\mathbb{L}$. La proposition 6.146 donne un isomorphisme de corps[64] $f:\mathbb{L}\to\mathbb{D}$ tel que f soit l'identité sur $\mathbb{K}$.

Si $\{v_i\}_{i\in I}$ est une base de $\mathbb{D}$ comme espace vectoriel sur $\mathbb{K}$, êtes-vous prêts à parier que $\{f(v_i)\}_{i\in I}$ est une base de $\mathbb{L}$ comme espace vectoriel sur $\mathbb{K}$[65] ? □

6.4.11 Non irréductible ou pas corps ?

SUBSECooEDMJooTXBfOu

Nous avons déjà mentionné que nous ne définissons le corps de rupture d'un polynôme que dans le cas de polynôme irréductible à coefficients dans un corps.

D'abord si P n'est pas irréductible, la question de chercher un corps de rupture n'a pas beaucoup de sens.

Si A est un anneau intègre et si P est un polynôme irréductible sur A, nous pouvons considérer le corps des fractions de A, dire $P\in\mathrm{Frac}(A)[X]$ et continuer. Étendre la définition de corps de rupture de cette façon aux polynômes à coefficients dans un anneau intègre n'est pas une grande révolution.

Au lieu de cela, nous pouvons nous demander dans quel cas nous aurions que $A[X]/(P)$ est directement un corps.

Exemple 6.151.
Soit le polynôme constant $P=2$ dans $\mathbb{Z}[X]$. Il y est irréductible parce qu'il ne peut pas être écrit comme produit de deux non inversibles. Ce polynôme a deux propriétés ennuyeuses :

— Il n'est plus irréductible sur $\mathbb{Q}$,

— Il n'existe aucun corps contenant une racine de P tout en contenant $\mathbb{Z}$ comme sous-anneau.

△

6.4.12 Clôture algébrique

THOooQFWWooMWXEhT

Théorème 6.152.
Tout corps $\mathbb{K}$ possède une clôture algébrique[66] *Ω. De plus si $\mathbb{L}$ est une extension de $\mathbb{K}$, alors $\mathbb{L}$ est $\mathbb{K}$-isomorphe à un sous corps de Ω.*

Les deux parties de ce théorème utilisent l'axiome du choix.

Notons en particulier que si Ω' est une autre clôture algébrique de $\mathbb{K}$, alors Ω et Ω' sont des sous corps l'un de l'autre et sont donc $\mathbb{K}$-isomorphes.

64. Un isomorphisme de corps est juste un isomorphisme d'anneaux.
65. Personnellement, je n'ai pas vérifié. Vérifiez vous-même et écrivez-moi pour dire si c'est bon ou non.
66. Définition 6.75.

Lemme 6.153.
Les polynômes $P, Q \in \mathbb{K}[X]$ ne sont pas premiers entre eux si et seulement s'ils ont une racine commune dans la clôture algébrique Ω de $\mathbb{K}$.

Démonstration. Soit A un polynôme non inversible divisant P et Q. Par définition de Ω, ce polynôme A a une racine dans Ω qui est alors une racine commune à P et Q dans Ω.

Pour le sens inverse, si α est une racine commune de P et Q, alors le polynôme $X - \alpha$ divise P et Q et donc P et Q ne sont pas premiers entre eux. □

ExfUqQXQ

Exemple 6.154.
Soit p un nombre premier. Montrons que le polynôme

$$Q(X) = X^p - X + 1 \tag{6.265}$$

est irréductible dans $\mathbb{F}_p$.

Nous supposons qu'il n'est pas irréductible, c'est-à-dire que

$$Q(X) = R(X)S(X) \tag{6.266}$$

avec R et S, des polynômes de degrés $\geqslant 1$ dans $\mathbb{F}_p[X]$

Soit $\bar{\mathbb{F}}_p$ une clôture algébrique[67] de $\mathbb{F}_p$ et $\alpha \in \bar{\mathbb{F}}_p$ tel que $R(\alpha) = 0$. Pour tout $a \in \mathbb{F}_p$, nous avons

$$\begin{aligned} Q(\alpha + a) &= (\alpha + a)^p - (\alpha + a) + 1 && \text{(6.267a)}\\ &= \alpha^p + a^p - \alpha - a + 1 && \text{(6.267b)}\\ &= \alpha^p - \alpha + 1 && \text{(6.267c)}\\ &= Q(\alpha) && \text{(6.267d)}\\ &= 0 && \text{(6.267e)} \end{aligned}$$

où nous avons utilisé le fait que $a^p = a$ et que α était une racine de Q. Ce que nous venons de prouver est que l'ensemble des racines de Q dans $\bar{\mathbb{F}}_p$ est donné par $\{\alpha + a \text{ tel que } a \in \mathbb{F}_p\}$.

Les polynômes R et S sont donc formés de produits de termes $X - (\alpha + a)$ avec $a \in \mathbb{F}_p$. L'un des deux –disons R pour fixer les idées– doit bien en avoir plus que 1. Nous avons alors

$$R(X) = \prod_{i=1}^{k} \big(X - (\alpha + a_i)\big) \tag{6.268}$$

où les a_i sont les éléments de $\mathbb{F}_p$. En développant un peu,

$$R(X) = X^k - \sum_{i=1}^{k} (\alpha + a_i) X^{k-1} + \text{termes de degré plus bas en } X. \tag{6.269}$$

Le coefficient devant X^{k-1} n'est autre que $k\alpha + \sum_i a_i$. Étant donné que $k \neq 0$ et que $R \in \mathbb{F}_p[X]$, nous devons avoir $\alpha \in \mathbb{F}_p$. Par conséquent nous avons $\alpha^p = \alpha$ et une contradiction :

$$Q(\alpha) = \alpha^p - \alpha + 1 = 1 \neq 0. \tag{6.270}$$

Le polynôme $X^p - X + 1$ est donc irréductible sur $\mathbb{F}_p$. △

6.4.13 Dérivée de polynômes

DEFooICHTooAggARW

Définition 6.155 ([178]).
Soit un anneau A ainsi que $a_i \in A$. Pour le polynôme $P = \sum_{k=0}^{n} a_k X^k$, nous définissons le ***polynôme dérivé***

$$P' = \sum_{k=1}^{n} k a_k X^{k-1}. \tag{6.271}$$

67. Définition 6.75. Pour l'existence c'est le théorème 6.152.

LEMooEFQQooTzDanH

Lemme 6.156 (Règle de Leibniz[178]).
Soit un anneau A. Nous considérons l'application dérivée[68]

$$\begin{aligned} D\colon A[X] &\to A[X] \\ P &\mapsto P'. \end{aligned} \tag{6.272}$$

(1) L'application D est A-linéaire.

(2) Elle vérifie $(PQ)' = PQ' + P'Q$.

LEMooFNFYooAflPRT

Lemme 6.157 ([178]).
Soit un corps $\mathbb{K}$. Nous considérons l'application de dérivation[69] $D\colon \mathbb{K}[X] \to \mathbb{K}[X]$.

(1) Si $\mathbb{K}$ est de caractéristique nulle, alors $\ker(D) = \mathbb{K}$.

(2) Si $\mathbb{K}$ est de caractéristique $p \neq 0$, alors $\ker(D) = \mathbb{K}[X^p]$.

Ici, $\mathbb{K}[X^p]$ est un gros abus de notations pour dire

$$\mathbb{K}[X^p] = \{Q(X^p) \text{ tel que } Q \in \mathbb{K}[X]\}. \tag{6.273}$$

6.4.14 Extensions séparables

Notons que dans ce qui va suivre nous allons parler de $\mathbb{K}[X]$, l'ensemble des polynômes sur un corps. Cela ne s'applique donc pas à $\mathbb{Z}[X]$ par exemple.

Une des choses intéressantes avec les extensions séparables c'est qu'elles vérifient le théorème de l'élément primitif 6.172.

LEMooGUUJooDSrQFU

Lemme 6.158 ([178]).
Soit un corps $\mathbb{K}$. Soient $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. Nous avons équivalence entre : ITEMooHMMMooWKzSLA

(1) a est une racine multiple[70] *de P.* ITEMooZXCTooPctrMm

(2) $(X-a) \mid P$ et $(X-a) \mid P'$. ITEMooBMSSooNOYfKn

(3) $P(a) = P'(a) = 0$

Démonstration. Soit n la multiplicité de la racine a de P. Éventuellement $n = 0$ si a n'est en réalité pas une racine. Nous avons

$$P = (X-a)^n Q \tag{6.274}$$ EQooRGYSooSkATBx

avec $Q(a) \neq 0$.

En ce qui concerne la dérivée, si $n = 0$ nous avons $P' = Q'$, et si $n \geqslant 1$ nous avons

$$P' = n(X-a)^{n-1}Q + (X-a)^n Q'. \tag{6.275}$$ EQooQLINooPKUENN

Nous faisons maintenant la preuve.

(1) **(1) implique (2)** Nous supposons que a est une racine multiple, c'est-à-dire $n \geqslant 2$. Les relations (6.274) et (6.275) nous montrent que $(X-a)$ divise aussi bien P que P'.

(2) **(2) implique (3)** Celle-ci est facile. Les hypothèses sont qu'il existe des polynômes A et B tels que $P = (X-a)A$ et $P' = (X-a)B$. Nous avons alors bien $P(a) = P'(a) = 0$.

(3) **(3) implique (1)** Nous avons $P = (X-a)^n Q$ avec $Q(a) \neq 0$. Par hypothèse $P(a) = 0$. Donc $n \geqslant 1$. Nous pouvons alors utiliser la formule (6.275) pour la dérivée :

$$P' = n(X-a)^{n-1}Q(a) + (X-a)^n Q'. \tag{6.276}$$

Nous avons donc $P'(a) = n0^{n-1}Q(a)$ avec $Q(a) \neq 0$. Par hypothèse $P'(a) = 0$ et donc $n - 1 \geqslant 0$, ce qui signifie $n \geqslant 2$ et donc que a est une racine multiple de P.

68. Définition 6.155.
69. Voir lemme 6.156.
70. Racine multiple, définition 3.124.

□

Proposition 6.159.
Soit P irréductible dans $\mathbb{K}[X]$ ayant des racines distinctes dans le corps de décomposition $\mathbb{L}$. Si $\mathbb{L}'$ est un autre corps de décomposition pour P, alors P a aussi ses racines distinctes dans $\mathbb{L}'$.

Démonstration. L'ingrédient est la proposition 6.146 qui donne l'unicité du corps de décomposition à $\mathbb{K}$-isomorphisme près. Soit donc $\psi\colon \mathbb{L} \to \mathbb{L}'$ un isomorphisme laissant invariant les éléments de $\mathbb{K}$. D'une part, étant donné que P est à coefficients dans $\mathbb{K}$, nous avons $\psi(P) = P$. D'autre part dans $\mathbb{L}$ le polynôme P s'écrit

$$P = a(X - \alpha_1)\dots(X - \alpha_n) \tag{6.277}$$

avec $a \in \mathbb{K}$ et $\alpha_i \in \mathbb{L}$. Nous avons donc

$$P = \psi(P) = a(X - \psi(\alpha_1))\dots(X - \psi(\alpha_n)). \tag{6.278}$$

Donc les racines de P dans $\mathbb{L}'$ sont les éléments $\psi(\alpha_i)$ qui sont distincts. □

Exemple 6.160.
Un polynôme peut être séparable sur un corps, mais non séparable sur un autre. Soit $\mathbb{L} = \mathbb{F}_p(T)$ et $\mathbb{K} = \mathbb{F}_p(T^p)$. Nous considérons le polynôme

$$P = X^p - T^p \tag{6.279}$$

dans $\mathbb{K}[X]$. Par le morphisme de Frobenius nous avons

$$P = (X - T)^p \tag{6.280}$$

dans $\mathbb{L}[X]$. Le polynôme P est irréductible sur $\mathbb{K}[X]$ parce que ses diviseurs sont de la forme $(X - T)^k$ qui contiennent T^k qui n'est pas dans $\mathbb{K}$ (sauf si $k = n$ ou $k = 0$).

Ce polynôme n'est pas séparable sur $\mathbb{K}$ parce que dans le corps de décomposition $\mathbb{L}$, la racine T est multiple. Notons bien le raisonnement : P étant irréductible, pour savoir si il est séparable, on le regarde dans un corps de décomposition.

Par contre si nous regardons P dans $\mathbb{L}[X]$ alors P n'est plus irréductible parce que ses facteurs irréductibles sont $(X - T)$. N'étant pas irréductible, nous regardons les racines de *ses facteurs irréductibles*. Or chacun des facteurs irréductibles étant $X - T$, les racines sont simples. △

Exemple 6.161.
Le polynôme $X^3 - 1$ est séparable sur $\mathbb{Q}$ parce que ses facteurs irréductibles dans $\mathbb{Q}[X]$ sont $X - 1$ et $X^2 + X + 1$, et ces deux polynômes ont des racines simples (dans $\mathbb{Q}(i)$). △

Exemple 6.162.
Le polynôme $(X^2 + 1)^2$ est séparable dans $\mathbb{Q}[X]$. En effet, il a pour facteurs irréductibles le polynôme $X^2 + 1$ (en deux exemplaires), et ce polynôme a pour racines $\pm i$ dans l'extension $\mathbb{Q}(i)$, racines qui sont simples pour ce polynôme. △

PropolyeZff

Proposition 6.163 ([179])**.**
Soit $P \in \mathbb{K}[X]$ un polynôme non constant. Les propriétés suivantes sont équivalentes.

ItemdqPFUi

(1) Il existe une extension de $\mathbb{K}$ dans laquelle P a une racine multiple.

ItemdqPFUib

(2) P a une racine multiple dans tout corps de décomposition .

ItemdqPFUii

(3) P et P' ont une racine commune dans une extension de $\mathbb{K}$.

ItemdqPFUiii

(4) le degré de $\operatorname{pgcd}(P, P')$ *est* $\geqslant 1$.

Démonstration. En plusieurs parties.

(1) **(1)⇒(2)** Soit a, une racine multiple de P dans une extension $\mathbb{L}$ de $\mathbb{K}$, et $\mathbb{E}$, un corps de décomposition de P. Alors nous voulons prouver que P ait une racine multiple dans $\mathbb{E}$.
Nous pouvons voir $P \in \mathbb{L}[X]$, et construire un corps de décomposition $\mathbb{E}'$ qui est une extension de $\mathbb{L}$. Vu que $\mathbb{E}$ et $\mathbb{E}'$ sont deux corps de décomposition de P nous avons un isomorphisme $\psi\colon \mathbb{E}' \to \mathbb{E}$. Si $a \in \mathbb{E}$ est une racine multiple de P, alors $\psi(a)$ est une racine multiple de P dans $\mathbb{E}'$ parce que

$$P\big(\psi(a)\big) = \psi\big(P(a)\big). \tag{6.281}$$

(2) **(2)⇒(3)** Soit $\mathbb{L}$ un corps de décomposition de P sur $\mathbb{K}$ et $a \in \mathbb{L}$, une racine multiple de P. On a alors $P = (X-a)^2Q$ avec $Q \in \mathbb{L}[X]$. En dérivant,

$$P' = 2(X-a)Q + (X-a)^2Q', \tag{6.282}$$

et donc a est également une racine de P'.

(3) **(3)⇒(4)** Soit D un pgcd de P et P'. D'après le théorème de Bézout il existe $A, B \in \mathbb{K}[X]$ tels que

$$AP + BP' = D. \tag{6.283}$$

Si a est une racine commune de P et P' dans une extension $\mathbb{L}$, alors c'est aussi une racine de D et donc $\deg(D) \geqslant 1$.

(4) **(4)⇒(1)** Si le degré de D est plus grand ou égal à 1, alors nous considérons une racine a de D dans $\mathbb{L}$ (une extension de $\mathbb{K}$). Étant donné que D divise P et P', l'élément a est une racine commune de P et P'. Nous montrons maintenant que a est alors une racine multiple de P. Vu que $P(a) = 0$ nous avons

$$P = (X-a)Q, \tag{6.284}$$

et $P' = Q + (X-a)Q'$. Mais alors $P'(a) = Q(a)$ et donc $Q(a) = 0$ et donc a est une racine double de P. Par conséquent a est une racine multiple de P dans $\mathbb{K}$.

□

Notons que si P est irréductible, cette proposition donne des conditions pour que P ne soit pas séparable.

LEMooVWBPooQITlUL

Lemme 6.164 ([167]).
Soient un corps $\mathbb{K}$ ainsi qu'un polynôme irréductible non nul $P \in \mathbb{K}[X]$. Soit un polynôme $Q \in \mathbb{K}[X]$. Alors P divise Q si et seulement si $\operatorname{pgcd}(P,Q) \neq 1$.

Démonstration. Nous notons $D = \operatorname{pgcd}(P,Q)$. C'est un polynôme qui divise P. De plus P est irréductible, donc non inversible. Vu que P est non nul, nous avons $\deg(P) \geqslant 1$.

(1) $\Rightarrow$ Nous supposons que P divise Q. Alors P est un diviseur commun de P et Q. Le lemme 6.55 (qui n'est en fait qu'une répétition de la définition du pgcd) nous dit alors que P divise $\operatorname{pgcd}(P,Q)$. Nous avons donc

$$0 < \deg(P) \leqslant \deg\big(\operatorname{pgcd}(P,Q)\big). \tag{6.285}$$

Donc le polynôme $\operatorname{pgcd}(P,Q)$ est un polynôme non constant et n'est en particulier pas 1.

(2) $\Leftarrow$ Le polynôme D est un diviseur commun de P et Q. En particulier $P = SD$ pour un certain polynôme S. Étant donné que P est irréductible, soit S soit D (ou les deux) est inversible. Par hypothèse $D \neq 1$, et comme D est unitaire, nous savons que D n'est pas un inversible. Bref, S est un inversible, c'est-à-dire que $S = k \in \mathbb{K}$. Nous notons que

$$P = kD. \tag{6.286}$$

EQooQESZooDHgkWn

Le polynôme D divise également Q. Il existe $T \in \mathbb{K}[X]$ tel que $Q = TD$. En remplaçant D par sa valeur déduite de (6.286) nous avons

$$Q = TD = k^{-1}TP, \tag{6.287}$$

ce qui signifie que P divise Q.

□

Nous rappelons à la très aimable lectrice que $\gcd(P,Q)$ est l'unique polynôme unitaire parmi les PGCD de P et Q. La définition des PGCD est 1.186 et l'unicité de l'unitaire est dans le lemme 6.52(2).

PROPooPRGQooOkEmag

Proposition-Définition 6.165 ([178, 180])**.**
Soient un corps $\mathbb{K}$, un polynôme non constant $P \in \mathbb{K}[X]$ ainsi qu'une extension algébriquement close Ω et un corps de décomposition[71] *$\mathbb{M}$ pour P. Alors les propriétés suivantes sont équivalentes*[72] *:*

ITEMooVMJLooUdZLQR

(1) $\gcd(P,P')=1$.

(2) P et P' n'ont de racines communes dans aucune extension de $\mathbb{K}$.

(3) P et P' n'ont pas de racines communes dans $\mathbb{M}$.

(4) P et P' n'ont pas de racines communes dans Ω.

ITEMooARPLooBoUrxX

(5) P n'a de racines multiples dans aucune extension de $\mathbb{K}$.

(6) P n'a que des racines simples dans $\mathbb{M}$.

(7) P n'a que des racines simples dans Ω.

(8) Il existe une extension $\mathbb{L}$ de $\mathbb{K}$ ainsi que $\beta, \alpha_i \in \mathbb{L}$ tels que

$$P = \beta \prod_{i=1}^{n} (X - \alpha_i) \tag{6.288}$$

où les α_i sont distincts.

(9) Il existe $\beta, \alpha_i \in \mathbb{M}$ tels que $P = \beta \prod_{i=1}^{n}(X-\alpha_i)$ où les α_i sont distincts.

(10) Il existe $\beta, \alpha_i \in \Omega$ tels que $P = \beta \prod_{i=1}^{n}(X-\alpha_i)$ où les α_i sont distincts.

Un polynôme irréductible qui vérifie ces conditions est dit ***séparable****.*

Si P est un polynôme non constant dont la décomposition en irréductibles est $P = P_1 \dots P_r$, nous disons qu'il est ***séparable*** *si tous les P_i le sont.*

PROPooMOZFooDGemBJ

Proposition 6.166 ([178, 180])**.**
Soit un corps $\mathbb{K}$ et un polynôme irréductible non nul $P \in \mathbb{K}[X]$. Les conditions suivantes sont équivalentes :

ITEMooDOTTooTZGhSq

(1) P est séparable[73]

ITEMooRTDCooJGEtum

(2) $P' \neq 0$

ITEMooZOMZooDPrNaC

(3) Il existe une extension $\mathbb{L}$ de $\mathbb{K}$ dans laquelle P a une racine simple.

Démonstration. En plusieurs parties.

(1) **(3) implique (2)** Soient une extension $\mathbb{L}$ de $\mathbb{K}$ ainsi qu'une racine simple $a \in \mathbb{L}$. Nous utilisons la proposition 3.128 : il existe un polynôme $Q \in \mathbb{L}[X]$ tel que $P = (X-a)Q(X)$ avec $\mathbb{L} \neq 0$.

Nous utilisons la règle de Leibniz (lemme 6.156) : $P' = (X-a)'Q(X) + (X-a)Q'(X) = Q(X) + (X-a)Q'(X)$. En évaluant en a,

$$P'(a) = Q(a) \neq 0. \tag{6.289}$$

71. Tout polynôme admet un corps de décomposition, proposition 6.143.
72. Pour lee polynôme dérivé P', définition 6.155.
73. Polynôme séparable, définition 6.165.

(2) **(2) implique** $\text{pgcd}(P, P') = 1$ Notons $D = \text{pgcd}(P, P')$. Vu que D divise P, il existe un polynôme Q tel que $P = QD$. Le polynôme P étant irréductible, soit Q soit D est inversible. Supposons que Q est inversible. Alors $P = kD$ avec $k \in \mathbb{K}$. En particulier $\deg(P) = \deg(D)$. Mais comme D est un diviseur de P', nous avons

$$\deg(D) \leqslant \deg(P') < \deg(P). \tag{6.290}$$

Nous avons une contradiction.

Nous en déduisons que dans l'égalité $P = DQ$, c'est D qui est inversible. Donc $D \in \mathbb{K}$. Vu que D est unitaire[74], nous avons $D = 1$.

(3) **(2) implique (1)** Nous avons déjà vu que dans le cas (2) nous avons $\text{pgcd}(P, P') = 1$. La proposition-définition 6.165(1) nous indique qu'alors P est séparable.

(4) **(1) implique (3)** La proposition 6.143 nous permet de considérer un corps de décomposition de P. Dans ce corps, P admet des racines (autant que son degré, mais c'est une autre histoire), et ces racines sont simples par la définition 6.165(5).

□

CorUjfJSE

Corolaire 6.167.
Si $\mathbb{K}$ est de caractéristique nulle, alors tout polynôme de $\mathbb{K}[X]$ est séparable.

Démonstration. Il suffit de montrer que les irréductibles sont séparables. Soit P un polynôme irréductible et unitaire de degré d. Le terme de plus haut degré de P' est alors dX^{d-1} qui est non nul parce que $d \neq 0$ en caractéristique nulle. Donc $P' \neq 0$ et donc P est séparable par la proposition 6.163. □

DEFooKTVHooTydOTM

Définition 6.168.
Soit $\mathbb{L}$ une extension algébrique de $\mathbb{K}$.

ITEMooOFYPooLYkIPr

*(1) On dit que l'élément $a \in \mathbb{L}$ est **séparable** sur $\mathbb{K}$ si son polynôme minimal dans $\mathbb{K}[X]$ est séparable sur $\mathbb{K}$ (définition 6.165).*

*(2) L'extension $\mathbb{L}$ est **séparable** si tous ses éléments sont séparables.*

PropUmxJVw

Proposition 6.169.
Soit $\mathbb{K}$ un corps. Les conditions suivantes sont équivalentes :

ITEMooUSKRooDmsGmw

(1) toutes les extensions algébriques de $\mathbb{K}$ sont séparables ;

ITEMooJGWLooKInxSG

(2) tout polynôme irréductible de $\mathbb{K}[X]$ est séparable.

En particulier les extensions algébriques des corps de caractéristique nulle sont toutes séparables.

Démonstration. En plusieurs parties.

(1) **(1) implique (2)**
Soit un polynôme irréductible P de $\mathbb{K}[X]$, et un corps de décomposition $\mathbb{L}$ de P. Cela est une extension algébrique par le corolaire 6.150. Elle est donc séparable par hypothèse.

Voilà une première chose de dite.

Maintenant, nous voudrions montrer que P est un polynôme séparable. Dans $\mathbb{L}$ nous avons

$$P = \prod_{i=1}^{n}(X - a_i), \tag{6.291}$$

et tout le défi est de prouver que les a_i sont tous distincts.

Soient deux racines $a, b \in \mathbb{L}$ de P. Nous considérons les polynômes minimaux μ_a et μ_b dans $\mathbb{K}[X]$. Ces deux polynômes divisent P parce que P est à la fois dans l'idéal annulateur de a

74. Le pgcd de deux polynômes est l'unitaire parmi tous les pgcd, voir 6.53.

et de b. Mais comme P est irréductible, il existe $k_a, k_n \in \mathbb{K}$ tels que $P = k_a\mu_a$ et $P = k_b\mu_b$. Donc les polynômes μ_a, μ_b et P sont multiples les uns des autres. Vu que μ_a et μ_n sont unitaires, $\mu_a = \mu_b$.

Nous avons :

$$P = k\mu = \prod_{i=1}^{n}(X - a_i). \tag{6.292}$$

Or le polynôme μ est irréductible par la proposition 6.67(1), et l'extension $\mathbb{L}$ est séparable, donc μ n'a que des racines simples, Donc tous les a_i sont distincts.

(2) **(2) implique (1)**

Soit une extension algébrique $\mathbb{L}$ de $\mathbb{K}$. Soit $a \in \mathbb{L}$. Nous devons prouver que le polynôme minimal de a dans $\mathbb{K}$ est séparable, c'est-à-dire qu'il n'a que des racines simples.

Le polynôme minimal $\mu_a \in \mathbb{K}[X]$ de a est irréductible et donc séparable. Donc $\mathbb{L}$ est séparable.

La dernière phrase est une conséquence du corolaire 6.167. □

CORooNZZMooIoBYXY

Corolaire 6.170.
Toute les extensions algébriques de $\mathbb{Q}$ sont séparables.

Démonstration. Le corps $\mathbb{Q}$ est de caractéristique nulle (définition 1.350). Le corolaire 6.167 dit alors que tout polynôme sur $\mathbb{Q}$ est séparable. La proposition 6.169 conclut en disant que toutes les extensions algébriques de $\mathbb{Q}$ sont séparables. □

ThobkwCMm

Théorème 6.171 ([100]).
Soit $\mathbb{K}$ un corps (pas spécialement fini). Tout sous-groupe fini de $\mathbb{K}^*$ est cyclique.

Démonstration. Soit G un sous-groupe fini de $\mathbb{K}^*$ et ω son exposant (qui est le PPCM des ordres des éléments de G). Étant donné que $|G|$ est divisé par tous les ordres, il est divisé par le PPCM des ordres. Bref, nous avons

$$x^\omega = 1 \tag{6.293}$$

pour tout $x \in G$. Mais ce polynôme possède au plus ω racines dans $\mathbb{K}$. Du coup $|G| \leqslant \omega$. Et comme on avait déjà vu que $\omega \mid |G|$, on a $\omega = |G|$. Il suffit plus que trouver un élément d'ordre effectivement ω. Cela est fait par le lemme 3.32. □

ThoORxgBC

Théorème 6.172 (Théorème de l'élément primitif[100]).
Toute extension de corps séparable finie admet un élément primitif[75].

Autrement dit, soient des éléments algébriques $\alpha_1, \dots, \alpha_n$ séparables[76] *sur $\mathbb{K}$, et soit l'extension engendrée $\mathbb{L} = \mathbb{K}(\alpha_1, \dots, \alpha_n)$. Alors $\mathbb{L}$ admet un élément primitif, c'est-à-dire un élément θ tel que $\mathbb{L} = \mathbb{K}(\theta)$.*

Démonstration. Si le corps $\mathbb{K}$ est fini, alors $\mathbb{L}$ est également fini. Donc $\mathbb{L}^*$ est cyclique par le théorème 6.171. Si θ est un générateur de $\mathbb{L}^*$, alors $\mathbb{L} = \mathbb{K}(\theta)$.

Passons au cas où $\mathbb{K}$ est infini. Il suffit d'examiner le cas $n = 2$; en effet pour $n = 1$ c'est trivial et si $n > 2$, alors

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \mathbb{K}(\alpha_1, \dots, \alpha_{n-1})(\alpha_n), \tag{6.294}$$

et donc si $\mathbb{K}(\alpha_1, \dots, \alpha_{n-1}) = \mathbb{K}(\theta)$, nous avons

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \mathbb{K}(\theta, \alpha_n) \tag{6.295}$$

et nous sommes réduit au cas $n = 2$ par récurrence.

75. Définition 6.93.
76. Définition 6.168(1).

Soit donc $\mathbb{L} = \mathbb{K}(\alpha, \beta)$; soit P le polynôme minimal de α sur $\mathbb{K}$ et Q celui de β. Nous nommons $\mathbb{E}$, un corps de décomposition de PQ. Nous avons $\mathbb{L} \subset \mathbb{E}$. Vu que P et Q sont polynômes minimaux d'éléments qui sont par hypothèse séparables, les polynômes P et Q sont séparables. Donc dans $\mathbb{E}$ les racines de P sont distinctes parce que P est irréductible (et idem pour Q). Soient les racines

$$\alpha_1 = \alpha, \alpha_2, \dots, \alpha_r \tag{6.296}$$

de P dans $\mathbb{E}$ et les racines

$$\beta_1 = \beta, \beta_2, \dots, \beta_s \tag{6.297}$$

de Q dans $\mathbb{E}$. Ici r et s sont les degrés de P et Q.

Si $s = 1$ alors $Q = X - \beta$ et donc $\beta \in \mathbb{K}$ (parce que $Q \in \mathbb{K}[X]$). Du coup nous avons $\mathbb{L} = \mathbb{K}(\alpha)$ et le théorème est démontré. Nous supposons donc maintenant que $s \geqslant 2$.

Pour chaque $(i, j) \in [\![1, r]\!] \times [\![2, s]\!]$, l'équation $\alpha_i + x\beta_k = \alpha_1 + x\beta_1$ pour $x \in \mathbb{K}$ a au plus[77] une solution donnée le cas échéant par

$$x = (\alpha_i - \alpha_1)(\beta_1 - \beta_k)^{-1} \tag{6.298}$$

Notons que cela est de toutes façons dans $\mathbb{L}$ et qu'étant donné que $\beta_1 \neq \beta_k$, cette solution a un sens (ici on utilise l'hypothèse de séparabilité). Étant donné que $\mathbb{K}$ est infini nous pouvons donc trouver un $c \in \mathbb{K}$ qui ne résout aucune des équations (6.298) :

$$\alpha_i + c\beta_k \neq \alpha_1 + c\beta_1. \tag{6.299}$$

Nous posons $\theta = \alpha_1 + c\beta_1$ et nous prétendons que $\mathbb{L} = \mathbb{K}(\theta)$.

Pour cela, commençons par montrer que $\beta_1 \in \mathbb{K}(\theta)$. On considère, dans $\mathbb{K}(\theta)[T]$, les polynômes $Q(T)$ et $S(T) = P(\theta - cT)$, et on nomme R le PGCD de ces deux polynômes. Alors, une racine de R doit être une racine de Q, et est donc un des β_i. Or, d'une part, le choix de θ fait que β_1 est une racine de R parce que

$$S(\beta_1) = P(\theta - c\beta_1) = P(\alpha_1 + c\beta_1 - c\beta_1) = P(\alpha_1) = 0. \tag{6.300}$$

D'autre part, si $k \geqslant 2$, alors

$$S(\beta_k) = P(\alpha_1 + c\beta_1 - c\beta_k) = P\big(\alpha_1 + c(\beta_1 - \beta_k)\big) \neq 0 \tag{6.301}$$

parce que $\alpha_1 + c(\beta_1 - \beta_k)$ ne vaut ni α_1 (le second terme est non-nul), ni un autre α_i (à cause de (6.299)).

Il s'ensuit que Q et S n'ont qu'une racine commune $\beta_1 = \beta$, qui est donc l'unique racine de R. Ainsi,

$$R = X - \beta \in \mathbb{K}(\theta)[T], \tag{6.302}$$

et donc $\beta \in \mathbb{K}(\theta)$.

Dès lors $\alpha = \alpha_1 = \theta - c\beta$ est alors immédiatement dans $\mathbb{K}(\theta)$; puisque les deux éléments α et β sont dans $\mathbb{K}(\theta)$, nous avons obtenu $\mathbb{L} = \mathbb{K}(\alpha, \beta) = \mathbb{K}(\theta)$.

□

Exemple 6.173.
Le théorème de l'élément primitif 6.172 ne tient pas pour les corps non commutatifs. Considérons par exemple le corps $\mathbb{K}$ des quaternions et le groupe à 8 éléments $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Ce dernier groupe n'est pas cyclique alors qu'il est un groupe fini dans $\mathbb{K}^*$. △

Exemple 6.174.
Il est aussi possible pour un groupe fini d'avoir $\omega(G) = |G|$ sans pour autant que G soit cyclique. Par exemple pour $G = S_3$, nous avons $|S_3| = 6$ alors que les éléments de S_3 sont soit d'ordre 2 soit d'ordre 3 et $\omega(G) = \text{ppcm}(2, 3) = 6$. Pourtant S_3 n'est pas cyclique. △

77. La solution (6.298) peut être dans $\mathbb{L}$ et non dans $\mathbb{K}$. L'équation peut donc très bien ne pas avoir de solutions $x \in \mathbb{K}$.

6.5 Idéal maximum

6.5.1 Idéal maximum

DefWHDdTrC

Définition 6.175.
*Une $\mathbb{K}$-algèbre est de **type fini** si elle est le quotient de $\mathbb{K}[X_1, \dots, X_n]$ par un idéal (pour un certain n).*

ThonoZyKa

Théorème 6.176 ([181]).
Soit $\mathbb{K}$ un corps et B, une $\mathbb{K}$-algèbre de type fini. Si B est un corps, alors c'est une extension algébrique finie de $\mathbb{K}$.

ThowgZYqx

Théorème 6.177 ([181]).
Si $\mathbb{K}$ est un corps algébriquement clos, les idéaux maximaux[78] *de $\mathbb{K}[X_1, \dots, X_n]$ sont de la forme*

$$(X_1 - a_1, \dots, X_n - a_n) \tag{6.303}$$

où les a_i sont des éléments de $\mathbb{K}$.

Démonstration. Nous commençons par montrer que

$$J = (X_1 - a_1, \dots, X_n - a_n) \tag{6.304}$$

est un idéal maximum. Pour cela nous considérons le morphisme surjectif d'anneaux

$$\begin{aligned} \phi\colon \mathbb{K}[X_1, \dots, X_n] &\to \mathbb{K} \\ P &\mapsto P(a_1, \dots, a_n). \end{aligned} \tag{6.305}$$

Soit $P \in \ker(\phi)$; nous écrivons la division euclidienne de P par $X - a_1$ puis celle du reste par $X - a_2$ et ainsi de suite :

$$P = (X - a_1)Q_1 + \cdots + (X_n - a_n)Q_n + R \tag{6.306}$$

EqDAkjH

où R doit être une constante parce que le premier reste est de degré zéro en X_1, le second est de degré zéro en X_1 et X_2, etc. Afin d'identifier cette constante, nous appliquons l'égalité (6.306) à $(a_1, \dots, a_n)$ et en nous rappelant que $P \in \ker(\phi)$ nous obtenons

$$0 = P(a_1, \dots, a_n) = R, \tag{6.307}$$

donc $R = 0$ et $P = (X_1 - a_1)Q_1 + \cdots + (X_n - a_n)Q_n$, c'est-à-dire $P \in J$. Nous avons donc $\ker(\phi) \subset J$. Par ailleurs $J \subset \ker(\phi)$ est évident, donc $J = \ker(\phi)$.

Vu que J est le noyau de l'application $\mathbb{K}[X_1, \dots, X_n] \to \mathbb{K}$, nous avons

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{J} = \mathbb{K}. \tag{6.308}$$

Donc J est un idéal maximal parce que tout polynôme n'étant pas dans J doit avoir un terme indépendant non nul et donc être dans $\mathbb{K}$ vis à vis du quotient $\mathbb{K}[X_1, \dots, X_n]/J$.

Nous montrons maintenant l'implication inverse. Nous supposons que I est un idéal maximum et nous montrons qu'il doit être égal à J (pour un certain choix de $a_1, \dots, a_n$).

Le quotient

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} \tag{6.309}$$

est une $\mathbb{K}$-algèbre de type fini (définition 6.175). De plus c'est un corps par la proposition 1.220. C'est donc une extension algébrique finie de $\mathbb{K}$ par le théorème 6.176. Mais $\mathbb{K}$ étant algébriquement clos, il est sa propre et unique extension algébrique ; nous en déduisons que

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} = \mathbb{K}. \tag{6.310}$$

78. Définition 1.219.

Donc pour tout $1 \leqslant i \leqslant n$, il existe $a_i \in \mathbb{K}$ tel que $X_i - a_i \in I$, sinon le monôme X_i ne se projetterait pas sur un élément dans $\mathbb{K}$ dans le quotient. Cela prouve que J est contenu dans I ; par maximalité nous avons donc $I = J$. □

Corolaire 6.178.
Soit $\mathbb{K}$ un corps algébriquement clos et I, un idéal de $\mathbb{K}[X_1, \ldots, X_n]$. Si nous notons

$$V(I) = \{x \in \mathbb{K}^n \text{ tel que } P(x_1, \ldots, x_n) = 0, \forall P \in I\} \tag{6.311}$$

l'ensemble des racines communes à tous les éléments de I, on a $V(I) = \varnothing$ si et seulement si $I = \mathbb{K}[X_1, \ldots, X_n]$.

Démonstration. Si $I = \mathbb{K}[X_1, \ldots, X_n]$ en particulier $1 \in I$ et nous avons évidemment $V(I) = \varnothing$. Le sens difficile est l'autre sens.

Supposons que $I \neq \mathbb{K}[X_1, \ldots, X_n]$ et que K est un idéal maximum contenant I (ça existe par le théorème de Krull 1.221). Nous savons déjà par le théorème 6.177 que K est de la forme $K = (X_1 - a_1, \ldots, X_n - a_n)$. Un élément de I est dans K, donc si $P \in I$ nous avons

$$P(a_1, \ldots, a_n) = 0, \tag{6.312}$$

c'est-à-dire que $(a_1, \ldots, a_n) \in V(I)$ et donc que $V(I) \neq \varnothing$. □

6.6 Polynômes symétriques et alternés

6.6.1 Polynômes symétriques, alternés ou semi-symétriques

Nous rappelons que le groupe symétrique S_n agit sur l'anneau des polynômes de n variables sur l'anneau A par le lemme 1.368.

Définition 6.179.
Un polynôme Q en n indéterminées est

(1) ***symétrique*** *si $Q = \sigma \cdot Q$ pour tout $\sigma \in S_n$;*
(2) ***alterné*** *si $\sigma \cdot Q = \epsilon(\sigma)Q$ pour tout $\sigma \in S_n$;*
(3) ***semi-symétrique*** *si $\sigma \cdot Q = Q$ pour tout $\sigma \in A_n$*

Le polynôme $T_1 + T_2$ est symétrique ; le polynôme $T_1 + T_2^2$ ne l'est pas.

6.6.2 Polynôme symétrique élémentaire

DEFooTREUooZKoXeg

Définition 6.180.
Le k-ième ***polynôme symétrique élémentaire*** *à n inconnues est le polynôme*

$$\sigma_k(T_1, \ldots, T_n) = \sum_{s \in F_k} \prod_{i=1}^{k} T_{s(i)} \tag{6.313}$$

où F_k est l'ensemble des fonctions strictement croissantes $\{1, 2, \ldots, k\} \to \{1, 2, \ldots, n\}$.

Une autre façon de décrire ces polynômes élémentaires est

$$\sigma_k = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} X_{i_1} \ldots X_{i_k}. \tag{6.314}$$

Par exemple

$$\sigma_1(T_1, \ldots, T_n) = T_1 + T_2 + \cdots + T_n \tag{6.315a}$$
$$\sigma_2(T_1, \ldots, T_n) = T_1T_2 + \cdots + T_1T_n + T_2T_3 + \cdots + T_2T_n + \cdots + T_{n-1}T_n \tag{6.315b}$$
$$\sigma_n(T_1, \ldots, T_n) = T_1 \ldots T_n. \tag{6.315c}$$

En particulier, $\sigma_2(x, y, z) = xy + yz + xz$.

TholReBiw

Théorème 6.181 ([182]).
Si Q est un polynôme symétrique en $T_1, \dots, T_n$, alors il existe un et un seul polynôme P en n indéterminées tel que

$$Q(T_1, \dots, T_n) = P\big(\sigma_1(T_1, \dots, T_n), \dots, \sigma_n(T_1, \dots, T_n)\big). \tag{6.316}$$

Exemple 6.182.
Nous voulons décomposer $P(x, y, z) = x^3 + y^3 + z^3$ en polynômes symétriques élémentaires, c'est-à-dire en

$$\begin{cases} \sigma_1 = x + y + z & \text{(6.317a)} \\ \sigma_2 = xy + xz + yz & \text{(6.317b)} \\ \sigma_3 = xyz. & \text{(6.317c)} \end{cases}$$

Étant donné que P est de degré 3, les seules combinaisons des σ_i qui peuvent intervenir sont σ_1^3, $\sigma_1\sigma_2$ et σ_3. Étant donné que dans P le coefficient de x^3 est un, il est obligatoire d'avoir un coefficient 1 devant σ_1^3. Nous le calculons :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y,z')
(x, y, z)
sage: P=x**3+y**3+z**3
sage: S1=x+y+z
sage: S2=x*y+x*z+y*z
sage: S3=x*y*z
sage: (S1**3).expand()
x^3 + 3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2 + y^3
                + 3*y^2*z + 3*y*z^2 + z^3
sage: (S1**3-P).expand()
3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2 + 3*y^2*z + 3*y*z^2
x^3 + 3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2
            + y^3 + 3*y^2*z + 3*y*z^2 + z^3
```

Dans la différence $\sigma_1^3 - P$ nous voyons que le terme en xyz est $6xyz$; par conséquent nous savons que le coefficient de σ_3 sera -6. Il nous reste :

```
sage: (S1**3+6*S3-P).expand()
3*x^2*y + 3*x^2*z + 3*x*y^2 + 12*x*y*z + 3*x*z^2 + 3*y^2*z + 3*y*z^2
```

que nous identifions facilement avec $3\sigma_1\sigma_2$. Nous avons donc

$$P = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \tag{6.318}$$

△

LemSoXCQH

Lemme 6.183 ([114]).
Soit $\mathbb{K}$ une extension de degré δ de $\mathbb{Q}$ et $P \in \mathbb{K}[T_1, \dots, T_m]$. Alors il existe $\bar{P} \in \mathbb{Q}[T_1, \dots, T_m]$ tel que

(1) $\deg \bar{P} = \delta \deg(P)$

(2) pour tout $(z_1, \dots, z_m) \in \mathbb{C}^m$ tel que $P(z_1, \dots, z_m) = 0$, on a $\bar{P}(z_1, \dots, z_m) = 0$.

Démonstration. En vertu de la proposition 6.169 et du corolaire 6.170, $\mathbb{K}$ est une extension séparable de $\mathbb{Q}$, et donc vérifie le théorème de l'élément primitif (6.172). Il existe $\theta \in \mathbb{K}$ tel que $\mathbb{K} = \mathbb{Q}(\theta)$. Soit $P_\theta \in \mathbb{Q}[X]$ le polynôme minimal de θ. L'extension $\mathbb{K}$ étant de degré δ, et θ étant un générateur, une base de $\mathbb{K}$ comme espace vectoriel sur $\mathbb{Q}$ est

$$\{1, \theta, \dots, \theta^{\delta-1}\}. \tag{6.319}$$

Mais par ailleurs la proposition 6.103(2) nous indique qu'une base de $\mathbb{Q}(\theta)$ sur $\mathbb{Q}$ est donnée par

$$\{1, \theta, \dots, \theta^{n-1}\} \tag{6.320}$$

où n est le degré de P_θ. Donc P_θ est de degré δ. Nous nommons $\theta_1, \dots, \theta_\delta$ les racines de P_θ dans un corps de décomposition. Ici nous notons $\theta = \theta_1$ et nous ne prétendons pas que $\theta_k \in \mathbb{K}$. Notons que ces θ_i sont toutes des racines simples de P_θ, sinon nous aurions un facteur irréductible $(X - \theta_k)^2$, et P_θ ne serait pas irréductible sur $\mathbb{Q}$.

Soit σ_k le morphisme canonique

$$\begin{aligned} \sigma_k \colon \mathbb{Q}(\theta) &\to \mathbb{Q}(\theta_k) \\ \sum_i q_i \theta^i &\mapsto \sum_i q_i \theta_k^i \end{aligned} \tag{6.321}$$

Nous avons $\sigma_1 \colon \mathbb{K} \to \mathbb{K}$ qui est l'identité.

Notons N le degré du polynôme $P \in \mathbb{K}[T_1, \dots, T_m]$ dont il est question dans l'énoncé. Nous le décomposons alors en

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_{il} T_i^l \tag{6.322}$$

avec $c_{il} \in \mathbb{K}$. Nous voyons $c_{i,}$ comme un élément de $\mathbb{K}^m$ et donc nous écrivons [79]

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_l(\theta)_i T_i^l \tag{6.323}$$

où $c_l \in \mathbb{Q}[X]^m$. Nous pouvons choisir $\deg(c_l) < \delta$ parce que les puissances plus grandes de θ ne génèrent rien de nouveau.

Nous posons aussi

$$P^{\sigma_k} = \sum_{l,i} c_l(\theta_k)_i T_i^l \in \mathbb{Q}(\theta_k)[T_1, \dots, T_m], \tag{6.324}$$

et $\bar{P} = P P^{\sigma_2} \dots P^{\sigma_k}$. Le coefficient de T_i^l dans $\bar{P}$ est

$$\bar{c}_l(\theta_1, \dots, \theta_\delta)_i = \sum_{l_1 + \dots + l_\delta = l} c_{l_1}(\theta_1)_i \dots c_{l_\delta}(\theta_\delta)_i. \tag{6.325}$$

Ce dernier est un polynôme en les θ_k à coefficients dans $\mathbb{Q}$. Qui plus est, c'est un polynôme symétrique. En effet un terme contenant $\theta_k^a \theta_l^b$ provenant de $c_{l_i}(\theta_k) c_{l_j}(\theta_l)$ a un terme correspondant $\theta_k^b \theta_l^a$ provenant de $c_{l_j}(\theta_k) c_{l_i}(\theta_l)$.

C'est donc le moment d'utiliser le théorème 6.181 à propos des polynômes symétriques élémentaires qui nous dit que les coefficients de $\bar{P}$ sont en réalité des polynômes en ceux de P_θ qui sont dans $\mathbb{Q}$. Donc $\bar{P} \in \mathbb{Q}[T_1, \dots, T_m]$. Par ailleurs nous avons que

$$\deg(\bar{P}) = \delta \deg(P) \tag{6.326}$$

parce que $\bar{P}$ est le produit de δ « copies » de P. De plus $P = P^{\sigma_1}$ divise $\bar{P}$ donc on a bien que si $P(z) = 0$ alors $\bar{P}(z) = 0$. Le polynôme $\bar{P}$ est celui que nous cherchions. □

79. Il me semble qu'il manque la somme sur i dans [114].

6.6.3 Relations coefficients racines

ThoOQRgjpl

Théorème 6.184 (Relations coeffitients-racines)**.**
Soit le polynôme $P = a_nX^n + \cdots + a_1X + a_0$ et r_i ses n racines. Alors nous avons pour chaque $1 \leqslant k \leqslant n$ la relation

$$\sigma_k(r_1, \ldots, r_n) = (-1)^k \frac{a_{n-k}}{a_n} \tag{6.327}$$

où σ_k est le k^epolynôme symétrique défini en 6.180.

ExHIfHhBr

Exemple 6.185.
Soit le polynôme

$$P(x) = x^3 + 2x^2 + 3x + 4 \tag{6.328}$$

et ses racines que nous nommons a, b, c. Nous voudrions calculer $a^2 + b^2 + c^2$. D'abord nous décomposons $Q(a, b, c) = a^2 + b^2 + c^2$ en polynômes symétriques élémentaires : $Q(a, b, c) = \sigma_1(a, b, c)^2 - 2\sigma_2(a, b, c)$.

Mais les relations coefficients-racines [80] nous donnent $\sigma_1(a, b, c) = -2$ et $\sigma_2(a, b, c) = 3$, donc

$$a^2 + b^2 + c^2 = (-2)^2 - 2 \cdot 3 = -2. \tag{6.329}$$

Cela nous assure déjà qu'au moins une des solutions n'est pas réelle.

Nous pouvons en avoir une vérification directe en calculant explicitement les racines (ce qui est possible pour le degré 3) :

```
sage: P(x)=x**3+2*x**2+3*x+4
sage: S=solve(  P(x)==0,x  )
sage: sols=[ s.rhs() for s in S  ]
sage: Q=[  s**2 for s in sols ]
sage: s=sum(Q)
sage: s.simplify_full()
-2
```

tex/frido/VAYVmNRpolynomeSym.py

Notez qu'il faut un peu chipoter pour isoler les solutions depuis la réponse de la fonction `solve`. △

En suivant le même cheminement que dans l'exemple, si P est un polynôme de degré n et si r_i sont ses racines, il est facile de calculer $Q(r_1, \ldots, r_n)$ pour n'importe quel polynôme symétrique Q

PropTETooGuBYQf

Proposition 6.186 (Annulation de fonctions polynomiales[183])**.**
Soit $\mathbb{K}$ un corps et P un polynôme à n indéterminées. Nous supposons que P s'annule sur un ensemble de la forme $A_1 \times \cdots \times A_n$ avec $\mathrm{Card}(A_j) > \deg_{X_j}(P)$ pour tout j. Alors $P = 0$.

De plus si $P = 0$ alors tous ses coefficients sont nuls [81].

Démonstration. Nous prouvons le résultat par récurrence sur le nombre n d'indéterminées. Si $n = 1$, cela est le théorème 6.111. Nous classons les monômes du polynôme P par ordre de puissance de X_n et nous le factorisons :

$$P = \sum_{i=1}^{m} P_i X_n^i \tag{6.330}$$

80. Théorème 6.184

81. L'intérêt de cela est qu'un polynôme de $\mathbb{Z}[X_1, \ldots, X_n]$ peut s'évaluer sur un élément de n'importe quel corps ; il restera le polynôme nul.

avec $P_i \in \mathbb{K}[X_1, \ldots, X_{n-1}]$. Soit $(a_1, \ldots, a_{n-1}) \in A_1 \times \ldots \times A_{n-1}$ et posons

$$Q(T) = P(a_1, \ldots, a_{n-1}, T) = \sum_{i=1}^{m} P_i(a_1, \ldots, a_{n-1})T^i. \tag{6.331}$$

Le polynôme Q s'annule sur A_n avec $\deg(Q) = \deg_{X_n}(P) < \text{Card}(A_n)$ et le théorème 6.111 nous donne $Q = 0$. Or les coefficients des différentes puissances de T dans $Q(T)$ sont les $P_i(a_1, \ldots, a_{n-1})$; ils sont donc nuls.

Nous avons montré que le polynôme P_i s'annule pour tout élément de $A_1 \times \ldots \times A_{n-1}$, mais nous avons

$$\deg_{X_j}(P_i) \leqslant \deg_{X_j} P < \text{Card}(A_j), \tag{6.332}$$

donc l'hypothèse de récurrence donne $P_i = 0$. Par suite, $P = 0$ également. □

6.7 Minuscule morceau sur la théorie de Galois

Vous trouverez des détails et des preuves à propos de la théorie de Galois dans [184, 100].

Définition 6.187.
Soit $\mathbb{K}$, un corps.

*Le **groupe de Galois** d'une extension $\mathbb{L}$ de $\mathbb{K}$ est le groupe des automorphismes de $\mathbb{L}$ laissant $\mathbb{K}$ invariant.*

Le groupe de Galois d'un polynôme sur $\mathbb{K}$ est le groupe de Galois de son corps de décomposition sur $\mathbb{K}$.

Définition 6.188.
*Des éléments $b_1, \ldots, b_n$ d'une extension de $\mathbb{K}$ sont **algébriquement indépendants** si ils ne satisfont à aucune relation du type*

$$\sum \alpha_{i_1 \ldots i_n} b_1^{i_1} \ldots b_n^{i_n} = 0 \tag{6.333}$$

avec $\alpha_{i_1 \ldots i_n} \in \mathbb{K}$.

Nous disons que l'équation

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{6.334}$$

est l'**équation générale** de degré n si les coefficients a_i sont algébriquement indépendants sur $\mathbb{K}$.

THOooDTCXooTIOSZe

Théorème 6.189.
Le groupe de Galois d'un polynôme de degré n est isomorphe au groupe symétrique S_n.

COROooBRKTooBWldRC

Corolaire 6.190 ([185])**.**
L'équation générale de degré n est résoluble par radicaux si et seulement si $n \leqslant 5$.

Chapitre 7

Topologie générale

7.1 Éléments généraux de topologie

7.1.1 Définitions et propriétés de base

DefTopologieGene

Définition 7.1 ([186]).
Soient X, un ensemble et $\mathcal{T}$, une partie de l'ensemble de ses parties qui vérifie les propriétés suivantes.

(1) Les ensembles $\varnothing$ et X sont dans $\mathcal{T}$,

(2) Une union quelconque[1] *d'éléments de $\mathcal{T}$ est dans $\mathcal{T}$.*

(3) Une intersection finie *d'éléments de $\mathcal{T}$ est dans $\mathcal{T}$.*

*Un tel choix $\mathcal{T}$ de sous-ensembles de X est une **topologie** sur X, et les éléments de $\mathcal{T}$ sont appelés des **ouverts**. On dit aussi que $(X, \mathcal{T})$ (voire simplement X lorsqu'il n'y a pas d'ambiguïté) est un **espace topologique**.*

Deux espaces topologiques sont isomorphes quand il existe une bijection continue d'inverse continue. Nous verrons ça en la définition 7.41.

7.1.2 Base de topologie

DEFooLEHPooIlNmpi

Proposition-Définition 7.2 (Base de topologie[187, 188]).
Soit un espace topologique $(X, \mathcal{T})$. Soit une partie $\mathcal{B}$ de $\mathcal{T}$. Les propriétés suivantes sont équivalentes :

ITEMooCTPEooRCaxvx

(1) Tout élément de $\mathcal{T}$ est une réunion d'éléments de $\mathcal{B}$.

ITEMooWOVWooRozYmM

(2) Pour tout $x \in X$ et pour tout ouvert $\mathcal{O}$ contenant x, il existe $B \in \mathcal{B}$ tel que

$$x \in B \subset \mathcal{O}. \tag{7.1}$$

*Une partie $\mathcal{B}$ de $\mathcal{T}$ qui vérifie ces propriétés est une **base de topologie** pour X.*

Démonstration. En deux parties.

(1) **(1) implique (2)** Soient $x \in X$ et $\mathcal{O}$ un ouvert contenant x. Étant donné que $\mathcal{O}$ est une réunion d'éléments de $\mathcal{B}$, il y a au moins un $B \in \mathcal{B}$ contenant x. Ce B vérifie $x \in B \subset \mathcal{O}$.

(2) **(2) implique (1)** Soit $\mathcal{O}$ un ouvert de X ; pour chaque $x \in \mathcal{O}$ nous considérons un ouvert $B(x) \in \mathcal{B}$ tel que $x \in B(x) \subset \mathcal{O}$. Nous avons alors $\mathcal{O} = \bigcup_{x \in \mathcal{O}} B(x)$.

□

1. Par « quelconque » nous entendons vraiment quelconque : c'est-à-dire indicée par un ensemble qui peut autant être $\mathbb{N}$ que $\mathbb{R}$ qu'un ensemble encore considérablement plus grand.

7.1.3 Fermés

DEFFermeooNSAAooHxZbAo

Définition 7.3.
*Si X est un espace topologique, un sous-ensemble F de X est dit **fermé** si son complémentaire, $X\backslash F$, est ouvert.*

DEFVoisinageooGHZCooLRcpXY

Définition 7.4.
*Si $a \in X$, on dit que $V \subset X$ est un **voisinage** de a si il existe un ouvert $\mathcal{O} \in \mathcal{T}$ tel que $a \in \mathcal{O}$ et $\mathcal{O} \subset V$.*

DEFooBWZIooXotZLA

Définition 7.5 (Base de voisinage[189])**.**
*Soient un espace topologique X ainsi que $a \in X$. Un ensemble $\{U_i\}_{i\in I}$ de voisinages de a est une **base de voisinages** de a si pour tout voisinage V de a, il existe $k \in I$ tel que $U_k \subset V$.*

LemQYUJwPC

Lemme 7.6.
Union et intersection de fermés.

ITEMooBHIGooMvkUtX

(1) Une intersection quelconque de fermés est fermée.

ItemKJYVooMBmMbG

(2) Une union finie de fermés est fermée.

Démonstration. Soit $\{F_i\}_{i\in I}$ un ensemble de fermés ; nous avons

$$\left(\bigcap_{i\in I} F_i\right)^c = \bigcup_{i\in I} F_i^c. \tag{7.2}$$

Le membre de droite est une union d'ouverts, c'est donc un ouvert ; donc l'intersection qui apparaît dans le membre de gauche est le complémentaire d'un ouvert : c'est donc un fermé.

De la même manière, le complémentaire d'une union finie de fermés est une intersection finie de complémentaires de fermés, et est donc ouvert [2]. □

LEMooSFMZooBguLdf

Lemme 7.7.
Ouverts et fermés.

ITEMooFPGVooUZHzkA

(1) Si F est fermé et si $\mathcal{O}$ est ouvert, alors $F\backslash\mathcal{O}$ est fermé.

ITEMooULNTooKQyrJb

(2) Si F est fermé dans l'ouvert $\mathcal{O}$, alors $\mathcal{O}\backslash F$ est ouvert.

Démonstration. Nous savons par le lemme 1.29 que $(F\backslash\mathcal{O})^c = F^c \cup \mathcal{O}$. Vu que F^c et $\mathcal{O}$ sont ouverts, l'union est ouverte. Le complémentaire de $F\backslash\mathcal{O}$ étant ouvert, il est fermé.

Pour le second point, nous considérons le complémentaire. Si nous nommons X l'espace topologique, nous avons

$$X\backslash(\mathcal{O}\backslash F) = (X\backslash\mathcal{O}) \cup F. \tag{7.3}$$

Étant une union de fermés, cela est fermé (lemme 7.6). Donc $\mathcal{O}\backslash F$ est ouvert. □

Dans un espace topologique, nous avons une caractérisation très importante des ouverts.

ThoPartieOUvpartouv

Théorème 7.8.
Une partie d'un espace topologique est ouverte si et seulement si elle contient un ouvert autour de chacun de ses éléments.

Démonstration. Soit X un espace topologique et $A \subset X$. Le sens direct est évident : A lui-même est un ouvert autour de $x \in A$, qui est inclus dans X.

Pour le sens inverse, nous supposons que A contienne un ouvert autour de chacun de ses points. Pour chaque $x \in A$, choisissons $\mathcal{O}_x \subset A$ un ouvert autour de x. Alors,

$$A = \bigcup_{x\in A} \mathcal{O}_x \tag{7.4}$$

EqAUniondesOx

2. Un bon exercice est d'écrire ces unions et intersections, pour se convaincre que ça fonctionne.

en effet, d'une part, $A \subset \bigcup_{x\in A} \mathcal{O}_x$ parce que chaque élément x de A est dans le $\mathcal{O}_x$ correspondant, par construction ; et d'autre part, $\bigcup_{x\in A} \mathcal{O}_x \subset A$ parce que chacun des $\mathcal{O}_x$ est inclus dans A.

Ainsi, A est égal à une union d'ouverts, cela prouve que A est un ouvert. □

Le lemme 7.232 est une version particulière de celui-ci, pour l'espace topologique $\mathbb{R}$. Une autre application typique est la proposition 7.2 et le théorème 7.205.

7.1.4 Quelques exemples

7.1.4.1 Une première vague

DefTopologieGrossiere

Exemple 7.9.
Soit un ensemble quelconque X. L'ensemble de parties $\mathcal{T} = \{\emptyset, X\}$ est une topologie sur X.

La topologie ainsi définie sur X est appelée **topologie grossière**. △

DefTopologieDiscrete

Exemple 7.10.
Pour un ensemble X quelconque, on considère l'ensemble $\mathcal{T}$ constitué de toutes les parties de X. Avec cet ensemble, on confère à nouveau une structure d'espace topologique à X ; toutes les parties sont des ouverts, et aussi des fermés. La topologie ainsi posée sur X est appelée **topologie discrète**. △

EXooLAOSooJtjJnu

Exemple 7.11 (Toutes les topologies d'un ensemble à 3 éléments)**.**
On pose $X = \{1, 2, 3\}$. Alors on peut munir X de 29 topologies différentes[190] ; saurez-vous les retrouver toutes ? △

7.1.4.2 Topologie engendrée

DefTopologieEngendree

Proposition-Définition 7.12 (Topologie engendrée, prébase[191])**.**
Soient un ensemble X, et $\mathcal{T}_0$ un ensemble de parties de X. Nous définissons $\tau(\mathcal{T}_0)$ comme étant l'union quelconque d'intersections finies d'éléments de $\mathcal{T}_0$.

Plus précisément, nous faisons les constructions suivantes :

(1) Nous notons $\{\mathcal{O}_i\}_{i\in I}$ les éléments de $\mathcal{T}_0$ indexés par l'ensemble I.

(2) Soit $B(\mathcal{T}_0)$ l'ensemble des intersections finies d'éléments de $\mathcal{T}_0$:

$$B(\mathcal{T}_0) = \Big\{ \bigcap_{j\in J} \mathcal{O}_j \Big\}_{J \text{ fini dans } I} \tag{7.5}$$

où nous convenons que $\bigcap_{j\in\emptyset} \mathcal{O}_j = X$ [3].

(3) Soit A un ensemble qui indexe $B(\mathcal{T}_0)$:

$$B(\mathcal{T}_0) = \{B_\alpha\}_{\alpha\in A}. \tag{7.6}$$

(4) Nous posons

$$\tau(\mathcal{T}_0) = \Big\{ \bigcup_{\alpha\in S} B_\alpha \Big\}_{S\subset A}. \tag{7.7}$$

Alors :

ITEMooTCGTooJwfpel

(1) $\tau(\mathcal{T}_0)$ est une topologie sur X.

ITEMooBJVVooEVRgdq

(2) Toute topologie sur X contenant $\mathcal{T}_0$ contient $\tau(\mathcal{T}_0)$.

*La topologie $\tau(\mathcal{T}_0)$ est appelée la **topologie engendrée** par $\mathcal{T}_0$. La partie $\mathcal{T}_0$ est appelée **prébase** de la topologie $\tau(\mathcal{T}_0)$.*

Démonstration. Pour (1), nous devons montrer les différents points de la définition 7.1 d'une topologie.

3. Bref, nous mettons X dans $B(\mathcal{T}_0)$.

(1) L'ensemble vide est dans $\tau(\mathcal{T}_0)$ parce que $\emptyset = \bigcup_{\alpha\in\emptyset} B_\alpha$. L'ensemble X est également dans $\tau(\mathcal{T}_0)$ parce que $X \in B(\mathcal{T}_0)$.

(2) Soient $\{D_l\}_{l\in L}$ des éléments de $\tau(\mathcal{T}_0)$ indexés par un ensemble L. Pour chaque l nous avons un ensemble $S \subset A$ tel que $D_l = \bigcup_{\alpha\in S_l} B_\alpha$. En posant $S = \bigcup_{l\in L} S_l$ nous avons

$$\bigcup_{l\in L} D_l = \bigcup_{\alpha\in S} B_\alpha \in \tau(\mathcal{T}_0). \tag{7.8}$$

Donc $\tau(\mathcal{T}_0)$ est stable par union quelconque.

(3) Soient D_1 et D_2 des éléments de $\tau(\mathcal{T}_0)$. Nous posons $D_i = \bigcup_{\alpha\in S_i} B_\alpha$. Alors nous avons

$$\bigcup_{\alpha\in S_1} B_\alpha \cap \bigcup_{\beta\in S_2} B_\beta = \bigcup_{\alpha,\beta\in S_1\times S_2} (B_\alpha \cap B_\beta). \tag{7.9}$$ EQooUCJOooCbKVpw

Mais B_α et B_β sont dans $B(\mathcal{T}_0)$. Donc $B_\alpha \cap B_\beta \in B(\mathcal{T}_0)$. Donc (7.9) est une union d'éléments de $B(\mathcal{T}_0)$.

Au final nous avons prouvé que $\tau(\mathcal{T}_0)$ est une topologie sur X.

Nous démontrons à présent le point (2). Soit une topologie μ sur X contenant $\tau(\mathcal{T}_0)$. Puisque μ est une topologie, les intersections finies d'éléments de μ sont dans μ, donc, en suivant les notations de 7.12, $B(\mathcal{T}_0) \subset \mu$.

Comme toutes les unions d'éléments de μ sont dans μ, l'inclusion de $B(\mathcal{T}_0)$ dans μ implique celle de $\tau(\mathcal{T}_0)$. □

Dès que nous avons une topologie, nous avons une notion de convergence de suite.

DefXSnbhZX

Définition 7.13 (Convergence de suite)**.**
*Une suite (x_n) d'éléments de X **converge** vers un élément x de X si pour tout ouvert $\mathcal{O}$ contenant x , il existe un $K \in \mathbb{N}$ tel que $k \geqslant K$ implique $x_k \in \mathcal{O}$.*

*Nous disons alors que x est une **limite** de (x_k).*

La proposition suivante montre que vérifier la convergence d'une suite sur une prébase suffit pour vérifier la convergence.

PROPooJTJBooNtczsO

Proposition 7.14.
Soit $\mathcal{T}_0$ un ensemble de parties de l'ensemble X. Soient une suite (x_n) dans X ainsi que $x \in X$. Nous supposons que la suite (x_n) satisfait la propriété suivante : pour tout $A \in \mathcal{T}_0$ tel que $x \in A$, il existe $K \in \mathbb{N}$ tel que $k \geqslant K$ implique $x_k \in A$.

Alors nous avons la convergence de suite[4]

$$x_n \xrightarrow{(X,\tau(\mathcal{T}_0))} x. \tag{7.10}$$

Démonstration. Nous considérons la topologie $\tau(\mathcal{T}_0)$ sur X. Soit un ouvert $\mathcal{O}$ contenant x. Nous le décomposons en suivant (à l'envers) la construction de la définition 7.12 :

$$\mathcal{O} = \bigcup_{\alpha\in S} B_\alpha \tag{7.11}$$

avec $B_\alpha \in B(\mathcal{T}_0)$. Donc pour chaque α, il existe un ensemble fini J_α tel que

$$B_\alpha = \bigcap_{j\in J_\alpha} A_j \tag{7.12}$$

avec $A_j \in \mathcal{T}_0$. Puisque $x \in \mathcal{O}$, nous avons un α_0 tel que $x \in B_{\alpha_0}$. Donc $x \in A_j$ pour tous les $j \in J_{\alpha_0}$.

Pour chaque $j \in J_{\alpha_0}$, il existe $K_j \in \mathbb{N}$ tel que $k \geqslant K_j$ implique $x_k \in A_j$. Comme J_{α_0} est un ensemble fini, nous pouvons poser $K = \max_{j\in J_{\alpha_0}} K_j$.

Maintenant, si $k \geqslant K$, nous avons $x_k \in A_j$ pour tout j, et donc $x_k \in B_{\alpha_0}$. Par conséquent, $x_k \in \mathcal{O}$. □

4. Définition 7.13.

7.1.5 Topologie produit

DefIINHooAAjTdY

Proposition-Définition 7.15 (Produit d'espaces topologiques, thème 27)**.**
Soient $\{(X_i, \tau_i)\}_{i=1,\dots,n}$ des espaces topologiques. Nous posons

$$X = \prod_{i=1}^{n} X_i \tag{7.13}$$

et

$$\mathcal{T} = \{\mathcal{O} \subset X \text{ tel que } \forall x \in U, \exists U_i \in \tau_i \text{ tel que } x \in U_1 \times \dots \times U_n \subset \mathcal{O}\}. \tag{7.14}$$

EQooVECGooKLUkTy

Alors $(X, \mathcal{T})$ est un espace topologique.

Cette topologie est la ***topologie produit****.*

Dans le cas d'espaces normés, nous verrons dans le lemme 7.227 que la topologie produit est la même que celle obtenue par la norme produit.

Proposition 7.16 (Convergence composante par composante)**.**
Soient des espaces topologiques X_i $(i = 1, \dots, n)$ et une suite $(a_k^{(1)}, \dots, a_k^{(n)})$ dans $X_1 \times \dots \times X_n$. Nous avons la convergence

$$(a_k^{(1)}, \dots, a_k^{(n)}) \xrightarrow{X_1 \times \dots \times X_n} (a^{(1)}, \dots, a^{(n)}) \tag{7.15}$$

si et seulement si $a_k^{(i)} \to a^{(i)}$ pour chaque i.

Démonstration. En deux parties.

(1) **Sens direct** Soient des ouverts $\mathcal{O}_i$ autour de $a^{(i)}$ dans X_i. Puisque $\mathcal{O}_1 \times \dots \times \mathcal{O}_n$ est un ouvert autour de $(a^{(1)}, \dots, a^{(n)})$, il existe $K \in \mathbb{N}$ tel que si $k \geqslant K$ nous avons $(a_k^{(1)}, \dots, a_k^{(n)}) \in \mathcal{O}_1 \times \dots \times \mathcal{O}_n$. Pour ce K nous avons séparément $a_k^{(i)} \in \mathcal{O}_i$ pour chaque i.

(2) **Sens inverse** Une prébase de la topologie sur $X_1 \times \dots \times X_n$ est donnée par les $\mathcal{O}_1 \times \dots \times \mathcal{O}_n$ où $\mathcal{O}_i$ est un ouvert de X_i. Voir la définition 7.15 de la topologie produit et la définition 7.12 de ce qu'est une prébase.

La proposition 7.14 nous permet de ne vérifier la convergence de $(a_k^{(1)}, \dots, a_k^{(n)})$ que sur la prébase. Soit donc $\mathcal{O} = \mathcal{O}_1 \times \dots \times \mathcal{O}_n$ avec $(a^{(1)}, \dots, a^{(n)}) \in \mathcal{O}$. Puisque $(a_k^{(i)})_{k \in \mathbb{N}} \to a^{(i)}$, pour chaque i, il existe $K_i \in \mathbb{N}$ tel que si $k \geqslant K_i$ alors $a_k^{(i)} \in \mathcal{O}_i$.

En posant $K = \max_i(K_i)$, nous avons $(a_k^{(1)}, \dots, a_k^{(n)}) \in \mathcal{O}_1 \times \dots \mathcal{O}_n$ pour tout $k \geqslant K$.

La proposition 7.14 permet de conclure.

□

7.1.6 Adhérence, fermeture, intérieur, point d'accumulation et point isolé

DEFooSVWMooLpAVZRInt

Définition 7.17.
Soient un espace topologique X et une partie A de X.

(1) Un point $x \in X$ est ***intérieur*** *à A si il existe un ouvert $\mathcal{O}$ tel quex $\in \mathcal{O} \subset A$.*

*(2) L'****intérieur*** *de A, notée $\mathrm{Int}(A)$, est l'union de tous les ouverts de X contenus dans A.*

Lemme 7.18.
Quelques propriétés en vrac.

ITEMooHLIMooJEacKt

(1) L'intérieur de A est l'ensemble de tous les points intérieurs de A.

ITEMooYTXSooMyiBpMgzK

(2) Pour tout $A \subset X$, l'ensemble $\mathrm{Int}(A)$ est un ouvert.

ITEMooYYFDooHgsRfV

(3) On a $\mathrm{Int}(A) \subset A$

ITEMooTDXFooFdyLeO

(4) Nous avons $\mathrm{Int}(A) = A$ si et seulement si A est un ouvert.

Démonstration. En plusieurs morceaux.

(1) **(1)** Si $a \in \text{Int}(A)$, alors a est dans un ouvert contenu dans A, et donc a est un point intérieur à A. Dans l'autre sens, si a est un point intérieur à A, alors il existe un ouvert $\mathcal{O} \subset A$ contenant a. Puisque $\mathcal{O}$ est un ouvert dans A, nous avons $\mathcal{O} \subset \text{Int}(A)$, et en particulier $a \in \text{Int}(A)$.

(2) **(2)** L'ensemble $\text{Int}(A)$ est une union d'ouverts.

(3) **(3)** L'ensemble $\text{Int}(A)$ est une union d'ensembles contenus dans A.

(4) **(4)** Supposons que $\text{Int}(A) = A$. Puisque $\text{Int}(A)$ est ouvert (point (3)), A est ouvert aussi. Dans l'autre sens, nous supposons que A est ouvert. Puisque A est un ouvert contenu dans A, nous avons $A \subset \text{Int}(A)$. Mais comme $\text{Int}(A) \subset A$, nous avons l'égalité.

□

7.1.6.1 Adhérence et fermeture

7.19.
Les mots « adhérence » et « fermeture » sont synonymes. Dans le Frido, nous allons utiliser les notations $\text{Adh}(A)$ et $\bar{A}$ de façon opportuniste. La notation $\bar{z}$ définissant le complexe conjugué de z, si A est une partie de $\mathbb{C}$, il est plus sûr d'écrire $\text{Adh}(A)$ pour la fermeture, plutôt que $\bar{A}$.

Au contraire, pour éviter une quantité excessive de parenthèses, nous écrirons $\overline{B(a,r)}$ pour la boule fermée.

DEFooSVWMooLpAVZR

Définition 7.20.
*Soient un espace topologique X et une partie A de X. Un point $x \in X$ est **adhérent** à A si tout ouvert de X contenant x a une intersection non vide avec A. L'ensemble des points d'adhérence de A est noté $\text{Adh}(A)$ ou $\bar{A}$.*

LEMooILNCooOFZaTe

Lemme 7.21.
À propos d'adhérence.

(1) L'adhérence de A est l'intersection de tous les fermés de X contenant A.

(2) Nous avons l'égalité

$$\text{Int}(A)^c = \text{Adh}(A^c). \tag{7.16}$$

Démonstration. Commençons par prouver la dernière égalité d'ensembles. On a les équivalences entre les affirmations suivantes, pour tout $x \in X$:

— x n'est pas dans $\text{Int}(A)$;

— il n'y a aucun ouvert contenant x et inclus dans A ;

— tout ouvert contenant x a une intersection non-vide avec A^c ;

— x est dans $\text{Adh}(A^c)$.

Nous allons à présent montrer l'égalité d'ensembles $\text{Int}(A)^c = \text{Adh}(A^c)$ en prouvant la double inclusion.

(1) **Si $\text{Int}(A)^c \subset \text{Adh}(A^c)$** Soit $x \in \text{Int}(A)^c$. Nous devons prouver que $x \in \text{Adh}(A^c)$. Soit un ouvert $\mathcal{O}$ contenant x. Vu que x n'est pas dans l'intérieur de A, l'ouvert $\mathcal{O}$ est pas inclus dans A, et donc $\mathcal{O} \cap A^c$ est non vide.
Nous avons montré que tout ouvert contenant x intersecte A^c. Autrement dit : $x \in \text{Adh}(A^c)$.

(2) $\text{Adh}(A^c) \subset \text{Int}(A)^c$
Soit $x \in \text{Adh}(A^c)$. Tout ouvert contenant x intersecte A^c, et ne peut donc pas être inclus dans A. Si aucun ouvert contenant x n'est inclus dans A, alors x n'est pas dans $\text{Int}(A)$.

□

RemAdhFerme

Remarque 7.22.
Comme corolaire du lemme 7.21, et grâce aux remarques faites pour les intérieurs, on obtient que pour $A \subset X$:

(1) l'ensemble $\bar{A}$ est fermé : c'est en effet le complémentaire d'un ouvert, précisément l'intérieur de A^c ;

(2) A est fermé si et seulement si $\bar{A} = A$: en effet, A est fermé si et seulement si A^c est ouvert, si et seulement si l'intérieur de A^c est A^c lui-même ; or, l'intérieur de A^c est le complémentaire de $\bar{A}$ par le lemme 7.21, si bien que A est fermé si et seulement si $(\bar{A})^c = A^c$, ou encore... ce qu'on affirmait au début.

DefEnsembleDense

Définition 7.23.
*Soit X un espace topologique. Un sous-ensemble A de X est **dense** dans X si $\bar{A} = X$.*

DEFooQUNJooWZasqV

Définition 7.24.
*Une partie A d'un espace topologique X est **discrète** si pour tout $a \in A$, il existe un ouvert $\mathcal{O}$ de X tel que $\mathcal{O} \cap A = \{a\}$.*

7.1.6.2 Frontière

Définition 7.25.
*Soit X un espace topologique, et $A \subset X$. La **frontière** de A, notée ∂A, est l'ensemble des points adhérents de A qui ne sont pas intérieurs à A. Ainsi,*

$$\partial A = \mathrm{Adh}(A) \setminus \mathrm{Int}(A). \tag{7.17}$$

7.1.6.3 Topologie induite

DefVLrgWDB

Proposition-Définition 7.26 (Topologie induite[192])**.**
Soit un espace topologique (X, τ_X), et soit $Y \subset X$. Nous définissons

$$\tau_Y = \{Y \cap \mathcal{O} \text{ tel que } \mathcal{O} \in \tau_X\}. \tag{7.18}$$

L'ensemble τ_Y est une topologie sur Y.
*Elle est la **topologie induite**.*

Démonstration. Il s'agit de vérifier les conditions de la définition 7.1.

(1) $Y \in \tau_Y$ Parce que $Y = X \cap Y$ et que X est un ouvert de X.

(2) $\varnothing \in \tau_Y$ Parce que $\varnothing = Y \cap \varnothing$ et que $\varnothing$ est un ouvert de X.

(3) **Union quelconque** Soient des ouverts A_i de X. Nous avons

$$\bigcup_{i\in I} Y \cap A_i = Y \cap \big(\bigcup_{i\in I} A_i\big). \tag{7.19}$$

EQooJUYHooIugQXG

Comme les A_i sont des ouverts de X, leur union est encore un ouvert de X. Donc (7.19) est encore dans τ_Y.

(4) **Intersection finie** Nous avons

$$\bigcap_{i\in I} Y \cap A_i = Y \cap \left(\bigcap_{i\in I} A_i\right). \tag{7.20}$$

□

LemBWSUooCCGvax

Lemme 7.27 ([1])**.**
Soit (X, τ_X) un espace topologique et $S \subset X$, un fermé de X sur lequel nous considérons la topologie induite τ_S. Si F est un fermé de (S, τ_S) alors F est fermé de (X, τ_X).

Démonstration. Nous savons que $F \subset S$ et que le complémentaire de F dans S est un ouvert de (S, τ_S) : il existe un ouvert $\Omega \in \tau_X$ tel que $S\backslash F = S \cap \Omega$. Si maintenant nous considérons le complémentaire de F dans X nous avons

$$F^c = (S\backslash F) \cup (X\backslash S) = (S \cap \Omega) \cup S^c = (S \cap \Omega) \cup (S^c \cap \Omega) \cup S^c = \Omega \cup S^c. \tag{7.21}$$

Puisque Ω et S^c sont des ouverts de X, l'union est un ouvert. Donc $F^c \in \tau_X$ et F est un fermé de X. □

DEFooEXHFooKlKcQD

Définition 7.28 ([193]).
*Si X est un espace topologique, un **sous-espace topologique** de X est une partie de X munie de la topologie induite*[5].

PROPooQFIMooYmgjhA

Proposition 7.29 ([1]).
Soit une partie $\mathcal{O}$ de l'espace topologique X. Les fermés de $\mathcal{O}$ sont les parties de la forme $F \cap \mathcal{O}$ où F est fermé de X.

Démonstration. En deux parties.

(1) **Dans un sens** Soit un fermé A de $\mathcal{O}$. La partie $\mathcal{O}\backslash A$ est ouverte dans $\mathcal{O}$, de telle sorte qu'il existe un ouvert S de X satisfaisant $\mathcal{O}\backslash A = S \cap \mathcal{O}$. La partie $F = X\backslash S$ est fermée dans X, et nous allons prouver que $A = (X\backslash S) \cap \mathcal{O}$. D'abord nous avons

$$(X\backslash S) \cap \mathcal{O} = \mathcal{O}\backslash S. \tag{7.22}$$

EQooORVBooCvgCgw

Nous avons l'union disjointe $\mathcal{O} = (\mathcal{O}\backslash S) \cup (\mathcal{O} \cap S)$. Tenant compte du fait que $\mathcal{O} \cap S = \mathcal{O}\backslash A$, nous avons l'union disjointe

$$\mathcal{O} = (\mathcal{O}\backslash S) \cup (\mathcal{O}\backslash A). \tag{7.23}$$

EQooXKPVooMoYFGB

Par ailleurs, vu que $A\backslash\mathcal{O}$, nous avons également l'union disjointe

$$\mathcal{O} = A \cup (\mathcal{O}\backslash A) \tag{7.24}$$

EQooQEMTooLEZwuy

En comparant (7.23) avec (7.24), nous trouvons $A = \mathcal{O} \cap S$.
En repartant de (7.22), nous avons maintenant

$$F \cap \mathcal{O} = (X\backslash S) \cap \mathcal{O} = \mathcal{O}\backslash S = A. \tag{7.25}$$

(2) **Dans l'autre sens** Soit un fermé F de X. Montrons que $F \cap \mathcal{O}$ est fermé dans $\mathcal{O}$. Pour cela nous prouvons que le complémentaire (dans $\mathcal{O}$) est ouvert dans $\mathcal{O}$. Nous avons :

$$\mathcal{O}\backslash(F \cap \mathcal{O}) = \mathcal{O} \cap (X\backslash F). \tag{7.26}$$

Comme F est fermé, $X\backslash F$ est ouvert.

□

LemkUYkQt

Lemme 7.30.
Si $B \subset A$ alors la fermeture de B pour la topologie de A (induite de X) que nous noterons $\tilde{B}$ est

$$\tilde{B} = \bar{B} \cap A \tag{7.27}$$

où $\bar{B}$ est la fermeture de B pour la topologie de X.

Démonstration. Si $a \in \bar{B} \cap A$, un ouvert de A autour de a est un ensemble de la forme $\mathcal{O} \cap A$ où $\mathcal{O}$ est un ouvert de X. Comme $a \in \bar{B}$, l'ensemble $\mathcal{O}$ intersecte B et donc $(\mathcal{O} \cap A) \cap B \neq \emptyset$. Donc a est bien dans l'adhérence[6] de B au sens de la topologie de A.

Pour l'inclusion inverse, soit $a \in \tilde{B}$, montrons que $a \in \bar{B} \cap A$. Par définition $a \in A$, parce que $\tilde{B}$ est une fermeture dans l'espace topologique A. Il faut donc seulement montrer que $a \in \bar{B}$. Soit donc $\mathcal{O}$ un ouvert de X contenant a ; par hypothèse $\mathcal{O} \cap A$ intersecte B (parce que tout ouvert de A contenant a intersecte B). Donc $\mathcal{O}$ intersecte B. Cela signifie que tout ouvert (de X) contenant a intersecte B, ou encore que $a \in \bar{B}$. □

5. Définition 7.26.
6. Définition 7.20.

7.1.7 Topologie sur $\mathbb{Q}$

Si A est un ouvert de X, on pourrait croire que la topologie induite n'a rien de spécial. Il est vrai que B sera ouvert dans A si et seulement si il est ouvert dans X, mais certaines choses surprenantes se produisent tout de même.

LEMooIGQCooOrroHT

Lemme 7.31.
La partie $\mathbb{Q}$ dans $\mathbb{R}$ est d'intérieur vide, et sa fermeture est $\mathbb{R}$.

ExloeyoR

Exemple 7.32.
Prenons $X = \mathbb{R}$ et $A =]0, 1[$. Si $B =]\frac{1}{2}, 1[$, alors la fermeture de B dans A sera $\tilde{B} = [\frac{1}{2}, 1[$ et non $[\frac{1}{2}, 1]$ comme on l'aurait dans $\mathbb{R}$. △

Prendre la topologie induite de $\mathbb{R}$ vers un fermé de $\mathbb{R}$ donne des boules un peu spéciales comme le montre l'exemple suivant.

ExKYZwYxn

Exemple 7.33.
Quid de la boule ouverte $B(1, \epsilon)$ dans le fermé $[0, 1]$? Par définition c'est

$$B(1, \epsilon) = \{x \in [0, 1] \text{ tel que } |x - 1| < \epsilon\} =]1 - \epsilon, 1]. \tag{7.28}$$

Oui, c'est *ouvert* dans $[0, 1]$. C'est d'ailleurs un des ouverts de la topologie induite de $\mathbb{R}$ sur $[0, 1]$.

Donc pour la topologie de $[0, 1]$, toutes les boules ouvertes $B(x, \delta)$ avec $x \in [0, 1]$ sont incluses dans $[0, 1]$. Bref, vous pouvez écrire

$$B(\frac{1}{2}, 10) \subset [0, 1], \tag{7.29}$$

mais vous avez intérêt à être très clair sur la topologie sous-entendue. △

7.1.7.1 Points d'accumulation et isolés

DEFooGHUUooZKTJRi

Définition 7.34.
*Soient un espace topologique X et une partie A de X. Un point $s \in X$ est un **point d'accumulation** de A si tout ouvert de X contenant s contient au moins un élément de $A\backslash\{s\}$.*

Quelle est la différence entre un point d'accumulation et un point d'adhérence ? La différence est que tous les points de A sont des points d'adhérence de A, parce que tout voisinage de $a \in A$ contient au moins a lui-même, alors que certains points de A peuvent ne pas être des points d'accumulation de A. Voir l'exemple 7.237.

Notons qu'un point d'accumulation de A dans X n'est pas spécialement dans A.

DEFooXIOWooWUKJhN

Définition 7.35.
*Soient un espace topologique X et une partie A de X. Un point $s \in A$ est un **point isolé** de A si il existe un voisinage ouvert $\mathcal{O}$ de s dans X tel que $A \cap \mathcal{O} = \{s\}$.*

7.1.8 Applications continues

La définition suivante est *la* définition de la continuité dans tous les cas.

DefOLNtrxB

Définition 7.36 (Application continue[194])**.**
Deux définitions :

ITEMooXARPooNzsWLr

*(1) Soient une application $f : X \to Y$ entre les espaces topologiques X et Y et un point $a \in X$. Nous disons que f est **continue** en a si pour tout ouvert W contenant $f(a)$, il existe un voisinage V de a dans X tel que $f(V) \subset W$.*

ITEMooEHGWooDdITRV

*(2) Une fonction $f : X \to Y$ est **continue** sur X si pour tout ouvert $\mathcal{O}$ de Y, l'ensemble*

$$f^{-1}(\mathcal{O}) = \{x \in X \text{ tel que } f(x) \in \mathcal{O}\} \tag{7.30}$$

defFminus1ofaset

est ouvert dans X.

PROPooOVKEooCkJmmO

Proposition 7.37 ([1]).
Une application $f\colon X \to Y$ entre deux espaces topologiques est continue sur X si et seulement si elle est continue en chacun des points de X.

Démonstration. Dans les deux sens.

(1) $\Rightarrow$ Nous supposons que $f\colon X \to Y$ est continue. Soit $a \in X$. Si W est un ouvert contenant $f(a)$, alors $V = f^{-1}(W)$ est un ouvert contenant a et vérifiant $f(V) \subset W$.

(2) $\Leftarrow$ Nous supposons que $f\colon X \to Y$ est continue en chaque point de X. Soit un ouvert $\mathcal{O}$ de Y. Pour $a \in f^{-1}(\mathcal{O})$, la partie $\mathcal{O}$ est un ouvert contenant $f(a)$. Donc il existe un ouvert V_a contenant a et tel que $f(V_a) \subset \mathcal{O}$.

Nous posons à présent $V = \bigcup_{a\in f^{-1}(\mathcal{O})} V_a$. C'est un ouvert comme union d'ouverts. Ensuite nous avons $V = f^{-1}(\mathcal{O})$. En effet d'une part nous avons $f(V_a) \subset \mathcal{O}$, donc $V_a \subset f^{-1}(\mathcal{O})$ pour tout $a \in X$. Donc $V \subset f^{-1}(\mathcal{O})$.

D'autre part, pour chaque $a \in f^{-1}(\mathcal{O})$ nous avons $a \in V_a$, et donc $f^{-1}(\mathcal{O}) \subset V$.

□

LEMooYTLSooKhetml

Lemme 7.38.
Soient deux espaces topologiques X et Y et une application $f\colon X \to Y$. Soit une base de topologie $\{A_i\}_{i\in I}$ de Y. Si $f^{-1}(A_i)$ est ouvert dans X pour tout $i \in I$ alors f est continue.

Exemple 7.39 ([1]).
Un truc bien avec la définition 7.36(1) est que la continuité de f en un point est définie pour tout point du domaine ; pas seulement les points d'accumulation. Soit par exemple une fonction simple

$$\begin{aligned} f\colon \{a\} &\to \mathbb{R} \\ a &\mapsto 4. \end{aligned} \tag{7.31}$$

Si W est un ouvert de $\mathbb{R}$ contenant 4, nous avons l'ouvert $V = \{a\}$ tel que $f(V) \subset W$. Donc f est continue au point 4.

Mais f est également continue sur $\{4\}$ en tant qu'espace topologique. En effet, si W est un ouvert de $\mathbb{R}$, l'ensemble $f^{-1}(W)$ est soit $\varnothing$ soit $\{a\}$. Dans les deux cas c'est un ouvert. △

LEMooATLRooEKnlro

Lemme 7.40.
Une application $f\colon X \to Y$ est continue si et seulement si pour tout fermés F de Y, la partie $f^{-1}(F)$ est fermée dans X.

Démonstration. Supposons que f est continue. Si F est fermé dans Y, alors F^c est ouvert et donc $f^{-1}(F^c) = f^{-1}(F)^c$ est ouvert. Donc $f^{-1}(F)$ est fermé.

Dans l'autre sens, si $\mathcal{O}$ est ouvert dans Y, alors $f^{-1}(\mathcal{O})^c = f^{-1}(\mathcal{O}^c)$ est fermé, de telle sorte que $f^{-1}(\mathcal{O})$ est ouvert. L'application f est alors continue. □

7.1.8.1 Isomorphismes

DEFooYPGQooMAObTO

Définition 7.41 (Isomorphisme d'espaces topologiques).
*Un **isomorphisme** d'espaces topologiques est une application bijective continue*[7] *entre deux espaces topologiques dont la réciproque est continue. On dit également **homéomorphisme**.*

*Un isomorphisme d'un espace avec lui-même est un **automorphisme**.*

LEMooMJSHooOszteq

Lemme 7.42 ([1]).
Si $f\colon X \to Y$ est un homéomorphisme[8] *et si F est fermé dans X, alors $f(F)$ est fermé dans Y.*

7. Application continue, définition 7.36.
8. Définition 7.41.

Démonstration. Le complémentaire F^c est ouverte. Vu que f^{-1} est continue, la partie $f(F^c)$ est ouverte. Comme f est une bijection, nous avons $f(F^c) = f(F)^c$. D'où le fait que $f(F)^c$ est ouvert, et donc que $f(F)$ est fermé. □

LEMooNMPGooRlgppQ

Lemme 7.43.
Soient un homéomorphisme $\varphi\colon X \to Y$ ainsi qu'une partie $S \subset X$. Alors

$$\varphi(\bar{S}) = \overline{\varphi(S)}. \tag{7.32}$$

Démonstration. Nous prouvons que $\overline{\varphi(S)} \subset \varphi(\bar{S})$; pour l'inclusion inverse, la lectrice fera elle-même.

Soit $x \in \overline{\varphi(S)}$. Nous allons prouver que $x \in \varphi(\bar{S})$, c'est à dire que $\varphi^{-1}(x) \in \bar{S}$. Pour cela nous considérons un ouvert $\mathcal{O}$ autour de $\varphi^{-1}(x)$. La partie $\varphi(\mathcal{O})$ est un ouvert contenant x. Donc $\varphi(\mathcal{O}) \cap \varphi(S) \neq \emptyset$, et donc $\mathcal{O} \cap S \neq \emptyset$, ce qui prouve que $\varphi^{-1}(x) \in \bar{S}$. □

7.2 Topologie rendant continue

PROPooGOEVooZBAOQh

Proposition-Définition 7.44 (Topologie initiale[195])**.**
Soient un ensemble X, des espaces topologiques $(Y_i, \tau_i)_{i\in I}$, et des applications $\varphi_i\colon X \to Y_i$. Nous notons

$$\Lambda = \{(i, \omega_i) \text{ tel que } i \in I, \omega_i \in \tau_i\}. \tag{7.33}$$

Pour chaque Γ fini dans Λ, nous posons

$$\Phi_\Gamma = \bigcap_{(i,\omega_i)\in\Gamma} \varphi_i^{-1}(\omega_i). \tag{7.34}$$

Enfin nous posons

$$\tau_I = \bigcup_{\Gamma \text{ fini dans } \Lambda} \Phi_\Gamma. \tag{7.35}$$

Nous avons :

(1) τ_I est une topologie sur X.

(2) Toutes les applications φ_i sont continues[9] *pour cette topologie.*

(3) La topologie τ_I est la plus faible topologie sur X pour laquelle toutes les φ_i sont continues.

*La topologie ainsi définie est souvent référée comme la plus petite topologie qui rend les applications φ_i continues. Elle est aussi appelée **topologie initiale** des φ_i.*

La topologie quotient d'un espace topologique par une relation d'équivalence est définie comme la plus petite topologie rendant continue la projection. Voir la définition 7.49.

LEMooVRGFooNgbwKu

Lemme 7.45 ([195, 196])**.**
Soient des espaces topologiques $\{(Y_i, \tau_i)\}_{i\in I}$, et X, un ensemble. Pour chaque i nous considérons une base de topologie $\mathcal{B}_i$ de τ_i. Nous considérons des applications $\varphi_i\colon X \to Y_i$.

L'ensemble

$$\mathcal{B} = \left\{ \bigcap_{j=1}^{n} \varphi_{\alpha_j}^{-1}(U_j) \text{ tel que } n \in \mathbb{N}, \alpha_1, \dots, \alpha_n \in I, U_k \in \mathcal{B}_k \right\} \tag{7.36}$$

est une base de la topologie initiale[10] *des φ_i.*

PROPooOPJCooAIUXCW

Proposition 7.46 ([1])**.**
Soient un espace topologique (Y, τ_Y) et une application $f\colon X \to Y$. Soient $x_0 \in X$ et $y_0 = f(y_0)$. Si $\{B_i\}_{i\in I}$ est une base de voisinages de y_0, alors $\{f^{-1}(B_i)\}_{i\in I}$ est une base de voisinages de x_0 pour la topologie initiale[11] *de f.*

9. Application continue, définition 7.36.
10. Proposition 7.44.
11. Définition 7.44.

Démonstration. Soit un ouvert $\mathcal{O}$ autour de x_0 pour la topologie initiale de f. En adaptant le lemme 7.45 au cas d'une seule application,

$$\mathcal{B} = \{f^{-1}(U) \text{ tel que } U \in \tau_Y\} \tag{7.37}$$

est une base de topologie sur X. Donc $\mathcal{O}$ est une union de la forme [12]

$$\mathcal{O} = \bigcup_{\alpha\in A} f^{-1}(U_\alpha) \tag{7.38}$$

avec $U_\alpha \in \tau_Y$. Soit $\alpha \in A$ tel que $x_0 \in f^{-1}(U_\alpha)$. Nous avons $y_0 = f(x_0) \in U_\alpha$. Vu que U_α est un ouvert contenant y_0, et vu que $\{B_i\}_{i\in I}$ est une base de voisinages de y_0, nous avons un $i \in I$ tel que

$$y_0 \in B_i \subset U_\alpha. \tag{7.39}$$

Donc nous avons aussi

$$f^{-1}(B_i) \subset f^{-1}(U_\alpha) \subset \mathcal{O}. \tag{7.40}$$

Donc tout voisinage de x_0 contient un élément de $\{f^{-1}(B_i)\}_{i\in I}$. Cela prouve que ce dernier est une base des voisinages de x_0. □

LEMooADPLooYylNsj

Lemme 7.47 ([195]).
Soient des espaces topologiques $\{(Y_i, \tau_i)\}_{i\in I}$, et X, un ensemble. Nous considérons des applications $\varphi_i\colon X \to Y_i$ ainsi que τ, la topologie minimale sur X telle que les applications φ_i soient continues.

Soit une suite (x_n) dans X. Nous avons $x_n \xrightarrow{\tau} x$ si et seulement si $\varphi_i(x_n) \xrightarrow{\tau_i} \varphi_i(x)$ pour tout $i \in I$.

Démonstration. Dans le sens direct, c'est seulement le fait que les φ_i sont continues.

Dans le sens réciproque, nous supposons que $\varphi_i(x_n) \xrightarrow{\tau_i} \varphi_i(x)$ pour tout i et nous devons prouver que $x_n \xrightarrow{\tau} x$.

Soit un voisinage U de x. Vue la base de topologie donnée dans le lemme 7.45, il existe un J fini dans I ainsi que des ouverts $\{V_j\}_{j\in J}$ tels que

$$x \in W = \bigcap_{j\in J} \varphi_j^{-1}(V_j) \subset U \tag{7.41}$$

Par hypothèse, $\varphi_i(x_n) \to \varphi_i(x)$. Mais V_j est un ouvert qui contient $\varphi_j(x)$. Donc il existe $N_j \in \mathbb{N}$ tel que $\varphi_j(x_n) \in V_j$ pour tout $n \geqslant N_j$. En posant [13] $N = \max\{N_j\}_{j\in J}$, nous avons que $\varphi_j(x_n) \in V_j$ pour tout $n \geqslant N$ et pour tout $j \in J$.

Dans ce cas nous avons aussi $x_n \in W \subset U$. La convergence est prouvée. □

Proposition 7.48 ([195]).
Soient des espaces topologiques (Y_i, τ_i), un ensemble X, et des applications $\varphi_i\colon X \to Y_i$. Nous considérons sur X la plus petite topologie rendant continues [14] les φ_i.

Soit un espace topologique Z. Une application $\psi\colon Z \to (X, \tau_I)$ est continue si et seulement si les applications $\varphi_i \circ \psi\colon Z \to Y_i$ sont continues pour tout $i \in I$.

Démonstration. Le sens direct est une simple composée de fonctions continues. Pour l'autre sens, nous supposons que les $\phi_i \circ \psi$ sont continues, nous considérons $U \in \tau_I$ et nous devons montrer que $\psi^{-1}(U)$ est un ouvert de Z.

En posant

$$\Lambda = \{(i, \omega) \text{ tel que } i \in I, \omega \in \tau_i\}, \tag{7.42}$$

il existe un Γ fini dans Λ tel que

$$U = \bigcap_{(i,\omega)\in\Gamma} \varphi_i^{-1}(\omega). \tag{7.43}$$

12. Définition 7.2(1).
13. Notez l'utilisation du lemme 1.74 pour justifier que le maximum existe.
14. Proposition 7.44.

Nous avons donc

$$\psi^{-1}(U) = \bigcap_{(i,\omega_i)\in\Gamma} (\psi \circ \varphi_i^{-1})(\omega_i). \tag{7.44}$$

Vu que $\psi \circ \varphi^{-1}$ est continue, chacun des $(\psi \circ \varphi_i^{-1})(\omega)$ est ouvert. Donc $\psi^{-1}(U)$ est ouvert comme intersection finie d'ouverts. □

7.2.1 Topologie quotient

DEFooHWSYooZZLXQU

Définition 7.49 ([197]).
*Soit un espace topologique X ainsi qu'une relation d'équivalence $\sim$ sur X. La **topologie quotient** sur l'ensemble $X/\sim$ est la plus petite topologie qui rend continue*[15] *la projection canonique $p\colon X \to X/\sim$.*

PROPooDRPLooONCwYs

Proposition 7.50.
Soit un espace topologique X muni d'une relation d'équivalence $\sim$. Une partie $\mathcal{O} \in X/\sim$ est ouverte[16] *si et seulement si $p^{-1}(\mathcal{O})$ est ouverte dans X.*

PROPooYKLBooQuqnfA

Proposition 7.51 ([197]).
Soient des espaces topologiques X et Y ainsi qu'une relation d'équivalence $\sim$ sur X. Soit la projection canonique $p\colon X \to X/\sim$. Une application $f\colon X/\sim \to Y$ est continue si et seulement si l'application composée $f \circ p\colon X \to Y$ est continue.

DEFooBXGJooOBQaNw

Définition 7.52 (Passage d'une application aux classes).
*Soient deux ensembles X et E ainsi qu'une relation d'équivalence $\sim$ sur X. Nous disons qu'une application $f\colon X \to E$ **passe aux classes** si f est constante sur chaque classe d'équivalence de X. Dans ce cas nous considérons l'**application quotient***

$$\begin{aligned} \tilde{f}\colon X/\sim &\to E \\ [x] &\mapsto f(x). \end{aligned} \tag{7.45}$$

LEMooKTINooKDjNeX

Lemme 7.53 ([198]).
Soient deux espaces topologiques X et Y. Soit une relation d'équivalence $\sim$ sur X. Nous considérons une application continue[17] *$f\colon X \to Y$ capable de descendre aux classes*[18]*. Alors l'application quotient $\tilde{f}\colon X/\sim \to Y$ est continue.*

Démonstration. Nous considérons la projection canonique $p\colon X \to X/\sim$. Soit un ouvert $\mathcal{O}$ dans Y. La partie $\tilde{f}^{-1}(\mathcal{O})$ sera ouverte si $p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big)$ est ouverte (proposition 7.50). Nous avons

$$p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big) = (\tilde{f} \circ p)^{-1}(\mathcal{O}). \tag{7.46}$$

Mais $\tilde{f} \circ p = f$ parce que $(\tilde{f} \circ p)(x) = \tilde{f}([x]) = f(x)$. Donc

$$p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big) = (\tilde{f} \circ p)^{-1}(\mathcal{O}) = f^{-1}(\mathcal{O}) \tag{7.47}$$

qui est ouvert parce que f est continue. □

Lemme 7.54 ([198]).
Soient deux espaces topologiques X et Y ainsi qu'une relation d'équivalence $\sim$ sur X. Nous considérons une application continue $g\colon X/\sim \to Y$. Alors

(1) L'application $g \circ p$ est continue.

(2) Si nous posons $f = g \circ p$, alors f descend aux classes et $\tilde{f} = g$.

15. Définition 7.44.
16. La topologie est définie en 7.49.
17. Définition 7.36.
18. Voir la définition 7.52.

Démonstration. La projection p est toujours continue par définition de la topologie quotient. Donc $g \circ p$ est continue par composition d'applications continues.

Voyons que f descend aux classes. Si $x \sim y$, alors $p(x) = p(y)$ et donc $f(x) = f(y)$. En ce qui concerne l'application $\tilde{f}$, nous avons

$$\tilde{f}\big([x]\big) = (g \circ p)(x) = g\big([x]\big), \tag{7.48}$$

donc $\tilde{f} = g$ et le lemme est démontré. □

7.3 Suites et convergence

7.55.
À propos de notations. La pire notation possible pour une suite est $(a_n)_n$. Mais que vient faire le second indice n ? Il peut être raisonnable d'écrire $(a_n)_{n\in I}$ lorsqu'on veut dire dans quel ensemble se déplace n. Si nous parlons de *suite*, il faut une sérieuse raison de prendre autre chose que $\mathbb{N}$ comme ensemble d'indices.

Une suite étant une fonction, de la même façon qu'on ne devrait pas dire « la fonction $f(x)$ », mais « la fonction f » ou « la fonction $x \mapsto f(x)$ », nous devrions simplement écrire a pour désigner la suite dont les éléments sont a_n.

Par conséquent, il est parfaitement légal, et même conseillé, d'écrire « $a + b$ » pour la somme des suites a et b. Et il est tout aussi légal d'écrire « $\lim a$ » au lieu de $\lim_{n\to\infty} a_n$.

Le hic est que nous écrivons souvent x la limite de la suite $n \mapsto x_n$. Dans ce cas, nous sommes évidemment obligé d'écrire l'indice n pour parler de la suite.

Tout cela pour dire qu'il faut être souple avec les notations.

7.3.1 Convergence dans un fermé

PROPooBBNSooCjrtRb

Proposition 7.56 ([1])**.**
Une suite contenue dans un fermé ne peut converger[19] *que vers un élément de ce fermé.*

Démonstration. Soient un espace topologique X et un fermé F dans X. Nous supposons que la suite (x_k) soit contenue dans F. Nous allons prouver qu'aucun élément de F^c ne peut être limite.

Soit $a \in F^c$. Puisque le complémentaire de F est un ouvert, et d'après le théorème 7.8, il existe un ouvert $\mathcal{O}_a$ contenant a, et contenu dans F^c. Le voisinage $\mathcal{O}_a$ de a ne contient donc aucun élément de la suite (x_k), qui ne peut donc pas converger vers a. □

CorLimAbarA

Corolaire 7.57.
Soit A un sous-ensemble d'un espace topologique X. Toute suite d'éléments de A qui converge, admet pour limite un élément de $\mathrm{Adh}(A)$.

Démonstration. Une fois la suite (x_n) fixée, il suffit de remarquer que tous les x_n sont dans $\mathrm{Adh}(A)$, et puis d'appliquer la proposition 7.56. □

LemPESaiVw

Lemme 7.58.
Soit $A \subset X$ muni de la topologie induite de X et (x_n) une suite dans A. Si (x_n) converge vers un élément x dans A, alors elle converge aussi vers x dans X.

Démonstration. Soit $\mathcal{O}$ un ouvert autour de x dans X. Alors $A \cap \mathcal{O}$ est un ouvert autour de x dans A et il existe $N \in \mathbb{N}$ tel que si $n \geqslant N$, alors $x_n \in A \cap \mathcal{O} \subset \mathcal{O}$. □

19. Définition 7.13.

7.3.2 Pour des limites uniques : séparabilité

Notons que l'on a parlé d'*une* limite de suite jusqu'à présent : en effet, si il existe deux éléments distincts x et y tels que tout ouvert contenant x contient y, alors la définition 7.13 dit que toute suite convergeant vers y converge aussi vers x...

EXooSHKAooZQEVLB

Exemple 7.59.
Oui, il y a moyen de converger vers plusieurs points distincts si l'espace n'est pas super cool. Nous pouvons par exemple [199] considérer la droite réelle munie de sa topologie usuelle et y ajouter un point $0'$ (qui clone le réel 0) dont les voisinages sont les voisinages de 0 dans lesquels nous remplaçons 0 par $0'$. Dans cet espace, la suite $(1/n)$ converge à la fois vers 0 et $0'$.

En fait, on « voit » le problème : on ne peut pas distinguer d'un point de vue topologique le 0 et le $0'$. △

Nous posons la définition suivante, qui nous permettra de donner une assez grande classe d'espaces topologiques dans lesquels nous avons unicité de la limite [20].

DefYFmfjjm

Définition 7.60 (Espace topologique Hausdorff)**.**
Si deux points distincts admettent toujours deux voisinages disjoints [21]*, nous disons que l'espace est* ***séparé*** *ou* ***de Hausdorff****.*

Attention, cette notion est à ne pas confondre avec :

DefUADooqilFK

Définition 7.61 (Espace topologique séparable)**.**
Un espace topologique est ***séparable*** *si il possède une partie dénombrable* [22] *dense* [23]*.*

PropUniciteLimitePourSuites

Proposition 7.62.
Dans un espace topologique séparé, si une suite converge, alors sa limite est unique.

Démonstration. Supposons que la suite (x_k) converge vers deux éléments distincts x et y. L'espace étant séparé, il existe deux ouverts $\mathcal{O}_x$ et $\mathcal{O}_y$, disjoints, contenant respectivement x et y. La suite convergeant à la fois vers x et y, il existe k_x et k_y, tels que, si $k \geqslant \max\{k_x, k_y\}$, l'élément x_k est (à la fois) dans $\mathcal{O}_x$ et $\mathcal{O}_y$. Cela est en contradiction avec le fait que ces deux ensembles sont disjoints. □

7.63.
Donc, on pourra parler, avec des espaces séparés, de « la limite d'une suite ». On notera $x_n \to a$, ou $\lim_{n\to\infty} x_n = a$, pour signifier que la suite (x_n) converge vers a.

LEMooMDTNooThlHJl

Lemme 7.64 ([1])**.**
Soit $a \neq 0$ dans un espace vectoriel topologique [24] *Hausdorff* [25]*. Il existe un voisinage V de 0 tel que $a \notin \bar{V}$.*

Démonstration. Étant donné que l'espace topologique est Hausdorff, nous pouvons considérer des voisinages V de 0 et W de a tels que $V \cap W = \varnothing$.

Dans ce cas nous avons $a \notin \bar{V}$ (voir la définition 7.20 de la fermeture de V). □

PROPooNRRIooCPesgO

Proposition 7.65 ([1])**.**
La convergence de suite pour la topologie de l'espace produit [26] *est équivalente à la convergence des suites « composante par composante ».*

20. Voir la proposition 7.113.
21. Définition 1.3.
22. Définition 1.129.
23. Définition 7.23.
24. Définition 7.170.
25. Définition 7.60
26. Définition 7.15.

Démonstration. En deux parties

(1) **Sens direct** Pour simplifier les notations, nous allons considérer le produit de deux espaces. Soit donc $(x_k, y_k) \xrightarrow{X\times Y} (x, y)$ et des ouverts $\mathcal{O}_1$ dans X autour de x et $\mathcal{O}_2$ autour de y dans Y.
La partie $\mathcal{O}_1 \times \mathcal{O}_2$ est ouverte dans $X \times Y$. Donc il existe K tel que $k \geqslant K$ implique $(x_k, y_k) \in \mathcal{O}_1 \times \mathcal{O}_2$.
Nous avons prouvé que pour tout ouvert $\mathcal{O}_1$ autour de x il existe K tel que $k \geqslant K$ implique $x_k \in \mathcal{O}_1$. Donc $x_k \xrightarrow{X} x$. Idem pour y.

(2) **Dans l'autre sens** Nous considérons l'espace produit[27] $X = \prod_{i=1}^{n} X_i$. Nous supposons pour chaque i, avoir une suite convergente $(x_i)_k \xrightarrow{X_i} x_i$.
Nous allons prouver que

$$\big((x_1)_k, \dots, (x_n)_k\big) \xrightarrow{X} (x_1, \dots, x_n). \tag{7.49}$$

Soit un ouvert $\mathcal{O}$ de X autour de $(x_1, \dots, x_n)$. Nous considérons des ouverts U_i de X_i tels que $x_i \in U_i$ et $U_1 \times \dots U_n \subset \mathcal{O}$.
Vu que $(x_i)_k \xrightarrow{X_i} x_i$, il existe $K_i \in \mathbb{N}$ tel que $k > K_i$ implique $(x_i)_k \in U_i$. Si $k \geqslant \max_i\{K_i\}$, alors $(x_i)_k \in U_i$ pour tout i et nous avons

$$\big((x_1)_k, \dots, (x_n)_k\big) \in U_1 \times \dots \times U_n \subset \mathcal{O}. \tag{7.50}$$

□
LEMooSJKMooKSiEGq

Lemme 7.66 ([1])**.**
Soit un espace topologique X. Soient dans X une suite (x_n) et un élément x tels que toute sous-suite de (x_n) contient une sous-suite convergente vers x. Alors $x_n \to x$.

Démonstration. Supposons que (x_n) ne converge pas vers x. Il existe alors un ouvert $\mathcal{O}$ autour de x tel que pour tout $N > 0$, il existe $n \geqslant N$ tel que x_n n'est pas dans $\mathcal{O}$.

Cela nous permet de construire une sous-suite de (x_n) composée d'éléments hors de $\mathcal{O}$. Aucune sous-suite de cette sous-suite ne peut converger vers x. □

7.3.3 Topologie induite

PROPooAHIMooYhOwUf

Proposition 7.67 ([1])**.**
Soient un espace topologique X et une partie $A \subset X$. Nous munissons A de la topologie induite depuis X.
ITEMooLXGWooYgPqoT

(1) Si X est Hausdorff[28]*, alors A est Hausdorff.*
ITEMooOQZDooGztljW

(2) Si X est à base dénombrable de topologie[29]*, alors A est à base dénombrable.*

Démonstration. En deux parties.

(1) **Pour (1)** Soient $a, b \in A$. Vu que X est Hausdorff, il existe des voisinages ouverts $\mathcal{O}_1$ de a et $\mathcal{O}_2$ de b tels que $\mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$. Les paries $A \cap \mathcal{O}_1$ et $A \cap \mathcal{O}_2$ sont des ouverts disjoints de A.

(2) **Pour (2)** Soit une base $\{B_i\}_{i\in I}$ de la topologie de X. Nous supposons que I est dénombrable. Nous démontrons que $\{B_i \cap A\}_{i\in I}$ et une base de la topologie de A.
En effet soit $a \in A$ et un ouvert S de A tel que $a \in S$. Il existe un ouvert $\mathcal{O}$ de X tel que $S = \mathcal{O} \cap A$. Vu que $\{B_i\}_{i\in I}$ est une base de la topologie de X, il existe $i \in I$ tel que $a \in B_i \subset \mathcal{O}$.
Vu que $\mathcal{O} \subset B_i$, nous avons

$$a \in S = A \cap \mathcal{O} \subset B_i \cap \mathcal{O}. \tag{7.51}$$

27. Pour les notations, ça va être le sport : $(x_i)_k$ désigne une suite dans X_i, mais x_i désigne la limite de cette suite.
28. Définition 7.60.
29. Base de topologie, proposition 7.2.

□

7.3.4 Fonctions équivalentes

DEFooWDSAooKXZsZY

Proposition-Définition 7.68 ([200]).
Soit un espace topologique X et $D \subset X$. Soient encore des fonctions $f, g : D \to \mathbb{C}$ et un point $a \in \mathrm{Adh}(D)$ [30].

Nous définissons sur $\mathrm{Fun}(D, \mathbb{C})$ la relation $f \sim g$ lorsque qu'il existe un voisinage V de a dans X et une fonction $\alpha : V \to \mathbb{R}$ telles que

(1) $\lim_{x \to a} \alpha(x) = 0$,

(2) pour tout $x \in (V \cap D) \backslash \{a\}$,

$$f(x) = \big(1 + \alpha(x)\big)g(x). \tag{7.52}$$

EQooQXKYooSDPpNq

Cette relation est une relation d'équivalence.

*Lorsque $f \sim g$, nous disons que f et g sont **équivalentes** en a.*

Démonstration. Nous devons prouver les trois conditions de la définition 1.32 de relation d'équivalence.

(1) **Réflexive** Il suffit de poser $\alpha(x) = 0$.

(2) **Symétrique** Si $f \sim g$, il existe une fonction α vérifiant ce qu'il faut telle que

$$f(x) = \big(1 + \alpha(x)\big)g(x). \tag{7.53}$$

Comme $\lim_{x \to a} \alpha(x) = 0$, il y a un voisinage de a sur lequel $|\alpha(x)| < 1$; il n'y a donc pas de problème de dénominateur en écrivant

$$g(x) = \frac{1}{1 + \alpha(x)} f(x). \tag{7.54}$$

Nous posons alors $\beta(x) = -\alpha(x)/\big(1 + \alpha(x)\big)$. Cela vérifie

$$g(x) = \big(1 + \beta(x)\big)f(x). \tag{7.55}$$

Et

$$\lim_{x \to a} \beta(x) = 0 \tag{7.56}$$

parce que $\lim_{x \to a} \big(1 + \alpha(x)\big) = 1$ et $\lim_{x \to a} -\alpha(x) = 0$.

(3) **Transitive** Soit $f \sim g$ et $g \sim h$. Sur un voisinage V de a nous avons

$$f(x) = \big(1 + \alpha(x)\big)g(x), \tag{7.57}$$

sur un voisinage W de a nous avons

$$g(x) = \big(1 + \beta(x)\big)h(x). \tag{7.58}$$

Sur le voisinage $V \cap W$ nous avons

$$f(x) = \big(1 + \beta(x) + \alpha(x) + (\alpha\beta)(x)\big)h(x). \tag{7.59}$$

Donc la fonction $\gamma(x) = \beta(x) + \alpha(x) + (\alpha\beta)(x)$ fait l'affaire.

□

Notons que la notion d'équivalence de fonctions, de même que la notion de limite, ne dépend pas des valeurs exactes atteintes par les fonctions au point.

30. Adhérence ou fermeture, c'est la même chose. Voir la définition 7.20 et le lemme 7.21.

Lemme 7.69.
Si f et g sont équivalentes en a, et si g ne s'annule pas sur un voisinage de a, alors pour tout $\epsilon > 0$, il existe r tel que

$$\frac{f(x)}{g(x)} \in B(1, \epsilon) \tag{7.60}$$

pour tout $x \in B(a, r)$.

Démonstration. Nous considérons un voisinage V de a sur lequel en même temps :

— la fonction α de la définition d'équivalence est définie,

— $|\alpha(x)| < \epsilon$ pour tout $x \in V$,

— $g(x) \neq 0$, pour tout $x \in V$.

Ensuite nous considérons $r > 0$ tel que $B(a, r) \subset V$. En divisant la condition (7.52) par $g(x)$ nous trouvons

$$\frac{f(x)}{g(x)} = 1 + \alpha(x). \tag{7.61}$$

Donc

$$|\frac{f(x)}{g(x)} - 1| = |\alpha(x)| \leqslant \epsilon, \tag{7.62}$$

ce qu'il fallait prouver. □

7.4 Connexité

L'idée de la connexité, c'est de s'assurer qu'un ensemble est « d'un seul tenant ».

DefIRKNooJJlmiD

Définition 7.70.
*Lorsque X est un espace topologique, nous disons qu'un sous-ensemble A est **non connexe** quand on peut trouver des ouverts O_1 et O_2 disjoints tels que*

$$A = (A \cap O_1) \cup (A \cap O_2), \tag{7.63}$$

EqDefunCon

*et tels que $A \cap O_1 \neq \emptyset$, et $A \cap O_2 \neq \emptyset$. Si un sous-ensemble n'est pas non-connexe, alors on dit qu'il est **connexe**.*

Une autre façon d'exprimer la condition (7.63) est de dire que A n'est pas connexe quand il est contenu dans la réunion de deux ouverts disjoints qui intersectent tous les deux A.

LEMooKXHQooAyVQsT

Lemme 7.71 ([1]).
Si C est un connexe de l'espace topologique X, alors C muni de la topologie induite de X est connexe.

PROPooBCFXooRlMvch

Proposition 7.72 ([201]).
Soit un espace topologique X. Soient $(C_i)_{i\in I}$ des connexes de X tels que $\bigcup_{i\in I} C_i \neq \emptyset$. Alors $\bigcup_{i\in I} C_i$ est connexe.

Démonstration. Nous notons $C = \bigcup_{i\in I} C_i$. Soient des ouverts disjoints A et B recouvrant C. Nous devons démontrer que soit $A \cap C$ soit $B \cap C$ est vide.

Nous notons $A_i = C_i \cap A$ et $B_i = C_i \cap B$. Les parties A_i et B_i recouvrent C_i et nous avons $A_i \cup B_i = C_i$. Le lemme 7.71 nous dit que C_i est connexe pour la topologie induite de X ; or dans cette topologie, les parties A_i et B_i sont ouvertes. Nous en déduisons que soit A_i soit B_i est vide. Mais lequel ?

Par hypothèse, $\bigcup_{i\in I} C_i$ est non vide. Soit x un élément de cette intersection. Soit $x \in A$ et alors $x \in A_i$ pour tout i ; soit $x \in B$ et alors $x \in B_i$ pour tout i.

Dans le premier cas, $x \in A_i$ pour tout i, de telle sorte que $B_i = \varnothing$ pour tout i. Nous avons alors

$$\varnothing = \bigcup_{i\in I} B_i = \bigcup_{i\in I}(C_i \cap B) = B \cap C. \tag{7.64}$$

Dans le second cas nous obtenons de même que $A \cap C = \varnothing$. □

DEFooFHXNooJGUPPn

Proposition-Définition 7.73 ([202]).
Soient un espace topologique X et un point $x \in X$.

ITEMooBZAQooNwuzaS

(1) La réunion de toutes les parties connexes de X contenant x est connexe.

(2) Cette réunion est la plus grande (au sens de la relation d'inclusion) de toutes les parties connexes de X contenant x.

*La réunion de toutes les parties connexes de X contenant x est nommée **composante connexe** de x dans X.*

Démonstration. La réunion de toutes les parties connexes contenant x est connexe par la proposition 7.72.

Nous notons C la réunion de toutes les parties connexes contenant x. Si D est un connexe contenant x, alors $D \subset C$ parce que C est une union de tous les connexes, y compris D. □

PropHSjJcIr

Proposition 7.74 ([1]).
Soit X un espace topologique. Les conditions suivantes sont équivalentes.

ITEMooXHIKooGqrgTs

(1) L'espace X est connexe.

ITEMooRTNPooADKVnw

(2) Si $X = O_1 \cup O_2$ avec O_1 et O_2 des ouverts disjoints, alors soit $O_1 = \varnothing$ soit $O_2 = \varnothing$.

ITEMooOEZYooFBNaOZ

(3) Si $X = F_1 \cup F_2$ avec F_1 et F_2 fermés disjoints dans X, alors $F_1 = \varnothing$ ou $F_2 = \varnothing$.

ITEMooNIPZooIDPmEf

(4) Si $A \subset X$ avec A ouvert et fermé en même temps, alors $A = \varnothing$ ou $A = X$.

Démonstration. En quatre parties.

(1) **(1) implique (2)** Par rapport à la définition 7.70, nous prenons la partie X de l'espace X. Supposons que O_1 et O_2 sont tout deux non vides. Dans ce cas nous avons

$$X = O_1 \cup O_2 = (O_1 \cap X) \cup (O_2 \cap X), \tag{7.65}$$

ce qui prouverait que X est non connexe. Contradiction. Un des O_i est vide.

(2) **(2) implique (3)** Soit une union disjointe de fermés $X = F_1 \cup F_2$. Puisque l'union est disjointe, nous avons $F_1 = X\backslash F_2$ et $F_2 = X\backslash F_1$, ce qui fait que F_1 et F_2 sont également ouverts. Nous en déduisons que $X = F_1 \cup F_2$ est une union disjointe d'ouverts. L'hypothèse indique que $F_1 = \varnothing$ ou $F_2 = \varnothing$.

(3) **(3) implique (4)** Soit A une partie ouverte et fermée de X. Nous supposons que A est ouvert et fermé, donc $X\backslash A$ est également ouvert et fermé : c'est la définition 7.3 d'un fermé. Nous avons évidemment l'union $X = A \cup (X\backslash A)$ qui est une union disjointe de fermés. Par hypothèse nous avons soit $A = \varnothing$ soit $X\backslash A = \varnothing$.

(4) **(4) implique (1)** Supposons que X ne soit pas connexe. Il existe donc des ouverts disjoints O_1 et O_2 tels que $X = O_1 \cup O_2$. Étant donné que $O_1 = X\backslash O_2$, la partie O_1 est fermée comme complémentaire d'ouvert. Donc O_1 est fermé et ouvert (et O_2 aussi d'ailleurs). Par hypothèse nous concluons que O_1 est soit X soit $\varnothing$.

□

Nous verrons plus tard (proposition 7.216) une autre caractérisation de la connexité basée sur la continuité des fonctions $X \to \mathbb{Z}$.

PROPooSCKNooRbewdv

Proposition 7.75 ([1]).
Soient un espace topologique X ainsi que $S \subset X$. Si $U \subset X$ est connexe[31] *et si $U \subset S \subset \bar{U}$, alors S est connexe.*

Démonstration. Supposons que S ne soit pas connexe. Il existe des ouverts disjoints A et B tels que $S \subset A \cup B$ et $S \cap A \neq \emptyset$, $S \cap B \neq \emptyset$. Nous prouvons que ces ouverts fonctionnent aussi pour prouver que U est non connexe (donc on aura une contradiction).

D'abord A et B recouvrent U parce que $U \subset S \subset A \cup B$. Ensuite prouvons que $U \cap A \neq \emptyset$. Soit $x \in S \cap A$. Vu que $x \in S \subset \bar{U}$, tout voisinage de x intersecte U (c'est la définition 7.20 de l'adhérence). De plus $x \in A$ et A est ouvert. Soit donc un voisinage V de x contenu dans A. Nous avons $V \subset A$ et $V \cap U \neq \emptyset$. Donc $A \cap U \neq \emptyset$.

Le même raisonnement tient pour B. □

PropIWIDzzH

Proposition 7.76.
Stabilité de la connexité par union.

ITEMooLVSSooTGstBz

(1) Une union quelconque de connexes ayant une intersection non vide est connexe.

(2) Pour tout $n \in \mathbb{N}, n > 0$, si $A_1, \ldots, A_n$ sont des connexes de X avec $A_i \cap A_{i+1} \neq \emptyset$, alors l'union $\bigcup_{i=1}^{n} A_i$ est connexe.

Démonstration. Point par point.

(1) Soient $\{C_i\}_{i\in I}$ un ensemble de connexes et un point p dans l'intersection : $p \in \bigcap_{i\in I} C_i$. Supposons que l'union ne soit pas connexe. Alors nous considérons A et B, deux ouverts disjoints recouvrant tous les C_i et ayant chacun une intersection non vide avec l'union.

Supposons pour fixer les idées que $p \in A$ et prenons $x \in B \cap \bigcup_{i\in I} C_i$. Il existe un $j \in I$ tel que $x \in C_j$. Avec tout cela nous avons

(1a) $C_j \subset A \cup B$ parce que $A \cup B$ recouvre tous les C_i,

(1b) $C_j \cap A \neq \emptyset$ parce que p est dans l'intersection,

(1c) $C_j \cap B \neq \emptyset$ parce que x est dans cette intersection.

Cela contredit le fait que C_j soit connexe.

(2) Pour la seconde partie nous procédons de proche en proche[32]. D'abord $A_1 \cup A_2$ est connexe par la première partie, ensuite $(A_1 \cup A_2) \cup A_3$ est connexe parce que les connexes $A_1 \cup A_2$ et A_3 ont un point d'intersection par hypothèse, et ainsi de suite.

□

PROPooCZJGooRlyEOV

Proposition 7.77 ([203]).
Soit un espace topologique X. Nous avons équivalence entre les points suivants :

(1) X est localement convexe.

(2) Pour tout ouvert U de X, les composantes connexes de U sont des ouverts de X.

(3) Les ouverts connexes forment une base des ouverts de E.

LEMooWGOCooHSoCzb

Lemme 7.78 ([204]).
Soit un espace topologique localement connexe X. Soient un fermé F de X, et D, une composante connexe de $X\backslash F$. Alors la frontière de D est dans F :

$$\partial D \subset F. \tag{7.66}$$

Démonstration. En plusieurs parties.

31. Définition 7.70.
32. Parce qu'on a la flemme de rédiger correctement une récurrence.

(1) **∂D est fermé** Par définition, $\partial D = \bar{D}\backslash D$. Utilisant 1.29(3) nous écrivons

$$X\backslash\partial D = (X\backslash\bar{D}) \cap D. \tag{7.67}$$

Vu que $\bar{D}$ est fermé, $X\backslash\bar{D}$ est ouvert. De plus D est ouvert par la proposition 7.77. Donc $X\backslash\partial D$ est ouvert.

(2) **Par l'absurde** Supposons qu'il existe $p \in \partial D\backslash F$. Vu que F est fermé, il existe $r > 0$ tel que $B(p,r) \cap F = \varnothing$. Étant donné que $p \in \partial D$, tout voisinage de p contient un point de D : il existe $x \in B(p,r) \cap D$.
Donc $B(p,r)$ et D sont des ouverts connexes qui ont une intersection non vide. La proposition 7.76 nous indique que $D \cup B(p,r)$ est un ouvert connexe strictement plus grand que D.
Cela contredit la maximalité de D en tant que composante connexe.

□

7.5 Compacité

La compacité est le thème 33.

7.5.1 Définition et notions connexes

Soit E, un sous-ensemble de $\mathbb{R}$. Nous pouvons considérer les ouverts suivants :

$$\mathcal{O}_x = B(x,1) \tag{7.68}$$

pour chaque $x \in E$. Évidemment,

$$E \subseteq \bigcup_{x\in E} \mathcal{O}_x. \tag{7.69}$$

Cette union contient en général de nombreuses redondances. Si par exemple $E = [-10,10]$, l'élément $3 \in E$ est contenu dans $\mathcal{O}_{3.5}$, $\mathcal{O}_{2.7}$ et bien d'autres. Pire : même si on enlève par exemple $\mathcal{O}_2$ de la liste des ouverts, l'union de ce qui reste continue à être tout E. La question est : *est-ce qu'on peut en enlever suffisamment pour qu'il n'en reste qu'un nombre fini ?*

Définition 7.79.
Soit E, un sous-ensemble de $\mathbb{R}$. Une collection d'ouverts $\mathcal{O}_i$ est un ***recouvrement*** *de E si $E \subseteq \bigcup_i \mathcal{O}_i$.*

DefJJVsEqs

Définition 7.80.
Une partie A d'un espace topologique est ***compacte*** *si elle vérifie la propriété de Borel-Lebesgue : pour tout recouvrement de A par des ouverts (c'est-à-dire une collection d'ouverts dont la réunion contient A) on peut extraire un recouvrement fini.*

Remarque 7.81.
Certaines sources (dont wikipédia) disent que pour être compact il faut aussi être séparé [33]. Pour ces sources, un espace qui ne vérifie que la propriété de Borel-Lebesgue est alors dit **quasi-compact**.

7.82.
La définition 7.80 en cache deux. En effet, si la partie A est l'espace topologique lui-même, cela définit un espace topologique compact. Un espace topologique est compact *en soi* lorsque de tout recouvrement par des ouverts, nous pouvons extraire un sous-recouvrement fini. Dans ce cas, si X est l'espace et si $\{A_i\}_{i\in I}$ est le recouvrement, nous avons $X = \bigcup_{i\in I} A_i$ et non une simple inclusion $X \subset \bigcup_{i\in I} A_i$.

LEMooNNHYooITNbyz

Lemme 7.83.
Si $a, b \in \mathbb{R}$, alors la partie $[a,b]$ est compacte[34] *dans $\mathbb{R}$.*

33. Définition 7.60.
34. Définition 7.80.

LEMooVYTRooKTIYdn

Lemme 7.84.
Si K est une partie compacte de l'espace topologique X, alors K est un espace topologique compact pour la topologie induite[35] *de X.*

Démonstration. Nous notons τ la topologie de X et τ_K la topologie induite de X vers K, c'est-à-dire

$$\tau_K = \{\mathcal{O} \cap K \text{ tel que } \mathcal{O} \in \tau\}. \tag{7.70}$$

Soient des ouverts $A_i \in \tau_K$ ($i \in I$ où I est un ensemble quelconque) tels que $\bigcup_i A_i = K$. Pour chaque $i \in I$, il existe un $\mathcal{O}_i \in \tau$ tel que $A_i = K \cap \mathcal{O}_i$. Nous avons

$$K = \bigcup_{i\in I}(K \cap \mathcal{O}_i) \subset \bigcup_{i\in I} \mathcal{O}_i. \tag{7.71}$$

Donc les $\mathcal{O}_i$ forment un recouvrement de K par des ouverts de X. Puisque K est une partie compacte de X, il existe un sous-ensemble fini J de I tel que

$$K \subset \bigcup_{j\in J} \mathcal{O}_j. \tag{7.72}$$

Nous avons donc aussi

$$K \subset \bigcup_{j\in J} K \cap \mathcal{O}_j = \bigcup_{j\in J} A_j. \tag{7.73}$$

Nous avons prouvé que $\{A_j\}_{j\in J}$ est un recouvrement fini de K par des ouverts de K. Donc K est un espace topologique compact. □

PROPooOXKSooEDzCRZ

Proposition 7.85 ([205])**.**
Soit K compact dans $\mathbb{C}$. Si $\mathcal{O}$ est une composante connexe[36] *de $\mathbb{C}\backslash K$, alors $\partial\mathcal{O} \subset K$.*

Démonstration. Supposons que $\partial\mathcal{O}$ n'est pas inclus dans K, et considérons $w \in \partial\mathcal{O}\backslash K$. Nous posons $A = \mathcal{O} \cup \{w\}$. Vu que $w \in \partial\mathcal{O}$, nous avons $w \in \bar{\mathcal{O}}$. Donc

$$\mathcal{O} \subset A \subset \bar{\mathcal{O}}. \tag{7.74}$$

La proposition 7.75 fait que A est connexe parce que $\mathcal{O}$ est connexe. Donc A est un connexe de $\mathbb{C}\backslash K$ contenant strictement $\mathcal{O}$. Impossible parce que $\mathcal{O}$ est une composante connexe de $\mathbb{C}\backslash K$. Contradiction.

Nous en déduisons que $\partial\mathcal{O} \subset K$. □

7.5.2 Espace localement compact

DefEIBYooAWoESf

Définition 7.86.
*Un espace topologique est **localement compact** si tout élément possède un voisinage compact.*

LEMooAXESooYvyesg

Lemme 7.87.
Si X est un espace topologique localement compact et si K est compact dans X, il existe un ouvert V tel que $K \subset V$ et $\bar{V}$ est compact.

LEMooKYMKooPxZjWN

Lemme 7.88.
Soient un espace localement compact X, un compact K et un ouvert $\mathcal{O}$ tel que $K \subset \mathcal{O}$. Il existe un ouvert relativement compact V tel que

$$K \subset V \subset \bar{V} \subset \mathcal{O}. \tag{7.75}$$

35. Définition 7.26.
36. Définition 7.73.

7.5.3 Autres compacité

DEFooTVDOooZbwOFK

Définition 7.89 (Séquentiellement compact)**.**
Nous disons qu'un espace topologique est ***séquentiellement compact*** *si toute suite admet une sous-suite convergente.*

DefFCGBooLpnSAK

Définition 7.90.
Un espace topologique est ***dénombrable à l'infini*** *si il est réunion dénombrable de compacts.*

Définition 7.91.
Une famille $\mathcal{A}$ de parties de X a la ***propriété d'intersection finie non vide*** *si tout sous-ensemble fini de $\mathcal{A}$ a une intersection non vide.*

PropXKUMiCj

Proposition 7.92.
Soient X un espace topologique et $K \subset X$. Les propriétés suivantes sont équivalentes : ItemXYmGHFai

(1) K est compact. ItemXYmGHFaii

(2) Si $\{F_i\}_{i\in I}$ est une famille de fermés telle que $\bigcap_{i\in I} F_i \cap K = \emptyset$, alors il existe une partie finie non vide A de I tel que $\bigcap_{i\in A} F_i \cap K = \emptyset$. ItemXYmGHFaiii

(3) Si $\{F_i\}_{i\in I}$ est une famille de fermés telle que pour tout choix de A fini dans I, $\bigcap_{i\in A} F_i \cap K \neq \emptyset$, alors l'intersection complète est non vide : $\bigcap_{i\in I} F_i \cap K \neq \emptyset$. ItemXYmGHFaiv

(4) Toute famille de fermés de X, à laquelle K est joint, et qui a la propriété d'intersection finie non vide, a une intersection non vide.

Démonstration. Les propriétés (3) et (2) sont équivalentes par contraposition. De plus le point (4) est une simple[37] reformulation en français de la propriété (3).

Prouvons (1) ⇒ (2). Soit $\{F_i\}_{i\in I}$ une famille de fermés tels que $K\bigcap_{i\in I} F_i = \emptyset$. Les complémentaires $\mathcal{O}_i$ de F_i dans X recouvrent K et donc on peut en extraire un sous-recouvrement fini :

$$K \subset \bigcup_{i\in A} \mathcal{O}_i \tag{7.76}$$

pour un certain sous-ensemble fini A de I. Pour ce même choix A, nous avons alors aussi

$$\bigcap_{i\in A} F_i \cap K = \emptyset. \tag{7.77}$$

L'implication (2) ⇒ (1) est la même histoire de passage aux complémentaires. □

Le théorème 7.299 est en général celui qu'on nomme « théorème des fermés emboîtés », mais le corolaire suivant en mériterait également le nom.

CORooQABLooMPSUBf

Corolaire 7.93 ([1])**.**
Soient un espace topologique compact X et une suite $(F_i)_{i\in\mathbb{N}}$ de fermés emboîtés[38] *dans X telle que*

$$\bigcap_{i\in\mathbb{N}} F_i = \emptyset. \tag{7.78}$$

Alors il existe $j_0 \in \mathbb{N}$ tel que $F_i = \emptyset$ pour tout $i \geqslant j_0$.

Démonstration. La proposition 7.92 nous dit qu'il existe une partie finie non vide J de $\mathbb{N}$ telle que $\bigcup_{j\in J} F_j = \emptyset$. Si $j_0 = \min(J)$, alors $F_j \subset F_{j_0}$ pour tout $j \in J$ et nous avons

$$\emptyset = \bigcap_{j\in J} F_j = F_{j_0}. \tag{7.79}$$

Dès que $F_{j_0} = \emptyset$, tous les suivants sont également vides. □

37. Enfin, simple... il faut remarquer que dans la formulation de (4), les intersections peuvent ne pas faire intervenir K, mais, au final, on s'en moque.
38. C'est-à-dire que $F_{i+1} \subset F_i$.

7.5.4 Quelques propriétés

LemOWVooZKndbI

Lemme 7.94.
Une partie K d'un espace topologique est compacte si et seulement si de tout recouvrement par des ouverts d'une base de topologie nous pouvons extraire un sous-recouvrement fini.

Remarquons que la partie qui est réellement à prouver est que, si « ça marche » pour des ouverts d'une base de topologie, alors « ça marche » pour tous types d'ouverts.

Démonstration. Soit K une partie d'un espace topologique et $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de K par des ouverts. Chacun des $\mathcal{O}_i$ est une union d'éléments de la base de topologie par la proposition 7.2 : disons $\mathcal{O}_i = \bigcup_{j\in J_i} A_{(i,j)}$. Soit $J = \{j = (i, j_i) | i \in I, j_i \in J_i\}$; alors nous obtenons $\bigcup_{j\in J} A_j = \bigcup_{i\in I} \mathcal{O}_i$.

Par hypothèse nous pouvons extraire un ensemble fini $J_0 \subset J$ tel que $K \subset \bigcup_{j\in J_0} A_j$. Par construction chacun des A_j est inclus dans (au moins) un des $\mathcal{O}_i$. Le choix d'un élément de I pour chacun des éléments de J_0 donne une partie finie I_0 de I telle que $K \subset \bigcup_{j\in J_0} A_j \subset \bigcup_{i\in I_0} \mathcal{O}_i$. □

Exemple 7.95 (Un compact non fermé)**.**
En général, un compact n'est pas toujours fermé. Si nous prenons par exemple un ensemble X de plus de deux points muni de la topologie grossière $\{\emptyset, X\}$. Toutes les parties de cet espace sont compactes, mais les seuls fermés sont $\{\emptyset, X\}$. Toutes les autres parties sont alors compactes et non fermées. △

LemnAeACf

Lemme 7.96 (Compacts et fermés[206])**.**
À propos de parties fermées dans un compact.

ITEMooNKAKooQoNddr

(1) Une partie fermée d'un compact est compacte.

ITEMooAZWVooLyPDeY

(2) Tout compact d'un espace topologique séparé est fermé.

Démonstration. En deux parties.

(1) **Pour (1)** Soient F fermé dans un compact K et $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de F par des ouverts. Puisque F est fermé, F^c est ouvert et $\{\mathcal{O}_i\}_{i\in I} \cup \{K\backslash F\}$ est un recouvrement de K par des ouverts. Si nous en extrayons un sous-recouvrement fini, c'est un recouvrement de F, et en supprimant éventuellement l'ouvert $K\backslash F$, ça reste un sous-recouvrement fini de F tout en étant extrait de $\{\mathcal{O}_i\}_{i\in I}$.

(2) **Pour (2)** Soient X un espace séparé et K compact dans X. Nous considérons $y \in K^c$ et, par hypothèse de séparation, pour chaque $x \in K$ nous considérons un voisinage ouvert V_x de x et un voisinage ouvert [39] W_x de y tels que $V_x \cap W_x = \emptyset$. Bien entendu les V_x forment un recouvrement de K par des ouverts dont nous pouvons extraire un sous-recouvrement fini : soit S fini dans K tel que

$$K \subset \bigcup_{x\in S} V_x. \tag{7.80}$$

L'ensemble $W = \bigcap_{x\in S} W_x$ est une intersection finie d'ouverts autour de y et est donc un ouvert autour de y.

Montrons que $W \cap K = \emptyset$. Soit $a \in K$; par définition de S, il existe $s \in S$ tel que $a \in V_s$. Par conséquent, a n'est pas dans W_s et donc pas non plus dans W.

L'ouvert W prouve que y est dans l'intérieur du complémentaire de K, et comme y est arbitraire, nous concluons que le complémentaire de K est ouvert (théorème 7.8), en d'autres termes, que K est fermé.

□

CORooSSFFooNkNmlS

Corolaire 7.97.
Dans un espace séparé, une intersection d'un compact avec un fermé est compacte.

39. Oui, la notation du voisinage peut surprendre, mais elle est quand même pratique pour ce qu'on veut en faire.

Démonstration. Soient un compact K et un fermé F. Par 7.96(2), K est fermé. La partie $K \cap F$ est donc fermée en tant qu'intersection de fermés (lemme 7.6(1)).

La partie $K \cap F$ est un fermé dans le compact K. Le lemme 7.96(1) dit alors que $K \cap F$ est compact. □

LEMooFJZDooSxYWVW

Lemme 7.98 ([1])**.**
Toute union finie de compacts est compacte.

Démonstration. Soient $(K_i)_{i=1,\dots,n}$ des compacts dans X. Si $\{\mathcal{O}_s\}_{i\in S}$ est un recouvrement de $\bigcup_{i\in I} K_i$ par des ouverts, à fortiori, ce sera un recouvrement de chacun des K_i. Pour chaque i, il existera donc une partie finie S_i de S telle que $\{\mathcal{O}_s\}_{s\in S_i}$ recouvre K_i.

L'union finie de parties finies S_i est une partie finie de S, et nous avons

$$\bigcup_i K_i \subset \bigcup_{s\in\bigcup_i S_i} \mathcal{O}_s. \tag{7.81}$$

□

PROPooQWHSooXeJOkT

Proposition 7.99 ([207])**.**
Dans un espace séparé, toute intersection de compacts est compacte.

Démonstration. Soit un espace topologie séparé X et des compacts $\{K_i\}_{i\in I}$ dans X (I est un ensemble quelconque). Chacun des K_i est fermé par le lemme 7.96(2). Donc l'intersection

$$K = \bigcap_{i\in I} K_i \tag{7.82}$$

est un fermé de X par le lemme 7.6(1). Soit i dans I. Nous avons $K \subset K_i$. Donc K est un fermé dans le compact K_i ; il est donc compact par le lemme 7.96. □

Exemple 7.100 (Intersection de compacts non compacte[207])**.**
Un exemple d'intersection de compacts qui n'est pas compacte. Vu la proposition 7.99, il va falloir chercher un espace non séparé. Soit $X = \mathbb{N} \cup \{x_1, x_2\}$ où x_1 et x_2 sont deux éléments distincts hors de $\mathbb{N}$. Nous définissons une topologie sur X en disant que les ouverts sont les parties suivantes :

— les parties de $\mathbb{N}$,
— la partie $\mathbb{N} \cup \{x_1\}$,
— la partie $\mathbb{N} \cup \{x_2\}$,
— la partie $\mathbb{N} \cup \{x_1, x_2\}$.

Nous considérons les parties $K_1 = \mathbb{N} \cup \{x_1\}$ et $K_2 = \mathbb{N} \cup \{x_2\}$.

(1) K_i **est compact** Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de K_1 par des ouverts de X. Alors il existe $i_0 \in I$ tel que $x_1 \in \mathcal{O}_{i_0}$. Vue la liste des ouverts, $\mathcal{O}_{i_0}$ est soit $\mathbb{N} \cup \{x_1\}$ soit $\mathbb{N} \cup \{x_1, x_2\}$. Dans les deux cas, $\{\mathcal{O}_{i_0}\}$ est un sous-recouvrement fini de K_1.

(2) $K_1 \cap K_2 = \mathbb{N}$ C'est immédiat parce que x_1 et x_2 sont distincts.

(3) $\mathbb{N}$ **n'est pas compact** Il peut être recouvert par les ouverts $\{\{i\}\}_{i\in\mathbb{N}}$ dont on ne peut pas extraire de sous-recouvrements finis.

△

PropGBZUooRKaOxy

Proposition 7.101.
Si V est une partie de l'espace topologique X muni de la topologie induite[40] *τ_V de celle de X, et si K est un compact de (V, τ_V) alors K est un compact de (X, τ_X).*

Démonstration. Soient $\{\mathcal{O}_\alpha\}_{\alpha\in A}$ des ouverts de X recouvrant K. Alors les ensembles $V \cap \mathcal{O}_\alpha$ recouvrent également K, mais sont des ouverts de V. Donc il en existe un sous-recouvrement fini. Soient donc $\{V \cap \mathcal{O}_i\}_{i\in I}$ recouvrant K avec I un sous-ensemble fini de A. Les ensembles $\{\mathcal{O}_i\}_{i\in I}$ recouvrent encore K et sont des ouverts de X. □

40. Définition 7.26.

Proposition 7.102 ([1]).
Soient des espaces topologiques X et Y. Nous considérons des ouverts A de X et B de Y. Soit un compact M dans $A \times B$[41]*. Il existe des compacts K et L dans A et B tels que $M \subset K \times L$.*

Démonstration. Nous considérons les « projections » de M sur A et B :

$$K = \{a \in A \text{ tel que } \exists b \in B \text{ tel que } (a,b) \in M\}, \tag{7.83}$$

et

$$L = \{b \in B \text{ tel que } \exists a \in A \text{ tel que } (a,b) \in M\}. \tag{7.84}$$

Nous avons $M \subset K \times L$; il reste à montrer que K et L sont des compacts de leurs espaces respectifs. Soit un recouvrement $\{U_i\}_{i\in I}$ de K par des ouverts de X et $\{V_j\}_{j\in J}$ de L par des ouverts de Y. Alors

$$\{U_i \times V_j\}_{\substack{i\in I\\ j\in J}} \tag{7.85}$$

est un recouvrement de M par des ouverts de $X \times Y$. Puisque M est un compact de $X \times Y$, nous pouvons en extraire un sous-recouvrement fini, c'est-à-dire I_0 fini dans I et J_0 fini dans J tels que

$$\{U_i \times V_j\}_{\substack{i\in I_0\\ j\in J_0}} \tag{7.86}$$

soit encore un recouvrement de $K \times L$. Nous prouvons à présent que $\{U_i\}_{i\in I_0}$ est un recouvrement de K, ce qui montrera que K est un compact.

Soit $a \in K$. Il existe $b \in B$ tel que $(a,b) \in M$. Donc il existe $i_0 \in I_0$ et $j_0 \in J_0$ tels que $(a,b) \in U_{i_0} \times V_{j_0}$. En particulier $a \in U_i$.

Le même raisonnement montre que $\{V_j\}_{j\in J_0}$ est un recouvrement de L. □

7.5.5 Espace relativement compact

DEFooBODRooEFhzeT

Définition 7.103 (Partie relativement compacte).
*Une partie d'un espace topologique est **relativement compact** si sa fermeture est compacte.*

LEMooJXGUooXDGIXG

Lemme 7.104 (Union finie de relativement compacts[1]).
Dans un espace séparé[42]*, une union finie de parties relativement compactes*[43] *est relativement compacte.*

Démonstration. Soit des paries relativement compactes A_i. Notre objectif est de prouver que $\overline{\bigcup_{i=1}^n A_i}$ est compact. La partie $\bigcup_{i=1}^n \bar{A}_i$ est compacte comme union finie de compacts (lemme 7.98) ; elle est donc fermée par le lemme 7.96(2) :

$$\overline{\bigcup_{i=1}^n \bar{A}_i} = \bigcup_{i=1}^n \bar{A}_i. \tag{7.87}$$

Nous avons donc

$$\overline{\bigcup_{i=1}^n A_i} \subset \overline{\bigcup_{i=1}^n \bar{A}_i} = \bigcup_{i=1}^n \bar{A}_i. \tag{7.88}$$

Autrement dit, la partie $\overline{\bigcup_{i=1}^n A_i}$ est un fermé dans un compact. Elle est donc compacte par le lemme 7.96(1). □

LEMooNIJEooMdvkuA

Lemme 7.105 ([1]).
Soit un espace topologique Hausdorff X. Nous supposons que $K \subset U$ avec K compact et U ouvert. Si $a \notin U$, alors il existe un voisinage A de a tel que $A \cap K = \varnothing$.

41. Topologie produit, définition 7.15
42. Séparé et Hausdorff sont synonymes, définition 7.60.
43. Définition 7.103.

Démonstration. Pour chaque k dans K, nous séparons k et a par des ouverts : A_k est un ouvert autour de a et B_k est un ouvert autour de k tels que $A_k \cap B_k = \varnothing$. Vu que $\{B_k\}_{k\in K}$ est un recouvrement du compact K par des ouverts, nous extrayons un sous-recouvrement fini $\{B_k\}_{k\in I}$, et nous posons $A = \bigcap_{k\in I} A_k$.

En tant qu'intersection finie d'ouverts, A est un ouvert. De plus $a \in A$. Nous avons $A \cap K = \varnothing$; en effet supposons que $s \in K$. Nous avons $s \in B_k$ pour un certain $k \in I$. Dans ce cas nous avons

$$A \cap B_k \subset A_k \cap B_k = \varnothing. \tag{7.89}$$

□

7.5.6 Compactifié d'Alexandrov

PROPooHNOZooPSzKIN

Proposition-Définition 7.106 ([208])**.**
Soit un espace topologique séparé localement compact[44] X*. Nous considérons un élément* $\omega \notin X$ *et l'ensemble* $\hat{X} = X \cup \{\omega\}$*. Nous nommons « ouverts de* $\hat{X}$ *» les parties suivantes :*

— *les ouverts de* X*,*

— *les parties de la forme* $A \cup \{\omega\}$ *avec* $X\backslash A$ *compact dans* X*.*

Alors $\hat{X}$ *est un espace topologique compact (cela justifie le nom « ouvert » donné aux parties susdéfinies).*

Démonstration. La première chose à faire est de prouver que $\hat{X}$ est bien un espace topologique (définition 7.1). Nous notons τ la topologie sur X et $\hat{\tau}$ l'ensemble des « ouverts » de $\hat{X}$. Le but est de prouver que $\hat{\tau}$ est une topologie.

(1) **L'espace lui-même** $\hat{X} \in \hat{\tau}$ parce que $\hat{X} = X \cup \{\omega\}$ et que $X\backslash X = \varnothing$ est compact.

(2) **Le vide** $\varnothing \in \tau \subset \hat{\tau}$.

(3) **Union quelconque** Soient A_i $(i \in I)$ des éléments de $\hat{\tau}$. Nous posons $I_1 = \{i \in I \text{ tel que } A_i \subset X\}$ et $I_2 = I\backslash I_1$. Nous avons

$$\bigcup_{i\in I} A_i = \big(\bigcup_{i\in I_1} A_i\big) \cup \big(\bigcup_{i\in I_2} A_i\big) = B \cup \big(\bigcup_{i\in I_2} B_i \cup \{\omega\}\big) \tag{7.90}$$

où B et les B_i sont des ouverts de X tels que $X\backslash B_i$ est compact dans X. Nous récrivons ça sous la forme

$$\bigcup_{i\in I} A_i = B \cup \big(\bigcup_{i\in I_2} B_i\big) \cup \{\omega\}. \tag{7.91}$$

La question est de savoir si

$$X\backslash\Big(B \cup \big(\bigcup_{i\in I_2} B_i\big)\Big) \tag{7.92}$$

est compact dans X. Un peu de réécriture :

$$X\backslash\Big(B \cup \big(\bigcup_{i\in I_2} B_i\big)\Big) = (X\backslash B) \cap X\backslash\big(\bigcup_{i\in I_2} B_i\big) = (X\backslash B) \cap \big(\bigcap_{i\in I_2} (X\backslash B_i)\big). \tag{7.93}$$

La partie $X\backslash B$ est fermée dans X parce que B est ouverte. La proposition 7.99 dit qu'une intersection de compacts est compacte (parce que X est séparé). Nous sommes donc en présence de l'intersection entre un compact et un fermé.

Tout compact d'un espace séparé est fermé [45]. Donc nous sommes en présence de l'intersection de deux fermés. Donc $(X\backslash B) \cap \big(\bigcap_{i\in I_2}(X\backslash B_i)\big)$ est fermé. Mais c'est contenu dans le compact $\bigcap_{i\in I_2}(X\backslash B_i)$. Fermé dans un compact, donc compact (lemme 7.96).

44. Définition 7.86.
45. Lemme 7.96(2).

(4) **Intersection finie** Nous considérons les « ouverts » $(A_i)_{i=1,\dots,n}$ de $\hat{X}$. Si ce sont tous des ouverts de X, l'intersection est un ouvert de X et on est bon.
Supposons que tous les A_i soient de la forme $A_i = B_i \cup \{\omega\}$ avec $X\backslash B_i$ compact. Alors

$$\bigcap_{i=1}^{n}(B_i \cup \{\omega\}) = \left(\bigcap_{i=1}^{n} B_i\right) \cup \{\omega\} \tag{7.94}$$

Mais le lemme 1.26 (appliqué un nombre fini de fois) donne

$$X\backslash\left(\bigcap_{i=1}^{n} B_i\right) = \bigcup_{i=1}^{n}(X\backslash B_i) \tag{7.95}$$

qui est compact en tant qu'union finie de compacts [46].
Enfin, nous supposons que les A_i sont un mélange des deux types, nous les séparons entre ceux qui sont directement des ouverts de X et les autres :

$$A_i = \begin{cases} B_i & \text{si } i \leqslant q \\ C_i \cup \{\omega\} & \text{si } q < i \leqslant n \end{cases} \tag{7.96}$$

où B_i sont ouverts et C_i sont des parties de X telles que $X\backslash C_i$ est compacte.
Nous avons

$$\bigcap_{i=1}^{n} A_i = \left(\bigcap_{i=1}^{q} B_i\right) \cap \left(\bigcap_{i=q+1}^{n} (C_i \cup \{\omega\})\right) \tag{7.97a}$$

$$= B \cap \left(\bigcap_{i=q+1}^{n} C_i\right). \tag{7.97b}$$

Justifications.

— Nous avons posé B est l'intersection des B_i.

— Vu que ω n'est pas dans B, nous pouvons l'oublier dans les $C_i \cup \{\omega\}$.

C'est le moment d'étudier $E = \bigcap_{i=q+1}^{n} C_i$. Nous avons

$$X\backslash E = X\backslash\left(\bigcap_{i=q+1}^{n} C_i\right) = \bigcup_{i=q+1}^{n}(X\backslash C_i). \tag{7.98}$$

Vu que $X\backslash C_i$ est compact, il est fermé [47]. La partie $X\backslash E$ est donc fermée comme union finie de fermés [48]. Et donc E est ouvert. Et finalement

$$\bigcap_{i=1}^{n} A_i = B \cap E \tag{7.99}$$

est un ouvert de X comme intersection d'ouverts. C'est donc aussi un ouvert de $\hat{X}$.

Nous avons fini de prouver que $(\hat{X}, \hat{\tau})$ est un espace topologique. Nous montrons à présent que $\hat{X}$ est compact.

Soit $\{A_i\}_{i\in I}$ un recouvrement de $\hat{X}$ par des ouverts. Pour au moins un $i_0 \in I$ nous avons $\omega \in A_{i_0}$. Nous posons $A_{i_0} = B \cup \{\omega\}$ avec $X\backslash B$ compact dans X.

Les ouverts $\{A_i\}_{i\in I}$ forment un recouvrement de $X\backslash B$ par des ouverts. Nous pouvons en extraire un sous-recouvrement fini :

$$X\backslash B \subset \bigcup_{i\in I_1} A_i. \tag{7.100}$$

46. Lemme 7.98.
47. Par le lemme 7.96(2) et le fait que nous considérons un espace séparé.
48. Par le lemme 7.6(2).

Nous avons alors

$$\hat{X} \subset \bigcup_{i\in I_1\cup\{i_0\}} A_i. \tag{7.101}$$

EQooJQFMooKPXLkh

Et voilà que $\hat{X}$ est recouvert par un nombre fini des A_i. Notez que (7.101) est une égalité, mais nous n'en avons pas besoin. □

7.107.
Oh bien entendu, les plus férus de questions embarrassantes demanderont, si X est l'espace considéré, où prendre ce ω ? Quel « objet » existe en-dehors de X ? Qui m'assure que X n'est pas tellement grand que tout est dedans ? Le fait est qu'il n'existe pas d'ensemble contenant tous les ensembles (c'est le corolaire 1.151). Nous pouvons donc toujours trouver un ensemble ω qui n'est pas dans X.

En ce qui concerne $\mathbb{R}$ auquel nous pouvons attacher deux infinis ($+\infty$ et $-\infty$), ce sera la définition 12.24.

Pour $\mathbb{C}$, nous donnerons une caractérisation de la limite en ∞ dans le lemme 12.85.

7.5.7 Propriété d'intersection finie

DEFooCESGooZkACqs

Définition 7.108 (Propriété d'intersection finie[209])**.**
*Soit un ensemble X. Une famille non vide $\mathcal{A}$ de parties de X a la **propriété d'intersection finie** si toutes intersection finie d'éléments de $\mathcal{A}$ est non vide.*

THOooCQSQooDuasqo

Théorème 7.109 ([209])**.**
Un espace est compact si et seulement si toute famille de parties fermées ayant la propriété d'intersection finie[49] *a une intersection non vide.*

7.6 Limite de fonction

DefYNVoWBx

Définition 7.110 (Limite d'une fonction, thème 29[210])**.**
*Soient des espaces topologiques X et Y ainsi que $\Omega \subset X$ et $a \in \mathrm{Adh}(\Omega)$. Soit une application $f : \Omega \to Y$. Nous disons que l'élément ℓ de Y est une **limite** de f en a lorsque pour tout ouvert V contenant ℓ, il existe un voisinage ouvert U de a tel que*

$$f\Big(\big(U \cap \Omega\big)\backslash\{a\}\Big) \subset V. \tag{7.102}$$

EQooXLJJooZDcOtU

Si un tel élément est unique[50]*, alors nous disons que cet élément est la **limite** de f et nous notons*

$$\lim_{x\to a} f(x) = \ell. \tag{7.103}$$

7.111.
Il aurait été tout aussi bien de définir la limite d'une fonction $f : X \to Y$ définie sur tout X, puis de considérer Ω avec la topologie induite depuis X.

Dans ce cas, nous aurions écrit (7.102) sous la forme

$$f\Big(U\backslash\{a\}\Big) \subset V \tag{7.104}$$

en disant que U est un voisinage de a, et en laissant la lectrice deviner que ici, « voisinage » signifie « voisinage au sens de la topologie induite ». Vu que nous considérons la fonction f uniquement définie sur Ω, c'est la seule interprétation possible, et il n'y aurait pas au d'ambigüité[211]. Mais bon...si ça va sans dire, ça va encore mieux en le disant.

49. Définition 7.108.
50. Rappelons que ce n'est pas toujours le cas, mais que ça l'est si l'espace topologique est séparé – définition 7.60.

Remarque 7.112.
Nous ne saurions trop insister sur le fait que la valeur de f en a n'intervient pas dans la définition de la limite de f en a. Il n'est même pas nécessaire que f soit définie en a pour que l'on puisse parler de limite de f en a. Par exemple nous avons

$$\lim_{x\to 1}\frac{x^2-1}{x-1}=2, \tag{7.105}$$

alors que la fonction n'est pas définie en $x=1$.

Plus généralement, un peu par principe, toutes les fois que la notion de limite apporte une information, le point où l'on prend la limite est spécial. Sinon on ne calculerait pas la limite, mais on regarderait directement la valeur de la fonction. Cela est typiquement le cas lorsque nous aborderons les dérivées. En effet, regardons (en faisant semblant d'anticiper) la définition (10.245). Dans la formule

$$f'(a)=\lim_{x\to a}\frac{f(x)-f(a)}{x-a}, \tag{7.106}$$

la fonction sur laquelle nous prenons la limite n'est *jamais* définie en $x=a$.

PropFObayrf

Proposition 7.113 (Unicité de la limite pour un espace séparé)**.**
Soient X un espace topologique, A une partie de X et Y un espace topologique séparé[51]*. Nous considérons une fonction $f\colon A\to Y$. Si $a\in \mathrm{Adh}(A)$, alors f admet au plus une limite en a.*

Démonstration. Soient y et y' des limites de f en a, ainsi que des voisinages V et V' de y et y'. Nous prenons également les voisinages W et W' correspondants :

$$\begin{cases} f(W\cap A)\subset V & \text{(7.107a)}\\ f(W'\cap A)\subset V'. & \text{(7.107b)}\end{cases}$$

Quitte à prendre des sous-ensembles nous pouvons supposer que W et W' sont ouverts. Il s'ensuit alors que :

— l'ensemble $W\cap W'$ est un ouvert contenant a et intersecte donc A ;

— l'ensemble $(W\cap W')\cap A$ est donc non vide ;

— et donc, $f(W\cap W'\cap A)$ est, lui aussi, non vide.

Mais

$$f(W\cap W'\cap A)\subset f(W\cap A)\subset V, \tag{7.108}$$

et

$$f(W\cap W'\cap A)\subset f(W'\cap A)\subset V', \tag{7.109}$$

d'où V et V' ont une intersection. Puisque ces ensembles sont arbitraires, nous avons prouvé que tout voisinage de y et tout voisinage de y' ont une intersection non vide ; étant donné que Y est séparé, nous devons avoir $y=y'$. □

Proposition 7.114.
À propos de séparation.

(1) Tout espace métrique est séparé.

(2) Si une suite dans un espace métrique possède une limite, alors elle est unique.

Démonstration. Si deux éléments x et y sont distincts, alors en posant $r=d(x,y)/3>0$, les boules $B(x,r)$ et $B(y,r)$ sont disjointes.

En ce qui concerne les limites, ce sont les propositions 7.62 et 7.113. □

51. Définition 7.60.

7.7 Topologie, distances et normes

Certains ensembles ont plus de structures qu'une topologie. Nous fixons quelques bases maintenant, et nous détaillerons certains résultats plus tard.

7.7.1 Distance et topologie métrique

DefMVNVFsX

Définition 7.115.
*Si E est un ensemble, une **distance** sur E est une application $d\colon E\times E\to\mathbb{R}$ telle que pour tout $x,y\in E$,*

(1) $d(x,y)\geqslant 0$
(2) $d(x,y)=0$ *si et seulement si* $x=y$,
(3) $d(x,y)=d(y,x)$
(4) $d(x,y)\leqslant d(x,z)+d(z,y)$.

*La dernière condition est l'**inégalité triangulaire**.*

*Un couple (E,d) formé d'un ensemble et d'une distance est un **espace métrique**.*

Le lemme suivant est similaire à la proposition 7.159.

LEMooXCXHooVtrkvl

Lemme 7.116.
Si (E,d) est un espace métrique et si $x,y,z\in E$, nous avons

$$|d(x,z)-d(y,z)|\leqslant d(x,y). \tag{7.110}$$

La définition-théorème suivante donne une topologie sur les espaces métriques en partant des boules.

ThoORdLYUu

Théorème-Définition 7.117 (Topologie métrique)**.**
*Soit (E,d) un espace métrique. Nous définissons les **boules ouvertes** par*

$$B(a,r)=\{x\in E \text{ tel que } d(a,x)<r\}. \tag{7.111}$$

EQooYCWSooIhibyd

pour tout $a\in E$ et $r>0$. Alors en posant

$$\mathcal{T}=\big\{\mathcal{O}\subset E \text{ tel que } \forall a\in\mathcal{O},\exists r>0 \text{ tel que } B(a,r)\subset\mathcal{O}\big\} \tag{7.112}$$

EqGDVVooDZfwSf

nous définissons une topologie sur E.

*Cette topologie sur E est la **topologie métrique** de (E,d). En présence d'une distance, sauf mention explicite du contraire, c'est toujours cette topologie-là que nous utiliserons.*

Démonstration. D'abord $\varnothing\in\mathcal{T}$ parce que tout élément de l'ensemble vide ...heu ...enfin parce que, d'accord hein [52]. Ensuite si les $\{A_i\}_{i\in I}$ sont des éléments de $\mathcal{T}$ et si $x\in\bigcup_{i\in I}A_i$ alors il existe $k\in I$ tel que $x\in A_k$. Par hypothèse il existe une boule $B(x,r)\subset A_k\subset\bigcup_{i\in I}A_i$.

Enfin si les $\{A_i\}_{i\in\{1,\dots,n\}}$ sont des éléments de $\mathcal{T}$ alors pour tout i il existe $r_i>0$ tel que $B(x,r_i)\subset A_i$. En prenant $r=\min\{r_i\}_{i=1,\dots,n}$ nous avons $B(x,r)\subset\bigcap_{i=1}^{n}A_i$. □

PROPooZXTXooEMLgMn

Proposition 7.118.
La topologie sur un espace métrique [53] est la topologie engendrée [54] par ses boules ouvertes.

NORMooJBMXooLHfAJK

7.119.
Si vous avez un peu de temps, vous pouvez vérifier que si $\mathbb{K}$ est un corps totalement ordonné, alors avec toutes les définitions de 1.379, en posant $d(x,y)=|x-y|$ nous avons une distance sur $\mathbb{K}$.

De plus, les boules définies en 1.379 sont alors les mêmes que celles définies en (7.111), ce qui donne à tout corps totalement ordonné une structure d'espace topologique.

52. Pour qui ne serait pas d'accord, ajoutez $\varnothing$ dans la définition des ouverts et puis c'est tout.
53. Définition 7.117.
54. Définition 7.12

PROPooUXDJooCrWBbd

Proposition 7.120 ([1]).
Soient un espace métrique (E,d), ainsi qu'une suite convergente $a_n \xrightarrow{d} \ell$. Il existe $r>0$ tel que pour tout n nous ayons $d(\ell,a_n)<r$.

Démonstration. Soit $r_1>0$ et $N\in\mathbb{N}$ tel que $n\geqslant N$ implique $d(\ell,a_n)<r_1$. Ensuite nous posons $r_2=\max\{d(\ell,a_n)\}_{n=0,\dots,N}$.

Pour tout n nous avons $d(a_n,\ell)\leqslant r_1+r_2$. □

PROPooNOHQooTqBQLk

Proposition 7.121 ([1]).
Soient un espace topologique X, un espace topologique normé[55] Y. Soit une application continue $f\colon X\to Y$. Pour tout $x\in X$, il existe un voisinage ouvert de x sur lequel f est bornée.

Démonstration. Soit $\delta>0$. Nous considérons dans Y la boule $B\big(f(x),\delta\big)$. Par continuité de f, la partie $f^{-1}\Big(B\big(f(x),\delta\big)\Big)$ est ouverte dans X. Si y est dans cette partie, alors

$$f(y)\in B\big(f(x),\delta\big). \tag{7.113}$$

L'application f y est donc bornée. □

7.7.2 Topologie métrique et induite

LEMooKDMYooMIcFRI

Lemme 7.122.
Soit un espace vectoriel normé $\big(V,\|.\|\big)$ muni d'un sous-espace vectoriel M. La topologie induite de M depuis V est la même que la topologie métrique $\big(M,\|.\|\big)$.

7.7.3 Intérieur, adhérence et frontière

7.123.
Choses déjà faites :

— Intérieur, définition 7.17.

— Adhérence, qui est la même chose que fermeture, définition 7.20, et précisé par le lemme 7.21.

Dans le cas de $\mathbb{R}^n$ dans lequel les boules forment une base de la topologie nous pouvons encore préciser de la façon suivante :

$$x\in \operatorname{Adh} A \overset{\text{def}}{\iff} \forall \epsilon>0, B(x,\epsilon)\cap A\neq\varnothing \tag{7.114}$$

Proposition 7.124.
Pour $A\subset\mathbb{R}^n$, nous avons

$$\operatorname{Int} A\subseteq A\subseteq \operatorname{Adh} A$$

DEFooACVLooRwehTl

Définition 7.125.
*La **frontière** ou le **bord** de A est défini par $\partial A=\operatorname{Adh} A\backslash \operatorname{Int} A$.*

LEMooMPZWooGrqYIX

Lemme 7.126.
Une partie A d'un espace topologique est ouverte si $A=\operatorname{Int} A$, et fermée si $A=\operatorname{Adh} A$.

LEMooEUYEooYcUfKr

Lemme 7.127 (Caractérisation équivalente de la frontière).
Soient X un espace topologique et $S\subset X$. Un point $x\in X$ est dans ∂S si et seulement si tout voisinage de x contient un point de S et un point de S^c.

55. Topologie en 7.117.

Démonstration. Supposons que tout voisinage de x contienne un point de S et un point de S^c. Alors $x \in Adh(S)$ (définition 7.20), mais pas dans l'intérieur de S parce que x ne possède pas de voisinage contenu dans S. Donc $x \in \partial S$.

À l'inverse, si $x \in \partial S$ alors x est dans l'adhérence de S et tout voisinage de x contient un point de S. Mais x n'est pas dans l'intérieur de S et tout voisinage de x contient un point qui n'est pas dans S, aka un point de S^c. □

Corolaire 7.128.
Un ensemble et son complémentaire ont même frontière.

Démonstration. Conséquence du lemme 7.127. Les points de $\partial(S^c)$ sont caractérisés par le fait que tout voisinage contient un point de S^c et un point de $(S^c)^c = S$. □

Exemple 7.129.
Soit $X = [0, 1]$ muni de la topologie de la distance $|x - y|$ (définition 7.117). Les points 0 et 1 *ne sont pas* dans la frontière de X. En effet une boule ouverte autour de 1 est un ensemble de la forme

$$B(1, r) = \{x \in X \text{ tel que } |x - 1| < r\} =]1 - r, 1] \tag{7.115}$$

où nous avons supposé $r < 1$.

Les points 0 et 1 sont par contre sur la frontière de $[0, 1]$ lorsque cet ensemble est vu comme partie de l'espace métrique $\mathbb{R}$. △

LEMooLKWEooItGnkP

Lemme 7.130 (Passage de douane[212, 213])**.**
Dans un espace topologique, toute partie connexe qui rencontre à la fois une partie A et son complémentaire rencontre nécessairement la frontière de A.

Démonstration. Nommons γ la partie connexe qui intersecte A et A^c. Les ouverts $\text{Int}(A)$ et $X\backslash\bar{A}$ ne peuvent pas recouvrir γ parce que ce sont deux ouverts disjoints alors que γ est connexe (voir la définition 7.70 de la connexité). Donc γ doit contenir des points qui sont dans $\bar{A}$ mais pas dans $\text{Int}(A)$. C'est-à-dire des points de ∂A. □

On vérifiera que les notations et les dénominations sont cohérentes en prouvant la proposition suivante.

Proposition 7.131.
Pour $\epsilon > 0$,

(1) l'adhérence de $B(x, \epsilon)$ est $\bar{B}(x, \epsilon)$,

(2) l'intérieur de $\bar{B}(x, \epsilon)$ est $B(x, \epsilon)$,

(3) la boule ouverte $B(x, \epsilon)$ est un ouvert,

(4) la boule fermée $\bar{B}(x, \epsilon)$ est un fermé.

Nous avons également les liens suivants entre intérieur, adhérence, ouvert, fermé et passage au complémentaire (noté ^{c}) :

Proposition 7.132.
Si $A \subset \mathbb{R}^n$ et $A^c = \mathbb{R}^n\backslash A$, nous avons

(1) $(\text{Int } A)^c = \text{Adh}(A^c)$ et $(\text{Adh } A)^c = \text{Int}(A^c)$,

(2) A est ouvert si et seulement si A^c est fermé,

(3) $\text{Int } A$ est le plus grand ouvert contenu dans A,

(4) $\text{Adh } A$ est le plus petit fermé contenant A,

ExBFLooUNyvbw

Exemple 7.133.
Il n'est en général pas vrai que $\overline{A \cap B} = \bar{A} \cap \bar{B}$. Par exemple si $A = [0, 1[$ et $B =]1, 2]$. Dans ce cas, $A \cap B = \varnothing$ alors que $\bar{A} \cap \bar{B} = \{1\}$. △

7.7.4 Boules ouvertes, fermées, sphères

DEFooPDSJooFcUqKH

Définition 7.134.
Soit un espace métrique (E, d).

(1) *Nous nommons* ***boule fermée*** *la fermeture de la boule ouverte, c'est-à-dire les parties de la forme $\overline{B(a,r)}$.*

(2) *La* ***sphère*** *de centre $a \in E$ et de rayon $r \in \mathbb{R}^+$ est la frontière*[56] *de la boule : $S(a,r) = \partial B(a,r)$.*

7.135.
Les différences entre boules ouverts, fermées et sphères sont très importantes. D'abord, les *boules* sont pleines tandis que la *sphère* est creuse. En comparant à une pomme, la boule ouverte serait la pomme « sans la peau », la boule fermée serait « avec la peau » tandis que la sphère serait seulement la peau.

LEMooDYYYooHZitMZ

Lemme 7.136.
Quelques liens entre les boules et les sphères.

(1) *La sphère est donnée par $S(a,r) = \{x \in V \text{ tel que } d(x,a) = r\}$.*

(2) *La fermeture de la boule est $\overline{B(a,r)} = \{x \in V \text{ tel que } d(x,a) < r\}$;*

(3) *Nous avons $\overline{B(a,r)} = B(a,r) \cup S(a,r)$.*

7.7.5 Continuité séquentielle

PropFnContParSuite

Corolaire 7.137 (Caractérisation séquentielle de la continuité en un point[1])**.**
Une application entre deux espaces topologiques continue en un point y est séquentiellement continue.

Démonstration. Soit une application $f\colon X \to Y$ entre les espaces topologiques X et Y. Nous supposons que f est continue en $a \in X$. Soit une suite convergente $x_k \xrightarrow{X} a$. Nous devons prouver que $f(x_k) \to f(a)$.

Soit un voisinage V de $f(a)$ dans Y. Le fait que f soit continue en a signifie[57] que $f(a)$ est une limite de f en a, c'est-à-dire[58] qu'il existe un voisinage W de a tel que $f(W\backslash\{a\}) \subset V$.

Puisque $x_k \to a$, il existe N tel que $x_k \in W$ pour tout $k \geqslant N$. Pour ces valeurs de k, nous avons $f(x_k) \in V$.

Nous avons prouvé que pour tout voisinage V de $f(a)$ dans Y, il existe N tel que $f(x_k) \in V$ dès que $k \geqslant N$. Cela signifie exactement que $f(x_k) \to f(a)$. □

7.7.5.1 Les boules, une base de topologie

PropNBSooraAFr

Proposition 7.138.
Un espace métrique séparable[59] *accepte une base de topologie*[60] *dénombrable.*

Soit A dense et dénombrable dans l'espace métrique séparable (E,d). Si $\{a_i\}_{i\in\mathbb{N}}$ est une énumération de A et $\{r_i\}_{i\in\mathbb{N}}$ une énumération de $\mathbb{Q}$, alors

$$\mathcal{B} = \{B(a_i, r_j)\}_{i,j\in\mathbb{N}} \tag{7.116}$$

est une base de la topologie[61] *de E.*

56. Frontière, définition 7.125.
57. C'est la définition 7.36 de la continuité en un point.
58. Définition 7.110 d'une limite.
59. Qui possède une partie dense dénombrable, définition 7.61.
60. Base de topologie, définition 7.2.
61. Définition 7.2.

Démonstration. Soient $x \in E$ et V un voisinage de x. Ce dernier contient une boule $B(x,r)$ et quitte à prendre r un peu plus petit nous supposons que $r \in \mathbb{Q}$ (existence d'un tel rationnel par le lemme 1.441).

Soit $a \in A$ avec $\|a-x\| < \frac{r}{3}$ (existe par densité de A dans E) ; nous avons $B(a, \frac{2r}{3}) \subset B(x,r)$ parce que si $y \in B(a, \frac{2r}{3})$ alors

$$\|y-x\| \leqslant \|y-a\| + \|a-x\| < \frac{2}{3}r + \frac{1}{3}r = r. \tag{7.117}$$

La seconde inégalité est stricte parce que les boules sont ouvertes. Le tout montre que $y \in B(x,r)$. Par ailleurs $x \in B(a, \frac{2}{3}r)$ et nous avons trouvé un élément de $\mathcal{B}$ contenant x tout en étant inclus dans V. Cela prouve que $\mathcal{B}$ est bien une base de la topologie de E. □

RemIPVLooHUXyeW

Remarque 7.139.
Il est vite vu que les cubes ouverts forment aussi une base de la topologie de $\mathbb{R}^n$. Cela est à mettre en rapport avec le fait que toutes les normes sont équivalentes sur $\mathbb{R}^n$ (proposition 10.49).

Voir aussi le corolaire 14.273 qui donnera tout ouvert comme union de pavés presque disjoints.

DefEnsembleBorne

Définition 7.140.
*Soit (X,d) un espace métrique. Un sous-ensemble $A \subset X$ est **borné** si il existe une boule de X contenant A.*

PROPooJIOAooWqzKMu

Proposition 7.141.
Toute réunion finie d'ensembles bornés est un ensemble borné. Toute partie d'un ensemble borné est un ensemble borné.

7.7.6 Continuité et compacité

Un résultat important dans la théorie des fonctions sur les espaces vectoriels normés est qu'une fonction continue sur un compact est bornée et atteint ses bornes. Ce résultat sera énormément utilisé pour trouver des maximums et minimums de fonctions. Le théorème exact est le suivant.

LemQFXOWyx

Lemme 7.142 (de Lebesgue[214])**.**
Soit (X,d) un espace métrique tel que toute suite ait une sous-suite convergente à l'intérieur de l'espace. Si $\{V_i\}$ est un recouvrement par des ouverts de X, alors il existe ϵ tel que pour tout $x \in X$, nous ayons $B(x,\epsilon) \subset V_i$ pour un certain i.

Démonstration. Par l'absurde, nous supposons que pour tout n, il existe un $x_n \in X$ tel que la boule $B(x_n, \frac{1}{n})$ n'est contenue dans aucun des V_i. De ces x_n, nous extrayons une sous-suite convergente (que nous nommons encore (x_n)) et nous posons $x_n \to x$. Pour n assez grand ($\frac{1}{n} < \epsilon$) nous avons $x_n \in B(x,\epsilon)$, donc tous les x_n suivants sont dans le V_i qui contient x. □

LemMGQqgDG

Lemme 7.143 ([214])**.**
Soit (X,d) un espace métrique tel que toute suite possède une sous-suite convergente. Pour tout $\epsilon > 0$, il existe un ensemble fini $\{x_i\}_{i\in I}$ tel que les boules $B(x_i,\epsilon)$ recouvrent X.

Démonstration. Soit par l'absurde un $\epsilon > 0$ contredisant le lemme. Il n'existe pas de parties finies de X autour des points desquels les boules de taille ϵ recouvrent X.

Nous construisons par récurrence une suite ne possédant pas de sous-suite convergente. Le premier terme, x_0 est pris arbitrairement dans X. Ensuite si nous avons déjà N termes de la suite, nous savons que les boules de rayon ϵ centrées sur les points $\{x_i\}_{i=1,\dots,N}$ ne recouvrent pas X. Donc nous prenons x_{N+1} hors de l'union de ces boules.

Ainsi nous avons une suite (x_n) dont tous les termes sont à distance plus grande que ϵ les uns des autres. Une telle suite ne peut pas contenir de sous-suite convergente. Contradiction. □

ThoBWFTXAZNH

Théorème 7.144 (Bolzano-Weierstrass[214], thème 33)**.**
Un espace métrique est compact si et seulement si toute suite admet une sous-suite qui converge à l'intérieur de l'espace.

Démonstration. Soient X un espace métrique compact et (x_n) une suite dans X. Nous considérons la suite de fermés emboîtés

$$X_n = \overline{\{x_k \text{ tel que } k > n\}}. \tag{7.118}$$

Ce sont des fermés ayant la propriété d'intersection finie non vide, et donc la proposition 7.92 nous dit qu'ils ont une intersection non vide. Un élément de cette intersection est automatiquement un point d'accumulation de la suite[62].

Nous passons à l'autre sens. Nous supposons que toute suite dans X contient une sous-suite convergente, et nous considérons $\{V_i\}_{i\in I}$, un recouvrement de X par des ouverts. Par le lemme 7.142, nous considérons un ϵ tel que pour tout x, il existe un $i \in I$ avec $B(x,\epsilon) \subset V_i$. Par le lemme 7.143, nous considérons un ensemble fini $\{y_i\}_{i\in A}$ tel que les boules $B(y_i,\epsilon)$ recouvrent X.

Par construction, chacune de ces boules $B(y_i,\epsilon)$ est contenue dans un des ouverts V_i. Nous sélectionnons donc parmi les V_i le nombre fini qu'il faut pour recouvrir les $B(y_i,\epsilon)$ et donc pour recouvrir X. □

ExEFYooTILPDk

Exemple 7.145 (Non compacité de la boule unité en dimension infinie)**.**
Le théorème de Bolzano-Weierstrass permet de voir tout de suite que la boule unité n'est pas compacte dans un espace vectoriel de dimension infinie : la suite des vecteurs de base ne possède pas de sous-suite convergente. △

Le théorème de Bolzano–Weierstrass 7.144 a l'importante conséquence suivante.

ThoWeirstrassRn

Théorème 7.146 (Weierstrass)**.**
Une fonction continue à valeurs réelles définie sur un compact est bornée et atteint ses bornes.

Démonstration. Soient K un compact et $f\colon K \to \mathbb{R}$ une fonction continue. Nous désignons par A l'ensemble des valeurs prises par f sur K :

$$A = f(K) = \{f(x) \text{ tel que } x \in K\}. \tag{7.119}$$

Nous considérons le supremum $M = \sup A = \sup_{x\in K} f(x)$ avec la convention suivante : si A n'est pas borné supérieurement, nous posons $M = \infty$ (voir définition 1.461).

Nous allons maintenant construire une suite (x_n) de deux façons différentes selon que $M = \infty$ ou non.

(1) Si $M = \infty$, nous choisissons, pour chaque $n \in \mathbb{N}$, un $x_n \in K$ tel que $f(x_n) > n$. C'est certainement possible parce que si A n'est pas borné, nous pouvons y trouver des nombres aussi grands que nous voulons.

(2) Si $M \neq \infty$, nous savons que pour tout ε, il existe un $y \in A$ tel que $y > M - \varepsilon$. Pour chaque n, nous choisissons donc $x_n \in K$ tel que $f(x_n) > M - \frac{1}{n}$.

Quel que soit le cas dans lequel nous sommes, la suite (x_n) est une suite dans K qui est compact, et donc nous pouvons en extraire une sous-suite convergente à l'intérieur de K par le théorème de Bolzano-Weierstrass 7.144. Afin d'alléger la notation, nous allons noter (x_n) la sous-suite convergente. Nous avons donc

$$x_n \to x \in K. \tag{7.120}$$

Par la proposition 7.137, nous savons que f prend en x la valeur

$$f(x) = \lim_{n\to\infty} f(x_n). \tag{7.121}$$

62. Définition 7.34.

Donc $f(x) < \infty$. Évidemment, si nous avions été dans le cas où $M = \infty$, la suite x_n aurait été choisie pour avoir $f(x_n) > n$ et donc il n'aurait pas été possible d'avoir $\lim_{n\to\infty} f(x_n) < \infty$. Nous en concluons que $M < \infty$, et donc que f est bornée sur K.

Afin de prouver que f atteint sa borne, c'est-à-dire que $M \in A$, nous considérons les inégalités

$$M - \frac{1}{n} < f(x_n) \leqslant M. \tag{7.122}$$

En passant à la limite $n \to \infty$, ces inégalités deviennent

$$M \leqslant f(x) \leqslant M, \tag{7.123}$$

et donc $f(x) = M$, ce qui prouve que f atteint sa borne M au point $x \in K$. □

LEMooQLVAooICaPvR

Lemme 7.147 ([1]).
Soient des compacts A, B et une fonction continue $f\colon A \times B \to \mathbb{R}$. Alors

$$\sup_{(x,y)\in A\times B} |f(x,y)| = \sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big). \tag{7.124}$$

Démonstration. Pour chaque $x \in A$, la fonction $f_x\colon B \to \mathbb{R}$ donnée par $f_x(y) = |f(x,y)|$ est continue et atteint donc sa borne [63] en $y_M(x)$. Notons que cela ne définit pas de façon univoque $y_M(x)$ parce que f_x peut atteindre son maximum en plusieurs points. L'important est que pour tout x, le nombre $|f\big(x, y_M(x)\big)|$ ne dépend pas du choix de $y_M(x)$ parmi les y qui réalisent le maximum.

Notons (x_0, y_0) un point de $A \times B$ sur lequel $|f|$ réalise son maximum [64] :

$$\sup_{(x,y)\in A\times B} |f(x,y)| = |f(x_0,y_0)|. \tag{7.125}$$

EQooDDXDooVsnlKG

Nous avons d'une part

$$\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big) = \sup_{x\in A}|f\big(x,y_M(x)\big)| \leqslant |f(x_0,y_0)| \tag{7.126}$$

Et d'autre part, quelques calculs avec justifications en-dessous :

SUBEQooPYJPooBJyEgN

$$\begin{aligned}
\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big) &\leqslant \sup_{x\in A}\sup_{y\in B}|f(x_0,y_0)| && \text{(7.127a)}\\
&= |f(x_0,y_0)| && \text{(7.127b)}\\
&\leqslant |f\big(x_0,y_M(x_0)\big)| && \text{(7.127c)}\\
&\leqslant \sup_{x\in A}|f\big(x,y_M(x)\big)| && \text{(7.127d)}\\
&\leqslant \sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big). && \text{(7.127e)}
\end{aligned}$$

SUBEQooAKPOooPdkxMJ
SUBEQooDYVPooUgQpfD
SUBEQooVOEAooNtzSpt

Justifications.

— Pour (7.127a). Le point (x_0, y_0) est un maximum de $|f|$.
— Pour (7.127c). y_M est définie pour maximiser, en fonction de x, la quantité $|f(x, y_M(x))|$.
— Pour (7.127d). Au lieu de conserver la valeur x_0 fixé, nous prenons le maximum sur tous les x possibles.

Vu que les premiers et derniers termes des inégalités (7.127) sont égaux, toutes les inégalités sont en réalité des égalités. En particulier, en reprenant (7.125),

$$\sup_{(x,y)\in A\times B} |f(x,y)| = |f(x_0,y_0)| = \sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big). \tag{7.128}$$

□

63. Théorème 7.146.
64. Encore une fois, ce point n'est pas déterminé de façon unique par cette propriété.

7.7.7 Distance à un ensemble

DEFooGNNUooFUZINs

Définition 7.148.
Si A est une partie de l'espace métrique (X,d), et si $b\in X$, nous définissons

$$d(b,A)=\inf_{y\in A} d(b,y). \tag{7.129}$$

LEMooAIARooQADaxM

Lemme 7.149 ([1]).
Si A est fermé dans (X,d), et si $b\in X$ vérifie $d(b,A)=0$, alors $b\in A$.

Démonstration. Puisque A est fermé, le complémentaire A^c est ouvert (c'est la définition 7.3). Supposons que $b\in A^c$. Alors il existe $r>0$ tel que $B(b,r)\subset A^c$. Si $a\in A$ nous avons alors $d(b,A)\geqslant d(b,a)\geqslant r>0$. Cela contredit l'hypothèse $d(b,A)=0$.

Nous en déduisons que b n'est pas dans A^c et qu'il est donc dans A. □

Exemple 7.150 (Pas avec un ouvert).
En prenant l'ouvert $A=]0,1[$ dans $\mathbb{R}$ nous avons $d(0,A)=0$, alors que 0 n'est pas dans A. △

LEMooJNRTooZyKiFC

Lemme 7.151 ([1]).
Soient un espace métrique (X,d) ainsi qu'une partie $A\subset X$. Soit $r>0$. La partie

$$\mathcal{O}=\{x\in X \text{ tel que } d(x,A)<r\} \tag{7.130}$$

est ouverte.

Démonstration. Soit $y\in\mathcal{O}$; nous avons $d(y,A)<r$. Autrement dit,

$$\inf_{a\in A} d(y,a)<r \tag{7.131}$$

et donc il existe $a\in A$ tel que $d(y,a)<r$. Soit $\delta=d(y,a)<r$. Nous montrons à présent que $B(y,r-\delta)$ est dans $\mathcal{O}$. En effet si $z\in B(y,r-\delta)$, alors

$$d(z,a)\leqslant d(z,y)+d(y,a)<r-\delta+\delta=r. \tag{7.132}$$

□

LEMooEQIZooLpsbOe

Lemme 7.152 ([1]).
Soit un fermé F de l'espace métrique (X,d). Si $a\in X$ vérifie $d(a,F)=0$, alors $a\in F$.

Démonstration. Supposons que $d(a,F)=0$, c'est-à-dire que $\inf_{x\in F} d(a,x)=0$. Il existe donc une suite (x_k) dans F telle que $d(a,x_k)\to 0$.

Cela signifie que $x_k\xrightarrow{(X,d)} a$. La proposition 7.56 nous dit alors que $a\in F$. □

LEMooCFGTooIfdcfk

Lemme 7.153 ([1]).
Si A est une partie de (X,d), alors la fonction

$$\begin{aligned} f\colon \Omega &\to [0,\infty[\\ x&\mapsto d(x,A)\end{aligned} \tag{7.133}$$

est continue.

7.7.8 Convexité

DEFooQQEOooAFKbcQ

Définition 7.154 (Partie convexe).
*Une partie A d'un espace vectoriel est **convexe** si pour tout $a,b\in A$ et pour tout $t\in[0,1]$, le point $ta+(1-t)b$ est dans A.*

Autrement dit, une partie est convexe lorsqu'elle contient tous les segments joignant ses points.

PROPooJOCEooUKhkqQ

Proposition 7.155 ([215]).
Toute intersection de convexes est convexe.

Démonstration. Soit un espace vectoriel E ainsi que des parties convexes $\{C_i\}_{i\in I}$ indexées par un ensemble quelconque I. Nous prouvons que $C = \bigcap_{i\in I} C_i$ est convexe.

Soient $x, y \in C$, ainsi que $i \in I$. Nous avons $x, y \in C_i$ et donc $\{tx + (1-t)y\}_{t\in[0,1]} \subset C_i$. Vu que cela est vrai pour tout i, nous avons

$$\{tx + (1-t)y\}_{t\in[0,1]} \subset \bigcap_{i\in I} C_i, \tag{7.134}$$

et donc le résultat attendu. □

PROPooJOXUooEZELna

Proposition 7.156 ([216]).
Si A et B sont des convexes d'un espace vectoriel topologique, alors $A - B$ est convexe.

Démonstration. Soient $a_1, a_2 \in A$ et $b_1, b_2 \in B$. Nous devons prouver que $t(a_1-b_1)+(1-t)(a_2-b_2)$ est dans $A - B$. C'est un simple calcul :

$$t(a_1 - b_1) + (1-t)(a_2 - b_2) = \underbrace{ta_1 + (1-t)a_2}_{\in A} - \underbrace{\big(tb_1 + (1-t)b_2\big)}_{\in B} \in A - B. \tag{7.135}$$

□

LEMooEFCCooOuStrb

Lemme 7.157.
L'intérieur d'un convexe est convexe.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

7.7.9 Norme

DefNorme

Définition 7.158 ([217], thème 25).
*Soit E un espace vectoriel (pas spécialement de dimension finie) sur le corps $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). Une **norme** sur E est une application $N\colon E \to [0,\infty[$ telle que*

(1) $N(x) \geqslant 0$
(2) $N(x) = 0$ *si et seulement si* $x = 0$;
(3) $N(\lambda x) = |\lambda| N(x)$
(4) $N(x+y) \leqslant N(x) + N(y)$

pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$.

ItemDefNormeii

ItemDefNormeiii

*La propriété (4) est appelée **inégalité triangulaire**.*
*Un espace vectoriel muni d'une norme est un **espace vectoriel normé**.*

En prenant $\lambda = -1$ dans la propriété (3), nous trouvons immédiatement que $N(-x) = N(x)$.

PropNmNNm

Proposition 7.159.
Toute norme N sur l'espace vectoriel E vérifie l'inégalité

$$\big|N(x) - N(y)\big| \leqslant N(x-y) \tag{7.136}$$

pour tout $x, y \in E$.

Démonstration. Nous avons, en utilisant le point (4) de la définition 7.158,

subEqNNNxxyyya
$$N(x) = N(x - y + y) \leqslant N(x-y) + N(y), \tag{7.137a}$$
subEqNNNxxyyyb
$$N(y) = N(y - x + x) \leqslant N(y-x) + N(x). \tag{7.137b}$$

Supposons d'abord que $N(x) \geqslant N(y)$. Dans ce cas, en utilisant (7.137a),

$$\big|N(x) - N(y)\big| = N(x) - N(y) \leqslant N(x-y) + N(y) - N(y) = N(x-y). \tag{7.138}$$

Si par contre $N(x) \leqslant N(y)$, alors nous utilisons (7.137b) et nous trouvons

$$\big|N(x) - N(y)\big| = N(y) - N(x) \leqslant N(y-x) + N(x) - N(x) = N(y-x) = N(x-y). \tag{7.139}$$

Dans les deux cas, nous avons retrouvé l'inégalité annoncée. □

Cette proposition signifie aussi que

$$-N(x-y) \leqslant N(x) - N(y) \leqslant N(x-y). \tag{7.140}$$

EqNleqNNleqNvqlobs

Le lemme suivant dit que nous pouvons remplacer l'inégalité triangulaire par la convexité de la boule unité dans la définition de norme.

LEMooAVIJooFhdXXr

Lemme 7.160 ([218])**.**
Soit une application $N\colon E \to [0,\infty[$ telle que

(1) $N(x) \geqslant 0$ pour tout $x \in E$,

(2) $N(x) = 0$ si et seulement si $x = 0$,

(3) $N(\lambda x) = |\lambda| N(x)$.

Alors N est une norme si et seulement si la partie

$$B = \{x \in E \text{ tel que } N(x) \leqslant 1\} \tag{7.141}$$

est convexe[65]*.*

Démonstration. Dans les deux sens.

(1) $\Rightarrow$ Nous supposons que N est une norme et nous prouvons que la boule B est convexe. Soient $x, y \in B$ et $\lambda \in [0,1]$. Nous avons

$$\begin{aligned} N\big(\lambda x + (1-\lambda)y\big) &\leqslant N(\lambda x) + N\big((1-\lambda)y\big) && \text{(7.142a)}\\ &= \lambda N(x) + (1-\lambda)N(y) && \text{(7.142b)}\\ &\leqslant \lambda + (1-\lambda) && \text{(7.142c)}\\ &= 1. && \text{(7.142d)} \end{aligned}$$

Nous avons utilisé diverses propriétés de la norme, ainsi que la majoration $N(x), N(y) \leqslant 1$.

(2) $\Leftarrow$ Nous supposons que B est convexe, et nous prouvons que N vérifie l'inégalité triangulaire. Soient $x, y \in E$ que nous choisissons tous deux non nuls (sinon c'est trop facile). Nous posons

$$z = \frac{x+y}{N(x)+N(y)} \tag{7.143}$$

EQooCIMBooFeOtWe

et la subtilité sera d'écrire z de telle sorte à être une somme de deux éléments de B. L'astuce est de poser

$$\lambda = \frac{N(x)}{N(x)+N(y)}. \tag{7.144}$$

Une simple vérification montre qu'alors

$$z = \lambda \frac{x}{N(x)} + (1-\lambda)\frac{y}{N(y)}. \tag{7.145}$$

Nous avons évidemment $x/N(x) \in B$ (et de même avec y). Puisque B est convexe, nous avons $z \in B$. Exprimons le fait que $z \in B$ à partir de la définition (7.143) :

$$\frac{N(x+y)}{N(x)+N(y)} \leqslant 1. \tag{7.146}$$

Cela signifie exactement $N(x+y) \leqslant N(x) + N(y)$.

65. Définition 7.154.

□

7.161.
Afin de suivre une notation proche de celle de la valeur absolue, à partir de maintenant, la norme d'un vecteur v sera notée $\|v\|$ au lieu de $N(v)$. La proposition 7.159 s'énoncera donc

$$\big|\|x\| - \|y\|\big| \leqslant \|x - y\|. \tag{7.147}$$

Un espace vectoriel E muni d'une norme est, on l'a déjà dit, un **espace vectoriel normé** ; on le notera $(E, \|.\|)$ pour distinguer la norme fixée.

Une autre inégalité utile de temps en temps.

CORooDFBGooAqVRfS

Corolaire 7.162.
Si a et b sont dans un espace vectoriel normé, alors

$$\big|\|a - b\| - \|b\|\big| \leqslant \|a\|. \tag{7.148}$$

Démonstration. Il s'agit seulement de la proposition 7.159 avec $x = a - b$ et $y = b$. □

LEMooWGBJooYTDYIK

Lemme-Définition 7.163 (Distance induite par une norme)**.**
Soit un espace vectoriel normé $(E, \|.\|)$. Nous posons

$$d(x, y) = \|x - y\|. \tag{7.149}$$

EQooZYJRooAHnsJG

Alors

ITEMooLITDooPeReOk

(1) d est invariante par translations : $d(a, b) = d(a + u, b + u)$

(2) d est une distance[66] *sur E.*

*C'est la **distance induite** par la norme.*

Démonstration. Le fait que la formule (7.149) soit invariante par translations est immédiat. En ce qui concerne le fait que ce soit une distance, le seul point délicat à vérifier est l'inégalité triangulaire. Mais, pour tous $x, y, z \in E$, on a

$$d(x, y) = \|x - y\| = \|x - z + z - y\| \leqslant \|x - z\| + \|z - y\| = d(x, z) + d(z, y). \tag{7.150}$$

□

DEFooPMVFooPSYVNQ

Définition 7.164.
La topologie associée à une norme est celle associée à la distance donnée en 7.163 par le théorème 7.117.

CORooMWCUooKyoyZV

Corolaire 7.165.
Un espace vectoriel normé est un espace vectoriel topologique : en d'autres mots, l'addition et la multiplication par un élément du corps sont continues.

PROPooYMCUooERvDpk

Proposition 7.166.
La norme est une application continue sur un espace vectoriel normé.

Plus précisément, si $(E, \|.\|)$ est un espace vectoriel normé, alors l'application

$$\begin{aligned} f\colon E &\to \mathbb{R} \\ x &\mapsto \|x\| \end{aligned} \tag{7.151}$$

est continue.

PROPooQUAZooGXskwF

Proposition 7.167.
La norme sur un espace vectoriel normé est une fonction de classe C^∞.

66. Définition 7.115.

LEMooGCJEooOAynZW

Lemme 7.168 ([1]).
Soient un espace vectoriel normé E ainsi qu'une partie libre $\{a_i\}$ dans E. Si nous avons $\|\sum_i \lambda_i a_i\| < M$, alors nous avons

$$|\lambda_i|\|a_i\| < M \tag{7.152}$$

pour chaque i.

LEMooSCIIooRyRrHA

Lemme 7.169.
Soient un espace vectoriel normé $(V, \|.\|)$ ainsi qu'une isométrie $f : V \to V$. Si A est une partie de V telle que $f(A) \subset A$, alors

$$\bar{A} = f(\bar{A}). \tag{7.153}$$

Nous étudierons plus en détail les espaces vectoriels topologiques à partir de la définition 7.170.

7.8 Espaces vectoriels topologiques

DefEVTopologique

Définition 7.170.
Un espace vectoriel V sur le corps valué[67] *$\mathbb{K}$ muni d'une topologie est un* ***espace vectoriel topologique*** *si*

(1) la somme de deux vecteurs est une application continue $V \times V \to V$; et

(2) la multiplication par un scalaire est une application continue $\mathbb{K} \times V \to V$.

Ici, sur $V \times V$ et sur $\mathbb{K} \times V$ nous avons la topologie produit.

Dans toute la suite, nous supposons que $\mathbb{K}$ est un corps avec une topologie métrique.

On le redit quand même : le corps[68] lui-même doit avoir sa topologie. Dans la grande majorité des cas, ce corps est $\mathbb{R}$ ou $\mathbb{C}$ muni de la topologie usuelle.

Mine de rien, le fait que les deux opérations usuelles soient continues a de belles conséquences sur la topologie de l'espace...

PROPooDXLFooFghbWk

Proposition 7.171 ([219]).
Soit un espace vectoriel topologique V. Pour $x \in V$ et $\lambda \in \mathbb{K}$, $\lambda \neq 0$ fixés, les fonctions T_x et M_λ définies par :

$$T_x : V \to V \qquad \text{et} \qquad M_\lambda : V \to V \tag{7.154}$$

$$y \mapsto x + y \qquad\qquad y \mapsto \lambda y \tag{7.155}$$

sont des automorphismes[69] *de l'espace topologique V.*

Démonstration. Ce sont des bijections continues, dont les inverses sont respectivement T_{-x} et $M_{1/\lambda}$. □

PROPooJYLVooPpLWFX

Proposition 7.172.
Soit un espace vectoriel topologique X. Si $\mathcal{O}$ est un ouvert dans X, alors pour tout $x \in X$ et pour tout $\lambda \in \mathbb{K}$, les parties $\mathcal{O} + x$ et $\lambda\mathcal{O}$ sont des ouverts.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

LEMooEZIYooBBxdJj

Lemme 7.173 ([1]).
Soit un espace vectoriel topologique V sur le corps $\mathbb{K}$. Si $\mathcal{O}$ est un ouvert autour de 0 dans V et si $\lambda \neq 0 \in \mathbb{K}$, il existe un ouvert U autour de 0 tel que $\lambda U \subset \mathcal{O}$.

67. Définition 1.475.
68. Définition 1.209
69. Définition 7.41.

Démonstration. La réponse est $U = \lambda^{-1}\mathcal{O}$. En effet par définition d'un espace vectoriel topologique, la fonction donnée par $f(x) = \lambda x$ est continue ; donc $U = f^{-1}(\mathcal{O})$ est un ouvert. De plus $\lambda U = \mathcal{O}$. □

7.8.1 Base de topologie

Toute base de voisinage de 0 se transporte en tout point de l'espace vectoriel topologique.

PropInvarianceTopologie

Corolaire 7.174 (Invariance de la topologie[219])**.**
Soit un espace vectoriel topologique (X, τ). Si $\mathcal{B}_0$ est une base de topologie en 0, alors $\mathcal{B}_0 + a$ est une base de topologie en a.

Démonstration. Soit un ouvert U contenant a. Alors $U - a$ est un ouvert contenant 0 (proposition 7.172). Il existe donc $B \in \mathcal{B}_0$ tel que $0 \in B \subset U - a$. Nous avons alors $a \in B + a \subset U$. □

7.8.1.1 Somme directe topologique

PropKZDqTR

Proposition-Définition 7.175 ([220, 221, 1])**.**
Soit V un espace vectoriel topologique et une décomposition en somme directe[70] *$V = V_1 \oplus V_2$. Alors les trois conditions suivantes sont équivalentes.*

ITEMooMUELooWdJQeW

(1) La bijection

$$\begin{aligned}\psi\colon V_1 \times V_2 &\to V\\ (x_1, x_2) &\mapsto x_1 + x_2\end{aligned} \tag{7.156}$$

est un homéomorphisme[71]*.*

ITEMooDKOYooUpEfOR

(2) Les parties V_1 et V_2 sont fermées dans V et la projection $s\colon V_2 \to V/V_1$ est un homéomorphisme.

ITEMooFSSMooCQzTIc

(3) Les parties V_1 et V_2 sont fermées dans V et la projection $\pi_2\colon V \to V_2$ est continue.

Lorsqu'une décomposition en somme directe vérifie ces conditions, nous disons que la décomposition est ***topologique****.*

Démonstration. Avant de commencer, nous rappelons les topologies.

— Sur V_1 et V_2 nous avons la topologie induite, définition 7.26.
— La topologie produit sur $V_1 \times V_2$ est la définition 7.15.
— La topologie quotient sur V/V_1 est la définition 7.49.

Voici quelques points qui ne dépendent pas des hypothèses (1), (2) ou (3).

(1) **La projection $p\colon V \to V/V_1$ est continue** La projection est toujours continue pour la topologie quotient ; c'est la définition même, voir 7.49.

(2) **ψ est une bijection continue** Le fait que ψ soit continue fait partie de la définition 7.170 d'un espace vectoriel topologique. Pour que ce soit une bijection, c'est le lemme 4.140.

(3) **s est injective** Soient $v, w \in V_2$ tels que $s(v) = s(w)$. Alors $\{v + x_1\}_{x_1 \in V_1} = \{w + y_1\}_{y_1 \in V_1}$. En particulier $v \in \{w + y_1\}_{y_1 \in V_1}$. Il existe donc $y_1 \in V_1$ tel que $v = w + y_1$. Donc $v - w \in V_1$. Mais comme V_2 est un espace vectoriel, nous avons aussi $v - w \in V_2$. Donc $v - w \in V_1 \cap V_2 = \{0\}$.

(4) **s est surjective** Soit $x \in V$. Nous devons trouver $v \in V_2$ tel que $s(v) = [x]$. Nous savons qu'il existe $x_1 \in V_1$ et $x_2 \in V_2$ tels que $x = x_1 + x_2$. Nous avons alors $[x] = s(x_2)$.

(5) **s est continue** Soit un ouvert $\mathcal{O}$ de V/V_1. La partie $p^{-1}(\mathcal{O})$ est ouverte dans V (proposition 7.50), et donc

$$s^{-1}(\mathcal{O}) = p^{-1}(\mathcal{O}) \cap V_2 \tag{7.157}$$

est ouvert dans V_2 parce que la topologie sur V_2 est celle induite[72] de la topologie de V.

70. Définition 4.139.
71. C'est à dire isomorphisme d'espaces vectoriels : bijection continue de réciproque continue. En ce qui concerne la topologie sur $V_1 \times V_2$, elle est donné par la définition 7.15.
72. Topologie induite, définition 7.26.

Et maintenant on prouve les équivalences.

(1) **Si (1) alors V_1 est fermé** La partie $V_1 \times \{0\}$ est fermée dans $V_1 \times V_2$. Vu que ψ^{-1} est continue, nous en déduisons que $\psi\big(V_1 \times \{0\}\big)$ est fermée dans V par le lemme 7.42. Mais comme $\psi\big(V_1 \times \{0\}\big) = V_1$, nous avons que V_1 est fermé dans V.

(2) **(1) ⇒ (3)** Nous supposons que $\psi\colon V_1 \times V_2 \to V$ est un homéomorphisme. Nous considérons l'application

$$\begin{aligned} \sigma\colon V_1 \times V_2 &\to V_1 \times V_2 \\ (x, y) &\mapsto (0, y). \end{aligned} \tag{7.158}$$

Cette application est continue et permet d'écrire π_2 sous la forme $\pi_2 = \psi \circ \sigma$. En tant que composée d'applications continues, l'application π_2 est continue.

(3) **(3) ⇒ (2)** L'application $\pi\colon V \to V_2$ est constante sur les classes (modulo V_1). Donc elle descend aux classes [73] en l'application

$$\begin{aligned} \tilde{\pi}\colon V/V_1 &\to V_2 \\ [x] &\mapsto \pi(x). \end{aligned} \tag{7.159}$$

Cette application est continue parce que π l'est et parce que le lemme 7.53 le dit. Le point à remarquer est que $s^{-1} = \tilde{\pi}$ parce que pour tout $x \in V_2$ nous avons

$$s\Big(\tilde{\pi}\big([x]\big)\Big) = s\big(\pi(x)\big) = s(x) = [x] \tag{7.160}$$

parce que $\pi(x) = x$ du fait que $x \in V_2$. Vu que $\tilde{\pi}$ est continue, l'application s^{-1} est également continue.

(4) **(2) ⇒ (1)** Nous pouvons écrire la projection $\pi_2\colon V \to V_2$ comme composée $\pi_2 = s^{-1} \circ p$. En effet pour $v_1 \in V_1$ et $v_2 \in V_2$ nous avons $(s^{-1} \circ p)(v_1 + v_2) = s^{-1}\big([v_2]\big) = v_2 = \pi_2(v_1 + v_2)$. Nous savon que p est continue (construction de la topologie et tout ça), et que s^{-1} est également continue par hypothèse. Donc π_2 est continue. Étant donné que $\pi_1 + \pi_2 = \mathrm{Id}$ et que l'identité est continue, nous déduisons que π_1 est également continue.

□

7.176 ([222])**.**
Si V est normé, il existe une façon plus directe (mais pas spécialement plus simple) de prouver l'implication (1) ⇒ (2), et en particulier la continuité de s^{-1}. Voyez 10.77.

Si V est de Banach, la continuité de s^{-1} peut venir du théorème d'isomorphisme de Banach 27.1 parce que s est une bijection continue entre espaces de Banach. Attention toutefois à vérifier que V/V_1 est de Banach [74].

7.8.1.2 Limite dans un espace vectoriel topologique

LEMooAHIGooJhpPvo

Lemme 7.177 (Changement de variables)**.**
Soient un espace vectoriel topologique [75] X *ainsi qu'un espace séparé* Y *et une application* $f\colon X \to Y$*. Nous supposons que* $\lim_{x\to a} f(x) = \ell$*. Alors* $\lim_{x\to b} f(x + a - b)$ *existe et vaut* ℓ*.*

Démonstration. Soit un voisinage V de ℓ dans Y. Il existe un voisinage U de a tel que $f(U) \subset V$. Nous posons $U' = U - a + b$. C'est un voisinage de b. En posant $g(x) = f(x + a - b)$ nous avons

$$g(U') = f(U - a + b + a - b) = f(U) \subset V. \tag{7.161}$$

Donc $\lim_{x\to b} g(x) = \ell$. C'est cette égalité qui signifie $\lim_{x\to b} f(x + a - b)$. □

73. Voir la définition 7.52.
74. Je ne l'ai pas fait, et au doigt mouillé je dirais que ça m'étonnerais que ce soit vrai pour tout choix de sous-espace vectoriel V_1. À vos risques et périls. Écrivez-moi si vous avez une idée.
75. Définition 7.170.

PROPooFGWXooFjvTYj

Proposition 7.178 (Limite de fonction composée[223]).
Soient des fonctions $f, g: \mathbb{R} \to \mathbb{R}$ telles que

$$\lim_{y \to l} f(y) = z \tag{7.162a}$$
$$\lim_{x \to a} g(x) = l. \tag{7.162b}$$

Nous supposons qu'il existe un intervalle ouvert I contenant l tel que $g(x) \neq l$ sur $I \backslash \{a\}$.
Alors

$$\lim_{x \to a} (f \circ g)(x) = \lim_{y \to l} f(y) = z. \tag{7.163}$$

PROPooKNVHooXlQyVA

Proposition 7.179 (Limite de fonction composée[223]).
Soient des fonctions $f, g: \mathbb{R} \to \mathbb{R}$ telles que

$$\lim_{y \to l} f(y) = z \tag{7.164a}$$
$$\lim_{x \to a} g(x) = l. \tag{7.164b}$$

Nous supposons que f est continue en l.
Alors

$$\lim_{x \to a} (f \circ g)(x) = \lim_{y \to l} f(y) = z. \tag{7.165}$$

PROPooBEHTooBrLWuh

Proposition 7.180.
Toute application linéaire entre espaces vectoriels topologiques de dimension finie est continue.

7.8.2 Corps topologique

DEFooWKLOooPdsxQl

Définition 7.181 (Anneau topologique[224]).
*Un **anneau topologique** est un anneau*[76] *muni d'une topologie dans laquelle l'addition et la multiplication sont continues*[77].

PROPooAWAKooKRmbGT

Proposition-Définition 7.182.
Si $(\mathbb{K}, |.|)$ est un corps valué[78]*, alors l'application*

$$\begin{aligned} d: \mathbb{K} \times \mathbb{K} &\to \mathbb{R}^+ \\ (x, y) &\mapsto |x - y| \end{aligned} \tag{7.166}$$

est une distance[79].
Un corps valué muni de sa topologie métrique[80] *est un corps topologique*[81].

LEMooCHDTooZsgXEK

Lemme 7.183.
Les corps $\mathbb{R}$ et $\mathbb{C}$ sont des corps valués. Leur topologie métrique (en tant que corps valués) est leur topologie usuelle.

7.8.3 Voisinage symétrique et équilibré

DEFooKBMAooAhYQAm

Définition 7.184 (Partie symétrique[219]).
*Une partie U d'un espace vectoriel topologique est **symétrique** si $x \in U$ implique $-x \in U$.*

76. Définition 1.42.
77. Définition 7.36(2).
78. Définition 1.475
79. Distance, définition 7.115.
80. Définition 7.117.
81. Définition 7.181.

DEFooPNRIooEFAlii

Définition 7.185 (Partie équilibrée[219]).
*Une partie U d'un espace vectoriel topologique V est **équilibrée** si pour tout $|\alpha| \leqslant 1$ dans $\mathbb{K}$, $\alpha U \subset U$.*

DEFooMNZJooSIAVGZ

Définition 7.186 (Partie absorbante[225]).
*Soit un espace vectoriel X. Une partie U est **absorbante** si pour tout $x \in X$, il existe $\rho_x > 0$ tel que $\lambda x \in U$ pour tout $|\lambda| \leqslant \rho_x$.*

DEFooYSWJooPkcrJe

Définition 7.187 (Partie absolument convexe).
*Une partie est **absolument convexe** si elle est convexe et équilibrée.*

LEMooXKOTooLWDXhp

Lemme 7.188.
Si A est équilibré ou absorbant, alors $0 \in A$.

Démonstration. En deux parties.

(1) **Si A est équilibré** Pour tout $|\alpha| < 1$, nous avons $\alpha A \subset A$. En particulier pour $\alpha = 0$ nous voyons que $0 \in A$.

(2) **Si A est absorbant** Si $x \in X$, alors il existe ρ tel que $\lambda x \in A$ pour tout $|\lambda| < \rho$. En particulier avec $\lambda = 0$ nous avons encore $0 \in A$.

□

LEMooEEKYooVCnBMq

Lemme 7.189 ([226]).
Dans un espace vectoriel topologique sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, tout ouvert contenant 0 est absorbant[82].

Démonstration. Soient un ouvert U contenant 0 et $x_0 \in X$. Nous considérons l'application

$$\begin{aligned} m\colon \mathbb{K} \times X &\to X \\ (t, x) &\mapsto tx \end{aligned} \tag{7.167}$$

qui est continue par définition 7.170. La partie $m^{-1}(U) \subset \mathbb{K} \times X$ est un ouvert contenant $(0, x_0)$. Par définition 7.15 de la topologie produit, il existe un $r > 0$ et un ouvert B de X contenant x_0 tels que

$$(0, x_0) \in B(0, r) \times B \subset m^{-1}(U). \tag{7.168}$$

Si $\lambda < r$, alors nous avons

$$\lambda x_0 = m(\lambda, x_0) \in m\big(B(0, r), B\big) \subset U. \tag{7.169}$$

□

LEMooKBRIooUAAPXV

Lemme 7.190.
Si A est équilibré et si $\xi \neq 0$ (dans $\mathbb{R}$ ou $\mathbb{C}$), alors

$$\xi A = |\xi| A \tag{7.170}$$

Démonstration. Vu que le corps $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, il existe $s \in \mathbb{K}$ tel que $|s| = 1$ et $\xi = s|\xi|$ (prendre $s = \xi/|\xi|$). Vu que $|s| = 1$ et que A est équilibré, $sA \subset A$ et donc $\xi A = s|\xi|A \subset |\xi|A$.

Pour l'inclusion dans l'autre sens, nous savons que $|s^{-1}| = 1$ et donc que $|\xi|A = s^{-1}\xi A \subset \xi A$. □

LEMooYSWXooNqAcOQ

Lemme 7.191 ([219, 1]).
Soit un espace vectoriel topologique.

ITEMooSWWQooTreWIE

(1) Soit un ouvert A autour de 0 dans l'espace vectoriel topologique V. Il existe $\delta > 0$ dans le corps $\mathbb{K}$ et un voisinage ouvert W de 0 tel que $\lambda W \subset A$ pour tout $|\lambda| < \delta$.

82. Définition 7.186.

ITEMooXZNHooGVplpu

(2) *Tout voisinage de* 0 *contient un ouvert équilibré.*

ITEMooRLVSooGihcLc

(3) *Tout voisinage de* 0 *contient un ouvert équilibré et symétrique.*

Démonstration. En plusieurs parties.

(1) **Pour (1)** Nous savons que $0 \cdot 0 = 0$ et que l'application

$$\begin{aligned} f\colon \mathbb{K} \times V &\to V \\ \lambda, x &\mapsto \lambda x \end{aligned} \tag{7.171}$$

est continue. La partie $f^{-1}(A)$ contient $(0,0)$. Il existe donc un voisinage ouvert S de 0 dans $\mathbb{K}$ et un voisinage ouvert W de 0 dans V tel que $S \times W \subset f^{-1}(A)$.
Puisque $\mathbb{K}$ est un corps dont la topologie est métrique [83], il existe une boule $S' = B(0,\delta) \subset S$. Donc nous avons $f(S' \times W) \subset A$ et pour tout $|\lambda| < \delta$, $\lambda W \subset A$.

(2) **Pour (2)** Soit un voisinage ouvert $\mathcal{O}$ de 0 dans V. Par le point (1), nous considérons un voisinage W de 0 et un $\delta > 0$ tel que $\lambda W \subset \mathcal{O}$ pour tout $|\lambda| < \delta$.
Nous posons

$$U = \{\lambda w < \text{ tel que } |\lambda| < \delta, w \in W\}. \tag{7.172}$$

Nous avons $U \subset \mathcal{O}$ par définition de W. De plus U est équilibré parce que si $|\mu| < 1$, et si $x \in U$, il existe $|\lambda| < \delta$ et $w \in W$ tels que $x = \lambda w$. Alors $\mu x = \mu\lambda w$. Nous avons $|\mu\lambda| < \delta$ et donc $\mu\lambda w \in U$.

(3) **Pour (3)** Nous considérons U équilibré comme dans (2). Ensuite nous posons $U' = U \cap (-U)$. La partie U' est symétrique, elle est ouverte (intersection d'ouverts). Et elle est équilibrée parce que si $x \in U'$ et $|\lambda| < 1$ alors :

— $x \in U$ et U est équilibré, donc $\lambda x \in U$.

— $x \in -U$ et U est équilibré, donc il existe $y \in U$ tel que $x = -y$. Pour ce y nous avons $\lambda y \in U$ et donc $\lambda x = -\lambda y \in -U$. Donc $\lambda x \in -U$.

— Au final, $\lambda x \in U \cap (-U) = U'$ et U' est équilibré.

□

LEMooQEFRooHAxOys

Lemme 7.192 ([227]).
Soit un espace vectoriel topologique V ainsi qu'un voisinage ouvert $\mathcal{O}$ de 0 *dans V. Il existe des voisinages ouverts U_1 et U_2 de* 0 *dans V tels que*

$$U_1 + U_2 \subset \mathcal{O}. \tag{7.173}$$

Démonstration. Par définition d'un espace vectoriel topologique, l'application

$$\begin{aligned} f\colon V \times V &\to V \\ x, y &\mapsto x + y \end{aligned} \tag{7.174}$$

est continue. Donc la partie $f^{-1}(\mathcal{O})$ est un ouvert de $V \times V$ (c'est la définition 7.36(2) de la continuité). La définition 7.15 de la topologie produit, appliquée au point $(0,0) \in V \times V$ implique qu'il existe des voisinages U_1 et U_2 de 0 dans V tels que

$$U_1 \times U_2 \subset f^{-1}(\mathcal{O}). \tag{7.175}$$

Donc $f(U_1 \times U_2) \subset \mathcal{O}$ et en particulier $U_1 + U_2 \subset \mathcal{O}$. □

PROPSommeTopologique

Proposition 7.193 ([219, 1]).
Soit V un espace vectoriel topologique, et $\mathcal{O}$ un voisinage ouvert de 0. *Il existe un voisinage ouvert U de* 0 *tel que*

83. Le corps $\mathbb{K}$ est un corps valué, et donc métrique par la définition 7.182.

(1) U *est symétrique*[84],

(2) U *est équilibré*[85]

(3) U *vérifie* $U+U\subset\mathcal{O}$.

(4) U *vérifie* $U+U+U+U\subset\mathcal{O}$.

Démonstration. En plusieurs petits pas.

(1) **Le point de départ** Le lemme 7.192 donne des voisinages ouverts U_1 et U_2 de 0 dans V tels que $U_1+U_2\subset\mathcal{O}$.

(2) **Symétrique** En posant $U'=U_1\cap U_2\cap(-U_1)\cap(-U_2)$, on a un sous-ensemble symétrique de U_1 et U_2 qui vérifie $U'+U'\subset U_1+U_2\subset\mathcal{O}$. De plus U' est encore un voisinage ouvert de 0 dans V.

(3) **équilibré** C'est le moment d'utiliser le lemme 7.191. La partie U' contient un voisinage ouvert U'' de 0 qui est symétrique et équilibré. Ce U'' vérifie encore $U''+U''\subset\mathcal{O}$.

(4) **En 4 parties** Maintenant nous ré-appliquons tout ce que nous venons de faire à U'' pour obtenir un voisinage symétrique et équilibré de 0 tel que $U+U\subset U'$. Nous avons alors $U+U+U+U\subset\mathcal{O}$.
Notons que ce U vérifie à fortiori $U+U\subset\mathcal{O}$.

□

LEMooKLOKooEfKhgN

Lemme 7.194 ([225, 228])**.**
Soit une partie B *d'un espace vectoriel topologique.*

(1) Si B *est équilibré, alors la fermeture* $\bar{B}$ *est équilibrée.*

ITEMooLZZEooAqoQVO

(2) Si B *est équilibré et si* $0\in\mathrm{Int}(B)$*, alors* $\mathrm{Int}(B)$ *est équilibré.*

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

7.8.4 Limite de suites

PROPooTJQPooMHtOAv

Proposition 7.195 ([1])**.**
Soit une suite (x_k) *dans l'espace vectoriel topologique* X*. L'élément* x *est une limite de* (x_k) *si et seulement si* 0 *est une limite de* $(x-x_k)$*.*

Démonstration. En deux parties.

(1) $\Rightarrow$ Supposons que x est une limite de (x_k). Soit un ouvert $\mathcal{O}$ autour de 0. La partie $\mathcal{O}+x$ est un ouvert autour de x (proposition 7.195). Donc il existe N tel que si $n\geqslant N$, alors $x_n\in\mathcal{O}+x$. Pour ces valeurs de n nous avons $x_n-x\in\mathcal{O}$.

(2) $\Leftarrow$ Soit un ouvert $\mathcal{O}$ autour de x. La partie $\mathcal{O}-x$ est un ouvert autour de 0. À partir de là la preuve fonctionne comme dans l'autre sens.

□

PROPooQIQSooHbNkIy

Proposition 7.196 ([1])**.**
Soient deux suites (x_k) *et* (y_k) *dans l'espace vectoriel topologique* X*. Si* x *est une limite*[86] *de* (x_k) *et si* y *est une limite de* (y_k)*, alors* $x+y$ *est une limite de* (x_k+y_k)*.*

Démonstration. En deux parties

(1) **Le cas** $x=y=0$ Nous commençons avec $x=y=0$. Soit un ouvert A autour de 0. La proposition 7.193 nous permet de considérer un voisinage U de 0 tel que $U\subset A$ et $U+U\subset A$. Vu que 0 est une limite de (x_k), il existe N tel que $x_n\in U$ pour tout $b\geqslant N$. Quitte à prendre le maximum, nous supposons aussi que $y_n\in U$ pour $n\geqslant N$. Avec ça, $x_n+y_n\in U+U\subset A$. Cela montre que 0 est une limite de (x_k+y_k).

84. Définition 7.184.
85. Définition 7.185.
86. Limite de suite, définition 7.13.

(2) **Le cas général** Nous supposons que x est une limite de (x_k) et que y est une limite de (y_k). La proposition 7.195 nous dit que 0 est une limite de $x - x_k$ et de $y - y_k$. La première partie dit alors que 0 est une limite de

$$k \mapsto x - x_k + y - y_k = x_k + y_k - (x + y). \tag{7.176}$$

La proposition 7.195, prise dans l'autre sens nous dit alors que $x + y$ est une limite de $k \mapsto x_k + y_k$.

□

7.8.5 Espace topologique localement convexe

DefPJokvAa

Définition 7.197 ([225]).
*Un espace topologique est **localement convexe** il admet une base de voisinages de l'origine dont les éléments sont des parties convexes.*

PROPooOAXHooOGhRbu

Proposition 7.198 ([225]).
Un espace vectoriel topologie localement convexe a une base de voisinage constituée d'éléments qui sont :

(1) ouverts
(2) absorbants
(3) convexes
(4) équilibrés

Démonstration. Soit un voisinage N de 0 dans l'espace vectoriel topologique X. Vu que X est localement convexe, il existe un voisinage convexe W de 0 contenu dans N. Par le lemme 7.191, W contient un ouvert équilibré U. Nous considérons enfin l'enveloppe convexe $S = \mathcal{C}\text{onv}(U)$ de U, et nous montrons que l'intérieur $\text{Int}(S)$ vérifie tout ce qu'il faut.

ouvert Parce que c'est un intérieur.

convexe L'intérieur d'un convexe est convexe par le lemme 7.157.

absorbant Tout ouvert contenant 0 est absorbant par le lemme 7.189

contient 0 La partie U contient 0 parce qu'elle est équilibrée et le lemme 7.188. Vu que U est ouvert, il contient un voisinage de tout ses points. Donc 0 est dans l'intérieur de U. Il est donc également dans l'intérieur de son enveloppe convexe.

équilibré Par le lemme (2) parce que c'est un ouvert contenant 0.

contenu dans N Nous avons les inclusions $U \subset W \subset N$. Étant donné que W est convexe, nous avons aussi $\mathcal{C}\text{onv}(U) \subset W$, et donc

$$\text{Int}(S) \subset S = \mathcal{C}\text{onv}(U) \subset W \subset N. \tag{7.177}$$

□

THOooPTLOooJLaHGE

Théorème 7.199 ([225]).
Soit un espace vectoriel topologique localement convexe[87]*. Il existe une base de voisinages $\mathcal{B}$ de 0 tels que :*

(1) les éléments sont absorbants et absolument convexes[88]*.*
(2) pour tout $A, B \in \mathcal{B}$, il existe $W \in \mathcal{B}$ tel que $W \subset A \cap B$,
(3) pour tout $A \in \mathcal{B}$, et pour tout $\rho > 0$, il existe $W \in \mathcal{B}$ tel que $W \subset \rho A$.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

87. Définition 7.197.
88. Absolument convexe : convexe et équilibrée.

Voici une réciproque du théorème 7.199.

THOooKNXGooEebnxI

Théorème 7.200 ([225]).
Soit en espace vectoriel X et une collection $\mathcal{B}$ de parties de X vérifiant

(1) les éléments sont absorbants et absolument convexes.

(2) pour tout $A, B \in \mathcal{B}$, il existe $W \in \mathcal{B}$ tel que $W \subset A \cap B$,

(3) pour tout $A \in \mathcal{B}$, et pour tout $\rho > 0$, il existe $W \in \mathcal{B}$ tel que $W \subset \rho A$.

Alors il existe une unique topologie τ sur X telle que

(1) τ est compatible avec la structure vectorielle de X,

(2) (X, τ) est localement convexe

(3) la collection $\mathcal{B}$ est une base de voisinages de 0

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

7.8.6 Limite de suites

Si (x_n) est une suite dans un espace vectoriel topologique, rien ne garantit qu'elle ait une limite, ni qu'elle soit unique. Donc lorsque nous écrivons

$$x_n \xrightarrow{V} x, \tag{7.178}$$

nous sous-entendons seulement que x est une limite.

De même, dans la proposition 7.202, nous montrerons que $x_n + y_n \xrightarrow{V} x + y$ et $\lambda x_n \xrightarrow{V} \lambda x$. Cela signifie que si x et y sont des limites de (x_n) et (y_n), alors $x + y$ est une limite de $(x_n + y_n)$ et que λx est une limite de (λx_n).

Si V est un espace vectoriel topologique dans lequel il n'y a pas unicité de la limite [89], nous ne pouvons pas exactement dire que le processus de limite est une opération linéaire sur l'ensemble des suites convergentes.

LEMooJDJVooHUKdSe

Lemme 7.201.
Soient un espace vectoriel topologique V ainsi qu'une suite (x_n) dans V. Nous avons

$$x_n \xrightarrow{V} x \tag{7.179}$$

si et seulement si

$$x_n - x \xrightarrow{V} 0. \tag{7.180}$$

Dans la proposition 7.202, nous considérons un espace vectoriel topologique. Il n'y a pas de produit. Pour la convergence

$$x_k y_k \xrightarrow{xy}, \tag{7.181}$$

voir la proposition 9.27.

PROPooZRCBooKiJhDg

Proposition 7.202 ([1]).
Soit V, un espace vectoriel topologique. Soient deux suites convergentes $x_n \xrightarrow{V} x$ et $y_n \xrightarrow{V} y$ ainsi que $\lambda \in \mathbb{K}$. Alors

ITEMooSHPAooQyEkgT

(1)

$$x_n + y_n \xrightarrow{V} x + y. \tag{7.182}$$

ITEMooYHHYooYATzWE

(2)

$$\lambda x_n \xrightarrow{V} \lambda x. \tag{7.183}$$

Démonstration. En deux parties.

89. La proposition 7.62 dit qu'il y a unicité de la limite dans les espaces topologiques séparés.

(1) **(1)** Nous allons montrer que $x_n + y_n - (x+y) \xrightarrow{V} 0$; ce sera suffisant par le lemme 7.201. Soit un ouvert $\mathcal{O}$ autour de 0. Soient des ouverts U_1 et U_2 autour de 0 tels que $U_1 + U_2 \subset \mathcal{O}$ (lemme 7.192).
Vues les convergences de (x_n) et de (y_n), il existe un N tel que $n \geqslant N$ implique $x_n - x \in U_1$ et $y_n - y \in U_2$. Dans ce cas, $x_n + y_n - (x+y) \in U_1 + U_2 \subset \mathcal{O}$.
Donc pour $n \geqslant N$ nous avons bien $x_n - y_n - (x+y) \in \mathcal{O}$, ce qui signifie que $x_n + y_n \xrightarrow{V} x+y$.

(2) **(2)** En plusieurs étapes.

(2a) $x_n - x \xrightarrow{V} 0$ C'est le lemme 7.201.

(2b) $\lambda x_n - \lambda x \xrightarrow{V} 0$ Soit un ouvert $\mathcal{O}$ autour de 0. Par le lemme 7.173, il existe un ouvert U autour de 0 tel que $\lambda U \subset \mathcal{O}$. Comme $x_n - x \xrightarrow{V} 0$, il existe N tel que $n \geqslant N$ implique $x_n - x \in U$.
Pour ces N et n nous avons aussi $\lambda(x_n - x) \in \lambda U \subset \mathcal{O}$. Nous avons donc démontré que $\lambda x_n - \lambda x \xrightarrow{V} 0$.

(2c) **Conclusion** Encore le lemme 7.201 nous permet de déduire que $\lambda x_n \xrightarrow{V} \lambda x$.

□

7.9 Applications continues

7.9.1 Continuité

La définition de la continuité d'une fonction est donnée en 7.36.

7.203.
Lorsque nous écrivons $f : X \to Y$, nous entendons que f est définie sur tout X, mais pas qu'elle soit surjective sur Y. En particulier, pour que f soit continue en a, il faut que a soit dans le domaine de définition de f.

Dans le cas de fonctions $\mathbb{R} \to \mathbb{R}$, l'espace X sera la partie de $\mathbb{R}$ sur laquelle f sera définie, et la topologie sera la topologie induite de $\mathbb{R}$.

PROPooOXBCooIzLaPe

Proposition 7.204 ([229]).
Soient deux espaces topologiques X et Y. Une application $f : X \to Y$ est continue[90] *si et seulement si pour tout $x \in X$ et pour tout voisinage*[91] *V de $f(x)$, la partie $f^{-1}(V)$ est un voisinage de x dans X.*

Démonstration. En deux parties.

(1) $\Rightarrow$ Soient $x \in X$ et un voisinage V de $f(x)$ dans Y. Il existe alors un ouvert $\mathcal{O}$ de Y tel que $f(x) \in \mathcal{O} \subset V$.
La partie $f^{-1}(\mathcal{O})$ vérifie :
- $f^{-1}(\mathcal{O})$ est un ouvert de X parce que f est continue.
- $x \in f^{-1}(\mathcal{O})$
- $f^{-1}(\mathcal{O}) \subset f^{-1}(V)$.

Donc $f^{-1}(V)$ contient un ouvert contenant x. Donc $f^{-1}(V)$ est un voisinage de x dans X.

(2) $\Leftarrow$ Soit un ouvert $\mathcal{O}$ de Y. Nous devons prouver que $f^{-1}(\mathcal{O})$ est un ouvert de X. Pour cela nous prouvons que $f^{-1}(\mathcal{O})$ contient un ouvert autour de chacun de ses éléments et utilisons le théorème 7.8.
Soit donc $x \in f^{-1}(\mathcal{O})$. La partie $\mathcal{O}$ est un voisinage de $f(x)$. Donc $f^{-1}(\mathcal{O})$ est un voisinage de x. Il existe donc un ouvert V de X tel que

$$x \in V \subset f^{-1}(\mathcal{O}). \tag{7.184}$$

90. Définition 7.36.
91. Définition 7.4

Nous en déduisons que $f^{-1}(\mathcal{O})$ contient bien un ouvert autour de chacun de ses points.

□

La proposition 7.298 donnera des détails sur ce qu'il se passe lorsque l'espace est métrique.

ThoESCaraB

Théorème 7.205.
Une fonction $f\colon X \to Y$ est une fonction continue si et seulement si elle est continue en chacun des points de X.

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que f est une fonction continue. Soient $a \in X$ et W un voisinage de $f(a)$. Nous considérons $\mathcal{O}$, un voisinage ouvert de $f(a)$ contenu dans W ; l'ensemble $f^{-1}(\mathcal{O})$ est alors un ouvert contenant a, et l'image de $f^{-1}(\mathcal{O})$ par f est bien entendu contenue dans W.

(2) $\Leftarrow$ Soit $\mathcal{O}$ un ouvert de Y. Pour prouver que $f^{-1}(\mathcal{O})$ est un ouvert de X, nous allons considérer un élément $a \in f^{-1}(\mathcal{O})$ et montrer qu'il existe un voisinage ouvert de a contenu dans $f^{-1}(\mathcal{O})$; le théorème 7.8 nous assurera alors que $f^{-1}(\mathcal{O})$ est ouvert.
L'ensemble $\mathcal{O}$ est un voisinage ouvert de $f(a)$ parce que a a été choisi dans $f^{-1}(\mathcal{O})$. Donc la continuité de f en a nous assure qu'il existe un voisinage W de a tel que $f(W) \subset \mathcal{O}$. En prenant un ouvert contenant a à l'intérieur de W nous avons un voisinage ouvert de a contenu dans $f^{-1}(\mathcal{O})$.

□

Remarque 7.206.
À cause de l'éventuelle non unicité de la limite, deux fonctions continues et égales sur un sous-ensemble dense ne sont pas spécialement égales. Ce sera vrai sur les espaces métriques et plus généralement pour les espaces séparés. Voir l'exemple 7.59 et la proposition 7.113.

LEMooCQQWooVSAWiy

Lemme 7.207 ([1])**.**
Soient une fonction $f\colon X \to Y$, et un point d'accumulation $a \in X$ [92]. La fonction f est continue en a si et seulement si $f(a)$ est une limite de f en a.

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que f est continue en $a \in X$. Soit un ouvert V de Y contenant $f(a)$. Par continuité de f au point [93] a, il existe un voisinage U de a tel que $f(U) \subset V$. À fortiori, $f\big(U\backslash a\big) \subset W$ comme le demande la définition de la limite.

(2) $\Leftarrow$ Nous supposons que $f(a)$ est une limite de $f(x)$ lorsque x tend vers a. Si W est un ouvert de Y contenant $f(a)$, il existe un voisinage V de a dans X tel que $f\big(V\backslash a\big) \subset W$. Mais puisque $f(a) \in W$, nous avons $f(V) \subset W$.

□

7.9.1.1 Continuité séquentielle

DefENioICV

Définition 7.208.
*Si X et Y sont deux espaces topologiques, une fonction $f\colon X \to \mathbb{R}$ est **séquentiellement continue** en un point a si pour toute suite convergente $x_n \to a$ dans X nous avons $f(x_n) \to f(a)$ dans Y.*

7.209.
Nous allons maintenant voir deux résultats disant que si une fonction est continue, alors elle peut être permutée avec une limite de suite. Dans le cas des espaces métriques, la proposition 7.255

92. Un point d'accumulation de X n'est pas spécialement dans X, si X est un sous-espace d'un autre. Par exemple 0 est un point d'accumulation de $]0,1[$ dans $\mathbb{R}$. Ici nous supposons que $a \in X$, sinon il n'y a de toute façon pas de continuité en a.

93. Continuité en un point, définition 7.36(1).

montrera la réciproque : si pour toute suite $x_n \to a$, nous avons $\lim_{n\to\infty} f(x_n) = y$, alors f a une limite en a qui vaut y.

fContEstSeqCont

Proposition 7.210 (Permuter limite et fonction continue[1])**.**
Soient deux espaces topologiques X et Y ainsi qu'une fonction $f\colon X \to Y$. Soit $a \in X$ et $\ell \in Y$. Si

$$\lim_{x\to a} f(x) = \ell, \tag{7.185}$$

alors, pour toute suite (x_k) telle que $x_k \to a$, on a

$$\lim f(x_k) = \ell. \tag{7.186}$$

Démonstration. Nous considérons une suite (x_k) qui converge vers a dans X. Soient V un voisinage de ℓ et W un voisinage de a tels que $f(W) \subset V$ (définition 7.110 de la continuité en un point). Par la convergence $x_k \to a$, il existe N tel que pour tout $k \geqslant N$, $x_k \in W$, et donc tel que $f(x_k) \in V$, ce qui donne la continuité séquentielle de f. □

7.9.1.2 Application réciproque

DEFooBFCQooPyKvRK

Définition 7.211 (injection, surjection, bijection)**.**
Soient des ensembles A et B ainsi qu'une application $f\colon A \to B$.

*(1) La fonction f est **injective** si $f(x_1) = f(x_2)$, implique $x_1 = x_2$.*

*(2) La fonction f est **surjective** si tous les éléments de B sont atteints, c'est-à-dire si pour tout $y \in B$ il existe $x \in A$ tel que $f(x) = y$.*

*(3) La fonction f est une **bijection** entre A et B si elle est injective et surjective, c'est-à-dire si pour tout $y \in B$ il existe un unique $x \in A$ tel que $f(x) = y$.*

La surjection et l'injection sont des propriétés bien différentes qu'il convient de prouver séparément. De plus une même « formule » peut définir une application injective, surjective, bijective ou non selon le domaine sur laquelle nous la considérons.

DEFooTRGYooRxORpY

Définition 7.212.
*Soit $f\colon A \to B$ une bijection. L'**application réciproque** de f est la fonction*

$$\begin{aligned} f^{-1}\colon B &\to A \\ y &\mapsto \text{le } x \in A \text{ tel que } f(x) = y. \end{aligned} \tag{7.187}$$

Plus généralement si $f\colon X \to Y$ est une application quelconque et si $S \subset Y$ nous notons

$$f^{-1}(S) = \{x \in X \text{ tel que } f(x) \in S\}, \tag{7.188}$$

et dans le cas où S est réduit à un unique élément y, nous notons $f^{-1}(y)$ au lieu de $f^{-1}\big(\{y\}\big)$. Si de plus $f^{-1}(S)$ est un singleton x, nous noterons $f^{-1}(S) = x$ et non $f^{-1}(S) = \{x\}$.

Les plus acharnés parmi les lecteurs se rendront compte de la différence ontologique fondamentale entre x et $\{x\}$.

7.9.2 Continuité et topologie induite

PROPooNPLBooPfmmym

Proposition 7.213 ([1])**.**
Soit une fonction $f\colon X \to Y$, continue sur l'ouvert A de X au sens où elle est continue en chaque point de A. Alors la fonction restriction $\tilde{f}\colon A \to Y$ est également continue pour la topologie sur A, induite[94] de X.

94. Définition 7.26.

Démonstration. Soit $a \in A$, et montrons que $\tilde{f}$ est continue en a, c'est-à-dire que $\tilde{f}(a) = f(a)$ soit une limite de $\tilde{f}$ en a. Soit un voisinage V de $\tilde{f}(a)$ dans Y. Par la continuité de f, nous avons un ouvert W de X tel que

$$f\big(W\backslash\{a\}\big) \subset V. \tag{7.189}$$

La partie $W \cap A$ est un voisinage de a pour la topologie de A, et vérifie

$$f\big(W \cap A\backslash\{a\}\big) \subset V. \tag{7.190}$$

donc $f(a)$ est une limite de $\tilde{f}$ pour $x \to a$. La fonction $\tilde{f}\colon A \to Y$ est continue en chaque point de A. □

Au niveau de la notion de continuité, il n'y a pas trop de changements en passant de $\mathbb{R}$ à $\mathbb{Q}$ muni de la topologie induite.

EXooHWIIooYYbfGE

Exemple 7.214.
Que signifie d'être continue pour une fonction $f\colon \mathbb{Q} \to \mathbb{R}$? D'après le théorème 7.205, il s'agit d'être continue en chaque point de $\mathbb{Q}$. Il s'agit donc, par la définition 7.36 que pour tout $q \in \mathbb{Q}$, le nombre $f(q)$ soit une limite de f pour $x \to q$.

L'espace d'arrivée étant $\mathbb{R}$, un voisinage de $f(q)$ est pris comme une boule de taille ϵ. La continuité de f exige qu'il y ait un voisinage W de q dans $\mathbb{Q}$ tel que pour tout $q' \in W$ (différent de q), $|f(q) - f(q')| < \epsilon$.

Qu'est-ce qu'un ouvert dans $\mathbb{Q}$? D'après la définition 7.26 de la topologie induite, ce sont les ensembles $\mathbb{Q} \cap \mathcal{O}$ avec $\mathcal{O}$ ouvert dans $\mathbb{R}$. Tout cela pour dire que pour tout $\epsilon > 0$, il doit exister $\delta > 0$ tel que pour tout $q' \in \mathbb{Q}$ tel que $0 < |q - q'| < \delta$, nous ayons $|f(q) - f(q')| < \epsilon$.

Bref, c'est exactement le mécanisme usuel de la continuité sur $\mathbb{R}$, sauf qu'il ne faut considérer que les rationnels. △

LEMooHAODooYSPmvH

Lemme 7.215 (Application partielle[1])**.**
Soient trois espaces topologiques X_1, X_2 et Y. Nous considérons une fonction continue $f\colon X_1 \times X_2 \to Y$ ainsi que $x_1 \in X_1$. Alors l'application

$$\begin{aligned} g\colon X_2 &\to Y \\ x_2 &\mapsto f(x_1, x_2) \end{aligned} \tag{7.191}$$

est continue.

Démonstration. Soit un ouvert $\mathcal{O}$ de Y ; par hypothèse sur f, la partie $f^{-1}(\mathcal{O})$ est ouverte dans $X_1 \times X_2$. Notre but est de prouver que $g^{-1}(\mathcal{O})$ est un ouvert de X_2. Nous avons :

$$g^{-1}(\mathcal{O}) = \{x_2 \in X_2 \text{ tel que } (x_1, x_2) \in f^{-1}(\mathcal{O})\}. \tag{7.192}$$

Nous considérons $x_2 \in g^{-1}(\mathcal{O})$ et nous prouvons qu'il existe dans X_2 un voisinage de x_2 entièrement contenu dans $g^{-1}(\mathcal{O})$.

Étant donné que (x_1, x_2) est dans $f^{-1}(\mathcal{O})$ qui est ouvert, la définition 7.15 de la topologie sur $X_1 \times X_2$ nous donne des ouverts A_1 dans X_1 et A_2 dans X_2 tels que

$$(x_1, x_2) \in A_1 \times A_2 \subset f^{-1}(\mathcal{O}). \tag{7.193}$$

Nous montrons à présent que $A_2 \subset g^{-1}(\mathcal{O})$. Soit $y_2 \in A_2$. Par construction $(x_1, y_2) \in A_1 \times A_2 \subset f^{-1}(\mathcal{O})$, donc

$$g(y_2) = f(x_1, y_2) \in \mathcal{O}. \tag{7.194}$$

Cela termine la démonstration. □

7.9.3 Continuité et connexité

PropConnexiteViaFonction

Proposition 7.216.
Un espace topologique X est connexe si et seulement si toute application continue[95] $X \to \mathbb{Z}$ *est constante.*

Démonstration. En deux parties.

(1) $\Rightarrow$ Soit une fonction continue $f\colon X \to \mathbb{Z}$. Supposons qu'elle ne soit pas constante. Nous allons en déduire que X n'est pas connexe. En effet supposons que $f(a) = u$ et $f(b) = v$ avec $u \neq v$. Nous posons

$$\begin{aligned} A &= f^{-1}(u) \\ B &= X \backslash A. \end{aligned} \tag{7.195}$$

La partie A est ouverte parce que $\{u\}$ est ouvert dans $\mathbb{Z}$. La partie B est également ouverte parce que c'est une union d'ouverts : $B = \bigcup_{n \neq u} f^{-1}(n)$. La parie A contient a, et B contient b.

Voila. Ce sont deux parties ouvertes non vides, disjointes qui recouvrent X. Donc X n'est pas connexe.

(2) $\Leftarrow$ Encore par l'absurde nous supposons que X n'est pas connexe. Soient deux ouverts A et B qui font ce qu'il faut. Alors en définissant

$$\begin{aligned} f\colon X &\to \mathbb{Z} \\ x &\mapsto \begin{cases} 0 & \text{si } x \in A \\ 1 & \text{si } x \in B, \end{cases} \end{aligned} \tag{7.196}$$

nous avons une fonction continue non constante sur X à valeurs dans $\mathbb{Z}$.

□

NORMooSCAWooPFnrVj

7.217.
Pour mettre les idées au clair, dire qu'une partie A n'est pas connexe[96] signifie qu'il existe des ouverts $\mathcal{O}_1$ et $\mathcal{O}_2$ vérifiant

ITEMooRACDooKLaVXP

(1) $\mathcal{O}_i \cap A \neq \varnothing$

ITEMooNCQVooNjAYCT

(2) $\mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$

ITEMooPIHJooNJYpQo

(3) $A \subset \mathcal{O}_1 \cup \mathcal{O}_2$.

LemConncontconn

Lemme 7.218.
L'image d'un connexe par une application continue est connexe.

Démonstration. Soit une application continue $f\colon X \to Y$ entre deux espaces topologiques.

Nous allons prouver la contraposée. Soit A une partie de Y telle que $f(A)$ ne soit pas connexe. Nous allons prouver que A elle-même n'est pas connexe. Vu que $f(A)$ n'est pas connexe, il existe des ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ recouvrant $f(A)$ et intersecant tous deux A. Nous prouvons que A n'est pas connexe en considérant les parties $A_1 = f^{-1}(\mathcal{O}_1)$ et $A_2 = f^{-1}(\mathcal{O}_2)$, et en vérifiant les propriétés de 7.217.

(1) **A_i est ouvert** Les parties A_i sont ouvertes parce qu'elles sont images inverses d'ouverts par une fonction continue (définition 7.36(2)).

(2) **Pour (1)** Nous devons prouver que $A_i \cap A \neq \varnothing$, c'est-à-dire qu'il existe un $x \in f^{-1}(\mathcal{O}_i) \cap A$. Nous commençons par considérer $y \in \mathcal{O}_i \cap f(A)$. Nous avons d'une part $f^{-1}(y) \subset f^{-1}(\mathcal{O}_i) = A_i$. D'autre part, vu que $y \in f(A)$, nous avons $f^{-1}(y) \cap A \neq \varnothing$. Nous prenons donc $x \in f^{-1}(y) \cap A$.
Ce x vérifie $x \in f^{-1}(y) \cap A \subset A_i \cap A$.

95. La topologie sur $\mathbb{Z}$ est celle de l'ensemble des parties. C'est également la topologie induite de $\mathbb{R}$, mais ça n'a aucune importance pour l'instant.
96. Connexe, définition 7.70.

(3) **Pour (2)** Si $x \in f^{-1}(\mathcal{O}_1) \cap f^{-1}(\mathcal{O}_2)$, alors $f(x) \in \mathcal{O}_1 \cap \mathcal{O}_2$, ce qui contredirait le fait que $\mathcal{O}_1$ et $\mathcal{O}_2$ sont disjoints. Il n'y a donc pas d'éléments dans l'intersection de $f^{-1}(\mathcal{O}_1)$ et de $f^{-1}(\mathcal{O}_2)$.

(4) **Pour (3)** Si $f^{-1}(\mathcal{O}_1)$ et $f^{-1}(\mathcal{O}_2)$ ne recouvrent pas A, il existe un x dans A qui n'est dans aucun des deux. Dans ce cas, $f(x)$ est dans $f(A)$, mais n'est ni dans $\mathcal{O}_1$, ni dans $\mathcal{O}_2$, ce qui contredirait le fait que ces deux derniers recouvrent $f(A)$.

Nous déduisons que A n'est pas connexe. Et donc le lemme. □

Une application de ce lemme sera le théorème des valeurs intermédiaires 9.93.

Exemple 7.219.
Les espaces topologiques $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes. △

Démonstration. Supposons par l'absurde que $f\colon \mathbb{R} \to \mathbb{R}^2$ soit un homéomorphisme. Nous posons $E = f\big(\mathbb{R}\backslash\{0\}\big)$ et $z_0 = f(0)$. Puisque f est bijective nous avons

$$E = \mathbb{R}^2\backslash\{z_0\}, \tag{7.197}$$

qui est connexe.

Comme E est connexe et que f^{-1} est continue, le lemme 7.218 nous dit que $f^{-1}(E)$ est connexe. Mais par définition, $f^{-1}(E) = \mathbb{R}\backslash\{0\}$ qui n'est pas connexe. □

7.9.4 Continuité et compacité

ThoImCompCotComp

Théorème 7.220.
L'image d'un compact[97] *par une fonction continue est un compact.*

Dans le cadre des espaces vectoriels normés, ce théorème est démontré en la proposition 7.294.

Démonstration. Soit $K \subset X$, un ensemble compact, et étudions $f(K)$; en particulier, nous considérons Ω, un recouvrement de $f(K)$ par des ouverts. Nous avons

$$f(K) \subseteq \bigcup_{\mathcal{O}\in\Omega} \mathcal{O}. \tag{7.198}$$

Par construction, nous avons aussi

$$K \subseteq \bigcup_{\mathcal{O}\in\Omega} f^{-1}(\mathcal{O}), \tag{7.199}$$

en effet, si $x \in K$, alors $f(x)$ est dans un des ouverts de Ω, disons $f(x) \in \mathcal{O}$, et évidemment, $x \in f^{-1}(\mathcal{O})$. Les $f^{-1}(\mathcal{O})$ recouvrent le compact K, et donc on peut en choisir un sous-recouvrement fini, c'est-à-dire un choix de $\{f^{-1}(\mathcal{O}_1), \dots, f^{-1}(\mathcal{O}_n)\}$ tels que

$$K \subseteq \bigcup_{i=1}^{n} f^{-1}(\mathcal{O}_i). \tag{7.200}$$

Dans ce cas, nous avons

$$f(K) \subseteq \bigcup_{i=1}^{n} \mathcal{O}_i, \tag{7.201}$$

ce qui prouve la compacité de $f(K)$. □

97. Définition 7.80.

7.9.5 Continuité de la réciproque sur un compact

LEMooNEEVooSeHYzx

Lemme 7.221.
Soit un espace compact K et un espace topologique séparé X. Si $f\colon K \to X$ est une bijection continue, alors f est un isomorphisme d'espaces topologiques.

LEMooPLGTooATIGov

Lemme 7.222 ([230]).
Soit une application continue et bijective $f\colon K \to X$ où K est compact et X est métrique. Alors la réciproque $f^{-1}\colon X \to K$ est continue.

Démonstration. Nous allons montrer que si F est fermé dans K, alors $(f^{-1})^{-1}(F)$ est fermé dans X. Le lemme 7.40 conclura. Si F est fermé dans K, alors F est compact (lemma 7.96(1)). Le théorème 7.220 dit que l'image d'un compact par une application continue est compacte. Donc $f(F)$ est compact dans X. Mais comme X est métrique, tout compact est fermé (lemme 7.96(2)). Bref, $f(F)$ est fermé. □

LEMooKSDKooDbKKeB

Lemme 7.223 ([1]).
Soit un espace vectoriel normé V ainsi qu'une application continue $f\colon [a,b] \to V$. Nous supposons que $f\colon [a,b[\to V$ est injective.

Alors en posant $S = f\big([a,b[\big)$, l'application réciproque

$$f^{-1}\colon S \to [a,b[\tag{7.202}$$

est continue.

Démonstration. Pour tout $x \in [a,b]$ nous notons $V_x = f\big([a,x]\big)$.

Par hypothèse d'injectivité, l'existence de f^{-1} sur V_b est assurée. En ce qui concerne sa continuité, pour chaque $x \in]a,b[$, l'application $f\colon [a,x] \to V_x$ vérifie le lemme 7.222. Donc l'application réciproque $f^{-1}\colon V_x \to [a,x]$ est continue.

Le théorème 7.205 dit alors que l'application f^{-1} est donc continue en chaque point de V_x pour tout $a < x < b$. Elle est donc continue en chaque point de $f\big([a,b[\big)$ parce que chacun de ces point est dans un V_x. Le théorème 7.205 (dans l'autre sens) montre alors que ${}^{-1}\colon S \to [a,b[$ est continue. □

7.9.6 Espace vectoriel topologique

PROPooHUWQooPWYBUu

Proposition 7.224 ([1]).
Soit un espace vectoriel topologique[98] *E, ainsi qu'une application linéaire f sur E. Si f est bornée sur un voisinage de 0, alors elle est continue.*

7.10 Produit fini d'espaces vectoriels normés

sec_prod

Dans cette section nous parlons de produits finis d'espaces. Cela ne signifie pas que chacun des espaces soient séparément de dimension finie.

7.10.1 Distance et norme produit

DefZTHxrHA

Proposition-Définition 7.225 (Distance produit).
Si $(E_1,d_1),\ldots,(E_n,d_n)$ sont des espaces métriques alors la formule

$$d(x,y) = \max_{i=1,\ldots,n} d_i(x_i,y_i) \tag{7.203}$$

*définit une distance sur le produit cartésien $E = E_1 \times \ldots \times E_n$. Elle est la **distance produit**.*

98. Définition 7.170.

La définition de la norme sur un produit d'espaces vectoriels normés découle immédiatement de la définition de la distance 7.225 :

DefFAJgTCE

Lemme-Définition 7.226 ([231]).
Soient V et W deux espaces vectoriels sur $\mathbb{K}$. Si (v_1, w_1) et (v_2, w_2) sont des éléments de $V \times W$ et si λ est un élément de $\mathbb{K}$, alors les opérations suivantes donnent une structure d'espace vectoriel au produit $V \times W$:

— $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$
— $\lambda(v_1, w_1) = (\lambda v_1, \lambda w_1)$.

Démonstration. Il faut seulement faire les vérifications d'usage. □

LEMooFQMSooLmdIvD

Lemme-Définition 7.227.
Soient deux espaces vectoriels normés V et W.

(1) L'opération

$$\|(v, w)\|_{V\times W} = \max\{\|v\|_V, \|w\|_W\}. \tag{7.204}$$

EqNormeVxWmax

est une norme[99] *sur $V \times W$.*

(2) La topologie de cette norme[100] *est la même que la topologie produit*[101] *sur $V \times W$.*

Démonstration. En plusieurs parties.

(1) **Norme** On doit vérifier les trois conditions de la définition 7.158.

— Soit (v, w) dans $V \times W$ tel que $\|(v, w)\|_{V\times W} = \max\{\|v\|_V, \|w\|_W\} = 0$. Alors $\|v\|_V = 0$ et $\|w\|_W = 0$, donc $v = 0_V$ et $w = 0_W$. Cela implique $(v, w) = (0_v, 0_w) = 0_{V\times W}$.

— Pour tout a dans $\mathbb{R}$ et (v, w) dans $V \times W$, la norme $\|a(v, w)\|_{V\times W}$ se calcule de la façon suivante :

$$\|a(v, w)\|_{V\times W} = \max\{\|av\|_V, \|aw\|_W\} = |a| \max\{\|v\|_V, \|w\|_W\} = |a|\|(v, w)\|_{V\times W}. \tag{7.205}$$

— Soient (v_1, w_1) et (v_2, w_2) dans $V \times W$.

$$\begin{aligned}\|(v_1, w_1) + (v_2, w_2)\|_{V\times W} &= \max\{\|v_1 + v_2\|_V, \|w_1 + w_2\|_W\}\\ &\leqslant \max\{\|v_1\|_V + \|v_2\|_V, \|w_1\|_W + \|w_2\|_W\}\\ &\leqslant \max\{\|v_1\|_V, \|w_1\|_W\} + \max\{\|v_2\|_V, \|w_2\|_W\}\\ &= \|(v_1, w_1)\|_{V\times W} + \|(v_2, w_2)\|_{V\times W}.\end{aligned} \tag{7.206}$$

(2) **Équivalence**
Dans cette preuve, nous considérons la « topologie de $V \times W$ » comme étant la topologie produit et « la topologie métrique de $V \times W$ » la topologie de la norme produit.

(2a) **Dans un sens** La définition 7.15 de la topologie produit dit qu'une prébase de $V \times W$ est donnée par

$$\big\{B(v, r) \times B(w, s) \text{ tel que } v \in V; w \in W; r, s > 0\big\}. \tag{7.207}$$

Nous prouvons maintenant que la partie $S = B(v_0, r) \times B(w_0, s)$ est un ouvert de l'espace $\big(V \times W, \|.\|_{V\times W}\big)$. Pour cela nous prouvons que tout élément de S contient un voisinage métrique contenu dans S.
Soit $(v_1, w_1) \in S$. Nous posons

$$d\big((v_1, w_1), (v_0, w_0)\big) = \delta < \max\{r, s\}. \tag{7.208}$$

99. Définition 7.158.
100. Topologie associée à une norme : c'est la topologie associée à la distance correspondante, définition 7.164.
101. Topologie produit, définition 7.15.

Nous considérons $\epsilon > 0$ et nous montrons que si ϵ est assez petit, $B\big((v_1, w_1), \epsilon\big) \subset S$. Pour cela nous considérons $(v, w) \in B\big((v_1, w_1), \epsilon\big)$ et nous calculons un tout petit peu :

$$\begin{aligned} d\big((v, w), (v_0, w_0)\big) &\leqslant d\big((v, w), (v_1, w_1)\big) + d\big((v_1, w_1), (v_0, w_0)\big) && \text{(7.209a)} \\ &< \epsilon + \delta. && \text{(7.209b)} \end{aligned}$$

Si ϵ est assez petit, le tout reste plus petit que $\max\{r, s\}$.

Donc S est bien un ouvert métrique par le théorème 7.8. Vu que la topologie métrique contient une prébase de la topologie produit, tout ouvert de la topologie produit est un ouvert de la topologie métrique.

(2b) **Dans l'autre sens** Soient un ouvert métrique $\mathcal{O}$ ainsi que $(v_0, w_0) \in \mathcal{O}$; il existe $r > 0$ tel que

$$B\big((v_0, w_0), r\big) \subset \mathcal{O}. \tag{7.210}$$

Nous affirmons que $B(v_0, r) \times B(w_0, r)$ est contenu dans $\mathcal{O}$, de telle sorte que $\mathcal{O}$ soit un ouvert de la topologie produit. Pour $(v_1, w_1) \in B(v_0, r) \times B(w_0, r)$ nous avons

$$d\big((v_1, w_1), (v_0, w_0)\big) = \max\{\|v_1 - v_0\|, \|w_1 - w_0\|\} < r \tag{7.211}$$

parce que $v_1 \in B(v_0, r)$ et $w_1 \in B(w_0, r)$.

Donc tout élément de $\mathcal{O}$ admet un voisinage « produit » contenu dans $\mathcal{O}$; donc $\mathcal{O}$ est ouvert pour le topologie produit.

□

7.228.
En particulier, pour la topologie de la norme maximum, la convergence d'une suite implique la convergence « composante par composante » par la proposition 7.65.

On remarque tout de suite que la norme $\|.\|_\infty$ sur $\mathbb{R}^2$ est la norme de l'espace produit $\mathbb{R} \times \mathbb{R}$. En outre cette définition nous permet de trouver plusieurs nouvelles normes dans les espaces $\mathbb{R}^p$. Par exemple, si nous écrivons $\mathbb{R}^4$ comme $\mathbb{R}^2 \times \mathbb{R}^2$ on peut munir $\mathbb{R}^4$ de la norme produit

$$\|(x_1, x_2, x_3, x_4)\|_{\infty,2} = \max\{\|(x_1, x_2)\|_\infty, \|(x_3, x_4)\|_2\}.$$

Les applications de projection de l'espace produit $V \times W$ vers les espaces «facteurs», V W sont notées $\mathtt{proj}_V$ et $\mathtt{proj}_W$ et sont définies par

$$\begin{aligned} \mathtt{proj}_V \colon V \times W &\to V \\ (v, w) &\mapsto v \end{aligned} \tag{7.212}$$

et

$$\begin{aligned} \mathtt{proj}_W \colon V \times W &\to W \\ (v, w) &\mapsto w. \end{aligned} \tag{7.213}$$

Les inégalités suivantes sont évidentes

$$\begin{aligned} \|\mathtt{proj}_V(v, w)\|_V &\leqslant \|(v, w)\|_{V \times W} \\ \|\mathtt{proj}_W(v, w)\|_W &\leqslant \|(v, w)\|_{V \times W}. \end{aligned} \tag{7.214}$$

La topologie de l'espace produit est induite par les topologies des espaces «facteurs». La construction est faite en deux passages : d'abord nous disons que une partie $A \times B$ de $V \times W$ est ouverte si A et B sont des parties ouvertes de V et de W respectivement. Ensuite nous définissons que une partie quelconque de $V \times W$ est ouverte si elle est une intersection finie ou une réunion de parties ouvertes de $V \times W$ de la forme $A \times B$.

Ce choix de topologie donne deux propriétés utiles de l'espace produit

(1) Les projections sont des **applications ouvertes**. Cela veut dire que l'image par $\texttt{proj}_V$ (respectivement $\texttt{proj}_W$) de toute partie ouverte de $V \times W$ est une partie ouverte de V (respectivement W).

(2) Pour toute partie A de V et B de W, nous avons $\text{Int}(A \times B) = \text{Int}\, A \times \text{Int}\, B$.[PgovlABeqbAbB]

Une propriété moins facile à prouver est que pour toute partie A de V et B de W nous avons $\overline{A \times B} = \bar{A} \times \bar{B}$. Voir le lemme 10.78.

Ce que nous avons dit jusqu'ici est valable pour tout produit d'un nombre fini d'espaces vectoriels normés. En particulier, pour tout $m > 0$ l'espace $\mathbb{R}^m$ peut être considéré comme le produit de m copies de $\mathbb{R}$.

Exemple 7.229.
Si V et W sont deux espaces vectoriels, nous pouvons considérer le produit $E = V \times W$. Les projections $\texttt{proj}_V$ et $\texttt{proj}_W$, définies dans la section 7.10, sont des applications linéaires.

En effet, la projection $\texttt{proj}_V : V \times W \to V$ est donnée par $\texttt{proj}_V(v, w) = v$. Alors,

$$\begin{aligned} \texttt{proj}_V\big((v, w) + (v', w')\big) &= \texttt{proj}_V\big((v + v'), (w + w')\big) \\ &= v + v' \\ &= \texttt{proj}_V(v, w) + \texttt{proj}_V(v', w'), \end{aligned} \tag{7.215}$$

et

$$\texttt{proj}_V\big(\lambda(v, w)\big) = \texttt{proj}_V\big((\lambda v, \lambda w)\big) = \lambda v = \lambda\, \texttt{proj}_V(v, w). \tag{7.216}$$

Nous laissons au lecteur le soin d'adapter ces calculs pour montrer que $\texttt{proj}_W$ est également une projection[102]. △

PropDXR_KbaLC

Proposition 7.230.
Si $\mathcal{O}$ est un voisinage de (a, b) dans $V \times W$ alors $\mathcal{O}$ contient un ouvert de la forme $B(a, r) \times B(b, r)$.

Démonstration. Puisque $\mathcal{O}$ est un voisinage, il contient un ouvert et donc une boule

$$B\big((a, b), r\big) = \{(v, w) \in V \times W \text{ tel que } \max\{\|v - a\|, \|w - b\|\} < r\}. \tag{7.217}$$

Évidemment l'ensemble $B(a, r) \times B(b, r)$ est dedans. □

7.11 Topologie réelle en dimension n

Nous considérons sur $\mathbb{R}$ la topologie donnée par la valeur absolue, et sur $\mathbb{R}^n$ celle de la topologie produit ou du maximum, qui sont identiques par le lemme 7.227.

En particulier, nous n'avons pas encore la norme donnée par $\|x\| = \sqrt{\sum_i x_i^2}$, parce qu'elle demande la racine carré, définie en 9.97.

7.11.1 Ouverts et fermés

La proposition suivante est évidemment à mettre en rapport avec le théorème 7.8.

PROPooEQYJooBbPiAj

Proposition 7.231.
Une partie A de $\mathbb{R}^n$ est ouverte si et seulement si pour tout $a \in A$ il existe $r > 0$ tel que $B(a, r) \subset A$.

Démonstration. C'est la définition 7.112 de la topologie métrique. □

LemMESSExh

Lemme 7.232.
Pour tout $x \in \mathbb{R}^n$ et tout $r > 0$ la boule[103] *$B(x, r)$ est ouverte.*

102. Écrivez-moi si ça pose un problème.
103. Définition 7.117.

Démonstration. Afin de prouver que la boule est ouverte, nous prenons un point $p \in B(x,r)$, et nous allons montrer qu'il existe une boule autour de p qui est contenue dans $B(x,r)$.

Étant donné que $p \in B(x,r)$, nous avons $d(p,x) < r$. Prouvons que la boule $B\big(p, r-d(p,x)\big)$ est contenue dans $B(x,r)$. Pour cela, nous prenons $p' \in B\big(p, r-d(p,x)\big)$, et nous essayons de prouver que $p' \in B(x,r)$. En effet, en utilisant l'inégalité triangulaire,

$$d(x,p') \leqslant d(x,p) + d(p,p') \leqslant d(x,p) + r - d(p,x) = r. \tag{7.218}$$

□

7.11.2 Point d'accumulation, point isolé

Les définitions de point d'accumulation et de point isolé sont 7.34 et 7.35. Nous voyons maintenant ce que ces définitions donnent dans le cas de l'espace topologique $\mathbb{R}$.

Lemme 7.233.
Soit $D \subset \mathbb{R}$. Un point $a \in D$ est isolé dans D si et seulement si il existe $\varepsilon > 0$ tel que

$$[a-\varepsilon, a+\varepsilon] \cap D = \{a\}. \tag{7.219}$$

Autrement dit, il existe un intervalle autour de a dans lequel a est le seul élément de D.

Lemme 7.234.
Un point $a \in \mathbb{R}$ est un point d'accumulation de D si pour tout $\varepsilon > 0$,

$$\Big([a-\varepsilon, a+\varepsilon]\backslash\{a\}\Big) \cap D \neq \varnothing. \tag{7.220}$$

Autrement dit, quel que soit l'intervalle autour de a que l'on considère, le point a n'est pas tout seul dans D.

Exemple 7.235.
Prenons $D = [0,1[\cup]2,3]$. Cet ensemble n'a pas de point isolé, et l'ensemble de ses points d'accumulation est $[0,1] \cup [2,3]$.

Notez que les points 1 et 2 sont des points d'accumulation de D qui ne font pas partie de D. Il est possible d'être un point d'accumulation de D sans être dans D, mais pour être un point isolé dans D, il faut être dans D. △

Exemple 7.236.
Soit $D = \{\frac{1}{n}\}_{n\in\mathbb{N}}$. Tous les points de cet ensemble sont des points isolés (vérifier !). Aucun point de D n'est point d'accumulation. Cependant 0 est un point d'accumulation. △

EXooWOYQooJolaTV

Exemple 7.237.
Soit $D =]1,2[\cup \{12\}$. Le point 12 est adhérence, mais pas d'accumulation parce que le voisinage $]9,14[$ n'intersectionne pas $D\backslash\{12\}$. △

7.11.3 Limite de suite

Définition 7.238 (Limite d'une suite dans $\mathbb{R}^m$)**.**
*Une suite de points (x_n) dans $\mathbb{R}^m$ est dite **convergente** si il existe un élément $\ell \in \mathbb{R}^m$ tel que*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geqslant N,\ \|x_n - \ell\| < \varepsilon. \tag{7.221}$$

EqCondLimSuite

*Dans ce cas, nous disons que ℓ est la **limite** de la suite (x_n) et nous écrivons $\lim x_n = \ell$ ou plus simplement $x_n \to \ell$.*

Remarque 7.239.
Nous n'écrivons pas « $\lim_{n\to\infty} x_n$ » parce que, lorsqu'on parle de suites, la limite est *toujours* lorsque n tend vers l'infini. Il n'y a aucun intérêt à chercher par exemple $\lim_{n\to 4} x_n$ parce que cela vaudrait x_4 et rien d'autre.

Ceci est une différence importante avec les limites de fonctions.

Lemme 7.240 (Unicité de la limite)**.**
Il ne peut pas y avoir deux nombres différents qui satisfont à la condition (7.221). En d'autres termes, si ℓ et ℓ' sont deux limites de la suite (x_n), alors $\ell = \ell'$.

Démonstration. Soit $\varepsilon > 0$. Nous considérons N tel que

$$\|x_n - \ell\| < \varepsilon \tag{7.222}$$

pour tout $n \geqslant N$, et $N' > 0$ tel que

$$\|x_n - \ell'\| < \epsilon \tag{7.223}$$

pour tout $n > N'$. Maintenant, nous prenons n plus grand que N et N' de telle façon que les deux équations pour x_n soient vérifiées en même temps. Alors

$$\|\ell - \ell'\| = \|\ell - x_n + x_n - \ell'\| \leqslant \|\ell - x_n\| + \|x_n - \ell'\| < 2\varepsilon. \tag{7.224}$$

Cela prouve que $\|\ell - \ell'\| = 0$. □

Le théorème de Bolzano-Weierstrass 7.144 dit que dans le cas métrique, la compacité séquentielle est équivalente à la compacité.

7.12 Topologie et distance

LemDUJXooWsnmpL

Lemme 7.241.
Soient (X_1, d_1) et (X_2, d_2) des espaces métriques séparables. Alors $X_1 \times X_2$ admet une base dénombrable de topologie constituée de produits de boules de X_1 par des boules de X_2. Plus précisément si A_i est dénombrable et dense dans X_i alors l'ensemble des produits

$$\big\{B(y_1, r_1) \times B(y_2, r_2)\big\}_{\substack{y_i \in A_i \\ r_i \in \mathbb{Q}^+}} \tag{7.225}$$

est une base de topologie pour $X_1 \times X_2$.

Démonstration. Soit $\mathcal{O}$ un ouvert de $X_1 \times X_2$ et $(x_1, x_2) \in \mathcal{O}$. Par définition de la topologie produit [104], il existe $r_1, r_2 \in \mathbb{Q}^+$ tels que $B(x_1, r_1) \times B(x_2, r_2) \subset \mathcal{O}$. Les parties A_i étant denses, il existe $y_i \in B(x_i, r_i/2) \cap A_i$. Avec ces choix nous avons $x_i \in B(y_i, \frac{r_i}{2})$. Nous avons donc

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}). \tag{7.226}$$

Il est facile de voir que $B(y_i, r_i/2) \subset B(x_i, r_i)$. En effet si $z_i \in B(y_i, r_i/2)$ alors

$$d_i(z_i, x_i) \leqslant d(z_i, y_i) + d(y_i, x_i) \leqslant \frac{r_i}{2} + \frac{r_i}{2} = r_i. \tag{7.227}$$

Au final,

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}) \subset \mathcal{O}. \tag{7.228}$$

□

Définition 7.242.
*Si (X, d_X) et (Y, d_Y) sont des espaces métriques, une **isométrie** est une application bijective $f : X \to Y$ telle que pour tout $x, y \in X$ nous ayons*

$$d_Y\big(f(x), f(y)\big) = d_X(x, y). \tag{7.229}$$

EQooVUOXooKJntMN

Remarque 7.243.
Une application vérifiant (7.229) est automatiquement injective. En pratique, il ne faut donc vérifier que la surjectivité.

104. Définition 7.15.

Exemple 7.244 (Manque de surjectivité)**.**
Si $X = [0, \infty[$ et $f(x) = x + 1$ alors f vérifie (7.229) pour la distance $d(x, y) = |x - y|$, mais n'est pas surjective. △

Proposition-Définition 7.245 (Groupe des isométries)**.**
Si (X, d) est un espace métrique,

(1) l'ensemble des isométries de X, noté $\mathrm{Isom}(X)$ *est un groupe pour la composition.*

(2) Ce groupe agit fidèlement[105] *sur X.*

PropLYMgVMJ

Proposition 7.246.
Une isométrie entre deux espaces métriques est continue.

Démonstration. Soient $f\colon X \to Y$ une application isométrique et $\mathcal{O}$ un ouvert de Y. Soit $a \in f^{-1}(\mathcal{O})$; si $d(a, b) < r$, alors $d\big(f(a), f(b)\big) < r$ et donc $b \in f^{-1}\big(B(f(a), r)\big)$. Donc autour de chaque point de $f^{-1}(\mathcal{O})$ nous pouvons trouver une boule ouverte contenue dans $f^{-1}(\mathcal{O})$, ce qui prouve que $f^{-1}(\mathcal{O})$ est ouvert. □

Exemple 7.247.
Si X est un ensemble, nous pouvons écrire la **distance discrète** :

$$d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y. \end{cases} \tag{7.230}$$

La topologie résultante est la topologie discrète, côtoyée dans l'exemple 7.10 [106].

Pour cette métrique, le groupe des isométries est le groupe symétrique de X, c'est-à-dire le groupe de toutes les bijections de X sur lui-même. △

7.12.0.1 Distance point-ensemble

Définition 7.248.
Si A est une partie de l'espace métrique (X, d) et si $x \in X$, nous disons que la ***distance*** *entre A et x est le nombre*

$$d(x, A) = \inf_{a \in A} d(x, a). \tag{7.231}$$

EqdefDistaA

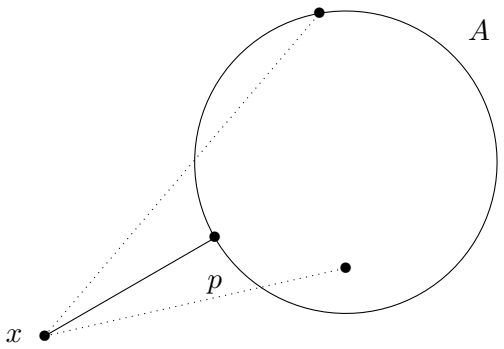


FIGURE 7.1: La distance entre x et A est donnée par la distance entre x et p. Les distances entre x et les autres points de A sont plus grandes que $d(x, p)$.

LabelFigDistanceEnsemble

105. Si vous ne savez pas ce que c'est, alors vous avez zappé la définition 2.30.
106. Vérifiez-le tout de même !

7.12.1 Suites et espaces métriques

PROPooJYOOooZWocoq

Proposition 7.249 (Caractérisation séquentielle de la limite[1]).
Soient deux espaces métriques X et Y ainsi qu'une fonction $f\colon X \to Y$. Soit $a \in X$ et $\ell \in Y$. On a

$$\lim_{x\to a} f(x) = \ell, \tag{7.232}$$

EqLimooJYOOooZWocoqG

si et seulement si, pour toute suite (x_k) telle que $x_k \to a$, on a

$$\lim f(x_k) = \ell. \tag{7.233}$$

EqLimooJYOOooZWocoqS

Par ailleurs, l'une des deux limites existe si et seulement si l'autre existe.

Démonstration. Le sens direct est la proposition 7.210. Pour la réciproque, nous passons par la contraposée. C'est-à-dire que nous supposons que ℓ n'est pas une limite de f pour $x \to a$. Il existe un ϵ tel que pour tout δ, il existe un x vérifiant $d_X(x;a) < \delta$ et $d_Y(f(x);\ell) > \epsilon$.

Nous construisons à présent une suite de la manière suivante. Pour $\delta = \frac{1}{n}$ nous considérons x_n tel que $d_X(x_n;a) < \delta$ et $d_Y(f(x_n);\ell) > \epsilon$. Cette suite converge vers a, mais la suite $f(x_n)$ ne converge manifestement pas vers ℓ : elle ne rentre jamais dans la boule $B(\ell,\epsilon)$. □

Une fonction continue est séquentiellement continue. Dans les espaces métriques la proposition suivante montre que la réciproque est également vraie et la continuité est équivalente à la continuité séquentielle. Cela n'est cependant pas vrai pour n'importe quel espace topologique.

PROPooBHRBooJMZYSg

Corolaire 7.250 (Caractérisation séquentielle de la continuité en un point[1]).
Si X et Y sont des espaces métriques, alors une fonction $f\colon X \to Y$ est continue en un point si et seulement si elle est séquentiellement continue en ce point.

Démonstration. Paraphrasons la preuve précédente. Nous supposons que X et Y sont métriques. Si f n'est pas continue en a, il existe $\epsilon > 0$ tel que pour tout $\delta > 0$, il existe x tel que $\|x - a\| \leqslant \delta$ et $\|f(x) - f(a)\| > \epsilon$. Nous considérons un tel ϵ et pour chaque $n \geqslant 1 \in \mathbb{N}$ nous considérons un x_n correspondant à $\delta = \frac{1}{n}$. Cela nous donne une suite $x_n \to a$ dans X mais $\|f(x_n) - f(a)\|$ reste plus grand que ϵ. Cela montre que f n'est pas non plus séquentiellement continue. □

Définition 7.251 (Fermeture séquentielle[232]).
*Une partie F d'un espace topologique X est dit **séquentiellement fermé** si la convergence d'une suite (x_n) de F vers x implique que x appartient à F.*

Les espaces métriques ont une propriété importante que la fermeture séquentielle est équivalente à la fermeture.

PropLFBXIjt

Proposition 7.252 (Caractérisation séquentielle d'un fermé).
Soient X un espace métrique et $F \subset X$. L'ensemble F est fermé si et seulement si toute suite contenue dans F et convergeant dans X converge vers un élément de F.

Démonstration. Une suite contenue dans un fermé ne peut converger que vers un élément de ce fermé : c'était la proposition 7.56. Le point le plus important est donc l'autre sens : si toute suite d'éléments de F converge dans F alors F est fermé.

Par contraposée, supposons que $X\backslash F$ ne soit pas ouvert. Alors il existe $x \in X\backslash F$ pour lequel tout voisinage intersecte F. En prenant $x_k \in B(x, \frac{1}{k})$, nous construisons une suite contenue dans F, convergeant vers x qui n'est pas dans F. □

Le lemme suivant est précisément la version « espace métrique » du corolaire 7.57 ; mais, donnons-en une preuve tout de même.

LemLimAbarA

Lemme 7.253.
Soit X un espace métrique, et soit (x_n) une suite convergente contenue dans un ensemble $A \subset X$. Alors la limite x_n appartient à $\bar{A}$.

Démonstration. Supposons que nous ayons une partie A de X, et une suite (x_n) dont la limite ℓ se trouve hors de $\bar{A}$. Dans ce cas, il existe un $r > 0$ tel que [107] $B(\ell, r) \cap A = \varnothing$. Si tous les éléments x_n de la suite sont dans A, il n'y en a donc aucun tel que $d(x_n, \ell) < r$. Cela contredit la notion de convergence $x_n \to \ell$. □

CorAdhEstLim

Corolaire 7.254.
Soit X un espace métrique, $A \subset X$ et $a \in \bar{A}$. Alors il existe une suite d'éléments dans A qui converge vers a.

Démonstration. Si $a \in A$, alors nous pouvons prendre la suite constante $x_n = a$. Si a n'est pas dans A, alors a est dans ∂A, et pour tout n, il existe un point de A dans la boule $B(a, \frac{1}{n})$. Si nous nommons x_n ce point, la suite ainsi construite est une suite contenue dans A et qui converge vers a (ce dernier point est laissé à la sagacité du lecteur.). □

En termes savants, ce corolaire signifie que la fermeture $\bar{A}$ est composé de A plus de toutes les limites de toutes les suites contenues dans A.

PropXIAQSXr

Proposition 7.255 (Caractérisation séquentielle de la continuité[1])**.**
Soient X et Y deux espaces topologiques séparés. Nous supposons que X est métrisable. Une application $f\colon X \to Y$ est continue sur X si et seulement si elle est séquentiellement continue sur X.

Démonstration. Le sens direct est déjà prouvé dans la proposition 7.210. Nous nous concentrons donc sur la réciproque.

Soit $\mathcal{O}$ un ouvert de Y ; nous allons voir que le complémentaire de $f^{-1}(\mathcal{O})$ est fermé dans X. Pour cela nous considérons une suite convergente $x_k \xrightarrow{X} x$ avec $x_k \in X\backslash f^{-1}(\mathcal{O})$ pour tout k. Nous allons montrer que $x \in X\backslash f^{-1}(\mathcal{O})$ et la caractérisation séquentielle [108] de la fermeture conclura que $X\backslash f^{-1}(\mathcal{O})$ est fermé.

Pour tout k, nous avons $f(x_k) \in X\backslash\mathcal{O}$, et $f(x_k) \xrightarrow{Y} f(x)$ parce que f est séquentiellement continue. Vu que $f(x_k)$ est une suite dans le fermé $Y\backslash\mathcal{O}$, la limite est également dans $Y\backslash\mathcal{O}$. Nous en déduisons que $f(x) \in Y\backslash\mathcal{O}$, de telle sorte que $x \in X\backslash f^{-1}(\mathcal{O})$. □

PropCJGIooZNpnGF

Proposition 7.256.
Si X et Y sont deux espaces métriques et $f, g\colon X \to Y$ sont deux fonctions continues égales sur une partie dense de X alors $f = g$.

Démonstration. Les fonctions f et g sont séquentiellement continues (proposition 7.137, ou proposition 7.250). Soient A un ensemble dense dans X sur lequel f et g sont égales, et $x \notin A$. Vu que A est dense, il existe une suite a_n dans A telle que $a_n \to x$. La séquentielle continuité de f et g donnent

$$f(a_n) \to f(x) \tag{7.234a}$$
$$g(a_n) \to g(x), \tag{7.234b}$$

mais pour tout n, $f(a_n) = g(a_n)$. Par unicité de la limite [109] dans Y, $f(x) = g(x)$. □

7.12.2 Espace métrisable

DEFooGLTUooJjaSmI

Définition 7.257 (Espace vectoriel topologique métrisable[233])**.**
*Un espace topologique est **métrisable** si il existe une distance compatible avec la topologie.*

107. Une autre manière de dire la même chose : si $\ell \notin \bar{A}$, alors $d(\ell, A) > 0$.
108. Proposition 7.252, valable parce que la topologie de X provient d'une métrique.
109. Proposition 7.113.

PROPooXWBTooCvGLOj

Proposition 7.258 ([234]).
Soit un espace topologique métrisable X.

ITEMooOXVRooBsKwuq

(1) Tout fermé de X est une intersection dénombrable d'ouverts.

(2) Tout ouvert de X est une union dénombrable de fermés.

Démonstration. Soit une métrique d compatible avec la topologie de X et un fermé A. Nous posons

$$V_n = \{x \in X \text{ tel que } d(x, A) < \frac{1}{n}\}. \tag{7.235}$$

Et juste pour faire simple nous notons $V_0 = X$.

(1) **Les parties V_n sont ouvertes** Soit $x \in V_n$. Trouvons un voisinage de x contenu dans V_n afin de pouvoir encore invoquer le théorème 7.8. D'abord, vu que $x \in V_n$, il existe $a \in A$ tel que $d(x, a) < \frac{1}{n}$ (ici les inégalités strictes sont importantes).
Soient $\epsilon > 0$ que nous fixerons plus bas, et $y \in B(x, \epsilon)$. L'inégalité triangulaire donne

$$d(y, a) \leqslant d(y, x) + d(x, a) < \epsilon + \frac{1}{n}. \tag{7.236}$$

Nous pouvons donc choisir ϵ de telle sorte que $d(y, a) < 1/n$. Avec ce ϵ, nous avons, pour tout $y \in B(x, \epsilon)$:

$$d(y, A) \leqslant d(y, a) < \frac{1}{n} \tag{7.237}$$

et donc $y \in V_n$.

(2) **A est l'intersection des V_n** Nous avons évidemment $A \subset V_n$ pour tout n. Et d'autre part, si $a \in \bigcap_{n\in\mathbb{N}} V_n$ alors $d(a, A) < \frac{1}{n}$ pour tout n. Cela implique $d(a, A) = 0$, et donc $a \in A$ par le lemme 7.149.

Ceci démontre le point (1).

En ce qui concerne la seconde partie, nous appliquons la première partie au complémentaire. Si $\mathcal{O}$ est ouvert, $\mathcal{O}^c$ est fermé et

$$\mathcal{O}^c = \bigcap_{n\in\mathbb{N}} V_n, \tag{7.238}$$

ce qui donne immédiatement

$$\mathcal{O} = \bigcup_{n\in\mathbb{N}} V_n^c \tag{7.239}$$

où les V_n^c sont fermés. □

CORooTWFYooCNMieM

Corolaire 7.259.
Si X est un espace topologique métrisable, alors X accepte une base dénombrable de topologie autour de chaque point.

Démonstration. Il s'agit seulement de remarquer que les singletons sont fermés et d'appliquer la proposition 7.258. □

7.13 Suites de Cauchy, métrique et espaces complets

7.13.1 Généralités

DefZSnlbPc

Définition 7.260 (Suite de τ-Cauchy, espace vectoriel topologique[216, 219]).
*Soit E un espace vectoriel topologique. Une suite (x_k) dans E est une **suite τ-Cauchy** si pour tout voisinage $\mathcal{U}$ de 0 il existe $N \in \mathbb{N}$ tel que $x_k - x_l \in \mathcal{U}$ pour tout $k, l \geqslant N$.*

DEFooVQDBooNxprFU

Définition 7.261 (Espace τ-complet)**.**
Nous disons qu'une partie A d'un espace vectoriel topologique est ***τ-complet*** *si toute suite τ-Cauchy d'éléments de A converge*[110] *vers un élément de A.*

DEFooXOYSooSPTRTn

Définition 7.262 (Suite de Cauchy, espace métrique)**.**
Une suite (a_k) dans un espace métrique (V,d) est ***de Cauchy*** *si pour tout $\epsilon > 0$ dans $\mathbb{R}$, il existe N tel que si $n, m \geqslant N$ alors $d(a_n, a_m) < \epsilon$.*

Notons qu'ici, même si l'espace V n'a rien à voir avec $\mathbb{R}$, nous prenons ϵ dans $\mathbb{R}$ et la distance à valeurs dans $\mathbb{R}$. Cela semble une évidence, mais il faut se rendre compte que $\mathbb{R}$ commence à prendre une place centrale dans nos constructions. Ce n'était pas le cas du temps où nous parlions de suites de Cauchy et de complétude dans des corps totalement ordonnés (définitions 1.379). Dans ce contexte, le ϵ était pris dans le corps lui-même.

DEFooHBAVooKmqerL

Définition 7.263 (Métrique complète)**.**
Soit (E,d) un espace métrique. Nous disons que la métrique d est ***complète*** *si toute suite de Cauchy dans (E,d) converge dans E.*

7.264.
Ces définitions méritent quelques remarques.

(1) Dans le cas des espaces vectoriels topologiques, nous définissons les notions de suite τ-Cauchy et d'espace topologique τ-complet. Nous ajoutons le préfixe τ pour indiquer que ce sont des notions topologiques.

(2) Dans le cas des espaces métriques, nous définissons la notion de *métrique* complète. C'est bien la métrique qui est complète, et non l'espace. En effet nous allons voir dans l'exemple 7.266 que le même espace topologique peut accepter plusieurs distances différentes (donnant la même topologie) donnant lieu à des suites de Cauchy différentes.

(3) Si un espace vectoriel a une topologie issue d'une distance, rien ne dit que ses suites τ-Cauchy et ses suites de Cauchy sont les mêmes. Ce sont deux notions à priori séparées. Si V est un espace vectoriel topologique que l'on peut munir de deux distances d_1, d_2 donnant toutes deux la topologie, dire que V est τ-complet, dire que d_1 est complète et dire que d_2 est complète sont trois choses différentes. Même si les trois topologies sont identiques.

(4) Nous allons bien entendu voir que dans de larges gammes d'exemples, les notions de suite de Cauchy et τ-Cauchy coincident.

7.13.2 Espace de Banach

DefVKuyYpQ

Définition 7.265 (espace de Banach, algèbre de Banach[235])**.**
Un ***espace de Banach*** *est un espace vectoriel normé complet*[111] *pour la topologie de la norme.*

Une ***algèbre de Banach*** *est une algèbre commutative et associative qui est un espace vectoriel normé complet.*

EXooNMNVooXyJSDm

Exemple 7.266 (La complétude n'est pas une propriété topologique[236])**.**
Le fait pour un espace d'être complet n'est pas une propriété topologique, mais une propriété métrique. Plus exactement, il existe des espaces topologiques isomorphes, mais dont l'un est complet et l'autre non.

Nous considérons la distance suivante sur $\mathbb{N}$:

$$d_1(x,y) = \left|\frac{1}{x} - \frac{1}{y}\right|. \tag{7.240}$$

110. Définition 7.13.
111. Définition 7.263.

Pour vérifier que cette formule définit bien une distance (définition 7.115), le seul point non immédiat est l'inégalité triangulaire :

$$d_1(x,y)=\left|\frac{1}{x}-\frac{1}{y}\right|\leqslant\left|\frac{1}{x}-\frac{1}{z}\right|+\left|\frac{1}{z}-\frac{1}{y}\right|=d_1(x,z)+d_1(z,y).\tag{7.241}$$

Au niveau de la topologie induite par cette distance, c'est la topologie discrète. En effet, soit $x\in\mathbb{N}$ et $\epsilon>0$; nous voulons déterminer la boule $B(x,\epsilon)$ en résolvant l'équation

$$\left|\frac{1}{x}-\frac{1}{y}\right|<\epsilon\tag{7.242}$$

pour $y\in\mathbb{N}$. Nous trouvons que $\frac{1}{y}>\frac{1}{x}-\epsilon$ et $\frac{1}{y}<\frac{1}{x}+\epsilon$, soit

$$\begin{cases}y>\dfrac{1}{\frac{1}{x}+\epsilon}&(7.243a)\\ y<\dfrac{1}{\frac{1}{x}-\epsilon}.&(7.243b)\end{cases}$$

Si ϵ est assez petit, la seule solution entière est $y=x$. Les ouverts sont donc toutes les parties parce que tous les singletons sont ouverts.

L'espace topologique associé à $(\mathbb{N},d_1)$ est donc la topologie discrète[112].

Si nous considérons par contre la distance usuelle sur $\mathbb{N}$, à savoir $d(x,y)=|x-y|$, nous obtenons encore la topologie discrète. Nous avons donc un isomorphisme d'espaces topologiques

$$(\mathbb{N},d)\simeq(\mathbb{N},d_1).\tag{7.244}$$

Nous pouvons même donner un isomorphisme explicite : $f(n)=n$.

La suite $(x_n)=n$ est une suite de Cauchy dans $(\mathbb{N},d_1)$ parce que si $\epsilon>0$ est donné, il suffit de prendre N assez grand pour avoir $\frac{1}{N}<\epsilon$ (possible par le lemme 1.441) nous avons, pour $n,m>N$:

$$\left|\frac{1}{n}-\frac{1}{m}\right|<\frac{1}{n}<\frac{1}{N}<\epsilon.\tag{7.245}$$

Or cette suite ne converge pas. Soit en effet un candidat limite k. Calculons

$$d_1(x_n,k)=\left|\frac{1}{n}-\frac{1}{k}\right|\to\frac{1}{k}\neq0.\tag{7.246}$$

L'espace $(\mathbb{N},d_1)$ n'est pas complet.

Notons que cette suite n'est pas de Cauchy dans $(\mathbb{N},d)$.

En résumé :

(1) Les espaces topologiques $(\mathbb{N},d)$ et $(\mathbb{N},d_1)$ sont isomorphes.
(2) Ils ont les mêmes notions de suites convergentes : une suite convergente pour l'un est convergente pour l'autre.
(3) Ils n'ont pas les mêmes notions de suites de Cauchy.
(4) Dans $(\mathbb{N},d_1)$, il existe des suites de Cauchy qui ne convergent pas (pas complet).
(5) L'espace $(\mathbb{N},d)$ est complet, mais $(\mathbb{N},d_1)$ n'est pas complet.
(6) Le fait pour un espace topologique métrique d'être complet n'est pas intrinsèque à sa topologie : la complétude est une propriété de la distance. La complétude est une propriété de la métrique, et non de la topologie qui s'en suit.

△

PROPooOVGGooNffWJW

Proposition 7.267 ([237]).
Le dual[113] *d'un espace de Banach*[114] *est de Banach.*

PROPooAPVCooFyMigL

Proposition 7.268 ([1]).
Soit un espace ce Banach E. Soient $v_k\in E$. Si $\sum_{k=0}^{\infty}v_k$ converge, alors $\{|v_k|\}_{k\in\mathbb{N}}$ est borné.

112. Celle dont toutes les parties sont des ouverts.
113. Définition 4.126.
114. Définition 7.265.

7.13.3 Espace topologique métrique

Dans les espaces vectoriels topologiques métriques, il n'y a pas d'ambiguïté.

PropooUEEOooLeIImr

Proposition 7.269 (Caractérisations avec la distance d).
Soit (E, d) un espace vectoriel topologique métrique.

ItemooROYMooAQCXnj

(1) Une suite (x_n) dans E est convergente[115] *vers x si et seulement si pour tout $\epsilon \in \mathbb{R}$ il existe N_ϵ tel que pour tout $n \geqslant N_\epsilon$ nous avons $d(x_n, x) \leqslant \epsilon$.*

(2) Une suite (x_n) dans E est de Cauchy[116] *si pour tout $\epsilon \in \mathbb{R}$, il existe un N_ϵ tel que si $p, q \geqslant N_\epsilon$, nous avons $d(x_p, x_q) \leqslant \epsilon$.*

Démonstration. En ce qui concerne la convergence :

(1) **Sens direct**

Nous supposons que $x_k \to x$ dans E. Soit $\epsilon > 0$; vu que $B(x, \epsilon)$ est un ouvert contenant x, il existe un $N_\epsilon > 0$ tel que $k > N_\epsilon$ implique $x_k \in B(x, \epsilon)$. Cela signifie $d(x, x_k) \leqslant \epsilon$.

(2) **Réciproque**

Nous supposons que pour tout $\epsilon > 0$, il existe $N_\epsilon > 0$ tel que si $k > N_\epsilon$ alors $x_k \in B(x, \epsilon)$. Soit un ouvert $\mathcal{O}$ autour de x. Nous sommes dans un espace métrique ; ergo la topologie est donné par le théorème 7.117 et en particulier la liste des ouverts est donnée par (7.112). Il existe donc une boule $B(x, \epsilon)$ incluse à $\mathcal{O}$. Pour tout $k > N_\epsilon$ nous avons alors $x_k \in B(x, \epsilon) \subset \mathcal{O}$.

En ce qui concerne les suites de Cauchy :

(1) **Sens direct** Si (x_n) est une suite de Cauchy et si $\epsilon > 0$ est donné, alors $B(0, \epsilon)$ est un voisinage de 0 et il existe N_ϵ tel que si $p, q \geqslant N_\epsilon$ alors $x_p - x_q \in B(0, \epsilon)$. Posons $u = x_p - x_q$; en utilisant l'invariance par translation (lemme 7.163(1)) nous avons

$$d(u, 0) = d(x_p - x_q, 0) = d(x_p, x_q). \tag{7.247}$$

Par conséquent $d(x_p, x_q) \leqslant \epsilon$.

(2) **Réciproque** Soit $\mathcal{O}$ un voisinage de 0. Il existe ϵ tel que $B(0, \epsilon) \subset \mathcal{O}$. Par hypothèse il existe N_ϵ tel que $d(x_p, x_q) \leqslant \epsilon$ dès que $p, q \geqslant N_\epsilon$. En utilisant encore l'invariance par translation nous avons

$$d(x_p, x_q) = d(x_p - x_q, 0), \tag{7.248}$$

et comme cela est plus petit que ϵ, nous avons $x_p - x_q \in B(0, \epsilon) \subset \mathcal{O}$.

□

PROPooZZNWooHghltd

Proposition 7.270 ([238]).
Toute suite convergente dans un espace métrique est de Cauchy.

Démonstration. Nous utilisons les caractérisations de la proposition 7.269 des suites convergentes et de Cauchy.

Soit un espace métrique (X, d) et $x_n \to \ell$ une suite convergente. Si $\epsilon > 0$, la proposition 7.269(1), dit qu'il existe N tel que pour tout $n > N$ nous ayons $d(x_n, \ell) < \epsilon$. Par conséquent si $n, m > N$ alors

$$d(x_n, x_m) \leqslant d(x_m, \ell) + d(l, x_m) \leqslant 2\epsilon. \tag{7.249}$$

Cela prouve que (x_n) est une suite de Cauchy. □

115. Définition 7.13.
116. Définition 7.260.

7.13.4 Compacts, fermés

PROPooRMAOooZIgXwy

Proposition 7.271 (Séparation [219, 239]).
Soit V, un espace vectoriel topologique. Soient des parties K et F de V telles que :

(1) K est compact,

(2) F est fermé,

(3) $K \cap F = \varnothing$.

Alors il existe un voisinage U de 0 tel que

$$(K + U) \cap (F + U) = \varnothing. \tag{7.250}$$

Autrement dit, la topologie d'un espace vectoriel topologique sépare les fermés des compacts.

Démonstration. Soit un élément $x \in K$. La partie F^c est un ouvert qui contient x ; donc $F^c - x$ est un voisinage ouvert de 0.

(1) **Quelque chose sans intersection avec F** Par la proposition 7.193, il existe un ouvert symétrique U_x autour de 0 tel que

$$U_x + U_x + U_x + U_x \subset F^c - x. \tag{7.251}$$

En enlevant un des éléments de la somme, nous gardons la même inclusion :

$$U_x + U_x + U_x \subset F^c - x. \tag{7.252}$$

Et en passant le x de l'autre côté :

$$x + U_x + U_x + U_x \subset F^c. \tag{7.253}$$

Cela pour dire que nous avons, pour chaque $x \in K$ un voisinage ouvert U_x de 0 tel que

$$(x + U_x + U_x + U_x) \cap F = \varnothing. \tag{7.254}$$
EQooPXUVooVvUEoN

(2) $(x + U_x + U_x) \cap (U_x + F) = \varnothing$ Supposons $z \in (x + U_x + U_x) \cap (U_x + F)$, et déduisons une contradiction. En particulier $z \in U_x + F$ et donc il existe $u \in U_x$, $f \in F$ tel que $z = v + f$. Nous avons alors (nous utilisons le fait que U_x soit symétrique)

$$f = z - v \in z + U_x \tag{7.255}$$

Mais z est aussi dans $x + U_x + U_x$. Donc

$$f \in z + U_x \subset x + U_x + U_x + U_x. \tag{7.256}$$

Cela prouve que $x + U_x + U_x + U_x \cap F \neq \varnothing$, en contradiction avec (7.254). Nous en concluons que

$$(x + U_x + U_x) \cap (U_x + F) = \varnothing. \tag{7.257}$$
EQooNLCRooWTEove

(3) **Un sous-recouvrement fini** Fini de rigoler. Si nous avons un compact dans les hypothèses, il fallait que ça arrive. Pour chaque $x \in K$ nous avons un voisinage ouvert U_x de 0 et donc un voisinage ouvert $x + U_x$ de x. Du coup $\{U_x + x\}_{x \in K}$ forme un recouvrement de K par des ouverts. Et hop, sous-recouvrement fini (c'est la définition 7.80 d'un compact) : nous avons $\{x_i\}_{i=1,\dots,n}$ tels que

$$K \subset \bigcup_{i=1}^{n} (x_i + U_{x_i}). \tag{7.258}$$

Nous posons $U = \bigcap_{i=1}^{n} U_{x_i}$. Même si c'est évident, remarquez que U est un ouvert autour de 0 et que U est plus petit que chacun des U_{x_i}. Récrivons (7.257) pour chacun des x_i :

$$(x + U_{x_i} + U_{x_i}) \cap (F + U_{x_i}) = \varnothing. \tag{7.259}$$
EQooGEAGooENFlCG

Lorsqu'on enlève des éléments dans une intersection vide, l'intersection reste vide. Donc en remplaçant des U_{x_i} par des U dans (7.259), nous conservons une intersection vide :

$$(x + U_{x_i} + U) \cap (F + U) = \varnothing. \tag{7.260}$$

(4) **Conclusion** Aucun ensemble de la forme $x + U_{x_i} + U$ n'intersecte $F + U$. L'union ne l'intersecte pas non plus :

$$\bigcup_{i=1}^{n}(x + U_{x_i}) + U \cap (F + U) = \varnothing. \tag{7.261}$$

Et donc

$$(K + U) \cap (F + U) = \varnothing, \tag{7.262}$$

comme nous le voulions.

□

DEFooGTOZooRcvGHg

Définition 7.272.
Une distance d sur un espace vectoriel topologique V est dite ***compatible*** *avec la topologie si la topologie induite*[117] *de d est celle de V.*

DEFooEUXVooEBYhNU

Définition 7.273.
Une distance d sur un espace vectoriel V est dite ***invariante*** *si pour tout $x, y, a \in V$ nous avons*

$$d(x + a, y + a) = d(x, y). \tag{7.263}$$

Notons que lorsque nous parlons d'une distance compatible avec un espace vectoriel topologique, nous parlons de compatibilité avec la topologie, pas avec la structure vectorielle.

LEMooZNBAooOhEwJd

Lemme 7.274 ([1])**.**
Soit un espace vectoriel muni d'une distance invariante. Alors

$$B(a, r) + x = B(a + x, r). \tag{7.264}$$

Démonstration. Un élément de $B(a, r) + x$ est de la forme $y + x$ avec $d(a, y) < r$. Nous avons alors $d(y + x, a + x) = d(y, a) < r$, de telle sorte que $y + x \in B(a + x, r)$.

Dans le sens inverse, si $y \in V$ vérifie $d(a + x, y) < r$, alors je prétend que $y - x \in B(a, r)$. En effet $d(y - x, a) = d(y - x + x, a + x) = d(y, a + x) < r$. □

Une version plus complète du lemme suivant sera dans la proposition 10.145 et ce qui suit.

LEMooIQBXooUEtdoy

Lemme 7.275.
Nous introduisons l'ensemble $\mathbb{D}_2$ des suites finies[118] $\mathbb{N}\backslash\{0\} \to \{0, 1\}$[119]*. Nous notons*

$$l(x) = \max\{i \in \mathbb{N} \text{ tel que } x_i \neq 0\}. \tag{7.265}$$

L'application

$$\begin{aligned} \varphi\colon \mathbb{D}_2 &\to [0, 1[\\ x &\mapsto \sum_{i=1}^{l(x)} \frac{x_i}{2^i}. \end{aligned} \tag{7.266}$$

est injective.

Démonstration. Le lemme 1.487 donne la somme partielle de la série géométrique. Dans notre cas, $q = 1/2$. Si $x \in D_n$, alors en majorant chacun des c_i par 1,

$$x \leqslant \sum_{i=1}^{n} \left(\frac{1}{2}\right)^i = \frac{1 - (1/2)^{n+1}}{1/2} - 1 = 1 - (1/2)^{n+1} < 1. \tag{7.267}$$

117. Définition 7.117.
118. Les suites donc nous parlons sont indexées en commençant à 1, et non à 0. C'est une précision importante.
119. C'est en gros ce qui se fera dans 10.143 pour les développements de nombres dans une base donnée, sauf qu'ici nous n'avons pas besoin de subtilités sur les queues de suites.

Donc c'est bon pour dire que φ prend ses valeurs dans $[0,1[$.

Il reste à voir l'injectivité. Nous supposons que $\varphi(x)=\varphi(y)$. Quitte à alonger x ou y par des zéros, nous supposons qu'elles ont même longueur N. Enfin nous définissons n_0 le plus petit indice pour lequel $x_i\neq y_i$. Nous avons :

$$0=\frac{1}{2^{n_0}}+\sum_{i=n_0+1}^{N}\frac{x_i-y_i}{2^i}. \tag{7.268}$$

EQooEUKNooVoJfEC

Le deuxième terme est majoré de la façon suivante :

$$\begin{aligned}
\left|\sum_{i=n_0+1}^{N}\frac{x_i-y_i}{2^i}\right| &\leqslant \sum_{i=n_0+1}^{N}\frac{|x_i-y_i|}{2^i} && \text{(7.269a)}\\
&\leqslant \sum_{i=n_0+1}^{N}\frac{1}{2^i} && \text{(7.269b)}\\
&= \sum_{i=1}^{N}\left(\frac{1}{2}\right)^i-\sum_{i=1}^{n_0}\left(\frac{1}{2}\right)^i && \text{(7.269c)}\\
&= \frac{1-(1/2)^{N+1}}{1/2}-\frac{1-(1/2)^{n_0+1}}{1/2} && \text{(7.269d)}\\
&= \frac{1}{2^{n_0}}-\frac{1}{2^N} && \text{(7.269e)}\\
&< \frac{1}{2^{n_0}}. && \text{(7.269f)}
\end{aligned}$$

L'égalité (7.268) n'est donc pas possible. Nous déduisons que n_0 n'existe en fait pas et que $x=y$. D'où l'injectivité de φ. □

Pour faire les sommes dans $\mathbb{D}_2$, nous introduisons $\mathbb{D}_2'$, l'ensemble des suites finies $\mathbb{N}\to\{0,1\}$ et

$$\begin{aligned}
\varphi'\colon \mathbb{D}_2' &\to [0,2[\\
x &\mapsto \sum_{i=0}^{l}(x)\frac{x_i}{2^i}.
\end{aligned} \tag{7.270}$$

DEFooRJSFooDvbkrx

Définition 7.276 ([1])**.**
Si $x,y\in\mathbb{D}_2$, nous définissons $z,r\in\mathbb{D}_2'$ par récurrence (à l'envers) de la façon suivantes[120]*. D'abord, quitte à ajouter des zéros, nous supposons que x et y aient la même longueur l.*

D'abord posons $r_l=0$.

Ensuite, pour tout k nous considérons les trois cas possibles.

Si $x_k+y_k+r_k$ **vaut** 0 **ou** 1 *Dans ce cas nous posons*

- $z_k=x_k+y_k+r_k$
- $r_{k-1}=0$

Si $x_k+y_k+r_k=2$ *Nous posons*

- $z_k=0$
- $r_{k-1}=1$

Si $x_k+y_k+r_k=3$ *Nous posons*

- $z_k=1$
- $r_{k-1}=1$.

Et enfin nous posons $z_0=r_0$.

120. z va être la somme en binaire et r va être la suite des reports.

LEMooIUDHooNDlVtg

Lemme 7.277 ([1]).
Si $r_k = 1$, alors $z_k \neq x_k + y_k$.

Démonstration. Il y a seulement quelques cas à traiter.

x_k	y_k	r_k	z_k
0	0	1	$z_k = 1 \neq x_k + y_k$
0	1	1	$z_k = 0 \neq x_k + y_k$
1	0	1	$z_k = 0 \neq x_k + y_k$
1	1	1	$z_k = 1 \neq x_k + y_k$

□

PROPooGHPZooEwXqMH

Proposition 7.278 ([1]).
Soient $x, y \in \mathbb{D}_2$ ainsi que z défini par 7.276. Nous avons

$$\varphi(x) + \varphi(y) = \varphi'(z). \tag{7.271}$$

Démonstration. Nous supposons que x et y sont de longueur l, et nous faisons une récurrence à l'envers. D'abord

$$\varphi(x) + \varphi(y) = \sum_{i=1}^{l} \frac{x_i + y_i}{2^i} = \sum_{i=1}^{l-2} \frac{x_i + y_i}{2^i} + \frac{x_{l-1} + y_{l-1}}{2^{l-1}} + \frac{x_l + y_l}{2^l}. \tag{7.272}$$

Il y a deux possibilités. Si $x_l + y_l$ vaut 0 ou 1, alors $r_{l-1} = 0$ et $z_l = x_l + y_l$, de telle sorte que

$$\frac{x_{l-1} + y_{l-1}}{2^{l-1}} + \frac{x_l + y_l}{2^l} = \frac{x_{l-1} + y_{l-1} + r_{l-1}}{2^{l-1}} + \frac{z_l}{2^l}. \tag{7.273}$$

Si $x_l + y_l = 2$, alors $r_{l-1} = 1$ et $z_l = 0$ de telle sorte que

$$\frac{x_{l-1} + y_{l-1}}{2^{l-1}} + \frac{x_l + y_l}{2^l} = \frac{x_{l-1} + y_{l-1}}{2^{l-1}} + \frac{2}{2^l} \tag{7.274a}$$

$$= \frac{x_{l-1} + y_{l-1} + 1}{2^{l-1}} + \frac{z_l}{2^l} \tag{7.274b}$$

$$= \frac{x_{l-1} + y_{l-1} + r_{l-1}}{2^{l-1}} + \frac{z_l}{2^l}. \tag{7.274c}$$

En ce qui concerne la récurrence, il faut faire les trois cas pour montrer que dans tous les cas,

$$\frac{x_k + y_k}{2^k} + \frac{x_{k+1} + y_{k+1} + r_{k+1}}{2^{k+1}} = \frac{x_k + y_k + r_k}{2^k} + \frac{z_{k+1}}{2^{k+1}}. \tag{7.275}$$

□

LEMooEAIBooMxCDHF

Lemme 7.279 ([1]).
Soit $z \in \mathbb{D}_2'$. Nous avons $\varphi'(z) < 1$ si et seulement si $z_0 = 0$.

Démonstration. D'une part, si $z_0 = 0$, alors

$$\varphi'(z) = \varphi'(0, z_1, \ldots, z_l) = \varphi(z_1, \ldots, z_l) \in [0, 1[. \tag{7.276}$$

D'autre part, si $z_0 \neq 0$, c'est que $z_0 = 1$. Et dans ce cas, nous avons

$$\varphi'(z) = z_0 + \sum_{i=1}^{l} (z) \frac{z_i}{2^i} \geqslant z_0 = 1. \tag{7.277}$$

Cela prouve que si $z_0 \neq 0$, alors $\varphi'(z) \geqslant 1$. □

Toujours avec l'injection φ du lemme 7.275 nous avons ceci.

LEMooVDPWooWLQzbF

Lemme 7.280 ([1]).
Soient $s,t\in\varphi(\mathbb{D}_2)$ tels que $s+t<1$. Alors $s+t\in\varphi(\mathbb{D}_2)$.

Démonstration. Nous considérons $x,y\in\mathbb{D}_2$ tels que $s=\varphi(x)$ et $t=\varphi(y)$. En considérant $z\in\mathbb{D}_2'$ et r comme avant, la proposition 7.278 donne $\varphi(x)+\varphi(y)=\varphi'(z)$. En particulier $\varphi'(z)<1$. Le lemme 7.279 dit alors que $z_0=0$. En posant $z'=(z_1,\ldots,z_l)$, nous avons

$$\varphi(z')=\varphi'(z)=\varphi(x)+\varphi(y). \tag{7.278}$$

Et donc z' est l'élément de $\mathbb{D}_2$ que nous cherchions : $s+t=\varphi(z')$. □

LEMooQYILooMimKHS

Lemme 7.281.
Soient $s,t\in\varphi(\mathbb{D}_2)$ tels que $s+t<1$. Nous posons $x=\varphi^{-1}(s)$, $y=\varphi^{-1}(t)$ et $z=\varphi^{-1}(s+t)$. Si k est le plus petit entier tel que $z_k\neq x_k+y_k$, alors $x_k=y_k=0$ et $z_k=1$.

Démonstration. Il suffit de traiter les 8 cas possible de x_k, y_k et r_k et voir ceux qui vont. Voici le tableau à remplir :

x_k	y_k	r_k	z_k	r_{k-1}
0	0	1		
0	1	1		
1	0	1		
1	1	1		
0	0	0		
0	1	0		
1	0	0		
1	1	0		

Et voici le tableau rempli :

x_k	y_k	r_k	z_k	r_{k-1}
0	0	1	0	0
0	1	1	1	0
1	0	1	1	0
1	1	1	0	1
0	0	0	1	0
0	1	0	0	1
1	0	0	0	1
1	1	0	1	1

Par le lemme 7.277, nous voulons une ligne qui satisfait $z_k\neq x_k+y_k$ et $r_{k-1}=0$. Cherchez un peu et vous verrez qu'il y a une seule ligne, et c'est celle qui donne $x_k=y_k=0$, $r_k=1$, $z_k=1$ et $r_{k-1}=0$. □

THOooAGBXooZnvQLK

Théorème 7.282 (Espace topologique métrisable[219, 1]).
Si V est un espace vectoriel topologique. Nous supposons :

(1) Tout point admet une base dénombrable de topologie[121].

(2) Le singleton $\{0\}$ est fermé[122].

Alors il existe une distance d sur V telle que

(1) d est compatible avec la topologie de V,

(2) d est invariante par translation.

121. Définition 7.2.

122. Attention : dans [219], cela fait partie de la définition d'un espace vectoriel topologique et n'est donc pas listé dans les hypothèses de ce théorème.

Démonstration. Vu que V admet une base dénombrable de topologie, nous en considérons une autour de 0 dans V : $\{A_i\}_{i\in\mathbb{N}}$.

(1) **Nouvelle base de topologie** Nous construisons une nouvelle base de topologie $\{U_i\}_{i\in\mathbb{N}}$ autour de 0 de la façon suivante. La proposition 7.193 donne un ouvert symétrique équilibré dans A_0. Nous le nommons U_0 :

$$U_0 \subset A_0. \tag{7.279}$$

Pour les suivants, U_{k+1} est un ouvert symétrique équilibré dans $U_k \cap A_k$ vérifiant

$$U_{k+1} + U_{k+1} + U_{k+1} + U_{k+1} \subset A_k \cap U_k. \tag{7.280}$$

Le fait est que $A_k \cap U_k$ est un ouvert autour de 0 ; donc la proposition 7.193 permet de construire un ouvert symétrique et équilibré U_{k+1} autour de 0 ayant la propriété demandée.

(2) **C'est bien une base** Notons d'abord que $U_k \subset A_k$. Si $\mathcal{O}$ est un ouvert autour de 0, il existe k tel que $A_k \subset \mathcal{O}$: c'est la définition 7.2 d'une base de topologie. Nous avons donc $U_k \subset A_k \subset \mathcal{O}$, et les $\{U_k\}$ forment une base dénombrable de la topologie autour de 0.

(3) **Quelques inclusions** Notons au passage quelques inclusions. Pour tout $n \in \mathbb{N}$ nous avons

$$U_{n+1} + U_{n+1} + U_{n+1} + U_{n+1} \subset U_n. \tag{7.281}$$

À fortiori nous avons

$$U_{n+1} \subset U_n \tag{7.282}$$

et pour tout $n, k \in \mathbb{N}$, nous obtenons alors

$$U_{n+1} + U_{n+2} + \cdots + U_{n+(k-1)} + U_{n+k} \subset U_{n+1} + U_{n+1} \subset U_n. \tag{7.283}$$

(4) **L'ensemble D** Nous considérons l'ensemble $\mathbb{D}_2$ et l'application φ du lemme 7.275. Nous notons $D = \varphi(\mathbb{D}_2) \subset [0,1[$. L'application $\varphi\colon \mathbb{D}_2 \to D$ est injective, et nous pouvons donc parler de $\varphi^{-1}(r)$ pour tout $r \in D$.

(5) **La fonction ϕ**
Nous définissons maintenant une fonction à valeurs dans la topologie τ_V de V :

$$\begin{aligned} \phi\colon D \cup [1,+\infty[&\to \tau_V \\ r &\mapsto \begin{cases} V & \text{si } r \geqslant 1; \\ \sum_i \varphi^{-1}(r)_i U_i & \text{si } r \in D. \end{cases} \end{aligned} \tag{7.284}$$

La dernière somme est toujours une somme finie et est un ouvert parce que la multiplication par un scalaire et l'addition sont des ouverts.
Notez aussi que plus r est petit, plus ce sont les grands i qui tendront à avoir $\varphi^{-1}(r) \neq 0$. En effet, pour $r = 1/8$, nous avons $\varphi^{-1}(1/8) = (0,0,0,1)$, et plus généralement pour $r < 1/2^N$, les N premiers termes de $\varphi^{-1}(r)$ seront nuls.
Quelques remarques sur cette fonction.

(5a) **Une égalité facile** Si je ne me trompe pas d'un off-by-one, nous avons

$$\phi(1/2^N) = U_N. \tag{7.285}$$

(5b) **$0 \in \phi(r)$ pour tout r** En effet, $\phi(r)$ n'est jamais vide, c'est toujours un voisinage de 0.

(5c) **$\phi(r) + \phi(s) \subset \phi(r+s)$** Si $r + s \geqslant 1$, c'est clair. Sinon, nous posons $u = \varphi^{-1}(r)$, $v = \varphi^{-1}(s)$ et $w = \varphi^{-1}(r+s)$. Dans la suite, nous prolongeons u et v pour qu'elles aient le même nombre d'éléments que nous notons N.
Deux cas se produisent : soit $w_i = u_i + v_i$ pour tout i soit non.

(5c.iii) **Premier cas** Si $w_i = u_i + v_i$ pour tout i, alors en particulier les u_i et v_i ne peuvent pas être 1 en même temps et nous pouvons séparer clairement les termes de la somme :

$$\phi(r+s) = \sum_i w_i U_i = \sum_i u_i U_i + \sum_i v_i U_i = \phi(r) + \phi(s); \tag{7.286}$$

(5c.iii) **Second cas** Sinon, posons k le plus petit entier tel que $w_k \neq u_k + v_k$. Alors le lemme 7.281 dit que $u_k = v_k = 0$ et $w_k = 1$. Nous avons d'abord

$$\phi(r) = \sum_{i=1}^{k-1} u_i U_i + \sum_{i=k+1}^{n} u_i U_i \subset \sum_{i=1}^{k-1} u_i U_i + U_{k+1} + U_{k+1} \tag{7.287}$$

où nous avons utilisé (7.282) pour tous les derniers termes. De même nous avons :

$$\phi(s) = \sum_{i=1}^{k-1} v_i U_i + \sum_{i=k+1}^{n} v_i U_i \subset \sum_{i=1}^{k-1} v_i U_i + U_{k+1} + U_{k+1}. \tag{7.288}$$

En combinant, et en utilisant (7.281),

$$\phi(r) + \phi(s) = \sum_{i=1}^{k-1} u_i U_i + \sum_{i=1}^{k-1} v_i U_i + \underbrace{U_{k+1} + U_{k+1} + U_{k+1} + U_{k+1}}_{\subset U_k} \tag{7.289a}$$

$$\subset \sum_{i=1}^{k-1} w_i U_i + U_k \tag{7.289b}$$

$$= \sum_{i=1}^{k} w_i U_i \tag{7.289c}$$

$$\subset \phi(r+s). \tag{7.289d}$$

Pour (7.289c) nous avons utilisé le fait que $w_k = 1$. Au final nous avons bien $\phi(r) + \phi(s) \subset \phi(r+s)$.

(5d) ϕ **est croissante** Dans notre contexte « croissante » signifie que si $r < s$ alors $\phi(r) \subset \phi(s)$. Il suffit d'écrire

$$\phi(r) \subset \phi(r) + \phi(s-r) \subset \phi(s). \tag{7.290}$$

(6) **Définition de la distance (enfin !)** Nous définissons l'application

$$\begin{aligned} f\colon V &\to \mathbb{R} \\ x &\mapsto \inf\{r \text{ tel que } x \in \phi(r)\}, \end{aligned} \tag{7.291}$$

et ensuite ce qui va être notre distance :

$$d(x,y) = f(y-x). \tag{7.292}$$

Notons que cet infimum est un réel, et qu'il n'est pas spécialement dans $\varphi(\mathbb{D}_2)$. Nous devons prouver que ce d est une distance invariante par translation et compatible avec la topologie.

(7) d **est invariante par translation** Il s'agit seulement d'écrire $d(x+a, y+a)$ et de remarquer que $(y+a) - (x+a) = y - x$.

$d(x,x) = 0$ Oui, car 0 est dans $\phi(r)$, pour tout r, puisque les U_i sont des voisinages de 0.

(8) $d(x,y) = 0$ **implique** $x = y$ Nous montrons que $f(x) \neq 0$ dès que $x \neq 0$. Rappellez-vous que nous avons posé l'hypothèse que $\{0\}$ est un fermé dans V. Vu que $\{x\}$ est un compact [123], la proposition 7.271 nous indique qu'il existe un voisinage A de 0 qui ne contient pas x.

123. Les singletons sont toujours des compacts dans les espaces topologiques.

Vu que les $\{U_k\}_{k\in\mathbb{N}}$ forment une base de la topologie autour de 0, il existe un n_0 tel que $x \notin U_{n_0}$. Et comme $U_{k+1} \subset U_k$ nous avons

$$x \notin U_k \tag{7.293}$$

pour tout $k \geqslant n_0$. Nous savons aussi que ϕ est croissante et que $\phi(1/2^N) = U_N$. Donc pour tout $\epsilon < \frac{1}{2^k}$, nous avons

$$\phi(\epsilon) \subset U_k \tag{7.294}$$

et donc $x \notin \phi(\epsilon)$. Donc

$$f(x) \geqslant \frac{1}{2^k} > 0 \tag{7.295}$$

la dernière inégalité est stricte et c'est important.

En ce qui concerne la distance, si $d(x,y) = 0$, alors $f(y-x) = 0$. Cela signifie que $y - x = 0$ et donc que $x = y$. Ok.

(9) $d(x,y) = d(y,x)$ Oui, car tous les voisinages considérés sont symétriques. Donc $x - y \in \phi(r)$ si et seulement si $y - x \in \phi(r)$.

(10) $d(x,z) \leqslant d(x,y) + d(y,z)$ Nous supposons que x, y et z sont trois points distincts, de telle sorte que les trois distances soient strictement positives.

Soit $\epsilon > 0$. Par définition des distances comme infimums, et grâce au corolaire 1.446, il existe r et s dans $\varphi(\mathbb{D}_2)$ tels que :

$$d(x,y) < r < d(x,y) + \frac{\epsilon}{2} \quad \text{et} \quad d(y,z) < s < d(y,z) + \frac{\epsilon}{2}. \tag{7.296}$$

En sommant :

$$d(x,y) + d(y,z) < r + s < d(x,y) + d(y,z) + \epsilon. \tag{7.297}$$

Vu que ϕ est croissante, on a $y - x \in \phi(r)$ et $z - y \in \phi(s)$; donc

$$z - x = (y - x) + (z - y) \in \phi(r) + \phi(s) \subset \phi(r + s) \tag{7.298}$$

Ainsi, pour tout $\epsilon > 0$, on a

$$d(x,z) \leqslant r + s < d(x,y) + d(y,z) + \epsilon. \tag{7.299}$$

(11) **Compatibilité avec la topologie** Nous devons montrer que les d-ouverts sont les mêmes que les ouverts de la topologie de V. Nous allons utiliser la proposition 7.2.

(11a) **Les ouverts sont des d-ouverts** Soit un ouvert $\mathcal{O}$ contenant $x \in V$. Alors $\mathcal{O} - x$ est un ouvert contenant 0, et il existe un n tel que

$$U_n \subset \mathcal{O} - x. \tag{7.300}$$

Nous avons alors

$$B(0, \frac{1}{2^n}) \subset U_n \subset \mathcal{O} - x. \tag{7.301}$$

Nous utilisant le lemme 7.274 pour additionner x dans toutes les inclusions :

$$B(x, \frac{1}{2^n}) = B(0, \frac{1}{2^n}) + x \subset \mathcal{O}. \tag{7.302}$$

Les boules forment donc une base de la topologie de V. Donc tous les ouverts de V sont des unions de boules et sont donc des d-ouverts.

(11b) **Les d-ouverts sont des ouverts**

Considérons une boule $B(0,r)$. Il existe un n tel que $B(0, 1/2^n) \subset B(0,r)$ et donc tel que

$$U_n \subset B(0,r). \tag{7.303}$$

En procédant par translations et tout ça, on en déduit que les $\{U_n + x\}_{n\in\mathbb{N}, x\in V}$ forment une base de la d-topologie. Donc les d-ouverts sont des ouverts.

□

7.13.5 Équivalence entre Cauchy et $\tau-$Cauchy

LEMooIAHSooFkXjvr

Lemme 7.283.
Soit un espace vectoriel topologique[124] V *et une distance* $d\colon V\times V\to\mathbb{R}^+$ *compatible*[125] *avec la topologie de* V*. Si* d *est invariante*[126]*, alors les suites de Cauchy pour* d *et les suites* τ*-Cauchy sont les mêmes.*

Démonstration. Nous avons deux implications à prouver.

(1) **Cauchy pour d implique τ-Cauchy** Soit (x_n), une suite de Cauchy dans V pour d, et un voisinage U de 0. Vu que d est compatible avec la topologie de V, il existe une boule ouverte $B(0,\epsilon)$ incluse à U. Soit $N>0$ tel que $m,n>N$ implique $d(x_n,x_m)<\epsilon$. Par invariance de la métrique, nous avons aussi
$$d(0,x_m-x_n)<\epsilon, \tag{7.304}$$
c'est-à-dire $x_m-x_n\in B(0,\epsilon)\subset U$. La suite (x_n) est donc τ-Cauchy.

(2) **τ-Cauchy implique Cauchy pour d** Soit (x_n), une suite τ-Cauchy dans V et $\epsilon>0$. Vu que $B(0,\epsilon)$ est un voisinage de 0 dans V, il existe N tel que $m,n>N$ implique $x_n-x_m\in B(0,\epsilon)$. Cela signifie que $d(0,x_n-x_m)<\epsilon$ et toujours par invariance, que $d(x_n,x_m)<\epsilon$.

□

Tout ceci nous mène à donner une large classe d'espaces vectoriels topologiques sur lesquelles les notions de suites de Cauchy pour une distance et τ-Cauchy coïncident.

THOooGQZSooAmQolf

Théorème-Définition 7.284.
Soit V *un espace vectoriel topologique métrisable*[127]*, alors il admet une métrique* d *compatible avec la topologie telle que une suite dans* V *est de Cauchy pour* d *si et seulement si elle est* τ*-Cauchy.*

Une ***suite de Cauchy*** *dans un espace vectoriel métrique* (E,d) *est une suite* τ*-Cauchy ou de Cauchy pour* d*.*

Démonstration. Soit d une métrique sur V satisfaisant au théorème 7.282. Vu qu'elle est invariante par translation, les suites d-Cauchy sont exactement les suites τ-Cauchy par le lemme 7.283. □

REMooUFQYooUVCCjs

Remarque 7.285.
Même si V est métrisable, si on choisit la métrique n'importe comment, on ne peut rien espérer.

7.286.
Sur les espaces vectoriels topologiques métrisables, nous pouvons donc parler de suite de Cauchy sans préciser si nous parlons de τ-Cauchy ou de d-Cauchy, parce que nous sous-entendons avoir choisi une métrique non seulement compatible avec la topologie, mais également invariante par translation.

Il reste cependant à traiter le cas d'un espace vectoriel topologique non métrisable. Dans ce cas, il n'y a pas de métrique, et la question de l'équivalence des définitions ne se pose pas.

Le théorème suivant donne la complétude de $\mathbb{R}$ et le critère de Cauchy pour les définitions métriques et topologiques usuelles. Lorsqu'on dit que $\mathbb{R}$ est complet, le plus souvent nous parlons de ce théorème, et non de 1.456 qui en est un lemme indispensable mais qui parle de notions différentes, bien que très liées.

THOooNULFooYUqQYo

Théorème 7.287 (Complétude de $\mathbb{R}$, critère de Cauchy[18])**.**
Nous avons :

(1) L'espace métrique $(\mathbb{R},d)$ *est complet (définition 7.263).*

124. Définition 7.170.
125. Définition 7.272.
126. Définition 7.273.
127. Voir la proposition 7.282 pour une condition suffisante.

ITEMooUUFCooIVtGgz

(2) *Une suite dans $\mathbb{R}$ est convergente (définition 7.13) si et seulement si elle est de Cauchy (définition 7.284).*

Démonstration. Tout ce théorème se base sur le fait que la définition de suite de Cauchy dans $(\mathbb{R}, d)$ et de suite convergente dans $(\mathbb{R}, d)$ coïncident avec les définitions correspondantes dans $\mathbb{R}$ vu comme simple corps ordonné (définitions 1.379).

Donc si (x_n) est de Cauchy dans $(\mathbb{R}, d)$, elle est de Cauchy dans le corps ordonné $(\mathbb{R}, \leqslant)$. Donc le théorème 1.456 nous dit que (x_n) est convergente dans $(\mathbb{R}, \leqslant)$. Et donc convergente dans $(\mathbb{R}, d)$.

Toutes les autres affirmations se prouvent de la même manière. □

Si vous n'êtes pas sûr ou si vous ne voulez pas étudier les notations de convergence et de suites de Cauchy dans les corps, vous pouvez simplement recopier la démonstration du théorème 1.456 en remplaçant partout $\mathbb{Q}$ par $\mathbb{R}$, et aussi en remplaçant les $|x - y|$ par $d(x, y)$.

7.288.
Nous pouvons également mettre une structure d'espace métrique sur $\mathbb{C}$ en posant

$$d(z, z') = |z - z'|. \tag{7.305}$$

PROPooOIQKooPtyXgm

Proposition 7.289.
L'espace métrique $(\mathbb{C}, d)$ est complet.

Démonstration. Commençons par nous rendre compte que pour tout $z \in \mathbb{C}$ nous avons $|\operatorname{Re}(z)| \leqslant |z|$. C'est bon ? Vous vous en êtes rendu compte ? Ok. Continuons.

Soit une suite de Cauchy (z_k) dans $\mathbb{C}$ et $\epsilon > 0$. Si nous posons $x_j = \operatorname{Re}(z_j)$, nous avons

$$|x_k - x_l| = |\operatorname{Re}(z_k - z_l)| \leqslant |z_k - z_l|. \tag{7.306}$$

Vu que (z_k) est de Cauchy, il existe un N tel que si $k, l \geqslant N$,

$$|x_k - x_l| \leqslant |z_k - z_l| \leqslant \epsilon. \tag{7.307}$$

Donc la suite des parties réelles converge par la complétude de $(\mathbb{R}, d)$ du théorème 7.287. Notez que le d ici n'est pas tout à fait le même, et que la démonstration fonctionne parce que la distance prise sur $\mathbb{R}$ est la restriction à $\mathbb{R}$ de la distance prise sur $\mathbb{C}$. Notons x la limite de (x_k).

De la même manière la suite des parties imaginaires $y_k = \operatorname{Im}(z_k)$ converge vers un réel que nous notons y. Avec tout cela, la suite z_k converge dans $\mathbb{C}$ vers $x + iy$. En effet pour ϵ donné et pour un k suffisament grand,

$$|z_k - (x + iy)| = \big|\operatorname{Re}(z_k) - x + i(\operatorname{Im}(z_k) - y)\big| \leqslant |x_k - x| + |y_k - y| \leqslant \epsilon. \tag{7.308}$$

□

7.14 Norme, espace vectoriel normé

SECooWKJNooKOqpsx

La valeur absolue est essentielle pour introduire les notions de limite et de continuité pour les fonctions d'une variable. Par exemple nous verrons dans la proposition 9.83 que la fonction $f : \mathbb{R} \to \mathbb{R}$ est continue en a si et seulement si pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$|x - a| \leqslant \delta \Rightarrow |f(x) - f(a)| \leqslant \varepsilon. \tag{7.309}$$

La quantité $|x - a|$ donne la « distance » entre x et a ; la définition de la continuité signifie que pour tout ε, il existe un δ tel que si a et x sont au plus à la distance δ l'un de l'autre, alors $f(x)$ et $f(a)$ ne seront éloignés au plus d'une distance ε.

La valeur absolue, dans $\mathbb{R}$, nous sert donc à mesurer des distances entre les nombres. Les principales propriétés de la valeur absolue sont :

(1) $|x| = 0$ implique $x = 0$,
(2) $|\lambda x| = |\lambda||x|$,
(3) $|x + y| \leqslant |x| + |y|$

pour tout $x, y \in \mathbb{R}$ et $\lambda \in \mathbb{R}$.

Afin de donner une notion de limite pour les fonctions de plusieurs variables, nous devons trouver un moyen de définir les notions de « taille » d'un vecteur et de distance entre deux points de $\mathbb{R}^n$, avec $n > 1$. La notion de « taille » doit satisfaire propriétés analogues à celles de la valeur absolue.

La première notion de « taille » pour un vecteur de $\mathbb{R}^2$ que nous vient à l'esprit est la longueur du segment entre l'origine et l'extrémité libre du vecteur. Cela peut être calculée à l'aide du théorème de Pythagore :

$$\text{taille de } (a, b) = \sqrt{a^2 + b^2}. \tag{7.310}$$

Nous pouvons introduire une notion de distance entre les éléments de $\mathbb{R}^2$ de façon similaire :

$$d\big((a_x, a_y), (b_x, b_y)\big) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}. \tag{7.311}$$

Cette définition a l'air raisonnable ; est-elle mathématiquement correcte ? Peut-elle jouer le rôle de la valeur absolue dans $\mathbb{R}^2$? Est-elle la seule définition possibles de « taille » et distance en $\mathbb{R}^2$?

Nous voulons formaliser les notions de « taille » et de distance dans $\mathbb{R}^n$, et plus généralement dans un espace vectoriel V de dimension finie. Pour cela nous nous inspirons des propriétés de la valeur absolue.

7.14.0.1 Critère de Cauchy

THOooRDYOooJHLfGq

Théorème 7.290 (Bolzano-Weierstrass, thème 33)**.**
Toute suite contenue dans un compact admet une sous-suite convergente.

Démonstration. Nous faisons la preuve par l'absurde en supposant que (x_k) n'admette pas de sous-suite convergente. Soit $a \in K$; aucune sous-suite de (x_k) ne converge vers a. En particulier, il existe un voisinage ouvert $\mathcal{O}_a$ de a et une partie finie I_a de $\mathbb{N}$ tel que $x_k \in \mathcal{O}_a$ seulement pour $k \in I_a$.

Les ouverts $\mathcal{O}_a$ recouvrent K ; nous pouvons en extraire un sous-recouvrement fini (c'est la définition 7.80 de la compactié). Nous avons donc des points $a_1, \dots, a_n$ tels que

$$K \subset \bigcup_{i=1}^{n} \mathcal{O}_{a_i} \tag{7.312}$$

et tels que pour chaque $\mathcal{O}_{a_i}$, nous avons $x_k \in \mathcal{O}_{a_i}$ seulement pour $k \in I_{a_i}$. Bien entendu, toute la suite est dans K et donc dans l'union.

En conclusion, nous avons $\mathbb{N} = \bigcup_{i=1}^{n} I_{a_i}$, ce qui prouve que $\mathbb{N}$ est un ensemble fini. Contradiction avec la proposition 1.121 qui dit que $\mathbb{N}$ est infini. □

CORooBTIPooUYWucb

Corolaire 7.291 ([1])**.**
Une suite dans un compact dont toutes les sous-suites convergentes convergent vers la même limite est convergente.

Démonstration. Soient un espace topologique X, un compact K et une suite (x_k) dans K. Nous supposons que toutes les sous-suites convergentes de (x_k) convergent vers $a \in X$. Nous devons montrer que $x_k \xrightarrow{X} a$.

Supposons que ce ne soit pas le cas. Alors il existe un voisinage V de a tel que pour tout $N \in \mathbb{N}$, il existe k tel que $x_k \notin V$. Cela produit une sous-suite hors de V. Cette sous-suite est encore dans K et possède donc une sous(-sous)-suite convergente (théorème 7.290). Par hypothèse, cette sous-sous-suite doit converger vers a, ce qui est impossible.

Contradiction et le corolaire est prouvé. □

Lemme 7.292.
Une suite de Cauchy[128] *dans un espace vectoriel normé admettant une sous-suite convergente est elle-même convergente vers la même limite.*

Démonstration. Soit (a_n) une suite de Cauchy dans un espace vectoriel normé E et ℓ la limite d'une sous-suite de (a_n). Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que $\|a_m - a_p\| < \epsilon$ dès que $m, p \geqslant N$. Nous allons montrer que si $k > N$ alors $\|a_k - \ell\| < 2\epsilon$. Pour cela nous considérons un $n > N$ tel que $\|a_n - \ell\| \leqslant \epsilon$ et nous calculons

$$\|a_k - \ell\| \leqslant \|a_k - a_n\| + \|a_n - \ell\| \leqslant 2\epsilon. \tag{7.313}$$

□

Dans le cas des espaces de dimension finie, le fait d'être complet est automatique, comme le montre la proposition suivante.

PROPooGJDTooXOoYfw

Proposition 7.293.
Soit $\big(E, \|.\|\big)$ un espace vectoriel normé de dimension finie sur un corps $\mathbb{K}$ qui est complet[129]*. Alors E est complet*[130]*.*

Pour rappel, la complétude de l'espace métrique $\mathbb{R}$ est la proposition 1.406.

Démonstration. Nous considérons une suite de Cauchy (f_n) dans E et si $\{e_\alpha\}$ est une base orthonormée de E nous définissons les coefficients $f_n = \sum_\alpha a_{n\alpha} e_\alpha$. La somme sur α est finie par hypothèse sur la dimension de E.

Nous avons

$$\|f_n - f_m\| = \|\sum_\alpha (a_{n\alpha} - a_{m\alpha}) e_\alpha\| = \sum_\alpha |a_{n\alpha} - a_{m\alpha}|^2. \tag{7.314}$$

Soit $\epsilon > 0$. Il existe N tel que si $m, n > N$ alors $\|f_n - f_m\| < \epsilon$. Avec ces conditions sur N, n et m nous avons

$$\sum_\alpha |a_{n\alpha} - a_{m\alpha}|^2 < \epsilon. \tag{7.315}$$

Pour chaque α nous avons donc $|a_{n\alpha} - a_{m\alpha}| < \sqrt{\epsilon}$.

Autrement dit, pour chaque α, la suite $(a_{n\alpha})_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{K}$ et converge donc dans $\mathbb{K}$. Soit a_α la limite et définissons $f = \sum_\alpha a_\alpha e_\alpha$. Nous avons alors

$$\|f_n - f\| = \|\sum_\alpha (a_{n\alpha} - a_\alpha) e_\alpha\|, \tag{7.316}$$

dont la limite $n \to \infty$ est bien zéro. Donc la suite (f_n) converge vers $f \in E$. L'espace E est alors complet. □

PropContinueCompactBorne

Proposition 7.294.
Soient V et W deux espaces vectoriels normés. Soient K une partie compacte de V et $f : K \to W$ une fonction continue. Alors l'image $f(K)$ est compacte dans W.

Ce résultat est démontré dans un cadre plus général par le théorème 7.220.

Démonstration. Nous allons prouver que $f(K)$ est fermée et bornée.

(1) **$f(K)$ est fermé** Nous allons prouver que si (y_n) est une suite convergente contenue dans $f(K)$, alors la limite est également contenue dans $f(K)$. Dans ce cas, nous aurons que l'adhérence de $f(K)$ est contenue dans $f(K)$ et donc que $f(K)$ est fermé. Pour chaque $n \in \mathbb{N}$, le vecteur y_n appartient à $f(K)$ et donc il existe un $x_n \in K$ tel que $f(x_n) = y_n$. La suite (x_n)

128. Définition 7.262.
129. La définition est 1.379, mais si vous n'avez pas envie de vous embarquer trop loin, dites juste « toutes les suites de Cauchy convergent ». Typiquement c'est $\mathbb{R}$ ou $\mathbb{C}$.
130. Définition 7.263.

ainsi construite est une suite dans le fermé K et possède donc une sous-suite convergente (proposition 7.290). Notons (x'_n) cette sous-suite convergente, et a sa limite : $\lim(x'_n) = a \in K$. Le fait que la limite soit dans K provient du fait que K est fermé.
Nous pouvons considérer la suite $f(x'_n)$ dans W. Cela est une sous-suite de la suite (y_n), et nous avons $\lim f(x'_n) = f(a)$ parce que f est continue. Par conséquent nous avons

$$f(a) = \lim f(x'_n) = \lim y_n. \tag{7.317}$$

Cela prouve que la limite de (y_n) est dans $f(K)$ et par conséquent que $f(K)$ est fermé.

(2) $f(K)$ **est borné** Si $f(K)$ n'est pas borné, nous pouvons trouver une suite (x_n) dans K telle que

$$\|f(x_n)\|_W > n \tag{7.318}$$

EqfxnWeeqn

Mais par ailleurs, l'ensemble K étant compact (et donc fermé), nous avons une sous-suite (x'_n) qui converge dans K. Disons $\lim(x'_n) = a \in K$.
Par la continuité de f nous avons alors $f(a) = \lim f(x'_n)$, et donc

$$\|f(a)\|_W = \lim \|f(x'_n)\|_W. \tag{7.319}$$

La suite $f(x'_n)$ est alors une suite bornée, ce qui n'est pas possible au vu de la condition (7.318) imposée à la suite de départ (x_n).

□

CorFnContinueCompactBorne

Corolaire 7.295.
Si $f\colon K \to \mathbb{R}$ est une application continue où K est une partie compacte d'un espace vectoriel normé, alors $f(K)$ est borné.

Démonstration. En effet, la proposition 7.294 montre que $f(K)$ est compact et donc borné. □

7.15 Espaces métriques

7.15.1 Espaces métrisables

Définition 7.296.
*Un espace topologique est **métrisable** si il est homéomorphe à un espace métrique.*

PROPooKNVUooMbLZoy

Proposition 7.297.
Une fonction séquentiellement continue sur un espace métrisable et à valeurs dans un espace métrique est continue.

Démonstration. Soient E un espace métrique et $\phi\colon X \to (E, d)$ un homéomorphisme. Nous supposons que $f\colon X \to Y$ est séquentiellement continue. Nous considérons l'application $\tilde{f} = f \circ \phi^{-1}$, c'est-à-dire

$$\begin{aligned} \tilde{f}\colon E &\to Y \\ a &\mapsto f\big(\phi^{-1}(a)\big). \end{aligned} \tag{7.320}$$

L'application ϕ^{-1} est continue et donc séquentiellement continue. De plus $\tilde{f}$ est séquentiellement continue. En effet si $a_k \xrightarrow{E} a$, alors

$$\tilde{f}(a_k) = f\big(\phi^{-1}(a_k)\big), \tag{7.321}$$

mais ϕ^{-1} est séquentiellement continue, donc $\phi^{-1}(a_k) \xrightarrow{X} \phi^{-1}(a)$, ce qui signifie que $\phi^{-1}(a_k)$ est une suite convergente dans X et donc

$$\lim_{k\to\infty} \tilde{f}(a_k) = \lim_{k\to\infty} f\big(\phi^{-1}(a_k)\big) = f\big(\phi^{-1}(a)\big) = \tilde{f}(a). \tag{7.322}$$

L'application $\tilde{f}$ est donc séquentiellement continue. Mais étant donné que $\tilde{f}$ est définie sur un espace métrique (E) et à valeurs dans un métrique, elle est continue par la proposition 7.255. L'application $f = \tilde{f} \circ \phi$ est donc continue en tant que composée d'applications continues. □

7.15.2 Fonctions continues

La propriété suivante donne des caractérisations importantes de la continuité dans le cas des espaces métriques.

PropQZRNpMn

Proposition 7.298 (Continuité, ouverts et voisinages et limite[240]).
Soient $f : E \to F$ une application entre espaces métriques et $a \in E$. Alors nous avons équivalence entre les choses suivantes :

ItemCBUoRWJi

(1) f est continue en a,

ItemCBUoRWJii

(2) Pour tout voisinage ouvert W de $f(a)$, il existe un voisinage ouvert V de a tel que $f(V) \subset W$.

ItemCBUoRWJiii

(3) Pour toute boule $W' = B\big(f(a), \epsilon\big)$, il existe une boule $V' = B(a, \delta)$ telle que $f(V) \subset W'$.

ItemCBUoRWJiv

(4) $\forall \epsilon > 0$, $\exists \delta > 0$ tel que $f\big(B(a, \delta)\big) \subset B\big(f(a), \epsilon\big)$.

ItemYNQpikrii

(5) $\lim_{x \to a} f(x) = f(a)$ où la limite est donnée par la définition 7.110,

ItemYNQpikriii

(6) Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $\|x - a\| < \delta$ implique $\|f(x) - f(a)\| < \epsilon$.

La proposition 7.348 nous montrera que ces équivalences tiennent encore lorsque l'espace a une topologie de seminormes.

Démonstration. L'équivalence (1) ⇔ (2) est la définition 7.36. L'équivalence (3) ⇔ (4) est une simple paraphrase.

Montrons (2) ⇒ (3). Si $W' = B\big(f(a), \delta\big)$, nous avons un voisinage V de a tel que $f(V) \subset W'$. L'ensemble V contenant une boule autour de chacun de ses points [131], il en contient un autour de a : $V' = B(a, \delta) \subset V$. A fortiori nous avons $f(V') \subset W$.

Montrons (3) ⇒ (2). Si W est un ouvert autour de $f(a)$, il contient une boule autour de $f(a)$: $B\big(f(a), \epsilon\big) \subset W$. Il existe donc une boule $V' = B(a, \delta)$ telle que $f(V') \subset B\big(f(a), \epsilon\big) \subset W$.

L'équivalence (1) ⇔ (5) est la définition 7.36 de la continuité en un point couplée à l'unicité de la limite due à la proposition 7.113 parce qu'un espace métrique est séparé.

Prouvons (5) ⇒ (6). Soient $\epsilon > 0$ et $V = B\big(f(a), \epsilon\big)$. Étant donné que $f(a)$ est une limite de f pour $x \to a$, il existe un voisinage W de a tel que $f(W) \subset V$. Soit $\delta > 0$ tel que $B(a, \delta) \subset W$; alors si $\|x - a\| < \delta$ nous avons $x \in B(x, \delta) \subset W$ et donc $f(x) \in B\big(f(a), \epsilon\big)$, c'est-à-dire $\|f(a) - f(x)\| < \epsilon$.

Enfin l'implication (2) ⇒ (5) est une réécriture de la définition de la limite en un point. □

Voici un théorème qui parle de fermés emboîtés dans un espace métrique. Le corolaire 7.93 parle du cas $\cap_i A_i = \varnothing$ dans un compact.

ThoCQAcZxX

Théorème 7.299 (Théorème [241]).
Soit (E, d) un espace métrique. Il est complet si et seulement si toute suite décroissante de fermés non vides dont le diamètre tend vers zéro a une intersection qui se réduit à un seul point.

Démonstration. En deux parties.

(1) **Condition suffisante**

Soit $\{F_n\}_{n \in \mathbb{N}}$ une telle suite de fermés emboités. Si nous choisissons des points $x_n \in F_n$, nous obtenons une suite (x_n) de Cauchy et qui est par conséquent convergente vu que l'espace est par hypothèse complet. De plus, pour chaque $N \geqslant n$, la queue de suite $(x_n)_{n \geqslant N}$ est contenue dans F_N et donc converge vers un élément de F_N (parce que ce dernier est fermé). Donc la limite de (x_n) est dans $\bigcap_{n \in \mathbb{N}} F_n$.

De plus cette intersection a diamètre nul parce que le diamètre de $\bigcap_{n \in \mathbb{N}} F_n$ est majoré par tous les diamètres des F_n, lesquels sont arbitrairement petits par hypothèse. Donc l'intersection est réduite a un point.

131. Cela est le théorème-définition 7.117 des ouverts dans un espace métrique, à ne pas confondre avec le théorème 7.8.

(2) **Condition nécessaire**

Soit (x_n) un suite de Cauchy. Nous considérons les ensembles

$$F_n = \overline{\{x_i \text{ tel que } i \geqslant n\}}. \tag{7.323}$$

Le fait que la suite soit de Cauchy implique que $\operatorname{diam}(F_n) \to 0$. Par hypothèse, nous avons alors

$$\bigcap_{n\in\mathbb{N}} F_n = \{a\}. \tag{7.324}$$

Pour s'assurer que a est bien la limite de (x_n), il suffit de remarquer que

$$d(x_n, a) \leqslant \operatorname{diam} F_n \to 0. \tag{7.325}$$

□

PropGULUooNzqZKj

Proposition 7.300.
Soient (X, d) un espace topologique métrique et F un fermé de X. Nous avons $d(x, F) = 0$ si et seulement si $x \in F$.

Démonstration. Si $x \in F$ alors $d(x, F) = 0$ parce que $d(x, x)$ fait partie de l'ensemble sur lequel nous prenons l'infimum.

Si réciproquement $d(x, F) = 0$, cela signifie que pour tout ϵ, il existe $x_\epsilon \in F$ tel que $d(x_\epsilon, x) \leqslant \epsilon$. En prenant $\epsilon = 1/k$ nous construisons une suite (x_k) d'éléments dans F vérifiant $d(x_k, x) = \frac{1}{k}$. Cela signifie que $\lim_{k\to\infty} x_k = x$ par la proposition 7.269(1).

Par la caractérisation séquentielle des fermés (un fermé contient les limites de toutes ses suites, proposition 7.252), la suite (x_k) étant dans F, la limite est dans F. Donc $x \in F$. □

LemooynkH

Lemme 7.301.
Soit A_n une suite décroissante de fermés dans un espace métrique[132] *compact K. Alors*

$$C = \bigcap_{n\in\mathbb{N}} A_n \tag{7.326}$$

est non vide.

Démonstration. Soit (x_n) une suite dans K telle que $x_n \in A_n$. La suite étant contenue dans A_1, et A_1 étant compact (lemme 7.96), elle possède une sous-suite $(y_n = x_{\sigma_1(n)})$ convergente dont la limite est dans A_1 par le théorème de Bolzano-Weierstrass 7.144. Une queue de la suite y_n est dans A_2 et nous considérons donc une sous-suite convergente dans A_2 donnée par

$$z_n = y_{\sigma_2(n)} = x_{\sigma_1\sigma_2(n)}. \tag{7.327}$$

En continuant ainsi nous construisons une suite convergente dans A_k. Nous considérons enfin la suite

$$y_n = x_{\sigma_1\ldots\sigma_n(n)}. \tag{7.328}$$

Pour tout k, une queue de cette suite est une sous-suite de $x_{\sigma_1\ldots\sigma_k(n)}$ et par conséquent cette suite converge dans A_k. La limite de cette suite est donc dans l'intersection demandée. □

Remarque 7.302.
Cette propriété est fausse pour les ouverts. Par exemple

$$\bigcap_{n>1}]0, \frac{1}{n}[= \varnothing. \tag{7.329}$$

132. L'hypothèse métrique provient de l'utilisation de Bolzano-Weierstrass, lequel est vrai pour les espaces séquentiellement compacts, dont les espaces métriques.

LemKIcAbic

Lemme 7.303.
Si K est un compact dans un espace métrique et F un fermé disjoint de K, alors $d(K,F) > 0$.

Démonstration. La fonction

$$\begin{aligned} K &\to \mathbb{R} \\ x &\mapsto d(x,F) \end{aligned} \tag{7.330}$$

est une fonction continue sur K, et donc atteint son minimum par le théorème de Weierstrass 7.146. Soit $x_0 \in K$ un point de K qui réalise ce minimum. Si $d(x_0,F) = 0$, alors on aurait une suite (x_n) dans F qui convergerait vers x_0, mais F étant fermé cela signifierait que x_0 serait dans F, ce qui contredirait l'hypothèse que F et K sont disjoints. □

PropLHWACDU

Proposition 7.304.
Soient (X,d) un espace métrique compact et (u_n) une suite de X telle que

$$\lim_{n\to\infty} d(u_n, u_{n+1}) = 0. \tag{7.331}$$

Alors l'ensemble des points d'accumulation[133] *de (u_n) est connexe.*

Démonstration. Nous notons Γ l'ensemble des points d'accumulation de la suite.

(1) **Γ est compact** Nous notons $A_p = \{u_n \text{ tel que } n \geqslant p\}$ et nous avons

$$\Gamma = \bigcap_{p\in\mathbb{N}} \overline{A_p} \tag{7.332}$$

parce que si $x \in \Gamma$, alors pour tout n, il existe $m > n$ tel que $x_m \in B(x,\epsilon)$, et donc tel que $x \in B(x_m,\epsilon)$. Donc pour tout ϵ et pour tout p, l'intersection $B(x,\epsilon) \cap A_p$ est non vide.
En tant qu'intersection de fermés, Γ est fermé (lemme 7.6). En tant que fermé dans un compact, Γ est compact (lemme 7.96).

(2) **Recouvrement par deux compacts**
Supposons que Γ ne soit[134] pas connexe. Nous pouvons alors considérer S et O, deux ouverts disjoints recouvrant Γ et intersectant tous deux Γ. Nous posons alors

$$A = S \cap \Gamma \tag{7.333a}$$
$$B = O \cap \Gamma, \tag{7.333b}$$

et nous avons évidemment $\Gamma = A \cup B$. Montrons que A est fermé (B le sera aussi par le même raisonnement). Soit une suite d'éléments de $S \cap \Gamma$ convergent dans X. Alors la limite est dans $\bar{\Gamma} = \Gamma$ et donc elle est donc O ou S, mais elle est certainement dans $\bar{S}$. Cependant $\bar{S}$ n'intersecte pas O. En effet si $x \in \bar{S} \cap O$, alors tout voisinage de x intersecterait S, mais il y a des voisinages de x étant inclus dans O parce que O est ouvert ; cela donnerait une intersection entre O et S, ce qui est impossible. Donc la limite n'est pas dans O et donc elle est dans S. Au final la limite est dans $S \cap \Gamma$, ce qui prouve son caractère fermé.
Comme d'habitude, $\Gamma \cap S$ est compact parce que fermé dans un compact[135].

(3) **Décomposition en trois morceaux**
Vu que A et B sont des compacts disjoints, nous avons $d(A,B) = \alpha > 0$ pour un certain α par le lemme 7.303. Nous notons

$$A' = \{x \in X \text{ tel que } d(x,A) < \frac{\alpha}{3}\} \tag{7.334a}$$
$$B' = \{x \in X \text{ tel que } d(x,B) < \frac{\alpha}{3}\} \tag{7.334b}$$

133. Définition 7.34.
134. est-ce qu'il faut vraiment un subjonctif ici ?
135. Lemme 7.96.

Nous avons $A' = \bigcup_{x\in A} B(x, \frac{\alpha}{3})$ et donc en tant qu'union d'ouverts, A' est ouvert (définition de la topologie). Même chose pour B'.

Enfin nous notons

$$K = X\backslash(A' \cup B') \tag{7.335}$$

qui est fermé en tant que complémentaire d'ouvert, et donc compact. Étant donné que $A \subset A'$ et $B \subset B'$, nous avons $K \cap \Gamma = \varnothing$.

L'idée est maintenant de montrer que K contient un point d'accumulation de (u_n).

(4) **Sous-suites de (u_n)**

L'hypothèse sur la suite (u_n) nous indique qu'il existe un N_0 tel que $\forall n \geqslant N_0$,

$$d(u_n, u_{n+1}) < \frac{\alpha}{3}. \tag{7.336}$$ EqIHioHjW

Soient $N > N_0$ et $x_0 \in A$. Étant donné que x_0 est point d'accumulation de la suite, il existe $n_1 > N$ tel que $d(x_0, u_{n_1}) < \frac{\alpha}{3}$. Même chose dans B : nous prenons $y_0 \in B$ et un naturel $n_2 > n_1$ tel que $d(y_0, u_{n_2}) < \frac{\alpha}{3}$. Nous avons $u_{n_1} \in A'$ et $u_{n_2} \in B'$.

Soit n_0 le plus petit naturel supérieur à n_1 tel que $u_{n_0} \notin A'$. Cela existe parce que $u_{n_2} \in B'$ et $B' \cap A' = \varnothing$, mais n_0 n'est pas n_2 lui-même parce que $d(A', B') \geqslant \frac{\alpha}{3}$ alors que nous considérons $n_0, n_1, n_2 > N_0$ et donc pour tous les i entre n_1 et n_2 (compris), $d(u_i, u_{i+1}) < \frac{\alpha}{3}$. Notons qu'ici le strict dans la condition (7.336) est important. Nous avons donc $N_0 < n_1 < n_0 < n_2$.

Nous allons maintenant montrer que u_{n_0} est dans K. C'est fait pour : il est loin en même temps de A' et de B'. En utilisant l'inégalité triangulaire à l'envers, nous avons

$$\begin{aligned} d(u_{n_0}, B) &\geqslant d(u_{n_0-1}, B) - d(u_{n_0-1}, u_{n0}) \\ &\geqslant d(A, B) - d(u_{n_0-1}, A) - d(u_{n_0-1}, u_{n_0}) \\ &\geqslant \alpha - \frac{\alpha}{3} - \frac{\alpha}{3} \\ &= \frac{\alpha}{3}. \end{aligned} \tag{7.337}$$

Pour la dernière inégalité nous avons utilisé le fait que u_{n_0-1} n'est pas dans A'. Bref, nous avons montré que u_{n_0} n'est pas dans B' (dans la définition de ce dernier nous avons bien une inégalité stricte). Vu que par définition u_{n_0} n'est pas non plus dans A', nous avons $u_{n_0} \in K$.

Nous avons montré jusqu'à présent que pour tout $N \geqslant N_0$, il existe un $n_0 \geqslant N$ tel que $u_{n_0} \in K$. Cela nous construit donc une sous-suite (v_n) de (u_n) contenue dans K. En tant que suite dans le compact K, la suite (v_n) admet un point d'accumulation dans K. Ce point est également point d'accumulation de la suite (u_n) complète, ce qui donne un point d'accumulation de (u_n) dans K et donc une contradiction.

Nous concluons que Γ est connexe. □

Encore une petite conséquence sans ambition du théorème de Bolzano-Weierstrass.

PropHNylIAW

Proposition 7.305.
Si (x_n) est une suite dans un compact telle que toute sous-suite convergente ait le même point x comme limite. Alors la suite entière converge vers x.

Démonstration. Supposons que ce ne soit pas le cas. Alors il existe un ϵ tel que pour tout $N > 0$, il existe $n > N$ avec $d(x_n, x) > \epsilon$. Cela nous donne une sous-suite de (x_n) composée d'éléments tous à une distance de x supérieure à ϵ. Nous la nommons (y_n) ; c'est une suite dans un compact qui admet donc une sous-suite convergente (et une telle sous-suite est une sous-suite de (x_n)) dont la limite devrait être x, mais c'est impossible par construction. □

LemGDeZlOo

Lemme-Définition 7.306 ([242])**.**
Soit Ω un ouvert dans un espace métrique E. Il existe une suite (K_n) de compacts tels que

(1) $K_n \subset \Omega$
(2) $\bigcup_{n=0}^{\infty} K_n = \Omega$
(3) $K_n \subset \text{Int}(K_{n+1})$.

Une telle suite de compacts vérifie alors

(1) *Il existe* δ_n *tel que pour tout* $z \in K_n$, $B(z, \delta_n) \subset K_{n+1}$.
(2) *Tout compact de* Ω *est inclus dans* $\text{Int}(K_n)$ *pour un certain* n.

ITEMooBPYPooEMhSmY

Une telle suite de compacts est une ***suite exhaustive*** *de compacts pour* Ω.

Démonstration. Nous considérons les ensembles

$$V_n = \{z \in E \text{ tel que } |z|\} \cup \bigcup_{a \notin \Omega} B(a, \frac{1}{n}), \tag{7.338}$$

et nous définissons $K_n = V_n^c$. Vérifions que ces ensembles vérifient tout ce qu'il faut.

(1) $\underline{K_n \subset \Omega}$ Si $a \notin \Omega$ alors a est dans tous les V_n et donc dans aucun des K_n ; nous avons donc bien $K_n \subset \Omega$.

(2) $\underline{\bigcup_{n=0}^{\infty} K_n = \Omega}$ Nous avons déjà prouvé que $\bigcup_{n=0}^{\infty} K_n \subset \Omega$. Pour avoir l'inclusion dans l'autre sens, soit $z \in \Omega$. Nous prenons $n_1 > |z|$ puis n_2 tel que $B(z, \frac{1}{n_2}) \subset \Omega$. Alors $z \in K_n$ avec $n > \max(n_1, n_2)$. Pour ce choix de n, nous avons $z \in K_n$. Cela prouve que $\Omega \subset \bigcup_{n=0}^{\infty} K_n$.

(3) $\underline{K_n \subset \text{Int}(K_{n+1})}$ Soit $z \in K_n$. L'élément z vérifie $d(z, \Omega^c) \geqslant \frac{1}{n}$. Du coup si nous prenons δ tel que

$$\frac{1}{n+1} < \delta < \frac{1}{n} \tag{7.339}$$

alors $B(z, \delta) \subset K_{n+1}$.

(4) $\underline{\textbf{Les } K_n \textbf{ sont compacts}}$ Enfin, les K_n sont tous compacts. En effet ils sont bornés parce que $K_n \subset B(0, n)$ et ensuite K_n est fermé en tant que complémentaire d'un ouvert (V_n est ouvert en tant qu'union d'ouverts).

Nous passons maintenant aux propriétés, qui sont indépendantes de la façon dont nous avons construit les K_n vérifiant les conditions.

(1) Nous pouvons considérer la fonction $K_n \to \mathbb{R}$ donnée par $z \mapsto d(z, K_{n+1}^c)$. Vu que $K_n \subset \text{Int}(K_{n+1})$, c'est une fonction (continue sur le compact K_n) prenant des valeurs strictement positives. Elle a donc un minimum strictement positif. Si δ_n est plus petit que ce minimum nous avons $B(z, \delta_n) \subset K_{n+1}$ pour tout $z \in K_n$.

(2) D'abord nous avons $\Omega = \bigcup_{n=0}^{\infty} \text{Int}(K_n)$. En effet nous avons

$$\Omega = \bigcup_{n=0}^{\infty} K_n \subset \bigcup_{n=0}^{\infty} \text{Int}(K_{n+1}) \subset \bigcup_{n=0}^{\infty} \text{Int}(K_n). \tag{7.340}$$

L'inclusion dans l'autre sens est facile.

Soit K compact dans Ω. Vu que Ω est l'union des $\text{Int}(K_n)$, nous avons

$$K \subset \bigcup_{n=0}^{\infty} \text{Int}(K_n). \tag{7.341}$$

Cela donne à K un recouvrement par des ouverts dont nous pouvons extraire un sous-recouvrement fini par compacité. Les K_n étant croissants, du recouvrement fini, il suffit de prendre le plus grand (disons K_m) et nous avons $K \subset \text{Int}(K_m)$.

□

Notons qu'avec la suite de K_n telle que construite, le dernier point est réglé en prenant

$$\frac{1}{n+1} < \delta_n < \frac{1}{n}. \tag{7.342}$$

LEMooWRIXooSBHavt

Lemme 7.307 ([243]).
Soient un compact $K \subset \mathbb{R}^d$ ainsi que des ouverts $\{\Omega_i\}_{i=1,\dots,n}$ tels que $K \subset \bigcup_{i=1}^n \Omega_i$.
Il existe des compacts $\{K_i\}_{i=1,\dots,n}$ tels que
— $K_i \subset \Omega_i$,
— $K \subset \bigcup_{i=1}^n K_i$.

Démonstration. Soit $x \in K$. Vu que les Ω_i recouvrent K, il existe un $k(x) \in \{1,\dots,n\}$ tel que $x \in \Omega_{k(x)}$. De plus, vu que $\Omega_{k(x)}$ est ouvert, il existe un voisinage de x contenu dans $\Omega_{k(x)}$ (théorème 7.8). Autrement dit, il existe $r(x) > 0$ tel que

$$\overline{B\big(x, r(x)\big)} \subset \Omega_{k(x)}. \tag{7.343}$$

Vu que l'ensemble $\{B\big(x, r(x)\big)\}_{x\in k}$ est un recouvrement de K par des ouverts, nous pouvons en extraire un sous-recouvrent fini [136]. Soient donc $x_1, \dots, x_m \in K$ tels que

$$K \subset \bigcup_{i=1}^m B\big(x_i, r(x_i)\big). \tag{7.344}$$ EQooETROooCMlJsx

Pour chaque $j = 1, \dots, n$, nous posons

$$A_j = \big\{l \in \{1,\dots,m\} \text{ tel que } k(x_l) = j\big\}. \tag{7.345}$$

Et enfin nous définisssons, pour $j = 1, \dots, n$ les parties

$$K_j = \bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)} \tag{7.346}$$

et il nous reste à prouver que ces ensembles répondent bien à la question.

(1) $\bigcup_{j=1}^n A_j = \{1,\dots,m\}$ **est une union disjointe** Un élément l de $A_i \cap A_j$ devrait vérifier $i = k(x_l) = j$. Si $s \in \{1,\dots,m\}$, alors $s \in A_{k(x_s)}$. Donc oui, l'union des A_j est tout $\{1,\dots,m\}$.

(2) $K_j \subset \Omega_j$ Nous avons

$$K_j = \bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)} \subset \bigcup_{l\in A_j} \Omega_{k(x_l)} = \bigcup_{l\in A_j} \Omega_j = \Omega_j. \tag{7.347}$$

(3) $K \subset \bigcup_{i=1}^n K_i$. Par (7.344), et vu que $\{1,\dots,m\} = \bigcup_{j=1}^n A_j$,

$$\begin{aligned}
K &\subset \bigcup_{i=1}^m B\big(x_i, r(x_i)\big) && \text{(7.348a)}\\
&= \bigcup_{j=1}^n \bigcup_{l\in A_j} B\big(x_l, r(x_l)\big) && \text{(7.348b)}\\
&= \bigcup_{j=1}^n \underbrace{\bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)}}_{K_j} && \text{(7.348c)}\\
&= \bigcup_{j=1}^n K_j. && \text{(7.348d)}
\end{aligned}$$

□ ThoFWXsQOZ

Théorème 7.308 (Tykhonov).
Un produit quelconque d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Nous n'allons donner la preuve que dans le cas d'un produit fini dans le théorème 7.314.

136. C'est la définition 7.80 d'un compact.

7.15.3 Ensembles enchainés

Soit (X, d) un espace métrique.

Définition 7.309.
Une ϵ***-chaine*** *joignant les points a et b de X est une suite finie $(u_0, \dots, u_n)$ dans X telle que $u_0 = a$, $u_n = b$ et pour tout $0 \leqslant i \leqslant n-1$ nous avons $d(u_n, u_{n+1}) \leqslant \epsilon$.*

Une partie A de X est ***bien enchainée*** *si pour tout $\epsilon > 0$ et pour tout $a, b \in A$, il existe une ϵ-chaine joignant a et b dans A.*

Lemme 7.310.
Les rationnels dans $\mathbb{R}$ sont bien enchainés.

Démonstration. Soient p et q des rationnels avec $p < q$, ainsi que $\epsilon > 0$. Le lemme 1.441 nous permet de considérer un rationnel δ vérifiant $0 < \delta < \epsilon$. Et nous définissons les rationnels

$$r_k = p + k\delta. \tag{7.349}$$

Vu que $\mathbb{Q}$ est archimédien [137], il existe K tel que $r_K > q$. D'autre part, $r_0 = p < q$. Donc il existe $N = \max\{k \in \mathbb{N} \text{ tel que } r_k < q\}$.

Nous considérons la chaine $(r_0, \dots, r_N, q)$. Elle débute à $r_0 = p$ et termine à q ; pas de problèmes avec ça. À part pour le dernier pas, nous avons

$$|r_n - r_{n-1}| = \delta < \epsilon, \tag{7.350}$$

donc c'est bien une ϵ-chaine. Il reste à voir $|q - r_N|$. Nous avons $r_N \leqslant q \leqslant r_{N+1}$, et donc

$$0 \leqslant q - r_N \leqslant r_{N+1} - r_N = \delta \leqslant \epsilon. \tag{7.351}$$

Donc ok aussi pour ce dernier pas. □

PROPooBUNOooIvfugn

Proposition 7.311 ([244, 1])**.**
Un espace métrique connexe [138] est bien enchainé.

Démonstration. Soit un espace métrique X et $\epsilon > 0$. La relation $x \sim y$ si et seulement si x et y peuvent être reliés par une ϵ-chaine est une relation d'équivalence.

Soit $x \in X$. Nous prouvons que la classe $[x]$ est ouverte. En effet soit $y \in [x]$, si $z \in B(y, \epsilon)$ nous avons $z \in [y]$, et donc $z \in [x]$. Nous en déduisons que $B(y, \epsilon) \subset [x]$, et donc que $[x]$ est ouvert par le théorème 7.8.

Donc les classes sont des ouverts.

Supposons que X n'est pas bien enchainé. Alors il existe ϵ pour lequel X possède plus qu'une classe d'équivalence. Soit $\{[x_k]\}_{k \in I}$ l'ensemble des classes d'équivalences.

Nous considérons un $i_0 \in I$ quelconque, et nous définissons les ouverts $A = [x_{i_0}]$ et

$$B = \bigcup_{k \in I \setminus \{i_0\}} [x_k]. \tag{7.352}$$

Ce sont deux ouverts disjoints qui recouvrent X qui n'est donc pas connexe. □

PROPooXHTWooZibddZ

Proposition 7.312.
La fermeture d'un ensemble bien enchainé dans un espace métrique compact (X, d) est connexe.

Démonstration. Soit $A \subset X$ un ensemble bien enchainé, et soient $a, b \in \bar{A}$. Nous construisons une suite (u_k) dans A de la façon suivante. Pour chaque $n > 0$ nous prenons $a' \in B(a, \frac{1}{n}) \cap A$ et $b' \in B(b, \frac{1}{n}) \cap A$. Ensuite nous considérons une $\frac{1}{n}$-chaine $\{v_i^{(n)}\}_{i \in I_n}$ dans A entre a' et b'. Ici

137. Proposition 1.399.
138. Définition 7.70.

l'ensemble I_n est fini. La suite (u_k) est simplement construite en mettant bout à bout les éléments $v_i^{(n)}$.

La suite ainsi construite est une suite dans A admettant a et b comme points d'accumulation (les autres points d'accumulation sont également dans $\bar{A}$) et telle que $\lim_{k\to\infty} d(u_k, u_{k+1}) = 0$. Par conséquent la proposition 7.304 nous dit que l'ensemble des points d'accumulation de (u_k) est connexe dans X. Nous le notons $C_{a,b}$.

Si nous fixons $a \in \bar{A}$, alors nous avons

$$\bigcup_{x\in\bar{A}} C_{a,x} = \bar{A}. \tag{7.353}$$

Vu que le membre de gauche est une union de connexes, c'est un connexe par la proposition 7.76. □

CORooSIKCooTncoQm

Corolaire 7.313.
Un espace métrique compact est connexe si et seulement si il est bien enchainé.

Démonstration. Dans le sens direct, c'est la proposition 7.311. Dans l'autre sens, si X est compact, alors X est fermé par le lemme 7.96(2). Et vu qu'il est fermé et bien enchainé, la proposition 7.312 implique qu'il est connexe. □

7.15.4 Produit fini d'espaces métriques

Pour rappel, la distance sur un espace produit est donnée par la définition 7.225.

THOIYmxXuu

Théorème 7.314 ([1]).
Un produit fini d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Démonstration. Soient K_1,..., K_n des compacts et $K = K_1 \times \ldots \times K_n$ le produit muni de sa métrique usuelle de la définition (7.225) (attention : chacun des K_i peut être de dimension infinie) :

$$d(\alpha, \beta) = \max\{d_i(\alpha_i, \beta_i)\} \tag{7.354}$$

où d_i est la distance sur K_i. Si (α_n) est une suite dans K alors la suite $(\alpha_n)_1$ est une suite dans le compact K_1 dont nous pouvons extraire une sous-suite convergente (Bolzano-Weierstrass 7.144). De la sous-suite de α correspondante nous extrayons la sous-suite pour la seconde composante, etc.

En fin de compte nous avons une sous-suite (que nous nommons α également) donc chacune des composantes est convergente. Notez que nous utilisons ici de façon cruciale le fait que nous ayons qu'un nombre fini de facteurs.

Autrement dit, pour chaque $i = 1, \ldots, n$, l'application $p \mapsto (\alpha_p)_i$ est une suite dans K_i, et cette suite converge vers ℓ_i.

Soit $\epsilon > 0$ pour chaque $i = 1, \ldots, n$, il existe $N_i > 0$ tel que si $p > N_i$ alors

$$d_i\big((\alpha_p)_i, \ell_i\big) \leqslant \epsilon. \tag{7.355}$$

En prenant $N = \max_k N_k$ et $n > N$ nous avons

$$d\big(\alpha_n, (\ell_1, \ldots, \ell_n)\big) \leqslant \epsilon. \tag{7.356}$$

Par conséquent de la suite (α) nous avons extrait une sous-suite convergente et la partie « réciproque » de Bolzano-Weierstrass nous assure alors que K est compact.

À l'inverse si un des facteurs n'est pas compact (mettons K_1) alors nous prenons un recouvrement $\{\mathcal{O}_i\}_{i\in I}$ de K_1 par des ouverts duquel il est impossible d'extraire un sous-recouvrement fini. Ensuite nous posons

$$\mathcal{P}_i = \mathcal{O}_i \times K_2 \times \ldots \times K_n, \tag{7.357}$$

qui est un recouvrement de K par des ouverts (de K) d'où aucun sous-recouvrement fini ne peut être extrait. □

Pour la culture générale, il y a bien entendu moyen de faire des produits dénombrables et pire d'espaces métriques.

Définition 7.315 ([245]).
Soient (E_n, d_n) des espaces métriques. Sur l'ensemble produit $E = \prod_{i=1}^{\infty} E_i$ nous définissons la métrique

$$d(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} d'_k(x_i, y_i) \tag{7.358}$$

où $d'_i = \min(d_i, 1)$.

On peut montrer que ce d est bien une distance et que (E,d) devient un espace métrique.

ThoKKBooNaZgoO

Théorème 7.316 (Tykhonov dénombrable[245]).
Un produit dénombrable d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Note : ce résultat est encore valable pour un produit quelconque, c'est le théorème de Tykhonov 7.308.

7.15.5 Équicontinuité

DEFooDHQDooFfIvsX

Définition 7.317 ([1, 246, 247]).
*Soient un espace topologique X et un espace vectoriel topologique Y. Une famille H d'applications $X \to Y$ est **équicontinue** en $a \in X$ si pour tout voisinage V de 0 dans Y, il existe un voisinage U de a dans X tel que*

$$h(U) \subset h(a) + V \tag{7.359}$$

pour tout $h \in H$.

Nous disons que H est équicontinue si elle est équicontinue en tout point.

LEMooMIHJooUhvPgM

Lemme 7.318 ([1]).
Soient un espace métrique X, un espace vectoriel normé Y ainsi qu'une famille H d'isométries linéaires $X \to Y$. Alors H est équicontinue.

Démonstration. Nous suivons la définition 7.317 de l'équicontinuité. Soient $a \in X$ et un voisinage V de 0 dans Y. Nous considérons $r > 0$ tel que $B(0,r) \subset V$, et nous posons $U = B(a,r)$.

Si $x \in U$ et $h \in H$, nous avons

$$\|h(x) - h(a)\| = \|h(x-a)\| = d(x,a) < r, \tag{7.360}$$

de telle sorte que $h(x) \in h(a) + B(0,r)$.

Donc H est équicontinue en a. Vu que a est arbitraire, H est équicontinue en tout point et donc équicontinue sur X. □

LEMooKEMRooYyqsBl

Lemme 7.319 ([247]).
Soit une famille de fonctions $f_i : X \to E$ indexée par un ensemble I où X est un espace topologique et E un espace métrique. Cette famille est équicontinue[139] *en $x \in X$ si pour tout $\epsilon > 0$, il existe un voisinage V de x tel que*

$$\|f_i(x) - f_i(y)\| < \epsilon \tag{7.361}$$

pour tout i dès que $y \in V$.

La proposition suivante permet de montrer que certaines fonctions définies par une limite sont continues. Ce sera par exemple le cas de la fonction puissance, proposition 12.407.

139. Définition 7.317.

PROPooICNNooAMjcut

Proposition 7.320 ([1, 247]).
Soit une suite équicontinue (f_i) de fonctions qui converge simplement vers f, alors f est continue.

Démonstration. Soit une suite équicontinue $f_i\colon X \to E$ convergeant simplement vers f. Soit $a \in X$. Nous prouvons que f est continue en a. Pour cela nous considérons $\epsilon > 0$ et, conformément à l'hypothèse équicontinuité un voisinage V de a tel que $|f_i(a) - f_i(x)| < \epsilon$ pour tout $x \in V$.

Nous avons la majoration

$$|f(x) - f(a)| \leqslant |f(x) - f_i(x)| + |f_i(x) - f_i(a)| + |f_i(a) - f(a)|. \tag{7.362a}$$

Plusieurs majorations.

— Vu que $f_i \to f$, il existe N_1 tel que $|f(x) - f_i(x)| < \epsilon$ pour tout $i > N_1$.
— De plus, par définition de V, nous avons aussi $|f_i(x) - f_i(a)| \leqslant \epsilon$.
— Vu que $f_i \to f$, il existe N_2 tel que $|f_i(a) - f(a)| < \epsilon$ pour tout $i > N_2$.

Donc en prenant $x \in V$ et $i > \max\{N_1, N_2\}$ nous avons

$$|f(x) - f(a)| \leqslant 3\epsilon. \tag{7.363}$$

□

¡¡ Avertissement/question au lecteur !! 7.321
Je n'ai pas du tout vérifié si le lemme 7.322 est correct. Sinon, il faudra trouver autre chose dans la preuve de la proposition 21.89.

LEMooSXWKooAmqeic

Lemme 7.322 ([1]).
Si A est une partie fermée de $\mathbb{R}^3$, alors $\texttt{proj}_{\mathbb{R}^+}(A)$ est fermé dans $\mathbb{R}$.

7.15.6 Continuité uniforme

DEFooYIPXooQTscbG

Définition 7.323 ([248]).
*Soient deux espaces métriques (E, d) et (E', d'). Une application $f\colon E \to E'$ est **uniformément continue** si pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $d(x, y) \leqslant \delta$ implique $d'\big(f(x), f(y)\big) \leqslant \epsilon$.*

Dans l'uniforme continuité, le α qui fait fonctionner ϵ doit le faire fonctionner pour tous les $x, y \in E$. C'est la différence avec la continuité simple dans laquelle nous pouvons choisir, pour un même ϵ, un δ différent en chaque point.

Nous parlons plus d'uniforme continuité dans la section 12.7.

7.16 Ensembles nulle part denses

Nous allons nous limiter au cas de $\mathbb{R}$, mais je crois que ça se généralise sans trop de peine aux espaces métriques, voire plus. Voir aussi la section 7.19 sur les espaces de Baire.

Définition 7.324.
*Un ensemble est dit **nulle part dense** si il n'est dense dans aucun intervalle.*

*Un ensemble dans $\mathbb{R}$ est de **première catégorie** ou **maigre** si il est une union dénombrable d'ensembles nulle part dense (c'est-à-dire d'ensembles denses sur aucun intervalle).*

ThoQGalIO

Théorème 7.325 (Baire[249]).
Une réunion dénombrable d'ensembles nulle part denses est d'intérieur vide.

Démonstration. Soient $a \in S$ et $\epsilon > 0$. Nous allons trouver un élément dans $B(a, \epsilon)$ qui n'est pas dans S. Nous commençons par choisir $x_1 \in B(a, \epsilon)$ et $r_1 < \frac{\epsilon}{2}$ tel que

$$B(x_1, r_1) \cap A_1 = \varnothing. \tag{7.364}$$

Ensuite nous choisissons $x_2 \in B(x_1, r_1)$ et $r_2 < \epsilon/4$ tel que $B(x_2, r_2) \subset B(x_1, r_1)$ et $B(x_2, r_2) \cap A_2 = \emptyset$. Notons que $B(x_2, r_2) \cap A_1 = \emptyset$ aussi, par construction.

Par récurrence nous construisons une suite d'éléments x_n et de rayons $r_n < \epsilon/2^n$ tels que

(1) $B(x_n, r_n) \cap A_j = \emptyset$ pour tout $j \leqslant n$,
(2) $\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1})$.

Cette suite étant de Cauchy (parce que contenue dans des intervalles emboîtés de rayon décroissant vers zéro), elle converge [140] donc vers un point qui en particulier appartient à $B(a, \epsilon)$. Mais la limite n'est dans aucun des A_n et donc pas dans S. □

7.17 Seminormes

Les principaux espaces topologiques construit avec des seminormes seront les espaces de fonctions de la définition 30.15. Nous verrons également la topologie $*$-faible sur $\mathscr{D}'(\Omega)$ en la définition 30.31.

7.17.1 Seminorme

DefPNXlwmi

Définition 7.326.
*Si E est un espace vectoriel sur le corps $\mathbb{K} = \mathbb{R}, \mathbb{C}$, une **seminorme** sur E est une application $p : E \to \mathbb{R}$ telle que*

ItemSHnimhDii

(1) $p(\lambda x) = |\lambda| p(x)$

ItemSHnimhDiii

(2) $p(x + y) \leqslant p(x) + p(y)$

pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$.

Remarque 7.327.
Deux remarques.

(1) La seule différence avec une norme est qu'une seminorme peut s'annuler en des éléments non-nuls de l'espace.
(2) Une seminorme prend ses valeurs dans $\mathbb{R}$. En particulier pour que p soit une seminorme, il faut que $p(x) < \infty$ pour tout x.

LEMooHTOAooGmRGZL

Lemme 7.328 ([250, 225])**.**
Soit une seminorme [141] p. Nous avons

ITEMooFRCRooZBWTxH

(1) $p(0) = 0$,

ITEMooHVWHooAVZsTf

(2) $|p(x) - p(y)| \leqslant p(x - y)$.

ITEMooXCHBooDoBvpL

(3) $p(x) \geqslant 0$ *pour tout $x \in X$.*

ITEMooQIGMooVNDakF

(4) $\ker(p)$ *est un sous-espace vectoriel.*

Démonstration. En plusieurs parties.

(1) **Pour (1)** Pour la première affirmation, il suffit de poser $\lambda = 0$ dans l'inégalité de définition $p(\lambda x) = |\lambda| p(x)$.
(2) **Pour (2)** Nous avons d'une part $p(x + h) \leqslant p(x) + p(h)$ et d'autre part $p(x) \leqslant p(x + h) + p(-h) = p(x + h) + p(h)$. En isolant $p(x + h) - p(x)$ dans chacune ce des deux inégalités,

$$-p(h) \leqslant p(x + h) - p(x) \leqslant p(h) \tag{7.365}$$

ou encore

$$|p(x + h) - p(x)| \leqslant p(h) \tag{7.366}$$

qui donne le résultat demandé en posant $h = y - x$.

140. Par la proposition 1.406
141. Définition 7.326.

(3) **Pour (3)** Soit $x \in X$. Nous appliquons le point (2) avec $y = 0$, en nous souvenant que $p(0) = 0$: $|p(x) - p(0)| \leqslant p(x - 0)$, c'est-à-dire $0 \leqslant |p(x)| \leqslant p(x)$.

(4) **Pour (4)** Soient $x, y \in \ker(p)$ ainsi que $\alpha, \beta \in \mathbb{K}$. En utilisant les points précédents,

$$0 \leqslant p(\alpha x + \beta y) \leqslant |\alpha| p(x) + |\beta| p(y) = 0. \tag{7.367}$$

Vu que les inégalités commencent par 0 et finissent par zéro, ce sont toutes les égalités, et en particulier $p(\alpha x + \beta y) = 0$.

□

DEFooNFOAooRAUuOk

Définition 7.329 ([251]).
*Soit un espace vectoriel complexe X. Une application $f : X \to \mathbb{C}$ est **dominée** par la seminorme p si pour tout x dans X nous avons $|f(x)| \leqslant p(x)$.*

*Si M est un sous-espace vectoriel de X, nous disons qu'une application $f : M \to \mathbb{R}$ est **dominée par le dessus** par la seminorme p si $f(m) \leqslant p(m)$ pour tout $m \in M$.*

LEMooKTSKooIteFGq

Lemme 7.330 ([251]).
Une application linéaire est dominée par une seminorme si et seulement si sa partie réelle est dominée par le dessus[142].

Démonstration. Soit une application linéaire $f : X \to \mathbb{C}$. Nous la décomposons en parties réelles et imaginaires par

$$f(x) = u(x) + iv(x) \tag{7.368}$$

où u et v sont des application à valeurs réelles.

(1) $\Rightarrow$ Nous supposons que $|f(x)| \leqslant p(x)$ pour tout $x \in X$. Nous avons les majorations

$$u(x) \leqslant |u(x)| \leqslant |f(x)| \leqslant p(x). \tag{7.369}$$

(2) $\Leftarrow$ Nous supposons que $u(x) \leqslant p(x)$, et nous devons prouver que $|f(x)| \leqslant p(x)$. Nous notons $z = f(x)$. Si $z = 0$ nous avons évidemment $|z| \leqslant p(x)$. Nous supposons que $z \neq 0$. Nous avons[143]

$$f(\frac{|z|}{z}x) = \frac{|z|}{z} f(x) = |z|. \tag{7.370}$$

En particulier $f(\frac{|z|}{z}x)$ est réel et est donc donné par u :

$$f(\frac{|z|}{z}x) = u(\frac{|z|}{z}x). \tag{7.371}$$

Nous avons alors le calcul

$$r = f(\frac{|z|}{z}x) = u(\frac{|z|}{z}x) \leqslant p(\frac{|z|}{z}x) = \big|\frac{|z|}{z}\big| p(x) = p(x). \tag{7.372}$$

□

7.17.2 Fonctionnelle de Minkowski

DEFooJQXHooGDaJtW

Définition 7.331 ([225]).
Soient un espace vectoriel X et une seminorme p sur X. Nous posons

$$\mathring{U}_p = \{x \in X \text{ tel que } p(x) < 1\} \tag{7.373a}$$

$$U_p = \{x \in X \text{ tel que } p(x) \leqslant 1\}. \tag{7.373b}$$

142. Définition 7.329.

143. Notez que si on écrit z en coordonnées polaires $z = re^{i\theta}$, le nombre $z/|z|$ est $e^{i\theta}$. Passez par les polaires pour simplifier les notations si cela vous chante, mais remarquez que le théorème 18.62 n'arrive que dans longtemps.

Notez qu'il n'y a aucune raison particulière de croie que U_p serait la fermeture de $\mathring{U}_p$ parce que pour l'instant nous n'avons aucune topologie sur X.

DEFooYQYBooRkAfjG

Définition 7.332 (Fonctionnelle de Minkowski[225]).
*Soient un espace vectoriel X ainsi qu'une partie $A \subset X$. La **fonctionnelle de Minkowski** de A est l'application*

$$\begin{aligned} p_A\colon X &\to \mathbb{R} \\ x &\mapsto \inf\{\lambda > 0 \text{ tel que } x \in \lambda A\}. \end{aligned} \tag{7.374}$$

Nous posons aussi $p_A(x) = \infty$ si l'ensemble est vide.

LEMooEIWEooPgoskd

Lemme 7.333 ([225]).
Soient un espace vectoriel X sur le corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) ainsi qu'une partie $A \subset X$ qui est non vide, absorbante[144] *et absolument convexe*[145]. *Alors*

(1) *La fonctionnelle de Minkowski p_A est une seminorme.*
(2) *Nous avons*[146]
$$\mathring{U}_{p_A} \subset A \subset U_{p_A}. \tag{7.375}$$

*Note : le lemme 7.334 donnera la même inclusion en mieux parce qu'il n'y aura pas besoin de l'*absolue *convexité.*

Démonstration. En plusieurs parties.

C'est une seminorme.

(1) $\underline{p_A(x) < \infty \text{ pour tout } x}$ Nous avons supposé que A est absorbant. Soit $x \in X$. Il existe $\rho > 0$ tel que $\lambda x \in A$ pour tout $|\lambda| < \rho$. En particulier $\rho x \in A$, et donc $x \in \frac{1}{\rho}A$. L'ensemble $\{\lambda > 0 \text{ tel que } x \in \lambda A\}$ est donc non vide parce qu'il contient $1/\rho$. Nous en déduisons que $p_A(x) \leqslant 1/\rho < \infty$.

(2) $\underline{p_A(0) = 0}$ Le lemme 7.188 nous dit que $0 \in A$, de telle sorte que $0 \in \lambda A$ pour tout $\lambda > 0$. Donc
$$p_A(0) = \inf\{\lambda > 0 \text{ tel que } 0 \in \lambda A\} = 0. \tag{7.376}$$

(3) $\underline{\textbf{Condition pour } \xi x \in \lambda A}$ SPITEMooGDCDooMBDsPB Soit $x \in X$. Nous prouvons que $\xi x \in \lambda A$ si et seulement si $x \in \frac{\lambda}{|\xi|}A$. En utilisant le lemme 7.190, nous avons
$$\lambda A = \frac{\lambda}{|\xi|}|\xi|A = \frac{\lambda}{|\xi|}\xi A. \tag{7.377}$$
Donc $\xi x \in \lambda A$ si et seulement si $\xi x \in \frac{\lambda}{|\xi|}\xi A$ si et seulement si $x \in \frac{\lambda}{|\xi|}A$.

(4) $\underline{p_A(\xi x) = |\xi| p_A(x)}$ Soient $x \in X$ et $\xi \in \mathbb{K}$ avec $\xi \neq 0$. Nous avons
$$\begin{aligned}
p_A(\xi x) &= \inf\{\lambda > 0 \text{ tel que } \xi x \in \lambda A\} && (7.378a)\\
&= \inf\{\lambda > 0 \text{ tel que } x \in \frac{\lambda}{|\xi|}A\} && \text{par (3)} \quad (7.378b)\\
&= \inf\{|\xi|\frac{\lambda}{|\xi|} > 0 \text{ tel que } x ß \frac{\lambda}{|x|}A\} && (7.378c)\\
&= |\xi| \inf\{\frac{\lambda}{|\xi|} > 0 \text{ tel que } x \in \frac{\lambda}{|\xi|}A\} && (7.378d)\\
&= |\xi| \inf\{\mu > 0 \text{ tel que } x \in \mu A\} && \text{cf. justif.} \quad (7.378e)\\
&= |\xi| p_A(x). && (7.378f)
\end{aligned}$$
SUBEQooIHJHooWkXRtA

Pour (7.378e). Si λ prend toutes les valeurs strictement positives, alors $\lambda/|\xi|$ aussi.

144. Définition 7.186.
145. Définition 7.187.
146. Avec les notations de 7.331.

(5) $p_A(x+y) \leqslant p_A(x) + p_A(y)$ Soit $\epsilon > 0$. Nous notons $S_A(x) = \{\lambda > 0 \text{ tel que } x \in \lambda A\}$. Vu que $p_A(x)$ et $p_A(y)$ sont définis comme des infimums, il existe $\lambda, \mu > 0$ tels que

$$\lambda \leqslant p_A(x) + \epsilon \tag{7.379a}$$
$$\mu \leqslant p_A(y) + \epsilon \tag{7.379b}$$

avec $x \in \lambda A$ et $y \in \mu A$.
Vu que A est convexe, et vu que

$$\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1, \tag{7.380}$$

pour tout $z, z' \in A$ nous avons

$$\frac{\lambda}{\lambda+\mu}z + \frac{\mu}{\lambda+\mu}z' \in A, \tag{7.381}$$

ou encore : $\lambda z + \mu z' \in (\lambda+\mu)A$. En particulier pour z et z' choisis tels que $\lambda z = x$ et $\mu z' = y$ nous avons $x + y \in (\lambda+\mu)A$. Nous avons alors

$$p_A(x+y) = \inf\{\delta > 0 \text{ tel que } x+y \in \delta A\} \leqslant \lambda + \mu \leqslant p_A(x) + p_A(y) + 2\epsilon. \tag{7.382}$$

Étant donné que ϵ est arbitraire, nous avons prouvé que $p_A(x+y) \leqslant p_A(x) + p_A(y)$.

(6) p_A **est une seminorme** Nous avons prouvé toutes les conditions de la définition 7.326.

$$\mathring{U}_{p_A} \subset A \subset U_{p_A}$$

(1) **Première inclusion** Soit $x \in \mathring{U}_{p_A}$, c'est-à-dire $p_A(x) < 1$, ou encore

$$\inf\{\lambda > 0 \text{ tel que } x \in \lambda A\} < 1. \tag{7.383}$$

Donc il existe $\lambda \in [0,1[$ tel que $x \in \lambda A$. Vu que A est équilibré (ça fait partie d'être absolument convexe) et que $\lambda < 1$, nous avons $\lambda A \subset A$. Bref,

$$x \in \lambda A \subset A. \tag{7.384}$$

Cela prouver que $x \in A$.

(2) **Seconde inclusion** Soit $x \in A$. Nous avons $1 \in \{\lambda > 0 \text{ tel que } x \in \lambda A\}$, et donc

$$p_A(x) = \inf\{\lambda > 0 \text{ tel que } x \in \lambda A\} \leqslant 1, \tag{7.385}$$

ce qui signifie que $x \in U_{p_A}$.

□
LEMooFUXVooJumScW

Lemme 7.334 ([252]).
Soit un espace vectoriel topologique X ainsi qu'une partie convexe et absorbante A. Alors

$$\{x \in X \text{ tel que } p_A(x) < 1\} \subset A \subset \{x \in X \text{ tel que } p_A(x) \leqslant 1\}. \tag{7.386}$$

Démonstration. Nous prouvons les deux inclusions séparément.

(1) **Première inclusion** Soit $x \in X$ vérifiant $p_A(x) < 1$. Nous avons

$$\inf\{\lambda > 0 \text{ tel que } x \in \lambda A\} < 1. \tag{7.387}$$

Soit $\lambda \in]0,1[$ tel que $x \in \lambda A$. En particulier $x/\lambda \in A$. Étant donné que A est convexe et que $0 \in A$ (toute partie absorbante contient 0), le segment joignant 0 à x/λ est contenu dans A. En particulier, nous avons

$$x = \lambda\left(\frac{1}{\lambda}x\right) + (1-\lambda)0 \in A \tag{7.388}$$

(2) **Seconde inclusion** Soit $x \in A$. Nous avons évidemment

$$1 \in \{\lambda > 0 \text{ tel que } x \in \lambda A\}. \tag{7.389}$$

Donc $p_A(x) = \inf\{\lambda > 0 \text{ tel que } x \in \lambda A\} \leqslant 1$.

□
LEMooFWCIooVwGutp

Lemme 7.335 ([225]).
Soit une seminorme q sur l'espace vectoriel X. Alors

(1) $\mathring{U}_q$ est absorbant.
(2) $\mathring{U}_q$ est absolument convexe.
(3) $q = p_{\mathring{U}_q}$.

Démonstration. En plusieurs points.

(1) **$\mathring{U}_q$ est absorbant** Soit $x \in X$. Nous devons montrer l'existence de $\rho_x > 0$ tel que pour tout $|\lambda| < \rho_x$, nous ayons $\lambda x \in \mathring{U}_q$.
Si $q(x) = 0$, alors $q(\lambda x) = |\lambda| q(x) = 0 < 1$ et donc n'importe quel $\rho_x > 0$ fait l'affaire.
Si $q(x) > 0$, alors il suffit de prendre $\rho_x \in]0, 1/q(x)[$. En effet si $|\lambda| < \rho_x$, nous avons

$$q(\lambda x) = |\lambda| q(x) \leqslant \frac{1}{q(x)} q(x) = 1, \tag{7.390}$$

et donc $\lambda x \in \mathring{U}_q$.

(2) **$\mathring{U}_q$ est équilibré** Soient $|\alpha| \leqslant 1$, et $x \in \mathring{U}_q$. Nous avons

$$q(\alpha x) = |\alpha| q(x) \leqslant q(x) < 1. \tag{7.391}$$

Donc $\alpha x \in \mathring{U}_q$.

(3) **$\mathring{U}_q$ est convexe** Soient $x, y \in \mathring{U}_q$ ainsi que $t \in [0, 1]$. Nous avons

$$\begin{aligned} q\big(tx + (1-t)y\big) &\leqslant q(tx) + q\big((1-t)y\big) && \text{(7.392a)}\\ &= |t| q(x) + |1-t| q(y) && \text{(7.392b)}\\ &\leqslant |t| + |1-t| && \text{(7.392c)}\\ &= t + (1-t) && \text{(7.392d)}\\ &= 1. && \text{(7.392e)} \end{aligned}$$

Donc $tx + (1-t)y \in \mathring{U}_q$.

(4) **$q = p_{\mathring{U}_q}$** Soit $x \in X$. Nous avons

$$p_{\mathring{U}_q}(x) = \inf\{\lambda > 0 \text{ tel que } x \in \lambda \mathring{U}_q\} = \inf\{\lambda > 0 \text{ tel que } q(x) < \lambda\} = q(x). \tag{7.393}$$

□
LEMooQSAJooCgPNnD

Lemme 7.336 ([253]).
Soit un espace vectoriel X, ainsi qu'une seminorme p sur X. Nous posons[147]

$$\begin{aligned} B_p(0,1) &= \{x \in X \text{ tel que } p(x) < 1\}\\ \bar{B}_p(0,1) &= \{x \in X \text{ tel que } p(x) \leqslant 1\}. \end{aligned} \tag{7.394}$$

Alors pour tout $\epsilon > 0$ nous avons

$$\begin{aligned} \epsilon B_p(0,1) &= \{x \in X \text{ tel que } p(x) < \epsilon\}\\ \epsilon \bar{B}_p(0,1) &= \{x \in X \text{ tel que } p(x) \leqslant \epsilon\}. \end{aligned} \tag{7.395}$$

147. Nous ne supposons pas de compatibilité particulière entre la topologie de X et p. Donc à priori $\bar{B}_p(0,1)$ n'est pas la fermeture de $B_p(0,1)$. Si ça avait été le cas, nous aurions écrit $\overline{B_p(0,1)}$.

PROPooDDURooJHfHLw

Proposition 7.337 ([253, 254]).
Soit un espace vectoriel topologique X, ainsi qu'une seminorme p. Les faits suivants sont équivalents.

ITEMooCAOXooGNBeSB

(1) p est continue.

ITEMooWWUSooEPkoWU

(2) p est continue en 0.

ITEMooPKFLooTRlKwo

(3) $B_p(0,1)$ est un voisinage de 0.

ITEMooRTFZooUXUVnA

(4) $\bar{B}_p(0,1)$ est un voisinage de 0.

Démonstration. Nous prouvons les implications dans l'ordre cyclique (1) ⇒ (3) ⇒ (4) ⇒ (2) ⇒ (1). Courage, c'est parti.

(1) **(1) ⇒ (3)** Nous avons

$$B_p(0,1) = \{x \in X \text{ tel que } p(x) < 1\} = p^{-1}\big(]-\infty,1[\big). \tag{7.396}$$

Vu que p est continue et que $]-\infty,1[$ est ouvert, la partie $B_p(0,1)$ est un ouvert contenant 0.

(2) **(3) ⇒ (4)** La partie $\bar{B}_p(0,1)$ contient $B_p(0,1)$ qui est un ouvert contenant 0.

(3) **(4) ⇒ (2)** Nous vérifions la condition de la définition 7.36(1). D'abord $p(0) = 0$. Nous considérons donc un ouvert W de $\mathbb{R}$ contenant 0, et nous allons chercher un voisinage V de 0 dans X tel que $p(V) \subset W$.
Soit $r > 0$ tel que $B(0,r) \subset W$. Par hypothèse $\bar{B}_p(0,1)$ est un voisinage de 0. Comme la topologie de X est vectorielle, la partie $\epsilon\bar{B}_p(0,1)$ est également un voisinage de 0 pour tout $\epsilon > 0$. Montrons que

$$p\Big(\frac{r}{2}\bar{B}_p(0,1)\Big) \subset B(0,r) \subset W. \tag{7.397}$$

Soit $x \in \frac{r}{2}\bar{B}_p(0,1)$. Nous avons $p(x) \leqslant r/2$ par le lemme 7.336, et donc $p(x) \in B(0,r) \subset W$.
Bref, en posant $V = \frac{r}{2}\bar{B}_p(0,1)$, nous avons bien $p(V) \subset W$.

(4) **(2) ⇒ (1)** Nous supposons que p est continue en $0 \in X$ et nous devons prouver qu'elle est continue en $a \in X$. Soient $x \in X$, et $r > 0$. Pour prouver la continuité de p en x (définition 7.36(1)), nous devons prouver l'existence d'un voisinage V de x tel que $f(V) \subset B\big(p(x),r\big)$.
Pour $\epsilon > 0$ nous posons

$$U_\epsilon = p^{-1}\big(B(0,\epsilon)\big). \tag{7.398}$$

Cette partie est ouverte parce que p est continue en 0 et que $p(0) = 0$. Nous allons prouver que si ϵ est assez petit, alors $p(x + E_\epsilon) \subset B\big(p(x),r\big)$.
Soit $z \in U_\epsilon$. Nous avons

$$\begin{aligned}
|p(x) - p(x+z)| &\leqslant p\big(x-(x+y)\big) && \text{lem.7.328} && (7.399a)\\
&= p(-z) && && (7.399b)\\
&= p(z) && && (7.399c)\\
&\subset B(0,\epsilon). && && (7.399d)
\end{aligned}$$

Donc $p(x+z) \in B\big(p(x),\epsilon\big)$ et nous avons montré que $p(x+U_\epsilon) \subset B\big(p(x),\epsilon\big)$. Donc en prenant $\epsilon \leqslant r$, c'est bon.

□

7.18 Topologie des seminormes

Soit $(p_i)_{i\in I}$ une famille de seminormes sur E. Il existe plusieurs façons de construire une topologie sur E à partir des seminormes.

7.18.1 Topologie initiale des seminormes

La première topologie qui nous tombe sous la main est celle de la topologie initiale [148] des p_i.
PROPooVUYTooJXkRYW

Proposition 7.338 ([1, 255]).
Soient un espace vectoriel X et une seminorme $p\colon X \to \mathbb{R}$. Soit $x_0 \in X$. En posant

$$B_\epsilon = \{x \in X \text{ tel que } |p(x) - p(x_0)| < \epsilon\}, \tag{7.400}$$

l'ensemble $\mathcal{B} = \{B_\epsilon\}_{\epsilon>0}$ est une base de voisinages de x_0 pour la topologie initiale de p.

Démonstration. Nous nommons τ la topologie initiale de p sur X. Posons $y_0 = p(x_0)$. Une base de voisinages de y_0 (dans $\mathbb{R}$) est $\{B(y_0, \epsilon)\}_{\epsilon>0}$. Donc la proposition 7.46 dit qu'une base de voisinages de x_0 dans X est donnée par

$$\{p^{-1}\Big(B\big(p(x_0), \epsilon\big)\Big)\}_{\epsilon>0}. \tag{7.401}$$

Or nous avons

$$\begin{aligned} p^{-1}\Big(B\big(p(x_0), \epsilon\big)\Big) &= \{x \in X \text{ tel que } p(x) \in B\big(p(x_0), \epsilon\big)\} && \text{(7.402a)}\\ &= \{x \in X \text{ tel que } |p(x) - p(x_0)| < \epsilon\}. && \text{(7.402b)} \end{aligned}$$

□

La topologie initiale des seminormes n'est pas une bonne idée, parce qu'elle ne donne en général pas lieu à une structure d'espace vectoriel topologique [149].

Exemple 7.339 ([255]).
Nous considérons $\mathbb{R}$ avec la seminorme

$$\begin{aligned} p\colon \mathbb{R} &\to \mathbb{R}\\ x &\mapsto |x|. \end{aligned} \tag{7.403}$$

Nous considérons sur $\mathbb{R}$ la topologie initiale de cette seminorme. En vertu de la proposition 7.338, une base de topologie en $x_0 \in \mathbb{R}$ est donnée par les parties

$$B_\epsilon(x_0) = \{x \in \mathbb{R} \text{ tel que } \big||x| - |x_0|\big| < \epsilon\}. \tag{7.404}$$

Considérons la suite constante $x_k = 1$. Le nombre 1 est évidemment une limite de (x_k). Mais en y regardant bien, le nombre -1 est également une limite. En effet une base de voisinages autour de -1 est donnée par

$$B_\epsilon(-1) = \{x \in \mathbb{R} \text{ tel que } \big||x| - 1\big| < \epsilon\}. \tag{7.405}$$

Et donc $1 \in B_\epsilon(-1)$ pour tout ϵ.

Si $\mathbb{R}$ muni de cette topologie était un espace vectoriel topologique, la proposition 7.172 dirait que $-1 + 1 = 0$ serait une limite de $(x_k + x_k)$. La suite $(x_k + x_k)$ est la suite constante égale à 2. Or un voisinage typique autour de 0 est

$$B_\epsilon(0) = \{x \in \mathbb{R}'|x| < \epsilon\}. \tag{7.406}$$

Autant dire que la suite $(x_k + x_k) = 2$ n'y rentre pas souvent. △

148. La topologie minimale rendant les p_i continue, définition 7.44.
149. Espace vectoriel topologique, définition 7.170.

7.18.2 Topologie avec des boules

DEFooROXMooCpmTwa

Définition 7.340.
Soit un espace vectoriel E sur lequel nous avons des seminormes $\{p_i\}_{i\in I}$. Pour chaque $i \in I$ nous posons

$$B_i(x,r) = \{y \in E \text{ tel que } p_i(y-x) < r\}. \tag{7.407}$$

Si J est une partie finie de I, nous considérons les parties

$$B_J(x,r) = \{y \in E \text{ tel que } p_j(y-x) < r \, \forall j \in J\}. \tag{7.408}$$

LEMooCNRNooWCejjZ

Lemme 7.341.
Soit un espace vectoriel E sur lequel nous avons des seminormes $\{p_i\}_{i\in I}$. Pour chaque $a \in E$, nous considérons la translation

$$\begin{aligned} t_a\colon E &\to E \\ x &\mapsto x + a. \end{aligned} \tag{7.409}$$

Pour toute partie finie $J \subset I$ nous avons

$$t_a\big(B_J(x,r)\big) = B_J(x+a,r) \tag{7.410}$$

Démonstration. D'une part si $y \in B_J(x+a,r)$, nous avons $p_j(y-x-a) < r$ pour tout $j \in J$. Dans ce cas, $y - a \in B_j(x,r)$, et nous avons donc bien $y = t_a(y-a) \in t_a\big(B_J(x,r)\big)$.

Pour l'inclusion dans l'autre sens, c'est le même raisonnement. □

DEFooZTKAooWYUyDa

Proposition-Définition 7.342 (Topologie et seminormes[256, 257, 225])**.**
Soient des seminormes $\{p_i\}_{i\in I}$ sur l'espace vectoriel E. Nous posons[150]

EQooNHMNooGJSeOr

$$\tau_B = \{\mathcal{O} \subset E \text{ tel que } \forall x \in \mathcal{O}, \exists J, r \text{ tel que } B_J(x,r) \subset \mathcal{O}\} \cup \{\varnothing\}. \tag{7.411}$$

Dans cette formule, les J sont des parties finies de I.

Pour chaque $i \in I$ et $x \in E$ nous posons

$$\begin{aligned} p_{i,x}\colon E &\to \mathbb{R} \\ y &\mapsto p_i(y-x). \end{aligned} \tag{7.412}$$

Alors

ITEMooQPVRooKyOBPi
(1) *Les parties τ_B forment une topologie sur E.*

ITEMooQIYWooAmPCdI
(2) *Les applications $p_i\colon E \to \mathbb{R}$ sont continues pour (E,τ_B).*

ITEMooYGNWooYhBzSa
(3) *Les applications $p_{i,x}\colon E \to \mathbb{R}$ sont continues pour τ_B.*

ITEMooMUUZooBKBhsH
(4) *Si τ est une topologie sur E dans laquelle tous les $p_{i,x}$ sont continus, alors $\tau_B \subset \tau$.*

ITEMooEGVFooNqDkdN
(5) *L'espace topologique (E,τ_B) est un espace vectoriel topologique localement convexe*[151]*.*

*Cette topologie est la **topologie des seminormes** sur E.*

Démonstration. En plusieurs parties.

(1) **Pour (1)** Vérifier les trois conditions de la définition 7.1. D'abord E et $\varnothing$ sont dans τ_B. Ensuite si $\{\mathcal{O}_s\}_{s\in T}$ sont des éléments de τ_B, nous considérons $x \in \bigcup_{s\in T} \mathcal{O}_s$, et nous montrons que x est contenu dans une boule contenue dans $\bigcup_{s\in T} \mathcal{O}_s$. Nous considérons $s \in T$ tel que $x \in \mathcal{O}_s$. Vu que $\mathcal{O}_s \in \tau_B$, il existe J, r tels que

$$x \in B_J(x,r) \subset \mathcal{O}_s \subset \bigcup_{s\in T} \mathcal{O}_s. \tag{7.413}$$

Enfin pour l'intersection finie, l'intersection $B_J(x,r) \cap B_{J'}(x,r')$ contient la boule $B_{J\cup J'}(x,\delta)$ avec $\delta = \min\{r, r'\}$.

150. Nous ajoutons explicitement l'ensemble vide parce que nous n'avons pas envie de vérifier des propriétés du genre $\forall x \in \varnothing$, blah blah...Mais bon. En principe c'est pas obligatoire parce que l'ensemble vide est obtenu avec $r = 0$.

151. Définition 7.197.

(2) **Pour (2)** Pour montrer que p_i est continue, nous montrons que l'image inverse d'un ouvert est ouverte. Soit $a \in \mathbb{R}$ ainsi que $r > 0$. Nous considérons l'ouvert $B(a,r)$ et nous montrons que $p_i^{-1}\big(B(a,r)\big) \in \tau_B$. Soit $x \in p_i^{-1}\big(B(a,r)\big)$. Nous allons montrer que si δ est assez petit, alors $B_i(x,\delta) \subset p_i^{-1}\big(B(a,r)\big)$.
Soit $y \in B_i(x,\delta)$. Par définition nous avons $p_i(y-x) < \delta$, et par le lemme 7.328,

$$|p_i(y) - p_i(x)| \leqslant p_i(x-y) < \delta. \tag{7.414}$$

D'autre part, vu que $x \in p_i^{-1}\big(B(a,r)\big)$, nous avons $|p_i(x) - a| < r$. Nous considérons $s > 0$ tel que $|p_i(x) - a| < r - s$. Avec tout ça,

$$\begin{aligned} |p_i(y) - a| &\leqslant |p_i(y) - p_i(x)| + |p_i(x) - a| && \text{(7.415a)}\\ &< \delta + r - s && \text{(7.415b)}\\ &< r && \text{(7.415c)} \end{aligned}$$

dès que nous choisissons $\delta < s$.
Nous avons donc montré que $p_i(y) \in B(a,r)$, et donc que $p_i\big(x,\delta\big) \subset B(a,r)$. Donc p_i est continue.

(3) **Les translations sont continues** Nous considérons les translations $t_a\colon E \to E$ données par $t_a(x) = x + a$. Soit un ouvert $\mathcal{O}$ dans E. Nous prouvons que $t_a^{-1}(\mathcal{O}) = t_{-a}(\mathcal{O})$ est ouvert. Pour cela nous considérons un élément $x \in t_{-a}(\mathcal{O})$ et nous prouvons qu'il existe une boule autour de x contenue dans $t_{-a}(\mathcal{O})$. Soit $y \in \mathcal{O}$ tel que $x = y - a$; vu que $\mathcal{O}$ est ouvert, il existe r, J tels que $B_J(y,r) \subset \mathcal{O}$. Nous avons alors

$$t_{-a}\big(B_J(y,r)\big) \subset t_{-a}(\mathcal{O}), \tag{7.416}$$

et donc

$$B_J(y-a,r) = t_{-a}\big(B_J(y,r)\big) \subset t_{-a}(\mathcal{O}). \tag{7.417}$$

(4) **Pour (3)** Nous avons $p_{i,x} = p_i \circ t_{-x}$. Nous avons déjà prouvé que p_i et t_{-x} sont continues.

(5) **Pour (4)** Soit une topologie τ telle que toutes les applications $p_{i,a}$ sont continues. Nous avons

$$B_i(a,r) = \{y \in E \text{ tel que } p_i(y-a) < r\} = \{y \in E \text{ tel que } p_{i,a}(y) < r\} = p_{i,a}^{-1}\big(B(0,r)\big). \tag{7.418}$$

Étant donné que $B(0,r)$ est ouverte et que $p_{i,a}$ est continue, la partie $p_{i,a}^{-1}\big(B(0,r)\big)$ est ouverte. Donc τ contient toutes les boules de la forme $B_i(a,r)$.
Vu que τ est une topologie, elle contient toutes les intersections finies d'ouvert. En particulier si J est une partie finie de I, τ doit contenir toutes les boules de la forme $B_J(a,r)$.

(6) **Pour (5)** Le point (5), ce sont les points suivants.

(7) **L'addition est continue** Nous prouvons que l'application

$$\begin{aligned} a\colon E \times E &\to E\\ (x,y) &\mapsto x + y \end{aligned} \tag{7.419}$$

est continue pour la topologie τ_B. Soit $\mathcal{O} \in \tau_B$; nous devons prouver que la partie

$$a^{-1}(\mathcal{O}) = \{(x,y) \in E \times E \text{ tel que } x + y \in \mathcal{O}\} \tag{7.420}$$

est ouvert dans $E \times E$ (topologie produit). Soit $(x_0, y_0) \in a^{-1}(\mathcal{O})$. Vu que $x_0 + y_0 \in \mathcal{O}$, il existe un r et un $i \in I$ tels que $B_i(x_0 + y_0, r) \subset \mathcal{O}$.
Nous prouvons que si ϵ est assez petit,

$$B_i(x_0,\epsilon) \times B_i(y_0,\epsilon) \subset a^{-1}\Big(B_i(x_0,y_0,r)\Big) \subset a^{-1}(\mathcal{O}). \tag{7.421}$$

Nous avons

$$\begin{aligned} p_i\big((x+y)-(x_0+y_0)\big) &= p_i\big((x-x_0)+(y-y_0)\big) && (7.422a)\\ &\leqslant p_i(x-x_0)+p_i(y-y_0) \qquad \text{def. 7.326(2)} && (7.422b)\\ &\leqslant 2\epsilon. && (7.422c)\end{aligned}$$

En prenant $\epsilon < r/2$ nous avons donc $a(x,y) \in B_i\big(x_0+y_0,r\big) \subset \mathcal{O}$.

(8) **La multiplication est continue** Nous considérons la multiplication

$$\begin{aligned} m\colon \mathbb{K}\times E &\to E\\ (\lambda,x) &\mapsto \lambda x,\end{aligned} \tag{7.423}$$

et nous montrons que c'est une application continue. Soit un ouvert $\mathcal{O}\subset E$, nous considérons $(\lambda_0,x_0)\in m^{-1}(\mathcal{O})$, et nous allons prouver que $\mathbb{K}\times E$ contient un ouvert contenant (λ_0,x_0) et contenu dans $m^{-1}(\mathcal{O})$.
Il existe $r>0$ et $i\in I$ tels que $B_i(\lambda_0x_0,r)\subset\mathcal{O}$. Nous allons trouver une valeur de $\epsilon>0$ telle que l'ouvert $B(\lambda_0,\epsilon)\times B_i(x_0,\epsilon)$ vérifie

$$(\lambda_0,x_0)\in B(\lambda_0,\epsilon)\times B_i(x_0,\epsilon)\subset m^{-1}\big(B_i(\lambda_0x_0,r)\big)\subset m^{-1}(\mathcal{O}). \tag{7.424}$$

Soit $(\lambda,x)\in B(\lambda_0,\epsilon)\times B_i(x_0,\epsilon)$. Nous avons $|\lambda|<|\lambda_0|+\epsilon$ et le calcul suivant :

$$\begin{aligned} p_i(\lambda x-\lambda_0x_0) &= p_i(\lambda x-\lambda x_0+\lambda x_0-\lambda_0x_0) && (7.425a)\\ &= p_i\big(\lambda(x-x_0)+(\lambda-\lambda_0)x_0\big) && (7.425b)\\ &\leqslant |\lambda|p_i(x-x_0)+|\lambda-\lambda_0|p_i(x_0) && (7.425c)\\ &\leqslant \big(|\lambda_0|+\epsilon\big)\underbrace{p_i(x-x_0)}_{<\epsilon}+\underbrace{|\lambda-\lambda_0|}_{<\epsilon}p_i(x_0) && (7.425d)\\ &< \epsilon\big(|\lambda_0|+\epsilon\big)+\epsilon p_i(x_0) && (7.425e)\\ &= \epsilon^2+\epsilon\big(|\lambda_0|+p_i(x_0)\big). && (7.425f)\end{aligned}$$

Il n'y a pas de problèmes pour trouver un $\epsilon>0$ pour que le tout soit plus petit que r.

(9) **Espace vectoriel topologique** Nous avons prouvé que (E,τ_B) était un espace vectoriel topologique.

(10) **Localement convexe en 0** Nous montrons que $B_i(0,r)$ est convexe. Soient $x,y\in B_i(0,r)$ ainsi que $t\in[0,1]$. Nous avons

$$\begin{aligned} p_i\big(tx+(1-t)y\big) &\leqslant |t|p_i(x)+|1-t|p_i(y) && (7.426a)\\ &\leqslant tr+(1-t)r && (7.426b)\\ &= r. && (7.426c)\end{aligned}$$

Donc $tx+(1-t)y\in B_i(0,r)$.

(11) **Localement convexe** Soit $a\in E$, et montrons que $B_i(a,r)$ est convexe. Si t_a est la translation de vecteur a nous avons $B_i(a,r)=t_a\big(B_i(0,r)\big)$. Soient $x,y\in B_i(a,r)$. Nous avons $x-a,y-a\in B_i(0,r)$, et vu que la boule $B_i(0,r)$ est convexe,

$$t(x-a)+(1-t)(y-a)\in B_i(0,r). \tag{7.427}$$

Un tout petit peu de calcul :

$$\begin{aligned} t(x-a)+(1-t)(y-a) &= tx+(1-t)y-ta-(1-t)a && (7.428a)\\ &= tx+(1-t)y-a. && (7.428b)\end{aligned}$$

Cela prouve que $tx+(1-t)y-a\in B_i(0,r)$, et donc que $tx+(1-t)y\in B_i(a,r)$, ce qu'iol fallait.

□

PROPooNAFAooShzixo

Proposition 7.343 ([225, 1]).
Soit un espace vectoriel X ainsi qu'un ensemble de seminormes $\{p_i\}_{i\in I}$. Il existe une unique topologie τ sur X telle que

(1) (X,τ) est un espace vectoriel topologique localement convexe.

(2) Une base de topologie de τ en 0 est donnée par

$$\mathcal{B}_0 = \{B_J(0,\epsilon) \text{ tel que } J \text{ est fini dans } I, \epsilon > 0\}. \tag{7.429}$$

Cette topologie est la topologie τ_B donnée en (7.411).

Démonstration. En ce qui concerne l'existence, la topologie τ_B donnée dans la proposition 7.342 fait l'affaire. Nous démontrons l'unicité. Soit une topologie τ sur X vérifiant les conditions. Nous allons prouver que $\tau = \tau_B$.

En deux inclusions.

(1) $\underline{\tau_B \subset \tau}$ La topologie τ_B a une base formée de parties de la forme $B_J(x,r)$. Nous montrons que ces parties sont dans τ. En utilisant le lemme 7.341 et la proposition 7.172, nous avons

$$B_J(x,r) = B_J(0,r) + x \subset \mathcal{B}_0 + x \subset \tau. \tag{7.430}$$

Vu que τ contient une base de la topologie τ_B, nous avons $\tau_B \subset \tau$.

(2) $\underline{\tau \subset \tau_B}$ Les $B_J(0,r)$ qui forment une base de topologie de τ en 0 sont des éléments de τ_B. Les $B_J(x,r) = B_J(0,r) + x$ forment une base de topologie de τ en x, et font aussi partie de τ_B.

□

PropQPzGKVk

Proposition 7.344.
Soit un ensemble E muni de la topologie des seminormes $\{p_i\}_{i\in I}$. Une suite (x_n) dans E converge vers x si et seulement si pour tout $i \in I$,

$$p_i(x - x_n) \to 0. \tag{7.431}$$

Démonstration. Si la suite (x_n) converge[152] vers x, alors pour tout ouvert $\mathcal{O}$ autour de x, il existe un N tel que si $n \geqslant N$, alors $x_n \in \mathcal{O}$. En particulier pour tout j et pour tout $\epsilon > 0$, il doit exister un $n \geqslant N_j$ tel que $x_n \in B_j(x,\epsilon)$.

Voyons l'implication inverse. Soit $\epsilon > 0$. Pour tout $i \in I$, il existe un N_i tel que $n \geqslant N_i$ implique $p_i(x - x_n) \leqslant \epsilon$. Si $\mathcal{O}$ est un ouvert, il doit contenir une boule du type $B_J(x,r)$ pour un certain ensemble fini $J \subset I$.

En prenant $N = \max\{N_j \text{ tel que } j \in J\}$, nous avons $p_j(x - x_n) \leqslant \epsilon$ pour tout j et donc $x_n \in B_J(x,r)$. □

¡¡ Avertissement/question à la lectrice !! 7.345
Je n'ai pas vérifié si la proposition 7.346 est correcte. D'ailleurs je même pas trouvé l'énoncé ; et j'avoue n'avoir pas trop cherché.

La preuve serait sans doute similaire à ce qu'on a pour le lemme 7.227.

PROPooNWFZooEFZbNW

Proposition 7.346 ([1]).
Soient des espaces vectoriels munis de seminormes $\big(E,\{p_i\}_{i\in I}\big)$ et $\big(F,\{q_j\}_{j\in J}\big)$. Nous posons

$$\begin{aligned} r_{ij}\colon E\times F &\to \mathbb{R} \\ (x,y) &\mapsto \max\big\{p_i(x), q_j(y)\big\}. \end{aligned} \tag{7.432}$$

EQooTFVXooQRZrFV

Alors :

152. Définition 7.13.

(1) *Les r_{ij} sont des seminormes.*

(2) *La topologie induite sur $E \times F$ par ces seminormes est la topologie produit.*

PROPooGXGQooLRTwvH

Proposition 7.347 ([258, 1])**.**
Un espace vectoriel muni de seminormes sur un corps valué est un espace vectoriel topologique[153].

Démonstration. Soit un espace vectoriel E muni des seminormes $\{p_i\}_{i\in I}$. Sa topologie est donnée par la définition 7.342. Sur le corps $\mathbb{K}$, nous avons la topologie métrique 7.182.

(1) **Somme**

Nous commençons par prouver que

$$\begin{aligned} f\colon E\times E &\to E\\ (x,y) &\mapsto x+y \end{aligned} \tag{7.433}$$

est continue. Soit un ouvert $\mathcal{O}$ de E ; nous allons prouver que $f^{-1}(\mathcal{O})$ est ouvert en prouvant qu'il contient une boule ouverte autour de chacun de ses points (théorème 7.8). Notez que $f^{-1}(\mathcal{O}) \subset E\times E$; la topologie sur cet ensemble est celle est seminormes r_{ij} données en (7.432). Nous allons en particulier utiliser la seminorme $q_i = r_{ii}$ donnée par

$$\begin{aligned} q_i\colon E\times E &\to \mathbb{R}\\ (x,y) &\mapsto \max\{p_i(x),p_i(y)\}. \end{aligned} \tag{7.434}$$

EQooPDRPooZMDeoY

Soit $(a,b)\in f^{-1}(\mathcal{O})$. Vu que $a+b\in\mathcal{O}$ et que $\mathcal{O}$ est ouvert, la partie $\mathcal{O}$ contient une boule ouverte autour de $a+b$ (définition 7.342). Soit $i\in I$ et $r>0$ tels que

$$B_i(a+b,r)\subset\mathcal{O}. \tag{7.435}$$

Nous allons prouver qu'il existe un $s>0$ tel que $B_i\big((a,b),s\big)\subset f^{-1}(\mathcal{O})$, et plus précisément que

$$f\Big(B_i\big((a,b),s\big)\Big)\subset B_i(a+b,r). \tag{7.436}$$

À gauche, B_i est la boule dans $E\times E$ pour la seminorme (7.434). Soit $(x,y)\in B_i\big((a,b),s\big)$, c'est-à-dire

$$q_i\big((a,b)-(x,y)\big)\leqslant s. \tag{7.437}$$

EQooKACCooWVoZAu

Pour savoir si $f(x,y)\in B_i(a+b,r)$, nous posons $x=a+h$ et $y=b+k$ et nous calculons

$$\begin{aligned} p_i\big(f(x,y)-(a+b)\big) &= p_i(x+y-a-b) && \text{(7.438a)}\\ &= p_i(h+k) && \text{(7.438b)}\\ &\leqslant p_i(h)+p_i(k) && \text{(7.438c)}\\ &\leqslant 2\max\{p_i(h),p_i(k)\} && \text{(7.438d)}\\ &= 2q_i(h,k) && \text{(7.438e)}\\ &\leqslant 2s \qquad \text{par (7.437).} && \text{(7.438f)} \end{aligned}$$

En posant $s=r/2$, nous avons bien $f(x,y)\in\mathcal{O}$, et donc $f^{-1}(\mathcal{O})$ est un ouvert ; f est alors continue.

(2) **Produit** Nous nommons $\mathbb{K}$ le corps de l'espace vectoriel E. Nous devons voir que l'application

$$\begin{aligned} f\colon \mathbb{K}\times E &\to E\\ (\lambda,x) &\mapsto \lambda x \end{aligned} \tag{7.439}$$

est continue.

153. Définition 7.170.

La topologie sur $\mathbb{K}$ est sa topologie métrique, c'est-à-dire la topologie de son unique seminorme $\lambda \mapsto |\lambda|$. La topologie sur $\mathbb{K} \times E$ est donc celle des seminormes

$$\begin{aligned} q_i \colon \mathbb{K} \times E &\to \mathbb{R} \\ (\lambda, x) &\mapsto \max\{|\lambda|, p_i(x)\}. \end{aligned} \tag{7.440}$$

EQooFSBEooWDsHot

Nous pouvons donc reprendre le même cheminement que celui que nous avons pris pour la somme. Soit un ouvert $\mathcal{O}$ dans E ; nous considérons $(\lambda, a) \in f^{-1}(\mathcal{O})$. Vu que $f(\lambda, a) \in \mathcal{O}$, et que $\mathcal{O}$ est ouvert pour la topologie des $\{p_i\}_{i\in I}$, il existe $i \in I$ et $r > 0$ tel que $B_i\big(f(\lambda, a), r\big) \subset \mathcal{O}$.
Nous allons prouver qu'il existe $s > 0$ tel que

$$f\Big(B_i\big((\lambda, a), s\big)\Big) \subset B_i(\lambda a, r). \tag{7.441}$$

Ici encore, à gauche B_i est la boule pour la seminorme q_i donnée en (7.440). Soit $(\mu, x) \in B_i\big((\lambda, a), s\big)$, c'est-à-dire

$$q_i\big((\mu, x) - (\lambda, a)\big) = \max\{|\lambda - \mu|, p_i(a - x)\} < s. \tag{7.442}$$

En particulier nous avons les deux inégalités

$$\begin{cases} |\lambda - \mu| < s, & \text{(7.443a)} \\ p_i(a - x) < s. & \text{(7.443b)} \end{cases}$$

Nous avons le calcul suivant : SUBEQSooHWMSooBJsRgy

$$\begin{aligned} p_i\big(f(\mu, x), \lambda a\big) &= p_i(\mu x - \lambda a) && \text{(7.444a)} \\ &= p_i(\mu x - \lambda x + \lambda x - \lambda a) && \text{(7.444b)} \\ &\leqslant p_i\big((\mu - \lambda)x\big) + |\lambda| p_i(x - a) && \text{(7.444c)} \\ &= |\mu - \lambda| p_i(x) + |\lambda| p_i(x - a) && \text{(7.444d)} \\ &\leqslant s p_i(x) + |\lambda| s. && \text{(7.444e)} \end{aligned}$$

C'est le moment de chercher une majoration pour $p_i(x)$:

$$p_i(x) = p_i\big(a + (x - a)\big) \leqslant p_i(a) + p_i(x - a) \leqslant p_i(a) + s. \tag{7.445}$$

Nous pouvons continuer la majoration (7.444) tout en ne nous posant pas de questions sur le sens de l'inégalité parce que nous cherchons $s > 0$:

$$\begin{aligned} p_i\big(f(\mu, x), \lambda a\big) &\leqslant s p_i(x) + |\lambda| s && \text{(7.446a)} \\ &\leqslant s\big(p_i(a) + s\big) + |\lambda| s && \text{(7.446b)} \\ &= s^2 + \big(|\lambda| + p_i(a)\big) s. && \text{(7.446c)} \end{aligned}$$

Nous devons prouver l'existence d'un $s > 0$ tel que $s^2 + \big(|\lambda| + p_i(a)\big)s < r$; autrement dit nous devons résoudre l'inéquation

$$s^2 + \big(|\lambda| + p_i(a)\big)s - r < 0. \tag{7.447}$$

Nous sommes en présence d'un polynôme du second degré en s qui vaut $-r < 0$ en $s = 0$. Par continuité, il existe un voisinage de $s = 0$ dans $\mathbb{R}$ sur lequel le polynôme reste strictement négatif. Il suffit de prendre un s positif dans ce voisinage.

□

La proposition suivante est pratiquement une copie de la proposition 7.298.

PropNGjQnqF

Proposition 7.348.
Soit $f \colon \mathbb{R} \to (E, p_i)_{i\in I}$ une application. Nous avons équivalence entre

ItemHNxGMpCi

(1) *la fonction f est continue en $t_0 \in \mathbb{R}$,*

ItemHNxGMpCii

(2) *si W est un voisinage ouvert de $f(t_0)$ il existe un voisinage ouvert V de t_0 (dans $\mathbb{R}$) tel que $f(V) \subset W$,*

ItemHNxGMpCiii

(3) *pour tout $i \in I$ et $\epsilon > 0$ il existe $\delta > 0$ tel que*

$$f\big(B(t_0, \delta)\big) \subset B_i\big(f(t_0), \epsilon\big). \tag{7.448}$$

Démonstration. L'équivalence (1) ⇔ (2) est la définition 7.36.

Prouvons (2) ⇒ (3). Soient $i \in I$ et $\epsilon > 0$. Considérons la boule $B_i\big(f(t_0), \epsilon\big)$, qui est un ouvert de E contenant $f(t_0)$. Il existe donc un ouvert V autour de t_0 tel que $f(V) \subset B_i\big(f(t_0), \epsilon\big)$. En particulier V contient une boule $B(t_0, \delta)$ et nous avons

$$f\big(B(t_0, \delta)\big) \subset f(V) \subset B_i\big(f(t_0), \epsilon\big). \tag{7.449}$$

Prouvons (3) ⇒ (2). Soit W un ouvert autour de $f(t_0)$. Il existe un $i \in I$ et $\epsilon > 0$ tel que $B_i\big(f(t_0), \epsilon\big) \subset W$. Nous avons alors un $\delta > 0$ tel que

$$f\big(B(t_0, \delta)\big) \subset B_i\big(f(t_0), \epsilon\big) \subset W. \tag{7.450}$$

□

Lorsqu'on a un espace E muni d'une quantité dénombrable de seminormes $\{p_k\}_{k\in I}$ nous définissons l'écart [154]

EqAAghiUR
$$d(x, y) = \sup_{k\geqslant 1} \min\Big\{\frac{1}{k}, p_k(x - y)\Big\}. \tag{7.451}$$

Notons que cet écart est invariant par translation au sens où pour tout x, y, h dans E nous avons

$$d(x + h, y + h) = \sup_{k\geqslant 1} \min\Big\{\frac{1}{k}, p_k(x - y)\Big\} = d(x, y). \tag{7.452}$$

7.18.3 Seminorme et métrisable

Définition 7.349 ([225])**.**
*Soit un espace vectoriel topologique X. Une famille $\{p_i\}_{i\in I}$ de seminormes est **séparante** si pour tout $x \in X\backslash\{0\}$, il existe $i \in I$ tel que $p_i(x) \neq 0$.*

PROPooMJEQooHtIyeX

Proposition 7.350.
Si X est un espace topologique dont la topologie est donnée par une famille dénombrable et séparante de seminormes [155], alors il est métrisable [156].

Démonstration. Nous utilisons le théorème 7.282. Nous considérons les seminormes $\{p_i\}_{i\in I}$ où I est dénombrable.

(1) **Base dénombrable** Vu que I est dénombrable, l'ensemble des parties finies de I est dénombrable [157]. Pour $a \in X$, l'ensemble des boules de la forme $B_J(a, 1/n)$ avec $n \in \mathbb{N}$ et J fini dans I est donc dénombrable et forme une base de topologie en a.

(2) **Singleton fermé** Prouvons que $\{0\}$ est fermé. Pour cela nous prouvons que le complémentaire est ouvert. Soit $a \neq 0$. Vu que $\{p_i\}_{i\in I}$ est séparante, il existe $i \in I$ tel que $p_i(a) \neq 0$. L'application p_i est continue par 7.342(2). Donc pour tout $\epsilon > 0$, la partie $p_i^{-1}\Big(B\big(p_i(a), \epsilon\big)\Big)$ est un ouvert contenant a.

154. Dans le cas de $E = \mathscr{D}(K)$, la première seminorme est numérotée à zéro, donc il faudra poser $d(\varphi_1, \varphi_2)$ avec p_{k-1} au lieu de p_k.
155. Topologie des seminormes, proposition 7.342.
156. Métrisable, définition 7.257.
157. Lemme 1.139.

En prenant $0 < \epsilon < p_i(a)$ et en tenant compte du fait que $p_i(0) = 0$,

$$|p_i(0) - p_i(a)| = |p_i(a)| > \epsilon, \tag{7.453}$$

et donc $p_i(0) \notin B\big(p_i(a), \epsilon\big)$ ou encore

$$0 \notin p_i^{-1}\Big(B\big(p_i(a), \epsilon\big)\Big). \tag{7.454}$$

Bref, l'ouvert $p_i^{-1}\Big(B\big(p_i(a), \epsilon\big)\Big)$ contient a mais pas 0.

□

7.18.4 Boules et topologie

PropLOwUvCO

Proposition 7.351 ([250]).
La topologie donnée par les boules

$$B_k(a, r) = \{x \in E \text{ tel que } \forall\, k \leqslant \frac{1}{r}, p_k(x - a) < r\} \tag{7.455}$$

EqCHfYlO

est la même que celle « usuelle » donnée par les seminormes. En disant « la même » nous entendons le fait que les ouverts sont les mêmes : A est ouvert pour une des deux topologies si et seulement si il est ouvert pour l'autre.

Démonstration. Pour cette démonstration nous allons préfixer par d les notions topologiques issues des boules (7.455) et par P celle des seminormes : P-continue, d-ouvert, etc.

D'abord nous avons

$$B(a, r) = \bigcap_{k \leqslant \frac{1}{k}} B_k(a, r). \tag{7.456}$$

EqRIURpQo

Si $\mathcal{O}$ est un d-ouvert, il contient une d-boule autour de chacun de ses points. Or d'après la formule (7.456), une d-boule est une intersection *finie* de P-ouverts et donc est un P-ouvert par définition. Donc $\mathcal{O}$ contient un P-ouvert autour de tous ses points et est donc P-ouvert.

Inversement nous supposons que $\mathcal{O}$ est un P-ouvert. Commençons par prouver que les seminormes p_k sont d-continues. En effet soient $k \in \mathbb{N}$, $\epsilon \leqslant \frac{1}{k}$ et $x, y \in E$ tels que $d(x, y) \leqslant \epsilon$; nous avons

$$\begin{aligned}
|p_k(y) - p_k(x)| &\leqslant p_k(x - y) && \text{(7.457a)}\\
&= \min\{\frac{1}{k}, p_k(x - y)\} && \text{(7.457b)}\\
&\leqslant d(x, y) && \text{(7.457c)}\\
&\leqslant \epsilon. && \text{(7.457d)}
\end{aligned}$$

Montrons à présent que $\mathcal{O}$ est d-ouverte. Si $a \in \mathcal{O}$, il existe k et r tels que $B_k(a, r) \subset \mathcal{O}$. Soit $x \in B_k(a, r)$. Montrons que si ϵ est suffisamment petit, la d-boule $B(x, \epsilon)$ est incluse à $B_k(a, r)$. Pour cela prenons $y \in B(x, \epsilon)$; nous avons

$$\big|p_k(a - x) - p_k(a - y)\big| \leqslant d(x, y) \leqslant \epsilon. \tag{7.458}$$

Par conséquent le nombre $p_k(a - y)$ est dans l'intervalle

$$p_k(a - x) \pm \epsilon \tag{7.459}$$

et il suffit de prendre $\epsilon < \frac{r - p_k(a-x)}{2}$. □

THOooTKWYooYYiBNa

Théorème 7.352 ([259]).
Soit une famille de seminormes $\{p_i\}_{i \in I}$ sur E. Pour une seminorme q sur E, les conditions suivantes sont équivalentes.

ITEMooHKNYooARXiXs

(1) *L'application $q\colon E \to \mathbb{R}$ est continue.*

ITEMooBBNCooGwHrUI

(2) *Il existe une partie finie J de I et un nombre $M \in \mathbb{R}$ tels que*

$$q(x) \leqslant M \max_{j\in J} p_j(x) \tag{7.460}$$

pour tout $x \in E$.

ITEMooWXSWooVfvSUy

(3) *L'application q est bornée dans un voisinage de 0.*

Démonstration. En trois parties.

(1) **(1) ⇒ (3)**
C'est la proposition 7.121.

(2) **(3) ⇒ (2)**
Soit un voisinage V de 0 sur lequel q est borné par μ. Ce voisinage contient une boule centrée en 0. Par la définition 7.342, nous avons une partie finie $J \subset I$ et $r > 0$ tels que q est bornée par μ sur $B_J(0, r)$.
Soit $x \in E$. Nous posons

$$t \in]0, \min_{j\in J} \frac{r}{p_j(x)}[. \tag{7.461}$$

Pour $l \in J$ nous avons alors

$$\begin{aligned} p_l(tx) &= tp_l(x) && (7.462a)\\ &< \min_{j\in J} \frac{r}{p_j(x)} p_l(x) && (7.462b)\\ &< \frac{r}{p_l(x)} p_l(x) && (7.462c)\\ &= r. && (7.462d) \end{aligned}$$

Donc $tx \in B_J(0, r)$. En ce qui concerne q nous avons alors $tq(x) = q(tx) < \mu$ et donc

$$q(x) < \frac{\mu}{t}. \tag{7.463}$$

En utilisant la proposition 1.451,

$$\begin{aligned} q(x) &< \mu \frac{1}{t} && (7.464a)\\ &= \mu \frac{1}{\min_{j\in J} \frac{r}{p_j(x)}} && (7.464b)\\ &= \mu \max_{j\in J} \frac{p_j(x)}{r} && (7.464c)\\ &= \frac{\mu}{r} \max_{j\in J} p_j(x). && (7.464d) \end{aligned}$$

En posant $M = \mu/r$ nous avons

$$q(x) \leqslant M \max_{j\in J} p_j(x), \tag{7.465}$$

comme demandé.

(3) **(2) ⇒ (1)**
Nous avons une partie finie $J \subset I$ et $M \in \mathbb{R}$ tels que $q(x) \leqslant M \max_{j\in J} p_j(x)$ pour tout $x \in E$.
Soit $x \in E$. Nous prouvons que q est continue en x. Pour tout y nous avons

$$|q(y) - q(x)| \leqslant q(x - y) \leqslant M \max_{j\in J} p_j(x - y). \tag{7.466}$$

Soit $\epsilon > 0$. Si nous prenons $y \in B_J(x, \epsilon)$, alors nous avons

$$|q(y) - q(x)| \leqslant M\epsilon. \tag{7.467}$$

Cela prouve que q est continue en x, et donc continue sur E par la proposition 7.37.

□

CORooSLUCooRaZyYK

Corolaire 7.353 ([259]).
Soient une famille de seminormes P sur E, ainsi qu'une famille Q de seminormes sur F. Une application linéaire $f : E \to F$ est continue si et seulement si $q \circ f$ est une seminorme continue sur E pour tout $q \in Q$.

CORooGMVHooNgYOaY

Corolaire 7.354 ([259]).
Soient des seminormes $P = \{p_i\}_{i\in I}$ sur E. Soit une application linéaire $f : E \to \mathbb{K}$. Elle est continue si et seulement si il existe une partie finie $J \subset I$ et $M \in \mathbb{R}$ tels que

$$|f(x)| < M \max_{j\in J} p_j(x) \tag{7.468}$$

pour tout $x \in E$.

7.18.4.1 Norme induite sur la topologie quotient

PROPooDUAVooEfrEGI

Proposition-Définition 7.355 (Norme quotient[260]).
Soient un espace vectoriel topologique normé E, et un sous-espace M. Pour $\alpha \in E/M$ nous posons

$$\|\alpha\|_{E/M} = d(\alpha, M) \tag{7.469}$$

où $d(\alpha, M)$ est la distance entre la partie α et la partie M.

Nous avons :

ITEMooQOZXooZiZdTn
(1) La formule $\|\alpha\| = \inf_{u\in\alpha} \|u\|_E$.

ITEMooFEOGooTTDgCj
(2) L'opération $\|.\|$ est une seminorme[158] *sur l'espace vectoriel E/M.*

ITEMooWCCSooMeqQHL
(3) C'est une norme si et seulement si M est fermé.

*Nous parlons de **(semi)norme quotient**.*

Démonstration. Point par point.

(1) **Pour (1)** La définition de $\|\alpha\|$ est $\|\alpha\| = d(\alpha, M) = \inf_{\substack{u\in\alpha\\ v\in M}} \|u - v\|$. Mais

$$\{u - v \text{ tel que } u \in \alpha, v \in M\} = \alpha, \tag{7.470}$$

donc

$$d(\alpha, M) = \inf_{u\in\alpha} \|u\|. \tag{7.471}$$

(2) **Pour (2)** Nous devons vérifier les propriétés de la définition 7.326. D'abord en tant que distance, nous avons $\|\alpha\| \geqslant 0$ pour tout α.
Si $\lambda \in \mathbb{R}$ nous avons aussi

$$\|\lambda\alpha\| = \inf_{u\in\lambda\alpha} \|u\| = \inf_{u\in\alpha} \|\lambda u\| = |\lambda| \inf_{u\in\alpha} \|u\| = |\lambda|\|\alpha\|. \tag{7.472}$$

Enfin si $\alpha, \beta \in E/M$, nous avons

$$\inf_{u\in\alpha+\beta} \|u\| \leqslant \inf_{\substack{u\in\alpha\\ v\in\beta}} \|u + v\| \leqslant \inf_{\substack{u\in\alpha\\ v\in\beta}} \big(\|u\| + \|v\|\big) = \inf_{u\in\alpha} \|u\| + \inf_{v\in\beta} \|v\| = \|\alpha\| + \|b\|. \tag{7.473}$$

158. Définition 7.326.

(3) **Pour (3) en supposant que M est fermé** Nous supposons que M est fermé et nous montrons que $\|.\|$ est une norme. Nous supposons donc que $d(\alpha, M) = 0$ et nous prouvons que $\alpha = 0$. Soit $x \in E$ tel que $\alpha = [x]$. Nous avons donc

$$\alpha = \{x - v \text{ tel que } v \in M\} \tag{7.474}$$

Notez qu'on a écrit $-v$ et non $+v$. De toutes façons M est vectoriel ; ça ne change rien et ça tombera mieux plus bas.

Nous avons donc

$$0 = \|\alpha\| = d(\alpha, M) = \inf_{u\in\alpha} \|u\| = \inf_{v\in M} \|x - v\|. \tag{7.475}$$

Il existe donc une suite (v_n) dans M telle que $\|x - v_n\| \to 0$. La suite est donc convergente : $v_n \to x$. Comme M est fermé, la proposition 7.252 nous indique que la limite doit être dans M. Autrement dit : $x \in M$. Par définition des classes nous avons alors $\alpha = [x] = 0$.

(4) **Pour (3) en supposant que $\|.\|$ est une norme**
Supposons que M n'est pas fermé. Il ne contient donc pas son adhérence. Soit $a \in \text{Adh}(M)\backslash M$. Vu que M est vectoriel, nous supposons que $a \neq 0$.

Étant donné que a est dans l'adhérence de M nous avons $d(a, M) = 0$ et donc $\|[a]\| = 0$.

□

¡¡ Avertissement/question à la lectrice !! 7.356
Je ne suis pas certain de la proposition 7.357. Peut-être qu'il faut ajouter l'hypothèse que M est fermé.

PROPooKLXSooSOUZkc

Proposition 7.357.
Soient un espace vectoriel topologique normé E, et un sous-espace M. La topologie quotient[159] *sur E/M est la même que celle de la seminorme induite*[160].

7.18.5 Espace dual

Nous parlerons plus en détail d'espace dual d'un espace normé en la section 10.15.

DefHUelCDD

Lemme-Définition 7.358.
Soient F un espace métrique et E un espace topologique vectoriel. Pour chaque $v \in E$, l'application

$$\begin{aligned} p_v \colon \mathcal{L}(E, F) &\to \mathbb{R} \\ T &\mapsto \|T(v)\|_F \end{aligned} \tag{7.476}$$

est une seminorme.

*La **topologie $*$-faible** sur $\mathcal{L}(E, F)$ est la topologie des ces seminormes.*

7.359.
C'est une famille de seminormes indicées par les éléments de E. Si E est un espace métrique, c'est cette topologie qui sera considérée sur son dual topologique E' des applications continues $E \to \mathbb{R}$.

La topologie ainsi définie est, dans l'idée, celle qui sera choisie pour les espaces de distributions, voir la définition 30.31.

La proposition suivante indique qu'elle est un peu la topologie de la convergence ponctuelle.

PROPooQWOOooUBhBiG

Proposition 7.360.
Soient E un espace muni de la topologie des seminormes $\{p_i\}_{i\in I}$ et F un espace métrique. Soient une suite (T_n) dans $\mathcal{L}(E, F)$ et $T \in \mathcal{L}(E, F)$. Nous avons $T_n \xrightarrow{*} T$ si et seulement si $T_n(v) \xrightarrow{F} T(v)$ pour tout $v \in E$.

159. Topologie quotient, définition 7.49.
160. Voir la définition 7.355.

Démonstration. Nous avons équivalence entre les lignes suivantes :

$$T_n \xrightarrow{*} T \tag{7.477a}$$
$$p_v(T_n - T) \to 0 \, \forall v \in E \qquad \text{proposition 7.344} \tag{7.477b}$$
$$\|T_n(v) - T(v)\|_F \to 0 \, \forall v \in E \tag{7.477c}$$
$$T_n(v) \xrightarrow{E} T(v). \tag{7.477d}$$

□

7.18.6 Espace $C^k(\mathbb{R}, E')$

Nous revenons à nos histoires de limites de la définition 7.13.

PropRBCiHbz

Proposition 7.361 (Unicité de la limite dans un dual topologique)**.**
Soient E un espace métrique et E' son dual topologique muni de sa topologie de la définition 7.358. Il y a unicité de l'élément de E' vers lequel une fonction $u\colon \mathbb{R} \to E'$ peut converger.

Démonstration. Soit T un élément vers lequel u_t converge lorsque $t \to t_0$. Soient $\epsilon > 0$ et $x \in E$. La boule $B_x(T, \epsilon)$ de E' subordonnée à la norme p_x et centrée en T est un ouvert de E'. Étant donné que u converge vers T il existe $\delta > 0$ tel que $u_t \in B_x(T, \epsilon)$ dès que $|t - t_0| \leqslant \delta$. Nous avons donc, pour tout $x \in E$, la limite (dans $\mathbb{R}$) :

$$\lim_{t \to t_0} u_t(x) = T(x). \tag{7.478}$$

Cela prouve que la convergence de u vers T implique l'existence pour tout x de la limite de $u_t(x)$ dans $\mathbb{R}$. Si T' est un autre élément vers lequel u_t converge, nous avons par le même raisonnement que

$$\lim_{t \to t_0} u_t(x) = T'(x). \tag{7.479}$$

Par unicité de la limite dans $\mathbb{R}$ nous devons alors avoir $T(x) = T'(x)$ pour tout x, c'est-à-dire $T = T'$. □

PropVKSNfiB

Proposition 7.362.
Soit $u\colon \mathbb{R} \to E'$ une fonction continue. Alors

ItemLSJjfZdi
(1) pour tout $x \in E$ la fonction $t \mapsto u_t(x)$ est continue,

ItemLSJjfZdii
(2) pour tout $x \in E$ nous avons la limite dans $\mathbb{R}$

EqWKdEPVQ
$$\lim_{t \to t_0} u_t(x) = u_{t_0}(x), \tag{7.480}$$

ItemLSJjfZdiii
(3) nous avons la limite dans E'

$$\lim_{t \to t_0} u_t = u_{t_0}. \tag{7.481}$$

Démonstration. Soient $x \in E$ et $\epsilon > 0$. Par la proposition 7.348 la continuité de u donne un $\delta > 0$ tel que

$$u_{B(t_0, \delta)} \subset B_x(u_{t_0}, \epsilon). \tag{7.482}$$

C'est-à-dire que si $|t - t_0| \leqslant \delta$ nous avons

$$\big|u_{t_0}(x) - u_t(x)\big| < \epsilon, \tag{7.483}$$

ce qui signifie bien que la fonction $t \mapsto u_t(x)$ est continue en tant que fonction $\mathbb{R} \to \mathbb{R}$. Cela est le point (1). Le théorème de limite et continuité dans $\mathbb{R}$ nous donne immédiatement la limite (7.480).

Nous passons à la preuve du point (3). Soit $\mathcal{O}$ un ouvert de E' contenant u_{t_0}. Il existe donc un $i \in I$ et $\epsilon > 0$ tel que $B_i(u_{t_0}, \epsilon) \subset \mathcal{O}$. Étant donné que u est continue, il existe $\delta > 0$ tel que

$$u_{B(t_0, \delta)} \subset B_i(u_{t_0}, \epsilon) \subset \mathcal{O}. \tag{7.484}$$

Cela signifie bien que

$$|t - t_0| \leqslant \delta \Rightarrow u_t \in \mathcal{O}, \tag{7.485}$$

c'est-à-dire que nous avons la limite $\lim_{t\to t_0} u_t = u_{t_0}$ dans E'. Pour dire cela nous avons utilisé la définition 7.110 de la limite et le résultat d'unicité 7.361. □

DefDZsypWu

Définition 7.363.
*Si nous avons une application $u\colon \mathbb{R} \to E'$ nous considérons sa **dérivée** donnée par la limite*

$$u'_{t_0} = \lim_{t\to t_0} \frac{u_t - u_{t_0}}{t - t_0}. \tag{7.486}$$

Cela est un nouvel élément de E' (pour peu que la limite existe). La fonction $u'\colon \mathbb{R} \to E'$ ainsi définie peut être continue ou non. Cela nous permet de définir les espaces $C^k(\mathbb{R}, E')$ et $C^\infty(\mathbb{R}, E')$.

Une des principales utilisations que nous ferons de ces espaces seront les espaces de fonctions à valeurs dans les distributions tempérées dont nous parlerons dans la section 30.4.

7.19 Espaces de Baire

DEFooYEMNecBSXLYz

Définition 7.364.
*Un **espace de Baire** est un espace topologique dans lequel toute intersection dénombrable d'ouverts denses est dense.*

LEMooTOJDooQDtWUC

Lemme 7.365 ([261])**.**
Un espace topologique est de Baire si et seulement si toute union dénombrable de fermés d'intérieur vides est d'intérieur vide.

ThoBBIljNM

Théorème 7.366 (Théorème de Baire[261])**.**
Les espaces suivants sont de Baire :

(1) les espaces topologiques localement compacts,

(2) les espaces métriques complets (donc ceux de Banach en particulier),

(3) tout ouvert d'un espace de Baire.

Démonstration. (1) **Espaces topologiques localement compacts**

(2) **Espaces métriques complets** Soit (E, d) un espace métrique complet. Soient V un ouvert quelconque de E et U_n une suite d'ouverts denses. Le but est de prouver que l'ensemble $\bigcap_{n\in\mathbb{N}} U_n$ intersecte V. Vu que V est ouvert dans un espace métrique, il contient une boule ouverte et donc une boule fermée B_0 de rayon strictement positif. L'ensemble U_1 est dense et intersecte donc un ouvert contenu dans B_0. L'intersection est un ouvert qui contient alors une boule fermée B_1 de rayon strictement positif. Continuant ainsi nous construisons une suite de fermés emboités B_n telle que

$$\bigcap_{n\in\mathbb{N}} U_n \cap V \tag{7.487}$$

contient l'intersection des B_n. Par le théorème 7.299 des fermés emboîtés (que nous utilisons parce que E est métrique et complet), cette intersection est non vide.

(3) **Ouvert d'un espace de Baire**

□

Parmi les applications du théorème de Baire, nous avons

— Le théorème de Banach-Steinhaus 10.152.

— Le théorème de l'application ouverte 10.161.

Chapitre 8

Espaces vectoriels (encore)

8.1 Déterminants

SecGYzHWs

8.2 Applications multilinéaires

DefFRHooKnPCT

Définition 8.1 (Application multilinéaire)**.**
*Soient des espaces vectoriels $V_1, \dots, V_k$ ainsi qu'un espace vectoriel W. Une application $T\colon V^k \to W$ est dite k-**linéaire** ou k-**multilinéaire** si pour tout $(v_1, \dots, v_k) \in V_1 \times \dots \times V_k$ les applications*

$$\begin{aligned} T_i\colon V_i &\to W \\ x &\mapsto T(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_k) \end{aligned} \tag{8.1}$$

sont linéaires.

*En particulier lorsque $k = 2$, nous parlons d'applications **bilinéaires**.*

Nous notons $\mathcal{L}(V_1, \dots, V_k; W)$ ou $\mathcal{L}_k(V_1, \dots, V_k; W)$ l'ensemble des applications k-linéaires de $V_1 \times \dots \times V_k$ vers W. Lorsque tout les V_i sont identiques nous notons $\mathcal{L}_k(V, W)$.

DEFooZFCKooGIoYrE

Définition 8.2.
*Soit E, un $\mathbb{K}$-espace vectoriel. Une forme multilinéaire $f\colon E^k \to \mathbb{K}$ est **antisymétrique** si*

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_k) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_k) \tag{8.2}$$

pour tout couple (i, j).

8.3.
Il serait une erreur[1] de noter $\mathcal{L}(V \times V, \mathbb{R})$ l'ensemble de applications bilinéaires sur V. En effet, $\mathcal{L}(V \times V, \mathbb{R})$ est l'ensemble des applications linéaires $V \times V \to \mathbb{R}$.

Si $S \in \mathcal{L}(V \times V, \mathbb{R})$, il n'y a pas de formules particulières pour $S(u, \lambda v)$. Par contre si $T \in \mathcal{L}_2(V, W)$, alors

$$T(u, \lambda v) = \lambda T(u, v). \tag{8.3}$$

Nous avons par contre

$$T(\lambda u, \lambda v) = \lambda^2 T(u, v) \tag{8.4}$$

alors que

$$S(\lambda u, \lambda v) = S\big(\lambda(u, v)\big) = \lambda S(u, v). \tag{8.5}$$

Certains[262] notent $\otimes^k V^*$ l'espace des formes k-linéaires sur V, voir la proposition 10.183.

Exemple 8.4.
Soit une application linéaire $A\colon \mathbb{R}^n \to \mathbb{R}^m$. L'application de $\mathbb{R}^m \times \mathbb{R}^n$ dans $\mathbb{R}$ associée à A définie par

$$\begin{aligned} T_A\colon \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R} \\ x, y &\mapsto x \cdot Ay \end{aligned} \tag{8.6}$$

1. Mais je crois que c'est fait à plusieurs endroits dans le Frido. Faire attention.

est une forme bilinéaire.

Elle est aussi donnée par la formule

$$T_A(x,y) = x \cdot Ay = \sum_{i,j} A_{ij}x_iy_j. \tag{8.7}$$

△

8.2.1 Continuité

LEMooGQWQooUfqURv

Lemme 8.5 ([263]).
Soit une forme n-multilinéaire $\alpha \in \mathcal{L}_n(E_1,\dots,E_n;F)$. Nous avons la formule

$$\alpha(x_1,\dots,x_n) - \alpha(a_1,\dots,a_n) = \sum_{i=1}^{n} \alpha(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n). \tag{8.8}$$

Démonstration. Nous y allons par récurrence.

(1) $\underline{n=2}$ Vu que α est multilinéaire, nous avons $\alpha(x_1,x_2) = \alpha(x_1-a_1,x_2) + \alpha(a_1,x_2)$. Avec ça, nous avons

$$\begin{aligned}\alpha(x_1,x_2) - \alpha(a_1,a_2) &= \alpha(x_1-a_1,x_2) + \alpha(a_1,x_2) - \alpha(a_1,a_2) &\text{(8.9a)}\\ &= \alpha(x_1-a_1,x_2) + \alpha(a_1,x_2-a_2). &\text{(8.9b)}\end{aligned}$$

(2) **La récurrence** Nous supposons que la formule est valable pour toute application n-multilinéaire, et nous considérons une application $n+1$-multilinéaire $\alpha \in \mathcal{L}_{n+1}(E,F)$. Nous considérons aussi des éléments $x_i, a_i \in E_i$ pour $i = 1,\dots,n+1$. Nous posons

$$\begin{aligned}\alpha' \colon E_1 \times \dots \times E_n &\to F\\ y_1,\dots,y_n &\mapsto \alpha(y_1,\dots,y_n,x_{n+1}).\end{aligned} \tag{8.10}$$

Cela est une application n-multilinéaire. Nous avons ce calcul :

$$\begin{aligned}&\alpha(x_1,\dots,x_{n+1}) - \alpha(a_1,\dots,a_{n+1})\\ &= \alpha'(x_1,\dots,x_n) - \alpha'(a_1,\dots,a_n) + \alpha'(a_1,\dots,a_n) - \alpha(a_1,\dots,a_{n+1}) &\text{(8.11a)}\\ &= \sum_{i=1}^{n} \alpha'(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n) + \alpha'(a_1,\dots,a_n) - \alpha(a_1,\dots,a_{n+1}) &\text{(8.11b)}\\ &= \sum_{i=1}^{n} \alpha'(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n) + \alpha(a_1,\dots,a_n,x_{n+1}) - \alpha(a_1,\dots,a_{n+1}) &\text{(8.11c)}\\ &= \sum_{i=1}^{n} \alpha'(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n) + \alpha(a_1,\dots,a_n,x_{n+1}-a_{n+1}) &\text{(8.11d)}\\ &= \sum_{i=1}^{n+1} \alpha(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n) &\text{(8.11e)}\end{aligned}$$

□

PROPooDQBOooByBvmj

Proposition 8.6 ([264]).
Soient des espaces vectoriels normés $E_1,\dots,E_n$ et F. Une application n-multilinéaire $\alpha\colon E_1 \times \dots E_n \to F$ est continue[2] *si et seulement si il existe $\lambda > 0$ tel que*

$$\|\alpha(x_1,\dots,x_n)\| \leqslant \lambda \|x_1\| \dots \|x_n\|. \tag{8.12}$$

EQooGTVEooZsvzAM

Démonstration. En deux parties.

2. Pour la topologie sur $E_1 \times \dots \times E_n$, c'est celle de la norme produit 7.227.

(1) $\Rightarrow$ Nous supposons que α est continue. En particulier, elle est continue en $(0,\dots,0)\in E_1\times\dots\times E_n$. Nous considérons $r>0$ tel que

$$\alpha\big(B(0,2r)\big)\subset B(0,1). \tag{8.13}$$

Nous posons $\lambda=1/r^n$, et nous allons voir que ça fait le boulot. Soit $(x_1,\dots,x_n)\in E_1\times\dots\times E_n$ avec $x_i\neq 0$ pour tout i. Nous posons $z_i=rx_i/\|x_i\|$, de telle sorte que $\|z_1\|=\dots=\|z_n\|=r$ et donc que $(z_1,\dots,z_n)\in B(0,2r)$ et donc

$$\|\alpha(z_1,\dots,z_n)\|<1. \tag{8.14}$$

Mais $\alpha(z_1,\dots,z_n)=r^n/\prod_i\|x_i\|$. Donc avec $\lambda=1/r^n$ nous trouvons

$$\|\alpha(x_1,\dots,x_n)\|\leqslant\lambda\prod_i\|x_i\|. \tag{8.15}$$

Si $x_i=0$, alors c'est encore plus simple parce que la relation (8.12) se réduit à $0\leqslant 0$.

(2) $\Leftarrow$ Nous supposons que $\lambda>0$ vérifie $\|\alpha(x_1,\dots,x_n)\|\leqslant\lambda\prod_i\|x_i\|$ pour tout $(x_1,\dots,x_n)\in E_1\times\dots\times E_n$. En utilisant le lemme 8.5, nous avons, pour tout x_i et a_i :

$$\|\alpha(x_1,\dots,x_n)-\alpha(a_1,\dots,a_n)\|=\|\sum_{i=1}^n\alpha(a_1,\dots,a_{i-1},x_i-a_i,x_{i+1},\dots,x_n)\| \tag{8.16a}$$

$$\leqslant\sum_{i=1}^n\lambda\|x_i-a_i\|\cdot\prod_{j=1}^{i-1}\|a_j\|\cdot\prod_{j=i+1}^n\|x_j\|. \tag{8.16b}$$

SUBEQooYDGCooJFBaxN

Et maintenant nous prenons $0<\epsilon<1$ et $(x_1,\dots,x_n)\in B\big((a_1,\dots,a_n),\epsilon\big)$, de telle sorte que $\|x_i-a_i\|<\epsilon$ pour tout i et que

$$\|x_i\|<\|a_i\|+\epsilon\leqslant\|a_i\|+1. \tag{8.17}$$

Nous posons $M=\max_{j=1,\dots,n}\|a_j\|:\|x_j\|\leqslant\|a_i\|+1\leqslant M+1$ et nous majorons sauvagement chacun des produits de (8.16b) par $(M+1)^n$:

$$\|\alpha(x_1,\dots,x_n)-\alpha(a_1,\dots,a_n)\|\leqslant\sum_{i=1}^n\lambda\|x_i-a_i\|(M+1)^{2n} \tag{8.18a}$$

$$\leqslant n\lambda\epsilon(M+1)^{2n}. \tag{8.18b}$$

Nous avons donc bien prouvé que $\alpha(x_1,\dots,x_n)\xrightarrow{F}\alpha(a_1,\dots,a_n)$ lorsque $(x_1,\dots,x_n)\xrightarrow{E_1\times\dots\times E_n}(a_1,\dots,a_n)$.

□

8.2.2 Formes multilinéaires alternées

DEFooYWOBooUGJojy

Définition 8.7.
*Soit E, un $\mathbb{K}$-espace vectoriel. Une application multilinéaire **alternée** sur E est une application k-linéaire $f\colon E^k\to\mathbb{K}$ telle que $f(v_1,\dots,v_k)=0$ dès que $v_i=v_j$ pour certains $i\neq j$.*

*Une application multilinéaire alternée est une **forme linéaire**.*

LemHiHNey

Lemme 8.8 ([265])**.**
Une forme k-linéaire alternée est antisymétrique[3]*. Si $\mathbb{K}$ est de caractéristique différente de 2, alors une forme antisymétrique est alternée*[4]*.*

3. Antisymétrique, définition 8.2.
4. Définition 8.7.

Démonstration. Soit f une forme alternée ; quitte à fixer toutes les autres variables, nous pouvons travailler avec une 2-forme et simplement montrer que $f(x,y) = -f(y,x)$. Pour ce faire nous écrivons

$$0 = f(x+y,x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y) = f(x,y) + f(y,x). \tag{8.19}$$

Pour la réciproque, si f est antisymétrique, alors $f(x,x) = -f(x,x)$. Cela montre que $f(x,x) = 0$ lorsque $\mathbb{K}$ est de caractéristique différente de deux. □

ProprbjihK

Proposition 8.9 ([266]).
Soit E, un $\mathbb{K}$-espace vectoriel de dimension n, où la caractéristique de $\mathbb{K}$ n'est pas deux. L'espace des n-formes multilinéaires alternées sur E est de $\mathbb{K}$-dimension 1.

Démonstration. Soient $\{e_i\}$, une base de E, une n-forme linéaire alternée $f\colon E \to \mathbb{K}$ ainsi que des vecteurs $(v_1,\dots,v_n)$ de E. Nous pouvons les écrire dans la base

$$v_j = \sum_{i=1}^{n} \alpha_{ij} e_i \tag{8.20}$$

et alors exprimer f par

$$f(v_1,\dots,v_n) = f\Big(\sum_{i_1=1}^{n} \alpha_{1i_1}e_{i_1},\dots,\sum_{i_n=1}^{n}\alpha_{ni_n}e_{i_n}\Big) \tag{8.21a}$$

$$= \sum_{i_1,\dots,i_n} \alpha_{1i_1}\dots\alpha_{ni_n} f(e_{i_1},\dots,e_{i_n}). \tag{8.21b}$$

Étant donné que f est alternée, les seuls termes de la somme sont ceux dont les i_k sont tous différents, c'est-à-dire ceux où $\{i_1,\dots,i_n\} = \{1,\dots,n\}$. Il y a donc un terme par élément du groupe des permutations S_n et

$$f(v_1,\dots,v_n) = \sum_{\sigma\in S_n} \alpha_{\sigma(1)1}\dots\alpha_{\sigma(n)n} f(e_{\sigma(1)},\dots,e_{\sigma(n)}). \tag{8.22}$$

En utilisant encore une fois le fait que la forme f soit alternée, $f = f(e_1,\dots,e_n)\Pi$ où

$$\Pi(v_1,\dots,v_n) = \sum_{\sigma\in S_n} \epsilon(\sigma)\alpha_{\sigma(1)1}\dots\alpha_{\sigma(n)n}. \tag{8.23}$$

Pour rappel, la donnée des v_i est dans les nombres α_{ij}.

L'espace des n-formes alternées est donc *au plus* de dimension 1. Pour montrer qu'il est exactement de dimension 1, il faut et suffit de prouver que Π est alternée. Par le lemme 8.8, il suffit de prouver que cette forme est antisymétrique [5].

Soient donc $v_1,\dots,v_n$ tels que $v_i = v_j$. En posant $\tau = (1i)$ et $\tau' = (2j)$ et en sommant sur $\sigma\tau\tau'$ au lieu de σ, nous pouvons supposer que $i = 1$ et $j = 2$. Montrons que $\Pi(v,v,v_3,\dots,v_n) = 0$ en tenant compte que $\alpha_{i1} = \alpha_{i2}$:

$$\Pi(v,v,v_3,\dots,v_n) = \sum_{\sigma\in S_n} \epsilon(\sigma)\alpha_{\sigma(1)1}\alpha_{\sigma(2)2}\alpha_{\sigma(3)3}\dots\alpha_{\sigma(n)n} \tag{8.24a}$$

$$= \sum_{\sigma\in S_n} \epsilon(\sigma\tau)\alpha_{\sigma\tau(1)1}\alpha_{\sigma\tau(2)2}\alpha_{\sigma\tau(3)3}\dots\alpha_{\sigma\tau(n)n} \qquad \text{où } \tau = (12) \tag{8.24b}$$

$$= -\sum_{\sigma\in S_n} \epsilon(\sigma)\alpha_{\sigma(1)1}\alpha_{\sigma(2)2}\alpha_{\sigma(3)3}\dots\alpha_{\sigma(n)n} \tag{8.24c}$$

$$= -\Pi(v,v,v_3,\dots,v_n). \tag{8.24d}$$

□

5. C'est ici que joue l'hypothèse sur la caractéristique de $\mathbb{K}$.

8.2.3 Déterminant d'une famille de vecteurs

Nous considérons un corps $\mathbb{K}$ et l'espace vectoriel E de dimension n sur $\mathbb{K}$.

DEFooODDFooSNahPb

Définition 8.10 (Déterminant d'une famille de vecteurs[11])**.**
*Le **déterminant** de la famille de vecteurs $(v_1, \dots, v_n)$ dans la base $B = \{b_1, \dots, b_n\}$ est l'élément de $\mathbb{K}$*

$$\det_{(b_1,\dots,b_n)}(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} b^*_{\sigma(i)}(v_i) \tag{8.25}$$

EQooOJEXooXUpwfZ

où

- *la somme porte sur le groupe symétrique,*
- *le nombre $\epsilon(\sigma)$ est la signature*[6] *de la permutation σ,*
- *les éléments $\{b_i^*\}$ sont la base duale*[7] *de $\{b_i\}$.*

8.11.
La base $\{e_i\}$ est la base canonique de $\mathbb{K}^n$, et l'élément e_k^* est la forme linéaire définie par

$$\begin{aligned} e_k^* \colon \mathbb{K}^n &\to \mathbb{K} \\ \sum_i x_i e_i &\mapsto x_k. \end{aligned} \tag{8.26}$$

Il n'est pas sous-entendu que $\mathbb{K}^n$ ait un produit scalaire. Il n'est donc pas autorisé de dire que $\{e_i\}$ est une base orthonormée et que $e_k^*(x) = \langle e_k, x \rangle$. Ce genre d'égalités sont vraies dans le cas $\mathbb{K} = \mathbb{R}$, mais n'ont pas de sens en général.

Le lemme 10.5 va un peu parler du cas où $\mathbb{K}^n$ est muni d'une base orthonormée.

LemJMWCooELZuho

Lemme 8.12 ([11])**.**
Les propriétés du déterminant. Soit B une base de E.

ITEMooAHOHooDZgtSB

(1) L'application $\det_B \colon E^n \to \mathbb{K}$ est n-linéaire.

ITEMooTXXBooBmDtzd

(2) L'application $\det_B \colon E^n \to \mathbb{K}$ est n-linéaire est antisymétrique et alternée[8]*.*

ITEMooNFJTooTqGoPr

(3) Pour toute base, $\det_B(B) = 1$.

ITEMooALRQooDvBzDQ

(4) Le déterminant ne change pas si on remplace un vecteur par une combinaison linéaire des autres :

$$\det_B(v_1, \dots, v_n) = \det_B\left(v_1 + \sum_{s=2}^{n} a_s v_s, v_2, \dots, v_n\right). \tag{8.27}$$

ITEMooQTTRooMbzqyW

(5) Si on permute les vecteurs,

$$\det_B(v_1, \dots, v_n) = \epsilon(\sigma) \det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)}). \tag{8.28}$$

ITEMooIPIDooTrerVF

(6) Si B' est une autre base :

$$\det_B = \det_B(B') \det_{B'} \tag{8.29}$$

EqAWICooBLTTQY
ITEMooXKTAooXynFTE

(7) Nous avons aussi la formule $\det_B(B') \det_{B'}(B) = 1$.

ItemDWFLooDUePAf

(8) Les vecteurs $\{v_1, \dots, v_n\}$ forment une base si et seulement si $\det_B(v_1, \dots, v_n) \neq 0$.

Démonstration. Point par point.

6. Définition 1.297.
7. Définition 4.127.
8. Alternée, définition 8.7. En caractéristique 2, alternée n'est pas équivalent à symétrique.

(1) **(1)** En posant $v_1 = x_1 + \lambda x_2$ nous avons

$$\det_B(x_1 + \lambda x_2, v_2, \dots, v_n) = \sum_\sigma \epsilon(\sigma) \prod_{i=1}^n e^*_{\sigma(i)}(v_i) \tag{8.30a}$$

$$= \sum_\sigma \epsilon(\sigma) \Big(e^*_{\sigma(1)}(x_1 + \lambda x_2) \Big) \prod_{i=2}^n e^*_{\sigma(i)}(v_i). \tag{8.30b}$$

À partir de là, la linéarité de $e^*_{\sigma(1)}$ montre que $\det_B$ est linéaire en son premier argument. Pour les autres arguments, le même calcul tient.

(2) **(2)**
Nous prouvons à présent que det est alternée. Si votre corps est de caractéristique différente de deux, vous pouvez lire la proposition 8.13.

Supposons $v_k = v_l$, et considérons la permutation $\beta = (k, l)$. Nous savons par la proposition 5.47 que $S_n = A_n \cup A_n\beta$. Cela nous permet de décomposer la somme sur S_n en deux parties :

$$\sum_{\sigma \in S_n} (-1)^\sigma \prod_i \epsilon^*_{\sigma(i)}(v_i) = \sum_{\sigma \in A_n} (-1)^\sigma \prod_i \epsilon^*_{\sigma(i)}(v_i) + \sum_{\sigma \in A_n} (-1)^{\sigma\beta} \prod_i \epsilon^*_{(\sigma\beta)(i)}(v_i). \tag{8.31}$$

EQoqWFHQooTrTTWl

D'abord $(-1)^\sigma = 1$ et $(-1)^{\sigma\beta} = -1$. Ensuite, pour un $\sigma \in A_n$ donné, nous avons

$$\prod_i \epsilon^*_{(\sigma\beta)(i)}(v_i) = \epsilon^*_{(\sigma\beta)(k)}(v_k)\epsilon^*_{(\sigma\beta)(l)}(v_l) \prod_{\substack{i \neq k \\ i \neq l}} \epsilon^*_{(\sigma\beta)(i)}(v_i) \tag{8.32a}$$

$$= \epsilon^*_{\sigma(l)}(v_k)\epsilon^*_{\sigma(k)}(v_l) \prod_{\substack{i \neq k \\ i \neq l}} \epsilon^*_{\sigma(i)}(v_i) \tag{8.32b}$$

$$= \epsilon^*_{\sigma(l)}(v_l)\epsilon^*_{\sigma(k)}(v_k) \prod_{\substack{i \neq k \\ i \neq l}} \epsilon^*_{\sigma(i)}(v_i) \tag{8.32c}$$

$$= \prod_i \epsilon^*_{\sigma(i)}(v_i). \tag{8.32d}$$

Donc les deux termes de la somme (8.31) ne diffèrent que par un signe. Elle est donc nulle, et la forme déterminant est alternée.

La fonction det est antisymétrique parce que alternée, voir le lemme 8.8.

(3) **(3)** Nous avons

$$\det_B(B) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^n \underbrace{e^*_{\sigma(i)}(e_i)}_{=\delta_{\sigma(i),i}}. \tag{8.33}$$

Si σ n'est pas l'identité, le produit contient forcément un facteur nul. Il ne reste de la somme que $\sigma = \mathrm{Id}$ et le résultat est 1.

(4) **(4)** Vu que $\det_B$ est linéaire en tous ses arguments,

$$\det_B\Big(v_1 + \sum_{s=2}^n a_s v_s, v_2, \dots, v_n\Big) = \det_B(v_1, \dots, v_n) + \sum_{s=2}^n a_s \det_B(v_s, v_2, \dots, v_n). \tag{8.34}$$

Chacun des termes de la somme est nul parce qu'il y a répétition de v_s parmi les arguments alors que la forme est alternée.

(5) **(5)** Nous devons calculer $\det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)})$, et pour y voir plus clair nous posons $w_i =$

$v_{\sigma(i)}$. Alors :

$$\det_B(v_{\sigma(1)},\dots,v_{\sigma(n)}) = \sum_{\sigma'}\epsilon(\sigma')\prod_{i=1}^{n}e^*_{\sigma'(i)}(w_i) \tag{8.35a}$$
$$= \sum_{\sigma'}\epsilon(\sigma')\prod_{i=1}^{n}e^*_{\sigma'(i)}(v_{\sigma(i)}) \tag{8.35b}$$
$$= \sum_{\sigma'}\epsilon(\sigma')\prod_{i=1}^{n}e^*_{\sigma^{-1}\sigma'(i)}(v_i) \tag{8.35c}$$
$$= \sum_{\sigma'}\epsilon(\sigma\sigma')\prod_{i=1}^{n}e^*_{\sigma'(i)}(v_i) \tag{8.35d}$$
$$= \epsilon(\sigma)\det_B(v_1,\dots,v_n). \tag{8.35e}$$

Justifications : nous avons d'abord modifié l'ordre des éléments du produit et ensuite l'ordre des éléments de la somme. Nous avons ensuite utilisé le fait que $\epsilon\colon S_n\to\{0,1\}$ était un morphisme de groupe (proposition 1.300).

(6) **(6)** Étant donné que l'espace des formes multilinéaires alternées est de dimension 1, il existe un $\lambda\in\mathbb{K}$ tel que $\det_B=\lambda\det_{B'}$. Appliquons cela à B' :

$$\det_B(B') = \lambda\det_{B'}(B'), \tag{8.36}$$

donc $\lambda=\det_B(B')$.

(7) **(7)** Il suffit d'appliquer l'égalité précédente à B en nous souvenant que $\det_B(B)=1$.

(8) **(8)** Si $B'=\{v_1,\dots,v_n\}$ est une base alors $\det_B(B')\neq 0$, sinon il n'est pas possible d'avoir $\det_B(B')\det_{B'}(B)=1$.
À l'inverse, si B' n'est pas une base, c'est que $\{v_1,\dots,v_n\}$ est liée par la proposition 4.19. Il y a donc moyen de remplacer un des vecteurs par une combinaison linéaire des autres. Le déterminant s'annule alors.

□
PROPooXNLDooGGkHpd

Proposition 8.13.
Si la caractéristique du corps de base n'est pas deux, le déterminant est antisymétrique et alterné.

Démonstration. Si la caractéristique du corps de base n'est pas deux, une forme antisymétrique est alternée (lemme 8.8).

Pour prouver que le déterminant est antisymétrique, remarquez que permuter v_k et v_l revient à calculer le nombre $\det_B(v_{\sigma_{kl}(1)},\dots,v_{\sigma_{kl}(n)})$ au lieu de $\det_B(v_1,\dots,v_n)$. Cela revient à changer la somme $\sum_\sigma$ en $\sum_{\sigma\circ\sigma_{kl}}$. Cela multiplie $\epsilon(\sigma)$ par -1 parce qu'on ajoute une permutation.

Donc le déterminant est antisymétrique. Nous en déduisons qu'il est alterné parce que, en permutant trivialement v_1 et v_1, nous obtenons $\det_B(v_1,v_1)=-\det_B(v_1,v_1)$. Si le corps est de caractéristique différente de deux, cela implique que $\det_B(v_1,v_1)=0$. □

D'après la proposition 8.9, il existe une unique forme n-linéaire alternée égale à 1 sur B, et c'est $\det_B\colon E^n\to\mathbb{K}$.

8.2.4 Déterminant d'un endomorphisme

L'interprétation géométrique du déterminant en termes d'aires et de volumes est donnée après le théorème 14.271. Pour un déterminant entre deux espaces différents, voir la proposition 10.75.

8.14.
Voici une définition du déterminant d'une application linéaire $f\colon E\to E$. Comme expliqué dans [267], il est peu probable qu'il existe une définition satisfaisante de déterminant pour une application linéaire $f\colon E\to F$, même si E et F ont une même dimension finie.

Cela soit dit en passant, ça rend plus compliqué que prévu la preuve que $\mathrm{GL}(E,F)$ est ouvert dans $\mathcal{L}(E,F)$. Il ne suffira pas de dire que le déterminant est une application continue. Voir la proposition 10.75.

LEMooQTRVooAKzucd

Lemme-Définition 8.15.
Si $f\colon E\to E$ est un endomorphisme, et si les parties B et B' sont deux bases, alors[9]

$$\det_B\big(f(B)\big)=\det_{B'}\big(f(B')\big). \tag{8.37}$$

Ce nombre, indépendant de la base choisie est nommé le ***déterminant*** *de f et est noté $\det(f)$.*

Démonstration. L'application

$$\begin{aligned}\varphi\colon E^n&\to\mathbb{K}\\ v_1,\dots,v_n&\mapsto\det_B\big(f(v_1),\dots,f(v_n)\big)\end{aligned} \tag{8.38}$$

est n-linéaire et alternée ; il existe donc $\lambda\in\mathbb{K}$ tel que $\varphi=\lambda\det_B$. En appliquant cela à B :

$$\det_B\big(f(B)\big)=\lambda\det_B(B)=\lambda. \tag{8.39}$$

Nous avons donc déjà prouvé que $\lambda=\det_B\big(f(B)\big)$, c'est-à-dire

$$\det_B\big(f(v)\big)=\det_B\big(f(B)\big)\det_B(v). \tag{8.40}$$

Nous allons maintenant introduire B' là où il y a du v en utilisant les formules (8.29) :

$$\det_B\big(f(v)\big)=\det_B(B')\det_{B'}\big(f(v)\big) \tag{8.41a}$$
$$\det_B(v)=\det_B(B')\det_{B'}(v). \tag{8.41b}$$

Nous obtenons

$$\det_{B'}\big(f(v)\big)=\det_B\big(f(B)\big)\det_{B'}(v). \tag{8.42}$$

Et on applique cela à $v=B'$:

$$\det_{B'}\big(f(B')\big)=\det_B\big(f(B)\big)\underbrace{\det_{B'}(B')}_{=1}. \tag{8.43}$$

□

PropYQNMooZjlYlA

Proposition 8.16.
Principales propriétés géométriques du déterminant d'un endomorphisme.

ItemUPLNooYZMRJy

(1) Si f et g sont des endomorphismes, alors $\det(f\circ g)=\det(f)\det(g)$.

ITEMooNZNLooODdXeH

(2) L'endomorphisme f est un automorphisme[10] *si et seulement si $\det(f)\neq 0$.*

ITEMooZMVXooLGjvCy

(3) Si $\det(f)\neq 0$ alors $\det(f^{-1})=\det(f)^{-1}$.

ItemooPJVYooYSwqaE

(4) L'application $\det\colon\mathrm{GL}(E)\to\mathbb{K}\backslash\{0\}$ est un morphisme de groupe.

Démonstration. Point par point.

(1) Nous considérons l'application

$$\begin{aligned}\varphi\colon E^n&\to\mathbb{K}\\ v&\mapsto\det_B\big(f(v)\big).\end{aligned} \tag{8.44}$$

Comme d'habitude nous avons $\varphi(v)=\lambda\det_B(v)$. En appliquant à B et en nous souvenant que $\det_B(B)=1$ nous avons $\det_B\big(f(B)\big)=\lambda$. Autrement dit :

$$\lambda=\det(f). \tag{8.45}$$

9. Définition de $\det_B(B')$, 8.10.
10. Endomorphisme inversible, définition 4.36.

Calculons à présent $\varphi\big(g(B)\big)$: d'une part,

$$\varphi\big(g(B)\big) = \det{}_B\big((f\circ g)(B)\big) \tag{8.46}$$

et d'autre part,

$$\varphi\big(g(B)\big) = \lambda \det{}_B\big(g(B)\big) = \lambda\det(g) \tag{8.47}$$

En égalisant et en reprenant la la valeur déjà trouvée de λ,

$$\det\big((f\circ g)(B)\big) = \det(f)\det(g), \tag{8.48}$$

ce qu'il fallait.

(2) Supposons que f soit un automorphisme. Alors si B est une base, $f(B)$ est une base. Par conséquent $\det(f) = \det_B\big(f(B)\big) \neq 0$ parce que $f(B)$ est une base (lemme 8.12(8)).
Réciproquement, supposons que $\det(f)\neq 0$. Alors si B est une base quelconque nous avons $\det_B\big(f(B)\big)\neq 0$, ce qui est uniquement possible lorsque $f(B)$ est une base. L'application f transforme donc toute base en une base et est alors un automorphisme d'espace vectoriel.

(3) Vu que le déterminant de l'identité est 1 et que f est inversible, $1 = \det(f\circ f^{-1}) = \det(f)\det(f^{-1})$.

□

PROPooFKDXooKMSolt

Proposition 8.17.
Soient deux espaces vectoriels E et F de dimension finies n et m sur le corps $\mathbb{K}$ munis de bases $\{e_i\}$ et $\{f_\alpha\}$. À une matrice $A\in\mathbb{M}(m\times n,\mathbb{K})$ nous associons l'application linéaire[11]

$$f_A(x) = \sum_{i\alpha} A_{\alpha i}x_i f_\alpha. \tag{8.49}$$

Alors, en ce qui concerne les déterminants[12]*, nous avons*

(1) $\det(f_A) = \det(A)$

(2) $\det(f_{AB}) = \det(f_A)\det(f_B)$

Démonstration. Nous devons étudier la formule

$$\det(f_A) = \sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n} e^*_{\sigma(i)}\big(f_A(e_i)\big). \tag{8.50}$$

En premier lieu nous avons

$$f_A(e_i) = \sum_{jk} A_{jk}(e_i)_k e_j = \sum_j A_{ji}e_j. \tag{8.51}$$

Nous avons alors

$$e^*_{\sigma(i)}\big(f_A(e_i)\big) = \sum_j A_{ji}\underbrace{e^*_{\sigma(i)}(e_j)}_{\delta_{j\sigma(i)}} = A_{\sigma(i)i}. \tag{8.52}$$

Au final,

$$\det(f_A) = \sum_\sigma \epsilon(\sigma)\prod_{i=1}^{n} A_{\sigma(i)i} = \det(A^t) = \det(A) \tag{8.53}$$

où la dernière égalité est autorisée par le lemme 4.81.

Cela prouve la formule $\det(f_A) = \det(A)$.

En ce qui concerne la seconde formule, il s'agit de se souvenir de la proposition 4.73 qui donne $f_{AB} = f_A\circ f_B$, et ensuite de la proposition 8.16(1) qui donne $\det(f_A\circ f_B) = \det(f_A)\det(f_B)$. □

11. Dont nous avons déjà beaucoup parlé entre autres dans la proposition 4.73.
12. Définition 8.15 pour les applications linéaires et 4.79 pour les matrices.

8.2.5 Déterminant de Vandermonde

PropnuUvtj

Proposition 8.18 ([114]).
*Le **déterminant de Vandermonde** est le polynôme en n variables donné par*

$$V(T_1,\dots,T_n)=\det\begin{pmatrix}1&1&\dots&1\\T_1&T_2&\dots&T_n\\\vdots&\ddots&\ddots&\vdots\\T_1^{n-1}&T_2^{n-1}&\dots&T_n^{n-1}\end{pmatrix}=\prod_{1\leqslant i<j\leqslant n}(T_j-T_i).\tag{8.54}$$

Notez que l'inégalité du milieu est stricte (sinon d'ailleurs l'expression serait nulle).

Démonstration. Nous considérons le polynôme

$$f(X)=V(T_1,\dots,T_{n-1},X)\in\big(\mathbb{K}[T_1,\dots,T_{n-1}]\big)[X].\tag{8.55}$$

C'est un polynôme de degré au plus $n-1$ en X et il s'annule aux points $T_1,\dots,T_{n-1}$. Par conséquent [13] nous pouvons factoriser les $X-T_i$, c'est-à-dire qu'il existe $\alpha\in\mathbb{K}[T_1,\dots,T_{n-1}]$ tel que

$$f=\alpha\prod_{i=1}^{n-1}(X-T_i).\tag{8.56}$$ EqeVxRwQ

Nous trouvons α en écrivant $f(0)$. D'une part la formule (8.56) nous donne

$$f(0)=\alpha(-1)^{n-1}T_1\dots T_{n-1}.\tag{8.57}$$ EqblwWMj

D'autre part la définition donne

$$\begin{aligned}f(0)&=\det\begin{pmatrix}1&\cdots&1&1\\T_1&&T_{n-1}&0\\\vdots&&\vdots&\vdots\\T_1^{n-1}&\cdots&T_{n-1}^{n-1}&0\end{pmatrix}&(8.58a)\\&=(-1)^{n-1}\det\begin{pmatrix}T_1&\dots&T_{n-1}\\\vdots&\ddots&\vdots\\T_1^{n-1}&\dots&T_{n-1}^{n-1}\end{pmatrix}&(8.58b)\\&=(-1)^{n-1}T_1\dots T_{n-1}\det\begin{pmatrix}1&\cdots&1\\\vdots&\ddots&\vdots\\T_1^{n-1}&\cdots&T_{n-1}^{n-1}\end{pmatrix}&(8.58c)\\&=(-1)^{n-1}T_1\dots T_{n-1}V(T_1,\dots,T_{n-1})&(8.58d)\end{aligned}$$

En égalisant avec (8.57), nous trouvons $\alpha=V(T_1,\dots,T_{n-1})$, et donc

$$f=V(T_1,\dots,T_{n-1})\prod_{j\leqslant n-1}(X-T_j)\tag{8.59}$$

Enfin, une récurrence montre que

$$\begin{aligned}V(T_1,\dots,T_n)&=f(T_n)&(8.60a)\\&=V(T_1,\dots,T_{n-1})\prod_{j\leqslant n-1}(T_n-T_j)&(8.60b)\\&=\prod_{k\leqslant n}\prod_{j\leqslant k-1}(T_k-T_j)&(8.60c)\\&=\prod_{1\leqslant j<k\leqslant n}(T_i-T_j).&(8.60d)\end{aligned}$$

□

13. Proposition 3.123.

Exemple 8.19.
Le déterminant de Vandermonde (proposition 8.18) est alterné, semi-symétrique et non symétrique. Le fait qu'il soit alterné est le fait qu'il soit un déterminant. Étant donné qu'il est alterné, il est semi-symétrique parce que sur A_n, nous avons $\epsilon = 1$. Étant donné qu'il est alterné, il change de signe sous l'action des éléments impairs de S_n et n'est donc pas symétrique. △

PropUDqXax

Proposition 8.20.
Un polynôme semi-symétrique $f \in \mathbb{K}[T_1, \dots, T_n]$ se décompose de façon unique en

$$f = P + VQ \tag{8.61}$$

où P et Q sont deux polynômes symétriques.

Démonstration. Nous commençons par prouver l'unicité en montrant que si $f = P + VQ$ avec P et Q symétrique, alors P et Q sont donnés par des formules explicites en termes de f.

Si σ_1 et σ_2 sont deux permutations impaires de $\{1, \dots, n\}$, alors $\sigma_1 \cdot f = \sigma_2 \cdot f$ parce que l'élément $\sigma_2^{-1}\sigma_1$ est pair (proposition 1.300), de telle sorte que $\sigma_2^{-1}\sigma_1 \cdot f = f$. Nous posons donc $g = \tau \cdot f$ où τ est une permutation impaire quelconque – par exemple une transposition.

Vu que V est alternée et que τ est une transposition nous avons

$$g = \tau \cdot f = P - VQ. \tag{8.62}$$

Donc $f + g = 2P$ et $f - g = 2VQ$. Cela donne P et Q en termes de f et g, et donc l'unicité.

Attention : cela ne donne pas un moyen de prouver l'existence parce que rien ne prouve pour l'instant que $f - g$ peut effectivement être écrit sous la forme VQ, c'est-à-dire que $f - g$ soit divisible par V. C'est cela que nous allons nous atteler à démontrer maintenant.

Nous commençons par prouver que $f + g$ est symétrique et $f - g$ alterné. Si σ est une transposition,

$$\sigma \cdot (f + g) = \sigma \cdot f + \sigma\tau \cdot f = g + f \tag{8.63}$$

parce que $\sigma\tau$ est pair. De la même façon,

$$\sigma \cdot (f - g) = g - f = \epsilon(\sigma)(f - g). \tag{8.64}$$

Dans les deux cas nous concluons en utilisant le fait que toute permutation est un produit de transpositions (proposition 1.295) et que ϵ est un homomorphisme.

Soient maintenant deux entiers $h < k$ dans $\{1, \dots, n\}$ et l'anneau

$$\big(\mathbb{K}[T_1, \dots, \hat{T}_k, \dots, T_n]\big)[T_k]. \tag{8.65}$$

Cet anneau contient le polynôme $T_k - T_h$ où T_k est la variable et T_h est un coefficient. Nous faisons la division euclidienne de $f - g$ par $T_k - T_h$ parce que nous avons dans l'idée de faire arriver le déterminant de Vandermonde et donc le produit de toutes les différences $T_k - T_h$:

$$f - g = (T_k - T_h)q + r \tag{8.66}$$

EqSHdgrG

où $\deg_{T_k} r < 1$, c'est-à-dire que r ne dépends pas de T_k. Nous revoyons maintenant l'égalité (8.66) dans $\mathbb{K}[T_1, \dots, T_n]$ et nous y appliquons la transposition τ_{kh}. Nous savons que $\tau_{kh}(f-g) = -(f-g)$ et $\tau_{kh}(T_k - T_h) = -(T_k - T_h)$, et donc

$$-(f - g) = -(T_k - T_h)\tau_{kh} \cdot q + \tau_{kh} \cdot r \tag{8.67}$$

EqVQhjKB

où $\tau_{kh} \cdot r$ ne dépend pas de T_h. Nous appliquons à (8.67) l'application

$$\begin{aligned} t\alpha\colon \mathbb{K}[T_1, \dots, T_n] &\to \mathbb{K}[T_1, \dots, \hat{T}_k, \dots, T_n] \\ \alpha(PT_1, \dots, \hat{T}_k, \dots, T_n) &= P(T_1, \dots, T_h, \dots, T_n). \end{aligned} \tag{8.68}$$

Cette application vérifie $\alpha\big(\tau_{kh}\cdot r\big)=\alpha(r)$ et nous avons

$$-\alpha(f-g)=\alpha(r). \tag{8.69}$$

Puis en appliquant α à la relation $f-g=(T_k-T_h)q+r$, nous trouvons

$$\alpha(f-g)=\alpha(r), \tag{8.70}$$

et par conséquent $\alpha(r)=0$. Ici nous utilisons l'hypothèse de caractéristique différente de deux. Dire que $\alpha(r)=0$, c'est dire que r est divisible par T_k-T_h, mais r étant de degré zéro en T_k, nous avons $r=0$. Par conséquent T_k-T_h divise $f-g$ pour tout $h<k$, et nous pouvons définir un polynôme Q par

$$f-g=2Q\prod_{h<k}\prod_{k\leqslant n}(T_k-T_h)=2Q(T_1,\ldots,T_n)V(T_1,\ldots,T_n), \tag{8.71}$$

où nous avons utilisé la formule du déterminant de Vandermonde de la proposition 8.18.

Étant donné que $f+g$ est un polynôme symétrique, nous allons aussi poser $f+g=2P$ avec P symétrique.

Montrons à présent que Q est un polynôme symétrique. Soit $\sigma\in S_n$; vu que nous savons déjà que $f-g$ est alternée, nous avons

$$\sigma\cdot(f-g)=\epsilon(\sigma)(f-g)=\epsilon(\sigma)2QV, \tag{8.72}$$

Mais en appliquant σ à l'équation (8.71),

$$\begin{aligned}\sigma\cdot(f-g)&=2(\sigma\cdot V)(T_1,\ldots,,T_n)(\sigma\cdot Q)(T_1,\ldots,T_n) &\text{(8.73a)}\\ &=2\epsilon(\sigma)V(T_1,\ldots,T_n)(\sigma\cdot Q)(T_1,\ldots,T_n). &\text{(8.73b)}\end{aligned}$$

Nous égalisons cela avec (8.72) et nous souvenant que l'anneau $\mathbb{K}[T_1,\ldots,T_n]$ est intègre par le théorème 3.104. Ensuite nous simplifions par $2\epsilon(\sigma)V$ pour obtenir

$$Q=\sigma\cdot Q, \tag{8.74}$$

c'est-à-dire que Q est symétrique.

Au final nous avons $f+q=2P$ et $f-g=2VQ$ avec P et Q symétriques. En faisant la somme,

$$f=P+VQ. \tag{8.75}$$

□

8.2.6 Déterminant de Gram

Si $x_1,\ldots,x_r$ sont des vecteurs d'un espace vectoriel, alors le **déterminant de Gram** est le déterminant

$$G(x_1,\ldots,x_r)=\det\big(\langle x_i,x_j\rangle\big). \tag{8.76}$$

Notons que la matrice est une matrice symétrique.

PropMsZhIK

Proposition 8.21.
Si F est un sous-espace vectoriel de base $\{x_1,\ldots,x_n\}$ et si x est un vecteur, alors le déterminant de Gram est un moyen de calculer la distance entre x et F par

$$d(x,F)^2=\frac{G(x,x_1,\ldots,x_n)}{G(x_1,\ldots,x_n)}. \tag{8.77}$$

8.2.7 Déterminant de Cauchy

Soient des nombres a_i et b_i $(i = 1, \dots, n)$ tels que $a_i + b_j \neq 0$ pour tout couple (i, j). Le **déterminant de Cauchy** est

$$D_n = \det\left(\frac{1}{a_i + b_j}\right). \tag{8.78}$$

ProptoDYKA

Proposition 8.22 ([268]).
Le déterminant de Cauchy est donné par la formule

$$D_n = \frac{\prod_{i<j}(a_j - a_i)\prod_{i<j}(b_j - b_i)}{\prod_{ij}(a_i + b_j)}. \tag{8.79}$$

8.2.8 Matrice de Sylvester

DEFooHGZDsecSQBaP

Définition 8.23 (Matrice de Sylvester, résultant[269]).
Soient P et Q deux polynômes non nuls, de degrés respectifs n et m :

$$P(x) = p_0 + p_1 x + \cdots + p_n x^n \tag{8.80a}$$
$$Q(x) = q_0 + q_1 x + \cdots + q_m x^m. \tag{8.80b}$$

*La **matrice de Sylvester** associée à P et Q est la matrice carrée $m + n \times m + n$ définie ainsi :*

(1) la première ligne est formée des coefficients de P, suivis de 0 :

$$\begin{pmatrix} p_n & p_{n-1} & \cdots & p_1 & p_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{8.81}$$

(2) la seconde ligne s'obtient à partir de la première par permutation circulaire vers la droite ;

(3) les $(m-2)$ lignes suivantes s'obtiennent en répétant la même opération ;

(4) la ligne $(m+1)$ est formée des coefficients de Q, suivis de 0 :

$$\begin{pmatrix} q_m & q_{m-1} & \cdots & q_1 & q_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{8.82}$$

(5) les $(n-1)$ lignes suivantes sont formées par des permutations circulaires.

*Le déterminant de la matrice de Sylvester associée à P et Q est appelé le **résultant** de P et Q et noté* $\operatorname{res}(P, Q)$.

Ainsi dans le cas $n = 4$ et $m = 3$, la matrice obtenue est

$$S_{p,q} = \begin{pmatrix} p_4 & p_3 & p_2 & p_1 & p_0 & 0 & 0 \\ 0 & p_4 & p_3 & p_2 & p_1 & p_0 & 0 \\ 0 & 0 & p_4 & p_3 & p_2 & p_1 & p_0 \\ q_3 & q_2 & q_1 & q_0 & 0 & 0 & 0 \\ 0 & q_3 & q_2 & q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_3 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & 0 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}. \tag{8.83}$$

EqPEgtle

Attention : si P est de degré n et Q de degré m, il y a m lignes pour P et n pour Q dans le déterminant du résultant (et non le contraire).

LemBFrhgnA

Lemme 8.24 ([270]).
Si P et Q sont deux polynômes de degrés n et m à coefficients dans l'anneau $\mathbb{A}$, alors pour tout $\lambda \in \mathbb{A}$,

$$\operatorname{res}(\lambda P, Q) = \lambda^m \operatorname{res}(P, Q) \tag{8.84a}$$
$$\operatorname{res}(P, \lambda Q) = \lambda^n \operatorname{res}(P, Q). \tag{8.84b}$$

Démonstration. Cela est simplement un comptage du nombre de lignes. Il y a m lignes contenant les coefficients de P ; donc prendre λP revient à multiplier m lignes dans un déterminant et donc le multiplier par λ^m. □

L'équation de Bézout (6.114) peut être traitée avec une matrice de Sylvester. Soient P et Q, deux polynômes donnés et à résoudre l'équation

$$xP + yQ = 0 \tag{8.85}$$

par rapport aux polynômes inconnus x et y dont les degrés sont $\deg(x) < \deg(Q)$ et $\deg(y) < \deg(P)$. Si nous notons $\tilde{x}$ et $\tilde{y}$ la liste des coefficients de x et y (dans l'ordre décroissant de degré), nous pouvons récrire l'équation (8.85) sous la forme

$$S_{PQ}^t \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = 0. \tag{8.86}$$

Pour s'en convaincre, écrivons pour les polynômes de l'exemple (8.83) :

$$\begin{pmatrix} p_4 & 0 & 0 & q_3 & 0 & 0 & 0 \\ p_3 & p_4 & 0 & q_2 & q_3 & 0 & 0 \\ p_2 & p_3 & p_4 & q_1 & q_2 & q_3 & 0 \\ p_1 & p_2 & p_3 & q_0 & q_1 & q_2 & q_3 \\ p_0 & p_1 & p_2 & 0 & q_0 & q_1 & q_2 \\ 0 & p_0 & p_1 & 0 & 0 & q_0 & q_1 \\ 0 & 0 & p_0 & 0 & 0 & 0 & q_0 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \\ x_0 \\ y_3 \\ y_2 \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_2p_4 + y_3q_3 \\ p_3x_2 + p_4x_1 + q_2y_3 + q_3y_2 \\ \\ \\ \vdots \\ \end{pmatrix} \tag{8.87}$$

Nous voyons que sur la ligne numéro k (en partant du bas et en numérotant de à partir de zéro) nous avons les produits p_ix_j et q_iy_j avec $i + j = k$. La colonne de droite représente donc bien les coefficients du polynôme $xP + yQ$.

PropAPxzcUl

Proposition 8.25.
Le résultant de deux polynômes est non nul si et seulement si les deux polynômes sont premiers entre eux.

Un polynôme P a une racine double en a si et seulement si P et P' ont a comme racine commune, ce qui revient à dire que P et P' ne sont pas premiers entre eux.

Une application importante de ces résultats sera le théorème de Rothstein-Trager 20.97 sur l'intégration de fractions rationnelles.

Exemple 8.26.
Si nous prenons $P = aX^2 + bX + c$ et $P' = 2aX + b$ alors la taille de la matrice de Sylvester sera $2 + 1 = 3$ et

$$S_{P,P'} = \begin{pmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{pmatrix}. \tag{8.88}$$

Le résultant est alors

$$\operatorname{res}(P, P') = -a(b^2 - 4ac). \tag{8.89}$$

Donc un polynôme du second degré a une racine double si et seulement si $b^2 - 4ac = 0$. Cela est un résultat connu depuis longtemps mais qui fait toujours plaisir à revoir. △

La matrice de Sylvester permet aussi de récrire l'équation de Bézout pour les polynômes ; voir le théorème 6.47 et la discussion qui s'ensuit.

Une proposition importante du résultant est qu'il peut s'exprimer à l'aide des racines des polynômes.

PropNDBOGNx

Proposition 8.27.
Si

$$P(X) = a_p \prod_{i=1}^{p} (X - \alpha_i) \tag{8.90a}$$

$$Q(X) = b_q \prod_{j=1}^{q} (X - \beta_i) \tag{8.90b}$$

alors nous avons les expressions suivantes pour le résultant :

$$\operatorname{res}(P,Q) = a_p^q b_q^p \prod_{i=1}^{p}\prod_{j=1}^{q}(\beta_j - \alpha_i) = b_q^p \prod_{j=1}^{q} P(\beta_j) = (-1)^{pq} a_p^q \prod_{i=1}^{p} Q(\alpha_i). \tag{8.91}$$

EqCFUumix

Démonstration. Si P et Q ne sont pas premiers entre eux, d'une part la proposition 8.25 nous dit que $\operatorname{res}(P,Q) = 0$ et d'autre part, P et Q ont un facteur irréductible en commun, ce qui signifie que nous devons avoir un des $X - \alpha_i$ égal à un des $X - \beta_j$. Autrement dit, nous avons $\alpha_i = \beta_j$ pour un couple (i,j). Par conséquent tous les membres de l'équation (8.91) sont nuls.

Nous supposons donc que P et Q sont premiers entre eux. Nous commençons par supposer que les polynômes P et Q sont unitaires, c'est-à-dire que $a_p = b_q = 1$. Nous considérons alors l'anneau

$$\mathbb{A} = \mathbb{Z}[\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q]. \tag{8.92}$$

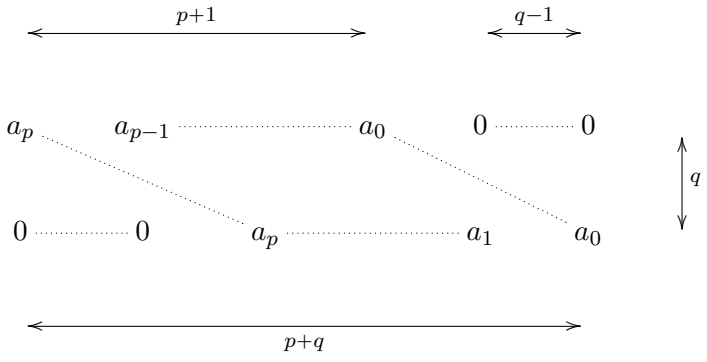
Dans cet anneau, l'élément $\beta_j - \alpha_i$ est irréductible (tout comme $X - Y$ est irréductible dans $\mathbb{Z}[X,Y]$). Le résultant $R = \operatorname{res}(P,Q)$ est un élément de $\mathbb{A}$ parce que tous leurs coefficients peuvent être exprimés à l'aide des α_i et des β_j. Dans $\mathbb{A}$, l'élément $\beta_j - \alpha_i$ divise R. En effet lorsque $\beta_j = \alpha_i$, le déterminant définissant le résultant est nul, ce qui signifie que $\beta_j - \alpha_i$ est un facteur irréductible de R.

Par conséquent il existe un polynôme $T \in \mathbb{A}$ tel que

$$R = T(\alpha_1, \dots, \beta_q) \prod_{i=1}^{p}\prod_{j=1}^{r}(\beta_j - \alpha_i). \tag{8.93}$$

Comptons les degrés. Pour donner une idée de ce calcul de degré, voici comment se présente, au niveau des dimensions, le déterminant :

EqJCaATOH



(8.94)

si les a_i sont les coefficients de P. Mais chacun des a_i est de degré 1 en les α_i, donc le déterminant dans son ensemble est de degré q en les α_i, parce que R contient q lignes telles que (8.94). Le même raisonnement montre que R est de degré p en les β_j. Par ailleurs le polynôme $\prod_{i=1}^{p}\prod_{j=1}^{r}(\beta_j - \alpha_i)$ est de degré p en les β_j et q en les α_i. Nous en déduisons que T doit être un polynôme ne dépendant pas de α_i ou de β_j.

Nous pouvons donc calculer la valeur de T en choisissant un cas particulier. Avec $P(X) = X^p$ et $Q(X) = X^q + 1$, il est vite vu que $R(P,Q) = 1$ et donc que $T = 1$.

Si les polynômes P et Q ne sont pas unitaires, le lemme 8.24 nous permet de conclure.

□

8.2.9 Théorème de Kronecker

Nous considérons K_n l'ensemble des polynômes de $\mathbb{Z}[X]$

(1) unitaires de degré n,
(2) dont les racines dans $\mathbb{C}$ sont de modules plus petits ou égaux à 1,
(3) et qui ne sont pas divisés par X.

Un tel polynôme s'écrit sous la forme

$$P = X^n + \sum_{k=0}^{n-1} a_k X^k. \tag{8.95}$$

ThoOWMNAVp

Théorème 8.28 (Kronecker[104])**.**
Les racines des éléments de K_n sont des racines de l'unité.

Démonstration. Vu que $\mathbb{C}$ est algébriquement clos nous pouvons considérer les racines $\alpha_1, \dots, \alpha_n$ de P dans $\mathbb{C}$. Nous les considérons avec leurs multiplicités.

Soit $R = X^n + \sum_{k=0}^{n-1} b_k X^k$ un élément de K_n dont nous notons $\beta_1, \dots, \beta_n$ les racines dans $\mathbb{C}$. Les relations coefficients-racines stipulent que

$$b_k = \sum_{1 \leqslant i_1 < \dots < i_{n-k} \leqslant n} \prod_{j=1}^{n-k} \beta_{i_j}. \tag{8.96}$$

En prenant le module et en se souvenant que $|\beta_l| \leqslant 1$ pour tout l, nous trouvons que

$$|b_k| \leqslant \binom{n}{n-k}. \tag{8.97}$$

Mais comme $b_k \in \mathbb{Z}$, nous avons

$$b_k \in \Big\{ -\binom{n}{n-k}, -\binom{n}{n-k} + 1, \dots, 0, \dots, \binom{n}{n-k} \Big\} \tag{8.98}$$

qui est de cardinal $2\binom{n}{n-k} + 1$. Nous avons donc

$$\mathrm{Card}(K_n) \leqslant \prod_{k=0}^{n-1} \big(1 + \binom{n}{n-k}\big) < \infty. \tag{8.99}$$

La conclusion jusqu'ici est que K_n est un ensemble fini.

Pour chaque $k \in \mathbb{N}^*$ nous considérons les polynômes

$$P_k = \prod_{i=1}^{n} (X - \alpha_i^k) \tag{8.100a}$$

$$Q_k = X^k - Y \in \mathbb{Z}[X, Y], \tag{8.100b}$$

et puis nous considérons le résultant $R_k = \mathrm{res}_X(P, Q_k) \in \mathbb{Z}[Y]$:

$$R_k = \mathrm{res}_X(P, Q_k) = \begin{pmatrix} 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & & & \\ 0 & \cdots & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 \\ & & & & & & & & \\ 1 & 0 & \dots & 0 & -Y & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & -Y & 0 & \dots & 0 \\ & & \ddots & & & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y \end{pmatrix} \tag{8.101}$$

Cela est un polynôme en Y dont le terme de plus haut degré est $(-1)^n Y^n$. Les petites formules de la proposition 8.27 nous permettent d'exprimer $R_k(Y)$ en termes des racines de P :

$$R_k(Y) = \prod_{i=1}^{n} Q_k(\alpha_i) = \prod_{i=1}^{n}(\alpha_i^k - Y) = (-1)^n \prod_{i=1}^{n}(Y - \alpha_i^k) = (-1)^n P_k(Y). \tag{8.102}$$

Vu que $P \in K_n$ nous savons que les α_i ne sont pas tous nuls ; donc $P_k \in K_n$. Cependant nous avons vu que K_n est un ensemble fini ; donc parmi les P_k, il y a des doublons (et pas un peu) [14]. Nous regardons même l'ensemble des P_{2^n} dans lequel nous pouvons en trouver deux les mêmes. Soit $l > k$ tels que $P_{2^k} = P_{2^l}$. Si α est racine de P_{2^k}, alors il est de la forme $\alpha = \beta^{2^k}$ pour une certaine racines β de P. Par conséquent

$$\alpha^{2^l/2^k} = \alpha^{2^{l-k}} \tag{8.103}$$ EqBEgJtzm

est racine de P_{2^l}. Notons que dans cette expression il n'y a pas de problèmes de définition d'exposant fractionnaire dans $\mathbb{C}$ parce que $l > k$. Vu que (8.103) est racine de P_{2^l}, il est aussi racine de P_{2^k}. Donc

$$\left(\alpha^{2^{l-k}}\right)^{2^{l-k}} = \alpha^{2^{2(l-k)}} \tag{8.104}$$

est racine de P_{2^l} et donc de P_{2^k}. Au final nous savons que tous les nombres de la forme $\alpha^{2^{n(l-k)}}$ sont racines de P_{2^k}. Mais comme P_{2^k} a un nombre fini de racines, nous pouvons en trouver deux égales. Si nous avons

$$\alpha^{2^{n(l-k)}} = \alpha^{2^{m(l-k)}} \tag{8.105}$$

pour certains entiers $m > n$, alors

$$\alpha^{2^{n(l-k)} - 2^{m(l-k)}} = 1, \tag{8.106}$$

ce qui prouve que α est une racine de l'unité. Nous avons donc prouvé que toutes les racines de P_{2^k} sont des racines de l'unité et donc que les racines de P sont racines de l'unité. □

8.3 Orientation

8.3.1 Cas vectoriel

DEFooNVRHooEBHUSu

Proposition-Définition 8.29 ([271]).
Soient deux bases $\mathcal{B}$ et $\mathcal{B}'$ d'un espace vectoriel réel E. Nous définissons la relation $\mathcal{B} \sim \mathcal{B}'$ si et seulement si $\det_{\mathcal{B}}(\mathcal{B}') > 0$ [15].

Cela est une relation d'équivalence [16] *sur l'ensemble des bases de E, et les classes sont les* ***orientations*** *de E.*

Démonstration. Tout est dans le lemme 8.12. D'abord quand $\mathcal{B}$ et $\mathcal{B}'$ sont des bases, $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$ ensuite, nous passons en revue les points qu'il faut pour être une relation d'équivalence.

(1) $\mathcal{B} \sim \mathcal{B}$ parce que $\det_{\mathcal{B}}(\mathcal{B}) = 1 > 0$.

(2) Vu que $\det_{\mathcal{B}}(\mathcal{B}') = \frac{1}{\det_{\mathcal{B}'}(\mathcal{B})}$, les deux sont positifs en même temps ou pas du tout.

(3) Si $\mathcal{B} \sim \mathcal{B}'$ et $\mathcal{B}' \sim \mathcal{B}''$, alors en utilisant la formule

$$\det_{\mathcal{B}}(\mathcal{B}'') = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B}''), \tag{8.107}$$

nous voyons que $\det_{\mathcal{B}}(\mathcal{B}'') > 0$.

□

14. Ici dans [104], il déduit qu'on a un k tel que $P_k = P_1 = P$. Moi je vois pourquoi on a un k et un l tels que $P_k = P_l$, mais pourquoi on peut en trouver un spécialement égal au premier ? Une réponse à cette question permettrait de solidement réduire la lourdeur de la suite de la preuve.

15. Définition 8.10.

16. Définition 1.32.

Lemme 8.30.
Soit un espace vectoriel réel E. L'ensemble des bases de E possède exactement deux orientations[17]

Démonstration. Nous considérons une base $\mathcal{B} = (e_1, \dots, e_n)$[18] à partir de laquelle nous définissons une autre base : $\mathcal{B}' = (-e_1, e_2, \dots, e_n)$. Nous allons prouver que ces deux bases ne sont pas équivalentes, et que toute base de E est équivalente soit à $\mathcal{B}$ soit à $\mathcal{B}'$.

(1) **Au moins deux classes** Le fait que $\det_{\mathcal{B}}(\mathcal{B}') = -1$ vient du fait que $\det_{\mathcal{B}}(\mathcal{B}) = 1$ et que l'application $\det_{\mathcal{B}}$ est n-linéaire ; en multipliant par -1 le premier argument, la valeur du déterminant est multipliée par -1.

Donc les bases $\mathcal{B}$ et $\mathcal{B}'$ ne sont pas équivalentes et il existe au moins deux classes.

(2) **Au plus deux classes** Nous montrons à présent que toute base est équivalente soit à $\mathcal{B}$ soit à $\mathcal{B}'$. Supposons que $\mathcal{B}''$ ne soit pas équivalente à $\mathcal{B}$, c'est-à-dire que $\det_{\mathcal{B}}(\mathcal{B}'') < 0$. Nous utilisons encore la formule (8.29),

$$\underbrace{\det_{\mathcal{B}}(\mathcal{B}'')}_{<0} = \underbrace{\det_{\mathcal{B}}(\mathcal{B}')}_{<0} \det_{\mathcal{B}'}(\mathcal{B}''), \tag{8.108}$$

et nous déduisons que $\det_{\mathcal{B}'}(\mathcal{B}'') > 0$.

□

8.31.
Vu qu'il n'y a que deux classes d'équivalence parmi les bases, nous pouvons utiliser le vocable « avoir la même orientation que » ou « avoir l'orientation contraire de ». Ce n'est pas ambigu.

Proposition 8.32 ([271])**.**
Si $\mathcal{B}$ est une base de l'espace vectoriel E de dimension n, et si τ est une transposition[19] *de S_n, alors la base $\tau(\mathcal{B})$ est de sens contraire.*

Démonstration. Le lemme 8.12(2) dit que $\det_{\mathcal{B}}$ est une forme anti-symétrique ; donc

$$\det_{\mathcal{B}}(\mathcal{B}') = -\det_{\mathcal{B}}\big(\tau(\mathcal{B})'\big). \tag{8.109}$$

Si l'un est positif, l'autre est négatif. Elles ont donc des orientations contraires. □

Corolaire 8.33.
Si $\mathcal{B}$ est une base de l'espace vectoriel E de dimension n, et si $\sigma \in S_n$, la base $\sigma(\mathcal{B})$ a même orientation que $\mathcal{B}$ si et seulement si $\sigma \in A_n$.

Démonstration. Notons c_1 la classe d'orientation de $\mathcal{B}$ et c_2 l'autre classe. La permutation σ se décompose en produit de transpositions dont la parité est fixée (proposition 1.296). Posons $\sigma = \tau_k \dots \tau_1$.

En posant $\mathcal{B}_0 = \mathcal{B}$ et $\mathcal{B}_{l+1} = \tau_{l+1}(\mathcal{B}_l)$, pour tout l, la base $\mathcal{B}_l$ est d'orientation contraire à celle de la base $\mathcal{B}_{l-1}$. Une base sur deux a l'orientation de $\mathcal{B}$ et l'autre sur deux a l'orientation contraire.

Donc $\sigma(\mathcal{B})$ a la même orientation que $\mathcal{B}$ si et seulement si k est pair. Mais $\sigma \in A_n$ si et seulement si k est pair. C'est bon. □

PROPooNBAXooKNUrnk

Proposition-Définition 8.34 ([271])**.**
Soit un espace vectoriel réel, et un endomorphisme f de E. Deux définitions. ITEMooOAXFooLIPHlW

*(1) L'endomorphisme f est **direct** si sont déterminant est strictement positif.* ITEMooNKYCooXTgKJA

*(2) L'endomorphisme **préserve l'orientation** si il transforme toute base de E en une base de même orientation.*

17. Définition 8.29.
18. Nous notons (e_1, e_2) et non $\{e_1, e_2\}$ parce que l'ordre est important.
19. Définition 1.293.

Un endomorphisme est direct si et seulement si il préserve l'orientation.

Démonstration. En deux sens.

(1) **Direct implique préserve l'orientation** Soit une base $\mathcal{B}$ de E et un endomorphisme direct u. D'abord, u est inversible du fait que son déterminant est non nul par la proposition 8.16(2). Donc u transforme une base en une base par le lemme 4.8.
La définition 8.15 du déterminant de u est que

$$\det(u) = \det_{\mathcal{B}}\big(u(\mathcal{B})\big) > 0. \tag{8.110}$$

Donc $\mathcal{B}$ et $u(\mathcal{B})$ ont même orientation.

(2) **Préserve l'orientation implique direct** Le fait que u préserve l'orientation signifie en particulier qu'il transforme une base en une base et qu'il est inversible par le lemme 4.8.
Donc si $\mathcal{B}$ est une base, $u(\mathcal{B})$ est encore une base et nous avons, parce que $\mathcal{B}$ et $u(\mathcal{B})$ ont même orientation,

$$0 < \det_{\mathcal{B}}\big(u(\mathcal{B})\big) = \det(u). \tag{8.111}$$

□

8.3.2 Cas affine

DEFooOTFPooIVkHFP

Définition 8.35.
*Nous disons qu'une application affine $f\colon \mathcal{E} \to \mathcal{E}$ **préserve l'orientation** si sa partie linéaire préserve l'orientation.*

8.4 Hermitien, orthogonal, adjoint

NORMooWGEJooCtGtqZ

8.36.
Une des choses à retenir de la définition de l'opérateur adjoint est que la notion de A^* dépend du produit scalaire considéré.

Il se fait que le plus souvent, sur $\mathbb{R}^n$, nous considérons le produit scalaire usuel et la base canonique. De ce fait, les notions d'opérateur adjoint et d'opérateur transposés se confondent avec la notion de matrice transposée. Ce sont pourtant, en général, trois notions distinctes.

DEFooROVNooFlTbSK

Proposition-Définition 8.37 (Définition de la transposée[1])**.**
Soient deux espaces vectoriels euclidiens ou hermitiens E et F et une application linéaire $A\colon E \to F$.

ITEMooRUZWooSZgGnf

(1) Il existe une unique application linéaire $B\colon F \to E$ telle que

$$\langle Ax, y\rangle_F = \langle x, By\rangle_E \tag{8.112}$$

EQooHWYKooFzAGgB

pour tout $x \in E$ et $y \in F$.

ITEMooXXEUooPtfPKY

(2) Si $\{e_i\}$ est une base orthonormée de E et $\{f_\alpha\}$ est une base orthonormée de F, alors la matrice de A et B pour ces bases sont liées par

$$B_{i\alpha} = A_{\alpha i}. \tag{8.113}$$

EQooUSNVooQtRNGL

*L'application B ainsi définie set nommée **adjoint** de A et sera notée $B = A^*$.*

Démonstration. Pour l'unicité, nous écrivons la condition avec $x = e_j$ pour obtenir :

$$\langle Ae_j, y\rangle = \langle e_j, By\rangle = (By)_j \tag{8.114}$$

c'est-à-dire que les coefficients $B(y)_j$ de $B(y)$ dans la base canonique sont fixés par la condition.

Pour l'existence, il suffit de vérifier que poser

$$B(y) = \sum_j \langle Ae_j, y\rangle e_j \tag{8.115}$$

fonctionne. Pour cela il faut utiliser la bilinéarité du produit scalaire et le fait que $\langle x, e_j\rangle = x_j$. Nous avons :

$$\begin{aligned}\langle x, B(y)\rangle &= \langle x, \sum_j \langle A(e_j), y\rangle e_j\rangle && (8.116a)\\ &= \sum_j \langle A(e_j), y\rangle\langle x, e_j\rangle && (8.116b)\\ &= \sum_j \langle A(x_j e_j), y\rangle && (8.116c)\\ &= \langle A(x), y\rangle. && (8.116d)\end{aligned}$$

En ce qui concerne la matrice de l'application B ainsi définie, nous écrivons la condition (8.112) avec $y = e'_\alpha$ et $x = e_i$, de telle sorte que

$$A(x) = A(e_i) = \sum_\beta A_{\beta i} e'_\beta \tag{8.117}$$

et

$$B(y) = B(e'_\alpha) = \sum_j B_{j\alpha} e_j. \tag{8.118}$$

Alors nous avons :

$$\sum_\beta A_{\beta i}\langle e'_\beta, e'_\alpha\rangle = \sum_j B_{j\alpha}\langle e_i, e_j\rangle, \tag{8.119}$$

donc

$$A_{\alpha i} = B_{i\alpha}. \tag{8.120}$$

□

8.38.
À cause de l'expression (8.113) pour la matrice de A^*, cette application est souvent appelé **transposé** de A et noté A^t. Nous allons cependant voir plus tard (définition8.189) que la transposée de A est une application $A^t\colon F^* \to E^*$. Il nous arrivera cependant d'écrire des égalités comme $\langle Ax, y\rangle = \langle x, A^t y\rangle$.

PROPooSHZMooGwdfBd

Proposition 8.39.
En ce qui concerne le déterminant,

$$\det(A^*) = \det(A)^* \tag{8.121}$$

où l'étoile à droite dénote la conjugaison complexe dans $\mathbb{C}$.

Démonstration. Écrivons l'expression explicite (8.25) du déterminant. Le tout avec la base canonique :

$$\det(A) = \det_{(e_1,\dots,e_n)}(Ae_1, \dots, Ae_n) = \sum_{\sigma\in S_n} \epsilon(\sigma)\prod_{i=1}^n e^*_{\sigma(i)}(Ae_i). \tag{8.122}$$

Mais nous pouvons développer :

$$e^*_{\sigma(i)}(Ae_i) = \langle e_{\sigma(i)}, Ae_i\rangle = \langle A^* e_{\sigma(i)}, e_i\rangle = \langle e_i, A^* e_{\sigma(i)}\rangle^* = e^*_i(A^* e_{\sigma(i)})^*. \tag{8.123}$$

Notez que dans la dernière expression, les trois * ont trois significations différentes. Par conséquent,

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma)\prod_{i=1}^n e^*_i(A^* e_{\sigma(i)})^*. \tag{8.124}$$

Mais $e_i^*(A^*e_{\sigma(i)}) = e_{\sigma(j)}^*(A^*e_j)$ pour $j = \sigma(i)$, donc le produit ne change pas si on déplace le σ :

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} e_{\sigma(i)}^*(A^*e_i)^* = \det(A^*)^*. \tag{8.125}$$

Nous avons donc $\det(A) = \det(A^*)^*$, c'est-à-dire $\det(A)^* = \det(A^*)$. Pour information, la dernière étoile est la conjugaison complexe. □

PROPooVPSYooRuoEFi

Proposition 8.40 ([1])**.**
Si $A\colon E_2 \to E_3$ et $B\colon E_1 \to E_2$ sont des applications linéaires, alors

$$(AB)^* = B^*A^* \tag{8.126}$$

où la « multiplication » est la composition.

Démonstration. L'existence de $(AB)^*$, de A^* et de B^* ne donne pas lieu à débat parce que la proposition 8.37 ne souffre pas de discussions. La propriété que $(AB)^*$ est unique a avoir est que

$$\langle ABx, y\rangle = \langle x, (AB)^*y\rangle \tag{8.127}$$

pour tout $x \in E_1$ et $y \in E_3$. Or l'application B^*A^* possède également cette propriété parce que

$$\langle x, B^*A^*y\rangle = \langle Bx, A^*y\rangle = \langle ABx, y\rangle. \tag{8.128}$$

La partie unicité de la proposition 8.37 nous impose donc d'accepter que les applications $(AB)^*$ et B^*A^* sont en réalité les mêmes[20]. □

8.41.
Un grand moment d'utilisation de la notion d'adjoint pour un opérateur non carré sera la définition d'une intégrale sur une variété ; en particulier dans la proposition 20.9.

Lemme 8.42.
Si E est un espace euclidien, un endomorphisme $f\colon E \to E$ est autoadjoint si et seulement si pour tout $x, y \in E$ nous avons $\langle x, f(y)\rangle = \langle f(x), y\rangle$.

Démonstration. Dans le sens direct, nous avons

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle = \langle x, f(y)\rangle. \tag{8.129}$$

La première égalité est la définition de f^* et la seconde est l'hypothèse $f = f^*$.

Dans l'autre sens, l'hypothèse est que l'endomorphisme f vérifie $\langle x, f(y)\rangle = \langle f(x), y\rangle$. Mais la proposition 8.37(1) spécifie que f^* est l'unique endomorphisme à satisfaire cette égalité. Donc $f = f^*$. □

8.4.1 Opérateur orthogonal, matrice orthogonale

DEFooYKCSooURQDoS

Définition 8.43.
*Un opérateur est **orthogonal** lorsque $A^* = A^{-1}$ où A^* est l'adjoint de A définit en 8.37.*

DEFooUHANooLVBVID

Définition 8.44.
*Une matrice U est **orthogonale** si $U^t = U^{-1}$. Le **groupe orthogonal** noté $\mathrm{O}(n)$ est l'ensemble des matrices orthogonales $n \times n$.*

LEMooSSALooSBFzJb

Lemme 8.45.
Soit un opérateur $A\colon \mathbb{R}^n \to \mathbb{R}^n$ muni du produit scalaire usuel. Il est orthogonal si et seulement si sa matrice dans la base canonique est orthogonale[21].

20. Et ce même si vous croyez les avoir déjà vu ensemble dans la même pièce.
21. Définition 8.44.

Démonstration. Soit la base canonique $\{e_i\}_{i=1,\dots,n}$ de $\mathbb{R}^n$. Nous avons

$$\langle AA^*e_i, e_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}, \tag{8.130}$$

donc $\big((AA^*)e_i\big)_j = \delta_{ij}$, ou encore $(AA^*)_{ij} = \delta_{ij}$, ce qui signifie que la matrice AA^* est l'identité. □

PropKBCXooOuEZcS

Proposition 8.46 (Thème 78)**.**
À propos de matrices orthogonales.

ITEMooHSTAooIbVrwa

(1) *L'ensemble des matrices réelles orthogonales forme un groupe noté* $\mathrm{O}(n,\mathbb{R})$.

(2) *Si A est une matrice orthogonale, alors* $\det(A) = \pm 1$.

ITEMooOWMBooHUatNb

(3) *Le groupe* $\mathrm{O}(n)$ *est le groupe des isométries linéaires*[22] *de* $\mathbb{R}^n$.

Démonstration. Commençons par prouver que $\mathrm{O}(n,\mathbb{R})$ est un groupe. La matrice $\mathbb{1}$ est orthogonale. De plus si A et B sont orthogonale, la matrice produit AB est orthogonale :

$$(AB)(AB)^t = ABB^tA^t = A\mathbb{1}A^t = \mathbb{1}. \tag{8.131}$$

Nous avons donc bien un groupe.

En ce qui concerne le déterminant, $AA^t = \mathbb{1}$ donne $\det(A)\det(A^t) = 1$, mais la proposition 8.39 dit que $\det(A) = \det(A^t)$, donc $\det(A)^2 = 1$. D'où le fait que $\det(A) = \pm 1$.

D'autre part si A est une isométrie de $\mathbb{R}^n$ alors pour tout $x, y \in \mathbb{R}^n$ nous avons $\langle Ax, Ay\rangle = \langle x, y\rangle$. En particulier,

$$\langle A^tAx, y\rangle = \langle x, y\rangle \tag{8.132}$$

pour tout $x, y \in \mathbb{R}^n$. En prenant $y = e_i$ nous trouvons

$$(A^tAx)_i = x_i, \tag{8.133}$$

ce qui signifie que pour tout x, $A^tAx = x$, ou encore que A^tA est l'identité.

Réciproquement si A^tA est l'identité nous avons

$$\langle x, y\rangle = \langle A^tAx, y\rangle = \langle Ax, Ay\rangle, \tag{8.134}$$

ce qui prouve que A est une isométrie. □

En ce qui concerne les valeurs propres des matrices de $\mathrm{O}(n)$ ainsi que leurs formes canoniques (avec des fonctions trigonométriques) pour $\mathrm{O}(3)$ et $\mathrm{SO}(3)$, ce sera pour la proposition 18.220 et ce qui s'ensuit.

DEFooJLNQooBKTYUy

Définition 8.47.
Le sous-groupe des matrices orthogonales de déterminant 1 *est le groupe* ***spécial orthogonal*** *noté* $\mathrm{SO}(n)$.

8.5 Topologie

8.5.1 Boules et sphères

Un espace vectoriel normé (définition 7.158) vient avec sa topologie métrique (théorème 7.117). Sphères et boules fermées viennent dans la définition 7.134.

Définition 8.48.
Une partie A de V est dite ***bornée*** *si il existe un réel R tel que $A \subset B(0_V, R)$.*

Une partie est donc bornée si elle est contenue dans une boule de rayon fini.

22. Au sens où, parmi les applications linéaires, les isométries sont les éléments de $\mathrm{O}(n)$. À part ça, il y a aussi les translations, mais c'est une autre histoire qui vous sera contée une autre fois.

Exemple 8.49.
Dans $\mathbb{R}$, les boules sont les intervalles ouverts et fermés tandis que la sphère est donnée par les points extrêmes des intervalles :

$$\begin{aligned} B(a,r) &=]a-r, a+r[, \\ \bar{B}(a,r) &= [a-r, a+b], \\ S(a,r) &= \{a-r, a+r\}. \end{aligned} \tag{8.135}$$

△

Exemple 8.50.
Si nous considérons $\mathbb{R}^2$, la situation est plus riche parce que nous avons plus de normes. Essayons de voir les sphères de centre $(0,0) \in \mathbb{R}^2$ et de rayon r pour les normes $\|.\|_1$, $\|.\|_2$ et $\|.\|_\infty$.

Pour la norme $\|.\|_1$, la sphère de rayon r est donnée par l'équation

$$|x| + |y| = r. \tag{8.136}$$

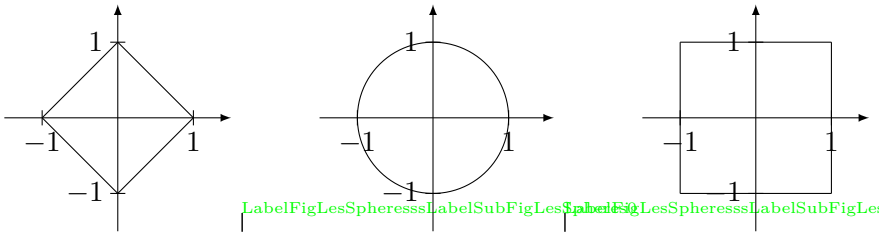
Pour la norme $\|.\|_2$, l'équation de la sphère de rayon r est

$$\sqrt{x^2 + y^2} = r, \tag{8.137}$$

et pour la norme supremum, la sphère de rayon r a pour équation

$$\max\{|x|, |y|\} = r. \tag{8.138}$$

Elles sont dessinées sur la figure 8.1



(a) La sphère unité pour la norme $\|.\|_1$ (b) La sphère unité pour la norme $\|.\|_2$ (c) La sphère unité pour la norme $\|.\|_\infty$

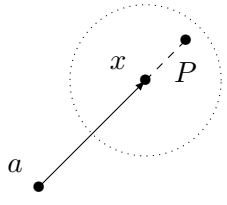
FIGURE 8.1: Les sphères de rayon 1 pour les trois normes classiques. LabelFigLesSpheres

△

PropBoitPtLoin

Proposition 8.51.
Soient V un espace vectoriel normé, a dans V et x tel que $d(a,x) = r$, c'est-à-dire $x \in S(a,r)$. Dans ce cas, toute boule centrée en x contient un point P tel que $d(P,a) > r$ et un point Q tel que $d(Q,a) < r$.

Démonstration. Soit une boule de rayon δ autour de x. Le but est de trouver un point P tel que $d(P,a) > r$ et $d(P,x) < \delta$. Pour cela, nous prenons P sur la même droite que x (en partant de a), mais juste « un peu plus loin », comme sur la figure suivante :



Plus précisément, nous considérons le point

$$P = x + \frac{v}{N} \tag{8.139}$$

où $v = x - a$ et N est suffisamment grand pour que $d(x, P)$ soit plus petit que δ. Cela est toujours possible parce que

$$d(P, x) = \|P - x\| = \frac{\|v\|}{N} \tag{8.140}$$

peut être rendu aussi petit que l'on veut par un choix approprié de N. Montrons maintenant que $d(a, P) > d(a, x)$:

$$\begin{aligned} d(a, P) &= \|a - x - \frac{v}{N}\| \\ &= \|a - x + \frac{a}{N} - \frac{x}{N}\| \\ &= \|\big(1 + \frac{1}{N}\big)(a - x)\| \\ &> \|a - x\| = d(a, x). \end{aligned} \tag{8.141}$$

Nous laissons en exercice le soin de trouver un point Q tel que $d(Q, a) < r$ et $d(Q, x) < \delta$. □

8.5.2 Ouverts, fermés, intérieur et adhérence

Définition 8.52.
*Soit $(V, \|.\|)$ un espace vectoriel normé et A, une partie de V. Un point a est dit **intérieur** à A si il existe une boule ouverte centrée en a et contenue dans A.*

*On appelle **l'intérieur** de A l'ensemble des points qui sont intérieurs à A. Nous notons* $\text{Int}(A)$ *l'intérieur de A.*

Notons que $\text{Int}(A) \subset A$ parce que si $a \in \text{Int}(A)$, nous avons $B(a, r) \subset A$ pour un certain r et en particulier $a \in A$.

Exemple 8.53.
Trouver l'intérieur d'un intervalle dans $\mathbb{R}$ consiste à « ouvrir là où c'est fermé ».

(1) $\text{Int}\big([0, 1[\big) =]0, 1[$.
Prouvons d'abord que $]0, 1[\subset \text{Int}([0, 1[)$. Si $a \in]0, 1[$, alors a est strictement supérieur à 0 et strictement inférieur à 1. Dans ce cas, la boule de centre a et de rayon $\frac{\min\{a, 1-a\}}{2}$ est contenue dans $]0, 1[$ (voir figure 8.2). Cela prouve que a est dans l'intérieur de $[0, 1[$.

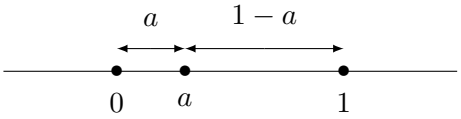


FIGURE 8.2: Trouver le rayon d'une boule autour de a. Une boule qui serait centrée en a avec un rayon strictement plus petit à la fois de a et de $1 - a$ est entièrement contenue dans le segment $]0, 1[$. LabelFigIntervalleUn

Prouvons maintenant que $\text{Int}\big([0, 1[\big) \subset]0, 1[$. Vu que l'intérieur d'un ensemble est inclus dans l'ensemble, nous savons déjà que $\text{Int}\big([0, 1[\big) \subset [0, 1[$. Nous devons donc seulement montrer que 0 n'est pas dans l'intérieur de $[0, 1[$. C'est le cas parce que toute boule du type $B(0, r)$ contient le point $-r/2$ qui n'est pas dans $[0, 1[$.

(2) $\text{Int}\Big([0, \infty[\Big) =]0, \infty[$.

(3) $\text{Int}\big(]2, 3[\big) =]2, 3[$.

△

ExempleIntBoules

Exemple 8.54.
Les intérieurs des boules et sphères sont importantes à savoir.

(1) $\text{Int}\big(B(a,r)\big) = B(a,r)$. Si $x \in B(a,r)$, nous avons $d(a,x) < r$. Alors la boule $B\big(x, r-d(x,a)\big)$ est incluse à $B(a,r)$, et donc x est dans l'intérieur de $B(a,r)$. Conseil : faire un dessin.

(2) $\text{Int}\big(\bar{B}(a,r)\big) = B(a,r)$. Par le point précédent, la boule $B(a,r)$ est certainement dans l'intérieur de la boule fermée. Il reste à montrer que les points de $\bar{B}(a,r)$ qui ne sont pas dans $B(a,r)$ ne sont pas dans l'intérieur. Ces points sont ceux dont la distance à a est *égale* à r. Le résultat découle alors de la proposition 8.51.

(3) $\text{Int}\big(S(a,r)\big) = \varnothing$. Si $x \in S(a,r)$, toute boule centrée en a contient des points qui ne sont pas à distance r de a.

Notez que la sphère est un exemple d'ensemble non vide mais d'intérieur vide.

△

Définition 8.55.
Une partie A de l'espace vectoriel normé $(V, \|.\|)$ est dite ***ouverte*** *si chacun de ses points est intérieur. La partie A est donc ouverte si $A \subset \text{Int}(A)$. Par convention, nous disons que l'ensemble vide $\varnothing$ est ouvert.*

Une partie est dite ***fermée*** *si son complémentaire est ouvert. La partie A est donc fermée si $V \backslash A$ est ouverte.*

Remarque : un ensemble A est ouvert si et seulement si $\text{Int}(A) = A$.

Définition 8.56.
Une partie A de l'espace vectoriel normé V est dite ***compacte*** *si elle est fermée et bornée.*

Nous verrons tout au long de ce cours que les ensembles compacts, et les fonctions définies sur ces ensembles ont de nombreuses propriétés agréables.

ExempleFermeIntevrR

Exemple 8.57.
En ce qui concerne les intervalles de $\mathbb{R}$,

— $]1,2[$ est ouvert ;

— $[3,4]$ est fermé ;

— $[5,6[$ n'est ni ouvert ni fermé ;

Les intervalles fermés de $\mathbb{R}$ sont toujours compacts. △

PropTopologieAx

Proposition 8.58.
Soit V un espace vectoriel normé.

(1) L'ensemble V lui-même et le vide sont à la fois fermés et ouverts.

(2) Toute union d'ouverts est ouverte.

(3) Toute intersection finie *d'ouverts est ouverte.*

ItemPropTopologieAxiv

(4) Le vide et V sont les seules parties de V à être à la fois fermées et ouvertes.

Démonstration. L'ingrédient principal de cette démonstration est que si a est un point d'un ouvert $\mathcal{O}$, alors il existe une boule autour de a contenue dans $\mathcal{O}$ parce que a doit être dans l'intérieur de $\mathcal{O}$.

(1) Nous avons déjà dit que, par définition, l'ensemble vide est ouvert. Cela implique que V lui-même est fermé (parce que son complémentaire est le vide). De plus, V est ouvert parce que toutes les boules sont inclues à V. Le vide est alors fermé (parce que son complémentaire est V).

(2) Soit une famille $(\mathcal{O}_i)_{i\in I}$ d'ouverts[23], et l'union

$$\mathcal{O} = \bigcup_{i\in I} \mathcal{O}_i. \tag{8.142}$$

Soit maintenant $a \in \mathcal{O}$. Nous devons prouver qu'il existe une boule centrée en a entièrement contenue dans $\mathcal{O}$. Étant donné que $a \in \mathcal{O}$, il existe $i \in I$ tel que $a \in \mathcal{O}_i$ (c'est-à-dire que a est au moins dans un des $\mathcal{O}_i$). Par hypothèse l'ensemble $\mathcal{O}_i$ est ouvert et donc tous ses points (en particulier a) sont intérieurs ; il existe donc une boule $B(a,r)$ centrée en a telle que $B(a,r) \subset \mathcal{O}_i \subset \mathcal{O}$.

(3) Soit une famille finie d'ouverts $(\mathcal{O}_k)_{k\in\{1,\ldots,n\}}$, et $a \in \mathcal{O}$ où

$$\mathcal{O} = \bigcap_{k=1}^{n} \mathcal{O}_k. \tag{8.143}$$

Vu que a appartient à chaque ouvert $\mathcal{O}_k$, nous pouvons trouver, pour chacun de ces ouverts, une boule $B(a,r_k)$ contenue dans $\mathcal{O}_k$. Chacun des r_k est strictement positif, et nous n'en avons qu'un nombre fini, donc le nombre $r = \min\{r_1,\ldots,r_n\}$ est strictement positif. La boule $B(a,r)$ est incluse dans toutes les autres (parce que $B(a,r) \subset B(a,r')$ lorsque $r \leqslant r'$), par conséquent

$$B(a,r) \subset \bigcap_{k=1}^{n} B(a,r_k) \subset \bigcap_{k=1}^{n} \mathcal{O}_k = \mathcal{O}, \tag{8.144}$$

c'est-à-dire que la boule de rayon r est une boule centrée en a contenue dans $\mathcal{O}$, ce qui fait que a est intérieur à $\mathcal{O}$.

(4) Je n'ai pas encore de démonstration de ce point[24].

□

La proposition dit que toute intersection *finie* d'ouvert est ouverte. Il est faux de croire que cela se généralise aux intersections infinies, comme le montre l'exemple suivant :

$$\bigcap_{i=1}^{\infty}]-\frac{1}{n},\frac{1}{n}[= \{0\}. \tag{8.145}$$

Chacun des ensembles $]-\frac{1}{n},\frac{1}{n}[$ est ouvert, mais le singleton $\{0\}$ est fermé (pourquoi ?).

Nous reportons à la proposition 1.461 la preuve du fait que tout ensemble borné de $\mathbb{R}$ possède un infimum et un supremum.

Définition 8.59.
*L'ensemble des ouverts de V est la **topologie** de V. La topologie dont nous parlons ici est dite **induite** par la norme $\|.\|$ de V (parce que cette norme définit la notion de boule et qu'à son tour la notion de boule définit la notion d'ouverts). Un **voisinage** de a dans V est un ensemble contenant un ouvert contenant a.*

Il existe de nombreuses topologies sur un espace vectoriel donné, mais certaines sont plus fameuses que d'autres. Dans le cas de $V = \mathbb{R}^n$, la topologie **usuelle** est celle induite par la norme euclidienne. Lorsque nous parlons de boules, de fermés, de voisinages ou d'autres notions topologiques (y compris de convergence, voir plus bas) dans $\mathbb{R}^n$, nous sous-entendons toujours la topologie de la norme euclidienne.

Exemple 8.60.
Les ensembles suivants sont des voisinages de 3 dans $\mathbb{R}$:

23. L'ensemble I avec lequel nous « numérotons » les ouverts $\mathcal{O}_i$ est *quelconque*, c'est-à-dire qu'il peut être $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}^n$ ou n'importe quel autre ensemble, fini ou infini.

24. Rédigez une preuve, photographiez votre feuille, et envoyez-la moi.

— $]1,5[$;
— $[0,10]$;
— $\mathbb{R}$.

Les ensembles suivants ne sont pas des voisinages de 3 dans $\mathbb{R}$:

— $]1,3[$;
— $]1,3]$;
— $[0,5[\backslash\{3\}$.

△

Proposition 8.61.
Dans un espace vectoriel normé,

(1) toute intersection de fermés est fermée ;

(2) toute union finie *de fermés est fermée.*

Encore une fois, l'hypothèse de finitude de l'intersection est indispensable comme le montre l'exemple suivant :

$$\bigcup_{n=1}^{\infty}[-1+\frac{1}{n},1-\frac{1}{n}]=]-1,1[. \tag{8.146}$$

Chacun des intervalles dont on prend l'union est fermé tandis que l'union est ouverte.

LEMooHPQTooHKdoiL

Lemme 8.62.
Soit A, une partie de l'espace vectoriel normé V. Un point $a\in V$ est adhérent[25] *à A dans V si et seulement si pour tout $\varepsilon>0$,*

$$B(a,\varepsilon)\cap A\neq\varnothing. \tag{8.147}$$

Un point peut être adhérent à A sans faire partie de A, et nous avons toujours $A\subset \mathrm{Adh}(A)$.

EXooICLBooJzQFNY

Exemple 8.63.
La terminologie « fermeture » de A pour désigner $\bar{A}$ provient de deux origines.

(1) L'ensemble $\bar{A}$ est le plus petit fermé contenant A. Cela signifie que si B est un fermé qui contient A, alors $\bar{A}\subset B$. Cela est fondamentalement le sens de la définition 7.20.

(2) Pour les intervalles dans $\mathbb{R}$, trouver $\bar{A}$ revient à fermer les extrémités qui sont ouvertes, comme on en a parlé dans l'exemple 8.57.

△

Exemple 8.64.
Dans $\mathbb{R}$, l'infimum et le supremum d'un ensemble sont des points adhérents. En effet si M est le supremum de $A\subset\mathbb{R}$, pour tout ε, il existe un $a\in A$ tel que $a>M-\varepsilon$, tandis que $M>a$. Cela fait que $a\in B(M,\varepsilon)$, et en particulier que pour tout rayon ε, nous avons $B(M,\varepsilon)\cap A\neq\varnothing$.

Le même raisonnement montre que l'infimum est également dans l'adhérence de A. △

ParlerEncoredeF

Exemple 8.65.
Il ne faut pas conclure de l'exemple précédent qu'un point limite ou adhérent est automatiquement un minimum ou un maximum. En effet, si nous regardons l'ensemble formé par les points de la suite $x_n=(-1)^n/n$, le nombre zéro est un point adhérent et une limite, mais pas un infimum ni un maximum. △

Lemme 8.66.
Si B est une partie fermée de V, alors $B=\bar{B}$.

25. Définition 7.20.

Démonstration. Supposons qu'il existe $a \in \bar{B}$ tel que $a \notin B$. Alors il n'y a pas d'ouverts autour de a qui soit contenu dans $\complement B$. Cela prouve que $\complement B$ n'est pas ouvert, et par conséquent que B n'est pas fermé. Cela est une contradiction qui montre que tout point de $\bar{B}$ doit appartenir à B lorsque B est fermé. □

Exemple 8.67.
Au niveau des intervalles dans $\mathbb{R}$, prendre l'adhérence consiste à « fermer là où c'est ouvert ». Attention cependant à ne pas fermer l'intervalle en l'infini.

(1) $\overline{[0,2[} = [0,2]$.
(2) $\overline{]3,\infty[} = [3,\infty[$.

Si au lieu de travailler dans $\mathbb{R}$, vous travaillez dans une extension comme le compactifié d'Alexandrov $\hat{\mathbb{R}}$ (définition 7.106) ou dans la droite réelle achevée (définition 12.24), vous devez un peu réfléchir avant de décider si il faut fermer les intervalles en $\pm\infty$ ou en ω. Bref, ne faites rien mécaniquement et posez vous des questions de topologie quand vous cherchez des fermetures. △

PROPooEURXooHHpkhd

Proposition 8.68.
Soit V un espace vectoriel normé et $a \in V$. Les trois conditions suivantes sont équivalentes :

(1) $a \in \bar{A}$;
(2) il existe une suite d'éléments x_n dans A qui converge vers a ;
(3) $d(a,A) = 0$.

Notez que dans cette proposition, nous ne supposons pas que a soit dans A.

PropComleIntBar

Proposition 8.69.
Pour toute partie A d'un espace vectoriel normé nous avons

(1) $V\backslash\bar{A} = \operatorname{Int}(V\backslash A)$,
(2) $V\backslash \operatorname{Int}(A) = \overline{V\backslash A}$.

En utilisant les notations du complémentaire (1.1.5), les deux points de la proposition se récrivent

(1) $\complement\bar{A} = \operatorname{Int}(\complement A)$,
(2) $\complement \operatorname{Int}(A) = \overline{\complement A}$.

ItemLemPropComplementiv

Démonstration. Nous avons $a \in V\backslash\bar{A}$ si et seulement si $a \notin \bar{A}$. Or ne pas être dans $\bar{A}$ signifie qu'il existe un rayon ε tel que la boule $B(a,\varepsilon)$ n'intersecte pas A. Le fait que la boule $B(a,\varepsilon)$ n'intersecte pas A est équivalent à dire que $B(a,\varepsilon) \subset V\backslash A$. Or cela est exactement la définition du fait que a est à l'intérieur de $V\backslash A$. Nous avons donc montré que $a \in V\backslash\bar{A}$ si et seulement si $a \in \operatorname{Int}(V\backslash A)$. Cela prouve la première affirmation.

Pour prouver la seconde affirmation, nous appliquons la première au complémentaire de A : $\complement(\overline{\complement A}) = \operatorname{Int}(\complement\complement A)$. En prenant le complémentaire des deux membres nous trouvons successivement

$$\begin{aligned} \complement\complement(\overline{\complement A}) &= \complement \operatorname{Int}(\complement\complement A), \\ \overline{\complement A} &= \complement \operatorname{Int}(A), \end{aligned} \tag{8.148}$$

ce qu'il fallait démontrer. □

Attention à ne pas confondre $\complement\bar{A}$ et $\overline{\complement A}$. Ces deux ensembles ne sont pas égaux. En effet, en tant que complément d'un fermé, l'ensemble $\complement\bar{A}$ est certainement ouvert, tandis que, en tant que fermeture, l'ensemble $\overline{\complement A}$ est fermé. Pouvez-vous trouver des exemples d'ensembles A tels que $\complement\bar{A} = \overline{\complement A}$?

PROPooRAZUooVMxPic

Proposition 8.70.
Soient A et B deux parties de l'espace vectoriel normé V.

(1) *Pour les inclusions, si* $A \subset B$, *alors* $\text{Int}(A) \subset \text{Int}(B)$ *et* $\bar{A} \subset \bar{B}$.
(2) *Pour les unions,* $\overline{A \cup B} = \bar{A} \cup \bar{B}$ *et* $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$.
(3) *Pour les intersections,* $\text{Int}(A) \cap \text{Int}(B) = \text{Int}(A \cap B)$ *et* $\text{Int}(A) \cup \text{Int}(B) \subset \text{Int}(A \cup B)$.

Démonstration. (1) Si a est dans l'intérieur de A, il existe une boule autour de a contenue dans A. Cette boule est alors contenue dans B et donc est une boule autour de a contenue dans B, ce qui fait que a est dans l'intérieur de B. Si maintenant a est dans l'adhérence de A, toute boule centrée en a contient un élément de A et donc un élément de B, ce qui prouve que a est dans l'adhérence de B.

(2) Nous avons $A \subset A \cup B$ et donc, en utilisant le premier point, $\bar{A} \subset \overline{A \cup B}$. De la même manière, $\bar{B} \subset \overline{A \cup B}$. En prenant l'union, $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$.
Réciproquement, soit $a \in \overline{A \cup B}$ et montrons que $a \in \bar{A} \cup \bar{B}$. Supposons par l'absurde que a ne soit ni dans $\bar{A}$ ni dans $\bar{B}$. Il existe donc des rayons ε_1 et ε_2 tels que

$$\begin{aligned} B(a, \varepsilon_1) \cap A &= \emptyset, \\ B(a, \varepsilon_2) \cap B &= \emptyset. \end{aligned} \tag{8.149}$$

En prenant $r = \min\{\varepsilon_1, \varepsilon_2\}$, la boule $B(a, r)$ est incluse aux deux boules citées et donc n'intersecte ni A ni B. Donc $a \notin \overline{A \cup B}$, d'où la contradiction.

(3) Si nous appliquons le second point à $\complement A$ et $\complement B$, nous trouvons

$$\overline{\complement A \cup \complement B} = \overline{\complement A} \cup \overline{\complement B}. \tag{8.150}$$

En utilisant les propriétés du lemme 1.29, le membre de gauche devient

$$\overline{\complement A \cup \complement B} = \overline{\complement(A \cap B)} = \complement \,\text{Int}(A \cap B), \tag{8.151}$$ Eq2707CACBCAB

tandis que le membre de droite devient

$$\overline{\complement A} \cup \overline{\complement B} = \complement \,\text{Int}(A) \cup \complement \,\text{Int}(A) = \complement\Big(\text{Int}(A) \cap \text{Int}(B)\Big). \tag{8.152}$$ Eq2707cAcBACAACB

En égalisant le membre de droite de (8.151) avec celui de (8.152) et en passant au complémentaire nous trouvons

$$\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B), \tag{8.153}$$

comme annoncé.
La dernière affirmation provient du fait que $\text{Int}(A) \subset \text{Int}(A \cup B)$ et de la propriété équivalente pour B.

□

Remarque 8.71.
Nous avons prouvé que $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$. Il arrive que l'inclusion soit stricte, comme dans l'exemple suivant. Si nous prenons $A = [0, 1]$ et $B =]1, 2]$, nous avons $A \cap B = \emptyset$ et donc $\overline{A \cap B} = \emptyset$. Par contre nous avons $\bar{A} \cap \bar{B} = \{1\}$.

Définition 8.72.
La ***frontière*** *d'un sous-ensemble A de l'espace vectoriel normé V est l'ensemble des points $a \in V$ tels que*

$$\begin{aligned} B(a, r) \cap A &\neq \emptyset, \\ B(a, r) \cap \complement A &\neq \emptyset, \end{aligned} \tag{8.154}$$

pour tout rayon r. En d'autres termes, toute boule autour de a contient des points de A et des points de $\complement A$. La frontière de A se note ∂A.

PropDescFrpbsmI

Proposition 8.73.
La frontière d'une partie A d'un espace vectoriel normé V s'exprime sous la forme

$$\partial A = \bar{A} \backslash \text{Int}(A). \tag{8.155}$$

Démonstration. Le fait pour un point a de V d'appartenir à $\bar{A}$ signifie que toute boule centrée en a intersecte A. De la même façon, le fait de ne pas appartenir à $\text{Int}(A)$ signifie que toute boule centrée en a intersecte $\complement A$. □

La description de la frontière donnée par la proposition 8.73 est celle qu'en pratique nous utilisons le plus souvent. Dans certains textes, elle est prise comme définition de la frontière.

Lemme 8.74.
La frontière de A peut également s'exprimer des façons suivantes :

$$\partial A = \bar{A} \cap \complement \,\text{Int}(A) = \bar{A} \cap \overline{\complement A}, \tag{8.156}$$

Démonstration. En partant de $\partial A = \bar{A} \setminus \text{Int}(A)$, la première égalité est une application de la propriété (2) du lemme 1.29. La seconde égalité est alors la proposition 8.69. □

Exemple 8.75.
Dans $\mathbb{R}$, la frontière d'un intervalle est la paire constituée des points extrêmes. En effet

$$\partial[a,b[= \overline{[a,b[} \setminus \text{Int}\big([a,b[\big) = [a,b] \setminus]a,b[= \{a,b\}. \tag{8.157}$$

Toujours dans $\mathbb{R}$ nous avons

$$\partial\mathbb{R} = \bar{\mathbb{R}} \setminus \text{Int}(\mathbb{R}) = \mathbb{R} \setminus \mathbb{R} = \varnothing, \tag{8.158}$$

et

$$\partial\mathbb{Q} = \bar{\mathbb{Q}} \setminus \text{Int}(\mathbb{Q}) = \mathbb{R} \setminus \varnothing = \mathbb{R}. \tag{8.159}$$

△

Exemple 8.76.
Dans $\mathbb{R}^n$, nous avons

$$\partial B(a,r) = \partial \bar{B}(a,r) = S(a,r). \tag{8.160}$$

Cela est un boulot pour la proposition 8.51. Si $x \in S(a,r)$ alors tout boule autour de x contient des points à distance strictement plus grande et plus petite que $d(a,x)$, c'est-à-dire des points dans $B(a,r)$ et hors de $B(a,r)$. Cela prouve que les points de $S(a,r)$ font partie de $\partial B(a,r)$, c'est-à-dire que $S(a,r) \subset \partial B(a,r)$; et idem pour $\bar{B}(a,r)$.

Pour prouver l'inclusion inverse, soit $x \in \partial B(a,r)$. Vu que toute boule autour de x contient des points intérieurs à $B(a,r)$, pour tout $\epsilon > 0$, $d(a,x) - \epsilon < r$, c'est-à-dire que $d(a,x) \leqslant r$. De la même manière toute boule autour de x contient des points hors de $B(a,r)$ signifie que pour tout ϵ, $d(a,x) + \epsilon > r$ ou encore que $d(a,x) \geqslant r$. Nous avons donc $d(a,x) = r$. △

Remarque 8.77.
Il serait toutefois faux de croire que $\partial A = \partial \bar{A}$ pour toute partie A de $\mathbb{R}^n$. En effet si $A = \mathbb{R}\setminus\{0\}$ nous avons $\partial A = \{0\}$ et $\bar{A} = \mathbb{R}$, donc $\partial \bar{A} = \varnothing$.

8.5.3 Point isolé, point d'accumulation

LEMooOMVNooVBlQMD

Lemme 8.78.
Soit un espace vectoriel normé V ainsi qu'une partie D de V.

(1) Un point $a \in D$ est dit isolé[26] *dans D relativement à V si il existe un $\varepsilon > 0$ tel que*

$$B(a,\varepsilon) \cap D = \{a\}. \tag{8.161}$$

(2) Un point $a \in V$ point d'accumulation[27] *de D si pour tout $\varepsilon > 0$,*

$$\Big(B(a,\varepsilon)\setminus\{a\}\Big) \cap D \neq \varnothing. \tag{8.162}$$

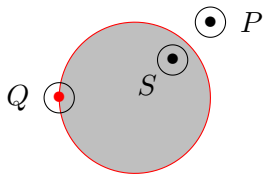


FIGURE 8.3: L'ensemble décrit par l'équation (8.163). Le point P est un point isolé de D, tandis que les points S et Q sont des points d'accumulation. LabelFigAccumulationIsole

Exemple 8.79.
Considérons la partie suivante de $\mathbb{R}^2$:

$$D = \{(x,y) \text{ tel que } x^2+y^2<1\} \cup \{(1,1)\}. \tag{8.163}$$ Eq2807BouleIso

Comme on peut le voir sur la figure 8.3, le point $P=(1,1)$ est un point isolé de D parce qu'on peut tracer une boule autour de P sans inclure d'autres points de D que P lui-même. Le point $Q=(-1,0)$ est un point d'accumulation de D parce que toute boule autour de Q contient des points de D.

Le point S, étant un point intérieur, est un point d'accumulation : toute boule autour de S intersecte D.

Notez cependant que le point Q lui-même n'est pas dans D parce que l'inégalité qui définit D est stricte. △

Remarque 8.80.
À propos de la position des points d'accumulation et des points isolés.

(1) Les points intérieurs sont tous des points d'accumulation.
(2) Les points isolés ne sont jamais intérieurs.
(3) Certains points d'accumulation ne font pas partie de l'ensemble. Par exemple le point 1 est un point d'accumulation de $E=]0,1[$.
(4) Les points de la frontière sont soit d'accumulation soit isolés.

Exemple 8.81.
Tous les points de $\mathbb{R}$ sont des points d'accumulation de $\mathbb{Q}$ parce que dans toute boule autour d'un réel, on peut trouver un nombre rationnel. △

Remarque 8.82.
L'ensemble des points d'accumulation d'un ensemble n'est pas exactement son adhérence. En effet, un point isolé dans A est dans l'adhérence de A, mais n'est pas un point d'accumulation de A.

8.5.4 Des exemples

ItemExoEVN3i

Exemple 8.83.
Nous considérons l'ensemble.

$$A_1 = \{(x,y) \in \mathbb{R}^2 \mid x^2-5x+6 < y \leqslant 2\}. \tag{8.164}$$

Si un point $(x,y) \in \mathbb{R}^2$ est tel que $x^2-5x+6<y$, alors dans une boule centrée en (x,y) (de rayon r_1), l'inégalité reste vraie (parce que la fonction $x^2-5x+6-y$ est une fonction continue). De la même manière, si nous avons $y<2$ en (x,y), alors nous avons encore l'inégalité dans une boule de rayon r_2. En prenant $r=\min\{r_1,r_2\}$, les deux inégalités restent vraies dans la boule de rayon r.

Donc les points (x,y) tels que $x^2-5x+6<y<2$ sont dans l'intérieur de A_1.

Pour les mêmes raisons, autour d'un point (x,y) tel que $x^2-5x+6>y$, nous pouvons trouver une boule dans laquelle l'inégalité reste stricte. Ces points ne sont donc pas dans l'adhérence de

26. Définition 7.35.
27. Définition 7.34.

A_1. Un point qui vérifie $x^2 - 5x + 6 = y = 2$ est par contre dans l'adhérence parce que dans toute boule, on pourra trouver un x tel que $x^2 - 5x + 6 < y$, et un y. L'adhérence est donc donnée par les inéquations

$$\bar{A}_1 \equiv x^2 - 5x + 6 \leqslant y \leqslant 2. \tag{8.165}$$

La frontière est donnée par les points de l'adhérence qui ne sont pas dans l'intérieur de A_1. Attention : **ne pas dire** que la frontière est alors donnée simplement en remplaçant les inégalités par des égalités : $\partial A_1 \equiv x^2 - 5x + 6 = y = 2$. Quel est cet ensemble ?

Trouver la frontière demande un peu plus de travail. Le point marqué sur la figure 8.4 est sur la frontière parce que toute boule intersecte l'intérieur et l'extérieur. Cela est dû au fait que, sur ce point, nous ayons $x^2 - 5x + 6 = y$ en même temps que $y < 2$. Donc si on prend une boule assez petite, on conserve $y < 2$, mais on obtient des points tels que $x^2 - 5x + 6 < y$.

En voyant le dessin, la chose à faire pour écrire la frontière est de trouver les deux points d'intersection entre la parabole et la droite horizontale. Ces points sont les points (x, y) qui satisfont au système

$$\begin{cases} x^2 - 5x + 6 = y & \text{(8.166a)} \\ y = 2. & \text{(8.166b)} \end{cases}$$

En substituant la seconde équation dans la première, il vient $x^2 - 5x + 6 = 2$, ce qui nous donne à résoudre le polynôme du second degré $x^2 - 5x + 4 = 0$. Les éventuelles solutions entières sont les diviseurs de 4. Par chance [28], on voit que $x = 1$ et $x = 4$ sont des solutions. Le théorème 3.130 nous assure qu'il n'y a pas d'autres racines. Les deux points d'intersection sont les points $P = (1, 2)$ et $Q = (4, 2)$. Les points de la frontière de A_1 sont donc donnés par

$$\begin{aligned} \partial A_1 = \{(x, y) \in \mathbb{R}^2 \text{ tels que } x^2 - 5x + 6 = y \text{ et } 1 \leqslant x \leqslant 4\} \\ \cup \{(x, y) \in \mathbb{R}^2 \text{ tels que } y = 2 \text{ et } 1 \leqslant x \leqslant 4\}. \end{aligned} \tag{8.167}$$

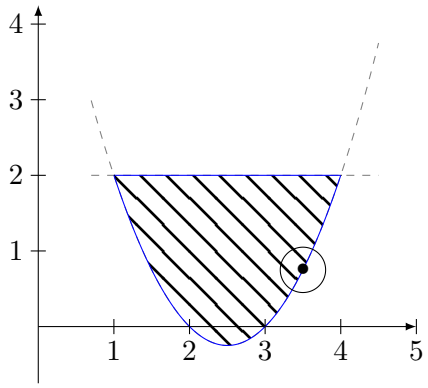


FIGURE 8.4: En hachuré : l'intérieur ; en trait plein : la frontière. L'adhérence est l'union des deux. Exemple8.83.

LabelFigAdhIntFr

Notez que les points de la parabole qui sont sur la frontière ne font pas partie de l'ensemble A_1 lui-même, tandis que ceux de la frontière qui sont sur la droite horizontale en font partie sauf $(4, 2)$ et $(1, 2)$.

L'intérieur de A_1 n'étant pas égal à A_1, cet ensemble n'est pas ouvert ; de la même manière, vu que $\bar{A}_1 \neq A_1$, l'ensemble n'est pas fermé. L'ensemble A_1 est par contre borné parce qu'il est contenu par exemple dans la boule de centre $(0, 0)$ et de rayon 5. Les points d'accumulation de A_1 sont les points de sa fermeture. △

ItemExoEVN3ii

Exemple 8.84.
Nous étudions

$$A_2 = \{(x, y) \in \mathbb{R}^2 \mid x + 1 < y < 2x\}. \tag{8.168}$$

28. Sinon, il aurait fallu utiliser la proposition 9.117.

Pour les mêmes raisons que dans l'exemple 8.83 l'intérieur est donné par

$$\text{Int}(A_2) \equiv x + 1 < y < 2x; \tag{8.169}$$

L'adhérence est donnée par

$$\overline{A_2} \equiv x + 1 \leqslant y \leqslant 2x, \tag{8.170}$$

Pour la frontière, les deux droites dont il est question dans la définition de A_2 (les droites $y = x+1$ et $y = 2x$) se coupent en $x = 1$ (refaire soi-même le dessin de la figure 8.5). Lorsque $x < 1$, les conditions $x + 1 < y$ et $y < 2x$ sont incompatibles : aucun point de A_2 n'est dans la partie $x < 1$ du plan. Lorsque $x > 1$, alors les points situés *entre* les deux droites font partie de A_2. La frontière est donc donnée par ces deux droites pour $x \geqslant 1$.

Étant donné que $\text{Int}(A_2) = A_2$, cet ensemble est ouvert (et donc pas fermé par la proposition 8.58(4)). Il n'est par contre pas borné parce qu'il contient des point (x, y) avec des x arbitrairement grands. △

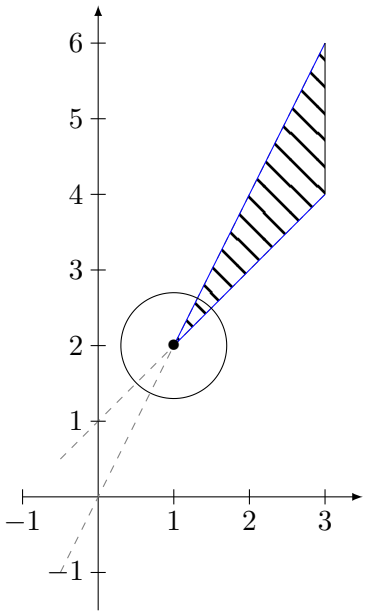


FIGURE 8.5: Notez que le point d'angle fait partie de la frontière, mais pas de l'ensemble. Exemple 8.84. LabelFigAdhIntFrDeux

Proposition 8.85 ([272]).
Tout partie de $\mathbb{R}$ *sans point d'accumulation est dénombrable.*

Démonstration. Soit une partie S de $\mathbb{R}$ ne contenant pas de points d'accumulation. Pour chaque $s \in S$, il existe un $\epsilon_s > 0$ tel que

$$B(s, \epsilon_s) \cap S = \{s\}. \tag{8.171}$$

Rien ne garantit cependant que $B(s, \epsilon_s) \cap B(t, \epsilon_t) = \varnothing$, alors que nous en aurons besoin pour la suite.

Le nombre

$$\inf\{|s - v| \text{ tel que } v \in S\backslash\{s\}\} \tag{8.172}$$

est au moins égal à ϵ_s et est don strictement positif. Nous posons

$$r_s = \frac{\inf\{|s - v| \text{ tel que } v \in S\backslash\{s\}\}}{4}, \tag{8.173}$$

et

$$I_s = B(s, r_s). \tag{8.174}$$

Nous avons maintenant $I_s \cap I_t = \varnothing$. Soit en effet $u \in I_s \cap I_t$. Alors

$$\begin{aligned} |s-t| &\leqslant |s-u| + |u-t| && (8.175a)\\ &\leqslant r_s + t_t && (8.175b)\\ &\leqslant 2r_s && (8.175c) \end{aligned}$$

où nous avons supposé $r_s \geqslant r_t$. Si ce n'est pas le cas, changer s et t dans ce qui suit ; les deux point ont des rôles symétrique. Nous avons donc

$$|s-t| \leqslant 2r_s = \frac{\inf\{|s-v| \text{ tel que } v \in S\backslash\{s\}\}}{2} \leqslant \frac{|s-t|}{2}. \tag{8.176}$$

Donc $|s-t| = 0$ et $s = t$.

Cela pour dire que I_s ne possède d'intersection avec I_t que si $s = t$.

Nous définissons alors une application

$$\begin{aligned} \varphi\colon S &\to \mathbb{Q}\\ s &\mapsto q_s \end{aligned} \tag{8.177}$$

où q_s est un choix de rationnel dans I_s. C'est le lemme 1.441 qui nous permet de choisir un tel rationnel.

La construction des intervalles I_s garantit que φ est une injection. Le fait qu'il existe une injection de $\mathbb{Q}$ vers S et le fait que $\mathbb{Q}$ est dénombrable impliquent que S est au plus dénombrable. □

8.6 Valeur propre et vecteur propre

8.6.1 Généralités

Nous savons qu'une application *linéaire* $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ est complètement définie par la donnée de son action sur les trois vecteurs de base, c'est-à-dire par la donnée de

$$Ae_1, \quad Ae_2 \quad \text{et} \quad Ae_3. \tag{8.178}$$

La matrice d'une application A se forme en mettant simplement les vecteurs Ae_1, Ae_2 et Ae_3 en colonne. Donc la matrice

$$A = \begin{pmatrix} 3 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0 \end{pmatrix} \tag{8.179}$$ EqExempleALin

signifie que l'application linéaire A envoie le vecteur e_1 sur $\begin{pmatrix} 3\\0\\0 \end{pmatrix}$, le vecteur e_2 sur $\begin{pmatrix} 0\\0\\1 \end{pmatrix}$ et le vecteur e_3 sur $\begin{pmatrix} 0\\1\\0 \end{pmatrix}$. Pour savoir comment A agit sur n'importe quel vecteur, on applique la règle de produit vecteur×matrice :

$$\begin{pmatrix} 1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x\\y\\z \end{pmatrix} = \begin{pmatrix} x+2y+3z\\ 4x+5y+6z\\ 7x+8y+9z \end{pmatrix}. \tag{8.180}$$

Une chose intéressante est de savoir quelles sont les directions invariantes de la transformation linéaire. Par exemple, on peut lire sur la matrice (8.179) que la direction $\begin{pmatrix} 1\\0\\0 \end{pmatrix}$ est invariante : elle est simplement multipliée par 3. Dans cette direction, la transformation est juste une dilatation.

Afin de savoir si v est un vecteur d'une direction conservée, il faut voir si il existe un nombre λ tel que $Av = \lambda v$, c'est-à-dire voir si v est simplement dilaté.

L'équation $Av = \lambda v$ se récrit $(A - \lambda\mathbb{1})v = 0$, c'est-à-dire qu'il faut résoudre l'équation

$$(A - \lambda\mathbb{1}) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{8.181}$$

Nous savons qu'une telle équation ne peut avoir de solutions que si $\det(A - \lambda\mathbb{1}) = 0$. La première étape est donc de trouver les λ qui vérifient cette condition.

8.6.2 Dans le vif du sujet

DefooMMKZooVcskCc

Définition 8.86.
*Soit un $\mathbb{K}$-espace vectoriel E et un endomorphisme $T : E \to E$. Un **vecteur propre** de T est un vecteur $v \neq 0$ tel que $T(v) = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Dans ce cas, λ est la **valeur propre** de v.*

*L'**espace propre** de T pour la valeur λ*[29] *est l'ensemble des vecteurs propres de T pour la valeur propre λ, et le vecteur nul.*

Définition 8.87.
*L'ensemble des valeurs propres de l'endomorphisme T est son **spectre** et est noté* $\mathrm{Spec}(T)$.

Remarque 8.88.
Le nombre zéro peut être une valeur propre ; c'est le vecteur zéro qui ne peut pas être vecteur propre. La matrice nulle est une matrice diagonalisable.

LEMooWHWUooFFXlzT

Lemme 8.89 ([273])**.**
Le spectre d'une matrice est égal au spectre de sa transposée : $\mathrm{Spec}(T) = \mathrm{Spec}(T^t)$.

Démonstration. Nous savons que T est inversible si et seulement si T^t l'est parce que si $TS = \mathbb{1}$ alors $S^tT^t = \mathbb{1}$. Nous avons équivalence entre les énoncés suivants :

— λ est une valeur propre de T
— il existe v tel que $(T - \lambda\mathbb{1})v = 0$
— $T - \lambda\mathbb{1}$ n'est pas inversible
— $(T^t - \lambda\mathbb{1})$ n'est pas inversible
— il existe w tel que $T^tw = \lambda w$.
— λ est valeur propre de T^t.

□

LemjcztYH

Lemme 8.90.
Soient un espace vectoriel E, un endomorphisme $T \in \mathrm{End}(E)$, ainsi que ses sous-espaces propres $\{E_\lambda\}_{\lambda \in \mathrm{Spec}(T)}$. Toute somme finie de la forme

$$E_{\lambda_1} + \ldots + E_{\lambda_p} \tag{8.182}$$

est directe[30].

Démonstration. Nous utilisons le lemme 4.141. Soient $v_i \in E_{\lambda_i}$ un choix de vecteurs tels que

$$\sum_{i=1}^{p} v_i = 0. \tag{8.183}$$

EQooROAXooFpgxxF

29. Nous laissons au lecteur le soin de vérifier que c'est bien un sous-espace vectoriel de E.
30. Définition 4.139.

Soit un entier j_0 entre 1 et p. Nous allons montrer que $v_{j_0} = 0$. Pour cela nous remarquons d'abord que, pour tout $i \neq j_0$,

$$\prod_{k\neq j_0} (\lambda_i - \lambda_k) = 0. \tag{8.184}$$

EQooGVPYooXRPEVU

Nous appliquons l'opérateur $\prod_{k\neq j_0}(T - \lambda_k \mathbb{1})$ à l'égalité (8.183) :

$$0 = \sum_{i=1}^{p} \prod_{k\neq j_0} (T - \lambda_k)v_i \tag{8.185a}$$

$$= \sum_{i=1}^{p} \prod_{k\neq j_0} (\lambda_i - \lambda_k)v_i \tag{8.185b}$$

SUBEQooHHGJooTvCcDb

$$= \prod_{k\neq j_0} (\lambda_{j_0} - \lambda_k)v_{j_0}. \tag{8.185c}$$

SUBEQooSHQYooNLVoVZ

Justifications.

— Pour (8.185b). Pour chaque k et i nous avons $(T - \lambda_k)v_i = Tv_i - \lambda_k v_i = \lambda_i v_i - \lambda_k v_i$ parce que v_i est un vecteur propre de T pour la valeur propre λ_i.

— Pour (8.185c). Dans la somme, seul le terme $i = j_0$ est non nul, à cause de (8.184).

Donc $v_{j_0} = 0$ parce que le produit $\prod_{k\neq j_0}(\lambda_{j_0} - \lambda_k)$, lui, est non nul. □

8.7 Polynômes d'endomorphismes

SECooUEQVooLBrRiE

Soit A un anneau commutatif et $\mathbb{K}$, un corps commutatif. L'injection canonique $A \to A[X]$ se prolonge en une injection

$$\mathbb{M}(A) \to \mathbb{M}\big(A[X]\big). \tag{8.186}$$

8.7.1 Polynômes d'endomorphismes

Soit $u \in \mathrm{End}(E)$ où E est un $\mathbb{K}$-espace vectoriel. Nous considérons l'application

$$\begin{aligned} \varphi_u \colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(u). \end{aligned} \tag{8.187}$$

EqOVKooeMJuv

L'image de φ_u est un sous-espace vectoriel. En effet si $A = \varphi_u(P)$ et $B = \varphi_u(Q)$, alors $A + B = \varphi_u(P + Q)$ et $\lambda A = (\lambda P)(u)$. En particulier c'est un espace fermé.

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E et P, un polynôme. Nous disons que P est un polynôme **annulateur** de u si $P(u) = 0$ en tant qu'endomorphisme de E.

LemQWvhYb

Lemme 8.91.
Si P et Q sont des polynômes dans $\mathbb{K}[X]$ et si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, nous avons

$$(PQ)(u) = P(u) \circ Q(u). \tag{8.188}$$

Démonstration. Si $P = \sum_i a_i X^i$ et $Q = \sum_j b_j X^j$, alors le coefficient de X^k dans PQ est

$$\sum_l a_l b_{k-l}. \tag{8.189}$$

EqCoefGPyVcv

Par conséquent $(PQ)(u)$ contient $\sum_l a_l b_{k-l} u^k$. Par ailleurs $P(u) \circ Q(u)$ est donné par

$$\sum_i a_i u^i \left(\sum_j b_j u^j \right)(x) = \sum_{ij} a_i b_j u^{i+j}(x). \tag{8.190}$$

Le coefficient du terme en u^k est bien le même que celui donné par (8.189). □

ThoDecompNoyayzzMWod

Théorème 8.92 (Décomposition des noyaux ou lemme des noyaux[274]).
Soit u un endomorphisme du $\mathbb{K}$-espace[31] *vectoriel E. Soit $P \in \mathbb{K}[X]$ un polynôme tel que $P(u) = 0$. Nous supposons que P s'écrive comme le produit $P = P_1 \dots P_n$ de polynômes deux à deux étrangers*[32]*. Alors*

(1)
$$E = \ker P_1(u) \oplus \dots \oplus \ker P_n(u). \tag{8.191}$$

(2) En posant $Q_i = \prod_{j \neq i} P_j$, il existe des polynômes R_i tels que $R_1Q_1 + \dots + R_nQ_n = 1$
ITEMooMOBPooPVXuXj

(3) Les projecteurs sur $\ker\big(P_j(u)\big)$ sont donnés par
$$\texttt{proj}_{\ker(P_i(i))} = (R_iQ_i)(u). \tag{8.192}$$

Démonstration. Dans ce qui suit, nous allons beaucoup utiliser le fait que $\mathbb{K}[X]$ soit commutatif (lemme 1.364). Nous posons
$$Q_i = \prod_{j \neq i} P_j. \tag{8.193}$$

(1) **Utilisation de Bézout**
Par le lemme 6.48 ces polynômes sont étrangers entre eux et le théorème de Bézout (théorème 6.47) donne l'existence de polynômes R_i tels que
$$R_1Q_1 + \dots + R_nQ_n = 1. \tag{8.194}$$
EQooMMCVooRzlXpA

(2) **Une première somme, pas directe** Si nous appliquons cette égalité à u et ensuite à $x \in E$ nous trouvons
$$\sum_{i=1}^{n} (R_iQ_i)(u)(x) = x, \tag{8.195}$$
EqqVcpUy
et en particulier si nous posons $E_i = \text{Image}\big(R_iQ_i(u)\big)$ nous avons
$$E = \sum_{i=1}^{n} E_i. \tag{8.196}$$
Cette dernière somme n'est éventuellement pas une somme directe.

(3) **Q_iQ_j est multiple de P** Si $i \neq j$, en utilisant la commutativité de $\mathbb{K}[X]$,
$$Q_iQ_j = \Big(\prod_{k\neq i} P_k\Big)\Big(\prod_{l\neq j} P_l\Big) = \Big(\prod_{\substack{k\neq i\\k\neq j}} P_k\Big)P_j\Big(\prod_{l\neq j} P_l\Big) = \prod_{\substack{k\neq i\\k\neq j}} P_k \prod_k P_l = S_{ij}P, \tag{8.197}$$
où S_{ij} est un polynôme. Nous voyons que Q_iQ_j est multiple de P.

(4) **Une somme directe** Toujours avec $i \neq j$, en utilisant le lemme 8.91,
$$(R_iQ_i)(u) \circ (R_jQ_j)(u) = \big(R_iQ_iR_jQ_j\big)(u) = \big(R_iR_j \underbrace{Q_iQ_j}_{=S_{ij}P}\big)(u) = (R_iR_jS_{ij})(u) \circ P(u) = 0 \tag{8.198}$$
Nous pouvons voir E comme un $\mathbb{K}$-module et appliquer le théorème 1.344. Les opérateurs $R_iQ_i(u)$ ont l'identité comme somme et sont orthogonaux, et nous avons donc la décomposition en somme directe :
$$E = \bigoplus_{i=1}^{n} R_iQ_i(u)E. \tag{8.199}$$
EQooJPQLooQZepwZ

31. Le corps $\mathbb{K}$ est commutatif comme tous les corps dans le Frido.
32. Définition 3.111.

(5) $R_iQ_i(u)E \subset \ker P_i(u)$ Attention : utilisation massive du lemme 8.91. Un élément de $R_iQ_i(u)E$ est de la forme $(R_iQ_i)(u)x$ avec $x \in E$. Nous appliquons l'endomorphisme $P_i(u)$ à cet élément, et nous vérifions que nous obtenons zéro :

$$\begin{aligned} P_i(u)\big((R_iQ_i)(u)x\big) &= (P_iR_iQ_i)(u)x && (8.200a)\\ &= (R_i \underbrace{P_iQ_i}_{=P})(u)x && (8.200b)\\ &= (R_iP)(u)x && (8.200c)\\ &= \big(R_i(u) \circ \underbrace{P(u)}_{=0}\big)x && (8.200d)\\ &= 0. && (8.200e) \end{aligned}$$

Par conséquent $\text{Image}(R_iQ_i(u)) \subset \ker P_i(u)$.

(6) **Et la somme qu'il nous fallait** Le fait que la somme (8.199) soit directe n'est en fait pas crucial. En effet, vu que chacun des termes est inclus dans $\ker P_i(u)$, nous avons la somme (pas directe à priori)

$$E = \sum_{i=1}^{n} R_iQ_i(u)E \subset \sum_{i=1}^{n} \ker P_i(u). \tag{8.201}$$

Mais cette fois, nous prouvons qu'elle est directe en utilisant la caractérisation du lemme 4.141(4). Supposons que, pour un certain k,

$$x \in \ker P_k(u) \cap \big(\sum_{j\neq k} \ker P_j(u)\big). \tag{8.202}$$

Nous allons montrer que $x = 0$.

(6a) $Q_i(u)x = 0$ **si** $i \neq k$ Si $i \neq k$, nous avons

$$Q_i(u)x = \Big(\prod_{\substack{j\neq i\\ j\neq k}} P_j\Big)P_k(u)x = 0 \tag{8.203}$$

parce que $x \in \ker P_k(u)$.

(6b) $Q_k(u)x = 0$ Nous savons qu'il existe $z_l \in \ker P_l(u)$ tel que $x = \sum_{l\neq k} z_l$. Nous avons alors

$$Q_k(u)x = \Big(\prod_{j\neq k} P_j\Big)\sum_{l\neq k} z_l = \sum_{l\neq k}\Big(\prod_{j\neq k} P_j(u)\Big)z_l = 0 \tag{8.204}$$

parce que parmi les $P_j(u)$ $(j \neq k)$, il y a $P_l(u)$ qui annule z_l.

(6c) **Et finalement** Nous avons prouvé que $Q_i(u)x = 0$ pour tout i. La formule de Bézout (8.194) donne alors

$$\sum_i R_i \subset Q_i(u)x_{=0} = x \tag{8.205}$$

et donc $x = 0$.

(7) **Les projecteurs**

□

8.93.
Ce résultat est utilisé pour prouver que toute représentation est décomposable en représentations irréductibles, proposition 16.10 ainsi que pour le théorème 8.219 qui dit que si le polynôme minimal d'un endomorphisme est scindé à racine simple alors il est diagonalisable.

CorKiSCkC

Corolaire 8.94.
Soit E, un $\mathbb{K}$-espace vectoriel de dimension finie et f, un endomorphisme semi-simple dont la décomposition du polynôme minimal μ_f en facteurs irréductibles sur $\mathbb{K}[X]$ est $\mu_f = M_1^{\alpha_1} \cdots M_r^{\alpha_r}$. Si F est un sous-espace stable par f, alors

$$F = \bigoplus_{i=1}^{r} \ker M_i^{\alpha_i}(f) \cap F \tag{8.206}$$

Démonstration. Nous posons $E_i = \ker M_i^{\alpha_i}(f)$ et $F_i = E_i \cap F$. Les polynômes $M_i^{\alpha_i}$ sont deux à deux étrangers et $\mu_f(f) = 0$, donc le lemme des noyaux (8.92) s'applique et

$$E = E_1 \oplus \ldots \oplus E_r. \tag{8.207}$$

Nous pouvons décomposer $x \in F$ en termes de cette somme :

EqbBbrdi

$$x = x_1 + \cdots + x_r \tag{8.208}$$

avec $x_i \in E_i$. Toujours selon le lemme des noyaux, les projections sur les espaces E_i sont des polynômes en f. Par conséquent F est stable sous toutes ces projections $\texttt{proj}_i : E \to E_i$, et en appliquant $\texttt{proj}_i$ à (8.208), $\texttt{proj}_i(x) = x_i$. Puisque $x \in F$, le membre de gauche est encore dans F et $x_i \in E_i \cap F$. Nous avons donc

$$F \subset \bigoplus_{i=1}^{r} F_i. \tag{8.209}$$

L'inclusion inverse est immédiate parce que $F_i \subset F$ pour chaque i. □

LemVISooHxMdbr

Lemme 8.95.
Si x est un vecteur propre de valeur propre λ pour l'endomorphisme u et si P est un polynôme, alors x est vecteur propre de $P(u)$ pour la valeur propre $P(\lambda)$.

Démonstration. C'est un simple calcul de $P(u)x$ en ayant noté [33] $P(X) = \sum_{k=0}^{n} c_k X^n$:

$$P(u)x = \sum_{k=0}^{n} c_k u^k(x) = \sum_{k=0}^{n} c_k \lambda^k x = P(\lambda)x. \tag{8.210}$$

□

8.7.2 Polynôme minimal et minimal ponctuel

Nous avons déjà vu la définition de polynôme minimal en 6.64. Le lemme suivant permet de parler de polynôme minimal d'endomorphisme.

LEMooEYPSooLCoPlY

Lemme 8.96 ([1])**.**
Si E est un $\mathbb{K}$-espace vectoriel, l'ensemble $\mathrm{End}(E)$ des endomorphismes de E est une extension du corps $\mathbb{K}$.

LEMooQJQGooRcAxmJ

Lemme 8.97.
Soit un endomorphisme $f : E \to E$ d'un $\mathbb{K}$-espace vectoriel de dimension finie. Il existe un unique polynôme annulateur unitaire de degré minimum [34].

Tout endomorphisme de $\mathbb{K}$-espace vectoriel de dimension finie possède un polynôme minimal [35].

33. En complète violation de ce qu'on disait dans 1.365.
34. Degré minimum au sens où il existe peut-être d'autres polynômes annulateurs, mais ils seront de degré plus élevé.
35. Définition 6.64.

Démonstration. Pour l'unicité, soient P et Q deux polynômes annulateurs de f de même degré minimum N et ayant tous deux 1 comme coefficient de x^N. Alors $P-Q$ est de degré $N-1$ tout en étant encore annulateur. Vu que nous avions dit que N était le degré minimum, le seul polynôme annulateur de degré $N-1$ est le polynôme nul. Donc $P-Q=0$.

Pour l'existence, les endomorphismes Id, f, f^2, ...ne peuvent pas être tous linéairement indépendants parce que la dimension de $\mathrm{End}(E)$ est finie. Il existe donc un nombre N et des coefficients a_k tels que $\sum_{k=0}^{N} a_k f^k = 0$. Le polynôme $P(X) = \sum_{k=0}^{N} a_k X^k$ est donc annulateur de f.

Une autre façon de le dire est que l'application linéaire $\varphi\colon \mathbb{K}[X] \to \mathrm{End}(E)$ donnée par $\varphi(P) = P(f)$ est un endomorphisme d'un espace vectoriel de dimension infinie vers un espace vectoriel de dimension finie. Il ne peut donc pas être injectif et possède donc un noyau non réduit à zéro.

L'existence d'un polynôme minimal est maintenant seulement dû au fait que, avec les notations de la définition 6.64, l'idéal I_f n'est pas réduit à $\{0\}$. □

Remarque 8.98.
La preuve donnée ci-dessus montre que $\deg(\mu) \leqslant \dim(E)^2$. Comme conséquence du théorème de Cayley-Hamilton 8.120 nous verrons qu'en réalité le degré du polynôme minimal est majoré par la dimension de l'espace.

PROPooZCUSooLUUrxi

Proposition 8.99 (Exemple en dimension infinie[1])**.**
L'endomorphisme de dérivation sur l'espace des fonctions dérivables $\mathbb{R} \to \mathbb{R}$ n'a pas de polynôme minimal.

Dans la suite, l'endomorphisme f du $\mathbb{K}$-espace vectoriel E de dimension n est fixé. Pour $x \in E$ nous notons

$$E_x = \{P(f)x \text{ tel que } P \in \mathbb{K}[X]\}. \tag{8.211}$$

EqooOAYDooEpZELo

Nous considérons le morphisme d'algèbres

$$\begin{aligned} \varphi\colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(f) \end{aligned} \tag{8.212}$$

et si $x \in E$ est donné nous considérons le morphisme de $\mathbb{K}$-espaces vectoriels

$$\begin{aligned} \varphi_x\colon \mathbb{K}[X] &\to E \\ P &\mapsto P(f)x. \end{aligned} \tag{8.213}$$

Les noyaux de ces applications sont des idéaux, entre autres par le lemme 8.91. Ils ont donc un unique générateur unitaire (chacun) par le théorème 6.43. En termes de vocabulaire, l'ensemble

$$\ker(\varphi) = \{P \in \mathbb{K}[X] \text{ tel que } P(f) = 0\} \tag{8.214}$$

est l'**idéal annulateur** de f et un polynôme P tel que $P(f) = 0$ est un polynôme annulateur de f.

DEFooUICRooBGYhqQ

Proposition-Définition 8.100.
La partie $\ker(\varphi_x)$ est un idéal de $\mathbb{K}[X]$ qui possède un unique générateur unitaire.

*Le générateur unitaire de $\ker(\varphi_x)$ est le **polynôme minimal ponctuel** de f en x. Il sera noté $\mu_{f,x}$ ou μ_x lorsque la dépendance en f est claire dans le contexte.*

Nous notons μ le générateur unitaire du noyau de φ et μ_x celui de φ_x. Puisque $\mu \in \ker(\varphi_x)$ pour tout x nous avons $\mu_x \mid \mu$ pour tout x.

ExooDTUJooIMqSKn

Exemple 8.101 (Pas en dimension infinie)**.**
En dimension infinie, il n'y a pas toujours de polynôme annulateur. Si E est un espace vectoriel de dimension infinie ayant une base dénombrable $\{e_i\}_{i\in\mathbb{N}}$ alors l'opérateur donné par $f(e_i) = e_{i+1}$ n'a pas de polynôme annulateur. Même pas ponctuel en quel que point que ce soit.

De même l'opérateur donné par $g(e_1) = 0$ et $g(e_i) = e_{i-1}$ si $i \neq 1$ n'a pas de polynôme annulateur, mais il a un polynôme annulateur ponctuel évident en $x = e_1$. L'exemple 15.92 donnera un habillage à peine subtil à cet exemple. △

PropAnnncEcCxj

Proposition 8.102.
Si P est un polynôme tel que $P(f) = 0$, alors le polynôme minimal μ_f divise P. Autrement dit, le polynôme minimal engendre l'idéal des polynômes annulateurs.

Démonstration. L'ensemble $\ker(\varphi) = \{Q \in \mathbb{K}[X] \text{ tel que } Q(f) = 0\}$ est un idéal par le lemme 8.91. Le polynôme minimal de f est un élément de degré plus bas dans I et par conséquent $I = (\mu_f)$ par le théorème 6.43. Nous concluons que μ_f divise tous les éléments de I. □

La proposition suivante permet de caractériser le polynôme minimal.

PROPooVUJPooMzxzjE

Proposition 8.103 ([130]).
Soit une application linéaire f sur un $\mathbb{K}$-espace vectoriel. Il existe un unique polynôme unitaire[36] *$P \in \mathbb{K}[X]$ tel que*

(1) $P(f) = 0$;

(2) l'application

$$\begin{aligned}\varphi\colon \frac{\mathbb{K}[X]}{(P)} &\to \mathrm{End}(E)\\ \bar{Q} &\mapsto Q(f)\end{aligned} \tag{8.215}$$

EQooIBMDooVTaEbf

est injective.

Démonstration. En ce qui concerne l'existence, il existe le polynôme minimal de f qui satisfait les conditions. Pour l'unicité nous y travaillons maintenant.

Supposons que l'application (8.215) soit injective. Alors pour tout $Q \in \mathbb{K}[X]$ tel que $Q(f) = 0$ nous avons $\bar{Q} = 0$, c'est-à-dire $Q = PR$ pour un certain $R \in \mathbb{K}[X]$. Autrement dit : P est un générateur unitaire de l'idéal annulateur de f. Le théorème 6.43(3) nous dit alors que $P = \mu$ parce que μ est également générateur unitaire. □

LemSYsJJj

Lemme 8.104 ([275]).
Soit $f\colon E \to E$ un endomorphisme de l'espace vectoriel E. Il existe un élément $x \in E$ tel que $\mu_{f,x} = \mu_f$.

Démonstration. Soit une décomposition en irréductibles du polynôme minimal $\mu = P_1^{\alpha_1} \dots P_r^{\alpha_r}$. Nous notons $E_i = \ker\big(P_i^{\alpha_i}(f)\big)$. Les polynômes P_i sont étrangers deux à deux (un diviseur commun aurait a fortiori été un diviseur et aurait contredit l'irréductibilité). Le lemme des noyaux 8.92 nous donne la somme directe

$$E = \bigoplus_{i=1}^{r} \ker\big(P_i^{\alpha_i}(f)\big). \tag{8.216}$$

Si $x_i \in E_i$ alors μ_{x_i} est une puissance de P_i. En effet $\mu_{x_i} \mid \mu$ et est donc un produit des puissances des P_j. Or si $(QP_j)(f)x_i = 0$ alors $(P_jQ)(f)x_i = 0$, ce qui donne $Q(f)x_i \in E_j \cap E_i = \{0\}$ si $j \neq i$. Donc μ_{x_i} n'est pas de la forme QP_j pour $j \neq i$. Nous en déduisons que μ_{x_i} est une puissance de P_i dès que $x_i \in E_i$. Nous choisissons $x_i \in E_i$ tel que $\mu_{x_i} = P_i^{\alpha_i}$.

Nous posons enfin $a = x_1 + \dots + x_r$; par définition du polynôme annulateur μ_a, nous avons

$$0 = \mu_a(f)a = \mu_a(f)x_1 + \dots + \mu_a(f)x_r. \tag{8.217}$$

EqooVIGGooSfuvwB

Mais $\mu_a(f)x_i \in E_i$, et la somme des E_i est directe, donc l'annulation de la somme (8.217) implique l'annulation de chacun des termes : $\mu_a(f)x_i = 0$ pour tout i. Cela prouve que $\mu_{x_i} \mid \mu_a$. Mais comme les μ_{x_i} sont premiers deux à deux (parce que ce sont les $P_i^{\alpha_i}$), nous concluons que le produit divise encore μ_a :

$$\prod_{i=1}^{r} \mu_{x_i} \mid \mu_a, \tag{8.218}$$

c'est-à-dire $\mu \mid \mu_a$. Comme nous avons aussi $\mu_a \mid \mu$, nous déduisons $\mu_a = \mu$. □

36. À mon avis, « unitaire » manque dans [130].

DEFooFEIFooNSGhQE

Définition 8.105 (Matrices, endomorphismes et vecteurs cycliques)**.**
Une matrice est ***cyclique*** *si elle est semblable à une matrice compagnon. Un endomorphisme $f: E \to E$ est* ***cyclique*** *si il existe un vecteur $x \in E$ tel que $\{f^k(x)\}_{k=0,\dots,n-1}$ est une base de E. Un vecteur ayant cette propriété est un* ***vecteur cyclique*** *pour f.*

LemAGZNNa

Lemme 8.106.
Soit E un espace vectoriel de dimension finie et un endomorphisme cyclique[37] *f de E. Soit un vecteur cyclique v de f, alors le polynôme minimal de f est égal au polynôme minimal de f au point v : $\mu_f = \mu_{f,v}$.*

Démonstration. Montrons que $\mu_{f,v}$ est un polynôme annulateur de f, ce qui prouvera que μ_f divise $\mu_{f,v}$ par la proposition 8.102. Étant donné que v est cyclique, tout élément de E s'écrit sous la forme $x = Q(f)v$. Prenons un polynôme P annulateur de f en v : $P(f)v = 0$. Nous montrons que P est alors un polynôme annulateur de f. En effet, nous avons

$$P(f)x = \big(P(f) \circ Q(f)\big)v = \big(Q(f) \circ P(f)\big)v = 0 \tag{8.219}$$

où nous avons utilisé le lemme 8.91. □

LEMooOWDAooWPbPda

Lemme 8.107 ([275])**.**
Soit $a \in E$ un vecteur cyclique pour f, tel que $\mu_a = \mu$. Alors E_a est un sous-espace stable par f pour lequel il existe un supplémentaire stable.

Démonstration. Soit $l = \deg(\mu) = \deg(\mu_a)$. L'espace E_a étant engendré par les $f^k(a)$ nous savons que $e_1 = a$, $e_2 = f(a)$,..., $e_l = f^{l-1}(a)$ forment une base de E_a. Nous pouvons la compléter en une base $\{e_1, \dots, e_n\}$ de E. Et nous posons[38]

$$\begin{aligned} G &= \{x \in E \text{ tel que } e_l^*\big(f^k(x)\big) = 0, \forall k \geqslant 0\} && \text{(8.220a)}\\ &= \bigcap_{k\geqslant 0} \ker\{e_l^* \circ f^k\} && \text{(8.220b)}\\ &= \bigcap_{k=0}^{l-1} \ker(e_l^* \circ f^k). && \text{(8.220c)} \end{aligned}$$

La dernière égalité est due au fait que l soit le degré de μ. Du coup f^l est une combinaison linéaire des f^i avec $i \leqslant l-1$.

Nous avons $f(G) \subset G$ et de plus $E_a \cap G = \{0\}$ parce qu'un élément de E_a est une combinaison linéaire d'éléments de la forme $f^j(a)$ $(j \leqslant l)$. Après application de f^{l-j}, ces éléments obtiennent une composante $f^l(a) = e_l$. De plus G est un sous-espace vectoriel du fait que $e_l^* \circ f^i$ est une application linéaire.

Montrons enfin que $\dim(G) = n - l$. Pour cela nous remarquons que G est une intersection d'hyperplans, et nous montrons que les équations définissant ces hyperplans sont linéairement indépendantes. Soit donc

$$\sum_{j=0}^{l-1} \lambda_j \big(e_l^* \circ f^j\big) = 0 \tag{8.221}$$

EqooOHESooRtBUfc

et montrons que $\lambda_j = 0$ pour tout j est l'unique solution. Soit $x \in E$ et appliquons l'opération (8.221) au vecteur $f^i(x)$; le résultat est zéro :

$$0 = \sum_{j=0}^{l-1} \lambda_j (e_l^* \circ f^i \circ f^j) = (e_l^* \circ f^i)P(u) \tag{8.222}$$

37. Voir la définition 8.105.

38. ici, comme presque partout, e_l^* est le dual de e_l, c'est-à-dire l'application linéaire sur E donnée par $e_l^*(e_i) = \delta_{li}$, voir la définition 4.127.

où nous avons posé $P(X)=\sum_{j=0}^{l-1}\lambda_j X^j$. Appliquons cela à a : pour tout i nous avons

$$(e_l^*\circ f^i)\big(P(f)a\big)=0. \tag{8.223}$$

Mais par définition de E_a, l'élément $P(f)a$ est dans E_a. Nous en déduisons que

$$P(f)a\in G\cap E_a=\{0\}, \tag{8.224}$$

c'est-à-dire que P est un polynôme annulateur de a. Mais P est de degré $l-1$ alors que le polynôme minimal de a est de degré l. Par conséquent $P=0$ et $\lambda_j=0$ pour tout j. □

DEFooBOHVooSOopJN

Définition 8.108.
*L'endomorphisme f d'un espace vectoriel est **semi-simple** si tout sous-espace stable par f possède un supplémentaire stable.*

LemrFINYT

Lemme 8.109.
Si le polynôme minimal d'un endomorphisme est irréductible, alors cet endomorphisme est semi-simple[39].

Démonstration. Soit f, un endomorphisme dont le polynôme minimal est irréductible et F, un sous-espace stable par f. Nous devons en trouver un supplémentaire stable. Si $F=E$, il n'y a pas de problème. Sinon nous considérons $u_1\in E\backslash F$ et

$$E_{u_1}=\{P(f)u_1 \text{ tel que } P\in\mathbb{K}[X]\}, \tag{8.225}$$

qui est un espace stable par f.

Montrons que $E_{u_1}\cap F=\{0\}$. Pour cela nous étudions l'idéal

$$I_{u_1}=\{P\in\mathbb{K}[X] \text{ tel que } P(f)u_1=0\}. \tag{8.226}$$

C'est un idéal non réduit à $\{0\}$ parce que le polynôme minimal de f par exemple est dans I_{u_1}. Soit P_{u_1} un générateur unitaire de I_{u_1}. Étant donné que $\mu_f\in I_{u_1}$, nous avons P_{u_1} divise μ_f et donc, $P_{u_1}=\mu_f$, parce que μ_f est irréductible par hypothèse.

Soit $y\in E_{u_1}\cap F$. Par définition il existe $P\in\mathbb{K}[X]$ tel que $y=P(f)u_1$ et si $y\neq 0$, cela signifie que $P\notin I_{u_1}$, c'est-à-dire que P_{u_1} ne divise pas P. Étant donné que P_{u_1} est irréductible cela implique que P_{u_1} et P sont premiers entre eux (ils n'ont pas d'autre pgcd que 1).

Nous utilisons maintenant des coefficient de Bézout (théorème 6.47) $A,B\in\mathbb{K}[X]$ tels que

$$AP+BP_{u_1}=1. \tag{8.227}$$

Nous appliquons cette égalité à f et puis à u_1 :

$$u_1=A(f)\circ\underbrace{P(f)u_1}_{=y}+B(f)\circ\underbrace{P_{u_1}(f)u_1}_{=0}=A(f)y. \tag{8.228}$$

Mais $y\in F$, donc $A(f)y\in F$. Nous aurions donc $u_1\in F$, ce qui est impossible par choix. Nous savons maintenant que l'espace $E_{u_1}\oplus F$ est stable sous f. Si cet espace est E alors nous arrêtons. Sinon nous reprenons le raisonnement avec $E_{u_1}\oplus F$ en guise de F et en prenant $u_2\in E\backslash(E_{u_1}\oplus F)$. Étant donné que E est de dimension finie, ce procédé s'arrête à un certain moment et nous aurons

$$E=F\oplus E_{u_1}\oplus\ldots\oplus E_{u_k} \tag{8.229}$$

où chacun des E_{u_i} sont stables. □

ThoFgsxCE

Théorème 8.110.
Un endomorphisme est semi-simple si et seulement si son polynôme minimal est produit de polynômes irréductibles distincts deux à deux.

39. Définition 8.108.

Démonstration. Supposons que f soit semi-simple et que son polynôme minimal soit donné par $\mu_f = M_1^{\alpha_1} \dots M_r^{\alpha_r}$ où les M_i sont des polynômes irréductibles deux à deux distincts. Nous devons montrer que $\alpha_i = 1$ pour tout i. Soit i tel que $\alpha_i \geqslant 1$ et $N \in \mathbb{K}[X]$ tel que $\mu_f = M^2 N$ où l'on a noté $M = M_i$. Nous étudions l'espace

$$F = \ker M(f) \tag{8.230}$$

qui est stable par f, et qui possède donc un supplémentaire S également stable par f. Nous allons montrer que MN est un polynôme annulateur de f.

D'abord nous prenons $x \in S$. Étant donné que F est le noyau de $M(f)$,

$$M(f)\big(MN(f)x\big) = \mu_f(f)x = 0, \tag{8.231}$$

ce qui signifie que $MN(f)x \in F$. Mais puisque S est stable par f nous avons aussi $MN(f)x \in S$. Finalement $MN(f)x \in F \cap S = \{0\}$. Autrement dit, $MN(f)$ s'annule sur S.

Prenons maintenant $y \in F$. Nous avons

$$MN(f) = N(f)\big(M(f)y\big) = 0 \tag{8.232}$$

parce que $y \in F = \ker M(f)$.

Nous avons prouvé que $MN(f)$ s'annule partout et donc que $MN(f)$ est un polynôme annulateur de f, ce qui contredit la minimalité de $\mu_f = M^2 N$.

Nous passons au sens inverse. Soit $m_f = M_1 \dots M_r$ une décomposition du polynôme minimal de l'endomorphisme f en irréductibles distincts deux à deux. Soit F un sous-espace vectoriel stable par f. Nous notons

$$E_i = \ker(M_i(f)) \tag{8.233}$$

et $f_i = f|_{E_i}$. Par le lemme 8.94 nous avons

$$F = \bigoplus_{i=1}^{r} (F \cap E_i). \tag{8.234}$$

Les espaces E_i sont stables par f et étant donné que M_i est irréductible, il est le polynôme minimal de f_i. En effet, M_i est annulateur de f_i, ce qui montre que le polynôme minimal de f_i divise M_i. Mais M_i étant irréductible, M_i est le polynôme minimal. Étant donné que $\mu_{f_i} = M_i$, l'endomorphisme f_i est semi-simple par le lemme 8.109.

L'espace $F \cap E_i$ étant stable par l'endomorphisme semi-simple f_i, il possède un supplémentaire stable que nous notons S_i :

$$E_i = S_i \oplus (F \cap E_i). \tag{8.235}$$

Étant donné que sur chaque S_i nous avons $f|_{S_i} = f_i$, l'espace $S = S_1 \oplus \dots \oplus S_r$ est stable par f. Par conséquent, nous avons

$$\begin{aligned}
E &= E_1 \oplus \dots \oplus E_r && \text{(8.236a)}\\
&= \big(S_1 \oplus (F \cap E_1)\big) \oplus \dots \oplus \big(S_r \oplus (F \cap E_r)\big) && \text{(8.236b)}\\
&= \big(\bigoplus_{i=1}^{r} S_i\big) \oplus \big(\bigoplus_{i=1}^{r} F \cap E_i\big) && \text{(8.236c)}\\
&= S \oplus F, && \text{(8.236d)}
\end{aligned}$$

ce qui montre que F a bien un supplémentaire stable par f et donc que f est semi-simple. □

Exemple 8.111 (L'espace engendré par $\mathbb{1}$, A, A^2,...)**.**
Soit A une matrice, et

$$E = \text{Span}\{A^k \text{ tel que } k \in \mathbb{N}\}. \tag{8.237}$$

Nous montrons que $\dim(E)$ est le degré du polynôme minimal de A.

D'abord l'idéal annulateur de A est engendré par le polynôme minimal[40] que nous notons $\mu = \sum_{k=0}^{p} a_k X^k$. La partie $\{\mathbb{1}, \dots, A^{p-1}\}$ est libre parce qu'une combinaison linéaire nulle de ces éléments serait un polynôme annulateur en A de degré plus petit que p. Donc $\dim(E) \geqslant p$.

La partie $\{\mathbb{1}, A, \dots, A^p\}$ est liée à cause du polynôme minimal. Isoler A^p dans $\mu(A) = 0$ donne un polynôme f de degré $p-1$ tel que $A^p = f(A)$.

Nous allons montrer à présent que la famille $\{\mathbb{1}, A, \dots, A^{p-1}\}$ est génératrice (alors $\dim(E) \leqslant p$). Soit un entier $q \geqslant p$ et de division euclidienne[41] $np + r = q$ avec $r < p$. Nous avons $A^q = A^{np}A^r$. D'une part

$$A^{np} = (A^p)^n = f(A)^n \tag{8.238}$$

est de degré $n(p-1)$. Par conséquent

$$A^q = f(A)^n A^r \tag{8.239}$$

qui est de degré $n(p-1)+r = q-n$. Autrement dit il existe un polynôme g_1 de degré $q-n$ tel que $A^q = g_1(A)$. Si $q-n > p-1$ alors nous pouvons recommencer et obtenir un polynôme g_2 de degré strictement inférieur à celui de g_1 tel que $A^q = g_2(A)$. Au bout du compte, il existe un polynôme g de degré au maximum $p-1$ tel que $A^q = g(A)$. Cela prouve que la partie $\{\mathbb{1}, A, \dots, A^{p-1}\}$ est génératrice de E.

La dimension de E est donc p, le degré du polynôme minimal. △

PropooCFZDooROVlaA

Proposition 8.112.
Soit f un endomorphisme d'un espace vectoriel de dimension finie. Nous avons l'isomorphisme d'espace vectoriel

$$\mathbb{K}[f] \simeq \frac{\mathbb{K}[X]}{(\mu_f)} \tag{8.240}$$

La dimension en est $\deg(\mu_f)$.

Démonstration. Notons avant de commencer que (μ) est l'idéal engendré par μ. Les classes dont il est question dans le quotient $\mathbb{K}[X]/(\mu)$ sont

$$\bar{P} = \{P + S\mu\}_{S \in \mathbb{K}[X]}. \tag{8.241}$$

Nous allons montrer que l'application suivante fournit l'isomorphisme :

$$\begin{aligned} \psi\colon \frac{\mathbb{K}[X]}{(\mu)} &\to \mathbb{K}[f] \\ \bar{P} &\mapsto P(f). \end{aligned} \tag{8.242}$$

(1) **ψ est bien définie** Si $Q \in \bar{P}$ alors $Q = P + S\mu$ pour un certain $S \in \mathbb{K}[X]$. Du coup nous avons

$$\psi(\bar{Q}) = P(f) + (S\mu)(f). \tag{8.243}$$

Mais $\mu(f) = 0$ donc le deuxième terme est nul. Donc $\psi(\bar{P})$ est bien défini.

(2) **Injectif** Si $\psi(\bar{P}) = 0$ nous avons $P(f) = 0$, ce qui signifie que $P = S\mu$ pour un polynôme S. Par conséquent $P \in (\mu)$ et donc $\bar{P} = 0$.

(3) **Surjectif** Soit $P \in \mathbb{K}[X]$. L'élément $P(f)$ de $\mathbb{K}[f]$ est dans l'image de ψ parce que c'est $\psi(\bar{P})$.

En ce qui concerne la dimension, le corolaire 6.44 en parle déjà : une base est donnée par les projections de $1, X, \dots, X^{\deg(\mu_f)-1}$. □

40. Proposition 8.102.
41. Théorème 1.222.

8.7.3 Polynôme caractéristique

DefOWQooXbybYD

Définition 8.113.
Soit un anneau commutatif A. Si $u \in \mathbb{M}(n,A)$, nous définissons le ***polynôme caractéristique de*** *u :*

$$\chi_u(X) = \det(u - X\mathbb{1}_n). \tag{8.244}$$

Eqksbdfu

Nous définissons de même le polynôme caractéristique d'un endomorphisme $u\colon E \to E$.

Remarque 8.114.
Quelques remarques à propos du signe [42].

- Certains auteurs définissent le polynôme caractéristique par $\det(X-u)$ au lieu de $\det(u-X)$.
- Wikipédia francophone[276] prend la définition $\det(X-u)$ (donc opposée de la notre). Allez lire la page de discussion.
- Sur les wikipédias en d'autres langues, ça varie.
- Un avantage de $\det(u-X)$ est que $\det(u) = \chi_u(0)$.
- Un avantage de $\det(X-u)$ est qu'il est unitaire.

LemooWCZMooZqyaHd

Lemme 8.115.
Le polynôme caractéristique χ_u est unitaire en dimension paire et a pour degré la dimension de l'espace vectoriel E.

ThoNhbrUL

Théorème 8.116.
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et un endomorphisme $u \in \mathrm{End}(E)$. Alors

(1) Le polynôme caractéristique divise $(\mu_u)^n$ dans $\mathbb{K}[X]$.
(2) Les polynômes caractéristiques et minimaux ont mêmes facteurs irréductibles dans $\mathbb{K}[X]$.
(3) Les polynômes caractéristiques et minimaux ont mêmes racines dans $\mathbb{K}[X]$.
(4) Le polynôme caractéristique est scindé si et seulement si le polynôme minimal est scindé.

ThoWDGooQUGSTL

Théorème 8.117.
Soit $u \in \mathrm{End}(E)$ et $\lambda \in \mathbb{K}$. Les conditions suivantes sont équivalentes

ItemeXHXhHi
(1) $\lambda \in \mathrm{Spec}(u)$
ItemeXHXhHii
(2) $\chi_u(\lambda) = 0$
ItemeXHXhHiii
(3) $\mu_u(\lambda) = 0$.

Démonstration. (1) ⇔ (2). Dire que λ est dans le spectre de u signifie que l'opérateur $u - \lambda\mathbb{1}$ n'est pas inversible, ce qui est équivalent à dire que $\det(u-\lambda\mathbb{1})$ est nul par la proposition 8.16(1) ou encore que λ est une racine du polynôme caractéristique de u.

(2) ⇔ (3). C'est une application directe du théorème 8.116 qui précise que le polynôme caractéristique a les mêmes racines dans $\mathbb{K}$ que le polynôme minimal. □

ExICOJcFp

Exemple 8.118.
Sur $\mathbb{R}^2$, nous considérons la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ qui a pour polynôme caractéristique [43] le polynôme $\chi_A = (X-1)^2$. Le nombre $\lambda = 1$ est une racine double de ce polynôme, et pourtant il n'y a qu'une seule dimension d'espace propre :

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.245}$$

entraine $x = 0$.

Ici la multiplicité algébrique est différente de la multiplicité géométrique. △

42. Attention : je crois qu'il y a des incohérences dans le Frido à propos de ce choix
43. Définition 8.113.

PropNrZGhT

Proposition 8.119 ([277]).
Soit f, un endomorphisme de E et $x \in E$. Alors

ItemfzKOCo

(1) L'espace $E_{f,x}$ est stable par f.

(2) L'espace $E_{f,x}$ est de dimension

$$p_{f,x} = \dim E_{f,x} = \deg(\mu_{f,x}) \tag{8.246}$$

où $\mu_{f,x}$ est le générateur unitaire de $I_{f,x}$.

ItemKHNExH

(3) Le polynôme caractéristique de $f|_{E_{f,x}}$ est $\mu_{f,x}$.

ItemHMviZw

(4) Nous avons

$$\chi_{f|_{E_{f,x}}}(f)x = \mu_{f,x}(f)x = 0. \tag{8.247}$$

Démonstration. Le fait que $E_{f,x}$ soit stable par f est classique. Le point (4) est une application du point (3). Les deux gros morceaux sont donc les points (2) et (3).

Étant donné que $\mu_{f,x}$ est de degré minimal dans $I_{f,x}$, l'ensemble

$$B = \{f^k(x) \text{ tel que } 0 \leqslant k \leqslant p_{f,x} - 1\} \tag{8.248}$$

est libre. En effet une combinaison nulle des vecteurs de B donnerait un polynôme en f de degré inférieur à $p_{f,x}$ annulant x. Nous écrivons

$$\mu_{f,x}(X) = X^{p_{f,x}} - \sum_{i=0}^{p_{f,x}-1} a_i X^i. \tag{8.249}$$

Étant donné que $\mu_{f,x}(f)x = 0$ et que la somme du membre de droite est dans $\mathrm{Span}(B)$, nous avons $f^{p_{f,x}}(x) \in \mathrm{Span}(B)$. Nous prouvons par récurrence que $f^{p_{f,x}+k}(x) \in \mathrm{Span}(B)$. En effet en appliquant f^k à l'égalité

$$0 = f^{p_{f,x}}(x) - \sum_{i=0}^{p_{f,x}-1} a_i f^i(x) \tag{8.250}$$

nous trouvons

$$f^{p_{f,x}+k}(x) = \sum_{i=0}^{p_{f,x}-1} a_i f^{i+k}(x), \tag{8.251}$$

alors que par hypothèse de récurrence le membre de droite est dans $\mathrm{Span}(B)$. L'ensemble B est alors générateur de $E_{f,x}$ et donc une base d'icelui. Nous avons donc bien $\dim(E_{f,x}) = p_{f,x}$.

Nous montrons maintenant que $\mu_{f,x}$ est annulateur de f au point x. Nous savons que

$$\mu_{f,x}(f)x = 0. \tag{8.252}$$

En y appliquant f^k et en profitant de la commutativité des polynômes sur les endomorphismes (proposition 8.91), nous avons

$$0 = f^k\big(\mu_{f,x}(f)x\big) = \mu_{f,x}(f)f^k(x), \tag{8.253}$$

de telle sorte que $\mu_{f,x}(f)$ est nul sur B et donc est nul sur $E_{f,x}$. Autrement dit,

$$\mu_{f,x}\big(f|_{E_{f,x}}\big) = 0. \tag{8.254}$$

Montrons que $\mu_{f,x}$ est même minimal pour $f|_{E_{f,x}}$.

Supposons avoir Q, un polynôme non nul de degré $p_{f,x} - 1$ annulant $f|_{E_{f,x}}$. En particulier $Q(f)x = 0$. Cela signifie que B est un système lié, alors que nous avons montré que c'était un système libre. Contradiction. Nous concluons que $\mu_{f,x}$ est le polynôme minimal de $f|_{E_{f,x}}$. □

Cette histoire de densité permet de donner une démonstration alternative du théorème de Cayley-Hamilton.

ThoCalYWLbJQ

Théorème 8.120 (Cayley-Hamlilton)**.**
Le polynôme caractéristique est un polynôme annulateur.

Une démonstration plus simple via la densité des diagonalisables est donnée en théorème 13.25.

Démonstration. Nous devons prouver que $\chi_f(f)x = 0$ pour tout $x \in E$. Pour cela nous nous fixons un $x \in E$, nous considérons l'espace $E_{f,x}$ et $\chi_{f,x}$, le polynôme caractéristique de $f|_{E_{f,x}}$. Étant donné que $E_{f,x}$ est stable par f, le polynôme caractéristique de $f|_{E_{f,x}}$ divise χ_f, c'est-à-dire qu'il existe un polynôme Q_x tel que

$$\chi_f = Q_x \chi_{f,x}, \tag{8.255}$$

et donc aussi

$$\chi_f(f)x = Q_x(f)\big(\chi_{f,x}(f)x\big) = 0 \tag{8.256}$$

parce que la proposition 8.119 nous indique que $\chi_{f,x}$ est un polynôme annulateur de $f|_{E_{f,x}}$. □

Corolaire 8.121.
Le degré du polynôme minimal est majoré par la dimension de l'espace.

Démonstration. Le polynôme minimal divise le polynôme caractéristique parce qu'il engendre l'idéal des polynômes annulateurs par la proposition 8.102. Or le degré du polynôme caractéristique est la dimension de l'espace par le lemme 8.115. □

Exemple 8.122 (Calcul de l'inverse d'un endomorphisme)**.**
Le théorème de Cayley-Hamilton donne un moyen de calculer l'inverse d'un endomorphisme inversible pourvu que l'on connaisse son polynôme caractéristique. En effet, supposons que

$$\chi_f(X) = \sum_{k=0}^{n} a_k X^k. \tag{8.257}$$

Nous aurons alors

$$0 = \chi_f(f) = \sum_{k=0}^{n} a_k f^k. \tag{8.258}$$

Nous appliquons f^{-1} à cette dernière égalité en sachant que $f^{-1}(0) = 0$:

$$0 = a_0 f^{-1} + \sum_{k=1}^{n} a_k f^{k-1}, \tag{8.259}$$

et donc

$$f^{-1} = -\frac{1}{\det(f)} \sum_{k=1}^{n} a_k f^{k-1} \tag{8.260}$$

où nous avons utilisé le fait que $a_0 = \chi_f(0) = \det(f)$. △

PropooBYZCooBmYLSc

Proposition 8.123.
Si $(X - z)^l$ $(l \geqslant 1)$ est la plus grande puissance de $(X - z)$ dans le polynôme caractéristique d'un endomorphisme f alors

$$1 \leqslant \dim(E_z) \leqslant l. \tag{8.261}$$

C'est-à-dire que nous avons au moins un vecteur propre pour chaque racine du polynôme caractéristique.

Démonstration. Si $(X-z)$ divise χ_f alors en posant $\chi_f = (X-z)P(X)$ nous avons

$$\det(f - X\mathbb{1}) = (X-z)P(X), \tag{8.262}$$

ce qui, évalué en $X = z$, donne $\det(f - z\mathbb{1}) = 0$. L'annulation du déterminant étant équivalente à l'existence d'un noyau non trivial, nous avons $v \neq 0$ dans E tel que $(f - z\mathbb{1})v = 0$. Cela donne $f(v) = zv$ et montre que v est vecteur propre de f pour la valeur propre z. Et aussi que $\dim(E_z) \geqslant 1$.

Si $\dim(E_z) = k$ alors le théorème de la base incomplète 4.14 nous permet d'écrire une base de E dont les k premiers vecteurs forment une base de E_z. Dans cette base, la matrice de f est de la forme

$$\begin{pmatrix} z & & & * \\ & \ddots & & \vdots \\ & & z & * \\ & & & * \end{pmatrix} \tag{8.263}$$

où les étoiles représentent des blocs à priori non nuls. En tout cas, sous cette forme, il est visible que $(X-z)^k$ divise χ_f. □

8.8 Formes bilinéaires et quadratiques

Plus à propos de formes bilinéaires dans le thème 16.

DEFooEEQGooNiPjHz

Définition 8.124 ([278])**.**
*Soient trois espaces vectoriels E, F et V sur le même corps commutatif $\mathbb{K}$. Une application $b\colon E \times F \to V$ est **bilinéaire** si elle est séparément linéaire en ses deux variables, c'est-à-dire si*

(1) $b(u_1 + u_2, v) = b(u_1, v) + b(u_2, v)$,

(2) $b(u, v_1 + v_2) = b(u, v_1) + b(u, v_2)$

(3) $b(\lambda u, v) = b(u, \lambda v) = \lambda b(u, v)$

pour tout $u, u_1, u_2 \in E$, $v, v_1, v_2 \in F$ et pour tout $\lambda \in \mathbb{K}$.

*(1) Dans le cas $E = F$ et $V = \mathbb{K}$, nous parlons de **forme bilinéaire** sur E.*

*(2) Nous parlons de forme bilinéaire **symétrique** si de plus $b(u, v) = b(v, u)$.*

DefBSIoouvuKR

Définition 8.125 ([279])**.**
*Soit un espace vectoriel E et $\mathbb{F}$ un corps de caractéristique différente de 2. Une **forme quadratique** sur E est une application $q\colon E \to \mathbb{F}$ pour laquelle il existe une forme bilinéaire symétrique $b\colon E \times E \to \mathbb{F}$ satisfaisant $q(x) = b(x, x)$ pour tout $x \in E$.*

L'ensemble des formes quadratiques réelles sur E est noté $Q(E)$.

La topologie sur $Q(E)$ sera la topologie métrique donnée dans le lemme 8.140.

DEFooJIAQooZkBtTy

Définition 8.126 (Application bilinéaire définie positive, thème 39)**.**
Si b est une application bilinéaire[44] *sur un espace vectoriel E nous disons qu'elle est*

(1) ***définie positive*** *si $b(x, x) \geqslant 0$ pour tout $x \in E$ et $b(x, x) = 0$ si et seulement si $x = 0$.*

(2) ***semi-définie positive*** *si $b(x, x) \geqslant 0$ pour tout $x \in E$. Nous dirons aussi parfois qu'elle est simplement « positive ».*

8.127.
Une application bilinéaire $E \times E \to \mathbb{K}$ n'est pas une application linéaire ; la distinction est importante. La linéarité est

$$b(\lambda u, \lambda v) = b\big(\lambda(u, v)\big) = \lambda b(u, v) \tag{8.264}$$

44. Définition 8.124.

et la bilinéarité est

$$b(\lambda u, v) = b(u, \lambda v) = \lambda b(u, v). \tag{8.265}$$

En réalité la seule forme qui soit à la fois linéaire et bilinéaire est la forme identiquement nulle : la condition

$$b(\lambda u, \lambda v) = \lambda^2 b(u, v) = \lambda b(u, v) \tag{8.266}$$

pour tout $\lambda \in \mathbb{K}$ implique $b(u, v) = 0$.

EXooKFBOooJjXZwu

Exemple 8.128 ([280]).
L'application

$$\begin{aligned} b\colon \mathbb{M}(n, \mathbb{K}) \times \mathbb{M}(n, \mathbb{K}) &\to \mathbb{K} \\ (A, B) &\mapsto \mathrm{Tr}(AB) \end{aligned} \tag{8.267}$$

est une forme bilinéaire symétrique.

La vérification est un calcul :

$$\mathrm{Tr}(BA) = \sum_i (BA)_{ii} = \sum_{ik} B_{ik} A_{ki} = \sum_{ik} A_{ki} B_{ik} = \sum_k (AB)_{kk} = \mathrm{Tr}(AB). \tag{8.268}$$

△

8.8.1 Dégénérescence d'une forme bilinéaire

Soit b, une forme bilinéaire symétrique non dégénérée sur l'espace vectoriel E de dimension n sur $\mathbb{K}$ où $\mathbb{K}$ est un corps de caractéristique différente de 2. Nous notons q la forme quadratique associée.

DEFooNUBFooLfCqaK

Définition 8.129.
Une forme bilinéaire est ***non dégénérée*** *si $b(x, z) = 0$ pour tout z implique $x = 0$.*

LemyKJpVP

Lemme 8.130.
Soit b une forme bilinéaire non dégénérée. Si x et y sont tels que $b(x, z) = b(y, z)$ pour tout z, alors $x = y$.

Démonstration. C'est immédiat du fait de la linéarité en le premier argument et de la non-dégénérescence : si $b(x, z) - b(y, z) = 0$ alors

$$b(x - y, z) = 0 \tag{8.269}$$

pour tout z, ce qui implique $x - y = 0$. □

8.8.2 Orthogonal pour une forme bilinéaire

DEFooENZGooXMWfUy

Définition 8.131.
Soit un espace vectoriel E et une forme bilinéaire symétrique b. Si $A \subset E$ nous notons

$$A^\perp = \{z \in E \text{ tel que } b(z, x) = 0 \, \forall x \in A\}. \tag{8.270}$$

*C'est l'****orthogonal*** *de A par rapport à la forme b.*

DEFooQQBQooKJdwxO

Définition 8.132.
Soit un espace vectoriel E et une forme bilinéaire symétrique b. Le ***noyau*** *est la partie*

$$\ker(b) = \{z \in E \text{ tel que } b(z, x) = 0 \forall x \in E\}. \tag{8.271}$$

PROPooWPKRooUAnVzd

Proposition 8.133 ([281]).
Soit un espace vectoriel E sur le corps $\mathbb{K}$. Soit une forme bilinéaire symétrique b sur E. Si A et B sont des parties de E nous avons :

ITEMooCGEIooKpSNXS

(1) $A^\perp$ *est un sous-espace vectoriel.*

ITEMooRRAOooTaVYgm

(2) *Si* $A \subset B$ *alors* $B^\perp \subset A^\perp$.

ITEMooBKXFooEpObWC

(3) $A^\perp = \text{Span}(A)^\perp$.

Démonstration. Point par point.

(1) **Pour (1)**
Soient $x, y \in A^\perp$, $a \in A$ et $\lambda \in \mathbb{K}$. Nous avons alors

$$b(x + \lambda y, a) = b(x, a) + \lambda b(y, a) = 0 \tag{8.272}$$

(2) **Pour (2)**
Soient $x \in B^\perp$ et $a \in A$. Nous avons $b(x, a) = 0$ parce que $a \in A \subset B$ et $x \in B^\perp$.

(3) **Pour (3)** Deux inclusions à prouver.

(3a) **Première inclusion** Nous avons $A \subset \text{Span}(A)$, donc le point (2) montre que $\text{Span}(A)^\perp \subset A^\perp$.

(3b) **Deuxième inclusion** Soit $x \in A^\perp$ et $y \in \text{Span}(A)$. Nous avons $y = \sum_i \lambda_i a_i$ pour certains $a_i \in A$ et $\lambda_i \in \mathbb{K}$. Alors

$$b(x, y) = \sum_i \lambda_i b(x, a_i) = 0. \tag{8.273}$$

□

PROPooSACFooQTsiJL

Proposition 8.134 ([281]).
Soit un espace vectoriel E sur le corps $\mathbb{K}$. Soit une forme bilinéaire symétrique b sur E. Si V est un sous-espace de E, alors :

ITEMooYHJUooQnbKQc

(1) *Si* $b_V : V \times V \to \mathbb{K}$ *est la restriction de* b, *alors* $V \cap V^\perp = \ker(b_V)$ [45].

ITEMooUZZFooDbuddF

(2) $(V + W)^\perp = V^\perp \cap W^\perp$.

ITEMooSZAGooLgkzYw

(3) $V \subset (V^\perp)^\perp$.

Démonstration. En plusieurs parties.

(1) **Pour (1)**
En deux parties.

(1a) $V \cap V^\perp \subset \ker(b_V)$ Soit $x \in V \cap V^\perp$. Si $v \in V$, nous avons $b(x, v) = 0$ parce que $x \in V^\perp$.

(1b) $\ker(b_V) \subset V \cap V^\perp$ Soit $x \in \ker(b_V)$. Nous avons $x \in V$ parce que $x \in \ker(b_V) \subset V$. Pour montrer que $x \in V^\perp$ nous considérons $y \in V$ et nous remarquons que $b(x, y) = 0$ parce que $x \in V^\perp$.

(2) **Pour (2)** En deux parties.

(2a) $(V + W)^\perp \subset V^\perp \cap W^\perp$ Soit $x \in (V+W)^\perp$. Si $v \in V$, alors $v \in V+W$, donc $f(x, v) = 0$. Cela prouve que $x \in V^\perp$. Nous prouvons que $x \in W^\perp$ de même.

(2b) $V^\perp \cap W^\perp \subset (V + W)^\perp$ Un élément de $V + W$ est de la forme $v + w$ avec $v \in V$ et $w \in W$. Nous avons

$$b(x, v + w) = b(x, v) + b(x, w) = 0. \tag{8.274}$$

(3) **Pour (3)** Si $x \in V$, alors $b(x, z) = 0$ pour tout $V^\perp$.

□

45. Définition du noyau de b, définition 8.132.

PROPooMAZZooRZZoHj

Proposition 8.135.
Soit une forme bilinéaire non dégénérée b sur l'espace vectoriel E de dimension finie. L'application

$$\begin{aligned}\varphi_b\colon E &\to E^*\\ x &\mapsto b(x,.)\end{aligned} \tag{8.275}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. En plusieurs parties.

(1) **Linéaire**
Le fait que φ_b soit linéaire est une conséquence de la bilinéarité de b.

(2) **Injectif** Si $\varphi_b(v) = \varphi_b(w)$, alors pour tout $x \in E$ nous avons $\varphi_b(v)x = \varphi_b(w)x$, c'est-à-dire

$$b(v - w, x) = 0. \tag{8.276}$$

Vu que b est non dégénérée, cela implique $v - w = 0$.

(3) **Isomorphisme** Nous savons que $\dim(E) = \dim(E^*)$ par le lemme 4.127. Une application linéaire injective entre espaces de même dimension est toujours une bijection par le corolaire 4.49.

□

LEMooRXMMooAvvOjF

Lemme 8.136 ([282]).
Soit une forme bilinéaire non dégénérée b sur l'espace vectoriel E de dimension finie. Soit un sous-espace vectoriel V. Nous avons

$$\dim(E) = \dim(V) + \dim(V^\perp). \tag{8.277}$$

EQooLFJBooNNJHgP

Démonstration. Nous considérons les applications

$$\begin{aligned}\varphi_b\colon E &\to E^*\\ x &\mapsto b(x,.).\end{aligned} \tag{8.278}$$

et l'application de restriction

$$\begin{aligned}r\colon E^* &\to V^*\\ \alpha &\mapsto \alpha|_F\end{aligned} \tag{8.279}$$

(1) **Théorème du rang** Nous appliquons le théorème du rang 4.50 à l'application r :

$$\mathrm{rk}(r) + \dim\big(\ker(r)\big) = \dim(E). \tag{8.280}$$

EQooSABBooUdWxZL

(2) $\mathrm{rk}(r) = \dim(V)$
Étant donné que r est surjective, $\mathrm{rk}(r) = \dim(V^*) = \dim(V)$. La seconde égalité est l'égalité des dimensions du lemme 4.127.

(3) $\varphi_b(V^\perp) \subset \ker(r)$ Soit $x \in V^\perp$. Pour tout $v \in V$ nous avons

$$r\big(\varphi_b(x)\big)v = \varphi_b(x)v = b(x,v) = 0. \tag{8.281}$$

Donc $\varphi_b(x) \in \ker(r)$.

(4) $\varphi_b\colon V^\perp \to \ker(r)$ **est surjective** Soit $\alpha \in \ker(r)$. Vu que φ_b est une bijection, il existe $x \in E$ tel que $\alpha = \varphi_b(x)$, c'est-à-dire $\alpha = b(x, \ \cdot\ )$. Nous avons $\alpha(y) = 0$ pour tout $y \in V$, c'est-à-dire $b(x,y) = 0$ pour tout $y \in V$, ce qui signifie que $x \in V^\perp$. Donc $\alpha \in \varphi_b(V^\perp)$.

(5) $\dim(V^\perp) = \dim\big(\ker(r)\big)$ L'application φ_b étant globalement injective, elle est une bijection entre $V^\perp$ et $\ker(r)$. Donc ces deux ont la même dimension.

(6) **Conclusion** Nous pouvons reprendre l'équation (8.280) et y mettre nos découvertes : $\text{rk}(r) = \dim(V)$ et $\dim\big(\ker(r)\big) = \dim(V^\perp)$. Cela donne la formule demandée

$$\dim(V) + \dim(V^\perp) = \dim(E). \tag{8.282}$$

□

LEMooYYGLooYIDmoa

Lemme 8.137.
Si V est un sous-espace de E de dimension finie, alors

$$(V^\perp)^\perp = V. \tag{8.283}$$

Démonstration. Nous avons déjà vu dans la proposition 8.134 que $V \subset (V^\perp)^\perp$.

Nous écrivons maintenant la formule (8.277) pour V, et pour $V^\perp$ (qui est un sous-espace vectoriel par la proposition 8.133(1)) :

$$\dim(V) + \dim(V^\perp) = \dim(E) \tag{8.284a}$$
$$\dim(V^\perp) + \dim\Big((V^\perp)^\perp\Big) = \dim(E). \tag{8.284b}$$

En soustrayant ces deux équations membre à membres, nous trouvons $\dim(V) = \dim\Big((V^\perp)^\perp\Big)$. Vu que V s'injecte (par l'identité) dans $(V^\perp)^\perp$ et qu'ils ont la même dimension, ils sont égaux. □

8.8.3 Formes quadratiques

LEMooOQIDooPSOeXL

Lemme 8.138 ([283]).
Soit une forme quadratique Q sur E. Si F est un sous-espace de E, alors

$$\dim(F) + \dim(F^\perp) \geqslant n \tag{8.285}$$

où $F^\perp$ est l'orthogonal par rapport à Q.

Démonstration. Nous posons $p = \dim(F)$. Nous considérons une base $\{f_i\}_{i=1,\dots,n}$ de E telle que $\{f_i\}_{i=1,\dots,p}$ est une base de F [46]. Nous posons

$$\begin{aligned}\phi\colon E &\to F\\ x &\mapsto \sum_{i=1}^{p} B(x, f_i) f_i\end{aligned} \tag{8.286}$$

où B est la forme bilinéaire associée à Q. Ce ϕ est une application linéaire à qui nous appliquons le théorème du rang (4.50) :

$$\dim(E) = \text{rk}(\phi) + \dim\big(\ker(\phi)\big). \tag{8.287}$$

EQooCLWLooCFxVDq

Mais vu que l'image de ϕ est dans F, nous avons $\text{rk}(\phi) \leqslant \dim(F)$. De plus, $\ker(\phi) = F^\perp$. Donc (8.287) devient

$$\dim(E) \leqslant \dim(F) + \dim(F^\perp). \tag{8.288}$$

□

LEMooUOZOooYvEcji

Lemme 8.139 ([283]).
Soit un espace vectoriel E de dimension finie et un sous-espace F sur lequel la forme quadratique Q est strictement définie positive ou négative. Alors

$$E = F \oplus F^\perp. \tag{8.289}$$

46. Théorème de la base incomplète, 4.25.

Démonstration. D'abord nous montrons que $F \cap F^\perp = \{0\}$. Si $v \neq 0$ est dans F, alors $Q(v) > 0$, et donc v n'est pas dans $F^\perp$. Donc $F \cap F^\perp \subset \{0\}$. L'inclusion inverse est immédiate.

Nous avons vu dans le lemme 8.138 que

$$\dim(E) \leqslant \dim(F) + \dim(F^\perp). \tag{8.290}$$

Vu que F et $F^\perp$ n'ont pas d'intersection autre que $\{0\}$, nous avons

$$\dim(E) \geqslant \dim(F \oplus F^\perp) = \dim(F) + \dim(F^\perp) \geqslant \dim(E). \tag{8.291}$$

Toutes ces inégalités sont donc des égalités et $\dim(E) = \dim(F) + \dim(F^\perp)$. □

LEMooWRCIooWkMpoF

Lemme 8.140.
La topologie considérée sur $Q(E)$[47] *est celle de la norme L'application*

$$\begin{aligned} N\colon Q(E) &\to \mathbb{R} \\ q &\mapsto \sup_{\|x\|_E=1} |q(x)|, \end{aligned} \tag{8.292}$$

EqZYBooZysmVh

est une norme.

L'application

$$\begin{aligned} N\colon S(n,\mathbb{R}) &\to \mathbb{R} \\ A &\mapsto \sup_{\|x\|_E=1} |x \cdot Ax| \end{aligned} \tag{8.293}$$

EQooJETQooIjxRWu

est une norme. À droite, nous trouvons le produit scalaire usuel sur $\mathbb{R}^n$ et la valeur absolue usuelle sur $\mathbb{R}$.

Ces normes sont celles que nous considérons pour la topologie sur $Q(E)$ et $S(n,\mathbb{R})$.

PROPooZLXVooOsXCcB

Proposition 8.141.
*Soit une forme bilinéaire b et la forme quadratique associée q. Alors nous avons l'**identité de polarisation** :*

$$b(x,y) = \frac{1}{2}\big(q(x) + q(y) - q(x-y)\big). \tag{8.294}$$

EqMrbsop

Démonstration. Il suffit de substituer dans le membre de droite $q(x) = b(x,x)$ et d'utiliser la bilinéarité :

$$\begin{aligned} q(x) + q(y) - q(x-y) &= b(x,x) + b(y,y) - b(x-y,x-y) && \text{(8.295a)} \\ &= b(x,x) + b(y,y) - b(x,x) + b(x,y) + b(y,x) - b(y,y) && \text{(8.295b)} \\ &= 2b(x,y) && \text{(8.295c)} \end{aligned}$$

où nous avons utilisé le fait que b est symétrique : $b(x,y) = b(y,x)$. □

PropFSXooRUMzdb

Proposition 8.142 (Matrice associée à une forme bilinéaire)**.**
Soit $\{e_i\}$ une base de E. L'application

$$\begin{aligned} \phi\colon Q(E) &\to S(n,\mathbb{R}) \\ q &\mapsto \big(b(e_i,e_j)\big)_{i,j} \end{aligned} \tag{8.296}$$

où b est forme bilinéaire associée à q est une bijection linéaire et continue[48]*.*

Démonstration. Si $\phi(q) = \phi(q')$; alors

$$q(x) = \sum_{i,j} \phi(q)_{ij} x_i x_j = \sum_{i,j} \phi(q')_{ij} x_i x_j = q'(x). \tag{8.297}$$

47. $Q(E)$ sont les formes quadratiques réelles sur E, définition 8.125.
48. Pour les topologies des normes données dans le lemme 8.140.

Donc $q = q'$. L'application ϕ est donc injective

De plus elle est surjective parce que si $B \in S(n, \mathbb{R})$ alors la forme quadratique

$$q(x) = \sum_{i,j} B_{ij} x_i x_j \tag{8.298}$$

a évidemment B comme matrice associée. L'application ϕ est donc surjective.

Notre application ϕ est de plus linéaire parce que l'association d'une forme quadratique à la forme bilinéaire associée est linéaire.

En ce qui concerne la continuité, nous la prouvons en zéro en considérant une suite convergente $q_n \xrightarrow{Q(E)} 0$. C'est-à-dire que

$$\sup_{\|x\|=1} |q_n(x)| \to 0. \tag{8.299}$$

Nous rappelons l'identité de polarisation [49] :

$$b_n(x, y) = \frac{1}{2}\big(q_n(x - y) - q_n(x) - q_n(y)\big). \tag{8.300}$$

En ce qui concerne deux des trois termes, il n'y a pas de problèmes :

$$\big|\phi(q_n)_{ij}\big| = \big|b_n(e_i, e_j)\big| \leqslant \frac{1}{2}\big|q_n(e_i - e_j)\big| + \frac{1}{2}\big|q_n(e_i)\big| + \frac{1}{2}\big|q_n(e_j)\big|. \tag{8.301}$$

Si n est assez grand, nous avons tout de suite

$$\big|\phi(q_n)_{ij}\big| \leqslant \frac{1}{2}\big|q_n(e_i - e_j)\big| + \epsilon. \tag{8.302}$$

Nous définissons e_{ij} et α_{ij} de telle sorte que $e_i - e_j = \alpha_{ij} e_{ij}$ avec $\|e_{ij}\| = 1$. Si $\alpha = \max\{\alpha_{ij}, 1\}$ alors nous avons

$$q_n(e_i - e_j) = \alpha_{ij}^2 q_n(e_{ij}) \leqslant \alpha^2 q_n(e_{ij}). \tag{8.303}$$

Il suffit maintenant de prendre n assez grand pour avoir $\sup_{\|x\|=1} |q_n(x)| \leqslant \frac{\epsilon}{\alpha^2}$ pour avoir

$$\big|\phi(q_n)_{ij}\big| \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{\alpha^2}. \tag{8.304}$$

□

8.8.4 Isotropie

DefVKMnUEM

Définition 8.143 (Isotropie[284])**.**
Quelques définitions.

(1) *Un vecteur est* ***isotrope*** *pour b si il est perpendiculaire à lui-même. En d'autres termes, x est isotrope si et seulement si $b(x, x) = 0$.*

(2) *Un sous-espace $W \subset E$ est* ***isotrope*** *si $W \cap W^\perp \neq \{0\}$.*

(3) *Un sous-espace $W \subset E$ est* ***totalement isotrope*** *si pour tout $x, y \in W$, nous avons $b(x, y) = 0$.*

Le ***cône isotrope*** *de b est l'ensemble de ses vecteurs isotropes :*

$$C(b) = \{x \in E \text{ tel que } b(x, x) = 0\}. \tag{8.305}$$

Nous introduisons quelques notations. D'abord pour $y \in E$ nous notons

$$\begin{aligned} \Phi_y \colon E &\to \mathbb{R} \\ x &\mapsto b(x, y) \end{aligned} \tag{8.306}$$

et ensuite

$$\begin{aligned} \Phi \colon E &\to E^* \\ y &\mapsto \Phi_y. \end{aligned} \tag{8.307}$$

49. Proposition 8.141.

DEFooHKQVooPuWZiP

Définition 8.144.
*Le fait pour une forme bilinéaire b d'être dégénérée signifie que l'application Φ n'est pas injective. Le **noyau** de la forme bilinéaire est celui de Φ, c'est-à-dire*

$$\ker(b) = \{z \in E \text{ tel que } b(z, y) = 0 \, \forall y \in E\}. \tag{8.308}$$

Autrement dit, $\ker(b) = E^\perp$ où le perpendiculaire est pris par rapport à b.

Notons tout de même que nous utilisons la notation $\perp$ même si b est dégénérée et éventuellement pas positive ; c'est-à-dire même si la formule $(x, y) \mapsto b(x, y)$ ne fournit pas un produit scalaire.

LEMooOTVGooRAOyaD

Lemme 8.145.
Si E est de dimension finie, et si V est un sous-espace non isotrope[50] *de E, alors $V^\perp$ est également non-isotrope.*

Démonstration. En dimension finie, le lemme 8.137 nous indique que $(V^\perp)^\perp = V$. Si $V^\perp$ était isotrope nous aurions $V^\perp \cap (V^\perp)^\perp \neq \varnothing$. Cela donne alors immédiatement $V^\perp \cap V \neq \varnothing$, qui serait que V est isotrope. □

8.9 Formes bilinéaires et quadratiques

Plus à propos de formes bilinéaires dans le thème 16.

LEMooLKNTooSfLSHt

Lemme 8.146.
Si q est une forme quadratique, il existe une unique forme bilinéaire b telle que $q(x) = b(x, x)$.

Démonstration. L'existence n'est pas en cause : c'est la définition d'une forme quadratique. Pour l'unicité, étant donné une forme quadratique, la forme bilinéaire b doit forcément vérifier l'identité de polarisation de la proposition 8.141. Elle est donc déterminée par q. □

Notons la division par 2 qui est le pourquoi de la demande de la caractéristique différente de 2 pour $\mathbb{F}$ dans la définition de forme quadratique.

DEFooGECOooCCGVXG

Définition 8.147.
*Soit une forme quadratique q sur E. Nous disons que $v, w \in E$ sont q-**orthogonaux** si $b(v, w) = 0$ où b est la forme bilinéaire associée à q par le lemme 8.146.*

8.9.1 Matrice associée à une forme bilinéaire

DEFooAOGPooXWXUcN

Définition 8.148.
Soit une forme bilinéaire[51] *$b : E \times E \to \mathbb{K}$ et une base quelconque $\{f_\alpha\}$ de E. Nous définissons les nombres*

$$B_{\alpha\beta} = b(f_\alpha, f_\beta), \tag{8.309}$$

EQooCUGFooRlKUtu

*qui forment une matrice symétrique dans $\mathbb{M}(n, \mathbb{K})$. Cette matrice est la **matrice associée** à la forme bilinéaire b.*

La matrice d'une forme quadratique est celle associée à sa forme bilinéaire associée.

LEMooDCIOooTlVZMR

Lemme 8.149.
Soit une forme bilinéaire $b : E \times E \to \mathbb{K}$ et une base quelconque $\{f_\alpha\}$ de E. Nous notons B la matrice de b (definition 8.148) et q la forme quadratique associée.

Alors nous avons

$$b(x, y) = \sum_{\alpha\beta} B_{\alpha\beta} x_\alpha y_\beta. \tag{8.310}$$

EQooQFMWooVKYLMx

50. Définition 8.143.
51. Définition 8.124.

et

$$b(x,y) = x \cdot By. \tag{8.311}$$

où le point est le produit scalaire usuel (composante par composante).

Démonstration. Si $x = \sum_\alpha x_\alpha f_\alpha$ et $y = \sum_\beta y_\beta f_\beta$:

En utilisant la convention (4.89) et les choses autour (voir aussi -1.1),

$$b(x,y) = \sum_\alpha x_\alpha \sum_\beta B_{\alpha\beta} y_\beta = \sum_\alpha x_\alpha (By)_\alpha = x \cdot By. \tag{8.312}$$

□

8.9.2 Changement de base : matrice d'une forme bilinéaire

PROPooOKTGooOYukoB

Proposition 8.150 ([1])**.**
Soient une forme quadratique q sur $\mathbb{R}^n$ et une application linéaire $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$. La matrice de $q \circ \phi$ est

$$\phi^t q \phi \tag{8.313}$$

où l'on a noté q la matrice de q et ϕ celle de ϕ.

Démonstration. C'est un calcul direct de $(q \circ \phi)(x)$:

$$\begin{aligned} q\big(\phi(x)\big) &= \sum_{ij} q_{ij} \phi(x)_i \phi(x)_j && (8.314a)\\ &= \sum_{ij}\sum_k\sum_l q_{ij\phi_{ik}} x_k \phi_{jl} x_l && \text{prop. 4.71} \quad (8.314b)\\ &= \sum_{kl} \Big(\sum_{ij} \phi_{ik} q_{ij} \phi_{jl} \Big) x_k x_l && = \sum_{kl} (\phi^t q \phi)_{kl} x_l x_l. \quad (8.314c) \end{aligned}$$

Et voila. □

PROPooLBIOooUpzxXA

Proposition 8.151 (Voir la section -1.1)**.**
Soit une forme bilinéaire[52] *$b\colon V \times V \to \mathbb{K}$ dont la matrice*[53] *dans la base $\{e_i\}$ est A et celle dans la base $\{f_\alpha\}$ est B. Nous supposons que les bases sont liées par $f_\alpha = \sum_i Q_{i\alpha} e_i$. Alors*

EQooZUVTooKjqnIj

$$B = Q^t A Q. \tag{8.315}$$

Démonstration. Soit $x, x' \in V$ de coordonnées (x_i) et (x'_i) dans la base $\{e_i\}$ et (y_α), (y'_α) dans la base $\{f_\alpha\}$. Par définition de la matrice associée à une forme bilinéaire,

$$b(x,x') = \sum_{ij} A_{ij} x_i x'_j = \sum_{\alpha\beta} B_{\alpha\beta} y_\alpha y'_\beta. \tag{8.316}$$

En remplaçant les x_i et x'_i par leurs valeurs en fonction de y_α et y'_β données par la proposition 4.116, nous trouvons

$$\begin{aligned} b(x,x') &= \sum_{ij\alpha\beta} A_{ij} Q_{i\alpha} y_\alpha Q_{j\beta} y'_\beta && (8.317a)\\ &= \sum_{\alpha\beta} (Q^t A Q)_{\alpha\beta} y_\alpha y'_\beta && (8.317b) \end{aligned}$$

où Q^t désigne la transposée de la matrice Q : $Q^t_{ij} = Q_{ji}$. Vu que les nombres y_α et y'_β sont arbitraires nous déduisons[54] que $B = Q^t A Q$. □

52. Définition 8.124
53. Définition 8.309.
54. Lemme 4.67.

REMooNEJLooSqgeih

Remarque 8.152.
Notons que cette « loi de transformation » n'est pas la même que celle pour une application linéaire[55]. Ici nous avons Q^t alors que pour les applications linéaires nous avions Q^{-1}.

Pour cette raison, tant que nous travaillons avec des bases orthonormées, c'est-à-dire tant que Q est orthogonale[56], nous pouvons confondre une application linéaire avec une application bilinéaire en passant par la matrice. Mais cette identification n'est pas du tout canonique : elle repose sur le fait que les bases soient orthonormées.

Il en découle que la réduction des endomorphismes et la réduction des formes bilinéaires ne sont pas tout à fait les mêmes théories. Par exemple la pseudo-diagonalisation simultanée (corollaire 10.40) est un résultat de réduction de forme bilinéaire et non d'endomorphismes.

8.9.3 Isométrie, forme quadratique et bilinéaire

Exemple 8.153.
La forme quadratique $q(x) = x_1^2 + x_2^2$ donne la norme euclidienne. La forme bilinéaire associée est $b(x, y) = x_1y_1 + x_2y_2$, qui est le produit scalaire usuel. △

Il ne faudrait pas déduire trop vite que la formule $\|x\|^2 = q(x)$ donne une norme dès que q est non dégénérée. En effet q peut ne pas être définie positive. La forme $q(x) = x_1^2 - x_2^2$ prend des valeurs positives et négatives. A fortiori $d(x, y) = q(x - y)$ ne donne pas toujours une distance.

DEFooECTUooRxBhHf

Définition 8.154.
*Une **isométrie** pour la forme quadratique q est une application bijective $f : V \to V$ telle que*

$$q(x - y) = q\big(f(x) - f(y)\big). \tag{8.318}$$

Dans les cas où q donne une distance, alors c'est une isométrie au sens usuel.

DEFooIQURooMeQuqX

Définition 8.155 (Thème 78)**.**
*Soit un espace vectoriel E muni d'une forme bilinéaire b. Une **isométrie** pour b est une bijection $f : E \to E$ telle que*

$$b\big(f(x), f(y)\big) = b(x, y) \tag{8.319}$$

pour tout $x, y \in E$.

LemewGJmM

Lemme 8.156.
Soient q une forme quadratique et b la forme bilinéaire associée par le lemme 8.146. Une application $f : E \to E$ telle que $f(0) = 0$ est une isométrie pour b si et seulement si elle est une isométrie pour q.

Démonstration. Pour une application bijective $f : E \to E$ telle que $f(0) = 0$, nous devons prouver l'équivalence des propriétés suivantes :

(1) $b\big(f(x), f(y)\big) = b(x, y)$ pour tout $x, y \in E$;

(2) $q\big(f(x) - f(y)\big) = q(x - y)$ pour tout $x, y \in E$.

Dans le sens direct, en posant $x = y$ nous trouvons tout de suite $q(f(x)) = q(x)$; ensuite en utilisant la distributivité de b,

$$\begin{aligned}
q\big(f(x) - f(y)\big) &= b\big(f(x) - f(y), f(x) - f(y)\big) && \text{(8.320a)}\\
&= q\big(f(x)\big) - 2b\big(f(x), f(y)\big) + q\big(f(y)\big) && \text{(8.320b)}\\
&= q(x) + q(y) - 2b(x, y) && \text{(8.320c)}\\
&= q(x - y). && \text{(8.320d)}
\end{aligned}$$

55. Proposition 4.119.
56. Définition 4.101.

Dans l'autre sens, nous commençons par remarquer que l'hypothèse $f(0) = 0$ donne $q(x) = q\big(f(x)\big)$. Ensuite nous utilisons l'identité de polarisation (8.294) :

$$b\big(f(x), f(y)\big) = \frac{1}{2}\big[q\big(f(x)\big) + q\big(f(y)\big) - q\big(f(x-y)\big)\big] \tag{8.321a}$$
$$= \frac{1}{2}\big[q(x) + q(y) - q(x-y)\big] \tag{8.321b}$$
$$= b(x, y). \tag{8.321c}$$

□

8.9.4 Isométries

Voici un théorème pas toujours bien énoncé dans les cours de physique qui font de la relativité. Au moment de « prouver » les transformations de Lorentz [57], beaucoup oublient de justifier pourquoi elles devraient être linéaires.

ThoDsFErq

Théorème 8.157 ([285]).
Une isométrie[58] *d'une forme bilinéaire non dégénérée est linéaire.*

Démonstration. Soient une forme bilinéaire non-dégénérée b sur l'espace vectoriel E ainsi qu'une isométrie f pour icelle. Soit $z \in E$; étant donné que f est bijective nous pouvons considérer l'élément $f^{-1}(z) \in E$ et calculer

$$b\big(f(x+y), z\big) = b\big(f(x+y), f(f^{-1}(z))\big) \tag{8.322a}$$
$$= b(x+y, f^{-1}(z)) \tag{8.322b}$$
$$= b(x, f^{-1}(z)) + b(y, f^{-1}(z)) \tag{8.322c}$$
$$= b(f(x), z) + b(f(y), z) \tag{8.322d}$$
$$= b\big(f(x) + f(y), z\big), \tag{8.322e}$$

donc $f(x+y) = f(x) + f(y)$ par le lemme 8.130.

De la même façon on trouve $b\big(f(\lambda x), z\big) = b\big(\lambda f(x), z\big)$ qui prouve que $f(\lambda x) = \lambda f(x)$ et donc que f est linéaire. □

Exemple 8.158.
Une isométrie peut ne pas être linéaire quand la forme bilinéaire est dégénérée. Par exemple pour la forme bilinéaire sur $\mathbb{R}^2$ donnée par

$$b\big((a,b),(x,y)\big) = ax, \tag{8.323}$$

nous pouvons faire

$$f(x,y) = \begin{pmatrix} x \\ \lambda(x,y) \end{pmatrix} \tag{8.324}$$

où λ est n'importe quoi. △

Remarque 8.159.
Des preuves alternatives.

(1) En utilisant un peut plus d'indices et un peu plus de mots comme « tenseurs », peut être trouvée dans [286]. Le fait que la preuve donnée soit tensorielle me fait penser que le résultat peut encore être généralisé.

(2) Et encore une autre preuve, utilisant des techniques de groupes de Lie sera la proposition 53.73.

57. Théorème 18.199.
58. Définition 8.155.

DEFooVTXWooVXfUnc

Définition 8.160.
*Soient deux espaces vectoriels E et V. Une application $f\colon E \to V$ est **affine** si il existe une application linéaire $u\colon E \to V$ et un élément $\alpha \in V$ tel que*

$$f(x) = u(x) + \alpha \tag{8.325}$$

pour tout $x \in E$.

THOooHSGHooNCLebs

Théorème 8.161.
Soit un espace vectoriel E muni d'une forme quadratique q. Soit une isométrie $f\colon E \to E$ pour q. Alors

(1) si $f(0) = 0$, alors f est linéaire ;

(2) si $f(0) \neq 0$ alors f est affine[59].

Démonstration. Nous considérons la forme bilinéaire associée b. Si $f(0) = 0$, nous savons par le lemme 8.156 que $b\big(f(x), f(y)\big) = b(x, y)$. La proposition 8.157 nous dit alors que f est linéaire.

Si $f(0) \neq 0$, alors nous posons $g(x) = f(x) - f(0)$ qui vérifie $g(0) = 0$ et

$$q\big(g(x) - g(y)\big) = q\big(f(x) - f(0) - f(y) + f(0)\big) = q(x - y). \tag{8.326}$$

Nous pouvons donc appliquer le premier point à g, déduire que g est linéaire et donc que f est affine. □

PROPooHSWIooKwiibH

Proposition 8.162 ([1]).
Soit un espace vectoriel E. Soient des points $a, x_1, \dots, x_n$ tels que $\{x_i - a\}_{i=1,\dots,n}$ forme une base de E. Si $f\colon E \to E$ est une application affine fixant a et tous les x_i, alors f est l'identité.

Démonstration. Nous décomposons $f(x) = u(x) + \alpha$ où u est linéaire et $\alpha \in E$. Nous avons, pour tout i :

$$\begin{cases} u(a) - a = \alpha & \text{(8.327a)} \\ u(x_i) - x_i = \alpha. & \text{(8.327b)} \end{cases}$$

En faisant la soustraction, nous obtenons $u(x_i - a) = x_i - a$. Donc u est une application linéaire fixant une base. Elle est donc l'identité. À partir de là, il est immédiat que $\alpha = 0$. □

Nous pouvons maintenant particulariser tout cela au cas de $\mathbb{R}^n$ muni du produit scalaire usuel et de la norme associée pour voir quel résultat nous avons à peine prouvé.

Nous notons ici $T(n)$ le groupe des translations sur $\mathbb{R}^n$. Un élément de $T(n)$ est une translation τ_v donnée par un vecteur v et agissant sur $\mathbb{R}^n$ par

$$\begin{aligned} \tau_v\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto x + v. \end{aligned} \tag{8.328}$$

Ce groupe est isomorphe au groupe abélien $(\mathbb{R}^n, +)$, et nous allons souvent identifier τ_v à v.

Vous savez par culture générale que les isométries de $\mathbb{R}^n$ pour le produit scalaire usuel sont les matrices orthogonales. En voici une petite généralisation (pensez à $\eta = \mathbb{1}$ dans le cas du produit scalaire usuel).

PROPooSYQMooEnZFdp

Proposition 8.163.
Soit une forme bilinéaire b sur $\mathbb{R}^n$ de matrice symétrique η. Si A est la matrice d'une application linéaire $\mathbb{R}^n \to \mathbb{R}^n$ telle que

$$b(Ax, Ay) = b(x, y) \tag{8.329}$$

pour tout $x, y \in \mathbb{R}^n$, alors

$$A^t \eta A = \eta. \tag{8.330}$$

59. Définition 8.160.

Démonstration. En suivant la formule générale (8.310),

$$b(Ax, Ay) = \sum_{ij} \eta_{ij}(Ax)_i(Ay)_j = \sum_{ijkl} \eta_{ij}A_{ik}A_{jl}x_k y_l. \tag{8.331}$$

En imposant que ce soit égal à $\sum_{kl} \eta_{kl}x_k y_l$ pour tout x, y nous avons la contrainte

$$\sum_{ij} \eta_{ij}A_{ik}A_{jl} = \eta_{kl} \tag{8.332}$$

qui signifie exactement $A^t\eta A = \eta$. □

8.10 Signature, théorème de Sylvester

DEFooWDCLooDkRYLK

Définition 8.164 (Signature[287])**.**
Soit une forme quadratique[60] *Q sur un espace vectoriel E de dimension finie n. L'**indice d'inertie** de Q est le nombre*

$$q = \max\{\dim(F) \text{ tel que } Q(v) < 0\, \forall v \in F\backslash\{0\}\}. \tag{8.333}$$

Nous définissons aussi

$$p = \max\{\dim(G) \text{ tel que } Q(v) > 0\, \forall v \in G\backslash\{0\}\}. \tag{8.334}$$

*Le couple (p, q) est la **signature** de Q.*

DEFooVITQooQaMaTF

Définition 8.165 (Rang d'une forme quadratique)**.**
Si $Q\colon E \to \mathbb{K}$ est une forme quadratique, nous considérons l'application

$$\begin{aligned} f_Q\colon E &\to E^* \\ x &\mapsto \big[y \mapsto B(x, y)\big]. \end{aligned} \tag{8.335}$$

*Le **rang** de Q est le rang de l'application linéaire f_Q.*

PROPooLRZQooSfprff

Proposition 8.166.
Le rang d'une forme quadratique est le rang de sa matrice dans n'importe quelle base.

Démonstration. Nous considérons une forme quadratique Q sur l'espace vectoriel E. Sa trace est, par définition, la trace de l'application linéaire f_Q de la définition 8.165. Or cette dernière trace ne dépend pas des bases choisies sur E et E^*. Nous la calculons donc maintenant.

Soit une base $\{e_i\}$ de E ainsi que sa base duale $\{e_i^*\}$ de E^*. Si $v = \sum_k v_k e_k \in E$, alors

$$f_Q(e_i)v = \sum_k v_k B(e_i, e_k) = \sum_k Q_{ik}v_k \tag{8.336}$$

où nous avons noté B la forme bilinéaire associée à Q et utilisé la définition 8.148 de la matrice associée à la forme quadratique Q. Nous avons donc $f_Q(ei) = \sum_k Q_{ik}e_k^*$ ou encore

$$f_Q(e_i)_k = Q_{ik}, \tag{8.337}$$

ce qui signifie, par (1) que la matrice associée à f_Q est la matrice Q^t.

Le rang de f_Q est donc celui de Q^t, qui est le même que celui de la matrice Q (ici, nous avons noté Q la matrice de la forme quadratique Q). Le rang de f_Q est celui de sa matrice par la proposition 4.110. □

60. Définition 8.125.

LEMooISHCooVDJEKo

Lemme 8.167 ([287]).
Soient une forme quadratique Q ainsi que deux bases Q-orthogonales $\{e_1,\dots,e_n\}$ et $\{e'_1,\dots,e'_n\}$. Nous posons

$$r = \text{Card}\{e_i \text{ tel que } Q(e_i) > 0\} \tag{8.338a}$$
$$r' = \text{Card}\{e_i \text{ tel que } Q(e'_i) > 0\} \tag{8.338b}$$
$$s = \text{Card}\{e_i \text{ tel que } Q(e_i) < 0\} \tag{8.338c}$$
$$s' = \text{Card}\{e_i \text{ tel que } Q(e'_i) < 0\} \tag{8.338d}$$

Alors $r = r'$ et $s = s'$.

Démonstration. Nous posons

$$I = \{i \text{ tel que } Q(e_i) \geqslant 0\} \tag{8.339a}$$
$$J = \{j \text{ tel que } Q(e'_j) \leqslant 0\} \tag{8.339b}$$

Nous commençons par prouver que $\{e_i\}_{i\in I} \cup \{e'_j\}_{j\in J}$ est libre. Supposons pour cela que

$$\sum_{i\in I} x_i e_i + \sum_{j\in J} y_j e'_j = 0, \tag{8.340}$$

et posons $z = \sum_{i\in I} x_i e_i$. Nous avons

EQooWGKAooElpETd
$$Q(z) = \sum_{i\in I} x_i^2 Q(e_i) \geqslant 0. \tag{8.341}$$

Mais nous avons aussi $z = -\sum_{j\in J} y_j e'_j$, donc

EQooJYOCooZPXmTf
$$Q(z) = \sum_{j\in J} y_j^2 Q(e'_j) \leqslant 0. \tag{8.342}$$

Donc $Q(z) = 0$. Vu (8.341), et le fait que $Q(e_i) > 0$, avoir $Q(z) = 0$ impose $x_i = 0$ pour tout i. La relation (8.342) nous donne aussi immédiatement que les y_j sont nuls. Donc la partie $\{e_i\}_{i\in I}\cup\{e'_j\}_{j\in J}$ est libre.

Le lemme 4.12 nous indique qu'une partie libre est toujours de cardinal plus petit ou égal à la dimension de l'espace[61]. Tout ça pour dire que

$$\underbrace{\text{Card}(I)}_{=r} + \underbrace{\text{Card}(J)}_{=n-r'} \leqslant n, \tag{8.343}$$

et donc $r \leqslant r'$.

Le même raisonnement, en partant de $I = \{i \text{ tel que } Q(e_i) \leqslant 0\}$ et de $J = \{j \text{ tel que } Q(e'_j) > 0\}$, prouve que $r' \leqslant r$.

La preuve de $s = s'$ est du même tonneau. □

Pour l'équivalence de formes quadratiques, voir la définition 8.252 et la proposition 8.253.

8.11 Produit scalaire, produit hermitien

DefVJIeTFj

Définition 8.168.
*Un **produit scalaire** sur un espace vectoriel réel est une forme bilinéaire*[62] *symétrique strictement définie positive*[63].

61. Ici nous utilisons l'hypothèse que V est de dimension finie.
62. Définition 8.124.
63. Définition 8.126.

La définition suivante est utile pour celles qui veulent faire de la relativité[64].

DEFooLPBGooXLxubc

Définition 8.169.
*Un **produit pseudo-scalaire** sur un espace vectoriel réel est une forme bilinéaire et symétrique.*

DEFooZBWTooIqXwRp

Définition 8.170.
*Nous dirons que deux vecteurs sont **orthogonaux** lorsque leur produit scalaire*[65] *est nul. Nous écrivons que $u \perp v$ lorsque $\langle u, v\rangle = 0$.*

*Si $\{e_i\}_{i=1,\dots,n}$ est une base de E, nous disons qu'elle est **orthonormée** si*

$$\langle e_i, e_j\rangle = \delta_{ij}. \tag{8.344}$$

LEMooLPUFooVCvnwW

Lemme 8.171.
Un produit scalaire est toujours non dégénéré[66].

Vu que nous allons voir un pâté d'espaces avec des produits scalaires, nous leur donnons un nom.

DefLZMcvfj

Définition 8.172.
*Un espace vectoriel **euclidien** est un espace vectoriel réel de dimension finie muni d'un produit scalaire (définition 8.168).*

Avouez que c'est drôle qu'un espace vectoriel est euclidien lorsqu'il possède une *multiplication* alors qu'un anneau est euclidien lorsqu'il possède une *division* (voir la définition 1.251). C'est pas très profond, mais si ça peut vous servir de moyen mnémotechnique...

PROPooSKVRooDGVCYj

Proposition-Définition 8.173 (Produit scalaire dans $\mathbb{R}^n$, thème 25)**.**
Si $x, y \in \mathbb{R}^n$, nous définissons

$$x \cdot y = \sum_{i=1}^{n} x_i y_i = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{8.345}$$

EQooFITHooEXDCGd

Cela est un produit scalaire.

*Ce produit scalaire est le **produit scalaire** qui sera toujours considéré. C'est de lui qui découle toujours la norme, et la topologie de $\mathbb{R}^n$. Il sera aussi souvent noté $\langle x, y\rangle$.*

Démonstration. Il faut que $b(x, y) = \sum_i x_i y_i$ soit bilinéaire, symétrique, strictement définie positive. Ce sont toutes des vérifications immédiates. Par exemple pour la symétrique, nous avons $\sum_i x_i y_i = \sum_i y_i x_i$. □

Calculons par exemple le produit scalaire de deux vecteurs de la base canonique : $\langle e_i, e_j\rangle$. En utilisant la formule de définition et le fait que $(e_i)_k = \delta_{ik}$, nous avons

$$\langle e_i, e_j\rangle = \sum_{k=1}^{m} \delta_{ik}\delta_{jk}. \tag{8.346}$$

Nous pouvons effectuer la somme sur k en remarquant qu'à cause du δ_{ik}, seul le terme avec $k = i$ n'est pas nul. Effectuer la somme revient donc à remplacer tous les k par des i :

$$\langle e_i, e_j\rangle = \delta_{ii}\delta_{ji} = \delta_{ji}. \tag{8.347}$$

Une des propriétés intéressantes du produit scalaire est qu'il permet de décomposer un vecteur dans une base, comme nous le montre la proposition suivante.

64. Voir le théorème 18.199 qui établit les transformations de Lorentz.
65. Définition 8.168.
66. Définition 8.129.

PropScalCompDec

Proposition 8.174.
Si nous notons v_i les composantes du vecteur v, c'est-à-dire si $v = \sum_{i=1}^{m} v_i e_i$, alors nous avons $v_j = \langle v, e_j \rangle$.

Démonstration.

$$v \cdot e_j = \sum_{i=1}^{m} \langle v_i e_i, e_j \rangle = \sum_{i=1}^{m} v_i \langle e_i, e_j \rangle = \sum_{i=1}^{m} v_i \delta_{ij} \tag{8.348}$$

Eqvejscalcomp

En effectuant la somme sur i dans le membre de droite de l'équation (8.348), tous les termes sont nuls sauf celui où $i = j$; il reste donc

$$v \cdot e_j = v_j. \tag{8.349}$$

□

Le produit scalaire ne dépend en réalité pas de la base orthogonale choisie.

LEMooZMCWooDfbrIq

Lemme 8.175.
Si $\{e_i\}$ est la base canonique, et si $\{f_i\}$ est une autre base orthonormale[67]*, alors si u et v sont deux vecteurs de $\mathbb{R}^m$, nous avons*

$$\sum_i u_i v_j = \sum_i u'_i v'_j \tag{8.350}$$

où u_i sont les composantes de u dans la base $\{e_i\}$ et u'_i sont celles dans la base $\{f_i\}$.

Démonstration. La preuve demande un peu d'algèbre linéaire. Étant donné que $\{f_i\}$ est une base orthonormale, il existe une matrice A orthogonale ($AA^t = \mathbb{1}$) telle que $u'_i = \sum_j A_{ij} u_j$ et idem pour v. Nous avons alors

$$\begin{aligned}
\sum_i u'_i v'_j &= \sum_i \left(\sum_j A_{ij} u_j \right) \left(\sum_k A_{ik} v_k \right) \\
&= \sum_{ijk} A_{ij} A_{ik} u_j v_k \\
&= \sum_{jk} \underbrace{\sum_i (A^t)_{ji} A_{ik}}_{=\delta_{jk}} u_j v_k \\
&= \sum_{jk} \delta_{jk} u_j v_k \\
&= \sum_k u_k v_k.
\end{aligned} \tag{8.351}$$

□

Cette proposition nous permet de réellement parler du produit scalaire entre deux vecteurs de façon intrinsèque sans nous soucier de la base dans laquelle nous exprimons les vecteurs.

8.11.1 Hermitien, unitaire, etc.

DefMZQxmQ

Définition 8.176 ([288])**.**
Soit E est un espace vectoriel sur $\mathbb{C}$. Une application $\langle .,. \rangle : E \times E \to \mathbb{C}$ est ***sesquilinéaire à droite*** *si pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{C}$,*

(1) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \bar{\lambda} y \rangle$,

(2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,

(3) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

Définitions supplémentaires :

67. Définition 8.170.

— *La forme $\langle .,. \rangle$ est* ***hermitienne*** *si de plus*

$$\langle x, y\rangle = \overline{\langle y, x\rangle}. \tag{8.352}$$

— *Un* ***produit hermitien*** *est une forme hermitienne strictement définie positive, c'est-à-dire telle que $\langle x, x\rangle \geqslant 0$ pour tout $x \in E$ et $\langle x, x\rangle = 0$ si et seulement si $x = 0$.*

PROPooRPRRooRYEMCB

Proposition-Définition 8.177.
Soit un espace hermitien $(E, \langle .,. \rangle)$. Soit un opérateur linéaire $A \in \mathrm{End}(E)$. Il existe un unique opérateur $B \in \mathrm{End}(E)$ tel que

$$\langle Ax, y\rangle = \langle x, By\rangle \tag{8.353}$$

pour tout $x, y \in E$.

L'opérateur B ainsi défini est noté $A^\dagger$ et est nommé ***hermitien conjugué*** *de A.*

LEMooBOXMooSDyCfm

Lemme 8.178.
Au niveau des matrices, nous avons

$$A^\dagger_{ij} = A^*_{ji} \tag{8.354}$$

où z^* est le conjugué complexe.

DEFooKEBHooWwCKRK

Définition 8.179.
Quelque définitions.

(1) Un opérateur A est ***hermitien*** *si $A^\dagger = A$.*

(2) Un opérateur A est ***unitaire*** *si $A^\dagger = A^{-1}$.*

Note : le conjugué hermitien est parfois noté A^* au lieu de $A^\dagger$.

8.180.
Le mot « hermitien » est réservé aux opérateurs sur des espaces hermitiens, c'est-à-dire des espaces vectoriels sur $\mathbb{C}$. Le mot « autoadjoint » par contre est plutôt utilisé dans le cadre d'opérateurs sur les espaces réels. En conséquence de quoi, ces deux mots sont synonymes, mais il est préférable d'utiliser « hermitien » lorsque l'espace vectoriel est sur $\mathbb{C}$ et « autoadjoint » lorsqu'il est sur $\mathbb{R}$.

L'ensemble des opérateurs autoadjoints de E est noté $\mathrm{S}(E)$. Cette notation provient du fait que dans $\mathbb{R}^n$ muni du produit scalaire usuel, les opérateurs autoadjoints sont les matrices symétriques.

Remarque 8.181.
Le fait d'être hermitien n'implique en rien le fait d'être inversible.

8.182.
Les normes associées aux produits scalaires font intervenir une racine carré, et donc devront être données plus tard. Voir le thème 25.

PROPooMWUCooMbJuaJ

Proposition 8.183 ([1])**.**
Nous considérons $\mathbb{C}^n$ vu comme espace vectoriel de dimension n sur $\mathbb{C}$.

(1) La formule, pour $x, y \in \mathbb{C}^n$,

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \bar{y}_k \tag{8.355}$$

EqFormSesqQrjvPH

définit une forme sesquilinéaire sur $\mathbb{C}^n$.

(2) L'ensemble $\mathbb{C}^n$ devient un espace vectoriel hermitien.

8.11.2 Éléments de matrice

PROPooZKWXooWmEzoA

Proposition 8.184.
Soit une application linéaire $A\colon \mathbb{R}^n \to \mathbb{R}^n$. Nous considérons le produit scalaire usuel[68] *sur $\mathbb{R}^n$. Alors :*

68. Définition 8.173.

(1) Les éléments de matrice de A sont donnés par $A_{ij} = e_i \cdot Ae_j$.

(2) Nous avons la formule $x \cdot Ay = \sum_{kl} A_{kl}x_k y_l$.

Démonstration. Pour Ae_j nous utilisons la formule 4.88 avec des notations plus décontractées : $Ae_j = \sum_k A_{kj}e_k$. Ensuite nous faisons un calcul avec la formule (8.345) :

$$e_i \cdot Ae_j = e_i \cdot \sum_k A_{kj}e_k = \sum_k A_{kj}\delta_{i,k} = A_{ij}. \tag{8.356}$$

La seconde formule à prouver est du même tonneau, en utilisant cette fois la formule (4.89) :

$$x \cdot Ay = \sum_k x_k(Ay)_k = \sum_{kl} x_k A_{kl} y_l = \sum_{kl} A_{kl}x_k y_l. \tag{8.357}$$

□

La proposition suivante est une version plus « pragmatique » de la proposition 4.134.

PROPooNITTooCYcrrT

Proposition 8.185 ([138])**.**
Soient un espace euclidien[69] *de dimension finie V ainsi qu'un sous-espace M. Nous posons*

$$M^\perp = \{x \in V \text{ tel que } x \cdot y = 0 \forall y \in M\}. \tag{8.358}$$

Alors $M \oplus M^\perp = V$.

Démonstration. D'abord si $x \in M \cap M^\perp$, alors $x \cdot x = 0$ et donc $x = 0$. Donc nous avons déjà $M \cap M^\perp = \{0\}$. Nous considérons une base $\{b_1, \dots, b_k\}$ de M, et nous définissons l'application linéaire

$$\begin{aligned} f\colon V &\to \mathbb{R}^k \\ x &\mapsto (x \cdot b_1, \dots, x \cdot b_k). \end{aligned} \tag{8.359}$$

Nous avons que $M^\perp = \ker(f)$. Le théorème du rang 4.47 nous indique que

$$\dim(V) = \dim\big(\ker(f)\big) + \dim\big(\text{Image}(f)\big) \leqslant \dim(M^\perp) + k = \dim(M^\perp) + \dim(M). \tag{8.360}$$

Une justification : vu que f prend ses valeurs dans $\mathbb{R}^k$, la dimension de son image est majorée par k.

Nous en déduisons que

$$\dim(M) + \dim(M^\perp) \geqslant \dim(V), \tag{8.361}$$

et la proposition 4.142 nous permet de conclure que $M \oplus M^\perp = V$. □

8.11.3 Transposée : pas d'approche naïve

SUBSECooGPXVooEYwIiJ

Il est légitime, si $t\colon E \to E$ est une application linéaire, de dire que sa transposée soit l'application linéaire $t^t\colon E \to E$ dont la matrice est la matrice transposée de celle de t. Lorsque nous travaillons sur $\mathbb{R}^n$ muni de la base canonique, cela ne pose pas de problème et nous pouvons écrire des égalités du type $\langle x, Ay\rangle = \langle A^t x, y\rangle$.

PROPooNARVooEuhweD

Proposition 8.186 (Matrice transposée et produit scalaire)**.**
Soit une matrice réelle A. En utilisant l'application linéaire associée[70] *$f_A\colon \mathbb{R}^n \to \mathbb{R}^n$, nous avons*

$$x \cdot f_A(y) = f_{A^t}(x) \cdot y. \tag{8.362}$$

Cette formule est souvent écrite $x \cdot Ay = A^t x \cdot y$ ou $\langle x, Ay\rangle = \langle A^t x, y\rangle$.

69. Qui possède un produit scalaire, définition 8.172.
70. Définition 4.68. Ici nous considérons la base canonique sur $\mathbb{R}^n$

Démonstration. Il s'agit d'un calcul utilisant la formule (4.83) et le produit scalaire (8.345) :

$$x \cdot f_A(y) = \sum_i x_i \big(f_A(y)\big)_i = \sum_i x_i \sum_j A_{ij} y_j = \sum_{ij} A^t_{ji} x_i y_j = \sum_j f_{A^t}(x)_j y_j = f_{A^t}(x) \cdot y. \tag{8.363}$$

□

Hélas nous allons voir que cette façon de définir une transposée est mauvaise.

Soit une application linéaire $t\colon E \to E$ de matrice A dans la base $\{e_i\}_{i=1,\dots,n}$ et de matrice B dans la base $\{f_\alpha\}_{\alpha=1,\dots,n}$. Nous notons Q la matrice de passage d'une base à l'autre :

$$e_i = \sum_\alpha Q^{-1}_{\alpha i} f_\alpha. \tag{8.364}$$

Nous nommons t_1 l'application linéaire associée à A^t dans la base $\{e_i\}$ et t_2 l'application linéaire associée à la matrice B^t dans la base $\{f_\alpha\}$. Définir la transposée d'une application linéaire comme étant l'application linéaire associée à la transposée de sa matrice ne sera une bonne définition que si $t_1 = t_2$.

La première chose facile à voir est

$$t_1(e_i)_j = \sum_k (A^t)_{jk}(e_i)_k = A^t_{ji} = A_{ij}. \tag{8.365}$$

EQooAMHPooUQEkJo

Pour calculer $t_2(e_i)_j$, c'est un peu plus laborieux :

$$t_2(e_i) = \sum_\alpha Q^{-1}_{\alpha i} t_2(f_\alpha) = \sum_{\beta\gamma\alpha} Q^{-1}_{\alpha i} B^t_{\gamma\beta} \underbrace{(f_\alpha)_\beta}_{\delta_{\alpha\beta}} f_\gamma = \sum_{\beta\gamma} Q^{-1}_{\beta i} B^t_{\gamma\beta} f_\gamma \tag{8.366a}$$

$$= \sum_{j\gamma} (B^t Q^{-1})_{\gamma i} Q_{j\gamma} e_j \tag{8.366b}$$

$$= \sum_j (QB^tQ^{-1})_{ji} e_j. \tag{8.366c}$$

Donc $t_2(e_i)_j = (QB^tQ^{-1})_{ji}$. En tenant compte du fait que $B = Q^{-1}AQ$ nous avons

$$t_2(e_i)_j = (QQ^tA^t(Q^{-1})^tQ^{-1})_{ji}. \tag{8.367}$$

Ceci est égal à l'expression (8.365) lorsque $Q^t = Q^{-1}$. Nous voyons que confondre transposée d'une application linéaire avec transposée de la matrice associée n'est valable que si nous sommes certain de ne considérer que des changements de base par des matrices orthogonales.

C'est la situation typique dans laquelle nous nous trouvons lorsque nous considérons des applications linéaires sur $\mathbb{R}^n$ muni de la base canonique, et que nous n'avons aucune intention de changer de base, et encore moins de chercher une base non orthonormale. Cette situation est clairement la situation la plus courante.

Exemple 8.187 ([93])**.**
Soit la base canonique $\{e_1, e_2\}$ de $\mathbb{R}^2$. Nous considérons l'application linéaire $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$t(e_1) = e_1 \tag{8.368a}$$

$$t(e_2) = 0. \tag{8.368b}$$

La matrice de t dans cette base est

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{8.369}$$

Elle est symétrique : elle vérifie $A^t = A$. Si nous comptions sur la transposée de matrice pour définir la transposée de t, nous aurions $t^t = t$.

Soit maintenant la base $f_1 = e_1$, $f_2 = e_1 + e_2$. Nous avons $t(f_1) = f_1$ et

$$t(f_2) = t(e_1) + t(e_2) = e_1 = f_1. \tag{8.370}$$

Donc la matrice de t dans cette base est

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}. \tag{8.371}$$

Et là, nous avons $B^t \neq B$. Donc en comptant sur cette base pour définir la transposée de t nous aurions $t^t \neq t$. △

NooMZVRooExWVKJ

8.188.
Autrement dit, la façon « usuelle » de voir la transposée d'une application linéaire, ne fonctionne dans les livres pour enfants uniquement parce qu'on y considère toujours $\mathbb{R}^n$ muni de la base canonique ou de bases orthonormées.

Notons que nous avons tout de même les notions d'opérateur adjoint et autoadjoint pour parler d'application orthogonale sans passer par la transposée, voir 8.43.

8.11.4 Transposée : la bonne approche

DefooZLPAooKTITdd

Définition 8.189.
*Si $f : E \to F$ est une application linéaire entre deux espaces vectoriels, la **transposée** est l'application $f^t : F^* \to E^*$ donnée par*

$$f^t(\omega)(x) = \omega\big(f(x)\big). \tag{8.372}$$

pour tout $\omega \in F^*$ et $x \in E$.

PROPooFAVAooDevBrT

Proposition 8.190 (transposée et conjuguée[1])**.**
Soit une application linéaire $A : \mathbb{R}^n \to \mathbb{R}^n$ ainsi que l'isomorphisme canonique $\phi : \mathbb{R}^n \to (\mathbb{R}^n)^*$. Nous avons

$$A^\dagger = \phi \circ A^t \circ \phi \tag{8.373}$$

où A^t est l'opérateur transposé, défini en 8.189, et $A^\dagger$ est l'hermitien conjugué défini en 8.177.

NORMooFUZYooWeXjp

8.191.
Étant données les propositions 8.190, le fait que l'isomorphisme soit canonique, que la matrice de l'hermitien conjugué est la matrice transposée (dans le cas réel, par le lemme 8.178), nous écrirons souvent A^t pour $A^\dagger$, et donc nous écrirons des égalités comme

$$x \cdot Ay = A^t x \cdot y \tag{8.374}$$

alors qu'il faudrait écrire

$$x \cdot Ay = A^\dagger x \cdot y \tag{8.375}$$

Notons également que cette remarque est une raison pour laquelle je préfère noter A^t au lieu de ${}^tA$.

LEMooEMNNooPquZMg

Lemme 8.192.
Soit E muni de la base $\{e_i\}$ et F muni de la base $\{g_i\}$ et une application $f : E \to F$. Si A est la matrice de f dans ces bases, alors A^t est la matrice de f^t dans les bases $\{e_i^*\}$ et $\{g_i^*\}$ de E^* et F^*.

Autrement dit, en utilisant l'application $\psi : \mathbb{M} \to \mathcal{L}(F^*, E^*)$ de la proposition 4.71,

$$\psi(A^t) = f^t. \tag{8.376}$$

Démonstration. Attention aux indices, ça va chauffer [71].

71. Et merci à Alain Vigne pour m'avoir fait remarquer qu'il fallait mettre de l'ordre dans les indices.

Nous allons montrer que $f^t = \psi(A^t)$ sur la base $\{g_i^*\}$, et pour cela nous appliquons $f^t(g_i^*)$ à $x \in E$:

$$\begin{aligned}
f^t(g_i^*)x &= g_i^*\big(f(x)\big) && \text{Définition 8.189} && (8.377a)\\
&= g_i^*\big(\sum_k x_k f(e_k)\big) && && (8.377b)\\
&= g_i^*\big(\sum_{kl} x_k A_{lk} g_l\big) && \text{eq. (4.88)} && (8.377c)\\
&= \sum_{kl} x_k A_{lk} \underbrace{g_i^*(g_l)}_{=\delta_{i,l}} && && (8.377d)\\
&= \sum_k x_k A_{ik} && && (8.377e)\\
&= \sum_k (A^t)_{ki} x_k && && (8.377f)\\
&= \sum_k (A^t)_{ki} e_k^*(x). && && (8.377g)
\end{aligned}$$

Voila. Donc nous avons

$$f^t(g_i^*) = \sum_k (A^t)_{ki} e_k^* = \psi(A^t) g_i^*. \tag{8.378}$$

□

8.193.
Intuitivement, les rangs de f et de f^t sont égaux parce que le rang est donné par la plus grande matrice carrée de déterminant non nul.

Nous donnons maintenant une vraie preuve de ce résultat.

LemSEpTcW

Lemme 8.194 ([289]).
Si $f\colon E \to F$ est une application linéaire, alors

$$\operatorname{rk}(f) = \operatorname{rk}(f^t). \tag{8.379}$$

Démonstration. Soient $n = \dim(E)$ et $r = \dim(F)$. Nous posons $\dim\ker(f) = p$ et donc $\operatorname{rk}(f) = n - p$. Soit $\{e_1, \dots, e_p\}$ une base de $\ker(f)$ que l'on complète en une base $\{e_1, \dots, e_n\}$ de E. Nous considérons maintenant les vecteurs

$$g_i = f(e_{p+i}) \tag{8.380}$$

pour $i = 1, \dots, n-p$. C'est-à-dire que les g_i sont les images des vecteurs qui ne sont pas dans le noyau de f. Prouvons qu'ils forment une famille libre. Si

$$\sum_{k=1}^{n-p} a_k f(e_{p+k}) = 0, \tag{8.381}$$

alors $f\big(\sum_k a_k e_{p+k}\big) = 0$, c'est-à-dire $\sum_k a_k e_{p+k} \in \ker(f)$. Comme une base de $\ker(f)$ est $\{e_1, \dots, e_p\}$, il existe des b_l tels que $\sum_k a_k e_{p+k} = \sum_{l=1}^p b_l e_l$. Autrement dit,

$$\sum_k a_k e_{p+k} - \sum_{l=1}^p b_l e_l = 0, \tag{8.382}$$

qui est une combinaison linéaire nulle des e_i. Donc $a_k = b_l = 0$ pour tout k et l. Tout ça pour dire que les $\{g_i\}_{i=1,\dots,n-p}$ est libre.

Étant donné que les vecteurs $g_1, \dots, g_{n-p}$ sont libres, nous pouvons les compléter en une base de F :

$$\{\underbrace{g_1, \dots, g_{n-p}}_{\text{images}}, \underbrace{g_{n-p+1}, \dots, g_r}_{\text{complétion}}\}. \tag{8.383}$$

Nous prouvons maintenant que $\mathrm{rk}(f^t) \geqslant n-p$ en montrant que les formes $\{g_i^*\}_{i=1,\dots,n-p}$ forment une partie libre (et donc l'espace image de f^t est au moins de dimension $n-p$). Pour cela nous prouvons que $f^t(g_i^*) = e_{i+p}^*$. En effet

$$f^t(g_i^*)(e_k) = g_i^*(f(e_k)), \tag{8.384}$$

Si $k = 1, \dots, p$, alors $f(e_k) = 0$ et donc $g_i^*(f(e_k)) = 0$; si $k = p + l$ alors

$$f^t(g_i^*)(e_k) = g_i^*(f(e_{k+l})) = g_i^*(g_l) = \delta_{i,l} = \delta_{i,k-p} = \delta_{k,i+p}. \tag{8.385}$$

Donc $f^t(g_i^*) = e_{i+p}^*$. Cela prouve que les formes $f^t(g_i^*)$ sont libres et donc que

$$\mathrm{rk}(f^t) \geqslant n - p = \mathrm{rk}(f). \tag{8.386}$$

En appliquant le même raisonnement à f^t au lieu de f, nous trouvons

$$\mathrm{rk}\left((f^t)^t\right) \geqslant \mathrm{rk}(f^t) \tag{8.387}$$

et donc, sachant que $(f^t)^t = f$, nous obtenons $\mathrm{rk}(f) = \mathrm{rk}(f^t)$.

□

PropWOPIooBHFDdP

Proposition 8.195 ([290]).
Si f est une application linéaire entre les espaces vectoriels E et F, alors nous avons

$$\mathrm{Image}(f^t) = \ker(f)^\perp. \tag{8.388}$$

Démonstration. Soient donc l'application $f\colon E \to F$ et sa transposée $f^t\colon F^* \to E^*$. Nous commençons par prouver que $\mathrm{Image}(f^t) \subset (\ker f)^\perp$. Pour cela nous prenons $\omega \in \mathrm{Image}(f^t)$, c'est-à-dire $\omega = \alpha \circ f$ pour un certain élément $\alpha \in F^*$. Si $z \in \ker(f)$, alors $\omega(z) = (\alpha \circ f)(z) = 0$, c'est-à-dire que $\omega \in (\ker f)^\perp$.

Pour prouver qu'il y a égalité, nous n'allons pas démontrer l'inclusion inverse, mais plutôt prouver que les dimensions sont égales. Après, on sait que si $A \subset B$ et si $\dim A = \dim B$, alors $A = B$. Nous avons

$$\begin{aligned}
\dim\left(\mathrm{Image}(f^t)\right) &= \mathrm{rk}(f^t) && && \text{(8.389a)}\\
&= \mathrm{rk}(f) && \text{lemme 8.194} && \text{(8.389b)}\\
&= \dim(E) - \dim\ker(f) && \text{théorème 4.47} && \text{(8.389c)}\\
&= \dim\left((\ker f)^\perp\right) && \text{proposition 4.134.} && \text{(8.389d)}
\end{aligned}$$

□

Lemme 8.196 ([130]).
Soit $\mathbb{K}$ un corps, E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et une application linéaire $f\colon E \to F$. L'application f est injective si et seulement si sa transposée[72] *f^t est surjective.*

Démonstration. Supposons que f soit injective. Alors par le lemme 4.56, il existe $g\colon F \to E$ tel que $g \circ f = \mathrm{Id}\,|_E$. Nous avons alors aussi $(g \circ f)^t = \mathrm{Id}\,|_{E^*}$, mais $(g \circ f)^t = f^t \circ g^t$, donc f^t est surjective.

Inversement, nous supposons que $f^t\colon F^* \to E^*$ est surjective. Alors en nous souvenant que E et F sont de dimension finie et en faisant jouer les identifications $(f^t)^t = f$ et $(E^*)^* = E$ nous savons qu'il existe $s\colon E^* \to F^*$ tel que $f^t \circ s = \mathrm{Id}\,|_{E^*}$. En passant à la transposée,

$$s^t \circ f = \mathrm{Id}\,|_E, \tag{8.390}$$

qui implique que f est injective. □

72. Définition 8.189.

8.11.5 Polynômes de Lagrange

Lemme-Définition 8.197.
Soit $E = \mathcal{P}_n(\mathbb{R})$ l'ensemble des polynômes à coefficients réels de degré au plus n. Soient $n+1$ réels distincts $a_0, \dots, a_n$. Nous considérons les formes linéaires associées $f_i \in \mathcal{P}_n(\mathbb{R})^*$,

$$f_i(P) = P(a_i). \tag{8.391}$$

La partie $\{f_1, \dots, f_n\}$ est une base de $\mathcal{P}_n(\mathbb{R})^*$.

*Les **polynômes de Lagrange** aux points (a_i) sont les polynômes de la base préduale de la base $\{f_i\}$.*

Démonstration. Nous prouvons que l'orthogonal est réduit au singleton nul :

$$\text{Span}\{f_0, \dots, f_n\}^\perp = \{0\}. \tag{8.392}$$

La proposition 4.134 conclura. Si $P \in \text{Span}\{f_i\}^\perp$, alors $f_i(P) = 0$ pour tout i, ce qui fait que $P(a_i) = 0$ pour tout $i = 0, \dots, n$. Un polynôme de degré au plus n qui s'annule en $n+1$ points est automatiquement le polynôme nul. □

Proposition 8.198.
Les polynômes de Lagrange aux points $(a_i)_{i=1,\dots,n}$ sont donnés par

$$P_i = \prod_{k \neq i} \frac{X - a_k}{a_i - a_k}. \tag{8.393}$$

Démonstration. Il suffit de vérifier que $f_j(P_i) = \delta_{i,j}$. Nous avons

$$f_j(P_i) = P_i(a_j) = \prod_{k \neq i} \frac{a_j - a_k}{a_i - a_k}. \tag{8.394}$$

Si $j \neq i$ alors un des termes est nul. Si au contraire $i = j$, tous les termes valent 1. □

8.11.6 Dual de $\mathbb{M}(n, \mathbb{K})$

PropHOjJpCa

Proposition 8.199 ([104]).
Soit $\mathbb{K}$, un corps. Les formes linéaires sur $\mathbb{M}(n, \mathbb{K})$ sont les applications de la forme

$$\begin{aligned} f_A \colon \mathbb{M}(n, \mathbb{K}) &\to \mathbb{K} \\ M &\mapsto \text{Tr}(AM). \end{aligned} \tag{8.395}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} f \colon \mathbb{M}(n, \mathbb{K}) &\to \mathbb{M}(n, \mathbb{K})^* \\ A &\mapsto f_A \end{aligned} \tag{8.396}$$

et nous voulons prouver que c'est une bijection. Étant donné que nous sommes en dimension finie, nous avons égalité des dimensions de $\mathbb{M}(n, \mathbb{K})$ et $(\mathbb{M}(n, \mathbb{K}))^*$, et il suffit de prouver que f est injective. Soit donc A telle que $f_A = 0$. Nous l'appliquons à la matrice $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$:

$$0 = f_A(E_{ij}) = \sum_k (AE_{ij})_{kk} = \sum_{kl} A_{kl}(E_{ij})_{lk} = \sum_{kl} A_{kl}\delta_{il}\delta_{jk} = A_{ji}. \tag{8.397}$$

Donc $A = 0$. □

Corolaire 8.200 ([104]).
Soient un corps $\mathbb{K}$ ainsi qu'une application $\phi \in \mathbb{M}(n, \mathbb{K})^*$ telle que pour tout $M, N \in \mathbb{M}(n, \mathbb{K})$ on ait

$$\phi(MN) = \phi(NM). \tag{8.398}$$

Alors il existe $\lambda \in \mathbb{K}$ tel que $\phi = \lambda \,\text{Tr}$.

Démonstration. La proposition 8.199 nous donne une matrice $A \in \mathbb{M}(n,\mathbb{K})$ telle que $\phi = f_A$. L'hypothèse nous dit que $f_A(MN) = f_A(NM)$, c'est-à-dire

$$\mathrm{Tr}(AMN) = \mathrm{Tr}(ANM) \tag{8.399}$$

pour toutes matrices $M, N \in \mathbb{M}(n,\mathbb{K})$. L'invariance cyclique de la trace[73] appliqué au membre de droite nous donne $\mathrm{Tr}(AMN) = \mathrm{Tr}(MAN)$, ce qui signifie que

$$\mathrm{Tr}\big((AM - MA)N\big) = 0 \tag{8.400}$$

ou encore que $f_{AM-MA} = 0$, et ce, pour toute matrice M. La fonction f étant injective nous en déduisons que la matrice A doit satisfaire

$$AM = MA \tag{8.401}$$

pour tout $M \in \mathbb{M}(n,\mathbb{K})$. En particulier, en prenant pour M les fameuses matrices E_{ij} et en calculant un peu,

$$A_{li}\delta_{j,m} = \delta_{i,l}A_{jm} \tag{8.402}$$

pour tout i, j, l, m. Cela implique que $A_{ll} = A_{mm}$ pour tout l et m et que $A_{jm} = 0$ dès que $j \neq m$. Il existe donc $\lambda \in \mathbb{K}$ tel que $A = \lambda\mathbb{1}$. En fin de compte,

$$\phi(X) = f_{\lambda\mathbb{1}}(X) = \lambda\,\mathrm{Tr}(X). \tag{8.403}$$

□

CorICUOooPsZQrg

Corolaire 8.201 ([104]).
Soit $\mathbb{K}$ un corps. Tout hyperplan de $\mathbb{M}(n,\mathbb{K})$ coupe $\mathrm{GL}(n,\mathbb{K})$.

Démonstration. Soit $\mathcal{H}$ un hyperplan de $\mathbb{M}(n,\mathbb{K})$. Il existe une forme linéaire ϕ sur $\mathbb{M}(n,\mathbb{K})$ telle que $\mathcal{H} = \ker(\phi)$. Encore une fois la proposition 8.199 nous donne $A \in \mathbb{M}(n,\mathbb{K})$ telle que $\phi = f_A$; nous notons r le rang de A. Par le lemme 4.112 nous avons $A = PJ_rQ$ avec $P, Q \in \mathrm{GL}(n,\mathbb{K})$ et

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{8.404}$$

Pour tout $M \in \mathbb{M}(n,\mathbb{K})$ nous avons

$$\phi(M) = \mathrm{Tr}(AM) = \mathrm{Tr}(PJ_rQM) = \mathrm{Tr}(J_rQMP), \tag{8.405}$$

la dernière égalité découlant de l'invariance cyclique de la trace[74]. Ce que nous cherchons est $M \in \mathrm{GL}(n,\mathbb{K})$ telle que $\phi(M) = 0$. Nous commençons par trouver $N \in \mathrm{GL}(n,\mathbb{K})$ telle que $\mathrm{Tr}(J_rN) = 0$. Celle-là est facile : c'est

$$N = \begin{pmatrix} 0 & 1 \\ \mathbb{1}_{n-1} & 0 \end{pmatrix}. \tag{8.406}$$

Les éléments diagonaux de J_rN sont tous nuls. Par conséquent en posant $M = Q^{-1}NP^{-1}$ nous avons notre matrice inversible dans le noyau de ϕ. □

8.12 Diagonalisation et trigonalisation

Ici encore $\mathbb{K}$ est un corps commutatif.

73. Lemme 4.65.
74. Lemme 4.65.

8.12.1 Matrices semblables

DefCQNFooSDhDpB

Proposition-Définition 8.202 (matrices semblables[1])**.**
Nous définissons, sur l'ensemble $\mathbb{M}(n,\mathbb{K})$ des matrices $n\times n$ à coefficients dans $\mathbb{K}$, la relation $A\sim B$ si et seulement si il existe une matrice $P\in \mathrm{GL}(n,\mathbb{K})$ telle que $B=P^{-1}AP$.

Cette relation est une relation d'équivalence.

Deux matrices équivalentes en ce sens sont dites ***semblables****.*

PROPooXGXTooQNAAxw

Proposition-Définition 8.203 (Matrices équivalentes)**.**
Nous définissons, sur l'ensemble $\mathbb{M}(n,\mathbb{K})$ des matrices $n\times n$ à coefficients dans $\mathbb{K}$, la relation $A\sim B$ si et seulement si il existe des matrice inversibles P et Q telles que $A=PBQ^{-1}$.

Cette relation est une relation d'équivalence.

Deux telles matrice sont dite ***équivalentes****.*

PROPooIXFSooZsFWHm

Proposition-Définition 8.204.
Soit un espace vectoriel E. Nous définissons sur $\mathrm{End}(E)$ la relation $u\sim v$ si et seulement si il existe une application inversible $A\colon E\to E$ telle que $v=A^{-1}\circ u\circ A$.

Cette relation est une relation d'équivalence.

Deux endomorphismes équivalents en ce sens sont dits ***semblables****.*

PROPooBGJBooXlDYEv

Proposition 8.205.
Deux applications linéaires sont semblables si et seulement si leurs matrices sont semblables dans toute base.

LEMooKUHJooDfNKeg

Lemme 8.206.
Le polynôme caractéristique[75] *est un invariant sous les similitudes.*

Démonstration. En effet si P est une matrice inversible,

$$\begin{aligned}\chi_{P^{-1}AP} &= \det(P^{-1}AP-\lambda X\mathbb{1}) && (8.407a)\\ &= \det\left(P(P^{-1}AP-\lambda X\mathbb{1})P^{-1}\right) && (8.407b)\\ &= \det(A-\lambda X\mathbb{1}) && (8.407c)\\ &= \chi_A. && (8.407d)\end{aligned}$$

□

La permutation de lignes ou de colonnes ne sont pas des similitudes, comme le montrent les exemples suivants :

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix}\quad B=\begin{pmatrix}2&1\\4&3\end{pmatrix}. \tag{8.408}$$

Nous avons $\chi_A=X^2-5X-2$ tandis que $\chi_B=X^2-5X+2$ alors que le polynôme caractéristique est un invariant de similitude.

8.12.2 Trace de matrices semblables

PROPooRMYQooWkEpJJ

Proposition 8.207 ([1])**.**
Invariance de la trace par similarité.

(1) Soit une application linéaire f. Si la matrice de f dans une base est A et est B dans une autre base, alors

$$\mathrm{Tr}(A)=\mathrm{Tr}(B). \tag{8.409}$$

(2) Deux opérateurs semblables[76] *ont même trace.*

75. Définition 8.113.
76. Opérateurs semblables, définition 8.202

Démonstration. Les matrices A et B sont liées par la proposition 4.119 : $B = Q^{-1}AQ$ où Q est la matrice qui lie les vecteurs des deux bases. L'invariance cyclique de la trace donnée en le lemme 4.65 implique que

$$\mathrm{Tr}(B) = \mathrm{Tr}(Q^{-1}AQ) = \mathrm{Tr}(QQ^{-1}A) = \mathrm{Tr}(A). \tag{8.410}$$

□

LEMooXXEYooKHyQjb

Lemme 8.208.
Soit une matrice A. Nous avons

$$\mathrm{Tr}(A^\dagger A) = \sum_{ij} |A_{ij}|^2. \tag{8.411}$$

Démonstration. Utilisant la proposition 8.178 pour les éléments de la matrice $A^\dagger$, nous avons

$$\mathrm{Tr}(A^\dagger A) = \sum_k (A^\dagger A)_{kk} = \sum_{kl} A^*_{lk}A_{lk} = \sum_{kl} |A_{lk}|^2. \tag{8.412}$$

□

LEMooQXFQooLGPcIt

Lemme 8.209 ([291]).
Si les matrice A et B sont unitairement semblables[77]*, alors*

$$\sum_{ij} |A_{ij}|^2 = \sum_{ij} |B_{ij}|^2. \tag{8.413}$$

Démonstration. Soit une matrice unitaire U telle que $B = UAU^\dagger$. Nous savons par le proposition 8.207 que des matrice semblables ont même trace. Or $B^\dagger B$ est unitairement semblable à $A^\dagger A$ parce que

$$B^\dagger B = UA^\dagger U^\dagger UAU^\dagger = UA^\dagger AU^\dagger. \tag{8.414}$$

Donc $\mathrm{Tr}(A^\dagger A) = \mathrm{Tr}(B^\dagger B)$. Le lemme 8.208 conclut. □

8.12.3 Endomorphismes nilpotents

La **trace** d'une matrice $A \in \mathbb{M}(n, \mathbb{K})$ est la somme de ses éléments diagonaux :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}. \tag{8.415}$$

Une propriété importante est son invariance cyclique.

LemhbZTay

Lemme 8.210.
Quelques propriétés de la trace.

(1) Si A et B sont des matrices carrées, alors $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

(2) La trace est un invariant de similitude.

Démonstration. C'est un simple calcul :

$$\mathrm{Tr}(AB) = \sum_{ik} A_{ik}B_{ki} = \sum_{ik} A_{ki}B_{ik} = \sum_{ik} B_{ik}A_{ki} = \sum_i (BA)_{ii} = \mathrm{Tr}(BA) \tag{8.416}$$

où nous avons simplement renommé les indices $i \leftrightarrow k$.

En particulier, la trace est un invariant de similitude parce que $\mathrm{Tr}(ABA^{-1}) = \mathrm{Tr}(A^{-1}AB) = \mathrm{Tr}(B)$ par l'invariance cyclique démontrée en 4.65(2). □

La trace étant un invariant de similitude, nous pouvons donc définir la **trace** comme étant la trace de sa matrice dans une base quelconque. Si la matrice est diagonalisable, alors la trace est la somme des valeurs propres.

77. Définition 8.202.

LemzgNOjY

Lemme 8.211 ([114]).
L'endomorphisme $u \in \mathrm{End}(\mathbb{C}^n)$ est nilpotent si et seulement si $\mathrm{Tr}(u^p) = 0$ pour tout p.

Démonstration. Supposons que u est nilpotent. Alors ses valeurs propres sont toutes nulles et celles de u^p le sont également. La trace étant la somme des valeurs propres, nous avons alors tout de suite $\mathrm{Tr}(u^p) = 0$.

Supposons maintenant que $\mathrm{Tr}(u^p) = 0$ pour tout p. Le polynôme caractéristique (8.244) est

EqfnCqWq
$$\chi_u = (-1)^n X^\alpha (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_r)^{\alpha_r}. \tag{8.417}$$

où les λ_i $(i = 1, \dots, r)$ sont les valeurs propres non nulles distinctes de u.

Il est vite vu que le coefficient de X^{n-1} dans χ_u est $-\mathrm{Tr}(u)$ parce que le coefficient de X^{n-1} se calcule en prenant tous les X sauf une fois $-\lambda_i$. D'autre part le polynôme caractéristique de u^p est le même que celui de u, en remplaçant λ_i par λ_i^p ; cela est dû au fait que si v est vecteur propre de valeur propre λ, alors $u^p v = \lambda^p v$.

Par l'équation (8.417), nous voyons que le coefficient du terme X^{n-1} dans le polynôme caractéristique est

eqSoDSKH
$$0 = \mathrm{Tr}(u^p) = \alpha_1 \lambda_1^p + \dots + \alpha_r \lambda_r^p. \tag{8.418}$$

Donc les nombres $(\alpha_1, \dots, \alpha_r)$ sont une solution non triviale [78] du système EqDpvTnu

$$\begin{cases} \lambda_1 X_1 + \dots + \lambda_r X_r = 0 & \text{(8.419a)} \\ \quad \vdots & \text{(8.419b)} \\ \lambda_1^r X_1 + \dots + \lambda_r^r X_r = 0. & \text{(8.419c)} \end{cases}$$

Ce sont les équations (8.418) écrites pour $p = 1, \dots, r$. Le déterminant de ce système est

$$\lambda_1 \dots \lambda_r \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^{r-1} & \dots & \lambda_r^{r-1} \end{pmatrix} \neq 0, \tag{8.420}$$

qui est un déterminant de Vandermonde (proposition 8.18) valant

$$0 = \lambda_1 \dots \lambda_r \prod_{1 \leqslant i \leqslant j \leqslant r} (\lambda_i - \lambda_j). \tag{8.421}$$

Étant donné que les λ_i sont distincts et non nuls, nous avons une contradiction et nous devons conclure que $(\alpha_1, \dots, \alpha_r)$ était une solution triviale du système (8.419). □

PropMWWJooVIXdJp

Proposition 8.212 ([292]).
Soit un $\mathbb{K}$-espace vectoriel E. Un endomorphisme $u \in \mathrm{End}(E)$ est nilpotent si et seulement si il existe une base de E dans laquelle la matrice de u est strictement triangulaire supérieure.

Démonstration. (1) $\Rightarrow$ Nous faisons la démonstration par récurrence sur la dimension de E.

Lorsque $n = 1$ nous avons $u = (a)$ avec $a \in \mathbb{K}$. Puisque $a^k = 0$ pour un certain k nous avons $a = 0$ parce qu'un corps est toujours un anneau intègre [79].

Lorsque $\dim(E) = n$ nous savons que u a un noyau non réduit au vecteur nul (parce qu'il est nilpotent). Soit donc un vecteur non nul $x \in \ker(u)$ et une base

$$\{x, e_2, \dots, e_n\} \tag{8.422}$$

78. Si $\alpha_1 = \dots = \alpha_r = 0$, alors les valeurs propres sont toutes nulles et la matrice est en réalité nulle dès le départ.
79. Lemme 1.198.

donnée par le théorème de la base incomplète 4.14. La matrice de u dans cette base s'écrit

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A & \\ 0 & & & \end{array}\right). \tag{8.423}$$

Un tout petit peu de calcul de produit de matrice montre que la matrice de u^k est de la forme

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A^k & \\ 0 & & & \end{array}\right). \tag{8.424}$$

Étant donné que l'endomorphisme u est nilpotent, la matrice A l'est aussi. L'hypothèse de récurrence dit alors que A est strictement triangulaire supérieure (ou en tout cas peut le devenir par un changement de base adéquat).

(2) $\Leftarrow$ Soit une base $\{e_1, \dots, e_n\}$ dans laquelle la matrice de u est strictement triangulaire supérieure.

Alors $u(e_1) = 0$ et plus généralement, $u(e_k) \in \text{Span}\{e_1, \dots, e_{k-1}\}$. Voyez par récurrence que $u^l(e_k) \in \text{Span}\{e_1, \dots, e_{k-l}\}$. Donc $u^l(e_k) = 0$ dès que $l \geqslant k$.

□

LEMooKPWKooOacXju

Lemme 8.213 ([1])**.**
Si $N \in \mathbb{M}(n, \mathbb{C})$ est une matrice nilpotente d'ordre de nilpotence r, alors $\{N^k\}_{k=0,\dots,r-1}$ est libre dans $\mathbb{M}(n, \mathbb{C})$.

PROPooWTFWooXHlmhp

Proposition 8.214 (Thème 41)**.**
Soit E un espace de Banach (espace vectoriel normé complet[80]). Si $A \in \mathcal{L}(E, E)$ est nilpotente, alors $(\mathbb{1} - A)$ est inversible et son inverse est donné par

$$(\mathbb{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k, \tag{8.425}$$

où l'infini peut évidemment être remplacé par l'ordre de nilpotence de A.

Démonstration. En ce qui concerne la convergence de la somme, elle ne fait pas de doute parce que A étant nilpotente, la somme contient seulement une quantité finie de termes non nuls.

Montrons à présent que la somme est l'inverse de $\mathbb{1} - A$ en multipliant terme à terme :

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n} (A^k - A^{k+1}) = \mathbb{1} - A^{n+1}. \tag{8.426}$$

Par conséquent

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \to 0. \tag{8.427}$$

La dernière limite est en réalité une égalité pour n assez grand. □

8.12.4 Endomorphismes diagonalisables

DefCNJqsmo

Définition 8.215.
*Une matrice est **diagonalisable** si elle est semblable[81] à une matrice diagonale.*

80. Définition 7.265.
81. Définition 8.202.

Une application linéaire est diagonalisable si elle est semblable[82] *à une application linéaire diagonale.*

Autrement dit, A est diagonalisable si il existe un opérateur diagonal D et un opérateur inversible P tels que $PAP^{-1} = D$.

La proposition 8.205 nous assure que la notion de diagonalisabilité pour les matrices et pour les applications sont les mêmes.

PROPooDEETooSOMiGO

Proposition 8.216.
Si A est un opérateur diagonalisable dont les valeurs propres sont λ_i, alors il existe un opérateur inversible Q tel que

$$A = Q^{-1}DQ \tag{8.428}$$

où D est l'opérateur diagonal contenant les λ_i sur sa diagonale.

Lemme 8.217.
Une matrice triangulaire supérieure avec des 1 sur la diagonale n'est diagonalisable que si elle est diagonale (c'est-à-dire si c'est la matrice unité).

Démonstration. Si A est une matrice triangulaire supérieure de taille n telle que $A_{ii} = 1$, alors $\det(A - \lambda\mathbb{1}) = (1 - \lambda)^n$, ce qui signifie que $\text{Spec}(A) = \{1\}$. Pour la diagonaliser, il faudrait une matrice $P \in \text{GL}(n, \mathbb{K})$ telle que $\mathbb{1} = P^{-1}AP$, ce qui est uniquement possible si $A = \mathbb{1}$. □

LemgnaEOk

Lemme 8.218.
Soit F un sous-espace stable par u. Soit une décomposition du polynôme minimal

$$\mu_u = P_1^{n_1} \dots P_r^{n_r} \tag{8.429}$$

où les P_i sont des polynômes irréductibles unitaires distincts. Si nous posons $E_i = \ker P_i^{n_i}$, alors

$$F = (F \cap E_1) \oplus \dots \oplus (F \cap E_r). \tag{8.430}$$

ThoDigLEQEXR

Théorème 8.219.
Soit E, un espace vectoriel de dimension n sur le corps commutatif $\mathbb{K}$ et $u \in \text{End}(E)$. Les propriétés suivantes sont équivalentes.

ItemThoDigLEQEXRiv

(1) L'endomorphisme u est diagonalisable.

ItemThoDigLEQEXRi

(2) Il existe un polynôme $P \in \mathbb{K}[X]$ non constant, scindé sur $\mathbb{K}$, dont toutes les racines sont simples, tel que $P(u) = 0$.

ItemThoDigLEQEXRii

(3) Le polynôme minimal μ_u est scindé sur $\mathbb{K}$ et toutes ses racines sont simples[83].

ItemThoDigLEQEXRiii

(4) Tout sous-espace de E possède un supplémentaire stable par u.

ITEMooZNJFooEiqDYp

(5) Dans une base adaptée, la matrice de u est diagonale et les éléments diagonaux sont ses valeurs propres.

Démonstration. Plein d'implications à prouver.

(1) **(2) implique (3)** Étant donné que $P(u) = 0$, il est dans l'idéal des polynômes annulateurs de u, et le polynôme minimal μ_u le divise parce que l'idéal des polynômes annulateurs est généré par μ_u par le théorème 6.43.

82. Définition 8.204.

83. Le polynôme *caractéristique*, lui, n'a pas spécialement ses racines simples ; il peut encore être de la forme

$$\chi_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}, \tag{8.431}$$

mais alors $\dim(E_{\lambda_i}) = \alpha_i$.

(2) **(3) implique (1)** Étant donné que le polynôme minimal est scindé à racines simples, il s'écrit sous forme de produits de monômes tous distincts, c'est-à-dire

$$\mu_u(X) = (X - \lambda_1)\dots(X - \lambda_r) \tag{8.432}$$

où les λ_i sont des éléments distincts de $\mathbb{K}$. Étant donné que $\mu_u(u) = 0$, le théorème de décomposition des noyaux (théorème 8.92) nous enseigne que

$$E = \ker(u - \lambda_1) \oplus \dots \oplus \ker(u - \lambda_r). \tag{8.433}$$

Mais $\ker(u - \lambda_i)$ est l'espace propre $E_{\lambda_i}(u)$. Donc u est diagonalisable.

(3) **(1) implique (4)** Soit $\{e_1, \dots, e_n\}$ une base qui diagonalise u, soit F un sous-espace de E et $\{f_1, \dots, f_r\}$ une base de F. Par le théorème 4.18(2), nous pouvons compléter la base de F par des éléments de la base $\{e_i\}$. Le complément ainsi construit est stable par u.

(4) **(4) implique (1)** En dimension un, tout endomorphisme est diagonalisable, nous supposons donc que $\dim E = n \geqslant 2$. Nous procédons par récurrence sur le nombre de vecteurs propres connus de u. Supposons avoir déjà trouvé p vecteurs propres $e_1, \dots, e_p$ de u. Considérons H, un hyperplan qui contient les vecteurs $e_1, \dots, e_p$. Soit F un supplémentaire de H stable par u ; par construction $\dim F = 1$ et si $e_{p+1} \in F$, il doit être vecteur propre de u.

(5) **(1) implique (2)** Nous supposons maintenant que u est diagonalisable. Soient $\lambda_1, \dots, \lambda_r$ les valeurs propres deux à deux distinctes, et considérons le polynôme

$$P(x) = (X - \lambda_1)\dots(X - \lambda_r). \tag{8.434}$$

Alors $P(u) = 0$. En effet si e_i est un vecteur propre pour la valeur propre λ_i,

$$P(u)e_i = \prod_{j \neq i}(u - \lambda_j) \circ (u - \lambda_i)e_i = 0 \tag{8.435}$$

par le lemme 8.91. Par conséquent $P(u)$ s'annule sur la base $\{e_i\}$.

(6) **(5) implique (2)** Si la matrice A est diagonale alors le polynôme $P = \prod_{i=1}^{n}(X - A_{ii})$ est annulateur de A. En effet,

$$P(A)e_k = \prod_{i=1}^{n}(A - A_{ii})x = \prod_{i=1}^{n}\big(u(e_k) - A_{ii}e_k\big) = \prod_{i=1}^{n}\big(A_{kk}e_k - A_{ii}e_k\big) = 0 \tag{8.436}$$

parce que le facteur $i = k$ est nul.

(7) **(3) implique (5)** le polynôme minimal de u s'écrit

$$\mu = (X - \lambda_1)\dots(X - \lambda_r), \tag{8.437}$$

et les espaces E_i du lemme 8.218 sont les espaces propres $E_i = \ker(u - \lambda_i)$. Nous avons donc une somme directe

$$E = E_1 \oplus \dots \oplus E_r. \tag{8.438}$$

Dans chacun des espaces propres, u a une matrice diagonale avec la valeur propre correspondante sur la diagonale. Une base de E constituée d'une base de chacun des espaces propres est donc une base comme nous en cherchons.

□

CorQeVqsS

Corolaire 8.220.
Si u est diagonalisable et si F est un sous-espace stable par u, alors

$$F = \bigoplus_{\lambda} E_\lambda(u) \cap F \tag{8.439}$$

où $E_\lambda(u)$ est l'espace propre de u pour la valeur propre λ. En particulier la restriction de u à F, $u|_F$ est diagonalisable.

Démonstration. Par le théorème 8.219, le polynôme μ_u est scindé et ne possède que des racines simples. Notons le

$$\mu_u(X) = (X - \lambda_1)\dots(X - \lambda_r). \tag{8.440}$$

Les espaces E_i du lemme 8.218 sont maintenant les espaces propres.

En ce qui concerne la diagonalisabilité de $u|_F$, notons que nous avons une base de F composée de vecteurs dans les espaces $E_\lambda(u)$. Cette base de F est une base de vecteurs propres de u. □

Lemme 8.221.
Soient E un $\mathbb{K}$-espace vectoriel et $u \in \mathrm{End}(E)$. Si $\mathrm{Card}\big(\mathrm{Spec}(u)\big) = \dim(E)$ alors u est diagonalisable.

Démonstration. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres distinctes de u. Nous savons que les espaces propres correspondants sont en somme directe (lemme 8.90). Par conséquent $\mathrm{Span}\{E_{\lambda_i}(u)\}$ est de dimension $n = \dim(E)$ et u est diagonalisable. □

Voici un résultat de diagonalisation simultanée. Nous donnerons un résultat de trigonalisation simultanée dans le lemme 12.105.

PropGqhAMei

Proposition 8.222 (Diagonalisation simultanée)**.**
Soit $(u_i)_{i\in I}$ une famille d'endomorphismes qui commutent deux à deux.

ItemGqhAMei

(1) Si $i, j \in I$ alors tout sous-espace propre de u_i est stable par u_j. Autrement dit $u_j\big(E_\lambda(u_i)\big) \subset E_\lambda(u_i)$.

(2) Si les u_i sont diagonalisables, alors ils le sont simultanément.

Démonstration. Supposons que u_i et u_j commutent et soit x un vecteur propre de u_i : $u_i(x) = \lambda x$. Nous montrons que $u_j(x) \in E_\lambda(u_i)$. Nous avons

$$u_i\big(u_j(x)\big) = u_j\big(u_i(x)\big) = \lambda u_j(x). \tag{8.441}$$

Par conséquent $u_j(x)$ est vecteur propre de u_i de valeur propre λ.

Montrons maintenant l'affirmation à propos des endomorphismes simultanément diagonalisables. Si $\dim E = 1$, le résultat est évident. Nous supposons également qu'aucun des u_i n'est multiple de l'identité. Nous effectuons une récurrence sur la dimension.

Soit u_0 un des u_i et considérons ses valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_r$. Pour chaque k nous avons

$$E_{\lambda_k}(u_0) \neq E, \tag{8.442}$$

sinon u_0 serait un multiple de l'identité. Par contre le fait que u_0 soit diagonalisable permet de décomposer E en espaces propres de u_0 :

$$E = \bigoplus_k E_{\lambda_k}(u_0). \tag{8.443}$$

Ce que nous allons faire est de simultanément diagonaliser les $(u_i)_{i\in I}$ sur chacun des E_{λ_k} séparément. Par le point (1), nous avons $u_i\colon E_{\lambda_k}(u_0) \to E_{\lambda_k}(u_0)$, et nous pouvons considérer la famille d'opérateurs

$$\Big(u_i|_{E_{\lambda_k}(u_0)}\Big)_{i\in I}. \tag{8.444}$$

Ce sont tous des opérateurs qui commutent et qui agissent sur un espace de dimension plus petite. Par hypothèse de récurrence nous avons une base de $E_{\lambda_k}(u_0)$ qui diagonalise tous les u_i. □

ExewINgYo

Exemple 8.223.
Soit un espace vectoriel sur un corps $\mathbb{K}$. Un opérateur **involutif** est un opérateur différent de l'identité dont le carré est l'identité. Typiquement une symétrie orthogonale dans $\mathbb{R}^3$. Le polynôme caractéristique d'une involution est $X^2 - 1 = (X + 1)(X - 1)$.

Tant que $1 \neq -1$, X^2-1 est donc scindé à racines simples et les involutions sont diagonalisables (8.219). Cependant si le corps est de caractéristique 2, alors $X^2-1=(X+1)^2$ et l'involution n'est plus diagonalisable.

Par exemple si le corps est de caractéristique 2, nous avons

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix} \tag{8.445a}$$

$$A^2=\begin{pmatrix}1&2\\0&1\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}. \tag{8.445b}$$

Cette matrice A représente donc une involution, mais n'est pas diagonalisable. △

8.12.5 Diagonalisation : cas complexe, pas toujours

Il n'est pas vrai qu'une matrice de $\mathbb{M}(n,\mathbb{C})$ soit toujours diagonalisable. En effet le théorème 8.219(3) dit qu'une matrice est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples. Certes sur $\mathbb{C}$ le polynôme minimal sera scindé, mais il ne sera pas spécialement à racines simples.

Exemple 8.224.
La matrice

$$A=\begin{pmatrix}0&1\\0&0\end{pmatrix} \tag{8.446}$$

a pour polynôme caractéristique $\chi_A(X)=X^2$. C'est également son polynôme minimal, qui n'est pas à racine simple.

Il est par ailleurs facile de voir que le seul espace propre de A est $\mathrm{Span}\{(1,0)\}$ (ici le span est sur $\mathbb{C}$). Donc l'espace $\mathbb{C}^2$ ne possède pas de base de vecteurs propres de A. △

Ce qui est vrai, c'est que le polynôme caractéristique a des racines, et que ces racines correspondent à des vecteurs propres. Mais il n'y a pas toujours autant de vecteurs propres que la multiplicité des racines.

8.225.
Lorsque la diagonalisation n'est pas possible, il est souvent possible de trigonaliser. Les matrices triangulaires ne sont pas aussi faciles à manipuler que les matrices diagonales, mais c'est toujours ça de pris.

Nous étudierons ça plus tard, en 12.10.1 parce que ça va nécéssiter le théorème de d'Alembert.

8.12.6 Diagonalisation : cas réel

LemSchureRelnrqfiy

Lemme 8.226 (Lemme de Schur réel)**.**
Soit $A\in\mathbb{M}(n,\mathbb{R})$. Il existe une matrice orthogonale Q telle que $Q^{-1}AQ$ soit de la forme

$$Q^{-1}AQ=\begin{pmatrix}\lambda_1&*&*&*&*\\0&\ddots&\ddots&\ddots&\vdots\\0&0&\lambda_r&*&*\\0&0&0&\begin{pmatrix}a_1&b_1\\c_1&d_1\end{pmatrix}&*\\0&0&0&0&\begin{pmatrix}a_s&b_s\\c_s&d_s\end{pmatrix}\end{pmatrix}. \tag{8.447}$$

EqMtrTSqRTA

Le déterminant de A est le produit des déterminants des blocs diagonaux et les valeurs propres de A sont les $\lambda_1,\dots,\lambda_r$ et celles de ces blocs.

Démonstration. Si la matrice A a des valeurs propres réelles, nous procédons comme dans le cas complexe. Cela nous fournit le partie véritablement triangulaire avec les valeurs propres $\lambda_1, \dots, \lambda_r$ sur la diagonale. Supposons donc que A n'a pas de valeurs propres réelles. Soit donc $\alpha + i\beta$ une valeur propre ($\beta \neq 0$) et $u + iv$ un vecteur propre correspondant où u et v sont des vecteurs réels. Nous avons

$$Au + iAv = A(u + iv) = (\alpha + i\beta)(u + iv) = \alpha u - \beta v + i(\alpha v + \beta u), \tag{8.448}$$

et en égalisant les parties réelles et imaginaires,

$$Au = \alpha u - \beta v \tag{8.449a}$$
$$Av = \alpha v + \beta u. \tag{8.449b}$$

Sur ces relations nous voyons que ni u ni v ne sont nuls. De plus u et v sont linéairement indépendants (sur $\mathbb{R}$), en effet si $v = \lambda u$ nous aurions $Au = \alpha u - \beta\lambda u = (\alpha - \beta\lambda)u$, ce qui serait une valeur propre réelle alors que nous avions supposé avoir déjà épuisé toutes les valeurs propres réelles.

Étant donné que u et v sont deux vecteurs réels non nuls et linéairement indépendants, nous pouvons trouver une base orthonormée $\{q_1, q_2\}$ de $\text{Span}\{u, v\}$. Nous pouvons étendre ces deux vecteurs en une base orthonormée $\{q_1, q_2, q_3, \dots, q_n\}$ de $\mathbb{R}^n$. Nous considérons à présent la matrice orthogonale dont les colonnes sont formées de ces vecteurs : $Q = [q_1\, q_2\, \dots\, q_n]$.

L'espace $\text{Span}\{e_1, e_2\}$ est stable par $Q^{-1}AQ$, en effet nous avons

$$Q^{-1}AQe_1 = Q^{-1}Aq_1 = Q^{-1}(aq_1 + bq_2) = ae_1 + be_2. \tag{8.450}$$

La matrice $Q^{-1}AQ$ est donc de la forme

$$Q^{-1}AQ = \begin{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} & C_1 \\ 0 & A_1 \end{pmatrix} \tag{8.451}$$

où C_1 est une matrice réelle $2 \times (n-1)$ quelconque et A_1 est une matrice réelle $(n-2) \times (n-2)$. Nous pouvons appliquer une récurrence sur la dimension pour poursuivre.

Notons que si A n'a pas de valeurs propres réelles, elle est automatiquement d'ordre pair parce que les valeurs propres complexes viennent par couple complexes conjugués.

En ce qui concerne les valeurs propres, il est facile de voir en regardant (8.447) que les valeurs propres sont celles des blocs diagonaux. Étant donné que $Q^{-1}AQ$ et A ont même polynôme caractéristique, ce sont les valeurs propres de A. □

ThoeTMXla

Théorème 8.227 (Théorème spectral, matrice symétrique[104])**.**
Une matrice symétrique réelle,

ITEMooJWHLooSfhNSW

(1) a un spectre contenu dans $\mathbb{R}$

ITEMooMWWRooXxGONW

(2) est diagonalisable par une matrice orthogonale.

Si M est une matrice symétrique réelle alors $\mathbb{R}^n$ possède une base orthonormée de vecteurs propres de M.

Démonstration. Soit A une matrice réelle symétrique. Elle agit sur l'espace $\mathbb{C}^n$ par la définition 4.68, et en particulier la formule 4.84. Nous munissons de plus $\mathbb{C}^n$ de la forme sesquilinéaire définie en la proposition 8.183.

Si λ est une valeur propre complexe pour le vecteur propre complexe v, alors d'une part $\langle Av, v\rangle = \lambda\langle v, v\rangle$ et d'autre part $\langle Av, v\rangle = \langle v, Av\rangle = \bar{\lambda}\langle v, v\rangle$. Par conséquent $\lambda = \bar{\lambda}$, et λ est réelle.

Le lemme de Schur réel 8.226 donne une matrice orthogonale Q qui trigonalise A. Les valeurs propres étant toutes réelles, la matrice $Q^{-1}AQ$ est même triangulaire (il n'y a pas de blocs dans la forme (8.447)). Prouvons que $Q^{-1}AQ$ est symétrique :

$$(Q^{-1}AQ)^t = Q^tA^t(Q^{-1})^t = Q^{-1}A^tQ = Q^{-1}AQ \tag{8.452}$$

où nous avons utilisé le fait que Q était orthogonale ($Q^{-1}=Q^t$) et que A était symétrique ($A^t=A$). Une matrice triangulaire supérieure symétrique est obligatoirement une matrice diagonale.

En ce qui concerne la base de vecteurs propres, soit $\{e_i\}_{i=1,\dots,n}$ la base canonique de $\mathbb{R}^n$ et Q une matrice orthogonale telle que $A=Q^tDQ$ avec D diagonale. Nous posons $f_i=Q^te_i$ et en tenant compte du fait que $Q^t=Q^{-1}$ nous avons $Af_i=Q^tDQQ^te_i=Q^t\lambda_ie_i=\lambda_if_i$. Donc les f_i sont des vecteurs propres de A. De plus ils sont orthonormés parce qu'en utilisant la proposition 8.186,

$$\langle f_i,f_j\rangle=\langle Q^te_i,Q^te_j\rangle=\langle e_i,Q^tQe_j\rangle=\langle e_i,e_j\rangle=\delta_{ij}. \tag{8.453}$$

□

Le théorème spectral pour les opérateurs autoadjoints sera traité plus bas parce qu'il a besoin de notions sur les formes bilinéaires, théorème 10.8.

RemGKDZfxu

Remarque 8.228.
Une matrice symétrique est diagonalisable par une matrice orthogonale. Nous pouvons en réalité nous arranger pour diagonaliser par une matrice de $\mathrm{SO}(n)$. Plus généralement si A est une matrice diagonalisable par une matrice $P\in\mathrm{GL}^+(n,\mathbb{R})$ alors elle est diagonalisable par une matrice de $\mathrm{GL}^-(n,\mathbb{R})$ en changeant le signe de la première ligne de P. Et inversement.

En effet, si nous avons $P^tDP=A$, alors en notant $*$ les quantités qui ne dépendent pas de a, b ou c,

$$\begin{aligned}\begin{pmatrix}a&*&*\\b&*&*\\c&*&*\end{pmatrix}\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix}\begin{pmatrix}a&b&c\\*&*&*\\*&*&*\end{pmatrix}&=\begin{pmatrix}a&*&*\\b&*&*\\c&*&*\end{pmatrix}\begin{pmatrix}\lambda_1a&\lambda_1b&\lambda_1c\\*&*&*\\*&*&*\end{pmatrix}\\&=\begin{pmatrix}\lambda_1a^2+*&\lambda_1ab+*&\lambda_1ac+*\\\dots&\dots&\dots\\\dots&\dots&\dots\end{pmatrix}.\end{aligned} \tag{8.454}$$

Nous voyons donc que si nous changeons les signes de a, b et c en même temps, le résultat ne change pas.

PROPooQHHPooSqpgcb

Proposition 8.229.
Une forme bilinéaire est non-dénénérée[84] *si et seulement si sa matrice associée est inversible.*

Démonstration. Nous savons que la matrice associée est symétrique et qu'elle peut donc être diagonalisée (théorème 8.227). En nous plaçant dans une base de diagonalisation, nous devons prouver que la forme est non-dégénérée si et seulement si les éléments diagonaux de la matrice sont tous non nuls.

Écrivons $b(x,z)$ en choisissant pour z le vecteur de base e_k de composantes $(e_k)_j=\delta_{kj}$:

$$b(x,e_k)=\sum_{ij}x_i(e_k)_j=\sum_i b_{ik}x_i=b_{kk}x_k. \tag{8.455}$$

Si b est dégénérée et si x est un vecteur non nul (disons que la composante x_i est non nulle) de E tel que $b(x,z)=0$ pour tout $z\in E$, alors $b_{ii}=0$, ce qui montre que la matrice de b n'est pas inversible.

Réciproquement si la matrice de b est inversible, alors tous les b_{kk} sont différents de zéro, et le seul vecteur x tel que $b_{kk}x_k=0$ pour tout k est le vecteur nul. □

8.12.7 Matrice définie positive

DefAWAooCMPuVM

Définition 8.230 (Matrice définie positive, opérateur défini positif)**.**
Un opérateur sur un espace vectoriel sur $\mathbb{C}$ ou $\mathbb{R}$ est ***défini positif*** *si toutes ses valeurs propres*

84. Définition 8.129.

sont réelles et strictement positives. Il est ***semi-défini positif*** *si ses valeurs propres sont réelles positives ou nulles.*

Mêmes définitions pour une matrice.

Afin d'éviter l'une ou l'autre confusion, nous disons souvent *strictement* défini positif pour positif.

NORMooAJLHooQhwpvr

8.231.
Quelques ensembles de matrices symétriques.

(1) $S(n,\mathbb{R})$ est l'ensemble des matrices symétriques réelles $n\times n$.
(2) $S^+(n,\mathbb{R})$ l'ensemble des matrices réelles symétriques $n\times n$ définies positives.
(3) $S^{++}(n,\mathbb{R})$ le sous-ensemble de $S^+(n,\mathbb{R})$ des matrices strictement définies positives.

Remarque 8.232.
Nous ne définissons pas la notion de matrice définie positive pour une matrice non symétrique.

PropcnJyXZ

Proposition 8.233.
Soit M, une matrice symétrique. Nous avons

ITEMooTJVQooYmRkas

(1) $\det(M)>0$ *et* $\mathrm{Tr}(M)>0$ *implique M définie positive*[85],
(2) $\det(M)>0$ *et* $\mathrm{Tr}(M)<0$ *implique M définie négative,*

ItemluuFPN

(3) $\det(M)<0$ *implique ni semi-définie positive, ni définie négative*
(4) $\det(M)=0$ *implique M semi-définie positive ou semi-définie négative.*

Lorsqu'un énoncé parle d'une matrice symétrique, le premier réflexe est de la diagonaliser : considérer une matrice orthogonale Q telle que $Q^tMQ=D$ avec D diagonale. Et les valeurs propres sur la diagonale : $D_{kk}=\lambda_k$. Les matrices symétriques définies positives ont cependant des propriétés même en dehors de leur base de diagonalisation.

Pour rappel, $\langle x,y\rangle$ est le produit scalaire dans $\mathbb{R}^n$ défini par la proposition 8.173.

LemWZFSooYvksjw

Lemme 8.234.
Soit une matrice symétrique M.

ITEMooSKRAooOgHbGA

(1) Elle est strictement définie positive si et seulement si $\langle x,Mx\rangle>0$ pour tout x non nul dans $\mathbb{R}^n$.

ITEMooMOZYooWcrewZ

(2) Elle est semi-définie positive si et seulement si $\langle x,Mx\rangle\geqslant 0$ pour tout x non nul dans $\mathbb{R}^n$.

ITEMooRRMFooHSOHxZ

(3) Si elle est seulement définie positive, alors $\langle x,Mx\rangle\geqslant\lambda\|x\|^2$ dès que $\lambda\geqslant 0$ minore toutes les valeurs propres.

Démonstration. Démonstration en trois parties.

(1) **(1)** Soit $\{e_i\}_{i=1,\dots,n}$ une base orthonormée de vecteurs propres de M dont l'existence est assurée par le théorème spectral 8.227. Nous nommons x_i les coordonnées de x dans cette base. Alors,

$$\langle x,Mx\rangle=\sum_{i,j}x_i\langle e_i,x_jMe_j\rangle=\sum_{i,j}x_ix_j\langle e_i,\lambda_je_j\rangle=\sum_{ij}x_ix_j\lambda_j\delta_{ij}=\sum_i\lambda_ix_i^2 \tag{8.456}$$

où les λ_i sont les valeurs propres de M. Le produit$\langle x,Mx\rangle$ est strictement positif pour tout x si et seulement si tous les λ_i sont strictement positifs.

(2) **(2)** Nous avons encore

$$\langle x,Mx\rangle=\sum_i\lambda_ix_i^2 \tag{8.457}$$

qui est plus grand ou égal à zéro si et seulement si tous les λ_i sont plus grands ou égaux à zéro.

85. Définition 8.230.

(3) **(3)**

Soit une matrice orthogonale T diagonalisant M, c'est-à-dire telle que $T^tMT = D$ avec D diagonale. Nous allons vérifier que si $\lambda \leqslant \min\{\lambda_i\}$, alors

$$\langle Tx, MTx\rangle \geqslant \lambda\|Tx\|^2 \tag{8.458}$$

pour tout x. Si nous considérons la base de diagonalisation $\{e_k\}$ pour les valeurs propres λ_k, nous avons le calcul

$$\begin{aligned}
\langle Tx, MTx\rangle &= \langle x, T^tMTx\rangle && (8.459a)\\
&= \langle x, Dx\rangle && (8.459b)\\
&= \sum_k \langle x, x_kDe_k\rangle && (8.459c)\\
&= \sum_k \lambda_k x_k \underbrace{\langle x, e_k\rangle}_{=x_k} && (8.459d)\\
&\geqslant \sum_k \lambda|x_k|^2 \qquad \text{en posant} \lambda = \min\{\lambda_i\} && (8.459e)
\end{aligned}$$

Nous avons donc

$$\langle Tx, MTx\rangle \geqslant \sum_k \lambda|x_k|^2 = \lambda\|x\|^2 = \lambda\|Tx\|^2. \tag{8.460}$$

Au dernier passage nous avons utilisé le fait que T est une isométrie (proposition 8.46). L'inéquation (8.458) est démontrée.

Comme T est une bijection [86], cela implique le résultat pour tout x.

□

Les personnes qui aiment les vecteurs lignes et colonnes écriront des inégalités comme

$$x^tMx \geqslant x^tx. \tag{8.461}$$

Tout à l'autre bout du spectre des personnes névrosées des notations, on trouvera des inégalités comme

$$M(x \otimes x) \geqslant x \cdot x. \tag{8.462}$$

Le penchant personnel de l'auteur de ces lignes est la notation avec le produit tensoriel. Si vous aimez ça, vous pouvez lire la section 10.12.15 et en particulier ce qui suit (10.671).

La notation adoptée ici avec le produit scalaire $\langle x, Mx\rangle$, qui peut aussi être écrite $x \cdot Mx$ est entre les deux. Elle a l'avantage de n'être pas technologique comme le produit tensoriel (si vous y mettez les pieds, vous devez savoir ce que vous faites), tout en évitant de se casser la tête à savoir qui est un vecteur ligne ou un vecteur colonne.

Proposition 8.235.
Une application bilinéaire est définie positive [87] *si et seulement si sa matrice symétrique associée* [88] *l'est.*

Démonstration. La définition 8.126 dit que b est strictement définie positive lorsque $b(x,x) \geqslant 0$ et $b(x,x) = 0$ si et seulement si $x = 0$.

D'autre part, le lemme 8.234 dit que la matrice B est strictement définie positive lorsque $x \cdot Bx \geqslant 0$ et $x \cdot Bx = 0$ si et seulement si $x = 0$.

Le lien entre les deux est que le lemme 8.149 nous enseigne que pour tout x et y,

$$b(x,y) = x \cdot By \tag{8.463}$$

où B est la matrice de b. □

86. Une matrice orthogonale a un déterminant qui vaut ± 1.
87. Définition 8.126.
88. Par la proposition 8.142.

PROPooCIEUooODqfwm

Proposition 8.236.
Soit une forme quadratique $q: E \to \mathbb{K}$ et sa matrice[89] $(q_{ij}) \in \mathbb{M}(n, \mathbb{K})$. *Nous avons*

$$q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j = \sum_{i=1}^{n} q_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} q_{ij} x_i x_j. \tag{8.464}$$

EQooAJTUooBSAhKv

Démonstration. La première égalité est seulement l'expression de la matrice de q. Pour la seconde, on découpe la somme sur j en trois parties : $j < i$, $j = i$ et $j > i$:

$$q(x) = \sum_{i} \sum_{j<i} q_{ij} x_i x_j + \sum_{i} q_{ii} x_i^2 + \sum_{i} \sum_{j>i} q_{ij} x_i x_j. \tag{8.465}$$

EQooMKMDooGKqbJi

Nous pouvons utiliser le lemme 1.311 en posant $A = \{(i,j) \text{ tel que } j < i\}$, $B = \{(i,j) \text{ tel que } j > i\}$, et en considérant la bijection

$$\begin{aligned} \varphi: A &\to b \\ (i,j) &\mapsto (j,i). \end{aligned} \tag{8.466}$$

Nous avons alors

$$\sum_{(i,j)\in A} q_{ij} x_i x_j = \sum_{(i,j)\in B} q_{ji} x_j x_i = \sum_{(i,j)\in B} q_{ij} x_i x_j. \tag{8.467}$$

La première égalité est le lemme et la seconde est la symétrie de q. Bref, dans (8.465), le premier et le dernier termes sont égaux. □

8.237.
De nombreux auteurs préfèrent écrire des choses comme $x^t By$ ou $xB^t y$ ou xBy^t et se poser de longues questions sur qui est un « vecteur colonne » et qui est un « vecteur ligne », et si la matrice B soit être transposée ou non. Toutes ces notations servent(?) à cacher un bête produit scalaire.

8.238.
Notons que la matrice associée à une forme bilinéaire (ou quadratique associée) est uniquement valable pour une base donnée. Si nous changeons de base, la matrice change. Cependant lorsque nous travaillons sur $\mathbb{R}^n$, la base canonique est tellement canonique que nous allons nous permettre de parler de « la » matrice associée à une forme bilinéaire.

Corolaire 8.239.
Une matrice symétrique strictement définie positive est inversible.

Démonstration. Si $Ax = 0$ alors $\langle Ax, x\rangle = 0$. Mais dans le cas d'une matrice strictement définie positive, cela implique $x = 0$ par le lemme 8.234. □

Lemme 8.240.
Pour une base quelconque, les éléments diagonaux d'une matrice symétrique semi-définie positive sont positifs. Si la matrice est strictement définie positive, alors les éléments diagonaux sont strictement positifs.

Démonstration. Il s'agit d'une application du lemme 8.234. Si A est définie positive et que $\{e_i\}$ est une base, alors

$$A_{ii} = \langle Ae_i, e_i\rangle \geqslant \lambda \|e_i\|^2 = \lambda \geqslant 0. \tag{8.468}$$

Si A est strictement définie positive, alors λ peut être choisi strictement positif. □

89. Matrice associée à une forme quadratique, définition 8.148.

8.12.8 Réduction de Gauss

THOooOMMFooKxqICS

Théorème 8.241 (Réduction de Gauss[293, 294])**.**
Soit une forme quadratique non nulle q sur l'espace vectoriel E sur le corps $\mathbb{K}$. Il existe une base $\{l_i\}_{i=1,\dots,n}$ de E^* et des coefficients $\alpha_i \in \mathbb{K}$ tels que

$$q(x) = \sum_{i=1}^{n} \alpha_i l_i(x)^2. \tag{8.469}$$

Démonstration. Notre point de départ sont les formules (8.464) pour la forme quadratique. Nous allons faire la preuve par récurrence sur la dimension de l'espace. Si $n = 1$, alors nous avons seulement

$$q(x) = \alpha x^2 \tag{8.470}$$

et donc le théorème est fait avec $l(x) = x$.

Nous supposons que le théorème est prouvé pour tout espace de dimension n. Une forme quadratique pour un espace de dimension $n+1$ s'écrit

$$q(x) = \sum_{i=1}^{n+1} m_{ii}x_i^2 + 2 \sum_{1\leqslant i<j\leqslant n+1} m_{ij}x_ix_j. \tag{8.471}$$

Vu que q est non nulle, un des m_{ij} est non nul. Nous allons diviser en plusieurs cas.

— $m_{11} \neq 0$
— $m_{kk} \neq 0$ avec $k \neq 1$
— $m_{12} \neq 0$ et $m_{ii} = 0$ pour tout i.
— $m_{kl} \neq 0$ avec $(k,l) \neq (1,2)$ et $m_{ii} = 0$ pour tout i.

Ces cas ne sont pas exclusifs, mais ils couvrent toutes les possibilités.

(1) **Si $m_{11} \neq 0$** Nous écrivons q sous la forme

$$q(x) = m_{11}x_1^2 + \sum_{i=2}^{n+1} m_{ii}x_i^2 + 2\sum_{i=1}^{n}\Big(\sum_{j=i+1}^{n+1} m_{ij}x_ix_j\Big) \tag{8.472a}$$
$$= m_{11}x_1^2 + \sum_{i=2}^{n+1} m_{ii}x_i^2 + 2\sum_{j=2}^{n+1} m_{1j}x_1x_j + 2\sum_{i=2}^{n}\sum_{j=i+1}^{n+1}(m_{ij}x_ix_j) \tag{8.472b}$$
$$= m_{11}x_1^2 + 2x_1\sum_{j=2}^{n+1} m_{1j}x_k + R(x_2,\dots,x_{n+1}) \tag{8.472c}$$
$$= m_{11}\left(x_1^2 + 2x_1\sum_{j=2}^{n+1}\frac{m_{1j}}{m_{11}}x_j\right) + R(x_2,\dots,x_{n+1}) \tag{8.472d}$$
$$= m_{11}\big(x_1^2 + 2x_1 f(x_2,\dots,x_{n+1})\big) + R(x_2,\dots,x_{n+1}) \tag{8.472e}$$
$$= m_{11}\big(x_1 + f(x_2,\dots,x_{n+1})\big)^2 - m_{11}f(x_2,\dots,x_{n+1})^2 + R(x_2,\dots,x_{n+1}) \tag{8.472f}$$

où

— R est une forme quadratique de $n-1$ variables ;
— nous avons noté $f(x_2,\dots,x_{n+1}) = \sum_{j=2}^{n+1}\frac{m_{1j}}{m_{11}}x_j$.

Maintenant, toute la partie $-f(x_2,\dots,x_{n+1})^2 + R(x_2,\dots,x_{n+1})$ est une forme quadratique de n variables. Par hypothèse de récurrence, il existe des coefficients α_i et des formes linéairement indépendantes sur $\mathbb{K}^n$ $l'_i(x_2,\dots,x_{n+1})$ telles que

$$-f(x_2,\dots,x_{n+1})^2 + R(x_2,\dots,x_{n+1}) = \sum_{i=2}^{n+1}\alpha_i l'_i(x_2,\dots,x_{n+1})^2. \tag{8.473}$$

En posant ensuite $l_j(x_1,\dots,x_{n+1}) = l'_j(x_2,\dots,x_{n+1})$, ainsi que $l_1(x_1,\dots,x_{n+1}) = x_1 + f(x_2,\dots,x_{n+1})$, nous avons

$$q(x) = m_{11}l_1(x)^2 + \sum_{j=2}^{n+1}\alpha_j l_j(x)^2. \tag{8.474}$$

(2) **Si $m_{kk} \neq 0$ avec $k \neq 1$**
Nous nommons k le plus petit entier pour lequel $m_{kk} \neq 0$, et nous supposons que $k \neq 1$, parce que nous avons déjà couvert ce cas. Dans ce cas, nous avons

$$q(x) = m_{kk}x_k^2 + \sum_{j=k+1}^{n+1} m_{jj}x_j^2 + 2\sum_{i=1}^{n}\Big(\sum_{j=i+1}^{n+1} m_{ij}x_ix_j\Big), \tag{8.475}$$

et tout tourne comme dans le premier cas.

(3) $m_{ii} = 0$ **pour tout** i **et** $m_{12} \neq 0$ Nous écrivons q en séparant les termes m_{1k} :

$$\begin{aligned}
q(x) &= 2\sum_{1\leqslant i<j\leqslant n+1} m_{ij}x_ix_j && (8.476a)\\
&= 2m_{12}x_1x_2 + 2\sum_{2\leqslant j\leqslant n+1} m_{1j}x_1x_j + 2\sum_{2\leqslant i<j\leqslant n+1} m_{ij}x_ix_j && (8.476b)\\
&= 2m_{12}x_1x_2 + 2x_1\sum_{2\leqslant j\leqslant n+1} m_{1j}x_j + 2\sum_{3\leqslant j\leqslant n+1} m_{2j}x_2x_j + 2\sum_{3\leqslant i<j\leqslant n+1} m_{ij}x_ix_j && (8.476c)\\
&= 2m_{12}x_1x_2 + x_1f(x_2,\dots,x_{n+1}) + x_2g(x_3,\dots,x_{n+1}) + T(x_3,\dots,x_{n+1}) && (8.476d)
\end{aligned}$$

SUBEQooLBXBooXoLyuw

où f et g sont linéaires et T est multilinéaire.
À ce moment, nous tentons de factoriser toute la partie concernant x_1 et x_2. L'idée est d'utiliser ceci :

$$(x_1 + g)(x_2 + f) = x_1x_2 + x_1f + x_2g + fg, \tag{8.477}$$

mais en mettant les bons coefficients pour reproduire ce que nous avons dans (8.476d) :

$$(2m_{12} + 2g)(x_2 + \frac{f}{m_{12}}) - \frac{2fg}{m_{12}} = 2m_{12}x_1x_2 + 2x_1f + 2x_2g. \tag{8.478}$$

Cela pour dire que

$$q(x) = 2(m_{12}x_1 + g)(x_2 + \frac{f}{m_{12}}) - \frac{2fg}{m_{12}} + T \tag{8.479}$$

où $-2fg/m_{12} + T$ est une forme quadratique de $x_3,\dots,x_{n+1}$, c'est-à-dire de $n-1$ variables. L'hypothèse de récurrence nous donne des formes linéaires $(l_i)_{i=3,\dots,n+1}$ telles que

$$\frac{2fg}{m_{12}} + T = \sum_{i=3}^{n+1}\alpha_i l_i(x)^2. \tag{8.480}$$

Nous pouvons donc déjà écrire

$$q(x) = 2l'_1(x)l'_2(x) + \sum_{i=3}^{n+1}\alpha_i l_i(x)^2 \tag{8.481}$$

où

— Les forme l_i avec $i \geqslant 3$ ne dépendent pas de x_1 et x_2, et sont donc indépendantes de l_1 et l_2.
— La forme l'_1 ne dépend pas de x_2,
— La forme l'_2 ne dépend pas de x_1.

Ce sont donc $n+1$ formes linéaires indépendantes. Le seul problème résiduel est que les formes l'_1 et l'_2 arrivent en produit l'une de l'autre. Nous en définissons donc deux de plus :

$$\begin{aligned} l_1(x) &= \frac{1}{2}(l'_1 + l'_2) \\ l_2(x) &= \frac{1}{2}(l'_1 - l'_2), \end{aligned} \tag{8.482}$$

qui sont linéairement indépendantes l'une de l'autre et indépendantes des l_i ($i \geqslant 3$). Au final,

$$q(x) = l_1(x)^2 + l_2(x)^2 + \sum_{i=3}^{n+1} \alpha_i l_i(x)^2. \tag{8.483}$$

(4) **Si $m_{ii} = 0$ et $m_{12} = 0$ et $m_{kl} \neq 0$ avec $k < l$** Nous considérons la permutation

$$\begin{aligned} \sigma\colon \{1, \dots, n+1\} &\to \{1, \dots, n+1\} \\ i &\mapsto \begin{cases} 1 & \text{si } i = k \\ 2 & \text{si } i = l \\ k & \text{si } i = 1 \\ l & \text{si } i = 2 \\ i & \text{sinon,} \end{cases} \end{aligned} \tag{8.484}$$

c'est-à-dire que σ permute 1 et k ainsi que 2 et l. Ensuite nous posons

$$\begin{aligned} s\colon \mathbb{R}^{n+1} &\to \mathbb{R}^{n+1} \\ e_i &\mapsto e_{\sigma(i)}. \end{aligned} \tag{8.485}$$

Nous allons un peu considérer $q \circ s$, pour changer :

$$(q \circ s)(x) = \sum_{i,j} m_{ij} s(x)_i s(x)_j = \sum_{ij} x_{\sigma(i)} x_{\sigma(j)}. \tag{8.486}$$

EQooLVAWooAirEzP

parce que $s(x)_i = x_{\sigma(i)}$.
Utilisons un petit abus de notation pour considérer

$$\begin{aligned} \sigma\colon \{1, \dots, n+1\} \times \{1, \dots, n+1\} &\to \{1, \dots, n+1\} \times \{1, \dots, n+1\} \\ (i, j) &\mapsto \big(\sigma(i), \sigma(j)\big). \end{aligned} \tag{8.487}$$

Cela est une bijection ; nous pouvons utiliser le lemme 1.309 pour permuter les termes dans (8.486) :

$$(q \circ s)(x) = \sum_{ij} m_{\sigma(i)\sigma(j)} x_{\sigma\sigma(i)} x_{\sigma\sigma(j)} \tag{8.488a}$$

$$= \sum_{ij} a_{ij} x_i x_j \tag{8.488b}$$

EQooPCTCooFnMWat

où nous avons posé $a_{ij} = m_{\sigma(i)\sigma(j)}$ et utilisé le fait que $\sigma = \sigma^{-1}$. Le point intéressant de l'histoire est que dans (8.488b), $a_{12} = m_{kl} \neq 0$. La forme $q \circ s$ est donc dans le cas déjà traité et il existe des formes linéaires l'_i telles que

$$(q \circ s)(x) = \sum_{i=1}^{n+1} \alpha_i l'_i(x)^2. \tag{8.489}$$

En évaluant cela en $s(x)$, et en tenant compte de $s = s^{-1}$, nous trouvons

$$q(x) = \sum_i \alpha_i (l_i \circ s)(x)^2, \tag{8.490}$$

de telle sorte que $l_i = l'_i \circ s$ soit la réponse à notre théorème.

□

8.12.9 Orthogonalité pour une forme bilinéaire

PROPooYXMMooYIuGRd

Proposition 8.242 ([294]).
Soient un espace vectoriel $(E, \mathbb{K})$ et une forme quadratique[90] *q. Une base de E est q-orthogonale*[91] *si et seulement si la matrice de q dans cette base est diagonale.*

Démonstration. La matrice de q est donnée par $Q_{ij} = b(e_i, e_j)$. Donc oui, cette matrice est diagonale si et seulement si les e_i sont orthogonaux. □

PROPooRERSooFHwWtB

Proposition 8.243 ([1]).
Si b est une forme bilinéaire symétrique non dégénérée, et si q est la forme quadratique associée, alors il existe une base q-orthogonale.

Les vecteurs de cette base ne sont pas isotropes[92]

Démonstration. Nous considérons une base $\{e_i\}$ de E. Vu que b est symétrique, elle est donnée par une matrice symétrique selon la formule $b(x, y) = \sum_{kl} S_{kl} x_k y_l$. Le théorème 8.227 permet de diagonaliser S par une matrice orthogonale : il existe une matrice diagonale D et une matrice orthogonale A telles que $S = A^t D A$.

Nous posons alors $s_i = Ae_i$, et nous vérifions que c'est bon. Nous avons :

$$\begin{aligned} b(s_i, s_j) &= \sum_{kl} S_{kl}(s_i)_k (s_j)_l && \text{(8.491a)}\\ &= \sum_{kl} S_{kl}(Ae_i)_k (Ae_j)_l && \text{(8.491b)}\\ &= \sum_{kl} S_{kl} A_{ki} A_{lj} && \text{(8.491c)}\\ &= \sum_{kl} (A^t)_{ik} S_{kl} A_{lj} && \text{(8.491d)}\\ &= D_{ij}. && \text{(8.491e)} \end{aligned}$$

Vu que b n'est pas dégénérée, les éléments diagonaux de D ne sont pas nuls. Donc les vecteurs s_i vérifient $b(s_i, s_i) \neq 0$. □

PROPooZYISooQZNgeo

Proposition 8.244.
Soit une forme quadratique q. Si une base (e_i) de E est q-orthogonale, alors $\mathcal{B} = \{e_i \text{ tel que } q(e_i) = 0\}$ est une base de $\ker(q)$.

Démonstration. Nous considérons un vecteur de base e_j, et nous montrons que $q(e_j) = 0$ si et seulement si $e_j \in \ker(q)$. Nous savons par la proposition 8.242 que la matrice de q dans la base (e_i) est diagonale et que les éléments diagonaux sont les $q(e_i)$. Soit $K = \{i \text{ tel que } q(e_i) = 0\}$.

(1) $\underline{\text{Span}\{e_i\}_{i\in K} \subset \ker(q)}$ Si $x = \sum_{i\in K} x_i e_i$, alors

$$q(x) = b(x, x) = \sum_{i,j\in K} |x_i|^2 b(e_i, e_j) = \sum_{i,j\in K} |x_i|^2 \delta_{ij} q(e_i) = 0 \tag{8.492}$$

parce que $q(e_i) = 0$ dès que $i \in K$.

(2) $\underline{\ker(q) \subset \text{Span}\{e_i\}_{i\in K}}$ Soit $x \in \ker(q)$ et écrivons-le sous la forme $x = \sum_{i=1}^{n} x_i e_i$. Nous avons

$$0 = q(x) = \sum_i |x_i|^2 q(e_i). \tag{8.493}$$

Mais $|x_i|^2 \geqslant 0$ et $q(e_i) \geqslant 0$, donc si $q(e_i) \neq 0$, alors $x_i = 0$. Donc les seules composantes non nulles de x sont celles sur lesquelles q s'annule. En d'autres termes $x = \sum_i x_i e_i \in \text{Span}\{e_i\}_{i\in K}$.

90. Définition 8.125.
91. Définition 8.147.
92. Définition 8.143.

□

THOooIDMPooIMwkqB

Théorème 8.245 ([294, 1]).
Toute forme quadratique sur un espace vectoriel de dimension finie admet une base formée de vecteurs 2 à 2 orthogonaux (pour la forme considérée).

Démonstration. Nous considérons la base $\{l_i\}$ de E^* donnée par la réduction de Gauss (théorème 8.241). La forme quadratique q s'écrit

$$q(x) = \sum_{i=1}^{n} \alpha_i l_i(x)^2. \tag{8.494}$$

La base préduale[93] $\{e_i\}$ de $\{l_i\}$ répond aux conditions. Pour le vérifier, nous considérons la forme bilinéaire associée à q par l'identité de polarisation 8.141 :

$$b(e_i, e_j) = \frac{1}{2}\big(q(e_i) + q(e_j) - q(e_i - e_j)\big). \tag{8.495}$$

Vu que $l_k(e_i) = \delta_{ki}$, nous avons

$$q(e_i) = \sum_{k=1}^{n} \alpha_k l_k(e_i)^2 = \alpha_i. \tag{8.496}$$

En utilisant la linéarité,

$$\begin{aligned} q(e_i - e_j) &= \sum_k \alpha_k l_k(e_i - e_j)^2 && \text{(8.497a)}\\ &= \sum_k \alpha_k (\delta_{ki} - \delta_{kj})^2 && \text{(8.497b)}\\ &= \sum_k \alpha_k (\delta_{ki} + \delta_{kj} - 2\delta_{ki}\delta_{kj}) && \text{(8.497c)}\\ &= \alpha_i + \alpha_j - 2\delta_{ij}\alpha_i. && \text{(8.497d)} \end{aligned}$$

Donc

$$b(e_i, e_j) = \delta_{ij}\alpha_i. \tag{8.498}$$

Les vecteurs $\{e_i\}$ sont donc bien deux à deux q-orthogonaux. □

Notons qu'en l'absence de notion de racine carrée sur $\mathbb{K}$, il n'est pas possible de considérer $\sqrt{\alpha_i}$ et donc de base q-orthonormée.

PROPooPMYCooAAtHsB

Proposition 8.246 ([1]).
Si $A \in \mathbb{M}(n, \mathbb{K})$ est telle que $\det(A) = 0$, alors il existe des matrices de manipulation de lignes et de colonnes[94] *$G_1, \dots, G_p$ et $H_1, \dots, H_q$ telles que $G_1 \dots G_N A H_1 \dots H_q$ ait une colonne de zéros.*

Démonstration. Si la matrice A n'a pas de colonnes de zéros, il existe un i tel que $A_{i1} \neq 0$. Nous utilisons une matrice de permutation de ligne 1 et i :

$$\big(S(n, i, 1)A\big)_{11} \neq 0. \tag{8.499}$$

Nous notons s ce nombre. Alors en multipliant la première ligne par $1/s$ nous avons un 1 dans la première case :

$$\big(T(n, 1, 1/s)S(n, i, 1)A\big)_{11} = 1. \tag{8.500}$$

Pour la suite nous n'allons plus être aussi explicite.

93. Définition, existence, unicité dans la proposition 4.131.
94. Les opérations élémentaires sur les lignes sont résumées en 4.88.

Maintenant, en faisant des combinaisons entre la première colonne et les autres, nous pouvons annuler toutes les autre entrées 1_{1i}. Tout ça pour dire qu'il existe des matrices $G_1, \ldots, G_{p'}$ et $H_1, \ldots, H_{q'}$ telles que

$$G_1 \ldots G_{p'} A H_1 \ldots H_{q'} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A^{(1)} & \\ 0 & & & \end{pmatrix}. \tag{8.501}$$

Si $A^{(1)}$ ne possède pas de colonnes de zéros, nous pouvons continuer.

Si nous parvenons à faire n pas de la sorte, alors nous aurions

$$G_1 \ldots G_p A H_1 \ldots H_q = \delta, \tag{8.502}$$

et donc, par le lemme 4.89, nous aurions $\det(G_1 \ldots G_p)\det(A)\det(H_1 \ldots H_q) = 1$, ce qui est impossible lorsque $\det(A) = 0$. Nous en concluons que le processus doit s'arrêter et qu'une des matrices $A^{(k)}$ doit avoir une colonne de zéros [95]. □

PROPooVUDJooLWjmSI

Proposition 8.247.
Une matrice dont le déterminant est nul n'est pas inversible.

Démonstration. Par la proposition 8.246, il existe des matrices de manipulation de lignes et de colonnes $G_1, \ldots, G_p$ telles que la matrice $G_1 \ldots G_p\, A\, H_1 \ldots H_q$ ait une colonne de zéros. De là, la proposition 4.99 implique que la matrice

$$G_1 \ldots G_p A H_1 \ldots H_q \tag{8.503}$$

EQooQGXBooXxFOth

n'est pas inversible. Vu les déterminants des matrices G_i, la proposition 4.98 implique que $G_1 \ldots G_p$ et $H_1 \ldots H_q$ sont inversibles. Si A était inversible, tout le produit serait inversible, ce qui est faux. □

THOooSNXWooSRjleb

Théorème 8.248.
Une matrice sur un corps commutatif est inversible si et seulement si son déterminant est non nul.

Démonstration. Dans un sens c'est la proposition 4.98 et dans l'autre sens c'est la proposition 8.247. □

PROPooHQNPooIfPEDH

Proposition 8.249.
Soient des matrices A et B sur un corps commutatif. Alors

$$\det(AB) = \det(A)\det(B). \tag{8.504}$$

Démonstration. Les propositions 4.102 et 4.103 ont déjà fait une grosse partie du travail. Il ne reste que le cas où $\det(A) = \det(B) = 0$.

Dans ce cas, les matrices A et B ne sont pas inversibles (proposition 8.248). Le produit AB n'est alors pas inversible non plus [96]. La proposition 8.248, utilisée dans le sens inverse, nous dit alors que $\det(AB) = 0$.

Au final dans le cas $\det(A) = \det(B) = 0$ nous avons $0 = \det(AB) = \det(A)\det(B) = 0$. □

Faisons maintenant le cas général des manipulations de lignes et colonnes.

PROPooSLLGooSZjQrv

Proposition 8.250.
Soit une matrice carrée $A \in \mathbb{M}(n, \mathbb{K})$. La matrice B obtenue par la substitution simultanée

$$C_j \to \sum_k a_{kj} C_k \tag{8.505}$$

a pour déterminant

$$\det(B) = \det(a)\det(A). \tag{8.506}$$

95. En réalité, le processus tel que nous l'avons décrit ne s'arrête que lorsque la première colonne est remplie de zéros.

96. Citez le lemme 4.95 si vous voulez justifier ça.

Démonstration. L'élément B_{ij} de la matrice B est une combinaison linéaire de tous les éléments de sa ligne :

$$B_{ij} = \sum_k a_{kj} A_{ik} = (Aa)_{ij}. \tag{8.507}$$

Donc $B = Aa$. La proposition 8.249 nous dit alors que $\det(B) = \det(a)\det(A)$. □

ThoQFVsBCk

Théorème 8.251 (de Sylvester[287])**.**
Soit Q une forme quadratique réelle de signature[97] *(p, q) sur $\mathbb{R}^n$ $(n \geqslant p + q)$. Alors pour toute base Q-orthogonale $\{e_i\}$ de $\mathbb{R}^{p+q}$ nous avons les propriétés suivantes.*

ITEMooCFQHooRWfmpT

(1) Les nombres p et q sont donnée par

$$p = \text{Card}\{i \text{ tel que } Q(e_i) > 0\} \tag{8.508a}$$

SUBEQooONWLooNsgmQY

$$q = \text{Card}\{i \text{ tel que } Q(e_i) < 0\}. \tag{8.508b}$$

SUBEQooFKXMooOVwxKR

ITEMooWLPVooSTOOjL

(2) Si A est la matrice de Q dans une base, alors il existe une matrice inversible P telle que

$$P^t A P = \begin{pmatrix} -\mathbb{1}_q & & \\ & \mathbb{1}_p & \\ & & 0_{n-p-q} \end{pmatrix}. \tag{8.509}$$

ITEMooGOHCooPrNQwm

(3) Le rang de Q est $p + q$.

Démonstration. Soit F un sous-espace de dimension maximale q sur lequel Q est définie négative. Le fait que la dimension de F soit q est la définition 8.164 de la signature. Nous notons $F^\perp$ sont Q-orthogonal, c'est-à-dire que

$$F^\perp = \{v \in E \text{ tel que } B(v, x) = 0 \, \forall x \in F\}. \tag{8.510}$$

Le lemme 8.139 nous assure que $E = F \oplus F^\perp$.

Le théorème 8.245 sur l'existence de bases Q-orthogonales nous permet de considérer une base Q-orthogonale de F et une de $F^\perp$. En réunissant les deux, nous avons une base de E. Nous la notons $\{f_1, \dots, f_n\}$ avec

— La partie $\{f_1, \dots, f_q\}$ est une base de F,

— La partie $\{f_{q+1}, \dots, f_n\}$ est une base de $F^\perp$,

— Remarquez cependant qu'il n'est pas dit que $n = q + p$.

Notons que pour $i > q$, nous avons $Q(f_i) \geqslant 0$, sinon la maximalité de F serait contredite par $\text{Span}\{f_1, \dots, f_q, f_i\}$.

Cela prouve que

$$\text{Card}\{i \text{ tel que } Q(f_i) > 0\} = p. \tag{8.511}$$

Le lemme 8.167 nous dit alors que

$$\text{Card}\{i \text{ tel que } Q(e_i) > 0\} = \text{Card}\{i \text{ tel que } Q(f_i) > 0\} = p. \tag{8.512}$$

C'est l'égalité (8.508b). L'égalité (8.508a) se prouve de la même façon, en prenant F maximal pour la propriété que Q y est strictement définie positive.

Le point (1) est prouvé.

Dans une base Q-orthogonale, la matrice de Q est diagonale, et contient sur la diagonale les valeurs de $Q(e_i)$. Parmi celles-ci, on en a p strictement positives et q strictement négatives. Les $n - p - q$ autres sont nulles. Vu que Q est à valeur réelle, nous avons une notion de racine carré,

97. Définition 8.164.

et nous pouvons considérer $e_i/\sqrt{|Q(e_i)|}$ au lieu de e_i. De cette façon, $Q(e_i)$ est normalisé. Avec ça, la matrice de Q est

$$D = \begin{pmatrix} \mathbb{1}_p & & \\ & -\mathbb{1}_q & \\ & & 0 \end{pmatrix}. \tag{8.513}$$

EQooLQNRooCsgKVF

Nous venons de prouver qu'il existe une base $\{e_i\}$ dans laquelle la matrice de Q est (8.513). Si A est la matrice de Q dans une base quelconque $\{f_i\}$ et si P est la matrice de changement de base $f_j = \sum_i P_{ij}e_i$, la proposition 8.151 donne $D = P^tAP$.

Le point (2) est prouvé.

Pour (3), la proposition 8.166 nous permet de calculer le rang de Q par le rang de sa matrice dans n'importe quelle base. Nous choisissons la base qui donne la matrice (8.513). Le rang est alors bien $p + q$. □

8.12.10 Équivalence de formes quadratiques

DEFooOLWYooMwhMJp

Définition 8.252 (Équivalence de forme quadratique[279])**.**
*Deux formes quadratiques q et q' sont **équivalentes** si il existe une application linéaire inversible ϕ telle que $q' = q \circ \phi$.*

PROPooBWXMooLsgyKm

Proposition 8.253 ([287])**.**
Deux formes quadratiques sont équivalentes[98] *si et seulement si elles ont même signature.*

Démonstration. En deux parties, et en utilisant tout le temps le théorème de Sylvester 8.251. Dans la suite nous allons toujours noter de la même façon les formes quadratiques, les applications linéaires et leurs matrices correspondantes.

(1) $\Rightarrow$ Soient deux formes quadratiques équivalentes q et q'. Nous avons une application linéaire ϕ telle que $q' = q \circ \phi$. Soit une matrice inversible P telle que

$$P^tAP = \begin{pmatrix} -\mathbb{1}_{q'} & & \\ & \mathbb{1}_{p'} & \\ & & 0 \end{pmatrix}. \tag{8.514}$$

EQooTCIRooFtWiTz

où (p', q') est la signature de q'. Par équivalence et par le changement de base de la proposition 8.150, nous avons $q' = \phi^t q \phi$, et donc

$$P^tq'P = (\phi P)^t q(\phi P). \tag{8.515}$$

Dans la base des $(\phi \circ P)e_i$, la matrice de q est (8.514). Donc le point 8.251(1) du théorème de Sylvester montre que la signature de q est également (p', q').

(2) $\Rightarrow$ Nous supposons que q et q' ont la même signature. Il existe donc des matrices inversibles P et Q telles que $P^tq'P$ et Q^tqQ aient la même forme (celle de la matrice (8.514)). Autrement dit,

$$q' \circ P = q \circ Q, \tag{8.516}$$

d'où nous déduisons que $q' = q \circ Q \circ P^{-1}$ parce que P est inversible. Comme l'application $Q \circ P^{-1}$ est inversible, les formes quadratiques q et q' sont équivalentes.

□

8.12.11 Diagonalisation

DEFooGVGGooWQEIET

Lemme-Définition 8.254.
Soit une forme quadratique[99] *q sur l'espace vectoriel V sur $\mathbb{K}$. Soit A la matrice de q dans la base $\{e_i\}$ et B sa matrice dans la base $\{f_\alpha\}$. Nous supposons que le changement de base est orthogonal.*

98. Définition 8.252.
99. Définition 8.125.

Alors les valeurs propres de A et B sont les mêmes.

Ces valeurs sont les ***valeurs propres*** *de q.*

Démonstration. Nous nous rappelons de la définition 8.148 de la matrice associée à q, et à la proposition 8.151 qui parle de changement de base : $B = Q^tAQ$ où Q est orthogonale.

Soit un vecteur propre v de A, de valeur propre λ. Alors nous prouvons que Q^tv est un vecteur propre pour B, de même valeur propre λ. En effet,

$$BQ^tv = Q^tAQQ^tv = Q^tAv = \lambda Q^tv \tag{8.517}$$

où nous avons utilisé $QQ^t = \mathbb{1}$ et $Av = \lambda v$. □

PropFWYooQXfcVY

Proposition 8.255.
Dans la base de diagonalisation de sa matrice associée, une forme quadratique a la forme

$$q(x) = \sum_i \lambda_i x_i^2 \tag{8.518}$$

où les λ_i sont les valeurs propres de la matrice associée à q.

Démonstration. Soit q une forme quadratique et b la forme bilinéaire associée. Si $\{f_i\}$ est une base de diagonalisation [100] de la matrice de b alors dans cette base nous avons

$$q(x) = b(x,x) = \sum_{ij} x_ix_jb(f_i,f_j) = \sum_i \lambda_i x_i^2 \tag{8.519}$$

où les λ_i sont les valeurs propres de la matrice de b. □

Notons que si nous choisissons une autre base de diagonalisation, les λ_i ne changement pas (à part l'ordre éventuellement).

Cela justifie la définition pour dire que nous nous permettrons de parler des **valeurs propres** d'une forme quadratique comme étant les valeurs propres de la matrice associée.

Le théorème 8.245 a déjà donné une base orthogonale pour toute forme quadratique sur un espace vectoriel $(E, \mathbb{K})$ de dimension finie. Dans le cas de $\mathbb{R}^n$, nous pouvons en donner une preuve basée sur le théorème spectral, c'est la proposition 8.256.

PROPooUKRUooGRIDHt

Proposition 8.256.
Soit une forme bilinéaire symétrique b sur un $\mathbb{R}^n$. Il existe une matrice orthogonale Q telle que

(1) $D = Q^tbQ$ est diagonale

(2) $D(x,y) = b(Qx,Qy)$ pour tout $x, y \in \mathbb{R}^n$.

Il existe une base $(f_i)_{i=1,\dots,n}$ qui est b-orthogonale.

Dans cet énoncé, nous mélangeons sans vergogne les formes et les matrices, en supposant qu'une base soit fixée [101]*. Par exemple*

$$D(x,y) = \sum_{ij} D_{ij}x_iy_j. \tag{8.520}$$

Démonstration. Pour la matrice diagonale, c'est le théorème spectral 8.227(2) qui joue parce que la matrice d'une forme bilinéaire symétrique est symétrique (c'est vu de la définition (8.309)).

100. Qui existe parce que la matrice est symétrique, théorème 8.227.

101. Autrement dit, si vous avez en tête d'utiliser cette proposition pour $\mathbb{R}^n$ c'est bon ; mais sinon vous devez choisir une base et considérer toutes les matrices dans cette base.

Pour le reste c'est un calcul :

$$D(x,y) = \sum_{ijkl} Q^t_{ik} b_{kl} Q_{lj} x_i y_j \tag{8.521a}$$
$$= \sum_{ijkl} b_{kl}(Q_{ki}x_i)(Q_{lj}y_j) \tag{8.521b}$$
$$= \sum_{kl} b_{kl}(Qx)_k(Qy)_l \tag{8.521c}$$
$$= b(Qx,Qy). \tag{8.521d}$$

Nous avons utilisé le produit matrice fois vecteur donné par (4.84).

En ce qui concerne l'existence d'une base b-orthogonale, vu que D est diagonale, nous avons, pour $i \neq j$ que $D(e_i,e_j) = 0$. Donc en posant $f_i = Qe_i$, nous trouvons

$$0 = D(e_i,e_j) = b(Qe_i,Qe_j) = b(f_i,f_j). \tag{8.522}$$

La base $(Qe_i)_{i=1,\dots,n}$ est donc b-orthogonale. □

8.13 Fonctions

Sect_fonctions

Soient $(E,\|.\|_E)$ et $(F,\|.\|_F)$ deux espaces vectoriels normés, et une fonction f de E dans F. Il est maintenant facile de définir les notions de limites et de continuité pour de telles fonctions en copiant les définitions données pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et en changeant simplement les valeurs absolues par les normes sur E et F.

La proposition suivante explicite la définition 7.110 dans le cas où la topologie est donnée par des boules.

PropHOCWooSzrMjl

Proposition 8.257 (Caractérisation de la limite)**.**
Soient des espaces vectoriels normés. Soit $f\colon E \to F$ une fonction de domaine $\mathrm{Dom}(f) \subset E$ et soit a un point d'accumulation de $\mathrm{Dom}(f)$.

(1) Si F est séparé[102] *et si f admet une limite en a, alors cette limite est unique.* PTEMooSHKNooStKGKH

(2) La fonction f admet une limite en $a \in E$ si et seulement si il existe un élément $\ell \in F$ tel que pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $x \in D = \mathrm{Dom}(f)$,

$$0 < \|x-a\|_E < \delta \;\Rightarrow\; \|f(x)-\ell\|_F < \varepsilon. \tag{8.523}$$ EqDefLimzxmasubV

*Si la limite existe et est unique, nous écrivons $\lim_{x\to a} f(x) = \ell$ et nous disons que ℓ est la **limite** de f lorsque x tend vers a.*

Démonstration. L'unicité est la proposition 7.113.

(1) $\Rightarrow$ La définition 7.110 nous assure de l'existence d'un élément ℓ tel que pour tout voisinage S de ℓ, il existe un ouvert U autour de a tel que $f\big(U \cap D\backslash\{a\}\big) \subset S$.
Soit $\epsilon > 0$. Nous posons $S = B(\ell,\epsilon)$. Il existe un voisinage U de a tel que $f\big(U \cap D\backslash\{a\}\big) \subset B(\ell,\epsilon)$. Puisque U est un voisinage de a, il contient une boule centrée en a (c'est dans la définition 7.117 de la topologie métrique). Soit donc $\delta > 0$ tel que $B(a,\delta) \subset U$.
Un élément de D qui est dans $B(0,\delta)\backslash\{a\}$ est un élément de D qui vérifie $0 < \|x-a\| < \delta$. Nous avons donc, pour $x \in D$ que

$$0 < \|x-a\| < \delta \Rightarrow \|f(x)-\ell\| < \epsilon. \tag{8.524}$$

(2) $\Leftarrow$ C'est le même raisonnement.

□

102. C'est le cas en dimension finie et en particulier pour $\mathbb{R}^n$. En dimension infinie, il faut être très prudent.

Remarque 8.258.
Le fait que nous limitions la formule (8.523) aux x dans le domaine de f n'est pas anodin. Considérons la fonction $f(x) = \sqrt{x^2 - 4}$, de domaine $|x| \geqslant 2$. Nous avons

$$\lim_{x \to 2} \sqrt{x^2 - 4} = 0. \tag{8.525}$$

Nous ne pouvons pas dire que cette limite n'existe pas en justifiant que la limite à gauche n'existe pas. Les points $x < 2$ sont hors du domaine de f et ne comptent dons pas dans l'appréciation de l'existence de la limite.

Vous verrez plus tard que ceci provient de la topologie induite de $\mathbb{R}$ sur l'ensemble $[2, \infty[$.

8.14 Sous espaces caractéristiques

Lorsqu'un opérateur n'est pas diagonalisable, les valeurs propres jouent quand même un rôle important.

DefFBNIooCGbIix

Définition 8.259.
Soient E un $\mathbb{K}$-espace vectoriel et $f \in \mathrm{End}(E)$. Pour $\lambda \in \mathbb{K}$ nous définissons

$$F_\lambda(f) = \{v \in E \text{ tel que } (f - \lambda \mathbb{1})^n v = 0, n \in \mathbb{N}\} \tag{8.526}$$

*et nous appelons cet ensemble un **sous-espace caractéristique** de f.*

L'espace $F_\lambda(f)$ est l'ensemble de nilpotence de l'opérateur $f - \lambda\mathbb{1}$ et

LemBLPooHMAoyJ

Lemme 8.260.
L'ensemble $F_\lambda(f)$ est non vide si et seulement si λ est une valeur propre de f. L'espace $F_\lambda(f)$ est invariant sous f.

Démonstration. Si $F_\lambda(f)$ est non vide, nous considérons $v \in F_\lambda(f)$ et n le plus petit entier non nul tel que $(f - \lambda)^n v = 0$. Alors $(f - \lambda)^{n-1} v$ est un vecteur propre de f pour la valeur propre λ. Réciproquement, si v est un vecteur propre de f pour la valeur propre λ, alors $v \in F_\lambda(f)$.

En ce qui concerne l'invariance, remarquons que f commute avec $f - \lambda\mathbb{1}$. Si $x \in F_\lambda(f)$ il existe n tel que $(f - \lambda\mathbb{1})^n x = 0$. Nous avons aussi

$$(f - \lambda\mathbb{1})^n f(x) = f\big((f - \lambda\mathbb{1})^n x\big) = 0, \tag{8.527}$$

par conséquent $f(x) \in F_\lambda(f)$. □

Contrairement à ce que l'on pourrait croire, il n'est pas vrai que toute matrice à coefficient réel est diagonalisable, même pas sur $\mathbb{C}$. La raison est qu'une telle matrice peut très bien avoir des valeurs propres multiples.

ExBRXUooIlUnSx

Exemple 8.261.
Le théorème 8.219 nous donne une façon simple de trouver des matrices non diagonalisables sur $\mathbb{C}$: il suffit que le polynôme minimal ne soit pas scindé à racines simples. Par exemple

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{8.528}$$

dont le polynôme caractéristique est $\chi_A = (1 - X)^2$. Ce polynôme n'a manifestement pas des racines simples. Nous pouvons faire le calcul explicite pour montrer que A n'est pas diagonalisable. D'abord l'unique valeur propre de A est 1 et nous pouvons sans peine résoudre

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.529}$$

qui revient au système

$$\begin{cases} x+y=x & \text{(8.530a)}\\ y=y. & \text{(8.530b)}\end{cases}$$

La première équation donne directement $y=0$. Le seul espace propre est de dimension 1 et est engendré par $\begin{pmatrix}1\\0\end{pmatrix}$. △

Nous donnons maintenant un exemple un peu plus avancé de matrice réelle non diagonalisable, qui montre la multiplicité algébrique et géométrique d'une racine d'un polynôme caractéristique.

RemBOGooCLMwyb

Remarque 8.262.
Considérons l'endomorphisme $f\in\mathrm{End}(\mathbb{C}^3)$ donné par la matrice

$$\begin{pmatrix}a&\alpha&\beta\\0&a&\gamma\\0&0&b\end{pmatrix} \tag{8.531}$$

avec $a\neq b$, $a\neq 0$, $b\neq 0$, $\alpha\neq 0$, β et γ sont des nombres complexes quelconques. Son polynôme caractéristique est

$$\chi_f(X)=(a-X)^2(b-X), \tag{8.532}$$

et les valeurs propres sont donc a et b. Nous trouvons les vecteurs propres pour la valeur a en résolvant

$$\begin{pmatrix}a&\alpha&\beta\\0&a&\gamma\\0&0&b\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}ax\\ay\\az\end{pmatrix}. \tag{8.533}$$

La troisième équation est $bz=az$ qui oblige $z=0$ parce que $a\neq b$ et $0\neq a$. La première est $ax+\alpha y=ax$ qui implique $y=0$ parce que $\alpha\neq 0$. Enfin la première équation se réduit à $ax=ax$ qui ne donne pas de contraintes sur x. En résumé : l'espace propre $E_a(f)$ est réduit à une seule dimension générée par $(1,0,0)$.

De la même façon l'espace propre correspondant à la valeur propre b est donné par le système

$$\begin{pmatrix}a&\alpha&\beta\\&a&\gamma\\&&b\end{pmatrix}\begin{pmatrix}w\\y\\z\end{pmatrix}=\begin{pmatrix}bx\\by\\bz\end{pmatrix} \tag{8.534}$$

La seconde équation donne $ay+\gamma z=by$, et donc

$$y=\frac{\gamma}{b-a}z. \tag{8.535}$$

La première équation est $ax+\alpha y+\beta z=bx$ qui donne

$$x=\frac{1}{b-a}(\alpha y+\beta z). \tag{8.536}$$

En y remettant la valeur déjà trouvée de y, nous trouvons que l'espace propre pour la valeur propre b est engendré par le vecteur

$$\begin{pmatrix}\frac{1}{b-a}\left(\beta+\frac{\alpha\gamma}{b-a}\right)\\ \frac{\gamma}{b-a}\\ 1\end{pmatrix}. \tag{8.537}$$

Vu que nous savons que a et b sont les seules valeurs propres et que nous venons de voir que leurs espaces propres sont de dimension 1, il n'y a donc pas trois vecteurs propres linéairement indépendants, et l'opérateur f n'est pas diagonalisable.

Par contre nous pouvons voir que $(f-a\mathbb{1})^2e_2=0$. En effet :

$$(f-a\mathbb{1})^2\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}0&\alpha&\beta\\0&0&\gamma\\0&0&b-a\end{pmatrix}\begin{pmatrix}\alpha\\0\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},\tag{8.538}$$

de telle sorte que le vecteur $(0,1,0)$ est également dans l'espace caractéristique $F_a(f)$.

Dans cet exemple, la multiplicité algébrique de la racine a du polynôme caractéristique vaut 2 tandis que sa multiplicité géométrique vaut seulement 1.

8.14.1 Théorèmes de décomposition

ThoSpectraluRMLok

Théorème 8.263 (Théorème spectral, décomposition primaire)**.**
Soit E un espace vectoriel de dimension finie sur le corps algébriquement clos $\mathbb{K}$ et $f\in\mathrm{End}(E)$. Alors

EqCTFHooBSGbYK
$$E=F_{\lambda_1}(f)\oplus\ldots\oplus F_{\lambda_k}(f)\tag{8.539}$$

où la somme est sur les espaces caractéristiques engendrés par les valeurs propres distinctes de f.

Les projecteurs sur les espaces caractéristiques forment un système complet et orthogonal.

Démonstration. Soit P le polynôme caractéristique de f et une décomposition

$$P=(f-\lambda_1)^{\alpha_1}\ldots(f-\lambda_r)^{\alpha_r}\tag{8.540}$$

en facteurs irréductibles. Par le théorème des noyaux (8.92) nous avons

EqDeFVSaYv
$$E=\ker(f-\lambda_1)^{\alpha_1}\oplus\ldots\oplus\ker(f-\lambda_r)^{\alpha_r}.\tag{8.541}$$

Les projecteurs sont des polynômes en f et forment un système orthogonal. Il nous reste à prouver que $\ker(f-\lambda_i)^{\alpha_i}=F_{\lambda_i}(f)$. L'inclusion

EqzmNxPi
$$\ker(f-\lambda_i)^{\alpha_i}\subset F_{\lambda_i}(f)\tag{8.542}$$

est évidente. Nous devons montrer l'inclusion inverse.

(1) $F_{\lambda_i}(f)\cap F_{\lambda_j}(f)=0$
Soit $v\in F_{\lambda_i}(f)\cap F_{\lambda_j}(f)$. Le fait que $v\in F_{\lambda_i}(f)$ implique qu'il existe $n\in\mathbb{N}$ tel que $(f-\lambda_i)^nv\neq0$ et $(f-\lambda_i)^{n+1}v=0$ (éventuellement $n=0$ si v est un vecteur propre). Posons $v_1=(f-\lambda_i)^nv$. Étant donné que $(f-\lambda_i)$ commute avec $(f-\lambda_j)$, ce v_1 est encore dans $F_{\lambda_j}(f)$. En effet, si k est tel que $(f-\lambda_j)^kv=0$, alors

$$(f-\lambda_j)^kv_1=(f-\lambda_j)^k(f-\lambda_i)^nv=(f-\lambda_i)^n(f-\lambda_j)^kv=0.\tag{8.543}$$

Il existe donc $m\in\mathbb{N}$ tel que $(f-\lambda_j)^mv_1\neq0$ et $(f-\lambda_j)^{m+1}v_1=0$. En posant $w=(f-\lambda_j)^mv_1$, nous avons

$$\begin{cases}(f-\lambda_i)w=(f-\lambda_j)^m(f-\lambda_i)^{n+1}v=0&(8.544a)\\(f-\lambda_j)w=(f-\lambda_j)^{m+1}v_1=0.&(8.544b)\end{cases}$$

Ce w serait donc un vecteur propre simultané pour les valeurs propres λ_i et λ_j. Vu que les espaces propres sont linéairement indépendants, les seules possibilités sont $i=j$ ou $w=0$.

(2) **Questions de dimension** Étant donné que espaces F_{λ_i} sont en somme directe, la somme de leurs dimensions est au maximum la dimension de E :

$$\sum_i\dim F_{\lambda_i}(f)\leqslant\dim E.\tag{8.545}$$

En tenant compte de l'inclusion (8.542) nous avons même

$$\dim E = \sum_i \dim \ker(f-\lambda_i)^{\alpha_i} \leqslant \sum_i \dim F_{\lambda_i}(f) \leqslant \dim E. \tag{8.546}$$

Vu qu'il y a $\dim(E)$ des deux côtés des inégalités, toutes les inégalités sont des égalités et nous avons

$$\sum_i \dim \ker(f-\lambda_i)^{\alpha_i} = \sum_i \dim F_{\lambda_i}(f). \tag{8.547}$$

EQooQPLMooDXAgZi

L'inclusion (8.542) nous dit qu'il y a une inégalité terme à terme dans les sommes de (8.547). Vu qu'il y a égalité des sommes, il y a en réalité égalité de chacun des termes : $\dim\ker(f-\lambda_i)^{\alpha_i} = \dim F_{\lambda_i}(f)$ et l'égalité des deux espaces de (8.542) :

$$\ker(f-\lambda_i)^{\alpha_i} = F_{\lambda_i}(f). \tag{8.548}$$

□

PROBooJRDAooEKxPbJ

¡¡ Avertissement/question au lecteur !! 8.264
Dans le cas où le corps n'est pas algébriquement clos, il paraît qu'il faut remplacer « diagonalisable » par « semi-simple ».

Si vous connaissez un énoncé précis et une démonstration, écrivez-moi.

Il y a peut-être une réponse dans [295].

Si l'espace vectoriel est sur un corps algébriquement clos, alors les endomorphismes semi-simples [103] sont les endomorphismes diagonaux.

ThoRURcpW

Théorème 8.265 (Décomposition de Dunford)**.**
Soit E un espace vectoriel sur le corps algébriquement clos [104] $\mathbb{K}$ et $u \in \mathrm{End}(E)$ un endomorphisme de E.

(1) L'endomorphisme u se décompose de façon unique sous la forme

$$u = s + n \tag{8.549}$$

où s est diagonalisable, n est nilpotent et $[s,n] = 0$ [105].

(2) Les endomorphismes s et n sont des polynômes en u et commutent avec u.

ItemThoRURcpWiii

(3) Si notons $\{\lambda_i\}$ les valeurs propres distinctes de u, et $F_{\lambda_i}(u)$ les espaces caractéristiques correspondants, alors les parties s et n sont données par

$$s = \sum_i \lambda_i p_i \tag{8.550a}$$

$$n = \sum_i (s - \lambda_i \mathbb{1}) p_i \tag{8.550b}$$

$p_i \colon E \to F_{\lambda_i}(u)$ est la projection de E sur $F_{\lambda_i}(u)$.

Démonstration. Le théorème spectral 8.263 nous indique que

$$E = \bigoplus_i F_{\lambda_i}(f). \tag{8.551}$$

103. Définition 8.108.
104. Je crois quo'n peut remplacer l'ypothèse de corps alébriquement clos par une hypothèse de polynôme caractéristique scindé. Écrivez-moi si vous avez une idée à ce propos.
105. Lorsque a et b sont des opérateurs, la notation $[a,b]$ signifie le commutateur entre a et b, c'est-à-dire $a\circ b - b\circ a$. Dire que $[a,b]=0$ signifie que $ab=ba$.

Nous considérons l'endomorphisme s de E qui consiste à dilater d'un facteur λ_i l'espace caractéristique $F_{\lambda_i}(f)$:

$$s = \sum_i \lambda_i p_i \tag{8.552}$$

où $p_i : E \to F_{\lambda_i}(u)$ est la projection de E sur $F_{\lambda_i}(u)$.

Nous allons prouver que $[s, f] = 0$ et $n = f - s$ est nilpotent. Cela impliquera que $[s, n] = 0$.

Si $x \in F_\lambda(f)$, alors nous avons $sf(x) = \lambda f(x)$ parce que $f(x) \in F_\lambda(f)$ tandis que $fs(x) = f(\lambda x) = \lambda f(x)$. Par conséquent f commute avec s.

Pour montrer que $f - s$ est nilpotent, nous en considérons la restriction

$$f - s : F_\lambda(f) \to F_\lambda(f). \tag{8.553}$$

Cet opérateur est égal à $f - \lambda\mathbb{1}$ et est par conséquent nilpotent.

Prouvons à présent l'unicité. Soit $u = s' + n'$ une autre décomposition qui satisfait aux conditions : s' est diagonalisable, n' est nilpotent et $[n', s'] = 0$. Commençons par prouver que s' et n' commutent avec u. En multipliant $u = s' + n'$ par s' nous avons

$$s'u = s'^2 + s'n' = s'^2 + n's' = (s' + n')s' = us', \tag{8.554}$$

par conséquent $[u, s'] = 0$. Nous faisons la même chose avec n' pour trouver $[u, n'] = 0$. Notons que pour obtenir ce résultat nous avons utilisé le fait que n' et s' commutent, mais pas leur propriétés de nilpotence et de diagonalisibilité.

Si $s' + n' = s + n$ est une autre décomposition, s' et n' commutent avec u, et par conséquent avec tous les polynômes en u. Ils commutent en particulier avec n et s. Les endomorphismes s et s' sont alors deux endomorphismes diagonalisables qui commutent. Par la proposition 8.222, ils sont simultanément diagonalisables. Dans la base de diagonalisation simultanée, la matrice de l'opérateur $s' - s = n - n'$ est donc diagonale. Mais $n - n'$ est également nilpotent, en effet si A et B sont deux opérateurs nilpotents,

$$(A + B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}. \tag{8.555}$$

Si n est assez grand, au moins un parmi A^k ou B^{n-k} est nul.

Nous savons que $n - n'$ est diagonal et nilpotent. Le seul opérateur diagonal à être nilpotent est l'opérateur nul[106]. Nous en déduisons que $n = n'$. Nous avons alors immédiatement aussi $s = s'$. □

8.266.
Le théorème 12.413 montrera que $A^n x \to 0$ pour tout x si et seulement si $\rho(A) < 1$, mais ça demande un résultat de vitesse comparée entre l'exponentielle et la puissance.

Une application de la décomposition de Jordan est l'existence d'un logarithme pour les matrices. La proposition suivante va d'une certaine manière donner un logarithme pour les matrices inversibles complexes. Dans le cas des matrices réelles m telles que $\|m - \mathbb{1}\| < 1$, nous donnerons au lemme 15.152 une formule pour le logarithme sous forme d'une série ; ce logarithme sera réel.

8.14.2 Valeurs singulières

DEFooZSCYooQnBzix

Définition 8.267.
*Soit M une matrice $m \times n$ sur $\mathbb{K}$ ($\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$). Un nombre réel σ est une **valeur singulière** de M si il existe des vecteurs unitaires $u \in \mathbb{K}^m$, $v \in \mathbb{K}^n$ tels que*

$$Mv = \sigma u \tag{8.556a}$$

$$M^*u = \sigma v. \tag{8.556b}$$

106. Parce qu'une puissance d'un opérateur diagonal est diagonal.

THOooSCOLooSKlVvd

Théorème 8.268 (Décomposition en valeurs singulières)**.**
Soit $M \in \mathbb{M}(m \times n, \mathbb{K})$ où $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Alors M se décompose en

$$M = ADB \tag{8.557}$$

où il existe deux matrices unitaires $A \in \mathrm{U}(m \times m)$, $B \in \mathrm{U}(n \times n)$ et une matrice (pseudo)diagonale $D \in \mathbb{M}(m \times n)$ tels que

(1) $A \in \mathrm{U}(m \times m)$, $B \in \mathrm{U}(n \times n)$ sont deux matrices unitaires,

(2) D est (pseudo)diagonale,

(3) les éléments diagonaux de D sont les valeurs singulières de M,

(4) le nombre d'éléments non nuls sur la diagonale de D est le rang[107] *de M.*

Corolaire 8.269.
Soit $M \in \mathbb{M}(n, \mathbb{C})$. Il existe un isomorphisme $f : \mathbb{C}^n \to \mathbb{C}^n$ tel que fM soit autoadjoint.

Démonstration. Si $M = ADB$ est la décomposition de M en valeurs singulières, alors nous pouvons prendre $f = \overline{B}^t A^{-1}$ qui est une matrice inversible. Pour la vérification que ce f répond bien à la question, ne pas oublier que D est réelle, même si M ne l'est pas. □

8.15 Extension du corps de base

SECooAUOWooNdYTZf

Nous avons discuté dans la section 6.4 de ce qui arrive au corps lorsqu'on l'étend. Dans cette section nous allons étudier ce qui arrive aux applications linéaires entre deux $\mathbb{K}$-espaces vectoriels lorsque nous étendons le corps $\mathbb{K}$ en un corps $\mathbb{L}$.

Soient $\mathbb{K}$ un corps (commutatif) et une extension $\mathbb{L}$ de $\mathbb{K}$. Soient E et F, des $\mathbb{K}$-espaces vectoriels de dimension finie, et entrons dans le vif du sujet [108].

8.15.1 Extension des applications linéaires

Définition 8.270 ([296])**.**
L'espace vectoriel obtenu par ***extension du corps de base*** *de E est l'espace vectoriel*

$$E_{\mathbb{L}} = \mathbb{L} \otimes_{\mathbb{K}} E. \tag{8.558}$$

Ce dernier est le quotient $\mathbb{L} \otimes_{\mathbb{K}} E = (\mathbb{L} \times E)/\sim$ par la relation d'équivalence

$$(\lambda, v) \sim \big(a\lambda, \frac{1}{a}v\big) \tag{8.559}$$

pour tout $a \in \mathbb{K}\backslash\{0\}$. Nous noterons $[\lambda, v]$ ou $\lambda \otimes v$ ou encore $\lambda \otimes_{\mathbb{K}} v$ la classe de (λ, v) pour la relation d'équivalence $\sim$:

$$[\lambda, v] = \{(\mu, w) \in \mathbb{L} \times E \text{ tel que } (\mu, w) \sim (\lambda, v)\}. \tag{8.560}$$

Un élément de $E_{\mathbb{L}}$ est de la forme $\sum_k [\lambda_k, v_k]$ avec $\lambda_k \in \mathbb{L}$ et $v_k \in E$. Si $f : E \to F$ est une applications linéaire, nous définissons

$$\begin{aligned} f_{\mathbb{L}} : E_{\mathbb{L}} &\to F_{\mathbb{L}} \\ [\lambda, v] &\mapsto [\lambda, f(v)]. \end{aligned} \tag{8.561}$$

EQooYGVHooHYbMHX

Remarque 8.271.
Si deux vecteurs de $E_{\mathbb{L}}$ sont linéairement indépendants pour $\mathbb{K}$, ils ne le sont pas spécialement pour $\mathbb{L}$. Par exemple si $\mathbb{C}$ est vu comme $\mathbb{R}$-espace vectoriel, alors $\{1, i\}$ est une partie libre. Mais dans $\mathbb{C}$, vu comme $\mathbb{C}$-espace vectoriel, la partie $\{1, i\}$ n'est pas libre.

107. Définition 4.46.
108. Le sujet étant le corps étendu.

LEMooJIGTooMqiJSN

Lemme 8.272.
Soient trois $\mathbb{K}$-espaces vectoriels E, F et G ainsi que deux applications linéaires $f\colon E\to F$ et $g\colon G\to E$. Si $\mathbb{L}$ est une extension de $\mathbb{K}$, alors

$$f_{\mathbb{L}}\circ g_{\mathbb{L}}=(f\circ g)_{\mathbb{L}}. \tag{8.562}$$

Démonstration. Il suffit de composer la définition (8.561) :

$$\begin{aligned}(f_{\mathbb{L}}\circ g_{\mathbb{L}})\big([\lambda,v]\big)&=f_{\mathbb{L}}\big([\lambda,g(v)]\big) &(8.563a)\\ &=[\lambda,(f\circ g)(v)] &(8.563b)\\ &=(f\circ g)_{\mathbb{L}}\big([\lambda,v]\big). &(8.563c)\end{aligned}$$

□

Nous définissons aussi l'injection canonique

$$\begin{aligned}\tau\colon E&\to E_{\mathbb{L}}\\ v&\mapsto[1,v].\end{aligned} \tag{8.564}$$

PropooWECLooHPzIHw

Proposition 8.273 ([130]).
Injectivité et surjectivité respectées.

(1) L'application $f_{\mathbb{L}}$ est injective si et seulement si f est injective.

(2) L'application $f_{\mathbb{L}}$ est surjective si et seulement si f est surjective.

Démonstration. En plusieurs parties.

(1) **Si $f_{\mathbb{L}}$ est injective** Supposons pour commencer que $f_{\mathbb{L}}$ soit injective. Le diagramme

$$\begin{array}{ccc}E & \xrightarrow{f} & F\\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau}\\ E_{\mathbb{L}} & \xrightarrow[f_{\mathbb{L}}]{} & F_{\mathbb{L}}\end{array} \tag{8.565}$$

est un diagramme commutatif. En effet

$$(\tau\circ f)(v)=[1,f(v)] \tag{8.566}$$

tandis que

$$(f_{\mathbb{L}}\circ\tau)(v)=f_{\mathbb{L}}[1,v]=[1,f(v)]. \tag{8.567}$$

Donc si $f(v)=0$ avec $v\neq 0$ nous aurions $(\tau\circ f)(v)=0$ et donc aussi $(f_{\mathbb{L}}\circ\tau)(v)=0$, alors que $\tau(v)\neq 0$ dans $E_{\mathbb{L}}$.

(2) **Si f est injecive**
Nous supposons que f est injective et que $f_{\mathbb{L}}(\lambda\otimes v)=f_{\mathbb{L}}(\mu\otimes w)$. Cela signifie que $\lambda\otimes f(v)=\mu\otimes f(w)$, ou encore qu'il existe $a\neq 0$ dans $\mathbb{K}$ tel que

$$\begin{cases}\mu=a\lambda & (8.568a)\\ f(w)=\dfrac{1}{a}f(v). & (8.568b)\end{cases}$$

Par linéarité de f, $f(v)=f(aw)$ et par injectivité de f, nous déduisons que $v=aw$. Enfin nous avons l'injectivité de $f_{\mathbb{L}}$ parce que

$$\mu\otimes w=a\lambda\otimes\frac{v}{a}=\lambda\otimes v. \tag{8.569}$$

□

PROPooMHARooUycAts

Proposition 8.274 ([1, 297]).
Soit $\{e_i\}_{i=1,\dots,p}$ une base de E. Alors $\{1\otimes e_i\}_i$ est une base de $E_{\mathbb{L}}=\mathbb{L}\otimes_{\mathbb{K}}E$.

Démonstration. L'espace vectoriel E peut être écrit comme somme directe $E=\bigoplus_i \mathbb{K}e_i$. Si $\lambda\in\mathbb{L}$ et $k\in\mathbb{K}$ nous avons

$$\lambda\otimes ke_i=\frac{\lambda}{k}\otimes e_i=\frac{\lambda}{k}(1\otimes e_i). \tag{8.570}$$

Cela pour introduire l'application

$$\begin{aligned}\psi\colon \mathbb{L}\otimes_{\mathbb{K}}E&\to\bigoplus_i\mathbb{L}(1\otimes e_i)\\ \sum_k\lambda_k\otimes v_k&\mapsto\oplus_i\sum_k(\lambda_kv_{ik})(1\otimes e_i)\end{aligned} \tag{8.571}$$

où $v_k=\sum_i v_{ik}e_i$ avec $v_{ik}\in\mathbb{K}$, qui représente un isomorphisme de $\mathbb{L}$-espaces vectoriels. La surjectivité est facile. En ce qui concerne l'injectivité, si

$$\sum_i\sum_k(\lambda_kv_{ik})(1\otimes e_i)=0 \tag{8.572}$$

alors les quantités suivantes sont nulles également :

$$\sum_i\sum_k(\lambda_kv_{ik})(1\otimes e_i)=\sum_{ik}(\lambda_k\otimes v_{ik}e_i)=\sum_k(\lambda_k\otimes\sum_i v_{ik}e_i)=\sum_k(\lambda_k\otimes v_k). \tag{8.573}$$

La dernière est l'argument de ψ. Le fait qu'il soit nul implique que ψ est injective. □

Remarque 8.275.
Nous n'avons pas dû prouver que chacun des $\lambda_k\otimes v_k$ était nul. Et encore heureux, parce que cela pouvait très bien être faux, puisqu'il y a plusieurs façons de noter un élément de $E_{\mathbb{L}}$ sous la forme de tels termes.

CORooTQGHooIKhNtr

Corolaire 8.276.
La $\mathbb{L}$-dimension de $E_{\mathbb{L}}$ est égale à la $\mathbb{K}$-dimension de E.

8.15.2 Projections

¡¡ Avertissement/question à la lectrice !! 8.277
Nous allons définir $\texttt{proj}\colon\mathcal{L}(E_{\mathbb{L}},F_{\mathbb{L}})\to\mathcal{L}(E,F)$ en faisant appel à des bases et en prouvant que les quantités définies ne dépendent pas des bases choisies. Il y a surement une façon plus « intrinsèque » de faire.

Nous savons que $\mathbb{L}$ est un $\mathbb{K}$-espace vectoriel dans lequel nous pouvons voir $\mathbb{K}$ comme un sous-espace (lemme 6.60). Dans cette optique nous choisissons dans $\mathbb{L}$ un supplémentaire de $\mathbb{K}$, c'est-à-dire un sous-espace vectoriel de $\mathbb{L}$ tel que

$$\mathbb{L}=\mathbb{K}\oplus V. \tag{8.574}$$

Nous avons alors naturellement une projection $\texttt{proj}\colon\mathbb{L}\to\mathbb{K}$.

Soit $\{e_i\}$ une base de E et $\{e_a\}$ une de F. Nous noterons également e_i et e_a les éléments $\tau(e_i)$ et $\tau(e_a)$ correspondants. Grâce à la proposition 8.274, ce sont des bases de $E_{\mathbb{L}}$ et $F_{\mathbb{L}}$. Si la fonction $f\colon E_{\mathbb{L}}\to F_{\mathbb{L}}$ s'écrit dans ces bases comme

$$f(e_i)=\sum_a f_{ai}e_a \tag{8.575}$$

alors nous définissons $\texttt{proj}(f)$ par

$$(\texttt{proj}\,f)e_i=\sum_a\texttt{proj}(f_{ai})e_a. \tag{8.576}$$

EQooSAFRooJufkLQ

PROPooOEHTooHyjuZQ

Proposition 8.278 ([1]).
L'application `proj` *définie en* (8.576) *est indépendante du choix des bases.*

Démonstration. Notons que dans ce qui suit, les sommes sur a ou b et celles sur i ou j ne vont pas jusqu'au même indice (dimensions de E et F). De plus nous manipulons deux quantités qui se notent $\texttt{proj}$. La première est la projection $\texttt{proj}\colon \mathbb{L} \to \mathbb{K}$ qui ne dépend que d'un choix de supplémentaire et que nous supposons fixée ici. D'autre part il y a $\texttt{proj}\colon E_{\mathbb{L}} \to E$ qui dépend à priori des bases choisies.

Nous choisissons de nouvelles bases qui sont liées aux anciennes bases par

$$\begin{cases} e'_b = \sum_a B_{ab} e_a & \text{(8.577a)} \\ e'_j = \sum_i A_{ji} e_i. & \text{(8.577b)} \end{cases}$$

Les matrices A et B sont dans $\mathrm{GL}(\mathbb{K})$. Nous allons écrire l'opérateur $\texttt{proj}'$ qui correspond à ces bases et montrer que pour toute application linéaire $f\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$ nous avons $\texttt{proj}(f) = \texttt{proj}'(f)$. Nous avons :

$$\begin{aligned} f(e'_j) &= \sum_i A_{ji} f(e_i) && \text{(8.578a)} \\ &= \sum_a \sum_b \sum_i A_{ji} f_{ai} (B^{-1})_{ba} e'b && \text{(8.578b)} \\ &= \sum_b \Big(\sum_{ai} A_{ji} f_{ai} (B^{-1})_{ba} \Big) e'b, && \text{(8.578c)} \end{aligned}$$

ce qui donne

$$(\texttt{proj}'\, f) e'_j = \sum_b \Big(\texttt{proj}\, \big(A_{ji} f_{ai} (B^{-1})_{ba} \big) \Big) e'_b. \qquad \text{(8.579)}$$

EQooUQNBooMWHBbD

Nous calculons maintenant $(\texttt{proj}'\, f)e_j$ en substituant $e_j = \sum_l (A^{-1})_{lj} e'_l$ et en utilisant (8.579), la linéarité de $\texttt{proj}'$ et la $\mathbb{K}$-linéarité de $\texttt{proj}\colon \mathbb{L} \to \mathbb{K}$:

$$\begin{aligned} (\texttt{proj}'\, f)\Big(\sum_l (A^{-1})_{lj} e'_l \Big) &= \sum_l (A^{-1})_{lj} \sum_b \sum_{ai} \texttt{proj}\, \big(A_{li} f_{ai} (B^{-1})_{ba} \big) e_b && \text{(8.580a)} \\ &= \sum_a \texttt{proj}(f_{aj}) e_a && \text{(8.580b)} \\ &= (\texttt{proj}\, f) e_j. && \text{(8.580c)} \end{aligned}$$

Donc $\texttt{proj} = \texttt{proj}'$. □

Note au passage : il y a un abus systématique de notation entre $e_i \in E$ et $\tau(e_i) = 1 \otimes e_i \in E_{\mathbb{L}}$.

REMooBEXGooLgpHzg

Remarque 8.279 ([1]).
L'opération $\texttt{proj}\colon \mathcal{L}(E_{\mathbb{L}}, F_{\mathbb{L}}) \to \mathcal{L}(E, F)$ ne dépend pas des bases choisies un peu partout. Mais elle dépend de l'application $\texttt{proj}\colon \mathbb{L} \to \mathbb{K}$ déjà construite. Et celle-là dépend du choix d'un supplémentaire V qui fournit $\mathbb{L} = \mathbb{K} \oplus V$.

Si $\texttt{proj}(\lambda) = 0$ pour un de ces choix, cela n'implique nullement que $\lambda = 0$. Penser à $i \in \mathbb{C}$ si la projection $\texttt{proj}\colon \mathbb{C} \to \mathbb{R}$ est l'application $(x + iy) \mapsto x$ parallèle à l'axe des imaginaires.

Par contre si $\texttt{proj}(\lambda) = 0$ pour tout choix de V, alors nous avons bien $\lambda = 0$. Dans la suite nous « fixons » un choix de V générique, et lorsque nous rencontrerons l'égalité $\texttt{proj}(\lambda) = 0$ nous en déduirons $\lambda = 0$.

PROPooPWDKooFNFWRI

Proposition 8.280.
Si $f\colon E \to F$ *et si* $f_{\mathbb{L}}(e_j) = \sum_a (f_{\mathbb{L}})_{aj} e_a$ *et si* $f(e_j) = \sum_a f_{aj} e_a$ *alors*

(1) $\texttt{proj}\, f_{\mathbb{L}} = f$,

ITEMooNMPYooXosGhI

(2) $(f_{\mathbb{L}})_{aj} = f_{aj} \in \mathbb{K}$.

Démonstration. Nous avons

$$f_{\mathbb{L}}(e_i) = \sum_a f_{ai}(1 \otimes e_a) = \sum_a f_{ai}\tau(e_a), \tag{8.581}$$

donc

$$(\mathtt{proj}\, f_{\mathbb{L}})e_i = \sum_a \mathtt{proj}(f_{ai})e_a = \sum_a f_{ai}e_a = f(e_i). \tag{8.582}$$

Cela prouve que $\mathtt{proj}\, f_{\mathbb{L}} = f$.

Par ailleurs,

EQooIOTFooNAdkit
$$f_{\mathbb{L}}(\tau(e_i)) = f_{\mathbb{L}}(1 \otimes e_i) = 1 \otimes f(e_i) = \tau\big(f(e_i)\big) = \sum_a f_{ai}\tau(e_a) \tag{8.583}$$

alors que par définition,

EQooMYSCooPFWATG
$$f_{\mathbb{L}}(\tau(e_i)) = \sum_a (f_{\mathbb{L}})_{ai}\tau(e_a). \tag{8.584}$$

Les éléments $\tau(e_a)$ formant une base[109], la comparaison de (8.583) avec (8.584) donne $(f_{\mathbb{L}})_{ai} = f_{ai} \in \mathbb{K}$. □

LEMooWZGSooONEnjZ

Lemme 8.281.
Soient des espaces vectoriels E, F et G ainsi que des applications linéaires

(1) $f\colon E \to F$,

(2) $\tilde{h}\colon G_{\mathbb{L}} \to E_{\mathbb{L}}$.

Alors nous avons

$$\boldsymbol{proj}(f_{\mathbb{L}} \circ \tilde{h}) = \boldsymbol{proj}(f_{\mathbb{L}}) \circ \boldsymbol{proj}(\tilde{h}). \tag{8.585}$$

De même[110] *si $f\colon E \to F$ et $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$, alors*

$$\boldsymbol{proj}(\tilde{h} \circ f_{\mathbb{L}}) = \boldsymbol{proj}(\tilde{h}) \circ \boldsymbol{proj}(f_{\mathbb{L}}). \tag{8.586}$$

Démonstration. Nous considérons une base $\{e_i\}$ de E, une base $\{e_a\}$ de F, et une base $\{e_\alpha\}$ de G. Pour écrire $\mathtt{proj}(f_{\mathbb{L}} \circ \tilde{h})$ à partir de la définition (8.576) nous commençons par écrire

$$(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a (f_{\mathbb{L}} \circ \tilde{h})_{a\alpha}e_a = \sum_{ai} (f_{\mathbb{L}})_{ai}(\tilde{h})_{i\alpha}e_a = \sum_a \Big(\sum_i f_{ai}(\tilde{h})_{i\alpha}\Big)e_a \tag{8.587}$$

où nous avons utilisé le fait que $(f_{\mathbb{L}})_{ai} = f_{ai}$. Donc, en utilisant la $\mathbb{K}$-linéarité de $\mathtt{proj}$,

EQooZGCGooQsCBQH
$$\mathtt{proj}(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a \sum_i \mathtt{proj}\Big(f_{ai}(\tilde{h})_{i\alpha}\Big)e_a = \sum_a \sum_i f_{ai}\,\mathtt{proj}\Big((\tilde{h})_{i\alpha}\Big)e_a. \tag{8.588}$$

D'autre part,

$$\begin{aligned}\mathtt{proj}(f_{\mathbb{L}}) \circ \mathtt{proj}(\tilde{h})e_\alpha &= \mathtt{proj}(f_{\mathbb{L}}) \sum_i \mathtt{proj}\Big((\tilde{h})_{i\alpha}\Big)e_i \\ &= \sum_i \mathtt{proj}\Big((\tilde{h})_{i\alpha}\Big) \sum_a f_{ai}e_a \\ &= \sum_{ai} \mathtt{proj}\Big((\tilde{h})_{i\alpha}\Big) f_{ai}e_a,\end{aligned} \tag{8.589}$$

et c'est égal à (8.588). □

109. Encore la proposition 8.274.

110. Je n'ai pas vérifié cette deuxième partie. Soyez prudente.

Remarque 8.282.
Nous n'avons en général pas $\texttt{proj}(xy) = \texttt{proj}(x)\,\texttt{proj}(y)$ pour tout $x, y \in \mathbb{L}$. Par exemple si $\mathbb{K} = \mathbb{R}$ et $\mathbb{L} = \mathbb{C}$ avec la projection canonique,

$$\texttt{proj}(i \cdot i) = \texttt{proj}(-1) = -1 \tag{8.590}$$

alors que $\texttt{proj}(i) = 0$.

PROPooHJJCooTXlzdH

Proposition 8.283.
Soient trois $\mathbb{K}$-espaces vectoriels E, F et G. Nous considérons deux applications linéaires $f\colon E \to F$ et $g\colon E \to G$.

Il existe une application linéaire $h\colon F \to G$ telle que $g = h \circ f$ si et seulement si il existe une application linéaire $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$ telle que $g_{\mathbb{L}} = \tilde{h} \circ f_{\mathbb{L}}$.

Démonstration. Dans le sens direct, il suffit de poser $\tilde{h} = h_{\mathbb{L}}$ et d'invoquer le lemme 8.272.

Dans le sens inverse, si nous avons $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$ tel que $g_{\mathbb{L}} = \tilde{h} \circ f_{\mathbb{L}}$ alors en appliquant $\texttt{proj}$ des deux côtés et en utilisant le lemme 8.281,

$$\texttt{proj}(g_{\mathbb{L}}) = \texttt{proj}(\tilde{h}) \circ \texttt{proj}(f_{\mathbb{L}}) \tag{8.591}$$

c'est-à-dire

$$g = \texttt{proj}(\tilde{h}) \circ f, \tag{8.592}$$

c'est-à-dire que l'application $\texttt{proj}(\tilde{h})\colon F \to G$ est la réponse à la proposition. □

8.15.3 Rang, polynôme minimal, polynôme caractéristique

PROPooJFQDooZSsxMf

Proposition 8.284 (Stabilité du rang par extension des scalaires[130])**.**
Si $f\colon E \to F$ est linéaire alors nous avons

$$\mathrm{rk}(f) = \mathrm{rk}(f_{\mathbb{L}}). \tag{8.593}$$

où à droite nous considérons le rang de l'application $\mathbb{L}$-linéaire $f_{\mathbb{L}}\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$.

Démonstration. Il existe un supplémentaire V tel que $E = \ker(f) \oplus V$ avec $\dim(V) = \mathrm{rk}(f)$. Nous pouvons factoriser f en

$$f = f_2 \circ f_1 \tag{8.594}$$

avec $f_1\colon E \to V$ est la projection parallèle à $\ker(f)$ et est surjective (vers V) parce que $\dim(V) = \mathrm{rk}(f) = \dim\big(\mathrm{Image}(f)\big)$. De plus $f_2\colon V \to F$ est injective parce que si $v \in V$ est tel que $f_2(v) = 0$ alors on aurait

$$f(v) = (f_2 \circ f_1)(v) = f_2(v) = 0. \tag{8.595}$$

Cela donne $v \in \ker(f) \cap V = \{0\}$. Par la proposition 8.273, les applications $(f_1)_{\mathbb{L}}$ et $(f_2)_{\mathbb{L}}$ sont respectivement surjective et injective.

L'application $(f_2)_{\mathbb{L}}\colon V_{\mathbb{L}} \to F_{\mathbb{L}}$ est forcément surjective sur son image, donc

$$(f_2)_{\mathbb{L}}\colon V_{\mathbb{L}} \to \mathrm{Image}(f_{\mathbb{L}}) \tag{8.596}$$

est un isomorphisme de $\mathbb{L}$-espaces vectoriels. Nous avons alors les égalités

$$\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{L}}\big(\mathrm{Image}(f_{\mathbb{L}})\big) = \mathrm{rk}(f_{\mathbb{L}}). \tag{8.597}$$

EQooWLOIooKlYWTU

Mais aussi, par les définitions posées plus haut,

$$\dim(V) = \mathrm{rk}(f) = \dim\big(\mathrm{Image}(f)\big). \tag{8.598}$$

EQooEVCGooAGimoU

Mais le corolaire 8.276 nous dit que $\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{K}}(V)$. Donc il y a égalité des deux lignes (8.597) et (8.598), ce qui donne $\mathrm{rk}(f) = \mathrm{rk}(f_{\mathbb{L}})$. □

PROPooZAZFooUFdCUv

Proposition 8.285.
Nous avons

(1) $\det(f) = \det(f_{\mathbb{L}})$

(2) $\chi_f = \chi_{f_{\mathbb{L}}}$.

Démonstration. Dès que l'on a des bases nous avons $(f_{\mathbb{L}})_{ai} = f_{ai}$ par la proposition 8.280(2). Le nombre $\det(f) \in \mathbb{K}$ est un polynôme en les f_{ai}. Entendons nous : il existe un polynôme indépendant de f et de $\mathbb{K}$ et de $\mathbb{L}$ donnant le déterminant de n'importe quelle matrice. Donc $\det(f) = \det(f_{\mathbb{L}})$.

Même chose pour le polynôme caractéristique (définition 8.113) : les coefficients de ce polynôme sont des polynômes en les f_{ai} qui sont indépendants de $\mathbb{L}$, de $\mathbb{K}$ et de f.

Notons que $\chi_{f_{\mathbb{L}}}$ est un polynôme à coefficients dans $\mathbb{K}$. □

La situation est très différente avec le polynôme minimal [111]. Autant il existe une « recette » pour créer le polynôme caractéristique, il n'en n'existe pas pour le polynôme minimal (ou en tout cas, il ne suffit pas d'appliquer des polynômes en les coefficients de la matrice). La proposition suivante montre que le polynôme minimal est conservé par extension de corps, mais que pour le savoir, il faut travailler plus.

PROPooXVZMooXcJrsJ

Proposition 8.286 ([130, 1])**.**
Soit $\mathbb{L}$ une extension du corps $\mathbb{K}$ et une application linéaire $f : E \to F$ entre deux $\mathbb{K}$-espaces vectoriels. Alors $\mu_f = \mu_{f_{\mathbb{L}}}$.

Démonstration. Nous allons montrer que l'application

$$\begin{aligned} \tilde{g} : \frac{\mathbb{L}[X]}{(\mu)} &\to \mathrm{End}(E_{\mathbb{L}}) \\ [P] &\mapsto P(f_{\mathbb{L}}) \end{aligned} \tag{8.599}$$

est bien définie et injective. La proposition 8.103 nous dira alors que μ est le polynôme minimal de $f_{\mathbb{L}}$.

Pour prouver que l'application $\tilde{g}$ est bien définie, nous commençons par prouver que $P(f_{\mathbb{L}}) = P(f)_{\mathbb{L}}$:

$$P(f_{\mathbb{L}})\lambda \otimes v = \sum_k a_k f_{\mathbb{L}}^k \lambda \otimes v \tag{8.600a}$$
$$= \lambda \otimes \sum_k a_k f^k(v) \tag{8.600b}$$
$$= \lambda \otimes P(f)v \tag{8.600c}$$
$$= P(f)_{\mathbb{L}}\lambda \otimes v. \tag{8.600d}$$

Par conséquent $\mu(f_{\mathbb{L}}) = 0$ et l'application est bien définie.

Sur $\mathbb{L}[X]/(\mu)$ nous considérons la base $\{1, [X], \dots, [X^{\deg(\mu)-1}]\}$, et sur $\mathrm{End}(E_{\mathbb{L}})$ nous considérons une base qui commence [112] par $\{f_{\mathbb{L}}^k\}_{k=0,\dots,\deg(\mu)-1}$. Montrons tout de même que cette partie est libre (sinon le théorème de la base incomplète ne s'applique pas) : si $\sum_k \lambda_k f_{\mathbb{L}}^k = 0$ alors

$$\sum_k \mathtt{proj}\left(\lambda_k f_{\mathbb{L}}^k\right) = 0. \tag{8.601}$$

EQooSFHVooLxqUEl

Pour détailler ce que cela implique, nous calculons ceci :

$$(\lambda f_{\mathbb{L}})(\tau(e_i)) = \lambda f_{\mathbb{L}}(\tau(e_i)) = \sum_a \lambda f_{ai} e_a, \tag{8.602}$$

111. Définition 6.64.
112. Théorème de la base incomplète 4.14(2).

par conséquent $\texttt{proj}(\lambda f_{\mathbb{L}})e_i = \sum_a \texttt{proj}(\lambda f_{ai})e_a$, et comme proj est $\mathbb{K}$-linéaire et que $f_{ai} \in \mathbb{K}$,

$$\texttt{proj}(\lambda f_{\mathbb{L}})e_i = \texttt{proj}(\lambda)\sum_a f_{ai}e_a = \texttt{proj}(\lambda)\,\texttt{proj}(f_{\mathbb{L}})e_i = \texttt{proj}(\lambda)f(e_i). \tag{8.603}$$

Appliquer la projection proj à l'équation (8.601) donne alors $\sum_k \texttt{proj}(\lambda)_k f^k = 0$. Mais comme les f^k sont linéairement indépendants sur $\mathbb{K}$ nous avons pour tout k : $\texttt{proj}(\lambda_k) = 0$ (égalité dans $\mathbb{K}$). En nous souvenant de la remarque 8.279 nous en déduisons $\lambda_k = 0$ dans $\mathbb{L}$.

Dans les choix de bases effectués, l'application $\tilde{g}$ a la forme

$$\tilde{g} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ * & * & * & \\ * & * & * & \\ * & * & * & \end{pmatrix}, \tag{8.604}$$

qui est injective.

Puisque $\tilde{g}$ est injective, μ est le polynôme minimal de $f_{\mathbb{L}}$ et donc $\mu = \mu_{\mathbb{L}}$. □

8.16 Frobenius et Jordan

8.16.1 Matrice compagnon

DEFooOSVAooGevsda

Définition 8.287.
*Soit le polynôme $P = X^n - a_{n-1}X^{n-1} - \ldots - a_1X - a_0$ dans $\mathbb{K}[X]$. La **matrice compagnon** de P est la matrice donnée par*

$$C(P) = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & a_{n-1} \end{pmatrix} \tag{8.605}$$

si $n \geqslant 2$ et par (a_0) si $n = 1$.

Une matrice est dite compagnon si elle a cette forme.

PROPooNDCLooYMGHmX

Proposition 8.288.
Si f est l'endomorphisme associé à la matrice $C(P)$ nous avons

$$f(e_i) = \begin{cases} e_{i+1} & \text{si } 1 \leqslant i < n \\ (a_0, \ldots, a_{n-1}) & \text{si } i = n. \end{cases} \tag{8.606}$$

De plus l'endomorphisme f vérifie $P(f)e_1 = 0$.

LemkVNisk

Lemme 8.289 ([298])**.**
Un polynôme sur un corps commutatif est le polynôme caractéristique de sa matrice compagnon. En d'autres termes nous avons $\chi_{C(P)} = P$.

Démonstration. Nous notons f l'endomorphisme associé à $C(P)$. La propriété $P(f)e_1 = 0$ nous indique que le polynôme minimal ponctuel de f en e_1 divise P. L'ensemble des puissances de f appliquées à e_1, $\big(f^k(e_1)\big)_{k=1,\ldots,n-1}$ est libre, donc le polynôme minimal ponctuel en e_1 est de degré n au minimum. En reprenant les notations du théorème 6.43, nous avons $I_{e_1} = (P)$ parce que P est de degré minimum dans I_{e_1} et $\chi_f \in I_{e_1}$.

Donc P divise χ_f et est de degré égal à celui de χ_f. Étant donné qu'ils sont tous deux unitaires, ils sont égaux. □

RemmQjZOA

Remarque 8.290.
Les matrices compagnons ne sont pas les seules dont le polynôme caractéristique est égal au polynôme minimal. En fait les matrices dont le polynôme caractéristique est égal au polynôme minimal sont denses dans l'ensemble des matrices. En effet une matrice dont le polynôme minimal n'est pas égal au polynôme caractéristique a un polynôme caractéristique avec une racine double. Il est possible, en modifiant arbitrairement peu la matrice, de séparer la racine double en deux racines distinctes.

8.16.2 Réduction de Frobenius

LEMooKUQDooKFeIYq

Lemme 8.291.
Soit un endomorphisme $f : E \to E$ sur l'espace vectoriel de dimension finie n. Nous notons μ et χ les polynômes minimal et caractéristique. Si f est cyclique, alors $\mu = \chi$.

Le théorème 8.304 donnera une version plus complète de ce lemme.

Démonstration. Soit v un vecteur cyclique de f, c'est-à-dire que $\{f^k(v)\}_{k=0,\dots,n-1}$ est libre. Donc si P est un polynôme de degré jusqu'à $n-1$ nous ne pouvons pas avoir $P(f) = 0$ parce que, appliqué à v, ce serait une combinaisons nulle non triviale des $f^k(v)$. Donc le polynôme minimal est au minimum de degré n. Mais le polynôme caractéristique est annulateur de degré n (Cayley-Hamilton 8.120), donc il est le polynôme minimal. □

THOooDOWUooOzxzxm

Théorème 8.292 (Réduction de Frobenius [299, 300, 301])**.**
Soit E, un $\mathbb{K}$-espace vectoriel, et $f \in \mathrm{End}(E)$. Alors il existe une suite de sous-espaces $E_1, \dots, E_r$ stables par f tels que

ItemmpwjnSs

(1) $E = \bigoplus_{i=1}^{r} E_i$;
(2) pour chaque E_i, l'endomorphisme restreint $f_i = f|_{E_i}$ est cyclique ;
(3) si μ_i est le polynôme minimal de f_i alors μ_{i+1} divise μ_i ;

Une telle décomposition vérifie automatiquement $\mu_1 = \mu_f$ et $\mu_1 \cdots \mu_r = \chi_f$, et la suite $(\mu_i)_{i=1,\dots,r}$ ne dépend que de f, et non du choix de la décomposition du point (1).

Les polynômes μ_i sont les **invariants de similitude** de l'endomorphisme f.

Démonstration. Nous commençons par montrer que si une telle décomposition existe, alors | subEqzcGouz

$$\chi_f = \prod_{i=1}^{r} \mu_i \tag{8.607a}$$ EqTaxsvb

$$\mu_f = \mu_1 \tag{8.607b}$$

où χ_f est le polynôme caractéristique de f et μ_f est le polynôme minimal. D'abord le polynôme caractéristique de f devra être égal au produit des polynômes caractéristiques des $f|_{E_i}$, mais ces derniers endomorphismes étant cycliques [113], leurs polynômes caractéristiques sont égaux à leurs polynômes minimaux (lemme 8.291). Cela prouve l'égalité (8.607a). Ensuite tous les μ_i doivent diviser le polynôme minimal, donc $\mathrm{ppcm}(\mu_1, \dots, \mu_r)$ divise μ_f. Cependant le polynôme minimal doit contenir une et une seule fois chacun des facteurs irréductibles du polynôme caractéristique, et chacun de ces facteurs sont dans les polynômes μ_i. Par conséquent $\mathrm{ppcm}(\mu_1, \dots, \mu_r) = \mu_f$. Mais par ailleurs $\mu_1 = \mathrm{ppcm}(\mu_1, \dots, \mu_r)$ parce qu'on a supposé $\mu_{i+1} \mid \mu_i$, donc $\mu_1 = \mu_f$.

Soit d, le degré du polynôme minimal de f et $y \in E$ tel que $\mu_f = \mu_{f,y}$ (voir lemme 8.104). Le plus petit espace stable sous f contenant y est

$$E_y = \mathrm{Span}\{y, f(y), \dots, f^{d-1}(y)\}. \tag{8.608}$$

113. Définition 8.105.

Nous notons $e_i = f^{i-1}(y)$. Notons que les vecteurs donnés forment bien une base de E_y parce que si les e_i n'était pas linéairement indépendants, alors nous aurions des a_k tels que $\sum_k a_k e_k = 0$ et avec lesquels

$$\big(\sum_k a_k X^k\big)(f)y = 0, \tag{8.609}$$

ce qui contredirait la minimalité de $\mu_{f,y}$.

La difficulté du théorème est de trouver un complément de E_y qui soit également stable sous f. Nous commençons par étendre [114] $\{e_1, \dots, e_d\}$ en une base $\{e_1, \dots, e_n\}$ de E. Ensuite nous allons montrer que

$$E = E_y \oplus F \tag{8.610}$$

avec

$$F = \{x \in E \text{ tel que } e_d^*\big(f^k(x)\big) = 0, \forall k \in \mathbb{N}\}. \tag{8.611}$$

qui fait appel aux définitions d'espace dual (définition 4.126) et de base duale (définition 4.127).

Par construction, F est invariant sous f. Montrons pour commencer que $E_y \cap F = \{0\}$. Un élément de E_y s'écrit

$$z = a_1 e_1 + \cdots + a_k e_k \tag{8.612}$$

avec $k \leqslant d$. Étant donné que f décale les vecteurs de base, nous avons $e_d^*\big(f^{d-k}(z)\big) = a_k$. Du coup $z \in F$ si et seulement si $a_1 = \ldots = a_d = 0$, c'est-à-dire que $E_y \cap F = \{0\}$.

Nous montrons maintenant que $\dim F = n - d$. Pour cela nous considérons l'application

$$\begin{aligned} T\colon \mathbb{K}[f] &\to E^* \\ g &\mapsto e_d^* \circ g. \end{aligned} \tag{8.613}$$

Cette application est injective. En effet un élément général de $\mathbb{K}[f]$ est

$$g = a_1 \operatorname{Id} + a_2 f + \cdots + a_p f^{p-1} \tag{8.614}$$

avec $p \leqslant d$. Si $T(g) = 0$, alors nous avons en particulier

$$0 = T(g)e_{d-p+1} = e_d^*(a_1 e_{d-p+1} + a_2 e_{d-p+2} + \cdots + a_p e_d) = a_p. \tag{8.615}$$

Donc $a_p = 0$ et en appliquant maintenant $T(g)$ à e_{d-p} nous obtenons $a_{p-1} = 0$. Au final nous trouvons que $g = 0$ et donc que T est injective.

Étant donné que $\dim \mathbb{K}[f] = d$ et que T est injective, $\dim \text{Image}(T) = d$. Étudions l'orthogonal [115] de l'image de T :

$$\begin{aligned} (\text{Image}(T))^\perp &= \{x \in E \text{ tel que } T(g)x = 0, \forall g \in \mathbb{K}[f]\} && \text{(8.616a)} \\ &= \{x \in E \text{ tel que } e_d^*\big(g(x)\big) = 0, \forall g \in \mathbb{K}[f]\} && \text{(8.616b)} \\ &= F. && \text{(8.616c)} \end{aligned}$$

Par conséquent $F^\perp = \text{Image}(T)$. Puisque $\dim \text{Image}(T) = d$, nous avons donc $\dim F = n - d$ et il est établi que $E = E_y \oplus F$.

Nous avons donc trouvé F, stable par f et tel que $E = E_y \oplus F$. Nous devons maintenant nous assurer que cette décomposition tombe bien pour les polynômes minimaux. Si P_1 est le polynôme minimal de $f|_{E_y}$, alors par le lemme 8.106 nous avons $P_1 = \mu_{f,y} = \mu_f$ parce que $f|_{E_y}$ est cyclique sur E_y. Mettons P_2, le polynôme minimal de $f|_F$. Étant attendu que F est stable par f, le polynôme P_2 divise P_1. En recommençant la construction sur F, nous construisons un nouvel espace F' stable sous F et vérifiant $\mu_{f|_{F'}} = P_2$, etc.

Nous passons maintenant à la partie unicité du théorème. Soient deux suites $F_1, \dots, F_r$ et $G_1, \dots, G_s$ de sous-espaces stables par f et vérifiant

(1) $E = \bigoplus_{i=1}^r F_i$,

114. Pour autant que j'aie compris, cette extension manque dans [299]. Corrigez-moi si je me trompe.
115. Définition 4.133.

(2) $f|_{F_i}$ est cyclique,

(3) $\mu_{f|_{F_{i+1}}}$ divise $\mu_{f|_{F_i}}$,

et, *mutatis mutandis*, les mêmes conditions pour la famille $\{G_i\}$. Nous posons $P_i = \mu_{f|_{F_i}}$ et $Q_i = \mu_{f|_{G_i}}$. Nous allons montrer par récurrence que $P_i = Q_i$ et $\dim F_i = \dim G_i$. Il ne sera cependant pas garanti que $F_i = G_i$. D'abord, $P_1 = Q_1$ parce qu'ils sont tous deux égaux à μ_f par les relations (8.607). Nous supposons que $P_i = Q_i$ pour $1 \leqslant i \leqslant j-1$ et nous tentons de montrer que $P_j = Q_j$.

Nous avons

$$P_j(f) = P_j(f)|_{F_1} \oplus \ldots \oplus P_j(f)|_{F_{j-1}}. \tag{8.617}$$

En effet étant donné que P_{j+k} divise P_j, il existe un polynôme A tel que $P_j = AP_{j+k}$, et donc tel que [116] $P_j(f) = A(f) \circ P_{j+k}(f)$, mais $P_{j+k}(f)F_{j+k} = 0$, donc $P_j(f)F_{j+k} = 0$. Les espaces G_i n'ayant à priori aucun rapport avec les polynômes P_i, nous écrivons

$$P_j(f) = P_j(f)|_{G_1} \oplus \ldots \oplus P_j(f)|_{G_{j-1}} \oplus P_j(f)|_{G_j} \oplus \ldots \oplus P_j(f)|_{G_s}. \tag{8.618}$$

Pour $1 \leqslant i \leqslant j-1$, nous avons supposé $P_i = Q_i$. Étant donné que la matrice de $f|_{F_i}$ est semblable à C_{P_i} et celle de $f|_{G_i}$ est semblable à C_{Q_i}, la matrice de $f|_{F_i}$ est semblable à la matrice de $f|_{G_i}$. En particulier,

$$\dim P_j(f)F_i = \dim P_j(f)G_i. \tag{8.619}$$

En prenant les dimensions des images dans les égalités (8.617) et (8.618), nous trouvons que

$$P_j(f)|_{G_j} = \ldots = P_j(f)|_{G_s} = 0. \tag{8.620}$$

Par conséquent P_j est annulateur de $f|G_j$, et il divise donc Q_j, qui est générateur de l'idéal des polynômes annulateurs de $f|_{G_j}$. La situation étant symétrique entre P et Q, nous montrons de même que Q_j divise P_j et donc que $P_j = Q_j$.

Ceci achève la démonstration du théorème de réduction de Frobenius.

□

Remarque 8.293.
Sous forme matricielle, ce théorème dit que toute matrice est semblable à une matrice de la forme bloc-diagonale

$$f = \begin{pmatrix} C_{\mu_1} & & \\ & \ddots & \\ & & C_{\mu_r} \end{pmatrix} \tag{8.621}$$

où les C_{μ_i} sont des matrices compagnons (définition 8.287).

En particulier, et ceci est très important, deux applications sont semblables si et seulement si elles ont même suite d'invariants de similitude.

Remarque 8.294.
Si nous travaillons sur $\mathbb{R}$, la réduite de Frobenius restera une matrice réelle, même si les valeurs propres sont complexes. En effet le procédé de Frobenius ne regarde absolument pas les valeurs propres, mais seulement les facteurs irréductibles du polynôme caractéristique. La réduite de Frobenius ne tente pas de résoudre ces polynômes, mais se contente d'en utiliser les matrices compagnon.

La situation sera différente dans le cas de la forme normale de Jordan.

8.16.3 Forme normale de Jordan

Il existe une preuve directe de la réduction de Jordan ne nécessitant pas la réduction de Frobenius[302]. Cette dernière passe par les espaces caractéristiques [117] et est à mon avis plus compliquée que la démonstration de Frobenius elle-même. Nous allons donc nous contenter de donner la réduction de Jordan comme un cas particulier de réduction de Frobenius.

116. En vertu du lemme 8.91.
117. Aussi appelés « espaces propres généralisés ».

ThoGGMYooPzMVpe

Théorème 8.295 (Réduction de Jordan)**.**
Soit E un espace vectoriel sur $\mathbb{K}$, et $f \in \mathrm{End}(E)$ un endomorphisme dont le polynôme caractéristique χ_f est scindé[118]*. Il existe une base de E dans laquelle la matrice de f s'écrit sous la forme*

$$M = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \\ 0 & \ddots & \\ & & J_{n_k}(\lambda_k) \end{pmatrix} \tag{8.622}$$

où les λ_i sont les valeurs propres de f (avec éventuelles répétitions) et $J_{n_i}(\lambda)$ représente le bloc $n_i \times n_i$

$$J_{n_i}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & & \\ & \lambda & 1 & 0 & \\ 0 & & \lambda & & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}. \tag{8.623}$$

En d'autres termes, $J_{n_i}(\lambda)_{ii} = \lambda$ et $J_n(\lambda)_{i-1,i} = 1$.

Démonstration. Nous commençons par le cas où f est nilpotente ; nous notons M sa matrice. Dans ce cas la seule valeur propre est zéro et le polynôme caractéristique est X^m pour un certain m. Nous savons par le lemme 8.289 que (la matrice de) f est semblable à sa matrice compagnon. En l'occurrence pour f nous avons

$$C_{X^m} = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix}. \tag{8.624}$$

Ensuite le changement de base (qui est une similitude) $(e_1, e_2, \dots, e_n) \mapsto (e_n, \dots, e_2, e_1)$ montre que C_{X^m} est semblable à un bloc de Jordan $J_m(0)$.

Supposons à présent que f ne soit pas nilpotente. Par l'hypothèse de polynôme caractéristique scindé, nous supposons que f a m valeurs propres distinctes et que son polynôme caractéristique est

$$\chi_f = (X - \lambda_1)^{l_1} \dots (X - \lambda_m)^{l_m}. \tag{8.625}$$

Le lemme des noyaux (théorème 8.92) nous enseigne que

$$E = \bigoplus_{i=1}^{m} \underbrace{\ker(f - \lambda_i \mathbb{1})^{l_i}}_{F_i}. \tag{8.626}$$

La restriction de $f - \lambda_i \mathbb{1}$ à F_i est par construction un endomorphisme nilpotent, et donc peut s'écrire comme un bloc de Jordan avec des zéros sur la diagonale. En utilisant la décomposition

$$f|_{F_i} = (f - \lambda_i \mathbb{1})|_{F_i} + \lambda_i \mathbb{1}_{F_i}, \tag{8.627}$$

nous voyons que $f|_{F_i}$ s'écrit comme un bloc de Jordan avec λ_i sur la diagonale. □

Remarque 8.296.
Nous pouvons calculer la forme normale de Jordan pour une matrice complexe ou réelle, mais dans les deux cas nous devons nous attendre à obtenir une matrice complexe parce que les valeurs propres d'une matrice réelle peuvent être complexes. Cependant nous demandons que le polynôme caractéristique de f soit scindé sur $\mathbb{K}$. En pratique, la décomposition de Jordan n'est garantie que sur les corps algébriquement clos, c'est-à-dire sur $\mathbb{C}$.

La suite des invariants de similitude sur laquelle repose la réduction de Frobenius, elle, est disponible sur tout corps, y compris $\mathbb{R}$.

118. C'est à cause de cette hypothèse que $\mathbb{K} = \mathbb{R}$ n'est pas le bon exemple de contexte de travail.

8.17 Commutant et endomorphismes cycliques

8.17.1 Endomorphisme cyclique

LemSGmdnE

Lemme 8.297.
Si A est la matrice de l'endomorphisme f alors nous avons équivalence des propriétés suivantes :

(1) La matrice A est cyclique.

(2) L'endomorphisme f est cyclique.

Si f est un endomorphisme de l'espace vectoriel E et si $x \in E$, nous notons

$$E_{f,x} = \text{Span}\{f^k(x) \text{ tel que } k \in \mathbb{N}\}. \tag{8.628}$$

DEFooIDMMooHIpUmT

Définition 8.298 (Commutant $\mathcal{C}(f)$ de f)**.**
*Soit E un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et un endomorphisme $f\colon E \to E$. Le **commutant** de f est l'ensemble des endomorphismes de E qui commutent avec f :*

$$\mathcal{C}(f) = \{g \in \mathcal{L}(E, E) \text{ tel que } g \circ f = f \circ g\}. \tag{8.629}$$

Il n'est pas très compliqué de vérifier que $\mathcal{C}(f)$ est un sous-espace vectoriel de $\mathcal{L}(E, E)$.

Notons l'inclusion évidente $\mathbb{K}[f] \subset \mathcal{C}(f)$. L'inclusion inverse va un peu nous occuper durant les prochaines pages.

8.17.2 Commutant : cas diagonalisable

PROPooRHHEooIRGmtl

Proposition 8.299 ([303])**.**
Si f est diagonalisable, alors

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda \in \text{Spec}(f)} \dim(E_\lambda)^2. \tag{8.630}$$

EQooOTFLooPUKAos

où les E_λ sont les espaces propres de f.

Démonstration. D'abord, soit $g \in \mathcal{C}(f)$ alors E_λ est stable par g. En effet si $v \in E_\lambda$ alors $f\big(g(v)\big) = g\big(f(v)\big) = g(\lambda v) = \lambda g(v)$, ce qui montre que $g(v)$ est un vecteur propre de f pour la valeur propre λ, et donc que $g(v) \in E_\lambda$.

Nous considérons ensuite l'application

$$\begin{aligned} \psi\colon \mathcal{C}(f) &\to \text{End}(E_1) \times \ldots \times \text{End}(E_r) \\ g &\mapsto g|_{E_1} \times \ldots \times g|_{E_r} \end{aligned} \tag{8.631}$$

qui est bien définie parce que g se restreint aux espaces propres de f. Nous allons noter $\psi(g)_\lambda$ la restriction de g à E_λ.

(1) ψ **est injective**
Supposons que $g, h \in \mathcal{C}(f)$ tels que $\psi(g) = \psi(h)$. Puisque f est diagonalisable nous pouvons décomposer $x \in E$ en ses composantes sur les espaces propres [119] :

$$x = \sum_{\lambda \in \text{Spec}(f)} x_\lambda \tag{8.632}$$

avec $x_\lambda \in E_\lambda$. Nous avons alors

$$g(x) = \sum_\lambda g(x_\lambda) = \sum_\lambda \psi(g)_\lambda(x_\lambda). \tag{8.633}$$

Par hypothèse nous avons $\psi(g)_\lambda = \psi(h)_\lambda$, et donc aussi

$$g(x) = \sum_\lambda \psi(g)_\lambda(x_\lambda) = \sum_\lambda \psi(h)_\lambda(x_\lambda) = \sum_\lambda h(x_\lambda) = h(x). \tag{8.634}$$

Cela prouve $g = h$ et donc que ψ est injective.

119. Théorème 8.219(5).

(2) ψ **est surjective** Si nous avons pour chaque $\lambda \in \mathrm{Spec}(f)$ un endomorphisme g_λ de E_λ alors en posant

$$g(x) = \sum_{\lambda \in \mathrm{Spec}(f)} g_\lambda(x_\lambda) \tag{8.635}$$

nous avons bien

$$\psi(g) = \big(g_{\lambda_1}, \dots, g_{\lambda_r}\big). \tag{8.636}$$

Nous pouvons donc conclure en écrivant

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda \in \mathrm{Spec}(f)} \dim\big(\mathrm{End}(E_\lambda)\big) = \sum_{\lambda \in \mathrm{Spec}(f)} \dim(E_\lambda)^2. \tag{8.637}$$

□

REMooUGFQooVzCOvV

Remarque 8.300.
Nous avons alors immédiatement

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{8.638}$$

lorsque f est diagonalisable.

En suivant la notation (8.211), un endomorphisme est cyclique lorsqu'il existe $x \in E$ tel que $E_x = E$.

PropooQALUooTluDif

Proposition 8.301 ([303])**.**
Si f est un endomorphisme diagonalisable d'un espace vectoriel E de dimension n. Nous avons équivalence des points suivants.

ITEMooSOYYooZVibjrii

(1) *Le polynôme minimal est égal au polynôme caractéristique : $\mu_f = \chi_f$*

ITEMooSOYYooZVibjrvi

(2) *L'endomorphisme f est cyclique.*

ITEMooSOYYooZVibjrv

(3) $\mathcal{C}(f) = \mathbb{K}[f]$ [120].

ITEMooSOYYooZVibjriv

(4) $\dim\big(\mathcal{C}(f)\big) = n$

ITEMooSOYYooZVibjriii

(5) *L'endomorphisme f possède n valeurs propres distinctes.*

ITEMooSOYYooZVibjri

(6) $\dim\big(\mathbb{K}[f]\big) = n$

Démonstration. Le point important de cette proposition réside dans les équivalences (1)-(3). Les autres sont des résultats intermédiaires. En particulier, dans le cas diagonalisable, nous allons voir que le point (5) est essentiellement une reformulation de (1).

(1) **(4) implique (5)** Par la formule (8.630), les espaces propres de f ont tous une dimension de 1. Par conséquent f possède n valeurs propres distinctes.

(2) **(5) implique (6)** Le théorème 8.219 nous dit que le polynôme minimal est scindé à racines simples. Puisque f possède n valeurs propres distinctes, μ est de degré n. Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 8.112 nous avons $\dim\big(\mathbb{K}[f]\big) = \deg(\mu) = n$ par la proposition 6.44.

(3) **(6) implique (1)** Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 8.112 et la proposition 6.44 nous avons $n = \dim\big(\mathbb{K}[f]\big) = \deg(\mu)$. Comme χ est un polynôme annulateur (Caley-Hamilton 8.120), il est divisé par μ. Maintenant μ et χ sont des polynômes unitaires de degré n et μ divise χ. Ils sont donc égaux.

(4) **(1) implique (2)** Le fait que f soit diagonalisable permet d'utiliser le théorème 8.219 pour dire que μ est scindé à racines simples. L'égalisation avec χ nous permet de dire que f possède

120. Rappel : $\mathcal{C}(f)$ est le commutant de f, définition 8.298.

n valeurs propres distinctes. Soient $\{e_1,\dots,e_n\}$ une base de diagonalisation, et prouvons que le vecteur $v=e_1+\cdots+e_n$ est cyclique. Nous avons

$$f^k(v)=\sum_{i=1}^{n}\lambda_i^k e_i. \tag{8.639}$$

Pour prouver que cette famille (avec $k=0,\dots,n-1$) est libre[121] nous en prenons une combinaison linéaire nulle et nous prouvons que les coefficients sont tous nuls. Soit donc

$$0=\sum_{l=0}^{n-1}a_l f^l(v)=\sum_{l=0}^{n-1}a_l\sum_{i=1}^{n}\lambda_i^l e_i=\sum_{i=1}^{n}\Big(\sum_{l=0}^{n-1}a_l\lambda_i^l\Big)e_i. \tag{8.640}$$

Par hypothèse, la double somme est nulle, et nous avons pour tout i :

$$\sum_{l=0}^{n-1}a_l\lambda_i^l=0. \tag{8.641}$$

En posant la matrice $A_{ij}=\lambda_i^j$, cela revient à étudier le système $\sum_j A_{ij}a_j=0$. Ce système n'a des solutions non nulles que si $\det(A)=0$; sinon il possède une unique solution qui est $a_j=0$ pour tout j. Nous devons donc calculer le déterminant

$$\det\begin{pmatrix}1&\lambda_1&\lambda_1^2&\cdots&\lambda_1^{n-1}\\ \vdots&\vdots&\vdots&&\vdots\\ 1&\lambda_n&\lambda_n^2&\cdots&\lambda_n^{n-1}\end{pmatrix}. \tag{8.642}$$

Il s'agit du déterminant de Vandermonde déjà étudié par la proposition 8.18. Nous avons $\det(A)=\prod_{1\leqslant i<j\leqslant n}(\lambda_j-\lambda_i)$. Ce déterminant est bien non nul parce que toutes les valeurs propres sont distinctes.

(5) **(2) implique (3)** Soit v un vecteur cyclique[122] de f. Si g est un endomorphisme, nous définissons des coefficients $(a_k(g))_{k=0,\dots,n-1}$ par

$$g(v)=\sum_{k=0}^{n-1}a_k(g)f^k(v). \tag{8.643}$$

C'est une bonne définition parce que $\{f^k(v)\}$ est une base.

Nous définissons alors l'application

$$\begin{aligned}\psi\colon\mathcal{C}(f)&\to\mathbb{K}[f]\\ g&\mapsto\sum_{k=0}^{n-1}a_k(g)f^k.\end{aligned} \tag{8.644}$$

C'est une application injective parce que si $\psi(g)=0$ alors $g(v)=0$ et pour tout k nous avons $g\big(f^k(v)\big)=f^k\big(g(v)\big)=0$. L'endomorphisme g s'annulant sur une base, il est nul.

L'application ψ est surjective. En effet si un polynôme $P=\sum_{k=0}^{n-1}a_kf^k$ est donné, il suffit de poser

$$g(x)=\sum_k a_k f^k(x) \tag{8.645}$$

pour avoir $\psi(g)=P$.

121. Ce sera alors une base parce que n vecteurs libres dans un espace de dimension n est toujours une base, proposition 4.19.

122. Définition 8.298.

(6) **(3) implique (4)** Si $n_1, \dots, n_r$ sont les dimensions des différents espaces propres de f, nous avons les inégalités

$$\dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant n = n_1 + \cdots + n_r \leqslant n_1^2 + \cdots + n_r^2 = \dim\big(\mathcal{C}(f)\big). \tag{8.646}$$

Par hypothèse d'égalité entre le premier et le dernier terme de cette suite d'inégalités, toutes les inégalités sont des égalités et en particulier $\dim\big(\mathcal{C}(f)\big) = n$.

Nous avons fini de prouver toutes les équivalences demandées. □

Exemple 8.302.
Pour mieux comprendre pourquoi le fait d'avoir n valeurs propres distinctes est équivalent à être cyclique, notons que si deux valeurs propres sont identiques, alors un morceau de la matrice de f serait par exemple $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, et dans ce cas n'importe quelle combinaison $ae_i + be_j$ reste proportionnelle à elle-même après application de f. Si nous avons des valeurs propres différentes par contre, nous avons par exemple dans $\mathbb{R}^2$ la matrice $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ qui donne $f(e_1 + e_2) = e_1 + 2e_2$. La partie $\{e_1 + e_2, e_1 + 2e_2\}$ est une base. △

8.17.3 Commutant : cas général

Nous considérons encore un espace vectoriel E de dimension finie n et un endomorphisme $f\colon E \to E$. Nous notons μ son polynôme minimal et μ_x le polynôme minimal ponctuel en x.

LEMooDFFDooJTQkRu

Lemme 8.303 ([304, 299, 130])**.**
Nous avons

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{8.647}$$

Démonstration. Si f est donnée, l'espace $\mathcal{C}(f)$ est l'espace des solutions de $fg = gf$. Supposons avoir choisi une base de E et notons A la matrice de f et X celle de g. L'équation est $AX - XA = 0$.

(1) **Si A est trigonalisable** Nous supposons avoir choisi la base de telle sorte que A soit triangulaire supérieure, et nous allons nous contenter de chercher les solutions X qui sont également triangulaires supérieure. Si il y en a déjà plus que n, a fortiori le résultat sera vrai.
Le produit de deux matrices triangulaires supérieures étant une matrice triangulaire supérieure, l'équation $AX - XA$ contient, pour les coefficients de X, $n(n+1)/2$ équations. Mais il se fait que les termes diagonaux ne sont pas de vraies équations parce que

$$(AX - XA)_{kk} = \sum_i \big(A_{ki}X_{ik} - X_{ki}A_{ik}\big) = \sum_{k \leqslant i \leqslant k} (A_{ki}X_{ik} - X_{ki}A_{ik}) = 0. \tag{8.648}$$

Nous avons donc au maximum

$$\frac{n(n+1)}{2} - n \tag{8.649}$$

équations linéairement indépendantes pour un minimum de $n(n+1)/2$ inconnues. L'espace des solutions est donc de dimension au minimum n.
Cela a l'air d'être une majoration assez large, mais il existe des cas d'égalité.

(2) **Si A n'est pas trigonalisable**
La preuve que nous donnons ici est valable même pour les endomorphismes trigonalisables.
Nous considérons le résultat de Frobenius 8.292. Nous avons donc la structure suivante :

- une décomposition en somme directe $E = E_1 \oplus \dots \oplus E_r$,
- les espaces E_i sont fixés par f,
- les endomorphismes $f_i = f|_{E_i}$ sont cycliques
- le polynôme minimal de f_i est μ_i et $\prod_{i=1}^r \mu_i = \chi_f$.

Les endomorphismes f_i^k commutent évidemment avec f_j, et la partie $\{f_i^k\}_{k=0,\dots,\deg(\mu_i)-1}$ est libre. Libre en tout cas en tant que partie de $\mathrm{End}(E_i)$. Mais en prolongeant par 0 sur E, ça reste libre en tant que partie de $\mathrm{End}(E)$.
Bien entendu les f_j^k et les f_i^k $(i \neq j)$ sont linéairement indépendants dans $\mathrm{End}(E)$ parce qu'ils n'agissent pas sur les mêmes vecteurs. Donc les endomorphismes $f_i^{k_i}$ avec $k_i = 0, \dots, \deg(\mu_i)-1$ forment une partie libre de $\mathrm{End}(E)$ composée d'endomorphismes qui commutent avec f. Il y en a en tout

$$\sum_{i=1}^{r} \deg(\mu_i) = \deg(\chi_f) = n. \tag{8.650}$$

Par conséquent $\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E)$.

□
THOooGLMSooYewNxW

Théorème 8.304 ([275]).
Soit un endomorphisme $f\colon E \to E$ sur l'espace vectoriel de dimension finie n. Nous notons μ et χ les polynômes minimal et caractéristique. Nous avons équivalence entre les points suivants :
ITEMooLRXIooLWaYqJi

(1) $\mu = \chi$,
ITEMooLRXIooLWaYqJii

(2) f *est cyclique,*
ITEMooLRXIooLWaYqJiii

(3) $\mathcal{C}(f) = \mathbb{K}[f]$.

Démonstration. Plusieurs implications. Notons que (2) implique (1) a déjà été démontré par le lemme 8.291.

(1) **(1) implique (2)** Conformément à ce que nous permet le lemme 8.104 nous choisissons [123] $a \in E$ de telle sorte à avoir $\mu_a = \mu$. De plus pour $x \in E$ nous considérons l'application

$$\begin{aligned} \varphi_x\colon \mathbb{K}[X] &\to E \\ P &\mapsto P(f)x. \end{aligned} \tag{8.651}$$

Nous avons $\varphi_a(P) = P(f)a$. Étant donné que E_a est engendré par les $f^k(a)$ nous avons $\varphi_a\big(\mathbb{K}[X]\big) = E_a$. De plus l'application φ_a passe aux classes pour (μ_a). Pour rappel, un élément de $\mathbb{K}[X]/(\mu_a)$ est de la forme

$$[P] = \{P + Q\mu_a\}_{Q\in\mathbb{K}[X]}. \tag{8.652}$$

Nous considérons donc l'application quotient

$$\begin{aligned} \psi\colon \frac{\mathbb{K}[X]}{(\mu_a)} &\to E_a \\ [P] &\mapsto \varphi_a(P), \end{aligned} \tag{8.653}$$

et nous prouvons que c'est un isomorphisme d'espace vectoriel.

(1a) **Linéaire** Parce que $(\lambda P + Q)(f) = (\lambda P)(f) + Q(f)$.

(1b) **Injectif** Si $\psi([P]) = 0$, alors $\varphi_a(P) = 0$, c'est-à-dire que $P \in \ker(\varphi_a)$. Mais, par définition 8.100 du polynôme minimal ponctuel, μ_a est générateur de $\ker(\varphi_a)$; donc il existe $Q \in \mathbb{K}[X]$ tel que $P = Q\mu_a$. Par conséquent $[P] = 0$, ce que nous voulions.

(1c) **Surjectif** Si $x \in E_a$ alors il existe des coefficients $x_k \in \mathbb{K}$ tels que $x = \sum_{k=0}^{\deg(\mu_a)-1} x_k f^k(a)$, c'est-à-dire $x = P(f)a = \varphi_a(P) = \psi([P])$.

Mais par hypothèse et par choix de a nous avons $\mu_a = \mu = \chi$, donc en fait $E_a = \mathbb{K}[X]/(\chi)$. Nous savons aussi que $\deg(\chi) = \dim(E)$ et que $\dim\big(\mathbb{K}[X]/P\big) = \deg(P)$ par la proposition 8.112. Au final nous avons $\dim(E_a) = \deg(\chi) = \dim(E)$. Et par conséquent $E_a = E$. Cela prouve que a est un vecteur cyclique pour f.

123. Dans toute la suite, nous devrions écrire μ_f et $\mu_{f,a}$ mais nous omettons d'indiquer explicitement la dépendance en f.

(2) **(2) implique (3)** Soit $g \in \mathcal{C}(f)$; nous devons prouver que g est un polynôme de f. Par hypothèse nous avons un vecteur cyclique que nous notons v. Nous avons un polynôme P (dépendant de g) tel que $g(v) = P(f)v$. Nous allons voir que $g = P(f)$. Soient $y \in E$ et Q un polynôme tels que $y = Q(f)v$; en notant que g commute avec $P(f)$ nous avons

$$g(y) = g\big(Q(f)v\big) = Q(f)\big(g(v)\big) = Q(f)\big(P(f)v\big) = P(f)Q(f)v = P(f)y. \tag{8.654}$$

Donc $g = P(f)$.

(3) **(3) implique (1)**
Nous avons les inégalités :

$$n \leqslant \dim\big(\mathcal{C}(f)\big) = \dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant \deg(\chi) = n. \tag{8.655}$$

La première inégalité est le lemme 8.303. Ensuite, toutes les inégalités se trouvent être des égalités. En particulier $\deg(\mu) = n$, ce qui signifie que $\mu = \chi$ parce que μ est un polynôme diviseur de χ, de même degré que χ et unitaire tout comme χ.

□

CORooAKQEooSliXPp

Corolaire 8.305 ([130]).
En suivant les notations sur les extensions de corps de base de la section 8.15, l'endomorphisme $f : E \to F$ est cyclique si et seulement si l'endomorphisme $f_{\mathbb{L}} : E_{\mathbb{L}} \to F_{\mathbb{L}}$ est cyclique.

Démonstration. Nous savons par le théorème 8.304 qu'un endomorphisme est cyclique si et seulement si son polynôme minimal est égal à son polynôme caractéristique. Or par les propositions 8.285 et 8.286, nous savons que ces polynômes sont identiques pour f et pour $f_{\mathbb{L}}$. □

THOooHUFBooReKZWG

Théorème 8.306 (Similitude et extension de corps[130]).
Les applications linéaires $f, g : E \to E$ sont semblables si et seulement si $f_{\mathbb{L}}$ et $g_{\mathbb{L}}$ le sont.

Démonstration. En ce qui concerne le sens direct, si il existe $m \in \mathrm{GL}(E)$ tel que $f = mgm^{-1}$ alors il suffit d'appliquer le lemme 8.281 pour avoir $f_{\mathbb{L}} = m_{\mathbb{L}} g_{\mathbb{L}} m_{\mathbb{L}}^{-1}$.

Nous considérons les invariants de similitude de f du théorème 8.292. Il existe une unique suite de polynômes unitaires μ_i $(i = 1, \dots, s)$ tels que $\mu_{i+1} \mid \mu_i$ et pour laquelle nous avons une décomposition $E = E_1 \oplus \dots \oplus E_s$ pour laquelle $f|_{E_i} : E \to E_i$ est cyclique et de polynôme minimal μ_i.

Nous avons aussi $E_{\mathbb{L}} = (E_1)_{\mathbb{L}} \oplus \dots \oplus (E_s)_{\mathbb{L}}$ et les $(E_i)_{\mathbb{L}}$ sont stables sous $f_{\mathbb{L}}$ qui y sera également cyclique (corolaire 8.305). De plus le polynôme minimal de $f_{\mathbb{L}}|_{(E_i)_{\mathbb{L}}}$ est également μ_i.

Autrement dit, la suite μ_i est également la suite des invariants de similitude de $f_{\mathbb{L}}$. La remarque 8.293 nous permet de conclure que f et g sont semblables si et seulement si $f_{\mathbb{L}}$ et $g_{\mathbb{L}}$ le sont. □

8.18 Hyperplans et formes linéaires

DEFooEWDTooQbUQws

Définition 8.307.
*Si E est un espace vectoriel de dimension n, un **hyperplan** de E est un sous-espace vectoriel de dimension $n - 1$.*

PROPooVYJUooAWDQrZ

Proposition 8.308 ([305]).
À propos d'hyperplans et de formes linéaires sur un espace vectoriel E sur le corps $\mathbb{K}$.

(1) Si φ est une forme linéaire non nulle, alors $\ker(\varphi)$ est un hyperplan.

(2) Si H est un hyperplan de E, il existe une forme linéaire dont H est le noyau :

$$H = \ker(\varphi). \tag{8.656}$$

Démonstration. En deux parties.

(1) Soit un supplémentaire A de H. Nous considérons la restriction $\varphi_A\colon A\to\mathbb{K}$. Vu que les éléments non nuls de A sont hors de H, nous avons $\varphi(x)\neq 0$ dès que x est non nul dans A. Cela implique que φ_A est surjective.

D'autre part, φ_A est également injective : si $\varphi_A(x)=\varphi_A(y)$, alors $\varphi_A(x-y)=0$, ce qui signifie que $x-y=0$ ou encore que $x=y$.

Donc φ_A est un isomorphisme de $\mathbb{K}$-espaces vectoriels ; nous en déduisons par le corolaire 4.45 que A est de dimension 1 sur $\mathbb{K}$, parce que $\mathbb{K}$ est de dimension 1.

(2) Nous utilisons le théorème de la base incomplète 4.14(4) pour considérer une base $\{e_i\}_{i=1,\dots,n}$ de E telle que $\text{Span}\{e_1,\dots,e_{n-1}\}=H$. Nous pouvons alors considérer la forme linéaire définie par

$$\varphi(e_i)=\begin{cases}0 & \text{si } i=1,\dots,n-1\\ 1 & \text{si } i=n.\end{cases}\tag{8.657}$$

Cette forme vérifie $\ker(\varphi)=H$.

□

Proposition 8.309 ([129]).
Soit un espace vectoriel E de dimension finie $n\geqslant 2$. Soit un sous-espace vectoriel V de E de dimension s. Alors V est une intersection de $n-s$ hyperplans de E.

Démonstration. Nous considérons une base de V que nous complétons[124] en une base de E : si $x=\sum_{i=1}^n x_ie_i$, nous avons $x\in V$ si et seulement si $x_{s+1}=\dots=x_n=0$. Nous considérons les formes linéaires

$$\begin{aligned}\varphi_i\colon E&\to\mathbb{R}\\ x&\mapsto x_i,\end{aligned}\tag{8.658}$$

et nous considérons les parties $H_i=\ker(\varphi_i)$ qui sont de hyperplans par la proposition 8.308. Les H_i avec $s+1\leqslant i\leqslant n$ sont une famille de $n-s$ hyperplans qui vérifient

$$V=\bigcap_{i=s+1}^{n}\ker(\varphi_i)\tag{8.659}$$

parce que $x\in\ker(\varphi_i)$ si et seulement si $x_i=0$.

Donc V peut être écrit comme intersection de $n-s$ hyperplans de E. □

PROPooRCLNooJpIMMl

Proposition 8.310 ([129]).
Soit un $\mathbb{K}$-espace vectoriel E de dimension finie $n\geqslant 2$. Si H_i sont des hyperplans de E, alors

$$\dim\Big(\bigcap_{i=1}^{m}H_i\Big)\geqslant n-m.\tag{8.660}$$

Démonstration. N'oubliez pas de prouver que $\bigcap_{i=1}^m H_i$ est un espace vectoriel. À part ça, nous faisons une petite récurrence.

(1) **Pour $m=2$** Nous savons déjà par la proposition 4.48 que

$$\dim(H_1+H_2)=\dim(H_1)+\dim(H_2)-\dim(H_1\cap H_2).\tag{8.661}$$

De plus $\dim(H_1+H_2)\leqslant n$. En remplaçant, par les valeurs,

$$\begin{aligned}\dim(H_1\cap H_2)&=\dim(H_1)+\dim(H_2)-\dim(H_1+H_2)&\text{(8.662a)}\\ &=n-1+n-1-\dim(H_1+H_2)&\text{(8.662b)}\\ &\geqslant 2n-2-n&\text{(8.662c)}\\ &=n-2.&\text{(8.662d)}\end{aligned}$$

Donc $\dim(H_1\cap H_2)\geqslant n-2$.

124. Théorème de la base incomplète, 4.14(4).

(2) **La récurrence** Nous calculons $\dim(H_1 \cap \ldots \cap H_m \cap H_{m+1})$ en commençant encore par la proposition 4.48 :

$$\dim(H_1 \cap \ldots \cap H_m \cap H_{m+1}) = \underbrace{\dim(H_1 \cap \ldots \cap H_m)}_{\leqslant n-m} + \dim(H_{m+1}) \tag{8.663a}$$

$$- \underbrace{\dim\big((H_1 \cap \ldots H_m) + H_{m+1}\big)}_{\leqslant n} \tag{8.663b}$$

$$\geqslant n - m + (n-1) - n \tag{8.663c}$$

$$= n - m - 1. \tag{8.663d}$$

C'est bon pour la récurrence.

□

8.18.1 Trouver la matrice d'une symétrie donnée

SubSecMtrSym

Les notions de déterminants, produit scalaire et vectoriels [125] donnent une bonne intuition géométrique des matrices. Nous pouvons alors chercher les matrices de quelques symétries dans $\mathbb{R}^2$ ou $\mathbb{R}^3$.

8.18.1.1 Symétrie par rapport à un plan

Comment trouver par exemple la matrice A qui donne la symétrie autour du plan $z = 0$? La définition d'une telle symétrie est que les vecteurs du plan $z = 0$ ne bougent pas, tandis que les vecteurs perpendiculaires changent de signe. Ces informations vont permettre de trouver comment A agit sur une base de $\mathbb{R}^3$. En effet :

(1) Le vecteur $\begin{pmatrix}1\\0\\0\end{pmatrix}$ est dans le plan $z = 0$, donc il ne bouge pas,

(2) le vecteur $\begin{pmatrix}0\\1\\0\end{pmatrix}$ est également dans le plan, donc il ne bouge pas non plus,

(3) et le vecteur $\begin{pmatrix}0\\0\\1\end{pmatrix}$ est perpendiculaire au plan $z = 0$, donc il va changer de signe.

Cela nous donne directement les valeurs de A sur la base canonique et nous permet d'écrire

$$A = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{pmatrix}. \tag{8.664}$$

Pour écrire cela, nous avons juste mis en colonne les images des vecteurs de base. Les deux premiers n'ont pas changé et le troisième a changé.

Et si maintenant on donne un plan moins facile que $z = 0$? Le principe reste le même : il faudra trouver deux vecteurs qui sont dans le plan (et dire qu'ils ne bougent pas), et puis un vecteur qui est perpendiculaire au plan [126], et dire qu'il change de signe.

Voyons ce qu'il en est pour le plan $x = -z$. Il faut trouver deux vecteurs linéairement indépendants dans ce plan. Prenons par exemple

$$f_1 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad f_2 = \begin{pmatrix}1\\0\\-1\end{pmatrix}. \tag{8.665}$$

EqfudE

125. Définitions 8.15, 8.168 et 10.28.

126. Pour le trouver, penser au produit vectoriel.

Nous avons

$$\begin{aligned} Af_1 &= f_1 \\ Af_2 &= f_2. \end{aligned} \tag{8.666}$$

Afin de trouver un vecteur perpendiculaire au plan, calculons le produit vectoriel :

$$f_3 = f_1 \times f_2 = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{vmatrix} = -e_1 - e_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}. \tag{8.667}$$

Nous avons

$$Af_3 = -f_3. \tag{8.668}$$

Afin de trouver la matrice A, il faut trouver Ae_1, Ae_2 et Ae_3. Pour ce faire, il faut d'abord écrire $\{e_1, e_2, e_3\}$ en fonction de $\{f_1, f_2, f_3\}$. La première des équations (8.665) dit que

$$f_1 = e_2. \tag{8.669}$$

Ensuite, nous avons

$$\begin{aligned} f_2 &= e_1 - e_3 \\ f_3 &= -e_1 - e_3. \end{aligned} \tag{8.670}$$

La somme de ces deux équations donne $-2e_3 = f_2 + f_3$, c'est-à-dire

$$e_3 = -\frac{f_2 + f_3}{2} \tag{8.671}$$

Et enfin, nous avons

$$e_1 = \frac{f_2 - f_3}{2}. \tag{8.672}$$

Maintenant nous pouvons calculer les images de e_1, e_2 et e_3 en faisant

$$\begin{aligned} Ae_1 &= \frac{Af_2 - Af_3}{2} = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \\ Ae_2 &= Af_1 = f_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\ Ae_3 &= -\frac{f_2 - f_3}{2} = -\frac{1}{2}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{8.673}$$

La matrice A s'écrit maintenant en mettant les trois images trouvées en colonnes :

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{8.674}$$

8.18.1.2 Symétrie par rapport à une droite

Le principe est exactement le même : il faut trouver trois vecteurs f_1, f_2 et f_3 sur lesquels on connaît l'action de la symétrie. Ensuite il faudra exprimer e_1, e_2 et e_3 en termes de f_1, f_2 et f_3.

Le seul problème est de trouver les trois vecteurs f_i. Le premier est tout trouvé : c'est n'importe quel vecteur sur la droite. Pour les deux autres, il faut un peu ruser parce qu'il faut impérativement qu'ils soient perpendiculaire à la droite. Pour trouver f_2, on peut écrire

$$f_2 = \begin{pmatrix} 1 \\ 0 \\ x \end{pmatrix}, \tag{8.675}$$

et puis fixer le x pour que le produit scalaire de f_2 avec f_1 soit nul. Si il n'y a pas moyen (genre si f_1 a sa troisième composante nulle), essayer avec $\begin{pmatrix} x \\ 1 \\ 0 \end{pmatrix}$. Une fois que f_2 est trouvé (il y a des milliards de choix possibles), trouver f_3 est super facile : prendre le produit vectoriel entre f_1 et f_2.

8.18.1.3 En résumé

La marche à suivre est

(1) Trouver trois vecteurs f_1, f_2 et f_3 sur lesquels on connaît l'action de la symétrie. Typiquement : des vecteurs qui sont sur l'axe ou le plan de symétrie, et puis des perpendiculaires. Pour la perpendiculaire, penser au produit scalaire et au produit vectoriel.
(2) Exprimer la base canonique e_1, e_2 et e_3 en termes de f_1, f_2, f_3.
(3) Trouver Ae_1, Ae_2 et Ae_3 en utilisant leur expression en termes des f_i, et le fait que l'on connaisse l'action de A sur les f_i.
(4) La matrice s'obtient en mettant les images des e_i en colonnes.

8.19 Théorème de Burnside

LemwXXzIt

Lemme 8.311.
Soit P, un polynôme sur $\mathbb{K}$. Une racine de P est une racine simple si et seulement si elle n'est pas racine de P'.

ThoBurnsideoPuCtS

Théorème 8.312.
Toute représentation[127] *d'un groupe abélien d'exposant fini sur $\mathbb{C}^n$ a une image finie.*

Démonstration. Étant donné que G est d'exposant fini, il existe $\alpha \in \mathbb{N}^*$ tel que $g^\alpha = e$ pour tout $g \in G$. Le polynôme $P(X) = X^\alpha - 1$ est scindé à racines simples. En effet tout polynôme sur $\mathbb{C}$ est scindé. Le fait qu'il soit à racines simples provient du lemme 8.311 parce que si $a^\alpha = 1$, alors il n'est pas possible d'avoir $\alpha a^{\alpha-1} = 0$.

Par ailleurs $P(g) = 0$. Le fait que nous ayons un polynôme annulateur de g scindé à racines simples implique que g est diagonalisable (théorème 8.219). Le fait que G soit abélien montre qu'il existe une base de $\mathbb{C}^n$ dans laquelle tous les éléments de G sont diagonaux. Nous devons par conséquent montrer qu'il existe un nombre fini de matrices de la forme

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \tag{8.676}$$

Nous savons que $\lambda_i^\alpha = 1$ parce que $g^\alpha = \mathbb{1}$, par conséquent chacun des λ_i est une racine de l'unité dont il n'existe qu'un nombre fini. □

ThooJLTit

Théorème 8.313 (Burnside[114, 306])**.**
Un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$ est fini si et seulement si il est d'exposant[128] *fini.*

Démonstration. Soit G un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$. Si G est fini, l'ordre de ses éléments divise $|G|$ (corolaire 2.14 au théorème de Lagrange) et l'exposant est le PPCM qui est donc fini également. Le théorème est déjà démontré dans un sens.

Dans l'autre sens, nous notons $e < \infty$ l'exposant de G, et nous allons prouver que l'ensemble G est fini. Nous commençons par remarquer que tous les éléments de G sont des racines du polynôme $X^e - 1$, et ensuite nous nous lançons dans le travail.

127. Définition 4.135.
128. Définition 1.271.

(1) **Générateurs**

Le groupe G est une partie de $\mathbb{M}(n,\mathbb{C})$ dont nous considérons l'algèbre engendrée[129] $\mathcal{G}$. Soit $C_1,\dots,C_r$ une famille génératrice de $\mathcal{G}$ constituée d'éléments de G et la fonction

$$\begin{aligned}\tau\colon G &\to \mathbb{C}^r\\ A &\mapsto \big(\mathrm{Tr}(AC_1),\dots,\mathrm{Tr}(AC_r)\big).\end{aligned} \tag{8.677}$$

(2) **τ est injective** Soient $A,B\in G$ tels que $\tau(A)=\tau(B)$. Si C_i est un générateur de G, nous avons $\mathrm{Tr}(AC_i)=\mathrm{Tr}(BC_i)$ et par la linéarité de la trace, nous avons

$$\mathrm{Tr}(AM)=\mathrm{Tr}(BM) \tag{8.678}$$ EqnCYmKW

pour tout $M\in G$. Notons par ailleurs

$$N=AB^{-1}-\mathbb{1}, \tag{8.679}$$

qui est diagonalisable parce que $AB^{-1}\in G$ et donc est annulé par le polynôme X^e-1 qui est scindé à racines simples. Du coup AB^{-1} est diagonalisable ; posons $PAB^{-1}P^{-1}=D$, alors $P\big(AB^{-1}-\mathbb{1}\big)P^{-1}=D-\mathbb{1}$ qui est encore diagonale. Donc N est diagonalisable.

Par ailleurs nous avons

$$\begin{aligned}\mathrm{Tr}\left((AB^{-1})^p\right)&=\mathrm{Tr}\left(AB^{-1}(AB^{-1})^{p-1}\right) && && (8.680a)\\ &=\mathrm{Tr}\left(BB^{-1}(AB^{-1})^{p-1}\right) && (8.678) && (8.680b)\\ &=\mathrm{Tr}\left((AB^{-1})^{p-1}\right). && && (8.680c)\end{aligned}$$

En continuant nous obtenons

$$\mathrm{Tr}\left((AB^{-1})^p\right)=\mathrm{Tr}(\mathbb{1})=n. \tag{8.681}$$

D'autre part,

$$N^k=(AB^{-1}-\mathbb{1})^k=\sum_{p=0}^{k}\binom{p}{k}(-1)^{k-p}(AB^{-1})^p \tag{8.682}$$

En prenant la trace, et en tenant compte du fait que $\mathrm{Tr}\left((AB^{-1})^p\right)=n$,

$$\mathrm{Tr}(N^k)=\sum_{p=0}^{k}\binom{p}{k}(-1)^{k-p}n=n(1-1)^k=0. \tag{8.683}$$

Donc la trace de N^k est nulle et le lemme 8.211 nous enseigne que N est alors nilpotente. Étant donné qu'elle est aussi diagonalisable, elle est nulle. Nous en concluons que $AB^{-1}=\mathbb{1}$ et donc que $A=B$. La fonction τ est donc injective.

(3) **Nombre fini de valeurs**

Les éléments de G sont annulés par X^e-1 qui est un polynôme scindé à racines simples. Dons le polynôme minimal d'un élément de G est (a fortiori) scindé à racines simples et le théorème 8.219 nous assure alors que ces éléments sont diagonalisables. Du coup les valeurs propres des matrices de G sont des racines $e$ièmes de l'unité. Par conséquent les traces des éléments de G ne peuvent prendre qu'un nombre fini de valeurs : toutes les sommes de n racines $e$ièmes de l'unité. Mais vu que les C_i sont dans G, nous avons

$$\mathrm{Image}(\tau)=\{\mathrm{Tr}(A)\text{ tel que }A\in G\}^r, \tag{8.684}$$

qui est un ensemble fini. Par conséquent G est fini parce que τ est injective.

□

129. Définition 1.349.

8.20 Ellipsoïde

LemYVWoohcjIX

Lemme 8.314.
Toute matrice peut être décomposée de façon unique en une partie symétrique et une partie antisymétrique. Cette décomposition est donnée par

$$S = \frac{M + M^t}{2}, \quad A = \frac{M - M^t}{2} \tag{8.685}$$

subEqHIQoqvhjWM

Démonstration. L'existence est une vérification immédiate de $S + A = M$ en utilisant (8.685). Pour l'unicité, si $S + A = S' + A'$ alors $S - S' = A - A'$. Mais $S - S'$ est symétrique et $A - A'$ est antisymétrique ; l'égalité implique l'annulation des deux membres, c'est-à-dire $S = S'$ et $A = A'$. □

DefOEPooqfXsE

Définition 8.315.
*Un **ellipsoïde** dans $\mathbb{R}^n$ centré en v est le lieu des points x vérifiant l'équation*

$$\langle x - v, M(x - v)\rangle = 1 \tag{8.686}$$

EqSNWooXfbTH

où M est une matrice symétrique strictement définie positive[130].

Lorsque nous parlons d'ellipsoïde plein*, il suffit de changer l'égalité en une inégalité.*

Remarque 8.316.
Le fait que M soit symétrique n'est pas tout à fait obligatoire ; la chose important est que toutes les valeurs propres soient strictement positives. En effet si M a toutes ses valeurs propres strictement positives, nous nommons S la partie symétrique de M et A la partie antisymétrique (lemme 8.314). Alors pour tout $x \in \mathbb{R}^n$ nous avons

$$x^tAx = \langle x, Ax\rangle = \langle A^tx, x\rangle = -\langle Ax, x\rangle = -\langle x, Ax\rangle, \tag{8.687}$$

donc $x^tAx = 0$. L'équation $x^tMx = 1$ est donc équivalente à $x^tSx = 1$ (elles ont les mêmes solutions).

De plus S reste strictement définie positive parce que pour tout $x \in \mathbb{R}^n$ nous avons

$$0 < x^tMx = x^tSx. \tag{8.688}$$

PropWDRooQdJiIr

Proposition 8.317.
Si $\mathcal{E}$ est un ellipsoïde centrée à l'origine, il existe une base de $\mathbb{R}^n$ dans laquelle son équation est :

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} = 1. \tag{8.689}$$

Démonstration. Nous avons une matrice symétrique strictement définie positive S telle que l'équation soit $\langle x, Sx\rangle = 1$. Le théorème spectral 8.227 nous fournit une base orthonormale $\{e_i\}$ dans laquelle $Se_i = \lambda_i e_i$ avec $\lambda_i > 0$. En substituant dans l'équation $\langle x, Sx\rangle = 1$ nous trouvons l'équation

$$\sum_i \lambda_i x_i^2 = 1. \tag{8.690}$$

En posant $a_i = \frac{1}{\sqrt{\lambda_i}}$, nous trouvons le résultat. Cette définition des a_i est toujours possible parce que $\lambda_i > 0$. □

CorKGJooOmcBzh

Corolaire 8.318.
Un ellipsoïde plein centré en l'origine admet une équation de la forme $q(x) \leqslant 1$ où q est une forme quadratique strictement définie positive.

130. Définition 8.230.

Pour rappel de notation, l'ensemble des formes quadratiques strictement définies positives sur l'espace vectoriel E est noté $Q^{++}(E)$.

Démonstration. Soit $\{e_i\}$ une base de $\mathbb{R}^n$ telle que l'ellipsoïde $\mathcal{E}$ ait pour équation

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} \leqslant 1. \tag{8.691}$$

Nous considérons la forme quadratique

$$\begin{aligned} q\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \sum_{i=1}^{n} \frac{\langle x, e_i\rangle^2}{a_i^2}. \end{aligned} \tag{8.692}$$

Nous avons évidemment $\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant 1\}$. De plus la forme q est strictement définie positive parce que dès que $x \neq 0$, au moins un des produits scalaires $\langle x, e_i\rangle$ est non nul et $q(x) > 0$. □

8.21 Système d'équations linéaires : méthode de Gauss

Pour résoudre un système d'équations linéaires, on procède comme suit :

(1) Écrire le système sous forme matricielle.

$$\text{p.ex. } \begin{cases} 2x + 3y &= 5 \\ x + 2y &= 4 \end{cases} \Leftrightarrow \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right)$$

(2) Se ramener à une matrice avec un maximum de 0 dans la partie de gauche en utilisant les transformations admissibles :

(2a) Remplacer une ligne par elle-même + un multiple d'une autre ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right) \overset{L_1 - 2.L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right)$$

(2b) Remplacer une ligne par un multiple d'elle-même ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right) \overset{-L_1 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 2 & 4 \end{array}\right)$$

(2c) Permuter des lignes.

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 0 & -2 \end{array}\right) \overset{L_1 \mapsto L_2' \text{ et } L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right)$$

(3) Retransformer la matrice obtenue en système d'équations.

$$\text{p.ex. } \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right) \Leftrightarrow \begin{cases} x &= -2 \\ y &= 3 \end{cases}$$

Remarque 8.319. — Si on obtient une ligne de zéros, on peut l'enlever :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \end{array}\right)$$

— Si on obtient une ligne de zéros suivie d'un nombre non-nul, le système d'équations n'a pas de solution :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{array}\right) \Leftrightarrow \begin{cases} \cdots \\ \cdots \\ 0x + 0y + 0z = 7 \end{cases} \Rightarrow \textbf{Impossible}$$

— Si on a moins d'équations que d'inconnues, alors il y a une infinité de solutions qui dépendent d'un ou plusieurs paramètres :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 1 & 0 & -2 & 2 \\ 0 & 1 & 3 & 0 \end{array}\right) \Leftrightarrow \begin{cases} x - 2z = 2 \\ y + 3z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 2 + 2\lambda \\ y = -3\lambda \\ z = \lambda \end{cases}$$

Chapitre 9

Analyse réelle : topologie et continuité

9.1 Intervalles

DEFooKRRYooZlwiEo

Définition 9.1 (Intervalle)**.**
*Une partie I de $\mathbb{R}$ est un **intervalle** si pour tout $a, b \in I$ nous avons $t \in I$ dès que $a \leqslant t \leqslant b$.*

PROPooJJRZooACUmWN

Proposition 9.2.
À propos d'intervalles.

Un intervalle[1] *est ouvert si il est de la forme $]a, b[$ avec éventuellement $a = -\infty$ ou $b = +\infty$.*

Un intervalle est fermé si il est de la forme $[a, b]$ ou $]-\infty, b]$ ou $[a, +\infty[$ avec $a, b \in \mathbb{R}$.

Remarque 9.3.
L'ensemble $\mathbb{R}$ ne contient pas $+\infty$ et $-\infty$. L'intervalle $[-\infty, 5]$ par exemple, n'est pas une partie de $\mathbb{R}$.

Exemple 9.4. (1) Les ensembles $]3, 7[$ et $]-\infty, \pi[$ sont des intervalles ouverts.

(2) Les ensembles $[10, 15]$ et $[-1, +\infty[$ sont des intervalles fermés.

(3) L'ensemble $]-4, -2[\cup]2, 9[$ n'est pas un intervalle (il y a un « trou » entre -2 et 2).

(4) L'ensemble $\mathbb{R}$ lui-même est un intervalle ; par convention, il est à la fois ouvert et fermé.

Un intervalle peut n'être ni ouvert ni fermé ; par exemple $]4, 8]$. Cet intervalle est « ouvert en 4 et fermé en 8 » . △

Définition 9.5 (Fonction, domaine, image, graphe)**.**
*Soient X et Y deux ensembles. Une **fonction** f définie sur X et à valeurs dans Y est une partie de $X \times Y$ telle que pour tout $x \in X$, il existe un unique $y \in Y$ tel que $(x, y) \in f$.*

9.6.
Notons qu'il n'est pas demandé que pour tout y, il existe x tel que $(x, y) \in f$. Autrement dit, la notation « $f : X \to Y$ » ne suppose pas que f est surjective sur Y. Mais elle doit être définie sur tout X.

Nous écrivons $y = f(x)$ pour dire $(x, y) \in f$.

— La partie de X qui contient tous les x sur lesquels f peut opérer est dite **domaine** de f. Le domaine de f est indiqué par $\text{Domaine}\, f$.

— L'élément de $y \in Y$ associé par f à un élément $x \in \text{Domaine}\, f$ (c'est-à-dire $f(x) = y$) est appellé **image** de x par f. L'**image** de la fonction f est la partie de Y qui contient les images de tous les éléments de $\text{Domaine}\, f$. L'image de f est indiquée par $\Im f$.

1. Définition 9.1.

— Le **graphe** de f est l'ensemble de tous les couples $(x, f(x))$ pour $x \in \text{Domaine } f$. Le graphe de f est une partie de l'ensemble noté $X \times Y$ et il est indiqué par $\text{Graph } f$. Dans ce chapitre $X = \mathbb{R}$ et $Y = \mathbb{R}$, donc le graphe de f est contenu dans le plan cartésien.

Définition 9.7 (Fonction croissante, décroissante et monotone)**.**
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction définie sur un intervalle $I \subset \mathbb{R}$.

*(1) La fonction f est **croissante** sur I si pour tout $x < y$ dans I nous avons $f(x) \leqslant f(y)$. Elle est* strictement *croissante si $f(x) < f(y)$ dès que $x < y$.*

*(2) La fonction f est **décroissante** sur I si pour tout $x < y$ dans I nous avons $f(x) \geqslant f(y)$. Elle est* strictement *décroissante si $f(x) > f(y)$ dès que $x < y$.*

*(3) La fonction f est dite **monotone** sur I si elle est, soit croissante, soit décroissante, sur I.*

Exemple 9.8.
La fonction $x \mapsto x^2$ est décroissante sur l'intervalle $]-\infty, 0]$ et croissante sur l'intervalle $[0, \infty[$. Elle n'est par contre ni croissante ni décroissante sur l'intervalle $[-4, 3]$. △

9.2 Application réciproque

9.2.1 Définitions

Les définitions d'injection, surjection, bijection et d'application réciproque sont les définitions 7.211 et 7.212.

EXooCWYHooLEciVj

Exemple 9.9. (1) La fonction $x \mapsto x^2$ n'est pas une bijection de $\mathbb{R}$ vers $\mathbb{R}$ parce qu'il n'existe aucun x tel que $x^2 = -1$.

(2) Nous verrons un peu plus tard (12.380) que la fonction

$$\begin{aligned} f : [0, +\infty[&\to [0, +\infty[\\ x &\mapsto x^2 \end{aligned} \tag{9.1}$$

est une bijection. Notez que c'est la même fonction que celle de l'exemple précédent. Seul l'intervalle sur lequel nous nous plaçons a changé.

(3) La fonction

$$\begin{aligned} f : \mathbb{R} &\to [0, \infty[\\ x &\mapsto x^2 \end{aligned} \tag{9.2}$$

n'est pas une bijection parce qu'il existe plusieurs x pour lesquels $f(x) = 4$.

En conclusion : il est très important de préciser les domaines des fonctions considérées. △

Remarque 9.10.
Dire que la fonction $f : I \to J$ est bijective, c'est dire que l'équation $f(x) = y$ d'inconnue x peut être résolue de façon univoque pour tout $y \in J$.

LEMooSDMMooYYDDLs

Lemme 9.11.
Toute fonction strictement monotone sur un intervalle I est injective.

Exemple 9.12.
Trouvons la fonction réciproque de la fonction affine $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto 3x - 2$. Si $y \in \mathbb{R}$, le nombre $f^{-1}(y)$ est la valeur de x pour laquelle $f(x) = y$. Il s'agit donc de résoudre

$$3x - 2 = y \tag{9.3}$$

par rapport à x. La solution est $x = \frac{y+2}{3}$ et donc nous écrivons

$$f^{-1}(y) = \frac{y + 2}{3}. \tag{9.4}$$

△

9.2.2 Graphe de la fonction réciproque

Par définition le graphe de la fonction f est l'ensemble des points de la forme (x, y) vérifiant $y = f(x)$. Afin de déterminer le graphe de la bijection réciproque nous pouvons faire le raisonnement suivant.

Le point (x_0, y_0) est sur le graphe de f

$\Leftrightarrow$

La relation $f(x_0) = y_0$ est vérifiée

$\Leftrightarrow$

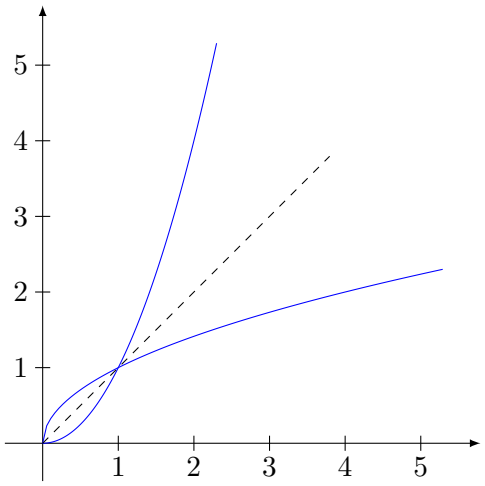
La relation $x_0 = f^{-1}(y_0)$ est vérifiée

$\Leftrightarrow$

Le point (y_0, x_0) est sur le graphe de f^{-1}.

À retenir 9.13
Dans un repère orthonormal, le graphe de la bijection réciproque est obtenu à partir du graphe de f en effectuant une symétrie par rapport à la droite d'équation $y = x$.

Le dessin suivant montre le cas de la courbe de la fonction carré comparé à celle de la racine carrée.



9.3 Topologie sur l'ensemble des réels

SECooGKHYooMwHQaD

Nous allons à présent donner la topologie sur $\mathbb{R}$ et ainsi résoudre les questions laissées en suspens lors de la construction des réels, voir 1.415.

Afin de pouvoir étudier la topologie des espaces métriques, il faut connaître quelques propriétés des réels, parce que nous allons étudier la fonction « distance » qui est une fonction continue à valeurs dans les réels.

La valeur absolue de la définition 1.379(2) permet de définir une norme sur $\mathbb{R}$.

LEMooBNAPooBTtXnX

Lemme 9.14.
L'application

$$x \mapsto |x| \tag{9.5}$$

est une norme[2] *sur* $\mathbb{R}$.

Démonstration. Grâce au lemme 1.383 et à la remarque 1.384, on a, pour tous $x,\ y,\ \lambda \in \mathbb{R}$:

(1) $|x| = 0$ implique $x = 0$,

2. Définition 7.158.

(2) $|\lambda x| = |\lambda||x|$,

(3) $|x + y| \leqslant |x| + |y|$,

et donc, les conditions de la définition 7.158 sont immédiatement vérifiées. □

DEFooNYGIooVGHSIA

Définition 9.15 (Topologie sur $\mathbb{R}$ et sur $\mathbb{Q}$)**.**
Le lemme 9.14 donne une norme sur $\mathbb{R}$ et $\mathbb{Q}$ à partir de la valeur absolue. La définition 7.117 donne alors une structure d'espace topologique. Hors cas rarissimes qui seront signalés, nous utiliserons toujours cette topologie sur $\mathbb{R}$ et sur $\mathbb{Q}$.

PropooUHNZooOUYIkn

Proposition 9.16.
Les rationnels sont denses dans les réels[3]*.*

Démonstration. Soient $x \in \mathbb{R}$ et $\epsilon \in \mathbb{R}^+$. Nous devons prouver l'existence d'un rationnel dans $B(x, \epsilon)$. Le lemme 1.441 dit qu'il existe un rationnel dans $]x-\epsilon/2, x+\epsilon/2[$ et donc dans $B(x, \epsilon)$. □

PropSLCUooUFgiSR

Proposition 9.17 ([1])**.**
Quel que soit le réel x, il existe une suite croissante de rationnels convergente vers x.

Démonstration. Soient $x \in \mathbb{R}$ et $\delta \in \mathbb{R}$; comme $x - \delta$ et x sont des réels, le lemme 1.441 donne un élément $q_\delta \in \mathbb{Q}$ tel que

$$x - \delta < q_\delta < x. \tag{9.6}$$

Il suffit alors de pêcher parmi ces q_δ pour trouver une suite croissante, et on montrera que cette suite converge vers x.

Soit x_0 un rationnel plus petit que x. Nous posons $\delta_0 = x - x_0$ et ensuite :

$$\begin{cases} \delta_i = x - x_i & \text{(9.7a)} \\ x_{i+1} = q_{\delta_i/2} \in \mathbb{Q}. & \text{(9.7b)} \end{cases}$$

Ainsi nous avons pour tout i les inégalités

$$x_i = x - \delta_i < x - \frac{\delta_i}{2} < x_{i+1} < x. \tag{9.8}$$

La suite (x_i) est donc une suite de rationnels, croissante et toujours plus petite que x. Mais nous avons à chaque étape $\delta_{i+1} < \frac{\delta_i}{2}$, ce qui implique que la suite des δ_i converge vers 0. Soit $\epsilon > 0$. Il existe k_0 tel que pour tout $k > k_0$, $\delta_k < \epsilon$. Pour un tel k, nous avons alors

$$x_{k+1} \in B(x, \frac{\delta_k}{2}) \subset B(x, \epsilon). \tag{9.9}$$

Tous les x_k, pour $k > k_0 + 1$, sont tels que $|x - x_k| < \epsilon$: la suite des x_k converge donc vers x. □

9.3.1 Compacité pour les réels

PROPooBFSAooKSugMj

Proposition 9.18.
Les parties compactes[4] *de $\mathbb{R}$ sont fermées et bornées.*

Démonstration. Prouvons d'abord qu'un ensemble compact est borné. Pour cela, supposons que K est un compact non borné vers le haut[5].

Nous posons $A_0 =]-\infty, 1[$ et $A_n =]n-1, n+1[$. Ces ouverts recouvrent $\mathbb{R}$ et donc K. Hélas tout choix d'un nombre fini des A_n est borné, et ne recouvre donc pas K. Les $\{A_i\}_{i=0,\dots,}$ forment un recouvrement de K par des ouverts dont on ne peut pas extraire un sous-recouvrement fini.

Cela prouve que K doit être borné.

3. Pour les topologies usuelles données en la définition 9.15.
4. Définition 7.80.
5. Nous laissons à titre d'exercice le cas où K est borné par le haut et pas par le bas.

Pour prouver que K est fermé, nous allons prouver que le complémentaire est ouvert. Et pour cela, nous allons prouver que si le complémentaire n'est pas ouvert, alors nous pouvons construire un recouvrement de K dont on ne peut pas extraire de sous-recouvrement fini.

Si $\mathbb{R}\backslash K$ n'est pas ouvert, il possède un point, disons x, tel que tout voisinage de x intersecte K. Soit $B(x,\epsilon_1)$, un de ces voisinages, et prenons $k_1 \in K \cap B(x,\epsilon_1)$. Ensuite, nous prenons ϵ_2 tel que k_1 ne soit pas dans $B(x,\epsilon_2)$, et nous choisissons $k_2 \in K \cap B(x,\epsilon_2)$. De cette manière, nous construisons une suite de $k_i \in K$ tous différents et de plus en plus proches de x. Prenons un recouvrement quelconque par des ouverts de la partie de K qui n'est pas dans $B(x,\epsilon_1)$. Les nombres k_i ne sont pas dans ce recouvrement.

Nous ajoutons à ce recouvrement les ensembles $\mathcal{O} =]k_i, k_{i+1}[$. Le tout forme un recouvrement (infini) par des ouverts dont il n'y a pas moyen de tirer un sous-recouvrement fini, pour exactement la même raison que la première fois. □

ThoBOrelLebesgue

Théorème 9.19 (Borel-Lebesgue)**.**
Un intervalle de $\mathbb{R}$ est compact si et seulement si il est de la forme $[a,b]$.

Démonstration. Tous les intervalles de $\mathbb{R}$ sont listés dans la proposition 1.466. Un compact est fermé et borné (proposition 9.18). Donc les intervalles dont une borne est $\pm\infty$ ne sont pas compacts. Parmi les intervalles $]a,b[$, $]a,b]$, $[a,b[$ et $[a,b]$, seul le dernier est fermé. Nous avons prouvé que si un intervalle est compact, alors il est de la forme $[a,b]$.

Nous prouvons à présent l'implication inverse : tous les intervalles de la forme $[a,b]$ sont compacts.

Soit Ω, un recouvrement du segment $[a,b]$ par des ouverts, c'est-à-dire que

$$[a,b] \subseteq \bigcup_{\mathcal{O}\in\Omega} \mathcal{O}. \tag{9.10}$$

Nous notons par M le sous-ensemble de $[a,b]$ des points m tels que l'intervalle $[a,m]$ peut être recouvert par un sous-ensemble fini de Ω. C'est-à-dire que M est le sous-ensemble de $[a,b]$ sur lequel le théorème est vrai. Le but est maintenant de prouver que $M = [a,b]$.

M **est non vide** En effet, $a \in M$ parce qu'il existe un ouvert $\mathcal{O} \in \Omega$ tel que $a \in \mathcal{O}$. Donc $\mathcal{O}$ tout seul recouvre l'intervalle $[a,a]$.

M **est un intervalle** Soient m_1, $m_2 \in M$. Le but est de montrer que si $m' \in [m_1,m_2]$, alors $m' \in M$. Il y a un sous-recouvrement fini de l'intervalle $[a,m_2]$ (par définition de $m_2 \in M$). Ce sous-recouvrement fini recouvre évidemment aussi $[a,m']$ parce que $[a,m'] \subseteq [a,m_2]$, donc $m' \in M$.

M **est un ensemble ouvert** Soit $m \in M$. Le but est de prouver qu'il y a un ouvert autour de m qui est contenu dans M. Admettons que Ω' soit un sous-recouvrement fini qui contienne l'intervalle $[a,m]$. Dans ce cas, on a un ouvert $\mathcal{O} \in \Omega'$ tel que $m \in \mathcal{O}$. Tous les points de $\mathcal{O}$ sont dans M, puisqu'ils sont tous recouverts par Ω'. Donc $\mathcal{O}$ est un voisinage de m contenu dans M.

M **est un ensemble fermé** M est un intervalle qui commence en a, en contenant a, et qui finit on ne sait pas encore où. Il est donc soit de la forme $[a,m]$, soit de la forme $[a,m[$. Nous allons montrer que M est de la première forme en démontrant que M contient son supremum s. Ce supremum est un élément de $[a,b]$, et donc il est contenu dans un des ouverts de Ω. Disons $s \in \mathcal{O}_s$. Soit c, un élément de $\mathcal{O}_s$ strictement plus petit que s ; étant donné que s est supremum de M, cet élément c est dans M, et donc on a un sous-recouvrement fini Ω' qui recouvre $[a,c]$. Maintenant, le sous-recouvrement constitué de Ω' et de $\mathcal{O}_s$ est fini et recouvre $[a,s]$.

Nous pouvons maintenant conclure : le seul intervalle non vide de $[a,b]$ qui soit à la fois ouvert et fermé est $[a,b]$ lui-même (proposition 7.74), ce qui prouve que $M = [a,b]$, et donc que $[a,b]$ est compact [6]. □

6. Si vous n'aimez pas le coup du fermé et ouvert, le lemme 9.20 donne une autre preuve.

LemOACGWxV

Lemme 9.20 ([307]).
Si $a < b \in \mathbb{R}$ alors le segment $[a, b]$ est compact[7].

Démonstration. Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de $[a, b]$ par des ouverts. Nous posons

$$M = \{x \in [a, b] \text{ tel que } [a, x] \text{ admet un sous-recouvrement fini extrait de } \{\mathcal{O}_i\}_{i\in I}\}. \tag{9.11}$$

Notre but est de prouver que $b \in M$.

(1) **a est dans M** Le point a est naturellement dans un des $\mathcal{O}_i$. L'intervalle $[a, a]$ est donc recouvert par un seul des $\mathcal{O}_i$.

(2) **M est un intervalle** Soient $m \in M$ et $m' \in [a, m[$. Le sous-recouvrement fini qui recouvre $[a, m]$ recouvre a fortiori $[a, m']$.

(3) **Les trois possibilités restantes** À ce niveau de la preuve, il reste trois possibilités pour M soit il est de la forme $[a, c]$ ou $[a, c[$ avec $c < b$, soit il est de la forme $[a, b]$. Nous allons maintenant éliminer les deux premiers cas.

(4) **Ce que M n'est pas** D'abord M n'est pas de la forme $[a, c[$ avec $c < b$. Par l'absurde, commençons par considérer $\mathcal{O}_{i_0}$ un ouvert du recouvrement qui contient c ; choisissons $m \in \mathcal{O}_{i_0}$ tel que $m < c$. Alors $m \in M$, et, si nous joignons $\mathcal{O}_{i_0}$ à un recouvrement fini de $[a, m]$ alors nous avons un recouvrement fini de $[a, c]$. On en déduit $c \in M$.

Ensuite M n'est pas de la forme $[a, c]$ avec $c < b$. En effet si on a un recouvrement fini de $[a, c]$ par des ouverts, alors un de ces ouverts contient c et donc contient des éléments de $[a, b]$ plus grands que c.

Nous déduisons que $M = [a, b]$ et qu'il est possible d'extraire un sous-recouvrement fini recouvrant $[a, b]$. □

LemCKBooXkwkte

Lemme 9.21 ([1]).
Si K_1 et K_2 sont des compacts dans $\mathbb{R}$ alors $K_1 \times K_2$ est compact dans $\mathbb{R}^2$.

Démonstration. Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de $K_1 \times K_2$ par des ouverts ; grâce au lemme 7.94 nous pouvons supposer que ce sont des carrés. Pour chaque $x \in K_1$, l'ensemble $\{x\} \times K_2$ est compact et donc recouvert par un nombre fini des $\mathcal{O}_i$. Soit R_x un ensemble fini des $\mathcal{O}_i$ recouvrant $\{x\} \times K_2$.

Comme R_x est une collection finie de carrés, nous pouvons considérer m_x, le minimum des rayons. L'ensemble K_1 est recouvert par les boules $B(x, m_x)$ et il existe donc une collection finie de $\{x_i\}_{i\in A}$ tels que $B(x_i, m_{x_i})$ recouvre K_1.

Alors $\{R_{x_i}\}_{i\in A}$ recouvre $K_1 \times K_2$ parce que R_{x_i} recouvre l'ensemble $B(x_i, m_{x_i}) \times \{K_2\}$. □

DEFooCNCJooWgCCrF

Définition 9.22 (Partie inductive[308, 309]).
*Soient $a < b$ dans $\mathbb{R}$. Une partie S de $[a, b]$ est **inductive** si*

(1) $a \in S$,

(2) Si $a \leqslant x < b$ et si $x \in S$ alors il existe $y > x$ tel que $[x, y] \subset S$.

(3) Si $a < x \leqslant b$ et si $[a, x[\subset S$, alors $x \in S$.

PROPooQVQKooIjWBrk

Proposition 9.23 ([308]).
L'unique partie inductive[8] *de $[a, b]$ est $[a, b]$ lui-même.*

Démonstration. Le fait que $[a, b]$ soit inductif dans $[a, b]$ est immédiat. Nous prouvons que c'est la seule. Soit S inductif dans $[a, b]$. Nous posons $S' = [a, b]\backslash S$. Si S' est non vide, nous pouvons considérer le nombre $\inf(S')$.

7. Définition 7.80
8. Définition 9.22.

(1) **Nous avons$\inf(S') > a$** D'abord nous savons que $\inf(S') \geqslant a$ parce que S' est une partie de $[a,b]$. De plus, vu que $a \in S$ il existe $y > a$ tel que $[a,y] \subset S$. Donc $S' \subset]y,b]$, de telle sorte que $\inf(S') \geqslant y > a$.

(2) **Si $\inf(S') \in S$** Si $\inf(S') = b$ alors $S = [a,b]$ et la preuve est faite. Supposons $\inf(S') < b$. Le nombre $\inf(S')$ vérifie $a < \inf(S') < b$ et $\inf(S') \in S$. Il existe donc $y > \inf(S')$ tel que $[\inf(S'), y] \subset S$. Nous avons alors $[a,y] \subset S$ et donc $\inf(S') > y$. Contradiction.

(3) **Si $\inf(S') \in S'$** Dans ce cas $[a, \inf(S')[\subset S$, et vu que S est inductif, $\inf(S') \in S$, contradiction.

Toutes les possibilités (à part $S' = \varnothing$) portant à des contradictions, nous déduisons que $S' = \varnothing$ et donc que $S = [a,b]$. □

9.3.2 Conséquence : les fermés bornés sont compacts

ThoXTEooxFmdI

Théorème 9.24 (Théorème de Borel-Lebesgue)**.**
Une partie d'un espace vectoriel normé réel de dimension finie est compacte si et seulement si elle est fermée et bornée.

Démonstration. Sens direct.

(1) **Compact implique borné** En effet si K est non borné dans E alors K contient une suite (x_n) avec $\|x_n\| > n$. Les boules $B_i(x_i, \frac{1}{3})$ sont disjointes. On pose $\mathcal{O}_0 = \complement \bigcup_i \overline{B(x_i, \frac{1}{5})}$, qui est ouvert comme complément d'un fermé. Pour $i \geqslant 1$ nous posons $\mathcal{O}_i = B(x_i, \frac{1}{4})$. Nous avons

$$K \subset \bigcup_{i \in \mathbb{N}} \mathcal{O}_i \tag{9.12}$$

mais puisque x_i est uniquement dans $\mathcal{O}_i$, nous ne pouvons pas extraire de sous-recouvrement fini.

(2) **Compact implique fermé** C'est le lemme 7.96(2).

Sens réciproque.

(1) **Un intervalle fermé et borné est compact dans $\mathbb{R}$** C'est le lemme 9.20.

(2) **Un produit de segments est compact** Le produit de deux compacts de $\mathbb{R}$ est un compact dans $\mathbb{R}^2$ par le lemme 9.21.

(3) **Un fermé et borné est compact** Soit K fermé et borné. Puisque K est borné, il est contenu dans un produit de segments. L'ensemble K est donc compact parce que fermé dans un compact, lemme 7.96.

□

Exemple 9.25 (Compacité de la boule unité)**.**
La boule unité fermée $\overline{B(0,1)}$ d'un espace vectoriel normé de dimension finie est compacte parce que fermée et bornée. En dimension infinie, cela n'est plus le cas. Certes la boule unité est encore fermée et bornée, mais elle n'est plus compacte. En effet nous allons donner un recouvrement par des ouverts duquel il ne sera pas possible d'extraire un sous-recouvrement fini.

Autour de chacune des extrémités des vecteurs de base, nous considérons la boule $A_i = B(e_i, \frac{1}{3})$. Ensuite nous considérons aussi l'ouvert

$$B(0,1) \backslash \bigcup_i \overline{B(e_i, \frac{1}{4})}. \tag{9.13}$$

Le tout recouvre $B(0,1)$ mais toutes les premières boules sont nécessaires. △

Le théorème de Bolzano-Weierstrass 7.290 nous permettra de prouver plus simplement la non compacité en dimension infinie. Voir l'exemple 7.145.

9.3.3 Suites et limites dans les réels

9.3.3.1 Limites, convergence

PROPooOSXCooJWXkWH

Proposition 9.26.
Une suite (x_n) dans un espace vectoriel normé E est convergente[9] *si et seulement si il existe un élément $\ell \in E$ tel que*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } n \geqslant N \Rightarrow \|x_n - \ell\| < \varepsilon. \tag{9.14}$$

EQooJRAMooEphgem

Dans ce cas, ℓ est la limite de la suite (x_n).

Démonstration. En deux parties.

(1) **Sens direct** Si $x_n \to \ell$ et si $\epsilon > 0$ il existe N_ϵ tel que pour tout $n \geqslant N$ nous avons $x_n \in B(\ell, \epsilon)$ (parce que cette boule est un ouvert contenant ℓ). Considérant la définition d'une boule, cette condition s'écrit bien $\|x_n - \ell\| < \epsilon$.

(2) **Sens inverse** Dans l'autre sens, soit $\mathcal{O}$ un ouvert contenant ℓ. Par définition de la topologie, il existe $\epsilon > 0$ tel que $B(\ell, \epsilon) \subset \mathcal{O}$. La condition (9.14) nous assure qu'il existe N_ϵ tel que pour tout $n \geqslant N_\epsilon$ nous ayons

$$x_n \in B(\ell, \epsilon) \subset \mathcal{O}, \tag{9.15}$$

ce qui assure que la suite (x_n) converge vers ℓ pour la topologie métrique de E.

□

Une façon équivalente d'exprimer le critère (9.14) est de dire que pour tout ϵ positif, il existe un rang $N \in \mathbb{R}$ tel que l'intervalle $[\ell - \epsilon, \ell + \epsilon]$ contient tous les termes x_n au-delà de N.

Il est à noter que le rang N dont il est question dans la définition de suite convergente dépend de ϵ.

9.3.4 Opérations sur les limites

PROPooIQOAooJPMoDD

Proposition 9.27 ([1])**.**
Soient des suites à valeurs réelles (a_i) et (b_j). Si elles sont convergentes, alors la suite ab est convergente et

$$\big(\lim_i a_i\big)\big(\lim_j b_j\big) = \lim_i (a_i b_i). \tag{9.16}$$

Démonstration. Nous nommons a et b les limites des suites (a_i) et (b_j). Soit $\epsilon > 0$ ainsi que $i \in \mathbb{N}$. Nous avons la majoration

$$|a_i b_i - ab| \leqslant |a_i b_i - a_i b| + |a_i b - ab| \tag{9.17a}$$
$$\leqslant |a_i||b_i - b| + |b||a_i - a|. \tag{9.17b}$$

Comme la suite (a_i) est convergente, elle est bornée[10]. Nous pouvons donc majorer $|a_i|$ par $R > 0$ qui ne dépend pas de i. Soit $\eta > 0$ tel que $(R + b)\eta < \epsilon$. Alors en prenant i assez grand pour que $|b_i - b| < \eta$ et $|a_i - a| < \eta$, nous avons bien

$$|a_i b_i - ab| \leqslant (R + b)\eta < \epsilon. \tag{9.18}$$

□

PROPooICZMooGfLdPc

Proposition 9.28.
Soient des suites (x_n) et (y_n) dans un espace vectoriel normé E. Si $x_n \xrightarrow{E} x$ et $y_n \xrightarrow{E} y$, alors

$$x_n + y_n \xrightarrow{E} x + y. \tag{9.19}$$

9. Définition 7.13.
10. Par 7.120. Attention : soyez capable d'adapter au cas présent.

Démonstration. Soit $\epsilon > 0$. Nous considérons N tel que si $n \geqslant N$, alors $\|x_n - x\| \leqslant \epsilon$ et $\|y_n - y\| \leqslant \epsilon$. En utilisant l'inégalité 7.158(4),

$$\|x_y + y_n - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\| \leqslant 2\epsilon. \tag{9.20}$$

Donc la suite $(x_n + y_n)$ converge vers $x + y$. □

LEMooGKIPooWgpFTB

Lemme 9.29.
La fonction

$$\begin{aligned} f\colon \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (x, y) &\mapsto x + y \end{aligned} \tag{9.21}$$

est continue.

Démonstration. Pour rappel, la topologie considérée sur $\mathbb{R}^n$ est celle de la définition 7.226. En vertu de la proposition 7.255, il est suffisant de prouver la continuité séquentielle. Soit donc une suite convergente

$$(x_n, y_n) \xrightarrow{\mathbb{R}\times\mathbb{R}} (x, y). \tag{9.22}$$

Nous devons prouver que

$$f(x_n, y_n) \xrightarrow{\mathbb{R}} f(x, y). \tag{9.23}$$

La proposition 7.65 nous permet de déduire la convergence composante par composante : $x_n \xrightarrow{\mathbb{R}} x$ et $y_n \xrightarrow{\mathbb{R}} x$. En permutant somme et limite (proposition 9.28) nous avons le calcul

$$f(x_n, y_n) = x_n + y_n \xrightarrow{\mathbb{R}} x + y = f(x, y). \tag{9.24}$$

D'où la convergence demandée. □

9.3.5 Exemples

LEMooNDSKooMsexOq

Lemme 9.30.
Quelques suites usuelles.

(1) La suite $x_n = \frac{1}{n}$ converge vers 0.

(2) La suite $x_n = (-1)^n$ ne converge pas.

9.3.6 Limites infinies

Deux limites pour voir comment ça fonctionne.

LEMooWCRSooWXVvcc

Lemme 9.31.
Si $r > 1$ nous avons :

(1) $\lim_{n\to\infty} r^n = \infty$.

(2) $\lim_{n\to\infty} \frac{r^n}{n} = \infty$.

Démonstration. Puisque $r > 1$ nous pouvons écrire $r = 1 + \delta$ avec $\delta > 0$. La formule du binôme de Newton (3.78) nous donne

$$(1 + \delta)^n = \sum_{k=0}^{n} \binom{n}{k} \delta^k > \binom{n}{1} \delta = n\delta. \tag{9.25}$$

La proposition 1.440 ($\mathbb{R}$ est archimédien) nous indique que $n\delta$ est arbitrairement grand lorsque n est grand, quelle que soit $\delta > 0$. Cela finit la preuve de la première limite.

Pour la seconde, nous posons $a_n = \frac{r^n}{n}$. Nous avons

$$\frac{a_{n+1}}{a_n} = \frac{n}{n+1} r. \tag{9.26}$$

Comme $\frac{n}{n+1} \to 1$, la suite $\frac{n}{n+1}r$ tend vers $r > 1$, et en particulier pour tout $\delta > 0$ tel que $r > 1+\delta$, il existe $N \in \mathbb{N}$ tel que, pour tout $n > N$,

$$\frac{n}{n+1}r > 1 + \delta. \tag{9.27}$$

Soit maintenant $k \in \mathbb{N}$. En utilisant un produit télescopique,

$$a_{N+k} = a_N \frac{a_{N+1}}{a_N}\frac{a_{N+2}}{a_{N+1}} \cdots \frac{a_{N+k}}{a_{N+k-1}} > a_N(1+\delta)^{k-1}. \tag{9.28}$$

Or $(1+\delta)^{k-1}$ tend vers ∞ lorsque $k \to \infty$ par le premier point. Donc nous avons $\lim_{n\to\infty} r^n/n = \infty$. □

DEFooEWRTooKgShmT

Définition 9.32.
*Nous disons que deux suites (u_n) et (v_n) sont **équivalentes** si il existe une application $\alpha\colon \mathbb{N} \to \mathbb{R}$ telle que*

(1) pour tout n à partir d'un certain rang, $u_n = v_n\alpha(n)$

(2) $\alpha(n) \to 1$.

9.3.7 Topologie et matrices

DEFooCQHDooYpUAhG

Lemme-Définition 9.33 (Topologie sur les matrices)**.**
Si $\mathbb{K}$ est un corps valué[11]*, alors l'opération*

$$\begin{aligned} \|.\|_{\mathbb{M}}\colon \mathbb{M}(n\times m,\mathbb{K}) &\to \mathbb{R}^+ \\ M &\mapsto \max_{kl} \|M_{kl}\|_{\mathbb{K}}. \end{aligned} \tag{9.29}$$

est une norme[12]*.*

*Cette norme est appelée **norme maximum** et nous considérons sur $\mathbb{M}(n\times m,\mathbb{K})$ la topologie associée à cette norme*[13]

PROPooOEETooPhqWuf

Proposition 9.34.
La multiplication matricielle

$$\begin{aligned} m\colon \mathbb{M}(n,s)\times\mathbb{M}(s,l) &\to \mathbb{M}(n,l) \\ (A,B) &\mapsto AB \end{aligned} \tag{9.30}$$

est une opération continue[14]*.*

Démonstration. D'après la proposition 7.225, nous pouvons tout faire avec la métrique produit. Supposons que $(A_k,B_k) \xrightarrow{\mathbb{M}(n,s)\times\mathbb{M}(s,l)} \mathbb{M}(n,l)$, et montrons que $\|m(A_k,B_k) - m(A,B)\| \to 0$. Notre hypothèse nous dit que

$$\|(A_k - A, B_k - B)\| \to 0, \tag{9.31}$$

c'est à dire (métrique produit)

$$\max\{\|A_k - A\|, \|B_k - B\|\} \to 0, \tag{9.32}$$

et donc $A_k \to A$ et $B_k \to B$ séparément. La définition 9.33 nous dit alors que pour tout (ij) nous avons

$$\begin{aligned} (A_k)_{ij} &\to A_{ij} \\ (B_k)_{ij} &\to B_{ij}. \end{aligned} \tag{9.33}$$

11. Définition 1.475.
12. Définition 7.158.
13. Définition 7.164.
14. Topologies de la norme 9.33.

Nous avons [SUBEQSooHPHLooXdtxYY]

$$\Big(m(A_k,B_k)-m(A,B)\Big)_{ij}=(A_kB_k)_{ij}-(AB)_{ij} \tag{9.34a}$$
$$=\sum_l\Big((A_k)_{il}(B_k)_{lj}-A_{il}B_{lj}\Big). \tag{9.34b}$$

Soit $\epsilon>0$. Par la proposition 9.27, pour chaque triple (i,j,l), il existe un $u_{ijk}\in\mathbb{N}$ tel que si $k\geqslant u_{ijl}$ nous avons

$$|(A_k)_{il}(B_k)_{lj}-A_{il}B_{lj}|<\epsilon. \tag{9.35}$$

Nous pouvons reprendre (9.34) avec un tel k. Nous avons :

$$\|(A_knB_k)-(A,B)\|<\sum_l\epsilon<s\epsilon. \tag{9.36}$$

□

9.3.8 Suites croissantes et bornées

Une suite est dite **contenue** dans un ensemble A si $x_n\in A$ pour tout n. Une suite est **bornée supérieurement** si il existe un M tel que $x_n\leqslant M$ pour tout n. De la même manière, la suite est bornée inférieurement si il existe un m tel que $x_n\geqslant m$ pour tout n.

Le lemme suivant est souvent utilisé pour prouver qu'une suite est convergente. Une version pour les fonctions $f:\mathbb{R}\to\mathbb{R}$ sera la proposition 12.36.

LemSuiteCrBorncv

Lemme 9.35 ([1]).
Une suite croissante et bornée supérieurement converge. Une suite décroissante bornée inférieurement est convergente.

Démonstration. Supposons que (x_n) soit une suite croissante non convergente. En particulier, par le théorème 7.287, cette suite n'est pas de Cauchy et il existe $\epsilon>0$ tel que pour tout $N\in\mathbb{N}$, il existe $p>N$ vérifiant

$$|x_p-x_N|>\epsilon. \tag{9.37}$$

Vu que $p>N$, et vu que la suite est croissante, nous pouvons récrire cette condition sous la forme $x_p\geqslant x_N+\epsilon$.

Nous définissons ainsi une application $p:\mathbb{N}\to\mathbb{N}$ telle que

$$x_{p(N)}>x_N+\epsilon. \tag{9.38}$$

Une telle application n'est pas du tout unique, mais nous en considérons une telle.

Il est vite vu par récurrence que

$$x_{p^k(0)}\geqslant x_0+k\epsilon. \tag{9.39}$$

La suite $n\mapsto x_{p^n(0)}$ est donc une sous-suite qui tend vers l'infini. Or cela n'est pas possible parce que la suite (x_n) est bornée. Donc contradiction, donc (x_n) est convergente. □

Une erreur courante est de croire que la borne est la limite : le lemme n'affirme pas ça. Par contre il est vrai que la borne donne ...hum ...une borne inférieure (ou supérieure) pour la limite.

PropCvRpComposante

Proposition 9.36.
Une suite (x_n) dans $\mathbb{R}^m$ est convergente dans $\mathbb{R}^m$ si et seulement si les suites de chaque composante sont convergentes dans $\mathbb{R}$. Dans ce cas nous avons

$$\lim x_n=\Big(\lim(x_n)_1,\lim(x_n)_2,\ldots,\lim(x_n)_m\Big) \tag{9.40}$$

où $(x_n)_k$ dénote la k-ième composante de (x_n).

Exemple 9.37.
La suite $x_n = \left(\frac{1}{n}, 1 - \frac{1}{n}\right)$ converge vers $(0,1)$ dans $\mathbb{R}^2$. En effet, en utilisant la proposition 9.36, nous devons calculer séparément les limites

$$\begin{aligned} \lim \frac{1}{n} &= 0 \\ \lim \left(1 - \frac{1}{n}\right) &= 1. \end{aligned} \tag{9.41}$$

△

Exemple 9.38.
Étant donné que la suite $(-1)^n$ n'est pas convergente, la suite $x_n = \left((-1)^n, \frac{1}{n}\right)$ n'est pas convergente dans $\mathbb{R}^2$. △

9.3.9 Suites adjacentes

DEFooDMZLooDtNPmu

Définition 9.39 ([310]).
*Les suites (a_n) et (b_n) sont **adjacentes** si l'une est croissante, l'autre décroissante et si $a_n - b_n \to 0$.*

THOooZJWLooAtGMxD

Théorème 9.40 (Théorème des suites adjacentes).
Nous considérons des suites adjacentes (a_n) et (b_n) avec (a_n) croissante et (b_n) décroissante. Alors

(1) $b_n \geqslant a_n$ pour tout n,

(2) $a_n \leqslant b_q$ pour tout n et q. C'est-à-dire que toute la suite a est plus petite que toute la suite b.

(3) les suites a et b sont convergentes,

(4) les suites a et b convergent vers la même limite, notée ℓ,

(5) nous avons $a_n \leqslant \ell \leqslant b_n$ pour tout n.

Démonstration. La suite $n \mapsto b_n - a_n$ est décroissante parce que $b_n - a_n \geqslant b_{n+1} - a_{n+1}$. Comme en plus $b_a - a_n \to 0$ nous avons

$$b_n - a_n \geqslant 0 \tag{9.42}$$

pour tout $n \in \mathbb{N}$. De plus $a_n \leqslant b_0$ pour tout n parce que si $a_N > b_0$ alors, b étant décroissante, $a_N > b_0 \geqslant b_N$ qui est contraire à ce que nous venons de prouver. La suite a étant croissante et majorée, elle est convergente [15] ; notons ℓ sa limite.

La suite b peut maintenant être écrite par

$$b_n = (b_n - a_n) + a_n \tag{9.43}$$

qui est une somme de deux suites convergentes. Elle est donc convergente et sa limite est la somme des limites [16], donc

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (b_n - a_n) + \lim_{n\to\infty} a_n = 0 + \ell = \ell. \tag{9.44}$$

Voilà. Donc les suites a et b convergent et ont la même limite.

Pour tout $n, q \in \mathbb{N}$ nous avons l'inégalité $a_n \leqslant b_q$. En prenant la limite $n \to \infty$ nous trouvons

$$\ell \leqslant b_q \tag{9.45}$$

pour tout q. Et de la même façon, $b_n \geqslant a_q$ donne $\ell \geqslant a_q$. L'un avec l'autre donne

$$a_q \leqslant \ell \leqslant b_q \tag{9.46}$$

pour tout $q \in \mathbb{N}$. □

15. Proposition 9.35.
16. Proposition 9.28.

PROPooXOOCooGMqJNe

Proposition 9.41 ([311]).
Soit une suite (a_n) dans $\mathbb{R}$. Nous supposons que les suites extraites (a_{2n}) et (a_{2n+1}) convergent vers la même limite notée ℓ.

Alors $a_n \to \ell$.

Démonstration. Soit $\epsilon > 0$. Il existe N_1 tel que $|a_{2n} - \ell| \leqslant \epsilon$ dès que $n \geqslant N_1$. Il existe également N_2 dès que $|a_{2n+1} - \ell| \leqslant \epsilon$ dès que $n \geqslant N_2$.

Nous posons $N = \max\{2N_1, 2N_2 + 2\}$ et nous avons, pour tout $n \geqslant N$:

$$|a_n - \ell| \leqslant \epsilon, \tag{9.47}$$

c'est-à-dire que $a \to \ell$. □

9.3.10 Limite supérieure et inférieure

ooMVZAooVVCOnP

Lemme-Définition 9.42.
Soit (a_n) une suite dans $\bar{\mathbb{R}}$. Les limites suivantes existent dans $\bar{\mathbb{R}}$

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \big(\sup_{k\geqslant n} a_k\big) \tag{9.48}$$

et

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \big(\inf_{k\geqslant n} a_k\big). \tag{9.49}$$

*Elles sont nommées **limite supérieure** et **limite inférieure** de la suite (a_n).*

Démonstration. Pour la limite supérieure, l'ensemble des $k \geqslant n$ est de plus en plus petit lorsque n grandit. Donc les ensembles $A_n = \{a_k \text{ tel que } k \geqslant n\}$ sont emboîtés et la suite $n \to \sup A_n$ est une suite décroissante. Elle a donc une limite dans $\bar{\mathbb{R}}$. □

ooEEQJooRMFzVR

9.43.
En ce qui concerne les suites d'ensembles, utiles en théorie des probabilités, nous définissons de même. Si les A_n sont des parties d'un ensemble Ω, nous définissons la **limite supérieure** et la **limite inférieure** de la suite A_n par

$$\limsup_{n\to\infty} A_n = \bigcap_{n\geqslant 1} \bigcup_{k\geqslant n} A_k \tag{9.50}$$

et

$$\liminf_{n\to\infty} A_n = \bigcup_{n\geqslant 1} \bigcap_{k\geqslant n} A_k \tag{9.51}$$

Nous avons

$$\limsup A_n = \{\omega \in \Omega \text{ tel que } \omega \in A_n \text{ pour une infinité de } n\}. \tag{9.52}$$

ooAQTEooYDBovS

Lemme 9.44.
Nous avons les formules pratiques suivantes :

$$\limsup a_n = \inf_{n\geqslant 1} \big(\sup_{k\geqslant n} a_k\big) \tag{9.53a}$$

$$\liminf a_n = \sup_{n\geqslant 1} \big(\inf_{k\geqslant n} a_k\big). \tag{9.53b}$$

Démonstration. La suite $n \mapsto \sup_{k\geqslant n} a_k$ est une suite décroissante, donc la limite est l'infimum. Même argument pour l'autre. □

ooIQIKooXWwAmM

Lemme 9.45.
La suite (a_n) dans $\mathbb{R}$ converge si et seulement si

$$\limsup a_n = \liminf a_n. \tag{9.54}$$

Dans ce cas, $\lim a_n = \limsup a_n = \liminf a_n$.

Démonstration. Nous commençons par supposer que $\limsup a_n = \liminf a_n = l$, et nous prouvons que $\lim a_n$ existe et vaut l. Soit $\epsilon > 0$. Il existe N tel que si $n \geqslant N$ nous avons

$$\left| \sup_{k\geqslant n} a_k - l \right| < \epsilon \tag{9.55}$$

et

$$\left| \inf_{k\geqslant n} a_k - l \right| < \epsilon. \tag{9.56}$$

Si $i \geqslant N$, alors [17] $a_i \leqslant \sup_{k\geqslant N}(a_k) \leqslant \ell + \epsilon$, et $a_i \geqslant \inf_{k\geqslant N}(a_k) \geqslant \ell - \epsilon$. Cela signifie que $a_n \in B(l, \epsilon)$, c'est-à-dire $a_k \to l$ par la proposition 9.26.

Dans l'autre sens, nous supposons que $\lim_n a_n = l$ et nous prouvons que la limite supérieure est égale à l [18]. Soit $\epsilon > 0$ et N_ϵ tel que $|a_n - l| < \epsilon$ pour tout $n \geqslant N_\epsilon$. Si $n \geqslant N_\epsilon$ nous avons

$$\left| \sup_{k\geqslant n} a_k - l \right| \leqslant \epsilon \tag{9.57}$$

et donc la limite de $\sup_{k\geqslant n} a_k$ lorsque $n \to \infty$ est bien l. □

LEMooHGJVooCbgOEK

Lemme 9.46.
Soit une suite (a_i) dans $\mathbb{R}$. Notons $L = \limsup_i(a_i)$. Pour tout $\epsilon > 0$, l'ensemble

$$S_\epsilon = \{n \in \mathbb{N} \text{ tel que } a_n \geqslant L + \epsilon\} \tag{9.58}$$

est fini.

Démonstration. Nous y allons par récurrence. Juste pour le sport, nous allons au passage montrer en détail comment on utilise le théorème 1.48.

Supposons que S_ϵ est infini. Alors pour tout n, la partie $S_\epsilon \backslash \{0, \dots, n\}$ est non vide (lemme 1.127) et nous pouvons considérer l'application

$$\begin{aligned} g\colon S_\epsilon &\to S_\epsilon \\ n &\mapsto \min\big(S_\epsilon \backslash \{0, \dots, n\}\big). \end{aligned} \tag{9.59}$$

Nous prenons $b > 1$ dans S_ϵ et considérons la fonction $f\colon \mathbb{N} \to S_\epsilon$ donnée par le théorème 1.48.

L'application f est strictement croissante parce que $f(n+1) = g\big(f(n)\big) \in \mathbb{N}\backslash\{0, \dots, f(n)\}$. En particulier $f(n) > n$ parce que nous avons décidé de commencer avec $f(0) = b > 1$.

Nous sommes maintenant armés pour contredire la définition 9.42 de la limite supérieure. Soit $n \in \mathbb{N}$. Vu que $f(n) \in S_\epsilon$ nous avons

$$a_{f(n)} \geqslant L + \epsilon, \tag{9.60}$$

et donc $\sup_{k\geqslant n} a_k \geqslant a_{f(n)} \geqslant L + \epsilon$ parce que $f(n) \geqslant n$.

Nous avons prouvé que $\sup_{k\geqslant n} a_k \geqslant L + \epsilon$ pour tout n, donc

$$\lim_{n\to\infty} \big(\sup_{k\geqslant n} a_k\big) \geqslant L + \epsilon > L. \tag{9.61}$$

Voila. Donc si S_ϵ est infini, $\limsup_i a_i \geqslant L + \epsilon > L$. □

17. Voir le lemme 1.385(1).
18. Je vous laisse faire la démonstration correspondante pour la limite inférieure. Contactez-moi si ça pose un problème.

LEMooMTRDooBMxFmn

Lemme 9.47.
Si (a_n) est une suite réelle,

$$\limsup_{n\to\infty}(a_n) = -\liminf_{n\to\infty}(-a_n). \tag{9.62}$$

Démonstration. En utilisant la proposition 1.463, nous avons

$$\limsup(a_n) = \lim_{n\to\infty}\Big(\sup_{k\geqslant n}(a_k)\Big) = \lim_{n\to\infty}\Big(-\inf_{k\geqslant n}(-a_k)\Big) = -\liminf(-a_n). \tag{9.63}$$

□

PROPooUOYTooGwGZHz

Proposition 9.48 ([1]).
Si $a_n \to a$, alors

$$\liminf_n |a_n| = |a|. \tag{9.64}$$

9.3.11 Ouverts, voisinage, topologie

Lorsque $x \in E$, nous rappelons qu'un voisinage [19] de x est n'importe quel sous-ensemble de E qui contient une boule ouverte centrée en x. La proposition 7.8 nous dit qu'un ensemble est ouvert si il contient un voisinage de chacun de ses points. Au passage, rappelons que l'ensemble vide est ouvert.

Pour rappel, la proposition 7.118 dit que l'ensemble des boules ouvertes d'un espace métrique génère la topologie de l'espace.

Nous rappelons qu'une partie A d'un espace métrique est dite bornée [20] si il existe une boule [21] qui contient A.

Mais revenons à $\mathbb{R}$...

LemSupOuvPas

Lemme 9.49.
Une partie ouverte de $\mathbb{R}$ ne contient pas son supremum.

Démonstration. Soit $\mathcal{O}$, un ensemble ouvert et s, son supremum. Si s était dans $\mathcal{O}$, on aurait un voisinage $B = B(s, r)$ de s contenu dans $\mathcal{O}$. Le point $s + r/2$ est alors à la fois dans $\mathcal{O}$ et plus grand que s, ce qui contredit le fait que s soit un supremum de $\mathcal{O}$. □

Par le même genre de raisonnement, on montre que l'union et l'intersection de deux ouverts, sont encore des ouverts.

Remarque 9.50.
L'intersection d'une infinité d'ouverts n'est pas spécialement un ouvert comme le montrent les parties $\{\mathcal{O}_k\}_{k\in\mathbb{N}^*}$ donnés par

$$\mathcal{O}_k =]1, 2 + \frac{1}{k}[. \tag{9.65}$$

Tous les ensembles $\mathcal{O}_k$ contiennent le point 2 qui est donc dans l'intersection. Mais nous allons montrer que pour tout $\epsilon > 0$, il existe n tel que $2 + \epsilon \notin \mathcal{O}_n$. Il suffit de prendre n tel que $\frac{1}{n} < \epsilon$ (lemme 1.441(2)).

PROPooANIOooIJHelX

Proposition 9.51.
Quelles que soient les parties A et B de $\mathbb{R}$, nous avons

$$\sup(A \cap B) \leqslant \sup A \leqslant \sup(A \cup B). \tag{9.66}$$

Démonstration. En deux parties.

19. Définition 7.4.
20. Définition 7.140.
21. À titre d'exercice, convainquez-vous que l'on peut dire boule *ouverte* ou *fermée* au choix sans changer la définition.

(1) $\sup(A \cap B) \leqslant \sup(A)$ Soit $s = \sup(A)$. En particulier, s est un majorant de A. Si $x \in A \cap B$, alors $x \in A$ et $s \geqslant x$. Donc s est également un majorant de $A \cap B$. Le lemme 1.462 conclut que $s \geqslant \sup(A \cap B)$.

(2) $\sup(A) \leqslant \sup(A \cup B)$ Soit $s = \sup(A \cup B)$. Par définition, s est un majorant de $A \cup B$. A fortiori, s est un majorant de A et donc est plus grand ou égal à $\sup(A)$.

□

9.3.12 Intervalles et connexité

Nous allons déterminer tous les sous-ensembles connexes[22] de $\mathbb{R}$. Pour cela nous relisons d'abord la notion d'intervalle donnée en 1.21 ainsi que la proposition 1.466 qui liste tous les intervalles de $\mathbb{R}$. La partie $I \subset \mathbb{R}$ est un intervalle si pour tout $a, b \in I$, tout nombre entre a et b est également dans I. Cette définition englobe tous les exemples connus d'intervalles ouverts, fermés avec ou sans infini : $[a, b]$, $[a, b[$, $]-\infty, a]$, ...L'ensemble $\mathbb{R}$ lui-même est un intervalle.

Si I est un intervalle, les nombres $\inf(I)$ et $\sup(I)$[23] sont les **extrémités** de I.

DefLISOooDHLQrl

Définition 9.52.
*Étant donnés deux points a et b dans $\mathbb{R}^p$ on appelle **segment** d'extrémités a et b, et on note $[a, b]$, l'image de $[0, 1]$ par l'application $s : [0, 1] \to \mathbb{R}^p$, $s(t) = (1-t)a + tb$. On pose $]a, b[= s(]0, 1[)$, et $]a, b] = s(]0, 1])$.*

Il faut observer que le segment $[a, b]$ est une courbe orientée : certes en tant que ensembles, $[a, b] = [b, a]$, mais si nous regardons la fonction de t correspondante à $[b, a]$, nous voyons qu'elle va dans le sens inverse de celle qui correspond à $[a, b]$. Nous approfondirons ces questions lorsque nous parlerons d'arcs paramétrés autour de la section 21.7.

Le segment $[b, a]$ est l'image de l'application $r : [0, 1] \to \mathbb{R}^p$ donnée par $r(t) = (1-t)b + ta$.

PropInterssiConn

Proposition 9.53.
Une partie de $\mathbb{R}$ est connexe si et seulement si c'est un intervalle[24].

Démonstration. La preuve est en deux parties. D'abord nous démontrons que si un sous-ensemble de $\mathbb{R}$ est connexe, alors c'est un intervalle ; et ensuite nous démontrons que tout intervalle est connexe.

Afin de prouver qu'un ensemble connexe est toujours un intervalle, nous allons prouver que si un ensemble n'est pas un intervalle, alors il n'est pas connexe. Prenons A, une partie de $\mathbb{R}$ qui n'est pas un intervalle. Il existe donc $a, b \in A$ et un x_0 entre a et b qui n'est pas dans A. Comme le but est de prouver que A n'est pas connexe, il faut couper A en deux ouverts disjoints. L'élément x_0 qui n'est pas dans A est le bon candidat pour effectuer cette coupure. Prenons M, un majorant de A et m, un minorant de A, et définissons

$$\begin{aligned} \mathcal{O}_1 &=]m, x_0[\\ \mathcal{O}_2 &=]x_0, M[. \end{aligned}$$

Si A n'a pas de minorant, nous remplaçons la définition de $\mathcal{O}_1$ par $]-\infty, x_0[$, et si A n'a pas de majorant, nous remplaçons la définition de $\mathcal{O}_2$ par $]x_0, \infty[$. Dans tous les cas, ce sont deux ensembles ouverts dont l'union recouvre tout A. En effet, $\mathcal{O}_1 \cup \mathcal{O}_2$ contient tous les nombres entre un minorant de A et un majorant sauf x_0, mais on sait que x_0 n'est pas dans A. Cela prouve que A n'est pas connexe.

22. Définition 7.70.
23. Qui existent par la proposition 1.461, quitte à poser $\pm\infty$ comme infimum et supremum lorsque I n'est pas borné.
24. Définition 1.21.

Jusqu'à présent nous avons prouvé que si un ensemble n'est pas un intervalle, alors il ne peut pas être connexe. Pour remettre les choses à l'endroit, prenons un ensemble connexe, et demandons-nous si il peut être autre chose qu'un intervalle ? La réponse est *non* parce que si il était autre chose, il ne serait pas connexe.

Prouvons à présent que tout intervalle est connexe. Pour cela, nous refaisons le coup de la contraposée. Nous allons donc prendre une partie A de $\mathbb{R}$, supposer qu'elle n'est pas connexe et prouver qu'elle n'est alors pas un intervalle. Nous avons deux ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ tels que $A \subset \mathcal{O}_1 \cup \mathcal{O}_2$. Notons $A_1 = A \cap \mathcal{O}_1$ et $A_2 = A \cap \mathcal{O}_2$; et prenons $a \in A_1$ et $b \in A_2$. Pour fixer les idées, on suppose que $a < b$. Maintenant, le jeu est de montrer qu'il existe un point x_0 entre a et b qui ne soit pas dans A (cela montrerait que A n'est pas un intervalle). Nous allons prouver que c'est le cas du point

$$x_0 = \sup\{x \in \mathcal{O}_1 \text{ tel que } x < b\}.$$

Étant donné que l'ensemble $\mathcal{A} = \{x \in \mathcal{O}_1 \text{ tel que } x < b\}$ est ouvert [25], le point x_0 n'est pas dans l'ensemble par le lemme 9.49. Nous avons donc

— soit x_0 n'est pas dans $\mathcal{O}_1$,

— soit $x_0 \leqslant b$,

— soit les deux en même temps.

Nous allons montrer qu'un tel x_0 ne peut pas être dans A. D'abord, remarquons que $\sup \mathcal{A} \leqslant \sup \mathcal{O}_1$ parce que $\mathcal{A}$ est une intersection de $\mathcal{O}$ avec quelque chose. Ensuite, il n'est pas possible que x_0 soit dans $\mathcal{O}_2$ parce que tout élément de $\mathcal{O}_2$ possède un voisinage contenu dans $\mathcal{O}_2$. Un point de $\mathcal{O}_2$ est donc toujours strictement plus grand que le supremum de $\mathcal{O}_1$.

Maintenant, en remarquant que si $x_0 \leqslant b$, alors $x_0 = b$ sinon b serait un majorant de $\mathcal{A}$ plus petit que x_0, ce qui n'est pas possible puisque x_0 est le supremum de $\mathcal{A}$ et donc le plus petit majorant. Oui mais si $x_0 = b$, c'est que $x_0 \in \mathcal{O}_2$, ce qu'on vient de montrer être impossible. Nous voilà déjà débarrassés des deuxièmes et troisièmes possibilités.

Si la première possibilité est vraie, alors x_0 n'est pas dans A parce qu'on a aussi prouvé que $x_0 \notin \mathcal{O}_2$. Or n'être ni dans $\mathcal{O}_1$ ni dans $\mathcal{O}_2$ implique de ne pas être dans A. Ce point $x_0 = \sup \mathcal{A}$ est donc hors de A.

Oui, mais comme $a \in \mathcal{A}$, on a obligatoirement $x_0 \geqslant a$. Mais par construction, on a aussi $x_0 \leqslant b$ (ici, l'inégalité est même stricte, mais ce n'est pas important). Donc

$$a \leqslant x_0 \leqslant b$$

avec a, $b \in A$, et $x_0 \notin A$. Cela finit de prouver que A n'est pas un intervalle. □

Le lemme suivant dit que si on recouvre un intervalle avec des ouverts, alors on peut ordonner ces ouverts de telle sorte qu'ils s'enchainent bien : on peut sauter de l'un à l'autre en passant par les intersections. C'est donc un lemme qui permet de passer du local au global.

LEMooNMGWooTfQDeO

Lemme 9.54.
Soient un intervalle I de $\mathbb{R}$ ainsi qu'un recouvrement $\{\mathcal{O}_i\}_{i=1,\dots,n}$ de I par des ouverts connexes tels que $\mathcal{O}_i \cap I \neq \varnothing$ pour tout i [26]. Alors il existe une bijection $\psi\colon \{1,\dots,n\} \to \{1,\dots,n\}$ telle que

(1)

$$\bigcup_{i=1}^{m} \mathcal{O}_{\psi(i)} \tag{9.67}$$

est connexe pour tout m.

(2)

$$\mathcal{O}_{\psi(m)} \cap \bigcup_{i=1}^{m-1} \mathcal{O}_{\psi(i)} \neq \varnothing. \tag{9.68}$$

25. C'est l'intersection entre l'ouvert $\mathcal{O}_1$ et l'ouvert $\{x \text{ tel que } x < b\}$.
26. Il est cependant possible que les $\mathcal{O}_i$ ne soient pas inclus dans I.

Démonstration. Nous allons construire ψ par récurrence ; plus précisément nous allons construire des applications $\psi_k : \{1,\dots,k\} \to \{1,\dots,n\}$ telle que

(1) ψ_k est injective.

(2) Si $i \leqslant k$ alors $\psi_k(i) = \psi_i(i)$.

(3) La partie
$$\bigcup_{i=1}^{m} \mathcal{O}_{\psi_k(i)} \tag{9.69}$$
est connexe pour tout $m \leqslant k$.

(4) Nous avons
$$\mathcal{O}_{\psi_k(m)} \cap \bigcup_{i=1}^{m-1} \mathcal{O}_{\psi_k(i)} \neq \varnothing \tag{9.70}$$
pour tout $m \leqslant k$.

Nous commençons en douceur par
$$\begin{aligned} \psi_1 : \{1\} &\to \{1,\dots,n\} \\ 1 &\mapsto 1. \end{aligned} \tag{9.71}$$
Ai-je besoin de vous prouver que c'est injectif ?

Nous supposons que les applications ψ_i sont correctement définies pour $i \leqslant k$, et nous construisons ψ_{k+1}. Nous posons
$$A = \psi_k\big(\{1,\dots,k\}\big) \tag{9.72a}$$
$$B = \{1,\dots,n\} \backslash A \tag{9.72b}$$
$$P = \bigcup_{i=1}^{k} \mathcal{O}_{\psi_k(i)} \tag{9.72c}$$
$$Q = \bigcup_{i=k+1}^{n} \mathcal{O}_{\psi_k(i)} \tag{9.72d}$$
En tant qu'unions d'ouverts, les parties P et Q sont ouvertes dans $\mathbb{R}$. Elles recouvrent l'intervalle I qui est connexe par la proposition 9.53. De plus $P \cap I \neq \varnothing$ et $Q \cap I \neq \varnothing$; donc, par définition de la connexité nous avons $P \cap Q \neq \varnothing$.

Il existe donc $i_0 \in B$ tel que $P \cap \mathcal{O}_{i_0} \neq \varnothing$. Nous posons
$$\begin{aligned} \psi_{k+1} : \{1,\dots,k+1\} &\to \{1,\dots,n\} \\ i &\mapsto \begin{cases} \psi_k(i) & \text{si } i \neq k+1 \\ i_0 & \text{si } i = k+1. \end{cases} \end{aligned} \tag{9.73}$$
Vérifions que ce ψ_{k+1} vérifie les conditions.

(1) ψ_{k+1} est injective. Soient i, j tels que $\psi_{k+1}(i) = \psi_{k+1}(j)$. Si $i = k+1$ et $j \neq k+1$ alors $\psi_{k+1}(i) = i_0 \in B$, alors que $\psi_{k+1}(j) = \psi_k(j) \in A$. Donc le cas $i = k+1$, $j \neq k+1$ n'est pas possible.

Si $i, j \neq k+1$, alors $\psi_{k+1}(i) = \psi_k(i)$ et $\psi_{k+1}(j) = \psi_k(j)$. L'injectivité de ψ_k implique que $i = j$.

(2) Si $i \leqslant k$, nous avons $\psi_{k+1}(i) = \psi_k(i) = \psi_i(i)$ en utilisant la récurrence.

(3) Nous séparons les cas $m = k+1$ et $m \neq k+1$. Si $m \neq k+1$ alors tous les ψ_{k+1} dans (9.69) [27] sont des ψ_k et la récurrence fonctionne. Si $m = k+1$ alors
$$\bigcup_{i=1}^{k+1} \mathcal{O}_{\psi_{k+1}(i)} = \bigcup_{i=1}^{k} \mathcal{O}_{\psi_k(i)} \cup \mathcal{O}_{\psi_{k+1}(k+1)} = P \cup \mathcal{O}_{i_0}. \tag{9.74}$$

27. Nous sommes en train de parler de cette équation avec $k+1$ au lieu de k, parce que nous sommes dans un processus de récurrence. Il est donc normal de dire qu'il y a des ψ_{k+1} dans cette équation.

Le nombre i_0 a été choisi pour avoir $\mathcal{O}_{i_0} \cap P \neq \emptyset$. Comme $\mathcal{O}_{i_0}$ et P sont des connexes, la proposition 7.76(1) implique que $P \cup \mathcal{O}_{i_0}$ est connexe.

(4) Encore une fois, si $m \neq k+1$, tous les ψ_{k+1} de (9.70) deviennent des ψ_k et la récurrence fonctionne. Avec $m = k+1$ nous avons

$$\mathcal{O}_{\psi_{k+1}(k+1)} \cap \bigcup_{i=1}^{k} \mathcal{O}_{\psi_{k+1}(i)} = \mathcal{O}_{i_0} \cap P. \tag{9.75}$$

Cette intersection est non vide, par choix du i_0.

Quand tous les ψ_k ($k = 1, \dots, n$) sont construits, en posant $\psi = \psi_n$ nous avons le résultat annoncé. □

ThoMKKooAbHaro

Théorème 9.55 (Théorème des bornes atteintes)**.**
Une fonction à valeurs réelles continue sur un compact est bornée et atteint ses bornes.

C'est-à-dire qu'il existe $x_0 \in K$ tel que $f(x_0) = \inf\{f(x) \text{ tel que } x \in K\}$ ainsi que x_1 tel que $f(x_1) = \sup\{f(x) \text{ tel que } x \in K\}$.

Démonstration. Soient un espace topologique compact K et une fonction continue $f\colon K \to \mathbb{R}$. Alors le théorème 7.220 indique que $f(K)$ est compact. Par conséquent $f(K)$ est un fermé borné de $\mathbb{R}$ par le théorème de Borel-Lebesgue 9.24. Puisque $f(K)$ est borné, la fonction f est bornée.

De plus $f(K)$ étant fermé, son infimum est un minimum et son supremum est un maximum : il existe $x \in K$ tel que $f(x) = \sup f(K)$ et il existe $y \in K$ tel que $f(y) = \inf f(K)$. □

DEFooIVFZooLLlahl

Définition 9.56 (propriété de valeurs intermédiaires[312])**.**
*Un espace topologique X a la **propriété de valeurs intermédiaires** si pour toute fonctions continue $f\colon X \to \mathbb{R}$, si pour tout $a, b \in X$, nous avons $[r_1, r_2] \subset f(X)$.*

PROPooGURQooAwKNUJ

Proposition 9.57 ([312])**.**
Un espace topologique a la propriété des valeurs intermédiaires si et seulement si il est connexe.

Démonstration. Les deux sens.

(1) $\Rightarrow$ Soit un espace topologique X que nous supposons ne pas être connexe. Nous considérons des ouverts A et B tels que $A \cup B = X$ et $A \cap B = \emptyset$. Ensuite nous prenons la fonction

$$\begin{aligned} f\colon X &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in B. \end{cases} \end{aligned} \tag{9.76}$$

Cette fonction ne vérifie manifestement pas la propriété des valeurs intermédiaires. Et pourtant elle est continue. En effet si $\mathcal{O}$ est un ouvert de $\mathbb{R}$, alors $f^{-1}(\mathcal{O})$ peut valoir $\emptyset$, A, B ou X suivant que 1 ou -1 sont ou non dans $\mathcal{O}$. Or toutes ces parties sont ouvertes. Donc $f^{-1}(\mathcal{O})$ est ouvert.

(2) $\Leftarrow$ Nous supposons maintenant que X est connexe, nous considérons une fonction continue $f\colon X \to \mathbb{R}$, et nous montrons qu'elle vérifie la propriété des valeurs intermédiaires. Le lemme 7.218 dit que $f(X)$ est un connexe de $\mathbb{R}$. La proposition 9.53 dit alors que $f(X)$ est un intervalle de $\mathbb{R}$.

□

9.3.13 Théorème de Bolzano-Weierstrass

Le théorème suivant est essentiellement inutile pour les raisons suivantes :

— Il est un cas particulier du théorème 7.144 qui donne pour tout espace métrique, l'équivalence entre la compacité et la compacité séquentielle.

— Il est un cas particulier du théorème 7.290 qui le donne pour tous les espaces compacts.

Bref, nous ne le laissons que pour le lecteur qui n'aurait pas en tête d'autres définitions de « compact » à part « fermé borné ».

ThoBolzanoWeierstrassRn

Théorème 9.58 (Théorème de Bolzano-Weierstrass)**.**
Toute suite contenue dans un compact de $\mathbb{R}^m$ admet une sous-suite convergente.

Démonstration. Nous rappelons qu'une partie compacte de $\mathbb{R}^m$ est fermée et bornée par le théorème de Borel-Lebesgue 9.24.

Soit (x_n) une suite contenue dans une partie bornée de $\mathbb{R}^m$. Considérons (a_n), la suite réelle des premières composantes des éléments de (x_n) : pour chaque $n \in \mathbb{N}$, le nombre a_n est la première composante de x_n. Étant donné que la suite (x_n) est bornée, il existe un M tel que $\|x_n\| < M$. La croissance de la fonction racine carrée donne

$$|a_n| \leqslant \|x_n\| \leqslant M. \tag{9.77}$$

La suite (a_n) est donc une suite réelle bornée et donc contient une sous-suite convergente par le théorème correspondant dans $\mathbb{R}$: 7.144. Soit a_{I_1} une sous-suite convergente de a_n. Nous considérons maintenant x_{I_1}, c'est-à-dire la suite de départ dont on a enlevé tous les éléments qu'il faut pour qu'elle converge en ce qui concerne la première composante.

Si nous considérons la suite b_{I_1} des *secondes* composantes de x_{I_1}, nous en extrayons, de la même façon que précédemment, une sous-suite convergente, c'est-à-dire que nous avons un $I_2 \subset I_1$ tel que b_{I_2} est convergent. Notons que a_{I_2} est une sous-suite de la (sous) suite convergente x_{I_1}, et donc a_{I_2} est encore convergente.

En continuant ainsi, nous construisons une sous-sous-sous-suite x_{I_3} telle que la suite des *troisièmes* composantes est convergente. Lorsque nous avons effectué cette procédure m fois, la suite x_{I_m} est une suite dont toutes les composantes convergent, et donc est une suite convergente par la proposition 9.36.

Le tableau suivant donne un petit schéma de la façon dont nous procédons. Les $\bullet$ sont les éléments de la suite que nous gardons, et les $\times$ sont ceux que nous « jetons ».

$$\begin{array}{lcccccccccc} x_{\mathbb{N}} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \dots \\ x_{I_1} & \times & \bullet & \bullet & \times & \bullet & \times & \times & \bullet & \bullet & \bullet & \dots \\ x_{I_2} & \times & \bullet & \times & \times & \bullet & \times & \times & \bullet & \bullet & \times & \dots \\ \vdots & & & & & & & & & & & \\ x_{I_m} & \times & \times & \times & \times & \bullet & \times & \times & \times & \bullet & \times & \dots \end{array} \tag{9.78}$$

La première ligne, $x_{\mathbb{N}}$, est la suite de départ. □

CorFHbMqGGyi

Corolaire 9.59.
Si une suite est croissante et bornée alors elle est convergente.

Démonstration. Nous nommons (x_n) la suite et nous prenons un majorant M. Toute la suite est alors contenue dans le compact $[x_0, M]$, ce qui donne une sous-suite $(x_{\alpha(n)})$ convergente par le théorème de Bolzano-Weierstrass 7.290. Si ℓ est la limite de cette sous-suite alors nous avons $\ell \geqslant x_n$ pour tout n.

Pour tout $\epsilon > 0$ il existe K tel que si $n > K$ alors $|\ell - x_{\alpha(n)}| < \epsilon$. Comme ℓ majore la suite nous avons même

$$x_{\alpha(n)} + \epsilon > \ell. \tag{9.79}$$

Puisque la suite est croissante pour tout $m > \alpha(K)$ nous avons $x_m + \epsilon > \ell$, ce qui signifie $|x_m - \ell| < \epsilon$. □

Nous aurons une version pour les fonctions croissantes et bornées en la proposition 12.56.

La proposition suivante dit que la notion d'ensemble non dénombrable ne prend pas réellement de force entre $\mathbb{R}$ et $\mathbb{R}^n$: il n'y a pas moyen de caser $\mathbb{R}$ dans $\mathbb{R}^n$ de façon à ce qu'il y tienne à son aise.

Proposition 9.60.
Une partie non dénombrable de $\mathbb{R}^n$ possède un point d'accumulation[28].

Démonstration. Soit une partie $A \subset \mathbb{R}^n$ sans point d'accumulation. Nous allons prouver que A est dénombrable.

Soient les compacts $K_n = \overline{B(0,n)}$. La partie $A \cap K_n$ est finie ; sinon elle aurait une partie en bijection avec $\mathbb{N}$ (proposition 1.142) et donc une suite. Or une suite dans un compact possède un point d'accumulation par le théorème 7.290.

Donc tous les $A \cap K_n$ sont finis. Puisque $A = \bigcup_n A \cap K_n$, l'ensemble A est une réunion dénombrable d'ensembles finis. Il est donc dénombrable. □

9.3.14 Recouvrement par des intervalles ouverts

Soit un ensemble E et un ensemble $\mathcal{A}$ de parties de E. Soit $A \in \mathcal{A}$. Nous aimerions savoir quels sont les éléments de $\mathcal{A}$ qui sont atteignables en partant de A et en ne « sautant » que d'intersection en intersection.

Nous notons $\mathcal{A} = \{B_i\}_{i\in I}$ où I est un ensemble d'indices (un ensemble quelconque).

$$s_1(A) = \{i \in I \text{ tel que } B_i \cap A \neq \varnothing\} \tag{9.80a}$$

$$\sigma_1(A) = \bigcup_{B\in s_1(A)} B. \tag{9.80b}$$

Et ensuite :

$$s_{k+1}(A) = \{i \in I \text{ tel que } B_i \cap \sigma_k(A) \neq \varnothing\} \tag{9.81a}$$

$$\sigma_{k+1}(A) = \bigcup_{B\in s_{k+1}(A)} B \tag{9.81b}$$

LEMooJHMTooXwxSAa

Lemme 9.61.
Soient un intervalle A de $\mathbb{R}$ et $\mathcal{A} = \{I_i\}_{i=1,\dots,N}$ un recouvrement de A par des intervalles ouverts. Si $I_1 \cap A \neq \varnothing$ alors

(1) $\sigma_N = \sigma_{N+1}$
(2) $A \subset \sigma_N(I_1)$.

Démonstration. Si $\sigma_{k+1} = \sigma_k$, alors tous les σ_{k+l} sont identiques. De plus si $\sigma_{k+1} \neq \sigma_k$, alors σ_{k+1} contient au moins un élément de plus que σ_k. Donc $\text{Card}(\sigma_k) \geqslant k$ et en particulier $N \leqslant \text{Card}(\sigma_N) \leqslant N$. Cela prouve le premier point.

L'ensemble $\sigma_N(I_1)$ est une union d'ouverts et est donc un ouvert. Quitte à renuméroter, nous écrivons

$$\sigma_N(I_1) = I_1 \cup \dots \cup I_n. \tag{9.82}$$

L'ensemble

$$\tau = \bigcup_{k=n+1}^{N} I_k \tag{9.83}$$

est ouvert et est disjoint de $\sigma_N(I_1)$ parce que si I_l $(l \geqslant n+1)$ intersectait $\sigma_N(I_1)$, nous aurions $l \in s_{N+1}$ ou encore $I_l \subset \sigma_{N+1} \backslash \sigma_N$.

Donc τ et σ_N sont deux ouverts disjoints qui recouvrent A. Puisque A est un intervalle, il est connexe[29]. Donc, soit $A \subset \tau$, soit $A \subset \sigma_N$. Comme $I_1 \cap A \neq \varnothing$ nous sommes dans le cas $A \subset \sigma_N$. □

LEMooGHPTooKgFvGb

Lemme 9.62.
Soit $x \in \mathbb{R}$. Si $\mathcal{A} = \{I_s\}_{s\in S}$ est un ensemble d'intervalles contenant x, alors $I = \bigcup_{s\in S} I_s$ est un intervalle[30].

28. Définition 7.34.
29. Définition 7.70 et proposition 9.53.
30. Définition 1.21.

Démonstration. Soient $a, b \in I$ (nous supposons $a < b$). Nous devons prouver que $[a, b] \subset I$. Pour cela nous considérons $y \in [a, b]$; il y a deux possibilités : soit $y < x$ soit $y > x$ (si $y = x$ alors $y \in I_s$).

Si $y < x$, alors $a \leqslant y < x$ et donc $y =\in I$. Si $y > x$, alors $x < y \leqslant b$ et $y \in I$. □

PROPooTXMBooZaSKFF

Proposition 9.63 ([313, 314])**.**
Un ouvert de $\mathbb{R}$ peut s'écrire comme union au plus dénombrable d'intervalles ouverts disjoints.

Plus précisément, si $\mathcal{O}$ est un ouvert de $\mathbb{R}$, il existe un ensemble $\mathcal{F} = \{I_s\}_{s\in S}$ où

(1) Chaque I_s est un intervalle ouvert contenu dans $\mathcal{O}$,

(2) Pour $s, t \in S$, si $I_s \neq I_t$, alors $I_s \cap I_t = \emptyset$.

(3) S est dénombrable,

Démonstration. Pour $x \in \mathcal{O}$, nous définissons J_x comme étant l'union de tous les intervalles ouverts contenus dans $\mathcal{O}$ et contenant x. Les J_x ne sont pas vides parce qu'ils contiennent toujours une boule centrée en x [31].

En tant qu'union d'intervalles, J_x est un intervalle par le lemme 9.62. De plus, J_x est ouvert parce que toute union d'ouverts est ouverte [32].

Nous notons $\mathcal{A}$ l'ensemble des intervalles ouverts contenus dans $\mathcal{O}$, et

$$\mathcal{A}_x = \{I \in \mathcal{A} \text{ tel que } x \in I\}. \tag{9.84}$$

(1) **Si $y \in J_x$, alors $J_x = J_y$** Puisque $y \in J_x$, nous pouvons considérer $J = \mathcal{A}_x \cap \mathcal{A}_y$. Nous avons

$$J_y = \bigcup_{I\in\mathcal{A}_y} I \subset \bigcup_{I\in\mathcal{A}_y} \underbrace{(I \cup J)}_{\in\mathcal{A}_x} \subset \bigcup_{I\in\mathcal{A}_x} I = J_x. \tag{9.85}$$

L'inclusion dans l'autre sens s'obtient en écrivant la même équation en échangeant x et y.

(2) **Les J_x sont disjoints** Nous prouvons à présent que pour $x, y \in \mathcal{O}$, nous avons $J_x = J_y$ ou $J_x \cap J_y = \emptyset$. En effet si $a \in J_x \cap J_y$, alors $J_a = J_x$ et $J_a = J_y$. Donc $J_x = J_y$.

(3) **Dénombrable** C'est le moment d'écrire $\mathcal{F} = \{J_x\}_{x\in\mathcal{O}}$. Comme tout intervalle contient au moins un rationnel (proposition 9.16), nous avons aussi

$$\mathcal{F} = \{J_x\}_{x\in\mathcal{O}} = \{J_q\}_{q\in\mathbb{Q}\cap\mathcal{O}}. \tag{9.86}$$

Cet ensemble $\mathcal{F}$ vérifie les conditions demandées.

□

9.4 Connexité par arcs

DEFooOXVCooBizpgK

Définition 9.64.
*Une partie A d'un espace topologique est **connexe par arcs** si pour tout $a, b \in A$, il existe une application continue $\gamma\colon [0, 1] \to A$ telle que $\gamma(0) = a$ et $\gamma(1) = b$.*

9.65.
Un exemple d'ensemble connexe mais pas connexe par arcs est donné par la proposition 21.56. L'idée de cet exemple est de construire un ensemble en deux parties reliées par un chemin de longueur infinie.

Un espoir fou nous prend alors de croire que nous pouvons produire un exemple plus simple avec $\mathbb{R} \cup \{+\infty\}$ parce que, dans cet ensemble, 1 et $+\infty$ sont reliés par un chemin de longueur infinie. La proposition 12.57 nous montrera que non.

31. C'est la définition 7.112 de la topologie métrique.
32. C'est dans la définition 7.1 d'une topologie.

LEMooQPYMooRKVSrv

Lemme 9.66.
Une partie connexe par arcs est connexe.

Beaucoup d'espaces connexe sont connexes par arcs. La proposition suivante couvre entre autres le cas de tout ouvert connexe d'un espace vectoriel normé comme l'espace euclidien.

PROPooYFDBooHbBjzF

Proposition 9.67 ([315])**.**
Tout espace connexe et localement connexe par arcs est connexe par arcs.

LEMooTVQMooFxrFaT

Lemme 9.68.
Soient deux espaces topologiques E et F, et $f : E \to F$ un homéomorphisme[33]*. E est connexe par arcs*[34] *si et seulement si F l'est.*

9.69.
Voici une idée de la preuve.

On montre en réalité que l'image d'un connexe par arcs par une application continue est un connexe par arcs, ce qui implique chaque sens de l'équivalence de l'énoncé.

Soient p et q des points de F. Il existe un chemin reliant un antécédent de p et un antécédent de q (dans E). L'image de ce chemin est un chemin reliant p et q (dans F) puisque composé d'applications continues.

LEMooQFQFooDlxkw

Lemme 9.70.
Une sphère de $\mathbb{R}^n$ est connexe par arcs si $n > 1$

9.71.
Une idée de la preuve.

On voit qu'un cercle est connexe par arcs car on a un paramétrage en sinus et cosinus. Pour une sphère S de centre a en dimension $n > 2$, on se donne p et q sur S et on définit P le plan affine passant par a, p et q. Alors $P \cap S$ est un cercle, donc on peut relier p à q par un chemin dans cette intersection.

Pour voir sur une formule que $P \cap S$ est un cercle, on peut écrire $x - a = \lambda(a - p) + \mu(a - q)$ l'équation (en x) du plan P, et $|x - a|^2 = R^2$ l'équation (en x) de la sphère. En injectant, on obtient une équation du second degré en λ, μ qui se révèle être l'équation d'un cercle à une transformation affine près.

LEMooDYNSooOmJbYq

Lemme 9.72.
Un ouvert connexe par arcs dans $\mathbb{R}^n$ ($n \geqslant 2$) reste connexe par arcs même si on lui enlève un point.

9.73.
Une idée de la preuve.

En effet, soit U un tel ouvert connexe par arcs, et p un point de U. Soient x et y sur $U\backslash\{p\}$. Il existe un chemin γ de x à y. Si le chemin ne passe pas par p, c'est gagné. Si il passe par p, on choisit une boule B fermée (de rayon non-nul) centrée en p qui ne contient ni x ni y. On note $E = \gamma^{-1}(B) \subset [0; 1]$ c'est un ensemble compact (fermé, par continuité de γ, et borné) dont on regarde le maximum $\bar{t}$ et le minimum $\underline{t}$.

Il reste enfin à définir un chemin entre p et q par morceaux

(1) Les points p et $\gamma(\underline{t})$ sont reliés par γ,

(2) Par connexité par arcs, il existe un chemin sur la sphère qui relie $\gamma(\underline{t})$ à $\gamma(\bar{t})$,

(3) et enfin $\gamma(\bar{t})$ et q sont reliés via γ ;

ce qui achève la construction d'un chemin continu entre p et q.

33. Définition 7.41.
34. Définition 9.64

Pour conclure l'exercice, par l'absurde, on prend un voisinage connexe et ouvert V de 0 dans le cône, homéomorphe à un ouvert connexe U de $\mathbb{R}^2$. Or $V\backslash\{0\}$ n'est pas connexe par arcs, alors que l'ouvert dont on retire un point reste connexe par arcs. C'est impossible, donc l'homéomorphisme n'existe pas, et le cône n'est pas une variété de dimension 2.

9.4.1 Des exemples

Exemple 9.74.
Nous étudions l'exemple suivant :

$$A_1 = \{(x,y) \in \mathbb{R}^2 \mid 2y^2 + 4y + 2 < x \leqslant \sqrt{4-y^2},\ y \in [-1.5, 0.5[\}. \tag{9.87}$$

On commence par tracer la parabole $x = 2y^2+4y+2$, la circonférence $x^2+y^2 = 4$ et les droites $y = -1.5$ et $y = 1/2$. On voit tout de suite que l'aire délimitée par les quatre courbes est donnée par l'union de deux parties. Dans la première $\sqrt{4-y^2} \leqslant x \leqslant 2y^2+4y+2$, $y \in [0, 0.5]$ et dans l'autre $2y^2+4y+2 \leqslant x \leqslant \sqrt{4-y^2}$, $y \in [-1.5, 0]$. L'ensemble A_1 est contenu dans la deuxième, 9.1. L'intérieur de A_1 est donné par $\text{Int}(A_1) = \{(x,y) \in \mathbb{R}^2 \mid 2y^2+4y+2 < x < \sqrt{4-y^2},\ y \in]-1.5, 0[\}$, et sa frontière est l'union de 3 morceaux de courbe ℓ_1, ℓ_2, ℓ_3 :

$$\begin{aligned}
\ell_1 &= \{(x,y) \,|\, x = 2y^2+4y+2,\ y \in [-1.5, 0]\} \\
\ell_2 &= \{(x,y) \,|\, x = \sqrt{4-y^2},\ y \in [-1.5, 0]\} \\
\ell_3 &= \{(x,y) \,|\, x \in [0.5, \sqrt{7/4}]\, y = -1.5\}.
\end{aligned} \tag{9.88}$$

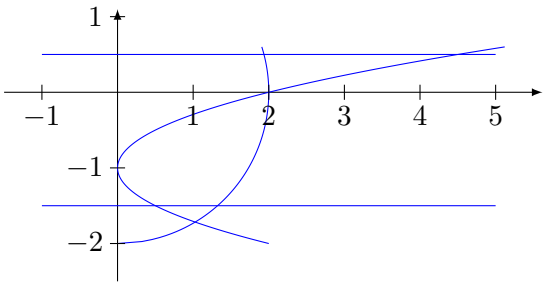


FIGURE 9.1: LabelFigLAfWmaN

△ EXooEJWBooDjBfKV

Exemple 9.75.
Nous étudions

$$A_3 = \mathbb{N} \times \mathbb{Q} = \{(x,y) \in \mathbb{R}^2 \mid x \in \mathbb{N}, y \in \mathbb{Q}\}. \tag{9.89}$$

L'ensemble A_3 n'est pas ouvert, ni fermé, ni borné dans la topologie de $\mathbb{R}^2$. Le lemme 7.31 dit que $\mathbb{Q}$ a un intérieur vide et sa fermeture est $\mathbb{R}$. L'ensemble $\mathbb{N}$, par contre est fermé et non borné. On peut remarquer que tous les points de $\mathbb{N}$ sont points isolés. La fermeture de A_3 est alors $\mathbb{N} \times \mathbb{R}$ et son intérieur est vide. On peut dessiner la fermeture de cet ensemble comme une famille de droites verticales $x = n$, pour tout n dans $\mathbb{N}$. △

Exemple 9.76.
Nous étudions l'ensemble

$$A_3 = \{(t, 2t) \in \mathbb{R}^2 \mid t \in [0,1]\ \}. \tag{9.90}$$

L'ensemble A_3 est un petit segment de droite. Son intérieur est vide parce que toute boule centrée en un point de la droite intersecte l'extérieur de la droite. Son adhérence et sa frontière sont A_3 lui-même parce que nous considérons les valeurs de t dans $[0,1]$ qui est un intervalle fermé. Si l'intervalle avait été ouvert, l'adhérence et la frontière auraient été trouvés en fermant :

$$\overline{\{(t,2t) \text{ tels que } t \in [0,1[\ \}} = \{(t,2t) \text{ tels que } t \in [0,1]\} \tag{9.91}$$

Étant donné que son adhérence est égal à lui-même, cet ensemble est fermé (et donc pas ouvert). Il est également borné parce qu'il est contenu dans une boule de rayon 3. △

Exemple 9.77.
Nous étudions l'ensemble

$$A_4 = \mathbb{Q} \times \mathbb{Q} = \{(x, y) \in \mathbb{R}^2 \mid x \in \mathbb{Q}, y \in \mathbb{Q}\}. \tag{9.92}$$

Dans $\mathbb{R}$ nous savons que $\bar{\mathbb{Q}} = \mathbb{R}$, $\text{Int}(\mathbb{Q}) = \varnothing$ et $\partial\mathbb{Q} = \mathbb{R}$ parce que toute boule centrée en un rationnel contient un irrationnel, et inversement, toute boule centrée en un irrationnel contient un rationnel. Dans $\mathbb{R}^2$ nous avons le même phénomène parce dans la boule $B\big((p, q), r\big)$ avec $(p, q) \in \mathbb{Q} \times \mathbb{Q}$, se trouvent en particulier les points de la forme (p, x) avec $x \in B(q, r) \subset \mathbb{R}$. Évidement, certains de ces x ne sont pas dans $\mathbb{Q}$ et par conséquent, la boule $B\big((p, q), r\big)$ contient les points $(p, x) \notin \mathbb{Q} \times \mathbb{Q}$.

De la même manière, si (x, y) est un point de $\mathbb{R}^2$, dans toute boule centrée en (x, y), il y aura un élément de $\mathbb{Q}^2$.

Par conséquent, $\text{Int}(\mathbb{Q} \times \mathbb{Q}) = \varnothing$, $\overline{\mathbb{Q} \times \mathbb{Q}} = \mathbb{R} \times \mathbb{R}$ et $\partial(\mathbb{Q} \times \mathbb{Q}) = \mathbb{R}^2$.

Il n'est ni ouvert ni fermé (parce qu'il n'est égal ni à son intérieur ni à sa fermeture). Il n'est pas borné non plus parce qu'il existe des nombres rationnels arbitrairement grands. △

Exemple 9.78.
Nous étudions l'ensemble

$$A_5 = \{(x, y) \in \mathbb{R}^2 \mid x \in]0, 1[, \sin\frac{1}{x} < y < 3\}. \tag{9.93}$$

La fonction $x \mapsto \sin(\frac{1}{x})$ est une des fonctions dont le graphe doit être connu. La figure 9.2 montre la situation. Comme d'habitude, il est fortement recommandé de refaire le dessin soi-même.

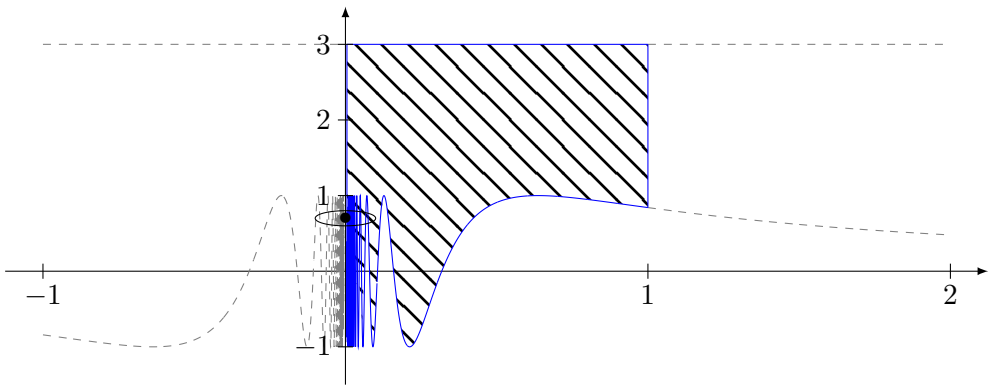


FIGURE 9.2: Les points qui sont sur l'axe vertical entre 0 et 3 sont sur la frontière, mais pas dans l'ensemble A_5. LabelFigAdhIntFrTrois

L'ensemble A_5 est ouvert parce que les conditions $x \in]0, 1[$ et $\sin\frac{1}{x} < y < 3$ sont des conditions « ouvertes » au sens où si un point les vérifient, alors on peut trouver une boule dans lequel ces conditions restent vérifiées. Cela prouve que $\text{Int}(A_5) = A_5$.

La fermeture de A_5 contient en outre les points tels que $\sin\frac{1}{x} = y$ entre $x = 0$ et $x = 1$ (les bornes étant incluses) ainsi que les points des trois segments de droites suivants :

$$\begin{aligned} \{(0, y) \text{ tels que } y \in [-1, 3]\} \\ \{(x, 3) \text{ tels que } x \in [0, 1]\} \\ \{(1, y) \text{ tels que } y \in [\sin(1), 3]\}. \end{aligned} \tag{9.94}$$

La frontière est composée de ces trois segments et du graphe de la fonction $\sin\frac{1}{x}$ entre 0 et 1.

L'ensemble A_5 est borné parce qu'il est contenu par exemple dans la boule centrée en $(0, 0)$ et de rayon 10. Il est ouvert et donc pas fermé. △

ItemexoEspVectoNorme0003iv

Exemple 9.79.
Nous étudions l'ensemble

$$A_6 = \bigcup_{n\in\mathbb{N}_0} \{(\frac{1}{n}, y) \mid y \in [0,1] \}. \tag{9.95}$$

L'ensemble A_6 est une union infinie de segments de droites verticaux, voir figure 9.3

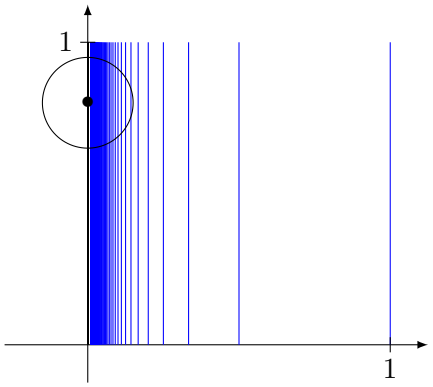


FIGURE 9.3: Le segment sur l'axe vertical entre $y = 0$ et $y = 1$ fait partie de l'adhérence et de la frontière, mais pas de l'ensemble A_6 lui-même.

LabelFigAdhIntFrSix

L'intérieur est vide parce qu'autour de tout réel de la forme $\frac{1}{n}$, il y a un réel qui n'est pas de cette forme. En ce qui concerne la frontière et l'adhérence, il s'agit de l'union de tous ces segments plus le segment en $x = 0$.

En effet, la boule de rayon r autour du point $(0, y)$ contient le point $(\frac{1}{n}, y)$ avec n assez grand pour que $\frac{1}{n} < r$. △

9.4.2 Quelques mots à propos de la droite réelle achevée

Définition 9.80.
*La **droite réelle achevée** est l'ensemble $\mathbb{R} \cup \{\pm\infty\}$ où $\pm\infty$ sont deux nouveaux éléments. Nous la notons $\overline{\mathbb{R}}$ pour des raisons que nous verrons à peine plus bas.*

Cette définition ne servirait à rien si nous n'y mettions pas une topologie pour positionner les éléments $\pm\infty$ par rapport à ceux qui existaient déjà dans $\mathbb{R}$.

Définition 9.81 (Topologie sur $\bar{\mathbb{R}}$)**.**
La topologie sur $\bar{\mathbb{R}}$ est celle sur $\mathbb{R}$ à laquelle nous ajoutons les voisinages de $\pm\infty$ de la façon suivante. Une partie V de $\bar{\mathbb{R}}$ est un voisinage de $+\infty$ si il existe $m > 0$ tel que $]m, +\infty] \subset V$.

Le lemme suivant justifie la notation $\overline{\mathbb{R}}$ pour la droite réelle achevée [35].

LEMooPZXHooEEXsTC

Lemme 9.82.
L'adhérence [36] de $\mathbb{R}$ dans $\overline{\mathbb{R}}$ est $\overline{\mathbb{R}}$.

Démonstration. Il suffit de prouver que $+\infty$ et $-\infty$ sont dans l'adhérence de $\mathbb{R}$. Nous le faisons pour $+\infty$. Ce n'est pas très compliqué : si A est un ouvert contenant $+\infty$, il contient une partie de la forme $]a, +\infty]$, et donc contient des éléments de $\mathbb{R}$. □

Pour la suite nous utilisons la notation (pratique en probabilité)

$$\{f < a\} = \{x \in S \text{ tel que } f(x) < a\}. \tag{9.96}$$

35. Notez que l'espace métrique $\mathbb{R}$ est déjà complet. Il ne s'agit donc pas d'une completion.
36. Définition 7.20.

9.5 Continuité

La définition de fonction continue est la définition 7.36.

(1) Dans le cas d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$, la définition de la continuité devient la proposition 9.83.

(2) Une application continue vers un espace normé est localement bornée, proposition 7.121.

PROPooVNGEooPwbxXP

Proposition 9.83.
Soient $A \subset \mathbb{R}$ et $a \in A$. La fonction $f\colon A \to \mathbb{R}$ est continue[37] *en a si et seulement si pour tout $\epsilon > 0$, il existe un $\delta > 0$ tel que si $x \in B(a,\delta) \cap A$ alors $|f(x) - f(a)| \leqslant \epsilon$.*

Démonstration. En deux parties.

(1) **Si f est continue** Soit $\epsilon > 0$. L'ouvert $W = B\big(f(a),\epsilon\big)$ contient $f(a)$. La définition de la continuité en a dit qu'il existe un ouvert V de A contenant a et tel que $f(V) \subset B\big(f(a),\epsilon\big)$. Vu que V est un ouvert de A[38], il contient une partie de la forme $B(a,\delta) \cap A$ pour un certain $\delta > 0$. Pour ce δ, nous avons bien que $f(x) \in B\big(f(a),\epsilon\big)$ dès que $x \in B(a,\sigma) \cap A$.

(2) **Si f dans l'autre sens** Soit un ouvert W de $\mathbb{R}$ contenant $f(a)$. Nous avons un $\epsilon > 0$ tel que $B\big(f(a),\epsilon\big) \subset W$. Il existe donc un $\delta > 0$ tel que

$$f\big(B(a,\delta) \cap A\big) \subset B\big(f(a),\epsilon\big) \subset W. \tag{9.97}$$

La partie $B(a,\delta) \cap A$ est un voisinage de a dans A.

□

Nous allons maintenant étudier quelques conséquences de la continuité sur $\mathbb{R}$.

(1) D'abord on voit que la continuité n'a été définie qu'en un point. On peut dire que la fonction f est continue *en tel point donné*, mais nous n'avons pas dit ce qu'est une fonction continue *dans son ensemble*.

(2) Le théorème 7.205 nous précise que si I est un intervalle de $\mathbb{R}$, la fonction f est continue sur I si et seulement si elle est continue en chaque point de I.

(3) Comme la définition de f continue en a fait intervenir $f(x)$ pour tous les x pas trop loin de a, il faut au moins déjà que f soit définie sur ces x. En d'autres termes, dire que f est continue en a demande que f existe sur un intervalle autour de a.
Ceci couplé à la définition précédente laisse penser qu'il est surtout intéressant d'étudier les fonctions qui sont continues sur un intervalle.

(4) L'intuition qu'une fonction continue doit pouvoir être tracée sans lever la main correspond aux fonctions continues sur des intervalles. Au moins sur l'intervalle où elle est continue, elle devrait être traçable en un coup. Cette intuition est complètement fausse (comme pratiquement toutes les intuitions), comme le montre l'exemple 9.84.

EXooJBGSooBOGSse

Exemple 9.84.
Il est très possible d'être continue en un seul point. Par exemple la fonction

$$f(x) = x(1 - \mathbb{1}_{\mathbb{Q}}(x)) \tag{9.98}$$

où $\mathbb{1}_{\mathbb{Q}}$ est la fonction indicatrice de $\mathbb{Q}$ dans $\mathbb{R}$. △

PROPooUBUAooNIxjfg

Proposition 9.85.
Si $f\colon \mathbb{R} \to \mathbb{R}$ est continue au point $a \in \mathbb{R}$ et si $f(a) \neq 0$, alors il existe un voisinage de a sur lequel f ne s'annule pas.

37. Définition 7.36(1).
38. Topologie induite et tout ça.

Démonstration. Si f s'annulait sur tout voisinage de a (mais pas en a lui-même), nous aurions, pour tout n un réel

$$x_n \in B\big(a, \frac{1}{n}\big)\backslash\{a\} \tag{9.99}$$

tel que $f(x_n) = 0$. Cela donnerait une suite $x_n \to a$ avec $f(x_n) \to 0$, ce qui contredit la continuité de f en a en vertu de la proposition 7.250 sur la continuité séquentielle en un point. □

Notons que ce résultat se généralise : si f est continue et pas égale à r en a, alors il existe un voisinage de a sur lequel elle ne prend pas la valeur r.

9.5.1 Opération sur la continuité

Nous allons démontrer maintenant une série de petits résultats qui permettent de simplifier la démonstration de la continuité de fonctions.

Théorème 9.86.
Si la fonction f est continue au point a, alors la fonction λf est également continue en a.

Démonstration. Commençons par exprimer la continuité de f en a. Soit $\epsilon_1 > 0$. Il existe $\delta_1 > 0$ tel que

$$(|x - a| \leqslant \delta_1) \Rightarrow |f(x) - f(a)| \leqslant \epsilon_1.$$

En travaillant avec λf au lieu de f,

$$(|x - a| \leqslant \delta_1) \Rightarrow |(\lambda f)(x) - (\lambda f)(a)| \leqslant |\lambda|\epsilon_1. \tag{9.100}$$

Passons à la continuité de λf. Soit $\epsilon > 0$. Nous posons $\epsilon_1 = \epsilon/|\lambda|$ et nous considérons le δ_1 correspondant :

$$(|x - a| \leqslant \delta_1) \Rightarrow |(\lambda f)(x) - (\lambda f)(a)| \leqslant |\lambda|\epsilon_1 = \epsilon.$$

Ce δ_1 est celui que l'on cherchait. □

PROPooBUJCooBHmKEG

Proposition 9.87 ([1])**.**
Si $f\colon \mathbb{R} \to \mathbb{R}$ est une fonction continue ne s'annulant pas, alors $x \mapsto 1/f(x)$ est une fonction continue.

Théorème 9.88.
Si f et g sont deux fonctions continues en a, alors la fonction $f + g$ est également continue en a.

Démonstration. La continuité des fonctions f et g au point a fait en sorte que pour tout choix de ϵ_1 et ϵ_2, il existe δ_1 et δ_2 tels que

$$(|x - a| \leqslant \delta_1) \Rightarrow |f(x) - f(a)| \leqslant \epsilon_1.$$

et

$$(|x - a| \leqslant \delta_2) \Rightarrow |g(x) - g(a)| \leqslant \epsilon_2.$$

La quantité que nous souhaitons analyser est $|f(x) + g(x) - f(a) - g(a)|$. Tout le jeu de la démonstration de la continuité est de triturer cette expression pour en tirer quelque chose en termes de ϵ_1 et ϵ_2. Si nous supposons avoir pris $|x - a|$ plus petit en même temps que δ_1 et que δ_2, nous avons

$$|f(x) + g(x) - f(a) - g(a)| \leqslant |f(x) - g(x)| + |g(x) - g(a)| \leqslant \epsilon_1 + \epsilon_2$$

en utilisant la formule générale $|a + b| \leqslant |a| + |b|$. Maintenant, si on choisit ϵ_1 et ϵ_2 tels que $\epsilon_1 + \epsilon_2 < \epsilon$, et les δ_1, δ_2 correspondants, on a

$$|f(x) + g(x) - f(a) - g(a)| \leqslant \epsilon,$$

pourvu que $|x - a|$ soit plus petit que δ_1 et δ_2. Le bon δ à prendre est donc le minimum de δ_1 et δ_2 qui eux-mêmes sont donnés par un choix de ϵ_1 et ϵ_2 tels que $\epsilon_1 + \epsilon_2 \leqslant \epsilon$. □

Pour résumer ces deux théorèmes, on dit que si f et g sont continues en a, alors la fonction $\alpha f + \beta g$ est également continue en a pour tout α, $\beta \in \mathbb{R}$.

PROPooVNKVooJvxarf

Proposition 9.89.
Soient des parties Ω_f et Ω_g dans $\mathbb{R}$. Soient $f\colon \Omega_f \to \mathbb{R}$ ainsi que $g\colon \Omega_g \to \mathbb{R}$ telles que

(1) g est continue en a et vaut $g(a) = \ell$.

(2) f est continue en ℓ et vaut $f(\ell) = b$.

(3) $g(\Omega_g) \subset \Omega_f$.

Alors $f \circ g$ est continue en a.

Démonstration. Soit $\epsilon > 0$. La continuité de f dit que il existe $\eta > 0$ tel que

$$y \in B(\ell, \eta) \cap \Omega_f \Rightarrow |f(y) - f(\ell)| < \epsilon. \tag{9.101}$$

La continuité de g donne $\delta > 0$ tel que

$$x \in B(a, \delta) \cap \Omega_g \Rightarrow |g(x) - g(a)| < \eta. \tag{9.102}$$

Si $x \in B(a, \delta) \cap \Omega_g$, alors $g(x) \in B(\ell, \eta) \cap \Omega_f$. Donc

$$|(f \circ g)(x) - f(\ell)| < \epsilon. \tag{9.103}$$

Mais $f(\ell) = (f \circ g)(a)$. Tout cela est la continuité de $f \circ g$ en a. □

Parmi les propriétés immédiates de la continuité d'une fonction, nous avons ceci qui est souvent bien utile.

CorNNPYooMbaYZg

Corolaire 9.90.
Si la fonction f est continue en a et si $f(a) > 0$, alors f est positive sur un intervalle autour de a.

Démonstration. Prenons $\epsilon < f(a)$ et voyons[39] ce que la continuité de f en a nous offre : il existe un δ tel que

$$(|x - a| \leqslant \delta) \Rightarrow |f(x) - f(a)| \leqslant \epsilon < f(a).$$

Nous en retenons que sur un intervalle (de largeur δ), nous avons $|f(x) - f(a)| \leqslant f(a)$. Par hypothèse, $f(a) > 0$, donc si $f(x) < 0$, alors la différence $f(x) - f(a)$ donne un nombre encore plus négatif que $-f(a)$, c'est-à-dire que $|f(x) - f(a)| > f(a)$, ce qui est contraire à ce que nous venons de démontrer. D'où la conclusion que $f(x) > 0$. □

9.5.2 La fonction la moins continue du monde

Parmi les exemples un peu sales de fonctions non continues, il y a celle-ci :

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 \text{ si } x \in \mathbb{Q} \\ 0 \text{ sinon.} \end{cases}$$

Par exemple, $\chi_{\mathbb{Q}}(0) = 1$, et[40] $\chi_{\mathbb{Q}}(\pi) = \chi_{\mathbb{Q}}(\sqrt{2}) = 0$. Bien que $\chi_{\mathbb{Q}}(0) = 1$, il n'existe *aucun* voisinage de 1 sur lequel la fonction reste proche de 1, parce que tout voisinage va contenir au moins un irrationnel. À chaque millimètre, cette fonction fait une infinité de bonds !

Cette fonction n'est donc continue nulle part.

À partir de là, nous pouvons construire la fonction suivante qui n'est continue qu'en un point :

$$f(x) = x\chi_{\mathbb{Q}}(x) = \begin{cases} x \text{ si } x \in \mathbb{Q} \\ 0 \text{ sinon.} \end{cases}$$

39. ici, nous insistons sur le fait que nous prenons ϵ *strictement* plus petit que $f(a)$.

40. Pour prouver que $\sqrt{2}$ n'est pas rationnel, c'est pas trop compliqué, mais pour prouver que π ne l'est pas non plus, il faudra encore manger de la soupe.

Cette fonction est continue en zéro. En effet, prenons $\delta > 0$; il nous faut un ϵ tel que $|x| \leqslant \epsilon$ implique $f(x) \leqslant \delta$ parce que $f(0) = 0$. Bon ben prendre simplement $\epsilon = \delta$ nous contente. Cette fonction est donc très facilement continue en zéro.

Et pourtant, dès que l'on s'écarte un tant soit peu de zéro, elle fait des bons une infinité de fois par millionième de millimètre ! Cette fonction est donc la plus discontinue du monde en tous les points, sauf un (zéro), où c'est une fonction continue !

9.5.3 Approche topologique

Nous avons vu que sur tout ensemble métrique, nous pouvons définir ce qu'est un ouvert : c'est un ensemble qui contient une boule ouverte autour de chacun de ses points. Quand on est dans un ensemble ouvert, on peut toujours un peu se déplacer sans sortir de l'ensemble.

Le théorème suivant est une très importante caractérisation des fonctions continues (de $\mathbb{R}$ dans $\mathbb{R}$) en termes de topologie, c'est-à-dire en termes d'ouverts.

¡¡ Avertissement/question au lecteur !! 9.91
Le théorème suivant n'a aucun sens parce que c'est pratiquement la définition de la continuité d'une fonction, définition 7.36. Peut-être que ce qu'on a en tête est la proposition 7.37 qui donne l'équivalence entre application continue et continu en chaque point.

ThoContInvOuvert

Théorème 9.92.
Si I est un intervalle ouvert contenu dans $\operatorname{dom} f$*, alors f est continue sur I si et seulement si pour tout ouvert $\mathcal{O}$ dans $\mathbb{R}$, l'image inverse $f|_I^{-1}(\mathcal{O})$ est ouvert.*

Démonstration. Dans un premier temps, nous allons transformer le critère de continuité en termes de boules ouvertes, et ensuite, nous passerons à la démonstration proprement dite. Le critère de continuité de f au point x dit que

$$\forall \delta > 0, \exists\, \epsilon > 0 \text{ tel que } \big(|x-a| < \epsilon\big) \Rightarrow |f(x) - f(a)| < \delta. \tag{9.104}$$

EqDEfCOntAn

Cette condition peut être exprimée sous la forme suivante :

$$\forall \delta > 0, \exists \epsilon \text{ tel que } a \in B(x,\epsilon) \Rightarrow f(a) \in B\big(f(x),\delta\big),$$

ou encore

$$\forall \delta > 0, \exists \epsilon \text{ tel que } f\big(B(x,\epsilon)\big) \subset B\big(f(x),\delta\big). \tag{9.105}$$

EqRedefContBoules

Jusque ici, nous n'avons fait que du jeu de notations. Nous avons exprimé en termes de topologie des inégalités analytiques. La condition (9.105) est le plus souvent utilisée comme définition de la continuité d'une fonction en x, lorsque le contexte ne demande pas de définitions plus générales. Si tel est le choix, il faut pouvoir retrouver (9.104) à partir de (9.105).

Passons maintenant à la démonstration proprement dite du théorème.

D'abord, supposons que f est continue sur I, et prenons $\mathcal{O}$, un ouvert quelconque. Le but est de prouver que $f|_I^{-1}(\mathcal{O})$ est ouvert. Pour cela, nous prenons un point $x_0 \in f|_I^{-1}(\mathcal{O})$ et nous allons trouver un ouvert autour de ce point, contenu dans $f|_I^{-1}(\mathcal{O})$. Nous écrivons $y_0 = f(x_0)$. Évidemment, $y_0 \in \mathcal{O}$, donc on a une boule autour de y_0 qui est contenue dans $\mathcal{O}$, soit donc $\delta > 0$ tel que

$$B(y_0,\delta) \subset \mathcal{O}.$$

Par hypothèse, f est continue en x_0, et nous pouvons donc y appliquer le critère (9.105). Il existe donc $\epsilon > 0$ tel que

$$f\big(B(x_0,\epsilon)\big) \subset B\big(f(x_0),\delta\big) \subset \mathcal{O}.$$

Cela prouve que $B(x_0,\epsilon) \subset f|_I^{-1}(\mathcal{O})$.

Dans l'autre sens, maintenant. Nous prenons $x_0 \in I$ et nous voulons prouver que f est continue en x_0, c'est-à-dire que pour tout δ, nous cherchons un ϵ tel que $f\big(B(x_0,\epsilon)\big) \subset B\big(f(x_0),\delta\big)$. Oui,

mais $B\big(f(x_0),\delta\big)$ est ouverte, donc par hypothèse, $f|_I^{-1}\Big(B\big(f(x_0),\delta\big)\Big)$ est ouvert, inclus dans I et contient x_0. Donc il existe un ϵ tel que

$$B(x_0,\epsilon)\subset f|_I^{-1}\Big(B\big(f(x_0),\delta\big)\Big),$$

et donc tel que

$$f\big(B(x_0,\epsilon)\big)\subset B\big(f(x_0),\delta\big),$$

ce qu'il fallait prouver. □

ThoValInter

Théorème 9.93 (Théorème des valeurs intermédiaires)**.**
Soit f, une fonction continue sur $[a,b]$, et supposons que $f(a)<f(b)$. Alors pour tout y tel que $f(a)\leqslant y\leqslant f(b)$, il existe un $x\in[a,b]$ tel que $f(x)=y$.

Démonstration. Nous savons que $[a,b]$ est connexe parce que c'est un intervalle (proposition 9.53). Donc $f\big([a,b]\big)$ est connexe (lemme 7.218) et donc est un intervalle (à nouveau la proposition 9.53). Étant donné que $f\big([a,b]\big)$ est un intervalle, il contient toutes les valeurs intermédiaires entre n'importe quels deux de ses éléments. En particulier toutes les valeurs intermédiaires entre $f(a)$ et $f(b)$. □

NORMooTQWWooQVPWIJ

9.94.
Une façon classique d'utiliser le théorème des valeurs intermédiaires 9.93. Si $f\colon[0,\infty[\to[0,\infty[$ est continue et vérifie $f(0)=0$ et $\lim_{x\to\infty}f(x)=\infty$, alors f est surjective.

En effet si $y\in[0,\infty[$, alors il existe $a\in[0,\infty[$ tel que $f(a)>y$. Donc il existe $x\in[0,a]$ tel que $f(x)=y$.

CorImInterInter

Corolaire 9.95.
L'image d'un intervalle par une fonction continue est un intervalle.

Démonstration. Soient I un intervalle, $\alpha<\beta\in f(I)$ et $\gamma\in\,]\alpha,\beta[$. Nous considérons $a,b\in I$ tels que $\alpha=f(a)$ et $\beta=f(b)$. Par le théorème des valeurs intermédiaires 9.93, il existe $t\in\,]a,b[$ tel que $f(t)=\gamma$. Par conséquent $\gamma\in f(I)$. □

9.5.4 Limite sur des chemins

PROPooREUUooUNDFRi

Proposition 9.96.
Soit une application $f\colon X\to Y$ entre espaces topologiques. Nous supposons que $\lim_{x\to a}f(x)=b$. Alors pour toute application continue $\gamma\colon\mathbb{R}\to X$ telle que $\gamma(0)=a$, nous avons

$$\lim_{t\to0}(f\circ\gamma)(t)=b. \tag{9.106}$$

9.5.5 Racine carrée

DEFooGQTYooORuvQb

Corolaire-Définition 9.97 (Existence de la racine carrée)**.**
*Si $x\geqslant0$ dans $\mathbb{R}$ alors il existe un unique réel $y\geqslant0$ tel que $y^2=x$. Ce nombre est noté $\sqrt{x}$ et est nommé **racine carrée** de x.*

Démonstration. La fonction $f\colon t\mapsto t^2$ est continue et strictement croissante. Nous avons $f(0)=0$ et[41] $f(x+1)>x$. Donc le théorème des valeurs intermédiaires 9.93 nous assure qu'il existe un unique $y\in[0,x+1]$ tel que $f(y)=x$. □

LEMooWSVNooKsymDy

Lemme 9.98.
Quelques formules.

(1) $\sqrt{xy}=\sqrt{x}\sqrt{y}$ *si* $x,y\geqslant0$.

41. Faites deux cas suivant $x\geqslant1$ ou non si vous le voulez, moi je prends $x+1$.

ITEMooEPHBooCEeJOD
(2) $\sqrt{\lambda^2 x} = |\lambda|\sqrt{x}$ *si* $x \geqslant 0$.

LEMooSBOAooOOIotR
Lemme 9.99.
La fonction racine carrée est strictement croissante.

Démonstration. Supposons que $x < y$. Si $\sqrt{x} > \sqrt{y}$, alors la croissance de la fonction carré donne $x > y$ qui est contraire à l'hypothèse. □

9.5.6 Module sur les nombres complexes

LEMooVHDAooJyoakR
Lemme-Définition 9.100.
Si $z \in \mathbb{C}$, alors $z\bar{z}$ est un réel positif.
Nous définissons le ***module*** *sur $\mathbb{C}$ par*[42]

$$|z| = \sqrt{z\bar{z}}. \tag{9.107}$$

Démonstration. Prouvons que $z\bar{z}$ est un réel positif. En effet si $z = a + bi$ alors

$$z\bar{z} = (a + bi)(a - bi) = a^2 - abi + abi + b^2 = a^2 + b^2 \geqslant 0. \tag{9.108}$$

□
LEMooJRLWooScVrkG
Lemme 9.101.
Si $z \in \mathbb{C}$ nous avons

EQooQDGTooBejPUE
$$z + \bar{z} = 2\operatorname{Re}(z). \tag{9.109}$$

Démonstration. Soit $z = a + bi$ avec $a, b \in \mathbb{R}$. Nous avons

$$z + \bar{z} = a + bi + a - bi = 2a = 2\operatorname{Re}(z). \tag{9.110}$$

□
PROPooUMVGooIrhZZg
Proposition 9.102.
Pour tout nombres complexes $z = a + bi$ et z', nous avons

ITEMooYBJVooGXiDSd
(1) $z\bar{z} = a^2 + b^2$;

ITEMooCGLSooKHbzkn
(2) $\bar{\bar{z}} = z$;

ITEMooDKWDooUjEuZA
(3) $|z| = |\bar{z}|$;

ITEMooFXKYooUOXbwH
(4) $|zz'| = |z|\,|z'|$;

ITEMooUJHPooUFdvqB
(5) $|z + z'| = \sqrt{|z|^2 + |z'|^2 + 2\operatorname{Re}(z\bar{z}')}$.

ITEMooDVMDooFDmOur
(6) $|z_1 + z_2| \leqslant |z_1| + |z_2|$

ITEMooHBIEooEhzlwI
(7) $|z_1 + z_2| = |z_1| + |z_2|$ *si et seulement si* $z_1\bar{z}_2 \in \mathbb{R}^+$.

ITEMooMCAAooTuUxLV
(8) *Nous avons* $|\operatorname{Re}(z)| \leqslant |z|$ *et nous avons l'égalité si et seulement si* $z \in \mathbb{R}$.

Démonstration. En plusieurs points.

(1) **Pour (1)** Calcul direct :

$$z\bar{z} = (a + bi)(a - bi) = a^2 - abi + abi + b^2 = a^2 + b^2. \tag{9.111}$$

(2) **Pour (2)** On a :

$$\bar{\bar{z}} = \overline{(a - bi)} = a + bi. \tag{9.112}$$

42. Définition de la racine carrée : 9.97.

(3) **Pour (3)** Même calcul que pour (1).

(4) **Pour (4)** Puisque les deux membres de l'égalité à prouver sont positifs, il est suffisant de prouver l'égalité des carrés (lemme 9.99). Nous avons

$$zz' = (a+bi)(a'+b'i) = aa' - bb' + i(ab' + ba'), \tag{9.113}$$

donc

$$\begin{aligned} |zz'|^2 &= (aa' - bb')^2 + (ab' + ba')^2 && (9.114a)\\ &= (aa')^2 + (bb')^2 - 2aa'bb' + (ab')^2 + (ba')^2 + 2ab'ba' && (9.114b)\\ &= a^2(a'^2 + b'^2) + b^2(b'^2 + a'^2) && (9.114c)\\ &= (a^2+b^2)(a'^2+b'^2). && (9.114d) \end{aligned}$$

D'autre part,

$$\big(|z||z'|\big)^2 = |z|^2|z'|^2 = (a^2+b^2)(a'^2+b'^2). \tag{9.115}$$ EQooRSGGooGfWTrS

(5) **Pour (5)** En utilisant la formule (9.109) pour $z'\bar{z}$, nous avons :

$$|z+z'|^2 = (z+z')(\bar{z}+\bar{z}') = z\bar{z} + z\bar{z}' + z'\bar{z} + z'\bar{z}' = |z|^2 + |z'|^2 + 2\operatorname{Re}(z\bar{z}'). \tag{9.116}$$

(6) **Pour (8)** En plusieurs points.

(6a) **L'inégalité** Par croissance de la fonction racine carrée nous avons, en posant $z = a + bi$:

$$|\operatorname{Re}(z)| = |a| = \sqrt{|a|^2} \leqslant \sqrt{|a|^2 + |b|^2} = |z|. \tag{9.117}$$ EQooBFANooKcSsWi

(6b) **Égalité dans un sens** Si $|\operatorname{Re}(z)| = |z|$, alors toutes les inégalités dans (9.117) sont des égalités. En particulier

$$\sqrt{|a|^2} = \sqrt{|a|^2 + |b|^2}. \tag{9.118}$$

Par stricte croissance de la fonction racine carré, nous déduisons que $|b|^2 = 0$ et donc que $b = 0$.

(6c) **Égalité dans l'autre sens** Si $z \in \mathbb{R}$, alors $\operatorname{Re}(z) = z$, et nous avons l'égalité.

(7) **Pour (6)** Nous posons $z_1 = a_1 + b_1 i$ et $z_2 = a_2 + b_2 i$. Ensuite nous calculons, en utilisant (9.115) :

$$\big(|z_1| + |z_2|\big)^2 = |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = a_1^2 + b_1^2 + a_2^2 + b_2^2 + 2\sqrt{(a_1^2+b_1^2)(a_2^2+b_2^2)}. \tag{9.119}$$

et

$$\begin{aligned} |z_1 + z_2|^2 &= |(a_1+a_2)^2 + i(b_1+b_2)|^2 && (9.120a)\\ &= (a_1+a_2)^2 + (b_1+b_2)^2 && (9.120b)\\ &= a_1^2 + a_2^2 + 2a_1a_2 + b_1^2 + b_2^2 + 2b_1b_2. && (9.120c) \end{aligned}$$

En faisant la différence,

$$\big(|z_1| + |z_2|\big)^2 - |z_1 + z_2|^2 = 2\sqrt{(a_1^2+b_1^2)(a_2^2+b_2^2)} - 2(a_1a_2 + b_1b_2). \tag{9.121}$$

Pour prouver que cette différence est positive, nous comparons les carrés des deux termes :

$$\begin{aligned} A &= (a_1^2+b_1^2)(a_2^2+b_2^2) && (9.122a)\\ B &= (a_1a_2 + b_1b_2)^2. && (9.122b) \end{aligned}$$

Nous avons :

$$A = a_1^2a_2^2 + a_1^2b_2^2 + b_1^2a_2^2 + b_1^2b_2^2 \tag{9.123}$$

et

$$B = a_1^2a_2^2 + b_1^2b_2^2 + 2a_1a_2b_1b_2. \tag{9.124}$$

Et un petit calcul montre enfin que

$$A - B = a_1^2b_2^2 + b_1^2a_2^2 - 2a_1a_2b_1b_2 = (a_1b_2 - b_1a_2)^2 \geqslant 0. \tag{9.125}$$

□
LEMooNIXZooDxfpNM

Lemme 9.103.
Soient des réels strictement positifs a_{ij} et des complexes $z_i \in \mathbb{C}$ tels que pour tout $j = 1, \dots, n$ nous ayons

$$\sum_{i=1}^{n} a_{ij}|z_i| = |\sum_i a_{ij} z_i|. \tag{9.126}$$

Alors les z_i sont colinéaires au sens où il existe des nombres réels positifs s_i et un $z_0 \in \mathbb{C}$ tels que $z_i = s_i z_0$.

LEMooXJBJooFDmhnV

Lemme 9.104.
Si deux nombres complexes $a, b \in \mathbb{C}$ vérifient $a\bar{b} \in \mathbb{R}$, alors nous sommes dans un des deux cas suivants :

- *$b = 0$*
- *il existe $\lambda \in \mathbb{R}$ tel que $a = \lambda b$.*

De façon équivalente, il existe $\alpha, \beta \in \mathbb{R}$ non tous deux nuls tels que $\alpha a + \beta b = 0$.

Démonstration. Nous écrivons $a = a_1 + ia_2$ et $b = b_1 + ib_2$ avec $a_i, b_i \in \mathbb{R}$. Nous supposons $b \neq 0$. Nous effectuons la multiplication $a\bar{b} = (a_1 + ia_2)(b_1 - ib_2)$ et nous annulons la partie imaginaire :

$$a_2 b_1 - a_1 b_2 = 0. \tag{9.127}$$

Si $b_1 = 0$ alors $a_1 b_2 = 0$ avec $b_2 \neq 0$, ce qui implique $a_1 = 0$. Donc $a = ia_2$, $b = ib_2$. Résultat obtenu.

Si $b_1 \neq 0$ alors

$$a_2 = \frac{a_1 b_2}{b_1}, \tag{9.128}$$

et nous avons alors

$$a = \frac{a_1}{b_1} b. \tag{9.129}$$

Mission accomplie.

Nous prouvons à présent la formulation équivalente. Si $b = 0$ il suffit de prendre $\alpha = 0$. Si $a = \lambda b$ il faut prendre $\alpha = \beta/\lambda$.

Dans l'autre sens, si $\alpha \neq 0$ alors $a = -(\beta/\alpha)b$ et si $\alpha = 0$ alors $\beta \neq 0$ et il reste $b = 0$. □
PROPooZJAXooYwSSvo

Proposition 9.105.
La paire $(\mathbb{C}, |.|)$ est un espace vectoriel normé[43].

Démonstration. Nous devons prouver les différents points de la définition 7.158.

(1) $|z| \geqslant 0$ parce que la racine carrée prend ses valeurs dans $\mathbb{R}^+$.
(2) Si $|z| = 0$, alors, en notant $z = a + bi$ nous avons $a^2 + b^2 = 0$. Cela implique $a = b = 0$ (vous pouvez soit invoquer le lemme 1.433, soit ne rien dire et faire comme si c'était évident).
(3) Si $\lambda \in \mathbb{R}$, alors $(\lambda z)(\overline{\lambda z}) = \lambda^2 z\bar{z}$ et nous avons

$$|\lambda z| = \sqrt{\lambda^2 z\bar{z}} = |\lambda|\sqrt{z\bar{z}} = |\lambda||z|. \tag{9.130}$$

(4) L'inégalité $|x + y| \leqslant |x| + |y|$ est la proposition 9.102(7).

□
PROPooXLARooYSDCsF

Proposition 9.106.
Si z_1 et z_2 sont des nombres complexes, alors

$$|z_1 z_2| = |z_1||z_2|. \tag{9.131}$$

Nous avons aussi, pour tout $n \in \mathbb{N}$,

$$|z^n| = |z|^n. \tag{9.132}$$
EQooATTQooRpIeCo

43. Définition d'une norme : 7.158.

Démonstration. D'abord $(a+bi)(c+di) = ac - db + (ad+bc)i$, de telle sorte que

$$|(a+bi)(c+di)|^2 = (ac-bd)^2 + (ad+bc)^2. \tag{9.133}$$

Mais en calculant d'autre part $|a+bi|^2|c+di|^2$, nous tombons sur la même valeur.

Une simple récurrence permet de conclure que $|z^n| = |z|^n$. □

Voilà. Vous êtes déjà content d'apprendre que l'on peut démontrer $|z^n| = |z|^n$ sans faire appel à la forme trigonométrique des nombres complexes.

LEMooONLNooXLNbtB

Lemme 9.107.
Pour tout $z \in \mathbb{C}$ nous avons $z\bar{z} = \bar{z}z = |z|^2$.

9.5.7 Théorème de Perron-Frobenius

Définition 9.108.
*Pour une matrice carrée T, nous notons $T \geqslant 0$ si $T_{ij} \geqslant 0$ pour tout i et j. Dans ce cas nous disons que T est **positive**.*

*Une matrice carré positive est **primitive** si il existe un entier $k > 0$ tel que $T^k > 0$.*

DEFooRFQCooQrLPVw

Définition 9.109.
Si T est une matrice $n \times n$ et si $x \in \mathbb{R}^n$, nous notons

$$xT = \sum_{ij} x_i T_{ij} e_j. \tag{9.134}$$

*Et nous disons que $x \in \mathbb{R}^n$ est un **vecteur propre à gauche** pour la valeur propre à gauche $\lambda \in \mathbb{C}$ si $xT = \lambda x$.*

THOooRSPJooMCFeeP

Théorème 9.110 (Perron-Frobenius[316])**.**
Soit une matrice positive et primitive T. Il existe une valeur propre ρ telle que

(1) $\rho > 0$,

(2) La matrice T a des vecteurs propres strictement positifs à gauche et à droite pour la valeur propre ρ.

(3) Pour toute valeur propre $\lambda \neq \rho$ nous avons $\rho > |\lambda|$

(4) Les espaces propres à gauche et à droite de T pour la valeur propre ρ sont de dimension 1.

Démonstration. Nous divisons la preuve en trois parties. La première va tout démontrer pour les vecteurs propres à gauche, la seconde fera la même chose pour les vecteurs propre à droite, et la troisième fera la synthèse.

– À gauche –

Soit $D = \{x \in \mathbb{R}^n \text{ tel que } x \neq 0, x \geqslant 0\}$. Nous considérons l'application

$$\begin{aligned} r\colon D &\to \mathbb{R} \\ x &\mapsto \min_j \frac{(xT)_j}{x_j} \end{aligned} \tag{9.135}$$

où la fraction vaut ∞ si $x_j = 0$.

(1) **r est majorée supérieurement** Pour chaque j (y compris ceux pour qui $x_j = 0$) nous avons l'inégalité suivante dans $\mathbb{R}$:

$$x_j r(x) \leqslant (xT)_j. \tag{9.136}$$

En multipliant par e_j et en faisant la somme, $r(x)x \leqslant xT$. Nous posons $u = \sum_i e_i$ et nous prenons le produit scalaire :

$$r(x)x \cdot u \leqslant xT \cdot u = x \cdot Tu. \tag{9.137}$$

EQooBRRSooUtIPXc

En posant $K = \max_j \sum_i T_{ij}$ nous avons $Tu \leqslant Ku$ parce que

$$(Tu)_k = \sum_l T_{kl} \underbrace{u_l}_{=1} = \sum_l T_{kl} \leqslant K = (Ku)_k. \tag{9.138}$$

Pour tout $x \in D$ nous avons $r(x) \leqslant K$. Nous utilisons ça pour continuer (9.137) :

$$r(x)x \cdot u \leqslant xT \cdot u = x \cdot Tu \leqslant Kx \cdot u. \tag{9.139}$$

Vu que $x \geqslant 0$ et $u \geqslant 0$, nous avons $x \cdot u \geqslant 0$ nous nous pouvons simplifier : $r(x) \leqslant K$ pour tout $x \in D$.

(2) **Une minoration** Nous savons que T est une matrice positive et même primitive, de telle sorte que T n'ait aucune colonne nulle. Donc

$$r(u) = \min_j \frac{\sum_i u_i T_{ij}}{u_j} = \min_j \sum_i T_{ij} > 0. \tag{9.140}$$

Nous avons donc

$$r(u) > 0. \tag{9.141}$$

(3) **Définition de ρ** Nous posons

$$\rho = \sup_{x \in D} r(x). \tag{9.142}$$

Vu que que nous avons déjà prouvé nous avons $0 < \rho \leqslant K$.

Notons que si $\lambda \in \mathbb{R}$, nous avons $r(x) = r(\lambda x)$, de telle sorte que

$$\rho = \sup_{\substack{x \in D \\ \|x\| \leqslant 1}} r(x). \tag{9.143}$$

(4) **Un compact** En posant

$$D' = \big(D \cap \overline{B(0,1)}\big) \backslash B(0, \epsilon) \tag{9.144}$$

avec $\epsilon < 1$, nous avons $r(D') = r(D)$ parce que $r(x) = r(\lambda x)$. La partie D' est évidemment bornée.

Vu que D et $\overline{B(0,1)}$ sont fermés, l'intersection est fermée (lemme 7.6(1)). La différence entre un fermé et un ouvert est fermée (lemme 7.7). Donc D' est fermé. Le théorème de Borel-Lebesgue 9.24 nous dit alors que D' est compact.

(5) **r est continue** Sur D', chacune des applications $x \mapsto (xT)_j / x_j$ est continue. Donc le minimum est continu et $r \colon D' \to \mathbb{R}$ est une application continue sur un compact.

(6) **v et ρ** Le théorème de Weierstrass 7.146 nous indique que r est bornée et atteint ses bornes sur D'. Il existe $v \in D'$ tel que $r(v) \geqslant r(x)$ pour tout $x \in D'$. Vu que par définition

$$\rho = \sup_{\substack{x \in D \\ \|x\| \leqslant 1}} r(x) = \sup_{x \in D'} r(x), \tag{9.145}$$

nous avons en fait $r(v) = \rho$.

(7) **Un mini sous-résultat** SPITEMoqBKGPooQDQWdK Nous démontrons un sous-résultat qui sera utilisé quelques fois dans la suite. Si $x \in \mathbb{R}^n$ vérifie $x \geqslant 0$ et $(xT)_j \geqslant \rho x_j$ pour tout j, alors

(7a) $xT = \rho x$, c'est-à-dire que x est un vecteur propre à gauche de T.

(7b) $x > 0$.

Par hypothèse du théorème, il existe $k > 0$ tel que $T^k > 0$. Nous posons $y = xT^k$. Alors, si $xT - \rho x \neq 0$ nous avons

$$yT - \rho y = (xT - \rho x)T^k > 0. \tag{9.146}$$

La dernière inégalité est parce que $xT - \rho x \geqslant 0$ et $T^k > 0$. En développant, cette inégalité signifie que pour tout j,

$$\sum_i y_i T_{ij} > \rho y_j, \tag{9.147}$$

ou encore que $\rho < \frac{\sum_i y_i T_{ij}}{y_j}$ pour tout j. En particulier pour le j qui réalise le minimum, $\rho < r(y)$, ce qui contredit ma maximalité de ρ. Nous en déduisons que $xT - \rho x = 0$.
À partir de là, nous déduisons que $x > 0$. En effet, en appliquant k fois l'égalité $xT = \rho x$, nous avons $xT^k = \rho^k x$. Comme $x \geqslant 0$ et $T^k > 0$, nous avons $xT^k > 0$. Vu que $\rho \neq 0$ nous en déduisons que $x > 0$.

(8) **v et ρ, suite** Nous avons $r(v) = \rho$. En particulier pour tout j nous avons $\rho v_j \leqslant \sum_i v_i T_{ij}$. Et comme $v \in D'$ nous avons aussi $v \geqslant 0$. En vertu de notre sous-résultat, nous déduisons que v est un vecteur propre à gauche de T : $vT = \rho v$, et que $v > 0$.
Cela prouve que T a un vecteur propre à gauche strictement positif.

(9) **Le plus grand** Nous prouvons que pour toute valeur propre $\lambda \in \mathbb{C}$, nous avons $\rho > |\lambda|$. Soit une valeur propre $\lambda \in \mathbb{C}$. Vu que le spectre « à gauche » et le spectre « à droite » sont identiques[44], T possède un vecteur propre à gauche : $xT = \lambda x$. Pour chaque j nous avons

$$\sum_i x_i T_{ij} = \lambda x_j, \tag{9.148}$$

et en passant au module, $|\lambda x_j| \leqslant |x_i| T_{ij}$. En posant $y = \sum_i |x_i| e_i$ nous avons $y \geqslant 0$ et

$$|\lambda| y_j \leqslant \sum_i y_i T_{ij}, \tag{9.149}$$

et donc

$$|\lambda| \leqslant \sum_i \frac{y_i T_{ij}}{y_j} \tag{9.150}$$

En passant au minimum sur j, cela donne $|\lambda| \leqslant r(y) \leqslant \rho$.
Nous avons montré que $|\lambda| \leqslant \rho$. Nous prouvons à présent que $|\lambda| \neq \rho$ en supposant que $|\lambda| = \rho$. Toujours en posant $y_i = |x_i|$, nous avons

$$\rho y_j \leqslant \sum_i y_i T_{ij} = (yT)_j. \tag{9.151}$$

Le sous-résultat (7) nous indique qu'alors y est un vecteur propre à gauche de T pour la valeur propre ρ : $yT = \rho u$. En itérant et en déballant, pour chaque j nous avons

$$\sum_i |x_i| T_{ij}^k = \rho^k |x_j|. \tag{9.152}$$

En faisant de même à partir de $xT^k = \lambda^k x$, nous trouvons

$$|\sum_i x_i T_{ij}^k| = |\lambda|^k |x_j|. \tag{9.153}$$

Vu que $|\lambda| = \rho$, nous pouvons égaliser les membres de droite de (9.152) et (9.153) :

$$\sum_i |x_i| T_{ij}^k = |\sum_i x_i T_{ij}^k|. \tag{9.154}$$

Le lemme 9.103 nous indique que les x_i sont colinéaires (dans $\mathbb{C}$), c'est-à-dire qu'il existe des réels $s_i \geqslant 0$ et $z_0 \in \mathbb{C}$ tels que $x_i = s_i z_0$ pour tout i.
En utilisant ce fait dans l'équation $xT = \lambda x$, nous trouvons $\sum_i x_i T_{ij} = \lambda x_j$ et

$$\sum_i s_i T_{ij} = \lambda s_j. \tag{9.155}$$

Cela prouve que λ est réel et positif parce que l'équation (9.155) est valable pour tout j et qu'il y en a bien un pour lequel $s_j \neq 0$. Le nombre réel positif λ vérifiant $|\lambda| = \rho$, nous déduisons que $\lambda = \rho$.

44. Lemme 8.89.

(10) **Espace propre de dimension 1** Nous avons déjà la vecteur propre v : $vT = \rho v$. Considérons un second vecteur propre à gauche $y \neq 0$: $yT = \rho y$. Supposons que y et v ne sont pas colinéaires et voyons les fonctions

$$\begin{aligned}\eta_i\colon \mathbb{R} &\to \mathbb{R}\\ c &\mapsto (v - cy)_i.\end{aligned} \tag{9.156}$$

Pour tout i nous avons $\eta_i(0) = v_i > 0$. Supposons pour fixer les idées qu'au moins un des y_i est positif. Alors pour chaque i, il existe un $c_i > 0$ tel que $v_i - cy_i = 0$. Nous prenons pour valeur de c le plus petit de ceux-là. Nous avons donc

$$\eta = v - cy \geqslant 0, \tag{9.157}$$

alors qu'au moins une des composantes de η est nulle. Et en plus $\eta T = \rho\eta$. Notre sous-résultat préféré nous dit qu'alors $\eta > 0$. Ah tiens, ça c'est une contradiction. Nous en déduisons que v et y sont colinéaires dans $\mathbb{R}^n$.

Nous avons terminé de prouver tout ce que nous devions faire « à gauche ».

À droite

En reprenant exactement les mêmes étapes [45], nous prouvons qu'il existe une valeur propre à droite ρ' telle que

(1) $\rho' > 0$
(2) ρ' a des vecteurs propres strictement positifs à droite.
(3) $\rho' > |\lambda|$ pour toute valeur propre $\lambda \neq \rho'$.
(4) L'espace propre à droite pour la valeur propre ρ' est de dimension 1.

Conclusion

Les nombres réels strictement positifs ρ et ρ' sont des valeurs propres. Si $\rho \neq \rho'$, alors les propriétés de ρ disent que $\rho > \rho'$. De même les propriétés de ρ' disent que $\rho' > \rho$. Bref $\rho = \rho'$.

Le nombre ρ vérifie toutes les propriétés du théorème. □

9.5.8 Limite à gauche et à droite

DEFooIBBQooFEwyMc

Définition 9.111 ([1]).
Soient un espace topologique X et une application $\gamma\colon I \to X$ où I est un voisinage ouvert de $a \in \mathbb{R}$.

*L'élément $b \in E$ est une **limite à gauche** de γ quand $t \to a$ lorsque b est une limite de la restriction de γ à $]-\infty, a[$. Dans ce cas nous parlons de limite lorsque t tend vers a^-.*

*L'élément $b \in E$ est une **limite à droite** de γ quand $t \to a$ lorsque b est une limite de la restriction de γ à $]a, \infty[$. Dans ce cas nous parlons de limite lorsque t tend vers a^+.*

Lorsqu'il y a unicité de la limite, nous notons les limites à gauche et à droite par

$$\lim_{t\to a^-} \gamma(t) \quad et \quad \lim_{t\to a^+} \gamma(t) \tag{9.158}$$

PROPooGUNLooSQrJKg

Proposition 9.112 ([1]).
Soient un espace topologique X et une application $\gamma\colon I \to X$ où I est un voisinage ouvert de $a \in \mathbb{R}$. L'élément $b \in E$ est une limite à gauche [46] et à droite de γ quand $t \to a$ si et seulement si il est une limite.

Dans le cas où la limite est unique [47], nous pouvons dire que

$$\begin{cases} \lim_{t\to a^+} \gamma(t) = b & \text{(9.159a)}\\ \lim_{t\to a^-} \gamma(t) = b & \text{(9.159b)}\end{cases}$$

45. Je n'ai pas vérifié ; écrivez-moi si cela pose problème.
46. Définition 9.111.
47. Par exemple lorsque X est séparé, voir la proposition 7.113.

si et seulement si[48]

$$\lim_{t\to a}\gamma(t)=b. \tag{9.160}$$

Démonstration. Nous nommons $I^+=I\cap]a,\infty[$ et $I^-=I\cap]-\infty,a[$. Nous supposons que γ admet b comme limite à gauche et à droite. Soit un voisinage U de b dans X. La limite à gauche donne l'existence de $\delta_1>0$ tel que

$$\gamma\big(B(a,\delta_1)\cap I^-\big)\subset U, \tag{9.161}$$

EQooEGOHooHauKRX

et la limite à droite donne un $\delta_2>0$ tel que

$$\gamma\big(B(a,\delta_2)\cap I^+\big)\subset U. \tag{9.162}$$

EQooXTMBooYBGxqi

Notez que nous avons choisi δ_1 et δ_2 de telle sorte à ce que $B(a,\delta_1)$ et $B(a,\delta_2)$ soient contenus dans I.

Nous considérons $\delta<\min\{\delta_1,\delta_2\}$, et nous prenons $x\in B(a,\delta)\backslash\{a\}$. Il y a deux possibilités : $x>a$ et $x<a$. Si $x>a$, alors nous avons $x\in B(a,\delta)\cap I^+\subset B(a,\delta_1)\cap I^+$, de telle sorte que $\gamma(x)\in U$ par (9.162). Si $x<a$, alors nous avons également $\gamma(x)\in U$, en utilisant (9.161). □

9.6 Norme à partir d'un produit scalaire

PROPooJLWSooNicxQV

Proposition 9.113 ([1])**.**
Soit un espace vectoriel complexe E muni d'une forme sesquilinéaire $\langle .,.\rangle$ telle que la forme quadratique réelle $x\mapsto\langle x,x\rangle$ soit définite positive. La formule[49]

$$\|x\|=\sqrt{\langle x,x\rangle} \tag{9.163}$$

EQooZIXRooMGcsXY

est une norme sur E.

DEFooGUXNooXwCsrq

Définition 9.114.
Dans le cas de $\mathbb{C}^n$, nous considérons toujours la norme associée à la forme (8.355) par la proposition 9.113, c'est-à-dire, pour rappel :

$$\langle x,y\rangle=\sum_{k=1}^{n}x_k\overline{y_k}, \tag{9.164}$$

et

$$\|x\|=\sqrt{\langle x,x\rangle} \tag{9.165}$$

pour tout $x\in\mathbb{C}^n$.

9.6.1 Continuité de la racine carrée, invitation à la topologie induite

Pourquoi nous intéresser particulièrement à la fonction racine carrée ? Parce qu'elle a une sale condition d'existence : son domaine de définition n'est pas ouvert. Or dans tous les théorèmes de continuité d'approche topologique que nous avons vus, nous avons donné des conditions *pour tout ouvert*. Nous nous attendons donc à avoir des difficultés avec la continuité de $\sqrt{x}$ en zéro.

Prenons I, n'importe quel intervalle ouvert dans $\mathbb{R}^+$, et voyons que la fonction[50]

$$\begin{aligned}f\colon\mathbb{R}^+&\to\mathbb{R}^+\\ x&\mapsto\sqrt{x}\end{aligned} \tag{9.166}$$

est continue sur I. Remarquons déjà que si I est un ouvert dans $\mathbb{R}^+$, il ne peut pas contenir zéro. Avant de nous lancer dans notre propos, nous prouvons un lemme qui fera tout le travail[51].

48. Cet énoncé doit être un tout petit peu modifié si vous utilisez la limite pointée au lieu de la limite épointée.
49. Pour la racine carrée, définition 9.97.
50. La racine carrée est définie en 9.97.
51. C'est toujours ingrat d'être un lemme : on fait tout le travail et c'est toujours le théorème qui est nommé.

Lemme 9.115.
Soit $\mathcal{O}$, un ouvert dans $\mathbb{R}^+$. Alors $\mathcal{O}^2 = \{x^2 \text{ tel que } x \in \mathcal{O}\}$ est également ouvert.

Démonstration. Un élément de $\mathcal{O}^2$ s'écrit sous la forme x^2 pour un certain $x \in \mathcal{O}$. Le but est de trouver un ouvert autour de x^2 qui soit contenu dans $\mathcal{O}^2$. Étant donné que $\mathcal{O}$ est ouvert, on a une boule centrée en x contenue dans $\mathcal{O}$. Nous appelons δ le rayon de cette boule :

$$B(x, \delta) \subset \mathcal{O}.$$

Étant donné que cet ensemble est connexe, nous savons par le lemme 7.218 que $B(x,\delta)^2$ est également connexe (parce que la fonction $x \mapsto x^2$ est continue). Son plus grand élément est $(x+\delta)^2 = x^2 + \delta^2 + 2x\delta > x^2 + \delta^2$, et son plus petit élément est $(x-\delta)^2 = x^2 + \delta^2 - 2x\delta$.

Ce qui serait pas mal, c'est que ces deux bornes entourent x^2 ; de cette façon elles définiraient un ouvert autour de x^2 qui soit dans $\mathcal{O}^2$. Hélas, c'est pas gagné que $x^2 + \delta^2 - 2x\delta$ soit plus petit que x^2.

Heureusement, en fait c'est vrai, parce que d'une part, comme $\mathcal{O} \subset \mathbb{R}^+$, on a $x > 0$, et d'autre part, pour que $\mathcal{O}$ soit positif, il faut que $\delta < x$. Donc on a évidemment $\delta < 2x$, et donc

$$x^2 + \delta^2 - 2x\delta = x^2 + \delta \underbrace{(\delta - 2x)}_{<0} < x^2.$$

Et nous en avons fini : l'ensemble

$$B(x,\delta)^2 =]x^2 + \delta^2 - 2x\delta, x^2 + \delta^2 + 2x\delta[\subset \mathcal{O}^2$$

est un intervalle qui contient x^2, et donc qui contient une boule ouverte centrée en x^2. □

Maintenant nous pouvons nous attaquer à la continuité de la racine carrée sur tout ouvert positif en utilisant le théorème 9.92.

Proposition 9.116.
L'application

$$\begin{aligned} f\colon [0,\infty[&\to \mathbb{R} \\ x &\mapsto \sqrt{x} \end{aligned} \tag{9.167}$$

est continue.

Démonstration. Soit un ouvert $\mathcal{O}$ de $\mathbb{R}$. Il existe un $r > 0$ tel que $B(y,r) \subset \mathcal{O}$. Nous allons restreindre plus fortement r plus tard. Soit $x \in f^{-1}(\mathcal{O})$; nous allons prouver que $B(x,\epsilon) \cap [0,\infty[\subset f^{-1}(\mathcal{O})$ pour un ϵ assez petit. Le théorème 7.8 dira alors que $f^{-1}(\mathcal{O})$ est ouvert.

Nous notons $y = f(x)$. Par croissance de la fonction racine carrée, nous avons

$$f\Big(B(x,\epsilon)\Big) =]\sqrt{x-\epsilon}, \sqrt{x+\epsilon}[. \tag{9.168}$$

Pour avoir $f\Big(B(x,\epsilon)\Big) \subset \mathcal{O}$, nous devons avoir

$$\begin{cases} \sqrt{x-\epsilon} > y - r & \text{(9.169a)} \\ \sqrt{x+\epsilon} < y + r. & \text{(9.169b)} \end{cases}$$

La première condition donne $x - \epsilon > (y-r)^2$. En développant le carré et en tenant compte de $y^2 = x$ nous demandons

$$\epsilon < 2ry - r^2. \tag{9.170}$$

Nos choix sont donc de demander $0 < r < 2y$, et ensuite $0 < \epsilon < 2ry - r^2$. Le premier est fait pour avoir $2ry - r^2 > 0$. □

9.6.2 Second degré

Nous résolvons à présent le polynôme du second degré.

PROPooEZIKooKjJroH

Proposition 9.117 ([317]).
Soit la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto ax^2+bx+c \end{aligned} \tag{9.171}$$

avec $a\neq 0$. Nous notons $\Delta=b^2-4ac$.

ITEMooMKUSooWwNTba

(1) Nous avons la formule

$$f(x)=a\big(x+\frac{b}{2a}\big)^2+c-\frac{b^2}{4a}. \tag{9.172}$$

EQooFKPOooAbIhCx

ITEMooHQTBooZuaPAs

(2) Si $a>0$, alors f a un minimum global en $x_m=-b/2a$.

ITEMooQMXVooWsqiXz

(3) Si $a<0$, alors f a un maximum global en $x_M=-b/2a$.

ITEMooMAMHooNWZVQI

(4) Si $\Delta<0$ alors f ne possède pas de racine réelle.

ITEMooKUUJooTsIHhI

(5) Si $\Delta=0$, alors f possède une unique racine $x_0=-b/2a$.

ITEMooQZGFooEGhMkX

(6) Si $\Delta>0$ alors f possède exactement deux racines distinctes données par

$$x_1=\frac{-b+\sqrt{\Delta}}{2a}\quad x_2=\frac{-b-\sqrt{\Delta}}{2a}. \tag{9.173}$$

EQooGHDPooVkoJNr

Démonstration. En plusieurs parties.

(1) **Pour (1)** C'est un calcul immédiat.

(2) **Pour (2)** Nous partons de la formule du point (1). Puisque $c-\frac{b^2}{4a}$ est constant, minimiser f revient à minimiser $x\mapsto\big(x+\frac{b}{2a}\big)^2$. Comme cette dernière fonction est toujours positive, elle a un minimum global là où elle est nulle, c'est-à-dire en $x_m=-b/2a$.

(3) **Pour (3)** Idem que pour (2).

Pour la suite nous effectuons quelques manipulations à partir de (9.172). Nous avons $f(x)=0$ lorsque

$$(x+\frac{b}{2a})^2=\frac{b^2-4ac}{4a^2}. \tag{9.174}$$

EQooRHNGooVsKBNt

(1) **Pour (4)** À gauche de (9.174) nous avons un nombre toujours positif ou nul. À droite, $4a^2>0$. Donc si $b^2-4ac<0$, l'égalité est impossible et il n'y a pas de x vérifiant $f(x)=0$.

(2) **Pour (5)** Si $b^2-4ac=0$, alors la condition (9.174) devient

$$\left(x+\frac{b}{2a}\right)^2=0, \tag{9.175}$$

et donc $x=-b/2a$ est l'unique solution.

(3) **Pour (6)** Si $b^2-4ac>0$, nous pouvons prendre la racine carré [52] des deux côtés de (9.174), et la condition devient

$$x+\frac{b}{2a}=\pm\sqrt{\frac{b^2-4ac}{4a^2}}, \tag{9.176}$$

ce qui donne

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}. \tag{9.177}$$

Ce sont là les deux seuls candidats pour vérifier $f(x)=0$.

52. Définition 9.97.

Un calcul direct montre que

$$f\left(\frac{-b+\sqrt{b^2-4ac}}{2a}\right)=0 \tag{9.178}$$

et que

$$f\left(\frac{-b-\sqrt{b^2-4ac}}{2a}\right)=0. \tag{9.179}$$

Donc ce sont bien des racines de f et ce sont les seules. Notez aussi qu'elles sont distinctes parce que $\Delta \neq 0$.

□

Chapitre 10

Espaces vectoriels normés

Plusieurs notions sur les espaces vectoriels normés (dont la définition 7.158) ont déjà été abordées dans la section 7.14. Voir aussi le thème 25.

10.0.1 Norme, produit scalaire et Cauchy-Schwarz (cas réel)

Dans la suite, le produit scalaire de x et y pourra être noté indifféremment par $x \cdot y$, $\langle x, y\rangle$ ou $b(x, y)$ lorsque une forme bilinéaire est donnée.

Nous rappelons au passage que les espaces vectoriels réels sont susceptibles de recevoir un produit scalaire, alors que les espaces vectoriels complexes sont susceptibles de recevoir un produit hermitien. Bien que de nombreux résultats soient identiques ou très similaires, ces deux notions sont à ne pas confondre.

Nous commençons par prouver qu'un produit scalaire étant donné, nous pouvons définir une norme par la formule $\|x\|^2 = \langle x, x\rangle$. Pour cela nous aurons besoin de l'inégalité de Cauchy-Schwarz.

ThoAYfEHG

Théorème 10.1 (Inégalité de Cauchy-Schwarz, cas réel)**.**
Soit un espace vectoriel réel E muni d'un produit scalaire[1] $(x, y) \mapsto x \cdot y$*. Nous posons*[2]

$$\|x\| = \sqrt{x \cdot x}. \tag{10.1}$$

Alors :

(1) Il y a l'inégalité

$$|x \cdot y| \leqslant \|x\|\|y\|. \tag{10.2}$$

EQooZDSHooWPcryG

pour tout $x, y \in E$.

(2) Il y a égalité $|x \cdot y| = \|x\|\|y\|$. si et seulement si x et y sont multiples l'un de l'autre.

(3) L'opération $\|.\|$ est une norme[3]*.*

(4) Cette norme vérifie l'identité du parallélogramme :

$$\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{10.3}$$

EqYCLtWfJ

Démonstration. Étant donné que les deux membres de l'inéquation sont positifs, nous allons travailler en passant au carré afin d'éviter les racines carrés dans le second membre.

Nous considérons l'application

$$\begin{aligned} P\colon \mathbb{R} &\to \mathbb{R}^+ \\ t &\mapsto \|x + ty\|, \end{aligned} \tag{10.4}$$

1. Produit scalaire, définition 8.168.
2. Attention à la notation : pour l'instant nous ne savons pas que c'est une norme ; c'est justement un des points de ce théorème. Par ailleurs, la racine carrée est définie par 9.97.
3. Définition 7.158.

et nous calculons un peu :

$$\begin{aligned}
0 &\leqslant \|x+ty\|^2 && (10.5a)\\
&= (x+ty)\cdot(x+ty) && (10.5b)\\
&= x\cdot x + x\cdot ty + ty\cdot x + t^2 y\cdot y && (10.5c)\\
&= \|y\|^2t^2 + 2(x\cdot y)t + \|x\|^2. && (10.5d)
\end{aligned}$$

Nous avons utilisé la bilinéarité (pour sortir les t) et la symétrique du produit scalaire.

Nous voyons que P est un polynôme du second degré en t à valeurs dans $[0,\infty[$. Par la proposition 9.117 nous en déduisons que le fameux b^2-4ac doit être négatif ou nul. Nous avons donc

$$\Delta = 4(x\cdot y)^2 - 4\|x\|^2\|y\|^2 \leqslant 0, \tag{10.6}$$

ce qui donne immédiatement

$$(x\cdot y)^2 \leqslant \|x\|^2\|y\|^2. \tag{10.7}$$

En ce qui concerne le cas d'égalité, si nous avons $x\cdot y = \|x\|\|y\|$, alors le discriminant Δ ci-dessus est nul et le polynôme P admet une racine double t_0. Pour cette valeur nous avons

$$P(t_0) = |x+t_0y| = 0, \tag{10.8}$$

ce qui implique $x+t_0y=0$ et donc que x et y sont liés.

(1) **C'est une norme** Nous allons nous contenter de prouver l'inégalité triangulaire. Si $x,y\in E$, nous avons

$$\begin{aligned}
\|x+y\| &= \sqrt{\|x\|^2+\|y\|^2+2x\cdot y} && (10.9a)\\
&\leqslant \sqrt{\|x\|^2+\|y\|^2+2|x\cdot y|} && (10.9b)\\
&\leqslant \sqrt{\|x\|^2+\|y\|^2+2\|x\|\|y\|} && (10.9c)\\
&= \sqrt{\big(\|x\|+\|y\|\big)^2} && (10.9d)\\
&= \|x\|+\|y\|. && (10.9e)
\end{aligned}$$

SUBEQooRISGooRUoBzj SUBEQooPQPYooBRRBzK

Justifications.

— Pour (10.9b). La fonction racine carrée est croissante, lemme 9.99.

— Pour (10.9c). Inégalité de Cauchy-Schwarz 10.1.

(2) **Inégalité du parallélogramme** Cette assertion est seulement un calcul :

$$\begin{aligned}
\|x-y\|^2+\|x+y\|^2 &= (x-y)\cdot(x-y)+(x+y)\cdot(x+y)\\
&= x\cdot x - x\cdot y - y\cdot x + y\cdot y\\
&\quad + x\cdot x + x\cdot y + y\cdot x + y\cdot y\\
&= 2x\cdot x + 2y\cdot y\\
&= 2\|x\|^2+2\|y\|^2.
\end{aligned} \tag{10.10}$$

□

Toute norme dérivant d'un produit scalaire vérifie l'identité du parallélogramme. Ce résultat sert souvent à prouver que des normes ne dérivent pas d'un produit scalaire. C'est le cas de la norme $N(x,y)=|x|+|y|$ du lemme 10.14 ainsi que du théorème de Weinersmith 27.46.

PROPooVSVMooZrqxdc

Proposition 10.2.
La norme euclidienne[4] *a les propriétés suivantes :*

4. Un espace euclidien est un espace vectoriel muni d'un produit scalaire, définition 8.172. Ici nous considérons la norme associée par le théorème 10.1.

(1) *Pour tout vecteur x et réel λ,*

$$\|\lambda x\| = |\lambda|\|x\|. \tag{10.11}$$

(2) *Pour tout vecteurs x et y,*

$$\|x + y\| \leqslant \|x\| + \|y\| \tag{10.12}$$

Démonstration. Un produit scalaire est en particulier une forme bilinéaire, et vérifie les conditions de la définition 8.124.

Pour le premier point nous avons

$$\|\lambda x\| = \sqrt{(\lambda x) \cdot (\lambda x)} = \sqrt{\lambda^2 x \cdot x} = |\lambda|\|x\|. \tag{10.13}$$

Nous avons utilisé la formule du lemme 9.98(2).

Pour le second point, nous avons les inégalités suivantes :

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2x \cdot y && \text{(10.14a)}\\ &\leqslant \|x\|^2 + \|y\|^2 + 2|x \cdot y| && \text{(10.14b)}\\ &\leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = \big(\|x\| + \|y\|\big)^2 && \text{(10.14c)} \end{aligned}$$

Nous avons utilisé d'abord la majoration $|x| \geqslant x$ qui est évidente pour tout nombre x ; et ensuite l'inégalité de Cauchy-Schwarz 10.1. □

DEFooJAGXooMgaUsR

Définition 10.3 (Norme sur $\mathbb{R}^n$)**.**
Sauf mention du contraire, nous considérons toujours sur $\mathbb{R}^n$ la norme (et donc la topologie) associée au produit scalaire de la proposition 8.173 par le théorème 10.1, c'est à dire

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}. \tag{10.15}$$

PropHIWjdMX

Proposition 10.4 ([318])**.**
Soit b une forme bilinéaire et symétrique. Alors

(1) $\ker(b) \subset C(b)$ *(cône d'isotropie, définition 8.143)*

(2) *si b est positive alors* $\ker(b) = C(b)$.

Démonstration. En deux parties.

(1) Si $z \in \ker(b)$ alors pour tout $y \in E$ nous avons $b(z, y) = 0$. En particulier pour $y = z$ nous avons $b(z, z) = 0$ et donc $z \in C(b)$.

(2) Soit b positive et $x \in C(b)$. Par l'inégalité de Cauchy-Schwarz (proposition 10.1) nous avons

$$|b(x, y)| \leqslant \sqrt{b(x, x)b(y, y)} = 0. \tag{10.16}$$

Donc pour tout y nous avons $b(x, y) = 0$.

□

LEMooEZFIooXyYybe

Lemme 10.5.
Soit un espace vectoriel euclidien[5] *E. Si $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de E et si $f : E \to E$ est un endomorphisme, alors*

$$\det(f) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \langle e_{\sigma(i)}, f(e_i)\rangle. \tag{10.17}$$

EQooQAZLooZutFUz

5. C'est-à-dire qu'il possède un produit scalaire, voir la définition 8.172.

Démonstration. Nous utilisons la définition 8.15 du déterminant d'un endomorphisme $\det(f) = \det_B\big(f(B)\big)$ en prenant la liste des vecteurs $\{e_i\}$ comme B. En l'occurrence, le i^{e} vecteur de la famille $f(B)$ est $f(e_i)$.

Puisque la base est orthonormée, nous avons $e_k^*(v) = \langle e_k, v\rangle$ et donc aussi

$$e^*_{\sigma(i)}(v_i) = \langle e^*_{\sigma(i)}, f(e_i)\rangle. \tag{10.18}$$

□

Et si vous avez tout suivi, vous aurez remarqué que les produits scalaires impliqués dans la formule (10.17) sont les éléments de la matrice de f dans la base $\{e_i\}$ parce que $\langle e_i, f(e_j)\rangle$ est la composante i de l'image de e_j par f. Si la matrice est composée en mettant en colonne les images des vecteurs de base, le compte est bon.

10.0.2 Topologie

PROPooEKTIooEKmWfn

Proposition-Définition 10.6 ([1])**.**
Si $(E, \|.\|)$ est un espace vectoriel normé, alors

$$\begin{aligned} d\colon E \times E &\to \mathbb{R} \\ (x, y) &\mapsto \|x - y\| \end{aligned} \tag{10.19}$$

est une distance[6]*.*

La ***topologie*** *associée à la norme $\|.\|$ est la topologie métrique de d*[7]*.*

PROPooWIGWooXPyDjp

Proposition 10.7 ([1])**.**
Soit un espace vectoriel normé $(E, \|.\|)$ de dimension finie. Soit un sous-espace vectoriel $F \subset E$. Nous posons

$$\begin{aligned} n\colon F &\to \mathbb{R} \\ x &\mapsto \|x\|. \end{aligned} \tag{10.20}$$

Alors

(1) (F, n) est un espace vectoriel normé.

(2) La topologie métrique[8] *de (F, n) est la topologie induite de E.*

Démonstration. L'application n est une norme parce que toutes les propriétés de la définition 7.158 passent directement de $\|.\|$ à n. Pour le reste, remarquons d'abord que

$$B_F(x, r) = B_E(x, r) \cap F. \tag{10.21}$$

Nous notons τ_i la topologie induite de E vers F, et τ_n la topologie sur F pour la norme n. Nous allons prouver que $\tau_i = \tau_n$.

(1) $\underline{\tau_i \subset \tau_n}$ Soit $A \in \tau_i$. Pour prouver que $A \in \tau_n$, nous prenons $x \in A$ et nous prouvons qu'il existe $r > 0$ tel que $B_F(x, r) \subset A$. Par définition de la topologie induite, il existe un ouvert A' de E tel que $A = F \cap A'$. Vu que A' est ouvert pour la topologie métrique, il existe $r > 0$ tel que $B_E(x, r) \subset A'$. Pour ce r nous avons

$$B_F(x, r) = B_E(x, r) \cap F \subset A' \cap F = A. \tag{10.22}$$

(2) $\underline{\tau_n \subset \tau_i}$ Soient $A \in \tau_n$ et $x \in A$. Il existe $r > 0$ tel que $x \in B_F(x, r) \subset A$, et donc aussi

$$x \in B_E(x, r) \cap F \in \tau_i. \tag{10.23}$$

□

6. Définition 7.115.
7. Définition 7.117.
8. Voir la définition 7.117.

10.1 Théorème spectral autoadjoint

ThoRSBahHH

Théorème 10.8 (Théorème spectral autoadjoint)**.**
Un endomorphisme autoadjoint d'un espace euclidien

(1) est diagonalisable dans une base orthonormée,

(2) a son spectre réel.

Démonstration. Nous procédons par récurrence sur la dimension de E, et nous commençons par $n = 1$[9]. Soit donc $f : E \to E$ avec $\langle f(x), y\rangle = \langle x, f(y)\rangle$. Étant donné que f est également linéaire, il existe $\lambda \in \mathbb{R}$ tel que $f(x) = \lambda x$ pour tout $x \in E$. Tous les vecteurs de E sont donc vecteurs propres de f.

Passons à la récurrence. Nous considérons $\dim(E) = n + 1$ et $f \in \mathrm{S}(E)$. Nous considérons la forme bilinéaire symétrique Φ_f et la forme quadratique associée ϕ_f. Pour rappel,

$$\Phi_f(x, y) = \langle x, f(y)\rangle \tag{10.24a}$$
$$\phi_f(x) = \Phi_f(x, x). \tag{10.24b}$$

Et nous allons laisser tomber les indices f pour noter simplement Φ et ϕ. Étant donné que $\overline{B(0,1)}$ est compacte et que ϕ est continue, il existe $x_0 \in \overline{B(0,1)}$ tel que

$$\lambda = \phi(x_0) = \sup_{x \in \overline{B(0,1)}} \phi(x). \tag{10.25}$$

Notons aussi que $\|x_0\| = 1$: le maximum est pris sur le bord. Nous posons

$$g = \lambda \mathrm{Id} - f \tag{10.26}$$

ainsi que

$$\Phi_1(x, y) = \langle x, g(y)\rangle. \tag{10.27}$$

C'est une forme bilinéaire et symétrique parce que

$$\Phi_1(y, x) = \langle y, g(x)\rangle = \langle g(y), x\rangle = \langle x, g(y)\rangle = \Phi_1(x, y) \tag{10.28}$$

où nous avons utilisé le fait que g était autoadjoint et la symétrie du produit scalaire. De plus Φ_1 est semi-définie positive parce que

$$\Phi_1(x, x) = \langle x, \lambda x - f(x)\rangle = \lambda\|x\|^2 - \phi(x). \tag{10.29}$$

Puisque λ est le maximum, nous avons tout de suite $\Phi_1(x) \geqslant 0$ tant que $\|x\| = 1$. Et si x n'est pas de norme 1, c'est le même prix parce qu'on se ramène à $\|x\| = 1$ en multipliant par un nombre positif. Attention cependant :

$$\Phi_1(x_0, x_0) = \lambda\|x_0\|^2 - \phi(x_0) = 0. \tag{10.30}$$

Donc Φ_1 a un noyau contenant x_0 par la proposition 10.4. Nous en déduisons que $\mathrm{Image}(g) \neq E$ en effet, $x_0 \in \mathrm{Image}(g)^\perp$, mais nous avons la proposition 4.134 sur les dimensions :

$$\dim E = \dim(\mathrm{Image}(g)) + \dim(\mathrm{Image}(g)^\perp). \tag{10.31}$$

Comme $\mathrm{Image}(g)^\perp$ est un espace vectoriel non réduit à $\{0\}$, la dimension de $\mathrm{Image}(g)$ ne peut pas être celle de E. L'endomorphisme g n'étant pas surjectif, il ne peut pas être injectif non plus parce que nous sommes en dimension finie ; il existe donc $e_1 \in E$ tel que $g(e_1) = 0$ et tant qu'à faire nous choisissons $\|e_1\| = 1$ (ici la norme est bien celle de l'espace euclidien considéré). Par définition,

$$f(e_1) = \lambda e_1, \tag{10.32}$$

9. Dans [104], l'auteur commence avec $n = 0$ mais moi je n'en ai pas le courage..

c'est-à-dire que $\lambda \in \mathrm{Spec}(f)$. Et ϕ étant une forme quadratique réelle nous avons $\lambda \in \mathbb{R}$.

Nous posons à présent $H = \mathrm{Span}\{e_1\}^\perp$. C'est un sous-espace stable par f parce que si $x \in H$ alors

$$\langle e_1, f(x)\rangle = \langle f(e_1), x\rangle = \lambda\langle e_1, x\rangle = 0. \tag{10.33}$$

Nous pouvons donc considérer la restriction de f à H : $f_H\colon H \to H$. Cet endomorphisme est bilinéaire et symétrique sur l'espace H de dimension inférieure à celle de E, donc la récurrence nous donne une base orthonormée

$$\{e_2, \dots, e_n\} \tag{10.34}$$

de vecteurs propres de f_H. De plus les valeurs propres sont réelles, toujours par récurrence. Donc

$$\mathrm{Spec}(f) = \{\lambda\} \cup \mathrm{Spec}(f_H) \subset \mathbb{R}. \tag{10.35}$$

Notons pour être complet que si $i \geqslant 2$ alors

$$\langle e_1, e_i\rangle = 0 \tag{10.36}$$

parce que le vecteur e_i est par construction choisi dans l'espace $H = e_1^\perp$. Nous avons donc bien une base orthonormée de E construite sur des vecteurs propres de f. □

CorSMHpoVK

Corolaire 10.9.
Soit E un espace vectoriel ainsi que des formes bilinéaires ϕ et ψ sur E. Si ψ est définie positive, alors il existe une base ψ-orthonormale dans laquelle ϕ est diagonale.

Démonstration. Il suffit de considérer l'espace euclidien E muni du produit scalaire $\langle x, y\rangle = \psi(x, y)$. Ensuite nous diagonalisons la matrice (symétrique) de ϕ pour ce produit scalaire à l'aide du théorème 10.8. □

DefNormeEucleApp

Définition 10.10.
*La **norme euclidienne** d'un élément de $\mathbb{R}^m$ est définie par $\|u\| = \sqrt{u \cdot u}$*[10].

Cette définition est motivée par le fait que le produit scalaire $u \cdot u$ donne exactement la norme usuelle donnée par le théorème de Pythagore :

$$u \cdot u = \sum_{i=1}^{m} u_i u_i = \sum_{i=1}^{m} u_i^2 = u_1^2 + u_2^2 + \cdots + u_m^2. \tag{10.37}$$

Le fait que $e_i \cdot e_j = \delta_{ij}$ signifie que la base canonique est **orthonormée**, c'est-à-dire que les vecteurs de la base canonique sont orthogonaux deux à deux et qu'ils ont tout 1 comme norme.

LemSclNormeXi

Lemme 10.11.
Pour tout $u \in \mathbb{R}^m$, il existe un $\xi \in \mathbb{R}^m$ tel que $\|u\| = \xi \cdot u$ et $\|\xi\| = 1$.

Démonstration. Vérifions que le vecteur $\xi = u/\|u\|$ ait les propriétés requises. D'abord $\|\xi\| = 1$ parce que $u \cdot u = \|u\|^2$. Ensuite

$$\xi \cdot u = \frac{u \cdot u}{\|u\|} = \frac{\|u\|^2}{\|u\|} = \|u\|. \tag{10.38}$$

□

10. La racine carrée est définie en 9.97.

10.1.1 Inégalité de Minkowski

Ce qui est couramment nommé « inégalité de Minkowski » est la proposition 27.39 dans les espaces L^p. Nous allons en donner ici un cas très particulier.

PropACHooLtsMUL

Proposition 10.12.
Si q est une forme quadratique[11] *sur $\mathbb{R}^n$ et si $x, y \in \mathbb{R}^n$ alors*

$$\sqrt{q(x+y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}. \tag{10.39}$$

Démonstration. La proposition 8.255 nous permet de « diagonaliser » la forme quadratique q. Quitte à ne plus avoir une base orthonormale, nous pouvons renormaliser les vecteurs de base pour avoir

$$q(x) = \sum_i x_i^2. \tag{10.40}$$

Le résultat n'est donc rien d'autre que l'inégalité triangulaire pour la norme euclidienne usuelle, laquelle est démontrée dans le théorème 10.1. □

10.13.
Un produit scalaire fournit donc toujours une norme et donc une topologie. Il ne faudrait cependant pas croire que toute norme dérive d'un produit scalaire, même pas en dimension finie. Et ce, malgré l'équivalence de toutes les normes du théorème 10.49 dont vous avez déjà peut-être entendu parler.

L'intérêt du lemme suivant sera apparent en 10.51.

LEMooRWJYooOIJkZc

Lemme 10.14.
Sur $\mathbb{R}^2$, l'application $N(x,y) = |x| + |y|$ est une norme[12] *qui ne dérive pas d'un produit scalaire*[13].

Démonstration. Nous commençons par montrer que N est une norme. Il faut vérifier les trois conditions de la définition 7.158.

(1) Il faut utiliser le lemme 1.383(1) dans les deux sens. Si $(x,y) = (0,0)$, alors évidemment $N(x,y) = 0$. Dans l'autre sens, si $N(x,y) = 0$ nous avons

$$0 = |x| + |y| \geqslant |x|. \tag{10.41}$$

Donc $|x| \leqslant 0$, mais comme $|x| \geqslant 0$, nous avons $|x| = 0$ et donc $x = 0$. Le même raisonnement tient pour y.

(2) En tenant compte du fait que $|\lambda x| = |\lambda||x|$, nous avons

$$N\big(\lambda(x,y)\big) = N(\lambda x, \lambda y) = |\lambda||x| + |\lambda||y| = |\lambda|(|x| + |y|) = |\lambda| N(x,y). \tag{10.42}$$

(3) Nous avons le calcul

$$\begin{aligned} N\big((x,y) + (a,b)\big) &= N(x+a, y+b) && \text{(10.43a)} \\ &= |x+a| + |y+b| && \text{(10.43b)} \\ &\leqslant |x| + |a| + |y| + |b| = N(x,y) + N(a,b) && \text{(10.43c)} \end{aligned}$$

SUBEQooIXKWooTNQFnu

Justification : pour (10.43c) nous avons utilisé $|a+b| \leqslant |a| + |b|$, du lemme 1.383.

Pour voir qu'elle ne dérive pas d'un produit scalaire, nous montrons qu'elle ne vérifie pas l'identité du parallélogramme du théorème 10.1.

Voici un petit bout de code qui nous permet de ne pas faire de recherches à la main :

11. Définition 8.125.
12. Définition 7.158.
13. La norme d'un produit scalaire est le théorème 10.1.

```sage
# Dans un cas réel, vous avez nettement intérêt à
# créer une classe 'Vecteur' qui implémente somme, différence
# et norme.
def N(v):
    return abs( v[0] )+abs(v[1])

def parall(v,w):
    # La différence v-w
    d=( v[0]-w[0],v[1]-w[1] )
    # La somme v+w
    s=( v[0]+w[0],v[1]+w[1] )

    return N(d)**2+N(s)**2-2*N(v)**2-2*N(w)**2
```

tex/sage/sageSnip018.sage

Il est vite vu qu'avec $v = (-1, 1)$ et $w = (1, 1)$, l'identité du parallélogramme n'est pas vérifiée. □

LemLPOHUme

Lemme 10.15 ([104]).
Soit V un espace vectoriel muni d'un produit scalaire[14] *et de la norme associée. Si $x, y \in V$ satisfont à $\|x + y\| = \|x\| + \|y\|$, alors il existe $\lambda \geqslant 0$ tel que $x = \lambda y$.*

Démonstration. Quitte à raisonner avec $x/\|x\|$ et $y/\|y\|$, nous supposons que $\|x\| = \|y\| = 1$. Dans ce cas l'hypothèse signifie que $\|x+y\|^2 = 4$. D'autre part en écrivant la norme en termes de produit scalaire,

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle, \tag{10.44}$$

ce qui nous mène à affirmer que $\langle x, y\rangle = 1 = \|x\|\|y\|$. Nous sommes donc dans le cas d'égalité de l'inégalité de Cauchy-Schwarz[15], ce qui nous donne un λ tel que $x = \lambda y$. Étant donné que $\|x\| = \|y\| = 1$ nous avons obligatoirement $\lambda = \pm 1$, mais si $\lambda = -1$ alors $\langle x, y\rangle = -1$, ce qui est le contraire de ce qu'on a prétendu plus haut. Par souci de cohérence, nous allons donc croire que $\lambda = 1$. □

PropVectsOrthLibres

Proposition 10.16 ([319]).
si $v_1, \cdots, v_k$ sont des vecteurs non nuls, orthogonaux deux à deux, alors ces vecteurs forment une famille libre.

Démonstration. Soit une combinaison linéaire nulle des v_i : $\sum_{i=1}^{k} \lambda_i v_i = 0$. Nous multiplions scalairement par v_k :

$$0 = \sum_i \lambda_i v_k \cdot v_i = \sum_i \lambda_i \delta_{ki} \|v_i\|^2 = \lambda_k \|v_k\|^2. \tag{10.45}$$

Donc $\lambda_k = 0$. □

LEMooYXJZooWKRFRu

Lemme 10.17.
Une isométrie d'un espace euclidien fixe l'origine.

Démonstration. Soit une isométrie f d'un espace euclidien : $f(x) \cdot f(y) = x \cdot y$ pour tout $x, y \in E$. En particulier pour $x = 0$ nous avons

$$f(0) \cdot f(y) = 0 \tag{10.46}$$

pour tout y. Parce que f est une bijection, nous avons $f(0) \cdot x = 0$ pour tout x. Comme le produit scalaire est non dégénéré[16] cela implique que $f(0) = 0$. □

14. Définition 8.168.
15. Théorème 10.1.
16. Lemme 8.171.

10.1.2 Cauchy-Schwarz etc. cas complexe

THOooSUCBooFnpkaF

Théorème 10.18 (Inégalité de Cauchy-Schwarz, cas complexe[320])**.**
Soit un espace vectoriel complexe muni d'un produit hermitien $\langle .,. \rangle$. Alors pour tous vecteurs x, y nous avons

$$|\langle x, y\rangle| \leqslant \|x\|\|y\| \tag{10.47}$$

où nous avons posé $\|x\| = \sqrt{\langle x, x\rangle}$.

Démonstration. Si $\langle x, y\rangle = 0$, le résultat est évident. Nous nous concentrons donc sur le cas où $\langle x, y\rangle \neq 0$. Nous posons

$$\theta = \frac{\langle x, y\rangle}{|\langle x, y\rangle|}. \tag{10.48}$$

C'est un élément de $\mathbb{C}$ de norme 1. Nous avons

$$\langle \frac{1}{\theta}x, y\rangle = \frac{|\langle x, y\rangle|}{\langle x, y\rangle}\langle x, y\rangle = |\langle x, y\rangle| \geqslant 0 \tag{10.49}$$

où le symbole « $\geqslant$ » signifie « est réel et positif ». Nous posons $x' = \frac{1}{\theta}x$ et nous considérons $t \in \mathbb{R}$. Remarquons que $\|x'\|^2 = \|x\|^2$:

$$\|x'\|^2 = \langle x', x'\rangle = \frac{1}{\theta\bar{\theta}}\langle x, x\rangle = \|x\|^2 \tag{10.50}$$

parce que $|\theta| = 1$.

En utilisant le fait que $\langle a, b\rangle + \langle b, a\rangle = \mathrm{Re}(\langle a, b\rangle)$ nous avons :

$$0 \leqslant \|x' + ty\|^2 = \|x'\|^2 + t\langle x', y\rangle + t\langle y, x'\rangle + t^2\|y\|^2 \tag{10.51a}$$

$$= \|y\|^2t^2 + 2\,\mathrm{Re}(\langle x', y\rangle)t + \|x'\|^2. \tag{10.51b}$$

C'est un polynôme de degré 2 en t qui n'est jamais strictement négatif. Autrement dit, il a au maximum une seule racine, ce qui signifie que son discriminant est négatif ou nul :

$$\mathrm{Re}(\langle x', y\rangle)^2 - \|y\|^2\|x'\|^2 \leqslant 0. \tag{10.52}$$

Mais nous avons choisi x' de telle sorte que $\langle x', y\rangle = |\langle x, y\rangle| \in \mathbb{R}$ et $\|x'\|^2 = \|x\|^2$; nous avons donc

$$|\langle x, y\rangle|^2 \leqslant \|x\|^2\|y\|^2, \tag{10.53}$$

comme il se devait. □

PROPooSSYJooHAXAnC

Proposition 10.19 (Identité du parallélogramme[321])**.**
Soit un espace vectoriel complexe E muni d'un produit hermitien $\langle .,. \rangle$. Nous posons $\|x\| = \sqrt{\langle x, x\rangle}$. Nous avons

(1) $\|.\|$ est une norme.

(2) Elle vérifie l'identité du parallélogramme :

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{10.54}$$

pour tout $x, y \in E$.

Démonstration. En ce qui concerne le fait que $\|.\|$ soit une norme, tout est essentiellement dans la définition 8.176 d'un produit hermitien. Voyons tout de même l'inégalité triangulaire. Nous avons :

$$\|x + y\|^2 = \langle x + y, x + y\rangle \tag{10.55a}$$

$$= \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle \tag{10.55b}$$

$$= \|x\|^2 + \|y\|^2 + 2\Re\big(\langle x, y\rangle\big) \tag{10.55c}$$

$$\leqslant \|x\|^2 + \|y\|^2 + 2|\Re\big(\langle x, y\rangle\big)| \tag{10.55d}$$

$$\leqslant \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \tag{10.55e}$$

$$\leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \tag{10.55f}$$

SUBEQooQGQBooMRIcUc

$$= \big(\|x\| + \|y\|\big)^2. \tag{10.55g}$$

Pour (10.55f) nous avons utilisé Cauchy-Schwarz 10.18. □

10.1.3 Diagonalisation : cas complexe, ce qu'on a

LEMooVCEOooIXnTpp

Lemme 10.20 (Théorème spectral hermitien)**.**
Nous considérons un espace vectoriel complexe hermitien. Pour un opérateur hermitien[17]*,*

(1) le spectre est réel,

(2) deux vecteurs propres pour des valeurs propres distinctes sont orthogonaux[18]*.*

Démonstration. Soit v un vecteur de valeur propre λ. Nous avons d'une part

$$\langle Av, v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2, \tag{10.56}$$

et d'autre part, en utilisant le fait que A est hermitien,

$$\langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \bar{\lambda}\|v\|^2, \tag{10.57}$$

par conséquent $\lambda = \bar{\lambda}$ parce que $v \neq 0$.

Soient λ_i et v_i $(i = 1, 2)$ deux valeurs propres de A avec leurs vecteurs propres correspondants. Alors d'une part

$$\langle Av_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle, \tag{10.58}$$

et d'autre part

$$\langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle = \lambda_2\langle v_1, v_2\rangle. \tag{10.59}$$

Nous avons utilisé le fait que λ_2 était réel. Par conséquent, soit $\lambda_1 = \lambda_2$, soit $\langle v_1, v_2\rangle = 0$. □

REMooMLBCooTuKFmz

Remarque 10.21.
Un opérateur de la forme A^*A est évidemment hermitien. De plus ses valeurs propres sont toutes positives parce que si $A^*Av = \lambda v$ alors

$$0 \leqslant \langle Av, Av\rangle = \langle A^*Av, v\rangle = \lambda\langle v, v\rangle. \tag{10.60}$$

Donc $\lambda \geqslant 0$.

10.1.4 Projection et orthogonalité

La définition du produit scalaire dans $\mathbb{R}^n$ est 8.173 et le lien avec la matrice d'une application linéaire est la proposition 8.184.

Remarque 10.22.
Outre l'orthogonalité, le produit scalaire permet de connaître l'angle entre deux vecteurs à travers la définition 18.52. D'autres interprétations géométriques du déterminant sont listées dans le thème 45.

Nous sommes maintenant en mesure de déterminer, pour deux vecteurs quelconques u et v, la projection orthogonale de u sur v. Ce sera le vecteur $\bar{u}$ parallèle à v tel que $u - \bar{u}$ est orthogonal à v. Nous avons donc

$$\bar{u} = \lambda v \tag{10.61}$$

et

$$(u - \lambda v) \cdot v = 0. \tag{10.62}$$

La seconde équation donne $u \cdot v - \lambda v \cdot v = 0$, ce qui fournit λ en fonction de u et v :

$$\lambda = \frac{u \cdot v}{\|v\|^2}. \tag{10.63}$$

17. Définition 8.179.
18. Pour la forme (8.355).

Nous avons par conséquent

$$\bar{u} = \frac{u \cdot v}{\|v\|^2} v. \tag{10.64}$$

Armés de cette interprétation graphique du produit scalaire, nous comprenons pourquoi nous disons que deux vecteurs sont orthogonaux lorsque leur produit scalaire est nul.

Nous pouvons maintenant savoir quel est le coefficient directeur d'une droite orthogonale à une droite donnée. En effet, supposons que la première droite soit parallèle au vecteur X et la seconde au vecteur Y. Les droites seront perpendiculaires si $X \cdot Y = 0$, c'est-à-dire si

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0. \tag{10.65}$$

Cette équation se développe en

$$x_1 y_1 = -x_2 y_2. \tag{10.66}$$ Eqxuyuklisca

Le coefficient directeur de la première droite est $\frac{x_2}{x_1}$. Isolons cette quantité dans l'équation (10.66) :

$$\frac{x_2}{x_1} = -\frac{y_1}{y_2}. \tag{10.67}$$

Donc le coefficient directeur de la première est l'inverse et l'opposé du coefficient directeur de la seconde.

Exemple 10.23.
Soit la droite $d \equiv y = 2x + 3$. Le coefficient directeur de cette droite est 2. Donc le coefficient directeur d'une droite perpendiculaires doit être $-\frac{1}{2}$. △

PROBooLOEKooZxNpJZ

¡¡ Avertissement/question au lecteur !! 10.24
La « preuve alternative » qui suit . . .je ne sais plus du tout à quoi elle faisait référence. Si vous avez une idée de l'énoncé, faites-le moi savoir.

Sinon, vous pouvez certainement ne pas la lire.

Preuve alternative. La preuve peut également être donnée en ne faisant pas référence au produit scalaire. Il suffit d'écrire toutes les quantités en termes des coordonnées de X et Y. Si nous posons

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \tag{10.68}$$

l'inégalité à prouver devient

$$(x_1 y_1 + x_2 y_2 + x_3 y_3)^2 \leqslant (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2). \tag{10.69}$$

Nous considérons la fonction

$$\varphi(t) = (x_1 + t y_1)^2 + (x_2 + t y_2)^2 + (x_3 + t y_3)^2 \tag{10.70}$$

En tant que norme, cette fonction est évidemment positive pour tout t. En regroupant les termes de chaque puissance de t, nous avons

$$\varphi(t) = (y_1^2 + y_2^2 + y_3^2)t^2 + 2(x_1 y_1 + x_2 y_2 + x_3 y_3)t + (x_1^2 + x_2^2 + x_3^2). \tag{10.71}$$

C'est un polynôme du second degré en t. Par conséquent le discriminant doit être négatif[19]. Nous avons donc

$$4(x_1 y_1 + x_2 y_2 + x_3 y_3)^2 - (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) \leqslant 0. \tag{10.72}$$

La thèse en découle aussitôt. □

19. Proposition 9.117.

10.1.5 Théorème de Pythagore

THOooHXHWooCpcDan

Théorème 10.25 (Pythagore[1]).
Soient un espace euclidien[20] *E ainsi que trois points $a, b, c \in E$ formant un triangle rectangle en a, c'est-à-dire tel que*

$$(b-a) \cdot (a-c) = 0 \tag{10.73}$$

EQooRAWAooBxlBsZ

Alors

$$\|b-c\|^2 = \|b-a\|^2 + \|a-c\|^2. \tag{10.74}$$

Démonstration. D'abord pour développons l'hypothèse (10.73) :

$$b \cdot a - b \cdot c - \|a\|^2 + a \cdot c = 0, \tag{10.75}$$

et nous isolons un bout qui va nous servir plus tard :

$$b \cdot a + a \cdot c = b \cdot c + \|a\|^2. \tag{10.76}$$

EQooWPWZooLjlVJk

Maintenant nous calculons un peu :

$$\begin{aligned}
\|b-a\|^2 + \|a-c\|^2 &= \|b\|^2 - 2b \cdot a + \|a\|^2 + \|a\|^2 - 2a \cdot c + \|c\|^2 && \text{(10.77a)}\\
&= 2\|a\|^2 + \|b\|^2 + \|c\|^2 - 2(b \cdot c + \|a\|^2) && \text{(10.77b)}\\
&= \|b\|^2 + \|c\|^2 - 2b \cdot c && \text{(10.77c)}\\
&= \|b-c\|^2. && \text{(10.77d)}
\end{aligned}$$

SUBEQooHCWXooQHpGTQ

Pour (10.77b), nous avons substitué (10.76). □

10.26.
Je profite de l'occasion pour montrer mon scepticisme quant aux preuves de Pythagore basées sur différents pliages et découpages des carrés construits sur les côtés du triangle.

Si, comme ici, nous considérons la géométrie dans $\mathbb{R}^2$ muni de son produit scalaire, alors le théorème 10.25 est le théorème de Pythagore et il n'est pas loin d'être la définition de la distance entre deux points. Ce serait exactement la définition pour le triangle $A = (0,0)$, $B = (a,0)$, $C = (a,b)$.

Pour autant que je le sache, la géométrie dans « le plan » (celle du collège) ne définit pas « longueur » et « aire ». Donc bon . . .Il y a peut-être un moyen de s'en sortir, mais je ne le connais pas.

Bref, soit on se met d'accord sur les définition (et dans ce cas je serais étonné qu'il existe une démonstration de Pythagore très différente de ce qu'on a ici), soit il faudrait se calmer avec les soi-disant preuves du théorème de Pythagore.

LEMooKDGRooMmJqnn

Lemme 10.27.
Soit un carré de côté c. Alors deux points du carré sont au maximum éloignés d'une distance égale à $c\sqrt{2}$.

10.1.6 Produit vectoriel

DEFooTNTNooRjhuJZ

Définition 10.28.
*Soient u et v, deux vecteurs de $\mathbb{R}^3$. Le **produit vectoriel** de u et v est le vecteur $u \times v$ défini par*

$$u \times v = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} \tag{10.78}$$

EQooCUJRooFuFPaZ

où les vecteurs e_1, e_2 et e_3 sont les vecteurs de la base canonique de $\mathbb{R}^3$.

20. Définition 8.172.

LEMooYKNPooGVMuse

Lemme 10.29.
Le produit vectoriel $u \times v$ est également exprimé par EQSooOWGZooNYruoy

$$u \times v = (u_2v_3 - u_3v_2)e_1 + (u_3v_1 - u_1v_3)e_2 + (u_1v_2 - u_2v_1)e_3 \quad \text{(10.79a)}$$ SEBEQooVROKooRpUOIr

$$= \sum_{i,j,k} \epsilon_{ijk} v_i w_j e_k \quad \text{(10.79b)}$$

où ϵ_{ijk} est défini par $\epsilon_{xyz} = 1$ et ensuite ϵ_{ijk} est 1 ou -1 suivant que la permutation des x, y et z est paire ou impaire. C'est-à-dire que ϵ_{ijk} est la signature de la permutation qui amène $(1, 2, 3)$ sur (i, j, k).

Démonstration. Il s'agit seulement de développer explicitement le déterminant (10.78). □

NORMooARQCooCAapiy

10.30.
Admettons que $a \times b = v$. En calculant le même produit vectoriel dans la base $f_i = -e_i$, les composantes de a et b changent de signe et la formule (10.79) dit que le produit vectoriel ne change pas. On serait tenté d'écrire, dans la base $\{f_i\}$

$$(-a) \times (-b) = v, \quad \text{(10.80)}$$

tout en pleurant parce que dans la base des f_i, le vecteur v devient $-v$.

Il y a des personnes que cela tracasse tellement qu'on entend parler de « le produit vectoriel est un pseudo-vecteur sous SO(2) ». Les physiciens en théorie quantique des champs sont terribles sur ce sujet [21].

Il suffit d'être clair. Le produit vectoriel n'est défini que sur $\mathbb{R}^3$, et est défini par sa formule dans la base canonique, point barre. Si vous avez des vecteurs a et b dont vous connaissez les composantes dans une autre base, vous devez calculer les composantes dans la base canonique, utiliser la formule pour trouver les composantes de $a \times b$ dans la base canonique. Ensuite, si ça vous chante, vous pouvez calculer à nouveau les composantes de $a \times b$ dans une autre base.

Tout cela pour dire que le produit vectoriel n'est pas une opération très généralisable. Il est possible, pour sembler plus intrinsèque, de tenter cette définition : le produit vectoriel $a \times b$ est le vecteur perpendiculaire à a et b, de longueur égale à l'aire du parallélogramme construit sur a et b.

Cette « définition » a plusieurs inconvénients.

— Elle demande quand même un produit scalaire et des aires ; bref, elle demande une structure métrique,

— Elle ne donne pas le sens. En effet, dans $\mathbb{R}^3$, il y a deux vecteurs de longueur donnée perpendiculaires à a et b. Il faut donc préciser le sens. Cela revient à donner une orientation et donc, fondamentalement, à choisir une base.

Bref, on retiendra que le produit vectoriel est une opération accrochée à $\mathbb{R}^3$ et à sa base canonique.

Lemme-Définition 10.31.
Nous avons l'égalité suivante pour tout $u, v, w \in \mathbb{R}^3$:

$$(u \times v) \cdot w = \det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}. \quad \text{(10.81)}$$ EQooKJYUooSQgfXU

*Le résultat est nommé le **produit mixte** de trois vecteurs de $\mathbb{R}^3$.*

21. D'autant plus que, au lieu d'écrire un produit scalaire $a \cdot b$, il écrivent $a_i b^i$ en disant que b^i est un covecteur. Avec toutes les complications que ça implique sur les « lois de transformation » d'un covecteur sous tel ou tel groupe.

10.32.
Nous avons donné un nom à la combinaison $(u \times v) \cdot w$. J'imagine que vous voyez pourquoi nous ne considérons pas la combinaison $(u \cdot v) \times w$.

Le lemme suivant donne un moyen compliqué et peu pratique de calculer la valeur absolue du produit mixte. La formule (10.82) ne sera utilisée que pour faire le lien entre un jacobien et un élément de volume en dimension trois lorsque nous verrons les intégrales sur des variétés. Voir l'équation (20.38).

LEMooSMWNooCmEZeY

Lemme 10.33 ([1])**.**
Le produit mixte peut également être exprimé par

$$|(u \times v) \cdot w|^2 = \det \begin{pmatrix} \|u\|^2 & u \cdot v & u \cdot w \\ v \cdot u & \|v\|^2 & v \cdot w \\ w \cdot u & w \cdot v & \|w\|^2 \end{pmatrix}. \tag{10.82}$$

EQooWZUQooYxdphW

Démonstration. Si nous notons

$$a = \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}, \tag{10.83}$$

il faut simplement remarquer que

$$\begin{pmatrix} \|u\|^2 & u \cdot v & u \cdot w \\ v \cdot u & \|v\|^2 & v \cdot w \\ w \cdot u & w \cdot v & \|w\|^2 \end{pmatrix} = aa^t. \tag{10.84}$$

Donc au niveau des déterminants, en utilisant les propositions 8.249 et le lemme 4.81 nous avons

$$\det \begin{pmatrix} \|u\|^2 & u \cdot v & u \cdot w \\ v \cdot u & \|v\|^2 & v \cdot w \\ w \cdot u & w \cdot v & \|w\|^2 \end{pmatrix} = \det(aa^t) = \det(a)\det(a^t) = \det(a)^2. \tag{10.85}$$

Et maintenant, par définition, $\det(a) = (u \times w) \cdot w$. Donc le résultat annoncé. □

PropScalMixtLin

Proposition 10.34.
Les applications produit scalaire, vectoriel et mixte sont multilinéaires. Spécifiquement, nous avons les propriétés suivantes.

(1) *Les applications produit scalaire et vectoriel sont bilinéaires. C'est-à-dire que pour tout vecteurs a, b, c et pour tout nombre α et β nous avons*

$$\begin{aligned} a \times (\alpha b + \beta c) &= \alpha(a \times b) + \beta(a \times c) \\ (\alpha a + \beta b) \times c &= \alpha(a \times c) + \beta(b \times c). \end{aligned} \tag{10.86}$$

(2) *Le produit mixte est trilinéaire.*

(3) *Le produit vectoriel est antisymétrique, c'est-à-dire $u \times v = -v \times u$.*

(4) *Nous avons $u \times v = 0$ si et seulement si u et v sont colinéaires, c'est-à-dire si et seulement si l'équation $\alpha u + \beta v = 0$ a une solution différente de la solution triviale $(\alpha, \beta) = (0, 0)$.*

PROPooMXAIooJureOD

Proposition 10.35 (Identité de Lagrange[322])**.**
Si $x, y \in \mathbb{R}^n$, alors

$$\|x\|^2\|y\|^2 - (x \cdot y)^2 = \sum_j \sum_{i<j} (x_i y_j - x_j y_i)^2. \tag{10.87}$$

Et si $n = 3$ alors

$$\|x \times y\|^2 = \|x\|^2\|y\|^2 - (x \cdot y)^2. \tag{10.88}$$

Démonstration. C'est un calcul. D'abord nous avons

$$\|x\|^2\|y\|^2 - (x \cdot y)^2 = \sum_i x_i^2 \sum_j y_j^2 - \big(\sum_k x_k y_k\big)^2 = \sum_{ij} x_i^2 y_j^2 - \sum_{kl} x_k y_k x_l y_l. \tag{10.89}$$

Ensuite nous coupons les sommes de la façon suivante

$$\sum_{ij} = \sum_j \sum_{i<j} + \sum_j (i=j) + \sum_j \sum_{i>j} \tag{10.90}$$

pour obtenir

$$\begin{aligned}\|x\|^2\|y\|^2 - (x \cdot y)^2 = &\sum_j \sum_{i<j} x_i^2 y_j^2 + \sum_j x_j^2 y_j^2 + \sum_j \sum_{i>j} x_i^2 y_j^2 \\ &- \sum_l \sum_{k<l} x_k y_k x_l y_l - \sum_k x_k^2 y_k^2 - \sum_l \sum_{k>l} x_k y_k x_l y_l.\end{aligned} \tag{10.91}$$

Il y a deux termes qui se simplifient. Notez que si A_{kl} est symétrique en kl nous avons

$$\sum_l \sum_{k<l} A_{kl} = \sum_k \sum_{l<k} A_{lk} = \sum_k \sum_{l<k} A_{kl}. \tag{10.92}$$

La première égalité était seulement un renommage des indices. Le coup des indices symétriques est justement ce qu'il se passe dans les deux termes en$x_k y_k x_l y_l$, donc nous les regroupons :

$$\|x\|^2\|y\|^2 - (x \cdot y)^2 = \sum_j \big(\sum_{i<j} x_i^2 x_j^2 + \sum_{i>j} x_i^2 y_j^2 - 2 \sum_{i>j} x_i y_i x_j y_j \big) \tag{10.93a}$$

$$= \sum_j \sum_{i<j} (x_i^2 y_j^2 + x_j^2 y_i^2 - 2 x_i y_i x_j y_j) \tag{10.93b}$$

$$= \sum_j \sum_{i<j} (x_i y_j - x_j y_i)^2. \tag{10.93c}$$

Voilà qui prouve la première formule. Pour la seconde, il faut seulement poser $n = 3$ et écrire les sommes explicitement.

— Pour $j = 1$, la somme sur i est $\sum_{i<1}$, c'est-à-dire aucun termes.
— Pour $j = 2$, il y a seulement $i = 1$, donc le terme $(x_1 y_2 - x_2 y_1)^2$.
— Pour $j = 3$, il y a les termes $i = 1$ et $i = 2$, donc les termes $(x_1 y_3 - x_3 y_1)^2 + (x_2 y_3 - x_3 y_2)^2$.

Ces trois termes collectés sont justement les composants (au carré) de $x \times y$ données dans la formule (10.79a). □

Les trois vecteurs de base e_x, e_y et e_y ont des produits vectoriels faciles à retenir :

$$\begin{aligned} e_x \times e_y &= e_z \\ e_y \times e_z &= e_x \\ e_z \times e_x &= e_y \end{aligned} \tag{10.94}$$

Les deux formules suivantes, qui mêlent le produit scalaire et le produit vectoriel, sont souvent utiles en analyse vectorielle :

$$\begin{aligned} (u \times v) \cdot w &= u \cdot (v \times w) \\ (u \times v) \times w &= -(v \cdot w)u + (u \cdot w)v \end{aligned} \tag{10.95}$$ EqFormExpluxxx

pour tout vecteurs u, v et w dans $\mathbb{R}^3$. Nous les admettons sans démonstration. La seconde formule est parfois appelée **formule d'expulsion**.

Exemple 10.36.
Calculons le produit vectoriel $v \times w$ avec

$$v = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \quad w = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}. \tag{10.96}$$

Les vecteurs s'écrivent sous la forme $v = 3e_x - e_y + e_z$ et $w = e_x + 2e_y - e_z$. Le produit vectoriel s'écrit

$$\begin{aligned}(3e_x - e_y + e_z) \times (e_x + 2e_y - e_z) &= 6e_x \times e_y - 3e_x \times e_z \\ &\quad - e_y \times e_x + e_y \times e_z \\ &\quad + e_z \times e_x + 2e_z \times e_y \\ &= 6e_z + 3e_y + e_z + e_x + e_y - 2e_x \\ &= -e_x + 4e_y + 7e_z.\end{aligned} \tag{10.97}$$

△

10.1.7 Produit mixte

Si a, b et c sont trois vecteurs, leur **produit mixte** est le nombre $a \cdot (b \times c)$. En écrivant le produit vectoriel sous forme de somme de trois déterminants 2×2, nous avons

$$\begin{aligned}&a \cdot (b \times c) \\ &= (a_1e_x + a_2e_y + a_3e_z) \cdot \left(\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} e_x - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} e_y + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \right) \\ &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.\end{aligned} \tag{10.98}$$

Le produit mixte s'écrit donc sous forme d'un déterminant. Nous retenons cette formule :

$$a \cdot (b \times c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{10.99}$$ EqProduitMixteDet

Un grand intérêt du produit vectoriel est qu'il fournit un vecteur qui est simultanément perpendiculaire aux deux vecteurs donnés.

PROPooTUVKooOQXKKl

Proposition 10.37.
Le produit vectoriel[22] $a \times b$ *est un vecteur orthogonal à* a *et* b.

Démonstration. Vérifions que $a \perp (a \times b)$. Pour cela, nous calculons $a \cdot (a \times b)$, c'est-à-dire le produit mixte

$$a \cdot (a \times b) = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0. \tag{10.100}$$

L'annulation de ce déterminant est due au fait que deux de ses lignes sont égales. □

Ces résultats admettent une intéressante généralisation.

LEMooFRWKooVloCSM

Lemme 10.38.
Soit $X \in \mathbb{R}^n$ *ainsi que* $v_1, \dots, v_{n-1} \in \mathbb{R}^n$. *Alors*

22. Définition 10.28.

(1) Nous avons

$$\det(X, v_1, \dots, v_{n-1}) = X \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} \tag{10.101}$$

EQooMQNPooRHHBjz

(2) Le vecteur

$$\det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} \tag{10.102}$$

est orthogonal à tous les v_i.

Démonstration. Puisque les deux côtés de (10.101) vus comme fonctions de X, sont des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$, il suffit de vérifier l'égalité sur une base.

Nous posons $\tau_i \colon \mathbb{R}^n \to \mathbb{R}^{n-1}$,

$$\tau_i(v)_k = \begin{cases} v_k & \text{si } k < i \\ v_{k+1} & \text{si } k \geqslant i. \end{cases} \tag{10.103}$$

et nous avons d'une part

$$e_k \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} = \det \begin{pmatrix} \tau_k v_1 \\ \vdots \\ \tau_k v_{n-1} \end{pmatrix} \tag{10.104}$$

et d'autre part,

$$\det(e_k, v_1, \dots, v_{n-1}) = \det \begin{pmatrix} 0 & & & \\ \vdots & & & \\ 1 & v_1 & \cdots & v_{n-1} \\ \vdots & & & \\ 0 & & & \end{pmatrix} = \det(\tau_k v_1, \dots, \tau_k v_{n-1}). \tag{10.105}$$

La première assertion est démontrée.

En ce qui concerne la seconde, il suffit d'appliquer la première et se souvenir qu'un déterminant est nul lorsque deux lignes sont égales [23]. En effet :

$$v_k \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} = \det(v_k, v_1, \dots, v_n) = 0. \tag{10.106}$$

□

10.1.8 Procédé de Gram-Schmidt

PropUMtEqkb

Proposition 10.39 (Procédé de Gram-Schmidt)**.**
Un espace euclidien possède une base orthonormée.

23. Corolaire 4.84.

Démonstration. Soit E un espace euclidien et $\{v_1, \dots, v_n\}$, une base quelconque de E. Nous posons d'abord

$$f_1 = v_1, \quad e_1 = \frac{f_1}{\|f_1\|}. \tag{10.107}$$

Ensuite

$$f_2 = v_2 - \langle v_2, e_1\rangle e_1, \quad e_2 = \frac{f_2}{\|f_2\|}. \tag{10.108}$$

Notons que $\{e_1, e_2\}$ est une base de $\text{Span}\{v_1, v_2\}$. De plus elle est orthogonale :

$$\langle e_1, f_2\rangle = \langle e_1, v_2\rangle - \langle v_2, e_1\rangle \underbrace{\langle e_1, e_1\rangle}_{=1} = 0. \tag{10.109}$$

Le fait que $\|e_1\| = \|e_2\| = 1$ est par construction. Nous avons donc donné une base orthonormée de $\text{Span}\{v_1, v_2\}$.

Nous continuons par récurrence en posant

$$f_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i\rangle e_i, \quad e_k = \frac{f_k}{\|f_k\|}. \tag{10.110}$$

Pour tout $j < k$ nous avons

$$\langle e_j, f_k\rangle = \langle e_j, v_k\rangle - \sum_{i=1}^{k-1} \langle v_k, e_i\rangle \underbrace{\langle e_i, e_j\rangle}_{=\delta_{ij}} = 0 \tag{10.111}$$

□

Cet algorithme de Gram-Schmidt nous donne non seulement l'existence de bases orthonormée pour tout espace euclidien, mais aussi le moyen d'en construire à partir de n'importe quelle base.

10.1.9 Pseudo-réduction simultanée

CorNHKnLVA

Corolaire 10.40 (Pseudo-réduction simultanée[323])**.**
Soient $A, B \in \text{S}(n, \mathbb{R})$ avec A définie positive[24]*. Alors il existe $Q \in \text{GL}(n, \mathbb{R})$ telle que Q^tBQ soit diagonale et $Q^tAQ = \mathbb{1}$.*

Démonstration. Nous allons noter $x \cdot y$ le produit scalaire usuel de $\mathbb{R}^n$ et $\{e_i\}_{i=1,\dots,n}$ sa base canonique.

Comme A est définie positive, l'expression $\langle x, y\rangle = x \cdot Ay$ donne un produit scalaire sur $\mathbb{R}^n$. Nous avons donc deux produits scalaires sur $\mathbb{R}^n$, et nous allons travailler avec les deux.

La proposition 10.39 appliquée à l'espace euclidien $(\mathbb{R}^n, \langle ., .\rangle)$ dit qu'il existe une base de $\mathbb{R}^n$ orthonormée $(f_i)_{i=1,\dots,n}$ pour ce produit scalaire. Nous considérons l'application linéaire P définie par

$$Pe_i = f_i. \tag{10.112}$$

Nous démontrons à présent que $P^tAP = \mathbb{1}$. Pour cela, nous calculons

SUBEQooGZDJooVMuWNn

$$\begin{aligned}
\delta_{ij} &= \langle f_i, f_j\rangle && (10.113a)\\
&= f_i \cdot Af_j && (10.113b)\\
&= Pe_i \cdot APe_j && (10.113c)\\
&= e_i \cdot P^tAPe_j && (10.113d)\\
&= (P^tAP)_{ij}. && (10.113e)
\end{aligned}$$

SUBEQooQNYUooNbyIzM

SUBEQooITBKooCEmqxx

Justifications :

24. Définition 8.230.

— Pour (10.113a), la base (f_j) est orthonormée pour le produit scalaire $\langle .,. \rangle$.

— Pour (10.113d), la proposition 8.186 sur la transposée.

— Pour (10.113e), la formule du produit scalaire usuel pour avoir les éléments de matrice, proposition 8.184.

La matrice P^tBP est une matrice symétrique, donc le théorème spectral 8.227 nous donne une matrice $R \in \mathrm{O}(n, \mathbb{R})$ telle que R^tP^tBPR soit diagonale. En posant maintenant $Q = PR$ nous avons la matrice cherchée. □

REMooDDUEooJXZIfE

Remarque 10.41.
Plusieurs remarques

(1) Nous n'avons pas prouvé l'existence d'une matrice P telle que $P^{-1}BP$ et $P^{-1}AP$ soient diagonales. Au contraire, nous avons Q^tBQ et Q^tAQ qui sont diagonales. Tant que Q n'est pas orthogonale, ce n'est pas la même chose.

Autrement dit, nous n'avons pas ici une vraie diagonalisation, parce que les matrices A et B ne sont pas semblables à des matrices diagonales. Voir les définitions 8.215 (diagonalisable) et 8.202 (semblable).

C'est pour cela que nous parlons de *pseudo*-diagonalisation.

(2) Dans le même ordre d'idée, la démonstration de la pseudo-diagonalisation simultanée parle clairement de formes bilinéaires, et non d'endomorphismes. Or en comparant les lois de transformations (4.233) et (8.315), nous voyons bien que la réduction en passant par Q^tAQ est bien une réduction de forme bilinéaire et non une réduction d'endomorphismes.

(3) Nous avons prouvé la pseudo-réduction simultanée comme corolaire du théorème de diagonalisation des matrices symétriques 8.227. Il aurait aussi pu être vu comme un corolaire du théorème spectral 10.8 sur les opérateurs autoadjoints via son corolaire 10.9.

10.2 Approximations

Le lemme suivant est surtout intéressant en dimension infinie.

Lemme 10.42.
Soit un espace vectoriel normé V et un sous-espace vectoriel dense A. Soit $v \in V$; il existe une suite (v_n) dans A telle que $v_n \xrightarrow{V} v$ et $\|v_n\| \leqslant \|v\|$ pour tout n.

Démonstration. Puisque A est dense, il existe une suite a_n dans A telle que $a_n \to v$. Ensuite il suffit de poser

$$v_n = \frac{n}{n+1}\frac{\|v\|}{\|a_n\|}a_n. \tag{10.114}$$

Par construction nous avons toujours

$$\|v_n\| = \frac{n}{n+1}\|v\| \leqslant \|v\|. \tag{10.115}$$

Et de plus, la norme étant continue[25],

$$\lim_{n\to\infty} v_n = \lim_{n\to\infty} \frac{n}{n+1} \lim_{n\to\infty} \frac{\|v\|}{\|a_n\|} \lim_{n\to\infty} a_n = v. \tag{10.116}$$

Le fait que v_n soit dans A est dû au fait que A soit vectoriel. □

PROPooVEMGooYKhMFy

Proposition 10.43.
Soit un espace vectoriel normé V et un sous-espace vectoriel dense A. Soit $v \in V$. Nous avons

$$\sup\{|v \cdot a| \text{ tel que } a \in A \text{ et } \|a\| \leqslant \lambda\} = \lambda\|v\|. \tag{10.117}$$

25. Où dans le calcul suivant utilisons-nous la continuité de la norme ? Posez vous la question.

Démonstration. D'abord pour tout $a \in A$ vérifiant $\|a\| \leqslant \lambda$ l'inégalité de Cauchy-Schwarz 10.1 donne

$$|v \cdot a| \leqslant \|v\|\|a\| \leqslant \lambda\|v\|. \tag{10.118}$$

Donc le supremum dont on parle est majoré par $\lambda\|v\|$.

Il nous faut l'inégalité dans l'autre sens. Par densité nous pouvons choisir une suite $v_n \in A$ tel que $v_n \to v$. Ensuite nous posons

$$a_n = \frac{\lambda}{\|v_n\|}v_n. \tag{10.119}$$

Nous avons $\|a_n\| = \lambda$ pour tout n et

$$|v \cdot a_n| = \frac{\lambda}{\|v_n\|}|v \cdot v_n|, \tag{10.120}$$

et en passant à la limite,

$$\lim_{n\to\infty} |v \cdot a_n| = \frac{\lambda}{\|v\|}\|v \cdot v\| = \lambda\|v\|. \tag{10.121}$$

Donc l'ensemble sur lequel nous prenons le supremum contient une suite convergente vers $\lambda\|v\|$. Le supremum est donc au moins aussi grand que cela. □

10.2.1 Quelques exemples de normes sur $\mathbb{R}^n$

Il est possible de définir de nombreuses normes sur $\mathbb{R}^n$. Citons-en quelques-unes parmi les normes $\|.\|_p$. Le cas général $p \geqslant 1$ sera fait dans 17.105.

PROPooCLZRooIRxCnZ

Proposition-Définition 10.44 ([324, 218])**.**
Les formules suivantes définissent des normes[26] *sur* $\mathbb{R}^n$*.*

ITEMooQBLGooPQKSev

(1) $\|x\|_1 = \sum_{i=1}^n |x_i|$,

ITEMooXQUFooLHrITI

(2) $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$,

ITEMooSOVDooTuhEik

(3) $\|x\|_\infty = \max_i |x_i|$.

La norme $\|.\|_\infty$ *est nommée* ***norme supremum****.*

Démonstration. Point par point.

(1) **Pour (1)** Déjà fait dans le lemme 10.14.

(2) **Pour (2)** Le cas $p = 2$ provient de l'inégalité

$$\sqrt{(a+b)^2} \leqslant \sqrt{a^2} + \sqrt{b^2}, \tag{10.122}$$

laquelle se démontre en passant au carré :

$$(a+b)^2 = a^2 + b^2 + 2ab \leqslant a^2 + b^2 + 2|ab| = \left(\sqrt{a^2} + \sqrt{b^2}\right)^2. \tag{10.123}$$

EQooRYNYooTzZpPz

(3) **Pour (3)** Pour l'inégalité triangulaire, nous faisons

$$\|x+y\|_\infty = \max_i |x_i + y_i| \leqslant \max_i \left(|x_i| + |y_i|\right) \leqslant \max_i |x_i| + \max_i |y_i| = \|x\|_\infty + \|y\|_\infty. \tag{10.124}$$

Les autres points de la définition 7.158 sont faciles quand on se rappelle que $x = 0$ si et seulement si $x_i = 0$ pour tout i.

□

26. Définition 7.158.

Parmi ces normes, celles qui seront le plus souvent utilisées sont les normes $\|.\|_1$ et $\|.\|_2$ qui, pour des raisons que nous verrons beaucoup plus tard [27] sont souvent notées $\|.\|_{L^p}$:

$$\begin{aligned}\|x\|_{L^1} &= \sum_{i=1}^{n} |x_i|,\\ \|x\|_{L^2} &= \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2}.\end{aligned} \tag{10.125}$$

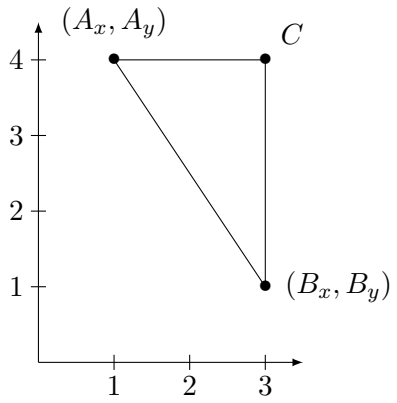


FIGURE 10.1: La *norme* euclidienne induit la *distance* euclidienne. D'où son nom. Le point C est construit aux coordonnées (A_x, B_y). LabelFigDistanceEuclide

Soient $A = (A_x, A_y)$ et $B = (B_x, B_y)$ deux éléments de $\mathbb{R}^2$. La distance [28] euclidienne entre A et B est donnée par $\|A - B\|_2$. En effet, sur la figure 10.1, la distance entre les points A et B est donnée par

$$|AB|^2 = |AC|^2 + |CB|^2 = |A_x - B_x|^2 + |A_y - B_y|^2, \tag{10.126}$$

par conséquent,

$$|AB| = \sqrt{|A_x - B_x|^2 + |A_y - B_y|^2} = \|A - B\|_2. \tag{10.127}$$

Remarque 10.45.
Si A, B et C sont trois points dans le plan $\mathbb{R}^2$, alors l'inégalité triangulaire $|AB| \leqslant |AC| + |CB|$ est précisément la propriété (4) de la norme (définition 7.158). En effet l'inégalité triangulaire s'exprime de la façon suivante en termes de la norme $\|.\|_2$:

$$\|A - B\|_2 \leqslant \|A - C\|_2 + \|C - B\|_2. \tag{10.128}$$ EqNDeuxAmBNNdd

En notant $u = A - C$ et $v = C - B$, l'équation (10.128) devient exactement la propriété de définition de la norme :

$$\|u + v\|_2 \leqslant \|u\|_2 + \|v\|_2. \tag{10.129}$$

Ceci explique pourquoi cette propriété des normes est appelée « inégalité triangulaire ».

10.3 Équivalence des normes

normes_equiv

Au premier coup d'œil, les notions dont nous parlons dans ce chapitre ont l'air très générales. Nous prenons en effet n'importe quel espace vectoriel V de dimension finie, et nous le munissons de n'importe quelle norme. Rien que dans $\mathbb{R}^m$ nous allons en définir une infinité par l'équation (17.429)). À partir de ces données, nous définissons les boules, la topologie, l'adhérence, etc.

27. La proposition 27.16, par exemple.
28. Ne pas confondre « distance » et « norme ».

10.3.1 En dimension finie

Dans $\mathbb{R}^n$, les normes $\|.\|_{L^1}$, $\|.\|_{L^2}$ et $\|.\|_\infty$ ne sont pas égales. Cependant elles ne sont pas complètement indépendantes au sens où l'on sent bien que si un vecteur sera grand pour une norme, il sera également grand pour les autres normes ; les normes « vont dans le même sens ». Cette notion est précisée par le concept de norme équivalente.

LEMooHAFEqWdNoAN DéfEquivNoAN

Proposition-Définition 10.46 ([325]).
Soient deux normes N_1 et N_2 sur l'espace vectoriel E de dimension finie. Nous disons que $N_1 \sim N_2$ si et seulement si il existe des réels strictement positifs k_1, k_2 tels que

$$k_1 N_1(x) \leqslant N_2(x) \leqslant k_2 N_1(x). \tag{10.130}$$

La relation $\sim$ est une relation d'équivalence[29] *sur l'ensemble des normes de E.*
*Si $N_1 \sim N_2$ nous disons que les normes N_1 et N_2 sur $\mathbb{R}^m$ sont **équivalentes**.*

Démonstration. Vu que tous les nombres sont strictement positifs, nous pouvons multiplier et diviser sans changer le sens des inégalités. Gardant cela en tête, nous faisons les trois vérifications.

(1) **réflexive** Nous avons $N_1 \sim N_1$ avec $k_1 = k_2 = 1$.

(2) **Symétrique** Supposons avoir k_1, k_2 tels que $k_1 N_1(x) \leqslant N_2(x) \leqslant k_2 N_1(x)$. Alors nous avons aussi

$$\frac{1}{k_2} N_2(x) \leqslant N_1(x) \leqslant \frac{1}{k_1} N_2(x). \tag{10.131}$$

(3) **Transitive** Supposons $N_1 \sim N_2$ et $N_2 \sim N_3$. Nous avons donc des nombres k_1, k_2, l_1 et l_2 tels que

$$\begin{aligned} k_1 N_1(x) &\leqslant N_2(x) \\ N_2(x) &\leqslant k_2 N_1(x) \\ l_1 N_2(x) &\leqslant N_3(x) \\ N_3(x) &\leqslant l_2 N_2(x). \end{aligned} \tag{10.132}$$

En combinant,

$$k_1 l_1 N_1(x) \leqslant l_1 N_2(x) \leqslant N_3(x) \leqslant l_2 N_2(x) \leqslant l_2 k_2 N_1(x). \tag{10.133}$$

En particulier, $k_1 l_1 N_1(x) \leqslant N_3(x) \leqslant l_2 k_2 N_1(x)$. □

PropLJEJooMOWPNi

Proposition 10.47.
Pour $\mathbb{R}^n$, nous avons les équivalences de normes $\|.\|_{L^1} \sim \|.\|_{L^2}$, $\|.\|_{L^1} \sim \|.\|_\infty$ et $\|.\|_{L^2} \sim \|.\|_\infty$. Plus précisément nous avons les inégalités[30]

ItemABSGooQODmLNi

(1) $\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n}\|x\|_2$

ItemABSGooQODmLNii

(2) $\|x\|_\infty \leqslant \|x\|_1 \leqslant n\|x\|_\infty$

ItemABSGooQODmLNiii

(3) $\|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n}\|x\|_\infty$

Démonstration. En mettant au carré la première inégalité nous voyons que nous devons vérifier l'inégalité

$$|x_1|^2 + \cdots + |x_n|^2 \leqslant \big(|x_1| + \cdots + |x_n|\big)^2 \tag{10.134}$$

qui est vraie parce que le membre de droite est égal au carré de chaque terme plus les double produits. La seconde inégalité provient de l'inégalité de Cauchy-Schwarz (théorème 10.1) sur les vecteurs

$$v = \begin{pmatrix} 1/n \\ \vdots \\ 1/n \end{pmatrix}, \quad w = \begin{pmatrix} |x_1| \\ \vdots \\ |x_n| \end{pmatrix}. \tag{10.135}$$

29. Définition 1.32.
30. Les racines carrés sont définies en 9.97.

Nous trouvons

$$\frac{1}{n}\sum_i |x_i| \leqslant \sqrt{n \cdot \frac{1}{n^2}}\sqrt{\sum_i |x_i|^2}, \tag{10.136}$$

et par conséquent

$$\sum_i |x_i| \leqslant \sqrt{n}\|x\|_2. \tag{10.137}$$

La première inégalité de (3) se démontre en remarquant que si a et b sont positifs, $a \leqslant \sqrt{a^2+b}$. En appliquant cela à $a = \max_i |x_i|$, nous avons

$$\max_i |x_i| \leqslant \sqrt{|x_1|^2 + \cdots + |x_n|^2} \tag{10.138}$$

parce que $\max_i |x_i|$ est évidemment un des termes de la somme. Pour la seconde inégalité de (3), nous avons

$$\sqrt{\sum_k |x_k|^2} \leqslant \left(\sum_k \max_i |x_i|^2\right)^{1/2} = \sqrt{n}\|x\|_\infty. \tag{10.139}$$

Pour obtenir cette inégalité, nous avons remplacé tous les termes $|x_k|$ par le maximum. □

Pour les autres normes $\|.\|_p$, il y a des inégalités dans 17.106 et 17.98 ; voir aussi le thème 25.
Une dernière avant l'équivalence de toutes les normes.

Proposition-Définition 10.48.
Les topologies suivantes sont égales sur $\mathbb{R}^n$.

ITEMooWACPooFBAWhx

(1) La topologie produit $\mathbb{R} \times \ldots \times \mathbb{R}$ des espaces topologiques $(\mathbb{R}, |.|)$,

ITEMooJPJHooGTuLen

(2) La topologie de la norme

$$\|(x_1, \ldots, x_n)\|_\infty = \max_i\{|x_i|\}, \tag{10.140}$$

ITEMooEBYQooXiOOtb

(3) La topologie de la norme

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{\sum_{i=1}^n x_i^2}. \tag{10.141}$$

Elle est la topologie que nous allons toujours considérer sur $\mathbb{R}^n$ (sauf mention très explicite du contraire).

Démonstration. Les topologies (1) et (2) sont identiques par le lemme 7.227. Les topologies (2) et (3) sont identiques par la proposition 10.47. □

En réalité, toutes les normes $\|.\|_{L^p}$ et $\|.\|_\infty$ sont équivalentes et, plus généralement, en dimension finie, toutes les normes sont équivalentes.

ThoNormesEquiv

Théorème 10.49 ([326, 327])**.**
Sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes.

Démonstration. Soit un espace vectoriel V de dimension finie. Nous allons montrer que toutes les normes sont équivalentes à la norme $\|.\|_\infty$. Soit donc une norme N sur V.

(1) **Inégalité dans un sens** Vu que V est de dimension finie, il accepte une base $\{e_1, \ldots, e_n\}$. En posant $D = \sum_{i=1}^n N(e_i)$, nous avons

$$N(x) = N(\sum_i x_i e_i) \leqslant \sum_i |x_i| N(e_i) \leqslant n\|x\|_\infty D. \tag{10.142}$$

Nous avons donc l'inégalité [31]

$$N(x) \leqslant nD\|x\|_\infty. \tag{10.143}$$

31. Dans 10.49, le n n'apparaît pas dans la majoration. C'est lui ou moi qui fait une erreur ? Pourquoi ?

(2) **Dans l'autre sens** La proposition 7.159 donne

$$\big|N(x)-N(y)\big| \leqslant N(x-y) \leqslant Dn\|x-y\|_\infty. \tag{10.144}$$

Donc l'application $N\colon (V,\|.\|_\infty) \to \mathbb{R}$ est continue. Vu que la sphère

$$S = \{x \in V \text{ tel que } \|x\|_\infty = 1\} \tag{10.145}$$

est compacte[32], la fonction continue N y prend un minimum[33]. Il existe donc $m>0$ tel que $m \leqslant N(x)$ pour tout $x \in S$.

Soit $x \in V$. Nous avons $x/\|x\|_\infty \in S$, de telle sorte que

$$m \leqslant N(\frac{x}{\|x\|_\infty}) = \frac{1}{\|x\|_\infty}N(x), \tag{10.146}$$

ou encore $N(x) \geqslant m\|x\|_\infty$.

(3) **Conclusion** Pour tout $x \in V$ nous avons les inégalités

$$m\|x\|_\infty \leqslant N(x) \leqslant nD\|x\|_\infty. \tag{10.147}$$

Donc toutes les normes sont équivalentes à la norme $\|.\|_\infty$.

Comme l'équivalence de norme est transitive, toutes les normes sont équivalentes.

□

CORooBRDYooLmGJDE

Corolaire 10.50.
Soit V un espace vectoriel de dimension finie et $\|.\|_1$, $\|.\|_2$ deux normes sur V. Alors l'identité $i\colon V \to V$ est un isomorphisme d'espace topologique $(V,\|.\|_1) \to (V,\|.\|_2)$.

De plus les ouverts sont les mêmes : une partie de V est ouverte dans $(V,\|.\|_1)$ si et seulement si elle est ouverte dans $(V,\|.\|_2)$.

Démonstration. Le théorème 10.49 nous dit qu'il existe des nombres $\alpha,\beta>0$ tels que

$$\alpha\|x\|_1 \leqslant \|x\|_2 \leqslant \beta\|x\|_1 \tag{10.148}$$

pour tout $x \in V$. Soit un 2-ouvert $\mathcal{O}$. Nous prouvons que $i^{-1}(\mathcal{O}) = \mathcal{O}$ est un 1-ouvert. Pour cela, soit $a \in \mathcal{O}$. Vu que $\mathcal{O}$ est un 2-ouvert, il existe $r>0$ tel que $B_2(a,r) \subset \mathcal{O}$. Prouvons que $B_1(a,r/\beta) \subset B_2(a,r)$. En effet si $y \in B_1(a,r/\beta)$, alors

$$\|y-a\|_2 \leqslant \beta\|y-a\|_1 \leqslant r. \tag{10.149}$$

Nous avons donc

$$B_1(a,r/\beta) \subset B_2(a,r) \subset \mathcal{O} = i^{-1}(\mathcal{O}). \tag{10.150}$$

Donc $i^{-1}(\mathcal{O})$ contient un 1-ouvert autour de chacun de ses points. Il est donc 1-ouvert.

Pour prouver que i^{-1} est continue, c'est la même chose. □

NORMooNKBCooKziIjx

10.51.
Le lemme 10.14 donnera une norme sur $\mathbb{R}^2$ qui ne dérive pas d'un produit scalaire. Vu que toutes les normes sur $\mathbb{R}^2$ produisent la même topologie (c'est le corolaire 10.50), il y a parfaitement moyen pour deux espaces vectoriels topologiques d'être isomorphes alors que l'un a une norme dérivant d'un produit scalaire et l'autre non.

32. La partie S est fermée et bornée de $(V,\|.\|_\infty)$, voir le théorème 9.24.
33. Théorème de Weierstrass 7.146.

10.52.
Le théorème d'équivalence de norme sera utilisé pour montrer que l'ensemble des formes quadratiques non dégénérées de signature (p,q) est ouvert dans l'ensemble des formes quadratiques, proposition 17.117. Plus généralement il est utilisé à chaque fois que l'on fait de la topologie sur les espaces de matrices en identifiant $\mathbb{M}(n,\mathbb{R})$ à $\mathbb{R}^{n^2}$, pour se rassurer en se disant que ce qu'on fait ne dépend pas de la norme choisie.

Voir aussi ce qu'on en fait en 12.291 pour démontrer la différentiabilité à partir des dérivées partielles.

PROPooNTCFooEcwZwt

Proposition 10.53 ([1]).
Soit un espace vectoriel V de dimension finie sur $\mathbb{C}$. Pour une base $B = \{e_i\}$ de V nous définissons

$$\|\sum_k v_k e_k\|_B = \sqrt{\sum_k |v_k|^2}. \tag{10.151}$$

EQooEGXVooLASQIC

(1) La formule (10.151) définit une norme sur V.

(2) Si B et B' sont des bases de V, alors les topologies induites par la norme $\|.\|_B$ et $\|.\|_{B'}$ sont égales.

Démonstration. Nous commençons par fixer une base $B = \{e_i\}_{i=1,\dots,n}$ de V. Cette base nous permet de définir

$$\begin{aligned}\varphi\colon V &\to \mathbb{C}^n\\ \sum_k v_k e_k &\mapsto (v_1,\dots,v_n).\end{aligned} \tag{10.152}$$

Cette application linéaire permet d'écrire

$$\|v\|_V = \|\varphi(v)\|_{\mathbb{C}^n}. \tag{10.153}$$

À partir de là, la vérification des propriétés de la définition 7.158 est immédiate. Par exemple :

$$\|v+w\| = \|\varphi(v+w)\| = \|\varphi(v)+\varphi(w)\| \leqslant \|\varphi(v)\| + \|\varphi(w)\| = \|v\| + \|w\|. \tag{10.154}$$

En ce qui concerne la seconde assertion, c'est le théorème 10.49. □

10.3.2 Contre-exemple en dimension infinie

SubSecPOlynomesCE

Lorsque nous considérons des espaces vectoriels de dimension infinie, les choses ne sont plus aussi simples. Nous voyons ici sur l'exemple de l'espace des polynômes que le théorème 10.49 n'est plus valable si on enlève l'hypothèse de dimension finie.

On considère l'ensemble des fonctions polynomiales à coefficients réels sur l'intervalle $[0,1]$.

$$\mathcal{P}_{\mathbb{R}}([0,1]) = \{p : [0,1] \to \mathbb{R} \,|\, p : x \mapsto a_0 + a_1 x + a_2 x^2 + \dots,\ a_i \in \mathbb{R},\ \forall i \in \mathbb{N}\}. \tag{10.155}$$

Cet ensemble, muni des opérations usuelles de somme entre polynômes et multiplications par les scalaires, est un espace vectoriel.

Sur $\mathcal{P}(\mathbb{R})$ on définit les normes suivantes

$$\begin{aligned}\|p\|_\infty &= \sup_{x\in[0,1]}\{p(x)\},\\ \|p\|_1 &= \int_0^1 |p(x)|\,dx,\\ \|p\|_2 &= \left(\int_0^1 |p(x)|^2\,dx\right)^{1/2}.\end{aligned} \tag{10.156}$$

Les inégalités suivantes sont immédiates

$$\begin{aligned}\|p\|_1 &= \int_0^1 |p(x)|\,dx \leqslant \|p\|_\infty,\\ \|p\|_2 &= \left(\int_0^1 |p(x)|^2\,dx\right)^{1/2} \leqslant \|p\|_\infty,\end{aligned} \tag{10.157}$$

mais la norme $\|\cdot\|_\infty$ n'est équivalente ni à $\|\cdot\|_1$, ni à $\|\cdot\|_2$. Soit $p_k(x) = x^k$. Alors

$$\begin{aligned}\|p_k\|_\infty &= 1,\\ \|p_k\|_1 &= \int_0^1 x^k\,dx = \frac{1}{k+1},\\ \|p_k\|_2 &= \left(\int_0^1 x^{2k}\,dx\right)^{1/2} = \sqrt{\frac{1}{2k+1}}.\end{aligned} \tag{10.158}$$

Pour $k \to \infty$ les normes $\|p_k\|_1$, $\|p_k\|_2$ tendent vers zéro, alors que la norme $\|p_k\|_\infty$ est constante, donc les normes ne sont pas équivalentes parce que il n'existe pas un nombre positif m tel que

$$\begin{aligned}m\|p_k\|_\infty &\leqslant \|p_k\|_1,\\ m\|p_k\|_\infty &\leqslant \|p_k\|_2,\end{aligned} \tag{10.159}$$

uniformément pour tout k dans $\mathbb{N}$.

PROPooILSOooPdUyFu

Proposition 10.54 ([1])**.**
Soit un espace vectoriel normé E de dimension n ainsi qu'un ouvert $U \subset E$. Il existe un ouvert V de $\mathbb{R}^n$ et une bijection isométrique $g\colon V \to U$.

10.4 Norme opérateur

La proposition suivante donne une norme (au sens de la définition 7.158) sur $\mathcal{L}(V,W)$ dès que V et W sont des espaces vectoriels normés.

DefNFYUooBZCPTr

Définition 10.55 (Norme opérateur[328], thème 40)**.**
Soient des espaces vectoriels normés $(V, \|.\|_V)$ et $(W, \|.\|_W)$. Soient une application linéaire $T\colon V \to W$, et le nombre

$$\|T\|_{\mathcal{L}} = \sup_{\substack{x\in V\\ x\neq 0}} \frac{\|T(x)\|_W}{\|x\|_V}. \tag{10.160}$$

*Le nombre $\|T\|_{\mathcal{L}}$ est la **norme opérateur** de T. Nous disons que cette norme est **subordonnée** aux normes choisies sur V et W.*

Si $\|T\| < \infty$, nous disons que T est borné[34]*. Nous notons*

$$\mathcal{L}^0(V,W) = \{T \in \mathcal{L}(V,W) \text{ tel que } \|T\| < \infty\}. \tag{10.161}$$

PROPooQMZLooOxBrmt

Proposition 10.56.
Si V est de dimension finie, alors $\|T\|_{\mathcal{L}} < \infty$ pour tout $T \in \mathcal{L}(V,W)$.

Démonstration. Si V est de dimension finie alors l'ensemble $\{x \in V \text{ tel que } \|x\| = 1\}$ est compact par le théorème de Borel-Lebesgue 9.24. Alors la fonction

$$x \mapsto \frac{\|T(x)\|}{\|x\|} \tag{10.162}$$

est une fonction continue sur un compact. Le corolaire 7.295 nous dit alors qu'elle est bornée. Le supremum est donc un nombre réel fini. □

34. Nous verrons plus tard que si V est de dimension infinie, ce n'est pas garanti.

PROPooZBJXooNDsqax

Proposition 10.57.
À propos d'opérateurs bornés.

(1) *La partie $\mathcal{L}^0(V,W)$ des applications linéaires bornées est un espace vectoriel.*

(2) *Le couple $\big(\mathcal{L}^0(V,W), \|.\|\big)$ est un espace vectoriel normé*[35].

Démonstration. Nous vérifions que l'application $\|.\|$ de $\mathcal{L}(V,W)$ dans $\mathbb{R}$ ainsi définie est effectivement une norme.

(1) $\|T\|_{\mathcal{L}} = 0$ signifie que $\|T(x)\| = 0$ pour tout x dans V. Comme $\| \cdot \|_W$ est une norme nous concluons que $T(x) = 0$ pour tout x dans V, donc T est l'application nulle.

(2) Pour tout $\lambda \in \mathbb{R}$ et tout T dans $\mathcal{L}(V,W)$ nous avons

$$\|\lambda T\|_{\mathcal{L}} = \sup_{\substack{x\in V\\ x\neq 0}} \frac{\|(\lambda T)(x)\|_W}{\|x\|_V} = \sup \frac{|\lambda|\|T(x)\|_W}{\|x\|} = |\lambda|\|T\|_{\mathcal{L}}. \tag{10.163}$$

(3) Pour tous T_1 et T_2 dans $\mathcal{L}(V,W)$ nous avons

$$\begin{aligned}
\|T_1 + T_2\|_{\mathcal{L}} &= \sup_{\substack{x\in V\\ x\neq 0}} \frac{\|T_1(x) + T_2(x)\|}{\|x\|} && \text{(10.164a)}\\
&\leqslant \sup \frac{\|T_1(x) + T_2(x)\|}{\|x\|} && \text{(10.164b)}\\
&= \sup \left(\frac{\|T_1(x)\|}{\|x\|} + \frac{\|T_2(x)\|}{\|x\|} \right) && \text{(10.164c)}\\
&=\leqslant \sup \frac{\|T_1(x)\|}{\|x\|} + \sup \frac{\|T_2(x)\|}{\|x\|} && \text{(10.164d)}\\
&= \|T_1\| + \|T_2\|. && \text{(10.164e)}
\end{aligned}$$

□

PROPooWHJCooXxCSWp

Proposition 10.58.
Nous avons la formule

$$\|T\|_{\mathcal{L}} = \sup_{x\in V} \frac{\|T(x)\|_W}{\|x\|_V} = \sup_{\|x\|_V=1} \|T(x)\|_W \tag{10.165}$$

EqFZPooIoecGH

Démonstration. Nous allons montrer que les ensembles sur lesquels ont prend le supremum sont en réalité les mêmes :

$$\underbrace{\left\{ \frac{\|Ax\|}{\|x\|} \right\}_{x\neq 0}}_{A} = \underbrace{\{\|Ax\| \text{ tel que } \|x\| = 1\}}_{B}. \tag{10.166}$$

Attention : ce sont des sous-ensembles de réels ; pas de sous-ensembles de $\mathbb{M}(\mathbb{R})$ ou des sous-ensembles de $\mathbb{R}^n$.

Pour la première inclusion, prenons un élément de A, et prouvons qu'il est dans B. C'est-à-dire que nous prenons $x \in V$ et nous considérons le nombre $\|Ax\|/\|x\|$. Le vecteur $y = x/\|x\|$ est un vecteur de norme 1, donc la norme de Ay est un élément de B, mais

$$\|Ay\| = \frac{\|Ax\|}{\|x\|}. \tag{10.167}$$

Nous avons donc $A \subset B$.

L'inclusion $B \subset A$ est immédiate. □

35. Espace vectoriel normé, définition 7.158.

LEMooHGCKooBzfAtg

Lemme 10.59.
Si A est une matrice bloc-diagonale formée des sous-matrices (A_i), alors A^k $(k \in \mathbb{Z})$ est une matrice bloc-diagonale formée des sous-matrices $(A_i)^k$.

PROPooJUYCooHnlFef

Proposition 10.60 ([1])**.**
Si A est une matrice bloc-diagonale, alors la norme de A majore la norme de chacun de ses blocs.

En d'autres termes, il y a autant de normes opérateur sur $\mathcal{L}(E,F)$ qu'il y a de paires de choix de normes sur E et F. En particulier, cela donne lieu à toutes les normes $\|A\|_p$ qui correspondent aux normes $\|.\|_p$ sur $\mathbb{R}^n$.

EXooXPXAooYyBwMX

Exemple 10.61.
Voyons la norme opérateur subordonnée à la norme $\|x\|_\infty = \max_i |x_i|$ sur $\mathbb{C}^n$. Par la proposition 10.58,

$$\|A\|_\infty = \sup_{\|x\|_\infty=1} \|Ax\|_\infty. \tag{10.168}$$

Vu que $(Ax)_i = \sum_k A_{ik}x_k$, lorsque $\|x\|_\infty \leqslant 1$ nous avons $|(Ax)_i| \leqslant \sum_k |A_{ik}|$. Donc nous avons toujours

EQooPLCIooVghasD
$$\|A\|_\infty \leqslant \max_i \sum_k |A_{ik}|. \tag{10.169}$$

△

Définition 10.62.
*La **topologie forte** sur l'espace des opérateurs est la topologie de la norme opérateur.*

Lorsque nous considérons un espace vectoriel d'applications linéaires, nous considérons toujours[36] dessus la topologie liée à cette norme.

10.4.1 Norme d'algèbre

DefJWRWQue

Définition 10.63 (Norme d'algèbre[328])**.**
Si A est une algèbre[37]*, une **norme d'algèbre** sur A est une norme telle que pour toute $x, y \in A$,*

$$\|xy\| \leqslant \|x\|\|y\|. \tag{10.170}$$

La norme opérateur est une norme d'algèbre, comme nous le verrons dans le lemme 10.69.

Un des intérêts d'utiliser une norme d'algèbre est que l'on a l'inégalité $\|x^k\| \leqslant \|x\|^k$. Cela sera particulièrement utile lors de l'étude des séries entières, voir par exemple 15.13.

DEFooEAUKooSsjqaL

Définition 10.64 ([329])**.**
*Le **rayon spectral** d'une matrice carrée A, noté $\rho(A)$, est défini de la manière suivante :*

EQooNVNQooNjIbSS
$$\rho(A) = \max_i |\lambda_i| \tag{10.171}$$

où les λ_i sont les valeurs propres de A.

10.65.
Quelques remarques sur la définition du rayon spectral.

- Même si A est une matrice réelle, les valeurs propres sont dans $\mathbb{C}$. Donc dans (10.171), $|\lambda_i|$ est le module dans $\mathbb{C}$ de λ_i.
- Puisque les valeurs propres de A sont les racines de son polynôme caractéristique (théorème 8.117), il y en a un nombre fini et le maximum est bien défini.

36. Sauf lorsque les événements nous forceront à trahir.
37. Définition 1.347.

— La définition s'applique uniquement pour les espaces de dimension finie.

LEMooIBLEooLJczmu

Lemme 10.66.
Soient des espaces vectoriels normés E et F, sur les corps $\mathbb{R}$ ou $\mathbb{C}$. Pour tout $A \in \mathcal{L}(E,F)$, et pour tout $u \in E$ nous avons la majoration

$$\|Au\| \leqslant \|A\|\|u\| \tag{10.172}$$

où la norme sur A est la norme opérateur subordonnée à la norme sur u.

Démonstration. Si $u \in E$ alors, étant donné que le supremum d'un ensemble est plus grand ou égal à chacun des éléments qui le compose,

$$\|A\| = \sup_{x\in E} \frac{\|Ax\|}{\|x\|} \geqslant \frac{\|Au\|}{\|u\|}, \tag{10.173}$$

donc le résultat annoncé : $\|Au\| \leqslant \|A\|\|u\|$. □

Le lemme suivant est valable en dimension infinie. Nous en toucherons un mot dans l'exemple 10.74.

LEMooWFNXooLyTyyX

Lemme 10.67.
Soient des espaces vectoriels normés E et F. Soit $x \in E$. Alors l'application d'évaluation

$$\begin{aligned} ev_x\colon \mathcal{L}(E,F) &\to F \\ f &\mapsto f(x) \end{aligned} \tag{10.174}$$

est continue.

Démonstration. Si $x = 0$, alors par linéarité de f nous avons $ev_0(f) = 0$ pour tout f. Donc d'accord pour la continuité.

Soit une suite convergente $f_k \xrightarrow{\mathcal{L}(E,F)} f$. Nous voulons prouver que $ev_x(f_k) \xrightarrow{F} ev_x(f)$, c'est-à-dire que

$$\lim_{k\to\infty} \|f_k(x) - f(x)\| = 0. \tag{10.175}$$

Par hypothèse si k est grand, alors $\|f_k - f\|_{\mathcal{L}(E,F)} \leqslant \epsilon$, c'est-à-dire que[38]

$$\sup_{y\in E} \frac{\|f_k(y) - f(y)\|}{\|y\|} \leqslant \epsilon. \tag{10.176}$$

En particulier pour notre x nous avons

$$\frac{\|f_k(x) - f(x)\|}{\|x\|} \leqslant \epsilon, \tag{10.177}$$

c'est-à-dire $\|f_k(x) - f(x)\| \leqslant \|x\|\epsilon$. Vu que $\|x\|$ est une simple constante et que ϵ est arbitraire, cela implique $f_k(x) \to f(x)$. □

10.5 Application linéaire continue et bornée

NORMooUKURooIkcRoS

10.68.
Quelque notations. Par analogie avec C^0 qui est l'ensemble des applications continues, nous notons $\mathcal{L}^0(E,F)$ l'ensemble des applications linéaires continues de E vers F. De même, $\mathrm{End}^0(E)$ est l'ensemble des endomorphismes continus de E.

38. Définition 10.55 de la norme sur $\mathcal{L}(E,F)$.

LEMooFITMooBBBWGI

Lemme 10.69 (La norme opérateur est une norme d'algèbre[1])**.**
Soient des espaces vectoriels normés E, F et G. Soient des opérateurs linéaires bornés $B\colon E \to F$, $A\colon F \to G$. Alors

$$\|A \circ B\| \leqslant \|A\|\|B\|. \tag{10.178}$$

C'est-à-dire que la norme opérateur est une norme d'algèbre[39]*.*

Démonstration. Nous avons les (in)égalités suivantes :

$$\|AB\| = \sup_{x\in E} \frac{\|ABx\|_G}{\|x\|_E} \tag{10.179a}$$

$$= \sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|x\|}\frac{\|Bx\|_F}{\|Bx\|_F} \tag{10.179b}$$

$$= \sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|Bx\|}\frac{\|Bx\|}{\|x\|} \tag{10.179c}$$

$$\leqslant \underbrace{\sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|Bx\|}}_{\leqslant\|A\|} \underbrace{\sup_{\substack{y\in E\\ By\neq 0}} \frac{\|Bx\|}{\|y\|}}_{=\|B\|} \tag{10.179d}$$

$$\leqslant \|A\|\|B\|. \tag{10.179e}$$

La dernière inégalité provient que dans $\sup_{\substack{x\in E\\ Bx\neq 0}} \|ABx\|/\|x\|$, le supremum est pris sur un ensemble plus petit que celui sur lequel porte la définition de la norme de A : seulement l'image de B au lieu de tout l'espace de départ de A. □

La proposition suivante dit qu'une application linéaire est continue si et seulement si elle est bornée. Un résultat similaire pour les applications multilinéaires est dans la proposition 8.6.

PROPooQZYVooYJVlBd

Proposition 10.70 (Bornée si et seulement si continue[330])**.**
Soient E et F des espaces vectoriels normés. Une application linéaire $E \to F$ est bornée si et seulement si elle est continue.

Démonstration. Nous commençons par supposer que A est bornée. Par le lemme 10.69, pour tout $x, y \in E$, nous avons

$$\|A(x) - A(y)\| = \|A(x - y)\| \leqslant \|A\|\|x - y\|. \tag{10.180}$$

En particulier si $x_n \xrightarrow{E} x$ alors

$$0 \leqslant \|A(x_n) - A(x)\| \leqslant \|A\|\|x_n - x\| \to 0 \tag{10.181}$$

et A est continue en vertu de la caractérisation séquentielle de la continuité, proposition 7.137.

Nous supposons maintenant que $\|A\|$ n'est pas borné : l'ensemble $\{\|A(x)\| \text{ tel que } \|x\| = 1\}$ contient des valeurs arbitrairement grandes. Alors pour tout $k \geqslant 1$ il existe $x_k \in B(0,1)$ tel que $\|A(x_k)\| > k$. La suite x_k/k tend vers zéro parce que $\|x_k\| = 1$, mais $\|A(x_k)\| \geqslant 1$ pour tout k. Cela montre que A n'est pas continue. □

Nous avons vu dans la proposition 10.70 que pour une application linéaire, être bornée est équivalent à être continue. Nous allons maintenant voir un certain nombre d'exemples illustrant ce fait.

ExHKsIelG

Exemple 10.71 (Une application linéaire non continue)**.**
Soit V l'espace vectoriel normé des suites *finies* de réels muni de la norme usuelle $\|c\| = \sqrt{\sum_{i=0}^{\infty} |c_i|^2}$ où la somme est finie. Nous nommons $\{e_k\}_{k\in\mathbb{N}}$ la base usuelle de cet espace, et nous considérons

39. Définition 10.63.

l'opérateur $f\colon V\to V$ donnée par $f(e_k)=ke_k$. C'est évidemment linéaire, mais ce n'est pas continu en zéro. En effet la suite $u_k=e_k/k$ converge vers 0 alors que $f(u_k)=e_k$ ne converge pas. △

Cet exemple aurait pu également être donnée dans un espace de Hilbert, mais il aurait fallu parler de domaine.

EXooDMVJooAJywMU

Exemple 10.72 (Une autre application linéaire non continue[331])**.**
En dimension infinie, une application linéaire n'est pas toujours continue. Soit E l'espace des polynômes à coefficients réels sur $[0,1]$ muni de la norme uniforme. L'application de dérivation $\varphi\colon E\to E$, $\varphi(P)=P'$ n'est pas continue.

Pour la voir nous considérons la suite $P_n=\frac{1}{n}X^n$. D'une part nous avons $P_n\to 0$ dans E parce que $P_n(x)=\frac{x^n}{n}$ avec $x\in[0,1]$. Mais en même temps nous avons $\varphi(P_n)=X^{n-1}$ et donc $\|\varphi(P_n)\|=1$.

Nous n'avons donc pas $\lim_{n\to\infty}\varphi(P_n)=\varphi(\lim_{n\to\infty}P_n)$ et l'application φ n'est pas continue en 0. Elle n'est donc continue nulle part par linéarité.

Nous avons utilisé le critère séquentiel de la continuité, voir la définition 7.208 et la proposition 7.137. △

RemOAXNooSMTDuN

Remarque 10.73.
Cette proposition permet de retrouver l'exemple 10.71 plus simplement. Si $\{e_k\}_{k\in\mathbb{N}}$ est une base d'un espace vectoriel normé formée de vecteurs de norme 1, alors l'opérateur linéaire donné par $u(e_k)=ke_k$ n'est pas borné et donc pas continu.

C'est également ce résultat qui montre que le produit scalaire est continu sur un espace de Hilbert par exemple.

EXooTQPEooRRdddt

Exemple 10.74.
Nous avons vu dans le lemme 10.67 que pour un $x\in E$ donné, l'application

$$\begin{aligned}ev_x\colon\mathcal{L}(E,F)&\to F\\ f&\mapsto f(x)\end{aligned}\tag{10.182}$$

est continue. Puisque ev_x est linéaire, la proposition 10.70 nous indique que ev_x est bornée. Vérifions-le directement. Le calcul n'est pas très compliqué :

$$\|ev_x\|=\sup_{\|f\|=1}\|ev_x(f)\|=\sup_{\|f\|=1}\|f(x)\|\leqslant\sup_{\|f\|=1}\|x\|\|f\|=\|x\|\tag{10.183}$$

où nous avons utilisé le lemme 10.66 en passant. Donc la norme de ev_x est majorée par $\|x\|$.

Elle est même égale à $\|x\|$. En effet, pour chaque $f\in\mathcal{L}(E,F)$ tel que $\|f\|=1$, nous avons

$$\|ev_x\|\geqslant\|ev_x(f)\|=\|f(x)\|.\tag{10.184}$$

En prenant $f=\mathrm{Id}$ nous trouvons $\|ev_x\|\geqslant\|x\|$. △

10.5.1 Déterminant entre espaces différents

Il n'existe à priori pas de façons naturelles de définir le déterminant d'une application linéaire $f\colon E\to F$ lorsque E et F sont des espaces vectoriels différents[267].

PROPooQCPJooFSzaPc

Proposition 10.75 ([267, 1])**.**
Soient E et F, deux espaces vectoriels de même dimension. Nous considérons une bijection linéaire $\varphi\colon E\to F$, et nous définissons

$$\begin{aligned}\det{}_\varphi\colon\mathcal{L}(E,F)&\to\mathbb{R}\\ f&\mapsto\det(\varphi^{-1}\circ f).\end{aligned}\tag{10.185}$$

ITEMooERPIooBvkAPo

(1) L'application $f\colon E\to F$ est inversible si et seulement si $\det_\varphi(f)\neq 0$.

ITEMooTLJEooKyPOVs

(2) *Si E et F sont des espaces vectoriels normés, alors l'application $\det_\varphi\colon \mathcal{L}(E,F)\to\mathbb{R}$ est continue.*

Démonstration. En deux parties.

(1) **Pour (1)** Vu que φ est une bijection, nous avons les équivalences

$$\begin{aligned} f \text{ est inversible} &\Leftrightarrow \varphi^{-1}\circ f \text{ est inversible} && (10.186a)\\ &\Leftrightarrow \det(\varphi^{-1}\circ f)\neq 0 \qquad \text{prop. 8.16(2)} && (10.186b)\\ &\Leftrightarrow \det_\varphi(f)\neq 0. && (10.186c)\end{aligned}$$

(2) **Pour (2)** Nous posons

$$\begin{aligned} s\colon \mathcal{L}(E,F)&\to\mathcal{L}(E,E)\\ f&\mapsto\varphi^{-1}\circ f,\end{aligned} \tag{10.186d}$$

et nous commençons par prouver que s est continue. Le lemme 10.69 dit que

$$\|\varphi^{-1}\circ f\|\leqslant\|\varphi^{-1}\|\|f\|, \tag{10.187}$$

de telle sorte que

$$\begin{aligned}\|s\|&=\sup_{\|f\|=1}\|s(f)\| && (10.188a)\\ &\leqslant\sup_{\|f\|=1}\|\varphi^{-1}\|\|f\| && (10.188b)\\ &=\|\varphi^{-1}\|. && (10.188c)\end{aligned}$$

Donc s est bornée, et donc elle est continue (proposition 10.70). Finalement l'application $\det_\varphi=\det\circ s$ est continue comme composée d'applications continues.

□

PROPooPNMXooLaJDGc

Proposition 10.76 ([1])**.**
Soient deux espaces vectoriels normés E et F de même dimension finie. La partie $\mathrm{GL}(E,F)$ des applications linéaires inversibles est ouverte dans $\mathcal{L}(E,F)$.

Pour la preuve, il faudra utiliser la proposition 10.75.

10.5.1.1 Mini bonus

NORMooUDZOooKxAPit

10.77 ([222])**.**
Vous vous souvenez de la définition 7.175 d'une somme directe topologique ? Si V est normé, il existe une façon plus directe (mais pas spécialement plus simple) de prouver l'implication (1) ⇒ (2), et en particulier la continuité de s^{-1}. Rappelons que dans le cas normé, nous avons plusieurs façons équivalentes de décrire les topologies [40].

— Sur V_1 et V_2 nous avons la topologie induite, définition 7.26, qui est la même que celle de la restriction de la norme de V (lemme 7.122).

— La topologie sur $V_1\times V_2$ est la définition 7.15. C'est la même que celle de la norme produit par le lemme 7.227.

— La topologie sur V/V_1 est la topologie quotient 7.49. La proposition 7.357 dit que la norme induite donne la même topologie.

40. Voir le thème 27.

L'espace $V_1 \times V_2$ est muni de sa norme produit de la définition 7.227 : $\|(v_1, v_2)\| = \max\{\|v_1\|, \|v_2\|\}$. Par hypothèse $\psi^{-1} : V_1 \oplus V_2 \to V_1 \times V_2$ est linéaire et continue. La proposition 10.70 nous indique qu'elle est alors bornée. Il existe donc un nombre $C \in \mathbb{R}^+$ tel que

$$\|\psi^{-1}(v_1 + v_2)\| = \max\{\|v_1\|, \|v_2\|\} \leqslant C\|v_1 + v_2\|. \tag{10.189}$$

Nous allons montrer que s^{-1} est séquentiellement continue [41] en partant d'une suite $\alpha_n \xrightarrow{V/V_1} 0$. En vertu de la proposition 7.357, nous avons $d(\alpha_n, V_1) \xrightarrow{\mathbb{R}} 0$.

Chacun des α_n contient un unique élément x_n de V_1, qui est donné par s^{-1} : $x_n = s^{-1}(\alpha_n)$. Nous avons

$$\|\alpha_n\| = \inf_{u \in \alpha_n} \|u\| = \inf_{v \in V_1} \|x_n + v\|. \tag{10.190}$$

Vu que $\|\alpha_n\|$ est donné par un infimum, nous pouvons, pour chaque n, choisir $v \in V_1$ de telle sorte que $x_n + v_n$ ne soit pas trop loin d'être l'infimum :

$$\big|\|x_n + v_n\| - \|\alpha_n\|\big| \leqslant \frac{1}{n}. \tag{10.191}$$

Un tel choix nous donne une suite (v_n) dans V_1 telle que

$$\big|\|x_n + v_n\| - \|\alpha_n\|\big| \to 0. \tag{10.192}$$

Vu que $\|\alpha_n\| \to 0$, cela implique $\|x_n + v_n\| \to 0$.

En remettant tout ensemble,

$$\|x_n\| \leqslant \max\{\|x_n\|, \|v_n\|\} \leqslant C\|x_n + v_n\| \to 0. \tag{10.193}$$

Donc $s^{-1}(\alpha_n) = x_n \to 0$, et s^{-1} est séquentiellement continue.

10.5.2 Suites

Nous allons maintenant parler de suites dans $V \times W$. Nous noterons (v_n, w_n) la suite dans $V \times W$ dont l'élément numéro n est le couple (v_n, w_n) avec $v_n \in V$ et $w_n \in W$. La notion de convergence de suite découle de la définition de la norme via la proposition 9.26. Il se fait que dans le cas des produits d'espaces, la convergence d'une suite est équivalente à la convergence des composantes. Plus précisément, nous avons le lemme suivant.

LemCvVxWcvVW

Lemme 10.78.
La suite (v_n, w_n) converge vers (v, w) dans $V \times W$ si et seulement les suites (v_n) et (w_n) convergent séparément vers v et w respectivement dans V et W.

Démonstration. Pour le sens direct, nous devons étudier le comportement de la norme de $(v_n, w_n) - (v, w)$ lorsque n devient grand. En vertu de la définition de la norme dans $V \times W$ nous avons

$$\big\|(v_n, w_n) - (v, w)\big\|_{V \times W} = \max\big\{\|v_n - v\|_V, \|w_n - w\|_W\big\}. \tag{10.194}$$

Soit $\varepsilon > 0$. Par définition de la convergence de la suite (v_n, w_n), il existe un $N \in \mathbb{N}$ tel que $n > N$ implique

$$\max\big\{\|v_n - v\|_V, \|w_n - w\|_W\big\} < \varepsilon, \tag{10.195}$$

et donc en particulier les deux inéquations

$$\|v_n - v\| < \varepsilon \tag{10.196a}$$
$$\|w_n - w\| < \varepsilon. \tag{10.196b}$$

De la première, il ressort que $(v_n) \to v$, et de la seconde que $(w_n) \to w$.

41. Définition 7.208. La continuité séquentielle est équivalente à la continuité par la proposition 7.255.

Pour le sens inverse, nous avons pour tout ε un N_1 tel que $\|v_n - v\|_V \leqslant \varepsilon$ pour tout $n > N_1$ et un N_2 tel que $\|w_n - w\|_W \leqslant \varepsilon$ pour tout $n > N_2$. Si nous posons $N = \max\{N_1, N_2\}$ nous avons les deux inégalités simultanément, et donc

$$\max\big\{\|v_n - v\|_V, \|w_n - w\|_W\big\} < \varepsilon, \tag{10.197}$$

ce qui signifie que la suite (v_n, w_n) converge vers (v, w) dans $V \times W$. □

PROPooKDGOooDjWQct

Proposition 10.79 ([1])**.**
Soit un espace V muni d'un produit scalaire à valeurs dans $\mathbb{K}$ (si $\mathbb{K} = \mathbb{C}$ nous supposons le produit hermitien, mais ce n'est pas très important ici). Alors l'application

$$\begin{aligned} a\colon V \times V &\to \mathbb{K} \\ (x, y) &\mapsto \langle x, y\rangle \end{aligned} \tag{10.198}$$

est continue.

Démonstration. Nous ne disons pas que l'espace $V \times V$ est muni d'un produit scalaire. Mais en tout cas c'est un espace métrique, et $\mathbb{K}$ l'est aussi. Donc a est une application entre deux espaces métriques et elle sera continue si et seulement si elle est séquentiellement continue (propositions 7.137 et 7.250).

Soit donc une suite convergente dans $V \times V$, c'est-à-dire $(x_k, y_k) \xrightarrow{V\times V} (x, y)$. Nous devons démontrer que $\langle x_k, y_k\rangle \xrightarrow{\mathbb{R}} \langle x, y\rangle$. Les majorations usuelles donnent

$$\begin{aligned} \big|\langle x_k, y_k\rangle - \langle x, y\rangle\big| &\leqslant \big|\langle x_k, y_k\rangle - \langle x, y_k\rangle\big| + \big|\langle x, y_k\rangle - \langle x, y\rangle\big| && \text{(10.199a)} \\ &= \big|\langle x_k - x, y_k\rangle\big| + \big|\langle x, y_k - y\rangle\big|. && \text{(10.199b)} \end{aligned}$$

Nous savons du lemme 10.78 que les suites (x_k) et (y_k) sont séparément convergentes : $x_k \xrightarrow{V} x$ et $y_k \xrightarrow{V} y$. En utilisant l'inégalité de Cauchy-Schwarz 10.2 nous trouvons

$$\big|\langle x_k - x, y_k\rangle\big| \leqslant \|x_k - x\|\|y_k\|. \tag{10.200}$$

Nous avons $\|x_k - x\| \to 0$ et $\|y_k\| \to \|y\|$, et par la règle du produit de limites dans $\mathbb{R}$ nous avons que $\big|\langle x_k - x, y_k\rangle\big| \to 0$. □

RemTopoProdPasRm

Remarque 10.80.
Il faut remarquer que la norme (7.204) est une norme *par défaut*. C'est la norme qu'on met quand on ne sait pas quoi mettre. Or il y a au moins un cas d'espace produit dans lequel on sait très bien quelle norme prendre : les espaces $\mathbb{R}^m$. La norme qu'on met sur $\mathbb{R}^2$ est

$$\|(x, y)\| = \sqrt{x^2 + y^2}, \tag{10.201}$$

et non la norme « par défaut » de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ qui serait

$$\|(x, y)\| = \max\{|x|, |y|\}. \tag{10.202}$$

Les théorèmes que nous avons donc démontré à propos de $V \times W$ ne sont donc pas immédiatement applicables au cas de $\mathbb{R}^2$.

Cette remarque est valable pour tous les espaces $\mathbb{R}^m$. À moins de mention contraire explicite, nous ne considérons jamais la norme par défaut (7.204) sur un espace $\mathbb{R}^m$.

Étant donné la remarque 10.80, nous ne savons pas comment calculer par exemple la fermeture du produit d'intervalle $]0, 1, [\times [4, 5[$. Il se fait que, dans $\mathbb{R}^m$, les fermetures de produits sont quand même les produits de fermetures.

PropovlAxBbarAbraB

Proposition 10.81.
Soient des espaces vectoriels normés V et W. Soient $A \subset V$ et $B \subset W$. Alors

$$\overline{A \times B} = \bar{A} \times \bar{B} \tag{10.203}$$

pour la norme produit[42] *sur $V \times W$.*

Démonstration. Nous commençons par prouver que $\overline{A \times B} \subset \bar{A} \times \bar{B}$. Soit donc $(x, y) \in \overline{A \times B}$. La proposition 7.252 nous assure de l'existence d'une suite dans $A \times B$ convergent vers (x, y). Nous considérons donc une suite (x_n, y_n) dans $A \times B$ telle que

$$(x_n, y_n) \xrightarrow{V \times W} (x, y). \tag{10.204}$$

Soit $\epsilon > 0$. Si n est assez grand nous avons

$$\|(x_n, y_n) - (x, y)\| \leqslant \epsilon. \tag{10.205}$$

Mais

$$\|(x_n, y_n) - (x, y)\| = \|(x_n - x, y_n - y)\| = \max\big(\|x_n - x\|_V, \|y_n - y\|_W\big). \tag{10.206}$$

Donc pour ce n nous avons $\|x_n - x\| \leqslant \epsilon$ et $\|y_n - y\| \leqslant \epsilon$. Nous en déduisons que $x_n \xrightarrow{V} x$ et $y_n \xrightarrow{W} y$. Donc $x \in \bar{A}$ et $y \in \bar{B}$.

Dans l'autre sens maintenant. Soit $(x, y) \in \bar{A} \times \bar{B}$. Nous considérons deux suites (x_n) dans A et (y_n) dans B telles que $x_n \xrightarrow{V} x$ et $y_n \xrightarrow{W} y$. Soit n suffisamment grand pour que $\|x_n - x\| \leqslant \epsilon$ et $\|y_n - y\| \leqslant \epsilon$. Nous avons alors

$$\|(x_n, y_n) - (x, y)\| = \max\{\|x_n - x\|, \|y_n - y\|\} \leqslant \epsilon. \tag{10.207}$$

□

exoEspVectoNorme0002

Proposition 10.82.
Soit $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0\}$.

(1) La partie A est ouverte dans $\mathbb{R}^2$.

(2) L'adhérence de A est donnée par

$$\bar{A} = A \cup \{(x, 0) \text{ tels que } x \geqslant 0\} \cup \{(0, y) \text{ tels que } y \geqslant 0\}. \tag{10.208}$$

(3) La frontière de A est donnée par

$$\partial A = \{(x, 0) \text{ tels que } x \geqslant 0\} \cup \{(0, y) \text{ tels que } y \geqslant 0\}. \tag{10.209}$$

Démonstration. Point par point.

(1) Prenons $(x, y) \in A$, et tâchons de trouver une boule autour de (x, y) qui soit contenue dans A. Par définition, $x > 0$ et $y > 0$. Donc si nous prenons $r = \min\{x, y\}/2$, la boule $B\big((x, y), r\big)$ est encore contenue dans A.

Notez que l'ensemble des boules du type $\mathcal{O}_{(x,y)} = B\big((x, y), \frac{\min\{x,y\}}{2}\big)$ est un recouvrement de A par des ouverts.

(2) L'adhérence est constituée des points qui « touchent » presque l'ensemble. Intuitivement, nous devinons que l'adhérence de A va être l'ensemble des points (x, y) tels que $x \geqslant 0$ et $y \geqslant 0$. D'abord, un point (a, b) avec $a < 0$ n'est pas dans $\bar{A}$ parce qu'il existe une boule autour de (a, b) telle que $x < 0$ pour tout (x, y) dans la boule (même chose pour les points avec $y < 0$).

42. Définition 7.227.

Ensuite, prouvons que les points de la forme $(0,y)$ et $(x,0)$ avec $x,y\geqslant 0$ sont dans $\bar{A}$. Pour cela, rien de tel qu'une suite. La suite $(\frac{1}{n},y)$ avec $y\geqslant 0$ est contenue dans A, et sa limite est clairement le point $(0,y)$. Nous en concluons que $(0,y)$ est un point de $\bar{A}$.
De la même façon, la suite $(x,\frac{1}{n})$ montre que le point $(x,0)$ est dans $\bar{A}$. Donc

$$\bar{A}=A\cup\{(x,0)\text{ tels que }x\geqslant 0\}\cup\{(0,y)\text{ tels que }y\geqslant 0\}. \tag{10.210}$$

En particulier, le point $(0,0)$ est dans $\bar{A}$.

(3) En ce qui concerne la frontière, nous utilisons la caractérisation $\partial A=\bar{A}\setminus\text{Int}(A)$. Étant donné que A est ouvert, $\text{Int}(A)=A$. Les points qui sont dans $\bar{A}$ et pas dans A sont les points avec $x=0$ ou $y=0$. Donc

$$\partial A=\{(x,0)\text{ tels que }x\geqslant 0\}\cup\{(0,y)\text{ tels que }y\geqslant 0\}. \tag{10.211}$$

L'adhérence peut aussi être trouvée en utilisant la proposition 10.81. Nous avons $A=]0,\infty[\times]0,\infty[$, et par conséquent, la fermeture de A est la produit des fermetures :

$$\bar{A}=[0,\infty[\times[0,\infty[. \tag{10.212}$$

Nous insistons sur le fait que la fermeture de $]0,\infty[$ n'est pas $[0,\infty]$. Ce dernier ensemble n'est pas une partie de $\mathbb{R}$. □

10.5.3 Continuité du produit de matrices

SUBSECooOAWAooFcyUfI

Nous avons introduit des normes sur $\mathbb{M}(n,\mathbb{K})$, entre autres la norme opérateur de la définition 10.55. Qui dit norme dit topologie. Il advient alors la question évidente : est-ce que des opérations aussi élémentaires que le produit de matrices sont continues pour ces topologies ?

Une façon simple de répondre à cela est d'introduire sur $\mathbb{M}(n,\mathbb{K})$ une nouvelle norme très simple : celle de $\mathbb{K}^n$. C'est la topologie par composante. Pour cette topologie, il est simple de voir que le produit matriciel est continu parce que les éléments de AB sont des polynômes en les éléments de A B. Ensuite il suffit d'invoquer l'équivalence de toutes les normes (théorème 10.49).

Voyons comment montrer cela de façon plus directe (bien que le raisonnement précédent soit une démonstration qui devrait déjà avoir convaincu les plus sceptiques). La preuve suivante va donc s'amuser à bien préciser les topologies et caractérisations utilisées.

LEMooRGNRooPovBQw

Lemme 10.83.
Si $\|.\|$ est une norme algébrique sur $\mathbb{M}(n,\mathbb{K})$ ($\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$) alors l'application

$$\begin{aligned}p\colon\mathbb{M}(n,\mathbb{K})\times\mathbb{M}(n,\mathbb{K})&\to\mathbb{M}(n,\mathbb{K})\\(A,B)&\mapsto AB\end{aligned} \tag{10.213}$$

est continue.

Démonstration. L'espace $\mathbb{M}(n,\mathbb{K})\times\mathbb{M}(n,\mathbb{K})$ est métrique (définition 7.226), donc la caractérisation séquentielle de la continuité (proposition 7.255) s'applique. Nous considérons donc une suite (A_k,B_k) dans $\mathbb{M}(n,\mathbb{K})\times\mathbb{M}(n,\mathbb{K})$ convergente vers (A,B), et nous allons prouver que $p(A_k,B_k)\to p(A,B)$.

Nous savons que la topologie sur $\mathbb{M}(n,\mathbb{K})\times\mathbb{M}(n,\mathbb{K})$ est la topologie produit (lemme 7.226) et que celle-ci donne la convergence composante par composante dès que nous avons convergence d'une suite ; c'est la proposition 7.65. Nous avons donc $A_k\xrightarrow{\mathbb{M}(n,\mathbb{K})}A$ et $B_k\xrightarrow{\mathbb{M}(n,\mathbb{K})}B$.

Voilà pour le contexte. Maintenant, la preuve de la continuité. Nous effectuons les majorations suivantes :

$$\begin{aligned}\|p(A_k,B_K)-AB\|&\leqslant\|p(A_k,B_k)-p(A_k,B)\|+\|p(A_k,B)-AB\| &\text{(10.214a)}\\&=\|A_kB_k-A_kB\|+\|A_kB-AB\| &\text{(10.214b)}\\&=\|A_k(B_k-B)\|+\|(A_k-A)B\| &\text{(10.214c)}\\&\leqslant\underbrace{\|A_k\|}_{\to\|A\|}\underbrace{\|B_k-B\|}_{\to 0}+\underbrace{\|A_k-A\|}_{\to 0}\|B\|. &\text{(10.214d)}\end{aligned}$$

□

10.5.4 Continuité de la composition

PROPooIPILooZMtYSP

Proposition 10.84 ([1]).
Soit un espace vectoriel normé E. L'application

$$\begin{aligned}\alpha\colon \operatorname{End}(E)^n &\to \operatorname{End}(E)\\ (A_1,\dots,A_n) &\mapsto A_1\circ\dots\circ A_n\end{aligned} \tag{10.215}$$

est une application multilinéaire continue.

Démonstration. Le fait que α soit multilinéaire est une simple vérification [43]. Pour la continuité, nous utilisons le lemme 10.69 qui donne

$$\|\alpha(A_1,\dots,A_n)\|\leqslant\|A_1\|\dots\|A_n\|. \tag{10.216}$$

Donc $\lambda=1$ fait fonctionner la proposition 8.6. □

10.6 Applications multilinéaires

Voir la définition 8.7 des applications multilinéaires.

DefKPBYeyG

Lemme-Définition 10.85.
Pour $T\in\mathcal{L}_k(E_1,\dots,E_k,F)$ nous posons

$$\|T\|=\sup_{\|v_i\|=1}\|T(v_1,\dots,v_k)\|_W. \tag{10.217}$$

Cela définit une norme.

PropUADlSMg

Proposition 10.86.
Soient des espaces vectoriels normés E_i et F. Une application n-linéaire [44]

$$T\colon E_1\times\dots\times E_n\to F \tag{10.218}$$

est est continue si et seulement si il existe un réel $L\geqslant 0$ tel que

$$\|T(x_1,\dots,x_n)\|_F\leqslant L\|x_1\|_{E_1}\cdots\|x_n\|_{E_n},\qquad\forall x_i\in E_i. \tag{10.219}$$

limitatezza

Démonstration. Pour simplifier l'exposition nous nous limitons au cas $n=2$ et nous notons $T(x,y)=x*y$

Supposons que l'inégalité (10.219) soit satisfaite.

$$\begin{aligned}\|x*y-x_0*y_0\|&=\|(x-x_0)*y-x_0*(y-y_0)\|\\ &\leqslant\|(x-x_0)*y\|+\|x_0*(y-y_0)\|\\ &\leqslant L\|x-x_0\|\|y\|+L\|x_0\|\|y-y_0\|.\end{aligned} \tag{10.220}$$

LimImplCont

Si $x\to x_0$ et $y\to y_0$ on voit que T est continue en passant à la limite aux deux côtés de l'inégalité (10.220).

Soit T continue en $(0,0)$. Évidemment [45] $0*0=0$, donc il existe $\delta>0$ tel que si $x\in B_{E_1}(0,\delta)$ et $y\in B_{E_2}(0,\delta)$ alors $\|x*y\|\leqslant 1$. En particulier si $(x,y)\in B_{E_1\times E_2}(0,\delta)$ nous sommes dans ce cas. Soient maintenant $x\in E_1\backslash\{0\}$ et $y\in E_2\backslash\{0\}$

$$x*y=\left(\frac{\|x\|}{\delta}\frac{\delta x}{\|x\|}\right)*\left(\frac{\|y\|}{\delta}\frac{\delta y}{\|y\|}\right)=\frac{\|x\|\|y\|}{\delta^2}\left(\frac{\delta x}{\|x\|}\right)*\left(\frac{\delta y}{\|y\|}\right). \tag{10.221}$$

43. Que je n'ai pas faite, alors faites-la.
44. Définition 8.1.
45. Dans la formule suivante, les trois zéros sont les zéros de trois espaces différents.

On remarque que $\delta x/\|x\|_m$ est dans la boule de rayon δ centrée en 0_m et que $\delta y/\|y\|_n$ est dans la boule de rayon δ centrée en 0_n. On conclut

$$x * y \leqslant \frac{\|x\|_m\|y\|_n}{\delta^2}.$$

Il faut prendre $L = 1/\delta^2$. □

PROPooOPCVooEWeWha

Proposition 10.87.
Si T est une application 2-linéaire, nous avons

$$\|T(u,v)\| \leqslant \|T\|\|u\|\|v\|. \tag{10.222}$$

EqYLnbRbC

Et nous notons que cette norme est uniquement définie pour les applications linéaires continues. Ce n'est pas très grave parce qu'alors nous définissons $\|T\| = \infty$ si T n'est pas continue. Cela pour retrouver le principe selon lequel on est continue si et seulement si on est borné.

isom_isom

Proposition 10.88.
On définit les fonctions

$$\begin{array}{rccc} \omega_g: & \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p) & \to & \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)), \\ \omega_d: & \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p) & \to & \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)), \end{array} \tag{10.223}$$

par

$$\omega_g(T)(x) = T(x, \ \cdot\ ), \qquad \forall x \in \mathbb{R}^m,$$

et

$$\omega_d(T)(y) = T(\ \cdot\ , y), \qquad \forall y \in \mathbb{R}^n.$$

Les fonctions ω_g et ω_d sont des isomorphismes qui préservent les normes.

10.7 Séries

SECooYCQBooSZNXhd

La notion de somme sur un ensemble infini est donnée en 10.112, voir aussi le thème 53 pour plus de notions de sommes finies et infinies.

DefGFHAaOL

Définition 10.89.
*Soit (a_k) une suite dans un espace vectoriel normé $(V, \|.\|)$. La suite des **sommes partielles** associée est la suite (s_k) définie par*

$$s_k = \sum_{i=0}^{k} a_i \tag{10.224}$$

*La **série** associée est la limite des sommes partielles*

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k \tag{10.225}$$

si elle existe.

*Si une telle limite existe nous disons que $\sum_{k=0}^{\infty} a_k$ **converge** dans V. Si la limite de la suite des sommes partielles n'existe pas nous disons que la série **diverge**.*

Remarque 10.90.
Si la limite de la suite des sommes partielles n'existe pas dans V, alors elle peut parfois exister dans des extensions de V. Par exemple une série de rationnels convergeant vers $\sqrt{2}$ dans $\mathbb{R}$ ne converge pas dans $\mathbb{Q}$. Autre exemple : avec une bonne topologie sur $\bar{\mathbb{R}}$, une série peut ne pas converger dans $\mathbb{R}$ mais converger vers $\pm\infty$ dans $\bar{\mathbb{R}}$.

Dans le cas des espaces de fonctions, nous avons une norme importante : la norme uniforme définie par $\|f\|_\infty = \sup\{f(x)\}$ où le supremum est pris sur l'ensemble de définition de f.

LEMooHUZEooSyPipb

Lemme 10.91.
Soit une suite (a_k) dans un espace métrique complet[46] *dont la série converge.*

ITEMooPFSQooDhKFGL

(1) Pour tout N nous avons

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{N} a_k + \sum_{k=N+1}^{\infty} a_k. \tag{10.226}$$

ITEMooQNHMooUPjupB

(2) La suite des queues de série converge vers 0, c'est-à-dire que

$$\lim_{N\to\infty} \sum_{k=N}^{\infty} a_k = 0. \tag{10.227}$$

Démonstration. Voici un petit calcul :

SUBEQooZRSHooSjsmK

$$\lim_{n\to\infty} \sum_{k=0}^{n} a_k = \lim_{n\to\infty} \Big(\sum_{k=0}^{N} a_k + \sum_{k=N+1}^{n} a_k \Big) \tag{10.228a}$$

SUBEQooTLVKooQfYXam

$$= \lim_{n\to\infty} \sum_{k=0}^{N} a_k + \lim_{n\to\infty} \sum_{k=N+1}^{n} a_k \tag{10.228b}$$

$$= \sum_{k=0}^{N} a_k + \sum_{k=N+1}^{\infty} a_k. \tag{10.228c}$$

Justifications :

— Pour (10.228a). Pour chaque n, la somme est finie et nous pouvons la décomposer. Si vous voulez vraiment couper les cheveux en quatre, vous devez fixer un ϵ, et un n de telle sorte à avoir $n > N$, parce que N est fixé dans l'énoncé du lemme.

— Pour (10.228b). Nous sommes dans un cas $\lim_{n\to\infty}(u_n + v_n)$ où (u_n) est constante et où $(u_n + v_n)$ converge. Nous pouvons donc permuter limite et somme[47].

Voilà que (1) est prouvé.

Nous écrivons $s_n = \sum_{k=0}^{n} a_k$; l'hypothèse est que la suite (s_n) est une suite convergente dans un espace métrique. Elle est donc de Cauchy par la proposition 7.270.

Soit $\epsilon > 0$. Il existe $N \in \mathbb{N}$ tel que pour tout $p, q > N$, nous ayons $\|s_p - s_q\| \leqslant \epsilon$. Soit $p > N$. Pour tout $n \geqslant 0$ nous avons

$$\epsilon > \|s_{p+n} - s_{p+1}\| = \| \sum_{k=p}^{p+n} a_k \|. \tag{10.229}$$

En prenant la limite $n \to \infty$ nous avons

$$\| \sum_{k=p}^{\infty} a_k \| \leqslant \epsilon. \tag{10.230}$$

Nous avons donc démontré qu'il existe N tel que si $p > N$, alors $\| \sum_{k=p}^{\infty} a_k \| \leqslant \epsilon$. Cela signifie exactement que $\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = 0$. □

10.7.1 Les trois types de convergence

Trois notions de convergence à ne pas confondre :

(1) La convergence absolue,

(2) la convergence normale. C'est la même que la convergence absolue, mais dans le cas particulier d'un espace de fonctions muni de la norme uniforme.

46. Définition 7.263.
47. Pour rappel, la proposition 9.28 demande la convergence des deux suites pour fonctionner.

(3) la convergence uniforme.

Voici les définitions.

DefVFUIXwU

Définition 10.92 (Convergence absolue)**.**
*Nous disons que la série $\sum_{n=0}^{\infty} a_n$ dans l'espace vectoriel normé V **converge absolument** si la série $\sum_{n=0}^{\infty} \|a_n\|$ converge dans $\mathbb{R}$.*

DefVBrJUxo

Définition 10.93 (Convergence normale)**.**
*Une série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge **normalement** si la série de nombres $\sum_n \|u_n\|_\infty$ converge. C'est-à-dire si la série converge absolument pour la norme $\|f\|_\infty$.*

DEFooPABSooPMXMOV

Définition 10.94 (Convergence uniforme)**.**
*La somme $\sum_n f_n$ **converge uniformément** vers la fonction F si la suite des sommes partielles converge uniformément, c'est-à-dire si*

$$\lim_{N\to\infty} \| \sum_{n=1}^{N} f_n - F\|_\infty = 0. \tag{10.231}$$

EqLNCJooVCTiJw

LEMooJZTBooIopLok

Lemme 10.95.
Soient un espace topologique X, un espace vectoriel normé, et une suite de fonctions $f_n\colon X \to V$. Si la série $\sum_n f_n$ converge normalement, alors elle converge uniformément.

PropAKCusNM

Proposition 10.96.
Une série absolument convergente dans un espace de Banach[48] y converge au sens usuel.

Démonstration. Soit (a_k) une suite dans un espace vectoriel normé complet dont la série converge absolument. Nous allons montrer que la suite des sommes partielles est de Cauchy. Cela suffira à montrer sa convergence par hypothèse de complétude.

Nous avons

$$\|s_p - s_l\| = \| \sum_{k=l+1}^{p} a_k\| \leqslant \sum_{k=l+1}^{p} \|a_k\| = \bar{s}_p - \bar{s}_l \tag{10.232}$$

où $\bar{s}_n = \sum_{k=0}^{n} \|a_k\|$ est la suite des sommes partielles de la série des normes (qui converge). Vu que la suite $(\bar{s}_n)$ converge dans $\mathbb{R}$, elle y est de Cauchy par la proposition 1.406. Donc il existe un N tel que $p, l > N$ implique

$$\|s_p - s_l\| = \bar{s}_p - \bar{s}_l \leqslant \epsilon. \tag{10.233}$$

Cela signifie que (s_n) est une suite de Cauchy et donc convergente. □

Exemple 10.97 (Si l'espace n'est pas complet[1])**.**
Dans un espace qui n'est pas complet, il est possible de construire un série qui converge absolument sans converger au sens usuel.

Nous allons trouver dans $\mathbb{Q}$ une série qui converge simplement vers $\sqrt{2}$ (et donc ne converge pas dans $\mathbb{Q}$) mais absolument vers 4.

La base est que si $A, B \in \mathbb{Q}$ avec $A < B$ il est possible de résoudre

$$\begin{cases} r_1 + r_2 = A & \text{(10.234a)} \\ |r_1| + |r_2| = B & \text{(10.234b)} \end{cases}$$

pour $r_1, r_2 \in \mathbb{Q}$. Ce n'est pas très compliqué : la solution est $r_1 = (A + B)/2$ et $r_2 = (A - B)/2$.

Nous considérons l'espace $\mathbb{Q}$ qui n'est pas complet dans $\mathbb{R}$. Soit une série (a_k) dans $\mathbb{Q}$ qui converge vers $\sqrt{2}$ (convergence dans $\mathbb{R}$) nous nommons (s_k) la suite des ses sommes partielles. Soit aussi la suite (b_k) qui converge vers 4 (zéro serait encore plus facile mais bon, juste pour faire un peu de généralité).

48. Un espace vectoriel normé complet. Typiquement $\mathbb{R}$.

Nous supposons que $a_k < b_k$ pour tout k et que les deux suites sont constituées de rationnels positifs. Nous nommons (s_k) et (s'_k) les sommes partielles. En particulier $s_n < s'_n$ et ce sont des suites croissantes.

Nous savons comment trouver $r_1, r_2 \in \mathbb{Q}$ tels que $r_1 + r_2 = s_1$ et $|r_1| + |r_2| = s'_1$. Par récurrence, si nous savons $r_1, \dots, r_k$ tels que

$$\begin{cases} r_1 + \dots + r_k = s_n & (10.235a) \\ |r_1| + \dots + |r_k| = s'_n & (10.235b) \end{cases}$$

(avec, soit dit en passant $k = 2n$), alors nous pouvons trouver des rationnels r_{k+1}, r_{k+2} tels que

$$\begin{cases} r_1 + \dots + r_k + r_{k+1} + r_{k+2} = s_{n+1} & (10.236a) \\ |r_1| + \dots + |r_k| + |r_{k+1}| + |r_{k+2}| = s'_{n+1}, & (10.236b) \end{cases}$$

en effet il s'agit de résoudre

$$\begin{cases} r_{k+1} + r_{k+2} = s_{n+1} - r_1 - \dots - r_k = s_{n+1} - s_n > 0 & (10.237a) \\ |r_{k+1}| + |r_{k+2}| = s'_{n+1} - |r_1| - \dots - |r_k| = s'_{n+1} - s'_n > 0. & (10.237b) \end{cases}$$

Cela se résout comme ci-dessus. Au final nous pouvons construire une suite (r_k) dans $\mathbb{Q}$ telle que

$$\sum_{k=0}^{2n} r_k = s_n \tag{10.238}$$

et

$$\sum_{k=0}^{2n} |r_k| = s'_n. \tag{10.239}$$

△

Remarque 10.98.
Nous savons que sur les espaces vectoriels de dimension finie toutes les normes sont équivalentes (théorème 10.46). La notion de convergence de série ne dépend alors pas du choix de la norme. Il n'en est pas de même sur les espaces de dimension infinie. Une série peut converger pour une norme mais pas pour une autre.

Lorsque nous verrons la convergence de séries, nous verrons que la convergence normale est la convergence absolue pour la norme uniforme.

LemCAIPooPMNbXg

Lemme 10.99.
Si E et F sont des espaces de Banach[49]*, l'espace $\mathcal{L}(E, F)$ est également de Banach.*

Démonstration. Soit (u_n) une suite de Cauchy dans $\mathcal{L}(E, F)$; si $x \in E$ il existe N tel que si $l, m > N$ alors $\|u_l - u_m\| < \epsilon$, c'est-à-dire que pour tout $\|x\| = 1$ on a $\|u_l(x) - u_n(x)\| < \epsilon$. Cela signifie que $u_n(x)$ est une suite de Cauchy dans l'espace complet F. Cette suite converge et nous pouvons définir l'application $u\colon E \to F$ par

$$u(x) = \lim_{n\to\infty} u_n(x). \tag{10.240}$$

Il suffit maintenant de prouver que u est linéaire, ce qui est une conséquence directe de la linéarité de la limite :

$$u(\alpha x + \beta y) = \lim_{n\to\infty} \big(\alpha u_n(x) + \beta u_n(y)\big). \tag{10.241}$$

□

PROPooYDFUooTGnYQg

Proposition 10.100.
Si une série converge dans un espace complet, la norme de son terme général converge vers 0 : si $\sum_{n=0}^{\infty} a_n$ converge, alors

$$\|a_n\| \to 0. \tag{10.242}$$

49. Je crois qu'il ne faut pas que E soit complet.

Démonstration. Soit une suite (a_n) dont la série converge vers s. Soit $\epsilon > 0$. La suite des sommes partielles (s_n) est de Cauchy et converge vers s : $s_n \to s$. En particulier il existe un N tel que si $n > N$, nous avons $\|s_n - s_{n-1}\| < \epsilon$. Pour de telles valeurs de n nous avons :

$$\|a_n\| = \|s_n - s_{n-1}\| \leqslant \epsilon. \tag{10.243}$$

Cela prouve que $a_n \to 0$. □

Dans le même ordre d'idée nous avons la convergence des queues de suites.

LEMooFUCOooCOqLRj

Lemme 10.101 ([1])**.**
Soit une suite $a\colon \mathbb{N} \to V$ où V est un espace vectoriel normé. Si $\sum_{k=0}^{\infty} a_k$ converge, alors

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = 0. \tag{10.244}$$

Démonstration. Ne nous en voulez pas si on décale l'énoncé de 1 : nous allons prouver que $\sum_{k=n+1}^{\infty} a_k \to 0$. Nous nommons (s_n) la suite des sommes partielles de a : $s_n = \sum_{k=0}^{n} a_k$. Soit n fixé dans $\mathbb{N}$; pour tout $N > n$ nous avons

$$\sum_{k=0}^{N} a_k = s_n + \sum_{k=n+1}^{N} a_k \tag{10.245}$$

En prenant la limite $N \to \infty$ nous trouvons

$$\sum_{k=0}^{\infty} a_k = s_n + \lim_{N\to\infty} \sum_{k=n+1}^{N} a_k = s_n + \sum_{k=n+1}^{\infty} a_k. \tag{10.246}$$

Cela étant valable pour tout n, nous prenons la limite $n \to \infty$. Par définition $s_n \to \sum_{k=0}^{\infty} a_k$; pour le reste

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_k + \lim_{n\to\infty} \sum_{k=n+1}^{\infty} a_k. \tag{10.247}$$

La dernière somme est donc nulle et nous avons prouvé le lemme. □

PROPooUEBWooUQBQvP

Proposition 10.102.
Si la série converge alors la somme est associative : $\sum_k (a_k + b_k) = \sum_k a_k + \sum_k b_k$.

Démonstration. Supposons que $\sum_k a_k$ et $\sum_k b_k$ convergent tous deux. Alors nous avons pour tout N :

$$\sum_{k=0}^{N} (a_k + b_k) = \sum_{k=0}^{N} a_k + \sum_{k=0}^{N} b_k. \tag{10.248}$$

Mais si deux limites existent alors la somme commute avec la limite. C'est le cas pour la limite $N \to \infty$, donc

$$\lim_{N\to\infty} \sum_{k=1}^{N} (a_k + b_k) = \lim_{N\to\infty} \sum_{k=0}^{N} a_k + \lim_{N\to\infty} \sum_{k=0}^{N} b_k = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k. \tag{10.249}$$

□

¡¡ Avertissement/question au lecteur !! 10.103
Je n'ai pas vérifié la proposition suivante. Écrivez-moi pour m'envoyer une preuve, ou pour me dire qu'elle n'est pas correcte. En tout cas ne la prenez pas comme vraie trop facilement.

PROPooQCOBooAMnljj

Proposition 10.104.
Si $a_n \geqslant 0$ et $b_n \geqslant 0$ alors

$$\sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} (a_n + b_n). \tag{10.250}$$

10.7.2 Somme téléscopique

PROPooQLOUooTDWfFF

Proposition 10.105 ([1]).
Soit un espace vectoriel normé E. Soit une série convergente $\sum_{k=0}^{\infty} a_k$. Alors

$$\sum_{k=0}^{\infty}(a_k - a_{k+1}) = a_0. \tag{10.251}$$

Démonstration. Pour chaque N nous avons $\sum_{k=0}^{N}(a_k - a_{k+1}) = a_0 - a_{N+1}$. Donc

$$\sum_{k=0}^{\infty}(a_k - a_{k+1}) = \lim_{N\to\infty}(a_0 - a_{N+1}). \tag{10.252}$$

Vu que le terme général tend vers zéro (proposition 10.100), nous avons le résultat annoncé. □

10.7.3 Distributivité

Nous allons parler d'exponentielle de matrice. Avant cela, quelques propriétés qui sont valables sur des algèbres normées. Le principal exemple que nous avons en tête est $\mathbb{A} = \mathbb{M}(n, \mathbb{C})$.

PROPooMZZQooEhQsgQ

Proposition 10.106 (Distributivité de la somme infinie[1]).
Soit une algèbre normée $\mathbb{A}$. Soient une suite d'éléments $A_k \in \mathbb{A}$ et un élément B. Nous supposons que la somme $\sum_{k=0}^{\infty} A_k$ converge. Alors

$$B\sum_k A_k = \sum_k (BA_k). \tag{10.253}$$

Démonstration. Soit $N \in \mathbb{N}$. Nous avons :

$$\| \sum_{k=0}^{N} BA_k - B\sum_{k=0}^{\infty} A_k \| = \| B \sum_{k=N+1}^{\infty} A_k \| \tag{10.254a}$$

SUBEQooDTNAooWpXOKP

$$\leqslant \|B\| \| \sum_{k=N+1}^{\infty} A_k \| \tag{10.254b}$$

SUBEQooJPSJooAqXtOJ

Justifications :

— Pour (10.254a). Linéarité du produit matriciel.

— Pour (10.254b). La norme est une norme d'algèbre [50].

À droite, la limite $N \to \infty$ donne zéro car $\|B\|$ est un simple nombre, et $\| \sum_{k=N+1}^{\infty} A_k \|$ est une queue de suite convergente par hypothèse.

Nous avons donc bien convergence

$$\lim_{N\to\infty} \sum_{k=0}^{N} BA_k = B \sum_{k=0}^{\infty} A_k. \tag{10.255}$$

□

10.7.4 Produit de Cauchy

PROPooFMEXooCNjdhS

Proposition 10.107 (Produit de Cauchy dans une algèbre normée[1]).
Soient une algèbre normée $\mathbb{A}$, un élément $A \in \mathbb{A}$, ainsi que des séries convergentes $\sum_{k=0}^{\infty} a_k A^k$ et $\sum_{l=0}^{\infty} b_l A^l$. Alors

$$\left(\sum_k a_k A^k \right) \left(\sum_l b_l A^l \right) = \sum_{n=0}^{\infty} \Big(\sum_{m=0}^{n} a_m b_{n-m} \Big) A^n. \tag{10.256}$$

50. Définition 10.63. Pour rappel, la norme opérateur en est une par le lemme 10.69.

Démonstration. Un calcul :

$$\Big(\sum_k a_k A^k\Big)\Big(\sum_l b_l A^l\Big) = \sum_k \big(\sum_l b_l A^l\big) a_k A^k \tag{10.257a}$$
$$= \sum_k \big(\sum_l b_l a_k A^{l+k}\big) \tag{10.257b}$$
$$= \lim_{K\to\infty} \sum_{k=0}^{K} \big(\lim_{L\to\infty} \sum_{l=0}^{L} b_l a_k A^{k+l}\big) \tag{10.257c}$$
$$= \lim_{K\to\infty} \lim_{L\to\infty} \sum_{k=0}^{K} \sum_{l=0}^{L} b_l a_k A^{k+l} \tag{10.257d}$$
$$= \lim_{K\to\infty} \lim_{L\to\infty} \sum_{n=0}^{K+L} \sum_{m=0}^{n} a_m b_{n-m} A^n \tag{10.257e}$$
$$= \lim_{K\to\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{n} a_m b_{n-m} A^m \tag{10.257f}$$
$$= \sum_{n=0}^{\infty} \sum_{m=0}^{n} a_m b_{n-m} A^m \tag{10.257g}$$

Justifications :

— Pour (10.257a), la proposition 10.106 nous permet d'entrer l'élément $\sum_l b_l A^l \in \mathbb{A}$ dans la somme sur k.

— Pour (10.257b), c'est la même chose.

— Pour (10.257d), la somme sur k étant finie (pour chaque K), elle commute avec la limite sur L.

— Pour (10.257e), c'est une manipulation de sommes finies. On regroupe les termes selon les puissances de A.

— Pour (10.257f), c'est effectuer la limite sur L pour K fixé.

— Pour (10.257g), l'expression dans la limite sur K ne dépend pas de K. Donc nous pouvons simplement supprimer la limite.

□

10.8 Sommes de familles infinies

10.8.1 Convergence commutative

Définition 10.108.
*Soit x_k une suite dans un espace vectoriel normé E. Nous disons que la suite **converge commutativement** vers $x \in E$ si $\lim_{n\to\infty} \|x_n - x\| = 0$ et si pour toute bijection $\tau : \mathbb{N} \to \mathbb{N}$ nous avons aussi*

$$\lim_{n\to\infty} \|x_{\tau(n)} - x\| = 0. \tag{10.258}$$

La notion de convergence commutative est surtout intéressante pour les séries. La somme

$$\sum_{k=0}^{\infty} x_k \tag{10.259}$$

converge commutativement vers x si $\lim_{N\to\infty} \|x - \sum_{k=0}^{N} x_k\| = 0$ et si pour toute bijection $\tau : \mathbb{N} \to \mathbb{N}$ nous avons

$$\lim_{N\to\infty} \|x - \sum_{k=0}^{N} x_{\tau(k)}\| = 0. \tag{10.260}$$

PopriXWvIY

Proposition 10.109.
Soit $(a_i)_{i\in\mathbb{N}}$ une suite absolument convergente[51] *dans $\mathbb{C}$. Alors elle converge commutativement.*

Démonstration. Soit $\epsilon > 0$. Nous posons $\sum_{i=0}^{\infty} a_i = a$ et nous considérons N tel que

$$|\sum_{i=0}^{N} a_i - a| < \epsilon. \tag{10.261}$$

Étant donné que la série des $|a_i|$ converge, il existe N_1 tel que pour tout $p, q > N_1$ nous ayons $\sum_{i=p}^{q} |a_i| < \epsilon$. Nous considérons maintenant une bijection $\tau\colon \mathbb{N} \to \mathbb{N}$. Prouvons que la série $\sum_{i=0}^{\infty} |a_{\tau(i)}|$ converge. Nous choisissons M de telle sorte que pour tout $n > M$, $\tau(n) > N_1$. Si s_k est la somme partielle de la suite $(a_{\tau(i)})_{i\in\mathbb{N}}$ et si $M < p < q$ nous avons

$$|s_q - s_p| = |\sum_{i=p}^{q} a_{\tau(i)}| \leqslant \sum_{i=p}^{q} |a_{\tau(i)}| < \epsilon. \tag{10.262}$$

Cela montre que (s_k) est une suite de Cauchy. Elle est alors convergente et nous en déduisons que la série

$$\sum_{i=0}^{\infty} a_{\tau(i)} \tag{10.263}$$

converge. Nous devons montrer à présent qu'elle converge vers la même limite que la somme « usuelle » $\lim_{N\to\infty} \sum_{i=0}^{N} a_i$.

Soit $n > \max\{M, N\}$. Alors

$$\sum_{k=0}^{n} a_{\tau(k)} - \sum_{k=0}^{n} a_k = \sum_{k=0}^{M} a_{\tau(k)} - \sum_{k=0}^{N} a_k + \underbrace{\sum_{M+1}^{n} a_{\tau(k)}}_{<\epsilon} - \underbrace{\sum_{k=N+1}^{n} a_k}_{<\epsilon}. \tag{10.264}$$

Par construction les deux derniers termes sont plus petits que ϵ parce que M et N sont les constantes de Cauchy pour les séries $\sum a_{\tau(i)}$ et $\sum a_i$. Afin de traiter les deux premiers termes, quitte à redéfinir M, nous supposons que $\{1, \dots, N\} \subset \tau\{1, \dots, M\}$; par conséquent tous les a_i avec $i < N$ sont atteints par les $a_{\tau(i)}$ avec $i < M$. Dans ce cas, les termes qui restent dans la différence

$$\sum_{k=0} a_{\tau(k)} - \sum_{k=0}^{N} a_k \tag{10.265}$$

sont des a_k avec $k > N$. Cette différence est donc en valeur absolue plus petite que ϵ, et nous avons en fin de compte que

$$\left|\sum_{k=0}^{n} a_{\tau(k)} - \sum_{k=0}^{n} a_k\right| < \epsilon. \tag{10.266}$$

□

PROPooVVHKooZjXQiP

Proposition 10.110 ([1])**.**
Si (a_k) est une suite réelle avec $a_k \geqslant 0$ et si $\sum_{k=0}^{\infty} a_k = \infty$, alors $\sum_{k=0}^{\infty} a_{\tau(k)} = \infty$ pour toute permutation τ de $\mathbb{N}$.

Démonstration. Soient $m > 0$. Il existe $N > 0$ tel que si $n \geqslant N$, alors $\sum_{k=0}^{n} a_k > m$. Vu que τ est une permutation, il existe N_1 tel que $\tau\{0, \dots, N_1\}$ contient $\{0, \dots, N\}$.

Alors pour tout $n \geqslant N_1$ nous avons

$$\sum_{k=0}^{n} a_{\tau(k)} \geqslant \sum_{k=0}^{N} a_k > m. \tag{10.267}$$

□

51. Définition 10.92.

PropyFJXpr

Proposition 10.111 (Théorème de réarrangement de Riemann[332, 333, 334, 1])**.**
Soit $\sum_{k=0}^{\infty} a_k$ une série réelle qui converge mais qui ne converge pas absolument. Alors pour tout $b \in \mathbb{R}$, il existe une bijection $\tau\colon \mathbb{N} \to \mathbb{N}$ telle que $\sum_{i=0}^{\infty} a_{\tau(i)} = b$.

Démonstration. Nous posons

$$N^+ = \{n \in \mathbb{N} \text{ tel que } a_n \geqslant 0\} \tag{10.268a}$$
$$N^- = \{n \in \mathbb{N} \text{ tel que } a_n < 0\} \tag{10.268b}$$

(1) <u>N^+ **ou** N^- **est infini**</u>
Vu que $N^+ \cup N^- = \mathbb{N}$, au moins un des deux est infini (lemme 1.1281.128), ce qui n'est pas possible.

(2) <u>N^- **est infini**</u>
Supposons que N^- est fini. Posons $n_0 = \max(N^-)$. Pour tout $k > n_0$ nous avons $a_k \geqslant 0$. Soit $N > n_0$. Nous avons

$$\sum_{k=0}^{N} a_k = \sum_{k=0}^{n_0} a_k + \sum_{n_0+1}^{N} a_k \tag{10.269a}$$
$$= \sum_{k=0}^{n_0} a_k - \sum_{k=0}^{n_0} |a_k| + \sum_{k=0}^{n_0} |a_k| + \sum_{k=n_0+1}^{N} |a_k| \tag{10.269b}$$
$$= \sum_{k=0}^{n_0} \big(a_k - |a_k|\big) + \sum_{k=0}^{N} |a_k|. \tag{10.269c}$$

En prenant la limite $N \to \infty$, à gauche nous avons ∞ par hypothèse, alors que nous avons un nombre fini à droite (aussi par hypothèse). Contradiction, nous en concluons que N^- est infini.

(3) <u>N^+ **est infini**</u> Même type de raisonnement [52].

(4) <u>**Construction de $\tau\colon \mathbb{N} \to \mathbb{N}$**</u> ITEMooKATRooVrnUBv
Nous allons construire par récurrence. D'abord

$$T_0 = \{l \in \mathbb{N} \text{ tel que } a_l \geqslant 0\} = N^+ \tag{10.270a}$$
$$S_0 = \{l \in \mathbb{N} \text{ tel que } a_l < 0\} = N^-. \tag{10.270b}$$

Si $\sum_{k=0}^{n} a_{\tau(k)} \leqslant b$, alors nous posons
— $\tau(n+1) = \min(T_n) \in N^+$
— $T_{n+1} = T_n \backslash \tau(n+1)$
— $S_{n+1} = S_n$
Si $\sum_{k=0}^{n} a_{\tau(k)} > b$, alors nous posons
— $\tau(n+1) = \min(S_n) \in N^-$
— $S_{n+1} = S_n \backslash \tau(n+1)$
— $T_{n+1} = T_n$.
Les parties T_n et S_n sont les parties de $\mathbb{N}$ qui restent à attribuer lorsque $\tau(n)$ est fixé.
Nous prouvons maintenant quelques propriétés qui disent que τ prends les éléments de N^+ et N^- dans l'ordre sans en oublier.

(5) <u>$T_n = T_0 \backslash \tau\{0, \dots, n\}$</u> ITEMooLBLUooYypSZd
Cela se prouve par récurrence. Si $T_n = T_0 \backslash \{\tau(1), \dots, \tau(n)\}$, il y a deux possibilités pour T_{n+1}. Soit $\tau(n+1) \in N^+$, soit $\tau(n+1) \in N^-$. Dans le premier cas, T_{n+1} est défini comme étant $T_n \backslash \{\tau(n+1)\}$. Dans le second cas, $\tau(n+1) \notin T_n$, et donc $T_{n+1} = T_n = T_n \backslash \tau(n+1)$.

52. Je n'ai pas vérifié. Alors soyez prudent.

(6) $S_n = S_0 \backslash \tau\{0, \dots, n\}$
Même raisonnement.

(7) $k_1 < k_2 \in N^+$ **avec** $k_2 \in \tau(\mathbb{N})$ **alors** $k_1 \in \tau(\mathbb{N})$
Soient $k_1 < k_2$ dans N^+. Nous supposons que $k_2 = \tau(n+1)$ pour un certain n. Donc $k_2 = \min(T_n)$, ce qui signifie que k_1 n'est pas dans T_n. Or $T_n = T_0 \backslash \{\tau(0), \dots, \tau(n)\}$. Nous en déduisons que $k_1 \in \{\tau(0), \dots, \tau(n)\}$.

(8) $k_1 < k_2 \in N^-$ **avec** $k_2 \in \tau(\mathbb{N})$ **alors** $k_1 \in \tau(\mathbb{N})$
Même raisonnement.

(9) τ **est injective** Supposons que $\tau(k) = \tau(l)$ avec $k < l$. Vu que $S_n \subset S_0$ et $T_n \subset T_0$ et $S_0 \cap T_0 = \varnothing$, nous devons avoir tous les deux dans T_0 ou tous les deux S_0.
Supposons pour fixer les idées que $\tau(k)$ et $\tau(l)$ sont tous les deux dans T_0. Nous avons $\tau(k) \in T_{k-1}$, mais, par construction $\tau(k) \notin T_k$. Vu que $T_l \subset T_k$, il est impossible d'avoir $\tau(k) = \tau(l) \in T_l$.
La même chose tient si $\tau(k)$ et $(\tau(l))$ sont dans S_0.

(10) τ **est surjective**
Nous supposons que τ n'est pas surjective. Supposons que n_0 n'est pas dans l'image de τ, et supposons que $n_0 \in T_0$. Alors aucun élément de N^+ plus grand que n_0 n'est dans $\tau(\mathbb{N})$ par le point (5).
Soit $N = \max\{k \in \mathbb{N} \text{ tel que } \tau(k) \in N^+\}$. Ça existe et c'est fini parce que $\tau(\mathbb{N})$ ne prend aucun élément de N^+ au-delà de n_0. Donc pour tout $k \geqslant N+1$ nous avons $a_{\tau(k)} < 0$. Donc

$$\sum_{k=0}^{\infty} a_{\tau(k)} = \sum_{k=0}^{N} a_{\tau(k)} + \sum_{k=N+1}^{\infty} a_{\tau(k)} \tag{10.271a}$$

$$= \sum_{k=0}^{N} a_{\tau(k)} + \sum_{k=0}^{N} |a_{\tau(k)}| - \sum_{k=0}^{N} |a_{\tau(k)}| - \sum_{k=N+1}^{\infty} |a_{\tau(k)}| \tag{10.271b}$$

$$= \sum_{k=0}^{N} \Big(a_{\tau(k)} + |a_{\tau(k)}| \Big) - \sum_{k=0}^{\infty} |a_{\tau(k)}|. \tag{10.271c}$$

Par la proposition 10.110, la dernière somme diverge parce que $\sum_{k=0}^{\infty} |a_{\tau(k)}| = \infty$. Nous venons de prouver que $\sum_{k=0}^{\infty} a_{\tau(k)} = -\infty$.
Prenons $K > N$. Nous devons avoir

$$\sum_{k=0}^{K} a_{\tau(k)} \geqslant b \tag{10.272}$$

parce que sinon on aurait $\tau(K+1) \in N^+$. Vu qu'à partir de K les sommes partielles sont décroissantes et bornées vers le bas, nous en déduisons la convergence de $\sum_{k=0}^{\infty} a_{\tau(k)}$. Contradiction.

(11) **La somme converge vers** b

(11a) **Trouver un** N **assez grand**
Il est temps de montrer que $\sum_{k=0}^{\infty} a_{\tau(k)} = b$. Soit $\epsilon > 0$. Vu que $\sum_{k=0}^{\infty} a_k$ converge, il existe $n_1 \in \mathbb{N}$ tel que $|a_n| < \epsilon$ pour tout $n \geqslant n_1$ (proposition 10.100).
Vu que τ est bijective, la partie $\{n \in \mathbb{N} \text{ tel que } \tau(n) \leqslant n_1\}$ est finie, et même de cardinal $n_1 + 1$ parce qu'elle est en bijection avec $\{0, \dots, n_1\}$. Nous posons

$$n_0 = \max\{k \in \mathbb{N} \text{ tel que } \tau(k) \leqslant n_1\}. \tag{10.273}$$

Si $n > n_0$, alors $\tau(n) > n_1$ et alors $|a_{\tau(n)}| < \epsilon$. En posant

$$N_0 = \min\{k > n_0 \text{ tel que } \tau(k) \in N^+\} \tag{10.274a}$$

$$N = \min\{k > N_0 \text{ tel que } \tau(k) \in N^-\} - 1, \tag{10.274b}$$

nous avons un $N > n_0$ tel que

$$\begin{cases} \tau(N) \in N^+ & (10.275a) \\ \tau(N+1) \in N^- & (10.275b) \\ |a_{\tau(n)}| < \epsilon, \forall n \geqslant N. & (10.275c) \end{cases}$$

(11b) **Travail sur les sommes partielles**
Nous posons[53] $S_n = \sum_{k=0}^{n} a_{\tau(k)}$. Vu que $\tau(N) \in N^+$, nous avons $S_{N-1} \leqslant b$. Vu que $\tau(N+1) \in N^-$, nous avons $S_N > b$. Or

$$|S_n - S_{n-1}| = |a_{\tau(N)} < \epsilon|. \tag{10.276}$$

Nous avons donc les inégalités suivantes :

$$\begin{cases} |S_N - S_{N-1}| < \epsilon & (10.277a) \\ S_{N-1} \leqslant b & (10.277b) \\ S_N > b. & (10.277c) \end{cases}$$

(11c) **Si $S_n > b + \epsilon$, alors contradiction** Soit $n > N$, et supposons que $S_n > b + \epsilon$. Vu que $|S_n - S_{n-1}| < \epsilon$, nous avons encore $S_{n-1} > b$, et donc, par construction de τ (point (4)), nous avons $S_n < S_{n-1}$. Au final,

$$b + \epsilon < S_n < S_{n-1} < S_{N-1}. \tag{10.278}$$

Et là, il y a une contradiction parce que $S_{N-1} \leqslant b$.
Nous en déduisons que nous avons toujours $S_n \leqslant b + \epsilon$.

(11d) **Idem : $S_n \geqslant b - \epsilon$**
Cela se prouve de la même façon.

Au final nous avons prouvé que pour tout $n > N$, $|S_n - b| \leqslant \epsilon$, ce qui montre bien que $S_n \to b$.

□

Pour une revue des définitions de sommes dans le cas de $\{0, \dots, n\}$, des ensembles finis quelconques, voir le thème 53. Maintenant nous donnons la définition pour une somme sur un ensemble infini.

DefIkoheE

Définition 10.112.
*Si $\{v_i\}_{i \in I}$ est une famille de vecteurs dans un espace vectoriel normé indexée par un ensemble quelconque I. Nous disons que cette famille est **sommable** de somme v si pour tout $\epsilon > 0$, il existe un J_0 fini dans I tel que pour tout ensemble fini K tel que $J_0 \subset K$ nous avons*[54]

$$\| \sum_{j \in K} v_j - v \| < \epsilon. \tag{10.279}$$

Si tel est le cas, nous notons[55]

$$\sum_{i \in I} v_i = v. \tag{10.280}$$

Dans cette définition, il faut comprendre que J_0 est l'ensemble minimum de termes qu'il faut prendre pour être ϵ-proche de v. Ensuite, K est là pour dire qu'en prenant plus de termes, on ne s'éloignera pas tellement.

53. Je sais, la notation S_n est déjà prise plus haut. Mais nous n'en avons plus besoin.
54. La somme sur la partie finie K est la définition 1.309.
55. Attention que pour définir ça, il faut prouver l'unicité ; je n'ai pas vérifié. Écrivez-moi si vous avez une preuve.

LEMooHFNXooFHfwzf

Lemme 10.113.
Soit une famille sommable $(v_i)_{i\in I}$ dans un espace vectoriel normé sur $\mathbb{C}$. Pour tout $\lambda \in \mathbb{C}$, la famille $(\lambda v_i)_{i\in I}$ est sommable et

$$\sum_{i\in I} \lambda v_i = \lambda \sum_{i\in I} v_i. \tag{10.281}$$

LEMooSBYEooNXzqJU

Lemme 10.114.
Soit une famille sommable $(a_i)_{i\in I}$. Si $J \subset I$, alors la famille $(a_j)_{j\in J}$ est sommable

LEMooAYFUooLMBBDn

Lemme 10.115.
Soit une famille sommable $(a_i)_{i\in I}$ de réels positifs. Si $J \subset I$, alors la famille $(a_j)_{j\in J}$ est sommable et

$$\sum_{j\in J} a_j \leqslant \sum_{i\in I} a_i. \tag{10.282}$$

Dans le cas de familles de nombres réels positifs, nous avons une caractérisation plus commode.

PROPooOYNRooQFpBly

Proposition 10.116 (Somme de réels positifs[1])**.**
Soit $(a_i)_{i\in I}$ une famille de nombres réels positifs indexés par un ensemble quelconque I.

ITEMooAYKKooVxXp

(1) La famille $(a_i)_{i\in I}$ est sommable[56] *si et seulement si*

$$\sup_{J \text{ fini dans } I} \sum_{j\in J} a_j < \infty. \tag{10.283}$$

ITEMooSDCYooNsbHez

(2) Si la famille est sommable, alors

$$\sum_{i\in I} a_i = \sup_{J \text{ fini dans } I} \sum_{j\in J} a_j. \tag{10.284}$$

Démonstration. En plusieurs parties.

(1) **Pour (1), $\Rightarrow$** Soit v comme dans la définition. Soit $\epsilon > 0$. Nous considérons une partie finie $J_0 \subset I$ comme il faut. Vu que tous les termes sont positifs,

$$\sup_{J \text{ fini dans } I} \sum_{j\in J} a_j = \sup_{\substack{J \text{ fini dans } I \\ J_0\subset J}} \sum_{j\in J} a_j. \tag{10.285}$$

Pour chaque J fini contenant J_0 dans I nous avons $\| \sum_{j\in J} a_j - v\| \leqslant \epsilon$ et donc

$$\| \sum_{j\in J} a_j\| \leqslant \|v\| + \epsilon. \tag{10.286}$$

Donc le supremum sur les J est majoré par $\|v\| + \epsilon$.
Vu que cela est valable pour tout ϵ, c'est aussi majoré par $\|v\|$ lui-même, mais ce n'est pas utile pour notre propos.

(2) **Pour (1), $\Leftarrow$** Nous supposons maintenant que $\sup_{J \text{ fini dans } I} \sum_{j\in J} a_j < \infty$. Soit $v < \infty$ ce supremum. Il existe un J_0 fini dans I tel que

$$\| \sum_{j\in J_0} a_j - v\| \leqslant \epsilon. \tag{10.287}$$

Étant donné que tous les termes sont positifs, nous pouvons même dire que

$$\sum_{j\in J_0} a_j \in [v - \epsilon, v]. \tag{10.288}$$

Pour tout K contenant J_0 nous avons la même majoration et donc

$$v - \epsilon \leqslant \sum_{j\in J_0} a_j \leqslant \sum_{k\in K} a_k \leqslant v. \tag{10.289}$$

56. Définition 10.112.

(3) **Pour (2)** Nous nommons $v = \sup_{J \text{ fini dans } I} \sum_{j\in J} a_j$. Soit $\epsilon > 0$; il existe J_0 fini dans I tel que

$$v - \epsilon \leqslant \sum_{j\in J_0} a_j \leqslant v. \tag{10.290}$$

Pour tout K fini contenant J_0 nous avons aussi $\sum_{j\in J_0} a_j \leqslant \sum_{k\in K} a_k$ et donc

$$\sum_{k\in K} a_k \in [v-\epsilon, v]. \tag{10.291}$$

Au final nous avons bien prouvé que

$$\| \sum_{k\in K} a_k - v\| \leqslant \epsilon. \tag{10.292}$$

EQooTWURooBsVfQS

Vu que pour tout ϵ, il existe un K qui réalise (10.292), la définition 10.112 nous dit que $\sum_{i\in I} a_i = v$.

□

LEMooGXPGooZTJPoN

Lemme 10.117 ([1]).
Soient un espace vectoriel normé V ainsi qu'une suite sommable $(a_k)_{k\in\mathbb{N}}$ dans V. Alors

$$\sum_{k\in\mathbb{N}} a_k = \lim_{N\to\infty} \sum_{k=0}^{N} a_k. \tag{10.293}$$

EQooXQWBooLymndN

La somme à gauche est celle de la définition 10.112 et celle de droite est donnée par la définition 1.85.

Démonstration. Par hypthèse, le membre de gauche existe, et nous le nommons v. Soit $\epsilon > 0$. Il existe J_0 fini dans $\mathbb{N}$ tel que pour tout K fini dans $\mathbb{N}$ tel que $J_0 \subset K$, nous avons

$$\| \sum_{k\in K} a_k - v\| < \epsilon. \tag{10.294}$$

Ici, la somme est celle de la définition 1.309. Nous nommons $m_0 = \max(J_0)$; pour tout $m > m_0$ nous avons

$$J_0 \subset \{0, \dots, m_0\} \subset \{0, \dots, m\} \tag{10.295}$$

et donc

$$\| \sum_{k\in\{0,\dots,m\}} a_k - v\| < \epsilon. \tag{10.296}$$

La bijection à prendre dans la définition (1.309) pour faire la somme sur $\{0, \dots, m\}$ peut être l'identité, de telle sorte que $\sum_{k\in\{0,\dots,m\}} a_k = \sum_{k=0}^{m} a_k$.

Nous notons s_i la somme partielle $s_i = \sum_{k=0}^{i} a_k$. Ce que nous avons prouvé jusqu'ici est que pour tout ϵ, il existe m_0 tel que si $m \geqslant m_0$, alors $\|s_i - v\| \leqslant \epsilon$. Cela signifie exactement que $\lim_{m\to\infty} \sum_{k=0}^{m} a_k = v$, c'est-à-dire que le membre de droite de (10.293) vaut v, ce qu'il fallait. □

Exemple 10.118.
La suite $a_i = (-1)^i$ n'est pas sommable parce que quel que soit J_0 fini dans $\mathbb{N}$, nous pouvons trouver J fini contenant J_0 tel que $\sum_{j\in J}(-1)^j > 10$. Pour cela il suffit d'ajouter à J_0 suffisamment de termes pairs. De la même façon en ajoutant des termes impairs, on peut obtenir $\sum_{j\in J'}(-1)^i < -10$. △

Exemple 10.119.
De temps en temps, la somme peut sortir d'un espace. Si nous considérons l'espace des polynômes $[0,1] \to \mathbb{R}$ muni de la norme uniforme, la somme de l'ensemble

$$\{1, -1, \pm\frac{x^n}{n!}\}_{n\in\mathbb{N}} \tag{10.297}$$

est zéro.

Par contre la somme de l'ensemble $\{1, \frac{x^n}{n!}\}_{n\in\mathbb{N}}$ est l'exponentielle qui n'est pas un polynôme. △

PROPooJLQAooAEbIvZ

Proposition 10.120 ([1]).
Soient un espace vectoriel normé V, deux ensembles disjoints A et B ainsi que $v\colon A\cup B\to V$. Si $\sum_{k\in A}v_k$ et $\sum_{k\in B}v_k$ sont sommables[57]*, alors $(v_k)_{k\in A\cup B}$ est sommable et*

$$\sum_{k\in A\cup B}v_k=\sum_{k\in A}v_k+\sum_{k\in B}v_k. \tag{10.298}$$

Démonstration. Nous nommons

$$\begin{aligned}s&=\sum_{k\in A}v_k\\ t&=\sum_{k\in B}v_k,\end{aligned} \tag{10.299}$$

et nous prouvons que $s+t$ vérifie la propriété qu'il faut (définition 10.112) pour être $\sum_{k\in A\cup B}v_k$. Soit $\epsilon>0$. Il existe I_0 fini dans A tel que $\|\sum_{k\in I}v_k-s\|<\epsilon$ dès que $I_0\subset I\subset A$. Il existe également J_0 fini dans B tel que $\|\sum_{k\in J}v_k-t\|<\epsilon$ dès que $J_0\subset J\subset B$.

Nous posons $S_0=I_0\cup J_0$. Si S est fini dans $A\cup B$ et vérifie $S_0\subset S\subset A\cup B$, nous notons $S_a=S\cap A$ et $S_b=S\cap B$. Étant donné que $I_0\subset S_a$ et $J_0\subset S_b$, nous avons

$$\begin{aligned}\|\sum_{k\in S}v_k-(s+t)\|&=\|\sum_{k\in S_a}v_k+\sum_{k\in S_b}v_k-(s+t)\| && \text{cf. justif.} && \text{(10.300a)}\\ &\leqslant\|\sum_{k\in S_a}v_k-s\|+\|\sum_{k\in S_b}v_k-t\| && && \text{(10.300b)}\\ &\leqslant 2\epsilon. && && \text{(10.300c)}\end{aligned}$$

SUBEQooPZSMooRNVSOt

Justifications.

— Pour 10.300a. Vu que les sommes sont finies et que S_a et S_b sont disjoints d'union S, nous pouvons couper la somme en utilisant le lemme 1.312.

□

10.8.2 Somme non dénombrables

Nous allons voir que les sommes non dénombrables ne sont pas intéressantes : si le nombre de valeurs non nulles parmi les $(x_i)_{i\in I}$ est non dénombrable, alors la somme est infinie. La bonne généralisation de somme infinie dans le cas non dénombrable est l'intégrale qui viendra seulement avec la définition 14.167 et la mesure de Lebesgue 14.143.

LEMooYJCVooHajEbg

Lemme 10.121.
Si A est non dénombrable dans $\mathbb{R}$, alors il existe $\delta>0$ tel que $A\cap\{|x|\geqslant\delta\}$ est non dénombrable.

Démonstration. Nous y allons par l'absurde, et nous supposons que A ne contient pas zéro (sinon il faut ajouter zéro aux A_n ci-dessous, et ça alourdit les notations). Nous supposons donc que les parties

$$A_n=A\cap\{|x|\geqslant\frac{1}{n}\} \tag{10.301}$$

sont dénombrables. Mais

$$A\subset\bigcup_{n=1}^{\infty}A_n. \tag{10.302}$$

Une union dénombrable d'ensembles dénombrables est dénombrable[58]. Vu qu'un ensemble non dénombrable ne peut être inclus dans un ensemble dénombrable[59], nous avons une contradiction. □

57. Définition 10.112.
58. Proposition 1.138.
59. Proposition 1.143.

LEMooQIMGooOUpZjk

Lemme 10.122.
Soit un ensemble I et une « suite » $(x_i)_{i\in I}$ avec $x_i \geqslant 0$ pour tout i. Si l'ensemble

$$F = \{i \in I \text{ tel que } x_i > 0\} \tag{10.303}$$

est non dénombrable, alors

$$\sum_{i\in I} x_i = \infty. \tag{10.304}$$

Démonstration. Nous considérons l'ensemble des valeurs non nulles atteintes par x :

$$V = \{x_i \text{ tel que } i \in F\}. \tag{10.305}$$

Il y a deux possibilités : soit V est dénombrable (ou fini), soit il est non dénombrable.

(1) **V est fini ou dénombrable** Dans ce cas, l'application $x\colon F \to [0,\infty[$ est une application d'un ensemble indénombrable vers un ensemble dénombrable. Le lemme 1.144 nous indique qu'il existe $y \in \mathbb{R}$ tel que $x^{-1}(y)$ est indénombrable et en particulier infini. La somme $\sum_{i\in x^{-1}(y)} x_i$ est une somme indénombrable de termes tous égaux et strictement positifs. Elle est infinie.

(2) **V est indénombrable** La partie V de $\mathbb{R}$ est non dénombrable ; elle est donc sujette au lemme 10.121 : il existe $\delta > 0$ tel que $W = V \cap \{x \geqslant \delta\}$ est indénombrable. Vu que $x_i \geqslant \delta$ pour tout i dans $x^{-1}(W)$ nous avons

$$\sum_{i\in x^{-1}(W)} x_i = \infty. \tag{10.306}$$

□

10.8.3 Sommes dénombrables

La proposition suivante nous enseigne que les sommes infinies se comportent normalement au moins en ce qui concerne les majorations termes à termes.

PropMpBStL

Proposition 10.123 ([1]).
Soit I un ensemble dénombrable. ITEMooZSDSooFUqXDO

(1) Soient $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$, deux familles de réels positifs telles que $a_i \leqslant b_i$ et telles que (b_i) est sommable[60]*. Alors (a_i) est sommable.* ITEMooREEYooOtklRb

(2) Si $(a_i)_{i\in I}$ est une famille de réels telle que $(|a_i|)$ est sommable, alors (a_i) est sommable. ITEMooFIGGooWADdLs

(3) Si $(z_i)_{i\in I}$ est une famille de complexes telle que $(|z_i|)$ est sommable, alors (z_i) est sommable.

Démonstration. En plusieurs parties.

(1) **Pour (1)** Nous pouvons utiliser la caractérisation 10.116 ; dans un premier temps nous avons

$$\sum_{i\in I} b_i = \sup_{J \text{ fini dans } I} b_i < \infty. \tag{10.307}$$

Pour chaque J fini dans I, le lemme 1.314 nous assure que

$$\sum_{j\in J} a_j \leqslant \sum_{j\in J} b_j \tag{10.308}$$

Donc le supremum existe et est plus petit pour (a_j) que pour (b_j). En utilisant à nouveau la caractérisation de la proposition 10.116(1) (mais dans l'autre sens), nous concluons que $(a_i)_{i\in I}$ est sommable.

60. Définition 10.112.

(2) **Pour (2)** Nous posons $I^+ = \{i \in I \text{ tel que } a_i \geqslant 0\}$ et $I^- = \{i \in I \text{ tel que } a_i < 0\}$.
Vu que $(|a_i|)_{i\in I}$ est sommable, la famille $(a_i)_{i\in I^+}$ est sommable par le lemme 10.114. La famille $(a_i)_{i\in I^-}$ est également sommable par le même lemme appliqué à $(-a_i)_{i\in I^-}$ qui est sommable par le lemme 10.113. Nous posons

$$S^+ = \sum_{i\in I^+} a_i \tag{10.309a}$$
$$S^- = \sum_{i\in I^-} a_i, \tag{10.309b}$$

et nous allons prouver que $(a_i)_{i\in I}$ est sommable de somme $S^+ + S^-$. Soit $\epsilon > 0$. Nous choisissons J_0^+ et J_0^- tels que $J_0^+ \subset I^+$, $J_0^- \subset I^-$ et

$$\| \sum_{k\in K^+} a_k - S^+ \| \leqslant \epsilon \tag{10.310a}$$
$$\| \sum_{k\in K^-} a_k - S^- \| \leqslant \epsilon \tag{10.310b}$$

pour tout K^+ et K^- vérifiant $J_0^+ \subset K^+ \subset I^+$ et $J_0^- \subset K^- \subset I^-$.
Nous posons $J_0 = J_0^+ \cup J_0^-$ (qui est une union disjointe). Pour tout K fini dans I contenant J_0, nous avons

$$\| \sum_{k\in K} a_k - (S^+ + S^-) \| = \| \sum_{k\in K\cap I^+} a_k + \sum_{k\in K\cap I^-} a_k - S^+ - S^- \| \tag{10.311a}$$
$$\leqslant \| \sum_{k\in K\cap I^+} a_k - S^+ \| + \| \sum_{k\in K\cap I^-} a_k - S^- \| \tag{10.311b}$$
$$\leqslant 2\epsilon \tag{10.311c}$$

parce que $J_0^+ \subset K \cap I^+$ et $J_0^- \subset K \cap I^-$.

(3) **Pour (3)** Similaire à (2), mais en coupant en 4 morceaux au lieu de 2 : les parties réelles et imaginaires en plus des parties positives et négatives.

□

EXooULLXooTDFYqf

Exemple 10.124.
Au sens de la définition 10.112 la famille

$$\frac{(-1)^n}{n} \tag{10.312}$$

n'est pas sommable. En effet la somme des termes pairs est ∞ alors que la somme des termes impairs est $-\infty$. Quel que soit $J_0 \subset \mathbb{N}$, nous pouvons concocter, en ajoutant des termes pairs, un J avec $J_0 \subset J$ tel que $\sum_{j\in J}(-1)^j/j$ soit arbitrairement grand. En ajoutant des termes négatifs, nous pouvons également rendre $\sum_{j\in J}(-1)^j/j$ arbitrairement petit. △

PropVQCooYiWTs

Proposition 10.125.
Si (a_{ij}) est une famille de nombres positifs indexés par $\mathbb{N}\times\mathbb{N}$ alors

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} = \sum_{i=1}^{\infty} \Big(\sum_{j=1}^{\infty} a_{ij} \Big). \tag{10.313}$$

Démonstration. Nous allons utiliser la proposition 10.116 pour traiter la somme de gauche. Nous considérons $J_{m,n} = \{0,\dots,m\}\times\{0,\dots,n\}$ et nous avons pour tout m et n :

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant \sum_{(i,j)\in J_{m,n}} a_{ij} = \sum_{i=0}^{m} \Big(\sum_{j=0}^{n} a_{ij} \Big). \tag{10.314}$$

Si nous fixons m et que nous prenons la limite $n \to \infty$ (qui commute avec la somme finie sur i) nous trouvons

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant \sum_{i=0}^{m} \Big(\sum_{j=0}^{\infty} a_{ij} \Big). \tag{10.315}$$

Cette inégalité étant valable pour tout m, elle est encore valable à la limite $m \to \infty$ et donc

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant \sum_{i=0}^{\infty} \Big(\sum_{j=0}^{\infty} a_{ij} \Big). \tag{10.316}$$

Pour l'inégalité inverse, il faut remarquer que si J est fini dans $\mathbb{N}^2$, il est forcément contenu dans $J_{m,n}$ pour m et n assez grands. Alors

$$\sum_{(i,j)\in J} a_{ij} \leqslant \sum_{(i,j)\in J_{m,n}} a_{ij} = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} \leqslant \sum_{i=0}^{\infty} \Big(\sum_{j=0}^{\infty} a_{ij} \Big). \tag{10.317}$$

Cette inégalité étant valable pour tout ensemble fini $J \subset \mathbb{N}^2$, elle reste valable pour le supremum. □

La définition générale de la somme 10.112 est compatible avec la définition usuelle dans les cas où cette dernière s'applique.

PropoWHdjw

Proposition 10.126 (commutative sommabilité)**.**
Soit I un ensemble dénombrable et une bijection $\tau\colon \mathbb{N} \to I$. Soit $(a_i)_{i\in I}$ une famille dans un espace vectoriel normé. Si $\sum_{i\in I} a_i$ existe, alors il est donné par

$$\sum_{i\in I} a_i = \lim_{N\to\infty} \sum_{k=0}^{N} a_{\tau(k)}. \tag{10.318}$$

Démonstration. Nous posons $a = \sum_{i\in I} a_i$. Soit $\epsilon > 0$ et J_0 comme dans la définition. Nous choisissons

$$N > \max_{j\in J_0}\{\tau^{-1}(j)\}. \tag{10.319}$$

En tant que sommes sur des ensembles finis, nous avons l'égalité

$$\sum_{k=0}^{N} a_{\tau(k)} = \sum_{j\in J} a_j \tag{10.320}$$

où J est un sous-ensemble de I contenant J_0. Soit J fini dans I tel que $J_0 \subset J$. Nous avons alors

$$\| \sum_{k=0}^{N} a_{\tau(k)} - a \| = \| \sum_{j\in J} a_j - a \| < \epsilon. \tag{10.321}$$

Nous avons prouvé que pour tout ϵ, il existe N tel que $n > N$ implique $\| \sum_{k=0}^{n} a_{\tau(k)} - a \| < \epsilon$. □

La réciproque n'est pas vraie. Même en supposant que $\lim_{N\to\infty} \sum_{n=0}^{N} a_n$ existe, il n'est pas forcé que $\sum_{n\in\mathbb{N}} a_n$ existe. Cela est une conséquence de l'exemple 10.124.

PROPooWLEDooJogXpQ

Proposition 10.127 ([1])**.**
Soit un espace vectoriel normé E et une famille sommable[61] *$\{v_i\}_{i\in I}$ d'éléments de E. Soit $f\colon E \to \mathbb{C}$ une application sur laquelle nous supposons*

(1) f est linéaire et continue ;

(2) la partie $\{f(v_i)_{i\in I}\}$ est sommable.

61. Définition 10.112.

Alors nous pouvons permuter la somme et f :

EQooONHXooKqJEbY
$$f\big(\sum_{i\in I} v_i\big) = \sum_{i\in I} f(v_i). \tag{10.322}$$

Démonstration. Soit $\epsilon > 0$; vu que les familles $\{v_i\}_{i\in I}$ et $\{f(v_i)\}_{i\in I}$ sont sommables, nous pouvons considérer les parties finies J_1 et J_2 de I telles que

$$\Big\| \sum_{j\in J_1} v_j - \sum_{i\in I} v_i \Big\| \leqslant \epsilon \tag{10.323}$$

et

$$\Big\| \sum_{j\in J_2} f(v_j) - \sum_{i\in I} f(v_i) \Big\| \leqslant \epsilon \tag{10.324}$$

Ensuite nous posons $J = J_1 \cup J_2$. Avec cela nous calculons un peu avec les majorations usuelles :

$$\|f(\sum_{i\in I} v_i) - \sum_{i\in I} f(v_i)\| \leqslant \|f(\sum_{i\in I} v_i) - f(\sum_{j\in J} v_j)\| + \|f(\sum_{j\in J} v_j) - \sum_{i\in I} f(v_i)\|. \tag{10.325}$$

Le second terme est majoré par ϵ, tandis que le premier, en utilisant la linéarité de f possède la majoration

$$\|f(\sum_{i\in I} v_i) - f(\sum_{j\in J} v_j)\| = \|f(\sum_{i\in I} v_i - \sum_{j\in J} v_j)\| \leqslant \|f\| \Big\| \sum_{i\in I} v_i - \sum_{j\in J} v_j \Big\| \leqslant \epsilon \|f\|. \tag{10.326}$$

Donc pour tout $\epsilon > 0$ nous avons

$$\|f(\sum_{i\in I} v_i) - \sum_{i\in I} f(v_i)\| \leqslant \epsilon(1 + \|f\|). \tag{10.327}$$

D'où l'égalité (10.322). □

10.9 Série réelle

secseries

La notion de série formalise le concept de somme infinie[62]. L'absence de certaines propriétés de ces objets (problèmes de commutativité et même d'associativité) incite à la prudence et montre à quel point une définition précise est importante.

10.9.1 Critères de convergence absolue

Étant donné le terme général d'une série, il est souvent –dans les cas qui nous intéressent– difficile de déterminer la somme de la série. L'exemple de la série géométrique est particulier[63], puisqu'on connaît une formule pour chaque somme partielle, mais pour l'exemple des séries de Riemann il n'y a aucune formule simple pour un α général. D'où l'intérêt d'avoir des critères de convergence ne nécessitant aucune connaissance de l'éventuelle limite de la série.

LemgHWyfG

Lemme 10.128 (Critère de comparaison)**.**
Soient $\sum_i a_i$ et $\sum_j b_j$ deux séries à termes positifs vérifiant

$$0 \leqslant a_i \leqslant b_i$$

alors

ITEMooBBWTooYDHnuH
(1) si $\sum_i a_i$ diverge, alors $\sum_j b_j$ diverge,

(2) si $\sum_j b_j$ converge, alors $\sum_i a_i$ converge (absolument).

62. La caractérisation qui nous intéresse est celle de la proposition 10.116.
63. Voir la proposition 10.135.

PROPooEEACooOYmtxd

Proposition 10.129 (Critère d'équivalence[326]).
Soient $\sum_i a_i$ et $\sum_j b_j$ deux séries à termes positifs. Supposons l'existence de la limite (éventuellement infinie) suivante

$$\lim_{i\to\infty} \frac{a_i}{b_i} = \alpha \tag{10.328}$$

avec $\alpha \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $\alpha \neq 0$ et $\alpha \neq \infty$, alors

$$\sum_i a_i \text{ converge} \iff \sum_j b_j \text{ converge}, \tag{10.329}$$

(2) si $\alpha = 0$ et $\sum_j b_j$ converge, alors $\sum_i a_i$ converge (absolument),

(3) si $\alpha = +\infty$ et $\sum_j b_j$ diverge, alors $\sum_i a_i$ diverge.

Démonstration. (1) Le fait que la suite a_n/b_n converge vers α signifie que tant sa limite supérieure que sa limite inférieure convergent vers α. En particulier la suite $\frac{a_n}{b_n}$ est bornée vers le haut et vers le bas. À partir d'un certain rang N, il existe M tel que

$$\frac{a_n}{b_n} < M \tag{10.330}$$

et il existe m tel que

$$\frac{a_n}{b_n} > m. \tag{10.331}$$

Nous avons donc $a_n < Mb_n$ et $a_n > mb_n$. La série de (a_n) converge donc si et seulement si la série de (b_n) converge.

(2) Si $\alpha = 0$, cela signifie que pour tout ϵ, il existe un rang tel que $\frac{a_n}{b_n} < \epsilon$, et donc tel que $a_n < \epsilon b_k$. La série de (a_i) converge donc dès que la série de (b_i) converge.

(3) Pour tout M, il existe un rang dans la suite à partir duquel on a $\frac{a_i}{b_i} > M$, et donc $a_k > Mb_k$. Si la série de (b_k) diverge, la série de (a_k) doit également diverger.

□

PropOXKUooQmAaJX

Proposition 10.130 (Critère du quotient[335]).
Soit $\sum_i a_i$ une série. Supposons l'existence de la limite (éventuellement infinie) suivante

$$\lim_{i\to\infty} \left|\frac{a_{i+1}}{a_i}\right| = L \tag{10.332}$$

avec $L \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $L < 1$, la série converge absolument,

(2) si $L > 1$, la série diverge,

(3) si $L = 1$ le critère échoue : il existe des exemples de convergence et des exemples de divergence.

Démonstration. (1) Soit b tel que $L < b < 1$. À partir d'un certain rang K, on a $\left|\frac{a_{i+1}}{a_i}\right| < b$. En particulier,

$$|a_{K+1}| < b|a_K|, \tag{10.333}$$

et pour a_{K+2} nous avons

$$|a_{K+2}| < b|a_{K+1}| < b^2|a_K|. \tag{10.334}$$

Au final,

$$|a_{K+n}| < b^n|a_K|. \tag{10.335}$$

Étant donné que la série $\sum_{n\geqslant K} b^n$ converge (parce que $b < 1$), la queue de suite $\sum_{i\geqslant K} a_i$ converge, et par conséquent la suite au complet converge.

(2) Si $L > 1$, on a
$$|a_K| < |a_{K+1}| < |a_{K+2}| < \dots \tag{10.336}$$
Il est donc impossible que la suite (a_i) converge vers zéro. La série ne peut donc pas converger.

(3) Par exemple la suite harmonique $a_n = \frac{1}{n}$ vérifie $L = 1$, mais la série ne converge pas. Par contre, la suite $a_n = \frac{1}{n^2}$ vérifie aussi le critère avec $L = 1$ tandis que la série $\sum_n \frac{1}{n^2}$ converge. □

Proposition 10.131 (Critère de la racine[326])**.**
Soit $\sum_i a_i$ une série, et considérons
$$\limsup_{i\to\infty} \sqrt[i]{|a_i|} = L$$
avec $L \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $L < 1$, la série converge absolument,

(2) si $L > 1$, la série diverge,

(3) si $L = 1$ le critère échoue.

Démonstration. (1) Si $L < 1$, il existe un $r \in]0, 1[$ tel que $|a_n|^{1/n} < r$ pour les grands n. Dans ce cas, $|a_n| < r^n$, et la série converge absolument parce que la série $\sum_n r^n$ converge du fait que $r < 1$.

(2) Si $L > 1$, il existe un $r > 1$ tel que $|a_n|^{1/n} > r > 1$. Cela fait que $|a_n|$ prend des valeurs plus grandes que n pour une infinité de termes. Le terme général a_n ne peut donc pas être une suite convergente. Par conséquent la suite diverge au sens où elle ne converge pas.

□

10.9.2 Critères de convergence simple

Les critères de comparaison, d'équivalence, du quotient et de la racine sont des critères de convergence absolue. Pour conclure à une convergence simple qui n'est pas une convergence absolue, le critère d'Abel sera notre outil principal.

10.9.2.1 Critère d'Abel

Proposition 10.132 (Critère d'Abel)**.**
Soit la série $\sum_i c_i z_i$ avec

(1) (c_i) est une suite réelle décroissante qui tend vers zéro,

(2) (z_i) est une suite dans $\mathbb{C}$ dont la suite des sommes partielles est bornée dans $\mathbb{C}$, c'est-à-dire qu'il existe un $M > 0$ tel que pour tout n,
$$\left|\sum_{i=1}^{n} z_i\right| \leqslant M. \tag{10.337}$$
Alors la série $\sum_i c_i z_i$ est convergente.

Remarquons que ce critère ne donne pas de convergence absolue.

10.9.3 Quelques séries usuelles

Proposition 10.133 (Série harmonique)**.**
*La **série harmonique** converge vers l'infini*[64] *:*
$$\sum_{k=1}^{\infty} \frac{1}{k} = +\infty. \tag{10.338}$$

64. Vous pouvez aussi dire qu'elle diverge, mais si on met la bonne topologie sur $\mathbb{R}$, la convergence vers $+\infty$ est plus précise que la non-convergence.

Démonstration. Considérons la sommes partielles $H_n = \sum_{k=1}^{n} \frac{1}{k}$. Considérons la différence

$$H_{2n} - H_n = \sum_{k=n+1}^{2n} \frac{1}{k}. \tag{10.339}$$

Cette somme contient $n+1$ termes, tous plus grands que $\frac{1}{2n}$, donc

$$H_{2n} - H_n > n \times \frac{1}{2n} = \frac{1}{2}. \tag{10.340}$$

Nous prouvons donc par récurrence que $H_{2^n} \geqslant \frac{n}{2}$. D'abord pour $n = 1$ nous avons

$$H_2 = 1 + \frac{1}{2}. \tag{10.341}$$

Ensuite la récurrence :

$$H_{2^n} > H_{2^{n-1}} + \frac{1}{2} \geqslant \frac{n-1}{2} + \frac{1}{2} = \frac{n}{2}. \tag{10.342}$$

□

10.134.
À quel point la série harmonique diverge-t-elle lentement ? Allez regarder
https://www.youtube.com/watch?v=_AtkIpi6KP0.

PROPooWOWQooWbzukS

Proposition-Définition 10.135 (Série géométrique)**.**
*La **série géométrique** de raison $q \in \mathbb{C}$ est*

$$\sum_{i=0}^{\infty} q^i. \tag{10.343}$$

EqZQTGooIWEFxL

ITEMooAFAMooGuXqBm

(1) Elle converge si et seulement si $|q| < 1$.

ITEMooBJHBooBMEmiG

(2) Si $|q| < 1$ alors

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}. \tag{10.344}$$

EqRGkBhrX

ITEMooVZHKooNGpDkx

(3) Si la somme commence en $n = 1$ au lieu de $n = 0$ alors

$$\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}. \tag{10.345}$$

EqPZOWooMdSRxY

Démonstration. Plusieurs points.

(1) **Pour (1) et (2)** La somme partielle est déjà donnée dans le lemme 1.487 :

$$S_N = \sum_{n=0}^{N} q^n = \frac{1 - q^{N+1}}{1-q}. \tag{10.346}$$

En vertu de (9.132), la limite $\lim_{N\to\infty} S_N$ existe si et seulement si $|q| \leqslant 1$ et dans ce cas nous avons le résultat parce que $q^{N+1} \to 0$.

(2) **Pour (3)** Il s'agit seulement du calcul

$$\sum_{n=1}^{\infty} q^n = \frac{1}{1-q} - 1 = \frac{q}{1-q}. \tag{10.347}$$

□

Un cas particulier de la formule (1.687) est le calcul de $\sum_{j=1}^{N} q^{-j}$ bien utile lorsque l'on joue avec des nombres binaires (voir l'exemple 34.12). Nous avons

$$\sum_{j=1}^{N} q^{-j} = \sum_{j=0}^{N} q^{-j} - 1 = \frac{1-q^{-N}}{q-1}. \tag{10.348}$$

EQooFMBAooEJkHWT

La série de Riemann est très liée aux intégrales impropres de la proposition 14.300.

PEXoCTYNZoCYQrs

Proposition 10.136 (Série de Riemann)**.**
*Pour $\alpha \in \mathbb{R}$, la **série de Riemann***

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} \tag{10.349}$$

EqSerRiem

converge (absolument, puisque réelle et positive) si et seulement si $\alpha > 1$, et diverge sinon.

ExIJMHooOEUKfj

Lemme 10.137 (Série exponentielle)**.**
La série

$$\exp(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}. \tag{10.350}$$

converge pour tout $t \in \mathbb{R}$.

Elle est nommée . Pour tout savoir de l'exponentielle et de ses variations, voir le thème 50.

Démonstration. Si $a_k = t^k/k!$ alors $\frac{a_{k+1}}{a_k} = \frac{t}{k}$ dont la limite $k \to \infty$ est zéro (quel que soit t). En vertu du critère du quotient 10.130 la série exponentielle converge (absolument) pour tout $t \in \mathbb{R}$. □

LEMooCVIQooUtuzgE

Lemme 10.138 (Série arithmético-géométrique[336])**.**
*Une **suite arithmético-géométrique** est une suite vérifiant pour tout n la relation*

$$u_{n+1} = au_n + b \tag{10.351}$$

avec a et b non nuls. Nous nommons u_0 le premier terme.

(1) Si la suite possède une limite, alors elle est donnée par

$$l = \frac{b}{1-a}. \tag{10.352}$$

EQooMZSSooQWpZUp

(2) Pour tout n nous avons la formule

$$u_n = a^n(u_0 - l) + l. \tag{10.353}$$

Démonstration. Si la suite possède une limite, cette dernière doit résoudre $l = al + b$, et donc être égale à (10.352).

Il n'est pas très compliqué de trouver le terme général de la suite en fonction de a et de b. Il suffit de considérer la suite $v_n = u_n - l$, et de remarquer que cette suite est géométrique :

$$v_{n+1} = av_n. \tag{10.354}$$

Par conséquent $v_n = a^n v_0$, ce qui donne pour la suite (u_n) la formule

$$u_n = a^n(u_0 - l) + l. \tag{10.355}$$

□

LEMooKDHPooPlFTIT

Lemme 10.139 ([337]).
Nous avons :

$$\sum_{k=1}^{N} \frac{1}{k(k+1)} = \frac{N}{N+1}. \tag{10.356}$$

et

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1. \tag{10.357}$$

Démonstration. Nous posons

$$f(n) = \sum_{k=1}^{n} \frac{1}{k(k+1)} \tag{10.358a}$$

$$g(n) = \frac{n}{n+1} \tag{10.358b}$$

et nous montrons par récurrence que $f(n) = g(n)$. Pour $n = 1$ nous avons $f(1) = g(1) = \frac{1}{2}$.

Nous supposons que $f(n) = g(n)$ et nous prouvons que $f(n+1) = g(n+1)$. Facile :

$$f(n+1) = f(n) + \frac{1}{(n+1)(n+2)} \tag{10.359a}$$

$$= \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} \tag{10.359b}$$

$$= \frac{n(n+2)+1}{(n+1)(n+2)} \tag{10.359c}$$

$$= \frac{n^2+2n+1}{(n+1)(n+2)} \tag{10.359d}$$

$$= \frac{(n+1)^2}{(n+1)(n+2)} \tag{10.359e}$$

$$= \frac{n+1}{n+2} \tag{10.359f}$$

$$= g(n+1). \tag{10.359g}$$

En ce qui concerne la seconde formule, par définition [65]

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k(k+1)} = \lim_{n\to\infty} \frac{n}{n+1} = 1. \tag{10.360}$$

□

10.9.4 Séries alternées

THOooOHANooHYfkII

Théorème 10.140 (Critère des séries alternées[311]).
Soit $(a_n)_{n\in\mathbb{N}}$ est une suite positive décroissante à limite nulle. Nous notons S_n la somme partielle de la série alternée : $S_n = \sum_{k=0}^{n}(-1)^k a_k$. Alors

(1) les sous-suites (S_{2n}) et (S_{2n+1}) sont adjacentes [66].

(2) La série $\sum_n (-1)^n a_n$ converge.

ITEMooWEPWooXhLMYL

(3) Si nous considérons le reste

$$R_n = \sum_{k=n+1}^{\infty} (-1)^k a_k, \tag{10.361}$$

65. Définition d'une série, 10.89.
66. Définition 9.39.

nous avons

$$\operatorname{sgn}(R_n) = (-1)^{n+1} \tag{10.362a}$$
$$|R_n| \leqslant a_{n+1}. \tag{10.362b}$$

Démonstration. En termes de notations, nous allons écrire (S_n) la suite des sommes partielles de $\sum_{k=0}^{\infty}(-1)^k a_k$. Nous notons (S_{2n}) la suite des termes pairs de cette suite. C'est donc la suite $n \mapsto S_{2n}$. Nous divisons en plusieurs morceaux.

(1) S_{2n} **est croissante** Nous avons simplement

$$S_{2n+2} - S_{2n} = a_{2n+2} - a_{2n+1} \leqslant 0. \tag{10.363}$$

(2) (S_{2n+1}) **est décroissante** Même calcul.

(3) **Les suites** (S_{2n}) **et** S_{2n+1} **sont adjacentes** Nous avons simplement

$$S_{2n+1} - S_{2n} = a_{2n+1} \to 0. \tag{10.364}$$

Nous concluons par le théorème des suites adjacentes 9.40 que les sous-suites des termes pairs et impairs sont convergentes et convergent vers la même limite.

C'est le moment d'utiliser la proposition 9.41 qui convaincra le lecteur que (S_n) converge vers la même limite, que nous notons S. Le théorème des suites adjacentes nous dit encore que

$$S_{2n+1} \leqslant S \leqslant S_{2n} \tag{10.365}$$

et donc que $R_{2n} = S - S_{2n} \leqslant 0$. Cela donne la majoration

$$|R_{2n}| = |S - S_{2n}| = S_{2n} - S \leqslant S_{2n} - S_{2n+1} = a_{2n+1}. \tag{10.366}$$

Nous faisons le même genre de majorations pour R_{2n+1}. □

10.9.5 Moyenne de Cesàro

La moyenne de Cesàro est le premier pas dans la direction des supersommes[2] qui permettent de sommer des choses de moins en moins convergentes, jusqu'à sommer la fameuse série $1 + 2 + 3 + 4 + \ldots = -1/12$.

DEFooLVRLooTeowkn

Définition 10.141.
Si $a\colon \mathbb{N} \to V$ est une suite dans l'espace vectoriel V, alors sa ***moyenne de Cesàro*** *est la limite (si elle existe) de la suite*

$$\sigma_n(a) = \frac{1}{n}\sum_{k=1}^{n} a_k. \tag{10.367}$$

En un mot, c'est la limite des moyennes partielles.

LemyGjMqM

Lemme 10.142.
Si la suite (a_n) converge vers la limite ℓ alors la suite admet une moyenne de Cesàro qui vaudra ℓ.

Démonstration. Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que $|a_n - \ell| < \epsilon$ pour tout $n > N$. En remarquant que

$$\frac{1}{n}\sum_{k=1}^{n} a_k - \ell = \frac{1}{n}\sum_{k=1}^{n}(a_k - \ell), \tag{10.368}$$

nous avons

$$\Big|\frac{1}{n}\sum_{k=1}^{n} a_k - \ell\Big| \leqslant \Big|\frac{1}{n}\sum_{k=1}^{N}|a_k - \ell|\Big| + \Big|\frac{1}{n}\sum_{k=N+1}^{n}\underbrace{|a_k - \ell|}_{\leqslant\epsilon}\Big| \tag{10.369a}$$
$$\leqslant \epsilon + \frac{n - N - 1}{n}\epsilon \tag{10.369b}$$
$$\leqslant 2\epsilon. \tag{10.369c}$$

Dans ce calcul nous avons redéfini N de telle sorte que le premier terme soit inférieur à ϵ. □

10.9.6 Écriture décimale d'un réel

Nous avons déjà vu la fonction (1.119) qui permet d'écrire des naturels dans une base $b \geqslant 2$ donnée. Nous allons maintenant construire une fonction du même type, pour la partie décimale d'un réel.

NORMALooTZWYooPMgOIm

10.143.
Soit $b \geqslant 2$ un entier qui sera la base dans laquelle nous allons écrire les nombres. Nous considérons l'ensemble $\mathbb{D}_b$ des suites dans $\{0, 1, \dots, b-1\}$ qui n'ont pas une queue de suite uniquement formée de $b-1$. Autrement dit une suite (c_n) est dans $\mathbb{D}_b$ lorsque pour tout N, il existe $k > N$ tel que $c_k \neq b-1$. Associé à cet ensemble nous considérons la fonction

$$\begin{aligned} \varphi_b \colon \mathbb{D}_b &\to [0,1[\\ c &\mapsto \sum_{n=1}^{\infty} \frac{c_n}{b^n}. \end{aligned} \tag{10.370}$$

EqXXXooOTsCK

Lemme 10.144.
La fonction φ_b est bien définie au sens où elle converge et prend ses valeurs dans $[0,1[$.

Démonstration. Tout se base sur la somme de la série géométrique (10.344) sous la forme

$$\sum_{k=0}^{\infty} \frac{1}{b^k} = \frac{b}{b-1}. \tag{10.371}$$

EqWZGooXIgwI

La somme (10.370) est donc majorée par $\sum_n \frac{b-1}{b^n}$ qui converge.

Pour prouver que l'image de φ_b est bien $[0,1[$, nous savons qu'au moins un des c_n (en fait une infinité) est plus petit que $b-1$, donc nous avons la majoration stricte [67]

$$\varphi_b(c) < \sum_{n=1}^{\infty} \frac{b-1}{b^n} = (b-1)\left(\sum_{n=1}^{\infty} \frac{1}{b^n} - 1 \right) = 1 \tag{10.372}$$

□

Le fait d'introduire l'ensemble $\mathbb{D}$ au lieu de l'ensemble de toutes les suites est justifié par la proposition suivante. Elle explique pourquoi un nombre possède au maximum deux écritures décimales distinctes et que ces deux sont obligatoirement de la forme, par exemple en base 10 :

$$0.34599999999\ldots = 0.34600000\ldots \tag{10.373}$$

mais qu'un nombre commençant par 0.347 ne peut pas être égal. C'est pour cela que dans la définition de $\mathbb{D}_b$ nous avons exclu les suites qui terminent par tout des $b-1$.

La proposition suivante complète ce qui est déjà dit dans le lemme 7.275.

PropSAOoofRlQR

Proposition 10.145.
Soit la fonction

$$\begin{aligned} \varphi \colon \{0, \dots, b-1\}^{\mathbb{N}} &\to [0,1[\\ x &\mapsto \sum_{n=1}^{\infty} \frac{x_n}{b^n}. \end{aligned} \tag{10.374}$$

Si $\varphi(x) = \varphi(y)$ et si n_0 est le plus petit entier tel que $x_{n_0} \neq y_{n_0}$ alors soit

$$x_{n_0} - y_{n_0} = 1 \tag{10.375}$$

et $x_n = 0$, $y_n = b-1$ pour tout $n > n_0$, soit le contraire : $y_{n_0} - x_{n_0} = 1$ avec $y_n = 0$ et $x_n = b-1$ pour tout $n > n_0$.

67. Notez que la somme (10.370) commence à un tandis que la série géométrique (10.371) commence à zéro.

Démonstration. Nous nous basons sur la formule (facilement dérivable depuis (10.371)) suivante :

$$\sum_{k=n_0+1}^{\infty} \frac{1}{b^k} = \frac{1}{b^{n_0+1}} \frac{b}{b-1}. \tag{10.376}$$

Nous avons

$$0 = \varphi(x) - \varphi(y) = \frac{x_{n_0} - y_{n_0}}{b^{n_0}} + \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} \geqslant \frac{x_{n_0} - y_{n_0}}{b^{n_0}} - \sum_{n=n_0+1}^{\infty} \frac{b-1}{b^n} = \frac{x_{n_0} - y_{n_0} - 1}{b^{n_0}}. \tag{10.377}$$

Le dernier terme étant manifestement positif[68], il est nul et nous avons $x_{n_0} - y_{n_0} = 1$.

Nous avons donc maintenant

$$0 = \varphi(x) - \varphi(y) = \frac{1}{b^{n_0}} + \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n}. \tag{10.378}$$ EqHWQoottPnb

Nous majorons la dernière somme de la façon suivante, en supposant que $|x_n - y_n| \neq b-1$ pour un certain $n > n_0$:

$$\left| \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} \right| \leqslant \sum_{n=n_0+1}^{\infty} \frac{|x_n - y_n|}{b^n} < \sum_{n=n_0+1}^{\infty} \frac{b-1}{b^n} = \frac{1}{b^{n_0}}. \tag{10.379}$$

Étant donné cette inégalité stricte, l'équation (10.378) ne peut pas être correcte (valoir zéro). Nous avons donc $|x_n - b_n| = b-1$ pour tout $n > n_0$. Donc pour chaque $n > n_0$ nous avons soit $x_n = 0$ et $y_n = b-1$, soit $a_n = b-1$ et $b_n = 0$. Pour conclure il faut encore prouver que le choix doit être le même pour tout n.

Nous nous mettons dans le cas $x_{n_0} - y_{n_0} = 1$; dans ce cas nous avons bien l'égalité (10.378) sans petites nuances de signes. Nous écrivons

$$\sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} = (b-1) \sum_{n=n_0+1}^{\infty} \frac{(-1)^{s_n}}{b^n} \tag{10.380}$$

où s_n est pair ou impair suivant que $x_n = 0$, $y_n = b-1$ ou le contraire. Si un des $(-1)^{s_n}$ est pas -1 alors nous avons l'inégalité stricte

$$(b-1) \sum_{n=n_0+1}^{\infty} \frac{(-1)^{s_n}}{b^n} > (b-1) \sum_{n=n_0+1}^{\infty} \frac{-1}{b^n} = -\frac{1}{b^{n_0}}. \tag{10.381}$$

Dans ce cas il est impossible d'avoir $\varphi(x) - \varphi(y) = 0$. Nous en concluons que $(-1)^{s_n}$ est toujours -1, c'est-à-dire $x_n - y_n = 1 - b$, ce qui laisse comme seule possibilité $x_n = 0$ et $y_n = b-1$. □ ThoRXBootpUpd

Théorème 10.146.
L'application $\varphi_b \colon \mathbb{D}_b \to [0,1[$ *est bijective.*

Démonstration. En ce qui concerne l'injection, nous savons de la proposition 10.145 que si $\varphi_b(x) = \varphi_b(y)$ pour $x, y \in \{0, \dots, b-1\}^{\mathbb{N}}$, alors soit x soit y a une queue de suite composée uniquement de $b-1$, ce qui est exclu dans $\mathbb{D}_b$. Nous en déduisons que φ_b est bien injective en prenant $\mathbb{D}_b$ comme ensemble départ.

La partie lourde est la surjectivité. Nous prenons $x \in [0,1[$ et nous allons construire par récurrence une suite $a \in \mathbb{D}_b$ telle que $\varphi_b(a) = x$. Si il existe $a_1 \in \{0, \dots, b-1\}$ tel que $x = a_1/b$ alors nous prenons la suite $(a_1, 0, \dots,)$ et nous avons évidemment $\varphi(a) = x$. Sinon il existe $a_1 \in \{0, \dots, b-1\}$ tel que

$$\frac{a_1}{b} < x < \frac{a_1 + 1}{b} \tag{10.382}$$

68. C'est ici qu'intervient la subdivision entre le cas $x_{n_0} - y_{n_0} = 1$ ou le contraire. En effet si « ce dernier terme était manifestement *négatif* », il aurait fallu majorer avec de $1 - b$ au lieu de $1 - b$.

parce que les autres possibilités pour x sont dans l'ensemble $[0,1[\backslash\{\frac{k}{b}\}_{k=0,\dots,b-1}$ que nous subdivisons en

$$]0,\frac{1}{b}[\cup]\frac{1}{b},\frac{2}{b}[\cup\dots\cup]\frac{b-1}{b},1[. \tag{10.383}$$

Pour la récurrence nous supposons avoir trouvé $a_1,\dots,a_n$ tels que

$$\sum_{k=1}^{n}\frac{a_k}{b^k}<x<\sum_{k=1}^{n-1}\frac{a_k}{b^k}+\frac{a_n+1}{b^n}. \tag{10.384}$$

Encore une fois si il existe $a_{n+1}\in\{0,\dots,b-1\}$ tel que $\sum_{k=1}^{n+1}\frac{a_k}{b^k}=x$ alors nous prenons ce a_{n+1} et nous complétons la suite avec des zéros pour avoir $\varphi(a)=x$. Sinon , pour simplifier les notations nous notons $x'=x-\sum_{k=1}^{n}\frac{a_k}{b^k}$ et nous avons

$$0<x'<\frac{a_n+1}{b^n}. \tag{10.385}$$

Le nombre x' est forcément dans un des intervalles

$$]\frac{s}{b^{n+1}},\frac{s+1}{b^{n+1}}[\tag{10.386}$$

avec $s\in\{0,\dots,b-1\}$. Nous prenons le s correspondant à x' comme a_{n+1}. Dans ce cas nous avons

$$\sum_{k=1}^{n+1}\frac{a_k}{b^k}<x<\sum_{k=1}^{n+1}\frac{a_k}{b^k}+\frac{1}{b^{n+1}}. \tag{10.387}$$

Note : les deux inégalités sont strictes. La première parce que si il y avait égalité, nous nous serions déjà arrêté en complétant avec des zéros. La seconde parce que

$$\sum_{k=n+2}^{\infty}\frac{a_k}{b^k}\leqslant\sum_{k=n+2}^{\infty}\frac{b-1}{b^k}=\frac{1}{b^{n+1}} \tag{10.388}$$

où l'égalité n'est possible que si $a_k=b-1$ pour tout $k\geqslant n+2$. Dans ce cas nous aurions eu

$$x=\sum_{k=1}^{n}\frac{a_k}{b^k}+\frac{a_{n+1}+1}{b^{n+1}} \tag{10.389}$$

et nous aurions choisi le nombre a_{n+1} autrement et complété la suite par des zéros à partir de là. Notons que cela prouve au passage que la suite que nous sommes en train de construire est bien dans $\mathbb{D}_b$ parce qu'elle ne contiendra pas de queue de suite composée de $b-1$.

Ceci termine la construction par récurrence de la suite $a\in\mathbb{D}_b$. Par construction nous avons pour tout $N\geqslant 1$,

$$\sum_{k=1}^{N}\frac{a_k}{b^k}\leqslant x\leqslant\sum_{k=1}^{N}\frac{a_k}{b^k}+\frac{1}{b^{N+1}}, \tag{10.390}$$

autrement dit : $\varphi_b(a_1,\dots,a_N)\in B(x,\frac{1}{b^{N+1}})$. Nous avons donc bien convergence

$$\lim_{N\to\infty}\varphi_b(a_1,\dots,a_N)=x \tag{10.391}$$

et l'application φ_b est surjective. □

L'application $\varphi_b^{-1}\colon[0,1[\to\mathbb{D}_b$ est la **décomposition décimale** en base b des nombres de $[0,1[$.

Tout cela nous permet de montrer entre autres que $\mathbb{R}$ n'est pas dénombrable. Vu qu'il y a une bijection entre $[0,1[$ et $\mathbb{D}_b$, il suffit de prouver que $\mathbb{D}_b$ est non dénombrable. De plus il suffit de démontrer que $\mathbb{D}_b$ est non dénombrable pour un entier $b\geqslant 2$ donné.

PropNNHooYTVFw

Proposition 10.147 ([338]).
Il n'existe pas de surjection $\mathbb{N} \to \mathbb{D}_b$. Autrement dit $\mathbb{D}_b$ est non dénombrable.

Démonstration. Nous prenons $b \neq 2$ pour des raisons qui seront claires plus tard. Soit $f : \mathbb{N} \to \mathbb{D}_b$. Pour $i \in \mathbb{N}$ nous notons

$$f(n) = (c_i^{(n)})_{i \geqslant 1}, \tag{10.392}$$

et nous définissons la suite

$$c_k = \begin{cases} 0 & \text{si } c_k^{(k)} \neq 0 \\ 1 & \text{si } c_k^{(k)} = 0. \end{cases} \tag{10.393}$$

C'est une suite dans $\mathbb{D}_b$ parce que $b \neq 2$ et que la suite ne contient que des 0 et des 1. Mais nous n'avons $f(n) = c$ pour aucun $n \in \mathbb{N}$ parce que nous avons $c_n \neq f(n)_n$.

Si $b = 2$ alors nous savons que $\mathbb{D}_2 \sim [0, 1[\sim \mathbb{D}_3$. Donc $\mathbb{D}_2 \sim \mathbb{D}_3$ et $\mathbb{D}_2$ ne peut pas plus être mis en bijection avec $\mathbb{N}$ que $\mathbb{D}_3$. □

Remarque 10.148.
Le cas de la base $b = 2$ doit être fait à part parce que rien n'empêche d'avoir une queue de 1. Il y a alors toutefois moyen de se débrouiller en construisant la suite c de façon plus subtile. Si $b = 2$ et $n \in \mathbb{N}$ alors $f(n)$ est une suite de 0 et 1 contenant une infinité de 0 (parce qu'il n'y a pas de queue de suite ne contenant que des 1). Nous construisons alors c de la façon suivante : d'abord nous recopions $f(0)$ jusqu'à son *deuxième* zéro que nous changeons en 1 ; nommons n_0 le rang de ce deuxième zéro. Ensuite nous recopions les éléments de $f(1)$ à partir du rang $n_0 + 1$ jusqu'au second zéro que nous changeons en 1, etc.

Le fait de prendre le deuxième zéro nous garantit que la suite c n'aura pas de queue de suite ne contenant que des 1.

Notons que cette construction s'adapte à tout b ; il suffit de prendre le second terme qui n'est pas $b - 1$ et le remplacer par $b - 1$.

Corolaire 10.149.
L'ensemble $[0, 1[$ n'est pas dénombrable.

Démonstration. L'ensemble $[0, 1[$ est en bijection avec $\mathbb{D}_b$ que nous venons de prouver n'être pas dénombrable. □

10.9.7 Théorème de Banach-Steinhaus

LEMooPIPLooMppGSO

Lemme 10.150 ([339]).
Soient des espaces vectoriels normés X et Y ainsi qu'une application linéaire bornée $T : X \to Y$. Pour tout $a \in X$ et pour tout $r > 0$ nous avons

$$\sup_{x \in B(a,r)} \|Tx\| \geqslant r\|T\| \tag{10.394}$$

Démonstration. Nous commençons avec $a = 0$. En utilisant la définition 10.55 de la norme opérateur,

$$\|T\| = \sup_{x \in X} \frac{\|Tx\|}{\|x\|} = \sup_{x \in B(0,r)} \frac{\|Tx\|}{\|x\|} \leqslant \frac{1}{r} \sup_{x \in B(0,r)} \|Tx\|. \tag{10.395}$$

Donc

$$\sup_{x \in B(0,r)} \|Tx\| \geqslant r\|T\|. \tag{10.396}$$

Il y a maintenant une astuce. Nous considérons un maximum :

$$\max\{\|T(a+x),\|T(a-x)\|\} \geqslant \frac{1}{2}\big(\|T(a+x)\|+\|T(a-x)\|\big) \tag{10.397a}$$
SUBEQooPJPMooDkqBHs

$$\geqslant \frac{1}{2}\big(\|T(a+x)-T(a+x)\|\big) \tag{10.397b}$$
SUBEQooEZUUooVlK1fn

$$= \frac{1}{2}\|T(2x)\| \tag{10.397c}$$

$$= \|Tx\|. \tag{10.397d}$$

Justifications :

— Pour (10.397a), la moyenne est plus petite que le maximum.

— Pour (10.397b), inégalité triangulaire : $\|\alpha-\beta\| \leqslant \|\alpha\|+\|\beta\|$.

Si maintenant $y \in B(a,r)$, nous avons $y = a+x$ pour un certain $x \in B(0,r)$, donc

$$\sup_{y\in B(a,r)} \|Ty\| = \sup_{x\in B(0,r)} \|T(a+x)\| \tag{10.398a}$$

$$= \sup_{x\in B(0,r)} \max\{\|T(a+x)\|,\|T(a-x)\|\} \tag{10.398b}$$
SUBEQooACJSooTHCAWs

$$\geqslant \sup_{x\in B(0,r)} \|Tx\| \tag{10.398c}$$

$$\geqslant r\|T\|. \tag{10.398d}$$

Pour (10.398b), l'ensemble sur lequel nous prenons le supremum n'est pas modifié fondamentalement si nous regroupons les éléments deux à deux en prenant le maximum : les éléments exclus sont majorés. □

Une version avec des seminormes sera le théorème 27.5.

THOooJHVNooIDDxyT

Théorème 10.151 (Théorème de Banach-Steinhaus[339])**.**
Soient un espace de Banach[69] X *et un espace vectoriel normé* Y*. Soit une famille* $\mathcal{F}$ *d'opérateurs linéaire bornés. Si pour tout* $x \in X$*,*

$$\sup_{T\in\mathcal{F}} \|Tx\| < \infty, \tag{10.399}$$

alors

$$\sup_{T\in\mathcal{F}} \|T\| < \infty. \tag{10.400}$$

Démonstration. Nous supposons que $\sup_{T\in\mathcal{F}}\|T\| = \infty$, de telle sorte que nous pouvons choisir une suite (T_n) dans $\mathcal{F}$ telle que $\|T_n\| \to \infty$. Cette suite peut diverger arbitrairement vite, et nous fixerons exactement cela plus tard.

Soit par ailleurs une suite $\alpha_n > 0$ d'éléments petits et tels que $\alpha_n \to 0$. Nous supposons que $\sum_{n=0}^{\infty}\alpha_n < \infty$.

Si $a \in X$, le lemme 10.150 dit que

$$\sup_{x\in B(a,\alpha_n)} \|T_n x\| \geqslant \|T_n\|\alpha_n. \tag{10.401}$$

En posant $x_0 = 0$, nous construisons une suite (x_n) par récurrence en imposant

(1) $x_n \in B(x_{n-1},\alpha_n)$

(2) $\|T_n x_n\| \geqslant \|T_n\|\alpha_n$.

En utilisant une série télescopique et l'inégalité triangulaire $\|x_k - x_{k+1}\| \leqslant \alpha_n$ à chaque étage,

$$\|x_p - x_q\| \leqslant \sum_{k=p}^{q}\alpha_k \leqslant \sum_{k=p}^{\infty}\alpha_k. \tag{10.402}$$

69. Définition 7.265.

Mais puisque la somme des α_n converge, la suite des queues de somme converge vers zéro[70] : $\lim_{p\to\infty}\sum_{k=p}^{\infty}\alpha_n=0$. Cela implique que (x_n) est une suite de Cauchy[71]. Vu que X est de Banach, la suite (x_n) a une limite dans X. Soit x cette limite.

Nous avons $\beta_n=\|x_n-x\|\to 0$. Il y aurait moyen de calculer β_n en fonction de α_n (surtout si nous avions donné une forme explicite à α_n), mais c'est sans importance ici. L'important est que c'est une suite qui tend vers zéro.

Nous avons

$$x\in B(x_n,\beta_n),\tag{10.403}$$

et donc il existe $a_n\in B(0,\beta_n)$ tel que $x=x_n+a_n$. Avec cela, pour chaque n nous avons :

$$\begin{aligned}\|T_nx\|&=\|T_n(x_n+a_n)\\ &\geqslant\|T_nx_n\|-\|T_na_n\|\\ &\geqslant\|T_nx_n\|-\|T_n\|\beta_n\\ &\geqslant\|T_n\|\alpha_n-\|T_n\|\beta_n=\|T_n\|(\alpha_n-\beta_n).\end{aligned}\tag{10.404}$$

SUBEQooPLVQooChVCLU

Pour 10.404c, nous avons utilisé $\|T_na_n\|\leqslant\|T_n\|\beta_n$. En résumé,

$$\|T_nx\|\geqslant\|T_n\|(\alpha_n-\beta_n).\tag{10.405}$$

Il suffit de choisir $\|T_n\|$ suffisamment rapidement croissant pour que[72]

$$\|T_n\|(\alpha_n-\beta_n)\to\infty,\tag{10.406}$$

et nous avons $\|T_nx\|\to\infty$, qui est contraire aux hypothèses. □

ThoPFBMHBN

Théorème 10.152 (Théorème de Banach-Steinhaus[104, 340])**.**
Soit E un espace de Banach[73] et F un espace vectoriel normé. Nous considérons une partie $H\subset\mathcal{L}^0(E,F)$ (espace des applications linéaires continues). Alors H est uniformément borné si et seulement si il est simplement borné.

Démonstration. Si H est uniformément borné, il est borné ; pas besoin de rester longtemps sur ce sens de l'équivalence. Supposons donc que H soit borné. Pour chaque $k\in\mathbb{N}^*$ nous considérons l'ensemble

$$\Omega_k=\{x\in E\text{ tel que }\sup_{f\in H}\|f(x)\|>k\}.\tag{10.407}$$

(1) **Les Ω_k sont ouverts** Soit $x_0\in\Omega_k$; nous avons alors une fonction $f\in H$ telle que $\|f(x_0)\|>k$, et par continuité de f il existe $\rho>0$ tel que $\|f(x)\|>k$ pour tout $x\in B(x_0,\rho)$. Par conséquent $B(x_0,\rho)\subset\Omega_k$ et Ω_k est ouvert par le théorème 7.8.

(2) **Les Ω_k ne sont pas tous denses dans E** Nous supposons que les ensembles Ω_k soient tous dense dans E. Le théorème de Baire 7.366 nous indique que E est un espace de Baire (parce que de Banach) et donc que

$$\overline{\bigcap_{k\in\mathbb{N}}\Omega_k}=E.\tag{10.408}$$

En particulier l'intersection des Ω_k n'est pas vide. Soit $x_0\in\bigcap_{k\in\mathbb{N}}\Omega_k$. Nous avons alors

$$\sup_{f\in H}\|f(x)\|=\infty,\tag{10.409}$$

ce qui est contraire à l'hypothèse. Donc les ouverts Ω_k ne sont pas tous denses dans E.

70. Lemme 10.91(2).
71. Proposition 7.270.
72. Le point important ici est que α_n (et donc β_n) est choisi sans référence à $\|T_n\|$.
73. Définition 7.265.

(3) **La majoration** Il existe $k \geqslant 0$ tel que Ω_k ne soit pas dense dans E, et nous voulons prouver que $\{\|f\| \text{ tel que } f \in H\}$ est un ensemble borné. Soit donc $k \geqslant 0$ tel que Ω_k ne soit pas dense dans E ; il existe un $x_0 \in E$ et $\rho > 0$ tels que

$$B(x_0, \rho) \cap \Omega_k = \varnothing. \tag{10.410}$$

Si $x \in B(x_0, \rho)$ alors x n'est pas dans Ω_k et donc

$$\sup_{f \in H} \|f(x)\| \leqslant k. \tag{10.411}$$

Afin d'évaluer $\|f\|$ nous devons savoir ce qu'il se passe avec les vecteurs sur une boule autour de 0. Pour tout $x \in B(0, \rho)$ et pour tout $f \in H$, la linéarité de f donne

$$\|f(x)\| = \|f(x + x_0) - f(x_0)\| \leqslant \|f(x + x_0) + f(x_0)\| \leqslant 2k. \tag{10.412}$$

Par continuité nous avons alors $\|f(x)\| \leqslant 2k$ pour tout $x \in \overline{B(0, \rho)}$. Si maintenant $x \in F$ vérifie $\|x\| = 1$ nous avons

$$\|f(x)\| = \frac{1}{\rho}\|f(\rho x)\| \leqslant \frac{2k}{\rho}, \tag{10.413}$$

et donc $\|f\| \leqslant \frac{2k}{\rho}$, ce qui montre que $2k/\rho$ est un majorant de l'ensemble $\{\|f\| \text{ tel que } f \in H\}$.

□

Une application du théorème de Banach-Steinhaus est l'existence de fonctions continues et périodiques dont la série de Fourier ne converge pas. Ce sera l'objet de la proposition 28.21.

10.9.8 Convergence forte

Lorsque nous avons une suite d'opérateurs linéaires, nous pouvons considérer la convergence d'une suite pour la norme opérateur : $A_k \to A$ lorsque $\|A_k - A\| \to 0$.

DEFooNREQooElLvec

Définition 10.153 ([341]).
*Soient un espace vectoriel E et un espace vectoriel normé V. Nous disons que la suite d'opérateur $T_k : E \to V$ **converge fortement** vers l'opérateur T si pour tout $x \in E$ nous avons*

$$\|T_k x - Tx\| \to 0. \tag{10.414}$$

Cette notion s'appelle *forte* par opposition à la convergence *faible* dont nous ne parlerons pas. Elle est cependant moins forte que la convergence en norme dont nous avons déjà parlé.

PROPooRFBLooUjSirP

Proposition 10.154.
Soient des espaces vectoriels normés E et F et une suite d'opérateurs $T_k : E \to F$ convergeant vers T[74]*. Alors cette suite converge également fortement.*

Démonstration. Soit $x \in E$ que nous supposons non nul. Soit $\lambda \in \mathbb{C}$ tel que $x = \lambda y$ avec $\|y\| = 1$. Nous avons

$$\|T_k x - Tx\| = |\lambda|\|T_k y - Ty\| \leqslant |\lambda| \sup_{\|z\|=1} \|T_k z - Tz\| = |\lambda|\|T_k - T\| \to 0. \tag{10.415}$$

La dernière étape est la convergence en norme $T_k \to T$. □

PROPooESZVooJNCidk

Proposition 10.155.
Soient E et F, des espaces vectoriels normés de dimension finie. Soit une suite (A_n) d'applications linéaires $E \to F$. Si elle converge fortement vers A, alors elle converge en norme vers A.

Démonstration. En plusieurs coups.

74. Sans précisions, ce sera toujours la convergence en norme.

(1) **Si une sous-suite converge** Commençons par montrer que si (B_n) est une sous-suite de (A_n) qui converge vers B, alors $B = A$. Autrement dit, A est le seul candidat limite pour A_n.
Soit $\|x\| = 1$. Nous avons

$$\|B_n x - Bx\| \leqslant \|B_n - B\|\|x\| = \|B_n - B\|, \tag{10.416}$$

mais pour la sous-suite (B_n) nous avons supposé $\|B_n - B\| \to 0$. Donc $\|B_n x - Bx\| \to 0$, ce qui signifie que $B_n x \to Bx$. Mais par hypothèse, $B_n x \to Ax$. Par unicité de la limite, $Bx = Ax$ pour tout x de norme 1. Pour les autres x, c'est la linéarité qui conclut.

(2) **Utilisation de deux gros résultats** Par l'hypothèse de convergence, pour chaque x nous avons $\sup_n \|A_n x\| < \infty$. Le théorème de Banach-Steinhaus 10.151 nous indique alors que l'ensemble $\mathcal{F} = \{A_n\}_{n\in\mathbb{N}}$ est borné. Il existe donc $M > 0$ tel que $\|A_n\| < M$ pour tout n.
Nous utilisons à présent l'hypothèse de dimension finie en disant que l'espace des applications linéaires $E \to F$ est de dimension finie, de telle sorte que ses boules fermées soient compactes.
Donc la suite (A_n) est contenue dans un compact.

(3) **Les sous-suite convergentes** La suite (A_n) est contenue dans un compact. Toutes ses sous-suites sont dans ce compact et possèdent donc une sous-suite convergente (théorème 7.144). Toutes ces sous-sous-suites convergent nécessairement vers A par ce que nous avons dit dans la première étape de la preuve. Le lemme 7.66 nous dit alors que $A_n \to A$.

□

10.10 Application ouverte

Définition 10.156 (application ouverte)**.**
*Soient deux espaces topologiques X et Y. Une application $f\colon X \to Y$ est **ouverte** si l'image de tout ouvert de X par f est un ouvert de Y.*

Nous disons que f est ouverte en $a \in X$ si l'image de tout ouvert contenant a est ouverte.

PROPooXGEGooHoMsne

Proposition 10.157.
Une application bijective est ouverte si et seulement si son inverse est continue.

Démonstration. Ce n'est seulement que la définition, mais pour le sport nous démontrons le sens direct.

Soit donc une application $f\colon X \to Y$ bijective et ouverte entre les espaces topologiques X et Y. Prouvons que $f^{-1}\colon Y \to X$ est continue. Pour cela nous considérons un ouvert $\mathcal{O}$ dans X, et nous prouvons que $(f^{-1})^{-1}(\mathcal{O})$ est ouvert dans Y. Par définition de l'inverse, $(f^{-1})^{-1}(\mathcal{O}) = f(\mathcal{O})$ et vu que f est ouverte, $f(\mathcal{O})$ est ouvert. □

LEMooHHIPooEpGfCg

Lemme 10.158.
Une application $f\colon X \to Y$ est ouverte si et seulement si pour tout $x \in X$ et pour tout voisinage U de x, la partie $f(U)$ est un voisinage de $f(x)$.

Démonstration. La preuve suit celle de la proposition 7.204. Le sens direct est un à fortiori.

Dans l'autre sens. Soit un ouvert $\mathcal{O}$ de X. Pour prouver que $f(\mathcal{O})$ est ouvert, nous considérons $y \in f(\mathcal{O})$, ainsi que $x \in \mathcal{O}$ tel que $f(x) = y$. Vu que $\mathcal{O}$ est un voisinage de x, la partie $f(\mathcal{O})$ est un voisinage de $y = f(x)$.

Il existe donc un ouvert V de Y tel que $y = f(x) \in V \subset f(\mathcal{O})$. Donc la partie $f(\mathcal{O})$ contient un ouvert autour de chacun de ses points, et elle est ouverte par le théorème 7.8. □

Lemme 10.159 ([342])**.**
Une application linéaire entre espaces vectoriels topologiques est ouverte si et seulement si elle est ouverte en 0.

Démonstration. Le sens direct est un à fortiori.

Soit un ouvert $\mathcal{O}$ et $a \in \mathcal{O}$. La partie $\mathcal{O} - a$ est ouverte et contient 0. Donc $f(\mathcal{O} - a)$ est un ouvert parce que f est ouverte en 0. Nous en déduisons, par linéarité, que $f(\mathcal{O}) - f(a)$ est ouvert et donc que $f(\mathcal{O})$ est ouverte. □

Lemme 10.160 ([342]).
Soient des espaces vectoriels normés E et F. Une application linéaire ouverte $f : E \to F$ est surjective.

Démonstration. Soit un ouvert $B(0,r)$ dans E. Puisque f est ouverte, la partie $f\big(B(0,r)\big)$ est ouverte dans F, et contient donc une boule $B_F(0,r')$ pour un certain $r' > 0$.

Soit $v \in F$. Nous considérons

$$v' = r' \frac{v}{2\|v\|}. \tag{10.417}$$

Nous avons $\|v'\| = r'/2$, et donc $v' \in B_F(0,r')$. Il existe donc $x \in E$ (et même dans $B_E(0,r)$) tel que $f(x) = v'$. Nous avons alors

$$f\big(\frac{2\|v\|}{r'}x\big) = v, \tag{10.418}$$

ce qui prouve que v est dans l'image de f, et donc que f est surjective. □

THOooATZKooXHWCRD

Théorème 10.161 (théorème de l'application ouverte[342, 343, 344]).
Soient des espaces de Banach[75] *E et F. Si l'application $f : E \to F$ est linéaire, surjective et continue, alors elle est ouverte.*

Démonstration. En plusieurs étapes.

(1) **Une union de fermés** Soit $y \in F$. Comme f est surjective, il existe $x \in E$ tel que $y = f(x)$. Soit $n \in \mathbb{N}$ tel que $x \in B(0,n)$. Nous avons alors

$$y \in f\big(B(0,n)\big) \subset \overline{f\big(B(0,n)\big)} \tag{10.419}$$

En notant

$$F_n = \overline{f\big(B_E(0,n)\big)}, \tag{10.420}$$

nous avons

$$F = \bigcup_{n=0}^{\infty} F_n. \tag{10.421}$$

(2) **Théorème de Baire** Le théorème 7.366 nous indique que F est un espace de Baire. Le lemme 7.365 nous dit alors qu'il existe un n tel que F_n soit d'intérieur non vide. Mettons F_N d'intérieur non vide.

(3) **Boule unité** Puisque F_N est d'intérieur non vide, il existe $y \in F_N$ et $\eta > 0$ tels que $B_F(y,\eta) \subset F_N$. Nous avons aussi

$$B_F(0,\eta) = B_F(y,\eta) - y, \tag{10.422}$$

et comme $y \in F_N$ nous avons $B_F(0,\eta) \subset F_N - F_N$, et vu qu'en plus $-F_N = F_N$, nous avons

$$B_F(0,\eta) \subset 2F_N = \overline{f\big(B_E(0,2N)\big)}. \tag{10.423}$$

Nous avons ensuite

$$B_F(0,1) = \frac{1}{\eta}B_F(0,\eta) \subset \frac{1}{\eta}\overline{f\big(B_E(0,2N)\big)} = \overline{f\big(B_E(0,2N/\eta)\big)}. \tag{10.424}$$

Ceci pour dire qu'il existe un $M \in \mathbb{R}$ tel que

$$B_F(0,1) \subset \overline{f\big(B_E(0,M)\big)}. \tag{10.425}$$

75. Espace de Banach : vectoriel, normé, complet. Définition 7.265.

Nous avons de même que

$$B_F\big(0, \frac{1}{2^n}\big) \subset \overline{f\big(B_E(0, M/2^n)\big)}. \tag{10.426}$$

Nous voudrions maintenant avoir la même inclusion sans la fermeture.

(4) **Une suite par récurrence** Soit $z \in B_F(0,1)$. Nous allons définir par récurrence une suite (x_n) dans E telle que

$$\begin{cases} x_n \in B_E\big(0, \frac{M}{2^{n-1}}\big) & \text{(10.427a)} \\ \|z - f(x_1 + \ldots + x_n)\| < \frac{1}{2^n}. & \text{(10.427b)} \end{cases}$$

(4a) **Le premier élément** Puisque $z \in B_F(0,1) \subset \overline{f\big(B_E(0,M)\big)}$, nous avons

$$B(z, \frac{1}{2}) \cap f\big(B_E(0,M)\big) \neq \varnothing. \tag{10.428}$$

Nous pouvons donc considérer $x_1 \in B_E(0,M)$ tel que $f(x_1) \in B(z, \frac{1}{2})$.
Ce x_1 vérifie les conditions (10.427).

(4b) **La récurrence** En utilisant l'hypothèse de récurrence et (10.426),

$$z - f(x_1 + \ldots + x_n) \in B_F\big(0, \frac{1}{2^n}\big) \subset \overline{f\big(B_E(0, M/2^n)\big)}, \tag{10.429}$$

de telle sorte que

$$B_F\big(z - f(x_1, \ldots, x_n), \frac{1}{2^{n+1}}\big) \cap f\big(B_E(0, M/2^n)\big) \neq \varnothing. \tag{10.430}$$

Nous pouvons donc considérer $x_{n+1} \in B_E(0, M/2^n)$ tel que

$$f(x_{n+1}) \in B_F\big(z - f(x_1 + \ldots + x_n), \frac{1}{2^{n+1}}\big). \tag{10.431}$$

Donc

$$z - f(x_1 + \ldots + x_n) - f(x_{n+1}) \in B_F\big(0, \frac{1}{2^{n+1}}\big). \tag{10.432}$$

Nous avons donc bien

$$\|z - f(x_1 + \ldots + x_{n+1})\| < \frac{1}{2^{n+1}}. \tag{10.433}$$

(5) **Convergence** Nous avons, pour tout n, $\|x_n\| < \frac{M}{2^{n-1}}$. Donc la série

$$\sum_{n=1}^{\infty} \|x_n\| \leqslant \sum_{n=1}^{\infty} \frac{M}{2^{n-1}} \tag{10.434}$$

converge. Autrement dit, la série des x_n converge absolument [76]. Puisque E est un espace de Banach, la proposition 10.96 nous dit que $\sum_{n=1}^{\infty} x_n$ converge dans E. Nous posons

$$x = \sum_{n=1}^{\infty} x_n. \tag{10.435}$$

En utilisant la série géométrique de la proposition 10.135(2), nous trouvons

$$\|x\| \leqslant \sum_{k=1}^{\infty} \|x_k\| \leqslant M \sum_{k=1}^{\infty} \frac{1}{2^{k+1}} = M \sum_{k=0}^{\infty} \frac{1}{2^k} = 2M. \tag{10.436}$$

76. Définition 10.92.

(6) **Passage à la limite** Nous avons $x \in B_E(0,2M)$, et

$$\lim_{n\to\infty} \|z - f(x_1 + \ldots + x_n)\| = 0. \tag{10.437}$$

Puisque les applications $\|.\|$, $t \mapsto z - t$ et f sont continues[77], nous pouvons rentrer la limite de partout et écrire

$$\|z - f(x)\| = 0, \tag{10.438}$$

ce qui signifie que $z = f(x)$. Comme z est un élément arbitraire de $B_F(0,1)$ nous avons prouvé que

$$B_F(0,1) \subset f\big(B_E(0,2M)\big). \tag{10.439}$$

Nous avons donc aussi que pour tout $r > 0$, il existe r' tel que

$$B_F\big(0,r\big) \subset f\big(B_E(0,r')\big). \tag{10.440}$$

En l'occurrence, $r' = r/2M$.

(7) **Passage aux voisinages** Nous montrons que l'image de tout voisinage de $x \in E$ contient un voisinage de $f(x)$ dans F. Soit $x \in E$ ainsi qu'un voisinage V de x. Il existe $r > 0$ tel que $B(x,r) \subset V$. Vu que f est linéaire,

$$f\big(B(x,r)\big) = f(x) + f\big(B(0,r)\big), \tag{10.441}$$

et il existe un r' tel que $B_F(0,r') \subset f\big(B_E(0,r)\big)$. Cela pour dire que

$$f(x) + B_F(0,r') \subset f\big(B(x,r)\big) \subset f(V). \tag{10.442}$$

Vu que $f(x) + B_F(0,r')$ est un ouvert autour de $f(x)$, nous avons prouvé que $f(V)$ contient un ouvert autour de $f(x)$, c'est-à-dire que $f(V)$ est un voisinage de $f(x)$.

(8) **Conclusion** Le lemme 10.158 conclut que f est ouverte.

□

10.11 Produit tensoriel d'espaces vectoriels

Si vous êtes pressés, vous pouvez aller lire la définition 10.170 de produit tensoriel d'espaces vectoriels. Mais si vous étiez vraiment pressés, vous ne seriez pas en train de lire des choses sur le produit tensoriel (il vous suffit de croire que $x \otimes y$ n'est finalement que la concatenation de x et y).

PROPooYONEooWvwPZT

Proposition-Définition 10.162.
*Soient un espace vectoriel V et un sous-espace N. Le **quotient** de V par N, noté V/N est l'ensemble des classes d'équivalence*[78] *pour la relation $x \sim y$ si et seulement si $x - y \in N$.*

Les définitions

(1) $[v] + [w] = [v + w]$

(2) $\lambda[v] = [\lambda v]$

ont un sens et définissent une structure d'espace vectoriel sur V/N.

En ce qui concerne la topologie, ce sera la définition 7.49.

Démonstration. Un élément général de la classe $[v]$ est de la forme $v + n$ avec $n \in N$. Le calcul suivant montre que la somme fonctionne :

$$[v + n_1] + [w + n_2] = [v + w + n_1 + n_2] = [v + w] \tag{10.443}$$

77. Oui, la continuité de f est une hypothèse en plus de sa linéarité parce que nous n'avons pas d'hypothèses sur la dimension de E et F.

78. Définition 1.33.

parce que $n_1 + n_2 \in N$. De même,

$$\lambda[v+n] = [\lambda v + \lambda n] = [\lambda v] \tag{10.444}$$

toujours parce que $\lambda n \in N$.

Notons que nous avons utilisé de façon on ne peut plus cruciale le fait que N soit un sous-espace vectoriel. □

Proposition 10.163.
Si $\{e_i\}$ est une base de V et si N est un sous-espace de V, alors $\{[e_i]\}$ est une partie génératrice de V/N.

Démonstration. Si $x = \sum_k x_k e_k$, alors $[x] = \sum_k x_k [e_k]$, donc oui. □

10.12 Produit tensoriel

Nous allons faire les choses suivantes :

— Définir ce qu'est *un* produit tensoriel de V et W, définition 10.165.

— Prouver que tous les produits tensoriels sont isomorphes, définition 10.166 et proposition 10.167.

— Construire un produit tensoriel relativement abstrait à base d'espace librement engendré par $V \times W$, définition 10.170 et proposition 10.172.

— Construire un produit tensoriel assez concret en tant qu'espace de formes bilinéaires sur $V \times W$, proposition 10.192.

10.164.
Dans le Frido, la notation $V \otimes W$ pourra désigner (au moins) trois choses différentes.

— Un produit tensoriel quelconque, c'est-à-dire un espace quelconque vérifiant les conditions de la définition 10.165.

— L'espace $F_{\mathbb{K}}(V \times W)/N$ défini en 10.170.

— L'espace $\mathcal{L}_2(V \times W, \mathbb{K})$ des formes bilinéaires.

Dans tous les cas l'espace T qui veut prétendre être le produit tensoriel $V \otimes W$ doit venir avec une application h. Lorsque (T, h) est un produit tensoriel de V et W, nous notons $v \otimes w$ l'élément $h(v, z) \in T$.

10.12.1 Définition générale

DEFooXKKQooAvWRNp

Définition 10.165 ([345])**.**
*Soient deux espaces vectoriels V et W. Un **produit tensoriel** de V et W est un couple (T, h) où T est un espace vectoriel et $h : V \times W \to T$ est une application*

(1) bilinéaire[79]

ITEMooJCNYooGvjjtL

(2) telle que pour tout espace vectoriel Z et toute application bilinéaire $f : V \times W \to Z$, il existe une unique application linéaire $\tilde{f} : T \to Z$ telle que $f = \tilde{f} \circ h$.

*La propriété (2) est appelée **propriété universelle** du produit tensoriel.*

L'élément $h(v, w)$ est souvent noté $v \otimes w$.

DEFooPLHTooRiHjlE

Définition 10.166.
*Un **morphisme** entre (T, h) et (T', h') est une application linéaire $\psi : T \to T'$ telle que $h' = \psi \circ h$.*

*Nous parlons d'**isomorphisme** si ψ a un inverse qui est également un morphisme.*

79. Définition 8.124.

PROPooROPHooQXqNzZ

Proposition 10.167 ([345]).
Si V et W sont des espaces vectoriels, tous les produits tensoriels entre V et W sont isomorphes entre eux au sens de la définition 10.166.

Plus précisément, si (T, h) et (T', h') sont deux produits tensoriels de V et W, alors

(1) il existe une unique application linéaire $g: T \to T'$ telle que $h' = g \circ h$,

(2) cette application g est inversible.

En particulier, l'application g est un isomorphisme d'espaces vectoriels.

Démonstration. Soient deux produits tensoriels (T, h) et (T', h').

(1) **Existence**
L'application $h': V \oplus W \to T'$ est bilinéaire, et (T, h) est un produit tensoriel. Donc il existe $g: T \to T'$ tel que $h' = g \circ h$. De même, il existe une application $g': T' \to T$ telle que $h = g' \circ h$.

(2) **Unicité**
En ce qui concerne l'unicité, vu que $h: V \oplus W \to T$ est surjective, la relation $h' = g \circ h$ prescrit les valeurs de g sur tous les éléments de T.

(3) **Inversible**
Ces deux applications g et g' vérifient $h' = gg'h$ et $h = g'gh$, et de plus $h: V \oplus W \to T$ est surjective. Soient $t \in T$ et $x \in V \oplus W$ tel que $t = h(x)$. Nous avons $h(x) = g'gh(x)$. C'est-à-dire $t = (g' \circ g)(t)$. De même dans l'autre sens, il existe $x' \in V \oplus W$ tel que $t = h'(x')$. En appliquant l'égalité $h' = gg'h'$ à x', nous trouvons $t = (g \circ g')(t)$.
Tout cela pour dire que $g' = g^{-1}$. Cette application g est donc un isomorphisme de produits tensoriels entre (T, h) et (T', h').

Au final, l'application $g: T \to T'$ étant linéaire et inversible, elle est un isomorphisme d'espaces vectoriels. □

Proposition 10.168 ([346]).
Soient deux espaces vectoriels V et W sur le corps $\mathbb{K}$ ainsi que $u, a \in V$ et $v, b \in W$ tous non nuls. Si

$$u \otimes v = a \otimes b, \tag{10.445}$$

alors il existe λ et μ tels que $a = \lambda u$ et $b = \mu v$.

Démonstration. Pour toute applications linéaires $f: V \to \mathbb{K}$ et $g: W \to \mathbb{K}$ nous posons

$$\begin{aligned} f * g: V \times W &\to \mathbb{K} \\ (v, w) &\mapsto f(v)g(w). \end{aligned} \tag{10.446}$$

Cette application $f * g$ étant bilinéaire, il existe une application $s_{fg}: V \otimes W \to \mathbb{K}$ telle que $s_{fg} \circ h = f * g$.

Nous considérons maintenant des bases $\{e_i\}$ de V et $\{e'_j\}$ de W, et nous considérons l'application linéaire $f_i: V \to \mathbb{K}$ donnant la i^{e}composante dans la base $\{e_i\}$; autrement dit

$$f_i(e_k) = \delta ik. \tag{10.447}$$

De même nous considérons l'application $g_j: W \to \mathbb{K}$. Comme dit plus haut, l'application $f_i * g_j: V \times W \to \mathbb{K}$ se relève en une application $s_{ij}: V \otimes W \to \mathbb{K}$ vérifiant

$$s_{ij}(u \otimes v) = (f_i * g_j)(u, v) = f_i(u)g_j(v). \tag{10.448}$$

Par hypothèse, $s_{ij}(u \otimes v) = s_{ij}(a \otimes b)$, donc

$$f_i(u)g_j(v) = f_i(a)g_j(b). \tag{10.449}$$

EQooBEXAooDMSEnG

Étant donné que $u \neq 0$, il existe i_0 tel que $f_{i_0}(u) \neq 0$. De même nous considérons j_0 tel que $g_{j_0}(v) \neq 0$. Dans l'égalité

$$f_{i_0}(u)g_{j_0}(v) = f_{i_0}(a)g_{j_0}(b), \tag{10.450}$$

le membre de gauche est non nul. Donc les deux facteurs du membre de droite sont non nuls. Donc en posant

$$\mu = \frac{f_{i_0}(a)}{f_{i_0}(u)}, \tag{10.451}$$

nous définissons bien un élément non nul $\mu \in \mathbb{K}$.

En repartant de (10.449) avec $i = i_0$ nous trouvons

$$g_i(v) = \mu g_j(b) \tag{10.452}$$

pour tout i. Donc $v = \mu b$, ce qu'il fallait démontrer. Un raisonnement similaire montre que $u = \lambda a$ pour un certain λ non nul dans $\mathbb{K}$. □

PROPooMMOPooEkqkpk

Proposition 10.169.
Tous les espaces vectoriels sont sur le corps $\mathbb{K}$. Si $h\colon V \times W \to T$ est un produit tensoriel et si $f\colon T \to Z$ est un isomorphisme d'espaces vectoriels, alors

$$f \circ h\colon V \times W \to Z \tag{10.453}$$

est un produit tensoriel.

Démonstration. En plusieurs parties.

(1) **$f \circ h$ est bilinéaire** Soient $v_1, v_2 \in V$ et $w \in W$ ainsi que $\lambda \in \mathbb{K}$. Nous avons

$$\begin{aligned}(f \circ h)(\lambda v_1 + v_2, w) &= f(\lambda v_1 \otimes w + v_2 \otimes w) && h \text{ est bilin.} && (10.454a)\\ &= \lambda f(v_1 \otimes w) + f(v_2 \otimes w) && f \text{ est lin.} && (10.454b)\\ &= \lambda(f \circ h)(v_1, w) + (f \circ h)(v_2, w), && && (10.454c)\end{aligned}$$

ce qui montre que $f \circ h$ est linéaire en sa première variable. La vérification de la linéarité en la seconde variable est le même type de vérification.

(2) **Propriété universelle** Soient un espace vectoriel Y et une application bilinéaire $b\colon V \times W \to Y$. Nous devons trouver une application linéaire $b'\colon Z \to Y$ telle que $b' \circ (f \circ h) = b$. Vu que (T, h) est un produit tensoriel, il existe une application bilinéaire $\tilde{b}\colon T \to Y$ telle que $\tilde{b} \circ h = b$. Vu que f est linéaire et inversible, l'application $b' = \tilde{b} \circ f^{-1}$ est linéaire et fait l'affaire :

$$b' \circ f \circ h = \tilde{b} \circ f^{-1} \circ f \circ h = \tilde{b} \circ h = b. \tag{10.455}$$

□

10.12.2 Constructions du produit tensoriel

Tout cela est fort bien : nous avons unicité à isomorphisme près du produit tensoriel d'espaces vectoriels. Mais nous n'avons pas encore de certitudes à propos de l'existence d'un couple (T, h) vérifiant les propriétés demandées pour être un produit tensoriel.

Nous allons maintenant construire un produit tensoriel.

10.12.3 Construction abstraite

C'est le moment pour vous de relire la définition 4.27 d'espace vectoriel librement engendré, et surtout le lemme 4.28 qui en donne une base.

DEFooKTVDooSPzAhH

Définition 10.170 ([345]).
Soient deux espaces vectoriels V et W sur le corps commutatif[80] $\mathbb{K}$. *Dans* $F_{\mathbb{K}}(V\times W)$[81] *nous considérons les sous-espaces suivants :*

$$A_1=\{\delta_{(v_1,w)}+\delta_{(v_2,w)}-\delta_{(v_1+v_2,w)} \text{ tel que } v_1,v_2\in V,w\in W\} \tag{10.456a}$$
$$A_2=\{\delta_{(v,w_1)}+\delta_{(v,w_2)}-\delta_{(v,w_1+w_2)} \text{ tel que } v\in V,w_1,w_2\in W\} \tag{10.456b}$$
$$A_3=\{\lambda\delta_{v,w}-\delta_{(\lambda v,w)} \text{ tel que } v\in V,w\in W,\lambda\in\mathbb{K}\} \tag{10.456c}$$
$$A_4=\{\lambda\delta_{v,w}-\delta_{(v,\lambda w)} \text{ tel que } v\in V,w\in W,\lambda\in\mathbb{K}\}. \tag{10.456d}$$

SUBEQooSHBJooJLPVbK

Nous considérons alors $N=\mathrm{Span}(A_1,A_2,A_3,A_4)$ *et le quotient*

$$V\otimes_{\mathbb{K}}W=F_{\mathbb{K}}(V\times W)/N. \tag{10.457}$$

Ce dernier espace vectoriel est le ***produit tensoriel*** *de V par W.*

REMooSLEGooWEiutz

Remarque 10.171.
Quelques remarques.

(1) Les éléments de $V\otimes W$ ne s'écrivent pas tous sous la forme $v\otimes w$. Certains ont vraiment besoin d'être écrits avec des sommes. En cela, la situation de $V\otimes W$ est réellement différente de celle de $V\times W$. Dans ce dernier, tous les éléments sont des couples.
(2) La classe de l'élément $\delta_{(v,w)}\in F(V\times W)$ sera d'habitude noté $v\otimes w$.
(3) Pour insister sur la notion de classe, nous allons aussi noter $[x]$ la classe de $x\in F(V\times W)$.
(4) L'arithmétique dans $V\otimes W$ est relativement simple. En ajoutant et soustrayant le même élément de A_3 nous avons par exemple

$$(\lambda v)\otimes w=(\lambda v)\otimes w+\lambda(v\otimes w)-(\lambda v)\otimes w. \tag{10.458}$$

ITEMooPVWHooMkgQoT

Nous obtenons de cette façon

$$\lambda(v\otimes w)=(\lambda v)\otimes w=v\otimes(\lambda w), \tag{10.459}$$

que nous noterons $\lambda v\otimes w$ sans plus de précision.

PROPooIWZDooRRZNCf

Proposition 10.172 ([345]).
L'espace vectoriel $V\times W$ muni de

$$\begin{aligned}h\colon V\oplus W&\to V\otimes W\\(v,w)&\mapsto v\otimes w\end{aligned} \tag{10.460}$$

est un produit tensoriel entre V et W.

Démonstration. Nous devons prouver les conditions de la définition 10.165.

(1) h **est bilinéaire**
Ce sont des calculs tels que faits dans la remarque 10.171(4) qui font le travail.
(2) h **est surjective**
Un élément de $V\otimes W$ est la classe d'un élément de $F(V\times W)$, c'est-à-dire de la forme

$$\Big[\sum_{i\alpha}\delta_{(v_i,w_\alpha)}\Big]=\sum_{i\alpha}v_i\otimes w_\alpha. \tag{10.461}$$

Cet élément est dans l'image de h comme le montre le calcul suivant[82] :

$$h\Big(\sum_{i\alpha}(v_i,w_\alpha)\Big)=\sum_{i\alpha}h(v_i,w_\alpha)=\sum_{i\alpha}v_i\otimes w_\alpha. \tag{10.462}$$

80. À part mention du contraire, tous les corps du Frido sont commutatifs.
81. Espace vectoriel librement engendré, voir la définition 4.28 et le lemme 4.28.
82. Faites bien la distinction entre $\delta_{v,w}$, (v,w) et $v\otimes w$. Sachez dans quel ensemble se trouvent chacun de ces trois objets.

(3) **Propriété universelle**

Soient un espace vectoriel U et une application linéaire $f\colon V\oplus W\to U$. Nous devons trouver une application linéaire $g\colon V\otimes W\to U$ telle que $f=g\circ h$. Pour cela nous commençons par considérer l'application

$$\begin{aligned} g\colon F(V\times W)&\to U\\ \delta_{(v,w)}&\mapsto f(v,w) \end{aligned} \tag{10.463}$$

définie sur tout $F(V\times W)$ par linéarité sans encombres parce que les $\delta_{v,w}$ forment une base par le lemme 4.28.

Nous démontrons que $g(N)=0$ pour avoir le droit de passer g aux classes et le considérer comme application partant de $V\otimes W$ au lieu de $F(V\times W)$. Prenons par exemple

$$\begin{aligned} g\big(\delta_{(v_1,w)}+\delta_{(v_2,w)}-\delta_{(v_1+v_2,w)}\big)&=g(\delta_{(v_1,w)})+g(\delta_{(v_2,w)})-g(\delta_{v_1+v_2,w}) &\text{(10.464a)}\\ &=f(v_1,w)+f(v_2,w)-f(v_1+v_2,w) &\text{(10.464b)}\\ &=0 &\text{(10.464c)} \end{aligned}$$

par la bilinéarité de f. Cela montre que $g(A_1)=0$. Nous montrons de même que $g(A_2)=g(A_3)=g(A_4)=0$, et enfin toujours par linéarité que $g(N)=0$. Pour rappel, les éléments de N sont les combinaisons linéaires finies d'éléments de A_1, A_2, A_3 et A_4.

Par passage aux classes, nous avons une application (que nous notons également g)

$$g\colon F(V\times W)/N\to U \tag{10.465}$$

vérifiant $g(v\otimes w)=f(v,w)$. Mais comme $h(v,w)=v\otimes w$, nous avons $g\circ h\colon V\oplus W\to U$ vérifiant $g\circ h=f$.

L'espace vectoriel $V\otimes W$ est donc un produit tensoriel. □

10.173.

Vu que $V\otimes W$ est un produit tensoriel de V et W, et vu qu'il y a unicité par la proposition 10.167, nous avons bien le droit de dire que $V\otimes W$ est *le* produit tensoriel. Cela justifie le titre.

10.174.

Les prochains lemmes et propositions vont nous dire que l'application

$$\begin{aligned} \varphi\colon V^*\otimes W&\to\mathcal{L}(V,W)\\ \alpha\otimes w&\mapsto\big(v\mapsto\alpha(v)w\big) \end{aligned} \tag{10.466}$$

est un isomorphisme d'espaces vectoriels lorsque V est de dimension finie. Vu que nous aimons les énoncés très explicites, ça va être découpé en plusieurs morceaux, l'énoncé va devenir un peu long ; mais c'est pour la bonne cause.

LEMooOJEBooQruWEp

Lemme 10.175.

Soient deux espaces vectoriels V et W dont W est de dimension finie. Alors l'application définie par

$$\begin{aligned} \varphi\colon F(V^*\times W)&\to\mathcal{L}(V,W)\\ \delta_{(\alpha,w)}&\mapsto\big(v\mapsto\alpha(v)w\big) \end{aligned} \tag{10.467}$$

sur la base « canonique » de $F(V^*\times W)$ passe aux classes.

Démonstration. Avec les notations de la définition 10.170 nous devons prouver que $\varphi(N)=0$. Nous montrons que $\varphi(A_4)=0$, et nous vous laissons faire les autres. Pour $\lambda\in\mathbb{K}$, $\alpha\in V^*$ et $w\in W$ en utilisant la linéarité de φ nous avons :

$$\begin{aligned} \varphi\big(\lambda\delta_{(\alpha,w)}-\delta_{(\alpha,\lambda w)}\big)v&=\lambda\varphi(\delta_{(\alpha,w)})(v)-\varphi(\delta_{(\alpha,\lambda w)})(v) &\text{(10.468a)}\\ &=\lambda\alpha(v)w-\alpha(v)(\lambda w) &\text{(10.468b)}\\ &=0 &\text{(10.468c)} \end{aligned}$$

parce que $\alpha(v)(\lambda w) = \lambda\alpha(v)w$ du fait que $\mathbb{K}$ est commutatif. La commutativité de $\mathbb{K}$ est ce qui permet de permuter le produit $\lambda\alpha(v)$.

Nous laissons à la lectrice le soin de prouver que $\varphi(A_1) = \varphi(A_2) = \varphi(A_3) = 0$. □

LEMooUQZHooWjIGsy

Lemme 10.176.
Si W est de dimension finie, alors $\mathcal{L}(V,W)$ muni de

$$\begin{aligned} h\colon V^* \oplus W &\to \mathcal{L}(V,W) \\ (\alpha, w) &\mapsto \big(v \mapsto \alpha(v)w\big) \end{aligned} \tag{10.469}$$

est un produit tensoriel[83] *de V^* par W.*

Démonstration. Nous devons prouver que

— h est bilinéaire,

— h est surjective

— pour tout espace vectoriel U, et pour toute application bilinéaire $f\colon V^* \oplus W \to U$, il existe une application linéaire $g\colon \mathcal{L}(V,W) \to U$ tel que $f = g \circ h$.

(1) **Bilinéaire** Le fait que h soit bilinéaire est une simple vérification.

(2) **Surjective** L'espace W étant de dimension finie, nous pouvons en considérer une base $\{z_i\}_{i\in I}$. Soit $\alpha \in \mathcal{L}(V,W)$. Si $v \in V$, l'élément $\alpha(v)$ peut être décomposé dans la base $\{z_i\}$, ce qui définit des applications linéaires $\alpha_i\colon V \to \mathbb{K}$ par

$$\alpha(v) = \sum_{i\in I} \alpha_i(v) z_i. \tag{10.470}$$

Notons que $\alpha_i \in V^*$. En comparant avec la définition de h, nous voyons que

$$\alpha(v) = \sum_i h(\alpha_i, z_i)(v), \tag{10.471}$$

c'est-à-dire $\alpha = \sum_i h(\alpha_i, w_i) = h\big(\sum_i(\alpha_i, z_i)\big)$. Nous avons donc bien $\alpha \in h(V^* \oplus W)$.

(3) **Propriété universelle**
Soient un espace vectoriel U et une application bilinéaire $f\colon V^* \oplus W \to U$. Pour $\alpha \in \mathcal{L}(V,W)$ nous définissons $g(\alpha)$ comme suit. D'abord nous écrivons α sous la forme

$$\alpha(v) = \sum_i \alpha_i(v) z_i, \tag{10.472}$$

et nous posons

$$g(\alpha) = \sum_i f(\alpha_i, z_i). \tag{10.473}$$

Avec cette définition, en posant $w = \sum_i w_i z_i$, nous avons

$$\begin{aligned} (g \circ h)(\alpha, w) &= g\big(v \mapsto \alpha(v)w\big) && \text{(10.474a)} \\ &= g\big(v \mapsto \sum_i \alpha(v) w_i z_i\big) && \text{(10.474b)} \\ &= \sum_i f(w_i\alpha, z_i) && \text{(10.474c)} \\ &= \sum_i f(\alpha, w_i z_i) && \text{(10.474d)} \\ &= f(\alpha, \sum_i w_i z_i) && \text{(10.474e)} \\ &= f(\alpha, w). && \text{(10.474f)} \end{aligned}$$

Cela prouve que $g \circ h = f$.

83. Définition 10.165.

□

PROPooKJTCooVTXWAQ

Proposition 10.177 ([347]).
Soient deux espaces vectoriels V et W dont V est de dimension finie. Alors l'application

$$\begin{aligned}\varphi\colon V^*\otimes W &\to \mathcal{L}(V,W)\\ \alpha\otimes w &\mapsto \big(v\mapsto \alpha(v)w\big)\end{aligned} \tag{10.475}$$

est bien définie[84] *et est un isomorphisme d'espaces vectoriels.*

Démonstration. Le lemme 10.176 donne une structure de produit tensoriel de V^* par W sur $\mathcal{L}(V,W)$. Rappelons les structures :

$$\begin{aligned}h\colon V^*\oplus W &\to V^*\otimes W\\ (\alpha,w) &\mapsto \alpha\otimes w\end{aligned} \tag{10.476}$$

et

$$\begin{aligned}h'\colon V^*\oplus W &\to \mathcal{L}(V,W)\\ (\alpha,w) &\mapsto \big[v\mapsto \alpha(v)w\big].\end{aligned} \tag{10.477}$$

La proposition 10.167 a déjà fait tout le boulot. La seule chose à faire est de vérifier qu'il existe une application $\varphi\colon V^*\otimes W\to\mathcal{L}(V,W)$ vérifiant simultanément les deux conditions suivantes :

ITEMooVNNSooNIXRoG

(1) $\varphi(\alpha\otimes w)=\big[v\mapsto\alpha(v)w\big]$
(2) $h'=\varphi\circ h$.

La seconde condition assure que φ sera un isomorphisme d'espaces vectoriels.

L'existence de φ vérifiant la condition (1) est un effet du lemme 10.175 qui donne une fonction sur $F(V^*\times W)$ dont le φ qui nous concerne est un quotient. Il reste à voir que cette application vérifie $h'=\varphi\circ h$.

En nous rappelant que $\alpha\otimes w=[\delta_{(\alpha,w)}]$ et en écrivant φ à la fois l'application et son passage au quotient,

$$(\varphi\circ h)(\alpha,w)=\varphi(\alpha\otimes w)=\varphi\big([\delta_{(\alpha,w)}]\big)=\varphi(\delta_{(\alpha,w)}). \tag{10.478}$$

En appliquant à $v\in V$ nous avons :

$$(\varphi\circ h)(\alpha,w)v=\varphi(\delta_{(\alpha,w)})v=\alpha(v)w=h'(\alpha,w)v. \tag{10.479}$$

Et voilà. Nous avons $\varphi\circ h=h'$. □

Une conséquence de la proposition 10.177 est que

$$\dim(V\otimes W)=\dim(V)\dim(W) \tag{10.480}$$

via le lemme 4.42(2).

10.12.4 Bases

LEMooXFIMooDkTSrq

Lemme 10.178 ([1]).
Si $\tau\colon V_1\to V_2$ est un isomorphisme d'espaces vectoriels, alors il existe un isomorphisme d'espaces vectoriels $\varphi\colon V_1\otimes W\to V_2\otimes W$ tel que $\varphi(v\otimes w)=\tau(v)\otimes w$[85].

84. Au sens où il existe une fonction φ définie sur tout $V^*\otimes W$ qui se réduit à cela pour les éléments de la forme $\alpha\otimes w$.

85. La proposition 10.179 dira que cette condition fixe complètement φ, mais c'est une autre histoire qui vous sera contée une autre fois.

Démonstration. L'application

$$\begin{aligned}\varphi_0\colon F(V_1\times W)&\to F(V_2\times W)\\ \delta_{(v,w)}&\mapsto \delta_{\big(\tau(v),w\big)}\end{aligned}\tag{10.481}$$

est un isomorphisme.

Cette application passe aux classes, mais pas au sens où $x\in[y]$ impliquerait $\varphi_0(x)=\varphi_0(y)$; au sens où si $x\in[y]$, alors $\varphi_0(x)\in[\varphi_0(y)]$. Par exemple

$$\varphi_0\big(\lambda\delta_{(v,w)}-\delta_{(v,\lambda w)}\big)=\lambda\delta_{\big(\tau(v),w\big)}-\delta_{\big(\tau(v),\lambda w\big)}\in[0].\tag{10.482}$$

Nous vous laissons le soin de vérifier les égalités correspondantes pour les autres parties de N.

Le passage au classes de φ_0 signifie que l'on considère l'application

$$\begin{aligned}\varphi\colon V_1\otimes W&\to V_2\otimes W\\ [x]&\mapsto[\varphi_0(x)]\end{aligned}\tag{10.483}$$

où vous aurez noté que la prise de classe à gauche n'est pas la même que celle à droite.

Il faut prouver que ce φ est un isomorphisme. En ce qui concerne la linéarité,

$$\begin{aligned}\varphi\big([x]+[y]\big)&=\varphi\big([x+y]\big)&(10.484a)\\ &=[\varphi_0(x+y)]&(10.484b)\\ &=[\varphi_0(x)+\varphi_0(y)]&(10.484c)\\ &=[\varphi_0(x)]+[\varphi_0(y)]&(10.484d)\\ &=\varphi([x])+\varphi([y]).&(10.484e)\end{aligned}$$

Je vous laisse le reste de la linéarité. Et en ce qui concerne le fait que ce soit une bijection, allez-y. □

PROPooTHDPooWgjUwk

Proposition 10.179 ([347])**.**
Soient des espaces vectoriels de dimension finie V et W. Soient une base $\{e_i\}$ de V et une base $\{f_\alpha\}$ de W.

Alors :

ITEMooQCILooUncdGl

(1) La partie $\{e_i\otimes f_\alpha\}$ est une base de $V\otimes W$.

ITEMooGCRHooJjAYPr

(2) Au niveau des dimensions, $\dim(V\otimes W)=\dim(V)\dim(W)$.

Démonstration. Vu que V est de dimension finie, nous avons un isomorphisme d'espaces vectoriels $V^*=V$, et même un isomorphisme d'espaces vectoriels

$$\begin{aligned}\tau\colon V&\to(V^*)^*\\ \tau(v)\mu&=\mu(v).\end{aligned}\tag{10.485}$$

Recopions l'isomorphisme de la proposition 10.177 en utilisant V^* au lieu de V :

$$\begin{aligned}\psi_0\colon(V^*)^*\otimes W&\to\mathcal{L}(V^*,W)\\ \tau(v)\otimes w&\mapsto\big(\mu\mapsto\tau(v)(\mu)w=\mu(v)w\big).\end{aligned}\tag{10.486}$$

En écrivant cela, nous avons tenu compte du fait que tout élément de $(V^*)^*$ peut être écrit de façon univoque sous la forme $\tau(v)$ pour un certain $v\in V$.

Vu que τ est un isomorphisme, l'application suivante est encore un isomorphisme [86] :

$$\begin{aligned}\psi\colon V\otimes W&\to\mathcal{L}(V^*,W)\\ v\otimes w&\mapsto\big(\mu\mapsto\mu(v)w\big).\end{aligned}\tag{10.487}$$

EQooAEFRooPfmAnj

86. Lemme 10.178.

Nous avançons. Vu que nous avons un isomorphisme, nous pouvons faire passer des bases. Le lemme 4.42 nous donne une base de $\mathcal{L}(V^*, W)$ en les éléments $\beta_{i\alpha}\colon V^* \to W$ définies par

$$\beta_{ij}(\mu) = \mu(e_i)f_\alpha. \tag{10.488}$$

Donc $\{\psi^{-1}(\beta_{i\alpha})\}$ est une base de $V \otimes W$.

Pour $a = \sum_i a_i e_i^*$ (base duale, définition 4.127) nous avons :

$$\psi(e_i \otimes f_\alpha)a = a(e_i)f_\alpha = \beta_{i\alpha}(a). \tag{10.489}$$

Cela prouve que $\psi^{-1}(\beta_{i\alpha}) = e_i \otimes f_\alpha$, et donc que ces $e_i \otimes f_\alpha$ est une base de $V \otimes W$.

La formule concernant les dimensions est simplement la définition 4.16 de la dimension : le nombre d'éléments dans une base. □

LEMooYJIQooRCkHMq

Lemme 10.180.
Dans le produit tensoriel $\mathbb{R} \otimes \mathbb{R}$, nous avons

(1) $x \otimes 1 = 1 \otimes x = x(1 \otimes 1)$ pour tout $x \in \mathbb{R}$.

(2) Si $x \geqslant 0$ nous avons aussi $x \otimes 1 = \sqrt{x} \otimes \sqrt{x}$.

LEMooKPWNooNjkAru

Lemme 10.181.
Le produit tensoriel est associatif.

Plus précisément, si U, V, W sont des espaces vectoriels, nous avons des isomorphismes d'espaces vectoriels

$$U \otimes (V \otimes W) = (U \otimes V) \otimes W. \tag{10.490}$$

Dans la suite, nous écrirons $U \otimes V \otimes W$ sans plus de précisions et même $\otimes^k V$ pour le produit de k copies de V.

Définition 10.182.
*Un k-**tenseur** sur V est un élément de $\otimes^k V^*$.*

PROPooIODGooYajpiy

Proposition 10.183.
Il y a un isomorphisme d'espaces vectoriels $\otimes^k V^* \to \mathcal{L}_k(V, \mathbb{R})$.

Exemple 10.184.
Pour $k = 2$, l'isomorphisme de la proposition 10.183 est

$$\begin{aligned} \psi\colon \mathcal{L}(V, V^*) &\to \mathcal{L}_2(V, \mathbb{R}) \\ \psi(\alpha)(u, v) &= \alpha(u)v. \end{aligned} \tag{10.491}$$

Il faut se rappeler que le proposition 10.177 donne déjà un isomorphisme entre $V^* \otimes V^*$ et $\mathcal{L}(V, V^*)$. △

10.12.5 Construction par les formes multilinéaires

LEMooCXPRooGfJMMV

Lemme 10.185.
Soient des espaces vectoriels de dimension finie V, W et Z. Soient une application bilinéaire $f\colon V \times W \to Z$ ainsi que des bases $\{e_i\}$ de V et $\{d_j\}$ de W.

Il existe des éléments f_{ij} dans Z tels que

$$f(v, w) = \sum_{ij} v_i w_j f_{ij} \tag{10.492}$$

pour tout $v \in V$ et $w \in W$.

10.186.
Ce que nous allons noter $h(T, S)$ est souvent noté $T \otimes S$, parce que la proposition 10.192 montrera qu'il s'agit bien d'un produit tensoriel. Cependant nous avons besoin de pas mal de propriétés de h avant de pouvoir l'affirmer.

DEFooUUMYooOUCzWk

Proposition-Définition 10.187.
Si $T \in \mathcal{L}_k(V,\mathbb{R})$ et $S \in \mathcal{L}_l(V,\mathbb{R})$, nous définissons

$$\begin{aligned} h(T,S)\colon V^{k+l} &\to \mathbb{R} \\ (v_1,\dots,v_{k+l}) &\mapsto T(v_1,\dots,v_k)S(v_{k+1},\dots,v_{k+l}). \end{aligned} \tag{10.493}$$

(1) L'application $h(T,S)$ est multilinéaire[87] *: $h(T,S) \in \mathcal{L}_{k+l}(V,\mathbb{R})$.*
(2) L'application h est associative : $h\big(h(T,S),R\big) = h\big(T,h(S,R)\big)$.
(3) L'application $h\colon \mathcal{L}_k(V,\mathbb{R}) \times \mathcal{L}_l(V,\mathbb{R}) \to \mathcal{L}_{k+l}(V,\mathbb{R})$ est bilinéaire.

Démonstration. En plusieurs parties.

(1) **multilinéaire** Nous fixons des $v_i \in V$, nous considérons $h(T,S)_i$ et nous calculons (en supposant $i \leqslant k$, sinon on adapte)

$$h(T,S)_i(v)a = h(T,S)(v_1,\dots,v,\dots,v_{k+l}) \tag{10.494a}$$
$$= T(v_1,\dots,v,\dots,v_k)S(v_{k+1},\dots,v_{k+l}) \tag{10.494b}$$
$$= T_i(v)S(\dots). \tag{10.494c}$$

Vu que T_i est linéaire, nous en déduisons que $h(T,S)_i$ est linéaire et donc que $h(T,S)$ est multilinéaire.

Associativité Facile de vérifier que

$$h\big(h(T,S),R\big)(v_1,\dots,v_{k+l+m}) = T(v_1,\dots,v_k)S(v_{k+1},\dots,v_{k+l})R(v_{k+l+1},\dots,v_{k+l+m}). \tag{10.495}$$

Bilinéaire En utilisant directement la définition, on vérifie que $h(T,\lambda S) = \lambda h(T,S)$ et $h(T,S_1+S_2) = h(T,S_1) + h(T,S_2)$.

□

10.188 (Quelques notations)**.**
Si V_i sont des espaces vectoriels, si $v_i \in V_i$, et si $\xi_i \in V_i^*$, nous notons

$$h(\xi_1,\dots,\xi_n)(v_1,\dots,v_n) = \prod_{i=1}^{n} \xi_i(v_i). \tag{10.496}$$

Si I est un multiindice ($I = (i_1,\dots,i_n)$), et si v est quelque chose qui peut avoir des indices, alors nous écrivons $v_I = (v_{i_1},\dots,v_{i_n})$.

Nous écrivons en particulier $\delta_{IJ} = \prod_{k=1}^{n} \delta_{i_k j_k}$.

LEMooPUNAooUnQTpJ

Lemme 10.189 (égalité sur une base[1])**.**
Soient des espaces vectoriels $(V_i)_{i=1,\dots,N}$ de dimensions n_i. Soient deux formes N-linéaires $S,T\colon V_1 \times \dots \times V_N \to \mathbb{R}$. Si pour tout multiindice I nous avons $S(e_I) = T(e_I)$, alors $S = T$.

Si les ξ_i sont des formes linéaires, nous notons ξ_I la forme multilinéaire $h(\xi_1,\dots,\xi_n)$.

Démonstration. Nous considérons les bases $\{e_i^{(k)}\}_{i=1,\dots,n_i}$ de V_k. Un élément générique de $V_1 \times \dots V_N$ est de la forme

$$\Big(\sum_{j_1}^{n_1} x_{j_1}^{(1)} e_{j_1}^{(1)},\dots,\sum_{j_N=1}^{n_N} x_{j_N}^{(N)} e_{j_N}^{(N)}\Big). \tag{10.497}$$

Nous appliquons S à cela en tenant compte de la multilinéarité :

$$S\Big(\sum_{j_1}^{n_1} x_{j_1}^{(1)} e_{j_1}^{(1)},\dots,\sum_{j_N=1}^{n_N} x_{j_N}^{(N)} e_{j_N}^{(N)}\Big) = \sum_{j_1=1}^{n_1} \dots \sum_{j_N=1}^{n_N} \prod_{i=1}^{N} x_{j_i}^{(i)} S(e_{j_1}^{(1)},\dots,e_{j_N}^{(N)}) \tag{10.498a}$$
$$= \sum_{J} \prod_{i=1}^{N} x_{j_i}^{(i)} S(e_J). \tag{10.498b}$$

87. Application multilinéaire, définition 8.1.

Le même calcul avec T donne la même chose, compte tenu de l'hypothèse $S(e_J) = T(e_J)$. □

THOooTAGKooDscwFG

Théorème 10.190 ([262]).
Soient des espaces vectoriels $(V_i)_{i=1,\dots,N}$ de dimensions n_i. Nous notons, pour chaque i une base $\{e_j^{(i)}\}_{j=1,\dots,n_i}$ de V_i. Nous notons $\{\alpha_j^{(i)}\}$ les bases duales.
Alors $\{\alpha_I\}_I$ est une base de $\mathcal{L}_N(V_1 \times \dots V_N, \mathbb{R})$.

Démonstration. Nous commençons par prouver que $\{\alpha_I\}$ est libre. Supposons que $\sum_I a_I\alpha_I = 0$. En l'appliquant à e_J, nous avons

$$0 = \sum_I a_I\alpha_I(e_J) = \sum_I a_I\delta_{IJ} = a_J. \tag{10.499}$$

Cela prouve que tous les a_I sont nuls, et donc que la partie est libre.

Pour prouver que la partie est génératrice nous considérons $T \in \mathcal{L}_N(V_1 \times \dots \times V_N, \mathbb{R})$ et nous posons $T_I = T(e_I)$ pour tout multiindice I. En posant $S = \sum_I T_I\alpha_I$, nous

$$S(e_J) = \sum_I T_I\alpha_I(e_J) = \sum_I T_I\delta_{IJ} = T_J. \tag{10.500}$$

Donc $S(e_J) = T(e_J)$. Le lemme 10.189 conclu. □

LEMooPBLRooJePqlk

Lemme 10.191.
Soient deux espaces vectoriels de dimension finie V et W. Nous notons $\{\alpha_i\}$ une base de V^* et $\{\beta_j\}$ une base de W^*. Pour chaque $\xi \in V^*$ et $\sigma \in W^*$ nous considérons l'application bilinéaire

$$\begin{aligned} h(\xi,\sigma)\colon V \times W &\to \mathbb{R} \\ (v,w) &\mapsto \xi(v)\sigma(w). \end{aligned} \tag{10.501}$$

La partie $\{h(\alpha_i,\beta_j)\}$ forment une base de l'espace des applications bilinéaires sur $V \times W$.

PROPooIFVBooGiMskq

Proposition 10.192 ([1, 348]).
Nous considérons l'application $h\colon V^* \times W^* \to \mathcal{L}_2(V \times W, \mathbb{R})$ donnée par

$$h(\xi,\sigma)(v,w) = \xi(v)\sigma(w). \tag{10.502}$$

Le couple $\big(\mathcal{L}_2(V \times W, \mathbb{R}), h\big)$ est un produit tensoriel de V^* et W^*.

Démonstration. Nous devons prouver que l'application h satisfait à la propriété universelle 10.165(2). Nous considérons une base $\{\alpha_i\}$ de V^* et $\{\beta_j\}$ de W^*. Si $f\colon V^* \times W^* \to Z$ est bilinéaire, le lemme 10.185 nous dit que f est déterminée par les éléments $f_{ij} = f(\alpha_i,\beta_j)$.

(1) **Existence** Nous posons

$$\begin{aligned} \tilde{f}\colon \mathcal{L}_2(V \times W, \mathbb{R}) &\to Z \\ \sum_{ij} a_{ij}h(\alpha_i,\beta_j) &\mapsto \sum_{ij} a_{ij}f_{ij}. \end{aligned} \tag{10.503}$$

Notez deux choses.

(1a) C'est une bonne définition parce que les $h(\alpha_i,\beta_j)$ forment une base de $\mathcal{L}_2(V \times W, \mathbb{R})$ (lemme 10.191).

(1b) L'application $\tilde{f}$ est linéaire.

Nous devons prouver que $\tilde{f} \circ h = f$. Soient $\xi = \sum_i \xi_i\alpha_i$ et $\sigma = \sum_j \sigma_j\beta_j$. Par bilinéarité nous avons

$$h(\xi,\sigma) = \sum_{ij} \xi_i\sigma_j h(\alpha_i,\beta_j). \tag{10.504}$$

Nous avons donc

$$(\tilde{f} \circ h)(\xi,\sigma) = \sum_{ij} \xi_i\sigma_j f_{ij} = \sum_{ij} \xi_i\sigma_j(\alpha_i,\beta_j) = f(\xi,\sigma), \tag{10.505}$$

et donc bien $\tilde{f} \circ h = f$.

(2) **Unicité** Nous savons que les éléments $h(\alpha_i, \beta_j)$ forment une base de $\mathcal{L}_2(V \times W, \mathbb{R})$. Si $\tilde{f}\colon \mathcal{L}_2(V \times W, \mathbb{R}) \to Z$ est une application linéaire vérifiant $\tilde{f} \circ h = f$, alors nous avons

$$\tilde{f}\big(h(\alpha_i, \beta_j)\big) = (\tilde{f} \circ h)(\alpha_i, \beta_j) = f(\alpha_i, \beta_j) = f_{ij}. \tag{10.506}$$

Donc l'application $\tilde{f}$ est entièrement déterminée sur une base par les éléments $f_{ij} \in Z$.

□

10.193.
Lorsque V et W sont des espaces vectoriels, si rien n'est précisé, le produit tensoriel

$$V \otimes W \tag{10.507}$$

sera celui de la proposition 10.192, et lorsque $\xi_1 \in V^*$ et $\xi_2 \in W^*$ nous notons

$$\xi_1 \otimes \xi_2 \tag{10.508}$$

l'élément $h(\xi_1, \xi_2)$.

Ce serait mieux de garder la notation $V \otimes W$ pour un produit tensoriel abstrait (définition 10.165), mais bon ...Nous n'allons pas non plus nous tuer sur les notations.

10.12.6 Décomposabilité

DEFooTSSVooUGybzL

Définition 10.194 ([262]).
*Une application $T \in \mathcal{L}_k(V, \mathbb{R})$ est **décomposable** si il existe des formes $\xi_1, \dots, \xi_k \in V^*$ telles que*

$$T(v_1, \dots, v_k) = \prod_{i=1}^{k} \xi_i(v_i) \tag{10.509}$$

autrement dit $T = h(\xi_1, \dots, \xi_k)$, ou encore $T = \xi_1 \otimes \dots \otimes \xi_k$.

LEMooLEIBooROVcim

Lemme 10.195 ([262, 1]).
Une forme 2-multilinéaire $T = \sum_{ij} a_{ij} h(\alpha_i, \alpha_j)$ est décomposable[88] *si et seulement si la matrice a est de rang 1.*

Démonstration. En deux parties.

(1) $\Rightarrow$ Soient $f_1, f_2 \in V^*$ tels que $T(v_1, v_2) = f_1(v_1) f_2(v_2)$ pour tout $v_1, v_2 \in V$. Nous appliquons T à $(e_1), e_2$:

$$T(e_k, e_l) = \sum_{ij} a_{ij} (\alpha_i \otimes \alpha_j)(e_k, e_l) = \sum_{ij} a_{ij} \alpha_i(e_i) \alpha_j(e_l) = a_{kl}, \tag{10.510}$$

mais aussi

$$T(e_k, e_l) = f_1(e_k) f_2(e_l). \tag{10.511}$$

Donc $a_{kl} = f_1(e_k) f_2(e_l)$. Notons A la matrice (et l'application linéaire) formée par ces nombres. Nous avons :

$$A(e_i) = \sum_k a_{ki} e_k = \sum_k f_1(e_k) f_2(e_i) e_k = f_2(e_i) \sum_k f_1(e_k) e_k. \tag{10.512}$$

En notons v le vecteur $\sum_k f_1(e_k) e_k$, nous avons

$$A(\mathbb{R}^n) = \mathrm{Span}(v). \tag{10.513}$$

Cela montre que la dimension de l'image de A est 1, et donc que A est de rang 1.

88. Définition 10.194.

(2) $\Leftarrow$ Nous supposons que a est de rang 1, et nous posons

$$\begin{aligned} T\colon V^2 &\to \mathbb{R} \\ (v,w) &\mapsto \sum_{kl} a_{kl} v_k w_l. \end{aligned} \tag{10.514}$$

Vu que $\mathrm{rk}(a) = 1$, il existe $u \in V$ tel que $a(\mathbb{R}^n) = \mathrm{Span}(u)$. En particulier pour chaque i, il existe λ_i tel que $a(e_i) = \lambda_i u$. Donc

$$a(v) = \sum_i v_i a(e_i) = \sum_i v_i \lambda_i u = (v \cdot \lambda)u, \tag{10.515}$$

et

$$a_{kl} = a(e_k)_l = \lambda_k u_l. \tag{10.516}$$

Avec tout ça,

$$T(v,w) = \sum_{kl} a_{kl} v_k w_l = (\lambda \cdot v)(u \cdot w). \tag{10.517}$$

Donc nous avons $T = f_1 \otimes f_2$ avec

$$\begin{aligned} f_1\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \lambda \cdot x \end{aligned} \tag{10.518}$$

et $f_2(x) = u \cdot x$.

□

10.196.
Il est dit dans [262] que l'étude de la décomposabilité devient très compliquée à partir des 3-tenseurs, et que c'est lié à l'intrication quantique. Personnellement je ne suis pas sûr de ce que ça veut dire, mais si ça vous parle, c'est sans doute intéressant.

10.12.7 Produit tensoriel d'applications linéaires

PROPooKNNJooLSqzID

Proposition 10.197 ([349, 350])**.**
Soit une application bilinéaire $f\colon V \times W \to Z$ avec $\dim(Z) = \dim(V)\dim(W) < \infty$. Nous supposons que Z est engendré par l'image de f :

$$Z = \mathrm{Span}\big(\mathrm{Image}(f)\big). \tag{10.519}$$

Alors f est un produit tensoriel de V et W.

Démonstration. En plusieurs parties.

(1) **Un autre produit tensoriel** Nous considérons un produit tensoriel $h\colon V \times W \to T$ (existence : proposition 10.172). Étant donné que f est bilinéaire, il existe une application linéaire $\tilde{f}\colon T \to Z$ telle que $f = \tilde{f} \circ h$.

$$\begin{array}{ccc} V \times W & \xrightarrow{h} & T \\ {\scriptstyle f}\downarrow & \swarrow_{\tilde{f}} & \\ Z & & \end{array} \tag{10.520}$$

(2) **$\tilde{f}$ est surjective** Nous prouvons que $\tilde{f}$ est surjective. Soit $z \in Z$. Étant donné que Z est engendré par $f(V \times W)$, il existe $a_{ij} \in \mathbb{R}$, $v_i \in V$ et $w_j \in W$ tels que $z = \sum_{ij} a_{ij} f(v_i, w_j)$. En

utilisant la linéarité de $\tilde{f}$ et le fait que $f = \tilde{f} \circ h$ nous avons :

$$z = \sum_{ij} a_{ij} f(v_i, w_j) \tag{10.521a}$$
$$= \sum_{ij} a_{ij} (\tilde{f} \circ h)(v_i, w_j) \tag{10.521b}$$
$$= \sum_{ij} a_{ij} \tilde{f}(v_i \otimes w_j) \tag{10.521c}$$
$$= \tilde{f}\big(\sum_{ij} a_{ij} v_i \otimes w_j\big). \tag{10.521d}$$

Nous voyons donc que z est dans l'image de $\tilde{f}$, et donc que $\tilde{f}$ est surjective.

(3) **Questions de dimensions** L'espace vectoriel T est de dimension $\dim(V)\dim(W)$ (proposition 10.179(2)). Vu que Z est également, par hypothèse, de dimension $\dim(V)\dim(W)$, nous avons une application linéaire surjective entre espaces de même dimensions. Elle est donc une bijection par le corolaire 4.49.

(4) **Conclusion** La proposition 10.169 dit que $\tilde{f} \circ h = f$ est un produit tensoriel.

□

PROPooLFHFooRmTAjY

Proposition 10.198 ([349, 1]).
Soient des espaces vectoriels de dimension finie V_1, V_2, W_1 et W_2. L'application

$$\begin{aligned} s\colon \mathcal{L}(V_1, W_1) \times \mathcal{L}(V_2, W_2) &\to \mathcal{L}(V_1 \otimes V_2, W_1 \otimes W_2) \\ s(A, B)(e_i \otimes e_j) &= Ae_i \otimes Be_j \end{aligned} \tag{10.522}$$

est un produit tensoriel.

C'est toujours cette structure que nous considérons lorsque nous parlerons du produit tensoriel d'applications linéaires.

Démonstration. Notre objectif est d'utiliser la proposition 10.197, c'est-à-dire de montrer que l'image de s engendre $\mathcal{L}(V_1 \otimes V_2, W_1 \otimes W_2)$. Nous considérons des bases pour tous les espaces vectoriels en présence : $\{e_i\}$ pour V_1, $\{e'_j\}$ pour V_2, $\{\epsilon_k\}$ pour W_1 et $\{\epsilon'_l\}$ pour W_2.

La proposition 10.179(1) dit que $\{e_i \otimes e'_j\}$ est une base de $V_1 \otimes V_2$ et que $\{\epsilon_k \otimes \epsilon'_l\}$ est une base de $W_1 \otimes W_2$, de sorte qu'une base de $\mathcal{L}(V_1 \otimes V_2, W_1 \otimes W_2)$ est donnée par les applications [89]

$$\begin{aligned} f_{ij,kl}\colon V_1 \otimes V_2 &\to W_1 \otimes W_2 \\ \sum_{st} a_{st} e_s \otimes e'_t &\mapsto a_{ij} \epsilon_k \otimes \epsilon'_l. \end{aligned} \tag{10.523}$$

Notre objectif est de montrer que ces applications sont dans l'image de s. Pour cela nous considérons les applications $A\colon V_1 \to W_1$ donnée par $A(e_s) = \delta_{si}\epsilon_k$ et $B\colon V_2 \to W_2$ donnée par $B(e'_t) = \delta_{tj}\epsilon'_l$. Nous avons alors

$$s(A, B)\big(\sum_{st} a_{st} e_s \otimes e'_t\big) = \sum_{st} a_{st} A(e_s) \otimes B(e'_t) \tag{10.524a}$$
$$= \sum_{st} \delta_{si}\epsilon_k \otimes \delta_{tj}\epsilon'_l \tag{10.524b}$$
$$= a_{ij}\epsilon_k \otimes \epsilon'_l. \tag{10.524c}$$

Cela prouve que $s(A, B) = f_{ij,kl}$.

Donc l'image de s engendre $\mathcal{L}(V_1 \otimes V_2, W_2 \otimes W_2)$ et la proposition 10.197 conclut que s est un produit tensoriel pour $\mathcal{L}(V_1, W_1)$ et $\mathcal{L}(V_2, W_2)$. □

89. Proposition 4.77(2).

10.12.8 Contraction

DEFooMQDHooBsyVpr

Définition 10.199.
Un (l,k)-tenseur sur V est une application $k+l$-multilinéaire sur $(V^*)^l \times V^k$.

DEFooIOJMooFtSTtK

Proposition-Définition 10.200 ([1]).
Soit un (l,k)-tenseur[90] *T sur V. Soient une base $\{e_i\}$, et sa base duale $\{\alpha_i\}$. Nous considérons une bijection linéaire $A\colon V \to V$, la base $\{Ae_i\}$ et la duale $\{\beta_i\}$.*

Nous considérons l'application d'insertion $\iota\colon V^n \to V^{n+1}$ donnée par

$$\iota_k(v)(v_1,\dots,v_n) = (v_1,\dots,v_{k-1},v,v_k,\dots,v_n). \tag{10.525}$$

Alors :

ITEMooSROIooGtTXCL

(1) Pour tout $r \leqslant l$, pour tout $s \leqslant k$, pour tout $\xi_i \in V^*$ et $v_j \in V$ nous avons

$$\sum_i T(\iota_r(\alpha_i)(\xi_1,\dots,\xi_{l-1}),\iota_s(e_i)(v_1,\dots,v_{k-1})) = \sum_i T(\iota_r(\beta_i)(\xi_1,\dots,\xi_{l-1}),\iota_s(Ae_i)(v_1,\dots,v_{k-1})). \tag{10.526}$$

ITEMooPWLBooOYrluK

(2) En posant

$$\begin{aligned} C^r_s\colon (V^*)^{l-1}\times V^{k-1} &\to \mathbb{R} \\ (\xi_1,\dots,\xi_{l-1},v_1,\dots,v_{k-1}) &\mapsto \sum_i T(\iota_r(\alpha_i)(\xi_1,\dots,\xi_{l-1}),\iota_s(e_i)(v_1,\dots,v_{k-1})), \end{aligned} \tag{10.527}$$

l'application $C^r_s(T)$ est un $(l-1,k-1)$-tenseur.

*Ce $C^r_s(T)$ est la (r,s)-**contraction** de T.*

Démonstration. Le point (2) est facile parce que c'est juste la multilinéarité. Nous faisons le point (2). Nous partons du lemme 4.128 qui dit que $\beta_i = \sum_j A^{-1}_{ij}\alpha_j$. Remarquez que

$$T\big(\iota_r(\sigma)(\xi_1,\dots,\xi_{l-1}),\dots\big) \tag{10.528}$$

est linéaire en σ. Cela nous permet de sortir plein de sommes de T dans le calcul suivant :

$$\begin{aligned} &\sum_i T\Big(\iota_r(\beta_i)(\xi_1,\dots,\xi_{l-1}),\iota_s(Ae_i)(v_1,\dots,v_{k-1})\Big) && \text{(10.529a)}\\ &= \sum_i T\Big(\iota_r(\sum_j A^{-1}_{ij}\alpha_j)(\xi_1,\dots,\xi_{l-1}),\iota_s(\sum_l A_{li}e_l)(v_1,\dots,v_{k-1})\Big) && \text{(10.529b)}\\ &= \sum_{ijl} A^{-1}_{ij}A_{li}T\big(\iota_r(\alpha_j)(\dots),\iota_s(e_l)(\dots)\big) && \text{(10.529c)}\\ &= \sum_i T\big(\iota_r(\alpha_i)(\dots),\iota_s(e_l)(\dots)\big) && \text{(10.529d)} \end{aligned}$$

□

Proposition 10.201 ([262]).
Soient $v,w \in V$ ainsi que $\beta,\gamma,\sigma \in V^*$. Nous avons

$$C^1_2(v\otimes w\otimes w\otimes\beta\otimes\gamma\otimes\sigma) = \gamma(v)w\otimes\beta\otimes\sigma. \tag{10.530}$$

Démonstration. Nous choisissons une base $\{e_i\}$ de V et sa base duale $\{\alpha_i\}$. Ensuite allons appliquer $C^1_2(v\otimes w\otimes\beta\otimes\gamma\otimes\sigma)$ à (ξ,v_1,v_2) avec $\xi \in V^*$, $v_1,v_2 \in V$. Remarquons d'abord que si nous adoptons

90. Définition 10.199.

la décomposition $v = \sum_k v_k e_k$ et $\gamma = \sum_l \gamma_l \alpha_l$, nous avons $\alpha_i(v) = v_i$ et $\gamma(e_i) = \gamma_i$. Voici le calcul :

$$C_1^1(v \otimes w \otimes \beta \otimes \gamma \otimes \sigma)(\xi, v_1, v_2) = \sum_i (v \otimes w \otimes \beta \otimes \gamma \otimes \sigma)(\alpha_i, \xi, v_1, e_i, v_2) \tag{10.531a}$$

$$= \sum_i \alpha_i(v)\xi(w)\beta(v_1)\gamma(e_i)\sigma(v_2) \tag{10.531b}$$

$$= \sum_i v_i\gamma_i\xi(w)\beta(v_1)\sigma(v_2) \tag{10.531c}$$

$$= \gamma(v)(w \times \beta \otimes \sigma)(\xi, v_1, v_2). \tag{10.531d}$$

Pour la dernière ligne, $\sum_i v_i\gamma_i = \gamma(v)$ parce que

$$\gamma(v) = \sum_{kl} \gamma_l\alpha_l(v_k e_k) = \sum_{kl} \gamma_l v_k \delta_{lk} = \sum_k \gamma_k v_k. \tag{10.532}$$

□

10.12.9 Tenseurs symétriques et alternés

Si T est une application k-multilinéaire sur V, et si $\sigma\colon \{1, \dots, k\} \to \{1, \dots, k\}$ est une permutation, nous notons T^σ l'application k-multilinéaire donnée par

$$\begin{aligned} T^\sigma\colon V^k &\to \mathbb{R} \\ (v_1, \dots, v_k) &\mapsto T(v_{\sigma(1)}, \dots, v_{\sigma(k)}). \end{aligned} \tag{10.533}$$

Définition 10.202 ([262]).
Un k-tenseur[91] *T sur V est **symétrique** si pour toute permutation $\sigma\colon \{1, \dots, k\} \to \{1, \dots, k\}$, nous avons*

$$T(v_1, \dots, v_k) = T(v_{\sigma(1)}, \dots, v_{\sigma(k)}), \tag{10.534}$$

c'est-à-dire $T^\sigma = T$.

Définition 10.203 ([262]).
*Un k-tenseur T est **alterné** si $T^\sigma = (-1)^\sigma T$. L'espace vectoriel des k-tenseurs alternés est noté $\Lambda^k V^*$. Nous notons aussi $\Lambda^0 V^* = \mathbb{R}$.*

LEMooJEZYooMmrtgu

Lemme 10.204.
Si T est un k-tenseur, et si σ et π sont des permutations, alors

(1) L'application $T \mapsto T^\sigma$ est linéaire en T.

(2) $(T^\pi)^\sigma = T^{\pi\circ\sigma}$.

(3) $(-1)^\pi(-1)^\sigma = (-1)^{\pi\circ\sigma}$.

LEMooZMTPooKnqSuz

Lemme 10.205.
Si V est un espace vectoriel réel de dimension n, et si B est une base de V, alors le déterminant[92] *$\det_B\colon V^n \to \mathbb{R}$ est un n-tenseur alterné.*

DEFooHAKAooIfbsEy

Définition 10.206.
Si T est un k-tenseur, nous définissons

$$\mathrm{Alt}(T) = \frac{1}{k!} \sum_{\pi \in S_k} (-1)^\pi T^\pi \tag{10.535}$$

où la somme porte sur le groupe symétrique[93].

91. C'est-à-dire une application k-multilinéaire $T\colon V^k \to \mathbb{R}$.
92. Définition 8.10.
93. Groupe symétrique, définition 1.274.

LEMooWYCAooKoFpRu

Lemme 10.207 ([262]).
L'application Alt: $\mathcal{L}_k(V) \to \Lambda^k V^*$ *est une projection linéaire, c'est-à-dire*

ITEMooLADTooNfbVlO

(1) $\text{Alt}\big(\mathcal{L}_k(V)\big) \subset \Lambda^k V^*$

ITEMooQPJKooTvriBv

(2) $\text{Alt}|_{\Lambda^k V*} = \text{Id}$.

ITEMooYOXTooTmzziO

(3) L'application Alt *est linéaire*

Démonstration. Pour (1), nous prouvons que $\big(\text{Alt}(T)\big)^\sigma = (-1)^\sigma \text{Alt}(T)$. Pour cela nous utilisons les formules du lemme 10.204. Nous avons

$$\begin{aligned}
\text{Alt}(T)^\sigma &= \left(\frac{1}{k!}\sum_{\pi\in S_k}(-1)^\pi T^\pi\right)^\sigma && (10.536a)\\
&= \frac{1}{k!}\sum_{\pi\in S_k}(-1)^\pi T^{\pi\circ\sigma} && (10.536b)\\
&= (-1)^\sigma\frac{1}{k!}\sum_{\pi\in S_k}(-1)^\sigma(-1)^\pi T^{\pi\circ\sigma} && (10.536c)\\
&= (-1)^\sigma\frac{1}{k!}\sum_{\pi\in S_k}(-1)^{\pi\circ\sigma} T^{\pi\circ\sigma} && (10.536d)\\
&= (-1)^\sigma\frac{1}{k!}\sum_{\pi\in S_k}(-1)^\pi T^\pi && (10.536e)\\
&= (-1)^\sigma \text{Alt}(T). && (10.536f)
\end{aligned}$$

SUBEQooTWZPooNpAijd

Pour (10.536e), nous avons utilisé le fait que $\pi \mapsto \pi\circ\sigma$ est une bijection de S_k et la proposition 1.310.

Pour le point (2), nous considérons un tenseur alterné T. Il vérifie donc $T^\pi = (-1)^\pi T$. Nous avons alors le calcul suivant :

$$\begin{aligned}
\text{Alt}(T) &= \frac{1}{k!}\sum_{\pi\in S_k}(-1)^\pi T^\pi && (10.537a)\\
&= \frac{1}{k!}\sum_{\pi\in S_k}(-1)^\pi(-1)^\pi T && (10.537b)\\
&= \frac{1}{k!}\sum_{\pi\in S_k} T && (10.537c)\\
&= T && (10.537d)
\end{aligned}$$

parce que S_k contient exactement $k!$ éléments par le lemme 1.277.

Pour le point (3), il s'agit de la linéarité de $T \mapsto T^\pi$ donnée par le lemme 10.204. □

LEMooIMYIooZIxYRp

Lemme 10.208 ([1]).
Si $T \in \mathcal{L}_k(V)$ *et* $S \in \mathcal{L}_l(V)$*, nous avons*

$$\text{Alt}(T\otimes S) = (-1)^{kl}\text{Alt}(S\otimes T). \tag{10.538}$$

Démonstration. Nous introduisons la permutation $\sigma \in S_{k+l}$ qui permute $\{1,\dots,k\}$ avec $\{k+1,\dots,k+l\}$; par exemple pour $k=3$ et $l=1$, ce serait $12345 \to 45123$.

Nous avons

$$\begin{aligned}
(T\otimes S)(v_1,\dots,v_k,u_1,\dots,u_l) &= T(v_1,\dots,v_k)S(u_1,\dots,u_l) && (10.539a)\\
&= (S\otimes T)\big(\sigma(v_1,\dots,v_k,u_1,\dots,u_l)\big). && (10.539b)
\end{aligned}$$

Donc $T\otimes S = (S\otimes T)\circ\sigma$.

Passons au calcul de $\mathrm{Alt}(T\otimes S)$. Nous avons

$$\mathrm{Alt}(T\otimes S) = \frac{1}{k!}\sum_{\pi\in S_{k+l}}(-1)^\pi (T\otimes S)\circ\pi \tag{10.540a}$$
$$= \frac{1}{k!}\sum_{\pi\in S_{k+l}}(-1)^\pi (S\otimes T)\circ\sigma\circ\pi \tag{10.540b}$$
$$= \frac{1}{k!}\sum_{s\in S_{k+l}}(-1)^{\sigma^{-1}\circ s}(S\otimes T)\circ s \tag{10.540c}$$
SUBEQooLJCJooLboUjv
$$= \frac{1}{k!}(-1)^{kl}\sum_s(-1)^s(S\otimes T)\circ s \tag{10.540d}$$
SUBEQooJLNGooEEnzci
$$= (-1)^{kl}\,\mathrm{Alt}(S\otimes T). \tag{10.540e}$$

Justifications :

— Pour (10.540c), nous avons changé de variables dans la somme : $s=\sigma\circ\pi$.

— Pour (10.540d), nous avons utilisé le fait que $(-1)^\sigma=(-1)^{kl}$ parce que qu'il faut kl transpositions pour réaliser σ.

□

LEMooERYDooAifUdG

Lemme 10.209 ([1]).
Soient $T\in\Lambda^k V^*$ et $S\in\Lambda^l V^*$. Nous avons

$$\sum_{\pi\in S_{k+l}}(-1)^\pi(T^\sigma\otimes S)^\pi = (-1)^\sigma\sum_{\pi\in S_{k+k}}(-1)^\pi(T\otimes S)^\pi. \tag{10.541}$$

Démonstration. Dans un premier temps, nous fixons un π, nous oublions le $(-1)^\pi$, et nous ne faisons pas la somme. Nous appliquons à $(v_1,\dots,v_{k+l})\in V^{k+l}$:

$$(T^\sigma\otimes S)^\pi(v_1,\dots,v_{k+l}) = (T^\sigma\otimes S)(v_{\pi(1)},\dots,v_{\pi(k+l)}) \tag{10.542a}$$
$$= T^\sigma(v_{\pi(1),\dots,v_{\pi(k)}})S(v_{\pi(k+1),\dots,v_{\pi(k+l)}}) \tag{10.542b}$$
$$= T\big(v_{\pi(\sigma(1))},\dots,v_{\pi(\sigma(k))}\big)S\big(v_{\pi(k+1)},\dots,v_{\pi(k+l)}\big). \tag{10.542c}$$

Arrêtez vous un moment pour vous assurer d'avoir bien compris pourquoi ce qui arrive dans T sont bien des $v_{\pi(\sigma(i))}$ et non des $v_{\sigma(\pi(i))}$. Une fois que c'est fait, nous définissons $\sigma'\in S_{k+l}$ par

$$\sigma'(x)=\begin{cases}\sigma(x) & \text{si } 1\leqslant x\leqslant k\\ x & \text{si } k+1\leqslant x\leqslant k+l.\end{cases} \tag{10.543}$$
EQooGWUUooNiLifY

Avec ça nous pouvons continuer le calcul :

$$(T^\sigma\otimes S)^\pi(v_1,\dots,v_{k+l})T\big(v_{(\pi\circ\sigma')(1)},\dots,v_{(\pi\circ\sigma'(k))}\big)S\big(v_{(\pi\circ\sigma')(k+1)},\dots,v_{(\pi\circ\sigma'(k+l))}\big) \tag{10.544a}$$
$$= (T\otimes S)^{\pi\circ\sigma'}(v_1,\dots,v_{k+l}). \tag{10.544b}$$
(10.544c)

En résumé, jusqu'ici nous avons prouvé que

$$(T^\sigma\otimes S)^\pi = (T\otimes S)^{\pi\circ\sigma'} \tag{10.545}$$

où σ' est donné en fonction de σ par la formule (10.543).

Nous faisons maintenant le calcul avec la somme et tout (en notant ϵ la signature) :

$$\sum_\pi\epsilon(\pi)(T^\sigma\otimes S)^\pi = \sum_\pi\epsilon(\pi)(T\otimes S)^{\pi\circ\sigma'} \tag{10.546a}$$
$$= \epsilon(\sigma')\sum_\pi\epsilon(\sigma')\epsilon(\pi)(T\otimes S)^{\pi\circ\sigma'} \tag{10.546b}$$
$$= \epsilon(\sigma')\sum_\pi\epsilon(\pi\circ\sigma')(T\otimes S)^{\pi\circ\sigma'} \tag{10.546c}$$
$$= \epsilon(\sigma')\sum_\pi\epsilon(\pi)(T\otimes S)^{\pi}. \tag{10.546d}$$

La dernière ligne est la proposition 1.310 et le fait que $\pi \mapsto \pi \circ \sigma'$ est une bijection de S_{k+l}. □

LEMooKINCooHUEtaT

Lemme 10.210 ([1]).
Soient $T \in \Lambda^k V^*$, $S \in \Lambda^l V^*$ et $R \in \Lambda^m V^*$. Nous avons[94]

$$\mathrm{Alt}\big(\mathrm{Alt}(T \otimes S) \otimes R\big) = \mathrm{Alt}(T \otimes S \otimes R) = \mathrm{Alt}\big(T \otimes \mathrm{Alt}(S \otimes R)\big). \tag{10.547}$$

Démonstration. Nous prouvons la première égalité[95]. Nous avons :

$$\mathrm{Alt}\big(\mathrm{Alt}(T \otimes S) \otimes R\big) = \frac{1}{(k+l+m)!} \sum_{\pi \in S_{k+l+m}} (-1)^\pi \big(\mathrm{Alt}(T \otimes S) \otimes R\big)^\pi \tag{10.548a}$$

$$= \frac{1}{(k+l+l)!} \sum_{\pi \in S_{k+l+m}} (-1)^\pi \Big(\frac{1}{(k+l)!} \sum_{\sigma \in S_{k+k}} (-1)^\sigma (T \otimes S)^\sigma \otimes R\Big)^\pi \tag{10.548b}$$

$$= \frac{1}{(k+l+m)!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} (-1)^\sigma \sum_{\pi \in S_{k+l+m}} (-1)^\pi \big((T \otimes S)^\sigma \otimes R\big)^\pi. \tag{10.548c}$$

C'est le moment d'utiliser le lemme 10.209. Nous continuons :

$$\mathrm{Alt}\big(\mathrm{Alt}(T \otimes S) \otimes R\big) = \frac{1}{(k+l+l)!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} (-1)^\sigma (-1)^\pi \big((T \otimes S) \otimes R\big)^\pi \tag{10.549a}$$

$$= \frac{1}{(k+l+m)!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \sum_{\pi} \big((T \otimes S) \otimes R\big)^\pi. \tag{10.549b}$$

Notez que ce qui est dans la somme sur σ ne dépend pas de σ. Nous avons donc $|S_{k+l}| = (k+k)!$ termes identiques, et nous continuons

$$\mathrm{Alt}\big(\mathrm{Alt}(T \otimes S) \otimes R\big) = \frac{1}{(k+l+m)!} \sum_{\pi \in S_{k+l+m}} (T \otimes S \otimes R)^\pi \tag{10.550a}$$

$$= \mathrm{Alt}(T \otimes S \otimes R). \tag{10.550b}$$

Et voilà. □

DEFooCTSPooZRIufr

Définition 10.211 (Produit extérieur).
Pour $T \in \Lambda^k V^*$ et $S \in \Lambda^l V^*$ nous définissons

$$T \wedge S = \frac{(k+l)!}{k!l!} \mathrm{Alt}(T \otimes S). \tag{10.551}$$

Cette opération est le ***produit extérieur****.*

LEMooKEOWooNDXqgr

Lemme 10.212 ([262]).
Propriétés du produit extérieur.

ITEMooELVSooHlORJy

(1) L'opération $\wedge$ est une application $\wedge : \Lambda^k V^* \times \Lambda^l V^* \to \Lambda^{k+l} V^*$.

ITEMooRUBEooKVHFSz

(2) L'application $\wedge : \Lambda^k V^* \times \Lambda^l V^* \to \Lambda^{k+l} V^*$ est linéaire en ses deux arguments.

ITEMooTCWPooIJYaRE

(3) Nous avons $T \wedge S = (-1)^{kl} S \wedge T$.

ITEMooKFPZooFenmCT

(4) Le produit extérieur est associatif : $(T \wedge S) \wedge R = T \wedge (S \wedge R)$.

Démonstration. En plusieurs parties.

(1) **Pour (1)** Nous savons que $T \otimes S \in \mathcal{L}_{k+l}(V)$. Le lemme 10.207 conclut.

94. Nous avons le droit d'écrire $T \otimes S \otimes R$ parce que le produit tensoriel est associatif, lemme 10.181.
95. Je vous conseille d'essayer de prouver la seconde ; je ne l'ai pas vérifiée.

(2) **Pour (2)** C'est la bilinéarité du produit tensoriel $\otimes$ et celle de l'antisymétrisation Alt donnée dans le lemme 10.207(3).

Notez que la bilinéarité du produit tensoriel $\otimes$ remonte à la définition d'un produit tensoriel (T,h) 10.165 qui demande à h d'être bilinéaire. La proposition 10.187 prouve que le produit $T\otimes S$ entre applications multilinéaires donne bien un produit tensoriel. Bref. Tout cela pour nous rappeler que ce que nous notons $T\otimes S$ pourrait tout aussi bien se noter $h(T,S)$.

(3) **Pour (3)** Conséquence du lemme 10.208.

(4) **Pour (4)** Soient $T\in\Lambda^k V^*$, $S\in\Lambda^l V^*$ et $R\in\Lambda^m V^*$. En déballant les définitions,

$$\begin{aligned}
(T\wedge S)\wedge R &= \frac{(k+l+m)!}{(k+l)!m!}\operatorname{Alt}\big((T\wedge S)\otimes R\big) && (10.552a)\\
&= \frac{(k+l+m)!}{(k+l)!m!}\operatorname{Alt}\left(\frac{(k+l)!}{k!l!}\operatorname{Alt}(T\times S)\otimes R\right) && (10.552b)\\
&= \frac{(k+l+m)!}{k!l!m!}\operatorname{Alt}\big(\operatorname{Alt}(T\otimes S)\otimes R\big) && (10.552c)\\
&= \frac{(k+l+m)!}{k!l!m!}\operatorname{Alt}\big(T\otimes(S\otimes R)\big) \qquad \text{lem. 10.210} && (10.552d)\\
&= T\wedge(S\wedge R). && (10.552e)
\end{aligned}$$

□

LEMooZJJLooFGuguy

Lemme 10.213 ([1]).
Si $\xi_i\in\bigwedge^{n_i}V^*$, alors nous avons[96]

$$\xi_1\wedge\ldots\wedge\xi_k=\frac{(\sum_{i=1}^k n_i)!}{\prod_{i=1}^k(n_i!)}\operatorname{Alt}(\xi_1\otimes\ldots\xi_k). \tag{10.553}$$

Démonstration. Nous y allons par récurrence. Pour $k=1$, l'égalité à prouver se réduit à

$$\xi_1=\operatorname{Alt}(\xi_1), \tag{10.554}$$

qui est vraie par le lemme 10.207(2).

Pour la récurrence nous faisons ceci :

$$\begin{aligned}
(\xi_1\wedge\ldots\wedge\xi_k)\wedge\xi_{k+1} &= \frac{(\sum_{i=1}^k n_i)!}{\prod_{i=1}^k(n_i!)}\operatorname{Alt}(\xi_1\otimes\ldots\xi_k)\wedge\xi_{k+1} && (10.555a)\\
&= \frac{(\sum_{i=1}^k n_i)!}{\prod_{i=1}^k(n_i!)}\frac{(\sum_{i=1}^{k+1} n_i)!}{(\sum_{i=1}^k n_i)!n_{k+1}!}\operatorname{Alt}(\xi_1\otimes\ldots\otimes\xi_{k+1}) && (10.555b)\\
&= \frac{(\sum_{i=1}^k n_i)!}{\prod_{i=1}^k(n_i!)}\operatorname{Alt}(\xi_1\otimes\ldots\xi_k). && (10.555c)
\end{aligned}$$

□

PROPooRRSZooJXOApq

Proposition 10.214 ([262]).
Soient $\xi_1,\ldots,\xi_k\in V^*$ et $v_1,\ldots,v_k\in V$. Nous avons

$$(\xi_1\wedge\ldots\wedge\xi_k)(v_1,\ldots,v_k)=\det\big(\xi_i(v_j)\big). \tag{10.556}$$

96. L'alterné est donné dans la définition 10.206.

Démonstration. Nous pouvons utiliser le lemme 10.213 dans lequel $n_i = 1$ pour tout i : $(\xi_1 \wedge \ldots \wedge \xi_k) = k! \operatorname{Alt}(\xi_1 \otimes \ldots \otimes \xi_k)$. Cela donne le calcul suivant :

$$(\xi_1 \wedge \ldots \wedge \xi_k)(v_1, \ldots, v_k) = k! \operatorname{Alt}(\xi_1 \otimes \ldots \otimes \xi_k)(v_1, \ldots, v_k) \tag{10.557a}$$
$$= k! \frac{1}{k!} \sum_{\pi \in S_k} (-1)^\pi (\xi_1 \otimes \ldots \otimes \xi_k)^\pi (v_1, \ldots, v_k) \tag{10.557b}$$
$$= \sum_{\pi \in S_k} (-1)^\pi \prod_{i=1}^{n} \xi_i(v_{\pi(i)}) \tag{10.557c}$$
$$= \det\big(\xi_i(v_j)\big) \qquad \text{def. 4.79.} \tag{10.557d}$$

Et voila. □

CORooDQSNooZtMuOJ

Corolaire 10.215.
Soient une 1-forme α, une k-forme ω ainsi que des vecteurs $\{v_i\}_{i=1,\ldots,k+1}$. Alors nous avons

SUBEQooTNYNooAeubHJ

$$(\alpha \wedge \omega)(v_1, \ldots, v_{k+1}) = \frac{1}{k!} \sum_{\pi \in S_{k+1}} (-1)^\pi (\alpha \otimes \omega)(v_{\pi(1)}, \ldots, v_{\pi(k+1)}) \tag{10.558a}$$

SUBEQooSNQTooZqxesM

$$= \frac{1}{k!} \sum_{\pi \in S_{k+1}} (-1)^\pi \alpha(v_{\pi(1)}) \omega(v_{\pi(2)}, \ldots, v_{\pi(k+1)}). \tag{10.558b}$$

Démonstration. Il s'agit simplement de déballer les définitions 10.211 et 10.206 :

$$(\alpha \wedge \omega)(v_1, \ldots, v_{k+1}) = \frac{(k+1)!}{k!} \operatorname{Alt}(\alpha \otimes \omega)(v_1, \ldots, v_{k+1}) \tag{10.559a}$$
$$= \frac{(k+1)!}{k!} \frac{1}{(k+1)!} \sum_{\pi \in S_{k+1}} (-1)^\pi (\alpha \otimes \omega)^\pi (v_1, \ldots, v_{k+1}) \tag{10.559b}$$
$$= \frac{1}{k!} \sum_{\pi \in S_{k+1}} (-1)^\pi (\alpha \otimes \omega)(v_{\pi(1)}, \ldots, v_{\pi(k+1)}) \tag{10.559c}$$
$$= \frac{1}{k!} \sum_{\pi \in S_{k+1}} (-1)^\pi \alpha(v_{\pi(1)}) \omega(v_{\pi(2)}, \ldots, v_{\pi(k+1)}). \tag{10.559d}$$

□

10.12.10 Décomposition

PROPooUGLOooTULnDK

Proposition 10.216 ([262]).
Si $\{\alpha_i\}_{i=1,\ldots,n}$ est une base de V^*, alors

$$\{\alpha_{i_1} \wedge \ldots \wedge \alpha_{i_k} \text{ tel que } 1 \leqslant i_1 < \ldots < i_k \leqslant n\} \tag{10.560}$$

est une base de $\bigwedge^k V^*$.

Nous avons

$$\dim(\bigwedge^k V^*) = \binom{n}{k}. \tag{10.561}$$

Démonstration. En plusieurs parties.

(1) **Partie libre** Nous notons $\{e_i\}_{i=1,\ldots,n}$ la base (pré-)duale de V. Pour $I = (i_1, \ldots, i_k)$ avec $i_1 < \ldots < i_k$, nous notons $\Omega_I = \alpha_{i_1} \wedge \ldots \wedge \alpha_{i_k}$. Soit $J = (j_1, \ldots, j_k)$ avec $j_1 < \ldots < j_k$. Nous avons

$$\Omega_I(e_I) = (\alpha_{i_1} \wedge \ldots \wedge \alpha_{i_k})(e_{j_1}, \ldots, e_{j_k}) \tag{10.562a}$$
$$= \operatorname{Alt}(\alpha_{i_1} \otimes \ldots \otimes \alpha_{i_k})(e_{j_1}, \ldots, e_{j_k}) \qquad \text{lem. 10.213} \tag{10.562b}$$
$$= k! \frac{1}{k!} \sum_{\pi \in S_k} (-1)^\pi \alpha_{i_1}(e_{\pi(j_1)}) \ldots \alpha_{i_k}(e_{\pi(j_k)}) \tag{10.562c}$$
$$= \sum_{\pi} (-1)^\pi \delta_{I\pi(J)}. \tag{10.562d}$$

Vu que les indices I et J sont ordonnées, si $\pi \neq \mathrm{Id}$, nous aurons forcément un des $\pi(j_l) \neq i_l$. Donc la somme sur π se réduit au terme $\pi = \mathrm{Id}$:

$$\Omega_I(e_J) = \delta_{IJ}. \tag{10.563}$$

EQooFTUIooDDgxLK

Cela prouve que les Ω_I sont linéairement indépendants.

(2) **Partie génératrice**

Nous devons encore prouver qu'ils sont générateurs de $\bigwedge^k V^*$.
Soit $T \in \bigwedge^k V^*$; en particulier nous avons $T \in \mathcal{L}_k(V, \mathbb{R})$, et le théorème 10.190 nous permet d'écrire $T = \sum_I T_I h(e_I)$ où la somme porte sur les $I = (i_1, \dots, i_k)$ avec $1 \leqslant i_j \leqslant n$ pour tout j, et $h(e_I) = e_{i_1} \otimes \dots \otimes e_{i_k}$.
En utilisant encore le lemme 10.213, nous avons

$$\Omega_I = k! \operatorname{Alt}\big(h(e_I)\big), \tag{10.564}$$

et donc

$$T = \operatorname{Alt}(T) = \sum_I T_I \operatorname{Alt}\big(h(e_I)\big) = \sum_I T_I \frac{1}{k!}\Omega_I. \tag{10.565}$$

Il est temps de noter que si I et I' sont les mêmes indices, mais ordonnés différemment, $\Omega_I = \pm\Omega_{I'}$. Si I est un multiindice, nous notons $r(I)$ le multiindice contenant les mêmes indices que I, mais ordonné. Ensuite nous notons

$$s(I) = \begin{cases} 1 & \text{si } \Omega_{r(I)} = \Omega_I \\ -1 & \text{si } \Omega_{r(I)} = -\Omega_I. \end{cases} \tag{10.566}$$

Avec ça nous avons toujours $\Omega_I = s(I)\Omega_{r(I)}$, et donc

$$T = \frac{1}{k!} \sum_I T_I s(I) \Omega_{r(I)}. \tag{10.567}$$

Donc T est bien une combinaison linéaire des $\Omega_{r(I)}$.

(3) **Cardinal** Nous posons

$$A = \{I \text{ tel que } 1 \leqslant i_1 < \dots < i_k \leqslant n\} \tag{10.568a}$$
$$B = \{\Omega_I \text{ tel que } I \in A\} \tag{10.568b}$$
$$C = \big\{\text{parties de cardinal } k \text{ dans } \{1, \dots, n\}\big\}. \tag{10.568c}$$

Nous devons prouver que $\operatorname{Card}(B) = \binom{n}{k}$. Le lemme 3.42 donne déjà $\operatorname{Card}(C) = \operatorname{Card}(A) = \binom{n}{k}$. Nous prouvons à présent que

$$\begin{aligned} \psi\colon A &\to B \\ I &\mapsto \Omega_I \end{aligned} \tag{10.569}$$

est une bijection. Cette application est surjective par définition de B. En ce qui concerne l'injectivité, si $\Omega_I = \Omega_J$, alors pour tout multiindice L nous avons $\Omega_I(e_L) = \Omega_J(e_L)$. En vertu de (10.563), cela donne $\delta_{IL} = \delta_{JL}$ pour tout multiindice L. En prenant en particulier $L = I$, nous trouvons $1 = \delta_{IJ}$ et donc $I = J$.

□

10.217.
Si $I = (i_1, \dots, i_k)$ est un multiindice et si $a = (a_1, \dots, a_n)$, alors nous notons

$$a_{\wedge I} = a_{i_1} \wedge \dots \wedge a_{i_k} \tag{10.570}$$

dans tous les cas où la formule a un sens.

PROPooSDVXooLvwPbg

Proposition 10.218.
Soit $k \in \mathbb{N}$. Nous notons

$$C_k = \{(i_1, \ldots, i_k) \in \mathbb{N}^k \text{ tel que } 1 \leqslant i_1 < \ldots < i_k \leqslant n\}. \tag{10.571}$$

Si $\{e_i\}_{i=1,\ldots,n}$ est une base de V, alors tout élément de $\bigwedge^k V$ peut être écrit sous la forme

$$\omega = \sum_{i \in C_k} \omega_I e_{\wedge I} \tag{10.572}$$

avec $\omega_I \in \mathbb{R}$.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

PROPooDXTOooKYDOiI

Proposition 10.219.
Soit un ouvert U dans $\mathbb{R}^n$. Si l'application

$$\omega \colon U \to \bigwedge^k (\mathbb{R}^n)^*$$

est de classe C^k, alors il existe des applications $\omega_I \in C^k(U)$ telles que

$$\omega(s) = \sum_{I \in C_k} \omega_I(s) e^*_{\wedge I}. \tag{10.574}$$

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

10.12.11 Règle de Leibniz pour le produit extérieur

Dans la suite nous allons intensément étudier la partie

$$S_{(k,l)} = \{\pi \in S_{k+l} \text{ tel que } \pi(1) < \ldots < \pi(k),\ \pi(k+1) < \ldots < \pi(k+l)\}. \tag{10.575}$$

Nous commençons avec $k = 1$. Dans ce cas, le lemme 10.220 donnera la forme générale d'un élément de $S_{(1,l)}$. En ce qui concerne l'intuition, si $\tau(1) = m$, toutes les autres valeurs de τ sont fixées à cause de la croissance de $j \mapsto \tau(j)$. Nous avons forcément que, lorsque j va de 2 à $l+1$, le nombre $\tau(j)$ prend dans l'ordre toutes les valeurs entre 1 et $l+1$, en sautant m.

Autrement dit, un élément de $S_{(1,l)}$ doit prendre cette forme :

$$\begin{array}{c|cccccc} j & 1 & 2 & 3 & 4 & 5 \\ \tau(j) & 3 & 1 & 2 & 4 & 5. \end{array} \tag{10.576}$$

La première colonne est séparée parce qu'elle joue un rôle spécial : les nombres sont ordonnés de chacun des deux côtés. À droite, nous avons simplement l'énumération $1, 2, 3, 4, 5$ en sautant 3.

En ce qui concerne la signature, pour tout remettre en ordre, il suffit de bouger le 3 vers sa place. Pour cela il faut 2 transpositions.

LEMooDBIEooYcjlDP

Lemme 10.220 ([1, 351])**.**
À propos de $S_{(1,l)}$.

(1) L'application $\phi \colon \{1, \ldots, l+1\} \to S_{(1,l)}$ donnée par

$$\phi(m)j = \begin{cases} m & si\ j = 1 \\ j-1 & si\ j \leqslant m \\ j & si\ j > m \end{cases} \tag{10.577}$$

EQooRALSooPWvZPd

est une bijection.

ITEMooRIHAooItlkCh
(2) *En ce qui concerne la signature, nous avons* $\epsilon\big(\phi(m)\big) = (-1)^{m-1}$.

ITEMooCOPZooLiUPWE
(3) *Pour l'inverse de* ϕ *nous avons* $\tau = \phi\big(\tau(1)\big)$ *pour tout* $\tau \in S_{(1,l)}$.

ITEMooHMNFooENphGK
(4) *Pour tout* $\tau \in S_{(1,l)}$ *nous avons*

$$\epsilon(\tau) = -(-1)^{\tau(1)} \tag{10.578}$$

Démonstration. Le fait que ϕ prenne effectivement ses valeurs dans $S_{(1,l)}$ est une simple vérification de la croissance.

(1) **Injective** Si $\phi(m) = \phi(n)$, alors en particulier $m = \phi(m)1 = \phi(n)1 = n$ et donc $m = n$.

(2) **Surjective** Soit $\tau \in S_{(1,l)}$. Nous posons $m = \tau(1)$ et nous prouvons que $\tau = \phi(m)$.
Par croissance, si $k \geqslant 2$, nous avons $\tau(k) \leqslant k$. En effet nous avons $\tau(l+1) \geqslant \tau(k) + (l+1-k)$. Au cas où nous aurions $\tau(k) > k$, nous aurions $\tau(l+1) \geqslant \tau(k) + l - 1 - k > l + 1$, ce qui est impossible.
D'autre part, pour tout $k \geqslant 2$ nous avons

EQooZLUJooGgoFqx
$$\tau(k) = \min\Big(\{1, \ldots, l+1\} \backslash \tau\{1, \ldots, k-1\}\Big). \tag{10.579}$$

Cela étant dit, commençons doucement en supposant $m1 = 1$. Dans ce cas l'équation (10.579) donne tout de suite $\tau(2) = 2$. Une récurrence montre qu'alors $\tau(k) = k$ pour tout k. Autrement dit, si $\tau(1) = 1$, nous avons $\tau = \mathrm{Id}$. Simple vérification que $ph(1) = \mathrm{Id}$ aussi.
Passons au cas général $\tau(1) = m > 1$. Nous avons

$$\tau(2) = \min\Big(\{1, \ldots, l+1\} \backslash \{m\}\Big) = 1 \tag{10.580}$$

parce que $m \neq 1$. Nous faisons une récurrence en supposant $\tau(j) = j - 1$ pour un certain $j < m$. Nous prouvons que $\tau(j+1) = j$. Pour cela,

$$\tau(j+1) = \min\Big(\{1, \ldots, l+1\} \backslash \{1, \ldots, j-1, m\}\Big). \tag{10.581}$$

Étant donné que $j < m$, nous avons $j - 1 < m$ et donc le nombre j est encore dans la différence dont nous prenons le minimum. Donc $\tau(j+1) = j$. Cela prouve que $\tau(j) = \phi(m)j$ pour tout $j < m$.
En ce qui concerne $\tau(m)$, nous avons

$$\tau(m) = \min\Big(\{1, \ldots, l+1\} \backslash \{1, \ldots, m-2, m\}\Big) = m - 1. \tag{10.582}$$

De même nous voyons que $\tau(m+1) = m + 1$. De là, la croissance en tenant compte de $\tau(k) \leqslant k$ fait le reste.

(3) **Pour (2)** Nous montrons que

$$\prod_{i=1}^{m-1} (m, i) \circ \phi(m) = \mathrm{Id}\,. \tag{10.583}$$

(3a) **Pour $j = 1$** Nous avons

$$\begin{aligned}
\prod_{i=1}^{m-1} (m, i)\phi(m)1 &= \prod_{i=1}^{m-1} (m, i)m && \text{(10.584a)}\\
&= \prod_{i=2}^{m-1} (m, i) \circ (m, 1)m && \text{(10.584b)}\\
&= \prod_{i=2}^{m-1} (m, i)1 && \text{(10.584c)}\\
&= 1 && \text{(10.584d)}
\end{aligned}$$

(3b) **Si $j \leqslant m-1$** Dans ce cas nous avonns $\phi(m)j = j-1$. Nous avons alors le calcul

$$\prod_{i=1}^{m-1}(m,i)\phi(m)j = \prod_{i=1}^{m-1}(m,i)(j-1) = \prod_{i=j}^{m-1}(m,i)\circ(m,j-1)(j-1) \tag{10.585a}$$

$$= \prod_{i=j}^{m-1}(m,i)m = \prod_{i=j+1}^{m-1}(m,i)(m,j)m = \prod_{i=j+1}^{m-1}(m,i)j = j. \tag{10.585b}$$

(3c) **Pour $j = m$** Même type de calculs :

$$\prod_{i=1}^{m-1}(m,i)\phi(m)m = (m,m-1)(m-1) = m. \tag{10.586}$$

(3d) **Pour $j > m$** Nous avons

$$\prod_{i=1}^{m-1}(m,i)\phi(m)j = \prod_{i=1}^{m-1}(m,i)j = j \tag{10.587}$$

parce que $j > m$, de telle sorte qu'il n'est atteint par aucune des transpositions.

Nous avons donc prouvé que $\prod_{i=1}^{m-1}(m,i)\circ\phi(m) = \mathrm{Id}$. Vu que ϵ est un morphisme[97] et que $\epsilon(\mathrm{Id}) = 1$, nous déduisons que

$$\epsilon\big(\phi(m)\big) = \epsilon\Big(\prod_{i=1}^{m-1}(m,i)\Big) = (-1)^{m-1}. \tag{10.588}$$

(4) **Pour (3)** Soit $\tau \in S_{(1,l)}$. Vu que ϕ est une bijection, il existe un unique $m \in \{1,\dots,l+1\}$ tel que $\tau = \phi(m)$. En appliquant cette égalité à 1 nous avons $\tau(1) = \phi(m)1$. Mais la définition (10.577) de ϕ donne $\phi(m)1 = m$. Nous avons donc prouvé que

$$\tau(1) = \phi(m)1 = m, \tag{10.589}$$

c'est-à-dire $m = \tau(1)$ et donc $\tau = \phi\big(\tau(1)\big)$.

(5) **Pour (4)** Soit $\tau \in S_{(1,l)}$. En utilisant les points (3) et (2), nous avons

$$\epsilon(\tau) = \epsilon\Big(\phi\big(\tau(1)\big)\Big) = (-1)^{\tau(1)-1} = -(-1)^{\tau(1)}. \tag{10.590}$$

□

LEMooFHCHooQgrwHD

Lemme 10.221 ([1]).
L'application

$$\begin{aligned}\varphi\colon B_{(1,l)} &\to S_{(1,l-1)}\\ \pi &\mapsto \phi\big(\pi(1)-1\big)\end{aligned} \tag{10.591}$$

est une bijection et vérifie

$$\epsilon\big(\varphi(\pi)\big) = -\epsilon(\pi). \tag{10.592}$$

Démonstration. En plusieurs parties.

(1) **Surjectif** Lorsque π parcours $B_{(1,l)}$, la valeur de $\pi(1)$ parcours les valeurs de 2 à $l+1$, et donc $\pi(1)-1$ va de 1 à l. Le lemme 10.220 dit que $\phi\big(\{1,\dots,l\}\big) = S_{1,l-1}$.

(2) **Injectif** Si $\varphi(\pi) = \varphi(\sigma)$, alors par injectivité de ϕ (le ϕ de $S_{(1,l-1)}$) nous avons $\pi(1) = \sigma(1)$. Or nous avons $B_{(1,l)} \subset S_{(1,l)}$. L'injectivié du ϕ de $S_{(1,l)}$ dir que si $\pi(1) = \sigma(1)$, alors $\pi = \sigma$.

97. Proposition 1.300.

(3) **Signature** Il s'agit d'abord le lemme 10.220(2). Soit $\pi \in B_{(1,l)}$. En posant $m = \pi(1) - 1$ nous avons

$$\epsilon\big(\varphi(\pi)\big) = \epsilon\Big(\phi\big(\pi(1) - 1\big)\Big) = (-1)^{m-1} = (-1)^{\pi(1)}. \tag{10.593}$$

Ensuite le lemme 10.220(4) nous donne $\epsilon(\pi) = -(-1)^{\pi(1)}$.

□

10.222.
Si I est une partie de $\{1, \dots, n\}$, nous notons $\mathcal{O}(I)$ l'élément de $\mathbb{N}^k$ donné par l'ordonnement des éléments de I. Si E est un ensemble, nous notons $\mathcal{P}_k(E)$ l'ensemble des parties de cardinal k de E.

LEMooKGETooKhomPn

Lemme 10.223 ([1]).
Nous notons $E = \{1, \dots, k + l\}$.

(1) L'application $\phi\colon \mathcal{P}_k(E) \to S_{(k,l)}$ donnée par

$$\phi(I)(1, \dots, l + k) = \big(\mathcal{O}(I), \mathcal{O}(E\backslash I)\big) \tag{10.594}$$

est une bijection.

(2) Pour $\tau \in S_{(k,l)}$ nous avons

$$\epsilon(\tau) = (-1)^{\sum_{n=1}^{k} \big(\tau(n)-n\big)}. \tag{10.595}$$

EQooPWREooEeHcBv

Démonstration. Le fait que $\phi(I)$ soit dans $S_{(k,l)}$ est visible dans la définition : les k premiers éléments de $\big(\phi(I)1, \dots, \phi(I)(k + l)\big)$ et les l suivants sont ordonnés.

(1) **Injective** Soient $I, J \in \mathcal{P}_k(E)$ tels que $\phi(I) = \phi(J)$. En particulier pour $j = 1, \dots, k$ nous avons$\phi(I)j = \phi(J)j$, et donc $\mathcal{O}(I)_j = \mathcal{O}(J)_j$. Cela prouve que $\mathcal{O}(I) = \mathcal{O}(J)$ et donc que $I = J$.

(2) **Surjective** Soit $\tau \in S_{(k,l)}$. Nous notons $I = \tau\{1, \dots, k\}$, et nous prouvons que $\tau = \phi(I)$. Vu que $\tau \in S_{(k,l)}$, le vecteur $\big(\tau(1), \dots, \tau(k)\big)$. Étant donné que τ est une bijection de E, nous avons aussi

$$\tau\{k + 1, \dots, k + l\} = E\backslash I, \tag{10.596}$$

et comme $\big(\tau(k + 1), \dots, \tau(k + l)\big)$ est ordonné, il est égal à $\mathcal{O}(E\backslash I)$.

Nous devons déterminer combien de transpositions sont nécessaires pour transformer $\big(\tau(1), \dots, \tau(k+l)\big)$ en $(1, \dots, k + l)$. Étant donné que $\tau \in S_{(k,l)}$, le plus grand parmi $\big(\tau(1), \dots, \tau(k)\big)$ est $\tau(k)$, il il faut $\tau(k) - k)$ permutations pour l'amener en k. Ce sont les transpositions

$$\prod_{j=k+1}^{\tau(k)} \big(\tau(k), \tau(j)\big). \tag{10.597}$$

Il faut ensuite $\tau(k - 1) - (k - 1)$ transpositions pour ramener $\tau(k - 1)$ en position $k - 1$, etc.

Il faut donc bien composer τ avec $\sum_{n=1}^{k} \big(\tau(n) - n\big)$ transpositions pour obtenir l'identité. □

LEMooEHTLooUSBBKL

Lemme 10.224 ([1]).
Bijection entre $B_{(k,l)}$ et $S_{(k,l-1)}$.

(1) L'application

$$\begin{aligned} \varphi\colon B_{(k,l)} &\to S_{(k,l-1)} \\ \pi &\mapsto \phi\big(\pi\{1, \dots, k\} - 1\big) \end{aligned} \tag{10.598}$$

est une bijection.

(2) Pour tout $\pi \in B_{(k,l)}$, nous avons

$$\epsilon\big(\varphi(\pi)\big) = (-1)^k \epsilon(\pi). \tag{10.599}$$

Démonstration. D'abord notons que φ est bien définie. En vu que $\pi(k+1)=1$, aucun élément de $\pi\{1,\dots,k\}$ n'est égal à 1. Et comme $\pi(j)\leqslant k+l$, les éléments de $\pi\{1,\dots,k\}-1$ sont entre 1 et $k+l-1$. L'ensemble $\pi\{1,\dots,k\}-1$ est donc un argument valable pour l'application $\phi\colon \mathcal{P}_k\{1,\dots,k+l-1\}\to S_{(k,l-1)}$.

(1) **φ est injective** Soient $\pi,\sigma\in B_{(k,l)}$ tels que $\varphi(\pi)=\varphi(\sigma)$. Nous notons $I_\pi=\pi\{1,\dots,k\}$ et $I_\sigma=\sigma\{1,\dots,k\}$. Pour tout $j=1,\dots,k$ nous avons

$$\varphi(\pi)j=\mathcal{O}(I_\pi-1)_j=\mathcal{O}(I_\pi)_j-1, \tag{10.600}$$

et similairement pour σ. Nous en déduisons que $\mathcal{O}(I_\pi)=\mathcal{O}(I_\sigma)$, et donc que $I_\pi=I_\sigma$. Vu que $\pi=\phi(I_\pi)$, nous en déduisons que $\pi=\sigma$.

(2) **Surjectif** Soit $\tau\in S_{(k,l-1)}$. Nous posons $I_\tau=\tau\{1,\dots,k\}$, et nous considérons l'élément $\pi\in B_{(k,l)}$ donné par

$$\pi(j)=\begin{cases}\tau(j)+1 & \text{si } j\leqslant k\\ 1 & \text{si } j=k+1\\ \mathcal{O}\big(E\backslash(I_\tau\cup\{1\})\big)_j & \text{si } j>k+1.\end{cases} \tag{10.601}$$

La dernière ligne signifie « pour les autres, vous prenez ce qui reste, dans l'ordre ». Nous montrons à présent que $\tau=\varphi(\pi)$. Nous avons

$$\varphi(\pi)=\phi\big(\pi\{1,\dots,k\}-1\big)=\phi\big(\tau\{1,\dots,k\}\big). \tag{10.602}$$

Cela prouve que $\varphi(\pi)=\phi(I_\tau)=\tau$.

En ce qui concerne la signature, il s'agit de manipuler la formule (10.595). Vu que $B_{(k,l)}\subset S_{(k,l)}$, elle peut être utilisée tant pour π que pour $\tau(\pi)$. Pour π nous avons[98]

$$|\pi|=\sum_{n=1}^{k}\big(\tau(n)-n\big). \tag{10.603}$$

En ce qui concerne $\varphi(\pi)$ nous avons

$$|\varphi(\pi)|=\sum_{n=1}^{k}\big(\varphi(\pi)n-n\big)=\sum_{n=1}^{k}\big(\pi(n)-1-n\big)=\sum_{n=1}^{k}\big(\pi(n)-n\big)-\sum_{n=1}^{k}1=|\pi|-k. \tag{10.604}$$

□

LEMooRKWHooBNxZBL

Lemme 10.225 ([1]).
Soient des vecteurs $v_1,\dots,v_{k+l}$. Si $\pi\in S_{k+l}$, si $T\in\mathcal{L}_k(V)$ et si $\sigma\in S_k$, alors nous avons

$$T^\sigma(v_{\pi(1)},\dots,v_{\pi(k)})=T\big(v_{\pi\sigma(1)},\dots,v_{\pi\sigma(k)}\big). \tag{10.605}$$

EQooQJXHooTfvclD

Démonstration. Poser $w_j=v_{\pi(j)}$. Nous avons

$$T^\sigma(v_{\pi(1)},v_{\pi(k)})=T^\sigma(w_1,\dots,w_k)=T(w_{\sigma(1)},\dots,w_{\sigma(k)})=T(v_{\pi\sigma(1)},\dots,v_{\pi\sigma(k)}). \tag{10.606}$$

□

LEMooTGNUooZnxkrc

Lemme 10.226 ([63, 1]).
Soient des tenseurs alternés $T\in\Lambda^kV^*$ et $S\in\Lambda^lV^*$. Alors nous avons

$$(T\wedge S)(v_1,\dots,v_{k+l})=\sum_{\pi\in S_{(k,l)}}\epsilon(\pi)T\big(v_{\pi(1)},\dots,v_{\pi(k)}\big)S\big(v_{\pi(k+1)},\dots,v_{\pi(k+l)}\big) \tag{10.607a}$$

$$=\sum_{\pi\in S_{(k,l)}}(T\otimes S)^\pi(v_1,\dots,v_{k+l}). \tag{10.607b}$$

SUBEQooUDLIooWwGFoe

98. Nous notons $|\pi|$ le nombre de transpositions avec lequel il faut composer π pour obtenir l'identité. C'est-à-dire $\epsilon(\pi)=(-1)^{|\pi|}$. Sinon l'expression est trop lourde.

où

$$S_{(k,l)} = \{\pi \in S_{k+l} \text{ tel que } \pi(1) < \ldots < \pi(k),\ \pi(k+1) < \ldots < \pi(k+l)\}, \tag{10.608}$$

et $\epsilon(\pi)$ *est la signature*[99] *de* π*, c'est-à-dire ce que nous notons souvent* $(-1)^{\pi}$*.*

Démonstration. En plusieurs parties.

(1) **Une relation d'équivalence** Nous considérons la relation d'équivalence suivante sur S_{k+l}. Nous posons $\pi_1 \sim \pi_2$ si $\{\pi_1(1), \ldots, \pi_1(k)\} = \{\pi_2(1), \ldots, \pi_2(k)\}$. Pour $\pi_0 \in S_{k+l}$ nous notons

$$A(\pi_0) = \{\pi \in S_{k+l} \text{ tel que } \{\pi(1), \ldots, \pi(k)\} = \{\pi_0(1), \ldots, \pi_0(k)\}\} \tag{10.609}$$

la classe d'équivalence de π_0.

(2) **Permutations partielles**

Lorsque $\pi \sim \pi_0$ nous notons $\sigma^{(\pi,\pi_0)}$ la permutation qui permet de « remettre dans l'ordre π_0 les éléments mélangés par π ». Plus précisément :

$$\begin{aligned} \sigma^{(\pi,\pi_0)} \colon \{1, \ldots, k+l\} &\to \{1, \ldots, k+l\} \\ i &\mapsto \begin{cases} (\pi_0 \circ \pi^{-1})(i) & \text{si } i \in \pi\big(\{1, \ldots, k\}\big) \\ i & \text{sinon.} \end{cases} \end{aligned} \tag{10.610}$$

Notez que $\sigma^{(\pi,\pi_0)}$ est une bijection de $\{1, \ldots, k+l\}$, c'est-à-dire que $\sigma^{(\pi,\pi_0)} \in S_{k+l}$. De même nous considérons $\tau^{(\pi,\pi_0)}$ donné par

$$\begin{aligned} \tau^{(\pi,\pi_0)} \colon \{1, \ldots, k+l\} &\to \{1, \ldots, k+l\} \\ i &\mapsto \begin{cases} (\pi_0 \circ \pi^{-1})(i) & \text{si } i \in \pi\big(\{k+1, \ldots, k+l\}\big) \\ i & \text{sinon.} \end{cases} \end{aligned} \tag{10.611}$$

Ces deux permutations vérifient

$$\pi_0 = \tau^{(\pi,\pi_0)} \circ \sigma^{(\pi,\pi_0)} \circ \pi. \tag{10.612}$$

En effet, si $i = 1, \ldots, k$, nous avons $\sigma^{(\pi,\pi_0)}\pi(i) = \pi_0(i)$ et donc

$$\tau^{(\pi,\pi_0)}\sigma^{(\pi,\pi_0)}\pi(i) = \pi_0(i) \tag{10.613}$$

parce que $\pi_0(i) \notin \pi\big(\{k+1, \ldots, k+l\}\big)$. Même raisonnement pour prouver (10.612) pour les $i = k+1, \ldots, k+l$.

(3) $A(\pi_0) \simeq S_k \times S_l$ Soit $\pi_0 \in S_{k+l}$. Nous notons $P = \pi_0\big(\{1, \ldots, k\}\big)$ et $Q = \pi_0\big(\{k+1, \ldots, k+l\}\big)$. Nous montrons que

$$\begin{aligned} \psi \colon A(\pi_0) &\to S_P \times S_Q \\ \pi &\mapsto \big(\sigma^{(\pi,\pi_0)}|_P, \tau^{(\pi,\pi_0)}|_Q\big) \end{aligned} \tag{10.614}$$

est une bijection.

(3a) **Injective** Soient $\pi_1, \pi_2 \in A(\pi_0)$. Nous supposons que $\sigma^{(\pi_1,\pi_0)}|_P = \sigma^{(\pi_2,\pi_0)}|_P$. Vu que $\sigma^{(\pi_1,\pi_0)}|_Q$ et $\sigma^{(\pi_2,\pi_0)}|_Q$ sont l'identité, nous avons en réalité $\sigma^{(\pi_1,\pi_0)} = \sigma^{(\pi_2,\pi_0)}$. Le même raisonnement est valable pour $\tau^{(\pi_1,\pi_0)} = \tau^{(\pi_2,\pi_0)}$.

En utilisant (10.612) pour π_1 et pour π_2, nous écrivons

$$\tau^{(\pi_1,\pi_0)} \circ \sigma^{(\pi_1,\pi_0)} \circ \pi_1 = \pi_0, \tag{10.615}$$

et

$$\tau^{(\pi_2,\pi_0)} \circ \sigma^{(\pi_2,\pi_0)} \circ \pi_2 = \pi_0, \tag{10.616}$$

et donc $\pi_1 = \pi_2$. L'injectivité est prouvée.

99. Signature, définition 1.297.

(3b) **Surjective** Soient $\sigma' \in S_P$ et $\tau' \in S_Q$. Nous les prolongeons par l'identité pour créer $\sigma, \tau \in S_{k+l}$. Nous posons

$$\pi_1 = \sigma^{-1} \circ \tau^{-1} \circ \pi_0. \tag{10.617}$$

Nous voulons prouver que $\psi(\pi_1) = (\sigma, \tau)$, mais d'abord nous devons rapidement vérifier que $\pi_1 \sim \pi_0$. Pour $i = 1, \dots, k$ nous avons

$$(\sigma^{-1} \circ \tau^{-1} \circ \pi_0)(i) = \sigma^{-1}\big(\pi_0(i)\big) \in \pi_0\{1, \dots, k\}. \tag{10.618}$$

Même raisonnement pour $i = k+1, \dots, k+l$, et nous déduisons que $\pi_1 \sim \pi_0$.
En utilisant la définition 10.610, pour $i \in \pi_0\{1, \dots, k\}$ nous avons

$$\begin{aligned}
\sigma^{(\pi_1, \pi_0)}(i) &= (\pi_0 \pi_1^{-1})(i) && \text{(10.619a)}\\
&= (\pi_0 \pi_0^{-1} \tau\sigma)(i) && \text{(10.619b)}\\
&= (\tau\sigma)(i) && \text{(10.619c)}\\
&= \sigma(i). && \text{(10.619d)}
\end{aligned}$$

Pour la dernière ligne, nous savons que $\sigma \in S_P$ et donc que $\sigma(i) \in P$, ce qui permet de dire que $\tau\big(\sigma(i)\big) = \sigma(i)$.

Nous avons donc une bijection $A(\pi_0) \to S_P \times S_Q$. Vu que P contient k élément et que Q en contient l, nous avons aussi une bijection $A(\pi_0) \to S_k \times S_l$.

(4) **Cardinal de $A(\pi_0)$** Par ce que nous venons de dire, nous pouvons écrire

$$\text{Card}(\pi_0) = k!l!, \tag{10.620}$$

et je vous prie de remarquer que cela ne dépend pas de π_0. Toutes les classes d'équivalence ont même cardinal.

(5) **Développer le produit extérieur** Développons le produit extérieur. Nous avons

$$(T \wedge S) = \frac{1}{k!l!} \sum_{\pi \in S_{l+k}} \epsilon(\pi)(T \otimes S)^\pi = \frac{1}{k!l!} \sum_{A \in S_{k+l}/\sim} \sum_{\pi \in A} \epsilon(\pi)(T \otimes S)^\pi. \tag{10.621}$$

EQooEQNJooDYrITi

Nous sélectionnons $\pi_0 \in S_{k+l}$ et nous regardons plus en détail la somme sur $A(\pi_0)$. Nous allons montrer que tous les éléments de la somme

$$\sum_{\pi \in A(\pi_0)} \epsilon(\pi)(T \otimes S)^\pi \tag{10.622}$$

sont égaux, et qu'il y a $k!l!$ tels éléments.

(6) **Un élément de la somme** Soient $\pi_0 \in S_{k+l}$ et $\pi \in A(\pi_0)$. Nous utilisons encore (10.612) : $\pi_0 = \tau^{(\pi,\pi_0)} \circ \sigma^{(\pi,\pi_0)} \circ \pi$. Nous allons provisoirement arrêter d'écrire les exposant ; nous écrivons σ pour $\sigma^{(\pi,\pi_0)}$. Donc

$$\begin{aligned}
(T \otimes S)^{\pi_0}(v_1, \dots, v_{k+l}) &= T\big(v_{\tau\sigma\pi(1)}, \dots, v_{\tau\sigma\pi(k)}\big) S\big(v_{\tau\sigma\pi(k+1)}, \dots, v_{\tau\sigma\pi(k+l)}\big) && \text{(10.623a)}\\
&= T\big(v_{\sigma\pi(1)}, \dots, v_{\sigma\pi(k)}\big) S\big(v_{\tau\pi(k+1)}, \dots, v_{\tau\pi(k+l)}\big). && \text{(10.623b)}
\end{aligned}$$

SUBEQooLEWCooTmElAn

Pour le premier facteur $i = 1, \dots, k$, nous avons $(\tau\sigma\pi)(i) = (\sigma\pi)(i)$. Nous introduisons $\sigma' = \pi^{-1} \circ \sigma \circ \pi|_{\{1,\dots,k\}}$. Avec tout ce que nous avons fait, $\sigma' \in S_k$, et il est donc légitime de regarder ce que fait $T^{\sigma'}$. Notez qu'il n'est pas légitime de regarder T^σ ; bref. Nous avons, par la formule (10.605) :

$$T^{\sigma'}(v_{\pi(1)}, \dots, v_{\pi(k)}) = T\big(v_{\pi\sigma'(1)}, \dots, v_{\pi\sigma'(k)}\big) = T\big(v_{\sigma\pi(1)}, v_{\sigma\pi(k)}\big). \tag{10.624}$$

Cela nous permet de réexprimer le premier facteur de (10.623b). Pour le second nous faisons le même genre de choses en introduisant τ' :

SUBEQSooEMHQooDgCgkp

$$\begin{aligned}
(T \otimes S)^{\pi_0}(v_1, \dots, v_{k+l}) &= T^{\sigma'}(v_{\pi(1),\dots,v_{\pi(k)}}) S^{\tau'}\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big) && \text{(10.625a)}\\
&= \epsilon(\sigma')\epsilon(\tau') T\big(v_{\pi(1)}, \dots, v_{\pi(k)}\big) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big) && \text{(10.625b)}\\
&= \epsilon(\sigma')\epsilon(\tau')(T \otimes S)^\pi(v_1, \dots, v_{k+l}). && \text{(10.625c)}
\end{aligned}$$

(7) **Signature** Vu que σ est l'identité sur $\pi\{k+1,\dots,k+l\}$, et vu que $\sigma' = \pi^{-1}\circ\sigma\circ\pi$, nous avons

$$\epsilon(\sigma') = \epsilon(\pi^{-1}\sigma\pi) = \epsilon(\sigma) \tag{10.626}$$

parce que ϵ est un morphisme.
Comme toujours, nous avons $\epsilon(\tau') = \epsilon(\tau)$ pour la même raison.
Nous avons $\pi_0 = \tau\circ\sigma\circ\pi$, et donc

$$\epsilon(\pi_0) = \epsilon(\tau)\epsilon(\sigma)\epsilon(\pi), \tag{10.627}$$

et en repartant de (10.625),

$$\epsilon(\pi_0)(T\otimes S)^{\pi_0} = \epsilon(\pi_0)\epsilon(\sigma)\epsilon(\tau)(T\otimes S)^{\pi} = \epsilon(\pi)(T\otimes S)^{\pi}. \tag{10.628a}$$

(8) **Résumé** Jusqu'à présent nous avons prouvé que pour tout $\pi\in A(\pi_0)$,

$$\epsilon(\pi)(T\otimes S)^{\pi} = \epsilon(\pi_0)(T\otimes S)^{\pi_0}. \tag{10.629}$$

(9) **Utilisation des classes** Nous repartons de (10.621). Pour chaque classe $A\in S_{k+l}/\sim$, nous choisissons un représentant $\pi_A\in A$. Pour tout $\pi\in A$ nous avons $\epsilon(\pi)(T\otimes S)^{\pi} = \epsilon(\pi_A)(T\otimes S)^{\pi_A}$. Nous avons alors

$$\begin{aligned}(T\wedge S) &= \frac{1}{k!l!}\sum_{A\in S_{k+l}/\sim}\sum_{\pi\in A}\epsilon(\pi)(T\otimes S)^{\pi} && (10.630a)\\ &= \frac{1}{k!l!}\sum_{A\in S_{k+l}/\sim}\sum_{\pi\in A}\epsilon(\pi_A)(T\otimes S)^{\pi_A} && (10.630b)\\ &= \frac{1}{k!l!}\sum_{A\in S_{k+l}/\sim}k!l!\epsilon(\pi_A)(T\otimes S)^{\pi_A} && (10.630c)\\ &= \sum_{A\in S_{k+l}/\sim}\epsilon(\pi_A)(T\otimes S)^{\pi_A}. && (10.630d)\end{aligned}$$

(10) **Conclusion** Chaque classe $A\in S_{k+l}/\sim$ contient un unique élément de $S_{(k,l)}$. Il suffit de prendre celui-là comme choix de π_A. Et nous avons finalement

$$T\wedge S = \sum_{\pi\in S_{(k,l)}}\epsilon(\pi)(T\otimes S)^{\pi}, \tag{10.631}$$

comme annoncé.

□

10.12.12 Produit intérieur

Pour la définition de $i_X(\omega)$ lorsque ω est une forme différentielle, voir la définition 56.30.

DEFooFWIHooJiTnwM

Définition 10.227.
*Si $v\in V$ et si $T\in\bigwedge^k V^*$, nous définissons le **produit intérieur** $i_v\colon\bigwedge^k V^*\to\bigwedge^{k-1}V^*$ par*

$$(i_vT)(X_1,\dots,X_{k-1}) = T(v,X_1,\dots,X_{k-1}). \tag{10.632}$$

PROPooETGJooMjxAFp

Proposition 10.228 ([1]).
Soit $\omega\in\bigwedge^k V^*$. L'application

$$\begin{aligned}f\colon V&\to\bigwedge^{k-1}V^*\\ v&\mapsto i_v(\omega)\end{aligned} \tag{10.633}$$

est linéaire.

Démonstration. Soit $\lambda \in \mathbb{R}$. Nous avons

$$\begin{aligned} i_{\lambda v}(\omega)(v_1,\dots,v_{k-1}) &= \omega(\lambda v, v_1,\dots,v_{k-1}) && (10.634a)\\ &= \lambda\omega(v,v_1,\dots,v_{k-1}) && (10.634b)\\ &= \lambda i_v(\omega)(v_1,\dots,v_{k-1}). && (10.634c)\end{aligned}$$

Le même genre de vérifications montre que

$$i_{v+w}(\omega)(v_1,\dots,v_{k-1}) = i_v(\omega)(v_1,\dots,v_{k-1}) + i_w(\omega)(v_1,\dots,v_{k-1}). \tag{10.635}$$

□

Le lemme suivant donne les coordonnées de $i_v(\omega)$ en fonction de celles de ω.

LEMooOBUVooBHpDZX

Lemme 10.229 ([1])**.**
Soient $v \in \mathbb{R}^n$ et $\omega \in \bigwedge^k(\mathbb{R}^n)^*$. Nous supposons que

$$\omega = \sum_{I=(i_1,\dots,i_k)} \omega_I e^*_{\otimes I} \tag{10.636}$$

et

$$i_v(\omega) = \sum_{J=(j_1,\dots,j_{k_1})} \omega'_J e^*_{\otimes J}. \tag{10.637}$$

Alors

$$\omega'_J = \sum_{i=1}^{n} \omega_{iJ} v_i. \tag{10.638}$$

Démonstration. Soient des vecteurs $v_2,\dots,v_k$. En ce qui concerne la somme sur J, nous notons le multiindice $J = (j_2,\dots,j_k)$ au lieu de $J = (j_1,\dots,j_{k-1})$ pour que ce soit plus simple plus bas :

$$\omega'(v_2,\dots,v_k) = \sum_{J=(j_2,\dots,j_k)} \omega'_J e^*_{\otimes J}(v_2,\dots,v_k) = \sum_{J=(j_2,\dots,j_k)} \omega'_J (v_2)_{j_1}\dots(v_k)_{j_{k-1}}. \tag{10.639}$$

EQooOCFQooUQqclB

et

SUBEQSooKJHQooGrKPwB

$$\begin{aligned} i_v(\omega)(v_2,\dots,v_k) &= \omega(v,v_2,\dots,v_k) && (10.640a)\\ &= \sum_I \omega_I e^*_{\otimes I}(v,v_2,\dots,v_k) && (10.640b)\\ &= \sum_I v_{i_1}(v_2)_{i_2}\dots(v_k)_{i_k} && (10.640c)\\ &= \sum_{i=1}^{n}\sum_{J=(j_2,\dots,j_k)} \omega_{iJ} v_i (v_2)_{j_2}\dots(v_k)_{j_k} && \text{cf. justif.} \quad (10.640d)\\ &= \sum_J \big(\sum_i \omega_{iJ} v_i\big)(v_2)_{j_2}\dots(v_k)_{j_{k-1}} && \text{cf. justif} \quad (10.640e)\end{aligned}$$

SUBEQooRYSHooDoXrHQ

SUBEQooLBXNooNMZggQ

Justifications.

— Pour (10.640d). Nous décomposons la somme sur $I = (i_1,\dots,i_k)$ en une somme sur le premier élément (que nous nommons i) et une somme sur les autres (que nous nommons $J = (j_2,\dots,j_k)$). Dans la notation ω_{iJ}, le iJ signifie le multiindice $(i,j_2,\dots,j_k)$.

— Pour (10.640e). Permuter les sommes.

Vu que $i_v(\omega) = \omega'$, en comparant (10.639) avec (10.640), nous trouvons

$$\omega'_J = \sum_i \omega_{iJ} v_i \tag{10.641}$$

comme il le fallait. □

PROPooYCRXooSOsqCb

Proposition 10.230 (Règle de Leibniz pour le produit extérieur[63, 1]).
Soient $T \in \Lambda^k V^*$ et $S \in \Lambda^l V^*$. Pour tout $v \in V$ nous avons

$$i_v(T \wedge S) = (i_v T) \wedge S + (-1)^k T \wedge (i_v S). \tag{10.642}$$

EQooGMHPooBjTtlv

Démonstration. Pour des soucis de lisibilité nous notons v_1 pour v et nous calculons en utilisant l'expression du lemme 10.226 pour le produit extérieur :

$$i_{v_1}(T \wedge S)(v_2, \dots, v_{k+l}) = \sum_{\pi \in S_{(k,l)}} \epsilon(\pi) T\big(v_{\pi(1)}, \dots, v_{\pi(k)}\big) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big). \tag{10.643}$$

EQooVYOYooUBenpQ

Vu que les éléments de $\pi \in S_{(k,l)}$ ordonnent séparément $\{\pi(1), \dots, \pi(k)\}$ et $\{\pi(k+1), \dots, \pi(k+l)\}$, la valeur de $\pi^{-1}(1)$ est soit 1 soit $k+1$. Nous définissons donc deux parties :

$$A_{(k,l)} = \{\pi \in S_{(k,l)} \text{ tel que } \pi(1) = 1\} \tag{10.644a}$$
$$B_{(k,l)} = \{\pi \in S_{(k,l)} \text{ tel que } \pi(k+1) = 1\}. \tag{10.644b}$$

Ces parties ont déjà été introduites et étudiées dans la définition 1.303 et le lemme 1.304. Nous poursuivons (10.643) :

$$\begin{aligned} i_{v_1}(T \wedge S)(v_2, \dots, v_{k+l}) = &\sum_{\pi \in A} \epsilon(\pi) T\big(v_{\pi(1)}, \dots, v_{\pi(k)}\big) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big) \\ &+ \sum_{\pi \in B} \epsilon(\pi) T\big(v_{\pi(1)}, \dots, v_{\pi(k)}\big) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big) \end{aligned} \tag{10.645}$$

EQooHZYQooHxqAYR

(1) **Première somme, sur $A_{k,l}$** Nous nous concentrons maintenant sur la première somme, c'est-à-dire les termes avec $\pi(1) = 1$.

$$\begin{aligned} &\sum_{\pi \in A} \epsilon(\pi) T(v_{\pi(1)}, \dots, v_{\pi(k)}) S(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}) \\ &\quad = \sum_{\pi \in A} \epsilon(\pi) (i_{v_1} T)(v_{\pi(2)}, \dots, v_{\pi(k)}) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big). \end{aligned} \tag{10.646}$$

C'est le moment de profiter des lemmes 1.303 et 1.311 pour changer la somme. Nous notons

$$f(\pi) = \epsilon(\pi)(i_{v_1} T)(v_{\pi(2),\dots,v_{\pi(k)}}) S\big(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\big), \tag{10.647}$$

et nous considérons la bijection $\varphi\colon A_{(k,l)} \to S_{(k-1,l)}$. Nous avons aussi $\varphi^{-1}(\tau) i = \tau(i-1) + 1$. Et donc, en posant $w_j = v_{j+1}$ pour y voir plus clair :

$$f\big(\varphi^{-1}(\tau)\big) = \epsilon\big(\varphi^{-1}(\tau)\big)(i_{v_1} T)\big(v_{\tau(1)+1}, \dots, v_{\tau(k-1)+1}\big) S\big(v_{\tau(k)+1}, \dots, v_{\tau(k+l-1)+1}\big) \tag{10.648a}$$
$$= \epsilon(\tau)(i_{v_1} T)\big(w_{\tau(1)}, \dots, w_{\tau(k-1)}\big) S\big(w_{\tau(k)}, \dots, w_{\tau(k+l-1)}\big) \tag{10.648b}$$
$$= \big(i_{v_1} T \otimes S\big)\big(w_{\tau(1)}, \dots, w_{\tau(k+l-1)}\big) \tag{10.648c}$$
$$= \big(i_{v_1} T \otimes S\big)^{\tau}(w_1, \dots, w_{k+l-1}) \tag{10.648d}$$
$$= \big(i_{v_1} T \otimes S\big)^{\tau}(v_2, \dots, v_{k+l}) \tag{10.648e}$$

En remettant dans la somme et en utilisant le lemme 1.311, | SUBEQSooSDMCooHmPCZr

$$\sum_{\pi \in A} f(\pi) = \sum_{\tau \in S_{(k-1,l)}} f\big(\varphi^{-1}(\tau)\big) \tag{10.649a}$$
$$= \sum_{\tau \in S_{(k-1,l)}} \big(i_{v_1} T \otimes S\big)^{\tau}(v_2, \dots, v_{l+l}) \tag{10.649b}$$
$$= (i_{v_1} T \wedge S)(v_2, \dots, v_{k+l}) \tag{10.649c}$$

Cela est déjà le premier terme de (10.642).

(2) **Seconde somme, sur $B_{(k,l)}$** Nous étudions maintenant la somme sur $B_{(k,l)}$ dans (10.645). Nous posons encore

$$f(\pi) = \epsilon(\pi)T\big(v_{\pi(1)},\dots,v_{\pi(k)}\big)S\big(v_{\pi(k+1)},\dots,v_{\pi(k+l)}\big). \tag{10.650}$$

Nous utilisons la bijection $\varphi\colon B_{(k,l)} \to S_{(k,l-1)}$ du lemme 10.224

$$\sum_{\pi\in B_{(k,l)}} f(\pi) = \sum_{\tau\in S_{(k,l-1)}} f\big(\varphi^{-1}(\tau)\big). \tag{10.651}$$

Nous avons d'abord $\epsilon\big(\varphi^{-1}(\tau)\big) = (-1)^k\epsilon(\tau)$. Ensuite, pour $j = 1,\dots,k$ nous avons $\varphi^{-1}(\tau)j = \tau(j) + 1$. Donc

$$T\big(v_{\varphi^{-1}(\tau)1},\dots,v_{\varphi^{-1}(\tau)k}\big) = T\big(v_{\tau(1)+1},\dots,v_{\tau(k)+1}\big). \tag{10.652}$$ EQooRVUBooGsYfNE

Ça, c'était la partie pas très subtile. Nous devons maintenant voir la partie avec S, c'est-à-dire déterminer en termes de τ les valeurs de $\big(\varphi^{-1}(\tau)(k+1),\dots,\varphi^{-1}(\tau)(k+l)\big)$.

Pour la suite, soyez attentifs à la distinction entre des égalités d'ensembles du type $\{a,b\} = \{c,d\}$ qui signifie que a est c ou d et que b est l'autre et les égalités de vecteurs du type $(a,b) = (c,d)$ qui signifie $a = c$ et $b = d$.

Vu que $\varphi^{-1}(\tau)$ est une bijection, nous avons $\varphi^{-1}(\tau)\{1,\dots,k+l\} = \{1,\dots,k+l\}$. Et nous avons déjà vu que

$$\varphi^{-1}(\tau)(1,\dots,k) = \big(\tau(1)+1,\dots,\tau(k)+1\big). \tag{10.653}$$

Mais par ailleurs, vu que τ prend ses valeurs dans $\{1,\dots,k+l-1\}$, nous avons

$$\{1,\dots,k+1\} = \{1,\tau(1)+1,\dots,\tau(k+l-1)+1\}. \tag{10.654}$$

Nous avons donc

$$\varphi^{-1}(\tau)\{k+1,\dots,k+l\} = \{1,\tau(k+1)+1,\dots,\tau(k+l-1)+1\}, \tag{10.655}$$

et comme les valeurs de τ sont ordonnées, nous avons même l'égalité de vecteurs

$$\varphi^{-1}(\tau)(k+1,\dots,k+l) = \big(1,\tau(k+1)+1,\dots,\tau(k+l-1)+1\big). \tag{10.656}$$

En utilisant cela,

$$\begin{aligned} S\big(v_{\varphi^{-1}(\tau)(k+1)},\dots,v_{\varphi^{-1}(\tau)(k+l)}\big) &= S\big(v_1,v_{\tau(k+1)+1},\dots,v_{\tau(k+l-1)+1}\big) && \text{(10.657a)}\\ &= (i_{v_1}S)\big(w_{\tau(k+1)},\dots,w_{\tau(k+l-1)}\big) && \text{(10.657b)}\end{aligned}$$

où nous avons posé $w_j = v_{j+1}$.

En écrivant aussi (10.652) en termes de w, nous avons

$$\begin{aligned} f\big(\varphi^{-1}(\tau)\big) &= (-1)^k\epsilon(\tau)T\big(w_{\tau(1)},\dots,w_{\tau(k)}\big)(i_{v_1}S)\big(w_{\tau(k+1)},\dots,w_{\tau(k+l-1)}\big) && \text{(10.658a)}\\ &= (-1)^k\epsilon(\tau)(T\otimes i_{v_1}S)\big(w_{\tau(1)},\dots,w_{\tau(k+l-1)}\big) && \text{(10.658b)}\\ &= (-1)^k\epsilon(\tau)(T\otimes i_{v_1}S)^\tau(w_1,\dots,w_{k+l-1}) && \text{(10.658c)}\\ &= (-1)^k\epsilon(\tau)(T\otimes i_{v_1}S)^\tau(v_2,\dots,v_{k+l}). && \text{(10.658d)}\end{aligned}$$

Et au final, en utilisant l'expression (10.607b) du produit extérieur, EQooOCIZooXZrqiQ

$$\begin{aligned} \sum_{\pi\in B_{(k,l)}} f(\pi) &= (-1)^k \sum_{\tau\in S_{(k,l-1)}} \epsilon(\tau)(T\otimes i_{v_1}S)^\tau(v_2,\dots,v_{k+l}) && \text{(10.659a)}\\ &= (-1)^k(T\wedge i_{v_1}S)(v_2,\dots,v_{k+l}). && \text{(10.659b)}\end{aligned}$$

(3) **Conclusion** Les deux termes de (10.645) sont (10.649) et (10.659). Nous avons donc finalement

$$i_{v_1}(T\wedge S)(v_2,\dots,v_{k+l}) = (i_{v_1}T\wedge S)(v_2,\dots,v_{k+l}) + (-1)^k(T\wedge i_{v_1}S)(v_2,\dots,v_{k+l}), \tag{10.660}$$

comme demandé.

□

10.12.13 Pull back

Définition 10.231.
*Si $f\colon W \to V$ est linéaire, nous définissons le **pull back** $f^*\colon \bigwedge^k V^* \to \bigwedge^k W^*$ par*

$$(f^* * T)(w_1, \dots, w_k) = T\big(f(w_1), \dots, f(w_k)\big). \tag{10.661}$$

Proposition 10.232 ([63]).
Pour toute application linéaire $f\colon W \to V$ nous avons

$$f^*(T \wedge S) = (f^*T) \wedge (f^*S). \tag{10.662}$$

Démonstration. Soient $T \in \bigwedge^k V^*$ et $S \in \bigwedge^l V^*$. En déballant les définitions, nous avons

$$\begin{aligned}
f^*(T \wedge S)(v_1, \dots, v_k) &= (T \wedge S)\big(f(v_1), \dots, f(v_{k+k})\big) && \text{(10.663a)}\\
&= \frac{(k+k)!}{k!l!}\frac{1}{k!}\sum_{\pi \in S_{k+k}} (-1)^\pi (T \otimes S)^\pi\big(f(v_1), \dots, f(v_k + l)\big) && \text{(10.663b)}\\
&= \frac{(k+k)!}{k!l!}\frac{1}{k!} && \text{(10.663c)}\\
&\quad \sum_{\pi \in S_{k+k}} (-1)^\pi T\big(f(v_{\pi(1)}), \dots, f(v_{\pi(k)})\big) S\big(f(v_{\pi(k+1)}), \dots, f(v_{\pi(k+l)})\big)\\
&= \frac{(k+k)!}{k!l!}\frac{1}{k!}\sum_{\pi \in S_{k+k}} (f^*T)(v_{\pi(1)}, \dots, v_{\pi(k)})(f^*S)(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}) && \text{(10.663d)}\\
&= \frac{(k+k)!}{k!l!}\frac{1}{k!}\sum_{\pi \in S_{k+k}} (f^*T) \otimes (f^*S)(v_{\pi(1)}, \dots, v_{\pi(k+l)}) && \text{(10.663e)}\\
&= (f^*T) \wedge (f^*S)(v_1, \dots, v_{k+l}). && \text{(10.663f)}\\
& && \text{(10.663g)}
\end{aligned}$$

□

10.12.14 Norme

Nous considérons des espaces vectoriels V et W de dimension finie. L'application (10.487) donne un isomorphisme d'espaces vectoriels

$$\begin{aligned}
\psi\colon V \otimes W &\to \mathcal{L}(V^*, W)\\
v \otimes w &\mapsto \big(\alpha \mapsto \alpha(v)w\big).
\end{aligned} \tag{10.664}$$

Et ça, c'est très bien, parce que nous connaissons une norme sur $\mathcal{L}(V^*, W)$: la norme opérateur 10.55.

DEFooEXXNooMgIpSV

Définition 10.233 ([1]).
Soient deux espaces vectoriels normés de dimension finie V et W. Sur $V \otimes W$ nous définissons, pour $t \in V \otimes W$

$$\|t\| = \|\psi(t)\|_{\mathcal{L}(V^*, W)}. \tag{10.665}$$

LEMooQPXHooJWfpmk

Lemme 10.234 ([1]).
La norme sur $V \otimes W$ vérifie

$$\|v \otimes w\| = \|v\|\|w\| \tag{10.666}$$

pour tout $v \in V$ et $w \in W$.

Démonstration. C'est un simple(?) calcul :

$$\|v\otimes w\| = \|\psi(v\otimes w)\| = \|\alpha\mapsto\alpha(v)w\| = \sup_{\|\alpha\|=1}\|\alpha(v)w\| = \sup_{\|\alpha\|=1}|\alpha(v)|\|w\|. \tag{10.667}$$

Étant donné que V est de dimension finie, $\sup_{\|\alpha\|=1}|\alpha(v)| = \|v\|$ [100]. Nous avons donc

$$\|v\otimes w\| = \|v\|\|w\|. \tag{10.668}$$

□

Le lemme suivant montre que $\mathbb{R}\otimes\mathbb{R}$ n'est pas du tout $\mathbb{R}\times\mathbb{R} = \mathbb{R}^2$. Au contraire, $\mathbb{R}\otimes\mathbb{R}$ est isomorphe à $\mathbb{R}$.

LEMooVONEooQpPgcn

Lemme 10.235 ([1])**.**
L'application

$$\begin{aligned}\varphi\colon \mathbb{R}\otimes\mathbb{R} &\to \mathbb{R}\\ 1\otimes 1 &\mapsto 1\end{aligned} \tag{10.669}$$

prolongée par linéarité est un isomorphisme isométrique.

Démonstration. D'abord une base de $\mathbb{R}$ est $\{1\}$; donc une base de $\mathbb{R}\otimes\mathbb{R}$ est $\{1\otimes 1\}$ par la proposition 10.179. Donc l'application proposée se prolonge par linéarité à tout $\mathbb{R}\otimes\mathbb{R}$.

Le fait que φ soit une bijection provient du fait que φ transforme une base en une base ; si vous n'y croyez pas, la vérification de l'injectivité et de la surjectivité est facile.

Pour que φ soit isométrique, nous faisons le calcul

$$\|\varphi(x\otimes y)\| = \|xy(1\otimes 1)\| = |xy|\|1\otimes 1\| = |xy| = \|x\otimes y\|. \tag{10.670}$$

Nous avons utilisé la propriété 7.158(3) d'une norme ainsi que le lemme 10.234 pour la norme sur $\mathbb{R}\otimes\mathbb{R}$. □

10.12.15 Applications bilinéaires, matrices et produit tensoriel

SECooUKRYooZjagcX

Soit E, un espace vectoriel de dimension finie. Si α et β sont deux formes linéaires sur un espace vectoriel E, nous définissons $\alpha\otimes\beta$ comme étant la 2-forme donnée par

$$(\alpha\otimes\beta)(u,v) = \alpha(u)\beta(v). \tag{10.671}$$

EQooUNRYooKBrXvK

Si a et b sont des vecteurs de E, ils sont vus comme des formes sur E via le produit scalaire et nous avons

$$(a\otimes b)(u,v) = (a\cdot u)(b\cdot v). \tag{10.672}$$

Cette dernière équation nous incite à pousser un peu plus loin la définition de $a\otimes b$ et de simplement voir cela comme la matrice de composantes

$$(a\otimes b)_{ij} = a_ib_j. \tag{10.673}$$

Cette façon d'écrire a l'avantage de ne pas demander de se souvenir qui est un vecteur ligne, qui est un vecteur colonne et où il faut mettre la transposée. Évidemment $(a\otimes b)$ est soit ab^t soit a^tb suivant que a et b soient ligne ou colonne.

100. Cela est une des raisons pour lesquelles nous sommes en dimension finie : je ne sais pas si cette égalité est vraie en dimension inifinie.

10.12.16 Application d'opérateurs

LemMyKPzY

Lemme 10.236.
Soient $x, y \in E$ et A, B deux opérateurs linéaires sur E vus comme matrices. Alors

$$(Ax \otimes By) = A(x \otimes y)B^t. \tag{10.674}$$

EqXdxySu

Démonstration. Calculons la composante ij de la matrice $(Ax \otimes By)$. Nous avons

$$\begin{aligned}(Ax \otimes By)_{ij} &= (Ax)_i(By)_j && \text{(10.675a)}\\ &= \sum_{kl} A_{ik}x_kB_{jl}y_l && \text{(10.675b)}\\ &= A_{ik}(x \otimes y)_{kl}B_{jl} && \text{(10.675c)}\\ &= \big(A(x \otimes y)B^t\big)_{ij}. && \text{(10.675d)}\end{aligned}$$

□

Le fait que les applications linéaires soient continues [101] est valable dans une assez large gamme d'espaces vectoriels[352]. Nous voyons ici dans le cas des espaces vectoriels normés de dimension finies.

PROPooADPDooOtukQP

Proposition 10.237.
Soient des espaces vectoriels normés E et F. Si $f : E \to F$ est une application linéaire et si E est de dimension finie, alors f est continue.

Démonstration. La proposition 10.56 nous dit que $\|f\| < \infty$, c'est-à-dire que f est borné. Donc la proposition 10.70 conclut. □

LemWWXVSae

Lemme 10.238.
Soit F un espace de Banach et deux suites $A_k \to A$ et $B_k \to B$ dans $\mathcal{L}(F, F)$. Alors $A_k \circ B_k \to A \circ B$ dans $\mathcal{L}(F, F)$, c'est-à-dire

$$\lim_{k\to\infty}(A_kB_k) = \left(\lim_{k\to\infty} A_k\right)\left(\lim_{k\to\infty} B_k\right). \tag{10.676}$$

Démonstration. Il suffit d'écrire

$$\|A_kB_k - AB\| \leqslant \|A_kB_k - A_kB\| + \|A_kB - AB\|. \tag{10.677}$$

Le premier terme tend vers zéro pour $k \to \infty$ parce que

$$\begin{aligned}\|A_kB_k - A_kB\| &= \|A_k(B_k - B)\| && \text{(10.678a)}\\ &\leqslant \|A_k\|\|B_k - B\| \to \|A\| \cdot 0 && \text{(10.678b)}\\ &= 0 && \text{(10.678c)}\end{aligned}$$

où nous avons utilisé la propriété fondamentale de la norme opérateur : la proposition 10.70. Le second terme tend également vers zéro pour la même raison. □

10.12.17 Convergence en norme et par composante

En dimension infinie, la convergence en norme et la convergence composante par composante ne s'impliquent ni dans un sens ni dans l'autre.

L'exemple suivant devrait être formalisé dans l'espace ℓ^2 des suites de carré sommable, mais vous voyez l'idée.

101. Proposition 10.70.

Exemple 10.239.
Nous considérons l'ensemble des suites réelle munie de la norme $\|x\| = \sqrt{\sum_{k=0}^{\infty} |x_k|^2}$. Dedans nous considérons les vecteurs de base e_i donnés par

$$(e_i)_n = \delta_{in}. \tag{10.679}$$

Ensuite nous considérons la base

$$f_i = e_1 + \frac{1}{2^i} e_i. \tag{10.680}$$

La suite $x_n = f_n - f_1$, dans cette base a toujours -1 comme première composante [102]. Et pourtant elle converge en norme vers 0. △

10.13 Calcul différentiel dans un espace vectoriel normé

10.13.1 Limite de fonctions

La limite de fonctions est la définition 7.110.

LEMooTFLVooMMNmQr

Lemme 10.240.
Soient deux espaces vectoriels normés V et E sur le corps $\mathbb{K}$. Si $f: V \to E$ est constante, $f(x) = u$ pour tout $x \in V$, alors pour tout $a \in V$, nous avons la limite $\lim_{x\to a} f(x) = u$.

PROPooBHHCooPbRggh

Proposition 10.241.
Soient deux espaces vectoriels normés V et E. Soient deux fonctions $f, g: V \to E$. Nous supposons que $\lim_{x\to a} f(x) = \alpha$ et $\lim_{x\to a} g(x) = \beta$. Alors

ITEMooDPBHooKRccPZ

(1) Si $\lambda \in \mathbb{K}$, alors $\lim_{x\to a} \big(f(x) + \lambda\big) = \alpha + \mu$.

ITEMooDKJPooMynOcG

(2) La fonction $f + g$ a une limite $x \to a$ qui vaut $\lim_{x\to a}(f + g)(x) = \alpha + \beta$,

(3) La fonction fg a une limite en a, qui vaut $\lim_{x\to a}(fg)(x) = \alpha\beta$,

LEMooYJGLooVBaglB

Lemme 10.242.
Soient un espace vectoriel normé E ainsi qu'une fonction $f: \mathbb{R} \to E$ telle que $\lim_{t\to 0} f(t) = v$. Alors pour tout $\lambda \in \mathbb{R}$ nous avons

$$\lim_{t\to 0} f(\lambda t) = v. \tag{10.681}$$

Démonstration. Petite clarification pour commencer : E est normé et donc séparé. La proposition 7.113 nous dit que la limite est unique ; nous voilà rassurés sur ce point. Ensuite nous utilisons la définition 7.110 de la limite.

Nous posons $g(t) = f(\lambda t)$. Soit un ouvert V autour de v dans E. Par l'hypothèse de limite pour f, il existe un voisinage U de 0 dans $\mathbb{R}$ tel que $f\big(U\backslash\{0\}\big) \subset V$.

En posant $U' = U/\lambda$, nous avons bien $g\big(U'\backslash\{0\}\big) \subset V$. En effet si $t \in U'\backslash\{0\}$, il existe $x \in U\backslash\{0\}$ tel que $t = x/\lambda$ et donc $g(t) = g(x/\lambda) = f(x) \in V$. □

10.13.2 Définition de la différentielle

Quelques motivations pour la notion de différentielle sont données dans 12.22.1.

DefDifferentielllePta

Proposition-Définition 10.243 ([1]).
Soient deux espaces vectoriels normés [103] de dimension finie E et F ainsi qu'une fonction $f: \mathcal{U} \to F$ où $\mathcal{U}$ est un ouvert de E. Si il existe une application linéaire $T \in \mathcal{L}(E, F)$ satisfaisant

$$\lim_{\substack{h\to 0\\ h\in E}} \frac{f(a+h) - f(a) - T(h)}{\|h\|_E} = 0, \tag{10.682}$$

EqCritereDefDiff

102. N'essayez pas de faire un dessin : ça ne fonctionne qu'en dimension infinie.
103. Définition 7.158.

alors il en existe une seule.

*Dans ce cas nous disons que f est **différentiable au point** a et l'application T ainsi définie est appelée **différentielle** de f au point a, et nous la notons df_a.*

Démonstration. Soient deux applications linéaires T_1, T_2 satisfaisant la condition (10.682). Nous avons

$$\frac{\|T_1(h)-T_2(h)\|_F}{\|h\|_E} \leqslant \frac{\|T_1(h)-f(a+h)+f(a)\|}{\|h\|}+\frac{\|f(a+h)-f(a)-T_2(h)\|}{\|h\|}\to 0. \tag{10.683}$$

Nous avons donc

$$\lim_{h\to 0}\frac{\|(T_1-T_2)(h)\|_F}{\|h\|_E}=0. \tag{10.684}$$

Soit $\epsilon>0$. Ce que signifie la limite est qu'il existe un $r>0$ tel que pour tout $u\in B_E(0,r)$, nous ayons

$$\frac{\|(T_1-T_2)(u)\|_F}{\|u\|_E}<\epsilon. \tag{10.685}$$

Soit $v\in E$. Nous considérons $\lambda\in\mathbb{R}$ tel que $\lambda v\in B(0,r)$, par exemple $\lambda<r/\|v\|$. Nous avons

$$\epsilon>\frac{\|(T_1-T_2)(\lambda v)\|_F}{\|\lambda v\|_E}=\frac{\|(T_1-T_2)(v)\|}{\|v\|}. \tag{10.686}$$

Cela donne

$$\|(T_1-T_2)(v)\|<\|v\|\epsilon. \tag{10.687}$$

Nous avons donc $\|(T_1-T_2)(v)\|=0$, soit $T_1(v)=T_2(v)$. □

L'application différentielle

$$\begin{aligned} df\colon E&\to\mathcal{L}(E,F)\\ a&\mapsto df_a \end{aligned} \tag{10.688}$$

est également très importante.

10.13.3 Pas en dimension infinie

Pourquoi ne pas définir la différentielle en dimension infinie[1] ? Dans le cas d'une application linéaire, en posant $T=f$, nous avons toujours la limite (10.682). Autrement dit, une application linéaire est toujours différentiable. Mais si f n'est pas bornée (et donc pas continue en dimension infinie), la limite (10.682) n'implique pas la continuité, parce qu'il n'y a pas de garanties sur la valeur de

$$\lim_{h\to 0}\|T(h)\|. \tag{10.689}$$

Autrement dit, nous pourrions avoir des fonctions différentiables non continues. Pour cela, nous nous abstenons de définir la différentielle d'une application dans le cas de dimension infinie.

10.13.4 Dérivée directionnelle

DEFooCATTooTPLtpR

Définition 10.244 (Dérivée directionnelle)**.**
*Soient des espaces vectoriels normés V et E. Soient une application $f\colon V\to E$ et $u\in V$. La **dérivée directionnelle** de f en $a\in V$ dans la direction u est l'élément de E défini par*

$$(\partial_u f)(a)=\lim_{\epsilon\to 0}\frac{f(a+\epsilon u)-f(a)}{\epsilon} \tag{10.690}$$

pourvu que la limite existe.

*Soient f une application de $U\subset\mathbb{R}^m$ dans $\mathbb{R}$ et u un vecteur de $\mathbb{R}^m$. La fonction f est **dérivable sur** U **suivant le vecteur** u, si f est dérivable suivant le vecteur u en tout point de U.*

DEFooOYFZooFWmcAB

Définition 10.245 (Dérivée)**.**
*Soit une application $f\colon \mathbb{R} \to V$ où V est un espace vectoriel normé. La **dérivée** de f en $a \in \mathbb{R}$ est le nombre*

$$f'(a) = (\partial_1 f)(a) \tag{10.691}$$

où ∂_1 est la dérivée directionnelle définie en 10.244.

*Dans le cas où la limite définissant $f'(a)$ existe, nous disons que la fonction f est dérivable en a. La **fonction dérivée** de f est*

$$\begin{aligned} f'\colon A' &\to \mathbb{R} \\ a &\mapsto f'(a) \end{aligned} \tag{10.692}$$

définie sur l'ensemble noté A' des points a où f est dérivable.

10.246.
En ce qui concerne les notations, si une base $\{e_i\}$ de V est donnée, nous notons $\partial_i f$ la dérivée de f dans la direction de e_i. La fonction $\partial_i f$ est la **dérivée partielle** de f. Dans le cas de $V = \mathbb{R}^n$, cela est souvent noté

$$\frac{\partial f}{\partial x_i}(a) = \frac{d}{dt}\Big[f(a + te_i)\Big]_{t=0}. \tag{10.693}$$

10.247.
Même si une fonction est dérivable en un point dans toutes les directions, on n'est pas sûr qu'elle soit continue en ce point. La dérivabilité directionnelle n'est donc pas une notion suffisante pour assurer la continuité. C'est pourquoi on introduit le concept de *différentiabilité*.

Remarque 10.248.
De nombreuses sources parlent de dérivée **dans la direction** du vecteur v en définissant (avec une certaine raison) une **direction** dans $\mathbb{R}^m$ comme étant un vecteur de norme 1.

Ces personnes ne définissent alors $\partial_u f$ que pour $\|u\| = 1$. Pourquoi ? Le but de la dérivée directionnelle dans la direction u est de savoir à quelle vitesse la fonction monte lorsque l'on se déplace en suivant la direction u. Cette information n'aura un caractère « objectif » que si l'on avance à une vitesse donnée. En effet, si on se déplace deux fois plus vite, la fonction montera deux fois plus vite. Par convention, on demande alors d'avancer à vitesse 1.

Ici, pour être plus souple en termes de notations et de manipulations, nous définissons $\partial_u f$ pour tout u (non nul). Nous devons cependant garder en tête que le nombre $(\partial_v f)(a)$ ne peut pas être interprété comme étant une « vitesse de croissance de f en a » de façon trop sérieuse.

Cas particulier où $n = 2$:

$a = (a_1, a_2)$, $u = (u_1, u_2)$ et

$$\frac{\partial f}{\partial u}(a_1, a_2) = \lim_{t \to 0} \frac{f(a_1 + tu_1, a_2 + tu_2) - f(a_1, a_2)}{t} \tag{10.694}$$

Un cas particulier des dérivées directionnelles est la dérivée partielle. Si nous considérons la base canonique e_i de $\mathbb{R}^n$, nous notons

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial e_i}. \tag{10.695}$$

Dans le cas d'une fonction à deux variables, nous avons donc les deux dérivées partielles

$$\frac{\partial f}{\partial x}(a) \quad \text{et} \quad \frac{\partial f}{\partial y}(a) \tag{10.696}$$

qui correspondent aux dérivées directionnelles dans les directions des axes. Ces deux nombres représentent de combien la fonction f monte lorsqu'on part de a en se déplaçant dans le sens des axes X et Y.

LEMooVOTHooPJcrWH

Lemme 10.249.
Nous notons $\mathbb{K}$ le corps $\mathbb{R}$ ou $\mathbb{C}$. Soient deux espaces vectoriels E et F sur $\mathbb{K}$. Soient une application $f\colon E \to F$ ainsi que $a, u \in E$ tels que $\frac{\partial f}{\partial u}(a)$ existe.

Alors pour tout $\lambda \in \mathbb{K}$, $(\partial_{\lambda u} f)(a)$ existe et

$$\frac{\partial f}{\partial(\lambda u)}(a) = \lambda \frac{\partial f}{\partial u}(a). \tag{10.697}$$

Démonstration. Nous allons utiliser le lemme 10.242. D'abord nous avons, pour tout t, λ et a :

$$\frac{f(a+t\lambda u) - f(a)}{t} = \lambda \frac{f(a+t\lambda u) - f(a)}{\lambda t}. \tag{10.698}$$

EQooRDUEooScpIZa

En posant $g(t) = \frac{f(a+tu)-f(a)}{t}$ (t est une variable dans $\mathbb{R}$), l'hypothèse est que $\lim_{t\to 0} g(t)$ existe et vaut $\frac{\partial f}{\partial u}(a)$. Le lemme 10.242 indique que $\lim_{t\to 0} g(\lambda t)$ existe aussi et vaut la même chose. Donc

$$\lim_{t\to 0} \frac{f(a+\lambda t u) - f(a)}{\lambda t} = \frac{\partial f}{\partial u}(a). \tag{10.699}$$

En prenant la limite dans (10.698), nous avons le résultat. □

Exemple 10.250.
Considérons la fonction $f(x, y) = 2xy^2$. Lorsque nous calculons $\partial_x f(x, y)$, nous faisons comme si y était constant. Nous avons donc $\partial_x f(x, y) = 2y^2$. Par contre lors du calcul de $\partial_y f(x, y)$, nous prenons x comme une constante. La dérivée de y^2 par rapport à y est évidemment $2y$, et par conséquent, $\partial_y f(x, y) = 4xy$. △

Pour les fonctions d'une seule variable, la dérivabilité en un point a implique la continuité en a. Cela n'est pas vrai pour les fonctions de plusieurs variables : il existe des fonctions f qui sont dérivables suivant tout vecteur au point a sans pour autant être continue en a.

Exemple 10.251.
Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases} \tag{10.700}$$

Pour voir que f n'est pas continue en $(0, 0)$ il suffit de calculer la limite de f restreinte à la parabole $y = x^2$

$$\lim_{x\to 0} f(x, x^2) = \frac{1}{2} \neq 0.$$

Pourtant la fonction f est dérivable en $(0, 0)$ dans toutes les directions. En effet, soit $v = (v_1, v_2)$. Si $v_2 \neq 0$, alors

$$\partial_v f(a) = \lim_{\substack{t\to 0 \\ t\neq 0}} \frac{t^3 v_1^2 v_2}{t^5 v_1^4 + t^3 v_2^2} = \frac{v_1^2}{v_2}, \tag{10.701}$$

EQooXOCWooVbCiRd

tandis que si $v_2 = 0$, alors la valeur de $f(tv_1, 0)$ est 0 pour tout t et v_1, donc la dérivée partielle de f par rapport à x en l'origine existe et est nulle. △

Exemple 10.252.
Pour une fonction réelle à variable réelle, la dérivabilité entraine la continuité. Il n'en va pas de même pour les fonctions à plusieurs variables, comme le montre l'exemple suivant :

$$f(x, y) = \begin{cases} 0 & \text{si } x = 0 \\ \dfrac{y}{x}\sqrt{x^2 + y^2} & \text{sinon.} \end{cases} \tag{10.702}$$

Nous avons tout de suite

$$\frac{\partial f}{\partial y}(0,0) = 0. \tag{10.703}$$

De plus si $u_x \neq 0$ nous avons

$$\frac{\partial f}{\partial u}(0,0) = \frac{u_y}{u_x}\|u\|. \tag{10.704}$$

Donc toutes les dérivées directionnelles de f en $(0,0)$ existent alors que la fonction n'y est manifestement pas continue. En effet sous forme polaire,

$$f(r,\theta) = \frac{r\sin(\theta)}{\cos(\theta)}, \tag{10.705}$$

et quelle que soit la valeur de r, en prenant θ suffisamment proche de $\pi/2$, la fraction peut être arbitrairement grande.

Nous verrons par la proposition 12.252 que la différentiabilité d'une fonction implique sa continuité. △

DEFooWZGTooKCWBJn

Définition 10.253 (gradient)**.**
Soit une application $f\colon V \to E$ entre deux espaces vectoriels normés. Soient $a \in V$ et une base finie $\{e_i\}_{i=1,\dots,n}$ de V. Nous définissons le ***gradient*** *de f en a comme étant le vecteur*

$$(\nabla f)(a) = \begin{pmatrix} (\partial_i f)(a) \\ \vdots \\ (\partial_n f)(a). \end{pmatrix} \tag{10.706}$$

PROPooYCLBooOLoDnM

Proposition 10.254.
Soit une application $f\colon V \to E$ entre deux espaces vectoriels normés. Soient $u \in V$, $a \in V$ ainsi qu'une base finie $\{e_i\}_I$ de V. Nous supposons que $(\partial_u f(a))$ et $(\partial_i f)(a)$ existent pour tout $i \in I$. Alors

$$(\partial_u f)(a) = \sum_{i\in I} u_i(\partial_i f)(a) = u \cdot (\nabla f)(a). \tag{10.707}$$

PROPooZOFLooBmcZqN

Proposition 10.255.
Soit une application $f\colon V \to E$ différentiable en $a \in V$. Alors toutes les dérivées directionnelles $(\partial_u f)(a)$ existent et

$$(df_a)(u) = \sum_i (\partial_u f)(a). \tag{10.708}$$

Démonstration. L'hypothèse de différentiabilité de f en a dit que

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - df_a(h)}{h} = 0. \tag{10.709}$$

Soit $u \in V$. Nous utilisons la proposition 9.96 avec le chemin

$$\begin{aligned} \gamma\colon]-1,1[&\to V \\ t &\mapsto tu. \end{aligned} \tag{10.710}$$

Cela donne :

$$\lim_{t\to 0} \frac{f(a+tu) - f(a) - df_a(tu)}{|t|} = 0. \tag{10.711}$$

nous pouvons dire que la limite $t \to 0$ existant, les limites à gauche et à droite existent et sont toutes deux égales à zéro [104]. À partir de

$$\lim_{t\to 0^+} \frac{f(a+tu) - f(a) - df_a(tu)}{|t|} = 0, \tag{10.712}$$

104. Cela n'est même pas un résultat en plus. C'est la proposition 9.96, avec le chemin γ pris sur $[-1,0[$ et sur $]0,1]$ séparément. Sinon, vous pouvez invoquer la proposition 9.112

en utilisant la linéarité $df_a(tu) = tdf_a(u)$, nous déduisons que

$$\lim_{t\to 0^+} \frac{f(a+tu)-f(a)}{t} = df_a(u). \tag{10.713}$$

De même à partir de la limite à gauche nous déduisons

$$\lim_{t\to 0^-} \frac{f(a+tu)-f(a)}{t} = df_a(u). \tag{10.714}$$

Et maintenant on utilise la proposition 9.112 pour dire que les limites à gauche et à droite existant et étant égales, la limite existe et est égale :

$$\lim_{t\to 0} \frac{f(a+tu)-f(a)}{t} = df_a(u). \tag{10.715}$$

Cela signifie bien que $(\partial_u f)(a) = df_a(u)$. □

PROPooEOXLooNKHGHj

Proposition 10.256 ([1])**.**
Si $f\colon V \to E$ est une application différentiable et injective, alors pour tout $x \in v$, l'application linéaire $df_x\colon V \to E$ est injective.

10.13.5 Théorème des accroissements finis

Si une fonction est dérivable en a alors elle peut être approximée « au premier ordre » par une formule simple qui sera généralisée pour des dérivées d'ordre supérieurs avec les séries de Taylor, théorème 12.461. Pour trouver des versions avec des dérivations partielles, voir le thème 55.

PropUTenzfQ

Proposition 10.257 (Développement limité au premier ordre)**.**
Soit un espace vectoriel normé E sur le corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). Si $f\colon \mathbb{R} \to E$ est une fonction dérivable[105]*, alors il existe une fonction $\alpha\colon \mathbb{R} \to E$ telle que*

EQooHBDHooPrVjJD

$$\begin{cases} f(a+h) = f(a) + hf'(a) + \alpha(h) & \text{(10.716a)} \\ \lim_{h\to 0} \frac{\alpha(h)}{h} = 0. & \text{(10.716b)} \end{cases}$$

Il existe aussi une fonction $\beta\colon \mathbb{R} \to E$ telle que

EQooPWIZooVuhjmt

$$\begin{cases} f(a+h) = f(a) + hf'(a) + h\beta(h) & \text{(10.717a)} \\ \lim_{h\to 0} \beta(h) = 0. & \text{(10.717b)} \end{cases}$$

SUBEQooMJTJooOpBfqv

Démonstration. La fonction f étant dérivable en a nous avons l'existence de la limite suivante :

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}. \tag{10.718}$$

Nous définissons $\beta\colon \mathbb{R} \to E$ par

$$\beta(h) = \frac{f(a+h)-f(a)}{h} - f'(a). \tag{10.719}$$

La proposition 10.241(1) donne directement $\lim_{h\to 0} \beta(h) = 0$. En multipliant par h (qui est différent de 0), et en isolant $f(a+h)$, nous avons la formule (10.717a). En nommant $\alpha(h) = h\beta(h)$ nous trouvons la fonction α de la formule (10.716) :

$$f(a+h) = f(a) + hf'(a) + \alpha(h) \tag{10.720}$$

avec

$$\lim_{h\to 0} \frac{\alpha(h)}{h} = \lim_{h\to 0} \beta(h) = 0. \tag{10.721}$$

□

105. Définition 10.245.

PROPooPLWZooCGmjfi

Proposition 10.258.
Soient deux espaces vectoriels normés V et E. Soit une application $f\colon V \to E$ dérivable dans le sens de h en $a \in V$. Alors nous avons une application $\beta\colon \mathbb{R} \to E$ telle que

$$\begin{cases} f(a+th) = f(a) + t(\partial_h f)(a) + \beta(t) & \text{(10.722a)} \\ \lim_{t\to 0} \beta(t) = 0. & \text{(10.722b)} \end{cases}$$

Démonstration. Nous posons

$$\begin{aligned} \varphi\colon \mathbb{R} &\to E \\ t &\mapsto f(a+th). \end{aligned} \tag{10.723}$$

et nous lui appliquons les accroissements finis de la proposition 10.257 : $\varphi(t) = \varphi(0) + t\varphi'(0) + t\beta(t)$. En ce qui concerne $\varphi'(0)$ nous avons :

$$\begin{aligned} \varphi'(0) &= \lim_{\epsilon\to 0} \frac{\varphi(\epsilon) - \varphi(0)}{\epsilon} && \text{(10.724a)} \\ &= \lim_{\epsilon\to 0} \frac{f(a+\epsilon h) - f(a)}{\epsilon} && \text{(10.724b)} \\ &= (\partial_h f)(a) && \text{(10.724c)} \end{aligned}$$

par la définition 10.244. □

val_medio_1

Théorème 10.259 (Accroissements finis pour les dérivées suivant un vecteur)**.**
Soit U un ouvert dans $\mathbb{R}^m$ et soit $f : U \to \mathbb{R}^n$ une fonction. Soient a et b deux points distincts dans U, tels que le segment[106] *$[a,b]$ soit contenu dans U. Soit u le vecteur*

$$u = \frac{b-a}{\|b-a\|_m}.$$

Si $\partial_u f(x)$ existe pour tout x dans $[a,b]$ on a

$$\|f(b) - f(a)\|_n \leqslant \sup_{x\in[a,b]} \|\partial_u f(x)\|_n \|b-a\|_m.$$

Démonstration. Nous considérons la fonction $g(t) = f\big((1-t)a + tb\big)$. Elle décrit la droite entre a et b parce que $g(0) = a$ et $g(1) = b$. En ce qui concerne la dérivée,

$$\begin{aligned} g'(t) &= \lim_{h\to 0} \frac{g(t+h) - g(t)}{h} \\ &= \lim_{h\to 0} \frac{f\big((1-t-h)a + (t+h)b\big) - f\big((1-t)a + tb\big)}{h} \\ &= \lim_{h\to 0} \frac{f\big(a + (t+h)(b-a)\big) - f\big(a + t(b-a)\big)}{h} \\ &= \frac{\partial f}{\partial u}\big(a + t(b-a)\big)\|b-a\|. \end{aligned} \tag{10.725}$$

Le dernier facteur $\|b-a\|$ apparaît pour la normalisation du vecteur u. En effet dans la limite, il apparaît $h(b-a)$, ce qui donnerait la dérivée le long de $b-a$, tandis que u vaut $(b-a)/\|b-a\|$.

Par le théorème des accroissements finis pour g, il existe $t_0 \in]0,1[$ tel que

$$g(1) = g(0) + g'(t_0)(1-0). \tag{10.726}$$

Donc

$$\|g(1) - g(0)\| \leqslant \sup_{t_0} \|g'(t_0)\| = \sup_{t_0\in]0,1[} \left\| \frac{\partial f}{\partial u}(a + t_0(b-a)) \right\| \|b-a\|. \tag{10.727}$$

Mais lorsque t_0 parcourt $]0,1[$, le point $a + t_0(b-a)$ parcourt le segment $]a,b[$, d'où le résultat. □

106. Définition 9.52.

10.13.6 Différentielle d'ordre 2

PROPooBODMooJHyzup

Proposition 10.260 ([1]).
Soient deux espaces vectoriels normés V et E. Soit une application $f\colon V \to E$ deux fois différentiable et dont les différentielles partielles secondes existent[107]*. Alors nous avons la formule*

$$(d^2f)_a(v)w = \sum_{ij}(\partial^2_{ij}f)(a)v_iw_j. \tag{10.728}$$

Démonstration. Rappel à propos des espaces concernés :

$$f\colon V \to E \tag{10.729a}$$
$$df\colon V \to \mathcal{L}(V,E) \tag{10.729b}$$
$$d^2f\colon V \to \mathcal{L}\big(V,\mathcal{L}(V,E)\big). \tag{10.729c}$$

En ce qui concerne la différentielle première nous avons

$$\begin{aligned} df\colon V &\to \mathcal{L}(V,E) \\ a &\mapsto \sum_i(\partial_i f)(a)e_i^*.. \end{aligned} \tag{10.730}$$

□

10.13.7 Quelques mots à propos des différentielles d'ordre supérieur

Soient deux espaces vectoriels normés V et W ainsi qu'une application $f\colon V \to W$. La différentielle est une application $df\colon V \to \mathcal{L}(V,W)$. Pour être clair, la différentielle seconde consiste à différentier df_x par rapport à x. C'est-à-dire que la différentielle seconde est une application $d(df)\colon V \to \mathcal{L}\big(V,\mathcal{L}(V,W)\big)$.

Et c'est là que commencent les problèmes. Les différentielles successives font intervenir des emboîtements de plus en plus profonds d'espaces comme $d^3f\colon V \to \mathcal{L}\Big(V,\mathcal{L}\big(V,\mathcal{L}(V,W)\big)\Big)$.

Nous introduisons quelques notations pour traiter ces espaces.

DefPNjMGqy

Définition 10.261 ([353, 1]).
Soient deux espaces vectoriels normés V et W ainsi qu'une application $f\colon V \to W$. Nous disons que f est

- *de classe C^0 si elle est continue,*
- *de classe C^1 si l'application différentielle $df\colon V \to \mathcal{L}(V,W)$ est continue,*
- *de classe C^k si sa différentielle est de classe C^{k-1}.*
- *de classe C^∞ si elle est de classe C^k pour tout k.*

Lorsque nous demandons que la différentielle de f soit continue, nous entendons bien la continuité de $df\colon V \to \mathcal{L}(V,W)$, c'est-à-dire la continuité de df_x par rapport à x. Sur$\mathcal{L}(V,E)$, nous considérons la topologie de la norme opérateur 10.55.

Le lien entre classe C^k et dérivées partielles d'ordre k sera le théorème 12.330.

DEFooJYOPooBzditG

Définition 10.262.
Soient des espaces vectoriels normés V et E. Nous définissons les espaces emboîtés par récurrence :

$$\begin{cases} E_0 = E & \text{(10.731a)} \\ E_{k+1} = \mathcal{L}(V,E_k). & \text{(10.731b)} \end{cases}$$

107. Ces deux hypothèses sont redondantes, mais nous n'avons pas envie de nous prendre la tête avec ça ici.

Définition 10.263.
Si $\{e_i\}$ est une base d'un espace vectoriel V, nous allons noter

$$\begin{aligned} \omega_i : V &\to \mathbb{R} \\ x &\mapsto x_i. \end{aligned} \tag{10.732}$$

Ce ω_i est ce qu'on appelle souvent e_i^*. Plus généralement, si I est le multiindice $(i_1, \dots, i_l)$ nous notons $\omega_I \in \mathcal{L}^l(V, \mathbb{R})$ par

$$\begin{aligned} \omega_I : V^l &\to \mathbb{R} \\ (x^{(1)}, \dots, x^{(l)}) &\mapsto x^{(1)}_{i_1} \dots x^{(l)}_{i_l}. \end{aligned} \tag{10.733}$$

Ce seront nos formes multilinéaires de base.

Afin de garder des notations très explicites, nous ne pouvons pas écrire des formules comme

$$df_a = \sum_i \frac{\partial f}{\partial x_i}(a)\omega_i$$

parce que si f prend ses valeurs dans $\mathcal{L}(V, \mathbb{R})$, lorsqu'on écrit $df_a(v)$, il n'y a aucune raison à priori de vouloir que v soit pris par ω_i au lieu de $\partial_i f(a)$.

Nous introduisons donc un produit fait exprès pour dire que « c'est celui de droite qui prend ».

DEFooLULCooYjBEaZ

Définition 10.264 ([1])**.**
Si W est un espace vectoriel, nous définissons le produit $\times_n$ par

$$\begin{aligned} \times_1 : W \times \mathcal{L}(V, \mathbb{R}) &\to \mathcal{L}(V, W) \\ (w \times_1 \alpha)(v) &= \alpha(v)w \end{aligned} \tag{10.734}$$

et par[108]

$$\begin{aligned} \times_n : W \times \mathbb{R}_n &\to W_n \\ (w \times_n \alpha)(v) &= w \times_{n-1} \alpha(v) \end{aligned} \tag{10.735}$$

Cette notation sera utilisée dans la proposition 12.292 pour écrire correctement df_a. Pour l'instant nous n'en avons pas besoin.

DefAQIQooYqZdya

Définition 10.265 (difféomorphisme)**.**
Soient U et V, deux ouverts d'un espace vectoriel normé. Une application f de U dans V est un ***difféomorphisme*** *si elle est bijective, différentiable*[109] *et dont l'inverse $f^{-1} : V \to U$ est aussi différentiable.*

Un C^k-difféomorphisme est un difféomorphisme qui est C^k et dont l'inverse est C^k.

10.266.
Truc marrant : un C^1-difféomorphisme n'est pas seulement un difféomorphisme qui est C^1. L'inverse doit également être C^1. Comment nommer un difféomorphisme qui est par ailleurs un application de classe C^1 ? Je ne sais pas.

10.267.
Sur le wikipédia anglophone[110], on demande que la différentielle de l'inverse soit continue. Ce que wikipédia anglophone est nommé *diffeomorphism* est ici un C^1-difféomorphisme.

RemATQVooDnZBbs

Remarque 10.268.
L'application norme étant continue, le critère du théorème 7.146 est en réalité assez général. Par

108. Définition 10.262 pour les espaces E_n et $\mathbb{R}_n$.
109. Différentiables, définition 10.243.
110. https://en.wikipedia.org/wiki/Diffeomorphism

exemple à partir d'une application différentiable[111] $f\colon X \to Y$ nous pouvons considérer la fonction réelle

$$a \mapsto \|df_a\| \tag{10.736}$$

où la norme est la norme opérateur[112]. Si f est de classe C^1 alors cette application est continue et donc bornée sur un compact K de X.

10.13.8 Différentielle d'applications linéaires

LEMooZSNMooCfjzOB

Lemme 10.269 (Différentielle d'une application linéaire)**.**
Soient deux espaces vectoriels normés E et F. Soit une application linéaire $f\colon E \to F$.

(1) Si $f\colon E \to F$ est linéaire, alors sa différentielle est

$$\begin{aligned} df\colon E &\to \mathcal{L}(E,F) \\ a &\mapsto f. \end{aligned} \tag{10.737}$$

(2) Si $f\colon E \to F$ est linéaire, toutes les différentielles d'ordre supérieures sont nulles.

(3) Toute application linéaire est de classe C^∞.

(4) Toute application affine est de classe C^∞.

Démonstration. Pour rappel, toujours bon à avoir en tête : $df\colon E \to \mathcal{L}(E,F)$. Soit $a \in E$; nous avons

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - f(h)\|_F}{\|h\|_E} = 0 \tag{10.738}$$

parce que le numérateur est nul pour tout h. Donc $h \mapsto f(h)$ est la différentielle de f au point a parce que elle vérifie la condition (10.243).

Nous avons prouvé que la différentielle de f est l'application constante

$$\begin{aligned} df\colon E &\to \mathcal{L}(E,F) \\ a &\mapsto f \end{aligned} \tag{10.739}$$

En ce qui concerne la différentielle seconde, nous prouvons que $d(df)_a = 0$ pour tout $a \in E$. En effet,

$$\lim_{h\to 0} \frac{\|df_{a+h} - df_a\|_{\mathcal{L}(E,F)}}{\|h\|_E} = 0 \tag{10.740}$$

parce que le numérateur vaut $f - f = 0$.

Maintenant il n'est pas compliqué de faire une récurrence : si f est de classe C^k et si $d^k(f) = 0$, alors $d^k(f)$ est de classe C^1 et $d^k f = 0$. □

LEMooAJDLooIPcmIV

Lemme 10.270.
Soient $a < b$ dans $\mathbb{R}$. L'application

$$\begin{aligned} f\colon [a,b] &\to [0,1] \\ x &\mapsto \frac{x-a}{b-a} \end{aligned} \tag{10.741}$$

EQooIINJooAlSqKF

est un C^∞-difféomorphisme[113].

Démonstration. En plusieurs parties.

(1) **Valeurs dans $[0,1]$** ITEMooLRECooOSEqJL Nous devons prouver que pour tout $x \in [a,b]$, nous avons $f(x) \in [0,1]$. D'une part si $x \in [a,b]$, alors $x - a \geqslant 0$ et donc $(x-a)/(b-a) \geqslant 0$.
Dans l'autre sens, si $(x-a)/(b-a) > 1$, alors $x - a > b - a$ et donc $x > b$. Donc $f(x) > 1$ n'arrive jamais pour $x \in [a,b]$.

111. Définition 10.243.
112. Définition 10.55.
113. Définition 10.261.

(2) **Injectif** Si $f(x) = f(t)$, alors en simplifiant par $b - a \neq 0$, nous trouvons $x - a = t - a$ et donc $x = t$ (ne citez le lemme 1.450 que si vous êtes capables de le prouver, sinon faites comme si c'était évident et il ne vous arrivera rien).

(3) **Surjectif** Il est vite vérifié que

$$f^{-1}(t) = t(b - a) + a, \tag{10.742}$$

EQooPSAWooNEJFih

et en procédant de même qu'au point (1), nous voyons que pour tout $t \in [0, 1]$, $f^{-1}(t) \in [a, b]$.

(4) **De classe C^∞** C'est le lemme 10.269 qui fait le travail parce que (10.741) est affine.

(5) **Inverse de classe C^∞** Encore le lemme 10.269 parce que (10.742) est affine.

□

10.13.9 Accroissements finis

LEMooYQZZooVybqjK

Lemme 10.271.
Soit une fonction $f : E \to V$ (espaces vectoriels normés) différentiable en $a \in E$. Alors il existe une fonction $\alpha : E \to V$ telle que

$$\begin{cases} \lim_{h \to 0} \dfrac{\alpha(h)}{\|h\|} = 0 & \text{(10.743a)} \\ f(a + h) = f(a) + df_a(h) + \alpha(h). & \text{(10.743b)} \end{cases}$$

Démonstration. Il s'agit seulement de poser

$$\alpha(h) = f(a + h) - f(a) - df_a(h). \tag{10.744}$$

Le fait que $\alpha(h)/\|h\| \to 0$ est alors la définition de la différentiabilité de f. □

10.13.10 Notations pour les applications linéaires

DEFooTLQUooJvknvi

Définition 10.272 ([354]).
Soient E et F deux espaces vectoriels normés.

(1) Si E et F sont deux espaces vectoriels (pas spécialement normés) nous notons $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F.

*(2) Un **morphisme d'espaces vectoriels normés** est une application linéaire $E \to F$ continue pour la topologie de la norme opérateur.*
L'ensemble des applications linéaires bornées[114] entre E et F est noté $\mathcal{L}^0(E, F)$.

*(3) Un **isomorphisme** est un morphisme continu inversible dont l'inverse est continu. Nous notons $\mathrm{GL}(E, F)$ l'ensemble des isomorphismes entre E et F.*

Le groupe $\mathrm{GL}(E, F)$ a la topologie des espaces normés grâce à la norme opérateur définie en 10.55.

10.13.11 (non ?) Différentiabilité des applications linéaires

LemooXXUGooUqCjmp

Lemme 10.273.
Soit une application linéaire f.

(1) Si f est continue, alors elle est différentiable et $df_a(u) = f(u)$ pour tout a et u.

(2) Si f n'est pas continue, alors elle n'est pas différentiable.

114. Nous avons vu dans la proposition 10.70 que la continuité était équivalente à être bornée pour la norme opérateur de la définition 10.55.

Démonstration. La linéarité de f donne :

$$f(a+h)-f(a)-f(h)=0, \tag{10.745}$$

et donc prendre $T=f$ dans la définition 10.243 fait fonctionner la limite. De plus T est alors continue par hypothèse ; elle est donc bien la différentielle de f.

Supposons que f ne soit pas continue, prenons une application linéaire continue T, et calculons

$$\frac{f(a+h)-f(a)-T(h)}{\|h\|}=\frac{(f-T)(h)}{\|h\|}=(f-T)(e_h) \tag{10.746}$$ EQooFLYMooEKTeQC

où e_h est le vecteur unitaire dans la direction de h. Vu que f n'est pas continue et que T l'est, l'application $f-T$ n'est pas continue. Elle n'est pas pas bornée par la proposition 10.70. Il existe alors un vecteur h tel que $\|(f-T)(e_h)\|>1$ (et même plus grand que ce qu'on veut).

Donc la limite de (10.746) pour $h\to 0$ ne peut pas être nulle. □

LemLLvgPQW

Lemme 10.274.
Une application linéaire continue est de classe C^∞.

Démonstration. Soit $a\in E$. Étant donné que f est linéaire et continue, elle est différentiable et

$$\begin{aligned} df\colon E&\to\mathcal{L}^0(E,F)\\ a&\mapsto f \end{aligned} \tag{10.747}$$

est une fonction constante et en particulier continue ; nous avons donc $f\in C^1$. Pour la différentielle seconde nous avons $d(df)_a=0$ parce que $df(a+h)-df(a)=f-f=0$. Toutes les différentielles suivantes sont nulles. □

10.13.12 Dérivation en chaine et formule de Leibniz

PropOYtgIua

Proposition 10.275.
Soient $f_i\colon U\to F_i$, des fonctions de classe C^r où U est ouvert dans l'espace vectoriel normé E et les F_i sont des espaces vectoriels normés. Alors l'application

$$\begin{aligned} f=f_1\times\cdots\times f_n\colon U&\to F_1\times\cdots\times F_n\\ x&\mapsto\big(f_1(x),\dots,f_n(x)\big) \end{aligned} \tag{10.748}$$

est de classe C^r et

$$d^rf=d^rf_1\times\dots d^rf_n. \tag{10.749}$$

Démonstration. Soit $x\in U$ et $h\in E$. La différentiabilité des fonctions f_i donne

$$f_i(x+h)=f_i(x)+(df_i)_x(h)+\alpha_i(h) \tag{10.750}$$

avec $\lim_{h\to0}\alpha_i(h)/\|h\|=0$. Par conséquent

$$f(x+h)=\big(\dots,f_i(x)+(df_i)_x(h)+\alpha_i(h),\dots\big) \tag{10.751a}$$

$$=\big(\dots,f_i(x),\dots\big)+\big(\dots,(df_i)_x(h),\dots\big)+\big(\dots,\alpha_i(h),\dots\big). \tag{10.751b}$$

Mais la définition 7.226 de la norme dans un espace produit donne

$$\lim_{h\to0}\frac{\|\big(\alpha_1(h),\dots,\alpha_n(h)\big)\|}{\|h\|}=0, \tag{10.752}$$

ce qui nous permet de noter $\alpha(h)=\big(\alpha_1(h),\dots,\alpha_n(h)\big)$ et avoir $\lim_{h\to0}\alpha(h)/\|h\|=0$. Avec tout ça nous avons bien

$$f(x+h)=f(x)+\big((df_1)_x(h),\dots,(df_n)_x(h)\big)+\alpha(h), \tag{10.753}$$

ce qui signifie que f est différentiable et

$$df_x=\big(df_1,\dots,df_n\big). \tag{10.754}$$

□

THOooIHPIooIUyPaf

Théorème 10.276.
Soient des espaces vectoriels normés E, V et W. Nous considérons deux fonctions $f\colon E \to V$ et $g\colon V \to W$. Nous supposons que :

(1) f est différentiable en $a \in E$

(2) g est différentiable en $f(a) \in V$

(3) df_a est de norme finie[115].

Alors $g \circ f\colon E \to W$ est différentiable en a et

$$d(g \circ f)_a(u) = dg_{f(a)}\big(df_a(u)\big), \tag{10.755}$$

ou encore

$$d(g \circ f)_a = dg_{f(a)} \circ df_a. \tag{10.756}$$

Démonstration. En utilisant le lemme 10.271 pour les fonctions f et g, nous avons

$$f(a+h) = f(a) + df_a(h) + \alpha(h) \tag{10.757}$$

EQooXNWZooJSPjBS

et

$$g\big(f(a)+k\big) = g\big(f(a)\big) + dg_{f(a)}(k) + \beta(k). \tag{10.758}$$

EQooIQZZooWPvMbE

L'application $dg_{f(a)} \circ df_a$ est une application linéaire, et est notre candidat différentielle. En suivant la définition 10.243, nous allons calculer

$$\lim_{h\to 0} \frac{(g \circ f)(a+h) - (g \circ f)(a) - (dg_{f(a)} \circ df_a)(h)}{\|h\|}. \tag{10.759}$$

Si cette limite existe et vaut zéro, alors nous aurons prouvé que le candidat différentielle est correct.

Pour cela, nous emboîtons les formules (10.757) et (10.758) l'une dans l'autre pour avoir :

$$\begin{aligned}(g \circ f)(a+h) &= g\big(f(a) + df_a(h) + \alpha(h)\big) && \text{(10.760a)}\\ &= g\big(f(a)\big) + dg_{f(a)}\big(df_a(h) + \alpha(h)\big) + \beta\big(df_a(h) + \alpha(h)\big). && \text{(10.760b)}\end{aligned}$$

Vu que $dg_{f(a)}$ est linéaire, le deuxième terme peut être coupé en deux et après recombinaisons,

$$(g \circ f)(a+h) - (g \circ f)(a) - (dg_{f(a)} \circ df_a)(h) = dg_{f(a)}\big(\alpha(h)\big) + \beta\big(df_a(h) + \alpha(h)\big). \tag{10.761}$$

Étant donné que $dg_{f(a)}$ est linéaire,

$$\frac{dg_{f(a)}\big(\alpha(h)\big)}{\|h\|} = dg_{f(a)}\left(\frac{\alpha(h)}{\|h\|}\right) \to 0. \tag{10.762}$$

Il nous reste à voir que

$$\lim_{h\to 0} \frac{\beta\big(df_a(h) + \alpha(h)\big)}{\|h\|} \tag{10.763}$$

EQooUQNUooFsNyJp

existe au vaut zéro. Vu que df_a est linéaire, il existe $M > 0$ tel que[116] $\|df_a(h)\| \leqslant M\|h\|$. D'autre part, vu que $\alpha(h)/\|h\| \to 0$, nous avons $\|\alpha(h)\| \leqslant \|h\|$ pour tout h suffisamment petit.

Donc si h est assez petit, nous avons

$$\|df_a(h) + \alpha(h)\| \leqslant (M+1)\|h\|. \tag{10.764}$$

EQooEQJBooSmacrD

Soit $\epsilon > 0$. Soit $\delta > 0$ tel que $\|h\| \leqslant \delta$ implique $\beta(h)/\|h\| \leqslant \epsilon$ et (10.764) en même temps. Soit r tel que $(M+1)r < \delta$; et notons que $r < \delta$. Nous considérons alors $h \in B(0,r)$ et nous calculons :

$$\frac{\beta\big(df_a(h) + \alpha(h)\big)}{\|h\|} = \frac{\beta\big(df_a(h) + \alpha(h)\big)}{\|df_a(h) + \alpha(h)\|} \frac{\|df_a(h) + \alpha(h)\|}{\|h\|} \leqslant (M+1)\epsilon. \tag{10.765}$$

La limite (10.763) existe donc et vaut zéro. □

115. Je ne suis pas totalement certain que cette hypothèse soit nécessaire, mais en tout cas, elle est utilisée.

116. Ce M est par exemple la norme opérateur de df_a, comme nous l'assure le lemme 10.66. C'est pour ce passage-ci que nous avons supposé que df_a était de norme finie.

10.13.13 Composition

ThoAGXGuEt

Théorème 10.277 (Différentielle de fonctions composées[355]).
Soient E, F et G des espaces vectoriels normés, U ouvert dans E et V ouvert dans F. Soient des applications de classe C^r ($r \geqslant 1$)

$$f\colon U \to V \tag{10.766a}$$
$$g\colon V \to G. \tag{10.766b}$$

Nous supposons que df_x et $dg_{f(x)}$ sont de norme finie.

Alors l'application $g \circ f\colon U \to G$ est de classe C^r et

$$d(g \circ f)_x = dg_{f(x)} \circ df_x. \tag{10.767}$$

EqHFmezmr

Démonstration. Nous nous fixons $x \in U$. La fonction f est différentiable en $x \in U$ et g en $f(x)$, donc nous pouvons écrire

$$f(x+h) = f(x) + df_x(h) + \alpha(h) \tag{10.768}$$

et

$$g\big(f(x)+u\big) = g\big(f(x)\big) + dg_{f(x)}(u) + \beta(u) \tag{10.769}$$

où la fonction α a la propriété que

$$\lim_{h\to 0} \frac{\|\alpha(h)\|}{\|h\|} = 0; \tag{10.770}$$

et la même chose pour β. La fonction composée en $x+h$ s'écrit donc

$$(g\circ f)(x+h) = g\big(f(x)+df_x(h)+\alpha(h)\big) = g\big(f(x)\big) + dg_{f(x)}\big(df_x(h)+\alpha(h)\big) + \beta\big(df_x(h)+\alpha(h)\big). \tag{10.771}$$

EqCXcfbIH

Nous montrons que tous les « petits » termes de cette formule peuvent être groupés. D'abord si h est proche de 0, nous avons [117]

$$\frac{\|df_x(h)+\alpha(h)\|}{\|h\|} \leqslant \frac{\|df_x\|\|h\|}{\|h\|} + \frac{\|\alpha(h)\|}{\|h\|}. \tag{10.772}$$

Si h est petit, le second terme est arbitrairement petit, donc en prenant n'importe que $M > \|df_x\|$ nous avons

$$\frac{\|df_x(h)+\alpha(h)\|}{\|h\|} \leqslant M. \tag{10.773}$$

Par ailleurs, nous avons

$$\frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|h\|} = \frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|df_x(h)+\alpha(h)\|}\frac{\|df_x(h)+\alpha(h)\|}{\|h\|} \leqslant M\frac{\|\beta\big(df_x(h)+\alpha(h)\big)\|}{\|df_x(h)+\alpha(h)\|}. \tag{10.774}$$

Vu que la fraction est du type $\frac{\beta(f(h))}{f(h)}$ avec $\lim_{h\to 0} f(h) = 0$, la fraction tend vers zéro lorsque $h \to 0$. En posant

$$\gamma_1(h) = \beta\big(df_x(h)+\alpha(h)\big) \tag{10.775}$$

nous avons $\lim_{h\to 0}\gamma_1(h)/\|h\| = 0$.

L'autre candidat à être un petit terme dans (10.771) est traité en utilisant le lemme 10.69 :

$$\|dg_{f(x)}\big(\alpha(h)\big)\| \leqslant \|dg_{f(x)}\|\|\alpha(h)\|. \tag{10.776}$$

Donc

$$\frac{\|dg_{f(x)}\big(\alpha(h)\big)\|}{\|h\|} \leqslant \|dg_{f(x)}\|\frac{\|\alpha(h)\|}{\|h\|}, \tag{10.777}$$

ce qui nous permet de poser

$$\gamma_2(h) = dg_{f(x)}\big(\alpha(h)\big) \tag{10.778}$$

117. Ici nous utilisons l'hypothèse de norme finie pour la différentielle.

avec γ_2 qui a la même propriété que γ_1. Avec tout cela, en posant $\gamma = \gamma_1 + \gamma_2$ nous récrivons

$$(g \circ f)(x+h) = g\big(f(x)\big) + dg_{f(x)}\big(df_x(h)\big) + \gamma(h) \tag{10.779}$$

avec $\lim_{h\to 0} \frac{\gamma(h)}{\|h\|} = 0$. Tout cela pour dire que

$$\lim_{h\to 0} \frac{(g \circ f)(x+h) - (g \circ f)(x) - \big(dg_{f(x)} \circ df_x\big)(h)}{\|h\|} = 0, \tag{10.780}$$

ce qui signifie que

$$d(g \circ f)_x = dg_{f(x)} \circ df_x. \tag{10.781}$$

Nous avons donc montré que si f et g sont différentiables, alors $g \circ f$ est différentiable avec différentielle donnée par (10.767).

Nous passons à la régularité. Nous supposons maintenant que f et g sont de classe C^r et nous considérons l'application

$$\begin{aligned} \varphi\colon L(F,G) \times L(E,F) &\to L(E,G) \\ (A,B) &\mapsto A \circ B. \end{aligned} \tag{10.782}$$

Montrons que l'application φ est continue en montrant qu'elle est bornée[118]. Pour cela nous écrivons la norme opérateur

$$\|\varphi\| = \sup_{\|(A,B)\|=1} \|\varphi(A,B)\| = \sup_{\|(A,B)\|=1} \|A \circ B\| \leqslant \sup_{\|(A,B)\|=1} \|A\|\|B\| \leqslant 1. \tag{10.783}$$

Justifications : d'une part la norme opérateur est une norme algébrique[119], et d'autre part la définition 7.226 de la norme sur un espace produit pour la dernière majoration. L'application φ est donc continue et donc C^∞ par le lemme 10.274. Nous considérons également l'application

$$\begin{aligned} \psi\colon U &\to L(F,G) \times L(E,F) \\ x &\mapsto \big(dg_{f(x)}, df_x\big). \end{aligned} \tag{10.784}$$

Vu que f et g sont C^1, l'application ψ est continue. Ces deux applications φ et ψ sont choisies pour avoir

$$(\varphi \circ \psi)(x) = \varphi\big(dg_{f(x)}, df_x\big) = dg_{f(x)} \circ df_x, \tag{10.785}$$

c'est-à-dire $\varphi \circ \psi = d(g \circ f)$. Les applications φ et ψ étant continues, l'application $d(g \circ f)$ est continue, ce qui prouve que $g \circ f$ est C^1.

Si f et g sont C^r alors $dg \in C^{r-1}$ et $dg \circ f \in C^{r-1}$ où il ne faut pas se tromper : $dg\colon F \to L(F,G)$ et $f\colon U \to F$; la composée est $dg \circ f\colon x \mapsto dg_{f(x)} \in L(F,G)$.

Pour la récurrence nous supposons que $f, g \in C^{r-1}$ implique $g \circ f \in C^{r-1}$ pour un certain $r \geqslant 2$ (parce que nous venons de prouver cela avec $r = 1$ et $r = 2$). Soient $f, g \in C^r$ et montrons que $g \circ f \in C^r$. Par la proposition 10.275 nous avons

$$\psi = dg \circ f \times df \in C^{r-1}, \tag{10.786}$$

et donc $d(g \circ f) = \varphi \circ \psi \in C^{r-1}$, ce qui signifie que $g \circ f \in C^r$. □

PROPooRCZOooSgvpSE

Proposition 10.278 ([1]).
Soit une application $f\colon E \to V$ de classe C^p. Soit une application linéaire $\varphi\colon V \to W$. Alors $\varphi \circ f$ est de classe C^p.

118. Proposition 10.70.
119. Lemme 10.69.

Démonstration. Toute la preuve est un grand jeu de cohérence des espaces en présence, alors soyez attentifs et capable de dire précisément à quel espace appartient chacun de objets entrant en jeu.

Nous posons $V_0 = V$ et $V_{k+1} = \mathcal{L}(E, V_k)$. Idem pour les espaces W_k. Ensuite nous posons

$$\begin{aligned}\varphi_1 \colon \mathcal{L}(E,V) &\to \mathcal{L}(E,W)\\ \alpha &\mapsto \varphi\circ\alpha.\end{aligned} \tag{10.787}$$

et

$$\begin{aligned}\varphi_k \colon \mathcal{L}(E,V_{k-1}) &\to \mathcal{L}(E,W_{k-1})\\ \alpha &\mapsto \varphi_{k-1}\circ\alpha.\end{aligned} \tag{10.788}$$

Notez la cohérence : si $a \in E$, $\alpha(a) \in V_{k-1} = \mathcal{L}(E, V_{k-2})$, et donc

$$(\varphi_{k-1}\circ\alpha)(a) = \varphi_{k-1}\big(\alpha(a)\big). \tag{10.789}$$

À droite nous avons $\varphi_{k-1}\big(\alpha(a)\big) \in \mathcal{L}(E, W_{k-2}) = V_{k-1}$.

De plus, φ est linéaire ; ça se prouve par récurrence en partant de φ_1 et en se basant sur le fait que φ est linéaire.

C'est parti pour une récurrence.

(1) **Énoncé** Nous allons prouver par récurrence que

$$d^k(\varphi\circ f) = \varphi_k \circ d^k f. \tag{10.790}$$

pour tout $k \leqslant p$.

(2) **Initialisation**

D'abord, f est de classe C^p, donc différentiable et φ est linéaire donc différentiable. Donc la composée est différentiable et le théorème 10.276 nous donne la différentiabilité de $\varphi\circ f$ ainsi que la formule

$$d(\varphi\circ f)_a(u) = d\varphi_{f(a)}\big(df_a(u)\big) = (\varphi\circ df_a)(u) = \varphi_1(df_a)(u). \tag{10.791}$$

Donc $d(\varphi\circ f)_a = \varphi_1(df_a)$, ce qui signifie

$$d(\varphi\circ f) = \varphi_1 \circ df. \tag{10.792}$$

C'est bon pour $k = 1$.

(3) **Le pas de récurrence**

Vu que f est de classe C^p, $d^k f$ est encore différentiable. Vu que φ_k est encore linéaire, nous pouvons encore utiliser la règle de différentiation de fonctions composées sur l'application $\varphi_k \circ d^k f$. Nous avons :

$$d^{k+1}(\varphi\circ f)_a(u) = d\big(d^k(\varphi\circ f)\big)_a(u) = d(\varphi_k\circ d^k f)_a(u). \tag{10.793}$$

C'est le moment d'utiliser la formule de différentiation en chaine :

$$d^{k+1}(\varphi\circ f)_a(u) = \big((d\varphi_k)_{d^k f_a}\circ d^{k+1}f_a\big)(u). \tag{10.794}$$

Mais φ_k étant linéaire, $(d\varphi_k)_{d^k f_a} = \varphi_k$, donc

$$d^{k+1}(\varphi\circ f)_a(u) = (\varphi_k\circ d^{k+1}f_a)(u). \tag{10.795}$$

Donc, en oubliant l'application au vecteur u,

$$d^{k+1}(\varphi\circ f)_a = \varphi_k\circ d^{k+1}f_a = \varphi_{k+1}\big(d^{k+1}f_a\big) = (\varphi_{k+1}\circ d^{k+1}f)(a). \tag{10.796}$$

Nous avons donc bien

$$d^{k+1}(\varphi\circ f) = \varphi_{k+1}\circ d^{k+1}f. \tag{10.797}$$

□

PROPooGJGVooWnbeAe

Proposition 10.279 ([1]).
Soient des espaces vectoriels E, V et W de dimension finie, et une fonction $f\colon E \to V$ de classe C^p. Si $\varphi\colon V \to W$ est linéaire, alors

$$\varphi \circ f\colon E \to W \tag{10.798}$$

est de classe C^p.

Démonstration. En utilisant le théorème de différentiation de fonctions composées 10.276,

$$d(\varphi \circ f)_a(u) = d\varphi_{f(a)} df_a(u), \tag{10.799}$$

et donc, parce que φ est linéaire,

$$d(\varphi \circ f)_a = \varphi \circ df_a. \tag{10.800}$$

Nous pouvons exprimer cela de façon un peu différente en posant $\varphi_1\colon \mathcal{L}(E,V) \to \mathcal{L}(E,W)$,

$$\varphi_1(\alpha)(a) = (\varphi \circ \alpha)(a). \tag{10.801}$$

Cela nous permet d'écrire $\varphi \circ df_a = (\varphi_1 \circ df)(a)$ et donc

$$d(\varphi \circ f) = \varphi_1 \circ df \tag{10.802}$$

EQooUJPWooTzgSJx

où φ_1 est encore une application linéaire. Une récurrence semble possible. Nous posons $V_0 = V$ et $W_0 = W$ puis

$$V_{k+1} = \mathcal{L}(E, V_k) \tag{10.803a}$$
$$W_{k+1} = \mathcal{L}(E, W_k) \tag{10.803b}$$

et

$$\begin{aligned} \varphi_k\colon \mathcal{L}(E, V_{k-1}) &\to \mathcal{L}(E, W_{k-1}) \\ g &\mapsto \varphi_{k-1} \circ g. \end{aligned} \tag{10.804}$$

Avec tout cela, nous prétendons que $d^k(\varphi \circ f) = \varphi_k \circ d^k f$ avec φ_k linéaire.

(1) φ_k **est linéaire** Soient $\alpha_1, \alpha_2 \in \mathcal{L}(E, V_{k-1})$, ainsi que $\lambda, \mu \in \mathbb{K}$. Nous avons, en utilisant la linéarité de φ_{k-1} :

$$\varphi_k(\lambda\alpha_1 + \mu\alpha_2)(a) = \varphi_{k-1}\big((\lambda\alpha_1 + \mu\alpha_2)(a)\big) \tag{10.805a}$$
$$= \varphi_{k-1}\big(\lambda\alpha_1(a)\big) + \mu\varphi_{k-1}\big(\alpha_2(a)\big) \tag{10.805b}$$
$$= \lambda\varphi_k(\alpha_1)a + \mu\varphi_k(\alpha_2)a. \tag{10.805c}$$

Donc φ_k est linéaire pour tout k.

(2) **La relation** La relation

$$d^k(\varphi \circ f) = \varphi_k \circ d^k f \tag{10.806}$$

se démontre par récurrence, chaque pas étant justifié de la même manière que (10.802).

□

PROPooLRRMooRzrTNE

Proposition 10.280 (Composition).
Soient trois espaces vectoriels normés E, V, W de dimension finie. Si les applications $f\colon E \to V$ et $g\colon V \to W$ sont de classe C^p, alors la composée $g \circ f\colon E \to W$ est de classe C^p.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

10.13.14 Règle de Leibniz

PROPooNUAJooMtJBySPROPooNUAJooMtJBySPROPooNUAJooMtJBySP

$\mathcal{L}(E, V \otimes W)$. Et je ne vous dit même pas à quel point il n'est pas évident, si $f \in C^\infty(E, V)$ et $g \in C^\infty(E, W)$ que nous aurions $f \otimes g \in C^\infty(E, V) \otimes C^\infty(E, W) = C^\infty(E, V \otimes W)$.

Tout cela pour dire que nous n'allons pas nous lancer dans des abus de notations. Non. Au lieu de cela, nous introduisons une notation. Pour rappel, dans tout le Frido, $\mathrm{Fun}(A, B)$ désigne l'ensemble de toutes les application de A vers B sans suppositions de régularité. Pour les puristes, nous précisions que si $f \in \mathrm{Fun}(A, B)$, nous supposons que f est définie sur tout A. hum ...sauf mention du contraire.

DEFooMVNDooFWFtRn

Définition 10.286.
Si $f \in \mathrm{Fun}(E, V)$ et $g \in \mathrm{Fun}(E, W)$, alors nous définissons

$$\begin{aligned} f\tilde{\otimes}g\colon E &\to V \otimes W \\ a &\mapsto f(a) \otimes g(a). \end{aligned} \tag{10.814}$$

PROPooCRVXooEGxdZl

Proposition 10.287.
Soient des applications continues $f\colon E \to V$ et $g\colon E \to W$ entre espaces vectoriels de dimension finies. Alors la fonction $f\tilde{\otimes}g\colon E \to V \otimes W$ est continue.

Démonstration. Soient $a \in E$ ainsi qu'une suite $x_k \to a$ dans E. Nous voulons prouver que $f\tilde{\otimes}g(x_k) \xrightarrow{V\otimes W} f(a) \otimes g(a)$. Nous avons :

$$\|f(x_k)\otimes g(x_k) - f(a)\otimes g(a)\| \leqslant \|f(x_k)\otimes g(x_k) - f(x_k)\otimes g(a)\| + \|f(x_k)\otimes g(a) - f(a)\otimes g(a)\|. \tag{10.815}$$

EQooSNXUooXrYOgY

Ensuite en utilisant la classe d'équivalence (10.456b),

$$f(x_k) \otimes g(x_k) - f(x_k) \otimes g(a) = f(x_k) \otimes \big(g(x_k) - g(a)\big), \tag{10.816}$$

et en ce qui concerne les normes,

$$\|f(x_k) \otimes g(x_k) - f(x_k) \otimes g(a)\| = \|f(x_k)\|\|\big(g(x_k) - g(a)\big)\|. \tag{10.817}$$

Mais par hypothèse, $f(x_k) \to f(a)$ et $g(x_k) \to g(a)$. Donc le tout tend vers zéro lorsque $k \to \infty$.

Le même raisonnement fonctionne avec le second terme de (10.815). □

Lorsque nous parlons de différentielle de produit de fonctions, nous voulons étudier la différentiabilité de $f\tilde{\otimes}g$ sous l'hypothèse de différentiabilité de f et g. Et aussi, si f et g sont de classe C^p, est-ce que $f\tilde{\otimes}g$ est également de classe C^p ?

Nous voudrions avoir une formule du type

$$d(f\tilde{\otimes}g) = df\tilde{\otimes}g + f\tilde{\otimes}dg, \tag{10.818}$$

mais ça ne colle pas au niveau des espaces. En effet, en évaluant cela en $a \in E$, nous avons à gauche $d(f\tilde{\otimes}g)_a \in \mathcal{L}(E, V \otimes W)$, tandis qu'à droite nous avons $df_a \otimes g(a) \in \mathcal{L}(E, V) \otimes W$ et $f(a) \otimes dg_a \in V \otimes \mathcal{L}(E, W)$.

Nous pourrions bien entendu dire que $V \otimes \mathcal{L}(E, W)$ est isomorphe à $\mathcal{L}(E, V \otimes W)$ et hop voilà, on n'en parle plus. Ce serait passer sur deux points importants. D'abord est-ce que $V \otimes \mathcal{L}(E, W)$ est vraiment isomorphe à $\mathcal{L}(E, V \otimes W)$? Et ensuite, l'isomorphisme implique une utilisation du théorème 10.276 qui est tout sauf une trivialité.

Bref, fidèle au principe fridesque de ne pas cacher des difficultés techniques sous des abus de notations, nous allons écrire les choses explicitement.

Lemme 10.288.
Si E, V et W sont de dimension finie, les applications

$$\begin{aligned} \psi\colon \mathcal{L}(E, V) \otimes W &\to \mathcal{L}(E, V \otimes W) \\ f \otimes w &\mapsto \Big(u \mapsto f(u) \otimes w\Big) \end{aligned} \tag{10.819}$$

EQooVWXBooCesUqH

et

$$\begin{aligned}\varphi\colon V\otimes\mathcal{L}(E,W)&\to\mathcal{L}(E,V\otimes W)\\ v\otimes g&\mapsto\big(a\mapsto v\otimes g(a)\big).\end{aligned}\tag{10.820}$$

sont des isomorphismes d'espaces vectoriels.

Dans le meilleur des mondes, ces applications devraient être affublés d'indices V et W.

Démonstration. Nous donnons des détails à propos de ψ. Pour φ c'est la même chose.

(1) **Linéaire** La formule (10.819) définit ψ en particulier sur une base de $\mathcal{L}(E,V)\otimes W$ par la proposition 10.179(1). Ce que signifie réellement la formule (10.819) est que ψ est ainsi définie sur la base et est prolongée par continuité.

(2) **Injective** Si pour un f et un w fixé nous avons $\psi(f\otimes w)=0$, alors il y a deux cas : soit $w=0$ soit $w\neq 0$. Dans le premier cas, $f\otimes w=0$, et dans le second cas, nous remarquons que

$$0=\psi(f\otimes w)(a)=f(a)\otimes w\tag{10.821}$$

pour tout $a\in E$. Cela implique $f(a)=0$ pour tout a et donc $f=0$, ce qui signifie que $f\otimes w=0$.

(3) **Bijective** En utilisant la proposition 10.179 et le lemme 4.42(2), nous avons égalité des dimensions entre $\mathcal{L}(E,V)\otimes W$ et $\mathcal{L}(E,V\otimes W)$.
Une application linéaire injective entre deux espaces vectoriels de même dimension (finie) est une bijection.

□

PROPooZOAFooRMeBgI

Proposition 10.289.
Soient des espaces vectoriels normés de dimension finie. Soient $f\colon E\to V$ et $g\colon E\to W$ des fonctions de classe C^1. Alors $f\tilde{\otimes}g\colon E\to V\otimes W$ est de classe C^1 nous avons les formules

$$d(f\tilde{\otimes}g)_a(u)=df_a(u)\otimes g(a)+f(a)\otimes dg_a(u)\tag{10.822}$$

EQooSUSCooBhZXFC

ainsi que

$$d(f\tilde{\otimes}g)=\psi\circ(df\tilde{\otimes}g)+\varphi\circ(f\tilde{\otimes}dg).\tag{10.823}$$

EQooOCEFooUrsIDd

Démonstration. Nous commençons par prouver que $f\tilde{\otimes}g$ est différentiable en injectant le candidat (10.822) dans la définition. Au numérateur nous avons :

$$(f\tilde{\otimes}g)(a+h)-(f\tilde{\otimes}g)(a)-df_a(h)\otimes g(a)-f(a)\otimes dg_a(h).\tag{10.824}$$

EQooOMXSooYsAiKh

Le lemme 10.271 assure qu'il existe une fonction $\alpha\colon E\to V$ telle que $\lim_{h\to0}\alpha(h)/\|h\|=0$ et $f(a+h)=f(a)+df_a(h)+\alpha(h)$. Même chose pour g. Nous avons donc

$$(f\tilde{\otimes}g)(a+h)=f(a+h)\otimes g(a+h)=\big(f(a)+df_a(h)+\alpha(h)\big)\otimes\big(g(a)+dg_a(h)+\beta(h)\big)\tag{10.825}$$

qui se développe en 9 termes. En effectuant les différences dans (10.824), nous nous retrouvons avec un numérateur qui vaut

$$f(a)\otimes\beta(h)+df_a(h)\otimes dg_a(h)+df_a(h)\otimes\beta(h)+\alpha(h)\otimes g(a)+\alpha(h)\otimes dg_a(h)+\alpha(h)\otimes\beta(h).\tag{10.826}$$

Nous pouvons prouver terme à terme qu'en divisant par $\|h\|$ nous avons une limite qui vaut zéro. Par exemple,

$$\lim_{h\to0}\frac{f(a)\otimes\beta(h)}{\|h\|}\tag{10.827}$$

se calcule en prenant la norme du numérateur et en utilisant le lemme 10.234 :

$$\frac{\|f(a)\otimes\beta(h)\|}{\|h\|}=\frac{\|f(a)\|\|\beta(h)\|}{\|h\|}\to0.\tag{10.828}$$

Tous les termes contenant $\alpha(h)$ ou $\beta(h)$ se traitent de la même manière. Le dernier terme à traiter est

$$\lim_{h\to 0} \frac{df_a(h)\otimes dg_a(h)}{\|h\|}. \tag{10.829}$$

En prenant la norme du numérateur, en utilisant encore le lemme 10.234 et en utilisant le lemme 10.66, nous avons

$$\|df_a(h)\otimes dg_a(h)\| = \|df_a(h)\|\|dg_a(h)\| \leqslant \|df_a\|\|dg_a\|\|h\|^2, \tag{10.830}$$

donc

$$\lim_{h\to 0} \frac{df_a(h)\otimes dg_a(h)}{\|h\|} = 0. \tag{10.831}$$

Notons que l'utilisation du lemme 10.66 requière que df_a soit continue, ce qui n'est pas évident en dimension infinie : une application linéaire n'est pas spécialement continue. C'est donc ici que nous utilisons le fait que E, V et W sont de dimension finie [121].

Ceci prouve que $f\tilde{\otimes}g$ est différentiable et nous donne la formule (10.822) pour appliquer sa différentielle à un élément de E. La formule (10.823) est un corolaire : elle se vérifie en l'appliquant à a puis à u.

Pour terminer nous devons prouver que $d(f\tilde{\otimes}g)$ est continue. Vu que f et g sont de classe C^1, les applications f, g, df et dg sont continues. Les applications ψ et φ sont également continues parce que linéaires sur des espaces de dimension finie. La proposition 10.287 appliquée à df et g montre que $df\tilde{\otimes}g$ est continue. La composition avec ψ qui est linéaire conserve la continuité.

Donc le membre de droite de (10.823) est continu et $f\tilde{\otimes}g$ a une différentielle continue. Elle est donc de classe C^1. □

Il est temps de démontrer le truc difficile, à savoir que si f et g sont de classe C^p, alors $f\tilde{\otimes}g$ est également de classe C^p.

PROPooAWZFooMIhoCN

Proposition 10.290.
Nous appelons P_k la propriété suivante :

> *Pour tout espaces vectoriels normés E, V, W de dimension finies et pour toutes applications $f\colon E\to V$ et $g\colon E\to W$ de classe C^k, la fonction $f\tilde{\otimes}g$ est de classe C^k.*

ITEMooDQRYooAEdxrW

(1) La propriété P_k est vraie pour tout k.

ITEMooUUIFooGDyTMM

(2) Si $f\colon E\to V$ et $g\colon E\to W$ sont de classe C^p, alors $f\tilde{\otimes}g\colon E\to V\otimes W$ est de classe C^p.

Démonstration. Il est compliqué de prouver le point (2) directement par récurrence pour des f, g, E, V et W fixés. La raison de cette difficulté est que les espaces en jeu dans les différentielles de f et g aux différents ordres sont tous différents. Il faut donc, pour chaque ordre de différentielle, repartir sur de tout nouveaux espaces vectoriels. C'est pour cela que nous incluons « pour tout espaces vectoriels » dans la propriété de récurrence.

Bref, une fois que (1) sera démontrée, le point (2) sera immédiat.

(1) **Pour $k=0$** Pour $k=0$. Nous supposons que f et g sont de classe C^0, c'est-à-dire continues. La proposition 10.287 montre alors que $f\tilde{\otimes}g$ est continue, c'est-à-dire de classe C^0.

(2) **Pour $k=1$** Bien que ce ne soit pas tout à fait nécessaire, nous prouvons que P_1 est également vraie avant de passer à la récurrence. Si f et g sont de classe C^1, alors la proposition 10.289 s'applique : $f\tilde{\otimes}g$ est de classe C^1.

(3) **Pour $k+1$** Nous faisons la récurrence en supposant que P_k est vraie, et en prouvant que P_{k+1} est vraie. Nous considérons des applications $f\colon E\to V$ et $g\colon E\to W$ de classe C^{k+1}. La proposition 10.289 dit que $f\tilde{\otimes}g$ est de classe C^1 et que

$$d(f\tilde{\otimes}g) = \psi\circ(df\tilde{\otimes}g) + \varphi\circ(f\tilde{\otimes}dg). \tag{10.832}$$

121. Il y a surement moyen de paufiner, et d'affaiblir cette hypothèse, mais je ne me lance pas là-dedans.

À droite, df et g sont de classe C^k parce que f et g sont de classe C^{k+1}. Donc $df\tilde{\otimes}g$ est de classe C^k par l'hypothèse de récurrence appliquée aux espaces $\mathcal{L}(E,V)$ et W. La proposition 10.278 nous assure alors que $\psi\circ(df\tilde{\otimes}g)$ est de classe C^k également.
Nous avons prouvé que $d(f\tilde{\otimes}g)$ est de classe C^k, donc $f\tilde{\otimes}g$ est de classe C^{k+1}. Cela nous fait la récurrence.

□

10.13.17 Formule des accroissements finis

PropDQLhSoy

Proposition 10.291.
Soient $a<b$ dans $\mathbb{R}$ et deux fonctions

$$f\colon [a,b]\to E \tag{10.833a}$$
$$g\colon [a,b]\to \mathbb{R} \tag{10.833b}$$

continues sur $[a,b]$ et dérivables sur $]a,b[$. Si pour tout $t\in]a,b[$ nous avons $\|f'(t)\|\leqslant g'(t)$ alors

$$\|f(b)-f(a)\|\leqslant g(b)-g(a). \tag{10.834}$$

Démonstration. Soit $\epsilon>0$ et la fonction

$$\begin{aligned}\varphi_\epsilon\colon [a,b]&\to\mathbb{R}\\ t&\mapsto \|f(t)-f(a)\|-g(t)-\epsilon t.\end{aligned} \tag{10.835}$$

Cela est une fonction continue réelle à variable réelle. En particulier pour tout $u\in]a,b[$ la fonction φ_ϵ est continue sur le compact $[u,b]$ et donc y atteint son minimum en un certain point $c\in[u,b]$; c'est le bon vieux théorème de Weierstrass 7.146. Nous commençons par montrer que pour tout u, ledit minimum ne peut être que b. Pour cela nous allons montrer que si $t\in[u,b[$, alors $\varphi_\epsilon(s)<\varphi_\epsilon(t)$ pour un certain $s>t$. Par continuité si s est proche de t nous avons

$$\left\|\frac{f(s)-f(t)}{s-t}\right\|-\frac{\epsilon}{2}<\|f'(t)\|<g'(t)+\frac{\epsilon}{2}=\frac{g(s)-g(t)}{s-t}+\frac{\epsilon}{2}. \tag{10.836}$$

Ces inégalités proviennent de la limite

$$\lim_{s\to t}\frac{f(s)-f(t)}{s-t}=f'(t), \tag{10.837}$$

donc si s et t sont proches,

$$\left\|\frac{f(s)-f(t)}{s-t}-f'(t)\right\| \tag{10.838}$$

est petit. Si $s>t$ nous pouvons oublier des valeurs absolues et transformer l'inégalité en

$$\|f(s)-f(t)\|<g(s)-g(t)+\epsilon(s-t). \tag{10.839}$$

Utilisant cela et l'inégalité triangulaire,

$$\varphi_\epsilon(s)\leqslant\|f(s)-f(t)\|+\|f(t)-f(a)\|-g(s)-\epsilon s \tag{10.840a}$$
$$\leqslant g(s)-g(t)+\epsilon s-\epsilon t+\|f(t)-f(a)\|-g(s)-\epsilon s \tag{10.840b}$$
$$=\varphi_\epsilon(t). \tag{10.840c}$$

Donc nous avons bien $\varphi_\epsilon(s)<\varphi_\epsilon(t)$ avec l'inégalité stricte. Par conséquent pour tout $u\in]a,b[$ nous avons $\varphi_\epsilon(b)<\varphi_\epsilon(u)$ et en prenant la limite $u\to a$ nous avons

$$\varphi_\epsilon(b)\leqslant\varphi_\epsilon(a). \tag{10.841}$$

Cette inégalité donne immédiatement

$$\|f(b)-f(a)\|\leqslant g(b)-g(a)+\epsilon(b-a) \tag{10.842}$$

pour tout $\epsilon>0$ et donc

$$\|f(b)-f(a)\|\leqslant g(b)-g(a). \tag{10.843}$$

□

ThoNAKKght

Théorème 10.292 (Théorème des accroissements finis)**.**
Soient E et F des espaces vectoriels normés, U ouvert dans E et une application différentiable $f\colon U \to F$. Pour tout segment $[a,b] \subset U$ nous avons

$$\|f(b)-f(a)\| \leqslant \left(\sup_{x\in[a,b]} \|df_x\| \right) \|b-a\|. \tag{10.844}$$

Démonstration. Nous prenons les applications

$$\begin{aligned} k\colon [0,1] &\to E \\ t &\mapsto f\big((1-t)a+tb\big) \end{aligned} \tag{10.845}$$

et

$$\begin{aligned} g\colon [0,1] &\to \mathbb{R} \\ t &\mapsto t \sup_{x\in[a,b]} \|df_x\|\|b-a\|. \end{aligned} \tag{10.846}$$

Pour tout t nous avons $g'(t) = M\|b-a\|$ où il n'est besoin de dire ce qu'est M. D'un autre côté nous avons aussi

$$\begin{aligned} k'(t) &= \lim_{\epsilon\to 0} \frac{f\big((1-t-\epsilon)a+(t+\epsilon)b\big) - f\big((1-t)a+tb\big)}{\epsilon} \\ &= \frac{d}{d\epsilon}\Big[f\big((1-t)a+tb+\epsilon(b-a)\big)\Big]_{\epsilon=0} \\ &= df_{(1-t)a+tb}(b-a) \end{aligned} \tag{10.847}$$

où nous avons utilisé l'hypothèse de différentiabilité de f sur $[a,b]$ et donc en $(1-t)a+tb$. Nous avons donc

$$\|k'(t)\| \leqslant \|b-a\|\|df_{(1-t)a+tb}\| \leqslant M\|b-a\| = g'(t) \tag{10.848}$$

La proposition 10.291 est donc utilisable et

$$\|k(1)-k(0)\| \leqslant g(1)-g(0), \tag{10.849}$$

c'est-à-dire

$$\|f(b)-f(a)\| \leqslant M\|b-a\| \tag{10.850}$$

comme il se doit. □

ProFSjmBAt

Proposition 10.293.
Soient E et F des espaces vectoriels normés, U ouvert dans E et une application $f\colon U \to F$. Soient $a,b \in U$ tels que $[a,b] \subset U$. Nous posons $u = (b-a)/\|b-a\|$ et nous supposons que pour tout $x \in [a,b]$, la dérivée directionnelle

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} \tag{10.851}$$

existe. Nous supposons de plus que $\frac{\partial f}{\partial u}(x)$ est continue en $x=a$. Alors

$$\|f(b)-f(a)\| \leqslant \left(\sup_{x\in[a,b]} \|\frac{\partial f}{\partial u}(x)\| \right) \|b-a\|. \tag{10.852}$$

Démonstration. Nous posons évidemment

$$M = \sup_{x\in[a,b]} \|\frac{\partial f}{\partial u}(x)\| \tag{10.853}$$

et nous considérons les fonctions

$$k(t) = f\big((1-t)a + tb\big) \tag{10.854}$$

et

$$g(t) = tM\|b-a\|. \tag{10.855}$$

Pour alléger les notations nous posons $x = (1-t)a+tb$ et nous calculons avec un petit changement de variables dans la limite :

$$k'(t) = \frac{d}{d\epsilon}\Big[f\big(x+\epsilon(b-a)\big)\Big]_{\epsilon=0} = \|b-a\|\frac{d}{d\epsilon}\Big[f\big(x+\frac{\epsilon}{\|b-a\|}(b-a)\big)\Big]_{\epsilon=0} = \|b-a\|\frac{\partial f}{\partial u}(x), \tag{10.856}$$

et donc encore une fois nous avons

$$\|k'(t)\| \leqslant g'(t), \tag{10.857}$$

ce qui donne

$$\|k(1)-k(0)\| = g(1)-g(0), \tag{10.858}$$

c'est-à-dire

$$\|f(b)-f(a)\| \leqslant \sup_{x\in[a,b]}\|\frac{\partial f}{\partial u}(x)\|\|b-a\|. \tag{10.859}$$

□

ThoOYwdeVt

Théorème 10.294.
Soient E, V deux espaces vectoriels normés, une application $f\colon E \to V$, un point $a \in E$ tel que pour tout $u \in E$, la dérivée

$$\frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} \tag{10.860}$$

existe pour tout $x \in B(a,r)$ et est continue (par rapport à x) en $x = a$. Nous supposons de plus que

$$\frac{\partial f}{\partial u}(a) = 0 \tag{10.861}$$

pour tout $u \in E$. Alors f est différentiable en a et

$$df_a = 0 \tag{10.862}$$

Démonstration. Soit $\epsilon > 0$. Pourvu que $\|h\|$ soit assez petit pour que $a+h \in B(a,r)$, la proposition 10.293 nous donne

$$\|f(a+h)-f(a)\| \leqslant \sup_{x\in[a,a+h]}\|\frac{\partial f}{\partial u}(x)\||h| \tag{10.863}$$

où $u = h/\|h\|$. Par continuité de $\partial_u f(x)$ en $x = a$ et par le fait que cela vaut 0 en $x = a$, il existe un $\delta > 0$ tel que si $\|h\| < \delta$ alors

$$\|\frac{\partial f}{\partial u}(a+h)\| \leqslant \epsilon. \tag{10.864}$$

Pour de tels h nous avons

$$\|f(a+h)-f(a)\| \leqslant \epsilon\|h\|, \tag{10.865}$$

ce qui prouve que l'application linéaire $T(u) = 0$ convient parfaitement pour faire fonctionner la définition 10.243. □

10.13.18 Applications multilinéaires

Nous avons déjà parlé d'applications multilinéaires dans la définition 8.1.

LemFRdNDCd

Lemme 10.295 (Leibniz pour les formes bilinéaires[355]).
Si $B\colon E\times F\to G$ est bilinéaire et continue, elle est C^∞ et

$$dB_{(x,y)}(u,v)=B(x,v)+B(u,y). \tag{10.866}$$

EqXYJgDBt

Démonstration. D'abord le membre de droite de (10.866) est une application linéaire et continue, donc c'est un bon candidat à être différentielle. Nous allons prouver que ça l'est, ce qui prouvera la différentiabilité de B. Avec ce candidat, le numérateur de la définition (10.243) s'écrit dans notre cas

$$B\big((x,y)+(u,v)\big)-B(x,y)-B(x,v)-B(u,y)=B(u,v). \tag{10.867}$$

Il reste à voir que

$$\lim_{(u,v)\to(0,0)}\frac{B(u,v)}{\|(u,v)\|}=0 \tag{10.868}$$

Par l'équation (10.222) nous avons

$$\frac{\|B(u,v)\|}{\|(u,v)\|}\leqslant\frac{\|B\|\|u\|\|v\|}{\|u\|}=\|B\|\|v\| \tag{10.869}$$

parce que $\|(u,v)\|\geqslant\|u\|$. À partir de là il est maintenant clair que

$$\lim_{(u,v)\to(0,0)}\frac{\|B(u,v)\|}{\|(u,v)\|}=0, \tag{10.870}$$

ce qu'il fallait. □

Proposition 10.296 (Règle de Leibniz[355]).
Soient E,F_1,F_2 des espaces vectoriels normés, U ouvert dans E et des applications de classe C^r ($r\geqslant 1$)

$$f_1\colon U\to F_1 \tag{10.871a}$$
$$f_2\colon U\to F_2 \tag{10.871b}$$
$$\tag{10.871c}$$

et $B\in\mathcal{L}^0(F_1\times F_2,G)$. Alors l'application

$$\begin{aligned}\varphi\colon U&\to G\\ x&\mapsto B\big(f_1(x),f_2(x)\big)\end{aligned} \tag{10.872}$$

est de classe C^r et

$$d\varphi_x(u)=\varphi\big((df_1)_x(u),f_2(x)\big)+\varphi\big(f_1(x),(df_2)_x(u)\big). \tag{10.873}$$

EqMNGBXWc

Démonstration. Par hypothèse B est continue (c'est la définition de l'espace $\mathcal{L}^0$), et donc C^∞ par le lemme 10.295. Par ailleurs la fonction $f_1\times f_2$ est de classe C^r parce que f_1 et f_2 le sont et parce que la proposition 10.275 le dit. L'application composée $B\circ(f_1\times f_2)$ est donc également de classe C^r par le théorème 10.277.

Il ne nous reste donc qu'à prouver la formule 10.873. En utilisant la différentielle du produit cartésien [122] nous avons

$$f\big(B\circ(f_1\times f_2)\big)_x(h)=dB_{(f_1\times f_2)(x)}\big((df_1)_x(h),(df_2)_x(h)\big). \tag{10.874}$$

Nous développons cela en utilisant le lemme 10.295 :

$$\begin{aligned}d\big(B\circ(f_1\times f_2)\big)_x(h)&=dB_{\left(f_1(x),f_2(x)\right)}\big((df_1)_x(h),(df_2)_x(h)\big) &(10.875a)\\ &=B\big(f_1(x),(df_2)_x(h)\big)+B\big((df_1)_x(h),f_2(x)\big), &(10.875b)\end{aligned}$$

comme souhaité. □

122. Proposition 10.275.

10.13.19 Différentielle partielle

VJM_CtSKT

Définition 10.297 (Différentielle partielle)**.**
*Soient E, F et G des espaces vectoriels normés et une fonction $f\colon E\times F\to G$. Nous définissons sa **différentielle partielle** sur l'espace E par*

$$\begin{aligned} d_1f_{(x_0,y_0)}\colon E&\to G\\ u&\mapsto \frac{d}{dt}\Big[f(x_0+tu,y_0)\Big]_{t=0}. \end{aligned} \tag{10.876}$$

La différentielle d_2 se définit de la même façon.

PropLDN_nHWDF

Proposition 10.298 ([355])**.**
Soient E_1, E_2 et F des espaces vectoriels normés, soit un ouvert $U\subset E_1\times E_2$ et une fonction $f\colon U\to F$.

ItemRDD_oPmXVi

(1) Si f est différentiable alors les différentielles partielles existent et

$$d_1f_{(x_0,y_0)}(u)=df_{(x_0,y_0)}(u,0) \tag{10.877a}$$
$$d_2f_{(x_0,y_0)}(v)=df_{(x_0,y_0)}(0,v) \tag{10.877b}$$

où $u\in E_1$ et $v\in E_2$.

(2) Si f est différentiable alors

$$df_{(x_0,y_0)}(u,v)=d_1f_{(x_0,y_0)}(u)+d_2f_{(x_0,y_0)}(v). \tag{10.878}$$

Démonstration. Nous posons $\alpha=(x_0,y_0)\in U$ et

$$\begin{aligned} j_\alpha^{(1)}\colon E_1&\to E_1\times E_2\\ x&\mapsto (x,y_0). \end{aligned} \tag{10.879}$$

C'est une fonction de classe C^∞ et

$$(dj_\alpha^{(1)})_{x_0}(u)=\frac{d}{dt}\Big[j_\alpha^{(1)}(x_0+tu)\Big]_{t=0}=\frac{d}{dt}\Big[(x_0+tu,y_0)\Big]_{t=0}=(u,0). \tag{10.880}$$

D'autre part

$$(d_1f)_\alpha(u)=\frac{d}{dt}\Big[f(x_0+tu,y_0)\Big]_{t=0} \tag{10.881a}$$
$$=\frac{d}{dt}\Big[(f\circ j_\alpha^{(1)})(x_0+tu)\Big]_{t=0} \tag{10.881b}$$
$$=\big(d(f\circ j_\alpha^{(1)})\big)_{x_0}(u). \tag{10.881c}$$

À ce moment nous utilisons la règle des différentielles composées 10.277 pour dire que

$$(d_1f)_\alpha(u)=df_{j_\alpha^{(1)}(x_0)}\circ(dj_\alpha^{(1)})_{x_0}(u)=df_\alpha(u,0). \tag{10.882}$$

Voilà qui prouve déjà le point (1).

Pour la suite nous considérons les fonctions

$$\begin{aligned} P_1(x,y)&=x, & J_1(u)&=(u,0),\\ P_2(x,y)&=y, & J_2(v)&=(0,v) \end{aligned} \tag{10.883}$$

et nous avons l'égalité évidente

$$J_1\circ P_1+J_2\circ P_2=\mathbb{1} \tag{10.884}$$

sur $E_1\times E_2$. En appliquant df_α à cette dernière égalité, en appliquant à (u,v) et en utilisant la linéarité de df_α nous trouvons

$$df_\alpha(u,v)=df_\alpha\big((J_1\circ P_1)(u,v)\big)+df_\alpha\big((J_2\circ P_2)(u,v)\big) \tag{10.885a}$$
$$=df_\alpha(u,0)+df_\alpha(0,v) \tag{10.885b}$$
$$=(d_1f)_\alpha(u)+(d_2f)_\alpha(v) \tag{10.885c}$$

où nous avons utilisé le point (1) pour la dernière égalité. □

10.13.20 L'inverse, sa différentielle

Si E est un espace de Banach, nous sommes intéressés à l'espace $\mathrm{GL}(E)$ des endomorphismes inversibles de E sur E. Cet ensemble est métrique par la formule usuelle

$$\|T\| = \sup_{\|x\|=1} \|T(x)\|_E. \tag{10.886}$$

PropQAjqUNp

Proposition 10.299 (Thème 41)**.**
Soit E un espace de Banach (espace vectoriel normé complet). Si A est un endomorphisme de E satisfaisant $\|A\| < 1$ pour la norme opérateur, alors $(\mathbb{1} - A)$ est inversible et son inverse est donné par

$$(\mathbb{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{10.887}$$

Démonstration. Étant donné que la norme opérateur est une norme algébrique (lemme 10.69), nous avons $\|A^k\| \leqslant \|A\|^k$. Par conséquent la série $\|A^k\|$ est majorée par la série géométrique qui converge [123]. Par conséquent $\sum_k A^k$ est une série absolument convergente et donc convergente par la proposition 10.96 et le fait que $\mathcal{L}(E)$ est complet (proposition 10.99).

Montrons à présent que la somme est l'inverse de $\mathbb{1} - A$ en utilisant le produit terme à terme autorisé par la proposition 10.106 :

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n} (A^k - A^{k+1}) = \mathbb{1} - A^{n+1}. \tag{10.888}$$

Par conséquent

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \leqslant \|A\|^{n+1} \to 0. \tag{10.889}$$

□

ThoCINVBTJ

Théorème 10.300 (Inverse dans $\mathrm{GL}(E)$[356, 355])**.**
Soient E et F des espaces vectoriels normés.

(1) L'ensemble $\mathrm{GL}(E)$ est ouvert dans $\mathrm{End}(E)$.

(2) L'application inverse

$$\begin{aligned} i\colon \mathrm{GL}(E,F) &\to \mathrm{GL}(F,E) \\ u &\mapsto u^{-1} \end{aligned} \tag{10.890}$$

est de classe C^∞ et

$$di_{u_0}(h) = -u_0^{-1} \circ h \circ u_0^{-1} \tag{10.891}$$

pour tout $h \in \mathrm{End}(E)$

Démonstration. Nous supposons que $\mathrm{GL}(E,F)$ n'est pas vide, sinon ce n'est pas du jeu.

(1) **Cas de dimension finie**
Si la dimension de E et F est finie, elles doivent être égales, sinon il n'y a pas de fonctions inversibles $E \to F$. L'ensemble $\mathrm{GL}(E,F)$ est donc naturellement $\mathrm{GL}(n,\mathbb{R})$. Un élément de $\mathbb{M}(n,\mathbb{R})$ est dans $\mathrm{GL}(n,\mathbb{R})$ si et seulement si son déterminant est non nul. Le déterminant étant une fonction continue (polynomiale) en les entrées de la matrice, l'ensemble $\mathrm{GL}(n,\mathbb{R})$ est ouvert dans $\mathbb{M}(n,\mathbb{R})$.
Même idée pour la régularité de la fonction $i\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$, $X \mapsto X^{-1}$. Les entrées de X^{-1} sont les cofacteurs de X divisé par $\det(X)$, et donc des polynômes en les entrées de X divisés par un polynôme qui ne s'annule pas sur $\mathrm{GL}(n,\mathbb{R})$, et donc sur un ouvert autour de X et de X^{-1}. Bref, tout est C^∞.
Le reste de la preuve parle de la dimension infinie.

123. Proposition 10.135.

(2) **Ouvert autour de l'identité**
Nous commençons par prouver que $B(\mathbb{1},1)\subset \mathrm{GL}(E)$. Pour cela il suffit de remarquer que si $\|u\|<1$ alors le lemme 10.299 nous donne un inverse de $(1+u)$ en la personne de $\sum_{k=0}^{\infty}(-u)^k$.

(3) **Ouvert en général**
Soit maintenant $u_0\in \mathrm{GL}(E)$. Si $\|u\|<\frac{1}{\|u_0^{-1}\|}$ alors $\|u_0^{-1}u\|<1$, ce qui signifie que

$$\mathbb{1}+u_0^{-1}u \tag{10.892}$$

est inversible. Mais $u_0+u=u_0(\mathbb{1}+u_0^{-1}u)$, donc $u_0+u\in\mathrm{GL}(E)$ ce qui signifie que

$$B\left(u_0,\frac{1}{\|u_0^{-1}\|}\right)\subset \mathrm{GL}(E). \tag{10.893}$$

(4) **Différentielle en l'identité**
Nous commençons par prouver que $di_{\mathbb{1}}(u)=-u$. Pour cela nous posons

$$\alpha(h)=\sum_{k=2}^{\infty}(-1)^k h^k \tag{10.894}$$

et nous calculons

$$di_{\mathbb{1}}(u)=\frac{d}{dt}\Big[i(\mathbb{1}+tu)\Big]_{t=0}=\frac{d}{dt}\Big[\mathbb{1}-tu+\alpha(tu)\Big]_{t=0}. \tag{10.895}$$

Il suffit de prouver que $\frac{d}{dt}\Big[\alpha(tu)\Big]_{t=0}=0$ pour conclure que $di_{\mathbb{1}}(u)=-u$. Pour cela, nous remarquons que $\alpha(0)=0$ et donc que

$$\frac{d}{dt}\Big[\alpha(tu)\Big]_{t=0}=\lim_{t\to0}\frac{\alpha(tu)-\alpha(0)}{t} \tag{10.896a}$$

$$=\lim_{t\to0}\sum_{k=2}^{\infty}(-1)^k\frac{(tu)^k}{t} \tag{10.896b}$$

$$=-\lim_{t\to0}u\sum_{k=1}^{\infty}(-1)^k t^k u^k. \tag{10.896c}$$

La norme de ce qui est dans la limite est majorée par

$$\|u\|\sum_{k=1}^{\infty}\|tu\|^k=\|u\|\left(\frac{1}{1-\|tu\|}-1\right), \tag{10.897}$$

et cela tend vers zéro lorsque $t\to\infty$. Nous avons utilisé la somme 10.344 de la série géométrique. Nous avons bien prouvé que $di_{\mathbb{1}}(u)=-u$.

(5) **Différentielle en général** Soit maintenant $u_0\in\mathrm{GL}(E)$ et $h\in\mathrm{End}(E)$ tel que $u_0+h\in\mathrm{GL}(E)$; par le premier point, il suffit de prendre $\|h\|$ suffisamment petit. Vu que $u_0+h=u_0(\mathbb{1}+u_0^{-1}h)$ nous avons

$$(u_0+h)^{-1}=(\mathbb{1}+u_0^{-1}h)^{-1}u_0^{-1}. \tag{10.898}$$

Nous pouvons donc calculer

$$(u_0+h)^{-1}=\big(\mathbb{1}-u_0^{-1}h+\alpha(u_0^{-1}h)\big)u_0^{-1}=u_0^{-1}-u_0^{-1}hu_0^{-1}+\alpha(u_0^{-1}h)u_0^{-1}, \tag{10.899}$$

et ensuite

$$di_{u_0}(h)=\frac{d}{dt}\Big[i(u_0+th)\Big]_{t=0}=\frac{d}{dt}\Big[u_0^{-1}-tu_0^{-1}hu_0^{-1}+\alpha(tu_0^{-1}h)u_0^{-1}\Big]_{t=0}, \tag{10.900}$$

mais nous avons déjà vu que

$$\frac{d}{dt}\Big[\alpha(th)\Big]_{t=0}=0, \tag{10.901}$$

donc

$$di_{u_0}(h)=-u_0^{-1}hu_0^{-1} \tag{10.902}$$

Cela donne la différentielle de l'application inverse.

(6) **Continuité de l'inverse**
L'application i est continue parce que différentiable.

(7) **L'inverse est C^∞**
Nous allons écrire la fonction inverse comme une composée. Soient les applications

$$\begin{aligned} B\colon \mathcal{L}^0(F,E)\times\mathcal{L}^0(F,E) &\to \mathcal{L}^0\big(\mathcal{L}^0(E,F),\mathcal{L}^0(F,E)\big)\\ B(\psi_1,\psi_2)(A) &= -\psi_1\circ A\circ\psi_2 \end{aligned} \tag{10.903}$$

et

$$\begin{aligned} \Delta\colon \mathcal{L}^0(F,E) &\to \mathcal{L}^0(F,E)\times\mathcal{L}^0(F,E)\\ \varphi &\mapsto (\varphi,\varphi) \end{aligned} \tag{10.904}$$

Nous avons alors

$$di = B\circ\Delta\circ i. \tag{10.905}$$

L'application Δ est de classe C^∞. Nous devons voir que B l'est aussi. Pour le voir nous commençons par prouver qu'elle est bornée :

$$\begin{aligned} \|B\| &= \sup_{\|\psi_1\|,\|\psi_2\|=1} \|B(\psi_1,\psi_2)\|_{\mathcal{L}\left(L(E,F),L(F,E)\right)}\\ &= \sup_{\|\psi_1\|,\|\psi_2\|=1}\sup_{\|A\|=1}\|\psi_1\circ A\circ\psi_2\|_{L(F,E)}\\ &\leqslant \sup_{\|\psi_1\|,\|\psi_2\|=1}\sup_{\|A\|=1}\|\psi_1\|\|A\|\|\psi_2\|\\ &\leqslant 1. \end{aligned} \tag{10.906}$$

Donc B est bien bornée et par conséquent continue. Une application bilinéaire continue est C^∞ par le lemme 10.295. La décomposition $di = B\circ\Delta\circ i$ nous donne donc que $i\in C^\infty$ dès que i est continue, ce que nous avions déjà montré.

□

10.14 Exponentielle de matrice

PropPEDScAOmSxQM

Proposition 10.301.
Soit V un espace vectoriel de dimension finie et $A\in\mathrm{End}(V)$. La série

$$\exp(A) = \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{k=0}^{\infty}\frac{A^k}{k!}. \tag{10.907}$$

converge normalement[124] *dans $\big(\mathrm{End}(V),\|.\|_{op}\big)$. L'**exponentielle** de la matrice A est cette matrice.*

Démonstration. Vu que la norme opérateur est une norme d'algèbre par le lemme 10.69, nous avons pour tout k la majoration $\|A^k\|\leqslant\|A\|^k$. Nous avons donc

$$\sum_{k=0}^{\infty}\frac{\|A^k\|}{k!}\leqslant\sum_{k}\frac{\|A\|^k}{k!}. \tag{10.908}$$

La dernière somme converge en vertu de la convergence de la série exponentielle donnée en le lemme 10.137. □

Étant donné que c'est une limite, il y a une question de convergence et donc de topologie. C'est pour cela que nous ne pouvions pas introduire l'exponentielle de matrice avant d'avoir introduit la norme des matrices. La convergence de la série pour toute matrice sera prouvée au passage dans la proposition 10.302.

La fonction exponentielle $x\mapsto e^x$ n'est pas un polynôme en x, mais nous avons le résultat marrant suivant.

124. Convergence normale, définition 10.93.

PropFMqsIE

Proposition 10.302.
Si u est un endomorphisme, alors $\exp(u)$ *est un polynôme en* u [125].

Démonstration. Nous considérons l'application

$$\begin{aligned}\varphi_u\colon \mathbb{K}[X] &\to \mathrm{End}(E)\\ P &\mapsto P(u)\end{aligned} \tag{10.909}$$

Étant donné que l'image de φ_u est un fermé dans $\mathrm{End}(E)$, il suffit de montrer que la série

$$\sum_{k=0}^{\infty} \frac{\varphi_u(X)^k}{k!} \tag{10.910}$$

converge dans $\mathrm{End}(E)$ pour qu'elle converge dans $\mathrm{Image}(\varphi_u)$. Pour ce faire nous nous rappelons de la norme opérateur [126] et de la propriété fondamentale $\|A^k\| \leqslant \|A\|^k$. En notant $A = \varphi_u(X)$,

$$\left\| \sum_{k=n}^{m} \frac{A^k}{k!} \right\| \leqslant \sum_{k=n}^{m} \frac{\|A^k\|}{k!} \leqslant \sum_{k=n}^{m} \frac{\|A\|^k}{k!}, \tag{10.911}$$

ce qui est une morceau du développement de $e^{\|A\|}$. La limite $n \to \infty$ est donc zéro par la convergence de l'exponentielle réelle. La suite des sommes partielles de e^A est donc de Cauchy. La série converge donc parce que nous sommes dans un espace vectoriel réel de dimension finie ($\mathrm{End}(E)$). □

10.303.
Pourquoi $\exp(u)$ est-il un polynôme d'endomorphisme alors que exp n'est pas un polynôme ? Lorsque nous disons que la fonction $x \mapsto \exp(x)$ n'est pas un polynôme, nous sommes en train de localiser la fonction exp à l'intérieur de l'espace de toutes les fonctions $\mathbb{R} \to \mathbb{R}$, c'est-à-dire à l'intérieur d'un espace de dimension infinie. Au contraire lorsqu'on parle de $\exp(u)$ et qu'on le compare aux endomorphismes $P(u)$, nous sommes en train de repérer $\exp(u)$ à l'intérieur de l'espace des matrices qui est de dimension finie. Il n'est donc pas étonnant que l'on parvienne moins à faire la distinction.

Si par contre nous considérons exp en tant qu'application $\exp\colon \mathrm{End}(E) \to \mathrm{End}(E)$, ce n'est pas un polynôme.

Si u et v sont des endomorphismes, nous aurons des polynômes P et Q tels que $e^u = P(u)$ et $e^v = Q(v)$; mais nous n'aurons en général évidemment pas $P = Q$. En cela, exp n'est pas un polynôme.

10.15 Espace dual

SECooKOJNooQVawFY

Définition 10.304.
Soit un espace vectoriel normé $(V, \|.\|)$ sur le corps $\mathbb{C}$ ou $\mathbb{R}$ (que nous nommons $\mathbb{K}$). Son ***dual topologique****, noté V' est l'ensemble des applications linéaires continues $V \to \mathbb{K}$.*

10.15.1 Topologies

Il est possible de mettre sur V' (au moins) deux topologies distinctes. La première est la topologie de la norme opérateur ; rien de nouveau pour elle. La seconde est la topologie $*$-faible dont nous avons déjà un peu parlé dans la définition 7.358.

En termes de notations, nous allons noter les seminormes de la topologie faible par

$$p_x(\varphi) = |\varphi(x)| \tag{10.912}$$

pour $x \in V$ et $\varphi \in V'$. À droite, les barres dénotent soit la valeur absolue (si $\mathbb{K} = \mathbb{R}$), soit le module (si $\mathbb{K} = \mathbb{C}$).

125. Nan, mais j'te jure : exp n'est pas un polynôme, mais $\exp(u)$ est un polynôme de u.
126. Définition 10.55.

LEMooFMAUooQBIeTh

Lemme 10.305.
Soit $\varphi \in V'$ et $x \in V$. Alors

$$p_x(\varphi) \leqslant \|x\|\|\varphi\|. \tag{10.913}$$

Si $\varphi_0 \in V'$, si $r > 0$ et si $x \in V$ nous avons aussi :

$$B(\varphi_0, r) \subset B_x(\varphi_0, \|x\|r). \tag{10.914}$$

Démonstration. En posant $x' = x/\|x\|$, le vecteur x' est de norme 1 et nous avons

$$p_x(\varphi) = |\varphi(x)| = \|x\||\varphi(x')| \leqslant \|x\|\|\varphi\|. \tag{10.915}$$

En ce qui concerne la seconde affirmation, si $\varphi \in B(\varphi_0, r)$ alors en notant $x' = x/\|x\|$ nous avons :

$$p_x(\varphi_0 - \varphi) = |\varphi_0(x) - \varphi(x)| = |(\varphi_0 - \varphi)(x)| \leqslant \|x\|\|\varphi_0 - \varphi\| \leqslant r\|x\|. \tag{10.916}$$

Donc $\varphi \in B_x\big(\varphi_0, r\|x\|\big)$. □

PROPooMINIooGCWsPR

Proposition 10.306.
Si une suite(φ_k) dans V' vérifie

$$\varphi_k \xrightarrow{\|.\|} \varphi \tag{10.917}$$

alors elle vérifie aussi

$$\varphi_k \xrightarrow{*} \varphi. \tag{10.918}$$

Démonstration. Soit une suite (φ_k) dans V', convergente vers φ pour la topologie de la norme. Soit $x \in V$, et $x' = x/\|x\|$. Nous avons

$$p_x(\varphi_k - \varphi) = \|x\||\varphi_k(x') - \varphi(x)| \leqslant \|x\|\|\varphi_k - \varphi\| \to 0. \tag{10.919}$$

□

LEMooEAVEooAFveHn

Lemme 10.307.
La translation dans V' est une opération continue pour la topologie de la norme opérateur et pour celle de la topologie $*$.

Démonstration. Soit une suite φ_k tendant vers 0 ; nous devons prouver que $\tau_\sigma(\varphi_k) \to \tau_\sigma(0) = \sigma$. Et ce, pour chacune des deux topologies.

(1) **Norme opérateur**

L'hypothèse $\varphi_k \xrightarrow{\|.\|} 0$ signifie que $\|\varphi_k\| \to 0$, c'est-à-dire que

$$\sup_{\|v\|=1} |\varphi_k(v)| \to 0. \tag{10.920}$$

Nous avons alors

$$\|\tau_\sigma(\varphi_k) - \sigma\| = \sup_{\|v\|=1} |\tau_\sigma(\varphi_k)v - \sigma(v)| = \sup_{\|v\|=1} |\varphi_k(v)| \to 0. \tag{10.921}$$

Donc d'accord pour $\tau_\sigma(\varphi) \to \sigma$.

(2) **Topologie $*$**

Nous supposons maintenant que $\varphi_k \xrightarrow{*} 0$. Pour tout $v \in V$ nous avons

$$p_v\big(\tau_\sigma(\varphi_k) - \sigma\big) = \big|\tau_\sigma(\varphi_k)v - \sigma(v)\big| = |\varphi_k(v)| = p_v(\varphi_k). \tag{10.922}$$

Mais par hypothèse, $p_v(\varphi_k) \to 0$.

□

Pour la suite, nous allons préfixer par N les concepts liés à la topologie de V' associée à la norme opérateur et par $*$, les concepts de la topologie $*$.

PROPooFGXAooFRWweD

Proposition 10.308.
Soit un espace vectoriel normé V. Un $*$-ouvert est toujours un N-ouvert.

Démonstration. Soit un $*$-ouvert $\mathcal{O}$ de V'. Il existe donc $x \in V$ et $r > 0$ tels que $B_x(\varphi, r) \subset \mathcal{O}$. Nous avons alors, en utilisant le lemme 10.305,

$$B(\varphi, r\|x\|) \subset B_x(\varphi, r) \subset \mathcal{O}. \tag{10.923}$$

Donc $\mathcal{O}$ est un N-ouvert. □

Corolaire 10.309.
Soit un espace topologique X. Si $f\colon (V', *) \to X$ est continue, alors $f\colon (V', \|.\|) \to X$ est continue.

Démonstration. Soit un ouvert $\mathcal{O}$ de X. Vu que f est $*$-continue, la partie $f^{-1}(\mathcal{O})$ est un $*$-ouvert de V'. Il est onc un N-ouvert de V' par la proposition 10.308. □

10.15.2 Module de continuité

DEFooYARJooYyzMMP

Définition 10.310.
*Soient deux espaces topologiques normés X et Y, ainsi qu'une application $f\colon X \to Y$. Le **module de continuité** de f est la fonction*

$$\begin{aligned} \omega_f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \cup \{\infty\} \\ h &\mapsto \sup_{\substack{x,y\in X \\ d_X(x,y)<h}} d_Y\big(f(x), f(y)\big). \end{aligned} \tag{10.924}$$

Écrite de façon plus compacte,

$$\omega_f(h) = \sup_{|x-y|<h} \|f(x) - f(y)\|. \tag{10.925}$$

EQooKWUVooSORHXN

Nous définissons aussi $\omega_f(h) = 0$ pour $h \leqslant 0$ parce que le lemme 10.312 fera grand cas de la continuité en zéro du module de continuité.

Notons que le module de continuité est une fonction croissante.

LemLUbgYeo

Lemme 10.311.
Soit $f \in C^0\big([0,1], \mathbb{C}\big)$ et ω son module de continuité. Si λ et h sont strictement positifs avec $\lambda h \in [0,1]$ alors

$$\omega(\lambda h) \leqslant (\lambda + 1)\omega(h). \tag{10.926}$$

Démonstration. La fonction ω est croissante, et pour $h, k > 0$ nous avons $\omega(h+k) \leqslant \omega(h) + \omega(k)$. Par récurrence pour tout $k \in \mathbb{N}$ nous avons

$$\omega(kh) \leqslant k\omega(h). \tag{10.927}$$

En écrivant cela pour $k = \lceil \lambda \rceil$, nous avons

$$\omega(\lambda h) \leqslant \omega(kh) \leqslant k\omega(h) \leqslant (\lambda + 1)\omega(h). \tag{10.928}$$

□

LemeERapq

Lemme 10.312.
Une fonction $f\colon X \to Y$ est uniformément continue[127] *si et seulement si son module de continuité vérifie*

$$\lim_{h\to 0} \omega_f(h) = 0. \tag{10.929}$$

Autrement dit, si et seulement si sont module de continuité est continu en zéro.[128].

Démonstration. Nous commençons par supposer que f est uniformément continue. Soit $\epsilon > 0$. Par uniforme continuité, il existe $\delta > 0$ tel que $d\big(f(x), f(y)\big) \leqslant \epsilon$ dès que $d(x,y) \leqslant \delta$. Si $h \in B(0,\delta)$, alors

$$\omega_f(h) \leqslant \omega_f(\delta) = \sup_{\substack{x,y\in X\\ d(x,y)\leqslant\delta}} d\big(f(x), f(y)\big) \leqslant \epsilon. \tag{10.930}$$

Cela prouve que $\lim_{h\to 0} \omega_f(h) = 0$.

Dans l'autre sens, si $\epsilon > 0$ est fixé, il suffit de prendre δ tel que $\omega_f(h) \leqslant \epsilon$ pour tout $h \leqslant \delta$ pour faire fonctionner la définition de l'uniforme continuité. □

LEMooKPPSooPIncvn

Lemme 10.313 ([357])**.**
Soient des espaces métriques E et E' et une suite de fonctions $(f_i)_{i\geqslant 0}$ qui converge uniformément vers f. Alors pour chaque $\delta > 0$ nous avons

$$\limsup_{i\to\infty} \omega_{f_i}(\delta) \leqslant \omega_f(\delta). \tag{10.931}$$

Démonstration. Soient $\delta > 0$ ainsi que $x, y \in E$ tels que $d(x,y) \leqslant \delta$. Pour chaque i nous avons

$$\begin{aligned} |f_i(x) - f_i(y)| &\leqslant |f_i(x) - f(x)| + |f(x) - f(y)| + |f(y) - f_i(y)| && \text{(10.932a)}\\ &\leqslant |f(x) - f(y)| + 2\|f_i - f\|_\infty && \text{(10.932b)}\\ &\leqslant \omega_f(\delta) + 2\|f_i - f\|_\infty. && \text{(10.932c)} \end{aligned}$$

Nous prenons le supremum de cela sur $\{x, y \in E \text{ tel que } \|x - y\| \leqslant \delta\}$ pour obtenir :

$$\omega_{f_i}(\delta) \leqslant \omega_f(\delta) + 2\|f_i - f\|_\infty. \tag{10.933}$$

La tentation est grande à ce point de prendre la limite des deux côtés pour $i \to \infty$. Cependant, rien ne nous permet de dire que la suite $i \mapsto \omega_{f_i}(\delta)$ ait une limite. Nous pouvons cependant prendre la limite supérieure[129] et obtenir

$$\limsup_{i\to\infty} \omega_{f_i}(\delta) \leqslant \omega_f(\delta). \tag{10.934}$$

□

10.16 Mini introduction aux nombres p-adiques

10.16.1 La flèche d'Achille

s:un

C'est un grand classique que je donne ici juste comme introduction pour montrer que des séries infinies peuvent donner des nombres finis de manière tout à fait intuitive.

Achille tire une flèche vers un arbre situé à 10 m de lui. Disons que la flèche avance à une vitesse constante de 1 m/s. Il est clair que la flèche mettra 10 s pour toucher l'arbre. En 5 s, elle aura parcouru la moitié de son chemin. On le note :

$$\text{temps} = 5s + \dots$$

127. Définition 7.323.
128. Dans ce lemme, nous avons deux espaces métriques, mais nous allons noter d la distance des deux côtés.
129. Définition 9.42.

Reste 5 m à faire. En 2.5 s, elle aura fait la moitié de ce chemin, soit $2.5m = \frac{10}{4}m$. On le note :

$$\text{temps} = \frac{10}{2}s + \frac{10}{4}s+$$

Reste $2.5m$ à faire. La moitié de ce trajet, soit $\frac{10}{8}m$, est parcouru en $\frac{10}{8}s$; on le note encore, mais c'est la dernière fois !

$$\text{temps} = \frac{10}{2}s + \frac{10}{4}s + \frac{10}{8}s+$$

En continuant ainsi à regarder la flèche qui parcourt des demi-trajets puis des moitiés de demi-trajets et encore des moitiés de moitiés de demi-trajets, et en sachant que le temps total est $10s$, on trouve :

$$10\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots\right) = 10.$$

On doit donc croire que la somme jusqu'à l'infini des inverses des puissances de deux vaut 1 :

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Cela peut être démontré à la loyale.

10.16.2 La tortue et Achille

Maintenant qu'on est convaincu que des sommes infinies peuvent représenter des nombres tout à fait normaux, passons à un truc plus marrant.

Achille, qui marche peinard à 10 m/h, part avec $1m$ d'avance sur une tortue qui avance à 1 m/h. Le temps que la tortue arrive au point de départ d'Achille, Achille aura parcouru $10m$, et le temps que la tortue mettra pour arriver à ce point, eh bien, Achille ne sera déjà plus là : il sera à $100m$. Si la tortue tient bon pendant un temps infini, et si l'on est confiant en le genre de raisonnements faits à la section 10.16.1, elle rattrapera Achille dans

$$1m + 10m + 100m + 1000m + \dots$$

Autant dire que ça ne risque pas d'arriver. Et pourtant, mettons en équations :

$$\begin{cases} x_{\text{Achile}}(t) = 1 + 10t & \text{(10.935a)} \\ x_{\text{tortue}}(t) = t. & \text{(10.935b)} \end{cases}$$

La tortue rejoint Achille au temps t tel que $x_{\text{Achille}(t)} = x_{\text{tortue}}(t)$. Un mini calcul donne $t = -1/9$. Physiquement, c'est une situation logique. Peut-on en déduire une égalité mathématique du style de

$$1 + 10 + 100 + 1000 + \dots = -\frac{1}{9} \text{ ???}$$

Là où les choses deviennent jolies, c'est quand on cherche à voir ce que peut bien être la valeur d'un hypothétique $x = 1 + 10 + 100 + 1000 + \dots$. En effet, logiquement on devrait avoir

$$\begin{aligned} \frac{x}{10} &= \frac{1}{10} + 1 + 10 + 100 + \dots \\ &= \frac{1}{10} + x. \end{aligned}$$

Reste à résoudre l'équation du premier degré : $\frac{x}{10} = x + \frac{1}{10}$. Ai-je besoin de donner la solution ?

10.16.3 Dans les nombres p-adiques, c'est vrai

Nous nous proposons d'apprendre sur les nombres p-adiques juste ce qu'il faut pour montrer que l'égalité

$$\sum_{k=0}^{\infty} 10^k = -\frac{1}{9} \tag{10.936}$$

est vraie dans les nombres 5-adiques. Tout ce qu'il faut est sur wikipedia.

Soit $a \in \mathbb{N}$ et p, un nombre premier. La **valuation** p-adique de a est l'exposant de p dans la décomposition de a en nombres premiers. On la note $v_p(a)$. Pour un rationnel on définit

$$v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b) \tag{10.937}$$

La **valeur absolue** p-adique de $r \in \mathbb{Q}$ est

$$|r|_p = p^{-v_p(r)}. \tag{10.938}$$

Nous posons $|0|_p = 0$. De là nous considérons la distance

$$d_p(x, y) = |x - y|_p. \tag{10.939}$$

LEMooQNHZooXlHdUc

Lemme 10.314.
L'espace $(\mathbb{Q}, d_p)$ est un espace métrique[130].

Nous considérons maintenant $p = 5$. Étant donné que $a = 5 \cdot 2$ nous avons $v_5(10) = 1$ et

$$v_5\left(\frac{1}{9}\right) = v_5(1) - v_5(9) = 0. \tag{10.940}$$

Nous avons

$$\sum_{k=0}^{N} 10^k + \frac{1}{9} = \frac{10^{N+1}}{9} \tag{10.941}$$

mais

$$v_p\left(\frac{10^{N+1}}{9}\right) = v_5(10^{N+1}) - v_5(9) = N + 1. \tag{10.942}$$

Par conséquent

$$d_5\Big(\sum_{k=0}^{N} 10^k, -\frac{1}{9}\Big) = \Big|\frac{10^{N+1}}{9}\Big|_p = p^{-(N+1)}. \tag{10.943}$$

En passant à la limite,

$$\lim_{N\to\infty} d_5\Big(\sum_{k=0}^{N} 10^k, -\frac{1}{9}\Big) = 0, \tag{10.944}$$

ce qui signifie que[131]

$$\sum_{k=0}^{\infty} 10^k = -\frac{1}{9}. \tag{10.945}$$

130. Définition 7.115
131. Voir la définition 10.89 de la convergence d'une série dans un espace métrique.

Chapitre 11

Espaces affines

11.1 Vecteurs agissant sur un espace

Définition 11.1.
Soit E, un espace vectoriel. Un ***espace affine modelé sur*** *E est un ensemble $\mathcal{E}$ sur lequel le groupe $(E,+)$ agit à droite transitivement et librement*[1].

Étant donné que E est un groupe commutatif, l'action peut être vue indifféremment à gauche ou à droite. Si $M \in \mathcal{E}$ et si $x \in E$ nous notons $M+x$ au lieu de $x \cdot M$ le résultat de l'action de x sur M.

NORMooZANAooQdXqlh

11.2.
Lorsque nous écrivons « $M+x$ », le symbole plus n'est pas une loi de composition interne de $\mathcal{E}$, mais une action.

DEFooWAYTooMLbqEE

Proposition-Définition 11.3.
Soient $N, M \in \mathcal{E}$. Il existe un unique $x \in E$ tel que $M+x=N$.
Nous noterons $\overrightarrow{MN}$ ce vecteur.

Démonstration. La transitivité de l'action assure l'existence et la liberté assure l'unicité. □

LEMooFZCRooQxzObv

Lemme 11.4.
Pour tout élément A nous avons $A+0=A$.

Démonstration. Soit $B \in \mathcal{E}$. Nous avons :

$$B+0=B+(0+0)=(B+0)+0 \tag{11.1}$$

parce que le $+$ dénote une action.
En appliquant cette égalité à l'élément $B=A-0$ nous trouvons l'égalité demandée. □

PROPooCOZCooCghwaR

Proposition 11.5.
Soit un espace affine $\mathcal{E}$ modelé sur l'espace vectoriel E. Soient $A, B, C \in \mathcal{E}$. Nous avons les égalités suivantes dans E :

ITEMooSDMIooUQiKeW

(1) $\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}$ *(relations de Chasles),*

ITEMooWZAVooGfGBwd

(2) $\overrightarrow{AA}=0$,

ITEMooLDVXooFZMbsQ

(3) $\overrightarrow{BA}=-\overrightarrow{AB}$.

Démonstration. Point par point.

1. Définition 2.45.

(1) **Pour (1)** Nous avons, par définition 11.3 les égalités

$$\begin{cases} C = A + \overrightarrow{AC} & \text{(11.2a)} \\ B = A + \overrightarrow{AB} & \text{(11.2b)} \\ C = B + \overrightarrow{BC} & \text{(11.2c)} \end{cases}$$

En substituant les deux premières dans la troisième, nous trouvons $A + \overrightarrow{AB} + \overrightarrow{BC} = A + \overrightarrow{AC}$. Par liberté de l'action, nous pouvons « simplifier » par A et trouver la relation de Chasles.

(2) **Pour (2)** Nous avons $A + \overrightarrow{AA} = A$, mais aussi $A + 0 = A$. Par unicité nous avons $\overrightarrow{AA} = 0$.

(3) **Pour (3)** Nous avons $B + \overrightarrow{BA} = A$ et $A + \overrightarrow{AB} = B$. En mettant bout à bout,

$$B + \overrightarrow{BA} + \overrightarrow{AB} = B. \tag{11.3}$$

Donc $\overrightarrow{BA} + \overrightarrow{AB} = 0$.

□

NORMooXAJLooIupekj

11.6.
Si E est un espace vectoriel, le groupe $(E, +)$ agit sur E par l'action $t_y(x) = y + x$. Utilisant cette action nous construisons l'**espace affine canonique** de E. En particulier nous notons $\mathcal{E}_n(\mathbb{K})$ l'espace affine canonique de $\mathbb{K}^n$ vu comme espace vectoriel sur $\mathbb{K}$.

— En tant qu'ensembles, $\mathcal{E}_n(\mathbb{K}) = \mathbb{K}^n$.

— Sur cet espace en particulier, si $M, N \in \mathcal{E}_n(\mathbb{K})$, nous avons $\overrightarrow{MN} = N - M$ où à droite, la différence est la différence vectorielle dans $\mathbb{K}^n$.

Ces deux points se généralisent immédiatement à un espace vectoriel E au lieu de $\mathbb{K}^n$.

11.2 Repères cartésiens affines

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{E}$ un espace affine construit sur E.

DEFooQELZooEXvxgw

Définition 11.7.
*Un multiplet $(A, e_1, \dots, e_n)$ où A est un point de $\mathcal{E}$ et $\{e_i\}$ est une base de E est un **repère cartésien** de $\mathcal{E}$.*

*Nous disons que $\{e_i\}$ est la **base associée** au repère.*

Proposition 11.8.
Si $\mathcal{E}$ est un espace affine modelé sur l'espace vectoriel E de dimension n sur le corps $\mathbb{K}$, et si $\big(A, \{e_i\}_{i=1,\dots,n}\big)$ est un repère cartésien, alors

$$\begin{aligned} \phi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1, \dots, x_n) &\mapsto A + \sum_i x_i e_i. \end{aligned} \tag{11.4}$$

est une bijection.

*Ces nombres x_i sont les **coordonnées** du point $A + \sum_i x_i e_i$ dans le repère (A, e_i).*

Démonstration. L'application ϕ est surjective parce que l'action de E sur $\mathcal{E}$ est transitive et injective parce que l'action est libre. □

11.3 Classification affine des coniques

Soit une conique $f(x, y) = 0$ avec

$$f(x, y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f \tag{11.5}$$

dans le repère $R = (A, e_i)$.

LEMooXZURooSVySRT

Lemme 11.9.
La signature de la quadratique

$$q(x,y) = ax^2 + 2bxy + cy^2 \tag{11.6}$$

ne dépend pas de la base choisie et un changement de variables

$$\begin{cases} \tilde{x} = \alpha x + \beta y & \text{(11.7a)} \\ \tilde{y} = \gamma x + \delta y & \text{(11.7b)} \end{cases}$$

peut nous amener dans trois cas :

$$q(x,y) = \begin{cases} \tilde{x}^2 + \tilde{y}^2 & \text{genre ellipse} \\ \tilde{x}^2 - \tilde{y}^2 & \text{genre hyperbole} \\ \tilde{x}^2 & \text{genre parabole.} \end{cases} \tag{11.8}$$

Dans le troisième cas, la matrice de q est de rang 1.

Nous cherchons maintenant à savoir si un point $I = (x_0, y_0)$ est un centre de symétrie de $f(x,y) = 0$. Pour cela nous choisissons le repère centré en I, c'est-à-dire que nous posons

$$\begin{cases} x = x_0 + \tilde{x} & \text{(11.9a)} \\ y = y_0 + \tilde{y}. & \text{(11.9b)} \end{cases}$$

Un peu de calcul montre qu'alors la conique s'écrit

$$f(x_0,y_0) + q(\tilde{x},\tilde{y}) + (2ax_0 + 2by_0 + 2d)\tilde{x} + (2bx_0 + 2cy_0 + 2e)\tilde{y} = 0. \tag{11.10}$$

LEMooMVIDooVEUJsp

Lemme 11.10.
Le point I sera un centre de symétrique si les termes linéaires en $\tilde{x}$ et $\tilde{y}$ s'annulent, c'est-à-dire si SyskhiOvW

$$\begin{cases} ax_0 + by_0 + d = 0 & \text{(11.11a)} \\ bx_0 + cy_0 + e = 0. & \text{(11.11b)} \end{cases}$$

Nous supposons que $(d,e) \neq (0,0)$, sinon la conique de départ serait déjà centrée. Le déterminant du système (11.11) est

$$\delta = ac - b^2. \tag{11.12}$$

Si ce dernier est différent de zéro, le système possède une unique solution et la conique aura alors un unique centre de symétrie.

Si le déterminant du système est nul, il y a soit aucun centre de symétrie, soit une infinité. Dans le premier cas nous sommes en présence d'une parabole, et dans le second cas de deux droites parallèles.

Exemple 11.11.
Soit

$$f(x,y) = x^2 + 2xy - y^2 - 6x + 2y - 1 = 0 \tag{11.13}$$ EqOgsEcz

donnée dans le repère affine $R = (A, \{e_i\})$. Nous commençons par étudier la signature de $q(x,y) = x^2 + 2xy - y^2$ dont la matrice symétrique est

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{11.14}$$

Son polynôme caractéristique est $\lambda^2 - 2$ dont les racines sont $\pm\sqrt{2}$. La signature est donc $(1,1)$ et nous sommes en présence d'une conique de genre hyperbole. Nous cherchons le centre en suivant le lemme 11.10. Nous posons $x = \tilde{x} + x_0$, $y = \tilde{y} + y_0$, et nous cherchons à résoudre le système

$$\begin{cases} x_0 + y_0 - 3 = 0 & \text{(11.15a)} \\ x_0 - y_0 + 1 = 0. & \text{(11.15b)} \end{cases}$$

L'unique solution est $(x_0, y_0) = (1, 2)$. Nous considérons le repère centré en (x_0, y_0), c'est-à-dire le repère

$$R' = (I, \{e_i\}) \tag{11.16}$$

avec $I = A + x_0 e_1 + y_0 e_2$ où A est l'origine du repère dans lequel l'équation (11.13) était donnée.

Par construction dans ce repère nous avons la conique

$$f(x_0, y_0) + q(\tilde{x}, \tilde{y}) = 0, \tag{11.17}$$

c'est-à-dire

$$\tilde{x}^2 + 2\tilde{x}\tilde{y} - \tilde{y}^2 - 2 = 0. \tag{11.18}$$

Maintenant, nous avons une quadrique centrée que nous voulons mettre sous une forme plus canonique :

$$\left(\frac{1}{\sqrt{2}}(\tilde{x} + \tilde{y})\right)^2 - \tilde{y}^2 - 1 = 0. \tag{11.19}$$

Nous posons donc

$$\begin{cases} X = \frac{1}{\sqrt{2}}(\tilde{x} + \tilde{y}) & \text{(11.20a)} \\ Y = \tilde{y}, & \text{(11.20b)} \end{cases}$$

pour trouver l'hyperbole

$$X^2 - Y^2 - 1 = 0. \tag{11.21}$$

Cherchons le changement de base correspondant. Pour trouver les coordonnées de e'_1 dans la base (e_1, e_2) nous cherchons pour quelles valeurs de x, y nous avons $e'_1 = xe_1 + ye_2$. Le point e'_1 étant caractérisé par $X = 1$, $Y = 0$ nous avons à résoudre

$$\begin{cases} \frac{1}{\sqrt{2}}(x + y) = 1 & \text{(11.22a)} \\ y = 0, & \text{(11.22b)} \end{cases}$$

ce qui donne $x = \sqrt{2}$ et $y = 0$. Donc

$$e'_1 = \sqrt{2}e_1. \tag{11.23}$$

Pour trouver e'_2, c'est le même raisonnement en posant $X = 0$ et $Y = 1$. Le résultat est :

$$e'_2 = -e_1 + e_2. \tag{11.24}$$

Résumons : EqfiVwym

$$\begin{cases} e'_1 = \sqrt{2}e_1 & \text{(11.25a)} \\ e'_2 = -e_1 + e_2. & \text{(11.25b)} \end{cases}$$

Il y a un dicton qui dit que les vecteurs de base se transforment avec la matrice inverse des coefficients. Prenons la matrice M donnée par

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = M \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \tag{11.26}$$

Calculons la matrice inverse.

```
sage: M=matrix([  [1/sqrt(2), 1/sqrt(2)],[0,1]  ])
sage: M
[1/2*sqrt(2) 1/2*sqrt(2)]
[           0            1]
sage: M.inverse()
[sqrt(2)      -1]
[       0       1]
```

tex/sage/sageSnip023.sage

Nous voyons que les colonnes de la matrice M^{-1} donnent les coordonnées des vecteurs e'_1 et e'_2.

△

11.4 Applications affines

Voici la définition d'une application affine entre deux espaces affines. La définition 8.160 donnera le définition d'une application affine entre espaces vectoriels.

DEFooUAWZooXcMKve

Définition 11.12.
*Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines sur les espaces vectoriels E et E' (sur le même corps $\mathbb{K}$). Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est dite **affine** si pour tout $M \in \mathcal{E}$, il existe une application linéaire[2] $u_M\colon E \to E'$ telle que*

$$f(M+x) = f(M) + u_M(x) \tag{11.27}$$

EqMqIoWX

pour tout $x \in E$.

La définition suivante permet de décomposer une application affine en une partie linéaire et une translation. À partir de là, la proposition 11.62 nous donnera une structure de groupe sur $\mathrm{Aff}(\mathbb{R}^n)$.

LEMooYJCDooOGAHkF

Lemme-Définition 11.13 (partie linéaire d'une application affine[1])**.**
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines sur les espaces vectoriels E et E' (sur le même corps $\mathbb{K}$). Nous considérons une application affine $f\colon \mathcal{E} \to \mathcal{E}'$.

Il existe une unique application linéaire $u\colon E \to E'$ telle que

$$f(M+x) = f(M) + u(x) \tag{11.28}$$

pour tout $x \in E$ et pour tout $M \in \mathcal{E}$.

*Cette application linéaire est appelée **partie linéaire** de f. Pour varier les notations, nous noterons souvent $f = \alpha \circ \tau_v$ pour une application linéaire α et la translation τ_v de vecteur v.*

Démonstration. En plusieurs étapes.

(1) **Unicité** Supposons que u_1 et u_2 vérifient la propriété, alors pour tout $x \in E$ et tout $M \in \mathcal{E}$ nous avons $f(M+x) = f(M) + u_1(x)$ et $f(M+x) = f(M) + u_2(x)$. Cela suffit à nous convaincre que $u_1 = u_2$.

(2) $u_M = u_N$ Avant de prouver l'existence, nous considérons $M, N \in \mathcal{E}$ et les applications linéaires u_M et u_N vérifiant l'équation (11.27) pour M et N respectivement. Prouvons que $u_M = u_N$.
Posons

$$f(M+x) = f(M) + u_M(x) \tag{11.29a}$$
$$f(N+y) = f(N) + u_N(y). \tag{11.29b}$$

Définissons $a \in E$ par $N = M + a$; nous avons d'une part

$$f(N+y) = f(M+y+a) = f(M) + u_M(y+a), \tag{11.30}$$

et d'autre part

$$f(N+y) = f(M+a) + u_N(y) = f(M) + u_M(a) + u_N(y). \tag{11.31}$$

Par conséquent $u_M(y+a) = u_M(a) + u_N(y)$. Par linéarité $u_N = u_M$.

(3) **Existence** Soit $M \in \mathcal{E}$. Nous affirmons que u_M fait l'affaire. En effet, soient $N \in \mathcal{E}$ et $x \in E$. Puisque $u_M = u_N$ nous avons

$$f(N+x) = f(N) + u_N(x) = f(N) + u_M(x). \tag{11.32}$$

Donc effectivement u_M peut être utilisé en tout point de $\mathcal{E}$.

□

Ce lemme est important car il permet de démontrer qu'une application est affine en prouvant la linéarité des u_M séparément sans devoir prouver qu'elles sont égales.

2. Définition 4.29.

11.4.1 Autres propriétés

LEMooXXTPooKYFGGM

Lemme 11.14 ([1]).
Soient $M \in \mathcal{E}$ et $A, B \in \mathcal{E}$ deux points donnés par $A = M + x_a$, $B = M + x_b$. Soit encore une application affine f sur $\mathcal{E}$. Alors

$$\overrightarrow{f(A)f(B)} = u_f(x_b - x_a). \tag{11.33}$$

Démonstration. En appliquant f à $A = M + x_a$ et $B = M + x_b$,

$$f(A) = f(M) + u_f(x_a) \tag{11.34a}$$
$$f(B) = f(M) + u_f(x_b). \tag{11.34b}$$

Donc $f(B) = f(A) - u_f(x_a) + u_f(x_b)$ ou encore

$$f(B) = f(A) + u_f(x_b - x_a). \tag{11.35}$$

□

Remarque 11.15.
La condition (11.27) pour tout $M \in \mathcal{E}$ est équivalente à demander

$$f \circ t_x = t_{u(x)} \circ f \tag{11.36}$$

pour tout $x \in E$.

PROPooALXYooHoMdqQ

Proposition 11.16 ([1]).
Soit une application affine $f : \mathcal{E} \to \mathcal{E}'$.

ITEMooSKCYooHyRZYN

(1) Il existe une unique application linéaire u_f telle que $f(M + x) = f(M) + u_f(x)$ pour tout $M \in \mathcal{E}$ et tout $x \in E$.

(2) L'application u_f est injective si et seulement si f est injective.

(3) L'application u_f est surjective si et seulement si f est surjective.

Démonstration. En plein de parties.

(1) **Pour (1)**
La partie (1) est le lemme 11.13.

(2) **Si u_f est injective** Soient $M, N \in \mathcal{E}$ tels que $f(M) = f(N)$. Nous avons

$$f(M) = f(N) = f\big(M + (N - M)\big) = f(M) + u_f(N - M), \tag{11.37}$$

donc $u_f(N - M) = 0$. Vu que u_f est injective, nous déduisons que $N - M = 0$.

(3) **Si f est injective** Soient $x, y \in E$ tels que $u_f(x) = u_f(y)$. Soit M quelconque dans $\mathcal{E}$; nous avons

$$f(M + x) = f(M) + u_f(x) = f(M) + u_f(y) = f(M + y). \tag{11.38}$$

L'injectivité de f nous indique alors que $M + x = M + y$ et donc que $x = y$ parce que l'action de E sur $\mathcal{E}$ est libre.

(4) **Si u_f est surjective** Soit $M \in \mathcal{E}$. Nous allons trouver un élément de $\mathcal{E}$ dont l'image par f est M. Soient $N \in \mathcal{E}$ et $x \in E$ tels que $u_f(x) = M - f(N)$.
Alors nous avons $f(N + x) = f(N) + u_f(x) = M$.

(5) **Si f est surjective** Soit $a \in E$. Nous voulons $x \in E$ tel que $u_f(x) = a$. Soit $M \in E$. Vu que f est surjective, il existe $N \in \mathcal{E}$ tel que $f(N) = f(M) + a$.
Posons $x = N - M$. Nous avons d'une part

$$f(M + x) = f(M) + u_f(x) \tag{11.39}$$

et d'autre part

$$f(M + x) = f\big(M + (N - M)\big) = f(N) = f(M) + a. \tag{11.40}$$

En égalisant nous trouvons $u_f(x) = a$.

□

Proposition 11.17.
Soient des espaces affines $\mathcal{E}$ et $\mathcal{E}'$ de même dimension. Une application affine $f\colon \mathcal{E} \to \mathcal{E}'$ est injective si et seulement si elle est surjective.

Démonstration. Nous allons utiliser les équivalences de la proposition 11.16, ainsi que le corolaire 4.49 pour la partie linéaire. Nous avons les équivalences :

$$f \text{ est injective} \Leftrightarrow u_f \text{ est injective} \Leftrightarrow u_f \text{ est surjective} \Leftrightarrow f \text{ est surjective.} \tag{11.41}$$

□

EXooAGINooYmvPML

Exemple 11.18.
L'espace $\mathbb{R}^n$ est très particulier parce qu'il agit sur lui-même ; il est donc un espace affine à lui tout seul : $\mathcal{E} = E = \mathbb{R}^n$.

Dans le cas de $\mathbb{R}^n$, en posant $M = 0$ dans la condition (11.27), si f est une application affine il existe une application linéaire α et un vecteur v tel que $f = \tau_v \circ \alpha$.

Notons que ça n'a pas de sens de poser $M = 0$, et la décomposition $f = \tau_v \circ \alpha$ n'a aucun sens en général. En particulier, nous ne pouvons pas appliquer une application linéaire à un élément d'un espace affine général. △

PROPooOUNEooQUZetW

Proposition 11.19.
Si $f\colon \mathcal{E} \to \mathcal{E}'$ et $g\colon \mathcal{E}' \to \mathcal{E}''$ sont des applications affines, alors $g \circ f\colon \mathcal{E} \to \mathcal{E}''$ est affine et $u_{g\circ f} = u_g \circ u_f$.

Démonstration. Si $M \in \mathcal{E}$ et $x \in E$ nous avons

$$\begin{aligned}(g \circ f)(M + x) &= g\big(f(M) + u_f(x)\big)\\ &= g\big(f(M)\big) + u_g\big(u_f(x)\big)\\ &= (g \circ f)(M) + (u_g \circ u_f)(x).\end{aligned} \tag{11.42}$$

□

THOooBAPDooEUtBgF

Théorème 11.20.
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines de dimensions finies p et q sur $\mathbb{K}$. Soient les repères cartésiens $R = (O, \{e_i\})$ et $R' = (O', \{e'_i\})$. Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est affine si et seulement si il existe une matrice $a \in \mathbb{M}_{p,q}(\mathbb{K})$ et $b \in \mathbb{K}^q$ tels que[3]

$$f(x) = b + ax. \tag{11.43}$$

EqCmNHis

11.5 Isomorphismes

Définition 11.21.
*Un **isomorphisme** entre les espaces affines $\mathcal{E}$ et $\mathcal{E}'$ est une application affine $f\colon \mathcal{E} \to \mathcal{E}'$ inversible dont l'inverse est affine.*

PropxtFeDE

Proposition 11.22.
Une application affine bijective est un isomorphisme. Si f est un isomorphisme d'espaces affines, alors $u_{f^{-1}} = (u_f)^{-1}$.

Proposition 11.23.
Un espace affine de dimension finie n sur un corps $\mathbb{K}$ est isomorphe à l'espace affine canonique $\mathcal{E}_n(\mathbb{K})$.

3. L'équation (11.43) est écrite en utilisant un abus de notation entre le vecteur $x \in \mathbb{K}^p$ et le point de $\mathcal{E}$ qui est représenté par x dans le repère $(A, \{e_i\})$.

Démonstration. Si nous considérons le repère $R = (A, \{e_i\})$ de l'espace affine $\mathcal{E}$ alors l'application

$$\begin{aligned} \varphi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1, \ldots, x_n) &\mapsto A + \sum_i x_i e_i \end{aligned} \tag{11.44}$$

est un isomorphisme. □

11.6 Sous espaces affines

DEFooJSUHooJjtpwz

Définition 11.24.
*Soit $\mathcal{E}$ un espace affine sur l'espace vectoriel E. Un **sous-espace affine** de $\mathcal{E}$ est une orbite de l'action d'un sous-espace vectoriel de E.*

Si $\mathcal{F}$ est un sous-ensemble de $\mathcal{E}$, il sera un sous-espace affine de $\mathcal{E}$ si et seulement si l'ensemble

$$F = \{AB \text{ tel que } A, B \in \mathcal{F}\} \tag{11.45}$$

est un sous-espace vectoriel de E. Dans ce cas nous disons que F est la **direction** de $\mathcal{F}$. Si $A \in \mathcal{F}$, alors l'orbite de A sous F est $\mathcal{F}$. La **dimension** de $\mathcal{F}$ est la dimension de sa direction.

Si $\mathcal{F}$ et $\mathcal{G}$ sont des sous-espaces affines de $\mathcal{E}$ de directions F et G, nous disons que $\mathcal{F}$ est **parallèle** à $\mathcal{G}$ si $F \subset G$.

Proposition 11.25.
Soit $\mathcal{F}$ un sous-espace affine de dimension k dans l'espace affine $\mathcal{E}$ de dimension n. Alors il existe une application affine $f\colon \mathcal{E} \to \mathbb{K}^{n-k}$ telle que $\mathcal{F} = f^{-1}(0)$.

Démonstration. Soient F la direction de $\mathcal{F}$ et $A \in \mathcal{F}$. Nous considérons une base $\{e_i\}$ adaptée à F au sens $\{e_1, \ldots, e_k\}$ est une base de F. Nous considérons maintenant le repère cartésien $(A, \{e_i\})$ et nous construisons l'application affine

$$\begin{aligned} f\colon \mathcal{E} &\to \mathbb{K}^{n-k} \\ A + \sum_{i=1}^{n} x_i e_i &\mapsto \begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix}. \end{aligned} \tag{11.46}$$

Par construction nous avons $f(M) = 0$ si et seulement si $M \in \mathcal{F}$. □

PropomhBwi

Proposition 11.26 ([59])**.**
Soit σ une partie de l'espace affine $\mathcal{E}$.

(1) L'intersection de tous les sous-espaces affines contenant σ est un sous-espace affine, noté $\mathcal{F}$.

(2) Si $A \in \sigma$, alors la direction de $\mathcal{F}$ est le sous-espace vectoriel

$$F = \text{Span}\{\overrightarrow{AM} \text{ tel que } M \in \sigma\}. \tag{11.47}$$

EqnRAUfg

Le sous-espace affine donné par la proposition 11.26 est le sous-espace affine **engendré** par la partie σ, et il est noté $\text{eae}(\sigma)$.

PROPooAKJBooMkmsiV

Proposition 11.27.
Soit $\mathcal{E}$ un espace affine de dimension n sur $\mathbb{K}$, soit $f\colon \mathcal{E} \to \mathbb{K}^r$ une fonction affine. Pour tout $a = (a_1, \ldots, a_r) \in \mathbb{K}^r$, l'ensemble $f^{-1}(a)$ est un sous-espace affine[4] de dimension $\dim \ker(u_f)$.

4. Définition 11.24.

Démonstration. Nous considérons le repère $(A, \{e_i\})$ de $\mathcal{E}$. Étant donné que f est affine nous avons

$$f\big(A + \sum_i x_i e_i\big) = f(A) + u_f\big(\sum_i x_i e_i\big). \tag{11.48}$$

Nous avons donc $f\big(A + \sum_i x_i e_i\big) = a$ lorsque

$$u_f(\sum_i x_i e_i) = a - f(A). \tag{11.49}$$

En utilisant le lemme 4.40, nous avons donc

$$f^{-1}(a) = A + (u_f)^{-1}\big(a - f(A)\big) = A + m + \ker(u_f) \tag{11.50}$$

où m est n'importe que élément de $u_f^{-1}\big(a - f(A)\big)$. □

PROPooUQLUooDQfYLT

Proposition 11.28.
Soit un espace vectoriel normé[5] $(V, \|.\|)$. *Pour tout* $a \in V$ *et* $r > 0$, *la boule* $B(a, r)$ *est convexe*[6]. *La boule fermée* $\overline{B(a, r)}$ *également.*

Démonstration. En deux parties.

(1) **La boule centrée en zéro** Soient $x, y \in B(0, r)$ et $\lambda \in]0, 1[$. Alors

$$\|\lambda x + (1 - \lambda)y\| \leqslant |\lambda|\|x\| + |1 - \lambda|\|y\| < (|\lambda| + |1 - \lambda|)r \leqslant r \tag{11.51}$$

où nous avons utilisé le fait que $|\lambda| = \lambda$ et $|1 - \lambda| = 1 - \lambda$.
Cela prouve que $\lambda x + (1 - \lambda)y \in B(0, r)$. Notez l'inégalité stricte due au fait que $\|x\| < r$ et $\|y\| < r$. Dans le cas de la boule fermée, nous avons une inégalité large.

(2) **La boule centrée autre part**
Soient $x, y \in B(a, r)$. Alors $x - a$ et $y - a$ sont dans $B(0, r)$, de telle sorte que

$$\lambda(x - a) + (1 - \lambda)(y - a) \in B(0, r) \tag{11.52}$$

par la première partie. En développant et simplifiant,

$$\lambda x + (1 - \lambda)y - a \in B(0, r), \tag{11.53}$$

ce qui signifie que $\lambda x + (1 - \lambda)y \in B(a, r)$.

□

PropPoNpPz

Proposition 11.29.
Soit A *un ensemble convexe*[7] *dans un espace vectoriel et* $v_1, \dots, v_n$ *des éléments de* A. *Alors toute combinaison*

$$a_1 v_1 + \cdots + a_n v_n \tag{11.54}$$

telle que $a_1 + \cdots + a_n = 1$ *et* $a_i \in [0, 1]$ *appartient à* A.

Démonstration. Nous prouvons la proposition pour $n = 3$. Nous devons trouver des nombres $t_1, t_2 \in [0, 1]$ tels que

$$t_2\big(t_1 v_1 + (1 - t_1)v_2\big) + (1 - t_2)v_3 = a v_1 + b v_2 + c v_3. \tag{11.55}$$

La réponse est immédiatement donnée par

$$t_2 a = 1 - c \tag{11.56a}$$
$$t_1 = a/t_2. \tag{11.56b}$$

Étant donné que $c \in [0, 1]$ nous avons $t_2 \in [0, 1]$. En ce qui concerne t_1 nous avons

$$t_1 = \frac{a}{t_2} \leqslant \frac{1 - c}{1 - c} = 1. \tag{11.57}$$

□

5. Définition 7.158.
6. Définition 7.154.
7. Définition 7.154.

11.7 Barycentre

Soit $\mathcal{E}$ un espace affine sur le $\mathbb{K}$-espace vectoriel E. Un couple (A, λ) avec $A \in \mathcal{E}$ et $\lambda \in \mathbb{K}$ est un **point pondéré**.

LemtEwnSH

Lemme-Définition 11.30 ([358]).
Soit une famille de points pondérés $\{(A_i, \lambda_i)\}_{i=1\ldots r}$. Si $\sum_i \lambda_i \neq 0$, alors il existe un unique $G \in \mathcal{E}$ tel que

$$\sum_{i=1}^{r} \lambda_i \overrightarrow{GA_i} = 0. \tag{11.58}$$

*Le point G donné par le lemme 11.30 est le **barycentre** des points pondérés (A_i, λ_i).*

Notons que l'on peut toujours supposer que $\sum_i \lambda_i = 1$ parce que le barycentre ne change pas lorsque tous les λ_i sont multipliés par un même nombre.

DefIMZooLFdIUB

Définition 11.31 (Combinaison convexe).
*Des nombres positifs ou nuls $\lambda_1, \ldots, \lambda_n$ vérifiant $\sum_i \lambda_i = 1$ forment une **combinaison convexe**.*

Le théorème suivant donne quelques caractérisations équivalentes du barycentre.

ThoIJVzxr

Théorème 11.32 ([358]).
Soient $\{(A_i, \lambda_i)\}_{i=1,\ldots,r}$ une famille de points pondérés. Les conditions suivantes sur le point $G \in \mathcal{E}$ sont équivalentes.

(1) Le point G est le barycentre de la famille.
(2) Pour tout $\alpha \in \mathbb{R}^*$, $\sum_i (\alpha\lambda_i)\overrightarrow{GA_i} = 0$.
(3) Il existe $A \in \mathcal{E}$ tel que $\big(\sum_i \lambda_i\big)\overrightarrow{AG} = \sum_i \lambda_i \overrightarrow{AA_i}$.
(4) Pour tout $B \in \mathcal{E}$, nous avons $\big(\sum_i \lambda_i\big)\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}$.

ItemEgOQBX

Définition 11.33.
*Si $A, B \in \mathcal{E}$, le **segment** $[AB]$ est l'ensemble des barycentres de A et B pondérés par des poids positifs (ouvert ou fermé suivant que l'on accepte que l'un ou l'autre des poids soit nul).*

Lorsque tous les λ_i sont égaux, nous parlons d'**isobarycentre**. Autrement dit, l'isobarycentre des points A_i est le barycentre des points pondérés $(A_i, 1)$.

11.7.1 Sous-espaces affines

Proposition 11.34.
Une partie $\mathcal{F}$ de $\mathcal{E}$ est un sous-espace affine si et seulement si elle est stable par barycentrisation.

Démonstration. Soit $\mathcal{F}$ un sous-espace affine de direction F et $A_1, \ldots, A_n$ des points de $\mathcal{F}$. Nous devons voir que le barycentre des points A_i pondérés de n'importe quelles masses appartient à $\mathcal{F}$. Pour ce faire nous faisons appel à la caractérisation (4) du théorème 11.32 : pour tout $B \in \mathcal{F}$,

$$\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}. \tag{11.59}$$

Puisque B et A_i sont dans $\mathcal{F}$, nous avons $\overrightarrow{BA_i} \in F$ et donc $\overrightarrow{BG} \in F$. Mais comme $B \in \mathcal{F}$, le point G est à son tour dans $\mathcal{F}$.

Réciproquement, nous supposons que $\mathcal{F}$ est stable par barycentrisme. Nous voudrions montrer que l'ensemble

$$F = \{\overrightarrow{AB} \text{ tel que } A, B \in \mathcal{F}\} \tag{11.60}$$

EqCmyWGi

est un sous-espace vectoriel. Soit $A \in \mathcal{F}$. Nous commençons par prouver que les vecteurs de la forme $\overrightarrow{AX}$ ($X \in \mathcal{F}$) forment un espace vectoriel. Considérons $\overrightarrow{AX} + \overrightarrow{AY}$ qui est un élément de E ; il existe donc $V \in \mathcal{E}$ tel que

$$\overrightarrow{AV} = \overrightarrow{AX} + \overrightarrow{AY}. \tag{11.61}$$

Par les relations de Chasles,

$$\overrightarrow{AV} = \overrightarrow{AV} + \overrightarrow{VX} + \overrightarrow{AV} + \overrightarrow{VY}, \tag{11.62}$$

donc

$$0 = \overrightarrow{VX} - \overrightarrow{VA} + \overrightarrow{VY}, \tag{11.63}$$

ce qui prouve que V est un barycentre de X, A, Y, et donc que $V \in \mathcal{F}$. De la même manière si $W \in \mathcal{E}$ est défini par $\overrightarrow{AW} = \mu\overrightarrow{AX}$, alors

$$\overrightarrow{AW} = \mu\overrightarrow{AX} = \mu(\overrightarrow{AW} + \overrightarrow{WX}), \tag{11.64}$$

ce qui signifie que

$$(1-\mu)\overrightarrow{AW} + \mu\overrightarrow{XW} = 0 \tag{11.65}$$

et que W est un barycentre.

Afin de montrer que (11.60) est bien un espace vectoriel, nous devons considérer $A, B, X, Y \in \mathcal{F}$ et prouver que $\overrightarrow{AX} + \overrightarrow{BY} \in F$. Nous avons

$$\begin{aligned}
\overrightarrow{AX} + \overrightarrow{BY} &= \overrightarrow{AX} + \overrightarrow{BA} + \overrightarrow{AY} && && (11.66a)\\
&= \overrightarrow{AV} + \overrightarrow{BA} && V \text{ est celui donné plus haut} && (11.66b)\\
&= \overrightarrow{AV} - \overrightarrow{AB} && && (11.66c)\\
&= \overrightarrow{AV} + \overrightarrow{AW} && W \text{ est donné par } \mu = -1. && (11.66d)\\
&= \overrightarrow{AV'}. && && (11.66e)
\end{aligned}$$

□

PropBVbCOS

Proposition 11.35 ([358]).
Soient $A_0, \dots, A_r$ des points de $\mathcal{E}$. L'ensemble des barycentres de ces points (avec des masses de somme 1) est le sous-espace affine engendré par les A_i que nous nommons $\mathcal{F}$.

Démonstration. Soit G le barycentre associé aux poids λ_i. Nous avons

$$G = A_0 + \overrightarrow{A_0G} = A_0 + \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i}. \tag{11.67}$$

Notons que les vecteurs $\overrightarrow{A_0A_i}$ sont dans la direction du sous-espace affine engendré par les A_i par (11.47). Donc G est bien dans $\mathcal{F}$.

Inversement si X est dans $\mathcal{F}$, on a

$$X = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i} \tag{11.68}$$

parce que $\sum_i \lambda_i \overrightarrow{A_0A_i}$ est un élément général de la direction de $\mathcal{F}$. Donc

$$\overrightarrow{A_0X} = \sum_i \lambda_i \overrightarrow{A_0A_i}, \tag{11.69}$$

et en utilisant la relation de Chasles sur chacun des $\overrightarrow{A_0A_i}$,

$$\overrightarrow{A_0X} = \sum_i \lambda_i \big(\overrightarrow{A_0X} + \overrightarrow{XA_i}\big). \tag{11.70}$$

De là nous concluons que

$$\big(1 - \sum_i \lambda_i\big)\overrightarrow{A_0X} + \sum_i \lambda_i \overrightarrow{A_iX} = 0, \tag{11.71}$$

ce qui signifie précisément que X est un barycentre des A_i. □

Proposition 11.36.
Soient $r+1$ point $A_0, \dots, A_r$ dans $\mathcal{E}$. Le sous-espace affine engendré par les A_i est au plus de dimension r.

Démonstration. La direction de l'espace engendré $\text{Aff}\{A_i\}$ est l'espace

$$\text{Span}\{\overrightarrow{A_0A_i}_{i=1,\dots,r}\} \tag{11.72}$$

qui est engendré par r vecteurs et donc est au plus de dimension r. □

En deux mots, la proposition suivante signifie que le barycentre des barycentres est le barycentre.

PropSFvjFZb

Proposition 11.37 (Associativité des barycentres[359])**.**
Soit une partition $J_0, \dots, J_r$ de $I = \{0, 1, \dots, n\}$. Soient des points $a_0, \dots, a_n \in \mathcal{E}$ et $\lambda_0, \dots, \lambda_n$ des nombres tels que $\sum_i \lambda_i \neq 0$. Nous supposons que $\mu_k = \sum_{i\in J_k} \lambda_i \neq 0$ pour tout k, et enfin nous nommons b_k le barycentre de la famille $\{(a_i, \lambda_i), i \in J_k\}$.

Alors le barycentre de la famille $\{(b_k, \mu_k)\}_{k=1,\dots,r}$ est le barycentre de la famille $\{(a_i, \lambda_i)\}_{i\in I}$.

Démonstration. Nous nommons b le barycentre des b_k pondérés par les μ_k, donc par définition

$$0 = \sum_{k=0}^{r} \mu_k \overrightarrow{bb_k} \tag{11.73a}$$

$$= \sum_k \sum_{i\in J_k} \lambda_i \overrightarrow{bb_k} \tag{11.73b}$$

$$= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i (\overrightarrow{ba_i} + \overrightarrow{a_ib_k}) \tag{11.73c}$$

$$= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i \overrightarrow{ba_i} + \sum_{k=0}^{r} \underbrace{\sum_{i\in J_k} \lambda_i \overrightarrow{a_ib_k}}_{=0} \tag{11.73d}$$

$$= \sum_{i\in I} \lambda_i \overrightarrow{ba_i}. \tag{11.73e}$$

Donc b est bien barycentre des a_i avec les poids λ_i. □

11.7.2 Enveloppe convexe

DefNLYYooXUHFUY

Définition 11.38.
*Soit A une partie d'un espace vectoriel E. L'**enveloppe convexe** de A, notée $\mathcal{C}\text{onv}(A)$ est l'intersection de tous les convexes contenant A.*

L'enveloppe convexe est un convexe. En effet soit C un convexe contenant A et $x, y \in \mathcal{C}\text{onv}(A)$; alors x et y sont dans C et par conséquent le segment $[x, y]$ est inclus dans C. Ce segment étant inclus dans tout convexe contenant A, il est inclus dans $\mathcal{C}\text{onv}(A)$.

PropSVvAQzi

Proposition 11.39 ([360])**.**
Soit C un convexe dans l'espace affine $\mathcal{E}$ et une famille de points pondérés $\{(a_i, \lambda_i)\}_{i=1,\dots,r}$ dont tous les poids sont positifs (et non tous nuls). Alors le barycentre est aussi dans C.

En d'autre termes, un convexe est stable par barycentrage à poids positifs[8]*.*

Démonstration. Nous prouvons par récurrence. D'abord pour $r = 2$. Le barycentre des points pondérés (a_1, λ_1), (a_2, λ_2) est le point b tel que

$$\lambda_1 \overrightarrow{ba_1} + \lambda_2 \overrightarrow{ba_2} = 0. \tag{11.74}$$

EqFWEErRX

8. Sauf si on prend tous les poids nuls ; mais contre ce genre d'idées, on ne peut rien faire.

Par définition, ce qui est noté $\overrightarrow{ab}$ n'est rien d'autre que $b-a$; en déballant (11.74), nous trouvons

$$\lambda_1(a_1-b)+\lambda_2(a_2-b)=0 \tag{11.75}$$

et donc

$$b=\frac{\lambda_1}{\lambda_1+\lambda_2}a_1+\frac{\lambda_2}{\lambda_1+\lambda_2}a_2, \tag{11.76}$$

qui est bien un point du segment $[a_1,a_2]$ parce que c'est une combinaison à coefficients positifs de somme 1.

Nous passons maintenant à la vraie récurrence avec un ensemble de points pondérés

$$A_r=\{(a_1,\lambda_1),\dots,(a_r,\lambda_r)\} \tag{11.77}$$

de masse totale non nulle ; et en vous laissant deviner ce que va désigner A_{r-1}. Si une des masses est nulle (disons λ_r), alors le barycentre de A_r est le même que celui de A_{r-1} et l'hypothèse de récurrence nous enseigne que ledit barycentre est dans C. Nous supposons donc que $\lambda_i\neq 0$ pour tout i. Dans ce cas le théorème d'associativité des barycentres 11.37 dit que le barycentre de A_r est le barycentre entre le barycentre de A_{r-1} et (a_r,λ_r), qui sont deux points de C par hypothèse de récurrence. □

Si E est un espace vectoriel et si $x_i\in E$ et $\lambda_i\in\mathbb{R}$, alors le barycentre des couples (x_i,λ_i) est le point g tel que $\sum_i\lambda_i\overrightarrow{gx_i}$, c'est-à-dire $\sum_i\lambda_i(x_i-g)=0$ ou encore

$$\sum_i\lambda_i x_i=\sum_i\lambda_i g. \tag{11.78}$$

Donc, quitte à diviser tous les λ_i par la somme, nous pouvons supposer que la somme des poids est 1. C'est pourquoi lorsque nous parlerons de barycentre dans un espace vectoriel sans contexte affine, nous allons toujours supposer $\sum_i\lambda_i=1$ et avoir le barycentre

$$g=\sum_i\lambda_i x_i. \tag{11.79}$$

PropYHMTmZX

Proposition 11.40.
Soit E, un espace vectoriel et $A\subset E$. L'enveloppe convexe $\mathcal{C}\mathrm{onv}(A)$ est l'ensemble des barycentres de familles finies de points affublés de masses positives.

Démonstration. Nous notons $\mathcal{B}$ l'ensemble des dits barycentres. Par la proposition 11.39, ces barycentres sont dans l'enveloppe convexe et donc $\mathcal{B}\subset\mathcal{C}\mathrm{onv}(A)$. A contrario, si nous prouvons que $\mathcal{B}$ était convexe, alors nous aurions $\mathcal{C}\mathrm{onv}(A)\subset\mathcal{B}$ parce que l'enveloppe convexe est l'intersection des convexes contenant A.

Soient $a,b\in\mathcal{B}$, c'est-à-dire que l'on a $a_0,\dots,a_n$ et $b_0,\dots,b_m$ dans A ainsi que les nombres strictement positifs $\lambda_0,\dots,\lambda_n$ et $\mu_0,\dots,\mu_m$ tels que

$$a=\sum_i\lambda_i a_i \qquad \sum_{i=1}^{n}\lambda_i=1 \tag{11.80a}$$

$$b=\sum_j\mu_j b_j \qquad \sum_{j=1}^{n}\mu_j=1 \tag{11.80b}$$

Un point du segment $[a,b]$ est de la forme $p=ta+(1-t)b$ avec $t\in[0,1]$. En développant,

$$p=\sum_{i=0}^{n}(t\lambda_i)a_i+\sum_{j=0}^{m}(1-t)\mu_j b_j. \tag{11.81}$$

C'est le barycentre de la famille $\{(a_i,\lambda_i),(b_j,\mu_j)\}$, parce que la somme des coefficients vaut bien 1 :

$$\sum_i(t\lambda_i)+\sum_j(1-t)\mu_j=t+(1-t)=1. \tag{11.82}$$

□

ThoJLDjXLe

Théorème 11.41 (Carathéodory[104])**.**
Dans un espace affine de dimension n, l'enveloppe convexe[9] *de A est l'ensemble des barycentres à coefficients positifs ou nuls de familles de $n+1$ points.*

Démonstration. Soit $x \in \mathcal{C}\text{onv}(A)$; on sait par la proposition 11.40 que x est barycentre de points de A avec des coefficients positifs :

EqWJDwQTH
$$x = \sum_{k=1}^{p} \lambda_k x_k \tag{11.83}$$

avec $\sum_k \lambda_k = 1$. Nous supposons que $p > n+1$ (sinon le théorème est réglé), et nous allons faire une récurrence à l'envers en montrant qu'on peut aussi écrire x sous forme d'un barycentre de strictement moins de p points.

Étant donné que $p-1 > n$, la famille $\{x+i-x_1\}_{i=2,\dots,p}$ est liée et il existe donc $\alpha_1, \dots, \alpha_p \in \mathbb{R}$ tels que $\sum_{i=2}^{p} \alpha_i(x_i - x_1) = 0$, c'est-à-dire telle que

$$\sum_{i=2}^{p} \alpha_i x_i = \sum_{i=2}^{p} \alpha_i x_1. \tag{11.84}$$

Nous posons $\alpha_1 = -\sum_{i=2}^{p} \alpha_1$. Remarquons qu'alors $\sum_{i=1}^{p} \alpha_i x_i = 0$ parce que

$$\sum_{i=1}^{p} \alpha_i x_i = \alpha_1 x_1 + \sum_{i=2}^{p} \alpha_i x_i = \alpha_1 x_1 + \sum_{i=2}^{p} \alpha_i x_1 = \sum_{i=1}^{p} \alpha_i x_1 = 0. \tag{11.85}$$

Par conséquent ça ne coûte rien de récrire (11.83) sous la forme

$$x = \sum_{i=1}^{p} (\lambda_i + t\alpha_i) x_i. \tag{11.86}$$

Les α_i ne sont pas tous nuls, mais leur somme est nulle, donc il y en a au moins un négatif. Nous notons

$$\tau = \min\{-\frac{\lambda_i}{\alpha_i} \text{ tel que } \alpha_i < 0\}, \tag{11.87}$$

et J l'ensemble de i pour lesquels ce minimum est atteint. Nous considérons aussi les nombres $\mu_i = \lambda_i + \tau\alpha_i$. Plusieurs remarques.

(1) Si $j \in J$, alors $\mu_j = 0$

(2) Si $\alpha_i > 0$ alors $\mu_i \geqslant 0$, mais si $\alpha_i < 0$ alors

$$\lambda_i + \tau\alpha_i \geqslant \lambda_i + (-\frac{\lambda_i}{\alpha_i})\alpha_i = 0 \tag{11.88}$$

donc $\mu_i \geqslant 0$ quand même.

(3) $\sum_{i=1}^{p} \mu_i = 1$, toujours parce que $\sum_{i=1}^{p} \alpha_i = 0$.

Avec tout ça, nous avons

$$\sum_{i \notin J} \mu_i x_i = \sum_{i=1}^{p} \mu_i x_i = x. \tag{11.89}$$

Et voilà, nous avons écrit x comme un barycentre à coefficients positifs de moins de p éléments parce que J n'est pas vide. □

CorOFrXzIf

Corolaire 11.42.
Dans un espace affine de dimension finie, l'enveloppe convexe d'un compact est compacte.

9. Définition 11.38.

Démonstration. Soit A une partie compacte de l'espace vectoriel E, et $\mathcal{C}\text{onv}(A)$ son enveloppe convexe. Nous allons montrer que toute suite dans $\mathcal{C}\text{onv}(A)$ admet une sous-suite convergente en écrivant un point de $\mathcal{C}\text{onv}(A)$ comme le théorème de Carathéodory 11.41 nous le suggère. Pour cela nous considérons le simplexe

$$\Lambda = \left\{ \lambda \in \mathbb{R}^{n+1} \text{ tel que } \sum_{k=1}^{n+1} \lambda_k = 1 \text{ et } \lambda_k \geqslant 0 \forall k \right\}. \tag{11.90}$$

Montrons en passant que Λ est compact. Si $\lambda_k \in \Lambda$ est une suite, alors chacun des λ_k est un $(n+1)$-uple de nombres dans $[0,1]$:

$$k \mapsto (\lambda_k)_i \tag{11.91}$$

est une suite qui possède une sous-suite convergente. En passant $n+1$ fois à une sous-suite, nous tombons sur une suite convergente vers $\lambda \in \Lambda$, grâce à la convergence composante par composante. De plus pour chaque k nous avons $\sum_{i=1}^{n+1}(\lambda_k)_i = 1$, et en passant à la limite, la somme étant une application continue, $\sum_i \lambda_i = 1$.

Considérons l'application

$$\begin{aligned} f\colon \Lambda \times A^{n+1} &\to \mathcal{C}\text{onv}(A) \\ (\lambda, x) &\mapsto \sum_{k=1}^{n+1} \lambda_k x_k. \end{aligned} \tag{11.92}$$

C'est une application continue parce qu'elle est bilinéaire en dimension finie ; son image est contenue dans $\mathcal{C}\text{onv}(A)$ par la proposition 11.39, et elle est surjective par le théorème de Carathéodory 11.41. Bref, $\mathcal{C}\text{onv}(A) = f(\Lambda \times A^{n+1})$ est donc l'image d'un compact par une application continue ; elle est donc compacte par le théorème 7.220. □

Notons que sans le théorème de Carathéodory, peut être que le nombre de points utiles pour décomposer les différents a_k n'était pas borné ; dans ce cas nous aurions du prendre une infinité de sous-suites et rien n'aurait été sûr.

11.7.3 Applications affines et barycentre

PROPooGSPZooRnVgiU

Proposition 11.43 ([361]).
Une application $f\colon \mathcal{E} \to \mathcal{E}'$ entre deux espaces affines est affine si et seulement si pour tout système $\{(A_i, \lambda_i)\}_{i=1,\dots,k}$ de barycentre G et de poids total non nul, le point $f(G)$ est barycentre du système $\{\big(f(A_i), \lambda_i\big)\}$.

Démonstration. En deux parties.

(1) **Si f est affine**
Par définition d'un barycentre,

$$\sum_i \lambda_i \overrightarrow{GA_i} = 0. \tag{11.93}$$

Nous considérons un point arbitraire $O \in \mathcal{E}$ et nous écrivons $A_i = O + x_i$, $G = O + x_g$. Ensuite nous utilisons le lemme 11.14 pour le calcul suivant :

$$\begin{aligned} \sum_i \lambda_i \overrightarrow{f(G)f(A_i)} &= \sum_i \lambda_i u_f(x_i - x_g) && \text{(11.94a)} \\ &= u_f\big(\sum_i \lambda_i (x_i - x_g)\big) && \text{(11.94b)} \\ &= u_f\big(\sum_i \lambda_i \overrightarrow{GA_i}\big) && \text{(11.94c)} \\ &= u_f(0) = 0. && \text{(11.94d)} \end{aligned}$$

Donc $f(G)$ est bien le barycentre du nouveau système.

(2) **Si f conserve les barycentres**

Nous définissons u par $f(O+x) = f(O) + u(x)$. À priori, ce u dépend de O et n'est pas linéaire.

(2a) **u est linéaire**

Soient $M, N \in \mathcal{E}$ et les éléments $x_m, x_n \in E$ tels que $\overrightarrow{OM} = x_m$ et $\overrightarrow{ON} = x_n$. Nous définissons enfin P par

$$\overrightarrow{OP} = \alpha\overrightarrow{OM} + \beta\overrightarrow{ON}, \tag{11.95}$$

et $P = O + x_p$. En décomposant $\overrightarrow{MO}$ et $\overrightarrow{NO}$ par les relations de Chasles de la proposition 11.5(1) nous avons

$$(\alpha+\beta-1)\overrightarrow{PO} - \alpha\overrightarrow{PM} - \beta\overrightarrow{PN} = 0 \tag{11.96}$$

et donc P est barycentre du système

$$\big\{(O, \alpha+\beta-1), (M,\alpha), (N,\beta)\big\}. \tag{11.97}$$

Le point $f(P)$ sera barycentre du système

$$\Big\{\big(f(O), \alpha+\beta-1\big), \big(f(M),\alpha\big), \big(f(N),\beta\big)\Big\}. \tag{11.98}$$

Cela signifie que

$$(\alpha+\beta-1)\overrightarrow{f(P)f(O)} - \alpha\overrightarrow{f(P)f(M)} - \beta\overrightarrow{f(P)f(N)} = 0. \tag{11.99}$$

En y substituant $\overrightarrow{f(P)f(O)} = u(-x_p)$, $\overrightarrow{f(P)f(M)} = u(x_m - x_p)$ et $\overrightarrow{f(P)f(N)} = u(x_n - x_p)$ ainsi que $x_p = \alpha x_m + \beta x_b$ nous trouvons

$$u(\alpha x_m + \beta x_n) = \alpha u(x_m) + \beta u(x_n). \tag{11.100}$$

Donc u est linéaire.

(2b) **u ne dépend pas du point O**

Il n'est pas besoin de démontrer cela parce que la définition 11.12 ne le demande pas. Note : c'est le lemme 11.13 qui dit que c'est par ailleurs vrai.

□

11.8 Repères, coordonnées cartésiennes et barycentriques

Définition 11.44.
*On dit que les points $A_0, \dots, A_r \in \mathcal{E}$ sont **affinement indépendants** si le sous-espace affine engendré est de dimension r.*

PropGAneHg

Proposition 11.45 ([358])**.**
Pour $r+1$ points $A_0, \dots, A_r$ dans $\mathcal{E}$, les propriétés suivantes sont équivalentes.

1. *Les A_i sont affinement indépendants.*
2. *Pour tout $i = 0, \dots, r$, le point A_i n'est pas dans $\mathrm{Aff}\{A_0, \dots, A_{i-1}, A_{i+1}, \dots, A_r\}$.* ItemrAzkIl
3. *Les points $A_0, \dots, A_{r-1}$ sont affinement indépendants et $A_r \notin \mathrm{Aff}\{A_0, \dots, A_{r-1}\}$.*
4. *Il existe i tel que les vecteurs $\overrightarrow{A_kA_i}$ $(k \neq i)$ sont linéairement indépendants.* ItemFBfcuq
5. *Pour tout $i \in \{1, \dots, r\}$, les vecteurs $\overrightarrow{A_kA_i}$ $(k \neq i)$ sont linéairement indépendants.*

Notons à propos de la condition (3) que l'existence d'un i pour lequel A_i n'est pas dans $\mathrm{Aff}\{A_0, \dots, A_{i-1}, A_{i+1}, \dots, A_r\}$ n'implique pas l'indépendance des $r+1$ points. En effet dans $\mathbb{R}^2$ nous considérons les 4 points $A_0 = (0,0)$, $A_1 = (1,0)$, $A_2 = (2,0)$ et $A_3 = (0,1)$. Évidemment le point A_3 n'est pas dans l'espace engendré par les trois autres ; il n'empêche que ces points ne sont pas affinement indépendants parce que la direction est de dimension 2 au lieu de 3.

DefguuwEO

Définition 11.46.
Soit $\mathcal{E}$ un espace affine de dimension n et $\mathcal{F}$ un sous-espace affine de dimension k. Un ***repère affine*** *de $\mathcal{F}$ est la donnée de $k+1$ points affinement indépendants de $\mathcal{F}$.*

Si $\{A_0, \dots, A_n\}$ est un repère affine, le point A_0 est l'**origine**. C'est un choix complètement arbitraire ; et c'est bien cet arbitraire qui nous amènera à considérer les coordonnées barycentriques au lieu des coordonnées cartésiennes.

Soit $M \in \mathcal{E}$; par définition nous avons

$$M = A_0 + \overrightarrow{A_0M}. \tag{11.101}$$

Mais nous savons que les vecteurs $\overrightarrow{A_0A_i}$ forment une base de E, nous avons donc des nombres λ_i tels que

$$\overrightarrow{A_0M} = \sum_{i=1}^{n} \lambda_i \overrightarrow{A_0A_i}. \tag{11.102}$$

Les nombres λ_i ainsi construits sont les **coordonnées cartésiennes** du point M dans le repère $\{A_0, \dots, A_n\}$ d'origine A_0.

À partir de ces coordonnées, le point $M \in \mathcal{E}$ se retrouve par la formule

$$M = A_0 + \sum_{i=1}^{n} \lambda_i \overrightarrow{A_0A_i}. \tag{11.103}$$

PROPooIXVBooPpKsDE

Proposition 11.47 ([1])**.**
La paire $\big(O, \{e_1, \dots, e_n\}\big)$ est un repère cartésien de $\mathcal{E}$ si et seulement si $\{O, O+e_1, \dots, O+e_n\}$ est un repère affine.

Démonstration. En deux parties.

(1) **Sens direct** Vue la proposition 11.45, il suffit de prouver que les vecteurs $\overrightarrow{O(O+e_i)}$ sont linéairement indépendants. Mais $\overrightarrow{O(O+e_i)} = e_i$, donc oui, ils sont linéairement indépendants.

(2) **Sens inverse** Il s'agit d'utiliser la même proposition 11.45 qui est encore applicable parce que c'est une équivalence.

□

11.48.
Soient (A, e_i) et (A', e'_i) deux repères cartésiens pour l'espace affine $\mathcal{E}$. Soit (s_{ij}) la matrice de changement de base entre $\{e_i\}$ et $\{e'_i\}$ dans E. Nous voudrions trouver les x_i en termes des x'_i.

Pour cela nous considérons un point M dans $\mathcal{E}$ et nous l'écrivons dans les deux bases. Cela fournit l'égalité

$$A + \sum_i x_i e_i = A' + \sum_i x'_i e'_i. \tag{11.104}$$

EqcYfuMg

Nous considérons les coordonnées (a_i) de A' dans le repère (A, e_i), c'est-à-dire

$$A' = A + \sum_i a_i e_i. \tag{11.105}$$

EqZNwPHE

En substituant $e'_i = \sum_k s_{jk} e_k$ et (11.105) dans (11.104) nous trouvons

$$\sum_k x_k e_k = \sum_k a_k e_k + \sum_{jk} s_{jk} x'_j e_k, \tag{11.106}$$

et par conséquent

$$x_k = a_k + \sum_j s_{jk} x'_j. \tag{11.107}$$

Les coordonnées barycentriques sont données par la proposition suivante.

PROPooTIRXooLAipRa

Proposition 11.49 ([358]).
Soient $A_0, \dots, A_r$ des points affinement indépendants dans $\mathcal{E}$ et $\mathcal{F} = \text{Aff}\{A_0, \dots, A_r\}$. Tout point $M \in \mathcal{F}$ s'écrit de façon unique comme barycentre[10] *des A_i affectés de poids λ_i tels que $\sum_{i=0}^{r} \lambda_i = 1$.*

Démonstration. Nous avons vu plus haut (définition 11.46) que l'affine indépendance des points A_i assurait que $(A_0, \dots, A_r)$ était un repère de $\mathcal{F}$.

En ce qui concerne l'existence de l'écriture de M comme barycentre, nous savons que les sous-espace affines sont exactement les ensembles de barycentres (proposition 11.35), c'est-à-dire que si on a des points dans un sous-espace affine, alors les barycentres de ces points est encore dans le sous-espace affine.

L'unicité est comme suit. Si M est barycentre des A_i avec poids λ_i, nous écrivons la caractérisation (4) du théorème 11.32 avec $B = A_0$:

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} \tag{11.108}$$

où la somme à droite s'étend à priori de 0 à r, mais comme $\overrightarrow{A_0A_0} = 0$, nous l'avons limitée à 1. Si M s'écrit comme barycentre de deux façons différentes, nous aurions

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} = \sum_{i=1}^{r} \mu_i \overrightarrow{A_0A_i} \tag{11.109}$$

avec $\sum_i \lambda_i = \sum_i \mu_i = 1$. Étant donné que les points $A_0, \dots, A_r$ forment un repère, les vecteurs $\overrightarrow{A_0A_i}$ sont linéairement indépendants (point (5) de la proposition 11.45) et donc $\lambda_i = \mu_i$ pour $i = 1, \dots, r$. La condition de somme des poids égale à 1 impose alors immédiatement $\lambda_0 = \mu_0$. □

DEFooTXPPooQdacbO

Définition 11.50.
Soit un espace affine $\mathcal{E}$ de dimension n. Soient des points affinement indépendants $A_1, \dots, A_n$. Pour $M \in \mathcal{E}$, la proposition 11.49 indique qu'il existe un unique choix de λ_i tel que

$$\begin{cases} \sum_i \lambda_i = 1 & \text{(11.110a)} \\ \sum_i \lambda_i \overrightarrow{MA_i} = 0. & \text{(11.110b)} \end{cases}$$

Ces λ_i sont les ***coordonnées barycentriques*** *de M dans le repère $\{A_i\}_{i=1,\dots,n}$.*

NORMooOGHBooMjmouu

11.51.
Soit $\mathbb{R}^2$ et les points non alignés A, B, C. Les coordonnées barycentriques (α, β, γ) dans ce système correspondent à l'unique $X \in \mathbb{R}^2$ tel que

$$\alpha\overrightarrow{XA} + \beta\overrightarrow{XB} + \gamma\overrightarrow{XC} = 0. \tag{11.111}$$

Exemple 11.52.
Soient les points $A = (3, 1)$, $B = (-1, 2)$ et $C = (0, -1)$ dans $\mathbb{R}^2$. Nous allons montrer qu'il forment un repère affine de $\mathbb{R}^2$. L'espace engendré par ces trois points est l'espace des

$$A + \alpha\overrightarrow{AB} + \beta\overrightarrow{AC}, \tag{11.112}$$

et la direction correspondante est l'espace vectoriel donné par $\alpha\overrightarrow{AB} + \beta\overrightarrow{AC}$ qui est de dimension deux. Donc l'espace affine engendré par A, B et C est de dimension 2. △

10. Définition 11.30.

Exemple 11.53.
Dans le repère (A, B, C), quel est le point de coordonnées barycentriques $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$? D'abord nous vérifions que

$$\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1. \tag{11.113}$$

Ensuite nous cherchons $X \in \mathbb{R}^2$ tel que

$$\frac{1}{6}\overrightarrow{AX} + \frac{1}{3}\overrightarrow{BX} + \frac{1}{2}\overrightarrow{CX} = 0, \tag{11.114}$$

c'est-à-dire

$$\frac{1}{6}\begin{pmatrix} x-3 \\ y-1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} x+1 \\ y-2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x \\ y+1 \end{pmatrix} = 0. \tag{11.115}$$

Nous trouvons immédiatement $x = 1/6$ et $y = 1/3$. Le point cherché est donc le point $\begin{pmatrix} 1/6 \\ 1/3 \end{pmatrix}$. △

LEMooDUMVooFtfFOe

Lemme 11.54 ([1]).
Une application affine $f : \mathcal{E} \to \mathcal{E}$ qui préserve les points d'une base affine de $\mathcal{E}$ est l'identité.

Démonstration. Une base affine de $\mathcal{E}$ consiste en $n+1$ points $\{A_0, \dots, A_n\}$ affinement indépendants. Nous utilisons la proposition 11.47 pour dire que $\big(A_0, \{\overrightarrow{A_0A_i}\}_{i=1,\dots,n}\big)$ est un repère cartésien.

En utilisant la formule du lemme 11.13,

$$f(A_i) = f(A_0 + \overrightarrow{A_0A_i}) = f(A_0) + u(\overrightarrow{A_0A_i}). \tag{11.116}$$

Donc $A_i = A_0 + u(\overrightarrow{A_0A_i})$, ce qui signifie que

$$u(\overrightarrow{A_0A_i}) = \overrightarrow{A_0A_i} \tag{11.117}$$

EQooAJYWooBTucpp

Par ailleurs, tout point M [11] de $\mathcal{E}$ peut être écrit sous la forme

$$M = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i}. \tag{11.118}$$

En appliquant f, et en utilisant (11.117),

$$f(M) = f(A_0) + \sum_i \lambda_i u(\overrightarrow{A_0A_i}) = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i} = M. \tag{11.119}$$

Donc tout point de $\mathcal{E}$ est fixé par f, ce qui signifie que f est l'identité. □

11.8.1 Équation de droite

DEFooCYDPooEdRbyl

Proposition-Définition 11.55.
Soit $\mathcal{E}$ un espace affine de dimension trois muni d'un repère barycentrique [12] *(A_1, A_2, A_3). Nous notons $D(a, b, c)$ l'ensemble des éléments de $\mathcal{E}$ dont les coordonnées barycentriques (normalisées) (x, y, z) vérifient $ax + by + cz = 0$, c'est-à-dire l'ensemble des $M \in \mathcal{E}$ tels que*

$$\begin{cases} x + y + z = 1 & \text{(11.120a)} \\ x\overrightarrow{MA_1} + y\overrightarrow{MA_2} + z\overrightarrow{MA_3} = 0. & \text{(11.120b)} \end{cases}$$

Alors l'ensemble $D(a, b, c)$ est un sous-espace de dimension 1 de $\mathcal{E}$.

*Nous nommons **droite affine** une telle partie de $\mathcal{E}$.*

Idée de preuve : ne pas oublier la condition $x + y + z = 1$ parce que la somme des coordonnées barycentriques doit valoir 1.

11. Même les points qui ne s'appellent pas « M » en fait.
12. Définition 11.50.

Exemple 11.56.
La droite $D(1,1,1)$ n'existe pas parce que ce serait $x+y+z=0$, qui est incompatible avec $x+y+z=1$. △

Les droites $D(a,b,c)$ et $D(a',b',c')$ s'intersectent selon les solutions du système

$$\begin{cases} x+y+z=1 & \text{(11.121a)}\\ ax+by+cz=0 & \text{(11.121b)}\\ a'x+b'y+c'z=0 & \text{(11.121c)} \end{cases}$$

Donc deux droites affines ont un unique point d'intersection si et seulement si

$$d=\begin{vmatrix} 1 & 1 & 1\\ a & b & c\\ a' & b' & c' \end{vmatrix}\neq 0. \tag{11.122}$$

Elles seront parallèles ou confondues si et seulement si $d=0$.

11.8.2 Associativité, coordonnées barycentriques dans un triangle

Lemme 11.57 ([362])**.**
Soient trois points non alignés A, B, C ainsi que des nombres α, β, γ tels que $\alpha+\beta\neq 0$ et $\alpha+\beta+\gamma\neq 0$.

Soit H le barycentre du système $\{(A,\alpha),(B,\beta)\}$ et G le barycentre de $\{(A,\alpha),(B,\beta),(C,\gamma)\}$.

Alors G est barycentre de $\{(H,\alpha+\beta),(C,\gamma)\}$.

Démonstration. Vues les définitions de H et G nous avons

$$\alpha\overrightarrow{HA}+\beta\overrightarrow{HB}=0 \tag{11.123a}$$
$$\alpha\overrightarrow{GA}+\beta\overrightarrow{GB}+\gamma\overrightarrow{GC}=0. \tag{11.123b}$$

En utilisant les relations de Chasles nous introduisons H dans la seconde relation :

$$\alpha(\overrightarrow{GH}+\overrightarrow{HA})+\beta(\overrightarrow{GH}+\overrightarrow{HB})+\gamma\overrightarrow{GC}=0 \tag{11.124a}$$
$$(\alpha+\beta)\overrightarrow{GH}+\underbrace{\alpha\overrightarrow{HA}+\beta\overrightarrow{HB}}_{=0}+\gamma\overrightarrow{GC}=0 \tag{11.124b}$$
$$(\alpha+\beta)\overrightarrow{GH}+\gamma\overrightarrow{GC}=0. \tag{11.124c}$$

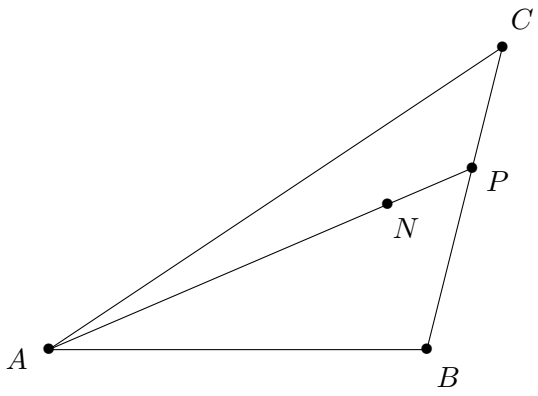
□

Les coordonnées barycentriques dans un triangle (et plus généralement en fait) permettent de faire des projections.

PROPooBCUVooWKttiH

Proposition 11.58.
Soient trois points non alignés A, B, C ainsi qu'un point N de coordonnées barycentriques (α,β,γ) dans le système (A,B,C). Si P est l'intersection $(AN)\cap(BC)$ alors les coordonnées de P sont $(0,\beta,\gamma)$.

Démonstration. Un dessin de la situation :



Dire que les coordonnées de N sont (α,β,γ) signifie que

$$\alpha\overrightarrow{NA}+\beta\overrightarrow{NB}+\gamma\overrightarrow{NC}=0. \tag{11.125}$$

Nous voudrions montrer que le point P est bien le point de coordonnées $(0,\beta,\gamma)$. Soit donc le point P tel que

$$\beta\overrightarrow{PB}+\gamma\overrightarrow{PC}=0 \tag{11.126}$$ EQooYLGDooWqKKQM

et montrons que ce point est l'intersection $(BC)\cap(NA)$.

D'abord la relation (11.126) nous dit immédiatement que P est sur la droite (BC). Ensuite, en utilisant les relations de Chasles pour introduire N :

$$\beta(\overrightarrow{PN}+\overrightarrow{NB})+\gamma(\overrightarrow{PN}+\overrightarrow{NC})=0. \tag{11.127}$$

Nous remplaçons $\beta\overrightarrow{NB}+\gamma\overrightarrow{NC}$ par $-\alpha\overrightarrow{NA}$ pour obtenir :

$$(\beta+\gamma)\overrightarrow{PN}-\alpha\overrightarrow{NA}=0. \tag{11.128}$$

Cela montre que les vecteurs $\overrightarrow{PN}$ et $\overrightarrow{NA}$ sont colinéaires, et donc que P, N et A sont alignés. □

11.9 Applications affines sur $\mathbb{R}^n$

Soit $v\in\mathbb{R}^n$; nous notons $\tau_v\colon\mathbb{R}^n\to\mathbb{R}^n$ la translation donnée par $\tau_v(x)=x+v$. Le groupe de toutes les translations de $\mathbb{R}^n$ est noté $T(n)$ et est isomorphe au groupe abélien $(\mathbb{R}^n,+)$.

Nous avons déjà discuté de la structure d'un espace vectoriel (en particulier $\mathbb{R}^n$) comme espace affine en 11.6.

LEMooZZAIooOMiayy

Lemme 11.59.
Décomposition d'une application affine.

ITEMooSJBFooYHURto

(1) *Une application $f\colon E\to E$ est affine si et seulement si il existe $v\in E$ et une application linéaire α sur E telle que $f=\tau_v\circ\alpha$.*

ITEMooPYEOooKIesYm

(2) *Dans ce cas, le choix de (v,α) est unique.*

ITEMooHIAUooRxfTqx

(3) *Si f est bijective, alors α est bijective.*

Démonstration. Nous supposons d'abord que f est affine. Alors il existe une application linéaire u_f sur E telle que

$$f(M+x)=f(M)+u_f(x)=(\tau_{f(M)}\circ u_f)(x) \tag{11.129}$$

pour tout x et M. De plus l'application u_f ne dépend ni de M ni de x (c'est la proposition 11.16(1)). En posant $M=0$ nous avons :

$$f(x)=(\tau_{f(0)}\circ u_f)(x). \tag{11.130}$$

Dans l'autre sens nous supposons avoir $v\in E$ et α linéaire sur E telles que

$$f(M)=(\tau_v\circ\alpha)(M). \tag{11.131}$$

Notons qu'il y a un abus de notation entre α qui est linéaire sur l'espace *vectoriel* E et l'application α qui est une application sur l'espace *affine* E. Cet abus est légitime parce que les deux espaces sont identiques en tant qu'ensembles. Ce qui est vraiment abuser par contre, c'est de se poser ce genre de questions.

Nous avons :

$$\begin{aligned} f(M+x) = \tau_v\big(\alpha(M+x)\big) = \alpha(M+x) + v = \alpha(M) + v + \alpha(x) \\ = (\tau_v \circ \alpha)(M) + \alpha(x) = f(M) + \alpha(x). \end{aligned} \tag{11.132}$$

Donc la fonction f vérifie la définition 11.12. La partie (1) est prouvée.

Pour prouver l'unicité de la partie (2), nous supposons que $\tau_v \circ \alpha = \tau_w \circ \alpha$. En appliquant cela à 0 nous trouvons $v = w$. Nous avons donc $\tau_v \circ \alpha = \tau_v \circ \beta$. Comme τ_v est inversible, nous en déduisons $\alpha = \beta$.

Enfin le point (3) est relativement évident du fait que τ_v, elle, est surement bijective. □

CORooATCNooUwEPNI

Corolaire 11.60.
Une application affine qui conserve l'origine est linéaire.

Démonstration. Conserver l'origine demande de poser $v = 0$ dans l'expression du lemme 11.59. □

PROPooYRCJooIcmUVI

Proposition 11.61.
Soit une application affine $f : \mathbb{R}^n \to \mathbb{R}^n$. L'ensemble des points fixes

$$\mathrm{Fix}(f) = \{x \in \mathbb{R}^n \text{ tel que } f(x) = x\} \tag{11.133}$$

est soit vide soit un sous-espace affine de $\mathbb{R}^n$.

Démonstration. Soit $f = \tau_v \circ \alpha$; nous avons $x \in \mathrm{Fix}(f)$ si et seulement si

$$x = \tau_v\big(\alpha(x)\big) = \alpha(x) + v, \tag{11.134}$$

autrement dit, en considérant l'application linéaire $\beta = \mathrm{Id} - \alpha$, si et seulement si $\beta(x) = v$. Nous écrivons $\mathrm{Fix}(f) = \beta^{-1}(v)$. Supposons que cet ensemble soit non vide et considérons $x_0 \in \beta^{-1}(v)$. Nous avons

$$\begin{aligned} \beta^{-1}(v) &= \{x \in \mathbb{R}^n \text{ tel que } \beta(x) = \beta(x_0)\} && \text{(11.135a)} \\ &= \{x \text{ tel que } \beta(x - x_0) = 0\} && \text{(11.135b)} \\ &= \{x \text{ tel que } x - x_0 \in \ker(\beta)\} && \text{(11.135c)} \\ &= \ker(\beta) + x_0 && \text{(11.135d)} \\ &= \tau_{x_0}\big(\ker(\beta)\big). && \text{(11.135e)} \end{aligned}$$

Mais comme $\ker(\beta)$ est un sous-espace vectoriel, $\beta^{-1}(v)$ est le translaté d'un sous-espace vectoriel, c'est-à-dire un sous-espace affine. □

11.9.1 Structure de groupe pour les applications affines

PROPooBPKKooJRAMeT

Proposition-Définition 11.62 ([1])**.**
L'ensemble des applications affines bijectives de $\mathbb{R}^n$ forment un groupe pour la composition. Les lois de groupe sont données par les formules suivantes :

(1) Le neutre est l'identité.

ITEMooGUFRooMuhXds

(2) Le produit est donné par

$$(\tau_v \circ \alpha)(\tau_w \circ \beta) = \tau_{\alpha(w)+v} \circ \alpha\beta. \tag{11.136}$$

EQooMIFSooKIvPmW

ITEMooYOMSooRUDSdm

(3) L'inverse est donné par

$$(\tau_v \circ \alpha)^{-1} = \tau_{-\alpha^{-1}(v)} \circ \alpha^{-1}. \tag{11.137}$$

Ce groupe est noté $\mathrm{Aff}(\mathbb{R}^n)$.

Démonstration. Pour l'identité, oui, composer par l'identité est neutre.

Le fait que la formule (11.136) soit vraie est un simple calcul :

$$(\tau_v \circ \alpha) \circ (\tau_w \circ \beta)(x) = (\alpha\beta)(x) + \alpha(w) + v = \big(\tau_{\alpha(w)+v} \circ \alpha\beta\big)x. \tag{11.138}$$

Le fait que la formule (11.136) donne bien un produit pour tous les éléments de $\mathrm{Aff}(\mathbb{R}^n)$ est le lemme 11.59.

En ce qui concerne l'inverse, c'est un calcul :

$$\begin{aligned}(\tau_{-\alpha^{-1}(v)}\alpha^{-1})(\tau_v\alpha)(x) &= (\tau_{-\alpha^{-1}(v)}\alpha^{-1})\big(\alpha(x)+v\big) && \text{(11.139a)}\\ &= \tau_{-\alpha^{-1}(v)}\big(x+\alpha^{-1}(v)\big) && \text{(11.139b)}\\ &= x. && \text{(11.139c)}\end{aligned}$$

□

Si $f\colon \mathbb{R}^n \to \mathbb{R}^n$ est une application affine, la proposition 11.59 affirme qu'il existe une application linéaire u telle que

$$f(x+y) = f(x) + u(y). \tag{11.140}$$

En écrivant cela pour $x = 0$,

$$f(y) = f(0) + u(y), \tag{11.141}$$

ou encore $f = \tau_{f(0)} \circ u$.

PROPooTPFZooKtFxhg

Proposition 11.63.
L'ensemble $\mathrm{Aff}(\mathbb{R}^n)$ *est isomorphe au produit semi-direct*[13]

$$\mathrm{Aff}(\mathbb{R}^n) \simeq T(n) \times_{\boldsymbol{Ad}} \mathrm{GL}(n,\mathbb{R}) \tag{11.142}$$

où $\boldsymbol{Ad}$ *est l'action adjointe, c'est-à-dire*

$$\begin{aligned}\boldsymbol{Ad}\colon \mathrm{GL}(n,\mathbb{R}) &\to \mathrm{Aut}\big(T(n)\big)\\ \alpha &\mapsto \big(\tau_v \mapsto \alpha\circ\tau_v\circ\alpha^{-1}\big).\end{aligned} \tag{11.143}$$

Démonstration. En plusieurs points.

(1) **Égalité d'ensembles** Il faut que $\mathrm{Aff}(\mathbb{R}^n)$ soit en bijection avec $T(n)\times\mathrm{GL}(n,\mathbb{R})$. En effet si $f \in \mathrm{Aff}(\mathbb{R}^n)$, la décomposition $f = \tau_v\circ\alpha$ est unique. D'abord en appliquant à 0, $f(0) = \tau_v\big(\alpha(v)\big) = v$. Donc v est fixé par la valeur de $f(0)$. Ensuite $\alpha = f\circ\tau_v^{-1}$, donc α fixé.

(2) **L'action adjointe fonctionne** Il faut vérifier que $\alpha\circ\tau_v\circ\alpha^{-1}$ est bien dans $T(n)$. Pour cela, en agissant sur $x\in\mathbb{R}^n$ nous trouvons

$$\alpha\tau_v\alpha^{-1}(x) = \alpha\big(\alpha^{-1}(x)+v\big) = x + \alpha(v) = \tau_{\alpha(v)}(x). \tag{11.144}$$

Le fait que $\mathbf{Ad}(\alpha)$ soit un automorphisme est toujours correct.

(3) **Morphisme** Nous montrons que

$$\begin{aligned}\psi\colon T(n)\times\mathrm{GL}(n,\mathbb{R}) &\to \mathrm{Aff}(\mathbb{R}^n)\\ (\tau_v,\alpha) &\mapsto \tau_v\circ\alpha\end{aligned} \tag{11.145}$$

est un isomorphisme de groupes. D'abord la loi de groupe sur $\mathrm{Aff}(\mathbb{R}^n)$ est donnée par

$$(\tau_v\circ\alpha)\circ(\tau_w\circ\beta) = \tau_{v+\alpha(w)}\circ(\alpha\circ\beta). \tag{11.146}$$

Ensuite la loi de groupe du produit semi-direct est donnée par

$$(\tau_v,\alpha)\cdot(\tau_w,\beta) = \big(\tau_v\,\mathbf{Ad}(\alpha)\tau_w,\alpha\beta\big) = \big(\tau_v\tau_{\alpha(w)},\alpha\beta\big) = \big(\tau_{\alpha(w)+v},\alpha\beta\big). \tag{11.147}$$

Nous avons donc bien

$$\psi\big((\tau_v,\beta)\cdot(\tau_w,\beta)\big) = \psi(\tau_v,\beta)\circ\psi(\tau_w,\beta). \tag{11.148}$$

□

13. Définition 2.47.

11.10 Isométries

DEFooZGKBooGgjkgs

Définition 11.64 (Isométrie d'espace affine)**.**
Si $\mathcal{E}$ est un espace affine muni d'une distance d, une isométrie de $\mathcal{E}$ est une application $f\colon \mathcal{E} \to \mathcal{E}$ préservant d, c'est-à-dire telle que

$$d(x,y) = d\big(f(x), f(y)\big). \tag{11.149}$$

Notons que toutes les applications affines ne sont pas des isométries : par exemple les homothéties.

PROPooHSOGooBbFTYt

Proposition 11.65.
Si $\mathcal{E}$ est modelé sur un espace euclidien $(E, \|.\|)$ alors la formule

$$d(A,B) = \|\overrightarrow{AB}\| \tag{11.150}$$

définit une distance[14] *sur $\mathcal{E}$.*

Démonstration. Étant donné la proposition 11.3, la formule a un sens parce qu'à A et B donnés dans $\mathcal{E}$, il est associé un unique vecteur $\overrightarrow{AB} \in E$.

Pour vérifier que d est une distance, nous vérifions les points de la définition 7.115 et nous utilisons les propriétés correspondantes dans la définition 7.158 d'une norme.

(1) $d(A,B) = \|\overrightarrow{AB}\| \geqslant 0$.

(2) Si $d(A,B) = 0$, alors $\|\overrightarrow{AB}\| = 0$, ce qui implique que $\overrightarrow{AB} = 0$. Nous avons donc

$$\begin{aligned} B &= A + \overrightarrow{AB} && \text{proposition 11.3} && (11.151a)\\ &= A + 0 && && (11.151b)\\ &= A && \text{lemme 11.4.} && (11.151c) \end{aligned}$$

(3) En utilisant la proposition 11.5(3),

$$d(A,B) = \|\overrightarrow{AB}\| = \|-\overrightarrow{BA}\| = \|\overrightarrow{BA}\| = d(B,A) \tag{11.152}$$

(4) En utilisant les relation de Chasles 11.5(1) ainsi que l'inégalité triangulaire 7.158(4)

$$d(A,B) = \|\overrightarrow{AB}\| = \|\overrightarrow{AC} + \overrightarrow{CB}\| \leqslant \|\overrightarrow{AC}\| + \|\overrightarrow{CB}\| = d(A,C) + d(C,B). \tag{11.153}$$

□

Nous parlons d'isométries affines ou linéaires dans le thème 78.

14. Définition 7.115.

Chapitre 12

Analyse réelle : limites et dérivation

12.1 Limite de fonctions

12.1.1 Définition

La définition générale de la limite est 7.110. Dans le cas de fonctions $\mathbb{R} \to \mathbb{R}$, elle peut s'écrire de façon plus efficace. La proposition suivante montre comment fonctionne la limite pour une fonction définie sur tout $\mathbb{R}$.

PropAJQQooQQClfp

Proposition 12.1 (Caractérisation de la limite)**.**
Soit une fonction $f : \mathbb{R} \to \mathbb{R}$ définie sur $\mathbb{R}$ et $a \in \mathbb{R}$. La fonction f admet la limite ℓ pour $x \to a$ si et seulement si il existe un réel ℓ tel que pour tout $\epsilon > 0$, il existe un $\delta > 0$ tel que

$$0 < |x - a| < \delta \Rightarrow |f(x) - \ell| < \epsilon. \tag{12.1}$$

EqDefLimiteFonction

Démonstration. Il s'agit de montrer l'équivalence avec la définition 7.110. Nous allons faire un usage intensif de la proposition 7.118.

(1) **Sens direct** Soient $\epsilon > 0$ et $V = B(\ell, \epsilon)$. Alors il existe un voisinage W de a dans $\mathbb{R}$ tel que

$$f\big(W\backslash\{a\}\big) \subset V. \tag{12.2}$$

Soit δ tel que $B(a, \delta) \subset W$. Nous avons encore

$$f\big(B(a, \delta)\backslash\{a\}\big) \subset V. \tag{12.3}$$

Soit maintenant $x \in \mathbb{R}$ tel que $0 < |x - a| < \delta$. Cela signifie $x \in B(a, \delta)\backslash\{a\}$. Pour un tel x nous avons donc $f(x) \in B(\ell, \epsilon)$, c'est-à-dire $|f(x) - \ell| < \epsilon$.

(2) **Dans l'autre sens** Soient un voisinage V de ℓ et $\epsilon > 0$ tel que $B(\ell, \epsilon) \subset V$. Nous considérons δ tel que $|f(x) - \ell| < \epsilon$ pour tout x satisfaisant $0 < |x - a| < \delta$.
Avec tout cela nous posons $W = B(x, \delta)$, et nous avons

$$f\big(W\backslash\{a\}\big) \subset B(\ell, \epsilon) \subset V. \tag{12.4}$$

□

Si aucun nombre ℓ ne vérifie la condition de la définition, alors on dit que la fonction n'admet pas de limite en a. Lorsque f possède la limite ℓ en a, nous notons

$$\lim_{x \to a} f(x) = \ell. \tag{12.5}$$

La proposition suivante a déjà été démontrée dans la proposition 7.113. Nous en donnons ici une démonstration adaptée au cas $\mathbb{R} \to \mathbb{R}$.

Proposition 12.2.
Soit une fonction $f : D \to \mathbb{R}$. Si a est un point d'accumulation de D et si il existe une limite de f en a, alors il en existe une seule.

Démonstration. Nous prouvons qu'il ne peut pas exister deux nombres $\ell \neq \ell'$ vérifiant tous les deux la condition (12.1).

Soient ℓ et ℓ' deux limites de f au point a. Par définition, pour tout ϵ nous avons des nombres δ et δ' tels que

EqsContf2307Right
$$\begin{aligned} |x-a| < \delta &\Rightarrow \big|f(x)-\ell\big| < \epsilon \\ |x-a| < \delta' &\Rightarrow \big|f(x)-\ell'\big| < \epsilon \end{aligned} \tag{12.6}$$

Pour fixer les idées, supposons que $\delta < \delta'$ (le cas $\delta \geqslant \delta'$ se traite de la même manière).

Étant donné que a est un point d'accumulation du domaine D de f, il existe un $x \in D$ tel que $|x-a| < \delta$. Évidemment, nous avons aussi $|x-a| < \delta'$. Les conditions (12.6) signifient alors que ce x vérifie en même temps

$$|f(x)-\ell| < \epsilon, \tag{12.7}$$

et

$$|f(x)-\ell'| < \epsilon. \tag{12.8}$$

Afin de prouver que $\ell = \ell'$, nous allons maintenant calculer $|\ell - \ell'|$ et montrer que cette distance est plus petite que tout nombre. Nous avons (voir remarque 12.3)

EqInesq2307ellellepr
$$|\ell-\ell'| = |\ell - f(x) + f(x) - \ell'| \leqslant |\ell - f(x)| + |f(x) - \ell'| < \epsilon + \epsilon. \tag{12.9}$$

En résumé, pour tout $\epsilon > 0$ nous avons

$$|\ell - \ell'| < 2\epsilon, \tag{12.10}$$

et donc $|\ell - \ell'| = 0$, ce qui signifie que $\ell = \ell'$. □

RemTechniqueIneqs
Remarque 12.3.
Les inégalités (12.9) utilisent deux techniques très classiques en analyse qu'il convient d'avoir bien compris. La première est de faire

$$|A - B| = |A - C + C - B|. \tag{12.11}$$

Il s'agit d'ajouter $-C + C$ dans la norme. Évidemment, cela ne change rien.

La seconde technique est l'inégalité

$$|A + B| \leqslant |A| + |B|. \tag{12.12}$$

Exemple 12.4.
Considérons la fonction $f(x) = 2x$, et calculons la limite $\lim_{x\to 3} f(x)$. Vu que $f(3) = 6$, nous nous attendons à avoir $\ell = 6$. C'est ce que nous allons prouver maintenant. Pour chaque $\epsilon > 0$ nous devons trouver un $\delta > 0$ tel que $|x-3| < \delta$ implique $|f(x)-6| < \epsilon$. En remplaçant $f(x)$ par sa valeur en fonction de x et avec quelques manipulations nous trouvons :

$$\begin{aligned} |f(x)-6| &< \epsilon \\ |2x-6| &< \epsilon \\ 2|x-3| &< \epsilon \\ |x-3| &< \frac{\epsilon}{2} \end{aligned} \tag{12.13}$$

Donc dès que $|x-3| < \frac{\epsilon}{2}$, nous avons $|f(x)-6| < \epsilon$. Nous posons donc $\delta = \frac{\epsilon}{2}$.

Plus généralement, nous avons $\lim_{x\to a} f(x) = 2a$, et cela se prouve en étudiant $|f(x) - 2a|$ exactement de la même manière. △

12.1.2 Quelques règles de calcul

LEMooSPXNooPSufAn

Lemme 12.5 ([1]).
Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{\epsilon\to 0} f(\epsilon) = 0$. Soit $a \in \mathbb{R}$ tel que $|a| \leqslant f(\epsilon)$ pour tout ϵ. Alors $a = 0$.

Démonstration. Soit $\epsilon' > 0$. Il existe $\delta > 0$ tel que $|f(\epsilon)| < \epsilon'$ pour tout $|\epsilon| < \delta$. Dès que $\epsilon < \delta$, nous avons

$$0 \leqslant |a| \leqslant f(\epsilon) \leqslant |f(\epsilon)| < \epsilon'. \tag{12.14}$$

Donc $a = 0$ par le lemme 1.441(3). □

Les opérations simples passent à la limite, sauf la division pour laquelle il faut faire attention au dénominateur. La proposition suivante est en grande partie un cas particulier de la proposition 10.241.

PropOpsSimplesLimites

Proposition 12.6.
Soient deux fonctions $f, g\colon \mathbb{R} \to \mathbb{R}$. Nous supposons que $\lim_{x\to a} f(x) = \alpha$ et $\lim_{x\to a} g(x) = \beta$. Alors

ITEMooOJUWooQpqqnQ

(1) La fonction $f + g$ a une limite $x \to a$ qui vaut $\lim_{x\to a}(f + g)(x) = \alpha + \beta$,

ITEMooNGWTooSDlEkp

(2) Pour tout $\lambda \in \mathbb{R}$, la fonction λf a une limite en a qui vaut $\lim_{x\to a}(\lambda f)(x) = \lambda\alpha$.

(3) La fonction fg a une limite en a, qui vaut $\lim_{x\to a}(fg)(x) = \alpha\beta$,

(4) si il existe un voisinage de a sur lequel g ne s'annule pas, et si $\beta \neq 0$, alors la fonction f/g a une limite en a qui vaut $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\alpha}{\beta}$.

Démonstration. En plusieurs points.

(1) **La somme** Soit $\epsilon > 0$. Soient $\delta_1 > 0$ et $\delta_2 > 0$ tels que

$$0 < |x - a| < \delta_1 \Rightarrow |f(x) - \alpha| < \epsilon \tag{12.15a}$$
$$0 < |x - a| < \delta_2 \Rightarrow |g(x) - \beta| < \epsilon. \tag{12.15b}$$

En prenant $\delta \leqslant \min\{\delta_1, \delta_2\}$, si $0 < |x - a| \leqslant \delta$ alors

$$|f(x) + g(x) - \alpha - \beta| \leqslant |f(x) - \alpha| + |g(x) - \beta| \leqslant 2\epsilon. \tag{12.16}$$

Cela étant valable pour tout ϵ, nous avons $|f(x) + g(x) - \alpha - \beta| = 0$ par le lemme 1.441(3).

(2) **Produit par un scalaire**
Soit $\epsilon > 0$. Nous devons trouver un δ tel que pour tout x dans $[a-\delta, a+\delta]$, on ait $|(\lambda f)(x) - \lambda\alpha| \leqslant \epsilon$. Cette dernière inégalité est équivalente à $|\lambda||f(x)-\alpha| \leqslant \epsilon$. Nous devons donc trouver un δ tel que

$$|f(x) - \alpha| \leqslant \frac{\epsilon}{|\lambda|}. \tag{12.17}$$

soit vraie pour tout x dans $[a - \delta, a + \delta]$. Mais l'hypothèse $\lim_{x\to a} f(x) = \alpha$ (prise avec $\epsilon/|\lambda|$ en guise de ϵ) dit précisément qu'il existe un δ tel que pour tout x dans $[a - \delta, a + \delta]$ on ait cette inégalité.

(3) **Le produit** Nous prenons ϵ, δ_1 et δ_2 dans les mêmes conditions que pour la somme. Nous majorons comme ceci :

$$|f(x)g(x) - \alpha\beta| \leqslant |f(x)g(x) - \alpha g(x)| + |\alpha g(x) - \alpha\beta| \tag{12.18a}$$
$$= |g(x)||f(x) - \alpha| + |\alpha||g(x) - \beta| \tag{12.18b}$$
$$\leqslant \epsilon\big(|g(x)| + |\alpha|\big) \tag{12.18c}$$
$$\leqslant \epsilon\big(|\beta| + \epsilon + |\alpha|\big). \tag{12.18d}$$

La dernière ligne majore $|g(x)| \leqslant |\beta| + \epsilon$.
Vu que $\lim_{\epsilon\to 0} \epsilon\big(|\beta| + \epsilon + |\alpha|\big) = 0$, le lemme 12.5 dit que $f(x)g(x) - \alpha\beta = 0$, ce que nous voulions.

(4) **Le quotient** Soit $\epsilon > 0$. Nous considérons $\delta > 0$ tel que nous ayons simultanément $|f(x) - \alpha| < \epsilon$, $|g(x) - \beta| < \epsilon$ et $g(x) \neq 0$ pour tout $x \in B(a, \delta)$. Nous avons alors le calcul

$$|\frac{f(x)}{g(x)} - \frac{\alpha}{\beta}| \leqslant |\frac{f(x)}{g(x)} - \frac{\alpha}{g(x)}| + |\frac{\alpha}{g(x)} - \frac{\alpha}{\beta}| \tag{12.19a}$$

$$\leqslant |\frac{1}{g(x)}||f(x) - \alpha| + |\alpha||\frac{1}{g(x)} - \frac{1}{\beta}| \tag{12.19b}$$

SUBEQooOHFHooWZrwcA

Nous avons la majoration

$$\frac{1}{|g(x)|} \leqslant \frac{1}{|\beta| - \epsilon}, \tag{12.20}$$

et comme $\beta \neq 0$, nous pouvons supposer que ϵ est assez petit pour que le dénominateur ne s'annule pas. En ce qui concerne le second terme de (12.19b),

$$\frac{1}{g(x)} - \frac{1}{\beta} = \frac{\beta - g(x)}{g(x)\beta}. \tag{12.21}$$

Ici aussi nous pouvons majorer $1/g(x)$. Le numérateur se majore par ϵ. En remettant tout ça dans (12.19b), nous avons une majoration qui tend vers zéro lorsque $\epsilon \to 0$.

□

PROPooVLBWooVttvFK

Proposition 12.7.
La limite est linéaire : pour toutes fonctions f et g admettant une limite en a et pour tout réels λ et μ.

$$\lim_{x \to a} (\lambda f + \mu g)(x) = \lambda \lim_{x \to a} f(x) + \mu \lim_{x \to a} g(x). \tag{12.22}$$

Démonstration. Tout est dans la proposition 12.6. □

Le résultat suivant est pratique pour le calcul des limites.

PropChmVarLim

Proposition 12.8.
Quand la limite existe, nous avons

$$\lim_{x \to a} f(x) = \lim_{\epsilon \to 0} f(a + \epsilon),$$

ce qui correspond à un « changement de variables » dans la limite.

Démonstration. Si $A = \lim_{x \to a} f(x)$, par définition,

$$\forall \epsilon' > 0, \exists \delta \text{ tel que } |x - a| \leqslant \delta \Rightarrow |f(x) - A| \leqslant \epsilon'. \tag{12.23}$$

EqCondFaplusespLim

La seule subtilité de la démonstration est de remarquer que si $|x - a| \leqslant \delta$, alors x peut être écrit sous la forme $x = a + \epsilon$ pour un certain $|\epsilon| \leqslant \delta$. En remplaçant x par $a + \epsilon$ dans la condition 12.23, nous trouvons

$$\forall \epsilon' > 0, \exists \delta \text{ tel que } |\epsilon| \leqslant \delta \Rightarrow |f(a + \epsilon) - A| \leqslant \epsilon', \tag{12.24}$$

ce qui signifie exactement que $\lim_{\epsilon \to 0} f(a + \epsilon) = A$. □

Il y a une petite différence de point de vue entre $\lim_{x \to a} f(x)$ et $\lim_{\epsilon \to 0} f(a + \epsilon)$. Dans le premier cas, on considère $f(x)$, et on regarde ce qu'il se passe quand x se rapproche de a, tandis que dans le second, on considère $f(a)$, et on regarde ce qu'il se passe quand on s'éloigne un tout petit peu de a. Dans un cas, on s'approche très près de a, et dans l'autre on s'en éloigne un tout petit peu. Le contenu de la proposition 12.8 est de dire que ces deux points de vue sont équivalents.

PROPooOUPNooTrClHw

Proposition 12.9 ([326])**.**
Soient des fonctions $f, g : \mathbb{R} \to \mathbb{R}$ telles que $\lim_{x \to a} f(x) = \ell$ et $\lim_{x \to a} g(x) = \ell' \neq 0$. Alors

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\ell}{\ell'}. \tag{12.25}$$

Démonstration. Nous avons :

$$\left|\frac{f(x)}{g(x)}-\frac{\ell}{\ell'}\right|=\frac{|\ell' f(x)-g(x)\ell|}{|g(x)\ell'|}. \tag{12.26}$$

Soit s, un minorant de $|g(x)|$ sur un voisinage de a ; puisque la limite en a est $\ell'\neq 0$, nous pouvons prendre par exemple $s=|\ell'|/2$: $|g(x)|>|\ell'|/2$ sur $B(a,\delta)$ dès que δ est assez petit. Nous considérons $x\in B(a,\delta)$. Avec cela nous avons :

$$\begin{aligned}\left|\frac{f(x)}{g(x)}-\frac{\ell}{\ell'}\right|&=\frac{|\ell' f(x)-g(x)\ell|}{|g(x)\ell'|} &(12.27a)\\ &\leqslant\frac{2}{|\ell'|}\left(\frac{|\ell' f(x)-g(x)\ell|}{|\ell'|}\right) &(12.27b)\\ &\leqslant\frac{2}{|\ell'|^2}\big(|\ell' f(x)-\ell\ell'|+|\ell\ell'-g(x)\ell|\big) &(12.27c)\\ &=\frac{2}{|\ell'|^2}\big(|\ell'||f(x)-\ell|+|\ell||\ell'-g(x)|\big). &(12.27d)\end{aligned}$$

Nous avons utilisé la règle du produit de valeurs absolues du lemme 1.383(7).

Soient $\epsilon>0$ et δ tel que $|f(x)-\ell|<\epsilon$ et $|g(x)-\ell'|<\epsilon$ pour tout $x\in B(a,\delta)$. Avec cela nous avons

$$\left|\frac{f(x)}{g(x)}-\frac{\ell}{\ell'}\right|\leqslant\frac{2}{|\ell'|^2\big(|\ell'|+|\ell|\big)}\epsilon. \tag{12.28}$$

D'où la limite attendue. □

LemLimMajorableVois

Lemme 12.10.
Si $\lim_{x\to a} f(x)=b$ avec a, $b\in\mathbb{R}$, alors il existe un $\delta>0$ et un $M>0$ tels que

$$(|x-a|\leqslant\delta)\Rightarrow|f(x)|\leqslant M.$$

Ce que signifie ce lemme, c'est que quand la fonction f admet une limite finie en un point, alors il est possible de majorer la fonction sur un intervalle autour du point.

Démonstration. Cela va être démontré par l'absurde. Supposons qu'il n'existe pas de δ ni de M qui vérifient la condition. Dans ce cas, pour tout δ et pour tout M, il existe un x tel que $|x-a|\leqslant\delta$ et $|f(x)|>M$. Ceci est valable pour tout M, donc, prenons par exemple, $b+1000$. Ainsi

$$\forall\delta>0,\exists x \text{ tel que } |x-a|\leqslant\delta \text{ et } |f(x)|>b+1000. \tag{12.29}$$

Cela signifie qu'aucun δ ne peut convenir dans la définition de $\lim_{x\to a} f(x)=b$, ce qui contredit les hypothèses. □

Dans le même ordre d'idée, on peut prouver que si la limite de la fonction en un point est positive, alors elle est positive autour de ce point. Plus précisément, nous avons la proposition suivante.

PropoLimPosFPos

Proposition 12.11.
Si f est une fonction telle que $\lim_{x\to a} f(x)>0$, alors il existe un voisinage de a sur lequel f est positive.

Démonstration. Supposons que $\lim_{x\to a} f(x)=y_0$. Par la définition de la limite, nous avons pour tout x dans un voisinage autour de a, $|f(x)-a|<\epsilon$. Ceci est valable pour tout ϵ, pourvu que le voisinage soit assez petit. Si nous choisissons un voisinage pour lequel $|f(x)-a|<\frac{y_0}{2}$, alors, sur ce voisinage, f est positive. □

Tholimfgabab

Théorème 12.12.
Si

$$\lim_{x\to a} f(x) = b_1 \qquad \text{et} \qquad \lim_{x\to a} g(x) = b_2, \tag{12.30}$$

alors

$$\lim_{x\to a}(fg)(x) = b_1 b_2. \tag{12.31}$$

Démonstration. Soit $\epsilon > 0$, et tentons de trouver un δ tel que $|f(x)g(x) - b_1b_2| \leqslant \epsilon$ dès que $|x - a| \leqslant \delta$. Nous avons

$$\begin{aligned} |f(x)g(x) - b_1b_2| &= |f(x)g(x) - b_1b_2 + f(x)b_2 - f(x)b_2| \\ &= \big|f(x)\big(g(x) - b_2\big) + b_2\big(f(x) - b_1\big)\big| \\ &\leqslant \big|f(x)\big(g(x) - b_2\big)\big| + \big|b_2\big(f(x) - b_1\big)\big| \\ &= |f(x)||g(x) - b_2| + |b_2||f(x) - b_1|. \end{aligned} \tag{12.32}$$

EqfgbunbdeuxMin

Maintenant, nous savons par le lemme 12.10 que pour un certain δ_1, la quantité $|f(x)|$ peut être majorée par un certain M dès que $|x-a| \leqslant \delta_1$. Prenons donc un tel δ_1, et supposons que $|x-a| \leqslant \delta_1$. Nous savons aussi que pour n'importe quel choix de ϵ_2 et ϵ_3, il existe des nombres δ_2 et δ_3 tels que $|f(x) - b_1| \leqslant \epsilon_2$ et $|g(x) - b_1| \leqslant \epsilon_3$ dès que $|x - a| \leqslant \delta_2$ et $|x - a| \leqslant \delta_3$. Dans ces conditions, la dernière expression (12.32) se réduit à

$$|f(x)g(x) - b_1b_2| \leqslant M\epsilon_2 + |b_2|\epsilon_3. \tag{12.33}$$

Pour terminer la preuve, il suffit de choisir ϵ_2 et ϵ_3 tels que $M\epsilon_2 + |b_2|\epsilon_3 \leqslant \epsilon$, et puis prendre $\delta = \min\{\delta_1, \delta_2, \delta_3\}$.

Remettons les choses dans l'ordre. On se donne un ϵ au départ. La première chose est de trouver un δ_1 qui permette de majorer $|f(x)|$ par M selon le lemme 12.10, et puis, choisissons ϵ_2 et ϵ_3 tels que $M\epsilon_2 + |b_2|\epsilon_3 \leqslant \epsilon$. Ensuite, nous prenons, en vertu des hypothèses de limites pour f et g, les nombres δ_2 et δ_3 tels que $|f(x) - b_1| \leqslant \epsilon_2$ et $|g(x) - b_2| \leqslant \epsilon_3$ dès que $|x - a| \leqslant \delta_2$ et $|x - a| \leqslant \delta_3$.

Si avec tout ça on prend $\delta = \min\{\delta_1, \delta_2, \delta_3\}$, alors la majoration et les deux inégalités sont valables en même temps, et au final

$$|f(x)g(x) - b_1b_2| \leqslant M\epsilon_2 + |b_2|\epsilon_3 \leqslant \epsilon,$$

ce qu'il fallait prouver. □

À l'aide de ces petits résultats, nous pouvons déjà calculer pas mal de limites. Nous pouvons déjà par exemple, calculer les limites de tous les polynômes en tous les nombres réels. En effet, nous connaissons la limite de la fonction $f(x) = x$. La fonction $x \mapsto x^2$ n'est rien d'autre que le produit de f par elle-même. Donc

$$\lim_{x\to a} x^2 = \big(\lim_{x\to a} x\big) \cdot \big(\lim_{x\to a} x\big) = a^2.$$

De la même façon, nous trouvons facilement que

$$\lim_{x\to a} x^n = a^n. \tag{12.34}$$

LEMooLSJZooDauTkc

Lemme 12.13.
Soit $t \in \mathbb{R}$. Nous avons

$$\lim_{x\to\infty} \frac{1}{x} \operatorname{int}(xt) = t \tag{12.35}$$

où int *est la fonction partie entière définie en 1.476.*

Démonstration. En vertu du lemme 1.476, nous avons $y = \text{int}(y) + \text{frac}(y)$ avec $\text{frac}(y) \in [0,1[$. Nous avons donc

$$\frac{1}{x}\text{int}(xt) = \frac{1}{x}\big(xt - \text{frac}(xt)\big) \tag{12.36a}$$
$$= \frac{xt}{x} + \frac{\text{frac}(xt)}{x} \tag{12.36b}$$
$$= t + \frac{\text{frac}(xt)}{x}. \tag{12.36c}$$

Puisque $\text{frac}(xt) \in [0,1[$, nous avons

$$\lim_{x\to\infty} \frac{\text{frac}(xt)}{x} = 0. \tag{12.37}$$

D'où la limite demandée. □

12.2 Limites pointées et épointées

SECooNJSGooGSAtdV

La limite d'une fonction en un point a déjà été définie en 7.110. Nous introduisons maintenant une notion très ressemblante.

DEFooBAPHooUtIaRS

Définition 12.14 (limite pointée[363])**.**
*Soient X et Y deux espaces topologiques, A une partie de X, f une application de A dans Y, a un point de X adhérent à A et ℓ un point de Y. On dit que ℓ est une **limite pointée** de f au point a si pour tout voisinage V de ℓ, il existe un voisinage W de a tel que pour tout point x de $W \cap A$, l'image $f(x)$ appartient à V.*

Quelques notations et vocabulaire.

(1) Nous allons limiter notre discussion au cas des fonctions $\mathbb{R} \to \mathbb{R}$.
(2) La limite de la définition 7.110 sera provisoirement nommée **limite épointée**, pour ne pas causer de confusion.
(3) Pour bien distinguer la limite pointée de la limite épointée, nous allons noter $LP_{x\to a}f(x)$ pour la limite pointée et $LE_{x\to a}f(x)$ pour la limite épointée.
(4) Nous allons utiliser la caractérisation 9.83 de la continuité de f en un point.
(5) Nous allons utiliser la caractérisation 12.1 de la limite épointée.

LEMooWAZLooDPvemu

Lemme 12.15.
Si une fonction $f\colon \mathbb{R} \to \mathbb{R}$ vérifie $LP_{x\to a}f(x) = \ell$, alors elle vérifie $LE_{x\to a}f(x) = \ell$.

Démonstration. Soit $\epsilon > 0$. L'existence de la limite pointée dit qu'il existe $\delta > 0$ tel que $|x-a| < \delta$ implique $|f(x) - \ell| < \epsilon$. À fortiori, si $0 < |x-a| < \delta$ nous avons aussi $|f(x) - \ell| < \epsilon$. Donc la limite épointée, par la caractérisation 12.1. □

La réciproque n'est pas vraie, comme le montre le lemme suivant.

LEMooOSNGooJpiXbK

Lemme 12.16 ([364])**.**
Soit la fonction

$$f\colon \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \begin{cases} 0 & si\ x \neq 0 \\ 1 & si\ x = 0. \end{cases} \tag{12.38}$$

Nous avons

ITEMooNRNCooFhbZwB

(1) $LE_{x\to 0}f(x) = 0$.

ITEMooUSWMooMNPMCT

(2) La limite pointée de f en 0 n'existe pas.

Démonstration. En deux parties.

(1) **Pour (1)** Pour tout $x \in B(0,\delta)\backslash\{0\}$ nous avons $f(x) = 0$. Donc la limite épointée suit.

(2) **Pour (2)** Soit $\delta > 0$. Le lemme 1.441 nous assure qu'il existe $x \in]-\delta, 0[$. Ce x est dans $B(0,\delta)$ et vérifie $f(x) = 0$. Nous avons aussi $x = 0$ dans la même boule. Donc $f\big(B(0,\delta)\big)$ contient au moins les nombres 1 et 0.
Il n'existe donc pas de ℓ tel que tout $f\big(B(0,\delta)\big)$ soit dans $B(\ell,\epsilon)$.

□
LEMooTPNEooRurTJJ

Lemme 12.17.
Soit une fonction $f\colon \mathbb{R}\backslash\{a\} \to \mathbb{R}$. Alors la limite pointée de f en a existe si et seulement si la limite épointée existe. Dans ce cas, elles sont égales.

Démonstration. En trois parties. Nous notons $\Omega = \mathbb{R}\backslash\{a\}$ le domaine de f.

(1) **Si la limite pointée existe, alors l'épointée existe** Soit $\epsilon > 0$. L'existence de la limite pointée dit qu'il existe $\delta > 0$ tel que $f\big(B(a,\delta) \cap \Omega\big) \subset B(\ell,\epsilon)$. Comme $B(a,\delta) \cap \Omega = B(a,\delta)\backslash\{a\}$, nous avons la condition épointée.

(2) **Si la limite épointée existe, alors la pointée existe** Du même tonneau.

(3) **Égalité** Le jeu des boules du premier point prouve l'égalité au passage.

□

Que faut-il donc ajouter à la limite épointé pour obtenir une limite pointée ? Réponse dans le lemme suivant.

Lemme 12.18.
Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$. Nous avons

$$LP_{x\to a}f(x) = \ell \tag{12.39}$$

si et seulement si les deux conditions suivantes sont réunies :

(1) $LE_{x\to a}f(x) = \ell$

(2) Soit f n'existe pas en a, soit f est continue en a.

Démonstration. En deux parties.

(1) $\Rightarrow$ L'existence d'une limite pointée implique celle de la limite épointée, et l'égalité entre les deux. Donc $LE_{x\to a}f(x) = \ell$. Supposons que f existe en a. Soit $\epsilon > 0$. L'existence d'une limite pointée signifie qu'il existe $\delta > 0$ tel que $|x-a| < \delta$ implique $|f(x) - \ell| < \epsilon$.
En particulier avec $x = a$ nous avons $|f(a) - \ell| < \epsilon$ pour tout ϵ. Donc $f(a) = \ell$.

(2) $\Leftarrow$ Nous supposons que $LE_{x\to a}f(x) = \ell$. Si f n'existe pas en a, alors les limites pointée et épointée coïncident [1]. Si par contre f existe en a et $f(a) = \ell$, alors nous travaillons.
Soit $\epsilon > 0$. Il existe $\delta > 0$ tel que $x \in B(a,\delta)\backslash\{a\}$ implique $|f(x) - \ell| < \epsilon$. Mais si $x = a$ nous avons $|f(x) - \ell| = 0 < \epsilon$. Au final nous avons $|f(x) - \ell| < \epsilon$ pour tout $x \in B(a,\delta)$. Donc $LP_{x\to a}f(x) = \ell$.

□
LEMooNEGOooCllIMN

Lemme 12.19 (limite et continuité)**.**
Supposons que $LE_{x\to a}f(x) = \ell$ et que f est continue en a. Alors $f(a) = \ell$.

Démonstration. Soit $\epsilon > 0$. L'hypothèse sur la limite dit qu'il existe $\delta_1 > 0$ tel que $0 < |x-a| < \delta_1$ implique $|f(x) - \ell| < \epsilon$.

L'hypothèse de continuité, avec la caractérisation 9.83, dit qu'il existe $\delta_2 > 0$ tel que $|x-a| < \delta_2$ implique $|f(x) - f(a)| < \epsilon$.

1. Lemme 12.17.

Nous considérons un $\delta > 0$ plus petit que δ_1 et que δ_2. Soit aussi un x satisfaisant $0 < |x-a| < \delta$. Nous avons

$$|f(a) - \ell| \leqslant |f(a) - f(x)| + |f(x) - \ell| \leqslant 2\epsilon. \tag{12.40}$$

Puisque cela est valide pour tout x, nous déduisons que $f(a) = \ell$. □

12.2.1 Théorèmes de composition de limites

La proposition suivante formalise le fait que la limite pointée soit stable par composition.

THOooOYXDooKDPkuW

Théorème 12.20.
Soient des fonctions $g\colon \mathbb{R} \to \mathbb{R}$ et $f\colon \Omega \to \mathbb{R}$ avec $\Omega \subset \mathbb{R}$ telles que

(1) $LP_{x\to a} g(x) = \ell$

(2) $LP_{y\to \ell} f(y) = b$.

(3) $g(\mathbb{R}) \subset \Omega$

Alors

$$LP_{x\to a}(f \circ g)(x) = b. \tag{12.41}$$

Démonstration. Soit $\epsilon > 0$. L'hypothèse de limite pour f donne $\eta > 0$ tel que

EQooLWGIooLqKThv

$$|y - \ell| < \eta \Rightarrow |f(y) - b| < \epsilon. \tag{12.42}$$

Soit $\delta > 0$ tel que $|x - a| < \delta$ implique $|g(x) - \ell| < \eta$.

Avec tout ça, si $|x - a| < \delta$ nous avons $|g(x) - \ell| < \eta$ et en appliquant l'implication (12.42) à $y = g(x)$ nous trouvons $|f\big(g(x)\big) - b| < \epsilon$. □

Le théorème suivant, qui traite de la composition de limites pointées, montre que la limite épointée ne passe pas gratuitement à la composition.

THOooNPBQooEMOYpd

Théorème 12.21.
Soient des fonctions $g\colon \mathbb{R} \to \mathbb{R}$ et $f\colon \Omega \to \mathbb{R}$ avec $\Omega \subset \mathbb{R}$ telles que

(1) $LE_{x\to a} g(x) = \ell$

(2) $LE_{y\to \ell} f(y) = b$.

(3) $g(\mathbb{R}) \subset \Omega$

ITEMooUNJAooCDOKcO

(4) Soit $\ell \notin \Omega$, soit f est continue en ℓ.

Alors

EQooTHTVooCyrFdN

$$LE_{x\to a}(f \circ g)(x) = b. \tag{12.43}$$

Démonstration. Nous notons Ω le domaine de f. Ce sera $\mathbb{R}$ ou $\mathbb{R}\backslash\{\ell\}$ selon le cas traité dans (4).

Soit $\epsilon > 0$. L'hypothèse de limite épointée pour f nous dit qu'il existe $\eta > 0$ tel que

$$y \in \Omega \cap B(\ell, \eta)\backslash\{\ell\} \tag{12.44}$$

implique $|f(y) - b| < \epsilon$.

L'hypothèse de limite épointée pour g, appliquée à ce η dit qu'il existe $\delta > 0$ tel que $0 < |x - a| < \delta$ implique $|g(x) - \ell| < \eta$.

(1) **Si f n'existe pas en ℓ** Supposons avoir $0 < |x - a| < \eta$. Alors nous avons $|g(x) - \ell| < \delta$. Notons qu'il est impossible d'avoir $g(x) = \ell$ parce que nous avons supposé $g(\mathbb{R}) \subset \Omega$ et que ℓ n'est pas dans Ω.

Nous avons quand même $0 < |g(x) - \ell| < \delta$. La condition de limite pour f appliquée à $y = g(x)$ donne alors $|f\big(g(x)\big) - b| < \epsilon$, ce qui signifie la limite épointée (12.43).

(2) **Si f est continue en ℓ** Le lemme 12.19 dit qu'alors $f(\ell) = b$. L'hypothèse de limite épointée sur f dit que

$$0 < |y - \ell| < \eta \Rightarrow |f(y) - b| < \epsilon. \tag{12.45}$$

Mais puisque $f(\ell) = b$, nous avons en réalité

$$|y - \ell| < \eta \Rightarrow |f(y) - b| < \epsilon. \tag{12.46}$$

EQooHAHHooRiAABt

Supposons donc que $0 < |x - a| < \delta$. Alors $|g(x) - \ell| < \eta$. En appliquant (12.46) à $y = g(x)$ nous trouvons

$$|f(y) - b| < \epsilon. \tag{12.47}$$

□

NORMooSLAJooLfDreV

12.22.
Les deux théorèmes sont incomplets. En effet, le théorème « pointé » 12.20 ne traite pas le cas où seules des limite épointées existent. Il est donc moins général que le théorème « épointé » 12.21. En contrepartie, le théorème « épointé » ne parvient pas à conclure à l'existence d'une limite pointée dans le cas où elle existe. Sa conclusion est donc moins forte.

Nous devons donc nous atteler à écrire un théorème qui traite tous les cas en obtenant la conclusion la plus forte possible dans chaque cas. Nous allons supposer que

$$\begin{cases} LE_{x\to a} g(x) = \ell & \text{(12.48a)} \\ LE_{y\to \ell} f(y) = b. & \text{(12.48b)} \end{cases}$$

Ensuite, les différents cas seront divisés selon quatre critères :

(1) g existe en a ou pas.

(2) f existe en ℓ ou pas.

(3) g est continue ou pas en a.

(4) f est continue ou pas en ℓ.

Cela fait $2^4 = 16$ combinaisons. Heureusement certaines sont impossibles : si une fonction n'existe pas en un point, elle ne peut pas y être continue.

Nous avons donc 9 cas résumés dans le théorème suivant.

$g(a)$	1	1	1	1	1	1	0	0	0
$f(\ell)$	1	1	1	1	0	0	1	1	0
$g \in C^0$	1	1	0	0	1	0	0	0	0
$f \in C^0$	1	0	1	0	0	0	1	0	0

(12.49)

THOooHXGIooBclAHA

Théorème 12.23 ([1])**.**
Soient des fonctions f et g telles que

$$\begin{cases} LE_{x\to a} g(x) = \ell & \text{(12.50a)} \\ LE_{y\to \ell} f(y) = b. & \text{(12.50b)} \end{cases}$$

ITEMooDXBLooVfhSWg

(1) Si f est continue en ℓ et si g est continue en a, alors $f \circ g$ est continue en a.

ITEMooIXBQooMDknwN

(2) Nous supposons :

(2a) g définie en a
(2b) f définie en ℓ
(2c) g continue en a
(2d) f non continue en ℓ.

Alors nous ne disons rien.

ITEMooHTIEooMKDrqx

(3) Nous supposons :

(3a) g *définie en* a *(3c)* g *non continue en* a
(3b) f *définie en* ℓ *(3d)* f *continue en* ℓ.

Alors $LE_{x\to a}(f\circ g)(x)=b$.

ITEMooVQMDooEtHfwC

(4) Nous supposons :

(4a) g *définie en* a *(4c)* g *non continue en* a
(4b) f *définie en* ℓ *(4d)* f *non continue en* ℓ.

Alors nous ne disons rien.

ITEMooANFQooWVrfTd

(5) Nous supposons :

(5a) g *définie en* a *(5c)* g *continue en* a
(5b) f *non définie en* ℓ *(5d)* f *non continue en* ℓ.

Alors $LE_{x\to a}(f\circ g)(x)=b$.

ITEMooDJBHooSlqpOO

(6) Nous supposons :

(6a) g *définie en* a *(6c)* g *non continue en* a
(6b) f *non définie en* ℓ *(6d)* f *non continue en* ℓ.

Alors $LE_{x\to a}(f\circ g)(x)=b$.

ITEMooUFJHooRzLglZ

(7) Nous supposons :

(7a) g *non définie en* a *(7c)* g *non continue en* a
(7b) f *définie en* ℓ *(7d)* f *continue en* ℓ.

Alors $LE_{x\to a}(f\circ g)(x)=b$.

ITEMooOAAVooSjoYOv

(8) Nous supposons :

(8a) g *non définie en* a *(8c)* g *non continue en* a
(8b) f *définie en* ℓ *(8d)* f *non continue en* ℓ.

Alors nous ne disons rien.

ITEMooPVZKooBXJARI

(9) Nous supposons :

(9a) g *non définie en* a *(9c)* g *non continue en* a
(9b) f *non définie en* ℓ *(9d)* f *non continue en* ℓ.

Alors $LE_{x\to a}(f\circ g)(x)=b$.

Démonstration. Cas par cas.

(1) **Cas (1)** C'est le théorème 9.89 de composition de la continuité.
(2) **Cas (2)** l'exemple du lemme 12.16.
(3) **Cas (3)** Théorème 12.21.
(4) **Cas (4)** Le contre-exemple dans ce cas est $g=\mathbb{1}_0$ et $f=\mathbb{1}_0$.
(5) **Cas (5)** Théorème 12.21.
(6) **Cas (6)** Théorème 12.21.
(7) **Cas (7)** Théorème 12.21.

(8) **Cas (8)** Contre-exemple, un peu artificiel, avec $g(x) = \frac{\mathbb{1}_0(x)}{x}$. C'est une fonction qui vaut 0 partout sauf en 0 où elle n'existe pas. Ensuite pour f, nous prenons l'indicatrice de $\{0\}$: $f = \mathbb{1}_0$. Pour tout $x \neq 0$ nous avons

$$(f \circ g)(x) = \mathbb{1}_0\left(\frac{\mathbb{1}_0(x)}{x}\right) = \mathbb{1}_0(0) = 1. \tag{12.51}$$

Donc $LE_{x\to 0}(f \circ g)(x) = 1$.
Mais nous avons pourtant

$$\begin{cases} LE_{x\to 0} g(x) = 0 & \text{(12.52a)} \\ LE_{y\to 0} f(x) = 0. & \text{(12.52b)} \end{cases}$$

(9) **Cas (9)** Théorème 12.21.

□

12.2.2 Discussion pointée Vs épointée

Résumé :

(1) Dans l'éducation nationale et dans les programmes en France, c'est la limite pointée qui est donnée.

(2) Dans le Frido ce sera la limite épointée. Autrement dit, nous réserverons la notation lim et le mot « limite » pour la limite épointée.

(3) De toutes façons, ça ne change pratiquement rien nulle part. Vous pourriez terminer l'agrégation sans vous en rendre compte. Vous pouvez sauter toute la discussion et reprendre une vie normale.

Le débat pour savoir quelle est la « meilleure » notion a déjà fait couler de nombreux octets[365, 366, 364, 367, 368].

12.2.2.1 La limite pointée est plus simple au départ

Il est vrai que la limite pointée est plus simple de premier abord.

12.2.2.2 Le théorème de composition

Le théorème de composition des limites pointées 12.20 est plus propre que le théorème de composition épointé 12.21. Intuitivement, on voudrait que la limite d'une fonction composée soit la composée des limites. Et ça c'est vrai pour la limite pointée, pas pour l'épointée.

12.2.2.3 Limite d'une fonction discontinue mais qui existe

Que pensez-vous que la limite en l'infini de la fonction suivante « devrait » être ?

$$\begin{aligned} f\colon [0,\infty] &\to \mathbb{R} \\ x &\mapsto \begin{cases} x^2 & \text{si } x \neq \infty \\ 0 & \text{si } x = \infty \end{cases} \end{aligned} \tag{12.53}$$

Là, tant que x s'approche de ∞ sans l'atteindre, il n'y a vraiment que de la croissance à perte de vue ; jusqu'à l'horizon et au-delà.

On pourrait faire la même remarque avec la fonction indicatrice de $\{0\}$. Qu'est-ce que la limite en zéro « devrait » être ?

12.2.2.4 Point d'étape

Aucune des deux limites ne donne le résultat « attendu » dans les deux cas. Toutes deux font une chose bien, et une chose pas bien.

12.2.2.5 Limite d'une fonction discontinue mais qui existe

Prenons la fonction indicatrice de $\{0\}$:

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{12.54}$$

En utilisant la limite pointée, on peut exprimer deux choses :

— la limite pointée en zéro n'existe pas

— la fonction n'est pas continue en zéro.

En utilisant la limite épointée, on peut exprimer deux choses :

— la limite épointée en zéro existe et vaut zéro.

— la fonction n'est pas continue en zéro.

La première paire d'informations est compatible avec la fonction $1/x$. Autrement dit, la valeur « inattendue » que prend f en zéro casse tout ce qu'on aurait pu dire sur un voisinage de zéro.

La seconde paire d'informations donne au moins une idée de ce qu'il se passe autour de zéro. On peut en déduire, par exemple, que f est intégrable sur un voisinage de zéro parce qu'elle y est bornée et que la valeur en un point ne fait rien à l'intégrabilité [2].

12.2.2.6 L'enseignement du cas précédent

La limite épointée donne une information sur ce qu'il se passe « autour » du point sans rien dire de ce qu'il se passe « sur » le point. Si nécessaire, la continuité complète l'information en précisant ce qu'il se passe « sur » le point.

Certaines questions n'ont pas besoin de savoir ce qu'il se passe en un seul point.

La limite épointée « refuse » de dire ce qu'il se passe autour du point parce qu'il y a un problème juste sur ledit point. Un seul point se comporte mal et tout le voisinage passe sous le radar.

12.2.2.7 La limite épointée est plus riche

La classe des fonction admettant une limite pointée est plus grande que celle admettant une limite épointée (lemme 12.15). L'utilisation de la limite épointée permet de décrire quelques cas supplémentaires par rapport à ce que l'on peut faire seulement avec la limite pointée.

Pour être plus précis, comme je le disait précédemment, en 12.22, aucune des deux notions n'est satisfaisante seule :

— mettez de la limite pointée dans les hypothèses, vous aurez un théorème moins général ;

— mettez de la limite pointée dans la thèse, vous aurez un résultat plus fort.

Le vrai intérêt de la limite épointée est que *en combinaison avec la notion de continuité* permet d'être plus général et plus précis que ce qu'on peut obtenir avec la limite pointée. Dit autrement, le couple (limite épointée, continuité) est plus fort que le couple (limite pointée, continuité).

D'un certain point de vue, oui, la limite pointée est plus simple, mais elle est plus simple parce qu'elle donne moins d'informations.

12.2.2.8 Retour sur le théorème de composition

Le *vrai* théorème de composition est le théorème 12.23. Lui, il passe en revue tous les cas possibles et donne le plus de conclusions possibles dans chaque cas.

Ce théorème s'exprime de façon à peu près convenable à l'aide de limite épointée et de continuité. J'attends de voir le même avec une limite pointée et la continuité.

2. Peut-être qu'il faut ajouter que f est mesurable ? Mais peut-être que l'existence de la limite implique la mesurabilité ? Dites-moi ce que vous en pensez.

Je suis très ouvert à la discussion si c'est pour avoir quelque chose de plus simple produisant les mêmes résultats. Je ne suis par contre pas très ouvert pour avoir quelque chose de plus simple, mais donnant moins de résultats. C'est toujours facile d'avoir des résultats plus courts, plus simples et plus intuitifs quand on se contente de moins.

12.2.2.9 En français

La limite épointée rend l'idée de « s'approcher sans atteindre ». En français l'expression « être à la limite de tel résultat » signifie le plus souvent être très proche du résultat, mais ne pas y être.

12.2.2.10 La question est pédagogique

Tant qu'on ne m'a pas montré comment on exprime le théorème de composition 12.23 avec des limites pointées, je resterai sur cette idée : la limite pointée est plus simple, mais elle dit moins.

Cela n'est cependant pas spécialement bloquant. Après tout, ça dépend de ce qu'on veut. D'un point de vue pédagogique, la limite pointée introduit autant de ϵ et de δ qu'on le veut, et permet d'introduire tous les concepts utiles en analyse.

La question est de savoir à quel point on est prêt à se compliquer la vie pour avoir des théorèmes un micro-cheveu plus complets. Le choix du Frido est de recevoir la difficulté avec résignation et de l'endurer avec courage, pour le plaisir d'avoir des théorèmes qui donnent un peu plus d'information [3].

12.2.2.11 En fait ça ne change presque rien

Certains craignent qu'utiliser la limite pointée demande d'ajuster beaucoup de résultats un peu partout[369]. Voici la liste des résultats du Frido qui doivent être modifiés pour être exprimés à l'aide de limite pointée.

— Le fameux théorème de composition 12.21,

— La limite à gauche et à droite implique la limite, proposition 9.112.

— Le lemme 12.215.

Le fait est que l'on ne calcule presque jamais de limites en une valeur où la fonction existe. Si on calcule une limite, c'est précisément parce qu'on regarde un point où la fonction n'existe pas.

Exemples :

— Quand on calcule une dérivée, on calcule

$$\lim_{\epsilon\to 0} \frac{f(a+\epsilon)-f(a)}{\epsilon}. \tag{12.55}$$

Cette fonction de ϵ n'existe pas lorsque $\epsilon = 0$. Donc les limites pointées et épointées sont identiques.

— De même, l'étude du sinus cardinal $f(x) = \sin(x)/x$ (lemme 20.211) est une fonction dont ça ne viendrait à l'idée de personne de calculer la limite pour $x \to 4$. Et ça tombe bien : la seule limite que ça donne envie de calculer est

$$\lim_{x\to 0} \frac{\sin(x)}{x}. \tag{12.56}$$

Et encore une fois, la fonction dans le limite n'existe pas au point limite.

— Beaucoup de conditions d'intégrabilité demandent des limites à l'infini. Là encore, ce sont des limites vers des points où la fonction n'existe pas. Franchement, qui va vouloir définir

$$\begin{aligned} f\colon [0,\infty] &\to \mathbb{R} \\ x &\mapsto \begin{cases} x^2 & \text{si } x \neq \infty \\ 0 & \text{si } x = \infty \end{cases} \end{aligned} \tag{12.57}$$

3. C'est une de mes citation préférées. Comme nous sommes entre adultes, je vous donne la référence : [369]. Si vous n'avez pas 18 ans, on peut vraiment se demander si le Frido est vraiment une lecture de votre âge.

sans rigoler ? OK. Pour cette fonction, il y a une différence entre la limite pointée et épointée. Mais franchement, c'est bien la limite épointée qui donne le résultat « intuitif ».

12.2.2.12 Si vous avez quand même envie de discuter

Essayez de garder votre salive pour des sujets importants.

— L'intégrale de Kurzweil-Henstock[370] contre Lebesgue. Là au moins, il y a des choses non triviales à dire, et des vrais résultats mathématiques à la clef.

— L'utilisation de τ au lieu de π.

12.2.2.13 En très résumé

Si vous ne voulez pas lire toute ma prose, voici ce dont vous devez être conscient :

(1) La limite épointée est celle utilisée partout sauf en France.

(2) La limite épointée est un peu plus compliquée que la limite pointée, mais elle permet de prouver plus de choses. En témoigne le théorème « complet » de composition 12.23 que je doute être facile à exprimer à l'aide des limites pointées et de la continuité [4].

(3) La limite pointée « cache » l'information sur tout le voisinage de a si la fonction se comporte mal juste en a.

Une fois que vous êtes conscients de ces quelque points, vous faites comme vous voulez ; ça n'a pratiquement aucune importance. La seule position indéfendable est celle de prendre la limite pointée et de ne pas prévenir la lectrice que les sources autres que françaises donnent une définition différente.

12.3 Limites en l'infini

Non, sur $\mathbb{R}$ nous n'allons pas ajouter ∞ avec la topologie d'Alexandrov de la définition 7.106. Nous n'allons pas considérer $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$.

DEFooRUyiBSUooALDDOa

Définition 12.24 (Droite réelle achevée[372])**.**
Nous considérons l'ensemble

$$\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\} \tag{12.58}$$

où $+\infty$ et $-\infty$ ne sont pas des éléments de $\mathbb{R}$.

Nous mettons sur $\bar{\mathbb{R}}$ la relation d'ordre en prenant celle de $\mathbb{R}$ à laquelle nous ajoutons les règles

(1) $-\infty < x$ pour tout $x \in \mathbb{R} \cup \{+\infty\}$

(2) $+\infty > x$ pour tout $x \in \mathbb{R} \cup \{-\infty\}$.

Nous mettons une topologie sur $\bar{\mathbb{R}}$ en donnant la base[5] suivante :

— $]a, b[$,

— $]a, +\infty]$,

— $[-\infty, b[$

pour tous réels a et b.

12.25.
La notation « ∞ » peut désigner l'unique élément ajouté dans du compactifié d'Alexandrov [6] aussi bien que l'élément infini positif dans la droite réelle achevée. Si vous faites attention au contexte [7], ça ne devrait pas poser de problèmes.

Certaines autrices réservent « ∞ » pour Alexandrov et écrivent toujours « $+\infty$ » et « $-\infty$ » pour la droite réelle achevée.

4. Il y a bien entendu moyen. Voir par exemple [371]. Sans ironie, je trouve ce théorème fascinant.
5. Base de topologie, définition 7.2.
6. Le compactifié d'Alexandrov $\hat{\mathbb{R}}$, définition 7.106.
7. Vous devez toujours avoir parfaitement clairement en tête la topologie que vous manipulez.

Lemme 12.26 ([1]).
La topologie sur $\mathbb{R}$ induite de celle sur $\bar{\mathbb{R}}$ est la topologie usuelle.

Démonstration. Nous notons $\tau_{\mathbb{R}}$ la topologie de $\mathbb{R}$, $\tau_{\bar{\mathbb{R}}}$ celle de $\bar{\mathbb{R}}$ et τ_i celle induite de $\bar{\mathbb{R}}$ sur $\mathbb{R}$. Nous devons prouver que $\tau_i = \tau_{\mathbb{R}}$.

(1) $\tau_i \subset \tau_{\mathbb{R}}$ Un élément de τ_i est de la forme $\mathcal{O} = \mathbb{R} \cap A$ où A est un élément de $\tau_{\bar{\mathbb{R}}}$. Vu que A est un ouvert de $\bar{\mathbb{R}}$, il est une réunion d'éléments de la base de topologie [8] ; donc $A = \bigcup_{i\in I} A_i$ où les A_i sont des trois types listés dans la définition 12.24.

(1a) Si $A_i =]a, b[$ alors $\mathbb{R} \cap A =]a, b[$ est un ouvert de $\mathbb{R}$.

(1b) Si $A_i =]a, +\infty]$, alors $\mathbb{R} \cap A_i =]a, +\infty[$ est un ouvert de $\mathbb{R}$.

(1c) Si $A_i = [-\infty, b[$, même chose.

Donc $\mathbb{R} \cap A = \bigcup_{i\in I}(\mathbb{R} \cap A_i)$ est une union d'ouverts de $\mathbb{R}$.

(2) $\tau_{\mathbb{R}} \subset \tau_i$ Comme les $]a, b[$ forment une base de topologie de $\mathbb{R}$, l'ensemble τ_i contient une base de topologie de $\mathbb{R}$ et donc contient tout $\tau_{\mathbb{R}}$.

□

Proposition 12.27 ([1]).
Soit une suite (x_k) dans $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Nous avons $x_k \xrightarrow{\bar{\mathbb{R}}} +\infty$ si et seulement si pour tout $M > 0$ il existe un $N > 0$ tel que $n \geqslant N$ implique $x_n > M$.

Démonstration. En deux parties.

(1) $\Rightarrow$ Pour tout voisinage A de $+\infty$, il existe un N tel que $n \geqslant N$ implique $x_n \in A$. Soit donc le voisinage $]M, +\infty]$, et le N correspondant. Nous avons alors, pour tout $n \geqslant N$, $x_n \in]M, +\infty]$ et donc $x_n \geqslant M$.

(2) $\Leftarrow$ Soit un ouvert A contenant $+\infty$. Nous avons $A = \bigcup_{i\in I} A_i$ où les A_i sont des trois types listés dans la définition 12.24. Comme $+\infty \in A$, pour au moins un des i, nous avons $A_i =]a, +\infty[$.
Prenons N tel que $n \geqslant N$ implique $x_n > a$. Alors pour $n \geqslant N$ nous avons $x_n \in A$.

□
LEMooUBFAooEmquQQ

Lemme 12.28 ([1]).
Soient $x > 1$ dans $\mathbb{R}$ et $n \geqslant 1$ dans $\mathbb{N}$. Alors $x^n \geqslant x$.

Démonstration. Pour $n = 1$, nous avons $x^n = x$ donc d'accord. Supposons que $x^n \geqslant x$ pour un certain $n \in \mathbb{N}$, et prouvons que $x^{n+1} \geqslant x$.

Calcul avec justifications en-dessous :

$$\begin{aligned} x^{n+1} &= xx^n && \text{(12.59a)} \\ &\geqslant xx && \text{(12.59b)} \\ &\geqslant x. && \text{(12.59c)} \end{aligned}$$
SUBEQooBBJYooIYaErs
SUBEQooCEJWooMUMchE

Justifications :

— Pour (12.59b) hypothèse de récurrence.

— Pour (12.59c), lemme 1.434.

□
LEMooFCIXooJuHFqk

Lemme 12.29.
Nous considérons l'espace topologique de la droite réelle achevée [9] $\bar{\mathbb{R}}$. Si $n \geqslant 1$ nous avons

$$\lim_{x\to+\infty} x^n = +\infty \qquad \text{(12.60)}$$
EQooRRFEooLYcuRP

8. C'est la proposition 7.2 qui dit ça.
9. Définition 12.24.

et

$$\lim_{x\to+\infty}\frac{1}{x^n}=0. \tag{12.61}$$

Démonstration. Si V est un voisinage de $+\infty$, alors nous devons montrer qu'il existe un voisinage W de $+\infty$ tel que $x^n\in V$ pour tout $x\in W$.

Un ouvert est une union d'éléments de la base de topologie [10]. Nous voyons que V contient au moins une partie de la forme $]R,+\infty]$. Nous supposons que $R>1$.

Si $x>R>1$, alors nous avons $x^n\geqslant x$ (lemme 12.28) et donc

$$x^n\geqslant x>R, \tag{12.62}$$

ce qui signifie $x\in V$.

En prenant $W=]R,+\infty]$, nous avons bien $W^n\subset V$. Cela prouve (12.60).

En ce qui concerne la seconde limite, la démonstration est du même type. Remarquez seulement que vous n'avez pas formellement le droit d'utiliser la proposition 12.9 en invoquant $\frac{1}{+\infty}=0$. □

12.3.1 Limite en des nombres

Nous posons la définition suivante.

DefInfNombre

Définition 12.30.
*Lorsque $a\in\mathbb{R}$, on dit que la fonction f **tend vers l'infini quand** x **tend vers** a si*

$$\forall M\in\mathbb{R},\exists\delta \text{ tel que } (|x-a|\leqslant\delta)\Rightarrow f(x)\geqslant M \text{ quand } x\in\operatorname{dom}f.$$

Cela signifie que l'on demande que dès que x est assez proche de a (c'est-à-dire dès que $|x-a|\leqslant\delta$), alors $f(x)$ est plus grand que M, et que l'on peut trouver un δ qui fait ça pour n'importe quel M. Une autre façon de le dire est que pour toute hauteur M, on peut trouver un intervalle de largeur δ autour de a [11] tel que sur cet intervalle, la fonction f est toujours plus grande que M.

Montrons sur un dessin pourquoi je disais que la fonction $x\to 1/x$ n'est pas de ce type.

Le problème est qu'il n'existe par exemple aucun intervalle autour de 0 sur lequel f serait toujours plus grande que 10. En effet n'importe quel intervalle autour de 0 contient au moins un nombre négatif. Or quand x est négatif, f n'est certainement pas plus grande que 10. Nous y reviendrons.

Pour l'instant, montrons que la fonction $f(x)=1/x^2$ est une fonction qui vérifie la définition 12.30. Avant de prendre n'importe quel M, prenons par exemple 100. Nous avons besoin d'un intervalle autour de zéro sur lequel f est toujours plus grande que 100. C'est vite vu que $f(0.1)=f(-0.1)=100$, donc l'intervalle $[-\frac{1}{10},\frac{1}{10}]$ convient. Partout dans cet intervalle, f est plus grande que 100. Partout ? Ben non : en $x=0$, la fonction n'est même pas définie, donc c'est un peu dur de dire qu'elle est plus grande que 100. C'est pour cela que nous avons ajouté la condition « quand $x\in\operatorname{dom}f$ » dans la définition de la limite.

Prenons maintenant un $M\in\mathbb{R}$ arbitraire, et trouvons un intervalle autour de 0 sur lequel f est toujours plus grande que M. La réponse est évidemment l'intervalle de largeur $1/\sqrt{M}$, c'est-à-dire

$$\left[-\frac{1}{\sqrt{M}},\frac{1}{\sqrt{M}}\right].$$

12.3.2 Limites quand tout va bien

D'abord définissons ce qu'on entend par la limite d'une fonction en un point quand il n'y a aucun infini en jeu.

10. C'est la définition 7.2.
11. C'est-à-dire un intervalle de la forme $[a-\delta,a+\delta]$.

DefLimPointSansInfini

Définition 12.31.
*On dit que la fonction f **tend vers** b **quand** x **tend vers** a si*

$$\forall \epsilon > 0, \exists \delta \text{ tel que } (|x-a| \leqslant \delta) \Rightarrow |f(x)-b| \leqslant \epsilon \text{ quand } x \in \operatorname{dom} f.$$

Dans ce cas, nous notons

$$\lim_{x \to a} f(x) = b. \tag{12.63}$$

Commençons par un exemple très simple : prouvons que $\lim_{x\to 0} x = 0$. C'est donc $a = b = 0$ dans la définition. Prenons $\epsilon > 0$, et trouvons un intervalle autour de zéro tel que partout dans l'intervalle, $x \leqslant \epsilon$. Bon ben c'est clair que $\delta = \epsilon$ fonctionne.

Plus compliqué maintenant, mais toujours sans surprises.

Proposition 12.32.

$$\lim_{x \to 0} x^2 = 0.$$

Démonstration. Soit $\epsilon > 0$. On veut un intervalle de largeur δ autour de zéro tel que x^2 soit plus petit que ϵ sur cet intervalle. Cette fois-ci, le δ qui fonctionne est $\delta = \sqrt{\epsilon}$. En effet un élément de l'intervalle $[-\delta, \delta]$ est un r de valeur absolue plus petite ou égale à δ :

$$|r| \leqslant \delta = \sqrt{\epsilon}.$$

En prenant le carré de cette inégalité on a :

$$r^2 \leqslant \epsilon,$$

ce qu'il fallait prouver. □

Calculer et prouver des valeurs de limites, mêmes très simples, devient vite de l'arrachage de cheveux à essayer de trouver le bon δ en fonction de ϵ si on n'a pas quelques théorèmes généraux. Heureusement nous en avons déjà quelques uns : 12.7, 12.6, et 12.9.

PROPooWXBAooAEweSF

Proposition 12.33 ([1])**.**
Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une application continue dont la variable y varie dans un compact I de $\mathbb{R}$. Alors la fonction

$$\begin{aligned} d\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sup_{y \in I} f(x,y) \end{aligned} \tag{12.64}$$

est continue.

Démonstration. Soit x_0 fixé. Prouvons que d est continue en x_0. Nous notons y_0 la valeur de y qui réalise le maximum (par le théorème 9.55 et le fait que les fonctions projection soient continues, lemme 7.215). Soit aussi $\epsilon > 0$ tellement fixé que même avec un tournevis hydraulique, il ne bougerait pas. Nous considérons δ tel que si $\|(x,y) - (x_0,y_0)\| \leqslant \delta$ alors $\|f(x,y) - f(x_0,y_0)\| < \epsilon$.

Si $|x - x_0| < \delta$ alors pour y assez proche de y_0 nous avons $\|(x,y) - (x_0,y_0)\| \leqslant \delta$, et donc $\|f(x,y) - f(x_0,y_0)\| \leqslant \epsilon$. Cela montre qu'il existe δ tel que $|x - x_0| \leqslant \delta$ implique $d(x) \geqslant d(x_0) - \epsilon$.

Nous devons encore trouver un δ tel que si $|x - x_0| \leqslant \delta$ alors $d(x) \leqslant d(x_0) + \epsilon$. Supposons que non. Alors pour tout δ il existe un x tel que $|x - x_0| \leqslant \delta$ et $d(x) > d(x_0) + \epsilon$. Cela nous donne une suite $x_i \to x_0$.

Pour chaque x_i nous notons y_i la valeur de y qui réalise le supremum correspondant. La suite (y_i) étant contenue dans un compact nous supposons prendre une sous-suite de (x_i) telle que la suite (y_i) converge. Nous nommons a la limite (et non y_0 parce que nous ne savons pas si $y_i \to y_0$). Pour chaque i nous avons

$$f(x_i, y_i) > \sup_{y \in I} f(x_0, y) + \epsilon. \tag{12.65}$$

En prenant la limite et en utilisant la continuité de f,

$$f(x_0, a) > \sup_{y \in I} f(x_0, y) + \epsilon, \tag{12.66}$$

ce qui est impossible. □

12.3.3 Limites de fonctions

Tentons de comprendre ce que signifie qu'un nombre ℓ *ne soit pas* la limite de f lorsque $x \to a$. Il s'agit d'inverser la condition de la proposition 8.257(2). Le nombre ℓ n'est pas une limite de f pour $x \to a$ lorsque

$$\exists \varepsilon > 0 \text{ tel que } \forall \delta > 0, \exists x \text{ tel que } 0 < \|x - a\| < \delta \text{ et } \|f(x) - \ell\| > \varepsilon, \tag{12.67}$$ EqCaractNonLim

c'est-à-dire qu'il existe un certain seuil ε tel qu'on a beau s'approcher aussi proche qu'on veut de a (distance δ), on trouvera toujours un x tel que $f(x)$ n'est pas ε-proche de ℓ.

Lemme 12.34 (Unicité de la limite)**.**
Si ℓ et ℓ' sont deux limites de $f(x)$ lorsque x tend vers a, alors $\ell = \ell'$.

Démonstration. Soit $\varepsilon > 0$. Nous considérons δ tel que $\|f(x)-\ell\| < \varepsilon$ pour tout x tel que $\|x-a\| < \delta$. De la même manière, nous prenons δ' tel que $\|x - a\| < \delta'$ implique $\|f(x) - \ell'\| < \varepsilon$. Pour les x tels que $\|x - a\|$ est plus petit que δ et δ' en même temps, nous avons

$$\|\ell - \ell'\| = \|\ell - f(x) + f(x) - \ell'\| \leqslant \|\ell - f(x)\| + \|f(x) - \ell'\| < 2\varepsilon, \tag{12.68}$$

et donc $\|\ell - \ell'\| = 0$ parce que c'est plus petit que 2ε pour tout ε. □

PROPooKPOXooEHIXJs

Proposition 12.35 ([1])**.**
Soient un espace vectoriel normé $(V, \|.\|)$ et $a \in V$. Soient encore un voisinage A de a et deux fonctions $f, g \colon A\backslash\{a\} \to \mathbb{R}$ qui admettent une limite en a.

Si $f(x) \leqslant g(x)$ pour tout $x \in A\backslash\{a\}$ alors

$$\lim_{x \to a} f(x) \leqslant \lim_{x \to a} g(x). \tag{12.69}$$

PROPooGQHKooWgykjW

Proposition 12.36.
Si $f \colon \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ est continue et croissante, alors il existe $\ell \in \mathbb{R} \cup \{\pm\infty\}$ tel que

$$\lim_{x \to \infty} f(x) = \ell. \tag{12.70}$$

12.3.4 Limite à gauche et à droite

Si a est à l'intérieur du domaine de f, nous savons ce que signifie $\lim_{x \to a} f(x)$. Nous donnons également une définition des limites à gauche et à droite.

Définition 12.37.
*Soient $D \subset \mathbb{R}$ et une fonction $f \colon D \to \mathbb{R}$. Si $a \in \mathrm{Adh}(D)$ nous définissons la **limite à droite** de f en a par*

$$\lim_{x \to a^+} f(x) = \lim_{x \to a} \tilde{f}(x) \tag{12.71}$$ EQooQKHLooMoSXVe

où $\tilde{f}$ est la fonction f restreinte à $D \cap \{x \text{ tel que } x > a\}$. La limite (12.71) est souvent écrite sous la forme condensée

$$\lim_{\substack{x \to a \\ x > a}} f(x). \tag{12.72}$$

Pour la limite à gauche c'est un peu la même chose :

$$\lim_{x \to a^-} f(x) = \lim_{\substack{x \to a \\ x < a}} f(x). \tag{12.73}$$

LEMooXJMFooCkzoVi

Lemme 12.38.
Soient $D \subset \mathbb{R}$ et une fonction $f\colon D \to \mathbb{R}$. Si $a \in \mathrm{Adh}(D)$ nous avons $\lim_{x\to a^+} f(x) = \ell$ si et seulement si pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $x \in]a, a+\delta[\cap D$ implique $f(x) \in B(\ell, \epsilon)$.

Démonstration. Nous avons les équivalences entre les propriétés suivantes, en utilisant la définition 7.110 de la limite :

(1) $\lim_{x\to a^+} f(x) = \ell$

(2) $\lim_{x\to a} \tilde{f}(x) = \ell$

(3) Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que si $x \in B(a, \delta) \cap D \cap \{x > a\}$ alors $f(x) \in B(\ell, \epsilon)$

(4) Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que si $x \in]a, a+\delta[\cap D$ alors $f(x) \in B(\ell, \epsilon)$

□

La proposition suivante est un cas particulier de la proposition 9.112

PROPooGDDJooDCmydE

Proposition 12.39 ([373]).
Soit une fonction $f\colon D \to \mathbb{R}$ où D est une partie de $\mathbb{R}$. Si $a \in \mathrm{Adh}(D)$ alors la limite $\lim_{x\to a} f(x)$ existe si et seulement si les limites à gauche et à droite existent et sont égales. Dans ce cas nous avons égalité :

$$\lim_{x\to a} f(x) = \lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x). \tag{12.74}$$

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous disons que $\lim_{x\to a} f(x) = \ell$. Si V est un voisinage de ℓ, il existe un voisinage U de a tel que $f\big(U \cap D\backslash\{a\}\big) \subset V$. En particulier il existe un $\delta > 0$ tel que si $x \in]a, a+\delta[\cap D$, alors $|f(x) - \ell| < \epsilon$. Cela est la limite à droite (lemme 12.38).

(2) $\Leftarrow$ Soit $\epsilon > 0$. Par la limite à droite, il existe $\delta_1 > 0$ tel que $f\big(]a, a+\delta_1[\cap D\big) \subset B(\ell, \epsilon)$. Par la limite à gauche, il existe δ_2 tel que $f\big(]a-\delta_2, a[\cap D\big) \subset B(\ell, \epsilon)$
En prenant $\delta = \min\{\delta_1, \delta_2\}$ nous avons bien $f\big(B(a, \delta) \cap D\backslash\{a\}\big) \subset B(\ell, \epsilon)$ comme le demande la définition de la limite.

□

12.40.
Quelques remarques à propos de la proposition 12.39.

(1) Cette proposition ne se généralise pas aux dimensions supérieures. Dans $\mathbb{R}^2$ par exemple, il ne faudrait pas croire que si les limites suivant toutes les directions existent alors la limite existe.

(2) Cette proposition est souvent utilisée pour calculer des limites dans lesquelles interviennent des valeurs absolues. Par exemple, durant la démonstration de la proposition 12.349.

12.4 Limite en compactifié d'Alexandrov

Nous considérons l'espace topologique localement compact $\mathbb{R}$, et son compactifié d'Alexandrov défini en 7.106. Nous avons donc un point supplémentaire noté ∞. Ce point n'est ni du côté des grands nombres positifs, ni du côté des grands nombres négatifs. Il n'est ni $+\infty$ ni $-\infty$.

Proposition 12.41.
Dans cet espace topologique $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$,

$$\lim_{x\to 0} \frac{1}{x} = \infty. \tag{12.75}$$

Démonstration. Soit un voisinage V de ∞ dans $\hat{\mathbb{R}}$. Il s'écrit $V = K^c \cup \{\infty\}$ pour un certain compact de $\mathbb{R}$. Le théorème 9.24 nous assure que K est borné. Donc il existe $R > 0$ tel que $K \subset B(0, R)$. Pour $x \in B(0, 1/R)$ nous avons

$$\left|\frac{1}{x}\right| > R, \tag{12.76}$$

et donc $1/x \in K^c$. Donc aussi $\frac{1}{x} \in V$. □

De la même façon, dans $\mathbb{C} \cup \{\infty\}$ nous avons

$$\lim_{z \to 0} \frac{1}{z} = \infty. \tag{12.77}$$

12.42.
Je vous laisse deviner la topologie à considérer sur $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$. Dans cet espace topologique la limite $\lim_{x \to 0} \frac{1}{x}$ n'existe pas.

12.4.1 Prolongement par continuité

12.4.1.1 Discussion avec mon ordinateur

Voici un extrait de ce que peut donner Sage. Nous lui donnons la fonction

$$f(x) = \frac{x+4}{3x^2 + 10x - 8}. \tag{12.78}$$

Cette fonction est inventée exprès pour que le dénominateur s'annule en -4. En fait $3x^2 + 10x - 8 = (x+4)(3x-2)$, et la fraction peut se simplifier en

$$f(x) = \frac{1}{3x-2}. \tag{12.79}$$

Et avec cela nous écririons $f(-4) = -\frac{1}{14}$. Voyons comment cela passe dans Sage.

```
----------------------------------------------------------------------
| Sage Version 5.2, Release Date: 2012-07-25                         |
| Type "notebook()" for the browser-based notebook interface.        |
| Type "help()" for help.                                            |
----------------------------------------------------------------------
sage: f(x)=(x+4)/(3*x**2+10*x-8)
sage: f(-4)
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
ValueError: power::eval(): division by zero
```

Il produit donc une erreur de division par zéro. Cela n'est pas étonnant. Pourtant si on lui demande, il est capable de simplifier. En effet :

```
sage: f.simplify_full()
x |--> 1/(3*x - 2)
sage: f.simplify_full()(-4)
-1/14
```

Nous considérons la question suivante : étant donné une fonction f définie sur $I \backslash \{x_0\}$, est-il possible de définir f en x_0 de telle façon à ce qu'elle soit continue ?

Exemple 12.43.
La fonction

$$\begin{aligned} f\colon \mathbb{R}\backslash\{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned} \tag{12.80}$$

n'est pas définie pour $x = 0$ et il n'y a pas moyen de définir $f(0)$ de telle sorte que f soit continue parce que $\lim_{x \to 0} \frac{1}{x}$ n'existe pas. △

12.4.1.2 Limite et prolongement

Reprenons l'exemple de la fonction (12.78) que mon ordinateur refusait de calculer en zéro :

$$f(x) = \frac{x+4}{3x^2+10x-8} = \frac{x+4}{(x+4)(3x-2)}. \tag{12.81}$$

Cette fonction a une condition d'existence en $x = -4$. Et pourtant, tant que $x \neq -4$, cela a un sens de simplifier les $(x+4)$ et d'écrire

$$f(x) = \frac{1}{3x-2}.$$

Étant donné que pour toute valeur de x différente de -4, la fonction f s'exprime de cette façon, nous avons

$$\lim_{x\to -4} f(x) = \lim_{x\to -4} \left(\frac{1}{3x-2}\right).$$

Oui, mais la fonction [12] $g(x) = 1/(3x-2)$ est continue en -4 et donc sa limite vaut sa valeur. Nous en déduisons que

$$\lim_{x\to -4} f(x) = -\frac{1}{14}.$$

Que dire maintenant de la fonction ainsi définie ?

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq -4 \\ -1/14 & \text{si } x = -4. \end{cases} \tag{12.82}$$

Cette fonction est continue en -4 parce qu'elle y est égale à sa limite. Les étapes suivies pour obtenir ce résultat sont :

— Repérer un point où la fonction n'existe pas,

— calculer la limite de la fonction en ce point, et en particulier vérifier que cette limite existe, ce qui n'est pas toujours le cas,

— définir une nouvelle fonction qui vaut partout la même chose que la fonction originale, sauf au point considéré où l'on met la valeur de la limite.

C'est ce qu'on appelle **prolonger la fonction par continuité** parce que la fonction résultante est continue. La prolongation de f par continuité est donc en général définie par

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq y \\ \lim_{y\to x} f(y) & \text{si } x = y \end{cases} \tag{12.83}$$

Dans le cas que nous regardions,

$$f(x) = \frac{x+4}{3x^2+10x-8},$$

le prolongement par continuité est donné par

$$\tilde{f} = \frac{1}{3x-2}. \tag{12.84}$$

Remarquons que cette fonction n'est toujours pas définie en $x = 2/3$.

12. Cette fonction g n'est pas f parce que g a en plus l'avantage d'être définie en -4.

12.4.2 Prolongement par continuité

Proposition-Définition 12.44 (Prolongement par continuité)**.**
Soit $f\colon I\backslash\{x_0\}\to\mathbb{R}$ telle que $\lim_{x\to x_0} f(x)=\ell\in\mathbb{R}$. La fonction

$$\begin{aligned}\tilde{f}\colon I&\to\mathbb{R}\\ \tilde{f}(x)&=\begin{cases}f(x) & si\ x\neq x_0\\ \ell & si\ x=x_0\end{cases}\end{aligned}\tag{12.85}$$

*est une fonction continue sur I et est appelée le **prolongement par continuité** de f en x_0.*

Vous noterez que dans cet énoncé nous demandons $\ell\in\mathbb{R}$. Les cas $\ell=\pm\infty$ sont donc exclus.

12.45.
Le lemme 12.61 donnera un autre gros morceau de prolongement par continuité. Là, ce ne sera pas juste une valeur qui manquera, mais carrément la majorité des valeurs ; mais par contre, ce ne sera pas vraiment de la prolongation par continuité, mais de la prolongation par Cauchy-continuité.

Exemple 12.46.
La fonction

$$\begin{aligned}f\colon\mathbb{R}\backslash\{-3,2\}&\to\mathbb{R}\\ x&\mapsto\frac{x^2+2x-3}{(x+3)(x-2)}\end{aligned}\tag{12.86}$$

admet pour limite $\lim_{x\to-3} f(x)=\frac{4}{5}$. Son prolongement par continuité en $x=-3$ est donné par

$$\tilde{f}(x)=\frac{x-1}{x-2}.\tag{12.87}$$

Notons que les fonctions f et $\tilde{f}$ ne sont pas identiques : l'une est définie pour $x=-3$ et l'autre pas. Lorsqu'on fait le calcul

$$\frac{x^2+2x-3}{(x+3)(x-2)}=\frac{(x-1)(x+3)}{(x+3)(x-2)}=\frac{x-1}{x-2},\tag{12.88}$$

la simplification n'est pas du tout un acte anodin. Le dernier signe « = » est discutable parce que les deux dernières expressions ne sont pas égales pour tout x ; elles ne sont égales « que » pour les x pour lesquels les deux expressions existent. △

Les fonctions trigonométriques donneront quelques exemples intéressants de prolongements par continuité. Voir l'exemple 18.240. Et une avec la fonction logarithme dans l'exemple 15.114.

12.4.3 Théorème de la bijection

À propos de la continuité de l'application réciproque dans divers cas, voir le thème 35.

LEMooYUOOooByoQdr

Lemme 12.47 ([1])**.**
Si I est un intervalle de $\mathbb{R}$, et si $x\in\mathrm{Adh}(I)$, alors x est dans une des trois situations suivantes[13] *:*

(1) $x\in\mathrm{Int}(I)$
(2) $x=\inf(I)$
(3) $x=\sup(I)$

Démonstration. Soit $a\in\mathrm{Adh}(I)$. Nous supposons que a n'est pas dans l'intérieur de I, et nous prouvons que nous sommes dans un des deux autres cas. Soit $\epsilon>0$.

13. Définitions de suprémum et d'infimum 1.461.

(1) $a+\epsilon$ **n'est pas un minorant** Vu que a est dans l'adhérence de I, pour tout n, il existe x_n dans $B(a,1/n)\cap I$. En choisissant n assez grand, nous avons

$$x_n < a + \frac{1}{n} < a + \epsilon, \tag{12.89}$$

de telle sorte que $a+\epsilon$ ne soit pas un minorant de I.

(2) $a-\epsilon$ **n'est pas un majorant** De même nous avons $x_n > a - \frac{1}{n} > a - \epsilon$.

(3) $a \geqslant x$ **ou** $a \leqslant x$ **pour tout** $x \in I$ Nous devons encore prouver que soit $a \geqslant x$ pour tout $x \in I$, soit que $a \leqslant x$ pour tout $x \in I$. Supposons que ce ne soit pas le cas. Il existe $x_-, x_+ \in I$ tels que

$$x_- < a < x_+. \tag{12.90}$$

Vu que I est un intervalle nous avons $[x_-, x_+] \subset I$. Alors en prenant $r < (x_+ - x_-)/2$ nous avons

$$B(a,r) \subset]x_-, x_+[\subset I, \tag{12.91}$$

ce qui signifierait que a est dans l'intérieur de I.

□

LEMooBDGGooSOdOXb

Lemme 12.48 ([374])**.**
Soient un intervalle I et une application continue $f\colon I \to \mathbb{R}$. Elle est injective si et seulement si elle est strictement monotone.

Démonstration. En deux parties.

(1) **Si f est injective** Nous prouvons que f est strictement monotone par l'absurde en supposant qu'elle ne l'est pas. Si f n'est pas strictement monotone sur I, il existe x_1, x_2, x_3 et x_4 dans I tels que

$$\begin{aligned} x_1 < x_1 \qquad & x_3 < x_4 \\ f(x_1) \leqslant f(x_2) \quad & f(x_3) \geqslant f(x_4). \end{aligned} \tag{12.92}$$

Vu que I est un intervalle, tous les points entrex_1, x_2, x_3 et x_4 sont dans I. Nous pouvons donc définir

$$\begin{aligned} g\colon [0,1] &\to \mathbb{R} \\ t &\mapsto f\big(tx_1 + (1-t)x_2\big) - f\big(tx_2 + (1-t)x_4\big). \end{aligned} \tag{12.93}$$

Nous avons

$$\begin{aligned} g(0) &= f(x_3) - f(x_4) \geqslant 0 \\ g(1) &= f(x_1) - f(x_2) \leqslant 0. \end{aligned} \tag{12.94}$$

Le théorème des valeurs intermédiaires 9.93 nous indique qu'il existe $t_0 \in [0,1]$ tel que $g(t_0) = 0$. Nous posons alors $u = t_0x_1 + (1-t_0)x_3$ et $v = t_0x_2 + (1-t_0)x_4$, de telle sorte que

$$0 = g(t_0) = f(u) - f(v). \tag{12.95}$$

Mais $u \neq v$ parce que $x_1 < x_2$ et $x_3 < x_4$. Nous avons contredit l'injectivité de f.

(2) **Si f est strictement monotone** Supposons que f soit strictement croissante. Nous devons prouver que f est injective. Pour cela supposons que $f(u) = f(v)$ avec $u, v \in I$. Nous avons alors

$$\begin{cases} f(u) \leqslant f(v) & \text{(12.96a)} \\ f(u) \geqslant f(v). & \text{(12.96b)} \end{cases}$$

Par stricte monotonie nous en déduisons que $u \leqslant v$ et que $u \geqslant v$, autrement dit, que $u = v$.

□

PropOARooUuCaYT

Proposition 12.49 ([375, 1])**.**
Une surjection monotone d'une partie X de $\mathbb{R}$ vers un intervalle est continue[14]*. Nous sommes dans une des trois situations du lemme 12.47.*

Démonstration. Soit une application monotone $g: X \to I$ où I est un intervalle. Nous supposons dans un premier temps que g est croissante. Soit un ouvert V dans I. Nous allons montrer que $g^{-1}(V)$ est ouvert dans X.

Pour cela nous considérons $x \in g^{-1}(V)$ et nous posons $y = g(x)$. Nous allons prouver qu'il existe un voisinage U de x vérifiant $g(U) \subset V$.

Le nombre y est dans un des trois cas du lemme 12.47.

(1) y **est à l'intérieur de** I Il existe $r > 0$ tel que $B(y, r) \subset V$. Nous posons alors $0 < \alpha < r$, et choisissons $x_- \in g^{-1}(y - \alpha)$ et $x_+ \in g^{-1}(y + \alpha)$. Vu que g est surjective sur I, ces nombres existent dans X. Par croissance de g, nous avons

$$x_- < x < x_+. \tag{12.97}$$

Enfin nous considérons la partie $U =]x_-, x_+[\cap X$. Ce U est un ouvert contenant x.
Avec tout ça nous avons $g(U) \subset V$.

(2) **Si** $y = \sup(I)$ Il existe r tel que $]y - r, y] \subset V$. Nous considérons $x_- \in g^{-1}(y - r)$, et nous posons $U =]x_-, \infty[\cap X$. Avec cela nous avons

$$g(U) \subset]y - \epsilon, y] \subset V. \tag{12.98}$$

(3) **Si** $y = \inf(I)$ Dans ce cas nous avons un r tel que $[y, y + r[\subset V$. Nous considérons $x_+ \in g^{-1}(y + r)$ et l'ouvert $]-\infty, x_+[\cap X$. Nous avons encore $g(U) \subset V$.

Nous devons encore prouver le cas où g est décroissante. Dans ce cas la fonction $h(x) = -g(x)$ est croissante et surjective sur l'intervalle $-I$. Elle est donc continue. Si h est continue, alors g est continue. □

Proposition 12.50.
Soient $f: I \to J$ une bijection et $f^{-1}: J \to I$ sa réciproque. Alors pour tout $x_0 \in I$ nous avons

$$f^{-1}\big(f(x_0)\big) = x_0 \tag{12.99}$$

EqHQRooNmLYbF

et pour tout $y_0 \in J$ nous avons

$$f\big(f^{-1}(y_0)\big) = y_0. \tag{12.100}$$

EqIYTooQPvZDr

Démonstration. Nous prouvons la relation (12.99) et nous laissons (12.100) comme exercice au lecteur.

Soit $x_0 \in I$. Posons $y_0 = f(x_0)$. La définition de l'application réciproque est que pour $y \in J$, $f^{-1}(y)$ est l'unique élément x de I tel que $f(x) = y$. Donc $f^{-1}(y_0)$ est l'unique élément de I dont l'image est y_0. C'est donc x_0 et nous avons $f^{-1}(y_0) = x_0$, c'est-à-dire

$$f^{-1}\big(f(x_0)\big) = x_0. \tag{12.101}$$

□

ThoKBRooQKXThd

Théorème 12.51 (Théorème de la bijection)**.**
Soit I un intervalle et f une fonction continue et strictement monotone de I dans $\mathbb{R}$. Nous avons alors :

(1) $f(I)$ est un intervalle de $\mathbb{R}$;

ITEMooMAWXooZXmVwA

(2) La fonction $f: I \to f(I)$ est bijective

14. Pour les topologies induites de X et de I. Dans le cas de I, si il est ouvert, c'est la topologie usuelle.

(3) *La fonction $f^{-1}\colon f(I)\to I$ est strictement monotone de même sens que f ;* ItemEJZooKuFoeFiv

(4) *La fonction $f\colon I\to f(I)$ est un homéomorphisme, c'est-à-dire que $f^{-1}\colon f(I)\to I$ est continue.*

Démonstration. Prouvons les choses point par point.

(1) Supposons pour fixer les idées que f est monotone croissante [15].
Soient $a<b$ dans $f(I)$. Par définition il existe $x_1,x_2\in I$ tels que $a=f(x_1)$ et $b=f(x_2)$. La fonction f est continue sur l'intervalle $[x_1,x_2]$ et vérifie $f(x_1)<f(x_2)$. Donc le théorème des valeurs intermédiaires 9.93 nous dit que pour tout t dans $[f(x_2),f(x_2)]$, il existe un $x_0\in[x_1,x_2]$ tel que $f(x_0)=t$. Cela montre que toutes les valeurs intermédiaires entre a et b sont atteintes par f et donc que $f(I)$ est un intervalle.

(2) Nous prouvons maintenant que f est bijective en prouvant séparément qu'elle est surjective et injective.

(2a) **f est surjective** Une fonction est toujours surjective depuis un intervalle I vers l'ensemble Image f.

(2b) **f est injective** Soit $x\neq y$ dans I ; pour fixer les idées nous supposons que $x<y$. La stricte monotonie de f implique que $f(x)<f(y)$ ou que $f(x)>f(y)$. Dans tous les cas $f(x)\neq f(y)$.

La fonction f est donc bijective.

(3) Comme d'accoutumée nous supposons que f est croissante. Soient $y_1<y_2$ dans $f(I)$; nous devons prouver que $f^{-1}(y_1)\leqslant f^{-1}(y_2)$. Pour cela nous considérons les nombres $x_1,x_2\in I$ tels que $f(x_1)=y_1$ et $f(x_2)=y_2$. Nous allons en prouver la contraposée en supposant que $f^{-1}(y_1)>f^{-1}(y_2)$. En appliquant f (qui est croissante) à cette dernière inégalité il vient

$$f\big(f^{-1}(y_1)\big)\geqslant f\big(f^{-1}(y_2)\big),\tag{12.102}$$

ce qui signifie

$$y_1\geqslant y_2\tag{12.103}$$

par l'équation (12.100).

(4) La fonction $f^{-1}\colon f(I)\to I$ est une fonction monotone et surjective, donc continue par la proposition 12.49.

□

Exemple 12.52.
La fonction

$$\begin{aligned}f\colon[2,3]&\to[4,9]\\ x&\mapsto x^2\end{aligned}\tag{12.104}$$

est une bijection. Sa réciproque est la fonction

$$\begin{aligned}f^{-1}\colon[4,9]&\to[2,3]\\ x&\mapsto\sqrt{x}.\end{aligned}\tag{12.105}$$

△

15. Traitez en tant qu'exercice le cas où f est décroissante.

12.5 Limite et continuité

SecLimiteFontion

Voir les remarques dans l'index thématique 29 pour comprendre la place et la portée de ce qui va venir à propos de limite et de continuité.

ThoLimCont

Théorème 12.53 (Limite et continuité)**.**
La fonction f est continue au point a si et seulement si $\lim_{x\to a} f(x) = f(a)$.

Démonstration. Nous commençons par supposer que f est continue en a, et nous prouvons que $\lim_{x\to a} f(x) = f(a)$. Soit $\epsilon > 0$; ce qu'il nous faut c'est un δ tel que $|x - a| \leqslant \delta$ implique $|f(x) - f(a)| \leqslant \epsilon$. La caractérisation 9.83 de la continuité donne l'existence d'un δ comme il nous faut.

Dans l'autre sens, c'est-à-dire prouver que f est continue au point a sous l'hypothèse que $\lim_{x\to a} f(x) = f(a)$, la preuve se fait de la même façon. □

Nous en déduisons que si nous voulons gagner quelque chose à parler de limites, il faut prendre des fonctions non continues. En effet, si une fonction est continue en un point, la limite ne donne aucune nouvelle information que la valeur de la fonction elle-même en ce point.

Prenons une fonction qui fait un saut. Pour se fixer les idées, prenons celle-ci :

$$f(x) = \begin{cases} 2x & \text{si } x \in]-\infty, 2[\\ x/2 & \text{si } x \in [2, +\infty[\end{cases} \tag{12.106}$$

EqnCtOEt

Essayons de trouver la limite de cette fonction lorsque x tend vers 2. Étant donné que f n'est pas continue en 2, nous savons déjà que $\lim_{x\to 2} f(x) \neq f(2)$. Donc ce n'est pas 1. Cette limite ne peut pas valoir 4 non plus parce qu'en prenant n'importe quel ϵ, la valeur de $f(2 + \epsilon)$ est très proche de 2, et donc, ne peut pas s'approcher de 4. En fait, on peut facilement vérifier que *aucun nombre ne vérifie la condition de limite pour f en* 2. Nous disons que la limite n'existe pas.

Il ne faudrait pas en déduire trop vite que si une fonction n'est pas continue en a, alors la limite $x \to a$ n'existe pas. Ce que dit le théorème 12.53 est que si une fonction n'est pas continue en a, alors sa limite (si elle existe) ne vaut pas $f(a)$.

EXooKREUooLeuIlv

Exemple 12.54 (Un exemple de continuité ; Thème 29)**.**
Soit la fonction

$$f(x) = \begin{cases} x & \text{si } x \neq 0 \\ 4 & \text{si } x = 0. \end{cases} \tag{12.107}$$

EQooSYSWooSGsUfR

Cette fonction n'est pas continue en $x = 0$, et pourtant la limite existe : $\lim_{x\to 0} f(x) = 0$. Etudions cela en détail, pour nous assurer de ce qu'il se passe.

Considérons l'ouvert $]3, 5[$. L'image réciproque de cet ouvert par f est la partie $]3, 5[\cup \{0\}$ qui n'est pas ouvert. Donc la fonction f n'est pas continue comme fonction $\mathbb{R} \to \mathbb{R}$.

Considérons pour comprendre la restriction $f\colon [-1, 1] \to \mathbb{R}$. L'image inverse de $]3, 5[$ par cette fonction est $\{0\}$ qui n'est pas un ouvert.

Plus généralement tant qu'on considère des restrictions de f sur des parties contenant un voisinage de 0, la fonction ne peut pas être continue [16].

Voyons ce qu'il en est de la continuité ponctuelle de f en $x = 0$. La définition 7.36 est celle de la continuité en un point ; elle dit que f sera continue en 0 si $f(0) = 4$ est une limite de f. Nous voilà parti vers la définition 7.110.

Soit le voisinage $V =]3, 5[$ de $f(0)$. Quel que soit le voisinage W de 0 dans $\mathbb{R}$, il existe un $\epsilon > 0$ tel que $W \subset B(0, \epsilon)$. Nous avons alors

$$f\big(W\backslash\{0\}\big) \subset f\big(B(0, \epsilon)\backslash\{0\}\big). \tag{12.108}$$

Mais le nombre $\epsilon/2$ fait partie de $f\big(B(0, \epsilon)\backslash\{0\}\big)$ et n'est pas dans V. Donc $f(0)$ n'est pas une limite de f en zéro. Cette fonction n'est donc pas continue en zéro. △

16. Les plus acharnés se demanderont ce qu'il se passe pour la restriction de f à la partie $\{0\}$ munie de la topologie induite de $\mathbb{R}$.

Exemple 12.55 (Même exemple, limite)**.**
Nous avons vu que, pour la fonction (12.107), le nombre 4 n'est pas une limite de f en zéro. Nous montrons à présent que 0 est une limite (et même la seule par la proposition 7.113 que nous ne rappellerons plus à chaque fois) de f.

Montrons que 0 est une limite de f en zéro, c'est-à-dire que $\lim_{x\to 0} f(x) = 0$.

Nous suivons la définition 7.110. Soit un voisinage V de 0 dans $\mathbb{R}$. Il existe δ tel que $B(0,\delta) \subset V$. En posant $\epsilon = \delta$ et en définissant $W = B(0,\epsilon)$ nous avons

$$f\big(B(0,\epsilon)\backslash\{0\}\big) = B(0,\epsilon)\backslash\{0\} \subset B(0,\delta) \subset V. \tag{12.109}$$

Donc 0 est une limite de f en zéro. △

Nous avons déjà vu par le corolaire 9.59 qu'une suite croissante et bornée était convergente. Il en va de même pour les fonctions.

PropMTmBYeU

Proposition 12.56 ([1])**.**
Si la fonction réelle $f\colon I = [a,b[\to \mathbb{R}$ est croissante et bornée, alors la limite

$$\lim_{x\to b} f(x) \tag{12.110}$$

existe et est finie.

Démonstration. Commençons par prouver que si (x_n) est une suite dans I convergeant vers b, alors $f(x_n)$ est une suite convergente. Dans un second temps, nous allons prouver que si (x_n) et (x'_n) sont deux suites qui convergent vers b, alors les suites convergentes $f(x_n)$ et $f(x'_n)$ convergent vers la même limite. Alors le critère séquentiel de la limite d'une fonction conclura (proposition 7.249).

Nous pouvons extraire de x_n une sous-suite croissante $(x_{\alpha(n)})$. Alors la suite $f\big(x_{\alpha(n)}\big)$ est une suite croissante et majorée, donc convergente par le corolaire 9.59 [17]. Nommons ℓ la limite et montrons qu'elle est aussi limite de f sur la suite originale.

Pour tout $\epsilon > 0$, il existe K tel que si $n > K$ alors $\big|f\big(x_{\alpha(n)}\big) - \ell\big| < \epsilon$. Soit K' tel que pour tout $n > K'$ nous ayons $x_n > x_{\alpha(K)}$. Cela est possible parce que la suite est bornée par b et converge vers b : il suffit de prendre K' de telle sorte que $|x_n - b| \leqslant |x_{\alpha(n)} - b|$. Si $n > K'$ alors $x_n > x_{\alpha(K)}$ et

$$f(x_n) \geqslant f(x_{\alpha(n)}) \geqslant \ell - \epsilon; \tag{12.111}$$

en résumé si $n > K$ alors $|f(x_n) - \ell| < \epsilon$. Cela prouve que $f(x_n) \to \ell$.

Soit maintenant une autre suite (x'_n) qui converge également vers b. Comme nous venons de le voir la suite $f(x'_n)$ est convergente et nous nommons ℓ' la limite. Si nous considérons (x''_n) la suite « alternée » $(x_1, x'_1, x_2, x'_2, \cdots)$ alors nous avons encore une suite qui converge vers b et donc $f(x''_n) \to \ell'$.

Mais étant donné que $f(x_n)$ et $f(x'_n)$ sont des sous-suites, elles doivent converger vers la même valeur. Donc $\ell = \ell' = \ell''$. □

PROPooLOQVooULDhZz

Proposition 12.57 ([1])**.**
L'ensemble $\mathbb{R} \cup \{\infty\}$ muni de la topologie de la compactification en un point [18] *est connexe par arcs.*

Démonstration. Nous allons montrer que le chemin

$$\begin{aligned} \gamma\colon [0,1] &\to [0,\infty] \\ x &\mapsto \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ \infty & \text{si } x = 0 \end{cases} \end{aligned} \tag{12.112}$$

est continu au sens de la définition 7.36(2) (qui est le seul sens possible au mot « continu »).

17. En gros nous sommes en train de dire que toute la théorie des fonctions convexes est un vulgaire corolaire de Bolzano-Weierstrass.
18. Définition 7.106.

Soit un ouvert $\mathcal{O}$ dans $\mathbb{R}\cup\{\infty\}$. Si cet ouvert ne contient pas ∞, alors $\gamma^{-1}(\mathcal{O})$ est ouvert dans $\mathbb{R}$ parce que la fonction $x\mapsto 1/x$ est continue [19].

Si $\infty\in\mathcal{O}$, alors $\mathcal{O}=\{\infty\}\cup\mathcal{O}'$ où $\mathcal{O}'$ est un ouvert de $\mathbb{R}$ ayant la propriété que $\mathbb{R}\backslash\mathcal{O}'$ est compact.

Nous avons $\gamma^{-1}(\mathcal{O})=\{0\}\cup\gamma^{-1}(\mathcal{O}')$. Et aussi que $\gamma^{-1}(\mathcal{O}')$ est un ouvert de $\mathbb{R}$ contenu dans $[0,1]$.

Puisque le complémentaire de $\mathcal{O}'$ est compact, il existe $a\in\mathbb{R}$ tel que $]a,\infty[\subset\mathcal{O}'$. Donc $\mathcal{O}'=]a,\infty[\cup\mathcal{O}''$ où $\mathcal{O}''$ est un ouvert.

Nous avons :

$$\gamma^{-1}(\mathcal{O})=\{\gamma^{-1}(\infty)\}\cup\gamma^{-1}\big(]a,\infty[\big)\cup\gamma^{-1}(\mathcal{O}'') \tag{12.113a}$$
$$=\{0\}\cup\big]0,\frac{1}{a}\big[\cup\gamma^{-1}(\mathcal{O}'') \tag{12.113b}$$
$$=\big[0,\frac{1}{a}\big[\cup\gamma^{-1}(\mathcal{O}''). \tag{12.113c}$$

Vous noterez le point essentiel où la topologie de la compactification agit : comme $\{0\}$ n'est pas un ouvert de $[0,1]$, nous devons nous assurer que la partie $\gamma^{-1}(\mathcal{O}')$ vienne se coller à $\{0\}$ pour compléter en un ouvert.

L'ensemble $\gamma^{-1}(\mathcal{O}'')$ est un ouvert de $\mathbb{R}$ contenu dans $[0,1]$. Nous avons donc

$$\gamma^{-1}(\mathcal{O})=\Big(\big]-1,\frac{1}{a}\big[\cup\gamma^{-1}(\mathcal{O}'')\Big)\cap[0,1]. \tag{12.114}$$

Cela est un ouvert de $[0,1]$ par définition de la topologie induite [20]. □

12.5.1 Prolongement des rationnels vers les réels

Si $f:\mathbb{Q}\to\mathbb{R}$ est une fonction continue pour la topologie induite, est-ce qu'on peut la prolonger en une fonction continue sur $\mathbb{R}$? La réponse est hélas non.

EXooWZNCooQkKdtJ

Exemple 12.58 ([376]).
Puisque $\sqrt{2}$ est irrationnel [21], ceci définit bien une fonction sur $\mathbb{Q}$:

$$\begin{aligned}f:\mathbb{Q}&\to\mathbb{R}\\ q&\mapsto\begin{cases}0&\text{si } q<\sqrt{2}\\ 1&\text{si } q>\sqrt{2}.\end{cases}\end{aligned} \tag{12.115}$$

C'est une fonction continue sur $\mathbb{Q}$. En effet, soient $q\in\mathbb{Q}$ et $\epsilon>0$. Nous prenons $\delta>0$ tel que $\sqrt{2}$ ne soit pas dans $B(q,\delta)$. Alors si $p\in B_{\mathbb{Q}}(q,\delta)$ nous avons $f(q)=f(p)$ et donc

$$|f(p)-f(q)|<\epsilon. \tag{12.116}$$

Il n'est cependant pas possible de la prolonger en une fonction continue sur $\mathbb{R}$. △

Pour qu'une fonction sur $\mathbb{Q}$ puisse être prolongée en une fonction continue sur $\mathbb{R}$, il faut un peu plus que la continuité. Il faut la Cauchy-continuité, que nous définissons pas plus tard qu'immédiatement.

DEFooXXOGooXblyKP

Définition 12.59 ([376]).
Soient X et Y deux espaces métriques. Une application $f:X\to Y$ est dite ***Cauchy-continue*** *si pour toute suite de Cauchy (x_n) dans X, la suite $\big(f(x_n)\big)$ est de Cauchy dans Y.*

19. Voir par exemple la proposition 12.6.
20. Définition 7.26.
21. Proposition 1.409. Le fait que $\sqrt{2}$ existe dans $\mathbb{R}$ est la proposition 1.474.

PROPooPUNFooFMytOY

Proposition 12.60 ([377]).
Soient X et Y des espaces métriques[22]*. Nous supposons que Y est complet*[23]*. Nous considérons $A \subset X$ et une application Cauchy-continue $f\colon A \to Y$.*

Alors :

(1) Il existe un unique prolongement continu $g\colon \bar{A} \to Y$ de f.

ITEMooDEZGooUAvwvF

(2) Ce prolongement est lui-même Cauchy-continu.

Démonstration. En plusieurs parties.

(1) **Une suite**

Soit $x \in \bar{A}$. Nous considérons une suite (a_n) dans A telle que $a_n \xrightarrow{X} x$. Alors (a_n) est de Cauchy[24] et donc $\big(f(a_n)\big)$ est de Cauchy dans Y. Vu que Y est complet, $\big(f(a_n)\big)$ est une suite convergente.

(2) **Indépendance**

Montrons que si (b_n) est une autre suite dans A convergeant vers x, alors $\lim_n f(a_n) = \lim_n f(b_n)$. Pour cela nous considérons la suite$(a_0, b_0, a_1, b_1, \dots,)$. Cela est une suite qui converge vers x. Elle est donc de Cauchy et son image par f est encore de Cauchy. Toutes les sous-suites de la suite image sont dont convergentes vers la même limite. En prenant les sous-suites paires et impaires, nous avons $\lim_n f(a_n) = \lim_n f(b_n)$.

(3) **Définition de g** Nous pouvons donc définir

$$\begin{aligned} g\colon \bar{A} &\to Y \\ x &\mapsto \lim_n f(a_n) \end{aligned} \tag{12.117}$$

où (a_n) est une quelconque suite dans A convergeant vers x.

(4) **g est séquentiellement continue** Soit une suite (x_n) dans $\bar{A}$. Nous devons prouver que $g(x_n) \xrightarrow{Y} g(x)$.

Nous pouvons choisir $a_n \in A$ tel que $d(a_n, x_n) < 1/n$. Si nous considérons une suite de Cauchy dans A convergeant vers x_n, et que nous prenons a_n assez loin dans cette suite, alors nous pouvons même choisir a_n tel que $d\big(f(a_n), g(x_n)\big) < 1/n$. Nous choisissons (a_n) de telle sorte à vérifier les deux en même temps.

Nous avons

$$d(a_n, x) \leqslant d(a_n, x_n) + d(x_n, x) \to 0. \tag{12.118}$$

Donc (a_n) est de Cauchy et par définition de g, nous avons

EQooWQYNooKIQAjc

$$g(x) = \lim_{n\to\infty} f(a_n). \tag{12.119}$$

À partir de là nous avons la majoration

$$d\big(g(x_n), g(x)\big) \leqslant d\big(g(x_n), f(a_n)\big) + d\big(f(a_n), g(x)\big). \tag{12.120}$$

Le premier terme tend vers zéro par choix de la suite (a_n) et le second tend vers zéro par (12.119).

(5) **g est continue** La proposition 7.255 donne l'équivalence entre la continuité et la continuité séquentielle.

(6) **Unicité** Soient g_1 et g_2, deux prolongations continues de f sur $\bar{A}$. Si $a_n \to x$ dans A, alors nous avons, par continuité de g_1 et g_2 : $g_1(a_n) \to g_1(x)$ et $g_2(a_n) \to g_2(x)$.

Mais $g_1(a_n) = g_2(a_n) = f(a_n)$; donc les limites sont égales.

22. Il a une distance, définition 7.115.
23. Définition 7.263.
24. Proposition 7.270.

(7) *g* **est Cauchy continue**

□

En terme de prolongement continu, nous avons ce lemme qui demande à une fonction d'être Cauchy-continue. Vous pouvez comparer avec le principe de prolongement analytique 17.138 qui donne un énoncé similaire pour un prolongement analytique.

LEMooUAFBooAwiXxj

Lemme 12.61 ([1, 378, 377])**.**
Soit une fonction Cauchy-continue[25] $f\colon \mathbb{Q} \to \mathbb{R}$.

(1) La limite $\lim_{q\to x} f(q)$ *existe pour tout* $x \in \mathbb{R}$.

(2) Il existe un unique prolongement continu $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$.

(3) Ce prolongement est donné par

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in \mathbb{Q} \\ \lim_{q\to x} f(q) & \text{sinon} \end{cases} \tag{12.121}$$

Démonstration. C'est tout dans la proposition 12.60 en sachant que

— $\mathbb{Q}$ est dense dans $\mathbb{R}$, proposition 9.16.

— $\mathbb{R}$ est complet, théorème 7.287.

□

PROPooXWHYooFiVYfi

Proposition 12.62.
Soient des fonctions continues $f, g\colon \mathbb{R} \to \mathbb{R}$. *Si* f *et* g *sont égales sur* $\mathbb{Q}$*, alors elles sont égales sur* $\mathbb{R}$.

Démonstration. Nous pouvons utiliser les propriétés fondamentales des réels et de la continuité. Soit $x \in \mathbb{R}$; nous voulons montrer que $f(x) = g(x)$. En prenant par exemple le lemme 1.453, il existe une suite q_i de rationnels telle que $q_i \xrightarrow{\mathbb{R}} x$.

Par ailleurs, f et g sont continues sur $\mathbb{R}$ et donc en chaque point de $\mathbb{R}$ (théorème 7.205). Par la caractérisation séquentielle 7.137 de la continuité, nous avons

$$f(x) = \lim_{i\to\infty} f(q_i) = \lim_{i\to\infty} g(q_i) = g(x). \tag{12.122}$$

□

PROPooTNIAooNAJDzL

Proposition 12.63 ([1])**.**
Soit une fonction strictement croissante $f\colon \mathbb{Q} \to \mathbb{R}$. *Alors la prolongation continue* $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ *est également strictement croissante.*

Démonstration. Soient $x, y \in \mathbb{R}$ avec $x < y$. Notons $d = y - x$. Nous considérons des suites de rationnels $x_k \to x$ et $y_l \to y$ telles que pour tout k, $x_k \in B(x, d/3)$ et $y_k \in B(y, d/3)$. En particulier, $x_k < y_l$ pour tout k et l.

Soient des rationnels q et q' tels que pour tout k,

$$x_k < q < q' < y_k. \tag{12.123}$$

Pour trouver de tels rationnels, il suffit de les chercher dans $]x + \frac{d}{3}, y - \frac{d}{3}[$. Cet intervalle étant de longueur $d/3$, il contient des rationnels.

Vue la croissance de f sur $\mathbb{Q}$, nous avons, pour tout k :

$$f(x_k) < f(q) < f(q') < f(y_k), \tag{12.124}$$

et à la limite :

$$\tilde{f}(x) \leqslant f(q) < f(q') \leqslant \tilde{f}(y). \tag{12.125}$$

Notez que les inégalités strictes se changent en inégalités larges au passage à la limite. D'où l'utilité de prendre *deux* rationnels entre x_k et y_k pour maintenir une inégalité stricte entre $\tilde{f}(x)$ et $\tilde{f}(y)$. □

25. Définition 12.59 ; nous en avons discuté dans l'exemple 12.58.

12.6 Espace des fonctions continues

Définition 12.64.
*Soit I, un intervalle de $\mathbb{R}$. L'**oscillation** sur I est le nombre*

$$\omega_f(I) = \sup_{x\in I} f(x) - \inf_{x\in I} f(x). \tag{12.126}$$

Pour chaque x fixé, la fonction

$$x \mapsto \omega_f\big(B(x,\delta)\big) \tag{12.127}$$

est une fonction positive, croissante et a donc une limite (pour $\delta \to 0$). Nous notons $\omega_f(x)$ cette limite qui est l'**oscillation** de f en ce point. Une propriété immédiate est que f est continue en x_0 si et seulement si $\omega_f(x_0) = 0$.

LemuaPbtQ

Lemme 12.65.
L'ensemble des points de discontinuité d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est une réunion dénombrable de fermés.

Démonstration. Soit D l'ensemble des points de discontinuité de f. Nous avons

$$D = \bigcup_{n=1}^{\infty} \{x \text{ tel que } \omega_f(x) \geqslant \frac{1}{n}\}. \tag{12.128}$$

Il nous suffit donc de montrer que pour tout ϵ, l'ensemble

$$\{x \text{ tel que } \omega_f(x) < \epsilon\} \tag{12.129}$$

est ouvert. Soit en effet x_0 dans cet ensemble. Il existe δ tel que $\omega_f\big(B(x_0,\delta)\big) < \epsilon$. Si $x \in B(x_0,\delta)$, alors si on choisit δ' tel que $B(x,\delta') \subset B(x_0,\delta)$, nous avons $\omega_f\big(B(x,\delta')\big) < \epsilon$, ce qui justifie que $\omega_f(x) < \epsilon$ et donc que x est également dans l'ensemble considéré. □

Théorème 12.66.
L'ensemble des points de discontinuité d'une limite simple de fonctions continues est de première catégorie.

Démonstration. Soit (f_n) une suite de fonctions qui converge simplement vers f. Nous devons écrire l'ensemble des points de discontinuité de f comme une union dénombrable d'ensembles tels que sur tout intervalle I, aucun de ces ensembles n'est dense. Nous savons déjà par le lemme 12.65 que l'ensemble des points de discontinuité de f est donné par

$$D = \bigcup_{n=1}^{\infty} \{x \text{ tel que } \omega_f(x) \geqslant \frac{1}{n}\}. \tag{12.130}$$

Nous essayons donc de prouver que pour tout ϵ, l'ensemble

$$F = \{x \text{ tel que } \omega_f(x) \geqslant \epsilon\} \tag{12.131}$$

est nulle part dense. Soit

$$E_n = \bigcap_{i,j>n} \{x \text{ tel que } |f_i(x) - f_j(x)| < \epsilon\}. \tag{12.132}$$

Nous montrons que cet ensemble est fermé en étudiant le complémentaire. Soit $x \notin E_n$; alors il existe un couple (i,j) tel que

$$|f_i(x) - f_j(x)| > \epsilon. \tag{12.133}$$

Par continuité, cette inégalité reste valide dans un voisinage de x. Donc il existe un voisinage de x contenu dans $\complement E_n$ et E_n est donc fermé.

De plus nous avons $E_n \subset E_{n+1}$ et $\bigcup_n E_n = \mathbb{R}$. Ce dernier point est dû au fait que pour tout x, il existe N tel que $i, j > N$ implique $|f_i(x) - f_j(x)| \leqslant \epsilon$. Cela est l'expression du fait que la suite $\big(f_n(x)\big)_{n\in\mathbb{N}}$ est de Cauchy.

Soit I, un intervalle fermé de $\mathbb{R}$. Nous voulons trouver un intervalle $J \subset I$ sur lequel f est continue. Nous écrivons I sous la forme

$$I = \bigcup_{n=1}^{\infty} (E_n \cap I). \tag{12.134}$$

Tous les ensembles $J_n = E_n \cap I$ ne peuvent être nulle part dense en même temps (à cause du théorème de Baire 7.325). Il existe donc un n tel que J_n contienne un ouvert J. Le but est de montrer que f est continue sur J. Pour ce faire, nous n'allons pas simplement majorer $|f(x)-f(x_0)|$ par ϵ lorsque $|x - x_0|$ est petit. Nous allons majorer l'oscillation de f sur $B(x_0, \delta)$ lorsque δ est petit. Pour cela nous prenons x_0 et x dans J et nous écrivons

$$|f(x) - f(x_0)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)|. \tag{12.135}$$

À ce niveau nous rappelons que n est fixé par le choix de J, dans lequel ϵ est déjà inclus. Nous choisissons évidemment $|x - x_0| \leqslant \delta$ de telle sorte que le second terme soit plus petit que ϵ en vertu de la continuité de f_n. Pour le premier terme, pour tout $i, j \geqslant n$ nous avons

$$|f_i(x) - f_j(x)| < \epsilon. \tag{12.136}$$

Si nous posons $j = n$ et $i \to \infty$, en tenant compte du fait que $f_i \to f$ simplement,

$$|f(x) - f_n(x)| \leqslant \epsilon. \tag{12.137}$$

Nous avons donc obtenu $|f(x)-f_n(x_0)| \leqslant 2\epsilon$. Cela signifie que dans un voisinage de rayon δ autour de x_0, les valeurs extrêmes prises par $f(x)$ sont $f_n(x_0) \pm 4\epsilon$. Nous avons donc prouvé que pour tout ϵ, il existe δ tel que

$$\omega_f\big([x_0 - \delta, x_0 + \delta]\big) \leqslant 4\epsilon. \tag{12.138}$$

De là nous concluons que

$$\lim_{\delta\to 0} \omega_f\big([x_0 - \delta, x_0 + \delta]\big) = 0, \tag{12.139}$$

ce qui signifie que f est continue en x_0. □

Exemple 12.67.
Une fonction discontinue sur $\mathbb{Q}$ et continue ailleurs. La fonction

$$f(x) = \begin{cases} 0 & \text{si } x \notin \mathbb{Q} \\ \frac{1}{q} & \text{si } x = p/q \end{cases} \tag{12.140}$$

où par « $x = p/q$ » nous entendons que p/q est la fraction irréductible [26].

Cette fonction est discontinue sur $\mathbb{Q}$ parce que si $q \in \mathbb{Q}$ alors $f(q) \neq 0$ alors que dans tous voisinage de q il existe un irrationnel sur qui la fonction vaudra zéro.

Montrons que f est continue sur les irrationnels. Si $x_0 \notin \mathbb{Q}$ alors $f(x_0) = 0$. Mais si on prend un voisinage suffisamment petit de x_0, nous pouvons nous arranger pour que tous les rationnels aient un dénominateur arbitrairement grand. En effet si nous nous fixons un premier rayon $r_0 > 0$ alors il existe un nombre fini de fractions de la forme 1, $\frac{k}{2}$, $\frac{k}{3}$,..., $\frac{k}{N}$ dans $B(x_0, r_0)$. Il suffit maintenant de choisir $0 < r \leqslant r_0$ tel que ces fractions soient toutes hors de $B(x_0, r)$. Dans cette boule nous avons $f < \frac{1}{N}$. Du coup f est continue en x_0. △

Définition 12.68 (Point périodique[379])**.**
Soit $f\colon I \to I$ une application d'un ensemble I dans lui-même. Si $x \in I$ vérifie $f^n(x) = x$ et $f^k(x) \neq x$ pour $k = 1, \ldots, n-1$ alors on dit que x est un point n-périodique.

26. Proposition 3.33.

LemAONBooGZBuYt

Lemme 12.69.
Soit I un segment[27] *de $\mathbb{R}$ et une fonction continue $f\colon I \to I$. Si K est un segment fermé avec $K \subset f(I)$ alors il existe un segment fermé $L \subset I$ tel que $K = f(L)$.*

Démonstration. Mentionnons immédiatement que f est continue sur I qui est compact[28]. Par conséquent tous les nombres dont nous allons parler sont finis parce que f est bornée par le théorème 9.55.

Soit $K = [\alpha, \beta]$. Si $\alpha = \beta$ alors le segment $L = \{a\}$ convient. Nous supposons donc que $\alpha \neq \beta$ et nous considérons $a, b \in I$ tels que $\alpha = f(a)$ et $\beta = f(b)$. Puisque $a \neq b$ nous supposons $a < b$ (le cas $a > b$ se traite de façon similaire).

Nous posons

$$A = \{x \in [a, b] \text{ tel que } f(x) = \alpha\}. \tag{12.141}$$

C'est un ensemble borné par a et b. De plus il est fermé ; ce dernier point n'est pas tout à fait évident parce que f n'est pas définie sur $\mathbb{R}$, mais sur I, qui est fermé. Le corolaire 9.90 n'est donc pas immédiatement utilisable. Prouvons donc que $Z = \{x \in \mathbb{R} \text{ tel que } f(x) = \alpha\}$ est fermé. Si x_0 est hors de Z alors, soit x_0 est dans I, soit il est hors de I. Dans ce second cas, le complémentaire de I étant ouvert, on a un voisinage de x_0 hors de I, et par conséquent hors de Z. Si au contraire $x_0 \in I$ alors il y a (encore) deux cas : soit $x_0 \in \text{Int}(I)$, soit x_0 est sur le bord de I. Dans le premier cas, le théorème des valeurs intermédiaires[29] fonctionne. Pour le second cas, nous supposons $x_0 = \max(I)$ (le cas $x_0 = \min(I)$ est similaire). Le théorème des valeurs intermédiaires dit que sur $[x_0 - \epsilon, x_0]$, $f \neq \alpha$ et en même temps, sur $]x_0, x_0 + \epsilon]$, nous sommes en dehors du domaine. Au final $\{f(x) = \alpha\}$ est fermé et A est alors fermé en tant que intersection de deux fermés.

L'ensemble A étant non vide ($a \in A$), il possède donc un maximum que nous nommons u :

$$u = \max(A). \tag{12.142}$$

Nous posons aussi

$$B = \{x \in [u, b] \text{ tel que } f(x) = \beta\} \tag{12.143}$$

qui est encore fermé, borné et non vide. Nous pouvons donc définir

$$v = \min(B). \tag{12.144}$$

Nous prouvons maintenant que $f\big([u, v]\big) = [\alpha, \beta]$. D'abord $f\big([u, v]\big)$ est un intervalle compact[30] contenant $f(u) = \alpha$ et $f(v) = \beta$. Par conséquent $[\alpha, \beta] \subset f\big([u, v]\big)$. Pour l'inclusion inverse supposons $t \in [u, v]$ tel que $f(t) > \beta$. Vu que $f(a) = \alpha$ et $\alpha < \beta$, par le théorème des valeurs intermédiaires, il existe $t_0 \in [a, t]$ tel que $f(t_0) = \beta$. Cela donne $t_0 < v$ et donc contredit la minimalité de v dans B. Nous en déduisons que $f\big([u, v]\big)$ ne contient aucun élément plus grand que β. Même jeu pour montrer que l'intervalle ne contient aucun élément plus petit que α.

En définitive, le segment $L = [u, v]$ satisfait toutes les exigences. □

Lorsque $I_2 \subset f(I_1)$ nous notons $I_1 \to I_2$ ou, si une ambiguïté est à craindre, $I_1 \xrightarrow{f} I_2$. Cette flèche se lit « recouvre ».

LemSSPXooMkwzjb

Lemme 12.70 ([380, 379])**.**
Soient les segments $I_0, \dots, I_{n-1}$ tels que nous ayons le cycle

$$I_0 \to I_1 \to \dots \to I_{n-1} \to I_0. \tag{12.145}$$

Alors f^n admet un point fixe $x_0 \in I_0$ tel que $f^k(x_0) \in I_k$ pour tout $k = 0, \dots, n-1$.

Démonstration. Nous prouvons les cas $n = 1$ et $n = 2$ séparément.

27. définition 9.52. Un segment est un intervalle fermé borné.
28. Par le lemme 9.20.
29. Théorème 9.93.
30. Corolaire 9.95 et théorème 7.220.

(1) $\underline{n=1}$ Nous avons $I_0 \to I_0$, c'est-à-dire que $I_0 \subset f(I_0)$. Si $I_0 = [a,b]$ alors nous posons $a = f(\alpha)$ et $b = f(\beta)$ pour certains $\alpha, \beta \in I_0$. Nous posons ensuite $g(x) = f(x) - x$.
Dans un premier temps, $g(\alpha) = a - \alpha \leqslant 0$ parce que $a = \min(I_0)$ et $\alpha \in I_0$. Pour la même raison, $g(\beta) = b - \beta \geqslant 0$. Le théorème des valeurs intermédiaires donne alors $t_0 \in [\alpha, \beta] \subset I_0$ tel que $g(t_0) = 0$. Nous avons donc $f(t_0) = t_0$.

(2) $\underline{n=2}$ Nous avons $I_0 \to I_1 \to I_0$. Puisque $I_1 \subset f(I_0)$, le lemme 12.69 donne un segment $J_1 \subset I_0$ tel que $f(J_1) = I_1$. Mézalors

$$J_1 \subset I_0 \subset f(I_1) = f^2(J_1). \tag{12.146}$$

Nous avons donc $J_1 \xrightarrow{f^2} J_1$ et par le cas $n = 1$ traité plus haut, la fonction f^2 a un point fixe x_0 dans J_1. De plus

$$f(x_0) \in f(J_1) = I_1, \tag{12.147}$$

le point x_0 est donc bien celui que nous cherchions.

(3) **Cas général** Nous avons

$$I_0 \to I_1 \to \ldots \to I_{n-1} \to I_0. \tag{12.148}$$

Puisque $I_1 \subset f(I_0)$, il existe $J_1 \subset I_0$ tel que $f(J_1) = I_1$. Mais

$$I_2 \subset f(I_1) = f^2(J_1), \tag{12.149}$$

donc il existe $J_2 \subset J_1$ tel que $I_2 = f^2(J_2)$. En procédant ainsi aussi longtemps qu'il le faut, nous construisons les ensembles $J_1, \ldots, J_{n-1}$ tels que

$$J_{n-1} \subset J_{n-2} \subset \ldots \subset J_1 \subset J_0 \tag{12.150}$$

tels que $I_k = f^k(J_k)$ pour tout $k = 1, \ldots, n-1$. La dernière de ces inclusions est $I_{n-1} = f^{n-1}(J_{n-1})$, mais $I_{n-1} \to I_0$, c'est-à-dire que

$$I_0 \subset f(I_{n-1}) = f^n(J_{n-1}), \tag{12.151}$$

et il existe $J_n \subset J_{n-1}$ tel que $I_0 \subset f^n(J_n)$. Mais comme $J_n \subset J_0$ nous avons en particulier $J_n \subset f^n(J_n)$.
Cela donne un point fixe $x_0 \in J_n$ pour f^n. Par construction, nous avons $J_n \subset J_{n-1} \subset \ldots \subset J_1 \subset J_0$ et donc $x_0 \in J_k$ pour tout k. En particulier

$$f^k(x_0) \in f^k(J_k) = I_k \tag{12.152}$$

pour tout k.

□

Théorème 12.71 (Théorème de Sarkowski[380, 379])**.**
Soit I, un segment de $\mathbb{R}$ et une application continue $f : I \to I$. Si f admet un point 3-périodique, alors f admet des points n-périodiques pour tout $n \geqslant 1$.

Démonstration. Soit $a \in I$ un point 3-périodique pour f et notons $b = f(a)$, $c = f(b)$. Les points b et c sont également des points 3-périodiques. Quitte à renommer, nous pouvons supposer que a est le plus petit des trois. Il reste deux possibilités : $a < b < c$ et $a < c < b$. Nous traitons d'abord le premier cas.

Supposons $a < b < c$. Nous posons $I_0 = [a,b]$ et $I_1 = [b,c]$. Nous avons immédiatement $I_1 \subset f(I_0)$ et comme $f(b) = c$ et $f(c) = a$, $f(I_1)$ recouvre $[a,c]$ et donc recouvre en même temps I_1 et I_2. Nous avons donc $I_0 \to I_1$, $I_1 \to I_0$ et $I_1 \to I_1$.

(1) **Un point 1-périodique** Nous avons $I_1 \to I_1$ qui prouve que f a un point fixe dans I_1. C'est le cas $n = 1$ du lemme 12.70. Voilà un point 1-périodique.

(2) **Un point 2-périodique** Nous avons $I_0 \to I_1 \to I_0$. Par conséquent, le lemme 12.70 dit que f^2 a un point fixe $x_0 \in I_0$ tel que $f(x_0) \in I_1$. Montrons que $f(x_0) \neq x_0$. Pour avoir $x_0 = f(x_0)$, il faudrait $x_0 \in I_0 \cap I_1 = \{b\}$. Mais b est un point 3-périodique, donc ne vérifiant certainement pas $f^2(b) = b$. Nous en déduisons que $f(x_0) \neq x_0$ et donc que x_0 est 2-périodique.

(3) **Un point 3-périodique** On en a par hypothèse.

(4) **Un point n-périodique pour $n \geqslant 4$** Nous avons le cyle

$$I_0 \to \underbrace{I_1 \to I_1 \to \ldots \to I_1}_{\text{n-1}fois} \to I_0. \tag{12.153}$$

Le lemme donne alors un point fixe $x \in I_0$ pour f^n tel que $f^k(x) \in I_1$ pour $k = 1, \ldots, n-1$. Est-ce possible que $x = b$? Non parce que $f^2(b) = a \in I_0$ alors que $f^2(x) \in I_1$. Mais $I_0 \cap I_1 = \{b\}$.

Par conséquent la relation $f^k(x) \in I_1$ exclut d'avoir $f^k(x) = x$, et le point x est bien n-périodique.

Passons au cas $a < c < b$. Alors nous posons $I_0 = [a, c]$ et $I_1 = [c, b]$. Encore une fois $f(I_0)$ contient a et b, donc $I_0 \to I_0$ et $I_0 \to I_1$. Mais en même temps $f(I_1)$ contient a et c, donc $I_1 \to I_0$.

Nous pouvons donc refaire comme dans le premier cas, en inversant les rôles de I_0 et I_1. En particulier nous pouvons considérer le cycle

$$I_1 \to I_0 \to I_0 \to \ldots \to I_0 \to I_1. \tag{12.154}$$

□

12.7 Uniforme continuité

SECooZSZUniffCYSDRO

L'uniforme continuité est la définition 7.323.

Définition 12.72.
*Une partie $A \subset \mathbb{R}^m$ est dite **bornée** si il existe un $M > 0$ tel que $A \subset B(0, M)$. Le **diamètre** de la partie A est le nombre*

$$\mathrm{Diam}(A) = \sup_{x,y \in A} \|x - y\| \in [0, \infty]. \tag{12.155}$$

Lorsque A est bornée, il existe un M tel que $\|x\| \leqslant M$ pour tout $x \in A$.

LEMooFABOooZYFjhv

Lemme 12.73.
Si A est une partie non vide et bornée de $\mathbb{R}^m$, alors $\mathrm{Diam}(A) = \mathrm{Diam}(\bar{A})$.

Démonstration. Vu que $A \subset \bar{A}$, nous avons l'inégalité $\mathrm{Diam}(A) \leqslant \mathrm{Diam}(\bar{A})$. Nous nommons $d = \mathrm{Diam}(\bar{A})$. Soit $\epsilon > 0$. Soient $x, y \in \bar{A}$ tels que $\|x - y\| > d - \epsilon$. Nous prenons $a \in B(x, \epsilon) \cap A$ et $b \in B(y, \epsilon) \cap A$.

Nous avons

$$\mathrm{Diam}(A) \geqslant \|a - b\| \geqslant \|x - y\| - 2\epsilon \geqslant d - 3\epsilon. \tag{12.156}$$

Nous avons donc $\mathrm{Diam}(A) \geqslant \mathrm{Diam}(\bar{A}) - 3\epsilon$ pour tout ϵ. Donc $\mathrm{Diam}(A) \geqslant \mathrm{Diam}(\bar{A})$. □

Exemple 12.74.
Prenons la fonction $f(x) = \frac{1}{x}$, et demandons nous pour quel δ nous sommes sûr d'avoir

$$|f(a + \delta) - f(a)| = \left| \frac{1}{a + \delta} - \frac{1}{a} \right| < \varepsilon. \tag{12.157}$$

Pour simplifier, nous supposons que $a > 0$. Nous calculons

$$\begin{aligned} \frac{1}{a} - \frac{1}{a+\delta} &< \varepsilon \\ \frac{\delta}{a(a+\delta)} &< \varepsilon \\ \delta &< \varepsilon a^2 + \varepsilon a\delta \\ \delta(1-\varepsilon a) &< \varepsilon a^2 \\ \delta &< \frac{\varepsilon a^2}{1-\varepsilon a}. \end{aligned} \tag{12.158}$$

Notons que, à ε fixé, plus a est petit, plus il faut choisir δ petit. La fonction $x \mapsto \frac{1}{x}$ n'est donc pas uniformément continue. Cela correspond au fait que, proche de zéro, la fonction monte très vite. Une fonction uniformément continue sera une fonction qui ne montera jamais très vite. △

PROPooVOUTooOtiGLG

Proposition 12.75.
Quelques propriétés des fonctions uniformément continues.

(1) *Toute application uniformément continue est continue ;*
(2) *la composée de deux fonctions uniformément continues est uniformément continue ;*

Nous verrons qu'une application lipschitzienne est uniformément continue (proposition 12.321).

Une fonction peut être uniformément continue sur un domaine et pas sur un autre. Le théorème suivant donne une importante indication à ce sujet.

ThoHeineContinueCompact

Théorème 12.76 (Heine)**.**
Soit K un compact de $\mathbb{R}^n$. Une fonction continue $f\colon \mathbb{R}^n \to \mathbb{R}^m$ est uniformément continue sur K.

Démonstration. Nous allons prouver ce théorème par l'absurde. Nous commençons par écrire la condition qui exprime que f n'est pas uniformément continue sur le compact K :

$$\exists \varepsilon > 0 \text{ tel que } \forall \delta > 0,\ \exists x, y \in K \text{ tels que } \|x-y\| < \delta \text{ et } \big|f(x) - f(y)\big| > \varepsilon. \tag{12.159}$$

En particulier (en prenant $\delta = \frac{1}{n}$ pour tout n), pour chaque n nous pouvons trouver x_n et y_n dans K qui vérifient simultanément les deux conditions suivantes :

$$\begin{cases} \|x_n - y_n\| < \dfrac{1}{n} & \text{(12.160a)} \\ \big|f(x_n) - f(y_n)\big| > \varepsilon. & \text{(12.160b)} \end{cases}$$

EqCond3107fxfyepsppt

Nous insistons que c'est le même ε pour chaque n. L'ensemble K étant compact, l'ensemble $K \times K$ est compact (théorème 7.314) et nous pouvons trouver une sous-suite convergente *du couple* (x_n, y_n) dans $K \times K$. Quitte à passer à ces sous-suites, nous supposons que (x_n, y_n) converge dans $K \times K$ et en particulier, que les suites (x_n) et (y_n) sont convergentes. Étant donné que pour chaque n elles vérifient $\|x_n - y_n\| < \frac{1}{n}$, les limites sont égales :

$$\lim x_n = \lim y_n = x. \tag{12.161}$$

L'ensemble K étant fermé, la limite x est dans K. Par continuité de f, nous avons finalement

$$\lim f(x_n) = \lim f(y_n) = f(x), \tag{12.162}$$

mais alors

$$\lim_{n\to\infty} \big|f(x_n) - f(y_n)\big| = 0, \tag{12.163}$$

ce qui est en contradiction avec le choix (12.160b).

Tout ceci prouve que $f(K)$ est bornée supérieurement et que f atteint son supremum (qui est donc un maximum). Le fait que $f(K)$ soit bornée inférieurement se prouve en considérant la fonction $-f$ au lieu de f. □

Remarque 12.77.
Nous pouvons ne pas utiliser le fait que le produit de compacts est compact.

Pour choisir les sous-suites (x_n) et (y_n), il suffit de prendre une sous-suite convergente de (x_n) et d'invoquer le fait que $\|x_n - y_n\| \leqslant \frac{1}{n}$. Les suites (x_n) et (y_n) étant adjacentes[31], la convergence de (x_n) implique la convergence de (y_n) vers la même limite.

Il est donc un peu superflu de parler de la convergence du couple (x_n, y_n).

PROPooBWUFooYhMlDp

Proposition 12.78 (Heine[381])**.**
Toute application continue d'un espace métrique compact dans un espace métrique quelconque est uniformément continue[32].

Démonstration. Soient un espace métrique compact X et un espace métrique quelconque E. Nous considérons une application continue $f\colon X \to E$.

(1) **Un ensemble** Soit $\epsilon > 0$. Nous considérons l'ensemble

$$K = \{(x, y) \in X \times X \text{ tel que } d\big(f(x), f(y)\big) \geqslant \epsilon\}. \tag{12.164}$$

(2) **Il est compact** L'espace X étant compact, $X \times X$ est également compact par le théorème 7.314. Les fonctions f et d étant continues, l'application

$$\begin{aligned} \varphi\colon X \times X &\to \mathbb{R}^+ \\ (x, y) &\mapsto d\big(f(x), f(y)\big) \end{aligned} \tag{12.165}$$

est continue, de telle sorte que la partie $\varphi \geqslant \epsilon$ est fermée. Un fermé dans un compact est compact par le lemme 7.96.

(3) **Une borne atteinte** Nous considérons l'application distance $d\colon K \to \mathbb{R}^+$. C'est une application continue sur le compact K ; donc elle atteint ses bornes (théorème des bornes atteintes, 9.55). Elle a un minimum que nous notons δ.

Comme $(x, x) \notin K$, nous avons $d(x, y) > 0$ pour tout $(x, y) \in K$. Et donc $\delta > 0$.

(4) **Conclusion** Si $x, y \in X$ sont tels que $d(x, y) < \delta$, alors $(x, y) \notin K$. De ce fait nous avons

$$d\big(f(x), f(y)\big) < \epsilon. \tag{12.166}$$

D'où l'uniforme continuité de f sur X.

□

LEMooIVAKooUiEENr

Lemme 12.79 ([1])**.**
Soit une fonction continue $f\colon [a, b] \times [c, d] \to \mathbb{R}$. Pour chaque $x \in [a, b]$ nous définissons

$$\begin{aligned} f_x\colon [c, d] &\to \mathbb{R} \\ y &\mapsto f(x, y). \end{aligned} \tag{12.167}$$

Alors l'application

$$\begin{aligned} g\colon [a, b] &\to \mathbb{R} \\ x &\mapsto \|f_x\|_\infty \end{aligned} \tag{12.168}$$

est continue.

Démonstration. Soit $\alpha \in [a, b]$, et prouvons la continuité de g en α.

(1) **Le décor** Nous considérons l'espace vectoriel normé $\big(C^0([a, b]), \|.\|_\infty\big)$ des fonctions continues sur $[a, b]$ muni de la norme uniforme. Prouvons que si $x_k \xrightarrow{[a,b]} \alpha$, alors $f_{x_k} \xrightarrow{unif} f_\alpha$.

31. Définition 9.39.
32. Uniforme continuité, définition 7.323.

(2) **Module de continuité** Pour cela nous avons le calcul suivant, avec justifications juste en-dessous :

$$\|f_{x_k} - f_\alpha\|_\infty = \sup_{y\in[c,d]} \|f(x_k,y) - f(\alpha,y)\| \tag{12.169a}$$
$$\leqslant \sup_{y\in[c,d]} \left|\omega_f\big(\|(x_k,y)-(\alpha,y)\|\big)\right| \tag{12.169b}$$
SUBEQooCOJQooWlvHUa
$$= \sup_{y\in[c,d]} |\omega_f(|x_k-\alpha|)| \tag{12.169c}$$
SUBEQooGLYMooEZKRKn
$$= \omega_f(|x_k-\alpha|). \tag{12.169d}$$

Justifications.

— Pour (12.169b). Utilisation du module de continuité, définition 10.310.

— Pour (12.169c). La norme dans $\mathbb{R}^2$ de $(x_k,y)-(\alpha,y)$.

(3) **Uniforme continuité** La fonction f est continue sur le compact $[a,b]\times[c,d]$. Elle est donc uniformément continue par le théorème de Heine 12.78, et donc son module de continuité vérifie $\lim_{h\to 0}\omega_f(h)=0$ par 10.312.
Nous avons donc

$$\|f_{x_k} - f_\alpha\|_\infty \leqslant \omega_f(|x_k-\alpha|) \xrightarrow{\mathbb{R}} 0. \tag{12.170}$$

Nous avons donc prouvé que si $x_k \xrightarrow{\mathbb{R}} \alpha$, alors $f_{x_k} \xrightarrow{\|.\|_\infty} f_\alpha$.

(4) **Conclusion** La norme étant une application continue, nous en déduisons que si $x_k \to \alpha$, alors $\|f_{x_k}\|_\infty \to \|f_\alpha\|_\infty$.
Ceci est la continuité séquentielle de la fonction g, et donc la continuité tout court.

□

12.8 Fonctions sur un compact

Par le théorème des valeurs intermédiaires 9.93, l'image d'un intervalle par une fonction continue est un intervalle, et nous avons l'importante propriété suivante des fonctions continues sur un compact.

Le théorème suivant est un cas particulier du théorème 9.55.

Théorème 12.80.
Si f est une fonction continue sur l'intervalle compact $[a,b]$. Alors f est bornée sur $[a,b]$ et elle atteint ses bornes.

Démonstration. Étant donné que $[a,b]$ est un intervalle compact, son image est également un intervalle compact, et donc est de la forme $[m,M]$. Ceci découle du théorème 7.220 et le corolaire 9.95. Le maximum de f sur $[a,b]$ est la borne M qui est bien dans l'image (parce que $[m,M]$ est fermé). Idem pour le minimum m. □

12.9 Polynômes, théorème de d'Alembert

L'algèbre des polynômes sur un anneau est définie en 1.359. Si $P\in A[X]$ et si $\alpha\in A$ nous avons également défini l'évaluation de P en α ; c'est la définition 1.362. Dans le cadre de l'analyse, lorsque nous considérons des polynômes, nous allons complètement confondre le polynôme avec la fonction qu'il définit.

12.9.1 Polynômes sur les réels

PROPooJKYJooFqbQMr

Proposition 12.81.
Tout polynôme à coefficients réels de degré impair possède au moins une racine réelle.

Démonstration. Nous mettons le plus haut degré en facteur :

$$P(x) = \sum_{k=0}^{n} a_k x^k = x^n \sum_{k=0}^{n} \frac{a_k}{x^{n-k}}. \tag{12.171}$$

Le terme en $k = n$ vaut $a_n x^n$ tandis que les autres sont de la forme (à un coefficient près) $\frac{1}{x^l}$ pour un $l \geqslant 1$. Lorsque $x \to \infty$, chacun de ces termes s'annule (lemme 12.29). Nous avons donc

$$\lim_{x\to\infty} P(x) = \infty, \tag{12.172}$$

et de même, n étant impair, $\lim_{x\to-\infty} P(x) = -\infty$. Le théorème des valeurs intermédiaires 9.93 nous donne alors l'existence d'un réel sur lequel P s'annule. □

PROPooARLAooMNnNsz

Proposition 12.82 ([1])**.**
À propos de réductibilité

(1) *Si $z \in \mathbb{C}$ est une racine d'un polynôme à coefficients réels, alors le conjugué complexe $\bar{z}$ est également une racine.*

(2) *Tout polynôme à coefficients réels de degré au moins 3 est réductible*[33] *sur $\mathbb{R}$.*

La proposition suivante donne une utilisation amusante de la notion de polynôme caractéristique[34].

PROPooKJWOooOjSFaA

Proposition 12.83 ([382])**.**
Soit un espace vectoriel E de dimension finie pour lequel il existe un endomorphisme $f : E \to E$ tel que $(f \circ f)(x) = -x$ pour tout $x \in E$. Alors la dimension de E est paire.

Démonstration. Cherchons les valeurs propres de f en résolvant l'équation $f(x) = \lambda x$. Nous appliquons f à cette égalité :

$$-x = \lambda f(x) = \lambda^2 x. \tag{12.173}$$

Donc λ ne peut pas être réel. Nous avons montré que f n'a pas de valeur propre réelle. Or le polynôme caractéristique de f est de degré égal à la dimension. Si la dimension est impaire, le polynôme caractéristique est de degré impair, et possède donc une racine réelle[35]. Autrement dit, l'absence de racines réelles au polynôme caractéristique indique une dimension paire. □

12.84.
Une autre preuve possible pour 12.83 est d'utiliser le déterminant : si la dimension de E est n nous avons :

$$\det(f^2) = \det(-\mathrm{Id}) = (-1)^n. \tag{12.174}$$

Donc $(-1)^n$ est positif, ce qui montre que n est pair.

12.9.2 Polynômes sur les complexes

Nous allons parler de comportement asymptotique de polynômes définis sur $\mathbb{C}$. La topologie que nous considérons est celle de la compactification en un point, décrite en 7.106.

Le lemme suivant donne une caractérisation de la limite en l'infini dans le compactifié $\hat{\mathbb{C}}$. Dans beaucoup de cas, cette caractérisation est prise comme la définition de la limite. Hélas, dans le Frido nous sommes des extrémistes et nous ne parvenons pas à dire le mot « limite » si il n'y a pas une topologie.

33. Définition 1.189.
34. Définition 8.113.
35. Proposition 12.81.

LEMooERABooQjLBzW

Lemme 12.85 ([1]).
Nous considérons la compactification en un point d'Alexandrov[36]. *Soit une fonction* $f\colon \mathbb{C} \to \mathbb{C}$. *Nous avons* $\lim_{z\to\infty} f(z) = \infty$ *si et seulement si pour tout* $M > 0$, *il existe* $R > 0$ *tel que* $|z| > R$ *implique* $|f(z)| > M$.

Démonstration. Souvenons-nous que, en général[37], nous avons

$$\lim_{x\to a} f(x) = b \tag{12.175}$$

si pour tout voisinage V de b, il existe un voisinage W de a tel que $z \in W\backslash\{a\}$ implique $f(z) \in V$.

Précisons encore un point de notation. Si K est une partie de $\mathbb{C}$, nous notons K^c son complémentaire dans $\mathbb{C}$, pas dans $\hat{\mathbb{C}}$.

Ceci étant dit, nous passons à la preuve.

(1) **Sens direct** Nous supposons que $\lim_{z\to\infty} f(z) = \infty$. Soit $M > 0$; nous considérons le voisinage $V = \overline{B(0,M)}^c \cup \{\infty\}$. Par définition de la limite, il existe un voisinage W de ∞ tel que $z \in W \Rightarrow f(z) \in V\backslash\{\infty\} = \overline{B(0,M)}^c$. Ce voisinage est de la forme $K^c \cup \{\infty\}$. Puisque K est compact, il est borné, et il existe $R > 0$ tel que $K \subset B(0,R)$.

Avec tout cela nous avons la chaine suivante d'implications :

$$|z| > R \Rightarrow z \in K^c \Rightarrow z \in W \Rightarrow f(z) \in V\backslash\{\infty\} = \overline{B(0,M)}^c \Rightarrow |f(z)| > M. \tag{12.176}$$

C'est bien la propriété que nous voulions.

(2) **Sens réciproque** Soit un voisinage V de ∞. Nous avons $V = K^c \cup \{\infty\}$ où K est compact dans $\mathbb{C}$. Il existe $M > 0$ tel que $K \subset B(0,M)$.

Par hypothèse, il existe R tel que $|z| > R \Rightarrow |f(z)| > M$. Soit $W = \overline{B(0,R)}^c \cup \{\infty\}$. Nous avons la chaine

$$z \in W \Rightarrow |z| > R \Rightarrow |f(z)| > M \Rightarrow f(z) \in K^c \Rightarrow f(z) \in V. \tag{12.177}$$

□

PROPooPWVWooGuftxZ

Proposition 12.86 ([1]).
Soit le polynôme

$$\begin{aligned} P\colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \sum_{i=0}^{n} a_i z^i \end{aligned} \tag{12.178}$$

où nous sous-entendons que $a_n \neq 0$. *La fonction* $z \mapsto |P(z)|$ *est équivalente*[38] *en l'infini à la fonction*

$$\begin{aligned} w\colon \mathbb{C} &\to \mathbb{R}^+ \\ z &\mapsto |a_n z^n|. \end{aligned} \tag{12.179}$$

Démonstration. Nous voudrions prouver qu'il existe une fonction $\alpha\colon \mathbb{C} \to \mathbb{R}$ telle que [EQooGXWZooDJZNzE]

$$\begin{cases} |\sum_{i=0}^{n} a_i z^i| = \big(1 + \alpha(z)\big)|a_n z^n|. & \text{(12.180a)} \\ \lim_{z\to\infty} \alpha(z) = 0. & \text{(12.180b)} \end{cases}$$

Nous trouvons un candidat pour être une telle fonction en isolant simplement $\alpha(z)$ de cette égalité. Nous trouvons

$$\alpha(z) = \left|\sum_{i=0}^{n} \frac{a_i}{a_n} z^{i-n}\right| - 1. \tag{12.181}$$

36. Définition 7.106.
37. Définition 7.110.
38. Définition 7.68.

Elle vérifie immédiatement (12.180). Le point qui fait intervenir la topologie de $\hat{\mathbb{C}}$ est de vérifier que $\lim_{z\to\infty}\alpha(z)=0$. Le terme $i=n$ de la somme vaut 1. Il suffit donc de montrer que pour $i\neq n$ nous avons

$$\lim_{z\to\infty}\frac{1}{z^{n-i}}=0. \tag{12.182}$$

Soit $\epsilon>0$. Nous devons prouver qu'il existe un voisinage V de ∞ dans $\hat{\mathbb{C}}$ tel que

$$\left|\frac{1}{z^{n-i}}-0\right|\leqslant\epsilon \tag{12.183}$$

pour tout $z\in V$.

En utilisant la proposition 9.106 nous avons déjà

$$\left|\frac{1}{z^{n-i}}\right|=\frac{1}{|z^{n-i}|}=\frac{1}{|z|^{n-i}}. \tag{12.184}$$

Soit $R>0$ tel que $\frac{1}{R}<\epsilon$. Nous considérons le voisinage $\{|z|>R\}\cup\{\infty\}$ de ∞. Dans ce voisinage, nous avons

$$\frac{1}{|z|^{n-i}}\leqslant\frac{1}{|z|}\leqslant\frac{1}{R}<\epsilon. \tag{12.185}$$

Et voilà. □

Le lemme suivant parle de polynôme sur $\mathbb{C}$. Vous pouvez l'adapter à $\hat{\mathbb{R}}$ et $\bar{\mathbb{R}}$.

LEMooYZVGooXZvBAc

Lemme 12.87.
Si $P\colon\mathbb{C}\to\mathbb{C}$ est un polynôme, alors $|P|$ atteint une borne inférieure globale.

Démonstration. Nous savons, par l'équivalence de fonctions prouvée dans la proposition 12.86 que $\lim_{z\to\infty}P(z)=\infty$. Soit $a>0$ dans $\mathbb{R}$. Par le lemme 12.85 il existe un $R>a$ tel que $|z|>R\Rightarrow|f(z)|>|f(a)|$.

La fonction $|P|$ est continue sur le compact $\overline{B(0,R)}$. Soit z_0 le point de minimum [39] de $|P|$ sur $\overline{B(0,R)}$.

Nous devons prouver que z_0 donne même un minimum global. Comme $a\in\overline{B(0,R)}$ nous avons

$$|P(z_0)|\leqslant|P(a)|. \tag{12.186}$$

Si $z\in\overline{B(0,R)}^c$, nous avons

$$|P(z)|>|P(a)|\geqslant|P(z_0)|. \tag{12.187}$$

Donc ce z_0 est un minimum sur $B(0,R)$ et sur $\overline{B(0,R)}^c$. Bref, un minimum global. □

LEMooTTOYooXaukuH

Lemme 12.88.
Soit le polynôme

$$\begin{aligned}P\colon\mathbb{C}&\to\mathbb{C}\\ z&\mapsto\sum_{i=0}^{n}a_iz^i.\end{aligned} \tag{12.188}$$

À part si $a_0=a_1=0$, la fonction P est équivalente à a_0+a_1z en $z=0$.

Note : si $a_0=a_1=0$, alors évidemment P devrait être équivalente à 0, ce qui est évidemment faux : la fonction identiquement nulle tend vers zéro plus vite que n'importe quoi.

Démonstration. En posant $g(z)=a_0+a_1z$, nous devons trouver une fonction α telle que

$$P(z)=\big(1+\alpha(z)\big)g(z). \tag{12.189}$$

EQooZFJBooVAYVBv

39. Théorème de Weierstrass 7.146.

Si $a_0 \neq 0$, il existe un voisinage épointé de $z = 0$ sur lequel la fonction

$$\alpha(z) = \frac{z^2 \sum_{i=2}^{n} a_i z^{i-2}}{a_0 + a_1 z} \tag{12.190}$$

EQooVCOVooAKWJxF

existe. Il n'y a aucun problème à ce que $\alpha(z) \to 0$ pour $z \to 0$[40], et un simple calcul[41] donne (12.190).

Si par contre $a_0 = 0$, nous faisons le calcul intermédiaire suivant :

$$\alpha(z)g(z) = P(z) - g(z) = z^2 \sum_{i=2}^{n} a_i z^{i-2}, \tag{12.191}$$

et donc, en isolant $\alpha(z)$ et en simplifiant par z, nous voyons que la fonction α définie par

$$\alpha(z) = \frac{z}{a_1} \sum_{i=2}^{n} a_i z^{i-2} \tag{12.192}$$

convient. □

PROPooLBBLooQwEiHr

Proposition 12.89 ([383, 1]).
Soient $a, b \in \mathbb{R}$.

ITEMooSPSWooKLtqzZ

(1) L'équation $z^2 = a + bi$ a une solution dans $\mathbb{C}$.

ITEMooQOJDooWjfGXv

(2) Pour tout l, l'équation $z^{2^l} = a + bi$ a une solution dans $\mathbb{C}$.

Nous ne disons pas que ces solutions sont uniques[42].

Démonstration. Pour prouver (1), l'équation $z^2 = a + bi$ a pour solution $\pm\xi$ où[43]

$$\xi = \sqrt{\frac{1}{2}a + \frac{1}{2}\sqrt{a^2 + b^2}} + i \operatorname{sgn}(b) \sqrt{-\frac{1}{2}a + \frac{1}{2}\sqrt{a^2 + b^2}}. \tag{12.193}$$

Nous n'avons en fait pas besoin de montrer que $\pm\xi$ sont toutes deux des solutions, ni que ce sont les seules. Un calcul direct montre que $\xi^2 = a + bi$ et nous sommes contents.

Pour (2), nous faisons une récurrence sur l. Nous savons que

$$z^{2^{k+1}} = (z^{2^k})^2. \tag{12.194}$$

Soit $\xi \in \mathbb{C}$ tel que $\xi^{2^k} = a + bi$; un tel ξ existe par hypothèse de récurrence. Alors si z est tel que $z^2 = \xi$, nous avons

$$z^{2^{k+1}} = a + bi. \tag{12.195}$$

□

Le théorème de d'Alembert possède de nombreuses démonstrations. En voici une qui à ma connaissance est celle demandant le moins d'analyse ; une démonstration à base de théorie de Galois peut être trouvée dans [100, 384]. Si vous lisez ces lignes pour savoir qu'un polynôme de degré n possède au *maximum* n racines, ce n'est pas ici qu'il faut regarder, mais le corolaire 3.131.

Une très jolie démonstration sur numberphile :
https://www.youtube.com/watch?v=shEk8sz1oOw.

THOooIRJYooBiHRyW

Théorème 12.90 (d'Alembert[383]).
Le corps $\mathbb{C}$ est algébriquement clos[44] *: tout polynôme non constant à coefficients complexes admet au moins une racine complexe*[45].

40. En remarquant toutefois que c'est une limite à deux dimensions. Sachez la définir.
41. En fait, la formule (12.190) est obtenue en isolant $\alpha(z)$ dans (12.189).
42. Comme vous en conviendrez en pensant à $z^2 = 1$ qui a déjà les solutions 1 et -1.
43. Si vous vous demandez où sont définies les racines carrés, c'est 9.97.
44. Définition 6.74.
45. C'est la définition 6.74 d'être algébriquement clos.

Démonstration. Nous effectuons une preuve tout à la fois par l'absurde et par récurrence en supposant que le polynôme

$$\begin{aligned} f\colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto z^n + a_1 z^{n-1} + \ldots + a_n \end{aligned} \tag{12.196}$$

n'a pas de racine dans $\mathbb{C}$, et que n soit le plus petit entier pour lequel un tel polynôme existe. Nous notons

$$n = 2^k m \tag{12.197}$$

où m est impair.

Le lemme 12.87 donne un point z_0 qui réalise le minimum global de $|f|$ sur $\mathbb{C}$. Nous posons $g(z) = f(z_0 + z)$ et nous définissons ses coefficients A_i par

$$g(z) = \sum_{i=0}^{n} A_i z^i. \tag{12.198}$$

Nous avons $A_n = 1$ et $|A_0| = |f(z_0)|$. Soit A_r le premier à être non nul parmi les A_1, A_2, ...

(1) **Si $r < n$** Par hypothèse de récurrence, il existe $\xi \in \mathbb{C}$ tel que $\xi^r = -A_1/A_r$. Nous avons

$$g(t\xi) = A_0 + \frac{-A_r t^r A_0}{A_r} + t^{r+1} \sum_{i=r+1}^{n} A_i \xi^i t^{i-r-1}. \tag{12.199}$$

En notant $P(t)$ le dernier polynôme, nous pouvons écrire cela sous forme compacte :

$$g(t\xi) = A_0 - t^r A_0 + t^{r+1} P(t). \tag{12.200}$$

Puisque

$$\lim_{t\to 0} \frac{t^{r+1}P(t)}{t^r |A_0|} = \lim_{t\to 0} tP(t) = 0, \tag{12.201}$$

il existe $t_0 > 0$ tel que

$$|t_0^{r+1} P(t_0)| < |A_0 t_0^r|. \tag{12.202}$$

Nous choisissons de plus $t_0 < 1$, de telle sorte que $1 - t^r > 0$. Avec cela nous avons

$$|g(t\xi)| \leqslant |A_0|(1 - t^r) + |t^{r+1}P(t)| = |A_0| \underbrace{-t^r|A_0| + |t^{r+1}P(t)|}_{<0} < |A_0|. \tag{12.203}$$

Or $|A_0|$ était un minimum global de $|g|$. Contradiction.

(2) **Si $r = n$**

Dans ce cas,

$$g(z) = f(z_0 + z) = A_0 + z^n, \tag{12.204}$$

et nous rappelons que $n = 2^k m$ où m est impair. Nous allons trouver une contradiction dans les quatre cas $\operatorname{Re}(A_0) > 0$, $\operatorname{Re}(A_0) < 0$, $\operatorname{Im}(A_0) > 0$ et $\operatorname{Im}(A_0) < 0$. Bien entendu ces cas se recouvrent largement, mais en toute généralité, nous avons besoin des quatre.

(2a) **Si $\operatorname{Re}(A_0) > 0$** La proposition 12.89 nous permet de considérer $v \in \mathbb{C}$ tel que $v^{2^k} = -1$. Nous avons alors

$$g(tv) = A_0 + (tv)^n = A_0 + t^n (v^{2^k})^m = A_0 + t^n (-1)^m = A_0 - t^n \tag{12.205}$$

parce que m est impair. Nous avons $\operatorname{Im}\big(g(tv)\big) = \operatorname{Im}(A_0)$. Si t est assez petit pour que $t^n < |\operatorname{Re}(A_0)|$ nous avons aussi $|\operatorname{Re}\big(g(tv)\big)| < |\operatorname{Re}(A_0)|$. Donc

$$|g(tv)|^2 = |\operatorname{Re}\big(g(tv)\big)|^2 + |\operatorname{Im}\big(g(tv)\big)|^2 < |\operatorname{Re}(A_0)|^2 + |\operatorname{Im}(A_0)|^2 = |A_0|^2. \tag{12.206}$$

Donc $|g(tv)| < |A_0|$. Contradiction.

(2b) **Si $\operatorname{Re}(A_0) < 0$** Nous prenons $v = 1$, et même histoire.

(2c) **Si** $\text{Im}(A_0) < 0$ Nous prenons $w \in \mathbb{C}$ tel que

$$w^{2^k} = i(-1)^{\frac{1}{2}(m-1)}. \tag{12.207}$$

Là, il y a un peu d'arrachage de cheveux pour bien voir les cas. La difficulté est que les puissances de i alternent entre 1, -1, i et $-i$. Puisque m est impair, nous avons un l tel que $m = 2l + 1$. Nous subdivisons les cas l pair et l impair.

(2c.iii) **Si l est pair** Alors d'une part $\frac{1}{2}(m-1) = l$ est pair et donc

$$(-1)^{\frac{1}{2}(m-1)} = 1. \tag{12.208}$$

Et d'autre part, $i^{2l+1} = (-1)^l i = i$. En tout,

$$i^m(-1)^{\frac{1}{2}(m-1)} = i. \tag{12.209}$$

(2c.iii) **Si l est impair** Alors $\frac{1}{2}(m-1) = l$ et $(-1)^{\frac{1}{2}(m-1)} = -1$. Mais en même temps, $i^{2l+1} = -i$, ce qui donne encore une fois

$$i^m(-1)^{\frac{1}{2}(m-1)} = i. \tag{12.210}$$

Bref, que l soit pair ou impair, nous avons $i^m(-1)^{\frac{1}{2}(m-1)} = i$.

Nous avons donc $\text{Re}\big(g(tw)\big) = \text{Re}(A_0)$ et $\text{Im}\big(g(tw)\big) < \text{Im}(A_0)$. Encore contradiction.

(3) **Si** $\text{Im}(A_0) > 0$ Même chose que ce que nous venons de faire, mais avec

$$w^{2^k} = -i(-1)^{\frac{1}{2}(m-1)}. \tag{12.211}$$

□

CORooKKNWooWEQukb

Corolaire 12.91 ([1])**.**
Tout polynôme de degré 3 à coefficients réels possède au moins une racine réelle.

Démonstration. Soient les racines λ_1, λ_2 et λ_3 du polynôme en question. Toutes trois sont dans $\mathbb{C}$. Supposons que λ_1 ne soit pas réelle. Alors λ_2 ou λ_3 doit être égale à $\bar{\lambda}_1$. Disons λ_2. Nous avons donc les racines λ_1, $\bar{\lambda}_1$ et λ_3. Le polynôme se factorise alors en

$$a(X - \lambda_1)(X - \bar{\lambda}_1)(X - \lambda_3). \tag{12.212}$$

EQooELMMooNbpBgg

Le coefficient a doit être réel parce qu'il est le coefficient du terme en X^3 (réel par hypothèse). Si λ_3 n'est pas réel, alors ce polynôme ne peut pas avoir des coefficients réels. Entre autres parce que le terme indépendant est $a|\lambda_1|^2\lambda_3$, qui est réel si et seulement si λ_3 est réel [46]. □

Tant que vous y êtes, vous pouvez voir que le polynôme (12.212) est à coefficients réels si et seulement si $a \in \mathbb{R}$ et $\lambda_3 \in \mathbb{R}$.

EXooIPLOooSNfiWg

Exemple 12.92.
Toute application linéaire $\mathbb{R}^3 \to \mathbb{R}^3$ a un vecteur propre. En effet si $R\colon \mathbb{R}^3 \to \mathbb{R}^3$ est linéaire, son polynôme caractéristique χ_R est de degré 3. Le corolaire 12.91 indique qu'un tel polynôme possède au moins une racine réelle. Une telle racine est une valeur propre de R par le théorème 8.117. △

DEFooVFTYooJyDePn

Définition 12.93.
Si $\lambda \in \mathbb{K}$ est une racine de χ_u, l'ordre de l'annulation est la ***multiplicité algébrique*** *de la valeur propre λ de u. À ne pas confondre avec la* ***multiplicité géométrique*** *qui sera la dimension de l'espace propre.*

46. Notez l'utilisation du lemme 9.107.

LEMooPJKEooQwuWMb

Lemme 12.94.
Si $\alpha \in \mathbb{C}$ n'est pas réel, alors les polynômes $X-\alpha$ et $X-\bar{\alpha}$ sont premiers entre eux dans $\mathbb{C}[X]$.

Démonstration. Un polynôme diviseur commun de $X-\alpha$ et $X-\bar{\alpha}$ doit être de la forme $aX+b$ parce qu'un diviseur ne peut pas être de degré plus haut. Si nous voulons que le diviseur commun soit non inversible, il faut $a \neq 0$.

Dire que $aX+b$ divise $X-\alpha$ signifie qu'il existe $s \in \mathbb{C}$ tel que $(aX+b)s = X-\alpha$. Cela donne $s = 1/a$ et $b = -\alpha/s$. En substituant la première dans la seconde, nous trouvons $b = -\alpha a$.

En recommençant avec $X-\bar{\alpha}$, nous trouvons également la relation $b = -\bar{\alpha}a$. En égalisant,

$$-\alpha a = -\bar{\alpha}a. \tag{12.213}$$

En simplifiant, $\alpha = \bar{\alpha}$. Cela contredit l'hypothèse que α n'est pas réel. □

PROPooUMDQooVmfDYU

Proposition 12.95.
Un polynôme irréductible à coefficients réels est, soit de degré un, soit de degré 2 avec un discriminant négatif.

Démonstration. Soit un polynôme P à coefficients réels de degré plus grand que 1. Alors le théorème de d'Alembert-Gauss (théorème 12.90) implique l'existence d'une racine $\alpha \in \mathbb{C}$. Si α est un réel, P est réductible. Si α n'est pas réel, alors son conjugué complexe $\bar{\alpha}$ est également une racine. Par conséquent les polynômes $(X-\alpha)$ et $(X-\bar{\alpha})$ divisent P dans $\mathbb{C}[X]$..

Ces deux polynômes sont premiers entre eux par le lemme 12.94. Par conséquent le produit

$$X^2 - (\alpha+\bar{\alpha})X + \alpha\bar{\alpha} \tag{12.214}$$

divise également P. Ce dernier est un polynôme à coefficients réels de degré 2. Donc tout polynôme de degré 3 ou plus est réductible. □

PROPooLXGSooXmVcVG

Proposition 12.96.
Si E est un espace vectoriel sur $\mathbb{C}$, tout endomorphisme possède au moins une valeur propre.

Démonstration. Soit un endomorphisme u sur E. Le théorème 8.117 dit que $\lambda \in \mathbb{C}$ est une valeur propre si et seulement si λ est une racine du polynôme caractéristique χ_u. Or ce polynôme possède au moins une racine dans $\mathbb{C}$ par le théorème de d'Alembert 12.90. □

12.10 Trigonalisation

12.10.1 Trigonalisation : généralités

SUBSECooMCOGooEoQCsz

Définition 12.97 ([385])**.**
*Une matrice dans $\mathbb{M}(n,\mathbb{K})$ est **trigonalisable** lorsqu'elle est semblable*[47] *à une matrice triangulaire supérieure.*

PropKNVFooQflQsJ

Proposition 12.98 (Trigonalisation et polynôme caractéristique scindé)**.**
Soit u un endomorphisme d'un espace vectoriel E sur le corps $\mathbb{K}$. Les points suivants sont équivalents.

ItemZKDMooOrTHkwi

(1) L'endomorphisme u est trigonalisable (auquel cas les valeurs propres sont sur la diagonale).

ItemZKDMooOrTHkwii

(2) Le polynôme caractéristique de u est scindé[48]*.*

Démonstration. En deux parties.

47. Définition 8.202.
48. Définition 6.39.

(1) **(2)⇒(1)** Nous avons par hypothèse

$$\chi_u(X) = \prod_{i=1}^{r}(X-\lambda_i)^{\alpha_i} \tag{12.215}$$

où les λ_i sont les valeurs propres de u. Le théorème de Cayley-Hamilton 8.120 dit que $\chi_u(u) = 0$, ce qui permet d'utiliser le théorème de décomposition des noyaux 8.92 :

$$E = \ker(X-\lambda_1)^{\alpha_1} \oplus \ldots \oplus \ker(X-\lambda_r)^{\alpha_r}. \tag{12.216}$$

Les espaces $F_{\lambda_i}(u) = \ker(u-\lambda_i)^{\alpha_i}$ sont les espaces caractéristiques de u, ce qui fait que $u-\lambda_i\mathbb{1}$ est nilpotent sur $F_{\lambda_i}(u)$. L'endomorphisme $u-\lambda_i\mathbb{1}$ est donc strictement trigonalisable supérieur sur son bloc [49]. Cela signifie que u est triangulaire supérieure avec les valeurs propres sur la diagonale.

(2) **(1)⇒(2)**
C'est immédiat parce que le déterminant d'une matrice triangulaire est le produit des éléments de sa diagonale.

□

Remarque 12.99.
La méthode des pivots de Gauss [50] certes permet de trigonaliser n'importe quelle matrice, mais elle ne correspond pas à un changement de base. Autrement dit, les pivots de Gauss ne sont pas des similitudes.

C'est là qu'il faut bien avoir en tête la différence entre *équivalence* et *similarité* [51]. Lorsqu'on parle de changement de base, de matrice trigonalisable ou diagonalisable, nous parlons de similarité et non d'équivalence.

12.10.2 Trigonalisation : cas complexe

La proposition 12.98 dit déjà que tous les endomorphismes sont trigonalisables sur $\mathbb{C}$. Nous allons aller plus loin et montrer que la trigonalisation peut être effectuée à l'aide d'une matrice unitaire.

Une démonstration alternative passant par le polynôme caractéristique sera présentée dans la remarque 12.108 utilisant la proposition 12.98.

LemSchurComplHAftTq

Lemme 12.100 (Lemme de Schur complexe, trigonalisation[386])**.**
Si $A \in \mathbb{M}(n,\mathbb{C})$, il existe une matrice unitaire [52] *U telle que UAU^{-1} soit triangulaire supérieure* [53]*. La diagonale de la matrice triangulaire contient alors les valeurs propres de A.*

Démonstration. Étant donné que $\mathbb{C}$ est algébriquement clos [54], nous pouvons toujours considérer un vecteur propre v_1 de A, de valeur propre λ_1. Nous pouvons utiliser un procédé de Gram-Schmidt pour construire une base orthonormée $\{v_1, u_2, \ldots, u_n\}$ de $\mathbb{C}^n$, et la matrice (unitaire)

$$Q = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & u_2 & \cdots & u_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}. \tag{12.217}$$

49. Proposition 8.212.
50. Le lemme 4.112.
51. Définition 4.111.
52. Définition 8.179.
53. « triangulaire supérieure » ne signifie pas « strictement triangulaire supérieure ». Ici, il est possible que la diagonale soit non nulle ; non seulement possible, mais même très probable en pratique.
54. Algébriquement clos, définition 6.74. Le fait que $\mathbb{C}$ le soit est le théorème de d'Alembert 12.90.

Nous avons $Q^{-1}AQe_1 = Q^{-1}Av_1 = \lambda Q^{-1}v = \lambda_1 e_1$, par conséquent la matrice $Q^{-1}AQ$ est de la forme

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix} \tag{12.218}$$

où $*$ représente une ligne quelconque et A_1 est une matrice de $\mathbb{M}(n-1,\mathbb{C})$. Nous pouvons donc répéter le processus sur A_1 et obtenir une matrice triangulaire supérieure (nous utilisons le fait qu'un produit de matrices orthogonales est une matrice orthogonale [55]). □

DefWQNooKEeJzv

Définition 12.101.
Un endomorphisme est ***normal*** *si il commute avec son conjugué hermitien* [56].

Les opérateurs normaux comprennent évidemment les opérateurs hermitiens, mais également les anti-hermitiens, et ça c'est bien parce que c'est le cas de l'algèbre associée à SU(2).

ThogammwA

Théorème 12.102 (Théorème spectral pour les matrices normales[302, 387, 388])**.**
Soit $A \in \mathbb{M}(n,\mathbb{C})$ une matrice de valeurs propres $\lambda_1,\dots,\lambda_n$ (non spécialement distinctes). Alors les conditions suivantes sont équivalentes :

ItemJZhFPSi

(1) A est normale [57],

ItemJZhFPSii

(2) A se diagonalise par une matrice unitaire,

ITEMooIIQTooQORrXP

(3) $\sum_{i,j=1}^n |A_{ij}|^2 = \sum_{j=1}^n |\lambda_j|^2$,

ITEMooJRKSooNfsQJb

(4) il existe une base orthonormale de vecteurs propres de A.

Démonstration. En plusieurs parties.

(1) **(1) ⇒ (2)** Soit Q la matrice unitaire donnée par la décomposition de Schur (lemme 12.100) : $A = QTQ^{-1}$. Étant donné que A est normale nous avons

$$QTT^*Q^{-1} = QT^*TQ^{-1}, \tag{12.219}$$

ce qui montre que T est également normale. Or une matrice triangulaire supérieure normale est diagonale. En effet nous avons $T_{ij} = 0$ lorsque $i > j$ et

$$(TT^*)_{ii} = (T^*T)_{ii} = \sum_{k=1}^n |T_{ki}|^2 = \sum_{k=1}^n |T_{ik}|^2. \tag{12.220}$$

Écrivons cela pour $i = 1$ en tenant compte de $|T_{k1}|^2 = 0$ pour $k = 2,\dots,n$,

$$|T_{11}|^2 = |T_{11}|^2 + |T_{12}|^2 + \cdots + |T_{1n}|^2, \tag{12.221}$$

ce qui implique que T_{11} est le seul non nul parmi les T_{1k}. En continuant de la sorte avec $i = 2,\dots,n$ nous trouvons que T est diagonale.

(2) **(2) ⇒ (1)** Si A se diagonalise par une matrice unitaire, $UAU^* = D$, nous avons

$$DD^* = UAA^*U^* \tag{12.222}$$

et

$$D^*D = UA^*AU^*, \tag{12.223}$$

qui ce prouve que A est normale.

55. Proposition 8.46(1).
56. Définition 8.177.
57. Définition 12.101.

(3) **(2) ⇒ (3)** Il existe une matrice unitaire U et une diagonale D telles que $D = U^\dagger AU$. La proposition 8.216 nous dit que la diagonale de D contient les valeurs propres de A, c'est-à-dire $D_{kl} = \delta_{kl}\lambda_k$.
Nous avons $D^\dagger = U^\dagger A^\dagger U$, et donc

$$D^\dagger D = U^\dagger A^\dagger UU^\dagger AU = U^\dagger A^\dagger AU, \tag{12.224}$$

c'est-à-dire que $D^\dagger D$ est semblable à $A^\dagger A$. La proposition 8.207 implique que leurs traces sont donc égales. Petit calcul en utilisant le lemme 8.178 ($A^\dagger_{ij} = A^*_{ji}$) :

$$\mathrm{Tr}(A^\dagger A) = \sum_k (A^\dagger A)_{kk} = \sum_{kl} A^*_{lk}A_{lk} = \sum_{kl} |A_{lk}|^2. \tag{12.225}$$ EQooAGPLooYXVqGy

Deuxième calcul :

$$\mathrm{Tr}(D^\dagger D) = \sum_{kl} D^\dagger_{kl}D_{lk} = \sum_{kl} \delta_{lk}\lambda^*_k\delta_{lk}\lambda_k = \sum_k |\lambda_k|^2. \tag{12.226}$$ EQooTUWNooIgwmx

Nous avons le résultat en égalisant (12.226) avec (12.225).

(4) **(3) ⇒ (2)** Le lemme de Schur complexe 12.100 nous indique qu'il existe une matrice unitaire U et une matrice triangulaire supérieure T telles que $T = UAU^\dagger$ et telle que $T_{ii} = \lambda_i$. Le lemme 8.209 nous indique que $\sum_{ij} |T_{ij}|^2 = \sum_{ij} |A_{ij}|^2$. Donc

$$\sum_{ij} |A_{ij}|^2 = \sum_{ij} |T_{ij}|^2 = \sum_j |\lambda_j|^2 + \sum_{j>i} |T_{ij}|^2. \tag{12.227}$$

Mais l'hypothèse nous dit que $\sum_{ij} |A_{ij}|^2 = \sum_j |\lambda_j|^2$. Nous en déduisons que $\sum_{j>i} |T_{ij}|^2 = 0$, et donc que T est diagonale.

(5) **(2)⇒ (4)** Soit une base $\{e_i\}$ de $\mathbb{C}^n$. Soit $A = UDU^\dagger$. Il suffit de poser $f_i = Ue_i$ pour avoir

$$Af_i = UDU^\dagger Ue_i = UDe_i = \lambda_i Ue_i = \lambda_i f_i. \tag{12.228}$$

Donc f_i est un vecteur propre de A. Le fait que $\{f_i\}$ soit une base de $\mathbb{C}^n$ est dû au fait que U est inversible.

(6) **(4) ⇒ (2)** Soit la base canonique $\{e_i\}$ de $\mathbb{C}^n$. Si $\{f_i\}$ est une base de vecteurs propres normalisés de A, alors il existe une matrice unitaire U telle que $f_i = Ue_i$. Alors la matrice $UAU^\dagger$ est diagonale.

□

Tant que nous en sommes à parler de spectre de matrices hermitiennes...Soit une matrice inversible $A \in \mathrm{GL}(n, \mathbb{C})$. La matrice A^*A est hermitienne [58] et le théorème 10.20 nous assure que ses valeurs propres sont réelles. Par la remarque 10.21, ses valeurs propres sont même positives. LEMooHUGEooVYhZdZ

Lemme 12.103 ([389]).
Si A est une matrice carrée et inversible,

$$\mathrm{Spec}(A^*A) = \mathrm{Spec}(AA^*) \tag{12.229}$$

Démonstration. Nous allons montrer l'égalité des polynômes caractéristiques. D'abord une simple multiplication montre que

$$(A^*A - \lambda\mathbb{1})A^{-1} = A^{-1}(AA^* - \lambda\mathbb{1}). \tag{12.230}$$

Nous prenons le déterminant de cette égalité en utilisant les propriétés 8.16(1) et (3) :

$$\det(A^*A - \lambda\mathbb{1})\det(A^{-1}) = \det(A^{-1})\det(AA^* - \lambda\mathbb{1}). \tag{12.231}$$

En simplifiant par $\det(A^{-1})$ (qui est non nul parce que A est inversible) nous obtenons l'égalité des polynômes caractéristiques et donc l'égalité des spectres. □

58. Définition 8.179.

En particulier les matrices hermitiennes, anti-hermitiennes et unitaires sont trigonalisables par une matrice unitaire, qui peut être choisie de déterminant 1.

LEMooRCFGooPPXiKi

Lemme 12.104.
Soient $A \in \mathbb{M}(n,\mathbb{C})$ et une matrice unitaire U telle que $A = UTU^{-1}$ où T est triangulaire.

(1) En ce qui concerne les polynômes caractéristiques, $\chi_A = \chi_T$.

(2) Pour les spectres, $\mathrm{Spec}(A) = \mathrm{Spec}(T)$.

(3) Les valeurs propres de A sont les éléments diagonaux de T.

Démonstration. Puisque U commute évidemment avec $\mathbb{1}$, nous avons

$$\chi_A(\lambda) = \det(A - \lambda\mathbb{1}) = \det(UTU^{-1} - \lambda\mathbb{1}) = \det\big(U(T - \lambda\mathbb{1})U^{-1}\big). \tag{12.232}$$

À ce niveau nous utilisons le fait que le déterminant soit multiplicatif 8.16 pour conclure :

$$\chi_A(\lambda) = \det\big(U(T - \lambda\mathbb{1})U^{-1}\big) = \det(U)\det(T - \lambda\mathbb{1})\det(U^{-1}) = \det(T - \lambda\mathbb{1}) = \chi_T(\lambda). \tag{12.233}$$

Pour les spectres, l'égalité des polynômes caractéristiques implique l'égalité des spectres parce que les valeurs propres sont les racines du polynôme caractéristique par le théorème 8.117.

Les valeurs propres d'une matrice triangulaire sont les valeurs sur la diagonale. □

LemSLGPooIghEPI

Lemme 12.105 (Trigonalisation simultanée)**.**
Une famille de matrices de $\mathrm{GL}(n,\mathbb{C})$ commutant deux à deux est simultanément trigonalisable.

Démonstration. Commençons par enfoncer une porte ouverte par la proposition 12.98 : toutes les matrices de $\mathrm{GL}(n,\mathbb{C})$ sont trigonalisables parce que tous les polynômes sont scindés.

Nous effectuons la démonstration par récurrence sur la dimension. Si $n = 1$ alors toutes les matrices sont triangulaires et nous ne nous posons pas de questions. Nous supposons donc $n > 1$.

Soit la famille $(A_i)_{i\in I}$ dans $\mathrm{GL}(n,\mathbb{C})$ et A_0 un de ses éléments. Nous nommons $\lambda_1,\dots,\lambda_r$ les valeurs propres distinctes de A_0. Le théorème de décomposition primaire 8.263 nous donne la somme directe d'espaces caractéristiques [59]

$$E = F_{\lambda_1}(A_0) \oplus \dots \oplus F_{\lambda_r}(A_0). \tag{12.234}$$

Nous pouvons supposer que cette somme n'est pas réduite à un seul terme. En effet si tel était le cas, A_0 serait un multiple de l'identité parce que A_0 n'aurait qu'une seule valeur propre et les sommes dans la décomposition de Dunford 8.265(3) se réduisent à un seul terme (et $p_i = \mathrm{Id}$). En particulier les dimensions des espaces $F_\lambda(A_0)$ sont strictement plus petites que n.

Puisque tous les A_i commutent avec A_0, les espaces $F_\lambda(A_0)$ sont stables par les A_i et nous pouvons trigonaliser les A_i simultanément sur chacun des $F_\lambda(A_0)$ en utilisant l'hypothèse de récurrence. □

ThoUWQBooCvutTO

Théorème 12.106 (Lie-Kolchin[380, 390])**.**
Tout sous-groupe connexe et résoluble [60] de $\mathrm{GL}(n,\mathbb{C})$ est conjugué à un groupe de matrices triangulaires.

Démonstration. Soit G un sous-groupe connexe et résoluble de $\mathrm{GL}(n,\mathbb{C})$. Nous faisons une récurrence sur n.

59. Définition 8.259.
60. Définition 2.22.

(1) **Si sous-espace non trivial stable par G** Nous commençons par voir ce qu'il se passe si il existe un sous-espace vectoriel non trivial V de $\mathbb{C}^n$ stabilisé par G. Pour cela nous considérons une base de $\mathbb{C}^n$ dont les premiers éléments forment une base de V (base incomplète, théorème 4.14). Les éléments de G s'écrivent, dans cette base,

$$\begin{pmatrix} g_1 & * \\ 0 & g_2 \end{pmatrix}. \tag{12.235}$$

Les matrices g_1 et g_2 sont carrées. Nous considérons alors l'application ψ définie par

$$\begin{aligned} \psi\colon G &\to \mathrm{GL}(V) \\ g &\mapsto g_1. \end{aligned} \tag{12.236}$$

Cela est un morphisme de groupes parce que

$$\begin{pmatrix} g_1 & * \\ 0 & g_2 \end{pmatrix} \begin{pmatrix} h_1 & * \\ 0 & h_2 \end{pmatrix} = \begin{pmatrix} g_1h_1 & * \\ 0 & g_2h_2 \end{pmatrix}, \tag{12.237}$$

de telle sorte que $\psi(gh) = \psi(g)\psi(h)$.

Le groupe $\psi(G)$ est connexe et résoluble. En effet $\psi(G)$ est connexe en tant qu'image d'un connexe par une application continue (lemme 7.218). Et il est résoluble en tant qu'image d'un groupe résoluble par un morphisme par la proposition 2.25.

Nous avons prouvé que $\psi(G)$ est un sous-groupe résoluble et connexe de $\mathrm{GL}(V)$. Vu que V est de dimension strictement plus petite que $\mathbb{C}^n$, nous utilisons l'hypothèse de récurrence et nous déduisons que $\psi(G)$ est conjugué à un groupe de matrices triangulaire, c'est-à-dire qu'il existe une base de V dans laquelle toutes les matrices g_1 (avec $g \in G$) sont triangulaires supérieures.

On fait de même avec l'application $g \mapsto g_2$, ce qui donne une base du supplémentaire de V dans laquelle les matrices g_2 sont triangulaires supérieures.

En couplant ces deux bases, nous obtenons une base de $\mathbb{C}^n$ dans laquelle toutes les matrices (12.235) (c'est-à-dire toutes les matrices de G) sont triangulaires supérieures.

(2) **Sinon** Nous supposons à présent que $\mathbb{C}^n$ n'a pas de sous-espaces non triviaux stables sous G. Nous posons $m = \min\{k \text{ tel que } D^k(G) = \{e\}\}$, qui existe parce que G et résoluble et que sa suite dérivée termine sur e (proposition 2.24).

(3) **Si $m = 1$** Si $m = 1$ alors G est abélien et il existe une base de G dans laquelle toutes les matrices de G sont triangulaires (lemme 12.105). Le premier vecteur d'une telle base serait stable par G, mais comme nous avons supposé qu'il n'y avait pas de sous-espaces non triviaux stabilisés par G, ce vecteur tout seul forme un sous-espace trivial, c'est-à-dire que $n = 1$. Dans ce cas, le théorème est démontré.

(4) **Si $m > 1$** Nous devons maintenant traiter le cas où $m > 1$. Nous posons $H = D^{m-1}(G)$; cela est un sous-groupe normal et abélien de G. Encore une fois le résultat de trigonalisation simultanée 12.105 donne une base dans laquelle tous les éléments de H sont triangulaires. En particulier le premier élément de cette base est un vecteur propre commun à toutes les matrices de H.

Soit V le sous-espace engendré par tous les vecteurs propres communs de H. Nous venons de voir que V n'est pas vide. Nous allons montrer que V est stable par G. Soient $h \in H$, $v \in V$ et $g \in G$:

$$h\big(g(v)\big) = g\underbrace{g^{-1}hg}_{\in H}(v) = g(\lambda v) = \lambda g(v) \tag{12.238}$$

parce que v est vecteur propre de $g^{-1}hg$. Ce que le calcul (12.238) montre est que $g(v)$ est vecteur propre de h pour la valeur propre λ. Donc $g(v) \in V$ et V est stabilisé par G. Mais comme il n'existe pas d'espaces non triviaux stabilisés par G, nous en déduisons que $V = \mathbb{C}^n$. Donc tous les vecteurs de $\mathbb{C}^n$ sont vecteurs propres communs de H. Autrement dit on a une base de diagonalisation simultanée de H.

(5) H **est dans le centre de** G

Montrons à présent que H est dans le centre de G, c'est-à-dire que pour tout $g \in G$ et $h \in H$ il faut $ghg^{-1} = h$. D'abord ghg^{-1} est une matrice diagonale (parce qu'elle est dans H) ayant les mêmes valeurs propres que h. En effet si λ est valeur propre de ghg^{-1} pour le vecteur propre v, alors

$$(ghg^{-1})(v) = \lambda v \tag{12.239a}$$
$$h\big(g^{-1}v\big) = \lambda\big(g^{-1}v\big), \tag{12.239b}$$

c'est-à-dire que λ est également valeur propre de h, pour le vecteur propre $g^{-1}v$. Mais comme h a un nombre fini de valeurs propres, il n'y a qu'un nombre fini de matrices diagonales ayant les mêmes valeurs propres que h. L'ensemble $\mathbf{Ad}(G)h$ est donc un ensemble fini. D'autre part, l'application $g \mapsto g^{-1}hg$ est continue, et G est connexe, donc l'ensemble $\mathbf{Ad}(G)h$ est connexe. Un ensemble fini et connexe dans $\mathrm{GL}(n, \mathbb{C})$ est nécessairement réduit à un seul point. Cela prouve que $ghg^{-1} = h$ pour tout $g \in G$ et $h \in H$.

(6) **Espaces propres stables pour tout** G

Soit $h \in H$ et W un espace propre de h (ça existe non vide, parce que H est triangularisé, voir plus haut). Alors nous allons prouver que W est stable pour tous les éléments de G. En effet si $w \in W$ avec $h(w) = \lambda w$ alors en permutant g et h,

$$hg(w) = g(hw) = \lambda g(w), \tag{12.240}$$

donc $g(w)$ est aussi vecteur propre de h pour la valeurs propre λ, c'est-à-dire que $g(w) \in W$. Comme nous supposons que $\mathbb{C}^n$ n'a pas d'espaces invariants non triviaux, nous devons conclure que $W = \mathbb{C}^n$, c'est-à-dire que H est composé d'homothéties. C'est-à-dire que pour tout $h \in H$ nous avons $h = \lambda_h \mathbb{1}$.

(7) **Contradiction sur la minimalité de** m

Les éléments d'un groupe dérivé sont de déterminant 1 parce que $\det(g_1 g_2 g_1^{-1} g_2^{-1}) = 1$. Par conséquent pour tout h, le nombre λ_h est une racine n^{e}de l'unité. Vu qu'il n'y a qu'une quantité finie de racines n^{e} de l'unité, le groupe H est fini et connexe et donc une fois de plus, réduit à un élément, c'est-à-dire $H = \{e\}$. Cela contredit la minimalité de m et donc produit une contradiction. Nous devons donc avoir $m = 1$.

(8) **Conclusion** Nous avons vu que si $\mathbb{C}^n$ avait un sous-espace non trivial fixé par G alors le théorème était démontré. Par ailleurs si $\mathbb{C}^n$ n'a pas un tel sous-espace, soit $m = 1$ (et alors le théorème est également prouvé), soit $m > 1$ et alors on a une contradiction.

Bref, le théorème est prouvé sous peine de contradiction.

□

Remarque 12.107.
Le lemme mentionne le fait que les valeurs propres de A sont les éléments diagonaux de T. Mais attention : ceci ne dit rien au niveau des multiplicités géométriques. Un nombre peut être cinq fois sur la diagonale de T alors que l'espace propre correspondant pour A n'est que de dimension 1. Exemple : la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{12.241}$$

a deux 1 sur la diagonale. Le nombre 1 est bien une valeur propre de A, mais le système

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{12.242}$$

donne $y = 0$ et donc un espace propre de dimension seulement 1.

RemXFZTooXkGzQg

Remarque 12.108.
Si $\mathbb{K}$ est algébriquement clos (comme $\mathbb{C}$ par exemple), alors tous les polynômes sont scindés et toutes les matrices sont trigonalisables[61]. Un exemple un peu simple de cela est la matrice

$$u = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{12.243}$$

Le polynôme caractéristique est $\chi_u(X) = X^2 + 1$ et les valeurs propres sont $\pm i$. Il est vite vu que dans la base

$$\{\begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}\} \tag{12.244}$$

de $\mathbb{C}^2$, la matrice u se note $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

RemREOSooGEDJWX

Remarque 12.109.
Cela nous donne une autre façon de prouver qu'une matrice nilpotente de $\mathbb{M}(n, \mathbb{C})$ ou $\mathbb{M}(n, \mathbb{R})$ est trigonalisable[391]. D'abord dans $\mathbb{M}(n, \mathbb{C})$, toutes les matrices sont trigonalisables[62], et les valeurs propres arrivent sur la diagonale. Mais comme les valeurs propres d'une matrice nilpotente valent zéro, elle est triangulaire stricte. Par ailleurs, son polynôme caractéristique est alors X^n.

Ensuite si $u \in \mathbb{M}(n, \mathbb{R})$ nous pouvons voir u comme une matrice dans $\mathbb{M}(n, \mathbb{C})$ et y calculer son polynôme caractéristique qui sera tout de même X^n. Ce polynôme étant scindé, la proposition 12.98 nous assure que u est trigonalisable. Une fois de plus, les valeurs propres étant sur la diagonale, elle est triangulaire supérieure stricte.

CorUNZooAZULXT

Corolaire 12.110.
Le polynôme caractéristique[63] *sur $\mathbb{C}$ d'une matrice s'écrit sous la forme*

$$\chi_A(X) = \prod_{i=1}^{r} (X - \lambda_i)^{m_i} \tag{12.245}$$

où les λ_i sont les valeurs propres distinctes de A et m_i sont les multiplicités correspondantes.

Démonstration. Le lemme 12.100 nous donne l'existence d'une base de trigonalisation ; dans cette base les valeurs propres de A sont sur la diagonale et nous avons

$$\chi_A(X) = \det(A - X\mathbb{1}) = \det \begin{pmatrix} X - \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & X - \lambda_r \end{pmatrix}, \tag{12.246}$$

qui vaut bien le produit annoncé. □

CORooTPDHooXazTuZ

Corolaire 12.111.
Si $A \in \mathbb{M}(n, \mathbb{C})$ et $k \in \mathbb{N}$ alors

$$\mathrm{Spec}(A^k) = \{\lambda^k \text{ tel que } \lambda \in \mathrm{Spec}(A)\}. \tag{12.247}$$

Démonstration. Par le lemme 12.100 nous avons une matrice unitaire U et une triangulaire T telles que $A = UTU^{-1}$. En passant à la puissance k nous avons aussi

$$A^k = UT^kU^{-1}. \tag{12.248}$$

Donc le spectre de A^k est celui de T^k (lemme 12.104 et le fait qu'une puissance d'une matrice triangulaire est encore triangulaire). Or les éléments diagonaux de T^k sont les puissances k^{e} des éléments diagonaux de T, qui sont les valeurs propres de A. □

61. La proposition 12.98 montre cela, et le lemme de Schur complexe 12.100 va un peu plus loin, et précise que la trigonalisation peut être obtenue par une matrice unitaire.
62. Parce que le polynôme caractéristique est scindé, voir la proposition 12.98.
63. Définition 8.113.

Pour le cas complexe, c'est le lemme 10.20 et le théorème 12.102.

12.11 Matrices, spectre et norme

Le lien entre la norme opérateur d'une matrice et son spectre sera entre autres utilisé pour étudier le conditionnement de problèmes numériques. Voir la définition 34.107 et par exemple son lien avec la résolution numérique de systèmes linéaires dans la proposition 34.112.

PROPooKLFKooSVnDzr

Proposition 12.112 ([329]).
Soit une matrice $A \in \mathbb{M}(n,\mathbb{C})$ de rayon spectral $\rho(A)$. Soit une norme $\|.\|$ sur $\mathbb{C}^n$ et la norme opérateur correspondante. Alors

$$\rho(A) \leqslant \|A^k\|^{1/k} \tag{12.249}$$

pour tout $k \in \mathbb{N}$.

Démonstration. Soit $v \in \mathbb{C}^n$ et $\lambda \in \mathbb{C}$ un couple vecteur-valeur propre. Nous avons $\|Av\| = |\lambda|\|v\|$ et aussi

$$|\lambda|^k\|v\| = \|\lambda^k v\| = \|A^k v\| \leqslant \|A^k\|\|v\|. \tag{12.250}$$

La dernière inégalité est due au fait que nous avons choisi sur $\mathbb{M}(n,\mathbb{C})$ la norme subordonnée à celle choisie sur $\mathbb{C}^n$, via le lemme 10.66. Nous simplifions par $\|v\|$ et obtenons $|\lambda| \leqslant \|A^k\|^{1/k}$. Étant donné que $\rho(A)$ est la maximum de tous les λ possibles, la majoration passe au maximum :

$$\rho(A) \leqslant \|A^k\|^{1/k}. \tag{12.251}$$

□

PROPooJGNFooEwtNmJ

Proposition 12.113.
Soient deux espaces vectoriels normés E et V. Soient des applications continues $f,g\colon E \to \mathrm{End}(V)$. Alors l'application

$$\begin{aligned}\psi\colon E &\to \mathrm{End}(V)\\ x &\mapsto f(x)\circ g(x)\end{aligned} \tag{12.252}$$

est continue.

Démonstration. Soit une suite $x_k \xrightarrow{E} x$. Nous devons montrer que $\psi(x_k) \xrightarrow{\mathrm{End}(V)} \psi(x)$. Pour cela nous utilisons le lemme 10.69 qui indique que la norme opérateur est une norme d'algèbre. Nous avons :

$$\begin{aligned}\|\psi(x_k)-\psi(x)\| &= \|f(x_k)\circ g(x_k) - f(x)\circ g(x)\| && \text{(12.253a)}\\ &\leqslant \|f(x_k)\circ g(x_k) - f(x_k)\circ g(x)\| + \|f(x_k)\circ g(x) - f(x)\circ g(x)\| && \text{(12.253b)}\\ &= \|f(x_k)\circ\big(g(x_k)-g(x)\big)\| + \|\big(f(x_k)-f(x)\big)\circ g(x)\| && \text{(12.253c)}\\ &\leqslant \|f(x_k)\|\|g(x_k)-g(x)\| + \|f(x_k)-f(x)\|\|g(x)\|. && \text{(12.253d)}\end{aligned}$$

Pour $k \to \infty$ nous avons $\|f(x_k)\| \to \|f(x)\|$, $\|f(x_k)-f(x)\| \to 0$ (parce que f est continue) et similaire avec g. Donc le tout tend vers zéro. □

12.11.1 Rayon spectral

La chose impressionnante dans la proposition suivante est que $\rho(A)$ est défini indépendamment du choix de la norme sur $\mathbb{M}(n,\mathbb{K})$ ou sur $\mathbb{K}$. Lorsque nous écrivons $\|A\|$, nous disons implicitement qu'une norme a été choisie sur $\mathbb{K}$, et que nous avons pris la norme subordonnée sur $\mathbb{M}(n,\mathbb{K})$.

PROPooWZJBooTPLSZp

Proposition 12.114 ([392]).
Soit A une matrice de $\mathbb{M}(n,\mathbb{R})$ ou $\mathbb{M}(n,\mathbb{C})$. Alors

$$\rho(A) \leqslant \|A\|. \tag{12.254}$$

Démonstration. Nous devons séparer les cas, suivant que le corps de base soit $\mathbb{R}$ ou $\mathbb{C}$.

(1) **Pour $A \in \mathbb{M}(n,\mathbb{C})$** Soit λ une valeur propre de A telle que $|\lambda|$ soit la plus grande. Nous avons donc $\rho(A) = |\lambda|$. Soit un vecteur propre $u \in \mathbb{C}^n$ pour la valeur propre λ. En prenant la norme sur l'égalité $Au = \lambda u$, et en utilisant le lemme 10.66,

$$|\lambda|\|u\| = \|Au\| \leqslant \|A\|\|u\|. \tag{12.255}$$

Donc $|\lambda| \leqslant \|A\|$ et $\rho(A) \leqslant \|A\|$.

(2) **Pour $A \in \mathbb{M}(n,\mathbb{R})$**
L'endroit qui coince dans le raisonnement effectué pour $\mathbb{M}(n,\mathbb{C})$ est que, certes $A \in \mathbb{M}(n,\mathbb{R})$ possède une plus grande valeur propre en module et qu'un vecteur propre lui est associé. Mais ce vecteur propre est, à priori, dans $\mathbb{C}^n$, et non dans $\mathbb{R}^n$. Nous pouvons donc écrire $Au = \lambda u$, mais pas $\|Au\| = |\lambda|\|u\|$ parce que nous ne savons pas quelle norme prendre sur $\mathbb{C}^n$.

Il n'est pas certain que nous ayons une norme sur $\mathbb{C}^n$ qui se réduit sur $\mathbb{R}^n$ à celle choisie implicitement dans l'énoncé. Nous allons donc ruser un peu.

Soit une norme N sur $\mathbb{C}^n$ [64]. Nous nommons également N la norme subordonnée sur $\mathbb{M}(n,\mathbb{C})$ et la norme restreinte sur $\mathbb{M}(n,\mathbb{R})$. Vu que N est une norme sur $\mathbb{M}(n,\mathbb{R})$ et que ce dernier est de dimension finie, le théorème 10.49 nous indique que N est équivalente à $\|.\|$. Il existe donc $C > 0$ tel que

$$N(B) \leqslant C\|B\| \tag{12.256}$$

EQooBNWMooNgnMxC

pour tout $B \in \mathbb{M}(n,\mathbb{R})$. Nous avons maintenant

$$\rho(A)^m \leqslant N(A^m) \leqslant C\|A^m\| \leqslant C\|A\|^m. \tag{12.257}$$

Justifications

— Par la proposition 12.112.
— Parce que $A^m \in \mathbb{M}(n,\mathbb{R})$ et la relation (12.256).
— Par itération du lemme 10.69.

Nous avons donc $\rho(A) \leqslant C^{1/m}\|A\|$ pour tout $m \in \mathbb{N}$. En prenant $m \to \infty$ et en tenant compte de $C^{1/m} \to 1$ nous trouvons $\rho(A) \leqslant \|A\|$.

□

LEMooGBLJooCPvxNl

Lemme 12.115 ([392])**.**
Soit $A \in \mathbb{M}(n,\mathbb{K})$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $\epsilon > 0$. Il existe une norme algébrique sur $\mathbb{M}(n,\mathbb{K})$ telle que

$$N(A) \leqslant \rho(A) + \epsilon. \tag{12.258}$$

Démonstration. Soit par le lemme 12.100 une matrice inversible U telle que $T = UAU^{-1}$ soit triangulaire supérieure, avec les valeurs propres sur la diagonale. Notons que même si $A \in \mathbb{M}(n,\mathbb{R})$, les matrices U et T sont, à priori, complexes.

Soit $s \in \mathbb{R}$ ainsi que les matrices

$$D_s = \text{diag}(1, s^{-1}, s^{-2}, \dots, s^{1-n}) \tag{12.259}$$

et $T_s = D_s T D_s^{-1}$. Nous fixerons un choix de s plus tard.

La norme que nous considérons est :

$$N(B) = \|(D_sU)B(D_sU)^{-1}\|_\infty \tag{12.260}$$

où $\|.\|_\infty$ est la norme sur $\mathbb{M}(n,\mathbb{K})$ subordonnée à la norme $\|.\|_\infty$ sur $\mathbb{K}^n$ dont nous avons déjà parlé dans l'exemple 10.61. Cela est bien une norme parce que

64. Il y en a plein, par exemple celle du produit scalaire $\langle x, y\rangle = \sum_k x_k\bar{y}_k$.

— Nous avons $\|B\|_\infty = 0$ si et seulement si $B = 0$, et comme (D_sU) est inversible, nous avons $(D_sU)B(D_sU)^{-1} = 0$ si et seulement si $B = 0$.
— $N(\lambda B) = |\lambda|N(B)$.
— Pour l'inégalité triangulaire :

$$N(B+C) = \|(D_sU)B(D_sU)^{-1} + (D_sU)C(D_sU)^{-1}\|_\infty \tag{12.261a}$$
$$\leqslant \|(D_sU)B(D_sU)^{-1}\|_\infty + \|(D_sU)C(D_sU)^{-1}\|_\infty \tag{12.261b}$$
$$= N(B) + N(C). \tag{12.261c}$$

En ce qui concerne la matrice A elle-même, nous avons

$$N(A) = \|(D_sU)A(D_sU)^{-1}\|_\infty = \|T_s\|_\infty. \tag{12.262}$$

C'est le moment de se demander comment se présente la matrice T_s. En tenant compte du fait que $(D_s)_{ik} = \delta_{ik}s^{1-i}$ nous avons

$$(T_s)_{ij} = \sum_{kl}(D_s)_{ik}T_{kl}(D_s^{-1})_{lj} = T_{ij}s^{j-i}. \tag{12.263}$$

La matrice T est encore triangulaire supérieure avec les valeurs propres de A sur la diagonale. Les éléments au-dessus de la diagonale sont tous multipliés par au moins s. Il est donc possible de choisir s suffisamment petit pour avoir [65]

$$\sum_{j=i+1}^{n} |(T_s)_{ij}| < \epsilon \tag{12.264}$$

EQooSIEIooTWAXQD

Avec ce choix, la formule 10.169 donne

$$N(T_s) \leqslant \max_i \sum_k |(T_s)_{ik}| \leqslant \epsilon + \rho(A). \tag{12.265}$$

En effet le ϵ vient de la somme sur toute la ligne sauf la diagonale (c'est-à-dire la partie $k \neq i$) et du choix (12.264) pour s. Le $\rho(A)$ provient du dernier terme de la somme (le terme sur la diagonale) qui est une valeur propre de A, donc majorable par $\rho(A)$.

Nous devons encore prouver que N est une norme algébrique. Pour cela nous allons montrer qu'elle est subordonnée à la norme

$$\begin{aligned} n\colon \mathbb{K}^n &\to \mathbb{R}^+ \\ v &\mapsto \|(UD_s)v\|_\infty. \end{aligned} \tag{12.266}$$

Cela sera suffisant pour avoir une norme algébrique par le lemme 10.69. La norme n sur $\mathbb{K}^n$ produit la norme suivante sur $\mathbb{M}(n,\mathbb{K})$:

$$n(B) = \sup_{v\neq 0}\frac{n(B)}{n(v)} = \sup_{v\neq 0}\frac{\|(UD_s)Bv\|_\infty}{\|UD_sv\|_\infty}. \tag{12.267}$$

Puisque UD_s est inversible, nous pouvons effectuer le changement de variables $v \mapsto (UD_s)^{-1}v$ pour écrire

$$n(B) = \sup_{v\neq 0}\frac{\|(UD_s)B(UD_s)^{-1}v\|_\infty}{\|(UD_s)(UD_s)^{-1}v\|_\infty} = \sup_{v\neq 0}\frac{\|(UD_s)B(UD_s)^{-1}v\|_\infty}{\|v\|_\infty} = \|(UD_s)B(UD_s)^{-1}\|_\infty = N(B). \tag{12.268}$$

□

PROPooYPLGooWKLbPA

Proposition 12.116.
Si $A \in \mathbb{M}(n,\mathbb{R})$ alors $\rho(A)^m = \rho(A^m)$ pour tout $m \in \mathbb{N}$.

65. Il me semble qu'il manque un module dans [392].

Démonstration. La matrice A peut être vue dans $\mathbb{M}(n,\mathbb{C})$ et nous pouvons lui appliquer le corolaire 12.111 :

$$\text{Spec}(A^k) = \{\lambda^k \text{ tel que } \lambda \in \text{Spec}(A)\}. \tag{12.269}$$

EQooJJIYooDBacjn

À noter qu'il n'y a pas de magie : le spectre de la matrice réelle A est déjà défini en voyant A comme matrice complexe. Le spectre dont il est question dans (12.269) est bien celui dont on parle dans la définition du rayon spectral.

Nous avons ensuite :

$$\begin{aligned}\rho(A^k) &= \max\{|\lambda| \text{ tel que } \lambda \in \text{Spec}(A^k)\} && (12.270a)\\ &= \max\{|\lambda^k| \text{ tel que } \lambda \in \text{Spec}(A)\} && (12.270b)\\ &= \max\{|\lambda|^k \text{ tel que } \lambda \in \text{Spec}(A)\} && (12.270c)\\ &= \rho(A)^k. && (12.270d)\end{aligned}$$

□

PROPooXEQLooHvzVVm

Proposition 12.117.
Soient des espaces vectoriels normés V de dimension n et W de dimension m sur $\mathbb{K}$ (corps normé). Nous considérons une base $\{e_s\}_{s=1,\dots,n}$ de V et $\{f_\alpha\}_{\alpha=1,\dots,m}$ de W.

Alors l'application

$$\begin{aligned}\psi\colon \mathbb{M}(n\times m,\mathbb{K}) &\to \mathcal{L}(V,W)\\ \psi(A)v &= \sum_{s\alpha} A_{s\alpha} v_s f_\alpha\end{aligned} \tag{12.271}$$

est un isomorphisme d'espaces topologiques.

Pour rappel, la topologie sur $\mathbb{M}(n,\mathbb{K})$ est donnée par la définition 9.33.

Démonstration. Nous savons déjà que ψ est une bijection. De plus, elle est linéaire et donc continue par la proposition 10.237. En ce qui concerne son inverse, c'est également une application linéaire (lemme 4.35) ; elle est alors également continue. □

PROPooDRHMooYzXbkl

Proposition 12.118.
Soit un espace vectoriel normé V de dimension finie. Soit une suite d'opérateurs $T_n \in \text{End}(V)$. Si $\{e_i\}$ est une base de V et si $T_n(e_i) \xrightarrow{V} e_i$ pour tout i, alors $T_n \xrightarrow{\text{End}(V)} \text{Id}$.

Démonstration. Nous utilisons l'application $\psi\colon \mathbb{M}(n,\mathbb{K}) \to \text{End}(V)$ définie en 4.68. Elle nous permet d'écrire

$$T_n(x) = \sum_{kl} \psi^{-1}(T_i)_{kl} x_l e_k, \tag{12.272}$$

que nous allons particulariser à $x = e_j$. Nous avons

$$\begin{aligned}e_j &= \lim_{n\to\infty} T_n(e_j) && (12.273a)\\ &= \lim_{n\to\infty} \sum_{kl} \psi^{-1}(T_n)_{kl}\delta_{jl} e_k && (12.273b)\\ &= \sum_k \big(\lim_{n\to\infty} \psi^{-1}(T_n)_{kj}\big) e_k && (12.273c)\end{aligned}$$

En identifiant les coefficients de e_j, on trouve

$$\lim_{n\to\infty} \psi^{-1}(T_n)_{kj} = \delta_{kj}. \tag{12.274}$$

Pour chaque k et l, à gauche nous avons une limite dans $\mathbb{K}$. Vue la topologie sur $\mathbb{M}(n,\mathbb{K})$ [66], nous pouvons écrire cela comme une limite dans $\mathbb{M}(n,\mathbb{K})$:

$$\lim_{n\to\infty} \psi^{-1}(T_n) = \mathbb{1}. \tag{12.275}$$

66. Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes (théorème 10.49). Sur $\mathbb{M}(n,\mathbb{K})$, nous avons convenu dans la définition 9.33 de considérer la norme maximum.

Nous savons que ψ^{-1} est continue (proposition 12.117) de telle sorte que nous pouvons la commuter avec la limite :

$$\mathbb{1} = \lim_{n\to\infty} \psi^{-1}(T_n) = \psi^{-1}\big(\lim_{n\to\infty} T_n\big). \tag{12.276}$$

Appliquant maintenant ψ des deux côtés, $\psi(\mathbb{1}) = \mathrm{Id}$ et

$$\mathrm{Id} = \lim_{n\to\infty} T_n. \tag{12.277}$$

□

Le point important de la définition 10.272 est la continuité. En dimension infinie, la continuité n'est par exemple pas équivalente à l'inversibilité (penser à $e_k \mapsto ke_k$).

Si V est un espace vectoriel normé, nous avons déjà défini son dual topologique V' comme étant l'ensemble des applications linéaires continues $V \to \mathbb{C}$ ou $V \to \mathbb{R}$ selon le corps de base de V. C'est la définition 4.126.

Proposition 12.119.
Soient un espace vectoriel normé V et un élément $v \in V$ vérifiant $\|v\| = 1$. Il existe une forme $\varphi \in V'$ telle que $\|\varphi\| = 1$ et $\varphi(v) = 1$.

Démonstration. Nous allons utiliser le théorème de la base incomplète 4.25. Pour cela nous considérons $I = V$ et la partie clairement génératrice $G = \{e_i = i\}_{i\in I}$ (si vous avez bien suivi, $G = V$ en fait ; rien de bien profond). Nous considérons ensuite $I_0 = \{v\}$. Le théorème de la base incomplète nous donne l'existence de I_1 tel que $I_0 \subset I_I \subset I$ et tel que $B = \{e_i\}_{i\in I_1}$ est une base.

Tout cela pour dire que $B = \{e_i\}_{i\in I_1}$ est une base contenant v. Nous allons aussi éventuellement redéfinir la norme de e_i pour avoir $\|e_i\| = 1$. Cette renormalisation n'affecte pas le fait que $v \in B$.

Nous passons maintenant à la définition de $\varphi\colon V \to \mathbb{K}$. Pour $x \in V$ nous commençons par écrire

$$x = \sum_{j\in J} x_j e_j \tag{12.278}$$

et nous posons

$$\varphi(x) = \begin{cases} x_v & \text{si } v \in J \\ 0 & \text{sinon.} \end{cases} \tag{12.279}$$

Cette définition a un sens par la partie unicité de la proposition 4.7 de décomposition d'un élément dans une base.

Nous devons calculer la norme de φ. Par la proposition 10.58 nous avons

$$\|\varphi\| = \sup_{\|x\|=1} |\varphi(x)|. \tag{12.280}$$ EQooEFLLooQWPSev

Avec $x = v$ nous avons $\varphi(x) = 1$ et donc $\|\varphi\| \geqslant 1$.

Nous devons encore montrer que $\|\varphi\| \leqslant 1$. Un élément $x \in V$ s'écrit toujours sous la forme

$$x = \sum_{i\in J} x_j e_j \tag{12.281}$$

pour un certain J fini dans I_1 et pour certains $x_j \in \mathbb{K}$. Pour un tel x nous avons $\varphi(x) = x_v$. Si $|\varphi(x)| \geqslant 1$, alors $|x_v| \geqslant 1$, mais alors

$$\|x\| \leqslant \sum_{j\in J} |x_j|\|e_j\| = \sum_{j\in J} |x_j| \geqslant |x_v| > 1, \tag{12.282}$$

ce qui fait que ce x ne participe pas au supremum (12.280).

Notons que φ est continue (et donc bien dans V') parce qu'elle est bornée (proposition 10.70).

□

12.11.2 Normes de matrices et d'applications linéaires

THOooNDQSooOUWQrK

Théorème 12.120 (Norme matricielle et rayon spectral[393])**.**
La norme 2 d'une matrice est liée au rayon spectral de la façon suivante :

$$\|A\|_2 = \sqrt{\rho(A^tA)} \tag{12.283}$$

ou plus généralement par $\|A\|_2 = \sqrt{\rho(A^*A)}$.

LEMooNESTooVvUEOv

Lemme 12.121.
Soit une matrice $A \in \mathbb{M}(n,\mathbb{R})$ *qui est symétrique, strictement définie positive. Soient* λ_{min} *et* λ_{max} *les plus petites et plus grandes valeurs propres. Alors*

$$\|A\|_2 = \lambda_{max} \qquad et \qquad \|A^{-1}\|_2 = \frac{1}{\lambda_{min}}. \tag{12.284a}$$

Démonstration. Soient les vecteurs $v_1,\dots,v_n$ formant une base orthonormée de vecteurs propres [67] de A. Nous notons v_{max} celui de λ_{max}. Nous avons :

$$\|A\|_2 \geqslant \|Av_{max}\| = |\lambda_{max}|\|v_{max}\| = |\lambda_{max}| = \lambda_{max}. \tag{12.285}$$

Voilà l'inégalité dans un sens. Montrons l'inégalité dans l'autre sens. Soit $x = \sum_i x_i v_i$ avec $\|x\|_2 = 1$. Alors

$$\|Ax\| = \left|\sum_i x_i\lambda_i v_i\right| \leqslant \sqrt{\sum_i x_i^2\lambda_i^2} \leqslant \lambda_{max}\sqrt{\sum_i x_i^2} = \lambda_{max}. \tag{12.286}$$

En ce qui concerne l'affirmation pour la norme de A^{-1}, il suffit de remarquer que ses valeurs propres sont les inverses des valeurs propres de A. □

LEMooCSBVooZzqxqg

Lemme 12.122 ([1, 394])**.**
Soit une matrice diagonale $D \in \mathbb{M}(n,\mathbb{C})$ *dont nous notons* $\lambda_i \in \mathbb{C}$ *les éléments diagonaux. Alors la norme opérateur*[68] *de* D *est donnée par*

$$\|D\|_2 = \max_i\{|\lambda_i|\}. \tag{12.287}$$

Démonstration. En plusieurs points.

(1) **Le compact** Puisque la partie $\{x \in \mathbb{C}^n \text{ tel que } \|x\|_2 = 1\}$ est compacte, nous pouvons utiliser un maximum au lieu d'un supremum dans la définition de la norme opérateur (théorème de Weierstrass 7.146.).

(2) **Notations pour $\mathbb{C}^n$** Pour se mettre d'accord sur les notations, si $x \in \mathbb{C}^n$, alors $x = \sum_i x_i e_i$ où $e_1 \in \mathbb{C}^n$ est le vecteur $(1,0,\dots,0)$. C'est un vecteur de base de $\mathbb{C}^n$ comme espace vectoriel sur $\mathbb{C}$. Et d'ailleurs $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de $\mathbb{C}^n$.

(3) **Norme dans $\mathbb{C}^n$** Lorsque A est un opérateur sur $\mathbb{C}^n$, nous avons

$$\|Ax\|_2 = \left(\sum_i |(Ax)_i|^2\right)^{1/2} = \left(\sum_i \left|\sum_j A_{ij}x_j\right|^2\right)^{1/2}. \tag{12.288}$$

Nous avons utilisé les conventions (4.89).

67. Possible par le théorème spectral 8.227.
68. Norme opérateur, définition 10.55. La notation $\|D\|_2$ signifie la norme opérateur de $D\colon \mathbb{C}^n \to \mathbb{C}^n$ où l'on a mis la norme euclidienne sur $\mathbb{C}^n$, c'est-à-dire la norme de la définition 9.114.

(4) **Le calcul** Si c'est bon pour vous, je me lance dans le calcul :

$$\|D\|_2 = \max_{\|x\|_2=1} \|Dx\|_2 \tag{12.289a}$$

$$= \max_{\|x\|_2=1} \left(\sum_{i=1}^{n} \lambda_i x_i \right) \tag{12.289b}$$

$$\leqslant \max_{\|x\|_2=1} \max_i |\lambda_i| \underbrace{\left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}}_{=\|x\|_2} = \max_i |\lambda_i|. \tag{12.289c}$$

L'inégalité $\|D\|_2 \leqslant \max_i |\lambda_i|$ est prouvée. Nous démontrons à présent l'inégalité dans l'autre sens. Appliquons D au vecteur de base e_i : $De_i = \lambda_i e_i$. Donc

$$\|D\|_2 \geqslant \|De_i\|_2 = |\lambda_i e_i| = |\lambda_i|. \tag{12.290}$$

Cela étant valable pour tout i, nous avons $\|D\|_2 \geqslant \max_i |\lambda_i|$. □

PropMAQoKAg

Proposition 12.123.
La fonction

$$\begin{aligned} f\colon \mathbb{M}(n,\mathbb{R}) \times \mathbb{M}(n,\mathbb{R}) &\to \mathbb{R} \\ (X,Y) &\mapsto \operatorname{Tr}(X^tY) \end{aligned} \tag{12.291}$$

est un produit scalaire sur $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Il faut vérifier la définition 8.168.

— La bilinéarité est la linéarité de la trace.

— La symétrie de f est le fait que $\operatorname{Tr}(A^t) = \operatorname{Tr}(A)$.

— L'application f est définie positive parce que si $X \in \mathbb{M}(n,\mathbb{R})$, alors X^tX est symétrique définie positive, donc diagonalisable avec des nombres positifs sur la diagonale. La trace étant un invariant de similitude, nous avons $f(X,X) = \operatorname{Tr}(X^tX) \geqslant 0$. De plus si $\operatorname{Tr}(X^tX) = 0$, alors $X^tX = 0$ (pour la même raison de diagonalisation). Mais alors $\|Xu\| = 0$ pour tout $u \in E$, ce qui signifie que $X = 0$.

□

EXooHDCWooHragNA

Exemple 12.124.
Soient $m = n$, un point λ dans $\mathbb{R}$ et T_λ l'application linéaire définie par $T_\lambda(x) = \lambda x$. La norme de T_λ est alors

$$\|T_\lambda\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m} \leqslant 1} \|\lambda x\|_{\mathbb{R}^n} = |\lambda|.$$

Notez que T_λ n'est rien d'autre que l'homothétie de rapport λ dans $\mathbb{R}^m$. △

EXooVXENooZbtBNi

Exemple 12.125.
Toutes les isométries de $\mathbb{R}^n$ ont norme 1. En effet si T est une isométrie, $\|Tx\| = \|x\|$. En ce qui concerne la norme de T nous avons alors

$$\|T\| = \sup_{x\in\mathbb{R}^n} \frac{\|T(x)\|}{\|x\|} = \sup_{x\in\mathbb{R}^n} \frac{\|x\|}{\|x\|} = 1. \tag{12.292}$$

△

LEMooNWQWooBrpXgn

Lemme 12.126.
Soit $b \in \mathbb{R}^m$. *Nous considérons l'application*

$$\begin{aligned} T_b\colon \mathbb{R}^m &\to \mathbb{R} \\ x &\mapsto b \cdot x. \end{aligned} \tag{12.293}$$

(1) L'application T_b est linéaire.

(2) Sa norme est $\|T_b\|_{\mathcal{L}} = \|b\|_{\mathbb{R}^m}$.

Démonstration. La norme de T_b satisfait les inégalités suivantes

$$\|T_b\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} |b \cdot x| \leqslant \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|b\|_{\mathbb{R}^m}\|x\|_{\mathbb{R}^m} \leqslant \|b\|_{\mathbb{R}^m},$$

$$\|T_b\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|b \cdot x\|_{\mathbb{R}^n} \geqslant \left|b \cdot \frac{b}{\|b|}\right| = \frac{\|b\|^2}{\|b\|} = \|b\|_{\mathbb{R}^n},$$

donc $\|T_b\|_{\mathcal{L}} = \|b\|_{\mathbb{R}^n}$. □

Proposition 12.127.
Une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est continue.

Démonstration. Soit x un point dans $\mathbb{R}^m$. Nous devons vérifier l'égalité

$$\lim_{h\to 0_m} T(x+h) = T(x). \tag{12.294}$$

Cela revient à prouver que $\lim_{h\to 0_m} T(h) = 0$, parce que $T(x+h) = T(x) + T(h)$. Nous pouvons toujours majorer $\|T(h)\|_n$ par $\|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|h\|_{\mathbb{R}^m}$ (lemme 10.66). Quand h s'approche de 0_m sa norme $\|h\|_m$ tend vers 0, ce que nous permet de conclure parce que nous savons que de toutes façons, $\|T\|_{\mathcal{L}}$ est fini. □

Note : dans un espace de dimension infinie, la linéarité ne suffit pas pour avoir la continuité : il faut de plus être borné (ce que sont toutes les applications linéaires $\mathbb{R}^m \to \mathbb{R}^n$). Voir la proposition 10.70.

PROPooHXJAooGaDtme

Proposition 12.128 ([1])**.**
Soit $A \in \mathrm{GL}(n,\mathbb{C})$. La suite[69] *$(A^k)_{k\in\mathbb{Z}}$ est bornée*[70] *si et seulement si A est diagonalisable*[71] *et $\mathrm{Spec}(A) \subset \mathrm{S}^1$.*

Démonstration. Nous commençons par supposer que A est diagonalisable et que $\mathrm{Spec}(A) \subset \mathrm{S}^1$. Nous nommons λ_i ses valeurs propres. Par hypothèse $\lambda_i \in \mathrm{S}^1$. Nous avons donc $|\lambda_i| = 1$ et, par la proposition 9.102(4), nous avons $|\lambda_i|^k = 1$.

La proposition 8.216 permet de considérer une matrice inversible Q telle que $A = Q^{-1}DQ$ où D est la matrice diagonale $D_{ii} = \lambda_i$. Nous avons donc aussi

$$A^k = Q^{-1}D^kQ. \tag{12.295}$$

La matrice D^k est diagonale et $D^k_{ii} = \lambda_i^k$.

(1) **Pour $k \geqslant 0$** La norme matricielle étant une norme d'algèbre[72],

$$\|A^k\|_2 = \|Q^{-1}D^kQ\|_2 \leqslant \|Q^{-1}\|_2\|D^k\|_2\|Q\|_2. \tag{12.296}$$

En ce qui concerne la norme $\|D\|_2$, nous avons le lemme 12.122 qui nous annonce que $\|D\|_2 = \max_i\{|\lambda_i|\} = 1$. Dans notre cas, nous avons donc $\|D\|_2 = 1$ et

$$\|A^k\|_2 \leqslant \|Q^{-1}\|_2\|Q\|_2. \tag{12.297}$$

Autrement dit, la suite $(A^k)_{k\in\mathbb{N}}$ est majorée en norme par le nombre $\|Q^{-1}\|_2\|Q\|_2$.

69. Oui, c'est avec $k \in \mathbb{Z}$. Vu que A est dans GL, elle est inversible, donc pas de soucis à considérer A^{-1}.
70. Nous considérons sur $\mathrm{GL}(n,\mathbb{C})$ la norme opérateur dérivant de la norme hermitienne sur $\mathbb{C}^n$ donnée par la formule (9.163).
71. Définition 8.215.
72. Lemme 10.69.

(2) **Pour $k \leqslant 0$** Alors il suffit de poser $B = A^{-1}$. La matrice B est autant diagonalisable que A et le même raisonnement s'applique : il existe une matrice inversible P telle que

$$\|B^k\|_2 \leqslant \|P^{-1}\|_2\|P\|_2. \tag{12.298}$$

(3) **Pour tous les k** La suite $(A^k)_{k\in\mathbb{Z}}$ est donc majorée par le maximum entre $\|P^{-1}\|_2\|P\|$ et $\|Q^{-1}\|_2\|Q\|$.

Dans l'autre sens, maintenant.

Puisque nous travaillons sur $\mathbb{C}$, le polynôme caractéristique de A est scindé et la réduction de Jordan [73] s'applique. Nous considérons une matrice inversible Q telle que $A = Q^{-1}MQ$ avec

$$M = \begin{pmatrix} \lambda_1\mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_s\mathbb{1} + N_s \end{pmatrix}, \tag{12.299}$$

où les N_i sont nilpotents. Notez qu'ici les « $\mathbb{1}$ » sont de différentes tailles.

(1) **Juste un bloc** Nous considérons un bloc de Jordan $\lambda\mathbb{1} + N$. Nous supposons que la suite $(\lambda\mathbb{1} + N)^k$ est bornée, et nous allons montrer que $|\lambda| = 1$ et $N = 0$.

Nous nommons s la borne, et r le plus petit entier tel que $N^r = 0$. En utilisant la formule du binôme [74], pour $k > r$ nous avons

$$(\lambda\mathbb{1} + N)^k = \sum_{l=0}^{r-1} \binom{k}{l} \lambda^{k-l}N^l < s. \tag{12.300}$$

Le lemme 8.213 indique que la partie $\{N^k\}$ est libre. En utilisant le lemme 7.168, nous en déduisons que

$$\left|\binom{k}{l}\lambda^{k-l}N^l\right| < s \tag{12.301}$$

EQooZWGCooGEsobY

pour tout $k > r$ et pour tout $l \leqslant k$.

En particulier, pour tout k nous pouvons considérer le terme $l = 0$. Cela donne

$$|\lambda|^k < s \tag{12.302}$$

qui implique $|\lambda| \leqslant 1$.

Nous avons donc deux possibilités : $|\lambda| < 1$ et $|\lambda| = 1$. Supposons $|\lambda| = 1$, et considérons l'inéquation (12.301) avec $l = 1$: $\|k\lambda^{k-1}N\| < s$. Cela implique que

$$k\|N\| < s \tag{12.303}$$

pour tout k. Cela n'est possible que si $\|N\| = 0$ parce que $\mathbb{R}$ est archimédien (théorème 1.440). Nous restons donc avec les deux possibilités

— $|\lambda| < 1$

— $|\lambda| = 1$ et $N = 0$.

Nous nous tournons maintenant sur la contrainte que $(\lambda\mathbb{1} + N)^k$ doive rester borné pour $k < 0$. Nous avons

$$\lambda\mathbb{1} + N = \lambda(\mathbb{1} + \lambda^{-1}N), \tag{12.304}$$

73. Théorème 8.295.
74. Proposition 3.43.

et nous pouvons appliquer la proposition 8.214 à l'opérateur nilpotent $-\lambda^{-1}N$ pour avoir

$$\big(\mathbb{1}-(-\lambda^{-1}N)\big)^{-1}=\sum_{l=0}^{\infty}(-\lambda)^{-l}N^l \tag{12.305a}$$

$$=\mathbb{1}+\sum_{l=1}^{\infty}(-\lambda)^{-l}N^l \tag{12.305b}$$

$$=\mathbb{1}+\sum_{l=1}^{r-1}(-\lambda)^{-l}N^l. \tag{12.305c}$$

Ceci pour dire que $(\lambda\mathbb{1}+N)^{-1}=\lambda^{-1}(\mathbb{1}+\lambda^{-1}N')$ pour une autre matrice nilpotente[75] $N'=\sum_{l=1}^{r-1}(-\lambda)^{-l}N^l$. Le travail déjà fait, appliqué à λ^{-1} et N', nous donne deux possibilités :

— $|\lambda^{-1}|<1$

— $|\lambda^{-1}|=1$ et $N'=0$.

La possibilité $|\lambda^{-1}|<1$ est exclue parce qu'elle impliquerait $|\lambda|>1$ qui avait déjà été exclu. Il ne reste donc que la possibilité $|\lambda|=1$ et $N=N'=0$.

(2) **Pour la matrice M** Nous supposons que $\{M^k\}$ est borné : $\|M^k\|\leqslant s$ pour tout s. En utilisant le lemme 10.59 et la proposition 10.60, pour tout n et pour tout i nous avons :

$$\|(\mathbb{1}+\lambda_iN_i)^k\|<s. \tag{12.306}$$

Nous appliquons ce que nous venons de montrer pour les blocs et nous obtenons $|\lambda_i|=1$ et $N_i=0$.

(3) **La matrice A** Nous pouvons enfin parler de la matrice $A=Q^{-1}MQ$. Nous avons $A^k=Q^{-1}M^kQ$, et donc aussi

$$M^k=QA^kQ^{-1}. \tag{12.307}$$

En ce qui concerne la norme, si la suite (A^k) est bornée par le réel s, alors

$$\|M^k\|=\|QA^kQ^{-1}\|\leqslant\|Q\|\|Q^{-1}\|\|A^k\|\leqslant s\|Q\|\|Q^{-1}\|. \tag{12.308}$$

Donc la suite (M^k) est bornée et nous pouvons appliquer à M^k ce que nous avons fait sur M. Nous avons donc

$$A=Q^{-1}MQ=Q^{-1}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_s\end{pmatrix}Q \tag{12.309}$$

avec $|\lambda_i|=1$. Nous avons prouvé que A est diagonalisable et que $\mathrm{Spec}(A)\subset S^1$.

□

12.12 Géométrie dans l'espace

12.129.
Les notions de droites, plans et parallélisme sont des notions vectorielles qui auraient pu être traitées beaucoup plus haut. La chose qui rend la géométrie un peu piquante est la notion de perpendicularité. Cette notion demande un produit scalaire et fait intervenir ici et là des polynômes du second degré. Travailler avec le second degré demande la connaissance des racines carrés[76] et donc d'un peu de topologie réelle et de continuité. La résolution dans $\mathbb{R}$ du polynôme du second degré est la proposition 9.117.

75. Notez que la somme part de $l=1$, sinon ce serait raté pour la nilpotence de N'.
76. Définition 9.97.

12.12.1 Droites dans l'espace

La notion d'application affine entre espaces vectoriels est la définition 8.160.

Une application affine n'est pas linéaire, mais presque au sens où des increments égaux dans le paramètre donne des increments égaux dans la valeur.

LEMooEQEAooQaFCMW

Lemme 12.130.
Si f est affine alors pour tout $a, b, v \in V$ nous avons

$$f(a+v) - f(a) = f(b+v) - f(b). \tag{12.310}$$

Démonstration. Simple calcul :

$$f(a+v) - f(a) = u(a+v) + \alpha - u(a) - \alpha = u(a) + u(v) - u(a) = u(v). \tag{12.311}$$

Le même calcul partant de $f(b+v)$ donnera évidement aussi $u(v)$. □

DEFooTQIFooKcloeY

Définition 12.131.
Soit un espace vectoriel E.

1. *Une **droite vectorielle** dans E est un sous-espace vectoriel de dimension 1.*
2. *Une **droite affine** est une partie de E de la forme $a + V$ où $a \in E$ et V est un sous-espace vectoriel de dimension 1 de E.*
3. *Un **plan vectoriel** est un sous-espace vectoriel de dimension 2.*
4. *Une partie P est un **plan affine** si il existe un $v \in E$ tel que $P - v$ soit un plan vectoriel.*

Le plus souvent, nous parlerons de « droite » et « plan » sans préciser « vectoriel » ou « affine ». Dans ces cas, le plus souvent, ce sera « affine ».

Définition 12.132 (Perpendiculaires et parallèles)**.**
Deux notions importantes.

1. *Nous disons que les droites $a + V$ et $b + W$ sont **parallèles** lorsque $V = W$.*
2. *Nous disons que les droites $a + V$ et $b + W$ sont **perpendiculaires** si pour tout $v \in V$ et $w \in W$ nous avons $v \cdot w = 0$.*

Vous noterez que le parallélisme est une notion vectorielle alors que la perpendicularité dépend du produit scalaire ; c'est une notion comme qui dirait « métrique ».

PROPooADJNooMyXUxG

Proposition 12.133.
Les propriétés usuelles.

1. *Deux droites parallèles ayant une intersection sont confondues.*
2. *Le parallélisme est une relation d'équivalence sur l'ensemble des droites de E.*
3. *Si la droite d_1 est parallèle à la droite d_2, alors une droite est perpendiculaire à d_1 si et seulement si elle est perpendiculaire à d_2.*

LEMooRLFQooJADark

Lemme 12.134.
Deux droites perpendiculaires ont un unique point d'intersection.

PROPooPWNWooYuyrOc

Proposition 12.135.
Soient une droite d et un point p.

1. *Il existe une unique droite parallèle à d contenant p.*
2. *Il existe une unique droite perpendiculaire à d contenant p.*

LEMooQQFFooEZYeck

Lemme 12.136.
Si D est une droite et si $a, b \in D$, alors $D - a = D - b$ et $D - a$ est une droite vectorielle.

Démonstration. Vu que D est une droite, il existe $v \in V$ tel que $D - v$ soit une droite vectorielle que nous notons L. Nous allons montrer que $D - a = D - v$. Comme a est arbitraire, cela suffit.

(1) $\underline{D - a \subset D - v}$ Un élément de $D - a$ est de la forme $x - a$ avec $x \in D$. Nous écrivons $x - a$ sous la forme $y - v$ et nous espérons que $y \in D$. Allons-y : d'abord nous isolons y dans $x - a = y - v$:

$$y = x - a + v = (x - v) - (a - v) + v. \tag{12.312a}$$

Puisque $x - v$ et $a - v$ sont des éléments de L, la somme est dans L et donc $y = l + v$ pour un certain élément de $l \in L$. Nous avons donc prouvé que $y \in D$ et donc que $x - a = y - v \in D - v$.

(2) $\underline{D - v \subset D - a}$ Nous notons $x - v$ un élément générique de $D - v$ ($x \in D$). En posant $y - a = x - v$, nous trouvons

$$y = x - v + a = \underbrace{x - v}_{\in L} + \underbrace{(a - v)}_{\in L} + v \tag{12.313}$$

Donc $y \in D$ et $x - v = y - a \in D - a$.

□

PROPooNTHVooWWyafJ

Proposition 12.137.
L'image d'une droite par une application affine[77] *est une droite.*

LEMooRWASooOfwjdw

Lemme 12.138.
Soit un espace vectoriel V sur le corps $\mathbb{K}$.

ITEMooYQCIooOrhRwj

(1) Si L est une droite vectorielle, alors pour tout $a \neq 0$ dans L, nous avons $L = \text{Image}(f)$ où f est l'application linéaire donnée par

$$\begin{aligned} f\colon \mathbb{K} &\to V \\ \lambda &\mapsto \lambda a. \end{aligned} \tag{12.314}$$

ITEMooZIGMooGruFMP

(2) Si D est une droite, alors pour tout $a \neq b$ sur D nous avons $D = \text{Image}(f)$ où f est l'application affine donnée par

$$\begin{aligned} f\colon \mathbb{K} &\to V \\ \lambda &\mapsto a + \lambda(b - a). \end{aligned} \tag{12.315}$$

ITEMooOKJZooIHYDIk

(3) Une partie $D \subset V$ est une droite (affine) si et seulement si il existe $a, v \in V$ tels que

$$D = \{a + \lambda v\}_{\lambda \in \mathbb{K}}. \tag{12.316}$$

Démonstration. En deux parties.

(1) **Pour (1)** Comme L est un sous-espace de dimension 1, il possède une base contenant un unique élément, disons $\{b\}$. En particulier $a = \mu b$ pour un certain $\mu \in \mathbb{K}$. Si $x \in L$ nous avons $x = \lambda_x b$ pour un certain λ_x, et donc

$$x = \frac{\lambda_x}{\mu} a. \tag{12.317}$$

Donc $x = f(\lambda_x/\mu)$. Cela prouve que $L \subset \text{Image}(f)$.

L'inclusion inverse est simplement le fait que $\lambda a \in L$ dès que $a \in L$ parce que L est vectoriel.

(2) **Pour (2)** Le lemme 12.136 nous indique qu'il existe une droite vectorielle L telle que $D - x = L$ pour tout $x \in D$.

77. Définition 8.160.

(2a) $\underline{D \subset \text{Image}(g)}$ Nous nommons $f\colon \mathbb{K} \to V$ l'application linéaire qui donne L. Puisque $b-a \in L$ nous avons

$$f(\lambda) = \lambda(b-a), \tag{12.318}$$

et tout élément de L est de la forme $f(\lambda)$. Nous avons aussi $D = L + a$; donc un élément de D est de la forme $f(\lambda) + a$ et donc de la forme $\lambda(b-a) + a = g(\lambda)$.

(2b) $\underline{\text{Image}(g) \subset D}$ Un élément de Image(g) est de la forme $a + \lambda(b-a)$ avec $\lambda \in \mathbb{K}$. Mais $b - a \in L$, donc $\lambda(b-a) \in L$ et

$$g(\lambda) = a + \lambda(b-a) \in a + L = D. \tag{12.319}$$

□

PROPooRSKLooVrxrFz

Proposition 12.139.
Si a et b sont deux points distincts de $\mathbb{R}^n$, alors il existe une unique droite contenant a et b.

PROPooTFTJooJejuBU

Proposition 12.140.
Soient deux droites d_1 et d_2 dans $\mathbb{R}^n$. Alors nous sommes dans une des trois situations suivantes :

(1) $d_1 \cap d_2 = \varnothing$.

(2) $\text{Card}(d_1 \cap d_2) = 1$

(3) $d_1 = d_2$.

Exemple 12.141.
Les exemples les plus courants d'applications affines sont les droites et les plans ne passant pas par l'origine.

Les droites Une droite dans $\mathbb{R}^2$ (ou $\mathbb{R}^3$) qui ne passe pas par l'origine est l'image d'une fonction de la forme $s(t) = ut + v$, avec $t \in \mathbb{R}$, et u et v dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ selon le cas.

En choisissant des coordonnées adéquates, les droites peuvent être aussi vues comme graphes de fonctions affines. Dans le cas de $\mathbb{R}^2$, on retrouve la fonction de l'exemple 4.33, pour $n = m = 1$.

Les plans De la même façon nous savons que tout plan qui ne passe pas par l'origine dans $\mathbb{R}^3$ est le graphe d'une application affine, $P(x,y) = (a,b)^T \cdot (x,y)^T + (c,d)^T$, lorsque les coordonnées sont bien choisies.

△

12.12.2 Projection orthogonale

Le théorème suivant n'est pas indispensablissime parce qu'il est le même que le théorème de la projection sur les espaces de Hilbert [78]. Cependant la partie existence est plus simple en se limitant au cas de dimension finie.

ThoWKwosrH

Théorème-Définition 12.142 (Théorème de la projection)**.**
Soit E un espace vectoriel réel ou complexe de dimension finie, $x \in E$, et C un sous-ensemble fermé convexe de E.

(1) Les deux conditions suivantes sur $y \in E$ sont équivalentes :

zzETsfYCSItemi

(1a) $\|x-y\| = \inf\{\|x-z\| \text{ tel que } z \in C\}$,

zzETsfYCSItemii

(1b) pour tout $z \in C$, $\text{Re}\langle x-y, z-y\rangle \leqslant 0$.

(2) Il existe un unique $y \in E$, noté $y = \texttt{proj}_C(x)$ vérifiant ces conditions.

Démonstration. Nous commençons par prouver l'existence et l'unicité d'un élément dans C vérifiant la première condition. Ensuite nous verrons l'équivalence.

78. Théorème 25.7

(1) **Existence**
Soit $z_0 \in C$ et $r = \|x - z_0\|$. La boule fermée $\overline{B(x,r)}$ est compacte[79] et intersecte C. Vu que C est fermé, l'ensemble $C' = C \cap \overline{B(x,r)}$ est compact. Tous les points qui minimisent la distance entre x et C sont dans C' ; la fonction

$$\begin{aligned} C' &\to \mathbb{R} \\ z &\mapsto d(x,z) \end{aligned} \tag{12.320}$$

est continue sur un compact et donc a un minimum qu'elle atteint[80]. Un point P réalisant ce minimum prouve l'existence d'un point vérifiant la première condition.

(2) **Unicité** Soient y_1 et y_2, deux éléments de C minimisant la distance avec x, et soit d ce minimum. Nous avons par l'identité du parallélogramme (10.3) que

$$\|y_1 - y_2\|^2 = -4\left\|\frac{y_1 + y_2 - x}{2}\right\|^2 + 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 \leqslant -4d + 2d + 2d = 0. \tag{12.321}$$

Par conséquent $y_1 = y_2$.

(3) **(1a)⇒ (1b)** Soit $z \in C$ et $t \in]0,1[$; nous notons $P = \texttt{proj}_C x$. Vu que y et P sont dans C et que C est convexe[81], le point $z = ty + (1-t)P$ est également dans C, et par conséquent,

$$\|x - P\|^2 \leqslant \|x - tz - (1-t)P\|^2 = \|(x-P) - t(z-P)\|^2. \tag{12.322}$$

Nous sommes dans un cas $\|a\|^2 \leqslant |a-b|^2$, qui implique $2\operatorname{Re}\langle a, b\rangle \leqslant \|b\|^2$. Dans notre cas,

$$2\operatorname{Re}\langle x - P, t(z-P)\rangle \leqslant t^2\|z-P\|^2. \tag{12.323}$$

En divisant par t et en faisant $t \to 0$ nous trouvons l'inégalité demandée[82] :

$$2\operatorname{Re}\langle x - P, z - P\rangle \leqslant 0. \tag{12.324}$$

(4) **(1b)⇒ (1a)**
Soit un point $P \in C$ vérifiant

$$\operatorname{Re}\langle x - P, z - P\rangle \leqslant 0 \tag{12.325}$$

pour tout $z \in C$. Alors en notant $a = x - P$ et $b = P - z$,

$$\begin{aligned} \|x - z\|^2 = \|x - P + P - z\|^2 &= \|a + b\|^2 \\ &= \|a\|^2 + \|b\|^2 + 2\operatorname{Re}\langle a, b\rangle \\ &= \|a\|^2 + \|b\|^2 - 2\operatorname{Re}\langle x - P, z - P\rangle \\ &\geqslant \|b\|^2, \end{aligned} \tag{12.326}$$

ce qu'il fallait.

□

PROPooHGGIooIssaTA

Proposition 12.143.
Soient une droite d dans $\mathbb{R}^3$ ainsi qu'un point p. La projection[83] $\texttt{proj}_d(p)$ *est le point d'intersection*[84] *entre d et la perpendiculaire à d passant par p.*

79. C'est ceci qui ne marche plus en dimension infinie.
80. Théorème 9.55.
81. Définition 7.154.
82. Ici nous utilisons la proposition 12.35, et c'est une des choses qui font que cette partie sur la « géométrie élémentaire » demande en réalité d'être placée après déjà une partie de l'analyse réelle.
83. Définition 12.142.
84. Lemme 12.134.

Démonstration. Nous considérons la droite $d = \{a + \lambda v\}_{\lambda\in\mathbb{R}}$ et un point $p \in \mathbb{R}^3$. Nous notons $x(\lambda) = a + \lambda v$ le point courant dans d. Conformément à la définition 12.142 de la projection orthogonale, nous allons minimiser la distance $\|p - x(\lambda)\|$ par rapport à λ.

Puisque $\|p - x(\lambda)\|$ est toujours positif, nous pouvons chercher à minimiser le carré :

$$\|p - x(\lambda)\|^2 = \|p\|^2 - 2p \cdot a - 2\lambda p \cdot v + \|a\|^2 + |\lambda|^2\|v\|^2 + 2\lambda a \cdot v. \tag{12.327}$$

Quitte à minimiser ça par rapport à λ, nous pouvons oublier les termes ne contenant pas λ. Nous posons donc

$$f(\lambda) = \|v\|^2\lambda^2 + 2(a - p) \cdot v\lambda \tag{12.328}$$

Comme le coefficient de λ^2 est positif, la proposition 9.117 nous dit que cette fonction aura un minimum (et non un maximum). La valeur λ_0 pour laquelle f est minimal se découvre grâce à 9.117(2) :

$$\lambda_0 = \frac{-2(a - p) \cdot v}{2\|v\|^2}. \tag{12.329}$$

Cela est la valeur de λ pour laquelle

$$\texttt{proj}_d(p) = x(\lambda_0); \tag{12.330}$$

nous avons donc

$$x(\lambda_0) = a - \frac{(a - p) \cdot v}{\|v\|^2}v. \tag{12.331}$$

Nous devons voir maintenant que $\big(p - x(\lambda_0)\big) \cdot v = 0$. Il suffit d'un peu déballer :

$$\big(p - x(\lambda_0)\big) \cdot v = p \cdot v - a \cdot v + \frac{(a - p) \cdot v}{\|v\|^2}\|v\|^2 = p \cdot v - a \cdot v + (a - p) \cdot v = 0. \tag{12.332}$$

□

LEMooGUVMooPXtXnV

Lemme 12.144.
Soit $v_1 \in \mathbb{R}^3$. Il existe des vecteurs v_2 et v_3 tels que les v_i sont deux à deux perpendiculaires.

Démonstration. Nous considérons $w \neq v$ dans $\mathbb{R}^3$ et nous profitons de la proposition 10.37 pour poser $v_2 = v_1 \times w$. Enfin nous définissons $v_3 = v_1 \times v_2$. □

LEMooGXGCooDfgbqG

Lemme 12.145.
Soient trois éléments $v_1, v_2, v_3 \in \mathbb{R}^3$ deux à deux perpendiculaires. Si $x \perp v_1$, alors $x \in \text{Span}\{v_2, v_3\}$.

Démonstration. Il faut se rappeler de la proposition 10.16 qui fait de $\{v_1, v_2, v_3\}$ une partie libre. Elle est donc une base par la proposition 4.19(2).

Soit $x \perp v_1$. Nous le décomposons dans la base $\{v_1, v_2, v_3\}$: $x = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. En prenant le produit scalaire par v_1, et en tenant compte du fait que $v_1 \cdot v_2 = v_2 \cdot v_3 = 0$ nous trouvons $0 = v_1 \cdot x = \lambda_1\|v_1\|^2$. Donc $\lambda_1 = 0$ et $x \in \text{Span}\{v_2, v_3\}$. □

12.12.3 Plan médiateur

PROPooSNUDooTxovMz

Proposition 12.146 (plan médiateur[1])**.**
Soient un espace euclidien V ainsi que deux points distincts $a, b \in V$. Si $m = (a+b)/2$, nous avons

$$\{x \in V \text{ tel que } x - m \perp b - a\} = \{x \in V \text{ tel que } \|x - a\| = \|x - b\|\}. \tag{12.333}$$

Dans le cas de $V = \mathbb{R}^3$, alors cet ensemble est un plan[85]*.*

*Ce plan est le **plan médiateur** du segment $[a, b]$.*

85. Définition 12.131.

Démonstration. Nous notons

$$M = \{x \in V \text{ tel que } x - m \perp b - a\}, \tag{12.334a}$$
$$N = \{x \in V \text{ tel que } \|x - a\| = \|x - b\|\}. \tag{12.334b}$$

(1) $M \subset N$ Soit $x \in M$. Nous avons $(x - m) \cdot (b - a) = 0$, et nous pouvons utiliser Pythagore 10.25 dans les triangles xbm et xma : SUBEQSooVEPCooKnyPoq

$$\|x - a\|^2 = \|x - m\|^2 + \|a - m\|^2 \tag{12.335a}$$
$$\|x - b\|^2 = \|x - m\|^2 + \|m - b\|^2. \tag{12.335b}$$

Vu que m est le milieu, nous avons $a-m = m-b$ et donc $\|a-m\| = \|m-b\|$. Nous voyons donc que les membres de droites des deux équations (12.335) sont égaux. Donc $\|x-a\|^2 = \|x-b\|^2$. Comme une norme est toujours positive, les carrés peuvent être simplifiés : $\|x-a\| = \|x-b\|$. Donc $x \in N$.

(2) $N \subset M$ Soit $x \in N$. Nous posons $h = \texttt{proj}_{(ab)}(x)$, la projection de x sur la droite (ab). La proposition 12.143 nous dit que h est l'unique point de (ab) tel que $x - h \perp b - a$.
Le théorème de Pythagore 10.25 dans le triangle ahx donne

$$\|x - a\|^2 = \|a - h\|^2 + \|x - h\|^2 \tag{12.336}$$

et dans le triangle bhx il donne :

$$\|b - x\|^2 = \|b - h\|^2 + \|h - x\|^2. \tag{12.337}$$

Par hypothèse nous avons $\|x - a\|^2 = \|x - b\|^2$ et donc

$$\|a - h\| = \|b - h\|. \tag{12.338}$$

Nous cherchons à présent quel(s) point(s) h de la droite (ab) vérifie(nt) cette condition, et nous espérons que ce sera $(a + b)/2$.
Nous cherchons h sous la forme $h = a+\lambda(b-a)$. D'une part nous avons $\|a-h\|^2 = \|\lambda(b-a)\|^2 = \lambda^2\|b - a\|^2$, et d'autre part

$$\|b - h\|^2 = \|b - a - \lambda(b - a)\|^2 = |1 - \lambda|^2\|b - a\|^2 \tag{12.339}$$

Nous en déduisons que $|\lambda| = |1 - \lambda|$. Cela laisse deux possibilités : la première est $\lambda = 1 - \lambda$ qui donne $\lambda = 1/2$ et la seconde est $\lambda = -(1 - \lambda)$ qui est impossible. Donc $\lambda = 1/2$ et

$$h = a + \frac{b - a}{2} = \frac{a + b}{2}. \tag{12.340}$$

Donc en posant $m = (a + b)/2$ nous avons bien $b - a \perp x - m$.

(3) **C'est un plan** Nous nous mettons maintenant dans le cas où V est l'espace $\mathbb{R}^3$ muni de sa norme usuelle. Posons $f_1 = b - a$ et considérons deux vecteurs f_2, f_3 tels que les f_i soient deux à deux perpendiculaires (lemme 12.144).
Nous allons prouver que $M = \text{Span}\{f_2, f_3\} + m$.

(3a) **Une inclusion** Si $x \in \text{Span}(f_2, f_3) + m$, alors $x = \alpha f_2 + \beta f_3 + m$ et nous avons bien $x - m \perp b - a$.

(3b) **L'autre inclusion** Soit $x \in M$. Donc $x - m \perp b - a$. Le lemme 12.145 nous indique alors que $x - m \in \text{Span}\{f_2, f_3\}$, ce qu'il fallait.

□

12.12.4 Tétraèdre

DEFooMUUMooFVxKyb

Définition 12.147 ([1]).
*Un **tétraèdre régulier** est un ensemble de 4 points A, B, C et D de $\mathbb{R}^3$ deux à deux équidistants.*

Nous allons nommer $\{a_i\}$ les segments entre les points, $\{d_i\}$ les droites sur ces segments, et $\{s_i\}$ les sommets.

Lemme 12.148.
Un tétraèdre régulier existe.

Démonstration. Prenez un triangle équilatéral ABC dans le plan $(.,.,0)$, et prenez ensuite un point D à la verticale du centre, placé à la bonne hauteur pour que les longueurs $\|AD\|$, $\|BD\|$ et $\|CD\|$ soient égales à $\|AB\|$. □

LEMooNWELooZeSEMN

Lemme 12.149.
Si T est un tétraèdre régulier, nous avons $d_i \cap T = a_i$.

LEMooUSKVooQJiBuz

Lemme 12.150.
Les droites $\{d_i\}_{i=1,\dots,6}$ ne sont pas confondues ni parallèles.

Démonstration. Si trois points A, B, C sont alignés, il n'est pas possible d'avoir $\|AB\| = \|AC\| = \|BC\|$. Donc il n'y a pas deux droites parmi les $\{d_i\}$ qui sont confondues.

Supposons que deux des droites AB et CD sont parallèles. En particulier, les points A, B, C et D sont dans un même plan : le plan $A + \text{Span}\{B - A, C - A\}$. Il n'est pas possible d'avoir 4 points dans un plan, tous équidistants deux à deux. □

Dans la suite, quand nous parlerons du « tétraèdre », nous parlerons de ses six points et six segments les joignant. L'ensemble $T \subset \mathbb{R}^3$ ne contient pas les surfaces et les volumes.

LEMooJCMKooOjMqtw

Lemme 12.151.
Soit un tétraèdre régulier T. Un point de $\mathbb{R}^3$ est un sommet si et seulement si il est l'intersection de deux des droites $\{d_i\}$ différentes.

Démonstration. En deux parties.

(1) **Sens direct**
Par définition les sommets sont les points A, B, C, D ; et les droites d_i sont les droites (AB), (AC), (AD), (BC), (DB) et (CD). Donc oui, les sommets sont à des intersections de ces droites.

(2) **Sens inverse** Soit un point $X \in \mathbb{R}^3$ à l'intersection entre deux des d_i. Nous avons déjà vu dans le lemme 12.150 que ces droites ne sont ni parallèles ni confondues. Donc elles ont au plus un point d'intersection. Voyons les couples possibles de droites.
On a une série de possibilités comme $(AB) \cap (AC)$. Dans ce cas, l'intersection entre ces deux droits est A qui est un des sommets. Ensuite nous avons une série de possibilités comme $(AB) \cap (CD)$. Ces deux droites n'ont pas d'intersection parce que si elles en avaient, les points A, B, C et D seraient dans le même plan, ce qui est impossible. Donc deux droites d_i ont soit, pas d'intersection, soit, une intersection qui est un sommet.

□

12.13 Géométrie dans le plan

LEMooYIHXooEwmlPo

Lemme 12.152 (Équation de droite).
Si D est une droite[86] dans $\mathbb{R}^2$, alors D est d'une des deux formes suivantes :

86. Définition 12.131, mais c'est surtout la caractérisation du lemme 12.138(3) que nous devons avoir en tête.

— *Soit il existe $\alpha \in \mathbb{R}$ tel que*

$$D = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x = \alpha\}, \tag{12.341}$$

— *soit il existe $\alpha, \beta \in \mathbb{R}$ tels que*

$$D = \{(x, y) \in \mathbb{R}^2 \text{ tel que } y = \alpha x + \beta\}. \tag{12.342}$$

Le premier cas correspond aux droites verticales.

Démonstration. Nous prenons une droite D et une application affine f donnée par (12.138)(3). Il y a deux possibilités : soit $b - a$ est un multiple de e_2, sois pas.

(1) **Si $b - a = \mu e_2$** Alors $f(\lambda) = a + \lambda\mu e_2$. Nous avons alors

$$\begin{aligned} D &= \{a + \lambda e_2\}_{\lambda\in\mathbb{R}} && (12.343a)\\ &= \{\lambda e_2 + a_1 e_1\} && \text{cf. justif.} \quad (12.343b)\\ &= \{(a_1, \lambda)\}_{\lambda\in\mathbb{R}} && (12.343c)\\ &= \{(x, y) \in \mathbb{R}^2 \text{ tel que } x = a_1\}. && (12.343d) \end{aligned}$$

SUBEQooEKMNooQgkSeC

Justifications.

— Pour 12.343b. Parce que $a_2 e_2 + \lambda e_2$ prend toutes les valeurs de la forme λe_2.

(2) **Sinon** Nous avons

$$f(\lambda) = \begin{pmatrix} a_1 + \lambda(b - a)_1 \\ a_2 + \lambda(b - a)_2 \end{pmatrix} \tag{12.344}$$

En posant $\alpha = (b - a)_2/(b - a)_1$ et $\beta = -\alpha a_1 + a_2$, nous avons bien

$$f(\lambda)_2 = \alpha f(\lambda)_1 + \beta. \tag{12.345}$$

□

PROPooOFKXooARiQZK

Proposition 12.153.
Si D est une droite dans $\mathbb{R}^2$, il existe une application affine[87] *$f : \mathbb{R}^2 \to \mathbb{R}$ telle que*

$$D = \{x \in \mathbb{R}^2 \text{ tel que } f(x) = 0\}. \tag{12.346}$$

Nous disons qu'une telle application affine est une application associée à D.

Si f est une application affine quelle que $f(x) = 0$ donne la droite D, alors pour tout réel non nul λ, l'application affine λf donnent également D. Il n'y a donc pas d'unicité.

Définition 12.154.
*Soit une application affine $f : \mathbb{R}^2 \to \mathbb{R}$. Nous appelons **demi-plans** associés à f les parties*

$$H_f^+ = \{x \in \mathbb{R}^2 \text{ tel que } f(x) > 0\} \tag{12.347}$$

et

$$H_f^- = \{x \in \mathbb{R}^2 \text{ tel que } f(x) < 0\}. \tag{12.348}$$

LEMooEZEDooSDbrkq

Lemme 12.155 ([1])**.**
Les demi-plans sont convexes[88]*.*

87. Définition 8.160.
88. Définition 7.154.

Démonstration. Soit une applications affine $f\colon \mathbb{R}^2 \to \mathbb{R}$ ainsi que $a, b \in \mathbb{R}^2$ tels que $f(a) > 0$ et $f(b) > 0$. Vu que f est affine, il existe une application linéaire $l\colon \mathbb{R}^2 \to \mathbb{R}$ ainsi que $\alpha \in \mathbb{R}$ tels que $f(x) = l(x) + \alpha$ pour tout x.

Nous considérons

$$\begin{aligned}\gamma\colon \mathbb{R} &\to \mathbb{R}^2\\ t &\mapsto a + t(b-a).\end{aligned} \tag{12.349}$$

Nous devons prouver que $(f \circ \gamma)(t) > 0$ pour tout $t \in [0,1]$.

Nous avons $(f \circ \gamma)(0) > 0$ et $(f \circ \gamma)(1) > 0$. Nous avons d'abord

$$\begin{aligned}(f \circ \gamma)(t) &= l\big(a + t(b-a)\big) + \alpha && (12.350a)\\ &= l(a) + fl(b) - tl(a) + \alpha && (12.350b)\\ &= (1-t)l(a) + tl(b) + t\alpha + (1-t)\alpha && (12.350c)\\ &= (1-t)f(a) + tf(b). && (12.350d)\end{aligned}$$

SUBEQooXZIPooMoICRm

Les nombres $f(a)$ et $f(b)$ sont strictement positifs. Les nombres $(1-t)$ et t sont positifs, mais ne s'annulent pas en même temps. Donc dans (12.350d), au moins un des deux termes est strictement positifs tandis que l'autre est positif ou nul. Bref, $(f \circ \gamma)(t) > 0$.

Cela prouve que le demi-plan $f(x) > 0$ est convexe. Le même raisonnement tient pour le demi-plan $f(x) < 0$. □

LEMooXLNZooFyqPxG

Lemme 12.156 ([1])**.**
Si $f, g\colon \mathbb{R}^2 \to \mathbb{R}$ sont affines et si $\ker(f) = \ker(g)$, alors

$$\{H_f^+, H_f^-\} = \{H_g^+, H_g^-\}. \tag{12.351}$$

Démonstration. Il existe un $a \in \mathbb{R}^2$ tel que $f(a) > 0$. En effet $f(x) = l(x) + \alpha$ où $l\colon \mathbb{R}^2 \to \mathbb{R}$ est linéaire et $\alpha \in \mathbb{R}$. Il suffit de prendre x tel que $l(x) < -\alpha$. Soit $a \in \mathbb{R}^2$ tel que $f(a) > 0$. Il y a deux possibilités : $g(a) > 0$ ou $g(a) < 0$ parce que $g(a) = 0$ n'est pas possible du fait que f et g s'annulent aux mêmes points..

(1) **Si $g(a) > 0$** Nous allons prouver qu'alors $H_f^+ = H_g^+$ et $H_f^- = H_g^-$. Soit $b \in H_f^+$. Nous savons que H_f^+ est convexe (lemme 12.155), de telle sorte que $f\big([a,b]\big) > 0$. En particulier f ne s'annule pas sur le segment $[a,b]$, et g non plus. Autrement dit, la fonction

$$\begin{aligned}s\colon [0,1] &\to \mathbb{R}\\ t &\mapsto g\big((1-t)a + tb\big)\end{aligned} \tag{12.352}$$

ne s'annule pas. Vu que $s(0) = g(a) > 0$, le théorème des valeurs intermédiaires[89] nous indique que $g(b) = s(1) > 0$. Donc $b \in H_g^+$.

Nous avons prouvé que $H_f^+ \subset H_g^+$. En inversant les rôles de f et g nous prouvons que $H_g^+ \subset H_f^+$.

(2) **Si $g(a) < 0$** Il se prouve de même que $H_f^+ = H_g^-$ et $H_f^- = H_g^+$.

□

PROPooSIANooTOKlBI

Proposition 12.157.
Si $f\colon \mathbb{R}^2 \to \mathbb{R}$ est une application affine non constante, alors $\ker(f)$ est une droite[90]*.*

Démonstration. Posons $f(x) = l(x) + \alpha$ où $l\colon \mathbb{R}^2 \to \mathbb{R}$ est linéaire et $\alpha \in \mathbb{R}$.

Commençons avec $\alpha = 0$. Considérons $a \in \ker(f)$ et prouvons que $\ker(f) = \{\lambda a\}_{\lambda \in \mathbb{R}}$. D'abord $f(\lambda a) = l(\lambda a) = \lambda l(a) = 0$. Donc $\{\lambda a\}_{\lambda \in \mathbb{R}} \subset \ker(f)$. D'autre part si $f(x) = 0$ alors que x n'est pas de la forme λa. Dans ce cas, $\{x, a\}$ forment une base de $\mathbb{R}^2$ et nous concluons que $\ker(f) = \mathbb{R}^2$, ce qui est contraire à l'hypothèse comme quoi $f \neq 0$.

Nous ne supposons plus que $\alpha = 0$. Soit $a \in \mathbb{R}^2$ tel que $l(a) = -\alpha$. Nous allons prouver que $\ker(f) = a + \ker(l)$.

89. Théorème 9.93.
90. La caractérisation 12.138(3) est plus pratique que la définition.

(1) $a+\ker(l)\subset\ker(f)$ Soit $z\in\ker(l)$. Nous avons $f(a+z)=l(a)+l(z)+\alpha=-\alpha+0+\alpha=0$. Donc $a+z\in\ker(f)$.

(2) $\ker(f)\subset a+\ker(l)$. Soit $b\in\ker(f)$. Nous prouvons que $b-a\in\ker(l)$. Nous avons $0=f(b)=l(b)+\alpha$ et donc $l(b)=-\alpha$. Donc

$$l(b-a)=l(b)-l(a)=-\alpha+\alpha=0. \tag{12.353}$$

parce que $l(a)=-\alpha$.

La conclusion est que $\ker(f)=a+\ker(l)$. La première partie ayant déjà montré que $\ker(l)$ est une droite, nous avons fini. □

LEMooTOUOooJbpLVB

Lemme 12.158 ([1, 162]).
Soit une fonction affine $f\colon\mathbb{R}^2\to\mathbb{R}$. Soient $a,b\in\mathbb{R}^2$ tels que $f(a),f(b)\neq0$.

ITEMooXLCFooQiDCrR

(1) L'intersection $[a,b]\cap\ker(f)$ contient 0 ou 1 point.

ITEMooMORGooTzNefp

(2) Si $[a,b]\cap\ker(f)=\varnothing$ alors $f(a)$ et $f(b)$ ont même signe.

ITEMooDWIQooXnoMJT

(3) Si $[a,b]\cap\ker(f)\neq\varnothing$ alors $f(a)$ et $f(b)$ sont de signe opposés.

Démonstration. Point par point.

(1) **Pour (1)** Nous savons que $\ker(f)$ est une droite (proposition 12.157). Vu que $f(a)\neq0$, l'unique droite passant par a et b (proposition 12.139) n'est pas la droite $\ker(f)$. Donc parmi les trois possibilités de la proposition 12.139, nous sommes forcément dans le cas où l'intersection est vide ou réduite à un unique point.

(2) **Pour (2)** Nous supposons que $\ker(f)\cap[a,b]=\varnothing$. Considérons la fonction

$$\begin{aligned}s\colon[0,1]&\to\mathbb{R}\\ t&\mapsto f\big((1-t)a+tb\big).\end{aligned} \tag{12.354}$$

La fonction s est continue et ne s'annule pas parce que les points $(1-t)a+tb$ sont ceux de $[a,b]$. Si $s(0)$ et $s(1)$ étaient de signe différents, le théorème des valeurs intermédiaires 9.93 donnerait un $t_0\in]0,1[$ tel que $s(t_0)=0$. Donc $s(0)$ et $s(1)$ ont le même signe.

(3) **Pour (3)** Nous prouvons la contraposée. Supposons que $f(a)$ et $f(b)$ aient le même signe (strict), et disons que $f(a),f(b)>0$ pour fixer les idées. Cela signifie que a et b sont dans le demi-plan H_f^+. Par convexité de ce dernier (lemme 12.155), nous avons $f(x)>0$ pour tout $x\in[a,b]$.

□

LEMooBWFSooBkIcOg

Lemme 12.159 ([1]).
Soient une application affine f. Soient $a\in\ker(f)$ et $v\in\mathbb{R}^2$ tels que[91] $\ker(f)=\{a+\lambda v\}_{\lambda\in\mathbb{R}}$. *Soit un vecteur non nul $w\in\mathbb{R}^2$ non parallèle à v.*

ITEMooYDYPooRzupcO

(1) $\{v,w\}$ est une base de $\mathbb{R}^2$.

ITEMooHKVWooLAEMTm

(2) Nous considérons l'application

$$\begin{aligned}p\colon\mathbb{R}^2&\to\mathbb{R}\\ \lambda v+\mu w&\mapsto\mu.\end{aligned} \tag{12.355}$$

Alors l'application

$$\begin{aligned}g\colon\mathbb{R}^2&\to\mathbb{R}\\ x&\mapsto p(x-a)\end{aligned} \tag{12.356}$$

est affine.

ITEMooGRZGooMFQPjM

(3) Les demi-plans de f et de g sont les mêmes.

91. Si on en croit la proposition 12.157, ça existe parce que le noyau d'une application affine est une droite.

Démonstration. Point par point.

(1) **Pour (1)** Les vecteurs v et w ne sont pas colinéaire et forment donc une base par le lemme 4.10(1).

(2) **Pour (2)** L'application p est linéaire. Nous avons

$$g(x) = p(x - a) = p(x) - p(a). \tag{12.357}$$

L'application g est donc bien affine.

(3) **Pour (3)** Les applications f et g sont affines. Si $x \in \ker(g)$, alors $p(x - a) = 0$, c'est-à-dire que $x - a = \lambda v$ et donc $x = \lambda v + a \in \ker(f)$. Donc $\ker(g) \subset \ker(f)$. De même nous trouvons que $\ker(f) \subset \ker(g)$.
Deux applications affines ayant le même noyau ont les mêmes demi-plans par le lemme 12.156.

□

LEMooURPNooYRsuaI

Lemme 12.160 ([1]).
Soit une application affine f ainsi qu'une application continue $g: \mathbb{R}^2 \to \mathbb{R}$ telles que $\ker(f) = \ker(g)$. Supposons l'existence de $a, b \in \mathbb{R}^2$ tels que $g(a) > 0$ et $g(b) < 0$.
Alors les demi-plans de f sont les parties $\{x \text{ tel que } g(x) > 0\}$ et $\{x \text{ tel que } g(x) < 0\}$.

PROPooTPHKooXnpJaV

Proposition 12.161 ([1]).
Soit une droite d dans $\mathbb{R}^2$ et une chemin dérivable $\gamma: \mathbb{R} \to \mathbb{R}^2$. Nous supposons que $\gamma(t_0) \in d$ et que $\gamma'(t_0)$ est non nul et non parallèle à d.
Alors il existe δ tel que pour tout $\epsilon < \delta$, $\gamma(t_0 + \epsilon)$ est dans un demi-plan de d et $\gamma(t_0 - \epsilon)$ est dans l'autre demi-plan de d.

Démonstration. Posons $a = \gamma(t_0)$. Nous considérons un vecteur $b_1 \in \mathbb{R}^2$ tel que $d = \{a + \lambda b_1\}_{\lambda \in \mathbb{R}}$, ainsi que b_2 tel que $\{b_1, b_2\}$ soit une base de $\mathbb{R}^2$.

Tout élément de $\mathbb{R}^2$ peut être écrit de façon unique sous la forme

$$x = a + x_1 b_1 + x_2 b_2. \tag{12.358}$$

Cela nous donne des fonctions continues $\sigma_i: \mathbb{R} \to \mathbb{R}$ telles que

$$\gamma(t) = a + \sigma_1(t) b_1 + \sigma_2(t) b_2, \tag{12.359}$$

et que

$$\gamma'(t) = \sigma_1'(t) b_1 + \sigma_2'(t) b_2. \tag{12.360}$$

Vu que $\gamma'(t_0)$ n'est pas parallèle à la droite d, nous avons $\sigma_2'(t_0) \neq 0$. De plus $\gamma(t_0) = a$, de telle sorte que $\sigma_1(t_0) = \sigma_2(t_0) = 0$.

Que dites-vous ? La fonction dérivable $\sigma_2: \mathbb{R} \to \mathbb{R}$ vaut zéro en t_0 et sa dérivée y est non nulle ? Supposons pour fixer les idées que $\sigma_2'(t_0) > 0$. Il existe donc un $\epsilon > 0$ tel que σ_2 est strictement positive sur $[t_0 - \epsilon, t_0[$ et strictement négative sur $]t_0, t_0 + \epsilon]$.

En vertu du lemme 12.159, les points de γ tels que $\sigma_2 > 0$ sont dans un demi-plan de d et les points de γ avec $\sigma_2 < 0$ sont dans l'autre demi-plan. □

THOooFMMLooLmAnAd

Théorème 12.162 (Théorème de Thalès[395]).
Soient trois points A, B, C non alignés dans $\mathbb{R}^2$. Soient $D \in (AB)$ et $E \in (AC)$. Nous supposons que (DE) est parallèle à BC.
Alors

(1)

$$\frac{\|D - A\|}{\|B - A\|} = \frac{\|E - A\|}{\|C - A\|} = \frac{\|E - D\|}{\|C - B\|} \tag{12.361}$$

(2) Il existe une homothétie $\phi: \mathbb{R}^2 \to \mathbb{R}^2$ centrée en A telle que $\phi(B) = D$ et $\phi(C) = E$.

THOooGFTWooACQLFJ

Théorème 12.163 (Théorème de Thalès dans le cercle[396]).
Soient des points O, A, B, C dans $\mathbb{R}^2$ tels que

(1) $\|A-O\| = \|B-O\| = \|C-O\|$.

(2) *A, O et B sont alignés.*

Alors le triangle ABC est rectangle en C.

12.14 Dérivée : exemples introductifs

12.14.1 La vitesse

Lorsqu'un mobile se déplace à une vitesse variable, nous obtenons la *vitesse instantanée* en calculant une vitesse moyenne sur des intervalles de plus en plus petits. Si le mobile a un mouvement donné par $x(t)$, la vitesse moyenne entre $t=2$ et $t=5$ sera

$$v_{\text{moy}}(2 \to 5) = \frac{x(5)-x(2)}{5-2}.$$

Plus généralement, la vitesse moyenne entre 2 et $2+\Delta t$ est donnée par

$$v_{\text{moy}}(2 \to 2+\Delta t) = \frac{x(2+\Delta t)-x(2)}{\Delta t}.$$

Cela est une fonction de Δt. Oui, mais rappellons qu'on a dans l'idée de calculer une vitesse instantanée, c'est-à-dire de voir ce que vaut la vitesse moyenne sur un intervalle très très très très petit. La notion de limite semble toute indiquée pour décrire mathématiquement l'idée physique de vitesse instantanée.

Nous allons dire que la vitesse instantanée d'un mobile est la limite quand Δt tend vers zéro de sa vitesse moyenne sur l'intervalle de temps Δt, ou en formule :

$$v(t_0) = \lim_{\Delta t \to 0} \frac{x(t_0)-x(t_0+\Delta t)}{\Delta t}. \tag{12.362}$$ Eqvinstlimite

12.14.2 La tangente à une courbe

Passons maintenant à tout autre chose, mais toujours dans l'utilisation de la notion de limite pour résoudre des problèmes intéressants. Comment trouver l'équation de la tangente à la courbe $y=f(x)$ au point $(x_0, f(x_0))$?

Essayons de trouver la tangente au point P donné de la courbe donnée à la figure 12.1.

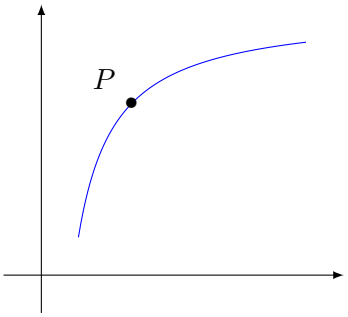


FIGURE 12.1: Comment trouver la tangente à la courbe au point P ? LabelFigTangenteQuestion

La tangente est la droite qui touche la courbe en un seul point sans la traverser. Afin de la construire, nous allons dessiner des droites qui touchent la courbe en P et un autre point Q, et nous allons voir ce qu'il se passe quand Q est très proche de P. Cela donnera une droite qui, certes, touchera la courbe en deux points, mais en deux points *tellement proches que c'est comme si c'étaient les mêmes*. On sent que la notion de limite va encore intervenir.

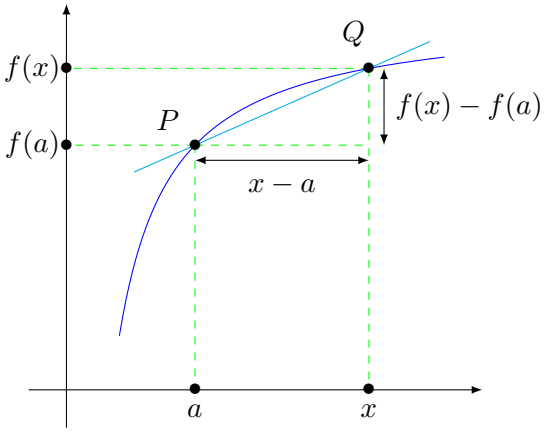


FIGURE 12.2: Traçons d'abord une corde entre le point P et un point Q un peu plus loin.LabelFigTangenteDetail

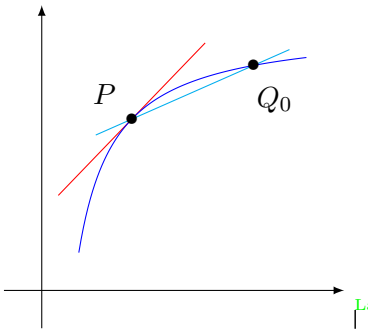
Nous avons placé le point, sur la figure 12.2, le point P en a et le point Q un peu plus loin, en x. En d'autres termes leurs coordonnées sont

$$P = \big(a, f(a)\big) \qquad\qquad Q = \big(x, f(x)\big). \tag{12.363}$$

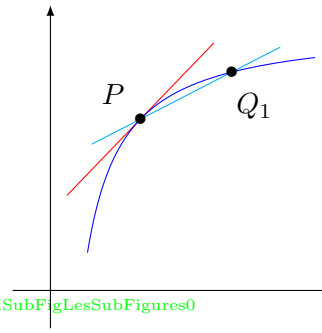
En regardant par exemple la figure 12.2, le coefficient directeur de la droite qui passe par ces deux points est donné par

$$\frac{f(x) - f(a)}{x - a}, \tag{12.364}$$

et bang ! Encore le même rapport que celui qu'on avait trouvé à l'équation (12.362) en parlant de vitesses. En regardant la figure 12.3, on constate réellement qu'en faisant tendre x vers a, on obtient la tangente.

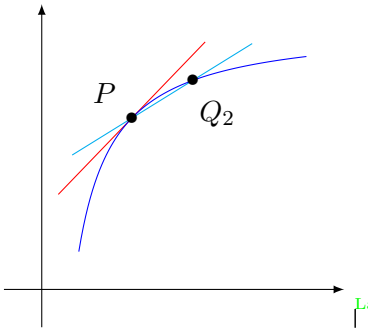


(a) Pas très bon ...

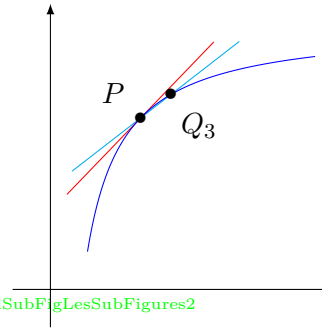


LabelFigLesSubFigureslabelSubFigLesSubFigures0 LabelFigLesSubFigures

(b) ...de mieux en mieux ...



(c) ...de mieux en mieux ...



LabelFigLesSubFigureslabelSubFigLesSubFigures2 LabelFigLesSubFigures

(d) ...presque parfait

FIGURE 12.3: Recherche de la tangente par approximations successives.LabelFigLesSubFigures

12.14.3 L'aire en dessous d'une courbe

SubSecAirePrimInto

Encore un exemple. Nous voudrions bien pouvoir calculer l'aire en dessous d'une courbe. Nous notons $S_f(x)$ l'aire en dessous de la fonction f entre l'abscisse 0 et x, c'est-à-dire l'aire bleue de la figure 12.4.

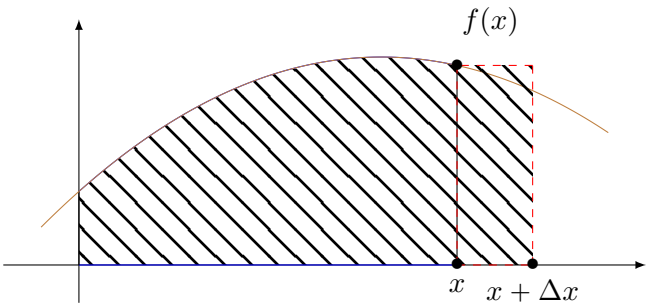


FIGURE 12.4: L'aire en dessous d'une courbe. Le rectangle rouge d'aire $f(x)\Delta x$ approxime l'augmentation de l'aire lorsqu'on passe de x à $x + \Delta x$.

LabelFigNOCGooYRHLCn

Si la fonction f est continue et que Δx est assez petit, la fonction ne varie pas beaucoup entre x et $x+\Delta x$. L'augmentation de surface entre x et $x+\Delta x$ peut donc être approximée par le rectangle de surface $f(x)\Delta x$. Ce que nous avons donc, c'est que quand Δx est très petit,

$$S_f(x + \Delta x) - S_f(x) = f(x)\Delta x, \tag{12.365}$$

c'est-à-dire

$$f(x) = \lim_{\Delta x \to 0} \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{12.366}$$

Donc, la fonction f est la dérivée de la fonction qui représente l'aire en dessous de f. Calculer des surfaces revient donc au travail inverse de calculer des dérivées.

Nous avons déjà vu que calculer la dérivée d'une fonction n'est pas très compliqué. Aussi étonnant que cela puisse paraitre, il se fait que le processus inverse est très compliqué : il est en général extrêmement difficile (et même souvent impossible) de trouver une fonction dont la dérivée est une fonction donnée.

Une fonction dont la dérivée est la fonction f s'appelle une **primitive** de f, et la fonction qui donne l'aire en dessous de la fonction f entre l'abscisse 0 et x est notée

$$S_f(x) = \int_0^x f(t)dt. \tag{12.367}$$

Nous pouvons nous demander si, pour une fonction f donnée, il existe une ou plusieurs primitives, c'est-à-dire si il existe une ou plusieurs fonctions F telles que $F' = f$. La réponse viendra par le corolaire 12.199.

12.15 Dérivation de fonctions réelles

seccontetderiv

Exemple 12.164.
Montrons que la fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto x$ est continue et dérivable [92]. Exceptionnellement (bien qu'on sache que la dérivabilité implique la continuité), montrons ces deux assertions séparément.

Continuité Pour prouver la continuité au point $a \in \mathbb{R}$ nous devons montrer que

$$\lim_{x \to a} x = a \tag{12.368}$$

c'est-à-dire

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in \mathbb{R} \, |x - a| < \delta \Rightarrow |x - a| < \epsilon \tag{12.369}$$

ce qui est clair en prenant $\delta = \epsilon$.

92. Définition 10.245.

Dérivabilité Soit $a \in \mathbb{R}$. Calculons la limite du quotient différentiel

$$\lim_{\substack{x\to a\\ x\neq a}} \frac{x-a}{x-a} = \lim_{\substack{x\to a\\ x\neq a}} 1 = 1 \tag{12.370}$$

ce qui prouve que f est dérivable et que sa dérivée vaut 1 en tout point a de $\mathbb{R}$.

On a donc montré que la fonction $x \mapsto x$ est continue, dérivable, et que sa dérivée vaut 1 en tout point a de son domaine.

△

PropSFyxOWF

Proposition 12.165.
Une fonction dérivable sur un intervalle est continue sur cet intervalle.

Démonstration. Soit I un intervalle sur lequel la fonction f est dérivable, et soit $x_0 \in I$. Nous allons prouver la continuité de f en x_0. Le fait que la limite

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} \tag{12.371}$$

existe implique a fortiori que

$$\lim_{h\to 0} f(x_0+h) - f(x_0) = 0. \tag{12.372}$$

Cela signifie la continuité de f en vertu du critère 12.53. □

THOooFFOZooCYGets

Théorème 12.166.
Toute fonction dérivable en un point est continue en ce point.

Démonstration. Soient $f\colon \mathbb{R} \to \mathbb{R}$ et $a \in \mathbb{R}$. Nous supposons que f n'est pas continue en a et nous allons en déduire qu'elle n'est pas non plus dérivable en a. Pour cela nous considérons le lien entre limite et continuité donné dans le théorème 12.53. Nier que f est continue en a revient à dire qu'il existe un voisinage V de $f(a)$ tel que

$$\forall r > 0,\ \exists \epsilon < r \text{ tel que } f(a+\epsilon) \notin V. \tag{12.373}$$

Si $B\big(f(a), R\big) \subset V$ [93], et si $r = 1/n$, nous construisons une suite $\epsilon_n \to 0$ telle que

$$|f(a+\epsilon_n) - f(a)| > R. \tag{12.374}$$

Avec cela nous avons

$$\frac{|f(a+\epsilon_n)-f(a)|}{\epsilon_n} > \frac{R}{\epsilon_n} \to \infty. \tag{12.375}$$

Donc la fonction f ne peut pas être dérivable en a. □

Remarque 12.167.
La réciproque du théorème précédent n'est pas vraie : il existe bien des fonctions qui sont continues en un point x_0, mais qui ne sont pas dérivables en x_0. La fonction valeur absolue, $x \mapsto |x|$, par exemple est continue sur tout $\mathbb{R}$ mais elle n'est pas dérivable en 0.

Si f est une fonction dérivable, il peut arriver que la fonction dérivée f' soit elle-même dérivable. Dans ce cas nous notons f'' ou $f^{(2)}$ la dérivée de la fonction f'. Cette fonction f'' est la **dérivée seconde** de f. Elle peut encore être dérivable ; dans ce cas nous notons $f^{(3)}$ sa dérivée, et ainsi de suite. Nous définissons $f^{(n)} = (f^{(n-1)})'$ la dérivée n^{e} de f. Nous posons évidemment $f^{(0)} = f$.

93. Existence par la définition de la topologie métrique 7.117.

12.15.1 Exemples

Exemple 12.168.
Commençons par la fonction $f(x) = x$. Dans ce cas nous avons

$$\frac{f(x) - f(a)}{x - a} = \frac{x - a}{x - a} = 1. \tag{12.376}$$

La dérivée est donc 1. △

Proposition 12.169.
La dérivé de la fonction $x \mapsto x$ vaut 1, en notations compactes : $(x)' = 1$.

Démonstration. D'après la définition de la dérivée, si $f(x) = x$, nous avons

$$f(x) = \lim_{\epsilon \to 0} \frac{(x + \epsilon) - x}{\epsilon} = \lim_{\epsilon \to 0} \frac{\epsilon}{\epsilon} = 1, \tag{12.377}$$

et c'est déjà fini. □

12.15.1.1 La fonction carré

Prenons ensuite $f(x) = x^2$. En utilisant le produit remarquable $(x^2 - a^2) = (x - a)(x + a)$ nous trouvons

$$\frac{f(x) - f(a)}{x - a} = x + a. \tag{12.378}$$

Lorsque $x \to a$, cela devient $2a$. Nous avons par conséquent

$$f'(x) = 2x. \tag{12.379}$$

LemDeccCarr

Lemme 12.170.
Si $f(x) = x^2$, alors $f'(x) = 2x$.

Démonstration. Utilisons la définition, et remplaçons f par sa valeur :

$$\begin{aligned}
f'(x) &= \lim_{\epsilon \to 0} \frac{f(x + \epsilon) - f(x)}{\epsilon} && \text{(12.380a)}\\
&= \lim_{\epsilon \to 0} \frac{(x + \epsilon)^2 - x^2}{\epsilon} && \text{(12.380b)}\\
&= \lim_{\epsilon \to 0} \frac{x^2 + 2x\epsilon + \epsilon^2 - x^2}{\epsilon} && \text{(12.380c)}\\
&= \lim_{\epsilon \to 0} \frac{\epsilon(2x + \epsilon)}{\epsilon} && \text{(12.380d)}\\
&= \lim_{\epsilon \to 0} (2x + \epsilon) && \text{(12.380e)}\\
&= 2x, && \text{(12.380f)}
\end{aligned}$$

ce qu'il fallait prouver. □

12.15.1.2 La fonction racine carré

Considérons maintenant la fonction $f(x) = \sqrt{x}$. Nous avons

$$\begin{aligned}
\frac{f(x) - f(a)}{x - a} &= \frac{\sqrt{x} - \sqrt{a}}{x - a}\\
&= \frac{(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a})}{(x - a)(\sqrt{x} + \sqrt{a})}\\
&= \frac{1}{\sqrt{x} + \sqrt{a}}.
\end{aligned} \tag{12.381}$$

Lorsque $x \to a$, nous obtenons

$$f'(a) = \frac{1}{2\sqrt{a}}. \tag{12.382}$$

Notons que la dérivée de $f(x) = \sqrt{x}$ n'existe pas en $x = 0$. En effet elle serait donnée par le quotient

$$f'(0) = \lim_{x\to 0} \frac{\sqrt{x} - \sqrt{0}}{x} = \lim_{x\to 0} \frac{\sqrt{x}}{x} = \lim_{x\to 0} \frac{1}{\sqrt{x}}. \tag{12.383}$$

Mais si x devient très petit, la dernière fraction tend vers l'infini.

12.15.2 Interprétation géométrique de la dérivée : tangente

Considérons le **graphe** de la fonction f sur I, c'est-à-dire l'ensemble

$$\big\{\big(x, f(x)\big) \text{ tel que } x \in I\big\}. \tag{12.384}$$

Le nombre

$$\frac{f(x) - f(a)}{x - a} \tag{12.385}$$

est la pente de la droite qui joint les points $\big(x, f(x)\big)$ et $\big(a, f(a)\big)$, voir la figure 12.5.

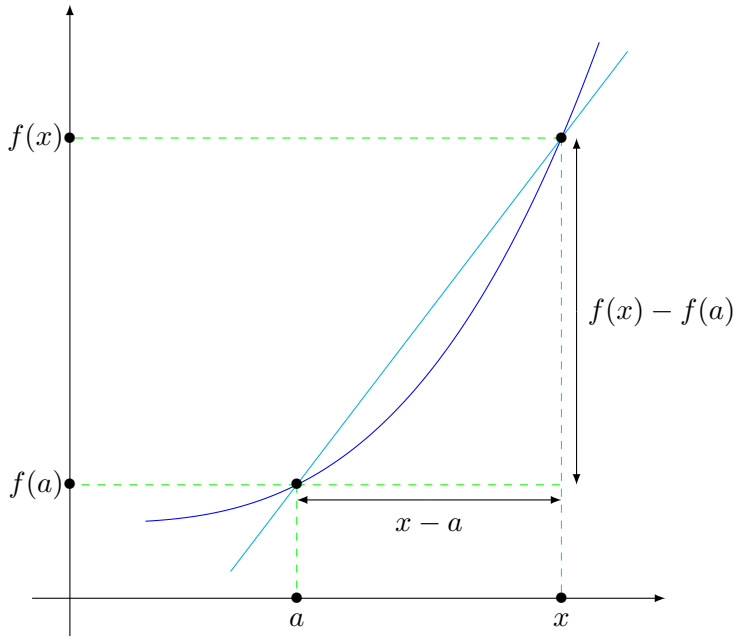


FIGURE 12.5: Le coefficient directeur de la corde entre a et x.LabelFigGWOYooRxHKSm

Étant donné que $f'(a)$ est le coefficient directeur de la tangente au point $\big(a, f(a)\big)$, l'équation de la tangente est

$$y - f(a) = f'(a)(x - a). \tag{12.386}$$EqTgfaen

12.15.3 Interprétation géométrique de la dérivée : approximation affine

Le fait que la fonction f soit dérivable au point $a \in I$ signifie que

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \ell \tag{12.387}$$

pour un certain nombre ℓ. Cela peut être réécrit sous la forme

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} - \ell = 0, \tag{12.388}$$

ou encore

$$\lim_{x\to a}\frac{f(x)-f(a)-\ell(x-a)}{x-a}=0. \tag{12.389}$$

Introduisons la fonction

$$\alpha(t)=\frac{f(a+t)-f(a)-t\ell}{t}. \tag{12.390}$$

Cette fonction est faite exprès pour que

$$\alpha(x-a)=\frac{f(x)-f(a)-\ell(x-a)}{x-a}; \tag{12.391}$$

par conséquent $\lim_{x\to a}\alpha(x-a)=0$. Nous récrivons l'équation (12.391) sous la forme

$$f(x)-f(a)-\ell(x-a)=(x-a)\alpha(x-a). \tag{12.392}$$

Le second membre tend vers zéro lorsque x tend vers a avec une « vitesse au carré » : c'est le produit de deux facteurs tous deux tendant vers zéro. Si x n'est pas très loin de a, il n'est donc pas une mauvaise approximation de dire

$$f(x)-f(a)-\ell(x-a)\simeq 0, \tag{12.393}$$

c'est-à-dire

$$f(x)\simeq f(a)+f'(a)(x-a). \tag{12.394}$$

Nous avons retrouvé l'équation (12.386). La manipulation que nous venons de faire revient donc à dire que la fonction f, au voisinage de a, est bien approximée par sa tangente.

L'équation (12.394) peut être aussi écrite sous la forme

$$f(x+\Delta x)\simeq f(x)+f'(x)\Delta x \tag{12.395}$$

qui est une approximation d'autant meilleure que Δx est petit.

12.16 Règles de calcul

D'abord une dérivée facile, qui sera utile pour démontrer la formule de dérivation d'un quotient.

Lemme 12.171.
Nous avons :

$$\left(\frac{1}{x}\right)'=-\frac{1}{x^2}. \tag{12.396}$$

Démonstration. En posant $f(x)=1/x$, nous avons le calcul

$$\frac{f(x+\epsilon)-f(x)}{\epsilon}=\frac{\frac{1}{x+\epsilon}-\frac{1}{x}}{\epsilon}=\frac{x-(x+\epsilon)}{\epsilon x(x+\epsilon)}=\frac{-1}{x(x+\epsilon)}. \tag{12.397}$$

Nous trouvons le résultat en passant à la limite et en tenant compte de la proposition 12.9 sur la limite d'un quotient. □

Proposition 12.172 ([397, 398, 399])**.**
Nous avons les règles suivantes.

(1) Si $f,g\colon\mathbb{R}\to\mathbb{R}$ sont dérivables en $a\in\mathbb{R}$, alors $f+g$ est dérivable en a et

$$(f+g)'(a)=f'(a)+g'(a). \tag{12.398}$$

(2) Si $f\colon\mathbb{R}\to\mathbb{R}$ est dérivable en $a\in\mathbb{R}$ et si $\lambda\in\mathbb{R}$, alors (λf) est dérivable en a et

$$(\lambda f)'(a)=\lambda f'(a). \tag{12.399}$$

ITEMooMQERooBCqnvS

(3) *Si $f, g\colon \mathbb{R} \to \mathbb{R}$ sont dérivables en $a \in \mathbb{R}$, alors fg est dérivable en a et*

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a). \tag{12.400}$$

Cette formule est appelée ***règle de Leibniz***[94].

ITEMooLYZCooVUPTyh

(4) *Soient deux intervalles I, J dans $\mathbb{R}$. Soient des fonctions $f\colon I \to J$ et $g\colon J \to \mathbb{R}$. Soit encore $a \in I$. Si f est dérivable en a et si g est dérivable en $f(a)$, alors $g \circ f$ est dérivable en a et*

$$(g \circ f)'(a) = g'\big(f(a)\big)f'(a). \tag{12.401}$$

ITEMooMUNQooLiKffz

(5) *Soient $f, g\colon I \to \mathbb{R}$ des fonctions sur un intervalle ouvert I. Soit $a \in I$; supposons que $g(a) \neq 0$. Alors la fonction $\frac{f}{g}$ est dérivable en a et*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \tag{12.402}$$

En particulier, la dérivation est une opération linéaire sur l'espace des fonctions infiniment dérivables.

Démonstration. Point par point.

(1) **Pour (1)** Soit $\epsilon > 0$. Nous avons

$$\frac{(f+g)(a+\epsilon) - (f+g)(a)}{\epsilon} = \frac{f(a+\epsilon) - f(a)}{\epsilon} + \frac{g(a+\epsilon) - g(a)}{\epsilon}. \tag{12.403}$$

Par hypothèse, les deux termes de droite ont une limite lorsque $\epsilon \to 0$. Donc le membre de gauche a une limite qui vaut la somme des deux limites[95], c'est-à-dire $f'(a) + g'(a)$.

(2) **Pour (2)** Écrivons la définition de la dérivée avec (λf) au lieu de f, et calculons un petit peu :

$$\begin{aligned}
(\lambda f)'(x) &= \lim_{\epsilon\to 0} \frac{(\lambda f)(x+\epsilon) - (\lambda f)(x)}{\epsilon} && \text{(12.404a)}\\
&= \lim_{\epsilon\to 0} \frac{\lambda\big(f(x+\epsilon)\big) - \lambda f(x)}{\epsilon} && \text{(12.404b)}\\
&= \lim_{\epsilon\to 0} \lambda \frac{f(x+\epsilon) - f(x)}{\epsilon} && \text{(12.404c)}\\
&= \lambda \lim_{\epsilon\to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} && \text{(12.404d)}\\
&= \lambda f'(x). && \text{(12.404e)}
\end{aligned}$$

(3) **Pour (3), règle de Leibniz** La définition de la dérivée dit que

$$(fg)'(x) = \lim_{\epsilon\to 0} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x)}{\epsilon}. \tag{12.405}$$

Eqfgrimeepsfgx

La subtilité est d'ajouter au numérateur la quantité $-f(x)g(x+\epsilon) + f(x)g(x+\epsilon)$, ce qui est permis parce que cette quantité est nulle[96]. Le numérateur de (12.405) devient donc

$$\begin{aligned}
&f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon) + f(x)g(x+\epsilon) - f(x)g(x)\\
&\qquad = g(x+\epsilon)\big(f(x+\epsilon) - f(x)\big) + f(x)\big(g(x+\epsilon) - g(x)\big),
\end{aligned} \tag{12.406}$$

94. Pour des formes plus générales, voir la proposition 30.11 et le lemme 30.12.
95. Limite de sommes, proposition 12.6(1).
96. Nous avons déjà fait le coup d'ajouter et enlever la même chose durant la démonstration du théorème 12.12. C'est une technique assez courante en analyse.

où nous avons effectué deux mises en évidence. Étant donné que nous avons deux termes, nous pouvons couper la limite en deux :

$$\begin{aligned}(fg)'(x) &= \lim_{\epsilon\to 0} g(x+\epsilon)\frac{f(x+\epsilon)-f(x)}{\epsilon} + \lim_{\epsilon\to 0} f(x)\frac{g(x+\epsilon)-g(x)}{\epsilon}\\ &= \lim_{\epsilon\to 0} g(x+\epsilon)\lim_{\epsilon\to 0}\frac{f(x+\epsilon)-f(x)}{\epsilon} + f(x)\lim_{\epsilon\to 0}\frac{g(x+\epsilon)-g(x)}{\epsilon},\end{aligned} \tag{12.407}$$

où nous avons utilisé le théorème 12.12 pour scinder la première limite en deux, ainsi que la proposition 12.6(2) pour sortir le $f(x)$ de la limite dans le second terme. Maintenant, dans le premier terme, nous avons évidemment [97] $\lim_{\epsilon\to 0} g(x+\epsilon) = g(x)$. Les limites qui restent sont les définitions classiques des dérivées de f et g au point x :

$$(fg)'(x) = g(x)f'(x) + f(x)g'(x), \tag{12.408}$$

ce qu'il fallait démontrer.

(4) **Pour (4)** Nous posons $b = f(a)$ et nous considérons la fonction suivante :

$$\begin{aligned}u\colon J &\to \mathbb{R}\\ y \mapsto u(y) &= \begin{cases}\dfrac{g(y)-g(b)}{y-b} & \text{si } y\neq b\\ g'(b) & \text{si } y=b.\end{cases}\end{aligned} \tag{12.409}$$

Vu que g est dérivable en b, la seconde ligne existe et u est continue en $y = b = f(a)$. C'est la définition de la dérivée.

Mais f est continue en a, donc $u\circ f$ est également continue en a, et nous avons

$$\lim_{x\to a}(u\circ f)(x) = u\big(f(a)\big) = u(b) = g'(b). \tag{12.410}$$

En réécrivant la définition de u en $f(x)$, l'expression suivante est une fonction continue de x :

$$u\big(f(x)\big) = \begin{cases}\dfrac{g\big(f(x)\big)-g(b)}{f(x)-b} & \text{si } f(x)\neq b\\ g'(b) & \text{si } y=b.\end{cases} \tag{12.411}$$

Si $f(x)\neq b$ nous avons :

$$g\big(f(x)\big) - g(b) = u\big(f(x)\big)\big(f(x)-b\big). \tag{12.412}$$

EQooKHQZooJdbmlT

Si par contre $f(x) = b$, en réalité, l'égalité (12.412) est encore valable parce qu'elle se résume à $0 = 0$. Nous divisons par $x-a$ et nous avons l'égalité

$$\frac{g\big(f(x)\big)-g\big(f(a)\big)}{x-a} = u\big(f(x)\big)\frac{f(x)-f(a)}{x-a} \tag{12.413}$$

qui est valable sur $I\backslash\{a\}$.

Il ne s'agit pas maintenant de prendre la limite $x\to a$ des deux côtés, parce que la limite du membre de gauche est précisément ce que ce théorème s'efforce de prouver exister. Nous montrons que la limite du membre de gauche existe en montrant que celle de droite existe.

D'une part, $u\circ f$ est continue et

$$\lim_{x\to a} u\big(f(x)\big) = u\big(f(a)\big) = u(b) = g'(b). \tag{12.414}$$

D'autre part, f est dérivable en a, donc

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = f'(a). \tag{12.415}$$

97. Pas tout à fait évidemment : selon le théorème 12.53, *limite et continuité*, il faut que g soit continue.

Tout cela pour dire qu'à droite, la limite existe et vaut $g'(b)f'(a)$. Donc nous avons l'existence de la limite que nous définissons $(g\circ f)'(a)$, et la valeur

$$\lim_{x\to a}\frac{g\big(f(x)\big)-g\big(f(a)\big)}{x-a}=g'\big(f(a)\big)f'(a). \tag{12.416}$$

Le résultat est prouvé.

(5) **Pour (5)** Nous considérons la fonction

$$\begin{aligned} i\colon \mathbb{R}\backslash\{0\}&\to\mathbb{R}\\ x&\mapsto\frac{1}{x}. \end{aligned} \tag{12.417}$$

La fonction g est dérivable en a, la fonction i est dérivable en $g(a)$. Donc par le théorème de dérivation des fonctions composées [98], la fonction $i\circ g$ est dérivable en a et

$$(i\circ g)'(a)=i'\big(g(a)\big)g'(a)=-\frac{g'(a)}{g(a)^2}. \tag{12.418}$$

Pour le quotient, nous utilisons la formule de la dérivée du produit sur $\frac{f}{g}(x)=f(x)\frac{1}{g(x)}$ pour dire que f/g est dérivable en a et

$$\left(\frac{f}{g}\right)'(a)=f'(a)\frac{1}{g(a)}+f(a)\left(\frac{1}{g}\right)'(a)=\frac{f'(a)}{g(a)}-\frac{f(a)g'(a)}{g(a)^2}=\frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}, \tag{12.419}$$

ce qu'il fallait démontrer.

□

Remarque 12.173.
Nous ne pouvons pas dire que la dérivée est une opération linéaire sur l'espace des fonctions dérivables. Certes la proposition 12.172 implique entre autres que l'ensemble des fonctions dérivables est un espace vectoriel. Mais la dérivée d'une fonction dérivable n'est pas spécialement dérivable.

Remarque 12.174.
La formule $(1/u)'=-u'/u^2$ ne peut pas être vue comme un cas particulier de $(u^\alpha)'=\alpha u^{\alpha-1}$ (proposition 12.435) parce que cette formule est utilisée dans la démonstration de la formule générale.

Pour les fonctions à valeurs dans $\mathbb{R}^n$, nous posons la définition suivante.

DEFooNEPUooWOMpyu

Définition 12.175.
Soit une fonction $f\colon\mathbb{R}\to\mathbb{R}^n$ dont les composantes $f_i\colon\mathbb{R}\to\mathbb{R}$ sont dérivables. Nous définissons la fonction f' par

$$f'(x)=\sum_i f_i'(x)e_i, \tag{12.420}$$

c'est-à-dire une dérivation composante par composante.

Cette définition est celle pour une fonction $\mathbb{R}\to\mathbb{R}^n$, et elle est facile. Très différente est la situation d'une fonction $\mathbb{R}^n\to\mathbb{R}$ dans laquelle il faudra introduire la notion de différentielle [99].

LEMooXHVBooHYjXdq

Lemme 12.176.
Soit une fonction dérivable $f\colon\mathbb{R}\to\mathbb{R}$. Nous posons

$$\begin{aligned} g\colon\mathbb{R}&\to\mathbb{R}\\ t&\mapsto f(a+\lambda t) \end{aligned} \tag{12.421}$$

où $a\in\mathbb{R}$. Alors g est dérivable et $g'(0)=\lambda f'(a)$.

98. Proposition 12.172(4).
99. Ce sera pour la définition 10.243.

Démonstration. Nous devons prouver que la limite

$$\lim_{\epsilon\to 0}\frac{g(\epsilon)-g(0)}{\epsilon} \tag{12.422}$$

existe et vaut $\lambda f'(a)$. Nous y allons avec les accroissements finis 10.257 :

$$g(\epsilon)-g(0)=f(a+\lambda\epsilon)-f(a)=f(a)+\lambda\epsilon f'(a)-f(a)=\lambda\epsilon f'(a). \tag{12.423}$$

Le quotient différentiel devient donc

$$\frac{g(\epsilon)-g(0)}{\epsilon}=\frac{\lambda\epsilon f'(a)}{\epsilon}=\lambda f'(a). \tag{12.424}$$

Il n'y a donc pas de problème à passer à la limite et nous avons $g'(0)=f'(a)$. □

Par rapport à la dérivation, les produits scalaire et vectoriel vérifient une règle de Leibniz.

PROPooFKKHooQZGXhE

Proposition 12.177.
Soit I un intervalle de $\mathbb{R}$. Si u et v sont dans $C^1(I,\mathbb{R}^3)$, alors

$$\begin{aligned}\frac{d}{dt}\big(u(t)\cdot v(t)\big)&=\big(u'(t)\cdot v(t)\big)+\big(u(t)\cdot v'(t)\big)\\ \frac{d}{dt}\big(u(t)\times v(t)\big)&=\big(u'(t)\times v(t)\big)+\big(u(t)\times v'(t)\big).\end{aligned} \tag{12.425}$$

EqFormLeibProdscalVect

Démonstration. Nous considérons des fonctions dérivables $f,g\colon\mathbb{R}\to\mathbb{R}^3$, et nous posons $\varphi(t)=f(t)\cdot g(t)$. En ce qui concerne la dérivée de la fonction $f\cdot g\colon\mathbb{R}\to\mathbb{R}$, nous devons étudier la limite

$$\lim_{\epsilon\to 0}\frac{\varphi(t+\epsilon)-\varphi(t)}{\epsilon}=\lim_{\epsilon\to 0}\frac{f(t+\epsilon)\cdot g(t+\epsilon)-f(t)\cdot g(t)}{\epsilon}. \tag{12.426}$$

EQooGRFKooNHceiW

La fonction f étant dérivable, la proposition 10.257 nous donne une fonction $\alpha\colon\mathbb{R}\to\mathbb{R}^3$ telle que

$$f(t+\epsilon)=f(t)+\epsilon f'(t)+\epsilon\alpha(\epsilon) \tag{12.427}$$

et $\lim_{\epsilon\to 0}\alpha(\epsilon)=0$. En substituant cela dans le numérateur de (12.426) nous calculons un peu : |

SUBEQSooMNAZoc

$$\begin{aligned}f(t+\epsilon)\cdot g(t+\epsilon)-f(t)\cdot g(t)&=\big(f(t)+\epsilon f'(t)+\epsilon\alpha(\epsilon)\big)\cdot\big(g(t)+\epsilon g'(t)+\epsilon\beta(\epsilon)\big) && (12.428a)\\ &\quad-f(t)\cdot g(t) && (12.428b)\\ &=\epsilon f(t)\cdot g'(t)+\epsilon f(t)\cdot\beta(\epsilon) && (12.428c)\\ &\quad+\epsilon f'(t)\cdot g(t)+\epsilon^2 f'(t)\cdot g'(t)+\epsilon^2 f'(t)\cdot\beta(\epsilon) && (12.428d)\\ &\quad+\epsilon\alpha(\epsilon)\cdot g(t)+\alpha(\epsilon)\epsilon^2\cdot g'(t)+\epsilon^2\alpha(\epsilon)\cdot\beta(\epsilon). && (12.428e)\end{aligned}$$

En divisant cela par ϵ et en prenant la limite $\epsilon\to 0$, il nous reste

$$f(t)\cdot g'(t)+f'(t)\cdot g(t). \tag{12.429}$$

□

12.16.1 Dérivée de la réciproque

PropMRBooXnnDLq

Proposition 12.178 ([400]).
Soit $f\colon I\to J$ une fonction bijective, continue et dérivable[100]*. Soient $x_0\in I$ et $y_0=f(x_0)$. Si $f'(x_0)\neq 0$ alors la fonction réciproque f^{-1} est dérivable en y_0 et sa dérivée est donnée par*

$$(f^{-1})'(y_0)=\frac{1}{f'(x_0)}. \tag{12.430}$$

EQooXGJEooSRaQBe

100. Définition 10.245.

Démonstration. Pour rappel, une fonction dérivable est toujours continue (proposition 12.165).

Prouvons que f^{-1} est dérivable au point $b = f(a) \in J$. Étant donné que f est dérivable en a, nous avons

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}. \tag{12.431}$$

Par ailleurs, étant donnée la continuité de f^{-1} donnée par la proposition 12.51(4), nous avons

$$\lim_{\epsilon\to 0} f^{-1}(b + \epsilon) = f^{-1}(b) = a. \tag{12.432}$$

Nous pouvons donc remplacer dans (12.431) tous les x par $f^{-1}(b + \epsilon)$ et prendre la limite $\epsilon \to 0$ au lieu de $x \to a$:

$$\begin{aligned} f'(a) &= \lim_{\epsilon\to 0} \frac{f\big(f^{-1}(b+\epsilon)\big) - f(a)}{f^{-1}(b+\epsilon) - a} \\ &= \lim_{\epsilon\to 0} \frac{b + \epsilon - f(a)}{f^{-1}(b+\epsilon) - f^{-1}(b)} \\ &= \lim_{\epsilon\to 0} \frac{\epsilon}{f^{-1}(b+\epsilon) - f^{-1}(b)} \\ &= \frac{1}{\lim_{\epsilon\to 0} \frac{f^{-1}(b+\epsilon)-f^{-1}(b)}{\epsilon}} \\ &= \frac{1}{(f^{-1})'(b)}. \end{aligned} \tag{12.433}$$

Nous avons utilisé le fait que $f(a) = b$ et $a = f^{-1}(b)$. □

Nous cherchons à construire une application dérivable et d'inverse dérivable entre $\mathbb{R}$ (en entier) et un ouvert borné de $\mathbb{R}$. Il serait tentant de prendre l'application arc tangente

$$\begin{aligned} \arctan\colon \mathbb{R} &\to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\\ x &\mapsto \arctan(x) \end{aligned}, \tag{12.434}$$

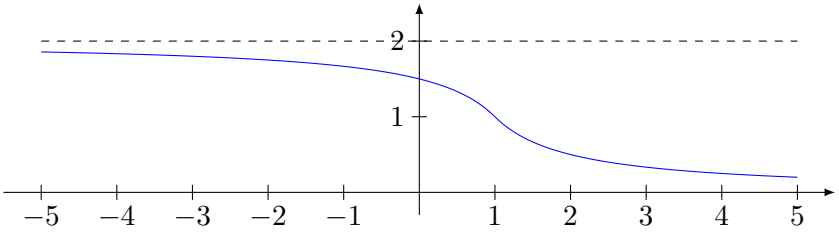
mais elle ne sera définie que dans le théorème 18.38.

Nous aimerions donner le logarithme comme exemple, mais l'exponentielle ne sera définie que dans longtemps, à partir des séries entières. Allez voir l'exemple 15.93 pour le logarithme comme réciproque de l'exponentielle.

Proposition 12.179 (difféomorphisme entre $\mathbb{R}$ et un ouvert borné)**.**
L'application

$$\begin{aligned} f\colon \mathbb{R} &\to \left]0, 2\right[\\ x &\mapsto \begin{cases} 2 + \frac{1}{x-2} & si\ x \leqslant 1 \\ \frac{1}{x} & si\ x > 1. \end{cases} \end{aligned} \tag{12.435}$$

est un C^1-difféomorphisme[101].



Démonstration. En plusieurs parties.

101. Définition 10.265.

(1) **Bijection** Nous voyons vite que cette fonction est strictement décroissante ; et un calcul de limite nous dit qu'il s'agit d'une bijection dérivable

$$f : \mathbb{R} \to]0, 2[. \tag{12.436}$$

La proposition 12.178 s'applique et la bijection réciproque est également dérivable (donc continue aussi).

(2) **Continue en $x = 1$** Il suffit de calculer les deux valeurs. Dans les deux cas nous obtenons $f(1) = 1$.

(3) **Dérivable en $x = 1$**
En ce qui concerne la dérivabilité en $x = 1$, nous devons étudier

$$\lim_{\epsilon \to 0} \frac{f(1+\epsilon) - f(1)}{\epsilon}. \tag{12.437}$$

La limite à gauche est égale à la dérivée de $x \mapsto 2 + \frac{1}{x-2}$ en $x = 1$ et la limite à droite est égale à la dérivée de $x \mapsto 1/x$ en $x = 1$. Dans les deux cas nous trouvons -1.

(4) **Dérivée continue** La proposition 12.178 dit que

$$(f^{-1})'(y) = \frac{1}{(f' \circ f^{-1})(y)}. \tag{12.438}$$ EQooNTHEooUQoMHT

Vu que f' et f^{-1} sont continues, $f' \circ f^{-1}$ est continue [102]. Et comme elle ne s'annule pas, la formule (12.438) donne une fonction continue [103].

□ PROPooFEJXooMjsvFo

Proposition 12.180 ([401]).
L'application

$$\begin{aligned} \psi : \mathbb{R}^n &\to B(0,1) \\ x &\mapsto \frac{x}{\sqrt{1+\|x\|^2}} \end{aligned} \tag{12.439}$$

est un C^∞-difféomorphisme [104].

Démonstration. En plusieurs parties.

(1) **Injective** Notons pour commencer que $\psi(x) = 0$ si et seulement si $x = 0$.
Soient $x, y \in \mathbb{R}^n$ tels que $\psi(x) = \psi(y)$. Nous avons

$$\frac{x}{\sqrt{1+\|x\|^2}} = \frac{y}{\sqrt{1+\|y\|^2}}, \tag{12.440}$$ EQooYGIWooBxSxyd

ce qui donne immédiatement que $y = \lambda x$. Nous devons prouver que $\lambda = 1$. Nous remplaçons y par λx dans le membre de droite de (12.440). En posant $r = \|x\|^2$ nous avons

$$\frac{\lambda}{\sqrt{1+\lambda^2 r}} = \frac{1}{\sqrt{1+r}}. \tag{12.441}$$

Nous passons au carré et nous résolvons par rapport à r : $(1+r)\lambda^2 = 1 + \lambda^2 r$, ce qui donne $\lambda^2 = 1$ et donc $\lambda = \pm 1$. Étant donné que

$$\frac{x}{\sqrt{1+r}} = \frac{\lambda x}{\sqrt{1+\lambda^2 r}}, \tag{12.442}$$

nous ne pouvons avoir $\lambda = -1$ que poir $x = 0$.

102. Proposition 9.89
103. Proposition 9.87.
104. Définition 10.265.

(2) **surjectif** D'abord $0 = \psi(0)$. Nous considérons $y \in B(0,1)$ et nous cherchons un $x \in \mathbb{R}^n$ tel que $\psi(x) = y$. Nous le cherchons sous la forme $x = \lambda y$ (où λ est une fonction de y). Nous avons à résoudre pour λ l'équation

$$\frac{\lambda^2}{1 + \lambda^2 \|y\|^2} = 1. \tag{12.443}$$

Cela nous mène à

$$\lambda = \frac{\pm 1}{\sqrt{1 - \|y\|^2}}, \tag{12.444}$$

et nous vérifions qu'en posant $\lambda = 1$, c'est bon. Notons d'abord qu'il n'y a aucun problèmes de zéros parce que $y \in B(0,1)$ donne toujours $1 - \|y\|^2 > 0$.

Maintenant, la vérification :

$$\psi\left(\frac{y}{\sqrt{1 - \|y\|^2}}\right) = \frac{y}{\sqrt{1 - \|y\|^2}} \frac{1}{\sqrt{1 + \frac{\|y\|^2}{1 - \|y\|^2}}}. \tag{12.445}$$

En travaillant un peu nous trouvons que le tout est égal à y.

(3) **ψ est C^∞** Ça se voit sur la formule parce que ce qui est dans la racine carrée est toujours strictement positif (et même borné par 1 vers le bas).

(4) **Calcul de la réciproque** Le calcul de la réciproque est déjà fait dans la partie sur la surjectivité. Nous notons juste le résultat :

$$\begin{aligned} \psi^{-1}\colon B(0,1) &\to \mathbb{R}^n \\ y &\mapsto \frac{y}{\sqrt{1 - \|y\|^2}}. \end{aligned} \tag{12.446}$$

EQooOTQZooBXiZGb

(5) **Inverse est C^∞** Cela se voit encore sur la formule (12.446), en gardant en tête que $y \in B(0,1)$.

□

12.16.2 Dérivée de fonction composée

PROPooDONLooWthqRR

Proposition 12.181 ([402]).
Soient des intervalles I et J dans $\mathbb{R}$ ainsi que des fonctions $f\colon I \to \mathbb{R}$ et $g\colon J \to \mathbb{R}$ tels que $g(J) \subset I$. Soit $a \in J$. Nous supposons que f est dérivable en $g(a)$ et g est dérivable en a.

Alors $f \circ g$ est dérivable en a et

$$(f \circ g)'(a) = f'\big(g(a)\big) g'(a). \tag{12.447}$$

Démonstration. Nous considérons la formule des accroissements finis sous la forme (10.717). Pour la fonction g, nous écrivons

$$g(a + \epsilon) = g(a) + \epsilon g'(a) + \epsilon\alpha(\epsilon) \tag{12.448}$$

avec $\alpha(\epsilon) \to 0$. Et de même pour $f\big(g(a)\big)$:

$$\begin{aligned} f\big(g(a + \epsilon)\big) &= f\big(g(a) + \epsilon g'(a) + \epsilon\alpha(\epsilon)\big) && \text{(12.449a)} \\ &= f\big(g(a)\big) + \big(\epsilon g'(a) + \epsilon\alpha(\epsilon)\big) f'\big(g(a)\big) + \epsilon\beta\big(\epsilon g'(a) + \epsilon\alpha(\epsilon)\big) && \text{(12.449b)} \end{aligned}$$

avec $\beta(\epsilon) \to 0$. Nous avons donc, pour ϵ assez petit pour que tout reste dans I et J [105], que

$$\frac{(f \circ g)(a + \epsilon) - (f \circ g)(a)}{\epsilon} = \big(g'(a) + \alpha(\epsilon)\big) f'\big(g(a)\big) + \beta\big(\epsilon g'(a) + \epsilon\alpha(\epsilon)\big). \tag{12.450}$$

En ce qui concerne la limite $\epsilon \to 0$, nous avons entre autres,

$$\lim_{\epsilon \to 0} \beta\big(\epsilon g'(a) + \epsilon\alpha(\epsilon)\big) = 0, \tag{12.451}$$

et donc bien $(f \circ g)'(a) = g'(a) f'\big(g(a)\big)$. □

105. Et c'est là qu'on utilise la continuité de f et g garantie par la proposition 12.165.

12.16.3 Dérivée de fonction périodique

DEFooHUZAooYyBmwe

Définition 12.182.
*Une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est **périodique** si il existe $T > 0 \in \mathbb{R}$ tel que*

$$f(x+T) = f(x) \tag{12.452}$$

*pour tout $x \in \mathbb{R}$. Un tel T est une **période** de f. Nous disons que T est* la *période de f si il est le minimum vérifiant la propriété.*

LEMooOGFGooCnTDjO

Lemme 12.183 ([1])**.**
Si f est une fonction périodique de période T, alors toutes les périodes sont de la forme kT avec $k \in \mathbb{N}^*$.

Démonstration. Considérons une période t. Nous avons $t > T$ par hypothèse. Si t n'est pas un multiple de T, la division euclidienne 1.222 permet d'écrire $t = kT + l$ avec $l < T$. Nous avons alors, pour tout $x \in \mathbb{R}$:

$$f(x) = f(x+t) = f(x+kT+l) = f(x+l). \tag{12.453}$$

Donc l est une période de f. Cela n'est pas possible parce que T est la plus petite. □

LEMooHWQYooXcNLts

Lemme 12.184.
Si $f\colon \mathbb{R} \to \mathbb{R}$ est périodique et si T est une période de f, alors f' est périodique et T en est une période.

Démonstration. Soit $a \in \mathbb{R}$. Par hypothèse f est dérivable en a et les limites qui suivent existent :

$$f'(a+T) = \lim_{\epsilon \to 0} \frac{f(a+T+\epsilon) - f(a+T)}{\epsilon} = \lim_{\epsilon \to 0} \frac{f(a+\epsilon) - f(a)}{\epsilon} = f'(a). \tag{12.454}$$

Nous avons utilisé la condition de périodicité en a et en $a+\epsilon$. □

12.17 Dérivation et croissance

Supposons une fonction dont la dérivée est positive. Étant donné que la courbe est « collée » à ses tangentes, tant que les tangentes montent, la fonction monte. Or, une tangente qui monte correspond à une dérivée positive, parce que la dérivée est le coefficient directeur de la tangente.

Ce résultat très intuitif peut être prouvé rigoureusement. C'est la tache à laquelle nous allons nous atteler maintenant.

PropGFkZMwD

Proposition 12.185.
Si f et f' sont des fonctions continues sur l'intervalle $[a,b]$ et si f' est strictement positive sur $[a,b]$, alors f est strictement croissante sur $[a,b]$.

De la même manière, si f' est strictement négative sur $[a,b]$, alors f est strictement décroissante sur $[a,b]$.

Démonstration. Nous n'allons prouver que la première partie. La seconde partie se prouve en considérant $-f$ et en invoquant alors la première [106]. Prenons x_1 et x_2 dans $[a,b]$ tels que $x_1 < x_2$. Par hypothèse, pour tout x dans $[x_1, x_2]$, nous avons

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} > 0. \tag{12.455}$$

Maintenant, la proposition 12.11 dit que quand une limite est positive, alors la fonction dans la limite est positive sur un voisinage. En appliquant cette proposition à la fonction

$$r(\epsilon) = \frac{f(x+\epsilon) - f(x)}{\epsilon}, \tag{12.456}$$

106. Méditer cela.

dont la limite en zéro est positive, nous trouvons que $r(\epsilon) > 0$ pour tout ϵ pas trop éloigné de zéro. En particulier, il existe un $\delta > 0$ tel que $\epsilon < \delta$ implique $r(\epsilon) > 0$; pour un tel ϵ, nous avons donc

$$r(\epsilon) = \frac{f(x+\epsilon) - f(x)}{\epsilon} > 0. \tag{12.457}$$

Étant donné que $\epsilon > 0$, nous avons que $f(x+\epsilon) - f(x) > 0$, c'est-à-dire que f est strictement croissante entre x et $x+\epsilon$.

Jusqu'ici, nous avons prouvé que la fonction f était strictement croissante dans un voisinage autour de chaque point de $[a,b]$. Cela n'est cependant pas encore tout à fait suffisant pour conclure. Ce que nous voudrions faire, c'est de prendre un voisinage $]a, m_1[$ autour de a sur lequel f est croissante. Donc, $f(m_1) > f(a)$. Ensuite, on prend un voisinage $]m_1, m_2[$ de m_1 sur lequel f est croissante. De ce fait, $f(m_2) > f(m_1) > f(a)$. Et ainsi de suite, nous voulons construire des m_3, m_4,..., jusqu'à arriver en b. Hélas, rien ne dit que ce processus va fonctionner. Il faut trouver une subtilité. Le problème est que les voisinages sur lesquels la fonction est croissante sont peut-être de plus en plus petits, de telle sorte à ce qu'il faille une infinité d'étapes avant d'arriver à bon port (en b).

Heureusement, nous pouvons drastiquement réduire le nombre d'étapes en nous souvenant du théorème de Borel-Lebesgue 9.19. Nous notons par $\mathcal{O}_x$, un ouvert autour de x tel que f soit strictement croissante sur $\mathcal{O}_x$. Un tel voisinage existe. Cela fait une infinité d'ouverts tels que

$$[a,b] \subseteq \bigcup_{x\in[a,b]} \mathcal{O}_x. \tag{12.458}$$

Ce que le théorème dit, c'est qu'on peut en choisir un nombre fini qui recouvre encore $[a,b]$. Soient $\{\mathcal{O}_{x_1}, \dots, \mathcal{O}_{x_n}\}$, les heureux élus, que nous supposons pris dans l'ordre : $x_1 < x_2 < \dots < x_n$. Nous avons

$$[a,b] \subseteq \bigcup_{i=1}^{n} \mathcal{O}_i. \tag{12.459}$$

Quitte à les rajouter à la collection, nous supposons que $x_1 = a$ et que $x_n = b$. Maintenant nous allons choisir encore un sous-ensemble de cette collection d'ouverts. On pose $\mathcal{A}_1 = \mathcal{O}_{x_1}$. Nous savons que $\mathcal{A}_1$ intersecte au moins un des autres $\mathcal{O}_{x_i}$. Cette affirmation vient du fait que $[a,b]$ est connexe (proposition 9.53), et que si $\mathcal{O}_{x_1}$ n'intersectait personne, alors

$$\mathcal{O}_{x_1} \quad \text{et} \quad \bigcup_{i=2}^{n} \mathcal{O}_{x_i} \tag{12.460}$$

forment une partition de $[a,b]$ en deux ouverts disjoints, ce qui n'est pas possible parce que $[a,b]$ est connexe. Nous nommons $\mathcal{A}_2$, un des ouverts $\mathcal{O}_{x_i}$ qui intersecte $\mathcal{A}_1$. Disons que c'est $\mathcal{O}_k$. Notons que $\mathcal{A}_1 \cup \mathcal{A}_2$ est un intervalle sur lequel f est strictement croissante. En effet, si y_{12} est dans l'intersection, $f(a) < f(y_{12})$ parce que f est strictement croissante sur $\mathcal{A}_1$, et pour tout $x > y_{12}$ dans $\mathcal{A}_2$, $f(x) > f(y_{12})$ parce que f est strictement croissante dans $\mathcal{A}_2$.

Maintenant, nous éliminons de la liste des $\mathcal{O}_{x_i}$ tous ceux qui sont inclus dans $\mathcal{A}_1 \cup \mathcal{A}_2$. Dans ce qu'il reste, il y en a automatiquement un qui intersecte $\mathcal{A}_1 \cup \mathcal{A}_2$, pour la même raison de connexité que celle invoquée plus haut. Nous appelons cet ouvert $\mathcal{A}_3$, et pour la même raison qu'avant, f est strictement croissante sur $\mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$.

En recommençant suffisamment de fois, nous finissons par devoir prendre un des $\mathcal{O}_{x_i}$ qui contient b, parce qu'au moins un des $\mathcal{O}_{x_i}$ contient b. À ce moment, nous avons fini la démonstration. □

PROPooKZPZooWjIsWg

Proposition 12.186 (Dérivée et croissance)**.**
Soit un intervalle $I \subset \mathbb{R}$. Une fonction dérivable $f : I \to \mathbb{R}$ est

(1) Si $f' > 0$ sur I, alors f est strictement croissante sur I.

ITEMooKMDXooPBnVoi

(2) Si $f' \geqslant 0$ sur I, alors f est croissante sur I.

(3) *Si $f' < 0$ sur I, alors f est strictement décroissante sur I.*

(4) *Si $f' \leqslant 0$ sur I, alors f est décroissante sur I.*

PROPooSGTBooFxUuXK

Proposition 12.187 ([403]).
Soient des intervalles I et J dans $\mathbb{R}$. Soit $f\colon I \to J$ une fonction dérivable et strictement monotone. Si f' ne s'annule pas sur I alors

ITEMooFXHYooRNHYPI

(1) *la fonction f est injective de I vers J,*

ITEMooWPPIooDUAYsH

(2) *la fonction f^{-1} est dérivable sur J,*

ITEMooTVMLooQgjLEB

(3) *et nous avons la formule*

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}. \tag{12.461}$$

EQooELIHooDxUExH

Démonstration. Nous supposons que f est strictement croissante sur I. Les adaptations à faire en cas de fonction strictement décroissante sont laissées à la lectrice. En plusieurs points.

(1) **Pour (1)** Une fonction strictement croissante est injective.

(2) **Pour (2)** Vu que f' ne s'annule pas, la la proposition 12.178 dit que f^{-1} est dérivable.

(3) **Pour (3)** Formule 12.430.

□

12.188.
Très souvent on préfère retenir la formule

$$(f^{-1})'(y_0) = \frac{1}{f'\big((f^{-1})(y_0)\big)} \tag{12.462}$$

EqWWAooBRFNsv

Elle est très simple à retrouver : il suffit d'écrire

$$f^{-1}\big(f(x)\big) = x \tag{12.463}$$

puis de dériver les deux côtés par rapport à x en utilisant la règle de dérivation des fonctions composées de la proposition 12.181 :

$$(f^{-1})'\big(f(x)\big)f'(x) = 1. \tag{12.464}$$

12.17.1 Théorèmes de Rolle et des accroissements finis

ThoRolle

Théorème 12.189 (Théorème de Rolle[404, 405]).
Soit f, une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$. Si $f(a) = f(b)$, alors il existe un point $c \in]a,b[$ tel que $f'(c) = 0$.

Démonstration. Étant donné que $[a,b]$ est un intervalle compact, l'image de $[a,b]$ par f est un intervalle compact, soit $[m,M]$ (théorème 7.220). Si $m = M$, alors le théorème est évident : c'est que la fonction est constante, et la dérivée est par conséquent nulle. Supposons que $M > f(a)$ (il se peut que $M = f(a)$, mais alors si f n'est pas constante, il faut avoir $m < f(a)$ et le reste de la preuve peut être adaptée).

Comme M est dans l'image de $[a,b]$ par f, il existe $c \in]a,b[$ tel que $f(c) = M$. Considérons maintenant la fonction

$$\tau(x) = \frac{f(c+x) - f(c)}{x}. \tag{12.465}$$

Par définition, $\lim_{x\to 0} \tau(x) = f'(c)$. Par hypothèse, si $u < c$,

$$\tau(u-c) = \frac{f(u) - f(c)}{u - c} > 0 \tag{12.466}$$

parce que $u-c<0$ et $f(u)-f(c)<0$. Par conséquent, $\lim_{x\to 0}\tau(x)\geqslant 0$. Nous avons aussi, pour $v>c$,

$$\tau(v-c)=\frac{f(v)-f(c)}{v-c}<0 \tag{12.467}$$

parce que $v-c>0$ et $f(v)-f(c)<0$. Par conséquent, $\lim_{x\to 0}\tau(x)\leqslant 0$. Mettant les deux ensemble, nous avons $f'(c)=\lim_{x\to 0}\tau(x)=0$, et c est le point que nous cherchions. □

Voici une généralisation du théorème de Rolle, dans le cas où nous n'aurions pas deux points sur lesquels la fonction est identique, mais deux points en lesquels la limite de la fonction est identique. Typiquement, lorsque les points en question sont $\pm\infty$.

THOooXDTBooFeSZoK

Théorème 12.190 (Généralisation du théorème de Rolle[404])**.**
Soient $-\infty\leqslant a<b\leqslant+\infty$. Soit une fonction dérivable $f\colon]a,b[\to\mathbb{R}$ telle que

$$\lim_{x\to a}f(x)=\lim_{x\to b}f(x)=\ell \tag{12.468}$$

avec $\ell\in\bar{\mathbb{R}}$. Alors il existe $x\in]a,b[$ tel que $f'(x)=0$.

Démonstration. Soit un difféomorphisme[107] strictement croissant $\varphi\colon\mathbb{R}\to]\alpha,\beta[$. Pour cela vous pouvez bricoler quelque chose à partie de la proposition 12.179. Mais n'utilisez pas la fonction arc tangente, parce qu'elle n'est définie qu'au théorème 18.38.

Nous posons $\tilde{a}=\varphi(a)$, $\tilde{b}=\varphi(b)$ et

$$g=\varphi\circ f\circ\varphi^{-1}\colon]\tilde{a},\tilde{b}[\to]\alpha,\beta[. \tag{12.469}$$

Cela est une fonction dérivable et continue sur $[\tilde{a},\tilde{b}]$ en posant $g(\tilde{a})=g(\tilde{b})=\varphi(\ell)$.

Donc il existe $\tilde{c}\in]\tilde{a},\tilde{b}[$ tel que $g'(\tilde{c})=0$. En posant $c=\varphi^{-1}(\tilde{c})$ nous avons $c\in]a,b[$ et, en utilisant de nombreuses fois la règle de dérivation des fonctions composées 12.172(4),

$$\begin{aligned}
f'(c)&=f'\big(\varphi^{-1}(\tilde{c})\big) && \text{(12.470a)}\\
&=(\varphi^{-1})'\Big(\big(g\circ\varphi\big)\big(\varphi^{-1}(\tilde{c})\big)\Big)(g\circ\varphi)'\big(\varphi^{-1}(\tilde{c})\big) && \text{(12.470b)}\\
&=(\varphi^{-1})'\big(g(\tilde{c})\big)\underbrace{g'(\tilde{c})}_{=0}\varphi'\big(\varphi^{-1}(\tilde{c})\big) && \text{(12.470c)}\\
&=0. && \text{(12.470d)}
\end{aligned}$$

□

Une autre généralisation du théorème de Rolle, avec des dérivées d'ordre supérieur :

PROPooCPCAooJjOZNy

Proposition 12.191 ([406])**.**
Soit un intervalle ouvert $I\subset\mathbb{R}$ contenant a,b $(a\neq b)$. Soit une fonction $f\in C^{n+1}(I,\mathbb{R})$. Si $f(a)=f(b)$ et si $f^{(j)}(a)=0$ pour $j=1,\dots,n$, alors il existe $c\in]a,b[$ tel que $f^{(n+1)}(c)=0$.

Démonstration. Le théorème de Rolle 12.189 nous dit qu'il existe $c_1\in]a,b[$ tel que $f'(c_1)=0$. Mais alors $f'(a)=f'(c_1)=0$, et le théorème de Rolle appliqué à f' donne $c_2\in]a,c_1[$ tel que $f''(c_2)=0$. Continuant ainsi n fois, il existe $c\in]a,b[$ tel que $f^{(n+1)}(c)=0$. □

Le théorème suivant est le théorème des **accroissements finis**. Une version avec des dérivées partielles sera la proposition 12.236.

ThoAccFinis

Théorème 12.192 (Accroissements finis)**.**
Soit f, une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$.

107. Définition 10.265.

ITEMooFZONooXJqLyX

(1) *Il existe au moins un réel* $c \in]a,b[$ *tel que*

$$f'(c) = \frac{f(b)-f(a)}{b-a} \tag{12.471}$$

Autrement dit, la tangente en c *est parallèle à la corde entre* a *et* b.

ITEMooXRQKooDBFpdQ

(2) *Nous avons la majoration*

$$\left|\frac{f(b)-f(a)}{b-a}\right| \leqslant \sup_{x\in]a,b[} |f'(x)||b-a|. \tag{12.472}$$

Démonstration. Considérons la fonction

$$\tau(x) = f(x) - \Big(\frac{f(b)-f(a)}{b-a}x + f(a) - a\frac{f(b)-f(a)}{b-a}\Big), \tag{12.473}$$

c'est-à-dire la fonction qui donne la distance entre f et le segment de droite qui lie $(a,f(a))$ à $(b,f(b))$. Par construction, $\tau(a)-\tau(b)=0$, donc le théorème de Rolle s'applique à τ pour laquelle il existe donc un $c\in]a,b[$ tel que $\tau'(c)=0$.

En utilisant les règles de dérivation, nous trouvons que la dérivée de τ vaut

$$\tau'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}, \tag{12.474}$$

donc dire que $\tau'(c)=0$ revient à dire que $f(b)-f(a)=(b-a)f'(c)$, ce qu'il fallait démontrer.

La majoration est une conséquence immédiate, parce que le supremum de $|f'(x)|$ est forcément plus grand que $|f'(c)|$. □

Une généralisation pour une fonction sur un intervalle $]a,b[$ où a et b peuvent être infinis.

THOooRIIBooOjkzMa

Théorème 12.193 (Généralisation des accroissements finis)**.**
Soient $-\infty \leqslant a < b \leqslant +\infty$ *et* f,g *des fonctions continues sur* $[a,b]$ *et dérivables sur* $]a,b[$.

Si $a=-\infty$:

— *Nous demandons la continuité sur* $]-\infty,b]$ *et la dérivabilité sur* $]-\infty,b[$.

— *Nous notons* $f(a)$ *la limite* $\lim_{x\to-\infty} f(x)$*, et nous supposons qu'elle est finie.*

Mêmes conventions si $b=+\infty$.

Alors il existe $c\in]a,b[$ *tel que*

$$\big(f(b)-f(a)\big)g'(c) = \big(g(b)-g(a)\big)f'(c). \tag{12.475}$$

Démonstration. Nous posons

$$h(t) = \big(g(b)-g(a)\big)f(t) - \big(f(b)-f(a)\big)g(t). \tag{12.476}$$

Nous avons $\lim_{t\to a} h(t) = \lim_{t\to b} h(t)$, de telle sorte que le théorème de Rolle généralisé 12.190 s'applique et il existe $c\in]a,b[$ tel que $h'(c)=0$. Pour ce c nous avons

$$0 = h'(c) = \big(f(b)-f(a)\big)g'(c) - \big(g(b)-g(a)\big)f'(c), \tag{12.477}$$

et donc

$$\big(f(b)-f(a)\big)g'(c) = \big(g(b)-g(a)\big)f'(c). \tag{12.478}$$

□

12.17.2 Règle de l'Hospital

PROPooBZHTooHmyGsy

Proposition 12.194 (Règle de l'Hospital pour $\frac{0}{0}$[407]).
Soient des fonctions f, g dérivables sur $]a,b[$ et dont la limite en a est nulle. Si g' ne s'annule pas sur $]a,b[$ et si

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell \tag{12.479}$$

alors

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell. \tag{12.480}$$

EQooJHWYooLGdbPH

Ici $\ell \in \bar{\mathbb{R}}$, et les hypothèses garantissent l'existence de la limite (12.480).

Démonstration. Soit $x \in]a,b[$. Les fonctions f et g sont dérivables sur $]a,x[$ et continues sur $[a,x]$, de telle sorte que le théorème 12.193 s'applique et nous avons $c_x \in]a,x[$ tel que

$$\big(f(x)-f(a)\big)g'(c_x) = \big(g(x)-g(a)\big)f'(c_x). \tag{12.481}$$

EQooMALUooNagayh

Nous nous souvenons de ce que signifient les notations dans le théorème : les notations $f(a)$, $f(x)$, $g(a)$ et $g(x)$ désignent en réalité les limites. Donc dans (12.481), nous avons $f(a) = g(a) = 0$.

D'autre part nous avons $g(x) \neq g(a)$, sinon le théorème de Rolle 12.190 annulerait g' quelque part dans $]a,x[$. Nous pouvons donc réécrire (12.481) sous la forme

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}. \tag{12.482}$$

EQooUCLVooFgAfwC

Mais $\lim_{x\to a^+} c_x = a$ parce que $c_x \in]a,x[$. Donc la limite du membre de droite de (12.482) lorsque $x \to a^+$ existe et vaut ℓ. La même limite à gauche doit alors exister et valoir la même valeur. □

PROPooTJVCooMeUhIy

Proposition 12.195 (L'Hospital pur $\frac{\infty}{\infty}$).
Soient f et g deux fonctions

(1) dérivables sur $]a,b[$,

(2) dont les limites en a sont toutes deux ∞,

(3) $g' \neq 0$ sur $]a,b[$.

(4)

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell \in \bar{\mathbb{R}}. \tag{12.483}$$

EQooVFYCooMjQGtJ

Alors

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell. \tag{12.484}$$

Cette dernière égalité signifie « la limite existe et vaut ℓ ».

Démonstration. Soit un intervalle $]x,y[$ strictement inclus dans $]a,b[$ avec $x,y \in \mathbb{R}$. Par le théorème des accroissements finis généralisés 12.193 il existe $c \in]x,y[$ tel que

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(c)}{g'(c)}. \tag{12.485}$$

Notons que le dénominateur à gauche n'est pas nul à cause du théorème de Rolle et de l'hypothèse que g' ne s'annule pas sur $[x,y]$. Nous isolons $f(x)$:

$$f(x) = \frac{f'(c)}{g'(c)}\Big(g(x)-g(y)\Big) + f(y). \tag{12.486}$$

EQooDFXNooJhdUca

Avant de diviser par $g(x)$ nous devons prendre quelques précautions. Soit V, un voisinage de ℓ[108]. Vu la limite (12.483), il existe $y \in]a,b[$ tel que

$$\frac{f'(t)}{g'(t)} \in V \tag{12.487}$$

pour tout $t \in]a,y[$. Nous utilisons ici avec subtilité le fait que ces intervalles sont une base de la topologie autour de ∞. Maintenant $f(y)$ et $g(y)$ sont fixés et sont des nombres réels. Vu que $\lim_{x\to a} g(x) = \infty$ nous pouvons choisir $r < y$ tel que nous ayons simultanément

(1) $g(x) \neq 0$ sur $]a,r[$,

(2)

$$\left|\frac{g(y)}{g(x)}\right| \leqslant \epsilon \tag{12.488}$$

et

$$\left|\frac{f(y)}{g(x)}\right| \leqslant \epsilon \tag{12.489}$$

pour tout $x \in]a,r[$.

Nous sommes maintenant armés de y et r satisfaisant tout cela et nous pouvons traiter avec la formule (12.486) en ne la considérant que pour $x \in]a,r[$. Soit $x \in]a,r[$; il existe $c_x \in]a,x[$ tel que

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}. \tag{12.490}$$ EQooNEZQooYGJmFW

Nous avons :

(1) $\lim_{x\to a^+} c_x = a$,

(2)

$$\lim_{x\to a^+} \frac{f'(c_x)}{g'(c_x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} = \ell, \tag{12.491}$$

(3)

$$\lim_{x\to a^+}\left(1 - \frac{g(y)}{g(x)}\right) = 1, \tag{12.492}$$

(4) $\lim_{x\to a^+} \frac{f(y)}{g(x)} = 0$.

Donc chaque partie du membre de droite de (12.490) a une limite bien déterminée pour $x \to a^+$. Les règles de calcul s'appliquent et nous avons

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell \times 1 + 0 = \ell. \tag{12.493}$$

□

12.17.3 Dérivée et primitive

COROOEOERooYprteX

Corolaire 12.196.
Soit f une fonction dérivable sur $[a,b]$ telle que $f'(x) = 0$ pour tout $x \in [a,b]$. Alors f est constante sur $[a,b]$.

Démonstration. Si f n'était pas constante sur $[a,b]$, il existerait un $x_1 \in]a,b[$ tel que $f(a) \neq f(x_1)$, et dans ce cas, il existerait, par le théorème des accroissements finis 12.192, un $c \in]a,x_1[$ tel que

$$f'(c) = \frac{f(x_1) - f(a)}{x_1 - a} \neq 0, \tag{12.494}$$

ce qui contredirait les hypothèses. □

108. Vous savez ce que signifie un « voisinage de ∞ » ? Allez voir la définition 12.24.

CorNErEgcQ

Corolaire 12.197.
Soient f et g, deux fonctions dérivables sur $[a,b]$ telles que

$$f'(x) = g'(x) \tag{12.495}$$

pour tout $x \in [a,b]$. Alors il existe un réel C tel que $f(x) = g(x) + C$ pour tout $x \in [a,b]$.

Démonstration. Considérons la fonction $h(x) = f(x) - g(x)$, dont la dérivée est, par hypothèse, nulle. L'annulation de la dérivée entraine par le corolaire 12.197 que h est constante. Si $h(x) = C$, alors $f(x) = g(x) + C$, ce qu'il fallait prouver. □

DefXVMVooWhsfuI

Définition 12.198.
*Soient I un intervalle ouvert de $\mathbb{R}$ et une fonction $f : I \to \mathbb{R}$. La fonction $F : I \to \mathbb{R}$ est une **primitive** de f si F est dérivable sur I et si $F'(x) = f(x)$ pour tout x dans I.*

Exprimé en termes des primitives, le corolaire 12.197 signifie que

CorZeroCst

Corolaire 12.199.
Si F et G sont deux primitives de la même fonction f sur un intervalle, alors il existe une constante C pour laquelle $F(x) = G(x) + C$.

Cela signifie qu'il n'y a, en réalité, pas des milliards de primitives différentes à une fonction. Il y en a essentiellement une seule, et puis les autres, ce sont juste les mêmes, mais décalées d'une constante.

Remarque 12.200.
L'hypothèse de se limiter à un intervalle est importante parce que si on considère la fonction sur deux intervalles disjoints, nous pouvons choisir la constante indépendamment dans l'un et dans l'autre. Par exemple la fonction

$$F(x) = \begin{cases} \ln(x) + 1 & \text{si } x > 0 \\ \ln(x) - 7 & \text{si } x < 0 \end{cases} \tag{12.496}$$

est une primitive de $\frac{1}{x}$ sur l'ensemble $\mathbb{R}\backslash\{0\}$.

Certains ne s'en privent pas. Le logiciel Sage par exemple fait ceci :

```
sage: f(x)=1/x
sage: F=f.integrate(x)
sage: A=F(x)-F(-x)
sage: A.full_simplify()
I*pi
```

En réalité lorsque $x > 0$, Sage définit $\ln(-x) = \ln(x) + i\pi$. Cela a une certaine logique parce que $\ln(-1) = i\pi$ (du fait que $e^{i\pi} = -1$), mais si on ne le sait pas, ça peut étonner.

12.201.
Il existe plusieurs primitives à une fonction donnée. En physique, la constante arbitraire est souvent fixée par une condition initiale, comme nous le verrons dans la section 42.1.

12.18 Fonctions de plusieurs variables

La physique, et les sciences en général, regorgent de fonctions à plusieurs variables.

Accélération centripète [109] Si une masse m tourne sur un cercle, elle subira une accélération dirigée vers l'intérieur égale à

$$F(v,r) = \frac{mv^2}{r} \tag{12.497}$$

où r est le rayon du cercle et v est la vitesse.

Pression dans un gaz Si on a n moles de gaz dans un volume V a une température T, alors la pression sera donnée par la fonction de trois variables

$$p = \frac{nRT}{V} \tag{12.498}$$

où R est la constante des gaz parfaits.

En mathématique, on peut inventer de nombreuses fonctions de plusieurs variables. La fonction

$$f(x,y) = x^2 + xy\cos(x^2+y^3) \tag{12.499}$$

est définie sur $\mathbb{R}^2$. La fonction

$$f(x,y,z) = \frac{x+y-2z}{1-x^2-y^2-z^2} \tag{12.500}$$

est définie sur $\mathbb{R}^3$ moins la sphère unité $\{x^2+y^2+z^2=1\}$.

Consacrons nous à l'étude des fonctions de plusieurs variables, en donnant tout d'abord quelques indications sur comment «dessiner» une telle fonction. Vous connaissez déjà la définition de graphe pour une fonction f d'une seule variable à valeurs dans $\mathbb{R}$: c'est l'ensemble des points du plan de la forme $(x,f(x))$. Vous voyez que cet ensemble n'est pas vraiment un gros morceau de $\mathbb{R}^2$ parce que son intérieur est vide : il y a une seule valeur de f qui correspond au point x, donc une boule de $\mathbb{R}^2$ centrée en $(x,f(x))$ de n'importe quel rayon contient toujours des points qui ne font pas partie du graphe de f.

Nous voulons donner une définition assez générale pour le graphe d'une fonction.

Définition 12.202.
*Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Le **graphe** de f est la partie de $\mathbb{R}^m \times \mathbb{R}^n$ de la forme*

$$\operatorname{Graph} f = \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n \,|\, y = f(x)\}. \tag{12.501}$$

Cette définition se spécialise de la façon suivante dans les cas communs. Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}$. Le graphe de f est la partie de $\mathbb{R}^m \times \mathbb{R}$ donné par

$$\operatorname{Graph} f = \{(x,y) \in \mathbb{R}^m \times \mathbb{R} \,|\, y = f(x)\}. \tag{12.502}$$

Et pour les fonctions $\mathbb{R}^2 \to \mathbb{R}$:

$$\operatorname{Graph} f = \{(x,y,z) \in \mathbb{R}^3 \text{ tel que } z = f(x,y)\}. \tag{12.503}$$

C'est cette définition qu'il faut garder à l'esprit lorsqu'on travaille sur des dessins en trois dimensions.

Si f est une fonction de deux variables indépendantes x et y à valeurs dans $\mathbb{R}$, alors un point dans le graphe de f est un point $(x,y,z) \in \mathbb{R}^3$ tel que

$$z = f(x,y), \tag{12.504}$$

ou encore, un point de la forme

$$\big(x,y,f(x,y)\big). \tag{12.505}$$

Nous avons parfois besoin de donner des représentations graphiques d'une fonction. Nous pouvons, par exemple, penser à la fonction qui associe à un point de la Terre, son altitude. Lorsqu'on

109. Appelez la « centrifuge » si vous voulez ; ça ne me fait ni chaud ni froid.

part pour une promenade en montagne on a envie de connaitre le graphe de cette fonction qui correspond en fait à la surface de la montagne. Bien sûr, nous ne voulons pas amener avec nous un modèle en 3D de la montagne, donc il nous faut une méthode efficace pour projeter le graphe de f sur le plan x-y tout en gardant les informations fondamentales. Pour cela nous avons besoin de deux définitions (à ne pas confondre !)

Définition 12.203.
*Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et soit c dans $\mathbb{R}$. La **z-section** de* Graph f *à la hauteur c est donnée par*

$$S_c^z = \{(x, y, c) \in \mathbb{R}^3 \mid f(x, y) = c\}.$$

def_niveau

Définition 12.204.
*Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ et soit c dans $\mathbb{R}$. La **courbe de niveau** de f à la hauteur c est l'ensemble*

$$N_c = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid f(x_1, \ldots, x_n) = c\}. \tag{12.506}$$

On peut représenter la fonction f d'une façon très précise en traçant quelques-unes de ses courbes de niveau. Dans la suite, on pourra considérer aussi les x-sections et les y-sections du graphe d'une fonction de deux variables. La x-section de Graph f à la hauteur a est

$$S_a^x = \{(a, y, z) \in \mathbb{R}^3 \mid f(a, y) = z\}.$$

Comme vous avez peut être déjà compris, S_a^x est le graphe de la fonction de y qu'on obtient de f en fixant $x = a$. Cette fonction est appelée x-section de f pour $x = a$.

Certaines surfaces dans $\mathbb{R}^3$ sont le graphe d'une fonction.

Exemple 12.205.
Quelques graphes importants.

Un plan non vertical Tout plan dans $\mathbb{R}^3$ peut être décrit par une équation de la forme

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = r,$$

où, (x_0, y_0, z_0) est vecteur dans $\mathbb{R}^3$, et a, b, c et r sont des nombres réels. Si $c \neq 0$ alors le plan n'est pas vertical et on peut dire qu'il est le graphe de la fonction

$$P(x, y) = \frac{r + cz_0 - a(x - x_0) - b(y - y_0)}{c},$$

quitte à choisir des nouvelles constantes s, t, q,

$$P(x, y) = sx + ty + q.$$

Un paraboloïde elliptique Pour tous α et β dans $\mathbb{R}$ les graphes des fonctions

$$PE_1(x, y) = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$$

ou de la fonction

$$PE_2(x, y) = -\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$$

sont des paraboloïdes elliptiques. Le premier est contenu dans le demi-espace $z \geqslant 0$, l'autre dans $z \leqslant 0$. Le nom de cette surface vient de la forme de ses sections. En fait toutes sections S_c^z sont des ellipses, alors que les sections S_a^x et S_b^y sont des paraboles.

Un paraboloïde hyperbolique (selle) Pour tous α et β dans $\mathbb{R}$ les graphes des fonctions

$$PH_1(x, y) = \frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$$

ou de la fonction

$$PH_2(x,y) = -\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$$

sont des paraboloïdes hyperboliques. Remarquez que les sections S_c^z de ce graphe sont des hyperboles, alors que les sections S_a^x et S_b^y sont des paraboles.

Une demi-sphère La fonction $S^+(x,y) = \sqrt{R^2 - x^2 - y^2}$ a pour graphe la demi-sphère supérieure centrée en l'origine et de rayon R. Le dernier de ces exemples nous signale une chose très importante : une sphère entière n'est pas le graphe d'une fonction de x et y. Par contre, une demi-sphère est bien le graphe de la fonction $f(x,y) = \sqrt{1 - x^2 - y^2}$.

L'équation que nous utilisons pour décrire une sphère de rayon R centrée en l'origine est

$$x^2 + y^2 + z^2 = R^2$$

Donc, à chaque point (x,y) dans le disque $x^2 + y^2 \leqslant R^2$ (notez que ce disque est contenu dans la section S_0^z), on peut associer deux valeurs de z : $z_1 = \sqrt{R^2 - x^2 - y^2}$ et $z_2 = -\sqrt{R^2 - x^2 - y^2}$. Par définition, une fonction n'associe qu'une seule valeur à chaque point de son domaine, d'où l'impossibilité de décrire cette sphère comme le graphe d'une fonction de x et y.

△

Considérons la fonction $Sp : \mathbb{R}^3 \to \mathbb{R}$ qui associe à (x,y,z) la valeur $x^2 + y^2 + z^2$. La sphère de rayon R centrée en l'origine est l'ensemble de niveau N_{R^2} de Sp. L'ensemble de niveau N_0 de Sp est l'origine, et tous les ensembles de niveau de hauteur négative sont vides. La même chose est vraie pour les ellipsoïdes centrées en l'origine avec les axes x, y et z comme axes principaux et comme longueurs de demi-axes a, b et c. Voici la fonction dont il sont les ensembles de niveau

$$El(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}.$$

Exemple 12.206.
Des ensembles de niveau importants.

Tout graphe Le graphe de toute fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ peut être considéré comme l'ensemble de niveau zéro de la fonction $F(x,y,z) = z - f(x,y)$.

Hyperboloïdes Les hyperboloïdes, comme les ellipsoïdes, sont une famille d'ensemble de niveau. En particulier, nous considérons des hyperboloïdes dont l'axe de symétrie est l'axe des z et qui sont symétriques par rapport un plan x-y. Une fois que les paramètres a, b et c sont fixés la fonction que nous intéresse est

$$Hyp(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}.$$

Les ensembles de niveau N_d pour $d > 0$ sont connexes, on les appelle *hyperboloïdes à une feuille*. L'ensemble de niveau N_0 est un *cône (elliptique)*, les deux moitiés du cône se touchent en l'origine. Enfin, les ensembles de niveau N_d pour $d < 0$ ne sont pas connexes et pour cette raison on les appelle *hyperboloïdes à deux feuilles*.

△

12.18.1 Graphes de fonctions à plusieurs variables

Le **graphe** d'une fonction de deux variables $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$ est l'ensemble

$$\Big\{\big(x,y,f(x,y)\big) \text{ tel que } (x,y) \in D\Big\} \subset \mathbb{R}^3. \tag{12.507}$$

Ce graphe est une surface dans $\mathbb{R}^3$.

ExempleTroisDxxyy

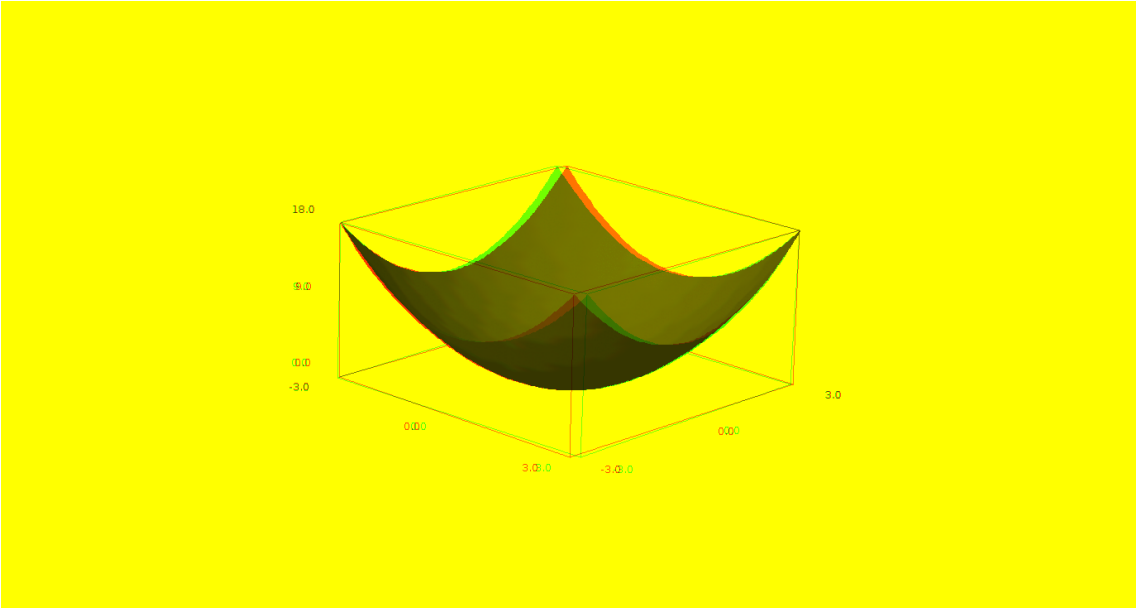
Exemple 12.207.
Tracer le graphe de la fonction

$$(x, y) \mapsto x^2 + y^2. \tag{12.508}$$

Le plus simple est de demander à Sage de nous fournir une représentation 3D

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2+y**2
sage: plot3d(f,(x,-3,3),(y,-3,3))
```

Voici ce que cela donne [110] : (à regarder avec des lunettes bleues et rouges) :



À part que l'ordinateur l'a dit, est-ce qu'on peut comprendre pourquoi le graphe de la fonction $x^2 + y^2$ ressemble à un bol ? En coordonnées cylindriques, le graphe s'écrit

$$z = r^2. \tag{12.509}$$

Donc il apparaît que plus on s'éloigne du point $(0, 0)$ dans le plan XY, plus le graphe va monter. Et il monte à quelle vitesse ? Il monte à la vitesse r^2. Il s'agit donc de dessiner la fonction $z = r^2$ dans le plan et de la « faire tourner ». △

12.18.2 Courbes de niveau

Une technique utile pour se faire une idée de la forme d'une fonction en trois dimensions est de tracer les **courbes de niveau**. La courbe de niveau de hauteur h est la courbe dans le plan donnée par l'équation

$$f(x, y) = h. \tag{12.510}$$

Exemple 12.208.
Dessinons par exemple les courbes de niveau de la fonction

$$f(x, y) = x + y + 2. \tag{12.511}$$

110. En vrai, ce que Sage donne est un objet qu'on peut même faire bouger.

La courbe de niveau h est donnée par l'équation $x + y + 2 = h$, c'est-à-dire

$$y(x) = -x + h - 2. \tag{12.512}$$

Par conséquent la courbe de niveau de hauteur 0 est $y = -x - 2$, celle de hauteur 5 est $y = -x + 3$, etc.

Nous pouvons également nous aider de Sage pour ce faire :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x+y+2
sage: var('h')
h
sage: niveau(h,x)=solve(f(x,y)==h,y)[0].rhs()
sage: g1(x)=niveau(1,x)
sage: g1
x |--> -x - 1
```

Ici la fonction `g1` est la courbe de niveau 1.

Si on veut faire tracer une courbe de niveau, Sage peut le faire :

```
sage: implicit_plot(f(x,y)==1,(x,-3,3),(y,-4,4))
```

Cela tracera la courbe de niveau $h = 1$ dans la partie du plan $x \in [-3, 3]$ et $y \in [-4, 4,]$.

△

Il est bien entendu possible de créer automatiquement 50 courbes de niveau et de demander de les tracer toutes sur le même graphe.

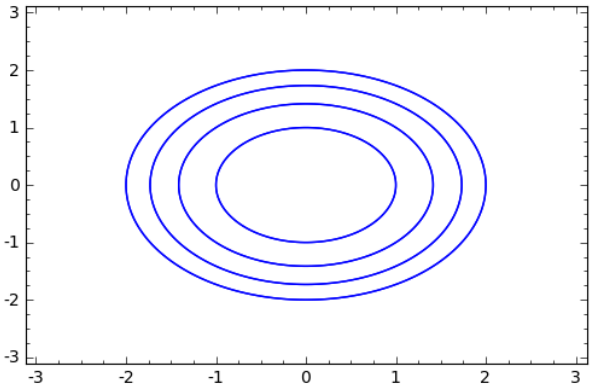
```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

var('x,y')
f=x**2+y**2
G=Graphics()
a=3
for i in range(0,5):
  G=G+implicit_plot(f==i,(x,-a,a),(y,-a,a))
show(G)
```

tex/frido/courbeNiveau.py

Le résultat est :



Notez que les courbes sont censées être des cercles : les axes X et Y n'ont pas la même échelle.

Exemple 12.209.
Un exemple plus riche en enseignements est celui de la fonction

$$f(x, y) = x^2 - y^2. \tag{12.513}$$

La courbe de niveau h est donnée par l'équation $x^2 - y^2 = h$.

Commençons par $h = 0$. Dans ce cas nous avons $(x + y)(x - y) = 0$ et par conséquent les courbes de niveau de hauteur zéro sont les deux droites $x + y = 0$ et $x - y = 0$.

Voyons ensuite la courbe de niveau $h = 1$. Cela est l'équation $x^2 - y^2 = 1$, c'est-à-dire

$$y(x) = \pm\sqrt{x^2 - 1}. \tag{12.514}$$

C'est une fonction qui n'est définie que pour $|x| \geqslant 1$. Avec $x = 1$ nous avons $y = 1$. Ensuite, lorsque x grandit, y grandit également, mais la courbe ne peut pas croiser la courbe de niveau $h = 0$. Donc, suivant les notations de la figure 12.6, la courbe de niveau « part » de P et doit monter sans croiser les diagonales.

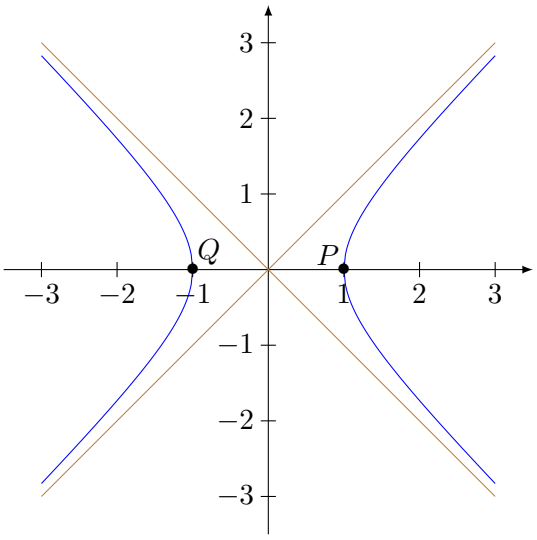


FIGURE 12.6: La courbe de niveau $h = 1$ de $x^2 - y^2$. Notez qu'elle est en deux morceaux. LabelFigCQIXooBEDnfK

En ce qui concerne la courbe de niveau $h = -1$, elle correspond à la courbe $y = \pm\sqrt{1 + x^2}$ qui est définie pour tous les $x \in \mathbb{R}$. Le même raisonnement que précédemment nous amène à la figure 12.7. △

Une autre façon de voir les courbes de niveau est de dire que la courbe de niveau de hauteur h est la projection dans le plan XY de la section du graphe de f par le plan $z = h$.

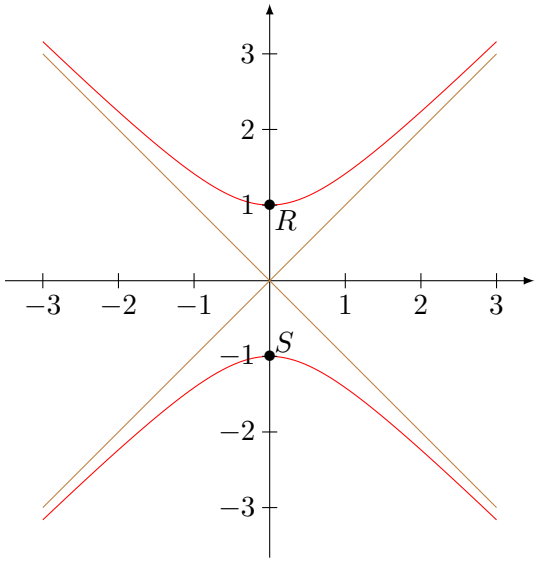


FIGURE 12.7: La courbe de niveau $x^2 - y^2 = -1$. LabelFigKGQXooZFNVnW

On peut également définir le graphe de fonctions de trois (ou plus) variables. Le graphe de la fonction $f\colon D \subset \mathbb{R}^3 \to \mathbb{R}$ est l'ensemble

$$\big\{\big(x, y, z, f(x, y, z)\big) \text{ tel que } (x, y, z) \in D\big\} \subset \mathbb{R}^4. \tag{12.515}$$

De tels graphes ne peuvent pas être représentés sur une feuille de papier. Il est toutefois possible de définir les ensembles de niveaux :

$$E_h = \big\{(x, y, z) \in D \text{ tel que } f(x, y, z) = h\big\}. \tag{12.516}$$

Ce sont des surfaces dans $\mathbb{R}^3$ que l'on peut dessiner.

Exemple 12.210.
Les surfaces de niveau de la fonction $f(x, y, z) = x^2 + y^2 + z^2$ sont des sphères. Il n'y a pas de surfaces de niveau pour les « hauteurs » négatives. △

Exemple 12.211.
Considérons la fonction $f(x, y, z) = x^2 + y^2 - z^2$. En coordonnées cylindriques, cette fonction s'écrit

$$f(r, \theta, z) = r^2 - z^2. \tag{12.517}$$

La surface de niveau 0 est donnée par l'équation $r = |z|$. Cela fait un cercle à chaque hauteur, dont le rayon grandit linéairement avec la hauteur ; le tout est donc un cône. C'est d'ailleurs le cône obtenu par rotation de la courbe de niveau $h = 0$ que nous avions obtenu pour la fonction $x^2 - y^2$.

En ce qui concerne les ensembles de niveau positifs, ils sont donnés par

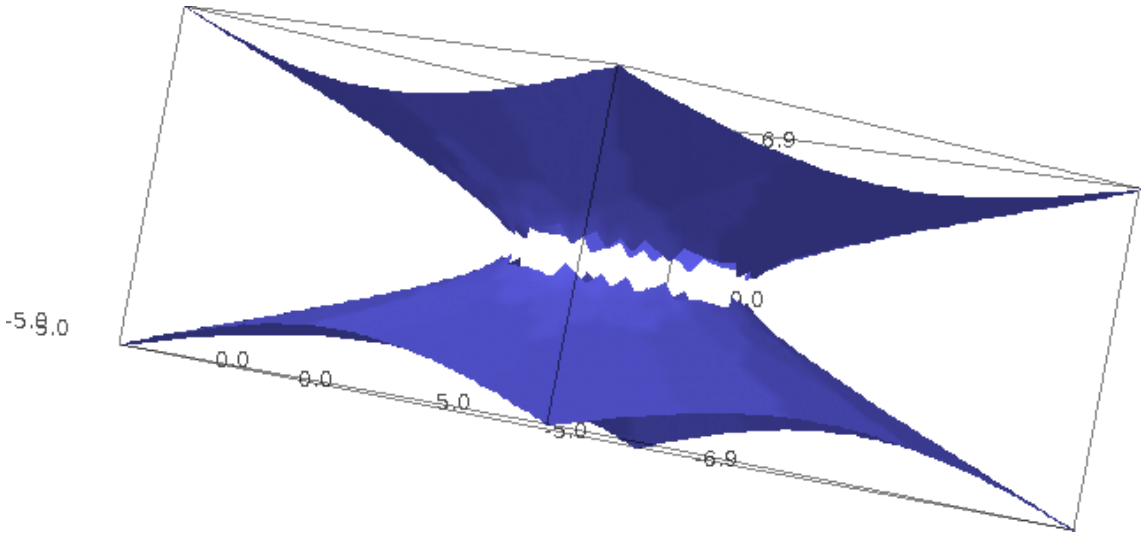
$$z = \pm\sqrt{x^2 + y^2 - h}. \tag{12.518}$$

Notez qu'ils ne sont pas définis pour $r \geqslant h$. Cela pose un petit problème quand on veut le tracer à l'ordinateur :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y')
(x, y)
sage: f(x,y)=sqrt(x**2+y**2-3)
```

```
sage: F=plot3d(f(x,y),(x,-5,5),(y,-5,5))
sage: G=plot3d(-f(x,y),(x,-5,5),(y,-5,5))
sage: F+G
```

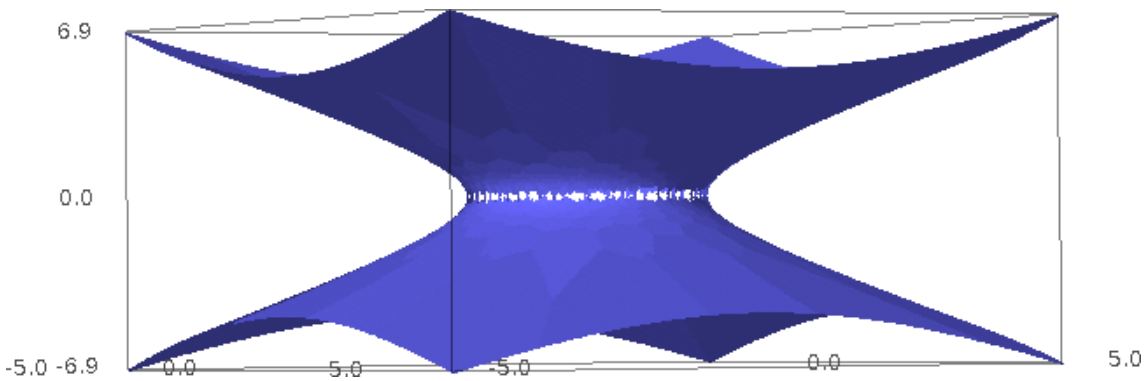
Le résultat est [111] :



On voit qu'il y a un grand trou au centre correspondant aux z proches de zéro. Or d'après l'équation, il n'en est rien : en $z = 0$ il y a bel et bien tout un cercle. Afin d'obtenir une meilleur image, il faut demander de tracer avec un maillage plus fin :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y')
(x, y)
sage: f(x,y)=sqrt(x**2+y**2-3)
sage: F=plot3d(f(x,y),(x,-5,5),(y,-5,5),plot_points=300)
sage: G=plot3d(-f(x,y),(x,-5,5),(y,-5,5),plot_points=300)
sage: F+G
```

Le temps de calcul est un peu plus long, mais le résultat est meilleur :



△

111. Encore une fois : ça donne mieux à l'écran, et vous pouvez le faire bouger ; je vous encourage à le faire !

12.19 Limites à plusieurs variables

SecLimVarsPlus

Prenons une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$. Nous disons que

$$\lim_{x\to x_0} f(x) = l \in \mathbb{R} \tag{12.519}$$

lorsque $\forall \epsilon > 0$, $\exists \delta$ tel que $\|x - x_0\| \leqslant \delta$ implique $|f(x) - l| \leqslant \epsilon$.

Remarquez qu'ici, $x \in \mathbb{R}^n$, et sachez distinguer $\|.\|$, la norme dans $\mathbb{R}^n$ de $|.|$ qui est la valeur absolue dans $\mathbb{R}$. Une autre façon d'exprimer cette définition est que l'ensemble des valeurs atteintes par f dans une boule de rayon δ autour de x_0 n'est pas très loin de l. Nous définissons donc

$$E_\delta = \{f(x) \text{ tel que } x \in B(x_0, \delta)\}. \tag{12.520}$$

Notez que si f n'est pas définie en x_0, il n'y a pas de valeurs correspondantes au centre de la boule dans E_δ. Ceci est évidemment la situation générique lorsqu'il y a une indétermination à lever dans le calcul de la limite. Nous avons alors

$$\lim_{x\to x_0} f(x) = l \tag{12.521}$$

lorsque $\forall \epsilon > 0$, $\exists \delta$ tel que

$$\sup\{|v - l| \text{ tel que } v \in E_\delta\} \leqslant \epsilon. \tag{12.522}$$

Eqvmoinsrapplimdeux

Une façon classique de montrer qu'une limite n'existe pas, est de prouver que, pour tout δ, l'ensemble E_δ contient deux valeurs constantes. Si par exemple $0 \in E_\delta$ et $1 \in E_\delta$ pour tout δ, alors aucune valeur de l (même pas $l = \pm\infty$) ne peut satisfaire à la condition (12.522) pour toute valeur de ϵ.

Nous laissons à la sagacité de l'étudiant le soin d'adapter tout ceci pour le cas $\lim_{x\to x_0} f(x) = \pm\infty$.

La proposition suivante semble évidente, mais nous sera tellement utile qu'il est préférable de l'expliciter :

PROPooPOAQooPmxEtb

Proposition 12.212.
Soient $f : D \to \mathbb{R}$ une fonction de domaine D, $a \in \mathrm{Adh}(D)$ et un voisinage V de a. Nous supposons que $V \cap D$ s'écrive comme une intersection finie :

$$V \cap D = \bigcap_{i=1}^{k} A_i$$

telle que $a \in \mathrm{Adh}\, A_i$ pour tout $i \leqslant k$. Alors, la limite

$$\lim_{x\to a} f(x) \tag{12.523}$$

EQooEXELooHccCGw

existe et vaut $b \in \mathbb{R}$ si et seulement si chacune des limites

$$\lim_{\substack{x\to a\\ x\in A_i}} f(x) \tag{12.524}$$

existe et vaut b.

Démonstration. On sait déjà que si la limite de $f : D \to \mathbb{R}$ existe, alors toute restriction à A_i admet la même limite [112]. Il suffit donc de prouver la réciproque.

Fixons provisoirement un entier i entre 1 et k ainsi que $\epsilon > 0$. Vu que $\lim_{\substack{x\to a\\ x\in A_i}} f(x) = b$, il existe $\delta_i > 0$ tel que si $x \in A_i$ et si $0 < |x - a| < \delta_i$, alors

$$|f(x) - f(a)| < \epsilon. \tag{12.525}$$

EQooUCAlooRpUguE

Quitte à prendre δ_i un peu plus petit, nous supposons que $V \subset B(a, \delta_i)$.

Nous posons $\delta = \min\{\delta_i\}_{i=1,\dots,k}$, et nous considérons $x \in D$ tel que $0 < |x - a| < \delta$. Nous avons alors

112. C'est une conséquence de la caractérisation séquentielle de la continuité 7.210.

(1) $x \in V \cap D$,
(2) il existe i tel que $x \in A_i$.

Ce x est donc un élément de A_i vérifiant $0 < |x-a| < \delta \leqslant \delta_i$. Il vérifie donc (12.525) : $|f(x)-f(a)| < \epsilon$.

Cela prouve la limite (12.523). □

Exemple 12.213. (1) Pour qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ admette une limite en $a \in \mathbb{R}$, il faut et il suffit qu'elle y admette une limite à droite et une limite à gauche qui soient égales.

Cela est une application de la proposition 12.212 avec $\mathbb{R} =]-\infty, a[\cup]a, \infty[$.

(2) Une suite (x_k) admet une limite si et seulement si les sous-suites (x_{2k}) et (x_{2k+1}) convergent vers la même limite. Ceci n'est pas une application directe de la proposition, mais la teneur est la même.

△

Lemme 12.214 ([1]).
Soient deux espaces vectoriels E et F. Soit une fonction $f\colon \mathbb{R} \to F$ telle que $\lim_{t\to 0} f(t) = y \in F$. Nous posons

$$\begin{aligned} \varphi\colon E &\to F \\ h &\mapsto f(\|h\|). \end{aligned} \tag{12.526}$$

Alors φ admet une limite pour $h \to 0$ et elle est donnée par

$$\lim_{h\to 0} \varphi(h) = \lim_{t\to 0} f(t). \tag{12.527}$$

Démonstration. Soit $\epsilon > 0$. Il existe δ tel que si $t < \delta$ alors $\|f(t) - y\|_F < \epsilon$. Si $\|h\| < \delta$ nous avons

$$\|\varphi(h) - y\| = \|f(\|h\|) - y\| < \epsilon. \tag{12.528}$$

Donc c'est bon. □

Voici, dans le même ordre d'idée, un autre résultat qui permet de réduire le nombre de variables dans une limite lorsque la fonction ne dépend pas de certaines variables.

LEMooYLIHooFBQyzC

Lemme 12.215 ([1]).
Soit une fonction $g\colon \mathbb{R} \to \mathbb{R}$ vérifiant

$$\lim_{t\to a} g(t) = \ell. \tag{12.529}$$

Alors pour tout $b \in \mathbb{R}$, la fonction

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto g(x) \end{aligned} \tag{12.530}$$

vérifie

$$\lim_{\substack{(x,y)\to(a,b)\\ x\neq a}} f(x, y) = \ell. \tag{12.531}$$

Démonstration. Soit $\epsilon > 0$. Par hypothèse sur la limite de g en a, il existe $\delta > 0$ tel que $0 < |t - a| < \delta$ implique $|g(t) - \ell| < \epsilon$.

Attention : passage subtil [113]. Si $0 < \|(x, y) - (a, b)\| < \delta$, alors nous avons évidemment aussi $|x - a| < \delta$, mais pas spécialement $0 < |x - a| < \delta$ comme le requis pour utiliser la limite de g.

Dans le calcul de la limite restreinte à $x \neq a$, les points qui interviennent sont les valeurs de (x, y) dans $B\big((a, b), \delta\big)\backslash\{x = a\}$. Or pour celles-là nous avons bien $0 < |x-a| < \delta$. Le calcul suivant fonctionne donc :

$$|f(x, y) - \ell| = |g(x) - \ell| < \epsilon. \tag{12.532}$$

□

113. Je rejette déjà en bloc et d'un revers de main toute tentative de dire « la limite épointée, c'est mieux ». Voir aussi l'exemple 12.216.

EXooHSYNooBZhDbE

Exemple 12.216.
Pourquoi prendre la limite $(x,y)\to(a,b)$ avec $x\neq a$ dans l'énoncé du lemme 12.215 ? Imaginons la fonction

$$g(x)=\begin{cases}0 & \text{si } x\neq 0\\ 1 & \text{si } x=0.\end{cases}\tag{12.533}$$

Dans ce cas, le graphe de la fonction $f(x,y)=g(x)$ est tout plat sauf la ligne $x=0$ qui est en hauteur. Nous avons donc $f(0,t)=1$ pour tout t et donc nous n'avons pas $\lim_{(x,y)\to(0,0)}f(x,y)=0$: tout voisinage de $(0,0)$ contient des points (x,y) tels que $f(x,y)=1$ et des points (x,y) tels que $f(x,y)=0$. △

12.217.
Il existe de nombreuses façons de calculer des limites à plusieurs variables. Plus nous connaîtrons de mathématiques, plus nous aurons de techniques à notre disposition. Nous allons tout de suite voir quelques méthodes. Voir le thème 67 pour plus de techniques et d'exemples.

12.19.1 Caractérisation de la limite par les suites

ExFNExempleMethodeTrigigi

Exemple 12.218.
Considérons la fonction

$$f(x,y)=\frac{xy}{x^2+y^2},\tag{12.534}$$

et remarquons que, quelle que soit la valeur de y, cette fonction est nulle lorsque $x=0$. De la même manière, nous voyons que si $x=y$, alors la fonction vaut [114] $\frac{1}{2}$.

Il est impossible que la fonction ait une limite en $(0,0)$ parce qu'on ne peut pas trouver un ℓ dont on s'approche à la fois en suivant la ligne $x=0$ et la ligne $x=y$.

Deux autres chemins avec encore deux autres valeurs sont dessinés sur la figure 12.8.

Cet exemple pourra être formalisé en utilisant le théorème 12.219. Voir l'exemple 12.220. △

ThoLimSuite

Théorème 12.219 (Caractérisation de la limite par les suites)**.**
Une fonction $f\colon D\subset\mathbb{R}^m\to\mathbb{R}^n$ admet une limite ℓ en un point d'accumulation a de D si et seulement si pour toute suite (x_n) dans $D\backslash\{a\}$ convergente vers a, la suite $\big(f(x_n)\big)$ dans $\mathbb{R}^n$ converge vers ℓ.

Démonstration. Supposons d'abord que la fonction ait une limite ℓ lorsque $x\to a$, et considérons une suite (x_n) dans $D\backslash\{a\}$ convergente vers a. Nous devons montrer que la suite $y_n=f(x_n)$ converge vers ℓ, c'est-à-dire que si nous choisissons $\varepsilon>0$ nous devons montrer qu'il existe un N tel que $n>N$ implique $\|y_n-\ell\|=\|f(x_n)-\ell\|<\varepsilon$.

Nous avons deux hypothèses. La première est la convergence de la fonction et la seconde est la convergence de la suite (x_n). L'hypothèse de convergence de la fonction nous dit que (le ε a déjà été choisi dans le paragraphe précédent)

$$\exists\delta \text{ tel que } 0<\|x-a\|<\delta\Rightarrow\|f(x)-\ell\|<\varepsilon.\tag{12.535}$$

Une fois choisi ce δ qui « va avec » le ε qui a été choisi précédemment, la définition de la convergence de la suite nous enseigne que

$$\exists N \text{ tel que } n>N\Rightarrow\|x_n-a\|<\delta.\tag{12.536}$$

Récapitulons ce que nous avons fait. Nous avons choisi un ε, et puis nous avons construit un N. Lorsque $n>N$, nous avons $\|x_n-a\|<\delta$. Mais alors, par construction de ce δ, nous avons $\|f(x_n)-\ell\|<\varepsilon$. Au final, $n>N$ implique bien $\|y_n-\ell\|<\varepsilon$, ce qu'il nous fallait.

Nous supposons maintenant que la fonction f *ne* converge *pas* vers ℓ, et nous allons construire une suite d'éléments x_n qui converge vers a sans que $(y_n)=(f(x_n))$ ne converge vers ℓ. La fonction

114. En fait ce que nous sommes en train de faire est de poser $\theta=\pi/2$ et $\theta=\pi/4$ dans (18.821).

f vérifie la condition (12.67). Nous prenons donc un ε tel que $\forall\delta$, il existe un x qui vérifie *en même temps* les deux conditions

$$\begin{cases} 0 < \|x-a\| < \delta & (12.537a)\\ \|f(x)-\ell\| > \varepsilon. & (12.537b)\end{cases}$$

Un tel x existe pour tout choix de δ. Choisissons un n arbitraire et $\delta = \frac{1}{n}$. Nous nommons x_n le x correspondant à ce choix de n. La suite (x_n) ainsi construite converge vers a parce que

$$\|x_n - a\| < \delta_n = \frac{1}{n}, \tag{12.538}$$

donc dès que n est grand, $\|x_n - a\|$ est petit. Mais la suite $y_n = f(x_n)$ ne converge pas vers ℓ parce que

$$\|f(x_n) - \ell\| > \varepsilon \tag{12.539}$$

pour tout n. La suite y_n ne s'approche donc jamais à moins d'une distance ε de ℓ. □

EXooNBTYooFyKRTB

Exemple 12.220.
Reprenons l'exemple 12.218. Considérons les deux suites $x_n = (0, \frac{1}{n})$ et $y_n = (\frac{1}{n}, \frac{1}{n})$. Ce sont deux suites dans $\mathbb{R}^2$ qui tendent vers $(0,0)$. Si la fonction f convergeait vers ℓ, alors nous aurions au moins

Eq3007Lixxyyell

$$\lim f(x_n) = \ell \tag{12.540a}$$
$$\lim f(y_n) = \ell, \tag{12.540b}$$

mais nous savons que pour tout n, $f(x_n) = f(0, \frac{1}{n}) = 0$ et $f(y_n) = f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2}$. Il n'y a donc aucun nombre ℓ qui vérifie les deux équations (12.540) parce que $\lim f(x_n) = 0$ et $\lim f(y_n) = \frac{1}{2}$. △

12.19.2 Règle de l'étau

Une première façon de calculer la limite d'une fonction est de la « coincer » entre deux fonctions dont nous connaissons la limite.

ThoRegleEtau

Théorème 12.221 (Règle de l'étau[408])**.**
Soit $\mathcal{O}$, un ouvert de $\mathbb{R}^m$ contenant le point a. Soient f, g et h, trois fonctions définies sur $\mathcal{O}\backslash\{a\}$. Supposons que

(1)

$$g(x) \leqslant f(x) \leqslant h(x). \tag{12.541}$$

pour tout $x \in \mathcal{O}\backslash\{a\}$

(2)

$$\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = \ell. \tag{12.542}$$

Alors la limite $\lim_{x\to a} f(x)$ existe et vaut ℓ.

Démonstration. Soit $\epsilon > 0$. Nous devons trouver un voisinage V de a tel que $|f(x) - \ell| < \epsilon$ pour tout $x \in V\backslash\{a\}$. Étant donné que $\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = \ell$, il existe un voisinage ouvert V de a tel que [115]

$$|g(x) - \ell| < \epsilon \tag{12.543a}$$
$$|h(x) - \ell| < \epsilon. \tag{12.543b}$$

pour tout $x \in V\backslash\{a\}$. Pour tout $x \neq a$ dans V, nous avons

$$\ell - \epsilon \leqslant g(x) \leqslant f(x) \leqslant h(x) \leqslant \ell + \epsilon. \tag{12.544}$$

Donc $|f(x) - \ell| < \epsilon$, comme nous l'avions annoncé. □

115. Si vous ne voyez pas comment avoir les deux conditions en même temps, prenez V_1 pour g et V_2 pour h, et considérez $V = V_1 \cap V_2$ qui sera encore un ouvert.

Cette méthode est très pratique lorsqu'on a des fonctions trigonométriques qui se factorisent parce qu'elles sont toujours majorables par 1 ; voir l'exemple 18.73.

Exemple 12.222.
Prouver la continuité en $(0,0)$ de la fonction

$$f(x,y)=\begin{cases}\frac{x|y|}{\sqrt{x^2+y^2}} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases}\tag{12.545}$$

Considérons une suite $(x_n,y_n)\in\mathbb{R}^2$ qui tend vers $(0,0)$. Étant donné que $\frac{|y|}{\sqrt{x^2+y^2}}<1$ pour tout x et y, nous avons

$$0\leqslant|f(x_n,y_n)|=\left|\frac{x_n|y_n|}{\sqrt{x_n^2+y_n^2}}\right|\leqslant|x_n|\to 0.\tag{12.546}$$

Donc nous avons

$$\lim_{(x,y)\to(0,0)}f(x,y)=0=f(0,0),\tag{12.547}$$

ce qui prouve que la fonction est continue en $(0,0)$ par la proposition 7.137. Nous avons utilisé la règle de l'étau (théorème 12.221). △

12.223.
Nous notons $f\sim g$ pour $x\to a$ lorsque $\lim_{x\to a}\frac{f(x)}{g(x)}=1$.

Cela signifie que f et g tendent vers la même limite, à la même vitesse.

12.19.3 Méthode des chemins

Lorsque la limite n'existe pas, il y a une façon en général assez simple de le savoir, c'est la **méthode des chemins**.

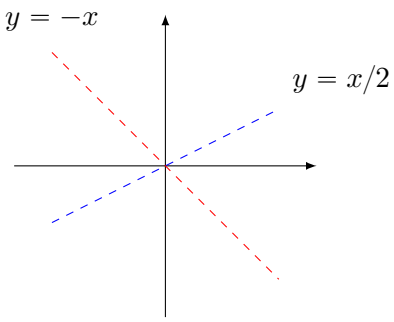


FIGURE 12.8: Sur toute la droite $y=-x$, la fonction vaut $-1/2$, tandis que sur toute la droite $y=x/2$, elle vaut $\frac{2}{5}$. Il est donc impossible que la fonction ait une limite en $(0,0)$, parce que dans toute boule autour de zéro, il y aura toujours un point de chacune de ces deux droites. LabelFigMethodeChemin

C'est la proposition suivante qui va faire une grosse partie du travail.

PROPooSAFIooWvmSiT

Proposition 12.224 ([1]).
Soit $f\colon D\subset\mathbb{R}^m\to\mathbb{R}^n$ et a un point d'adhérence de D. Alors nous avons

$$\lim_{x\to a}f(x)=\ell\tag{12.548}$$

si et seulement si pour toute fonction $\gamma\colon\mathbb{R}\to\mathbb{R}^m$ telle que $\lim_{t\to0}\gamma(t)=a$, nous avons

$$\lim_{t\to0}(f\circ\gamma)(t)=\ell.\tag{12.549}$$

Démonstration. En deux parties.

(1) **Sens direct** Soit une fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ telles que $\lim_{t\to 0}\gamma(t) = a$. Par le théorème 12.219, il suffit de montrer que $(f\circ\gamma)(t_n) \to \ell$ pour toute suite $t_n \to 0$ dans $\mathbb{R}$.
Nous savons que la suite $n \mapsto \gamma(t_n)$ est une suite qui converge vers a. Mais l'hypothèse $\lim_{x\to a} f(x) = \ell$ implique que pour toute suite $x_n \to a$ nous avons $f(x_n) \to \ell$. Cela est en particulier vrai pour la suite $n \mapsto \gamma(t_n)$. Donc :

$$\lim_{n\to\infty} f\big(\gamma(t_n)\big) = \ell, \tag{12.550}$$

ce qu'il fallait prouver.

(2) **Réciproque** Pour les mêmes raisons de caractérisation séquentielle que précédemment, il faut prouver que $\lim_{n\to\infty} f(x_n) = \ell$ pour tout suite $x_n \to a$.

(2a) **Un chemin** Soit la fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ affine par morceaux et telle que

$$\gamma\left(\frac{1}{n}\right) = x_n. \tag{12.551}$$

Nous prolongeons γ par $\gamma(t) = a$ pour $t \leqslant 0$.

(2b) $\gamma(t) \to a$ Nous montrons que $\lim_{t\to 0}\gamma(t) = a$. Soient $\epsilon > 0$ et N tel que $x_n \in B(a,\epsilon)$ pour tout $n \geqslant N$. Si $t < \frac{1}{N}$ alors $t \in [\frac{1}{k+1}, \frac{1}{k}]$ pour un certain $k > N$. Donc

$$\gamma(t) \in [\gamma(\frac{1}{k+1}), \gamma(\frac{1}{k})] \tag{12.552}$$

et donc $\gamma(t) \in [x_{k+1}, x_k]$ parce que γ est formé de ces segments de droites. Mais comme $B(a,\epsilon)$ est convexe [116], nous avons

$$\gamma(t) \in [x_{k+1}, x_k] \subset B(a,\epsilon). \tag{12.553}$$

Nous avons donc bien $\lim_{t\to 0}\gamma(t) = a$.

(2c) **Conclusion** L'hypothèse nous donne alors $\lim_{t\to 0}(f\circ\gamma)(t) = \ell$. En particulier le critère de la caractérisation séquentielle de la limite dit que

$$\lim_{n\to\infty} f\big(\gamma(\frac{1}{n})\big) = \ell, \tag{12.554}$$

ce qui signifie $\lim_{n\to\infty} f(x_n) = \ell$.

□

CorMethodeChemin

Corolaire 12.225.
Soient $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et a un point d'accumulation de D. Si nous avons deux fonctions $\gamma_1, \gamma_2\colon \mathbb{R} \to \mathbb{R}^m$ telles que

$$\lim_{t\to 0}\gamma_1(t) = \lim_{t\to 0}\gamma_2(t) = a \tag{12.555}$$

tandis que

$$\lim_{t\to 0}(f\circ\gamma_1)(t) \neq \lim_{t\to 0}(f\circ\gamma_2)(t), \tag{12.556}$$

ou bien que l'une des deux limites n'existe pas, alors la limite de $f(x)$ lorsque $x \to a$ n'existe pas.

CorMethodeChemoinNegatif

Corolaire 12.226.
Soient $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et a un point d'accumulation de D. Si il existe une fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ avec $\gamma(0) = a$ telle que la limite $\lim_{t\to 0}(f\circ\gamma)(t)$ n'existe pas, alors la limite $\lim_{x\to a} f(x)$ n'existe pas.

En ce qui concerne le calcul de limites, la méthode des chemins peut être utilisé de trois façons :

116. C'est la proposition 11.28.

(1) Dès que l'on trouve une fonction $\gamma\colon \mathbb{R} \to \mathbb{R}^m$ telle que $\lim_{t\to 0}(f\circ\gamma)(t) = \ell$, alors nous savons que *si la limite* $\lim_{x\to a} f(x)$ *existe*, alors cette limite vaut ℓ.

(2) Dès que l'on a trouvé deux fonctions γ_i qui tendent vers a, mais dont les limites de $\lim_{t\to 0}(f\circ\gamma_i)(t)$ sont différentes, alors la limite $\lim_{x\to a} f(x)$ n'existe pas.

(3) Dès qu'on trouve une chemin le long duquel il n'y a pas de limite, alors la limite n'existe pas (corolaire 12.226).

La méthode des chemins ne permet donc pas de calculer une limite quand elle existe. Elle permet uniquement de la « deviner », ou bien de prouver que la limite n'existe pas.

Exemple 12.227.
Soit à calculer

$$\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}. \tag{12.557}$$

Eq3007ExempleLimiche

Si nous prenons le chemin $\gamma_1(t) = (t,t)$, nous avons bien $\lim_{t\to 0}\gamma_1(t) = (0,0)$, et nous avons

$$\lim_{t\to 0}(f\circ\gamma_1)(t) = \lim_{t\to 0}\frac{t-t}{t+t} = 0. \tag{12.558}$$

Donc si la limite (12.557) existait, elle vaudrait obligatoirement 0. Mais si nous considérons $\gamma_2(t) = (0,t)$, nous avons

$$(f\circ\gamma_2)(t) = \frac{-t}{t} = -1, \tag{12.559}$$

donc si la limite existe, elle doit obligatoirement valoir -1. Ne pouvant être égale à 0 et à -1 en même temps, la limite (12.557) n'existe pas. △

12.20 Dérivée directionnelle

Nous sommes capables de dériver une fonction de deux variables $f(x,y)$ par rapport à x et par rapport à y. C'est-à-dire que nous sommes capables de donner la variation de la fonction lorsqu'on bouge le long des axes horizontal et vertical. Il est évidemment souhaitable de parler de la variation de la fonction lorsqu'on se déplace le long d'autre droites.

Soit donc $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ un vecteur unitaire (c'est-à-dire $u_1^2 + u_2^2 = 1$), et considérons la fonction de une variable

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(a+tu_1, b+tu_2). \end{aligned} \tag{12.560}$$

La fonction φ n'est rien d'autre que la fonction f vue le long de la droite de direction donnée par le vecteur u. Nous pouvons aussi l'écrire $\varphi(t) = f(p+tu)$.

Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables et soit $(a,b) \in \mathbb{R}^2$. La façon la plus naturelle de définir une dérivée à deux variables est de considérer les **dérivées partielles** définies par

$$\begin{aligned} \frac{\partial f}{\partial x}(a,b) &= \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \\ \frac{\partial f}{\partial y}(a,b) &= \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b}. \end{aligned} \tag{12.561}$$

Ces nombres représentent la façon dont le nombre $f(x,y)$ varie lorsque soit seul x varie soit seul y varie. Les dérivées partielles se calculent de la même façon que les dérivées normales. Pour calculer $\partial_x f$, on fait « comme si » y était une constante, et pour calculer $\partial_y f$, on fait comme si x était une constante.

12.20.1 Dérivée partielle et directionnelles

Lorsque f est différentiable, la dérivée directionnelle est donnée par

$$\frac{\partial f}{\partial u}(p) = \nabla f(p) \cdot u. \tag{12.562}$$

EqDerDirnablau

LEMooASGNooCWqAmN

Lemme 12.228.
Les projections canonniques sont des applications différentiables.

LEMooGYJUooOudNLH

Lemme 12.229.
Toute fonction polynômiale à n variables est différentiable comme application de $\mathbb{R}^n$ vers $\mathbb{R}$.

LEMooUMRXooDUQHpK

Lemme 12.230.
Toute fonction rationnelle, du type $f(x) = \frac{P(x)}{Q(x)}$ où P et Q sont des polynômes, est différentiable en tout point a tel que $Q(a) \neq 0$.

LEMooTPDSooGbxfhP

Lemme 12.231.
Pour une fonction d'une variable $f : D \subset \mathbb{R} \to \mathbb{R}$, le caractère différentiable et le caractère dérivable coïncident. De plus, on a

$$df_a(u) = f'(a)u. \tag{12.563}$$

deriveepartielles

12.232.
En pratique, ayant une formule pour la fonction f, nous la dérivons par rapport à la variable x_i en utilisant les règles usuelle de dérivation en considérant que les autres (x_j avec $j \neq i$) sont des constantes.

Exemple 12.233.
Pour $f(x, y) = xy + x^2$, les dérivées partielles s'écrivent

$$\frac{\partial f}{\partial x} = y + 2x \quad \text{et} \quad \frac{\partial f}{\partial y} = x$$

△

12.20.2 Gradient : direction de plus grande pente

Étant donné que u est de norme 1, l'inégalité de Cauchy-Schwarz donne

$$\left|\nabla f(a, b) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right| \leqslant \|\nabla f(a, b)\|. \tag{12.564}$$

Donc

$$-\|\nabla f(p)\| \leqslant \nabla f(p) \cdot u \leqslant \|\nabla f(p)\|. \tag{12.565}$$

La norme de la dérivée directionnelle (qui est la valeur absolue du nombre au centre) est donc « coincée » entre $-\|\nabla f(p)\|$ et $\|\nabla f(p)\|$. Prenons par exemple

$$u = \frac{\nabla f(p)}{\|\nabla f(p)\|}. \tag{12.566}$$

Dans ce cas, nous avons exactement

$$\nabla f(p) \cdot u = \|\nabla f(p)\|, \tag{12.567}$$

qui est la valeur maximale que la dérivée directionnelle peut prendre.

La direction du gradient est donc la direction suivant laquelle la dérivée directionnelle est la plus grande. Pour la même raison, la dérivée directionnelle est la plus petite dans le sens opposé au gradient.

En termes bien clairs : lorsqu'on veut aller le plus vite possible au ski, on prend la direction du gradient de la piste de ski. C'est dans cette direction que ça descend le plus vite. Dans quelle direction vont les débutants ? Ils vont perpendiculairement à la pente (ce qui ennuie tout le monde, mais c'est un autre problème). Les débutants vont donc dans la direction perpendiculaire au gradient. Prenons donc $u \perp \nabla f(p)$ et calculons la dérivée directionnelle de f dans la direction u en utilisant la formule 12.562 :

$$\frac{\partial f}{\partial u}(p) = \nabla f(p) \cdot u = 0 \tag{12.568}$$

parce que nous avons choisi $u \perp \nabla f(p)$. Nous voyons donc que les débutants en ski ont eu la bonne intuition que la direction dans laquelle la piste ne descend pas, c'est la direction perpendiculaire au gradient.

C'est aussi pour cela que l'on a tendance à faire du zig-zag à vélo lorsqu'on monte une pente très forte et qu'on est fatigué. C'est toujours pour cela que les routes de montagne font de longs lacets. La montée est moins rude en suivant une direction proche d'être perpendiculaire au gradient !

Théorème 12.234.
Le gradient des fonctions suit à peu près les mêmes règles que les dérivées. Soient f et g deux fonctions différentiables. Nous avons entre autres

(1) $\nabla(f+g) = \nabla f + \nabla g$;

(2) $\nabla(fg)(a,b) = g(a,b)\nabla f(a,b) + f(a,b)\nabla g(a,b)$;

(3) Dès que $g(a,b) \neq 0$, nous avons

$$\nabla\frac{f}{g} = \frac{g(a,b)\nabla f(a,b) - f(a,b)\nabla g(a,b)}{g(a,b)^2}. \tag{12.569}$$

12.20.3 Gradient : orthogonal au plan tangent

Vu que le gradient d'une fonction est la direction de plus grande pente et que le plan tangent est le plan de plus petite pente, quoi de plus naturel que de penser que le gradient est orthogonal au plan tangent ?

Lemme 12.235.
Soit $\phi\colon \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 et la partie

$$\Gamma = \{x \in \mathbb{R}^n \text{ tel que } \phi(x) = C\} \tag{12.570}$$

pour une certaine constante C.

Soit $x_0 \in \Gamma$. Le gradient de ϕ en x_0 est orthogonal au plan tangent à Γ en x_0.

Démonstration. Un vecteur tangent à Γ en x_0 est de la forme $\gamma'(0)$ où $\gamma\colon \mathbb{R} \to \Gamma$ vérifie $\gamma(0) = x_0$. Puisque ϕ est constante sur Γ nous avons

$$\frac{d}{ds}\Big[\phi\big(\gamma(s)\big)\Big]_{s=0} = 0, \tag{12.571}$$

ce qui donne

$$\sum_i \frac{\partial\phi}{\partial x_i}\big(\gamma(0)\big)\gamma_i'(0) = 0, \tag{12.572}$$

ce qui signifie exactement $\langle(\nabla\phi)(x_0), \gamma'(0)\rangle = 0$. Le vecteur $(\nabla\phi)(x_0)$ est donc perpendiculaire à tout vecteur tangent de Γ en x_0. □

12.20.4 Mise en bouche en dimension 2

Nous savons déjà comment dériver les fonctions composées de $\mathbb{R}$ dans $\mathbb{R}$. C'est la proposition 12.181. Si nous avons deux fonctions $f\colon \mathbb{R} \to \mathbb{R}$ et $u\colon \mathbb{R} \to \mathbb{R}$, nous formons la composée $\varphi = f \circ u\colon \mathbb{R} \to \mathbb{R}$ dont la dérivée vaut

$$\varphi'(a) = f'\big(u(a)\big)u'(a). \tag{12.573}$$

Considérons maintenant le cas un peu plus compliqué des fonctions $f\colon \mathbb{R}\to\mathbb{R}$ et $u\colon \mathbb{R}^2\to\mathbb{R}$, et de la composée

$$\begin{aligned}\varphi\colon \mathbb{R}^2 &\to \mathbb{R}\\ \varphi(x,y) &= f\big(u(x,y)\big).\end{aligned} \tag{12.574}$$

Afin de calculer la dérivée partielle de φ par rapport à x, nous admettons que pour tout a, b et t, il existe $c\in[a,a+t]$ tel que

$$u(a+t,b)=u(a,b)+t\frac{\partial u}{\partial x}(c,b). \tag{12.575}$$

Cela est une généralisation immédiate du théorème 12.192. Nous devons calculer

$$\frac{\partial\varphi}{\partial x}(a,b)=\lim_{t\to0}\frac{\varphi(a+t,b)-\varphi(a,b)}{t}=\lim_{t\to0}\frac{f\big(u(a+t,b)\big)-f\big(u(a,b)\big)}{t}. \tag{12.576}$$

EqPremPasDiffxyp

Étant donné l'hypothèse que nous avons faite sur u, nous avons

$$f\big(u(a+t,b)\big)=f\big(u(a,b)+t\frac{\partial u}{\partial x}(c,b)\big). \tag{12.577}$$

En utilisant le théorème des accroissements finis pour f, nous avons un point d entre $u(a,b)$ et $u(a,b)+t\frac{\partial u}{\partial x}(c,b)$ tel que

$$f\big(u(a,b)+t\frac{\partial u}{\partial x}(c,b)\big)=f\big(u(a,b)\big)+t\frac{\partial u}{\partial x}(c,b)f'(d). \tag{12.578}$$

Le numérateur de (12.576) devient donc

$$t\frac{\partial u}{\partial x}(c,b)f'(d). \tag{12.579}$$

Certes les points c et d sont inconnus, mais nous savons que c est entre a et $a+t$ ainsi que d se situe entre $u(a,b)$ et $u(a,b)+t\frac{\partial u}{\partial x}(c,b)$. Lorsque nous prenons la limite $t\to0$, nous avons donc $\lim_{t\to0}c=a$ et $\lim_{t\to0}d=u(a,b)$. Nous avons alors

$$\lim_{t\to0}\frac{t\frac{\partial u}{\partial x}(c,b)f'(d)}{t}=\frac{\partial u}{\partial x}(a,b)f'\big(u(a,b)\big). \tag{12.580}$$

La formule que nous avons obtenue (de façon pas très rigoureuse) est

$$\frac{\partial}{\partial x}f\big(u(x,y)\big)=\frac{\partial u}{\partial x}(x,y)f'\big(u(x,y)\big). \tag{12.581}$$

Prenons maintenant un cas un peu plus compliqué où nous voudrions connaître les dérivées partielles de la fonction φ donnée par

$$\varphi(x,y,z)=f\big(u(x,y),v(x,y,z)\big) \tag{12.582}$$

où $f\colon\mathbb{R}^2\to\mathbb{R}$, $u\colon\mathbb{R}^2\to\mathbb{R}$ et $v\colon\mathbb{R}^3\to\mathbb{R}$.

Commençons par la dérivée partielle par rapport à z. Étant donné que φ ne dépend de z que via la seconde entrée de f, il est normal que seule la dérivée partielle de f par rapport à sa seconde entrée arrive dans la formule :

$$\frac{\partial\varphi}{\partial z}(x,y,z)=\frac{\partial f}{\partial v}\big(u(x,y),v(x,y,z)\big)\frac{\partial v}{\partial z}(x,y,z). \tag{12.583}$$

La dérivée partielle par rapport à y demande de tenir compte en même temps de la façon dont f varie avec sa première entrée et la façon dont elle varie avec sa seconde entrée ; cela nous fait deux termes :

$$\frac{\partial\varphi}{\partial y}(x,y,z)=\frac{\partial f}{\partial u}\big(u(x,y),v(x,y,z)\big)\frac{\partial u}{\partial y}(x,y)+\frac{\partial f}{\partial v}\big(u(x,y),v(x,y,z)\big)\frac{\partial v}{\partial y}(x,y,z). \tag{12.584}$$

Cette formule a une interprétation simple. Lançons un caillou du sommet d'une falaise. Son mouvement est une chute libre avec une vitesse initiale horizontale :

$$\begin{cases} x(t) = v_0 t & (12.585a) \\ y(t) = h_0 - \frac{gt^2}{2} & (12.585b) \end{cases}$$

où v_0 est la vitesse initiale horizontale et h_0 est la hauteur de la falaise. Si nous sommes intéressés à la distance entre le caillou et le bas de la falaise (point $(0,0)$), le théorème de Pythagore nous dit que

$$d(t) = \sqrt{x^2(t) + y^2(t)}. \tag{12.586}$$

Pour trouver la variation de la distance par rapport au temps il faut savoir de combien la distance varie lorsque x varie et multiplier par la variation de x par rapport à t, et puis faire la même chose avec y.

12.20.5 Accroissements finis et dérivées partielles

PROPooCAWBooINcNxj

Proposition 12.236 (Accroissements finis)**.**
Soient des espaces vectoriels normés V et W. Soit une application $f\colon V \to W$. Soient des points $a, b \in V$ tels que f est continue sur le segment $[a,b]$ et partiellement dérivable dans la direction $b-a$ sur $]a,b]$. Alors il existe $c \in]a,b[$ tel que

$$f(b) = f(a) + (\partial_\beta f)(c) \tag{12.587}$$

où $c \in [a,b]$ et $\beta = b - a$.

Démonstration. Nous considérons la fonction

$$\begin{aligned} \varphi\colon [0,1] &\to W \\ t &\mapsto f\big(a + t(b-a)\big). \end{aligned} \tag{12.588}$$

Par le théorème des accroissements finis 12.192, il existe $s \in [0,1]$ tel que [117]

$$\varphi(1) = \varphi(0) + \varphi'(s)(1-0). \tag{12.589}$$

Autrement dit,

$$f(b) = f(a) + (\partial_\beta f)\big(a + s(b-a)\big). \tag{12.590}$$

Nous avons le résultat en posant $c = a + s(b-a)$. □

LEMooNMTAooLgMkgH

Lemme 12.237 (Accroissements finis[1])**.**
Soient des espaces vectoriels normés V et W. Soit une fonction $f\colon V \to W$ qui est différentiable sur un voisinage $\mathcal{O}$ de $a \in V$. Soient $v \in V$ et $\epsilon > 0$ tels que $a + \epsilon v$ reste dans $\mathcal{O}$.

Nous considérons une base de V pour donner un sens aux dérivées partielles $\partial_k f$. Alors il existe une fonction $\alpha\colon V \to W$ telle que

$$f(a+\epsilon v) = f(a) + \sum_{k=1}^{n} \epsilon v_k (\partial_k f)\big(a + \sum_{i=k+1}^{n} \epsilon v_i e_i\big) + \epsilon\alpha(\epsilon) \tag{12.591}$$

où la somme sur i est nulle dans le cas $k = n$.

Démonstration. Nous commençons par nous attaquer à la dérivation par rapport à la première variable. La définition 10.244 de la dérivation partielle nous invite à poser

$$\varphi(t) = f\big(a + \sum_{k=2}^{n} \epsilon v_k e_k + t v_1 e_1\big). \tag{12.592}$$

117. Les a et b dans l'énoncé de 12.192 sont les valeurs $s = 0$ et $s = 1$ ici. Rien à voir avec les a et b d'ici qui sont des éléments de V.

Nous avons :

$$\varphi'(0) = v_1(\partial_1 f)\big(a + \sum_{k=2}^{n} \epsilon v_k e_k\big). \tag{12.593}$$

Nous appliquons les accroissements finis 10.257 à la fonction φ en $t = 0$. Nous avons une fonction $\alpha\colon \mathbb{R} \to W$ telle que $\lim_{t\to 0} \alpha_1(t) = 0$ et

$$\varphi(t) = \varphi(0) + t\varphi'(0) + t\alpha_1(t). \tag{12.594}$$

Nous écrivons cette égalité pour $t = \epsilon$, tout en rappelant que $\varphi(\epsilon) = f(a + \epsilon v)$:

$$f(a + \epsilon v) = \varphi(\epsilon) = f\big(a + \sum_{k=2}^{n} \epsilon v_k e_k\big) + \epsilon v_1(\partial_1 f)\big(a + \sum_{k=2}^{n} \epsilon v_k e_k\big) + \epsilon\alpha_1(\epsilon). \tag{12.595}$$

Pour la suite, il suffit de recommencer en écrivant $\sum_{k=2}^{n} \epsilon v_k e_k = \epsilon v_2 e_2 + \sum_{k=3}^{n} \epsilon v_k e_k$ dans le second terme. □

Voici une version un peu moins technologique.

PROPooYYSMooUDxtlB

Proposition 12.238.
Soit une fonction $f\colon V \to \mathbb{R}$ où V est un espace vectoriel métrique. Soit $a \in V$ tel que $(\partial_i f)(a)$ existe. Alors il existe une fonction $\beta\colon \mathbb{R} \to \mathbb{R}$ tel que

$$f(a + \epsilon e_i) = f(a) + \epsilon(\partial_i f)(a) + \epsilon\beta(\epsilon) \tag{12.596}$$

et $\lim_{\epsilon\to 0} \beta(\epsilon) = 0$.

Démonstration. Nous posons

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(a + te_i). \end{aligned} \tag{12.597}$$

Par hypothèse (et définition 10.244 de la dérivée partielle), la fonction φ est dérivable et

$$\varphi'(0) = (\partial_i f)(a). \tag{12.598}$$

Nous appliquons le théorème des accroissements finis 10.257 sur la fonction φ :

EQooGSLNooIcrLIb

$$\varphi(t) = \varphi(0) + t\varphi'(0) + t\beta(t) \tag{12.599}$$

pour une certaine fonction $\beta\colon \mathbb{R} \to \mathbb{R}$ qui vérifie $\lim_{t\to 0} \beta(t) = 0$. En remplaçant φ par sa valeur en termes de f dans (12.599),

$$f(a + te_i) = f(a) + t(\partial_i f)(a) + t\beta(t). \tag{12.600}$$

□

12.21 Formes différentielles

SecFormDiffRappel

Nous parlerons de formes différentielles exactes et fermées dans la section 20.84.

12.21.1 Décomposition dans la base duale

DEFooMGXSooWioKie

Définition 12.239.
*Soit U, un ouvert dans $\mathbb{R}^n$. Une 1-**forme différentielle** ω sur U est une application*

$$\begin{aligned} \omega\colon U &\to (\mathbb{R}^n)^* \\ x &\mapsto \omega_x. \end{aligned} \tag{12.601}$$

*Une k-**forme différentielle** est une application*

$$\omega\colon U \to \mathcal{L}_k(\mathbb{R}^n, \mathbb{R}) \tag{12.602}$$

telle que ω_x est alternée pour tout x.

Étant donné que $\mathcal{L}_k(\mathbb{R}^n, \mathbb{R})$ est un espace vectoriel normé (définition 10.85), nous savons ce qu'est une forme différentielle continue ou de classe C^k.

REMooSKKSooDZFMAr

Remarque 12.240.
L'ensemble des 1-formes différentielles devient un espace vectoriel avec les définitions

$$\begin{aligned}(\lambda\omega)_x(v) &= \lambda\omega_x(v)\\ (\omega+\mu)_x(v) &= \omega_x(v) + \mu_x(v).\end{aligned} \tag{12.603}$$

Nous connaissons la base de $(\mathbb{R}^n)^*$ définie en 4.127. Nous allons noter ces formes par dx_i :

$$\begin{aligned}e_1^* = dx_1\colon v &\mapsto v_1\\ &\vdots\\ e_n^* = dx_n\colon v &\mapsto v_n\end{aligned} \tag{12.604}$$

EQooITHKooDzigPY

Toute forme différentielle s'écrit

$$\omega_x = \sum_{i=0}^{n} a_i(x)dx_i \tag{12.605}$$

où $a_1, \ldots, a_n$ sont les composantes de ω dans la base usuelle, et sont des fonctions à valeurs réelles.

LEMooNSFKooLOKGGy

Lemme 12.241.
*Une 1-forme différentielle est **continue** si les fonctions a_i sont continues. La forme sera C^k quand les a_i seront C^k.*

Pour un vecteur $v = (v_1, \ldots, v_n)$ on a donc par définition de dx_i

$$\omega_x(v) = \sum_{i=0}^{n} a_i(x)v_i. \tag{12.606}$$

Ces fonctions a_i peuvent être trouvées en appliquant ω aux éléments de la base canonique de $\mathbb{R}^n$:

$$a_j(x) = \omega_x(e_j) \tag{12.607}$$

parce que $\omega_x(e_j) = \sum_i a_i(x)dx_i(e_i) = \sum_i a_i(x)\delta_{ij} = a_j(x)$.

12.21.2 L'isomorphisme musical

Si G est un champ de vecteurs sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$, nous pouvons définir

$$\begin{aligned}G_x^\flat\colon \mathbb{R}^n &\to \mathbb{R}\\ v &\mapsto \langle G(x), v\rangle\end{aligned} \tag{12.608}$$

EqDefBemol

Pour chaque x, l'application $G_x^\flat$ est une forme sur $\mathbb{R}^n$, c'est-à-dire une application linéaire de $\mathbb{R}^n$ vers $\mathbb{R}$. Nous écrivons que

$$G_x^\flat \in \big(\mathbb{R}^n\big)^*. \tag{12.609}$$

Nous pouvons ainsi déterminer le développement de $G^\flat$ dans la base des dx_i en faisant le calcul

$$G_x^\flat(e_i) = \langle G(x), e_i\rangle = G_i(x), \tag{12.610}$$

donc les composantes de $G^\flat$ dans la base dx_i sont exactement les composantes de G dans la base e_i :

$$G_x^\flat = G_1(x)dx_1 + \cdots + G_n(x)dx_n. \tag{12.611}$$

La construction inverse existe également. Si ω est une 1-forme différentielle, nous pouvons définir le champ de vecteurs $\omega^\sharp$ par la formule (implicite)

$$\omega_x(v) = \langle \omega^\sharp(x), v \rangle \tag{12.612}$$

pour tout $v \in \mathbb{R}^n$. Par définition, $(\omega^\sharp)^\flat = \omega$.

Lemme 12.242.
En composantes nous avons :

$$\omega^\sharp(x) = \big(a_1(x), \dots, a_n(x)\big). \tag{12.613}$$

Si G est un champ de vecteurs, alors $(G^\flat)^\sharp = G$.

12.22 Différentielle

Nous avons déjà donné une définition abstraite de la différentielle dans la définition 10.243. Nous en voyons maintenant quelques motivations dans le cas de fonctions sur $\mathbb{R}^2$ ou $\mathbb{R}^n$.

12.22.1 Exemples introductifs

SEBSECooLPRQooJRQCFL

La notion de dérivée est associée à la recherche de la droite tangente à une courbe. Reprenons rapidement le cheminement. La dérivée de $f\colon \mathbb{R} \to \mathbb{R}$ au point a est un nombre $f'(a)$, qui définit donc une application linéaire dont le coefficient angulaire est $f'(a)$, et que nous notons df_a :

$$\begin{aligned} df_a\colon \mathbb{R} &\to \mathbb{R} \\ u &\mapsto f'(a)u. \end{aligned} \tag{12.614}$$

La droite donnée par l'équation

$$y(a+u) = f'(a)u \tag{12.615}$$

est parallèle à la tangente en a. Pour trouver la tangente, il suffit de la décaler de la hauteur qu'il faut. L'équation de la droite tangente au graphe de f au point $\big(a, f(a)\big)$ devient

$$y(x) = f(a) + f'(a)(x-a) = f(a) + df_a(x-a). \tag{12.616}$$ EqDiffRapTgDer

Nous nous proposons de généraliser cette formule au cas de la recherche du plan tangent à une surface.

Exemple 12.243.
Considérons $f(x,y) = x^2y + y^2e^x$. Les dérivées partielles sont

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy + y^2e^x \\ \frac{\partial f}{\partial y} &= x^2 + 2ye^x. \end{aligned} \tag{12.617}$$

△

Cet exemple était l'exemple facile où tout se passe bien.

Exemple 12.244.
Les choses sont moins simples lorsqu'on considère la fonction suivante :

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases} \tag{12.618}$$

On voit que pour tout x et tout y, nous avons $f(x,0)=f(0,y)=0$. Donc cette fonction est nulle sur les axes horizontaux et verticaux. Nous avons en particulier

$$\begin{aligned}\frac{\partial f}{\partial x}(0,0)&=0\\ \frac{\partial f}{\partial y}(0,0)&=0.\end{aligned}\tag{12.619}$$

Donc ces dérivées partielles existent.

Il n'est par contre pas question de dire que cette fonction « va bien » autour du point $(0,0)$. En effet si nous regardons sa valeur sur la droite diagonale $y=x$, nous avons

$$f(x,x)=\frac{x^2}{2x^2}=\frac{1}{2}.\tag{12.620}$$

Par conséquent si nous suivons la fonction le long de la droite $y=x$, la hauteur vaut $\frac{1}{2}$ en permanence, sauf juste en $(0,0)$ où la fonction fait un grand plongeon !

```
sage: var('x,y')
(x, y)
sage: f(x,y)=(x*y)/(x**2+y**2)
sage: plot3d(f,(x,-2,2),y(-2,2))
```

D'ailleurs elle fait un plongeon le long de toutes les droites (sauf verticale et horizontale). En effet si nous regardons la fonction le long de la droite $y=mx$, nous avons

$$f(x,mx)=\frac{mx^2}{x^2+m^2x^2}=\frac{m}{1+m^2}.\tag{12.621}$$

La fonction est donc *constante* sur chacune de ces droites. Il n'est donc pas question de dire que cette fonction est « dérivable » en $(0,0)$, vu qu'elle fait des grands sauts dans presque toutes les directions. △

Nous devons donc trouver mieux que les dérivées partielles pour étudier le comportement des fonctions un peu problématiques.

12.22.2 Différentielle

Nous nous souvenons de l'équation (12.392) qui nous dit que pour une fonction d'une variable la dérivabilité signifiait qu'il existait un nombre ℓ et une fonction α tels que

$$f(x)=f(a)+\ell(x-a)+(x-a)\alpha(x-a)\tag{12.622}$$

et $\lim_{t\to 0}\alpha(t)=0$.

En nous inspirant de cela, nous comprenons peut-être un peu le pourquoi de la définition 10.243.

12.245.
L'objet df_a est *en soi* une application $df_a\colon\mathbb{R}^m\to\mathbb{R}^n$. Nous notons $df_a(u)$ la valeur de df_a sur le vecteur $u\in\mathbb{R}^m$. En particulier, l'application df est une forme différentielle au sens de la définition 12.239.

12.246.
Les propositions 12.257 et 12.260 vont montrer qu'en étudiant bien les dérivées partielles, nous pouvons conclure à la différentiabilité d'une fonction. Attention cependant, nous verrons dans l'exemple 12.266 que l'existence des dérivées directionnelles partielles ne permettait pas de conclure à la différentiabilité.

12.22.3 Matrice de la différentielle

La différentielle est une application linéaire. Elle possède donc une matrice lorsque des bases sont fixées.

PROPooBMROooThgLuU

Proposition 12.247.
Soient une application différentiable $f\colon \mathbb{R}^m \to \mathbb{R}^n$ et $a \in \mathbb{R}^m$. Dans les bases canoniques de $\mathbb{R}^m$ et $\mathbb{R}^n$, la matrice de df_a est

$$(df_a)_{ij} = \frac{\partial f_i}{\partial x_j}(a). \tag{12.623}$$

Démonstration. Le lien entre matrice et application linéaire est vu dans la proposition 4.71. Dans le cas des bases canoniques de $\mathbb{R}^m$ et $\mathbb{R}^n$ nous savons qu'extraire une composante revient à prendre le produit scalaire. Nous avons donc

$$(df_a)_{ij} = \big(df_a(e_j)\big)_i = df_a(e_j) \cdot e_i. \tag{12.624}$$

La linéarité de la dérivation donne alors

$$(df_a)_{ij} = df_a(e_j) \cdot e_i = \frac{d}{dt}\Big[f(a+te_j)\Big]_{t=0} \cdot e_i = \frac{d}{dt}\Big[f_i(a+te_j)\Big]_{t=0} = \frac{\partial f_i}{\partial x_j}(a). \tag{12.625}$$

□

LEMooDDUZooLwXkRp

Lemme 12.248 ([1]).
Soit une fonction $g\colon \mathbb{R} \to \mathbb{R}$ de classe C^∞. Nous posons

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto g(x). \end{aligned} \tag{12.626}$$

Alors f est de classe C^∞ sur $\mathbb{R}^2$.

Démonstration. Le problème lorsqu'il faut démontrer qu'une fonction est de classe C^∞, c'est que $d^k f$ sera une application de $\mathbb{R}^2$ vers un espace qui est un terrible emboîtement de $\mathcal{L}(\mathbb{R}^2, \dots)$. Pour traiter cette difficulté, nous considérons les espaces suivants : $V_0 = \mathbb{R}$ et par récurrence $V_{k+1} = \mathcal{L}(\mathbb{R}^2, V_k)$.

Et nous considérons également les éléments

$$\begin{aligned} \alpha_1\colon \mathbb{R}^2 &\to \mathbb{R} \\ (u,v) &\mapsto u \end{aligned} \tag{12.627}$$

et plus généralement $\alpha_k \in V_k$ donné par

$$\begin{aligned} \alpha_k\colon \mathbb{R}^2 &\to V_{k-1} \\ (u,v) &\mapsto u\alpha_{k-1}. \end{aligned} \tag{12.628}$$

Notons que dans l'expression $u\alpha_{k-1}$, il s'agit d'un produit entre un scalaire $u \in \mathbb{R}$ et un vecteur $\alpha_{k+1} \in V_{k-1}$.

Nous prouvons maintenant par récurrence que $d^k f_{(a,b)} = g^{(k)}(a)\alpha_k$, en utilisant directement la définition.

(1) **Initialisation**
Pour $k = 1$, nous calculons

$$\frac{|f(a+h_1, b+h_2) - f(a,b) - g'(a)\alpha_1(h)|}{\|h\|} = \frac{|g(a+h_1) - g(a) - g'(a)h_1|}{\|h\|} \tag{12.629}$$

Notre but est de calculer la limite de cela lorsque $h \xrightarrow{\mathbb{R}^2} 0$ avec $h \neq 0$. L'hypothèse sur la dérivabilité de g nous indique que si $0 < |t| < \delta$, alors

$$\frac{|g(a+t) - g(a) - tg'(a)|}{|t|} < \epsilon. \tag{12.630}$$

EQooQLWNooLRKhUv

Nous considérons donc la boule épointée de $\mathbb{R}^2$ de rayon δ : $B = B\big((0,0),\delta\big)\backslash\{(0,0)\}$, et nous considérons $h \in B$. Deux cas sont à distinguer : $h_1 = 0$ et $h_1 \neq 0$.
Si $h_1 = 0$, alors

$$\frac{|g(a+h_1) - g(a) - g'(a)h_1|}{\|h\|} = 0. \tag{12.631}$$

Sinon nous avons $0 < h_1 \leqslant \|h\| < \delta$ et donc

$$\frac{|g(a+h_1) - g(a) - g'(a)h_1|}{\|h\|} \leqslant \frac{|g(a+h_1) - g(a) - g'(a)h_1|}{|h_1|} < \epsilon \tag{12.632}$$

par la relation (12.630). Nous avons donc bien

$$\lim_{h\to 0} \frac{|f(a+h_1, b+h_2) - f(a,b) - g'(a)\alpha_1(h)|}{\|h\|} = 0. \tag{12.633}$$

(2) **Récurrence**
Nous supposons que $d^k f_{a,b} = g^{(k)}(a)\alpha_k$, et nous devons prouver que $d^k f$ est différentiable et que $d^{k+1} f_{(a,b)} = g^{(k+1)}(a)\alpha_{k+1}$. Pour cela nous introduisons tout dans la définition de la différentielle pour voir ce qui arrive.
Nous avons :

$$\begin{aligned}&\frac{d^k f_{(a+h_1,b+h_2)} - d^k f_{(a,b)} - g^{(k+1)}(a)\alpha_{k+1}(h_1,h_2)}{\|h\|}\\ &= \frac{g^{(k)}(a+h_1)\alpha_k - g^{(k)}(a)\alpha_k - g^{(k+1)}(a)h_1\alpha_k}{\|h\|}.\end{aligned} \tag{12.634}$$

Cela est, pour chaque $h \neq 0$, un élément V_k, mais le coefficient α_k se factorise de telle sorte que nous devons seulement calculer la limite (si elle existe)

$$\lim_{h\to 0} \frac{g^{(k)}(a+h_1) - g^{(k)}(a) - h_1 g^{(k+1)}(a)}{\|h\|}. \tag{12.635}$$

Le même jeu de séparation entre $h_1 = 0$ et $h_1 \neq 0$ que dans le cas $k = 1$ nous permet de déduire que cette limite existe et vaut zéro, grace à la définition de $g^{(k+1)}$.

Nous avons donc prouvé que f est différentiable autant que fois que souhaité. Elle est donc de classe C^∞ comme annoncé. □

12.22.4 Différentielle, dual et forme différentielle

12.22.4.1 Dans la base duale

Nous avons déjà parlé en (12.604) de la base $\{dx_i\}_{i=1,\dots,n}$ des formes différentielles sur $\mathbb{R}^n$.

Proposition 12.249.
La forme de base dx_i est la différentielle de la fonction de projection

$$\begin{aligned}\texttt{proj}_i \colon \mathbb{R}^n &\to \mathbb{R}\\ v &\mapsto v_i.\end{aligned} \tag{12.636}$$

Autrement dit nous avons

$$d(\texttt{proj}_i)_a = dx_i \tag{12.637}$$

pour tout i et pour tout a.

Démonstration. Le quotient

$$\frac{\texttt{proj}_i(a+h) - \texttt{proj}_i(a) - dx_i(h)}{\|h\|} \tag{12.638}$$

est toujours nul. La limite est a fortiori nulle. □

Nous avons donc $(d\,\mathtt{proj}_i)_a = dx_i$ pour tout a. Notons que les fonctions dx_i et $\mathtt{proj}_i$ sont les mêmes. Cela justifie la notation « dx_i » pour les formes différentielles de base, parce que ce sont les différentielles des fonctions « coordonnées » que nous pouvons noter x_i.

Étant donnée une fonction f, il est légitime de nous demander comment (si elle existe) la différentielle se décompose en chaque point dans la base duale. C'est-à-dire fixer les fonctions a_i en termes des dérivées de f pour avoir

$$df_a = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i. \tag{12.639}$$

C'est ce que nous allons faire dans le corolaire 12.254.

Exemple 12.250.
Si $F\colon \mathbb{R}^2 \to \mathbb{R}$ est une fonction C^2, sa différentielle est la forme

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy. \tag{12.640}$$

Si nous nommons f et g les fonctions $\partial_x F$ et $\partial_y F$, nous avons donc

$$Df = f dx + g dy, \tag{12.641}$$

qui vérifie

$$\partial_y f = \partial_x g, \tag{12.642}$$

parce que $\frac{\partial f}{\partial y} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial g}{\partial x}$. Ce que nous avons donc prouvé, c'est que △

Lemme 12.251.
Si $f dx + g dy$ est la différentielle d'une fonction de classe C^2 sur $\mathbb{R}^2$, alors $\partial_y f = \partial_x g$.

12.22.5 Ce n'est pas la différentielle extérieure

Il existe une notion de différentielle extérieure, mais ce n'est pas celle-là que nous utilisons la majorité du temps. En particulier si E et F sont des espaces vectoriels normés, lorsque $f\colon E \to F$ est une fonction, df est une application

$$df\colon E \to \mathcal{L}(E, F) \tag{12.643}$$

et la différentielle seconde est la différentielle de cette application-là. Chose faisable parce que $\mathcal{L}(E, F)$ est un espace vectoriel on ne peut plus respectable.

Soit $D \subset \mathbb{R}^n$. Par définition de la différentielle extérieure d'une 1-forme, nous avons une formule de Leibniz

$$d(f\omega) = df \wedge \omega + f d\omega. \tag{12.644}$$

En particulier,

$$d(f dx) = df \wedge dx + f \underbrace{d(dx)}_{=0} = \frac{\partial f}{\partial x} \underbrace{dx \wedge dx}_{=0} + \frac{\partial f}{\partial y} dy \wedge dx. \tag{12.645}$$

Attention : la différentielle extérieure n'est pas la différentielle usuelle. Certes dans le cas d'une 0-forme (c'est-à-dire d'une fonction), les deux notions coïncident, mais ça ne va pas plus loin. La différentielle extérieure vérifie $d^2\omega = 0$ pour tout ω, y compris pour les fonctions : si $\omega = df$ alors $d\omega = 0$.

Nous mentionnerons la différentielle extérieure dans le cas de

(1) Théorème de Stockes 20.71.

12.22.6 Continuité, dérivabilité et différentiabilité

Le théorème suivant reprend les principales propriétés d'une fonction différentiable. Il est à ne pas confondre avec le théorème 12.294 qui dira que si les dérivées partielles sont continues sur un voisinage de a, alors f est différentiable en a.

ThoRapPropDiffSi

Proposition 12.252.
Soit un espace vectoriel normé V et une fonction $f\colon \mathbb{R}^n \to V$. Si f est différentiable au point $a \in \mathbb{R}^n$ alors

(1) elle est continue en a,

(2) elle admet une dérivée dans toutes les directions de $\mathbb{R}^n$,

(3) toutes les dérivées directionnelles $\partial_u f(a)$ existent et nous avons l'égalité

$$\begin{aligned} df_a\colon \mathbb{R}^n &\to \mathbb{R}^m \\ u &\mapsto df_a(u) = \frac{\partial f}{\partial u}(a) = \sum_i \frac{\partial f}{\partial x_i}(a)u_i, \end{aligned} \tag{12.646}$$

EqDiffPartRap

si les u_i sont les composantes de u dans la base canonique $\mathbb{R}^n$.

La dernière égalité sera de temps en temps utilisée sous la forme

$$df_a(u) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0}. \tag{12.647}$$

EqOWQSoMA

Démonstration. La limite

$$\lim_{h\to 0_m} \frac{\|f(a+h) - f(a) - T(h)\|_n}{\|h\|_m} = 0,$$

implique que

$$\lim_{h\to 0_m} \|f(a+h) - f(a) - T(h)\|_n = 0.$$

Comme T est dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, on a $\lim_{h\to 0} T(h) = 0$, d'où la continuité de f au point a.

Si u est un vecteur non nul, la différentiabilité de f au point a implique

$$\lim_{t\to 0} \frac{\|f(a+tu) - f(a) - T(tu)\|_n}{\|tu\|_m} = 0,$$

par la linéarité de T et par l'égalité $\|tu\|_m = |t|\|u\|_m$ on obtient

$$\lim_{t\to 0} \frac{f(a+tu) - f(a)}{|t|} = T(u).$$

Donc f est dérivable suivant le vecteur u et $\partial_u f(a) = T(u) = df_a(u)$. □

CORooTBUMooHPncPH

Corolaire 12.253 (Différentielle et dérivée)**.**
Soit une application différentiable $f\colon \mathbb{R} \to V$ où V est un espace vectoriel normé. Alors $f'(u) = df_u(1)$.

Démonstration. En vertu de la proposition 12.252, nous avons

$$df_u(1) = \frac{d}{dt}\Big[f(u+t1)\Big]_{t=0} = f'(u). \tag{12.648}$$

Nous avons utilisé le fait que pour une fonction sur $\mathbb{R}$, l'unique dérivée partielle est la dérivée normale. □

CORooXURPooQMKvBl

Corolaire 12.254.
Si f est différentiable, alors la forme différentielle df_a se décompose en

$$df_a = \sum_i (\partial_i f)(a)dx_i. \tag{12.649}$$

Démonstration. Vue la définition des formes dx_i nous pouvons remplacer u_i par $dx_i(u)$ dans l'égalité (12.646) et écrire

$$df_a(u) = \sum_i (\partial_i f)(a) dx_i(u) \tag{12.650}$$

et donc écrire l'égalité demandée. □

Le lemme suivant regroupe quelques égalités avec lesquelles nous allons souvent travailler. Il explique comment sont liées les dérivées directionnelles, les dérivées partielles et la différentielle.

LemdfaSurLesPartielles

Lemme 12.255.
Si $f\colon \mathbb{R}^m \to \mathbb{R}^n$ est une fonction différentiable, alors

$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot u \tag{12.651}$$

pour tout vecteur $u \in \mathbb{R}^m$

Démonstration. La première égalité est la proposition 12.252, et la seconde est seulement la définition de la dérivée directionnelle avec des notations un peu plus snob. En particulier nous avons

$$df_a(e_i) = \frac{\partial f}{\partial x_i}(a). \tag{12.652}$$

Pour le reste c'est la linéarité de la différentielle qui joue : le vecteur u peut être écrit de façon unique comme combinaison linéaire des vecteurs de base

$$u = \sum_{i=1}^{m} u_i e_i, \qquad u_i \in \mathbb{R}, \forall i \in \{1, \ldots, m\}.$$

Alors, la linéarité de df_a nous donne

$$df_a(u) = df_a\left(\sum_{i=1}^{m} u_i e_i\right) = \sum_{i=1}^{m} u_i \left(df_a e_i\right) = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a). \tag{12.653}$$

Le lien avec le gradient est la définition du produit scalaire (8.173). □

La formule $df_a(u) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0}$ est bien utile pour calculer des différentielles, mais elle ne permet pas de prouver que f est différentiable. Autrement dit, même si le calcul de la dérivée $\frac{d}{dt}\Big[f(a+tu)\Big]_{t=0}$ donne un résultat pour tout u, nous ne pouvons pas en déduire que f est différentiable au point a.

PROPooJMBJooQzjXYf

Proposition 12.256 ([1])**.**
Si $f\colon \mathbb{R}^n \to \mathbb{R}^m$ est une application différentiable, alors sa matrice est donnée par

$$(df_a)_{ij} = df_a(e_j) \cdot e_i = (\partial_j f_i)(a). \tag{12.654}$$

Démonstration. En suivant la formule de la proposition 4.71(1), nous devons calculer

$$(df_a)_{ij} = \big(df_a(e_j)\big)_i. \tag{12.655}$$

En utilisant les formules du lemme 12.255, nous avons $df_a(e_i) = (\partial_i f)(a)$, et donc

$$(df_a)_{ij} = \big(df a_a(e_j)\big)_i = (\partial_j f)(a)_i = (\partial_j f_i)(a). \tag{12.656}$$

□

PropExistDiffUn

Proposition 12.257.
Soient f une fonction de x et y et un point $(a,b) \in \mathbb{R}^2$. Si les nombres $\partial_x f(a,b)$ et $\partial_y f(a,b)$ existent et si il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que

$$\begin{aligned} f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \\ + \|(x,y)-(a,b)\|\alpha\Big(\|(x,y)-(a,b)\|\Big) \end{aligned} \tag{12.657}$$

et

$$\lim_{t\to 0} \alpha(t) = 0, \tag{12.658}$$

alors f est différentiable en (a,b).

12.258.
Dans cet énoncé nous avons écrit $d\big((x,y),(a,b)\big)$ la distance entre (x,y) et (a,b), c'est-à-dire le nombre $\sqrt{(x-a)^2+(y-b)^2}$. Afin d'écrire l'équation (12.657) sous forme plus compacte, nous introduisons le vecteur

$$\nabla f(a,b) = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b). \end{pmatrix} \tag{12.659}$$

L'équation (12.657) devient alors

$$f(X) = f(P) + \nabla f(a,b) \cdot (X-P) + \|X-P\|\alpha\big(\|X-P\|\big). \tag{12.660}$$

Le vecteur $(\nabla f)(a,b)$ est appelé le **gradient** de f au point (a,b).

Remarque 12.259.
Nous avons introduit la notation ∇f pour le gradient d'une fonction f. Nous allons par la suite introduire $\nabla \cdot F$ pour la divergence du champ de vecteurs F et $\nabla \times F$ pour son rotationnel.

Toutes les formules pour ∇f, $\nabla \cdot F$ et $\nabla \times F$ peuvent facilement être mémorisées en pensant à ∇ comme étant le vecteur

$$\nabla = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}. \tag{12.661}$$

Nous allons ici cependant seulement penser à (12.661) comme un moyen mnémotechnique ; nous ne donnons pas de définition à « ∇ » tout seul.

PropExistDiffDeux

Proposition 12.260.
Soit f une fonction de deux variables admettant des dérivées partielles $\partial_x f(x,y)$ et $\partial_y f(x,y)$ qui sont elles-mêmes des fonctions continues de x et y. Alors la fonction f est différentiable partout.

Proposition 12.261.
Si f est différentiable en (a,b) alors pour tout vecteur u, la fonction

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(a+tu_1, b+tu_2) \end{aligned} \tag{12.662}$$

est dérivable en 0 et on a

$$\varphi'(0) = \nabla f(p) \cdot u \tag{12.663}$$

où nous avons noté $p = (a,b)$.

Démonstration. Réécrivons la formule (12.660) sous la forme

$$f(x) = f(p) + \nabla f(p) \cdot (x-p) + \|x-p\|\alpha(\|x-p\|). \tag{12.664}$$

Cela étant vrai pour tout x, nous l'écrivons en particulier pour $x = p + tu$ où t est un réel et u est le vecteur unitaire choisi. Nous avons donc

$$f(p + tu) = f(p) + t\nabla f(p) \cdot u + \|tu\|\alpha(\|tu\|). \tag{12.665}$$

En utilisant le fait que u est unitaire, $\|tu\| = |t|\|u\| = |t|$. La dérivée de φ en 0 est alors donnée par

$$\lim_{t\to 0} \frac{f(p + tu) - f(p)}{t} = \lim_{t\to 0} \nabla f(p) \cdot u + \alpha(|t|). \tag{12.666}$$

Lorsque nous prenons la limite, le membre de gauche devient $\varphi'(0)$ tandis que dans le membre de droite, le second terme disparaît. Nous avons finalement

$$\varphi'(0) = \nabla f(p) \cdot u \tag{12.667}$$

□

12.22.7 Calcul de valeurs approchées

Si nous remplaçons les accroissements $x - a$ et $y - b$ par h et k, le critère de différentiabilité s'écrit

$$\begin{aligned} f(a + h, b + k) = f(a, b) + \frac{\partial f}{\partial x}(a, b)h + \frac{\partial f}{\partial y}(a, b)k \\ + \sqrt{h^2 + k^2}\alpha\big(\sqrt{h^2 + k^2}\big). \end{aligned} \tag{12.668}$$

Le dernier terme du membre de droite tend vers zéro à une vitesse double lorsque h et k tendent vers zéro : d'une part parce que $\sqrt{h^2 + k^2}$ tend vers zéro et d'autre part parce que $\alpha\big(\sqrt{h^2 + k^2}\big)$ tend vers zéro. Nous avons donc la « bonne » approximation

$$f(x, y) \simeq f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b). \tag{12.669}$$ EqFormApproxfxyab

lorsque (x, y) n'est pas trop loin de (a, b). Cette expression est évidemment une généralisation immédiate de l'équation (12.395). Elle exprime que l'on peut obtenir des informations sur la valeur d'une fonction en (x, y) si on peut calculer la fonction et ses dérivées en un point (a, b) non loin de (x, y).

Cette formule peut aussi être vue sous la forme suivante, plus pratique dans certains calculs :

$$f(a + \Delta x, b + \Delta y) \simeq f(a, b) + \Delta x\frac{\partial f}{\partial x}(a, b) + \Delta y\frac{\partial f}{\partial y}(a, b). \tag{12.670}$$ EqFormApproxfxyabDF

Exemple 12.262.
Prenons la fonction $f(x, y) = \cos(x)\sin(y)$ et calculons une approximation de

$$f\big(\frac{\pi}{3} + 0.01, \frac{\pi}{2} + 0.03\big). \tag{12.671}$$

D'abord les dérivées partielles sont

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= -\sin(x)\sin(y) \\ \frac{\partial f}{\partial y}(x, y) &= \cos(x)\cos(y). \end{aligned} \tag{12.672}$$

Nous allons utiliser l'approximation

$$f\big(\frac{\pi}{3} + 0.01, \frac{\pi}{2} + 0.03\big) \simeq f\big(\frac{\pi}{3}, \frac{\pi}{2}\big) + 0.01\frac{\partial f}{\partial x}\big(\frac{\pi}{3}, \frac{\pi}{2}\big) + 0.03\frac{\partial f}{\partial y}\big(\frac{\pi}{3}, \frac{\pi}{2}\big). \tag{12.673}$$

Nous avons

$$\begin{aligned}\frac{\partial f}{\partial x}\big(\frac{\pi}{3},\frac{\pi}{2}\big) &= -\sin\frac{\pi}{3}\sin\frac{\pi}{2} = -\frac{\sqrt{3}}{2}\\ \frac{\partial f}{\partial y}\big(\frac{\pi}{3},\frac{\pi}{2}\big) &= \cos\frac{\pi}{3}\cos\frac{\pi}{2} = 0.\end{aligned} \tag{12.674}$$

Par conséquent

$$f\big(\frac{\pi}{3}+0.01,\frac{\pi}{2}+0.03\big) \simeq \frac{1}{2} - 0.01\frac{\sqrt{3}}{2} = \frac{1}{2} - \frac{\sqrt{3}}{200}. \tag{12.675}$$

```
sage: var('x,y')
(x, y)
sage: f(x,y)=cos(x)*sin(y)
sage: a=f(pi/3+0.01,pi/2+0.03)
sage: numerical_approx(a)
0.491093815387986
sage: b=1/2-sqrt(3)/200
sage: numerical_approx(b)
0.491339745962156
sage: numerical_approx(a-b)
-0.000245930574169814
```

Cela fait une erreur de l'ordre du dix millième.

△

Remarque 12.263.
Les esprits les plus critiques diront que cette vérification par Sage n'en est pas une parce que Sage a certainement utilisé un algorithme d'approximation qui se base sur la même idée que ce que nous venons de faire, et que par conséquent le fait qu'il obtienne le même résultat que nous est un peu tautologique.

Ils n'auront pas tort. Cependant, le code source de Sage est disponible publiquement [118] ; vous pouvez aller le lire et vérifier qu'il y a effectivement une *preuve* que le résultat fourni par Sage possède une bonne dizaine de décimales correctes.

Cette disponibilité publique du code source est une des nombreuses différences fondamentales entre Sage et votre calculatrice [119]. Dois-je vous rappeler qu'un des principes fondamentaux de l'éthique scientifique est que les résultats et les méthodes utilisés doivent être absolument ouverts à la vérification et à la critique de tous ?

12.22.8 Différentielle et tangente

La notion de dérivée partielle (ou de dérivée suivant un vecteur) pour une fonction de plusieurs variables n'est pas une généralisation de la notion de dérivée en une variable d'espace. En fait, du point de vue géométrique, la dérivée de la fonction $g : \mathbb{R} \to \mathbb{R}$ au point a est la pente de la ligne droite tangente au graphe de g au point $(a, g(a))$. Cette ligne, d'équation $r(x) = g'(a)x + g(a)$, est la meilleure approximation affine du graphe de g au point a, comme à la figure 12.9.

Le graphe d'une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ est une surface de deux paramètres dans $\mathbb{R}^3$. Si l'approximation affine d'une telle surface au point $(x, y, f(x, y))$ existe, alors elle est un plan tangent. En dimension plus haute, le graphe de la fonction $f : \mathbb{R}^m \to \mathbb{R}$ est une surface de m paramètres dans $\mathbb{R}^{m+1}$ et son approximation affine (si elle existe) est un hyperplan de $\mathbb{R}^m$.

Nous allons voir que si f prend ses valeurs dans $\mathbb{R}^n$ l'approximation affine de f au point a est l'élément de $f(a) + \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ qui ressemble le plus à f au voisinage de a. Plus précisément, on utilise les définitions suivantes.

118. Voir http://www.sagemath.org
119. et les autres logiciels de type fenêtre, pomme ou feuille d'érable.

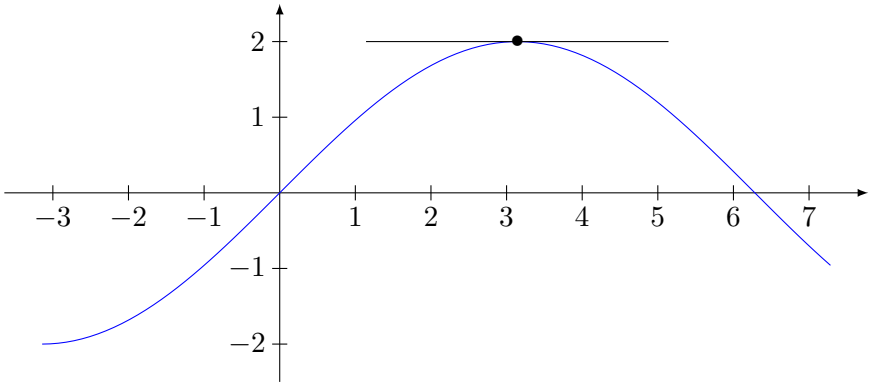


FIGURE 12.9: Tangentes au graphe d'une fonction d'une variable LabelFigTangentSegment

Définition 12.264.
*Soient f et g deux applications d'un ouvert U de $\mathbb{R}^m$ dans $\mathbb{R}^n$. On dit que g est **tangente** à f au point $a \in U$ si $f(a) = g(a)$ et*

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - g(x)\|_n}{\|x - a\|_m} = 0.$$

La relation de tangence est une relation d'équivalence. Nous sommes particulièrement intéressés par le cas où f admet une application affine tangente au point a.

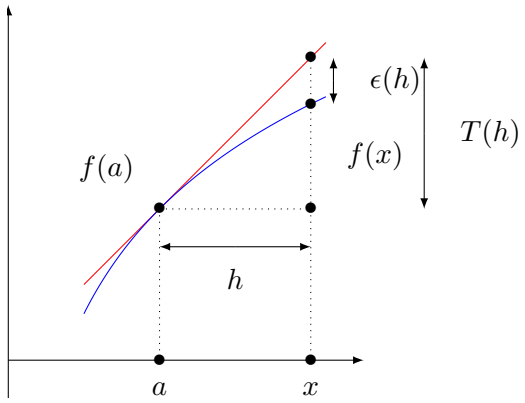


FIGURE 12.10: Interprétation géométrique de la différentielle. LabelFigDifferentielle

En ce qui concerne l'interprétation géométrique, si nous regardons la figure 12.10, et d'ailleurs aussi en voyant la définition 10.682, la fonction est différentiable et la différentielle est T si il existe une fonction α telle que

$$f(a+u) - f(a) - T(u) = \alpha(u) \tag{12.676}$$

où la fonction α satisfait

$$\lim_{u \to 0} \frac{\|\alpha(u)\|}{\|u\|} = 0 \tag{12.677}$$ EqPresqueTa

C'est cela qui fait écrire $f(a+u) - f(a) - df_a(u) = o(\|u\|)$ à ceux qui n'ont pas peur de la notation o.

La différentielle df_a est donc la partie linéaire de l'application affine qui approxime au mieux la fonction f autour du point a. La notion de différentielle est la vraie généralisation du concept de dérivée pour fonctions de plusieurs variables, en outre elle nous permet d'expliciter la relation qui associe au vecteur u la dérivée $\partial_u f(a)$, pour f et a fixés.

Remarque 12.265.
Si on remplace les normes $\|\cdot\|_m$ et $\|\cdot\|_n$ par d'autres normes, l'existence et la valeur de la différentielle de f au point a ne sont pas remises en cause. En effet, soient $\|\cdot\|_M$ une norme sur

$\mathbb{R}^m$ et $\|\cdot\|_N$ une norme sur $\mathbb{R}^n$. Par le théorème 10.49, ces normes sont équivalentes à $\|.\|_m$ et $\|.\|_n$ respectivement ; il existe donc des constantes k, K, l, $L > 0$ telles que pour tout vecteur u de $\mathbb{R}^m$ et tout vecteur v de $\mathbb{R}^n$

$$k\|u\|_M \leqslant \|u\|_m \leqslant K\|u\|_M,$$
$$l\|v\|_N \leqslant \|v\|_n \leqslant L\|v\|_N.$$

Les éléments de $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ sont les mêmes et on a

$$\begin{aligned}\frac{l}{K}\frac{\|f(a+h)-f(a)-T(h)\|_N}{\|h\|_M} \leqslant \frac{\|f(a+h)-f(a)-T(h)\|_n}{\|h\|_m} \leqslant \\ \leqslant \frac{L}{k}\frac{\|f(a+h)-f(a)-T(h)\|_N}{\|h\|_M}.\end{aligned} \tag{12.678}$$

Il est donc possible, pour démontrer la différentiabilité ou pour calculer la différentielle, d'utiliser le critère (10.682) avec une norme au choix. Parfois c'est utile.

12.22.9 Prouver qu'une fonction n'est pas différentiable

Chacun des points du théorème 12.252 est en soi un critère pour montrer qu'une fonction n'est pas différentiable en un point.

12.22.9.1 Continuité

Le premier critère à vérifier est donc la continuité. Si une fonction n'est pas continue en un point, alors elle n'y sera pas différentiable. Pour rappel, la continuité en a se teste en vérifiant si $\lim_{x\to a} f(x) = f(a)$.

12.22.9.2 Linéarité

Un second test est la linéarité de la dérivée directionnelle par rapport à la direction : l'application $u \mapsto \frac{\partial f}{\partial u}(a)$ doit être linéaire, sinon df_a n'existe pas.

Exemple0046Diff

Exemple 12.266.
Examinons la fonction

$$\begin{aligned}f\colon \mathbb{R}^2 &\to \mathbb{R}\\ (x,y) &\mapsto \begin{cases}\dfrac{xy^2}{x^2+y^4} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases}\end{aligned} \tag{12.679}$$

Prenons $u = (u_1, u_2)$ et calculons la dérivée de f dans la direction de u au point $(0,0)$:

$$\begin{aligned}\frac{\partial f}{\partial u}(0,0) &= \lim_{t\to 0}\frac{f(tu_1,tu_2)-f(0,0)}{t}\\ &= \lim_{t\to 0}\frac{1}{t}\left(\frac{tu_1t^2u_2^2}{t^2u_1^2+t^4u_2^4}\right)\\ &= \lim_{t\to 0}\left(\frac{u_1u_2^2}{u_1^2+t^2u_2^4}\right)\\ &= \begin{cases}\dfrac{u_2^2}{u_1} & \text{si } u_1\neq 0\\ 0 & \text{si } u_1 = 0.\end{cases}\end{aligned} \tag{12.680}$$

Cette application n'est pas linéaire par rapport à u. En effet, notons

$$\begin{aligned}A\colon \mathbb{R}^n &\to \mathbb{R}\\ u &\mapsto \frac{\partial f}{\partial u}(0,0),\end{aligned} \tag{12.681}$$

et vérifions que pour tout u et v dans $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$, nous ayons $A(\lambda u) = \lambda A(u)$ et $A(u+v) = A(u) + A(v)$. La première égalité est vraie, parce que

$$A(\lambda u) = A(\lambda u_1, \lambda u_2) = \frac{\lambda^2 u_2^2}{\lambda u_1} = \lambda \frac{u_2^2}{u_1} = \lambda A(u). \tag{12.682}$$

Mais nous avons par exemple

$$A\big((0,1) + (2,3)\big) = A(2,4) = \frac{16}{2} = 8, \tag{12.683}$$

tandis que

$$A(0,1) + A(2,3) = 0 + \frac{9}{2} \neq 8. \tag{12.684}$$

La fonction f n'est donc pas différentiable en $(0,0)$, parce que la candidate différentielle, $df_{(0,0)}(u) = \frac{\partial f}{\partial u}(0,0)$, n'est même pas linéaire.

△

Voici une autre façon de traiter la fonction de l'exemple 12.266.

ExeFHmCLII

Exemple 12.267.
La figure 12.11 représente le domaine d'une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$, et sur chacune des parties, elle est définie différemment.

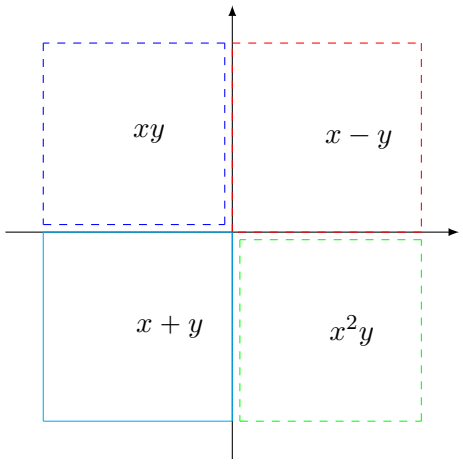


FIGURE 12.11: La fonction de l'exemple 12.267.

LabelFigFWJuNhU

L'expression de f est ici

$$f(x,y) = \begin{cases} xy & \text{si } x < 0 \text{ et } y > 0 \\ x - y & \text{si } x \geqslant 0 \text{ et } y \geqslant 0 \\ x^2 y & \text{si } x > 0 \text{ et } y < 0 \\ x + y & \text{sinon.} \end{cases} \tag{12.685}$$

On note que les deux axes forment une zone à problèmes. La zone hors des axes est un ouvert sur lequel f est différentiable car composée de polynômes. Analysons chacun des points de la forme (a,b) dans la zone à problèmes (c'est-à-dire si $ab = 0$).

(1) **Si $a = 0$ et $b > 0$** Un tel point $(0,b)$ est sur l'axe vertical, dans la moitié supérieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y > 0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} x - y = 0 - b = -b$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} xy = 0b = 0$$

qui sont différentes puisque b est supposé non nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b>0$.

(2) **Si $a=0$ et $b<0$** Un tel point $(0,b)$ est sur l'axe vertical, dans la moitié inférieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} x^2y = 0^2b = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} x+y = 0+b = b$$

qui sont différentes puisque b est supposé non nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b<0$.

(3) **Si $a>0$ et $b=0$** Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié droite. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x>0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} x-y = a-0 = a$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} x^2y = a^20 = 0$$

qui sont différentes puisque a est supposé non nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a>0$.

(4) **Si $a<0$ et $b=0$** Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié gauche. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y>0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y>0}} xy = a0 = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y\leqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y\leqslant 0}} x+y = a+0 = a$$

qui sont différentes puisque a est supposé non nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a<0$.

(5) **Si $a=0$ et $b=0$** Le cas du point $(0,0)$ est particulier, puisque il est adhérent aux quatre composantes du domaine où la fonction est définie différemment. Pour étudier la continuité, il faut donc étudier quatre limites. Ces limites ont déjà été étudiées ci-dessus et valent toutes 0, ce qui prouve la continuité de f en $(0,0)$.

En ce qui concerne la différentiabilité, on sait qu'il est nécessaire que toutes les dérivées directionnelles existent. Calculons la dérivée dans la direction $(0,1)$ (au point $(0,0)$) :

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f((0,0)+t(0,1))-f(0,0)}{t} = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(0,t)}{t} = \dots$$

qu'on sépare en deux cas, car $f(0,t)$ possède une formule différente si $t<0$ ou si $t\geqslant 0$:

$$\lim_{\substack{t\to 0\\ t\neq 0}}\frac{f(0,t)}{t}=\begin{cases}\lim_{\substack{t\to 0\\ t<0}}\frac{f(0,t)}{t}=\lim_{\substack{t\to 0\\ t<0}}\frac{0+t}{t}=1\\ \lim_{\substack{t\to 0\\ t\geqslant 0}}\frac{f(0,t)}{t}=\lim_{\substack{t\to 0\\ t\geqslant 0}}\frac{0-t}{t}=-1\end{cases}$$

ce qui prouve que la limite n'existe pas, donc que la dérivée directionnelle n'existe pas, et finalement que la fonction n'est pas différentiable.

Conclusion : La fonction donnée est continue hors des axes et au point $(0,0)$, mais discontinue partout ailleurs sur les axes. Elle est différentiable hors des axes, mais ne l'est pas sur les axes.

△

12.22.9.3 Cohérence des dérivées partielles et directionnelle

Dans la pratique, nous pouvons calculer $\partial_u f(a)$ pour une direction u générale, et puis en déduire $\partial_x f$ et $\partial_y f$ comme cas particuliers en posant $u=(1,0)$ et $u=(0,1)$. Une chose incroyable, mais pourtant possible est qu'il peut arriver que

$$\frac{\partial f}{\partial u}(a)\neq\sum_i\frac{\partial f}{\partial x_i}(a)u_i. \tag{12.686}$$

Ceci se produit lorsque f n'est pas différentiable en a.

12.22.9.4 Un candidat dans la définition (marche toujours)

Lorsqu'une fonction est donnée, un candidat différentielle au point (a_1,a_2) est souvent assez simple à trouver en un point :

$$T(u_1,u_2)=\frac{\partial f}{\partial x}(a_1,a_2)u_1+\frac{\partial f}{\partial y}(a_1,a_2)u_2. \tag{12.687}$$

L'application T est la candidate différentielle en ce sens que si la différentielle existe, alors elle est égale à T. Ensuite, il faut vérifier si

$$\lim_{(x,y)\to(a_1,a_2)}\frac{f(x,y)-f(a_1,a_2)-T\big((x,y)-(a_1,a_2)\big)}{\|(x,y)-(a_1,a_2)\|}=0 \tag{12.688}$$

EqLimDefDiff

ou non. Si oui, alors la différentielle existe et $df_{(a,b)}(u)=T(u)$, sinon [120], la différentielle n'existe pas.

Attention : dans la ZAP, les dérivées partielles $\partial_x f$ et $\partial_y f$ ne peuvent en général pas être calculées en utilisant les règles de calcul (c'est bien pour ça que la ZAP est une zone à problèmes). Il faut d'office utiliser la définition

$$\frac{\partial f}{\partial x}(a_1,a_2)=\lim_{t\to 0}\frac{f(a_1+t,a_2)-f(a_1,a_2)}{t}, \tag{12.689}$$

et la définition correspondante pour $\partial_y f$.

Conclusion

Soient $f:A\subset\mathbb{R}^n\to\mathbb{R}^m$, et $a\in\mathrm{Int}(A)$. Si f est différentiable en a,

$$(df_a(e_j))_i=d(f_i)_a(e_j)=\frac{\partial f_i}{\partial x_j}(a)=[Jac(f)_{|a}]_{ij} \tag{12.690}$$

et la matrice de l'application linéaire df_a est la matrice jacobienne $m\times n$ de f en a notée $Jac(f)_{|a}$.

120. y compris si la limite (12.688) n'existe même pas.

12.22.10 Gradient

Définition 12.268.
Soit f une fonction différentiable de $\mathbb{R}^m$ dans $\mathbb{R}$. On appelle ***gradient*** *de f la fonction $\nabla f : \mathbb{R}^m \to \mathbb{R}^m$ de composantes*

$$\partial_1 f, \dots, \partial_m f.$$

Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$, $f(a) = (f_1(a), \dots, f_n(a))^T$. On appelle ***matrice jacobienne*** *de f la fonction $J(f) : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^n$ définie par*

$$a \mapsto \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_m f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(a) & \dots & \partial_m f_n(a) \end{pmatrix} \tag{12.691}$$

12.22.11 Linéarité

La proposition suivante signifie que la différentiation est une opération linéaire sur l'ensemble des fonctions différentiables.

PropDiffLineaire

Proposition 12.269.
Soient f et g deux fonctions de $U \subset \mathbb{R}^m$ dans $\mathbb{R}^n$ différentiables au point $a \in U$, et soit λ dans $\mathbb{R}$. Alors les fonctions $f + g$ et λf sont différentiables au point a et on a

$$\begin{aligned} d(f+g)(a) &= df(a) + dg(a), \\ d(\lambda f)(a) &= \lambda df(a), \end{aligned} \tag{12.692}$$

Démonstration.

$$\begin{aligned} &\lim_{h \to 0_m} \frac{\|(f(a+h) + g(a+h)) - (f(a) + g(a)) - df(a).h - dg(a).h\|_n}{\|h\|_m} \leqslant \\ &\lim_{h \to 0_m} \frac{\|f(a+h) - f(a) - df(a).h\|_n}{\|h\|_m} + \lim_{h \to 0_m} \frac{\|g(a+h) - g(a) - dg(a).h\|_n}{\|h\|_m} = 0. \end{aligned} \tag{12.693}$$

De même on démontre la propriété $d(\lambda f)(a) = \lambda df(a)$. □

12.23 Produit

Soient f et g deux fonctions de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Nous notons $f \cdot g$ la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ donnée par le produit scalaire point par point, c'est-à-dire

$$(f \cdot g)(x) = f(x) \cdot g(x) \tag{12.694}$$

pour tout $x \in \mathbb{R}^m$. Le point dans le membre de droite est le produit scalaire dans $\mathbb{R}^n$. Le cas particulier $n = 1$ revient au produit usuel de fonctions :

$$(fg)(x) = f(x)g(x). \tag{12.695}$$

LemDiffProsuid

Lemme 12.270.
Si f et g sont des fonctions différentiables sur $\mathbb{R}^m$ à valeurs dans $\mathbb{R}$, alors la fonction produit fg est également différentiable et

$$(dfg)_a = g(a)df_a + f(a)dg_a \tag{12.696}$$

EqDifffgProd

au sens où pour chaque u dans $\mathbb{R}^m$,

$$(dfg)_a(u) = g(a)df_a(u) + f(a)dg_a(u). \tag{12.697}$$

Démonstration. Ce que nous devons faire pour vérifier la formule 12.696, c'est de vérifier le critère (10.682) en remplaçant f par fg et $T(h)$ par $g(a)df(a).h + f(a)dg(a).h$.

Ce que nous avons au numérateur est

$$\begin{aligned}\clubsuit &= (fg)(a+h) - (fg)(a) - g(a)df(a).h - f(a)dg(a).h \\ &= f(a+h)g(a+h) - f(a)g(a) - g(a)df(a).h - f(a)dg(a).h.\end{aligned} \tag{12.698}$$

Maintenant, nous allons faire apparaitre $\big(f(a+h)-f(a)-df(a)\big)g(a+h)$ en ajoutant et soustrayant ce qu'il faut pour conserver $\clubsuit$:

$$\begin{aligned}\clubsuit = &\big(f(a+h) - f(a) - df(a).h\big)g(a+h) \\ &+ f(a)g(a+h) + g(a+h)df(a).h \\ &- f(a)g(a) - g(a)df(a).h - f(a)dg(a).h.\end{aligned} \tag{12.699}$$

Nous mettons maintenant $f(a)$ et $df(a).h$ en évidence là où c'est possible :

$$\begin{aligned}\clubsuit = &\big(f(a+h) - f(a) - df(a).h\big)g(a+h) \\ &+ f(a)\big(g(a+h) - g(a) - dg(a).h\big) \\ &+ \big(g(a+h) - g(a)\big)df(a).h.\end{aligned} \tag{12.700}$$

Nous devons maintenant considérer la limite

$$\lim_{h\to 0} \frac{\|\clubsuit\|}{\|h\|}. \tag{12.701}$$

Étant donné que f et g sont différentiables, les deux premiers termes sont nuls :

$$\begin{aligned}\lim_{h\to 0} \frac{\big(f(a+h) - f(a) - df(a).h\big)}{\|h\|} g(a+h) = 0 \\ \lim_{h\to 0} f(a) \frac{\big(g(a+h) - g(a) - dg(a).h\big)}{\|h\|} = 0.\end{aligned} \tag{12.702}$$

En ce qui concerne le troisième terme, en utilisant la norme d'une application linéaire, nous avons

$$\lim_{h\to 0} \frac{\|df(a).h\|}{\|h\|} \leqslant \sup_{h\in\mathbb{R}^m} \frac{\|df(a).h\|}{\|h\|} = \|df(a)\|, \tag{12.703}$$

et par conséquent

$$\begin{aligned}0 &\leqslant \lim_{h\to 0} \|g(a+h) - g(a)\| \frac{\|df(a).h\|\|h\|}{\|h\|} \\ &\leqslant \lim_{h\to 0} \|g(a+h) - g(a)\|\|df(a)\| = 0\end{aligned} \tag{12.704}$$

parce que g est continue (la limite du premier facteur est nulle tandis que la norme de $df(a)$ est un nombre constant). Nous avons donc bien prouvé que la formule (12.696) est la différentielle de fg au point a. □

Ce résultat se généralise pour des fonctions f et g de $\mathbb{R}^m$ dans $\mathbb{R}^n$ dans la proposition suivante qui généralise tout en même temps la proposition 12.177.

Proposition 12.271.
Soient f et g deux fonctions de $U \subset \mathbb{R}^m$ dans $\mathbb{R}^n$ différentiables au point $a \in U$. Alors la fonction $f \cdot g$ est différentiable au point a et on a

$$d(f \cdot g)(a) = g(a) \cdot df(a) + f(a) \cdot dg(a) \tag{12.705}$$

au sens où

$$d(f \cdot g)_a(u) = g(a) \cdot \big(df_a(u)\big) + f(a) \cdot \big(dg_a(u)\big) \tag{12.706}$$

Eqdfdotgexpl

pour tout $u \in \mathbb{R}^m$.

Démonstration. La preuve du cas $n = 1$ est déjà faite ; c'est la formule (12.696). Pour le cas général $n \geqslant 2$, nous passons aux composantes en nous rappelant que

$$(f \cdot g)(a) = \sum_{i=1}^{n} f_i(a)g_i(a) = \sum_{i=1}^{n}(f_i g_i)(a). \tag{12.707}$$

En utilisant la linéarité de la différentiation, nous nous réduisons donc au cas des produits $f_i g_i$ qui sont des fonctions de $\mathbb{R}^m$ dans $\mathbb{R}$:

$$\begin{aligned} d(f \cdot g)(a) &= d\left(\sum_{i=1}^{n} f_i g_i\right)(a) \\ &= \sum_{i=1}^{n} \big(df_i(a)g_i(a) + f_i(a)dg_i(a)\big) \\ &= g(a) \cdot df(a) + f(a) \cdot dg(a). \end{aligned} \tag{12.708}$$

Ceci termine la preuve. □

12.23.1 Difficulté d'ordre supérieur

12.272.
Il serait tentant de faire une récurrence sur le lemme 12.270 pour dire que si f et g sont de classe C^p, alors le produit fg est également de classe C^p, parce que la formule de $d(fg)$ contient des produits de fonctions de classe C^p et C^{p-1}.

Le problème est que le lemme 12.270 est énoncé et prouvé pour des fonctions à valeurs dans $\mathbb{R}$, alors que déjà la formule

$$d(fg) = gdf + fdg \tag{12.709}$$

contient le produit de $g\colon E \to \mathbb{R}$ par $df\colon E \to \mathcal{L}(E, \mathbb{R})$. Lorsque nous montons dans les différentielles, la situation empire, et les produits dont sont composés les formules sont réellement à définir...

Oublions un instant les questions de régularité, et calculons sans ménagement, pour voir ce qu'il se passe. Nous considérons un espace vectoriel E ainsi que des fonctions $f\colon E \to \mathbb{R}$ et $g\colon E \to V$ où V est un autre espace vectoriel.

Nous avons

$$d(fg)_a(u) = df_a(u)g(a) + f(a)dg_a(u). \tag{12.710}$$

Les deux termes sont des produits $\mathbb{R} \times V \to \mathbb{R}$. Montons un coup :

$$d(gdf)_a(u) = \frac{d}{dt}\Big[(gdf)(a+tu)\Big]_{t=0} = \frac{d}{dt}\Big[g(a+tu)df_{a+tu}\Big]_{t=0} = dg_a(u)df_a + g(a)(d^2f)_a(u). \tag{12.711}$$

Un autre pour voir comment ça se passe plus haut :

$$d(dfdg)_a(u) = \frac{d}{dt}\Big[(dfdg)(a+tu)\Big]_{t=0} = \frac{d}{dt}\Big[df_{a+tu}dg_{a+tu}\Big]_{t=0} = (d^2f)_a(u)dg_a + df_a(d^2g)_a(u). \tag{12.712}$$

Là déjà vous noterez que nous sommes passés par le produit

$$df_{a+tu}df_{a+tu} \tag{12.713}$$

qui pour chaque t est un produit $\mathcal{L}(E, \mathbb{R}) \times \mathcal{L}(E, V)$ que nous n'avons pas réellement défini.

En continuant le calcul ainsi nous trouvons par exemple

$$\begin{aligned} (d^3fg)_a(u) = {} & d^3f_a(u)g(a) + d^2f_a dg_a(u) \\ & + d^2f_a(u)dg_a + df_a d^2g_a(u) \\ & + d^2f_a(u)dg_a + df_a(d^2g)_a(u) \\ & + df_a(u)d^2g_a + f(a)(d^3g)_a(u). \end{aligned} \tag{12.714}$$

Vous noterez que cette formule contient trois termes que nous aurions eu envie de noter d^2fdg. Or ces trois termes ne sont pas identiques : deux sont $d^2f_a(u)dg_a$ et un est $(d^2f)_a dg_a(u)$.

12.23.2 Solution : produit tensoriel

Afin de donner un sens à tous les produits, nous allons passer par les produits tensoriels. Nous avons déjà le théorème 10.290 qui fait pratiquement tout.

PROPooWNCGooHbmcVb

Proposition 12.273 ([1]).
Soient des fonctions $f\colon \mathbb{R}^n \to \mathbb{R}$ et $g\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^p. Alors fg est de classe C^p.

Démonstration. Nous considérons l'application

$$\begin{aligned}\varphi\colon \mathbb{R}\otimes\mathbb{R} &\to \mathbb{R}\\ 1\otimes 1 &\mapsto 1\end{aligned} \tag{12.715}$$

dont nous avons déjà parlé dans le lemme 10.235. En utilisant la notation $\tilde{\otimes}$ de la définition 10.286, nous avons

$$fg = \varphi\circ(f\tilde{\otimes}g). \tag{12.716}$$

La proposition 10.290 nous dit que $f\tilde{\otimes}g\colon \mathbb{R}^n \to \mathbb{R}\otimes\mathbb{R}$ est de classe C^p. Vu que φ est un isomorphisme d'espaces vectoriels, la proposition 10.278 nous dit que $\varphi\circ(f\tilde{\otimes}g)$ est encore de classe C^p.

Et voilà. □

12.23.3 Formes bilinéaires

Nous avons aussi une formule importante pour la différentielle des formes bilinéaires.

bilin_diff

Lemme 12.274.
Toute application bilinéaire

$$\begin{aligned}B\colon \mathbb{R}^m\times\mathbb{R}^n &\to \mathbb{R}^p\\ B(a_1,a_2) &= a_1\star a_2\end{aligned} \tag{12.717}$$

est différentiable en tout point (a_1,a_2) de $\mathbb{R}^m\times\mathbb{R}^n$, et on a

$$dB(a_1,a_2).(h_1,h_2) = h_1\star a_2 + a_1\star h_2.$$

Démonstration.

$$\begin{aligned}&\frac{\|B(a_1+h_1,a_2+h_2)-B(a_1,a_2)-(h_1\star a_2+a_1\star h_2)\|_p}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} =\\ &= \frac{\|(a_1+h_1)\star(a_2+h_2)-a_1\star a_2-(h_1\star a_2+a_1\star h_2)\|_p}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = \spadesuit\end{aligned} \tag{12.718}$$

on rajoute et on enlève la quantité $(a_1+h_1)\star a_2$ dans le numérateur, et on obtient

$$\begin{aligned}\spadesuit &= \frac{\|(a_1+h_1)\star h_2 + h_1\star a_2 - (h_1\star a_2+a_1\star h_2)\|_p}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} =\\ &= \frac{\|h_1\star h_2\|_p}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} \leqslant C\frac{\|h_1\|_m\|h_2\|_n}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}}\\ &\leqslant C\frac{\|(h_1,h_2)\|^2_{\mathbb{R}^m\times\mathbb{R}^n}}{\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}} = C\|(h_1,h_2)\|_{\mathbb{R}^m\times\mathbb{R}^n}.\end{aligned} \tag{12.719}$$

Si on prend la limite de cette expression pour $(h_1,h_2)\to(0_m,0_n)$ on obtient 0, donc la preuve est complète. À noter, que dans l'avant-dernier passage on a utilisé la continuité des applications linéaires $\texttt{proj}_m : \mathbb{R}^m\times\mathbb{R}^n \to \mathbb{R}^m$ et $\texttt{proj}_n : \mathbb{R}^m\times\mathbb{R}^n \to \mathbb{R}^n$ qui à chaque point (a_1,a_2) de $\mathbb{R}^m\times\mathbb{R}^n$ associent a_1 et a_2 respectivement. □

PropEKLTooSvZjdW

Proposition 12.275.
Soit V et W deux espaces vectoriels et $\varphi\colon V\to W$ un isomorphisme. Soit $f\colon \mathbb{C}\to V$ une application telle que $\varphi\circ f\colon \mathbb{C}\to W$ soit différentiable.

Alors f est différentiable et $df = \varphi^{-1}\circ d(\varphi\circ f)$.

Démonstration. Si T est la différentielle de $\varphi \circ f$ au point z nous avons

$$0 = \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{(\varphi\circ f)(z+h)-(\varphi\circ f)(z)+T(h)}{h} \tag{12.720a}$$

$$= \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{(\varphi\circ f)(z+h)-(\varphi\circ f)(z)+\varphi\big((\varphi\circ T)(h)\big)}{h} \tag{12.720b}$$

$$= \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \varphi\left(\frac{f(z+h)-f(z)+(\varphi\circ T)(h)}{h}\right) \tag{12.720c}$$

$$= \varphi \lim_{h\to 0} \frac{f(z+h)-f(z)+\varphi^{-1}T(h)}{h}. \tag{12.720d}$$

Pour la dernière ligne, nous avons permuté φ avec la limite (parce que φ est continue). Nous avons prouvé que f est différentiable et que $df = \varphi^{-1}\circ T = \varphi^{-1}\circ d(\varphi\circ f)$. □

12.24 Différentielle de fonction composée

Une importante règle de différentiation est la règle de différentiation d'une fonction composée (*chain rule* dans les livres anglais et américains). Cette règle généralise la règle de dérivation pour fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Cette règle a déjà été donnée dans le théorème 10.276, mais si vous avez seulement envie d'entendre parler de $\mathbb{R}^n$, vous pouvez lire le lemme 12.276 suivi de la proposition 12.278.

Le lemme suivant est essentiellement une reformulation du lemme 10.271.

Def_diff2

Lemme 12.276.
Soit U un ouvert de $\mathbb{R}^m$. La fonction $f : U \to \mathbb{R}^n$ est différentiable au point a dans U, si et seulement si il existe une fonction $\sigma_f : U\times U \to \mathbb{R}^n$ telle que SubEqsDiff2

$$\sigma_f(a,a) = \lim_{x\to a}\sigma_f(a,x) = 0 \tag{12.721a}$$

$$f(x) = f(a) + T(x-a) + \sigma_f(a,x)\|x-a\|_m, \tag{12.721b}$$

def_diff2

pour une certaine application linéaire $T \in \mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$.

Démonstration. Si les conditions (12.721) sont satisfaites alors T est la différentielle de f en a. En effet, dans ce cas nous avons

$$f(a+h) = f(a) + T(h) + \sigma_f(a,a+h)\|h\|, \tag{12.722}$$

et la condition (10.682) devient

$$\lim_{h\to 0}\frac{\|\sigma_f(a,a+h)\|\|h\|}{\|h\|} = \lim_{h\to 0}\|\sigma_f(a,a+h)\| = 0 \tag{12.723}$$

Si f est différentiable au point a il suffit de prendre $T = df(a)$ et

$$\sigma_f(a,x) = \frac{f(x)-f(a)-df(a).(x-a)}{\|x-a\|_m}.$$

□

Remarque 12.277.
La fonction $\sigma_f(a,x)\|x-a\|_m$ est ce qui avait été appelé $\epsilon(h)$ sur la figure 12.10.

PropDiffCompose

Proposition 12.278.
Soient U un ouvert de $\mathbb{R}^m$ et V un ouvert de $\mathbb{R}^n$. Soient $f : U \to V$ et $g : V \to \mathbb{R}^p$ deux fonctions différentiables respectivement au point a dans U et $b = f(a)$ dans V. Alors la fonction composée $g\circ f : U \to \mathbb{R}^p$ est différentiable au point a et

$$d(g\circ f)_a = dg_{f(a)}\circ df_a. \tag{12.724}$$

EqDiffCompose

Démonstration. En tenant compte du lemme 12.276 on peut écrire

$$f(a+h)-f(a)=df_a(h)+\sigma_f(a,a+h)\|h\|_m, \qquad \forall h\in U\backslash\{0\}, \tag{12.725a}$$
$$g(b+k)-g(b)=dg_b(k)+\sigma_g(b,b+k)\|k\|_n, \qquad \forall k\in V\backslash\{0\}. \tag{12.725b}$$

On sait que $f(a)=b$ et que $f(a+h)$ est un élément de V et $f(a+h)=f(a)+k$ pour $k=df(a).h+\sigma_f(a,a+h)\|h\|_m$. Par substitution dans la deuxième équation on obtient

$$\begin{aligned}
&g\big(f(a+h)\big)-g\big(f(a)\big)\\
&\quad=dg_{f(a)}\Big(df_a(h)+\sigma_f(a,a+h)\|h\|_m\Big)\\
&\qquad+\sigma_g\left(f(a),f(a+h)\right)\left\|df_a(h)+\sigma_f(a,a+h)\|h\|_m\right\|_n\\
&\quad=g\circ f(a+h)-g\circ f(a)\\
&\quad=dg_{f(a)}\circ df_a(h)\\
&\qquad+\|h\|_m\Big[dg_{f(a)}\sigma_f(a,a+h)\\
&\qquad+\sigma_g\left(f(a),f(a+h)\right)\left\|df_a\frac{h}{\|h\|_m}+\sigma_f(a,a+h)\right\|_n\Big],
\end{aligned} \tag{12.726}$$

donc

$$(g\circ f)(a+h)-(g\circ f)(a)=dg_f(a)\circ df_a(h)+S(a,a+h)\|h\|_m \tag{12.727}$$

où S représente le contenu du dernier grand crochet. Il ne reste plus qu'à prouver que $S(a,a+h)$ est $o(\|h\|_m)$. En tenant compte du fait que $\sigma_f(a,a+h)$ et $\sigma_g\left(f(a),f(a+h)\right)$ sont $o(\|h\|_m)$,

$$\begin{aligned}
&\lim_{h\to 0_m}\frac{S(a,a+h)}{\|h\|_m}=\lim_{h\to 0_m}\frac{dg_{f(a)}\sigma_f(a,a+h)}{\|h\|_m}+\\
&+\lim_{h\to 0_m}\frac{\sigma_g\left(f(a),f(a+h)\right)\left\|df_a\frac{h}{\|h\|_m}+\sigma_f(a,a+h)\right\|_n}{\|h\|_m}=0.
\end{aligned} \tag{12.728}$$

□

Remarque 12.279.
Note : la formule (12.724) est à comprendre de la façon suivante. Si $u\in\mathbb{R}^m$, alors

$$d(g\circ f)_a(u)=\underbrace{dg_{f(a)}}_{\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}\Big(\underbrace{df_a(u)}_{\in\mathbb{R}^n}\Big)\in\mathbb{R}^p. \tag{12.729}$$

Le lemme suivant sert à prouver les théorèmes 15.9 et 20.72. Il est fondamentalement la raison de la formule définissant l'intégrale d'une forme sur un chemin (définition 20.47).

LEMooKNBVooQSowos

Lemme 12.280 ([1]).
Soient un espace vectoriel normé F de dimension finie, ainsi que E, un espace vectoriel normé. Nous considérons un chemin de classe C^1

$$\gamma\colon[a,b]\to E \tag{12.730}$$

et une application de classe C^1

$$f\colon E\to F. \tag{12.731}$$

Si $g=f\circ\gamma$, alors

$$g'(t)=(df)_{\gamma(t)}\big(\gamma'(t)\big) \tag{12.732}$$

pour tout $t\in[a,b]$.

Démonstration. Nous écrivons la dérivée de g de la façon suivante : SUBEQSooWKRYooGyVgNl

$$\begin{aligned} g'(t) &= \frac{d}{ds}\Big[g(t+s)\Big]_{s=0} && \text{(12.733a)}\\ &= \frac{d}{ds}\Big[(f\circ\gamma)(t+s)\Big]_{s=0} && \text{(12.733b)}\\ &= d(f\circ\gamma)_t(1) && \text{(12.733c)}\\ &= (df)_{\gamma(t)}\big(d\gamma_t(1)\big) && \text{(12.733d)}\\ & && \text{(12.733e)} \end{aligned}$$

SUBEQooHUZVooWECnVZ SUBEQooOAYLooIzltaY

Justifications :

— Pour (12.733c) : les formules (12.255). Notez que le 1 à qui s'applique la différentielle de $f\circ\gamma$ est le vecteur de $\mathbb{R}$ qui est multiplié par s dans l'expression $(f\circ\gamma)(t+s)$.

— Pour (12.733d) : la différentiation de fonctions composées de la proposition 12.278.

Mais

$$d\gamma_t(1) = \frac{d}{ds}\Big[\gamma(t+s)\Big]_{s=0} = \gamma'(t). \tag{12.734}$$

En remettant au bout de (12.733), nous obtenons le résultat. □

12.24.1 Fonctions composées

Cette façon de voir la différentielle nous permet de jeter un nouveau regard sur la formule de différentiation des fonctions composées. Soient

$$\begin{aligned} f&\colon \mathbb{R}^p \to \mathbb{R}^n\\ g&\colon \mathbb{R}^n \to \mathbb{R}, \end{aligned} \tag{12.735}$$

et $h\colon \mathbb{R}^p \to \mathbb{R}$ définie par

$$h(u) = g\big(f(u)\big) = (g\circ f)(u). \tag{12.736}$$

Nous allons noter x les coordonnées de $\mathbb{R}^p$, a un point de $\mathbb{R}^p$ et u, un vecteur de $\mathbb{R}^p$ accroché au point a. Pour $\mathbb{R}^n$, les notations seront y pour les coordonnées, b pour un point de $\mathbb{R}^n$ et v, un vecteur « accroché » au point b.

Nous avons

$$dg_b(v) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}(b) dy_i(v). \tag{12.737}$$

Ici $dy_i(v)$ signifie la $i$ème composante de v. C'est simplement v_i. Cette formule étant valable pour tout point $b\in\mathbb{R}^n$ et pour tout vecteur v, nous pouvons l'écrire en particulier pour

$$\begin{cases} b = f(a) & \text{(12.738a)}\\ v = df_a(u). & \text{(12.738b)} \end{cases}$$

Cela donne

$$dg_{f(a)}\big(df_a(u)\big) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) dy_i\big(df_a(u)\big). \tag{12.739}$$

Eodgfadfau

Mais

$$df_a(u) = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a) dx_j(u), \tag{12.740}$$

donc la $i$ème composante de ce vecteur est

$$\big(df_a(u)\big)_i = \sum_{j=1}^{p} \frac{\partial f_i}{\partial x_j}(a) dx_j(u). \tag{12.741}$$

En remplaçant $dy_i\big(df_a(u)\big)$ par cela dans l'expression (12.739), nous trouvons

$$dg_{f(a)}\big(df_a(u)\big) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) \sum_{j=1}^{p} \frac{\partial f_i}{\partial x_j}(a) dx_j(u). \tag{12.742}$$

Nous pouvons vérifier que c'est la différentielle de $g \circ f$ au point a, appliquée au vecteur u. En effet

$$d(g \circ f)_a(u) = \sum_{j=1}^{p} \frac{\partial (g \circ f)}{\partial x_j}(a) dx_j(u), \tag{12.743}$$

tandis que, par la dérivation de fonctions composées,

$$\frac{\partial (g \circ f)}{\partial x_j}(a) = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i}\big(f(a)\big) \frac{\partial f_i}{\partial x_j}(a). \tag{12.744}$$ EqDerCompofg

Au final, ce que nous avons prouvé est que

$$d(g \circ f)_a(u) = dg_{f(a)}\big(df_a(u)\big). \tag{12.745}$$ EQooWSIYooBmsBDU

12.25 Autres trucs sur la différentielle

12.25.1 Différentielle et dérivées partielles

Étant donné que pour tout vecteur u dans $\mathbb{R}^m$ on a $\partial_u f(a) = \nabla f(a) \cdot u$, le gradient de f nous donne la direction dans laquelle la croissance de f est maximale.

La matrice jacobienne calculée au point a est la matrice associée canoniquement à l'application linéaire $df_a : \mathbb{R}^m \to \mathbb{R}^n$.

12.25.2 Plan tangent

On a dit au début de cette section que si f est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ alors le graphe de f est une surface à deux paramètres et que l'application affine tangente au graphe de f au point $(a, f(a))$ est un plan. Maintenant on sait que ce plan est celui d'équation

$$T_a(x, y) = f(a_1, a_2) + \frac{\partial f}{\partial x}(a_1, a_2)(x - a_1) + \frac{\partial f}{\partial y}(a_1, a_2)(y - a_2). \tag{12.746}$$

Le plan tangent au graphe de f au point a est le graphe de cette fonction T_a.

PROPooJPRUooNOcXPJ

Proposition 12.281.
Il existe des fonctions différentiables dont les dérivées partielles ne sont pas continues.

Retenez que si vous obtenez que les dérivées partielles d'une fonction ne sont pas continues, vous ne pouvez pas immédiatement en conclure que la fonction ne sera pas différentiable.

12.25.3 Notes idéologiques quant au concept de plan tangent

ssecConceptPlanTag

Notons G, le graphe d'une fonction f, c'est-à-dire

$$G = \{(x, y, z) \in \mathbb{R}^3 \text{ tel que } z = f(x, y)\}. \tag{12.747}$$

Première affirmation : si $\gamma\colon \mathbb{R} \to G$ est une courbe telle que $\gamma(0) = \big(a, f(a)\big)$, alors $\gamma'(0) \in \mathbb{R}^n$ est dans le plan tangent à G au point $\big(a, f(a)\big)$.

Plus fort : tous les éléments du plan tangent sont de cette forme.

Le plan tangent à G en un point $x \in G$ est donc constitué des vecteurs vitesse de tous les chemins qui passent par x.

Prenons maintenant S, une courbe de niveau de G, c'est-à-dire

$$S = \{(x, y) \in \mathbb{R}^2 \text{ tel que } f(x, y) = C\}. \tag{12.748}$$

Si nous prenons un chemin dans G qui est, de plus, contraint à S, c'est-à-dire tel que $\gamma(t) \in S$, alors $\gamma'(0)$ sera tangent à G (ça, on le savait déjà), mais en plus, $\gamma'(0)$ sera tangent à S, ce qui est logique.

La morale est que si vous prenez un chemin qui se ballade dans n'importe quoi, alors la dérivée du chemin sera un vecteur tangent à ce n'importe quoi.

En outre, si $\gamma(t) \in S$ et $\gamma(0) = a$, alors

$$\langle \nabla f(a), \gamma'(0) \rangle = 0, \tag{12.749}$$

c'est-à-dire que le vecteur tangent à la courbe de niveau est perpendiculaire au gradient. Cela est intuitivement logique parce que la tangente à la courbe de niveau correspond à la direction de *moins* grande pente.

12.25.4 Gradient et recherche du plan tangent

Nous avons maintenant en main les concepts utiles pour trouver l'équation du plan tangent à une surface.

De la même manière que la tangente à une courbe était la droite de coefficient directeur donné par la dérivée, maintenant, le plan tangent à une surface est le plan dont les vecteurs directeurs sont les dérivées partielles :

La généralisation de l'équation (12.616) est

$$T_a(x) = f(a) + \sum_k \frac{\partial f}{\partial x_k}(a)(x - a)_k \tag{12.750}$$

EqDefPlanTag

Nous introduisons aussi souvent l'opérateur différentiel abstrait **nabla**, noté ∇ et qui est donné par le vecteur

$$\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right). \tag{12.751}$$

Les égalités suivantes sont juste des notations, sommes toutes logiques, liées à ∇ :

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right), \tag{12.752}$$

et

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right). \tag{12.753}$$

EqDefGradient

Ce dernier est un élément de $\mathbb{R}^n$: chaque entrée est un nombre réel.

Définition 12.282.
Le vecteur gradient de f au point a est le vecteur donné par la formule (12.753).

La notation ∇ permet d'écrire la différentielle sous forme un peu plus compacte. En effet, la formule (12.646) peut être notée

$$df_a(u) = \langle \nabla f(a), u \rangle. \tag{12.754}$$

En utilisant ce produit scalaire, l'équation (12.750) peut se réécrire

$$T_a(x) = f(a) + \sum_i \frac{\partial f}{\partial x_i}(a)(x - a)^i = f(a) + \langle \nabla f(a), x - a \rangle. \tag{12.755}$$

Afin d'éviter les confusions, il est souhaitable de bien mettre les parenthèses et noter $(\nabla f)(a)$ au lieu de $\nabla f(a)$.

12.25.4.1 Plan tangent en dimension deux

Le plan T_a avec $a = (a_1, a_2)$ a pour équation dans $\mathbb{R}^3$:

$$z = f(a_1, a_2) + \frac{\partial f}{\partial x}(a_1, a_2)\,(x - a_1) + \frac{\partial f}{\partial y}(a_1, a_2)\,(y - a_2). \tag{12.756}$$

EqPlanTgEnDimDeux

Définition 12.283.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable en un point a. Le plan tangent *au graphe de f en $(a, f(a))$ est l'ensemble des points*

$$\begin{aligned} T_a f &= \{(x, z) \in \mathbb{R}^n \times \mathbb{R} \text{ tel que } z = f(a) + df_a(x - a)\} \\ &= \{(x, z) \in \mathbb{R}^n \times \mathbb{R} \text{ tel que } z = f(a) + \langle \nabla f(a), x - a\rangle\} \end{aligned}$$

EXOooMWGTooLzZp0035

Exemple 12.284.
Étudiez la continuité, la dérivabilité et la continuité de la dérivée de chacune des fonctions suivantes :

(1) $x \mapsto \begin{cases} 0 & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

(2) $x \mapsto \begin{cases} \sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(3) $x \mapsto \begin{cases} x\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

Item0035d

(4) $x \mapsto \begin{cases} x^2\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

Dans ces exercices, les fonctions données sont dérivables et à dérivée continue sur $\mathbb{R}_0$ car pour $a \in \mathbb{R}_0$, il existe toujours une boule autour de a dans laquelle la fonction est composée de fonctions dérivables (sin, $\frac{1}{x}$, ...). L'intérêt de l'exercice est donc d'établir (ou réfuter) la continuité et la dérivabilité en 0.

(1) Notons f cette fonction. f n'est pas continue en 0 car

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x) = \lim_{\substack{x\to 0\\ x\neq 0}} 0 = 0 \neq f(0)$$

En particulier f n'est pas dérivable en 0 (et donc la continuité de sa dérivée n'a pas de sens en 0).

(2) Dans ce cas-ci, la limite « restreinte »

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x) = \lim_{\substack{x\to 0\\ x\neq 0}} \sin\left(\frac{1}{x}\right)$$

n'existe pas puisque, par exemple, pour la suite de terme général

$$x_k = \frac{1}{\frac{\pi}{2} + 2k\pi}$$

on a bien $\lim_{k\to\infty} x_k = 0$ mais

$$\lim_{k\to\infty} f(x_k) = 1 \neq f(0)$$

puisque pour tout $k \in \mathbb{N}$, $f(x_k) = 1$. Donc la fonction n'est pas continue.

(3) Montrons que la fonction, notée f, est continue en 0. Pour tout x réel, nous avons

$$0 \leqslant |f(x)| = |x| \left|\sin\left(\frac{1}{x}\right)\right| \leqslant |x|$$

ce qui montre que $\lim_{x\to 0} f(x) = 0$ par la règle de l'étau.

Par ailleurs, f n'est pas dérivable en 0 car la limite

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{f(x)-f(0)}{x-0} = \lim_{\substack{x\to 0\\ x\neq 0}} \sin\left(\frac{1}{x}\right)$$

n'existe pas, comme on l'a vu précédemment.

(4) Montrons que cette fonction, notée f, est dérivable en 0 (ce qui prouvera qu'elle y est continue). Calculons

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{f(x)}{x} = \lim_{\substack{x\to 0\\ x\neq 0}} x\sin\left(\frac{1}{x}\right) = 0$$

où la dernière égalité a été montrée à l'exercice précédent. Nous avons donc $f'(0) = 0$.

Par ailleurs, en utilisant les règles de calcul usuelles sur les dérivées, nous obtenons pour $x \neq 0$

$$f'(x) = \sin\left(\frac{1}{x}\right) - \frac{1}{x}\cos\left(\frac{1}{x}\right)$$

qui est une fonction ne possédant pas de limite en 0 puisque, par exemple, si x_k est tel que

$$\frac{1}{x_k} = \frac{\pi}{4} + 2k\pi$$

alors la suite $(x_k)_{k\in\mathbb{N}}$ tend vers 0, mais $f'(x_k)$ tend vers $+\infty$ lorsque $k \to +\infty$. La dérivée n'est donc pas continue en zéro.

△

EXooELTHooDdJyJE

Exemple 12.285.

Dessiner les courbes de niveaux des fonctions suivantes. Représenter ensuite leur graphe dans l'espace. Donner l'équation du plan tangent en l'origine.

(1) $f(x,y) = \sqrt{x^2+y^2}$.

(2) $f(x,y) = \sqrt{1-x^2-y^2}$.

(3) $f(x,y) = (x^2+y^2)^2 - 8xy$.

Item0041

Les courbes de niveau de l'exercice (3) sont les *ovales de Cassini* ; en particulier, la courbe de niveau 0 est la *lemniscate de Bernouilli*.

(1) $f(x) = \sqrt{x^2+y^2}$. Les courbes de niveau $f(x) = C$ sont des cercles (sauf $f(x) = 0$ qui se réduit à un point). Les sections horizontales étant des cercles, et le rayon de ces cercles augmentant linéairement, le graphe est une cône. Nous pouvons nous en convaincre en vérifiant par exemple que la droite $t \mapsto (t, 0, t)$ est bien entièrement contenue dans $z = f(x,y)$.

Afin de déterminer la différentielle, nous calculons les dérivées partielles

$$\frac{\partial f}{\partial x} = \frac{1}{2}(x^2+y^2)^{-1/2}\cdot 2x = \frac{x}{\sqrt{x^2+y^2}}, \tag{12.757}$$

EqDerrPaert0041x

et

$$\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}. \tag{12.758}$$

EqDerrPart0041y

Pour le plan tangent, nous essayons d'utiliser la formule (12.756). Pour cela, nous devons trouver les dérivées partielles en zéro.

Il est vite vu que la formule (12.757) n'a pas de limite pour $(x,y) \to (0,0)$. Ceci *ne prouve pas* que la différentielle de f n'existe pas en $(0,0)$. L'existence de la différentielle implique la continuité de la fonction, et non de la différentielle elle-même. En effet, une différentielle peut exister en un point sans qu'elle soit la limite de la différentielle aux autres points. Nous

avons vu par exemple, dans l'exercice 12.284(4), un exemple de fonction dérivable [121] en 0, mais dont la dérivée n'est pas continue en zéro.

Il ne suffit donc pas de calculer les limites de 12.757 et de 12.758 pour trouver la différentielle de f en $(0,0)$. Il n'est par contre pas très compliqué de remarquer que les dérivées partielles n'existent pas en $(0,0)$, par exemple parce que

$$\lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} \tag{12.759}$$

n'existe pas pour cause de limite différente pour $t>0$ et $t<0$. Il n'y a donc pas de plan tangent. Ceci est conforme à l'intuition : il n'y a pas de plan tangent à un cône en son sommet.

Nous pouvons faire une petite vérification du fait que le graphe est bien un cône : la droite reliant $(0,0,0)$ à $(x,y,\sqrt{x^2+y^2})$ est entièrement contenue dans le graphe de f. En effet si nous posons

$$\gamma(t)=(tx,ty,t\sqrt{x^2+y^2}), \tag{12.760}$$

pour tout t, nous avons $\gamma_z(t)=f\big(\gamma_x(t)^2+\gamma_y(t)^2\big)$.

(2) $f(x,y)=\sqrt{1-x^2-y^2}$. Les courbes de niveau $f(x,y)=C$ n'existent que pour $C\leqslant 1$, et ce sont des cercles

$$x^2+y^2=1-C. \tag{12.761}$$

Cette fois, le graphe est une coupole ce sphère. Nous allons en effet vérifier que l'arc de cercle centré en $(0,0,0)$ joignant se sommet $(0,0,1)$ à $(1,0,0)$ dans le plan $y=0$ est entièrement contenu dans le graphe de f. La symétrie de f sous les rotations dans le plan $x-y$ fait le reste. L'arc de cercle en question est le chemin

$$\gamma(t)=\big(1-t,0,\sqrt{1-(1-t)^2}\big). \tag{12.762}$$

Chaque point de ce chemin vérifie bien la relation

$$f\big(\gamma_x(t),\gamma_y(t)\big)=\gamma_z(t). \tag{12.763}$$

Le plan tangent à la coupole de sphère en $(0,0,1)$ est évidement horizontal. Nous nous attendons donc à trouver que la différentielle de f en $(0,0)$ est nulle. Simple calcul :

$$\frac{\partial f}{\partial x}(x,y)=\frac{1}{2}\frac{-2x}{\sqrt{1-x^2-y^2}}, \tag{12.764}$$

et

$$\frac{\partial f}{\partial y}(x,y)=\frac{1}{2}\frac{-2y}{\sqrt{1-x^2-y^2}}. \tag{12.765}$$

Évaluées en $(0,0)$, ce deux dérivées partielles sont nulles. Donc *si la différentielle existe* en $(0,0)$, elle sera nulle (voir l'expression (12.646)). Afin de voir qu'elle existe, il faut juste montrer que $df_{(0,0)}(x,y)=0$ fonctionne dans la définition 10.243.

(3) $f(x,y)=(x^2+y^2)^2-8xy$. La courbe de niveau zéro, en coordonnées polaires est donnée par

$$r=2\sqrt{\sin(2\theta)}. \tag{12.766}$$

Les dérivées partielles sont données par

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y)&=4(x^2+y^2)x-8y\\ \frac{\partial f}{\partial y}(x,y)&=4(x^2+y^2)y-8x\end{aligned} \tag{12.767}$$

△

121. Pour rappel, en dimension un, la dérivée est *exactement* la notion de différentielle.

12.25.4.2 Plan tangent en dimension trois

Nous avons vu que, de la même façon qu'en deux dimensions nous avions l'approximation (12.394) d'une fonction par sa tangente, en trois dimensions nous avons l'approximation suivante d'une fonction de deux variables :

$$f(x,y) \simeq f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \tag{12.768}$$

lorsque (x,y) n'est pas trop loin de (a,b). Cela signifie que le graphe de f ressemble au graphe de la fonction $T_{(a,b)}$ donnée par

$$T_{(a,b)}(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(x-a). \tag{12.769}$$

En notations compactes :

$$T_p(x) = f(p) + \nabla f(p) \cdot (x-p). \tag{12.770}$$

Le graphe de la fonction T_p sera le **plan tangent** au graphe de f au point p. L'équation du plan tangent sera donc

$$z - f(p) = \nabla f(p) \cdot (x-p). \tag{12.771}$$

Remarque 12.286.
Lorsque nous utilisons la notation vectorielle, la lettre « x » désigne le vecteur (x,y). Il faut être attentif. Dans un cas x est un vecteur dans l'autre c'est une composante d'un vecteur.

12.26 Jacobienne

12.26.1 Rappels et définitions

Dans cette section nous considérons des fonctions $f : D \to \mathbb{R}^m$ où $D \subset \mathbb{R}^n$, et un point $a \in \operatorname{Int} D$ où f est différentiable.

Remarque 12.287.
La définition de continuité (resp. différentiabilité) pour une fonction à valeurs vectorielles est celle introduite précédemment, et on remarque que pour avoir la continuité (resp. différentiabilité) de f en un point, il faut et il suffit que chacunes des composantes de $f = (f_1, \dots, f_m)$, vues séparément comme fonctions à n variables et à valeurs réelles, soient continues (resp. différentiables) en ce point.

Définition 12.288.
*La **jacobienne** de f en a est la matrice de l'application linéaire donnée par la différentielle. Elle a de nombreuses notations*

$$J_f(a) = \frac{\partial(f_1,\dots,f_m)}{\partial x_1,\dots,x_m} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \tag{12.772}$$

*Autrement dit, c'est la matrice composée de l'ensemble des dérivées partielles de f. Le **jacobien** de f au point a est le déterminant de cette matrice.*

*Si $m = 1$, cette matrice ne contient qu'une ligne ; c'est donc un vecteur appelé le **gradient** de f au point a et noté $\nabla f(a)$.*

Remarque 12.289. (1) Si la fonction est supposée différentiable, calculer la jacobienne revient à connaitre la différentielle. En effet, par linéarité de la différentielle et par définition des dérivées partielles, nous avons

$$df_a(u) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

où $u = (u_1, \dots, u_n)$ et où le membre de droite est un produit matriciel

(2) Remarquons que la jacobienne peut exister en un point donné sans que la fonction soit différentiable en ce point !

NORMooKBJVooDDDmOa

12.290.
Le théorème de différentiation de fonctions composées 12.278 peut également se lire au niveau des matrices jacobiennes. La matrice jacobienne de $g \circ f$ au point a est le produit matriciel des matrices jacobiennes de g et de f. Plus précisément, nous avons

$$J_{g\circ f}(a) = J_g\big(f(a)\big)J_f(a). \tag{12.773}$$

Remarquez que nous considérons la matrice jacobienne de g au point $f(a)$.

Dans le cas particulier où $m = 1$ et f est une fonction d'un intervalle I dans $\mathbb{R}^n$, dérivable au point a, on trouve que la fonction composée $g \circ f$ est dérivable au point a si g est différentiable, et alors

$$(g \circ f)'(a) = dg\,(f(a))\,.f'(a).$$

En fait, pour les fonctions d'une seule variable la dérivabilité coïncide avec la différentiabilité.

12.27 Fonctions de classe C^1

La définition des fonctions de classe C^k est la définition 10.261.

Soit f une fonction différentiable de U, ouvert de $\mathbb{R}^m$, dans $\mathbb{R}^n$. L'application différentielle de f est une application de $\mathbb{R}^m$ dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$

$$\begin{aligned} df\colon \mathbb{R}^m &\to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ a &\mapsto df_a \end{aligned} \tag{12.774}$$

Nous savons que $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ est un espace vectoriel normé avec la définition 4.31. Si T est un élément dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ alors la norme de T est définie par

$$\|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} = \sup_{x\in\mathbb{R}^m} \frac{\|T(x)\|_n}{\|x\|_m} = \sup_{\substack{x\in\mathbb{R}^m \\ \|x\|_m\leqslant 1}} \|T(x)\|_n.$$

Lorsqu'il existe un $M > 0$ tel que $\|df(a)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} < M$ pour tout a dans U, nous disons que la différentielle de f est **bornée** sur U.

NORMooDAZZooDiGFoW

12.291.
Quelques précisions sur l'énoncé de la proposition 12.292. Lorsque nous parlons de $\partial_i f$, nous supposons donnée une base de V. Il n'y a aucune raison que la norme sur V soit adaptée à cette base. Nous allons donc utiliser une norme « euclidienne » adaptée à la base, et invoquer l'équivalence de toutes les normes pour dire que si une fonction est différentiable pour une norme elle l'est pour toutes les normes.

12.27.0.1 Différentielle et dérivées partielles

PROPooUDJLooHwzjQF

Proposition 12.292 ([326]).
Soient des espaces vectoriels normés de dimensions finies V et E ainsi qu'une application $f\colon V \to E$. Nous supposons que les dérivées partielles $\frac{\partial f}{\partial x_i}$ existent sur un voisinage de $a \in V$ et qu'elles sont continues en a.

Alors

(1) l'application f est différentiable en a,

(2) la différentielle est donnée par[122]

$$df_a = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) \times_1 \omega_j. \tag{12.775}$$

Démonstration. L'espace V a une norme que nous notons $\|.\|_V$; nous n'allons presque pas l'utiliser. Nous nommons $\{e_i\}$ la base de V sous-entendue lorsque nous parlons des dérivées partielles, et nous considérons la norme $\|\|$ sur V donnée par $\|v\| = \sqrt{\sum_{i=1}^{n} x_i^2}$ pour $x = \sum_i x_i e_i$. Toutes les normes et boules dont nous allons parler dans la suite seront par rapport à cette norme.

Nous posons $\epsilon > 0$. La continuité des dérivées partielles en a nous permet de considérer $\delta > 0$ tel que

$$\|\frac{\partial f}{\partial x_j}(s) - \frac{\partial f}{\partial x_j}(a)\| < \epsilon \tag{12.776}$$

pour tout $s \in B(a, \delta)$, et pour tous les j en même temps[123].

Soit $x \in B(a, \delta)$. Nous considérons les points

$$s_j = (x_1, \dots, x_j, a_{j+1}, \dots, a_n). \tag{12.777}$$

Précision : $s_n = x$.

(1) **Un peu de convexité** Nous montrons que le segment $[s_j, s_{j-1}]$ est dans $B(a, \delta)$. Nous avons $s_j - a = (x_1 - a_1, \dots, x_j - a_j, 0, \dots, 0)$, et donc

$$\|s_j - a\| \leqslant \|x - a\| < \delta, \tag{12.778}$$

et donc $s_j \in B(a, \delta)$ pour tout j. Vu que la sphère $B(a, \delta)$ est convexe[124], tout le segment entre s_j et s_{j-1} est dedans.

(2) **Accroissements finis** En nous souvenant que x et a ont été fixés, nous considérons les fonctions intermédiaires suivantes :

$$\begin{cases} g_1(t) = f(t, a_2, \dots, a_n) & \text{(12.779a)} \\ g_j(t) = f(x_1, \dots, x_{j-1}, t, a_{j+1}, \dots, a_n) & \text{(12.779b)} \\ g_n(t) = f(x_1, \dots, x_{n-1}, t). & \text{(12.779c)} \end{cases}$$

Elles vérifient en particulier $g_j(x_j) = f(s_j)$ et $g_j(a_j) = f(s_{j-1})$. Donc le théorème des accroissements finis 12.192 nous donne l'existence de τ_j entre x_j et a_j tel que

$$f(s_j) - f(s_{j-1}) = g_j(x_j) - g_j(a_j) = g_j'(\tau_j)(x_j - a_j). \tag{12.780}$$

(3) **Dérivées partielles** Les fonctions g_j ont été construites de telle sorte à donner les dérivées partielles de f via une simple dérivation :

$$g_j'(t_0) = \frac{\partial f}{\partial x_j}(x_1, \dots, x_{j-1}, t_0, a_{j+1}, \dots, a_n). \tag{12.781}$$

Or lorsque[125] $t_0 \in [x_j, a_j]$, le point $(x_1, \dots, x_{j-1}, t_0, a_{j+1}, \dots, a_n)$ est dans $[s_j, s_{j-1}]$. Cela est le cas pour $t_0 = \tau_j$. Nous notons

$$\bar{s}_j = (x_1, \dots, x_{j-1}, \tau_j, a_{j+1}, \dots, a_n) \tag{12.782}$$

et nous avons

$$g_j'(\tau_j) = \frac{\partial f}{\partial x_j}(\bar{s}_j) \tag{12.783}$$

avec $\bar{s}_j \in [s_{j-1}, s_j]$. Tout cela pour récrire (12.780) sous la forme

$$f(s_j) - f(s_{j-1}) = \frac{\partial f}{\partial x_j}(\bar{s}_j)(x_j - a_j). \tag{12.784}$$

122. Voir la définition 10.264 pour $\times_n$.

123. Ce qui peut n'être pas possible en dimension infine. Je dis ça comme ça, juste pour faire remarquer que cette proposition n'est peut-être pas vraie en dimension infine. Voir aussi l'exemple de la fonction (25.152).

124. Proposition 11.28.

125. Retournez éventuellement les intervalles sur $a_j < x_j$.

(4) **Une belle somme télescopique** En nous souvenant que $s_0 = a$ et que $s_n = x$, nous avons cette somme télescopique

$$f(x) - f(a) = \sum_{j=1}^{n} \big(f(s_j) - f(s_{j-1})\big) \tag{12.785a}$$

$$= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\bar{s}_j)(x_j - a_j) \tag{12.785b}$$

$$= \sum_{j=1}^{n} \big[(\partial_j f)(\bar{s}_j) - (\partial_j f)(a)\big](x_j - a_j) + \sum_{j=1}^{n} (\partial_j f)(a)(x_j - a_j). \tag{12.785c}$$

Nous isolons les termes qui nous intéressent dans la définition de la différentielle :

$$f(x) - f(a) - \sum_{j=1}^{n} (\partial_j f)(a)(x_j - a_j) = \sum_{j=1}^{n} \big[(\partial_j f)(\bar{s}_j) - (\partial_j f)(a)\big](x_j - a_j). \tag{12.786}$$

(5) **Et enfin : le quotient différentiel** En utilisant la majoration (12.776), et en remarquant que $\|x - a\|$ majore $|x_j - a_j|$ pour chaque j, nous avons l'inégalité

$$\|f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)\| \leqslant n\epsilon\|x - a\|. \tag{12.787}$$

Autrement dit, en posant $T = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\omega_j$ nous avons

$$\|f(a + h) - f(a) - T(h)\| \leqslant n\epsilon\|h\|, \tag{12.788}$$

et donc

$$\lim_{h\to 0} \frac{\|f(a + h) - f(a) - T(h)\|}{\|h\|} = 0. \tag{12.789}$$

(6) **Et donc ?** La limite (12.789) nous indique que f serait différentiable de différentielle T si $\|.\|$ était la norme sur V. C'est l'équivalence de toutes les normes [126] qui fait en sorte que la norme sur V n'est pas importante.

□

Proposition 12.293.
Soient des espaces vectoriels normés de dimensions finies V et E ainsi qu'une application $f : V \to E$. Nous supposons que les dérivées partielles $\frac{\partial f}{\partial x_i}$ existent et sont continues sur un ouvert $\mathcal{U}$ de V.

Alors f est de classe C^1 sur $\mathcal{U}$.

Démonstration. Nous notons $\mathcal{U}$ le voisinage sur lequel les dérivées partielles de f existent et sont continues. Il s'agit d'appliquer la proposition 12.292 en chaque point de $\mathcal{U}$. Nous avons alors

$$df_x = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \times_1 \omega_j \tag{12.790}$$

pour tout $x \in \mathcal{U}$. Pour $a \in V$ et $h \in V$ supposé petit, nous avons

$$\|df_a - df_{a+h}\|_{\mathcal{L}(V,E)} = \sup_{\|v\|=1} \|df_a(v) - df_{a+h}(v)\|_E \tag{12.791a}$$

$$= \sup_{\|v\|=1} \| \sum_{j=1}^{n} \big(\frac{\partial f}{\partial x_j}(a) - \frac{\partial f}{\partial x_j}(a + h)\big)v_j\| \tag{12.791b}$$

$$\leqslant \sup_{\|v\|=1} \| \sum_{j=1}^{n} \big(\frac{\partial f}{\partial x_j}(a) - \frac{\partial f}{\partial x_j}(a + h)\big)\|v_i| \tag{12.791c}$$

$$\leqslant \| \sum_{j=1}^{n} \big(\frac{\partial f}{\partial x_j}(a) - \frac{\partial f}{\partial x_j}(a + h)\big)\| \tag{12.791d}$$

Justifications :

126. Sur un espace de dimension finie, 10.49.

— Pour (12.791d), nous avons majoré $|v_i| \leqslant \|v\| = 1$.

— Pour (12.791c), nous avons utilisé le lemme 10.66.

La continuité de $\partial_j f$ conclut que $\lim_{h\to 0} \|df_a - df_{a+h}\| = 0$, ce qui signifie que df est continue et donc que f est de classe C^1. □

THOooBEAOooBdvOdr

Théorème 12.294.
Soient des espaces vectoriels normés de dimensions finies V et E. Soit un ouvert $\mathcal{U}$ de V. Une application $f\colon V \to E$ est de classe C^1 sur $\mathcal{U}$ si et seulement si ses dérivées partielles existent et sont continues sur $\mathcal{U}$.

Dans ce cas, nous avons la formule

$$df_a = \sum_i \frac{\partial f}{\partial x_i}(a) \times_1 \omega_i. \tag{12.792}$$

EQooPREWooVHVIAF

Démonstration. Le premier sens, y compris la formule (12.792), a été fait dans la proposition 12.293. Nous supposons maintenant que f est de classe C^1 sur $\mathcal{U}$ et nous allons prouver que ses dérivées partielles existent et sont continues. Pour tout $t \neq 0$ nous avons l'égalité

$$\frac{f(a+te_i) - f(a)}{t} = \frac{f(a+te_i) - f(a) - df_a(te_i)}{t} + df_a(e_i). \tag{12.793}$$

Par définition de la différentielle, prendre la limite $t \to 0$ à droite donne $df_a(e_i)$ parce que la fraction tend vers zéro. La limite définissant la dérivée partielle existe donc et vaut

$$\frac{\partial f}{\partial x_i}(a) = df_a(e_i). \tag{12.794}$$

Il nous reste à prouver que les dérivées partielles sont continues :

$$\begin{aligned}
\|\frac{\partial f}{\partial x_i}(a+h) - \frac{\partial f}{\partial x_i}(a)\| &= \|df_a(e_i) - df_{a+h}(e_i)\| && \text{(12.795a)}\\
&= \|(df_{a+h} - df_a)(e_i)\| && \text{(12.795b)}\\
&\leqslant \|df_{a+h} - df_a\|\|e_i\| && \text{(12.795c)}\\
&= \|df_{a+h} - df_a\|. && \text{(12.795d)}
\end{aligned}$$

Vu que f est de classe C^1, la limite $h \to 0$ de cela donne zéro. □

Voici une version du même théorème, démontré dans le cas seulement de dimension deux. Il rend peut-être aussi plus clairement pourquoi ça ne marche pas en dimension infinie.

Diff_totale

Proposition 12.295.
Soit U un ouvert dans $\mathbb{R}^m$ et a un point dans U. Soit f une application de U dans $\mathbb{R}^n$. Si toutes les dérivées partielles de f existent sur U et sont continues au point a alors f est différentiable au point a.

Démonstration. On se limite au cas $m = 2$. Pour rendre les calculs plus simples on utilise ici la norme $\|\cdot\|_\infty$ dans l'espace $\mathbb{R}^2$, mais comme on a vu plus en haut, cela ne peut pas avoir des conséquences sur la différentiabilité de f. Si la différentielle de f au point a existe alors elle est définie par la formule

$$df_a(v) = \frac{\partial f}{\partial x}(a)v_1 + \frac{\partial f}{\partial y}(a)v_2$$

pour tout v dans $\mathbb{R}^m$.

On commence par prouver le résultat en supposant que les dérivées partielles de f au point a sont nulles. La différentiabilité de f signifie que pour toute constante $\varepsilon > 0$ il y a une constante $\delta > 0$ telle que si $\|v\|_\infty \leqslant \delta$ alors

$$\frac{\|f(a_1+v_1, a_2+v_2) - f(a_1,a_2)\|_n}{\|v\|_\infty} \leqslant \varepsilon.$$

On écrit alors

$$\begin{aligned}&\|f(a_1+v_1,a_2+v_2)-f(a_1,a_2)\|_n=\\&=\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)+f(a_1+v_1,a_2)-f(a_1,a_2)\|_n\leqslant\\&\leqslant\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)\|_n+\|f(a_1+v_1,a_2)-f(a_1,a_2)\|_n.\end{aligned}\tag{12.796}$$

Comme la dérivée partielle $\partial_x f$ est nulle au point a on sait que pour toute constante $\varepsilon>0$ il y a une constante $\delta_1>0$ telle que si $|v_1|\leqslant\delta_1$ alors

$$\|f(a_1+v_1,a_2)-f(a_1,a_2)\|_n\leqslant\varepsilon|v_1|.$$

Pour l'autre terme on a, par la proposition 10.259,

$$\|f(a_1+v_1,a_2+v_2)-f(a_1+v_1,a_2)\|_n\leqslant\sup\{\|\partial_y f(x)\|_n\,|\,x\in S\}|v_2|.\tag{12.797}$$

où S est le segment d'extrémités (a_1+v_1,a_2) et (a_1+v_1,a_2+v_2). Comme la dérivée partielle $\partial_y f$ est continue et nulle au point a on sait que pour toute constante $\varepsilon>0$ il existe une constante $\delta_2>0$ telle que si $\|(u_1,u_2)\|_\infty\leqslant\delta_2$ alors $\|\partial_y f(a_1+u_1,a_2+u_2)\|_n\leqslant\varepsilon$. Si on choisit $\delta=\min\{\delta_1,\,\delta_2\}$ le segment S est contenu dans la boule de rayon δ centrée au point a et on obtient

$$\|f(a_1+v_1,a_2+v_2)-f(a_1,a_2)\|_n\leqslant\varepsilon|v_1|+\varepsilon|v_2|\leqslant2\varepsilon\|v\|_\infty.$$

Cela prouve que f est différentiable en (a_1,a_2) et que la différentielle est nulle :

$$df_{(a_1,a_2)}=0.\tag{12.798}$$

Dans le cas général, où les dérivées partielles de f au point a ne sont pas spécialement nulles, on peut considérer la fonction [127]

$$g(x,y)=f(x,y)-\partial_1 f(a)x-\partial_2 f(a)y,\tag{12.799}$$

EqXHVooJeQKrR

qui a dérivées partielles nulles au point a. La fonction g est donc différentiables. La fonction f est maintenant la somme de g et de la fonction linéaire et continue $(x,y)\mapsto\partial_1 f(a)x-\partial_2 f(a)y$. On verra dans la prochaine section que la somme de deux fonctions différentiables est une fonction différentiable. Par conséquent, la fonction f est différentiable. □

Remarque 12.296.
En dimension infinie, il n'est pas vrai que l'existence et la continuité de toutes les dérivées partielles en un point implique la différentiabilité en ce point. Pour donner un exemple, nous allons continuer l'exemple 10.71 avec la fonction 25.152 sur un espace de Hilbert.

En dimension infinie nous aurons le théorème 10.294 qui donnera quelque chose de moins fort.

12.27.0.2 Dérivée partielle de fonctions composées

12.297.
Une petite note sur les notations du théorème de dérivation partielle de fonctions composées. La formule (12.811) est plus souvent écrite sous la forme

$$\frac{\partial(f\circ g)}{\partial x_i}(a)=\sum_{k=1}^{n}\frac{\partial f}{\partial y_k}\big(g(a)\big)\frac{\partial g_k}{\partial x_i}(a).\tag{12.800}$$

Du point de vue du névrosé des notations précises que je suis, cette façon d'écrire est délicate à exprimer parce qu'il faudrait décider ce que signifie « noter x les variables de $\mathbb{R}^n$ et y celles de $\mathbb{R}^m$ ».

Et je ne vous parle même pas des problèmes que ça pose si x est justement le nom d'une fonction quand on considère un changement de variable.

127. Vous verrez dans la discussion à propos de la fonction (25.152) pourquoi cette fonction ne fonctionne pas dans le cas de la dimension infinie.

12.298.
À propos de changement de variables ...J'ai une bonne nouvelle : il n'y a pas de notions de « différentielle partielle » et de « différentielle totale ». Ces notions sont introduites par les personnes qui utilisent de mauvaises notations pour distinguer deux notions différentes qu'ils sont incapables de distinguer par des notations claires [128].

Les choses mal faites en une dimension Voici comment on présente (mal) les choses. Soit une fonction $y(x)$. Si on effectue un changement de variables $x = x(t)$, on peut voir y comme une fonction de t au lieu de x, et parler de la dérivée de y à travers x.

Si $y(x) = x^2$ et si $x(t) = \sin(t)$, on se retrouve à écrire

$$y'(t) = 2\sin(t)\cos(t) \tag{12.801}$$

et

$$y'(x) = 2x. \tag{12.802}$$

Comment l'objet y' peut dépendre du nom de la variable ? ! ? Notez qu'en substituant $x = \sin(t)$ dans (12.802), le compte n'est pas bon : on n'a pas (12.801).

Pour résoudre ce problème, on peut dire que la bonne quantité à regarder n'est pas y' mais bien dy. Alors on doit en fait regarder non $y'(x)$ mais $y'(x)dx$ où $dx = \cos(t)dt$. Avec ça, on a en effet

$$dy = \frac{dy}{dx}\frac{dx}{dt}dt = 2\sin(t)\cos(t)dt, \tag{12.803}$$

et

$$dy = \frac{dy}{dt}dt = 2\sin(t)\cos(t)dt. \tag{12.804}$$

Il faut alors bien comprendre que le « y » dans $\frac{dy}{dt}$ n'est pas le même que le y dans $\frac{dy}{dx}$.

Les choses bien faites en une dimension Soit une fonction $y\colon \mathbb{R} \to \mathbb{R}$. Nous considérons une application (appelée « changement de variable » si on veut, mais surtout appelée « difféomorphisme de $\mathbb{R}$ vers $\mathbb{R}$ ») $x\colon \mathbb{R} \to \mathbb{R}$.

Maintenant que la notation « x » est prise pour désigner une fonction, il n'est plus possible d'écrire $y(x)$, parce que y est une application qui prend en argument un élément de $\mathbb{R}$ et non de $C^\infty(\mathbb{R})$.

On introduit donc une nouvelle fonction $z = y \circ x$. Donc $z(t) = y\big(x(t)\big) = \sin(t)^2$. Le nom de la variable importe peu tant que ce n'est pas un nom déjà utilisé. Nous avons alors aucune ambigüité :

$$y'(u) = 2u \tag{12.805}$$

et

$$z'(u) = 2\sin(u)\cos(u). \tag{12.806}$$

Notez que les égalités suivante sont parfaitement correctes et ne souffrent d'aucun problème d'interprétation :

— $y(t) = t^2$
— $z(\xi) = \sin(\xi)^2$
— $x(s) = \sin(s)$
— $x\big(y(u)\big) = \sin\big(y(u)\big) = \sin(u^2)$.

Si les fonctions y et x représentent des quantités physiques, certaines de ces formules sont peut-être idiotes à écrire, mais elles sont correctes.

128. Là je vise la quasi totalité des sources parlant de changement de variable dans les équations différentielles et les physiciens, en particulier les mécaniciens.

Les choses mal faites en dimension 2 Soit une fonction f de la position et du temps qu'on note $f(x,t)$. Cette fonction vérifie une équation aux dérivées partielles, mettant en jeu $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial t}$.

Coup de théâtre : on veut savoir la valeur de f le long d'une trajectoire $x(t)$. Nous voici avec

$$f\big(x(t),t\big). \tag{12.807}$$

Et là quand on veut parler de la dérivée de f par rapport à t, on doit distinguer la dérivée partielle $\frac{\partial f}{\partial t}$ qui consiste à ne dériver l'expression $f\big(x(t),t\big)$ que par rapport aux t qui apparaissent « vraiment », de la « dérivée totale » qui consiste à dériver par rapport à toutes les occurrences de t, y compris à travers x.

Les choses bien faites en dimension 2 Soit une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$. Prenons par exemple $f(x,t) = x\sin(t)$. Soit une trajectoire $s\colon \mathbb{R} \to \mathbb{R}$ que nous nous gardons bien de noter « x ». Ce qu'on entend par « voir f sur la trajectoire s » signifie en réalité considérer la fonction

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f\big(s(t),t\big). \end{aligned} \tag{12.808}$$

Ce qu'on aurait appelé la dérivée totale de f par rapport à t est simplement la dérivée usuelle φ' de φ en tant que bête fonction $\mathbb{R} \to \mathbb{R}$. Si par exemple $s(t) = t^2$, nous avons

$$\varphi(t) = t^2\sin(t). \tag{12.809}$$

Nous avons donc toutes les dérivées sans ambigüité :

$$\frac{\partial f}{\partial x}(x,t) = \sin(t) \tag{12.810a}$$

$$\frac{\partial f}{\partial t}(x,t) = x\cos(t) \tag{12.810b}$$

$$\varphi'(t) = 2t\sin(t) + t^2\cos(t). \tag{12.810c}$$

Le mieux serait même d'écrire $\partial_1 f$ et $\partial_2 f$ au lieu de $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial t}$, parce que ce sont des fonctions complètement déterminées par f et non par la notation x et t qu'on a choisie pour nommer les variables au moment d'écrire $f(x,t) = x\sin(t)$.

12.299.
Et d'ailleurs en mathématique, il n'y a rien qui s'appelle « changement de variable ». Il y a seulement des choses qui s'appellent « composition de fonction ».

Cela n'est pas limité à l'analyse. Il n'y a par exemple pas de concept de « choisir une base dans laquelle une matrice est diagonale ». Si S est une matrice symétrique, il existe une matrice diagonale D et une matrice orthogonale A telles que $D = ASA^{-1}$. Quand on veut démontrer une théorème sur S, on commence par démontrer le théorème dans le cas particulier où S est diagonale, puis on espère que le résultat se généralise facilement aux matrices de la forme ADA^{-1} où D est diagonale et A est orthogonale.

12.300.
Tout cela pour dire que nous allons maintenant prouver le théorème de dérivation partielle de fonctions composées. C'est ce théorème qui est utilisé chaque fois qu'on fait un « changement de variables » dans une équation aux dérivées partielles.

THOooKBTYooWFtoSF

Théorème 12.301 ([1, 409])**.**
Soient :

(1) trois espaces vectoriels normés U, V et W de dimension finie

(2) $a \in V$, un voisinage Ω de a

(3) des applications $g\colon \Omega \to U$ et $f\colon g(\Omega) \to W$

(4) des bases de U, V et W de telle sorte que parler des dérivées partielles ait un sens. Toutes les bases vont être notées $\{e_i\}$ sans précisions. Sinon, on ne va pas s'en sortir.

Nous supposons que

(1) f est de classe C^1 sur un voisinage de $g(a)$ [129],

(2) g admet une dérivée partielle dans la direction de e_i en a.

Alors

(1) La fonction $f \circ g$ a une dérivée partielle dans la direction de e_i en a,

(2) nous avons la formule

$$\partial_i(f\circ g)(a) = \sum_k (\partial_k f)\big(g(a)\big)(\partial_i g_k)(a). \tag{12.811}$$

EQooZMAJooIusxgD

Démonstration. Vu que f est différentiable en $g(a)$, nous pouvons y utiliser le lemme 12.255 et écrire

EQooFZBVooTwNexc

$$df_{g(a)}\left(\frac{\partial g}{\partial x_i}(a)\right) = \frac{d}{dt}\Big[f\big(g(a)+t\frac{\partial g}{\partial x_i}(a)\big)\Big]_{t=0} \tag{12.812a}$$

$$= \lim_{\epsilon\to 0}\frac{f\big(g(a)+\epsilon\frac{\partial g}{\partial x_i}(a)\big)-f\big(g(a)\big)}{\epsilon}. \tag{12.812b}$$

Maintenant, le jeu est de travailler $f\big(g(a)+\epsilon\partial_i g(a)\big)$ de deux façons différentes. D'une part en effectuant un développement de f autour de $g(a)$ et d'autre part dé-développant $g(a)+\epsilon\partial_i g(a)$ pour le changer en $g(a+\epsilon e_i)$.

(1) **Développer f** En utilisant le lemme 12.237 pour f autour de $g(a)$, nous avons

$$f\big(g(a)+\epsilon\partial_i g(a)\big) = f\big(g(a)+\epsilon\sum_k(\partial_i g_k)(a)e_k\big) \tag{12.813a}$$

$$= f\big(g(a)\big)+\epsilon\sum_k(\partial_i g_k)(a)(\partial_k f)\big(g(a)+\epsilon u_k\big) \tag{12.813b}$$

où $u_k = \sum_{l=k+1}^n(\partial_i g_l)(a)e_l$. La forme exacte de u_k n'est pas importante pour notre histoire. En mettant cela dans la dernière fraction de (12.812),

$$\frac{f\big(g(a)+\epsilon(\partial_i g)(a)\big)-f\big(g(a)\big)}{\epsilon} \tag{12.814a}$$

$$= \frac{f\big(g(a)\big)+\epsilon\sum_k(\partial_i g_k)(a)(\partial_k f)\big(g(a)+\epsilon u_k\big)+\epsilon\alpha(\epsilon)-f\big(g(a)\big)}{\epsilon} \tag{12.814b}$$

$$= \sum_k(\partial_i g_k)(a)(\partial_k f)\big(g(a)+\epsilon u_k\big)+\alpha(\epsilon). \tag{12.814c}$$

Nous prenons la limite $\epsilon\to 0$ en tenant compte du fait que $\lim_{\epsilon\to 0}\alpha(\epsilon)=0$ et du fait que les dérivées partielles de f sont continues en $g(a)$ (théorème 12.294). Nous avons

$$df_{g(a)}\big((\partial_i g)(a)\big) = \sum_k(\partial_i g_k)(a)(\partial_k f)\big(g(a)\big). \tag{12.815}$$

(2) **Développer g** Nous développons maintenant g à l'intérieur de f. Pour cela nous utilisons les accroissements finis 12.238 pour g :

$$g(a)+\epsilon\frac{\partial f}{\partial x_i}(a) = g(a+\epsilon e_i)-\epsilon\alpha(\epsilon). \tag{12.816}$$

Pour avancer dans (12.812), nous considérons le terme

$$f\big(g(a)+\epsilon\frac{\partial g}{\partial x_i}(a)\big) = f\big(g(a+\epsilon e_i)-\epsilon\alpha(\epsilon)\big) \tag{12.817}$$

129. Classe C^1, définition 10.261.

dans lequel nous allons développer f autour de $g(a+\epsilon e_i)$. Vu que la fonction $\epsilon \mapsto g(a+\epsilon e_i)$ est continue (elle est dérivable par hypothèse), pourvu que ϵ soit assez petit, le point $g(a+\epsilon e_i)$ est encore dans le voisinage de $g(a)$ sur lequel f est de classe C^1, de telle sorte qu'un développement de f ne pose pas de problèmes. Nous pouvons appliquer le lemme (12.237) :

$$f\big(g(a+\epsilon e_i)-\epsilon\alpha(\epsilon)\big)=f\big(g(a+\epsilon e_i)\big)-\sum_k \epsilon\alpha(\epsilon)_k(\partial_k f)\big(g(a+\epsilon e_i)+\epsilon u_k\big)+\epsilon\beta(\epsilon). \tag{12.818}$$

Ce que nous avons dans la limite dans (12.812) est

$$\frac{f\big(g(a+\epsilon e_i)\big)-\sum_k \epsilon\alpha(\epsilon)_k(\partial_k f)\big(g(a+\epsilon e_i)+\epsilon u_k\big)-f\big(g(a)\big)}{\epsilon} \tag{12.819a}$$

$$=\frac{(f\circ g)(a+\epsilon e_i)-(f\circ g)(a)}{\epsilon}-\sum_k \alpha(\epsilon)_k(\partial_k f)\big(f(a+\epsilon e_i)+\epsilon u_k\big). \tag{12.819b}$$

Nous allons passer à la limite. Vu que f est de classe C^1, ses dérivées partielles sont continue. Le vecteur u_k ne dépendant pas de ϵ, toute la partie $(\partial_k f)\big(g(a+\epsilon e_i)+\epsilon u_k\big)$ peut être majorée uniformément en ϵ. Et comme $\alpha(\epsilon)\to 0$, tout le second terme disparait.
Ce qui reste à la limite est

$$df_{g(a)}\left(\frac{\partial g}{\partial x_i}(a)\right)=\frac{\partial(f\circ g)}{\partial x_i}(a). \tag{12.820}$$

□
CORooNQURooAWjfcA

Corolaire 12.302.
Soient $a\in\mathbb{R}$, un voisinage ouvert I de a, une application $\gamma\colon I\to\mathbb{R}^n$ et une fonction $f\colon \gamma(I)\to\mathbb{R}$. Nous supposons que f est de classe C^1 sur $\gamma(I)$ et que γ est dérivable en a.
Alors

(1) La fonction $f\circ\gamma$ est dérivable en a.

(2) Nous avons la formule

$$(f\circ\gamma)'(a)=\gamma'(a)\cdot(\nabla f)\big(\gamma(a)\big). \tag{12.821}$$

Démonstration. Toutes les hypothèses sont calquées sur celles du théorème 12.301. L'application $f\circ\gamma$ est donc dérivable en a. Pour les application sur $\mathbb{R}$, la seule dérivée partielle est la dérivée usuelle, donc la formule (12.811) s'écrit

$$(f\circ\gamma)'(t)=\sum_k(\partial_k f)\big(\gamma(a)\big)\gamma_k'(a)=(\nabla f)\big(\gamma(a)\big)\cdot\gamma'(a). \tag{12.822}$$

□

Donnons un exemple d'utilisation de cette formule. Si

$$\begin{aligned} g&\colon\mathbb{R}^2\to\mathbb{R}^3\\ f&\colon\mathbb{R}^3\to\mathbb{R}, \end{aligned} \tag{12.823}$$

nous avons $\varphi\colon\mathbb{R}^2\to\mathbb{R}$. Les dérivées partielles de φ sont données par les formules

$$\frac{\partial\varphi}{\partial x}(x,y)=\frac{\partial f}{\partial x_1}\big(g(x,y)\big)\frac{\partial g_1}{\partial x}(x,y)+\frac{\partial f}{\partial x_2}\big(g(x,y)\big)\frac{\partial g_2}{\partial x}(x,y)+\frac{\partial f}{\partial x_3}\big(g(x,y)\big)\frac{\partial g_3}{\partial x}(x,y) \tag{12.824}$$

et

$$\frac{\partial\varphi}{\partial y}(x,y)=\frac{\partial f}{\partial x_1}\big(g(x,y)\big)\frac{\partial g_1}{\partial y}(x,y)+\frac{\partial f}{\partial x_2}\big(g(x,y)\big)\frac{\partial g_2}{\partial y}(x,y)+\frac{\partial f}{\partial x_3}\big(g(x,y)\big)\frac{\partial g_3}{\partial y}(x,y) \tag{12.825}$$

Notez que les dérivées de φ et des composantes de g sont calculées en (x,y), tandis que celles de f sont calculées en $g(x,y)$.

12.28 Différentielle et dérivée complexe

SECooJWNOooOgMiWR

12.303.
Nous commençons par donner quelques éléments à propos de dérivée et de différentielle pour des fonctions $\mathbb{C} \to \mathbb{C}$ parce que les séries entières vont souvent être des fonctions complexes. Le gros du chapitre sur les fonctions holomorphes est le chapitre 26.

Nous identifions $\mathbb{R}^2$ à $\mathbb{C}$ par l'application

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{C} \\ (x,y) &\mapsto x+iy. \end{aligned} \tag{12.826}$$

Dans cette partie, nous désignons par Ω un ouvert de $\mathbb{C}$.

DEFooVJVXooKlnFkh

Définition 12.304.
*Une fonction $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-**dérivable** si la limite*

$$\lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{f(a+h)-f(a)}{h} \tag{12.827}$$

existe. Dans ce cas, cette limite est la dérivée de f et est notée f'.

DefMMpjJZ

Définition 12.305.
*Soit Ω un ouvert dans $\mathbb{C}$. Une fonction $f\colon \Omega \to \mathbb{C}$ est **holomorphe** si elle est $\mathbb{C}$-dérivable sur Ω.*

Une avalanche de conditions équivalentes à l'holomorphie est donnée dans le théorème 26.82.

DEFooQSMCooOoWVZk

Définition 12.306.
*Si K est un compact de $\mathbb{C}$, nous disons qu'une fonction est **holomorphe** sur K si il existe un ouvert contenant K sur lequel f est holomorphe.*

Et si f n'est réellement définie que sur K, elle est holomorphe sur K si il existe une extension holomorphe de f vers un ouvert contenant K.

Définition 12.307.
Une matrice de la forme

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \tag{12.828}$$

*avec $\alpha, \beta \in \mathbb{R}$ est une **similitude**.*

LEMooJNFEooZCbJMo

Lemme 12.308.
En tant qu'application linéaire $\mathbb{C} \to \mathbb{C}$, l'opération de multiplication par $\alpha+\beta i$ est la matrice

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{12.829}$$

Démonstration. Cela est vite remarqué en calculant explicitement $(\alpha+\beta i)(u_1+iu_2)$. □

LEMooKIOZooEUxfsB

Lemme 12.309.
Une application $A\colon \mathbb{C} \to \mathbb{C}$ est $\mathbb{C}$-linéaire si et seulement si elle est une similitude en tant qu'application $\mathbb{R}^2 \to \mathbb{R}^2$.

Dans ce cas, il existe $z_0 \in \mathbb{C}$ tel que $A(z) = z_0 z$ pour tout $z \in \mathbb{C}$.

Démonstration. Commençons par considérer l'application A sur $\mathbb{R}^2$. Elle est en particulier une application $\mathbb{R}$-linéaire et par conséquent il existe une matrice $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ telle que

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{12.830}$$

Nous voulons maintenant imposer la $\mathbb{C}$-linéarité, c'est-à-dire que nous voulons

$$A\big((a+bi)(x+iy)\big) = (a+bi)A(x+iy) \tag{12.831}$$

pour tout $a,b,x,y\in\mathbb{R}$. À gauche nous avons

$$A\big(ax-by+i(bx+ay)\big) \tag{12.832}$$

et à droite nous avons

$$(a+bi)\big(\alpha x+\beta y+i(\gamma x+\delta y)\big). \tag{12.833}$$

En égalant les deux expressions nous obtenons les équations

$$\begin{cases} \beta b = -b\gamma & \text{(12.834a)}\\ -\alpha b+\beta a = a\beta - b\delta & \text{(12.834b)}\\ \delta b = b\alpha & \text{(12.834c)}\\ -\gamma b+\delta a = b\beta + a\delta, & \text{(12.834d)} \end{cases}$$

dont nous tirons immédiatement que $\gamma=-\beta$ et $\delta=\alpha$. La matrice de A est donc de la forme demandée.

Inversement nous devons prouver que la fonction

$$f(x+iy) = \alpha x+\beta y+i(-\beta x+\alpha y) \tag{12.835}$$

EqOEWYooMaHCNb

est $\mathbb{C}$-linéaire, c'est-à-dire qu'elle vérifie $f(z_0z)=z_0f(z)$ pour tout $z_0,z\in\mathbb{C}$. Cela est un simple calcul que nous confions à Sage : le code suivant affiche « 0 ».

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

def f(z):
    var('alpha,beta')
    x=z.real_part()
    y=z.imag_part()
    return alpha*x+beta*y+I*( -beta*x+alpha*y  )


var('a,b,x,y')

A=a+b*I
Z=x+y*I

z1=f( A*Z  )
z2=A*f( Z )

rep=z1-z2
print(rep.full_simplify())
```

tex/frido/code_sage3.py

Pour conclure, notons que la fonction (12.835) est la fonction de multiplication par $\alpha-i\beta$. □

NORMooMKNDooBeoGRN

12.310.
Soient une fonction $f\colon\mathbb{C}\to\mathbb{C}$ et l'isomorphisme canonique $\varphi\colon\mathbb{C}\to\mathbb{R}^2$. La fonction f définit la fonction

$$F=\varphi\circ f\circ\varphi^{-1}\colon\mathbb{R}^2\to\mathbb{R}^2. \tag{12.836}$$

Cela est la fonction $\mathbb{R}^2 \to \mathbb{R}^2$ associée à f. Il serait tentant de croire que tout ce qui est vrai pour F est également vrai pour f. Eh bien non.

Par exemple, F peut être différentiable sans que f le soit. La proposition suivante donne une condition sur dF pour que f soit différentiable.

La proposition suivante donne le lien entre df et la dérivée complexe f'. Pour avoir le lien avec $\partial_z f$, il faudra voir la proposition 26.4.

PropKJUDooJfqgYS

Proposition 12.311.
Une fonction $f : \Omega \to \mathbb{C}$ est $\mathbb{C}$-dérivable en $a \in \Omega$ si et seulement si elle est différentiable en a et si df_a est une similitude.

Plus précisément avec les notations de 12.310, la fonction f est $\mathbb{C}$-dérivable (donc holomorphe) au point $z_0 = x_0 + iy_0$ si et seulement si la fonction F est différentiable en (x_0, y_0) et si la matrice de dF est de la forme

$$dF = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \tag{12.837}$$

EQooWZGKooLDEHCr

c'est-à-dire si $dF_{(x_0,y_0)}$ fournit une application $\mathbb{C}$-linéaire.

Dans ce cas, le lien entre $\mathbb{C}$-dérivée et différentielle est donné par

$$(df_{z_0})(z) = f'(z_0)z. \tag{12.838}$$

EqPAEFooYNhYpz

Démonstration. Nous décomposons f en parties réelles et imaginaires :

$$f(x + iy) = P(x, y) + iQ(x, y) \tag{12.839}$$

où P et Q sont des fonctions réelles. La jacobienne de F est la matrice

$$\begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} \\ \frac{\partial Q}{\partial x} & \frac{\partial Q}{\partial y} \end{pmatrix}, \tag{12.840}$$

et la condition dont nous parlons s'écrit comme le système

EqFDUrXBP

$$\begin{cases} \dfrac{\partial P}{\partial x} = \dfrac{\partial Q}{\partial y} & \text{(12.841a)} \\ \dfrac{\partial P}{\partial y} = -\dfrac{\partial Q}{\partial x}. & \text{(12.841b)} \end{cases}$$

Si F est différentiable en (x_0, y_0) alors nous avons

$$F\big((x_0, y_0) + (h, k)\big) = F(x_0, y_0) + dF_{(x_0,y_0)} \begin{pmatrix} h \\ k \end{pmatrix} + s(|h| + |k|) \tag{12.842}$$

EqwlV6R

où s est une fonction vérifiant $\lim_{t\to 0} \frac{s(t)}{t} = 0$. Soit

$$dF_{(x_0,y_0)} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{12.843}$$

Si nous posons $\sigma = \alpha - i\beta$ et $w = h + ik$, l'équation (12.842) s'écrit dans $\mathbb{C}$ sous la forme

$$f(z_0 + w) = f(z_0) + \sigma w + s(|w|), \tag{12.844}$$

EqYFmoiM

ce qui implique que f est $\mathbb{C}$-dérivable en z_0.

Supposons maintenant que f soit $\mathbb{C}$-dérivable en z_0. Alors nous avons

$$f'(z_0) = \lim_{w\to 0} \frac{f(z_0 + w) - f(z_0)}{w} = \sigma \in \mathbb{C}, \tag{12.845}$$

ce qui se récrit sous la forme

$$\lim_{w\to 0} \frac{f(z_0+w)-f(z_0)-\sigma w}{w} = 0. \tag{12.846}$$

Si nous posons $z_0 = x_0 + iy_0$, $w = h + ik$ et $\sigma = \alpha - i\beta$ nous avons

$$\lim_{(h,k)\to(0,0)} \left| \frac{F\big((x_0,y_0)+(h,k)\big) - F(x_0,y_0) - \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}}{|w|} \right| = 0, \tag{12.847}$$

ce qui signifie que F est différentiable et que sa différentielle est la matrice

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{12.848}$$ EqMLthLD

La matrice (12.848) est, vue dans $\mathbb{R}^2$, la matrice de multiplication dans $\mathbb{C}$ par $\alpha - i\beta = f'(z_0)$. En d'autre termes, dans $\mathbb{C}$ nous avons

$$df_{z_0}(z) = f'(z_0)z, \tag{12.849}$$

et en particulier la différentielle est donnée par

$$df_{z_0} = f'(z_0)dz. \tag{12.850}$$ EqPropZOkfmQ

□

Exemple 12.312 (Une application C^∞ mais pas $\mathbb{C}$-dérivable)**.**
Nous considérons la fonction

$$\begin{aligned} f\colon \mathbb{C} &\to \mathbb{C} \\ x+iy &\mapsto x. \end{aligned} \tag{12.851}$$

Vu que c'est une application linéaire, elle est différentiable une infinité de fois et sa différentielle est elle-même. C'est donc une application C^∞.

Elle n'est cependant pas $\mathbb{C}$-dérivable. En effet le quotient différentiel est, pour $\epsilon \in \mathbb{C}$:

$$\frac{f(x+iy+\epsilon_x+i\epsilon_y)-f(x+iy)}{\epsilon} = \frac{\epsilon_x}{\epsilon}. \tag{12.852}$$

Cela n'a pas de limite lorsque $\epsilon \to 0$. Pour voir cela nous invoquons la méthode des chemins du corolaire 12.225 avec les chemins $\epsilon_1(t) = t$ et $\epsilon_2(t) = it$. Dans le premier cas, le quotient différentiel vaut 1 pour tout t, tandis que dans le second il vaut zéro pour tout t. △

12.28.1 Quelques règles de calcul

LEMooVDXOooUyFHXZ

Lemme 12.313.
Si f et g sont deux fonctions holomorphes sur un ouvert $\Omega \subset \mathbb{C}$ et si g ne s'annule pas sur Ω, alors f/g est holomorphe sur Ω.

12.29 Théorèmes des accroissements finis

SecThoAccrsFinis

Nous avons déjà démontré (lemme 12.255) que si f est différentiable au point x alors $df_x(u) = \partial_u f(x)$. Une importante conséquence est le théorème des accroissements finis

val_medio_2

Théorème 12.314 (Accroissements finis, inégalité de la moyenne)**.**
Soit U un ouvert dans $\mathbb{R}^m$ et soit $f : U \to \mathbb{R}^n$ une fonction différentiable. Soient a et b deux points dans U, $a \neq b$, tels que le segment $[a,b]$ soit contenu dans U. Alors

$$\|f(b)-f(a)\|_n \leqslant \sup_{x\in[a,b]} \|df_x\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \|b-a\|_m. \tag{12.853}$$

Démonstration. On utilise le théorème 10.259 et le fait que

$$\|\partial_u f(x)\|_n \leqslant \|df_x\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|u\|_m,$$

pour tout u dans $\mathbb{R}^m$. □

La proposition suivante est une application fondamentale du théorème des accroissements finis 12.314.

PropAnnulationEtConstance

Proposition 12.315.
Soit U un ouvert connexe par arcs de $\mathbb{R}^m$ et une fonction $f\colon U \to \mathbb{R}^n$. Les conditions suivantes sont équivalentes :

ItemPropCstDiffZeroi

(1) f est constante ;

ItemPropCstDiffZeroii

(2) f est différentiable et $df(a) = 0$ pour tout $a \in U$;

ItemPropCstDiffZeroiii

(3) les dérivées partielles $\partial_1 f, \dots, \partial_m f$ existent et sont nulles sur U.

Démonstration. Nous allons démonter les équivalences en plusieurs étapes. D'abord (1) ⇒ (2), puis (2) ⇒ (3), ensuite (3) ⇒ (2) et enfin (2) ⇒ (1).

Commençons par montrer que la condition (1) implique la condition (2). Si $f(x)$ est constante, alors la condition (10.682) est vite vérifiée en posant $T(h) = 0$.

Afin de voir que la condition (2) implique la condition (3), remarquons d'abord que la différentiabilité de f implique que les dérivées partielles existent (proposition 12.252) et que nous avons l'égalité $df(a).u = \sum_i u_i \partial_i f(a)$ pour tout $u \in \mathbb{R}^m$ (lemme 12.255). L'annulation de $\sum_i u_i \partial_i f(a)$ pour tout u implique l'annulation des $\partial_i f(a)$ pour tout i.

Prouvons maintenant que la propriété (3) implique la propriété (2). D'abord, par le théorème 12.294, l'existence et la continuité des dérivées partielles $\partial_i f(a)$ implique la différentiabilité de f. Ensuite, la formule $df(a).u = \sum_i u_i \partial_i f(a)$ implique que $df(a) = 0$.

Il reste à montrer que (2) implique la condition (1), c'est-à-dire que l'annulation de la différentielle implique la constance de la fonction. C'est ici que nous allons utiliser le théorème des accroissements finis. En effet, si a et b sont des points de U, le théorème 12.314 nous dit que

$$\|f(b) - f(a)\|_n \leqslant \sup_{x\in[a,b]} \|df(x)\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|b - a\|_m. \tag{12.854}$$

Mais $\|df(x)\| = 0$ pour tout $x \in U$, donc ce supremum est nul et $f(b) = f(a)$, ce qui signifie la constance de la fonction. □

12.30 Fonctions Lipschitziennes

DEFooQHVEooDbYKmz

Définition 12.316.
Soient (E, d_E) et (F, d_F) deux espaces métriques[130]*, $f\colon E \to F$ une application et un réel k strictement positif. Nous disons que f est **Lipschitzienne** de constante k sur E si pour tout $x, y \in E$,*

$$d_F\big(f(x), f(y)\big) \leqslant k d_E(x, y). \tag{12.855}$$

Soit f une fonction k-Lipschitzienne. Si $y \in \overline{B(x,\delta)}$ alors $\|x-y\| \leqslant \delta$ et donc $\big\|f(x)-f(y)\big\| \leqslant k\delta$. Cela signifie que la condition Lipschitz peut s'énoncer en termes de boules fermées par

$$f\big(\overline{B(x,\delta)}\big) \subset \overline{B\big(f(x), k\delta\big)} \tag{12.856}$$

EqDZytUbn

tant que $\overline{B(x,\delta)}$ est contenue dans le domaine sur lequel f est Lipschitz.

130. Pour rappel, les espaces métriques sont définis par la définition 7.115 et le théorème 7.117 ; je précise que nous ne supposons pas que E soit vectoriel ; en particulier il peut être un ouvert de $\mathbb{R}^n$.

Proposition 12.317.
Soit U un ouvert convexe de $\mathbb{R}^m$, et soit $f : U \to \mathbb{R}^n$ une fonction différentiable. La fonction f est Lipschitzienne sur U si et seulement si df est bornée sur U.

Démonstration. Le fait que l'application différentielle df soit bornée signifie qu'il existe un $M > 0$ dans $\mathbb{R}$ tel que $\|df_a\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leqslant M$, pour tout a dans U. Si cela est le cas, alors le théorème 12.314 et la convexité [131] de U impliquent évidemment que f est de Lipschitz de constante plus petite ou égale à M.

Inversement, si f est Lipschitz de constante k, alors pour tout a dans U et u dans $\mathbb{R}^m$ on a

$$\left\| \frac{f(a+tu) - f(a)}{t} \right\|_n \leqslant k\|u\|_m,$$

En passant à la limite pour $t \to 0$ on a

$$\|\partial_u f(a)\|_n = \|df_a(u)\|_n \leqslant k\|u\|_m,$$

donc la norme de df_a est majorée par k pour tout a dans U. □

Notez cependant qu'une fonction peut être Lipschitzienne sans être différentiable.

PropFZgFTEW

Proposition 12.318.
Une fonction Lipschitzienne $f : \mathbb{R} \to \mathbb{R}$ est continue.

Démonstration. Nous utilisons la caractérisation de la continuité donnée par le théorème 7.205. Prouvons donc la continuité en $a \in \mathbb{R}$. Pour tout x nous avons

$$\big|f(x) - f(a)\big| \leqslant k|x - a|. \tag{12.857}$$

Si $\epsilon > 0$ est donné, il suffit de prendre $\delta < \frac{\epsilon}{k}$ pour avoir

$$\big|f(x) - f(a)\big| \leqslant k\frac{\epsilon}{k} = \epsilon. \tag{12.858}$$

Donc f est continue en a. □

DefJSFFooEOCogV

Définition 12.319.
Une fonction

$$\begin{aligned} f : \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^p \\ (t, y) &\mapsto f(t, y) \end{aligned} \tag{12.859}$$

*est **localement Lipschitz** en y au point (t_0, y_0) si il existe des voisinages V de t_0 et W de y_0 et un nombre $k > 0$ tels que pour tout $(t, y) \in V \times W$ on ait*

$$\big\|f(t_0, y_0) - f(t, y)\big\| \leqslant k\|y - y_0\|. \tag{12.860}$$

La fonction est localement Lipschitz sur un ouvert U de $\mathbb{R}^n \times \mathbb{R}^m$ si elle est localement Lipschitz en chaque point de U.

NORMooYNRAooBgobcK

12.320.
Autrement dit, une fonction est localement Lipschitzienne en sa deuxième variable lorsque tout point admet un voisinage sur lequel elle est Lipschitzienne.

PROPooVZSAooUneOQK

Proposition 12.321.
Toute application Lipschitz [132] est uniformément continue [133].

131. La convexité de U sert à assurer que la droite reliant a à b est contenue dans U ; c'est ce que nous utilisons dans la démonstration du théorème 12.314.
132. Définition 12.316.
133. Définition 7.323.

Démonstration. Soit une application k-lipschitzienne $f\colon E \to F$ entre deux espaces métriques. Soient $\epsilon > 0$ ainsi que $a \in E$. Nous considérons $\delta = \epsilon/k$. Si $x \in B(a,\delta)$ nous avons

$$d_F\big(f(a),f(x)\big) \leqslant kd_E(a,x) \leqslant k\delta \leqslant \epsilon. \tag{12.861}$$

□

PropGIBZooVsIqfY

Proposition 12.322.
Si f et g sont deux fonctions localement Lipschitz alors $f+g$ l'est.

Démonstration. Il s'agit d'un simple calcul avec une majoration standard :

$$\begin{aligned}\|(f+g)(t_0,y_0)-(f+g)(t,y)\| &\leqslant \|f(t_0,y_0)-f(t,y)\|+\|g(t_0,y_0)-g(t,y)\| &\text{(12.862a)}\\ &\leqslant k_f\|y-y_0\|+k_g\|y-y_0\| &\text{(12.862b)}\\ &= (k_f+k_g)\|y-y_0\|. &\text{(12.862c)}\end{aligned}$$

□

LemCFZUooVqZmpc

Lemme 12.323.
La fonction donnée par

$$f(t,(x,y)) = xy \tag{12.863}$$

est localement Lipschitz en tout point.

Démonstration. Nous avons la majoration classique

$$|f\big(t,(x_0,y_0)\big)-f\big(t,(x,y)\big)| = |x_0y_0-xy| \leqslant |x_0y_0-x_0y|+|x_0y-xy| \leqslant |x_0||y_0-y|+|y||x_0-x|. \tag{12.864}$$

Vu que nous parlons de fonction *localement Lipschitzienne*, nous pouvons majorer $|y|$ et $|x_0|$ par un même nombre k dans un voisinage de (x_0,y_0). Cela donne

$$|f\big(t,(x_0,y_0)\big)-f\big(t,(x,y)\big)| \leqslant k\big(|y_0-y|+|x_0-x|\big) \leqslant \sqrt{2}k\| \begin{pmatrix} x_0-x\\ y_0-y\end{pmatrix}\|. \tag{12.865}$$

Nous avons utilisé l'équivalence de norme de la proposition 10.47(1). □

12.31 Différentielles d'ordre supérieur

SecDiffOrdSup

12.31.1 Différentielle et dérivées partielles

C'est le moment de relire les définitions 10.261 et 10.262.

Définition 12.324.
Une sorte de projection. Pour $u \in V$ nous définissons

$$\begin{aligned}\texttt{proj}_u\colon \mathcal{L}^n(V,E) &\to \mathcal{L}^{n-1}(V,E)\\ \texttt{proj}_u(\omega)(v_1,\dots,v_{n-1}) &= \omega(u,v_1,\dots,v_{n-1}).\end{aligned} \tag{12.866}$$

LEMooSMZQooJBVySP

Lemme 12.325 ([1, 410])**.**
À propos de dimensions[134]*,*

ITEMooUWEBooSzFseN

(1) $\dim(E_n) = \dim(V)^n \dim(E)$.

ITEMooFMKQooFSMpgF

(2) $\dim\big(\mathcal{L}^n(V,E)\big) = \dim(E)\dim(V)^n$.

134. E_n, définition 10.262.

LEMooTDLNooTcPkLg

Lemme 12.326 ([1]).
Nous définissons les ϕ_k par récurrence. D'abord

$$\begin{aligned}\phi_1\colon \mathcal{L}^n(V,E) &\to \mathcal{L}\big(V,\mathcal{L}^{n-1}(V,E)\big)\\ \phi_1(\omega)u &= \texttt{proj}_u(\omega),\end{aligned} \tag{12.867}$$

et ensuite

$$\begin{aligned}\phi_k\colon \mathcal{L}^n(V,E) &\to \mathcal{L}^{n-k}(V,E)_k\\ \phi_k(\omega)u &= \phi_{k-1}\big(\texttt{proj}_u(\omega)\big).\end{aligned} \tag{12.868}$$

Les applications ϕ_k sont bijectives.

Démonstration. Nous prouvons par récurrence.

(1) **Injective, $k=1$** Soit $\omega\in\mathcal{L}^n(V,E)$ tel que $\phi_1(\omega)=0$. Pour tout $u\in V$ nous avons $\phi_1(\omega)u=0$, ce qui signifie que $\texttt{proj}_u(\omega)=0$ ou encore que pour tout $v_1,\dots,v_{n-1}\in V$ nous avons $\omega(u,v_1,\dots,v_{n-1})=0$. Nous avons donc bien $\omega=0$.

(2) **Surjective, $k=1$** Soit $\alpha\in\mathcal{L}\big(V,\mathcal{L}^{n-1}(V,E)\big)$. Nous allons construire $\omega\in\mathcal{L}^n(V,E)$ tel que $\phi_1(\omega)=\alpha$. Nous posons

$$\omega(v_0,\dots,v_{n-1})=\alpha(v_0)(v_1,\dots,v_{n-1}). \tag{12.869}$$

Avec lui nous avons bien $\texttt{proj}_{v_0}(\omega)=\alpha(v_0)$.

(3) **Injective, $k=k$** Nous supposons que $\phi_k(\omega)=0$, c'est-à-dire que pour tout $u\in V$ nous avons

$$\phi_{k-1}\big(\texttt{proj}_u(\omega)\big)=0. \tag{12.870}$$

Cela implique $\texttt{proj}_u(\omega)=0$ parce que ϕ_{k-1} est injective par hypothèse de récurrence. Nous déduisons que $\omega=0$, et que ϕ_k est injective.

(4) **Surjective, $k=k$** Nous allons montrer que $\phi_k\colon\mathcal{L}^n(V,E)\to\mathcal{L}^{n-k}(V,E)_k$ est une application linéaire injective entre deux espaces de même dimension.
Il s'agit d'utiliser le lemme 12.325. D'abord

$$\dim\big(\mathcal{L}^{n-k}(V,E)_k\big)=\dim(V)^k\dim\big(\mathcal{L}^{n-k}(V,E)\big)=\dim(V)^n\dim(E). \tag{12.871}$$

Ensuite $\dim\big(\mathcal{L}^n(V,E)\big)=\dim(V)^n\dim(E)$. Le compte est bon.
Le théorème du rang (formule (4.50)) avec $\dim\big(\ker(\phi_k)\big)=0$ nous dit que le rang de ϕ_k est maximal et donc que ϕ_k est surjective.

□

LEMooLTYCooAoMJKD

Lemme 12.327.
Soit $w\in W$. Nous avons

$$\big[w\times_k\phi_k(\omega_I)\big]\times_1\omega_l=w\times_{k+1}\phi_{k+1}(\omega_{l,I}). \tag{12.872}$$

Démonstration. Il s'agit d'appliquer les deux membres à un élément $v\in V$. D'abord

$$\begin{aligned}\big[w\times_k\phi_k(\omega_I)\big]\times_1\omega_l(v)&=\omega_l(v)w\times_k\phi_k(\omega_I) &\text{(12.873a)}\\ &=w\times_k\big(\phi_k(\omega_I)v_l\big) &\text{(12.873b)}\end{aligned}$$

SUBEQooMWPMooYAsYCR

Pour (12.873b) nous avons utilisé le fait que $\omega_{l,I}(v)=v_l\in\mathbb{R}$ et que le produit $\times_k$ est linéaire. De l'autre côté, en utilisant les définitions de $\times_k$ et de ϕ_k, nous avons

$$\begin{aligned}\big[w\times_{k+1}\phi_{k+1}(\omega_{l,I})\big]v&=w\times_k\big(\phi_{k+1}(\omega_{l,I})v\big) &\text{(12.874a)}\\ &=w\times_k\phi_k\big(\texttt{proj}_v(\omega_{l,I})\big) &\text{(12.874b)}\\ &=w\times_k\phi_k(v_l\omega_I) &\text{(12.874c)}\end{aligned}$$

parce que $\texttt{proj}_v(\omega_{l,I})=v_l\omega_I$. □

PROPooVGRRooHSwcPl

Proposition 12.328 ([1]).
Soit une application $f\colon V \to E$ de classe C^n. Alors

(1) pour tout n-multiindice I, la dérivée partielle $\partial_I^n f$ existe,

(2) nous avons la formule

$$(d^n f)_a = \sum_{i_1,\dots,i_n} \frac{\partial^n f}{\partial x_{i_1} \dots \partial x_{i_n}}(a) \times_n \phi_n(\omega_{i_1,\dots,i_n}) \tag{12.875}$$

où la somme porte sur tous les multiindices de taille n, le produit $\times_n$ est la définition 10.264 et l'application ϕ_n est donnée dans l'énoncé du lemme 12.326.

Démonstration. Nous faisons tout cela par récurrence. Le cas $n = 1$ est déjà fait par le théorème 12.294 parce que $\phi_1(\omega) = \omega$. Nous supposons que le résultat est démontré pour une valeur k, et nous considérons f de classe C^n avec $n \geqslant k + 1$. L'hypothèse de récurrence dit que

EQooBOZUooCiFIYR
$$(d^k f)_a = \sum_I (\partial_I f)(a) \times_k \phi_k(\omega_I) \tag{12.876}$$

est continuement différentiable par rapport à a. Nous pouvons donc utiliser la formule (12.792) pour calculer la différentielle de (12.876) :

$$\begin{aligned}
(d^{k+1} f)_a &= \sum_l \frac{\partial (d^k f)}{\partial x_l}(a) \times_1 \omega_l && \text{(12.877a)}\\
&= \sum_l \sum_I \big[(\partial_l \partial_I f)(a) \times_k \phi_k(\omega_I)\big] \times_1 \omega_l && \text{(12.877b)}\\
&= \sum_{l,I} (\partial_{l,I} f)(a) \times_{k+1} \phi_{k+1}(\omega_{l,I}) && \text{(12.877c)}
\end{aligned}$$

SUBEQooOXUHooFDJiow

où nous avons utilisé le lemme 12.327 pour obtenir (12.877c). □

PROPooKOBVooQhrAeJ

Proposition 12.329.
Soit une application $f\colon V \to E$ telle que les dérivées partielles $\frac{\partial f}{\partial x_{i_1} \dots \partial x_{i_n}}$ existent et sont continues sur un ouvert $\mathcal{U}$. Alors f est de classe C^n sur $\mathcal{U}$.

Démonstration. Pour $n = 1$, cette proposition est déjà contenue dans le théorème 12.294. Nous pouvons donc directement passer au pas de récurrence. Nous définissons les objets

$$T_k(a) = \sum_I (\partial_I f)(a) \times_k \phi_k(\omega_I) \tag{12.878}$$

où la somme porte sur tous les k-multiindices. Ces applications T_k existent et sont continues par hypothèse.

Le théorème 12.294 nous indique que $T_1(a) = df_a$. Nous supposons maintenant que $(d^k f)_a = T_k(a)$ et nous allons prouver que T_{k+1} est la différentielle de $d^k f$. Pour cela nous devons prouver que

EQooPKXDooWWVjkT
$$\lim_{h \to 0} \frac{T_k(a+h) - T_k(a) - T_{k+1}(a)h}{\|h\|} = 0. \tag{12.879}$$

En utilisant le lemme 12.327 nous pouvons dégrossir le terme $T_{k+1}(a)h$:

$$\begin{aligned}
T_{k+1}(a)h &= \sum_{l,I} \big[(\partial_{l,I} f)(a) \times_{k+1} \phi_{k+1}(\omega_{l,I})\big]h && \text{(12.880a)}\\
&= \sum_{l,I} \big[(\partial_{l,I} f)(a) \times_k \phi_k(\omega_I)\big] \times_1 \omega_l(h) && \text{(12.880b)}\\
&= \sum_{l,I} h_l (\partial_{l,I} f)(a) \times_k \phi_k(\Omega_I) && \text{(12.880c)}\\
&= \sum_l h_l \big[\partial_l \sum_I (\partial_I f)(a)\big] \times_k \phi_k(\omega_I). && \text{(12.880d)}
\end{aligned}$$

SUBEQooACCZooXrxHpY

Vu que $\partial_I f$ est une fonction dont les dérivées partielles existent et sont continues, le théorème 12.294 (toujours lui) nous dit qu'elle est différentiable et que

$$d(\partial_I f)_a(h) = \sum_l h_l \partial_l(\partial_I f)(a). \tag{12.881}$$

En mettant la somme sur I tout devant dans (12.880d), nous trouvons

$$T_{k+1}(a)h = \sum_I d(\partial_I f)_a(h) \times_k \phi_k(\omega_I). \tag{12.882}$$

C'est pas beau la vie ? Nous pouvons écrire le numérateur du quotient différentiel (12.879) :

$$T_k(a+h) - T_k(a) - T_{k+1}(a)h = \sum_I \big[(\partial_I f)(a+h) - (\partial_I f)(a)\big] \times_k \phi_k(\omega_I) \tag{12.883a}$$

$$- \sum_I d(\partial_I f)_a(h) \times_k \phi_k(\omega_I) \tag{12.883b}$$

$$= \sum_I \big[(\partial_I f)(a+h) - (\partial_I f)(a) - d(\partial_I f)(a)h\big] \times_l \phi_k(\omega_I). \tag{12.883c}$$

Nous utilisons le fait que l'application $w \mapsto w \times_k \phi_k(\omega_I)$ est linéaire et commute avec la limite, c'est-à-dire que

$$\lim_{h\to 0} \big(w(h) \times_k \phi_k(\omega_I)\big) = \big(\lim_{h\to 0} w(h)\big) \times_k \phi_k(\omega_I). \tag{12.884}$$

Et bien sûr, la somme sur les multiindices I commute également avec la limite. Donc

$$\lim_{h\to 0} \frac{T_k(a+h) - T_k(a) - T_{k+1}(a)h}{\|h\|} = \sum_I \left(\lim_{h\to 0} \frac{(\partial_I f)(a+h) - (\partial_I f)(a) - d(\partial_I f)_a h}{\|h\|}\right) \times_k \phi_k(\omega_I) \tag{12.885a}$$

$$= 0. \tag{12.885b}$$

Cela prouve que T_{k+1} est bien la différentielle de T_k. Par récurrence, la fonction f est bien n fois continuement différentiable. □

Maintenant que nous avons plein de lemmes et de résultats, il est facile de démontrer un très gros résultat en peu de lignes.

THOooPZTAooTASBhZ

Théorème 12.330.
Soit une application $f\colon V \to E$ entre deux espaces vectoriels normés de dimension finies. Nous avons équivalence entre

ITEMooBOWTooXgxhpS

(1) f est de classe C^n

ITEMooPVZHooHihSRD

(2) les dérivées partielles $\partial_i f$ sont de classe C^{n-1}

ITEMooVBQMooBleazN

(3) les dérivées partielles $\partial_I^n f$ existent et sont continues pour tout multiiindice de longueur n.

Démonstration. En plusieurs implications.

(1) **(3) implique (1)** C'est la proposition 12.329.

(2) **(1) implique (3)** C'est la proposition 12.328.

(3) **(3) implique (2)** En posant $g_i = \partial_i f$, l'hypothèse est que les $\partial_I^{n-1} g_i$ existent et sont continues pour tout multiindices I de longueur $n-1$. En appliquant « (3) implique (1) » à la fonction g_i, la fonction g_i est de classe C^{n-1}.

(4) **(2) implique (3)** Les fonctions g_i déjà définies sont de classe C^{n-1}. Nous leur appliquons « (1) implique (3) », nous savons que les fonctions $\partial_I^{n-1} g_i$ sont continues, c'est-à-dire que les $\partial_I \partial_i f$. Vu que tout multiiindice de longueur n peut être écrit sous la forme iI où I est de longueur $n-1$, nous avons prouvé que les $\partial_J^n f$ existent et sont continues pour tout multiindice J de longueur n.

□

PROPooCKROooYutfEX

Proposition 12.331 ([1]).
Soit un espace vectoriel V de dimension finie n sur K ($= \mathbb{C}$ ou $\mathbb{R}$). Soit une base $\{e_i\}_{i\in I}$ de V. Pour chaque i, la projection

$$\begin{aligned} \texttt{proj}_i\colon V &\to \mathbb{K} \\ \sum_k e_k &\mapsto a_i \end{aligned} \tag{12.886}$$

est de classe C^∞.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

PROPooMOMMooDYIayv

Proposition 12.332 ([1]).
Soient deux espaces vectoriels normés de dimension finie V et W. Soient des bases $\{e_i\}$ et $\{e'_j\}$ de V et W. Une application $f\colon V \to W$ est de classe C^p si et seulement si chacune de ses composantes le sont, c'est-à-dire si en écrivant

$$f(s) = \sum_j f_j(s)e'_j, \tag{12.887}$$

les applications $f_j\colon V \to \mathbb{R}$ sont de classe C^p.

Démonstration. En deux parties.

(1) **Sens direct** Considérant la proposition 12.331, la projection

$$\begin{aligned} \texttt{proj}_i\colon W &\to \mathbb{R} \\ \sum_k a_k e'_k &\mapsto a_i \end{aligned} \tag{12.888}$$

est de classe C^∞. Vu que f est de classe C^p, la composée $\texttt{proj}_i \circ f$ est de classe C^p par la proposition 10.280.

(2) **Sens réciproque** Si I est un multiindice de longueur p, nous avons

$$(\partial_I^p f)(s) = \sum_k (\partial_I^p f_k)(s)e'_k, \tag{12.889}$$

qui est continue parce que les f_k sont de classe C^p. Donc f est de classe C^p par le théorème 12.330(3).

□

12.31.2 Application à valeurs dans les opérateurs

PROPooUCRNooaqsebj

Proposition 12.333 ([1]).
Soient des espaces vectoriels normés V et W. Soient un ouvert $U \subset \mathbb{R}^n$ et une application $A\colon U \to \mathrm{End}(V, W)$. L'application A est de classe C^k si et seulement si toutes les applications $s \mapsto A(s)_{ij}$ sont de classe C^k.

Démonstration. Nous utilisons les applications f_{ij} de la proposition 4.77, en nous souvenant que les éléments de matrice de $A(s)$ sont les coordonnées de $A(s) \in \mathrm{End}(V, W)$ dans ces coordonnées (proposition 4.78). Le résultat est alors seulement un cas particulier de la proposition 12.332. □

PROPooYORGooDmwrJg

Proposition 12.334 ([1]).
Soient des espaces vectoriels normés V et W. Soient un ouvert $U \subset \mathbb{R}^n$ et une application $A\colon U \to \mathrm{End}(V, W)$. Nous supposons que pour chaque $v \in V$, l'application $s \mapsto A(s)v$ est C^k.

Alors A est C^k.

Démonstration. Soient une base $\{e_i\}$ de V et $\{e'_j\}$ de W. Nous considérons les applications

$$\begin{aligned} f_i\colon U &\to W \\ s &\mapsto A(s)e_i. \end{aligned} \tag{12.890}$$

Par hypothèse, f_i est une application C^k. Nous avons

$$f_i(s) = A(s)e_i = \sum_j A(s)_{ji}e'_j \tag{12.891}$$

Nous profitons de la proposition 12.331 qui dit que la projection est de classe C^∞, et de la proposition 10.280 à propos des composées, l'application $(\mathtt{proj}_j \circ f_i)$ est de classe C^p. Or nous avons

$$(\mathtt{proj}_j \circ f_i)(s) = A(s)_{ji}. \tag{12.892}$$

La proposition 12.333 conclut que f est de classe C^p. □

12.31.3 Espaces d'applications multilinéaires et identifications

Si V et E sont des espaces vectoriels de dimensions finies, la différentielle de $f\colon V \to E$ est une application $df\colon V \to \mathcal{L}(V,E)$. La différentielle seconde est une application $d(df)\colon V \to \mathcal{L}\big(V,\mathcal{L}(V,E)\big)$ et ainsi de suite.

Une grande difficulté de la manipulation des différentielles d'ordre supérieurs provient de cet emboîtement d'espaces d'applications linéaires. Nous nous attaquons à présent à la description de ces espaces emboîtés [135].

Pour la suite, nous considérerons des espaces vectoriels normés V et E de dimension finie. Nous notons $\mathcal{L}^n(V,E)$ l'espace des applications multilinéaires $V^n \to E$.

Nous introduisons le produit suivant[1] :

$$\begin{aligned} \cdot\ : W \times \mathcal{L}\big(V,\mathcal{L}(V,\mathbb{R})\big) &\to \mathcal{L}\big(V,\mathcal{L}(V,W)\big) \\ \big((w\ \cdot\ \psi)(u)\big)v &= \big(\psi(u)v\big)w. \end{aligned} \tag{12.893}$$

Dans la suite, pour économiser des parenthèses et des maux de tête, nous allons noter $\psi(u,v)$ le nombre $\psi(u)v$. Il n'est cependant pas question de dire que ψ est une application bilinéaire. En effet, identifier $\mathcal{L}\big(V,\mathcal{L}(V,W)\big)$ à l'espace des applications bilinéaires $V \times V \to W$ ne sert pas à grand chose pour l'instant parce qu'une telle identification a le prix de devoir prouver que toutes les propriétés des différentielles passent à travers l'identification, tâche qui est à priori (conservation de la difficulté) de la même nature que celle à laquelle nous nous attachons à présent.

LEMooHCUSooXYHuBR

Lemme 12.335 ([1]).
Soient deux espaces vectoriels normés V et E ainsi que $\psi \in \mathcal{L}\big(V,\mathcal{L}(V,\mathbb{R})\big)$. Pour tout $a \in E$ nous avons

$$\|a\ \cdot\ \psi\|_{\mathcal{L}\left(V,\mathcal{L}(V,E)\right)} = \|\psi\|_{\mathcal{L}\left(V,\mathcal{L}(V,\mathbb{R})\right)}\|a\|_E. \tag{12.894}$$

Démonstration. Il s'agit seulement d'un calcul :

$$\|a\ \cdot\ \psi\| = \sup_{\|v\|=1} \|(a\ \cdot\ \psi)(v)\|_{\mathcal{L}(V,E)} \tag{12.895a}$$

$$= \sup_{\|v\|=1} \sup_{\|w\|=1} \|(a\ \cdot\ \psi)(v)w\|_E \tag{12.895b}$$

$$= \sup_{\|v\|=1} \sup_{\|w\|=1} \|\big(\psi(v)w\big)a\|_E \tag{12.895c}$$

SUBEQooDVSVooFkgDQb

$$= \sup_{\|v\|=1} \sup_{\|w\|=1} |\psi(v)w|\,\|a\|_E \tag{12.895d}$$

SUBEQooBJQDooDyZMQy

$$= \sup_{\|v\|=1} \|\psi(v)\|\,\|a\| \tag{12.895e}$$

$$= \|\psi\|\|a\|. \tag{12.895f}$$

135. Toutes les constructions, tous les énoncés et les preuves qui suivent sont de l'invention personnelle de l'auteur de ces lignes. Je n'ai trouvé nulle part une source qui s'attaque réellement à la récurrence.

Notez que dans (12.895c), $|\psi(v)w|$ est un simple réel ; c'est pourquoi nous le retrouvons hors de la norme $\|.\|_E$ dans (12.895d), muni d'une simple valeur absolue. □

Lemme 12.336 ([1]).
Soient $\psi \in \mathcal{L}\big(V, \mathcal{L}(V, \mathbb{R})\big)$ et une fonction continue $f : V \to E$. Alors la fonction

$$\begin{aligned} g : V &\to \mathcal{L}\big(V, \mathcal{L}(V, E)\big) \\ x &\mapsto f(x) \cdot \psi \end{aligned} \tag{12.896}$$

est continue.

Démonstration. Pour des raisons de notations, nous allons écrire g_x pour $g(x)$. Cela étant dit nous considérons $a \in E$, une suite $x_k \xrightarrow{E} a$ et nous calculons :

$$\|g_a - g_{x_k}\| = \sup_{\|u\|=1} \|g_a(u) - g_{x_k}(u)\|_{\mathcal{L}(V,E)} \tag{12.897a}$$

$$= \sup_{\|u\|=1} \sup_{\|v\|=1} \|g_a(u)v - g_{x_k}(u)v\|_E \tag{12.897b}$$

$$= \sup_{\|u\|=1} \sup_{\|v\|=1} \|f(a)\psi(u)v - f(x_k)\psi(u)v\| \tag{12.897c}$$

$$= \sup_{\|u\|=1} \sup_{\|v\|=1} \|f(a) - f(x_k)\|\big|\psi(u)v\big| \tag{12.897d}$$

$$= \big\|f(a) - f(x_k)\big\| \sup_{u,v} |\psi(u)v| \tag{12.897e}$$

$$= \|f(a) - f(x_k)\| \, \|\psi\|. \tag{12.897f}$$

En prenant la limite $k \to \infty$, et en considérant que f est continue en a, nous obtenons

$$\lim_{k\to\infty} \|g_a - g_{x_k}\| = 0. \tag{12.898}$$

□

Lemme 12.337 ([1]).
Soient une application différentiable $f : V \to E$ ainsi que $\psi \in \mathcal{L}\big(V, \mathcal{L}(V, \mathbb{R})\big)$. Soit

$$\begin{aligned} g : V &\to \mathcal{L}\big(V, \mathcal{L}(V, E)\big) \\ x &\mapsto f(x) \cdot \psi. \end{aligned} \tag{12.899}$$

Alors g est différentiable et pour tout $a \in V$ nous avons

$$dg_a(h) = df_a(h) \cdot \psi. \tag{12.900}$$

Notons que nous n'avons pas $dg_a = df_a \cdot \psi$. En effet, $df_a \in \mathcal{L}(V, W)$, de telle sorte que $df_a \cdot \psi \in \mathcal{L}\big(V, \mathcal{L}\big(V, \mathcal{L}(V, W)\big)\big)$. Les espaces ne s'emboîtent pas dans le bon ordre.

Démonstration. Il s'agit de vérifier que $h \mapsto df_a(h) \cdot \psi$ vérifie la condition de la définition 10.243. En utilisant le fait que $(a + b) \cdot \psi = a \cdot \psi + b \cdot \psi$ ainsi que le lemme 12.335 nous écrivons

$$\frac{\|f(a+h) \cdot \psi - f(a) \cdot \psi - df_a(h) \cdot \psi\|}{\|h\|} = \|\frac{f(a+h) - f(a) - df_a(h)}{h} \cdot \psi\| \tag{12.901a}$$

$$= \|\frac{f(a+h) - f(a) - df_a(h)}{h}\|\|\psi\|. \tag{12.901b}$$

Vu que f est différentiable en a et que df_a est la différentielle, nous avons bien

$$\lim_{h\to 0} \|\frac{f(a+h) - f(a) - df_a(h)}{h}\|\|\psi\| = 0. \tag{12.902}$$

□

Démonstration. Nous faisons (1) par récurrence. D'abord $\dim(E_0) = \dim(E)$ et ensuite

$$\dim(E_{k+1}) = \dim \mathcal{L}(V, E_k) = \dim(V) \dim(E_k) = \dim(V)^{n+1} \dim(E). \tag{12.903}$$

Pour (2), si $\{e_i\}$ est une base de V, un élément $\omega \in \mathcal{L}^n(V, E)$ est déterminé par les valeurs de $\omega(e_{i_1}, \dots, e_{i_n})$ qui peuvent être n'importe quel vecteur de E. Donc la dimension est $\dim(V)^n \dim(E)$. □

Lemme 12.338 ([1]).
Soit $n \in \mathbb{N}$ nous définissons par récurrence

$$\begin{aligned} \psi_{n,0} \colon E_n &\to E_n \\ \alpha &\mapsto \alpha. \end{aligned} \tag{12.904}$$

et

$$\begin{aligned} \psi_{n,k} \colon E_n &\to \mathcal{L}^k(V, E_{n-k}) \\ \psi_{n,k}(\alpha)(v_1, \dots, v_k) &= \big(\psi_{n,k-1}(\alpha)(v_1, \dots, v_{k-1})\big)v_k. \end{aligned} \tag{12.905}$$

L'application

$$\psi_{n,n} \colon E_n \to \mathcal{L}^n(V, E) \tag{12.906}$$

est un isomorphisme isométrique.

Démonstration. Nous allons démontrer par récurrence sur k que tous les $\psi_{n,k}$ sont des isomorphismes isométriques. Pour $k = 0$ c'est évident parce que $\psi_{n,0}$ est l'identité.

(1) **Injective** Soient $\alpha \in E_n$ tels que $\psi_{n,k+1}(\alpha) = 0$. Cela signifie que pour tout $v_1, \dots, v_{k+1}$ nous avons

$$\psi_{n,k}(\alpha)(v_1, \dots, v_k)v_{k+1} = 0. \tag{12.907}$$

c'est-à-dire $\psi_{n,k}(\alpha)(v_1, \dots, v_k) = 0$. Vu que $\psi_{n,k}$ est injective (hypothèse de récurrence), nous avons $\alpha = 0$.

(2) **Surjective** Soit $\omega \in \mathcal{L}^{k+1}(V, E_{n-k})$; nous cherchons $\alpha \in E_n$ tel que $\psi_{n,k+1}(\alpha) = \omega$, c'est-à-dire tel que

$$\psi_{n,k+1}(\alpha)(v_1, \dots, v_{k+1}) = \psi_{n,k}(\alpha)(v_1, \dots, v_k)v_{k+1} = \omega(v_1, \dots, v_{k+1}). \tag{12.908}$$

EQooACUOooKCKame

Notez que

$$\psi_{n,k}(\alpha) \in \mathcal{L}^k(V, E_{n-k}) = \mathcal{L}^k\big(V, \mathcal{L}(V, E_{n-k-1})\big). \tag{12.909}$$

En considérant $\sigma \in \mathcal{L}^k\big(V, \mathcal{L}(V, E_{n-k-1})\big)$ donné par

$$\sigma(v_1, \dots, v_k)v_{k+1} = \omega(v_1, \dots, v_{k+1}), \tag{12.910}$$

il existe (hypothèse de récurrence sur k) un $\alpha \in E_n$ tel que $\psi_{n,k}(\alpha) = \sigma$.
Pour ce α, la condition (12.908) est satisfaite.

(3) **Isométrique** Encore une fois par récurrence. Soit $\alpha \in E_n$. Nous avons

$$\begin{aligned} \|\psi_{n,k}(\alpha)\|_{\mathcal{L}^k(V,E_{n-k})} &= \sup_{\|v_i\|=1} \|\psi_{n,k}(\alpha)(v_1, \dots, v_k)\|_{E_{n-k}} && \text{(12.911a)} \\ &= \sup_{\|v_i\|=1} \|\psi_{n,k-1}(\alpha)(v_1, \dots, v_{k-1})v_k\|_{E_{n-k}} && \text{(12.911b)} \\ &= \sup_{\substack{\|v_i\|=1 \\ i=1,\dots,k-1}} \|\psi_{n,k-1}(v_1, \dots, v_{k-1})\|_{\mathcal{L}(V,E_{n-k-1})} && \text{(12.911c)} \\ &= \|\psi_{n,k-1}(\alpha)\| && \text{(12.911d)} \\ &= \|\alpha\|. && \text{(12.911e)} \end{aligned}$$

La dernière égalité est l'hypothèse de récurrence. Notez la subtile utilisation du lemme 7.147 qui permet de donner un sens aux supremums, grace au fait que $\{v \in V \text{ tel que } \|v\| = 1\}$ est compact.

□

LEMooFBEGooCqrzxH

Lemme 12.339.
Nous avons $\psi_{n,n} = \phi_n^{-1}$.

Démonstration. Pour nous échauffer nous posons $\omega \in \mathcal{L}^n(V, \mathbb{R})$, et nous calculons : EQooPBQIooUValDA

$$\begin{aligned}
\psi_{n,2}\big(\phi_n(\omega)\big)(v_1, v_2) &= \Big(\psi_{n,1}\big(\phi_n(\omega)\big)v_1\Big)v_2 && \text{(12.912a)}\\
&= \big(\phi_n(\omega)v_1\big)v_2 && \text{(12.912b)}\\
&= \phi_{n-1}\big(\texttt{proj}_{v_1}(\omega)\big) && \text{(12.912c)}\\
&= \phi_{n-2}\big(\texttt{proj}_{v_2}\texttt{proj}_{v_1}(\omega)\big). && \text{(12.912d)}
\end{aligned}$$

Cela étant dit, nous allons prouver ceci par récurrence :

$$\psi_{n,k}\big(\phi_n(\omega)\big)(v_1, \ldots, v_k) = \phi_{n-k}\Big(\prod_{i=1}^{k}\texttt{proj}_{v_i}(\omega)\Big). \tag{12.913}$$

Notez l'ordre du produit des projections. En ce qui concerne cet ordre, pour fixer les idées voici un exemple :

$$\texttt{proj}_{v_2}\texttt{proj}_{v_1}(\omega) = \big(\texttt{proj}_{v_1}(\omega)\big)(v_2) = \omega(v_1, v_2). \tag{12.914}$$

Faisons maintenant la récurrence.

(1) **Pour $k = 2$** C'est le calcul (12.912).

(2) **Pour $k + 1$** C'est encore un calcul, en faisant attention à l'ordre dans lequel viennent les projections :

$$\begin{aligned}
\psi_{n,k+1}\big(\phi_n(\omega)\big)(v_1, \ldots, v_{k+1}) &= \Big(\psi_{n,k}\big(\phi_n(\omega)\big)(v_1, \ldots, v_k)\Big)v_{k+1} && \text{(12.915a)}\\
&= \Big(\phi_{n-k}\big(\prod_{i=1}^{k}\texttt{proj}_{v_i}(\omega)\big)\Big)v_{k+1} && \text{(12.915b)}\\
&= \phi_{n-k-1}\big(\texttt{proj}_{v_{k+1}}\prod_{i=1}^{k}\texttt{proj}_{v_i}(\omega)\big) && \text{(12.915c)}\\
&= \phi_{n-(k+1)}\big(\prod_{i=1}^{k+1}\texttt{proj}_{v_i}(\omega)\big). && \text{(12.915d)}
\end{aligned}$$

Il ne reste qu'à écrire la formule démontrée avec $k = n$:

$$\psi_{n,n}\big(\phi_n(\omega)\big)(v_1, \ldots, v_n) = \phi_{n-n}\big(\prod_{i=1}^{n}\texttt{proj}_{v_i}(\omega)\big) = \omega(v_1, \ldots, v_n). \tag{12.916}$$

Donc nous avons bien que $\psi_{n,n}\big(\phi_n(\omega)\big) = \omega$. □

LEMooDGUGooTXOtih

Lemme 12.340.
Soient deux espaces vectoriels normés V et E. Nous considérons une application linéaire $A\colon V \to E$, une forme $\omega \in \mathcal{L}^n(V, \mathbb{R})$ ainsi que la forme

$$\begin{aligned}
\alpha\colon V &\to E_n\\
u &\mapsto A(u)\psi^{-1}(\omega).
\end{aligned} \tag{12.917}$$

Alors

$$\psi_{n+1,n+1}(\alpha)(v_1, \ldots, v_{n+1}) = A(v_1)\omega(v_2, \ldots, v_{n+1}). \tag{12.918}$$

Démonstration. Nous prouvons la formule suivante par récurrence :

$$\psi_{n+1,k}(\omega)(v_1,\dots,v_k) = A(v_1)\phi_{n-k+1}\big(\prod_{i=2}^{k}\mathtt{proj}_{v_i}(\omega)\big). \tag{12.919}$$

Nous commençons la récurrence avec $k=2$, en tenant compte que $\psi^{-1}=\phi_n$ (lemme 12.339) :

$$\begin{aligned}
\psi_{n+1,2}(\alpha)(v_1,v_2) &= \big(\psi_{n+1,1}(\alpha)v_1\big)v_2 && (12.920a)\\
&= \alpha(v_1)v_2 && (12.920b)\\
&= A(v_1)\big(\psi^{-1}(\omega)v_2\big) && (12.920c)\\
&= A(v_1)\big(\phi_n(\omega)v_2\big) && (12.920d)\\
&= A(v_1)\phi_{n-1}\big(\mathtt{proj}_{v_2}(\omega)\big). && (12.920e)
\end{aligned}$$

C'est bon pour $k=2$. Nous passons à la récurrence :

$$\begin{aligned}
\psi_{n+1,k+1}(\omega)(v_1,\dots,v_{k+1}) &= \big(\phi_{n+1,k}(\omega)(v_1,\dots,v_k)\big)v_{k+1} && (12.921a)\\
&= A(v_1)\Big(\phi_{n-k+1}\big(\prod_{i=2}^{k}\mathtt{proj}_{v_i}\omega\big)\Big)v_{k+1} && (12.921b)\\
&= A(v_1)\phi_{n-k}\big(\mathtt{proj}_{v_{k+1}}\prod_{i=2}^{k}\mathtt{proj}_{v_i}\omega\big) && (12.921c)\\
&= A(v_1)\phi_{n-(k+1)}\big(\prod_{i=2}^{k+1}\mathtt{proj}_{v_i}\omega\big). && (12.921d)
\end{aligned}$$

La récurrence est terminée. Nous écrivons la formule pour $k=n+1$ en tenant compte de $\phi_0=\mathrm{Id}$ pour terminer la preuve de ce lemme :

$$\psi_{n+1,n+1}(\omega)(v_1,\dots,v_{n+1}) = A(v_1)\phi_0\big(\prod_{i=2}^{n+1}\mathtt{proj}_{v_i}\omega\big) = A(v_1)\omega(v_2,\dots,v_{n+1}). \tag{12.922}$$

□

12.31.4 Fonctions différentiables plusieurs fois

Définition 12.341.
*Un C^k-**difféomorphisme** est une application inversible de classe C^k dont l'inverse est également de classe C^k.*

bilin_2diff

Exemple 12.342.
Soit $B:\mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}^n$ une application bilinéaire. On définit $f:\mathbb{R}^m\to\mathbb{R}^n$ par $f(x)=B(x,x)$. Le lemme 12.274 nous dit que B est différentiable. Cela implique la différentiabilité de f. Pour trouver la différentielle de la fonction f, nous écrivons $f=B\circ s$ où $s\colon\mathbb{R}^m\to\mathbb{R}^m\times\mathbb{R}^m$ est l'application $s(x)=(x,x)$. En utilisant la règle de différentiation de fonctions composées,

$$df(a) = dB\big(s(a)\big)\circ ds(a). \tag{12.923}$$

Mais $ds(a).u=(u,u)$ parce que $s(a+h)-s(a)-(h,h)=0$. Par conséquent,

$$df(a).u = dB\big(s(a)\big)(u,u) = B(u,a)+B(a,u) \tag{12.924}$$

EqdBsaExp

où nous avons utilisé la formule du lemme 12.274. La formule (12.924) peut être écrite sous la forme compacte

$$df(a) = B(\,\cdot\,,a)+B(a,\,\cdot\,) \tag{12.925}$$

La fonction $df(a)$ ainsi écrite est linéaire par rapport à a, donc différentiable. En outre elle coïncide avec sa différentielle, comme on a vu dans le lemme 10.269, au sens que la différentielle de df au

point a sera l'application que à chaque x dans $\mathbb{R}^m$ associe l'application linéaire $B(x,\ \cdot\ )+B(\ \cdot\ ,x)$. On voit bien que d^2f au point a est une application de $\mathbb{R}^m$ vers l'espace des applications linéaires $\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$. On peut utiliser d'autre part l'isomorphisme des espaces $\mathcal{L}(\mathbb{R}^m,\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n))$ et $\mathcal{L}(\mathbb{R}^m\times\mathbb{R}^m,\mathbb{R}^n)$ et dire que, une fois que a est fixé, l'application $d^2f(a)$ est une application bilinéaire sur $\mathbb{R}^m\times\mathbb{R}^m$. On écrit alors $d^2f(a)(x,y)=B(x,y)+B(y,x)$. △

12.31.5 Différentielle seconde, fonction de classe C^2

La différentielle seconde dans l'exemple 12.342 est symétrique, c'est-à-dire que $d^2f(a)(x_1,x_2)=d^2f(a)(x_2,x_1)$. En fait toute différentielle seconde est symétrique.

Le théorème de Schwarz permet de permuter les dérivées partielles d'ordre 2 dans le cas d'une fonction de classe C^2. Il y aura une généralisation pour les fonctions de classe C^m dans la proposition 12.352.

Schwarz

Théorème 12.343 (Schwarz)**.**
Soit U un ouvert de $\mathbb{R}^m$ et $f:U\subset\mathbb{R}^m\to\mathbb{R}^n$ une fonction de classe $\mathcal{C}^2$. Alors, pour tout couple i,j d'indices dans $\{1,\dots,m\}$ et pour tout point a dans U, on a

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(a)=\frac{\partial^2 f}{\partial x_j\partial x_i}(a).$$

Démonstration. Pour simplifier nous nous limitons ici au cas $m=2$. Soit (h,g) un vecteur fixé dans $\mathbb{R}^2$. Pour tout $v=(x,y)$ dans $\mathbb{R}^2$ on note

$$\begin{aligned}\Delta_h f(v)&=f(v+he_1)-f(v)=f(x+h,y)-f(x,y),\\ \Delta_g f(v)&=f(v+ge_2)-f(v)=f(x,y+g)-f(x,y),\end{aligned}\tag{12.926}$$

Nous avons

$$\begin{aligned}\Delta_g\Delta_h f(v)&=\big(f(x+h,y+g)-f(x,y+g)\big)-\big(f(x+h,y)-f(x,y)\big),\\ \Delta_h\Delta_g f(v)&=\big(f(x+h,y+g)-f(x+h,y)\big)-\big(f(x,y+g)-f(x,y)\big),\end{aligned}\tag{12.927}$$

donc,

$$\frac{1}{g}\Delta_g\left(\frac{1}{h}\Delta_h f(v)\right)=\frac{1}{h}\Delta_h\left(\frac{1}{g}\Delta_g f(v)\right).\tag{12.928}$$

On utilise alors le théorème des accroissements finis 12.192

$$\frac{1}{h}\Delta_h f(v)=\frac{1}{h}\big(f(x+h,y)-f(x,y)\big)=\frac{1}{h}\partial_1 f(x+t_1h,y)h=\partial_1 f(x+t_1h,y),\tag{12.929}$$

pour un certain t_1 dans $]0,1[$. De même on obtient

$$\frac{1}{g}\Delta_g f(v)=\partial_2 f(x,y+t_2g),$$

pour un certain t_2 dans $]0,1[$. Alors

$$\frac{1}{g}\Delta_g\big(\partial_1 f(x+t_1h,y)\big)=\frac{1}{h}\Delta_h\big(\partial_2 f(x,y+t_2g)\big).\tag{12.930}$$

En appliquant encore une fois le théorème des accroissements finis on a

$$\partial_2\partial_1 f(x+t_1h,y+s_1g)=\partial_1\partial_2 f(x+s_2h,y+t_2g).\tag{12.931}$$

Il suffit maintenant de passer à la limite pour $(h,g)\to(0,0)$ et de se souvenir du fait que f est $\mathcal{C}^2$ seulement si ses dérivées partielles secondes sont continues [136] pour avoir $\partial_2\partial_1 f(v)=\partial_1\partial_2 f(v)$. □

136. Théorème 12.330.

Si f est deux fois différentiable $d^2f(a)$ est l'application bilinéaire associée avec la matrice symétrique

$$H_f(a) = \begin{pmatrix} \partial_1^2 f(a) & \dots & \partial_1\partial_m f(a) \\ \vdots & \ddots & \vdots \\ \partial_1\partial_m f(a) & \dots & \partial_m^2 f(a), \end{pmatrix} \tag{12.932}$$

Cette matrice est dite la matrice **hessienne** de f.

Exemple 12.344.
Montrons qu'il n'existe pas de fonctions f de classe $\mathcal{C}^2$ telles que

$$\begin{cases} \partial_x f(x,y) = 5\sin x & \text{(12.933a)} \\ \partial_y(x,y) = 6x + y. & \text{(12.933b)} \end{cases}$$

Ceci est vite fait en appliquant le théorème de Schwarz, 12.343 ; ce que nous trouvons est

$$\partial_y(\partial_x f) = 0$$

alors que

$$\partial_x(\partial_y f) = 6. \tag{12.934}$$

Vu que $\partial_x(\partial_y f) \neq \partial_y(\partial_x f)$, le théorème 12.343 dit que f ne peut pas être de classe C^2. △

NORMooKSAVooCeILmI

12.345.
Soit une fonction de classe C^2 $f\colon V \to \mathbb{R}$ où V est un espace vectoriel de dimension $n < \infty$. Nous avons

$$f\colon V \to \mathbb{R} \tag{12.935a}$$
$$df\colon V \to \mathcal{L}(V,\mathbb{R}) \tag{12.935b}$$
$$d^2f\colon V \to \mathcal{L}\Big(V, \mathcal{L}(V,\mathbb{R})\Big), \tag{12.935c}$$

avec, en suivant les différentes formules du lemme 12.255,

$$df_a(u) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} \tag{12.936}$$

et

$$(d^2f)_a(u) = \frac{d}{dt}\Big[df_{a+tu}\Big]_{t=0} \tag{12.937}$$

pour tout $a, u \in V$. Notons que dans le deuxième cas, il s'agit d'une limite dans $\mathcal{L}(V,\mathbb{R})$. Si $\dim(V) = n$, alors $\dim\mathcal{L}(V,\mathbb{R}) = n$ et avec un choix de base, nous pouvons trouver une matrice $n \times n$ pour $(d^2f)_a$.

Soit une base $\{e_i\}$ de V et la base duale $\{e_i^*\}$ de $\mathcal{L}(V,\mathbb{R})$. Nous allons chercher la matrice de $(d^2f)_a$ pour ces bases. L'élément de matrice

$$\big[(d^2f)_a\big]_{ij} \tag{12.938}$$

est la composante e_j^* de $(d^2f)_a$ appliqué à e_i. Trouver cette composante e_j^* revient à appliquer l'élément $(d^2f)_a e_i$ de $\mathcal{L}(V,\mathbb{R})$ à e_j. Le calcul est donc :

$$\begin{aligned} \big[(d^2f)_a\big]_{ij} &= \big((d^2f)_a e_i\big)(e_j) && \text{(12.939a)} \\ &= \frac{d}{dt}\Big[df_{a+te_i}(e_j)\Big]_{t=0} && \text{(12.939b)} \\ &= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(a+te_i+se_j)\Big]_{s=0}\Big]_{t=0} && \text{(12.939c)} \\ &= \frac{\partial^2 f}{\partial x_i \partial x_j}(a). && \text{(12.939d)} \end{aligned}$$

SUBEQooDRZFooAuuad

Attention : le passage à (12.939b) n'est pas une trivialité. Le fait est que si $t \mapsto A(t)$ est une application continue $\mathbb{R} \to \mathcal{L}(V, \mathbb{R})$ alors

$$\lim_{t\to 0} \big(A(t)v\big) = \big(\lim_{t\to 0} A(t)\big)v. \tag{12.940}$$

Donc la matrice de d^2f est la matrice des dérivées secondes. Il s'agit d'une matrice symétrique par le théorème de Schwarz 12.343.

NORMooZAOEooGqjpLH

12.346.
Si $a \in V$, nous pouvons aussi voir $(d^2f)_a$ comme une forme bilinéaire sur V grâce à la proposition 10.88. Si $u, v \in V$ nous notons

$$(d^2f)_a(u, v) = (d^2f)_a(u)v. \tag{12.941}$$

À droite, il s'agit de la définition réelle de d^2f sans abus de notations, et à gauche, il s'agit d'une notation. Cette application bilinéaire $(d^2f)_a \in \mathcal{L}^{(2)}(V, \mathbb{R})$ a pour matrice symétrique la matrice des dérivées secondes calculées en a.

ExZHZYcNH

Exemple 12.347.
Voyons comment la différentielle seconde fonctionne entre deux espaces vectoriels. Soient deux espaces vectoriels de dimension finie V et W. Pour que les choses soient claires, nous avons :

$$f\colon V \to W \tag{12.942a}$$
$$df\colon V \to \mathcal{L}(V, W) \tag{12.942b}$$
$$d^2f\colon V \to \mathcal{L}\Big(V, \mathcal{L}(V, W)\Big). \tag{12.942c}$$

Si $a \in V$, alors $(d^2f)_a$ est une application $V \to \mathcal{L}(V, W)$. Il faut donc l'appliquer à $u \in V$ et ensuite à $v \in V$ pour obtenir un élément de W :

$$(d^2f)_a(u)v = \frac{d}{dt}\Big[df_{a+tu}\Big]_{t=0} v \tag{12.943a}$$
$$= \frac{d}{dt}\Big[df_{a+tu}(v)\Big]_{t=0} \tag{12.943b}$$
$$= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(a + tu + sv)\Big]_{s=0}\Big]_{t=0} \tag{12.943c}$$
$$= \frac{\partial^2 f}{\partial u \partial v}(a). \tag{12.943d}$$

Par conséquent nous voyons

$$d^2f\colon V \to \mathcal{L}^{(2)}(V, W)$$
$$d^2f_a(u, v) = \frac{\partial^2 f}{\partial u \partial v}(a). \tag{12.944}$$

EaQHINNtD

Dans le cas d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$, nous avons une seule direction et par linéarité de (12.944) par rapport à u et v, nous avons

$$d^2f_a(u, v) = f''(a)uv \tag{12.945}$$

EQooSOCGooLiNGmG

où les produits sont des produits usuels dans $\mathbb{R}$ et f'' est la dérivée seconde usuelle. △

Tout ceci est un peu résumé dans la proposition suivante.

PROPooFWZYooUQwzjW

Proposition 12.348.
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^2. Alors en désignant par H_af sa matrice hessienne au point a nous avons

$$(d^2f)_a(u, v) = \frac{\partial^2 f}{\partial u \partial v}(a) = \langle (H_af)u, v\rangle \tag{12.946}$$

pour tout $u, v \in \mathbb{R}^n$.

Démonstration. La première égalité est l'équation (12.945) déjà faite. Pour la seconde, il faut se rappeler du lien entre dérivée partielle et dérivée directionnelle, donné en le lemme 12.255. En particulier ici nous avons

$$\frac{\partial^2 f}{\partial u \partial v} = \sum_{kl} \frac{\partial^2 f}{\partial x_k \partial x_l}(a) u_k v_l = \langle (H_a f)u, v \rangle. \tag{12.947}$$

□

En particulier, la matrice hessienne $H_a f$ est symétrique et donc diagonalisable (théorème spectral 8.227). Si e_i est un vecteur propre unitaire pour la valeur propre λ_i nous avons

$$(d^2 f)_a(e_i, e_i) = \langle (H_a f)e_i, e_i \rangle = \lambda \langle e_i, e_i \rangle = \lambda. \tag{12.948}$$

Enfin pour celles qui aiment les notations matricielles de tout poil, il y a cette façon-ci d'écrire :

$$(d^2 f)_a(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \end{pmatrix} \begin{pmatrix} \partial_x^2 f(a) & \partial_{xy}^2 f(a) \\ \partial_{xy}^2 f(a) & \partial_y^2 f(a) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{12.949}$$

12.31.6 Ordre supérieur

Intuitivement, une fonction est de classe C^p si elle est p fois continûment différentiable. Cela est la définition 10.261.

Ce qui est terrible avec les différentielles d'ordre supérieurs, c'est l'emboîtement des espaces. Pour une fonction $f : E \to F$, nous allons souvent poser

$$V_0 = F \tag{12.950a}$$
$$V_{k+1} = \mathcal{L}(E, V_k), \tag{12.950b}$$

de telle sorte à avoir

$$df : E \to \mathcal{L}(E, F) = V_1 \tag{12.951}$$

et

$$d^2 f : E \to \mathcal{L}(E, V_1) = V_2, \tag{12.952}$$

ce qui donne en général

$$d^k f : E \to \mathcal{L}(E, v_{k-1}) = V_k. \tag{12.953}$$

La proposition suivante lie une bonne fois pour toute dérivée et différentielle dans le cadre de fonctions $\mathbb{R} \to \mathbb{R}$.

PROPooCNDHooKRwils

Proposition 12.349 ([1]).
Une fonction $f : \mathbb{R} \to \mathbb{R}$ est de classe C^p si et seulement si elle est p fois continûment dérivable.

Démonstration. Nous commençons par poser un certain nombre de notations. Comme souvent nous posons $V_0 = \mathbb{R}$ et

$$V_{k+1} = \mathcal{L}(\mathbb{R}, V_k). \tag{12.954}$$

De plus nous considérons $M_1 \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ donnée par $M_1(t) = t$. Et par récurrence

$$M_{k+1}(t) = tM_k. \tag{12.955}$$

Nous avons $M_1 \in V_1$ et $M_k : \mathbb{R} \to V_{k-1}$, c'est-à-dire $M_k \in V_k$.

Les formules que nous allons prouver sont que d'une part,

$$df_a = f'(a)M_1. \tag{12.956}$$

et que d'autre part, plus généralement,

$$(d^k f)_a = f^{(k)}(a)M_k. \tag{12.957}$$

En plusieurs parties et par récurrence.

(1) **Si f est continûment dérivable, alors f est C^1** Nous montrons que l'application T donnée par $T(h) = hf'(a)$ vérifie la condition pour être la différentielle de f en a :

$$\frac{f(a+h) - f(a) - f'(a)h}{\|h\|} = \frac{f(a+h) - f(a)}{\|h\|} - 1_h f'(a). \quad \text{EQooCPWKooWdgbED} \tag{12.958}$$

où nous avons noté 1_h le vecteur unité dans la direction de h, c'est-à-dire $1_h = h/\|h\|$. Vu que $h \in \mathbb{R}$, c'est simplement

$$1_h = \begin{cases} 1 & \text{si } h > 0 \\ -1 & \text{si } h < 0 \end{cases} \tag{12.959}$$

et nous ne définissons pas 1_h si $h = 0$.

C'est le moment de prendre la limite de (12.958) pour $h \to 0^+$ et $h \to 0^-$ séparément. Lorsque $h \to 0^+$, nous avons $\|h\| = h$ et $1_h = 1$. Vu que f est supposée dérivable, la limite du quotient existe et vaut $f'(a)$. Donc le tout a une limite nulle :

$$\lim_{h\to 0^+} \frac{f(a+h) - f(a) - f'(a)h}{\|h\|} = \lim_{h\to 0^+} \frac{f(a+h) - f(a)}{h} - f'(a) = 0. \tag{12.960}$$

En ce qui concerne la limite $h \to 0^-$, nous avons $\|h\| = -h$ et $1_h = -1$, et à nouveau une limite nulle. La proposition 12.39 nous permet alors de conclure que la limite existe et est nulle. Les limites à gauche et à droite étant nulles, la limite existe et est nulle par la proposition 12.39.

(2) **Si $f^{(p)}$ est continue alors $d^p f$ aussi** Nous passons à la récurrence de notre preuve. Sous l'hypothèse que $f^{(p)}$ existe et est continue, nous allons montrer que $d^p f$ existe, est continue et vaut

$$(d^p f)_a = f^{(p)}(a) M_p. \tag{12.961}$$

Soit $k < p$ tel que ce soit bon (pour $k = 1$ c'est déjà fait). Nous avons $(d^k f)_a = f^{(k)}(a)M_k$, et pour prouver que $(d^{k+1}f)_a = f^{(k+1)}(a)M_{k+1}$ nous l'y mettons dans la définition de la différentielle. Nous avons :

$$\frac{(d^k f)_{a+h} - (d^k f)_a - f^{(k+1)}(a)M_{k+1}(h)}{\|h\|} = \frac{f^{(k)}(a+h)M_k - f^{(k)}(a)M_k - hf^{(k+1)}(a)M_k}{\|h\|}. \tag{12.962}$$

La limite $h \to 0$ est une limite dans V_k, et elle se traite comme précédemment. Elle vaut zéro parce que $f^{(k+1)}$ est la dérivée de $f^{(k)}$. Cela justifie les faits que $d^k f$ est différentiable en a et que la différentielle est donné par la formule voulue.

Par hypothèse, $k + 1 \leqslant p$, donc $f^{(k+1)}$ est continue. Par conséquent l'application $a \mapsto f^{(k+1)}(a)M_{k+1}$ est continue.

(3) **Si f est de classe C^1 alors f' existe et est continue** Dire que f est de classe C^1 revient à dire que la différentielle $df \colon \mathbb{R} \to \mathcal{L}(\mathbb{R}, \mathbb{R})$ existe et est continue. Soyons conscient que $df_a(\epsilon) = \epsilon df_a(1)$ et calculons

$$\frac{f(a+\epsilon) - f(a) - df_a(\epsilon)}{\epsilon} = \frac{f(a+\epsilon) - f(a)}{\epsilon} - df_a(1). \tag{12.963}$$

La définition de la différentielle est que la limite de cela pour $\epsilon \to 0$ soit nulle. Cela implique que la limite suivante existe et vaut

$$\lim_{\epsilon\to 0} \frac{f(a+\epsilon) - f(a)}{\epsilon} = df_a(1). \tag{12.964}$$

Nous avons prouvé que $f'(a) = df_a(1)$.

La fonction $a \mapsto df_a$ est continue. Pouvons-nous en déduire que f' est continue ? Nous avons

$$f' = ev_1 \circ df \tag{12.965}$$

où ev_1 est l'application d'évaluation dont le lemme 10.67 a déjà donné la continuité. Donc f' est continue comme composée d'applications continues.

(4) *f est C^p.* **Récurrence** Nous supposons que f est de classe C^p, et nous allons montrer par récurrence que $f^{(k)}$ existe et est continue pour tout $k \leqslant p$. Posons exactement l'énoncé de notre récurrence.
Pour $k = 1$ c'est fait. Nous supposons que la formule soit correcte pour un certain $k < p$ et nous y allons pour $k + 1$. Nous avons SUBEQSooUPLAooQhueCl

$$\frac{(d^k f)(a+h) - (d^k f)(a) - f^{(k+1)}(a)M_{k+1}(h)}{\|h\|} = \frac{\big[f^{(k)}(a+h) - f^k(a) - hf^{(k+1)}(a)\big]M_k}{\|h\|} \tag{12.966a}$$

$$= \big[\frac{f^{(k)}(a+h) - f^{(k)}(a)}{\|h\|} - 1_h f^{(k+1)}(a)\big]M_k. \tag{12.966b}$$

où nous avons aussi tenu compte que $M_{k+1}(h) = hM_k$.
C'est le moment de calculer séparément les limites $h \to 0^+$ et $h \to 0^-$. Cela fonctionne comme toutes les autres fois.

□

Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ différentiable l fois. L'application

$$d^l f\colon \mathbb{R}^n \to \mathcal{L}\Big(\mathbb{R}^n, \mathcal{L}\big(\mathbb{R}^n, \mathcal{L}(\dots)\big)\Big) \tag{12.967}$$

au point x appliquée à $v^{(1)}$ appliquée au point $v^{(2)}$, …, appliquée à $v^{(l)}$ est notée

$$(d^l f)_x(v^{(1)}, \dots, v^{(l)}) \in \mathbb{R}. \tag{12.968}$$
EQooITQLooQUUIJ
PROPooQKZIooXTvkIr

Proposition 12.350.
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ différentiable l fois. Avec la notation (12.968) nous avons

$$(d^l f)_x(v^{(1)}, \dots v^{(l)}) = \sum_{k_1,\dots,k_l} v^{(1)}_{k_1} \dots v^{(l)}_{k_l} \frac{\partial^l f}{\partial x_{k_1} \dots \partial x_{k_l}}(x). \tag{12.969}$$

Démonstration. La preuve se fait par récurrence sur l, en sachant que la formule est déjà vraie pour $l = 1$ et $l = 2$. Si la formule est valable pour l, nous avons

$$(d^{l+1} f)_x(v^{(1)}, \dots, v^{(l+1)}) = \frac{d}{dt}\Big[(d^l f)_{x+tv^{(l+1)}}(v^{(1)}, \dots, v^{(l)})\Big]_{t=0} \tag{12.970a}$$

$$= \sum_{k_1 \dots k_l} v^{(1)}_{k_1} \dots v^{(l)}_{k_l} \frac{d}{dt}\Big[\frac{\partial^l f}{\partial x_{k_1} \dots \partial x_{k_l}}(x + tv^{l+1})\Big]_{t=0} \tag{12.970b}$$

$$= \sum_{k_1 \dots k_l} v^{(1)}_{k_1} \dots v^{(l)}_{k_l} \sum_i \frac{\partial}{\partial x_i} \frac{\partial^l f}{\partial x_{k_1} \dots \partial x_k}(x). \tag{12.970c}$$

Cela donne le résultat attendu. □

12.351.
La formule de la proposition 12.350 nous permet d'écrire de jolies formules comme

$$(d^3 f)_x(h, h, h) = \sum_{ijk} h_i h_j h_k (\partial^3_{ijk} f)(x). \tag{12.971}$$
EQooXRWWooMoKoQR
PROPooYGJDooOqibbh

Proposition 12.352 ([1]).
Soit un espace vectoriel V. Si $f\colon \mathbb{R}^n \to V$ est une application de classe C^m, alors pour tout $p \leqslant m$ nous pouvons permuter les dérivées partielles. Plus précisément, si $i_1, \dots, i_p$ sont dans $\{1, \dots, n\}$ et si $\sigma \in S_p$, nous avons

$$\partial_{i_1} \dots \partial_{i_p} f = \partial_{i_{\sigma(1)}} \dots \partial_{i_{\sigma(p)}} f. \tag{12.972}$$

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

À mon avis la proposition 12.352 se fait par récurrence avec 12.343.

PROPooMGFBooHWGXyC

Proposition 12.353.
Tout polynôme $P: \mathbb{R}^n \to \mathbb{R}^m$ est de classe C^∞.

12.32 Suites et séries : généralités

SECooTDZNooJvjPks

12.32.1 Norme suprémum

Définition 12.354 (norme suprémum[326])**.**
Soient un ensemble Ω, une partie A de Ω et un espace normé V. Lorsque g est une fonction $g: \Omega \to V$, nous notons

$$\|g\|_A = \sup_{x \in A} \|g(x)\| \tag{12.973}$$

*C'est la **norme suprémum** limitée à la partie A.*

*Nous disons qu'une suite de fonctions (f_n) définies sur un ensemble A **converge uniformément sur** A vers la fonction f si*

$$\lim_{n \to \infty} \|f_n - f\|_A = 0. \tag{12.974}$$

LEMooLPRZooUPsWTR

Lemme 12.355.
Soient un espace compact K et une fonction continue $f: K \to \mathbb{C}$ qui se décompose en partie réelle et complexe comme $f(x) = u(x) + iv(x)$. Alors

(1) Les fonctions u et v sont continues sur K.
(2) $\|u\|_\infty \leqslant \|f\|_\infty$,
(3) $\|v\|_\infty \leqslant \|f\|_\infty$,

12.32.2 Convergence uniforme

12.32.2.1 Critère de Cauchy uniforme

PropNTEynwq

Proposition 12.356 (Critère de Cauchy uniforme[411])**.**
Soit X un espace topologique et (Y, d) un espace topologique complet. La suite de fonctions $f_n: X \to Y$ converge uniformément sur A si et seulement si pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que si $k, l > N$ alors

$$d\big(f_k(x), f_l(x)\big) \leqslant \epsilon \tag{12.975}$$

pour tout $x \in A$.

Grosso modo, cela dit que si qu'une suite de Cauchy pour la norme uniforme est une suite uniformément convergente. Le fait que la suite converge fait partie du résultat et n'est pas une hypothèse. Ce critère sera utilisé pour montrer que $\big(C(K), \|.\|_\infty\big)$ est complet, proposition 12.359.

Démonstration. Si $f_n \xrightarrow{unif} f$ alors le critère est satisfait ; c'est dans l'autre sens que la preuve est intéressante.

Soit donc une suite de fonctions satisfaisant au critère et montrons qu'elle converge uniformément. Pour tout $x \in X$ la suite $n \mapsto f_n(x)$ est de Cauchy dans l'espace complet Y ; nous avons donc convergence ponctuelle $f_n \to f$. Nous devons prouver que cette convergence est uniforme. Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que si $k, l > N$ alors

$$d\big(f_k(x), f_l(x)\big) \leqslant \epsilon \tag{12.976}$$

pour tout $x \in X$. Si nous nous fixons un tel k et un $x \in A$ nous considérons l'inégalité

$$d\big(f_k(x), f_l(x)\big) \leqslant \epsilon \tag{12.977}$$

qui est vraie pour tout l. En passant à la limite $l \to \infty$ (limite qui commute avec la fonction distance par définition de la topologie) nous avons

$$d\big(f_k(x), f(x)\big) \leqslant \epsilon. \tag{12.978}$$

Cette inégalité étant valable pour tout $x \in X$, cela signifie que $f_n \xrightarrow{unif} f$. □

ThoUnigCvCont

Théorème 12.357 (Limite uniforme de fonctions continues)**.**
Soit A, un ensemble mesuré et $f_n \colon A \to \mathbb{R}^n$, une suite de fonctions continues convergeant uniformément vers f. Si les fonctions f_n sont toutes continues en $x_0 \in A$, alors f est continue en x_0.

Démonstration. Soit $\epsilon > 0$. Si $x \in A$ nous avons, pour tout n, la majoration

$$\begin{aligned} \|f(x) - f(x_0)\| &\leqslant \|f(x) - f_n(x)\| + \|f_n(x) - f_n(x_0)\| + \|f_n(x_0) - f(x_0)\| && \text{(12.979a)}\\ &\leqslant \|f_n(x) - f_n(x_0)\| + 2\|f_n - f\|_\infty. && \text{(12.979b)} \end{aligned}$$

Grâce à l'uniforme convergence, nous considérons $N \in \mathbb{N}$ tel que $\|f_n - f\| \leqslant \epsilon$ pour tout $n \geqslant N$. Pour de tels n, nous avons

$$\|f(x) - f(x_0)\| \leqslant 2\epsilon + \|f_n(x) - f_n(x_0)\|. \tag{12.980}$$

La continuité de f_n nous fournit un $\delta > 0$ tel que $\|f_n(x_0) - f_n(x)\| < \epsilon$ dès que $\|x - x_0\| < \delta$. Pour ce δ, nous avons alors $\|f_n(x) - f_n(x_0)\| < \epsilon$.

Donc lorsque $n \geqslant N$ (N est choisi en fonction de ϵ), et $\|x - x_0\| < \delta$ (δ est choisi en fonction de N), nous avons

$$\|f(x) - f(x_0)\| \leqslant 3\epsilon, \tag{12.981}$$

ce qui prouve la continuité de f en x_0. □

12.32.2.2 Complétude avec la norme uniforme

PropCZslHBx

Proposition 12.358 (Limite uniforme de fonctions continues)**.**
Soit X un espace topologique et (Y, d) un espace métrique. Si une suite de fonctions $f_n \colon X \to Y$ continues converge uniformément, alors la limite est séquentiellement continue[137]*.*

Démonstration. Soit $a \in X$ et prouvons que f est séquentiellement continue en a. Pour cela nous considérons une suite $x_n \to a$ dans X. Nous savons que $f(x_n) \xrightarrow{Y} f(x)$. Pour tout $k \in \mathbb{N}$, tout $n \in \mathbb{N}$ et tout $x \in X$ nous avons la majoration

$$\begin{aligned} \big\|f(x_n) - f(x)\big\| &\leqslant \big\|f(x_n) - f_k(x_n)\big\| + \big\|f_k(x_n) - f_k(x)\big\| + \big\|f_k(x) - f(x)\big\| && \text{(12.982a)}\\ &\leqslant 2\|f - f_k\|_\infty + \big\|f_k(x_n) - f_k(x)\big\|. && \text{(12.982b)} \end{aligned}$$

Soit $\epsilon > 0$. Si nous choisissons k suffisamment grand la premier terme est plus petit que ϵ. Et par continuité de f_k, en prenant n assez grand, le dernier terme est également plus petit que ϵ. □

PropSYMEZGU

Proposition 12.359.
Soit X un espace topologique métrique (Y, d) un espace métrique complet. Alors les espaces

(1) $\big(C_b^0(X,Y), \|.\|_\infty\big)$ des fonctions continues et bornées $X \to Y$,
(2) $\big(C_0^0(X,Y), \|.\|_\infty\big)$ des fonctions continues et s'annulant à l'infini
(3) $\big(C_0^k(X,Y), \|.\|_\infty\big)$ des fonctions de classe C^k et s'annulant à l'infini

sont complets.

137. Si X est métrique, alors c'est la continuité usuelle par la proposition 7.137.

Démonstration. Soit (f_n) une suite de Cauchy dans $C(X,Y)$, c'est-à-dire que pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que si $k, l > N$ nous avons $\|f_k - f_l\|_\infty \leqslant \epsilon$. Cette suite vérifie le critère de Cauchy uniforme 12.356 et donc converge uniformément vers une fonction $f\colon X \to Y$. La continuité (ou l'aspect C^k) de la fonction f découle de la convergence uniforme et de la proposition 12.358 (c'est pour avoir l'équivalence entre la continuité séquentielle et la continuité normale que nous avons pris l'hypothèse d'espace métrique).

Si les fonctions f_k sont bornées ou s'annulent à l'infini, la convergence uniforme implique que la limite le sera également. □

Notons que si X est compact, les fonctions continues sont bornées par le théorème 7.220 et nous pouvons simplement dire que $C^0(X,Y)$ est complet, sans préciser que nous parlons des fonctions bornées.

LemdLKKnd

Lemme 12.360.
Soient un espace topologique compact A et un espace complet B. L'ensemble des fonctions continues de A vers B muni de la norme uniforme est complet.

Dit de façon courte : $\big(C(A,B), \|.\|_\infty\big)$ est complet.

Démonstration. Soit (f_k) une suite de Cauchy de fonctions dans $C(A,B)$. Pour chaque $x \in A$ nous avons

$$\|f_k(x) - f_l(x)\|_B \leqslant \|f_k - f_l\|_\infty, \tag{12.983}$$

de telle sorte que la suite $(f_k(x))$ est de Cauchy dans B et converge donc vers un élément de B. La suite de Cauchy (f_k) converge donc ponctuellement vers une fonction $f\colon A \to B$. Nous devons encore voir que cette fonction est continue ; ce sera l'uniformité de la norme qui donnera la continuité. En effet soit $x_n \to x$ une suite dans A qui converge vers $x \in A$. Pour chaque $k \in \mathbb{N}$ nous avons

$$\|f(x_n) - f(x)\| \leqslant \|f(x_n) - f_k(x_n)\| + \|f_k(x_n) - f_k(x)\| + \|f_k(x) - f(x)\|. \tag{12.984}$$

En prenant k et n assez grands, cette expression peut être rendue aussi petite que l'on veut ; le premier et le troisième terme par convergence ponctuelle $f_k \to f$, le second terme par continuité de f_k. La suite $f(x_n)$ est donc convergente vers $f(x)$ et la fonction f est continue. □

¡¡ Avertissement/question à la lectrice !! 12.361
Il serait sans doute bon de revoir cette preuve à la lumière du critère de Cauchy uniforme 12.356.

12.362 ([412]).
Le théorème de Stone-Weierstrass indique que les polynômes sont denses pour la topologie uniforme dans les fonctions continues. Donc il existe des limites uniformes de fonctions C^∞ qui ne sont même pas dérivables. Les espaces de type C^p munis de $\|.\|_\infty$ ne sont donc pas complets sans quelques hypothèses. Voir la proposition 12.359 et le thème 37.

ThoUFPLEZh

Théorème 12.363 (Théorème de Dini[413]).
Soient un espace métrique complet D et une suite de fonctions $f_n \in C(D,\mathbb{R})$ telle que

(1) $f_n \to g$ ponctuellement,

(2) $g \in C(D,\mathbb{R})$,

(3) la suite (f_n) est croissante, c'est-à-dire que pour tout $x \in D$ et pour tout $n \geqslant 0$ nous avons $f_{n+1}(x) \geqslant f_n(x)$.

Alors la convergence est uniforme.

Démonstration. Soit $x \in D$ et $\epsilon > 0$. Il existe $N(x) \in \mathbb{N}$ tel que

$$g(x) - \epsilon \leqslant f_{N(x)} \leqslant g(x). \tag{12.985}$$

De plus g et $f_{N(x)}$ sont des fonctions continues, donc il existe $\eta(x)$ tel que si $y \in B\big(x, \eta(x)\big)$ alors

$$g(y) \in B\big(g(x), \epsilon\big) \tag{12.986a}$$
$$f_{N(x)}(y) \in B\big(f_{N(x)}(x), \epsilon\big). \tag{12.986b}$$

Si $n \geqslant N(x)$ et si $y \in B(x, \eta(x))$ alors nous avons les majorations

$$g(y) \geqslant f_n(y) \geqslant f_{N(x)}(y) \geqslant f_{N(x)}(x) - \epsilon \geqslant g(x) - 2\epsilon \geqslant g(y) - 3\epsilon. \tag{12.987}$$

Justifications :

(1) Les deux premières inégalités sont la croissance de la suite.
(2) La suivante est (12.986b).
(3) Ensuite il y a le choix de $N(x)$.
(4) Et enfin il y a (12.986a).

Nous retenons que si $x \in D$ et si $n \geqslant N(x)$ alors

$$g(y) \geqslant f_n(y) \geqslant g(y) - 3\epsilon \tag{12.988}$$

pour tout $y \in B(x, \eta(x))$.

Nous utilisons maintenant la compacité de D. Pour chaque $x \in D$ nous pouvons considérer la boule ouverte $B\big(x, \eta(x)\big)$; ces boules recouvrent D. Nous en extrayons un sous-recouvrement fini, c'est-à-dire un ensemble fini d'éléments $x_1, \ldots, x_K$ tels que

$$D = \bigcup_{k=1}^{K} B\big(x_k, \eta(x_k)\big). \tag{12.989}$$

Si à ce moment vous ne comprenez pas pourquoi c'est une égalité au lieu d'une inclusion, il faut lire l'exemple 7.33. Considérons

$$n \geqslant N = \max\{N(x_1), \ldots, N(x_K)\}. \tag{12.990}$$

Pour tout $y \in D$ il existe $k \in \{1, \ldots, K\}$ tel que $y \in B\big(x_k, \eta(x_k)\big)$, et vu que $n \geqslant N(x_k)$ nous reprenons la majoration (12.988) :

$$g(y) \geqslant f_n(y) \geqslant g(y) - 3\epsilon. \tag{12.991}$$

Pour le n choisi nous avons ces inégalités pour tout $y \in D$, c'est-à-dire que nous avons $\|f_n - g\| \leqslant 3\epsilon$ et donc la convergence uniforme. □

Proposition 12.364 ([1]).
Soient une suite de fonctions continues $u_i : \mathbb{R} \to \mathbb{R}$ et une fonction continue u telle que $u_i \to u$ simplement. Alors la convergence est uniforme sur tout compact.

Démonstration. Soit un compact K ; nous notons $\|.\|$ la norme uniforme sur K. Supposons que la limite ne soit pas uniforme, c'est-à-dire qu'il existe un $\epsilon > 0$ tel que

$$\|u_i - u\| > 2\epsilon \tag{12.992}$$

pour tout i. Cela permet de considérer pour tout i un élément $x_i \in K$ tel que [138]

$$\|u_i(x_i) - u(x_i)\| > \epsilon. \tag{12.993}$$

Pour cela, il faut noter que K est compact et que la fonction $x \mapsto \|u_i(x) - u(x)\|$ est continue sur K. Elle est donc bornée et atteint son maximum (c'est le théorème de Weierstrass 7.146).

138. Notez l'inégalité stricte, obetenue en considérant 2ϵ plus haut.

La suite $i \mapsto x_i$ est une suite dans un compact, et quitte à prendre une sous-suite, nous supposons qu'elle converge vers $a \in K$ (ça, c'est Bolzano-Weierstrass 7.143).

La convergence ponctuelle $u_i \to u$, prise en a, dit qu'il existe un N tel que $|u_i(a) - u(a)| < \epsilon$ pour tout $i \geqslant N$. Pour un tel i, nous avons aussi

$$|u_i(x) - u(x)| < \epsilon \tag{12.994}$$

sur un voisinage de a, parce que $u_i - u$ est continue. Mais tout voisinage de a contient un élément x_j pour lequel

$$|u_i(x_j) - u(x_j)| > \epsilon. \tag{12.995}$$

Contradiction. □

12.32.3 Série de fonctions

Les séries de fonctions sont des cas particuliers de suites.

DEFooYEIUooCAgrxI

Définition 12.365.
Si (f_n) est une suite de fonctions, nous définissons la somme des f_n de la façon suivante :

$$\sum_{n=1}^{\infty} f_n = \lim_{N\to\infty} \sum_{n=1}^{N} f_n. \tag{12.996}$$

Le membre de droite est une définition de la notation introduite dans le membre de gauche.

Avant de vous lancer, relisez une bonne fois les définitions de convergence absolue (définition 10.92) et de convergence uniforme (équation 10.231).

Lemme 12.366.
Soient des fonctions $u_n : \Omega \to \mathbb{C}$. Si il existe une suite réelle positive $(a_n)_{n\in\mathbb{N}}$ telle que

(1) pour tout $z \in \Omega$ et pour tout $n \in \mathbb{N}$ nous avons $|u_n(z)| \leqslant a_n$ (c'est-à-dire $a_n \geqslant \|u_n\|_\infty$),

(2) la somme $\sum_n a_n$ converge,

alors la série de fonctions $\sum_{n=0}^{\infty} u_n$ converge normalement[139].

Démonstration. Découle du lemme de comparaison 10.128. □

ThoSerCritAbel

Théorème 12.367.
Soit $\sum_{k=1}^{\infty} g_k(x)$, une série de fonctions complexes où $g_k(x) = \varphi_k(x)\psi_k(x)$. Supposons que

(1) $\varphi_k : A \to \mathbb{C}$ et $|\sum_{k=1}^{K} \varphi_k(x)| \leqslant M$ où M est indépendant de x et K,

(2) $\psi_k : A \to \mathbb{R}$ avec $\psi_k(x) \geqslant 0$ et pour tout x dans A, $\psi_{k+1}(x) \leqslant \psi_k(x)$, et enfin supposons que $\psi_k(x)$ converge uniformément vers 0.

Alors $\sum_{k=1}^{\infty} g_k$ est uniformément convergente.

ThoAbelSeriePuiss

Théorème 12.368.
Si la série de puissances (réelle) converge en $x = x_0 + R$, alors elle converge uniformément sur $[x_0 - R - \epsilon, x_0 + R]$ $(\epsilon > 0)$ vers une fonction continue.

PropOMBbwst

Proposition 12.369.
Soient E et F, deux espaces vectoriels normés, Ω une partie ouverte de E et une suite de fonctions continues $u_n : \Omega \to F$ convergeant normalement sur Ω, c'est-à-dire que $\sum_n \|u_n\|_\infty$ converge, la norme $\|.\|_\infty$ devant être comprise comme la norme supremum sur Ω. Alors la fonction $u = \sum_n u_n$ est continue sur Ω.

139. Définition 10.93.

Démonstration. Soit $x, x' \in \Omega$ en supposant que $\|x - x'\|$ est petit. Soit encore $\epsilon > 0$. Nous allons montrer la continuité en x. Pour cela nous savons que pour tout N l'inégalité suivante est correcte :

$$\|u(x) - u(x')\| \leqslant \left\| \sum_{n=0}^{N} u_n(x) - \sum_{n=0}^{N} u_n(x') \right\| + \sum_{n=N+1}^{\infty} \|u_n(x)\| + \sum_{n=N+1}^{\infty} \|u_n(x')\|. \tag{12.997}$$

Les deux derniers termes sont majorés par $\sum_{n=N+1}^{\infty} \|u_n\|_\infty$ qui, par hypothèse, peut être rendu aussi petit que souhaité en choisissant N assez grand. Nous choisissons donc un N tel que ces deux termes soient plus petits que ϵ. Ce N étant fixé, la fonction $\sum_{n=0}^{N} u_n$ est continue et nous pouvons choisir x' assez proche de x pour que le premier terme soit majoré par ϵ. □

ThoSerUnifCont

Théorème 12.370 (Série uniforme de fonctions continues[1])**.**
Soit un espace topologique X ainsi qu'un espace vectoriel normé V. Soient des fonctions continues $f_n : X \to V$. Si la série $\sum_{n=0}^{\infty} f_n$ converge uniformément[140]*, alors la fonction somme est séquentiellement continue.*

Si X est métrisable, alors la somme est continue.

Démonstration. Nous notons f la somme (qui existe par hypothèse) et par (s_N) la suite des sommes partielles. En tant que somme finie de fonctions continues, chacune des fonctions s_N est continue. Ce que dit la définition 10.94, c'est que la convergence des sommes partielles est uniforme :

$$s_N \xrightarrow{unif} f. \tag{12.998}$$

La proposition 12.358 dit alors que f est séquentiellement continue.

Nous en déduisons la continuité de f dans le cas d'un espace métrisable avec la proposition 7.255. □

Le corolaire suivant permet de considérer des séries de fonctions indexées par exemple par $\mathbb{Z}$ plutôt que par $\mathbb{N}$.

Corolaire 12.371.
Une famille dénombrable de fonctions continues convergeant normalement converge vers une fonction continue.

Démonstration. Soit I dénombrable. Considérons une famille de fonctions continues $(f_n)_{n\in I}$ telles que la famille $(\|f_i\|_\infty)_{i\in I}$ soit sommable. Le proposition 10.126 nous permet d'utiliser une bijection entre I et $\mathbb{N}$. La proposition 12.369 s'applique alors. □

ThoCritWeierstrass

Théorème 12.372 (Critère de Weierstrass)**.**
Soit une suite de fonctions $f_k : A \to \mathbb{C}$ telles que $|f_k(x)| \leqslant M_k \in \mathbb{R}$, $\forall x \in A$. Si $\sum_{k=1}^{\infty} M_k$ converge, alors $\sum_{k=1}^{\infty} f_k$ converge absolument et uniformément.

Démonstration. La convergence normale est facile : l'hypothèse dit que $\|f_k\|_\infty \leqslant M_k$, et donc que

$$\sum_{k=1}^{\infty} \|f_k\|_\infty \leqslant \sum_{k} M_k < \infty. \tag{12.999}$$

La convergence uniforme est à peine plus subtile. Nous nommons F la fonction somme. Pour tout x et pour tout N, nous avons

$$\left\| \sum_{n=1}^{N} f_n(x) - F(x) \right\| = \| \sum_{n=N}^{\infty} f_n(x) \| \tag{12.1000a}$$

$$\leqslant \sum_{n=N}^{\infty} \|f_k(x)\| \tag{12.1000b}$$

$$\leqslant \sum_{n=N}^{\infty} \|f_n\|_\infty. \tag{12.1000c}$$

140. Définition 10.94.

La convergence normale étant assurée, la série $\sum_{n_1}^{\infty} \|f_n\|_\infty$ est finie, ce qui implique que la queue de somme $\sum_{n=N}^{\infty} \|f_n\|_\infty$ tend vers zéro lorsque $N \to \infty$. Pour tout ϵ, il existe donc un N (non dépendant de x) tel que

$$\| \sum_{n=1}^{N} f_n(x) - F(x) \| \leqslant \epsilon. \tag{12.1001}$$

En prenant le supremum sur $x \in A$ nous trouvons la convergence uniforme. □

Remarque 12.373.
Il n'y a pas de critère correspondant pour les suites. Il n'est pas vrai que si $\lim_{n\to\infty} \|f_n\|$ existe, alors $\lim_{n\to\infty} f_n$ existe, comme le montre l'exemple

$$f_n(x) = \begin{cases} 1 & \text{si } x \in [0,1] \text{ et } n \text{ est pair} \\ 1 & \text{si } x \in [1,2] \text{ et } n \text{ est impair} \\ 0 & \text{sinon.} \end{cases} \tag{12.1002}$$

12.33 Permuter limite et dérivée

Une version avec intégrales de la démonstration qui suit est dans 14.329. Le même pour les dérivées partielles sera le théorème 12.377.

THOooXZQCooSRteSr

Théorème 12.374 ([326, 357, 1], thème 62)**.**
Soient une suite de fonctions $f_i : \mathbb{R} \to \mathbb{R}$, une fonction $f : \mathbb{R} \to \mathbb{R}$ et une fonction $g : \mathbb{R} \to \mathbb{R}$ telles que

(1) f_i est de classe C^1 pour tout i,
(2) $f_i \to f$ simplement,
(3) $f_i' \to g$ uniformément sur tout compact.

Alors

ITEMooYSWDooFFeQCd

(1) $f_i \to f$ uniformément sur tout compact.

ITEMooFAWUooVQJPZh

(2) $f' = g$,
(3) f est de classe C^1,

Démonstration. Un point à la fois.

(1) **Pour (1)** Soit un compact $K \subset \mathbb{R}$. Dans cette partie, toutes les fonctions en jeu sont restreintes à K. En particulier, lorsque nous parlerons du module de continuité [141] ω_g pour g ou ω_i pour f_i', nous parlerons en réalité des fonctions $g|_K$ et $f_i'|_K$.
Ceci dit, nous allons montrer que (f_i) est une suite de Cauchy pour la norme uniforme [142].
Soit $\epsilon > 0$. On note ω_i le module de continuité de f_i'. Soient $y \in K$, $n \in \mathbb{N}$ et posons $\alpha_n = \frac{y-x}{n+1}$. Pour tout $i \geqslant 0$, nous avons la somme télescopique

$$f_i(y) = f_i(x) + \sum_{k=0}^{n} \Big[f_i(x + (k+1)\alpha_n) - f_i(x + k\alpha_n) \Big]. \tag{12.1003}$$

Par le théorème des accroissements finis 12.192, il existe pour tout $0 \leqslant k \leqslant n$ un réel $u_{n,i,k} \in [k\alpha_n, (k+1)\alpha_n]$ tel que

$$f_i(x + (k+1)\alpha_n) - f_i(x + k\alpha_n) = |\alpha_n| f_i'(x + u_{n,i,k}), \tag{12.1004}$$

de sorte que

$$f_i(y) = f_i(x) + |\alpha_n| \sum_{k=0}^{n} f_i'(x + u_{n,i,k}). \tag{12.1005}$$

141. Définition 10.310.
142. Et si vous avez bien suivi l'avertissement, c'est bien de la norme uniforme sur K que nous parlons.

Et pour tout $i,j\geqslant 0$, on obtient SUBEQSooUYFTooWPVfWt

$$|f_i(y)-f_j(y)|\leqslant |f_i(x)-f_j(x)|+|\alpha_n|\sum_{k=0}^{n}\left|f_i'(x+u_{n,i,k})-f_j'(x+u_{n,j,k})\right| \tag{12.1006a}$$

$$\leqslant |f_i(x)-f_j(x)| \tag{12.1006b}$$
$$+|\alpha_n|\sum_{k=0}^{n}\left|f_i'(x+u_{n,i,k})-f_i'(x+u_{n,j,k})\right|$$
$$+|\alpha_n|\sum_{k=0}^{n}\left|f_i'(x+u_{n,j,k})-f_j'(x+u_{n,j,k})\right|$$

$$\leqslant |f_i(x)-f_j(x)| \tag{12.1006c}$$
$$+|\alpha_n|\sum_{k=0}^{n}\left|f_i'(x+u_{n,i,k})-f_i'(x+u_{n,j,k})\right|$$
$$+|\alpha_n|\sum_{k=0}^{n}\|f_i'-f_j'\|$$

$$\leqslant |f_i(x)-f_j(x)|+|\alpha_n|\sum_{k=0}^{n}\omega_i(|\alpha_n|)+|\alpha_n|\sum_{k=0}^{n}\|f_i'-f_j'\| \tag{12.1006d}$$

$$\leqslant |f_i(x)-f_j(x)|+|x-y|\left(\omega_i(|\alpha_n|)+\|f_i'-f_j'\|_K\right) \tag{12.1006e}$$

$$\leqslant |f_i(x)-f_j(x)|+M\Big(\omega_i\big(\frac{M}{n+1}\big)+\|f_i'-f_j'\|_K\Big) \tag{12.1006f}$$

Justifications :

— Pour (12.1006c) La norme supremum est forcément plus grande que la valeur en un point.

— Pour (12.1006d) Remarquer que $\|x+u_{n,i,k}-(x+u_{n,j,k})\|=\|u_{n,i,k}-u_{n,j,k}\|\leqslant|\alpha_n|$. Cela est donc une bonne occasion de prendre la définition (10.925) du module de continuité :

$$\left|f_i'(x+u_{n,i,k})-f_i'(x+u_{n,j,k})\right|\leqslant\omega_i(\alpha_n). \tag{12.1007}$$

— Pour (12.1006e), il y a $n+1$ termes dans les sommes et $\alpha_n=(y-x)/(n+1)$.

— Pour (12.1006f), nous travaillons uniquement sur le compact K. En particulier $x,y\in K$ et il existe un nombre M ne dépendant que de K tel que $|y-x|<M$. De plus ω_i est décroissante. Donc en remplaçant $|y-x|$ par M nous majorons.

Recopions notre dernière inéquation :

$$|f_i(y)-f_j(y)|\leqslant|f_i(x)-f_j(x)|+M\Big(\omega_i\big(\frac{M}{n+1}\big)+\|f_i'-f_j'\|_K\Big) \tag{12.1008}$$

Vu que f_i est de classe C^1, la fonction f_i' est continue. Et vu que nous travaillons sur le compact K, elle est même uniformément continue (proposition 12.78). Le lemme 10.312 dit qu'une fonction uniformément continue a un module de continuité continu en zéro : $\lim_{\delta\to0}\omega_i(\delta)=0$. Nous pouvons donc prendre la limite $n\to0$ pour nous supprimer le module de continuité :

$$|f_i(y)-f_j(y)|\leqslant|f_i(x)-f_j(x)|+M\|f_i'-f_j'\|_K. \tag{12.1009}$$

Nous prenons maintenant le supremum par rapport à y :

$$\|f_i-f_j\|_K\leqslant|f_i(x)-f_j(x)|+M\|f_i'+f_j'\|_K. \tag{12.1010}$$

Par hypothèse nous avons la convergence simple $f_i\to f$, c'est-à-dire la convergence $f_i(x)\xrightarrow{\mathbb{R}}f_i(x)$ pour tout x. Pour le x que nous nous sommes fixés, la suite $i\mapsto f_i(x)$ est donc une suite de Cauchy dans $\mathbb{R}$.

Soit $\epsilon > 0$. Il existe $N \in \mathbb{N}$ tels que si $i, j > N$ nous avons $|f_i(x) - f_j(x)| < \epsilon$. De même la convergence uniforme $f_i' \xrightarrow{\|.\|_K} g$ implique que f_k' est également de Cauchy pour la norme uniforme. Donc pour un $i, j > N$ (éventuellement un autre N, mais on prend le maximum entre les deux), nous avons $\|f_i' - f_j'\| < \epsilon$.
Donc si $i, j > N$ nous avons

$$\|f_i - f_j\|_K \leqslant \epsilon + M\epsilon = (M+1)\epsilon. \tag{12.1011}$$

Nous avons donc prouvé que (f_i) est une suite de Cauchy pour la norme $\|.\|_K$. Cela implique que (f_i) a une limite uniforme sur K. Vu que nous avons déjà $f_i \to f$, nous en déduisons que sur K, cette limite est uniforme :

$$f_i \xrightarrow{\|.\|_K} f. \tag{12.1012}$$

Voilà qui prouve la convergence uniforme sur tout compact.

(2) **Pour (2)** Nous ne savons encore rien de la fonction limite f. Nous montrons qu'elle est dérivable et que $f' = g$.
Soient $y \in \mathbb{R}$ et un voisinage compact $K = \overline{B(y, \delta)}$ de y avec $\delta > 0$. Pour tout $i > 0$ nous avons :

$$\Big|\frac{f(y+\delta) - f(y)}{\delta} - g(y)\Big| \tag{12.1013a}$$
$$\leqslant \frac{|f(y) - f_i(y)|}{\delta} + \frac{|f(y+\delta) - f_i(y+\delta)|}{\delta} + \Big|\frac{f_i(y+\delta) - f_i(y)}{\delta} - g(y)\Big| \tag{12.1013b}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + \Big|\frac{f_i(y+\delta) - f_i(y)}{\delta} - g(y)\Big| \tag{12.1013c}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + |f_i'(u) - g(y)| \tag{12.1013d}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + |f_i'(u) - f_i'(y)| + |f_i'(y) - g(y)| \tag{12.1013e}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + |f_i'(u) - f_i'(y)| + \|f_i' - g\|_K \tag{12.1013f}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + \omega_i(|u-y|) + \|f_i' - g\|_K \tag{12.1013g}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + \|f_i' - g\|_K + \omega_i(2\delta) \tag{12.1013h}$$
$$\leqslant \frac{2}{\delta}\|f_i - f\|_K + \|f_i' - g\|_K + \omega_g(2\delta) \tag{12.1013i}$$

Justifications :

— Pour 12.1013d. Nous avons utilisé le théorème des accroissements finis 12.192, qui assure l'existence de $u \in B(y, \delta)$ tel que

$$\frac{f_i(y+\delta) - f_i(y)}{\delta} = f_i'(u). \tag{12.1014}$$

— Pour 12.1013f. Nous faisons un supremum sur le $y \in K$ dans le dernier terme.

— Pour 12.1013h. Nous majorons $|u - y|$ par le diamètre 2δ du compact $K = \overline{B(y, \delta)}$.

— Pour 12.1013i. Le lemme 10.313 et le fait que $f_i' \xrightarrow{\|.\|_K} g$.

Recopions la dernière inégalité :

$$\Big|\frac{f(y+\delta) - f(y)}{\delta} - g(y)\Big| \leqslant \frac{2}{\delta}\|f_i - f\|_K + \|f_i' - g\|_K + \omega_g(2\delta). \tag{12.1015}$$

Nous prenons la limite $i \to \infty$. Par le point (1) nous savons que $\|f_i - f\|_K \to 0$. Par hypothèse nous savons aussi que $\|f_i' - g\|_K \to 0$. Nous restons donc avec

$$\Big|\frac{f(y+\delta) - f(y)}{\delta} - g(y)\Big| \leqslant \omega_g(2\delta). \tag{12.1016}$$

Or, par uniforme continuité de g, nous avons $\lim_{\delta\to 0}\omega_g(\delta)=0$, donc la limite dans le membre de gauche se passe bien et

$$\lim_{\delta\to\delta}\frac{f(y+\delta)-f(y)}{\delta}=g(y),\tag{12.1017}$$

ce qui signifie que f est dérivable en y et que la dérivée est $g(y)$.

□

¡¡ Avertissement/question au lecteur !! 12.375
Aussi incroyable que cela puisse paraitre, je n'ai pas trouvé d'énoncés du théorème 12.377. Donc soyez prudent. C'est donc une adaptation personnelle du cas sur $\mathbb{R}$. Écrivez-moi si vous avez un problème ou un doute.

¡¡ Avertissement/question à la lectrice !! 12.376
De plus, l'énoncé de 12.377 demande la convergence uniforme des dérivées directionnelles dans toutes les directions. Je ne serais pas étonné que la convergence uniforme seulement des dérivées partielles dans les directions « de base » suffise.

ThoSerUnifDerr

Théorème 12.377 ([326, 357, 1], thème 62)**.**
Soient des espaces vectoriels normés V et W. Soient un ouvert U de V et des fonctions $f_k\colon U\to W$, une autre fonction $f\colon U\to W$ ainsi que, pour toute direction[143] *$\alpha\in V$, des fonctions $g_\alpha\colon U\to W$. Nous supposons que*

(1) Les f_k sont de classe C^1.
(2) $f_k\to f$ simplement.
(3) $\partial_\alpha f_k\to g_\alpha$ uniformément sur tout compact.

Alors

ITEMooQOSUooQGSUXC

(1) Nous avons la convergence $f_i\to f$ uniformément sur tout compact.

ITEMooGFPLooGYEvkh

(2) Pour toute direction α, nous avons $\partial_\alpha f=g_\alpha$.
(3) La fonction f est de classe C^1 sur U.

Démonstration. En plusieurs parties.

(1) **Uniforme convergence sur les boules fermées** Soit $a\in U$. Nous considérons $r>0$ tel que $\overline{B(a,r)}\subset U$ (ça existe parce que U est ouvert ; il suffit de prenre r plus petit qu'un qui fait que $B(a,r)\subset U$), et nous posons $K=\overline{B(a,r)}$. Nous allons prouver l'uniforme convergence $f_i\xrightarrow{\|.\|_{\overline{B(a,r)}}}f$, et nous verrons plus tard comment faire pour l'uniforme convergence sur un compact général dans U.

Nous restreignons toutes les fonctions à K. Nous notons $\omega_{\alpha,i}$ le module de continuité[144] de $\partial_\alpha f_i$. Pour chaque $n\in\mathbb{N}$ nous définissons encore

$$\begin{aligned}\alpha_n\colon V&\to V\\ x&\mapsto\frac{a-x}{n+1}.\end{aligned}\tag{12.1018}$$

Soit $x\in\overline{B(a,r)}$. Nous écrivons la somme télescopique

$$f_i(x)=f_i(a)+\sum_{k=0}^{n}\big[f_i\big(a+(k+1)\alpha_n(x)\big)-f_i\big(a+k\alpha_n(x)\big)\big]\tag{12.1019}$$

EQooMXVLooXFcsGH

143. Ici le mot « direction » n'a pas de sens particulier ; c'est juste un élément quelconque. Si nous faisions de la géométrie différentielle hard-core, ce serait un vecteur tangent.
144. Définition 10.310.

Notez que les points auxquels sont évalués f_i sont dans $\overline{B(a,r)}$ parce que, si $l \in [0, n+1]$, nous avons

$$\|a + l\alpha_n(x) - a\| = \|l\alpha_n(x)\| = l\frac{\|a-x\|}{n+1} \leqslant \|a-x\| \leqslant r. \tag{12.1020}$$

Ce point est important parce que rien ne nous dit que U est convexe ; pour la suite nous avons besoin que tous les points sur les segments entre a et les différents points que nous allons considérer restent dans $\overline{B(a,r)}$. C'est d'ailleurs pour cette convexité de la boule que nous commençons notre preuve par le cas où K est une boule. Bref.

Le théorème des accroissements finis 12.236 nous assure l'existences d'éléments

$$y_{k,i} \in \big[a + (k+1)\alpha_n(x), a + k\alpha_n(x)\big] \tag{12.1021}$$

tels que

$$f_i\big(a + (k+1)\alpha_n(x)\big) - f_i\big(a + k\alpha_n(x)\big) = (\partial_\beta f)\big(y_{k,i}(x)\big)\alpha_n(x). \tag{12.1022}$$

Notez que les $y_{j,i}(x)$ sont dans $\overline{B(a,r)}$.

Soit $\epsilon > 0$. Nous allons calculer $\|f_i(x) - f_j(x)\|$ en substituant les valeurs de $f_i(x)$ et $f_j(x)$ données par (12.1019). Nous avons :

$$\clubsuit = \|f_i(x) - f_j(x)\| \leqslant \|f_i(a) - f_j(a)\| + \sum_{k=0}^{n} \|\alpha_n(x)\|\|(\partial_\beta f_i)\big(y_{k,i}(x)\big) - (\partial_\beta f_j)\big(y_{k,j}(x)\big)\| \tag{12.1023a}$$

$$\leqslant \epsilon + \frac{r}{n+1}\sum_{k=0}^{n} \|(\partial_\beta f_i)\big(y_{k,i}(x)\big) - (\partial_\beta f_j)\big(y_{k,j}(x)\big)\| \tag{12.1023b}$$

$$\leqslant \epsilon + \frac{r}{n+1}\sum_{k=0}^{n} \|(\partial_\beta f_i)\big(y_{k,i}(x)\big) - (\partial_\beta f_i)\big(y_{k,j}(x)\big)\| \tag{12.1023c}$$

$$+ \frac{r}{n+1}\sum_{k=0}^{n} \|(\partial_\beta f_i)\big(y_{k,j}(x)\big) - (\partial_\beta f_j)\big(y_{k,j}(x)\big)\|$$

$$\leqslant \epsilon + \frac{r}{n+1}\sum_{k=0}^{n} \big\|\omega_{\beta,i}\big(\|y_{k,i}(x) - y_{k,j}(x)\|\big)\big\| \tag{12.1023d}$$

$$+ \frac{r}{n+1}\sum_{k=0}^{n} \|\partial_\beta f_i - \partial_\beta f_j\|$$

$$\leqslant \epsilon + \frac{r}{n+1}\sum_{k=0}^{n} \big\|\omega_{\beta,i}\big(\|y_{k,i}(x) - y_{k,j}(x)\|\big)\big\| + \frac{r}{n+1}\sum_{k=0}^{n}\epsilon \tag{12.1023e}$$

Justifications :

— Pour (12.1023c). Majoration $\|\alpha_n(x)\| \leqslant \frac{r}{n+1}$. De plus nous considérons des i et j assez grands pour que $\|f_i(a) - f_j(a)\|_K \leqslant \epsilon$. Cela est possible parce que (f_i) est une suite de Cauchy pour la norme uniforme sur K.

— Pour (12.1023d). Utilisation du module de continuité, définition 10.310.

— Pour (12.1023e). Nous avons la convergence uniforme $\partial_\beta f_i \xrightarrow{\overline{B(a,r)}} g_\beta$, de sorte que $i \mapsto \partial_\beta f_i$ est une suite de Cauchy. Si i et j sont assez grands, nous pouvons majorer $\|\partial_\beta f_i - \partial_\beta f_j\| \leqslant \epsilon$.

Étudions deux minutes ce qui est dans le module de continuité de (12.1023e). Nous avons $y_{k,i}(x) \in \big[a + (k+1)\alpha_n(x), a + k\alpha_n(x)\big]$, donc la différence $\|y_{k,i}(x) - y_{k,j}(x)\|$ se majore par la taille de cet intervalle :

$$\|y_{k,i}(x) - y_{k,j}(x)\| \leqslant \|a - k\alpha_n(x) - \big(a + (k+1)\alpha_n(x)\big)\| = \|\alpha_n(x)\|. \tag{12.1024}$$

Vu que le module de continuité est une fonction croissante,

$$\omega_{\beta,i}\big(\|y_{k,i}(x)-y_{k,j}(x)\|\big)\leqslant\omega_{\beta,i}\big(\|\alpha_n(x)\|\big)\leqslant\omega_{\beta,i}\left(\frac{\|a-x\|}{n+1}\right)\leqslant\omega_{\beta,i}\left(\frac{r}{n+1}\right).\tag{12.1025}$$

En substituant tout ça dans (12.1023e), nous continuons :

$$\clubsuit\leqslant\epsilon+\frac{r}{n+1}\sum_{k=0}^{n}\omega_{\beta,i}\big(\frac{r}{n+1}\big)+r\epsilon\leqslant\epsilon+r\omega_{\beta,i}\left(\frac{r}{n+1}\right)+r\epsilon\tag{12.1026}$$

Résumons. Pour tout $x\in\overline{B(a,r)}$, pour tout $n\in\mathbb{N}$, si i et j sont assez grands, nous avons la majoration

$$\|f_i(x)-f_j(x)\|\leqslant\epsilon+r\omega_{\beta,i}\big(\frac{r}{n+1}\big)+r\epsilon.\tag{12.1027}$$

Nous pouvons prendre la limite $n\to\infty$ de deux côtés. Vu que $\partial_\beta f_i$ est uniformément continue[145], le module de continuité tend vers zéro (lemme 10.312).
Si i et j sont assez grands, nous avons donc

$$\|f_i(x)-f_j(x)\|\leqslant\epsilon(1+r),\tag{12.1028}$$

et donc aussi

$$\|f_i-f_j\|_{\overline{B(a,r)}}\leqslant\|f_i(x)-f_j(x)\|\leqslant\epsilon(1+r).\tag{12.1029}$$

Cela prouve que (f_i) est une suite de Cauchy pour $\|.\|_K$. Donc f_i converge uniformément vers une certaine fonction. Vu qu'elle converge simplement vers f, elle converge uniformément vers f.
Nous avons donc prouvé que

$$f_i\xrightarrow{\|.\|_{\overline{B(a,r)}}}f,\tag{12.1030}$$

c'est-à-dire la convergence uniforme sur toute boule compacte.

(2) **Convergence uniforme sur tout compact** Soient un compact K de U, et $\epsilon>0$. L'ensemble $\{B(x,r)\}_{x\in K,r>0}$ est un recouvrement de K par des ouverts. On en extrait un sous-recouvrent fini, et on ferme les boules sans changer le fait que ce soit un recouvrement :

$$K\subset\bigcup_{i=1}^{n}\overline{B(a_i,r_i)}.\tag{12.1031}$$

Vue la convergence uniforme sur toute boule fermée, pour chaque i, il existe N_i tel que $n>N_i$ implique

$$\|f_n-f\|_{\overline{B(a_i,r_i)}}<\epsilon.\tag{12.1032}$$

En prenant $N>\max\{N_i\}$, nous avons

$$\|f_n-f\|_K<\epsilon\tag{12.1033}$$

pour tout $n>N$.

(3) **Pour (2)** Soit $a\in U$ et un voisinage compact $K=\overline{B(a,\delta)}$ de a. Nous considérons une direction u avec $\|u\|=1$. Nous calculons un peu :

$$\begin{aligned}\heartsuit=\|\frac{f(a+\delta u)-f(a)}{\delta}-g_u(a)\|&\leqslant\|\frac{f(a+\delta u)-f_i(a+\delta u)}{\delta}\|&(12.1034a)\\&+\|\frac{f_i(a+\delta u)-f_i(a)}{\delta}-g_u(a)\|\\&+\|\frac{f_i(a)-f(a)}{\delta}\|\\&\leqslant\frac{2}{\delta}\|f_i-f\|+\|\frac{f_i(a+\delta u)-f_i(a)}{\delta}-g_u(a)\|&(12.1034b)\end{aligned}$$

145. Elle est continue sur un compact, proposition 12.78.

Ici, la norme $\|f_i-f\|$ est une norme supremum sur K (vous devriez l'avoir deviné du contexte). C'est le moment d'utiliser le théorème des accroissements finis 12.236 : il existe $y\in[a+\delta u,a]$ tel que

$$\frac{f_i(a+\delta u)-f_i(a)}{\delta}=(\partial_i f_i)(y) \tag{12.1035}$$

Nous continuons :

$$\heartsuit \leqslant \frac{2}{\delta}\|f_i-f\|+\|(\partial_u f_i)(y)-g_u(a)\| \tag{12.1036a}$$
$$\leqslant \frac{2}{\delta}\|f_i-f\|+\|(\partial_u f_i)(y)-(\partial_u f_i)(a)\|+\|(\partial_u f_i)(a)-g_u(a)\| \tag{12.1036b}$$

SUBEQooCLCQooGcFAZD

Nous introduisons le module de continuité[146] $\omega_{i,u}$ de $(\partial_u f_i)$ pour traiter le premier terme :

$$\|(\partial_u f_i)(y)-(\partial_u f_i)(a)\|\leqslant\omega_{i,u}\big(\|y-a\|\big)\leqslant\omega_{i,u}(\delta). \tag{12.1037}$$

Nous utilisons aussi la convergence uniforme sur tout compact (point (1)) $\partial_u f_i\xrightarrow{\|.\|_K}g_u$ pour majorer le second terme de (12.1036b) par ϵ lorsque i est grand.

Nous continuons. Pour tout i assez grand, nous avons

$$\|\frac{f(a+\delta u)-f(a)}{\delta}-g_u(a)\|\leqslant\frac{2}{\delta}\|f_i-f\|+\epsilon+\omega_{i,u}(\delta). \tag{12.1038}$$

Nous prenons la limite $i\to\infty$ en tenant compte du lemme 10.313 : $\lim_{i\to\infty}\omega_{i,u}(\delta)=\omega_g(\delta)$:

$$\|\frac{f(a+\delta u)-f(a)}{\delta}-g_u(a)\|\leqslant\epsilon+\omega_g(\delta). \tag{12.1039}$$

Et enfin en prenant la limite $\delta\to0$ nous trouvons que pour tout ϵ,

$$\lim_{\delta\to0}\|\frac{f(a+\delta u)-f(a)}{\delta}-g_u(a)\|\leqslant\epsilon, \tag{12.1040}$$

et donc nous avons prouvé que

$$\lim_{\delta\to0}\|\frac{f(a+\delta u)-f(a)}{\delta}-g_u(a)\|=0. \tag{12.1041}$$

Cela prouve que $(\partial_u f)(a)$ existe et vaut $g_u(a)$.

Vu que les g_u sont continues, la fonction f est également de classe C^1 par le théorème 12.294. □

12.34 La fonction puissance

Si x et y sont des réels, définir x^y n'est pas une mince affaire. Pour l'instant nous savons déjà définir x^n lorsque $x\in\mathbb{R}$ et $n\in\mathbb{N}$. Voir la définition 1.206 et le thème 52.

Pour la suite nous notons

$$\begin{cases} f_\alpha(x)=x^\alpha & \text{(12.1042a)}\\ g_a(x)=a^x & \text{(12.1042b)}\end{cases}$$

pour autant que ces fonctions sont définies[147].

146. Définition 10.310.

147. L'objet des pages suivantes est de déterminer pour quelles valeurs de a, α et x nous pouvons trouver des définitions raisonnables pour ces fonctions.

12.34.1 Sur les naturels

DEFooKEBIooZtPkac

Définition 12.378.
La fonction puissance définie sur $\mathbb{N}$ s'étend à $\mathbb{Z}$ de la façon suivante :

$$x^{-n} = \frac{1}{x^n} \tag{12.1043}$$

pour $n \geqslant 0$. Cela donne donne donc x^n pour $x \in \mathbb{R}$ et $n \in \mathbb{Z}$ a l'exception de $x = 0$ lorsque $n < 0$.

Nous étudions quelques propriétés de cette fonction pour $n > 0$ fixé.

12.379.
La limite

$$\lim_{x \to \infty} x^n = \infty \tag{12.1044}$$

demande la topologie sur la droite réelle achevée. C'est le lemme 12.29.

PROPooXQYFooPxoEHE

Proposition 12.380.
Soit $n \in \mathbb{N}\backslash\{0\}$; nous posons $f_n(x) = x^n$.

Si n est pair,

$$f_n \colon [0, \infty[\to [0, \infty[\tag{12.1045}$$

est bijective.

Si n est impair,

$$f_n \colon \mathbb{R} \to \mathbb{R} \tag{12.1046}$$

est bijective.

Toutes les fonctions f_n sont continues sur $\mathbb{R}$.

Démonstration. En plusieurs morceaux, pas spécialement dans l'ordre auquel on s'attend.

(1) **Continuité**
Soit $x \in \mathbb{R}$. En vertu de 12.53 nous allons prouver que $\lim_{\epsilon \to 0} f_n(x + \epsilon) = f_n(x)$. Pour cela nous utilisons la formule du binôme 3.43 avec $x, h > 0$:

$$f_n(x + h) = (x + h)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} h^k. \tag{12.1047}$$

Nous fixons $x_0 \in \mathbb{R}$. Calcul :

$$|f_n(x_0 + h) - f_n(x)| = |\sum_{k=1}^{n} \binom{n}{k} x_0^{n-k} h^k| \tag{12.1048a}$$

$$\leqslant \sum_{k=1}^{n} \binom{n}{k} |x_0|^{n-k} |h|^k \tag{12.1048b}$$

$$= h \sum_{k=1}^{n} \binom{n}{k} |x_0|^{n-k} |h|^{k-1} \tag{12.1048c}$$

$$\leqslant h \sum_{k=1}^{n} \binom{n}{k} |x_0|^{n-k}. \tag{12.1048d}$$

Justifications :

— Le terme $k = 0$ est égal à $x^n = f_n(x)$ parce que $\binom{n}{0} = 1$.

— Dans la somme nous avons majoré $|h|$ par 1, opération justifiée par le fait que nous ayons dans l'idée de faire $h \to 0$.

Nous avons donc

$$\lim_{h \to 0} |f_n(x_0 + h) - f_n(x)| \leqslant \lim_{h \to 0} h \sum_{k=1}^{n} \binom{n}{k} |x_0|^{n-k} = 0. \tag{12.1049}$$

D'où la continuité de f_n en tout point $x_0 \in \mathbb{R}$.

(2) **Pour n pair ou impair, bijection sur les positifs** Ceci sera déjà le résultat complet pour les n pairs, et a moitié du résultat pour les n impairs.

(2a) **Stricte croissance** Soit $n \neq 0$ dans $\mathbb{N}$. Nous commençons par prouver que f_n est strictement croissante sur $[0,\infty[$. Nous repartons de la formule du binôme, mais cette fois, nous séparons les termes $k = 0$ et $k = n$ des autres (si $n = 1$, il y a un peu de réécriture) en tenant compte de $\binom{n}{0} = \binom{n}{n} = 1$:

$$f_n(x+h) = x^n + h^n + \sum_{k=1}^{n-1} \binom{n}{k} x^{n-k}h^k > x^n = f_n(x). \tag{12.1050}$$

Vous noterez que l'inégalité est stricte même si $n = 1$.

Vu que nous avons stricte monotonie, le théorème 12.51(2) nous dit que

$$f_n\colon [0,\infty[\to f_n\big([0,\infty[\big) \tag{12.1051}$$

est une bijection.

(2b) **Bijection**

Nous prouvons que $f_n\big([0,\infty[\big) = [0,\infty[$. Si $x > 0$ alors $f_n(x) > 0$, cela prouve une inclusion.

Pour l'autre inclusion nous savons que $\lim_{x\to\infty} f_n(x) = \infty$ par le lemme 12.29. Si $y \in [0,\infty[$, alors il existe x_0 tel que $f_n(x_0) > y$. Étant donné que $f_n(0) = 0$ et que nous avons déjà prouvé que f_n était continue (proposition 12.380), le théorème des valeurs intermédiaires 9.93 nous indique l'existence de $x_1 \in [0,x_0[$ tel que $f_n(x_1) = y$.

Nous avons prouvé que pour tout n, la fonction

$$f_n\colon [0,\infty[\to [0,\infty[\tag{12.1052}$$

EQooYWHGooJWMTUI

est une bijection.

(3) **Pour n impair**

Nous montrons à présent que si n est impair, alors

$$f_n\colon]-\infty,0] \to]-\infty,0] \tag{12.1053}$$

EQooTSLJooMAAUXH

est une bijection.

Tout se base sur le fait que si $x > 0$ alors $f_n(-x) = -f_n(x)$. Le fait que (12.1052) soit injective et surjective montre alors tout de suite le fait que (12.1053) soit également injective et surjective.

□

Vous noterez que la continuité de f_n démontrée dans la proposition 12.380 est indépendant de la proposition 12.61 qui sera invoquée plus tard pour définir a^x lorsque $a > 0$ dans $\mathbb{R}$.

12.34.2 Sur les rationnels, racines

L'existence, pour tout réel $a \geqslant 0$, d'un réel r tel que $r^2 = a$ est déjà faite en la proposition 1.474.

DEFooJWQLooWkOBxQ

Définition 12.381 (Exposant rationnels)**.**
La proposition 12.380 nous dit entre autres que pour tout $n \in \mathbb{N}$, la fonction

$$\begin{aligned} f_n\colon [0,\infty[&\to [0,\infty[\\ x &\mapsto x^n \end{aligned} \tag{12.1054}$$

est bijective. Nous définissions alors, pour $a \in [0,\infty[$,

$$a^{1/n} = f_n^{-1}(a). \tag{12.1055}$$

Autrement dit, le nombre $a^{1/n}$ est l'unique solution positive de

$$x^n = a. \tag{12.1056}$$

NORMooDUNZooUNdUKg

12.382.
Nous ne définissons pas $a^{1/n}$ pour $a < 0$, du moins pas encore. Vu que f_3 est bijective sur $\mathbb{R}$, il serait tentant de définir $(-1)^{1/3} = f_3^{-1}(-1) = -1$.

Cela causera un certain nombre de problèmes plus tard vu que nous aurons envie de deux choses en même temps :

— d'une part $\ln(-1) = i\pi$,

— d'autre part, $a^x = e^{x\ln(a)}$.

De cette façon, nous devrions avoir

$$(-1)^{1/3} = e^{i\pi/3}, \tag{12.1057}$$

qui est un nombre complexe non réel. Voici un exemple de ce que ça donne avec Sage :

```
sage: a=(-1)**(1/3)
sage: a.real_part()
1/2
sage: a.imag_part()
1/2*sqrt(3)
```

tex/sage/sageSnip019.sage

DEFooPOELooPouwtD

Définition 12.383 (Racine)**.**
Pour $n \in \mathbb{N}$ nous définissons $\sqrt[n]{x} = f_n^{-1}(x)$. Lorsque n est pair, la fonction $x \mapsto \sqrt[n]{x}$ n'est définie que sur $\mathbb{R}^+$, et lorsque n est impair, elle est définie sur tout $\mathbb{R}$.

NORMooYPRNooWCjEgR

12.384.
Notons que les fonctions $x \mapsto \sqrt[3]{x}$ et $x \mapsto x^{1/3}$ ne sont pas les mêmes : la première est définie sur tout $\mathbb{R}$ et donne des valeurs réelles tandis que la seconde n'est (pour l'instant) définie que sur les positifs, et donnera (quand on l'aura définie par l'exponentielle) des nombres complexes sur les négatifs.

En suivant cette convention, c'est-à-dire en réservant la notation $\sqrt{\ }$ pour l'inverse de f_2, nous ne devrions pas écrire des choses comme « $\sqrt{-1} = i$ », mais plutôt « $(-1)^{1/2} = i$ ». En effet, $\sqrt{-1}$ n'est pas défini et ne sera jamais défini alors que $(-1)^{1/2}$ n'est pas encore défini, mais sera défini par

$$(-1)^{1/2} = e^{\frac{1}{2}\ln(-1)} = e^{i\pi/2} = i. \tag{12.1058}$$

En résumé, nous avons les fonctions suivantes :

(1) $\sqrt[n]{\ }: \mathbb{R} \to \mathbb{R}$ si n est impair,

(2) $\sqrt[n]{\ }: [0, \infty[\to [0, \infty[$ si n est pair,

(3) $x^{1/n}: [0, \infty[\to [0, \infty[$ pour tout $n \in \mathbb{N}$.

Cependant nous n'hésiterons pas à utiliser la notation $\sqrt{x}$ pour $x^{1/2}$ même lorsque x est négatif, parce c'est une notation très pratique. Il faut garder en tête que cette façon de faire est incohérente parce qu'elle inciterait à penser que $\sqrt[3]{-1} = e^{i\pi/3}$ au lieu de $\sqrt[3]{-1} = -1$.

Pour toute la suite de cette section, nous allons considérer a^x uniquement pour $a > 0$.

DEFooHXUFooOJTVXA

Définition 12.385.
Pour $m, n \in \mathbb{N}$ nous définissons ($a^{1/n}$ est la définition 12.381)

$$a^{m/n} = (a^m)^{1/n}, \tag{12.1059}$$

EQooZFQAooTsMbub

ce qui définit la fonction puissance sur $\mathbb{Q}^+$. Enfin nous posons

$$a^{-q} = \frac{1}{a^q} \tag{12.1060}$$

DEFooTUCVooXjkxRb

lorsque $q \in \mathbb{Q}^+$.

Et avec tout ça, lorsque $a > 0$ nous avons défini a^q pour tout $q \in \mathbb{Q}$.

Nous allons souvent noter la définition (12.1059) sous la forme

$$f_{m/n}(x)^n = x^m. \tag{12.1061}$$

EQooZUKKooVfikZo

Lemme 12.386 ([1]).
Si $a, b > 0$, nous avons

LEMooOFPMooIEmSNA

ITEMooEFUAooYBeJza

(1) Pour $n \in \mathbb{N}$ nous avons $(ab)^{1/n} = a^{1/n}b^{1/n}$.

ITEMooHGPPooDBzWKx

(2) Pour $n, m \in \mathbb{N}$ nous avons $(ab)^{1/n} = a^{1/n}b^{1/n}$.

ITEMooUYTLooHzXwtf

(3) Pour $q \in \mathbb{N}$ nous avons $(ab)^q = a^q b^q$.

Démonstration. Point par point.

(1) **Pour (1)** Utilisant la formule pour les exposants entiers, nous avons

$$\left[(a^{1/n})(b^{1/n})\right]^n = (a^{1/n})^n(b^{1/n})^n = ab. \tag{12.1062}$$

Donc le nombre $x = a^{1/n}b^{1/n}$ vérifie $x^n = ab$. Autrement dit $x = f_n^{-1}(ab) = (ab)^{1/n}$

(2) **Pour (2)**

(3) **Pour (3)**

□

LEMooIDLJooZALNaD

Lemme 12.387 ([1]).
Pour $a > 0$ et $p, q \in \mathbb{Z}$ nous avons :

$$a^{p/q} = (a^p)^{1/q} = (a^{1/q})^p. \tag{12.1063}$$

Démonstration. Nous divisons la preuve en fonction de la positivité du numérateur et du dénominateur.

(1) **Numérateur et dénominateurs positifs**
Nous commençons avec $p, q \in \mathbb{N}$. La première égalité est la définition 12.381. Pour la seconde, la définition de $(a^p)^{1/q}$ est d'être le $x > 0$ tel que

$$x^q = a^p. \tag{12.1064}$$

La définition de $a^{1/q}$ est d'être le $y > 0$ tel que

$$y^q = a. \tag{12.1065}$$

Ce y vérifie donc aussi $y^{pq} = a^p$ et donc $(y^p)^q = a^p$. Autrement dit, $y^p = x$, c'est-à-dire exactement

$$(a^{1/q})^p = (a^p)^{1/q}. \tag{12.1066}$$

Le lemme est prouvé dans le cas où $p, q \in \mathbb{N}$.

(2) **Numérateur et dénominateur négatifs**
Si p et q sont tous les deux négatifs, nous remarquons que $p/q = (-p)/(-q)$ et nous sommes dans le même cas qu'avant.

(3) **Numérateur négatif, dénominateur positif**
Pour simplifier les notations nous supposons toujours $p, q \in \mathbb{N}$ mais nous considérons $a^{(-p)/q}$. Nous avons d'une part :

$$a^{(-p)/q} = a^{-(p/q)} = \frac{1}{a^{p/q}} = \frac{1}{(a^{1/q})^p} = (a^{1/q})^{-p}. \tag{12.1067}$$

Dans ce calcul, nous avons utilisé au dénominateur le résultat dans le cas positif.
Le lemme 7.250(3) nous dit que $(1/a^p) = (1/a)^p$. Nous faisons le calcul suivant :

$$(a^{-p})^{1/q} = \left(\left(\frac{1}{a}\right)^p\right)^{1/q} = \left(\left(\frac{1}{a}\right)^{1/q}\right)^p = \left(\frac{1}{a^{1/q}}\right)^p = \frac{1}{(a^{1/q})^p} = (a^{1/q})^{-p} \tag{12.1068}$$

où nous avons utilisé le résultat avec $1/a$ en guise de a.

(4) **Numérateur positif, dénominateur négatif**
Nous traitons maintenant $a^{p/(-q)}$. Nous avons d'une part

$$a^{p/(-q)} = a^{-(p/q)} = \frac{1}{a^{p/q}} = \frac{1}{(a^p)^{1/q}} = (a^p)^{-(1/q)} = (a^p)^{1/(-q)}. \tag{12.1069}$$

Et d'autre part :

$$a^{p/(-q)} = \frac{1}{a^{p/q}} = \frac{1}{(a^{1/q})^p} = \left(\frac{1}{a^{1/q}}\right)^p = \left(a^{-(1/q)}\right)^p = (a^{1/(-q)})^p. \tag{12.1070}$$

□

Le lemme suivant montre que la définition sur $\mathbb{Q}^-$ est cohérente avec celle sur $\mathbb{Q}^+$, au sens où finalement nous retrouvons que $a^{m/n}$ vérifie $x^n = a^m$ quel que soient les signes de m et n.

Lemme 12.388 ([1]).
Le nombre $y = a^{-m/n}$ vérifie l'équation $y^{-n} = a^m$

Démonstration. Nous posons $x = a^{m/n}$, c'est-à-dire $x^n = a^m$. Nous avons, par définition $y = a^{-m/n} = \frac{1}{x}$. Alors

$$y^{-n} = \frac{1}{\left(\frac{1}{x}\right)^n} = x^n = a^m, \tag{12.1071}$$

donc c'est bon. □
LEMooJYGUooHhLASp

Lemme 12.389 ([1]).
Pour $a > 0$ et $q, q' \in \mathbb{Q}$ nous avons

$$a^q a^{q'} = a^{q+q'}. \tag{12.1072}$$

Démonstration. Nous mettons q et q' au même dénominateur. Soient $q = s/c$ et $q' = r/c$ avec $s, r \in \mathbb{Z}$ et $c \in \mathbb{N}$. En utilisant les égalités du lemme 12.387 nous trouvons

$$a^{s/c}a^{r/c} = (a^{1/c})^s(a^{1/c})^r = (a^{1/c})^{s+r} = a^{(s+r)/c} = a^{q+q'}. \tag{12.1073}$$

□
LEMooXJXUooLoiTMo

Lemme 12.390 ([1]).
La fonction puissance prend les valeurs suivantes.

(1) Si $a = 1$ alors $a^q = 1$ pour tout $q \in \mathbb{Q}$.
ITEMooKZCGooKskUQx

(2) Si $a > 1$ alors
- *$a^q > 1$ si $q > 0$*
- *$a^q < 1$ si $q < 0$*
- *$a^0 = 1$.*

(3) Si $a < 1$ alors
- *$a^q < 1$ si $q > 0$*
- *$a^q > 1$ si $q < 0$*
- *$a^0 = 1$.*

Démonstration. Si $a = 1$ alors $a^k = 1$ pour tout $k \in \mathbb{N}$. Ensuite, pour $m, n \in \mathbb{N}$, $a^{n/m}$ est solution de $x^m = a^n = 1$, donc $x = 1$. En ce qui concerne les puissances négatives, $1/1 = 1$.

Si $a > 1$ alors $a^k > 1$ pour tout $k \in \mathbb{N}$. De plus pour $q > 0$ nous avons $q = m/n$ avec $m, n \in \mathbb{N}$. Alors $a^{m/n}$ est solution de $x^m = a^n > 1$. Or pour $x \leqslant 1$ nous avons $x^m \leqslant 1$, donc la solution à $x^m = a^n$ vérifie forcément $x > 1$.

Toujours avec $a > 1$, si $q < 0$ nous posons $q = -q'$ avec $q' > 0$. Alors

$$a^q = q^{-q'} = \frac{1}{a^{q'}}. \tag{12.1074}$$

Mais $a^{q'} > 1$, donc l'inverse est inférieur à 1.

En ce qui concerne les cas $a < 1$, ils sont obtenus en posant $b = 1/a$ et en calculant

$$a^q = \left(\frac{1}{b}\right)^q = \frac{1}{b^q} = b^{-q}. \tag{12.1075}$$

□

PROPooVXKBooQPPjMn

Proposition 12.391 ([1])**.**
Soit $a > 1$. Alors

$$\lim_{n\to\infty} a^n = \infty. \tag{12.1076}$$

Démonstration. Soient $a > 1$ et $M > 0$. Nous devons prouver qu'il existe $n \in \mathbb{N}$ tel que $a^n > M$. Nous posons $a = 1 + h$. Alors en utilisant la formule du binôme,

$$a^n = (1 + h)^n = \sum_{k=0}^{n} \binom{n}{k} h^{n-k}. \tag{12.1077}$$

Tous les termes de la somme sont strictement positifs. Prenons le terme $k = n - 1$. Il vaut

$$\binom{n}{n-1} h = nh. \tag{12.1078}$$

Donc $a^n \geqslant nh$, donc oui, cela peut être rendu arbitrairement grand avec n sans toucher à a parce que $\mathbb{N}$ est archimédien par la proposition 1.84. □

PROPooGCBZooTcyGtO

Proposition 12.392 ([1])**.**
Pour $a > 0$ nous considérons la fonction

$$\begin{aligned} g_a \colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto a^q. \end{aligned} \tag{12.1079}$$

(1) Si $a \in]0, 1[$ alors g_a est décroissante et

$$\lim_{q\to\infty} g_a(q) = 0, \qquad \lim_{q\to-\infty} g_a(q) = \infty. \tag{12.1080a}$$

ITEMooGOEVooKVoVpZ

(2) Si $a > 1$ alors g_a est croissante et

$$\lim_{q\to\infty} g_a(q) = \infty, \qquad \lim_{q\to-\infty} g_a(q) = 0. \tag{12.1081a}$$

Démonstration. Nous prouvons le cas $a > 1$. L'autre cas s'en déduit en posant $b = 1/a$. Pour la croissance, soient $q \in \mathbb{Q}$ et $r > 0$ dans $\mathbb{Q}$. En utilisant le lemme 12.389, nous avons

$$a^{q+r} = a^q a^r > a^q \tag{12.1082}$$

parce que $a^r > 1$ par le lemme 12.390.

En ce qui concerne la limite $q \to \infty$, la fonction g_a est croissante et non bornée par la proposition 12.391. Donc sa limite est ∞.

Pour la limite $q \to -\infty$, nous avons

$$\lim_{q\to-\infty} a^q = \lim_{q\to\infty} a^{-q} = \lim_{q\to\infty} \frac{1}{a^q} = 0. \tag{12.1083}$$

□

PROPooIIDGooTRtlUD

Proposition 12.393 ([1]).
Soit $a > 0$. Nous avons

$$\lim_{q\to 0} a^q = 1. \tag{12.1084}$$

Notons que cette limite est une limite dans $\mathbb{Q}$ parce que nous n'avons même pas encore défini a^x lorsque x est irrationnel.

Démonstration. Nous notons, comme à l'accoutumée, $g_a(x) = a^x$. Nous allons prouver que pour toute suite (x_k) dans $\mathbb{Q}$ telle que $x_k \to 0$, nous avons $a^{x_k} \to 1$.

(1) **Suite positive** Soit une suite de rationnels $x_k \to 0$ avec $x_k > 0$. En définissant y_k par $x_k = 1/y_k$ nous savons que a^{1/y_k} est la solution de $x^{y_k} = a$.
Nous posons $t_k = a^{x_k}$ et notre but est de prouver que $t_k \to 1$. Pour tout k nous avons la relation

$$t_k^{y_k} = a. \tag{12.1085}$$

Soit $s > 1$. Il existe un $M > 0$ tel que $y_k > M$ implique $s^{y_k} > a$ (proposition 12.391). Donc dès que $y_k > M$ nous avons $t_k < s$.
De la même manière, si $r < 1$, il existe un $R > 0$ tel que $y_k > R$ implique $r^{y_k} < a$. Donc dès que $y_k > R$ nous avons $t_k > r$.
Soit donc un voisinage $]r, s[$ de 1 (avec $r < 1$ et $s > 1$). Nous avons les nombres M et R correspondant et nous posons $L = \max\{M, R\}$. Soit K tel que $k > K$ implique $y_k > L$. Alors pour $k > K$ nous avons aussi $t_k < s$ et $t_k > r$, c'est-à-dire $t_k \in]r, s[$.
Cela prouve que $t_k \to 1$.

(2) **Suite négative** Supposons maintenant que (x_k) est une suite de rationnels vérifiant $x_k \to 0$ avec $x_k < 0$. En posant $z_k = -x_k$, la suite (z_k) est dans le cas précédent et nous avons

$$a^{x_k} = a^{-z_k} = \frac{1}{a^{z_k}} \to 1. \tag{12.1086}$$

(3) **Suite générale** Si $x_k \xrightarrow{\mathbb{Q}} 0$ avec $x_k \neq 0$. Nous considérons la suite $t_k = a^{x_k}$. Toute sous-suite contient une sous-suite avec $x_k > 0$ ou une sous-suite avec $x_k < 0$. Dans les deux cas, la sous-sous-suite converge vers 1. Le lemme 7.66 conclut que $t_k \to 1$.

(4) **Conclusion** Pour toute suite $x_k \to 0$ nous avons $g_a(x_k) \to 1$. Par le critère séquentiel de la limite (proposition 7.249) nous avons $\lim_{x\to 0} g_a(x) = 1$.

□

LEMooKDBPooLQwxMD

Lemme 12.394.
Soit $a > 0$. La fonction

$$\begin{aligned} g_a\colon \mathbb{Q} &\to \mathbb{R} \\ x &\mapsto a^x \end{aligned} \tag{12.1087}$$

est continue.

Démonstration. Soient $x \in \mathbb{Q}$ et une suite $x_k \to 0$ (toujours dans $\mathbb{Q}$) et utilisons le lemme 12.389 :

$$a^{x+x_k} = a^x a^{x_k}. \tag{12.1088}$$

Cela est, dans $\mathbb{R}$, le produit entre une constante (a^x) et une suite. La limite est donc le produit de cette constante et la limite de la suite (si elle existe). Par la proposition 12.393 nous avons la limite $a^{x_k} \to 1$, et donc

$$\lim_{k\to\infty} a^{x+x_k} = a^x, \tag{12.1089}$$

ce qui prouve la continuité (caractérisation séquentielle, proposition 7.137) de g_a. □

PROPooQRFSooVzYdJM

Proposition 12.395 ([1, 162]).
Soit $a > 0$ dans $\mathbb{R}$. La fonction

$$\begin{aligned} g_a\colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto a^q \end{aligned} \tag{12.1090}$$

est Cauchy-continue.

Démonstration. En quelque étapes.

(1) **Pour $a > 1$** Avant de nous lancer dans la preuve directe, nous prouvons une petite formule. Soit $\epsilon > 0$. Vu que, par la proposition 12.393, $\lim_{q\to 0} g_a(q) = 1$, il existe $\delta > 0$ tel que $0 < q < \delta$ implique $|1 - g_a(q)| < \epsilon$.
Soient maintenant $p, q \in \mathbb{Q}$ tels que $|p - q| < \delta$. En utilisant de plus la définition (12.1060) et la formule du lemme 12.389,

$$|g_a(q) - g_a(p)| = |g_a(q)| \left|1 - \frac{g_a(p)}{g_a(q)}\right| \tag{12.1091a}$$
$$= |g_a(q)||1 - g_a(p - q)| \tag{12.1091b}$$
$$\leqslant |g_a(q)|\epsilon. \tag{12.1091c}$$

Nous y allons pour la preuve directe. Soit une suite de Cauchy (q_n) dans $\mathbb{Q}$. Nous devons prouver que la suite $n \mapsto g_a(q_n)$ est de Cauchy dans $\mathbb{R}$. Soit $\epsilon > 0$.
La suite (q_n) étant de Cauchy dans $\mathbb{Q}$, elle l'est également dans $\mathbb{R}$, elle est bornée parce que convergente vu que $\mathbb{R}$ est complet [148]. Vu que g_a est croissante [149] et que (q_n) est bornée, il existe M tel que $|g_a(q_n)| \leqslant M$ pour tout n.
Nous considérons δ tel que $0 < q < \delta$ implique $|1 - g_a(q)| \leqslant \epsilon$, ainsi que N tel que $i, j > N$ implique $|q_i - q_j| \leqslant \delta$ (là nous utilisons le fait que (q_n) est de Cauchy). Pour de tels N, i, j nous avons

$$|g_a(q_i) - g_a(q_j)| \leqslant M\epsilon. \tag{12.1092}$$

Donc la suite $g_a(q_n)$ est de Cauchy.

(2) **Pour $a = 1$** La fonction g_a est constante.

(3) **Pour $a = 0$** Dans ce cas $a^q = 0$ pour tout q, et l'image de toute suite de Cauchy est de Cauchy.

(4) **Pour $0 < a < 1$** Considérons une suite de Cauchy (q_i) dans $\mathbb{Q}$. Nous posons $b = 1/a$, et nous utilisons la définition 12.385 ainsi que le lemme 12.386 :

$$a^{q_i} = \frac{1}{b^{q_i}} = b^{-q_i}. \tag{12.1093}$$

La suite $(-q_i)$ est de Cauchy et $b > 1$, donc la suite $i \mapsto b^{-q_i}$ est de Cauchy par le point précédent.

□

12.396.
L'ingrédient magique qui fait fonctionner la proposition 12.395 est le fait que $g_a(x+y) = g_a(x)g_a(y)$ couplé au fait que $\lim_{q\to 0} g_a(q) = 1$. C'est cela qui débloque la situation pour étendre la fonction puissance de $\mathbb{Q}$ vers $\mathbb{R}$ en utilisant le lemme 12.61.

Le chemin suivi par [414] pour étendre la fonction puissance de $\mathbb{Q}$ vers $\mathbb{R}$ est un peu différent : il définit $a^x = \sup\{a^q \text{ tel que } q < x, q \in \mathbb{Q}\}$. La preuve que cette définition donne $x \mapsto a^x$ continue sur $\mathbb{R}$ repose, elle aussi, essentiellement sur le fait que $\lim_{q\to 0} a^q = 1$.

Il y a donc une certaine justice.

148. Théorème 7.287.
149. Proposition 12.392(2).

DEFooOJMKooJgcCtq

Proposition-Définition 12.397 (Fonction puissance[1]).
Si $a > 0$ la fonction

$$\begin{aligned} g_a\colon \mathbb{Q} &\to \mathbb{R} \\ x &\mapsto a^x. \end{aligned} \tag{12.1094}$$

est Cauchy-continue par la proposition 12.395. Si $x \in \mathbb{R}$ nous définissons

$$a^x = \tilde{g}_a(x) \tag{12.1095}$$

où $\tilde{g}_a$ est l'extension de g_a donnée par le lemme 12.61.

Nous allons la noter g_a également, et écrire a^x la valeur de g_a même lorsque x n'est pas un rationnel.

PROPooVADRooLCLOzP

Proposition 12.398 ([1]).
Quelques propriétés de la fonction puissance.

ITEMooQHYRooJIewyp

(1) Pour $a > 0$, la fonction $g_a\colon x \mapsto a^x$ est continue sur $\mathbb{R}$.

ITEMooIZBLooSGtWIp

(2) Pour $a > 1$, la fonction $g_a\colon x \mapsto a^x$ est croissante.

ITEMooSCJBooNVJZah

(3) Pour $a > 0$ et $x, y \in \mathbb{R}$ nous avons

EQooEWIHooDRAQGR

$$a^x a^y = a^{x+y}. \tag{12.1096}$$

En particulier,

$$a^{-x} = \frac{1}{a^x}. \tag{12.1097}$$

Démonstration. La continuité de $x \mapsto a^x$ est par construction. Le point (1) est fait.

Pour le point (2), lorsque $a > 1$, la fonction $f_a\colon \mathbb{Q} \to \mathbb{R}$ est croissante (proposition 12.392). Donc par la proposition 12.63, la fonction $x \mapsto a^x$ est croissante sur $\mathbb{R}$.

Et enfin pour le point (3), il faut faire un peu plus attention. Soient des suites $x_i \to x$ et $y_i \to y$ dans $\mathbb{Q}$. Calculons : SUBEQSooMPNLooPoyjwJ

$$\begin{aligned} a^x a^y &= (\lim_i a^{x_i}) a^y && \text{SUBEQooOCIOooZcswMo (12.1098a)} \\ &= \lim_i \left(a^{x_i} a^y\right) && \text{SUBEQooEKQXooPLqzrG (12.1098b)} \\ &= \lim_i \left(\lim_k a^{x_i} a^{y_k}\right) && \text{SUBEQooZEXDooRytDvS (12.1098c)} \\ &= \lim_i \left(\lim_k a^{x_i + y_k}\right) && \text{SUBEQooSYNBooIQZJzl (12.1098d)} \\ &= \lim_i a^{x_i + y} && \text{SUBEQooKHKCooGwaPDQ (12.1098e)} \\ &= a^{x+y}. && \text{SUBEQooMZBFooSoSgKU (12.1098f)} \end{aligned}$$

Justifications :

— Pour 12.1098a. Définition de a^x lorsque $x \in \mathbb{R}$.

— Pour 12.1098b. Nous entrons le nombre a^y dans la limite. Entrer un facteur dans une limite convergente dans $\mathbb{R}$ est un acte anodin.

— Pour 12.1098c. Définition de a^y, et renter le nombre réel a^{x_i} dans la limite sur k.

— Pour 12.1098d. Utilisation du lemme 12.389, valable pour $x_i, y_k \in \mathbb{Q}$.

— Pour 12.1098e. Pour i fixé, la suite $k \mapsto x_i + x_k$ est une suite de rationnels qui converge vers le réel $x_i + y$. Par définition 12.397 de la fonction puissance nous avons alors $\lim_k a^{x_i+y_k} = a^{x_i+y}$.

— Pour 12.1098f. La suite de réels $i \mapsto x_i + y$ converge dans $\mathbb{R}$ vers le réel $x + y$. Par la continuité de $t \mapsto a^t$ (ça fait partie du lemme 12.61 définissant la fonction puissance sur $\mathbb{R}$) nous avons $\lim_i a^{x_i+y} = a^{x+y}$.

Vous remarquerez que les limites sur k et sur i ne s'enlèvent pas tout à fait avec la même justification. Nous aurions pu invoquer la continuité sur $\mathbb{R}$ de $t \mapsto a^t$ pour les deux limites. Mais cette continuité, dans le cas d'une suite purement constituée de rationnels, est la définition de la prolongation vers $\mathbb{R}$. □

LEMooIPLLooCgpCIn

Lemme 12.399.
Soient $a, b > 0$. Si $1 < x < y$ alors

$$a - b < ay - bx. \tag{12.1099}$$

Démonstration. Nous posons $y = x + s$ avec $s > 0$. Alors

$$ay - bx = a(x + s) - bx = (a - b)x + as > (a - b)x > a - b \tag{12.1100}$$

parce que $as > 0$ et $x > 1$. □

PROPooJXHFooCDwxCS

Proposition 12.400 ([1]).
Pour $q > 0$ dans $\mathbb{Q}$, la fonction

$$\begin{aligned} f_q \colon \mathbb{Q}^+ &\to \mathbb{R} \\ x &\mapsto x^q \end{aligned} \tag{12.1101}$$

est strictement croissante.

Démonstration. Division selon la généralité de q.

(1) **Si q est entier positif** Soit $q = n \in \mathbb{N}$. Si $s > 0$ alors l'inégalité $(x + s)^n > x^n$ découle du binôme de Newton de la proposition 3.43.

(2) **Si q est rationnel** Soient un rationnel $q = m/n$ et un nombre strictement positif s. Nous avons, par la définition 12.381 sous la forme (12.1061) :

$$f_{m/n}(x + s)^n = (x + s)^m > x^m = f_{m/n}(x)^n. \tag{12.1102}$$

Nous avons utilisé la stricte croissance de $x \mapsto x^m$. Cela donne

$$f_{m/n}(x + s)^n > f_{m/n}(x)^n. \tag{12.1103}$$

En utilisant encore la stricte croissance de $x \mapsto x^n$, nous avons le résultat.

□

CORooYWNNooLwKmiD

Corolaire 12.401.
Soient $1 < b < a$ dans $\mathbb{R}$ et des rationnels strictement positifs $p < q$. Alors

$$a^p - b^p < a^q - b^q \tag{12.1104}$$

Démonstration. Nous notons $q = p + r$ avec $r > 0$ dans $\mathbb{Q}$. Par la proposition 12.400,

$$a^r > b^r. \tag{12.1105}$$

Cela nous permet d'utilise le lemme 12.399 pour écrire

$$a^p - b^p < a^p a^r - b^p b^r = a^q - b^q. \tag{12.1106}$$

□

PROPooKWRGooMTbRdU

Proposition 12.402 ([1]).
Soient $a, b > 0$ et $\alpha \in \mathbb{R}$. Nous avons :

$$a^\alpha b^\alpha = (ab)^\alpha. \tag{12.1107}$$

Démonstration. Nous supposons que c'est bon pour $\alpha \in \mathbb{N}$ et $\alpha \in \mathbb{Z}$. Pour les autres, nous donnons plus de détails.

(1) $\mathbb{Q}^+$ Soit $q = m/n$ avec $m, n \in \mathbb{N}$. Si $a^{m/n} = x$ et $b^{m/n} = y$, alors

EQooGNMAooQJMNsL

$$x^n = a^m \tag{12.1108a}$$
$$y^n = b^m \tag{12.1108b}$$

par (12.1061). Nous multiplions (12.1108a) par y^n à gauche et par b^m à droite : $x^n y^n = a^m b^m$. En tenant compte du résultat pour $\mathbb{N}$, nous avons

$$(xy)^n = (ab)^m, \tag{12.1109}$$

ce qui signifie que le nombre xy est $(ab)^{m/n}$.

(2) **Pour $\mathbb{Q}^-$** Soit $q \in \mathbb{Q}^+$, nous avons le calcul

$$a^{-q}b^{-q} = \frac{1}{a^q b^q} = \frac{1}{(ab)^q} = (ab)^{-q}. \tag{12.1110}$$

(3) **Pour $\mathbb{R}$** Soit une suite de rationnels $\alpha_i \to \alpha$. Nous avons

$$a^\alpha b^\alpha = \big(\lim_i a^{\alpha_i}\big)\big(\lim_j b^{\alpha_j}\big) = \lim_i \big(a^{\alpha_i} b^{\alpha_i}\big) = \lim_i (ab)^{\alpha_i} = (ab)^\alpha. \tag{12.1111}$$

Justifications :

— la proposition 9.27 pour le produit des limites,
— le résultat dans $\mathbb{Q}$ que nous venons de prouver,
— la définition de $(ab)^\alpha$ comme limite de $(ab)^{\alpha_i}$.

□

Pour rappel, la proposition suivante, dans le cas de $\alpha \in \mathbb{Q}^+$ est la proposition 12.400.

PROPooUOFKooYyGwIr

Proposition 12.403 ([1]).
Pour $\alpha > 0$, la fonction

$$\begin{aligned} f_\alpha \colon]0, \infty[&\to \mathbb{R} \\ x &\mapsto x^\alpha \end{aligned} \tag{12.1112}$$

est strictement croissante.

Démonstration. Nous rappellons que le cas $\alpha \in \mathbb{Q}^+$ est déjà traité par la proposition 12.400. Soient $x \in]0, \infty[$ et $s > 0$. Nous allons montrer que $f_\alpha(x + s) - f_\alpha(x) > 0$. Pour cela nous décomposons en plusieurs cas.

(1) $x > 1$ Par la proposition 1.455, nous considérons une suite strictement croissante de rationnels strictement positifs $\alpha_i \to \alpha$. Pour tout i nous avons $\alpha_i > \alpha_0$.
En utilisant la stricte croissance de f_{α_0} et le lemme 12.390(2), nous avons les inégalités $1 < x^{\alpha_0} < (x + s)^{\alpha_0}$, et en particulier

$$0 < (x + s)^{\alpha_0} - x^{\alpha_0}. \tag{12.1113}$$

De plus nous avons $1 < x < x + s$ et $\alpha_0 < \alpha_i$ pour tout i. Donc le corolaire 12.401 s'applique et nous avons, pour tout i :

$$0 < (x + s)^{\alpha_0} - x^{\alpha_0} < (x + s)^{\alpha_i} - x^{\alpha_i}. \tag{12.1114}$$

C'est le moment de passer à la limite $i \to \infty$. La seconde inégalité devient non stricte, mais la première reste :

$$0 < (x + s)^{\alpha_0} - x^{\alpha_0} \leqslant (x + s)^\alpha - x^\alpha. \tag{12.1115}$$

Nous avons donc bien la stricte croissance de f_α sur $]1, \infty[$.

(2) $\underline{0 < x \leqslant 1}$ Nous choisissons encore $\alpha_i \to \alpha$ strictement croissante dans $\mathbb{Q}$. Pour chaque i, nous avons encore

$$(x+s)^{\alpha_i} - x^{\alpha_i} > 0. \tag{12.1116}$$

Le passage à la limite change l'inégalité stricte en inégalité large, et ne permet donc pas de conclure immédiatement. Nous devons donc ruser. Soit $k \in \mathbb{N}$ tel que $k(x+s) > 1$ et $kx > 1$ (existence parce que $\mathbb{R}$ est archimédien, proposition 1.440). Nous avons :

$$\big(k(x+s)\big)^{\alpha} - (kx)^{\alpha} > 0 \tag{12.1117}$$

par la partie « $x > 1$ » que nous venons de prouver. Grâce à la proposition 12.402 nous pouvons factoriser k^α :

$$0 < \big(k(x+s)\big)^{\alpha} - (kx)^{\alpha} = k^{\alpha}\big((x+s)^{\alpha} - x^{\alpha}\big). \tag{12.1118}$$

Vu que $k^\alpha > 0$, cela implique $(x+s)^\alpha - x^\alpha > 0$, ce qu'il fallait.

Nous avons fini de prouver que la fonction f_α était strictement croissante sur $]0,\infty[$. □

LEMooJVXQooDPUuuJ

Lemme 12.404.
Si $p \geqslant 1$ et si $x \in [0,1]$ alors $x^p \leqslant x$.

Démonstration. Vu que $p \geqslant 1$, nous avons $p = 1+\alpha$ avec $\alpha \geqslant 0$. Nous pouvons donc écrire[150]

$$x^p = x^{1+\alpha} = xx^\alpha = xf_\alpha(x). \tag{12.1119}$$

Vu que f_α est croissante (proposition 12.403), que $f_\alpha(0) = 1$ et que $f_\alpha(1) = 1$, nous avons $f_\alpha(x) \in [0,1]$ dès lors que $x \in [0,1]$. Donc

$$xf_\alpha(x) \leqslant x. \tag{12.1120}$$

□

Nous prouvons à présent que f_α est localement injective ; nous en avons besoin pour prouver la continuité. Or cette continuité est nécessaire à prouver que f_α est localement bijective. Donc nous ne pouvons pas énoncer la bijectivité ici. Ce sera la proposition 12.410.

PROPooHKTKooCUEBjh

Proposition 12.405.
Soient $\alpha \neq 0$ et $x \in \mathbb{R}\backslash\{0\}$. Il existe un voisinage V de x sur lequel

$$f_\alpha : V \to f_\alpha(V) \tag{12.1121}$$

est injective.

Démonstration. Soit $x > 0$; nous considérons un voisinage V de x inclu à $]0,\infty[$. Soit $y \in V$; pour fixer les idées nous supposons $y < x$. Par la stricte croissance de f_α sur $]0,\infty[$ (proposition 12.403), nous avons $f_\alpha(y) < f_\alpha(x)$ et en particulier $f_\alpha(x) \neq f_\alpha(y)$.

Le cas $x < 0$ se traite de façon analogue, avec la stricte décroissance de f_α sur $]-\infty,0[$. □

Notons que les voisinages sur lesquels f_α est injective sont assez grands. Ils peuvent être toute une demi-droite, si l'on veut.

LEMooQTNKooLVEytN

Lemme 12.406.
Soient $\alpha > 0$, une suite de rationnels strictement décroissante $\alpha_i \to \alpha$ ainsi que les fonctions

$$\begin{aligned} f_{\alpha_i} :]1,\infty[&\to \mathbb{R} \\ x &\mapsto x^{\alpha_i}. \end{aligned} \tag{12.1122}$$

La famille $\{f_{\alpha_i}\}_{i\in\mathbb{N}}$ est équicontinue[151].

150. En utilisant la proposition 12.398(3).
151. Définition 7.317.

Démonstration. Soient $x > 1$, et $\alpha > 0$. Nous allons montrer que $\{f_{\alpha_i}\}$ est équicontinue en x. Soit s tel que $1 < x < x + s$; le corolaire 12.401 nous enseigne que

$$(x+s)^p - x^p < (x+s)^q - x^q \tag{12.1123}$$

dès que $p < q$. En particulier, f_p étant croissante par la proposition 12.403,

$$0 < (x+s)^{\alpha_i} - x^{\alpha_i} < (x+s)^{\alpha_0} - x^{\alpha_0}. \tag{12.1124}$$

Soit $\epsilon > 0$ et δ tel que $s < \delta$ implique $|(x+s)^{\alpha_0} - x^{\alpha_0}| < \epsilon$. Alors nous avons aussi, pour de tels δ et s :

$$|(x+s)^{\alpha_i} - x^{\alpha_i}| < |(x+s)^{\alpha_0} - x^{\alpha_0}| < \epsilon. \tag{12.1125}$$

En procédant de même pour $s < 0$, nous trouvons bien que

$$|y^{\alpha_i} - x^{\alpha_i}| \leqslant \epsilon \tag{12.1126}$$

pour tout $y \in B(x, \delta)$.

Cela signifie que $\{f_i\}$ est équicontinue. □

PROPooUQNZooSSHLqr

Proposition 12.407 ([1])**.**
Soit $\alpha > 0$ dans $\mathbb{R}$. La fonction

$$\begin{aligned} f_\alpha \colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^\alpha \end{aligned} \tag{12.1127}$$

est continue.

Démonstration. Nous allons subdiviser quelque cas.

(1) **Pour $\alpha \in \mathbb{N}$** Nous supposons que ce cas va bien.

(2) **Pour $\alpha \in \mathbb{Q}^+$** Soit $q = m/n$ avec $m, n \in \mathbb{N}$. Soit aussi $\epsilon > 0$. Nous avons :

$$f_{m/n}(x)^n = x^m \tag{12.1128a}$$
$$f_{m/n}(x+\epsilon)^n = (x+\epsilon)^m. \tag{12.1128b}$$

SUBEQooGNCSooWAeRcL

L'équation (12.1128b) s'écrit aussi bien sous la forme

$$f_n\big(f_{m/n}(x+\epsilon)\big) = (x+\epsilon)^m. \tag{12.1129}$$

En prenant la limite,

$$\lim_{\epsilon \to 0} \big[f_n\big(f_{m/n}(x+\epsilon)\big)\big] = x^m = f_{m/n}(x)^n. \tag{12.1130}$$

Vu que f_n est continue, nous pouvons la permuter avec la limite dans le membre de gauche tout en écrivant $f_{m/n}(x)^n = f_n\big(f_{m/n}(x)\big)$ dans le membre de droite :

$$f_n\big[\lim_{\epsilon \to 0} f_{m/n}(x+\epsilon)\big] = f_n\big(f_{m/n}(x)\big). \tag{12.1131}$$

La fonction f_n étant injective dans un voisinage autour de x (proposition 12.405),

$$\lim_{\epsilon \to 0} f_{m/n}(x+\epsilon) = f_{m/n}(x), \tag{12.1132}$$

ce qui est la continuité de $f_{m/n}$ en x.

(3) **Pour $\alpha \in \mathbb{R}^+$**

Nous prouvons séparément le cas $x < 1$ et le cas $x \geqslant 1$.

(3a) **Si $x \in]1, \infty[$**

Commençons par $x \in]1, \infty[$.

Soit une suite $\alpha_i \to \alpha$ strictement décroissante dans $\mathbb{Q}^+$. Le lemme 12.406 nous dit que l'ensemble de fonctions $\{f_{\alpha_i} \colon]1, \infty[\to \mathbb{R}\}_{i \in \mathbb{N}}$ est équicontinu. La convergence simple $f_{\alpha_i} \to f_\alpha$ étant par définition, la proposition 7.320 nous dit que la fonction $f_\alpha \colon]1, \infty[\to \mathbb{R}$ est continue.

(3b) **Si $x \in]0,1]$**
Soient maintenant $x \in]0,1]$, et $\epsilon > 0$. Il existe $k \in \mathbb{N}$ tel que $kx > 1$, et $k^\alpha > 1$. Soit $\delta_1 > 0$ tel que $B(x,\delta_1) > 0$. Si $y \in B(x,\delta_1)$, alors

$$|y^\alpha - x^\alpha| \leqslant k^\alpha |y^\alpha - x^\alpha| = \big|(ky)^\alpha - (kx)^\alpha\big|. \tag{12.1133}$$

Par ailleurs, vu que $kx > 1$, la fonction f_α est continue en kx. Il existe donc δ_2 tel que

$$z \in B(kx,\delta_2) \Rightarrow \big|z^\alpha - (kx)^\alpha\big| < \epsilon. \tag{12.1134}$$

Nous considérons $\delta < \min(\delta_1, \delta_2/k)$. Si $y \in B(x,\delta)$, alors d'une part

$$|ky - kx| = k|y - x| \leqslant k\delta < \delta_2. \tag{12.1135}$$

Pour un tel y nous avons donc $ky \in B(kx,\delta_2)$ et donc

$$|y^\alpha - x^\alpha| \leqslant \big|(ky)^\alpha - (kx)^\alpha\big| < \epsilon. \tag{12.1136}$$

Donc la continuité de f_α en x.

(4) **Pour $\alpha \in \mathbb{R}^-$**
Si $\alpha > 0$, la fonction $f_{-\alpha}$ est donnée par

$$f_{-\alpha}(x) = \frac{1}{f_\alpha(x)} \tag{12.1137}$$

et est donc continue (sauf en $x = 0$ où elle n'existe pas).

□
PROPooDWZKooNwXsdV

Proposition 12.408 ([1]).
Soient $a > 0$ ainsi que $x, y \in \mathbb{R}$. Alors

$$(a^x)^y = (a^y)^x = a^{xy}. \tag{12.1138}$$

Démonstration. Nous découpons en fonction de la nature de x et y.

(1) **x rationnel, y naturel** Si $q \in \mathbb{Q}$ et $n \in \mathbb{N}$ alors la formule

$$(a^q)^n = a^{nq} \tag{12.1139}$$

découle seulement d'une récurrence sur la formule 12.1096.

(2) **$x, y \in \mathbb{Q}$** Soient $y = m/n$ avec $n \in \mathbb{Z}$, $m \in \mathbb{N}$ et $q \in \mathbb{Q}$. Nous avons, en utilisant la partie déjà démontrée et le lemme 12.387,

$$(a^q)^y = (a^q)^{m/n} = \big((a^q)^m\big)^{1/n} = (a^{mq})^{1/n} = a^{mq/n} = a^{yq}. \tag{12.1140}$$

(3) **x, y irrationnels**
Soient des suites des rationnels $x_i \to x$ et $y_i \to y$. En utilisant les définitions,

$$(a^x)^y = \lim_i (a^x)^{y_i} = \lim_i \big(\lim_j a^{x_j}\big)^{y_i}. \tag{12.1141}$$
EQooXITUooHYNSPU

Fixons un i pour commencer. Nous avons, par la continuité de f_{y_i} (proposition 12.407)

$$\big(\lim_j a^{x_j}\big)^{y_i} = f_{y_i}\big(\lim_j a^{x_j}\big) = \lim_j \big(f_{y_i}(a^{x_j})\big) = \lim_j a^{x_j y_i}. \tag{12.1142}$$

Nous avons utilisé le résultat déjà démontré dans le cas des rationnels. La suite $j \mapsto x_j y_i$ est une suite dans $\mathbb{Q}$ qui converge vers le réel xy_i, donc la limite sur j redonne la fonction puissance :

$$\big(\lim_j a^{x_j}\big)^{y_i} = \lim_j a^{x_j y_i} = a^{xy_i}. \tag{12.1143}$$
EQooWOBSooFoRBod

Le résultat découle maintenant de la prise de limite dans (12.1141) qui revient à prendre la limite $i \to \infty$ de l'expression dans (12.1143) :

$$(a^x)^y = \lim_i \big(\lim_j a^{x_j}\big)^{y_i} = \lim_i a^{xy_i} = a^{xy}. \tag{12.1144}$$

□
PROPooJRWCooGiXAYt

Proposition 12.409.
Soit $\alpha > 0$ dans $\mathbb{R}$. Nous avons

$$\lim_{x\to\infty} x^\alpha = \infty. \tag{12.1145}$$

Démonstration. Nous séparons la preuve en fonction de la nature de α.

(1) **Si $\alpha \in \mathbb{N}$** C'est le lemme 12.29.

(2) **Si $\alpha = 1/n$** Soit $n \neq 0$ dans $\mathbb{N}$, et prouvons que $\lim_{x\to\infty} x^{1/n} = \infty$. La proposition 12.398 nous indique que $x \mapsto x^{1/n} = f_n^{-1}(u)$ est croissante et continue. Elle possède donc une limite ℓ éventuellement infinie par la proposition 12.36. Posons

$$\lim_{x\to\infty} f_n^{-1}(x) = \ell. \tag{12.1146}$$

Nous voulons appliquer f_n des deux côtés et profiter de la continuité de f_n pour permuter avec la limite. Si vous avez peur du cas $\ell = +\infty$, supposez $\ell \neq +\infty$ et considérez ce qui suit comme une preuve par l'absurde que $\ell = +\infty$. Nous avons :

$$f_n\big(\lim_{x\to\infty} f_n^{-1}(x)\big) = \lim_{x\to\infty} f_n\big(f_n^{-1}(x)\big) = \lim_{x\to\infty} x = \infty. \tag{12.1147}$$

donc $f_n(\ell) = \infty$, et nous concluons que $\ell = \infty$.

(3) **Si $\alpha \in \mathbb{Q}$** Nous posons $\alpha = p/q$ avec $p, q \in \mathbb{N}$. Alors

$$x^\alpha = (x^{1/q})^p \tag{12.1148}$$

par la proposition 12.408. Autrement dit,

$$\lim_{x\to\infty} x^\alpha = \lim_{x\to\infty} (f_p \circ f_{1/q})(x) = \infty \tag{12.1149}$$

parce que tant f_p que $f_{1/q}$ ont une limite $+\infty$.

(4) **Le cas général** Nous considérons enfin $\alpha > 0$ dans $\mathbb{R}$. Le lemme 1.441 nous permet de considérer $q \in \mathbb{Q}$ tel que $0 < q < \alpha$. La proposition 12.398(2) nous dit que, pour chaque x, $x^q < x^\alpha$. Donc

$$\lim_{x\to\infty} x^\alpha \geqslant \lim_{x\to\infty} x^q = \infty \tag{12.1150}$$

en utilisant le point précédent.

□
PROPooEXGKooCqzLor

Proposition 12.410.
Soit $\alpha > 0$. La fonction

$$\begin{aligned} f_\alpha\colon [0,\infty[&\to [0,\infty[\\ x &\mapsto x^\alpha \end{aligned} \tag{12.1151}$$

est bijective.

Démonstration. La proposition 12.403 nous dit que $f_\alpha\colon]0,\infty[\to \mathbb{R}$ est strictement croissante. Vu que $f_\alpha(0) = 0$, nous savons que $f_\alpha([0,\infty[) \subset [0,\infty[$. La stricte croissance nous dit également que f_α est injective.

Il reste à voir que f_α est surjective. Rappelons quelque faits.

— D'abord une facile : $f_\alpha(0) = 0$.

— Nous avons $\lim_{x\to\infty} x^\alpha = \infty$ par la proposition 12.409.

— La fonction f_α est continue par la proposition 12.407.

Le théorème des valeurs intermédiaires [152] conclut que f_α est surjective sur $[0,\infty[$. □

152. Voir 9.94.

Le lemme suivant montre en gros que x^y croît plus rapidement en y qu'en x.

LemLJOSooEiNtTs

Lemme 12.411.
Pour tout $\alpha > 0$ et $a < 1$ nous avons la limite

$$\lim_{n\to\infty} n^\alpha a^n = 0 \tag{12.1152}$$

Démonstration. Soit $k \in \mathbb{N}$ plus grand que α. Soit la suite numérique $s_n = n^k a^n$. Tous ses termes sont positifs et

$$\frac{s_n}{s_{n+1}} = \left(\frac{n}{n+1}\right)^k \frac{1}{a}. \tag{12.1153}$$

Étant donné que $n/n+1 \to 1$ et que $a < 1$, il existe un certain rang à partir duquel la suite (s_n) est décroissante. Deux conclusions :

— Elle est majorée par une constante M.

— Elle est convergente par le lemme 9.35.

Soit l tel que $ka^l < 1$ et $n > l$ alors

$$s_{n+l} = (n+l)^k a^{n+l} \leqslant kn^k a^n a^l = ka^l s_n \leqslant ka^l M. \tag{12.1154}$$

La majoration est due au fait que dans $(n+l)^k$ nous avons k termes tous plus petits que n^k. De la même façon,

$$s_{2n+2l} \leqslant ka^{2l} s_{2n} \leqslant ka^{2l} M. \tag{12.1155}$$

En posant $\varphi(i) = in + il$ nous avons

$$s_{\varphi(i)} \leqslant ka^i M, \tag{12.1156}$$

qui est une sous-suite convergente vers 0. Or si une suite est convergente (ce qui est le cas de (s_n)), toutes les sous-suites convergent vers la même limite. Nous en concluons que $s_n \to 0$. □

12.412.
Une conséquence est que si vous voulez choisir un mot de passe fort, la longueur du mot est plus importante que la taille de l'alphabet choisit : il est plus efficace de choisir une combinaison longue qu'une combinaisons mélangeant des lettres, chiffres et symboles spéciaux.

Exemple : si vous choisissez un mot de passe contenant majuscules, minuscules, chiffres et symboles spéciaux complètement mélangés (ne mentez pas, vous ne le faites pas), mais que vous ne le choisissez que de taille 6, vous avez 72^6 possibilités (en supposant un jeu de 10 symboles spéciaux).

Eh bien, en seulement 8 lettres minuscules, vous avez plus de possibilités : $26^8 > 72^6$.

De nombreux sites font l'erreur de considérer que

— « ggzxzheaiynshunxuydajkwyohgqxz » est un mot de passe faible,

— « azerty.2019A » est un mot de passe fort.

Il n'en est rien. Le premier est considérablement meilleur que le second, même si le second, très superficiellement, mélange les lettres majuscules, minuscules, chiffres et symboles spéciaux.

Voilà voilà. La prochaine fois qu'un site vous refusera un mot de passe de 30 lettres minuscules mélangées, vous saurez pourquoi il n'y a rien qui marche en informatique, et en particulier pourquoi la sécurité générale de nos systèmes d'information est désastreuse.

THOooHWNWooTewPvt

Théorème 12.413.
Soit une matrice $A \in \mathbb{M}(N, \mathbb{C})$. La suite $(A^n x)$ tend vers zéro pour tout x si et seulement si $\rho(A) < 1$ où $\rho(A)$ est le rayon spectral de A

Démonstration. Dans le sens direct, il suffit de prendre comme x, un vecteur propre de A. Dans ce cas nous avons $A^n x = \lambda^n x$. Mais $\lambda^k x$ ne tend vers zéro que si $\lambda < 1$. Donc toutes les valeurs propres de A doivent être plus petites que 1 et $\rho(A) < 1$.

Pour l'autre sens nous utilisons la décomposition de Dunford (théorème 8.265) : il existe une matrice inversible P telle que

$$A = P^{-1}(D + N)P \tag{12.1157}$$

où D est diagonale, N est nilpotente et $[D, N] = 0$. Étant donné que $D + N$ est triangulaire, son polynôme caractéristique est

$$\chi_{D+N}(X) = \prod_i (D_{ii} - X). \tag{12.1158}$$

Par similitude, c'est le même polynôme caractéristique que celui de A et nous savons alors que la diagonale de D contient les valeurs propres de A.

Vu que $A^n = P^{-1}(D + N)^n P$, nous allons montrer que $\|(D + N)^n\| \to 0$, et ce sera suffisant. Notons r l'ordre de nilpotence de N (c'est à dire $N^r = 0$), et prenons $n > r$. En utilisant le fait que D et N commutent, pour tout $n \geqslant r$ nous avons :

$$\|(D + N)^n\| \leqslant \| \sum_{k=0}^{n} \binom{n}{k} D^{n-k} N^k \| \tag{12.1159a}$$

$$\leqslant \sum_{k=0}^{n} \binom{n}{k} \|D\|^{n-k} \|N^k\| \tag{12.1159b}$$

$$= \sum_{k=0}^{r-1} \binom{n}{k} \rho(D)^{n-k} \|N\|^k \tag{12.1159c}$$

$$\leqslant c \sum_{k=0}^{r-1} \binom{n}{k} \rho(D)^n \tag{12.1159d}$$

$$= c\rho(D)^n \sum_{k=0}^{r-1} \binom{n}{k} \tag{12.1159e}$$

$$\leqslant c\rho(D)^n \sum_{k=0}^{r-1} \frac{n^{k-1}}{k!} \tag{12.1159f}$$

$$\leqslant c\rho(D)^n \sum_{k=0}^{r-1} n^{r-2} \tag{12.1159g}$$

$$= cr\rho(D)^n n^{r-2}. \tag{12.1159h}$$

Justifications.

— Pour (12.1159d). Nous avons posé $c = \max_{k=1,\dots,r-1} \|N\|^k \rho(D)^{-k}$.

— Pour (12.1159f). Lemme 3.44.

— Pour (12.1159g). On oublie le $k!$ et on remplace k par $r - 1$.

Récapitulons ces inéquations :

$$\|(D + N)^n\| \leqslant c'\rho(D)^n n^{r-2} \tag{12.1160}$$

où c' est une nouvelle constante. Du coup si $\rho(D) < 1$ alors $\|(D+N)^k\| \to 0$ par lemme 12.411. □

12.35 Densité des polynômes

12.35.1 Théorème de Stone-Weierstrass

Voir le thème 34.

Note : le lemme 12.414 est utilisé dans la démonstration du théorème 12.418 ; c'est pour cela que nous l'avons isolé.

LemYdYLXb

Lemme 12.414.
Il existe une suite de polynômes sur $[0,1]$ convergeant uniformément vers la fonction racine carrée.

Démonstration. Nous donnons cette suite par récurrence :

$$P_0(t) = 0 \tag{12.1161a}$$
$$P_{n+1}(t) = P_n(t) + \frac{1}{2}\big(t - P_n(t)^2\big). \tag{12.1161b}$$

Nous commençons par montrer que pour tout $t \in [0,1]$, $P_n(t) \in [0, \sqrt{t}]$. Pour P_0, c'est évident. Ensuite nous avons

$$P_{n+1}(t) - \sqrt{t} = P_n(t) - \sqrt{t} + \frac{1}{2}(t - P_n(t)^2) \tag{12.1162a}$$
$$= \big(P_n(t) - \sqrt{t}\big)\left(1 - \frac{1}{2}\frac{t - P_n(t)^2}{P_n(t) - \sqrt{t}}\right) \tag{12.1162b}$$
$$= \big(P_n(t) - \sqrt{t}\big)\left(1 - \frac{\sqrt{t} + P_n(t)}{2}\right) \tag{12.1162c}$$
$$\leqslant 0 \tag{12.1162d}$$

parce que $\sqrt{t} \leqslant 1$ et $P_n(t) \leqslant 1$ par hypothèse de récurrence.

Nous savons au passage que $P_n(t)$ est une suite réelle croissante parce que $t - P_n(t)^2 \geqslant t - (\sqrt{t})^2 = 0$. La suite $P_n(t)$ est donc croissante et majorée par $\sqrt{t}$; elle converge donc. Les candidats limites sont déterminés par l'équation

$$\ell = \ell + \frac{1}{2}(t - \ell^2), \tag{12.1163}$$

dont les solutions sont $\ell = \pm\sqrt{t}$. La suite étant positive, nous avons une convergence ponctuelle de P_n vers la racine carrée. Cette suite étant une suite croissante de fonctions continues sur un compact, convergeant ponctuellement vers une fonction continue, la convergence est uniforme par le théorème de Dini 12.363. □

LemUuxcqY

Lemme 12.415.
Soit K, un compact de $\mathbb{R}$ et f_n une suite de fonctions sur K convergeant uniformément vers f. Soient un espace topologique X et une application $g\colon X \to K$. Alors $f_n \circ g$ converge uniformément vers $f \circ g$.

Démonstration. En effet, pour tout $x \in X$ nous avons

$$\|(f_n \circ g) - (f \circ g)\|_\infty = \sup_{x\in X} \|f_n\big(g(x)\big) - f\big(g(x)\big)\| \leqslant \|f_n - f\|_\infty. \tag{12.1164}$$

Par conséquent, si $\epsilon > 0$ est donné, il suffit de choisir n de telle sorte à avoir $\|f_n - f\|_\infty < \epsilon$ et nous avons $\|(f_n \circ g) - (f \circ g)\|_\infty \leqslant \epsilon$. □

Définition 12.416.
*Nous disons qu'une algèbre A de fonctions sur un espace X **sépare les points** de X si pour tout $x_1 \neq x_2$ il existe $g \in A$ telle que $g(x_1) \neq g(x_2)$.*

Nous pouvons maintenant énoncer et démontrer une forme nettement plus générale du théorème de Stone-Weierstrass. Le théorème 12.418 le donne pour $C(X, \mathbb{C})$ et le théorème 12.417 le donne pour $C(X, \mathbb{R})$.

THOooMDILooGPXbTW

Théorème 12.417 (Stone-Weierstrass[415])**.**
Soient X, un espace compact et Hausdorff. Soit A, une sous-algèbre de $C(X, \mathbb{R})$ contenant une fonction constante non nulle. Alors A est dense dans $\Big(C(X, \mathbb{R}), \|.\|_\infty\Big)$ si et seulement si A sépare les points de X.

Démonstration. Nous allons écrire la démonstration en plusieurs étapes (dont la première est le lemme 12.414). Nous commençons par la première partie, sur les réels.

Première étape Pour tout $x \neq y \in X$ et pour tout $\alpha, \beta \in \mathbb{R}$, il existe une fonction $f \in A$ telle que $f(x) = \alpha$ et $f(y) = \beta$.

En effet, vu que A sépare les points nous pouvons considérer une fonction $g \in A$ telle que $g(x) \neq g(y)$ et ensuite poser

$$f(z) = \alpha - \frac{\alpha - \beta}{g(y) - g(x)}\big(g(z) - g(x)\big). \tag{12.1165}$$

Les constantes faisant partie de A, cette fonction f est encore dans A.

Seconde étape Pour tout n-uples de fonctions $f_1, \dots, f_n$ dans $\bar{A}$, les fonctions $\min(f_1, \dots, f_n)$ et $\max(f_1, \dots, f_n)$ sont dans $\bar{A}$.

Nous le démontrons pour $n = 2$; le reste allant évidemment par récurrence. Soient $f, g \in \bar{A}$. Étant donné que

$$\max(f, g) = \frac{f + g}{2} + \frac{|f - g|}{2} \tag{12.1166a}$$

$$\min(f, g) = \frac{f + g}{2} - \frac{|f - g|}{2}, \tag{12.1166b}$$

if suffit de montrer que si $f \in \bar{A}$ alors $|f| \in \bar{A}$. Si f est nulle, c'est évident ; supposons que $f \neq 0$ et posons $M = \|f\|_\infty \neq 0$. Pour tout $x \in X$ nous avons

$$\frac{f(x)^2}{M^2} \in [0, 1]. \tag{12.1167}$$

Nous considérons alors la suite

$$h_n = P_n \circ \frac{f^2}{M^2} \tag{12.1168}$$

où P_n est une suite de polynômes convergent uniformément vers la racine carrée (voir lemme 12.414). Le lemme 12.415 nous assure que h_n converge uniformément vers $\frac{|f|}{M}$ dans $C(X, \mathbb{R})$. Étant donné que $\bar{A}$ est également une algèbre, h_n est dans $\bar{A}$ pour tout n et la limite s'y trouve également (pour rappel, la fermeture $\bar{A}$ est celle de la topologie de la convergence uniforme).

Troisième étape Soit $\epsilon > 0$, $f \in C(X, \mathbb{R})$ et $x \in X$. Il existe une fonction $g_x \in \bar{A}$ telle que

$$\begin{cases} g_x(x) = f(x) & \text{(12.1169a)} \\ g_x(y) \leqslant f(y) + \epsilon & \text{(12.1169b)} \end{cases}$$

pour tout $y \in X$.

Soit $z \in X \backslash \{x\}$ et une fonction h_z telle que $h_z(x) = f(x)$ et $h_z(z) = f(z)$. Une telle fonction existe par une des étapes précédentes. Étant donné que f et h_z sont continues, il existe un voisinage ouvert V_z de z sur lequel

$$h_z(y) \leqslant f(y) + \epsilon \tag{12.1170}$$

pour tout $y \in V_z$. Nous pouvons sélectionner un nombre fini de points $z_1, \dots, z_n$ tels que les ouverts $V_{z_1}, \dots, V_{z_n}$ recouvrent X (parce que X est compact, de tout recouvrement par des ouverts, nous extrayons un sous recouvrement fini.). Nous posons

$$g_x = \min(h_{z_1}, \dots, h_{z_n}) \in \bar{A}. \tag{12.1171}$$

Si $y \in X$, nous sélectionnons le i tel que $h_{z_i}(y) \leqslant f(y) + \epsilon$ et nous avons

$$g_x(y) \leqslant h_{z_i}(y) \leqslant f(y) + \epsilon. \tag{12.1172}$$

Étape finale Soit $\epsilon > 0$ et $f \in C(X, \mathbb{R})$. Pour chaque $x \in X$ nous considérons une fonction $g_x \in \bar{A}$ telle que

$$\begin{cases} g_x(x) = f(x) & \text{(12.1173a)} \\ g_x(y) \leqslant f(y) + \epsilon & \text{(12.1173b)} \end{cases}$$

pour tout $y \in X$. Les fonctions f et g_x sont continues, donc il existe un voisinage ouvert W_x de x sur lequel

$$g_x(y) \geqslant f(y) - \epsilon. \tag{12.1174}$$

De ces W_x nous extrayons un sous recouvrement fini de X : $W_{x_1}, \dots, W_{x_m}$ et nous posons

$$\varphi = \max(g_{x_1}, \dots, g_{x_n}) \in \bar{A}. \tag{12.1175}$$

Si $y \in X$, il existe un i tel que

$$\varphi(y) \geqslant g_{x_i}(y) \geqslant f(y) - \epsilon. \tag{12.1176}$$

La première inégalité est le fait que φ est le maximum des g_{x_k}, et la seconde est le choix de i. Donc pour tout $y \in X$ nous avons

$$f(y) - \epsilon \leqslant \varphi(y) \leqslant f(y) + \epsilon. \tag{12.1177}$$ EqJMxHaF

La première inégalité est ce que l'on vient de faire. La seconde est le fait que pour tout i nous ayons $g_{x_i}(y) \leqslant f(y) + \epsilon$; le fait que φ soit le maximum sur les i ne change pas l'inégalité.

Le fait que les inégalités (12.1177) soient vraies pour tout $y \in X$ signifie que $\|\varphi - f\|_\infty \leqslant \epsilon$, et donc que $f \in \mathrm{Adh}\big(\mathrm{Adh}(A)\big) = \mathrm{Adh}(A)$.

Tout cela prouve que $C(X, \mathbb{R}) \subset \mathrm{Adh}(A)$. L'inclusion inverse est le fait que $C(X, \mathbb{R})$ est fermé pour la norme $\|.\|_\infty$, étant donné qu'une limite uniforme de fonctions continues est continue.

Nous pouvons maintenant nous tourner vers l'énoncé concernant $C(X, \mathbb{C})$. □ ThoWmAzSMF

Théorème 12.418 (Stone-Weierstrass[1])**.**
Soit X, un espace compact et Hausdorff. Soit une sous-algèbre[153] *A stable par conjugaison*[154] *de $C(X, \mathbb{C})$ contenant une fonction constante non nulle. Alors A est dense dans $\Big(C(X, \mathbb{C}), \|.\|_\infty\Big)$ si et seulement si A sépare les points de X.*

Entendons-nous bien : ici A et $C(X, \mathbb{C})$ sont des algèbres à coefficients dans $\mathbb{C}$.

Démonstration. La preuve de cette version dans $C(X, \mathbb{C})$ va bien entendu fortement reposer sur le cas dans $C(X, \mathbb{R})$ que nous venons de prouver. Soit donc A, une sous-algèbre vérifiant les hypothèses.

(1) $\mathrm{Re}(A) \subset A$ Nous prouvons que si $f \in A$, alors $\mathrm{Re}(f) \in A$. En effet, vu que A est stable par conjugaison, si $f \in A$, alors $\bar{f} \in A$ et $f + \bar{f} = 2\,\mathrm{Re}(f) \in A$.

Nous posons

$$A_1 = \{\mathrm{Re}(g) \text{ tel que } g \in A\}. \tag{12.1178}$$

(1) A_1 **est une sous-algèbre de** A Le fait que les élément de A_1 soient dans A est déjà fait. Pour le produit, si $g_1, h_1 \in A_1$, alors il existe $g, h \in A$ tels que $g_1 = \mathrm{Re}(g)$ et $h_1 = \mathrm{Re}(h)$. Nous avons

$$(g_1 + ig_2)(h_1 + ih_2) = g_1h_1 - g_2h_2 + i(g_1h_2 + g_2h_1) \in A. \tag{12.1179}$$

La partie réelle de cela est dans A_1, donc

$$g_1h_2 - g_2h_1 \in A_1. \tag{12.1180}$$ EQooYAGUooJVnaEa

153. Algèbre, définition 1.347.
154. Pour tout $g \in A$, nous avons $\bar{g} \in A$.

Mais comme $g_1 + ig_2 \in A$, nous avons aussi $g_1 - ig_2 \in A$ et donc

$$(g_1 - ig_2)(h_1 + ih_2) = g_1h_1 + g_2h_2 + i(g_1h_2 - g_2h_1) \in A. \tag{12.1181}$$

La partie réelle de cela est dans A_1. Donc

$$g_1h_1 + g_2h_2 \in A_1. \tag{12.1182}$$

En comparant avec (12.1180), nous avons $g_1h_1 \in A_1$.

(2) A_1 **sépare les points de** X Soient $x, y \in X$ ainsi que $f \in A$ séparant les points x et y, c'est-à-dire

$$f(x) \neq f(y). \tag{12.1183}$$

Supposons $f_1(x) = f_1(y)$. Vu que f sépare, si ce ne sont pas les parties réelles, ce sont les parties imaginaires. C'est-à-dire que $f_2(x) \neq f_2(y)$. Mais d'autre part, $if = -f_2 + if_1 \in A$, donc en réalité $f_2 \in A_1$ également.

Le partie A_1 dans $C(X, \mathbb{R})$ vérifie les hypothèses de Stone-Weierstrass réel 12.417, donc A_1 est dense dans $C(X, \mathbb{R})$. Le même raisonnement montre que A_2 est également dense dans $C(X, \mathbb{R})$ [155]

Soit maintenant le vif de la preuve : $f \in C(X, \mathbb{C})$ avec $f = u + iv$, les fonctions u et v étant dans $C(X, \mathbb{R})$. Nous avons des suites $u_k \xrightarrow{unif} u$ et $v_k \xrightarrow{unif} v$ pour des suites (u_k) et (v_k) dans $C(X, \mathbb{R})$.

Par le même genre de raisonnements que nous avons déjà fait, nous nous convainquons que $u_k + iv_k \in A$ pour chaque k. Nous avons

$$\|u_k + iv_k - u - iv\|_\infty \leqslant \|u_k - u\|_\infty + \|v_k - v\|_\infty \tag{12.1184}$$

En prenant k assez grand, les deux termes peuvent être rendus plus petit que ϵ. □

CORooNIUJooLDrPSv

Corolaire 12.419 ([1]).
Soit B, la boule fermée de centre 0 et de rayon 1 dans $\mathbb{R}^n$. La partie $C^\infty(B, \mathbb{R}^n)$ est dense dans $\big(C(B, B), \|.\|_\infty\big)$.

Démonstration. Soit $f \in C(B, B)$ et $\epsilon > 0$. La fonction donnant la composante i est une fonction $f_i \in C(B, \mathbb{R})$ et il existe donc, par le théorème de Stone-Weierstrass 12.418, une fonction $g_i \in C^\infty(B, \mathbb{R})$ telle que $\|g_i - f_i\|_\infty \leqslant \epsilon$.

La fonction g dont les composantes sont les g_i ainsi construits vérifie $\|g - f\|_\infty \leqslant n\epsilon$. □

Attention toutefois que rien n'assure que les fonctions construites par le corolaire 12.419 prennent leurs valeurs dans B.

Le théorème suivant est un des énoncés les plus classiques de Stone-Weierstrass. Il découle évidemment du théorème général 12.418 (encore qu'il faut alors bien comprendre qu'il faut traiter la fonction $x \mapsto \sqrt{x}$ séparément). Il en existe cependant une preuve indépendante via les polynômes de Bernstein, dans le théorème 36.185. Par contre, n'allez pas croire que c'est plus simple.

ThoGddfas

Théorème 12.420 (Stone-Weierstrass).
Soit f, une fonction continue sur un compact $K \subset \mathbb{R}$ à valeurs dans $\mathbb{R}$. Alors pour tout $\epsilon > 0$, il existe un polynôme P tel que $\|P - f\|_\infty < \epsilon$.

Autrement dit, les polynômes sont denses dans $\big(C^0(K, \mathbb{R}), \|.\|_\infty\big)$.

Démonstration. Nous allons prouver que les polynômes sur $\mathbb{R}$, restreins à K satisfont les hypothèses du théorème de Stone-Weierstrass 12.417.

(1) **Partie de** $C^0(K, \mathbb{R})$ Les polynômes sont des fonctions continues par la proposition 12.407.

(2) **Sous-algèbre** Les produits et sommes de polynômes restent des polynômes.

155. Il me semble même que $A_1 = A_2$ et qu'il y a un raccourci possible dans cette preuve en exploitant ce fait. Écrivez-moi pour dire ce que vous en pensez.

(3) **Contient une fonction constante non nulle** Les fonctions constantes sont des polynômes.

(4) **Sépare les points de K** Le polynôme $P(x) = x$ sépare tous les points que vous voulez.

□

THOooJGAJooBcTtDH

Théorème 12.421 (Stone-Weierstrass)**.**
Soit $f \in C^0(K, \mathbb{R}^n)$ où K est un compact $K \subset \mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Alors pour tout $\epsilon > 0$, il existe un polynôme $P\colon \mathbb{R}^n \to \mathbb{R}^n$ tel que $\|P - f\|_K < \epsilon$.

Autrement dit, les polynômes sont denses dans $\big(C^0(K, \mathbb{R}^n), \|.\|_K\big)$.

12.36 Primitive de fonction continue

PropQACVooBnHtRJ

Proposition 12.422 ([416])**.**
Soit un intervalle compact K de $\mathbb{R}$ et une suite (f_n) de fonctions continues sur K telles que $f_n \xrightarrow{unif} f$. Si chacune des fonctions f_n a une primitive sur K alors f également.

Démonstration. Soit $x_0 \in K$ et les primitives F_n choisies[156] pour avoir $F_n' = f_n$ et $F_n(x_0) = 0$. Nous allons voir que (F_n) est une suite de Cauchy dans $\big(K, \|.\|_\infty\big)$. Soient $n, m \in \mathbb{N}$ et $x \in K$. Nous avons

$$\begin{aligned}
\|F_n - F_m\|_\infty &= \sup_{x\in K} \|F_n(x) - F_m(x)\| && (12.1185a)\\
&= \sup_{x\in K} \|(F_n - F_m)(x)\| && (12.1185b)\\
&\leqslant \sup_{x\in K} \Big(\|F_n' - F_m'\|_{[x,x_0]}\|x - x_0\|\Big) && (12.1185c)\\
&\leqslant \sup_{x\in K} \Big(\|f_n - f_m\| \operatorname{diam}(K)\Big) && (12.1185d)\\
&\leqslant \operatorname{diam}(K)\|f_n - f_m\|_K. && (12.1185e)
\end{aligned}$$

SUBEQooDABTooNgvVaG

SUBEQooKTTOooMJVsEZ

Justifications.

— Pour (12.1185c). Théorème des accroissements finis 10.292.

— Pour (12.1185d). Parce que $x \in K$ et que K est borné, $\|x - x_0\|$ est majoré par $\operatorname{diam}(K)$.

Vu que (f_n) est de Cauchy, si n et m sont assez grands, cela tend vers zéro. La suite (F_n) converge donc vers une certaine fonction F.

Le théorème 12.374 nous permet de permuter la limite et la dérivée pour conclure que $F' = f$ et donc que f a une primitive sur K. □

PropKKGAooDQYGKg

Proposition 12.423 ([416])**.**
Soit un intervalle ouvert I de $\mathbb{R}$ et une fonction $f\colon I \to \mathbb{R}$ qui admet une primitive sur tout compact de I. Alors f a une primitive sur I.

Démonstration. Nous considérons une suite exhaustive[157] de compacts K_n pour I et $x_0 \in K_0$. Nous considérons aussi F_n la primitive de f sur K_n telle que $F_n(x_0) = 0$ (possible parce que $x_0 \in K_n$ pour tout n). Les fonctions F_n sont des restrictions les unes des autres, et nous pouvons définir

$$\begin{aligned} F\colon I &\to \mathbb{R}\\ x &\mapsto F_n(x) \text{ si } x \in K_n. \end{aligned} \tag{12.1186}$$

Nous avons évidemment $F(x_0) = 0$ et nous allons prouver que F est une primitive de f sur I. Soit $x \in I$ vu que I est ouvert, nous pouvons choisir n_0 tel que $x \in \operatorname{Int}(K_{n_0})$. Les fonctions F et

156. Les fonctions F_n étant dérivables sont continues.
157. Voir le lemme 7.306.

F_{n_0} sont égales sur K_n et donc sur un ouvert autour de x. Par conséquent F est dérivable en x et $F'(x) = F'_{n_0}(x) = f(x)$. □

ThoEXXyooCLwgQg

Théorème 12.424.
Soit I un intervalle ouvert de $\mathbb{R}$. Une fonction continue sur I admet une primitive[158] *sur I.*

Démonstration. Sur chaque compact de I, la fonction f est limite uniforme de polynômes[159] (théorème de Stone-Weierstrass 12.420). Donc f est primitivable sur tout compact de I (proposition 12.422) et donc sur I par la proposition 12.423. □

12.36.1 Dérivation de la fonction puissance (première)

Nous n'allons pas complètement résoudre la question de la dérivation de la fonction $x \mapsto a^x$; il faudrait des logarithmes, et nous ne les avons pas encore définis. Le logarithme sera introduit comme fonction inverse de l'exponentielle en 15.83.

PROPooMXCDooBffXbl

Proposition 12.425 ([414]).
Soit la fonction puissance

$$\begin{aligned} g_a : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto a^x. \end{aligned} \tag{12.1187}$$

(1) La fonction g_a est dérivable.

(2) La dérivée vérifie l'équation

$$g'_a(x) = g'_a(0) g_a(x). \tag{12.1188}$$

EQooNIUJooPqDnax

Démonstration. La fonction g_a est continue par 12.398(1). La proposition 12.424 nous dit donc que la fonction g_a admet une primitive sur $\mathbb{R}$. Nous notons F une telle primitive.

Soit $x \in \mathbb{R}$. En posant $F_x(t) = F(x + t)$, nous avons une primitive de $t \mapsto a^x a^t$. En effet

$$F'_x(t) = F'(x + t) \frac{d}{dt}\Big[x + t\Big]_{t=0} = a^{x+t} = a^x a^t. \tag{12.1189}$$

Par ailleurs la fonction $t \mapsto a^x F(t)$ est également une primitive de $t \mapsto a^x a^t$. Donc il existe un nombre $C(x)$ tel que

$$F_x(t) = F(x + t) = a^x F(t) + C(x). \tag{12.1190}$$

Le nombre $F(1) - F(0)$ est un nombre sans histoires. Nous avons :

SUBALIGNooVARJooIcPEHN

$$\begin{aligned} g_a(x)\big(F(1) - F(0)\big) &= g_a(x)F(1) - g_a(x)F(0) && \text{(12.1191a)} \\ &= F_x(1) - C(x) - F_x(0) + C(x) && \text{(12.1191b)} \\ &= F_x(1) - F_x(0) && \text{(12.1191c)} \\ &= F(1 + x) - F(x). && \text{(12.1191d)} \end{aligned}$$

La fonction F étant dérivable, nous en déduisons que g_a est dérivable.

Vu que nous n'avons aucune idée de la forme de F, nous ne pouvons pas tirer beaucoup d'informations d'une dérivation des membres de gauche et de droite de (12.1191).

En ce qui concerne la formule, nous écrivons la fameuse équation fonctionnelle[160]

$$g_a(x + y) = g_a(x) g_a(y) \tag{12.1192}$$

Nous fixons x et dérivons par rapport à y en $y = y_0$:

$$g'_a(x + y_0) = g_a(x) g'_a(y_0). \tag{12.1193}$$

En posant $y_0 = 0$ nous trouvons le résultat demandé. □

158. Définition 12.198.

159. Si tu veux te passer de Stone-Weierstrass, tu peux prouver que toute fonction continue sur un compact est limite uniforme de fonctions affines par morceaux, par exemple. Voir [416].

160. Pour rappel, proposition 12.398(3).

12.426.
La démonstration donnée dans [414] s'assure d'abord de l'existence d'une intégrale (lemme 14.256), pose ensuite $A = \int_0^1 g_a(t)dt$ et fait le calcul suivant :

$$Ag_a(x) = \int_0^1 g_a(x)g_a(t)dt = \int_0^1 g_a(x+t)dt = \int_x^{x+1} g_a(t)dt. \tag{12.1194}$$

Vu que le membre de droite est une fonction dérivable de x, nous concluons que g_a est dérivable. Cela demande donc toute la théorie de l'intégration pour prouver la *dérivabilité* d'une fonction.

La démonstration donnée ici est à peine mieux. Elle utilise l'existence d'une primitive et donc tout le théorème de Stone-Weierstrass 12.420.

Dans les deux cas, je trouve que la situation n'est pas fameuse. Si vous êtes capable de montrer l'existence de la limite

$$\lim_{\epsilon\to 0} \frac{a^\epsilon - 1}{\epsilon} \tag{12.1195}$$

sans recourir à autre chose que des astuces sur les limites, je suis preneur. Ou, au contraire, si vous avez un argument pour dire que c'est impossible, dites-le moi également. Écrivez-moi.

Nous posons une définition

DEFooPJKMooOfZzgy

Définition 12.427.
*Soit $a > 0$. Nous nommons l'**équation fonctionnelle** l'équation* EQooULHBooByFVec

$$\begin{cases} f(x+y) = f(x)f(y) & \text{(12.1196a)} \\ f(1) = a & \text{(12.1196b)} \end{cases}$$

pour la fonction inconnue $f\colon \mathbb{R} \to \mathbb{R}$.

Note : une équation du même type, avec $f(x+y) = f(x) + f(y)$ sera dans le lemme 17.139.

DEFooXMQTooSbZzqJ

Définition 12.428.
*Soit $a > 0$. Nous nommons l'**équation exponentielle** l'équation* EQooGDBYooUoAFPW

$$\begin{cases} y' = y & \text{(12.1197a)} \\ y(1) = a & \text{(12.1197b)} \end{cases}$$

pour la fonction inconnue $y\colon \mathbb{R} \to \mathbb{R}$.

LEMooWUQBooXzjmZv

Lemme 12.429 ([1])**.**
Si elle existe, la solution au problème

$$\begin{cases} y' = y & \text{(12.1198a)} \\ y(0) = 1 & \text{(12.1198b)} \end{cases}$$

vérifie $y(x) = 1/y(-x)$.

En particulier une telle solution ne s'annule pas sur $\mathbb{R}$.

Démonstration. Supposons que y soit une solution, et posons $h(x) = y(x)y(-x)$. Une dérivation montre que $h'(x) = 0$. Vu que $h(0) = 1$, nous avons $h(x) = 1$ pour tout x. Par conséquent $y(x)y(-x) = 1$ pour tout x. □

PropDJQSooYIwwhy

Proposition 12.430 (Unicité de l'exponentielle)**.**
Si elle existe, la solution au problème

$$\begin{cases} y' = y & \text{(12.1199a)} \\ y(0) = 1 & \text{(12.1199b)} \end{cases}$$

est unique.

Démonstration. Soient y et g deux solutions et considérions la fonction $h(x) = g(x)y(-x)$. Un calcul immédiat donne

$$h'(x) = 0 \tag{12.1200}$$

et donc h est constante. Vu que $h(0) = 1$ nous avons $g(x)y(-x) = 1$ pour tout x, c'est-à-dire

$$g(x) = \frac{1}{y(-x)} = y(x). \tag{12.1201}$$

La seconde égalité est le lemme 12.429. □

12.431.
Nous savons qu'il existe une unique solution de l'équation exponentielle avec $a = 1$. Avec la relation

$$g_a'(x) = g_a(x)g_a'(0), \tag{12.1202}$$

de la proposition 12.425, nous n'en sommes pas loin. Il faut encore savoir si il existe un $a > 0$ tel que $g_a'(0) = 1$. Notre culture générale nous dit qu'un tel réel existe et est la fameuse constante e.

Nous nous attelons maintenant à la tâche de montrer l'existence de la chose.

12.36.2 Équation fonctionnelle

Il n'est un secret pour personne (proposition 12.398(3)) que la fonction

$$\begin{aligned} g_a\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto a^x \end{aligned} \tag{12.1203}$$

vérifie l'équation fonctionnelle (12.1196). Nous pouvons nous demander à quel point cette propriété caractérise la fonction puissance.

PROPooJDPEooYTDVtU

Proposition 12.432 ([414]).
Encore plusieurs résultats sur la fonction g_a avec $a > 0$.

ITEMooZJUEooVoqKul

(1) La fonction g_a vérifie l'équation fonctionnelle.

ITEMooCSQXooUDyiMq

(2) La dérivée vérifie $g_a'(0) \neq 0$.

ITEMooCKIHooNuDwrk

(3) Pour tout a, en posant $\alpha = 1/g_a'(0)$ nous avons

$$g_{a^\alpha}'(0) = 1. \tag{12.1204}$$

ITEMooQQFRooWtlViJ

(4) Il existe un unique $e > 0$ tel que

$$g_e' = g_e. \tag{12.1205}$$

ITEMooERTLooWLjlnZ

(5) Pour la valeur de e donnée en (4), la fonction g_e vérifie l'équation exponentielle (12.1197)

$$\begin{cases} g_e' = g_e & \text{(12.1206a)} \\ g_e(1) = e. & \text{(12.1206b)} \end{cases}$$

Démonstration. Un point à la fois.

(1) **Pour (1)** Le fait que g_a vérifie l'équation fonctionnelle est la proposition 12.398(3).

(2) **Pour (2)**
La formule (12.1188) de la proposition 12.425 nous assure que

$$g_a'(x) = g_a(x)g_a'(0). \tag{12.1207}$$

EQooSCDJooTvniEp

Donc $g_a'(0) = 0$ impliquerait que g_a est constante, ce qui n'est pas le cas.

(3) **Pour (3)** Par ailleurs la proposition 12.408 nous permet d'écrire

$$g_a(\alpha x) = g_{a^\alpha}(x). \tag{12.1208}$$

En dérivant des deux côtés,

$$\alpha g'_a(\alpha x) = g'_{a^\alpha}(x). \tag{12.1209}$$

EQooIHCDooGSEGNm

En posant donc $\alpha = g'_a(0)$ et en évaluant (12.1209) en $x = 0$ nous trouvons le résultat.

(4) **Pour (4), existence**
Pour les valeurs de α données par le point (3), nous avons $g'_{a^\alpha}(0) = 1$, et l'équation (12.1207) nous donne alors

$$g'_{a^\alpha}(x) = g_{a^\alpha}(x). \tag{12.1210}$$

Comme de plus $g_{a^\alpha}(0) = 1$, cette fonction vérifie bien l'équation exponentielle.

(5) **Pour (4), unicité**
Si a et b font en sorte que $g'_a = g_a$ et $g'_b = g_b$, alors nous avons aussi $g'_a(0) = g'_b(0) = 1$ à cause de (12.1207). Donc g_a et g_b vérifient l'équation de la proposition 12.430 dont la solution est unique. Donc $g_a = g_b$.
Pour tout x nous avons $g_a(x) = g_b(x)$. En particulier pour $x = 1$ nous avons $a = b$.

□

PROPooGBUPooWtWaFI

Proposition 12.433 ([414]).
Soit $a > 0$. Nous considérons l'équation fonctionnelle 12.427 et l'équation exponentielle 12.428 pour une fonction $f : \mathbb{R} \to \mathbb{R}$.

ITEMooYHAVooWzJqBj

(1) Si f vérifie l'équation fonctionnelle, alors

$$f(q) = a^q \tag{12.1211}$$

pour tout $q \in \mathbb{Q}$.

ITEMooQHOMooNVzSxn

(2) Si f vérifie l'équation fonctionnelle et est monotone, alors $f = g_a$.

ITEMooCNXOooZcrxeB

(3) Si f vérifie l'équation fonctionnelle et est continue, alors $f = g_a$.

Démonstration. En beaucoup de parties. Nous commençons par prouver (1). Nous supposons que $f : \mathbb{R} \to \mathbb{R}$ est une fonction vérifiant l'équation fonctionnelle [161].

(1) $f(x) \geqslant 0$ Quel que soit $x \in \mathbb{R}$ nous avons

$$f(x) = f(\frac{x}{2} + \frac{x}{2}) = f(\frac{x}{2})^2 \geqslant 0. \tag{12.1212}$$

Vous noterez que cet argument ne fonctionne pas si f est à valeurs dans $\mathbb{C}$ au lieu de $\mathbb{R}$.

(2) **Pour $n \in \mathbb{N}$** Soit $n \in \mathbb{N}$. Je vous laisse rédiger la récurrence correctement, mais l'idée est que $f(1) = a$ et ensuite que

$$f(n+1) = f(n)f(1) = f(1)^n f(1) = f(1)^{n+1}. \tag{12.1213}$$

(3) **Pour $m \in \mathbb{Z}$** Nous avons d'une part que $f(-m+m) = f(0) = 1$, mais d'autre part que $f(-m+m) = f(-m)f(m)$. Donc $1 = f(-m)f(m)$; et nous concluons que

$$f(-m) = \frac{1}{f(m)}. \tag{12.1214}$$

161. Bien que ce ne soit pas strictement nécessaire ici, nous rappelons qu'une telle fonction existe par la proposition 12.432.

(4) **Pour $q = 1/n$** Nous savons que $f(1) = a$, mais $1 = \frac{1}{n} + \dots \frac{1}{n}$ (avec n termes), donc

$$a = f(\frac{1}{n} + \dots \frac{1}{n}) = f(\frac{1}{n})^n. \tag{12.1215}$$

Cela implique que $f(\frac{1}{n})^n = a$. La proposition 12.380 indique qu'il existe un unique $x > 0$ tel que $x^n = a$. Vu que nous savons déjà que f est partout positive[162], cette contrainte fixe $f(1/n)$ et la définition 12.381 nous permet d'écrire

$$f(\frac{1}{n}) = a^{1/n}. \tag{12.1216}$$

(5) **Pour $q \in \mathbb{Q}$** Nous posons $q = m/n$. Le nombre q peut être écrit sous la forme $\frac{m}{n} = \frac{1}{n} + \dots + \frac{1}{n}$ avec m termes. Donc

$$f(\frac{m}{n}) = f(\frac{1}{n} + \dots \frac{1}{n}) = f(\frac{1}{n})^m = (a^{1/n})^m = a^{m/n} \tag{12.1217}$$

où nous avons utilisé le lemme 12.387.

La preuve de (1) est terminée.

(1) **Démonstration de (2)** Nous faisons maintenant la preuve de (2). Nous supposons que f vérifie l'équation fonctionnelle et qu'elle est monotone. Pour fixer les idées, nous supposons qu'elle est monotone croissante[163].
Nous considérons les parties[164]

$$A = \{q \in \mathbb{Q} \text{ tel que } q < x\} \tag{12.1218a}$$
$$B = \{q \in \mathbb{Q} \text{ tel que } q > x\}. \tag{12.1218b}$$

Vu que f est croissante, nous avons $f(x) \geqslant f(q)$ pour tout $q \in A$ et $f(x) \leqslant f(q)$ pour tout $q \in B$. En passant au supremum et à l'infimum,

$$\sup_{q\in A} f(q) \leqslant f(x) \leqslant \inf_{q\in B} f(q). \tag{12.1219}$$

Mais il existe dans A une suite strictement croissante convergente q_i vers x (parce que $x = \sup(A)$), donc

$$a^x = \lim_i a^{q_i} \tag{12.1220}$$

par la définition 12.397. Et de même, il existe une suite r_i décroissante dans B telle que $x = \lim r_i$. Cette suite donne aussi

$$a^x = \lim_i a^{r_i}. \tag{12.1221}$$

Nous avons donc l'encadrement

$$a^x \leqslant f(x) \leqslant a^x, \tag{12.1222}$$

qui implique que $f(x) = a^x$.

(2) **Démonstration de (3)**
Nous ne supposons plus que f est monotone. Au lieu de cela nous supposons qu'elle est continue. Nous avons déjà vu en (1) que $f = g_a$ sur $\mathbb{Q}$. Mais par hypothèse f est continue et par la proposition 12.398, g_a est continue. La proposition 12.62 conclut que $f = g_a$ sur $\mathbb{R}$.

□
PROPooLTLWooBGcXAZ

Proposition 12.434 ([414]).
Si y vérifie l'équation exponentielle, alors elle est continue, monotone et vérifie l'équation fonctionnelle.

162. C'est ici que l'hypothèse de fonction à valeurs dans $\mathbb{R}$ est cruciale. Pour $f : \mathbb{R} \to \mathbb{C}$ ceci ne fonctionne pas, et de loin.
163. Si f est monotone décroissante, soit vous adaptez la preuve, soit vous essayez de voir si on ne peut pas recycler le cas croissant en l'appliquant à $-f$.
164. Il du meilleur gout de citer le lemme 1.441 pour dire qu'ils sont non vides.

12.36.3 Dérivation de la fonction puissance (seconde)

La proposition suivante donne la dérivée de $x \mapsto x^q$ pour tout $q \in \mathbb{Q}$. La formule donnée est encore valable pour $x \mapsto x^\alpha$ pour tout $\alpha \in \mathbb{R}$, mais elle demandera plus de théorie pour être démontrée, voir la proposition 14.294.

PROPooSGLGooIgzque

Proposition 12.435 ([1]).
Pour tout $\alpha \in \mathbb{Q}$, si $f_\alpha(x) = x^\alpha$ alors

$$f'_\alpha(x) = \alpha x^{\alpha-1}. \tag{12.1223}$$

En particulier, f_α est de classe C^∞ sur $\mathbb{R}\backslash\{0\}$.

Démonstration. Petit à petit.

(1) **Naturel** Nous prouvons que $(x^n)' = nx^{n-1}$ par récurrence en utilisant la règle de Leibniz de la propositon 12.172(3).
D'abord pour $n = 1$ nous avons $f_1(x) = x$ et donc

$$f'_1(x) = \lim_{\epsilon\to 0} \frac{(x+\epsilon) - x}{\epsilon} = 1. \tag{12.1224}$$

Supposons que $f'_k(x) = kx^{k-1}$ pour un certain $k \in \mathbb{N}$. Nous prouvons que $f'_{k+1}(x) = (k+1)x^k$. Nous avons

$$x^{k+1} = xx^k. \tag{12.1225}$$

En utilisant la règle de Leibniz et l'hypothèse de récurrence,

$$\big(x^{k+1}\big)' = (x)'x^k + x\big(x^k\big)' = x^k + x\big(kx^{k-1}\big) = x^k + kx^k = (k+1)x^k, \tag{12.1226}$$

ce qu'il fallait démontrer.

(2) **Rationnel positif** Soit donc $\alpha = p/q$ avec $p, q \in \mathbb{N}$. Le lemme 12.387 nous permet d'écrire $f_{p/q}(x) = x^{p/q} = (x^p)^{1/q}$. Cela donne

$$f_{p/q}(x)^q = x^p. \tag{12.1227}$$

Nous dérivons cette relation par rapport à x en utilisant à la fois la règle pour les entiers et la règle des fonctions composées [165] :

$$qf'_{p/q}(x)^{q-1} f_{p/q}(x) = px^{p-1}. \tag{12.1228}$$

En isolant $f'_{p/q}(x)$ dans cette expression et en utilisant le fait que $\frac{x^a}{x^b} = x^{a-b}$, nous trouvons le résultat.

(3) **Rationnels négatifs**
Soit $\alpha = -p/q$ avec $p, q \in \mathbb{N}$. Nous avons $x^{-p/q} = \frac{1}{f_{p/q}(x)}$. En utilisant la proposition 12.172(5) et le point déjà prouvé sur les rationnels positifs,

$$f'_{p/q} = -\frac{f'_{-p/q}}{f^2_{p/q}} = -\frac{(-p/q)x^{-p/q-1}}{x^{-2p/q}} = (p/q)x^{p/q-1}. \tag{12.1229}$$

Notez l'utilisation de la proposition 12.408 au dénominateur.

(4) **Irrationnel**
Ah ah ! On vous a bien eu. Les irrationnels, c'est pour la proposition 14.294.

En ce qui concerne le fait que la fonction f_α est de classe C^∞ sur $\mathbb{R}\backslash\{0\}$, c'est simplement une récurrence. Attention : si le rationnel α est négatif, $f_\alpha(0)$ n'est pas défini. Mais, lorsque α est positif non entier, à partir d'un certain ordre, les dérivées font intervenir x^β avec $\beta < 0$. D'où la restriction à $\mathbb{R}\backslash\{0\}$ du domaine sur lequel f_α est de classe C^∞.

Si α est positif entier, alors f_α est de classe C^∞ sur tout $\mathbb{R}$ parce que toutes les dérivées sont nulles à partir d'un certain ordre. □

165. Proposition 12.172(4).

12.36.4 Vers les complexes

Nous avons déjà vu la proposition 12.432 qui dit essentiellement que si une fonction continue $f\colon \mathbb{R} \to \mathbb{R}$ vérifie $f(x+y) = f(x)f(y)$, alors $f(x) = a^x$. Comme indiqué durant la preuve, cette proposition (et en particulier sa preuve) ne fonctionne pas pour les fonctions à valeurs complexes. L'endroit où cela coinçait est que la contrainte

$$f(\frac{1}{n})^n = a \tag{12.1230}$$

n'implique pas grand chose lorsque f est à valeurs complexes.

Nous allons maintenant attaquer ce problème.

LEMooDEGEooXheixp

Lemme 12.436.
Soit $\alpha \in \mathbb{C}$. Si elle existe, la solution au problème

$$\begin{cases} y' = \alpha y & \text{(12.1231a)} \\ y(0) = 1 & \text{(12.1231b)} \end{cases}$$

pour $y\colon \mathbb{R} \to S^1$ est unique.

Démonstration. Soient deux solutions y_1 et y_2. Nous posons $h(x) = y_1(x)y_2(-x)$. Une dérivation donne

$$h'(x) = y_1'(x)y_2(-x) - y_1(x)y_2'(-x). \tag{12.1232}$$

En y substituant $y_1'(x) = \alpha y_1(x)$ et $y_2'(-x) = \alpha y_2(x)$ nous trouvons $h'(x) = 0$. Donc h est constante et nous avons

EQooTWBQooBLLKSt
$$y_1(x)y_2(-x) = 1 \tag{12.1233}$$

pour tout x. Notons que cette identité est encore valable avec $y_1 = y_2$. Nous avons en particulier les égalités $y_1(x)y_1(-x) = 1$ et $y_2(x)y_2(-x) = 1$, et nous notons au passage que $y_1(x)$ et $y_2(x)$ ne s'annulent pas.

En substituant dans (12.1233) la valeur $y_2(-x) = \frac{1}{y_2(x)}$ nous trouvons

$$\frac{y_1(x)}{y_2(x)} = 1, \tag{12.1234}$$

ce qui signifie $y_1(x) = y_2(x)$. □

Dans la proposition suivante, S^1 désigne l'ensemble des nombres complexes de norme 1, dont un paramétrage sera donné dans la proposition 18.61(1) :

$$S^1 = \{e^{ix} \text{ tel que } x \in \mathbb{R}\} = \{e^{ix} \text{ tel que } x \in [0, 2\pi[\}. \tag{12.1235}$$

PROPooVJLYooOzfWCd

Proposition 12.437 ([1]).
Soit une fonction continue $f\colon \mathbb{R} \to S^1$ vérifiant

EQooHANKooHirpTL
$$f(x+y) = f(x)f(y). \tag{12.1236}$$

Alors

(1) f est dérivable,

(2) f satisfait au système

$$\begin{cases} f'(x) = f'(0)f(x) & \text{(12.1237a)} \\ f(0) = 1, & \text{(12.1237b)} \end{cases}$$

(3) il existe $m \in \mathbb{R}$ tel que $f(x) = e^{imx}$.

Démonstration. Pour chaque $m \in \mathbb{R}$, la fonction

$$g_m(x) = e^{imx} \tag{12.1238}$$

vérifie évidemment toutes les conditions. Le but de cette démonstration est de montrer que les conditions imposées à f la déterminent de façon univoque (à part ce m).

La condition (12.1236) nous dit que $f(0) = 1$. Soit une primitive F de f. Il existe $s > 0$ tel que $F(s) > F(0)$ parce que $F' = f$ et $f(0) = 1$.

Soit $a \in \mathbb{R}$. La fonction G_a donnée par $G_a(x) = F(x + a)$ est une primitive de $x \mapsto f(x)f(a)$. Donc $G_a(x) = f(a)F(x)$. Cela dit nous avons

$$f(x)\big(F(s) - F(0)\big) = f(x)F(s) - f(x)F(0) = G_1(x) - G_0(x). \tag{12.1239}$$

Le membre de droite est évidemment dérivable, et $F(s) - F(0) \neq 0$. Donc f est dérivable.

Nous dérivons maintenant la relation $f(x+y) = f(x)f(y)$ par rapport à y en $y = 0$. Cela donne

$$f'(x) = f'(0)f(x). \tag{12.1240}$$

Donc il existe $\alpha \in \mathbb{C}$ tel que $f'(x) = \alpha f(x)$.

Jusqu'ici nous avons prouvé qu'il existe $\alpha \in \mathbb{C}$ tel que

$$\begin{cases} f'(x) = \alpha f(x) & \text{(12.1241a)} \\ f(0) = 1. & \text{(12.1241b)} \end{cases}$$

Or le lemme 12.436 donne l'unicité de la solution à ce système, et il ne faut pas chercher loin : la solution est

$$f(x) = e^{\alpha x}. \tag{12.1242}$$

Pour avoir $f(x) \in S^1$, nous devons de plus imposer que α soit imaginaire pur. Donc, en posant $\alpha = im$, nous avons $m \in \mathbb{R}$ tel que $f(x) = e^{imx}$. □

12.37 Applications analytique entre espaces de Banach

12.37.1 En général

DEFooIXGBooPgJTzB

Définition 12.438 ([417]).
Soient des espaces de Banach[166] *E et F. Une **série de puissance** sur E à valeurs dans F est une application de la forme*

$$x \mapsto \sum_{n=0}^{\infty} \alpha_n(x) \tag{12.1243}$$

où les α_n sont des applications

(1) n-multilinéaire, c'est à dire $\alpha_n \in \mathcal{L}_n(E, F)$ (définition 8.1),

(2) symétrique,

(3) continue.

En termes de notations, nous posons

$$\alpha_n(x) = \alpha(x, \dots, x) = \alpha_n x^n. \tag{12.1244}$$

Notons toutefois que la dernières notation est particulièrement vicieuse parce que, si E est un espace de Banach, il n'y a pas de sens à multiplier des éléments de E entre eux.

Notons aussi que α_0 est une application 0-multilinéaire, c'est à dire juste un élément de F.

DEFooZUYHooDwwRfp

Définition 12.439 ([417]).
*Une série de puissance qui est une somme finie est un **polynôme**.*

166. Définition 7.265.

DEFooAIHMooKbWsBt

Définition 12.440 ([417]).
Soient deux espaces de Banach E et F. Une application $f\colon E \to F$ définie sur un voisinage de $x_0 \in E$ est ***analytique*** *en x_0 si elle peut être écrite sous la forme*[167]

$$f(x) = \sum_{k=0}^{\infty} \alpha_k(x - x_0) \tag{12.1245}$$

où pour chaque $k \in \mathbb{N}$, l'application $\alpha_k \in \mathcal{L}_k(E, F)$ est symétrique et continue.

THOooNBNFooMsAorH

Théorème 12.441 ([417]).
Soient des espaces de Banach E et F sur $\mathbb{C}$. Soient des applications $\alpha_n \in \mathcal{L}(E, F)$. Nous supposons que $\{\alpha_n(x)\}$ est borné. Nous posons $z = tx$ avec $|t| < 1$. Alors

$$\sum_{n=0}^{\infty} \alpha_n(z) \tag{12.1246}$$

converge absolument.

Démonstration. Par hypothèse, il existe $K \in \mathbb{R}$ tel que $\|\alpha_n(x)\| < K$ pour tout n. Alors nous avons

$$\|\alpha_n(z)\| = \|\alpha_n(tz, \dots, tz)\| = \|t^n \alpha_n(z, \dots, z)\| \leqslant |t|^n K. \tag{12.1247}$$

Vu que $|t| < 1$, la série $\sum_{n=0}^{\infty} K|t|^n$ converge (proposition 10.135(1)). Donc la série $\sum_k \|\alpha_n(z)\|$ converge par le critère de comparaison[168] □

NORMooQDCKooXHtrHQ

12.442.
Dans la proposition suivante, nous devons faire appel à quelque astuces de notations. Si $\alpha \in \mathcal{L}_{k+n}(E, F)$ est symétrique, alors nous notons αx^n l'élément de $\mathcal{L}_k(E, F)$ donné par

$$(y_1, \dots, y_k) \mapsto \alpha(x, \dots, x, y_1, \dots, y_n). \tag{12.1248}$$

Notons au passage que l'ordre n'est pas important parce que α est symétrique.

THOooHSXLooMCzTTD

Théorème 12.443 ([417]).
Soient des espaces de Banach E et F. Nous supposons que la série de puissance

$$\begin{aligned} f\colon E &\to F \\ x &\mapsto \sum_{n=0}^{\infty} a_n(x - x_0)^n \end{aligned} \tag{12.1249}$$

ait un rayon de convergence $r_0 > 0$.

Nous posons[169]

$$b_k = \sum_{n=0}^{\infty} \binom{n+k}{n} a_{n+k}(x_1 - x_0)^n \in \mathcal{L}_k(E, F), \tag{12.1250}$$

Alors :

ITEMooWBTAooGrKzpo

(1) f est analytique[170] *en tout point de $B(x_0, r)$,*

(2) La série

$$\sum_{k=0}^{\infty} b_k(x - x_1)^k \tag{12.1251}$$

a un rayon de convergence $r_1 \geqslant r_0 - \|x_1 - x_0\|$.

167. Juste pour rappel, ce que nous notons ici $\alpha_k(x - x_0)$ est ce qui serait noté $\alpha_k(x - x_0)^k$.
168. Lemme 10.128(2).
169. Voir la notation expliquée en 12.442.
170. Définition 12.440.

(3) Pour tout $x \in B(x_1, r_1)$ nous avons

$$f(x) = \sum_{k=0}^{\infty} b_k (x - x_1)^k. \tag{12.1252}$$

PROPooFEMWooDwAYyO

Proposition 12.444 ([1]).
Soient un espace de Banach E ainsi que $a \in E$. L'application

$$x \mapsto x + a \tag{12.1253}$$

est analytique.

PROPooXKHBooQvvNmN

Proposition 12.445 (Composition d'applications analytiques[1]).
Soient des espaces de Banach E_1, E_2 et E_3. Soient des application analytiques $f : E_1 \to E_2$ et $g : E_2 \to E_3$. Alors la composée $g \circ f : E_1 \to E_2$ est analytique.

12.37.2 Série de puissance $\mathbb{C}$

PROPooBFGJooYVvYaw

Proposition 12.446 ([1]).
Si $f : \mathbb{C} \to \mathbb{C}$ est une série de puissances[171] *sur $\mathbb{C}$, alors il existe des $c_k \in \mathbb{C}$ tels que*

$$f(z) = \sum_{k=0}^{\infty} c_k z^k. \tag{12.1254}$$

Démonstration. En effet soit une application $\alpha \in \mathcal{L}_k(\mathbb{C}, \mathbb{C})$. Nous avons $\alpha(z, \dots, z) = z\alpha(1, z, \dots, z)$ et donc par récurrence,

$$\alpha(z, \dots, z) = z^k \alpha(1, \dots, 1). \tag{12.1255}$$

En posant $c_k = \alpha_k(1, \dots, 1)$, nous avons

$$f(z) = \sum_k \alpha_k(z, \dots, z) = \sum_k \alpha_k(1, \dots, 1) z^n = \sum_k c_k z^k. \tag{12.1256}$$

□

12.37.3 L'opérateur d'inversion est analytique

La proposition suivante indique que si A est un opérateur sur un espace de Banach, alors ce que nous noterions naïvement

$$\sum_k a_k A^k \tag{12.1257}$$

est effectivement une série de puissances au sens de la définition 12.438.

L'ensemble $\mathrm{End}^0(E)$ est l'ensemble des endomorphismes bornés de E. De façon équivalente[172], ce sont les endomorphismes continus de E. Je ne suis pas totalement certain que la proposition 12.447 ait besoin de se restreindre à $\mathrm{End}^0(E)$.

PROPooGCLEooBVPHcu

Proposition 12.447 ([1]).
Soit un espace de Banach E. Si $A \in \mathrm{End}^0(E)$ et si $\alpha \in \mathbb{C}$, il existe

$$\alpha \in \mathcal{L}_n\big(\mathrm{End}^0(E), \mathrm{End}^0(E)\big) \tag{12.1258}$$

qui est symétrique, continue et telle que

$$\alpha(A, \dots, A) = \lambda A^n \tag{12.1259}$$

où $A^n = A \circ \dots \circ A$.

171. Définition 12.438.
172. Proposition 10.70.

Démonstration. Nous posons

$$\begin{aligned}\alpha\colon (\mathrm{End}(E)^0)^n &\to \mathrm{End}(E)^0\\ A_1,\dots,A_n &\mapsto \frac{\lambda}{n!}\sum_{\sigma\in S_n} A_{\sigma(1)}\circ\dots\circ A_{\sigma(n)}.\end{aligned}\tag{12.1260}$$

Il est vite vu que α est n-linéaire. Et vu que $|S_n| = n!$, nous avons bien $\alpha(A,\dots,A) = A^n$.

Pour prouver que α est continue, nous utilisons la proposition 8.6. Nous avons

$$\|\alpha(A_1,\dots,A_n)\| \leqslant \frac{1}{n!}\sum_{\sigma\in S_n}\|A_{\sigma(1)}\circ\dots\circ A_{\sigma(n)}\| \tag{12.1261a}$$

$$\leqslant \frac{1}{n!}\sum_{\sigma}\|A_{\sigma(1)}\|\dots\|A_{\sigma(n)}\| \qquad \text{prop. 8.6} \tag{12.1261b}$$

$$= \|A_1\|\dots\|A_n\|. \tag{12.1261c}$$

Donc en posant $\lambda = 1$ avons l'inégalité (8.12). □

PROPooKZNWooLdjDTP

Proposition 12.448 ([1]).
Soit un espace de Banach E. Soient deux opérateur $A, T \in \mathrm{End}^0(E)$. Nous avons

$$T\circ\sum_{k=0}^{\infty}A^k = \sum_{k=0}^{\infty}(TA^k).\tag{12.1262}$$

Démonstration. L'application

$$\begin{aligned}\varphi\colon \mathrm{GL}^0(E) &\to \mathrm{GL}^0(E)\\ X &\mapsto T\circ X\end{aligned}\tag{12.1263}$$

est continue. Donc

$$\varphi\Big(\lim_{N\to\infty}\sum_{k=0}^{N}A^k\Big) = \lim_{N\to\infty}\varphi\Big(\sum_{k=0}^{N}A^k\Big) \tag{12.1264a}$$

$$= \lim_{N\to\infty}\sum_{k=0}^{N}\varphi(A^k) \tag{12.1264b}$$

$$= \sum_{k=0}^{\infty}(TA^k). \tag{12.1264c}$$

□

PROPooPVZBooESAiRv

Proposition 12.449 ([418]).
Soit un espace de Banach E.

(1) L'application

$$\begin{aligned}f\colon B(0,1) &\to \mathrm{End}(E)\\ A &\mapsto \sum_{k=0}^{\infty}A^k\end{aligned}\tag{12.1265}$$

où $B(0,1)$ est le boule de rayon 1 pour la norme opérateur est analytique.

(2) Soit $A \in \mathrm{End}(E)$ avec $\|A\| < 1$. Alors A est inversible et

$$A^{-1} = \sum_{k=0}^{\infty}A^k.\tag{12.1266}$$

Démonstration. En plusieurs parties.

(1) **Analytique** Nous posons

$$\begin{aligned} f\colon B(0,1) &\to \mathrm{End}(E) \\ A &\mapsto \sum_{k=0}^{\infty} A^k \end{aligned} \tag{12.1267}$$

Vu que $\|A\| < 1$, la série converge absolument [173]. Grâce à la proposition 12.447, l'application f est une série de puissances. Le théorème 12.443(1) nous dit alors que f est analytique sur $B(0,1)$.

(2) **Inverse** Nous calculons maintenant $(\mathbb{1} - A)\sum_{k=0}^{\infty} A^k$ en utilisant la proposition 10.135. Nous avons :

$$\begin{aligned} (\mathbb{1} - A)\sum_{k=0}^{\infty} A^k &= \sum_{k=0}^{\infty}(\mathbb{1} - A)A^k && (12.1268a)\\ &= \sum_{k=0}^{\infty}\left(A^k - A^{k+1}\right) && (12.1268b)\\ &= A^0 && \text{prop. 10.105} \quad (12.1268c)\\ &= \mathbb{1} && (12.1268d) \end{aligned}$$

Cela prouve que $(\mathbb{1} - A)$ est inversible et que son inverse est $\sum_{k=0}^{\infty} A^k$.

□

PROPooCHTEooWWzSqK

Proposition 12.450 ([418]).
Soit un espace de Banach E. Soit T inversible et continu dans $\mathrm{End}(E)$. *Nous posons*

$$\rho = \frac{1}{\|T^{-1}\|}. \tag{12.1269}$$

Pour tout $S \in B\big(0, 1/\|T^{-1}\|\big)$, nous avons :

(1) $(T + S)$ est inversible.

(2) La série $\sum_{k=0}^{\infty}(T^{-1}S)^k$ converge absolument.

(3) Nous avons la formule

$$(T + S)^{-1} = \sum_{k=0}^{\infty}(-T^{-1}S)^k T. \tag{12.1270}$$

(4) L'application

$$\begin{aligned} i\colon B(T,\rho) &\to \mathrm{GL}(E) \\ A &\mapsto A^{-1} \end{aligned} \tag{12.1271}$$

est analytique [174].

Démonstration. Nous commençons par considérer $T = \mathrm{Id}$. Nous passons maintenant au cas général. Soient un opérateur inversible T ainsi que $S \in B\big(T, 1/\|T^{-1}\|\big)$. Nous considérons

$$\phi(S) = \sum_{k=0}^{\infty}(-T^{-1}S)^k T^{-1}. \tag{12.1272}$$

(1) **Rayon de convergence** Nous prouvons que cette série a un rayon de convergence $\rho = 1/\|T^{-1}\|$. En effet supposons que $S \in B(0,\rho)$. Vu que T est borné (la continuité est dans les hypothèses), nous pouvons utiliser le lemme 10.69 :

$$\|T^{-1}S\| \leqslant \|T^{-1}\| \cdot \|S\| < 1. \tag{12.1273}$$

Nous avons donc convergence absolue de ϕ sur $B(0,\rho)$. La fonction ϕ est donc analytique sur $B(0,\rho)$.

173. Proposition 10.135(1).
174. Définition 12.438.

(2) **Inverse** Vu que $\| - T^{-1}S\| < 1$, nous pouvons utiliser la proposition 12.449 : l'opérateur $\mathbb{1} - (-T^{-1}S)$ est inversible et son inverse est $\sum_k (-T^{-1}S)^k$.

(3) **Lien avec $(T+S)^{-1}$** Petit calcul. Nous avons :

$$\sum_k (-T^{-1}S)^k T^{-1} = \Big[\sum_k (-T^{-1}S)^k\Big]T^{-1} \tag{12.1274a}$$
$$= \Big[\big(\mathbb{1} - (-T^{-1}S)\big)\Big]^{-1} T^{-1} \tag{12.1274b}$$
$$= \Big[T(\mathbb{1} + T^{-1}S)\Big]^{-1} \tag{12.1274c}$$
$$= (T+S)^{-1}. \tag{12.1274d}$$

Cela prouve que si $S \in B(0,\rho)$, alors $(T+S)^{-1} = \phi(S)$.

(4) **Analycité de i** Pour $A \in B(T,\rho)$, nous posons $S = A - T$, de telle sorte à avoir $\|S\| < \rho$ et $(T+S)^{-1} = \phi(S)$. Nous avons donc

$$i(A) = \phi(A - T) \tag{12.1275}$$

Vu que ϕ est analytique et que $A \mapsto A - T$ est analytique (proposition 12.444), par composition, l'application i est analytique (proposition 12.445).

□

PROPooNRXBooTdEKsu

Proposition 12.451 ([1]).
Soit une isométrie $\varphi\colon E \to F$ entre espaces de Banach. Les applications

$$\begin{aligned} u\colon \mathrm{GL}(E) &\to \mathrm{GL}(F,E) \\ g &\mapsto g \circ \varphi \end{aligned} \tag{12.1276}$$

et

$$\begin{aligned} s\colon \mathcal{L}(E,F) &\to \mathrm{End}(E) \\ f &\mapsto \varphi \circ f. \end{aligned} \tag{12.1277}$$

sont analytiques.

PROPooHJETooZHzodH

Proposition 12.452.
Soient deux espaces de Banach isométrique E et F. La partie $\mathrm{GL}(E,F)$ est ouverte dans $\mathcal{L}(E,F)$.

PROPooYJWGooEbbLVw

Proposition 12.453 ([1]).
Soient deux espaces de Banach isométrique E et F. L'opérateur d'inversion

$$\begin{aligned} j\colon \mathrm{GL}(E,F) &\to \mathrm{GL}(F,E) \\ f &\mapsto f^{-1} \end{aligned} \tag{12.1278}$$

est analytique.

Démonstration. Nous considérons une isométrie $\varphi\colon F \to E$, et l'application

$$\begin{aligned} s\colon \mathcal{L}(E,F) &\to \mathrm{End}(E) \\ f &\mapsto \varphi \circ f. \end{aligned} \tag{12.1279}$$

(1) **φ est linéaire** Vérification immédiate.

(2) **s est une injection** Suppose que $s(f) = s(g)$. Alors pour tout $x \in E$, nous avons $(\varphi \circ f)(x) = (\varphi \circ g)(x)$. Et comme φ est bijective, cela implique $f(x) = g(x)$ et donc que $f = g$.

(3) **φ est surjective** Soit $f \in \mathrm{End}(E)$. Nous avons $f = s(f \circ \varphi^{-1})$.

(4) **s est une isométrie** Nous avons

$$\begin{aligned}
\|s(f)\| &= \sum_{\|x\|=1} \|s(f)(x)\| && (12.1280a)\\
&= \sup_{\|x\|=1} \|(\varphi \circ f)(x)\| && (12.1280b)\\
&= \sup_{\|x\|=1} \|f(x)\| \qquad \varphi \text{ est isométrique} && (12.1280c)\\
&= \|f\| && (12.1280d)
\end{aligned}$$

Nous avons prouvé que $s\colon \mathcal{L}(E,F) \to \mathrm{End}(E)$ est une bijection isométrique.

L'application s se restreint en $s\colon \mathrm{GL}(E,F) \to \mathrm{GL}(E)$. Pour $f \in \mathrm{GL}(E,F)$, nous avons $j(f) = i\big(s(f)\big) \circ \varphi$. En effet

$$i\big(s(f)\big) \circ \varphi = i\big(\varphi \circ f\big) \circ \varphi = f^{-1} \circ \varphi^{-1} \circ \varphi = f^{-1}. \tag{12.1281}$$

En utilisant l'application u de la proposition 12.451, nous avons alors

$$j = u \circ i \circ s. \tag{12.1282}$$

Vu que s, i et u sont analytiques [175], j est analytique. □

12.38 Polynômes de Taylor

DEFooCAppZooTayEsIR

Définition 12.454.
*Soit un ouvert $\Omega \subset \mathbb{C}$. Une fonction $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-**analytique** sur Ω si pour tout $z_0 \in \Omega$, il existe une suite complexe (c_n) et $r > 0$ tels que*

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n \tag{12.1283}$$

pour tout $z \in B(z_0, r)$.

Définition 12.455.
*Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$. Si il existe, nous définissons le n^e **polynôme de Taylor** de f au point $a \in \mathbb{R}$ par*

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \tag{12.1284}$$

*Et la **série de Taylor** de f est la limite :*

$$T(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{12.1285}$$

dans la mesure où la somme converge.

Tant que f est n fois dérivable, le polynôme P_n existe et vérifie $P_n(a) = f(a)$. Nous ne pouvons rien en dire de plus pour l'instant. En particulier, si f est de classe C^∞ il ne faudrait pas croire que

$$\lim_{n\to\infty} P_n(x) = f(x) \tag{12.1286}$$

pour tout x dans un voisinage de a. Autrement dit, même si toutes les dérivées de f existent, la série entière T n'est pas garantie de

— un rayon de convergence [176] plus grand que zéro,

175. Propositions 12.451 et 12.450.
176. Définition 15.12.

— et même avec un grand rayon de convergence, que la limite soit les valeurs de f.

NORMooADIZooUmevqk

12.456.
Il n'est pas très compliqué de construire une fonction f telles que $f(0) = 0$ et telle que $f^{(k)}(0) = 0$ pour tout k, sans pour autant que f soit nulle partout (voir les fonctions plateaux 15.14.1). Les polynômes de Taylor d'une telle fonction sont tous identiquement nuls.

Ceci pour dire qu'en posant

$$T(x) = \sum_{n=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k, \tag{12.1287}$$

nous n'avons aucune garantie de $T = f$, même pas sur le rayon de convergence de la série entière définissant P. Et nous n'avons pas de garanties d'avoir un rayon de convergence plus grand que 0.

Notons toutefois que les polynômes étant denses pour la norme supremum parmi les fonctions continues [177], pour tout compact, il existe une suite de polynômes qui converge uniformément uniformément f. Mais ces polynômes ne sont pas spécialement ceux de Taylor.

12.457.
Ce que nous venons de dire en 12.456 n'est pas vrai pour les fonctions analytiques [178]. Une fonction analytique $f : \Omega \to \mathbb{R}$ s'écrit, autour de 0, sous la forme

$$f(x) = \sum_{n=0}^{\infty} a_n x^n. \tag{12.1288}$$

Demander $f^{(k)}(0) = 0$ pour tout k implique $a_n = 0$ pour tout n, et donc $f = 0$ sur un voisinage de 0.

La condition d'analycité est donc très rigide.

Le théorème de Taylor que nous démontrons à présent n'est pas un résultat que va dans le sens de $\lim_{n\to\infty} P_n(x) = f(x)$. C'est un résultat qui dit juste que $\lim_{x\to a} P_n(x) = f(a)$, et que la limite va d'autant plus vite que n est grand.

Le théorème de Taylor généralise le développement limité au premier ordre de la proposition 10.257.

12.458.
Lorsque le contexte n'est pas ambigu, nous notons simplement P_n le polynôme d'ordre n de f au point a. De même nous notons le reste

$$R_n(x) = f(x) - P_n(x). \tag{12.1289}$$

PROPooUYCMooQjeXpn

Proposition 12.459 ([419])**.**
Soit une fonction f qui est n fois dérivable sur l'intervalle ouvert $I \subset \mathbb{R}$ contenant a. Alors

$$\lim_{x\to a} \frac{R_n(x)}{(x-a)^n} = \lim_{x\to a} \frac{f(x) - P_n(x)}{(x-a)^n} = 0 \tag{12.1290}$$

où P_n est le n^e polynôme de Taylor de f autour de $x = a$.

Démonstration. Pour tout $k = 0, \dots, n$ nous avons $f^{(k)}(a) = P_n^{(k)}(a)$ et donc

$$R_n^{(k)}(a) = 0 \tag{12.1291}$$

pour $k = 0, \dots, n$. En posant d'autre part $s(x) = (x-a)^n$ nous avons $s^{(k)}(a) = 0$ pour tout $k = 0, \dots, n-1$. Par conséquent la règle de l'Hospital de la proposition 12.194 s'applique au quotient $R_n(x)/s(x)^n$. En l'utilisant n fois,

$$\lim_{x\to a} \frac{R_n(x)}{s(x)} = \lim_{x\to a} \frac{R^{(k)}(x)}{k!(x-a)^0} = \frac{0}{k!} = 0. \tag{12.1292}$$

□

177. Théorème 12.420.
178. Définition 12.454 qu'il faut écrire pour $\mathbb{R}$ au lieu de $\mathbb{C}$

Nous démontrons à présent que le polynôme de Taylor est le seul à avoir la propriété de la proposition 12.459.

Proposition 12.460 ([419]).
Soit f, une fonction n fois dérivable sur l'intervalle $I \subset \mathbb{R}$ contenant 0. Soit un polynôme Q de degré n (ou moins) tel que

$$\lim_{x\to a} \frac{f(x)-Q(x)}{(x-a)^n} = 0. \tag{12.1293}$$

Alors Q est le polynôme de Taylor de degré n pour f en a ci-après simplement noté P_n.

Démonstration. D'après la proposition 12.459, la fonction $f-P_n$ vérifie la même limite que $f-Q$. Donc P_n-Q vérifie également la limite

$$\lim_{x\to 0} \frac{(P_n-Q)(x)}{x^n} = 0. \tag{12.1294}$$

Nous notons $P_n(x) = \sum_{k=0}^{n} a_k x^k$ et $Q(x) = \sum_{k=0}^{n} b_k x^k$. La relation (12.1294) donne en particulier

$$\lim_{x\to 0}(P_n-Q)(x) = 0 \tag{12.1295}$$

qui donne $a_0 - b_0 = 0$. Nous continuons par récurrence en supposant que $a_i = b_i$ pour $i = 0, \dots, k$. Alors

$$0 = \lim_{x\to 0} \frac{(P_n-Q)(x)}{x^{k+1}} = \lim_{x\to 0} \sum_{l=k+1}^{n} (a_l - b_l)x^{l-(k+1)}. \tag{12.1296}$$

Le seul terme non nul à droite est celui vérifiant $l-(k+1)=0$. Et ce terme donne l'équation

$$a_{k+1} - b_{k+1} = 0, \tag{12.1297}$$

c'est-à-dire $a_{k+1} = b_{k+1}$. La récurrence continue ainsi jusqu'à $k = n$, et nous pouvons conclure que $Q = P_n$. □

L'intérêt de cette proposition est que si l'on trouve, par n'importe quel moyen, un polynôme Q vérifiant la condition (12.1293), alors nous savons que c'est le polynôme de Taylor.

Théorème 12.461 (Théorème de Taylor[326, 420]).
Soit $I \subset \mathbb{R}$ un intervalle non vide et non réduit à un point de $\mathbb{R}$ ainsi que $a \in I$. Soit une fonction $f\colon I \to \mathbb{R}$ telle que $f^{(n)}(a)$ existe. Alors il existe une fonction α définie sur I et à valeurs dans $\mathbb{R}$ vérifiant les deux conditions suivantes :

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \alpha(x)(x-a)^n, \tag{12.1298a}$$

$$\lim_{t\to a} \alpha(t) = 0 \tag{12.1298b}$$

pour tout $x \in I$. Ici $f^{(k)}$ dénote la k-ième dérivée de f (en particulier, $f^{(0)} = f$, $f^{(1)} = f'$).

Démonstration. Si $R_n(x) = f(x) - P_n(x)$, il suffit de poser

$$\alpha(x) = \frac{R_n(x)}{(x-a)^n} \tag{12.1299}$$

et d'utiliser la proposition 12.459. □

Remarque 12.462.
Quelque remarques.

(1) La formule (12.1298a) est une égalité, et non une approximation. Ce qui serait une approximation serait de réécrire la formule dans le terme contenant α.

(2) Nous avons l'égalité (12.1298a) uniquement sur I. Pour les x hors de I, le polynôme existe évidemment, mais nous n'avons pas spécialement de fonction α, et d'ailleurs la fonction f n'est pas spécialement définie.

12.463.
Les conditions (12.1298) sont souvent aussi énoncées sous la forme qu'il existe une fonction α telle que SUBEQooPYABooKpDgdu

$$\begin{cases} \lim_{t\to 0} \frac{\alpha(t)}{t^n} = 0 & (12.1300a) \\ f(a+h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a) + \cdots + \frac{h^n}{n!}f^{(n)}(a) + \alpha(h). & (12.1300b) \end{cases}$$

Le théorème suivant donne une expression pas tout à fait explicite, mais pas mal quand même pour le reste de Taylor.

THOooSIGRooJTLvlV

Théorème 12.464.
Soient un intervalle ouvert $I \subset \mathbb{R}$ ainsi que $a \in I$. Soit encore une fonction de classe C^{k+1} sur I.

(1) Pour tout $x \in I$, il existe un $c \in]a, x[$ tel que l'égalité

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \quad (12.1301)$$ EQooQFMFooBVpGzy

soit vérifiée.

ITEMooVGBVooGXXvIz

(2) Soit $h \in \mathbb{R}$ tel que $\overline{B(a,h)} \subset I$. Il existe $\theta \in [0,1]$ tel que

$$f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k + \frac{f^{(n+1)}(a+\theta h)}{(n+1)!}h^{n+1}. \quad (12.1302)$$

Démonstration. Pour les besoins de la preuve, nous allons démontrer la formule (12.1301) pour un $b \in I$ au lieu de x. C'est juste que nous allons écrire b au lieu de x parce que nous aurons besoin de la notation x dans le courant de la preuve.

Nous posons

$$R(x) = f(x) - P_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \quad (12.1303)$$

Cela vérifie $R(a) = f(a) - f(a) = 0$ et même

$$R^{(j)}(a) = 0 \quad (12.1304)$$

pour tout $j = 1, \ldots, n$. Nous posons encore

$$F(x) = R(x) - \frac{R(b)}{(b-a)^{n+1}}(x-a)^{n+1}. \quad (12.1305)$$

Nous avons $F^{(j)}(a) = 0$ pour tout $j = 0, \ldots, n$ ainsi que

$$F(b) = R(b) - \frac{R(b)}{(b-a)^{n+1}}(b-a)^{n+1} = 0. \quad (12.1306)$$

et aussi

$$F(a) = R(a) - 0 = 0. \quad (12.1307)$$

Bref, la fonction F vérifie les conditions de la généralisation 12.191 du lemme de Rolle. Il existe donc $c \in]a, b[$ tel que $F^{(n+1)}(c) = 0$. Mais vu que $P_n^{(n+1)}(x) = 0$, nous avons $R^{(n+1)}(x) = f^{(n+1)}(x)$, de telle sorte que

$$f^{(n+1)}(c) = \frac{R(b)(n+1)!}{(b-a)^{n+1}}. \quad (12.1308)$$

En injectant cela dans la définition de F

$$F(x) = R(x) - \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}. \tag{12.1309}$$

En évaluant en $x = b$, et en nous souvenant que $F(b) = 0$, nous trouvons

$$0 = R(b) - \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1} \tag{12.1310}$$

qui est ce que nous voulions prouver. □

Voici un énoncé pour les fonctions à plusieurs variables.

THOooTDFRooEkChgi

Théorème 12.465 ([421]).
Si $f\colon E \to \mathbb{R}$ est une application n fois différentiable en $a \in E$ alors il existe une fonction $\epsilon\colon \mathbb{R} \to \mathbb{R}$ telle que

$$\begin{cases} f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \ldots + & \text{(12.1311a)} \\ \qquad + \ldots + \frac{1}{n!}(d^nf)_a(h,\ldots,h) + \|h\|^n\epsilon(\|h\|) & \text{(12.1311b)} \\ \lim_{t\to 0}\epsilon(t) = 0. & \text{(12.1311c)} \end{cases}$$

12.38.1 Fonctions « petit o »

Nous voulons formaliser l'idée d'une fonction qui tend vers zéro « plus vite » qu'une autre. Nous disons que $f \in o\big(\varphi(x)\big)$ si

$$\lim_{x\to 0}\frac{f(x)}{\varphi(x)} = 0. \tag{12.1312}$$

En particulier, nous disons que $f \in o(x)$ lorsque $\lim_{x\to 0} f(x)/x = 0$.

En termes de notations, nous définissons l'ensemble $o(x)$ l'ensemble des fonctions f telles que

$$\lim_{x\to 0}\frac{f(x)}{x} = 0. \tag{12.1313}$$

Plus généralement si g est une fonction telle que $\lim_{x\to 0} g(x) = 0$, nous disons $f \in o(g)$ si

$$\lim_{x\to 0}\frac{f(x)}{g(x)} = 0. \tag{12.1314}$$

De façon intuitive, l'ensemble $o(g)$ est l'ensemble des fonctions qui tendent vers zéro « plus vite » que g.

Nous pouvons donner un énoncé alternatif au théorème 12.461 en définissant $h(x) = \epsilon(x+a)x^n$. Cette fonction est définie exprès pour avoir

$$h(x-a) = \epsilon(x)(x-a)^n, \tag{12.1315}$$

et donc

$$\lim_{x\to 0}\frac{h(x)}{x^n} = \lim_{x\to 0}\epsilon(x-a) = \lim_{x\to a}\epsilon(x) = 0. \tag{12.1316}$$

Donc $h \in o(x^n)$.

Le théorème dit donc qu'il existe une fonction $\alpha \in o(x^n)$ telle que

$$f(x) = T^a_{f,n}(x) + \alpha(x-a). \tag{12.1317}$$

pour tout $x \in I$.

Remarque 12.466.
À titre personnel, l'auteur de ces lignes déconseille d'utiliser cette notation qui est un peu casse-figure pour qui ne la maîtrise pas bien.

Exemple 12.467.
Le développement en série du cosinus sera traité dans la proposition 18.75. △

PROPooTOXIooMMIghF

Proposition 12.468 (Ordre deux sur $\mathbb{R}^n$[1])**.**
Soit un ouvert Ω de $\mathbb{R}^n$ et $a \in \Omega$ ainsi qu'une fonction $f : \Omega \to \mathbb{R}$ de classe C^2. Alors il existe une fonction $\alpha : \mathbb{R}^n \to \mathbb{R}$ telle que

$$\begin{cases} f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(h) & (12.1318a) \\ \lim_{h\to 0} \alpha(h) = 0. & (12.1318b) \end{cases}$$

Ici, la notation $(d^2f)_a(h,h)$ réfère à ce qui est expliqué en 12.346.

Démonstration. Dans la suite nous considérons t et h tels que toutes les expressions suivantes aient un sens, c'est-à-dire que tous les trucs comme $a + th$ restent dans Ω. Pour $h \in \mathbb{R}^n$ nous nommons e_h le vecteur unitaire dans la direction de h, c'est-à-dire $e_h = h/\|h\|$ et nous posons

$$k_h(t) = f(a + te_h). \tag{12.1319}$$

et nous lui appliquons Taylor 12.461 à l'ordre deux : il existe une fonction β_h telle que

$$k_h(x) = k_h(0) + xk_h'(0) + \frac{x^2}{2}k_h''(0) + x^2\beta_h(x). \tag{12.1320}$$

EQooETDFooAmiRcV

avec $\lim_{x\to 0} \beta_h(x) = 0$.

En ce qui concerne les dérivées de k_h nous avons

$$k_h'(0) = df_a(e_h) \tag{12.1321}$$

et

$$k_h''(0) = (d^2f)_a(e_h, e_h). \tag{12.1322}$$

Il est maintenant temps d'écrire $f(a+h) = k(\|h\|)$ et de substituer les dérivées de k par les différentielles de f dans (12.1320) :

$$f(a+h) = k(\|h\|) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h^2\|\beta_h(\|h\|). \tag{12.1323}$$

EQooUSUGooYPscxV

Il reste à voir que la fonction $\alpha : h \mapsto \beta_h(\|h\|)$ tend vers zéro pour $h \to 0$. En prenant la limite $h \to 0$ dans (12.1323), il est manifeste que la limite du membre de gauche existe et vaut $f(a)$. Donc la limite du membre de droite doit exister et valoir également $f(a)$. Nous en déduisons que la limite de

$$df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\beta_h(\|h\|) \tag{12.1324}$$

existe et vaut zéro. La limite des deux premiers termes existe et vaut zéro, donc la limite du troisième existe et vaut zéro :

$$\lim_{h\to 0} \|h\|^2\beta_h(\|h\|) = 0. \tag{12.1325}$$

□

PROPooWWMYooPOmSds

Proposition 12.469.
Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction $f : \Omega \to \mathbb{R}$ de classe C^1 et deux fois différentiable sur $]x, x+h[$. Alors il existe $\theta \in]0,1[$ tel que

$$f(x+h) = f(x) + df_x(h) + \frac{1}{2}(d^2f)_{x+\theta h}(h,h). \tag{12.1326}$$

12.38.2 Autres formulations

ExempleUtlDev

Exemple 12.470.
Une des façons les plus courantes d'utiliser les formules (12.1298) est de développer $f(a+t)$ pour des petits t en posant $x=a+t$ dans la formule :

$$f(a+t)=f(a)+f'(a)t+f''(a)\frac{t^2}{2}+\epsilon(a+t)t^2 \tag{12.1327}$$

EqDevfautouraens

avec $\lim_{t\to 0}\epsilon(a+t)=0$. Ici, la fonction T dont on parle dans le théorème est $T^a_{f,2}(a+t)=f(a)+f'(a)t+f''(a)\frac{t^2}{2}$.

Lorsque x et y sont deux nombres « proches [179] », nous pouvons développer $f(y)$ autour de $f(x)$:

$$f(y)=f(x)+f'(x)(y-x)+f''(x)\frac{(y-x)^2}{2}+\epsilon(y-x)(y-x)^2, \tag{12.1328}$$

Eqfydevfx

et donc écrire

$$f(x)-f(y)=-f'(x)(y-x)-f''(x)\frac{(y-x)^2}{2}-\epsilon(y-x)(y-x)^2. \tag{12.1329}$$

De cette manière nous obtenons une formule qui ne contient plus que y dans la différence $y-x$. △

12.38.3 Formule et reste

PropDevTaylorPol

Proposition 12.471.
Soient $f\colon I\subset\mathbb{R}\to\mathbb{R}$ et $a\in\mathrm{Int}(I)$. Soit un entier $k\geqslant 1$. Si f est k fois dérivable en a, alors il existe un et un seul polynôme P de degré $\leqslant k$ tel que

$$f(x)-P(x-a)\in o\big(|x-a|^k\big) \tag{12.1330}$$

lorsque $x\to a$, $x\neq a$. Ce polynôme est donné par

$$P(h)=f(a)+f'(a)h+\frac{f''(a)}{2!}h^2+\cdots+\frac{f^{(k)}(a)}{k!}h^k. \tag{12.1331}$$

Notons encore deux façons alternatives d'écrire le résultat. Si $f\in C^k$ il existe une fonction α telle que $\lim_{t\to 0}\alpha(t)=0$ et

$$f(x)=\sum_{n=0}^{k}\frac{f^{(n)}(a)}{n!}(x-a)^n+(x-a)^n\alpha(x-a). \tag{12.1332}$$

Si $f\in C^{k+1}$ alors

$$f(x)=\sum_{n=0}^{k}\frac{f^{(n)}(a)}{n!}(x-a)^n+(x-a)^{n+1}\xi(x-a) \tag{12.1333}$$

EquQtpoN

où ξ est une fonction telle que $\xi(t)$ tend vers une constante lorsque $t\to 0$.

La proposition suivante donne une intéressante façon de trouver le reste d'un développement de Taylor.

PropResteTaylorc

Proposition 12.472.
Soient I, un intervalle dans $\mathbb{R}$ et $f\colon I\to\mathbb{R}$ une fonction de classe C^k sur I telle que $f^{(k+1)}$ existe sur I. Soient $a\in\mathrm{Int}(I)$ et $x\in I$. Alors il existe $c\in\,]x,a[$ tel que

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}. \tag{12.1334}$$

179. par exemple dans une limite $(x,y)\to(h,h)$.

12.38.4 Reste intégral

Comme son nom l'indique, le « reste intégral » demande de savoir les intégrales. La formule du reste intégral sera donc pour après la définition des intégrales, proposition 20.149.

12.39 Développement limité autour de zéro

Dans cette sections nous supposons toujours que les fonctions sont définies sur un intervalle ouvert de $\mathbb{R}$, I, contenant 0.

12.39.1 Généralités

Définition 12.473.
*Soit $f : I \to \mathbb{R}$ une fonction définie sur un ouvert I autour de zéro. Nous disons que f admet un **développement limité** autour de 0 à l'ordre n si il existe une fonction $\alpha : I \to \mathbb{R}$ telle que*

$$\begin{cases} f(x) = P_n(x) + x^n\alpha(x) & (12.1335a) \\ \lim_{x\to 0} \alpha(x) = 0 & (12.1335b) \end{cases}$$

*où $P_n(x) = a_0 + a_1x + \cdots + a_nx^n$ est une polynôme de degré n. Le polynôme P_n est appelé la **partie régulière** du développement.*

La fonction α est appelé le **reste** du développement et sera parfois noté α_f. Lorsque P est la partie régulière d'un développement limité de f nous notons parfois $f \sim P$.

Proposition 12.474 (Troncature)**.**
Si f admet un développement limité d'ordre n alors il admet également un développement limité d'ordre n' pour tout $n' < n$. Ce dernier s'obtient en tronquant le polynôme d'ordre n à l'ordre n'.

Proposition 12.475 (Unicité)**.**
Si f admet une développement limité alors ce dernier est unique : il existe un unique polynôme P_n d'ordre n et une unique fonction α vérifiant simultanément les deux conditions

$$\begin{cases} f(x) = P_n(x) + x^n\alpha(x), & (12.1336a) \\ \lim_{x\to 0} \alpha(x) = 0. & (12.1336b) \end{cases}$$

ExTHGooCBcnAy

Exemple 12.476.
En ce qui concerne les séries géométriques de raison x nous savons les formules

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x} \tag{12.1337}$$

et

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1 - x} \tag{12.1338}$$

pour tout $x \in]-\infty, 1[$. Comparant les deux, il est naturel d'essayer de prendre $1 + x + x^2 + \cdots + x^n$ comme développement limité de la fonction $f(x) = \frac{1}{1-x}$. Pour voir si cela fonctionne, il faut vérifier si « le reste » est bien de la forme $x^n\alpha(x)$ avec $\lim_{x\to 0} \alpha(x) = 0$.

Le reste en question est donné par

$$\frac{1}{1 - x} - 1 - x - x^2 - \ldots - x^n = \frac{1}{1 - x} - \frac{1 - x^{n+1}}{1 - x} = \frac{x^{n+1}}{1 - x} = x^n\frac{x}{1 - x}. \tag{12.1339}$$

En posant $\alpha(x) = \frac{x}{1-x}$ nous avons donc bien

$$f(x) = \frac{1}{1 - x} = 1 + x + x^2 + \cdots + x^n + x^n\frac{x}{1 - x} \tag{12.1340}$$

et $\lim_{x\to 0} \frac{x}{1-x} = 0$. Cela est le développement limité de f à l'ordre n autour de 0. △

La formule des accroissements finis est un cas particulier de développement fini. Supposons que f soit dérivable en 0. En effet nous pouvons facilement trouver la fonction α qui convient. Sachant que $f(0)+xf'(0)$ donne l'approximation affine de f autour de 0, nous cherchons α en écrivant

$$f(x)=f(0)+xf'(0)+x\alpha(x). \tag{12.1341}$$

Cela nous pousse à définir

$$\alpha(x)=\frac{f(x)-f(0)}{x}-f'(0). \tag{12.1342}$$

EqDCEooKozKx

Notons que cette fonction n'est pas définie en $x=0$, mais cela n'a pas d'importance : seule la limite $\lim_{x\to 0}\alpha(x)$ nous intéresse. Par définition de la dérivée,

$$\lim_{x\to 0}\alpha(x)=\lim_{x\to 0}\frac{f(x)-f(0)}{x}-f'(0)=0. \tag{12.1343}$$

En conclusion si f est dérivable, son développement limité à l'ordre 1 est donné par

$$f(x)=f(0)+xf'(0)+x\alpha(x) \tag{12.1344}$$

où $\alpha(x)$ est donnée par la formule (12.1342).

12.39.2 Formule de Taylor-Young

Plus généralement nous avons la proposition suivante qui donne le développement limité de toute fonction dérivable n fois.

PropVDGooCexFwy

Proposition 12.477 (Formule de Taylor-Young)**.**
Soit f une fonction n fois dérivable sur un intervalle I contenant 0. Alors il existe une fonction $\alpha\colon I\to\mathbb{R}$ telle que

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2}x^2+\frac{f^{(3)}(0)}{3!}x^3+\cdots+\frac{f^{(n)}(0)}{n!}x^n+x^n\alpha(x) \tag{12.1345}$$

EQooBKZDooToYvIR

et

$$\lim_{x\to 0}\alpha(x)=0. \tag{12.1346}$$

Cette proposition nous permet de calculer facilement des développements limités tant que nous sommes capables de calculer les dérivées successives de la fonction à développer. Dans l'exemple 12.476 nous avons dû utiliser des astuces et des formules pour déterminer le développement limité de $\frac{1}{1-x}$. Au contraire la formule (12.1345) nous permet de trouver le polynôme en appliquant mécaniquement une formule simple.

Exemple 12.478.
Utilisation de la formule (12.1345) pour déterminer le développement limité de la fonction

$$f(x)=\frac{1}{1-x}. \tag{12.1347}$$

Il faut calculer les dérivées successives de f :

$$f(x)=\frac{1}{1-x} \tag{12.1348a}$$

$$f'(x)=\frac{1}{(1-x)^2} \tag{12.1348b}$$

$$f''(x)=\frac{2}{(1-x)^3} \tag{12.1348c}$$

Avec ces résultats, nous devinons que

$$f^{(n)}(x)=\frac{n!}{(1-x)^{n+1}}. \tag{12.1349}$$

Pour en être sûr nous le prouvons par récurrence. La dérivée de $\frac{n!}{(1-x)^{n+1}}$ est donnée par

$$\frac{n!(n+1)(1-x)^n}{(1-x)^{2n+2}} = \frac{(n+1)!}{(1-x)^{n+2}}. \tag{12.1350}$$

Évaluées en $x = 0$, les dérivées successives de f sont $f(0) = 0$, $f'(0) = 1$, $f''(0) = 2$,...,$f^{(n)}(0) = n!$. Utilisant la formule (12.1345) nous avons

$$f(x) = 1 + x + x^2 + \cdots + x^n + x^n\alpha(x), \tag{12.1351}$$

conformément à ce que nous avions déjà trouvé. △

EXooFLBJooYfuRsG

Exemple 12.479.
Soient $r \in \mathbb{Q}$ et la fonction donnée par

$$f(x) = (1 + x)^r. \tag{12.1352}$$

Nous notons I le domaine de cette fonction : c'est $\mathbb{R}$ si $r > 0$ ou $[-1, \infty]$ si $r < 0$. Si par contre $r = 0$, la fonction est constante et le domaine est $I = \mathbb{R}$.

En ce qui concerne les dérivées [180] : $f'(x) = r(1 + x)^{r-1}$ et plus généralement

$$f^{(k)}(x) = r(r - 1)\dots(r - k + 1)(1 + x)^{r-k} \tag{12.1353}$$

si $k > 0$. Pour $k = 0$ nous avons $f^{(k)}(0) = 1$. Le développement de Taylor-Young est alors

$$(1 + x)^r = 1 + \sum_{k=1}^{n} \frac{r(r - 1)\dots(r - k + 1)}{k!}x^k + x^n\alpha(x). \tag{12.1354}$$

Notons que que si r est un entier, pour $k = r$, le produit au numérateur s'annule et le développement s'arrête.

Dans le développement de $(1 + x)^r$, nous reconnaissons la formule de $\binom{r}{k}$, sauf que nous ne pouvons pas l'écrire avec cette notation lorsque r n'est pas entier. △

Cet exemple fonctionnera encore avec $r \in \mathbb{R}$ au lieu de $r \in \mathbb{Q}$, mais il faudra la proposition 15.95 pour la dérivée

Remarque 12.480.
Pour alléger la notation et ne pas écrire $\dots + x^n\alpha(x)$ nous pouvons aussi écrire

$$f(x) \sim 1 + x + x^2 + \cdots + x^n, \tag{12.1355}$$

mais il est interdit d'écrire

$$f(x) = 1 + x + x^2 + \cdots + x^n \tag{12.1356}$$

en mettant un signe d'égalité entre une fonction et son développement limité [181].

Notons cependant que la proposition 12.477 ne donne pas de moyen simple de trouver la fonction α. Si la fonction f est très régulière dans l'intervalle I on a le résultat suivant.

Proposition 12.481 (Reste dans la forme de Lagrange)**.**
Si la fonction f est dérivable $n + 1$ fois dans I alors il existe $\bar{x}$ dans l'intervalle $[0, x]$ tel que

$$f(x) = P_n(x) + \frac{1}{(n + 1)!}f^{n+1}(\bar{x})x^{n+1}. \tag{12.1357}$$

180. Nous utilisons la proposition 12.435.

181. Il faut cependant être très prudents avec la notation abrégée. Elle pourrait nous faire oublier des informations importantes, voir les développements des fonctions trigonométriques pour un exemple.

12.39.3 Règles de calcul

Les règles suivantes permettent de calculer les développements limités des fonctions qu'on peut écrire comme combinaison de fonctions dont nous savons déjà le développement.

Il est toujours possible de calculer le développement limité d'une fonction par la formule de Taylor-Young (proposition 12.477). Les règles suivantes peuvent nous économiser de l'effort et du temps.

12.39.3.1 Linéarité des développements limités

L'opération qui consiste à prendre le développement limité d'une fonction est une opération linéaire : connaissant les développements limités de f et de g, il suffit de les sommer pour obtenir celui de $f+g$. De même, si λ est une constante, le développement limité de λf est le développement limité de f fois λ.

PROPooAUVPooGRIZYF

Proposition 12.482.
Soient λ et μ dans $\mathbb{R}$. Si f et g sont deux fonctions acceptant des développements limités d'ordre n EqJPPooCHihNn

$$f(x) = P(x) + x^n\alpha_f(x) \tag{12.1358a}$$
$$g(x) = Q(x) + x^n\beta(x) \tag{12.1358b}$$

avec $\lim_{x\to 0}\alpha(x) = \lim_{x\to 0}\beta(x) = 0$, alors la fonction $\lambda f + \mu g$ admet le développement limité

$$(\lambda f + mug)(x) = (\lambda P + \mu Q)(x) + x^n(\lambda\alpha + \mu\beta)(x). \tag{12.1359}$$ EqCJFooVnvCtz

Remarque 12.483.
La forme explicite du reste ne nous interesse pas. Dans la pratique on écrira toujours $(f+g)(x) = (P+Q)(x) + x^n\alpha(x)$, où on appelle α une fonction opportune telle que $\lim_{x\to 0}\alpha(x) = 0$.

Démonstration. Vu les définitions (12.1358) des polynômes P, Q et des restes α et β, l'égalité (12.1359) est une conséquence de la linéarité de la dérivation et de la proposition 12.477

De plus $P+Q$ est un polynôme de degré n dès que P et Q sont des polynômes de degré n, et

$$\lim_{x\to 0}(\lambda\alpha + \mu\beta)(x) = \lim_{x\to 0}\lambda\alpha(x) + \lim_{x\to 0}\mu\beta(x) = 0. \tag{12.1360}$$

Par conséquent $\lambda\alpha + \mu\beta$ est la fonction de reste de $\lambda f + \mu g$. □

ExKPBooJmdFvY

Exemple 12.484.
Calculer le développement de la fonction

$$f(x) = 3\sqrt[3]{1+x} + e^{-2x}. \tag{12.1361}$$

Le développement de $\sqrt[3]{1+x}$ est donné par la formule de l'exemple 12.479 avec $\alpha = \frac{1}{3}$. Nous avons donc dans un premier temps

$$\sqrt[3]{1+x} = 1 + \frac{1}{3}x + \frac{\frac{1}{3}\left(\frac{1}{3}-1\right)}{2}x^2 + \frac{\frac{1}{3}\left(\frac{1}{3}-1\right)\left(\frac{1}{3}-2\right)}{6}x^3 + x^3\alpha(x) \tag{12.1362a}$$
$$= 1 + \frac{1}{3}x - \frac{1}{9}x^2 + \frac{5}{81}x^3 + x^3\alpha(x). \tag{12.1362b}$$

Nous avons alors

$$3\sqrt[3]{1+x} + e^{-2x} = 3\Big[1 + \frac{1}{3}x - \frac{1}{9}x^2 + \frac{5}{81}x^3 + x^3\alpha(x)\Big] + 1 - 2x + 2x^2 - \frac{4}{3}x^3 + x^3\beta(x) \tag{12.1363a}$$
$$= 4 - x + \frac{5}{3}x^2 - \frac{31}{27}x^3 + x^3\big(\alpha(x) + \beta(x)\big). \tag{12.1363b}$$

△

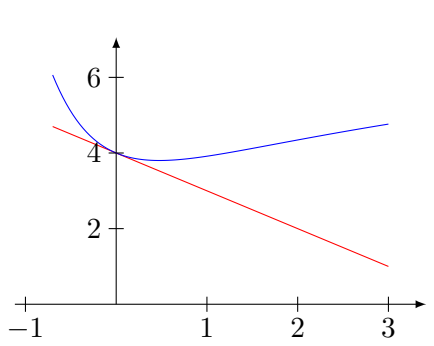
La condition $\lim_{x\to 0}\alpha(x)=0$ signifie que l'approximation qui consiste à remplacer $f(x)$ par le polynôme n'est pas une trop mauvaise approximation lorsque x est petit. Cela ne signifie rien de plus. En particulier si x est grand, l'approximation polynomiale peut-être (et est souvent) très mauvaise.

À ce propos, notez qu'un polynôme tend toujours vers $\pm\infty$ lorsque x est grand. Une approximation polynomiale d'une fonction bornée est donc toujours (très) mauvaise pour les grandes valeurs de x.

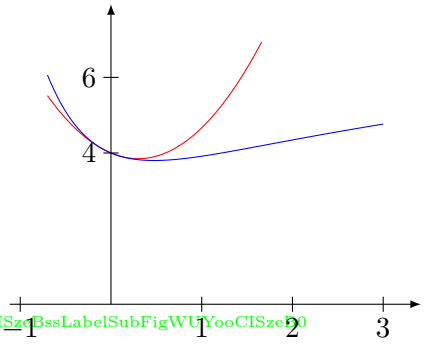
À titre d'exemple nous avons tracé sur la figure 12.12 la fonction

$$f(x)=3\sqrt[3]{x+1}+e^{-2x} \tag{12.1364}$$

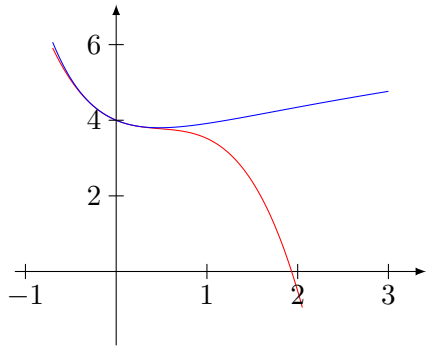
et ses développements limités d'ordre 1 à 3. Il est particulièrement visible que l'approximation est assez bonne pour la partie gauche du graphe sur laquelle la fonction est bien croissante, alors qu'elle est franchement mauvaise sur la droite où le graphe ressemble plutôt à une constante [182].



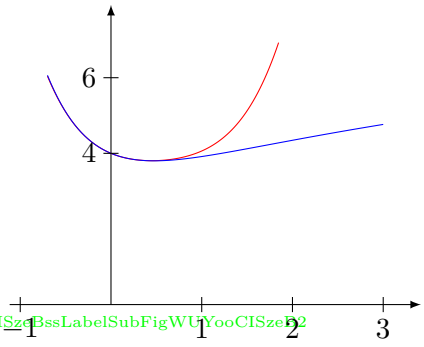
(a) Polyôme d'ordre 1



(b) Polyôme d'ordre 2



(c) Polyôme d'ordre 3



(d) Polyôme d'ordre 4

FIGURE 12.12: Les développements limités d'ordre de plus en plus grand de la fonction de l'exemple 12.484. La fonction est en bleu et les « approximations » sont en rouge.LabelFigWUYooCISzeB

12.39.3.2 Développement limité d'un quotient

PROPooMANAooXhuanS

Proposition 12.485.
Si P_f est le polynôme du développement limité de f à l'ordre n et P_g celui de g, alors nous obtenons le développement limité de f/g à l'ordre n en effectuant la division selon les puissances croissantes de P_f par P_g.

Attention : il s'agit bien de faire une division selon les puissances croissantes, et non une divisions euclidienne. La division euclidienne de A par B consiste à écrire $A=BQ+R$ avec le reste R de degré le plus *petit* possible. Ici nous voulons avoir un reste de degré le plus *grand* possible.

182. Pouvez-vous cependant dire que vaut $\lim_{x\to\infty}f(x)$?

12.39.3.3 Développement limité d'une fonction composée

Proposition 12.486.
Soient f et g des fonctions admettant des développements limités d'ordre n au voisinage de 0. Nous supposons que $\lim_{x\to 0} g(x) = 0$. Alors la composée $f\big(g(x)\big)$ admet un développement limité d'ordre n au voisinage de 0 qui s'obtient en substituant le développement de g à chaque «x» du développement de f, et en supprimant tous les termes de degré plus élevé que n.

12.40 Développement ailleurs qu'à l'origine

Il est intéressant de développer une fonction au voisinage de zéro lorsque nous nous intéressons à son comportement pour les x pas très grands. Il est toutefois souvent souhaitable de savoir le comportement d'une fonction au voisinage d'autres valeurs que zéro.

Pour développer la fonction f autour de x_0, nous considérons la fonction $h \mapsto f(x_0 + h)$ que nous développons autour de zéro (pour h). L'objectif est de trouver une polynôme P et une fonction α tels que

$$\begin{cases} f(x) = P(x) + (x - x_0)^n \alpha(x) & \text{(12.1365a)} \\ \lim\limits_{x\to x_0} \alpha(x) = 0. & \text{(12.1365b)} \end{cases}$$

En pratique, le développement limité à l'ordre n d'une fonction autour d'un point x_0 quelconque à l'intérieur de son domaine prend la forme suivante, qui généralise la formule de Taylor-Young vue dans la proposition 12.477

Proposition 12.487 (Formule de Taylor-Young, cas général)**.**
Soit f une fonction n fois dérivable sur un intervalle I contenant x_0. Alors il existe une fonction $\alpha\colon I \to \mathbb{R}$ telle que

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \\ + \frac{f^{(3)}(x_0)}{3!}(x - x_0)^3 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + (x - x_0)^n \alpha(x - x_0) \quad \text{(12.1366)}$$

EqTJRooUbsyzJ

et

$$\lim_{t\to 0} \alpha(t) = 0. \quad \text{(12.1367)}$$

12.41 Développement au voisinage de l'infini

Il est souvent utile de connaitre le comportement d'une fonction pour les grandes valeurs de x et de déterminer ses asymptotes éventuelles. La technique que nous allons utiliser consiste à poser $x = \frac{1}{h}$ et de développer la fonction "auxiliaire" $g(h) = f(1/h)$ autour de $h = 0$. La limite avec $h \to 0^+$ donnera le comportement pour $x \to \infty$ et la limite $h \to 0^-$ donnera le comportement pour $x \to -\infty$.

Dans le cas d'une développement autour de $\pm\infty$ nous ne parlons plus de développement *limité* mais de **développement asymptotique**.

12.41.1 La fonction puissance : remarques pour la suite

Il y a encore de nombreuses choses à dire sur la fonction puissance. Pour savoir lesquelles, voir le thème 52.

12.42 Fonctions réelles de deux variables réelles

Une **fonction réelle de 2 variables réelles** est une fonction $f : A \subset \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto z = f(x, y)$.

Le **graphe de** f, noté $\mathrm{Gr}\, f$, est un sous-ensemble de $\mathbb{R}^3$:

$$\mathrm{Gr}\, f = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in A \text{ et } z = f(x, y)\}$$

Les **courbes de niveau** de la fonction f sont obtenues en posant $f(x, y) = \lambda$.

12.42.1 Limites de fonctions à deux variables

Ici nous n'allons pas entrer dans tous les détails, mais simplement mentionner les quelques techniques les plus courantes.

ThoLimiteCompose

Théorème 12.488.
Soient deux fonctions $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^p \to \mathbb{R}^q$. Si a est un point adhérent au domaine de $g \circ f$ et si

$$\begin{aligned} \lim_{x \to a} f(x) &= b \\ \lim_{y \to b} g(y) &= c, \end{aligned} \tag{12.1368}$$

alors

$$\lim_{x \to a} (g \circ f)(x) = c. \tag{12.1369}$$

Les techniques usuelles sont

(1) La règle de l'étau. Cette technique demande un peu plus d'imagination parce qu'il faut penser à un « truc » différent pour chaque exercice. En revanche, la justification est facile : il y a un théorème qui dit que ça marche.

(2) Lorsqu'on applique la règle de l'étau, penser à

$$|x| = \sqrt{x^2} \leqslant \sqrt{x^2 + y^2}. \tag{12.1370}$$

Cela permet de majorer le numérateur. Attention : ce genre de majoration fonctionne seulement au numérateur : agrandir le dénominateur ferait diminuer la fraction.

(3) Il n'est pas vrai que

$$|x| = \sqrt{x^2} \leqslant \sqrt{x^4} \leqslant \sqrt{x^4 + 2y^4}. \tag{12.1371}$$

En effet, si x est petit, alors $x^2 > x^4$, et non le contraire.

Une technique très efficace pour les limites $(x, y) \to (0, 0)$ est le passage aux coordonnées polaires. Il s'agit de poser

$$\begin{cases} x = r\cos(\theta) & \text{(12.1372a)} \\ y = r\sin(\theta) & \text{(12.1372b)} \end{cases}$$

et puis de faire la limite $r \to 0$.

Si la limite obtenue **ne dépend pas de** θ, alors c'est la limite cherchée. Voici quelque exemples.

EXooASMWooCBSyVf

Exemple 12.489.
Calculer la limite

$$\lim_{(x,y) \to (0,0)} \frac{x - y}{x + y}. \tag{12.1373}$$

Ici la méthode des chemins pour est particulièrement éclairante. Regardons d'abord la fonction sur la droite $x = y$. Nous avons

$$f(x, y) = \frac{x - x}{2x} = 0. \tag{12.1374}$$

Donc la fonction est nulle sur toute la ligne.

Si nous regardons maintenant la ligne verticale $x = 0$, nous avons

$$f(0, y) = \frac{-y}{y} = -1, \tag{12.1375}$$

donc la fonction vaut -1 sur toute la ligne verticale. △

Beaucoup d'autres exemples seront donné en 18.233 lorsque nous aurons vu les coordonnées polaires.

12.42.2 Dérivées partielles

La **dérivée partielle** par rapport à x au point (x,y) est notée

$$\frac{\partial f}{\partial x}(x,y) \tag{12.1376}$$

et se calcule en dérivant f par rapport à x en considérant que y est constante.

De la même manière, la dérivée partielle par rapport à y au point (x,y) est notée

$$\frac{\partial f}{\partial y}(x,y) \tag{12.1377}$$

et se calcule en dérivant f par rapport à y en considérant que x est constante.

Pour les dérivées partielles secondes,

— $f''_{xx}(x,y) = (f'_x)'_x = \frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}(\frac{\partial f}{\partial x})$.

— $f''_{yy}(x,y) = (f'_y)'_y = \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}(\frac{\partial f}{\partial y})$.

— $f''_{xy}(x,y) = (f'_x)'_y = (f'_y)'_x = f''_{yx}(x,y)$ ou $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}) = \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

12.42.3 Différentielle et accroissement

La **différentielle totale** de f au point (a,b) est donnée, quand elle existe (!), par la formule

$$df(a,b) = \frac{\partial f}{\partial x}(a,b)dx + \frac{\partial f}{\partial y}(a,b)dy. \tag{12.1378}$$

De la même façon que la formule des accroissements finis disait que $f(x+a) \simeq f(x) + af'(x)$, en deux dimensions nous avons que l'**accroissement** approximatif de f au point (a,b) pour des accroissements Δx et Δy est

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta x \frac{\partial f}{\partial x}(x,y) + \Delta y \frac{\partial f}{\partial y}(x,y). \tag{12.1379}$$

Le **plan tangent** au graphe de f au point $\big(a,b,f(a,b)\big)$ est

$$T_{(a,b)}(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \tag{12.1380}$$

essayez d'écrire l'équation de la droite tangente au graphe de $f(x)$ au point $x=a$ en termes de la dérivée de f, et comparez votre résultat à cette formule.

Un des principaux théorèmes pour tester la différentiabilité d'une fonction est le suivant.

ThoProuverDiffable

Théorème 12.490.
Soit une fonction $f\colon \mathbb{R}^m \to \mathbb{R}^p$. Si les dérivées partielles existent dans un voisinage de a et donc continues en a, alors f est différentiable en a.

Le plus souvent, nous prouvons qu'une fonction est différentiable en calculant les dérivées partielles et en montrant qu'elles sont continues.

Dérivation implicite : Soit $F(x,f(x)) = 0$ la représentation implicite d'une fonction $y = f(x)$ alors

$$y' = f'(x) = -\frac{F'_x}{F'_y}.$$

12.43 Les fonctions à valeurs vectorielles

Jusqu'à présent nous avons vu des fonctions de plusieurs variables qui prenaient leurs valeurs dans $\mathbb{R}$. Nous allons maintenant voir ce qu'il se passe lorsque les fonctions prennent leurs valeurs dans $\mathbb{R}^3$.

Une fonction d'une variable est dite **à valeurs vectorielles** lorsque

$$\begin{aligned} f\colon I \subset \mathbb{R} &\to \mathbb{R}^3 \\ f(x) &= \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}. \end{aligned} \tag{12.1381}$$

Les fonctions $f_i\colon \mathbb{R} \to \mathbb{R}$ sont les **composantes** de f. Ce que nous avons raconté à propos des dérivées passe facilement :

$$\frac{f(a+\epsilon)-f(a)}{\epsilon} = \begin{pmatrix} \frac{f_1(a+\epsilon)-f_1(a)}{\epsilon} \\ \frac{f_2(a+\epsilon)-f_2(a)}{\epsilon} \\ \frac{f_3(a+\epsilon)-f_3(a)}{\epsilon} \end{pmatrix}. \tag{12.1382}$$

En particulier dès que les fonctions f_i sont dérivables, nous avons

$$f'(a) = \begin{pmatrix} f_1'(a) \\ f_2'(a) \\ f_3'(a) \end{pmatrix} \tag{12.1383}$$

comme dérivée de la fonction. Cette dérivée est un vecteur.

Exemple 12.491.
Si

$$f\colon x \in \mathbb{R} \mapsto \begin{pmatrix} x^2e^x \\ \cos(x^2) \\ x^3+x \end{pmatrix}, \tag{12.1384}$$

alors

$$f'(x) = \begin{pmatrix} 2xe^x + x^2e^x \\ -2x\sin(x^2) \\ 3x^2+1 \end{pmatrix}. \tag{12.1385}$$

△

12.44 Fonctions vectorielles de plusieurs variables

Ce sont les fonctions de la forme

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix}. \end{aligned} \tag{12.1386}$$

En ce qui concerne les dérivées, tout se passe comme avant. Si les dérivées partielles des composantes f_i existent au point $a \in \mathbb{R}^3$, alors

$$\frac{\partial f}{\partial x}(a) = \begin{pmatrix} \partial_x f_1(a) \\ \partial_x f_2(a) \\ \partial_x f_3(a) \end{pmatrix}, \quad \frac{\partial f}{\partial y}(a) = \begin{pmatrix} \partial_y f_1(a) \\ \partial_y f_2(a) \\ \partial_y f_3(a) \end{pmatrix}, \quad \frac{\partial f}{\partial z}(a) = \begin{pmatrix} \partial_z f_1(a) \\ \partial_z f_2(a) \\ \partial_z f_3(a) \end{pmatrix}. \tag{12.1387}$$

12.45 Limites à plusieurs variables

PropLimParcompos

Proposition 12.492.
Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$. Nous avons

$$\lim_{x\to a} f(x) = \ell \tag{12.1388}$$

si et seulement si

$$\lim_{x\to a} f_i(x) = \ell_i \tag{12.1389}$$

pour tout $i \in \{1,\dots,n\}$ où $f_i(x)$ dénote la i-ème composante de $f(x)$ et ℓ_i la i-ème composante de $\ell \in \mathbb{R}^n$.

Cette proposition revient à dire que la convergence d'une fonction est équivalente à la convergence de chacune de ses composantes.

Démonstration. L'élément clef de la preuve est le fait que pour tout vecteur $u \in \mathbb{R}^p$, nous ayons l'inégalité

$$|u_i| \leqslant \sqrt{\sum_{k=1}^{p} |u_k|^2} = \|u\|. \tag{12.1390}$$

Equilequnorme

La norme (dans $\mathbb{R}^p$) d'un vecteur est plus grande ou égale à la valeur absolue de chacune de ses composantes.

Supposons que nous ayons une fonction dont chacune des composantes a une limite en a : $\lim_{x\to a} f_i(x) = \ell_i$. Montrons que dans ce cas la fonction f tend vers ℓ. Si nous considérons $\varepsilon > 0$, par définition de la limite de chacune des fonctions f_i, il existent des δ_i tels que

$$\|x - a\|_{\mathbb{R}^m} < \delta_i \Rightarrow |f_i(x) - \ell_i| < \varepsilon. \tag{12.1391}$$

Notez que la norme à gauche est une norme dans $\mathbb{R}^m$ et que celle à droite est une simple valeur absolue dans $\mathbb{R}$. Considérons $\delta = \min\{\delta_i\}_{i=1,\dots n}$. Si $\|x - a\| < \delta$, alors

$$\|f(x) - \ell\| = \sqrt{\sum_{i=1}^{n} |f_i(x) - \ell_i|^2} < \sqrt{\sum_{i=1}^{n} \varepsilon^2} = \sqrt{n\varepsilon^2} = \sqrt{n}\varepsilon. \tag{12.1392}$$

Nous voyons qu'en choisissant les δ_i tels que $|f_i(x) - \ell_i| < \varepsilon$, nous trouvons $\|f(x) - \ell\| < \sqrt{n}\varepsilon$. Afin d'obtenir $\|f(x) - \ell\| < \varepsilon$, nous choisissons donc les δ_i de telle manière a avoir $|f_i(x) - \ell_i| < \varepsilon/\sqrt{n}$.

Nous avons donc prouvé que la limite composante par composante impliquait la limite de la fonction. Nous devons encore prouver le sens inverse.

Supposons donc que $\lim_{x\to a} f(x) = \ell$, et prouvons que nous ayons $\lim_{x\to a} f_i(x) = \ell_i$ pour chaque i. Soit $\varepsilon > 0$ et $\delta > 0$ tel que $\|x - a\| < \delta$ implique $\|f(x) - \ell\| < \varepsilon$. Avec ces choix, nous avons

$$|f_i(x) - \ell_i| \leqslant \|f(x) - \ell\| < \varepsilon \tag{12.1393}$$

où nous avons utilisé la majoration (12.1390) avec $f(x) - \ell$ en guise de u. □

De même, pour la continuité nous avons la proposition suivante :

Proposition 12.493.
Soit une fonction $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$ et $a \in D$. La fonction f est continue en a si et seulement si chacune de ses composantes l'est, c'est-à-dire si et seulement si chacune des fonctions $f_i\colon D \to \mathbb{R}$ est continue en a.

Essayez de prouver cette proposition directement par la définition de la continuité, en suivant pas à pas la démonstration de la proposition 12.492.

Propfaposfxposcont

Proposition 12.494.
Soit $f\colon \mathbb{R}^m \to \mathbb{R}$ et a, un point du domaine de f telle que $f(a) > 0$. Alors il existe un rayon r tel que $f(x) > 0$ pour tout x dans $B(a, r)$.

Cette proposition signifie que si la fonction est strictement positive en un point, alors elle restera strictement positive en tous les points « pas trop loin ».

Démonstration. Prenons $\varepsilon = f(a)/2$ dans la définition de la continuité. Il existe donc un rayon δ tel que pour tout x dans $B(a, \delta)$,

$$|f(x) - f(a)| \leqslant \frac{f(a)}{2}, \tag{12.1394}$$

en d'autres termes, $f(x) \in B\big(f(a), \frac{f(a)}{2}\big)$. évidemment aucun nombre négatif ne fait partie de cette dernière boule lorsque $f(a)$ est strictement positif. □

CorfneqzOuvert

Corolaire 12.495.
Si $f : \mathbb{R}^m \to \mathbb{R}$ est une fonction continue, alors l'ensemble

$$A = \{x \in \mathbb{R}^m \text{ tels que } f(x) \neq 0\} \tag{12.1395}$$

est ouvert.

Démonstration. Soit $x \in A$. Si $x > 0$ (le cas $x < 0$ est laissé en exercice), alors il existe une boule autour de x sur laquelle f reste strictement positive (proposition 12.494). Cette boule est donc contenue dans A. Étant donné qu'autour de chaque point de A nous pouvons trouver une boule contenue dans A, ce dernier est ouvert. □

ExBNOQEWe

Exemple 12.496.
Soit $GL_n(\mathbb{R})$ l'ensemble des matrices $n \times n$ inversibles. Nous allons montrer que $GL_n(\mathbb{R})$ est un ouvert de $\mathbb{R}^{n^2}$. L'identification entre les vecteurs et les matrices consiste simplement à « déplier » la matrice pour en faire un vecteur. Par exemple, en dimension deux,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \in \mathbb{R}^4. \tag{12.1396}$$

En dimension 3,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix} \in \mathbb{R}^9. \tag{12.1397}$$

Une matrice est inversible si et seulement si son déterminant est non nul. Or le déterminant est un polynôme en les composantes de la matrice. En dimension deux, nous avons

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc, \tag{12.1398}$$

mais en écriture « dépliée », nous pouvons aussi bien écrire

$$\det \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = ad - bc. \tag{12.1399}$$

En dimension 3, le déterminant est donc un polynôme des 9 variables qui apparaissent dans le vecteur « déplié ». En général, dans $\mathbb{R}^{n^2}$, nous considérons donc le polynôme $\det : \mathbb{R}^{n^2} \to \mathbb{R}$ qui

à un vecteur $X \in \mathbb{R}^{n^2}$ fait correspondre le déterminant de la matrice obtenue en « repliant » le vecteur X.

Donc dans $\mathbb{R}^{n^2}$, l'ensemble des matrices inversibles est donné par l'ensemble des vecteurs sur lesquels le polynôme det ne s'annule pas, c'est-à-dire

$$\{X \in \mathbb{R}^{n^2} \text{ tels que } \det(X) \neq 0\}. \tag{12.1400}$$

Mais le déterminant est un polynôme, et donc une fonction continue. Cet ensemble est par conséquence ouvert par le corolaire 12.495. △

La proposition suivante montre que la limite peut « passer à travers » les fonctions continues.

PropLimCompose

Proposition 12.497 (limite de fonction composée)**.**
Soit $f\colon \mathbb{R}^n \to \mathbb{R}^q$ *et* $g\colon \mathbb{R}^m \to \mathbb{R}^n$ *telles que*

$$\lim_{x\to a} g(x) = p \tag{12.1401a}$$
EqLimCompHypa

$$\lim_{y\to p} f(y) = q \tag{12.1401b}$$
EqLimCompHypb

Alors nous avons $\lim_{x\to a}(f\circ g)(x) = q$.

Démonstration. Comme presque toute preuve à propos de limite ou de continuité, nous commençons par choisir $\varepsilon > 0$. Nous devons montrer qu'il existe un δ tel que $\|x-a\| \leqslant \delta$ implique $\|f\big(g(x)\big) - q\| \leqslant \varepsilon$.

La limite (12.1401b) impose l'existence d'un $\tilde{\delta}$ tel que $\|y-p\| \leqslant \tilde{\delta}$ implique $\|f(y)-q\| \leqslant \varepsilon$, tandis que la limite (12.1401a) donne un δ tel que $\|x-a\| \leqslant \delta$ implique $\|g(x)-p\| \leqslant \tilde{\delta}$ (nous avons pris $\tilde{\delta}$ en guise de ε dans la définition de la limite pour g).

Avec ces choix, si $\|x-a\| \leqslant \delta$, alors $\|g(x)-p\| \leqslant \tilde{\delta}$, et par conséquent,

$$\|f\big(g(x)\big) - q\| \leqslant \varepsilon, \tag{12.1402}$$

ce que nous voulions. □

De façon pragmatique, la proposition 12.497 nous fournit une formule pour les limites de fonctions composée :

$$\lim_{x\to a}(f\circ g)(x) = \lim_{y\to \lim_{x\to a} g(x)} f(y) \tag{12.1403}$$
Eqlimfgvomp

lorsque f est continue.

Remarque 12.498.
La formule (12.1403) ne peut pas être utilisée à l'envers. Il existe des cas où $\lim_{x\to a}(g\circ f)(x) = q$, et $\lim_{x\to a} f(x) = p$ sans pour autant avoir $\lim_{y\to q} g(y) = q$. Par exemple

$$g(x) = \begin{cases} 2 & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{12.1404a}$$

$$f(x) = |x|. \tag{12.1404b}$$

Nous avons $(g\circ f)(x) = 2$ pour tout x, ainsi que $\lim_{x\to 0} f(x) = 0$, mais la limite $\lim_{y\to 0} g(y)$ n'existe pas.

En général, lorsqu'un ensemble est donné par des inégalités, prendre la fermeture consiste à transformer les inégalités strictes en inégalités non strictes ; prendre l'intérieur consiste à rendre stricte toutes les inégalités ; la frontière consiste *en gros* à transformer toutes les inégalités en égalités (nous allons voir que pour la frontière, c'est un peu plus de travail). Comprenez bien que cela n'est vrai que « en général ». Il faut toujours bien regarder sur chaque exemple si il n'y a pas l'un ou l'autre point problématique.

La proposition 12.494 sera une des clefs pour dire que si une inégalité stricte est satisfaite en un point, alors elle sera satisfaite en tout point dans un voisinage. Voir aussi l'exemple 12.499.

exoEspVectoNorme0008

Exemple 12.499.
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. Montrer que le sous-ensemble A de $\mathbb{R}^2$ défini par $A = \{(x, y) \in \mathbb{R}^2;\ y > f(x)\}$ est ouvert.

Prouver que l'ensemble $B = \{(x, y) \in \mathbb{R}^2$ tels que $y \geqslant f(x)\}$ est fermé.

Nous considérons la fonction $g : \mathbb{R}^2 \to \mathbb{R}$ donnée par $g(x, y) = f(x) - y$. Cela est une fonction continue parce que c'est une différence de fonctions continues. Par définition,

$$A = \{(x, y) \in \mathbb{R}^2 \text{ tel que } g(x, y) > 0\}. \tag{12.1405}$$

Par la proposition 12.494, autour de chaque point (x, y) tel que $g(x, y) > 0$ (c'est-à-dire autour de chaque point de A), il existe une boule sur laquelle g reste strictement positive. L'ensemble A est donc ouvert.

Pour prouver que l'ensemble B est fermé, prouver que le complémentaire est ouvert, c'est-à-dire que les points tels que $y - f(x) < 0$ forment un ouvert. Cela revient au même que ce que nous avons fait pour A.

△

12.46 Champs de vecteurs

Un champ de vecteur est une fonction $f : \mathbb{R}^3 \to \mathbb{R}^3$. Géométriquement, il s'agit simplement de mettre un vecteur en chaque point de l'espace. Cela arrive très souvent en physique.

Exemple 12.500.
Si un fluide (eau, gaz) coule dans un tube, en tout point le point a une vitesse, qui sera un vecteur généralement dirigé le long du tube. △

Exemple 12.501.
La force d'attraction de la Terre sur une masse m située au point $r = (x, y, z)$ est donnée par

$$F(r) = -G\frac{Mmr}{\|r\|^3}. \tag{12.1406}$$

Dans cette expression, tant r que $F(r)$ sont des vecteurs. Nous l'avons représenté sur la figure 12.13.

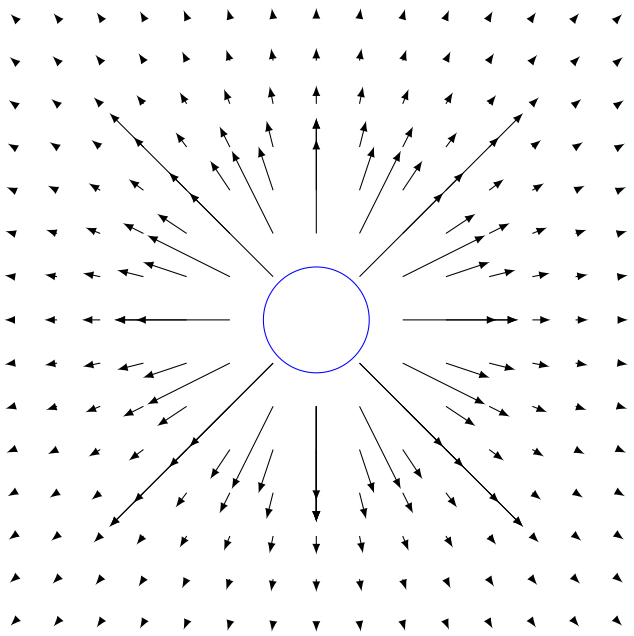


FIGURE 12.13: Le champ de gravitation de la Terre. LabelFigSQNPooPTrLRQ

L'application

$$\begin{aligned} F\colon \mathbb{R}^3 &\to \mathbb{R}^3 \\ r &\mapsto F(r) \end{aligned} \tag{12.1407}$$

est le champ gravitationnel de la Terre.

△

12.46.1 Matrice jacobienne

La **matrice jacobienne** de la fonction $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ au point $a \in \mathbb{R}^3$ est la matrice dont les colonnes sont les vecteurs $\frac{\partial f}{\partial x}(a)$, $\frac{\partial f}{\partial y}(a)$ et $\frac{\partial f}{\partial z}(a)$, c'est-à-dire

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(a) & \frac{\partial f_1}{\partial y}(a) & \frac{\partial f_1}{\partial z}(a) \\ \frac{\partial f_2}{\partial x}(a) & \frac{\partial f_2}{\partial y}(a) & \frac{\partial f_2}{\partial z}(a) \\ \frac{\partial f_3}{\partial x}(a) & \frac{\partial f_3}{\partial y}(a) & \frac{\partial f_3}{\partial z}(a) \end{pmatrix}. \tag{12.1408}$$

Exemple 12.502.
Si

$$f(x,y,z) = \begin{pmatrix} xye^z \\ x^2 + \cos(yz) \\ xyz \end{pmatrix}, \tag{12.1409}$$

alors

$$J_f(x,y,z) = \begin{pmatrix} ye^z & xe^z & xye^z \\ 2x & -z\sin(yz) & -y\sin(yz) \\ yz & xz & xy \end{pmatrix}. \tag{12.1410}$$

△

12.47 Divergence, rotationnel et l'opérateur nabla

Nous avons déjà vu le gradient d'une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$

$$\nabla f(x,y,z) = \begin{pmatrix} \partial_x f(x,y,z) \\ \partial_y f(x,y,z) \\ \partial_z f(x,y,z) \end{pmatrix} \tag{12.1411}$$

EqDefNablaf

Afin de définir la divergence et le rotationnel, nous introduisons ∇ sous une forme un peu plus abstraite comme le « vecteur »

$$\nabla = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}. \tag{12.1412}$$

Vue comme ça, la formule (12.1411) est claire.

Si F est un champ de vecteurs, nous introduisons la **divergence** de F par

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \tag{12.1413}$$

Cela est une fonction. Et nous introduisons le rotationnel du champ de vecteur F par

$$\begin{aligned} \nabla \times F &= \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix} \\ &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) e_x - \left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right) e_y + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) e_z. \end{aligned} \tag{12.1414}$$

Cela est un champ de vecteur. En utilisant le symbole complètement antisymétrique ϵ_{ijk}, le rotationnel d'un champ de vecteur peut s'écrire

$$\nabla \times F = \sum_{ijk} \epsilon_{ijk} \partial_i F_j e_k. \tag{12.1415}$$

Le gradient, la divergence et le rotationnel consistent à appliquer simplement à ∇ les trois produits qu'on peut effectuer sur un vecteur :

(1) Le produit d'un vecteur par un scalaire multiplie chacune des composantes :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} f = \begin{pmatrix} \partial_x f \\ \partial_y f \\ \partial_z f \end{pmatrix}. \tag{12.1416}$$

(2) Le produit scalaire d'un vecteur avec un autre vecteur donne lieu à la divergence :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \cdot \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \tag{12.1417}$$

(3) Le produit vectoriel de deux vecteurs :

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \times \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix}. \tag{12.1418}$$

Ces trois opérations joueront un rôle central en électromagnétisme dans les équations de Maxwell.

Exemple 12.503.
Soit $F(x, y, z) = xe_x + xye_y + e_z$, c'est-à-dire

$$F(x, y, z) = \begin{pmatrix} x \\ xy \\ 1 \end{pmatrix}. \tag{12.1419}$$

Son rotationnel est donné par

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial y} \\ x & xy & 1 \end{vmatrix} = (0-0)e_x - (0-0)e_y + (y-0)e_z = ye_z = \begin{pmatrix} 0 \\ 0 \\ y \end{pmatrix}. \tag{12.1420}$$

△

Afin d'étudier comment se comporte la composition de ces opérateurs, nous aurons besoin de ce lemme que nous n'énoncerons pas précisément.

LemPermDerrxyz

Lemme 12.504.
Si $f : \mathbb{R}^3 \to \mathbb{R}$ est une fonction de classe C^2, alors on peut permuter l'ordre des dérivées :

$$\begin{aligned} \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \\ \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) &= \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right) \\ \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) \end{aligned} \tag{12.1421}$$

La fonction

$$(x,y,z) \mapsto \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y,z) \tag{12.1422}$$

sera notée

$$\frac{\partial^2 f}{\partial x \partial y}. \tag{12.1423}$$

Il y a deux propriétés importantes :

Théorème 12.505.
Soit $f\colon \mathbb{R}^3 \to \mathbb{R}$ une fonction de classe C^2. Alors

$$\nabla \times (\nabla f) = 0. \tag{12.1424}$$

Si $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ est un champ de vecteurs de classe C^2, alors

$$\nabla \cdot (\nabla \times F) = 0. \tag{12.1425}$$

Démonstration. Ce sont seulement deux calculs qui manipulent les définitions. Pour le premier, la divergence de f est le champ de vecteurs

$$\nabla f = \frac{\partial f}{\partial x}e_x + \frac{\partial f}{\partial y}e_y + \frac{\partial f}{\partial z}e_z. \tag{12.1426}$$

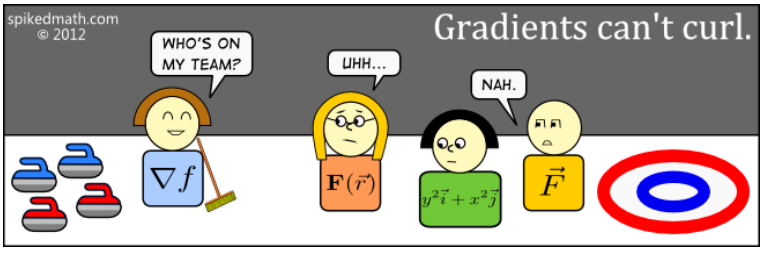
En mettant ce champ dans la définition du rotationnel,

$$\begin{aligned}\nabla \times (\nabla f) = \begin{vmatrix} e_x & e_y & e_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} &= \left[\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right)\right]e_x \\ &\quad - \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right)\right]e_y \\ &\quad + \left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right]e_z.\end{aligned} \tag{12.1427}$$

En utilisant le lemme 12.504, chacun des termes fait zéro.

La seconde propriété se démontre en utilisant le même type de calcul. □

Remarque 12.506.
Il n'y a pas de propriétés du même style pour la combinaison $\nabla \times (\nabla \cdot F)$ pour le rotationnel de la divergence. En effet la divergence d'un champ de vecteur est une fonction, et il n'y a pas de rotationnel pour une fonction.



http://spikedmath.com/501.htmlSpiked math, licence Creative Commons by-nc 2.5.

12.48 Interprétation géométrique et physique de la divergence

En physique, on dit qu'un champ de vecteurs à divergence nulle est **incompressible**. Nous allons essayer de comprendre pourquoi. Lorsqu'un fluide incompressible se déplace, il faut qu'en chaque point il y ait autant de fluide qui rentre que de fluide qui sort. Nous allons voir sur quelques

exemples que la divergence d'un champ de vecteurs est le « bilan de masse » d'un fluide qui se déplace selon le champ de vecteurs.

Si en un point la divergence est positive, cela signifie qu'il y a une perte de masse et si la divergence est négative, cela signifie qu'il y a une accumulation de masse.

Prenons par exemple un fluide qui se déplace selon le champ de vitesse montré à figure 12.14.

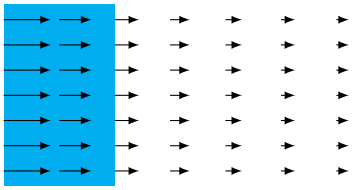


FIGURE 12.14: Le champ de vecteurs $F(x,y) = \frac{1}{x}(1,0)$. LabelFigBEHTooWsdrys

Étant donné que la vitesse diminue lorsque x avance, il y a une accumulation de fluide. Regardez en effet la quantité de fluide qui rentre dans le rectangle par rapport à la quantité de fluide qui en sort. Ce champ de vecteurs a pour équation :

$$F(x,y) = \frac{1}{x}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1/x\\0\end{pmatrix}. \tag{12.1428}$$

Sa divergence vaut donc

$$(\nabla \cdot F)(x,y) = \frac{\partial F_x}{\partial x}(x,y) + \underbrace{\frac{\partial F_y}{\partial y}(x,y)}_{=0} = -\frac{1}{x^2}. \tag{12.1429}$$

Cette divergence étant négative, il y a bien accumulation de fluide en tout point, et d'autant plus que x est petit.

ExamDivFrot

Exemple 12.507.
Prenons le champ de vecteurs tournant

$$F(x,y) = \frac{1}{\sqrt{x^2+y^2}}\begin{pmatrix}y\\-x\end{pmatrix} \tag{12.1430}$$

représenté à la figure 12.15. Cela est un vecteur qui est constamment perpendiculaire au rayon.

Un fluide dont la vitesse serait donné par ce champ de vecteur se contente de tourner. Intuitivement il ne devrait pas y avoir de divergence parce qu'il n'y a aucune accumulation de fluide. En effet,

$$\nabla \cdot F(x,y) = \frac{-2xy}{(x^2+y^2)^2} + \frac{2xy}{(x^2+y^2)^2} = 0. \tag{12.1431}$$

△

Exemple 12.508.
Prenons le cas du champ de force de gravitation :

$$F(x,y,z) = \frac{1}{(x^2+y^2+z^2)^{3/2}}\begin{pmatrix}x\\y\\z\end{pmatrix}. \tag{12.1432}$$

Nous pouvons rapidement remarquer que $\nabla \cdot F = 0$. Est-ce que cela peut se comprendre sur le dessin de la figure 12.16 ?

Essayons de voir combien de fluide entre dans la zone bleue et combien en sort. D'abord, il est certain que les vecteurs qui sortent sont plus courts que ceux qui rentrent, ce qui voudrait dire qu'il y a plus de fluide qui rentre. Mais on voit également que le *nombre* de vecteurs qui sortent

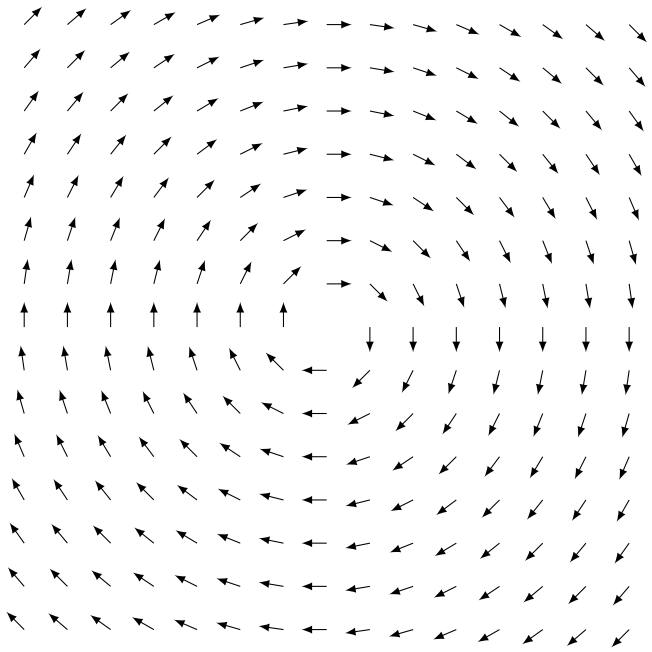


FIGURE 12.15: Le champ de vecteurs $F(x,y)=(y,-x)$. LabelFigYQVHooYsGLHQ

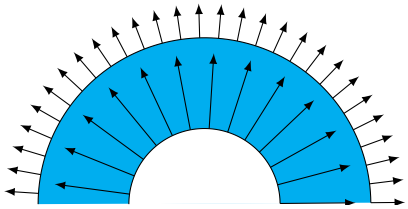


FIGURE 12.16: Le champ de vecteur de la gravité. Nous avons tracé, sur les deux cercles la même densité de vecteurs, c'est-à-dire le même nombre de vecteurs par unité de surface. LabelFigZGUDooEsqCWQ

est plus grand parce que la seconde sphère est plus grande et qu'il y a un vecteur en chaque point de la sphère.

Intuitivement nous pouvons dire que la quantité qui rentre dans la sphère de rayon r_1 donnée par la taille des vecteurs entrants multiplié par la surface de la sphère, c'est-à-dire

$$4\pi r_1^2\|F(x,y,z)\|, \tag{12.1433}$$ EqQpinormeVecto

mais $\|F(x,y,z)\|=\frac{1}{r_1^2}$, donc la quantité de fluide entrant est 4π. La quantité de fluide sortant sera la même.

Cela explique deux choses

(1) Pourquoi les forces de gravitation et électromagnétiques sont en $1/r^2$; c'est parce que nous vivons dans un monde avec trois dimensions d'espace. En étudiant très précisément le champ de gravitation, certains physiciens espèrent trouver des déviations expérimentales par rapport à la règle du $1/r^2$; cela *pourrait* être un signe que l'espace contient des dimensions supplémentaires.

(2) Pourquoi il y a un 4π comme coefficient dans beaucoup d'équations en électromagnétisme ; en particulier dans certaines anciennes unités de flux.

△

Remarque 12.509.
Nous allons voir plus loin comment s'assurer que l'équation (12.1433) représente bien la « quantité de fluide » qui rentre dans la zone délimitée

12.49 Quelques formules de Leibniz

La divergence étant une combinaison de dérivées, il n'est pas tellement étonnant que la divergence de produits donne lieux à des formules en deux termes. Si f est une fonction et si F et G sont des champs de vecteurs, nous avons la proposition suivante.

PROPooDMWEooNaJBCM

Proposition 12.510.
Si F et G sont des champs de vecteurs dont toutes les dérivées partielles existent, alors

ITEMooFDJIooKTnvKj

(1) $\nabla \cdot (fF) = f\nabla \cdot F + F \cdot \nabla f$

(2) $\nabla \cdot (F \times G) = G \cdot \nabla \times F - F \cdot \nabla \times G$

(3) $\nabla \times (fF) = f\nabla \times F + \nabla f \times F$.

Chapitre 13

Analyse sur des groupes

13.1 Action de groupe et connexité

ThojrLKZk

Théorème 13.1.
Soit G un groupe topologique localement compact et dénombrable à l'infini[1] *agissant continûment et transitivement sur un espace topologique localement compact*[2] *E. Soit $x \in E$. Alors l'application*

$$\begin{aligned} \varphi\colon G/\operatorname{Stab}(x) &\to E \\ [g] &\mapsto g \cdot x \end{aligned} \tag{13.1}$$

où $\operatorname{Stab}(x)$ est le stabilisateur[3] *de x est un homéomorphisme.*

LemkLRAet

Lemme 13.2.
Si G et H sont des groupes topologiques tels que G/H et H sont connexes[4]*, alors G est connexe.*

Démonstration. Soit $f\colon G \to \{0,1\}$ une fonction continue. Considérons l'application

$$\begin{aligned} \tilde{f}\colon G/H &\to \{0,1\} \\ [g] &\mapsto f(g). \end{aligned} \tag{13.2}$$

D'abord nous montrons qu'elle est bien définie. En effet si $h \in H$ nous aurions $\tilde{f}([gh]) = f(gh)$, mais étant donné que H est connexe, l'ensemble gH est également connexe ; la fonction continue f est donc constante sur gH. Nous avons donc $f(gh) = f(g)$.

Étant donné que G/H est également connexe, la fonction $\tilde{f}$ doit être constante. Si g_1 et g_2 sont deux éléments du groupe, nous avons $f(g_1) = \tilde{f}([g_1]) = \tilde{f}([g_2]) = f(g_2)$. Nous en déduisons que f est constante et que G est connexe. □

13.3.
La connexité de $\mathrm{SO}(3)$ peut être démontrée en suivant les lignes de [422]. Le corolaire 12.91 permet de dire que les éléments de $\mathrm{SO}(3)$ ont une valeur propre égale à 1.

THOooYQFNooPaYmaP

Théorème 13.4.
Pour tout $n \geqslant 2$, le groupe $\mathrm{SO}(n)$ est connexe, le groupe $\mathrm{O}(n)$ a deux composantes connexes.

Démonstration. La seconde assertion découle de la première parce que les matrices de déterminant 1 et celles de déterminant -1 ne peuvent pas être reliées par un chemin continu tandis que l'application

$$M \mapsto \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix} M \tag{13.3}$$

1. Cela signifie qu'il est une réunion dénombrable de compacts
2. Définition 7.86.
3. Définition 2.28.
4. Définition 7.70.

est un homéomorphisme entre les matrices de déterminant 1 et celles de déterminants -1. Montrons donc que $G = \mathrm{SO}(n)$ est connexe par arcs pour $n \geqslant 2$ en procédant par récurrence sur la dimension.

Nous acceptons le résultat pour $G = \mathrm{SO}(2)$. Notons que nous en avons besoin pour prouver que la sphère S^{n-1} est connexe.

Le groupe $\mathrm{SO}(n)$ agit, par définition, de façon transitive sur la sphère S^{n-1}. Soit $a \in S^{n-1}$, nous avons

$$G \cdot a = S^{n-1} \tag{13.4a}$$
$$G_a \simeq \mathrm{SO}(n-1) \tag{13.4b}$$

où G_a est le fixateur de a dans G. Pour montrer le second point, nous considérons $\{e_i\}$, la base canonique de $\mathbb{R}^n$ et $M \in G$ telle que $Ma = e_1$. Le fixateur de e_1 est évidemment isomorphe à $\mathrm{SO}(n-1)$ parce qu'il est constitué des matrices de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{11} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix} \tag{13.5}$$

où $(a_{ij}) \in \mathrm{SO}(n-1)$. L'application

$$\begin{aligned} \alpha\colon G_{e_1} &\to G_a \\ A &\mapsto M^{-1}AM \end{aligned} \tag{13.6}$$

est un isomorphisme entre G_a et $\mathrm{SO}(n-1)$. Le théorème 13.1 nous montre alors que, en tant qu'espaces topologiques,

$$G/G_a = S^{n-1}. \tag{13.7}$$

L'hypothèse de récurrence montre que $G_a = \mathrm{SO}(n-1)$ est connexe tandis que nous savons que S^{n-1} est connexe. Le lemme 13.2 conclut que $G = \mathrm{SO}(n)$ est connexe. □

LemIbrsFT

Lemme 13.5.
Une bijection continue entre un espace compact et un espace séparé est un homéomorphisme.

PROPooTVHJooBRmUCd

Proposition 13.6.
Les groupes $\mathrm{U}(n)$ *et* $\mathrm{SU}(n)$ *sont connexes.*

Démonstration. Soit $G(n)$ le groupe $\mathrm{SU}(n)$ ou $\mathrm{U}(n)$. Ce groupe opère transitivement sur la sphère complexe

$$S^{n-1}_{\mathbb{C}} = \{z \in \mathbb{C}^n \text{ tel que } |z| = 1\}. \tag{13.8}$$

En notons, pour chaque k, $z_k = x_k + iy_k$, nous considérons l'application continue

$$\begin{aligned} s\colon S^{n-1}_{\mathbb{C}} &\to S^{2n-1} \\ (z_1, \dots, z_n) &\mapsto (x_1, y_1, \dots, x_n, y_n). \end{aligned} \tag{13.9}$$

Vérifions que s prend bien ses valeurs dans S^{2n-1}. Nous avons

$$|s(z_1, \dots, z_n)| = \sqrt{\sum_k |x_k|^2 + |y_k|^2} = \sqrt{\sum_k |z_k|^2} = 1 \tag{13.10}$$

parce que $|z_k|^2 = |x_k|^2 + |y_k|^2$. Le lemme 13.5 dit que cette bijection continue est un homéomorphisme. Soit $a \in S^{n-1}_{\mathbb{C}}$, nous avons

$$G \cdot a = S^{n-1}_{\mathbb{C}} \tag{13.11a}$$
$$G_a \simeq G(n-1). \tag{13.11b}$$

La seconde ligne est un isomorphisme de groupe et un homéomorphisme. Il est donné de la façon suivante. D'abord le fixateur de e_1 dans $G(n)$ est donné par les matrices de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{11} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix} \tag{13.12}$$

où $(a_{ij}) \in G(n-1)$. Par ailleurs si M est une matrice de $G(n)$ telle que $Ma = e_1$, nous avons l'homéomorphisme

$$\begin{aligned} \alpha\colon G_{e_1} &\to G_a \\ A &\mapsto M^{-1}AM. \end{aligned} \tag{13.13}$$

Encore une fois, cela est un homéomorphisme par le lemme 13.5. Par composition nous avons $G_a \simeq G(n-1)$ et un homéomorphisme

$$G(n)/G_a = S_{\mathbb{C}}^{n-1}. \tag{13.14}$$

Le groupe G_a et l'ensemble $S_{\mathbb{C}}^{n-1}$ étant connexes, le groupe $G(n)$ est connexe par le lemme 13.2. □

Lemme 13.7 ([380])**.**
Si G est un sous-groupe connexe de $\mathrm{GL}(n,\mathbb{C})$ alors son groupe dérivé[5] *l'est également.*

Démonstration. Soit S_m l'ensemble des produits de m commutateurs de G :

$$S_m = \{g_1, \dots, g_m \text{ où les } g_i \text{ sont des commutateurs}\}. \tag{13.15}$$

La partie S_m est l'image de G par l'application continue

$$\begin{aligned} \underbrace{G \times \dots \times G}_{2m \text{ facteurs}} &\to G \\ (g_1, h_1, g_2, h_2, \dots, g_m, h_m) &\mapsto [g_1, h_1] \dots [g_m, h_m] \end{aligned} \tag{13.16}$$

En tant qu'image d'un connexe par une application continue, S_m est connexe par le lemme 7.218. Puisque les S_m ont l'identité en commun, le groupe dérivé

$$D(G) = \bigcup_{m=1}^{\infty} S_m \tag{13.17}$$

est également connexe. □

13.2 Espaces de matrices

13.2.1 Dilatations et transvections

Soit un corps commutatif $\mathbb{K}$ et $n \geqslant 2$.

ThoooAZKDooNDcznv

Théorème-Définition 13.8 ([380])**.**
Soit une application linéaire $u\colon E \to E$ dont les points fixes forment un hyperplan noté H d'équation $H = \ker(f)$ avec $f \in E^*$.

ITEMooGTKRooQSPNoI

(1) Les affirmations suivantes sont équivalentes :

ITEMooZHYRooFGKaifi

(1a) $\det(u) \neq 1$

ooXKLWooTfUMzV

(1b) L'application u est diagonalisable et a une valeur propre qui vaut $\det(u) \neq 1$.

5. Définition 2.2.

ooMZPTooCLylbh

(1c) $\text{Image}(u - \text{Id}) \not\subseteq H$.

ITEMooZHYRooFGKaifiv

(1d) Il existe une base de E dans laquelle la matrice de u est $\text{diag}(1, \dots, 1, \lambda)$ *avec* $\lambda \neq 1$.

ITEMooMSJXooUsLCHx

(2) Les affirmation suivantes sont équivalentes :

ITEMooRTIEooOoWCFsa

(2a) Il existe $a \in H$ tel que pour tout $x \in E$, $u(x) = x + f(x)a$.

ITEMooRTIEooOoWCFsb

(2b) Dans une base adaptée, la matrice de u est donnée par

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}. \tag{13.18}$$

EQooFXBDooTgZwMv

(3) Les conditions (1a)-(1d) sont respectées si et seulement si les conditions (2a)-(2b) ne sont pas respectées (elles sont les négations l'une de l'autre.).

*Une **dilatation** est soit l'identité soit un endomorphisme qui respecte les conditions (1).*
*Une **transvection** est soit l'identité soit un endomorphisme qui vérifie les conditions (2).*

Démonstration. Nous allons prouver plein d'implications ...

(1) **(1a) implique (1b)** Le théorème de la base incomplète (voir remarque 4.15) permet de considérer une base $\{e_1, \dots, e_n\}$ de E telle que $\{e_1, \dots, e_{n-1}\}$ soit une base de H. Dans cette base, la matrice de u est de la forme suivante (les cases non remplies sont nulles et les étoiles correspondent à des valeurs inconnues mais pas spécialement nulles) :

$$\begin{pmatrix} 1 & & & * \\ & \ddots & & \vdots \\ & & 1 & * \\ & & & \lambda \end{pmatrix} \tag{13.19}$$

EqooPQOEooGUyIwa

Le fait que le déterminant de u ne soit pas 1 implique que $\lambda \neq 1$. Par conséquent le polynôme caractéristique

$$\chi_u(X) = (1 - X)^{n-1}(\lambda - X) \tag{13.20}$$

possède une racine $\lambda \neq 1$, et donc u possède un vecteur propre v pour cette valeur [6]. Le vecteur v est linéairement indépendant de $\{e_1, \dots, e_{n-1}\}$ (parce que vecteur propre de valeur propre différente). Par conséquent l'ensemble $\{e_1, \dots, e_{n-1}, v\}$ est une base par la proposition 4.19. C'est une base de vecteurs propres et donc une base de diagonalisation [7].

(2) **(1b) implique (1c)** Nous nommons maintenant $\{e_1, \dots, e_n\}$ la base de diagonalisation. Nous avons $u(e_n) = \lambda e_n$ avec $\det(u) = \lambda \neq 1$. Nous avons

$$(u - \text{Id})(e_n) = (\lambda - 1)e_n \notin H, \tag{13.21}$$

ce qui prouve que l'image de e_n par $u - \text{Id}$ n'est pas dans H.

(3) **(1c) implique (1d)** Reprenons une base $\{e_1, \dots, , e_n\}$ donnant la matrice (13.19). Il existe $x \in E$ tel que $u(x) - x$ n'est pas dans H, c'est-à-dire tel que $u\big(u(x) - x\big) \neq u(x) - x$. Nous en déduisons que

$$u^2(x) - 2u(x) + x \neq 0 \tag{13.22}$$

ou encore que

$$(X - 1)^2(u)x \neq 0. \tag{13.23}$$

6. Proposition 8.123.

7. Nous pourrions en dire à peine un peu plus et prouver le point (1d), mais cela ne servirait à rien parce que nous voulons prouver les équivalences et qu'il faudra quand même prouver que (1c) implique (1d).

C'est-à-dire que $(X-1)^2$ n'est pas un polynôme annulateur de u. Or ce serait le cas si $X-1$ était le polynôme minimal (proposition 8.102). Le polynôme caractéristique étant $(X-1)^{n-1}(X-\lambda)$ (et étant annulateur[8]), le polynôme minimal est de la forme

$$\mu_u(X) = \begin{cases} (X-1)(X-\lambda) & \text{si } \lambda \neq 1 \\ X-1 & \text{si } \lambda = 1. \end{cases} \tag{13.24}$$

Dans notre cas nous venons de voir que ce n'est pas $X-1$ et donc c'est $(X-1)(X-\lambda)$ avec $\lambda \neq 1$.

Nous devons trouver une base de diagonalisation ...Supposons

$$u(e_n) = \sum_{k=1}^{n-1} a_k e_k + \lambda e_n, \tag{13.25}$$

dans lequel nous venons de prouver que $\lambda \neq 1$, et cherchons

$$e'_n = \sum_{j=1}^{n} p_j e_j \tag{13.26}$$

de telle sorte à avoir $u(e'_n) = \lambda e'_n$. Nous avons

$$u(e'_n) = \sum_{j=1}^{n-1} p_j u(e_j) + p_n u(e_n) = \sum_{j=1}^{n-1} (p_j + p_n a_j) e_j + p_n \lambda e_n. \tag{13.27}$$

En égalisant à $\lambda \sum_{j=1}^{n} p_j e_j$, il vient

$$p_j + p_n a_j = \lambda p_j \tag{13.28}$$

pour tout $j = 1, \ldots, n-1$ et la condition triviale $p_n \lambda = \lambda p_n$ pour $j = n$. Nous en déduisons que le choix

$$p_j = \frac{p_n a_j}{\lambda - 1} \tag{13.29}$$

fonctionne (parce que $\lambda \neq 1$ comme nous l'avons démontré plus haut). En bref, il suffit de poser

$$e'_n = \sum_{j=1}^{n-1} \frac{p_n a_j}{\lambda - 1} e_j + p_n e_n \tag{13.30}$$

avec p_n au choix pour avoir une base $\{e_1, \ldots, e_{n-1}, e'_n\}$ de diagonalisation de u avec $\lambda \neq 1$ comme dernière valeur propre.

(4) **(1d) implique (1a)** Évident ...encore faut-il se souvenir d'invoquer l'invariance du déterminant par changement de base.

Nous avons terminé la première série d'équivalences. Nous continuons avec la seconde.

(1) **(2a) implique (2b)** Nous prenons $e_{n-1} = a$ et nous complétons en une base de H. Pour e_n il suffit de prendre n'importe quel vecteur v tel que $f(v) \neq 0$ (qui existe parce que $f = 0$ est seulement un hyperplan), et de le normaliser.

Dans cette base, la matrice de u a la forme désirée parce que $u(e_n) = e_n + f(e_n)a = e_n + e_{n-1}$ du fait que $e_{n-1} = a$ et $f(e_n) = 1$.

(2) **(2b) implique (2a)** Soit $\{e_1, \ldots, e_n\}$ cette base. En prenant $a = e_{n-1}$ et en posant $x = \sum_k x_k e_k$ nous avons

$$u(x) = \sum_{k=1}^{n-1} x_k e_k + x_n(e_{n-1} + e_n) = x + x_n e_{n-1} = x + x_n a. \tag{13.31}$$

8. Théorème de Cayley-Hamilton 8.120.

Mais puisque $f(x) = \sum_i f_i x_i$, et que $f(e_i) = 0$ pour tout $i = 1, \dots, n-1$ nous avons $f(x) = f_n x_n$. Il n'y a cependant pas de raison d'avoir $f_n = 1$. Mais en définissant

$$e'_i = \frac{1}{f_n} e_i \tag{13.32}$$

nous avons bien $u(e'_n) = \frac{1}{f_n}(e_{n-1} + e_n) = e'_{n-1} + e'_n$. Donc dans cette base nous avons encore la matrice de u de la forme

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}, \tag{13.33}$$

mais cette fois avec $f(e'_n) = 1$.

Nous avons terminé avec la seconde série d'équivalences. Il nous reste à prouver que la première est équivalente à la négation de la seconde.

(1) **non (1c) implique (2a)** Considérons $x_0 \in E$ tel que $f(x_0) = 1$ et posons $a = u(x_0) - x_0 \in \text{Image}(u - \text{Id})$. Par la négation de (1c) nous avons $a \in H$. De plus $x_0 \notin H$ (sinon $f(x_0) = 0$) donc $u(x_0) \neq x_0$ et $a \neq 0$.

Nous montrons que ce choix de a fonctionne : $u(x) = x + f(x)a$ pour tout $x \in E$. Nous faisons cela séparément pour $x \in H$ et pour $x = x_0$.

Si $h \in H$ alors $u(h) = h$ et $f(h) = 0$ donc $h + f(h)a = h = u(h)$. Si $x = x_0$ alors $u(x_0) = a + x_0$ (c'est la définition de a) et $x_0 + f(x_0)a = x_0 + a$.

(2) **(2b) implique non (1a)** Dans une base adaptée nous avons

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}, \tag{13.34}$$

et donc $\det(u) = 1$, ce qui contredit (1a).

□

13.9.
Dans le cas des dilatations et des transvections, les points fixes forment un hyperplan.

Selon cette terminologie, l'application $x \mapsto \lambda x$ n'est pas une dilatation mais un produit de dilatations.

Remarque 13.10.
Nous notons E_{ij} la matrice qui possède uniquement 1 en position (i, j). C'est-à-dire que $\big(E_{ij}\big)_{kl} = \delta_{ik}\delta_{jl}$. Soit H l'hyperplan des points fixes de f. Dans une base contenant une base de H, la matrice d'une transvection a pour forme type :

$$T_{ij}(\lambda) = \mathbb{1} + \lambda E_{ij} \tag{13.35}$$

EqooZAKHooBiKlTd

avec $i \neq j$ et $\lambda \in \mathbb{K}$, et une dilatation a pour forme type la matrice diagonale

$$D_i(\alpha) = \mathbb{1} + (\alpha - 1)E_{ii} \tag{13.36}$$

avec $\alpha \in \mathbb{K}^*$.

Bien entendu, en choisissant une base quelconque, les matrices des dilatations et des translations peuvent avoir des formes différentes.

LemooTQJXooGoIxsI

Lemme 13.11.
Quelques manipulations de lignes et de colonnes pour les matrices.

ITEMooRWANooPAVjkm

(1) *La multiplication à gauche par $T_{ij}(\lambda)$ revient à effectuer le remplacement de ligne*

$$L_i \to L_i + \lambda L_j. \tag{13.37}$$

ITEMooHPSMooWBrSXP

(2) *La multiplication à droite par $T_{ij}(\lambda)$ revient à effectuer le remplacement de colonne*

$$C_j \to C_j + \lambda C_i. \tag{13.38}$$

ITEMooXUGFooKcbrxs

(3) *La multiplication à gauche par $T_{ij}(1)T_{ji}(-1)T_{ij}(1)$ revient à la substitution de lignes*

$$\begin{cases} L_i \to L_j & \text{(13.39a)} \\ L_j \to -L_i. & \text{(13.39b)} \end{cases}$$

Notons qu'il n'est pas possible d'inverser deux lignes à l'aide de transvections sans changer un signe parce que les transvections sont de déterminant 1 alors que l'inversion de lignes change le signe du déterminant.

Démonstration. Point par point.

(1) **Pour (1)** Nous devons prouver que

$$\big(T_{ij}(\lambda)A\big)_{kl} = \begin{cases} A_{kl} & \text{si } k \neq i \\ A_{il} + \lambda A_{jl} & \text{si } k = i. \end{cases} \tag{13.40}$$

Un peu de calcul matriciel avec utilisation modérée des indices donne :

$$\begin{aligned} \big(T_{ij}(\lambda)A\big)_{kl} &= \sum_s \big(T_{ij}(\lambda)\big)_{ks} A_{sl} && \text{(13.41a)} \\ &= \sum_s \delta_{ks}A_{sl} + \lambda\delta_{ik}\delta_{js}A_{sl} && \text{(13.41b)} \\ &= A_{kl} + \lambda\delta_{ik}A_{jl}. && \text{(13.41c)} \end{aligned}$$

(2) **Pour (2)** C'est la même chose.

(3) **Pour (3)** Si nous appliquons successivement ces trois matrices (de droite à gauche) nous effectuons les substitutions :

$$\begin{cases} L'_i = L_i + L_j \\ L'_j = L_j \end{cases} \text{ suivi de } \begin{cases} L''_i = L'_i \\ L''_j = L'_j - L'_i \end{cases} \text{ et de } \begin{cases} L'''_i = L''_i + L''_j \\ L'''_j = L''_j. \end{cases} \tag{13.42}$$

En effectuant ces substitutions,

$$L'''_i = L''_i + L''_j = L'_i + (L'_j - L'_i) = L'_j = L_j \tag{13.43}$$

et

$$L'''_j = L''_j = L'_j - L'_i = L_j - (L_i + L_j) = -L_i, \tag{13.44}$$

ce qu'il fallait.

□

PropooFDNRooWFfUDd

Proposition 13.12 ([423]).
Soient $n \geqslant 2$ et $\mathbb{K}$ un corps commutatif.

(1) *Si $A \in \mathrm{GL}(n,\mathbb{K})$, il existe des transvections $U_1, \dots, U_r$, $V_1, \dots, V_s$ telles que*

$$A = U_1 \dots U_r \, D_n\big(\det(A)\big) \, V_1 \dots V_s. \tag{13.45}$$

EQooKSQVooIpkdJE

ITEMooLRYXooSoKRiA

(2) *L'ensemble des transvections engendre le groupe spécial linéaire $\mathrm{SL}(n,\mathbb{K})$.*

(3) L'ensemble des transvections et des dilatations engendre le groupe linéaire $\mathrm{GL}(n,\mathbb{K})$.

Démonstration. Nous allons montrer que toutes les matrices de $\mathrm{SL}(n,\mathbb{K})$ peuvent être écrites comme produits de matrices de la forme (13.35). Cela montrera qu'étant donné un endomorphisme f et une base *pas spécialement liée à* f, il est possible d'écrire la matrice de f comme produit de transvections dont les hyperplans invariants sont « contenus » dans cette base. Cela suffit à prouver que les transvections engendrent $\mathrm{SL}(n,\mathbb{K})$ grâce au lemme 1.323.

Toutes les transvections ont un déterminant égal à 1. Donc le groupe engendré par les transvections est inclus dans $\mathrm{SL}(2,\mathbb{K})$. Soit $A\in\mathrm{GL}(n,\mathbb{K})$; nous allons utiliser le pivot de Gauss pour la diagonaliser. Étant donné que A est inversible, sa première colonne n'est pas nulle. Si $A_{i1}\neq 0$ alors une multiplication à gauche par $L_{1i}\big((A_{11}-1)/A_{i1}\big)$ effectue la substitution

$$L_1\to L_1-\frac{A_{11}-1}{A_{i1}}L_i \tag{13.46}$$

qui met un 1 en la position $(1,1)$. Notons que si la première colonne est de la forme

$$\begin{pmatrix} s\\0\\\vdots\\0\end{pmatrix} \tag{13.47}$$

avec $s\neq 0$ alors il faut plutôt faire les substitutions $L_2\to L_2+L_1$ et ensuite $L_1\to L_1-\frac{1}{s}L_2$ pour obtenir le même résultat. En effectuant le pivot avec A_{11}, une suite d'opérations sur les lignes et les colonnes donnent

$$M_1\dots M_pAN_1\dots N_q=\begin{pmatrix}1&0\\0&A_1\end{pmatrix} \tag{13.48}$$

où $A_1\in\mathrm{GL}(n-1,\mathbb{K})$ et $\det(A_1)=\det(A)$. En continuant de la sorte nous arrivons sur une matrice diagonale [9]

$$M_1\dots M_{p'}AN_1\dots N_{q'}=\begin{pmatrix}1&&&\\&\ddots&&\\&&1&\\&&&\alpha\end{pmatrix} \tag{13.49}$$

avec $\alpha=\det(A)$. En d'autres termes nous avons prouvé qu'il existe des transvections $U_1,\dots,U_r$ et $V_1,\dots,V_s$ telles que

$$A=U_1\dots U_r\,D_n\big(\det(A)\big)\,V_1\dots V_s. \tag{13.50}$$

EQooZYYFooQGCgxU

Cela prouve que les transvections et les translations engendrent $\mathrm{GL}(n,\mathbb{K})$. Si $A\in\mathrm{SL}(n,\mathbb{K})$ alors $D_n\big(\det(A)\big)=1$ et l'équation (13.50) est un produit de transvections. □

Proposition 13.13.
Le groupe $\mathrm{GL}(n,\mathbb{R})$ *est engendré par les endomorphismes inversibles diagonalisables.*

Démonstration. Par la proposition 13.12, le groupe $\mathrm{GL}(n,\mathbb{R})$ est engendré par les dilatations et les transvections. Il suffit donc de montrer qu'à leur tour, ces deux types d'endomorphismes sont engendrés par les endomorphismes inversibles et diagonalisables.

Les dilatations sont diagonalisables et inversibles. C'est bon pour elles.

Soit une transvection u, et une base $\{e_i\}_{i=1,\dots,n}$ dans laquelle u est de la forme (13.18). Nous considérons l'endomorphisme $d\colon E\to E$ défini par $d(e_k)=ke_k$. Cet endomorphisme est diagonalisable parce que son polynôme minimal, $\mu_d=\prod_{k=1}^n(X-k)$, est scindé à racines simples (voir le théorème 8.219).

9. Attention : les opérations sur les lignes et les colonnes ne sont pas des opérations de similitude. Il n'est pas question de prétendre ici que toutes les matrices de $\mathrm{GL}(n,\mathbb{K})$ sont diagonales, voir la définition 4.111.

Nous avons évidemment $u = d^{-1} \circ (d \circ u)$ où d^{-1} est diagonalisable et inversible. Voyons que $d \circ u$ est également diagonalisable en montrant que μ_d est son polynôme minimal (qui est scindé à racines simples).

Il suffit de montrer que $\mu_d(d \circ u)(e_k) = 0$ pour tout k. Ainsi μ_d sera un polynôme annulateur de $d \circ u$ de degré n, et donc minimal.

(1) **Si $k \leqslant n-1$** Alors $u(e_k) = e_k$ et $(d \circ u - n)e_k = (k-n)e_k$. Donc :

$$\mu_d(d\circ u)(e_k) = (d\circ u-1)(d\circ u-2)\dots(d\circ u-n)e_k = (k-1)(k-2)\dots(k-n)e_k = 0 \quad (13.51)$$

parce que dans le produit des $k-i$, il y en a forcément un de nul.

(2) **Si $k = n$** Dans un premier temps,

$$(d \circ u - n)e_n = d(e_n + e_{n-1}) - ne_n = ne_n + (n-1)e_{n-1} - ne_n = (n-1)e_{n-1}. \quad (13.52)$$

Ensuite

$$\begin{aligned}\big(d \circ u - (n-1)\big)e_{n-1} &= d(e_{n-1}) - (n-1)e_{n-1} && (13.53a)\\ &= d(e_{n-1}) - (n-1)e_{n-1} && (13.53b)\\ &= (n-1)e_{n-1} - (n-1)e_{n-1} && (13.53c)\\ &= 0 && (13.53d)\end{aligned}$$

Le polynôme μ_d est donc un polynôme scindé à n racines simples annulateur de $d \circ u$, qui est alors diagonalisable et inversible (parce que u et d le sont).

Donc sous la forme $u = d^{-1}(du)$, la transvection u est écrite comme produit de diagonalisables inversibles. □

PROPooSAOTooIlpJoY

Proposition 13.14 ([150]).
Soient $n \geqslant 3$ et $\mathbb{K}$ un corps de caractéristique différente de 2. Alors

(1) le groupe dérivé de $\mathrm{GL}(n,\mathbb{K})$ est $\mathrm{SL}(n,\mathbb{K})$;

(2) le groupe dérivé de $\mathrm{SL}(n,\mathbb{K})$ est $\mathrm{SL}(n,\mathbb{K})$.

La preuve utilise le fait que les transvections engendrent $\mathrm{SL}(n,\mathbb{K})$ et que les transvections avec les dilatations engendrent $\mathrm{GL}(n,\mathbb{K})$. Voir la proposition 13.12.

13.2.2 Connexité de certains groupes

LEMooIPOVooZJyNoH

Lemme 13.15.
Le groupe $\mathrm{O}(n,\mathbb{R})$ n'est pas connexe.

Démonstration. La non connexité par arcs est facile parce que les éléments de déterminant 1 ne peuvent pas être reliés aux éléments de déterminant -1 par un chemin continu restant dans $\mathrm{O}(n)$ à cause du théorème des valeurs intermédiaires 9.93.

En ce qui concerne la connexité, il faut en dire un peu plus.

Les éléments de $\mathrm{O}(n,\mathbb{R})$ ont des déterminants égaux à 1 ou à -1. Ces deux parties sont des ouverts (pour la topologie induite de $\mathbb{M}(n,\mathbb{R})$). En effet soit $A \in \mathrm{SO}(n,\mathbb{R})$ (la partie contenant les déterminants 1 ; ce que l'on va dire tient pour l'autre partie). Alors, parce que le déterminant est une fonction continue sur $\mathbb{M}(n,\mathbb{R})$, il existe un voisinage $\mathcal{O}$ de A dans $\mathbb{M}(n,\mathbb{R})$ dans lequel le déterminant reste entre $\frac{1}{2}$ et $\frac{3}{2}$ (c'est la définition de la continuité avec $\epsilon = 1/2$). L'ensemble $\mathcal{O} \cap \mathrm{O}(n,\mathbb{R})$ est par définition un ouvert de $\mathrm{O}(n,\mathbb{R})$ et ne contient que des éléments de déterminant 1.

La partie $\mathrm{O}(n,\mathbb{R})$ de $\mathbb{M}(n,\mathbb{R})$ est donc non-connexe selon la définition 7.70. □

LEMooQMXHooZQozMK

Lemme 13.16.
Les groupes $\mathrm{U}(n)$ et $\mathrm{SU}(n)$ sont connexes par arcs.

Démonstration. Soient A une matrice unitaire, et Q une matrice unitaire qui diagonalise A. Étant donné que les valeurs propres arrivent par paires complexes conjuguées,

$$QAQ^{-1} = \begin{pmatrix} e^{i\theta_1} & & & & \\ & e^{-i\theta_1} & & & \\ & & \ddots & & \\ & & & e^{i\theta_r} & \\ & & & & e^{-i\theta_r} \end{pmatrix}. \tag{13.54}$$

Le chemin $U(t)$ obtenu en remplaçant θ_i par $t\theta_i$ avec $t \in [0,1]$ joint QAQ^{-1} à l'identité. Par conséquent $Q^{-1}U(t)Q$ joint A à l'unité. □

Théorème 13.17.
Les matrices normales[10] *forment un espace connexe par arc.*

Démonstration. Soient A une matrice normale et U une matrice unitaire qui diagonalise A. Nous considérons $U(t)$, un chemin qui joint $\mathbb{1}$ à U dans $\mathrm{U}(n)$. Pour chaque t, la matrice

$$A(t) = U(t)^{-1}AU(t) \tag{13.55}$$

est normale. Nous avons donc trouvé un chemin dans les matrices normales qui joint A à une matrice diagonale. Il est à présent facile de la joindre à l'identité.

Toutes les matrices normales étant connexes à l'identité, l'ensemble des matrices normales est connexe. □

PROPooALQCooLZCKrH

Proposition 13.18.
Le groupe $\mathrm{SL}(n,\mathbb{K})$ *est connexe par arcs.*

Démonstration. Soit $A \in \mathrm{SL}(n,\mathbb{K})$; par la proposition 13.12(2) nous pouvons écrire

$$A = \prod_{c\in X} T_c(\lambda_c) \tag{13.56}$$

où X est une partie de l'ensemble des couples (i,j) dans $\{1,\dots,n\}$. En posant

$$\begin{aligned} \varphi\colon [0,1] &\to \mathrm{SL}(n,\mathbb{K}) \\ t &\mapsto \prod_{c\in X} T_c(t\lambda_c) \end{aligned} \tag{13.57}$$

nous avons une application continue de A vers $\mathbb{1}$, qui, à tout t fait correspondre la matrice $\varphi(t)$, inversible de déterminant 1.

Donc tous les éléments de $\mathrm{SL}(n,\mathbb{K})$ peuvent être reliés à $\mathbb{1}$. Par conséquent, $\mathrm{SL}(n,\mathbb{K})$ est connexe par arcs. □

PROPooVJNIooMByUJQ

Proposition 13.19 ([424])**.**
Le groupe $\mathrm{GL}(n,\mathbb{C})$ *est connexe par arcs.*

Démonstration. Soient $A \in \mathrm{GL}(n,\mathbb{C})$ et sa décomposition (13.45). Comme montré précédemment, chacune des transvections peut être reliée à $\mathbb{1}$ par un chemin continu dans $\mathrm{SL}(n,\mathbb{C})$. En ce qui concerne le facteur de translation, nous ne pouvons pas simplement prendre le chemin donné par $t \mapsto D_n\big(t\det(A)\big)$ parce que le résultat n'est pas inversible en $t = 0$.

Puisque $\mathbb{C}^*$ est connexe par arcs il existe une application continue $\alpha\colon [0,1] \to \mathbb{C}^*$ telle que $\alpha(0) = \det(A) \in \mathbb{C}^*$ et $\alpha(1) = 1$. Il suffit alors de prendre $D_n\big(\alpha(t)\big)$ et nous avons un chemin continu de A vers $\mathbb{1}$ restant dans $\mathrm{GL}(n,\mathbb{C})$. □

PROPooBIYQooWLndSW

Proposition 13.20.
Le groupe $\mathrm{GL}(n,\mathbb{R})$ *a exactement deux composantes connexes par arcs.*

10. Définition 12.101.

Démonstration. Nous notons $\mathrm{GL}^+(n,\mathbb{R})$ et $\mathrm{GL}^-(n,\mathbb{R})$ les parties de $\mathrm{GL}(n,\mathbb{R})$ formées des applications de déterminant strictement positifs et strictement négatifs respectivement. D'après le théorème des valeurs intermédiaires (théorème 9.93), il n'existe pas d'application continue dans $\mathrm{GL}(n,\mathbb{R})$ reliant $\mathrm{GL}^+(n,\mathbb{R})$ à $\mathrm{GL}^-(n,\mathbb{R})$ tout en restant dans les applications de déterminant non nul [11].

Montrons que $\mathrm{GL}^{\pm}(n,\mathbb{R})$ sont connexes par arcs. Si $A\in\mathrm{GL}^+(n,\mathbb{R})$ alors grâce à la décomposition (13.45), il existe un chemin continu de A vers $D_n\big(\det(A)\big)$. Puisque $\mathbb{R}^{\pm}$ sont connexes par arcs, il est possible de relier $D_n\big(\det(A)\big)$ à $D_n(\pm1)$ par un chemin continu. □

13.2.3 Densité

PropDigDensVxzPuo

Proposition 13.21.
Les matrices diagonalisables sont denses dans $\mathbb{M}(n,\mathbb{C})$.

Démonstration. D'après le lemme de Schur 12.100, une matrice de $\mathbb{M}(n,\mathbb{C})$ est de la forme

$$A=Q\begin{pmatrix}\lambda_1 & * & *\\ 0 & \ddots & *\\ 0 & 0 & \lambda_n\end{pmatrix}Q^{-1}. \tag{13.58}$$

Les valeurs propres sont sur la diagonale. La matrice est diagonalisable si les éléments de la diagonales sont tous différents. Il suffit maintenant de considérer n suites $(\epsilon_k^{(r)})_{k\in\mathbb{N}}$ convergentes vers zéro telles que pour chaque k les nombres $\lambda_r+\epsilon_k^{(r)}$ soient tous différents. La suite de matrices

$$A_k=Q\begin{pmatrix}\lambda_1+\epsilon_k^{(1)} & * & *\\ 0 & \ddots & *\\ 0 & 0 & \lambda_n+\epsilon_k^{(n)}\end{pmatrix}Q^{-1}. \tag{13.59}$$

est alors diagonalisable pour tout k et nous avons $\lim_{k\to\infty}A_k=A$. □

PropQGUPooVudelJ

Proposition 13.22.
Les matrices inversibles sont denses dans l'ensemble des matrices. C'est-à-dire que $\mathrm{GL}(n,\mathbb{R})$ *est dense dans* $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Soit $A\in\mathbb{M}(n,\mathbb{R})$; le lemme de Schur réel 8.226 nous permet d'écrire

$$A=Q^{-1}\begin{pmatrix}\lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & \begin{pmatrix}a & b\\ c & d\end{pmatrix} & & \\ & & & & \ddots & \end{pmatrix}Q \tag{13.60}$$

EQooJHPAooWFVSat

avec Q orthogonale.

Nous allons définir une suite (A_k) de matrices inversibles convergeant vers A. Pour cela nous considérons, pour chaque i, une suite $(\epsilon_k^{(i)})$ telle que $\lambda_i+\epsilon_k^{(i)}\neq0$ et $\epsilon_k^{(i)}\to0$. La matrice A_k est la matrice donnée par (13.60) dans laquelle nous remplaçons λ_i par $\lambda_i+\epsilon_k^{(i)}$.

En ce qui concerne les blocs, ceux dont le déterminant est non nul, nous n'y touchons pas, et ceux dont le déterminant est nul, nous remplaçons a par $a+\epsilon_k$.

Avec cela, $Q^{-1}A_kQ$ est une suite dans $\mathrm{GL}(n,\mathbb{R})$ qui converge vers A. □

11. Si $\varphi\colon[0,1]\to\mathrm{GL}(n,\mathbb{R})$ est le chemin, la fonction à mettre dans le théorème des valeurs intermédiaires est la fonction $f\colon[0,1]\to\mathbb{R}$ $t\mapsto\det\big(\varphi(t)\big)$.

PROPooZUHOooQBwfZq

Proposition 13.23.
Si $A \in \mathbb{M}(n,\mathbb{C})$ alors

$$e^{\mathrm{Tr}(A)} = \det(e^A). \tag{13.61}$$

Démonstration. Ici, e^A est l'exponentielle, soit d'endomorphisme, soit de matrice définie par la proposition 10.301.

Le résultat est un simple calcul pour les matrices diagonalisables. Si A n'est pas diagonalisable, nous considérons une suite de matrices diagonalisables A_k dont la limite est A (proposition 13.21). La suite

$$a_k = e^{\mathrm{Tr}(A_k)} \tag{13.62}$$

converge vers $e^{\mathrm{Tr}(A)}$ tandis que la suite

$$b_k = \det(e^{A_k}) \tag{13.63}$$

converge vers $\det(e^A)$. Mais nous avons $a_k = b_k$ pour tout k ; les limites sont donc égales. □

CORooOKKSooHrsYOs

Corolaire 13.24.
Si $X \in \mathbb{M}(n,\mathbb{C})$ alors

$$\frac{d}{dt}\Big[\det(e^{tX})\Big]_{t=0} = \mathrm{Tr}(X). \tag{13.64}$$

Démonstration. Nous écrivons la proposition 13.23 pour tX au lieu de X ; pour chaque t nous avons

$$\det\left(e^{tX}\right) = e^{\mathrm{Tr}(tX)} = e^{t\,\mathrm{Tr}(X)}. \tag{13.65}$$

La dérivation par rapport à t en $t = 0$ donne le résultat attendu. □

ThoHZTooWDjTYI

Théorème 13.25 (Cayley-Hamilton[425, 426])**.**
Tout endomorphisme d'un espace vectoriel de dimension finie sur un corps commutatif quelconque annule son propre polynôme caractéristique

Une autre démonstration est donnée par le théorème 8.120.

Démonstration. La preuve est divisée en plusieurs étapes.

(1) **Endomorphisme diagonalisable** Soit u un endomorphisme sur un espace vectoriel V de dimension n sur un corps $\mathbb{K}$ et χ_u sont polynôme caractéristique. Nous savons que si λ est une valeur propre de u alors $\chi_u(\lambda) = 0$ par le théorème 8.117(2). En combinant avec le lemme 8.95, si x est vecteur propre pour la valeur propre λ de u nous avons

$$\chi_u(u)x = \chi_u(\lambda)x = 0. \tag{13.66}$$

Donc tant que u possède une base de vecteurs propres nous avons $\chi_u(u) = 0$.

(2) **Le cas complexe**
Nous nous restreignons à présent (et provisoirement) au cas $\mathbb{K} = \mathbb{C}$, ce qui nous donne $u \in \mathbb{M}(n,\mathbb{C})$. Les matrices diagonalisables sont denses dans $\mathbb{M}(n,\mathbb{C})$ par la proposition 13.21. Si $A \in \mathbb{M}(n,\mathbb{C})$ nous considérons une suite de matrices diagonalisables $A_k \xrightarrow{\mathbb{M}(n,\mathbb{C})} A$. Pour chaque k nous avons par le point précédent

$$\chi_{u_k}(u_k) = 0. \tag{13.67}$$

Chacune des composantes de $\chi_{u_k}(u_k)$ est un polynôme en les composantes de u_k, ce qui légitime le passage à la limite :

$$\chi_u(u) = 0. \tag{13.68}$$

Le théorème est établi pour toutes les matrices de $\mathbb{M}(n,\mathbb{C})$ et donc aussi pour tous les sous-corps de $\mathbb{C}$ comme $\mathbb{R}$ ou $\mathbb{Z}$.

(3) **Le cas général**
Par définition, $\chi_u(X) = \det(u - X\mathbb{1})$; les coefficients de X sont des polynômes à coefficients entiers en les composantes de u. En substituant u à X nous obtenons une matrice dont chacune des entrées est un polynôme à coefficients entiers en les coefficients de u. Pour chaque i et j entre 1 et n il existe donc un polynôme $P_{ij} \in \mathbb{Z}(X_1, \dots, X_{n^2})$ tel que

$$\chi_u(u)_{ij} = P(u_{11}, \dots, u_{nn}). \tag{13.69}$$

Ces polynômes ne dépendent pas de u ni du corps sur lequel on travaille. Notre but est maintenant de prouver que $P_{ij} = 0$.
Étant donné que le cas complexe (et a fortiori entier) est déjà prouvé nous savons que pour tout $u \in \mathbb{M}(n, \mathbb{Z})$ nous avons $P(u_{11}, \dots, u_{nn}) = 0$. La proposition 6.186 nous donne effectivement $P = 0$, en conséquence de quoi l'endomorphisme $\chi_u(u)$ est nul.

□

Exemple 13.26.
Pour montrer que chaque composante $\chi_u(u)$ est bien un polynôme à coefficients entiers en les coefficients de u, voyons l'exemple 2×2 : $u = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. D'abord

$$\chi_u(X) = \det \begin{pmatrix} a - X & b \\ c & d - X \end{pmatrix} = X^2 - (a + d)X + ad - cb. \tag{13.70}$$

Le coefficient de X^2 est 1, celui de X est $-a - d$ et le terme indépendant est $ad - cb$; tout trois sont des polynômes à coefficients entiers en a, b, c, d. Après substitution de X par u,

$$\chi_u(u)_{ij} = (u^2)_{ij} - (a + d)u_{ij} + ad - cb. \tag{13.71}$$

C'est bien un polynôme à coefficients entiers en les entrées de la matrice u. △

13.2.4 Racine carrée d'une matrice hermitienne positive

PropVZvCWn

Proposition-Définition 13.27.
Si $A \in \mathbb{M}(n, \mathbb{C})$ est une matrice hermitienne[12] *positive, alors il existe une unique matrice hermitienne positive R telle que $A = R^2$. De plus R est un polynôme (de $\mathbb{R}[X]$) en A.*
*La matrice R ainsi définie est la **racine carrée** de A, et est notée $\sqrt{A}$.*

Démonstration. En plusieurs parties.

(1) **Existence** Étant donné que A est hermitienne, elle est diagonalisable par une matrice unitaire (proposition 12.102), et ses valeurs propres sont réelles et positives (parce que A est positive). Soit donc P une matrice unitaire telle que

$$P^*AP = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \tag{13.72}$$

avec $\alpha_i > 0$. Si on pose

$$R = P \begin{pmatrix} \sqrt{\alpha_1} & & \\ & \ddots & \\ & & \sqrt{\alpha_n} \end{pmatrix} P^*, \tag{13.73}$$

alors $R^2 = A$ parce que $P^*P = \mathbb{1}$.

12. Définition 8.179.

(2) **Hermitienne positive** La matrice R est hermitienne parce que, avec un peu de notation raccourcie, $R = P^*\sqrt{\alpha}P$ et $R^* = P^*\sqrt{\alpha}P$. D'autre part, elle est positive parce que ses valeurs propres sont les $\sqrt{\alpha_i}$ qui sont positives.

(3) **Polynôme** Nous montrons maintenant que la matrice R est un polynôme en A. Pour cela nous considérons un polynôme Q tel que $Q(\alpha_i) = \sqrt{\alpha_i}$ pour tout i. Soit $\{e_i\}$ une base de diagonalisation de A : $Ae_i = \alpha_i e_i$. Alors c'est encore une base de diagonalisation de $Q(A)$. En effet si $Q = \sum_k a_k X^k$, alors

$$Q(A)e_i = (\sum_k a_k A^k)e_i = (\sum_k a_k \alpha_i^k)e_i = Q(\alpha_i)e_i = \sqrt{\alpha_i}e_i. \tag{13.74}$$

Les valeurs propres de $Q(A)$ sont donc $\sqrt{\alpha_i}$. Nous savons maintenant que $Q(A)$ a la même base de diagonalisation de A (et donc la même matrice unitaire P qui diagonalise), c'est-à-dire que

$$Q(A) = P^* \begin{pmatrix} \sqrt{\alpha_1} & & \\ & \ddots & \\ & & \sqrt{\alpha_n} \end{pmatrix} = R. \tag{13.75}$$

Donc oui, R est un polynôme en A.

Notons que ce Q n'est pas du tout unique ; il existe une infinité de polynômes envoyant n nombres donnés sur n nombres donnés.

(4) **Unicité** Soit S une matrice hermitienne positive telle que $R^2 = S^2 = A$. D'abord S commute avec A parce que

$$SA = S^3 = S^2S = AS. \tag{13.76}$$

Donc S commute aussi avec $Q(A) = R$. Étant donné que S et R commutent et sont diagonalisables, ils sont simultanément diagonalisables par le corolaire 8.220. Soient $D_R = PRP^*$ et $D_S = PSP^*$ les formes diagonales de R et S dans une base de simultanée diagonalisation. Les carrés des valeurs propres de R et S étant identiques (ce sont les valeurs propres de A) et les valeurs propres de R et S étant positives, nous déduisons que $D_R = D_S$ et donc que $R = P^*D_RP = P^*D_SP = S$.

□

Une des applications usuelles de cette proposition est la décomposition polaire.

13.2.5 Racine carrée d'une matrice symétrique positive

LemTLlTAAf

Lemme 13.28 ([427]).
Le groupe orthogonal $\mathrm{O}(n,\mathbb{R})$ *est compact.*

Démonstration. Nous avons $O(n) = f^{-1}\big(\{\mathbb{1}_n\}\big)$ où f est l'application continue $A \mapsto A^tA$. En tant qu'image inverse d'un fermé par une application continue, le groupe $O(n)$ est fermé.

De plus il est borné parce que tous les coefficients d'une matrice orthogonale sont $\leqslant 1$, donc $\|A\|_\infty$ pour tout $A \in O(n)$. □

PropPEMDqVT

Proposition 13.29.
Une matrice symétrique définie ou semi définie positive, admet une unique racine carrée symétrique. Le spectre de la racine carrée est la racine carrée du spectre de la matrice de départ.

Démonstration. Propriétés de la racine carrée d'une matrice symétrique

(1) **Existence** Soit T une matrice symétrique et Q une matrice orthogonale qui diagonalise[13] T : $QTQ^{-1} = D$ avec $D = \mathrm{diag}(\lambda_i)$ et $\lambda_i \geqslant 0$. En posant $R = Q^{-1}\sqrt{D}Q$, il est vite vérifié que $R^2 = T$ et que R est symétrique. En ce qui concerne le spectre, R a pour valeurs propres les $\sqrt{\lambda_i}$.

13. Théorème 8.227.

(2) **Unicité**

Soit R une matrice symétrique de T : $R^2 = T$. Du coup R et T commutent : $RT = R^3 = TR$. Par conséquent les espaces propres de T sont stables sous R. Soit E_λ l'un d'eux de dimension d, et T_F, R_F les restrictions de T et R à E_λ. L'application T_F est une homothétie et $R_F^2 = T_F = \lambda\mathbb{1}$. Mais R_F est encore une matrice symétrique définie positive, donc nous pouvons considérer une base $\{e_1, \dots, e_d\}$ de E_λ qui diagonalise R_F avec les valeurs propres μ_i ; nous avons donc en même temps

$$R_f^2(e_i) = \mu_i^2 e_i \tag{13.77a}$$
$$T_F(e_i) = \lambda e_i, \tag{13.77b}$$

de telle sorte que $\mu_i^2 = \lambda$. Mais les valeurs propres de R_F sont positives, sont $\mu_i = \sqrt{\lambda}$ pour tout i. En conclusion R_F est univoquement déterminé par la donnée de T. Vu que cela est valable pour tous les espaces propres de T et que ces espaces propres engendrent tout E, l'opérateur R est déterminé de façon univoque par T.

□

Notons que nous n'avons démontré l'unicité qu'au sein des matrices symétriques.

13.2.6 Décomposition polaires : cas réel

Nous rappelons que $S^{++}(n, \mathbb{R})$ est l'ensemble des matrices symétriques strictement définies positives, d'après le paragraphe 8.231.

LemMGUSooPqjguE

Lemme 13.30.
La partie $S^+(n, \mathbb{R})$ est fermée dans $\mathbb{M}(n, \mathbb{R})$.

Démonstration. En effet si S_k est une suite de matrices symétriques convergeant dans $\mathbb{M}(n, \mathbb{R})$ vers la matrice A, les suites $(S_k)_{ij}$ et $(S_k)_{ji}$ des composantes ij et ji sont des suites égales, et donc leurs limites sont égales [14]. Donc la limite est symétrique.

En ce qui concerne le spectre, le théorème 8.227 nous permet de diagonaliser : $S_k = Q_k D_k Q_k^{-1}$ où les D_k sont des matrices diagonales remplies de nombres positifs ou nuls. Comme $O(n)$ est compact [15], nous avons une sous-suite $Q_{\varphi(k)}$ convergente : $Q_{\varphi(k)} \to Q$. Pour chaque k, nous avons

$$S_{\varphi(k)} = Q_{\varphi(k)} D_{\varphi(k)} Q_{\varphi(k)}^{-1}, \tag{13.78}$$

dont la limite existe et vaut A. Puisque pour tout k, $D_{\varphi(k)} = Q_{\varphi(k)}^{-1} S_{\varphi(k)} Q_{\varphi(k)}$ et que le produit matriciel est continu, la suite $k \mapsto D_{\varphi(k)}$ est une suite convergente dans $\mathbb{M}(n, \mathbb{R})$. Nous notons D sa limite qui est encore une matrice diagonale contenant des nombres positifs ou nuls sur la diagonale.

$$A = \lim_{k\to\infty} S_{\varphi(k)} = QDQ^{-1}, \tag{13.79}$$

et donc le spectre de A est la limite de ceux des matrices $D_{\varphi(k)}$. Chacun étant positif, la limite est positive. Donc $A \in S^+(n, \mathbb{R})$. □

LemZKJWqIP

Lemme 13.31.
La fermeture de l'ensemble des matrices symétriques strictement définies positives est l'ensemble des matrices définies positives : $\mathrm{Adh}(S^{++}(n, \mathbb{R})) = S^+(n, \mathbb{R})$.

14. Ici nous utilisons le critère de convergence composante par composante et le fait que nous ne sommes pas trop inquiétés par la norme que nous choisissons parce que toutes les normes sont équivalentes par le théorème 10.49.
15. Lemme 13.28.

Démonstration. Le lemme 13.30 nous dit que $S^+(n,\mathbb{R})$ était fermé. Nous devons prouver que pour tout élément de $S^+(n,\mathbb{R})$, il existe une suite (S_k) dans $S^{++}(n,\mathbb{R})$ convergeant vers S.

Si $S\in S^+(n,\mathbb{R})$ alors nous avons la diagonalisation

$$S=QDQ^{-1}=Q\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}Q^{-1} \tag{13.80}$$

où $\lambda_i\geqslant 0$ pour tout i. Nous définissons

$$D_k=\begin{pmatrix}\lambda_1+\epsilon_k^{(1)} & & \\ & \ddots & \\ & & \lambda_n+\epsilon_k^{(n)}\end{pmatrix} \tag{13.81}$$

où ϵ_k^i est une suite convergent vers 0 telle que $\lambda_i+\epsilon_n^{(i)}>0$ pour tout i et k. Typiquement si $\lambda_i>0$ alors $\epsilon_k^{(i)}=0$ et sinon $\epsilon_k^{(i)}=1/k$.

Pour tout k nous avons $QD_kQ^{-1}\in S^{++}(n,\mathbb{R})$ et de plus $QD_kQ^{-1}\to QDQ=S$. □

ThoLHebUAU

Théorème 13.32 (Décomposition polaire de matrices symétriques définies positives[427, 428, 429])**.**
En ce qui concerne les matrices inversibles :

$$\begin{aligned}f\colon O(n,\mathbb{R})\times S^{++}(n,\mathbb{R})&\to \mathrm{GL}(n,\mathbb{R})\\ (Q,S)&\mapsto SQ\end{aligned} \tag{13.82}$$

est un homéomorphisme[16]*.*

En ce qui concerne les matrices en général :

$$\begin{aligned}g\colon O(n,\mathbb{R})\times S^{+}(n,\mathbb{R})&\to \mathbb{M}(n,\mathbb{R})\\ (Q,S)&\mapsto SQ\end{aligned} \tag{13.83}$$

est une surjection mais pas une injection.

De plus les mêmes conclusions tiennent si nous prenons $(Q,S)\mapsto QS$ au lieu de SQ.

¡¡ Avertissement/question au lecteur !! 13.33
Je crois que les éléments de la décomposition polaire sont des polynômes en M. Écrivez-moi si vous pouvez confirmer ou infirmer.

Démonstration. Nous commençons par prouver les résultats concernant les matrices inversibles.

(1) **Existence et unicité**
Si $M=SQ$, alors $MM^t=SQQ^tS^t=S^2$, donc S doit être une racine carrée symétrique de la matrice définie positive MM^t. La proposition 13.29 nous dit que ça existe et que c'est unique. Donc S est univoquement déterminé par M. Maintenant avoir $Q=MS^{-1}$ est obligatoire (unicité) et fonctionne :

$$Q^tQ=(S^{-1})^tM^tMS^{-1}=S^{-1}S^2S^{-1}=\mathbb{1}, \tag{13.84}$$

donc Q ainsi défini est orthogonale.

Notons que ceci ne fonctionne pas lorsque M n'est pas inversible parce qu'alors S n'est pas inversible.

16. Cela est en réalité un difféomorphisme, voir la remarque 13.34.

(2) **Homéomorphisme**

Le fait que f soit continue n'est pas un problème : c'est un produit de matrices. Nous devons vérifier que f^{-1} est continue. Soit une suite convergente $M_k \to M$ dans $\mathrm{GL}(n,\mathbb{R})$. Si nous nommons (Q_k, S_k) la décomposition polaire de M_k et (Q,S) celle de M, nous devons prouver que $Q_k \to Q$ et $S_k \to S$. En effet dans ce cas nous aurions

$$\lim_{k\to\infty} f^{-1}(M_k) = \lim_{k\to\infty}(Q_k,S_k) = (Q,S) = f^{-1}(M). \tag{13.85}$$ EqJJkoaJv

Étant donné que $O(n)$ est compact (lemme 13.28), la suite (Q_k) admet une sous-suite convergente (Bolzano-Weierstrass, théorème 7.144) que nous nommons

$$Q_{\varphi(k)} \to F \in O(n). \tag{13.86}$$

Vu que la suite (M_k) converge, sa sous-suite converge vers la même limite : $M_{\varphi(k)} \to M$ et vu que pour tout k nous avons $S_k = M_k Q_k^{-1}$,

$$S_{\varphi(k)} \to G = MF^{-1}. \tag{13.87}$$

Vu que chacune des matrices $S_{\varphi(k)}$ est symétrique définie positive, la limite est symétrique et semi-définie positive [17]. Donc $G \in S^+(n,\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R})$ parce que de plus M et F étant inversibles, G est inversible. En ce qui concerne la sous-suite nous avons

$$M_{\varphi(k)} = S_{\varphi(k)} Q_{\varphi(k)} \to GF = M \tag{13.88}$$

où $F \in O(n)$ et $G \in S^+(n,\mathbb{R})$. Par unicité de la décomposition polaire de M (partie déjà démontrée), nous avons $G = S$ et $F = Q$.

Nous avons prouvé que toute sous-suite convergente de Q_k a Q pour limite. Donc la suite elle-même converge [18] vers Q. Donc $Q_k \to Q$. Du coup vu que $S_k = M_k Q_k^{-1}$ est un produit de suites convergentes, S_k converge également, vers S : $S_k \to S$.

Au final l'application f^{-1} est bien continue parce que les égalités (13.85) ont bien lieu.

Nous passons maintenant à la preuve dans le cas des matrices en général.

Soit $A \in \mathbb{M}(n,\mathbb{R})$; par densité (lemme 13.22), il existe une suite (A_k) dans $\mathrm{GL}(n,\mathbb{R})$ telle que $A_k \to A$. Pour chacun des k nous appliquons la décomposition polaire déjà prouvée : $A_k = Q_k S_k$. D'abord (Q_k) est une suite dans le compact [19] $\mathrm{O}(n,\mathbb{R})$ et accepte donc une sous-suite convergente. Quitte à redéfinir la suite de départ, nous supposons pour alléger les notations que $Q_k \to Q \in \mathrm{O}(n,\mathbb{R})$. Vu que Q_k est inversible,

$$S_k = Q_k^{-1} A_k \tag{13.89}$$

Le produit matriciel étant continu nous avons $S_k \to S$ dans $\mathbb{M}(n,\mathbb{R})$. Mais $S^+(n,\mathbb{R})$ étant fermé (lemme 13.30) nous avons aussi $S \in S^+(n,\mathbb{R})$. □

RemBJCBooGLiRmG

Remarque 13.34.
Pour démontrer que f est différentiable, nous devons utiliser le théorème d'inversion locale 17.47 ; cela est fait dans la proposition 17.57.

CorAWYBooNCCQSf

Corolaire 13.35.
Toute matrice peut être écrite sous la forme $Q_1 D Q_2$ où Q_1 et Q_2 sont orthogonales et D est diagonale.

Démonstration. Si $A \in \mathbb{M}(n,\mathbb{R})$ alors la décomposition polaire 13.32 nous donne $A = SQ$ où S est symétrique définie positive et Q est orthogonale. La matrice S peut ensuite être diagonalisée par le théorème 8.227 : $S = RDR^{-1}$ où D est diagonale et R est orthogonale. Avec ces deux décompositions en main, $A = SQ = RDR^{-1}Q$. La matrice $R^{-1}Q$ est orthogonale. □

17. Lemme 13.31
18. Proposition 7.305, pas difficile.
19. Lemme 13.28.

13.2.7 Enveloppe convexe

Définition 13.36.
*Un point a d'un ensemble convexe C est un **point extrémal** si $C\backslash\{a\}$ est convexe.*

ThoBALmoQw

Théorème 13.37 ([104]).
Soit E un espace euclidien de dimension $n \geqslant 1$ et $\mathcal{L}(E)$ l'espace des opérateurs linéaires sur E sur lequel nous considérons la norme subordonnée[20] à celle sur E. L'ensemble des points extrémaux de la boule unité fermée de $\mathcal{L}(E)$ est le groupe orthogonal $O(n,\mathbb{R})$.

Démonstration. Nous notons $\mathcal{B}$ la boule unité fermée de $\mathcal{L}(E)$. Montrons pour commencer que les éléments de $O(n)$ sont extrémaux dans $\mathcal{B}$. D'abord si $A \in O(E)$ alors $\|A\| = 1$ parce que $\|Ax\| = \|x\|$. Supposons maintenant que A n'est pas extrémal, c'est-à-dire qu'il est le milieu d'un segment joignant deux points (distincts) de la boule unité de $\mathcal{L}(E)$. Soient donc $T, U \in \mathcal{B}$ tels que $A = \frac{1}{2}(T+U)$. Pour tout $x \in E$ tel que $\|x\| = 1$ nous avons

$$1 = \|x\| = \|Ax\| = \frac{1}{2}\|Tx + Ux\| \leqslant \frac{1}{2}\big(\|Tx\| + \|Ux\|\big) \leqslant \frac{1}{2}\big(\|T\| + \|U\|\big) \leqslant 1 \tag{13.90}$$

EqKTuAIE

Toutes les inégalités sont en réalité des égalités. En particulier nous avons

$$\|Tx + Ux\| = \|Tx\| + \|Ux\|, \tag{13.91}$$

mais alors nous sommes dans un cas d'égalité dans l'inégalité de Cauchy-Schwarz (théorème 10.1) et donc il existe $\lambda \geqslant 0$ tel que $Tx = \lambda Ux$. Mais de plus les ~~inégalité~~ égalités (13.90) nous donnent

$$\frac{1}{2}\big(\|Tx\| + \|Ux\|\big) = 1 \tag{13.92}$$

alors que nous savons que $\|Tx\|, \|Ux\| \leqslant 1$, donc $\|Tx\| = \|Ux\| = 1$. La seule possibilité est d'avoir $\lambda = 1$. Nous avons prouvé que $Tx = Ux$ pour tout x de norme 1. Nous en déduisons que $T = U$.

Au final A n'est pas le milieu d'un segment dans $\mathcal{B}$.

Nous passons donc à l'inclusion inverse : nous prouvons que les points extrémaux de $\mathcal{B}$ sont dans $O(E)$. Pour cela nous prenons $U \in \mathcal{B}\backslash O(E)$ et nous allons montrer que U n'est pas un point extrémal : nous allons l'écrire comme milieu d'un segment dans $\mathcal{B}$.

Par la seconde partie du théorème de décomposition polaire 13.32, il existe $Q \in O(n,\mathbb{R})$ et $S \in S^+(n,\mathbb{R})$ tels que $U = QS$. Nous diagonalisons S à l'aide de la matrice orthogonale P :

$$S = PDP^{-1} \tag{13.93}$$

avec $D = \operatorname{diag}(\lambda_i)$. En termes de normes, nous avons

$$\|U\| = \|S\| = \|D\|. \tag{13.94}$$

En effet vu que Q est orthogonale, $\|Ux\| = \|QSx\| = \|Sx\|$ pour tout x, donc $\|U\| = \|S\|$. De plus pour tout x nous avons

$$\|Sx\| = \|PDP^{-1}x\| = \|DP^{-1}x\|. \tag{13.95}$$

Étant donné que P^{-1} est une bijection, le supremum des $\|Sx\|$ sera le même que celui des $\|Dx\|$ et donc $\|S\| = \|D\|$. Étant donné que par définition $\|U\| \leqslant 1$, nous avons aussi $\|D\| \leqslant 1$ et donc $0 \leqslant \lambda_i \leqslant 1$ (pour rappel, les valeurs propres de D sont positives ou nulles parce que S est ainsi).

Comme $U \notin O(E)$, au moins une des valeurs propres n'est pas 1, supposons que ce soit λ_1. Alors nous avons $\alpha, \beta \in [-1,1]$ avec $-1 \leqslant \alpha < \beta \leqslant 1$ et $\lambda_1 = \frac{1}{2}(\alpha+\beta)$. Nous posons alors

$$D_1 = \operatorname{diag}(\alpha, \lambda_2, \ldots, \lambda_n) \tag{13.96a}$$
$$D_2 = \operatorname{diag}(\beta, \lambda_2, \ldots, \lambda_n). \tag{13.96b}$$

20. Définition 10.55.

Nous avons bien $D_1 \neq D2$ et $D_1 + D_2 = D$. Par conséquent

$$U = \frac{1}{2}\big(QPD_1P^{-1} + QPD_2P^{-1}\big) \tag{13.97}$$

avec $QPD_1P^{-1} \neq QPD_2^{-1}$. La matrice U est donc le milieu d'un segment. Reste à montrer que ce segment est dans $\mathcal{B}$. Pour ce faire, prenons $x \in E$ et calculons :

$$\|QPD_iP^{-1}x\| = \|D_iP^{-1}x\| \leqslant \|P^{-1}x\| = \|x\| \tag{13.98}$$

parce que $\|D_i\| \leqslant 1$ et P^{-1} est orthogonale. Au final la norme de QPD_iP est plus petite que 1 et donc U est bien le milieu d'un segment dans $\mathcal{B}$, et donc non extrémal. □

ThoVBzqUpy

Théorème 13.38 ([430]).
L'enveloppe convexe de $O(n)$ dans $\mathbb{M}_n(\mathbb{R})$ est la boule unité pour la norme induite de $\|.\|_2$ sur $\mathbb{R}^n$.

Démonstration. Nous notons $\mathcal{B}$ la boule unité fermée de $\mathbb{M}(n,\mathbb{R})$ et $\mathcal{C}\text{onv}\big(O(n,\mathbb{R})\big)$ l'enveloppe convexe de $O(n,\mathbb{R})$. Vu que $\mathcal{B}$ est convexe nous avons $\mathcal{C}\text{onv}\big(O(n)\big) \subset \mathcal{B}$.

Maintenant nous devons prouver l'inclusion inverse. Pour ce faire nous supposons avoir un élément $A \in \mathcal{B}\backslash\mathcal{C}\text{onv}\big(O(n)\big)$ et nous allons dériver une contradiction.

Remarquons que $O(n)$ est compact par le lemme 13.28 et que par conséquent $\mathcal{C}\text{onv}(O(n))$ est compacte par le corolaire 11.42 et donc fermée. Nous considérons un produit scalaire $(X,Y) \mapsto X \cdot Y$ sur $\mathbb{M}$. Vu que $\mathcal{C}\text{onv}\big(O(n)\big)$ est un fermé convexe nous pouvons considérer la projection [21] sur $\mathcal{C}\text{onv}(A)$ relativement au produit scalaire choisi.

Nous notons $P = \texttt{proj}_{\mathcal{C}\text{onv}\left(O(n)\right)}(A)$. En vertu du théorème de projection, nous avons

$$(A - P) \cdot (M - P) \leqslant 0 \tag{13.99}$$

EqYSisLTL

pour tout $M \in \mathcal{C}\text{onv}\, O(n)$. Notons $B = A - P$ pour alléger les notations. L'équation (13.99) s'écrit

$$B \cdot M \leqslant B \cdot P. \tag{13.100}$$

EqQDLZqXQ

D'autre part vu que $B \neq 0$ nous avons $B \cdot B > 0$, c'est-à-dire $B \cdot (A - P) > 0$ et donc

$$B \cdot A > B \cdot P. \tag{13.101}$$

En combinant avec (13.100),

$$B \cdot M \leqslant B \cdot P < B \cdot A. \tag{13.102}$$

EqIQNlwql

Nous utilisons maintenant la décomposition polaire, théorème 13.32, pour écrire $B = QS$ avec $Q \in O(n)$ et $S \in S^+(n,\mathbb{R})$. Vu que l'inégalité (13.102) tient pour tout $M \in \mathcal{C}\text{onv}(O(n))$, elle tient en particulier pour $Q \in O(n)$. Donc

$$B \cdot Q < B \cdot A. \tag{13.103}$$

EQooOSAXooLbTQAG

Nous nous attachons à présent au produit scalaire $(X,Y) \mapsto \text{Tr}(X^tY)$ de la proposition 12.123. D'abord

$$B \cdot Q = \text{Tr}(B^tQ) = \text{Tr}(S^tQ^tQ) = \text{Tr}(S^t) = \text{Tr}(S), \tag{13.104}$$

EaHVxWdau

et ensuite l'inégalité (13.103) devient

$$\text{Tr}(S) < B \cdot A = \text{Tr}(S^tQ^tA). \tag{13.105}$$

21. Le théorème de projection : théorème 12.142.

Nous choisissons une basse $\{e_i\}$ diagonalisant S : $Se_i = \lambda_i e_i$ vérifiant automatiquement $\lambda_i \geqslant 0$ parce que S est semi-définie positive [22]. Alors

$$\begin{aligned}
\mathrm{Tr}(S) &< \mathrm{Tr}(S^tQ^tA) && (13.106a)\\
&= \sum_i \langle S^tQ^tAe_i, e_i\rangle && (13.106b)\\
&= \sum_i \langle Ae_i, QSe_i\rangle && (13.106c)\\
&\leqslant \sum_i \|Ae_i\||\lambda_i| \underbrace{\|Qe_i\|}_{=1} && (13.106d)\\
&\leqslant \sum_i \lambda_i && A\in\mathcal{B}\Rightarrow\|Ae_i\|\leqslant 1 \quad (13.106e)\\
&= \mathrm{Tr}(S). && (13.106f)
\end{aligned}$$

Il faut noter que la première inégalité est stricte, et donc nous avons une contradiction. □

13.2.8 Décomposition de Bruhat

ThoizlYJO

Théorème 13.39 (Décomposition de Bruhat)**.**
Soit $\mathbb{K}$ un corps ; un élément $M \in \mathrm{GL}(n,\mathbb{R})$ s'écrit sous la forme

$$M = T_1 P_\sigma T_2 \tag{13.107}$$

où T_1 et T_2 sont des matrices triangulaires supérieures inversibles et où P_σ est une matrice de permutations $\sigma \in S_n$. De plus il y a unicité de σ.

Démonstration. Afin de rendre les choses plus visuelles, nous nous permettons de donner des exemples au fur et à mesure de la preuve. Nous prenons l'exemple de la matrice

$$\begin{pmatrix} 1 & 3 & 4\\ 2 & 5 & 6\\ 0 & 7 & 8 \end{pmatrix}. \tag{13.108}$$

(1) **Existence** Soit $M \in \mathrm{GL}(n,\mathbb{R})$; puisqu'elle est inversible, on a un indice i_1 maximum tel que $M_{i_1,1} \neq 0$. Nous changeons toutes les lignes jusque là, c'est-à-dire que nous faisons, pour $1 \leqslant i < i_1$,

$$L_i \to L_i - \frac{M_{i1}}{M_{i_1 1}} L_{i_1}. \tag{13.109}$$

EqGHUbwR

Voir le lemme 13.11(3).
Nous avons donc obtenu une matrice dont la première colonne est nulle sauf la case numéro i_1. L'opération (13.109) revient à considérer la multiplication par la matrice de transvection

$$T_1^{(i)} = T_{ii_1}\left(-\frac{M_{i1}}{M_{i_1 1}}\right) \tag{13.110}$$

pour tout $i < i_1$. Pour rappel nous ne changeons que les lignes *au-dessus* de la i_1. Du coup les matrices $T_1^{(i)}$ sont triangulaires supérieures. Nous avons donc la nouvelle matrice $M_1 = \left(\prod_{i<i_1} T_1^{(i)}\right) M$ pour laquelle toute la première colonne est nulle sauf un élément.
Dans le cas de l'exemple, le « pivot » sera la ligne $(2,5,6)$ et la matrice se transforme à l'aide de la matrice $T_1 = T_{12}(-1/2)$:

$$\begin{pmatrix} 1 & -1/2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 4\\ 2 & 5 & 6\\ 0 & 7 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 1\\ 2 & 5 & 6\\ 0 & 7 & 8 \end{pmatrix}. \tag{13.111}$$

EqyjXIYf

22. Définition 8.230.

Maintenant nous faisons de même avec les colonnes (en renommant M la matrice obtenue à l'étape précédente) :

$$C_j \to C_j - \frac{M_{i_1 j}}{M_{i_1 1}} C_1, \tag{13.112}$$

qui revient à multiplier à droite par les matrices $T_{1j}(\frac{M_{i_1 i}}{M_{i_1 1}})$ avec $j > 1$. Encore une fois ce sont des matrices triangulaires supérieures.

Dans l'exemple, pour traiter la seconde colonne, nous multiplions (13.111) à droite par la matrice $T_{12}(-5/2)$:

$$\begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 5 & 6 \\ 0 & 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & -5/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 0 & 6 \\ 0 & 7 & 8 \end{pmatrix}. \tag{13.113}$$

Appliquer encore la matrice $T_{13}(-6/2)$ apporte la matrice

$$\begin{pmatrix} 0 & 1/2 & 1 \\ 2 & 0 & 0 \\ 0 & 7 & 8 \end{pmatrix}. \tag{13.114}$$

Enfin nous multiplions la matrice obtenue par $\frac{1}{M_{i_1 1}} \mathbb{1}$ pour normaliser à 1 l'élément « pivot » que nous avions choisi. Dans notre exemple nous multiplions par $1/2$ pour trouver

$$\begin{pmatrix} 0 & 1/4 & 1/2 \\ 1 & 0 & 0 \\ 0 & 7/2 & 4 \end{pmatrix}. \tag{13.115}$$

Fqduglwu

La matrice obtenue jusqu'ici possède une ligne et une colonne de zéros avec un 1 à leur intersection, et elle est de la forme

$$M' = T_1 M T_2 \tag{13.116}$$

où T_1 et T_2 sont triangulaires supérieures et inversibles, produits de matrices de transvection (et d'une matrice scalaire pour la normalisation).

Il reste à recommencer l'opération avec la seconde colonne (qui n'est pas toute nulle parce que le déterminant est encore non nul) puis la suivante, etc. Dans notre exemple de l'équation (13.115), nous éliminerions le 1/4 et le 4 en utilisant le 7/2.

Encore une fois tout cela se fait à l'aide de matrices supérieures parce qu'à chaque étape, les colonnes précédent le pivot sont déjà nulles (sauf un 1) et ne doivent donc pas être touchées.

À la fin de ce processus, ce qui reste est une matrice TMT' qui ne contient plus qu'un seul 1 sur chaque ligne et chaque colonne, c'est-à-dire une matrice de permutations : $P_\sigma = TMT'$ et donc

$$M = T_\sigma^{-1} (T')^{-1}. \tag{13.117}$$

(2) **Unicité**

Soient $\sigma, \sigma \in S'_n$ tels que $T_1 P_\sigma T_2 = S_1 P_\tau S_2$ avec T_i et S_i triangulaires supérieures et inversibles. En posant $T = T_2 S_2^{-1}$ et $S = T_1^{-1} S_1$, nous avons

$$P_\sigma T = S P_\tau \tag{13.118}$$

où S et T sont des matrices triangulaires supérieures et inversibles. Par les calculs de la preuve du lemme 4.105,

$$\begin{cases} (P_\sigma T)_{kl} = T_{\sigma^{-1}(k)l} & \text{(13.119a)} \\ (SP_\tau)_{kl} = S_{k\tau(l)}, & \text{(13.119b)} \end{cases}$$

et donc

$$T_{\sigma^{-1}(k)l} = S_{k\tau(l)}. \tag{13.120}$$

EqKlmgQT

En écrivant cette équation avec $k = \sigma(i)$ (nous rappelons que σ est bijective),

$$T_{il} = S_{\sigma(i)\tau(l)}. \tag{13.121}$$

Nous savons que les termes diagonaux de T sont non nuls parce que T est triangulaire supérieure et inversible (donc pas de colonnes entières nulles). Nous avons donc, en prenant $i = l = k$,

$$0 \neq T_{kk} = S_{\sigma(k)\tau(k)}. \tag{13.122}$$

La matrice étant triangulaire supérieure, cela implique

$$\sigma(k) \leqslant \tau(k). \tag{13.123}$$ EqEmiBTX

De la même manière en écrivant (13.120) avec $l = \tau^{-1}(i)$,

$$S_{ki} = T_{\sigma^{-1}(k)\tau^{-1}(i)} \tag{13.124}$$

et donc

$$\sigma^{-1}(k) \leqslant \tau^{-1}(k). \tag{13.125}$$

En écrivant cela avec $k = \sigma(j)$, nous avons $j \leqslant \tau^{-1}\sigma(j)$ et en appliquant enfin τ,

$$\tau(j) \leqslant \sigma(j). \tag{13.126}$$

En comparant avec (13.123), nous avons $\sigma = \tau$.

□

13.3 Sous-groupes du groupe linéaire

LemOCtdiaE

Lemme 13.40 ([104]).
Soit V un espace vectoriel de dimension finie muni d'une norme euclidienne $\|.\|$. Soit K un compact convexe de V et G, un sous-groupe compact de $\mathrm{GL}(V)$ tel que

$$u(K) \subset K \tag{13.127}$$

pour tout $u \in G$. Alors il existe $a \in K$ tel que $u(a) = a$ pour tout $u \in G$.

Démonstration. Avant de nous lancer dans la preuve, nous avons besoin d'un petit résultat.

(1) **Un pré-résultat**
Nous commençons par prouver que si $v \in \mathcal{L}(V)$ vérifie $v(K) \subset K$, alors v a un point fixe dans K. Pour cela nous considérons $x_0 \in K$ et la suite

$$x_k = \frac{1}{k+1} \sum_{i=0}^{k} v^i(x_0). \tag{13.128}$$

Étant donné que K est convexe et stable par v, la suite (x_k) est contenue dans K et accepte une sous-suite convergente[23] que nous allons noter $x_{\varphi(n)}$ avec $\varphi\colon \mathbb{N} \to \mathbb{N}$ strictement croissante. Soit $a \in K$ la limite :

$$\lim_{n\to\infty} x_{\varphi(n)} = a. \tag{13.129}$$

Tant que nous y sommes nous pouvons aussi calculer $v(x_k)$:

$$v(x_k) = v\left(\frac{1}{k+1}\sum_{i=1}^{k} v^i(x_0)\right) \tag{13.130a}$$

$$= \frac{1}{k+1}\sum_{i=0}^{k} v^{i+1}(x_0) \tag{13.130b}$$

$$= x_k + \frac{1}{k+1}\Big(v^{k+1}(x_0) - x_0\Big). \tag{13.130c}$$ EqUAfcaKG

23. C'est Bolzano-Weierstrass, théorème 7.144.

La norme $\|v^{k+1}(x_0)-x_0\|$ est bornée par le diamètre de K, donc en prenant la limite $k\to\infty$ le second terme de (13.130c) tend vers zéro. En prenant ces égalités en $k=\varphi(n)$ et en prenant $n\to\infty$, nous trouvons

$$v(a)=a, \tag{13.131}$$

c'est-à-dire le résultat que nous voulions dans un premier temps.

(2) **Une norme sur V**

Nous passons maintenant à la preuve du lemme. D'abord nous remarquons que le groupe G agit sur V par $u\cdot x=u(x)$ et de plus, considérant la fonction continue

$$\begin{aligned}\alpha\colon G&\to V\\ u&\mapsto u(x),\end{aligned} \tag{13.132}$$

nous voyons que les orbites de cette action sont compactes en tant qu'image par α du compact G (théorème 7.220). Nous posons

$$\begin{aligned}\nu\colon V&\to\mathbb{R}^+\\ x&\mapsto\max_{u\in G}\|u(x)\|.\end{aligned} \tag{13.133}$$

Cette définition a un sens parce que l'orbite $\{u(x)\text{ tel que }u\in G\}$ est compacte dans V et donc l'ensemble des normes est compact dans $\mathbb{R}$ et admet un maximum. De plus cela donne une norme sur V parce que nous vérifions les conditions de la définition 7.158 :

(2a) Pour tout $x,y\in V$ nous avons :

$$\nu(x+y)=\max_{u\in G}\|u(x)+u(y)\|\leqslant\max_{u\in G}(\|u(x)\|+\|u(y)\|)\leqslant\nu(x)+\nu(y). \tag{13.134}$$

(2b) Si $\nu(x)=0$, alors l'égalité $\max_{u\in G}\|u(x)\|=0$ nous enseigne que $\|u(x)\|=0$ pour tout $u\in G$ et donc en particulier avec $u=\mathrm{Id}$ nous trouvons $x=0$.

(2c) Pour tout $\lambda\in\mathbb{R}$ et $x\in V$,

$$\nu(\lambda x)=\max_{u\in G}\|u(\lambda x)\|=\max\|\lambda u(x)\|=\max|\lambda|\|u(x)\|=|\lambda|\nu(x). \tag{13.135}$$

De plus la fonction ν est constante sur les orbites de G.

(3) **Un point fixe**

Pour tout $u\in G$ nous posons

$$F_u=\{x\in K\text{ tel que }u(x)=x\}; \tag{13.136}$$

par le pré-résultat, aucun de ces ensembles n'est vide. Ils sont de plus tous fermés par continuité de u (le complémentaire est ouvert). Nous devons prouver que $\bigcap_{u\in G}F_u\neq\varnothing$ parce qu'une intersection serait un point fixe de tous les éléments de G. Supposons donc que $\bigcap_{u\in G}F_u=\varnothing$. Alors les complémentaires des F_u forment un recouvrement ouvert de K et nous pouvons en extraire un sous-recouvrement fini par compacité. Soient $\{u_i\}_{i=1,\dots,p}$ les éléments qui réalisent ce recouvrement. Alors

$$\bigcap_{i=1}^{p}F_{u_i}=\varnothing. \tag{13.137}$$

Nous considérons l'opérateur

$$v=\frac{1}{p}\sum_{i=1}^{p}u_i\in\mathcal{L}(V). \tag{13.138}$$

Puisque K est convexe et stable sous chacun des u_i, nous avons aussi $v(K) \subset K$ et donc il existe $a \in K$ tel que $v(a) = a$. Pour ce a, nous avons

$$\nu\big(v(a)\big) = \nu\left(\frac{1}{p}\sum_{i=1}^{p} u_i(a)\right) \tag{13.139a}$$
$$\leqslant \frac{1}{p}\sum_{i=1}^{p} \nu\left(u_i(a)\right) \tag{13.139b}$$
$$= \frac{1}{p}\sum_{i=1}^{p} \nu(a) \tag{13.139c}$$
$$= \nu(a) \tag{13.139d}$$

où nous avons utilisé la constance de ν sur les orbites de G. Par ailleurs nous savons que $v(a) = a$, donc en réalité à gauche dans (13.139a) nous avons $\nu(a)$ et toutes les inégalités sont des égalités. Nous avons en particulier

$$\nu\left(\sum_{i=1}^{p} u_i(a)\right) = \sum_{i=1}^{p} \nu\left(u_i(a)\right). \tag{13.140}$$

Notons $u_0 \in G$ l'élément qui réalise le maximum de la définition de ν pour le vecteur $\sum_i u_i(a)$:

$$\nu\left(\sum_i u_i(a)\right) = \|u_0\left(\sum_i u_i(a)\right)\| \leqslant \sum_i \|u_0 u_i(a)\| \leqslant \sum_i \nu\big(u_i(a)\big). \tag{13.141}$$

Mais nous venons de voir (équation (13.140)) que l'expression de gauche est égale à celle de droite. Donc les inégalités sont des égalités et en particulier la première inégalité devient l'égalité

$$\|\sum_i u_0 u_i(a)\| = \sum_i \|u_0 u_i(a)\|. \tag{13.142}$$

En vertu du lemme 10.15, il existe des nombres positifs λ_i tels que

$$u_0 u_1(a) = \lambda_2 u_0 u_2(a) = \ldots = \lambda_p u_0 u_p(a). \tag{13.143}$$

Du fait que u_0 est inversible nous avons aussi

$$u_1(a) = \lambda_2 u_2(a) = \ldots = \lambda_p u_p(a). \tag{13.144}$$

Mais par constance de ν sur les orbites nous avons $\nu(u_i(a)) = \nu(u_j(a))$ pour tout i et j ; en appliquant ν à la série d'égalités (13.144), nous trouvons que tous les λ_i doivent être égaux à 1. En particulier

$$u_1(a) = u_2(a) = \ldots = u_p(a). \tag{13.145}$$

Nous récrivons maintenant l'équation $v(a) = a$ avec la définition de v :

$$a = v(a) = \frac{1}{p}\sum_{i=1}^{p} u_i(a) = u_j(a) \tag{13.146}$$

pour n'importe quel j. Donc

$$a \in \bigcap_{i=1}^{p} F_{u_i}, \tag{13.147}$$

ce qui contredit notre hypothèse de départ.

□

Proposition 13.41 ([430, 104, 431]).
Soit G un sous-groupe compact de $\mathrm{GL}(n, \mathbb{R})$. Alors

(1) Il existe une forme quadratique définie positive q sur $\mathbb{R}^n$ telle que $G \subset \mathrm{O}(q)$.

(2) Le groupe G est conjugué à un sous-groupe de $\mathrm{O}(n,\mathbb{R})$.

Démonstration. Nous considérons le (pas tout à fait) morphisme de groupe

$$\begin{aligned}\rho\colon G &\to \mathrm{GL}\big(\mathrm{S}(n,\mathbb{R})\big)\\ u &\mapsto \rho_u\colon s \mapsto u^t s u,\end{aligned} \tag{13.148}$$

et tant que nous en sommes à considérer, nous considérons l'ensemble

$$H = \{M^tM \text{ tel que } M \in G\} \subset \mathrm{S}(n,\mathbb{R}). \tag{13.149}$$

Cet ensemble est constitué de matrices définies positives parce que si $\langle M^tMx, x\rangle = 0$, alors $0 = \langle Mx, Mx\rangle = \|Mx\|$, mais M étant inversible, cela implique que $x = 0$. Qui plus est, cet ensemble est compact dans $\mathrm{GL}(n,\mathbb{R})$ en tant qu'image du compact G par l'application continue $M \mapsto M^tM$. L'enveloppe convexe $K = \mathcal{C}\mathrm{onv}(H)$ est alors également compacte par le théorème 11.42. Enfin nous considérons $L = \rho(G)$, qui est un sous-groupe compact de $\mathrm{GL}\big(\mathrm{S}(n,\mathbb{R})\big)$ parce que $\rho_u\rho_v = \rho_{vu} \in \rho(G)$. Nous remarquons que ρ_u étant linéaire, elle préserve les combinaisons convexes et donc pour tout $u \in G$, $\rho_u(K) \subset K$.

Bref, L est un sous-groupe compact de $\mathrm{GL}(n,\mathbb{R})$ préservant le compact K de $\mathrm{S}(n,\mathbb{R})$. Par le lemme 13.40, il existe $s \in K$ tel que $\rho_u(s) = s$ pour tout $u \in G$. Ou encore :

$$u^t s u = s \tag{13.150}$$

pour tout $u \in G$. Fort de ce s bien particulier, nous considérons la forme quadratique associée : $q(x) = x^t s x$. Cette forme est définie positive parce que s l'est. Nous avons $G \subset \mathrm{O}(q)$ parce que si $u \in G$ alors

$$q\big(ux\big) = (ux)^t s u x = x^t \underbrace{u^t s u}_{=s} x = q(x). \tag{13.151}$$

Le premier point est prouvé.

La matrice s est symétrique et définie positive. Le théorème 8.227 nous permet donc de la diagonaliser en $\mathrm{diag}(\lambda_1,\dots,\lambda_n)$ avec $\lambda_i > 0$, et ensuite transformée en la matrice $\mathbb{1}_n$ par la matrice $\mathrm{diag}(1/\sqrt{\lambda_i})$. Nous avons donc une matrice $a \in \mathrm{GL}(n,\mathbb{R})$ telle que $a^tsa = \mathbb{1}_n$. Avec ça, si $u \in G$, nous avons

$$(a^{-1}ua)^t(a^{-1}ua) = (a^{-1}ua)^t\mathbb{1}_n(a^{-1}ua) = a^tu^t(a^t)^{-1}a^tsaa^{-1}ua = a^tu^tsua = a^tsa = \mathbb{1}, \tag{13.152}$$

ce qui prouve que $a^{-1}ua$ est dans $\mathrm{O}(n,\mathbb{R})$, et donc que $a^{-1}Ga \subset \mathrm{O}(n,\mathbb{R})$. □

Chapitre 14

Tribus, théorie de la mesure, intégration

14.1 Tribus

Vous pouvez vous reporter au thème 23 pour voir plus vite où sont les définitions associées.

14.1.1 Généralités

DefjRsGSy

Définition 14.1 (Tribu, espace mesurable[432])**.**
*Si Ω est un ensemble, un ensemble $\mathcal{A}$ de sous-ensembles de Ω est une **tribu** si*

(1) $\Omega \in \mathcal{A}$;

(2) $A^c \in \mathcal{A}$ pour tout $A \in \mathcal{A}$;

ItemooPEQNooYiYNtN

(3) si $(A_i)_{i\in\mathbb{N}}$ est une suite dénombrable d'éléments de $\mathcal{A}$, alors $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$.

*Le couple $(\Omega, \mathcal{A})$ est alors un **espace mesurable**.*

Remarque 14.2.
Nous trouvons parfois la notation

$$\bigcup_{k\in\mathbb{N}} A_k = \sup_{k\geqslant 0} A_k. \tag{14.1}$$

LemBWNlKfA

Lemme 14.3.
Opérations ensemblistes sur les tribus.

ITEMooTDXNooFszBzi

(1) Une tribu est stable par intersections au plus dénombrables.

ItemXQVLooFGBQNj

(2) Une tribu est stable par différence ensembliste.

Démonstration. Soit $(A_i)_{i\in I}$ une famille au plus dénombrable d'éléments de la tribu $\mathcal{A}$. Nous devons prouver que $\bigcap_{i\in I} A_i$ est également un élément de $\mathcal{A}$. Pour cela nous passons au complémentaire :

$$\complement\left(\bigcap_{i\in I} A_i\right) = \bigcup_{i\in I} \complement A_i. \tag{14.2}$$

La définition d'une tribu implique que le membre de droite est un élément de la tribu. Par stabilité d'une tribu par complémentaire, l'ensemble $\bigcap_{i\in I} A_i$ est également un élément de la tribu.

La seconde assertion est immédiate à partir de la première parce que $A\backslash B = A \cap \complement B$. □

Si $(\mathcal{A}_i)_{i\in I}$ est un ensemble de tribus (indexé par un ensemble I quelconque) alors

$$\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i \tag{14.3}$$

est également une tribu.

14.1.2 Tribu engendrée

DEFooJSAKooSDnGzt

Définition 14.4.
*Soit $\mathcal{D}$ un ensemble de parties de Ω. La **tribu engendrée** par $\mathcal{D}$ est l'intersection de toutes les tribus de Ω contenant $\mathcal{D}$. C'est la plus petite tribu contenant $\mathcal{D}$. Nous la noterons le plus souvent $\sigma(\mathcal{D})$*

LEMooCAUMooEdpjRs

Lemme 14.5.
Si $(\Omega, \mathcal{A})$ est un espace mesurable et si A est une partie mesurable de Ω, alors la tribu engendrée[1] *par A est*

$$\sigma(A) = \{\emptyset, A, \complement A, \Omega\}. \tag{14.4}$$

DefNOJWooLGKhmJ

Lemme-Définition 14.6 (Tribu engendrée[433])**.**
Soit une application $f : S_1 \to S_2$ et $\mathcal{F}$ une tribu de S_1. Alors

(1) L'ensemble

$$\mathcal{F}_f = \{B \subset S_2 \text{ tel que } f^{-1}(B) \in \mathcal{F}\} \tag{14.5}$$

est une tribu sur S_2.

(2) C'est la plus grande tribu de S_2 pour laquelle f est mesurable.

*Cette tribu est la est la **tribu engendrée** par f. Elle sera souvent notée $\sigma(f)$.*

Démonstration. Vérifier les trois propriétés de la définition 14.1.

(1) $S_2 \in \mathcal{F}$ parce que $f^{-1}(S_2) = S_1 \in \mathcal{F}_f$.

(2) Si $B \in \mathcal{F}_f$, alors nous prouvons que $B^c \in \mathcal{F}_f$. Nous avons $f^{-1}(B) \in \mathcal{F}$ et donc $f^{-1}(B)^c \in \mathcal{F}$. Pour conclure, il suffit de remarquer que $f^{-1}(B^c) = f^{-1}(B)^c$.

(3) Si $(B_i) \in \mathcal{F}_f$, alors

$$f^{-1}\left(\bigcup_i B_i\right) = \bigcup_i f^{-1}(B_i) \in \mathcal{F}_f. \tag{14.6}$$

Donc $\bigcup_i B_i \in \mathcal{F}_f$.

En ce qui concerne la maximalité, si $R \subset S_2$ n'est pas dans $\mathcal{F}_f$ alors $f^{-1}(R)$ n'est pas dans $\mathcal{F}$ et donc f ne serait pas mesurable. □

PropHYLooLgOCy

Proposition 14.7 ([434])**.**
Soit S un ensemble et $\mathcal{F}$ une tribu de S. Soit une classe $\mathcal{N} \subset \mathcal{P}(S)$ telle que

(1) Si $A \in \mathcal{N}$ alors il existe $Y \in \mathcal{F} \cap \mathcal{N}$ tel que $A \subset Y$.

(2) Si $A \in \mathcal{N}$ et $B \subset A$ alors $B \in \mathcal{N}$.

(3) La classe $\mathcal{N}$ est stable par union dénombrable.

Alors la classe

$$\mathcal{T} = \{X \cup A \text{ avec } A \in \mathcal{N} \text{ et } X \in \mathcal{F}\} \tag{14.7}$$

est une tribu.

Démonstration. L'ensemble $\mathcal{T}$ est stable par union dénombrable parce que $\mathcal{F}$ et $\mathcal{N}$ le sont. De plus S et $\emptyset$ sont dans $\mathcal{F}$ et donc dans $\mathcal{T}$. Nous devons voir que $\mathcal{T}$ est stable par complémentarité.

Soit donc $A \in \mathcal{N}$ et $X \in \mathcal{F}$; nous savons que $(A \cup X)^c = A^c \cap X^c$. De plus il existe $Y \in \mathcal{F} \cap \mathcal{N}$ tel que $A \subset Y$ et nous pouvons exprimer A^c en termes de Y : $A^c = Y^c \cup (Y \backslash A)$. Donc

$$(A \cup X)^c = \big(Y^c \cup (Y \backslash A)\big) \cap X^c = \underbrace{(Y^c \cap X^c)}_{\in \mathcal{F}} \cup \underbrace{\big((Y \backslash A) \cap X^c\big)}_{\in \mathcal{N}}. \tag{14.8}$$

Le fait que la seconde partie soit dans $\mathcal{N}$ est due au fait que ce soit une partie de $Y \in \mathcal{N}$. Nous avons donc bien $(A \cup X)^c \in \mathcal{T}$. □

1. Tribu engendrée par une partie, définition 14.4.

14.1.3 Tribu induite

DefDHTTooWNoKDP

Proposition-Définition 14.8 (Tribu induite, tribu-trace[435]).
Soit un espace mesurable $(S, \mathcal{F})$ et une partie $R \subset S$. L'ensemble

EQooNOQNooYIdDOz

$$\mathcal{F}_R = \{A \cap R \text{ tel que } A \in \mathcal{F}\} \tag{14.9}$$

*est une tribu. Elle est la **tribu induite** de R depuis S. Elle est aussi nommée « tribu trace ».*

Démonstration. D'abord R et $\varnothing$ dont dans $\mathcal{F}_R$. Si $C \in \mathcal{F}_R$ alors $C = A \cap R$ pour un certain $A \in \mathcal{F}$ et nous devons prouver que $R \cap C^c$ est dans $\mathcal{F}_R$ (le complémentaire de C dans R). Nous avons

$$R \cap C^c = R \cap (A \cap R)^c = R \cap A^c \in \mathcal{F}_R \tag{14.10}$$

parce que $A^c \in \mathcal{F}$. Enfin si $C_i \in \mathcal{F}_R$ alors $C_i = R \cap A_i$ pour des A_i dans $\mathcal{F}$. Nous avons

$$\bigcup_{i \in \mathbb{N}} C_i = \bigcup_{i \in \mathbb{N}} (R \cap A_i) = R \cap \big(\bigcup_{i \in \mathbb{N}} A_i\big), \tag{14.11}$$

mais $\bigcup_i A_i \in \mathcal{F}$ donc $\bigcup_i C_i \in \mathcal{F}_R$. □

PROPooUNNSooMUQKfp

Proposition 14.9.
Si R est mesurable dans $(\Omega, \mathcal{A})$ alors

$$\mathcal{A}_R = \{A \cap R \text{ tel que } A \in \mathcal{A}\} = \{S \in \mathcal{A} \text{ tel que } S \subset R\}, \tag{14.12}$$

où $\mathcal{A}_R$ est la tribu induite de $\mathcal{A}$ sur R.

Démonstration. La première égalité est simplement la définition (14.9) de la tribu induite. Pour le reste, nous notons $\mathcal{F} = \{S \in \mathcal{A} \text{ tel que } S \subset R\}$, et nous prouvons que $\mathcal{A}_R = \mathcal{F}$.

(1) $\underline{\mathcal{F} \subset \mathcal{A}_R}$ Si $S \in \mathcal{F}$, alors $S \in \mathcal{A}$ et $S \subset R$. Donc $S = S \cap R \in \mathcal{A}_R$.

(2) $\underline{\mathcal{F}_R \subset \mathcal{A}_R}$ Dans l'autre sens, si $S \in \mathcal{A}_R$, alors il existe $A \in \mathcal{A}$ tel que $S = A \cap R$. Donc $S \subset A$ et $S \in \mathcal{A}$ parce que R et A sont des éléments de $\mathcal{A}$ (stable par intersection).

□

14.2 Théorie de la mesure

DefSLWooleMaif

Définition 14.10 ([436]).
*Une **mesure extérieure** sur un ensemble S est une application $m^* \colon \mathcal{P}(S) \to [0, \infty]$ telle que*

(1) $m^*(\varnothing) = 0$,

(2) Si $A \subset B$ dans S alors $m^*(A) \leqslant m^*(B)$

ItemARKooppZfDaiii

(3) Si les A_n sont des parties de S alors

EqZLMooSxvaL

$$m^*\big(\bigcup_{n \in \mathbb{N}} A_n\big) \leqslant \sum_{n \in \mathbb{N}} m^*(A_n). \tag{14.13}$$

La différence avec une mesure est que nous ne demandons pas que (14.13) soit une égalité lorsque les A_n sont disjoints.

14.2.1 Mesure sur un ensemble de parties

DefWUPHooEklLmR

Définition 14.11 (Mesure sur un ensemble de parties[437]).
*Soient S un ensemble et $\mathcal{C}$ un ensemble de parties de S contenant $\varnothing$. Une **mesure positive** sur $(S, \mathcal{C})$ est une application $\mu \colon \mathcal{C} \to [0, \infty]$ telle que*

(1) $\mu(\varnothing) = 0$,

(2) Si $A_n \in \mathcal{C}$ sont des éléments deux à deux disjoints dans $\mathcal{C}$ et tels que $\bigcup_n A_n \in \mathcal{C}$ alors

$$\mu\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \sum_{n=0}^{\infty} \mu(A_n). \tag{14.14}$$

*La mesure est **finie** si $\mu(S) < \infty$ et σ**-finie** si il existe une suite (S_n) dans $\mathcal{C}$ telle que $S = \bigcup_n S_n$ et $\mu(S_n) < \infty$.*

Remarque 14.12.
La condition $\mu(\varnothing) = 0$ est nécessaire. Certes, si $A \in \mathcal{A}$ nous avons

$$\mu(A) = \mu(A \cup \varnothing) = \mu(A) + \mu(\varnothing) \tag{14.15}$$

EQooCDSQooZMjgQY

parce que A et $\varnothing$ sont disjoints. Cela semble indiquer que $\mu(\varnothing) = 0$, mais pas tout à fait : il est encore possible d'avoir $\mu(B) = \infty$ pour tout $B \in \mathcal{A}$, y compris $\mu(\varnothing) = \infty$. À cause de cette exception, la relation (14.15) n'implique pas $\mu(\varnothing) = 0$.

Sans hypothèse sur l'ensemble de parties considéré, nous ne pouvons pas dire grand chose de plus.

14.2.2 Mesure sur une algèbre de parties

DefTCUoogGDud

Définition 14.13 (Algèbre de parties[436])**.**
*Soit S, un ensemble. Une classe $\mathcal{D}$ de parties de S est une **algèbre de parties** de S si*

(1) $S \in \mathcal{D}$ et $\varnothing \in \mathcal{D}$,

(2) si $A \in \mathcal{D}$ alors $A^c \in \mathcal{D}$,

(3) si $A, B \in \mathcal{D}$ alors $A \cup B \in \mathcal{D}$.

14.14.
En anglais, ce sont des *field of sets*[438].

Les algèbres de parties ne sont pas des classes si sauvages que ça ; en témoigne le lemme suivant.

LemBFKootqXKl

Lemme 14.15.
Une algèbre de partie est stable par intersection (finie) et par différence ensembliste.

Démonstration. Il suffit de remarquer que $A \cap B = \big(A^c \cup B^c\big)^c$ et que $A \backslash B = A \cap B^c$. □

LemZQUooMdCpq

Lemme 14.16 ([436])**.**
Si $\mathcal{D}$ est une algèbre de parties de S et si μ est une mesure sur $(S, \mathcal{D})$ alors

(1) si $A, B \in \mathcal{D}$ avec $A \subset B$ alors $\mu(A) \leqslant \mu(B)$

ItemMFUooWCPNnii

(2) si $A_n \in \mathcal{D}$ et $\bigcup_n A_n \in \mathcal{D}$ alors

$$\mu\Big(\bigcup_n A_n\Big) \leqslant \sum_n \mu(A_n). \tag{14.16}$$

La propriété (2) est la σ**-sous-additivité**.

Démonstration. Si $A \subset B$ alors $B = A \cup (B \backslash A)$ avec A et $B \backslash A$ disjoints donc

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geqslant \mu(A). \tag{14.17}$$

Pour la seconde, on passe par les compléments deux à deux : nous posons

$$\begin{cases} B_0 = \varnothing & \text{(14.18a)} \\ B_n = A_n \backslash \bigcup_{k<n} B_k. & \text{(14.18b)} \end{cases}$$

Ces ensembles sont deux à deux disjoints et $\bigcup_n B_n = \bigcup_n A_n \in \mathcal{D}$, donc

$$\mu\big(\bigcup_n A_n\big) = \sum_n \mu\big(A_n \setminus \bigcup_{k<n} B_k\big) \leqslant \sum_n \mu(A_n), \tag{14.19}$$

où nous avons utilisé la première partie du lemme. □

PropIUOoobjfIB

Proposition 14.17 (Mesure extérieure à partir d'une algèbre de parties[436])**.**
Soient $\mathcal{D}$ une algèbre de partie sur l'ensemble S et μ une mesure sur $(S, \mathcal{D})$. Alors l'application

$$\begin{aligned}\mu^* : \mathcal{P}(S) &\to [0, +\infty] \\ X &\mapsto \inf\left\{\sum_n \mu(A_n) \text{ tel que } A_n \in \mathcal{D}, X \subset \bigcup_n A_n\right\}\end{aligned} \tag{14.20}$$

EqRNJooQrcoa

est une mesure extérieure[2] *sur S et pour tout $A \in \mathcal{D}$ nous avons $\mu^*(A) = \mu(A)$.*

Démonstration. En plusieurs parties.

(1) **La définition est bonne** L'ensemble sur lequel l'infimum est pris n'est pas vide : il suffit de prendre $A_1 = S$ et $A_{n\geqslant 2} = \varnothing$.

(2) **Le vide** D'abord $\mu^*(\varnothing) = 0$ parce que $\varnothing \in \mathcal{D}$. Prendre ensuite $A_n = \varnothing$.

(3) **μ^* est croissante** Soit $X \subset Y$ dans $\mathcal{P}(S)$. Si une suite (A_n) dans $\mathcal{D}$ vérifie $Y \subset \bigcup_n A_n$, alors la même suite vérifie $X \subset \bigcup_n A_n$. Par conséquent nous avons l'inclusion

$$\left\{\sum_n \mu(A_n) \text{ tel que } A_n \in \mathcal{D}, X \subset \bigcup_n A_n\right\} \subset \left\{\sum_n \mu(A_n) \text{ tel que } A_n \in \mathcal{D}, Y \subset \bigcup_n A_n\right\}, \tag{14.21}$$

et donc l'inégalité $\mu^*(X) \leqslant \mu^*(Y)$.

(4) **Inégalité par union dénombrable** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de parties de S. Si il existe n_0 tel que $\mu^*(X_{n_0}) = \infty$ alors nous avons automatiquement $\sum_n \mu^*(X_n) = \infty$ et l'inégalité demandée est évidente parce que n'importe quel nombre est plus petit ou égal à ∞. Nous supposons donc maintenant que $\mu^*(X_n) < \infty$ pour tout n.

Soit $\epsilon > 0$. Pour tout $n \geqslant 1$ il existe une suite $(B_k^{(n)})_{k\in\mathbb{N}}$ dans $\mathcal{D}$ telle que $X_n \subset \bigcup_k B_k^{(n)}$ et

$$\mu^*(X_n) + \frac{\epsilon}{2^n} \geqslant \sum_k \mu(B_k^{(n)}). \tag{14.22}$$

Étant donné que

$$\bigcup_n X_n \subset \bigcup_n \big(\bigcup_k B_k^{(n)}\big), \tag{14.23}$$

nous avons[3]

$$\mu^*\big(\bigcup_n X_n\big) \leqslant \sum_n \sum_k \mu^*(B_k^{(n)}) \leqslant \sum_n \Big(\mu^*(X_n) + \frac{\epsilon}{2^n}\Big) = \sum_n \mu^*(X_n) + \epsilon. \tag{14.24}$$

Cette inégalité étant valable pour tout ϵ, nous avons bien

$$\mu^*\big(\bigcup_n X_n\big) = \sum_n \mu^*(X_n). \tag{14.25}$$

(5) **Restriction** Soit $A \in \mathcal{D}$. Nous avons automatiquement $\mu^*(A) \leqslant \mu(A)$ parce que $\mu(A)$ est dans l'ensemble dont nous prenons l'infimum (prendre $A_1 = A$ et $A_{n\geqslant 2} = \varnothing$).

2. Définition 14.10.
3. Nous utilisons la somme $\sum_{n\in\mathbb{N}}(1/2^n) = 1$, proposition 10.135(2).

En ce qui concerne l'inégalité inverse nous considérons une suite A_n dans $\mathcal{D}$ telle que $A \subset \bigcup_n A_n$. Étant donné que $A \in \mathcal{D}$ et que $\mathcal{D}$ est une algèbre de parties nous avons $A \cap A_n \in \mathcal{D}$ et $\bigcup_n (A \cap A_n) = A \in \mathcal{D}$. Par conséquent

$$\mu(A) = \mu\Big(\bigcup_n (A \cap A_n)\Big) \leqslant \sum_n \mu(A \cap A_n) \leqslant \sum_n \mu(A_n). \tag{14.26}$$

Donc tous les éléments de l'ensemble sur lequel nous prenons l'infimum sont plus grands que $\mu(A)$. Nous en déduisons que $\mu^*(A) \geqslant \mu(A)$.

□

14.2.3 Mesure sur une tribu, espace mesuré

La définition suivante est une simple copie de la définition générale 14.11 d'une mesure sur un ensemble de parties. La seule différence est que l'union d'éléments d'une tribu est encore dans la tribu, et la condition (2) ne demande pas de le préciser.

DefBTsgznn

Définition 14.18 (Espace mesuré[437, 439])**.**
Soit un espace mesurable[4] $(\Omega, \mathcal{A})$*.*

Une ***mesure signée*** *sur* $(\Omega, \mathcal{A})$ *est une application* $\mu\colon \mathcal{A} \to]-\infty, \infty]$ *telle que*

(1) $\mu(\varnothing) = 0$,

ItemQFjtOjXiii

(2)

$$\mu\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mu(A_i) \tag{14.27}$$

EQooKJHQooHIQyUi

si les A_i *sont des éléments de* $\mathcal{A}$ *deux à deux disjoints.*

La mesure signée μ *est une* ***mesure positive*** *si elle prend ses valeurs dans* $[0, \infty]$*.*

Lorsque μ *est une mesure positive, le triplet* $(\Omega, \mathcal{A}, \mu)$ *est un* ***espace mesuré****.*

Une mesure est ***σ-finie*** *si il existe un recouvrement dénombrable de* Ω *par des ensembles de mesure finie. Si la mesure est* σ*-finie, nous disons que l'espace* $(\Omega, \mathcal{A}, \mu)$ *est un espace mesuré* σ*-fini.*

La mesure μ *sur* Ω *est* ***finie*** *si* $\mu(\Omega) < \infty$*.*

LEMooECDQooEBtMxL

Lemme 14.19 ([1])**.**
Soit un espace mesuré[5] $(\Omega, \mathcal{A}, \mu)$*. Si* $Z \in \mathcal{A}$ *est une partie de mesure nulle, et si* $A \subset Z$ *est mesurable alors* $\mu(A) = 0$*.*

Démonstration. Nous pouvons écrire $Z = A \cup (Z \backslash A)$. Vu que A et $Z \backslash A$ sont disjoints, nous avons

$$0 = \mu(Z) = \mu(A) + \mu(Z \backslash A). \tag{14.28}$$

La seule façon pour deux nombres positifs d'avoir une somme nulle est qu'ils soient tous les deux nuls. □

14.2.4 Mesures sigma-finies

LEMooXOPMooMmIbmD

Lemme 14.20 ([440])**.**
Soient deux mesures σ*-finies*[6] μ *et* λ *sur le même espace mesurable* $(\Omega, \mathcal{A})$*. Il existe des parties mesurables* Z_n *telles que*

(1) $\mu(Z_n) < \infty$,

(2) $\lambda(Z_n) < \infty$

4. Les définitions de tribus et d'espaces mesurables sont en 14.1.
5. Définition 14.18.
6. Définition 14.18.

(3) $\bigcup_{n=1}^{\infty} Z_n = \Omega$.

Nous pouvons de plus choisir ces Z_n pour qu'ils soient disjoints.

Démonstration. Le fait que les mesures soient σ-finies donne des parties mesurables (X_n) et(Y_n) telles que $\mu(X_n) < \infty$, $\lambda(Y_n) < \infty$, $\bigcup_n X_n = \bigcup_n Y_n = \Omega$.

Pour chaque $i, j \in \mathbb{N}^2$ nous posons $Z_{ij} = X_i \cap Y_j$. Nous avons $\lambda(Z_{ij}) \leqslant \lambda(Y_j) < \infty$. De même $\mu(Z_{ij}) \leqslant \mu(X_i) < \infty$. Et enfin

$$\bigcup_{ij} Z_{ij} = \bigcup_{n} (X_i \cap Y_j) = \bigcup_{i} \Big(\bigcup_{j} (X_i \cap Y_j) \Big) = \bigcup_{i} X_i = \Omega. \tag{14.29}$$

Comme $\mathbb{N}^2$ est dénombrable[7], nous considérons une bijection $\varphi\colon \mathbb{N} \to \mathbb{N}^2$ et les ensembles $Z'_i = Z_{\varphi(i)}$ répondent à la question.

En ce qui concerne la possibilité que les (Z_n) soient disjoints, supposons qu'ils ne le soient pas. Alors il suffit de poser $Z'_n = X_n \backslash \bigcup_{k=1}^{n-1} Z'_k$ pour que les Z'_n aient les bonnes propriétés tout en étant disjoints. □

14.2.5 Suite du texte

14.21 ([1])**.**
Dans le cas d'une mesure signée, μ prend ses valeurs dans $]-\infty, \infty]$ et non dans $[-\infty, \infty]$ parce que nous ne voulons pas avoir des sommes contenant $\infty - \infty$.

Notez aussi que la condition (14.27) interdit d'avoir, pour un $\alpha > 0$ fixé, une infinité de parties disjointes de mesure plus petite que α.

Si $(\Omega, \mathcal{A}, \mu)$ et $(S, \mathcal{F}, \nu)$ sont deux espaces mesurés, alors nous notons

$$(\Omega, \mathcal{A}, \mu) \subset (S, \mathcal{F}, \nu) \tag{14.30}$$

lorsque $\Omega \subset S$, $\mathcal{A} \subset \mathcal{F}$ et pour tout $A \in \mathcal{A}$, $\mu(A) = \nu(A)$.

DefHGsQxHB

Définition 14.22 (Ensemble mesurable)**.**
*Les éléments de $\mathcal{A}$ sont les ensembles **mesurables** pour la mesure μ.*

Si la mesure est σ-finie, nous pouvons choisir le recouvrement croissant pour l'inclusion. En effet si $(E_n)_{n\in\mathbb{N}}$ est le recouvrement, il suffit de considérer $F_n = \bigcup_{k\leqslant n} E_k$. Ces ensembles F_n forment tout autant un recouvrement dénombrable, mais ils vont évidemment croissants.

Le lemme suivant complète la propriété 14.18(2) lorsque les ensembles ne sont pas disjoints.

LemPMprYuC

Lemme 14.23 (Unions et différences ensemblistes dans un espace mesurable)**.**
Soit un espace mesuré $(\Omega, \mathcal{F}, \mu)$.

ITEMooSUIRooNDVOoB

(1) Soient $A, B \in \mathcal{F}$ avec $A \subset B$. Alors

$$\mu(B\backslash A) = \mu(B) - \mu(A) \tag{14.31}$$

ITEMooLEGKooWnYmlf

(2) Si $A, B \in \mathcal{F}$ avec $A \subset B$, alors

$$\mu(A) \leqslant \mu(B). \tag{14.32}$$

ITEMooMCNBooRGVGqA

(3) Si $A, B \in \mathcal{F}$ avec $A \subset B$, et si $\mu(B) < \infty$, alors $\mu(A) < \infty$.

ITEMooABPYooFQEzqE

(4) Si (M_n) est une suite d'éléments de $\mathcal{F}$ pas spécialement disjoints, alors

$$\mu\Big(\bigcup_k M_k\Big) \leqslant \sum_k \mu(M_k). \tag{14.33}$$

EqWWFooYPCTt

Démonstration. Point par point.

7. L'ensemble $\mathbb{N}$ est dénombrable par 1.132 et $\mathbb{N}^2$ l'est pas 1.135.

(1) **Pour (1)** Nous décomposons $B = A \cup (B\backslash A)$ en remarquant que l'union est disjointe et que $B\backslash A \in \mathcal{F}$ par le lemme 14.3. La propriété (2) de la définition de mesure nous donne alors

$$\mu(B) = \mu\big((B\backslash A) \cup A\big) = \mu(B\backslash A) + \mu(A) \tag{14.34}$$

et donc

$$\mu(B) = \mu(B\backslash A) + \mu(A) \tag{14.35}$$

EQooTWCWooGGQRiZ

comme demandé.

(2) **Pour (2)** Il s'agit de reprendre (14.35) :

$$\mu(A) = \mu(B) - \mu(B\backslash A) \leqslant \mu(B). \tag{14.36}$$

EQooXGVQooRucvoF

(3) **Pour (3)** Il s'agit de continuer (14.36) :

$$\mu(A) \leqslant \mu(B) < \infty. \tag{14.37}$$

(4) **Pour (4)** Nous considérons la suite disjointe

$$\begin{cases} M'_0 = \varnothing & \text{(14.38a)} \\ M'_k = M_k\backslash M'_{k-1}. & \text{(14.38b)} \end{cases}$$

Nous avons $\bigcup_k M'_k = \bigcup_k M_k$. Nous avons alors le calcul suivant :

$$\mu\Big(\bigcup_k M_k\Big) = \mu\Big(\bigcup_k M'_k\Big) = \sum_k \mu(M'_k) = \sum_k \mu(M_k\backslash M'_{k-1}) \leqslant \sum_k \mu(M_k). \tag{14.39}$$

La dernière inégalité utilise le point (2).

□

LemAZGByEs

Lemme 14.24 ([1])**.**
Résultats sur les unions croissantes d'ensembles mesurables dans $(S, \mathcal{A}, \mu)$.

ItemJWUooRXNPci

(1) Si (A_k) est une suite croissante d'ensembles μ-mesurables dont l'union est mesurable, alors

$$\lim_{n\to\infty} \mu(A_k) = \mu\Big(\bigcup_k A_k\Big). \tag{14.40}$$

ItemJWUooRXNPcii

(2) Soit K_n, une suite emboîtée d'éléments de $\mathcal{A}$ tels que $K_n \to S$. Si $A \in \mathcal{A}$ alors

$$\lim_{n\to\infty} \mu(A \cap K_n) = \mu(A). \tag{14.41}$$

Démonstration. Pour prouver (1), nous faisons le coup de l'union télescopique, en posant $A_0 = \varnothing$:

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} (A_k\backslash A_{k-1}). \tag{14.42}$$

Les ensembles $A_k\backslash A_{k-1}$ sont deux à deux disjoints, donc la propriété (2) de la définition d'une mesure donne

$$\begin{aligned} \mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) &= \mu\left(\bigcup_{k=1}^{\infty} (A_k\backslash A_{k-1})\right) && \text{(14.43a)} \\ &= \sum_{k=1}^{\infty} \mu(A_k\backslash A_{k-1}) && \text{(14.43b)} \\ &= \sum_{k=1}^{\infty} \big(\mu(A_k) - \mu(A_{k-1})\big) && \text{(14.43c)} \\ &= \lim_{k\to\infty} \mu(A_k) - \mu(A_0) && \text{(14.43d)} \\ &= \lim_{k\to\infty} \mu(A_k). && \text{(14.43e)} \end{aligned}$$

subEqMDRRorb

où pour obtenir 14.43c, nous avons utilisé le lemme 14.23.

Le point (2) est une application du point (1). □

Exemple 14.25.
L'intégration « à la Riemann » n'est pas dans la théorie des espaces mesurés. En effet l'ensemble

$$\mathcal{A} = \{A \subset [0,1] \text{ tel que } \mathbb{1}_A \text{ est intégrable au sens de Riemann}\} \tag{14.44}$$

n'est pas une tribu. Par exemple les singletons en font partie tandis que $[0,1] \cap \mathbb{Q}$ n'en fait pas partie bien que ce soit une union dénombrable de singletons. △

Définition 14.26.
*Si μ est une mesure nous disons qu'une propriété est vraie μ-**presque partout** si elle est fausse seulement sur un ensemble de mesure nulle.*

Par exemple la fonction de Dirichlet est presque partout égale à la fonction 1 (pour la mesure de Lebesgue).

Définition 14.27 (fonction mesurable)**.**
Une application entre espace mesurés

$$f : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}') \tag{14.45}$$

*est **mesurable** si pour tout $B \in \mathcal{A}'$, l'ensemble $f^{-1}(B)$ est dans $\mathcal{A}$.*

LemIDITgAy

Lemme 14.28.
Une union dénombrable de parties de mesure nulle est de mesure nulle.

Démonstration. C'est une conséquence du lemme 14.23(4) : si les A_i sont de mesure nulle,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mu(A_i) = 0 \tag{14.46}$$

□

Définition 14.29.
*Si (A_n) est une suite croissante d'ensembles alors la **limite** est*

$$\lim_n A_n = \bigcup_{i=0}^{\infty} A_i. \tag{14.47}$$

Si la suite est décroissante alors la limite est

$$\lim_n A_n = \bigcap_{i=0}^{\infty} A_i. \tag{14.48}$$

Pour une suite ni croissante ni décroissante d'ensembles, il y a la notion de limite inductive [8] qui sera un peu traitée à la section 47.2.

PropAFNPSsm

Proposition 14.30 ([441])**.**
Soient μ une mesure sur Ω et (S_n) une suite croissante d'ensembles μ-mesurables de Ω. Nous notons

$$S = \lim_n S_n. \tag{14.49}$$

Alors pour tout ensemble mesurable [9] $A \subset \Omega$ nous avons

$$\mu(A \cap S) = \lim_{n\to\infty} \mu(A \cap S_n). \tag{14.50}$$

8. *direct limit* en anglais.
9. Définition 14.22

Démonstration. L'inégalité $\lim \mu(A \cap S_n) \leqslant \mu(A \cap S)$ est simple à prouver. En effet pour tout n nous avons $A \cap S_n \subset A \cap S$ et donc par le lemme 14.23 nous avons

$$\mu(A \cap S_n) \leqslant \mu(A \cap S). \tag{14.51}$$

En passant à la limite (qui respecte les inégalités) nous avons l'inégalité.

Nous passons à l'inégalité dans l'autre sens.

(1) **Si ∞** D'abord si $\mu(A \cap S_n) = \infty$ pour un certain n, alors cela vaut encore ∞ pour tous les n suivants, et la limite est ∞ sans problème. Donc nous supposons que $\mu(A \cap S_n) < \infty$ pour tout $n \in \mathbb{N}$.

(2) **Une petite hypothèse en plus** Quitte à renommer les indices, nous supposons que $S_0 = \varnothing$.

(3) **S comme union de différences** Nous montrons à présent que $S = \bigcup_n (S_{n+1} \backslash S_n)$. Soit $x \in S$. Il existe $n \geqslant 0$ tel que $x \in S_n$. Et vu que $S_0 = \varnothing$, il existe même un $n \geqslant 0$ tel que $x \notin S_n$ et $x \in S_{n+1}$. Autrement dit, pour tout $x \in S$, il existe $n \geqslant 0$ tel que $x \in S_{n+1} \backslash S_n$.

Nous avons donc bien

$$S = \bigcup_{n=0}^{\infty} (S_{n+1} \backslash S_n) \tag{14.52}$$

Comme annoncé.

(4) **Un bon petit calcul** Par conséquent

$$A \cap S = A \cap \bigcup_{n=0}^{\infty} (S_{n+1} \backslash S_n) = \bigcup_{n=0}^{\infty} A \cap (S_{n+1} \backslash S_n) \tag{14.53}$$

Étant donné que les ensembles $A \cap (S_{n+1} \backslash S_n)$ sont disjoints,

$$\begin{aligned}
\mu(A \cap S) &= \sum_{n=0}^{\infty} \mu\big(A \cap (S_{n+1} \backslash S_n)\big) && \text{(14.54a)}\\
&= \sum_{n=0}^{\infty} \mu\Big((A \cap S_{n+1}) \backslash (A \cap S_n)\Big) && \text{(14.54b)}\\
&= \sum_{n=0}^{\infty} \big(\mu(A \cap S_{n+1}) - \mu(A \cap S_n)\big) && \text{(14.54c)}\\
&= \lim_{n \to \infty} \mu(A \cap S_{n+1}) - \underbrace{\mu(A \cap S_0)}_{=0} && \text{(14.54d)}\\
&= \lim_{n \to \infty} \mu(A \cap S_n). && \text{(14.54e)}
\end{aligned}$$

subeqLTvmTiQ

Dans ce calcul nous avons utilisé plusieurs fois le fait que les S_n et A étaient mesurables (et la propriété de tribu qui dit que $A \cap S_n$ est également mesurable) ainsi que le lemme 14.23. Nous avons aussi utilisé la série télescopique dans $\mathbb{R}$ pour obtenir (14.54d).

□

DefRECXooWwYgej

Définition 14.31 (λ-système[442])**.**
Soit E un ensemble. Un ensemble $\mathcal{D}$ de parties de E est un **λ-système** *lorsqu'il vérifie les conditions suivantes :*

(1) si $A, B \in \mathcal{D}$ avec $A \subset B$ alors $B \backslash A \in \mathcal{D}$,

(2) si $(A_k)_{k \geqslant 1}$ est une suite croissante d'éléments de $\mathcal{D}$ alors $\bigcup_k A_k \in \mathcal{D}$.

Note : une tribu est un λ-système.

Lemme 14.32 ([442])**.**
Une intersection quelconque de λ-systèmes dans E est un λ-système dans E.

Démonstration. Soient $\{\mathcal{D}_l\}_{l\in L}$ des λ-systèmes indicés par un ensemble L. Si $A, B \in \bigcap_{l\in L}\mathcal{D}_l$ alors $B\backslash A \in \mathcal{D}_l$ pour tout $l \in L$ et donc $B\backslash A \in \bigcap_{l\in L}\mathcal{D}_l$. De la même façon si (A_k) est une suite croissante dans $\bigcap_{l\in L}\mathcal{D}_l$ alors pour tout $l \in L$ nous avons $\bigcup_k A_k \in \mathcal{D}_l$. Donc $\bigcup_k A_k \in \bigcap_l \mathcal{D}_l$. □

Ce lemme est ce qui permet de définir le λ-système **engendré** par une classe $\mathcal{A}$ de parties de E : c'est l'intersection de tous les λ-systèmes de E contenant $\mathcal{A}$.

LemLUmopaZ

Lemme 14.33 ([442])**.**
Soit $\mathcal{C}$ une classe de parties de E (contenant E lui-même) qui soit stable par intersection finie. Alors le λ-système engendré par $\mathcal{C}$ coïncide avec la tribu engendrée par $\mathcal{C}$.

Démonstration. Nous notons $\mathcal{E}$ le λ-système engendré par $\mathcal{C}$ et $\mathcal{F}$ la tribu engendrée par $\mathcal{C}$. Étant donné que $\mathcal{F}$ est un λ-système nous avons $\mathcal{E} \subset \mathcal{F}$. Pour montrer l'inclusion inverse nous allons prouver que $\mathcal{E}$ est une tribu.

D'abord pour $C \in \mathcal{C}$ nous posons

$$\mathcal{G}_C = \{A \subset \mathcal{E} \text{ tel que } A \cap C \in \mathcal{E}\}. \tag{14.55}$$

et pour $F \in \mathcal{E}$,

$$\mathcal{H}_F = \{A \in \mathcal{E} \text{ tel que } A \cap F \in \mathcal{E}\}. \tag{14.56}$$

Nous allons montrer que $\mathcal{G}_C$ et $\mathcal{H}_F$ sont des λ-systèmes et que $\mathcal{G}_C = \mathcal{H}_F = \mathcal{E}$.

Nous commençons par $\mathcal{G}_C$. Si $A, B \in \mathcal{G}_C$ avec $A \subset B$ alors

$$(B\backslash A) \cap C = \underbrace{(B \cap C)}_{\in\mathcal{E}} \backslash \underbrace{(A \cap C)}_{\in\mathcal{E}}. \tag{14.57}$$

Puisque $\mathcal{E}$ est un λ-système et que $(A \cap C) \subset (B \cap C)$, nous avons bien $(B\backslash A) \cap C \in \mathcal{E}$ et donc $B\backslash A \in \mathcal{G}_C$. Soit maintenant (A_k) une suite croissante dans $\mathcal{G}_C$. Nous avons

$$\big(\bigcup_{k=1}^{\infty} A_k\big) \cap C = \bigcup_{k=1}^{\infty}(A_k \cap C) \tag{14.58}$$

qui est une union d'une suite croissante d'éléments de $\mathcal{E}$. Donc $\bigcup_{k=1}^{\infty}(A_k \cap C) \in \mathcal{E}$, ce qui signifie que $\bigcup_{k=1}^{\infty} A_k \in \mathcal{G}_C$. Cela termine la preuve du fait que $\mathcal{G}_C$ soit un λ-système.

Étant donné que $\mathcal{C}$ est stable par intersection finie, si $K \in \mathcal{C}$ nous avons $C \cap K \in \mathcal{C}$, ce qui signifie que $K \in \mathcal{G}_C$. Nous avons donc $\mathcal{C} \subset \mathcal{G}_C$. Donc $\mathcal{G}_C$ est un λ-système vérifiant $\mathcal{C} \subset \mathcal{G}_C \subset \mathcal{E}$. Mais comme $\mathcal{E}$ est le plus petit λ-système contenant $\mathcal{C}$ nous avons en fait $\mathcal{G}_C = \mathcal{E}$.

Nous montrons à présent que $\mathcal{H}_F$ est un λ-système. Si $A, B \in \mathcal{H}_F$ avec $A \subset B$ alors $(B\backslash A)\cap F = (B\cap F)\backslash(A\cap F)$. Comme $\mathcal{E}$ est un λ-système et que $A\cap F$ et $B\cap F$ sont dans $\mathcal{E}$ avec $A\cap F \subset B\cap F$, nous avons

$$(B \cap F)\backslash(A \cap F) \in \mathcal{H}_F. \tag{14.59}$$

Soit maintenant $(A_k)_{k\geqslant 1}$ une suite croissante dans $\mathcal{H}_F$. Pour tout k nous avons $A_k \cap F \in \mathcal{E}$, ce qui donne

$$\big(\bigcap_{k=1}^{\infty} A_k\big) \cap F = \bigcap_{k=1}^{\infty}(A_k \cap F) \in \mathcal{E}. \tag{14.60}$$

Donc $\mathcal{H}_F$ est un λ-système vérifiant $\mathcal{C} \subset \mathcal{H}_F \subset \mathcal{E}$. Nous en concluons que pour tout $C \in \mathcal{C}$ et pour tout $F \in \mathcal{E}$,

$$\mathcal{G}_C = \mathcal{H}_F = \mathcal{E}. \tag{14.61}$$

Nous allons maintenant prouver que $\mathcal{E}$ est une tribu [10].

(1) Si $F \in \mathcal{E}$ alors $E \cap F = F \in \mathcal{E}$, ce qui signifie que $E \in \mathcal{H}_F = \mathcal{E}$.

(2) Si $A \in \mathcal{E}$ alors $E\backslash A \in \mathcal{E}$ parce que $\mathcal{E}$ est un λ-système et $E \in \mathcal{E}$. Donc $\complement A \in \mathcal{E}$.

10. Définition 14.1.

(3) Montrons que $\mathcal{E}$ est stable par union finie en considérant $A, B \in \mathcal{E}$. Comme E est également un élément de $\mathcal{E}$ nous avons

$$E\backslash(A \cup B) = (E\backslash A) \cap (E\backslash B) \in \mathcal{E}. \tag{14.62}$$

Cela prouve que $\complement(A \cup B) \in \mathcal{E}$. Par complémentarité nous avons aussi $A \cup B \in \mathcal{E}$.

Soient $A_k \in \mathcal{E}$, et nommons $B_p = A_1 \cup \ldots \cup A_p$. Les ensembles B_p forment une suite croissante d'éléments de $\mathcal{E}$. L'union est donc dans $\mathcal{E}$ et ce dernier est, au final, stable par union dénombrable.

Maintenant que $\mathcal{E}$ est une tribu nous avons $\mathcal{F} \subset \mathcal{E}$ parce que $\mathcal{F}$ est la plus petite tribu contenant $\mathcal{C}$. Nous en déduisons que $\mathcal{E} = \mathcal{F}$, ce qu'il fallait démontrer. □

Le théorème suivant permet de prouver que la mesure de Lebesgue est l'unique mesure possible ayant les bonnes valeurs sur les intervalles (théorème 14.142). Si vous voulez il y a aussi le théorème de Dynkin qui est un peu similaire[443].

ThoJDYlsXu

Théorème 14.34 (Unicité des mesures[442])**.**
Soient μ et ν, deux mesures sur $(E, \mathcal{A})$ et un ensemble $\mathcal{E}$ de parties de E telles que

(1) La tribu engendrée par $\mathcal{E}$ soit $\mathcal{A}$.

(2) si $A, B \in \mathcal{E}$ alors $A \cap B \in \mathcal{E}$

(3) il existe une suite croissante (E_n) dans $\mathcal{E}$ telle que

(3a) $E = \lim E_n$,

(3b) $\mu(E_n)$ et $\nu(E_n)$ sont finis pour tout n.

Alors si les mesures μ et ν coïncident sur $\mathcal{E}$, elles coïncident sur $\mathcal{A}$ en entier.

Démonstration. Soit (E_n) une suite croissante dans $\mathcal{E}$ telle que $E = \lim E_n$.

(1) **Des restrictions** Nous considérons μ_n et ν_n, les restrictions de μ et ν à E_n, c'est-à-dire

$$\mu_n(A) = \mu(A \cap E_n) \tag{14.63a}$$
$$\nu_n(A) = \nu(A \cap E_n). \tag{14.63b}$$

Puisque les E_n sont dans $\mathcal{E} \subset \mathcal{A}$, ils sont mesurables au sens de μ et ν. Par la proposition 14.30, pour tout $A \in \mathcal{E}$ nous avons alors

$$\lim_{n\to\infty} \mu_n(A) = \mu(A) \tag{14.64a}$$
$$\lim_{n\to\infty} \nu_n(A) = \nu(A) \tag{14.64b}$$

(2) **Ce que nous devons prouver** Nous devons donc seulement montrer que pour tout $A \in \mathcal{A}$ et pour tout $n \in \mathbb{N}$, $\mu_n(A) = \nu_n(A)$. Pour cela nous nous fixons un n et nous considérons l'ensemble de parties

$$\mathcal{D} = \{A \in \mathcal{A} \text{ tel que } \mu_n(A) = \nu_n(A)\}. \tag{14.65}$$

Le but sera de prouver que $\mathcal{D} = \mathcal{A}$.

(3) **$\nu_n = \mu_n$ sur $\mathcal{E}$** Soit $A \in \mathcal{E}$. Vu que $E_n \in \mathcal{E}$, par hypothèse $A \cap E_n \in \mathcal{E}$ et donc

$$\begin{aligned} \mu_n(A) &= \mu(A \cap E_n) && (14.66a)\\ &= \nu(A \cap E_n) && (14.66b)\\ &= \nu_n(A). && (14.66c) \end{aligned}$$

SUBaEQooUJDPooQLSWmn

Pour (14.66b), nous avons utilisé l'hypothèse comme quoi $\mu = \nu$ sur $\mathcal{E}$.

(4) **Encore d'autres parties** Nous définissons $\mathcal{E}' = \mathcal{E} \cup \{E\}$. En particulier $\mathcal{E}' \subset \mathcal{D}$.

(5) $\mu_n = \nu_n$ **sur** $\mathcal{E}'$ Nous avons déjà vue l'égalité sur $\mathcal{E}$. Il suffit de vérifier l'égalité sur E. Vu que $E \cap E_n = E_n \in \mathcal{E}$, nous avons

$$\mu_n(E) = \mu(E \cap E_n) = \mu(E_n) = \nu(E_n) = \nu(E_n \cap E) = \nu_n(E). \tag{14.67}$$

(6) $\mathcal{D}$ **est un** λ**-système** Montrons que $\mathcal{D}$ est un λ-système. Soient $A, B \in \mathcal{D}$ avec $A \subset B$. Alors, étant donné que les mesures μ_n et ν_n sont finies, le lemme 14.23 nous donne

$$\mu_n(B\backslash A) = \mu_n(B) - \mu_n(A) \tag{14.68a}$$
$$\nu_n(B\backslash A) = \nu_n(B) - \nu_n(A). \tag{14.68b}$$

Donc $\mu_n(B\backslash A) = \nu_n(B\backslash A)$ et $B\backslash A \in \mathcal{D}$.

Soit par ailleurs une suite croissante $(A_k)_{k\geqslant 1}$ d'éléments de $\mathcal{D}$. Le lemme 14.24(1) nous donne

$$\mu_n(\bigcup_{k=1}^{\infty} A_k) = \lim_{p\to\infty} \mu_n(A_p). \tag{14.69}$$

Mais puisque pour chaque p nous avons $\mu_n(A_p) = \nu_n(A_p)$, nous avons aussi

$$\mu_n(\bigcup_{p=1}^{\infty} A_p) = \nu_n(\bigcup_{p=1}^{\infty} A_p). \tag{14.70}$$

Donc $\mathcal{D}$ est bel et bien un λ-système contenant $\mathcal{E}'$.

(7) **Conclusion** Par le lemme 14.33, le λ-système engendré par $\mathcal{E}'$ est égal à la tribu engendrée par $\mathcal{E}'$, mais par hypothèse la tribu engendrée par $\mathcal{E}$ est $\mathcal{A}$, donc le λ-système engendré par $\mathcal{E}'$ est $\mathcal{A}$. Comme $\mathcal{D}$ est un λ-système contenant $\mathcal{E}'$, nous avons alors $\mathcal{A} \subset \mathcal{D}$ et donc $\mathcal{A} = \mathcal{D}$, ce qu'il fallait.

□

PROPooAMTTooDfFqcP

Proposition 14.35 ([444]).
Si $\mathcal{E}$ est une base dénombrable de la topologie de l'espace X, alors $\mathcal{B}or(X) = \sigma(\mathcal{E})$.

ExDMPoohtNAj

Exemple 14.36.
La partie $\mathcal{E}$ des intervalles de $\mathbb{R}$ de la forme $]a, b[$ (avec $a, b \in \mathbb{Q}$) engendre les boréliens par les proposition 7.138 et 14.35. Par conséquent pour vérifier que deux mesures sont égales sur les boréliens de $\mathbb{R}$, il suffit de prouver qu'elles sont égales sur les intervalles ouverts. △

14.2.6 Mesure extérieure

Nous avons déjà défini la notion de mesure extérieure en la définition 14.10.

LemULSooBgZLI

Lemme 14.37 ([436]).
Soit $(S, \mathcal{F}, \mu)$ un espace mesuré et $X \subset S$. Alors

$$\inf_{\substack{A\in\mathcal{F}\\ X\in A}} \mu(A) = \inf\left\{\sum_n \mu(A_n) \text{ tel que } A_n \in \mathcal{F}, X \subset \bigcup_k A_k\right\}. \tag{14.71}$$

Démonstration. Pour montrer l'inégalité $\geqslant$, nous remarquons qu'il y a plus d'éléments dans l'ensemble du second membre que dans le premier. En effet si $A \in \mathcal{F}$ avec $X \subset A$ alors dans le membre de gauche nous pouvons prendre $A_1 = A$ et $A_{n\geqslant 1} = \varnothing$.

Pour l'inégalité dans l'autre sens, nous montrons que tout élément de

$$\left\{\sum_n \mu(A_n) \text{ tel que } A_n \in \mathcal{F}, X \subset \bigcup_k A_k\right\} \tag{14.72}$$

EqZRAooBCPFk

est plus grand qu'un élément de

$$\{\mu(A) \text{ tel que } A \in \mathcal{F}, X \subset A\}. \tag{14.73}$$

EqYNMooNvCtS

En effet si $A_n \in \mathcal{F}$ avec $X \subset \bigcup_k A_k$ alors en posant $A = \bigcup_k A_k$ nous avons $A \in \mathcal{F}$ avec $X \subset A$ ainsi que $\mu(A) \leqslant \sum_n \mu(A_n)$. Cela prouve que l'élément $\sum_n \mu(A_n)$ de (14.72) est plus grand que l'élément $\mu(A)$ de (14.73). □

14.38.
La proposition 14.39 pourrait être vue comme un cas particulier de la proposition 14.17 en utilisant 14.37. Nous en donnons cependant une preuve directe, qui est presque identique à celle de 14.17, mais avec une ou deux simplifications.

PropFDUooVxJaJ

Proposition 14.39 ([436])**.**
Soit un espace mesuré $(S, \mathcal{F}, \mu)$ et l'application

$$\begin{aligned} \mu^* \colon \mathcal{P}(S) &\to [0, \infty] \\ X &\mapsto \inf\{\mu(A) \text{ tel que } A \in \mathcal{F}, X \subset A\}. \end{aligned} \tag{14.74}$$

Alors μ^* est une mesure extérieure sur S et sa restriction à $\mathcal{F}$ est égale à μ.

Démonstration. Notons que la définition est bonne parce que l'ensemble sur lequel l'infimum est pris n'est pas vide : considérer $A = S$.

(1) **Le vide** D'abord $\mu^*(\varnothing) = 0$ parce que $\varnothing \in \mathcal{F}$.

(2) **μ^* est croissante** Soit $X \subset Y$ dans $\mathcal{P}(S)$. Si $Y \subset A$ alors $X \subset A$, donc

$$\inf\{\mu(A) \text{ tel que } A \in \mathcal{F}, X \subset A\} \leqslant \inf\{\mu(A) \text{ tel que } A \in \mathcal{F}, Y \subset A\}, \tag{14.75}$$

ce qui signifie que $\mu^*(X) \leqslant \mu^*(Y)$.

(3) **Inégalité par union dénombrable** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de parties de S. Si il existe n_0 tel que $\mu^*(X_{n_0}) = \infty$ alors nous avons automatiquement $\sum_n \mu^*(X_n) = \infty$ et l'inégalité demandée est évidente parce que n'importe quel nombre est plus petit ou égal à ∞. Nous supposons donc que $\mu^*(X_n) < \infty$ pour tout n.
Soit $\epsilon > 0$ et par définition pour chaque n, il existe un $A_n \in \mathcal{F}$ tel que $X_n \subset A_n$ et $\mu(A_n) \leqslant \mu^*(X_n) + \frac{\epsilon}{2^n}$. Bien entendu nous avons

$$\bigcup_n X_n \subset \bigcup_n A_n \in \mathcal{F}. \tag{14.76}$$

Nous en déduisons que

$$\mu^*\Big(\bigcup_n X_n\Big) \leqslant \mu\Big(\bigcup_n A_n\Big). \tag{14.77}$$

Mais $(S, \mathcal{F}, \mu)$ étant un espace mesuré,

$$\mu\Big(\bigcup_n A_n\Big) \leqslant \sum_n \mu(A_n). \tag{14.78}$$

Au final nous avons les inégalités

$$\mu^*\Big(\bigcup_n X_n\Big) \leqslant \mu\Big(\bigcup_n A_n\Big) \tag{14.79a}$$

$$\leqslant \sum_n \mu(A_n) \tag{14.79b}$$

$$\leqslant \sum_n \mu^*(X_n) + \epsilon \underbrace{\sum_n \frac{1}{2^n}}_{=1} \tag{14.79c}$$

$$= \sum_n \mu^*(X_n) + \epsilon. \tag{14.79d}$$

Ceci étant vrai pour tout ϵ,

$$\mu^*\big(\bigcup_n X_n\big) \leqslant \sum_n \mu^*(X_n), \tag{14.80}$$

ce qui prouve que μ^* est une mesure extérieure.

(4) **Restriction** Supposons que $X \in \mathcal{F}$. Alors si $X \subset A$ nous avons $\mu(X) \leqslant \mu(A)$; mais en même temps, $\mu(X)$ est dans l'infimum qui définit $\mu^*(X)$ donc

$$\mu^*(X) \leqslant \mu(X) \leqslant \inf\{\mu(A) \text{ tel que } A \in \mathcal{F}, X \subset A\} \leqslant \mu(X) \leqslant \mu^*(X). \tag{14.81}$$

Donc nous avons égalité de tous les éléments de cette chaine d'inégalité.

□

DefTRBoorvnUY

Définition 14.40.
*Soit S un ensemble et m^* une mesure extérieure sur S. Une partie $A \subset S$ est m^*-**mesurable** si pour tout $X \subset S$,*

$$m^*(X) = m^*(X \cap A) + m^*(X \cap A^c). \tag{14.82}$$

Remarque 14.41.
L'inégalité

$$m^*(X) \leqslant m^*(X \cap A) + m^*(X \cap A^c) \tag{14.83}$$

étant toujours vraie, pour prouver qu'un ensemble est m^*-mesurable, il est suffisant de prouver l'inégalité inverse :

$$m^*(X) \geqslant m^*(X \cap A) + m^*(X \cap A^c) \tag{14.84}$$

La définition 14.40 est motivée par la proposition suivante.

PropOJFoozSKAE

Proposition 14.42.
Soit un espace mesuré $(S, \mathcal{F}, \mu)$ et μ^* la mesure extérieure qui va avec. Alors tous les éléments de $\mathcal{F}$ sont μ^*-mesurables.

En d'autres termes, pour tout $A \in \mathcal{F}$ et tout $X \subset S$ nous avons

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{14.85}$$

Démonstration. Puisque $X = (X \cap A) \cup (X \cap A^c)$, et que μ^* est une mesure extérieure,

$$\mu^*(X) \leqslant \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{14.86}$$

Nous devons montrer l'inégalité inverse.

Soit $B \in \mathcal{F}$ tel que $X \subset B$. D'une part nous avons $X \cap A \subset B \cap A \in \mathcal{F}$, donc

$$\mu^*(X \cap A) \leqslant \mu^*(B \cap A) = \mu(B \cap A). \tag{14.87}$$

Et d'autre part, $X \cap A^c \subset B \cap A^c \in \mathcal{F}$, donc

$$\mu^*(X \cap A^c) \leqslant \mu(B \cap A^c). \tag{14.88}$$

En rassemblant,

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \mu(B \cap A) + \mu(B \cap A^c) = \mu(B). \tag{14.89}$$

EqLSMooTyHLB

La dernière égalité vient du fait que $B \cap A$ et $B \cap A^c$ sont disjoints et que μ est une mesure. L'inégalité (14.89) étant vraie pour tout $B \in \mathcal{F}$ tel que $X \subset B$, elle est encore vraie pour l'infimum :

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \inf\{\mu(B) \text{ tel que } B \in \mathcal{F}, X \subset B\} = \mu^*(X). \tag{14.90}$$

Nous avons donc prouvé que

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \mu^*(X). \tag{14.91}$$

□

Remarque 14.43.
Notons la duplicité du vocabulaire. Les ensembles μ-mesurables sont les éléments de $\mathcal{F}$, qui sont à priori les seuls sur lesquels μ est calculable[11], alors que les μ^*-mesurables sont les parties de S qui vérifient une certaine propriété (et μ^* est calculable sur toutes les parties de S).

14.3 Applications mesurables

14.3.1 Propriétés

DefQKjDSeC

Définition 14.44 (Fonction mesurable)**.**
*Soient $(E, \mathcal{A})$ et $(F, \mathcal{F})$ deux espaces mesurés. Une fonction $f\colon E \to F$ est **mesurable** si pour tout $\mathcal{O} \in \mathcal{F}$, l'ensemble $f^{-1}(\mathcal{O})$ est dans $\mathcal{A}$.*

PROPooEFHKooARJBwW

Proposition 14.45.
Soient $(S_i, \mathcal{F}_i)$ $(i = 1, 2, 3)$ des espaces mesurables et des fonctions mesurables $f\colon S_1 \to S_2$ et $g\colon S_2 \to S_3$. Alors la fonction $g \circ f\colon S_1 \to S_3$ est mesurable.

Démonstration. Soit $B \in \mathcal{F}_3$. Alors

$$(g \circ f)^{-1}(B) = f^{-1}\big(g^{-1}(B)\big) \in f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1. \tag{14.92}$$

□

14.3.2 D'une tribu à l'autre

LemooVDXJooZNYelH

Lemme 14.46 ([433])**.**
Soient une application $f\colon S_1 \to S_2$ et une tribu $\mathcal{F}_2$ sur S_2. Alors $f^{-1}(\mathcal{F}_2)$ est une tribu sur S_1

Démonstration. Il faut prouver les trois propriétés de la définition 14.1 d'une tribu.

(1) D'abord f est définie sur tout S_1, donc $f^{-1}(S_2) = S_1$ alors que $S_2 \in \mathcal{F}_2$.

(2) Soit $A \in f^{-1}(\mathcal{F}_2)$, c'est-à-dire $A = f^{-1}(B)$ pour un certain $B \in \mathcal{F}_2$. En ce qui concerne le complémentaire :

$$A^c = f^{-1}(B)^c = S_1 \backslash f^{-1}(B) = f^{-1}(S_2 \backslash B) = f^{-1}(B^c). \tag{14.93}$$

(3) Si $(A_i)_{i\in\mathbb{N}}$ sont des éléments de $f^{-1}(\mathcal{F}_2)$ avec $A_i = f^{-1}(B_i)$ alors

$$\bigcup_i A_i = \bigcup_i f^{-1}(B_i) = f^{-1}\big(\bigcup_i B_i\big). \tag{14.94}$$

Ce qui est dans la dernière parenthèse est dans $\mathcal{F}_2$ parce que cette dernière est une tribu.

□

PROPooJXMKooNNNXaT

Proposition 14.47 ([1])**.**
Si $f\colon (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ est une application mesurable et si $\mu\colon \mathcal{A}_1 \to [0, \infty]$ est une mesure, alors

$$\begin{aligned} \mu'\colon \mathcal{A}_1 &\to [0, \infty] \\ A &\mapsto (\mu \circ f^{-1})(A) \end{aligned} \tag{14.95}$$

est une mesure.

Le lemme suivant est également nommé « lemme de transfert ».

11. « calculable » au sens où μ y vaut un nombre bien défini ; après, que ce soit facile ou pas à calculer dans la pratique, c'est une autre histoire.

LemOQTBooWGYuDU

Lemme 14.48 (Lemme de transport)**.**
Soit $f\colon S_1 \to S_2$ une application et une classe $\mathcal{C}$ de parties de S_2. Alors

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big). \tag{14.96}$$

Démonstration. Puisque $\sigma(\mathcal{C})$ est une tribu dans S_2 alors le lemme 7.6 dit que $f^{-1}\big(\sigma(\mathcal{C})\big)$ est une tribu qui contient en particulier $f^{-1}(\mathcal{C})$. Nous en déduisons que $\sigma\big(f^{-1}(\mathcal{C})\big) \subset f^{-1}\big(\sigma(\mathcal{C})\big)$.

Réciproquement. Dans S_1 nous avons la tribu $\sigma\big(f^{-1}(\mathcal{C})\big)$. Nous pouvons alors considérer la tribu

$$\mathcal{F}_f = \{B \subset S_2 \text{ tel que } f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{C})\big)\}. \tag{14.97}$$

Montrons que $\mathcal{C} \subset \mathcal{F}_f$. Lorsque $B \in \mathcal{C}$ nous avons $f^{-1}(B) \in f^{-1}(\mathcal{C}) \subset \sigma\big(f^{-1}(\mathcal{C})\big)$. Du coup $B \in \mathcal{F}_f$. Nous avons alors, en passant aux tribus engendrées :

$$\sigma(\mathcal{C}) \subset \sigma(\mathcal{F}_f) = \mathcal{F}_f. \tag{14.98}$$

Si maintenant $B \in \sigma(\mathcal{C})$, nous avons $f^{-1}(B) \in \sigma\big(f^{-1}(\mathcal{C})\big)$, ce qui signifie que

$$f^{-1}\big(\sigma(\mathcal{C})\big) \subset \sigma\big(f^{-1}(\mathcal{C})\big). \tag{14.99}$$

□

Le théorème suivant est important pour prouver qu'une application est mesurable. En effet, il permet de ne tester si une application n'est mesurable uniquement que sur une partie génératrice de la tribu d'arrivée [12].

ThoECVAooDUxZrE

Théorème 14.49.
Soient des espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi qu'une application $f\colon S_1 \to S_2$. Si il existe un ensemble de parties $\mathcal{C}$ de S_2 tel que

- $\sigma(\mathcal{C}) = \mathcal{F}_2$
- $f^{-1}(B) \in \mathcal{F}_1$ *pour tout* $B \in \mathcal{C}$

alors f est mesurable.

Démonstration. Par hypothèse, $\sigma(\mathcal{C}) = \mathcal{F}_2$ et $f^{-1}(\mathcal{C}) \subset \mathcal{F}_1$ et nous pouvons utiliser le lemme de transfert 14.48 :

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big) \tag{14.100}$$

qui s'écrit ici

$$\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}(\mathcal{F}_2). \tag{14.101}$$

Mais comme $f^{-1}(\mathcal{C}) \subset \mathcal{F}_1$, nous avons aussi $\sigma\big(f^{-1}(\mathcal{C})\big) \subset \mathcal{F}_1$, ce qui signifie que

$$f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1. \tag{14.102}$$

Cela est exactement le fait que f soit mesurable. □

14.4 Tribu borélienne

14.4.0.1 Définition

DEFooQBQGooTqGdtY

Définition 14.50 (Tribu borélienne)**.**
*La tribu des **boréliens**, notée $\mathcal{B}or(\mathbb{R}^d)$ est la tribu engendrée par les ouverts de $\mathbb{R}^d$. Plus généralement si Y est un espace topologique, la tribu des boréliens est la tribu engendrée par les ouverts de Y.*

12. Typiquement les ouverts pour les boréliens.

PROPooYEkvbWBz

Proposition 14.51.
La tribu engendrée par une base dénombrable de la topologie est celle des boréliens.

Démonstration. Si une base de topologie est donnée, tout ouvert peut être écrit comme union d'élément de la base, proposition 7.2. Dans le cas d'une base dénombrable, cette union sera forcément dénombrable. Une tribu étant stable par union dénombrable, tout ouvert est dans la tribu engendrée par la base de topologie. Les autres boréliens suivent automatiquement.

Dit avec plus de lettres et moins de phrases, si $\mathcal{D}$ est une base dénombrable de la topologie de X, et si $\mathcal{O}$ est un ouvert de X, nous avons $\mathcal{O} = \bigcup_{i=1}^{\infty} A_i$ avec $A_i \in \mathcal{D}$. Puisqu'une tribu est stable par union dénombrable [13], nous avons $\mathcal{O} \in \sigma(\mathcal{D})$. En conséquence, $\mathcal{B}\text{or}(X) \subset \sigma(\mathcal{D})$.

Mais comme $\mathcal{D} \subset \mathcal{B}\text{or}(X)$ l'inclusion inverse est automatique. D'où l'égalité $\mathcal{B}\text{or}(X) = \sigma(\mathcal{D})$. □

14.4.0.2 Les boréliens de $\mathbb{R}$

Nous rappelons que la topologie de $\mathbb{R}$ est celle des boules donnée par le théorème 7.117. Nous rappelons (voir la proposition 7.138 et sa preuve) que les boules ouvertes de la forme $B(q,r)$ avec $q, r \in \mathbb{Q}$ forment une base dénombrable de la topologie de $\mathbb{R}$.

LemZXnAbtl

Lemme 14.52.
Soit $\{q_i\}$ une énumération des rationnels. La tribu engendrée par les ouverts $\sigma_i =]q_i, \infty[$ est la tribu des boréliens.

Démonstration. Si $a < b$ dans $\mathbb{Q}$ alors $\sigma_a \backslash \sigma_b =]a, b]$. Ensuite

$$\bigcup_{n\in\mathbb{N}^*} \sigma_a \backslash \sigma_{b-\frac{1}{n}} = \bigcup_{n\in\mathbb{N}^*}]a, b - \frac{1}{n}] =]a, b[. \tag{14.103}$$

Par union dénombrable, tous les intervalles $]a, b[$ avec $a, b \in \mathbb{Q}$ sont dans la tribu engendrée par les σ_i. Ces boules ouvertes forment une base de la topologie de $\mathbb{R}$ par la proposition 7.138 et la proposition 14.51 conclut. □

Exemple 14.53.
Les singletons sont des boréliens de $\mathbb{R}$ parce que

$$\{x\} = \Big( ]-\infty, x[\cup]x, +\infty[\Big)^c. \tag{14.104}$$

Puisqu'une tribu est stable par union dénombrable, l'ensemble $\mathbb{Q}$ est un borélien de $\mathbb{R}$. Et comme les tribus sont stables par différence ensembliste (14.3(2)), l'ensemble des irrationnels est un borélien de $\mathbb{R}$. △

14.4.0.3 Diverses expressions

LEMooUPYDooPVjscA

Lemme 14.54.
Soient un espace topologique X et un borélien B de X. Nous considérons sur B la topologie induite [14] de X et les boréliens $\mathcal{B}or(B)$ correspondants. Nous avons :

$$\mathcal{B}or(B) = \{A \in \mathcal{B}or(X) \text{ tel que } A \subset B\} = \{B \cap A \text{ tel que } A \in \mathcal{B}or(X)\}. \tag{14.105}$$

En particulier,

$$\mathcal{B}or(B) = \mathcal{B}or(X)_B. \tag{14.106}$$

EQooEUWVooCBUims

13. Définition 14.1(3)
14. Définition 7.26.

Démonstration. L'égalité

$$\{A \in \mathcal{B}\text{or}(X) \text{ tel que } A \subset B\} = \{B \cap A \text{ tel que } A \in \mathcal{B}\text{or}(X)\} \tag{14.107}$$

est déjà dans la proposition 14.9.

Nous démontrons maintenant que

$$\mathcal{B}\text{or}(B) = \{A \cap B \text{ tel que } A \in \mathcal{B}\text{or}(X)\}. \tag{14.108}$$

Pour ce faire, nous nous rappelons du lemme de transport 14.48. Soit l'injection canonique $f\colon B \to X$; pour tout $A \subset X$ nous avons $f^{-1}(A) = A \cap B$.

Nous considérons la classe $\mathcal{T}$ des ouverts de X. Par définition de la topologie induite, les ouverts de B sont les éléments de $f^{-1}(\mathcal{T})$. Donc

$$\sigma\big(f^{-1}(\mathcal{T})\big) = \mathcal{B}\text{or}(B). \tag{14.109}$$

Mais d'autre part,

$$f^{-1}\big(\sigma(\mathcal{T})\big) = \{A \cap B \text{ tel que } A \in \mathcal{B}\text{or}(X)\}. \tag{14.110}$$

Donc le lemme de transport 14.48 nous dit que

$$\mathcal{B}\text{or}(B) = \sigma\big(f^{-1}(\mathcal{T})\big) = f^{-1}\big(\sigma(\mathcal{T})\big) = \{A \cap B \text{ tel que } A \in \mathcal{B}\text{or}(X)\}. \tag{14.111}$$

Pour finir, l'égalité (14.106) se démontre :

$$\mathcal{B}\text{or}(X)_B = \{B \cap A \text{ tel que } A \in \mathcal{B}\text{or}(X)\} = \mathcal{B}\text{or}(B). \tag{14.112}$$

□

14.4.1 Applications continues et boréliennes

DefHHIBooNrpQjs

Définition 14.55 (Fonction borélienne)**.**
Une application $f\colon (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d))$ [15] *est* ***borélienne*** *si elle est mesurable, c'est-à-dire si pour tout $B \in \mathcal{B}or(\mathbb{R}^d)$ nous avons $f^{-1}(B) \in \mathcal{A}$.*

Si rien n'est précisé, une application entre deux espaces topologiques est borélienne lorsqu'elle est mesurable en considérant la tribu borélienne sur les deux *espaces.*

Si $\mathcal{A}$ est une tribu sur un ensemble E, nous notons $m(\mathcal{A})$ l'ensemble des fonctions qui sont $\mathcal{A}$-mesurables.

Le plus souvent lorsque nous parlerons de fonctions $f\colon X \to Y$ où Y est un espace topologique, nous considérons la tribu borélienne sur Y. Ce sera en particulier le cas dans la théorie de l'intégration.

Le théorème suivant est très important parce qu'en pratique c'est souvent lui, en conjonction avec la proposition 14.125 qui permet de déduire qu'une fonction est borélienne.

ThoJDOKooKaaiJh

Théorème 14.56 ([433])**.**
Soient X et Y deux espaces topologiques. Alors toute application continue $f\colon X \to Y$ est borélienne [16].

Démonstration. Pour vérifier que f est borélienne, nous devons prouver que $f^{-1}(B)$ est borélien pour tout borélien B de Y. Heureusement, le théorème 14.49 nous permet de limiter la vérification aux B appartenant à une classe engendrant les boréliens de Y.

La classe en question est toute trouvée : ce sont les ouverts. Si $\mathcal{O}$ est un ouvert de Y alors $f^{-1}(\mathcal{O})$ est un ouvert de X et donc un borélien de X. □

15. Tribu des boréliens, définition 14.50.
16. Définition 14.55.

Le théorème suivant donne une importante compatibilité entre l'induction de tribu et l'induction de topologie : la tribu induite à partir des boréliens sur un sous-espace topologique est la tribu des boréliens pour la topologie induite.

ThoSVTHooChgvYa

Théorème 14.57 ([433]).
Soit X, un espace topologique et $Y \subset X$ une partie munie de la topologie induite. Alors

$$\mathcal{B}or(Y) = \mathcal{B}or(X)_Y \tag{14.113}$$

où $\mathcal{B}or(X)_Y$ est la tribu sur Y induite de $\mathcal{B}or(X)$ par la définition 14.8.

Démonstration. Nous notons τ_X et τ_Y les topologies de X et Y.

(1) $\mathcal{B}\mathbf{or}(Y) \subset \mathcal{B}\mathbf{or}(X)_Y$ Si $A \in \tau_Y$ alors $A = Y \cap \Omega$ pour un $\Omega \in \tau_X$. Mais puisque Ω est un ouvert de X, il est un borélien de X, ce qui donne que $Y \cap \Omega$ est un élément de $\mathcal{B}\mathrm{or}(X)_Y$. Cela prouve que $\tau_Y \subset \mathcal{B}\mathrm{or}(X)_Y$, c'est-à-dire que $\mathcal{B}\mathrm{or}(X)_Y$ est une tribu sur Y contenant les ouverts de Y. Nous avons donc

$$\mathcal{B}\mathrm{or}(X) \subset \mathcal{B}\mathrm{or}(X)_Y. \tag{14.114}$$

(2) **Réciproquement** L'application $\mathrm{Id}\colon (Y, \tau_Y) \to (X, \tau_X)$ est continue parce que si Ω est ouvert de X alors $\mathrm{Id}^{-1}(\Omega) = \Omega \cap Y \in \tau_Y$. Par conséquent l'identité est une application borélienne (théorème 14.56), ce qui signifie que $\mathrm{Id}^{-1}\big(\mathcal{B}\mathrm{or}(X)\big) \subset \mathcal{B}\mathrm{or}(Y)$, ou encore que si $B \in \mathcal{B}\mathrm{or}(X)$, alors $\mathrm{Id}^{-1}(B) = B \cap Y \in \mathcal{B}\mathrm{or}(Y)$. Cela signifie que

$$\mathcal{B}\mathrm{or}(X)_Y \subset \mathcal{B}\mathrm{or}(Y). \tag{14.115}$$

□

CorooMJQYooFfwoTd

Corolaire 14.58.
Si U est un borélien de l'espace topologique X, alors les boréliens de U sont les boréliens de X inclus dans U :

$$\mathcal{B}or(U) = \{B \in \mathcal{B}or(X) \text{ tel que } B \subset U\}. \tag{14.116}$$

Démonstration. Si $B' \in \mathcal{B}\mathrm{or}(U)$, le théorème 14.57 donne un borélien $B \in \mathcal{B}\mathrm{or}(X)$ tel que $B' = B \cap U$. Mais U étant borélien de X, l'intersection $B \cap U$ est encore un borélien de X. □

Ce corolaire s'applique en particulier lorsque U est un ouvert.

La proposition suivante montre comment il est possible de construire un espace mesuré à partir d'une bijection avec un espace mesuré déjà connu. Attention cependant : la mesure construite dans cette proposition n'est pas celle qui est le plus adapté. Voir la proposition 14.283 et le blabla 14.284.

PROPooXQHTooUxJoyq

Proposition 14.59.
Soient un espace mesuré $(\Omega, \mathcal{A}, \mu)$, un ensemble Ω' et une bijection $\varphi\colon \Omega \to \Omega'$. Nous posons

(1) $\mathcal{A}' = \varphi(\mathcal{A})$,

(2) $\mu'(B) = \mu\big(\varphi^{-1}(B)\big)$ *pour tout* $B \in \mathcal{A}'$.

Alors $(\Omega', \mathcal{A}', \mu')$ est un espace mesuré.

Démonstration. En plusieurs points.

(1) **$\mathcal{A}'$ est une tribu** Il faut vérifier les différents points de la définition 14.1. D'abord, puisque $\Omega \in \mathcal{A}$, nous avons $\Omega' = \varphi(\Omega) \in \mathcal{A}'$. Pour le complémentaire, si $B \in \mathcal{A}'$ alors $B = \varphi(A)$ pour un certain $A \in \mathcal{A}$. Comme $\mathcal{A}$ est une tribu nous avons alors $\Omega \backslash A \in \mathcal{A}$ et donc $\varphi(\Omega \backslash A) \in \mathcal{A}'$. Mais comme φ est bijective,

$$\varphi(\Omega \backslash A) = \Omega' \backslash \varphi(A) = \Omega' \backslash B. \tag{14.117}$$

Le complémentaire de B est donc bien dans $\mathcal{A}'$. Pour la troisième condition, soient $B_i \in \mathcal{A}'$. Pour chaque i, il existe $A_i \in \mathcal{A}$ tel que $B_i = \varphi(A_i)$. Nous avons $\bigcup_i A_i \in \mathcal{A}$, donc

$$\bigcup_i B_i = \bigcup_i \varphi(A_i) = \varphi\Big(\bigcup_i A_i\Big) \in \mathcal{A}'. \tag{14.118}$$

Nous avons fini de prouver que $(\Omega', \mathcal{A}')$ était un espace mesurable.

(2) μ' **est une mesure positive** D'abord $\mu'(\varnothing) = \mu\big(\varphi^{-1}(\varnothing)\big) = \mu(\varnothing) = 0$. Ensuite si les A_i sont disjoints dans $\mathcal{A}'$ nous avons

$$\mu'\Big(\bigcup_{i=0}^{\infty} A_i\Big) = \mu\left(\varphi^{-1}\Big(\bigcup_{i=0}^{\infty} A_i\Big)\right) = \mu\left(\bigcup_i \varphi^{-1}(A_i)\right) = \sum_{i=0}^{\infty} \mu\big(\varphi^{-1}(A_i)\big) = \sum_i \mu'(A_i). \quad (14.119)$$

□

PROPooIKYYooCFaDaI

Proposition 14.60.
Soit une bijection continue d'inverse continue $\varphi\colon \Omega \to \Omega'$. Alors

$$\varphi\big(\mathcal{B}or(\Omega)\big) = \mathcal{B}or(\Omega'). \quad (14.120)$$

Démonstration. Si $A \in \mathcal{B}\text{or}(\Omega')$, alors $A = \varphi\big(\varphi^{-1}(A)\big) \in \varphi\big(\mathcal{B}\text{or}(\Omega)\big)$ parce que φ est continue et donc borélienne (proposition 14.56). Le même raisonnement fonctionne dans l'autre sens parce que nous avons supposé que φ est continue et d'inverse continu. □

14.4.2 Tribu de Baire

Définition 14.61.
Une partie d'un espace topologique est ***rare*** *si elle est contenue dans un fermé d'intérieur vide.*
Une partie est ***maigre*** *si elle est réunion finie ou dénombrable de parties rares.*

Exemple 14.62.
L'ensemble $\mathbb{Q}$ dans $\mathbb{R}$ est maigre mais n'est pas rare parce que $\bar{\mathbb{Q}} = \mathbb{R}$. △

PROPooCHTWooZFiSMf

Proposition 14.63 ([434])**.**
Soit X un espace topologique. L'ensemble de parties[17]

$$\mathcal{B}a(X) = \{B \cup A \text{ avec } B \text{ borélien et } A \text{ maigre}\} \quad (14.121)$$

est une tribu. Elle est appelée la ***tribu de Baire*** *de l'espace X.*

Démonstration. Nous allons montrer que les boréliens et les maigres vérifient les conditions de la proposition 14.7.

(1) Si A est maigre, il s'écrit comme $A = \bigcup_{i\in\mathbb{N}} R_i$ où les R_i sont rares. Il existe donc des fermés d'intérieur vide F_i tels que $R_i \subset F_i$; en particulier $A \subset \bigcup_i F_i$. En tant que fermés, $F_i \in \mathcal{B}\text{or}(X)$; de plus chaque F_i est rare, donc $\bigcup_i F_i$ est maigre. L'ensemble A est donc bien contenu dans un ensemble maigre et borélien.

(2) Soit A maigre et $B \subset A$. Nous avons, avec les mêmes notations, $A = \bigcup_i R_i$ et $B = \bigcup_i (R_i \cap B)$. Les ensembles $R_i \cap B$ sont encore rares, donc B est une union dénombrable d'ensembles rares. L'ensemble B est donc maigre.

(3) Si les ensembles (A_i) sont maigres, alors ils sont unions dénombrables de rares : $A_i = \bigcup_k R_k^{(i)}$. Nous avons alors

$$\bigcup_i A_i = \bigcup_{(i,k)\in\mathbb{N}^2} R_k^{(i)}, \quad (14.122)$$

et donc $\bigcup_i A_i$ est encore une union dénombrable d'ensembles rares.

□

PropGRHootvAWq

Proposition 14.64 ([434])**.**
Une partie B de l'espace topologique X est dans la tribu de Baire de X si et seulement si il existe un ouvert U tel que $B \Delta U$ est maigre.

17. Pour rappel, la tribu borélienne est définie en 14.50.

Démonstration. Nous définissons la relation d'équivalence[18] suivante sur $\mathcal{P}(X)$: nous disons que $A \sim B$ si et seulement si $A\Delta B$ est maigre.

(1) **Réflexive** Nous avons $A\Delta A = \varnothing$, donc $A \sim A$.

(2) **symétrique** Nous avons $A\Delta B = B\Delta A$, donc $\sim$ est symétrique.

(3) **transitive** Si A, B, C sont des parties de X alors nous avons toujours

$$A\Delta C \subset (A \cup B \cup C)\backslash(A \cap B \cap C) = (A\Delta B) \cup (B\Delta C). \tag{14.123}$$

Donc si $A \sim B$ et $B \sim C$ alors $A\Delta C$ est contenu dans une union de maigres et est donc maigre.

(4) **Autres propriétés de $\sim$** De plus la relation d'équivalence $\sim$ vérifie $A \sim B$ si et seulement si $A^c \sim B^c$, par le lemme 1.31(1).
Pour compléter les propriétés de $\sim$ mentionnons encore le fait que si F est fermé alors $F \sim \text{Int}(F)$. En effet $F \cup \text{Int}(F) = F$ et $F \cap \text{Int}(F) = \text{Int}(F)$, de telle sorte que $F\Delta\,\text{Int}(F) = F\backslash\,\text{Int}(F)$. Cet ensemble est un fermé parce que son complémentaire est $F^c \cup \text{Int}(F)$ qui est une union d'ouverts. De plus $F \subset \text{Int}(F)$ est d'intérieur vide, de telle sorte qu'il est rare et donc maigre.

Pour la suite de la preuve nous posons

$$\mathcal{F} = \{A \subset X \text{ tel que il existe un ouvert } U \text{ avec } U \sim A\}, \tag{14.124}$$

et nous devons prouver que $\mathcal{F} = \mathcal{B}\text{a}(X)$.

(1) $\mathcal{F} \subset \mathcal{B}\mathbf{a}(X)$
Soit $A \in \mathcal{F}$ et un ouvert U tel que $U \sim A$. Alors nous posons $M = U\Delta A$ qui est maigre. En vertu du lemme 1.31(2), nous avons

$$A = M\Delta U = (M \cup U)\backslash(M \cap U), \tag{14.125}$$

ce qui prouve que A est dans la tribu engendrée par les ouverts et les maigres, laquelle tribu est contenue dans $\mathcal{B}\text{a}(X)$.

(2) $\mathcal{B}\mathbf{a}(X) \subset \mathcal{F}$
Nous allons montrer que $\mathcal{F}$ est une tribu contenant tous les ouverts et tous les maigres. Alors en particulier $\mathcal{F}$ contiendra $\mathcal{B}\text{a}(X)$. Si U est ouvert, $U \sim U$ et donc $U \in \mathcal{F}$. Si M est maigre, alors $M \sim \varnothing$ et donc $M \in \mathcal{F}$. Il reste à prouver que $\mathcal{F}$ est une tribu.

(2a) **Vide et tout l'ensemble** C'est facile : $\varnothing$ et X sont dans $\mathcal{F}$.

(2b) **Comlémentaire** Commençons par nous souvenir que $F \sim \text{Int}(F)$ dès que F est fermé. Si $A \in \mathcal{F}$ alors il existe un ouvert U tel que $A \sim U$ et donc aussi $A^c \sim U^c$. D'autre part U^c est fermé, donc $U^c \sim \text{Int}(U^c)$, donc

$$A^c \sim U^c \sim \text{Int}(U^c), \tag{14.126}$$

ce qui implique que $A^c \in \mathcal{F}$.

(2c) **Union dénombrable** Soit $A_n \in \mathcal{F}$ et $M_n = A_n\Delta U_n$ avec M_n maigre et U_n ouvert. Nous allons prouver que

$$\bigcup_n A_n \sim \bigcup_n U_n. \tag{14.127}$$

Pour cela il faut remarquer que

$$\Big(\bigcup_n A_n\Big)\Delta\Big(\bigcup_n U_n\Big) \subset \bigcup_n (A_n\Delta U_n) = \bigcup_n M_n. \tag{14.128}$$

Le terme le plus à droite est maigre, ce qui signifie que celui le plus à gauche est contenu dans un maigre et donc est maigre lui-même.

18. Définition 1.32

□

Proposition 14.65.
Si B est un borélien de X, alors il existe un ouvert U et un maigre M tels que

(1) $B\Delta U$ est maigre,

(2) $M\Delta U = B$,

(3) $B\Delta M$ est ouvert.

Démonstration. Puisque B est borélien, il est aussi dans la tribu de Baire et il existe par la proposition 14.64 un ouvert U tel que $M = B\Delta U$ est maigre. En prenant ce U et ce M, les trois conditions sont vérifiées parce que

$$M\Delta U = (B\Delta U)\Delta U = B \tag{14.129}$$

et

$$B\Delta M = M\Delta B = (U\Delta B)\Delta B = U. \tag{14.130}$$

Tout ceci par le lemme 1.31(2). □

14.5 Espace mesuré complet

14.5.1 Partie négligeable

DefAVDoomkuXi

Définition 14.66.
*Soit un espace mesuré $(X, \mathcal{A}, \mu)$. Une partie N de X est **négligeable** pour μ si il existe $Y \in \mathcal{A}$ tel que $N \subset Y$ et $\mu(Y) = 0$.*

LemVKNooOCOQw

Lemme 14.67.
L'ensemble des parties négligeables est stable par union dénombrable.

Démonstration. Si les ensembles N_i sont négligeables, alors pour chaque i nous avons $Y_i \in \mathcal{A}$ tel que $N_i \subset Y_i$ et $\mu(Y_i) = 0$. Alors bien entendu $\bigcup_i N_i \subset \bigcup_i Y_i$ et en utilisant (14.33),

$$\mu\Big(\bigcup_i Y_i\Big) \leqslant \sum_i \mu(Y_i) = 0. \tag{14.131}$$

□

DefBWAoomQZcI

Définition 14.68.
*L'espace mesuré $(X, \mathcal{F}, \mu)$ est **complet** si tout ensemble μ-négligeable est dans $\mathcal{F}$.*

Notons que la proposition 14.7 s'applique si $(X, \mathcal{F}, \mu)$ est un espace mesuré et $\mathcal{N}$ est l'ensemble des parties μ-négligeables. C'est ce qui permet de donner le théorème suivant, que nous redémontrons de façon indépendante de la proposition 14.7.

thoCRMootPojn

Théorème 14.69 (Complétion d'espace mesuré[436, 445, 446])**.**
Soit un espace mesuré $(X, \mathcal{F}, \mu)$ et $\mathcal{N}$ l'ensemble des parties μ-négligeables de X.

(1) Les ensembles suivants sont égaux :

$$\mathcal{A} = \{A \subset X \text{ tel que } \exists B, C \in \mathcal{F} \text{ tel que } B \subset A \subset C, \mu(C\backslash B) = 0\} \tag{14.132a}$$

$$\mathcal{B} = \{B \cup N \text{ tel que } B \in \mathcal{F}, N \in \mathcal{N}\} \tag{14.132b}$$

EqFJoorxZNU

$$\mathcal{C} = \{A \subset X \text{ tel que } \exists B \in \mathcal{F} \text{ tel que } A\Delta B \in \mathcal{N}\}. \tag{14.132c}$$

Ici $A\Delta B$ est la différence symétrique de A et B, définition 1.30.

(2) L'ensemble $\hat{\mathcal{F}} = \mathcal{A} = \mathcal{B} = \mathcal{C}$ est une tribu.

(3) La définition

$$\begin{aligned}\mu' : \mathcal{B} &\to [0,\infty] \\ A \cup N &\mapsto \mu(A)\end{aligned} \tag{14.133}$$

est cohérente.

(4) L'application μ' ainsi définie est une mesure sur $(X,\mathcal{A})$.

(5) L'espace $(X,\mathcal{A},\mu')$ est complet.

(6) La mesure μ' prolonge μ.

thoCRMootPojnvii

(7) La mesure μ' est minimale au sens où toute mesure complète prolongeant μ prolonge μ'.

Démonstration. Commençons par prouver que les trois ensembles $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ sont égaux.

(1) $\underline{\mathcal{A} \subset \mathcal{B}.}$ Soit $A \in \mathcal{A}$. Alors nous avons des ensembles $B, C \in \mathcal{F}$ tels que $B \subset A \subset C$ avec $\mu(C\backslash B) = 0$. Alors nous avons aussi $A = B \cup (C\backslash B)$, ce qui prouve que $A \in \mathcal{B}$.

(2) $\underline{\mathcal{B} \subset \mathcal{C}.}$ Soit $A \in \mathcal{B}$, c'est-à-dire que $A = B \cup N$ avec $B \in \mathcal{F}$ et $N \in \mathcal{N}$. Nous avons évidemment $A \cup B = A$ et donc

$$A\Delta B = (A \cup B)\backslash(A \cap B) = A\backslash(A \cap B) = (B \cup N)\backslash(A \cap B) \subset N. \tag{14.134}$$

Pour comprendre la dernière inclusion, si x appartient à $A = B \cup N$ sans être dans N alors $x \in B$ et donc $x \in A \cap B$. Par conséquent nous avons $A\Delta B \subset N$ et donc $A\Delta B \in \mathcal{N}$.

(3) $\underline{\mathcal{C} \subset \mathcal{A}}$ Soit donc $A \in \mathcal{C}$; il existe $B \in \mathcal{F}$ tel que $A\Delta B \in \mathcal{N}$ ou encore, il existe $D \in \mathcal{F}$ tel que $A\Delta B \subset D$ avec $\mu(D) = 0$. Si nous posons $B' = B \cap D^c$ et $C' = B \cup D$ alors nous prétendons avoir

$$B' \subset A \subset C'. \tag{14.135}$$

Et nous le prouvons. En effet si $x \in B \cap D^c$ alors en remarquant que B se divise en

$$B = (B \cap A) \cup \big(B \cap (A\Delta B)\big), \tag{14.136}$$

et en nous souvenant que $B \cap (A\Delta B) \subset D$, il vient que $B \cap D^c \subset B \cap A$. Et en particulier $x \in A$. D'autre part

$$A \subset B \cup (A\Delta B) \subset B \cup D. \tag{14.137}$$

Nous avons donc bien $B' \subset A \subset C'$. Par stabilité de la tribu $\mathcal{F}$ sous les intersections et complémentaires, nous avons aussi $B', C' \in \mathcal{F}$. De plus

$$C'\backslash B' = (B \cup D)\backslash(B \cap D^c) \subset D, \tag{14.138}$$

et donc

$$\mu(C'\backslash B') \leqslant \mu(D) = 0. \tag{14.139}$$

Nous avons donc prouvé que $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C} \subset \mathcal{A}$, et donc que $\mathcal{A} = \mathcal{B} = \mathcal{C}$. Nous pouvons maintenant noter $\mathcal{A}$ indifféremment les trois ensembles.

Nous prouvons à présent que $\mathcal{A}$ est une tribu.

(1) **Tribu : le vide** Pas de problème à $\varnothing \in \mathcal{A}$

(2) **Tribu : complémentaire** Soit $A \in \mathcal{A}$. Alors il existe $B, C \in \mathcal{F}$ tels que $B \subset A \subset C$ avec $\mu(C\backslash B) = 0$. En passant au complémentaire,

$$C^c \subset A^c \subset B^c. \tag{14.140}$$

Mais $B^c\backslash C^c = C\backslash B$, donc $\mu(B^c\backslash C^c) = 0$.

(3) **Tribu : union dénombrable** Soit (A_n) des éléments de $\mathcal{A}$. Pour chaque n nous avons des ensembles $B_n, C_n \in \mathcal{F}$ tels que$B_n \subset A_n \subset C_n$ avec $\mu(C_n \backslash B_n) = 0$. En ce qui concerne les unions nous avons

$$\bigcup_n B_n \subset \bigcup_n A_n \subset \bigcup_n C_n, \tag{14.141}$$

et

$$(\bigcup_n C_n) \backslash (\bigcup_n B_n) \subset \bigcup_n (C_n \backslash B_n). \tag{14.142}$$

Par conséquent, en utilisant (14.33),

$$\mu\left((\bigcup_n C_n) \backslash (\bigcup_n B_n)\right) \leqslant \mu\left(\bigcup_n (C_n \backslash B_n)\right) \leqslant \sum_n \mu(C_n \backslash B_n) = 0. \tag{14.143}$$

Cela prouve que $\bigcup_n A_n \in \mathcal{A}$, et donc que $\mathcal{A}$ est une tribu.

(4) **Définition cohérente** Soient $A, A' \in \mathcal{F}$ et $N, N' \in \mathcal{N}$ tels que $A \cup N = A' \cup N'$. Nous considérons $Y, Y' \in \mathcal{F}$ tel que $N \subset Y$, $N' \subset Y'$ et $\mu(Y) = \mu(Y') = 0$. En vertu de (14.33) nous avons

$$\mu(A) \leqslant \mu(A \cup Y) \leqslant \mu(A' \cup Y \cup Y') \leqslant \mu(A') + \mu(Y) + \mu(Y') = \mu(A'). \tag{14.144}$$

En écrivant la même chose en échangeant les primes, nous prouvons également $\mu(A') \leqslant \mu(A)$. Au final $\mu(A) = \mu(A')$, c'est-à-dire

$$\mu'(A \cup N) = \mu'(A' \cup N'). \tag{14.145}$$

La définition de μ' est donc cohérente.

(5) **μ' est une mesure**
Le fait que μ' soit positive et que $\mu'(\varnothing)$ soit nul ne pose pas de problème. Il faut voir l'union dénombrable disjointe. Si les ensembles $A_i = B_i \cup N_i$ sont disjoints, alors les B_i et le N_i sont tous disjoints deux à deux. De plus l'ensemble $\bigcup_i N_i$ est négligeable parce que nous avons déjà vu que $\mathcal{N}$ était stable par union dénombrable (14.33). Donc

$$\mu'\left(\bigcup_i B_i \cup N_i\right) = \mu'\Big((\bigcup_i B_i) \cup \underbrace{(\bigcup_i N_i)}_{\in\mathcal{N}}\Big) = \mu(\bigcup_i B_i) = \sum_u \mu(B_i) = \sum_i \mu'(B_i \cup N_i). \tag{14.146}$$

(6) **Espace complet** Un ensemble μ'-négligeable est automatiquement μ-négligeable. En effet si H est μ'-négligeable, il existe $B \in \mathcal{F}$ et $N \in \mathcal{N}$ tels que $H \subset B \cup N$ avec $\mu(B) = 0$. Comme N est μ-négligeable, il existe $Y \in \mathcal{F}$ tel que $N \subset Y$ et $\mu(Y) = 0$. Donc $H \subset B \cup N \subset B \cup Y$ avec $\mu(B \cup Y) = 0$.

Tous les ensembles μ-négligeables faisant partie de $\mathcal{B}$, tous les ensembles μ'-négligeables font partie de $\mathcal{A}$.

(7) **Prolongement** La mesure μ' prolonge μ. En effet si $A \in \mathcal{F}$ alors $A = A \cup \varnothing \in \mathcal{B}$ et A est μ'-mesurable. De plus $\mu'(A) = \mu'(A \cup \varnothing) = \mu(A)$.

(8) **Minimalité**
Soit un espace mesuré complet $(X, \mathcal{M}, \nu)$ prolongeant $(X, \mathcal{F}, \mu)$. Pour $A \in \mathcal{A}$ nous devons prouver que $A \in \mathcal{M}$ et que $\mu'(A) = \nu(A)$. Il existe $B \in \mathcal{F}$ et $N \in \mathcal{N}$ tels que $A = B \cup N$. Puisque N est μ-négligeable, il est également ν-négligeable et donc ν-mesurable parce que ν est complète : $A \in \mathcal{M}$. Nous avons le calcul

$$\nu(B) \leqslant \nu(B \cup N) \leqslant \nu(B) + \nu(N) = \nu(B). \tag{14.147}$$

Vu que le premier et dernier termes de ces inégalités sont égaux, toutes les inégalités sont des égalités et nous avons $\nu(B) = \nu(B \cup N)$. Nous pouvons enfin faire le calcul

$$\begin{aligned}
\nu(A) &= \nu(B \cup N) && (14.148a)\\
&= \nu(B) && (14.148b)\\
&= \mu(B) && (14.148c)\\
&= \mu'(B \cup N) && (14.148d)\\
&= \mu'(A). && (14.148e)
\end{aligned}$$

SUBEQooVCBOooSeRAjw SUBEQooBDASooTBBBMs

Justifications.

— Pour (14.148c). La mesure ν prolonge μ.

— Pour (14.148d). Définition de μ'.

L'égalité $\mu'(A) = \nu(A)$ est prouvée.

□

Définition 14.70.
*L'espace mesuré complet $(X, \mathcal{A}, \mu')$ défini par le théorème 14.69 est l'**espace mesuré complété** de $(X, \mathcal{F}, \mu)$.*

Nous noterons le complété de $(S, \mathcal{F}, \mu)$ par $(S, \hat{\mathcal{F}}, \hat{\mu})$

ThoUUIooaNljH

Théorème 14.71 (Carathéodory[436])**.**
Soit S un ensemble et m^* une mesure extérieure sur S. Alors

RPPooHSWWsi

(1) l'ensemble $\mathcal{M}$ des parties m^*-mesurables est une tribu,

(2) la restriction de m^* est une mesure sur $(S, \mathcal{M})$,

(3) l'espace mesuré $(S, \mathcal{M}, m^*)$ est complet[19].

Démonstration. Une grosse partie de la preuve sera de prouver la stabilité de $\mathcal{M}$ par union dénombrable quelconque ; cela sera divisé en plusieurs parties.

(1) **Tribu : le vide** L'ensemble vide est m^*-mesurable.

(2) **Tribu : complémentaire** Soit $A \in \mathcal{M}$ et $X \in S$. La condition qui dirait $A^c \in \mathcal{M}$ est :

$$m^*(X) = m^*(X \cap A^c) + m^*(X \cap A), \tag{14.149}$$

qui est la même que celle qui dit que A est dans $\mathcal{M}$.

(3) **Tribu : union finie** Soient $A, B \in \mathcal{M}$ et $X \subset S$. Alors, comme m^* est une mesure extérieure,

$$\begin{aligned}
m^*(X) &\leqslant m^*\big(X \cap (A \cup B)\big) + m^*\big(X \cap (A \cup B)^c\big) && (14.150a)\\
&= m^*\big((X \cap A) \cup (X \cap B)\big) + m^*\big(X \cap A^c \cap B^c\big). && (14.150b)
\end{aligned}$$

Mais nous pouvons écrire la première union sous forme d'une union disjointe de la façon suivante :

$$(X \cap A) \cup (X \cap B) = (X \cap A) \cup (X \cap B \cap A^c), \tag{14.151}$$

ce qui donne

$$\begin{aligned}
m^*(X) &\leqslant m^*(X \cap A) + m^*(X \cap B \cap A^c) + m^*(X \cap A^c \cap B^c) && (14.152a)\\
&= m^*(X \cap A) + m^*(X \cap A^c) && (14.152b)\\
&= m^*(X) && (14.152c)
\end{aligned}$$

subeqLYNooRdrgCj

parce que les deux derniers termes de (14.152a) se somment à $m^*(X \cap A^c)$ parce que $B \in \mathcal{M}$. La dernière ligne est le fait que A soit m^*-mesurable.

19. Définition 14.68.

(4) **Union finie disjointe** Soient $\{A_1, \dots, A_n\}$ des éléments deux à deux disjoints de $\mathcal{M}$. Nous allons maintenant prouver par récurrence que

$$m^*\Big(X \cap \big(\bigcup_{k=1}^{n} A_k\big)\Big) = \sum_{k=1}^{n} m^*(X \cap A_k). \tag{14.153}$$

Si $n = 1$ le résultat est évident. Sinon, le fait que A_{n+1} soit m^*-mesurable donne

$$m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big)\Big) = m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}\Big) + m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}^c\Big). \tag{14.154}$$

Le fait que les A_k soient disjoints implique aussi que

$$X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1} = X \cap A_{n+1} \tag{14.155}$$

et

$$X \cap \big(\bigcup_{k=1}^{n+1} A_k\big) \cap A_{n+1}^c = X \cap \big(\bigcup_{k=1}^{n} A_k\big) \tag{14.156}$$

et donc

$$\begin{aligned} m^*\Big(X \cap \big(\bigcup_{k=1}^{n+1} A_k\big)\Big) &= m^*(X \cap A_{n+1}) + m^*\Big(X \cap \big(\bigcup_{k=1}^{n} A_k\big)\Big) && (14.157a)\\ &\overset{rec.}{=} m^*(X \cap A_{n+1}) + \sum_{k=1}^{n} m^*(X \cap A_k) && (14.157b)\\ &= \sum_{k=1}^{n+1} m^*(X \cap A_k). && (14.157c) \end{aligned}$$

La relation (14.153) est prouvée.

Notons qu'en particularisant à $X = S$ nous avons

$$m^*\big(\bigcup_{k=1}^{n} A_k\big) = \sum_{k=1}^{n} m^*(A_k) \tag{14.158}$$

dès que les A_k sont des éléments deux à deux disjoints de $\mathcal{M}$.

(5) **Union dénombrable disjointe** Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments deux à deux disjoints dans $\mathcal{M}$. Nous allons prouver les affirmations suivantes :

— $\bigcup_n A_n \in \mathcal{M}$

— $m^*\big(\bigcup_n A_n\big) = \sum_n m^*(A_n)$

où toutes les sommes et unions sur n sont entre 1 et ∞.

(5a) **Première affirmation**

Nous posons $A = \bigcup_k A_k$ et $B_n = \bigcup_{k=1}^{n} A_k$. Nous savons que $B_n \in \mathcal{M}$ pour tout n par le point précédent. Donc si $X \in S$ nous avons

$$\begin{aligned} m^*(X) &= m^*(X \cap B_n) + m^*(X \cap B_n^c) && (14.159a)\\ &= \sum_{k=1}^{n} m^*(X \cap A_k) + m^*(X \cap B_n^x) && (14.159b)\\ &\geqslant \sum_{k=1}^{n} m^*(X \cap A_k) + m^*(X \cap A^c) && (14.159c) \end{aligned}$$

où nous avons utilisé la relation (14.153) sur les B_n ainsi que le fait que $A^c \subset B_n^c$ (parce que $B_n \subset A$). L'inégalité (14.159a) étant vraie pour tout n, elle est vraie à la limite :

$$m^*(X) \geqslant \sum_{k=1}^{\infty} m^*(A \cap A_k) + m^*(X \cap A^c) \tag{14.160a}$$

$$\geqslant m^*\Big(\bigcup_k (X \cap A_k)\Big) + m^*(X \cap A^c) = m^*\Big(X \cap \big(\bigcup_k A_k\big)\Big) + m^*(X \cap A^c) \tag{14.160b}$$

$$\geqslant m^*(X \cap A) + m^*(X \cap A^c), \tag{14.160c}$$

ce qui signifie que $A \in \mathcal{M}$.

(5b) **Seconde affirmation** En particularisant à $X = A$ et en tenant compte des faits que $A \cap A_k = A_k$ et $A \cap A^c = \varnothing$,

$$m^*(A) \geqslant \sum_{k=1}^{\infty} m^*(A \cap A_k) + m^*(A \cap A^c), \tag{14.161}$$

c'est-à-dire que pour tout n nous avons

$$m^*\big(\bigcup_{k\in\mathbb{N}} A_k\big) \geqslant \sum_{k=1}^{n} m^*(A_k). \tag{14.162}$$

L'inégalité est encore vraie à la limite, et l'inégalité inverse étant toujours vraie pour une mesure extérieure,

$$m^*\big(\bigcup_{k\in\mathbb{N}} A_k\big) = \sum_{k=1}^{\infty} m^*(A_k). \tag{14.163}$$

(6) **Union dénombrable quelconque** Soit maintenant une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{M}$ que nous ne supposons plus être disjoints. Nous nous ramenons au cas disjoint en posant

$$\begin{cases} B_1 = A_1 & \text{(14.164a)} \\ B_n = A_n \cap \big(\bigcup_{k=1}^{n-1} A_k\big)^c, & \text{(14.164b)} \end{cases}$$

c'est-à-dire que nous mettons dans B_n les éléments de A_n qui ne sont dans aucun des A_k précédents. Autrement dit, nous posons $B_0 = \varnothing$ et $B_n = A_n \backslash B_{n-1}$. L'ensemble $\mathcal{M}$ étant stable par réunion finie, par complément et par intersection finie nous avons $B_n \in \mathcal{M}$. De plus les B_n sont disjoints, donc

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k \in \mathcal{M}. \tag{14.165}$$

La première égalité se justifie de la façon suivante : si $x \in \bigcup_{k=1}^{\infty} A_k$ alors nous notons n_0 le plus petit n tel que $x \in A_n$ et alors $x \in B_{n_0}$.

(7) **Espace complet** Nous prouvons à présent que $(S, \mathcal{M}, m^*)$ est un espace mesuré complet. Soit N une partie m^*-négligeable de S et $Y \in \mathcal{M}$ tel que $m^*(Y) = 0$ et $N \subset Y$. D'abord $m^*(N) = 0$ parce que

$$m^*(N) \leqslant m^*(Y) = 0. \tag{14.166}$$

Si $X \subset S$ nous avons

$$X \cap N \subset N \Rightarrow m^*(X \cap N) = 0 \tag{14.167a}$$

$$X \cap N^c \subset X \Rightarrow m^*(X \cap N^c) \leqslant m^*(X). \tag{14.167b}$$

Donc

$$m^*(X \cap N) + m^*(X \cap N^c) \leqslant m^*(X), \tag{14.168}$$

ce qui montre que N est m^*-mesurable.

□

14.72.
Ce théorème nous pousse à adopter des éléments de notation. Lorsqu'un espace mesuré $(S, \mathcal{F}, \mu)$ est donné, nous noterons

$$(S, \mathcal{M}, \mu^*) \tag{14.169}$$

l'espace mesuré construit de la façon suivante. D'abord μ^* est la mesure extérieure associée à μ par la proposition 14.39. Ensuite $\mathcal{M}$ est la tribu des parties μ^*-mesurables, qui est bien une tribu parce que μ^* est une mesure extérieure (14.71). La proposition (14.42) dit alors que $\mathcal{F} \subset \mathcal{M}$. De plus 14.71 nous explique que si $A \in \mathcal{F}$ alors $\mu(A) = \mu^*(A)$. Tout cela pour dire que

$$(S, \mathcal{F}, \mu) \subset (S, \mathcal{M}, \mu^*). \tag{14.170}$$ EqXDPooKwWAF

Et enfin, 14.71 nous dit que l'espace mesuré $(S, \mathcal{M}, \mu^*)$ est complet.

ExOIXoosScTC

Exemple 14.73.
Montrons un cas dans lequel $(S, \mathcal{M}, \mu^*)$ n'est pas σ-fini. Soit S un ensemble non dénombrable et $\mathcal{F}$ la tribu des parties de S qui sont, soit finis ou dénombrables, soit de complémentaire fini ou dénombrable. Nous y mettons la mesure

$$\mu(A) = \begin{cases} 0 & \text{si } A \text{ est au plus dénombrable} \\ \infty & \text{sinon.} \end{cases} \tag{14.171}$$

Cette mesure n'est pas σ-finie parce qu'aucune union de dénombrables est non dénombrable. De plus $(S, \mathcal{F}, \mu)$ est complet parce que toute partie contenue dans un ensemble fini ou dénombrable est fini ou dénombrable (1.143).

(1) $\mathcal{F}$ **n'est pas** $\mathcal{P}(S)$ La tribu $\mathcal{F}$ est différente de $\mathcal{P}(S)$. En effet S étant infini, il existe par 1.154 une bijection $\varphi\colon \{1,2\} \times S \to S$. Alors l'ensemble $\varphi\big(\{1\} \times S\big)$ est non dénombrable et son complémentaire

$$\varphi\big(\{1\} \times S\big)^c = \varphi\big(\{2\} \times S\big) \tag{14.172}$$

n'est pas dénombrable non plus. Cet ensemble n'est donc pas de $\mathcal{F}$.

(2) $\mathcal{M}$ **est** $\mathcal{P}(S)$ En effet, soit $A \subset S$; il faut prouver que pour tout $X \subset S$ nous avons

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{14.173}$$

Nous prouvons cela en séparant les cas, suivant que X est dénombrable ou non.
Si X est fini ou dénombrable, alors $X \cap A$ et $X \cap A^c$ le sont également, et nous avons $\mu^*(X) = \mu(X) = 0$ ainsi que $\mu^*(X \cap A) = \mu^*(X \cap A^c) = 0$.
Si au contraire X n'est pas dénombrable,

$$\mu^*(X) = \inf_{\substack{A \in \mathcal{F} \\ X \subset A}} \mu(A) = \infty, \tag{14.174}$$

parce que X n'étant pas dénombrable, l'ensemble A ne l'est pas non plus et $\mu(A) = \infty$. Mais comme X n'est pas dénombrable, soit $X \cap A$, soit $X \cap A^c$ (soit les deux) n'est pas dénombrable non plus ; par conséquent

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) = \infty. \tag{14.175}$$

Par conséquent $(S, \mathcal{F}, \mu) \neq (S, \mathcal{M}, \mu^*)$. Mais puisque $(S, \mathcal{F}, \mu)$ est complété nous devons avoir $(S, \mathcal{F}, \mu) = (S, \hat{\mathcal{F}}, \hat{\mu})$. Tout cela pour dire que nous avons un exemple avec

$$(S, \mathcal{M}, \mu^*) \neq (S, \hat{\mathcal{F}}, \hat{\mu}). \tag{14.176}$$

△

Nous avons deux façons de créer un espace complet à partir de $(S, \mathcal{F}, \mu)$.

(1) Partir de la mesure extérieure μ^* et construire $(S, \mathcal{M}, \mu^*)$.

(2) Partir des ensembles μ-négligeables, construire $\hat{\mathcal{F}}$ et ensuite $(S, \hat{\mathcal{F}}, \hat{\mu})$.

Ces deux façons ne sont pas équivalentes en général comme le montre l'exemple 14.73. Mais il sera montré par la proposition 14.77 que si $(S, \mathcal{F}, \mu)$ est σ-fini alors les deux sont équivalent.

LemAESoofkMpi

Lemme 14.74.
Soit $(S, \mathcal{F}, \mu)$ un espace mesuré. Alors pour tout $X \subset S$ tel que $\mu^*(X) < \infty$ il existe $A \in \mathcal{F}$ tel que $X \subset A$ et $\mu^*(X) = \mu(A)$.

C'est-à-dire que μ^* a beau être défini sur toutes les parties de S, ce qu'il faut rajouter pour être μ-mesurable, c'est pas grand chose.

Démonstration. Par définition de la mesure extérieure associée à μ en tant qu'infimum, pour tout $n \geqslant 1$, il existe $A_n \in \mathcal{F}$ tel que $X \subset A_n$ et $\mu(A_n) \leqslant \mu^*(X) + \frac{1}{2^n}$. Nous posons $A = \bigcap_{n \geqslant 1} A_n$ et nous vérifions que ce A fait l'affaire.

D'abord $A \in \mathcal{F}$ parce qu'une tribu est stable par union dénombrable. Ensuite pour tout $n \geqslant 1$ nous avons

$$\mu(A) \leqslant \mu(A_n) \leqslant \mu^*(X) + \frac{1}{2^n}, \tag{14.177}$$

et à la limite $\mu(A) \leqslant \mu^*(X)$. Mais $X \subset A$ implique $\mu^*(X) \leqslant \mu(A)$ parce que $\mu^*(X)$ l'infimum d'un ensemble contenant $\mu(A)$. □

LemXOUNooUbtpxm

Corolaire 14.75.
Soit une mesure μ et la mesure extérieure μ^* associée[20]*. Une partie N de X est négligeable si et seulement si $\mu^*(N) = 0$.*

Démonstration. Si μ^* est la mesure extérieure associée à μ et si N est μ-négligeable alors $\mu^*(N) = 0$ parce que

$$\mu^*(N) \leqslant \mu^*(Y) = \mu(Y) = 0 \tag{14.178}$$

pour un certain Y mesurable de mesure nulle contenant N.

D'autre part si $\mu^*(N) = 0$ alors le lemme 14.74 donne une partie mesurable A telle que $N \subset A$ et $\mu(A) = 0$, c'est-à-dire que N est négligeable. □

LemOAEoocBDaO

Lemme 14.76.
Si l'espace mesuré $(S, \mathcal{F}, \mu)$ est σ-fini alors l'espace mesuré $(S, \mathcal{M}, \mu^*)$ est également σ-fini.

Démonstration. Puisque $(S, \mathcal{F}, \mu)$ est σ-fini, nous avons une suite croissante A_n d'éléments de $\mathcal{F}$ tels que $\bigcup_n A_n = S$ et telle que $\mu(A_n) < \infty$ pour tout n. Étant donné que $\mathcal{F} \subset \mathcal{M}$, cette suite convient également pour montrer que $(S, \mathcal{M}, \mu^*)$ est σ-fini parce que $\mu^*(A_n) = \mu(A_n) < \infty$. □

La proposition suivante montre que si $(S, \mathcal{F}, \mu)$ est σ-finie alors nous avons l'égalité.

PropIIHooAIbfj

Proposition 14.77.
Soit $(S, \mathcal{F}, \mu)$ un espace mesuré σ-fini, μ^* la mesure extérieure associée et $\mathcal{M}$ la tribu des ensembles μ^*-mesurables[21]*. Alors*

$$(S, \mathcal{M}, \mu^*) = (S, \hat{\mathcal{F}}, \hat{\mu}). \tag{14.179}$$

Démonstration. La proposition 14.42 indique que tous les éléments de $\mathcal{F}$ sont μ^*-mesurables, c'est-à-dire que $\mathcal{F} \subset \mathcal{M}$. Mais l'espace $(S, \mathcal{M}, \mu^*)$ est complet par le théorème de Carathéodory 14.71, donc par minimalité du complété (14.69(7)),

$$(S, \hat{\mathcal{F}}, \hat{\mu}) \subset (S, \mathcal{M}, \mu^*) \tag{14.180}$$

20. Par la proposition 14.39.
21. C'est bien une tribu par 14.71(1).

au sens où $\hat{\mathcal{F}} \subset \mathcal{M}$ et si $A \in \hat{\mathcal{F}}$ alors $\hat{\mu}(A) = \mu^*(A)$. Notons que cette inclusion est vraie même si la mesure n'est pas σ-finie.

Nous passons à l'inclusion inverse. Soit $A \in \mathcal{M}$, c'est-à-dire que pour tout $Y \subset S$ nous avons

$$\mu^*(Y) = \mu^*(Y \cap A) + \mu^*(Y \cap A^c). \tag{14.181}$$

EqTZAooTCdGg

Nous allons montrer que $A \in \hat{\mathcal{F}}$ en séparant les cas suivant que $\mu^*(A) = \infty$, ou non.

(1) **Si $\mu^*(A) < \infty$**
Par le lemme 14.74, il existe $X \in \mathcal{F}$ tel que $A \subset X$ et $\mu^*(A) = \mu(X)$. Comme $(S, \mathcal{F}, \mu) \subset (S, \mathcal{M}, \mu^*)$ nous avons alors

$$\mu^*(A) = \mu(X) = \mu^*(X). \tag{14.182}$$

EqKFQooQaont

Nous écrivons la relation (14.181) avec ce X en guise de Y, et en nous souvenant que $X \cap A = A$ et $X \cap A^c = X \backslash A$:

$$\mu^*(X) = \mu^*(A) + \mu^*(X \backslash A). \tag{14.183}$$

En tenant compte de (14.182) et du fait que $\mu^*(A) < \infty$, nous pouvons simplifier et trouver $\mu^*(X \backslash A) = 0$. Le lemme 14.74 nous donne alors $B \in \mathcal{F}$ tel que $X \backslash A \subset B$ et $\mu(B) = \mu^*(X \backslash A) = 0$, c'est-à-dire que $X \backslash A$ est μ-négligeable. Par conséquent $X \backslash A \in \hat{\mathcal{F}}$. En écrivant

$$A = X \backslash (X \backslash A), \tag{14.184}$$

nous avons écrit A comme différence de deux éléments de $\hat{\mathcal{F}}$ et nous concluons que $A \in \hat{\mathcal{F}}$.

(2) **Si $\mu^*(A) = \infty$**
Le lemme 14.76 nous indique que $(S, \mathcal{M}, \mu^*)$ est σ-fini et il existe donc une suite $(S_n)_{n \geqslant 1}$ dans $\mathcal{M}$ telle que $\bigcup_n S_n = S$ et $\mu^*(S_n) < \infty$. L'ensemble $A \cap S_n$ est un élément de $\mathcal{M}$ vérifiant

$$\mu^*(A \cap S_n) \leqslant \mu^*(S_n) < \infty, \tag{14.185}$$

ce qui implique que $A \cap S_n \in \hat{\mathcal{F}}$ par la première partie. Maintenant $A = \bigcup_n (A \cap S_n) \in \hat{\mathcal{F}}$ par union dénombrable d'éléments de la tribu $\hat{\mathcal{F}}$.

□

PROPooAMIEooRomnMG

Proposition 14.78 ([1])**.**
Soit un espace mesuré $(\Omega, \mathcal{F}, \mu)$. Nous considérons un mesurable $M \in \mathcal{F}$ ainsi que

- *la tribu induite $\mathcal{F}_M = \{A \cap M \text{ tel que } A \in \mathcal{F}\}$,*
- *la tribu complétée $\hat{\mathcal{F}}$ de $\mathcal{F}$ dans Ω,*
- *la tribu complétée $\widehat{\mathcal{F}_M}$ de $\mathcal{F}_M$ dans M (où nous avons considéré la mesure restreinte[22] de μ).*
- *la tribu induite $(\hat{\mathcal{F}})_M = \{A \cap M \text{ tel que } A \in \hat{\mathcal{F}}\}$ de $\hat{\mathcal{F}}$ sur M.*

Alors

$$(\hat{\mathcal{F}})_M = \widehat{\mathcal{F}_M}. \tag{14.186}$$

Démonstration. L'utilisation de la proposition 14.9 nous donne déjà les expressions alternatives

$$(\hat{\mathcal{F}})_M = \{A \cap M \text{ tel que } A \in \hat{\mathcal{F}}\} = \{A \subset M \text{ tel que } A \in \hat{\mathcal{F}}\} \tag{14.187}$$

et

$$\mathcal{F}_M = \{A \cap M \text{ tel que } A \in \mathcal{F}\} = \{A \subset M \text{ tel que } A \in \mathcal{F}\}. \tag{14.188}$$

Pour prouver $(\hat{\mathcal{F}})_M = \widehat{\mathcal{F}_M}$ il faudra faire deux inclusions, et nous avons l'embarras du choix.

22. Ce n'est pas ce qu'il se passe dans le cas de S^1 par rapport à $\mathbb{C}$, voir la proposition 18.72(3) bien que S^1 soit un borélien de $\mathbb{C}$.

(1) **Première : $\widehat{\mathcal{F}_M} \subset \{M \cap A \text{ tel que } A \in \hat{\mathcal{F}}\}$** Un élément de $\widehat{\mathcal{F}_M}$ est de la forme $B \cup N$ où $B \in \mathcal{F}_M$ et où N est négligeable[23] dans M. Vu que $B \in \mathcal{F}_M$, il existe $A \in \mathcal{F}$ tel que $B = A \cap M$. Vu que B et N sont dans M nous pouvons « factoriser » l'intersection :

$$B \cup N = M \cap (A \cup N) \tag{14.189}$$

avec N négligeable dans M et donc également négligeable dans Ω. Donc $A \cup N \in \hat{\mathcal{F}}$.

(2) **Deuxième : $\{M \cap A \text{ tel que } A \in \hat{\mathcal{F}}\} \subset \widehat{\mathcal{F}_M}$**

Soit $A \in \hat{\mathcal{F}}$. Nous avons une partie négligeable N de Ω et un élément $B \in \mathcal{F}$ tels que $A = B \cup N$. Nous avons la décomposition

$$M \cap (B \cup N) = (M \cap B) \cup (M \cap N). \tag{14.190}$$ EQooWWQEooDRlnLN

Il s'agit maintenant de nous assurer que cette décomposition implique que $M \cap (B \cup N) \in \widehat{\mathcal{F}_M}$. Soit $N_1 \in \mathcal{F}$ tel que $\mu(N_1) = 0$ et $N \subset N_1$. Puisque $M \cap N_1 \in \mathcal{F}$ (intersections dans une tribu), nous pouvons écrire

$$M \cap N \subset M \cap N_1 \tag{14.191}$$

avec $\mu(M \cap N_1) = 0$. Cela pour dire que $M \cap N$ est négligeable dans M. La décomposition (14.190) est donc bien une union d'un élément de $\mathcal{F}_M$ avec un négligeable de M, et donc bien un élément de $\widehat{\mathcal{F}_M}$.

□

14.79.
La principale application de la proposition 14.78 est le cas où $\mathcal{F} = \mathcal{B}\text{or}(\mathbb{R}^n)$ et M est un borélien B de $\mathbb{R}^n$. Dans ce cas, la proposition explique que la tribu de Lebesgue sur B (complétée depuis les boréliens de la topologie induite) est donnée directement par l'intersection entre B et la tribu de Lebesgue de $\mathbb{R}^n$. Donc sans devoir passer par la topologie induite, les boréliens et la completion :

$$\mathcal{L}\text{eb}(\mathbb{R}^n)_M = \widehat{\mathcal{B}\text{or}(\mathbb{R}^n)_M}. \tag{14.192}$$

Exemple dans la proposition 18.70 qui donne une structure d'espace mesuré dans S^1 à partir de la mesure de Lebesgue sur $\mathbb{C}$.

14.5.2 Prolongement

Le théorème suivant est parfois nommé théorème d'extension de Carathéodory, par exemple sur Wikipédia. Le théorème de Carathéodory en étant un des ingrédients principaux, on comprend.

ThoLCQoojiFfZ

Théorème 14.80 (Prolongement de Hahn[436])**.**
Soit $\mathcal{A}$ une algèbre de parties d'un ensemble S et μ une mesure sur $(S, \mathcal{A})$. Soit $\mathcal{F} = \sigma(\mathcal{A})$ la tribu engendrée par $\mathcal{A}$. Alors

(1) La mesure μ se prolonge en une mesure m sur $\mathcal{F}$.

(2) Si μ est σ-finie alors le prolongement est unique et m est σ-finie.

(3) Si μ est finie, alors m l'est aussi.

Démonstration. La proposition 14.17 nous donne une mesure extérieure μ^* sur S dont la restriction à $\mathcal{A}$ est μ. Si $\mathcal{M}$ est la tribu des parties μ^*-mesurables de S alors le théorème de Carathéodory 14.71 nous dit que $(S, \mathcal{M}, \mu^*)$ est un espace mesuré.

23. Pour rappel, une partie est négligeable quand elle est incluse à une partie de mesure nulle.

(1) $\mathcal{A} \subset \mathcal{M}$ Cette partie est une adaptation de ce qui a déjà été fait dans la preuve de la proposition 14.42. Soit $A \in \mathcal{A}$ et $X \in S$; nous devons prouver la relation de la définition 14.40. Comme μ^* est une mesure extérieure nous avons automatiquement

$$\mu^*(X) \leqslant \mu^*(X \cap A) + \mu^*(X \cap A^c). \tag{14.193}$$

Il reste à prouver l'inégalité inverse. Soit une suite B_k d'éléments de $\mathcal{A}$ telle que $X \subset \bigcup_k B_k$; nous avons alors

$$\mu^*(X \cap A) \leqslant \mu^*\Big(\bigcup_{k=1}^{\infty} B_k \cap A\Big) \leqslant \sum_{k=1}^{\infty} \mu^*(B_k \cap A) = \sum_k \mu(B_k \cap A) \tag{14.194}$$

où nous avons utilisé la définition 14.10(3) ainsi que le lemme 14.15. De la même façon,

$$\mu^*(X \cap A^c) \leqslant \sum_k \mu(B_k \cap A^c). \tag{14.195}$$

Mettant les deux bouts ensemble, en remarquant que $B_k \cap A \in \mathcal{A}$ et donc que $\mu^*(B_k \cap A) = \mu(B_k \cap A)$,

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \sum_k \mu(B_k \cap A) + \mu(B_k \cap A^c) = \sum_k \mu(B_k). \tag{14.196}$$

La somme $\mu^*(X \cap A) + \mu^*(X \cap A^c)$ est donc inférieure à chacun des éléments de l'ensemble sur lequel on prend l'infimum pour définir [24] $\mu^*(X)$, donc

$$\mu^*(X \cap A) + \mu^*(X \cap A^c) \leqslant \mu^*(X). \tag{14.197}$$

A fortiori nous avons $\sigma(\mathcal{A}) \subset \mathcal{M}$ et donc $(S, \sigma(\mathcal{A}), \mu^*)$ est un espace mesuré. Cela prouve l'existence d'une mesure prolongeant μ à $\sigma(\mathcal{A})$.

(1) **Unicité**

Nous supposons à présent que μ est σ-finie. Soient m_1 et m_2 deux mesures prolongeant μ et définies sur une tribu contenant $\mathcal{A}$. Nous posons

$$\mathcal{C} = \{A \in \mathcal{A} \text{ tel que } \mu(A) < \infty\}. \tag{14.198}$$

Dans l'optique d'utiliser le théorème d'unicité des mesures 14.34, nous prouvons que $\sigma(\mathcal{A}) = \sigma(\mathcal{C})$. Vu que μ est σ-finie, il existe une suite croissante (S_n) d'éléments de $\mathcal{A}$ telle que $S = \bigcup_n S_n$ et $\mu(S_n) < \infty$. Alors si $A \in \mathcal{A}$ nous avons $A = \bigcup_n (A \cap S_n)$, et donc $A \in \sigma(\mathcal{C})$. Donc $\mathcal{A} \subset \sigma(\mathcal{C})$. Mais étant donné que $\mathcal{C} \subset \mathcal{A}$ nous avons aussi $\sigma(\mathcal{C}) \subset \sigma(\mathcal{A})$. Au final $\sigma(\mathcal{A}) = \sigma(\mathcal{C})$.

Les mesures m_1 et m_2 sont des mesures sur $\sigma(\mathcal{C})$ coïncidant sur $\mathcal{C}$ (parce que $\mathcal{C} \subset \mathcal{A}$). De plus la classe $\mathcal{C}$ est stable par intersection finie et contient une suite croissante dont l'union est S (parce que μ est σ-finie).

Le théorème 14.34 nous dit alors que m_1 et m_2 coïncident sur $\sigma(\mathcal{C}) = \sigma(\mathcal{A})$.

(2) **Extension finie et σ-finie**

Enfin si μ est σ-finie il existe $S_n \in \mathcal{A}$ avec $\mu(S_n) < \infty$ et $\bigcup_n S_n = S$. Ces ensembles vérifient tout autant $m(S_n) = \mu(S_n) < \infty$ pour tout prolongement m de μ.

Idem si μ est finie, tout prolongement est fini.

□

24. Définition 14.20.

ExKCEoolsZrL

Exemple 14.81 ([436]).
Soit $\mathcal{A}$, l'algèbre de parties de $\mathbb{R}$ formée par les réunions finies d'intervalles de la forme $]-\infty,a[$, $[a,b[$ et $[b,+\infty[$ avec $-\infty<a\leqslant b<+\infty$. Notons que les singletons ne font pas partie de $\mathcal{A}$ parce que $[a,a[=\varnothing$. Nous posons

$$\mu(A)=\begin{cases}0 & \text{si } A=\varnothing\\ \infty & \text{sinon.}\end{cases}\tag{14.199}$$

Cela donne une mesure (non σ-finie) sur $(\mathbb{R},\mathcal{A})$.

Nous allons prouver que la tribu engendrée par $\mathcal{A}$ est la tribu des boréliens et que μ accepte (au moins) deux prolongements distincts à $\sigma(\mathcal{A})$.

D'abord nous avons

$$]a,b[=\big(]-\infty,a[\cup[b,+\infty[\big)\cap[a,b[,\tag{14.200}$$

donc toutes les boules ouvertes appartiennent à $\sigma(\mathcal{A})$. Ces dernières comprenant une base dénombrable de la topologie de $\mathbb{R}$ (par la proposition 7.138), tous les ouverts de $\mathbb{R}$ sont dans $\sigma(\mathcal{A})$. Par conséquent $\mathcal{B}\text{or}(\mathbb{R})\subset(\mathbb{R}^d)$. Mais en même temps tous les éléments de $\mathcal{A}$ sont des boréliens, donc $\mathcal{B}\text{or}(\mathbb{R})=\sigma(\mathcal{A})$ parce que la fermeture en tant qu'algèbre de parties est plus petite que la fermeture en tant que tribu.

La mesure de comptage prolonge μ parce qu'à part l'ensemble vide, tous les éléments de $\mathcal{A}$ sont infinis. Notons que les singletons sont dans $\sigma(\mathcal{A})$, donc la mesure de comptage prend d'autres valeurs que 0 et $+\infty$.

Par ailleurs la mesure

$$\mu'(A)=\begin{cases}0 & \text{si } A=\varnothing\\ +\infty & \text{sinon}\end{cases}\tag{14.201}$$

est également une mesure prolongeant μ à $\sigma(\mathcal{A})=\mathcal{B}\text{or}(\mathbb{R})$.

La mesure de comptage et μ' sont deux prolongements distincts de μ. △

Exemple 14.82 ([436]).
Nous montrons maintenant une mesure non σ-finie qui se prolonge en deux mesures distinctes, toutes deux σ-finies.

Nous considérons la même algèbre $\mathcal{A}$ de parties que celle donnée dans l'exemple 14.81, mais cette fois vue sur $\mathbb{Q}$ uniquement. La mesure de comptage m sur $(\mathbb{Q},\mathcal{A})$ n'est pas σ-finie.

Puisque les singletons sont des boréliens, nous avons $\sigma(\mathcal{A})=\mathcal{P}(\mathbb{Q})$, ce qui fait que $(\mathbb{Q},\sigma(\mathcal{A}),m)$ est un prolongement σ-fini de m. L'espace mesuré $(\mathbb{Q},\sigma(\mathcal{A}),2m)$ est également σ-fini et est un prolongement distinct de $(\mathbb{Q},\mathcal{A},m)$. △

PROPooORDCooJEsjzR

Proposition 14.83.
Soient des espaces mesurés $(S_1,\mathcal{F}_1,\mu_1)$ et $(S_2,\mathcal{F}_2,\mu_2)$ ainsi qu'une application $\varphi\colon S_1\to S_2$ avec les hypothèses suivantes :

(1) φ est une bijection,

(2) φ est mesurable d'inverse mesurable,

(3) si $\mu_1(A)=0$ alors $\mu_2\big(\varphi(A)\big)=0$,

(4) si $\mu_2(A)=0$ alors $\mu_1\big(\varphi^{-1}(A)\big)=0$.

Alors

$$\hat{\mathcal{F}}_2=\varphi(\hat{\mathcal{F}}_1).\tag{14.202}$$

Démonstration. Nous prouvons que $\hat{\mathcal{F}}_2\subset\varphi(\hat{\mathcal{F}}_1)$. Vu la symétrie des hypothèses, l'inclusion inverse se fera de même.

Soit $A\in\hat{\mathcal{F}}_2$. Nous avons $A=B\cup N$ avec $B\in\mathcal{F}_2$ et N, une partie μ_2-négligeable. Nous considérons $N_1\in\mathcal{F}_2$ tel que $\mu_2(N_1)=0$ et $N\subset N_1$. Notre but est maintenant de prouver que $\varphi^{-1}(B\cup N)\in\hat{\mathcal{F}}_1$.

Comme φ est une bijection, nous avons

$$\varphi^{-1}(B \cup N) = \varphi^{-1}(B) \cup \varphi^{-1}(N). \tag{14.203}$$

Là-dedans, $\varphi^{-1}(B) \in \mathcal{F}_1$ parce que φ est borélienne. Il nous reste à voir que $\varphi^{-1}(N)$ est μ_1-négligeable. Puisque $N \subset N_1$, nous avons $\varphi^{-1}(N) \subset \varphi^{-1}(N_1)$ où $\varphi^{-1}(N_1) \in \mathcal{F}_1$.

Par construction, $\mu_2(N_1) = 0$ et par hypothèse, $\mu_1\big(\varphi^{-1}(N_1)\big) = 0$.

Au total,

$$\varphi^{-1}(B \cup N) = \underbrace{\varphi^{-1}(B)}_{\in \mathcal{F}_1} \cup \underbrace{\varphi^{-1}(N)}_{\mu_1\text{-négligeable}} \in \hat{\mathcal{F}}_1. \tag{14.204}$$

□

14.5.3 Mesure image

Le produit d'une mesure par une fonction est défini par la propriété 14.212.

PropJCJQooAdqrGA

Proposition-Définition 14.84 (Mesure image[433])**.**
Soient $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ des espaces mesurables. Soit $\varphi\colon S_1 \to S_2$ une application mesurable. Si m_1 est une mesure positive sur S_1 alors l'application définie par

$$m_2(A_2) = m_1\big(\varphi^{-1}(A_2)\big) \tag{14.205}$$

est une mesure positive sur $(S_2, \mathcal{F}_2)$.

*La mesure m_2 ainsi définie est la **mesure image** de m_1 par l'application φ. Elle est notée $\varphi(m_1)$.*

Démonstration. Il y a deux choses à vérifier pour avoir une mesure positive[25]. D'abord pour l'ensemble vide :

$$m_2(\varnothing) = m_1\big(\varphi^{-1}(\varnothing)\big) = m_1(\varnothing) = 0. \tag{14.206}$$

Ensuite pour l'additivité. Soient A_n dans $\mathcal{F}_2$ des parties deux à deux disjointes et telles que $\bigcup_n A_n \in \mathcal{F}_2$. Alors nous avons

$$\begin{aligned} m_2\big(\bigcup_n A_n\big) &= m_1\Big(\varphi^{-1}\big(\bigcup_n A_n\big)\Big) && \text{(14.207a)}\\ &= m_1\big(\bigcup_n \varphi^{-1}(A_n)\big) && \text{(14.207b)}\\ &= \sum_n m_1\big(\varphi(A_n)\big) && \text{(14.207c)}\\ &= \sum_n m_2(A_n). && \text{(14.207d)} \end{aligned}$$

□

Lemme 14.85.
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi que deux mesures μ et ν sur $(S_1, \mathcal{F}_1)$. Si $\varphi\colon S_1 \to S_2$ est mesurable et si $\mu \leqslant \nu$ alors $\varphi(\mu) \leqslant \varphi(\nu)$.

Démonstration. Soit B mesurable dans $(S_2, \mathcal{F}_2)$ (c'est-à-dire $B \in \mathcal{F}_2$). Alors

$$\varphi(\mu)(B) = \mu\big(\varphi^{-1}(B)\big) \leqslant \nu\big(\varphi^{-1}(B)\big) = \varphi(\nu)(B). \tag{14.208}$$

□

Il est naturel de se demander comment il faut intégrer par rapport à une mesure image. La réponse sera dans le théorème 14.219.

25. Définition 14.18

14.5.4 Régularité d'une mesure

Certaines mesures ont de la compatibilité avec la topologie. Nous allons étudier ça.

LEMooCGKXooYWjRwk

Lemme 14.86 ([447]).
Soit un espace topologique métrique (Ω, d). Nous considérons sa tribu des boréliens[26] *$\mathcal{B}or(\Omega)$ ainsi qu'une mesure finie μ sur $\big(\Omega, \mathcal{B}or(\Omega)\big)$.*

Soit un borélien A de Ω et $\epsilon > 0$.

Il existe un fermé F et un ouvert V de Ω tels que

(1) $F \subset A \subset V$

(2) $\mu(V \backslash F) < \epsilon$.

Démonstration. Soit la famille $\mathcal{D}$ des parties D de Ω qui vérifient la propriété suivante : pour tout $\epsilon > 0$, il existe un fermé F et un ouvert V de Ω tels que $F \subset D \subset V$ et $\mu(V \backslash F) < \epsilon$.

Nous allons prouver que $\mathcal{D}$ est une tribu qui contient tous les ouverts.

(1) **$\mathcal{D}$ contient les ouverts** Soit un ouvert D. Nous posons

$$F_n = \{x \in \Omega \text{ tel que } d(x, D^c) \geqslant 2^{-n}\}. \tag{14.209}$$

(1a) **F_n est fermé**
Le lemme 7.151 montre que le complémentaire F_n^c est ouvert. Donc F_n est fermé.

(1b) $D \subset \bigcup_{n\in\mathbb{N}} F_n$ Si $x \in D$, alors il existe $\delta > 0$ tel que $B(x, \delta) \subset D$ (parce que D est ouvert). Donc $d(x, V^c) \geqslant \delta$. Donc $x \in F_n$ pour $2^{-n} < \delta$.

(1c) $\bigcup_{n\in\mathbb{N}} F_n \subset D$ Si $x \in F_n$, nous avons $d(x, D^c) > 0$, c'est-à-dire que x n'est pas dans D^c. Autrement dit, $x \in D$.

(1d) $\bigcup_{n\in\mathbb{N}} F_n = D$ Nous avons donc l'égalité

$$D = \bigcup_{n\in\mathbb{N}} F_n. \tag{14.210}$$

Vu que $F_n \subset F_{n+1}$, le lemme 14.24(1) nous indique que

$$\lim_{n\to\infty} \mu(F_n) = \mu\Big(\bigcup_{k\in\mathbb{N}} F_k\Big) = \mu(D). \tag{14.211}$$

Étant donné que la mesure est finie, nous pouvons écrire cela sous la forme

$$\mu(D) - \mu(F_n) \to 0. \tag{14.212}$$

Pour chaque n nous avons l'encadrement

$$F_n \subset D \subset D \tag{14.213}$$

où F_n et D sont ouverts. Lorsque ϵ est donné, il suffit de prendre n assez grand pour avoir $\mu(D \backslash F_n) < \epsilon$ pour avoir un encadrement de D par un fermé et un ouvert (D lui-même) dont la différence des mesures est plus petite que ϵ.

Tout cela pour dire que $D \in \mathcal{D}$.

(2) **$\mathcal{D}$ est une tribu** Il faut vérifier les trois points de la définition 14.1.

(2a) $\Omega \in \mathcal{D}$ Nous venons de voir que les ouverts sont dans $\mathcal{D}$. Or Ω est un ouvert.

(2b) $D \in \mathcal{D}$ **implique** $D^c \in \mathcal{D}$ Soit F fermé et V ouvert tels que $F \subset D \subset V$. Nous avons aussi

$$V^c \subset D^c \subset F^c \tag{14.214}$$

où V^c est fermé et F^c est ouvert. De plus $F^c \backslash V^c = V \backslash F$ et donc

$$\mu(F^c \backslash V^c) = \mu(V \backslash F). \tag{14.215}$$

Nous pouvons donc choisir F et V pour avoir $\mu(F^c \backslash V^c) < \epsilon$.

26. Définition 14.50.

(2c) $\bigcup_{i\in\mathbb{N}} D_i \in \mathcal{D}$ Soient $D_i \in \mathcal{D}$. Pour chaque n nous posons

$$F_n \subset D_n \subset V_n \tag{14.216}$$

en choisissant V_n et F_n de telle sorte que $\mu(V_n\backslash F_n) < 2^{-n}\epsilon$.
Nous posons

$$Y_N = \bigcup_{n=0}^{N} F_n, \tag{14.217}$$

et

$$Y = \bigcup_{n=0}^{\infty} F_n. \tag{14.218}$$

Chacun des Y_N est fermé en tant qu'union finie de fermés (lemme 7.6(2)). Mais Y ne l'est pas spécialement [27]. Le lemme 14.24 nous dit cependant que $\mu(Y) = \lim_{N\to\infty}\mu(Y_N)$.
Nous posons

$$D = \bigcup_{n=0}^{\infty} D_n \tag{14.219}$$

ainsi que

$$V = \bigcup_{n\in\mathbb{N}} V_n. \tag{14.220}$$

La partie V est ouverte dans Ω comme union d'ouverts (c'est dans le définition d'une topologie). Nous avons, pour tout N, l'encadrement

$$Y_N = \bigcup_{n=0}^{N} F_n \subset Y \subset D \subset V. \tag{14.221}$$

Nous prouvons à présent que $\lim_{N\to\infty}\mu(V\backslash Y_N) = 0$, de telle sorte que l'encadrement (14.221) dise que $D \in \mathcal{D}$.
D'abord nous avons

$$V\backslash Y \subset \bigcup_n (V_n\backslash F_n) \tag{14.222}$$

parce que si $x \in V\backslash Y$, alors $x \in V_i$ pour un certain i, mais vu que x n'est pas dans Y, il n'est dans aucun des F_n donc en particulier pas dans F_i et $x \in V_n\backslash F_i$.
Un peu de calcul :

$$\begin{aligned}
\mu(V) - \mu(Y) &= \mu(V\backslash Y) && \text{(14.223a)}\\
&\leqslant \mu\Big(\bigcup_n (V_n\backslash F_n)\Big) && \text{(14.223b)}\\
&\leqslant \sum_{n=0}^{\infty} \mu(V_n\backslash F_n) && \text{(14.223c)}\\
&= \sum_{n=0}^{\infty} 2^{-n}\epsilon && \text{(14.223d)}\\
&= 2\epsilon. && \text{(14.223e)}
\end{aligned}$$

Justifications :

— Pour (14.223a), c'est le lemme 14.23.

— Pour (14.223b), c'est (14.222).

— Pour (14.223c), c'est le lemme 14.23(4).

— Pour (14.223e), c'est la série géométrique (10.345).

27. Par exemple $A_n = [1/n, 2]$ sont des fermés dont l'union est $]0, 2]$ qui n'est pas fermé.

Nous choisissons maintenant N assez grand pour que $\mu(Y)-\mu(Y_N)<\epsilon$. Nous avons alors l'encadrement

$$Y_N \subset Y \subset D \subset V \tag{14.224}$$

avec

$$\mu(V\backslash Y_N)=\mu(V)-\mu(Y_N)=\underbrace{\mu(V)-\mu(Y)}_{\leqslant 2\epsilon}+\mu(Y)-\mu(Y_N)\leqslant 2\epsilon+\epsilon=3\epsilon. \tag{14.225}$$

Nous avons donc montré que $\mathcal{D}$ était une tribu contenant les ouverts. Donc $\mathcal{D}$ contient tous les boréliens. □

LEMooZDFVooFUgFGZ

Lemme 14.87 ([447]).
Soit un espace topologique métrique (Ω,d). Nous considérons sa tribu des boréliens $\mathcal{B}or(\Omega)$ ainsi qu'une mesure μ sur $\big(\Omega,\mathcal{B}or(\Omega)\big)$.

Soient un ouvert $W\subset\Omega$ tel que $\mu(W)<\infty$ et un borélien A tel que $A\subset W$. Soit aussi $\epsilon>0$. Il existe un fermé F et un ouvert V tels que

(1) $\mu(V)<\infty$,

(2) $\mu(V\backslash F)<\epsilon$,

(3) et $F\subset A\subset V$.

Démonstration. Vu que la mesure de W est finie, nous considérons la mesure finie

$$\begin{aligned}\nu\colon \mathcal{B}\text{or}(\Omega)&\to[0,\mu(W)]\\ B&\mapsto\mu(B\cap W).\end{aligned} \tag{14.226}$$

La partie A étant borélienne ; par le lemme 14.86, nous avons un fermé F et un ouvert V_1 ouvert tels que

$$F\subset A\subset V_1 \tag{14.227}$$

et $\nu(V_1\backslash F)<\epsilon$. Nous posons $V=V_1\cap W$; vu que $A\subset W$ et $A\subset V_1$ nous avons aussi $A\subset V_1\cap W$ et donc l'encadrement

$$F\subset A\subset V\subset W. \tag{14.228}$$

En ce qui concerne la mesure :

$$\mu(V\backslash F)=\mu(V)-\mu(F)=\mu(V\cap W)-\mu(F\cap W)=\nu(B)-\nu(F)<\epsilon. \tag{14.229}$$

□

ThoPKGEooVrpsGU

Théorème 14.88 ([433]).
Soit X un espace métrique et m une mesure positive bornée sur $\big(X,\mathcal{B}or(X)\big)$. Alors si B est un borélien,

(1) Régularité extérieure : $m(B)=\inf\{m(\Omega) o\grave{u}\ \Omega$ *est un ouvert contenant* $B\}$

(2) Régularité intérieure : $m(B)=\sup\{m(F) o\grave{u}\ F$ *est un fermé,* $F\subset B\}$.

Démonstration. Soit $\mathcal{F}$ l'ensemble des $B\in\mathcal{B}\text{or}(X)$ tels que pour tout $\epsilon>0$, il existe Ω_ϵ ouvert et F_ϵ fermé tels que $F_\epsilon\subset B\subset\Omega_\epsilon$ et $m(\Omega_\epsilon\backslash F_\epsilon)\leqslant\epsilon$. Nous allons montrer que $\mathcal{F}$

— est une tribu

— contient les ouverts

— est inclus dans la tribu borélienne (ça c'est dans la définition de $\mathcal{F}$).

De ces trois points nous déduirons que $\mathcal{F}=\mathcal{B}\text{or}(X)$.

(1) **$\mathcal{F}$ contient les ouverts** Soit Ω un ouvert de X. Alors Ω^c est fermé et $d(x,\Omega^c)=0$ si et seulement si $x\in\Omega^c$ par la proposition 7.300. Nous pouvons donc écrire

$$\Omega^c=\bigcap_{n\geqslant 1}\{x\in X \text{ tel que } d(x,\Omega^c)<\frac{1}{n}\}. \tag{14.230}$$

En passant au complémentaire et en posant $F_n=\{x\in X \text{ tel que } d(x,\Omega^c)\geqslant\frac{1}{n}\}$ nous avons

$$\Omega=\bigcup_{n\geqslant 1}F_n. \tag{14.231}$$

Chacun des F_n est fermé parce que F_n est l'image réciproque du fermé $[\frac{1}{n},\infty[$ par l'application $x\mapsto d(x,\Omega^c)$ qui est continue. De plus les F_n forment une suite croissante, donc le lemme 14.24 nous assure que $m(\Omega)=\lim_{n\to\infty}m(F_n)$. Et le lemme 14.23 que $m(\Omega\backslash F_n)=m(\Omega)-m(F_n)$.

Soit $\epsilon>0$. Il existe alors $n_\epsilon\geqslant 1$ tel que

$$m(\Omega\backslash F_n)=m(\Omega)-m(F_n)\leqslant\epsilon. \tag{14.232}$$

Bref si Ω est ouvert nous considérons $\Omega_\epsilon=\Omega$ et $F_\epsilon=F_{n_\epsilon}$ et nous avons

$$F_\epsilon\subset\Omega\subset\Omega_\epsilon \tag{14.233}$$

avec $m(\Omega_\epsilon\backslash F_\epsilon)\leqslant\epsilon$.

L'ensemble $\mathcal{F}$ contient les ouverts.

(2) **$\mathcal{F}$ est une tribu** Il y a à vérifier les trois conditions de la définition 14.1.

(2a) **Les ensembles faciles** Les ensembles X et $\varnothing$ sont dans $\mathcal{F}$ parce qu'ils sont ouverts et fermés.

(2b) **Complémentaire** Soit $B\in\mathcal{F}$, soit $\epsilon>0$ et les ensembles F_ϵ et Ω_ϵ qui vont avec. Alors en passant au complémentaire nous avons

$$\Omega_\epsilon^c\subset B^c\subset F_\epsilon^c \tag{14.234}$$

De plus

$$F_\epsilon^c\backslash\Omega_\epsilon^c=F_\epsilon^c\cap(\Omega_\epsilon^c)^c=F_\epsilon^c\cap\Omega_\epsilon=\Omega_\epsilon\backslash F_\epsilon. \tag{14.235}$$

Par conséquent

$$m(F_\epsilon^c\backslash\Omega_\epsilon^c)=m(\Omega_\epsilon\backslash F_\epsilon)\leqslant\epsilon. \tag{14.236}$$

Cela montre que $B^c\in\mathcal{F}$.

(2c) **Union dénombrable** Soient (B_n) une suite d'éléments de $\mathcal{F}$ et $\epsilon>0$. Pour chaque n, le lemme 14.86 nous permet de choisir un ouvert Ω_n et un fermé F_n tels que $F_n\subset B_n\subset\Omega_n$ et

$$m(\Omega_n\backslash F_n)\leqslant\frac{\epsilon}{2^{n+2}}. \tag{14.237}$$

Puisque $\Omega_n\backslash B_n\subset\Omega_n\backslash F_n$ nous avons aussi

$$m(\Omega_n\backslash B_n)\leqslant m(\Omega_n\backslash F_n)\leqslant\frac{\epsilon}{2^{n+2}}. \tag{14.238}$$

Nous posons $\Omega=\bigcup_{n\geqslant 1}\Omega_n$ (un ouvert) et $B=\bigcup_{n\geqslant 1}B_n$ ainsi que $A=\bigcup_{n\geqslant 1}F_n$ (qui n'est pas spécialement fermé).

Le but est de majorer $m(\Omega\backslash F)$ où F est un fermé qui est encore à déterminer. Calculons déjà ceci :

$$\Omega\backslash B = \bigcup_n \Omega_n \cap \big(\bigcup_k B_k\big)^c \tag{14.239a}$$
$$= \Big(\bigcup_n \Omega_n\Big) \cap \Big(\bigcap_k B_k^c\Big) \tag{14.239b}$$
$$= \bigcup_n \Big(\Omega_n \cap \big(\bigcap_k B_k^c\big)\Big) \tag{14.239c}$$
$$\subset \bigcup_n \big(\Omega_n \cap B_n^c\big) \tag{14.239d}$$
$$= \bigcup_n (\Omega_n\backslash B_n). \tag{14.239e}$$

Justifications.

— Pour (14.239c). Pour toute suite d'ensembles on a $(\bigcup_k A_k)^x = \bigcap_k A_k^c$.

— Pour (14.239d). Si $x \in \Omega_n \cap (\bigcap_k B_k^c)$ pour un certain n, alors en particulier $x \in \bigcap_k B_k^c$ et $x \in B_n^c$. Donc $x \in \Omega_n \cap B_n^c$.

— Pour (14.239e), notez que l'union n'est pas spécialement disjointe.

Par conséquent,

$$m(\Omega\backslash B) \leqslant \sum_{n=1}^{\infty} m(\Omega_n\backslash B_n) \leqslant \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+2}} = \frac{\epsilon}{4}. \tag{14.240}$$

De la même façon nous avons

$$B\backslash A = \big(\bigcup_{n=1}^{\infty} B_n\big) \cap \big(\bigcup_{k=1}^{\infty} F_n\big)^c \subset \bigcup_{n=1}^{\infty} B_n\backslash F_n. \tag{14.241}$$

Nous avons alors le inégalités de mesures

$$m(B\backslash A) \leqslant \sum_{n=1}^{\infty} m(B_n\backslash F_n) \tag{14.242a}$$
$$\leqslant \sum_{n=1}^{\infty} m(\Omega_n\backslash F_n) \tag{14.242b}$$
$$\leqslant \frac{\epsilon}{4}. \tag{14.242c}$$

C'est vraiment dommage que A ne soit pas en général un fermé, sinon il répondrait à la question. Nous posons $F_1' = F_1$ et $F_n' = \bigcup_{k=1}^n F_k$. En tant qu'unions finies de fermés, les F_n' sont des fermés (lemme 7.6(2)). De plus la suite (F_n') est croissante et l'union est A. Par le lemme 14.24(1) nous avons

$$m(A) = m\Big(\bigcup_n F_n'\Big) = \lim_{n\to\infty} m(F_n'). \tag{14.243}$$

Il existe donc n_ϵ tel que

$$m(A) - m(F_{n_\epsilon}') \leqslant \epsilon \tag{14.244}$$

Nous posons $F = F_{n_\epsilon}'$. Comme $F \subset A$ nous avons aussi $m(A\backslash F) = m(A) - m(F) \leqslant \epsilon$. Et en plus $F \subset A \subset B \subset \Omega$, ce qui donne bien la propriété voulue $F \subset B \subset \Omega$. Il reste à nous assurer de $m(\Omega\backslash F)$. Nous avons d'abord

$$m(B\backslash F) = m\big((B\backslash A) \cup (A\backslash F)\big) = m(B\backslash A) + m(A\backslash F) \leqslant \frac{5\epsilon}{4}. \tag{14.245}$$

Et enfin :

$$m(\Omega\backslash F) = m\big((\Omega\backslash B) \cup (B\backslash F)\big) = m(\Omega\backslash B) + m(B\backslash F) \leqslant \frac{6\epsilon}{4}. \tag{14.246}$$

Et donc à redéfinition près de ϵ, c'est d'accord.

Il est donc établi que $\mathcal{F}$ est une tribu. Qui plus est, l'ensemble $\mathcal{F}$ est une tribu incluse aux boréliens et contenant les ouverts. Ergo $\mathcal{F} = \mathcal{B}\text{or}(X)$.

(3) **Régularité extérieure** Soit B un borélien et $\epsilon > 0$. Alors il existe F_ϵ fermé et Ω_ϵ ouvert tels que $F_\epsilon \subset B \subset \Omega_\epsilon$ et $m(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon$. Vu que $B \subset \Omega_\epsilon$ pour tout ϵ, nous avons aussi

$$m(B) \leqslant \inf_\epsilon m(\Omega_\epsilon). \tag{14.247}$$

Mais comme $m(\Omega_\epsilon) \geqslant m(B)$ pour tout ϵ, nous avons en réalité $m(B) = \inf_\epsilon m(\Omega_\epsilon)$.
Soit maintenant un ouvert Ω tel que $B \subset \Omega$. Nous devons prouver l'existence d'un $\epsilon > 0$ tel que $m(\Omega_\epsilon) \leqslant m(\Omega)$. Cela permettra de conclure que l'infimum sur tous les ouverts contenant B est égal à l'infimum sur les ouverts de la forme Ω_ϵ.
Nous posons $m(\Omega) = m(B) + \delta$ et avec $\epsilon \leqslant \delta$ nous avons

$$m(\Omega_\epsilon \backslash B) \leqslant m(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon \tag{14.248}$$

et donc aussi

$$m(\Omega_\epsilon) \leqslant m(B) + \epsilon \leqslant m(B) + \delta = m(\Omega). \tag{14.249}$$

(4) **Régularité intérieure** Elle se fait de même.

□

DefFMTEooMjbWKK

Définition 14.89.
Soit X un espace topologique et m une mesure positive sur $\big(X, \mathcal{B}or(X)\big)$.

ItemTTPTooStDcpw

*(1) m est une **mesure de Borel** si elle est finie sur tout compact.*

*(2) m est **régulière extérieurement** si $\forall B \in \mathcal{B}or(X)$,*

$$m(B) = \inf\{m(\Omega) \text{ tel que } \Omega \text{ est ouvert et } B \subset \Omega\} \tag{14.250}$$

*(3) m est **régulière intérieurement** si $\forall B \in \mathcal{B}or(X)$,*

$$m(B) = \sup\{m(K) \text{ tel que } K \text{ est compact et } K \subset B\} \tag{14.251}$$

*(4) m est une mesure **régulière** si elle est régulière dans les deux sens.*

*(5) m est une **mesure de Radon** si elle est de Borel et régulière.*

PropNCASooBnbFrc

Proposition 14.90.
Soit X un espace localement compact et dénombrable à l'infini[28] *Alors toute mesure de Borel sur $\big(X, \mathcal{B}or(X)\big)$ est de Radon.*

Démonstration. Nous avons une suite exhaustive[29] de compacts X_k tels que

$$X = \bigcup_{k\geqslant 1} X_k = \bigcup_{k\geqslant 1} \text{Int}(X_k). \tag{14.252}$$

(1) **Régularité intérieure** Soit B, un borélien de X ; nous avons $B = \bigcup_{k\geqslant 1}(B \cap X_k)$ et comme cette union est croissante,

$$m(B) = \lim_{k\to\infty} m(B \cap X_k) \tag{14.253}$$

par le lemme 14.24(1). Dans la suite, il va y avoir beaucoup de considérations sur les topologies induites. Nous nommons τ_k la topologie de X_k induite depuis celle de X. Il ne faudra pas confondre les expressions « un compact *de* X_k » et « un compact *dans* X_k ». La première parle d'un compact pour la topologie τ_k. La seconde parle d'un compact pour la topologie de X, inclus dans X_k.

28. Définitions 7.86 et 7.90.
29. Définition 7.306.

Si $a < m(B)$ alors il existe $k \geqslant 1$ tel que $a < m(B \cap X_k)$, c'est-à-dire

$$a < m(B \cap X_k) \leqslant m(B). \tag{14.254}$$

Mais (X_k, m) est un espace mesuré borné parce que m est de Borel et X_k est compact. Par conséquent la (restriction de la) mesure m est régulière sur l'espace mesuré $\big(X_k, \mathcal{B}\text{or}(X_k)\big)$ par le théorème 14.88. De plus l'ensemble $B \cap X_k$ est un borélien de (X_k, τ_k) parce que

$$B \cap X_k \in \mathcal{B}\text{or}(X)_{X_k} = \mathcal{B}\text{or}(X_k) \tag{14.255}$$

où nous avons utilisé la propriété de compatibilité entre topologie induite et tribu des borélien du théorème 14.57. Il existe donc un fermé F_ϵ de (X_k, τ_k) tel que

$$\begin{cases} F_\epsilon \subset B \cap X_k & \text{(14.256a)} \\ m(B \cap X_k) \leqslant m(F_\epsilon) + \epsilon. & \text{(14.256b)} \end{cases}$$

En mettant bout à bout les inégalités nous avons trouvé

$$a < m(B \cap X_k) \leqslant m(F_\epsilon) + \epsilon. \tag{14.257}$$ EQooXRESooGsGIEQ

L'ensemble F_ϵ est un compact de (X, τ_X). En effet X_k étant fermé de (X, τ_X), le lemme 7.27 nous dit que F_ϵ est un fermé de (X, τ_X). Mais X_k étant compact, F_ϵ est un fermé inclus dans un compact, il est donc compact (lemme 7.96).
Enfin nous prouvons la régularité intérieure de la mesure m, c'est-à-dire que

$$m(B) = \sup\{m(K) \text{ tel que } K \text{ est compact dans } B\} \tag{14.258}$$

en vérifiant les deux conditions de la définition 1.461. D'abord $m(B) \geqslant \{m(K) \text{ tel que } \ldots\}$ parce que $m(B) \geqslant m(K)$ pour tout $K \subset B$. Ensuite prenons $\epsilon > 0$, et considérons l'inégalité (14.257) avec $a = m(B) - \epsilon$. Alors nous avons

$$m(B) - 2\epsilon < m(F_\epsilon). \tag{14.259}$$

Cela prouve que $m(B) - 2\epsilon$ n'est pas un majorant de $\{m(K) \text{ tel que } \ldots\}$.

(2) **Régularité extérieure**
Soit un borélien B de X. Si $m(B) = \infty$ alors tous les ouverts contenant B ont mesure infinie et $m(B)$ en est évidemment l'infimum. Nous supposons donc que $m(B) < \infty$.
Nous notons τ_k la topologie induite de X sur $\text{Int}(X_k)$. Nous posons $B_k = B \cap \text{Int}(X_k)$. L'espace $\big(\text{Int}(X_k), m\big)$ est un espace mesuré borné et $B_k \in \mathcal{B}\text{or}\Big(\text{Int}(X_k)\Big)$. Il existe donc un ouvert Ω_k de $\big(\text{Int}(X_k), \tau_k\big)$ tel que $B_k \subset \Omega_k$ et

$$m(\Omega_k \backslash B_k) \leqslant \frac{\epsilon}{2^k}. \tag{14.260}$$

De plus $\text{Int}(X_k)$ est un ouvert de (X, τ_X), donc en réalité Ω_k est un ouvert de X. Nous posons

$$\Omega = \bigcup_{k=1}^{\infty} \Omega_k \tag{14.261}$$

qui est encore un ouvert de (X, τ_X).
Il est temps de voir que Ω vérifie $m(\Omega \backslash B) \leqslant \epsilon$. Pour cela,

$$\begin{aligned} \Omega \backslash B &= \big(\bigcup_k \Omega_k\big) \cap \big(\bigcup_l B_l\big)^c & \text{(14.262a)} \\ &= \big(\bigcup_k \Omega_k\big) \cap \big(\bigcap_l B_l^c\big) & \text{(14.262b)} \\ &= \subset \bigcup_k (\Omega_k \cap B_k^c) & \text{(14.262c)} \\ &= \bigcup_k (\Omega_k \backslash B_k), & \text{(14.262d)} \end{aligned}$$

ce qui donne au niveau des mesures :

$$m(\Omega\backslash B) \leqslant \sum_{k=1}^{\infty} m(\Omega_k\backslash B_k) \leqslant \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon. \tag{14.263}$$

□

RemooOAGCooRHpjxd

Remarque 14.91.
Exprimé sur $\mathbb{R}^N$, la proposition 14.90 s'exprime en disant que toute mesure de Borel sur $\mathbb{R}^N$ est régulière. Typiquement, l'espace X dont il est question est un ouvert de $\mathbb{R}^N$.

14.5.5 Théorème de récurrence

Soient X un espace mesurable, μ une mesure finie sur X et $\phi\colon X \to X$ une application mesurable[30] préservant la mesure, c'est-à-dire que pour tout ensemble mesurable $A \subset X$,

$$\mu\big(\phi^{-1}(A)\big) = \mu(A). \tag{14.264}$$

Si $A \subset X$ est un ensemble mesurable, un point $x \in A$ est dit **récurrent** par rapport à A si et seulement si pour tout $p \in \mathbb{N}$, il existe $k \geqslant p$ tel que $\phi^k(x) \in A$.

ThoYnLNEL

Théorème 14.92 (Théorème de récurrence de Poincaré.)**.**
Si A est mesurable dans X, alors presque tous les points de A sont récurrents par rapport à A.

Démonstration. Soit $p \in \mathbb{N}$ et l'ensemble

$$U_p = \bigcup_{k=p}^{\infty} \phi^{-k}(A) \tag{14.265}$$

des points qui repasseront encore dans A après p itérations de ϕ. C'est un ensemble mesurable en tant que union d'ensembles mesurables (pour rappel, les tribus sont stables par union dénombrable, comme demandé à la définition 14.1), et nous avons donc

$$\mu(U_p) \leqslant \mu(X) < \infty. \tag{14.266}$$

De plus $U_p = \phi^{-p}(U_0)$, donc $\mu(U_p) = \mu(U_0)$. Vu que $U_p \subset U_0$, nous avons

$$\mu(U_0\backslash U_p) = 0. \tag{14.267}$$

Étant donné que $A \subset U_0$ nous avons a fortiori que

$$\{x \in A \text{ tel que } x \notin U_p\} \subset U_0\backslash U_p, \tag{14.268}$$

et donc

$$\mu\{x \in A \text{ tel que } x \notin U_p\} = 0. \tag{14.269}$$

Cela signifie exactement que l'ensemble des points x de A tels que aucun des $\phi^k(x)$ avec $k \geqslant p$ n'est dans A est de mesure nulle. □

14.6 Mesurabilité des fonctions à valeurs réelles

Nous allons parler de la mesurabilité de fonctions

$$f\colon (S, \mathcal{F}) \to \big(\bar{\mathbb{R}}, \mathcal{B}\text{or}(\bar{\mathbb{R}})\big) \tag{14.270}$$

où $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

30. Définition 14.44.

normooGAAJooUPCbzG

14.93.
Nous convenons que $0 \times \pm\infty = 0$ parce que nous voulons qu'une droite (qui est un rectangle dont une mesure est 0 et l'autre ∞) soit de mesure nulle dans $\mathbb{R}^2$.

Les produits et sommes $\pm\infty \pm \pm\infty$ et $\pm\infty \times \pm\infty$ sont ceux que l'on croit. Sauf bien entendu $+\infty - \infty$ et $1/0$ qui ne sont toujours pas définis.

LEMooBLOLooAdNViv

Lemme 14.94.
L'ensemble B est un borélien de $\bar{\mathbb{R}}$ si et seulement si il existe un borélien B_0 de $\mathbb{R}$ tel que B soit B_0 ou $B_0 \cup \{+\infty\}$ ou $B_0 \cup \{-\infty\}$ ou $B_0 \cup \{+\infty, -\infty\}$.

Démonstration. Comme la topologie usuelle sur $\mathbb{R}$ est la topologie induite de celle sur $\bar{\mathbb{R}}$, la tribu induite l'est aussi par le théorème 14.56. Donc si B est un borélien de $\bar{\mathbb{R}}$, l'ensemble $B \cap \mathbb{R}$ est un borélien de $\mathbb{R}$. □

LemooCRVJooQosHPq

Lemme 14.95 ([433])**.**
Si $\mathcal{S}_0$ est l'ensemble des intervalles du type

$$]\alpha, \beta[, \quad [-\infty, \beta[, \quad]\alpha, +\infty] \tag{14.271}$$

avec $-\infty < \alpha < \beta < +\infty$ alors $\sigma(\mathcal{S}_0) = \mathcal{B}or(\bar{\mathbb{R}})$.

Démonstration. Les intervalles $]\alpha, \beta[$ engendrent la topologie de $\mathbb{R}$[31], donc $\mathcal{B}\text{or}(\mathbb{R}) \subset \sigma(\mathcal{S}_0)$. De plus le lemme 14.3 nous autorise à dire que

$$\bigcap_{n \geqslant 1} [n, +\infty] = \{+\infty\} \in \sigma(\mathcal{S}_0). \tag{14.272}$$

Par conséquent tous les ensembles énumérés dans le lemme 14.94 font partie de $\sigma(\mathcal{S}_0)$. Cela implique que $\mathcal{B}\text{or}(\bar{\mathbb{R}}) \subset \sigma(\mathcal{S}_0)$.

Pour l'inclusion inverse, $\sigma(\mathcal{S}_0)$ est engendré par des parties qui font partie de $\mathcal{B}\text{or}(\bar{\mathbb{R}})$, donc $\sigma(\mathcal{S}_0) \subset \mathcal{B}\text{or}(\bar{\mathbb{R}})$. □

14.6.1 Fonctions à valeurs réelles sur un espace mesurable

THOooWHFLooKYGsOm

Théorème 14.96.
Soient un espace mesurable $(S, \mathcal{F})$ et une fonction $f : S \to \bar{\mathbb{R}}$. Les propriétés suivantes sont équivalentes.

ITEMooHAMHooYLqUhVi

(1) La fonction f est mesurable.

ITEMooHAMHooYLqUhVii

(2) L'ensemble $\{f < a\}$ est dans $\mathcal{F}$ pour tout $a \in \mathbb{R}$

ITEMooHAMHooYLqUhViii

(3) L'ensemble $\{f \leqslant a\}$ est dans $\mathcal{F}$ pour tout $a \in \mathbb{R}$

Démonstration. Plusieurs implications à prouver.

(1) **(1)⇒(2)** Puisque f est mesurable et que $[-\infty, a[\in \mathcal{B}\text{or}(\bar{\mathbb{R}})$, nous avons $f^{-1}\big([-\infty, a[\big) \in \mathcal{F}$.

(2) **(2)⇒(1)** Nous posons $\mathcal{A} = \{[-\infty, a[\text{ tel que } a \in \mathbb{R}\}$.
Nous avons $\mathcal{A} \subset \mathcal{S}_0$ (le $\mathcal{S}_0$ du lemme 14.95). Et de plus,

$$]\alpha, \beta[= [-\infty, \beta[\backslash [-\infty, \alpha] = [-\infty, \beta[\backslash \bigcap_{n \geqslant 1} [-\infty, \alpha + \frac{1}{n}[. \tag{14.273}$$

Donc $]\alpha, \beta[\in \sigma(\mathcal{A})$.
Et aussi :

$$]\alpha, +\infty] = \bar{\mathbb{R}} \backslash \bigcap_{n \in \mathbb{N}} [-\infty, \alpha + \frac{1}{n}[, \tag{14.274}$$

31. Parce toutes les boules sont des intervalles de ce type et que les boules forment une base de topologie, proposition 7.138.

ce qui donne $]\alpha, +\infty] \in \sigma(\mathcal{A})$.

Au final, $\mathcal{S}_0 \subset \sigma(\mathcal{A})$ et donc $\sigma(\mathcal{S}_0) \subset \sigma(\mathcal{A})$. Le lemme 14.95 nous dit que $\sigma(\mathcal{S}_0) = \mathcal{B}\text{or}(\bar{\mathbb{R}})$. Nous avons donc bien $\sigma(\mathcal{S}_0) = \sigma(\mathcal{A}) = \mathcal{B}\text{or}(\bar{\mathbb{R}})$.

Par ailleurs, nous savons que $f^{-1}(\mathcal{A}) \subset \mathcal{F}$ parce que les éléments de $\mathcal{A}$ sont de la forme $\{f < a\}$. Cela donne $\sigma\big(f^{-1}(\mathcal{A})\big) = \mathcal{F}$. Mais $\sigma\big(f^{-1}(\mathcal{A})\big)$ peut aussi s'exprimer par le lemme de transport 14.48 : $\sigma\big(f^{-1}(\mathcal{A})\big) = f^{-1}\big(\sigma(\mathcal{A})\big)$. En combinant les deux,

$$f^{-1}\big(\sigma(\mathcal{A})\big) = \mathcal{F}, \tag{14.275}$$

et en remplaçant $\sigma(\mathcal{A})$ par $\mathcal{B}\text{or}(\bar{\mathbb{R}})$ nous avons ce que nous voulions :

$$f^{-1}\big(\mathcal{B}\text{or}(\bar{\mathbb{R}})\big) \in \mathcal{F}, \tag{14.276}$$

ce qui signifie que f est mesurable.

(3) **(3)⇒(2)** Nous avons

$$\{f < a\} = \bigcup_{n\geqslant 1} \{f \leqslant a - \frac{1}{n}\}. \tag{14.277}$$

donc ceci est une union dénombrable d'éléments de $\mathcal{F}$. Et $\{f < a\}$ est dans $\mathcal{F}$.

(4) **(1)⇒(3)** Nous avons

$$\{f \leqslant a\} = \{f < a\} \cup f^{-1}\big([-\infty, a]\big). \tag{14.278}$$

Le premier ensemble est dans $\mathcal{F}$ par (2). Ensuite $[-\infty, a]$ est un fermé de $\bar{\mathbb{R}}$ et donc un borélien de $\bar{\mathbb{R}}$. Son image réciproque est donc un élément de $\mathcal{F}$ parce que f est mesurable. Au final nous avons bien $\{f \leqslant a\} \in \mathcal{F}$.

□

LemFOlheqw

Lemme 14.97 ([448]).
Une fonction $f\colon X \to \mathbb{R}$ est mesurable si et seulement si $f^{-1}(I)$ est mesurable pour tout I de la forme $]a, \infty[$.

Démonstration. Nous devons prouver que $f^{-1}(A)$ est mesurable dans X pour tout borélien A de $\mathbb{R}$. Nous posons

$$S = \{A \subset \mathbb{R} \text{ tel que } f^{-1}(A) \text{ est mesurable dans } X\} \tag{14.279}$$

et nous prouvons que c'est une tribu. D'abord $f^{-1}(\mathbb{R}) = X$, et X est mesurable, donc $\mathbb{R} \in S$. Ensuite si $A \in S$ alors $f^{-1}(A^c) = f^{-1}(A)^c$. En tant que complémentaire d'un mesurable de X, l'ensemble $f^{-1}(A)^c$ est mesurable dans X. Et enfin si $A_n \in S$ alors $f^{-1}(\bigcup_n A_n) = \bigcup_n f^{-1}(A_n)$ qui est encore mesurable dans X en tant qu'union de mesurables.

Donc S est une tribu qui contient tous les ensembles de la forme $]a, \infty]$. Le lemme 14.52 conclut que S contient tous les boréliens de $\mathbb{R}$. □

LEMooMYUFooKqdDNc

Lemme 14.98.
Soit $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Soit $\lambda \in \bar{\mathbb{R}}$. Les parties $\{x \geqslant \lambda\}$, $\{x > \lambda\}$, $\{x \leqslant \lambda\}$ et $\{x < \lambda\}$ sont des boréliens de $\bar{\mathbb{R}}$.

LEMooAITEooMjHxvh

Lemme 14.99.
Soit une application mesurable $f\colon (\Omega, \mathcal{A}) \to \big(\mathbb{R} \cup \{\pm\infty\}, \mathcal{B}or\big)$. Soit $\lambda \in \mathbb{R}$; nous définissons

$$\begin{aligned} f_\lambda\colon \Omega &\to \mathbb{R} \\ \omega &\mapsto \min\big(f(\omega), \lambda\big). \end{aligned} \tag{14.280}$$

Alors f_λ est mesurable

Démonstration. Soit A mesurable (i.e. borélien) dans $\mathbb{R}\cup\{\pm\infty\}$. Nous devons montrer que $f_\lambda^{-1}(A)$ est mesurable. Pour cela nous écrivons

$$A=\Big(A\cap\{x>\lambda\}\Big)\cup\Big(A\cap\{x\leqslant\lambda\}\Big).\tag{14.281}$$

Par définition f_λ ne prend jamais de valeurs plus grandes que λ, donc $f_\lambda^{-1}\big(A\cap\{x>\lambda\}\big)=\varnothing$. D'autre part, $f_\lambda^{-1}\big(A\cap\{x\leqslant\lambda\}\big)=f^{-1}\big(A\cap\{x\leqslant\lambda\}\big)$.

Étant donné que A et $\{x\leqslant\lambda\}$ sont boréliens[32], l'intersection $A\cap\{x\leqslant\lambda\}$ est borélienne, et donc

$$f_\lambda^{-1}(A)=f^{-1}\big(A\cap\{x\leqslant\lambda\}\big)\in\mathcal{A}.\tag{14.282}$$

□

LemIGKvbNR

Lemme 14.100 ([448])**.**
Soit $f_n\colon X\to\mathbb{R}$ une suite de fonctions mesurables[33]*. Alors $\sup_n f_n$ est mesurable.*

Démonstration. Nous avons

$$\begin{aligned}(\sup f_n)^{-1}\big(]a,\infty]\big)&=\{x\in X\text{ tel que }(\sup f_n)(x)>a\}&(14.283a)\\&=\bigcup_n\{x\in X\text{ tel que }f_n(x)>a\}&(14.283b)\\&=\bigcup_n f_n^{-1}\big(]a,\infty]\big).&(14.283c)\end{aligned}$$

Étant donné que f_n est mesurable et que $]a,\infty]$ est mesurable, chacun des $f_n^{-1}\big(]a,\infty]\big)$ est mesurable dans X. L'ensemble $(\sup f_n)^{-1}\big(]a,\infty]\big)$ est donc une union dénombrable de parties mesurables. Il est donc mesurable.

Le lemme 14.97 conclut que $\sup f_n$ est mesurable. □

PropFYPEOIJ

Proposition 14.101.
Si $f_n\colon X\to\mathbb{R}$ est une suite de fonctions mesurables et positives, alors la fonction[34] *$\sum_n f_n$ est mesurable.*

Démonstration. Nous considérons les fonctions $s_k(x)=\sum_{n=0}^{k}f_n(x)$ qui valent éventuellement ∞ en certains points. Nous avons

$$\sum_n f_n(x)=\sup_k s_k(x),\tag{14.284}$$

donc le lemme 14.100 nous donne la mesurabilité de la somme de f_n. □

ooUDHFooJjKscR

Définition 14.102.
Soit $(S,\mathcal{F})$ un espace mesurable. Une ***partition mesurable dénombrable*** *de S est une suite $(S_n)_{n\geqslant1}$ de parties de S telles que*

(1) $S_n\in\mathcal{F}$ pour tout n,

(2) $S_n\cap S_k=\varnothing$ si $n\neq k$,

(3) $S=\bigcup_{n\geqslant1}S_n$.

LEMooXAPQooPpZUmP

Lemme 14.103 (Lemme de recollement)**.**
Soit (S_n) une partition mesurable dénombrable de l'espace mesurable $(S,\mathcal{F})$. Soit $(S',\mathcal{F}')$ un autre espace mesurable et des fonctions mesurables

$$f_n\colon(S_n,\mathcal{F}_{S_n})\to(S',\mathcal{F}')\tag{14.285}$$

32. Lemme 14.98.
33. Ici X est un espace mesuré et $\mathbb{R}$ est muni des boréliens.
34. Définition 12.365 pour la série de fonctions.

où $\mathcal{F}_{S_n}$ est la tribu induite[35]*. Alors la fonction*

$$\begin{aligned} f\colon (S,\mathcal{F}) &\to (S',\mathcal{F}') \\ x &\mapsto f_n(x) \text{ si } x\in S_n \end{aligned} \tag{14.286}$$

est mesurable.

Démonstration. Soit $A'\in\mathcal{F}'$; nous devons prouver que $f^{-1}(A')\in\mathcal{F}$. Nous savons que

$$f^{-1}(A') = \bigcup_{n\geqslant 1} f_n^{-1}(A'), \tag{14.287}$$

EqooGKFFooEwTdig

qui est une union dénombrable d'éléments $f_n^{-1}(A')\in\mathcal{F}_{S_n}$.

Puisque $S_n\in\mathcal{F}$ nous avons $\mathcal{F}_{S_n}\subset\mathcal{F}$ parce qu'un élément de $\mathcal{F}_{S_n}$ est de la forme $S_n\cap B$ avec $B\in\mathcal{F}$. Ainsi, pour chaque n nous avons

$$f_n^{-1}(A')\in\mathcal{F}_{S_n}\subset\mathcal{F}. \tag{14.288}$$

Au final l'égalité (14.287) écrit $f^{-1}(A')$ comme une union d'éléments de $\mathcal{F}$ et est donc un élément de $\mathcal{F}$. □

PROPooODDVooEEmmTX

Proposition 14.104.
Soit $(S,\mathcal{F})$ un espace mesurable et des applications mesurables $f,g\colon S\to\bar{\mathbb{R}}$. Alors les fonctions suivantes sont mesurables :

(1) λf pour tout $\lambda\in\mathbb{R}$
(2) $f+g$ si elle existe.
(3) $1/f$ si elle existe.
(4) fg.

Démonstration. Commençons par clarifier « si elle existe ». La fonction $f+g$ n'existe pas au point $x\in S$ si $f(x)=+\infty$ et $g(x)=-\infty$. La fonction $1/f$ n'existe pas au point $x\in S$ si $f(x)=0$. Voir le point 14.93.

(1) **La partie où $f+g$ existe est mesurable** La partie de S sur laquelle $f+g$ existe est

$$\{x\in S \text{ tel que } \big(f(x),g(x)\big)\neq(+\infty,-\infty), \big(f(x),g(x)\big)\neq(-\infty,+\infty)\}. \tag{14.289}$$

Nous avons

$$\{(f,g)=(+\infty,-\infty)\} = \{f=\infty\}\cap\{g=-\infty\} \tag{14.290}$$

qui est un ensemble mesurable parce que, par exemple,

$$\{+\infty\} = \bigcap_{n\geqslant 1}[n,+\infty]. \tag{14.291}$$

Le cas $(-\infty,+\infty)$ est identique, et au final la partie de S sur laquelle $f+g$ n'existe pas est mesurable. Par complémentarité la partie sur laquelle $f+g$ existe est également mesurable[36].

(2) **Idem pour la partie sur laquelle $1/f$ existe** Idem.

(3) **Mesurabilité de λf** Si $\lambda=0$, nous avons une fonction constante dont la mesurabilité est évidente[37]. Nous supposons $\lambda>0$. Alors

$$\{\lambda f<a\} = \{f<a/\lambda\}\in\mathcal{F}. \tag{14.292}$$

Pour $\lambda<0$ nous avons de la même manière

$$\{\lambda f<a\} = \{f>a/\lambda\}\in\mathcal{F}. \tag{14.293}$$

Ce dernier point est suffisant pour que λf soit mesurable par le théorème 14.96(3) et par complémentarité.

35. Définition 14.8.
36. Parfois on a envie de dire que l'affirmation « A est mesurable » ne passe pas le test de Popper.
37. Prenez quand même le temps d'y penser.

(4) **Mesurabilité de $f+g$** Soit $a \in \mathbb{R}$; le théorème 14.96 nous demande d'avoir envie de prouver que $\{f+g<a\} \in \mathcal{F}$. Nous avons

$$f(x)+g(x)<a \tag{14.294}$$

si et seulement si

$$f(x)<a-g(x) \tag{14.295}$$

si et seulement si

$$\exists q \in \mathbb{Q} \text{ tel que } f(x)<q<a-g(x). \tag{14.296}$$

Donc

$$\{f+g<a\}=\bigcup_{q\in\mathbb{Q}}\Big(\{f<q\}\cap\{g<a-q\}\Big), \tag{14.297}$$

qui est une union dénombrable d'éléments de $\mathcal{F}$. Donc $\{f+g<a\} \in \mathcal{F}$ et $f+g$ est mesurable. Note qu'en toute rigueur il faudrait « $\cap$là où $f+g$ est définie » un peu partout, mais cela ne change rien parce que l'intersection de deux parties mesurables est mesurable.

(5) **Mesurabilité de $1/f$** Soit $a \in \mathbb{R}$. Si $a>0$ alors

$$\{1/f<a\}=\{f<0\}\cup\{f>\frac{1}{a}\} \in \mathcal{F}. \tag{14.298}$$

et si $a<0$ alors

$$\{1/f<a\}=\{f<0\}\cap\{f>\frac{1}{a}\} \in \mathcal{F}. \tag{14.299}$$

(6) **Mesurabilité de fg** Nous allons la prouver en plusieurs fois.

(6a) **Si f est mesurable alors f^2 est mesurable** Si $a \leqslant 0$ alors $\{f^2<a\}=\varnothing$. Si $a>0$ nous avons

$$\{f^2<a\}=\{-\sqrt{a}<f<\sqrt{a}\} \in \mathcal{F}. \tag{14.300}$$

(6b) **$f\mathbb{1}_A$ est mesurable** Soit $A \in \mathcal{F}$, et prouvons que $f\mathbb{1}_A$ est mesurable. Par définition,

$$(f\mathbb{1}_A)(x)=\begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases} \tag{14.301}$$

Nous posons

$$\begin{aligned} f_1\colon A^c &\to \bar{\mathbb{R}} \\ x &\mapsto 0 \end{aligned} \tag{14.302}$$

et

$$\begin{aligned} f_2\colon A &\to \bar{\mathbb{R}} \\ x &\mapsto f(x). \end{aligned} \tag{14.303}$$

Alors nous avons

$$(\mathbb{1}_A f)(x)=\begin{cases} f_1(x) & \text{si } x \in A^c \\ f_2(x) & \text{si } x \in A. \end{cases} \tag{14.304}$$

Les ensembles A et A^c forment une partition mesurable dénombrable de S. La fonction f_1 est mesurable ; pour prouver que f_2 est mesurable, nous l'écrivons $f_2=f\circ j_A$ où $j_A\colon A\to S$ est l'injection canonique. L'application

$$j_A\colon (A,\mathcal{F}_A)\to(S,\mathcal{F}) \tag{14.305}$$

est mesurable parce que si $B \in \mathcal{F}$ alors $j_A^{-1}(B)=A\cap B \in \mathcal{F}_A$. D'autre part l'application

$$f\colon (S,\mathcal{F})\to\big(\bar{\mathbb{R}},\mathcal{B}\text{or}(\bar{\mathbb{R}})\big) \tag{14.306}$$

est mesurable par hypothèse. La composée $f_2=f\circ j_A$ est alors mesurable par la proposition 14.45. Le lemme de recollement 14.103 nous donne alors la mesurabilité de $f\mathbb{1}_A$.

(6c) **Le produit fg est mesurable** Nous posons

$$F = \{x \in S \text{ tel que } |f(x)| < +\infty, |g(x)| < \infty\}. \tag{14.307}$$

En tant qu'intersection de deux ensembles mesurables, F est mesurable. Par la partie précédente, les applications $f_1 = f\mathbb{1}_F$ et $g_1 = g\mathbb{1}_F$ sont mesurables. L'application $f_1 + g_1 \colon S \to \mathbb{R}$ est encore mesurable. Par conséquent l'application

$$f_1 g_1 = \frac{1}{2}\big((f_1 + g_1)^2 - f_1^2 - g_1^2\big) \tag{14.308}$$

est mesurable.

Voyons maintenant ce qui se passe en dehors de F. Nous allons utiliser le lemme de recollement sur la fonction

$$(fg)(x) = \begin{cases} (f_1 g_1)(x) & \text{si } x \in F \\ -\infty & \text{si } x \in \mathcal{U} \\ 0 & \text{si } x \in \mathcal{V} \\ +\infty & \text{si } x \in \mathcal{W} \end{cases} \tag{14.309}$$

où $F, \mathcal{U}, \mathcal{V}, \mathcal{W}$ forment une partition mesurable dénombrable[38] de S. Pour le sport nous montrons que $\mathcal{U}$ est mesurable :

$$\begin{aligned} \mathcal{U} = &\big(\{f = -\infty\} \cap \{g > 0\}\big) && \text{(14.310a)} \\ \cup &\big(\{f = +\infty\} \cap \{g < 0\}\big) && \text{(14.310b)} \\ \cup &\big(\{g = -\infty\} \cap \{f > 0\}\big) && \text{(14.310c)} \\ \cup &\big(\{g = +\infty\} \cap \{f < 0\}\big). && \text{(14.310d)} \end{aligned}$$

□

ooABKWooPbfSOZ

Proposition 14.105.
Si $f_n \colon S \to \bar{\mathbb{R}}$ est une suite de fonctions mesurables, alors les fonctions $\inf_n f_n$ et $\sup_n f_n$ sont mesurables.

Démonstration. Nous avons les découpages

$$\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\} \in \mathcal{F} \tag{14.311}$$

et

$$\{\sup_n f_n \leqslant a\} = \bigcap_n \{f_n \leqslant a\} \in \mathcal{F}. \tag{14.312}$$

EQooNYKVooDOjQXM

Le théorème 14.96 permet de conclure. □

Note : pour (14.312) nous ne pouvions pas utiliser les inégalités strictes parce que $\{\sup_n f_n < a\}$ n'est pas spécialement égal à $\bigcap_n \{f_n < a\}$.

LEMooMGUOooMGwknZ

Lemme 14.106.
Soient des fonctions mesurables $f_i \colon \Omega \to \mathbb{R}$ $(i = 1, \dots, n)$. Alors la fonction

$$\begin{aligned} g \colon \Omega &\to \mathbb{R} \\ x &\mapsto \max\{f_1(x), \dots, f_n(x)\} \end{aligned} \tag{14.313}$$

est mesurable.

38. Définition 14.102.

14.107.
La proposition 14.105 nous permet de définir les parties positives et négatives de f par $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$. Ce sont des applications mesurables. Nous avons les décompositions

$$f = f^+ - f^- \tag{14.314a}$$
$$|f| = f^+ + f^-. \tag{14.314b}$$

CORooNXYUooEcvDlP

Corolaire 14.108.
Si $f\colon S \to \bar{\mathbb{R}}$ est mesurable alors les applications f^+, f^- et $|f|$ sont mesurables en tant qu'applications $S \to \bar{\mathbb{R}}^+$.

Démonstration. Nous faisons la preuve pour f^+. Nous savons que $f^+\colon S \to \bar{\mathbb{R}}$ est mesurable par la proposition 14.105. Nous considérons l'injection canonique $j\colon \bar{\mathbb{R}}^+ \to \bar{\mathbb{R}}$ et

$$\begin{aligned} f_1^+\colon S &\to \bar{\mathbb{R}}^+ \\ x &\mapsto f^+(x). \end{aligned} \tag{14.315}$$

Alors $f_1^+ = j \circ f^+$ est mesurable. Et c'est bien cela que nous voulions.

□

Note : f^+ et f_1^+ sont exactement les mêmes fonctions. Elles ne diffèrent que par la tribu que nous considérons sur l'espace d'arrivée. Nous allons à partir de maintenant les noter toutes deux f^+.

Remarque 14.109.
L'application $|f|$ peut être mesurable sans que f le soit. Soit en effet une partie $A \notin \mathcal{F}$, et posons

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in A^c. \end{cases} \tag{14.316}$$

Alors $f^{-1}(\{1\}) = A$ n'est pas mesurable alors que $|f|(x) = 1$ pour tout x.

Il est temps d'aller relire les définitions 9.42.

PropooMFIBooJzaleK

Proposition 14.110.
Si les fonctions $f_n\colon S \to \bar{\mathbb{R}}$ sont mesurables alors les fonctions $\limsup f_n$ et $\liminf f_n$ sont mesurables.

Démonstration. Par le lemme 9.44 nous écrivons $\limsup_n f_n(x) = \inf_{n\geqslant 1} \sup_{k\geqslant n} f_k(x)$. Pour chaque k nous considérons la fonction $g_k = \sup_{n\geqslant k} f_n$. Par la proposition 14.105, les fonctions g_k sont mesurables. En utilisant encore la même proposition, $\inf_{n\geqslant 1} g_k$ est encore mesurable. □

PropooDXBGooSFqrai

Proposition 14.111 ([449])**.**
Si $f_n\colon S \to \bar{\mathbb{R}}$ est une suite de fonctions mesurables dont la limite ponctuelle existe, alors la limite est mesurable.

Démonstration. Si la limite existe, elle est égale à la limite supérieure par le lemme 9.45. Or la limite supérieure est mesurable par la proposition 14.110. □

14.6.2 Fonction étagée

DefBPCxdel

Définition 14.112 ([450])**.**
Soit $(S, \mathcal{F})$ un espace mesurable et une fonction $f\colon S \to \big(\bar{\mathbb{R}}, \mathcal{B}or(\bar{\mathbb{R}})\big)$. Il serait dommage de confondre les trois concepts suivants.

— *Une **fonction simple** est une fonction dont l'image est constituée d'un nombre fini de valeurs.*

— *Une **fonction étagée** est une fonction simple qui est elle-même une fonction mesurable.*
— *Une **fonction en escalier** est une fonction étagée dont les valeurs sont constantes sur des intervalles : ce sont donc des fonctions constantes par morceaux.*

Dans les trois cas, la fonction f peut être écrite comme somme de fonctions caractéristiques :

$$f(x) = \sum_{j=1}^{p} \alpha_j \mathbb{1}_{A_j}(x) \tag{14.317}$$

où $A_j = f^{-1}(\alpha_j)$. Ce qui change est la nature des A_j.

— Si f est simple, les A_j sont quelconques.
— Si f est étagée, les A_i peuvent être choisis mesurables parce que $\{\alpha_i\}$ est un borélien, ce qui fait de $A_i = f^{-1}(\alpha_i)$ un choix mesurable.
— Si f est en escalier, les A_i sont des intervalles.

Définition 14.113.
*La **forme canonique** d'une fonction simple f est la suivante. Soit $\{\alpha_i\}_{i=1,\dots,l}$ les valeurs distinctes prises par f et $A_i = f^{-1}(\alpha_i)$. La forme canonique de f est alors*

$$f = \sum_{i=1}^{l} \alpha_i \mathbb{1}_{A_i}. \tag{14.318}$$

LEMooNWLTooCDuRQI

Lemme 14.114.
Si f est une fonction simple dont la représentation canonique est

$$f = \sum_{i=1}^{l} \alpha_i \mathbb{1}_{A_i}, \tag{14.319}$$

alors

(1) les A_i sont disjoints,
(2) l'union est égale à tout l'ensemble : $S = \bigcup_i A_i$.

¡¡ Avertissement/question à la lectrice !! 14.115
Le lemme 14.116 et le théorème 14.118 disent la même chose alors que la preuve du théorème 14.118 est beaucoup plus compliquée.La démonstration du lemme serait fausse ?

M'est avis que ce que le théorème donne en plus est la convergence uniforme en cas de fonction bornée. La suite (14.320) ne va pas converger uniformément.

LemYFoWqmS

Lemme 14.116 (Limite croissante de fonctions étagées[1])**.**
Soit $f\colon (S, \mathcal{F}) \to \bar{\mathbb{R}}$ une fonction positive mesurable. Il existe une suite $f_n\colon S \to \mathbb{R}$ de fonctions étagées positives telles que $f_n \to f$ ponctuellement et $f_n \leqslant f$.

Démonstration. Nous considérons (q_n) une suite parcourant tous les rationnels positifs[39] avec $q_0 = 0$ pour être sûr. Pour $n \in \mathbb{N}$ nous définissons la fonction

$$f_n(x) = \max\{q_i \text{ tel que } i \leqslant n,\ q_i \leqslant f(x)\}. \tag{14.320}$$

EqooXQYJooSSJwtM

L'ensemble sur lequel le maximum est pris n'est pas vide parce que $q_0 = 0$. La fonction f_n est simple parce qu'elle ne prend que n valeurs différentes. Nous avons aussi, par construction, $f_n(x) \leqslant f(x)$. Et aussi pour tout $x \in S$, $f_n(x) \to f(x)$, parce que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

En ce qui concerne le fait que f_n soit mesurable, nous notons $\{r_0, \dots, r_n\}$ l'ensemble des $\{q_0, \dots, q_n\}$ classés dans l'ordre croissant. Nous posons en plus $r_{n+1} = +\infty$. Nous avons alors

$$f_n^{-1}(r_k) = \{x \in S \text{ tel que } f(x) \geqslant r_k, f(x) < r_{k+1}\} = \{f \geqslant r_k\} \cap \{f < r_{k+1}\}. \tag{14.321}$$

En tant qu'intersection de deux ensembles mesurables, le théorème 14.96 dit que $f_n^{-1}(r_k)$ est mesurable. □

39. Nous rappelons que $\mathbb{Q}$ est dénombrable et dense dans $\mathbb{R}$ par la proposition 9.16.

Remarque 14.117.
Pour avoir $f_n < |f|$ nous pouvons poser

$$f_n(x) = \begin{cases} \max\{q_i \text{ tel que } i \leqslant n,\ q_i \leqslant f(x)\} & \text{si } f(x) \geqslant 0 \\ \min\{q_i \text{ tel que } i \leqslant n,\ q_i \geqslant f(x)\} & \text{si } f(x) < 0. \end{cases} \tag{14.322}$$

THOooXHIVooKUddLi

Théorème 14.118 (Théorème fondamental d'approximation, thème 24[433, 447, 451])**.**
Soit un espace mesurable $(S, \mathcal{A})$.

(1) Soit une fonction mesurable $f : S \to [0, +\infty]$. Alors il existe une suite croissante de fonctions $\varphi_n : S \to [0, +\infty[$ étagées[40] positives dont la limite ponctuelle est f.

(2) Si de plus f est bornée, la convergence est uniforme.

(3) Idem pour f à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$.

Démonstration. Nous découpons l'intervalle $[0, n]$ en plusieurs morceaux.

$$I_{n,k} = \begin{cases} [\frac{k}{2^n}, \frac{k+1}{2^n}[& \text{si } 0 \leqslant k \leqslant n2^n - 1 \\ [n, \infty] & \text{si } k = n2^n. \end{cases} \tag{14.323}$$

Nous posons $S_{n,k} = f^{-1}(I_{n,k})$. Ce sont des ensembles mesurables parce que f est mesurable. Et de plus, pour chaque n, la suite $(S_{n,k})_{k\geqslant 0}$ est une partition mesurable finie de S. Nous posons

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{S_{n,k}}. \tag{14.324}$$

C'est-à-dire que sur chaque $S_{n,k}$ nous approximons f par le bas. La fonction φ_n est étagée et positive : $0 \leqslant \varphi_n(x) \leqslant f(x)$ par construction.

(1) **Croissance** Nous allons voir que $\varphi_n \leqslant \varphi_{n+1}$. Soit $k \neq n2^n$. Si $x \in S_{n,k}$ alors $\varphi_n(x) = \frac{k}{2^n}$ et nous avons aussi la décomposition

$$S_{n,k} = S_{n+1,2k} \cup S_{n+1,2k+1}. \tag{14.325}$$

Si $x \in S_{n+1,2k}$ alors $\varphi_{n+1}(x) = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = \varphi_n(x)$. Et si $x \in S_{n+1,2k+1}$ alors

$$\varphi_{n+1}(x) = \frac{2k+1}{2^{n+1}} = \frac{k + \frac{1}{2}}{2^n} > \varphi_n(x). \tag{14.326}$$

Il reste à traiter le cas $x \in \{f \geqslant n\}$. Dans ce cas nous avons $\varphi_n(x) = n$. Il y a encore deux cas à traiter :

$$\{f \geqslant n\} = \{f \in [n, n+1[\} \cup \{f \in [n+1, \infty]\}. \tag{14.327}$$

Pour plus de simplicité dans les notations, nous notons $\bar{n} = n2^n$, c'est-à-dire que $I_{n,\bar{n}}$ est le $I_{n,k}$ avec le k le plus grand possible. Nous avons

$$I_{n,\bar{n}} = [n, n+1[\cup [n+1, \infty]. \tag{14.328}$$

Le premier élément se décompose en $I_{n+1,k}$ avec $k < \bar{n} + 1$ (nous préciserons plus tard exactement les valeurs de k) tandis que le second est $[n+1, \infty] = I_{n+1,\overline{n+1}}$.
Pour $x \in S_{n+1,\overline{n+1}}$ nous avons

$$\varphi_{n+1}(x) = \frac{(n+1)2^{n+1}}{2^{n+1}} = n + 1 > \varphi_n(x). \tag{14.329}$$

Si au contraire $f(x) \in [n, n+1[$ nous devons précisément voir quels sont les k qui font en sorte que $I_{n+1,k}$ recouvre $[n, n+1[$. Le plus petit k est donné par $\frac{k}{2^{n+1}} = n$, c'est-à-dire

40. Définition 14.112.

$k = n2^{n+1}$ et le plus grand k est donné par $\frac{k}{2^{n+1}} < n+1$, c'est-à-dire $k = 2^{n+1}(n+1) - 1$. Donc si $f(x) \in [n, n+1[$ alors $x \in S_{n+1,k}$ avec

$$n2^{n+1} \leqslant k \leqslant (n+1)2^{n+1} - 1 \tag{14.330}$$

Dans ce cas

$$\varphi_{n+1}(x) = \frac{k}{2^{n+1}} \geqslant \frac{n2^{n+1}}{2^{n+1}} = n = \varphi_n(x). \tag{14.331}$$

(2) **Convergence ponctuelle** Si $f(x) < \infty$ alors il existe[41] $n_0 \in \mathbb{N}$ tel que $f(x) < n_0$. Pour $n \geqslant n_0$ nous avons $f(x) < n$ et donc $\varphi_n(x)$ se calcule à partir d'un des intervalles de taille $1/2^n$:

$$\varphi_n(x) = \frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}. \tag{14.332}$$

Donc

$$|\varphi_n(x) - f(x)| \leqslant \frac{1}{2^n}, \tag{14.333}$$

ce qui signifie que $\lim_{n\to\infty} \varphi_n(x) = f(x)$.

Si $f(x) = +\infty$ alors $f(x) > n$ pour tout n. Et alors $\varphi_n(x) = n$ pour tout n, ce qui donne bien $\varphi_n(x) \to \infty$.

(3) **Convergence uniforme** Soit f bornée : $0 \leqslant f(x) < M$ pour tout $x \in S$. Soit aussi $\epsilon > 0$. Nous prenons $n_0 > M$ tel que $\frac{1}{2^{n_0}} < \epsilon$. Alors pour tout $n \geqslant n_0$ nous avons

$$0 \leqslant f(x) - \varphi_n(x) \leqslant \frac{1}{2^n} \leqslant \frac{1}{2^{n_0}} \leqslant \epsilon. \tag{14.334}$$

Notez qu'il n'y a pas de valeurs absolues parce que nous savons déjà que la limite est croissante.

□

ThofrestemesurablesXYYX

Théorème 14.119 (Doob[432]).
Soit une application mesurable[42] $X : \Omega \to \mathbb{R}^d$. *Une application* $Y : \Omega \to \mathbb{R}^p$ *est* $\mathcal{A}_X$*-mesurable si et seulement si il existe une fonction borélienne* $f : \mathbb{R}^d \to \mathbb{R}^p$ *telle que* $Y = f(X)$.

Démonstration. En séparant Y par coordonnées, et en séparant, pour chacune, les parties positives et négatives, nous supposons que Y est à valeurs réelles positives. Le théorème 14.118 dit qu'il existe une suite croissante d'applications $\mathcal{A}_X$-mesurables et étagées $\varphi_n : \Omega \to [0, \infty[$ telles que $\varphi_n \to Y$ ponctuellement.

Vu que φ_n est étagée, il existe des $\mathcal{A}_X$-mesurables $A_{nk} \subset \Omega$ tels que $\varphi_n = \sum_{k=0}^{s_n} a_{ak} \mathbb{1}_{A_{nk}}$ avec $a_{nk} > 0$. Par la définition 14.6 de la tribu engendrée par une application, il existe un borélien $B_{nk} \subset \mathbb{R}$ tel que $A_{nk} = X^{-1}(B_{nk})$. Avec ça nous avons

$$\varphi_n = \sum_{k=0}^{s_n} a_{nk} \mathbb{1}_{A_{nk}} \tag{14.335a}$$

$$= \sum_k a_{nk} \mathbb{1}_{X^{-1}(B_{nk})} \tag{14.335b}$$

$$= \sum_k a_{nk} (\mathbb{1}_{B_{nk}} \circ X) \tag{14.335c}$$

$$= \left(\sum_k a_{nk} \mathbb{1}_{B_{nk}} \right) \circ X. \tag{14.335d}$$

41. Le vrai snob citera ici le lemme 1.438.
42. Elle est notée X parce que l'application usuelle de ce théorème est en théorie des variables aléatoires.

Nous posons

$$\begin{aligned} f_n\colon \mathbb{R} &\to \mathbb{R}^+ \\ f_n &= \sum_{k=0}^{s_n} a_{nk}\mathbb{1}_{B_{nk}}. \end{aligned} \tag{14.336}$$

Pour $x \in \mathbb{R}$, il n'y a que deux possibilités. Soit $x \in X(\Omega)$, soit non. Si x n'est pas dans $X(\Omega)$, alors il n'est dans aucun des B_{nk} et nous avons $f_n(x) = 0$ pour tout n. Si $x \in X(\Omega)$, alors il existe $\omega \in \Omega$ tel que $x = X(\omega)$. Dans ce cas

$$f_n(x) = f_n\big(X(\omega)\big) = \varphi_n(\omega), \tag{14.337}$$

qui est croissante en n. De plus si $x = X(\omega)$, nous avons

$$f_n(x) = \varphi_n(\omega) \to Y(\omega). \tag{14.338}$$

Nous avons donc montré que pour tout x, $n \mapsto f_n(x)$ est convergente :

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{si } x \notin X(\Omega) \\ Y(\omega) & \text{si } x = X(\omega). \end{cases} \tag{14.339}$$

Nous notons $f = \lim_{n\to\infty} f_n$, et nous avons

$$Y(\omega) = \lim_{n\to\infty} \varphi_n(\omega) = \lim_{n\to\infty} \Big((f_n \circ X)(\omega)\Big) = \lim_{n\to\infty} f_n\big(X(\omega)\big) = f\big(X(\omega)\big) = (f\circ X)(\omega). \tag{14.340}$$

Nous avons donc bien trouvé une application $f\colon \mathbb{R} \to \mathbb{R}$ telle que $Y = f \circ X$. □

14.6.3 Fonctions réelles à variables réelles

Nous nous focalisons à présent sur le cas des fonctions

$$f\colon \big(\mathbb{R}, \mathcal{B}\mathrm{or}(\mathbb{R})\big) \to \big(\bar{\mathbb{R}}, \mathcal{B}\mathrm{or}(\bar{\mathbb{R}})\big). \tag{14.341}$$

NORMooNFOMooYnaflN

14.120 ([1])**.**
Anticipons un peu pour expliquer pourquoi ce que nous allons faire maintenant est suffisant pour ce que nous avons en tête [43]. Toutes les fonctions mesurables

$$f\colon \big(\mathbb{R}, \mathcal{B}\mathrm{or}(\mathbb{R})\big) \to \big(\bar{\mathbb{R}}, \mathcal{B}\mathrm{or}(\bar{\mathbb{R}})\big) \tag{14.342}$$

seront a fortiori mesurables au sens de

$$f\colon \big(\mathbb{R}, \mathcal{L}\mathrm{eb}(\mathbb{R})\big) \to \big(\bar{\mathbb{R}}, \mathcal{B}\mathrm{or}(\bar{\mathbb{R}})\big) \tag{14.343}$$

où $\mathcal{L}\mathrm{eb}(\mathbb{R})$ est la tribu de Lebesgue sur $\mathbb{R}$, c'est-à-dire la tribu complétée de celle des boréliens (définition 14.143).

14.121.
Nous allons maintenant donner quelques conditions pour que des fonctions soient mesurables au sens de la tribu des boréliens sur l'espace d'arrivée et de départ. Ces résultats seront donc immédiatement applicables à la théorie de l'intégration où nous considérons la tribu de Lebesgue sur l'espace de départ.

Autrement dit, les résultats présentés ici sont un peu plus forts que ce dont nous avons réellement besoin . . .ou alors ce sont les hypothèses que nous allons poser en théorie de l'intégration, qui seront un peu plus fortes que nécessaires. C'est une question de point de vue.

43. Pour rappel, nous avons en tête de définir une théorie de la mesure afin d'y définir des intégrales. En particulier nous allons étudier l'intégrale de Lebesgue et en ce qui concerne $\mathbb{R}^n$, nous aurons la tribu de Lebesgue.

CorooJYDVooCrXVun

Corolaire 14.122.
Si I est un intervalle de $\mathbb{R}$, alors toute application monotone $f : I \to \mathbb{R}$ est borélienne.

Démonstration. Puisque f est monotone, l'ensemble $\{f < a\}$ est un intervalle. Or tous les intervalles sont boréliens, donc f est mesurable par le théorème 14.96. □

Définition 14.123.
*Si I est un intervalle de $\mathbb{R}$, une fonction $f : I \to \mathbb{R}$ est **monotone par morceaux** si il existe une suite strictement croissante de points $(x_i)_{i\in\mathbb{Z}}$ dans I telle que f ait la propriété sur chacun des ouverts $]x_j, x_{j+1}[$..*

Remarque 14.124.
Quelques remarques.

(1) Dans cette définition, les points sont numérotés par $\mathbb{Z}$ et non par $\mathbb{N}$ parce que nous nous laissons la liberté d'avoir une infinité de points de chacun des deux côtés.

(2) La notion de C^1 par morceaux sera la définition 20.37. Attention : ce ne sera pas la même.

PropooLNBHooBHAWiD

Proposition 14.125.
Soit I un intervalle de $\mathbb{R}$ et une fonction $f : I \to \mathbb{R}$. Si f est continue ou monotone par morceaux sur I alors elle y est borélienne.

Démonstration. L'ensemble $\{]x_j, x_{j+1}[\}_{j\in\mathbb{Z}} \cup \{x_i\}_{i\in\mathbb{Z}}$ forme une partition mesurable dénombrable de I (les singletons sont des boréliens). À une belle redéfinition près de la numérotation (deux fois $\mathbb{Z}$ va dans $\mathbb{N}$), nous les appelons $(I_n)_{n\in\mathbb{N}}$, et nous définissons les fonctions f_k comme étant les restrictions de f aux intervalles I_k.

Toute fonction sur un singleton est mesurable. Toute fonction continue sur un ouvert est mesurable (théorème 14.56). Toute fonction monotone sur un ouvert est mesurable (corolaire 14.122).

Le lemme de recollement 14.103 donne alors la mesurabilité de f. □

14.126.
Toutes les fonctions que nous pouvons écrire explicitement sont mesurables ...en tout cas toutes celles que l'on trouve en pratique. En effet nous avons déjà toutes les fonctions continues par morceaux via la proposition 14.125 et ensuite toutes les limites par la proposition 14.111. Cela donne les séries, les dérivées, les primitives, etc.

14.7 Tribu produit

14.7.1 Produit d'espaces mesurables

DefTribProfGfYTuR

Définition 14.127 (Tribu produit)**.**
*Si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux tribus sur deux ensembles Ω_1 et Ω_2, nous définissons la **tribu produit** $\mathcal{A}_1 \otimes \mathcal{A}_2$ comme étant la tribu engendrée par*

$$\{X \times Y \text{ tel que } X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}. \tag{14.344}$$

*Ces ensembles sont appelés **rectangles** de $(\Omega_1, \mathcal{A}_1) \otimes (\Omega_2, \mathcal{A}_2)$.*

PropLJJWooKqWlTr

Proposition 14.128 ([452])**.**
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$. Si $\mathcal{C}_i$ est une classe de parties de S_i avec $\mathcal{F}_i = \sigma(\mathcal{C}_i)$ et $S_i \in \mathcal{C}_i$. Alors

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{14.345}$$

Démonstration. Nous notons p_1 et p_2 les projections de $S_1 \times S_2$ vers S_1 et S_2. Nous commençons par prouver que

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\big(p_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)\big). \tag{14.346}$$

eqSGPBooLpQHfq

En effet cette union est dans $\mathcal{F}_1 \otimes \mathcal{F}_2$ parce que ce sont tous des produits de la forme $A_1 \times S_2$ et $S_1 \times A_2$ où $A_i \in \mathcal{F}_i$. Inversement, tous les produits de la forme $A_1 \times A_2$ sont dans la tribu engendrée par l'union parce que

$$A_1 \times A_2 = (A_1 \times S_2) \cap (S_1 \times A_2). \tag{14.347}$$

Par conséquent, la partie $p_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)$ engendre tous les produits qui engendrent la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2$. L'égalité (14.346) est donc correcte.

Si $C_1 \in \mathcal{C}_1$ alors

$$p_1^{-1}(C_1) = C_1 \times S_2 \in \mathcal{C}_1 \times \mathcal{C}_2 \tag{14.348}$$

et donc $p_1^{-1}(\mathcal{C}_1) \subset \mathcal{C}_1 \times \mathcal{C}_2$. En utilisant le lemme de transport 14.48 nous avons alors

$$p_1^{-1}(\mathcal{F}_1) = p_1^{-1}\big(\sigma(\mathcal{C}_1)\big) = \sigma\big(p_1^{-1}(\mathcal{C}_1)\big) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2) \tag{14.349}$$

EqDQLYooVOLqMZ

et de la même façon,

$$p_2^{-1}(\mathcal{F}_2) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{14.350}$$

EqMTRCooVHNTHJ

Vu les relations (14.349), (14.350) et (14.346) nous avons

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\big(p_1^{-1}(\mathcal{F}_1) \cup p_2^{-1}(\mathcal{F}_2)\big) \subset \sigma(\mathcal{C}_1 \times \mathcal{C}_2). \tag{14.351}$$

Réciproquement, si $C_1 \in \mathcal{C}_1$ et $C_2 \in \mathcal{C}_2$ alors

$$C_1 \times C_2 = (C_1 \times S_2) \cap (S_1 \times C_2) = p_1^{-1}(C_1) \cap p_2^{-1}(C_2) \in \mathcal{F}_1 \otimes \mathcal{F}_2. \tag{14.352}$$

□

PROPooSKIEooCpLVQT

Proposition 14.129 ([1, 453])**.**
Soient des espaces mesurables[44] $(\omega, \mathcal{F})$, $(\Omega_i, \mathcal{A}_i)$ $(i = 1, 2)$*. Nous considérons des applications mesurables* $f_i \colon \Omega \to \Omega_i$*. Alors l'application* $(f_1, f_2) \colon \Omega \to \Omega_1 \times \Omega_2$ *est mesurable pour la tribu produit*[45]*.*

Démonstration. Nous considérons l'ensemble

$$\mathcal{D} = \{B \subset \Omega_1 \times \Omega_2 \text{ tel que } (f_1, f_2)^{-1}(B) \in \mathcal{F}\}, \tag{14.353}$$

et nous allons prouver que $\mathcal{D}$ est une tribu contenant les parties de la forme $A_1 \times A_2$.

(1) **$\mathcal{D}$ est une tribu** Il faut vérifier la définition 14.1. D'abord $\Omega_1 \times \Omega_2 \in \mathcal{D}$ parce que

$$\{(f_1, f_2) \in \Omega_1 \times \Omega_2\} = \Omega \in \mathcal{F}. \tag{14.354}$$

Ensuite si $B \in \mathcal{D}$, nous avons

$$\begin{aligned}
(f_1, f_2)^{-1}(B^c) &= \{\omega \in \Omega \text{ tel que } (f_1, f_2)(\omega) \in B^c\} && \text{(14.355a)}\\
&= \{\omega \in \Omega \text{ tel que } (f_1, f_2)(\omega) \in B\}^c && \text{(14.355b)}\\
&\in \mathcal{F}^c && \text{(14.355c)}\\
&= \mathcal{F}. && \text{(14.355d)}
\end{aligned}$$

Nous avons utilisé le fait que $\{\omega \in \Omega \text{ tel que } (f_1, f_2)(\omega) \in B\} \in \mathcal{F}$. Enfin si $(B_i)_{i \in \mathbb{N}}$ sont des éléments de $\mathcal{D}$, nous avons

$$(f_1, f_2)^{-1}\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \bigcup_{i=1}^{\infty} (f_1, f_2)^{-1}(B_i) \in \bigcup_{i \in \mathbb{N}} \mathcal{F} = \mathcal{F}. \tag{14.356}$$

44. Définition 14.1.
45. Définition 14.127.

(2) $A_1 \times A_2 \in \mathcal{D}$ Soient $A_1 \in \mathcal{A}_1$ et $A_2 \in \mathcal{A}_2$. Nous prouvons que $A_1 \times A_2 \in \mathcal{D}$. Pour cela,

$$\begin{aligned}(f_1, f_2)^{-1}(A_1 \times A_2) &= \{\omega \in \Omega \text{ tel que } f_1(\omega) \in A_1, f_2(\omega) \in A_2\} && (14.357a)\\ &= \{\omega \in \Omega \text{ tel que } f_1(\omega) \in A_1\} \cap \{\omega \in \Omega \text{ tel que } f_2(\omega) \in A_2\} && (14.357b)\\ &= f_1^{-1}(A_1) \cap f_2^{-1}(A_2) && (14.357c)\\ &\in \mathcal{F} \cap \mathcal{F} && (14.357d)\\ &\subset \mathcal{F}. && (14.357e)\end{aligned}$$

(3) **Conclusion** La partie $\mathcal{D}$ est une tribu contenant les ensembles de la forme $A_1 \times A_2$. La définition 14.127 de la tribu produit dit alors que $\mathcal{A}_1 \otimes \mathcal{A}_2 \subset \mathcal{D}$.

□

14.7.2 Le cas des boréliens

Si X_1 et X_2 sont des espaces topologiques et si nous notons $\mathcal{O}_i$ l'ensemble de leurs ouverts, par définition $\mathcal{B}\text{or}(X_i) = \sigma(\mathcal{O}_i)$. De plus par la proposition 14.128 nous savons que

$$\sigma(\mathcal{O}_1 \times \mathcal{O}_2) = \mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2). \quad (14.358)$$

EqOHMSooRSLrDk

LemDEDQooJyzXgC

Lemme 14.130.
Si $(X_i, \mathcal{O}_i)$ sont des espaces topologiques, alors

$$\mathcal{B}or(X_1) \otimes \mathcal{B}or(X_2) \subset \mathcal{B}or(X_1 \times X_2) \quad (14.359)$$

Démonstration. Si $A_i \in \mathcal{O}_i$ alors $A_1 \times A_2$ est un ouvert de $X_1 \times X_2$ (voir la définition 7.15). Par conséquent, $\mathcal{O}_1 \times \mathcal{O}_2$ est contenu dans l'ensemble des ouverts de $X_1 \times X_2$ ou encore

$$\mathcal{O}_1 \times \mathcal{O}_2 \subset \mathcal{B}\text{or}(X_1 \times X_2), \quad (14.360)$$

et donc

$$\sigma(\mathcal{O}_1 \times \mathcal{O}_2) \subset \sigma\big(\mathcal{B}\text{or}(X_1 \times X_2)\big) \quad (14.361)$$

finalement, par (14.358)

$$\mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2) \subset \mathcal{B}\text{or}(X_1 \times X_2). \quad (14.362)$$

□

Il n'y a en général pas égalité, mais nous allons immédiatement voir que dans (presque) tous les cas raisonnables, les boréliens sur un produit sont le produit des boréliens.

PropNAAJooBPbjkX

Proposition 14.131 ([452]).
Soient (X_1, d_1) et (X_2, d_2) des espaces métriques séparables. Alors

$$\mathcal{B}or(X_1 \times X_2) = \mathcal{B}or(X_1) \otimes \mathcal{B}or(X_2). \quad (14.363)$$

Démonstration. Nous savons par le lemme 7.241 que tout ouvert de $X_1 \times X_2$ est une réunion dénombrable d'éléments de $\mathcal{O}_1 \times \mathcal{O}_2$. Donc tout ouvert de $X_1 \times X_2$ est dans $\mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2)$. Par conséquent

$$\mathcal{B}\text{or}(X_1 \times X_2) \subset \mathcal{B}\text{or}(X_1) \otimes \mathcal{B}\text{or}(X_2). \quad (14.364)$$

L'inclusion inverse étant déjà acquise par le lemme 14.130, nous avons l'égalité. □

CorWOOOooHcoEEF

Proposition 14.132.
Les boréliens sur $\mathbb{R}^N$ sont ceux qu'on croit.

(1) $\mathcal{B}or(\mathbb{R}^2) = \mathcal{B}or(\mathbb{R}) \otimes \mathcal{B}or(\mathbb{R})$

(2) $\mathcal{B}or(\mathbb{R}^{N+1}) = \mathcal{B}or(\mathbb{R}^N) \otimes \mathcal{B}or(\mathbb{R})$

Démonstration. Cela n'est rien d'autre que la proposition 14.131. □

Proposition 14.133.
Soit un espace mesurable $(S, \mathcal{F})$ et des applications $f_k : S \to \mathbb{R}$ $(k = 1, \dots, N)$. Alors l'application

$$\begin{aligned} f : (S, \mathcal{F}) &\to (\mathbb{R}^N, \mathcal{B}or(\mathbb{R}^N)) \\ x &\mapsto \big(f_1(x), \dots, f_N(x)\big) \end{aligned} \tag{14.365}$$

est mesurable si et seulement si chacun des f_i est mesurable.

Démonstration. Division en deux.

(1) **Condition nécessaire** Nous supposons que les f_i sont mesurables. Nous avons

$$f^{-1}\Big(\prod_{k=1}^{N}]a_k, b_k[\Big) = \{x \in S \text{ tel que } f_1(x) \in]a_1, b_1[, \cdots f_N(x) \in]a_N, b_N[\} \tag{14.366a}$$

$$= \bigcap_{k=1}^{N} f_k^{-1}\big(]a_k, b_k[\big). \tag{14.366b}$$

Cela est une intersection finie d'éléments de $\mathcal{F}$ et est donc un élément de $\mathcal{F}$. Mais les pavés ouverts engendrent $\mathcal{B}\text{or}(\mathbb{R}^N)$ parce qu'ils sont une base dénombrable de la topologie (proposition 14.51). Le théorème 14.49 nous assure alors que f est mesurable parce que l'image inverse d'une base de la tribu est mesurable.

(2) **Condition suffisante** Si f est mesurable alors en particulier

$$f_k^{-1}\big(]a, b[\big) = f^{-1}\big(\mathbb{R} \times \dots \times]a, b[\times \mathbb{R} \times \dots \times \mathbb{R}\big) \in \mathcal{F}. \tag{14.367}$$

Pour cela nous avons utilisé la proposition 14.132 qui nous indique que le produit dans la parenthèse est un borélien de $\mathbb{R}^N$ en tant que produit de boréliens de $\mathbb{R}$.
Encore une fois f_k^{-1} tombe dans $\mathcal{F}$ pour une base dénombrable de la topologie de $\mathbb{R}$ et est donc mesurable.

□

14.8 Mesure de Lebesgue sur $\mathbb{R}$

SecZTFooXlkwk

Nous notons $\mathcal{S}$ l'ensemble des intervalles[46] de $\mathbb{R}$.

Proposition 14.134.
L'ensemble des réunions finies d'éléments de $\mathcal{S}$ est une algèbre de parties de $\mathbb{R}$ que nous allons noter $\mathcal{A}_{\mathcal{S}}$.

Démonstration. Nous devons vérifier la définition 14.13. Les ensembles $\mathbb{R}$ et $\varnothing$ sont des intervalles et font donc partie de $\mathcal{A}_{\mathcal{S}}$.

Si $A \in \mathcal{A}_{\mathcal{S}}$ se décompose en union d'intervalles de la forme (a_k, b_k) avec $k = 1, \dots, n$ (ici nous mettons des parenthèses au lieu de crochets parce qu'à priori nous ne savons pas). Alors

$$A^c = \bigcup_{k=0}^{n} (b_k, a_{k+1}) \tag{14.368}$$

où nous avons posé $b_0 = -\infty$ et $a_{n+1} = +\infty$. Ici encore les parenthèses sont soit fermées soit ouvertes en fonction de ce qu'étaient celles dans la décomposition de A. Quoi qu'il en soit, cette décomposition de A^c montre que $A^c \in \mathcal{A}_{\mathcal{S}}$.

Enfin si $A, B \in \mathcal{A}_{\mathcal{S}}$ alors $A \cup B \in \mathcal{A}_{\mathcal{S}}$. □

46. Définition 1.21.

Lemme 14.135.
Tout élément de $\mathcal{A}_{\mathcal{S}}$ admet une décomposition minimale unique en réunion finie d'intervalles. Cette décomposition est formée d'intervalles deux à deux disjoints.

Démonstration. Nous allons montrer que si $A \in \mathcal{A}_{\mathcal{S}}$, alors la décomposition minimale consiste en les composantes connexes[47] de A. Pour cela nous rappelons que la proposition 9.53 dit qu'une partie de $\mathbb{R}$ est connexe si et seulement si elle est un intervalle. D'abord cela nous dit immédiatement que les composantes connexes de A forment une décomposition de A en intervalles. Nous devons prouver qu'elle est minimale.

Soit $\{C_k\}_{k=1,\dots,n}$ les composantes connexes de A. Aucun connexe de $\mathbb{R}$ contenu dans A ne peut intersecter plus d'un des C_k, et par conséquent nous ne pouvons pas décomposer A en moins de n intervalles.

Pour l'unicité, soit $\{I_k\}_{k=1,\dots,n}$ un ensemble de n intervalles tels que $\bigcup_{k=1}^{n} I_k = A$. Chacun des I_k intersecte un et un seul des C_k. En effet si $x \in I_k \cap C_i$ et $y \in I_k \cap C_j$, alors $[x,y] \subset I_k$ parce que I_k est un intervalle. Mais C_i étant le plus grand connexe contenant x, $[x,y] \subset C_i$ et de la même façon, $[x,y] \subset C_j$. Par conséquent C_i et C_j sont tous deux la composante connexe de x et y. Nous en déduisons que $C_i = C_j$, c'est-à-dire $i = j$.

Par ailleurs nous avons $I_k \cap I_l = \varnothing$ dès que $k \neq l$ parce que sinon l'ensemble $I_k \cup I_l$ serait connexe et la décomposition des $\{I_k\}_{k=1,\dots,n}$ ne serait pas minimale : en remplaçant I_k et I_l par $I_k \cup I_l$ on aurait eu une décomposition contenant moins d'éléments. Donc à renumérotation près nous pouvons supposer que I_k intersecte C_l si et seulement si $k = l$.

Dans ce cas nous devons avoir $I_k = C_k$, sinon les éléments de $C_k \backslash I_k$ ne seraient pas dans $\bigcup_{i=1}^{n} I_i$. □

Définition 14.136 (longueur d'intervalle[436])**.**
Si I est un intervalle d'extrémités a et b avec $-\infty \leqslant a \leqslant b \leqslant +\infty$ alors nous définissons la ***longueur*** *de I par*

$$\ell(I) = \begin{cases} b-a & \text{si } -\infty < a \leqslant b < +\infty \\ \infty & \text{si } a \text{ ou } b \text{ est infini} \end{cases} \tag{14.369}$$

Si $A \in \mathcal{A}_{\mathcal{S}}$ et si sa décomposition minimale est $A = \bigcup_{k=1}^{n} I_k$, alors on définit

$$\ell(A) = \sum_{k=1}^{n} \ell(I_k). \tag{14.370}$$

Le lemme suivant nous indique que nous pouvons calculer la longueur d'un élément de $\mathcal{A}_{\mathcal{S}}$ sans savoir la décomposition minimale, pourvu que l'on connaisse une décomposition disjointe.

LemIUQooEzHun

Lemme 14.137 ([436])**.**
Si

$$B = \bigcup_{r=1}^{p} J_r \tag{14.371}$$

est une décomposition de $B \in \mathcal{A}_{\mathcal{S}}$ en intervalles deux à deux disjoints alors

$$\ell(B) = \sum_{r=1}^{p} \ell(J_r). \tag{14.372}$$

Démonstration. Nous prouvons dans un premier temps le résultat dans le cas où $B = I$ est un intervalle. Soit I un intervalle et une décomposition en intervalles disjoints $I = \bigcup_{r=1}^{p} J_r$. Nous montrons qu'alors $\ell(I) = \sum_{r=1}^{p} \ell(J_r)$. Nous verrons ensuite comment passer au cas où B est un élément générique de $\mathcal{A}_{\mathcal{S}}$.

(1) **Si $B = I$ est un intervalle infini**
Si I est infini alors un des J_r soit l'être et donc $\sum_{r=1}^{p} \ell(J_r) = \infty = \ell(I)$.

47. Définition 7.73.

(2) **Si $B = I$ est un intervalle ininfini**

Pour chaque $r = 1, \dots, p$ nous notons a_r et b_r les extrémités de J_r. Vu que les J_r sont connexes et disjoints, si $a_k \leqslant a_l$ alors $b_k \leqslant a_l$, sinon l'ensemble (non vide) $]a_l, b_k[$ serait dans l'intersection $I_k \cap I_l$ qui, elle, est vide. Plus généralement, si $x \in J_k$ et $y \in J_l$ avec $x < y$ alors pour tout $x' \in J_k$ et tout $y' \in J_l$ nous avons $x' < y'$. Vu qu'il y a un nombre fini d'ensembles J_r, nous pouvons les classer dans l'ordre croissant :

$$a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \dots \leqslant b_{p-1} \leqslant a_p \leqslant b_p. \tag{14.373}$$

Vu que les J_r sont disjoints et que leur union est connexe nous avons en réalité

$$a = a_1 \leqslant b_1 = a_2 \leqslant b_2 = a_3 \leqslant \dots \leqslant b_{p-1} = a_p \leqslant b_p, \tag{14.374}$$

donc une somme télescopique donne

$$\ell(I) = b - a = \sum_{r=1}^{p} (b_r - a_r) = \sum_{r=1}^{p} \ell(J_r). \tag{14.375}$$

(3) **Si B n'est pas un intervalle** Soit $\{I_k\}_{k=1,\dots,n}$ la décomposition minimale de B. Alors

$$\spadesuit = \ell(B) = \sum_{k=1}^{n} \ell(I_k) = \sum_{k=1}^{n} \ell\Big(\bigcup_{r=1}^{p} (I_k \cap J_r) \Big). \tag{14.376}$$

Mais I_k est un intervalle et s'écrit comme union disjointe $I_k = \bigcup_{r=1}^{p} (I_k \cap J_r)$, donc par la première partie

$$\spadesuit = \sum_{k=1}^{n} \sum_{r=1}^{p} \ell(I_k \cap J_r) = \sum_{r=1}^{p} \sum_{k=1}^{n} \ell(I_k \cap J_r). \tag{14.377}$$

Ici J_r est un intervalle qui se décompose en $J_r = \bigcup_{k=1}^{n} (I_k \cap J_r)$, donc nous pouvons encore utiliser la première partie :

$$\spadesuit = \sum_{r=1}^{p} \ell(J_r), \tag{14.378}$$

ce qu'il fallait.

□

LemPIOooRLkbo

Lemme 14.138.
Si $A, B \in \mathcal{A}_{\mathcal{S}}$ avec $A \subset B$ alors $\ell(A) \leqslant \ell(B)$.

Démonstration. Nous avons évidemment $B = A \cup B\backslash A$. Notons que $B\backslash A \in \mathcal{A}_{\mathcal{S}}$ par le lemme 14.15. Si $\{I_k\}$ est une décomposition disjointe de A et $\{J_i\}$ une de $B\backslash A$ alors $\{I_k\} \cup \{J_i\}$ est une décomposition disjointe de $A \cup B\backslash A$ et le lemme 14.137 nous dit que

$$\ell(B) = \ell(A \cup B\backslash A) = \ell(A) + \ell(B\backslash A). \tag{14.379}$$

Par conséquent $\ell(B) \geqslant \ell(A)$. □

LemUMVooZJgMu

Lemme 14.139.
Si I est un intervalle et si il se décompose en

$$I = \bigcup_{n \in \mathbb{N}} I_n \tag{14.380}$$

où les I_n sont des intervalles disjoints, alors

$$\ell(I) = \sum_{n=1}^{\infty} \ell(I_n). \tag{14.381}$$

Démonstration. Nous allons encore diviser la preuve en deux parties suivant que I soit de longueur finie ou pas.

(1) **Si I est de longueur finie**

Soient a et b les extrémités de I : $-\infty < a \leqslant b < +\infty$. Pour tout $N \geqslant 1$ nous avons

$$\sum_{n=1}^{N} \ell(I_n) = \ell\Big(\bigcup_{n=1}^{N} I_n\Big) \leqslant \ell(I). \tag{14.382}$$

La première égalité est le lemme dans le cas d'une union finie 14.137. L'inégalité est le lemme 14.138. Cela étant vrai pour tout N, à la limite $N \to \infty$ nous conservons l'inégalité :

$$\sum_{n=1}^{\infty} \ell(I_n) \leqslant \ell(I). \tag{14.383}$$

Nous devons encore voir l'inégalité inverse. Pour cela nous supposons que $a < b$. Sinon $\ell(I) = 0$ et tous les I_n doivent être vide sauf un qui contiendra seulement $\{a\}$ (si I le contient).

Soit $\epsilon > 0$ avec $\epsilon < b - a$ et l'intervalle

$$[a + \frac{\epsilon}{4}, b - \frac{\epsilon}{4}] = [a', b'] \subset I. \tag{14.384}$$

Si les a_n et le b_n sont le extrémités des I_n alors

$$[a', b'] \subset I = \bigcup_{n\geqslant 1} I_n \subset \bigcup_{n\geqslant 1}]a_n - \frac{\epsilon}{2^{n+2}}, b_n + \frac{\epsilon}{2^{n+2}}[= \bigcup_{n\geqslant 1}]a'_n, b'_n[\tag{14.385}$$

où nous avons posé $a'_n = a_n - \epsilon/2^{n+2}$ et $b'_n = b_n + \epsilon/2^{n+2}$. Nous avons donc recouvert le compact [48] $[a', b']$ par des ouverts. Nous pouvons donc en extraire un sous-recouvrement fini (c'est la définition de la compacité), c'est-à-dire une partie finie F de $\mathbb{N}$ telle que

$$[a', b'] \subset \bigcup_{n\in F}]a'_n, b'_n[. \tag{14.386}$$

Le lemme 14.138 nous dit alors que

$$\heartsuit = b' - a' \leqslant \ell\Big(\bigcup_{n\in F}]a'_n, b'_n[\Big) \leqslant \sum_{n\in F} (b'_n - a'_n). \tag{14.387}$$

La seconde inégalité se prouve en recopiant [49] la preuve de 14.16. Nous continuons le calcul :

$$\heartsuit \leqslant \sum_{n\in F} (b_n - a_n) + \sum_{n\in F} \frac{\epsilon}{2^{n+1}} \leqslant \sum_{n\in F} (b_n - a_n) + \frac{\epsilon}{2}. \tag{14.388}$$

Mais $b' - a' = (b - a) - \frac{\epsilon}{2}$, donc

$$b - a - \frac{\epsilon}{2} \leqslant \sum_{n\in F} (b_n - a_n) + \frac{\epsilon}{2}. \tag{14.389}$$

D'où nous déduisons que

$$\ell(I) = b - a \leqslant \sum_{n\in F} (b_n - a_n) + \epsilon \leqslant \sum_{n\in \mathbb{N}} (b_n - a_n) + \epsilon = \sum_{n\in\mathbb{N}} \ell(I_n) + \epsilon. \tag{14.390}$$

Cela étant valable pour tout ϵ nous déduisons que

$$\ell(I) \leqslant \sum_{n\in\mathbb{N}} \ell(I_n). \tag{14.391}$$

48. Lemme 9.20.

49. Nous ne pouvons pas invoquer directement le lemme 14.16 parce que nous n'avons pas encore prouvé que ℓ était une mesure sur $(\mathbb{R}, \mathcal{A}_\mathcal{S})$.

(2) **Si I est de longueur infinie**

Étant donné que I est un intervalle de longueur infinie, il doit au moins contenir un ensemble du type $]-\infty,a]$ ou $[a,+\infty[$; donc pour tout $M>0$, il existe $N\geqslant 1$ tel que

$$\ell\big(I\cap[-N,N]\big)\geqslant M. \tag{14.392}$$

Mais $I\cap[-N,N]$ est un intervalle et

$$I\cap[-N,N]=\bigcup_{n\in\mathbb{N}} I_n\cap[-N,N] \tag{14.393}$$

qui est une union disjointe. Par conséquent,

$$M\leqslant\ell\big(I\cap[-N,N]\big)=\sum_n\ell\big(I_n\cap[-N,N]\big)\leqslant\sum_n\ell(I_n). \tag{14.394}$$

Cela étant vrai pour tout $M>0$, nous concluons que

$$\sum_{n\in\mathbb{N}}\ell(I_n)=\infty. \tag{14.395}$$

□

Remarque 14.140.
Pour la preuve de 14.139 nous ne pouvons pas classer les I_n en ordre croissant comme nous l'avons fait dans la preuve de 14.137. En effet si $I=[0,1]$ et que nous recouvrons $[0,\frac{1}{2}[$ et $]\frac{1}{2},1]$ par une infinité d'intervalles chacun, nous ne pouvons plus les classer par ordre croissant.

PropULFoodgXrR

Proposition 14.141 ([436])**.**
La fonction ℓ ainsi définie est une mesure σ-finie sur l'algèbre de parties $\mathcal{A}_{\mathcal{S}}$.

Démonstration. Le fait que ℓ soit σ-finie provient par exemple du fait que $\ell\big(]-n,n[\big)=2n$ tandis que $\bigcup_n]-n,n[=\mathbb{R}$.

Nous devons à présent prouver que ℓ est additive. Soient $(A_i)_{i\in\mathbb{N}}$ des éléments disjoints de $\mathcal{A}_{\mathcal{S}}$, avec leurs décomposition minimales

$$A_i=\bigcup_{k=1}^{n} I_k^{(i)}. \tag{14.396}$$

Pour chaque $i\in\mathbb{N}$, le lemme 14.139 nous indique que

$$\ell(A_i)=\sum_{k\in\mathbb{N}}\ell(I_k^{(i)}). \tag{14.397}$$

L'ensemble $\mathbb{N}\times\mathbb{N}$ est dénombrable et nous pouvons considérer la décomposition

$$\bigcup_{i\in\mathbb{N}}A_i=\bigcup_{(i,k)\in\mathbb{N}\times\mathbb{N}} I_k^{(i)}. \tag{14.398}$$

Cette décomposition n'est pas spécialement minimale [50] mais elle est disjointe. Le lemme 14.139 donne

$$\ell\Big(\bigcup_i A_i\Big)=\sum_{(i,k)\in\mathbb{N}\times\mathbb{N}}\ell(I_k^{(i)})=\sum_{i\in\mathbb{N}}\left(\sum_{k\in\mathbb{N}}\ell(I_k^{(i)})\right)=\sum_{i\in\mathbb{N}}\ell(A_i). \tag{14.399}$$

La décomposition de la somme sur $\mathbb{N}^2$ en deux sommes sur $\mathbb{N}$ est faite en vertu de la proposition 10.125. □

50. A_1 pourrait contenir $[0,1]$ et A_2 contenir $]1,2]$.

14.8.1 Mesure et tribu de Lebesgue

ThoDESooEyDOe

Théorème 14.142.
Il existe une unique mesure λ sur $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R})\big)$ telle que

$$\lambda\big(]a,b[\big) = b - a \tag{14.400}$$

pour tout $a \leqslant b$ dans $\mathbb{R}$.

Démonstration. L'existence provient du théorème de prolongement de Hahn 14.80 : la mesure ℓ sur $(\mathcal{A}_{\mathcal{S}})$ se prolonge à $\sigma(\mathcal{A}_{\mathcal{S}}) = \mathcal{B}\text{or}(\mathbb{R})$.

Nous ne pouvons pas prouver l'unicité en invoquant la partie unicité de Hahn (c'est tentant parce que ℓ est σ-finie) parce que dans ce théorème nous ne fixons la valeur de λ que sur une toute petite partie de $\mathcal{A}_{\mathcal{S}}$. Nous allons cependant voir que cette petite partie suffit à garantir l'unicité.

La classe

$$\mathcal{D} = \{]a,b[\text{ tel que } -\infty < a \leqslant b < +\infty\} \tag{14.401}$$

est stable par intersection finie et engendre la tribu borélienne. En effet $\mathcal{D}$ contient toutes les boules et donc une base dénombrable de la topologie de $\mathbb{R}$ (proposition 7.138). Donc tous les ouverts de $\mathbb{R}$ sont dans $\sigma(\mathcal{D})$ et $\sigma(\mathcal{D}) = \mathcal{B}\text{or}(\mathbb{R})$. Nous pouvons donc dire grâce au théorème 14.34 qu'il y a unicité de la mesure sur $\mathcal{B}\text{or}(\mathbb{R})$ lorsque les valeurs sur $\mathcal{D}$ sont fixées. □

DefooYZSQooSOcyYN

Définition 14.143.
*La mesure de l'espace mesuré $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda\big)$ donné par le théorème 14.142 est la **mesure de Lebesgue** sur $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R})\big)$.*

*Nous définissons aussi la **tribu de Lebesgue** par la proposition 14.77 : $\big(\mathbb{R}, \mathcal{L}eb(\mathbb{R}), \lambda\big)$ est l'espace mesuré complété de $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \lambda\big)$.*

Remarque 14.144.
Il n'est pas évident que la tribu de Lebesgue soit plus grande que celle des boréliens, ni que la tribu des parties soit plus grande que celle de Lebesgue. Nous mentionnons cependant les faits suivants.

(1) Il existe des ensembles mesurables non-boréliens, et cela ne nécessite pas l'axiome du choix. Un argument classique de cardinalité est donné dans [434]. La construction la plus explicite que j'aie trouvée est dans [454], mais ça a l'air de demander des connaissances précises sur les ordinaux.

(2) Vu que l'ensemble de Cantor C est mesurable de mesure nulle (proposition 14.163), tout sous-ensemble de Cantor est mesurable de mesure nulle parce que la tribu de Lebesgue est complète par définition. Le cardinal de $\mathcal{P}(C)$ est strictement supérieur à la puissance du continu, alors que le cardinal de l'ensemble des boréliens est au plus égal à la puissance du continu. Donc il existe des non boréliens contenus dans Cantor ; de tels non boréliens sont alors mesurables au sens de Lebesgue.

(3) Si nous admettons l'axiome du choix alors il existe des ensembles non mesurables au sens de Lebesgue. Nous en verrons un dans l'exemple 14.157.

Exemple 14.145 (Un ouvert contenant tous les rationnels et de mesure arbitrairement petite)**.**
Il est possible de construire un ouvert de $\mathbb{R}$ contenant $\mathbb{Q}$ et de mesure de Lebesgue plus petite que ϵ. Pour cela si (q_i) est une énumération des rationnels, il suffit de prendre

$$\mathcal{O} = \bigcup_{n=1}^{\infty} B(q_n, \frac{\epsilon}{2^{n+1}}). \tag{14.402}$$

Cela est un ouvert comme union d'ouverts, ça contient tous les rationnels, et sa mesure se majore. En effet le théorème 14.142 donne $\lambda\big(B(q_n, \frac{\epsilon}{2^n})\big) = \frac{\epsilon}{2^n}$. Vu que ces boules ne sont à priori pas disjointes, le lemme 14.23 donne

$$\lambda(\mathcal{O}) \leqslant \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon \tag{14.403}$$

par (10.345) avec $q = \frac{1}{2}$.

Par complémentarité, nous pouvons construire un ensemble fermé de mesure non nulle et ne contenant aucun rationnel. Et même un fermé dans $[0,1]$, de mesure $1-\epsilon$ ne contenant aucun rationnel.

Cela peut surprendre parce qu'il existe des tonnes de suites d'irrationnels qui convergent vers des rationnels [51], et il semble difficile de créer un ensemble contenant beaucoup d'irrationnels tout en préservant la propriété de fermeture vis à vis des suites convergentes. △

Exemple 14.146 (Mesure finie, non borné)**.**
Il existe des parties de $\mathbb{R}$ qui sont de mesure finie sans être bornés. Par exemple en posant

$$A = \bigcup_{n=1}^{\infty} B(n, \frac{1}{2^n}). \tag{14.404}$$

La partie A n'est pas bornée parce que que $\mathbb{N} \subset A$. Mais en termes de mesure,

$$\lambda(A) \leqslant \sum_{n=1}^{\infty} \frac{1}{2^n} < \infty \tag{14.405}$$

en vertu de la somme de la série géométrique, proposition 10.135. △

14.8.2 Propriétés de la mesure de Lebesgue

Proposition 14.147.
Tout ensemble dénombrable de $\mathbb{R}$ est mesurable de mesure nulle.

Démonstration. Un point de $\mathbb{R}$ est un intervalle de mesure nulle. Si D est dénombrable, il est union disjointes et dénombrable de points. Le lemme 14.139 nous dit alors que sa mesure est $\lambda(D) = \sum_{i=1}^{\infty} \lambda(\{a_i\}) = 0$. □

Remarque 14.148.
Il existe cependant des ensembles non dénombrables et tout de même de mesure nulle. Par exemple l'ensemble de Cantor (voir la proposition 14.163).

PropooOACLooLMIUuY

Proposition 14.149.
La mesure de Lebesgue est invariante par translation, c'est-à-dire que si A est mesurable alors $\lambda(A) = \lambda(A+\alpha)$ pour tout réel α.

Démonstration. Nous commençons par les intervalles ouverts :

$$\lambda\big(]a,b[+ \alpha\big) = \lambda\big(]a+\alpha, b+\alpha[\big) = (b+\alpha) - (a+\alpha) = b - a = \lambda\big(]a,b[\big). \tag{14.406}$$

D'après ce qui est dit dans l'exemple 14.36, la mesure de Lebesgue sur les boréliens est invariante par translation.

Si A est mesurable alors il existe un borélien B et un ensemble négligeable N tels que $A = B \cup N$ par la caractérisation 14.132b de la complétion. Alors $A + \alpha = B + \alpha \cup N + \alpha$ et $N + \alpha$ est encore un ensemble négligeable. Donc $\lambda(A+\alpha) = \alpha(B+\alpha) = \lambda(B)$. □

La mesure ℓ définie sur l'algèbre de parties $\mathcal{A}_{\mathcal{S}}$ (voir proposition 14.141). La proposition 14.17 nous donne donc une mesure extérieure par

$$\lambda^*(X) = \inf\{\sum_n \ell(A_n); A_n \in \mathcal{A}_{\mathcal{S}}, X \subset \bigcup_n A_n\}. \tag{14.407}$$

EqJGXoogdKqb

La proposition suivante montre que cette mesure extérieure peut être exprimée seulement avec des intervalles ouverts.

51. Si $q \in \mathbb{Q}$ et $r \in \mathbb{R}\backslash\mathbb{Q}$ alors la suite $(q + r/10^k)_k$ est une suite d'irrationnels convergente vers le rationnel q.

PropTNOooDcfwn

Proposition 14.150.
Nous avons

$$\lambda^*(X) = \inf\{\sum_{n\geqslant 1} \ell(I_n); I_n \text{ sont des intervalles ouverts et } X \subset \bigcup_n I_n\}. \tag{14.408}$$

Démonstration. Nous savons que dans la définition (14.407), chacun des A_n est une réunion disjointe d'intervalles (pas spécialement ouverts) deux à deux disjoints ; donc

$$\lambda^*(X) = \inf\{\sum_n \ell(I_n); I_n \in \mathcal{S}, X \subset \bigcup_n I_n\}. \tag{14.409}$$

Soit $\epsilon > 0$. Si $A \subset \bigcup_n I_n$, pour chaque $n \geqslant 1$ nous considérons un intervalle ouvert J_n tel que $I_n \subset J_n$ et $\ell(I_n) + \frac{\epsilon}{2^n} \leqslant \ell(J_n)$. Faisant cela pour chacun des découpages de X en intervalles nous trouvons

$$\lambda^*(X) \leqslant \inf\{\sum_n \ell(J_n) \ J_n \text{ est ouvert et } X \subset \bigcup_n J_n\} + \epsilon. \tag{14.410}$$

Étant donné que ϵ est arbitraire nous avons l'égalité. □

PropMXIoojpKvd

Proposition 14.151 ([436])**.**
Si $X \subset \mathbb{R}$ est tel que $\lambda^*(X) < \infty$ alors

ItemGJUoozrDILi

(1) Pour tout $\epsilon > 0$ il existe un ouvert Ω_ϵ tel que

$$\begin{cases} X \subset \Omega_\epsilon & \text{(14.411a)} \\ \lambda(\Omega_\epsilon) \leqslant \lambda^*(X) + \epsilon. & \text{(14.411b)} \end{cases}$$

ItemGJUoozrDILii

(2) Il existe une intersection dénombrable d'ouverts G telle que

$$\begin{cases} X \subset G & \text{(14.412a)} \\ \lambda(G) = \lambda^*(X). & \text{(14.412b)} \end{cases}$$

Démonstration. Pour (1), la proposition 14.150 nous a déjà dit que

$$\lambda^*(X) = \inf\{\sum_n \ell(I_n) \ I_n \text{ est un intervalle ouvert}, X \subset \bigcup_n I_n\}, \tag{14.413}$$

donc si $\epsilon > 0$, il existe des intervalles ouverts I_n tels que

$$\begin{cases} X \subset \bigcup_n I_n & \text{(14.414a)} \\ \sum_n \ell(I_n) \leqslant \lambda^*(X) + \epsilon. & \text{(14.414b)} \end{cases}$$

Si nous posons $\Omega_\epsilon = \bigcup_n I_n$, alors nous avons bien

$$\begin{cases} X \subset \Omega_\epsilon & \text{(14.415a)} \\ \lambda(\Omega_\epsilon) \leqslant \sum_n \ell(I_n) \leqslant \lambda^*(X) + \epsilon. & \text{(14.415b)} \end{cases}$$

En ce qui concerne (2), pour chaque $k \geqslant 1$ nous considérons l'ensemble $\Omega_{1/k}$ obtenu comme précédemment avec $\epsilon = 1/k$ et nous posons $G = \bigcap_{k\geqslant 1} \Omega_{1/k}$. Cela est une intersection dénombrable d'ouverts vérifiant $X \subset G$ (parce que $X \subset \Omega_{1/k}$ pour tout k) et donc $\lambda^*(X) \leqslant \lambda^*(G) = \lambda(G)$. De plus pour tout k nous avons

$$\lambda(G) \leqslant (\Omega_{1/k}) \leqslant \lambda^*(X) + \frac{1}{k} \tag{14.416}$$

pour tout k. En faisant $k \to \infty$ nous avons

$$\lambda(G) \leqslant \lambda^*(X). \tag{14.417}$$

Au final

$$\lambda(G) \leqslant \lambda^*(X) \leqslant \lambda(G), \tag{14.418}$$

d'où l'égalité. □

Corolaire 14.152.
Une partie $N \subset \mathbb{R}$ est négligeable[52] *si et seulement si $\lambda^*(N) = 0$.*

Démonstration. Nous savons que si N est négligeable il existe un borélien Y tel que $N \subset Y$ avec $\lambda(Y) = 0$. Par conséquent [53]

$$\lambda^*(N) \leqslant \lambda^*(Y) = \lambda(Y) = 0. \tag{14.419}$$

Pour l'implication inverse nous supposons que $\lambda^*(N) = 0$ et nous prenons l'ensemble G définit par la proposition 14.151(2) : c'est un borélien contenant N et tel que $\lambda(G) = \lambda^*(N) = 0$. L'ensemble N est donc négligeable. □

ThoHFXooONFRN

Théorème 14.153 (Régularité extérieure de la mesure de Lebesgue)**.**
Pour tout mesurable $A \subset \mathbb{R}$ nous avons

$$\lambda(A) = \inf\{\lambda(\Omega); \Omega \text{ ouvert contenant } A\}. \tag{14.420}$$

Démonstration. Nous commençons par le cas où B est un borélien.

(1) **Si B borélien, $\lambda(B) < \infty$**
Soit $\epsilon > 0$; par la proposition 14.151(1) il existe un ouvert Ω_ϵ contenant B tel que $\lambda(\Omega_\epsilon) \leqslant \lambda^*(B) + \epsilon$. Vu qu'ici B est borélien, $\lambda^*(B) = \lambda(B)$ et nous concluons que pour tout ϵ il existe un ouvert Ω_ϵ tel que

$$\begin{cases} B \subset \Omega_\epsilon & \text{(14.421a)} \\ \lambda(\Omega_\epsilon) \leqslant \lambda(B) + \epsilon, & \text{(14.421b)} \end{cases}$$

et donc

$$\lambda(B) = \inf\{\lambda(\Omega);\ \Omega \text{ ouvert contenant } B\ \}. \tag{14.422}$$

(2) **Si B borélien, $\lambda(B) = +\infty$**
Dans ce cas l'infimum est pris uniquement sur des ouverts Ω tels que $\lambda(\Omega) = \infty$.

(3) **Si A est mesurable non borélien**
Nous passons maintenant au cas où A est mesurable sans être borélien. Il s'écrit donc $A = B \cup N$ avec B borélien et N négligeable par la proposition 14.69, et par définition $\lambda(A) = \lambda(B)$. Si Y est un borélien tel que $N \subset Y$ et $\lambda(Y) = 0$ alors

$$\begin{aligned}
\lambda(A) = \lambda(B) &= \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \subset \Omega\} && \text{(14.423a)} \\
&\leqslant \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \cup N \subset \Omega\} && \text{(14.423b)} \\
&\leqslant \inf_{\Omega', Y'}\{\lambda(\Omega' \cup Y') \text{ tel que } \Omega', Y' \text{ ouverts}, B \subset \Omega', Y \subset Y'\} && \text{(14.423c)} \\
&\leqslant \inf_{\Omega', Y'}\{\lambda(\Omega') + \lambda(Y') \text{ tel que } \Omega', Y' \text{ ouverts}, B \subset \Omega', Y \subset Y'\} && \text{(14.423d)} \\
&\leqslant \inf_{\Omega'}\{\lambda(\Omega') \text{ tel que } \Omega' \text{ ouvert}, B \subset \Omega\} && \text{(14.423e)} \\
&= \lambda(B). && \text{(14.423f)}
\end{aligned}$$

subeqMTHoopkSKOi subeqMTHoopkSKOii subeqMTHoopkSKOiii subeqMTHoopkSKOiv subeqMTHoopkSKOv

Justifications :

— (14.423a) Le cas borélien déjà fait.

— (14.423b) Les ouverts Ω tels que $B \cup N \subset \Omega$ vérifient a fortiori $B \subset \Omega$; nous avons donc agrandit l'ensemble sur lequel l'infimum est pris.

— (14.423c) Parmi les ouverts Ω qui recouvrent $B \cup N$, il y a ceux de la forme $\Omega' \cup Y'$ où Ω' recouvre B et Y' est un ouvert contenant Y. Donc nous avons rétréci l'ensemble sur lequel l'infimum est pris et par conséquent agrandit l'infimum.

52. Définition 14.66.
53. Au péril d'être lourd nous rappelons que λ^* est défini sur toutes les parties de $\mathbb{R}$.

— (14.423d) Mesure d'une union majorée par la somme des mesures.
— (14.423e) Vu que Y est borélien, $\lambda(Y) = \inf_{Y' \text{ ouvert}}\{\lambda(Y') \text{ tel que } Y \subset Y'\} = 0$. Donc pour tout Ω' et tout $\epsilon > 0$, nous pouvons trouver un Y' vérifiant les conditions tel que $\lambda(\Omega') + \lambda(Y') \leqslant \lambda(\Omega') + \epsilon$.

Toutes les inégalités sont des égalités en en particulier (14.423b) donne

$$\lambda(A) = \inf\{\lambda(\Omega) \text{ tel que } \Omega \text{ ouvert}, B \cup N \subset \Omega\}, \tag{14.424}$$

ce qu'il fallait.

□
PropEZNoofLkVb

Proposition 14.154 ([436]).
Si A est mesurable dans $\mathbb{R}$ et si $\epsilon > 0$ alors il existe un ouvert Ω_ϵ et un fermé F_ϵ tels que subeqHNEooaNqDu

$$\begin{cases} F_\epsilon \subset A \subset \Omega_\epsilon & (14.425a)\\ \lambda(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon. & (14.425b) \end{cases}$$

Démonstration. Nous commençons par le cas où A est un borélien, que nous noterons B.

(1) **Première étape**
Montrons qu'il existe un ouvert U_ϵ tel que

$$\begin{cases} B \subset U_\epsilon & (14.426a)\\ \lambda(U_\epsilon \backslash B) \leqslant \frac{\epsilon}{2}. & (14.426b) \end{cases}$$

Si $\lambda(B) < \infty$ alors le théorème 14.153 nous donne un ouvert U_ϵ tel que $B \subset U_\epsilon$ et $\lambda(U_\epsilon) \leqslant \lambda(B) + \frac{\epsilon}{2}$. Nous avons alors

$$\lambda(\Omega_\epsilon \backslash B) = \lambda(\Omega_\epsilon) - \lambda(B) \leqslant \frac{\epsilon}{2}. \tag{14.427}$$

Si par contre $\lambda(B) = \infty$, nous posons $B_n = B \cap [-n, n]$ et $\epsilon_n = \epsilon/2^{n+1}$. Pour chaque n nous avons un ouvert Ω_n tel que

$$\begin{cases} B_n \subset \Omega_n & (14.428a)\\ \lambda(\Omega_n \backslash B_n) \leqslant \frac{\epsilon}{2^{n+1}} & (14.428b) \end{cases}$$

Par conséquent en posant $\Omega = \bigcup_{n \geqslant 1} \Omega_n$ nous avons [54]

$$\begin{cases} B \subset \Omega & (14.429a)\\ \lambda(\Omega \backslash B) \leqslant \lambda\big(\bigcup_n (\Omega_n \backslash B_n)\big) \leqslant \sum_{n \geqslant 1} \lambda(\Omega_n \backslash B_n) = \frac{\epsilon}{2}. & (14.429b) \end{cases}$$

La première étape est terminée.

(2) **Deuxième étape**
Nous prouvons à présent qu'il existe un ouvert Ω_ϵ et un fermé F_ϵ tels que

$$\begin{cases} F_\epsilon \subset B \subset \Omega_\epsilon & (14.430a)\\ \lambda(\Omega_\epsilon \backslash B) \leqslant \frac{\epsilon}{2} & (14.430b)\\ \lambda(B \backslash F_\epsilon) \leqslant \frac{\epsilon}{2}. & (14.430c) \end{cases}$$

L'ouvert Ω_ϵ, nous l'avons déjà de l'étape précédente. Pour le fermé, nous appliquons la première étape au borélien B^c ; ce qui nous trouvons est un ouvert G_ϵ tel que

$$\begin{cases} B^c \subset G_\epsilon & (14.431a)\\ \lambda(G_\epsilon \backslash B^c) \leqslant \frac{\epsilon}{2}. & (14.431b) \end{cases}$$

En posant $F_\epsilon = G_\epsilon^c$ nous avons un fermé tel que $F_\epsilon \subset B$ et

$$\lambda(B \backslash F_\epsilon) = \lambda(F_\epsilon^c \backslash B^c) = \lambda(G_\epsilon \backslash B^c) \leqslant \frac{\epsilon}{2}. \tag{14.432}$$

54. Nous utilisons la petite relation ensembliste $\big(\bigcup_n A_n\big)\backslash\big(\bigcup_n B_n\big) \subset \bigcup_n (A_n \backslash B_n)$.

(3) **Dernière étape**
Les ensembles F_ϵ et Ω_ϵ trouvés à la deuxième étape donnent bien les relations (14.425). En effet $\Omega_\epsilon \backslash F_\epsilon = (\Omega_\epsilon \backslash B) \cup (B \backslash F_\epsilon)$, donc

$$\lambda(\Omega_\epsilon \backslash F_\epsilon) \leqslant \lambda(\Omega_\epsilon \backslash B) + \lambda(B \backslash F_\epsilon) = \epsilon. \tag{14.433}$$

Nous passons au cas où $A = B \cup N$ est mesurable. Nous commençons par prendre les Ω_ϵ et F_ϵ qui correspondent à B :

$$\begin{cases} F_\epsilon \subset B \subset \Omega_\epsilon & \text{(14.434a)} \\ \lambda(\Omega_\epsilon \backslash F_\epsilon) \leqslant \epsilon. & \text{(14.434b)} \end{cases}$$

Soit Y un borélien tel que $N \subset Y$ et $\lambda(Y) = 0$ puis un ouvert Y' tel que $\lambda(Y') \leqslant \epsilon$ et $Y \subset Y'$. L'existence d'un tel Y' est assurée par la proposition 14.153 appliquée à Y. Nous vérifions que les ensembles F_ϵ et $\Omega_\epsilon \cup Y'$ fonctionnent. En effet $\Omega_\epsilon \cup Y' \backslash F_\epsilon \subset (\Omega_\epsilon \backslash F_\epsilon) \cup Y'$, donc

$$\begin{cases} F_\epsilon \subset B \cup N \subset \Omega_\epsilon \cup Y' & \text{(14.435a)} \\ \lambda\big((\Omega_\epsilon \backslash F_\epsilon)\big) \leqslant \lambda(\Omega_\epsilon \backslash F_\epsilon) + \lambda(Y') \leqslant 2\epsilon. & \text{(14.435b)} \end{cases}$$

Donc en réalité il faut choisir $\Omega_{\epsilon/2}$, $F_{\epsilon/2}$ et $\lambda(Y') \leqslant \epsilon/2$. □

THOooJNMCooPMvCDq

Théorème 14.155 (Régularité intérieure de la mesure de Lebesgue)**.**
Si A est mesurable dans $\mathbb{R}$ alors

$$\lambda(A) = \sup\{\lambda(K); K \text{ compact contenu dans } A\}. \tag{14.436}$$

Démonstration. Par la proposition 14.154 nous avons

$$\lambda(A) = \sup_{F \text{ fermé dans } A} \lambda(F). \tag{14.437}$$

EqTPEooUHTbH

Pour un tel F nous posons $K_n = F \cap [-n, n]$ qui est compact [55] et contenu dans B. De plus le lemme 14.24(2) nous dit que

$$\lambda(F) = \lim_{n\to\infty} \lambda(K_n) \tag{14.438}$$

Donc tous les $\lambda(F)$ peuvent être arbitrairement approchés par un $\lambda(K)$ avec K compact dans A, et le supremum (14.437) n'est pas affecté en nous restreignant à prendre des compacts contenus dans B :

$$\lambda(A) = \sup_{F \text{ fermé dans } A} \lambda(F) = \sup_{K \text{ compact dans } A} \lambda(K). \tag{14.439}$$

□

14.8.3 Fonctions mesurables

Lemme 14.156.
Soit une fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable telle que $\lambda(f \neq 0) > 0$. Alors il existe une partie mesurable M et $m > 0$ tels que $\lambda(M) > 0$ et $f(x) > m$ pour tout $x \in M$.

Démonstration. Nous notons

$$D = \{x \in \mathbb{R} \text{ tel que } f(x) > 0\}, \tag{14.440}$$

et nous supposons que $\lambda(D) > 0$ pour fixer les idées (si ce n'est pas le cas, nous prenons pour D la partie où f est strictement négative).

55. parce que fermé et borné, théorème de Borel-Lebesgue 9.24.

Nous posons

$$A_1 = [1, \infty[\tag{14.441a}$$
$$A_n = [\frac{1}{n}, \frac{1}{n-1}[. \tag{14.441b}$$

Ces parties A_n sont disjointes ; donc les parties

$$D_n = \{x \in \mathbb{R} \text{ tel que } f(x) \in A_n\} \tag{14.442}$$

sont également disjointes. Vu que $\bigcup_n A_n =]0, \infty[$, nous avons $D = \bigcup_{n\in\mathbb{N}} D_n$. Vu que

$$\lambda(D) = \sum_{n=1}^{\infty} \lambda(D_n) > 0, \tag{14.443}$$

il existe au moins un N tel que $\lambda(D_N) > 0$. Pour $x \in D_N$ nous avons

$$f(x) \in A_N = [\frac{1}{N}, \frac{1}{N-1}[. \tag{14.444}$$

Donc pour $x \in D_N$ nous avons $f(x) > \frac{1}{N}$. □

14.8.4 Ensemble de Vitali (non mesurable)

EXooCZCFooRPgKjj

Exemple 14.157 (Un ensemble non mesurable au sens de Lebesgue[455])**.**
Nous considérons l'ensemble quotient $\mathbb{R}/\mathbb{Q}$; chaque classe intersecte l'intervalle $[0, 1]$. Grâce à l'axiome du choix (voir 1.9) nous pouvons construire un ensemble V contenant un représentant dans $[0, 1]$ de chaque classe. Un tel ensemble est un **ensemble de Vitali**. Nous allons prouver que V n'est pas mesurable.

Supposons que V soit mesurable. Alors tous les ensembles de la forme $V + q$ ($q \in \mathbb{Q}$) sont mesurables et ont même mesure par la proposition 14.149. Nous posons

$$A = \bigcup_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} (V + q) \subset [-1, 2]. \tag{14.445}$$

Cela est une union disjointe d'ensembles mesurables. Donc

$$\lambda(A) = \sum_{\substack{q\in\mathbb{Q}\\ -1\leqslant q\leqslant 1}} \lambda(V + q). \tag{14.446}$$

Vu que $A \subset [-1, 2]$ nous avons $\lambda(A) \leqslant 3$ et donc tous les termes de la somme doivent être nuls. Nous avons donc $\lambda(A) = 0$.

Prouvons toutefois que $[0, 1] \subset A$, ce qui serait une contradiction. Soit $x \in [0, 1]$; il est dans une des classes de $\mathbb{R}/\mathbb{Q}$ et donc il existe $v \in V$ tel que $x - v \in \mathbb{Q}$. De plus $x, v \in [0, 1]$, donc

$$-1 \leqslant x - v \leqslant 1. \tag{14.447}$$

Cela fait que $x \in V + (x - v) \subset A$. Nous avons donc $x \in A$ et donc $[0, 1] \subset A$. En conséquence de quoi nous aurions $\lambda(A) \geqslant 1$. △

14.8.5 Ensemble de Cantor

Nous considérons la fonction donnant l'écriture décimale des nombres définie en (10.370).

DefIYDooVIDJs

Définition 14.158 (Ensemble de Cantor)**.**
Soit $K_0 = [0,1[$ et les ensembles K_n définis par la récurrence

$$K_{n+1} = \big(\frac{1}{3}K_n\big) \cup \big(\frac{1}{3}(K_n + 2)\big). \tag{14.448}$$

L'ensemble

$$K = \bigcap_{n\geqslant 0} K_n \tag{14.449}$$

*est l'**ensemble triadique de Cantor**.*

Les principales propriétés de l'ensemble de Cantor sont qu'il est non dénombrable (proposition 14.162) et borélien de mesure nulle (proposition 14.163).

14.159.
L'idée de base pour prouver que l'ensemble K est non dénombrable est que ses éléments sont les nombres qui s'écrivent en base 3 sans utiliser le chiffre 1. En prenant un nombre sans 1 écrit en base 3, en changeant tous les 2 en 1 et en lisant le résultat en base 2, nous obtenons tous les nombres possibles en base 2 et donc une quantité non dénombrable. L'idée est donc simple et astucieuse. La mise en musique est un peu plus délicate parce qu'il faut faire attention aux queues de suites ; c'est pour cela que nous avons construit l'ensemble de Cantor en partant de $[0,1[$ et non de $[0,1]$.

Le lemme suivant dit précisément ce que nous entendons en disant que les éléments de l'ensemble de Cantor sont les nombres qui s'écrivent en base 3 sans utiliser le chiffre 1. Nous rappelons que $\mathbb{D}_3$ est l'ensemble des suites constituées de 0, 1 et 2, et qui ne se terminent pas par une suite infinie de 2, voir 10.143 pour une définition précise.

LemAZGoosKzEm

Lemme 14.160 ([1])**.**
Soit $n \in \mathbb{N}$ et $x \in \mathbb{D}_3$ (définition 10.143) ; nous avons $\varphi_3(x) \in K_n \in$ si et seulement si $x_1, \dots, x_n \in \{0,2\}$.

Démonstration. Nous procédons par récurrence en commençant avec $n = 1$. Si $x_1 = 1$ alors

$$\varphi_3(x) = \frac{1}{3} + \sum_{k=2}^{\infty} \frac{x_k}{3^k} \in [\frac{1}{3}, \frac{2}{3}[. \tag{14.450}$$

Notons que $\varphi_3(x) = \frac{2}{3}$ est impossible parce que ça demanderait une queue de suite de 2. Par conséquent $\varphi_3(x) = [0,1[\backslash[\frac{1}{3}, \frac{2}{3}[= K_1$.

Nous passons à la récurrence.

(1) **Sens direct**

Nous supposons que $x_1, \dots, x_{n+1} \in \{0,2\}$ et nous montrons que $\varphi_3(x) \in K_{n+1}$. Nous considérons deux cas suivant que x_1 vaut 0 ou 1. Pour comprendre pourquoi nous divisons les cas suivant la valeur de x_1 et non de x_n, faire un dessin de comment K_n se transforme en K_{n+1} et remarquer dans K_2, les deux premiers segments ne sont pas une division du premier segment de K_1, mais bien une copie des *deux* segments de K_1.

Écrivons encore $\varphi_3(x)$:

$$\varphi_3(x) = \sum_{k=1}^{n+1} \frac{x_k}{3^k} + \sum_{k=n+2}^{\infty} \frac{x_k}{3^k}. \tag{14.451}$$

(1a) **Si $x_1 = 0$** Alors nous avons

$$3\varphi_3(x) = \sum_{k=2}^{\infty} \frac{x_k}{3^{k-1}} = \sum_{k=1}^{\infty} \frac{x_{k+1}}{3^k} = \varphi_3(x_2, \dots, x_n, x_{n+1}, \dots) \tag{14.452}$$

Vu que par hypothèse $x_2, \dots, x_{n+1}$ sont dans $\{0,2\}$ nous avons $3\varphi_3(x) \in K_n$ par hypothèse de récurrence. Cela implique que $\varphi_3(x) \in K_{n+1}$.

(1b) **Si $x_1 = 2$** Alors

$$\varphi_3(x) = \frac{2}{3} + \sum_{k=2}^{\infty} \frac{x_k}{3^k}, \tag{14.453}$$

et

$$3\varphi_3(x) - 2 = \sum_{k=1}^{\infty} \frac{x_{k+1}}{3^k} = \varphi(x_2, \ldots, x_{n+1}, \ldots), \tag{14.454}$$

et donc là nous avons $3\varphi_3(x) - 2 \in K_n$, ce qui implique encore $\varphi_3(x) \in K_{n+1}$.

(2) **Sens réciproque**
Nous devons maintenant prouver que $\varphi_3(x) \in K_{n+1}$ implique $x_1, \ldots, x_{n+1} \in \{0, 2\}$. Par le même calcul que précédemment nous avons soit

$$3\varphi_3(x) = \varphi_3(x_2, \ldots, x_{n+1}, \ldots), \tag{14.455}$$

si $x_1 = 0$, soit

$$3\varphi_3(x) - 2 = \varphi_3(x_2, \ldots, x_{n+1}, \ldots), \tag{14.456}$$

si $x_1 = 2$. Dans les deux cas, si $x_l = 1$ pour un certain $2 \leqslant l \leqslant n + 1$, alors l'hypothèse de récurrence donne que ces éléments ne sont pas dans K_n et donc $\varphi_3(x)$ pas dans K_{n+1}.

□
CorSEDooJmeXt

Corolaire 14.161 ([1]).
En posant $\mathbb{E} = \{x \in \mathbb{D}_3$ tel que $x_i \neq 1 \forall i\}$ nous avons $K = \varphi_3(\mathbb{E})$. Et plus précisément, $\varphi_3 \colon \mathbb{E} \to K$ est une bijection.

Démonstration. Nous divisons la preuve en trois étapes.

(1) **Image contenue dans K** Si $x \in \mathbb{E}$ et $n \in \mathbb{N}$ nous avons $x_1, \ldots, x_n \in \{0, 2\}$ et donc $\varphi_3(x) \in K_n$ par la proposition 14.160. Donc

$$\varphi_3(x) \in \bigcap_{n \geqslant 1} K_n = K. \tag{14.457}$$

(2) **Injective** L'application $\varphi_3 \colon \mathbb{E} \to K$ est injective parce qu'elle est déjà injective depuis $\mathbb{D}_3$.

(3) **Surjective** Soit $p \in K \subset [0, 1[$. Vu que $\varphi_3 \colon \mathbb{D}_3 \to [0, 1[$ est surjective (théorème 10.146), il existe $x \in \mathbb{D}_3$ tel que $\varphi_3(x) = p$. Pour tout n nous avons $\varphi_3(x) \in K_n$ et donc $x_1, \ldots, x_n \in \{0, 2\}$ et donc au final $x \in \mathbb{E}$.

□
PropTPPooDySbm

Proposition 14.162 ([1]).
L'ensemble de Cantor est non dénombrable.

Démonstration. Nous avons prouvé à la proposition 10.147 que l'ensemble $\mathbb{D}_2$ n'était pas dénombrable. Nous allons à présent prouver que l'application

$$\begin{aligned} \psi \colon \mathbb{D}_2 &\to K \\ c &\mapsto \varphi_3(c \text{ en remplaçant les 1 par des 2}) \end{aligned} \tag{14.458}$$

est une bijection. Le fait que ψ soit injective est une conséquence du fait que ce soit la composition de deux applications injectives (le remplacement et φ_3). Il faut par contre montrer que l'image est égale à K, en notant qu'il n'est pas évident à priori que l'image soit contenue dans K.

L'opération qui consiste à remplacer les 1 par des 2 est une bijection $\mathbb{D}_2 \to \mathbb{E}$. Le corolaire 14.161 nous dit aussi que $\varphi_3 \colon \mathbb{E} \to K$ est une bijection. En tant que composée de bijections, ψ est une bijection.

Étant en bijection avec $\mathbb{D}_2$ qui n'est pas dénombrable par la proposition 10.147, l'ensemble de Cantor n'est pas dénombrable. □

PropBEWooXZdKN

Proposition 14.163 (Ensemble de Cantor)**.**
L'ensemble de Cantor[56] *est borélien, non dénombrable et de mesure nulle.*

Démonstration. Nous reprenons les notations de la définition 14.158. Le fait que l'ensemble de Cantor soit non dénombrable a été prouvé dans la proposition 14.162.

L'ensemble de Cantor étant une intersection dénombrable de boréliens, il est borélien par le lemme 14.3. Vu que $K_n \subset [0,1[$ nous avons $\frac{1}{3}K_n \leqslant \frac{1}{3}$ et $\frac{1}{3}(K_n+2) \geqslant \frac{2}{3}$, donc K_n est une union disjointe de 2^n intervalles de mesure $2/3^n$. Nous avons donc

$$\lambda(K_n) = \left(\frac{2}{3}\right)^n. \tag{14.459}$$

L'ensemble de Cantor étant contenu dans chacun des K_n, sa mesure est plus petite que la mesure de chacun des K_n (lemme 14.23) et donc $\lambda(K) \leqslant \left(\frac{2}{3}\right)^n$ pour tout n ; ergo $\lambda(K) = 0$. □

14.8.6 Mesure positive sans intervalle

Vu que la mesure de Lebesgue est basée sur la mesure des intervalles et quelques extensions, nous sommes en droit de croire qu'une partie de mesure strictement positive de $\mathbb{R}$ doit toujours contenir un intervalle, éventuellement à partie de mesure nulle près. Eh bien non.

EXooVZVIooXZvDaE

Exemple 14.164 ([456])**.**
Soient une énumération (q_i) de $\mathbb{Q} \cap]0,1[$ et une suite (r_i) telle que $\sum_{i=0}^{\infty} r_i < \frac{1}{2}$. Quitte à prendre r_i plus petit, supposons de plus que $B(q_i, r_i) \subset [0,1]$.

Nous posons $J_n = B(q_n, r_n/2)$, $J = \bigcup_{n=0}^{\infty} J_n$ et

$$B = [0,1] \backslash J. \tag{14.460}$$

Les parties J_i ne sont pas disjointes, donc, en notant λ la mesure de Lebesgue,

$$0 < \lambda(J) \leqslant \sum_{i=0}^{\infty} \lambda(J_i) \leqslant \frac{1}{2}. \tag{14.461}$$

Mais, par définition, l'union $[0,1] = B \cup J$ est disjointe, donc

$$1 = \lambda\big([0,1]\big) = \lambda(J) + \lambda(B). \tag{14.462}$$

Nous en déduisons que

$$\frac{1}{2} \leqslant \lambda(B) \leqslant 1. \tag{14.463}$$

Je plaide que cette partie B ne contient non seulement aucun intervalle, mais qu'il est impossible de le compléter par une partie de mesure nulle pour obtenir un intervalle.

Soit un intervalle I dans $[0,1]$. Il existe $q_i \in I$ et donc[57]

$$J_i \subset B \backslash I. \tag{14.464}$$

Donc il n'existe pas de parties de mesure nulle qui, ajoutée à B, contiendrait I. △

Vous voulez un truc dingue à propos de la partie J de l'exemple 14.164 ? Le théorème 14.155 nous dit qu'il existe dans J des compacts de mesure arbitrairement proches de $\lambda(J)$. Il existe donc des compacts non seulement de mesure strictement positive mais même de mesure assez grande, tout en étant infiniment découpés.

56. Définition 14.158

57. C'est ici que nous utilisons le fait que r_i est choisi pour que $B(q_i, r_i)$ ne déborde pas de $[0,1]$. Sinon il aurait fallu chipoter et prendre seulement une partie de la boule.

14.9 Intégrale par rapport à une mesure

14.165.
Nous n'en avons pas encore terminé avec la théorie de la mesure, mais nous devons quand même définir les intégrales et voir quelques propriétés avant de continuer avec la mesure parce que la définition de la mesure sur un espace mesurable produit [58] passe par une intégrale.

NORMooFZEDooIxSgLe

14.166.
En théorie de l'intégration, la convention est la suivante : pour une fonction $f\colon X \to \mathbb{R}$, nous considérons sur X la tribu des ensembles mesurables au sens de Lebesgue sur X, *tout en gardant celle des boréliens sur l'ensemble d'arrivée.* C'est-à-dire qu'en théorie de l'intégration, c'est

$$f\colon \big(X, \mathcal{L}\mathrm{eb}(X)\big) \to \big(\mathbb{R}, \mathcal{B}\mathrm{or}(\mathbb{R})\big). \tag{14.465}$$

En particulier, $f\colon \mathbb{R}^n \to \mathbb{R}^m$ sera mesurable si pour tout borélien A de $\mathbb{R}^m$ l'ensemble $f^{-1}(A)$ est Lebesgue-mesurable dans $\mathbb{R}^n$.

Étant donné qu'il est franchement difficile de créer des ensembles non mesurables au sens de Lebesgue, il est franchement difficile de créer des fonctions non mesurables à valeurs réelles. L'hypothèse de mesurabilité est donc toujours satisfaite dans les cas pratiques.

Voir aussi le point 14.120, et les résultats qui suivent.

14.9.1 Définition pour les fonctions à valeurs positives

Voir le thème 30.

Une mesure μ sur un espace mesurable $(\Omega, \mathcal{A})$ permet de définir une fonctionnelle linéaire sur l'ensemble des fonctions mesurables $\Omega \to \mathbb{R}$. Cette fonctionnelle linéaire est l'intégrale que nous allons définir à présent.

DefTVOooleEst

Définition 14.167.
Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré ainsi que $Y \in \mathcal{A}$. Notre but est de définir

$$\int_Y f d\mu \tag{14.466}$$

*que nous nommons **intégrales de** f sur Y pour la mesure μ.*

(1) ***Fonction étagée*** ITEMooQVEAooZhRtUC *Si $f\colon \Omega \to [0, \infty[$ est une fonction étagée [59], et si sa forme canonique est $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ alors nous définissons*

$$\int_Y f d\mu = \sum_i \alpha_i \mu(Y \cap A_i). \tag{14.467}$$ EqooGAFMooZLziPs

(2) ***Fonction mesurable à valeurs positives*** ITEMooQGKEooZVhHHq *Pour une fonction $\mathcal{A}$-mesurable $f\colon \Omega \to [0, \infty]$ nous définissons l'intégrale de f sur Y par*

$$\int_Y f d\mu = \sup\Big\{ \int_Y \psi d\mu \text{ où } \psi \text{ est une fonction étagée finie telle que } 0 \leqslant \psi \leqslant f \Big\}. \tag{14.468}$$ EqDefintYfdmu

CORooZPLOooXgbtDs

Corolaire 14.168 ([1])**.**
Si $f\colon Z \to [0, \infty]$ sur $Z \subset \Omega$ et si $\mu(Z) = 0$, alors

$$\int_Z f = 0. \tag{14.469}$$

58. Théorème 14.246.
59. Définition 14.112.

Démonstration. Notons d'abord que si A est une partie mesurable de Z, alors $\mu(A) = 0$ par le lemme 14.19. Si f ne prend que des valeurs finies, tous les termes de (14.467) sont nuls.

Si une fonction prends des valeurs infinies, la partie (1) de la définition n'est pas applicable. La partie (2) demande alors de prendre un supremum sur des fonctions qui rentrent dans la partie (1).

Si ψ est une fonction étagée finie sur Z, alors $\int_Z \psi = 0$ parce que tous les termes de (14.467) sont nuls. □

Remarque 14.169 (Pourquoi se limiter aux fonctions positives)**.**
Toute fonction mesurable à valeurs dans $[0, +\infty]$ est intégrable (l'intégrale vaut éventuellement $+\infty$). Au moment où une fonction commence à prendre des valeurs positives et négatives, nous demandons à pouvoir intégrer séparément les parties positives et négatives. C'est pour cela que nous disons qu'une fonction f à valeurs dans $\mathbb{R}$ est intégrable si $|f|$ l'est.

14.170.
Le nombre $\int_0^\infty f$ est défini directement par (14.468) complètement indépendamment d'une éventuelle limite $\lim_{x\to\infty} \int_0^x f$. Cette limite sera traitée dans le lemme 14.261.

NORMooXTGBooKDnAhZ

14.171.
Si la fonction n'est pas mesurable ? Alors nous n'avons pas défini son intégrale. Supposons la plus simple des fonctions non mesurables sur Ω : la fonction indicatrice d'une partie non mesurable :

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon.} \end{cases} \tag{14.470}$$

où $A \subset \Omega$ n'est pas mesurable [60].

Nous supposons que l'espace mesuré $(\Omega, \mathcal{F}, \mu)$ est complet (définition 14.68). Vu que A n'est pas mesurable, il n'est pas contenu dans une partie négligeable (parce que l'espace est complet), et nous voulons que l'intégrale ne soit pas nulle ; sinon on se demande bien à quoi sert une intégrale.

Toute fonction étagée minorant f est forcément nulle en dehors de A. Dès que B est une partie mesurable de mesure non nulle dans A, le complémentaire de B dans A est encore non mesurable, et nous voulons encore que l'intégrale de f sur ce complémentaire soit non nulle.

Mais comme A n'est pas mesurable et que $\mathbb{1}_A$ n'est le supremum d'aucune suite de fonctions mesurables (lemme 14.100), bien que le supremum qui définirait l'intégrale de f existe (toute partie de $\mathbb{R}$ a un supremum), il est sans espoir que ce supremum ait un sens que l'on puisse interpréter en tant que mesure de f.

LEMooHAUGooWITETb

Lemme 14.172.
L'intégrale d'une fonction positive nulle presque partout est nulle.

Démonstration. Soient un espace mesuré $(\Omega, \mathcal{A}, \mu)$, et une fonction mesurable $f : \Omega \to [0, \infty]$. Nous posons

$$\Omega_+ = \{x \in \Omega \text{ tel que } f(x) \neq 0\}. \tag{14.471}$$

L'hypothèse est que $\mu(\Omega_+) = 0$. Nous devons prouver que $\int_\Omega f d\mu = 0$. Vu que f est positive, nous utilisons la définition 14.467. Soit une fonction étagée positive ψ minorant f. Nous la décomposons en

$$\psi = \sum_{k=1}^{n} \psi_k \mathbb{1}_{A_k} \tag{14.472}$$

où les A_k sont mesurables et $\psi_k \in [0, \infty[$. Nous allons prouver que $\psi_i \mu(A_k) = 0$ pour tout k, en séparant trois cas.

60. Ça existe, par exemple 14.157.

(1) **Si $A_k \cap \Omega_+ = \varnothing$** Soit $x \in A_k$. Nous avons

$$0 = \psi(x) = \sum_k \psi_k \mathbb{1}_{A_k}(x) = \psi_k. \tag{14.473}$$

Donc $\psi_k = 0$.

(2) **Si $A_k \subset \Omega_+$** Alors, par le lemme 14.23, $\mu(A_k) \leqslant \mu(\Omega_+) = 0$ et donc $\psi_k \mu(A_k) = 0$.

(3) **Si $A_k \cap \Omega_+ \neq A_k$** Soit $x \in A_k \backslash \Omega_+$. Nous avons

$$\psi_k = \psi(x) \leqslant f(x) = 0, \tag{14.474}$$

et donc encore $\psi_k = 0$.

Nous avons donc prouvé que pour toute fonction étagée positive minorant f,

$$\int_\Omega \psi d\mu = \sum_{k=1}^{n} \psi_k \mu(A_k) = 0. \tag{14.475}$$

Le supremum est donc nul. □

14.9.2 Premières propriétés

14.173.
Si $(\Omega, \mathcal{A}, \mu)$ est un espace mesurable, et si Y est un élément de $\mathcal{A}$, nous avons l'espace mesurable $(Y, \mathcal{A}_Y, \mu_Y)$ donné par

— $\mathcal{A}_y = \{B \cap Y \text{ tel que } B \in \mathcal{A}\}$,

— $\mu_Y = \mu$.

Et là, nous arrivons à un problème de notations parce que $\int_Y f d\mu$ peut désigner l'intégrale de f sur Y dans $(\Omega, \mathcal{A}, \mu)$ ou l'intégrale de f sur Y dans $(Y, \mathcal{A}_Y, \mu_Y)$.

Heureusement, nous allons tout de suite montrer que ces deux choses sont identiques.

Lemme 14.174.
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ ainsi que $Y \in \mathcal{A}$. Nous considérons une fonction $f : \Omega \to \mathbb{R}^+$ qui est $\mathcal{A}$-mesurable et intégrable sur Y.

Alors, avec des notations que j'espère être claires,

(1) f est $\mathcal{A}_Y$-mesurable,

(2) f est $(Y, \mathcal{A}_Y, \mu_Y)$-intégrable,

(3) nous avons l'égalité

$$\int_{(Y, \mathcal{A}_Y, \mu_Y)} f|_Y = \int_{(Y \subset \Omega, \mathcal{A}, \mu)} f. \tag{14.476}$$

Démonstration. Nous considérons les deux ensembles suivants :

$$S_1 = \{\psi \text{ étagées sur } Y \text{ et majorées par } f|_Y\} \tag{14.477a}$$

$$S_2 = \{\psi \text{ étagées sur } \Omega \text{ et majorées par } f\}. \tag{14.477b}$$

Nous considérons l'application suivante :

$$\begin{aligned} s &: S_1 \to S_2 \\ s(\psi)(x) &= \begin{cases} \psi(x) & \text{si } x \in Y \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{14.478}$$

L'application s est une bijection.

Pour $\psi \in S_1$ nous avons

$$\psi = \sum_{k=1}^{n} \psi_k \mathbb{1}_{A_k}|_Y \tag{14.479}$$

avec $A_k \in \mathcal{A}_Y \subset \mathcal{A}$ et $\mathbb{1}_{A_k}|_Y : Y \to \{0,1\}$. Nous avons aussi

$$s(\psi) = \sum_k \psi_k \mathbb{1}_{A_k} \tag{14.480}$$

avec $\mathbb{1}_{A_k} : \Omega \to \{0,1\}$.

En ce qui concerne les intégrales de ces fonctions étagées, nous avons

$$\begin{aligned}\int_{(Y,\mathcal{A}_Y,\mu_Y)} \psi &= \sum_{k=1}^{n} \psi_k \mu_Y(A_k \cap Y) && \text{(14.481a)}\\ &= \sum_{k=1}^{n} \psi_k \mu(A_k) && \text{(14.481b)}\\ &= \int_{(Y\subset\Omega,\mathcal{A},\mu)} s(\psi). && \text{(14.481c)}\end{aligned}$$

SUBEQooGWYGooQuucEo

Justifications. Pour passer à (14.481b) nous avons utilisé d'abord que $A_k \subset Y$ et ensuite que $\mu_Y(A_k) = \mu(A_k)$.

Nous sommes maintenant prêts à prouver l'égalité du lemme. Nous avons ceci :

$$\begin{aligned}\int_{(Y,\mathcal{A}_Y,\mu_Y)} f|_Y &= \sup\{\int_{(Y,\mathcal{A}_Y,\mu_Y)} \psi \text{ tel que } \psi \in S_1\} && \text{(14.482a)}\\ &= \sup\{\int_{(Y\subset\Omega,\mathcal{A},\mu)} s(\psi) \text{ tel que } \psi \in S_1\} && \text{(14.482b)}\\ &= \sup\{\int_{(Y\subset\Omega,\mathcal{A},\mu)} \varphi \text{ tel que } \varphi \in S_2\} && \text{(14.482c)}\\ &= \int_{(Y\subset\Omega,\mathcal{A},\mu)} f. && \text{(14.482d)}\end{aligned}$$

□

LemooPJLNooVKrBhN

Lemme 14.175.
Si $(\Omega,\mathcal{A},\mu)$ est un espace mesuré et si $B \in \mathcal{A}$ alors

$$\mu(B) = \int_B 1 d\mu = \int_\Omega \mathbb{1}_B. \tag{14.483}$$

Démonstration. La fonction caractéristique d'une partie mesurable est une fonction étagée dont la forme canonique est $\mathbb{1}_B = 1 \cdot \mathbb{1}_B + 0 \times \mathbb{1}_{B^c}$. Son intégrale est donc

$$\int \mathbb{1}_B d\mu = 1 \times \mu(B) + 0 \times \mu(B^c) = \mu(B) \tag{14.484}$$

parce que $0 \times \mu(B^c) = 0$, même si $\mu(B^c) = \infty$, comme nous l'avons convenu en 14.93. □

PROPooGTMVooPHcrRl

Proposition 14.176 ([1]).
Soient une fonction $f : (\Omega,\mathcal{A},\mu) \to \mathbb{R}^+$ et une fonction g intégrable sur Ω telle que $f \leqslant g$. Alors f est intégrable.

Démonstration. Une fonction étagée qui minore f minore également g. Donc l'ensemble sur lequel il faut faire le supremum pour définir $\int_\Omega f$ est inclus dans celui pour $\int_\Omega g$. Le second supremum étant fini, le premier l'est également. □

LEMooSPOFooBxDEAV

Lemme 14.177.
Soient un espace mesuré $(\Omega,\mathcal{A},\mu)$, une fonction $f : \Omega \to \mathbb{R}^+$ et $Y \in \mathcal{A}$. Nous avons :

$$\int_Y f d\mu = \int_\Omega f \mathbb{1}_Y d\mu. \tag{14.485}$$

EQooSBDKooPTDEcr

Démonstration. En plusieurs parties, selon la généralité.

(1) **Si f est étagée** Nous posons $f = \sum_{k=1}^{n} f_k \mathbb{1}_{A_k}$ avec $f_k \in \mathbb{R}^+$. Dans ce cas,

$$f\mathbb{1}_Y = \sum_k f_k \mathbb{1}_{A_k \cap Y} \tag{14.486}$$

est encore une fonction étagée. Donc nous avons d'une part

$$\int_\Omega f\mathbb{1}_Y = \int_\Omega \sum_k f_k \mathbb{1}_{A_k \cap Y} = \sum_k f_k \mu(A_k \cap Y), \tag{14.487}$$

et d'autre part,

$$\int_Y f d\mu = \sum_k f_k \mu(Y \cap A_k), \tag{14.488}$$

(2) **Si f est à valeurs positives** Nous posons

$$S_1 = \{\psi \text{ étagées sur } \Omega \text{ tel que } 0 \leqslant \psi \leqslant f\mathbb{1}_Y\} \tag{14.489}$$

et

$$S_2 = \{\psi\mathbb{1}_Y \text{ tel que } \psi \text{ étagée avec } 0 \leqslant \psi \leqslant f\}. \tag{14.490}$$

Nous prouvons que $S_1 = S_2$.
Si $\psi \in S_1$, alors

$$0 \leqslant \psi \leqslant f\mathbb{1}_Y \leqslant f. \tag{14.491}$$

De plus comme $\psi = 0$ hors de Y nous avons $\psi = \psi\mathbb{1}_Y$.
Pour l'autre inclusion, soit $0 \leqslant \psi \leqslant f$ pour une fonction étagée ψ et montrons que $\psi\mathbb{1}_Y \in S_1$. L'application $\psi\mathbb{1}_Y$ est étagée sur Ω et vérifie

$$0 \leqslant \psi\mathbb{1}_Y \leqslant f\mathbb{1}_Y \tag{14.492}$$

parce que $\psi \leqslant f$.

(3) **L'égalité à prouver** Dans l'égalité 14.485 à prouver, le membre de droite est, d'après la définition 14.468,

$$\int_\Omega f\mathbb{1}_Y = \sup\{\int \psi \text{ tel que } \psi \in S_1\}. \tag{14.493}$$

Il nous reste donc à prouver que $\int_Y f$ se calcule de la même façon avec les éléments de S_2. D'abord nous copions la définition :

$$\int_Y f = \sup\{\int_Y \psi \text{ tel que } 0 \leqslant \psi \leqslant f\}. \tag{14.494}$$

Ensuite nous réfléchissons un peu. Si $0 \leqslant \psi \leqslant f$ avec $\psi = \sum_k \psi_k \mathbb{1}_{A_k}$, alors

$$\int_Y \psi = \sum_k \mu(A_k \cap Y)\psi_k = \int_Y \psi\mathbb{1}_Y = \int_\Omega \psi\mathbb{1}_Y. \tag{14.495}$$

La dernière égalité est la partie déjà faite, à propos des fonctions étagées. Nous avons donc bien

$$\int_Y f = \sup\{\int_Y s \text{ tel que } s \in S_2\}. \tag{14.496}$$

□

14.9.3 Propriétés plus avancées

14.9.3.1 Convergence monotone

Le théorème suivant est très utile parce que le théorème fondamental d'approximation 14.118 donne les fonctions étagées qu'il faut.

ThoRRDooFUvEAN

Théorème 14.178 (Théorème de la convergence monotone ou de Beppo-Levi[457]).
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et (f_n) une suite croissante de fonctions mesurables à valeurs dans $[0, \infty]$. Alors la limite ponctuelle $\lim_{n\to\infty} f_n$ existe, est mesurable et

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega \lim_{n\to\infty} f_n d\mu, \tag{14.497}$$

EqFHqCmLV

cette intégrable valant éventuellement ∞.

Démonstration. La limite ponctuelle de la suite est la fonction à valeurs dans $[0, \infty]$ donnée par

$$f(x) = \lim_{n\to\infty} f_n(x). \tag{14.498}$$

Ces limites existent parce que pour chaque x la suite $f_n(x)$ est une suite numérique croissante. Nous notons

$$I_0 = \int_\Omega f d\mu. \tag{14.499}$$

Nous posons par ailleurs

$$I_n = \int_\Omega f_n. \tag{14.500}$$

Cela est une suite numérique croissante qui a par conséquent une limite que nous notons $I = \lim_{n\to\infty} I_n$. Notre objectif est de montrer que $I = I_0$. D'abord par croissance de la suite, pour tous n nous avons $I_n \leqslant I_0$, par conséquent $I \leqslant I_0$.

Nous prouvons maintenant l'inégalité dans l'autre sens en nous servant de la définition (14.468). Soit une fonction étagée h telle que $h \leqslant f$, et une constante $0 < C < 1$. Nous considérons les ensembles

$$E_n = \{x \in \Omega \text{ tel que } f_n(x) \geqslant Ch(x)\}. \tag{14.501}$$

Ces ensembles vérifient les propriétés $E_n \subset E_{n+1}$ et $\bigcup_{n=1}^{\infty} E_n = \Omega$. Pour chaque n nous avons les inégalités

$$\int_\Omega f_n \geqslant \int_{E_n} f_n \geqslant C \int_{E_n} h. \tag{14.502}$$

Si nous prenons la limite $n \to \infty$ dans ces inégalités,

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant C \lim_{n\to\infty} \int_{E_n} h = C \int_\Omega h. \tag{14.503}$$

Par conséquent $\lim_{n\to\infty} \int f_n \geqslant C \int_\Omega h$. Mais étant donné que cette inégalité est valable pour tout C entre 0 et 1, nous pouvons l'écrire sans le C :

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant \int_\Omega h. \tag{14.504}$$

EqzAKEaU

Par définition, l'intégrale de f est donné par le supremum des intégrales de h où h est une fonction simple dominée par f. En prenant le supremum sur h dans l'équation (14.504) nous avons

$$\lim_{n\to\infty} \int_\Omega f_n \geqslant \int_\Omega f, \tag{14.505}$$

ce qu'il nous fallait. □

Remarque 14.179.
La proposition 14.118 ainsi que le lemme 14.116 montrent qu'une fonction mesurable peut-être écrite comme limite croissante de fonctions simples. Cela permet de démontrer des théorèmes en commençant par prouver sur les fonctions simples et en utilisant Beppo-Levi pour généraliser.

Remarque 14.180.
Une des raisons de demander la positivité des fonctions f_n est de n'avoir pas d'ambiguïté à parler d'intégrales qui valent ∞. Si par exemple nous prenons $\Omega = [0,1]$ et que nous considérons

$$f_n(x) = \begin{cases} 0 & \text{si } x \leqslant \frac{1}{n} \\ \frac{1}{x} & \text{sinon.} \end{cases} \tag{14.506}$$

Ce sont des fonctions intégrables, mais la limite étant la fonction $1/x$, l'égalité (14.497) est une égalité entre deux intégrales valant ∞.

CorNKXwhdz

Corolaire 14.181 (Inversion de somme et intégrales)**.**
Si (u_n) est une suite de fonctions mesurables positives ou nulles, alors

$$\sum_{i=0}^{\infty} \int u_i = \int \sum_{i=0}^{\infty} u_i. \tag{14.507}$$

Démonstration. Nous considérons la suite des sommes partielles de (u_n) : $f_n(x) = \sum_{i=0}^{n} u_i(x)$. Le théorème de la convergence monotone (théorème 14.178) implique que

$$\lim_{n\to\infty} \int f_n = \int \lim_{n\to\infty} f_n. \tag{14.508}$$

Nous remplaçons maintenant f_n par sa valeur en termes des u_i et dans le membre de gauche nous permutons l'intégrale avec la somme finie :

$$\lim_{n\to\infty} \sum_{i=0}^{n} \int u_i = \int \sum_{i=0}^{\infty} u_n, \tag{14.509}$$

ce qu'il fallait démontrer. □

14.9.3.2 Lemme de Fatou

LemFatouUOQqyk

Lemme 14.182 (Lemme de Fatou)**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f_n \colon \Omega \to [0,\infty]$ une suite de fonctions mesurables. Alors la fonction $f(x) = \liminf f_n(x)$ est mesurable et

$$\int_\Omega \liminf f_n d\mu \leqslant \liminf \int_\Omega f_n d\mu. \tag{14.510}$$

Démonstration. Nous posons

$$g_n(x) = \inf_{i \geqslant n} f_i(x). \tag{14.511}$$

Cela est une suite croissance de fonctions positives mesurables telles que, par définition,

$$\lim_{n\to\infty} g_n(x) = \liminf f_n(x). \tag{14.512}$$

Nous pouvons y appliquer le théorème de la convergence monotone,

$$\lim_{n\to\infty} \int g_n(x) = \int \liminf f_n(x). \tag{14.513}$$

Par ailleurs, pour chaque $i \geqslant n$ nous avons

$$\int g_n \leqslant \int f_i, \tag{14.514}$$

en passant à l'infimum nous avons

$$\int g_n \leqslant \inf_{i\geqslant n} \int f_i, \tag{14.515}$$

et en passant à la limite nous avons

$$\int \liminf f_n = \lim_{n\to\infty} \int g_n \leqslant \lim_{n\to\infty} \inf_{i\geqslant n} \int f_i = \liminf_{i\to\infty} \inf f_i. \tag{14.516}$$

□

L'inégalité donnée dans ce lemme n'est en général pas une égalité, comme le montre l'exemple suivant :

$$f_i = \begin{cases} \mathbb{1}_{[0,1]} & \text{si } i \text{ est pair} \\ \mathbb{1}_{[1,2]} & \text{si } i \text{ est impair.} \end{cases} \tag{14.517}$$

Nous avons évidemment $g_n(x) = 0$ tandis que $\int_{[0,2]} f_i = 1$ pour tout i.

ThoooCZCXooVvNcFD

Théorème 14.183 ([436])**.**
Soient f, g des fonctions étagées positives[61] *sur $(\Omega, \mathcal{A}, \mu)$. Alors si $\alpha \in [0, \infty]$ nous avons*

(1)

$$\int_\Omega (\alpha f) d\mu = \alpha \int_\Omega f d\mu. \tag{14.518}$$

ITEMooBLEVooDznQTY

(2)

$$\int_\Omega (f + g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu. \tag{14.519}$$

ITEMooOJRAooQkoQyD

(3) Si $a_k \in \mathbb{R}^+$ et si les f_k sont étagées positives,

$$\int_\Omega \left(\sum_{k=1}^n a_k f_k \right) = \sum_{k=1}^n a_k \left(\int_\Omega f_k d\mu \right). \tag{14.520}$$

Démonstration. En ce qui concerne le produit par un nombre, tout repose sur le fait que

$$(\alpha f)^{-1}(\alpha a_i) = f^{-1}(a_i), \tag{14.521}$$

ce qui fait que si la représentation canonique de f est $f = \sum_i a_i \mathbb{1}_{A_i}$ alors la représentation canonique de αf est $\alpha f = \sum_i (\alpha a_i) \mathbb{1}_{A_i}$. Donc

$$\int_\Omega \alpha f d\mu = \sum_i \alpha a_i \mu(A_i) = \alpha \sum_i a_i \mu(A_i) = \alpha \int_\Omega f d\mu. \tag{14.522}$$

Pour la somme c'est plus lourd. Soient les formes canoniques

$$f = \sum_i a_i \mathbb{1}_{A_i} \tag{14.523a}$$

$$g = \sum_j b_j \mathbb{1}_{B_j}. \tag{14.523b}$$

Vu que l'union des B_j est Ω nous avons l'union disjointe $A_i = \bigcup_j A_i \cap B_j$ et donc $\mu(A_i) = \sum_j \mu(A_i \cap B_j)$. Nous avons donc pour les intégrales :

$$\int_\Omega f d\mu = \sum_i a_i \sum_j \mu(A_i \cap B_j) \tag{14.524a}$$

$$\int_\Omega g d\mu = \sum_k b_k \sum_l \mu(B_k \cap A_l). \tag{14.524b}$$

61. Pour le cas plus général, voir 14.193.

Pour la somme :

$$\int_\Omega f d\mu + \int_\Omega g d\mu = \sum_{k,l}(a_k + b_l)\mu(A_k \cap B_l). \tag{14.525}$$

Nous devons maintenant évaluer $\int_\Omega (f+g)d\mu$. Pour cela nous remarquons que si $c \in (f+g)(\Omega)$ (l'ensemble des valeurs atteintes pas $f+g$), alors nous notons

$$I_c = \{(k,l) \text{ tel que } a_k + b_l = c\} \tag{14.526}$$

et nous avons

$$\{f + g = c\} = \bigcup_{(k,l)\in I_c} (A_k \cap B_l), \tag{14.527}$$

et comme cette union est disjointe, nous pouvons faire la somme des mesures :

$$\mu(f + g = c) = \sum_{(k,l)\in I_c} \mu(A_k \cap B_l). \tag{14.528}$$

Cela nous permet de faire le calcul suivant :

$$\begin{aligned}
\int_\Omega (f+g)d\mu &= \sum_{c\in(f+g)(\Omega)} c\mu(f+g=c) && \text{(14.529a)}\\
&= \sum_{c\in(f+g)(\Omega)} c \sum_{(k,l)\in I_c} \mu(A_k \cap B_l) && \text{(14.529b)}\\
&= \sum_{c\in(f+g)(\Omega)} \sum_{(k,l)\in I_c} (a_k + b_l)\mu(A_k \cap B_l) && \text{(14.529c)}
\end{aligned}$$

Dans cette double somme, tous les couples (k,l) sont tirés une et une seule fois parce qu'ils sont tous dans un et un seul des I_c, donc

$$\begin{aligned}
\int_\Omega (f+g)d\mu &= \sum_{c\in(f+g)(\Omega)} \sum_{(k,l)\in I_c} (a_k + b_l)\mu(A_k \cap B_l) && \text{(14.530a)}\\
&= \sum_{(k,l)} (a_k + b_l)\mu(A_k \cap B_l) && \text{(14.530b)}\\
&= \int_\Omega f d\mu + \int_\Omega g d\mu. && \text{(14.530c)}
\end{aligned}$$

□

Remarque 14.184.
Si $f = \sum_k a_k \mathbb{1}_{A_k}$ n'est pas une décomposition canonique, il n'en reste pas moins que chacun des $\mathbb{1}_{A_k}$ est la forme canonique de lui-même. Donc le théorème 14.183 s'applique et nous avons quand même

$$\int_\Omega f d\mu = \sum_k a_k \mu(A_k). \tag{14.531}$$

PROPooOVDEooDJvOau

Proposition 14.185.
Soient deux fonctions mesurables $f, g : \Omega \to [0, +\infty]$. Alors

$$\int_\Omega (f+g) = \int_\Omega f + \int_\Omega g. \tag{14.532}$$

Démonstration. Soient des suites $f_n \to f$ et $g_n \to g$ fournies par le théorème fondamental d'approximation 14.118. Par le théorème de la convergence monotone 14.178 nous avons d'une part

$$\lim_{n\to\infty} \int_\Omega (f_n + g_n) = \int_\Omega (f+g), \tag{14.533}$$

et par le théorème 14.183 nous avons d'autre part

$$\lim_{n\to\infty}\int_\Omega (f_n+g_n) = \lim_{n\to\infty}\Big(\int f_n + \int g_n\Big) = \int f + \int g \tag{14.534}$$

où nous avons encore utilisé la convergence monotone.

En égalant les deux, nous avons notre résultat. □

14.9.4 Fonctions à valeurs réelles

L'intégrale d'une fonction à valeurs dans $[0,+\infty]$ étant faite, nous passons aux fonctions à valeurs dans $[-\infty,+\infty]$.

PROPooPVTCXoFZEZCK

Proposition 14.186 ([1]).
Soit une fonction mesurable $f\colon \Omega\to\bar{\mathbb{R}}$. Nous considérons les deux fonction suivantes à valeurs dans $[0,+\infty]$:

$$f^+(x) = \begin{cases} 0 & si\ f(x)<0 \\ f(x) & si\ f(x)\geqslant 0. \end{cases} \tag{14.535a}$$

$$f^-(x) = \begin{cases} 0 & si\ f(x)>0 \\ -f(x) & si\ f(x)\leqslant 0. \end{cases} \tag{14.535b}$$

Nous avons $\int_\Omega |f|<\infty$ si et seulement si $\int_\Omega f^+<\infty$ et $\int_\Omega f^-<\infty$.

Démonstration. Vu que f est mesurable, les fonctions f^+ et f^- sont également mesurables et nous avons l'égalité

$$|f| = f^+ + f^-. \tag{14.536}$$

La proposition 14.185 nous dit alors que

$$\int_\Omega |f| = \int_\Omega f^+ + \int_\Omega f^-. \tag{14.537}$$

Dans cette égalité, tous les nombres sont dans $[0,\infty]$. Le membre de gauche vaut $+\infty$ si et seulement si au moins un des deux de droite vaut $+\infty$. □

NORMooRDHAooKRniYt

14.187 (Intégrable, qui existe, etc.[458]).
Plusieurs définitions à ne pas confondre.

(1) Si

$$\min\left(\int_\Omega f^+, \int_\Omega f^-\right) < \infty, \tag{14.538}$$

EQooCUTSooBRpByt

alors nous disons que $\int_\Omega f$ **existe**, et nous posons

$$\int_\Omega f = \int_\Omega f^+ - \int_\Omega f^- \tag{14.539}$$

EQooOGFDooUSfCDg

où les intégrales de droite sont définies par 14.167. Une intégrale qui existe peut très bien valoir $+\infty$ ou $-\infty$.

(2) Si $\int_\Omega |f|<\infty$, alors nous disons que f est **intégrable**.

D'après ces définitions, en prenant fonction constante $f(x)=1$, nous avons :

— $\int_{\mathbb{R}} 1dx$ existe parce que sa partie négative est identiquement nulle, de telle sorte que $\int_{\mathbb{R}} f^- = 0<\infty$.

— La fonction f n'est pas intégrable parce que $\int_{\mathbb{R}} |1|dx = \infty$.

Nous verrons comment donner un sens à $\int_\Omega f$ dans certains cas où f n'est pas intégrable sur Ω dans la section 14.16.3 sur les intégrales impropres.

Nous définissons aussi

$$\mu(f) = \int_\Omega f \tag{14.540}$$

si f est une fonction mesurable sur Ω.

LEMooMWKTooIKomSw

Lemme 14.188.
Pour $f : \Omega \to \mathbb{R}$ nous avons $\int_\Omega |f| < \infty$ si et seulement si $\int_\Omega f$ existe et est finie.

Démonstration. Deux sens.

(1) $\Rightarrow$ La proposition 14.186 dit que si $\int_\Omega |f| < \infty$, alors $\int_\Omega f^+ < \infty$ et $\int_\Omega f^- < \infty$. La condition (14.538) est alors remplie et l'intégrale $\int_\Omega f$ existe. Et comme les deux bouts sont finis, la définition (14.539) dit que $\int_\Omega f < \infty$.

(2) $\Leftarrow$ Vu que $\int_\Omega f$ existe, au moins un des deux parmi $\int_\Omega f^+$ et $\int_\Omega f^-$ est finie. Comme la différence entre les deux est finie (définition (14.539)), nous déduisons que les deux sont finies. La proposition 14.186 dit alors que $\int_\Omega |f| < \infty$.

□

PROPooDOMBooIVNzlZ

Proposition 14.189.
Soit une fonction f sur l'espace mesuré Ω. Si $\int_\Omega |f| < \infty$, alors

$$\left|\int_\Omega f\right| \leqslant \int_\Omega |f| < \infty. \tag{14.541}$$

Démonstration. La proposition 14.186 dit immédiatement que $\int_\Omega f^+ < \infty$ et $\int_\Omega f^- < \infty$. Calcul :

$$\begin{aligned}
\left|\int_\Omega f\right| &= \left|\int_\Omega f^+ - \int_\Omega f^-\right| && (14.542a)\\
&\leqslant \left|\int_\Omega f^+\right| + \left|\int_\Omega f^-\right| && \text{lem. 1.383(6)} \quad (14.542b)\\
&= \int_\Omega f^+ + \int_\Omega f^- && (14.542c)\\
&= \int_\Omega (f^+ + f^-) && (14.542d)\\
&= \int_\Omega |f|. && (14.542e)
\end{aligned}$$

□

Remarque 14.190.
Dans $\mathbb{R}^d$, quasiment toutes les fonctions et ensembles sont mesurables. En effet la construction d'ensembles non mesurables demande obligatoirement l'utilisation de l'axiome du choix ; de tels ensembles doivent être construits « exprès pour ». Il y a très peu de chances pour que vous tombiez sur un ensemble non mesurable de $\mathbb{R}^d$ sans que vous ne vous en rendiez compte.

Il y en a un en l'exemple 14.157.

Remarque 14.191.
« Mesurable » ne signifie pas « intégrable ». Par exemple la fonction

$$\begin{aligned}
f : \mathbb{R} &\to \bar{\mathbb{R}} \\
\omega &\mapsto \begin{cases} \frac{1}{\omega} & \text{si } \omega \neq 0 \\ \infty & \text{si } \omega = 0. \end{cases}
\end{aligned} \tag{14.543}$$

est mesurable, mais non intégrable.

14.9.5 Additivité de l'intégrale

LemPfHgal

Lemme 14.192.
Soit une fonction $f\colon \Omega \to \mathbb{R}$ telle que $|f(x)| \leqslant g(x)$ pour tout $x \in \Omega$. Si g est intégrable, alors f est intégrable.

Démonstration. La fonction g est manifestement à valeurs réelles positives. La proposition 14.176 nous dit alors que $|f|$ est intégrable. Ensuite c'est au tour de la proposition 14.186 de conclure à l'intégrabilité de f. □

PROPooFIYEooCpdmwZ

Proposition 14.193.
Soient deux fonctions intégrables sur $(S, \mathcal{F}, \mu)$ et à valeurs dans $\mathbb{C}$. Alors $f + g$ est intégrable et

$$\int_S (f+g)d\mu = \int_S f d\mu + \int_S g d\mu. \tag{14.544}$$

Démonstration. En plusieurs étapes suivant la généralité de f et g.

(1) **Si f et g sont étagées et positives** C'est le théorème 14.183(2) déjà prouvé.

(2) **Si f et g sont à valeurs positives** Le théorème fondamental d'approximation 14.118 nous permet de considérer des suites croissantes de fonctions étagées positives (f_k) et (g_k) qui vérifient $f_k \to f$ et $g_k \to g$.
Pour chaque k nous avons

$$\int_S (f_k + g_k)d\mu = \int_S f_k d\mu + \int_S g_k d\mu. \tag{14.545}$$

EQooXXYOooUhkOJL

De plus, la suite $k \mapsto f_k + g_k$ est une suite croissante de fonctions étagées positives convergeant vers $f + g$. Le théorème de la convergence monotone 14.178 nous permet donc de passer à la limité dans (14.545) et de permuter toutes les limites avec toutes les intégrales, des deux côtés.

(3) **f et g à valeurs réelles** Il faut diviser le domaine en de nombreuses régions suivant les signes de f, g et $f + g$.

□

Nous prouvons à présent l'additivité de l'intégrale pour des unions finie. Une version pour les unions dénombrables sera donnée dans les propositions 14.210 et 14.211.

PropOPSCooVpzaBt

Proposition 14.194 (σ-additivité finie)**.**
Soient des parties $A, B \subset \Omega$ disjointes dans $(\Omega, \mathcal{A}, \mu)$. Soit une fonction $f\colon \Omega \to \mathbb{R}$. Alors les intégrales $\int_A f$ et $\int_B f$ existent si et seulement si l'intégrale $\int_{A\cup B} f$. Dans ce cas nous avons :

$$\int_{A\cup B} f = \int_A f + \int_B f. \tag{14.546}$$

Démonstration. Vu que A et B sont disjoints, $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B$. En utilisant alors le lemme 14.177 et la proposition 14.193 nous avons le calcul

$$\int_{A\cup B} f = \int_\Omega f\mathbb{1}_{A\cup B} = \int_\Omega f\mathbb{1}_A + \int_\Omega f\mathbb{1}_B = \int_A f + \int_B f. \tag{14.547}$$

□

14.9.6 Fonctions à valeurs vectorielles (dimension finie)

Nous voulons intégrer des fonctions du type

$$f\colon \Omega \to V \tag{14.548}$$

où Ω et V sont des espaces vectoriels. Nous expliquons à présent plus précisément le cadre.

NORMooTQBIooBaScjt

14.195.
Nous considérons à présent un espace vectoriel normé $(V, \|.\|)$ de dimension finie, et un espace mesuré $(\Omega, \mathcal{A}, \mu)$.

Attention à ne pas confondre espace de départ et espace d'arrivée. Vu que V est un espace topologique, nous avons bien entendu les boréliens de V, et pour peut que nous ayons une mesure sur V (qui qui n'est pas compliqué à créer à partir de celle canonique de $\mathbb{R}^n$ et un isomorphisme), nous avons déjà une définition de $\int_V f d\mu$ lorsque $f\colon V \to \mathbb{R}$.

Ici nous nous proposons non d'intégrer $f\colon V \to \mathbb{R}$ mais bien $f\colon (\Omega, \mathcal{A}, \mu) \to V$ où V est un espace vectoriel normé.

Le lemme suivant est le point de départ pour définir les intégrales de fonctions à valeurs dans un espace vectoriel de dimension finie. Pour les fonctions à valeurs dans un espace de dimension infinie (par exemple de Banach), il existe des choses, mais c'est un peu plus compliqué.

LEMooCVHDooLJASAs

Lemme 14.196 ([1])**.**
Soit un espace vectoriel V réel de dimension finie, muni de la norme N. Soient une base $\{e_i\}$ de V, et une fonction $f\colon (\Omega, \mathcal{A}, \mu) \to V$ telle que la norme $N(f)\colon \Omega \to \mathbb{R}^+$ soit intégrable. Nous notons f_i les composantes de f : $f(x) = \sum_i f_i(x)e_i$.

Alors pour chaque i,

(1) la fonction $|f_i|\colon \Omega \to \mathbb{R}^+$ est intégrable,

(2) la fonction $f_i\colon \Omega \to \mathbb{R}$ est intégrable.

Démonstration. Si V était un espace muni d'un produit scalaire, et si la base $\{e_i\}$ était orthonormée, ce serait facile parce que la norme majore toutes les composantes. Hélas, ce n'est pas spécialement le cas. La base $\{e_i\}$ n'est pas spécialement orthonormée et même la norme N ne dérive pas spécialement d'un produit scalaire.

Nous allons utiliser l'équivalence de toutes les normes en dimension finie (théorème 10.49) pour nous ramener au cas d'une norme euclidienne.

Nous considérons sur V la norme « euclidienne » construite sur la base $\{e_i\}$: $\| \sum_i v_i e_i\| = \sum_i |v_i|^2$. Par équivalence des normes nous avons des nombres non nuls λ_1 et λ_2 tels que

$$N(v) \leqslant \lambda_1 \|v\|, \tag{14.549}$$

et

$$\|v\| \leqslant \lambda_2 N(v) \tag{14.550}$$

pour tout $v \in V$. Pour un i fixé nous avons alors les majorations

$$N\big(f_i(x)e_i\big) \leqslant \lambda_1 \|f_i(x)e_i\| \leqslant \lambda_1 \|f(x)\| \leqslant \lambda_1\lambda_2 N\big(f(x)\big). \tag{14.551}$$

En posant $N_i = N(e_i)$ nous avons la majoration [62]

$$|f_i(x)| \leqslant \frac{\lambda_1\lambda_2}{N(e_i)} N\big(f(x)\big). \tag{14.552}$$

L'application

$$\begin{aligned} |f_i|\colon \Omega &\to \mathbb{R}^+ \\ x &\mapsto |f_i(x)| \end{aligned} \tag{14.553}$$

est donc une fonction à valeurs réelles positives, majorée par une fonction intégrable (la fonction $x \mapsto N\big(f(x)\big)$). Elle est donc intégrable par le lemme 14.192.

La fonction f_i elle-même est alors intégrable par la proposition 14.186. □

62. Vous notez l'utilisation de la condition (3) de la définition 7.158 de la norme pour « convertir » la norme N en valeur absolue.

Notons que ce lemme est en réalité très simple si V est un espace vectoriel normé dont la norme découle d'un produit scalaire, comme c'est le cas pour $\mathbb{C}$. D'ailleurs, il ne faut pas se voiler la face : le cas d'intégrales de fonctions à valeurs dans $\mathbb{C}$ sera dans le Frido le cas de loin le plus courant. À ce propos, nous n'avons pas encore défini ce que nous voulons noter $\int_\Omega f d\mu$ lorsque f est une fonction à valeurs vectorielles. Comblons vite ce manque ...

PROPooOFSMooLhqOsc

Proposition-Définition 14.197 ([1])**.**
Soit une fonction $f\colon \Omega \to V$ où V est un espace vectoriel normé de dimension finie. Soit une base $\{e_i\}$ de V. Si la fonction $\|f\|\colon \Omega \to \mathbb{R}^+$ est intégrable, alors

(1) toutes les composantes $f_i\colon \Omega \to \mathbb{R}$ sont intégrables,

(2) le vecteur

$$\sum_i (\int_\Omega f_i) e_i \tag{14.554}$$

EQooQCKMooZChybq

ne dépend pas de la base choisie.

*Dans ce cas, la fonction f est dite **intégrable** et nous définissons*

$$\int_\Omega f d\mu = \sum_i (\int_\Omega f_i) e_i. \tag{14.555}$$

Démonstration. Le fait que les composantes soient intégrables est le lemme 14.196. Soient deux bases de V, $\{e_i\}$ et $\{s_\alpha\}$, liées conformément à (4.219) par la relation $s_\alpha = \sum_i Q_{i\alpha} e_i$ pour une certaine matrice inversible Q. Nous avons pour tout $x \in \Omega$:

$$f(x) = \sum_i f_i(x) e_i = \sum_\alpha f_\alpha(x) s_\alpha \tag{14.556}$$

avec $f_\alpha(x) = \sum_i f_i(x) Q^{-1}_{\alpha i}$ par la proposition 4.116.

Notons pour être pointilleux que les ensembles $\{e_i\}$ et $\{s_\alpha\}$ ne sont pas indexés par le même ensemble, de telle sorte que f_i ne peut pas être confondu avec f_α, même lorsqu'on attribue des valeurs à i et à α.

Comme combinaisons linéaires des fonctions f_i qui sont intégrables, les fonctions f_α sont intégrables (proposition 14.193). En écrivant $\int_\Omega f$ par rapport à la base $\{s_\alpha\}$ nous trouvons :

$$\sum_\alpha (\int f_\alpha) s_\alpha = \sum_\alpha \Big(\int \sum_i f_i(x) Q^{-1}_{\alpha i} dx \Big) \sum_j Q_{j\alpha} e_j \tag{14.557a}$$

$$= \sum_j \int \sum_{\alpha i} f_i(x) Q^{-1}_{\alpha i} Q_{j\alpha} dx e_j \tag{14.557b}$$

$$= \sum_j \int f_j(x) dx e_j \tag{14.557c}$$

$$= \sum_j (\int f_j) e_j \tag{14.557d}$$

où nous avons permuté des sommes finies et des intégrales des fonctions f_i, à valeurs dans $\mathbb{R}$ en vertu de la proposition 14.193 □

La proposition suivante est, pour les intégrales à valeurs vectorielles, analogue à la proposition 14.186.

PROPooNSCPooCMkrZl

Proposition 14.198.
Soit une fonction mesurable $f\colon \Omega \to (V, \|.\|)$. Soit une base $\{e_i\}$ de V et la décomposition $f = \sum_i f_i e_i$.

Nous avons équivalence entre

ITEMooYLADooCXKEds

(1) $\int_\Omega \|f\| < \infty$

ITEMooLEYEooQTGwmt

(2) $\int_\Omega |f_i| < \infty$

ITEMooYDDAooMKwDIR

(3) $\int_\Omega f_i^+ < \infty$ *et* $\int_\Omega f_i^- < \infty$.

Démonstration. L'équivalence entre les points (2) et (3) est la proposition 14.186. Nous démontrons l'équivalence entre (1) et (2).

Vu que toutes les normes sont équivalentes sur V, nous considérons en particulier la norme associée à la base $\{e_i\}$ donnée par

$$N(x) = \sum_i |x_i|. \tag{14.558}$$

Il existe des constantes λ_1 et λ_2 telles que

$$\lambda_1\big(\sum_i |f_i(x)|\big) \leqslant \|f(x)\| \leqslant \lambda_2\big(\sum_i |f_i(x)|\big) \tag{14.559}$$

pour tout $x \in \Omega$.

La première inégalité dit que si $\int_\Omega \|f\| < \infty$, alors $\lambda_1\big(\sum_i \int_\Omega |f - i|\big) < \infty$. Et vu que chacun des termes est positif, ils sont tous finis.

La seconde inégalité donne l'implication réciproque. □

14.9.7 Quelques propriétés

Le lemme suivant nous aide à détecter des fonctions presque partout nulles.

Lemfobnwt

Lemme 14.199.
Soit f une fonction mesurable positive ou nulle telle que

$$\int_\Omega f d\mu = 0. \tag{14.560}$$

Alors $f = 0$ μ-presque partout.

Démonstration. L'ensemble des points $x \in \Omega$ tels que $f(x) \neq 0$ peut s'écrire comme une union dénombrable disjointe :

$$\{x \in \Omega \text{ tel que } f(x) \neq 0\} = \bigcup_{i=0}^{\infty} E_i \tag{14.561}$$

avec

$$E_0 = \{x \in \Omega \text{ tel que } f(x) > 1\} \tag{14.562a}$$

$$E_i = \{x \in \Omega \text{ tel que } \frac{1}{i+1} \leqslant f(x) < \frac{1}{i}\}. \tag{14.562b}$$

Si un des ensembles E_i est de mesure non nulle, alors nous pouvons considérer la fonction simple $h(x) = \frac{1}{i+1}\mathbb{1}_{E_i}$ dont l'intégrale sur Ω est strictement positive. Par conséquent le supremum de la définition (14.468) est strictement positif.

Nous savons donc que $\mu(E_i) = 0$ pour tout i. Étant donné que la mesure d'une union disjointe dénombrable est égale à la somme des mesures, nous avons

$$\mu\{x \in \Omega \text{ tel que } f(x) \neq 0\} = 0, \tag{14.563}$$

ce qui signifie que f est nulle μ-presque partout. □

CorjLYiSm

Corolaire 14.200.
Soit f une fonction mesurable sur l'espace mesuré $(\Omega, \mathcal{A}, \mu)$ telle que

$$\int_\Omega f \mathbb{1}_{f>0} d\mu = 0. \tag{14.564}$$

Alors $f \leqslant 0$ presque partout.

Démonstration. Nous avons l'égalité d'ensembles

$$\{f\mathbb{1}_{f>0} \neq 0\} = \{\mathbb{1}_{f>0} \neq 0\}. \tag{14.565}$$

Mais lemme 14.199 implique que $f\mathbb{1}_{f>0}$ est nulle presque partout, c'est-à-dire que la mesure de l'ensemble du membre de gauche est nulle par conséquent

$$\mu\{\mathbb{1}_{f>0} \neq 0\} = 0. \tag{14.566}$$

Cela signifie que la fonction f est presque partout négative ou nulle. □

LEMooLXVGooUDuQzc

Lemme 14.201 ([1])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ ainsi qu'une partie $X \in \mathcal{A}$ telle que $\mu(X) > 0$. Si $f : \Omega \to \mathbb{R}$ vérifie $f > 0$ sur X, alors

$$\int_X f > 0. \tag{14.567}$$

Démonstration. Nous posons, pour $n \in \mathbb{N}$,

$$A_n = \{x \in x \text{ tel que } \frac{1}{n} \geqslant f(x) > \frac{1}{n+1}\}, \tag{14.568}$$

$$A_0 = \{x \in X \text{ tel que } f(x) > 1\}. \tag{14.569}$$

Nous avons l'union disjointe $X = \bigcup_{i=0}^{\infty} A_i$, et donc $\mu(X) = \sum_{i=0}^{\infty} \mu(A_i)$. Nous en déduisons qu'il existe $i \in \mathbb{N}$ tel que $\mu(A_i) > 0$. Nous posons alors

$$\begin{aligned} \psi : X &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{si } x \notin A_i \\ 1/2i & \text{si } x \in A_i. \end{cases} \end{aligned} \tag{14.570}$$

Cette fonction ψ est étagée et vérifie $0 \leqslant \psi \leqslant f$. Comme l'intégrale de f sur X est le supremum des intégrales de ce type de fonctions, nous avons

$$\int_X f d\mu \geqslant \int_X \psi d\mu = \frac{1}{2i}\mu(A_i) > 0. \tag{14.571}$$

□

¡¡ Avertissement/question à la lectrice !! 14.202
Je n'ai pas vérifié le lemme suivant. Écrivez-moi si vous avez une démonstration.

LEMooVZXFooZAQWKY

Lemme 14.203.
Si $f \in L^1(\Omega, \mathcal{A}, \mu)$ et si $\int_A f d\mu = 0$, alors $\mu(A) = 0$

14.9.8 Permuter limite et intégrale

14.9.8.1 Convergence uniforme

PropbhKnth

Proposition 14.204 (Permuter limite et intégrale)**.**
Soit $f_n \to f$ uniformément sur un ensemble mesuré A de mesure finie. Alors si les fonctions f_n et f sont intégrables sur A, nous avons

$$\lim_{n\to\infty} \int_A f_n = \int_A \lim_{n\to\infty} f_n. \tag{14.572}$$

Démonstration. Notons f la limite de la suite (f_n). Pour tout n nous avons les majorations

$$\left|\int_A f_n d\mu - \int_A f d\mu\right| \leqslant \int_A |f_n - f| d\mu \tag{14.573a}$$

$$\leqslant \int_A \|f_n - f\|_\infty d\mu \tag{14.573b}$$

$$= \mu(A)\|f_n - f\|_\infty \tag{14.573c}$$

où $\mu(A)$ est la mesure de A. Le résultat découle maintenant du fait que $\|f_n - f\|_\infty \to 0$. □

Il existe un résultat considérablement plus intéressant que cette proposition. En effet, l'intégrabilité de f n'est pas nécessaire. Cette hypothèse peut être remplacée soit par l'uniforme convergence de la suite (théorème 14.205), soit par le fait que les normes des f_n sont uniformément bornées (théorème de la convergence dominée de Lebesgue 14.207).

ThoUnifCvIntRiem

Théorème 14.205 ([459])**.**
La limite uniforme d'une suite de fonctions intégrables sur un borné est intégrable, et on peut permuter la limite et l'intégrale.

Plus précisément, soit A un ensemble de μ-mesure finie et $f_n : A \to \mathbb{R}$ des fonctions intégrables sur A. Si la limite $f_n \to f$ est uniforme, alors f est intégrable sur A et nous pouvons inverser la limite et l'intégrale :

$$\lim_{n\to\infty} \int_A f_n = \int_A \lim_{n\to\infty} f_n. \tag{14.574}$$

Démonstration. Soit $\epsilon > 0$ et n tel que $\|f_n - f\|_\infty \leqslant \epsilon$ (ici la norme uniforme est prise sur A). Étant donné que f_n est intégrable sur A, il existe une fonction simple φ_n qui minore f_n telle que

$$\left|\int_A \varphi_n - \int_A f_n\right| < \epsilon. \tag{14.575}$$

La fonction $\varphi_n + \epsilon$ est une fonction simple qui majore la fonction f. Si ψ est une fonction simple qui minore f, alors

$$\int_A \psi \leqslant \int_A \varphi_n + \epsilon \leqslant \int_A f_n + \epsilon\mu(A). \tag{14.576}$$

Par conséquent le supremum qui définit $\int_A f$ existe, ce qui montre que f est intégrable. Le fait qu'on puisse inverser la limite et l'intégrale est maintenant une conséquence de la proposition 14.204. □

Remarque 14.206.
L'hypothèse sur le fait que A soit de mesure finie est importante. Il n'est pas vrai qu'une suite uniformément convergente de fonctions intégrables est intégrables. En effet nous avons par exemple la suite

$$f_n(x) = \begin{cases} 1/x & \text{si } x < n \\ 0 & \text{sinon} \end{cases} \tag{14.577}$$

qui converge uniformément vers $f(x) = 1/x$ sur $A = [1, \infty[$. Le limite n'est cependant guerre intégrable sur A.

14.9.8.2 Convergence dominée de Lebesgue

ThoConvDomLebVdhsTf

Théorème 14.207 (Convergence dominée de Lebesgue)**.**
Soit une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ sur $(\Omega, \mathcal{A}, \mu)$ à valeurs dans $\mathbb{C}$ ou $\mathbb{R}$. Nous supposons que

(1) Pour chaque n nous avons $f_n \in L^1(\Omega, \mathcal{A}, \mu)$,

(2) $f_n \to f$ simplement presque partout sur Ω,

(3) Il existe une fonction $g \in L^1(\Omega)$ telle que

$$|f_n(x)| \leqslant g(x) \tag{14.578}$$

pour presque[63] *tout $x \in \Omega$ et pour tout $n \in \mathbb{N}$.*

Alors

(1) f est intégrable[64]*,*

(2) $\lim_{n\to\infty} \int_\Omega f_n = \int_\Omega f$,

(3) $\lim_{n\to\infty} \int_\Omega |f_n - f| = 0$.

Démonstration. La fonction limite f est intégrable parce que $|f| \leqslant g$ et g est intégrable[65]. Par hypothèse nous avons

$$-g(x) \leqslant f_n(x) \leqslant g(x). \tag{14.579}$$

En particulier la fonction $g_n = f_n + g$ est positive et mesurable si bien que le lemme de Fatou 14.182 implique

$$\int_\Omega \liminf g_n \leqslant \liminf \int_\Omega g_n. \tag{14.580}$$

Évidemment nous avons $\liminf g_n = f + g$, de telle sorte que

$$\int f + \int g \leqslant \liminf \int g_n = \liminf \int f_n + \int g, \tag{14.581}$$

et le nombre $\int g$ étant fini, nous pouvons le retrancher des deux côtés de l'inégalité :

$$\int f \leqslant \liminf \int f_n. \tag{14.582}$$

EQooXGKRooDqUDlG

Afin d'obtenir une minoration de $\int f$ nous récrivons (14.582) pour la suite de fonctions $k_n = -f_n \to k = -f$:

$$\int k \leqslant \liminf \int k_n. \tag{14.583}$$

En utilisant le lemme 9.47, nous obtenons

$$-\int f \leqslant \liminf \int -f_n = -\limsup \int f_n, \tag{14.584}$$

et donc

$$\limsup \int f_n \leqslant \int f. \tag{14.585}$$

En combinant avec (14.582), nous avons la suite d'inégalités

$$\limsup \int f_n \leqslant \int f \leqslant \liminf \int f_n. \tag{14.586}$$

La limite supérieure étant plus grande ou égale à la limite inférieure, les trois quantités sont égales.

Nous prouvons maintenant le troisième point. Soit la suite de fonctions

$$h_n(x) = |f_n(x) - f(x)| \tag{14.587}$$

qui tend ponctuellement vers zéro. De plus

$$h_n(x) \leqslant |f_n(x)| + |f(x)| \leqslant 2g(x), \tag{14.588}$$

ce qui prouve que les h_n majorés par une fonction intégrable. Donc

$$\lim_{n\to\infty} \int_\Omega |f_n - f| = \lim_{n\to\infty} \int_\Omega h_n(x)dx = \int_\Omega \lim_{n\to\infty} |f_n(x) - f(x)| = 0 \tag{14.589}$$

□

63. Si il n'y avait pas le « presque » ici, ce théorème serait à peu près inutilisable en probabilité ou en théorie des espaces L^p, comme dans la démonstration du théorème de Fischer-Riesz 27.45 par exemple.

64. C'est à dire que $\int_\Omega |f| < \infty$, voir le vocabulaire en 14.187.

65. Par le lemme 14.192

Remarque 14.208.
Lorsque nous travaillons sur des problèmes de probabilités, la fonction g peut être une constante parce que les constantes sont intégrables sur un espace de probabilité.

CorCvAbsNormwEZdRc

Corolaire 14.209.
Soit $(a_i)_{i\in\mathbb{N}}$ une suite numérique absolument convergente. Alors elle est convergente. Il en est de même pour les séries de fonctions si on considère la convergence ponctuelle.

Démonstration. L'hypothèse est la convergence de l'intégrale $\int_{\mathbb{N}} |a_i| dm(i)$ où dm est la mesure de comptage. Étant donné que $|a_i| \leqslant |a_i|$, la fonction a_i (fonction de i) peut jouer le rôle de g dans le théorème de la convergence dominée de Lebesgue (théorème 14.207). □

14.9.9 Additivité de l'intégrale de Lebesgque

Les propositions 14.210 et 14.211 démontrent la même chose. La différence est la méthode utilisée pour permuter une somme et une intégrale. Dans le premier cas, nous utilisons la convergence monotone (et sommes obligés de séparer le cas où f est positive), alors que dans le second cas, nous utilisons la convergence dominée de Lebesgue, et nous ne devons pas faire de séparation d'après la positivité de f.

PROPooTFOAooJBwmCV

Proposition 14.210 (σ-additivité dénombrable[1])**.**
Si $(A_i)_{i\in\mathbb{N}}$ sont des parties mesurables disjointes de $(\Omega, \mathcal{A}, \mu)$ et si $f\colon \Omega \to \mathbb{R}$ est intégrable sur $\bigcup_{i=0}^{\infty} A_i$ alors les intégrales $\int_{A_i} f d\mu$ existent et

$$\int_{\bigcup_i A_i} f d\mu = \sum_{i=0}^{\infty} \int_{A_i} f d\mu. \tag{14.590}$$

Démonstration. En deux cas d'après la positivité de f.

(1) **Si f est positive** Nous posons $f_N = f\mathbb{1}_{\bigcup_{i=0}^{N} A_i}$. Cette suite de fonctions vérifie la limite

$$\lim_{N\to\infty} f_N = f\mathbb{1}_{\bigcup_{i=0}^{\infty} A_i}. \tag{14.591}$$

De plus, pour chaque N nous avons

$$\int_\Omega f_N = \int_\Omega f\mathbb{1}_{\bigcup_i A_i} = \int_{\bigcup_i A_i} f = \sum_{i=0}^{N} \int_{A_i} f \tag{14.592}$$

Justifications :

— La proposition 14.177 pour l'introduction de la fonction caractéristique de $\bigcup_i A_i$

— La proposition 14.194 qui traite le cas de la sous-additivité finie pour la dernière égalité.

La suite $(f_N)_{n\in\mathbb{N}}$ est une suite croissante de fonctions mesurables[66] et positives. Donc le théorème de la convergence monotone 14.178 s'applique et

$$\sum_{i=0}^{\infty} \int_{A_i} f = \lim_{N\to\infty} \sum_{i=0}^{N} \int_{A_i} f = \lim_{N\to\infty} \int_{\bigcup_{i=0}^{N} A_i} f = \lim_{N\to\infty} \int_\Omega f_N \tag{14.593a}$$

$$= \int_\Omega \lim_{N\to\infty} f_N = \int_\Omega f\mathbb{1}_{\bigcup_{i=0}^{\infty} A_i} = \int_{\bigcup_i A_i} f d\mu. \tag{14.593b}$$

66. La fonction f elle-même est mesurable ; c'est inclus dans la définition de « intégrable ».

(2) **Si f est à valeurs réelles** Si f est à valeurs dans $\mathbb{R}$, alors $f = f_+ - f_-$ où f_+ et f_- sont intégrables. Nous avons alors

$$\int_{\bigcup_{i=0}^{\infty} A_i} f d\mu = \int_{\bigcup_i A_i} f_+ - \int_{\bigcup_i A_i} f_- \tag{14.594a}$$
$$= \sum_{k=0}^{\infty} \int_{A_k} f_+ - \sum_{k=0}^{\infty} \int_{A_k} f_- \tag{14.594b}$$
$$= \sum_{k=0}^{\infty} \left(\int_{A_k} f_+ - \int_{A_k} f_- \right) \tag{14.594c}$$
SUBEQooMTZPooLqMHKP
$$= \sum_{k=0}^{\infty} \int_{A_k} f. \tag{14.594d}$$
SUBEQooVZNMooRmFoLq

Justifications :

— Pour (14.594c), c'est l'associativité de la somme, proposition 10.102.

— Pour (14.594d), c'est la proposition 14.193.

□

PROPooDWYNooWKJmEV

Proposition 14.211 (σ-additivité[460])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Nous considérons des parties disjointes $\{A_i\}_{i\in\mathbb{N}}$ de Ω telles que $\bigcup_{k=0}^{\infty} A_k = \Omega$. Si $f \in L^1(\Omega)$, alors

$$\int_{\Omega} f d\mu = \sum_{k=0}^{\infty} \int_{A_k} f d\mu. \tag{14.595}$$

Démonstration. Nous posons $\Omega_n = \bigcup_{k=0}^{n} A_k$ ainsi que $f_n = f \mathbb{1}_{\Omega_n}$. Pour chaque $N \in \mathbb{N}$ nous avons
EQSooBREOooWzviSK

$$\sum_{k=0}^{N} \int_{A_k} f d\mu = \int_{\bigcup_{k=0}^{N} A_k} f \tag{14.596a}$$
EQooCVVVooTIINmg
$$= \int_{\Omega_N} f \tag{14.596b}$$
$$= \int_{\Omega} f_N. \tag{14.596c}$$
SUBEQooJZLQooKJQoes

Justifications :

— Pour (14.596a), c'est la proposition 14.194 qui traite du cas de sommes finies.

— Pour (14.596c) c'est la proposition 14.210.

L'idée est maintenant de passer à la limite des deux côtés de (14.596). Voici le raisonnement :

— Nous montrons qu'à droite, la limite existe et vaut $\int_\Omega f d\mu$.

— Le fait que la limite du membre de droite existe implique l'existence de la limite du membre de gauche.

— La limite du membre de gauche vaut $\sum_{k=0}^{\infty} \int_{A_k} f d\mu$.

La limite du membre de droite s'établi avec le théorème de la convergence dominée de Lebesgue 14.205.

— Nous avons convergence simple $f_n \to f$ parce que $\bigcup_{n=0}^{\infty} A_i = \Omega$.

— La fonction $g = |f|$ est intégrable sur Ω parce que $f \in L^1(\Omega)$ par hypothèse.

— Pour tout $n \in \mathbb{N}$ et pour tout $x \in \Omega$ nous avons $|f_n(x)| \leqslant g(x)$ parce que $|f_n(x)|$ est soit égal à $g(x)$ soit égal à zéro suivant que $x \in \Omega_n$ ou non.

Donc le théorème de la convergence dominée est applicable. La limite du membre de droite de (14.596) existe et vaut :

$$\lim_{N\to\infty}\int_\Omega f_N=\int_\Omega f. \tag{14.597}$$

Nous pouvons alors prendre aussi la limite du membre de gauche dans (14.596) et obtenir le résultat attendu. □

14.9.10 Produit d'une mesure par une fonction (mesure à densité)

PropooVXPMooGSkyBo

Proposition-Définition 14.212 (Produit d'une mesure par une fonction[1, 461])**.**
Soit un espace mesuré $(\Omega,\mathcal{F},\mu)$ et une fonction mesurable positive $w\colon\Omega\to\bar{\mathbb{R}}^+$. Alors la formule

$$(w\cdot\mu)(A)=\int_A w d\mu \tag{14.598}$$

*pour tout $A\in\mathcal{F}$ définit une mesure positive sur $(\Omega,\mathcal{F})$ appelée **produit** de la mesure μ par la fonction w. La fonction w est la **densité** de la mesure $w\cdot\mu$ par rapport à la mesure μ.*

Démonstration. D'abord $(w\cdot\mu)(\varnothing)=0$ parce que le lemme 14.175 donne

$$(w\cdot\mu)(\varnothing)=\int_\Omega w\mathbb{1}_\varnothing d\mu=\int_\Omega 0d\mu=0\times\mu(\Omega)=0 \tag{14.599}$$

où nous avons (éventuellement) utilisé deux fois la convention $0\times\infty=0$.

Ensuite si les ensembles A_i sont des éléments deux à deux disjoints de $\mathcal{F}$ alors nous avons $\mathbb{1}_{\bigcup_{i=1}^\infty A_i}=\sum_{i=1}^\infty\mathbb{1}_{A_i}$, et donc

$$(w\cdot\mu)(\bigcup_{i=0}^\infty A_i)=\int_{\bigcup_{i=0}^\infty A_i}wd\mu=\sum_{i=0}^\infty\int_{A_i}wd\mu=\sum_{i=0}^\infty(w\cdot\mu)(A_i). \tag{14.600}$$

où nous avons utilisé la σ-additivité dénombrable de l'intégrale de la proposition 14.210. □

En particulier nous parlons souvent de mesure à densité par rapport à la mesure de Lebesgue. C'est alors la construction suivante.

DEFooJNQYooKUcezo

Définition 14.213.
*Si μ est une mesure sur $\mathbb{R}^d$, une fonction $f\colon\mathbb{R}^d\to\mathbb{R}$ est une **densité** pour μ si pour tout $A\subset\mathbb{R}^d$ nous avons*

$$\mu(A)=\int_A f(x)dx \tag{14.601}$$

où dx est la mesure de Lebesgue.

*Si la mesure μ admet une densité, nous disons que c'est une **mesure à densité** par rapport à la mesure de Lebesgue. Nous disons aussi que μ est **absolument continue**.*

Exemple 14.214.
Toutes les mesures n'admettent pas de densité. Par exemple la mesure de Dirac donnée par

$$\nu(A)=\begin{cases}1 & \text{si } 0\in A\\ 0 & \text{sinon.}\end{cases} \tag{14.602}$$

EQooDMFCooVEManF

n'a pas de densité par rapport à la mesure de Lebesgue. △

La mesure ν de l'exemple 14.602 admet, au sens des distributions, la mesure de Dirac δ comme densité, mais c'est une autre histoire qui vous sera contée une autre fois.

PropooJMWAooDzfpmB

Proposition 14.215 ([461])**.**
Soit une fonction mesurable $w\colon(S,\mathcal{F},\mu)\to\bar{\mathbb{R}}^+$.

(1) Si $f\colon S\to\bar{\mathbb{R}}^+$ est mesurable, alors $f\cdot(w\cdot\mu)=(fw)\cdot\mu$.
(2) Si $f\colon S\to\bar{\mathbb{R}}$ ou $\mathbb{C}$ est mesurable, elle est $w\cdot\mu$-intégrable si et seulement si fw est μ-intégrable. Dans ce cas, nous avons encore $f\cdot(w\cdot\mu)=(fw)\cdot\mu$.

Attention : dans le cas où f est à valeurs dans $\mathbb{C}$, alors il faut que w soit à valeurs finies dans $\mathbb{R}$ parce que nous n'avons pas défini $\infty\times z$ lorsque $z\in\mathbb{C}$.

Démonstration. Nous commençons par prouver le résultat pour la fonction caractéristique de l'ensemble mesurable A. Nous avons : $\mathbb{1}_A\cdot(w\cdot\mu)(B)=\int_B\mathbb{1}_A d(w\cdot\mu)$. Mais par définition, l'intégrale d'une fonction indicatrice est la mesure de l'ensemble indiqué. En passant sur le fait que $\mathbb{1}_A\mathbb{1}_B=\mathbb{1}_{A\cap B}$,

$$\int_B\mathbb{1}_A d(w\cdot\mu)=(w\cdot\mu)(A\cap B)=\int_S\mathbb{1}_{A\cap B}wd\mu=\int_S\mathbb{1}_A\mathbb{1}_B wd\mu=\int_B\mathbb{1}_A wd\mu=(\mathbb{1}_A w)\cdot\mu(B).\tag{14.603}$$

Supposons maintenant que f soit une fonction étagées qui s'écrit $f=\sum_k a_k\mathbb{1}_{A_k}$ où les A_k sont des ensembles mesurables disjoints. Alors le calcul est le suivant, en utilisant le fait que sur A_k, on a $a_k=f(x)$:

$$\begin{aligned}
f\cdot(g\cdot\mu)B&=\int_B fd(g\cdot\mu)&(14.604a)\\
&=\sum_k a_k(g\cdot\mu)(A_k\cap B)&(14.604b)\\
&=\sum_k a_k\int_{A_k\cap B}gd\mu&(14.604c)\\
&=\int_{A_k\cap B}f(x)g(x)d\mu(x)&(14.604d)\\
&=\sum_k(fg\cdot\mu)(A_k\cap B)&(14.604e)\\
&=(fg\cdot\mu)(B)&(14.604f)
\end{aligned}$$

parce que les $A_k\cap B$ forment une partition de l'ensemble B (voir le point (2) de la définition 14.18).

Si $f\colon S\to\bar{\mathbb{R}}^+$ est mesurable, le théorème 14.118 donne une suite croissante f_n de fonctions étagées positives convergeant (ponctuellement) vers f. Vu que la fonction w est positive, nous avons aussi la limite positive et croissante $wf_n\to wf$. Ainsi l'utilisation du théorème de la convergence monotone est justifié dans le calcul suivant :

$$\int_S fd(w\cdot\mu)=\lim_{n\to\infty}\int_S f_n d(w\cdot\mu)=\lim_{n\to\infty}\int_S(wf_n)d\mu=\int_S wfd\mu.\tag{14.605}$$

Nous passons maintenant au cas général où f est une fonction à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$ (avec w finie dans ce dernier cas). Nous avons la chaine d'équivalences

$\Leftrightarrow$ f est $(w\cdot\mu)$ intégrable
$\Leftrightarrow$ $|f|$ est $(w\cdot\mu)$-intégrable
$\Leftrightarrow$ $|f|w$ est μ-intégrable
$\Leftrightarrow$ $|fw|$ est μ-intégrable.

Si cela est le cas, la formule se démontre en se ramenant au cas déjà prouvé des fonctions positives en utilisant les $(fw)^+=f^+w$, $(fw)^-=f^-w$ etc. □

14.9.11 Mesure et topologie

EXooKQDRooVMWaEC

Exemple 14.216 (Un compact n'est pas toujours de mesure finie)**.**
Soit l'espace mesurable $(\mathbb{R},\mathcal{B}\text{or}(\mathbb{R}))$ réel avec ses boréliens et la fonction

$$\begin{aligned}w\colon\big(\mathbb{R},\mathcal{B}\text{or}(\mathbb{R})\big)&\to\big(\bar{\mathbb{R}},\mathcal{B}\text{or}(\bar{\mathbb{R}})\big)\\ x&\mapsto\begin{cases}\frac{1}{|x|}&\text{si } x\neq 0\\ +\infty&\text{si } x=0.\end{cases}\end{aligned}\tag{14.606}$$

Essayons d'étudier la mesure de densité w par rapport à la mesure de Lebesgue.

(1) **w est mesurable** Soit un borélien B de $\bar{\mathbb{R}}$. Si B ne contient pas ∞ alors $w^{-1}(B)$ est un borélien de $\mathbb{R}$ par continuité de l'application restreinte $w\colon \mathbb{R}\backslash\{0\} \to \mathbb{R}$. Ici nous avons par exemple appliqué la proposition 14.125 à chacun des deux intervalles $]-\infty,0[$ et $]0,\infty[$. Si $+\infty \in B$ alors

$$w^{-1}(B) = w^{-1}\big(B\backslash\{\infty\}\big) \cup w^{-1}(\{\infty\}) = w^{-1}\big(B\backslash\{\infty\}\big) \cup \{0\}, \tag{14.607}$$

qui est borélien par union de boréliens.

(2) **Mesure produit** La proposition 14.212 nous assure alors qu'en posant [67]

$$\mu(B) = \int_B \frac{1}{|x|} d\lambda(x) \tag{14.608}$$

où λ est la mesure de Lebesgue, nous avons une mesure.

(3) **Mesure du singleton**
Pour avoir les idées claires, nous pouvons nous demander la mesure $\mu\big(\{0\}\big)$. Nous cela nous devons calculer

$$\int_{\{0\}} \frac{1}{|x|} d\lambda(x) = \int_{\{0\}} w(x) d\lambda(x) \tag{14.609}$$

où là, l'abus de notation n'est plus possible. Mais quelle que soit la fonction étagée $h = \sum_i \alpha_i \mathbb{1}_{A_i}$ considérée,

$$\int_{\{0\}} h(x) d\lambda(x) = \sum_i \alpha_i \lambda\big(A_i \cap \{0\}\big) = 0. \tag{14.610}$$

Attention : ceci n'a rien de particulier à la fonction $x \mapsto 1/|x|$. Lorsqu'une mesure a une densité par rapport à Lebesgue, la mesure d'un singleton sera toujours nulle.

(4) **Mesure de la boule compacte**
Il n'en reste pas moins que $\mu\big([-1,1]\big) = \infty$.

△

14.217.
En réalité, il n'y a pas de liens forts entre mesure et topologie. Un espace topologique est une chose, et y mettre une mesure en est une autre. Bien entendu, une topologie étant donnée, nous pouvons considérer la tribu des boréliens et y mettre une mesure un peu quelconque. Il n'y a pas de choix canonique.

Notons que même dans l'exemple de compact de mesure infinie 14.216, la mesure introduite n'est pas sans lien avec la topologie de $\mathbb{R}$. En effet pour avoir une mesure à densité par rapport à Lebesgue, nous avons dû prendre une application mesurable par rapport à la tribu des boréliens, laquelle est éminemment liée à la topologie. Il y a donc parfaitement moyen de construire des espaces mesurés tenant compte de la topologie, et ayant des propriétés qui ne sont pas celle attendues.

Quand les choses sont faciles, ça se passe bien. La proposition suivante dit qu'une fonction continue sur un compact y est intégrable ; sauf que pour dire cela de façon précise, il faut un peu bosser parce qu'il y a de écueils à éviter, tels que l'exemple 14.216.

PROPooKFRSooANzglT

Proposition 14.218 ([1])**.**
Soit un espace mesuré $(K, \mathcal{A}, \mu)$ et une fonction $f\colon K \to \mathbb{R}$. Nous supposons pas mal de trucs techniques :

(1) La mesure est finie : $\mu(K) < \infty$.

(2) L'ensemble K est par ailleurs un espace topologique compact [68]*.*

67. Avec un mini abus de notation : si $0 \in B$, cette notation n'est pas tout à fait correcte.
68. Nous ne prétendons pas que la tribu $\mathcal{A}$ soit liée à la topologie de K.

ITEMooBKYHooWnxUGL

(3) La fonction f est continue pour les topologies de K et de $\mathbb{R}$.

ITEMooJCNUooJzIlKI

(4) La fonction f est mesurable pour la tribu $\mathcal{A}$ de K et la tribu des boréliens de $\mathbb{R}$.

Alors f est intégrable sur K et $\int_K |f| < \infty$.

L'hypothèse (4) ne se déduit pas nécessairement de l'hypothèse (3). Dans les cas usuels, nous avons bien « continue implique mesurable », mais si $\mathcal{A}$ n'a aucun rapport avec la topologie . . .hum . . .

Démonstration. Si nous écrivons $f(x) = f^+(x) - f^-(x)$ avec f^+ et f^- prenant des valeurs positives ou nulles[462], en vertu de la proposition 14.186, si nous devons prouver séparément $\int_K f^+ < \infty$ et $\int_K f^- < \infty$. Nous allons donc prouver cette proposition en plusieurs étapes.

(1) **Si f est positive** La fonction f est continue sur K qui est compact (même en tant qu'espace topologique en soi ; il n'est pas nécessaire d'être compact *dans* quelque chose), donc elle a un maximum par le théorème 7.146 nommons M ce maximum. Donc $f\colon K \to [0, M]$. De plus la mesure μ sur K est finie et vérifie disons $\mu(K) = m$.
Soit une fonction étagée $h\colon K \to \mathbb{R}^+$ majorée par f. Nous notons

$$h(x) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(x) \tag{14.611}$$

où les A_i sont des éléments de $\mathcal{A}$. Vu que $0 \leqslant h(x) \leqslant f(x) \leqslant M$, nous avons[69]

$$\int_K h = \sum_{i=1}^{n} \alpha_i \mu(K \cap A_i) \leqslant \sum_{i=1}^{n} M\mu(K \cap A_i) \leqslant M\mu(K) = Mm \tag{14.612}$$

parce que les A_i sont disjoints et vérifient $\bigcup_i A_i = K$ (lemme 14.114).
Donc tous les éléments de l'ensemble sur lequel nous prenons le supremum dans la définition (14.468) sont contenus dans $[0, Mm]$. Le supremum est donc dans $[0, Mm]$ et est alors strictement plus petit que l'infini.

(2) **Si f est positive ou négative** Nous appliquons la première partie séparément à f^+ et f^-. Et nous avons alors que f est intégrable et

$$\int_K |f| = \int_K f^+ + \int_K f^- < \infty. \tag{14.613}$$

□

THOooVADUooLiRfGK

Théorème 14.219 ([461]).
Soient deux espaces mesurables $(S_1, \mathcal{F}_1)$ et $(S_2, \mathcal{F}_2)$ ainsi qu'une application mesurable $\varphi\colon S_1 \to S_2$. Soit encore μ, une mesure positive sur $(S_1, \mathcal{F}_1)$.

Si $f\colon S_2 \to \bar{\mathbb{R}}$ ou $\mathbb{C}$ est mesurable alors,

ItemooKMBIooZpHJSS

(1) f est $\varphi(\mu)$-intégrable si et seulement si $f \circ \varphi$ est μ-intégrable.

ItemooLAPYooUreDEl

(2) dans le cas où f est $\varphi(\mu)$-intégrable, nous avons

$$\int_{S_2} f d\big(\varphi(\mu)\big) = \int_{S_1} (f \circ \varphi) d\mu. \tag{14.614}$$

EqooSOHXooXShdoy

Démonstration. L'intégrabilité est la définition 14.186, et demande que $|f|$ soit intégrable. L'égalité (14.614) a un sens si les deux membres sont infinis. Tant que les fonctions considérées sont positives, le point (1) est immédiat. Ce n'est qu'au moment où les fonctions considérées deviennent à valeurs dans $\mathbb{C}$ ou $\mathbb{R}$ que l'intégrabilité de $|f|$ commence à jouer parce qu'il faut que f^+ et f^- soient séparément intégrables.

Nous allons prouver la formule (14.614) pour des fonctions de plus en plus générales. Pour la suite nous notons $\mu' = \varphi(\mu)$.

69. Définition (14.467).

(1) **Pour $f = \mathbb{1}_B$, B mesurable** Soit $B \in \mathcal{F}_2$. Nous avons $\mathbb{1}_B \circ \varphi = \mathbb{1}_{\varphi^{-1}(B)}$. Donc en utilisant le lemme 14.175 nous avons

$$\int_{S_2} \mathbb{1}_B d\mu' = \mu'(B) = \mu\big(\varphi^{-1}(B)\big) = \int_{S_1} \mathbb{1}_{\varphi^{-1}(B)} d\mu = \int_{S_1} (\mathbb{1}_B \circ \varphi) d\mu. \tag{14.615}$$

(2) **f est étagée positive**
La fonction f peut être écrite sous la forme

$$f = \sum_{k=1}^{n} a_k \mathbb{1}_{B_k} \tag{14.616}$$

avec $B_k \in \mathcal{F}_2$ et $a_k \in \mathbb{R}^+$. Nous avons alors, en utilisant la sous-additivité de l'intégrale du théorème 14.183(3),

$$\int_{S_2} f d\mu' = \sum_k a_k \int_{S_2} \mathbb{1}_{B_k} d\mu' \tag{14.617a}$$

$$= \sum_k a_k \int_{S_1} (\mathbb{1}_{B_k} \circ \varphi) d\mu \tag{14.617b}$$

$$= \int_{S_1} \Big(\sum_k a_k \mathbb{1}_{B_k} \Big) \circ \varphi d\mu \tag{14.617c}$$

$$= \int_{S_1} (f \circ \varphi) d\mu. \tag{14.617d}$$

(3) **f à valeurs dans $\bar{\mathbb{R}}^+$**
Vu que f est mesurable, par le théorème 14.118 il existe une suite croissante de fonctions étagées positives convergeant vers f. Soit donc cette suite, $f_n \colon S_2 \to \mathbb{R}^+$. Les fonctions $f_n \circ \varphi$ sont étagées et positives et nous avons aussi la limite ponctuelle et croissante $f_n \circ \varphi \to f \circ \varphi$ parce que φ est continue. Le théorème de la convergence monotone (théorème 14.178) permet d'écrire ceci :

$$\int_{S_2} f d\mu' = \lim \int_{S_2} f_n d\mu' = \lim \int_{S_1} (f_n \circ \varphi) d\mu = \int_{S_1} (f \circ \varphi) d\mu. \tag{14.618}$$

(4) **Pour $f \colon S_2 \to \bar{\mathbb{R}}$ ou $\mathbb{C}$**
C'est maintenant que l'intégrabilité va jouer. Nous avons $|f| \circ \varphi = |f \circ \varphi|$, donc

$$\int_{S_2} |f| d\mu' = \int_{S_1} |f| \circ \varphi d\mu = \int_{S_1} |f \circ \varphi| d\mu, \tag{14.619}$$

ce qui montre que f est μ'-intégrable si et seulement si $f \circ \varphi$ est μ-intégrable.
De plus si $f = f^+ - f^-$ alors $f^+ \circ \varphi = (f \circ \varphi)^+$, $f^- \circ \varphi = (f \circ \varphi)^-$, et de façon similaire pour les parties imaginaires et réelles.

□

14.10 Mesure à densité

14.10.1 Théorème de Radon-Nikodym

Définition 14.220 ([463]).
*Soient μ et ν deux mesures sur l'espace mesurable $(\Omega, \mathcal{A})$. Nous disons que la mesure μ est **dominée** par ν si pour tout ensemble mesurable A, $\nu(A) = 0$ implique $\mu(A) = 0$.*

Nous disons que la mesure signée[70] *ν est **absolument continue** par rapport à la mesure positive μ si pour tout $A \in \mathcal{A}$, nous $\mu(A) = 0 \Rightarrow \nu(A) = 0$. Dans ce cas nous notons*

$$\nu \ll \mu. \tag{14.620}$$

70. Définition 14.18.

DEFooRZARooTbtJac

Définition 14.221 ([439]).
*Deux mesures signées μ et ν sur $(\Omega, \mathcal{A})$ sont **mutuellement singulières** si il existe $E, F \in \mathcal{A}$ tels que*

$$\begin{cases} E \cap F = \varnothing & (14.621a)\\ \mu(A) = \mu(A \cap E) & (14.621b)\\ \nu(A) = \nu(A \cap F). & (14.621c) \end{cases}$$

Nous notons ce fait par

$$\mu \perp \nu. \tag{14.622}$$

LEMooDIGPooJqEwGl

Lemme 14.222 ([1]).
Soient deux mesures positives λ et μ. Nous avons $\lambda \perp \mu$ si et seulement si il existe des parties mesurables E', F' telles que

(1) $E' \cup F' = \Omega$
(2) $E' \cap F' = \varnothing$
(3) $\mu(F') = \lambda(E') = 0$.

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que $\mu \perp \lambda$, c'est-à-dire qu'il existe E, F tels que $E \cap F = \varnothing$, $\mu(A) = \mu(A \cap E)$ et $\lambda(A) = \lambda(A \cap F)$ pour tout mesurables A.
Nous posons $R = \Omega \backslash (E \cup F)$. C'est la partie de Ω sur laquelle ni μ ni λ n'agit. Il suffit de poser $F' = F \cup R$ et $E' = E$ pour avoir ce qu'il faut. Par exemple,

$$\mu(F') = \mu(F \cup R) = \mu\big((F \cup R) \cap E\big) = \mu(\varnothing) = 0. \tag{14.623}$$

, et de même pour les autres vérifications.

(2) $\Leftarrow$ Il suffit de poser $E = E'$ et $F = F'$ et de vérifier. Par exemple, vu que l'union $E \cup F = \Omega$ est disjointe,

$$\mu(A) = \mu(A \cap E) + \mu(A \cap F) = \mu(A \cap E). \tag{14.624}$$

□

LEMooVOJDooFgbwSE

Lemme 14.223 ([464]).
Soit un espace mesurable $(\Omega, \mathcal{A})$. Nous considérons

(1) des mesures positives finies $(\lambda_n)_{n\in\mathbb{N}}$,
(2) une mesure positive μ

telles que $\lambda_n \perp \mu$ pour tout n. Alors

$$\sum_{n=1}^{\infty} \lambda_n \perp \mu. \tag{14.625}$$

Démonstration. Nous utilisons la caractérisation du lemme 14.222 pour des mesures mutuellement singulières.

Pour chaque n nous avons des mesurables A_n et B_n tels que $A_n \cup B_n = \Omega$, $A_n \cap B_n = \varnothing$ et $\mu(A_n) = \lambda_n(B_n) = 0$.

Nous posons $A = \bigcup_{n=1}^{\infty} A_n$ et $B = \Omega \backslash A$. Nous avons bien entendu $A \cup B = \Omega$ et $A \cap B = \varnothing$.

(1) $\lambda_k(B) = 0$ **pour tout** k Pour chaque k nous avons $B \subset B_k$ parce que si $x \in B$, alors en particulier $x \in \Omega \backslash A_k = B_k$. Nous avons donc

$$\Omega \backslash A = B \subset B_k = \Omega \backslash A_k. \tag{14.626}$$

Donc, par croissance de la mesure,

$$\lambda_k(B) = \lambda_k(\Omega \backslash A) \leqslant \lambda_k(\Omega \backslash A_k) = \lambda_k(B_k) = 0. \tag{14.627}$$

(2) $\lambda(B) = 0$ Il s'agit de sommer : $\lambda(B) = \sum_{n=1}^{\infty} \lambda_k(B) = 0$.
(3) $\mu(A) = 0$ Nous avons

$$\mu(A) = \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leqslant \sum_{n=1}^{\infty} \mu(A_n) = 0. \tag{14.628}$$

□

LEMooYYIIooCUfGxA

Lemme 14.224 ([1]).
Soient λ et ν des mesures positives σ-finies ainsi qu'une mesure positive μ telle que $\lambda \perp \mu$ et $\nu \perp \mu$. Alors

$$(\lambda + \nu) \perp \mu. \tag{14.629}$$

Démonstration. Nous considérons des mesurables disjoints (X_n) comme dans le lemme 14.20. Pour chaque n nous avons les mesures finies λ_n et ν_n données par

$$\lambda_n(A) = \lambda(A \cap X_n) \tag{14.630a}$$
$$\nu_n(A) = \nu(A \cap X_n). \tag{14.630b}$$

Ces mesures vérifient $\lambda_n \perp \mu$ et $\nu_n \perp \mu$. De plus les (X_n) étant disjoints, nous avons $\lambda(A) = \sum_n \lambda_n(A)$ et $\nu(A) = \sum_n \nu_n(A)$. Nous pouvons donc utiliser le lemme 14.223 :

$$\sum_{n=1}^{\infty} (\lambda_n + \nu_n) \perp \mu. \tag{14.631}$$

EQooWXCKooGrqShQ

Par associativité de la somme (tous les termes sont positifs), pour chaque mesurable A nous avons

$$\sum_n (\lambda_n + \nu_n)(A) = \sum_n \lambda_n(A) + \sum_n \nu_n(A) = \lambda(A) + \nu(A) = (\lambda + \nu)(A). \tag{14.632}$$

Cela pour dire que le membre de gauche de (14.631) est bien $\lambda + \nu$. □

Définition 14.225.
*La mesure μ est **portée** par la partie $E \in \mathcal{A}$ si pour tout $A \in \mathcal{A}$,*

$$\mu(A) = \mu(A \cap E). \tag{14.633}$$

Définition 14.226.
Soit une mesure signée μ et une partie mesurable A.

*(1) La partie A est **négative** si $\mu(A') \leqslant 0$ pour tout A' mesurable dans A.*
*(2) La partie A est **nulle** si $\mu(A') = 0$ pour tout A' mesurable dans A.*
*(3) La partie A est **positive** si $\mu(A') \geqslant 0$ pour tout A' mesurable dans A.*

PROPooBLNXooRRxxVv

Proposition 14.227 ([439]).
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ où μ est une mesure signée. Si une partie $D \in \mathcal{A}$ vérifie $\mu(D) \leqslant 0$, alors il existe une partie négative $A \subset D$ telle que $\mu(A) \leqslant \mu(D)$.

Démonstration. Nous définissons des parties A_i par récurrence en commençant par poser $A_0 = D$. Ensuite, si A_n est défini, nous posons

$$t_n = \sup\{\mu(B) \text{ tel que } B \subset A_n, B \in \mathcal{A}\}. \tag{14.634}$$

Nous avons $t_n \geqslant \mu(\varnothing) = 0$. Par définition du supremum, il existe B_n mesurable dans A_n vérifiant $\mu(B_n) \geqslant t_n/2 \geqslant 0$. Nous posons alors

$$A_{n+1} = A_n \backslash B_n. \tag{14.635}$$

Enfin nous posons

$$A = D \backslash \bigcup_{n \geqslant 0} B_n. \tag{14.636}$$

Vérifions que A possède les propriétés demandées.

(1) $A \subset D$ Par définition.

(2) $\mu(A) \leqslant \mu(D)$

(2a) **Les B_i sont disjoints** Nous avons

$$B_n \subset A_n = A = D \setminus \bigcup_{n\geqslant 0} B_n, \tag{14.637}$$

et donc $B_n \cap B_k = \varnothing$ pour tout $k < n$. Cela montre que les B_i sont disjoints.

(2b) $\mu(A) \leqslant \mu(D)$ Étant donné que $B_n \subset D$ pour tout n, l'égalité $A = D\setminus\bigcup_{n\geqslant 0} B_n$ peut être écrite

$$D = A \cup \bigcup_{n\geqslant 0} B_n \tag{14.638}$$

où à droite nous avons une union disjointe. Bref, au niveau des mesures nous avons

$$\mu(A) = \mu(D) - \sum_{n\geqslant 0} \mu(B_n) \leqslant \mu(D) \tag{14.639}$$

parce que $\mu(B_n) \geqslant 0$ pour tout n.

(3) $A \subset A_n$ Par définition $A = D\setminus\bigcup_{n\geqslant 0} B_n$ alors que $A_n = D\setminus\bigcup_{n=0}^{n-1} B_n$. Donc en partant de D, pour construire A, on enlève plus de choses que pour construire A_n. Donc $A \subset A_n$.

(4) **A est une partie négative** Supposons que A n'est pas une partie négative : il existe A' mesurable dans A tel que $\mu(A') > 0$. En particulier,

$$A' \subset A \subset A_n. \tag{14.640}$$

La partie A' est donc dans l'ensemble sur lequel on prend le supremum pour construire t_n. En d'autres termes, $t_n \geqslant \mu(A') > 0$, et donc

$$\mu(B_n) \geqslant \frac{t_n}{2} \geqslant \frac{\mu(A')}{2}. \tag{14.641}$$

Et avec tout ça,

$$\mu(A) = \mu(D) - \sum_{n\geqslant 0} \mu(B_n) \leqslant \mu(D) - \sum_{n\geqslant 0} \frac{\mu(A)}{2} = -\infty. \tag{14.642}$$

EQooYNJEooKJCfET

(5) **Conclusion** Nous avons dit, dans la définition d'une mesure négative que μ ne prenait jamais la valeur $-\infty$. L'inégalité (14.642) est donc une contradiction et nous concluons que A est une partie négative pour μ.

□

LEMooVFEWooOyXcPS

Lemme 14.228 ([439]).
Soient une mesure positive μ et une mesure signée ν. Si $\nu \perp \mu$ et si $\nu \ll \mu$, alors $\nu = 0$

Démonstration. Vu que $\mu \perp \nu$, il existe $E \in \mathcal{A}$ tel que

$$\begin{cases} \mu(A) = \mu(A \cap E) & \text{(14.643a)} \\ \nu(A) = \nu(A \cap E^c) & \text{(14.643b)} \end{cases}$$

pour tout $A \in \mathcal{A}$.

Soit $A \in \mathcal{A}$. Nous avons $\mu(A \cap E^c) = 0$ parce que $\mu(A) = \mu(A \cap E) + \mu(A \cap E^c)$ et $\mu(A \cap E) = \mu(A)$. Étant donné que $\nu \ll \mu$ nous avons aussi $\nu(A \cap E^c) = 0$. Donc

$$\nu(A) = \nu(A \cap E^c) = 0. \tag{14.644}$$

□

LEMooABVVooMvxlRo

Lemme 14.229.
Soient une mesure positive μ ainsi que deux mesures signées ν_1, ν_2 telles que $\nu_1 \perp \mu$ et $\nu_2 \perp \mu$. Alors $\nu_1 + \nu_2 \perp \mu$.

Démonstration. Nous avons des parties $E_1, F_1, E_2, F_2 \in \mathcal{A}$ telles que

$$\begin{cases} \nu_1(A) = \nu_1(A \cap F_1) & (14.645a) \\ \mu(A) = \mu(A \cap E_1), & (14.645b) \end{cases}$$

et

$$\begin{cases} \nu_2(A) = \nu_2(A \cap F_2) & (14.646a) \\ \mu(A) = \mu(A \cap E_2), & (14.646b) \end{cases}$$

et $E_1 \cap F_1 = \varnothing$ et $E_2 \cap F_2 = \varnothing$. Nous posons $E_0 = E_1 \cap E_2$ et $F_0 = F_1 \cup F_2$, et nous allons prouver que $\nu_1 + \nu_2 \perp \mu$ vérifiant que le couple (E_0, F_0) vérifie les conditions de la définition 14.221.

(1) $\underline{E_0 \cap F_0 = \varnothing}$ Nous avons :

$$E_0 \cap F_0 = (E_1 \cap E_2) \cap (F_1 \cup F_2) = (E_1 \cap E_2 \cap F_1) \cup (E_1 \cap E_2 \cap F_2) = \varnothing. \quad (14.647)$$

(2) $\underline{\mu(A) = \mu(A \cap E_0)}$ Nous prouvons maintenant que $\mu(A) = \mu(A \cap E_0)$ pour tout $A \in \mathcal{A}$. Nous écrivons d'abord

$$\mu(A) = \mu(A \cap E_1) = \mu(A \cap E_1 \cap E_2) + \mu\big(A \cap (E_1 \backslash E_2)\big), \quad (14.648)$$

EQooTMOKooEulEmx

et, de la même façon,

$$\mu(A) = \mu(A \cap E_2) = \mu(A \cap E_2 \cap E_1) + \mu\big(A \cap (E_2 \backslash E_1)\big). \quad (14.649)$$

En faisant la différence entre les deux, $0 = \mu\big(A \cap (E_1 \backslash E_2)\big) - \mu\big(A \cap (E_2 \backslash E_1)\big)$, c'est-à-dire

$$\mu\big(A \cap (E_1 \backslash E_2)\big) = \mu\big(A \cap (E_2 \backslash E_1)\big). \quad (14.650)$$

En particulier pour $A = \Omega$ nous avons

$$\mu(E_1 \backslash E_2) = \mu(E_2 \backslash E_1). \quad (14.651)$$

Nous savons, par définition de E_1 que $\mu(A) = \mu(A \cap E_1)$ pour tout $A \in \mathcal{A}$. Nous écrivons cette égalité pour $A = (E_2 \backslash E_1)$:

$$\mu(E_1 \backslash E_2) = \mu\big((E_2 \backslash E_1) \cap E_1\big) = \mu(\varnothing) = 0. \quad (14.652)$$

Nous en déduisons que $\mu(E_2 \backslash E_1) = 0$ et donc que $\mu\big(A \cap (E_2 \backslash E_1)\big) = 0$ pour tout $A \in \mathcal{A}$. Et enfin nous repartons de (14.648) :

$$\mu(A) = \mu(A \cap E_1) = \underbrace{\mu(A \cap E_1 \cap E_2)}_{=\mu(A \cap E_0)} + \underbrace{\mu\big(A \cap (E_1 \backslash E_2)\big)}_{=0}, \quad (14.653)$$

autrement dit, $\mu(A) = \mu(A \cap E_0)$, comme nous devions prouver.

(3) $\underline{\nu_i(A) = \nu_i(A \cap F_0)}$ Nous allons prouver que pour $i = 1, 2$ nous avons $\nu_i(A) = \nu_i(A \cap F_0)$. Nous commençons par écrire $A \cap F_0$ comme union disjointe en partant de 1.28(2) :

$$\begin{aligned} A \cap F_0 &= (A \cap F_1) \cup (A \cap F_2) && (14.654a) \\ &= (A \cap F_1) \cup \big((A \cap F_2) \backslash (A \cap F_1)\big). \qquad \text{cf. justif.} && (14.654b) \end{aligned}$$

Justification : pour tout ensembles X et Y nous avons $X \cup Y = X \cup (Y \backslash X)$.

Ayant écrit $A \cap F_0$ comme une union disjointe, nous pouvons calculer ν_1 dessus. Nous posons $R = (A \cap F_2) \backslash (A \cap F_1)$, et nous calculons un peu :

$$\nu_1(A \cap F_0) = \nu_1(A \cap F_1) + \nu_1(R) = \nu_1(A) + \nu_1(R \cap F_1) = \nu_1(A) + \nu_1(\varnothing) = \nu_1(A). \quad (14.655)$$

Le même jeu en permutant les rôles de 1 et 2 donne $\nu_2(A \cap F_0) = \nu_2(A)$.

(4) **Conclusion** Nous avons $\nu_1(A) = \nu_1(A \cap F_0)$ et $\nu_2(A) = \nu_2(A \cap F_0)$. En faisant la somme, c'est bon.

□

PROPooIBLHooTMfEJW

Proposition-Définition 14.230 (Décomposition de Hahn[439])**.**
Soit une mesure signée μ sur $(\Omega, \mathcal{A})$. Il existe des parties mesurables $N, P \subset \Omega$ telles que

(1) $N \cup P = \Omega$,

(2) $N \cap P = \varnothing$,

(3) N *est négatif pour* μ,

(4) P *est positive pour* μ.

Un tel couple (N, P) est une ***décomposition de Hahn*** *pour μ.*

De plus si (N, P) et (N', P') sont deux décompositions de Hahn pour μ, alors les parties[71] *$N \Delta N'$ et $P \Delta P'$ sont nulles pour μ.*

Démonstration. D'abord l'existence puis l'« unicité ».

Existence

Nous posons d'abord $N_0 = \varnothing$, et nous construisons des parties N_k par récurrence. Nous posons

$$s_n = \inf\{\mu(D) \text{ tel que } D \subset \Omega \backslash N_n, D \in \mathcal{A}\}. \tag{14.656}$$

Notes : il est possible d'avoir $s_n = -\infty$, et comme $D = \varnothing$ fonctionne dans l'infimum, nous avons $s_n \leqslant 0$. Vu que s_n est un infimum et qu'il est négatif, il existe une partie mesurable $D_n \subset \Omega \backslash N_n$ telle que $\mu(D_n) \leqslant \frac{s_n}{2}$. Pour cette partie nous avons

$$\mu(D_n) \leqslant \frac{s_n}{2} \leqslant 0. \tag{14.657}$$

La proposition 14.227 dit qu'il existe une partie mesurable $A_n \subset D_n$ qui est négative pour μ et qui vérifie $\mu(A_n) \leqslant \mu(D_n)$. Nous posons alors, pour notre récurrence :

$$N_{n+1} = N_n \cup A_n. \tag{14.658}$$

Nous posons ensuite

$$\begin{cases} N = \bigcup_{n \geqslant 0} A_n & \text{(14.659a)} \\ P = \Omega \backslash N, & \text{(14.659b)} \end{cases}$$

et nous prouvons que (N, P) fonctionne. Nous avons immédiatement que $N \cup P = \Omega$ et que $N \cap P = \varnothing$.

(1) N **est négative pour** μ Nous avons, pour tout n que $N_n = \varnothing \cup A_1 \cup \ldots \cup A_{n-1}$. Donc $N_n \subset N$ pour tout $n \geqslant 1$.

Notons aussi que les A_n sont disjoints. En effet

$$A_n \subset D_n \subset \Omega \backslash N_n = \Omega \backslash (A_1 \cup \ldots \cup A_{n-1}). \tag{14.660}$$

Donc A_n n'intersecte aucun des A_k avec $k < n$.

Nous pouvons maintenant que N est une partie négative. Soit $B \subset N$. Notons que chaque A_n est négatif, donc $\mu(A_n \cap B) \leqslant 0$ parce que, évidemment, $A_n \cap B \subset B$. Étant donné que l'union $N = \bigcup_{n \geqslant 0} A_n$ est disjointe, nous avons

$$\mu(B) = \sum_{n \geqslant 0} \mu(B \cap A_n) \leqslant 0. \tag{14.661}$$

Donc $m(B) \leqslant 0$ et nous avons montré que N est une partie négative.

71. La partie $A \Delta B$ est la différence symétrique (définition 1.30). Dire que $N \Delta N'$ est nulle pour μ signifie que pratiquement tout est dans l'intersection. Nous avons donc unicité de la décomposition à partie nulle près.

(2) **P est positive pour μ** Nous y allons par l'absurde. Supposons avoir $D \subset P$ avec $\mu(D) < 0$. Nous avons déjà vu que $N_n \subset N$, donc $\Omega\backslash N \subset \Omega\backslash N_n$. En ce qui concerne D,

$$D \subset P = \Omega\backslash N \subset \Omega\backslash N_n. \tag{14.662}$$

Donc $s_n \leqslant \mu(D) < 0$. Nous avons alors le calcul

$$\begin{aligned}
\mu(N) &= \sum_{n\geqslant 0} \mu(A_n) && && \text{(14.663a)}\\
&\leqslant \sum_{n\geqslant 0} \mu(D_n) && \text{cf. justif.} && \text{(14.663b)}\\
&\leqslant \sum_{n\geqslant 0} \frac{\mu(D)}{2} && \text{cf.justif.} && \text{(14.663c)}\\
&= -\infty. && && \text{(14.663d)}
\end{aligned}$$

Justifications.

— Pour (14.663b). Nous avons $\mu(A_n) \leqslant \mu(D_n)$ par choix de A_n.

— Pour (14.663c). Parce que[72] $\mu(D_n) \leqslant \frac{s_n}{2} \leqslant \frac{\mu(D)}{2}$.

Vu que μ ne prend pas la valeur $-\infty$, nous avons une contradiction. Nous en déduisons que P est positive pour μ.

« Unicité »

Supposons avoir des couples (N, P) et (N', P'). Nous avons d'abord

$$P\backslash P' = P \cap N' \tag{14.664}$$

parce que si $x \in P\backslash P'$, alors $x \in P$ (ça c'est facile), et $x \in P\backslash P' \subset \Omega\backslash P' = N'$. Dans l'autre sens si $x \in P \cap N'$, alors $x \in N' \subset \Omega \cap P'$. De même nous avons $P'\backslash P = P' \cap N$.

En utilisant la formule (1.12) de la différence symétrique,

$$P\Delta P' = (P\backslash P') \cup (P'\backslash P) = (P \cap N') \cup (P' \cap N). \tag{14.665}$$

Nous trouvons de la même façon

$$N\Delta N' = (P \cap N') \cup (P' \cap N). \tag{14.666}$$

En ce qui concerne les mesures, $P \cap N' \subset P$ et P est positif, donc $\mu(P \cap N') \geqslant 0$. Mais $P \cap N' \subset N'$ et N' est négatif, donc $\mu(P \cap N')) \leqslant 0$.

Nous en déduisons que $\mu(P \cap N') = 0$. Pour la même raison, $\mu(P' \cap N) = 0$. Nous déduisons alors de (14.665) et (14.666) que $\mu(P\Delta P') = \mu(N\Delta N') = 0$. □

Lemme 14.231.
Soit une décomposition de Hahn (N, P) pour la mesure signée μ. Nous posons

$$\mu_-(A) = -\mu(A \cap N) \tag{14.667a}$$
$$\mu_+(A) = \mu(A \cap P). \tag{14.667b}$$

Alors

(1) μ_- et μ_+ sont des mesures positives.

(2) μ_- prend ses valeurs dans $[0, \infty[$

72. Je n'ai vu ce passage correctement justifié nulle part. Voir ma question sur le wikipédia anglophone[465], qui fait la même « faute » que [439]. Écrivez-moi si vous comprenez leur démarche.

Démonstration. Vu que N est négative pour μ, pour toute partie mesurable A, nous avons $\mu(A\cap N)\leqslant 0$ et donc

$$\mu_-(A)=-\mu(A\cap N)\geqslant 0. \tag{14.668}$$

Par ailleurs μ prend ses valeurs dans $]-\infty,\infty]$. De cette façon $-\mu$ ne peut pas descendre jusqu'à $-\infty$. □

DEFooGICVooKZWLrB

Définition 14.232 (Décomposition de Jordan)**.**
Soit une mesure signée μ sur $(\Omega,\mathcal{A})$. Un couple de mesures positives (μ_-,μ_+) est une ***décomposition de Jordan*** *de μ si il existe une partie mesurable E telle que*

(1) $\mu_-(E)=0$
(2) $\mu_+(\Omega\backslash E)=0$
(3) $\mu=\mu_+-\mu_-$.

LEMooTVFRooAPdbuP

Lemme 14.233 ([1])**.**
Soit une mesure signée μ et une décomposition de Jordan (μ_-,μ_+). Si E est un mesurable vérifiant les conditions de la définition 14.232, alors $(\Omega\backslash E,E)$ est une décomposition de Hahn.

Démonstration. Supposons que $A\subset E$. Alors $\mu_-(A)=0$ parce que $\mu_-(E)=0$ et μ_- est une mesure positive. Nous avons alors

$$\mu(A)=\mu_+(A)-\mu_-(A)=\mu_+(A)\geqslant 0. \tag{14.669}$$

Cela prouve que E est une partie positive pour μ. Le même raisonnement montre que $\Omega\backslash E$ est une partie négative pour μ. □

DEFooDYFPooITNRlI

Proposition 14.234 (Décomposition de Jordan[439])**.**
Soit une mesure mesure signée μ sur $(\Omega,\mathcal{A})$.

(1) Il existe une unique décomposition de Jordan (μ_-,μ_+) pour μ. ITEMooDMVXooTXBFnM
(2) Si (N,P) est une décomposition de Hahn, alors la décomposition de Jordan de μ est donnée par

$$\mu_-(A)=-\mu(A\cap N) \tag{14.670a}$$
$$\mu_+(A)=\mu(A\cap P). \tag{14.670b}$$

Démonstration. Nous prouvons l'existence directement en prouvant le point (2). Vérifions les trois points de la définition 14.234. Nous posons $E=P$, et nous vérifions

$$\mu_-(E)=-\mu(E\cap N)=-\mu(E\cap N)=-\mu(\varnothing)=0. \tag{14.671}$$

, et de même $\mu_+(\Omega\backslash E)=\mu_+(N)=\mu(A\cap N)=0$. Enfin, vu que $\Omega=N\cup P$ est une union disjointe, nous avons $\mu(A)=\mu(A\cap N)+\mu(A\cap P)=-\mu_-(A)+\mu_+(A)$.

Il nous reste à prouver l'unicité de la décomposition de Jordan. Nous supposons que (μ_-,μ_+) et (ν_-,ν_+) sont des décompositions de Jordan de μ. Nous avons donc des mesurables E et F tels que $\mu=\nu_+-\nu_-=\mu_+-\mu_-$ et

$$\begin{cases}\mu_-(E)=0 & \text{(14.672a)}\\ \mu_+(\Omega\backslash E)=0 & \text{(14.672b)}\\ \nu_-(F)=0 & \text{(14.672c)}\\ \nu_+(\Omega\backslash F)=0. & \text{(14.672d)}\end{cases}$$

Le lemme 14.233 nous indique que $(\Omega\backslash E,E)$ et $(\Omega\backslash F,F)$ sont des décompositions de Hahn pour μ. Aussi, pour prendre des notations plus intuitives, nous posons

$$\begin{aligned}N&=\Omega\backslash E & P&=E\\ N'&=\Omega\backslash F & P'&=F\\ \mu'_-&=\nu_- & \mu'_+&=\nu_+\end{aligned} \tag{14.673}$$

La partie unicité de la décomposition de Hahn 14.230 implique que $N\Delta N'$ et $P\Delta P'$ sont des parties nulles pour μ. Étant donné que $N'\backslash N \subset N\Delta N'$, la partie $N'\backslash N$ est également nulle pour μ. Idem pour $N\backslash N'$. Enfin nous avons l'union disjointe

$$N' = (N' \cap N) \cup (N'\backslash N), \tag{14.674}$$

et nous pouvons calculer

$$\begin{aligned}
\mu'_-(A) &= -\mu(A \cap N') && (14.675a)\\
&= -\mu\big((A \cap N') \cap (N' \cap N)\big) - \mu\big((A \cap N') \cap (N'\backslash N)\big) && (14.675b)\\
&= -\mu(A \cap N \cap N') - \mu\big(A \cap (N'\backslash N)\big) && (14.675c)\\
&= -\mu\big(A \cap N \cap N'\big) - \mu\big(A \cap (N\backslash N')\big) && \text{cf.justif.} \quad (14.675d)\\
&= -\mu(A \cap N) && \text{cf. justif} \quad (14.675e)\\
&= \mu_-(A). && (14.675f)
\end{aligned}$$

SUBEQooFQVAooNNxDHX

SUBEQooKCMDooWeECXv

Justifications.

— Pour (14.675d). Nous avons

$$\mu\big(A \cap (N\backslash N')\big) = \mu\big(A \cap (N'\backslash N)\big) = 0 \tag{14.676}$$

parce que $N\backslash N'$ et $N'\backslash N$ sont des parties nulles pour μ. Bref le passage (14.675d) consiste à remplacer un truc nul par un autre.

— Pour (14.675e). Nous avons l'union disjointe $N = (N' \cap N) \cup (N\backslash N')$.

Nous avons prouvé que $\mu'_-(A) = \mu_-(A)$, et donc que $\mu'_- = \mu_-$. Nous faisons de même pour $\mu'_+ = \mu_+$. □

THOooKSISooPAqZcp

Théorème 14.235 (Radon-Nikodym[439, 1])**.**
Soit un espace mesurable $(\Omega, \mathcal{A})$. Nous considérons

(1) μ, une mesure positive σ-finie.

(2) ν, une mesure signée[73] *ou non telle que $|\nu|$ est σ-finie.*

Alors

(1) Il existe un unique couple de mesures signées ν_a, ν_s tel que

$$\begin{cases} \nu = \nu_a + \nu_s & (14.677a)\\ \nu_a \ll \mu & (14.677b)\\ \nu_s \perp \mu. & (14.677c) \end{cases}$$

(2) Il existe une unique (à égalité presque partout près) fonction mesurable et intégrable[74] *$f : \Omega \to \mathbb{R} \cup \{+\infty\}$ telle que*

$$\nu_a(A) = \int_A f d\mu \tag{14.678}$$

pour tout $A \in \mathcal{A}$.

(3) Si $\nu \geqslant 0$ alors $\nu_a \geqslant 0$, $\nu_s \geqslant 0$ et $f \geqslant 0$.

Démonstration. Nous commençons par l'unicité.

Unicité des mesures et de la fonction

Séparément.

73. Mesure signée, définition 14.18.

74. Intégrale ne signifie pas L^1. Nous demandons juste que l'intégrale existe, si elle vaut ∞, c'est bon pour nous.

(1) **Unicité des mesures** Nous supposons avoir deux tels couples (ν_a, ν_s) et (ν'_a, ν'_s). Nous avons donc $\nu_a + \nu_s = \nu$ et $\nu'_a + \nu'_s = \nu$. En faisant la différence,

$$\nu_a - \nu'_a = (\nu - \nu_s) - (\nu - \nu'_s) = \nu'_s - \nu_s. \tag{14.679}$$

D'autre part, vu que $\nu_s \perp \mu$ et $\nu'_s \perp \mu$, le lemme 14.229 donne $\nu'_s - \nu_s \perp \mu$. Enfin $\nu_a \ll \mu$ et $\nu'_a \ll \mu$, donc $\nu_a - \nu'_a \ll \mu$. Au final nous avons

$$\begin{cases} \nu_a - \nu'_a \ll \mu & \text{(14.680a)} \\ \nu'_s - \nu_s \perp \mu & \text{(14.680b)} \\ \nu_a - \nu'_a = \nu'_s - \nu_s. & \text{(14.680c)} \end{cases}$$

Tout cela avec le lemme 14.228 nous donne $\nu_a - \nu'_a = \nu_s - \nu'_s = 0$.

(2) **Unicité de la fonction**
Soient deux fonctions $f, g\colon \Omega \to \mathbb{R} \cup \{\infty\}$ vérifiant la contrainte. Pour toute partie $A \in \mathcal{A}$, nous avons

$$\nu(A) = \int_A f d\mu = \int_A g d\mu. \tag{14.681}$$

Nous écrivons cette égalité pour la partie $A = \{f > g\}$:

$$\int_{\{f>g\}} (f - g) d\mu = 0. \tag{14.682}$$

Vu que sur la partie $\{f > g\}$ nous avons $f - g > 0$, nous avons $\mu\big(\{f > g\}\big) = 0$ par le lemme 14.201. Nous trouvons de même que $\mu\big(\{f < g\}\big) = 0$. Au final, nous avons $f = g$ presque partout.

Existence, μ positive finie, ν positive et finie

Nous commençons par prouver la partie existence en supposant que μ est finie et que ν est positive et finie. Nous introduisons l'ensemble suivant :

$$\mathcal{H} = \left\{ f\colon \Omega \to \mathbb{R}^+ \cup \{\infty\} \text{ tel que } \begin{cases} f \text{est mesurable} \\ \int_A f d\mu \leqslant \nu(A) \text{ pour tout } A \in \mathcal{A} \end{cases} \right\} \tag{14.683}$$

L'ensemble $\mathcal{H}$ n'est pas vide parce que la fonction $f = 0$ est dedans (ici nous utilisons l'hypothèse supplémentaire que $\nu \geqslant 0$). Nous notons

$$\alpha = \sup\{\int_\Omega f d\mu \text{ tel que } f \in \mathcal{H}\}. \tag{14.684}$$

Nous avons $0 \leqslant \alpha \leqslant \mu(\Omega) < \infty$. En vertu du lemme 1.464 nous pouvons prendre une suite (f_n) dans $\mathcal{H}$ telle que

$$\alpha = \lim n \to \infty \int_\Omega f_n d\mu \tag{14.685}$$

et

$$\int_\Omega f_n d\mu \leqslant \int_\Omega f_{n+1} d\mu. \tag{14.686}$$

(1) **Les fonctions g_n**
Pour chaque $n \geqslant 1$ nous notons

$$\begin{aligned} g_n\colon \Omega &\to \mathbb{R} \\ x &\mapsto \max\{f_1(x), \dots, f_n(x)\}. \end{aligned} \tag{14.687}$$

Nous nous fixons provisoirement un n.

(2) $g_n \in \mathcal{H}$
Pour chaque $x \in \Omega$, il existe un $i \in \{1, \dots, n\}$ tel que $g_n(x) = f_i(x)$. Pour $A \in \mathcal{A}$ nous posons

$$A_1 = \{x \in \Omega \text{ tel que } g_n(x) = f_1(x)\} \tag{14.688a}$$
$$A_2 = \{x \in \Omega \text{ tel que } g_n(x) = f_2(x)\} \backslash A_1 \tag{14.688b}$$
$$\vdots \tag{14.688c}$$
$$A_{k+1} = \{x \in \Omega \text{ tel que } g_n(x) = f_{k+1}(x)\} \backslash (A_1 \cup \dots \cup A_k) \tag{14.688d}$$
$$\vdots \tag{14.688e}$$
$$A_n = \{x \in \Omega \text{ tel que } g_n(x) = f_n(x)\} \backslash (A_1 \cup \dots \cup A_{n-1}) \tag{14.688f}$$

Ces ensembles sont disjoints et $A = \bigcup_{i=1}^{n} A_n$, de telle sorte que

$$\int_A g_n d\mu = \sum_{i=1}^{n} \int_{A_i} g_n d\mu \tag{14.689a}$$
$$= \sum_{i=1}^{n} \int_{A_i} f_i d\mu \tag{14.689b}$$
$$\leqslant \sum_{i=1}^{n} \nu(A_i) \text{cf. justif} \tag{14.689c}$$
$$= \nu(A). \tag{14.689d}$$

SUBEQooIULRooCXJmzd

Justification de (14.689c). Vu que $f_i \in \mathcal{H}$, nous avons $\int_{A_i} f_i \leqslant \nu(A_i)$. Bref, nous avons prouvé que $\int_A g_n d\mu \leqslant \nu(A)$, et donc que $g_n \in \mathcal{H}$.

(3) g_n **est mesurable** Les applications f_n sont mesurables, et g_n est définie comme un maximum sur les $f_i(x)$. Donc g_n est mesurable (lemme 14.106). Nous ne fixons plus le n. La suite (g_n) est croissante. Donc la limite $f = \lim_{n\to\infty} g_n$ existe, et elle est mesurable par la proposition 14.110.

(4) $\int_\alpha f d\mu = \alpha$
SPITEMooGLPCooBjBAdm
Le théorème de la convergence monotone 14.178 va plus loin et nous dit que f est intégrable et

$$\int_A f d\mu = \lim_{n\to\infty} \int_A g_n d\mu \leqslant \nu(A). \tag{14.690}$$

EQooBGAQooCdaJID

Nous avons donc encore $f \in \mathcal{H}$ et donc aussi

$$\int_\Omega f d\mu \leqslant \alpha. \tag{14.691}$$

Mais cette inégalité tient aussi dans l'autre sens :

$$\int_\Omega f d\mu = \lim \int_\Omega g_n d\mu \tag{14.692a}$$
$$\geqslant \lim \int_\Omega f_n d\mu \tag{14.692b}$$
$$= \alpha. \tag{14.692c}$$

Nous avons donc prouvé que $\int_\Omega f d\mu = \alpha < \infty$, et donc que $f \in L^1(\Omega, \mathcal{A}, \mu)$.

(5) **Résumé pour l'instant** Nous avons construit une fonction $f \in L^1(\Omega, \mathcal{A}, \mu)$ telle que pour tout $A \in \mathcal{A}$,

$$\int_A f d\mu \leqslant \nu(A). \tag{14.693}$$

(6) **Les mesures** Nous définissons à présent les applications ν_a et ν_s en espérant qu'elles fonctionneront. D'abord nous posons

$$\nu_a(A) = \int_A f d\mu. \tag{14.694}$$

Et ensuite nous posons $\nu_s = \nu - \nu_a$. Il nous reste à prouver que f, ν_a et ν_s satisfont à toutes les conditions.

(7) $\nu_a \ll \mu$
Si $\mu(A) = 0$ nous avons

$$\nu_a(A) = \int_A f d\mu = 0 \tag{14.695}$$

par le lemme 14.203.

(8) ν_s **est positive**
Nous avons

$$\nu_s(A) = \nu(A) - \nu_a(A) = \nu(A) - \int_A f d\mu \geqslant 0, \tag{14.696}$$

par l'inégalité (14.690).

(9) ν_s **est finie**
Nous savons que ν_a et finie et que $\int_\Omega f d\mu = \alpha < \infty$. Donc ν_s est finie.

(10) $\nu_s \perp \mu$
Nous introduisons la mesure signée $\lambda_n = \nu_s - \frac{1}{n}\mu$, et nous considérons la décomposition de Hahn pour chacun des λ_n (proposition 14.230) : les couples (N_n, P_n). Nous considérons aussi les fonctions $h_n = f + \frac{1}{n}\mathbb{1}_{P_n}$.

(10a) $h_n \in \mathcal{H}$ La fonction h_n est mesurables parce que f et P_n le sont. En termes d'intégrale nous avons

$$\begin{aligned}
\int_A h_n d\mu &= \int_A f d\mu + \frac{1}{n}\mu(A \cap P_n) && (14.697a)\\
&= \nu_a(A) + \nu_s(A \cap P_n) - \lambda_n(A \cap P_n) && \text{cf. justif} \quad (14.697b)\\
&= \nu(A) - \nu_s(A) + \nu_s(A \cap P_n) - \lambda_n(A \cap P_n) && (14.697c)\\
&= \nu(A) - \big(\nu_s(A \cap P_n) + \nu_s(A \cap N_n)\big) + \nu_s(A \cap P_n) && (14.697d)\\
&\quad - \lambda_n(A \cap P_n) && \text{cf. justif.}\\
&= \nu(A) - \nu_s(A \cap N_n) - \lambda_n(A \cap P_n) && (14.697e)\\
&\leqslant \nu(A) && \text{cf. justif.} \quad (14.697f)
\end{aligned}$$

Justifications.

— Pour (14.697b). La définition (14.694) de ν_a).
— Pour (14.697d). Parce que $P_n \cup N_n = \Omega$ est une union disjointe.
— Pour (14.697f). Parce que ν_s est une mesure positive. Donc $\nu_s(A \cap N_n) \geqslant 0$. Et aussi parce que, par construction, λ_n est positive sur P_n.

Bref, nous avons déjà prouvé que $\int_A h_n d\mu \leqslant \nu(A)$. Cela signifie que $h_n \in \mathcal{H}$.

(10b) $\mu(P_n) = 0$ Nous avons vu en (4) que $\int_\Omega f d\mu = \alpha$. Nous pouvons alors faire le calcul suivant :

$$\begin{aligned}
\alpha &\geqslant \int_\Omega h_n d\mu && \text{cf.justi.} \quad (14.698a)\\
&= \int_\Omega df\mu + \frac{1}{n}\int_\Omega \mathbb{1}_{P_n} d\mu && (14.698b)\\
&= \alpha + \frac{1}{n}\mu(P_n). && (14.698c)
\end{aligned}$$

Justification pour (14.698a) : parce que $h_n \in \mathcal{H}$ et que α est le supremum.
En soustrayant α des deux côtés, nous avons $0 \geqslant \frac{1}{n}\mu(P_n)$ pour tout n. Comme μ est une mesure positive, nous en déduisons que $\mu(P_n) = 0$.

(10c) **Définition de P** Nous posons $P = \bigcup_{n=1}^{\infty} P_n$.

(10d) $\mu(P) = 0$ Vu que $\mu(P_n) = 0$ pour tout n, nous avons $\mu(P) \leqslant \sum_n \mu(P_n) = 0$.

(10e) $\nu_s(\Omega\backslash P) = 0$
Nous avons $\Omega\backslash P \subset \Omega\backslash P_n = N_n$. Donc pour tout n nous avons

$$\lambda_n(\Omega\backslash P) \leqslant \lambda_n(N_n) = 0. \tag{14.699}$$

En utilisant la définition de λ_n nous avons

$$0 \geqslant \lambda_n(\Omega\backslash P) = \nu_s(\Omega\backslash P) - \frac{1}{n}\mu(\Omega\backslash P), \tag{14.700}$$

et donc, pour tout n,

$$\nu_s(\Omega\backslash P) \leqslant \frac{1}{n}\mu(\Omega\backslash P). \tag{14.701}$$

En faisant $n \to \infty$ nous trouvons $\nu_s(\Omega\backslash P) \leqslant 0$. Et comme ν_s est une mesure positive, nous en déduisons que $\nu_s(\Omega\backslash P) = 0$.

Nous avons fini de prouver que $\nu_s \perp \mu$.

Nous avons donc terminé la première partie de la preuve, c'est-à-dire la partie où nous supposions que μ est finie et ν est positive et finie. Nous passons à présent à un cas un peu plus général.

——— Existence, μ positive σ-finie, ν positive et σ-finie ———

Vu que μ et ν sont σ-finie, le lemme 14.20 dit qu'il existe des X_n mesurables tels que $\Omega = \bigcup_{n\geqslant 1} X_n$, $\mu(X_n) < \infty$ et$\nu(X_n) < \infty$. Nous supposons de plus que ce X_n sont disjoints [75]
Nous considérons à présent les mesures finies μ_n et ν_n définies par

$$\mu_n(A) = \mu(A \cap X_n) \tag{14.702a}$$
$$\nu_n(A) = \nu(A \cap X_n) \tag{14.702b}$$
$$\tag{14.702c}$$

Par la partie déjà faite, nous pouvons faire une décomposition pour chaque n : il existe des mesures $\nu_{n,a}$, $\nu_{n,s}$ et une fonction $f_n \in L^1(\Omega, \mathcal{A}, \mu)$ telles que [SUBEQSooLONIooAksRgK]

$$\begin{cases} \nu_{n,a}(A) = \int_A f_n d\mu_n & \text{(14.703a)} \\ \nu_n = \nu_{n,a} + \nu_{n,s} & \text{(14.703b)} \\ \nu_{n,a} \ll \mu_n & \text{(14.703c)} \\ \nu_{n,s} \perp \mu_n. & \text{(14.703d)} \end{cases}$$

Nous allons directement changer les f_n pour qu'ils soient plus faciles à traiter. Nous posons [76]

$$g_n(x) = \begin{cases} f_n(x) & \text{si } x \in X_n \\ 0 & \text{sinon.} \end{cases} \tag{14.704}$$

Pour ces fonctions nous avons, parce que μ_n est nulle en dehors de X_n,

$$\int_A g_n d\mu_n = \int_{A\cap X_n} g_n d\mu_n = \int_{A\cap X_n} f_n d\mu_n = \int_A f_n d\mu_n = \nu_{n,a}(A). \tag{14.705}$$

Autrement dit, les fonctions g_n fonctionnent aussi bien que les fonctions f_n. Pour ne pas alourdir les notations, nous gardons (14.703), mais nous supposons que f_n est nulle en dehors de X_n.

75. Si ce n'est pas le cas, poser $Y_{n+1} = X_{n+1}\backslash Y_n$.
76. Cette partie est de l'invention personnelle. J'espère que c'est correct. Redoublez de vigilance.

Notez aussi que pour tout $A\in\mathcal{A}$ nous avons l'union disjointe $A=\bigcup A\cap X_n$, de telle sorte que

$$\mu(A)=\sum_n\mu(A\cap X_n)=\sum_n\mu_n(A),\tag{14.706}$$

et nous pouvons écrire

$$\mu=\sum_{n\geqslant 0}\mu_n.\tag{14.707}$$

Nous posons

$$\begin{cases} f=\sum_{n\geqslant 1}f_n & (14.708a)\\ \nu_a=\sum_{n\geqslant 1}\nu_{n,a} & (14.708b)\\ \nu_s=\sum_{n\geqslant 1}\nu_{s,a}, & (14.708c)\end{cases}$$

et nous prouvons que cela fait l'affaire. Notons que les fonctions f_n sont L^1 ; cela ne signifie pas que f soit L^1, mais nous ne le demandons pas. La fonction f est certainement intégrable parce qu'elle est positive. Au pire $\int_A f d\mu=\infty$ pour certains A. C'est la vie.

Nous devons vérifier les points suivants :

— $\nu_a(A)=\int_A f d\mu$
— $\nu=\nu_a+\nu_s$
— $\nu_a\ll\mu$
— $\nu_s\perp\mu$.

(1) $\int_A df\mu=\nu(A)$ Voici un calcul :

$$\begin{aligned}\int_A f d\mu&=\int_A\sum_n f_n d\mu && (14.709a)\\ &=\sum_k\int_{A\cap X_k}\big(\sum_n f_n\big)d\mu_k && \text{cf.justif.} \quad (14.709b)\\ &=\sum_k\sum_n\int_{A\cap X_k}f_n d\mu_k && \text{cf. justif} \quad (14.709c)\\ &=\sum_k\int_{A\cap X_k}f_k d\mu_k && \text{cf. justif.} \quad (14.709d)\\ &=\sum_k\nu_{k,a}(A) && (14.709e)\\ &=\nu_a(A). && (14.709f)\end{aligned}$$

Justifications.

— Pour (14.709b). Les X_k sont disjoints et leur union fait Ω. Nous pouvons donc tranquillement couper l'intégrale en morceaux [77].
— Pour (14.709c). Nous permutons la somme et l'intégrale avec le corolaire 14.181 :
— Pour (14.709d). Vu que $f_n=0$ sur X_k tant que $k\neq n$, tous les termes de la somme sur n sont nuls sauf le terme $n=k$.

Voila déjà une belle première vérification de faite.

(2) $\nu=\nu_a+\nu_s$

La proposition 10.104 nous permet de fusionner le sommes.

$$\begin{aligned}(\nu_a+\nu_s)(A)&=\sum_n\nu_{n,a}(A)+\sum_n\nu_{n,s}(A)\\ &=\sum_n\Big(\nu_{n,a}(A)+\nu_{n,s}(A)\Big)=\sum_n\nu_n(A)=\nu(A).\end{aligned}\tag{14.710}$$

77. Ce découpage est également de l'invention personnelle. Quadruplez de vigilance.

(3) $\nu_a \ll \mu$
Supposons que $\mu(A) = 0$. Vu que l'union $\Omega = \bigcup_n X_n$ est disjointe, nous avons

$$0 = \mu(A) = \sum_{n\geqslant 1} \mu(A \cap X_n) = \sum_n \mu_n(A). \tag{14.711}$$

Nous en déduisons que $\mu_n(A) = 0$ pour tout n parce que μ_n est une mesure positive. Étant donné que $\mu_{n,a} \ll \mu_n$, nous avons $\nu_{n,a}(A) = 0$ pour chaque n. En faisant la somme,

$$\nu_a(A) = \sum_n \nu_{n,a}(A) = 0. \tag{14.712}$$

(4) $\nu_s \perp \mu$ Nous décomposons en petites parties.

(4a) $\nu_{n,s} \perp \mu_k$ **pour tout** n **et** k Nous vérifions la définition 14.221 avec $E = X_k$ et $F = X_n$, qui sont disjoints. Nous avons d'une part que $\mu_k(A) = \mu(A \cap X_k)$, et d'autre part que $\mu_k(A \cap X_k) = \mu(A \cap X_k \cap X_k) = \mu(A \cap X_k)$. Donc $\mu_k(A) = \mu_k(A \cap X_k)$.
Et d'autre part [78],

$$\begin{aligned}
\nu_{n,s}(A) &= \nu_n(A) - \nu_{n,a}(A) && (14.713a)\\
&= \nu(A \cap X - n) - \int_A f_n d\mu_n && (14.713b)\\
&= \nu(A \cap X - n) - \int_{A\cap X_n} f_n d\mu_n \qquad f_n = 0 \text{ sur } \Omega\backslash X_n. && (14.713c)\\
&= \nu(A \cap X_n) - \nu_{n,a}(A \cap X_n). && (14.713d)
\end{aligned}$$

Le membre de droite ne change pas si nous prenons $A \cap X_n$ au lieu de A. Donc le membre de gauche non plus. Nous en déduisons que $\nu_{n,s}(A) = \nu_{n,s}(A \cap X_n)$.

(4b) $\nu_{n,s} \perp \mu$ Appliquer le lemme 14.223 à $\mu = \sum_n \mu_n$.

(4c) $\nu_s \perp \mu$ Appliquer le lemme 14.223 à $\nu_s = \sum_n \nu_{n,s}$.

——— Existence, μ positive σ-finie, ν signée, $|\nu|$ σ-finie ———

Nous sommes enfin dans le cas général. Vu que $|\nu|$ est σ-finie, les deux parties de la décomposition de Jordan [79] (ν_-, ν_+) de ν sont des mesures positives σ-finies. Nous pouvons donc appliquer tout ce que nous venons de voir aux mesures ν_- et ν_+ séparément : nous avons des mesures positives $\nu_{+,a}$, $\nu_{+,s}$, ν_-, a et ν_-, s telles que

$$\begin{cases}
\nu_{+,a} + \nu_{+,s} = \nu_+ & (14.714a)\\
\nu_{+,a} \ll \mu & (14.714b)\\
\nu_{+,s} \perp \mu & (14.714c)\\
\nu_{-,a} + \nu_{-,s} = \nu_- & (14.714d)\\
\nu_{-,a} \ll \mu & (14.714e)\\
\nu_{-,s} \perp \mu. & (14.714f)
\end{cases}$$

Nous posons alors

$$\begin{aligned}
\nu_a &= \nu_{+,a} - \nu_{-,a} && (14.715a)\\
\nu_s &= \nu_{+,s} - \nu_{-,s}, && (14.715b)
\end{aligned}$$

et nous vérifions que toutes les conditions sont respectée. D'abord

$$\nu_a + \nu_s = \nu_{+,a} - \nu_{-,a} + \nu_{+,s} - \nu_{-,s} = \nu_+ - \nu_- = \nu. \tag{14.716}$$

Ensuite si $\nu_{+,a} \ll \mu$ et $\nu_{-,a} \ll \mu$, nous avons $\nu_a = \nu_{+,a} - \nu_{-,a} \ll \mu$ parce que le fait d'être dominée passe à la somme finie [80]. Et enfin, le lemme 14.224 assure que $\nu_a \perp \mu$. □

78. Invention personnelle, octuplez de vigilance.
79. Proposition 14.234.
80. Je n'ai pas dit que ça ne passait pas aux sommes dénombrables. Je suis prêt à parier que non, mais j'y ai pas vraiment réfléchi.

CorZDkhwS

Corolaire 14.236.
Si μ est une mesure σ-finie dominée par la mesure σ-finie m, alors μ possède une unique fonction de densité.

CorDomDens

Corolaire 14.237.
Soient μ et m, deux mesures positives σ-finies sur $(\Omega, \mathcal{A})$. Alors m domine μ si et seulement si μ possède une densité par rapport à m.

Démonstration. Si μ est dominée par m, alors la décomposition $\mu = \mu + 0$ satisfait le théorème de Radon-Nikodym. Par conséquent il existe une fonction f telle que

$$\mu(A) = \int_A f dm. \tag{14.717}$$

Cette fonction est alors une densité pour μ par rapport à m.

Pour la réciproque, nous supposons que μ a une densité f par rapport à m, et que A est un ensemble de m-mesure nulle :

$$m(A) = \int_\Omega \mathbb{1}_A dm = 0. \tag{14.718}$$

Cela signifie que la fonction $\mathbb{1}_A$ est m-presque partout nulle. La fonction produit $\mathbb{1}_A f$ est également nulle m-presque partout, et par conséquent

$$\mu(A) = \int_\Omega \mathbb{1}_A f dm = 0. \tag{14.719}$$

□

¡¡ Avertissement/question à la lectrice !! 14.238
Est-ce que la démonstration de cela ne demande pas la convergence monotone d'une façon ou d'une autre ?

14.10.2 Mesure complexe

DefGKHLooYjocEt

Définition 14.239 (Mesure complexe[466])**.**
*Si $(\Omega, \mathcal{A})$ est un espace mesurable, une **mesure complexe** est une application $\mu : \mathcal{A} \to \mathbb{C}$ telle que*

(1) $\mu(\varnothing) = 0$,

(2) La σ-additivité : si les éléments $A_i \in \mathcal{A}$ sont disjoints, alors $\sum_i \mu(A_i) = \mu(\bigcup_i A_i)$.

Notons que la série $\sum_i \mu(A_i)$ est alors nécessairement absolument convergente. En effet changer l'ordre de la somme ne change pas l'union, et donc ne change pas la valeur de la somme. Si $\sigma : \mathbb{N} \to \mathbb{N}$ est une permutation,

$$\sum_i \mu(A_{\sigma(i)}) = \mu\Big(\bigcup_i A_{\sigma(i)}\Big) = \mu\Big(\bigcup_i A_i\Big) = \sum_i \mu(A_i). \tag{14.720}$$

Le théorème 10.109 dit alors que la somme doit être absolument convergente.

ThoZZMGooKhRYaO

Théorème 14.240 (Radon-Nikodym complexe [81])**.**
Soit μ une mesure positive sur $(\Omega, \mathcal{A})$ et ν une mesure complexe. Alors

(1) Il existe un unique couple de mesures complexes ν_a, ν_s sur $(\Omega, \mathcal{A})$ tel que

(1a) $\nu = \nu_a + \nu_s$

(1b) $\nu_a \ll \mu$

81. L'histoire du nom de ce théorème est intéressante. Lorsque monsieur et madame Rèmederdonnukodym apprirent que leurs amis, les Rèmedelaboulechevelue avaient appelé leur fils Théo, ils décidèrent d'en faire autant. C'est en souvenir de ces circonstances que monsieur Nikodym (prénommé Radon) décida de faire des math.

(1c) $\nu_s \perp \mu$.

(2) *Ces mesures satisfont alors* $\nu_a \perp \nu_s$.

(3) *Il existe une fonction intégrable* $h: \Omega \to \mathbb{C}$ *telle que* $\nu_a = h\mu$.

(4) *La fonction* h *est unique à* μ*-équivalence près.*

(5) *Si de plus* $\nu \ll \mu$ *alors* $\nu = h\mu$.

ItemDIXOooFqOkgGv

Démonstration. No proof. □

RemSYRMooZPBhbQ

Remarque 14.241.
Le point (5) est souvent utilisé sous la forme

$$\nu(A) = \int_\Omega \mathbb{1}_A(\omega)h(\omega)d\mu(\omega) = \int_A h(\omega)d\mu(\omega). \tag{14.721}$$

14.10.3 Théorème d'approximation

ThoAFXXcVa

Théorème 14.242 (Théorème d'approximation, thème 24[447])**.**
Soit un espace topologique métrique (Ω, d). *Nous considérons sa tribu des boréliens* $\mathcal{B}or(\Omega)$ *ainsi qu'une mesure* μ *sur* $\big(\Omega, \mathcal{B}or(\Omega)\big)$.

Soient un ouvert $W \subset \Omega$ *tel que* $\mu(W) < \infty$ *et un un borélien* A *tel que* $A \subset W$. *Soit aussi* $\epsilon > 0$.

Il existe un fermé $F \subset W$ *et une fonction* $f \in C^0(\Omega, \mathbb{R})$ *vérifiant*

(1) $F \subset A \subset W$,

ITEMooOZVJooSViuds

(2) $f|_F = 1$,

ITEMooIEFSooHXYZrK

(3) $f|_{W^c} = 0$

ITEMooSOQVooBbvfgy

(4) $\|f - \mathbb{1}_A\|_{L^1} < \epsilon$

Démonstration. Par le lemme 14.87, il existe un fermé F et un ouvert V tels que

$$F \subset A \subset V \subset W \tag{14.722}$$

et $\mu(V \backslash F) < \epsilon$. Nous posons alors

$$f(x) = \frac{d(x, V^c)}{d(x, V^c) + d(x, F)}. \tag{14.723}$$

Le dénominateur de cette expression ne s'annule jamais parce que si $d(x, V^c) = 0$, c'est que $x \in V^c$. Mais alors x n'est pas dans V et donc pas dans F non plus. La partie F étant fermée, $d(x, F) > 0$ par lemme 7.152. De plus la fonction f est continue par le lemme 7.153.

(1) **Pour (2)** Si $x \in F$, alors $d(x, F) = 0$, et f devient

$$f(x) = \frac{d(x, V^c)}{d(x, V^c)} = 1 \tag{14.724}$$

(2) **Pour (3)** Si $x \in W^c$, alors $x \in V^c$ et $d(x, V^c) = 0$ si bien que $f(x) = 0$.

(3) **Pour (4)** Les premiers points montrent que

$$\mathbb{1}_F \leqslant f \leqslant \mathbb{1}_V. \tag{14.725}$$

Mais nous avons aussi, par ailleurs,

$$\mathbb{1}_F \leqslant \mathbb{1}_A \leqslant \mathbb{1}_V. \tag{14.726}$$

Ces deux encadrement, par le lemme 1.447 donnent l'encadrement

$$|f - \mathbb{1}_A| \leqslant \mathbb{1}_V - \mathbb{1}_F. \tag{14.727}$$

En ce qui concernent les intégrales nous avons alors

$$\begin{aligned}\int_\Omega |\mathbb{1}_A - f| &\leqslant \int_\Omega (\mathbb{1}_V - \mathbb{1}_F) d\mu && \text{(14.728a)}\\ &= \mu(V) - \mu(F) && \text{(14.728b)}\\ &< \epsilon. && \text{(14.728c)}\end{aligned}$$

SUBEQooVJDXooFtCelQ

Pour (14.728b), c'est le lemme 14.175.

□

14.11 Produit de mesures

Si vous cherchez le produit d'une mesure par une fonction, c'est la définition 14.212.

LemAQmWEmN

Lemme 14.243 (Propriété des sections[448])**.**
Soient $\mathcal{A}_1$ et $\mathcal{A}_2$ des tribus sur les ensembles Ω_1 et Ω_2. Si $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ alors pour tout $x \in \Omega_1$ et $y \in \Omega_2$, les ensembles

subEqCTtPccK

$$A_1(y) = \{x \in \Omega_1 \text{ tel que } (x, y) \in A\} \tag{14.729a}$$
$$A_2(x) = \{y \in \Omega_2 \text{ tel que } (x, y) \in A\} \tag{14.729b}$$

sont mesurables.

Démonstration. Soit $y \in \Omega_2$; nous allons prouver le résultat pour $A_1(y)$. Pour cela nous notons

$$S = \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 \text{ tel que } \forall y \in \Omega_2, A_1(y) \in \mathcal{A}_1\}, \tag{14.730}$$

et nous allons noter que S est une tribu contenant les rectangles. Par conséquent, S sera égal à $\mathcal{A}_1 \otimes \mathcal{A}_2$.

(1) **Les rectangles**
Considérons le rectangle $A = X \times Y$ et si $y \in \Omega_2$ alors

$$A_1(y) = \{x \in \Omega_1 \text{ tel que } (x, y) \in X \times Y\}. \tag{14.731}$$

Donc soit $y \in Y$ alors $A_1(y) = X \in \mathcal{A}_1$, soit $y \notin Y$ et alors $A_1(y) = \varnothing \in \mathcal{A}_1$.

(2) **Tribu : ensemble complet**
Nous avons $\Omega_1 \times \Omega_2 \in S$ parce que c'est un rectangle.

(3) **Tribu : complémentaire** Soit $A \in S$. Montrons que $A^c \in S$. Nous avons d'abord

$$(A^c)_1(y) = \{x \in \Omega_1 \text{ tel que } (x, y) \in A^c\}. \tag{14.732}$$

D'autre part

$$A_1(y)^c = \{x \in \Omega_1 \text{ tel que } (x, y) \notin A\} = \{x \in \Omega_1 \text{ tel que } (x, y) \in A^c\} = (A^c)_1(y). \tag{14.733}$$

Vu que $\mathcal{A}_1$ est une tribu et que par hypothèse $A_1(y) \in \mathcal{A}_1$, nous avons aussi $A_1(y)^c \in S$, et donc $(A^c)_1(y) \in \mathcal{A}_1$, ce qui prouve que $A^c \in S$.

(4) **Tribu : union dénombrable** Soit une suite $A_n \in S$. Nous avons

$$(\bigcup_n A_n)_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in \bigcup_n A_n\} \tag{14.734a}$$
$$= \bigcup_n \{x \in \Omega_1 \text{ tel que } (x,y) \in A_n\} \tag{14.734b}$$
$$= \bigcup_n (A_n)_1(y), \tag{14.734c}$$

et ce dernier ensemble est dans $\mathcal{A}_1$ parce que c'est une union dénombrable d'éléments de $\mathcal{A}_1$.

Nous avons donc prouvé que S est une tribu contenant les rectangles, donc S contient au moins $\mathcal{A}_1 \otimes \mathcal{A}_2$. □

Corolaire 14.244.
Si $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$ est une fonction mesurable[82] *sur $X \times Y$ alors pour chaque y dans Ω_2, la fonction*

$$\begin{aligned} f_y\colon X &\to \mathbb{R} \\ x &\mapsto f(x,y) \end{aligned} \tag{14.735}$$

est mesurable.

Démonstration. Soit $\mathcal{O}$ un ensemble mesurable de $\mathbb{R}$ (i.e. un borélien), et $y \in \Omega_2$. Nous avons

$$f_y^{-1}(\mathcal{O}) = \{x \in X \text{ tel que } f(x,y) \in \mathcal{O}\} = A_1(y) \tag{14.736}$$

où

$$A = \{(x,y) \in \Omega_1 \times \Omega_2 \text{ tel que } f(x,y) \in \mathcal{O}\} = f^{-1}(\mathcal{O}). \tag{14.737}$$

Ce dernier est mesurable parce que f l'est. □

ThoCCIsLhO

Théorème 14.245 ([448][83])**.**
Soient $(\Omega_i, \mathcal{A}_i, \mu_i)$ $(i = 1, 2)$ deux espaces mesurés σ-finis. Soit $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Alors les fonctions[84]

$$x \mapsto \mu_2\big(A_2(x)\big) \tag{14.738a}$$
$$y \mapsto \mu_1\big(A_1(y)\big) \tag{14.738b}$$

sont mesurables et

$$\int_{\Omega_1} \mu_2\big(A_2(x)\big) d\mu_1(x) = \int_{\Omega_2} \mu_1\big(A_1(y)\big) d\mu_2(y). \tag{14.739}$$

EqRKXwsQJ

Démonstration. Nous supposons d'abord que μ_1 et μ_2 sont finies et nous notons $\mathcal{D}$ le sous-ensemble de $\mathcal{A}_1 \otimes \mathcal{A}_2$ sur lequel le théorème est correct. Nous allons commencer par prouver que $\mathcal{D}$ est un λ-système.

(1) **λ-système : différence ensembliste** Soient $A, B \in \mathcal{D}$ avec $A \subset B$. Nous avons

$$(B\backslash A)_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in B\backslash A\} \tag{14.740a}$$
$$= \{x \in \Omega_1 \text{ tel que } (x,y) \in B\}\backslash\{x \in \Omega_1 \text{ tel que } (x,y) \in A\} \tag{14.740b}$$
$$= B_1(y)\backslash A_1(y). \tag{14.740c}$$

Vu que $A_1(y) \subset B_1(y)$ et que les mesures sont finies le lemme 14.23 nous donne

$$\mu_1\big((B\backslash A)_1(y)\big) = \mu_1\big(B_1(y)\big) - \mu_1\big(A_1(y)\big), \tag{14.741}$$

82. Définition 14.44.
83. Modèle non contractuel : des notations et la définition de λ-système peuvent varier entre la référence et le présent texte.
84. Voir la notation du lemme 14.729.

et similairement pour $1 \leftrightarrow 2$. Les deux fonctions (de y) à droite étant mesurables, nous avons la mesurabilité de la fonction $y \mapsto \mu_1\big((B\backslash A)_1(y)\big)$.

Prouvons la formule intégrale en nous rappelant que la formule (14.739) est supposée correcte pour A et B séparément :

$$\int_{\Omega_2} \mu_1\big((B\backslash A)_1(y)\big)d\mu_2(y) = \int_{\Omega_2} \mu_1\big(B_1(y)\big)d\mu_2(y) - \int_{\Omega_2} \mu_1\big(A_1(y)\big)d\mu_2(y) \tag{14.742a}$$

$$= \int_{\Omega_1} \mu_2\big(B_2(x)\big)d\mu_1(x) - \int_{\Omega_1} \mu_2\big(A_2(x)\big)d\mu_1(x) \tag{14.742b}$$

$$= \int_{\Omega_1} \mu_2\big((B\backslash A)_2(x)\big)d\mu_1(x). \tag{14.742c}$$

(2) **λ-système : limite de suite croissante**

Soit (A_n) une suite croissante dans $\mathcal{D}$; nous posons $B_n = A_n\backslash A_{n-1}$ et $A_0 = \varnothing$ de telle sorte à travailler avec une suite d'ensembles disjoints qui satisfait $\bigcup_n A_n = \bigcup_n B_n$. Vu que la suite est croissante nous avons $A_{n-1} \subset A_n$ et donc $B_n \in \mathcal{D}$ par le point déjà fait sur la différence ensembliste. Nous avons :

$$(\bigcup_n B_n)_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in \bigcup_n B_n\} \tag{14.743a}$$

$$= \bigcup_n \{x \in \Omega_1 \text{ tel que } (x,y) \in B_n\} \tag{14.743b}$$

$$= \bigcup_n (B_n)_1(y). \tag{14.743c}$$

Par conséquent, par la propriété (2) d'une mesure nous avons

$$\mu_1\big((\bigcup_n B_n)_1(y)\big) = \sum_n \mu_1\big((B_n)_1(y)\big). \tag{14.744}$$

En tant que somme de fonctions positives et mesurables, la fonction

$$y \mapsto \sum_n \mu_1\big((B_n)_1(y)\big) \tag{14.745}$$

est mesurable par la proposition 14.101. Il faut encore vérifier la formule intégrale. Le gros du boulot est de permuter une somme et une intégrale par le corolaire 14.181 :

$$\int_{\Omega_2} \sum_n \mu_1\big((B_n)_1(y)\big)d\mu_2(y) = \sum_n \int_{\Omega_2} \mu_1\big((B_n)_1(y)\big)d\mu_2(y) \tag{14.746a}$$

$$= \sum_n \int_{\Omega_1} \mu_2\big((B_n)_2(x)\big)d\mu_1(x) \tag{14.746b}$$

$$= \int_{\Omega_1} \sum_n \mu_2\big((B_n)_2(x)\big)d\mu_1(x) \tag{14.746c}$$

$$= \int_{\Omega_1} \mu_2\big((\bigcup_n B_n)_1(y)\big)d\mu_1(x). \tag{14.746d}$$

Maintenant que $\mathcal{D}$ est un λ-système contenant les rectangles, le lemme 14.33 dit que la tribu engendrée par $\mathcal{D}$ (c'est-à-dire $\mathcal{A}_1 \otimes \mathcal{A}_2$) est le λ-système $\mathcal{D}$ lui-même.

La preuve est finie dans le cas de mesures finies. Nous commençons maintenant à prouver dans le cas où les mesures μ_1 et μ_2 sont seulement σ-finies. Nous considérons des suites croissantes $\Omega_{i,n} \to \Omega_i$ d'ensembles mesurables et de mesure finie : $\mu_i(\Omega_{i,n}) < \infty$. D'abord remarquons que

$$\mu_2\Big((A \cap \Omega_{1,j} \times E_{2,j})_2(x)\Big) = \mu_2\Big(A_2(x) \cap \Omega_{2,j}\Big)\mathbb{1}_{\Omega_{1,j}}. \tag{14.747}$$

EqNFuBzBF

En effet,

$$\begin{aligned}\heartsuit &= (A\cap\Omega_{1,j}\times E_{2,j})_2(x) && (14.748a)\\ &= \{y\in\Omega_2 \text{ tel que } (x,y)\in A\cap\Omega_{1,j}\times E_{2,j}\} && (14.748b)\\ &= \{y\in\Omega_2 \text{ tel que } (x,y)\in A\times\Omega_{2,j}\}\cap\{y\in\Omega_2 \text{ tel que } (x,y)\in\Omega_{1,j}\times\Omega_{2,j}\}. && (14.748c)\end{aligned}$$

Si $y\in\Omega_{1,j}$ alors $\{y\in\Omega_2 \text{ tel que } (x,y)\in\Omega_{1,j}\times\Omega_{2,j}\}=\Omega_{2,j}$ et dans ce cas

$$\heartsuit=\{y\in\Omega_2 \text{ tel que } (x,y)\in A\times\Omega_{2,j}\}\cap\Omega_{2,j}=A_2(x)\cap E_{2,j}. \tag{14.749}$$

Et inversement, si $x\notin\Omega_{1,j}$ alors $\heartsuit=\varnothing$. Dans les deux cas nous avons (14.747).

Les ensembles $A\cap\Omega_{1,j}\times\Omega_{2,j}$ étant de mesure finie, nous pouvons leur appliquer la première partie :

$$\int_{\Omega_1}\mu_2\Big((A\cap\Omega_{1,j}\times\Omega_{2,j})_2(x)\Big)d\mu_1(x)=\int_{\Omega_2}\mu_1\Big((A\cap\Omega_{1,j}\times\Omega_{2,j})_1(y)\Big)d\mu_2(u), \tag{14.750}$$

ou encore

$$\int_{\Omega_1}\mu_2\Big(A_2(x)\cap\Omega_{2,j}\Big)\mathbb{1}_{\Omega_{1,j}}(x)d\mu_1(x)=\int_{\Omega_2}\mu_1\Big(A_1(y)\cap\Omega_{1,j}\Big)\mathbb{1}_{\Omega_{2,j}}(y)d\mu_2(y). \tag{14.751}$$

Ce que nous avons dans ces intégrales sont (par rapport à j) des suites croissantes de fonction positives ; nous pouvons donc permuter une limite et une intégrale. En sachant que si $k\to\infty$, alors

$$\begin{aligned}\mathbb{1}_{1,j}(x)&\to 1 && (14.752a)\\ \mu_2\big(A_2(x)\cap\Omega_2,j\big)&\to\mu_2\big(A_2(x)\big), && (14.752b)\end{aligned}$$

nous trouvons le résultat demandé. □

ThoWWAjXzi

Théorème-Définition 14.246 ([467, 468])**.**
Soient μ_i des mesures σ-finies sur[85] *$(\Omega_i,\mathcal{A}_i)$ $(i=1,2)$.*

ITEMooMMAAooVnHmzc

(1) Il existe une et une seule mesure, notée $\mu_1\otimes\mu_2$, sur $(\Omega_1\times\Omega_2,\mathcal{A}_1\otimes\mathcal{A}_2)$ telle que

$$(\mu_1\otimes\mu_2)(A_1\times A_2)=\mu_1(A_1)\mu_2(A_2) \tag{14.753}$$

EqQIuWLQU

pour tout $A_1\in\mathcal{A}_1$ et $A_2\in\mathcal{A}_2$.

(2) Cette mesure est donnée par la formule[86]

$$(\mu_1\otimes\mu_2)(A)=\int_{\Omega_1}\mu_2\big(A_2(x)\big)d\mu_1(x)=\int_{\Omega_2}\mu_1\big(A_1(y)\big)d\mu_2(y) \tag{14.754}$$

EqDFxuGtH

où $A_1(y)$ et $A_2(x)$ sont définis en (14.729).
*Cette mesure est la **mesure produit** de μ_1 par μ_2.*

(3) La mesure $\mu_1\otimes\mu_2$ ainsi définie est σ-finie.

ITEMooUOATooQpdcND

(4) Le produit de mesures est associatif : si μ_1, μ_2 et μ_3 sont des mesures, alors σ-finies, alors

$$(\mu_1\otimes\mu_2)\otimes\mu_3=\mu_1\otimes(\mu_2\otimes\mu_3). \tag{14.755}$$

Démonstration. La partie « existence » sera divisée en deux parties : l'une pour prouver que les formules (14.754) donnent une mesure et une pour montrer que cette mesure vérifie la condition (14.753).

85. La tribu produit est la définition 14.127.
86. Voir les notations du lemme 14.243.

(1) **Unicité**
L'ensemble des rectangles de $\Omega_1 \times \Omega_2$ engendre la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$, est fermé par intersection et contient une suite croissante d'ensembles $P_n \times R_n$ de mesure finie ($\mu(P_n \times R_n) < \infty$) telle que $P_n \times R_n \to \Omega_1 \times \Omega_2$. Cette suite est donnée par le fait que μ_1 et μ_2 sont σ-finies. En effet si (X_n) et (Y_n) sont des recouvrements dénombrables de Ω_1 et Ω_2 par des ensembles de mesure finie, en posant $P_n = \bigcup_{k=1}^{n} X_k$ et $R_n = \bigcup_{k=1}^{n} Y_k$ nous avons bien une suite croissante de rectangles qui tendent vers $\Omega_1 \times \Omega_2$. Avec ces rectangles en main, le théorème 14.34 donne l'unicité.

(2) **Les formules définissent une mesure** Le théorème 14.245 dit que ces formules ont un sens et que l'égalité entre les deux intégrales est correcte. Nous prouvons à présent qu'elles déterminent effectivement une mesure sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$.
Pour tout $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $\mu(A) \geqslant 0$ parce que μ est donnée par l'intégrale d'une fonction positive.
En ce qui concerne la condition d'unions dénombrable disjointe, soient $A^{(i)}$ des éléments disjoints de $\mathcal{A}_1 \otimes \mathcal{A}_2$; nous commençons par remarquer que

$$\left(\bigcup_{i=1}^{\infty} A^{(i)}\right)_2 (x) = \{y \in \Omega_2 \text{ tel que } (x,y) \in \bigcup_{i=1}^{\infty} A^{(i)}\} \tag{14.756a}$$
$$= \bigcup_{i=1}^{\infty} \{y \in \Omega_2 \text{ tel que } (x,y) \in A^{(i)}\} \tag{14.756b}$$
$$= \bigcup_{i=1}^{\infty} A_2^{(i)}(x). \tag{14.756c}$$

Par conséquent,

$$\mu\left(\bigcup_{i=1}^{\infty} A^{(i)}\right) = \int_{\Omega_1} \mu_2\left(\left(\bigcup_{i=1}^{\infty} A^{(i)}\right)_2 (x)\right) d\mu_1(x) \tag{14.757a}$$
$$= \int_{\Omega_1} \sum_{i=1}^{\infty} \mu_2\big(A_2^{(i)}(x)\big) d\mu_1(x) \tag{14.757b}$$
$$= \int_{\Omega_1} \lim_{n\to\infty} \sum_{i=1}^{n} \mu_2\big(A_2^{(i)}(x)\big) d\mu_1(x). \tag{14.757c}$$

où nous avons utilisé l'additivité de la mesure μ_2. À ce niveau, il serait commode de permuter la somme et l'intégrale. Pour ce faire nous considérons la suite (croissante) de fonctions

$$f_n(x) = \sum_{i=1}^{n} \mu_2\big(A_2^{(i)}(x)\big). \tag{14.758}$$

Nous pouvons permuter la limite et l'intégrale grâce au théorème de la convergence monotone 14.178 ; ensuite la somme se permute avec l'intégrale en tant que somme finie :

$$\mu\left(\bigcup_{i=1}^{\infty} A^{(i)}\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \int_{\Omega_1} \big(A_2^{(i)}(x)\big) d\mu_1(x) \tag{14.759a}$$
$$= \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A^{(i)}) \tag{14.759b}$$
$$= \sum_{i=1}^{\infty} \mu(A^{(i)}). \tag{14.759c}$$

(3) **Elles vérifient la condition** Prouvons que les formules (14.754) se réduisent à (14.753) dans le cas des rectangles. Soit donc $A = X_1 \times X_2$ avec $X_i \in \mathcal{A}_i$. Alors

$$A_1(y) = \{x \in \Omega_1 \text{ tel que } (x,y) \in X_1 \times X_2\} \tag{14.760}$$

et

$$\mu_1\big(A_1(y)\big) = \mathbb{1}_{X_2}(y)\mu_1(X_1), \tag{14.761}$$

donc

$$\begin{aligned}(\mu_1 \otimes \mu_2)(A) &= \int_{\Omega_2} \mu_1\big(A_1(y)\big)d\mu_2(y) && \text{(14.762a)}\\ &= \int_{\Omega_2} \mu_1(X_1)\mathbb{1}_{X_2}(y)d\mu_2(y) && \text{(14.762b)}\\ &= \mu_1(X_1)\int_{\Omega_2} \mathbb{1}_{X_2}(y)d\mu_2(y) && \text{(14.762c)}\\ &= \mu_1(X_1)\mu_2(X_2). && \text{(14.762d)}\end{aligned}$$

Pour cela nous avons utilisé le fait que l'intégrale de la fonction caractéristique d'un ensemble mesurable est la mesure de cet ensemble.

□

DefUMlBCAO

Définition 14.247 (Produit d'espaces mesurés)**.**
*Si $(\Omega_i, \mathcal{A}_i, \mu_i)$ sont deux espaces mesurés, l'**espace produit** est l'ensemble $\Omega_1 \times \Omega_2$ muni de la tribu produit $\mathcal{A}_1 \otimes \mathcal{A}_2$ de la définition 14.127 et de la mesure produit $\mu_1 \otimes \mu_2$ définie par le théorème 14.246.*

Remarque 14.248.
Il n'est pas garanti que la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$ soit la tribu la plus adaptée à l'ensemble $S_1 \times S_2$. Dans le cas de $\mathbb{R}^N$, il se fait que c'est le cas : en prenant des produits des boréliens sur $\mathbb{R}$ on obtient bien les boréliens sur $\mathbb{R}^N$, voir proposition 14.132.

14.12 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$

DEFooSWJNooCSFeTF

Définition 14.249 (Mesure de Lebesgue)**.**
En plusieurs étapes.

(1) D'abord nous avons la mesure λ_N sur $\mathbb{R}^N$ définie sur

$$\big(\mathbb{R}^N, \mathcal{B}or(\mathbb{R}) \otimes \ldots \otimes \mathcal{B}or(\mathbb{R})\big) \tag{14.763}$$

comme le produit $\lambda \otimes \ldots \otimes \lambda$ via la définition 14.247 (et grâce au fait que le produit est associatif, théorème 14.246(4)).

(2) Ensuite nous nous souvenons du corolaire 14.132 qui donne λ_N comme une mesure sur

$$\big(\mathbb{R}^N, \mathcal{B}or(\mathbb{R}^N)\big). \tag{14.764}$$

(3) Et enfin nous considérons la complétion de la mesure λ_N (théorème 14.69), que nous notons encore λ_N.

PropSKXGooRFHQst

Proposition 14.250 ([451])**.**
Tout ouvert de $\mathbb{R}^n$ est une union dénombrable et disjointe de cubes semi-ouverts.

Démonstration. Nous allons même montrer que ces cubes peuvent être choisis sur un quadrillage.

Soit G un ouvert de $\mathbb{R}^n$. Soit $\{Q_i^1\}_{i\in\mathbb{N}}$ un découpage de $\mathbb{R}^n$ en cubes semi-ouverts de côté 1 et dont les sommets sont en les coordonnées entières. Ils sont de la forme

$$\prod_{i=1}^{n} [n_i, n_i + 1[\tag{14.765}$$

où les n_i sont des entiers. Ce sont des cubes disjoints. Nous considérons ensuite pour chaque $k > 1$ le découpage $\{Q_i^{(k)}\}_{i\in\mathbb{N}}$ de $\mathbb{R}^n$ en cubes de côtés 2^{-k} qui consiste à découper en 2 les côtés des

cubes du découpage $Q^{(k-1)}$. Ces cubes forment encore un découpage dénombrable de $\mathbb{R}^n$ en des cubes disjoints. Ils sont de la forme

$$\prod_{i=1}^{n}[\frac{n_i}{2^k},\frac{n_i+1}{2^k}[\tag{14.766}$$

où les n_i sont encore entiers. Ensuite nous considérons $\mathcal{E}$ l'union de tous les $Q_i^{(k)}$ contenus dans G.

Montrons que $\mathcal{E}=G$. D'abord $\mathcal{E}\subset G$ parce que $\mathcal{E}$ est une union d'ensembles contenus dans G. Ensuite si $x\in G$, il existe une boule de rayon r autour de x contenue dans G ; alors un des ensembles $Q_i^{(k)}$ avec $2^{-j}<\frac{r}{2}$ est contenue dans $B(x,r)$ et donc dans $\mathcal{E}$.

Bien entendu l'union qui donne $\mathcal{E}$ n'est pas satisfaisante par ce que les $Q_i^{(k+1)}$ sont contenus dans les $Q_i^{(k)}$; les intersections sont donc loin d'être vides.

Nous faisons ceci :

$$R^{(0)}=\{Q_i^{(1)}\text{contenu dans } G\} \tag{14.767a}$$
$$R^{(k+1)}=\{Q_i^{(k+1)}\text{contenus dans } G \text{ et pas dans } R^{(k)}\}. \tag{14.767b}$$

En fin de compte l'union de tous les ensembles contenus dans les $R^{(k)}$ forment encore $\mathbb{R}^n$, mais sont d'intersection vide. □

Les cubes dont il est question dans cette preuve, de côtés 2^{-k} sont souvent appelés des cubes **dyadiques**.

14.12.1 Propriétés d'unicité

CorMPDAooDJRrom

Corolaire 14.251.
La mesure λ_N est l'unique mesure sur $(\mathbb{R}^N,\mathcal{B}or(\mathbb{R}^N))$ à satisfaire

$$\mu(\prod_{i=1}^{N}[a_i,b_i])=\prod_{i=1}^{N}|a_i-b_i| \tag{14.768}$$

Démonstration. Par définition de la mesure produit, λ_N est l'unique mesure sur $(\mathbb{R}^N,\mathcal{B}\text{or}(\mathbb{R})\otimes\ldots\otimes\mathcal{B}\text{or}(\mathbb{R}))$ à satisfaire la condition. La proposition 14.132 conclut. □

Vu que les compacts de $\mathbb{R}^n$ sont les fermés bornés (théorème 9.24), et que tout borné est dans un tel produit d'intervalle, la mesure de Lebesgue est une mesure de Borel (définition 14.89(1)).

THOooTMWHooThsDHj

Théorème 14.252 ([469])**.**
La mesure de Lebesgue est invariante par translation. Autrement dit si A est mesurable dans $\mathbb{R}^n$ et si $a\in\mathbb{R}^n$ alors $A+a$ est mesurable et

$$\lambda_N(A+a)=\lambda_N(A). \tag{14.769}$$

Démonstration. Nous supposons que A est borélien ; sinon il l'est à ensemble négligeable près. Nous nommons μ la mesure donnée par

$$\mu(A)=\lambda_N(A+a). \tag{14.770}$$

Vu que

$$\mu(\prod_{n=1}^{N}[r_n,s_n[)=\lambda_N(\prod_i[r_n+a_n,s_n+a_n[)=\prod_i|s_n-r_n|. \tag{14.771}$$

Vu qu'il y a unicité de la mesure vérifiant cette propriété (corolaire 14.251), nous avons $\mu=\lambda_N$. □

Pour la suite nous notons Q_0 le cube unité de $\mathbb{R}^N$: $Q_0=\big([0,1[\big)^N$.

ThoCABFooHbUzWc

Théorème 14.253 ([469]).
Soit μ une mesure positive sur $\mathbb{R}^N$ telle que

(1) μ soit invariante par translation (des boréliens),

(2) $\mu(Q_0) = 1$.

Alors $\mu = \lambda_N$.

Démonstration. Pour simplifier l'écriture nous faisons $N = 2$. Notre but est de prouver que $\mu([0,r[\times [0,r'[) = rr'$ pour tout $r, r' \in \mathbb{R}$.

(1) **Longueur $=1/J$** Soient J, K des entiers. Nous pouvons diviser le cube Q_0 en rectangles de côtés $1/J$ et $1/K$:

$$Q_0 = \bigcup_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} [\frac{j-1}{J}, \frac{j}{J}[\times [\frac{k-1}{K}, \frac{k}{K}[\tag{14.772}$$

où l'union est disjointe. En ce qui concerne la mesure nous commençons par utiliser la sous-additivité :

$$\mu(Q_0) = \sum_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} \mu\left([\frac{j-1}{J}, \frac{j}{J}[\times [\frac{k-1}{K}, \frac{k}{K}[\right). \tag{14.773}$$

Nous utilisons ensuite, sur chacun des termes séparément l'invariance par translation selon les vecteurs $(\frac{j-1}{J}, 0)$ et $(0, \frac{k-1}{K})$:

$$1 = \mu(Q_0) = \sum_{\substack{1\leqslant j\leqslant J\\ 1\leqslant k\leqslant K}} \mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right) = JK\mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right), \tag{14.774}$$

et donc

$$\mu\left([0, \frac{1}{J}[\times [0, \frac{1}{K}[\right) = \frac{1}{J} \times \frac{1}{K}. \tag{14.775}$$

(2) **Longueur L/K**
Soient L, M des entiers et calculons :

$$\mu\left([\frac{0}{J}, \frac{L}{J}[\times [\frac{0}{K}, \frac{M}{K}[\right) = \sum_{\substack{0\leqslant l\leqslant L-1\\ 0\leqslant m\leqslant M-1}} \mu\left([\frac{l}{J}, \frac{l+1}{J}[\times [\frac{m}{K}, \frac{m+1}{K}[\right) \tag{14.776a}$$

$$= LM\mu\left([\frac{0}{J}, \frac{1}{J}[\times [\frac{0}{K}, \frac{1}{K}[\right) \tag{14.776b}$$

$$= LM \times \frac{1}{J} \times \frac{1}{K}. \tag{14.776c}$$

Nous avons donc, pour tout J, K, L, M :

$$\mu\left([0, \frac{L}{J}[\times [0, \frac{M}{K}[\right) = \frac{L}{J} \times \frac{M}{K}, \tag{14.777}$$

c'est-à-dire que pour tout $r, s \in \mathbb{Q}^+$ nous avons

$$\mu\big([0,r[\times [0,s[\big) = rs. \tag{14.778}$$

(3) **Longueur réelle** Nous passons au cas de longueur réelle. Soit $a > 0$ et une suite croissante de rationnels $r_n \to a$. Une telle suite existe par la proposition 9.17. L'intervalle $[0,a[$ s'écrit sous la forme d'une union croissante $[0,a[= \bigcup_{n\geqslant 1}[0,r_n[$; le lemme 14.24(1) peut être utilisé et nous avons

$$\mu\big([0,a[\big) = \mu\left(\bigcup_{n\geqslant 1}[0,r_n[\right) = \lim_{n\to\infty}\mu\big([0,r_n[\big) = \lim_{n\to\infty} r_n = a. \tag{14.779}$$

Enfin, si $a, a' \in \mathbb{R}$, l'invariance par translation donne

$$\mu\big([a,a'[\big) = \mu\big([0,a'-a[\big) = a'-a. \tag{14.780}$$

Par unicité de la mesure ayant cette propriété, nous avons $\mu = \lambda_N$. □

CorKGMRooHWOQGP

Corolaire 14.254.
Si μ est une mesure positive sur $\mathbb{R}^N$ invariante par translation et telle que $\mu(Q_0) = C < \infty$ alors $\mu = C\lambda_N$.

Démonstration. Si $C > 0$ nous considérons la mesure $\frac{1}{C}\mu$ qui vérifie $(\frac{1}{C}\mu)(Q_0) = 1$. En conséquence du théorème 14.253, $\frac{1}{C}\mu = \lambda_N$ et $\mu = C\lambda_N$.

Si au contraire $C = 0$ alors nous pouvons paver $\mathbb{R}^N$ avec des cubes Q_i de côté 1 qui ont tous mesure 0. Par conséquent, $\mathbb{R}^N = \bigcup_{i=1}^{\infty} Q_i$, donc $\mu(\mathbb{R}^N) = \sum_i \mu(Q_i) = 0$. Par conséquent $\mu = 0$ parce que toute partie de $\mathbb{R}^N$ a une mesure au maximum égale à celle de $\mathbb{R}^N$. □

14.12.2 Régularité

Les différentes notions de régularité pour une mesure sont données dans la définition 14.89. Ce sont essentiellement des questions de compatibilité entre la mesure et la topologie.

Proposition 14.255.
La mesure de Lebesgue est une mesure de Radon sur tout ouvert de $\mathbb{R}^N$.

Démonstration. Soit V un ouvert de $\mathbb{R}^N$. C'est localement compact et dénombrable à l'infini. Il suffit de prouver que λ_N est de Borel sur V pour que le théorème 14.90 conclue à la régularité de la mesure de Lebesgue.

Soit K un compact de V. Par la proposition 7.101 c'est également un compact de $\mathbb{R}^N$. Par conséquent K est dans un pavé fermé de $\mathbb{R}^N$ du type

$$K \subset \prod_{n=1}^{N} [a_n, b_n] \tag{14.781}$$

et donc en passant par le corolaire 14.251,

$$\lambda_N(K) \leqslant \prod_{i=1}^{N} (b_n - a_n) < \infty. \tag{14.782}$$

Nous avons démontré que λ_N reste fini sur tout compact de V. □

14.13 Propriétés de l'intégrale de Lebesgue

Un lemme qui a l'air de rien, mais qui au final est souvent utilisé ; tellement qu'on l'oublie un peu.

LEMooWKSWooPptdEm

Lemme 14.256 ([1]).
Soit un compact K de $\mathbb{R}$ et une fonction continue $f\colon \mathbb{R} \to \mathbb{R}$. Alors l'intégrale

$$\int_K f \tag{14.783}$$

existe et est finie.

Démonstration. Vu que f est continue sur le compact K, elle y atteint une borne supérieure [87] que nous nommons M.

87. Nous ne nous lasserons jamais de citer le théorème de Weierstrass 7.146.

Soit R tel que $B(0,R)$ contienne K. La fonction $(M+1)\mathbb{1}_{B(0,R)}$ majore strictement f sur le mesurable $B(0,R)$. L'ensemble sur lequel nous prenons le supremum dans la définition (14.468) de l'intégrale de f est majoré par le nombre fini $(M+1)\mu(K)$. Le supremum existe et est fini (proposition 1.461). □

Le lemme suivant est la contrepartie du côté des intégrales de l'invariance par translation de la mesure de Lebesgue démontrée dans le théorème 14.252.

LEMooGKOGooPLYaUO

Lemme 14.257 (Invariance par translation)**.**
Soient f intégrable sur $\mathbb{R}^d$ et $a \in \mathbb{R}^d$. Alors en posant

$$\begin{aligned} g\colon \mathbb{R}^d &\to \mathbb{C} \\ x &\mapsto f(x+a) \end{aligned} \tag{14.784}$$

nous avons

$$\int_{\mathbb{R}^d} f d\lambda = \int_{\mathbb{R}^d} g d\lambda. \tag{14.785}$$

Nous pouvons aussi écrire

$$\int_{\mathbb{R}^d} f(x+a) d\lambda(x) = \int_{\mathbb{R}^d} f(x) d\lambda(x). \tag{14.786}$$

14.13.1 Quelques limites dans les bornes

Dans le cas de l'intégrale de Lebesgue définie par 14.167, si f est une fonction sur $\mathbb{R}$ et si λ est la mesure de Lebesgue, nous avons une définition directe de

$$\int_0^\infty f d\lambda. \tag{14.787}$$

Nous sommes cependant en droit de nous demander si nous n'aurions pas également ceci :

$$\lim_{x\to\infty} \int_0^x f\lambda = \int_0^\infty f d\lambda. \tag{14.788}$$

EQooDVKKooCiFzmA

Lorsque l'intégrale considérée est celle de Riemann, l'égalité (14.788) est une définition. Ici, ça va être une propriété, voir le lemme 14.261.

14.258.
Tant que nous sommes à parler de limites dans les bornes, nous aurions pu vouloir, pour les séries, suivre le chemin suivant :

— Définir l'intégrale de Lebesgue sur un espace mesuré.

— Prendre au passage le cas particulier $\sum_{k=0}^\infty a_k = \int_{\mathbb{N}} a$ où $a\colon \mathbb{N} \to \mathbb{R}$ est une fonction mesurable pour la mesure de comptage.

— Démontrer qu'avec ces définitions, $\sum_{k=0}^\infty a_k = \lim_{N\to\infty} \sum_{k=0}^N a_k$.

Or le dernier point est pris comme définition et son égalité avec l'intégrale pour la mesure de comptage est une propriété [88]. Pourquoi ? Parce que la définition 14.18 de mesure positive demande déjà d'avoir défini les sommes sur $\mathbb{N}$.

LEMooOPKLooGucWjb

Lemme 14.259.
Soit une partie mesurable $A \subset \mathbb{R}^+$ de mesure finie. Alors

$$\lim_{M\to\infty} \lambda\big(A \cap [M,\infty[\big) = 0. \tag{14.789}$$

88. Proposition 14.265.

Démonstration. La fonction

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto \lambda\big(A\cap [x,\infty[\big) \end{aligned} \tag{14.790}$$

est décroissante et bornée vers le bas par 0. Elle possède dont une limite $\ell\geqslant 0$ (corolaire 9.59). Nous allons prouver que $\ell=0$ en calculant la limite sur les entiers.

Nous posons $J_k=[k,k+1[$. Pour $n\in\mathbb{N}$ nous avons :

$$f(n)=\lambda\big(A\cap[n,\infty[\big) \tag{14.791a}$$

$$=\lambda\Big(\bigcup_{k=n}^{\infty}(A\cap J_k)\Big) \tag{14.791b}$$

$$=\sum_{k=n}^{\infty}\lambda(A\cap J_k). \tag{14.791c}$$

Mais nous savons par hypothèse sur la mesure de A que

$$\lambda(A)=\sum_{k=0}^{\infty}\lambda(A\cap J_k)<\infty. \tag{14.792}$$

Donc $f(n)$ est une queue de série convergente. Elle tend donc vers zéro par le lemme 10.101. C'est-à-dire que

$$\lim_{n\to\infty}f(n)=0. \tag{14.793}$$

Et comme $\lim_{x\to\infty}f(x)$ existe et vaut ℓ, la seule possibilité est $\ell=0$. □

LEMooMUHWooZPbMDb

Lemme 14.260.
Soit une fonction mesurable $f\colon\mathbb{R}\to\mathbb{R}^+$. Alors

$$\lim_{x\to\infty}\int_x^{\infty}f d\lambda=0. \tag{14.794}$$

Démonstration. Nous posons

$$F(x)=\int_x^{\infty}f d\lambda. \tag{14.795}$$

Nous commençons par prouver que c'est une fonction décroissante. En effet,

$$F(x)-F(x+a)=\int_x^{\infty}f-\int_{x+a}^{\infty}f=\int_x^{x+a}f+\int_{x+a}^{\infty}f-\int_{x+a}^{\infty}=\int_x^{x+a}f\geqslant 0. \tag{14.796}$$

Nous avons utilisé 14.194.

Vu que f prend ses valeurs dans $\mathbb{R}^+$, nous avons $F(x)\geqslant 0$ pour tout x. La fonction F est décroissante et bornée vers le bas. Donc elle a une limite :

$$\lim_{x\to\infty}F(x)=\ell\geqslant 0. \tag{14.797}$$

Supposons $\ell>0$ et posons $0<\epsilon<\ell$. Soit M tel que pour tout $x>M$ nous ayons

$$\int_x^{\infty}f>m. \tag{14.798}$$

EQooZVKQooQDzHXp

Soit $a\in\mathbb{R}^+$ tel que

$$\Big|\int_a^{\infty}f(t)-\ell\,dt\Big|<\epsilon. \tag{14.799}$$

En vertu de (14.798) nous considérons $a_0>a$ tel que

$$\int_{a_0}^{\infty}f=I_0>m. \tag{14.800}$$

Nous construisons la suite strictement croissante (a_k) de la façon suivante :

$$\int_{a_k}^{\infty} f = I_k > m \tag{14.801}$$

et

$$|\int_{a_k}^{a_{k+1}} f - I_k| < \epsilon. \tag{14.802}$$

Donc pour k nous avons

$$\int_{a_k}^{a_{k+1}} f \geqslant I_k - \epsilon \geqslant m - \epsilon. \tag{14.803}$$

Mais

$$\int_{a_0}^{\infty} = \sum_{k=0}^{\infty} \int_{a_k}^{a_{k+1}} f \geqslant \sum_k (m - \epsilon) = \infty. \tag{14.804}$$

Nous avons une contradiction. □

LEMooKGZDooWiKiHR

Lemme 14.261 ([1]).
Soit $f : \mathbb{R} \to \mathbb{R}$ intégrable sur $[a, \infty]$ pour la mesure de Lebesgue. Alors la limite

$$\lim_{b\to\infty} \int_{[a,b]} f \tag{14.805}$$

existe et vaut

$$\lim_{b\to\infty} \int_{[a,b]} f = \int_{[a,\infty]} f. \tag{14.806}$$

De même nous avons

$$\lim_{a\to-\infty} \int_{[a,b]} f = \int_{]-\infty,b]} f \tag{14.807}$$

EQooNRHDooVzAEFD

Démonstration. Vu qu'on a le droit de découper les domaines d'intégration en domaines disjoints, et que l'ajout d'un point ne change rien à la mesure de Lebesgue, nous avons pour tout $b \in [a, \infty[$ que

$$\int_{[a,\infty]} f = \int_{[a,b]} f + \int_{[b,\infty]} f. \tag{14.808}$$

Nous avons l'intention de prendre la limite $b \to \infty$. Le lemme 14.260 nous assure que la limite de la dernière intégrale existe et vaut zéro. Donc la limite de l'intégrale du milieu existe et vaut celle à gauche.

En ce qui concerne la formule (14.807), nous posons $g(x) = f(-x)$, et nous considérons $-a < b$. Et nous avons, en utilisant la première partie sur la fonction g :

$$\lim_{a\to\infty} \int_{[-a,b]} f = \lim_{a\to\infty} \int_{[-b,a]g} \tag{14.809a}$$

$$= \int_{[-b,\infty[} g \tag{14.809b}$$

$$= \int_{]-\infty,b]} f. \tag{14.809c}$$

□

PROPooSBKEooOlxqRk

Proposition 14.262 ([1]).
Soit une fonction $f : \mathbb{R} \to \mathbb{R}$ intégrable sur $\mathbb{R}$. Alors

$$\lim_{R\to\infty} \int_{-R}^{R} f = \int_{\mathbb{R}} f \tag{14.810}$$

au sens où la limite de gauche existe et vaut celle de droite.

14.13.2 Mesure de comptage et série

DEFooILJRooByDzhs

Définition 14.263 (mesure de comptage)**.**
*Soit $(S,\mathcal{F})$ un ensemble mesurable. La **mesure de comptage** sur $(S,\mathcal{F})$ est la mesure définie par*

$$m(A)=\begin{cases}\text{Card}(A) & \text{si } A \text{ est fini}\\ +\infty & \text{sinon.}\end{cases}\tag{14.811}$$

Cette mesure est utilisée pour voir des séries comme des intégrales sur $(\mathbb{N},\mathcal{P}(\mathbb{N}),m)$.

LEMooDTFHooLVsvAw

Lemme 14.264.
Si m est la mesure de comptage sur $\mathbb{N}$ et si $\mathcal{P}(\mathbb{N})$ est l'ensemble des parties de $\mathbb{N}$, alors le triple $\big(\mathbb{N},\mathcal{P}(\mathbb{N}),m\big)$ est un espace mesuré[89] *σ-fini.*

Démonstration. Bien entendu l'ensemble des parties de $\mathbb{N}$ est une tribu sur $\mathbb{N}$. Nous devons donc seulement vérifier les conditions de la définition 14.18.

Si A est une partie de $\mathbb{N}$, alors $m(A)\in[0,\infty]$ parce que c'est soit le cardinal de A soit ∞.

Le cardinal de l'ensemble vide est zéro (ça fait partie de la définition du cardinal 1.126).

Soient des parties deux à deux disjointes $\{A_i\}_{i\in I}$ de $\mathbb{N}$. Posons $A=\bigcup_{i\in I}A_i$. Si I est infini dénombrable, alors l'union est infinie : prendre par exemple la partie $S=\{\min(A_i)\}_{i\in I}$ qui est en bijection avec I. Nous avons alors d'une part

$$m\big(\bigcup_i A_i\big)\geqslant m(S)=\infty,\tag{14.812}$$

et d'autre part

$$\sum_{i\in I}m(A_i)\geqslant\sum_{i\in I}1=\infty.\tag{14.813}$$

Si par contre I est fini, alors le lemme 1.128(5) dit que

$$m\big(\bigcup_{i\in I}A_i\big)=\text{Card}\big(\bigcup_{i\in I}A_i\big)=\sum_{i\in I}m(A_i).\tag{14.814}$$

Le fait que notre espace soi σ-fini se voit par exemple en posant $E_n=\{0,\dots,n\}$. Chaque E_n est de mesure finie, et leur union est $\bigcup_{n\in\mathbb{N}}E_n=\mathbb{N}$. □

PROPooPNQAooDRLcCm

Proposition 14.265 ([1])**.**
Soit l'espace mesuré $\big(\mathbb{N},\mathcal{P}(\mathbb{N}),m\big)$[90]*. Nous considérons une application $a\colon\mathbb{N}\to\mathbb{R}^+$ (c'est à dire une suite de nombres réels positifs).*

(1) L'intégrale $\int_{\mathbb{N}}a\,dm<\infty$ si et seulement si la série aussi : $\sum_{n=0}^{\infty}a_n<\infty$.
(2) Si $\int_{\mathbb{N}}a\,dm$ existe, alors

$$\int_{\mathbb{N}}a\,dm=\sum_{n=0}^{\infty}a_n.\tag{14.815}$$

EQooFIRMooKQEsAn

Démonstration. Nous allons travailler avec la définition 14.167 de l'intégrale et 14.112 pour les fonctions étagées.

(1) **Si l'intégrale est $<\infty$** Nous supposons que $\int_{\mathbb{N}}adm<\infty$. La proposition 14.210 nous permet de découper $\mathbb{N}$ en parties disjointes. Nous choisissons de voir $\mathbb{N}=\bigcup_{i\in\mathbb{N}}\{i\}$ et donc de considérer l'égalité

$$\int_{\mathbb{N}}adm=\sum_{i=0}^{\infty}\int_{\{i\}}adm.\tag{14.816}$$

Vu que $m\big(\{i\}\big)=1$ et que a est constante sur $\{i\}$, nous avons $\int_{\{i\}}adm=a_i$, et donc

$$\int_{\mathbb{N}}adm=\sum_{i=0}^{\infty}a_i.\tag{14.817}$$

89. Définition 14.18.
90. La mesure de comptage sur $\mathbb{N}$ est donnée en la définition 14.263.

(2) **Si la somme est $< \infty$** Le terme général a_n tend vers zéro[91]. En particulier, pour chaque $r \in \mathbb{R}$, il n'existe qu'un nombre fini d'éléments de la suite a plus grands que r.
Soit une fonction étagée $\psi\colon \mathbb{N} \to \mathbb{R}^+$ minorant a. Vu qu'elle est étagée, elle ne prend qu'un nombre fini de valeurs différentes que nous nommons $\{\alpha_i\}$, et les parties $A_i = \psi^{-1}(\alpha_i)$ sont finies parce que ψ minore a.
Avant de faire un petit calcul, posons quelque notations. D'abord $Z = \psi^{-1}(0)$. Nous avons $\mathbb{N} = \bigcup_i A_i \cup Z$. Nous posons aussi $A = \bigcup_i A_i$; c'est le support de ψ.
Nous pouvons maintenant faire un calcul pour l'intégrale de ψ sur $\mathbb{N}$: SUBEQSooPMCHooIdtHqG

$$\begin{aligned}
\int_{\mathbb{N}} \psi dm &= \sum_i \alpha_i \operatorname{Card}(A_i) && \text{def. (14.467)} && (14.818a)\\
&= \sum_i \Big(\sum_{k\in A_i} \psi(k) \Big) && && (14.818b)\\
&= \sum_{k\in\bigcup_i A_i} \psi(k) && && (14.818c)\\
&= \sum_{k\in A} \psi(k) + \sum_{k\in Z} \psi(k) && && (14.818d)\\
&= \sum_{k\in\mathbb{N}} \psi(k). && && (14.818e)
\end{aligned}$$

SUBEQooIJMMooPKLHVk SUBEQooGOBRooYhAufP

Justifications.

— Pour (14.818b). Pour tout k dans A_i nous avons $\psi(k) = \alpha_i$.

— Pour (14.818c). Les A_i sont disjoints et chacun est fini. La proposition 10.120 fait le boulot.

Étant donné que ψ minore a, nous avons donc montré que

$$\int_{\mathbb{N}} \psi dm = \sum_{k\in\mathbb{N}} \psi(k) \leqslant \sum_{k\in\mathbb{N}} a_n. \tag{14.819}$$

Donc l'intégrale de toutes les fonctions étagées minorant a est majorée par la série de a. Donc l'intégrale de a existe et est majorée par cette somme.

(3) **Valeur de l'intégrale** Nous supposons encore que la série est finie. Nous avons déjà prouvé que

$$\int_{\mathbb{N}} a dm \leqslant \sum_{n\in\mathbb{N}} a_n. \tag{14.820}$$ EQooNCMCooYMaASZ

Nous allons prouver l'inégalité inverse en trouvant une bonne suite de fonctions étagées. Soit, pour chaque $N \in \mathbb{N}$,

$$\begin{aligned}
\psi_N\colon \mathbb{N} &\to \mathbb{R}^+\\
k &\mapsto \begin{cases} a_k & \text{si } k \leqslant N\\ 0 & \text{sinon.} \end{cases}
\end{aligned} \tag{14.821}$$

Cela est une fonction étagée qui minore a. Nous pouvons donc reprendre le calcul (14.818) :

$$\int_{\mathbb{N}} \psi dm = \sum_{k\in\mathbb{N}} \psi_N(k) = \sum_{k=0}^{N} a_k. \tag{14.822}$$

Par définition de la série, nous avons

$$\lim_{N\to\infty} \int_{\mathbb{N}} \psi_N dm = \sum_{k=0}^{\infty} a_k. \tag{14.823}$$

Donc le supremum des intégrales de fonctions étagées est au moins égal à $\sum_{k=0}^{\infty} a_k$. Nous avons prouvé l'inégalité inverse de (14.820), et donc l'égalité (14.815).

91. Proposition 10.100.

□

EXooDMAUooWWVZbq

Exemple 14.266.
L'intervalle $I = [0,1]$ muni de la tribu de toutes ses parties et de la mesure de comptage est un espace mesuré non σ-fini. △

14.13.3 Théorème de la moyenne

LEMooOXNPooXpTxtx

Lemme 14.267 ([1]).
Soit A connexe par arcs[92]*, soit $f\colon A \to \mathbb{R}$ continue ainsi que $\alpha, \beta \in f(A)$ avec $\alpha < \beta$. Pour tout $y \in [\alpha, \beta]$, il existe $x \in A$ tel que $f(x) = y$.*

Démonstration. Soient $a, b \in A$ avec $f(a) = \alpha$ et $f(b) = \beta$. Vu que A est connexe par arcs, il existe une application continue $\gamma\colon [0,1] \to A$ telle que $\gamma(0) = a$ et $\gamma(1) = b$. Le théorème des valeurs intermédiaires appliqué à la composée $f \circ \gamma$ dit qu'il existe $t_0 \in [0,1]$ tel que $(f \circ \gamma)(t_0) = y$. Nous avons donc le résultat avec $x = \gamma(t_0)$. □

ThoooEZLGooMChwLT

Théorème 14.268 ([1, 162]).
Soit Q un compact connexe par arcs et une fonction continue $f\colon Q \to \mathbb{R}$. Si λ est la mesure de Lebesgue, alors il existe $a \in Q$ tel que

$$f(a) = \frac{1}{\lambda(Q)} \int_Q f d\lambda \tag{14.824}$$

Démonstration. En posant $I = \int_Q f d\lambda$ nous avons immédiatement

$$\min(f)\lambda(Q) \leqslant I \leqslant \max(f)\lambda(Q) \tag{14.825}$$

EqooTYQCooVxdazW

où le minimum et le maximum existent parce que f est continue sur un compact. Si une des deux inégalités est une égalité alors la fonction est constante. En effet supposons que la première inégalité soit une égalité ; si la fonction n'était pas constante, il existerait une boule sur laquelle f serait strictement supérieure à $\min(f)$. En intégrant d'abord sur cette boule et ensuite sur le complémentaire nous obtenons une intégrale plus grande que $\min(f)\lambda(Q)$.

Vu que f est continue sur le compact Q, il existe $\alpha, \beta \in Q$ tels que $f(\alpha) = \min(f)$ et $f(\beta) = \max(f)$. Vu que $\min(f) \leqslant I/\lambda(Q) \leqslant \max(f)$, le lemme 14.267 montre qu'il existe $a \in Q$ tel que $f(a) = I/\lambda(Q)$. Autrement dit,

$$f(a) = \frac{1}{\lambda(Q)} \int_Q f d\lambda. \tag{14.826}$$

□

14.14 Changement de variables dans une intégrale multiple

Dans ce qui suit, U et V sont des ouverts de $\mathbb{R}^N$ et $\phi\colon U \to V$ est un C^1-difféomorphisme. Nous notons $\mathcal{Q}$ l'ensemble des cubes fermés dans U dont les côtés sont parallèles aux axes.

14.14.1 Des lemmes

LemooJYCGooIkkDVn

Lemme 14.269 ([469]).
Soient μ et ν deux mesures de Borel sur l'ouvert U de $\mathbb{R}^N$. Si $\mu(Q) \leqslant \nu(Q)$ pour tout $Q \in \mathcal{Q}$ alors $\mu(B) \leqslant \nu(B)$ pour tout borélien B.

Démonstration. Si Q est un cube semi-ouvert, c'est-à-dire de la forme

$$Q = \prod_{n=1}^{N} [a_n, a_n + h[\subset U \tag{14.827}$$

92. Définition 9.64.

alors Q est une réunion croissante de cubes fermés du type $[a_n+\epsilon, a_n+h-\epsilon]$, et donc $\mu(Q) \leqslant \nu(Q)$ par le lemme 14.24(1). La propriété est donc vraie pour les cubes semi-ouverts.

Si Ω est un ouvert, alors il est réunion disjointe dénombrable de cubes semi-ouverts par la proposition 14.250. Donc pour tout ouvert $\Omega \subset U$ nous avons $\mu(\Omega) \leqslant \nu(\Omega)$. En vertu de la proposition 14.90 et de la remarque 14.91, les mesures μ et ν sont régulières, et l'inégalité au niveau des ouverts se répercute en inégalité pour tout boréliens de U :

$$\mu(B) \leqslant \nu(B) \tag{14.828}$$

pour tout $B \in \mathcal{B}\text{or}(U)$. Notons que U étant ouvert dans $\mathbb{R}^N$, les boréliens de U sont exactement les boréliens de $\mathbb{R}^N$ inclus dans U par le corolaire 14.58. □

LemooJCEDooBRyjRg

Lemme 14.270 ([469]).
Soit une application $\theta\colon U \to \mathbb{R}^N$ de classe C^1 où U est ouvert dans $\mathbb{R}^N$. Pour tout $Q \in \mathcal{Q}$ nous avons

$$\lambda_N\big(\theta(Q)\big) \leqslant \sup_{s\in Q} \|d\theta_s\|^N \lambda_N(Q). \tag{14.829}$$

Démonstration. Nous notons h la longueur du côté du cube. Le théorème des accroissements finis 12.314, pour la composante θ_i donne, pour $u, v \in Q$:

$$\big|\theta_i(u) - \theta_i(v)\big| \leqslant \sup_{s\in Q} \|(d\theta_i)_s\| \|u-v\| \leqslant \sup_{s\in Q} \|(d\theta_i)_s\| h. \tag{14.830}$$

EqooFZMAooKWdxxJ

D'autre part nous avons (nous écrivons pour $N=2$ pour être plus court) :

$$d\theta_s(u) = \frac{d}{dt}\Big[\theta_1(s+tu)e_1 + \theta_2(s+tu)e_2\Big]_{t=0} = (d\theta_1)_s(u)e_1 + (d\theta_2)_s(u)e_2. \tag{14.831}$$

Donc pour chaque i : $\|d\theta_s\| \geqslant \|(d\theta_i)_s\|$, et nous continuons la majoration (14.830) :

$$\big|\theta_i(u) - \theta_i(v)\big| \leqslant \sup_{s\in Q} \|(d\theta_i)_s\| h \leqslant \sup_{s\in Q} \|d\theta_s\| h. \tag{14.832}$$

Les points $\theta(u)$ et $\theta(v)$ sont donc dans un cube de côté $\sup_{s\in Q} \|d\theta_s\| h$, ce qui permet de majorer $\lambda_N\big(\theta(Q)\big)$ par

$$\lambda_N\big(\theta(Q)\big) \leqslant \left(\sup_{s\in Q} \|d\theta_s\| h\right)^N = \left(\sup_{s\in Q} \|d\theta_s\|\right)^N \lambda_N(Q) \tag{14.833}$$

où le dernier facteur provient de l'égalité $h^N = \lambda_N(Q)$. □

14.14.2 Déterminant et mesure de Lebesgue

Dans la suite, Q_0 désigne le cube unité : $Q_0 = \big([0,1[\big)^N$.

ThoBVIJooMkifod

Théorème 14.271 (Interprétation géométrique du déterminant[469]).
Soit une application linéaire $T\colon \mathbb{R}^N \to \mathbb{R}^N$. Alors pour tout borélien B de $\mathbb{R}^N$,

$$\lambda_N\big(T(B)\big) = |\det(T)| \lambda_N(B). \tag{14.834}$$

Démonstration. Nous considérons la mesure positive μ donnée par $\mu(B) = \lambda_N\big(T(B)\big)$, qui est bien une mesure par la proposition 14.84. Cette mesure est invariante par translation parce que λ_N l'est :

$$\mu(B+a) = \lambda_N\big(T(B)+a\big) = \lambda_N\big(T(B)\big) = \mu(B). \tag{14.835}$$

De plus, $T(Q_0)$ est borné et nous notons $\mu(Q_0) = C$. Nous avons $\mu = C\lambda_N$ par le corolaire 14.254.

(1) $C(T_1T_2) = C(T_1)C(T_2)$ Par définition,

$$C(T_1T_2)\lambda_N(B) = \lambda_N\big((T_1T_2)(B)\big) \tag{14.836a}$$
$$= \lambda_N\big(T_1(T_2B)\big) = C(T_1)\lambda_N\big(T_2(B)\big) = C(T_1)C(T_2)\lambda_N(B). \tag{14.836b}$$

Par conséquent la fonction C est multiplicative :

$$C(T_1T_2) = C(T_1)C(T_2). \tag{14.837}$$

Et en plus, $C(\mathrm{Id}) = 1$.

(2) **Matrice diagonale** En guise de T, nous considérons l'application linéaire diagonale donnée par $De_i = d_ie_i$, ou, sous forme matricielle, $D = \mathrm{diag}(d_1, \dots, d_N)$ qui fait

$$T(Q_0) = [0, d_1[\times \dots \times [0, d_N[\tag{14.838}$$

La mesure de cela est $|d_1 \cdots d_N|$, ce qui nous donne

$$C(D) = |d_1 \dots d_N| = |\det(D)|. \tag{14.839}$$

(3) **Matrice orthogonale** Nous considérons maintenant $T = U$ où U est une matrice orthogonale ($UU^t = 1$). Une matrice orthogonale est une isométrie [93] qui conserve donc la boule unité : $UB(0,1) = B(0,1)$. Nous avons

$$\lambda_N\big(B(0,1)\big) = \lambda_N\big(UB(0,1)\big) = C(U)\lambda_N\big(B(0,1)\big) \tag{14.840}$$

par conséquent $C(U) = 1$, et 1 est justement le déterminant de U.

(4) **Matrice quelconque** Nous savons par le corolaire 13.35 de la décomposition polaire que toute matrice peut être écrite sous la forme $T = U_1DU_2$ où U_i sont orthogonales et D est diagonale. Donc $C(T) = C(U_1)C(D)C(U_2) = \det(U_1)\det(D)\det(U_2) = \det(U_2DU_2) = \det(T)$ parce que le déterminant est multiplicatif (proposition 8.16(1)).

□

Ce théorème donne une interprétation géométrique du déterminant en tant que facteur de dilatation des volumes lors de l'utilisation d'une application linéaire. Si T est une application linéaire quelconque,

$$\lambda_N\big(T(Q_0)\big) = |\det(T)|\lambda_N(Q_0) = |\det(T)|. \tag{14.841}$$

Le déterminant de T est le volume de l'image du cube unité par l'application T.

De la même façon, en utilisant l'application linéaire $T(x) = ax$ nous avons pour tout borélien B :

$$\lambda_N(aB) = a^N\lambda_N(B). \tag{14.842}$$

Une dilatation d'un facteur a des longueurs provoque une multiplication par a^N des volumes.

LEMooOLSMooCimcIT

Lemme 14.272.
Tout hyperplan [94] *de $\mathbb{R}^n$ est de mesure de Lebesgue nulle.*

Démonstration. Dans le cas d'un hyperplan vectoriel, il s'agit de prendre le théorème 14.271. En posant $B = \mathbb{R}^n$, et en considérant l'application linéaire de projection de $\mathbb{R}^n$ vers l'hyperplan, on a le résultat parce que le déterminant d'une telle application est nul.

Si l'hyperplan n'est pas vectoriel, alors nous utilisons l'invariance par translation [95] et nous nous ramenons au cas vectoriel. □

93. Proposition 8.46.
94. Définition 8.307.
95. Théorème 14.252.

CorTHDQooWMSbJe

Corolaire 14.273 ([451]).
Tout ouvert de $\mathbb{R}^n$ est une union dénombrable de cubes presque disjoints[96].

Démonstration. Il suffit de prendre les cubes de la proposition 14.250 et de les fermer. Ce que l'on ajoute est de mesure nulle[97]. □

Remarque 14.274.
La proposition 14.250 est une propriété seulement de la topologie de $\mathbb{R}^n$ alors que le corolaire 14.273 fait intervenir la mesure de Lebesgue parce qu'il faut bien dire que les intersections sont de mesure (de Lebesgue) nulle.

14.14.3 Ensembles négligeables

LemWHKJooGPuxEN

Lemme 14.275 ([470]).
L'image d'une partie négligeable de $\mathbb{R}^d$ par une application Lipschitz est négligeable.

Démonstration. Soit N une partie négligeable de $\mathbb{R}^d$ et une application Lipschitz $f\colon N \to \mathbb{R}^d$. Soit $Q \subset \mathbb{R}^d$ un cube borné de côté r. Pour tout $x, x' \in N \cap Q$ nous avons

$$\|f(x) - f(x')\| \leqslant C\|x - x'\| \leqslant Cr. \tag{14.843}$$

Donc $f(N \cap Q)$ est dans une boule de rayon Cr. Mais comme toutes les normes sont équivalentes[98] sur $\mathbb{R}^d$ nous pouvons tout aussi bien prendre la norme $\|.\|_1$ au lieu de la norme $\|.\|_2$ (qui est toujours la norme prise implicitement lorsqu'on parle de $\mathbb{R}^n$), de telle sorte que les boules soient des cubes. Quoi qu'il en soit, $f(N \cap Q)$ est contenu dans un cube de côté $2Cr$ et au niveau de la mesure extérieure,

$$m^*\big(f(N \cap Q)\big) \leqslant (2Cr)^d = (2C)^d r^d, \tag{14.844}$$

ou encore

$$m\big(f(N \cap Q)\big) \leqslant (2C)^d m(Q) \tag{14.845}$$

parce que r^d est la mesure du cube Q.

Soit maintenant $\epsilon > 0$; vu que N est négligeable, il existe un ouvert U contenant N et tel que $m(U) < \epsilon$. Ce U est une union presque disjointe de cubes dyadiques (Q_n) par le corolaire 14.273. Nous avons alors

$$\begin{aligned}
m^*\big(f(N)\big) &= m^*\big(f(\bigcup_n N \cap Q_n)\big) && \text{(14.846a)}\\
&= m^*\big(\bigcup_n f(N \cap Q_n)\big) && \text{(14.846b)}\\
&\leqslant \sum_n m^*(f(N \cap Q_n)) && \text{(14.846c)}\\
&\leqslant \sum_n (2C)^d m(Q_n) && \text{(14.846d)}\\
&= (2C)^d m(U) && \text{(14.846e)}\\
&< (2C)^d \epsilon. && \text{(14.846f)}
\end{aligned}$$

Au final, $m^*\big(f(N)\big) \leqslant (2C)^d\epsilon$. L'ensemble N est donc négligeable parce que le lemme 14.75 le dit : $m^*(N) = 0$. □

Corolaire 14.276.
Un sous-espace vectoriel strict de $\mathbb{R}^N$ est négligeable.

96. « presque » au sens où les intersections éventuelles sont de mesure de Lebesgue nulle.
97. Voir le lemme 14.272.
98. Proposition 10.47

Démonstration. Un sous-espace vectoriel strict de $\mathbb{R}^N$ de dimension $k < N$ est l'image de

$$A = \{t_1 e_1 + \cdots + t_k e_k \text{ tel que } t_i \in \mathbb{R}\} \tag{14.847}$$

par une application linéaire. Ce A est un pavé de mesure de Lebesgue nulle. Donc l'image est négligeable par le lemme 14.275. □

14.14.4 Parties et fonctions mesurables

Pour rappel, la notion d'application de classe C^1 est donnée par la définition 10.261.

PropRDRNooFnZSKt

Proposition 14.277.
Soient U et V des ouverts de $\mathbb{R}^d$ et $\phi: U \to V$ un C^1-difféomorphisme. Si $E \subset U$ est mesurable, alors $\phi(E)$ est mesurable[99].

Démonstration. Si E est mesurable, il existe un borélien B et un ensemble négligeable N tels que $E = B \cup N$. Vu que ϕ est un homéomorphisme, l'application ϕ^{-1} est borélienne parce que continue (théorème 14.56). Nous avons

$$\phi(B) = (\phi^{-1})^{-1}(B), \tag{14.848}$$

c'est-à-dire que $\phi(B)$ est l'image inverse de B par ϕ^{-1}. L'ensemble $\phi(B)$ est donc borélien.

Il reste à voir que $\phi(N)$ est négligeable. Soit $Q \subset U$ un cube compact. L'application $d\phi: Q \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ est continue et donc bornée (par la remarque 10.268) sur le compact Q. Par les accroissements finis (théorème 10.292), l'application ϕ est donc Lipschitz sur Q. La partie $\phi(N \cap Q)$ est alors négligeable par le lemme 14.275. Pour conclure,

$$\phi(N) = \bigcup_i \phi(N \cap Q_i) \tag{14.849}$$

où les Q_i sont tous des cubes compacts. Donc $\phi(N)$ est une union dénombrable d'ensembles négligeables ; ergo négligeable lui-même par le lemme 14.67. □

Proposition 14.278.
Soient U et V des ouverts de $\mathbb{R}^N$ et $\phi: U \to V$ un C^1-difféomorphisme. Si $f: V \to \mathbb{C}$ est mesurable, alors $f \circ \phi: U \to \mathbb{C}$ l'est.

Démonstration. Soit A une partie mesurable de $\mathbb{C}$. Il nous faut prouver que

$$(f \circ \phi)^{-1}(A) = \phi^{-1}\big(f^{-1}(A)\big) \tag{14.850}$$

soit mesurable. Par hypothèse , $f^{-1}(A)$ est mesurable. Vu que ϕ est un C^1-difféomorphisme, elle et son inverse sont mesurables par la proposition 14.277. Donc l'image du mesurable $f^{-1}(A)$ par ϕ^{-1} est encore mesurable. □

PROPooSEJJooIfGUkW

Proposition 14.279.
Si $f: \mathbb{R}^n \to \mathbb{R}^n$ est de classe C^∞ et si N est une partie de mesure (de Lebesgue) nulle dans $\mathbb{R}^n$, alors $f(N)$ est également de mesure nulle.

PROPooIESAooBsThho

Proposition 14.280.
Soit une application $f: \mathbb{R}^n \to \mathbb{R}$ qui s'annule une quantité dénombrable de fois. Soit N de mesure nulle dans $\mathbb{R}^n$. Pour tout $\epsilon > 0$, il existe $\delta \in \mathbb{R}^n$ tel que $\|\delta\| < \epsilon$ et tel que la fonction $x \mapsto f(x + \delta)$ ne s'annule pas sur N.

Démonstration. Soient $\{x_i\}_{i \in \mathbb{N}}$ les racines de f. Pour chaque i, nous notons

$$F_i = \{\delta \in \mathbb{R}^n \text{ tel que } x_i + \delta \in N\}. \tag{14.851}$$

Chacun de F_i est de mesure nulle, et donc $\bigcup_{i=1}^{\infty} F_i$ est de mesure nulle. Il suffit de prendre $\delta \in B(0, \epsilon) \backslash \bigcup_i F_i$, et nous avons $x_i + \delta$ hors de N pour tout i. □

99. Ici « mesurable » parle de mesurabilité au sens de la tribu de Lebesgue, c'est-à-dire pas seulement les boréliens.

14.14.5 Le théorème et sa démonstration

THOooUMIWooZUtUSg

Théorème 14.281 (Changement de variable[470, 469]).
Soient U et V des ouverts de $\mathbb{R}^N$ ainsi qu'un C^1-difféomorphisme $\phi\colon U \to V$. Nous notons J_ϕ la fonction

$$\begin{aligned} J_\phi\colon \mathbb{R}^N &\to \mathbb{R} \\ a &\mapsto \det(d\phi_a). \end{aligned} \tag{14.852}$$

Alors :

ItemVWYDooOzwnyfi

(1) Si $E \subset U$ est borélien, alors $\phi(E)$ est borélien et

$$\lambda_N\big(\phi(E)\big) = \int_E |J_\phi| d\lambda_N, \tag{14.853}$$

c'est-à-dire $\phi^{-1}(\lambda_N) = |J_\phi| \cdot \lambda_N$.

ITEMooEZUBooGBuDOS

(2) Si $f\colon V \to [0, +\infty]$ est mesurable alors la fonction

$$\begin{aligned} (f \circ \phi) \times |J_\phi|\colon U &\to [0, \infty] \\ x &\mapsto (f \circ \phi)(x)|J_\phi(x)|d\lambda_N(x) \end{aligned} \tag{14.854}$$

l'est également et[100]

$$\int_{\phi(U)} f d\lambda_N = \int_U (f \circ \phi)(x)|J_\phi(x)|d\lambda_N(x). \tag{14.855}$$

EqRANEooQsFbbC

ITEMooAJGDooGHKnvj

(3) Si $f\colon V \to \mathbb{C}$ est mesurable alors elle est intégrable si et seulement si $(f \circ \phi) \times |J_\phi|\colon U \to \mathbb{C}$ est intégrable. Si c'est le cas, alors nous avons encore la formule de changement de variables :

$$\int_V f d\lambda_N = \int_{\phi^{-1}(V)} (f \circ \phi)|J_\phi| d\lambda_N. \tag{14.856}$$

EQooLYAWooTArAZR

Démonstration. Attention : la preuve va être longue.

(1) Le fait que $\phi(E)$ soit borélien lorsque E l'est est la proposition 14.277. En ce qui concerne la formule annoncée, il faut travailler.

(1a) **Inégalité dans un sens (cubes)** Nous commençons par prouver l'inégalité

$$\lambda_N\big(\phi(Q)\big) \leqslant \int_Q |J_\phi(x)|dx \tag{14.857}$$

EqooQCXXooSjGzks

pour tout $Q \in \mathcal{Q}$. On peut diviser le côté du cube Q en k éléments de longueurs égales. Le cube est alors divisé en k^N petits cubes d'intérieurs disjoints. Nous les nommons Q_i ($i = 1, \dots, k^N$) Nous avons alors

$$\sum_i \lambda_N(Q_i) = \sum_i \lambda_N\big(\operatorname{Int}(Q_i)\big) = \lambda_N\big(\bigcup_i \operatorname{Int}(Q_i)\big) \leqslant \lambda_N(Q) \leqslant \sum_i \lambda_N(Q_i). \tag{14.858}$$

La dernière inégalité est le fait que les intersections ne sont pas disjointes. Toutes ces inégalités sont en réalité des égalités et en particulier : $\lambda_N(Q) = \sum_i \lambda_N(Q_i)$.
Soit $a \in Q_i$. Posons

$$\begin{aligned} &\theta\colon U \to U \\ &\theta = (d\phi_a)^{-1} \circ \phi \end{aligned} \tag{14.859}$$

Cela appelle deux commentaires. D'abord l'application $d\phi_a\colon U \to V$ est inversible parce que ϕ est un difféomorphisme (lemme 10.284). Ensuite, l'application θ est la composée

100. L'intégrabilité d'une fonction est la définition 14.186 qui stipule que l'intégrale de $|f(x)|$ est finie. L'égalité proposée a un sens si les deux membres sont infinis. Il n'y a donc pas d'hypothèses d'intégrabilité obligatoire pour écrire une intégrale lorsque la fonction a des valeurs positives.

de $(d\phi_a)$ (qui est linéaire) et de ϕ qui est de classe C^1 ; donc θ est de classe C^1. Donc le lemme 14.270 s'applique. La différentielle de θ n'est pas trop compliquée à écrire parce que nous avons la formule de différentielle d'une composée (théorème 10.277) et le fait que $(d\phi_a)^{-1}$ qui est linéaire et donc sa propre différentielle (lemme 10.273). Nous avons donc $d\theta = (d\phi_a)^{-1} \circ d\phi$, et le lemme donne

$$\lambda_N\left((d\phi_a)^{-1}\phi(a)\right) \leqslant \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_i) \tag{14.860}$$

Étant donné que $(d\phi_a)^{-1}$ est une application linéaire, la proposition 14.271 s'applique, et donc

$$\lambda_N\left((d\phi_a)^{-1}\phi(a)\right) = |\det(d\phi_a)^{-1}|\lambda_N\big(\phi(a)\big). \tag{14.861}$$

Le déterminant d'une application réciproque est donné par la proposition 8.16(4) :

$$\det\left((d\phi_a)^{-1}\right) = \frac{1}{\det\big(d\phi_a\big)} = \frac{1}{J_\phi(a)}. \tag{14.862}$$

Recollant les morceaux,

$$\lambda_N\big(\phi(Q_i)\big)\frac{1}{J_\phi(a)} \leqslant \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_i), \tag{14.863}$$

ou encore :

$$\lambda_N\big(\phi(Q_i)\big) \leqslant |J_\phi(a)| \sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\|^N \lambda_N(Q_i). \tag{14.864}$$

Vu que a et s sont proches l'un de l'autre (on peut choisir encore la taille du cube), nous pouvons espérer que $(d\phi_a)^{-1}$ ne soit pas loin d'être l'inverse de $d\phi_s$. Et c'est en effet le cas. Pour s'en assurer, remarquons que l'application

$$d\phi\colon Q_i \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N) \tag{14.865}$$

est continue et même uniformément continue parce que Q_i est compact. De plus la composition de différentielles étant un produit de matrices nous pouvons permuter la limite dans le calcul suivant :

$$\lim_{s\to a}(d\phi_a)^{-1} \circ d\phi_s = (d\phi_a)^{-1} \circ \lim_{s\to a} d\phi_s = \mathbb{1}. \tag{14.866}$$

Donc si $\epsilon > 0$ est donné, il existe δ tel que pour tout $s \in B(a,\delta)$, $\|(d\phi_a)^{-1} \circ d\phi_s - \mathbb{1}\| \leqslant \epsilon$. En ce qui concerne les normes, si $\|A - \mathbb{1}\| \leqslant \epsilon$ alors $\|A\| \leqslant \|A - \mathbb{1}\| + \|\mathbb{1}\| \leqslant \epsilon + 1$.

Cela étant dit, nous nous souvenons que nous avions découpé U en un nombre fini de cubes Q_i d'égales dimensions ; il suffit de prendre k suffisamment grand pour que la diagonale des cubes sot plus petite que le minimum des δ_i. Avec un tel découpage,

$$\sup_{s\in Q_i} \|(d\phi_a)^{-1} \circ d\phi_s\| \leqslant 1 + \epsilon \tag{14.867}$$

et par conséquent

$$\lambda_N\big(\phi(Q_i)\big) \leqslant (1+\epsilon)^N |J_\phi(a_i)|\lambda_N(Q_i) \tag{14.868}$$

EqooQRMNooZduAkX

où nous avons ajouté un indice i au point a pour nous rappeler que nous avons choisi $a \in Q_i$.

Le théorème de la moyenne 14.268 appliqué à l'intégrale $\int_{Q_i} |J_\phi(t)|d\lambda_N(t)$ donne l'existence d'un $a_i \in Q_i$ tel que

$$|J_\phi(a_i)| = \frac{1}{\lambda_N(Q_i)} \int_{Q_i} |J_\phi| d\lambda_N. \tag{14.869}$$

Ce point a_i vérifie l'inégalité (14.868) comme tout point de Q_i. Nous sommons ces inégalités sur tous les i :

$$\lambda_N\big(\phi(Q)\big) \leqslant \sum_i \lambda_N\big(\phi(Q_i)\big) \tag{14.870a}$$

$$\leqslant (1+\epsilon)^N \sum_i \left(\frac{1}{\lambda_N(Q_i) \int_{Q_i} |J_\phi| d\lambda_N} \right) \lambda_N(Q_i) \tag{14.870b}$$

$$= (1+\epsilon)^N \sum_i \int_{Q_i} |J_\phi| d\lambda_N \tag{14.870c}$$

$$= (1+\epsilon)^N \int_Q |J_\phi| d\lambda_N \tag{14.870d}$$

où nous avons utilisé le fait que $\mathbb{1}_Q = \sum_i \mathbb{1}_{Q_i}$ presque partout. En prenant le limite $\epsilon \to 0$ nous trouvons

$$\lambda_N\big(\phi(Q)\big) \leqslant \int_Q |J_\phi| d\lambda_N. \tag{14.871}$$

L'inégalité (14.857) est prouvée.

(1b) **Inégalité pour les boréliens**

Soit B un borélien de U. Vu que U et V sont des ouverts de $\mathbb{R}^N$, les mesures de Lebesgue sur U et sur V sont les mêmes que celles sur $\mathbb{R}^N$ par le corolaire 14.58.

Par les définitions 14.212 et 14.84, les applications μ et ν définies par $\mu = \phi^{-1}(\lambda_N)$ et $\nu = |J_\phi|\lambda_N$ sont des mesures positives sur U (de Borel, qui plus est). L'inégalité (14.857) à peine prouvée s'écrit $\mu(Q) \leqslant \nu(Q)$ pour tout cube Q. Le lemme 14.269 nous dit alors que l'inégalité tient pour tout borélien.

(1c) **Inégalité dans l'autre sens**

En utilisant la notation de la mesure image et du produit d'une mesure par une fonction [101], nous pouvons écrire l'inégalité prouvée sous la forme $\phi^{-1}(\lambda_N) \leqslant |J_\phi|\lambda_N$. En inversant les rôles de U et V (et donc de ϕ et ϕ^{-1}) nous avons aussi

$$\phi(\lambda_N) \leqslant |J_{\phi^{-1}}|\lambda_N. \tag{14.872}$$

En y appliquant ϕ^{-1} et le lemme 14.84,

$$\lambda_N \leqslant \phi^{-1}\big(|J_{\phi^{-1}}|\lambda_N\big). \tag{14.873}$$

EqooHJCHooVIaheI

Nous prouvons à présent que $\phi^{-1}\big(|J_{\phi^{-1}}| \cdot \lambda_N\big) = \Big(|J_{\phi^{-1}}| \circ \phi\Big) \cdot \phi^{-1}(\lambda_N)$ en appliquant à un borélien B de U. D'une part

$$\phi^{-1}\big(|J_{\phi^{-1}}| \cdot \lambda_N\big)(B) = \big(|J_{\phi^{-1}}| \cdot \lambda_N\big)\phi(B) \tag{14.874a}$$

$$= \int_{\phi(B)} |J_{\phi^{-1}}| d\lambda_N, \tag{14.874b}$$

et d'autre part,

$$\Big(|J_{\phi^{-1}}| \circ \phi\Big) \cdot \phi^{-1}(\lambda_N)B = \int_{\mathbb{R}^N} \mathbb{1}_B(x)\big(|J_{\phi^{-1}}| \circ \phi\big)(x) d\big(\phi^{-1}(\lambda_N)\big)(x) \tag{14.875a}$$

$$= \int_{\mathbb{R}^N} \mathbb{1}_B\big(\phi^{-1}(x)\big)\big(|J_{\phi^{-1}}| \circ \phi\big)\big(\phi^{-1}(x)\big) d\lambda_N(x) \tag{14.875b}$$

$$= \int_{\mathbb{R}^N} \mathbb{1}_{\phi(B)} |J_{\phi^{-1}}| \tag{14.875c}$$

$$= \int_B |J_{\phi^{-1}}| d\lambda_N. \tag{14.875d}$$

ooDKSWooXwQwgQ

Justification :

101. Définition 14.84 et 14.212

— Pour (14.875b), le théorème 14.219(2).

L'équation (14.873) devient alors

$$\lambda_N \leqslant \big(|J_{\phi^{-1}}| \circ \phi\big) \cdot \phi^{-1}(\lambda_N). \tag{14.876}$$

Nous allons faire le produit de cette mesure par $|J_\phi|$ en nous souvenant que $J_\phi(x) = \det\big(d\phi_x\big)$. Par le lemme 10.284 nous avons aussi $(d\phi_x)^{-1} = d\phi^{-1}_{\phi(x)}$ et donc, par la propriété 8.16(3) du déterminant,

$$J_\phi(x) = \frac{1}{\det\big(d\phi^{-1}_{\phi(x)}\big)} = \frac{1}{J_{\phi^{-1}}\big(\phi(x)\big)}. \tag{14.877}$$

Nous avons

$$|J_\phi| \cdot \lambda_N \leqslant |J_\phi| \cdot \big(|J_{\phi^{-1}}| \circ \phi\big) \cdot \phi^{-1}(\lambda_N). \tag{14.878}$$

En utilisant la proposition 14.215, il s'agit de multiplier la mesure $\phi^{-1}(\lambda_N)$ par la fonction

$$x \mapsto |J_\phi(x) J_{\phi^{-1}}\big(\phi(x)\big)| = 1. \tag{14.879}$$

Nous avons donc bien

$$|J_\phi| \cdot \lambda_N \leqslant \phi^{-1}(\lambda_N), \tag{14.880}$$

et donc l'égalité

$$|J_\phi| \cdot \lambda_N = \phi^{-1}(\lambda_N), \tag{14.881}$$

c'est-à-dire le point (1).

(2) Le fait que la fonction proposée soit mesurable est le fait que la mesurabilité n'est pas affectée par produit et composition (propositions 14.104 et 14.45), et le fait que pour les mêmes raisons, l'application $J_\phi\colon U \to \mathbb{R}$ est également mesurable. En ce qui concerne la formule nous allons la démontrer dans le cas de fonctions de plus en plus générales.

(2a) **Pour les fonctions indicatrices** Soit B un borélien de U. Considérons la fonction $f = \mathbb{1}_{\phi(B)}$. Alors

$$\int_V f d\lambda_N = \int_{\mathbb{R}^N} \mathbb{1}_{\phi(B)}(y)\mathbb{1}_V(y) d\lambda_N(y) = \int_{\mathbb{R}^N} \mathbb{1}_{\phi(B)} d\lambda_N = \lambda_N\big(\phi(B)\big). \tag{14.882}$$

parce que $V = \phi(U)$ et $B \subset U$, donc $\mathbb{1}_{\phi(B)}\mathbb{1}_{\phi(U)} = \mathbb{1}_{\phi(B)}$. D'autre part, pour calculer l'autre membre de (14.855) nous remarquons que $f = \mathbb{1}_{\phi(B)} = \mathbb{1}_B \circ \phi^{-1}$, ce qui donne

$$\int_U f\big(\phi(x)\big)|J_\phi(x)| d\lambda_N(x) = \int_U \mathbb{1}_B |J_\phi| d\lambda_N = \int_B |J_\phi| d\lambda_N. \tag{14.883}$$

L'ensemble B étant borélien, il est extrêmement mesurable, ce qui fait que le point (1) s'applique : les expressions (14.882) et (14.883) sont égales.

(2b) **Pour les fonctions étagées**
Soit $f\colon V \to \mathbb{R}^+$ une fonction étagée :

$$f(x) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(x) \tag{14.884}$$

Nous pouvons faire le calcul suivant :

$$\int_V f d\lambda_N = \int_V \sum_i a_i \mathbb{1}_{A_i} d\lambda_N \tag{14.885a}$$
$$= \sum_i a_i \int_V \mathbb{1}_{A_i} d\lambda_N \tag{14.885b}$$
$$= \sum_i \int_U (\mathbb{1}_{A_i} \circ \phi)(x) |J_\phi(x)| d\lambda_N(x) \tag{14.885c}$$
$$= \sum_i a_i \int_U \mathbb{1}_{\phi^{-1}(A_i)} |J_\phi(x)| d\lambda_N(x) \tag{14.885d}$$
$$= \int_V \underbrace{\sum_i a_i \mathbb{1}_{\phi^{-1}(A_i)}(x)}_{=(f\circ\phi)(x)} |J_\phi(x)| d\lambda_N(x) \tag{14.885e}$$
$$= \int_V (f \circ \phi)|J_\phi| d\lambda_N. \tag{14.885f}$$

Justifications :

— Pour (14.885b) : linéarité de l'intégrale, théorème 14.183(2) [102]

— Pour (14.885c) : le cas des fonctions indicatrices est utilisé pour chaque i entre 1 et n.

(2c) **Fonction mesurable positive** Soit $f : V \to [0, \infty]$. Par le théorème fondamental d'approximation 14.118, il existe une suite croissante de fonctions étagées et mesurables $\varphi_n : V \to [0, \infty[$ dont la limite ponctuelle est f. Nous avons alors le calcul suivant :

$$\int_V f d\lambda_N = \lim_{n\to\infty} \int_V \varphi_n d\lambda_N \tag{14.886a}$$
$$= \lim_{n\to\infty} \int_U (\varphi_n \circ \phi)|J_\phi| d\lambda_N \tag{14.886b}$$
$$= \int_U \lim_{n\to\infty} (\varphi_n \circ \phi)|J_\phi| d\lambda_N \tag{14.886c}$$
$$= \int_U (f \circ \phi)|J_\phi| d\lambda_N. \tag{14.886d}$$

Justifications :

— Pour (14.886a), c'est le théorème de la convergence monotone 14.178.

— Pour (14.886b), c'est le présent théorème pour la fonction étagée φ_n.

— Pour (14.886c), c'est encore la convergence dominée, justifiée par le fait que $\varphi_n \circ \phi$ est également une suite croissante : si $x \in U$ alors $\varphi_{n+1}\big(\phi(x)\big) \geqslant \varphi_n\big(\phi(x)\big)$.

— Pour (14.886d), c'est la limite ponctuelle $\varphi_n\big(\phi(x)\big) \to f\big(\phi(x)\big)$.

(3) La partie sur l'intégrabilité repose sur le fait que $|f| \circ \phi = |f \circ \phi|$. Ici $|.|$ est le module et non une valeur absolue. Les faits suivants sont équivalents :

— la fonction $f : V \to \mathbb{C}$ est intégrable

— la fonction $|f| : V \to \mathbb{R}$ est intégrable.

— la fonction $(|f| \circ \phi)|J_\phi| : U \to \mathbb{R}$ est intégrable (par le point (2)).

— la fonction $(f \circ \phi)|J_\phi| : U \to \mathbb{R}$ est intégrable.

En ce qui concerne la formule, il s'agit seulement d'appliquer le point (2) aux parties positives, négatives, imaginaires et réelles de f.

□

102. Il est remarquable que nous n'utilisons cette linéarité que pour les fonctions étagées.

Notons que la formule peut être écrite sous la forme

$$\langle f, g\rangle_V = \langle f\circ\phi, (g\circ\phi)|J|\rangle_U, \tag{14.887}$$

EQooQKARooELPCEQ

qui est plus pratique lorsqu'on parle de produits scalaires. Pour rappel, $\phi\colon U\to C$ est un C^1-difféomorphisme.

14.282.
La formule de changement de variables peut être comprise de la façon suivante. Si ϕ est linéaire alors le facteur $|J_\phi|$ est la mesure de l'image par ϕ d'une portion de $\mathbb{R}^p$ de mesure 1, sinon $|J_\phi|$ est le rapport entre la mesure de l'image d'un élément infinitésimal de volume de $\mathbb{R}^p$ et sa mesure originale.

Soit $\phi(u,v) = g(u,v)e_1 + h(u,v)e_2$ un difféomorphisme dans $\mathbb{R}^2$. Soit (x_0,y_0) l'image par ϕ de (u_0,v_0). On considère le petit rectangle R de sommets (u_0,v_0), $(u_0+\Delta u, v_0)$, $(u_0+\Delta u, v_0+\Delta v)$ et $(u_0, v_0+\Delta v)$. L'image de R n'est pas un rectangle en général, mais peut être bien approximée par le rectangle de sommets (x_0,y_0), $(x_0,y_0)+\phi_u\Delta u$, $(x_0,y_0)+\phi_u\Delta u+\phi_v\Delta v$ et $(x_0,y_0)+\phi_v\Delta v$ et son aire est $\|\phi_u\times\phi_v\|\Delta u\Delta v$. La valeur $|\phi_u\times\phi_v|$ est exactement $|J_\phi|$

14.14.6 Exemples

Énormément d'exemples sont disponibles avec les coordonnées polaires et toutes leurs variations. Cependant les fonctions trigonométriques ne seront vues que plus tard ; les coordonnées polaires, cylindrique et sphériques seront vues en section 18.13 et les exemples d'utilisation pour les intégrales seront dans la section 18.15.

Un exemple avec une exponentielle sera donnée dans l'exemple 15.106.

14.15 Changement d'espace mesuré

PROPooILOEooBiumKD

Proposition 14.283 ([1]).
Soit un espace mesuré $(\Omega,\mathcal{A},\mu)$. Soit un ensemble Ω' et une bijection $\varphi\colon\Omega\to\Omega'$. Nous posons

(1) $\mathcal{A}' = \varphi(\mathcal{A})$,

(2) $\mu'(B) = \mu\big(\varphi^{-1}(B)\big)$ *pour tout* $B\in\mathcal{A}'$.

Soit enfin une fonction mesurable $f\colon\Omega\to X$.

Alors

(1) Le triple $(\Omega',\mathcal{A}',\mu')$ est un espace mesuré.

(2) L'application $f\circ\varphi^{-1}\colon\Omega'\to X$ est mesurable.

(3) Nous avons l'égalité

$$\int_\Omega f d\mu = \int_{\Omega'} (f\circ\varphi^{-1}) d\mu'. \tag{14.888}$$

Démonstration. La proposition 14.59 montre déjà que $(\Omega',\mathcal{A}',\mu')$ est un espace mesuré.

Soit une partie S mesurable dans X. Alors $f^{-1}(S)$ est mesurable dans Ω par hypothèse sur f, c'est-à-dire que $f^{-1}(S)\in\mathcal{A}$. Ensuite $(\varphi\circ f^{-1})(S)$ est mesurable dans Ω' par hypothèse sur φ. Cela prouve que $f\circ\varphi^{-1}$ est une application mesurable.

Nous avons encore à prouver l'égalité d'intégrale. Par la définition 14.167 nous avons

$$\int_\Omega f d\mu = \sup\{\sum_{i=1}^n a_i\mu(A_i)\} \tag{14.889}$$

où le supremum est sur tous les n et tous les choix de $A_i\in\mathcal{A}$, $a_i\in\mathbb{R}^+$ tels que $f|_{A_i} > a_i$. Vu que $\mathcal{A}' = \varphi(\mathcal{A})$, si $A_i\in\mathcal{A}$ et a_i sont choisis, nous avons aussi

$$f\circ\varphi^{-1}|_{\varphi(A_i)} \geqslant a_i \tag{14.890}$$

avec $\varphi(A_i) \in \mathcal{A}'$. Donc pour un choix de $\{(A_i, a_i)\}$ donné,

$$\sum_{i=1}^{n} a_i \mu(A_i) = \sum_{i=1}^{n} a_i \mu'\big(\varphi(A_i)\big). \tag{14.891}$$

Au final,

$$\int_\Omega f d\mu = \sup\{\sum_{i=1}^{n} a_i \mu(A_i)\} = \sup\{\sum_i a_i \mu'\big(\varphi(A_i)\big)\} = \int_{\varphi(\Omega)} f \circ \varphi^{-1} d\mu'. \tag{14.892}$$

□

REMooOMYYooNFiKOs

Remarque 14.284 (Ce n'est pas la mesure que nous voulons)**.**
La mesure donnée par la proposition 14.283 n'est pas celle que nous voulons d'habitude sur Ω'. Anticipons un peu pour comprendre. Prenons l'exemple de la partie C^2 de $\mathbb{R}$ donnée par

$$C = \{(x, y) \in \mathbb{R}^2 \text{ tel que } y = x^2, x \in]0, 3[\}. \tag{14.893}$$

(1) La façon correcte de définir la longueur de C est de prendre une limite d'approximations par des morceaux de droites, comme fait à la définition 21.3.

(2) Cette définition de la longueur peut être exprimée sous forme intégrale par le théorème 21.9 qui nous assure que

$$l(C) = \int_0^3 \|\varphi'(t)\| dt = \int_0^3 \sqrt{1 + 4t^2} dt \neq \mu'(C). \tag{14.894}$$

En effet, $\mu'(C) = \mu\big(\varphi^{-1}(C)\big) = \mu\big(]0, 3[\big) = 3$, alors que pour tout t nous avons $\sqrt{1 + 4t^2} > 1$ et donc $l(C) > 3$.

(3) Donc μ' n'est pas exactement ce que nous aurions pu vouloir appeler la « mesure » de C.

(4) La mesure à considérer sur C doit donc plutôt être quelque chose comme le produit de la mesure μ' par la fonction $\|\varphi'\|$. Mais cela est une autre histoire qui vous sera contée une autre fois.

ITEMooJTKCooYQknqo

(5) Dans le cas de S^1, nous avons $\varphi(x) = e^{ix}$, et $\|\varphi'(x)\| = 1$. Donc la mesure donné ici est probablement bien celle que nous voulons. Peut-être à coefficient $\frac{1}{2\pi}$ près pour avoir une normalisation $\mu'(S^1) = 1$. Cela est également une autre histoire qui vous sera contée une autre fois ; par exemple dans la proposition 18.70.

14.16 Primitives et intégrales

DEFooGLJDooFeZBBC

Définition 14.285.
Si $a < b$ nous posons

$$\int_a^b f(t)dt = \int_{[a,b]} f. \tag{14.895}$$

Si par contre $a > b$ nous posons $\int_a^b f = -\int_b^a f$.

Un petit cas particulier :

$$\int_{-\infty}^b f(t)dt = \int_{]-\infty,b]} f. \tag{14.896}$$

PropEZFRsMj

Proposition 14.286 (Primitive et intégrale[326])**.**
Soit f une fonction intégrable sur $[a, b]$ et continue sur $]a, b[$. Alors la fonction

$$\begin{aligned} F\colon [a, b] &\to \mathbb{R} \\ x &\mapsto \int_{[a,x]} f(t)dt. \end{aligned} \tag{14.897}$$

est l'unique primitive de f sur $]a, b[$ s'annulant en $x = a$.

Démonstration. Nous devons prouver que F est dérivable et que pour tout $x_0 \in]a,b[$ nous avons $F'(x_0) = f(x_0)$. Soit $\epsilon > 0$. Par continuité de f en x_0, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que

$$f(x_0 + h) = f(x_0) + \alpha(h) \tag{14.898}$$

avec $\lim_{h\to 0} \alpha(h) = 0$. Cette dernière limite signifie qu'il existe un $\delta > 0$ tel que $|\alpha(h)| < \epsilon$ pour tout h tel que $|h| < \delta$, c'est-à-dire pour tout $h \in B(0,\delta)$. À partir de maintenant nous ne considérons plus que de tels h.

Notre travail maintenant est de prouver que F est dérivable en x_0, et de montrer que la dérivée est $f(x_0)$. Pour cela,

$$\begin{aligned} F(x_0 + h) - F(x_0) &= \int_{x_0}^{x_0+h} f(t)dt & (14.899a)\\ &= \int_0^h f(x_0 + t)dt & (14.899b)\\ &= \int_0^h \big[f(x_0) + \alpha(t)\big]dt & (14.899c)\\ &= hf(x_0) + \int_0^h \alpha(t)dt. & (14.899d) \end{aligned}$$

Nous avons donc montré que pour tout $\epsilon > 0$, il existe un δ (défini via la fonction α) tel que $|h| < \delta$ implique

$$\left|\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0)\right| < \epsilon. \tag{14.900}$$

Cela signifie que

$$\lim_{h\to 0} \frac{F(x_0 + h) - F(x_0)}{h} = f(x_0), \tag{14.901}$$

qui n'est rien d'autre que le fait que F est dérivable en x_0 et que sa dérivée est $f(x_0)$.

Le fait que F s'annule en $x = a$ est par sa définition. L'unicité provient du corolaire 12.199. □

14.16.1 Théorème fondamental du calcul intégral

ThoRWXooTqHGbC

Théorème 14.287 (Théorème fondamental du calcul intégral)**.**
Soit f une fonction continue sur un intervalle ouvert I contenant strictement l'intervalle $[a,b] \subset \mathbb{R}$ et F une primitive de f sur I. Alors

$$\int_a^b f(t)dt = F(b) - F(a). \tag{14.902}$$

Démonstration. Nous avons vu par la proposition 14.286 que la fonction

$$\begin{aligned} G\colon [a,b] &\to \mathbb{R}\\ x &\mapsto \int_a^x f(t)dt \end{aligned} \tag{14.903}$$

était l'unique primitive de f sur $]a,b[$ à s'annuler pour $x = a$. Nous avons évidemment

$$\int_a^b f(t)dt = G(b). \tag{14.904}$$

Si F est une primitive quelconque, il suffit de soustraire sa valeur en $x = a$: $G(x) = F(x) - F(a)$ et donc

$$\int_a^b f(t)dt = G(b) = F(b) - F(a), \tag{14.905}$$

comme il fallait le prouver. □

Le théorème fondamental s'écrit souvent sous la forme [103]

$$f(x) = f(a) + \int_a^x f'(t)dt. \tag{14.906}$$

EqooBBCYooNweVrF

Sous cette forme, il faut penser que nous calculons $f(x)$ en un point pas trop éloigné de a, en sachant $f(a)$ et en intégrant la dérivée entre les deux.

Remarque 14.288.
Le lien entre primitive et intégrale est fondamentalement lié à l'invariance par translation de la mesure de Lebesgue, et non à la construction précise de cette mesure. Mais en même temps, la mesure de Lebesgue est l'unique à être invariante par translation.

Quelque remarques.

(1) Le théorème fondamental du calcul intégral est à utiliser pour calculer des intégrales des fonctions réelles lorsqu'on a des primitives sur un domaine strictement plus large que le domaine sur lequel nous voulons intégrer.
(2) Une version pour les intégrales impropres sera donnée au corolaire 14.299.
(3) Une primitive est forcément une fonction continue parce qu'une primitive est dérivable.
(4) Le théorème fondamental du calcul intégral ne sert pas qu'à calculer des intégrales à partir de primitives. Il sert aussi à démontrer des résultats plus théoriques, comme le théorème 12.374.
(5) En vertu du corolaire 12.199, une fonction ne possède qu'une seule primitive à constante près.

14.16.2 Exemples et applications

Si f est une fonction définie sur un intervalle I et y admettant des primitives, nous notons

$$\int f(x)dx \tag{14.907}$$

l'ensemble des primitives de f sur I :

$$\int f(x)dx = \{F(x) + C \text{ tel que } C \in \mathbb{R}\} \tag{14.908}$$

où F est une quelconque primitive de f.

Exemple 14.289.
Une primitive bien connue de $f : x \mapsto x^2$ est la fonction $F : x \to \frac{x^3}{3}$. Nous écrivons donc

$$\int x^2 dx = \frac{x^3}{3} + C. \tag{14.909}$$

Cela est un abus de notations terrible pour dire en réalité

$$\{x \mapsto \frac{x^3}{3} + C \text{ tel que } C \in \mathbb{R}\}. \tag{14.910}$$

△

En termes de notations, nous posons

$$\int_a^b f(t)dt = \Big[F(t)\Big]_{t=a}^{t=b} = F(b) - F(a). \tag{14.911}$$

Thfondcalc

103. Par exemple dans les théorèmes du reste des polynômes de Taylor 15.53 et de Cauchy-Lipschitz 17.39.

Remarque 14.290.
La valeur de l'intégrale ne dépend pas de la primitive qu'on choisi pour le calculer, car si F_1 et F_2 sont deux primitives de f alors $F_1 = F_2 + C$ et $F_1(b) - F_1(a) = (F_2(b) + C) - (F_2(a) + C) = F_2(b) - F_2(a)$.

Remarque 14.291.
Si l'intervalle d'intégration est réduit à un seul point alors la valeur de l'intégrale est zéro. Nous le savions déjà, et cela est cohérent avec le théorème fondamental car $\int_a^a f(t)dt = F(a) - F(a) = 0$.

Remarque 14.292.
Toute intégrale d'une fonction impaire sur un intervalle symétrique par rapport à l'origine est nulle.

PROPooJYIAooXLkbMx

Proposition 14.293 ([1]).
Soient des espaces vectoriels normés E et F où F est de dimension finie[104]*. Nous considérons une fonction $f\colon E \to F$ de classe C^1 ainsi qu'un chemin $\gamma\colon [0,1] \to E$ de classe C^1 également.*

Alors nous avons l'égalité

$$\int_0^1 (df)_{\gamma(t)}\big(\gamma'(t)\big) = f\big(\gamma(1)\big) - f\big(\gamma(0)\big). \tag{14.912}$$

Démonstration. Nous posons

$$\begin{aligned} g\colon [0,1] &\to F \\ t &\mapsto (f \circ \gamma)(t). \end{aligned} \tag{14.913}$$

Cette fonction vérifie $g'(t) = (df)_{\gamma(t)}\big(\gamma'(t)\big)$ par le lemme 12.280. Le théorème fondamental du calcul intégral[105] nous permet donc d'écrire

$$\int_0^1 (df)_{\gamma(t)}\big(\gamma'(t)\big)dt = \int_0^1 g'(t)df = g(1) - g(0). \tag{14.914}$$

Notons que g est continue grâce aux hypothèses de classe C^1 pour γ et f. □

La proposition suivante est la généralisation à $\mathbb{R}$ de la proposition 12.435.

PROPooKIASooGngEDh

Proposition 14.294.
Pour tout $\alpha \in \mathbb{R}$, si $f_\alpha(x) = x^\alpha$ alors

$$f'_\alpha(x) = \alpha x^{\alpha-1}. \tag{14.915}$$

Au niveau du domaine, c'est $\mathbb{R}$ auquel il faut enlever $\{0\}$ si $\alpha - 1 < 0$.

Démonstration. Soient $\alpha \in \mathbb{R}$ et une suite de rationnels α_i qui converge vers α. Les plus amateurs d'abstraction diront $(\alpha_i) \in \alpha$ en référence à la proposition 1.419.

Nous notons $f_\alpha(x) = x^\alpha$ et $f_i(x) = x^{\alpha_i}$. Par définition nous avons

$$f_i \to f_\alpha \tag{14.916}$$

ponctuellement. De plus en utilisant la proposition 12.435 nous savons que $f'_i(x) = \alpha_i x^{\alpha_i - 1}$. En posant $g(x) = \alpha x^{\alpha-1}$ nous avons donc

$$f'_i \to g. \tag{14.917}$$

ponctuellement. Mais f'_i est continue pour tout i et g également. Donc la convergence $f_i \to f_\alpha$ est uniforme sur tout compact[106]. Le théorème 12.374 nous permet de permuter limite et dérivée pour avoir $g = f'_\alpha$. □

104. Sinon l'intégrale dont nous allons parler n'est pas définie au sens où nous n'en avons pas donné de définition. Voir 14.195.
105. Théorème 14.287.
106. Proposition 12.364.

14.16.3 Intégrales impropres

SecGAVooBOQddU

Définition 14.295 ([326]).
*Une fonction $f\colon D \subset \mathbb{R} \to \mathbb{R}$ est **localement intégrable** sur un intervalle I si f est intégrable sur tout intervalle compact contenu dans I.*

PropCJAooQhNYkp

Proposition 14.296.
Soit $f\colon [a,b] \to \mathbb{R}$ une fonction intégrable. Alors

$$\int_{[a,b]} f = \lim_{x\to b^-} \int_a^x f. \tag{14.918}$$

EqPPMooBQDTYl

Démonstration. Notons que la valeur de f en b n'a strictement aucune importance parce que l'intégrale de Lebesgue ne dépend pas du choix de la valeur de la fonction en un ensemble de mesure nulle ; et en même temps la limite à gauche de (14.918) ne dépend pas non plus de la valeur de f en b. Bref si f n'est pas définie en b, nous pouvons poser $f(b) = 42$.

Notons de plus que du point de vue de l'intégrale de Lebesgue, $\int_{[a,b]}$ et $\int_{[a,b[}$ sont identiques et valent toutes les deux $\int_a^b$ (lorsque ça existe).

Supposons d'abord que f est positive. Alors nous posons $f_n = f\mathbb{1}_{[a,b-\frac{1}{n}]}$. Ponctuellement nous avons la limite croissante $f_n \to f$ et de plus

$$\lim_{x\to b^-} \int_{[a,x]} f = \lim_{n\to\infty} \int_{[a,b]} f_n. \tag{14.919}$$

Chacun des f_n est intégrable sur $[a,b]$. Le théorème de Beppo-Levi 14.178 implique que f est intégrable sur $[a,b]$ et que

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f. \tag{14.920}$$

Cela montre que dans le cas d'une fonction f positive nous avons bien (14.918).

Si f n'est pas positif, alors nous la décomposons en partie positive et négative $f = f^+ - f^-$ et par définition de l'intégrale d'une fonction non positive,

$$\lim_{x\to b^-} \int_{[a,x[} f = \lim \int f^+ - \lim \int f^-. \tag{14.921}$$

□

Il peut cependant arriver que la limite $\lim_{x\to b} \int_a^b f$ existe alors que f n'est pas intégrable sur $[a,b]$. C'est l'ennui des fonctions non positives. Un exemple classique est

$$\int_0^\infty \frac{\sin(t)}{t} dt \tag{14.922}$$

EqMMVooDSpgfz

La définition suivante est motivée par le lemme 14.261.

DEFooINPOooWWObEz

Définition 14.297 ([471]).
Si

$$\lim_{x\to b} \int_a^x f \tag{14.923}$$

*existe alors nous disons que l'intégrale est **convergente** en b. Ce procédé de limite est l'intégrale **impropre** de f sur $[a,b]$.*

Exemple 14.298 (Intégale impropre).
Nous considérons la fonction $f\colon [0,\infty[\to \mathbb{R}$ définie par

$$f(x) = \begin{cases} \frac{1}{n} & \text{si } x \in [2n-2, 2n-1[\\ -\frac{1}{n} & \text{si } x \in [2n-1, 2n[. \end{cases} \tag{14.924}$$

Par la divergence de la série harmonique, $\int_0^\infty |f|$ n'existe pas. La fonction f n'est donc pas intégrable au sens de Lebesgue (définition 14.186).

Cependant pour tout n pair nous avons

$$\int_0^n f = 0. \tag{14.925}$$

Du coup pour tout $x \geqslant 0$ nous avons

$$\int_0^x f = \int_{2n}^x f \tag{14.926}$$

où $2n$ est le plus grand nombre pair inférieur à x. Nous avons $|x - 2n| \leqslant 2$ et $|f(x)| \leqslant \frac{1}{n}$ pour $x \in [2n, x]$. Donc

$$\int_{2n}^x f \leqslant \frac{2}{n}. \tag{14.927}$$

Nous avons par conséquent

$$\lim_{x\to\infty} \int_0^x f = 0, \tag{14.928}$$

ce qui signifie que l'intégrale de f sur $[0, \infty[$ converge au sens des intégrales impropres. △

L'intégrale (14.922) est une intégrale convergente mais la fonction n'est pas intégrable (parce que pour être intégrale il faut que $|f|$ soit intégrable). Nous pouvons ainsi dire que cette intégrale converge mais n'existe pas.

Le corolaire suivant nous autorise à utiliser le théorème fondamental du calcul intégral 14.287 même dans les cas limites.

CorMUIooXREleR

Corolaire 14.299.
Si f est localement intégrable sur $[a, b]$ et si F est une primitive de f sur tout ouvert de $[a, b]$ alors

$$\int_a^b f = \lim_{x\to b^-} F(x) - F(a). \tag{14.929}$$

Démonstration. Pour chaque x dans $[a, b[$ nous avons

$$\int_a^x f = F(x) - F(a). \tag{14.930}$$

La proposition 14.296 nous explique que la limite $x \to b^-$ du membre de gauche existe et vaut $\int_a^b f$. Donc également le membre de droite :

$$\int_a^b f = \lim_{x\to b^-} \int_a^x f = \lim_{x\to b^-} F(x) - F(a). \tag{14.931}$$

□

La convergence des intégrales de fonctions $\frac{1}{x^\alpha}$ en 0 et ∞ est une question classique de l'intégration. De plus ces fonctions servent souvent à utiliser une théorème de comparaison (type intégrale dominée de Lebesgue).

PropBKNooPDIPUc

Proposition 14.300.
Deux intégrales remarquables.

(1) Nous avons

$$\int_0^1 \frac{1}{x^\alpha} = \infty \tag{14.932}$$

si et seulement si $\alpha \geqslant 1$.

ITEMooJFSXooHmgmEj

(2) *Nous avons*

$$\int_1^\infty \frac{1}{x^\alpha} = \infty \tag{14.933}$$

si et seulement si $\alpha \leqslant 1$.

Démonstration. La fonction $\frac{1}{x^\alpha}$ admet la primitive $F(x) = \frac{1}{1-\alpha}\frac{1}{x^{\alpha-1}}$ sur tout compact de $]0,\infty[$. Le corolaire 14.299 nous permet [107] de dire que $\int_0^1 \frac{1}{x^\alpha}$ vaudra

$$\lim_{x\to 0-+} \frac{1}{1-\alpha}\frac{1}{x^{\alpha-1}}. \tag{14.934}$$

Cela est strictement plus petit que ∞ si et seulement si $\alpha < 1$. □

14.17 Théorème de Fubini-Tonelli et de Fubini

Nous rappelons que $\mathbb{R}^n$ muni de la mesure de Lebesgue est un espace mesuré σ-fini, conformément à la définition 14.18.

Le théorème de Fubini-Tonelli parle de fonctions à valeurs réelles positives et non de fonctions à valeurs complexes. Le truc est que ce théorème va servir de base pour construire les autres. Si nous avons une fonction à valeurs complexes, elle se décompose en parties réelles et imaginaires qui elles-mêmes se décomposent en parties positives et négatives. Au final, les preuves pour $f\colon \Omega \to \mathbb{C}$ se ramènent à appliquer quatre fois le théorème pour $f\colon \Omega \to \bar{\mathbb{R}}^+$.

ThoWTMSthY

Théorème 14.301 (Fubini-Tonelli[448])**.**
Soient $(\Omega_i, \mathcal{A}_i, \mu_i)$ *deux espaces mesurés* σ*-finis, et* $(\Omega, \mathcal{A}, \mu)$ *l'espace produit. Soit une fonction* $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$ *une fonction mesurable et positive (valant éventuellement* ∞ *à certains endroits) Alors :*

ITEMooUTMNooVIBdpP

(1) *Les fonction*

$$F_1\colon x \mapsto \int_{\Omega_2} f(x,y)d\mu_2(y) \tag{14.935}$$

EQooWLADooQwNhEy

et

$$F_2\colon y \mapsto \int_{\Omega_1} f(x,y)d\mu_1(x) \tag{14.936}$$

sont mesurables.

ITEMooFKQUooCoCOLV

(2) *Toutes les intégrales imaginables existent et sont égales :* EqJRVtOGx

$$\int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\otimes\mu_2)(x,y) = \int_{\Omega_1}\left[\int_{\Omega_2} f(x,y)d\mu_2(y)\right]d\mu_1(x) \tag{14.937a}$$

$$= \int_{\Omega_2}\left[\int_{\Omega_1} f(x,y)d\mu_1(x)\right]d\mu_2(y) \tag{14.937b}$$

où tous les membres de l'égalité valent éventuellement $+\infty$.

Démonstration. Commençons par prouver le théorème dans le cas d'une fonction caractéristique d'un ensemble mesurable : $f(x,y) = \mathbb{1}_A(x,y)$ pour un certain ensemble $A \subset \Omega_1 \times \Omega_2$. Dans ce cas,

$$F_1(x) = \int_{\Omega_2} \mathbb{1}_A(x,y)d\mu_2(y) = \int_{\Omega_2} \mathbb{1}_{A_1(y)}(x)d\mu_2(y) = \mu_2\big(A_1(x)\big), \tag{14.938}$$

107. Tout ce que nous avons fait avec la borne b de l'intégrale $\int_a^b$ reste valable avec la borne a.

et nous avons déjà vu au théorème 14.245 que cette fonction F_1 était alors mesurable. En utilisant maintenant les égalités (14.754) ainsi que le fait que $\mathbb{1}_A(x,y) = \mathbb{1}_{A_2(x)}(y)$ nous avons

$$\int_{\Omega_1\times\Omega_2} \mathbb{1}_A(x,y)d(\mu_1\otimes\mu_2)(x,y) = (\mu_1\otimes\mu_2)(A) \tag{14.939a}$$
$$= \int_{\Omega_1} \mu_2\big(A_2(x)\big)d\mu_1(x) \tag{14.939b}$$
$$= \int_{\Omega_1}\left[\int_{\Omega_2} \mathbb{1}_{A_2(x)}(y)d\mu_2(y)\right]d\mu_1(x) \tag{14.939c}$$
$$= \int_{\Omega_1}\left[\int_{\Omega_2} \mathbb{1}_{A}(x,y)d\mu_2(y)\right]d\mu_1(x). \tag{14.939d}$$

Le théorème étant valable pour les fonctions caractéristiques, il est valable pour les fonctions simples (définition 14.112) par linéarité de l'intégrale.

Si f n'est pas une fonction simple, alors la proposition 14.118 nous donne une suite croissante de fonctions simples et positives convergeant ponctuellement vers f. La partie du théorème sur les fonctions simples dit que pour chaque n l'intégrale

$$\int_{\Omega_1\times\Omega_2} f_n(x,y)d(\mu_1\otimes\mu_2)(x,y) \tag{14.940}$$

peut être décomposée comme il faut en suivant la formule (14.937). Il faut pouvoir permuter la limite et l'intégrale dans chacun des cas. D'abord le théorème de la convergence monotone 14.178 appliqué à l'espace $\Omega_1\times\Omega_2$ dit que

$$\lim_{n\to\infty}\int_{\Omega_1\times\Omega_2} f_n(x,y)d(\mu_1\otimes\mu_2)(x,y) = \int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\otimes\mu_2)(x,y). \tag{14.941}$$

Ensuite, pour chaque $x\in\Omega_1$, les fonctions

$$\sigma_n(y) = \int_{\Omega_1} f_n(x,y)d\mu_1(x) \tag{14.942}$$

forment une suite croissante de fonctions mesurables ; nous leur appliquons encore le théorème de la convergence monotone :

$$\lim_{n\to\infty}\int_{\Omega_2}\left[\int_{\Omega_1} f_n(x,y)d\mu_1(x)\right]d\mu_2(y) = \lim_{n\to\infty}\int_{\Omega_2}\sigma_n(y)d\mu_2(y) \tag{14.943a}$$
$$= \int_{\Omega_2}\left[\lim_{n\to\infty}\int_{\Omega_1} f_n(x,y)d\mu_1(x)\right]d\mu_2(y) \tag{14.943b}$$
$$= \int_{\Omega_2}\left[\int_{\Omega_1} f(x,y)d\mu_1(x)\right]d\mu_2(y) \tag{14.943c}$$

où nous avons utilisé une seconde fois Beppo-Levi. □

Remarque 14.302.
Les formules (14.937) sont bien, mais ne garantissent en aucun cas que $f\in L^1(\Omega_1\times\Omega_2)$: il faut encore que ces intégrales soient finies.

CorTKZKwP

Corolaire 14.303 ([468])**.**
Soient $(\Omega_i,\mathcal{A}_i,\mu_i)$ deux espaces mesurés σ-finis, et $(\Omega,\mathcal{A},\mu)$ l'espace produit[108]*. Soit une fonction mesurable $f:\Omega\to\mathbb{R}$ ou $\mathbb{C}$. Alors les conditions suivantes sont équivalentes*

ITEMooZRAXooTRDIlZ

(1) $f\in L^1(\Omega_1\times\Omega_2)$,

108. Définition 14.247.

ITEMooJMPLooZKwxQC

(2)

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|d\mu_2\right]d\mu_1<\infty, \tag{14.944}$$

ITEMooLLBCooTRycwG

(3)

$$\int_{\Omega_2}\left[\int_{\Omega_1}|f|d\mu_1\right]d\mu_2<\infty. \tag{14.945}$$

Démonstration. Nous commençons par supposer que f est à valeurs dans $\mathbb{R}$. La notation $|f|$, pour l'instant, dénote donc bien la valeur absolue et non le module.

La fonction $|f|$ est mesurable et positive par hypothèse et par le fait que si f est mesurable, alors $|f|$ l'est également par le corolaire 14.108. Le théorème 14.301(2) nous dit alors que les intégrales suivantes existent et sont égales :

$$\int_{\Omega_1\times\Omega_2}|f|d(\mu_1\otimes\mu_2)=\int_{\Omega_1}\left[\int_{\Omega_2}|f(x,y)|d\mu_2(y)\right]d\mu_1(x)=\int_{\Omega_2}\left[\int_{\Omega_1}|f(x,y)|d\mu_1(x)\right]d\mu_2(y). \tag{14.946}$$

EQooAIQGooNtBQuC

Attention : rien ne dit encore que ces intégrales sont finies.

(1) **(1) implique (2) et (3)** Si $f\in L^1(\Omega_1\times\Omega_2)$ alors $|f|$ y est également. Cela implique que le membre de gauche de (14.946) est fini. Les deux autres sont alors également finis.

(2) **(2) ou (3) implique (1)** Les expressions à droite de (14.946) sont finies. Donc celle de gauche également. Cele signifie que $|f|\in L^1(\Omega_1\times\Omega_2)$. Par conséquent f est également dans $L^1(\Omega_2\times\Omega_2)$.

Nous passons maintenant au cas où f est à valeurs dans $\mathbb{C}$. Nous décomposons

$$f=f_R+if_I \tag{14.947}$$

où f_R et f_I sont des fonctions réelles. Nous avons

$$\int_\Omega|f|\leqslant\int_\Omega|f_R|+\int_\Omega|f_I|. \tag{14.948}$$

EQooZEOAooIMwKwk

Donc si f_R et f_I sont dans $L^1(\Omega)$, la fonction f le sera aussi. De même,

$$\int_\Omega|f_R|\leqslant\int_\Omega|f|, \tag{14.949}$$

qui donne l'inverse : si $f\in L^1(\Omega)$ alors $f_R,f_I\in L^1(\Omega)$. Bref, f est intégrable sur Ω si et seulement si f_R et f_I le sont.

Supposons que $f\in L^1(\Omega_1\times\Omega_2)$. Alors

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|d\mu_2\right]d\mu_1\leqslant\int_{\Omega_1}\left[\int_{\Omega_2}|f_R|\right]+\int_{\Omega_1}\left[\int_{\Omega_2}|f_I|\right]<\infty \tag{14.950a}$$

où nous avons appliqué (1) implique (2) aux fonctions f_R et f_I qui sont dans $L^1(\Omega_1\times\Omega_2)$ parce que f y est.

Dans l'autre sens, si

$$\int_{\Omega_1}\left[\int_{\Omega_2}|f|\right]<\infty, \tag{14.951}$$

alors en remplaçant $|f|$ par $|f_R|$ ou par $|f_I|$ nous restons fini. En appliquant alors « (2) implique (1) » nous trouvons que f_R et f_I sont dans $L^1(\Omega_1\times\Omega_2)$. Et cela implique que $f\in L^1(\Omega_1\times\Omega_2)$. □

ThoFubinioYLtPI

Théorème 14.304 (Fubini[468])**.**
Soient $(\Omega_i,\mathcal{A}_i,\mu_i)$ deux espaces mesurés σ-finis, et $(\Omega,\mathcal{A},\mu)$ l'espace produit. Soit

$$f\in L^1\big((\Omega,\mathcal{A}),\mathbb{C}\big), \tag{14.952}$$

c'est-à-dire une fonction à valeurs mesurable et intégrable sur Ω. Alors :

ITEMooVFGWooZTePQS

(1) *Pour presque tout $x \in \Omega_1$, la fonction $y \mapsto f(x,y)$ est $L^1(\Omega_2)$.*

ITEMooCYMKooUdizni

(2) *Si nous posons*

$$\varphi_f(x) = \int_{\Omega_2} f(x,y) d\mu_2(y); \tag{14.953}$$

alors $\varphi_f \in L^1(\Omega_1)$.

ItemQMWiolgiii

(3) *Nous avons la formule d'inversion d'intégrale*

$$\int_\Omega f d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \varphi_f d\mu_1 \tag{14.954a}$$

$$= \int_{\Omega_1} \left[\int_{\Omega_2} f(x,y) d\mu_2(y) \right] d\mu_1(x) \tag{14.954b}$$

$$= \int_{\Omega_2} \left[\int_{\Omega_1} f(x,y) d\mu_1(x) \right] d\mu_2(y). \tag{14.954c}$$

Démonstration. Nous commençons par supposer que f est à valeurs réelles : $f \in L^1\big((\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2), \mathbb{R}\big)$. Nous décomposons la fonction f en parties positives et négatives : $f = f^+ - f^-$ avec f^+ et f^- positives ou nulles. Nous avons évidemment

$$\int_{\Omega_1 \times \Omega_2} |f^+| \leqslant \int_{\Omega_1 \times \Omega_2} |f| < \infty. \tag{14.955}$$

Donc f^+ et f^- sont des éléments de $L^1(\Omega_1 \times \Omega_2)$.

(1) **Pour (1)**
Nous posons

$$\varphi_{f^+}(x) = \int_{\Omega_2} f^+(x,y) d\mu_2(y) \tag{14.956}$$

pour tous les $x \in \Omega_1$ pour lesquels cette intégrale est bien définie. Vu que f^+ est positive et mesurable, le théorème de Fubini-Tonelli 14.301(1) s'applique donc pour nous dire que φ_{f^+} est mesurable.
De plus le résultat (14.937) appliqué à f^+ donne

EQooSETWooRwkCnW

$$\int_{\Omega_1} \varphi_{f^+} d\mu_1 = \int_{\Omega_1 \times \Omega_2} f^+ d(\mu_1 \otimes \mu_2) < \infty. \tag{14.957}$$

Le fait que le tout soit fini est une conséquence du fait déjà mentionné que $f^+ \in L^1(\Omega_1 \times \Omega_2)$. Vu que φ_{f^+} est une fonction positive, l'inégalité (14.957) signifie que $\varphi_{f^+} \in L^1(\Omega_1, \mu_1)$.
En particulier, $\varphi_{f^+}(x) < \infty$ pour presque tout $x \in \Omega_1$. C'est-à-dire pour presque tout $x \in \Omega_1$:

$$\int_{\Omega_2} f^+(x,y) d\mu_2(y) < \infty, \tag{14.958}$$

et sachant que $f^+ \geqslant 0$ nous avons $f^+(x, \cdot) \in L^1(\Omega_2)$ pour presque tout x.

(2) **Pour (2)**
Partout où φ_{f^+} et φ_{f^-} sont finies nous avons

$$\varphi_f = \varphi_{f^+} - \varphi_{f^-}, \tag{14.959}$$

et comme cela a lieu presque partout, nous pouvons considérer une partie mesurable $A \subset \Omega_1$ telle que $\mu_1(A) = 0$ et $\varphi_f(x) = \varphi_{f^+}(x) - \varphi_{f^-}(x)$ pour tout x hors de A. Bref, nous posons

$$g(x) = \begin{cases} \varphi_{f^+} - \varphi_{f^-}(x) & \text{si } x \in A^c \\ 0 & \text{si } x \in A. \end{cases} \tag{14.960}$$

Cette fonction g est mesurable et $g = \varphi_f$ presque partout. De plus

$$\int_{\Omega_1} |g| d\mu_1 = \int_{A^c} |g| \leqslant \int_{A^c} \varphi_{f^+} + \int_{A^c} \varphi_{f^-} < \infty. \tag{14.961}$$

La dernière inégalité est le fait que $\varphi_{f^\pm}$ sont dans $L^1(\Omega_1)$. Et notons au passage que nous aurions pu laisser toutes les intégrales sur Ω_1 sans faire de précisions sur la distinction entre Ω_1 et A^c parce que la partie de Ω_1 sur laquelle $\varphi_{f^\pm}$ sont infinies est trop petite pour changer la valeur de l'intégrale.

Nous avons donc $g \in L^1(\Omega_1)$, et par conséquent également $\varphi_f \in L^1(\Omega_1)$ parce que ces deux fonctions sont égales presque partout (les classes sont égales).

(3) **Pour (3)**

En utilisant l'équation (14.957) nous avons

$$\begin{aligned}\int_{\Omega_1} \varphi_f d\mu_1 &= \int g d\mu_1 = \int_{\Omega_1} \varphi_{f^+} - \int_{\Omega_1} \varphi_{f^-} && \text{(14.962a)}\\ &= \int_{\Omega_1\times\Omega_2} f^+ d\mu - \int_{\Omega_1\times\Omega_2} f^- d\mu && \text{(14.962b)}\\ &= \int_{\Omega_1\times\Omega 2} f d\mu. && \text{(14.962c)}\end{aligned}$$

Et toutes ces intégrales sont finies.

Et c'est maintenant que nous considérons le cas complexe. Nous décomposons $f = f_R + if_I$ avec des fonctions réelles f_R et f_I. Comme déjà mentionné autour de (14.948), les fonctions f_R et f_I sont intégrables. Nous leur appliquons le théorème.

Les valeurs de x pour lesquelles $f_R(x, \cdot\,)$ et $f_I(x, \cdot\,)$ ne sont pas dans $L^1(\Omega_2)$ forment un ensemble de mesure nulle, nommons le A. En posant

$$g(x,y) = \begin{cases} f_R(x,y) + if_I(x,y) & \text{si } x \in A^c\\ 0 & \text{si } x \in A, \end{cases} \tag{14.963}$$

nous avons que $g(x, \cdot\,)$ est intégrable pour tout $x \in A^c$. Vu que pour ces valeurs de x nous avons $g(x,y) = f(x,y)$ nous en déduisons que pour $x \in A^c$ nous avons aussi $f(x, \cdot\,) \in L^1(\Omega_2)$.

Les autres points se traitent de la même façon [109]. □

NORMooKIRJooPvyPWQ

14.305.

En pratique, il n'est pas toujours évident qu'une fonction soit intégrable sur $\Omega_1 \times \Omega_2$. Pour permuter des intégrales sur une fonction à deux paramètres nous faisons comme suit.

(1) Nous testons l'intégrabilité en chaine de $|f|$, et si c'est bon, le corolaire 14.303 nous donne $f \in L^1(\Omega_1 \times \Omega_2)$.

(2) Nous utilisons le théorème de Fubini 14.304 pour séparer et permuter les intégrales comme des ingénieurs.

Si la fonction $(x,y) \mapsto f(x)g(y)$ satisfait aux hypothèse du théorème de Fubini alors

$$\int_{\Omega_1\times\Omega_2} f(x)g(y)dx \otimes dy = \left(\int_{\Omega_1} f(x)dx\right)\left(\int_{\Omega_2} g(y)dy\right). \tag{14.964}$$

EqTJEEsJW

Le théorème de Fubini est souvent utilisé sous cette forme.

109. Attention : je n'ai pas vérifié explicitement. C'est juste une intuition. Vérifiez et écrivez-moi pour dire si c'est bon ou non.

Exemple 14.306 (Nécessité d'avoir des mesures σ-finies)**.**
Nous montrons que le théorème ne tient pas si une des deux mesures n'est pas σ-finie. Soit $I = [0,1]$. Nous considérons l'espace mesuré

$$(I, \mathcal{B}\text{or}(I), \lambda) \tag{14.965}$$

où $\mathcal{B}\text{or}(I)$ est la tribu des boréliens sur I et λ est la mesure de Lebesgue (qui est σ-finie). D'autre part nous considérons l'espace mesuré

$$(I, \mathcal{P}(I), m) \tag{14.966}$$

où $\mathcal{P}(I)$ est l'ensemble des parties de I et m est la mesure de comptage. Cette dernière n'est pas σ-finie parce que les seuls ensembles de mesure finie pour la mesure de comptage sont des ensembles finis, or une union dénombrable d'ensemble finis ne peut pas recouvrir l'intervalle I.

Nous allons montrer que dans ce cadre, l'intégrale de la fonction indicatrice de la diagonale sur I^2 ne vérifie pas le théorème de Fubini. Étant donné que $\mathcal{B}\text{or}(I) \subset \mathcal{P}(I)$ nous avons

$$\mathcal{B}\text{or}(I^2) \subset \mathcal{B}\text{or}(I) \otimes \mathcal{P}(I). \tag{14.967}$$

Soit $\Delta = \{(x,x) \text{ tel que } x \in I\}$. La fonction

$$\begin{aligned} g\colon I^2 &\to \mathbb{R} \\ (x,y) &\mapsto x - y \end{aligned} \tag{14.968}$$

est continue et $\Delta = g^{-1}(\{0\})$ est donc fermé dans I^2. L'ensemble Δ est donc un borélien de I^2 et par conséquent un élément de la tribu $\mathcal{B}\text{or}(I)\otimes\mathcal{P}(I)$. La fonction indicatrice $\mathbb{1}_\Delta$ est alors mesurable pour l'espace mesuré

$$(I \times I, \mathcal{B}\text{or}(I) \otimes \mathcal{P}(I), \lambda \otimes m). \tag{14.969}$$

Pour x fixé nous avons

$$\mathbb{1}_\Delta(x,y) = \begin{cases} 1 & \text{si } y = x \\ 0 & \text{si } y \neq x \end{cases} = \mathbb{1}_{\{x\}}(y), \tag{14.970}$$

et donc

$$A_1 = \int_I \left(\int_I \mathbb{1}_\Delta(x,y) dm(y) \right) d\lambda(x) \tag{14.971a}$$

$$= \int_I \left(\int_I \mathbb{1}_{\{x\}}(y) dm(y) \right) d\lambda(x) \tag{14.971b}$$

$$= \int_I \Big(m(\{x\}) \Big) d\lambda(x) \tag{14.971c}$$

$$= \int_I 1 d\lambda(x) \tag{14.971d}$$

$$= 1. \tag{14.971e}$$

Par contre le support de $\mathbb{1}_\Delta$ étant de mesure nulle pour la mesure de Lebesgue, nous avons

$$\int_I \mathbb{1}_\Delta(x,y) d\lambda(x) = 0 \tag{14.972}$$

et par conséquent

$$A_2 = \int_I \left(\int_I \mathbb{1}_\Delta(x,y) d\lambda(x) \right) dm(y) = 0. \tag{14.973}$$

Nous voyons donc que le théorème de Fubini ne s'applique pas. △

EXooLUFAooGcxFUW

Exemple 14.307.
Nous nous proposons de calculer l'intégrale suivante en utilisant le théorème de Fubini :

$$G = \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} \tag{14.974}$$

alors que la fonction $x \mapsto e^{-x^2}$ n'a pas de primitives parmi les fonctions élémentaires.

Nous allons le faire de deux façons. Une première directe en utilisant le théorème de Fubini sur un domaine non borné, et une seconde en utilisant Fubini sur un domaine borné, et en passant à la limite ensuite.

(1) **Fubini, domaine non borné**

Par symétrie nous pouvons nous contenter de calculer

$$G_+ = \int_0^\infty e^{-x^2} dx. \tag{14.975}$$

L'astuce est de passer par l'intermédiaire

$$\begin{aligned} H &= \int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-(x^2+y^2)} dxdy && \text{(14.976a)} \\ &= \int_{\mathbb{R}^+} \left(\int_{\mathbb{R}^+} e^{-x^2} e^{-y^2} dx \right) dy && \text{(14.976b)} \\ &= \left(\int_{\mathbb{R}^+} e^{-x^2} dx \right)^2 && \text{(14.976c)} \\ &= G_+^2 && \text{(14.976d)} \end{aligned}$$

L'intégrale (14.976a) se calcule en passant aux coordonnées polaires et le résultat est $H = \frac{\pi}{4}$. Nous avons alors $G = \frac{\sqrt{\pi}}{2}$ et

$$\int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi}. \tag{14.977}$$

(2) **Fubini, domaine borné, puis limite** Une variante, qui n'applique pas Fubini sur un domaine non borné. Nous commençons par écrire

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx := \lim_{R \to +\infty} \int_{-R}^{+R} e^{-x^2} dx \tag{14.978}$$

et puis nous faisons le calcul

$$\begin{aligned} I^2 &= \lim_{R \to +\infty} \left((\int_{-R}^{+R} e^{-x^2} dx)(\int_{-R}^{+R} e^{-y^2} dy) \right) \\ &= \lim_{R \to +\infty} \left(\iint_{K_R} e^{-(x^2+y^2)} dxdy \right) && \text{(14.979)} \\ &= \lim_{R \to +\infty} \left(\iint_{C_R} e^{-(x^2+y^2)} dxdy \right) \end{aligned}$$

où K est le carré de demi-côté R centré à l'origine et de côtés parallèles aux axes et C_R est le cercle de rayon R centré à l'origine.

La première étape à justifier est simplement l'application de Fubini. Pour le passage de l'intégrale du carré vers le cercle, définissons

$$I_K(r) = \int_{K_r} f, \quad I_C(r) = \int_{C_r} f \tag{14.980}$$

où K_r est la carré de demi-côté r et C_r est le cercle de rayon r. Le demi-côté du carré inscrit à C_r est $\sqrt{2}$, donc pour tout r nous avons

$$I_K(\sqrt{2}r) \geqslant I_C(r) > I_K(r), \tag{14.981}$$

et en prenant la limite, nous avons évidement

$$\lim_{r\to\infty} I_K(\sqrt{2}r) = \lim_{r\to\infty} I_K(r), \tag{14.982}$$

et donc cette limite est également égale à $\lim_{r\to\infty} I_C(t)$.
Il ne reste qu'à calculer la dernière intégrale sur le cercle en passant aux coordonnées polaires :

$$\int_{C_R} e^{-(x^2+y^2)}dxdy = \int_0^{2\pi} d\theta \int_0^R re^{-r^2}dr = \pi(1-e^{-R^2}). \tag{14.983}$$

La limite donne π, nous en déduisons que

$$\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}. \tag{14.984}$$

EqFDyHTs

△

Le théorème de Fubini-Tonelli nous permet également d'inverser des sommes et des séries. En effet une somme n'est rien d'autre qu'une intégrale pour la mesure de comptage :

$$\sum_{n=0}^{\infty} a_n = \int_{\mathbb{N}} a_n dm(n). \tag{14.985}$$

La proposition suivante montre comment il faut faire.

PropInversSumIntFub

Proposition 14.308.
Soient les espaces mesurés $(\mathbb{N}, \mathcal{P}(\mathbb{N}), m)$, $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \lambda)$ où λ est la mesure de Lebesgue ainsi qu'une suite de fonctions positives $f_n\colon \mathbb{R}^d \to \mathbb{R}$. Nous supposons de plus que la fonction f_n soit intégrable pour tout n et que les résultats forment une suite sommable. Alors

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}^n} f_n(x)dx = \int_{\mathbb{R}^d} \sum_{n\in\mathbb{N}} f_n(x)dx. \tag{14.986}$$

Démonstration. Nous pouvons la récrire le membre de gauche sous la forme

$$\int_{\mathbb{N}} \left(\int_{\mathbb{R}^n} f(n,x)dx \right) dm(n) \tag{14.987}$$

avec la notation évidente $f(n,x) = f_n(x)$. Prouvons que la fonction $f\colon \mathbb{N}\times\mathbb{R}^d \to \mathbb{R}$ ainsi définie est une fonction mesurable pour l'espace mesuré

$$\big(\mathbb{N}\times\mathbb{R}^d, \mathcal{P}(\mathbb{N})\otimes\mathcal{B}\mathrm{or}(\mathbb{R}^d), m\otimes\lambda\big). \tag{14.988}$$

Si $A\subset\mathbb{R}$, nous avons

$$f^{-1}(A) = \bigcup_{n\in\mathbb{N}} \{n\}\times f_n^{-1}(A). \tag{14.989}$$

Chacun des ensembles dans l'union appartient à la tribu $\mathcal{P}(\mathbb{N})\times\mathcal{B}\mathrm{or}(\mathbb{R}^d)$ tandis que les tribus sont stables sous les unions dénombrables. La fonction f est donc mesurable. Comme nous avons supposé que f était positive, le théorème de Fubini-Tonelli s'applique et nous avons

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{N}} f(n,x)dm(n) \right) dx = \int_{\mathbb{R}^d} \sum_{n\in\mathbb{N}} f_n(x)dx. \tag{14.990}$$

□

ThoFubini

Théorème 14.309 (Fubini)**.**
Soit $(x,y)\mapsto f(x,y)\in\bar{\mathbb{R}}$ une fonction intégrable sur $B_n\times B_m\subset\mathbb{R}^{n+m}$ où B_n et B_m sont des ensembles mesurables de $\mathbb{R}^n$ et $\mathbb{R}^m$. Alors :

(1) *pour tout $x \in B_n$, sauf éventuellement en les points d'un ensemble $G \subset B_n$ de mesure nulle, la fonction $y \in B_m \mapsto f(x,y) \in \bar{\mathbb{R}}$ est intégrable sur B_m*

(2) *la fonction*

$$\begin{aligned} B_n \backslash G &\to \mathbb{R} \\ x &\mapsto \int_{B_m} f(x,y)dy \end{aligned} \tag{14.991}$$

est intégrable sur $B_n \backslash G$.

(3) *On a*

$$\int_{B_n \times B_m} f(x,y)dxdy = \int_{B_n} \left(\int_{B_m} f(x,y)dy \right) dx. \tag{14.992}$$

Notons en particulier que si $f(x,y) = \varphi(x)\phi(y)$, alors $\int_{B_m} \varphi(y)dy$ est une constante qui peut sortir de l'intégrale sur B_n, et donc

$$\int_{B_n \times B_m} \varphi(x)\phi(y)dxdy = \int_{B_n} \varphi(x)dx \int_{B_m} \phi(y)dy. \tag{14.993}$$ EqFubiniFactori

14.17.1 Utiliser Fubini sur $\mathbb{R}^n$

Nous montrons à présent une série de théorème qui disent que sur $\mathbb{R}^n$, nous pouvons intégrer dans l'ordre qu'on veut.

PROPooODWYooWSlHlU

Proposition 14.310 ([1]).
Si $f \in L^1(\mathbb{R}^n)$, alors

$$\int_{\mathbb{R}^n} f = \int_{\mathbb{R}^{n-1}} \left[\int_{\mathbb{R}} f_{(x_1,\dots,x_{n-1})}(t)t \right] dx_1, \dots, dx_{n-1} \tag{14.994}$$

où, pour $(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$, nous définissons

$$\begin{aligned} f_{(x_1,\dots,x_{n-1})} \colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(x_1, \dots, x_{n-1}, t) \end{aligned} \tag{14.995}$$

Démonstration. Il s'agit d'une application du théorème de Fubini 14.304(3) avec $\Omega_1 = \mathbb{R}^{n-1}$ et $\Omega_2 = \mathbb{R}$. □

LEMooEDQUooZudeex

Lemme 14.311 ([1]).
Soit un ouvert U de $\mathbb{R}^n$. Soit une fonction C^∞ à support compact $f \colon U \to \mathbb{R}$. Alors l'application

$$\begin{aligned} g \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0 & si\ x \notin U \\ f(x) & si\ x \in U \end{cases} \end{aligned} \tag{14.996}$$

est de classe C^∞.

PROPooCQEIooENbfux

Proposition 14.312 ([1]).
Soit un ouvert U de $\mathbb{R}^n$. Soit une fonction C^∞ à support compact $f \colon U \to \mathbb{R}$. Alors

$$\int_U \partial_i f = 0. \tag{14.997}$$

Démonstration. En deux parties très inégales.

(1) **Si $i = n$** Nous considérons l'application du lemme 14.311 :

$$\begin{aligned} g\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{si } x \notin U \\ f(x) & \text{si } x \in U \end{cases} \end{aligned} \tag{14.998}$$

Elle est de classe C^∞ et vérifie

$$\int_U \partial_n f = \int_{\mathbb{R}^n} \partial_n g. \tag{14.999}$$

Soit $a \in \mathbb{R}^n$. Nous posons

$$\begin{aligned} h\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto g(a + te_n). \end{aligned} \tag{14.1000}$$

Nous avons

$$\begin{aligned} (\partial_n g)(a + ue_n) &= \frac{d}{dt}\big[g(a + ue_n + te_n)\big]_{t=0} && \text{(14.1001a)} \\ &= \frac{d}{dt}\big[g\big(a + (t+u)e_n\big)\big]_{t=0} && \text{(14.1001b)} \\ &= \frac{d}{dt}\big[h(t+u)\big]_{t=0} && \text{(14.1001c)} \\ &= h'(u). && \text{(14.1001d)} \end{aligned}$$

Nous pouvons donc déjà calculer une intégrale :

$$\begin{aligned} \int_{\mathbb{R}} (\partial_n g)(a + te_n)dt &= \int_{\mathbb{R}} h'(t)dt && && \text{(14.1002a)} \\ &= \lim_{R\to\infty} \int_{-R}^{R} h'(t)dt && \text{prop. 14.262} && \text{(14.1002b)} \\ &= \lim_{R\to\infty} \Big(h(R) - h(-R)\Big) && \text{tho. 14.287} && \text{(14.1002c)} \\ &= 0. && && \text{(14.1002d)} \end{aligned}$$

La dernière égalité est parce que $h(R) = g(a + Re_t)$ et que g est à support compact.
Nous utilisons la proposition 14.310 pour isoler l'intégrale sur x_n dans $\int_{\mathbb{R}^n} \partial_i g$:

$$\int_{\mathbb{R}^n} \partial_n g = \int_{\mathbb{R}^{n-1}} \left[\int_{\mathbb{R}} (\partial g)(a + te_n)dt\right] da_1 \dots da_{n-1} = 0. \tag{14.1003}$$

(2) **Si $i \neq n$**
Nous considérons l'application $\sigma\colon \mathbb{R}^n \to \mathbb{R}^n$ qui permute e_i et e_n. Nous avons $\int_{\mathbb{R}^n} f \circ \sigma = \int_{\mathbb{R}^n} f$ par le changement de variables, théorème 14.281. Cela nous ramène au cas $i = n$.

□

14.17.2 Absolue continuité et variation bornée

DEFooLZWPooSSRaMv

Définition 14.313 ([472, 473]).
*Soit un intervalle I de $\mathbb{R}$. Une fonction $f\colon I \to \mathbb{R}$ est **absolument continue** si pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que pour tout suite finie $\{[a_n, b_n]\}_{n\leqslant N}$ de sous-intervalles de I d'intérieurs disjoints,*

$$\sum_{n\geqslant 0} (b_n - a_n) < \delta \Rightarrow \sum_{n\geqslant 0} |f(a_n) - f(b_n)| < \epsilon. \tag{14.1004}$$

PROPooNZYKooSlwFYK

Proposition 14.314 ([472]).
Soient des applications absolument continues f et g sur $[a, b]$. Alors

(1) $f+g$, $f-g$ et fg sont absolument continues.

(2) Si il existe $C>0$ tel que $|g(x)|\geqslant C$ pour tout $x\in[a,b]$, alors f/g est absolument continue.

PROPooHCCWooHaYitq

Proposition 14.315 ([472]).
Si $f\in L^1\big([a,b]\big)$, alors la fonction

$$\begin{aligned} F\colon [a,b]&\to\mathbb{R}\\ x&\mapsto\int_a^x f(t)dt \end{aligned} \tag{14.1005}$$

est absolument continue.

PROPooUMXIooUhbZvl

Proposition 14.316 ([472]).
Une application F sur $[a,b]$ est absolument continue[110] *si et seulement si il existe une application f intégrable sur $[a,b]$ telle que pour tout $x\in[a,b]$,*

$$F(x)-F(a)=\int_a^x f(t)dt. \tag{14.1006}$$

DEFooLANWooGnVGPV

Définition 14.317 (Variation totale, variation bornée[473, 474]).
Soit un intervalle I de $\mathbb{R}$. Pour $n\in\mathbb{N}$ nous définissons

$$\mathcal{P}_n=\{(x_0,\dots,x_n)\text{ tel que }a=x_0\leqslant x_1\leqslant\dots\leqslant x_n=b\}. \tag{14.1007}$$

et

$$\mathcal{P}=\bigcup_{n=1}^{\infty}\mathcal{P}_n. \tag{14.1008}$$

*Si $P\in\mathcal{P}$, nous notons $l(P)$ le $n\in\mathbb{N}$ tel que $P\in\mathcal{P}_n$. La **variation totale** d'une application application $f\colon I\to\mathbb{C}$ est le nombre*

$$V_a^b(f)=\sup_{P\in\mathcal{P}}\sum_{i=0}^{l(P)-1}|f(P_{i+1})-f(P_i)|. \tag{14.1009}$$

*Nous disons que f est à **variation bornée** sur $[a,b]$ si $V_a^b(f)<\infty$. L'ensemble des fonctions à variation bornée de $[a,b]$ est noté $BV\big([a,b]\big)$.*

PROPooDJTDooMMWMcS

Proposition 14.318 ([473]).
Si f est à variation bornée sur $[a,b]$ et si $c\in[a,b]$, alors

$$V_a^b(f)=V_a^c(f)+V_c^b(f). \tag{14.1010}$$

PROPooMONZooWVkYWb

Proposition 14.319 ([472]).
Soit f absolument continue sur $[a,b]$. Alors f est à variation bornées sur $[a,b]$.

PROPooVHGNooSDlrrq

Proposition 14.320 ([473]).
À propos de fonctions à variation bornées.

(1) *Une fonction $f\colon[a,b]\to\mathbb{R}$ est à variation bornée si et seulement si elle est la différence entre deux fonctions monotones sur $[a,b]$.*

(2) *Si $f\in BV\big([a,b]\big)$, alors $f'(x)$ existe pour presque tout $x\in[a,b]$.*

110. Définition 14.313.

14.17.3 Dérivation dans le bornes

La proposition suivante sera utilisée entre autres pour montrer que sous l'hypothèse d'une densité continue, la loi exponentielle est sans mémoire, proposition 36.150.

PropDerrFnAvecBornesFonctions

Proposition 14.321.
Soit $f(x,t)$ une fonction continue sur $[\alpha,\beta]\times[a,b]$, telle que $\frac{\partial f}{\partial x}$ existe et soit continue sur $]\alpha,\beta[\times[a,b]$. Soient $\varphi(x)$ et $\psi(x)$, des fonctions continues de $[\alpha,\beta]$ dans $\mathbb{R}$ et admettant une dérivée continue sur $]\alpha,\beta[$. Alors la fonction

$$F(x)=\int_{\varphi(x)}^{\psi(x)} f(x,t)dt \tag{14.1011}$$

admet une dérivée continue sur $]\alpha,\beta[$ et

$$\frac{dF}{dx}=\int_{\varphi(x)}^{\psi(x)}\frac{\partial f}{\partial x}(x,t)dt+f\big(x,\psi(x)\big)\cdot\frac{d\psi}{dx}-f\big(x,\varphi(x)\big)\cdot\frac{d\varphi}{dx}. \tag{14.1012}$$

EqFormDerrFnAvecBorneNInt

L'exemple qui suit devrait pouvoir être rendu rigoureux en utilisant des distributions correctement.

ExfYXeQg

Exemple 14.322.
Si g est une fonction continue, la fonction suivante est une primitive de g :

$$\int_0^x g(t)dt=\int_0^\infty g(t)\mathbb{1}_{t<x}(t)dt. \tag{14.1013}$$

Nous nous proposons de justifier *de façon un peu heuristique* le fait que ce soit bien une primitive de g en considérant la fonction

$$f(t,x)=g(t)\mathbb{1}_{t<x}(t). \tag{14.1014}$$

Nous posons

$$F(x)=\int_0^\infty f(x,t)dt, \tag{14.1015}$$

et nous calculons F' en permutant la dérivée et l'intégrale [111]. D'abord,

$$f(t,x)=\begin{cases}g(t) & \text{si } t\in[0,x]\\ 0 & \text{sinon.}\end{cases} \tag{14.1016}$$

La dérivée de f par rapport à x est donnée par la distribution

$$\frac{\partial f}{\partial x}(t_0,x_0)=g(t_0)\delta(t_0-x_0). \tag{14.1017}$$

Donc

$$F'(x_0)=\int_0^\infty\frac{\partial f}{\partial x}(t,x_0)dt=\int_0^\infty g(t)\delta(t-x_0)=g(x_0), \tag{14.1018}$$

comme attendu. △

Cet exemple sera rendu rigoureux par les propositions 14.327 et 14.328.

PROPooDOTQooWaMBrO

Proposition 14.323 ([1]).
Soit un espace mesuré $(\Omega,\mathcal{A},\mu)$. Soient des fonctions réelles et mesurables f_n et g sur Ω. Si

(1) $f_n(x)\leqslant g(x)$ pour presque tout $x\in\Omega$,

(2) $\lim_{n\to\infty}f_n(x)=f(x)$ pour tout $x\in\Omega$,

alors $f(x)\leqslant g(x)$ pour presque tout $x\in\Omega$.

111. Ceci n'est pas rigoureux : il faudrait avoir un théorème à propos de distributions qui permet de le faire.

Démonstration. Notons $Z_n = \{x \in \Omega \text{ tel que } f_n(x) > g(x)\}$. La partie $Z = \bigcup_{n\in\mathbb{N}} Z_n$ est de mesure nulle par le lemme 14.28. Pour chaque $x \in \Omega\backslash Z$, nous avons $f_n(x) \leqslant g(x)$ et nous pouvons passer à la limite pour obtenir $f(x) \leqslant g(x)$. Cette dernière égalité étant valable sur $\Omega\backslash Z$, elle est valable presque partout. □

PROPooRUSEooFYGLLU

Proposition 14.324 ([1]).
Si f est une fonction intégrable et continue sur un voisinage de a, alors

$$\lim_{n\to\infty} n\int_a^{a+1/n} f = f(a). \tag{14.1019}$$

Voir le reste du thème 63.

PROPooZOJHooKwZOFW

Proposition 14.325 ([1]).
Soit f intégrable sur $[a,b]$ vérifiant

$$\int_a^c f = 0 \tag{14.1020}$$

pour tout $c \in [a,b]$. Alors $f = 0$ presque partout.

PROPooWFSPooIBogJV

Proposition 14.326 ([473]).
Soit $f\colon [a,b] \to \mathbb{R}$ croissante et presque partout dérivable. Alors f' est mesurable et

$$\int_a^b f'(t)dt \leqslant f(b) - f(a) \tag{14.1021}$$

PROPooUMIPooQGXaPd

Proposition 14.327 ([472]).
Soit une fonction réelle $f \in L^1\big([a,b]\big)$. Nous définissons

$$\begin{aligned} F\colon [a,b] &\to \mathbb{R} \\ F(x) &\mapsto \int_a^x f(t)dt. \end{aligned} \tag{14.1022}$$

Alors F est dérivable presque partout sur $[a,b]$ et là où F est dérivable, nous avons $F' = f$.

Démonstration. Nous divisons en plusieurs cas de généralité croissante.

(1) **En supposant que f est bornée** Nous supposons que $|f(x)| < K$ pour tout $x \in [a,b]$. Pour chaque $n \geqslant 1$, nous posons

$$\begin{aligned} g_n\colon [a,b] &\to \mathbb{R} \\ x &\mapsto \frac{F(x+1/n) - F(x)}{1/n}. \end{aligned} \tag{14.1023}$$

En développant un peu la valeur de F là-dedans, nous trouvons

$$g_n(x) = n\int_x^{x+1/n} f. \tag{14.1024}$$

Vu que $|f(x)| < K$, nous avons

$$|g_n(x)| \leqslant n\int_x^{x+1/n} K = Kn\frac{1}{n} = K. \tag{14.1025}$$

Par la proposition 14.315, F est absolument continue ; la proposition 14.319 dit alors qu'elle est à variation bornée, et enfin 14.320 implique que F est dérivable presque partout. Autrement dit, la limite $\lim_{n\to\infty} g_n(x)$ existe pour presque tout x. Et là où elle existe, elle vaut forcément F'.

Soit $c \in [a,b]$. Utilisons à présent le théorème de la convergence dominée de Lebesgue 14.207 sur $[a,c]$. Vérifions les hypothèses :

— Pour tout n, nous avons $g_n \in L^1\big([a,b]\big)$.
— Nous avons $g_n \to F'$ ponctuellement presque partout.
— La fonction constante K est intégrable sur $[a,b]$ et majore $|g_n(x)|$ pour tout n et pour tout x.

Nous avons

$$\lim_{n\to\infty}\int_a^c g_n = \int_a^c F'. \tag{14.1026}$$

L'intégrale dans le membre de gauche est :

$$\int_a^c g_n = \int_a^c \frac{F(x+1/x)-F(x)}{1/n}dx = n\int_a^c \left[\int_a^{x+1/n} f - \int_a^x f\right]dx. \tag{14.1027}$$

Nous nous intéressons d'abord à la première intégrale avec le changement de variables [112] $u = x + 1/x$:

$$\int_a^c \left[\int_a^{x+1/n} f(t)dt\right] = \int_a^c F(x+1/n) = \int_{a+1/n}^{c+1/n} F(u)du. \tag{14.1028}$$

En remettant dans (14.1027),

$$\int_a^c g_n = n\int_{a+1/n}^{c+1/n} F(x)dx - n\int_a^c F(x)dx = n\int_c^{c+1/n} F - n\int_a^{a+1/n} F \tag{14.1029}$$

Et nous pouvons enfin passer tranquillement à la limite $n \to \infty$ en utilisant la proposition 14.324 :

$$\lim_{n\to\infty}\int_a^c g_n = F(c) - F(a). \tag{14.1030}$$

Et en reprenant depuis (14.1026),

$$\int_a^c (F' - f) = 0 \tag{14.1031}$$

pour tout $c \in [a,b]$. Cela prouve [113] que $F'(x) = f(x)$ presque partout [114].

(2) **Si $f \in L^1$ et $f \geqslant 0$** Nous supposons maintenant que $f \in L^1\big([a,b]\big)$ et que $f \geqslant 0$. Nous posons

$$f_n(x) = \begin{cases} f(x) & \text{si } 0 \leqslant f(x) \leqslant n \\ 0 & \text{sinon,} \end{cases} \tag{14.1032}$$

ainsi que

$$F_n(x) = \int_a^x f_n(t)dt \tag{14.1033}$$

et

$$G_n(x) = \int_a^x \big(f(t) - f_n(t)\big). \tag{14.1034}$$

Nous avons $F = F_n + G_n$, et, comme $f(t) - f_n(t) \geqslant 0$, la fonction G est croissante.
La fonction f_n étant majorée par n, la première partie d'applique et nous avons

$$F_n'(x) = f_n(x) \tag{14.1035}$$

pour presque tout x. Le même raisonnement que précédemment, mettant en scène les propositions 14.315, 14.319 et 14.320 montre que F est dérivable presque partout. Donc dans

112. Théorème 14.281.
113. Par la proposition 14.325.
114. Ici, le « presque partout » est l'intersection du « presque partout » où F est dérivable et le « presque partout » de la proposition 14.325.

l'égalité $F = F_n + G_n$, les fonctions F et F_n sont dérivables. Nous en déduisons que G_n est également dérivable et que

$$F'(x) = F_n'(x) + G_n'(x) \tag{14.1036}$$

pour presque tout x. Comme G_n est croissante, nous avons $G_n'(x) \geqslant 0$ et la majoration

$$F'(x) = F_n'(x) + G_n'(x) \geqslant F_n'(x) = f_n(x). \tag{14.1037}$$

En passant à la limite et en utilisant la proposition 14.323, nous avons

$$F'(x) \geqslant f(x) \tag{14.1038}$$

pour presque tout $x \in [a,b]$. En intégrant cette inégalité de a à b [115] nous avons

$$\int_a^b F'(x)dx \geqslant \int_a^b f(x)dx = F(b) - F(a). \tag{14.1039}$$

Ceci n'a aucun rapport avec cela, mais je te rappelle que la proposition 14.326 [116] dit que

$$\int_a^b F'(x)dx \leqslant F(b) - F(a). \tag{14.1040}$$

Nous avons donc

$$\int_a^b F' = F(b) - F(a) = \int_a^b f, \tag{14.1041}$$

et donc

$$\int_a^b \big[F' - f\big] = 0 \tag{14.1042}$$

Vu que $F' - f \geqslant 0$, le fait que l'intégrale soit nulle implique que $F' - f = 0$ presque partout [117]

(3) **Le cas générique** Si $f \in L^1\big([a,b]\big)$. Nous décomposons f en parties positives et négatives : $f = f^+ + f^-$. En posant $F_+ = \int_a^b f^+$ et $F_- = \int_a^b f^-$, nous avons $F = F_+ + F_-$.
En appliquant le cas positif à f^+ nous avons immédiatement $F_+' = f^+$ presque partout. Ensuite nous considérons $g(x) = -f^-(x)$ et $G(x) = \int_a^b G$. Nous avons alors $G' = g$. Mais comme $G = -F_-$ et $g = -f^-$, ça donne $F_-' = f^-$. En mettant tout ensemble,

$$F' = F_+' + F_-' = f^+ + f^- = f. \tag{14.1043}$$

□
PropJLnPpaw

Proposition 14.328.
Si $f \in L^1(\mathbb{R})$, alors la fonction

$$F(x) = \int_{-\infty}^x f(t)dt \tag{14.1044}$$

est presque partout dérivable et pour les points où elle l'est, nous avons $F'(x) = f(x)$.

14.17.4 Permuter limite et dérivée

NORMALooGYUEooKrYjyz

14.329 ([326]).
Voici une preuve alternative du théorème 12.374. Elle utilise des intégrales ; elle demande donc plus de dépendances.

115. En posant par exemple $F' = f = 0$ sur la partie où cette inégalité n'est pas vraie. De toutes façon cette partie est de mesure nulle.
116. Qui s'applique parce que F est presque partout croissante du fait que f est supposée positive.
117. Encore moins « presque partout » qu'avant parce que c'est l'intersection entre le « presque partout » d'avant et de celui-ci. Mais bon...Tant qu'on a une quantité dénombrable d'intersections, ça reste de mesure nulle. Alors une de plus ou une de moins...

Énoncé Soient une suite de fonctions $f_i\colon \mathbb{R} \to \mathbb{R}$, une fonction $f\colon \mathbb{R} \to \mathbb{R}$ et une fonction $g\colon \mathbb{R} \to \mathbb{R}$ telles que

(1) f_i est de classe C^1 pour tout i,

(2) $f_i \to f$ simplement,

(3) $f_i' \to g$ uniformément sur tout compact.

Alors

(1) f est de classe C^1,

(2) $f' = g$,

(3) $f_i \to f$ uniformément sur tout compact.

Preuve Nous commençons par considérer $x_0 \in \mathbb{R}$ et un intervalle compact K contenant x_0. Nous montrons que $f'(x_0) = g(x_0)$ en plusieurs étapes.

(1) **Une formule intégrale** Par hypothèse, les fonctions f_i sont continues (en particulier sur un ouvert contenant K), et le théorème fondamental de l'analyse 14.287 donne

$$f_i(x) = f_i(x_0) + \int_{x_0}^{x} f_i'(t)dt \tag{14.1045}$$

pour tout $x \in K$. Nous avons envie de prendre la limite $i \to \infty$ en permutant la limite avec l'intégrale. Pour cela nous allons utiliser la convergence dominée de Lebesgue.

(2) **Convergence dominée** La convergence uniforme sur tout compact des fonctions continues f_i' vers g donne la continuité de g, théorème 12.357. En particulier g est bornée et donc intégrable sur le compact $[x_0, x]$. Mais il en faut plus pour le théorème de la convergence dominée de Lebesgue (théorème 14.207). Soit $a > 0$; il existe N tel que pour tout $i > n$ nous ayons $\|f_i' - g\| < a$. Avec cela nous avons

$$|f_i'(x)| < |g(x)| + a \tag{14.1046}$$

pour tout $x \in K$. En particulier, la fonction $x \mapsto g(x) + a$ fonctionne pour la convergence dominée et nous pouvons permuter la limite et l'intégrale dans (14.1045).

(3) **Passage à la limite**
En passant à la limite $i \to \infty$ dans (14.1045) nous trouvons

$$f(x) = f(x_0) + \int_{x_0}^{x} g(t)dt. \tag{14.1047}$$

(4) **Premières conclusions**
Il suffit maintenant de prendre la dérivée de (14.1047) au point $x = x_0$ grâce à la proposition 14.286 :

$$f'(x_0) = g(x_0). \tag{14.1048}$$

Cela nous donne l'égalité $f' = g$ parce que x_0 était arbitraire.

De plus g est continue comme limite uniforme des fonctions continues f_i'. Plus précisément, pour voir la continuité de g en x_0, prendre un ouvert borné $B(x_0, r)$ autour de x_0, et ensuite un compact K contenant cet ouvert. La convergence uniforme $f_i' \to g$ sur K implique la convergence uniforme sur $B(x_0, r)$ et donc la continuité sur $B(x_0, r)$ (théorème 12.357).

(5) $f_i \to f$ **uniforme sur tout compact**
Un compact n'étant pas spécialement connexe, nous ne pouvons pas reprendre le travail fait jusqu'ici sans prendre une petite précaution. Soit un compact L. Cette partie de $\mathbb{R}$ étant bornée[118], nous pouvons prendre r assez grand pour que $L \subset \overline{B(0, r)}$. Nous posons

118. Par le théorème de Borel-Lebesgue 9.24

$K = \overline{B(0,r)}$ et nous prouvons la convergence uniforme $f_i \to f$ sur K. A fortiori, cela donnera la convergence uniforme sur L.

Prenons la différence entre (14.1047) et (14.1045) :

$$|f(x) - f_i(x)| = \Big|f(x_0) - f_i(x_0) + \int_{x_0}^{x} g(t) - f_i'(t)dt\Big| \tag{14.1049a}$$

$$\leqslant |f(x_0) - f_i(x_0)| + \Big|\int_{x_0}^{x} |g(t) - f_i'(t)|dt\Big| \tag{14.1049b}$$

$$\leqslant |(f - f_i)(x_0)| + |x - x_0|\|g - f_i'\|_K. \tag{14.1049c}$$

SUBEQooIWSJoqGckNmj

Notez les valeurs absolues autour de l'intégrale dans (14.1049b). Elles sont nécessaires parce que x est dans un voisinage de x_0, sans que nous sachions si $x \geqslant x_0$ ou $x \leqslant x_0$ (ça change le signe de l'intégrale).

Nous avons donc

$$\|f - f_i\| \leqslant |(f - f_i)(x_0)| + \operatorname{diam}(K)\|g - f_i'\| \tag{14.1050}$$

où $\operatorname{diam}(K)$ est le diamètre de K, c'est-à-dire la plus grande distance entre deux éléments de K c'est un nombre fini parce que K est borné. Il majore évidemment $|x - x_0|$. Le membre de droite tend vers zéro si $i \to \infty$ parce que nous avons convergence simple $f_i \to f$ et donc $(f - f_i)(x_0) \to 0$, et parce que nous avons convergence uniforme sur tout compact, donc $\|g - f_i'\| \to 0$.

Nous avons donc bien $\lim_{i\to\infty} \|f - f_i\| = 0$, c'est-à-dire convergence uniforme de (f_i) vers f sur K.

Chapitre 15

Suites et séries de fonctions

Les généralités sur les suites et séries de fonctions, c'est dans la section 12.32.

15.1 Séries de fonctions

15.1.1 Intégration de séries de fonctions

ThoCciOlZ

Théorème 15.1.
La somme uniforme de fonctions intégrables sur un ensemble de mesure fine est intégrable et on peut permuter la somme et l'intégrale.

En d'autres termes, supposons que $\sum_{n=0}^{\infty} f_n$ converge uniformément vers F sur A avec $\mu(A) < \infty$. Si F et f_n sont des fonctions intégrables sur A alors

$$\int_A F(x)d\mu(x) = \sum_{n=0}^{\infty} \int_A f_n(x)d\mu(x). \tag{15.1}$$

Démonstration. Ce théorème est une conséquence du théorème 14.205. En effet nous définissons la suite des sommes partielles

$$F_N = \sum_{n=0}^{N} f_n. \tag{15.2}$$

La limite $\lim_{N\to\infty} F_N = F$ est uniforme. Par conséquent la fonction F est intégrable et

$$\int_A F = \lim_{N\to\infty} \int_A F_N = \lim_{N\to\infty} \int_A \sum_{n=0}^{N} f_n = \lim_{N\to\infty} \sum_{n=0}^{N} \int_A f_n = \sum_{n=0}^{\infty} \int_A f_n. \tag{15.3}$$

La première égalité est le théorème 14.205, les autres sont de simples manipulations rhétoriques. □

Le théorème suivant est une paraphrase du théorème de la convergence dominée de Lebesgue (14.207).

ThoockMHn

Théorème 15.2.
Soient des fonctions $(f_n)_{n\in\mathbb{N}}$ telles que $\sum_{n=0}^{N} f_n$ soit intégrable sur $(\Omega, \mathcal{A}, \mu)$ pour chaque N. Nous supposons que la somme converge simplement vers

$$f(x) = \sum_{n=0}^{\infty} f_n(x) \tag{15.4}$$

et qu'il existe une fonction g telle que

$$\left| \sum_{n=0}^{N} f_n \right| < g \tag{15.5}$$

pour tout $N \in \mathbb{N}$. Alors

(1) $\sum_{n=0}^{\infty} f_n$ *est intégrable,*

(2) *on peut permuter somme et intégrale :*

$$\lim_{N\to\infty} \int_\Omega \sum_{n=0}^{N} f_n d\mu = \int_\Omega \sum_{n=0}^{\infty} f_n, \tag{15.6}$$

(3)

$$\lim_{N\to\infty} \int_\Omega \left| \sum_{n=0}^{N} f_n - \sum_{n=0}^{\infty} f_n \right| = \lim_{N\to\infty} \int_\Omega \left| \sum_{n=N}^{\infty} f_n \right| = 0. \tag{15.7}$$

ThoCSGaPY

Théorème 15.3.
Soit f_n des fonctions $C^1[a,b]$ telles que

(1) *la série $\sum_n f_n(x_0)$ converge pour un certain $x_0 \in [a,b]$,*

(2) *la série des dérivées $\sum_n f_n'$ converge uniformément sur $[a,b]$.*

Alors la série $\sum_n f_n$ converge vers une fonction F et

(1) *La convergence est uniforme sur $[a,b]$.*

(2) *La fonction F est dérivable*

(3) *$F'(x) = \sum_n f_n'(x)$.*

15.1.2 Différentiabilité

LEMooPIWYooZEofkW

Lemme 15.4.
Soient E et F deux espaces vectoriels normés. Si la suite (T_n) converge vers T dans $\mathcal{L}(E,F)$, alors pour tout $v \in E$ nous avons

$$\left(\sum_{n=0}^{\infty} T_n \right)(v) = \sum_{n=0}^{\infty} T_n(v). \tag{15.8}$$

LEMooEYARooNExmiw

Lemme 15.5 ([475])**.**
Soit un espace métrique V. Soient un compact K et un ouvert U contenant K. Alors il existe $\delta > 0$ tel que

$$\{x \in V \text{ tel que } d(x,K) < \delta\} \subset U. \tag{15.9}$$

Démonstration. Pour chaque $x \in K$, nous posons

$$A_x = \{\alpha \geqslant 0 \text{ tel que } B(x,\alpha) \subset U\}. \tag{15.10}$$

Le théorème 7.8 indique que A_x contient toujours des éléments strictement positifs (parce que K est dans U qui est ouvert). Nous considérons la fonction

$$\begin{aligned} f\colon K &\to \mathbb{R} \\ x &\mapsto \sup\{\alpha \geqslant 0 \text{ tel que } B(x,\alpha) \subset U\}. \end{aligned} \tag{15.11}$$

Nous prouvons que f est continue. Soit $\epsilon > 0$. Soit $y \in B(x,\epsilon)$.

(1) $\underline{f(x) - \epsilon \leqslant f(y)}$ Nous prouvons que pour tout $\alpha \in A_x$, le nombre $\alpha - \epsilon$ est dans A_y, c'est-à-dire que $B(y, \alpha-\epsilon) \subset U$. Soit $a \in B(y,\alpha-\epsilon)$. Nous avons

$$d(x,a) \leqslant d(x,y) + d(y,a) \leqslant \epsilon + (\alpha - \epsilon) = \alpha. \tag{15.12}$$

Donc $a \in U$.

Cela signifie que pour tout $\alpha \in A_x$, nous avons $\alpha - \epsilon \in A_y$, ou encore que $A_x - \epsilon \subset A_y$. En passant au supremum, $\sup(A_x) - \epsilon \leqslant \sup(A_y)$. Nous avons donc bien $f(x) - \epsilon \leqslant f(y)$.

(2) $f(y) \leqslant \epsilon + f(x)$ Soit $\epsilon' > 0$. Il existe un élément u hors de U tel que $d(x,y) < f(x) + \epsilon'$. Cet élément vérifie donc

$$f(y) \leqslant d(y,u) \leqslant d(y,x) + d(x,u) \leqslant \epsilon + f(x) + \epsilon'. \tag{15.13}$$

Pour tout $\epsilon' > 0$ nous avons $f(y) \leqslant \epsilon + f(x) + \epsilon'$. En prenant la limite $\epsilon' \to 0$ nous trouvons $f(y) \leqslant f(x) + \epsilon$.

(3) **f est continue** Nous avons montré que si $y \in B(x,\epsilon)$, alors $|f(x) - f(y)| \leqslant \epsilon$. Cela signifie que f est continue en $x \in K$.

(4) **Conclusion** La fonction f est continue sur le compact K. Elle est donc bornée et atteint ses bornes [1]. Soit $\delta = \min\{f(x) \text{ tel que } x \in K\}$. Comme dit plus haut, $f(x) > 0$ pour tout x. Donc $\delta > 0$.

Si $x \in V$ vérifie $d(x,K) < \delta$, alors il existe $y \in K$ tel que $d(x,y) < \delta$. Par minimalité de δ, nous avons $f(y) \geqslant \delta$ et donc $B(y,\delta) \subset U$. En particulier $x \in B(y,\delta) \subset U$.

□

PROPooGRXXooWcZyJG

Proposition 15.6 ([475]).
Tout ouvert connexe par arcs d'un espace vectoriel normé est connexe par arcs de classe C^1.

Démonstration. Soient un espace vectoriel normé V, et un ouvert connexe par arcs U. Nous considérons $x, y \in U$ et un chemin continu $\gamma\colon [0,1] \to U$ joignant x à y.

(1) **Utilisation du lemme** La partie $\gamma\big([0,1]\big)$ est compacte parce qu'elle est l'image par une application continue d'un compact (théorème 7.220). Par le lemme 15.5, il existe $\delta > 0$ tel que pour tout $t \in [0,1]$, $B\big(\gamma(t),\delta\big) \subset U$.

(2) **Uniforme continuité** L'application γ est continue sur le compact $[0,1]$. Le théorème de Heine 12.78 nous dit que γ est alors uniformément continue : il existe $\eta > 0$ tel que pour tout $t, t' \in [0,1]$ tels que $|t - t'| < \eta$, nous ayons $d\big(\gamma(t),\gamma(t')\big) < \delta$.

(3) **La courbe** Nous considérons une subdivision $0 = t_0 < t_1 < \ldots < t_n = 1$ telle que $|t_i - t_{i-1}| < \eta$ [2]. Nous notons $P_i = \gamma(t_i)$ et $Q_i = \gamma(t_{i+1})$. Soient les applications [3]

$$\begin{aligned} \varphi_i\colon \mathbb{R} &\to V \\ t &\mapsto t^3 Q_i + 3t^2(1-t)Q_i + 3t(1-t)^2 P_i + (1-t)^3 P_i. \end{aligned} \tag{15.14}$$

Nous considérons ensuite les reparamétrisations des φ_i de $[0,1]$ vers $[t_i, t_{i+1}]$:

$$\begin{aligned} \sigma_i\colon \mathbb{R} &\to V \\ t &\mapsto \varphi_i\left(\frac{t - t_i}{t_{i+1} - t_i}\right). \end{aligned} \tag{15.15}$$

Ces applications sont de classe C^1. Enfin nous posons

$$\begin{aligned} \psi\colon [0,1] &\to V \\ t &\mapsto \sigma_i(t) \quad \text{si } t \in [t_i, t_{i+1}]. \end{aligned} \tag{15.16}$$

EQooIQHHooGspmQJ

(4) **Ce qu'il faut voir** Petite pause pour lister les choses que nous devons vérifier.

— $\psi(0) = \gamma(0)$, $\psi(1) = \gamma(1)$

— ψ est continue.

— ψ est à valeurs dans U.

— ψ est de classe C^1.

1. Théorème de Weierstrass 7.146.
2. Êtes-vous capable d'en prouver l'existence ? Prenez un entier $z > \eta$ et $t_k = k/z$. Notez l'utilisation du lemme 1.438 que je me lasserai jamais de citer.
3. Ce sont des courbes de Bézier d'ordre 3 de points de contrôle P_i, P_i, Q_i et Q_i.

On y va...

(5) $\psi(0)$ **et** $\psi(1)$ Nous avons

$$\psi(0) = \sigma_0(0) = \varphi_0(0) = P_0 = \gamma(0), \tag{15.17}$$

et, utilisant le fait que $t_n = 1$,

$$\psi(1) = \sigma_{n-1}(1) = \varphi_{n-1}\left(\frac{1 - t_{n-1}}{t_n - t_{n-1}}\right) = \varphi_{n-1}(1) = Q_{n-1} = \gamma(t_n) = \gamma(1). \tag{15.18}$$

(6) ψ **est continue** Nous devons vérifier la continuité aux raccords, c'est-à-dire que nous devons prouver $\sigma_i(t_{i+1}) = \sigma_{i+1}(t_i)$. D'une part $\sigma_i(t_{i+1}) = \phi_i(1) = Q_i = \gamma(t_{i+1})$, et d'autre part $\sigma_{i+1}(t_i) = \varphi_{i+1}(0) = P_{i+1} = \gamma(t_{i+1})$. C'est bon.

(7) **À valeurs dans** U Nous devons vérifier que $\varphi_i\big([0,1]\big) \subset U$. Nous avons bien sûr $P_i \in B(P_i, \delta)$, mais nous avons aussi $Q_i \in B(P_i, \delta)$. En effet, $|t_{i+1} - t_i| < \eta$ de telle sorte que $Q_i = \gamma(t_{i+1}) \in B\big(\gamma(t_i), \delta\big) = B(P_i, \delta)$. Le nombre η a été choisi pour ça.
Le nombre δ a été choisi pour avoir également $B(P_i, \delta) \subset U$. Bref, P_i et Q_i sont dans $B(P_i, \delta)$ qui est convexe[4] et contenue dans U. Pour $t \in [t_i, t_{i+1}[$, le point $\varphi_i(t)$ est une combinaison convexe de P_i et Q_i :

$$t^3 + 3t^2(1-t) + 3t(1-t)^2 + (1-t)^3 = 1. \tag{15.19}$$

Donc $\varphi_i(t) \in B(P_i, \delta) \subset U$.

(8) ψ **est de classe** C^1 Nous avons $\varphi_i'(t) = 6t(1-t)(Q_i - P_i)$. Donc pour tout i nous avons $\varphi_i'(1) = \varphi'(0) = 0$. En ce qui concerne ψ' nous avons

$$\lim_{t \to t_i^+} \psi'(t) = \psi'(t_i) = 0 \tag{15.20}$$

et

$$\lim_{t \to t_i^-} \psi'(t) = \lim_{t \to 1} \varphi_{i-1}(t) = 0, \tag{15.21}$$

d'où la continuité de ψ'.
Nous en déduisons que ψ est de classe C^1 par le théorème 12.330.

□

Remarque 15.7.
Quelle est la forme de l'image de la courbe C^1 définie en (15.16) ? L'image est une ligne brisée : une suite de segments de droites. Elle fait un angle à chaque raccord. Il est contre-intuitif qu'une ligne faisant des angles puisse être de classe C^1, mais c'est pourtant le cas. Le fait est que ψ' s'annule à chacun de ces raccords, de telle sorte que ψ' est quand même continu.

¡¡ Avertissement/question à la lectrice !! 15.8
Le théorème 15.9 se démontre ici avec des intégrales. Je suis presque certain qu'on doit pouvoir adapter la démonstration du théorème 12.377 pour ne pas avoir à utiliser d'intégrales.

ThoLDpRmXQ

Théorème 15.9 ([476]).
Soient E et F deux espaces vectoriels normés, Ω un ouvert connexe par arcs. Soit (u_n) une suite de fonctions $u_n : \Omega \to F$ telle que

(1) pour tout n, la fonction u_n est de classe C^1 sur Ω,

(2) la série $\sum_n u_n$ converge simplement sur Ω,

(3) la série des différentielles $\sum_n (du_n)$ converge normalement sur tout compact de Ω.

4. Une boule est convexe, proposition 11.28.

Alors la somme $u = \sum_n u_n$ est de classe C^1 sur Ω et sa différentielle est donnée par

$$du = \sum_{n=0}^{\infty} du_n. \tag{15.22}$$

Démonstration. Pour chaque n, la fonction $du_n\colon \Omega \to \mathcal{L}(E,F)$ est une fonction continue parce que u_n est de classe C^1. La série convergeant normalement, la fonction $\sum_{n=0}^{\infty} du_n$ est également continue par la proposition 12.369. La difficulté de ce théorème est donc de prouver que cela est bien la différentielle de la fonction $\sum_n u_n$, c'est-à-dire que

$$d\left(\sum_{n=0}^{\infty} u_n\right) = \sum_{n=0}^{\infty} du_n. \tag{15.23}$$

Soient $a, x \in \Omega$. Nous considérerions bien le segment $[a,x]$, mais vu que Ω n'est supposé que connexe par arcs de classe C^1 (la proposition 15.6 pour que l'arc puisse être choisi C^1), nous ne pouvons pas faire mieux pour joindre a à x que choisir un chemin de classe C^1

$$\gamma\colon [0,1] \to \Omega \tag{15.24}$$

tel que $\gamma(0) = a$ et $\gamma(1) = x$.

L'astuce est de poser

$$\begin{aligned} f_n\colon [0,1] &\to \mathbb{R} \\ t &\mapsto (du_n)_{\gamma(t)}\big(\gamma'(t)\big), \end{aligned} \tag{15.25}$$

et d'en étudier l'intégrale[5].

(1) **Permuter somme et intégrale** Nous voudrions permuter la somme et l'intégrale dans l'expression $\int_0^1 \sum_i f_i(t)dt$. Pour cela nous commençons par regarder quelques majorations de normes.

D'abord γ est de classe C^1, ce qui fait que γ' est continue. Vu que la norme est une application continue, la fonction $t \mapsto \|\gamma'(t)\|$ est également continue sur le compact $[0,1]$. Elle est donc majorée par une constante que nous nommons M. C'est le théorème de Weierstrass 7.146.

Ensuite nous avons le calcul

$$\|f_i(t)\| = \|(du_i)_{\gamma(t)}\big(\gamma'(t)\big)\| \leqslant \|(du_i)_{\gamma(t)}\|\|\gamma'(t)\| \leqslant M\|du_i\|_\infty < \infty. \tag{15.26}$$

Justifications :

— Pour la première inégalité. C'est le lemme 10.66.

— Pour la seconde inégalité. Il s'agit de l'inégalité évidente

$$\|du_i\|_\infty = \sup_{x \in \gamma([0,1])} \|(du_i)_x\| \tag{15.27}$$

Notons que la norme $\|.\|_\infty$ ne réfère pas à un supremum sur E, mais seulement sur l'image de γ. Nous aurions pu faire preuve d'un peu de créativité dans les notations.

— L'application du_i est continue sur le compact $\gamma\big([0,1]\big)$. Donc le supremum est fini et atteint.

Maintenant nous posons

$$g_n(t) = \sum_{i=0}^{n} f_i(t). \tag{15.28}$$

Nous avons la majoration

$$\|g_n(t)\| \leqslant \sum_{i=0}^{n} \|f_i(t)\| \leqslant M \sum_{i=0}^{n} \|du_i\|_\infty < \infty. \tag{15.29}$$

5. Cela revient à étudier l'intégrale de la forme différentielle du_n sur le chemin γ. Voir la définition 20.47 et tout ce qui s'en suit.

Le fait que le tout soit fini est l'hypothèse de convergence normale sur tout compact. Le compact en question est $\gamma\big([0,1]\big)$.

C'est le moment d'utiliser le théorème de la convergence dominée de Lebesgue 14.207. Attention aux notations un peu décalées. Nous avons $g_n \to \sum_{i=1}^{\infty} f_i$ (convergence simple) et $\|g_n(t)\| \leqslant A$ où A est une constante que nous voyons comme une fonction constante intégrable sur le compact $[0,1]$. Nous permutons la limite et l'intégrale :

$$\begin{aligned}
\int_0^1 \sum_{i=0}^{\infty} f_i(t)dt &= \int_0^1 (\lim_{n\to\infty} g_n)(t)dt && (15.30a)\\
&= \lim_{n\to\infty} \int_0^1 g_n(t)dt && (15.30b)\\
&= \lim_{n\to\infty} \int_0^1 \sum_{i=0}^{n} f_i(t)dt && (15.30c)\\
&= \lim_{n\to\infty} \sum_{i=0}^{n} \int_0^1 f_i(t)dt && (15.30d)\\
&= \sum_{i=0}^{\infty} \int_0^1 f_i(t)dt. && (15.30e)
\end{aligned}$$

(2) **Accroissements**

Nous pouvons maintenant faire le petit calcul suivant :

$$\sum_n \int_0^1 (du_n)_{\gamma(t)}\big(\gamma'(t)\big)dt = \sum_n \Big(u_n\big(\gamma(1)\big) - u_n\big(\gamma(0)\big)\Big) = \sum_n \big(u_n(x) - u_n(a)\big) = u(x) - u(a) \tag{15.31}$$

où nous avons utilisé le théorème fondamental du calcul intégral sous la forme de la proposition 14.293.

Nous retenons l'égalité

$$u(x) = u(a) + \int_0^1 \sum_n (du_n)_{\gamma(t)}\big(\gamma'(t)\big)dt. \tag{15.32}$$

EQooTXYWooIDxVri

(3) **Remarque** La formule (15.32) n'est pas une forme de formule des accroissements finis qui parlerait d'évaluer une fonction u en x en partant de a et en intégrant du le long d'un chemin joignant a et x.

Ce serait le cas si nous pouvions permuter la somme et la différentielle qui se trouvent dans l'intégrale. Or permuter somme et différentielle est précisément l'objet du théorème que nous sommes en train de prouver.

(4) **Différentielle**

Forts de la formule (15.32), nous calculons $du_a(v)$, c'est-à-dire la différentielle de u au point a appliquée au vecteur $v \in E$. Pour cela, nous savons que Ω est ouvert, donc Ω contient une boule de rayon r autour de a, ce qui nous permet de dire que pour un a donné, le point $a+sv$ est dans Ω pour tout $s \in B(0,\epsilon)$ lorsque ϵ n'est pas trop grand. Pour chacun de ces s, nous considérons un chemin de classe C^1 joignant a à $a+sv$. Ce chemin sera noté

$$\gamma_s : [0,1] \to \Omega \tag{15.33}$$

et $\gamma(0) = a$, $\gamma(1) = a + sv$. Nous avons le calcul

$$\begin{aligned}
du_a(v) &= \frac{d}{ds}\Big[u(a+sv)\Big]_{s=0} && \text{(15.34a)}\\
&= \frac{d}{ds}\Big[\int_0^1 \sum_{n=0}^{\infty} (du_n)_{\gamma_s(t)}\big(\gamma_s'(t)\big)dt\Big]_{s=0} && \text{(15.34b)}\\
&= \frac{d}{ds}\Big[\sum_{n=0}^{\infty}\int_0^1 (du_n)_{\gamma_s(t)}\big(\gamma_s'(t)\big)dt\Big]_{s=0} && \text{(15.34c)}\\
&= \frac{d}{ds}\Big[\sum_n \big(u_n(a+sv) - u_n(a)\big)\Big]_{s=0} && \text{(15.34d)}\\
&= \frac{d}{ds}\Big[(\sum_n u_n)(a+sv)\Big]_{s=0} && \text{(15.34e)}\\
&= \sum_n \frac{d}{ds}\Big[u_n(a+sv)\Big]_{s=0} && \text{(15.34f)}\\
&= \sum_n (du_n)_a(v) && \text{(15.34g)}\\
&= \Big(\sum_n (du_n)_a\Big)(v) && \text{(15.34h)}\\
&= \Big(\sum_n du_n\Big)_a(v). && \text{(15.34i)}
\end{aligned}$$

Justifications :

— Pour 15.34c. Permuter la somme et l'intégrale comme plus haut.

— Pour 15.34f. Permuter une somme et une dérivée classique des fonctions $\mathbb{R} \to F$ données par $s \mapsto u_n(a+sv)$. Il s'agit d'utiliser le théorème 12.374 sur chaque composantes dans F.

— Pour 15.34i. Chaque du_n est une application $du_n \colon E \to \mathcal{L}(E,F)$. Au fait près que la notation est plus lourde, il s'agit simplement d'une définition de la somme ponctuelle d'une suite de fonctions : $\sum_n f_n(a) = (\sum_n f_n)(a)$. Dans ce cas-ci, le tout est encore un élément de $\mathcal{L}(E,F)$ que nous appliquons à v.

□

Lemme 15.10.
Sur $[\frac{1}{2}, 1]$, étudier la convergence de la série

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{x-1}{x}\right)^n. \tag{15.35}$$

Étudier la convergence de la série dérivée ; en déduire que

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln(2). \tag{15.36}$$

Démonstration. Nous posons $y(x) = \frac{x-1}{x}$, et nous regardons la série

$$\tilde{f}(y) = \sum_{n=1}^{\infty} \frac{1}{n} y^n. \tag{15.37}$$

Cette série de puissance converge absolument pour $|y| < 1$, voir l'exemple 4 de la page 123bis du cours de première. Cette série converge également simplement en $y = -1$, par le corolaire de la

page 123 du même cours[6]. Nous sommes dans le cas d'une série de puissance dont le disque de convergence est centré en 0, et dont le rayon est 1, mais qui converge (en plus) simplement sur un des bords du disque. Cela est le cadre du théorème 12.368 qui nous permet de dire que pour tout $\epsilon > 0$, la série (15.37) converge uniformément sur $[-1, 1-\epsilon]$.

La fonction $\tilde{f}(y)$ est donc continue sur $[-1, 1-\epsilon]$, et donc en particulier sur $[-1, 0]$. Par ailleurs, la fonction $y(x)$ est continue en $x \neq 0$. En tant que composée de fonctions continues, la fonction $f(x) = \tilde{f}\big(y(x)\big)$ est continue sur $[\frac{1}{2}, 1]$.

Nous la mettons la série des dérivées sous la forme d'une série de puissances :

$$g(x) = \frac{1}{x^2} \sum_{n=1}^{\infty} \left(\frac{x-1}{x} \right)^{n-1}. \tag{15.38}$$ EqSerieDerrTreize

Afin d'éviter tout malentendu, nous insistons sur le fait que g est la série des dérivée de la série f. Nous ne savons pas encore si g existe (c'est-à-dire si elle converge), ni si sa somme est la dérivée de f. C'est cela que nous allons tenter d'établir maintenant.

Nous posons à nouveau $y(x) = \frac{x-1}{x}$, et nous savons que la série de puissances $\sum_n y^n$ converge uniformément pour $y \in [-1+\epsilon, 1-\epsilon]$ pour tout $\epsilon > 0$. En repassant aux variables x, pour tout $\epsilon > 0$, nous avons convergence uniforme de la série

$$\sum_{n=0}^{\infty} \left(\frac{x-1}{x} \right)^{n}, \tag{15.39}$$

sur le compact $x \in [\frac{1}{2-\epsilon}, \frac{1}{\epsilon}]$, ou, pour parler plus simplement, sur $x \in [\frac{1}{2} + \epsilon, a]$ pour tout ϵ (petit) et a (grand). Nous avons donc également convergence uniforme de la série des dérivées (15.38) sur le même intervalle. Maintenant, le théorème 15.9 montre que la série des dérivée est bien la dérivée de la série, c'est-à-dire que

$$g(x) = f'(x) \tag{15.40}$$

sur $]\frac{1}{2}, 1]$. Notez que la convergence uniforme *sur tout compact* de la série des dérivées est suffisante.

Une bonne nouvelle est qu'il est possible de sommer explicitement la série $\sum_k y^k$. En effet, il est montré à la page 115 du cours de première que $\sum_{k=0}^{n} z^k = \frac{1-z^{n+1}}{1-z}$, donc

$$\sum_{k=0}^{\infty} y^n = \lim_{n\to\infty} \frac{1-y^{n+1}}{1-y} = \frac{1}{1-y}, \tag{15.41}$$ EqFormSomGeometrze

lorsque $|y| < 1$. Du coup, nous avons simplement

$$f'(x) = g(x) = \frac{1}{x^2} \sum_{n=1}^{\infty} \left(\frac{x-1}{x} \right)^{n-1} = \frac{1}{x^2} \sum_{n=0}^{\infty} \left(\frac{x-1}{x} \right)^{n} = \frac{1}{x^2} \left(\frac{1}{1-\left(\frac{x-1}{x}\right)} \right) = \frac{1}{x}, \tag{15.42}$$

donc la fonction f a la forme simple $f(x) = \ln(x) + C$. Notez bien le petit jeu de variables de sommation. Au départ $g(x)$ est une somme qui part de 1 avec un exposant $n-1$, et nous la transformons en une somme qui part de 0 avec un exposant n. C'est cela qui nous permet d'appliquer la formule (15.41).

Étant donné que $f(1) = 0$, nous avons

$$f(x) = \ln(x) \tag{15.43}$$

pour tout $x \in [\frac{1}{2} + \epsilon, 1]$. Mais nous avons vu que la fonction f était continue sur $[\frac{1}{2}, 1]$. Étant donné que $\ln(x)$ et $f(x)$ sont deux fonctions continues sur $[\frac{1}{2}, 1]$ qui sont égales sur tout compact $[\frac{1}{2} + \epsilon, 1]$, nous déduisons que ces deux fonctions sont en réalité égales sur tout l'entièreté du compact $[\frac{1}{2}, 1]$.

En particulier, en $x = \frac{1}{2}$, nous avons

$$f(\frac{1}{2}) = \sum_{n=1}^{\infty} \frac{1}{n}(-1)^n = \ln(1/2) = -\ln(2). \tag{15.44}$$

□

6. Un étudiant avait dit se souvenir qu'Abel s'appliquait seulement aux séries alternées ; c'est ce corolaire (critère des séries alternées) qui l'a induit en erreur. En effet, Abel (proposition 5, page 122) est plus général, mais s'applique particulièrement bien aux séries alternées.

15.2 Séries entières

Dans cette section nous allons parler de séries complexes autant que de séries réelles. L'étude des propriétés à proprement parler complexes des séries entières (holomorphie) sera effectuée dans le chapitre dédié, voir le théorème 26.60 et ses conséquences.

15.2.1 Disque de convergence

Une **série de puissance** est une série de la forme

$$\sum_{k=0}^{\infty} c_k(z-z_0)^k \tag{15.45}$$ eqseriepuissance

où $z_0 \in \mathbb{C}$ est fixé, (c_k) est une suite complexe fixée, et z est un paramètre complexe. Nous disons que cette série est *centrée* en z_0.

Définition 15.11.
*Une **série entière** est une somme de la forme*

$$\sum_{n=0}^{\infty} a_n z^n \tag{15.46}$$

avec $a_n, z \in \mathbb{C}$.

Une série entière peut définir une fonction

$$f(z) = \sum_n a_n z^n. \tag{15.47}$$

Le but de cette section est d'étudier des conditions sur la suite (a_n) qui assurent la continuité de f ou la possibilité de dériver ou intégrer la série terme à terme.

DefZWKOZOl

Définition 15.12.
*Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière. Le **rayon de convergence** de cette série est le nombre*

$$R = \sup\{r \in \mathbb{R}^+ \text{ tel que la suite } (a_n r^n) \text{ est bornée}\} \in [0,\infty]. \tag{15.48}$$

*La boule $B(0,R)$ est le **disque de convergence** de la série.*

Dans le cas d'une série de la forme $\sum_n a_n(z-z_0)^n$, le disque de convergence est l'ensemble $|z-z_0| < R$.

15.13.
Notez que le disque de convergence proprement dit est ouvert. Donc, pour être correct, on devrait parler de la *frontière* pour parler de la partie $|z| = R$. Toutefois, nous désignerons souvent cette partie en parlant du *bord* du disque.

15.14.
En réalité, il serait plus correct de parler du rayon de convergence de la suite (a_n) parce qu'au moment où on l'étudie, nous ne savons pas encore si la somme existera. Il ne devrait donc pas être autorisé d'écrire « étudions le rayon de convergence de $\sum_n a_n z^n$ ».

Le rayon de convergence d'une série ne dépend que des réels $|a_n|$, même si à la base $a_n \in \mathbb{C}$.

15.15.
Sur Wikipédia[477], le rayon de convergence est défini par le supremum des $|z|$ tels que la série $\sum_n a_n z^n$ converge. Je vous invite à vous étonner que cela est équivalent à la définition donnée ici.

Il est dingue que demander que la suite $(a_n r^n)$ soit bornée soit suffisant pour que la série converge. En réalité ce n'est pas tout à fait le cas ; les séries qui convergent sont celles pour $|z|$

strictement plus petit que le rayon de convergence. Et là ça marche. En effet, si $x < R$, il existe un r tel que $x < r < R$. Nous considérons ϵ donné par $x = \epsilon r$, qui vérifie $\epsilon < 1$ et nous avons

$$a_n x^n = a_n(\epsilon r)^n = (a_n r^n)\epsilon^n \to 0. \tag{15.49}$$

Le critère d'Abel 15.18 va formaliser ça.

REMooYOTEooKvxHSf

Remarque 15.16.
Si pour tout n nous avons $|b_n| \geqslant |a_n|$ alors le rayon de convergence de la série $\sum_n a_n z^n$ est au moins aussi grand que celui de la série $\sum_n b_n z^n$. Cela y compris lorsque l'un ou l'autre des rayons de convergences est infini.

LEMooVCTNooCQHkzs

Lemme 15.17 ([1])**.**
Le rayon de convergence pour la suite $b_n = a_{n+k}$ est le même que celui pour a_n.

Démonstration. Soit $r > 0$. Nous avons $r^k b_n r^n = a_{n+k} r^{n+k}$. Vu que les k premiers termes d'une suite ne changement pas le fait que la suite soit bornée, la suite $(a_n r^n)_{n \geqslant k}$ est bornée si et seulement si la suite $(a_n r^n)_{n \in \mathbb{N}}$ est bornée. □

LemmbWnFI

Lemme 15.18 (Critère d'Abel)**.**
Soit $R > 0$ le rayon de convergence de la somme $\sum_n a_n z^n$ et $z \in \mathbb{C}$.

(1) Si $|z| < R$ alors la série converge absolument.

(2) Si $|z| > R$ alors la série diverge.

Démonstration. Démonstration en deux parties.

(1) Si $|z| < R$ alors la suite $(a_n z^n)$ est bornée et il existe un nombre $M \in \mathbb{R}$ tel que $|a_n| r^n \leqslant M$ pour tout n. Nous considérons alors un r tel que $|z| < r < R$ et nous pouvons calculer :

$$|a_n z^n| = |a_n| r^n \big(\frac{|z|}{r}\big)^n \leqslant M \left(\frac{|z|}{r}\right)^n \tag{15.50}$$

Vu que $|z| < r$ nous tombons sur la série géométrique (10.343) qui converge. Par le critère de comparaison [7] la série $\sum_{n=0}^{\infty} |a_n z^n|$ converge.

(2) Par définition du rayon de convergence, la suite $(a_n z^n)$ n'est donc pas bornée et la série ne peut pas converger à cause de la proposition 10.100.

□

CORooCUDSooTfMvAB

Corolaire 15.19.
Soit une série entière $\sum_n a_n z^n$. Soit un nombre ρ tel que

(1) La série converge pour $|z| < \rho$.

(2) La série diverge pour $|z| > \rho$.

Alors le rayon de convergence est ρ.

Démonstration. Nous notons R le rayon de convergence de la série.

(1) $\underline{R \geqslant \rho}$ Si $R < \rho$, nous prenons r strictement entre R et ρ. Le critère d'Abel 15.18 nous dit, pour $|z| = r$, que la série diverge. Par hypothèse, elle converge ; contradiction.

(2) $\underline{R \leqslant \rho}$ De même si $R > \rho$, alors nous prenons $\rho < r < R$. Le critère d'Abel nous dit que la série diverge pour $|z| = r$. L'hypothèse nous dit le contraire. Nouvelle contradiction.

□

7. Lemme 10.128.

15.20.
Le critère d'Abel parle bien de convergence absolue, et non de convergence normale. Pour chaque t, la série $\sum_k |a_n t^k|$ converge. Si par contre nous posons $u_k(t) = a_k t^k$, nous n'avons à priori pas la convergence normale $\sum_k \|u_k\|_\infty$, même pas si la norme est la norme supremum sur $B(0,R)$ [8]. Prenons comme exemple simplement $a_k = 1$ pour tout k. Pour tout $|t| < 1$, la série $\sum_k t^k$ converge absolument (série géométrique), mais nous aurions $\|u_k\|_\infty = 1$ et donc divergence évidente de $\sum_k \|u_k\|_\infty$.

La proposition suivante sera surtout utile lorsqu'on parlera de dérivée.

PropHDIUooKTbVSX

Proposition 15.21 ([478]).
Quel que soit le nombre $\alpha \in \mathbb{R}$, les séries $\sum_n a_n z^n$ et $\sum_n n^\alpha a_n z^n$ ont même rayon de convergence.

Démonstration. Nous posons

$$E = \{r \in \mathbb{R}^+ \text{ tel que } (a_n r^n) \text{ est borné } \} \tag{15.51a}$$
$$E' = \{r \in \mathbb{R}^+ \text{ tel que } (n^\alpha a_n r^n) \text{ est borné } \} \tag{15.51b}$$

Nous allons prouver que $E = E'$ (et donc leurs supremum seront égaux).

(1) **Si $\alpha > 0$**
Dans ce cas $E' \subset E$ est facile. Nous prouvons l'inclusion dans l'autre sens. Si E est vide, c'est évident. Nous supposons que E n'est pas vide, et nous considérons $r \in E$, et nous montrons que $r \in E'$. Pour cela nous prenons un nombre s tel que $r < s < \sup(E)$. Nous avons

$$n^\alpha a_n r^n = n^\alpha a_n \left(\frac{r}{s}\right)^n s^n = n^\alpha \left(\frac{r}{s}\right)^n a_n s^n. \tag{15.52}$$

Mais $r/s < 1$, donc le lemme 12.411 dit que $n^\alpha (r/s)^n \to 0$. Cela est donc borné par une constante M. Donc

$$n^\alpha a_n r^n \leqslant M a_n s^n. \tag{15.53}$$

Mais la suite $(a_n s^n)$ est bornée. Donc la suite $n^\alpha a_n r^n$ est également bornée, ce qui prouve que $r \in E'$.

(2) **Si $\alpha < 0$** Même raisonnement, mais en inversant les rôles de E et E', et en adaptant un peu [9].

□

Remarque 15.22.
Au fond, cette proposition n'est rien d'autre que dire que dans $n^\alpha r^n$, l'effet « convergent » est r^n qui est une décroissance exponentielle tandis que l'effet « divergent » est n^α qui a une croissance seulement polynomiale.

LEMooNAWTooHWqKBK

Lemme 15.23 ([479]).
Soit une suite (u_n) de réels strictement positifs. Si

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \ell, \tag{15.54}$$

alors la suite $(u_n^{1/n})$ a une limite et elle vaut ℓ également.

Démonstration. Nous commençons par supposer que $\ell > 0$; nous ferons le cas $\ell = 0$ après. Soit $\epsilon < \ell$. L'existence de la limite dans l'hypothèse dit qu'il existe N tel que pour tout $n \geqslant N$ nous ayons

$$\frac{u_{n+1}}{u_n} \in B(\ell, \epsilon). \tag{15.55}$$

EQooDFKNooMQLVrW

8. Il y aurait par contre bien convergence sur tout compact ? Chère lectrice, dites moi ce que vous en pensez
9. Je n'ai pas vérifié ; écrivez-moi si ça pose un problème.

Écrivons un produit télescopique :

$$\frac{u_n}{u_N}=\prod_{p=N}^{n-1}\frac{u_{p+1}}{u_p}. \tag{15.56}$$

Mais chacun des facteurs du produit est soumis à l'encadrement (15.55), donc

$$(\ell-\epsilon)^{n-N}\leqslant\frac{u_n}{u_N}\leqslant(\ell+\epsilon)^{n-N}, \tag{15.57}$$

et donc, en multipliant par $u_N>0$ nous avons

$$u_N(\ell-\epsilon)^{n-N}\leqslant u_n\leqslant u_N(\ell+\epsilon)^{n-M}. \tag{15.58}$$

La fonction $t\mapsto t^{1/n}$ est croissante [10] ; nous pouvons l'appliquer aux inégalités sans changer le sens :

$$u_N^{1/n}(\ell-\epsilon)^{\frac{n-N}{n}}\leqslant u_n^{1/n}\leqslant u_N^{1/n}(\ell+\epsilon)^{\frac{n-N}{n}}. \tag{15.59}$$

Pour rappel, à ce point le N est fixé pour correspondre au $\epsilon<\ell$ que nous avons choisi.

Vu que $1/n\in\mathbb{Q}$ nous invoquons la proposition 12.393 pour dire que $\lim_{n\to\infty}u_N^{1/n}=1$. De plus $(n-N)/n\to 1$ de telle sorte que

$$(\ell\pm\epsilon)^{\frac{n-N}{n}}\to\ell\pm\epsilon. \tag{15.60}$$

Nous considérons donc un N' tel que pour tout $n\geqslant N'$ nous ayons

$$\left|u_N^{1/n}(\ell-\epsilon)^{\frac{n-N}{n}}-(\ell-\epsilon)\right|<\epsilon \tag{15.61a}$$

$$\left|u_N^{1/n}(\ell+\epsilon)^{\frac{n-N}{n}}-(\ell+\epsilon)\right|<\epsilon. \tag{15.61b}$$

Pour tout $n\geqslant\max(N,N')$ nous avons

$$(\ell-\epsilon)-\epsilon<u_N^{1/n}(\ell-\epsilon)^{\frac{n-N}{n}}\leqslant u_n^{1/n}\leqslant u_N^{1/n}(\ell+\epsilon)^{\frac{n-N}{n}}<(\ell+\epsilon)+\epsilon. \tag{15.62}$$

Donc,

$$\ell-2\epsilon\leqslant u_n^{1/n}\leqslant\ell+2\epsilon. \tag{15.63}$$

Cela prouve que la limite $\lim_{n\to\infty}u_n^{1/n}$ existe et vaut ℓ.

Et si $\ell=0$? Dans ce cas, il existe N tel que pour tout $n\geqslant N$ nous avons

$$0\leqslant\frac{u_{n+1}}{u_n}<\epsilon. \tag{15.64}$$

Ensuite il faut recommencer tous les calculs, avec pour seule différence que tous les membres de gauche sont 0. □

LEMooDWNZooXwejrF

Lemme 15.24 (Règle de Cauchy[480, 479])**.**
Soit une suite (x_n) dans un espace vectoriel normé. Nous posons

$$p=\limsup_{n\to\infty}\|x_n\|^{1/n}. \tag{15.65}$$

Alors :

ITEMooZZBIooUYrtYL

(1) Si $p<1$, la série $\sum_{n\in\mathbb{N}}x_n$ est absolument convergente.

ITEMooQGKNooOFeFRd

(2) Si $p>1$, alors la série $\sum_{n\in\mathbb{N}}x_n$ est ne converge pas.

Démonstration. En deux parties.

10. Proposition 12.403, en notant que les arguments sont tous bien positifs.

(1) **Pour (1)** Soit $\epsilon > 0$ tel que $p + \epsilon < 1$. Le lemme 9.46 nous dit que l'ensemble

$$S_\epsilon = \{n \in \mathbb{N} \text{ tel que } \|x_n\|^{1/n} \geqslant p + \epsilon\} \tag{15.66}$$

est fini. Quitte à prendre une queue de suite, nous supposons que S_ϵ est vide, c'est-à-dire que $\|x_n\| < (p+\epsilon)^n$ pour tout n. En posant $q = p + \epsilon$, nous voyons que $\|x_n\| < q^n$ avec $q < 1$. La comparaison avec la série géométrique 10.135(1) conclut.

(2) **Pour (2)** Alors en prenant ϵ tel que $p - \epsilon > 1$, nous avons une infinité de termes vérifiant $\|x_n\| > (p-\epsilon)^n > 1$.

□

ThoSerPuissRap

Théorème 15.25 (Formule de Hadamard[481])**.**
Le rayon de convergence[11] *de la série entière* $\sum_n a_n z^n$ *est donné par une des deux formules*

$$\frac{1}{R} = \limsup \sqrt[k]{|a_k|} \tag{15.67}$$

EqRayCOnySer

ou

$$\frac{1}{R} = \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| \tag{15.68}$$

EqAlphaSerPuissAtern

lorsque a_k *est non nul à partir d'un certain* k*.*

Si une de ces formules donne $1/R = 0$*, alors le rayon de convergence est infini.*

Démonstration. En deux, voire quatre parties. Nous allons utiliser le corolaire 15.19, en commençant par supposer que la limite supérieure n'est ni 0 ni ∞.

(1) **Première formule, si $|z| < R$** Posons $L = \limsup_{n\to\infty} \left(|a_n|^{1/n}\right)$. Si $|z| < \frac{1}{L}$, alors

$$\limsup |a_n z^n|^{1/n} = \limsup |a_n|^{1/n}|z| < L\frac{1}{L} = 1 \tag{15.69}$$

donc la règle de Cauchy du lemme 15.24 conclut à la convergence de la série.

(2) **Première formule, si $|z| > R$** C'est exactement le même calcul, mais l'inégalité arrive dans l'autre sens :

$$\limsup |a_n z^n|^{1/n} = \limsup |a_n|^{1/n}|z| > L\frac{1}{L} = 1 \tag{15.70}$$

donc la règle de Cauchy du lemme 15.24 conclut à la divergence de la série.

(3) **Première formule, si la limite est 0** Si $\limsup \sqrt{|a_k|}^{1/k} = 0$, alors toujours le même calcul donne :

$$\limsup |a_n z^n|^{1/n} = 0 \tag{15.71}$$

et donc la série converge pour tout z. D'où le rayon de convergence infini.

(4) **Première formule, si la limite est ∞** Soit $z \in \mathbb{C}$. La partie

$$S = \{n \in \mathbb{N} \text{ tel que } |a_n|^{1/n} > 1/|z|\} \tag{15.72}$$

est infinie. Pour chaque $n \in S$, nous avons $|a_n z^n| > 1$. La suite des $a_n z^n$ ne convergeant pas vers zéro (il y a même une infinité de termes plus grands que 1), la série ne peut pas converger (proposition 10.100).

(5) **Seconde formule** Le lemme 15.23 nous ramène au cas de la première formule.

□

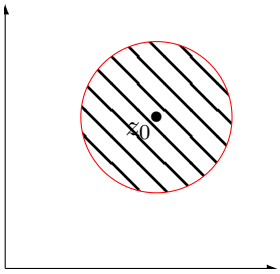


FIGURE 15.1: À l'intérieur du disque de convergence, la convergence est absolue. En dehors, la série diverge. Sur le cercle proprement dit, tout peut arriver. LabelFigDisqueConv

Notons que le critère d'Abel ne dit rien pour les points tels que $|z-z_0|=R$. Il faut traiter ces points au cas par cas. Et le pire, c'est qu'une série donnée peut converger pour certain des points sur le bord du disque, et diverger en d'autres. Le théorème d'Abel radial (théorème 15.40) nous donnera quelques informations sur le sujet.

Il y a un dessin à la figure 15.1.

Si les suites a_n et b_n sont équivalentes, alors les séries correspondantes auront le même rayon de convergence. Cela ne signifie pas que sur le bord du disque de convergence, elles aient même comportement. Par exemple nous avons

$$\frac{1}{\sqrt{n}} \sim \frac{1}{\sqrt{n}} + \frac{(-1)^n}{n}. \tag{15.73}$$

En même temps, en $z=-1$ la série

$$\sum_{n\geqslant 1} \frac{z^n}{\sqrt{n}} \tag{15.74}$$

converge par le critère des séries alternées [12]. Par contre la série

$$\sum_{n\geqslant 1} \left(\frac{1}{\sqrt{n}} + \frac{(-1)^n}{n} \right) z^n \tag{15.75}$$

ne converge pas pour $z=-1$.

Exemple 15.26.
Soit $\alpha\in\mathbb{R}$ et considérons la série $\sum_{n\geqslant 1} a_n z^n$ où a_n est la n-ième décimale de α. Si α est un nombre décimal limité, la suite (a_n) est finie et le rayon de convergence est infini. Sinon, pour tout N il existe un $n>N$ tel que $a_n\neq 0$ et la suite (a_n) ne tend pas vers zéro. Par conséquent la série

$$\sum_n a_n z^n \tag{15.76}$$

diverge pour $z=1$ et le rayon de convergence satisfait $R\leqslant 1$. Nous avons aussi $|a_n|\leqslant 9$, de telle manière à ce que la série soit bornée et par conséquent majorée en module par $9z^n$, ce qui signifie que $R\geqslant 1$.

Nous déduisons alors $R=1$. △

15.2.2 Somme et produit de séries

THOooSDQQooIawBOk

Théorème 15.27.
Soient $\sum_n a_n z^n$ et $\sum b_n z^n$ deux séries de rayon de convergences respectivement R_a et R_b. Si R_s est le rayon de convergence de $\sum_n (a_n+b_n)z^n$, nous avons

$$R_s \geqslant \min\{R_a, R_b\}. \tag{15.77}$$

11. Définition 15.12.
12. Théorème 10.140.

THOooINHDooZxErnp

Théorème 15.28.
Soit une série $\sum_n a_n z^n$ de rayon de convergence R.

(1) Si $\lambda \neq 0$ la série $\sum_n (\lambda a_n) z^n$ a le même rayon de convergence que la série $\sum_n a_n z^n$

(2) Si $|z| < R_a$ nous avons

$$\sum_{n=0}^{\infty} (\lambda a_n) z^n = \lambda \sum_{n=0}^{\infty} a_n z^n. \tag{15.78}$$

LEMooNYAXooKUuQFe

Lemme 15.29 ([1]).
Soient deux suites de nombres complexes (a_n) et (b_n). Soit $n \in \mathbb{N}$. Nous avons :

$$\sum_{k=0}^{n} \sum_{l=0}^{k} a_l b_{k-l} = \sum_{l=0}^{n} \big(\sum_{j=0}^{n-l} b_j \big) a_l. \tag{15.79}$$

Démonstration. Le problème est qu'à gauche la borne de la somme sur l dépend de k ; cela nous empêche de permuter les sommes [13]. Qu'à cela ne tienne : nous complétons la somme en introduisant

$$\sigma_{lk} = \begin{cases} 1 & \text{si } l \leqslant k \\ 0 & \text{sinon .} \end{cases} \tag{15.80}$$

Nous avons alors

$$\sum_{k=0}^{n} \sum_{l=0}^{k} a_l b_{k-l} = \sum_{k=0}^{n} \sum_{l=0}^{n} \sigma_{lk} a_l b_{k-l} = \sum_{l=0}^{n} \sum_{k=0}^{n} \sigma_{lk} a_l b_{k-l} = \sum_{l=0}^{n} \big(a_l \sum_{k=l}^{n} b_{k-l} \big) \tag{15.81a}$$

$$= \sum_{l=0}^{n} \big(a_l \sum_{j=0}^{n-l} b_j \big) = \sum_{l=0}^{n} \big(\sum_{j=0}^{n-l} b_j \big) a_l. \tag{15.81b}$$

□

LEMooLPBCooRWuvJB

Lemme 15.30 ([1]).
Si la série $\sum_{n=0}^{\infty} a_n z^n$ a un rayon de convergence R, alors nous avons

$$\big(\sum_{n=0}^{\infty} a_n z^n \big) z = \sum_{n=0}^{\infty} a_n z^{n+1} \tag{15.82}$$

et les rayons de convergences sont égaux à R.

PROPooPKGEooZKyxwo

Proposition 15.31 ([482]).
Soient (a_n) et (b_n) des suites dans $\mathbb{C}$. Nous supposons que $\sum_n a_n$ est absolument convergente [14] et que $\sum_n b_n$ est convergente.

Alors en posant

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}, \tag{15.83}$$

la série $\sum_n c_n$ est convergente et

$$\sum_{n=0}^{\infty} c_n = \big(\sum_i a_i \big) \big(\sum_j b_j \big). \tag{15.84}$$

13. Vu que toutes les sommes sont finies, ce ne sont certainement pas les questions de convergence qui nous retiennent.

14. Définition 10.92.

Démonstration. Nous commençons par quelques notations sur les sommes partielles et leurs limites. Nous posons $A_n = \sum_{k=0}^{n} a_n$, $B_n = \sum_{k=0}^{n} b_n$ et nous avons, par hypothèse, les convergences $A_n \xrightarrow{\mathbb{C}} A$ et $B_n \xrightarrow{\mathbb{C}} B$.

En ce qui concerne la somme partielle pour les (c_n), en appliquant le lemme 15.29,

$$C_n = \sum_{k=0}^{n} c_k = \sum_{k=0}^{n}\sum_{l=0}^{k} a_l b_{k-l} = \sum_{l=0}^{n}\big(\sum_{j=0}^{n-l} b_j\big)a_l = \sum_{l=0}^{n} a_l B_{n-l}. \tag{15.85}$$

Soit $\epsilon > 0$.

(1) **Des indices assez grands** Nous définissons $N_1, N_2 \in \mathbb{N}$ de la façon suivante :

— Vu que $B_j \to B$, il existe $N_1 \in \mathbb{N}$ tel que pour tout $j \geqslant N_1$, $|B_j - B| \leqslant \epsilon$.

— Vu que $\sum_n a_n$ converge, la proposition 10.100 nous dit que $|a_i| \to 0$ (ce n'est pas ici que nous utilisons la convergence absolue). Il existe donc $N_2 \in \mathbb{N}$ tel que $|a_i| \leqslant \epsilon/N_1$ pour tout $i \geqslant N_2$.

— Nous considérons $N \geqslant N_1 + N_2$.

(2) **Un majorant** Vu que la série $\sum_n a_n$ converge absolument, la somme $\sum_{n=0}^{\infty} |a_n|$ est bornée. De même, la suite $j \mapsto |B_j - B|$ est bornée et nous choisissons M assez grand pour majorer les deux en même temps :

$$\sum_{n=0}^{\infty} |a_n| < M \tag{15.86a}$$

$$|B_j - B| < M \quad \forall j. \tag{15.86b}$$

(3) **Et on calcule un peu** Nous avons assez préparé de notations et de majorations. C'est le moment de prouver que $C_n - A_n B \to 0$. Nous avons

$$|C_n - A_n B| = |\sum_{l=0}^{n} a_k(B_{n-l} - B)| \tag{15.87a}$$

$$\leqslant \sum_{l=0}^{n} |a_l||B_{n-l} - B| \tag{15.87b}$$

$$= \sum_{l=0}^{n-N_1} |a_l||B_{n-l}| + \sum_{l=n-N_1+1}^{n} |a_l||B_{n-l} - B| \tag{15.87c}$$

$$\leqslant \sum_{l=0}^{n-N_1} |a_l|\epsilon + M \sum_{l=n-N_1+1}^{n} |a_l| \tag{15.87d}$$

$$\leqslant \epsilon M + M \sum_{l=n-N_1+1}^{n} \frac{\epsilon}{N_1} \tag{15.87e}$$

$$= 2\epsilon M. \tag{15.87f}$$

Justifications :

— Pour (15.87d). Dans la première somme, $n-l \geqslant n-(n-N_1) = N_1$, donc $|B_{n-l}-B| \leqslant \epsilon$. Dans la seconde somme nous avons seulement majoré $|N_{n-l} - B| \leqslant M$.

— Pour (15.87e). Dans la première somme, il s'agit de la majoration (15.86a). Dans le seconde somme, $l \geqslant n - N_1 + 1 \geqslant N_1 + N_2 - N_1 + 1 \geqslant N_2 + 1$, ce qui implique $|a_l| \leqslant \epsilon/N_1$.

— Pour (15.87f). La somme contient $n - (n - N_1 + 1) + 1 = N_1$ termes. Chaque terme valant ϵ/N_1.

(4) **Conclusion** Nous avons prouvé que pour tout $\epsilon > 0$, il existe N tel que $|C_n - A_nB| \leqslant \epsilon$. Nous écrivons maintenant

$$C_n = (C_n - A_nB) + A_nB \tag{15.88}$$

Vu que $C_n - A_nB \to 0$ et que $A_nB \to AB$, la somme des deux suites converge vers[15] $0 + AB = A$ et donc

$$C_n \to AB. \tag{15.89}$$

□

Le théorème suivant donne une formule (dit « produit de Cauchy ») pour le produit de deux séries entières. Nous en donnons une adaptation dans le cas de séries de puissances dans une algèbre normée dans la proposition 10.107.

ThokPTXYC

Théorème-Définition 15.32 (Produit de Cauchy dans $\mathbb{C}$[483])**.**
Soient $\sum_n a_n z^n$ et $\sum b_n z^n$ deux séries de rayon de convergences respectivement R_a et R_b. La série entière

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) z^n. \tag{15.90}$$

EqFPGGooDQlXGe

*est le **produit de Cauchy** des séries $\sum_n a_n z^n$ et $\sum_n b_n z^n$.*

Nous notons R_p le rayon de convergence de la série.

ITEMooFOVPooBaVknN

(1) Nous avons l'inégalité $R_p \geqslant \min\{R_a, R_b\}$.

ITEMooHRNZooWviigD

(2) Si $|z| < \min\{R_a, R_b\}$ alors

$$\sum_{n=0}^{\infty} \left(\sum_{i+j=n} a_i b_j \right) z^n = \left(\sum_{n=0}^{\infty} a_n z^n \right) \left(\sum_{n=0}^{\infty} b_n z^n \right). \tag{15.91}$$

EQooSGXHooHwjQEV

Démonstration. En plusieurs parties.

(1) **Préambule** Nous allons fixer z, et utiliser la proposition 15.31. Ce que nous appelons a_n là-bas est $a_n z^n$ ici. Idem pour les b_n qui sont $b_n z^n$ et c_n qui devient $c_n z^n$. Vu que z sera fixé, tout cela n'est pas très profond.

Ces substitutions sont très courantes lorsque nous prouvons des résultats sur les séries entières comme corolaires de résultats généraux sur les séries.

(2) **Pour (1)** Si $|z| < \min\{R_a, R_b\}$, alors les séries $\sum_n a_n z^n$ et $\sum_n b_n z^n$ sont absolument convergentes par lemme d'Abel 15.18. La proposition 15.31 pour les suites $(a_n z^n)$ et $(b_n z^n)$ fait alors le boulot : en posant

$$c_n = \sum_{k=0}^{n} a_n b_{n-k}, \tag{15.92}$$

la série $\sum_n c_n z^n$ converge et vaut le produit des deux.

La série $\sum_n c_n z^n$ converge sur $\{z \in \mathbb{C} \text{ tel que } |z| < \min\{R_a, R_b\}\}$. Donc en posant $r < \min\{R_a, R_b\}$, la suite $(c_n r^n)$ est bornée (dans $\mathbb{C}$) et nous avons que $R_p \geqslant r$ (utilisation très littérale de la définition du rayon de convergence). Donc $R_p \geqslant \min\{R_a, R_b\}$.

(3) **Pour (2)** Le travail est déjà fait.

□

Exemple 15.33.
Montrons un produit de Cauchy dont le rayon de convergence est strictement plus grand que le minimum. D'abord nous considérons

$$A = 1 - z, \tag{15.93}$$

15. C'est la proposition 9.28.

c'est-à-dire $a_0 = 1$, $a_1 = -1$, $a_{n\geqslant 2} = 0$ avec $R_a = \infty$. Ensuite nous considérons

$$B = \sum_n z^n, \tag{15.94}$$

c'est-à-dire $B = (1-z)^{-1}$ et $R_b = 1$. Le produit de Cauchy de ces deux séries valant 1, le rayon de convergence est infini.

Notons qu'alors l'égalité (15.91) a lieu dans $B(0,1)$, mais pas au-delà.

Donc le « produit de Cauchy » de deux séries peut ne pas être égal au produit des deux séries, au sens où il est possible que le produit existe là où une des deux séries n'existe plus. △

Exemple 15.34.
Nous montrons que

$$\sum_{n=0}^{\infty}(n+1)x^n = \frac{1}{(1-x)^2} \tag{15.95}$$

pour $x \in]-1,1[$.

Étant donné que pour tout r dans $]-1,1[$ la suite $(n+1)r^n$ est bornée, le rayon de convergence est correct. Pour les x dans ce domaine nous avons

$$\frac{1}{(1-x)^2} = \frac{1}{(1-x)}\frac{1}{(1-x)} = \left(\sum_{n=0}^{\infty}x^n\right)\left(\sum_{m=0}^{\infty}z^m\right). \tag{15.96}$$ EqIwbuTk

Nous devons expliciter ce produit de Cauchy en utilisant le théorème 15.32. Pour tout i nous avons $a_i = b_i = 1$. Par conséquent le produit (15.96) devient

$$\sum_{n=0}^{\infty}\sum_{i+j=n}x^n = \sum_{n=0}^{\infty}(n+1)x^n. \tag{15.97}$$

△

Nous voulons maintenant faire le produit de Cauchy à plus que deux facteurs. Pour cela nous prouvons d'abord un certain nombre de lemmes traitant de la combinatoire du problème.

Nous posons, pour $n, N \in \mathbb{N}$:

$$V_n(N) = \{x \in \mathbb{N}^N \text{ tel que } \sum_{i=1}^{N}x_i = n\}. \tag{15.98}$$ EQooJCBSooMSbaCd

LEMooRKEVooDdpuHt

Lemme 15.35.
Nous avons

$$V_n(N+1) = \bigcup_{y=0}^{n}\bigcup_{x\in V_y(N)}(x, n-y). \tag{15.99}$$

Démonstration. En deux parties.

(1) **Première inclusion** Un élément de $\bigcup_{y=0}^{n}\bigcup_{x\in V_u(N)}(x,n-y)$ est un élément $z \in \mathbb{N}^{N+1}$ de la forme $z = (x, n-y)$ tel que $x \in V_y(N)$. Donc

$$\sum_{i=1}^{N+1}z_i = \sum_{i=1}^{N}x_i + (n-y) = y + n - y = n. \tag{15.100}$$

Donc $z \in V_n(N+1)$.

(1a) **L'autre inclusion** Un élément de $V_n(N+1)$ est de la forme $z = (x,y)$ avec $x \in \mathbb{N}^N$ et $y \in \mathbb{N}$. Posons $t = \sum_{i=1}^{N}x_i$, de telle sorte que $x \in V_t(N)$.
Vu que $z \in V_n(N+1)$ nous avons d'autre part $y = n - \sum_{i=1}^{N}x_i = n - t$.

□

La proposition suivante généralise le produit de Cauchy du théorème 15.32 au cas de plus de deux facteurs. Nous ne pouvons cependant pas considérer 15.32 comme un cas particulier de 15.36, parce que la démonstration va utiliser le cas à deux facteurs.

PROPooJPVVooLqSdSn

Proposition 15.36 (Produit de Cauchy[1])**.**
Soient des nombres complexes a_{ik} tels que les séries entières

$$s_i(z) = \sum_{k=0}^{\infty} a_{ik} z^k \tag{15.101}$$

soient convergentes avec un rayon de convergence R_i. Nous posons

$$c_n = \sum_{x \in V_n(N)} \prod_{i=1}^{N} a_{ix_i}, \tag{15.102}$$

et nous appelons R_p le rayon de convergence de la série $\sum_n c_n z^n$.

(1) Nous avons l'inégalité $R_p \geqslant \min\{R_i\}$.

ITEMooUVNXooLxlawx

(2) Pour $|z| \leqslant \min\{R_i\}$ nous avons l'égalité

$$\prod_{i=1}^{N} s_i(z) = \sum_{n=1}^{N} \left(\sum_{x \in V_n(N)} \prod_{i=1}^{N} a_{ix_i} \right) z^n. \tag{15.103}$$

EQooHCUGooDRhxzt

Démonstration. Nous prouvons cela par récurrence sur N. D'abord pour $N = 1$ nous avons $V_n(1) = \{n\}$. Donc $c_n = \prod_{i=1}^{1} a_{ix_i} = a_{1n}$ et donc la série à droite dans (15.103) est seulement $\sum_{n=0}^{\infty} a_{1n} z^n = s_1(z)$.

Nous supposons le théorème prouvé pour toutes valeurs jusqu'à N et nous prouvons pour $N+1$. Si $|z| < \min\{R_i\}$ alors toutes les séries convergent et en utilisant l'associativité du produit dans $\mathbb{C}$ nous avons :

$$\prod_{i=1}^{N+1} s_i(z) = \left(\prod_{i=1}^{N} s_i(z) \right) s_{N+1}(z) = \sum_{n=1}^{\infty} \left(\sum_{x \in V_n(N)} \prod_{i=1}^{N} a_{ix_i} \right) z^n s_{N+1}(z). \tag{15.104}$$

Nous allons maintenant utilliser le produit de Cauchy à deux termes du théorème 15.31. Notez que c'est bien l'utilisation de ce théorème qui nous permet d'obtenir la convegence dans notre pas de récurrence, et non l'hypothèse de réccurence actuelle. Bref, nous posons

$$b_{1k} = \sum_{x \in V_k(N)} \prod_{i=1}^{N} a_{ix_i} \tag{15.105a}$$

$$b_{2k} = a_{N+1,k}. \tag{15.105b}$$

L'utilisation du produit de Cauchy à deux facteurs donne le coefficient de z^n sous la forme suivante

(justifications plus bas) :

$$c_n = \sum_{y\in V_n(2)} b_{1y_1}b_{2y_2} \tag{15.106a}$$

$$= \sum_{y\in V_n(2)} \sum_{x\in V_{y_1}(N)} \prod_{i=1}^{N} a_{ix_i}a_{N+1,y_2} \tag{15.106b}$$

$$= \sum_{y=0}^{n} \sum_{x\in V_y(N)} \prod_{i=1}^{N} a_{ix_i}a_{N+1,n-y} \tag{15.106c}$$

$$= \sum_{(x,y)\in V_n(N+1)} \prod_{i=1}^{N} a_{ix_i}a_{N+1,y} \tag{15.106d}$$

$$= \sum_{x\in V_n(N+1)} \prod_{i=1}^{N+1} a_{ix_i}. \tag{15.106e}$$

Justifications :

— Pour (15.106c). L'ensemble $V_n(2)$ n'est pas très compliqué à expliciter :

$$V_n(2) = \{(y, n-y) \text{ tel que } y = 0, \dots, n\}. \tag{15.107}$$

— Pour (15.106d). La somme porte sur les (x,t) avec $x \in \mathbb{N}^N$ et $y \in \mathbb{N}$ tels que (x,y) est dans $V_n(N+1)$, et la justification de l'égalité est le lemme 15.35.

□

15.2.3 Convergence normale

Théorème 15.37.
Une série entière converge normalement sur tout disque fermé inclus au disque de convergence.

Démonstration. Toute boule fermée incluse à $B(0,R)$ est incluse à la boule $\overline{B(0,r)}$ pour un certain $r < R$. Nous nous concentrons donc sur une telle boule fermée.

Pour chaque n nous posons $u_n(z) = a_n z^n$ que nous voyons comme une fonction sur $\overline{B(0,r)}$. Pour tout $n \in \mathbb{N}$ et tout $z \in \overline{B(0,r)}$ nous avons

$$\|u_n\|_\infty \leqslant |a_n z^n| \leqslant |a_n| r^n. \tag{15.108}$$

Étant donné que $r < R$ la série $\sum_n |a_n| r^n$ converge et la série $\sum_n \|u_n\|$ est convergente. La série $\sum_n a_n z^n$ est alors normalement convergente. □

Exemple 15.38.
Encore une fois nous n'avons pas d'informations sur le comportement au bord. Par exemple la série $\sum_n z^n$ a pour rayon de convergence $R = 1$, mais $\sup_{z\in B(0,1)} |z^n| = 1$ et nous n'avons pas de convergence normale sur la boule fermée. △

La convergence normale n'est donc pas de mise sur tout l'intérieur du disque de convergence. La continuité, par contre est effective sur la boule. En effet si $z_0 \in B(0,R)$ alors il existe un rayon $0 < r < R$ tel que $B(z_0,r) \subset B(0,R)$. Sur $B(z_0,r)$ nous avons convergence normale et donc continuité en z_0.

La différence est que la continuité est une propriété locale tandis que la convergence normale est une propriété globale.

Proposition 15.39.
Soit $f(z) = \sum_n a_n z^n$ avec un rayon de convergence R. Si $\sum |a_n| R^n$ converge alors

(1) la série $\sum_n a_n z^n$ converge normalement sur $\overline{B(0,R)}$,

(2) f est continue sur $\overline{B(0,R)}$.

Démonstration. La conclusion est claire dans l'intérieur du disque de convergence. En ce qui concerne le bord, chacune des sommes partielles est une fonction continue. De plus nous avons $\|u_n\| \leqslant |a_n|R^n$, dont la série converge. Par conséquent nous avons convergence normale sur le disque fermé. □

Le théorème suivant permet de donner, dans le cas de fonctions réelle, des informations sur la convergence en une des deux extrémités de l'intervalle de convergence.

ThoLUXVjs

Théorème 15.40 (Convergence radiale de Abel)**.**
Soit $f(x) = \sum_n a_n x^n$ une série réelle de rayon de convergence $0 < R < \infty$.

(1) Si $\sum a_n R^n$ converge, alors f est continue sur $[0, R]$.

(2) Si $\sum_n a_n(-R)^n$ converge, alors f est continue sur $[-R, 0]$.

La proposition 15.102 donnera un exemple d'utilisation pour la série de $\ln(1-x)$ (qui n'est pas encore définie à ce moment).

Le résultat suivant permet d'identifier deux séries complexes lorsque leurs valeurs sur $\mathbb{R}$ sont identiques.

Proposition 15.41.
Soient les séries $f(z) = \sum a_n z^n$ et $g(z) = \sum b_n z^n$ convergentes dans $B(0,R)$. Si $f(x) = g(x)$ pour $x \in [0, R[$ alors $a_n = b_n$.

Démonstration. Soit n_0 le plus petit entier tel que $a_{n_0} \neq b_{n_0}$. Pour tout $z \in B(0,R)$ nous avons

$$f(z) - g(z) = \sum_{n=n_0}^{\infty} (a_n - b_n) z^n = z^{n_0} \varphi(z) \tag{15.109}$$

où

$$\varphi(z) = \sum_{n \geqslant 0} (a_{n+n_0} - b_{n+n_0}) z^n. \tag{15.110}$$

Par le théorème 15.27 le rayon de convergence de φ est plus grand que R et la fonction φ est continue en 0. Étant donné que $\varphi(0) = a_{n_0} - b_{n_0} \neq 0$ et que φ est continue nous avons un ρ tel que $\varphi \neq 0$ sur $B(0,\rho)$. Or cela n'est pas possible parce que au moins sur la partie réelle de cette dernière boule, φ doit être nulle. □

PropSNMEooVgNqBP

Proposition 15.42 ([484, 1])**.**
Si la série entière $\sum_{n \geqslant 0} a_n z^n$ a un rayon de convergence R alors

(1) La somme est une fonction holomorphe[16] *dans le disque de convergence.* ItemUULDooEGRNiA

(2) La somme est différentiable et

$$du_{z_0}(z) = \sum_{n=1}^{\infty} n a_n z_0^{n-1} z. \tag{15.111}$$

(3) De plus pour tout $z_0 \in B(0,R)$, on pose[17]

$$S(z) = \sum_{n \geqslant 0} a_n z^n \tag{15.112a}$$

$$T(z) = \sum_{n \geqslant 1} n a_n z^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} z^n. \tag{15.112b}$$

Alors nous avons

$$\lim_{z \to z_0} \frac{S(z) - S(z_0)}{z - z_0} = T(z_0). \tag{15.113}$$ EqVQDPooQPICwN

16. Définition 12.305.
17. Pour rappel, dans tout ce texte, $B(a,r)$ est une boule *ouverte*.

Démonstration. Nous allons prouver, en utilisant le théorème 15.9, que la somme est une fonction différentiable et que la différentielle est $\mathbb{C}$-linéaire. La proposition 12.311 nous dira alors que la somme est $\mathbb{C}$-dérivable.

Nous posons $u_n(z) = a_n z^n$, qui est une fonction de classe C^1. En ce qui concerne sa différentielle nous considérons $z_0 \in B(0,R)$ et nous avons (si $n = 0$ alors la différentielle est nulle)

$$(du_n)_{z_0}(z) = \frac{d}{dt}\Big[u_n(z_0+tz)\Big]_{t=0} \tag{15.114a}$$
$$= \frac{d}{dt}\Big[a_n(z_0+tz)^n\Big]_{t=0} \tag{15.114b}$$
$$= \frac{d}{dt}\Big[na_n(z_0^{n-1}tz)\Big]_{t=0} \tag{15.114c}$$
$$= na_n z_0^{n-1} z. \tag{15.114d}$$

En cours de calcul nous avons développé $(z_0+tz)^n$ et gardé seulement les termes de degré 1 en t. Il y en a n et ils sont tous égaux à $z_0^{n-1}tz$.

La convergence simple $\sum_n u_n$ est dans les hypothèses. Il reste à prouver que la somme des différentielles converge uniformément sur tout compact autour de z_0 ne débordant pas du disque ouvert de convergence. Soit K un compact autour de z_0. Dans le calcul suivant nous utilisons une première fois la norme uniforme de du_n vu comme fonction de K vers $\mathcal{L}(\mathbb{C},\mathbb{C})$ et une fois la norme opérateur [18] de $(du_n)_{z_0}$ comme application linéaire $\mathbb{C}\to\mathbb{C}$:

$$\|du_n\|_k = \sup_{z_0\in K}\|(du_n)_{z_0}\| \tag{15.115a}$$
$$= \sup_{z_0\in K}\sup_{|z|=1}|(du_n)_{z_0}(z)| \tag{15.115b}$$
$$= \sup_{z_0\in K}\sup_{|z|=1}|na_n z_0^{n-1}z| \tag{15.115c}$$
$$= \sup_{z_0\in K} n|a_n||z_0|^{n-1}. \tag{15.115d}$$

Vu que $z\mapsto|z|^{n-1}$ est une application continue sur le compact K, elle atteint son maximum (théorème 7.146). Nous considérons z_K, un point qui réalise le supremum. Ce nombre est dans le disque de convergence parce que K est un compact autour de z_0.

Nous devons prouver que $\sum_n n|a_n||z_K|^{n-1}$ converge. Vu que $|z_K|$ est une constante (par rapport à n) nous pouvons étudier la convergence en écrivant $|z_K|^n$ au lieu de $|z_K|^{n-1}$.

La suite $(a_n|z_K|^n)$ est une suite bornée. En considérant un majorant M de $\{|a_n|r^n\}_{n\in\mathbb{N}}$, nous avons en particulier $|a_n||z_K|^n < M$ pour tout n. Nous considérons de plus r de telle sorte que $K\subset B(0,r)\subset B(0,R)$. En particulier $|z_K|<r$ et nous avons

$$n|a_n||z_K|^n \leqslant n|a_n|r^n\left(\frac{|z_K|}{r}\right)^n \leqslant nM\left(\frac{|z_K|}{r}\right)^n. \tag{15.116}$$

Nous savons que ce qui est dans la parenthèse est plus petit que 1, mais que $\sum_n nx^n$ converge dès que $|x|<1$. Par conséquent

$$\sum_n \|du_n\|_K \tag{15.117}$$

converge et le théorème 15.9 fonctionne : $du = \sum_{n=1}^{\infty} du_n$ et la somme $\sum_n u_n$ est de classe C^1.

La différentielle de $\sum_n u_n$ s'exprime explicitement par

$$du_{z_0}(z) = \sum_{n=1}^{\infty} na_n z_0^{n-1} z. \tag{15.118}$$ EqJBFMqcMiSABz

Cette forme montre que du_{z_0} est une application $\mathbb{C}$-linéaire et donc la somme est $\mathbb{C}$-dérivable par la proposition 12.311. Ergo holomorphe sur le disque de convergence par définition 12.305.

18. Définition 10.55.

En ce qui concerne la formule (15.113), elle provient de la formule (12.838) : $f'(z_0)$ est donné par la facteur multiplicatif de du_{z_0}. En l'occurrence la formule (15.118) nous donne

$$f'(z_0) = \sum_{n\geqslant 1} n a_n z_0^{n-1}. \tag{15.119}$$

□

15.2.4 Dérivation

LemFVMaSD

Lemme 15.43.
Soit une série entière $\sum a_n z^n$ de rayon de convergence R. Les séries

$$\sum \frac{a_n}{n+1} z^{n+1} \tag{15.120}$$

et

$$\sum_{n\geqslant 1} n a_n z^{n-1} \tag{15.121}$$

ont même rayon de convergence R.

Notons toutefois que nonobstant ce lemme, les séries dont il est question peuvent se comporter différemment sur le bord du disque de convergence. En effet la série

$$\sum \frac{1}{n} z^n \tag{15.122}$$

diverge pour $z = 1$ alors que

$$\sum \frac{1}{n(n+1)} z^{n+1} \tag{15.123}$$

converge pour $z = 1$.

Les théorèmes de dérivation et d'intégration de séries de fonctions (théorèmes 15.1 et 15.3) fonctionnent bien dans le cas des séries entières. Ils donnent la proposition 15.44 pour la dérivation et 15.49 pour l'intégration.

ProptzOIuG

Proposition 15.44.
Soit la série entière

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{15.124}$$

de rayon de convergence R. Alors la fonction f est C^1 sur $]-R,R[$ et se dérive terme à terme :

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \tag{15.125}$$

pour tout $x \in]-R,R[$.

Démonstration. Nous savons que la série $\sum_{n=1}^{\infty} n a_n x^{n-1}$ a le même rayon de convergente que celui de la série f. En particulier cette série des dérivées converge normalement sur tout compact dans $]-R,R[$ et la somme est continue. Le théorème 15.3 conclut. □

Remarque 15.45.
À part lorsqu'on parle de fonction $\mathbb{R} \to \mathbb{R}$, la notion de classe C^k s'entend au sens de la différentielle, et non de la dérivée, voir les définitions 10.261. C'est cela qui explique la structure de la démonstration de la proposition 15.42.

CorCBYHooQhgara

Corolaire 15.46 ([484, 1])**.**
La somme d'une série entière est de classe C^∞ sur le disque ouvert de convergence.

Démonstration. La proposition 15.42 a démontré en réalité nettement plus : sur le disque ouvert de convergence, la somme est une fonction holomorphe. Il n'est cependant pas possible de conclure ainsi parce que le fait qu'une fonction holomorphe est C^∞ ne sera démontré qu'au coût de nombreux efforts dans le théorème 26.60(3).

(1) **Cas réel** Nous considérons la série entière $\sum_n a_n x^n$ pour $x \in \mathbb{R}$ de rayon de convergence R. Une simple récurrence sur la proposition 15.44 donne le résultat.

(2) **Cas complexe** Attention : le fait d'être de classe C^k est le fait d'être k fois *différentiable*. Rien à voir avec la $\mathbb{C}$-dérivabilité.

En ce qui concerne la différentiabilité nous avons la proposition 15.42 qui dit que dans le disque de convergence, la fonction $u(z) = \sum_n a_n z^n$ a pour différentielle l'application $du\colon \mathbb{C} \to \mathcal{L}_{\mathbb{C}}(\mathbb{C},\mathbb{C})$,

$$\begin{aligned} du\colon \mathbb{C} &\to \mathcal{L}_{\mathbb{C}}(\mathbb{C},\mathbb{C}) \\ du_{z_0}(z) &= \Big(\sum_{n=0}^{\infty}(n+1)a_{n+1}z_0^n\Big)z. \end{aligned} \tag{15.126}$$

Nous allons éviter de considérer la différentielle seconde comme une application

$$d^2u\colon \mathbb{C} \to \mathcal{L}\big(\mathbb{C},\mathcal{L}(\mathbb{C},\mathbb{C})\big) \tag{15.127}$$

parce que ça nous mènerait trop loin pour parler de la différentielle k^{e}. Au lieu de cela nous allons considérer l'isomorphisme d'espace vectoriel

$$\begin{aligned} \psi\colon \mathbb{C} &\to \mathcal{L}_{\mathbb{C}}(\mathbb{C},\mathbb{C}) \\ z_0 &\mapsto \psi(z_0)z = z_0 z. \end{aligned} \tag{15.128}$$

Dans cette optique nous écrivons :

$$du_{z_0} = \psi\Big(\sum_{n=0}^{\infty}(n+1)a_{n+1}z_0^n\Big) \tag{15.129}$$

ou encore :

$$(\psi^{-1}\circ d)u(z_0) = \sum_{n\geqslant 0}(n+1)a_{n+1}z_0^n. \tag{15.130}$$

Nous allons prouver par récurrence que l'égalité suivante est vraie (y compris le fait que la somme converge) :

$$(\psi^{-1}\circ d)^k u(z_0) = \sum_{n=0}^{\infty}\frac{(n+k)!}{n!}a_{n+k}z_0^n. \tag{15.131}$$

Prouvons d'abord que cette somme converge pour tout k. Nous avons $(n+k)!/n! < (n+k)^k$ et donc il suffit de prouver que la série de coefficients $n^k a_n$ converge. C'est le cas par la proposition 15.21.

Nous pouvons calculer la différentielle de $(\psi^{-1}\circ d)^k u$ en dérivant terme à terme en utilisant (encore) la proposition 15.42(2) :

$$d\big((\psi^{-1}\circ d)^k u\big)_{z_0}(z) = \sum_{n=1}^{\infty}\frac{(n+k)!}{n!}a_{n+k}na_0^{n-1}z \tag{15.132a}$$

$$= \sum_{n=0}^{\infty}\frac{(n+k+1)!}{n!}a_{n+k+1}z_0^n z. \tag{15.132b}$$

Nous appliquons ψ^{-1} à cela :

$$(\psi^{-1}\circ d)^{k+1}u(z_0) = \sum_{k=0}^{\infty}\frac{(n+k+1)!}{n!}a_{n+k+1}z_0^n. \tag{15.133}$$

(3) **Dérouler à l'envers**

Nous allons maintenant utiliser la proposition 12.275 pour montrer que u est de classe C^k pour tout k. Nous avons démontré que $(\psi^{-1}\circ d)^k u$ était différentiable. Par conséquent, $d\big((\psi^{-1}\circ d)^{k-1}u\big)$ est différentiable et donc $(\psi^{-1}\circ d)^{k-1}$ est de classe C^1. En continuant ainsi, $(\psi^{-1}\circ d)^{k-l}u$ est de classe C^l et u est de classe C^k.

□

Le lemme suivant est encore essentiellement valable dans un espace de Banach (proposition 10.299).

LemPQFDooGUPBvF

Lemme 15.47.
Plusieurs choses sur des séries entières.

(1) La série entière $\sum_{n\geqslant 0} z^{nk}$ a un rayon de convergence 1 et converge vers la fonction

$$\sum_{n\geqslant 0} z^{nk} = \frac{1}{1-z^k}. \tag{15.134}$$

(2) Lorsque $|\omega|=1$, la série

$$\frac{1}{\omega - z} = \sum_{k\geqslant 0} \frac{z^k}{\omega^{k+1}}. \tag{15.135}$$

EqSSHZooLwCBAZ

a un rayon de convergence égal à 1.

ITEMooHFVHooPCgzZV

(3) Si $|\omega|=1$, la série

$$\frac{1}{(\omega-z)^k} = \frac{1}{(k-1)!}\sum_{s=0}^{\infty}\frac{(s+k-1)!}{s!}\frac{z^s}{\omega^{s+k+1}} \tag{15.136}$$

a un rayon de convergence égal à 1.

Démonstration. Les coefficients de la série sont $a_n = 1$ lorsque n est multiple de k et $a_n = 0$ autrement. Donc pour $r = 1$ la suite $r^n a_n$ reste bornée[19]. Cela prouve que le rayon de convergence est au moins 1. Par ailleurs si $r > 1$ alors clairement la suite $(a_n r^n)$ n'est pas bornée. Cela prouve le rayon de convergence égal à 1.

Soit donc $z \in B(0,1)$. Nous avons, par le lemme 15.30,

$$\left(\sum_{n=0}^{\infty} z^{nk}\right)(1-z^k) = \sum_{n=0}^{\infty} z^{nk} - \sum_{n=0}^{\infty} z^{(n+1)k} \tag{15.137a}$$

$$= 1 + \sum_{n=1}^{\infty} z^{nk} - \sum_{n=0}^{\infty} z^{(n+1)k} \tag{15.137b}$$

$$= 1 + \sum_{n=0}^{\infty} z^{(n+1)k} - \sum_{n=0}^{\infty} z^{(n+1)k} \tag{15.137c}$$

$$= 1 \tag{15.137d}$$

En ce qui concerne la série (15.135), elle s'obtient facilement :

$$\frac{1}{\omega - z} = \frac{1}{\omega}\frac{1}{1-\frac{z}{\omega}} = \frac{1}{\omega}\sum_{s=0}^{\infty}\left(\frac{z}{\omega}\right)^s = \sum_s \omega^{-s-1}z^s. \tag{15.138}$$

La troisième série s'obtient en dérivant la seconde, ce qui est permis dans le disque de convergence par la proposition 15.44. □

Remarque 15.48.
Sur le bord du disque de convergence, la série $\sum_n z^{nk}$ ne converge pas. En effet le rayon étant 1, sur le bord nous avons la série $\sum_n e^{ink\theta}$ dont la norme du terme général ne tend pas vers zéro.

19. Utilisation directe de la définition 15.12.

15.2.5 Intégration

PropfeFQWr

Proposition 15.49.
Soit la série entière $\sum a_n x^n$ de rayon de convergence R.

(1) Pour tout segment $[a,b] \subset]-R,R[$ nous pouvons intégrer terme à terme :

$$\int_a^b \left(\sum_{n=0}^{\infty} a_n x^n \right) dx = \sum_{n=0}^{\infty} a_n \int_a^b x^n dx. \tag{15.139}$$

(2) La série entière obtenue en intégrant terme à terme a le même rayon de convergence que celui de la série de départ.

Démonstration. La première assertion est un cas particulier du théorème général 15.1. Pour le rayon de convergence, le lemme 15.43 fait le travail. □

Vu que le rayon de convergence ne varie pas par la dérivation ou par l'intégration et qu'une série entière est de classe C^∞ sur son disque de convergence, nous pouvons dériver terme à terme autant de fois que nous le voulons sans faire de fautes dans le disque de convergence.

15.3 Séries de Taylor

SECooDWRMooUKSuPh

15.50.
Avant de commencer, une petite formule de dérivation toute simple que nous allons utiliser souvent :

$$(z^k)^{(l)} = \begin{cases} 0 & \text{si } l > k \\ \frac{k!}{(k-l)!} z^{k-l} & \text{sinon.} \end{cases} \tag{15.140}$$

EqSOFdwhw

Dans les cas où il est permis de dériver terme à terme, nous avons la formule

$$f^{(p)}(x) = \sum_k a_k (x^k)^{(p)} = \sum_{k=p}^{\infty} a_k \frac{k!}{(k-p)!} x^{k-p} \tag{15.141}$$

EQooTNOMooJZClyE

15.3.1 Polynôme de Taylor d'une série entière

Le polynôme de Taylor d'une fonction définie par une série entière s'obtient en tronquant la série. Cela est une assez bonne nouvelle que nous allons démontrer maintenant.

PROPooQLHNooRsBYbe

Proposition 15.51 ([1]).
Soit une série entière

$$f(x) = \sum_k a_k x^k \tag{15.142}$$

de rayon de convergence $R > 0$.

Pour tout $n \in \mathbb{N}$, il existe une fonction α telle que

$$f(x) = \sum_{k=0}^{n} a_k x^k + \alpha(x) x^n \tag{15.143}$$

EQooSXUJooFjsVek

et

$$\lim_{t \to 0} \alpha(t) = 0. \tag{15.144}$$

Tout ceci étant convenu que

- *l'égalité (15.143) est uniquement valable sur le disque de convergence,*
- *La fonction α dépend de n.*

Démonstration. Le corolaire 15.46 nous indique que f est de classe C^∞ sur $]-R,R[$ et que nous pouvons dériver terme à terme.

En utilisant la formule (15.141) et en l'évaluant en $x = x_0$, tous les termes s'annulent sauf $k = p$:

$$f^{(p)}(0) = p!a_p. \tag{15.145}$$

Le théorème de Taylor 12.461 nous indique alors qu'il existe $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

$$f(x) = \sum_{k=0}^{n} a_k x^k + \alpha(x)x^n. \tag{15.146}$$

□

15.3.2 Une majoration pour le reste

LEMooOVPIooAPWFOm

Lemme 15.52.
Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$ dérivable $n+1$ fois sur $B(a,R)$. Alors pour tout $x \in B(a,r)$,

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \dots \int_a^{u_{n-1}} f^{(n)}(u_n) du_n \dots du_1. \tag{15.147}$$

Démonstration. Nous allons intensivement utiliser le théorème fondamental du calcul intégral 14.287 sous la forme de la formule (14.906). Nous avons d'abord

$$f(x) = f(a) + \int_a^x f'(u_1)du_1 = \int_a^x \big[f'(a) + \int_a^{u_1} f''(u_2)du_2\big]du_1. \tag{15.148}$$

Toute l'astuce de ce théorème est de continuer à substituer $f^{(k)}(t)$ par $f^{(k)}(a)$ plus une intégrale de a à t de $f^{(k+1)}(u)$. Nous démontrons ainsi par récurrence que

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \dots \int_a^{u_{n-1}} f^{(n)}(u_n) du_n \dots du_1. \tag{15.149}$$

EQooOWJMooHATpMV

La preuve de cela se fait en substituant

$$f^{(n)}(u_n) = f^{(n)}(a) + \int_a^{u_n} f^{(n+1)}(u_{n+1})du_{n+1} \tag{15.150}$$

et en remarquant (encore par récurrence par exemple) que

$$\int_a^x \dots \int_a^{u_{n-1}} du_n \dots du_1 = \frac{(x-a)^n}{n!}. \tag{15.151}$$

□

Le théorème suivant donne majoration du reste du polynôme de Taylor. Il est un premier pas dans la démonstration de formules comme

$$\lim_{n\to\infty} P_n(x) = f(x) \tag{15.152}$$

lorsque P_n est un polynôme de Taylor autour d'un point $a \neq x$. Nous ne saurions trop insister sur le fait que de telles formules ne seraient valables que pour une classe relativement restreintes de fonctions.

THOooEUVEooXZJTRL

Théorème 15.53 (Inégalité de Taylor[485])**.**
Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$ dérivable $n+1$ fois et telle que $|f^{(n+1)}(x)| \leqslant M_n$ sur $B(a,d)$. Alors

$$|R_n(x)| \leqslant \frac{M_n}{(n+1)!}|x-a|^{n+1} \tag{15.153}$$

où $R_n(x) = f(x) - P_n(x)$ et où P_n sont les polynômes de Taylor autour de $a \in \mathbb{R}$.

Démonstration. Nous pouvons écrire la formule du lemme 15.52 pour $n+1$ au lieu de n ; cela donne

$$f(x) = P_n(x) + \int \cdots, \tag{15.154}$$

et donc

$$|R_n(x)| = |P_n(x) - f(x)| = \int_a^x \dots \int_a^{u_n} f^{(n+1)}(u_{n+1}) du_{n+1} \cdots du_1 \tag{15.155}$$

En effectuant toutes les intégrales nous trouvons [20]

$$|R_n(x)| \leqslant \frac{M_n}{(n+1)!}|x-a|^{n+1}. \tag{15.156}$$

□

Cette formule pour le reste est très bien, mais pour l'exploiter au maximum de ses possibilités, il faudra la notion de convergence de suite de fonctions, et en particulier la notion de série de fonctions, pour pouvoir écrire

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} x^k \tag{15.157}$$

lorsque cela est possible. Nous renvoyons donc aux séries de Taylor, section 15.3, et en particulier aux fonctions analytiques de la sous-section 15.3.3.

15.3.3 Fonctions analytiques

SUBSECooXKHWooEzqGRJ

Fonction analytique : définition 12.454.

Nous avons vu les polynômes de Taylor et déjà noté qu'il n'est pas en général vrai que $\lim_{n\to\infty} P_n(x) = f(x)$ pour des x même proches du point autour duquel les polynômes de Taylor P_n sont calculés.

Nous allons maintenant étudier la classe des fonctions pour lesquelles la série de Taylor est égale à la fonction de départ. D'abord une proposition montrant que les coefficients de Taylor sont les seuls pour lesquels il est possible d'espérer avoir une telle propriété.

PROPooTRWVooETTtbP

Proposition 15.54 ([486])**.**
Soit une fonction donnée par la série entière

$$f(x) = \sum_{k=0}^{\infty} c_n (x-a)^n \tag{15.158}$$

sur la boule de convergence $B(a,R)$ avec $R>0$ (hypothèse : le rayon de convergence est strictement positif). Alors

$$c_n = \frac{f^{(n)}(a)}{n!}. \tag{15.159}$$

Démonstration. Par hypothèse, nous avons un rayon de convergence $R>0$, et le corolaire 15.46 nous indique que f y est de classe C^∞. Et nous pouvons dériver terme à terme par la proposition 15.44. Cela pour dire qu'il nous est autorisé d'utiliser la formule (15.141) pour calculer les dérivées de f au point a. Nous avons d'abord

$$f^{(p)}(x) = \sum_{n=p}^{\infty} c_n \frac{n!}{(n-p)!}(x-a)^{n-p}, \tag{15.160}$$

et donc

$$f^{(p)}(a) = c_p p! \tag{15.161}$$

qui donne immédiatement le résultat. □

20. Je me demande si je n'ai pas une faute entre n et $n+1$ quelque part. Relisez attentivement et écrivez-moi si vous trouvez une faute.

Proposition 15.55.
Soit l'intervalle $I = B(a,r)$. Si il existe M tel que

$$|f^{(n)}(x)| \leqslant \frac{M}{r^n}n! \tag{15.162}$$

pour tout $x \in B(a,r)$. Alors nous avons la convergence simple

$$P_n \to f \tag{15.163}$$

sur $B(a,r)$. Ici, P_n est le polynôme de Taylor d'ordre n pour la fonction f autour du point a.

Démonstration. Vu que nous avons $|f^{(n)}(x)| \leqslant \frac{M}{r^n}n!$ pour tout x, nous pouvons poser

$$M_n = \frac{M}{r^n}n! \tag{15.164}$$

dans le théorème 15.53 pour le faire fonctionner. Nous avons alors

$$|R_n(x)| \leqslant \frac{M}{r^n}n!\frac{1}{(n+1)!}|x-a|^{n+1} = \frac{M}{n+1}|x-a|\left|\frac{x-a}{r}\right|^n. \tag{15.165}$$

Vu que $x \in B(a,r)$ nous avons $|x-a| < r$ et donc $|(x-a)/r|^n < 1$. Nous pouvons aussi majorer $|x-a|$ par r et écrire

$$|R_n(x)| \leqslant \frac{rM}{n+1}. \tag{15.166}$$

Nous avons donc bien $\lim_{n\to\infty} R_n(x) \to 0$. □

THOooEAYXooTvVYXN

Théorème 15.56 ([487]).
Soit une fonction analytique[21] *$f\colon \mathbb{R} \to \mathbb{R}$ non identiquement nulle. La partie*

$$X = \{x \in \mathbb{R} \text{ tel que } f(x) \neq 0\} \tag{15.167}$$

est discrète[22].

Démonstration. Nous supposons que X n'est pas discret et nous prouvons que $f = 0$. Soit une suite convergente $a_i \to a$ avec $a_i \in X$. Étant donné que f est analytique en a, nous avons

$$f(x) = \sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(a)(x-a)^n, \tag{15.168}$$

et nous prouvons par induction que $f^{(n)}(a) = 0$.

D'abord $f(a) = 0$ parce que f est continue, et $f(a) = \lim_{n\to\infty} f(a_i)$.

Ensuite nous supposons que $f^{(k)}(a) = 0$ pour tout $k < n$. Nous avons

$$\lim_{x\to a}\frac{f(x)}{(x-a)^n} = \lim_{x\to a}\frac{f'(x)}{n(x-a)^{n-1}} \tag{15.169a}$$

$$= \dots \tag{15.169b}$$

$$= \lim_{x\to a}\frac{f^{(n-1)}(x)}{n!(x-a)} \tag{15.169c}$$

$$= \lim_{x\to a}\frac{f^{(n)}(x)}{n!}, \tag{15.169d}$$

21. Définition 12.454.
22. Définition 7.24.

parce que chaque étape est une indétermination du type 0/0 sur laquelle nous pouvons utiliser la règle de l'Hospital[23]. Par continuité de $f^{(n)}$, nous avons prouvé que $\lim_{x\to a} f(x)/(x-a)^n$ existe et vaut

$$\lim_{x\to a}\frac{f(x)}{(x-a)^n}=\frac{f^{(n)}(a)}{n!}. \tag{15.170}$$

En particulier, nous pouvons calculer cette limite sur la suite de a_i :

$$\frac{f^{(n)}(a)}{n!}=\lim_{x\to a}\frac{f(x)}{(x-a)^n}=\lim_{i\to\infty}\frac{f(a_i)}{(a_i-a)^n}=0. \tag{15.171}$$

Nous avons donc bien prouvé que $f^{(n)}(a)=0$. □

15.4 Algèbre engendrée par une matrice

Nous allons en dire le strict minimum indispensable pour notre propos. Pour plus de détails, voir [488], et pour nettement plus de détails, [489].

DEFooXPFVooZzPXEK

Définition 15.57.
*Si $A\in\mathbb{M}(n,\mathbb{K})$, alors l'**algèbre engendrée** par A est l'intersection de toutes les sous-algèbres de $\mathbb{M}(n,\mathbb{K})$ contenant A.*

LEMooZGFYooOHvxLy

Lemme 15.58 ([488, 1])**.**
Si le polynôme minimal[24] de A est de degré p, alors la partie $\{\mathbb{1},A,A^2,\dots,A^{p-1}\}$ est une base de $\mathrm{Alg}(A)$.

Démonstration. La partie proposée est libre parce qu'une combinaison linéaire de ses éléments est un polynôme de degré $p-1$ en A. Une annulation d'un tel polynôme serait contraire au fait que le polynôme minimal est de degré p.

Cette partie est génératrice parce que, étant elle-même une sous-algèbre de $\mathbb{M}(n)$ contenant A, elle contient $\mathrm{Alg}(A)$. □

CORooQTUQooDtjljc

Corolaire 15.59 ([1])**.**
L'algèbre $\mathrm{Alg}(A)$ est commutative.

Démonstration. Écrire deux combinaisons linéaires d'éléments de la base donnée par le lemme 15.58, et notez que les produits ne font intervenir que des produits de A. □

Une bonne question est de savoir si e^{tA} est dans $\mathrm{Alg}(A)$. Pour le savoir il va falloir d'abord définir l'exponentielle ; rendez-vous donc au lemme 15.62.

15.5 Exponentielle sur une algèbre normée

15.5.1 Définition

Dans ce qui suit, nous considérons une algèbre commutative.

DEFooSFDUooMNsgZY

Proposition-Définition 15.60 (Exponentielle[1])**.**
Soit $(A,\|.\|)$ une algèbre[25] commutative de dimension finie sur $\mathbb{C}$ munie d'une norme d'algèbre. Pour $x\in A$ nous définissons

$$\exp(x)=\sum_{k=0}^{\infty}\frac{x^k}{k!}. \tag{15.172}$$

EQooCUVTooGNQrFj

Cette définition a les propriétés suivantes :

23. Proposition 12.194.
24. Toute matrice a un polynôme minimal par le lemme 8.97.
25. Définition 1.347.

(1) *C'est bien défini pour tout $x \in A$. C'est-à-dire que pour chaque x, la série (15.172) converge.*

(2) *Cela donne une application continue* $\exp\colon A \to A$. ITEMooGGVAooVfhGuu

(3) *La fonction* exp *est différentiable et*

$$(d\exp)_x(y) = \exp(x)y, \tag{15.173}$$ EQooKWBUooLUdBAw

le dernier produit étant la structure d'algèbre sur A.

Démonstration. Pour la différentiabilité de exp, nous voulons utiliser le théorème 15.9. Pour cela nous posons

$$u_k(x) = \frac{x^k}{k!} \tag{15.174}$$

(1) **Convergence simple** Nous prouvons la convergence simple, c'est-à-dire pour chaque x séparément, de la série (15.172) dans deux buts. D'abord de nous assurer que la définition posée de exp a un sens, et ensuite pour commencer à vérifier les hypothèses du théorème 15.9.
Nous montrons que les sommes partielles forment une suite de Cauchy. Nous fixons $x \in A$ et nous posons

$$s_n = \sum_{k=0}^{n} \frac{x^k}{k!}. \tag{15.175}$$

Soient $p > q$, deux entiers. Nous avons :

$$\|s_p - s_q\| = \| \sum_{k=q+1}^{p} \frac{x^k}{k!} \| \leqslant \sum_{k=q+1}^{p} \frac{\|x^k\|}{k!} \leqslant \sum_{k=q+1}^{p} \frac{\|x\|^k}{k!} \tag{15.176}$$ EQooYNZNooDaiPhU

où nous avons utilisé le fait que la norme sur A soit une norme d'algèbre.
C'est le moment d'utiliser la série exponentielle donnée dans le lemme 10.137 que nous appliquons avec $t = \|x\|$. La série donnée par les coefficients $a_k = \|x\|^k/k!$ converge et ses sommes partielles forment en particulier une suite de Cauchy. Donc ce que nous avons à droite dans (15.176) peut être rendu arbitrairement petit lorsque p et q sont grands.

(2) u_k **est continue** Il s'agit de remarquer que $(x+h)^k = x^k + hC(x,h)$ où C est une fonction bornée de h (lorsque h est dans un voisinage de $0 \in A$). Donc

$$\|(x+h)^k - x^k\| \leqslant \|h\|\|C(x,h)\| \to 0. \tag{15.177}$$

(3) **Candidat différentielle de u_k** Nous trouvons à présent un candidat à être différentielle de u_k. Pour cela nous faisons le calcul suivant, sans trop nous soucier de la rigueur :

$$(du_k)_x(y) = \frac{d}{dt}\Big[u_k(x+ty)\Big]_{t=0} = k\frac{1}{k!}x^{k-1}y = u_{k-1}(x)y. \tag{15.178}$$

(4) u_k **est différentiable** Nous fixons $x \in A$ et nous posons $T(y) = u_{k-1}(x)y$. Ensuite nous vérifions que cela vérifie la définition de la différentielle : nous devons calculer

$$\lim_{h\to 0} \frac{u_k(x+h) - u_k(x) - T(h)}{\|h\|} = \lim_{h\to 0} \frac{(x+h)^k - x^k - kx^{k-1}h}{k!\|h\|} = \clubsuit. \tag{15.179}$$ EQooNPKGooVmEYAV

Vous vous souvenez de la formule pour $(x+h)^k$? Essayez de vous en souvenir. Le premier terme est x^k, et le second est $kx^{k-1}h$. Pour le reste c'est un polynôme dont tous les termes contiennent au moins h^2. Nous avons donc

$$\clubsuit = \lim_{h\to 0} \frac{h^2 P(x,h)}{k!\|h\|} = 0. \tag{15.180}$$

Nous en concluons que u_k est différentiable et que

$$(du_k)_x(y) = u_{k-1}(x)y. \tag{15.181}$$

(5) u_k **est de classe C^1** Nous devons démontrer que la différentielle est continue ; cela est la continuité de l'application

$$\begin{aligned} du_k\colon A &\to \mathcal{L}(A,A) \\ x &\mapsto (du_k)_x. \end{aligned} \tag{15.182}$$

La topologie sur A est celle de la norme, et celle sur $\mathcal{L}(A,A)$ est celle de la norme opérateur associée à la norme sur A. Nous avons[26] :

$$\lim_{h\to 0}\|(du_k)_{x+h}-(du_k)_x\| = \lim_{h\to 0}\sup_{\|y\|=1}\|u_{k-1}(x+h)y-u_{k-1}(x)y\| \tag{15.183a}$$

$$\leqslant \lim_{h\to 0}\sup_{\|y\|=1}\|u_{k+1}(x+h)-u_{k-1}(x)\|\|y\| \tag{15.183b}$$

$$= \lim_{h\to 0}\|u_{k+1}(x+h)-u_{k-1}(x)\|. \tag{15.183c}$$

Le fait que cette limite vaille zéro est maintenant la continuité de u_{k-1}.

(6) **Convergence normale sur tout compact**
Soit un compact K de A. Par le théorème de Borel-Lebesgue 9.24, K est fermé et borné. C'est pour ceci que nous avons supposé que A était de dimension finie sur $\mathbb{R}$. Soit donc $R>0$ tel que $\|y\|<R$ pour tout $y\in K$. Nous avons

$$\|du_k\|_K = \sup_{x\in K}\|(du_k)_x\| = \sup_{x\in K}\frac{\|x^{k-1}\|}{(k-1)!} \leqslant \sup_{x\in K}\frac{\|x\|^{k-1}}{(k-1)!} \leqslant \frac{R^{k-1}}{(k-1)!}. \tag{15.184}$$

Mais la série $\sum_{k=0}^{\infty}\frac{R^k}{k!}$ converge. Nous avons donc la convergence normale demandée.

(7) **Conclusion**
Le théorème 15.9 conclut que l'exponentielle est de classe C^1 et que sa différentielle est donnée par la formule

$$(d\exp)_x(y) = \sum_{k=0}^{\infty}(du_k)_x(y) = \sum_{k=1}^{\infty}(du_k)_x(y) = \sum_{k=0}^{\infty}u_k(x)y = \exp(x)y. \tag{15.185}$$

Notez le jeu d'indices : $du_k=0$ lorsque $k=0$ (ce qui permet de faire commencer la somme à 1) et ensuite du_k fait intervenir u_{k-1} (ce qui fait revenir le départ de la somme à $k=0$).

□

15.61.
Lorsque nous disons que la différentielle de l'exponentielle est l'exponentielle elle-même, nous référons au point 15.60(3) : la différentielle de exp en x est l'opérateur de multiplication par $\exp(x)$.

Nous pouvons comprendre maintenant que exp est même de classe C^∞ parce qu'à chaque différentiation nous tombons sur la même fonction, laquelle est de classe au moins C^1.

Cependant, pour formaliser ça, il faut un peut travailler. Le cauchemar des différentielles successives d'une application $A\to A$ est que les espaces en jeu sont des emboîtements terribles de $\mathcal{L}(A,\mathcal{L}(A,\mathcal{L}(A,A)))$.

Ce qui nous sauve est que l'espace $\mathcal{L}(A,V)$ est un A-module, quel que soit V. En particulier lorsque V est lui-même déjà un emboîtement. Faisons un lemme pour voir comment ça fonctionne.

LEMooCEVGooFVXndZ

Lemme 15.62 ([1])**.**
Si $A\in\mathbb{M}(n,\mathbb{K})$ avec $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$, nous considérons l'algèbre $\mathrm{Alg}(A)$ *engendrée par A*[27]*. Alors*

$$e^{\mathrm{Alg}(A)}\subset\mathrm{Alg}(A). \tag{15.186}$$

Démonstration. Le lemme 15.58 nous dit que $\mathrm{Alg}(A)$ est une algèbre de dimension finie. Elle est commutative par le corolaire 15.59. Donc la proposition 15.60 s'applique. □

26. N'oubliez pas de faire à part le cas $k=0$ parce que ce qui suit n'est correct que pour $k\geqslant 1$.
27. Définition 15.57.

15.5.2 Différentielles

LEMooHKQVooQTrHHW

Lemme 15.63 ([1]).
Soient deux espaces vectoriels normés E et V tels que V soit un E-module[28]*. Nous supposons les normes soient telles que $\|xv\|_V \leqslant \|x\|_E\|v\|_V$.*

Soit une fonction différentiable $f\colon E \to V$ telle que la différentielle $df\colon E \to \mathcal{L}(E,V)$ soit de la forme

$$df_x(y) = yg(x) \tag{15.187}$$

pour une certaine fonction différentiable $g\colon E \to V$.

Alors f est C^1, et deux fois différentiable telle que

$$\begin{aligned} d^2f\colon E &\to \mathcal{L}\big(E,\mathcal{L}(E,V)\big) \\ (d^2f)_x(y)z &= z(dg_x)(y) \end{aligned} \tag{15.188}$$

pour tout $x,y,z \in E$.

Démonstration. En plusieurs étapes.

(1) **f est C^1** Nous savons, par hypothèse, que f est différentiable. Il faut montrer que sa différentielle est continue, en remarquant déjà que g est continue parce que différentiable.
Soit $x_k \xrightarrow{E} x$, et calculons $\|df_{x_k} - df_x\|$:

$$\begin{aligned} \|df_{x_k} - df_x\| &= \sup_{\|y\|=1} \|df_{x_k}(y) - df_x(y)\| \\ &= \sup_{\|y\|=1} \|\big(g(x_k) - g(x)\big)y\| \\ &\leqslant \sup_{\|y\|=1} \|g(x_k) - g(x)\|\|y\| \\ &= \|g(x_k) - g(x)\|. \end{aligned} \tag{15.189}$$

Donc nous avons bien $df_{x_k} \xrightarrow{\mathcal{L}(E,V)} df_x$, ce qui signifie la continuité de df. Donc f est de classe C^1.

(2) **f est deux fois différentiable**
Pour montrer que df est différentiable, nous mettons directement dans la définition (10.243) le candidat

$$\begin{aligned} T_x(h)\colon R &\to V \\ T_x(h)z &= zdg_x(y). \end{aligned} \tag{15.190}$$

Nous devons vérifier la limite suivante :

$$\lim_{h \xrightarrow{E} 0} \frac{df_{x+h} - df_x - T_x(h)}{\|h\|} = 0. \tag{15.191}$$

EQooTBCKooRxBCun

Étudions la norme du numérateur :

$$\|df_{x+h} - df_x - T_x(h)\| = \sup_{\|y\|=1} \|df_{x+h}(y) - df_x(y) - T_x(h)y\| \tag{15.192a}$$

$$= \sup_{\|y\|=1} \|yg(x+h) - yg(x) - ydg_x(h)\| \tag{15.192b}$$

$$\leqslant \sup_{\|y\|=1} \|y\|\|g(x+h) - g(x) - dg_x(h)\|. \tag{15.192c}$$

La limite (15.191) se déduit donc de la différentiabilité de g.

Note : la partie démontrant que f est C^1 n'est pas strictement obligatoire parce qu'en vérifiant que f est deux fois différentiable, nous vérifions de facto que df est en particulier continue. □

28. Définition 1.330.

LEMooTUWQooMCCDcm

Lemme 15.64 ([1]).
Soient des algèbres normées A et V telles que V soit un A-module vérifiant $\|xv\| \leqslant \|x\|\|v\|$ pour tout $x \in A$ et $v \in V$. Alors $\mathcal{L}(A,V)$ est un A-module vérifiant $\|x\alpha\| \leqslant \|x\|\|\alpha\|$ pour tout $x \in A$ et $\alpha \in \mathcal{L}(A,V)$.

Démonstration. C'est un simple calcul utilisant la norme opérateur :

$$\|x\alpha\| = \sup_{\|y\|=1} \|(x\alpha)y\| = \sup_{\|y\|=1} \|x\alpha(y)\| \leqslant \sup_{\|y\|=1} \|x\|\|\alpha(y)\| = \|x\| \sup_{\|y\|=1} \|\alpha(y)\| = \|x\|\|\alpha\|. \tag{15.193}$$

□

PROPooTBDAooQouzSk

Proposition 15.65 ([1]).
La fonction $\exp\colon A \to A$ *est de classe* C^∞ *et vérifie, pour tout* $k \geqslant 1$ *la récurrence*

$$(d^k \exp)_x(y) = y(d^{k-1}\exp)_x. \tag{15.194}$$

Démonstration. La formule proposée fonctionne avec $k = 1$:

$$(d\exp)_x(y) = y\exp(x). \tag{15.195}$$

C'est la relation 15.173.

Nous considérons $k > 1$, nous supposons que exp est de classe C^{k-1} et k fois différentiable. Nous allons prouver que exp est alors de classe C^k et $k+1$ fois différentiable, et que la différentielle de $d^k\exp$ est donné par la formule

$$(d^{k+1}\exp)_x(y) = y(d^k\exp)_x. \tag{15.196}$$

Pour nous mettre au clair avec les espaces en présence, nous supposons que

$$d^{k-1}\exp\colon A \to \mathcal{L}(A,V) \tag{15.197a}$$

$$d^k\exp\colon A \to \mathcal{L}\big(A,\mathcal{L}(A,V)\big) \tag{15.197b}$$

pour un certain espace vectoriel normé V, lequel est un de ces terrifiants emboîtement de type $\mathcal{L}\Big(A,\mathcal{L}\big(A,\mathcal{L}(A,A)\big)\Big)$. Il est bien un espace vectoriel normé, et également un A-module parce qu'on peut toujours définir la multiplication d'un élément $v \in V$ par un élément $x \in A$ comme étant la multiplication par x du résultat final de l'évaluation emboîtée, laquelle se termine par un élément de A. Donc tout se met bien.

Quoi qu'il en soit, nous posons

$$T_x(y) = y(d^k\exp)_x \tag{15.198}$$

et nous vérifions ce que cela donne dans la définition de la différentielle. Si nous avons

$$\lim_{h\to 0} \frac{(d^k\exp)_{x+h} - (d^k\exp)_x - T_x(h)}{\|h\|} = 0 \tag{15.199}$$

alors nous aurons prouvé tout ce qu'il nous faut.

Le numérateur est une application $A \to \mathcal{L}(A,V)$; nous en écrivons la norme comme il se doit :

$$\|(d^k\exp)_{x+h} - (d^k\exp)_x - T_x(h)\| = \sup_{\|y\|=1} \|(d^k\exp)_{x+h}(y) - (d^k\exp)_x(y) - h(d^k\exp)_x y\| \tag{15.200a}$$

$$= \sup_{\|y\|=1} \|y(d^{k-1}\exp)_{x+h} - y(d^{k-1}\exp)_x - h(d^k\exp)_x y\| \tag{15.200b}$$

$$= \sup_{\|y\|=1} \|y(d^{k-1}\exp)_{x+h} - y(d^{k-1}\exp)_x - hy(d^{k-1}\exp)_x\| \tag{15.200c}$$

$$\leqslant \|(d^{k-1}\exp)_{x+h} - (d^{k-1}\exp)_x - h(d^{k-1}\exp)_x\| \tag{15.200d}$$

$$= \|(d^{k-1}\exp)_{x+h} - (d^{k-1}\exp)_x - (d^k\exp)_x(h)\|. \tag{15.200e}$$

Dans ce calcul nous avons utilisé le lemme 15.64 et $T_x(h)y = h(d^k \exp)_x y$. Maintenant, la limite

$$\lim_{h\to 0} \frac{\|(d^{k-1}\exp)_{x+h} - (d^{k-1}\exp)_x - (d^k\exp)_x(h)\|.}{\|h\|} \tag{15.201}$$

n'est rien d'autre que la limite arrivant dans la définition du fait que $d^k \exp$ est la différentielle de $d^{k-1}\exp$. Cette limite est donc zéro comme nous voulions le prouver. □

Le théorème suivant est très important parce qu'il permet de définir l'exponentielle d'une matrice. Et les exponentielles de matrices sont utiles, entre très nombreuses autres choses pour résoudre certaines équations différentielles.

THOooFGTQooZPiVLO

Théorème-Définition 15.66 ([1]).
Soit une algèbre normée A (pas spécialement commutative). La formule

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{15.202}$$

définit une fonction différentiable dont la différentielle est donnée par[29]

$$(d\exp)_x(y) = \sum_{i,j\in\mathbb{N}} \frac{x^i y x^j}{(i+j+1)!} \tag{15.203}$$

EQooFGPPooZKHeXU

15.67.
Nous ne démontrons pas cela ici.

Il s'agit d'une adaptation de la proposition 15.60. Là où il faut faire attention, c'est dans l'équation (15.179) : il n'y a pas k termes $x^{k-1}h$ dans $(x+h)^k$, mais k termes de la forme $x^i h x$. C'est pour cela que la différentielle n'est pas donnée par $T(y) = u_{k-1}(x)y$, mais bien par la somme (15.203).

M'est avis en réalité que toute la démonstration du théorème 15.151 passe facilement au cas présent.

15.5.3 Exponentielle de matrice

SECooBYQBooZifJsg

Proposition 15.68 ([490]).
Soient des matrices $A, P \in \mathbb{M}(n,\mathbb{C})$ telle que P soit inversible. Alors[30]

$$e^{P^{-1}AP} = P^{-1}e^A P. \tag{15.204}$$

Démonstration. Pour chaque $m \in \mathbb{N}$ nous avons $(P^{-1}AP)^m = P^{-1}A^m P$. Ensuite,

$$e^{P^{-1}AP} = \sum_k \frac{(P^{-1}AP)^k}{k!} = \sum_k \frac{P^{-1}A^kP}{k!} = P^{-1}\sum_k \frac{A^k}{k!}P. \tag{15.205}$$

Nous avons utilisé la proposition 10.106 pour sortir P^{-1} à gauche et P à droite de la somme. □

PROPooFLHPooRhLiZE

Proposition 15.69 ([490]).
L'exponentielle de matrice vérifie

ITEMooCVALooEfLQCyI

(1) $e^0 = \mathrm{Id}$

ITEMooNGPWooIyPEQt

(2) $A^m e^A = e^A A^m$

ITEMooEOSMooQWjcjA

(3) $\exp(A)^t = \exp(A^t)$

29. La fonction exponentielle est, j'en suis quasiment certain, de classe C^∞. Si vous connaissez un moyen pas trop douloureux de prouver cela, faites-le moi savoir.

30. La définition de l'exponentielle de matrice est 15.60 où la convergence de la somme est celle de la norme opérateur 10.55.

ITEMooROPJooMarenu

(4) *Si $AB = BA$ alors $Ae^B = e^B A$ et $e^A e^B = e^B e^A$.*

Démonstration. Point par point.

(1) **Pour (1)** Juste substituer $A = 0$ dans la définition. Tous les termes tombent sauf le premier. Il faut utiliser le fait que $A^0 = \mathrm{Id}$.

(2) **Pour (2)** Il faut utiliser la proposition 10.106 pour écrire

$$A^m \sum_k \frac{A^k}{k!} = \sum_k \frac{A^m A^k}{k!} = \sum_k \frac{A^k A^m}{k!} = \sum_k \frac{A^k}{k!} A^m. \tag{15.206}$$

EQooLUUVooCtUtIC

(3) **Pour (3)** Pour chaque k nous avons l'égalité $(A^k)^t = (A^t)^k$. En utilisant encore le coup de la queue de suite qui converge vers zéro,

$$\| \sum_{k=0}^{N} \frac{(A^t)^k}{k!} - (e^A)^t \| = \| \sum_{k=N+1}^{\infty} \frac{(A^t)^k}{k!} \| \to 0. \tag{15.207}$$

(4) **Pour (4)** Pour prouver $Ae^B = e^B A$, c'est le même genre de manipulations que (15.206). Maintenant, vu que A et e^B commutent, l'égalité à peine prouvée montre que e^A et e^B commutent.

□

PROPooKDKDooCUpGzE

Proposition 15.70 ([490]).
Soient $A \in \mathbb{M}(n, \mathbb{C})$ ainsi que $s, t \in \mathbb{C}$. Alors

$$e^{sA} e^{tA} = e^{(s+t)A}. \tag{15.208}$$

Démonstration. Nous calculons le produit $e^{sA} e^{tA}$ par le produit de Cauchy de la proposition 10.107 :

$$\clubsuit = \left(\sum_k \frac{t^k}{k!} A^k \right) \left(\sum_l \frac{s^l}{l!} A^l \right) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{t^m}{m!} \frac{s^{n-m}}{(n-m)!} A^n. \tag{15.209}$$

À ce point, nous multiplions et divisons par $n!$ et nous réarrangons la somme de la façon suivante :

$$\clubsuit = \sum_{n=0}^{\infty} \frac{A^n}{n!} \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} t^m s^{n-m}. \tag{15.210}$$

Nous reconnaissons la somme sur m comme étant un binôme de Newton[31] pour $(t+s)^n$. Nous avons donc finalement

$$\clubsuit = \sum_{n=0}^{\infty} \frac{\big((t+s)A\big)^n}{n!} = e^{(t+s)A}. \tag{15.211}$$

□

La proposition suivante dit que les exponentielles de matrices sont inversibles. Elle ne dit pas que toutes les matrices inversibles sont des exponentielles. Ce sera la proposition 15.131.

PROPooRERRooMutKcg

Proposition 15.71 ([490]).
Si $A \in \mathbb{M}(n, \mathbb{C})$, alors e^A est inversible et

$$(e^A)^{-1} = e^{-A}. \tag{15.212}$$

31. Proposition 3.43.

Démonstration. Il suffit de prendre $s = 1$ et $t = -1$ dans la proposition 15.70 et nous avons

$$e^A e^{-A} = e^0 = \mathbb{1}. \tag{15.213}$$

Cela prouve que e^A est inversible et que son inverse est e^{-A}. □

PROPooSDNNooQtHkhA

Proposition 15.72 ([490, 1])**.**
Soit $A \in \mathbb{M}(n, \mathbb{C})$. Nous considérons l'application

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{M}(n, \mathbb{C}) \\ t &\mapsto e^{tA}. \end{aligned} \tag{15.214}$$

Nous avons la formule de dérivation

$$\varphi'(t) = Ae^{tA}. \tag{15.215}$$

Démonstration. L'application φ est une fonction composée de

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{M}(n, \mathbb{R}) \\ t &\mapsto tA \end{aligned} \tag{15.216}$$

et

$$\begin{aligned} \exp\colon \mathbb{M}(n, \mathbb{R}) &\to \mathbb{M}(n, \mathbb{R}) \\ A &\mapsto e^A. \end{aligned} \tag{15.217}$$

Et nous avons $\varphi = \exp \circ f$. Il y a un choix difficile à faire. Soit nous travaillons dans $\mathbb{M}(n, \mathbb{R})$ et nous allons devoir invoquer le théorème 15.66, soit nous travaillons dans l'algèbre $\mathrm{Alg}(A)$ engendrée par A (définition 15.57) qui est une algèbre commutative [32] qui nous permet de n'utiliser que 15.173, qui est quand même plus basique.

Coup de théâtre, nous prenons la seconde solution et nous réécrivons les fonctions f et exp de la façon suivante :

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathrm{Alg}(A) \\ t &\mapsto e^{tA} \end{aligned} \tag{15.218}$$

se décompose en

$$\begin{aligned} f\colon \mathbb{R} &\to \mathrm{Alg}(A) \\ t &\mapsto tA \end{aligned} \tag{15.219}$$

et

$$\begin{aligned} \exp\colon \mathrm{Alg}(A) &\to \mathrm{Alg}(A) \\ A &\mapsto \exp(A). \end{aligned} \tag{15.220}$$

Le fait que ces applications soient bien définies est le lemme 15.62 qui assure que les espaces d'arrivée sont bien dans $\mathrm{Alg}(A)$. Et c'est parti pour le calcul :

$$\begin{aligned} \varphi'(u) &= d\varphi_u(1) && \text{SUBEQooFDPQooBgTDXF (15.221a)} \\ &= d\exp_{f(u)}\big(df_u(1)\big) && \text{SUBEQooVOFWooDoTQGy (15.221b)} \\ &= \exp\big(f(u)\big)df_u(1) && \text{SUBEQooCILLooHsEDOE (15.221c)} \\ &= e^{uA}A && \text{SUBEQooBMAQooNbzBI (15.221d)} \end{aligned}$$

Justifications :

— Pour 15.221a. Pour la dérivée, nous utilisons le corolaire 12.253.

— Pour 15.221b. La règle de la différentielle en chaine du théorème 10.276.

— Pour 15.221c. La formule (15.173).

— Pour 15.221d. Parce que $f(u) = uA$ et que $df_u(1) = A$.

32. Corolaire 15.59.

□

Le théorème suivant montre que le produit d'exponentielle de matrices suit la règle usuelle tant que les matrices commutent. Cela est cependant plutôt l'exception que la règle. À priori nous avons $e^A e^B \neq e^{A+B}$.

THOooXCPEooYGyLOp

Théorème 15.73 ([490]).
Soient $A, B \in \mathbb{M}(n, \mathbb{C})$ telles que $AB = BA$. Alors

$$e^{A+B} = e^A e^B. \tag{15.222}$$

Démonstration. Vu que A et B commutent nous avons $Ae^{tB} = e^{tB}A$ (proposition 15.69(4)). Ensuite nous posons

$$g(t) = e^{t(A+B)}e^{-tB}e^{-tA}. \tag{15.223}$$

Nous calculons la dérivée de g en utilisant la règle de Leibniz et la proposition 15.72 :

$$\begin{aligned} g'(t) = &(A+B)e^{t(A+B)}e^{-tB}e^{-tA} \\ &+ e^{t(A+B)}(-B)e^{-tB}e^{-tA} \\ &+ e^{t(A+B)}e^{-tB}(-A)e^{-tA}. \end{aligned} \tag{15.224}$$

Vu que A, B et $A+B$ commutent, nous pouvons réarranger les facteurs en

$$\begin{aligned} g'(t) = &(A+B)e^{t(A+B)}e^{-tB}e^{-tA} \\ &- Be^{t(A+B)}e^{-tB}e^{-tA} \\ &- Ae^{t(A+B)}e^{-tB}e^{-tA}. \end{aligned} \tag{15.225}$$

Enfin, cela fait

$$g'(t) = (A+B-B-A)e^{t(A+B)}e^{-tB}e^{-tA} = 0. \tag{15.226}$$

Donc g est constante et nous avons

$$e^{t(A+B)}e^{-tB}e^{-tA} = g(0) = \mathbb{1}. \tag{15.227}$$

En multipliant à droite par $e^{tA}e^{tB}$ nous trouvons

$$e^{t(A+B)} = e^{tA}e^{tB} \tag{15.228}$$

comme annoncé. □

15.6 Exponentielle et logarithme dans les réels

Pour avoir une vue synthétique du plan, voir le thème 50.

15.6.1 L'équation différentielle

Pour la suite nous notons y une solution de l'équation $y' = y$, $y(0) = 1$, et nous allons en donner des propriétés indépendamment de l'existence, donnée par le théorème 15.76.

PropTLECooEiLbPP

Proposition 15.74.
Quelques propriétés de y (si elle existe) :

(1) Pour tout $x \in \mathbb{R}$ nous avons $y(x)y(-x) = 1$.

(2) $y(x) > 0$ pour tout x.

(3) y est strictement croissante.

Démonstration. Nous posons $\varphi(x) = y(x)y(-x)$ et nous dérivons :

$$\varphi'(x) = y'(x)y(-x) - y(x)y'(-x) = 0. \tag{15.229}$$

Donc φ est constante[33]. Vu que $\varphi(0) = 1$ nous avons automatiquement $y(x)y(-x) = 1$ pour tout x.

Les deux autres allégations sont simples : si $y(x_0) < 0$ alors il existe $t \in]0, x_0[$ tel que $y(t) = 0$, ce qui est impossible parce que $y(t)y(-t) = 1$. La stricte croissance de y s'ensuit. □

PROPooGGUIooExVHPM

Proposition 15.75.
Quelques formules pour tout $a, b \in \mathbb{R}$ et $n \in \mathbb{Z}$:

ITEMooMPSUooWQpVQJ

(1) $y(a + b) = y(a)y(b)$
(2) $y(na) = y(a)^n$
(3) $y\left(\frac{a}{n}\right) = \sqrt[n]{y(a)}$.

Démonstration. Nous posons $h(x) = y(a + b - x)y(x)$ et nous avons encore $h'(x) = 0$ dont nous déduisons que h est constante. De plus

$$h(0) = y(a + b)y(0) = y(a + b) \tag{15.230}$$

et

$$h(b) = y(a)y(b). \tag{15.231}$$

Vu que h est constante, ces deux expressions sont égales : $y(a + b) = y(a)y(b)$.

Forts de cette relation, une récurrence donne $y(na) = y(a)^n$ pour tout $n \in \mathbb{N}$. De plus

$$y(a) = y\left(\frac{a}{n} \times n\right) = y\left(\frac{a}{n}\right)^n, \tag{15.232}$$

ce qui donne $y(a) = y(a/n)^n$ ou encore $y(a/n) = \sqrt[n]{y(a)}$.

Enfin pour les négatifs, si $n \in \mathbb{N}$,

$$y(-na) = \frac{1}{y(na)} = \frac{1}{y(a)^n} = y(a)^{-n}. \tag{15.233}$$

Et de la même façon,

$$y\left(-\frac{a}{n}\right) = \frac{1}{y\left(\frac{a}{n}\right)} = \sqrt[n]{\frac{1}{y(a)}} = \sqrt[-n]{y(a)}. \tag{15.234}$$

□

15.6.2 Existence

Jusqu'ici nous avons donné des propriétés d'une éventuelle fonction y qui vérifierait l'équation différentielle. Il est temps de montrer qu'une telle fonction existe.

ThoKRYAooAcnTut

Théorème 15.76.
La série entière

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{15.235}$$

EqEIGZooKWSyPS

définit une fonction dérivable solution de

$$\begin{cases} y' = y & \text{(15.236a)} \\ y(0) = 1. & \text{(15.236b)} \end{cases}$$

EQooSEIHooNmQKiC

33. Proposition 12.185.

Démonstration. La formule de Hadamard (théorème 15.25) donne le rayon de convergence de la série (15.235) par

$$\frac{1}{R} = \lim_{k\to\infty} \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} = \lim_{k\to\infty} \frac{1}{k+1} = 0. \tag{15.237}$$

Donc nous avons un rayon de convergence infini. La fonction y est définie sur $\mathbb{R}$ et la proposition 15.44 nous dit que y est dérivable. Nous pouvons aussi dériver terme à terme :

$$y'(x) = \sum_{k=0}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = y(x). \tag{15.238}$$

Notez le petit jeu d'indice de départ de k. Dans un premier temps, nous remarquons que $k = 0$ donne un terme nul et nous le supprimons, et dans un second temps nous effectuons la simplification des factorielles (qui ne fonctionne pas avec $k = 0$). □

15.77.
Nous savons que la fonction y existe parce qu'une solution de l'équation différentielle $y' = y$, $y(0) = 1$ est donnée par la fameuse série (théorème 15.76). À part cela, ce qui a été fait avec cette équation différentielle ne permet pas de prouver l'existence de y. Donc, du point de vue de « définir l'exponentielle par son équation différentielle », c'est pas encore gagné. Notons au passage que le nombre e n'est pas encore bien défini via l'équation différentielle.

15.6.3 Le nombre de Neper e

Nous savons par le théorème 15.76 que $x \mapsto \exp(x)$ est une solution de l'équation différentielle exponentielle (avec la bonne condition initiale). Or une telle solution est unique par la proposition 12.430.

Définition 15.78 (Le nombre de Neper)**.**
Nous notons e le nombre $\exp(1)$.

PropCELWooLBSYmS

Proposition 15.79.
Pour tout $x \in \mathbb{R}$, nous avons

$$\exp(x) = e^x. \tag{15.239}$$

EQooBFIHooKopcmf

Démonstration. Soit y vérifiant la fameuse équation différentielle. Nous savons que $y = \exp$ parce que c'est l'unique solution (proposition 12.430). Nous avons :

$$y(x) = y(1)^x. \tag{15.240}$$

Si $q \in \mathbb{Q}$ alors $q = a/b$ et

$$y(q) = y\left(\frac{a}{b}\right) = y\left(a \times \frac{1}{b}\right) = y\left(\frac{1}{b}\right)^a = \left(\sqrt[b]{y(1)}\right)^a = y(1)^{a/b} = y(1)^q. \tag{15.241}$$

Le résultat est prouvé pour les rationnels.

En ce qui concerne un élément général $x \in \mathbb{R}$, la fonction $x \mapsto y(x)$ est continue sur $\mathbb{R}$, et la fonction $x \mapsto e^x$ également (proposition 12.397). Ces deux fonctions étant égales sur $\mathbb{Q}$, elles sont égales sur $\mathbb{R}$ par la proposition 7.256). □

PROPooIKJBooLOipUM

Proposition 15.80.
Pour tout $\lambda \in \mathbb{R}$, l'équation différentielle

$$\begin{cases} y' = \lambda y & \text{(15.242a)} \\ y(0) = 1 & \text{(15.242b)} \end{cases}$$

a une unique solution donnée par

$$y(t) = \exp(\lambda t) = e^{\lambda t} \tag{15.243}$$

Une conséquence des propositions 15.79 et 12.392 est que

EqLOIUooHxnEDn

$$\lim_{x\to-\infty} e^x = 0 \tag{15.244a}$$
$$\lim_{x\to+\infty} e^x = +\infty, \tag{15.244b}$$

et en particulier,

$$\begin{aligned}\exp\colon \mathbb{R} &\to]0,\infty[\\ x &\mapsto e^x\end{aligned} \tag{15.245}$$

est une bijection.

LEMooXFAXooLVbebl

Lemme 15.81.
Nous avons l'encadrement $2.5 < e < 2.8$.

Démonstration. Vu que tous les termes de la série exponentielle sont strictement positifs, nous avons

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} > \frac{1}{1} + \frac{1}{1} + \frac{1}{2} = 2.5. \tag{15.246}$$

Pour la majoration, nous notons que $k! > 2^k$ à partir de $k = 4$. Donc nous gardons les termes $k = 0, 1, 2, 3$ sous la forme $1/k!$ et nous remplaçons les autres par $1/2^n$:

$$e < \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \sum_{k=4}^{\infty} \frac{1}{2^n}. \tag{15.247}$$

Nous pouvons calculer explicitement la somme en utilisant la série géométrique de la proposition 10.135(3) :

$$\sum_{k=4}^{\infty} \frac{1}{2^n} = 1 - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) = \frac{1}{8}. \tag{15.248}$$

En remettant tout ensemble,

$$e < 2 + \frac{2}{3} + \frac{1}{8} = \frac{67}{24} < 2.8. \tag{15.249}$$

□

PROPooFRKUooZyhHIC

Proposition 15.82 ([491])**.**
Le nombre e est irrationnel.

Démonstration. En vertu de la proposition 15.79, nous avons $e^x = \exp(x)$ pour tout réel x. En particulier pour $x = 1$,

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}. \tag{15.250}$$

Supposons que $e = p/q$ pour certains $p, q \in \mathbb{N}$ avec $q \neq 0$ (notez que ça signifie $q \geqslant 1$).

Nous avons

$$q!e = \sum_{k=0}^{q} \frac{q!}{k!} + \sum_{k=q+1}^{\infty} \frac{q!}{k!}. \tag{15.251}$$

Dans le membre de gauche,

$$q!e = q!\frac{p}{q} = (q-1)!p \in \mathbb{N}. \tag{15.252}$$

Dans la première somme du membre de droite, vu que $q \geqslant k$, nous avons

$$\sum_{k=0}^{q} \frac{q!}{k!} \in \mathbb{N}. \tag{15.253}$$

Nous devons donc avoir $\sum_{k=q+1}^{\infty} \frac{q!}{k!} \in \mathbb{N}$. Nous allons maintenant prouver que ce n'est pas le cas. Nous avons :

$$\sum_{k=1}^{\infty} \frac{q!}{(q+k)!} = \sum_{k=1}^{\infty} \frac{1}{\prod_{l=q+1}^{q+k} l} \tag{15.254a}$$
$$= \frac{1}{q+1} + \sum_{k=2}^{\infty} \frac{1}{\prod_{l=q+1}^{q+k} l} \tag{15.254b}$$
$$= \frac{1}{q+1} \left[1 + \sum_{k=2}^{\infty} \frac{1}{\prod_{l=q+2}^{q+k} l} \right] \tag{15.254c}$$
$$< \frac{1}{q+1} \left[1 + \sum_{k=2}^{\infty} \frac{1}{(q+1)^{(q+k)-(q+2)+1}} \right] \tag{15.254d}$$
$$= \frac{1}{q+1} \left[1 + \sum_{k=2}^{\infty} \frac{1}{(q+1)^{k-1}} \right] \tag{15.254e}$$
$$= \frac{1}{q+1} \left[1 + \sum_{k=1}^{\infty} \frac{1}{(q+1)^{k}} \right] \tag{15.254f}$$
$$= \frac{1}{q+1} \sum_{k=0}^{\infty} \frac{1}{(q+1)^k}. \tag{15.254g}$$

C'est le moment d'utiliser la série géométrique de la proposition 10.135(2). Vu que $q \geqslant 1$ nous avons $1/(q+1) < 1$ et la somme fonctionne :

$$\sum_{k=1}^{\infty} \frac{q!}{(q+k)!} < \frac{1}{q+1} \left(\frac{1}{1 - \frac{1}{q+1}} \right) = \frac{1}{q}. \tag{15.255}$$

Nous avons donc

$$0 < \sum_{k=1}^{\infty} \frac{q!}{(q+k)!} < \frac{1}{q} < 1. \tag{15.256}$$

Donc ce terme n'est pas un nombre entier, ce que nous avions énoncé. □

15.6.4 Application réciproque : logarithme

DEFooELGOooGiZQjt

Proposition-Définition 15.83.
L'application $\exp\colon \mathbb{R} \to]0, \infty[$ *est une bijection. L'application réciproque*

$$\ln\colon]0, \infty[\to \mathbb{R} \tag{15.257}$$

est le ***logarithme****.*

Démonstration. Le fonction exponentielle est dérivable, toujours strictement positive, donc strictement croissante. Les limites en $\pm\infty$ sont 0 et $+\infty$. Le théorème des valeurs intermédiaires 9.93 nous dit que c'est une bijection. En effet, l'injectivité est la stricte croissance. En ce qui concerne la surjection, soit $y \in]0, \infty[$. Vu que la limite en $-\infty$ est zéro, il existe $A \in \mathbb{R}$ tel que $\exp(x) < y$ pour tout $x < A$, et de la même façon, il existe $B \in \mathbb{R}$ tel que $\exp(x) > y$ pour tout $x > B$. Si $a < A$ et $b > B$ alors $\exp(a) < y$ et $\exp(b) > y$, donc y est dans l'image de $[a, b]$ par l'exponentielle. □

LEMooCYGTooEjXEUu

Lemme 15.84 ([1])**.**
Le logarithme est une fonction continue.

Démonstration. C'est une conséquence du théorème de la bijection 12.51(4), et de la continuité de l'exponentielle sur $\mathbb{R}$, qui est une partie du théorème 15.76. □

PROPooPDJLooXphpEM

Proposition 15.85 (Dérivée du logarithme)**.**
Pour tout $s \in \mathbb{R}^+$ nous avons

$$\ln'(s) = \frac{1}{s}. \tag{15.258}$$

Démonstration. L'application $\exp\colon \mathbb{R} \to]0,\infty[$ est une bijection continue dérivable[34]. La proposition 12.178 s'applique donc et pour tout $x \in \mathbb{R}$ nous avons

$$\ln'\big(\exp(x)\big) = \frac{1}{\exp'(x)}. \tag{15.259}$$

Le théorème 15.76 nous indique que $\exp'(x) = \exp(x)$. Donc pour tout x nous avons $\ln'\big(\exp(x)\big) = \frac{1}{\exp(x)}$. Vu que x est arbitraire et que exp est surjective sur $\mathbb{R}^+$, pour tout $s \in]0,\infty[$ nous avons

$$\ln'(s) = \frac{1}{s}. \tag{15.260}$$

□

PROPooLAOWooEYvXmI

Proposition 15.86 ([1])**.**
Pour tout $x, y \in \mathbb{R}$ et pour $a > 0$ nous avons

$$\ln(\frac{1}{x}) = -\ln(x), \tag{15.261}$$

et

$$\ln(xy) = \ln(x) + \ln(y), \tag{15.262}$$

EQooJVMUooVpIJKyo

et

$$\ln(a^x) = x\ln(a) \tag{15.263}$$

EQooEJQSooWCczXy

et

$$a^x = e^{x\ln(a)}. \tag{15.264}$$

Démonstration. Nous avons, par la proposition 12.398,

$$e^{-\ln(x)} = \frac{1}{e^{\ln(x)}} = \frac{1}{x}. \tag{15.265}$$

En prenant le logarithme des deux côtés nous trouvons

$$-\ln(x) = \ln\left(\frac{1}{x}\right). \tag{15.266}$$

Nous pouvons continuer avec la suivante.

Par définition, $\ln(xy)$ est donné par $\exp\big(\ln(xy)\big) = xy$. Mais nous avons aussi, par la proposition 12.398 :

$$e^{\ln(x)+\ln(y)} = e^{\ln(x)}e^{\ln(y)} = xy. \tag{15.267}$$

Nous avons donc démontré (15.262).

La relation (15.263) de démontre d'abord pour $x \in \mathbb{N}$, puis pour $x \in \mathbb{Q}$ et enfin pour $x \in \mathbb{R}$. Si $n \in \mathbb{N}$ alors la relation (15.262) donne immédiatement

$$\ln(a^n) = n\ln(a). \tag{15.268}$$

pour tout $a \in \mathbb{R}$.

Si $m, n \in \mathbb{N}$, le nombre $a^{n/m}$ est par définition le $x > 0$ tel que

$$x^m = a^n. \tag{15.269}$$

34. Dérivable (et donc continue) par le théorème 15.76. Bijection par la proposition 15.83.

En prenant le logarithme des deux côtés : $\ln(x^m) = \ln(a^n)$ et en utilisant la relation déjà démontrée pour $\mathbb{N}$ nous trouvons $m\ln(x) = n\ln(a)$ et donc

$$\ln(a^{m/n}) = \ln(x) = \frac{m}{n}\ln(a). \tag{15.270}$$

La relation est donc démontré pour $\ln(a^q)$ avec $q \in \mathbb{Q}^+$.

Nous passons à $q = -m/n \in \mathbb{Q}^-$, c'est-à-dire toujours $m, n \in \mathbb{N}$. Nous avons, en utilisant la proposition 15.86,

$$\ln(a^{-q}) = \ln(\frac{1}{a^q}) = -\ln(a^q) = -q\ln(a). \tag{15.271}$$

Enfin si $x \in \mathbb{R}$ nous considérons une suite de rationnels $x_k \to x$. Pour chaque k nous avons

$$\ln(a^{x_k}) = x_k \ln(a). \tag{15.272}$$

Nous prenons la limite deux deux côtés. À droite nous avons tout de suite $x\ln(a)$, et à gauche, par continuité de la fonction ln (lemme 15.84) et de la fonction puissance (définition 12.397) nous trouvons $\ln(a^x)$. □

LemPEYJooEZlueU

Lemme 15.87.
Si $a, b \in]0, \infty[$ alors

$$\ln(ab) = \ln(a) + \ln(b) \tag{15.273}$$

et

$$\ln\left(\frac{1}{b}\right) = -\ln(b). \tag{15.274}$$

EqOOZGooQWkGlA

Démonstration. Nous posons $f(x) = \ln(ax)$ qui est une fonction dérivable [35]. Alors $f'(x) = \frac{a}{ax} = \frac{1}{x}$. Cette fonction f est donc une primitive de $\frac{1}{x}$ et il existe une constante K telle que

$$f(x) = \ln(x) + K. \tag{15.275}$$

Vu que $\ln(1) = 0$ nous avons $K = f(1) = \ln(a)$. Donc

$$\ln(ax) = \ln(x) + \ln(a). \tag{15.276}$$

En ce qui concerne la seconde formule à démontrer, nous avons

$$\ln(1) = \ln\left(\frac{1}{b}b\right) = \ln\left(\frac{1}{b}\right) + \ln(b). \tag{15.277}$$

Étant donné que $\ln(1) = 0$ nous en déduisons la formule (15.274). □

15.88.
La formule (15.263) en particulier est pratique pour réexprimer des fonctions puissances compliquées en écrivant

$$a^x = e^{\ln(a^x)} = e^{x\ln(a)}. \tag{15.278}$$

EQooYEWCooKyrayP

Cela aide à calculer la dérivée de $x \mapsto a^x$.

Notons que certains prennent (15.278) comme définition de la fonction puissance.

35. Dérivée du logarithme, proposition 15.85.

15.6.5 Approximations numériques de e

Nous donnons maintenant quelques approximations numériques de e, particulièrement inefficaces.

Lemme 15.89.
Nous avons

$$2 < e < 3. \tag{15.279}$$

Démonstration. Nous savons que $y(0) = 1$ et $y'(0) = 1$. La fonction y est strictement croissante (et donc sa dérivée aussi). Nous avons donc $y'(x) > 1$ pour tout $x \in]0, 1]$, et donc

$$y(1) > 1 + 1 \times 1 = 2. \tag{15.280}$$

Sachant que $2 > y'(x)$ pour tout $x \in]0, 1[$ nous pouvons refaire le coup de l'approximation affine, cette fois en majorant :

$$y(1) < 1 + 2 \times 1 = 3. \tag{15.281}$$

□

De la même façon nous savons que

$$y(\frac{1}{n}) > 1 + \frac{1}{n} \tag{15.282}$$

parce que y' est minoré par 1 sur $]0, \frac{1}{n}[$. Avec cela nous avons aussi la majoration

$$y(\frac{1}{n}) < 1 + \frac{1}{n} \times \left(1 + \frac{1}{n}\right) = 1 + \frac{1}{n} + \frac{1}{n^2}. \tag{15.283}$$

Et enfin nous pouvons donner l'encadrement, valable pour tout n :

$$\left(1 + \frac{1}{n}\right)^n < y(1) < \left(1 + \frac{1}{n} + \frac{1}{n^2}\right)^n. \tag{15.284}$$

Pour $n = 10$ nous trouvons

$$2.50 < e < 2.83. \tag{15.285}$$

Bien que ce soit à mon avis humainement pas possible à faire à la main nous avons, pour $n = 100$:

$$2.70 < e < 2.7317 \tag{15.286}$$

Cela reste un encadrement très modeste.

Une méthode plus efficace consiste à calculer directement le développement de définition

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{n!}. \tag{15.287}$$

```
def u(k):
    """
    return the kth term in the expansion of 'e'
    """
    return 1/factorial(k)

def sum_u(n):
    """
    return the sum of the 'n' first terms, that is with
    k froim 0 to n-1.
    """
```

```
    L=[  u(k) for k in range(0,n)  ]
    return sum(L)

s = sum_u(5)          # This is a fraction
print(s)
print( numerical_approx(s)  )
```

tex/sage/sageSnip013.sage

¡¡ Avertissement/question à la lectrice !! 15.90
Comment trouver, avec cette méthode, un encadrement pour e ?

Ce petit programme, avec 5 termes donne $e \simeq 65/24 \simeq 2.708$. Avouez que c'est déjà bien mieux.

15.6.6 Résumé des propriétés de l'exponentielle

ThoRWOZooYJOGgR

Théorème 15.91.
Les choses que nous savons sur l'exponentielle :

ITEMooEIKKooLNoaRD

(1) Il y a unicité de la solution à l'équation différentielle subeqBKJNooJQtbBD

$$\begin{cases} y' = y & \text{(15.288a)} \\ y(0) = 1. & \text{(15.288b)} \end{cases}$$

(2) L'équation différentielle (15.288) possède une solution donnée par la série entière

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{15.289}$$ EqUARSooKXnQxu

(3) Cette solution est une bijection $y : \mathbb{R} \to]0, \infty[$.

ItemYTLTooSnfhOu

(4) La fonction y ainsi définie est de classe C^∞.

(5) Elle est également donnée par la formule

$$\exp(x) = e^x \tag{15.290}$$

où e est défini par $e = \exp(1)$.

(6) Elle vérifie

$$e^{a+b} = e^a e^b \tag{15.291}$$ EQooVFXUooBfmjJY

Nous nommons **exponentielle** cette fonction.

Démonstration. Point par point.

(1) C'est la proposition 12.430.

(2) C'est le théorème 15.76.

(3) Le rayon de convergence de la série (15.289) est infini (théorème 15.76) ; elle est donc définie sur $\mathbb{R}$. Le fait que ce soit une bijection est dû au fait qu'elle est strictement croissante (proposition 15.74) ainsi qu'aux limites (15.244).

(4) Vu que $y = y'$, y est dérivable. Mais comme y' est alors égale à une fonction dérivable, y' est dérivable. En dérivant l'égalité $y' = y$ nous obtenons $y'' = y'$ et le jeu continue.

(5) C'est la proposition 15.79.

(6) C'est la proposition 15.75(1).

□

ExooLRHCooMYLQTU

Exemple 15.92 (Un endomorphisme sans polynôme annulateur[277])**.**
l'exponentielle permet de donner un exemple d'un endomorphisme n'ayant pas de polynôme annulateur [36] : l'endomorphisme de dérivation

$$\begin{aligned} D\colon C^\infty(\mathbb{R},\mathbb{R}) &\to C^\infty(\mathbb{R},\mathbb{R}) \\ f &\mapsto f' \end{aligned} \tag{15.292}$$

n'a pas de polynôme annulateur. En effet supposons que $P = \sum_{k=0}^{p} a_k X^k$ en soit un, et considérons les fonctions $f_\lambda\colon t \mapsto e^{\lambda t}$. Nous avons

$$0 = P(D)f_\lambda = \sum_k a_k D^k(f_\lambda) = \sum_k a_k \lambda^k f_\lambda = P(\lambda)f_\lambda. \tag{15.293}$$

Par conséquent λ est une racine de P pour tout $\lambda \in \mathbb{R}$. Cela implique que $P = 0$.

D'ailleurs si on y pense bien, cet exemple n'est qu'un habillage de l'exemple 8.101. △

ExZLMooMzYqfK

Proposition 15.93.
Quelques propriétés du logarithme.

(1) *Le logarithme est une application dérivable et strictement croissante.*
(2) *Le logarithme est la primitive de $x \mapsto \frac{1}{x}$ qui s'annule en $x = 1$.*

Démonstration. Elle est donc bijective, d'inverse continue et dérivable par le théorème 12.51 et la proposition 12.178.

La dérivée de la fonction logarithme peut être calculée en utilisant la formule (12.462), mais aussi de façon plus piettone en écrivant l'expression suivante, valable pour tout $x \in \mathbb{R}$:

$$\ln\big(\exp(x)\big) = x, \tag{15.294}$$

que nous pouvons dériver en utilisant le théorème de dérivation des fonctions composées :

$$\ln'\big(\exp(x)\big)\exp'(x) = 1. \tag{15.295}$$

Mais $\exp'(x) = \exp(x)$, donc

$$\ln'(y) = \frac{1}{y} \tag{15.296}$$

pour tout y dans l'image de exp, c'est-à-dire pour tout y dans l'ensemble de définition de ln.

Par ailleurs, $\exp(0) = 1$ donc

$$\ln(1) = \ln\big(\exp(0)\big) = 0. \tag{15.297}$$

En ce qui concerne l'unicité d'une primitive s'annulant en $x = 1$, c'est le corolaire 12.199. □

15.6.7 Dérivée de la fonction puissance

EXooGMRIooUucRez

Exemple 15.94.
Soit la fonction $f(x, y) = x^y$, définie en 12.397. Nous allons en calculer les dérivées partielles au point $(1, 2)$. Notons que f n'est pas définie pour $x < 0$, mais que cela n'a pas d'importance parce que nous pouvons nous restreindre à un voisinage du point $(1, 2)$. La première dérivée partielle est facile :

$$\partial_x f(1,2) = (yx^{y-1})_{(x,y)=(1,2)} = 2.$$

Pour la seconde, il faut utiliser les propriétés de l'exponentielle et du logarithme. D'abord le logarithme est par définition l'application réciproque de l'exponentielle (définition 15.83), donc

$$x^y = \exp\big(\ln(x^y)\big). \tag{15.298}$$

36. En dimension finie, le lemme 8.97 dit qu'il y en a toujours un.

Ensuite nous calculons en utilisant la proposition 15.86 :

$$\partial_y f(1,2) = \partial_y \left(e^{y \ln x}\right)_{(x,y)=(1,2)} = \left(\ln x e^{y \ln x}\right)_{(x,y)=(1,2)} = \ln(1) e^{2\ln(1)}. \tag{15.299}$$

△

Cet exemple est facilement généralisable aux fonctions de la forme $x \mapsto u(x)^{v(x)}$. Voici une proposition qui dit comment faire.

PROPooKUULooKSEULJ

Proposition 15.95 ([1])**.**
Soit une fonction dérivable $u\colon \mathbb{R} \to \mathbb{R}$ et $a > 0$. Nous avons

$$(a^u)' = u' \ln(a) a^u. \tag{15.300}$$

Si de plus $u(x) > 0$ pour tout x, nous avons

$$(u^a)' = a u' u^{a-1}. \tag{15.301}$$

Démonstration. Nous considérons la fonction $f(x) = a^{u(x)}$. Vu que $f(x) > 0$ pour tout x, nous pouvons en prendre le logarithme et écrire l'égalité, valable pour tout x :

$$f(x) = e^{\ln(a^{u(x)})} = \exp\big(u(x)\ln(a)\big). \tag{15.302}$$

Sachant la dérivée de l'exponentielle, cela n'est rien d'autre que la dérivée d'une fonction composée :

$$f'(x) = \ln(a) u'(x) e^{u(x)\ln(a)}. \tag{15.303}$$

Pour l'autre, nous posons

$$g(x) = u(x)^a, \tag{15.304}$$

qui peut encore s'écrire sous la forme

$$g(x) = e^{a \ln\left(u(x)\right)}. \tag{15.305}$$

Ici encore, c'est la dérivée de fonctions composées qui donne le résultat. □

15.6.8 Dérivée du logarithme

LEMooTGCBooJdkLpg

Lemme 15.96.
Si $u\colon \mathbb{R} \to]0,\infty[$ est dérivable alors $\ln(u)' = \dfrac{u'}{u}$.

Démonstration. Cela est une conséquence du théorème de dérivation des fonctions composées : si $g(x) = \ln(u(x))$ alors

$$g'(x) = \ln'\big(u(x)\big) u'(x) = \frac{1}{u(x)} u'(x). \tag{15.306}$$

□

15.6.9 Taylor pour l'exponentielle

PROPooQBRGooAhGrvP

Proposition 15.97 (Développement de l'exponentielle)**.**
Pour tout entier n, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + x^n \alpha(x). \tag{15.307}$$

Démonstration. Il s'agit de la proposition 15.51 appliquée à la série entière (15.60). □

15.6.10 Analycité

Vu que $\exp(x)$ est défini par une série entière (définition 15.60) et vu la proposition 15.54, il n'est pas étonnant que exp soit analytique. Traitons ce cas.

Exemple 15.98 (Analycité de l'exponentielle)**.**
Soient $a \in \mathbb{R}$ et $R > 0$. Nous démontrons que exp est analytique sur $B(a, R)$. Si $f(x) = e^x$, alors $f^{(n)}(x) = e^x$ pour tout n (équation (15.236a)). Nous avons donc

$$|f^{(n)}(x)| < e^{a+R} \tag{15.308}$$

pour tout $x \in B(a, R)$. Nous partons de l'expression (15.53) du reste :

$$|R_n(x)| \leqslant \frac{M_n}{(n+1)!}|x-a|^{n+1} \leqslant \frac{e^{a+R}}{(n+1)!}R^{n+1}. \tag{15.309}$$

Mais nous avons la limite

$$\lim_{n\to\infty} \frac{R^{n+1}}{(n+1)!} = 0 \tag{15.310}$$

pour tout R.

Donc avec les polynômes de Taylor P_n calculés en a, nous avons $P_n \to \exp$ simplement sur $\mathbb{R}$.

Nous pouvons donc développer la fonction exponentielle autour de n'importe quel point, et avoir convergence des polynômes vers l'exponentielle sur tout $\mathbb{R}$. Vous accepterez cependant que si a et x sont éloignés, la convergence $P_n(x) \to \exp(x)$ peut être extrêmement lente. △

15.6.11 Autres propriétés et petits calculs

Lemme 15.99.
Si les suites (u_n) et (v_n) sont équivalentes[37] *et si (v_n) admet une limite l différente de 1, alors les suites $(\ln u_n)$ et $(\ln v_n)$ sont équivalentes.*

Démonstration. En effet si $u_n = v_n\alpha(n)$ alors en utilisant la formule du lemme 15.87,

$$\ln(u_n) = \ln(v_n) + \ln\big(\alpha(n)\big) = \ln(v_n)\left(1 + \frac{\ln\big(\alpha(n)\big)}{\ln(v_n)}\right), \tag{15.311}$$

et comme $\alpha(n) \to 1$, la parenthèse tend vers 1. □

15.6.12 Taylor pour le logarithme

Vu que $\ln(0)$ n'existe pas, il n'est pas question de développer ln autour de $x = 0$. À la place, nous allons le développer autour de $x = 1$ et plus précisément nous allons étudier Taylor pour la fonction $f(x) = \ln(1 + x)$. Les résultats seront résumés dans la proposition 15.102.

PROPooWCUEooJudkCV

Proposition 15.100 ([1])**.**
Soit la fonction

$$\begin{aligned} f\colon]-1, \infty[&\to \mathbb{R} \\ x &\mapsto \ln(1+x). \end{aligned} \tag{15.312}$$

Pour tout n, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0}\alpha(t) = 0$ et

$$f(x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}x^k + \alpha(x)x^n \tag{15.313}$$

pour tout x dans le domaine de f.

Notez la somme qui part de $k = 1$ et non $k = 0$.

37. Définition 9.32.

Démonstration. Nous utilisons la formule de Taylor-Young (proposition 12.477). La première dérivée de f se calcule en utilisant le lemme 15.96 :

$$f'(x) = \frac{1}{1+x}. \tag{15.314}$$

Pour les dérivées suivantes, c'est juste du calcul et nous pouvons prouver par récurrence que

$$f^{(k)}(x) = \frac{(k-1)!(-1)^{k+1}}{(1+x)^k}. \tag{15.315}$$

EQooKEAOooGmTLJF

En ce qui concerne l'évaluation en zéro :

$$f^{(k)}(0) = \begin{cases} 0 & \text{si } k = 0 \\ (k-1)!(-1)^{k+1} & \text{sinon.} \end{cases} \tag{15.316}$$

Du fait que $f^{(0)}(0) = \ln(1) = 0$, la somme commence à $k = 1$ et non $k = 0$. Nous avons

$$f(x) = \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!}x^k + \alpha(x)x^n = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}x^k + \alpha(x)x^n. \tag{15.317}$$

□

Nous étudions les polynômes de la série de Taylor pour

$$\begin{aligned} f\colon]-1,\infty[&\to \mathbb{R} \\ x &\mapsto \ln(1+x). \end{aligned} \tag{15.318}$$

Les dérivées successives de f ont déjà été calculées en (15.315). Nous développons autour de $x = 0$. Donc $f(0) = \ln(1) = 0$ et pour les autres,

$$f^{(k)}(0) = (-1)^{k+1}(k-1)!. \tag{15.319}$$

Pour les polynômes de Taylor, nous avons

$$P_n(x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}x^k \tag{15.320}$$

où vous noterez la somme qui part de $k = 1$ et non de $k = 0$. Nous avons aussi la série de Taylor de f donnée par

$$T(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}x^k. \tag{15.321}$$

EQooTAREooKfpTPo

La somme est une limite ponctuelle, là où elle existe.

Jusqu'à présent, la seule certitude à props de T est que $T(0) = f(0) = 0$. Pour le reste :

— Rien ne dit que $T(x)$ existe pour d'autres x que $x = 1$.

— Et même si $T(x)$ existait pour d'autres x (c'est-à-dire si le rayon de convergence de (15.321) était strictement plus grand que zéro), rien n'assurerait que la valeur serait celle de f.

— Et même si $T(x)$ convergeait vers f sur son disque de convergence, ce ne serait pas encore assez pour dire que f est analytique, parce que l'analycité demande que les séries de Taylor autour de *chaque* point converge vers f. Or ici nous ne parlons encore que de T qui est la série autour de $x = 0$.

LEMooWMGGooRpAxBa

Lemme 15.101.
La série de Taylor de $x \mapsto \ln(1+x)$ autour de $x = 0$ converge sur $]-1, 1]$. Elle ne converge pas pour $x = -1$.

Démonstration. En ce qui concerne le rayon de convergence de T, nous utilisons la formule de Hadamard [38] avec

$$a_k = \frac{(-1)^{k+1}}{k}. \tag{15.322}$$

Ce que nous trouvons est

$$\frac{1}{R} = \lim_{k\to\infty} |\frac{a_{k+1}}{a_k}| = \lim_{k\to\infty} \frac{k}{k+1} = 1. \tag{15.323}$$

Le rayon de convergence de T est donc 1. Nous avons donc que $P_n \to T$ sur $]-1,1[$, et peut-être que $P_n \to T$ en $x = \pm 1$.

Pour $x = -1$. L'intuition nous dit que ce serait $\ln(0)$ qui n'est pas défini. C'est le cas parce que

$$P_n(-1) = \sum_{k=1}^{n} \frac{(-1)^{k+1}(-1)^k}{k} = -\sum_{k=1}^{n} \frac{1}{k}. \tag{15.324}$$

La limite $n \to \infty$ diverge. Donc T n'est pas définie en $x = -1$.

Pour $x = 1$ par contre,

$$P_n(1) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}. \tag{15.325}$$

Le critère des séries alternées [39] nous donne la convergence de cette série. □

Nous savons maintenant que la série de Taylor T converge sur $]-1,1]$, et que $T(0) = f(0) = \ln(1) = 0$. Le premier gros morceau intéressant vient maintenant : nous allons prouver que $T(x)$ converge vers ce que nous croyons, c'est-à-dire $\ln(1+x)$ en personne.

PROPooKPBIooJdNsqX

Proposition 15.102.
Pour tout $x \in]-1,1]$ nous avons

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \tag{15.326}$$

EqweEZpV

De plus nous avons

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln(2). \tag{15.327}$$

EqKUQmQZ

Démonstration. Il s'agit d'utiliser l'expression du reste fourni par le théorème 12.464. Pour tout $x \in]-1,\infty[$, il existe un $c \in]0,x[$ (le c dépend de x) tel que

$$P_n(x) - f(x) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1}. \tag{15.328}$$

Cela est parce que f est de classe C^∞. Calculons un peu :

$$P_n(x) - f(x) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1} \tag{15.329a}$$

$$= \frac{(-1)^n n!}{(1+c)^{n+1}} \frac{1}{(n+1)!} x^{n+1} \tag{15.329b}$$

$$= \frac{(-1)^n}{n+1} \left(\frac{x}{1+c}\right)^{n+1}. \tag{15.329c}$$

Lorsque $x > 1$, il n'y a aucune garantie sur la convergence de cela pour $n \to \infty$. Pour rappel, $c \in]0,x[$. Si par contre $x \in]-1,1[$, alors nous savons que

$$\left|\frac{x}{1+c}\right| < 1, \tag{15.330}$$

38. Théorème 15.25.
39. Théorème 10.140.

et donc convergence $P_n(x) - f(x) \to 0$.

Jusqu'ici nous avons prouvé que pour la série de Taylor converge vers $\ln(1+x)$ pour $x \in]-1,1[$. Nous avons également vu que la série converge pour $x = 1$. Donc la fonction

$$g(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \tag{15.331}$$

est continue sur $]-1,1]$ et égale à $\ln(x+1)$ sur $]-1,1[$. Vu que $f\colon x \mapsto \ln(x+1)$ est continue sur $]-1,\infty[$, nous avons également $g(1) = f(1) = \ln(2)$.

Ceci nous mène au dernier point de notre proposition : $g(1) = \ln(2)$ s'écrit précisément

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln(2). \tag{15.332}$$

□

15.103.
La formule (15.327) peut sembler très chouette pour trouver des approximations de $\ln(2)$. Le problème est qu'elle ne donne aucune idée de l'erreur commise en tronquant la série.

Vous pouvez, certes, écrire ceci :

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} = \frac{47}{60} \simeq 0.78. \tag{15.333}$$

Hélas, ce calcul n'a aucune valeur pour affirmer que $\ln(2)$ doit être proche de 0.78. Ni même pour affirmer que $\ln(2) < 1$.

Avoir des valeurs numériques de $\ln(2)$ (c'est-à-dire que « chiffres corrects devant ou derrière la virgule ») demande d'avoir un encadrement. Cela doit donc se faire avec des formules de séries avec reste ; les formules exactes qui demandent de sommer jusqu'à l'inifini sont inutiles pour avoir des approximations numériques.

Dans le cas de $\ln(2)$, une approximation numérique sera donnée à l'aide de Taylor avec reste intégrale dans la proposition 20.151.

Lemme 15.104.
Soit la fonction[40]

$$f(x) = \frac{\ln(1+x)}{x} \tag{15.334}$$

(1) Elle admet un prolongement de classe C^∞ sur $]-1,\infty[$.
(2) $f(0) = 1$.

La seconde condition étant évidemment avec un abus de notation entre f et son prolongement, parce que f n'est pas définie en zéro.

Démonstration. La difficulté étant de voir que f a un prolongement en zéro et qu'elle y est de classe C^∞.

La proposition 15.102 nous donne l'égalité

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \tag{15.335}$$

pour tout $x \in]-1,1]$; en particulier pour $x = 0$. Nous faisons le petit calcul suivant : SUBEQooRLQOooEzNFDp

$$\frac{1}{x}\ln(1+x) = \frac{1}{x}\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \tag{15.336a}$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^{n-1} \tag{15.336b}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^{n}}{n+1} x^{n}. \tag{15.336c}$$

40. Pour la définition du logarithme, c'est la définition 15.83.

Ce calcul n'est pas valable pour $x = 0$, mais ça ne nous empêche pas de poser

$$T(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^n, \tag{15.337}$$

qui, lui, est bien définie en zéro. Le rayon de convergence de la série T est égal à 1, de telle sorte que

$$T :]-1, 1[\to \mathbb{R} \tag{15.338}$$

de classe C^∞, et est égale à f sur $]-1, 1[\backslash\{0\}$.

La série T est donc le prolongement demandé. En ce qui concerne $f(0)$, c'est un abus pour écrire $T(0)$ qui vaut immédiatement 1. □

Notons qu'un calcul de limite

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} \tag{15.339}$$

donnait la valeur $f(0) = 1$. Donc prolonger avec $f(0) = 1$ était la seule possibilité pour avoir une fonction continue. De là à dire que le prolongement ainsi créé est de classe C^∞, c'est une autre histoire, qui est résolue par les séries entières.

15.6.13 Développements et calcul de limites

Lors d'un calcul de limite, développer une partie d'une expression peut être utile.

Exemple 15.105.
À calculer :

$$\lim_{x\to 0} \frac{\ln(1+x)}{x}. \tag{15.340}$$

Cela est une indétermination de type $\frac{0}{0}$. Le développement limité du numérateur [41] nous donne une fonction $\alpha(x)$ telle que $\lim_{x\to 0} \alpha(x) = 0$ et

$$\frac{\ln(1+x)}{x} = \frac{x - \frac{x^2}{2} + x^2\alpha(x)}{x} = 1 - \frac{x}{2} + x\alpha(x). \tag{15.341}$$

Sur le membre de droite la limite est facile à calculer :

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = \lim_{x\to 0} \left(1 - \frac{x}{2} + x\alpha(x)\right) = 1. \tag{15.342}$$

△

15.6.14 Une petite intégrale

EXooNIOZooWxciAC

Exemple 15.106.
Soit V la région trapézoïdale de sommets $(0, -1)$, $(1, 0)$, $(2, 0)$, $(0, -2)$, comme à la figure 15.2(a). Calculons ensemble l'intégrale double

$$\int_V e^{\frac{x+y}{x-y}}\, dV,$$

avec le changement de variable [42] $\psi(x, y) = (x + y, x - y)$. C'est-à-dire que nous considérons les nouvelles variables

$$\begin{cases} u = x + y & \text{(15.343a)} \\ v = x - y. & \text{(15.343b)} \end{cases}$$

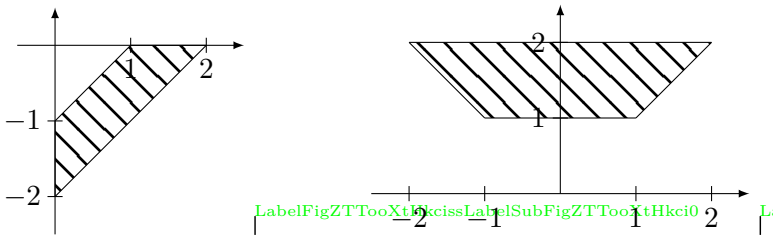
41. Proposition 15.100.
42. Théorème 14.281.

Il faut remarquer d'abord que le changement de variable proposé est dans le mauvais sens. On écrit alors $\phi(u,v) = \psi^{-1}(u,v) = \big((u+v)/2, (u-v)/2\big)$, c'est-à-dire

$$\begin{cases} x = \dfrac{u+v}{2} & \text{(15.344a)} \\ y = \dfrac{u-v}{2}. & \text{(15.344b)} \end{cases}$$

La région qui correspond à V est U, le trapèze de sommets $(-1,1)$, $(1,1)$, $(2,2)$ et $(-2,2)$, qu'on voit sur la figure 15.2(b) et qu'on décrit par

$$U = \{(u,v) \in \mathbb{R}^2 \,|\, 1 \leqslant v \leqslant 2,\ -v \leqslant u \leqslant v\}.$$



(a) La région V (b) La région $U = \phi^{-1}(V)$

FIGURE 15.2: Avant et après le changement de variables LabelFigZTTooXtHkci

Le déterminant de la matrice jacobienne de ψ^{-1} est $J_{\psi^{-1}}$,

$$J_{\psi^{-1}}(u,v) = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}. \tag{15.345}$$

On a alors, en utilisant le fait que $F(x) = ae^{x/a}$ est une primitive de $f(x) = e^{x/a}$ (proposition 15.76) ainsi que le théorème fondamental de l'analyse (théorème 14.287),

$$\int_V e^{\frac{x+y}{x-y}}\,dV = \int_U e^{\frac{u}{v}}\frac{1}{2}\,dV = \int_1^2\int_{-v}^{v} e^{\frac{u}{v}}\frac{1}{2}\,du\,dv = \frac{3}{4}(e - e^{-1}).$$

△

15.7 Vitesses des puissances, de l'exponentielle et du logarithme

15.7.1 Un peu de théorie

Voici une série de résultats qui lient les vitesses des polynômes, du logarithme et de l'exponentielle.

LEMooNYFVooXjFShk

Lemme 15.107.
Si P est un polynôme et si $a > 0$, alors

$$\lim_{x\to\infty} e^{-ax}P(x) = 0 \tag{15.346}$$

Démonstration. Nous prouvons par récurrence que pour tout n, nous avons $e^{-ax}x^n \to 0$. D'abord nous écrivons[43]

$$f(x) = e^{-ax}x = \frac{x}{e^{ax}}, \tag{15.347}$$

et ensuite la règle de l'Hospital 12.195 nous donne

$$\lim_{x\to\infty}\frac{x}{e^{ax}} = \lim_{x\to\infty}\frac{1}{ae^{ax}} = 0. \tag{15.348}$$

43. En utilisant 12.398(3).

En ce qui concerne la récurrence, c'est encore la règle de l'Hospital :

$$\lim_{x\to\infty}\frac{x^n}{e^{ax}}=\frac{n}{a}\lim_{x\to\infty}\frac{x^{n-1}}{e^{ax}}=0. \tag{15.349}$$

□

PROPooKVIFooGdKpfP

Proposition 15.108.
Nous avons :

ITEMooCDSQooSIctbz

(1) *Pour tout $\alpha>0$, il existe N tel que $\ln(n)\leqslant n^\alpha$ pour tout $n\geqslant N$.*

ITEMooZMAWooTbDNAd

(2) *Pour tout $p>0$ et tout $\alpha>0$, il existe N tel que*

$$\ln(n)^p<n^\alpha \tag{15.350}$$

pour tout $n\geqslant N$.

ITEMooBLNOooZQNTfd

(3) *Pour tout $n\geqslant 1$ nous avons la limite*

$$\lim_{x\to0^+}x^n\ln(x)=0. \tag{15.351}$$

ITEMooMLNMooAyJTox

(4) *Nous avons*

$$\lim_{x\to0^+}\frac{\ln(1-x)}{x}=-1. \tag{15.352}$$

ITEMooIQEKooBionsK

(5) *L'exponentielle croit plus vite que tout polynôme, et plus vite que logarithme :*

$$\lim_{x\to\infty}e^{-x}(\ln x)^n x^\alpha=0 \tag{15.353}$$

EqExpDecrtPlusVite

pour tout n et pour tout α.

ITEMooDUQWooNvAvmR

(6) *Pour tout $n>0$, nous avons la limite*

$$\lim_{x\to0^+}x^n e^{1/x}=\infty. \tag{15.354}$$

Le point (1) et sa généralisation (2) nous font dire que le logarithme croît moins vite que n'importe quel polynôme.

Démonstration. En plusieurs parties.

(1) **Pour (1)** En effet, nous avons, par la règle de l'Hospital (proposition 12.194),

$$\lim_{x\to\infty}\frac{x^\alpha}{\ln(x)}=\lim_{x\to\infty}\frac{\alpha x^{\alpha-1}}{1/x}=\lim_{x\to\infty}\alpha x^\alpha=\infty \tag{15.355}$$

quand $\alpha>0$. La dérivée du logarithme est dans la proposition 15.93.

(2) **Pour (2)** Il faut prendre le N qui convient à l'item (1) pour $n^{\alpha/p}$. Ainsi nous avons $\ln(n)<n^{\alpha/p}$ et donc $\ln(n)^p\leqslant n^\alpha$.

(3) **Pour (3)** Lorsque $x\neq0$ nous avons

$$x^n\ln(x)=\frac{\ln(x)}{1/x^n}, \tag{15.356}$$

qui est un cas $\frac{\infty}{\infty}$. Nous nous en remettons à la règle de l'Hospital 12.195. D'abord nous nous assurons de la limite des dérivées :

$$\lim_{x\to0^+}\frac{1/x}{-nx^{-n-1}}=\lim_{x\to0^+}-\frac{1}{n}\frac{x^{n+1}}{x}=0. \tag{15.357}$$

La règle de l'Hospital conclut à l'existence de la limite demandée et à son égalité à 0.

(4) **Pour (4)** En effet, par la règle de l'Hospital 12.194,

$$\lim_{x\to 0} -\frac{\ln(1-x)}{x} = \lim_{x\to 0} -\frac{\frac{-1}{1-x}}{1} = \lim_{x\to 0} \frac{1}{1-x} = 1 \tag{15.358}$$

EqGICpOX

(5) **Pour (5)** Notons $f_1(x) = e^{-x/2}(\ln x)^n$ et $f_2(x) = e^{-x/2}x^\alpha$. Le lemme 15.107 donne tout de suite $\lim_{x\to\infty} f_2(x) = 0$.

En ce qui concerne f_1, l'item (2) nous indique que nous avons

$$f_1(x) = e^{-x/2}(\ln x)^n \leqslant e^{-x/2}x \tag{15.359}$$

dès que x est assez grand. Le lemme 15.107 nous dit alors que $\lim_{x\to\infty} f_2(x) = 0$.

Enfin, nous avons

$$e^{-x}(\ln x)^n x^\alpha = f_1(x)f_2(x) \tag{15.360}$$

et donc la limite demandée.

(6) **Pour (6)** Nous passons au logarithme :

$$\ln(x^n e^{1/x}) = \ln(x^n) + \ln(e^{1/x}) = n\ln(x) + \frac{1}{x} = \frac{nx\ln(x)+1}{x}. \tag{15.361}$$

Grâce à la limite déjà prouvée en (3), le numérateur tend vers 1 lorsque $x \to 0^+$. Donc le tout tend vers $+\infty$. Au final,

$$\lim_{x\to 0^+} x^n e^{1/x} = \lim_{x\to 0^+} e^{\ln(x^n e^{1/x})} = \infty. \tag{15.362}$$

□

EXooQNCJooFpnvnf

Exemple 15.109.
Le lemme 12.411 a déjà prouvé la limite

$$\lim_{n\to\infty} n^\alpha a^n \tag{15.363}$$

pour tout $\alpha > 0$ et $a < 1$.

L'utilisation de propriétés de l'exponentielle nous permet de donner une nouvelle preuve, plus courte [44].

Le théorème 15.91 et la proposition 15.86 nous permettent de passer à l'exponentielle. Pour chaque n nous avons :

$$n^\alpha a^n = e^{\alpha\ln(n)+n\ln(a)}. \tag{15.364}$$

EqLKLQooLIWgm

Ce qui est dans l'exponentielle est

$$\alpha\ln(n) + n\ln(a) = n\big(\alpha\frac{\ln(n)}{n} + \ln(a)\big). \tag{15.365}$$

Dans la parenthèse, $\ln(a) < 0$ et $\frac{\ln(n)}{n} \to 0$. Donc ce qui est dans l'exponentielle (15.364) tend vers $-\infty$ et au final l'expression demandée tend vers zéro. △

Remarque 15.110.
Vous ne pouvez pas à priori considérer l'exemple 15.109 comme une preuve alternative au lemme 12.411, parce que vous n'êtes pas sûr que dans toute la théorie permettant de définir l'exponentielle (en particulier la convergence de $\sum_k x^k/k!$), le lemme n'est pas utilisé [45].

44. C'est toujours facile de prétendre qu'une preuve est plus courte qu'une autre lorsqu'on utilise en une ligne des très gros théorèmes qui ont mis dix pages à être démontrés.

45. Faites la vérification et dites moi si c'est bon.

Lemme 15.111.
*La **fonction de Riemann***

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \tag{15.366}$$

est de classe C^∞ sur $]1, \infty[$.

Démonstration. Afin de faire le coup du compact, nous étudions la convergence uniforme de la série sur tout compact de $]1, \infty[$. Soit $\epsilon > 0$, et regardons ce qu'il se passe sur un compact dont le minimum [46] est $1 + \epsilon$. Dans ce cas, $n^x \geqslant n^{1+\epsilon}$, et donc $f_n(x) = \frac{1}{n^x} \leqslant \frac{1}{n^{1+\epsilon}}$. Étant donné que la série numérique $\sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}}$ converge, la fonction de Riemann converge uniformément par le critère de Weierstrass (théorème 12.372). Nous avons donc convergence uniforme de la série sur tout compact de $]1, \infty[$, ce qui fait que ζ est une fonction continue pour $x > 1$ par le théorème 12.370.

Nous allons à présent utiliser le théorème 12.374 pour prouver que la fonction de Riemann est C^1. Il faut donc prouver que la série des dérivées $(n^{-x})' = -\ln(n)n^{-x}$ converge uniformément sur tout compact de $]1, \infty[$.

Nous prenons encore une fois un compact K dont le minimum est $1+\epsilon$. D'abord, nous majorons le logarithme par un x^α : lorsque n est assez grand, nous avons

$$\ln(n)n^{-x} \leqslant n^\alpha n^{-x}; \tag{15.367}$$

la proposition 15.108(1) nous dit que pour tout $\alpha > 0$, il existe un n à partir duquel cette inégalité est valide. Étant donné que $1 + \epsilon$ est le minimum du compact, nous pouvons encore majorer en remplaçant x par $1 + \epsilon$:

$$\ln(n)n^{-x} \leqslant \frac{1}{n^{1+\epsilon-\alpha}}. \tag{15.368}$$

Afin de pouvoir utiliser le critère de Weierstrass, nous devons nous assurer que la série $\sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon-\alpha}}$ converge. Cela n'est vrai que si $1 + \epsilon - \alpha > 1$, mais le choix de α étant encore arbitraire, nous choisissons $0 < \alpha < \epsilon$.

Ainsi, la série des dérivées converge uniformément sur tout compact et nous en déduisons que cette série est bien la dérivée de la fonction de Riemann qui est C^1.

Afin de traiter les dérivées d'ordre supérieur, il faut calculer

$$(n^{-x})^p = (-1)^p\big(\ln(n)\big)^p n^{-x}, \tag{15.369}$$

et remarquer que $\lim_{x\to\infty} x^\alpha/(\ln(x))^p = \infty$. Par conséquent y a encore moyen de remplacer le logarithme par un x^α. Le reste de la preuve est la même. □

Ici se termine la preuve de ce lemme. Nous restons cependant sur notre faim en ce qui concerne la convergence uniforme de la série sur l'ouvert $]1, \infty[$. En effet, nous avons prouvé la convergence uniforme sur tout compact (et cela nous a suffit pour prouver le lemme), mais nous n'avons pas prouvé que la série n'était pas uniformément convergente sur $]1, \infty[$ pour autant.

Nous allons montrer qu'il n'y a pas uniforme convergence en prouvant que si x est assez proche de 1, alors la suite des sommes partielles de $\zeta(x)$ est aussi proche que l'on veut de la suite des sommes partielles de $\sum_{n=1}^{\infty} \frac{1}{n}$ qui, elle, diverge.

Lemme 15.112.
Nous avons

$$\lim_{x\to 1^+} \zeta(x) = \infty \tag{15.370}$$

où la limite est une limite à droite : la limite à gauche n'existe pas.

Démonstration. Soit $M > 0$. Prouvons que $\exists \epsilon$ tel que $\zeta(1 + \epsilon) \geqslant M$. D'abord, choisissons un k tel que

$$\sum_{n=1}^{k} \frac{1}{n} > M, \tag{15.371}$$

46. Pour rappel, un compact dans $\mathbb{R}$ a *toujours* un minimum.

et choisissons un ϵ tel que

$$\max_{n\in\{1,\dots,k\}}\left|\frac{1}{n}-\frac{1}{n^{1+\epsilon}}\right|<\alpha. \tag{15.372}$$

Un tel choix de ϵ est possible pour tout α. Maintenant, nous choisissons α de façon à avoir $k\alpha<\sigma$. Avec ça, nous avons

$$\sum_{n=1}^{k}\frac{1}{n}-\sum_{n=1}^{k}\frac{1}{n^{1+\epsilon}}=\sum_{n=1}^{k}\left(\frac{1}{n}-\frac{1}{n^{1+\epsilon}}\right)<k\alpha<\sigma. \tag{15.373}$$

En prenant σ tel que $M-\sum_{n=1}^{k}(1/n)<\sigma$, nous trouvons ainsi un ϵ tel que $\sum_{n=1}^{k}\frac{1}{n^{1+\epsilon}}>M$. Cela prouve le lemme. □

Armé de ce lemme, il est maintenant aisé de prouver que la série définissant la fonction de Riemann n'est pas uniformément convergente sur $]1,\infty[$. Prenons la $k$ième somme partielle $s_k(x)=\sum_{n=1}^{k}(1/n^x)$. Pour chaque k, cela est une fonction bornée de x (y compris en $x=1$), donc $\sup_{x\in]1,\infty[}s_k(x)=M_k$. Armé de cette majoration, nous faisons

$$\|s_k-\zeta\|_\infty=\sup_{x\in]1,\infty[}\big(\zeta(x)-s_k(x)\big)>\sup\big(\zeta(x)-M_k\big)=\infty, \tag{15.374}$$

il n'y a donc pas moyen que la limite de $\|\zeta-s_k\|_\infty$ quand $k\to\infty$ soit nulle. Il n'y a donc pas uniforme convergence de ζ sur l'intervalle $]1,\infty[$.

PropBQGBooHxNrrf

Proposition 15.113.
Pour tout polynôme P et pour tout $a>0$ la fonction $f(x)=P(x)e^{-ax}$ est intégrable[47] *sur $[0,\infty[$.*

Démonstration. Nous avons $f(x)=P(x)e^{-ax/2}e^{-ax/2}$, et par la vitesse comparée des exponentielles et polynômes[48], pour un certain $M>0$ nous pouvons affirmer que $P(x)e^{-ax/2}<1$ sur $[M,\infty[$. Dès lors

$$|f(x)|<e^{-ax/2}, \tag{15.375}$$

qui est intégrable. □

EXooAGEOooQdQkrS

Exemple 15.114.
La fonction logarithme (définition 15.83) n'est pas définie pour $x\leqslant 0$. Par conséquent la fonction $f(x)=x\ln(|x|)$ n'est pas définie en $x=0$. Elle est bien définie pour $x<0$ et vérifie

$$\lim_{x\to 0}x\ln(|x|)=0. \tag{15.376}$$

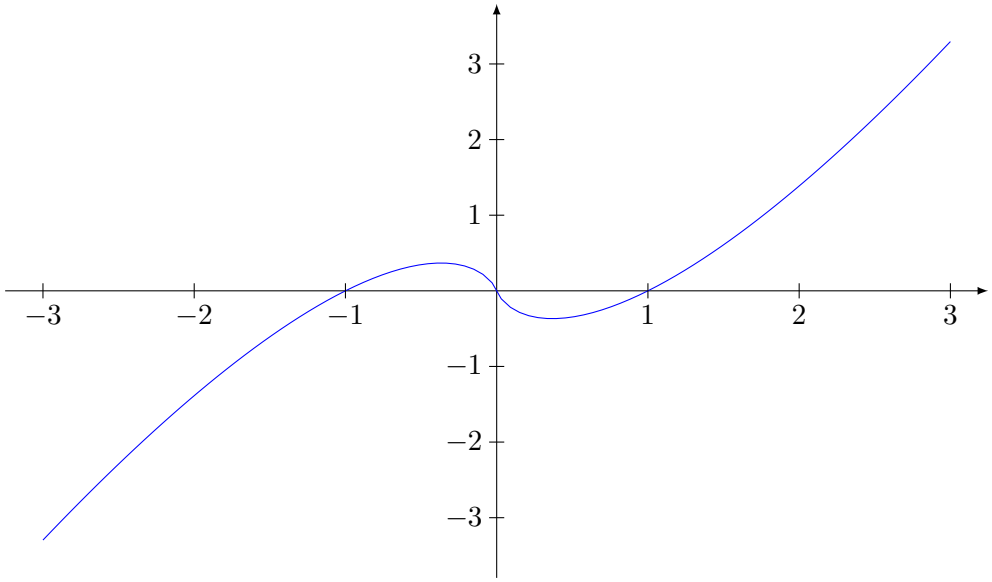
Nous pouvons donc définir la fonction

$$\begin{aligned}\tilde{f}\colon\mathbb{R}&\to\mathbb{R}\\ x&\mapsto\begin{cases}x\ln(|x|) & \text{si } x\neq 0\\ 0 & \text{si } x=0.\end{cases}\end{aligned} \tag{15.377}$$

Contrairement à la fonction initiale f, cette fonction $\tilde{f}$ est définie et continue en 0.

Notez que sur le graphe de la fonction $\tilde{f}$, la courbe est bien régulière en $x=0$.

47. Définition 14.186.
48. Voir 15.108 et 15.107.



△

DEFooFJZXooIVrzvk

Définition 15.115.
Soient deux suites (a_n) et (b_n) dans $\mathbb{R}$. Nous disons que

(1) $a_n \to \infty$ ***plus vite*** *que* (b_n) *si* $\lim n \to \infty a_n/b_n = \infty$.

(2) $a_n \to \infty$ ***aussi vite*** *que* (b_n) *si* $\lim n \to \infty a_n/b_n$ *existe et est finie.*

(3) $a_n \to 0$ ***plus vite*** *que* (b_n) *si* $\lim n \to \infty a_n/b_n = 0$.

(4) $a_n \to 0$ ***moins vite*** *que* (b_n) *si* $\lim n \to \infty a_n/b_n = \infty$.

Exemple 15.116.
Deux petits exemples.

(1) Les suites $a_n = n^2$ et $b_n = n$ tendent toutes deux vers ∞, mais (a_n) va plus vite parce que $a_n/b_n = n \to \infty$.

(2) Les suites $a_n = 1/n$ et $b_n = 1/n^2$ tendent toutes deux vers 0, mais (a_n) va moins vite parce que $a_n/b_n = n \to \infty$.

△

15.117.
Je vous laisse cogiter sur l'énoncé suivant : si A est un ensemble de suites tendant toutes vers ∞, alors il existe une suite qui converge vers ∞ plus vite que tous les éléments de A.

Lorsque A est dénombrable, à mon avis, il suffit de prendre $x_n = \big(\max_{i=1,\dots,n}\{a_n^{(i)}, 2\}\big)^2$.

Lorsque A n'est pas dénombrable, je ne suis pas sûr. Écrivez-moi si vous avez une idée.

Lemme 15.118.
Si $a_n \to \infty$, il existe une suite tendant vers ∞ plus vite.

Démonstration. Il suffit de prendre $b_n = a_n^2$. □

Lemme 15.119.
La suite $a_n = n!$ tend vers ∞ vite que l'exponentielle $b_n = e^n$.

Démonstration. Nous devons étudier le comportement de $n!/e^n$ lorsque n est grand. Étant donné que $e < 3$ (lemme 15.81), nous avons

$$\frac{n!}{e^n} \geqslant \frac{1 \times 2 \times 3 \times 4^{n-3}}{3^n} = \frac{5}{3^3}\left(\frac{4}{3}\right)^{n-3}. \tag{15.378}$$

Et comme $4/3 > 1$ nous avons bien $n!/e^n \to \infty$. □

15.7.2 Nombres premiers

Le théorème suivant dit que la somme des inverses des nombre premiers diverge. Cela est à comparer avec la proposition 10.136 qui dit que la somme des inverses des carrés converge.

ThonfVruT

Théorème 15.120.
Soit P, l'ensemble des nombres premiers. Alors la somme $\sum_{p\in P}\frac{1}{p}$ diverge et plus précisément,

$$\sum_{\substack{p\leqslant x\\ p\in P}}\frac{1}{p}\geqslant \ln(\ln(x))-\ln(2). \tag{15.379}$$

Démonstration. Nous posons

$$S_x=\{q\leqslant x \text{ avec } q \text{ sans facteurs carrés}\} \tag{15.380}$$

et

$$P_x=\{p\in P \text{ tel que } p\leqslant x\}. \tag{15.381}$$

Si

$$K_x=\{(q,m) \text{ tels que } q \text{ n'a pas de facteurs carrés et } qm^2\leqslant x\}, \tag{15.382}$$

alors nous avons

$$K_x=\bigcup_{q\in S_x}\bigcup_{m\leqslant\sqrt{x/q}}(q,m). \tag{15.383}$$

Par définition et par le lemme 3.24 nous avons aussi

$$\{n\leqslant x\}=\{qm^2 \text{ tel que } (q,m)\in K_x\}. \tag{15.384}$$

Tout cela pour décomposer la somme

$$\sum_{n\leqslant x}\frac{1}{n}=\sum_{q\in S_x}\sum_{m\leqslant\sqrt{x/q}}\frac{1}{qm^2}\leqslant\sum_{q\in S_x}\frac{1}{q}\underbrace{\sum_{m\geqslant 1}\frac{1}{m^2}}_{=C}. \tag{15.385}$$

EqpoJpuC

Nous avons aussi

$$\prod_{p\in P_x}\left(1+\frac{1}{p}\right)=1+\sum_{p\in P_x}\frac{1}{p}+\sum_{\substack{p,q\in P_x\\ p<q}}\frac{1}{pq}+\sum_{\substack{p,q,r\in P_x\\ p<q<r}}\frac{1}{pqr}+\ldots \tag{15.386a}$$

$$\geqslant 1+\sum_{p\in P_x}\frac{1}{p}+\sum_{\substack{p,q\in P_x\\ pq\leqslant x}}\frac{1}{pq}+\sum_{\substack{p,q,r\in P_x\\ pqr\leqslant x}}\frac{1}{pqr}+\ldots \tag{15.386b}$$

Les sommes sont finies. Les sommes s'étendent sur toutes les façons de prendre des produits de nombres premiers distincts de telle sorte de conserver un produit plus petit que x ; c'est-à-dire que les sommes se résument en une somme sur les éléments de S_x :

$$\exp\left(\sum_{p\in P_x}\frac{1}{p}\right)\geqslant\prod_{p\in P_x}\left(1+\frac{1}{p}\right)\geqslant\sum_{q\in S_x}\frac{1}{q}. \tag{15.387}$$

EqqoilQz

La première inégalité est simplement le fait que $1+u\leqslant e^u$ si $u\geqslant 0$ (directe de la définition 15.91). Les inégalités suivantes proviennent du fait que le logarithme est une primitive de la fonction inverse (proposition 15.93) :

$$\ln(x)\leqslant\sum_{n\geqslant x}\int_n^{n+1}\frac{dt}{t}\leqslant\sum_{n\geqslant x}\frac{1}{n}. \tag{15.388}$$

Nous prolongeons ces inégalités avec les inégalités (15.385) et (15.387) :

$$\ln(x) \leqslant \sum_{n \geqslant x} \frac{1}{n} \leqslant C \sum_{q \in S_x} \frac{1}{q} \leqslant C \exp\left(\sum_{p \in P_x} \frac{1}{p} \right). \tag{15.389}$$

En passant au logarithme,

$$\ln\big(\ln(x)\big) \leqslant \ln(C) + \sum_{p \in P_x} \frac{1}{p}. \tag{15.390}$$

Ceci montre la divergence de la série de droite. Nous cherchons maintenant une borne pour C. Pour cela nous écrivons

$$\sum_{n=1}^{N} \frac{1}{n^2} \leqslant 1 + \sum_{n=2}^{N} \frac{1}{n(n-1)} \tag{15.391a}$$

$$= 1 + \sum_{n=2}^{N} \left(\frac{1}{n-1} - \frac{1}{n} \right) \tag{15.391b}$$

$$= 1 + 1 - \frac{1}{N} \tag{15.391c}$$

$$\leqslant 2. \tag{15.391d}$$

Donc $C \leqslant 2$. □

Ce théorème prend une nouvelle force en considérant le théorème de Müntz 17.7 qui dit qu'alors l'ensemble $\mathrm{Span}\{x^p \text{ tel que } p \text{ est premier}\}$ est dense dans les fonctions continues sur $[0,1]$ muni de la norme uniforme ou $\|.\|_2$.

15.7.3 Quelques limites

Nous voyons à présent quelques calculs de limite et de développements mettant en scène des logarithmes et exponentielles.

composel

Exemple 15.121.
Pour trouver le développement de la fonction $f(x) = e^{-2x}$, il suffit d'écrire celui de e^t et de remplacer ensuite t par $-2x$. Le développement à l'ordre 3 de la fonction exponentielle est :

$$e^t = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + t^3\alpha(t). \tag{15.392}$$

Le développement de $f(x) = e^{-2x}$ sera donc

$$f(x) = 1 - 2x + \frac{4x^2}{2} - \frac{8x^3}{6} - 8x^3\alpha(-2x). \tag{15.393}$$

Donc le polynôme de degré 3 partie régulière de g est :

$$1 - 2x + 2x^2 - \frac{4}{3}x^3, \tag{15.394}$$

et la fonction reste correspondante est :

$$\alpha_g(x) = -8\alpha(-2x). \tag{15.395}$$

△

Exemple 15.122.
Nous savons les développements

$$f(x) = \ln(1+x) \sim x - \frac{x^2}{2} + \frac{x^3}{3} \tag{15.396}$$

et

$$\sin(x) \sim x - \frac{x^3}{6}. \tag{15.397}$$

Nous obtenons le développement d'ordre 3 de la fonction $x \mapsto \ln\big(1+\sin(x)\big)$ en écrivant

$$\ln\big(1+\sin(x)\big) \sim \big(x - \frac{x^3}{6}\big) - \frac{1}{2}\left(x - \frac{x^3}{6}\right)^2 + \frac{1}{3}\left(x - \frac{x^3}{6}\right)^3. \tag{15.398}$$

Il s'agit maintenant de trouver les termes qui sont de degré inférieur ou égal à 3.

D'abord

$$\left(x - \frac{x^3}{6}\right)^2 = x^2 - \frac{x^4}{3} + \frac{x^6}{36} \sim x^2 \tag{15.399}$$

Nous avons alors aussi

$$\left(x - \frac{x^3}{6}\right)^6 \sim x^2\left(x - \frac{x^3}{6}\right) \sim x^3. \tag{15.400}$$

En replaçant tout ça dans (15.398) nous trouvons

$$\ln\big(1+\sin(x)\big) \sim x - \frac{x^2}{2} + \frac{x^3}{6}. \tag{15.401}$$

△

Exemple 15.123.

Calculer

$$\lim_{x\to\infty} e^{1/x}\sqrt{1+4x^2} - 2x. \tag{15.402}$$

Nous allons effectuer un développement asymptotique de la partie « difficile » de l'expression posant d'abord $x = 1/h$. Si $f(x) = e^{1/x}\sqrt{1-4x^2}$ alors

$$g(h) = \frac{1}{|h|}e^h\sqrt{h^2+4} = \frac{1}{h}\big(1+h+h\alpha(h)\big)\big(2+h\beta(h)\big). \tag{15.403}$$

La première parenthèse est le développement de e^h et la seconde celui de $\sqrt{h^2+4}$. Nous nous apprêtons à faire la limite $x \to \infty$ qui correspond à $h \to 0^+$, nous pouvons donc supposer que $h > 0$ et omettre la valeur absolue. En effectuant le produit et en regroupant tous les termes contenant h^2, $\alpha(h)$ ou $\beta(h)$ dans un seul terme $h\gamma(h)$,

$$f(h) = \frac{1}{h}\big(2+2h+h\gamma(h)\big) = \frac{2}{h} + 2 + \gamma(h) = 2x + 2 + \gamma(1/x) \tag{15.404}$$

où γ est une fonction vérifiant $\lim_{t\to 0}\gamma(t) = 0$.

Nous sommes maintenant en mesure de calculer la limite (15.402) :

$$\lim_{x\to\infty} e^{1/x}\sqrt{1+x^2} - 2x = \lim_{x\to\infty}\big(2x + 2 + \gamma(1/x) - 2x\big) = 2. \tag{15.405}$$

△

15.8 Trigonométrie hyperbolique

Définition 15.124.

Les fonctions ***cosinus hyperbolique*** *et* ***sinus hyperbolique*** *sont les fonctions définies sur* $\mathbb{R}$ *par les formules suivantes :*

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \tag{15.406a}$$

$$\sinh(x) = \frac{e^x - e^{-x}}{2}. \tag{15.406b}$$

Si vous ne vous rappelez plus la définition de e^x*, c'est 15.60.*

PROPooUNHHooIksdoJ

Proposition 15.125.
Quelques propriétés algébriques des fonctions trigonométriques hyperboliques.

(1) $\cosh(-x) = \cosh(x)$
(2) $\sinh(-x) = -\sinh(x)$
(3) $\cosh^2(x) - \sinh^2(x) = 1$ ITEMooSJBDooAiRgjj
(4) $\cosh^2(x) + \sinh^2(x) = \cosh(2x)$
(5) $\cosh(x)\cosh(y) + \sinh(x)\sinh(y) = \cosh(x+y)$ ITEMooOJRFooUCUaDl
(6) $\cosh(x)\cosh(y) - \sinh(x)\sinh(y) = \cosh(x-y)$
(7) $\cosh(x)\sinh(y) + \sinh(x)\cosh(y) = \sinh(x+y)$
(8) $\cosh(x)\sinh(y) - \sinh(x)\cosh(y) = -\sinh(x-y)$.

Démonstration. Si s'agit simplement de remplacer les définitions et d'utiliser les formules concernant les puissances, dont la formule (12.1096). □

PROPooAOOHooXvLfrZ

Proposition 15.126.
Quelques propriétés analytiques des fonctions trigonométriques hyperboliques.

(1) $\cosh'(x) = \sinh(x)$
(2) $\sinh'(x) = \cosh(x)$. ITEMooZNZLooNMQFWr
(3) $\cosh(x) \geqslant 1$.

Démonstration. Pour les dérivées, il s'agit d'utiliser la dérivation de l'exponentielle, laquelle est facile par le théorème 15.91(1).

Pour (3), nous commençons par les $x \geqslant 0$. D'abord $\cosh(0) = 1$. Ensuite $\cosh'(x) = \sinh(x) = \frac{e^x - e^{-x}}{2}$. Vu que $x > 0$ nous avons $e^x > e^{-x} > 0$. Donc la dérivée de cosh est strictement positive sur $]0, \infty[$. La fonction y est donc partout plus grande que $\cosh(0) = 1$.

Pour les $x < 0$, nous avons la fait que cosh est paire. □

PROPooQLNYooIIOdvm

Proposition 15.127.
La fonction $\sinh\colon \mathbb{R} \to \mathbb{R}$ *est bijective.*

Démonstration. En deux parties.

(1) **Injective** Si $\sinh(a) = \sinh(b)$, alors le théorème de Rolle 12.189 affirme qu'il existe $c \in]a, b[$ tel que $\sinh'(c) = 0$. Mais la proposition 15.126 nous dit que $\sinh'(x) = \cosh(c) \geqslant 1$. Donc impossible.

(2) **Surjective** Nous avons

$$\lim_{x \to -\infty} \sinh(x) = -\infty \tag{15.407}$$

et

$$\lim_{x \to \infty} \sinh(x) = \infty. \tag{15.408}$$

Soit $y \in \mathbb{R}$. Il existe $m < 0$ tel que $\sinh(m) < y$ et $M > 0$ tel que $\sinh(M) > y$. Le théorème des valeurs intermédiaires 9.93 nous enseigne qu'il existe $x \in [m, M]$ tel que $\sinh(x) = y$.

□

PROPooWEHGooOBqSHY

Proposition 15.128 ([1])**.**
Soient $a, b \in \mathbb{R}$ *tels que* $a^2 - b^2 = 1$. *Il existe un unique* $(x, \sigma) \in \mathbb{R} \times \{\pm 1\}$ *tel que* SUBEQSooBIYDooIBuduV

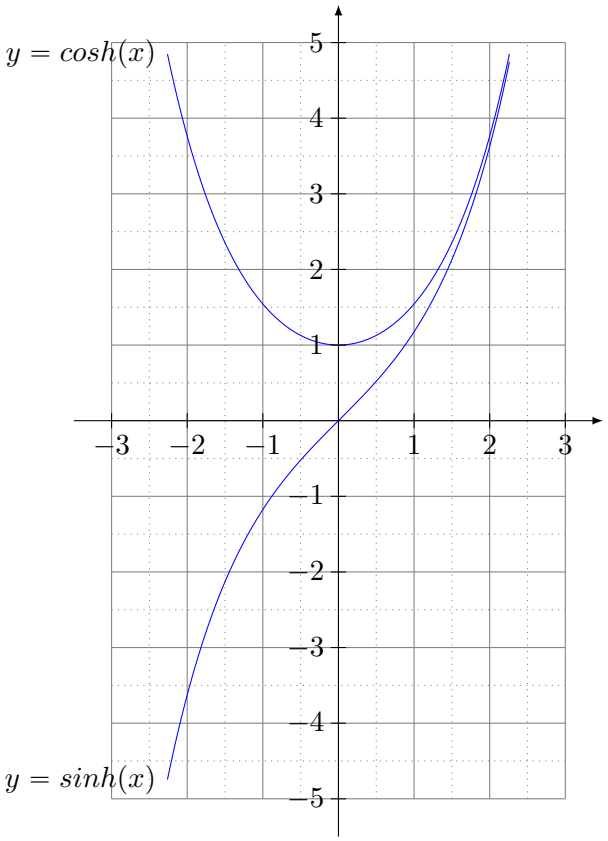
$$\begin{cases} a = \sigma \cosh(x) & \text{(15.409a)} \\ b = \sinh(x). & \text{(15.409b)} \end{cases}$$

Démonstration. Vu que le sinus hyperbolique est une bijection [49], il existe un unique $x \in \mathbb{R}$ tel que $\sinh(x) = b$. Maintenant un petit calcul :

$$a^2 = 1 + \sinh(x)^2 = 1 + \frac{e^{2x} + e^{-2x} - 2}{4} = \frac{e^{2x} + e^{-2x} + 2}{4} = \cosh(x)^2. \tag{15.410}$$

Vu que $\cosh(x)^2 = a^2$, il existe un unique $\sigma \in \{\pm 1\}$ tel que $\sigma \cosh(x) = a$. □

Les représentations graphiques sont ceci :



La **tangente hyperbolique** est donnée par le quotient

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}. \tag{15.411}$$

15.9 Séries entières de matrices

15.9.1 Différentiabilité

PropAMBXKgV

Proposition 15.129.
Soit (a_n) une suite dans $\mathbb{C}$ de rayon de convergence R et la fonction

$$\begin{aligned} f\colon \mathbb{M}(n,\mathbb{C}) &\to \mathbb{M}(n,\mathbb{C}) \\ A &\mapsto \sum_{k=0}^{\infty} a_k A^k \end{aligned} \tag{15.412}$$

Alors

(1) La différentielle de f sur $B(0,R)$ est

$$df_A(U) = \sum_{k=0}^{\infty} a_k \sum_{l=0}^{k-1} A^l U A^{k-1-l}, \tag{15.413}$$

EqRDVodDa

c'est-à-dire que l'on peut différentier terme à terme. (Ici c'est A qui est dans $B(0,R)$)

49. Proposition 15.127.

(2) La convergence de la somme 15.413 est absolue.

(3) La convergence de la somme 15.413 est normale sur tout compact.

(4) La fonction f est de classe C^1 sur $B(0,R)$, c'est-à-dire que la fonction $A \mapsto df_A$ est continue.

Notons que df_A n'est pas tout à fait une série entière. Cependant, en ce qui concerne les normes, c'est tout comme si ça l'était.

Démonstration. Nous posons $u_k(A) = a_k A^k$, qui est une fonction de classe C^∞ et dont la différentielle est donnée par

$$(du_k)_A(U) = \frac{d}{dt}\Big[u_k(A+tU)\Big]_{t=0} = a_k \frac{d}{dt}\Big[(A+tU)^k\Big]_{t=0}; \tag{15.414}$$

en distribuant le produit nous trouvons tout un tas de termes dont seuls ceux contenant exactement une fois tU ne vont pas s'annuler. Étant donné que U et A ne commutent pas nous avons l'expression un peu moche

$$(du_k)_A(U) = \sum_{l=0}^{k-1} a_k A^l U A^{k-1-l}. \tag{15.415}$$

En ce qui concerne la norme, nous regardons celle de $(du_k)_A$ pour un A fixé ; c'est-à-dire que nous en regardons la norme opérateur :

$$\|(du_k)_A\| = \sup_{\|U\|=1} \| \sum_{l=0}^{k-1} a_k A^l U A^{k-1-l} \| \leqslant \sum_{l=0}^{k-1} |a_k| \|A\|^l \|A\|^{k-1-l} \leqslant k|a_k|\|A\|^{k-1}. \tag{15.416}$$

Pour donner la convergence nous considérons un nombre r tel que $\|A\| < r < R$, de telle sorte que la suite $(a_n r^n)$ soit bornée par un nombre M et que nous puissions écrire

$$\|(du_k)_A\| \leqslant k|a_k|\|A\|^{k-1} = \frac{k|a_k|\|A\|^k}{\|A\|} = \frac{k|a_k|}{\|A\|} r^k \left(\frac{\|A\|}{r}\right)^k \leqslant \frac{M}{\|A\|} k \left(\frac{\|A\|}{r}\right)^k, \tag{15.417}$$

dont la série converge. Nous avons donc convergence absolue de la série

$$\sum_{k=0}^{\infty} (du_k)_A. \tag{15.418}$$

Passons à la convergence normale sur tout compact. Nous nous fixons $r < R$ et nous nous intéressons à la norme de du_k sur $\overline{B(0,r)}$, c'est-à-dire

$$\|du_k\|_\infty = \sup_{x \in \overline{B(0,r)}} \|(du_k)_A\|. \tag{15.419}$$

Vu que $\overline{B(0,r)}$ est compact, ce supremum est un maximum et nous pouvons noter A_k la matrice qui le réalise. Nous réalisons alors les mêmes manipulations que pour (15.417) :

$$\|du_k\|_\infty = \|(du_k)_{A_k}\| \leqslant k|a_k|\|A_k\|^{k-1} \leqslant k|a_k|r^{k-1} = \frac{1}{r}k|a_k|r^k. \tag{15.420}$$

Nous prenons maintenant $r < r_0 < R$ et M, un majorant de $(a_n r_0^n)$, de telle sorte qu'en multipliant et divisant par r_0^k,

$$\|du_k\|_\infty \leqslant \frac{k|a_k|r_0^k}{r}\frac{r^k}{r_0^k} \leqslant \frac{kM}{r}\left(\frac{r}{r_0}\right)^k, \tag{15.421}$$

dont la série converge. Nous avons donc convergence normale sur tout compact. Par voie de ~~fait~~ conséquences nous avons continuité de la série

$$\sum_{k=0}^{\infty} (du_k)_A \tag{15.422}$$

et convergence vers df_A par le théorème 15.9. □

PropQIIURAh

Proposition 15.130.
Si le rayon de convergence de la série $u(A) = \sum_{k=0}^{\infty} a_k A^k$ est R, alors

(1) elle converge normalement sur tout compact de $B(0,R)$;

(2) la fonction u y est de classe C^∞.

Démonstration. Nous posons

$$\begin{aligned} u_k \colon \mathbb{M}(n,\mathbb{R}) &\to \mathbb{M}(n,\mathbb{R}) \\ A &\mapsto a_k A^k \end{aligned} \tag{15.423}$$

qui est évidemment une fonction de classe C^∞. Nous étudions la j^{e} différentielle en m, pour $k > j$ (dans une série, nous ne nous intéressons pas aux premiers termes). La j^{e} différentielle appliquée à v_1 appliquée à v_2, etc s'exprime de la façon suivante :

$$(d^j u_k)_m(v_1,\dots,v_j) = \frac{d}{dt_1}\cdots\frac{d}{dt_j}\Big(u_k(m+t_1v_1+\cdots+t_jv_j)\Big)_{t_i=0}. \tag{15.424}$$

Dans le produit $(m+t_1v_1+\cdots+t_jv_j)^k$, seuls les termes contenant exactement une fois chacun des t_i ne s'annulera pas après avoir fait la dérivée et évalué en $t_i = 0$. Combien de termes cela fait ? Parmi les k facteurs, il faut en placer j qui ne sont pas m (cela fait $\binom{k}{j}$ possibilités), et puis il faut ordonner ces j termes, cela fait encore $j!$ possibilités. Au final,

$$\|(d^j u_k)_m\| \leqslant |a_k|\binom{k}{j} j! \|m\|^{k-j} = |a_k| P(k) \|m\|^{k-j} \tag{15.425}$$

où $P(k) = \frac{k!}{(k-j)!}$ est un polynôme de degré j.

Afin d'étudier la convergence normale sur tout compact de la série des $d^j u_k$, nous considérons $r < r_0 < R$ et nous allons prouver la convergence normale sur $\overline{B(0,r)}$. Vu que c'est un compact, il existe une matrice $m_k \in \overline{B(0,r)}$ telle que

$$\begin{aligned} \|d^j u_k\|_\infty &= \|(d^j u_k)_{m_k}\| && \text{(15.426a)} \\ &\leqslant |a_k| P(k) \|m_k\|^{k-j} && \text{(15.426b)} \\ &\leqslant |a_k| P(k) r^{k-j} && \text{(15.426c)} \\ &= \frac{|a_k| P(k)}{r^j} r^k && \text{(15.426d)} \\ &= \frac{|a_k| r_0^k P(k)}{r^j} \left(\frac{r}{r_0}\right)^k && \text{(15.426e)} \\ &\leqslant \frac{M}{r^j} P(k) \left(\frac{r}{r_0}\right)^k && \text{(15.426f)} \end{aligned}$$

où M est un majorant de $a_n r_0^n$. Vu que $r/r_0 < 1$, la somme sur k converge et nous avons convergence normale sur tout compact de

$$d^j \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} d^j(a_k A^k) \tag{15.427}$$

avec un peu d'abus de notation. □

15.10 Exponentielle de matrices

PropKKdmnkD

Proposition 15.131.
Une matrice complexe est inversible si et seulement si elle est une exponentielle.

Autrement dit :

$$\mathrm{GL}(n,\mathbb{C}) = e^{\mathbb{M}(n,\mathbb{C})}. \tag{15.428}$$

Démonstration. Nous avons déjà prouvé dans la proposition 15.71 que toutes les exponentielles étaient inversibles. Ici nous nous concentrons sur la réciproque.

Soit $A \in \mathrm{GL}(n,\mathbb{C})$; nous allons donner une matrice $B \in \mathbb{M}(n,\mathbb{C})$ telle que $A = \exp(B)$. D'abord remarquons qu'il suffit de prouver le résultat pour une matrice par classe de similitude. En effet si $A = \exp(B)$ et si M est inversible alors [EqqACuGK]

$$\exp(MBM^{-1}) = \sum_k \frac{1}{k!}(MBM^{-1})^k \tag{15.429a}$$

$$= \sum_k \frac{1}{k!} MB^kM^{-1} \tag{15.429b}$$

$$= M\exp(B)M^{-1}. \tag{15.429c}$$

Donc $MAM^{-1} = \exp(MBM^{-1})$. Nous pouvons donc nous contenter de trouver un logarithme pour les blocs de Jordan. Nous supposons donc que $A = (\mathbb{1} + N)$ avec $N^m = 0$. En nous inspirant de (15.326), nous posons [50]

$$D(t) = tN - \frac{t^2}{2}N^2 + \cdots + (-1)^m \frac{t^{m-1}}{m-1}N^{m-1} \tag{15.430}$$

et nous allons prouver que $e^{D(1)} = \mathbb{1} + N$. Notons que N étant nilpotente, cette somme ainsi que toutes celles qui viennent sont finies. Il n'y a donc pas de problèmes de convergences dans cette preuve (si ce n'est les passages des équations (15.429)).

Nous posons $S(t) = e^{D(t)}$ (la somme est finie), et nous avons

$$S'(t) = D'(t)e^{D(t)} \tag{15.431}$$

Afin d'obtenir une expression qui donne S' en termes de S, nous multiplions par $(\mathbb{1} + tN)$ en remarquant que $(\mathbb{1} + tN)D'(t) = N$ nous avons

$$(\mathbb{1} + tN)S'(t) = NS(t). \tag{15.432}$$

En dérivant à nouveau,

$$(\mathbb{1} + tN)S''(t) = 0. \tag{15.433}$$ [EqKiccqP]

La matrice $(\mathbb{1} + tN)$ est inversible parce que son noyau est réduit à $\{0\}$. En effet si $(\mathbb{1} + tN)x = 0$, alors $Nx = -\frac{1}{t}x$, ce qui est impossible parce que N est nilpotente. Ce que dit l'équation (15.433) est alors que $S''(t) = 0$. Si nous développons $S(t)$ en puissances de t nous nous arrêtons au terme d'ordre 1 et nous avons

$$S(t) = S(0) + tS'(0) = \mathbb{1} + tD'(0) = 1 + tN. \tag{15.434}$$

En $t = 1$ nous trouvons $S(1) = \mathbb{1} + N$. La matrice $D(1)$ donnée est donc bien un logarithme de $\mathbb{1} + N$. □

15.10.1 Diagonalisabilité d'exponentielle

[PropCOMNooIErskN]

Proposition 15.132 ([114])**.**
Si $A \in \mathbb{M}(n,\mathbb{R})$ a un polynôme caractéristique scindé, alors A est diagonalisable si et seulement si e^A est diagonalisable.

Démonstration. Si A est diagonalisable, alors il existe une matrice inversible M telle que $D = M^{-1}AM$ soit diagonale (c'est la définition 8.215). Dans ce cas nous avons aussi $(M^{-1}AM)^k = M^{-1}A^kM$ et donc $M^{-1}e^AM = e^{M^{-1}AM} = e^D$ qui est diagonale.

La partie difficile est donc le contraire.

50. Le logarithme d'un nombre n'est pas encore définit à ce moment, mais cela ne nous empêche pas de poser une définition ici pour une application des réels vers les matrices.

(1) **Qui est diagonalisable et comment ?** Nous supposons que e^A est diagonalisable et nous écrivons la décomposition de Dunford (théorème 8.265) :

$$A = S + N \tag{15.435}$$

où S est diagonalisable, N est nilpotente, $[S, N] = 0$. Nous avons besoin de prouver que $N = 0$.

Les matrices A est S commutent ; en passant au développement nous en déduisons que A et e^S commutent, puis encore en passant au développement que e^A et e^S commutent. Vu que S est diagonalisable, e^S l'est et par hypothèse e^A est également diagonalisable. Donc e^A et e^{-S} sont simultanément diagonalisables par la proposition 8.222.

Étant donné que A et S commutent, nous avons $e^N = e^{A-S} = e^A e^{-S}$, et nous en déduisons que e^N est diagonalisable vu que les deux facteurs e^A et e^{-S} sont simultanément diagonalisables.

(2) **Unipotence**

Si r est le degré de nilpotence de N, nous avons

$$e^N - \mathbb{1} = N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!}. \tag{15.436}$$ EqQHjyLZQ

Donc

$$(e^N - \mathbb{1})^k = \left(N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!} \right)^k \tag{15.437}$$

où le membre de droite est un polynôme en N dont le terme de plus bas degré est de degré k. Donc $(e^N - \mathbb{1})$ est nilpotente et e^N est unipotente.

Si M est la matrice qui diagonalise e^N, alors la matrice diagonale $M^{-1}e^N M$ est tout autant unipotente que e^N elle-même. En effet,

$$\begin{aligned}
(M^{-1}e^N M - \mathbb{1})^r &= \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} M^{-1} (e^N)^k M & (15.438a)\\
&= M^{-1} \left(\sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} (e^N)^k \right) M & (15.438b)\\
&= M^{-1}(e^N - \mathbb{1})^r M & (15.438c)\\
&= 0. & (15.438d)
\end{aligned}$$

La matrice $M^{-1}e^N M$ est donc une matrice diagonale et unipotente ; donc $M^{-1}e^N M = \mathbb{1}$, ce qui donne immédiatement que $e^N = \mathbb{1}$.

(3) **Polynômes annulateurs**

En reprenant le développement (15.436) sachant que $e^N = \mathbb{1}$, nous savons que

$$N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!} = 0. \tag{15.439}$$

Dit en termes pompeux (mais non moins porteurs de sens), le polynôme

$$Q(X) = X + \frac{X^2}{2} + \cdots + \frac{X^{r-1}}{(r-1)!} \tag{15.440}$$

est un polynôme annulateur de N.

La proposition 8.102 stipule que le polynôme minimal d'un endomorphisme divise tous les polynômes annulateurs. Dans notre cas, X^r est un polynôme annulateur et donc le polynôme minimal de N est de la forme X^k. Donc il est X^r lui-même.

Nous avons donc $X^r \mid Q$. Mais Q est un polynôme contenant le monôme X donc X^r ne peut diviser Q que si $r = 1$. Nous en concluons que X est un polynôme annulateur de N. C'est-à-dire que $N = 0$.

(4) **Conclusion**
Vu que Dunford[51] dit que $A = S + N$ et que nous venons de prouver que $N = 0$, nous concluons que $A = S$ avec S diagonalisable.

□

15.11 Étude d'asymptote

Lorsqu'une fonction tend vers l'infini pour $x \to \infty$, une question qui peut venir est : à quelle vitesse tend-t-elle vers l'infini ?

Il est « visible » que la fonction logarithme ne tend pas très vite vers l'infini : certes

$$\lim_{x\to\infty} \ln(x) = +\infty, \tag{15.441}$$

mais par exemple $\ln(100000) \simeq 11.5$ tandis que $e^{100000} \simeq 10^{43429}$. Sans contestations possibles, l'exponentielle croit plus vite que le logarithme.

Soient f et g deux fonctions dont la limite $x \to \infty$ est ∞. Si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0 \tag{15.442}$$

nous disons que g tend vers ∞ plus vite que f ; si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty \tag{15.443}$$

nous disons que f tend vers ∞ plus vite que g, et si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = a \in \mathbb{R} \tag{15.444}$$

avec $a \neq 0$ alors nous disons que f tend vers l'infini à la même vitesse que $ag(x)$.

Exemple 15.133.
La fonction $x \mapsto x^2$ tend vers l'infini plus vite que la fonction $x \mapsto \sqrt{x}$. △

Dans cette section nous allons nous contenter de déterminer les fonctions qui tendent vers l'infini aussi vite qu'une droite oblique, que nous appellons asymptote et que nous voulons déterminer.

Exemple 15.134.
Déterminer les asymptotes obliques (s'ils existent) de la fonction

$$f(x) = e^{1/x}\sqrt{1 + 4x^2}. \tag{15.445}$$

Tout d'abord nous remarquons que $\lim_{x\to\infty} f(x) = \infty$. Nous sommes donc en présence d'une branche du graphe qui tend vers l'infini. Ensuite,

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} e^{1/x}\sqrt{\frac{1}{x^2} + 4} = 2. \tag{15.446}$$

Donc le graphe de f tend vers l'infini à la même vitesse que le graphe de la fonction $y = 2x$. Nous aurons donc une asymptote oblique de coefficient directeur 2. De façon imagée, nous pouvons penser que le graphe de f et celui de $y = 2x$ sont presque parallèles si x est assez grand. Afin de déterminer l'ordonnée à l'origine de l'asymptote, il nous reste à voir quelle est la « distance » entre le graphe de f et celui de $y = 2x$:

$$\lim_{x\to\infty} f(x) - 2x = \lim_{x\to\infty} e^{1/x}\sqrt{1 + 4x^2} - 2x. \tag{15.447}$$

Cette limite a été calculée dans l'exemple 15.123 et vaut 2.

Nous concluons que le graphe de la fonction f admet l'asymptote

$$y = 2x + 2. \tag{15.448}$$

△

51. Théorème 8.265.

15.12 Développement en série

15.12.1 Série génératrice d'une suite

Soit u_n une suite telle que le rayon de convergence de

$$f(z) = \sum_{n=0}^{\infty} u_n z^n \tag{15.449}$$

soit strictement positif. Alors la série f est la **série génératrice** de la suite (u_n).

Grâce au théorème 15.44 nous pouvons la dériver terme à terme autour de $z = 0$. En utilisant la petite formule (15.140) nous trouvons

$$f^{(l)}(z) = \sum_{n=l}^{\infty} u_n \frac{n!}{(n-l)!} z^{n-l}, \tag{15.450}$$

EqNChVCpF

et donc

$$u_l = \frac{f^{(l)}(0)}{l!}. \tag{15.451}$$

D'où le nom de série génératrice. Cela est évidemment intéressant seulement si nous connaissons une autre forme pour f par ailleurs.

Nous en utiliserons une pour déterminer les partitions d'un nombre en parts fixes, proposition 26.99.

15.12.2 Développement en série et Taylor

DefwmRzKh

Définition 15.135.
*Soit une fonction $f : \mathbb{C} \to \mathbb{C}$ et $z_0 \in \mathbb{C}$. Nous disons que f est **développable en série entière** dans un voisinage de z_0 si il existe une série $\sum_n a_n z^n$ de rayon de convergence $R > 0$ et $r \leqslant R$ tel que*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{15.452}$$

pour tout $z \in B(z_0, r)$.

Proposition 15.136.
Si V est un ouvert dans $\mathbb{C}$ alors l'ensemble des fonctions $V \to \mathbb{C}$ développables en série entière forme une $\mathbb{C}$-algèbre.

Démonstration. Les séries entières passent aux sommes et aux produits en gardant des rayons de convergence non nuls. □

ThoTGPtDj

Proposition 15.137.
*Si f est développable en série entière à l'origine alors elle est C^∞ sur un voisinage de l'origine et le développement est celui de **Taylor** :*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{15.453}$$

pour tout x dans un voisinage de 0.

Démonstration. Si $f(x) = \sum a_n x^n$, nous savons que f est C^1 et que nous pouvons dériver terme à terme (au moins dans un voisinage). De plus le fait de dériver ne change pas le domaine. Par récurrence, la fonction est C^∞ sur le voisinage. En dérivant k fois la série $\sum a_n x^n$ nous trouvons

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) a_n x^{n-k}. \tag{15.454}$$

En calculant en $x = 0$ nous trouvons

$$f^{(k)}(0) = k!a_k, \tag{15.455}$$

d'où le terme général

$$a_k = \frac{f^{(k)}(0)}{k!}. \tag{15.456}$$

□

Si f est une fonction et si la série

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{15.457}$$

converge, alors cette série est la **série de Taylor** de f.

Remarque 15.138.
La série de Taylor d'une fonction n'est pas liée à sa fonction de façon aussi raide qu'on pourrait le croire. Même dans le cas d'une fonction C^∞ il peut arriver que $T_f(x) \neq f(x)$.

Il peut aussi arriver que f ne soit pas développable en série entière.

Exemple 15.139.
Nous considérons la fonction

$$f(x) = \begin{cases} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \tag{15.458}$$

Nous avons

$$f'(x) = \begin{cases} \frac{2}{x^3} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \tag{15.459}$$

Note : pour la seconde ligne nous devons faire explicitement le calcul

$$f'(0) = \lim_{t \to 0} \frac{f(t) - f(0)}{t} = \lim_{t \to 0} \frac{1}{t} e^{-1/t^2} = 0. \tag{15.460}$$

Plus généralement nous avons $f^{(k)}(0) = 0$, et par conséquent la série de Taylor converge (trivialement) vers la fonction identiquement nulle.

Cette fonction n'est donc pas développable en série entière vu qu'il n'existe aucun voisinage de zéro sur lequel la série de f coïncide avec f. △

ExwobBAW

Exemple 15.140.
Développement de $f(x) = \arctan(x)$. Nous savons que

$$f'(x) = \frac{1}{1 + x^2}, \tag{15.461}$$

alors que nous connaissons le développement

$$\frac{1}{1 - x} = \sum_{n=0}^{\infty} x^n \tag{15.462}$$

EqVmnagT

pour tout $x \in B(0, 1)$. Nous avons donc successivement

$$\frac{1}{1 + x} = \sum_{n=0} (-x)^n \tag{15.463a}$$

$$\frac{1}{1 + x^2} = \sum_{n=0} (-1)^n x^{2n} \tag{15.463b}$$

$$\arctan(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{2n + 1} + C. \tag{15.463c}$$

Notons que dans la dernière nous avons évité d'écrire la somme depuis $n = 0$ (qui serait un terme constant) et nous avons écris explicitement « $+C$ ». Étant donné que $\arctan(0) = 0$, nous devons poser $C = 0$ et donc

$$\arctan(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}. \tag{15.464}$$

△

15.12.3 Resommer une série

Nous avons vu comment trouver la série correspondant à une fonction donnée. Un exercice difficile consiste à trouver la fonction qui correspond à une somme donnée.

15.12.3.1 Les sommes du type $\sum_n P(n)x^n$

Pour calculer

$$\sum_{n=0}^{\infty} P(n)x^n \tag{15.465}$$

où P est un polynôme de degré m nous commençons par écrire

$$P(n) = \alpha_0 + \alpha_1(n+1) + \alpha_2(n+1)(n+2) + \cdots + \alpha_m(n+1)\dots(n+m). \tag{15.466}$$

Nous décomposons alors la somme en m sommes de la forme

$$\sum_{n=0}^{\infty} \alpha_k \frac{(n+k)!}{n!} x^n = \alpha_k \left(\sum_{n=0}^{\infty} x^{n+k} \right)^{(k)}. \tag{15.467}$$

Effectuons par exemple

$$\sum_{n=0}^{\infty} x^{n+3} = \frac{1}{1-x} - 1 - x - x^2 \tag{15.468}$$

Notons que dans un usage pratique, ce terme devra être ensuite dérivé trois fois, de telle manière que les termes « correctifs » n'interviennent pas. Cette méthode ne demande donc que de calculer les dérivées successives de $1/(1-x)$.

Exemple 15.141.
Calculons la fonction

$$f(x) = \sum_{n=0}^{\infty} n^3 x^n. \tag{15.469}$$

D'abord nous écrivons

$$n^3 = -1 + 7(n+1) - 6(n+1)(n+2) + (n+1)(n+2)(n+3). \tag{15.470}$$

Nous avons

$$\sum_{n=0}^{\infty} (n+1)x^n = \left(\sum_{n=0}^{\infty} x^{n+1} \right)' = \left(\frac{1}{1-x} - 1 \right)' = \frac{1}{(x-1)^2}. \tag{15.471}$$

De la même façon,

$$\sum_n (n+1)(n+2)x^n = \left(\sum x^{n+2} \right)'' = \frac{-2}{(x-1)^3} \tag{15.472a}$$

$$\sum_n (n+1)(n+2)(n+3)x^n = \frac{6}{(x-1)^4}. \tag{15.472b}$$

En remettant tout ensemble nous obtenons

$$\sum_{n=0}^{\infty} n^3 x^n = -\frac{1}{1-x} + \frac{7}{(x-1)^2} + \frac{12}{(x-1)^3} + \frac{6}{(x-1)^4}. \tag{15.473}$$

Nous pouvons vérifier ce résultat en traçant les deux courbes et en remarquant qu'elles coïncident.

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: n=var('n')
sage: S(x)=sum(  [ n**3*x**n for n in range(0,30)  ]   )
sage: f(x)=-1/(1-x)+7/((x-1)**2)+12/((x-1)**3)+6/( (x-1)**4  )
sage: S(0.1)
0.214906264288980
sage: f(0.1)
0.214906264288981
sage: f.plot(-0.5,0.5)+S.plot(-0.5,0.5)
```

△

15.12.3.2 Les sommes du type $\sum_n x^n/P(n)$

Si $P(n)$ a des racines entières, nous pouvons le décomposer en fractions simples et utiliser la somme

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x). \tag{15.474}$$

Nous avons par exemple

$$\sum_{n=0}^{\infty} \frac{x^n}{n+1} = \frac{1}{x} \sum_{n=0} \frac{x^{n+1}}{n+1} \tag{15.475a}$$

$$= \frac{1}{x} \sum_{n=1}^{\infty} \frac{x^n}{n} = -\frac{\ln(1-x)}{x}. \tag{15.475b}$$

Notez le changement de point de départ de la somme au passage.

Autre exemple :

$$\sum_{n=0}^{\infty} \frac{x^n}{n+3} = \frac{1}{x^3} \left(\sum_{n=1}^{\infty} \frac{x^n}{n} - x - \frac{x^2}{2} \right) \tag{15.476a}$$

$$= -\frac{\ln(x-1)}{x^3} - \frac{1}{x^2} - \frac{1}{2x}. \tag{15.476b}$$

Si le polynôme possède des racines non entières, les choses se compliquent.

Exemple 15.142.
Calculons

$$\sum_{n=0}^{\infty} \frac{x^n}{2n+1}. \tag{15.477}$$

Si $x \geqslant 0$, en posant $t = \sqrt{x}$ nous trouvons

$$\sum_{n=0}^{\infty} \frac{x^n}{2n+1} = \frac{1}{t} \sum_{n=0}^{\infty} \frac{t^{2n+1}}{2n+1}. \tag{15.478}$$

Étudions

$$H(t) = \sum_{n=0}^{\infty} \frac{t^{2n+1}}{2n+1}. \tag{15.479}$$

Nous avons

$$H'(t) = \sum_{n=0}^{\infty} t^{2n} = \sum_{n=0} (t^2)^n = \frac{1}{1-t^2}. \tag{15.480}$$

EqBuPicM

Une primitive de cette fonction est

$$\frac{1}{2}\ln\left|\frac{t+1}{t-1}\right|. \tag{15.481}$$

En $t=0$, cette fonction vaut 0 qui est la bonne valeur. Donc nous avons bien

$$H(t)=\frac{1}{2}\ln\left|\frac{t+1}{t-1}\right|. \tag{15.482}$$

Notons que ce que l'équation (15.480) nous dit est que $H(t)$ est une primitive de $1/(1-t^2)$. Il faut choisir la bonne primitive en fixant une valeur.

Nous avons donc

$$\sum_{n=0}^{\infty}\frac{x^n}{2n+1}=\frac{1}{2\sqrt{x}}\ln\left|\frac{\sqrt{x}+1}{\sqrt{x}-1}\right| \tag{15.483}$$

pour $x>0$. Nous devons encore trouver ce que cela vaut pour $x<0$.

Nous posons successivement $X=-x$ puis $g(X)=f(-X)$. Ce que nous devons calculer est

$$g(t)=\frac{1}{t}\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{2n+1}. \tag{15.484}$$

Si nous posons

$$h(t)=\sum\frac{(-1)^n t^{2n+1}}{2n+1}, \tag{15.485}$$

alors

$$h'(t)=\sum(-1)^n t^{2n}=\sum(-t^2)^n=\frac{1}{1+t^2}, \tag{15.486}$$

par conséquent $h(t)=\arctan(t)$ (cela avait déjà été déduit à l'envers dans l'exemple 15.140).

Au final

$$f(x)=\sum_{n=0}^{\infty}\frac{x^n}{2n+1}=\begin{cases}\frac{1}{2\sqrt{x}}\ln\left|\frac{\sqrt{x}+1}{\sqrt{x}-1}\right| & \text{si } x>0\\ \frac{\arctan(\sqrt{-x})}{\sqrt{-x}} & \text{si } x<0\\ 1 & \text{si } x=0.\end{cases} \tag{15.487}$$

Notons qu'elle est continue en zéro à gauche et à droite.

△

Exemple 15.143.
Nous considérons l'exemple suivant :

$$f(x)=\sum_{n=0}^{\infty}\frac{x^n}{3n+2}. \tag{15.488}$$

Nous posons $t=\sqrt[3]{x}$, et nous substituons :

$$\frac{x^n}{3n+2}=\frac{t^{3n}}{3n+2}=\frac{1}{t^2}\frac{t^{3n+2}}{3n+2}. \tag{15.489}$$

Nous devons étudier la fonction

$$g(t)=\sum_{n=0}^{\infty}\frac{t^{3n+2}}{3n+2} \tag{15.490}$$

Nous avons

$$g'(t)=\sum_{n=0}t^{3n+1}=t\sum_{n=0}t^{3n}=\frac{t}{1-t^3}. \tag{15.491}$$

Notons que $g(0)=0$. △

Exemple 15.144.
Calculer le nombre

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}. \tag{15.492}$$

EqgUyKYe

Nous aurions envie de dire que cela est $f(-1)$ pour la fonction f donnée en (15.487). Le problème est que le rayon de convergence de f étant 1, rien n'est garanti quand au fait que la fonction y soit continue en $x = -1$. En particulier nous devons justifier le fait que

$$\lim_{x\to -1} \sum_n \frac{x^n}{2n+1} = \lim_{x\to -1} \frac{1}{\sqrt{-x}} \arctan(\sqrt{-x}). \tag{15.493}$$

Ce qui nous sauve est le critère d'Abel radial (théorème 15.40). En effet la série

$$\sum \frac{r^n}{2n+1} \tag{15.494}$$

EqAFrXRB

étant convergente avec $r = -1$, la série correspondante est continue sur $[-1, 0]$. Nous pouvons donc calculer la série (15.492) en posant $x = -1$ dans (15.487) :

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}. \tag{15.495}$$

Note : la série (15.494) ne converge pas avec $r = 1$. La fonction f n'est pas continue en $x = 1$. △

ExGxzLlP

Exemple 15.145.
Nous avons

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}. \tag{15.496}$$

En effet si nous désignons par f la somme à gauche, nous trouvons que $f = g'$ avec

$$g(x) = \sum_{n=1}^{\infty} x^n. \tag{15.497}$$

Nous savons par ailleurs que $g(x) = 1/(1-x)$. Par conséquent

$$f(x) = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2}. \tag{15.498}$$

△

15.12.3.3 Sage, primitives et logarithme complexe

ooOPWYooDDSZWx

15.146.
Attention : Sage pourrait nous induire en erreur si nous n'y prenions pas garde. En effet ce que vous ne savez pas mais que Sage sait, c'est que

$$\ln(-1) = i\pi. \tag{15.499}$$

Par conséquent Sage se permet de donner des primitives sans valeurs absolues dans le logarithme :

```
sage: f(x)=1/x
sage: f.integrate(x)
x |--> log(x)
```

La primitive à laquelle on s'attend d'habitude est $\ln(|x|)$. Ici la réponse est correcte parce que si x est négatif nous avons

$$\ln(x) = \ln\big((-1)|x|\big) = \ln(-1) + \ln(|x|). \tag{15.500}$$

Cette fonction est donc décalée de la primitive usuelle seulement de la constante $\ln(-1)$.

Un exemple plus élaboré :

```
sage: h(x)=1/(1-x**2)
sage: H=h.integrate(x)
sage: H
x |--> -1/2*log(x - 1) + 1/2*log(x + 1)
sage: H(0)
-1/2*I*pi
```

Exemple 15.147.
Encore une fois il faut faire attention en demandant la primitive à Sage :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x)=x/(1-x**3)
sage: F=f.integrate(x)
sage: F(0)
-1/3*I*pi - 1/3*sqrt(3)*arctan(1/3*sqrt(3))
```

Cette fois la primitive proposée diffère de celle qu'on cherche de la constante complexe

$$-\frac{\pi}{3}i. \tag{15.501}$$

Mais il y a pire si nous voulons tracer. Nous voudrions définir la fonction $F_2(x) = F(x) - F(0)$. Mathématiquement c'est bien de cette fonction que nous parlons, mais :

```
sage: F2(x)=F(x)-F(0)
sage: F2(x)
1/3*I*pi - 1/3*sqrt(3)*arctan(1/3*(2*x + 1)*sqrt(3)) +
    +1/3*sqrt(3)*arctan(1/3*sqrt(3)) - 1/3*log(x - 1) + 1/6*log(x^2 + x + 1)
sage: F2.plot(x,-0.1,0.1)
verbose 0 (4101: plot.py, generate_plot_points)
        WARNING: When plotting, failed to evaluate function at 200 points.
verbose 0 (4101: plot.py, generate_plot_points)
        Last error message: 'unable to simplify to float approximation'
```

Il refuse de tracer. Pourquoi ? La partie complexe de l'expression de F_2 est mathématiquement nulle, mais elle est en deux parties :

$$\frac{\pi}{3} + \text{la partie imaginaire de} - \frac{1}{3}\ln(x-1). \tag{15.502}$$

Lorsque Sage tente de tracer, il donne à x un certain nombre de valeurs et calcule une *valeur approchée* de $\ln(x-1)$. Cette dernière ne se simplifie pas avec le nombre *exact* $\pi/3$. Sage reste donc avec une partie imaginaire qu'il ne peut pas tracer.

Notez la nuance :

```
sage: ln(-0.1)
-2.30258509299405 + 3.14159265358979*I
sage: ln(-1/10)
I*pi + log(1/10)
```

Du coup nous avons aussi

```
sage: F2(-0.1)
1/3*I*pi - 1/3*sqrt(3)*arctan(0.266666666666667*sqrt(3))
    + 1/3*sqrt(3)*arctan(1/3*sqrt(3)) - 0.0474885065133152 - 1.04719755119660*I
```

△

15.12.3.4 Nombres de Bell

Ici nous montrerions bien le théorème 15.170 sur les nombres de Bell parce que c'est essentiellement un résultat sur les séries entières et leurs manipulations. Hélas, il demande un tout petit peu d'équation différentielle (presque rien). Donc il est postposé jusqu'en page 1514.

15.13 Séries entières de matrices

secEVnZXgf

Nous nous proposons d'étudier des séries de la forme

$$\sum_{k=0}^{\infty} a_k A^k \tag{15.503}$$

où A est une matrice. L'essentiel de la théorie va rester. Nous considérons une norme algébrique (définition 10.63), c'est-à-dire $\|AB\| \leqslant \|A\|\|B\|$.

15.13.1 Rayon de convergence

La notion de rayon de convergence de cette série reste la même : c'est la définition 15.12 qui ne dépend que des coefficients a_k et pas du tout de ce qu'on met à côté dans la somme. Évidemment il faudra montrer que dans le cas des matrices, le nom « rayon de convergence » n'est pas usurpé.

PropFIPooSSmJDQ

Proposition 15.148.
Soit (a_n) une suite dans $\mathbb{C}$ de rayon de convergence R et $A \in \mathbb{M}(n,\mathbb{R})$ une matrice vérifiant $\|A\| < R$. Alors la série

$$\sum_{k=0}^{\infty} a_k A^k \tag{15.504}$$

converge absolument, c'est-à-dire que $\sum_k \|a_k A^k\| < \infty$.

Démonstration. Nous avons les majorations

$$\|a_n A^n\| \leqslant |a_n|\|A^n\| \leqslant |a_n|\|A\|^n. \tag{15.505}$$

Par hypothèse $\|A\| < R$ et R est un supremum, donc il existe r tel que $\|A\| < r < R$ avec $(a_n r^n)$ borné. Nommons M un majorant de la suite $(a_n r^n)$. Alors nous avons

$$\|a_n A^n\| \leqslant |a_n| r^n \frac{\|A\|^n}{r^n} \leqslant M \left(\frac{\|A\|}{r} \right)^n. \tag{15.506}$$

La série du membre de droite converge parce que c'est une série géométrique de raison plus petite que 1, proposition 10.135. □

15.13.2 Convergence et rayon spectral

Le concept de rayon spectral permet aussi de donner des informations sur la convergence de séries de matrices. Pour rappel le rayon spectral d'une matrice est le maximum du module de ses valeurs propres (définition 10.64). Le rayon spectral de la matrice A est noté $\rho(A)$.

La proposition suivante sera redémontrée indépendamment dans le théorème 15.150.

PROPooDJFLooBqqEPT

Proposition 15.149 ([114]).
Si $A \in \mathbb{M}(n,\mathbb{C})$ est telle que $\rho(A) < 1$, alors $A^n \to 0$.

Démonstration. Nous nous plaçons dans une base des espaces caractéristiques [52] de A, c'est-à-dire que nous supposons que la matrice A a la forme

$$A = \begin{pmatrix} \lambda_1\mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_s\mathbb{1} + N_s \end{pmatrix} \tag{15.507}$$

EqWMykgLo

où les λ_i sont les valeurs propres de A et les N_i sont nilpotentes. En effet nous savons que l'espace caractéristique F_{λ_i} est l'espace de nilpolence de $A - \lambda_i\mathbb{1}$. Si nous notons A_i la restriction de A à cet espace, la matrice $N_i = A_i - \lambda_i\mathbb{1}$ est nilpotente. Du coup $A_i = \lambda_I\mathbb{1} + N_i$ et nous avons bien la décomposition (15.507).

Nous avons donc $A^n \to 0$ si et seulement si $(N_i + \lambda_i\mathbb{1})^n \to 0$ pour tout i. Soit donc N nilpotente et $\lambda < 1$ (parce que nous savons que toutes les valeurs propres de A sont inférieures à un). Nous avons

$$(\lambda\mathbb{1} + N)^n = \sum_{k=0}^{n}\binom{n}{k}\lambda^{n-k}N^k = \sum_{k=0}^{r-1}\binom{n}{k}\lambda^{n-k}N^k. \tag{15.508}$$

Nous voyons que le nombre de termes dans la somme ne dépend pas de n. De plus pour chacun de termes, la puissance de N ne dépend pas non plus de n. Le terme

$$\binom{n}{k}\lambda^{n-k} \leqslant P(n)\lambda^{n-k} \tag{15.509}$$

où P est un polynôme tend vers zéro lorsque n devient grand parce que c'est un cas polynôme fois exponentielle. □

THOooMNLGooKETwhh

Théorème 15.150 (Série de Von Neumann, thème 41[392]).
Soit $A \in \mathbb{M}(n,\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Les affirmations suivantes sont équivalentes.

ITEMooCGLSooZsMXSt

(1) $\lim_{k\to\infty} A^k = 0$

ITEMooYBGEooXAzVbD

(2) $\rho(A) < 1$

ITEMooEJSQooTqkBbo

(3) $\sum_{k=0}^{\infty} A^k$ *converge.*

Dans le cas où ces conditions sont vérifiées, nous avons aussi

— $\mathbb{1} - A$ *est inversible,*
— $\sum_{k=0}^{\infty} A^k = (\mathbb{1} - A)^{-1}$

Démonstration. Nous supposons qu'une norme est donnée sur $\mathbb{K}^n$ et nous considérons sur $\mathbb{M}(n,\mathbb{K})$ la topologie associée à la norme subordonnée [53]. Nous subdivisons la preuves en différentes implications.

(1) **(1) implique (2)** Si $\rho(A) \geqslant 1$, en combinant la proposition 12.114 avec la proposition 12.116, nous avons

$$\|A^m\| \geqslant \big(\rho(A)\big)^m \geqslant 1 \tag{15.510}$$

Mais la limite $A^k \xrightarrow{\mathbb{M}(n,\mathbb{K})} 0$ signifie la limite $\|A^k\| \xrightarrow{\mathbb{R}} 0$. Le fait que tous les éléments de la suite soient plus grand que 1 empêche cette limite.

52. Voir le théorème 8.263
53. Si on parle de convergence d'une suite, c'est qu'il y a une topologie quelque part.

(2) **(2) implique (1)**
Vu que $\rho(A) < 1$, il existe $\epsilon > 0$ tel que $\rho(A) + \epsilon < 1$. Par le lemme 12.115 il existe une norme N sur $\mathbb{M}(n, \mathbb{K})$ telle que $N(A) \leqslant \rho(A) + \epsilon < 1$. Notons que cette norme N dépend de A et de ϵ.

Avec cette norme nous avons

$$N(A^k) \leqslant N(A)^k \xrightarrow{\mathbb{R}} 0. \tag{15.511}$$

Cela signifie que $A^k \xrightarrow{N} 0$. L'équivalence entre toutes les normes sur $\mathbb{M}(n, \mathbb{K})$ donne alors la convergence $A^k \xrightarrow{\|.\|} 0$.

Pour une preuve alternative de cette implication, voir la proposition 15.149.

(3) **(3) implique (1)**
La convergence d'une série implique que la norme du terme général converge vers zéro par la proposition 10.100. Nous avons donc $\|A^k\| \to 0$, ce qui signifie $A^k \to 0$, et donc $\rho(A) < 1$ parce que (1) implique (2).

(4) **$\rho(A) < 1$ implique $\mathbb{1} - A$ est inversible**
Si μ est une valeur propre de $\mathbb{1} - A$ alors

$$\det\big((\mathbb{1} - A) - \mu\mathbb{1}\big) = \det\big(A - (1 - \mu)\mathbb{1}\big), \tag{15.512}$$

donc $1 - \mu$ est une valeur propre de A. Donc les valeurs propres de $\mathbb{1} - A$ sont les nombres $1 - \lambda_i$ où les λß sont les valeurs propres de A. Par hypothèse, nous avons $\lambda_i < 1$ pour tout i, donc les valeurs propres de $\mathbb{1} - A$ sont toutes non nulles. Donc $\mathbb{1} - A$ est inversible (pas de noyau).

(5) **Le reste** Nous montrons à présent que si $\rho(A) < 1$ alors $\sum_{k=0}^{\infty} A^k$ converge vers $\mathbb{1} - A$. Pour cela nous savons déjà que $\mathbb{1} - A$ est inversible. Nous posons

$$B_m = \mathbb{1} + A + \ldots + A^m, \tag{15.513}$$

ce qui donne immédiatement $AB_m = A + A^2 + \ldots + A^{m+1}$. Nous avons donc

$$(\mathbb{1} - A)B_m = \mathbb{1} - A^{m+1}. \tag{15.514}$$

Nous savons que $\lim_{m\to 0} A^m = 0$, donc

$$(\mathbb{1} - A)\sum_{k=0}^{\infty} A^k = \lim_{k\to\infty}(\mathbb{1} - A)B_k = \lim_{k\to\infty}(\mathbb{1} - A^{k+1}) = \mathbb{1}. \tag{15.515}$$

Notez au passage que nous avons permuté la somme avec le produit matriciel (voir 10.5.3).

□

15.13.3 Exponentielle et logarithme de matrice

subsecXNcaQfZ

La définition de l'exponentielle dans le cas des matrices est celle sur les algèbres normées non commutatives, 15.66.

PropXFfOiOb

Proposition 15.151.
L'application

$$\begin{aligned}\exp\colon \mathbb{M}(n, \mathbb{R}) &\to \mathbb{M}(n, \mathbb{R})\\ A &\mapsto \sum_{k=0}^{\infty} \frac{A^k}{k!}\end{aligned} \tag{15.516}$$

est une application de classe C^∞. Sa différentielle en zéro est l'identité : $(d\exp)_0 = \mathrm{Id}$.

Démonstration. En ce qui concerne la continuité, nous savons que le rayon de convergence de la suite $\frac{1}{k!}$ est infini ; la proposition 15.130 conclut.

Pour la différentielle, c'est la proposition 15.129 qui nous permet d'écrire

$$d\exp_0(U) = \frac{d}{dt}\Big[\exp(tU)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_{k=0}^{\infty}\frac{t^kU^k}{k!}\Big]_{t=0} = \sum_{k=0}^{\infty}\frac{kt^{k-1}U^k}{k!}\Big|_{t=0} = U \tag{15.517}$$

parce que seul le terme $k = 1$ n'est pas nul. □

Nous avons vu par la proposition 15.131 que toute matrice complexe inversible a un logarithme. Nous allons maintenant parler de logarithme de matrices réelles avec une condition sur la norme. La formule ci-dessous montre explicitement que le logarithme est réel.

$$\begin{aligned}\ln\colon \{A\in\mathbb{M}(n,\mathbb{R}) \text{ tel que } \|A-\mathbb{1}\|<1\} &\to \mathbb{M}(n,\mathbb{R})\\ A &\mapsto \sum_{k=0}^{\infty}(-1)^k\frac{(A-\mathbb{1})^{k+1}}{k+1}.\end{aligned} \tag{15.518}$$

LemQZIQxaB

Lemme 15.152.
Si $\|m\| < 1$ dans $\mathbb{M}(n,\mathbb{R})$, alors nous posons

$$\ln(\mathbb{1}+m) = \sum_{k=0}^{\infty}(-1)^k\frac{m^{k+1}}{k+1}. \tag{15.519}$$

EqIKgMabb

Cette fonction a les propriétés suivantes.

(1) Elle est de classe C^∞.

(2) Elle est un bon logarithme au sens où

$$e^{\ln(\mathbb{1}+m)} = \mathbb{1}+m. \tag{15.520}$$

(3) Elle vérifie l'approximation

$$\ln(1+m) = m+\sigma(m) \tag{15.521}$$

où σ a la propriété que

$$\lim_{k\to\infty} k\sigma\left(\frac{m}{k}\right) = 0. \tag{15.522}$$

Démonstration. Le rayon de convergence de la suite $a_k = \frac{(-1)^k}{k+1}$ est 1. Donc l'application donnée est C^∞ sur $B(0,1)$ par le théorème 15.130.

D'après la formule (15.519) nous avons

$$\sigma(m) = \sum_{l=1}^{\infty}(-1)^l\frac{m^{l+1}}{l+1}. \tag{15.523}$$

Nous avons alors

$$k\sigma(\frac{m}{k}) = \sum_{l=1}^{\infty}(-1)^l\frac{m^{l+1}}{k^l(l+1)}, \tag{15.524}$$

et donc

$$\|k\sigma(\frac{m}{k})\| \leqslant \sum_{l=1}^{\infty}\frac{\|m\|^{l+1}}{k^l(l+1)} \leqslant \frac{1}{k}\sum_{l=1}^{\infty}\frac{\|m\|^{l+1}}{l+1} \overset{k\to\infty}{\to} 0 \tag{15.525}$$

Cela prouve la dernière assertion. □

Proposition 15.153.
Soit V un espace vectoriel de dimension finie et $A \in \mathrm{End}(V)$. Nous considérons la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathrm{End}(V) \\ t &\mapsto e^{tA}. \end{aligned} \tag{15.526}$$

Cette fonction vérifie

$$f'(t) = \big(e^{tA}\big)' = Ae^{tA}. \tag{15.527}$$

Démonstration. Si nous posons $f_k(t) = \frac{t^k A^k}{k!}$ alors la fonction f est la somme : $f = \sum_{k=0}^{\infty} f_k$. Nous allons permuter la somme et la dérivation à l'aide du théorème 15.9. Vu que

$$f_k'(t) = \frac{kt^{k-1}A^k}{k!}, \tag{15.528}$$

la suite des dérivées converge normalement sur $\mathbb{R}$, nous pouvons dériver terme à terme pour obtenir

$$\Big(\sum_{k=0}^{\infty} t^k \frac{A^k}{k!}\Big)' = \sum_{k=0}^{\infty} kt^{k-1}\frac{A^k}{k!} = \sum_{k=1}^{\infty} kt^{k-1}\frac{A^k}{k!} = A\sum_{k=1}^{\infty}\frac{A^{k-1}t^{k-1}}{(k-1)!} = Ae^{tA}. \tag{15.529}$$

Notez le jeu au niveau du point départ de la somme : elle passe de 0 à 1 parce que le terme zéro est nul, mais la simplification $\frac{k}{k!} = \frac{1}{(k-1)!}$ n'a pas de sens pour $k = 0$. □

LemQEARooLRXEef

Lemme 15.154 ([492])**.**
Soit $A \in \mathrm{End}(V)$ où V est un espace vectoriel réel de dimension finie. Si nous notons λ_i ($i = 1, \dots, r$) les valeurs propres distinctes de A alors il existe un polynôme $P \in \mathbb{R}[X]$ tel que

$$\|e^{tA}\| \leqslant P\big(|t|\big)\sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}. \tag{15.530}$$

Démonstration. Le polynôme caractéristique de A se note, d'après le corolaire 12.110 de la façon suivante :

$$\chi_A(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i} \tag{15.531}$$

où m_i est la multiplicité de la valeur propre λ_i. Le lemme des noyaux 8.92 nous dit qu'en posant

$$V_i = \ker(A - \lambda_i \mathbb{1})^{m_i} \tag{15.532}$$

nous avons $V = \bigoplus_{i=1}^{r} V_i$. Nous nommons $p_i\colon V \to V$ la projection canonique de E sur V_i ainsi que x_i la composante de $x \in V$ dans l'espace caractéristique V_i et nous posons $A_i = p_i \circ A$. Les espaces caractéristiques sont stables par A (lemme 8.260), donc $(Ax_i)_i = Ax_i$. Par conséquent $\sum_i A_i p_i = A$ parce que

$$\big(\sum_i p_i A p_i\big)(x) = \sum_i (Ax_i)_i = \sum_i Ax_i = A\sum_i x_i = Ax. \tag{15.533}$$

En ce qui concerne les puissances de A nous avons de même

$$A_i^n x_i = A_i \underbrace{A_i^{n-1}x_i}_{\in V_i} = AA_i^{n-1}x_i = A^n x_i, \tag{15.534}$$

et donc

$$\sum_{i=1}^{r} A_i^n p_i = A^n. \tag{15.535}$$

En particulier,

$$e^{tA} = \sum_i e^{tA_i}p_i. \tag{15.536}$$

EqPVIooGxwFBH

C'est de cette exponentielle de matrice que nous devons étudier la norme.

La décomposition de Dunford du théorème 8.265 est toujours un bon plan pour traiter avec les exponentielles : nous avons $A = s + n$ avec

$$s = \sum_k \lambda_k p_k, \quad n = \sum_k (A - \lambda_k \mathbb{1}) p_k. \tag{15.537}$$

Nous montrons que la décomposition de Dunford de $p_i A$ est $p_i A = p_i s + p_i n$. Nous avons

$$p_i s = \sum_k \lambda_k p_i p_k = \lambda_i p_i \tag{15.538}$$

qui est bien diagonalisable. De plus les espaces caractéristiques sont stables par n, donc $p_i n$ est nilpotent. Enfin ils commutent :

$$[p_i s, p_i n] = \lambda_i (p_i n - p_i n p_i). \tag{15.539}$$

EqNJJooDxKkxn

Vu que n préserve les espaces caractéristiques, lorsque $v \in V_k$ avec $k \neq i$ nous avons $p_i n p_i v = 0$ et $p_i n v = 0$. Mais si $v \in V_i$ alors

$$p_i n p_i v = p_i n v = n v \tag{15.540}$$

et $p_i n v = n v$, donc les opérateurs $p_i n$ et $p_i n p_i$ sont égaux et (15.539) donne bien zéro. En ce qui concerne l'exponentielle de A_i nous avons

$$e^{p_i A} = e^{p_i s} e^{p_i n} = e^{\lambda_i p_i} \exp\big((A - \lambda_i \mathbb{1}) p_i\big). \tag{15.541}$$

Nous pouvons maintenant sérieusement nous attaquer à la norme de e^{tA} de l'équation (15.536). D'abord nous avons $\|p_i\| = 1$ parce que l'opérateur p_i est l'identité sur au moins un vecteur (en fait tout ceux de l'espace caractéristique V_i). En utilisant les propriétés de la norme opérateur [54], nous trouvons dans un premier temps [55] :

$$\|e^{tA}\| \leqslant \sum_{i=1}^{r} \|e^{tA_i}\| \leqslant \sum_{i=1}^{r} |e^{t\lambda_i}| \underbrace{\sum_{k=0}^{m_i} \frac{|t|^k}{k!} \|A - \lambda_i \mathbb{1}_i\|^k}_{=P_i(|t|}} \tag{15.542}$$

où $\mathbb{1}_i$ est l'opérateur identité sur V_i. Petit détail dans le calcul :

$$\|e^{\lambda_i p_i}\| \leqslant \sum_{l=0}^{\infty} \frac{\lambda_i^l}{k!} \|p_i\|^l = e^{\lambda_i}. \tag{15.543}$$

Notons que tous les termes de $P_i(|t|)$ et $P_i\big(|t|\big)$ sont positifs, de telle sorte que nous pouvons majorer en ajoutant des termes partout. À la place d'avoir $P_i(|t|)$ comme coefficient de $|e^{t\lambda_i}|$ nous majorons en mettant $\sum_{j=1}^{r} P_j(|t|)$ comme coefficient :

$$\|e^{tA}\| \leqslant \sum_{i=1}^{r} |e^{t\lambda_i}| P_i\big(|t|\big) = \sum_{i=1}^{r} |e^{t\lambda_i}| \sum_{j=1}^{r} P_j\big(|t|\big) = P\big(|t|\big) \sum_{i=1}^{r} e^{t \operatorname{Re}(\lambda_i)}. \tag{15.544}$$

L'arrivée de la partie réelle est une égalité usuelle pour les nombres complexes : $|e^{a+bi}| = e^a |e^{bi}| = e^a$. □

54. Surtout le fait que ce soit une norme d'algèbre, lemme 10.69.
55. Si les valeurs propres de A sont λ_i, celles de tA sont $t\lambda_i$.

15.13.4 Calcul effectif de l'exponentielle d'une matrice

SUBSECooGAHVooBRUFub

Nous reprenons l'exemple de [493]. Soit A une matrice dont le polynôme minimum s'écrit

$$P(X) = (X-1)^2(X-2). \tag{15.545}$$

Par le théorème 8.92 de décomposition des noyaux nous avons

$$E = \ker(A-1)^2 \oplus \ker(A-2). \tag{15.546}$$

En suivant les notations de ce théorème nous avons $P_1(X) = (X-1)^2$, $P_2(X) = X-2$ et

$$Q_1(X) = X-2 \tag{15.547a}$$
$$Q_2(X) = (X-1)^2. \tag{15.547b}$$

En posant

$$\begin{aligned} R_1(X) &= -X \\ R_2(X) &= 1, \end{aligned} \tag{15.548}$$

nous avons une solution de Bezout :

$$R_1Q_1 + R_2Q_2 = 1. \tag{15.549}$$

Conformément au théorème 8.92(3), le projecteur p_i sur $\ker\big(P_i(A)\big)$ est $(R_iQ_i)(A)$:

$$\begin{aligned} p_1 &= \mathtt{proj}_{\ker(A-1)^2} = -A(A-2) \\ p_2 &= \mathtt{proj}_{\ker(A-2)} = (A-1)^2 \end{aligned} \tag{15.550}$$

Passons maintenant au calcul de l'exponentielle [56]. Nous avons évidemment

EQooVIOWopACVAoZ

$$e^A = e^Ap_1 + e^Ap_2. \tag{15.551}$$

Afin de nous adapter au projecteurs que nous avons en main, nous écrivons $A = \mathbb{1} + (A-\mathbb{1})$ et nous utilisons la proposition 15.69(4) :

$$e^A = e^{\mathbb{1}+(A-\mathbb{1})} = e^{\mathbb{1}}e^{A-\mathbb{1}} = ee^{A-\mathbb{1}}. \tag{15.552}$$

Pour tout x, nous avons $(A-\mathbb{1})^kp_1(x) = 0$ dès que $k \geqslant 2$. Donc nous avons

$$e^Ap_1(x) = ee^{A-1}p_1(x) = e\sum_k \frac{(A-\mathbb{1})^k}{k!}p_1(x) = ep_1(x) + e(A-\mathbb{1})p_1(x) = eAp_1(x). \tag{15.553}$$

En ce qui concerne p_2 nous faisons de même en partant de $e^A = e^{2\mathbb{1}}e^{A-2\mathbb{1}}$:

$$e^Ap_2(x) = e^2\sum_{k=0}^{\infty}\frac{(A-\mathbb{1})^k}{k!}p_2(x) = e^2p_2(x). \tag{15.554}$$

En recollant les morceaux avec (15.551),

$$e^A(x) = eAp_1(x) + e^2p_2(x) = -eA^2(A-2)x + e^2(A-1)^2x. \tag{15.555}$$

Au final nous avons prouvé que

$$e^A = -eA^2(A-2) + e^2(A-1)^2. \tag{15.556}$$

56. Définition 15.66. Thème 50

15.13.5 Opérateur d'inversion

PROPooSGVWooQupJxn

Proposition 15.155 ([494]).
Soit un espace de Banach réel E. Nous considérons l'espace vectoriel $\mathrm{End}(E)$ muni de la norme opérateur.

(1) *La partie Ω des opérateurs inversibles*[57] *est ouverte.*

(2) *L'opérateur d'inversion*

$$\begin{aligned} i\colon \Omega &\to \Omega \\ f &\mapsto f^{-1} \end{aligned} \tag{15.557}$$

est analytique[58].

Note : pour prouver ça, il faudra surement considérer une adaptation aux opérateur du théorème 15.150.

PROPooFXPZooKUEarD

Proposition 15.156 ([1]).
Soit un espace de Banach réel E. Nous considérons l'espace vectoriel $\mathrm{End}(E)$ muni de la norme opérateur.

(1) *La partie $\mathrm{GL}(E)$ des opérateurs inversibles est ouverte.*

(2) *L'opérateur d'inversion*

$$\begin{aligned} i\colon \mathrm{GL}(E) &\to \mathrm{GL}(E) \\ f &\mapsto f^{-1} \end{aligned} \tag{15.558}$$

est analytique.

Démonstration. La partie $\mathrm{GL}(E)$ est formée des opérateurs inversibles qui sont continus et dont l'inverse est continu. Or un opérateur est continu si et seulement si il est borné (proposition 10.70).

Donc $\mathrm{GL}(E)$ est une intersection de deux ouverts de $\mathrm{End}(E)$: les opérateurs linéaires continus et les opérateurs linéaires inversibles. Donc $\mathrm{GL}(E)$ est ouvert.

Vu que $\mathrm{GL}(E)$ est ouvert et contenu dans Ω, une application analytique sur Ω est analytique sur $\mathrm{GL}(E)$. □

15.14 Lemme de Borel

15.14.1 Fonctions plateaux, Urysohn, partition de l'unité

subsecOSYAooXXCVjv

Une fonction plateau sur $\mathbb{R}$ est donnée dans le lemme d'Urysohn 15.165. Sur $\mathbb{R}^n$, il y a le lemme d'Urysohn 27.74.

Vous voulez une fonction de classe C^∞ nulle sur un ouvert, mais qui n'est pas nulle partout ? En voici une.

LEMooFLUSooKaZRRY

Lemme 15.157.
La fonction

$$\varphi(x) = \begin{cases} e^{-1/x} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases} \tag{15.559}$$

est de classe C^∞.

Démonstration. Pour tout polynôme P nous avons la limite[59]

$$\lim_{x\to 0^+} \frac{e^{-1/x}}{P(x)} = 0. \tag{15.560}$$

57. En dimension finie, Ω est $\mathrm{GL}(E)$, mais en dimension infinie, la partie $\mathrm{GL}(E)$ contient seulement les opérateurs inversibles continue d'inverse continu. Voir la définition 10.272.

58. Je ne sais même pas comment on définit le mot « analytique » dans le cas d'une application entre espaces de Banach. Si vous le savez, écrivez-moi.

59. Voir la proposition 15.108(3).

De là, en écrivant les dérivées successives de φ, il est facile de voir qu'elles sont continues en $x = 0$. □

LEMooRVSIooKcpWoK

Lemme 15.158.
Soit $m > 0$. Il existe une application $\psi_m \in C^\infty(\mathbb{R})$ vérifiant

$$\psi_m(x) = \begin{cases} 1 & si\ x < 0 \\ 0 & si\ x > m \\ positive & si\ x \in [0, m]. \end{cases} \tag{15.561}$$

Démonstration. Nous partons de la fonction φ du lemme 15.157. Ensuite nous considérons

$$\psi_m(x) = 1 - \frac{\int_0^x \varphi(t)dt}{\int_0^m \varphi(t)dt} \tag{15.562}$$

Cette fonction est encore de classe C^∞. En effet, le dénominateur $\int_0^m \varphi(t)dt$ est un simple nombre strictement positif sans histoires tandis que la proposition 14.286 dit que $x \mapsto \int_0^x \varphi(t)dt$ est une primitive de φ. Vu que φ est déjà de classe C^∞, sa primitive l'est également. □

LEMooFFPVooDKGUAp

Lemme 15.159 ([243])**.**
Soient $a < b$ dans $\mathbb{R}$. Il existe une fonction $f \in C^\infty(\mathbb{R})$ telle que

(1) $0 \leqslant f \leqslant 1$,
(2) $f = 0$ *sur* $]-\infty, a]$,
(3) $f = 1$ *sur* $[b, \infty[$.

Démonstration. C'est une variation sur le thème de la fonction du lemme 15.158 ; il s'agit de la retourner, dilater et décaler. Posez successivement $f_1(x) = \psi_m(-x)$, $f_2(x) = f_1\big(mx/(b-a)\big)$ et $f_3(x) = f_2(x-a)$ et je crois que le compte est bon. La fonction f_3 est celle que nous cherchons. □

PROPooAZJZooTYWjzb

Proposition 15.160.
Soient $a < b < c < d$ dans $\mathbb{R}$. Il existe une fonction $f \in C^\infty(\mathbb{R})$ à valeurs positives telle que

(1) $f(x) = 1$ *si* $x \in [b, c]$
(2) $\text{supp}(f) \subset [a, d]$.

Démonstration. Nous considérons la fonction ψ_m du lemme 15.158, et nous considérons les fonctions

$$f_1(x) = \psi_{d-c}(x - c) = \begin{cases} 1 & \text{si } x < c \\ 0 & \text{si } x > d \\ \text{positive} & \text{si } x \in [c, d]. \end{cases} \tag{15.563a}$$

$$f_2(x) = \psi(b-a)(b-x) = \begin{cases} 0 & \text{si } x < a \\ 1 & \text{si } x > b \\ \text{positive} & \text{si } x \in [a, b]. \end{cases}, \tag{15.563b}$$

et finalement la fonction suivante répond à la question des fonctions plateaux sur $\mathbb{R}$:

$$f(x) = f_1(x)f_2(x). \tag{15.564}$$

EqIHAEoqXifcU

□

Une variation sur le même thème est l'existence de fonctions infiniment dérivables à support compact, c'est-à-dire des fonctions dans $C_c^\infty(\mathbb{R}^d) = \mathscr{D}(\mathbb{R}^d)$.

LEMooLHIFooWpbauN

Lemme 15.161 ([1]).
Pour tout $n \in \mathbb{N}$ nous avons

$$\lim_{x\to 1^-} \frac{e^{-1/(1-x^2)}}{(1-x^2)^n} = 0. \tag{15.565}$$

Démonstration. D'abord le lemme 15.107 nous indique que

$$\lim_{x\to\infty} x^n e^{-x} = 0. \tag{15.566}$$

Nous prouvons ensuite que

$$\lim_{x\to 0^+} \frac{e^{-1/x}}{x^n} = 0. \tag{15.567}$$

Pour cela nous considérons $\epsilon > 0$. Soit $M > 0$ tel que $x^n e^{-x} < \epsilon$ pour tout $x > M$. Nous considérons δ tel que $0 < x < \delta$ implique $1/x > M$.

Pour de tels x, nous avons $e^{-1/x}/x^n < \epsilon$.

Nous montrons enfin que

$$\lim_{x\to 1^+} \frac{e^{-1/(1-x^2)}}{(1-x^2)} = 0. \tag{15.568}$$

Pour cela, soit $\epsilon > 0$. Soit δ tel que $0 < x < \delta$ implique $e^{-1/x}/x^n < \epsilon$. Soit δ' tel que $1 < x < 1 + \delta'$ implique $1 - x^2 < \delta$. Avec ça, nous avons

$$1 < x < 1 + \delta' \Rightarrow \frac{e^{-1/(1-x^2)}}{(1-x^2)^n} < \epsilon. \tag{15.569}$$

□

PROPooAHLKooMFMgFq

Proposition 15.162.
La fonction $\xi\colon \mathbb{R}^d \to \mathbb{R}$ donnée par

$$\xi(x) = \begin{cases} e^{-1/(1-\|x\|^2)} & \text{si } x \in B(0,1) \\ 0 & \text{sinon.} \end{cases} \tag{15.570}$$

EqQBYNEMu

est de classe C^∞ et à support compact.

Démonstration. Le fait que le support soit compact est le fait qu'un support est toujours fermé (c'est dans la définition) et que le support de ξ est borné, contenu dans $B(0,1)$. Le vrai travail est de montrer que cette fonction est de classe C^∞.

Nous commençons par voir en dimension 1. C'est à dire la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} e^{-1/(1-x^2)} & \text{si } |x| < 1 \\ 0 & \text{sinon} \end{cases} \end{aligned} \tag{15.571}$$

Le lemme 15.161 dit que f est continue. En ce qui concerne les dérivées de f, vous pouvez montrer par récurrence que

$$f^{(n)}(x) = \begin{cases} \frac{P_(x)}{(x^2-1)} f(x) & \text{si } |x| < 1 \\ 0 & \text{sinon} \end{cases} \tag{15.572}$$

où P_n est un polynôme. Le lemme 15.161 (encore lui) nous indique que $f^{(n)}$ est continue.

Pour que ξ soit de classe C^∞, il suffit maintenant d'invoquer la proposition 7.167 qui dit que la norme est une application de classe C^∞. □

CORooHHZXooXmwGmC

Corolaire 15.163.
Il existe une fonction $\xi \in \mathscr{D}\big(B(0,R)\big)$ telle que $\int_{\mathbb{R}} \xi(x)dx = 1$.

Démonstration. Prenez la fonction de la proposition 15.162. Si $R < 1$, faites une redéfinition $\xi_2 = \xi(\lambda x)$ pour que le support soit dans $B(0, R)$. Ensuite, si l'intégrale n'est pas 1, encore une redéfinition $\xi_3 = \mu\xi_2$. □

Il ne faudrait pas croire pour autant que tout est toujours rose au pays des fonctions C^∞ à support compact.

Proposition 15.164.
Si ϕ est une fonction C^∞ non nulle à support compact sur $\mathbb{R}$, alors ϕ'/ϕ n'est pas bornée.

Plus présisément, nous posons $D = \{x \in \mathbb{R}$ tel que $\phi(x) \neq 0\}$. Alors la fonction

$$\begin{aligned} f\colon D &\to \mathbb{R} \\ x &\mapsto \phi'(x)/\phi(x) \end{aligned} \tag{15.573}$$

n'est pas bornée.

Démonstration. Quitte à décaller et à multiplier, nous supposons que $\phi(0) = 1$. Sinon vous considérez x_0 tel que $\phi(x_0) = y_0 \neq 0$ et vous adaptez tout le reste de la démonstration à vos frais.

Vu que ϕ est continue, la partie $Z = \{x \geqslant 0 \in \mathbb{R}$ tel que $\phi(x) = 0\}$ est fermée. Elle est également bornée vers le bas par 0. Donc elle possède un minimum que nous nommons a :

$$a = \min\{x \geqslant 0 \text{ tel que } \phi(x) = 0\}. \tag{15.574}$$

Les valeurs de ϕ qui vons nous intéresser sont :

$$\phi(x) = \begin{cases} > 0 & \text{si } x \in [0, a[\\ = 0 & \text{si } x = a. \end{cases} \tag{15.575}$$

Enfin, nous posons

$$M = \sup\{\phi'(x)/\phi(x)\}_{x\in[0,a[}. \tag{15.576}$$

Nous supposons que $M < \infty$ [60].

Nous posons

$$\begin{aligned} F\colon [0, a[&\to \mathbb{R} \\ x &\mapsto \int_0^x \frac{\phi'(t)}{\phi(t)}dt. \end{aligned} \tag{15.577}$$

Vu que ϕ est de classe C^∞ et que x est dans $[0, a[$ sur lequel ϕ ne s'annule pas, cette intégrale n'a rien d'exceptionnel.

Nous pouvons majorer F de la façon suivantes :

$$F(x) = \int_0^x \frac{\phi'(t)}{\phi(t)}dt \leqslant \int_0^x M dt = Mx < Ma. \tag{15.578}$$

La proposition 14.286 lie primitive et intégrale ; celle de $\phi'(x)/\phi(x)$ est $\ln\big(\phi(x)\big)$ nous avons donc

$$F(x) = \ln(\phi(x)) - \ln(\phi(0)) = \ln\big(\phi(x)\big). \tag{15.579}$$

EQooCEVKooLgXiaj

Mais vu que $\lim_{x\to a}\phi(x) = 0$, nous avons $\lim_{x\to a} F(x) = \infty$, d'où la contradiction. □

Notez que l'expression 15.579 montre que $\phi'(x)/\phi(x)$ tend vers $-\infty$, ce qui est logique : la dérivée est négative alors que ϕ reste positive. Ce que dit la proposition est qu'une fonction C^∞ à support compact tend plus vite vers zéro que sa dérivée.

60. Vous pouvez ne pas supposer cela et voir le reste de la preuve comme une démonstration que $M = \infty$. Ici nous allons faire par l'absurde et montrer une contradiction en supposant que $M < \infty$.

15.14.2 Lemme de Urysohn

Le lemme d'Urysohn comprend de nombreuses variantes plus ou moins générales[61]. En voici une parmi les plus simples.

LEMooECTNooKagaRU

Lemme 15.165 (Lemme d'Urysohn)**.**
Soient un ouvert $\mathcal{U}$ de $\mathbb{R}$ et un compact K inclus dans $\mathcal{U}$. Il existe une fonction $f\colon \mathbb{R} \to [0,1]$ de classe C^∞ telle que

(1) $f(x) = 1$ *pour* $x \in K$,

(2) $\operatorname{supp}(f) \subset \mathcal{U}$.

Démonstration. Vu que K est compact, il est borné (théorème 9.24). Nous posons $b = \min(K)$ et $c = \max(K)$. En particulier b et c sont des éléments de K et donc de $\mathcal{U}$. Comme $\mathcal{U}$ est ouvert, il existe des boules centrées en b et c contenues dans $\mathcal{U}$. Soit r le rayon de telles boules :

$$B(b,r) \subset \mathcal{U} \tag{15.580a}$$
$$B(c,r) \subset \mathcal{U}. \tag{15.580b}$$

Nous posons $a = b - r \in \mathcal{U}$ et $d = c + r \in \mathcal{U}$. Nous considérons à présent la fonction plateau f de la proposition 15.160. Elle est de classe C^∞ et vérifie $f(x) = 1$ pour $x \in [b,c] = K$ ainsi que $f(x) = 0$ hors de $[a,d] \subset \mathcal{U}$. □

15.166.
Notons que la fonction du lemme d'Urysohn 15.165 n'épouse pas spécialement très bien la forme de K. Si par exemple $K = [0,1] \cup [10,11]$, la fonction f sera égale à 1 au moins sur $[0,11]$.

LEMooGZXYooZqFzyc

Lemme 15.167 ([1])**.**
Si une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ est intégrable, et si $s \in \mathbb{R}$ alors l'application

$$\begin{aligned} f_s\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{s^n} f(x/s) \end{aligned} \tag{15.581}$$

est intégrable et

$$\int_{\mathbb{R}^n} f_s = \int_{\mathbb{R}^n} f. \tag{15.582}$$

Démonstration. Il s'agit d'appliquer le théorème 14.281(2) au changement de variable $\phi(x) = x/s$. Nous avons $f_s = \frac{1}{s^n} f \circ \phi$, et

$$(d\phi)_a v = \frac{1}{s} v, \tag{15.583}$$

de telle sorte que le déterminant de $d\phi$ soit une constante et vaille $\det(d\phi) = 1/s^n$. Enfin ϕ est une bijection de $\mathbb{R}^n$. Nous avons tous les ingrédients pour faire le changement de variables :

$$\int_{\mathbb{R}^n} f_s(x)dx = \int_{\mathbb{R}^n} (f \circ \phi)(x) \frac{1}{s^n} dx = \int_{\phi(\mathbb{R}^n)} f(xdx). \tag{15.584}$$

□

Pour une fonction plateau sur $\Omega \subset \mathbb{R}^n$, voir le lemme d'Urysohn, proposition 27.74.

61. Pour une version sur $\mathbb{R}^n$, voir la proposition 27.74.

15.14.3 Le lemme de Borel

LemRENIIEL

Lemme 15.168 (Lemme de Borel[104]).
Soit (a_n) une suite dans $\mathbb{R}$. Il existe une fonction $u \in C^\infty(\mathbb{R})$ telle que $u^{(k)}(0) = a_k$ pour tout $k \geqslant 0$.

Démonstration. Soit $\varphi \in C_c^\infty(\mathbb{R})$ une fonction telle que $\varphi(x) = 1$ si $|x| \leqslant \frac{1}{2}$ et telle que $\text{supp}(\varphi) \subset]-1,1[$.

Nous commençons par considérer une suite de réels strictement positifs (λ_k) dont nous fixerons une valeur précise plus tard, et nous posons

$$f_k(x) = \varphi(\lambda_k x)\frac{a_k}{k!}x^k. \tag{15.585}$$

Nous allons étudier la convergence et les propriétés de $u(x) = \sum_{k=0}^{\infty} f_k(x)$.

Calculons (formellement) la m^{e}dérivée de f_k :

$$f_k^{(m)}(x) = \frac{a_k}{k!}\sum_{l=0}^{m}\binom{m}{l}\lambda_k^{m-l}\varphi^{(m-l)}(\lambda_k x)(x^k)^{(l)} \tag{15.586a}$$

$$= a_k\sum_{l=0}^{m}\binom{m}{l}\lambda_k^{m-l}\varphi^{(m-l)}(\lambda_k x)\frac{x^{k-l}}{(k-l)!}. \tag{15.586b}$$

Notons que nous travaillons à m fixé et que nous ne nous intéressons qu'aux termes avec k assez grand ; nous pouvons donc supposer $k \geqslant m$. De toutes façons pour $\sum_{k=0}^{m} f_k$, on a la classe C^∞, et la permutation de la somme avec tout ce qu'on veut. Vu que φ est continue à support compact nous pouvons poser

$$M_m = \max_{0\leqslant j\leqslant m}\|\varphi^j\|_\infty = \max_{0\leqslant j\leqslant m}\max_{x\in\mathbb{R}}|\varphi^{(j)}(x)|. \tag{15.587}$$

Nous continuons en nous fixant un $x \in \mathbb{R}$ et un $k \geqslant m$.

Si $|x| > \frac{1}{\lambda_k}$, alors $\varphi^{(m)}(\lambda_k x) = 0$ parce que $\lambda_k x$ est strictement hors du support de φ qui est $]-1,1[$. Donc pour $|x| > \frac{1}{\lambda_k}$.

Si par contre $|x| \leqslant \frac{1}{\lambda_k}$, nous avons les majorations

$$|f_k^{(m)}(x)| \leqslant |a_k|\sum_{l=0}^{m}\binom{m}{l}|\lambda_k|^{m-l}\underbrace{\varphi^{(m-l)}(\lambda_k x)}_{\leqslant M_m}\frac{1}{(k-l)!}\underbrace{|x|^{k-l}}_{\leqslant(1/\lambda_k)^{k-l}} \tag{15.588a}$$

$$\leqslant |a_k|M_m|\lambda_k|^{m-k}\frac{1}{(k-m)!}\sum_{l=0}^{m}\binom{m}{l} \tag{15.588b}$$

$$\leqslant \frac{|a_k|M_m|\lambda_k|^{m-k}2^m}{(k-m)!} \tag{15.588c}$$

$$= \frac{|a_k|M_m 2^m}{(k-m)!|\lambda_k|^{k-m}} \tag{15.588d}$$

EqQSPUaun

où pour faire disparaitre la somme de coefficients binomiaux, nous avons remarqué que $\sum_{l=0}^{m}\binom{m}{l}$ est le nombre total de termes dans le développement de $(a+b)^m$, c'est-à-dire 2^m. Nous voulons, pour m fixé, étudier la convergence de la somme de cela. Notons que le 2^m n'a en particulier strictement aucune importance parce qu'on travaille à m fixé.

Nous fixons maintenant la valeur des λ_k :

$$\lambda_k = \max\{|a_k|, 1\}. \tag{15.589}$$

Avec cela, en nous souvenant que nous n'étudions que les termes $k > m$, le dénominateur de (15.588d) est réellement croissant en k, donc nous avons la majoration

$$|f_k^{(m)}(x)| \leqslant \frac{M_m 2^m}{(k-m)!}. \tag{15.590}$$

Au final nous avons

$$\|f_k^{(m)}\|_\infty \leqslant \frac{2^m M_m}{(k-m)!}. \tag{15.591}$$

Et la somme de cela converge sans difficultés. Donc la série

$$u(x) = \sum_{k=0}^{\infty} f_k^{(m)}(x) \tag{15.592}$$

converge normalement et donc uniformément sur $\mathbb{R}$. Nous pouvons alors permuter la somme et la dérivation par le théorème 15.3. Donc

$$u^{(m)} = \sum_{k=0}^{\infty} f_k^{(m)} \tag{15.593}$$

est continue. En particulier, pour évaluer en zéro, on peut faire

$$u^{(m)}(0) = \sum_{k=0}^{\infty} f_k^{(m)}(0). \tag{15.594}$$

Nous avons

$$f_k(x) = \varphi(\lambda_k x)\frac{a_k}{k!}x^k. \tag{15.595}$$

Pour calculer la dérivée en zéro, il suffit de la calculer sur un voisinage sur lequel $\varphi(\lambda_k x)$ est la constante 1 ; un tel voisinage existe pour tout k. À ce moment le calcul est classique :

$$f_k^{(m)}(x) = \begin{cases} a_k & \text{si } k = m \\ 0 & \text{sinon.} \end{cases} \tag{15.596}$$

Finalement nous avons bien

$$u^{(m)}(0) = \sum_{k=0}^{\infty} f_k^{(m)}(0) = a_k. \tag{15.597}$$

□

Remarque 15.169.
Pour prouver le lemme de Borel, la première chose qui passe par la tête est la fonction toute simple

$$u(x) = \sum_{k=0}^{\infty} \frac{a_k}{k!}x^k. \tag{15.598}$$

Évidemment si on calcule les dérivées successives de cette fonction, nous trouvons les bons résultats. Le problème est la convergence. Rien qu'en prenant $a_k = k!k^k$, la série ne converge pour aucun x positif. L'idée de multiplier chacun de f_k par une fonction plateau sur un petit intervalle autour de zéro a plusieurs avantages. D'abord on conserve les dérivées correctes parce qu'on ne touche pas à la valeur des f_k sur un petit voisinage. Ensuite cela ne modifie pas la continuité ; et enfin en multipliant par $\varphi(\lambda_k x)$, ça calme méchamment les divergences parce que $\lambda_k x$ passe vite au dessus de 1 (et donc en dehors du support de φ) si λ_k est grand. D'où le fait qu'il soit normal que les λ_k soient de l'ordre des a_k.

15.15 Nombres de Bell

ThoYFAzwSg

Théorème 15.170 (Nombres de Bell[104])**.**
Soient $n \geqslant 1$ et B_n le nombre de partitions distinctes de l'ensemble $\{1, \dots, n\}$ avec la convention que $B_0 = 0$. Alors

(1) *La série entière*

$$\sum_{n=0}^{\infty} \frac{B_n}{n!} x^n \tag{15.599}$$

a un rayon de convergence $R > 0$ et sa somme est donnée par

$$f(x) = e^{e^x - 1} \tag{15.600}$$

pour tout $x \in]-R, R[$.

(2) *Pour tout $k \in \mathbb{N}$,*

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}. \tag{15.601}$$

(3) *Le rayon de convergence de la série (15.599) est en réalité infini : $R = \infty$.*

Démonstration. (1) Soient $n \geqslant 1$ et $0 \leqslant k \leqslant n$. Nous notons E_k l'ensemble des partitions de $\{1, \dots, n+1\}$ pour lesquelles le « paquet » contenant $n+1$ soit de cardinal $k+1$. Calculons le cardinal de E_k.

Pour construire un élément de E_k, il faut d'abord prendre le nombre $n+1$ et lui adjoindre k éléments choisis dans $\{1, \dots, n\}$, ce qui donne $\binom{n}{k}$ possibilités. Ensuite il faut trouver une partition des $(n+1) - (k+1) = n - k$ éléments restants, ce qui fait B_{n-k} possibilités. Donc

$$\text{Card}(E_k) = \binom{n}{k} B_{n-k}. \tag{15.602}$$

L'intérêt des ensembles E_k est que $\{E_0, \dots, E_n\}$ est une partition de l'ensemble des partitions de $\{1, \dots, n+1\}$, c'est-à-dire que $B_{n+1} = \sum_{k=0}^{n} \text{Card}(E_k)$, ce qui va nous donner une relation de récurrence pour les B_n :

$$B_{n+1} = \sum_{k=0}^{n} \text{Card}(E_k) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} = \sum_{l=0}^{n} \binom{n}{n-l} B_l = \sum_{l=0}^{n} \binom{n}{l} B_l. \tag{15.603}$$

où nous avons utilisé un petit changement de variables $l = n-k$. Afin d'étudier la convergence de la série (15.599), nous allons montrer par récurrence que pour tout n, $B_n < n!$. D'abord pour $n = 0$ c'est bon : $B_1 = 1$ parce que la seule partition de $\{1\}$ est $\{1\}$. Supposons que l'inégalité soit vraie pour une certaine valeur k, et montrons qu'elle est vraie pour la valeur $k+1$.

$$B_{k+1} = \sum_{l=0}^{n} \binom{n}{l} B_l \leqslant \sum_{l=0}^{n} \binom{n}{l} l! = \sum_{l=0}^{n} \frac{n!}{l!(n-l)!} l! = n! \sum_{l=0}^{n} \underbrace{\frac{1}{(n-l)!}}_{\leqslant 1} \leqslant n!(n+1) = (n+1)! \tag{15.604}$$

où nous avons utilisé la formule (3.77) du coefficient binomial : $\binom{n}{l} = \frac{n!}{l!(n-l)!}$.

Donc pour tout $x \in \mathbb{R}$ nous avons

$$0 \leqslant \frac{B_n}{n!} |x^n| \leqslant |x|^n, \tag{15.605}$$

et donc la série a un rayon de convergence au moins aussi grand que celui de la série géométrique, c'est-à-dire que 1. Donc $R \geqslant 1$. Nous nommons R ce rayon de convergence.

(2) Soit $x \in]-R, R[$. Pour une telle valeur de x à l'intérieur du disque de convergence, la proposition 15.44 nous permet de dériver terme à terme la série [62]

$$f(x) = \sum_{k=0}^{\infty} \frac{B_k}{k!} x^k = 1 + \sum_{k=0}^{\infty} \frac{B_{k+1}}{(k+1)!} x^{k+1}, \tag{15.606}$$

62. C'est ici qu'on utilise la convention $B_0 = 0$ et ça aura une influence sur le choix de la constante K plus bas.

pour obtenir

$$f'(x) = \sum_{k=0}^{\infty} \frac{B_{k+1}}{(k+1)!}(k+1)x^k = \sum_{k=0}^{\infty} \frac{B_{k+1}}{k!}x^k \tag{15.607a}$$

$$= \sum_{k=0}^{\infty} \left(\sum_{l=0}^{k} \binom{k}{l} B_l \right) \frac{x^k}{k!} = \sum_{k=0}^{\infty} \left(\sum_{l=0}^{k} \frac{B_l}{l!(l-k)!} \right) x^k. \tag{15.607b}$$

En cette expression, nous reconnaissons un produit de Cauchy (proposition 15.32) avec $a_l = \frac{B_l}{l!}$ et $b_n = \frac{1}{n!}$. Vu que ce sont deux séries ayant un rayon de convergence plus grand que zéro, le produit a encore un rayon de convergence plus grand que zéro et nous pouvons prendre le produit des séries :

$$f'(x) = \left(\sum_{l=0}^{\infty} \frac{B_l}{l!}x^l \right) \left(\sum_{k=0}^{\infty} \frac{1}{k!}x^k \right) = f(x)e^x. \tag{15.608}$$

Étudions l'équation différentielle $y' = ye^x$. D'abord par un argument en lacet de chaussure [63], une solution est de classe C^∞. Ensuite si une solution est non nulle, elle est de signe constant. En effet si $y(x_0) < 0$ et $y(x_1) = 0$ (on choisit x_1 minimum pour cette propriété parmi les nombres plus grands que x_0) alors il existe [64] un $t \in]x_0, x_1[$ tel que $y'(t) > 0$, ce qui donnerait $y(t) > 0$, ce qui contredirait la minimalité de x_1.

Nous prétendons [65] que cette équation différentielle a un espace de solutions de dimension 1. En effet, si $y' = ye^x$ et $g' = ge^x$ alors en posant $\varphi = y/g$ nous obtenons tout de suite $\varphi' = 0$, ce qui signifie que φ est constante, ou encore que y et g sont multiples l'un de l'autre.

Si nous en trouvons une non nulle par n'importe quel moyen, c'est bon. Une solution étant dérivable est continue, donc l'équation $f' = fe^x$ nous indique que f' est continue. Une solution non nulle va automatiquement accepter un petit voisinage sur lequel la manipulation suivante a un sens :

$$\frac{f'(x)}{f(x)} = e^x, \tag{15.609}$$

donc $\ln\big(|f(x)|\big) = e^x + C$ et $f(x) = Ke^{e^x}$ pour une certaine constante. Il est vite vérifié que cette fonction est une solution de l'équation différentielle $y'(x) = y(x)e^x$ et par unicité, toutes les solutions sont de cette forme. Autrement dit, l'espace des solutions est l'espace vectoriel $\mathrm{Span}\{x \mapsto e^{e^x}\}$. Étant donné que $f(0) = 0$, nous devons choisir $K = \frac{1}{e}$ et donc

$$f(x) = \frac{1}{e}e^{e^x} = e^{e^x - 1}. \tag{15.610}$$

(3) Nous commençons par écrire la fonction f comme une série de puissance. La partie simple du calcul : pour $x \in]-R, R[$, nous avons

$$e^{e^x} = \sum_{k=0}^{\infty} \frac{(e^x)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{l=0}^{\infty} \frac{(kx)^l}{l!} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{k^l}{k!} \frac{x^l}{l!}. \tag{15.611}$$

EqQDjgjDN

Notons que cela n'est pas une série de puissance en x parce qu'il y a la double somme. Nous allons inverser les sommes au moyen du théorème de Fubini sous la forme du corolaire 14.303. Pour cela nous considérons la fonction

$$\begin{aligned} a\colon \mathbb{N} \times \mathbb{N} &\to \mathbb{R} \\ (k, l) &\mapsto \frac{(kx)^l}{k!l!} \end{aligned} \tag{15.612}$$

63. Genre ce qui est fait pour prouver 15.91(4).
64. Théorème de Rolle 12.189.
65. Ou alors on utilise le théorème 32.16 avec $M(x) = e^x$ dans les cas $n = 1$ et $I =]-R, R[$.

et nous mettons la mesure de comptage[66] sur $\mathbb{N}$ et $\mathbb{N}^2$. Nous commençons donc à vérifier l'intégrabilité variable par variable de $|a|$: SubEqsFHsBfhk

$$\int_{\mathbb{N}} \left(\int_{\mathbb{N}} |a(k,l)| dm(l) \right) dm(k) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{(k|x|)^l}{l!} \tag{15.613a}$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} e^{k|x|}. \tag{15.613b}$$

Nous devons montrer que cette dernière somme va bien. Pour cela nous posons $u_k = \frac{e^{k|x|}}{k!}$ et nous remarquons que $\frac{u_{k+1}}{u_k} \to 0$. Donc la double intégrale (15.613) converge, ergo $a \in L^1(\mathbb{N} \times \mathbb{N})$, ce qui nous permet d'utiliser le théorème de Fubini 14.304 pour inverser les ~~sommes~~ ~~intégrales~~ sommes dans l'équation (15.611) :

$$\frac{1}{e} e^{e^x} = \frac{1}{e} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{k!} \frac{1}{l!} (kx)^l = \sum_{l=0} \frac{1}{e} \frac{1}{l!} \left(\sum_{k=0}^{\infty} \frac{k^l}{k!} \right) x^l. \tag{15.614}$$

Cela est un développement en série entière pour la fonction $\frac{1}{e} e^{e^x}$, dont nous savions déjà le développement (15.599) ; par unicité du développement nous pouvons identifier les coefficients :

$$B_l = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^l}{k!}. \tag{15.615}$$

(4) Le développement (15.611) étant en réalité valable pour tout x et tous les calculs subséquents l'étant aussi, le développement

$$e^{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n \tag{15.616}$$

est en fait valable pour tout x, ce qui donne à la série entière un rayon de convergence infini. □

66. Nous passons outre les avertissements et menaces de Arnaud Girand.

Chapitre 16

Représentations et caractères

16.1 Représentations et caractères

Définition 16.1.
Si G est un groupe, l'ensemble des homomorphismes $\mathrm{Hom}(G, \mathbb{C}^*)$ est un groupe pour la multiplication. Un élément de $\mathrm{Hom}(G, \mathbb{C}^*)$ est un ***caractère abélien****. Le nom « abélien » vient du fait que le caractère prenne ses valeurs dans $\mathbb{C}^*$. Nous notons $\hat{G} = \mathrm{Hom}(G, \mathbb{C}^*)$.*

Théorème 16.2.
Soit G un groupe abélien fini. Alors G est isomorphe à $\hat{G}$.

L'isomorphisme n'est pas canonique.

Démonstration. Étant donné la structure des groupes abéliens finis donnée par le théorème 5.23, nous commençons par nous concentrer sur $G = \mathbb{Z}/n\mathbb{Z}$. Nous allons montrer que

$$\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}) \simeq \mathrm{U}_n = \{\xi \in \mathbb{C} \text{ tel que } \xi^n = 1\}. \tag{16.1}$$

Pour cela nous avons l'isomorphisme

$$\begin{aligned} \psi\colon \mathrm{Hom}(\mathbb{Z}, \mathbb{C}^*) &\to \mathbb{C}^* \\ f &\mapsto f(1). \end{aligned} \tag{16.2}$$

Notons que si $f \in \mathrm{Hom}(\mathbb{Z}, \mathbb{C}^*)$, alors $f(k) = f(1)^k$, donc ψ est bien un isomorphisme. Cela nous amène à définir

$$\begin{aligned} \varphi\colon \mathrm{Hom}\Big((\mathbb{Z}/n\mathbb{Z}, +), (\mathbb{C}^*, \cdot)\Big) &\to \mathrm{U}_n \\ f &\mapsto f(1). \end{aligned} \tag{16.3}$$

Remarquons que pour tout $f \in \mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^*)$ on a bien $f(1)^n = 1$. En effet si $[k] \in \mathbb{Z}/n\mathbb{Z}$, alors $f\big([k]\big) = f(1)^k$ et en particulier

$$f(1)^n = f([n]) = f(0) = 1. \tag{16.4}$$

Donc $f(1) \in \mathrm{U}_n$. Le φ est injective parce que si $f(1) = g(1)$ alors $f = g$ du fait que $f(k) = f(1)^k = g(1)^k = g(k)$.

Nous en sommes à avoir prouvé que $\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^*) \simeq \mathrm{U}_n$ (introduit au lemme 19.2). Il faudrait encore montrer que $\mathrm{U}_n \simeq \mathbb{Z}/n\mathbb{Z}$. Pour cela nous nous rappelons du lemme 19.5 nous ayant raconté que le groupe U_n des racines de l'unité était cyclique et d'ordre n. Il est donc bien isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Passons au cas où

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \ldots \times \mathbb{Z}/d_k\mathbb{Z}. \tag{16.5}$$

Dans ce cas nous montrons que

$$\begin{aligned}\alpha\colon \bigtimes_{i=1}^{k} \operatorname{Hom}(\mathbb{Z}/d_i\mathbb{Z}, \mathbb{C}^*) &\to \operatorname{Hom}(G, \mathbb{C}^*)\\ \alpha(\chi_1, \ldots, \chi_k)(g_1, \ldots, g_k) &= \chi_1(g_1)\ldots\chi_k(g_k).\end{aligned} \tag{16.6}$$

Ce α est injectif parce qu'en appliquant l'égalité

$$\alpha(\chi_1, \ldots, \chi_k) = \alpha(\chi'_1, \ldots, \chi'_k) \tag{16.7}$$

à l'élément $g = (0, \ldots, 1, \ldots, 0)$ alors nous trouvons $\chi_i(1) = \chi'_i(1)$ parce que $\chi_j(0) = 1$. Du coup $\chi_i = \chi'_i$.

L'application α est en plus surjective. En effet si $\chi \in \operatorname{Hom}(G, \mathbb{C}^*)$, alors nous définissons

$$\chi_i(g_i) = \chi(0, \ldots, g_i, \ldots, 0), \tag{16.8}$$

et nous avons alors $\alpha(\chi_1, \ldots, \chi_k) = \chi$.

Nous devons encore montrer que α est un homomorphisme. Si $\chi, \chi' \in \times_{i=1}^{k} \operatorname{Hom}(\mathbb{F}_{d_i}, \mathbb{C}^*)$, alors

$$\begin{aligned}\alpha(\chi\chi')(g_1, \ldots, g_k) &= (\chi_1\chi'_1)(g_1)\ldots(\chi_k\chi'_k)(g_k) && \text{(16.9a)}\\ &= \chi_1(g_1)\ldots\chi_k(g_k)\chi'_1(g_1)\ldots\chi'_k(g_k) && \text{(16.9b)}\\ &= \alpha(\chi)(g_1, \ldots, g_k)\alpha(\chi')(g_1, \ldots, g_k) && \text{(16.9c)}\\ &= \big(\alpha(\chi)\alpha(\chi')\big)(g_1, \ldots, g_k). && \text{(16.9d)}\end{aligned}$$

Donc $\alpha(\chi\chi') = \alpha(\chi)\alpha(\chi')$. □

Théorème 16.3.
Soit G un groupe abélien fini. Les groupes G et $\hat{\hat{G}}$ sont isomorphes et un isomorphisme canonique est donné par $\alpha\colon g \mapsto f_g$ donné par

$$f_g(\chi) = \chi(g). \tag{16.10}$$

Démonstration. D'abord f_g est bien un caractère de $\hat{G}$ parce que

$$f_g(\chi\chi') = (\chi\chi')(g) = \chi(g)\chi'(g) = f_g(\chi)f_g(\chi'). \tag{16.11}$$

Le fait que α soit un homomorphisme de groupes est direct :

$$f_{gg'}(\chi) = \chi(gg') = \chi(g)\chi(g') = f_g(\chi)f_{g'}(\chi) = (f_g f_{g'})(\chi). \tag{16.12}$$

D'autre part nous savon que G et $\hat{\hat{G}}$ ont le même cardinal. Il suffit donc de prouver l'injectivité de α pour être sûr de la bijectivité. Pour cela nous devons prouver que si $g \neq e$ alors $f_g \neq f_e$. Nous savons que pour tout caractère $\chi \in \hat{G}$, $f_e(\chi) = \chi(e) = 1$. Donc pour tout $g \in G\backslash\{e\}$, nous devons trouver $\chi \in \hat{G}$ tel que $\chi(g) \neq 1$.

En vertu de ce que nous connaissons sur la structure des groupes abéliens finis (théorème 5.23), nous commençons $G = \mathbb{Z}/n\mathbb{Z}$ et considérons le caractère donné par $\chi([1]) = e^{2i\pi/n}$. Ce χ est un isomorphisme entre G et $\mathrm{U}(n)$; nous n'avons $\chi([k]) = 0$ que si $[k] = [n] = [0]$. Pour rappel dans $\mathbb{Z}/n\mathbb{Z}$, le neutre est $e = 0$ et non $e = 1$.

Passons au cas général :

$$G \simeq \mathbb{Z}/n_1\mathbb{Z} \times \ldots \times \mathbb{Z}/n_k\mathbb{Z} \tag{16.13}$$

Si $g = (g_1, \ldots, g_k)$ est non nul dans G, alors il existe k tel que $g_k \neq 0$ et on prend

$$\chi(g_1, \ldots, g_k) = \chi_k(g_k) \tag{16.14}$$

où χ_k est le caractère $\chi_i([1]) = e^{2\pi i/n_k}$. Ce χ est alors un caractère non trivial de G.

□

16.1.1 Crochet de dualité et transformée de Fourier

Si G est un groupe abélien, nous définissons le crochet de dualité entre G et $\hat{G}$ par

$$\begin{aligned}\langle ., . \rangle \colon G \times \hat{G} &\to \mathbb{C}^* \\ \langle g, \chi \rangle &= \chi(g).\end{aligned} \tag{16.15}$$

Notons que l'image de ce crochet n'est pas $\mathbb{C}^*$ entier, mais seulement le groupe unitaire $\mathrm{U}(n)$ où n est l'exposant [1] de G.

Si f, g sont des applications de G dans $\mathbb{C}$, alors on leur associe le produit scalaire

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{s \in G} f(s)\overline{g(s)}. \tag{16.16}$$

LEMooCSTVooJfhzEU

Lemme 16.4.
Les caractères de G forment une base orthonormée de $\mathbb{C}^G$ pour ce produit scalaire.

Démonstration. Étant donné que les $\chi(s)$ sont des nombres complexe de module 1, nous avons $\chi(s)\overline{\chi(s)} = 1$ et par conséquent $\langle \chi, \chi \rangle = 1$.

Si par contre $\chi \neq \chi'$, alors il existe $s_\in G$ tel que $\chi(s_0) \neq \chi'(s_0)$. Dans ce cas en effectuant un changement de variable $s \to s_0 s$ dans la sommation,

$$\langle \chi, \chi' \rangle = \frac{1}{|G|} \sum_{s \in G} \chi(s)\overline{\chi'(s)} \tag{16.17a}$$

$$= \frac{1}{|G|} \sum_{s \in G} \chi(s_0 s)\overline{\chi'(s_0 s)} \tag{16.17b}$$

$$= \frac{1}{|G|} \chi(s_0)\overline{\chi'(s_0)} \sum_{s \in G} \chi(s)\overline{\chi'(s)}. \tag{16.17c}$$

Donc nous avons trouvé

$$\langle \chi, \chi' \rangle \big(1 - \chi(s_0)\overline{\chi'(s_0)}\big) = 0. \tag{16.18}$$

Mais vu que $\chi(s_0) \neq \chi'(s_0)$, la parenthèse est non nulle (pour rappel $\chi(s_0)$ est un complexe de module 1) et par conséquent $\langle \chi, \chi' \rangle = 0$.

Nous déduisons immédiatement que les caractères forment une famille libre parce que si $\sum_i a_i \chi_i = 0$ (la somme est sur tous les caractères), alors en prenant le produit scalaire avec χ_k,

$$\sum_i a_i \langle \chi_k, \chi_i \rangle = 0, \tag{16.19}$$

et donc $a_k = 0$.

Les caractères forment donc un système libre orthonormé. De plus l'espace engendré à la bonne dimension parce que le cardinal de l'ensemble des caractères est la dimension (complexe) de l'espace des fonctions de G dans $\mathbb{C}$ parce que, en utilisant l'isomorphisme entre G et $\hat{G}$,

$$\operatorname{Card} \hat{G} = \operatorname{Card}(G) = \dim_{\mathbb{C}} \mathbb{C}^G. \tag{16.20}$$

□

Du fait que les caractères forment une base orthonormée, nous pouvons écrire, pour toute application $f \colon G \to \mathbb{C}$,

$$f = \sum_{\chi \in \hat{G}} \langle \chi, f \rangle \chi. \tag{16.21}$$

EqnnsXWC

À une fonction $f \colon G \to \mathbb{C}$ nous associons la **transformée de Fourier**

$$\begin{aligned}\hat{f} \colon \hat{G} &\to \mathbb{C} \\ \chi &\mapsto \langle \chi, f \rangle.\end{aligned} \tag{16.22}$$

1. Définition 1.271.

Nous avons donc aussi une espèce de formule d'inversion

$$f = \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi \tag{16.23}$$

qui n'est qu'une réécriture de 16.21.

16.1.2 Groupes non abéliens

Nous avons vu que le groupe des caractères $\hat{G}$ contenait toute l'information sur un groupe abélien. Malheureusement, pour les groupes non abéliens, ça ne va pas suffire, et nous allons introduire la notion de représentations, dont les caractères seront un cas particulier de dimension un.

Proposition 16.5.
Soit G un groupe (pas spécialement abélien). Nous avons

$$\hat{G} \simeq \text{Hom}\big(G/D(G), \mathbb{C}^*\big). \tag{16.24}$$

Démonstration. Ce qui fait fonctionner la preuve est le fait que si $f\colon G \to \mathbb{C}^*$ est un homomorphisme, alors f s'annule sur $D(G)$. L'isomorphisme est

$$\begin{aligned} \psi\colon \hat{G} &\to \text{Hom}\big(G/D(G), \mathbb{C}^*\big) \\ \psi(f)[g] &= f(g). \end{aligned} \tag{16.25}$$

Cette application est bien définie parce que si f est un homomorphisme,

$$f(gklk^{-1}l^{-1}) = f(g). \tag{16.26}$$

D'autre part ψ est un homomorphisme de groupe parce que

$$\psi(f_1f_2)[g] = (f_1f_2)(g) = f_1(g)f_2(g) = \psi(f_1)[g]\psi(f_2)[g] = \big(\psi(f_1)\psi(f_2)\big)[g]. \tag{16.27}$$

Pour l'injectivité de ψ, soit f_1 et f_2 telles que $\psi(f_1) = \psi(f_2)$. Alors pour tout $g \in G$ nous avons

$$\psi(f_1)[g] = \psi(f_2)[g] \tag{16.28}$$

et donc $f_1(g) = f_2(g)$.

Enfin ψ est surjective. En effet, soit $\bar{f} \in \text{Hom}\big(G/D(G), \mathbb{C}^*\big)$. Alors nous obtenons $\psi(f) = \bar{f}$ en posant

$$f(g) = \bar{f}[g]. \tag{16.29}$$

Il faut juste vérifier que le f ainsi défini est dans $\hat{G}$, c'est-à-dire que $f(g_1g_2) = f(g_1)f(g_2)$. □

Cette proposition nous montre que

$$\hat{G} = \widehat{G/D(G)}, \tag{16.30}$$

alors que $G/D(G)$ est abélien ; il n'est donc pas tellement possible que $\hat{G}$ contienne beaucoup d'informations intéressantes sur G.

16.1.3 Représentations linéaires des groupes finis

Si $\dim V = 1$, alors $\text{GL}(V) = \mathbb{C}^*$ et les représentation sont les caractères abéliens.

ExKUAyUD

Exemple 16.6.
Considérons le triangle équilatéral A, B, C donné par les points

$$\begin{cases} A = 1 & (16.31a) \\ B = (-\frac{1}{2}, \frac{\sqrt{3}}{2}) & (16.31b) \\ C = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}) & (16.31c) \\ & (16.31d) \end{cases}$$

Dans la base (pas orthonormée) $\{A, B\}$ de $\mathbb{R}^2$, ces trois points sont donnés par

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} -1 \\ -1 \end{pmatrix}. \tag{16.32}$$

Le groupe symétrique S_3 agit sur le triangle par permutation des sommets. Vues dans la base $\{A, B\}$, les transpositions correspondent aux matrices

$$\rho(AB) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{16.33a}$$

$$\rho(AC) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \tag{16.33b}$$

$$\rho(BC) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}. \tag{16.33c}$$

Cherchons la matrice correspondante à la permutation (A, B, C). Vu que A est envoyé sur B, la première colonne sera [2] $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, et comme B est envoyé sur C, la deuxième colonne sera $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Nous avons donc

$$\rho(ABC) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}. \tag{16.34}$$

Nous vérifions que $\rho(ABC)C = A$:

$$\rho(ABC)C = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \tag{16.35}$$

Par ailleurs, la permutation (A, B, C) se décompose en $(A, B, C) = (A, C)(A, B)$ et nous pouvons vérifier que

$$\rho(AC)\rho(AB) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = \rho(ABC). \tag{16.36}$$

△

Définition 16.7.
Si (V, ρ) et (V', ρ') sont deux représentations du groupe G, alors nous définissons la ***somme directe*** *par $\big(V \oplus V', \rho \oplus \rho'\big)$ donné par*

$$(\rho \oplus \rho')(g) = \begin{pmatrix} \rho(g) & 0 \\ 0 & \rho'(g) \end{pmatrix} \in \mathrm{GL}(V \oplus V'). \tag{16.37}$$

Nous noterons souvent $2V$ pour la représentations $(V, \rho) \oplus (V, \rho)$ et plus généralement l'écriture

$$V = \bigoplus_i k_i W_i \tag{16.38}$$

signifiera la représentation somme de k_i termes de la représentation W_i. Ici encore un abus est commis entre la représentation (ρ_i, W_i) et l'espace W_i.

2. Les colonnes sont les images des vecteurs de base.

16.1.4 Module

Nous considérons la $\mathbb{C}$-algèbre $G[\mathbb{C}]$ des combinaisons (formelles) d'éléments de G à coefficients dans $\mathbb{C}$, c'est-à-dire l'ensemble

$$\mathbb{C}[G] = \{\sum_{s\in G} a_s s\} \tag{16.39}$$

avec le produit hérité de la bilinéarité :

$$\sum_{s\in G}\sum_{t\in G} a_s b_t st = \sum_s\sum_t a_s b_{s^{-1}t} t, \tag{16.40}$$

et la somme

$$(\sum_s a_s s) + \sum_t b_t t = \sum_{s\in G}(a_s + b_s)s. \tag{16.41}$$

Le tout est une $\mathbb{C}$-algèbre agissant sur V par

$$\left(\sum_s a_s s\right) v = \sum_{s\in G} a_s \rho(s)v \in V \tag{16.42}$$

Les sous-modules indécomposables seront les représentations irréductibles.

Définition 16.8.
*La représentation (V,ρ) du groupe G est **irréductible** si les seuls sous-espaces invariants de V sous $\rho(G)$ sont V et $\{0\}$.*

Exemple 16.9.
La représentation de S_3 sur $\mathbb{R}^2$ donnée par les permutations des sommets d'un triangle équilatéral donnée dans l'exemple 16.6 est irréductible. △

La question qui vient est de savoir si une représentation possédant des sous-espaces invariants peut être écrite comme la somme de représentations irréductibles.

PropHeyoAN

Proposition 16.10.
Soit (V,ρ) une représentation linéaire de dimension finie d'un groupe fini[3]*. Si W_1 est un sous-espace stable*[4]*, alors il existe un sous-espace W_2 également stable et tel que $V = W_1 \oplus W_2$.*

Toute représentation linéaire est décomposable en représentations irréductibles.

Démonstration. Soit $P\colon V \to V$ un projecteur sur W_1, c'est-à-dire que $P^2 = P$ et $P(V) = W_1$. Pour construire un tel projecteur, on peut par exemple prendre un supplémentaire de W_1 dans V puis utiliser la décomposition[5]. Nous considérons l'opérateur

$$P_G = \frac{1}{|G|}\sum_{g\in G} \rho(g)\circ P\circ \rho(g)^{-1}. \tag{16.43}$$

Prouvons que ce P_G est encore un projecteur. D'abord pour tout $g\in G$ nous avons

$$\rho(g)P_G\rho(g)^{-1} = \frac{1}{|G|}\sum_{s\in G}\rho(gs)P\rho(gs)^{-1} = P_G. \tag{16.44}$$

La dernière égalité est un changement de variables dans la somme[6]. Cela signifie que $P_G\rho = \rho P_G$.

3. La démonstration marche aussi pour les groupes compacts, mais il faudrait des intégrales.
4. c'est-à-dire si ρ n'est pas irréductible.
5. Ou encore prendre une base de W_1, l'étendre en une base de V et définir P comme l'annulation des coefficients des vecteurs « complétant » la base.
6. Et c'est ça qui demande un peu de technique pour écrire la preuve dans le cas d'un groupe compact : il faut une mesure de Haar.

Nous avons même $P_G P = P$ parce que si $v \in W_1$, alors

$$P_G(v) = \frac{1}{|G|} \sum_{s\in G} \rho(s) P \underbrace{\rho(s)^{-1} v}_{\in W_1} \tag{16.45a}$$

$$= \frac{1}{|G|} \sum_{s} \rho(s)\rho(s)^{-1} v \tag{16.45b}$$

$$= v. \tag{16.45c}$$

Avec cela nous pouvons conclure que $P_G^2 = P_G$ parce que

$$P_G \circ P_G = \frac{1}{|G|} \sum_{g} P_G \rho(g) P \rho(g)^{-1} \tag{16.46a}$$

$$= \frac{1}{|G|} \sum_{g} \rho(g) P_G P \rho(g)^{-1} \tag{16.46b}$$

$$= \frac{1}{|G|} \sum_{g} \rho(g) P \rho(g)^{-1} \tag{16.46c}$$

$$= P_G. \tag{16.46d}$$

Donc P_G est un projecteur, est stable sous les conjugaisons par $\rho(g)$ et commute avec $\rho(g)$. Nous décomposant Id de façon évidente en

$$\mathrm{Id} = P_G + (\mathrm{Id} - P_G). \tag{16.47}$$

Étant donné que l'opérateur P_G commute avec tous les $\rho(g)$, les noyaux de P_G et $\mathrm{Id} - P_G$ sont des sous-espaces invariants. Vu que P_G est un projecteur, nous avons $q(P_G) = 0$ avec $q(X) = X^2 - X$. Pour appliquer le lemme des noyaux (théorème 8.92), nous remarquons que $q(X) = X(X-1)$ et donc

$$V = \ker P_G \oplus \ker(P_G - \mathbb{1}). \tag{16.48}$$

Si nous posons $W_2 = \ker P_G$, il reste à voir que $\ker(P_G - \mathbb{1}) = W_1$. D'abord $W_1 \subset \ker(P_G - \mathrm{Id})$ parce que si $w \in W_1$, ce dernier étant stable,

$$P_G w = \frac{1}{|G|} \sum_{g\in G} \rho(g) P \underbrace{\rho(g)^{-1} w}_{\in W_1} \tag{16.49a}$$

$$= \frac{1}{|G|} \sum_{g\in G} w \tag{16.49b}$$

$$= w. \tag{16.49c}$$

Pour prouver l'inclusion inverse, nous savons que P_G et P sont des projecteurs tels que $P_G P = P$, ce qui signifie que l'image de P_G est incluse à celle de P, c'est-à-dire à W_1. Mais $\mathrm{Image}(P_G) = \ker(\mathbb{1} - P_G)$, donc

$$\ker(\mathbb{1} - P_G) = \mathrm{Image}(P_G) \subset \mathrm{Image}(P) = W_1. \tag{16.50}$$

La représentation ρ se décompose donc en deux sous-représentations (ρ, W_1) et ρ, W_2. Si l'une des deux n'est pas irréductible, le processus peut recommencer. Vu que la dimension de V est finie, toute représentation se décompose en une somme finie de représentation irréductibles. □

16.1.5 Structure hermitienne

Soit (ρ, V) une représentation de G sur un espace vectoriel complexe V. Nous voulons munir V d'un produit scalaire hermitien (définition 8.176) tel que les opérateurs $\rho(g)$ soient tous des isométries. C'est-à-dire que nous voudrions définir $\langle u, v\rangle_G$ de telle sorte à avoir

$$\langle \rho(g)u, \rho(g)v\rangle_G = \langle u, v\rangle_G \tag{16.51}$$

pour tout $g \in G$. Nous commençons par considérer un produit hermitien $\langle ., . \rangle$ quelconque et puis nous définissons

$$\langle u, v \rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)u, \rho(g)v \rangle. \tag{16.52}$$

Nous devons vérifier que c'est un produit. La seule des conditions dont la vérification n'est pas immédiate est celle de positivité. Pour tout $g \in G$ et tout $v \in V$, nous avons $\langle \rho(g)v, \rho(g)v \rangle$ est positif et nul si et seulement si $\rho(g)v = 0$. Étant donné que $\rho(e)v = v$, parmi les termes de la somme

$$\langle u, u \rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)v, \rho(g)v \rangle, \tag{16.53}$$

au moins un est strictement positif (pourvu que $v \neq 0$) ; les autres sont positifs ou nuls. Par conséquent $\langle v, v \rangle_G = 0$ si et seulement si $v = 0$.

Donc les groupes finis peuvent être vus comme des parties de groupes d'isométrie. De la même façon, en utilisant une mesure de Haar pour faire la moyenne, nous pouvons plonger les groupes compacts dans des groupes unitaires.

16.1.6 Caractères

Définition 16.11.
Soit (V, ρ) une représentation linéaire du groupe G. Le ***caractère*** *de ρ est la fonction*

$$\begin{aligned} \chi_\rho \colon G &\to \mathbb{C} \\ s &\mapsto \operatorname{Tr}\big(\rho(s)\big). \end{aligned} \tag{16.54}$$

Par invariance cyclique de la trace, nous avons

$$\chi_\rho(sts^{-1}) = \chi_\rho(t), \tag{16.55}$$

ce qui fait que le caractère est une fonction constante sur les classes de conjugaison.

D'après sa fiche wikipédia, le marquis de Sade, passionné de théâtre, faisait des représentations qui avaient du caractère.

Un **caractère irréductible** est un caractère d'une représentation irréductible.

Définition 16.12.
Une application $f \colon G \to \mathbb{C}$ est ***centrale*** *si elle est constante sur les classes de conjugaison.*

Les traces sont des applications centrales.

L'ensemble des fonctions centrales sur un groupe fini (ou tout au moins ayant un nombre fini de classes de conjugaison) est un $\mathbb{C}$-espace vectoriel de dimension égale au nombre de classes, et nous pouvons mettre le produit scalaire

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{s \in G} f(s)\overline{g(s)}. \tag{16.56}$$ EqIrEpVI

C'est une forme hermitienne sur l'espace des fonctions centrales.

16.2 Équivalence de représentations et caractères

Cette section prend des éléments des articles lemme de Schur, caractère d'une représentation, fonction centrale et trace de wikipédia.

Définition 16.13.
Nous disons que les deux représentations (V, ρ) et (V', ρ') sont ***équivalentes*** *si il existe une bijection linéaire $f \colon V \to V'$ telle que*

$$f \circ \rho = \rho' \circ f. \tag{16.57}$$

Nous disons alors que f ***entrelace*** *ρ et ρ'.*

ThoyftobH

Théorème 16.14 (Théorème de Schur)**.**
Si (V,ρ) et (V',ρ') sont des représentations irréductibles non équivalentes alors la seule application linéaire $f\colon V\to V'$ entrelaçant ρ et ρ' est la fonction nulle.

En d'autres termes, soit les représentations sont équivalentes (et il y a un isomorphisme), soit il n'y a même pas un homomorphisme.

Démonstration. Soit $f\in\mathcal{L}(V,V')$ telle que $f\circ\rho=\rho'\circ f$. Alors $\ker f$ est un sous-espace stable sous $\rho(G)$, et $\text{Image}(f)$ est un sous-espace de V' stable par $\rho'(G)$. Par irréductibilité, nous avons que $\ker(f)=\{0\}$ ou V. Même chose pour $\text{Image}(f)$. Il y a deux possibilités.

(1) Si $\ker(f)=\{0\}$, alors $\text{Image}(f)\neq\{0\}$ et alors $\text{Image}(f)=V'$. Du coup f est injective et surjective, c'est-à-dire est un isomorphisme.
(2) Si $\ker(f)=V$, alors $f=0$.

□

Corolaire 16.15 (Schur pour les représentations sur $\mathbb{C}$)**.**
Soit (V,ρ) une représentation irréductible, alors l'ensemble

$$\text{End}_G(V,\rho)=\{f\in\text{End}(V)\text{ tel que }\rho\circ f=f\circ\rho\}\tag{16.58}$$

est l'ensemble des homothéties.

Démonstration. Soit $f\in\text{End}_G(V,\rho)$. Vu que l'espace est sur $\mathbb{C}$, l'endomorphisme f a une valeur propre λ. L'opérateur $g=f-\lambda\mathbb{1}$ est aussi un opérateur d'entrelacement de ρ alors que $\ker(g)\neq\{0\}$ par définition de valeur propre. Du coup $\ker(g)=V$, ce qui signifie que f est l'isométrie de rapport λ : $f=\lambda\,\text{Id}$. □

LempUSOlo

Lemme 16.16.
Si (ρ,V) et (ρ',V') sont des représentations équivalentes de caractères χ et χ', alors $\chi=\chi'$.

Démonstration. Si $A\colon V\to V'$ est un isomorphisme d'espace vectoriel entrelaçant ρ et ρ', c'est-à-dire si pour tout g, $\rho'(g)A=A\rho(g)$, alors $\rho'(g)=A\rho(g)A^{-1}$ et

$$\chi'(g)=\text{Tr}\big(\rho'(g)\big)=\text{Tr}\big(A\rho(g)A^{-1}\big)=\text{Tr}\big(\rho(g)\big)\tag{16.59}$$

parce que la trace est un invariant de similitude (lemme 8.210). □

LemJqIZns

Lemme 16.17.
Si χ est le caractère de la représentation complexe (V,ρ) du groupe fini G, alors pour tout $g\in G$ nous avons $\chi(g^{-1})=\overline{\chi(g)}$.

Démonstration. Par le corolaire 2.14 au théorème de Lagrange, nous avons $g^{|G|}=e$ et donc en tant qu'opérateur, $\rho(g)^{|G|}=\mathbb{1}$. Les valeurs propres de $\rho(g)$ sont donc des racines de l'unité. Si nous notons λ_i ces valeurs propres, alors $\chi(g)=\sum_i\lambda_i$, et en considérant la matrice dans sa base de diagonalisation (lemme de Schur complexe, 12.100), nous voyons que

$$\chi(g^{-1})=\text{Tr}\big(\rho(g)^{-1}\big)=\sum_i\frac{1}{\lambda_i}.\tag{16.60}$$

Mais λ_i étant une racine de l'unité nous avons $\frac{1}{\lambda_i}=\bar{\lambda}_i$, ce qui fait que

$$\chi(g^{-1})=\sum_i\bar{\lambda}_i=\overline{\chi(g)}.\tag{16.61}$$

□

PropJzbfWi

Proposition 16.18.
Soient deux représentations irréductibles complexes (V,ρ) et (V',ρ') du même groupe fini G, et χ et χ' leurs caractères respectifs. Nous avons

(1) $\langle\chi,\chi'\rangle=0$ *si ρ et ρ' ne sont pas équivalentes.*

(2) $\langle\chi,\chi'\rangle=1$ *si les représentations sont équivalentes.*

Démonstration. Nous considérons les bases $\{e_1,\dots,e_n\}$ de V et $\{f_1,\dots,f_m\}$ de V'. Puis nous considérons la matrice $F(k,l)=E_{kl}\in\mathbb{M}_{m,n}(\mathbb{C})$ où pour rappel, E_{kl} est la matrice de composantes $(E_{kl})_{ij}=\delta_{ki}\delta_{lj}$. Nous posons

$$F_G(k,l)=\frac{1}{|G|}\sum_{g\in G}\rho(g)\circ F(k,l)\circ\rho'(g)^{-1}. \tag{16.62}$$

En nous permettant de ne pas réécrire les indices k et l de F et F_G, nous montrons que F_G entrelace ρ et ρ' :

$$\begin{aligned}
F_G\circ\rho'(t)&=\frac{1}{|G|}\sum_{s\in G}\rho(s)\circ F\circ\rho'(s^{-1})\circ\rho'(t) && (16.63a)\\
&=\frac{1}{|G|}\sum_{s}\rho(s)F\rho'(s^{-1}t) && (16.63b)\\
&=\frac{1}{|G|}\sum_{k}\rho(tk)F\rho'(k^{-1}) && (16.63c)\\
&=\frac{1}{|G|}\rho(t)\sum_{k}\rho(k)F\rho'(k^{-1}) && (16.63d)\\
&=\rho(t)\circ F_G. && (16.63e)
\end{aligned}$$

Dans ce calcul nous avons effectué le changement de variables $k=(s^{-1}t)^{-1}$ qui donne $s=tk$.

Par ailleurs nous avons

$$\begin{aligned}
\Big(\rho(g)F(k,l)\rho'(g^{-1})\Big)_{ij}&=\sum_{r=1}^{n}\sum_{s=1}^{m}\rho(g)_{ir}F(k,l)_{rs}\rho'(g^{-1})_{sj} && (16.64a)\\
&=\sum_{rs}\rho(g)_{ir}\delta_{kr}\delta_{ls}\rho'(g^{-1})_{sj} && (16.64b)\\
&=\rho(g)_{ik}\rho'(g^{-1})_{lj}, && (16.64c)
\end{aligned}$$

et par conséquent

$$F_G(k,l)_{ij}=\frac{1}{|G|}\sum_{g\in G}\rho(g)_{ik}\rho'(g^{-1})_{lj}. \tag{16.65}$$ Eqgvpzfz

Si χ et χ' sont les caractères de ρ et ρ', alors nous avons le produit (16.56) qui donne

$$\begin{aligned}
\langle\chi,\chi'\rangle&=\frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\chi'(g)} && (16.66a)\\
&=\frac{1}{|G|}\sum_{g}\chi(g)\chi'(g^{-1}) && \text{lemme 16.17} && (16.66b)\\
&=\frac{1}{|G|}\sum_{g}\sum_{i=1}^{n}\sum_{j=1}^{m}\rho(g)_{ii}\rho'(g^{-1})_{jj} && (16.66c)\\
&=\sum_{ij}F_G(i,j)_{ij} && \text{par (16.65).} && (16.66d)
\end{aligned}$$ sEqKYywTM

Si les représentations ρ et ρ' ne sont pas équivalentes, le fait que F_G en soit un opérateur d'entrelacement implique par le théorème de Schur 16.14 que $F_G=0$ et donc $\langle\chi,\chi'\rangle=0$.

Si au contraire les représentation sont équivalentes, alors le lemme 16.16 nous dit que $\chi = \chi'$ et nous reprenons la définition :

$$\langle \chi, \chi \rangle = \frac{1}{|G|} \sum_g \chi(g)\overline{\chi(g)} = \frac{1}{|G|} \sum_{g \in G} 1 = 1 \tag{16.67}$$

parce que les nombres $\chi(g)$ sont des racines de l'unité. □

16.2.1 Représentation régulière

DEFooGGPTooExTxWj

Définition 16.19.
*Nous notons λ la **représentation régulière gauche**, agissant sur le $\mathbb{K}$-espace vectoriel des fonctions $G \to \mathbb{K}$ par*

$$\Big(\lambda(g)f\Big)(h) = f(g^{-1}h). \tag{16.68}$$

D'autre part nous considérons les fonctions $\delta_g \colon G \to \mathbb{K}$ (ici $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$ ou pire) définie par

$$\delta_g(h) = \begin{cases} 1 & \text{si } g = h \\ 0 & \text{sinon.} \end{cases} \tag{16.69}$$

La représentation régulière agit sur les fonctions δ_s de la façon suivante :

$$\lambda(g)\delta_s = \delta_{gs} \tag{16.70}$$

parce que $\big(\lambda(g)\delta_s\big)(h) = \delta_s(g^{-1}h) = \delta_{gs}(h)$.

LEMooGYIKooRRUxHq

Lemme 16.20.
Le caractère de la représentation régulière gauche est donné par

$$\chi_\lambda = |G|\delta_e. \tag{16.71}$$

EqUuoVNa

Démonstration. Appliquer l'équation (16.71) fonctionne parce que $\chi_\lambda(e)$ est la dimension de l'espace des fonctions sur G, c'est-à-dire $|G|$. Si par contre $g \neq e$, alors $\lambda(g)$ est une matrice de permutation (dans la base des δ_h) et a donc tous ses éléments diagonaux nuls. □

Si ρ est une représentation et si f est une fonction sur le groupe, alors nous considérons l'opérateur

$$\rho_f = \sum_{g \in G} f(g)\rho(g). \tag{16.72}$$

PropEAXkAY

Proposition 16.21 ([495]).
Si (ρ, V) est une représentation irréductible et si f est une fonction centrale sur G, alors l'opérateur ρ_f est une homothétie de V de rapport

$$\frac{1}{\dim V} \sum_{g \in G} f(g)\chi(g) \tag{16.73}$$

où χ est le caractère de ρ.

Démonstration. Nous commençons par voir que ρ_f entrelace ρ. En effet,

$$\rho(t)^{-1} \circ \rho_f \circ \rho(t) = \sum_g f(g)\rho(t^{-1}gt) \tag{16.74a}$$

$$= \sum_h f(tht^{-1})\rho(h) \qquad \text{chm.var.} h = t^{-1}gt \tag{16.74b}$$

$$= \sum_h f(h)\rho(h) \tag{16.74c}$$

$$= \rho_f \tag{16.74d}$$

où en écrivant $f(tht^{-1}) = f(h)$, nous avons utilisé le fait que f était centrale. Étant donné que ρ_f entrelace une représentation irréductible, le lemme de Schur (16.14) nous indique que ρ_f est une homothétie. Soit k le facteur d'homothétie. Alors d'une part $\mathrm{Tr}(\rho_f) = nk$. D'autre part,

$$\mathrm{Tr}(\rho_f) = \mathrm{Tr}\big(\sum_g f(g)\rho(g)\big) \tag{16.75a}$$
$$= \sum_g f(g)\,\mathrm{Tr}\big(\rho(g)\big) \tag{16.75b}$$
$$= \sum_g f(g)\chi(g). \tag{16.75c}$$

Du coup effectivement

$$k = \frac{1}{n}\sum_{g\in G} f(g)\chi(g). \tag{16.76}$$

□

16.2.2 Caractères et représentations : suite et fin

Lemme 16.22.
Un groupe fini n'a (à équivalence près) qu'un nombre fini de représentations irréductibles.

Démonstration. Les caractères irréductibles forment un système orthonormé (proposition 16.18) et donc libre parmi les fonctions centrales. Donc il y a au plus autant de caractères irréductibles que la dimension de l'espace des fonctions centrales ; et ce dernier est de dimension finie donnée par le nombre de classes de conjugaison de G. □

Nous savons que les caractères de deux représentations équivalentes sont égaux. Étant donné qu'il n'existe qu'un nombre fini de représentations irréductibles, il existe un nombre fini de caractères irréductibles. Nous pouvons donc fixer les notations suivantes. Les caractères irréductibles seront notés $\{\varphi_i\}_{i=1,\dots,h}$ et nous noterons (σ_i, W_i) une représentation ayant le caractère φ_i.

Théorème 16.23 ([495])**.**
Soit (ρ, V) une représentation de G de caractère χ. Alors sa décomposition en représentations irréductibles est donnée par

$$(V,\rho) = \bigoplus_{i=1}^{h} k_i(W_i,\sigma_i) \tag{16.77}$$

avec $k_i = \langle\chi,\varphi_i\rangle$. En particulier, à permutation près des facteurs, la décomposition d'une représentation en représentations irréductibles est unique.

Démonstration. La décomposition de χ en caractères irréductibles est donnée par $\chi = \sum_i k_i\varphi_i$; en prenant le produit de cette égalité avec φ_j et en tenant compte de l'othonormalité des caractères irréductibles,

$$\langle\chi,\varphi_j\rangle = \sum_i k_i\langle\varphi_i,\varphi_j\rangle = k_j. \tag{16.78}$$

□

Le théorème suivant est ce qui nous permet de dire que l'étude des caractères et l'étude des représentations, c'est la même chose.

ThoWGkfADd

Théorème 16.24.
Soit G un groupe fini[7]*.*

ItemZReOWoHi

(1) Deux représentations sont équivalentes si et seulement si elles ont même caractères.

ItemZReOWoHii

(2) Si χ est le caractère d'une représentation, alors

7. Nous sommes depuis longtemps dans l'étude des représentations des groupes finis.

(2a) $\langle\chi,\chi\rangle\in\mathbb{N}$

(2b) $\langle\chi,\chi\rangle=1$ *si et seulement si la représentation est irréductible.*

Démonstration. Nous démontrons chaque point séparément.

(1) Le fait que deux représentations équivalentes aient même caractère est le lemme 16.16. Nous montrons l'autre sens. Si (ρ,V) et (ρ',V') sont deux représentations irréductibles de décompositions

$$V=\bigoplus_i k_iW_i \tag{16.79a}$$

$$V'=\bigoplus_i k_i'W_i, \tag{16.79b}$$

alors si $\chi=\chi'$, nous avons $k_i=k_i'$ et les représentations sont identiques.

(2) Soit (ρ,V) une représentation ayant χ comme caractère. En posant $k_i=\langle\chi,\varphi_i\rangle$ nous avons la décomposition en représentations irréductibles

$$V=\bigoplus_i k_iW_i, \tag{16.80}$$

et aussi

$$\langle\chi,\chi\rangle=\langle\sum_i k_i\varphi_i,\sum_j k_j\varphi_j\rangle=\sum_i k_i^2\in\mathbb{N}. \tag{16.81}$$

Ce nombre est de plus égal à 1 si et seulement si tous les termes de la somme sont nuls sauf un qui vaudrait 1. Ce cas donne une représentation irréductible.

□

PropYLnxIjk

Proposition 16.25.
Si (λ,R) est la représentation régulière gauche de décomposition en représentations irréductibles

$$R=\bigoplus_i k_iW_i, \tag{16.82}$$

alors

(1) $k_i=\dim W_i$, ITEMooLXIJooDxkGJh

(2) $\sum_i(\dim W_i)^2=|G|$, ItemEXAjTIh

(3) pour tout $g\in G$, $\sum_i(\dim W_i)\varphi_i(g)=0$ [8].

(4) Si $\{(n_i,\varphi_i)\}$ *est la liste des couples dimension,caractère des représentations irréductibles non équivalentes, alors pour tout* $s\in G\backslash\{e\}$ *nous avons* $\sum_{i=1}^p n_i\varphi_i(s)=0$ *où la somme porte sur les représentations irréductibles non équivalentes.*

Démonstration. Nous notons r le caractère de la représentation régulière gauche. Nous avons

$$k_i=\langle r,\varphi_i\rangle=\frac{1}{|G|}\sum_{s\in G}r(s)\overline{\varphi_i(s)}=\overline{\varphi_i(e)}. \tag{16.83}$$

Mais $\varphi_i(e)=\dim(W_i)\in\mathbb{N}$, donc nous avons bien $k_i=\dim(W_i)$. Le caractère de la représentation régulière peut alors s'exprimer de deux façons :

$$|G|\delta_e=\sum_i(\dim W_i)\varphi_i. \tag{16.84}$$

En évaluant cette égalité en e nous trouvons directement

$$|G|=\sum_i(\dim W_i)^2, \tag{16.85}$$

8. Cette propriété est appelée « orthogonalité des colonnes » pour une raison qui apparaîtra au moment de compléter le tableau (16.111).

et en l'évaluant en $s \neq e$, nous trouvons

$$0 = \sum_i (\dim W_i)\varphi_i(s). \tag{16.86}$$

□

fqsdf

Le théorème suivant est valable pour les groupes finis (comme toute cette section).

Thogocemg

Théorème 16.26 ([495])**.**
Les caractères irréductibles $\chi_1, \dots, \chi_h$ forment une base orthonormée des fonctions centrales sur G.

Démonstration. Nous savons déjà qu'ils forment un système orthonormé. Considérons le sous-espace $H = \text{Span}\{\varphi_i\}_{i=1,\dots,h}$ de l'espace des fonctions centrales sur G. En vertu de la proposition 4.134, il nous suffit de prouver que $H^\perp = \{0\}$. Soit donc f, une fonction centrale appartenant à $H^\perp$. Pour tout i, nous avons $\langle f, \varphi_i \rangle = 0$ et donc aussi $\langle \bar{f}, \bar{\varphi}_i \rangle = 0$.

Considérant une représentation irréductible (σ, W) de caractère φ, nous savons par la proposition 16.21 que l'opérateur

$$\sigma_{\bar{f}} = \sum_g \bar{f}(g)\varphi(g) \tag{16.87}$$

est une homothétie de rapport $\langle \bar{f}, \bar{\varphi} \rangle / \dim W = 0$. Étant donné que toutes les représentations sont des sommes directes de représentations irréductibles, en réalité l'opérateur $\rho_{\bar{f}}$ est nul pour toute représentation ρ. En particulier pour la représentation régulière,

$$0 = \lambda_{\bar{f}}(\delta_t) = \sum_{g \in G} \bar{f}(g)\lambda(g)(\delta_t) = \sum_g \bar{f}(g)\delta_{ft}. \tag{16.88}$$

En écrivant cette égalité avec $t = e$ et puis en appliquant à $k \in G$ nous trouvons

$$0 = \sum_g \bar{f}(g)\delta_g(k) = \bar{f}(k). \tag{16.89}$$

Donc $\bar{f} = 0$ et f est nulle. □

CorbdcVNC

Corolaire 16.27.
Le nombre de représentations irréductibles non équivalentes d'un groupe fini est égal à son nombre de classes de conjugaison.

Démonstration. Le nombre de classes de conjugaison est la dimension de l'espace des fonctions centrales qui elle-même est égale au nombre de caractères irréductibles par le théorème 16.26. Enfin deux caractères irréductibles sont égaux si et seulement si les représentations sous-jacentes sont équivalentes. □

CORooWAGXooByrelO

Corolaire 16.28.
Toutes les représentations irréductibles d'un groupe abélien sont de dimension 1.

Démonstration. Le corolaire 16.27 nous dit qu'il y a autant de représentations unitaires qu'il n'y a de représentations irréductibles (non équivalentes). Mais les classes de conjugaisons sont des singletons (lemme 1.169). Nous avons donc exactement $|G|$ représentations irréductibles lorsque G est abélien.

Mais d'autre part la proposition 16.25(2) donne $\sum_i (\dim W_i)^2 = |G|$ lorsque la somme parcours les représentations irréductibles. Il y a $|G|$ termes à la somme, donc tous les termes doivent être 1. □

16.3 Représentation produit tensoriel

Définition 16.29.
Soient ρ et ϕ, deux représentations d'un groupe G sur des espaces vectoriels V et W. La représentation ***produit tensoriel*** *est la représentation*

$$\begin{aligned}\rho\otimes\phi\colon G&\to \mathrm{GL}(V\otimes W)\\ (\rho\otimes\phi)(g)(v\otimes w)&=\rho(g)v\otimes\phi(g)w.\end{aligned}\tag{16.90}$$

L'espace $V\otimes W$ est le produit tensoriel de V et W, défini en 10.170.

Pour trouver son caractère, nous considérons une base $\{e_i\}$ de V et une base $\{e_\alpha\}$ de W, et la base $\{e_i\otimes e_\alpha\}$ de $V\otimes W$. Donc

$$(\rho\otimes\phi)(g)(e_i\otimes e_\alpha)=\rho(g)e_i\otimes\phi(g)e_\alpha.\tag{16.91}$$

Nous devons savoir quelle est la composante « $e_i\otimes e_\alpha$ » de cette dernière expression, et c'est évidemment

$$\rho(g)_{ii}\rho_{\alpha\alpha},\tag{16.92}$$

ce qui nous amène à dire que

$$\mathrm{Tr}(\rho\otimes\phi)(g)=\sum_i\sum_\alpha\rho(g)_{ii}\phi(g)_{\alpha\alpha}=\mathrm{Tr}\big(\rho(g)\big)\,\mathrm{Tr}\big(\phi(g)\big),\tag{16.93}$$

c'est-à-dire au final que

$$\chi_{\rho\otimes\phi}=\chi_\rho\chi_\phi.\tag{16.94}$$
EqOTmvfif

16.4 Exemple sur le groupe symétrique

Soit $G=S_3$, un des premiers groupes finis non abéliens. On en a une représentation de dimension deux en tant que permutation des sommets d'un triangle équilatéral, donnée dans l'exemple 16.6 ; nous notons ρ cette représentation.

Nous y avons aussi la représentation de signature donnée par

$$\begin{aligned}\epsilon\colon S_3&\to\mathrm{GL}(\mathbb{C})\\ \sigma&\mapsto\epsilon(\sigma)\,\mathrm{Id}\,.\end{aligned}\tag{16.95}$$

Et enfin il y a la représentation triviale. Ce sont les trois représentations irréductibles ; pour rappel il y a autant de représentations irréductibles que de classes de conjugaison (corolaire 16.27).

Classe de conjugaison	taille	χ_1	χ_ϵ	χ_ρ
Id	1	1	1	2
(A,B)	3	1	-1	0
(A,B,C)	2	1	1	-1

Nous calculons par exemple le produit scalaire

$$\langle\chi_1,\chi_\epsilon\rangle=\frac{1}{6}\big(1\cdot\chi_1(\mathrm{Id})\overline{\chi_\epsilon(\mathrm{Id})}+3\cdot\chi_1(A,B)\overline{\chi_\epsilon(A,B)}+2\cdot\chi_1(A,B,C)\overline{\chi_\epsilon(A,B,C)}\big)\tag{16.96a}$$

$$=0.\tag{16.96b}$$

D'autre part nous avons aussi

$$\langle\chi_\rho,\chi_\rho\rangle=\frac{1}{6}(1\cdot2\cdot2+3\cdot0+2\cdot1)=1.\tag{16.97}$$

16.5 Table des caractères du groupe symétrique S_4

SecUMIgTmO

Pour la table des caractères de S_4, voir [104]. Et si vous voulez la table des caractères du groupe diédral, vu que ce sont de isométries de $\mathbb{R}^n$, il faudra voir plus bas en la section 18.17.

16.5.1 Calculs à partir de rien ou presque

Nous savons que les classes de conjugaison dans S_4 sont caractérisées par la structure des décompositions en cycles (proposition 1.291). Elles sont données dans l'exemple 1.294.

Nous avons donc 5 classes de conjugaison, et il nous faut donc 5 représentations irréductibles non équivalentes (corolaire 16.27) dont nous allons chercher les caractères.

La première est la représentation triviale de dimension 1 ; nous notons χ_1 son caractère et nous avons la ligne

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1

(16.98)

Ensuite nous avons la signature qui est un morphisme non trivial $\epsilon\colon S_n \to \{-1,1\}$. Nous avons alors la ligne

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_ϵ	1	1	-1	1	-1	1

EqGNRavtl (16.99)

Une troisième représentation pas trop compliquée à trouver est celle

$$\begin{aligned}\rho_p\colon S_4 &\to \mathrm{GL}(4,\mathbb{C})\\ \rho_p(\sigma)e_i &= e_{\sigma(i)}.\end{aligned} \tag{16.100}$$

Cela n'est pas une représentation irréductible parce que $\mathbb{C}^4$ se décompose en deux sous-espaces stables :

$$D = \mathrm{Span}(1,1,1,1) \tag{16.101a}$$

$$H = \{x \in \mathbb{C}^4 \text{ tel que } x_1+x_2+x_3+x_4=0\}. \tag{16.101b}$$

La représentation induite sur D est la représentation triviale. Puis sur H, elle induit une autre représentation que nous allons noter ρ_s.

Nous allons à présent déduire le caractère de la représentation ρ_s et prouver qu'elle est irréductible. Il est cependant possible de sauter cette étape en échange d'un certain travail sur les isométries du tétraèdre. Voir la proposition 18.195 et ensuite

Nous avons la décomposition $\rho_p = \rho_1 \oplus \rho_s$ et donc

$$\chi_p = \chi_1 + \chi_s. \tag{16.102}$$

Nous savons déjà χ_1. Le caractère χ_p n'est pas très compliqué parce que $\chi_p(\sigma)$ est une matrice de permutations des vecteurs de base. Donc la matrice $\rho_p(\sigma)$ a un 1 sur la diagonale pour les i tels que $\sigma(i)=i$. Nous avons donc

$$\chi_p(\mathrm{Id}) = 4 \qquad \chi_p(12) = 2 \tag{16.103a}$$

$$\chi_p\big((12)(34)\big) = 0 \qquad \chi_p(123) = 1 \tag{16.103b}$$

$$\chi_p(1234) = 0. \tag{16.103c}$$

Le caractère χ_s peut être calculé par simple soustraction :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_s	3	3	1	0	-1	-1

EqILZsKfo (16.104)

Avant d'ajouter cette ligne au tableau des représentations irréductibles nous devons savoir si ρ_s en est une. Pour cela, tant que nous avons son caractère nous pouvons utiliser le critère du théorème 16.24 :

$$\langle \chi_s, \chi_s \rangle = \frac{1}{|S_4|} \sum_{\sigma \in S_4} \chi_s(\sigma)^2. \tag{16.105}$$

Nous avons tout de suite $|S_4| = 4 \cdot 3 \cdot 2 = 24$ et puis

$$24\langle\chi_s, \chi_s\rangle = 3^2 + 6 \cdot 1^2 + 8 \cdot 0^2 + 6 \cdot (-1)^2 + 3 \cdot (-1)^2 = 24, \tag{16.106}$$

donc oui, le caractère est irréductible parce que $\langle\chi_s, \chi_s\rangle = 1$. Et nous pouvons donc ajouter la ligne (16.104) à notre tableau. Par ailleurs, nous notons qu'elle est de dimension 3.

Pour le reste nous savons qu'il y a autant de représentations irréductibles que de classes de conjugaison, de telle sorte qu'il ne manque que deux représentations irréductibles. De plus la proposition 16.25 nous dit que si n_i est la dimension de la i^{e} représentation irréductible, alors

$$|S_4| = \sum_i n_i^2. \tag{16.107}$$

Dans notre situation, si nous nommons n_1 et n_2 les dimensions des deux représentations qui nous manquent, nous avons $24 = n_1^2 + n_2^2 + (1^2 + 1^2 + 3^2)$, c'est-à-dire $n_1^2 + n_2^2 = 13$. Il n'y a pas des tonnes de sommes de deux carrés qui font 13. Il y a $n_1 = 2$ et $n_2 = 3$, et c'est tout.

Nous recherchons donc encore une représentation de dimension 2 et une de dimension 3. Pour cela nous allons un peu regarder les produits tensoriels qui s'offrent à nous. Pour faire une dimension 3, il faut faire le produit d'une de dimension 1 par une de dimension 3. Là encore le choix est très limité et nous demande d'essayer

$$\rho_W = \rho_s \otimes \rho_\epsilon \tag{16.108}$$

qui agit sur l'espace $V_2 \otimes V_\epsilon$ par

$$\rho_W(g)(v \otimes x) = \rho_s(g)v \otimes \rho_\epsilon(g)x. \tag{16.109}$$

Pour savoir son caractère nous utilisons la petite formule toute simple (16.94) : nous multiplions case par case les tableaux (16.104) et (16.99) :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_W	3	3	-1	0	1	-1

(16.110)

Avant de réellement ajouter cette ligne au tableau, nous devons nous assurer qu'elle est bien irréductible. Nous utilisons le même critère : $\langle\chi_W, \chi_W\rangle = 1$, donc c'est bon.

Pour trouver le dernier caractère, que nous nommerons χ_u, il ne faut pas beaucoup d'imagination. Il suffit d'utiliser les relations d'orthogonalité du théorème 16.26, en sachant que la dimension est 2 et qu'alors $\chi_u(\text{Id}) = 2$, c'est pas trop compliqué :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1
χ_ϵ	1	1	-1	1	-1	1
χ_s	3	3	1	0	-1	-1
χ_W	3	3	-1	0	1	-1
χ_u	2	2	b	c	d	e

EqOKtZYEQ (16.111)

Les relations d'orthogonalité des colonnes de la propriété 16.25 nous permettent de calculer les coefficients manquants. En pratique, il suffit de prendre le produit scalaire de chaque ligne avec la première et d'égaler avec zéro. Nous trouvons $b = 0$, $c = 1$, $d = 0$, et $e = 2$. Le tableau final est :

	dimension	Id	(12)	(123)	(1234)	(12)(34)
χ_1	1	1	1	1	1	1
χ_ϵ	1	1	-1	1	-1	1
χ_s	3	3	1	0	-1	-1
χ_W	3	3	-1	0	1	-1
χ_u	2	2	0	-1	0	2

(16.112)

Notons que nous sommes parvenus à remplir la dernière ligne sans rien savoir de la représentation qui va avec. Nous allons cependant donner une interprétation géométrique et fixer cette représentation comme agissant sur le triangle équilatéral en 16.33.

16.5.2 À propos de la représentation ρ_s

Une des représentations trouvées (la représentation ρ_s) peut être vue comme le groupe $\mathrm{Iso}(T)$ des isométries affine du tétraèdre [9] grâce à la proposition 18.195 qui donne un isomorphisme de groupe $S_4 \simeq \mathrm{Iso}(T)$ lorsque T est un tétraèdre régulier de $\mathbb{R}^3$.

Nous verrons donc ça plus en détail dans la section 18.10.19.

16.5.3 À propos de la représentation ρ_u

Nous nous penchons à présent sur la représentations ρ_u dont nous ne savons rien à part qu'elle est de dimension 2 et son caractère.

Lemme 16.30.
Nous avons $\rho_u(s) = \mathrm{Id}$ pour tout $s \in V_4$.

Démonstration. Tous les éléments de V_4 sont conjugués (à part l'identité, mais pour elle le résultat est clair), donc il suffit de prouver le résultat pour un élément quelconque.

L'endomorphisme $\rho_u\big((12)(34)\big)$ est un endomorphisme d'ordre 2 sur $\mathbb{R}^2$ dont la trace est 2. Imposons donc

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{16.113}$$

sous la contrainte $a + d = 2$. La résolution est assez rapide et donne $b = c = 0$, $a = d = 1$.

Vous voulez une démonstration plus technologique ? Oui ? Alors commencez par remarquer que l'opérateur $A = \rho_u\big((12)(34)\big)$ vérifie $A^2 = 1$, donc le polynôme $X^2 - 1$ est un polynôme annulateur de A. Il est peut-être minimal ou peut être pas, mais en tout cas le polynôme minimal divise celui-là et donc est soit $X - 1$ soit $X + 1$ soit $X^2 - 1$. Dans les trois cas il est scindé à racines simples, et l'endomorphisme A est diagonalisable par le théorème 8.219(3).

Mais comme $A^2 = 1$, les valeurs propres (ce qui est sur la diagonale) de A ne peuvent être que ± 1. La trace étant 2, les éléments diagonaux ne peuvent être que 1. Et $A = \mathrm{Id}$. □

Le groupe V_4 définit en 5.56 est normal dans S_4, donc le quotient S_4/V_4 est un groupe par le lemme 2.10.

Lemme 16.31.
L'application

$$\begin{aligned} \tilde{\rho}_u \colon S_4/V_4 &\to \mathrm{GL}(2,\mathbb{R}) \\ [g] &\mapsto \rho_u(g) \end{aligned} \tag{16.114}$$

est bien définie et donne une représentation irréductible de S_4/V_4.

Démonstration. Montrons que c'est bien défini. Si $s \in V_4$ nous devons prouver que $\rho_u(gs) = \rho_u(g)$. Vu que ρ_u est un homomorphisme (c'est une représentation), et que $\rho_u(s) = \mathrm{Id}$ nous avons directement

$$\rho_u(gs) = \rho_u(g)\rho_u(s) = \rho_u(g). \tag{16.115}$$

Nous devons prouver que la représentation $\tilde{\rho}_u$ est irréductible. Si un sous-espace non trivial $\mathrm{Span}(x)$ était stabilisé par $\tilde{\rho}_u$, il serait également stabilisé par ρ_u. Mais comme ρ_u est irréductible, elle ne stabilise personne. □

Lemme 16.32.
Le groupe S_4/V_4 est un groupe non-abélien, isomorphe à S_3.

Démonstration. Le groupe S_4/V_4 a une représentation irréductible de dimension 2, et n'est donc pas abélien par le corolaire 16.28.

Il contient $|S_4|/|V_4| = 24/4 = 6$ éléments (théorème de Lagrange 2.13). Or $6 = 3 \times 2$, donc le groupe S_4/V_4 est dans le cas non-abélien du théorème 5.25(2). Cette partie parle d'unicité du

9. Définition 12.147.

groupe non-abélien d'ordre 6. Or S_3 est un groupe non-abélien d'ordre 6, donc S_4/V_4 est isomorphe à S_3. □

Attention : il n'est pas correct de dire que S_4/V_4 est un sous-groupe de S_4 juste parce que c'est un quotient de S_4 ; ce n'est en général pas vrai (exemple 2.11).

NORMooQQCYooILyOxc

16.33.
Nous sommes maintenant aptes à identifier la représentation ρ_u. D'abord nous nous rappelons de la représentation $\rho_s : S_4 \to \mathrm{Iso}(T)$ de S_4 sur le tétraèdre. Ensuite si A est un sommet dudit tétraèdre et que $S_3 \subset S_4$ est la partie qui fixe A alors nous avons une représentation

$$\rho_s : S_3 \to \mathrm{Iso}(T) \tag{16.116}$$

qui agit en réalité sur le triangle équilatéral T' opposé au sommet A.

Nous avons finalement la chaine d'homomorphismes de groupes

$$S_4 \xrightarrow{\texttt{proj}} S_4/V_4 \xrightarrow{\simeq} S_3 \xrightarrow{\rho_s} \mathrm{Iso}(T') \tag{16.117}$$

Cela est donc une représentation $S_4 \to \mathrm{Iso}(T')$. Elle est de dimension 2 et est irréductible (elle contient les rotations d'angle $2\pi/3$ qui ne fixent aucune direction). Elle est donc la représentation ρ_u qui est la seule irréductible de dimension 2.

Nous avons donc montré que la représentation ρ_u dont nous ne savions rien est la représentation de S_3 sur un triangle équilatéral obtenue à partir de celle de S_4 sur le tétraèdre, en fixant un point.

Chapitre 17

Encore de l'analyse (et c'est pas fini)

17.1 Densité des polynômes

CorRSczQD

Corolaire 17.1.
Si $X \subset \mathbb{R}$ est compact et de mesure finie[1]*, alors l'ensemble des polynômes est dense dans $\big(C(X,\mathbb{R}), \|.\|_2\big)$.*

Démonstration. Si f est une fonction dans $C(X,\mathbb{R})$ et si $\epsilon \geqslant 0$ est donné alors nous pouvons considérer un polynôme P tel que $\|f-P\|_\infty \leqslant \epsilon$. Dans ce cas nous avons

$$\|f-P\|_2^2 = \int_X |f(x)-P(x)|^2 dx \leqslant \int_X \epsilon^2 dx = \epsilon^2 \mu(X) \tag{17.1}$$

où $\mu(X)$ est la mesure de X (finie par hypothèse). □

17.2 Primitive et intégrale

Nous avons déjà parlé de primitive de fonction continue en la proposition 12.424.

PropHFWNpRb

Proposition 17.2.
Soit I un intervalle borné ouvert de $\mathbb{R}$. Une fonction $h \in C_c^\infty(I)$ admet une primitive dans $C_c^\infty(I)$ si et seulement si $\int_I h = 0$.

Démonstration. Si une primitive H de h est à support compact, alors

$$\int_I h = H(b) - H(a) = 0 - 0 = 0. \tag{17.2}$$

Pas de problèmes dans ce sens.

Supposons maintenant que $\int_I h = 0$. Le fait que h admette une primitive dans $C^\infty(I)$ est évident : toute fonction continue admet une primitive[2]. Soit H une telle primitive et $\tilde{H} = H - H(b)$. Alors $\tilde{H}(b) = 0$ et

$$\tilde{H}(a) = H(a) - H(b) = -\int_I h = 0. \tag{17.3}$$

Nous rappelons que le support d'une fonction est *la fermeture* de l'ensemble des points de non-annulation.

Supposons que le support de h soit inclus dans $[m,M] \subset]a,b[$. En prenant des nombres m' et M' tels que $a < m' < m$ et $M < M' < b$ (nous insistons sur le caractère strict de ces inégalités), la fonction h est nulle sur $[a,m']$ et sur $[M',b]$; la fonction $\tilde{H}$ doit donc y être constante. Mais nous avons déjà vu que $\tilde{H}(a) = \tilde{H}(b) = 0$. Donc l'ensemble des points sur lesquels $\tilde{H}$ n'est pas nul est inclus dans $]m',M'[$ et donc est strictement (des deux côtés) inclus dans I. □

1. Dans $\mathbb{R}$ cette hypothèse est évidemment superflue par rapport à l'hypothèse de compacité ; mais ça suggère des généralisations ...
2. Théorème 12.424.

17.2.1 Théorème taubérien de Hardy-Littlewood

Un théorème **taubérien** est un théorème qui compare les modes de convergence d'une série.

Lemme 17.3.
Si f et g sont des fonctions continues, alors $s(x) = \max\{f(x), g(x)\}$ est également une fonction continue.

Démonstration. Soit x_0 et prouvons que s est continue en x_0. Si $f(x_0) \neq g(x_0)$ (supposons $f(x_0) > g(x_0)$ pour fixer les idées), alors nous avons un voisinage de x_0 sur lequel $f > g$ et alors $s = f$ sur ce voisinage et la continuité provient de celle de f.

Si au contraire $f(x_0) = g(x_0) = s(x_0)$ alors si (a_n) est une suite tendant vers x_0, nous prenons N tel que $\big|f(a_n) - f(x_0)\big| \leqslant \epsilon$ pour tout $n > N$ et M tel que $\big|g(a_n) - g(x_0)\big| \leqslant \epsilon$ pour tout $n > M$. Alors pour tout $n > \max\{N, M\}$ nous avons

$$\big|s(a_n) - s(x_0)\big| \leqslant \epsilon, \tag{17.4}$$

d'où la continuité de s en x_0. □

La proposition suivante dit que si une fonction connaît un saut, alors on peut la lisser par une fonction continue.

PropTIeYVw

Proposition 17.4.
Soit f continue sur $[a, x_0[$ et sur $[x_0, b]$ avec $f(x_0^-) < f(x_0)$. En particulier nous supposons que $f(x_0^-)$ existe et est finie. Alors pour tout $\epsilon > 0$, il existe une fonction continue s telle que sur $[a, b]$ on ait $s \geqslant f$ et

$$\int_a^b (s - f)(x)\, dx \leqslant \epsilon. \tag{17.5}$$

Démonstration. Nous notons A la taille du saut :

$$A = f(x_0) - f(x_0^-). \tag{17.6}$$

Quitte à changer a et b, nous pouvons supposer que

$$f(x) < f(x_0) + \frac{A}{3} \tag{17.7}$$

pour $x \in [a, x_0[$ et

$$f > f(x_0) + \frac{2A}{3} \tag{17.8}$$

pour $x \in [x_0, b]$. C'est le théorème des valeurs intermédiaires qui nous permet de faire ce choix.

Soit $m(x)$ la droite qui joint le point $\big(x_0 - \epsilon, f(x_0 - \epsilon)\big)$ au point $\big(x_0, f(x_0^+)\big)$. Nous posons

$$s(x) = \begin{cases} f(x) & \text{si } x < x_0 - \epsilon \\ \max\{m(x), f(x)\} & \text{si } x_0 - \epsilon \leqslant x \leqslant x_0 \\ f(x) & \text{si } x > x_0. \end{cases} \tag{17.9}$$

En vertu des différents choix effectués, c'est une fonction continue. En effet

$$s(x_0 - \epsilon) = \max\{f(x_0 - \epsilon), f(x_0, \epsilon)\} = f(x_0 - \epsilon) \tag{17.10}$$

et

$$s(x_0) = \max\{m(x_0), f(x_0^+)\} = f(x_0^+) \tag{17.11}$$

parce que $m(x_0) = f(x_0^+)$. En ce qui concerne l'intégrale, si nous posons

$$M = \sup_{x,y \in [a,b]} |f(x) - f(y)|, \tag{17.12}$$

nous avons

$$\int_a^b s - f = \int_{x_0 - \epsilon}^{x_0} s - f \leqslant \epsilon M. \tag{17.13}$$

□

LemauxrKN

Lemme 17.5.
Pour tout polynôme P, nous avons la formule

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty}x^nP(x^n)=\int_0^1 P(x)dx. \tag{17.14}$$

Démonstration. D'abord pour $P=1$, la formule se réduit à la série géométrique connue (proposition 10.135). Ensuite nous prouvons la formule pour le polynôme $P=X^k$ et la linéarité fera le reste pour les autres polynômes. Nous avons

$$(1-x)\sum_n x^nx^{kn}=(1-x)\sum_n(x^{1+k})^n=\frac{1-x}{1-x^{1+k}}=\frac{1}{1+x+\cdots+x^k}. \tag{17.15}$$

Donc

$$\lim_{x\to 1^-}(1-x)\sum_n x^nP(x^n)=\frac{1}{1+k}. \tag{17.16}$$

Par ailleurs, c'est vite vu que

$$\int_0^1 x^k dx=\frac{1}{k+1}. \tag{17.17}$$

□

ThoPdDxgP

Théorème 17.6 (Hardy-Littlewood[496])**.**
Soit (a_n) une suite réelle telle que

(1) $\frac{a_n}{n}$ tend vers une constante,

(2) $F(x)=\sum_{n=0}^{\infty}a_nx^n$ a un rayon de convergence $\geqslant 1$,

(3) $\lim_{x\to 1^-}F(x)=l$.

Alors $\sum_{n=0}^{\infty}a_n=l$.

Démonstration. Quitte à prendre la suite $b_0=a_0-l$ et $b_n=a_n$, on peut supposer $l=0$.

Soit Γ l'ensemble des fonctions

$$\gamma\colon[0,1]\to\mathbb{R} \tag{17.18}$$

telles que

(1) $\sum_{n=0}^{\infty}a_n\gamma(x^n)$ converge pour $0\leqslant x<1$,

(2) $\lim_{x\to 1^-}\sum_{n\geqslant 0}a_n\gamma(x^n)=0$.

Ce Γ est un espace vectoriel.

(1) **Les polynômes sont dans Γ** Soit $\gamma(t)=t^s$. Pour $0\leqslant x<1$ nous avons

$$\sum_{n=0}^{\infty}a_n\gamma(x^n)=\sum_{n=0}^{\infty}a_nx^{ns}<\sum_{n=0}^{\infty}a_nx^n. \tag{17.19}$$

Donc la condition de convergence est vérifiée. En ce qui concerne la limite,

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty}a_nx^{ns}=\lim_{x\to 1^-}F(x^s)=0 \tag{17.20}$$

parce que par hypothèse, $\lim_{x\to 1^-}F(x)=0$.

(2) **Définition de la fonction qui va donner la réponse** Nous considérons la fonction

$$g(t) = \begin{cases} 0 & \text{si } 0 \leqslant t < 1/2 \\ 1 & \text{si } 1/2 \leqslant t \leqslant 1, \end{cases} \tag{17.21}$$

c'est-à-dire $g = \mathbb{1}_{[\frac{1}{2},1]}$. Nous montrons que si $g \in \Gamma$, alors le théorème est terminé. Si $0 \leqslant x \leqslant 1$, on a $0 \leqslant x^n < 1/2$ dès que

$$n > -\frac{\ln(2)}{\ln(x)} \tag{17.22}$$

avec une note comme quoi $\ln(x) < 0$, donc la fraction est positive. Nous désignons par N_x la partie entière de ce n adapté à x. L'idée est que la fonction $g(x^n)$ est la fonction indicatrice de $0 \leqslant n \leqslant N_x$, et donc

$$\sum_{n\geqslant 0} a_n g(x^n) = \sum_{n=0}^{N_x} a_n. \tag{17.23}$$

Mais si $x \to 1^-$, alors $N_x \to \infty$, donc

$$\lim_{N\to\infty} \sum_{n=0}^{N} a_n = \lim_{x\to 1^-} \sum_{n=0}^{N_x} a_n = \lim_{x\to 1^-} \sum_{n\in\mathbb{N}} a_n g(x^n), \tag{17.24}$$

et cela fait zéro si $g \in \Gamma$.

(3) **Approximation de g par des polynômes** Nous considérons la fonction

$$h(t) = \frac{g(t) - t}{t(1-t)} = \begin{cases} \frac{1}{t-1} & \text{si } t \in [0, 1/2[\\ \frac{1}{t} & \text{si } t \in [1/2, 1]. \end{cases} \tag{17.25}$$

La seconde égalité est au sens du prolongement par continuité. La fonction h est une fonction non continue qui fait un saut de -2 à 2 en $x = 1/2$. En vertu de la proposition 17.4 (un peu adaptée), nous pouvons considérer deux fonctions continues s_1 et s_2 telles que

$$s_1 \leqslant h \leqslant s_2 \tag{17.26}$$

et

$$\int_0^1 s_2 - s_1 \leqslant \epsilon. \tag{17.27}$$

Notons que l'inégalité $s_1 \leqslant s_2$ doit être stricte sur au moins un petit intervalle autour de $x = 1/2$. Soient P_1 et P_2, deux polynômes tels que $\|P_1 - s_1\|_\infty \leqslant \epsilon$ et $\|P_2 - s_2\|_\infty \leqslant \epsilon$ (ici la norme supremum est prise sur $[0, 1]$). C'est le théorème de Stone-Weierstrass (12.420) qui nous permet de le faire.
Nous posons aussi [3]

$$Q_1 = P_1 + \epsilon \tag{17.28a}$$
$$Q_2 = P_2 - \epsilon. \tag{17.28b}$$

Nous avons

$$\int_0^1 Q_1 - Q_2 \leqslant \int_0^1 Q_1 - P_1 + P_1 - P_2 + P_2 - Q_2. \tag{17.29}$$

Pour majorer cela, d'abord $Q_1 - P_1 = P_2 - Q_2 = \epsilon$, ensuite,

$$P_1 - P_2 = P_1 - s_1 + s_1 - s_2 + s_2 - P_2 \tag{17.30}$$

dans lequel nous avons $P_1 - s_1 \leqslant \epsilon$, $s_2 - P_2 \leqslant \epsilon$ et $\int_0^1 s_1 - s_2 \leqslant \epsilon$. Au final, nous posons $q = Q_2 - Q_1$ et nous avons

$$\int_0^1 q \leqslant 5\epsilon. \tag{17.31}$$

3. À ce niveau, je crois qu'il y a une faute de frappe dans [496].

Enfin nous posons aussi

$$R_i(x) = x + x(1-x)Q_i. \tag{17.32}$$

Ces polynômes vérifient $R_i(0) = 0$, $R_i(1) = 1$ et

$$R_1 \leqslant g \leqslant R_2 \tag{17.33}$$

parce que

$$Q_1 \leqslant P_1 \leqslant h \leqslant P_2 \leqslant Q_2 \tag{17.34}$$

et

$$t + t(1-t)Q_1 \leqslant \underbrace{t + t(1-t)h(t)}_{g(t)} \leqslant t + t(1-t)Q_2. \tag{17.35}$$

(4) **Preuve que g est dans Γ** D'abord si $0 \leqslant x < 1$, $x^N < \frac{1}{2}$ pour un certain N, et alors $g(x^N) = 0$. Du coup la série

$$\sum_{n=0}^{\infty} a_n g(x^n) = \sum_{n=0}^{N} a_n \tag{17.36}$$

est une somme finie qui converge donc.

D'autre part nous prenons M tel que $|a_n| < \frac{M}{n}$ pour tout n. Nous majorons $\sum_{n\in\mathbb{N}} a_n g(x^n)$ en utilisant R_1. Mais vu que R_1 est un polynôme, nous pouvons dire que $|\sum_{n=0}^{\infty} a_n R_1(x^n)| \leqslant \epsilon$ en prenant $x \in [\lambda, 1[$ et λ assez grand. Nous avons :

$$\left|\sum_{n=0}^{\infty} a_n g(x^n)\right| \leqslant \left|\sum_{n=0}^{\infty} a_n g(x^n) - \sum_{n=0}^{\infty} a_n R_1(x^n)\right| + \underbrace{\left|\sum_{n=0}^{\infty} a_n R_1(x^n)\right|}_{\leqslant\epsilon} \tag{17.37a}$$

$$\leqslant \epsilon + \sum_{n=0}^{\infty} |a_n|(g - R_1)(x^n) \tag{17.37b}$$

$$\leqslant \epsilon + \sum_{n=0}^{\infty} |a_n|(R_2 - R_1)(x^n) \tag{17.37c}$$

$$\leqslant \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n}(Q_2 - Q_1)(x^n) \tag{17.37d}$$

$$= \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n} q(x^n) \tag{17.37e}$$

$$\leqslant \epsilon + M(1-x)\sum_{n} x^n q(x^n). \tag{17.37f}$$

Justifications :

— La ligne (17.37d) vient du fait que $R_2 - R_1 = x(1-x)(Q_2 - Q_1)$.

— La ligne (17.37f) provient d'une majoration sauvage de $1/n$ par 1 et de $1-x^n$ par $1-x$.

Par le lemme 17.5, nous avons alors

$$\lim_{x\to 1^-} |\sum_n a_n g(x^n)| \leqslant \epsilon + M\int_0^1 q \leqslant 6\epsilon. \tag{17.38}$$

□

17.2.2 Théorème de Müntz

ThoAEYDdHp

Théorème 17.7 (Théorème de Müntz[497, 498, 499]).
Soit $C_0\big([0,1]\big)$, l'espace des fonctions continues sur $[0,1]$ muni de la norme $\|.\|_\infty$ ou $\|.\|_2$ et une suite (α_n) strictement croissante de nombres positifs. Nous notons ϕ_λ la fonction $x\mapsto x^\lambda$.

Alors

$$\overline{\text{Span}\{1,\phi_{\alpha_n}\}} \tag{17.39}$$

est dense dans $C_0\big([0,1]\big)$ si et seulement si

$$\sum_{n=2}^{\infty}\frac{1}{\alpha_n}=+\infty. \tag{17.40}$$

Nous prouvons le théorème pour la norme $\|.\|_2$.

Démonstration. Soit $m\in\mathbb{R}^+$; nous notons $\Delta_N(m)$ la distance entre ϕ_m et $\text{Span}\{\phi_{\alpha_1},\dots,\phi_{\alpha_N}\}$. Cette distance peut être évaluée avec le déterminant de Gram (proposition 8.21)

$$\Delta_N(m)^2=\frac{G(\phi_m,\phi_{\alpha_1},\dots,\phi_{\alpha_N})}{G(\phi_{\alpha_1},\dots,\phi_{\alpha_N})}. \tag{17.41}$$

Pour calculer cela, nous avons besoin des produits scalaires [4]

$$\langle\phi_a,\phi_b\rangle=\int_0^1 x^{a+b}dx=\frac{1}{a+b+1}. \tag{17.42}$$

Pour avoir des notations plus compactes, nous notons $\alpha_0=m$. Donc nous avons à calculer le déterminant

$$G(\phi_m,\phi_{\alpha_1},\dots,\phi_{\alpha_N})=\det\Big(\tfrac{1}{\alpha_i+\alpha_j+1}\Big) \tag{17.43}$$

où $i,j=0,\dots,N$. Nous reconnaissons un déterminant de Cauchy (proposition 8.22) en posant, dans $\frac{1}{\alpha_i+\alpha_j+1}$, $a_i=\alpha_i$ et $b_j=\alpha_j+1$. Étant donné que $b_j-b_i=a_j-a_i$, nous avons

$$G(\phi_m,\phi_{\alpha_1},\dots,\phi_{\alpha_N})=\frac{\prod_{0\leqslant i<j\leqslant N}(\alpha_j-\alpha_i)^2}{\prod_{i=0}^N\prod_{j=0}^N(\alpha_i+\alpha_j+1).} \tag{17.44}$$

Nous séparons maintenant les termes où i ou j sont nuls. En ce qui concerne le dénominateur, il faut prendre tous les couples (i,j) avec i et j éventuellement égaux à zéro. Nous décomposant cela en trois paquets. Le premier est $(0,0)$; le second est $(0,i)$ (chaque couple arrive en fait deux fois parce qu'il y a aussi $(i,0)$) ; et le troisième sont les i,j tous deux différents de zéro :

$$(2m+1)\prod_{ij}(\alpha_i+\alpha_j+1)\prod_i(\alpha_i+m+1)^2. \tag{17.45}$$

Notons que dans le produit central, le carré est contenu dans le fait qu'on écrit $\prod_{ij}$ et non $\prod_{i<j}$. Nous avons donc

$$G(\phi_m,\phi_{\alpha_1},\dots,\phi_{\alpha_N})=\frac{\prod_{i<j}(\alpha_i-\alpha_j)^2\prod_i(\alpha_i-m)^2}{(2m+1)\prod_{ij}(\alpha_i+\alpha_j+1)\prod_i(\alpha_i+m+1)^2}. \tag{17.46}$$

Le calcul de $G(\phi_{\alpha_1},\dots,\phi_{\alpha_N})$ est plus simple [5] :

$$G(\phi_{\alpha_1},\dots,\phi_{\alpha_N})=\frac{\prod_{i<j}(\alpha_i-\alpha_j)^2}{\prod_{ij}(\alpha_i+\alpha_j+1)}. \tag{17.47}$$

4. C'est ici qu'on se particularise à la norme $\|.\|_2$.
5. Je crois qu'il y a une faute de frappe dans le dénominateur de [497].

En divisant l'un par l'autre il ne reste que les facteurs comprenant m et en prenant la racine carrée,

$$\Delta_N(m) = \frac{1}{\sqrt{2m+1}} \prod_{i=1}^{N} \left| \frac{\alpha_i - m}{\alpha_i + m + 1} \right|. \tag{17.48}$$

Nous passons maintenant à la preuve proprement dite. Supposons que $V = \text{Span}\{\phi_{\alpha_i}, i \in \mathbb{N}\}$ est dense. Si m est un des α_i, il peut évidemment être approché par les ϕ_{α_i}. Mais vue la densité de V, un ϕ_m avec $m \neq \alpha_i$ (pour tout i) alors ϕ_m peut également être arbitrairement approché par les ϕ_{α_i}, c'est-à-dire que

$$\lim_{N\to\infty} \Delta_N(m) = 0. \tag{17.49}$$

Nous posons

$$u_n = \ln\left(\frac{\alpha_n - m}{\alpha_n + m + 1} \right) \tag{17.50}$$

et nous prouvons que la série $\sum_n u_n$ diverge. En effet nous nous souvenons de la formule $\ln(ab) = \ln(a) + \ln(b)$, de telle sorte que la N^{e}somme partielle de $\sum_n u_n$ est

$$\ln\left(\frac{\alpha_1 - m}{\alpha_1 + m + 1} \cdot \ldots \cdot \frac{\alpha_N - m}{\alpha_N + m + 1} \right) = \ln\left(\sqrt{2m+1}\Delta_N(m) \right), \tag{17.51}$$

qui tend vers $-\infty$ lorsque $N \to \infty$.

Si la suite (α_n) est majorée et plus généralement si nous n'avons pas $\alpha_n \to \infty$, alors évidemment la série $\sum_n \frac{1}{\alpha_n}$ diverge. Nous supposons donc que $\lim_{n\to\infty} \alpha_n = \infty$. Nous avons aussi [6]

$$u_n = \ln\left(\frac{\alpha_n - m}{\alpha_n + m + 1} \right) = \ln\left(1 - \frac{2m+1}{\alpha_n + m + 1} \right) \sim -\frac{2m+1}{\alpha_n}. \tag{17.52}$$

Une justification est donnée à l'équation (15.358). Ce que nous avons surtout est

$$\sum_n u_n \sim -(2m+1) \sum_n \frac{1}{\alpha_n}. \tag{17.53}$$

Étant donné que la série de gauche diverge, celle de droite diverge [7].

Nous prouvons maintenant le sens opposé : nous supposons que la série $\sum_n 1/\alpha_n$ diverge et nous posons

$$V = \text{Span}\{\phi_{\alpha_n} \text{ tel que } n \in \mathbb{N}\}. \tag{17.54}$$

Il suffit de prouver que $\phi_m \in \bar{V}$ pour tout m parce qu'un corolaire du théorème de Stone-Weierstrass 17.1 montre que $\text{Span}\{\phi_k \text{ tel que } k \in \mathbb{N}\}$ est dense dans C pour la norme $\|.\|_2$.

Si $\alpha_n \to \infty$, nous avons :

$$u_n \sim \frac{2m+1}{\alpha_n} \to 0 \tag{17.55}$$

et alors $\Delta_N(m) \to 0$. Dans ce cas nous avons immédiatement $\phi_m \in \bar{V}$.

Si par contre α_n ne tend pas vers l'infini, nous repartons de l'expression (17.48), nous posons $0 < \alpha = \sup_i \alpha_i$ et nous calculons :

$$\sqrt{2m+1}\Delta_N(m) = \prod_{i=1}^{N} \frac{|\alpha_i - m|}{\alpha_i + m + 1} \tag{17.56a}$$

$$\leqslant \prod_{i=1}^{N} \frac{\alpha_i + m}{\alpha_i + m + 1} = \prod_{i=1}^{N} \left(1 - \frac{1}{\alpha_i + m + 1} \right) \tag{17.56b}$$

$$\leqslant \prod_{i=1}^{N} \left(1 - \frac{1}{\alpha + m + 1} \right) = \left(1 - \frac{1}{\alpha + m + 1} \right)^N. \tag{17.56c}$$

Cette dernière expression tend vers 0 lorsque $N \to \infty$. □

6. Je crois qu'il y a une faute de signe dans la dernière expression de [498].
7. Nous utilisons le fait que si $u_n = \sum v_n$ en tant que suites et si $\sum_n u_n$ diverge, alors $\sum_n v_n$ diverge.

REMooGPYYooCQJwFa

Remarque 17.8.
Certaines sources[8] citent le théorème de Müntz comme ceci (avec un implicite que $\alpha_i \neq 0$) :

$$\overline{\text{Span}\{1, \phi_{\alpha_i}\}} = C\big([0,1]\big) \Leftrightarrow \sum_{i\geqslant 1} \frac{1}{\alpha_i} = +\infty. \tag{17.57}$$

EQooPCSZooUDSzwQ

Que penser de la présence explicite du 1 (c'est-à-dire de ϕ_0) ou non dans l'ensemble ?

Première chose : la présence éventuelle de ϕ_0 est la raison pour laquelle nous faisons commencer la somme à $i = 2$ et non $i = 1$. Dans le même ordre d'idée, si $\text{Span}\{\phi_{\alpha_i}\}$ est dense, alors en prenant n'importe quelle queue de suite, ça reste dense.

Prouvons donc l'énoncé (17.57). Si $\text{Span}\{1, \phi_{\alpha_i}\}$ est dense, alors en posant $\beta_1 = 0$, $\beta_i = \alpha_{i-1}$ notre théorème prouve que $\sum_{i=2}^{\infty} \frac{1}{\beta_i} = +\infty$, cela est exactement que $\sum_{i=1}^{\infty} \frac{1}{\alpha_i} = +\infty$. Dans l'autre sens, si $\sum_{i\geqslant 1} \frac{1}{\alpha_i} = +\infty$, alors nous avons aussi $\sum_{i\geqslant 2} \frac{1}{\alpha_i} = +\infty$ et notre théorème dit que $\text{Span}\{\phi_{\alpha_i}\}$ est dense. A fortiori, $\text{Span}\{1, \phi_{\alpha_i}\}$ est dense.

Exemple 17.9.
Nous savons depuis le théorème 15.120 que la somme des inverses des nombres premiers diverge. △

17.3 Intégrales convergeant uniformément

17.3.1 Définition et propriété

DEFooSHWAooWtswtp

Définition 17.10.
Soit (Ω, μ) un espace mesuré. Nous disons que l'intégrale

$$\int_\Omega f(x,\omega) d\mu(\omega) \tag{17.58}$$

converge uniformément *en x si pour tout $\epsilon > 0$, il existe un compact K_ϵ tel que pour tout compact K tel que $K_\epsilon \subset K$ nous avons*

$$\left| \int_{\Omega\setminus K} f(x,\omega) d\mu(\omega) \right| \leqslant \epsilon. \tag{17.59}$$

Le point important est que le choix de K_ϵ ne dépend pas de x.

LemOgQdpJ

Lemme 17.11.
Soit

$$F(x) = \int_\Omega f(x,\omega) d\mu(\omega), \tag{17.60}$$

une intégrale uniformément convergente. Pour chaque $k \in \mathbb{N}$ nous considérons un compact K_k tel que

$$\left| \int_{\Omega\setminus K_k} f(x,\omega) d\mu(\omega) \right| \leqslant \frac{1}{k}. \tag{17.61}$$

Alors la suite de fonctions F_k définie par

$$F_k(x) = \int_{K_k} f(x,\omega) d\mu(\omega) \tag{17.62}$$

converge uniformément vers F.

8. Dont le rapport du jury 2014

Démonstration. Nous avons

$$\big|F_k(x) - F(x)\big| = \left|\int_{K_k} f(x,\omega)d\mu(\omega) - \int_{\Omega} f(x,\omega)d\mu(\omega)\right| \tag{17.63a}$$

$$= |\int_{\Omega\setminus K_k} f(x,\omega)d\mu(\omega)| \tag{17.63b}$$

$$\leqslant \frac{1}{k}. \tag{17.63c}$$

□

17.3.2 Critères de convergence uniforme

Afin de tester l'uniforme convergence d'une intégrale, nous avons le **critère de Weierstrass** :

ThoCritWeiIntUnifCv

Théorème 17.12.
Soit $f(x,t)\colon [\alpha,\beta]\times[a,\infty[\to\mathbb{R}$, une fonction dont la restriction à toute demi-droite $x=c$ est mesurable. Si $|f(x,t)|<\varphi(t)$ et $\int_a^\infty \varphi(t)dt$ existe, alors l'intégrale

$$\int_0^\infty f(x,t)dt \tag{17.64}$$

est uniformément convergente.

Le théorème suivant est le **critère d'Abel** :

ThoAbelIntUnif

Théorème 17.13.
Supposons que $f(x,t)=\varphi(x,t)\psi(x,t)$ où φ et ψ sont bornées et intégrables en t au sens de Riemann sur tout compact $[a,b]$, $b\geqslant a$. Supposons que :

(1) $\left|\int_a^T \varphi(x,t)dt\right|\leqslant M$ *où M est indépendant de T et de x,*

(2) $\psi(x,t)\geqslant 0$,

(3) pour tout $x\in[\alpha,\beta]$, $\psi(x,t)$ est une fonction décroissante de t,

(4) les fonctions $x\mapsto\psi(x,t)$ convergent uniformément vers 0 lorsque $t\to\infty$.

Alors l'intégrale

$$\int_a^\infty f(x,t)dt \tag{17.65}$$

est uniformément convergente.

Remarque 17.14.
Étant donné que la fonction sinus est bornée, il est tentant de l'utiliser comme φ dans le critère d'Abel (théorème 17.13). Hélas,

$$\int_0^T \sin(xt)dt = -\frac{1}{x}\big(\cos(xT)-\cos(x)\big), \tag{17.66}$$

qui n'est pas bornée en x. Poser $\varphi(x,t)=\sin(xt)$ *ne fonctionne pas* pour assurer la convergence uniforme sur un intervalle qui contient des x arbitrairement proches de 0. Le critère d'Abel avec $\varphi(x,t)=\sin(xt)$ ne permet que de conclure à l'uniforme convergence *sur tout compact* ne contenant pas 0. C'est toutefois souvent suffisant pour étudier la continuité ou la dérivabilité en se servant coup du compact, voir 17.18.

17.4 Fonctions définies par une intégrale

SecCHwnBDj

Soit (Ω, μ) un espace mesuré. Nous nous demandons dans quel cas l'intégrale

$$F(x) = \int_\Omega f(x, \omega) d\omega \tag{17.67}$$

définit une fonction F continue, dérivable ou autre.

Dans la suite nous allons considérer des fonctions f à valeurs réelles. Quitte à passer aux composantes, nous pouvons considérer des fonctions à valeurs vectorielles. Par contre le fait que x soit dans $\mathbb{R}$ ou dans $\mathbb{R}^n$ n'est pas spécialement une chose facile à traiter.

17.4.1 Continuité sous l'intégrale

Nous allons présenter deux théorèmes donnant la continuité de F.

(1) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 17.15,

(2) si l'intégrale est uniformément convergente, nous avons le théorème 17.16.

ThoKnuSNd

Théorème 17.15.
Soit (Ω, μ) est un espace mesuré, soit $x_0 \in \mathbb{R}^m$ et $f : U \times \Omega \to \mathbb{R}$ où U est ouvert dans $\mathbb{R}^m$. Nous supposons que

(1) Pour chaque $x \in \mathbb{R}^m$, la fonction $\omega \mapsto f(x, \omega)$ est dans $L^1(\Omega, \mu)$.

(2) Pour chaque $\omega \in \Omega$, la fonction $x \mapsto f(x, \omega)$ est continue en x_0.

ItemNAuYNG

(3) Il existe une fonction $G \in L^1(\Omega)$ telle que

$$|f(x, \omega)| \leqslant G(\omega) \tag{17.68}$$

pour tout $x \in U$.

Alors la fonction

$$\begin{aligned} F : U &\to \mathbb{R} \\ x &\mapsto \int_\Omega f(x, \omega) d\mu(\omega) \end{aligned} \tag{17.69}$$

est continue en x_0.

Démonstration. Soit (x_n) une suite convergente vers x_0. Nous considérons la suite de fonctions $f_n : \Omega \to \mathbb{R}$ définies par

$$f_n(\omega) = f(x_n, \omega). \tag{17.70}$$

sur qui nous pouvons utiliser le théorème de la convergence dominée (théorème 14.207) pour obtenir

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \int_\Omega f(x_n, \omega) d\mu(\omega) \tag{17.71a}$$

$$= \int_\Omega \lim_{n\to\infty} f(x_n, \omega) d\mu(\omega) \tag{17.71b}$$

$$= \int_\Omega f(x_0, \omega) d\mu(\omega) \tag{17.71c}$$

$$= F(x_0). \tag{17.71d}$$

Nous avons utilisé la continuité de $f(., \omega)$. □

Si nous avons un peu de compatibilité entre la topologie et la mesure, alors nous pouvons utiliser l'uniforme convergence d'une intégrale pour obtenir la continuité d'une fonction définie par une intégrale.

ThotexmgE

Théorème 17.16.
Soit (Ω, μ) un espace topologique mesuré tel que tout compact est de mesure finie. Soit une fonction $f : \mathbb{R} \times \Omega \to \mathbb{R}$ telle que

(1) Pour chaque $x \in \mathbb{R}$, la fonction $f(x, .)$ est $L^1(\Omega, \mu)$.

(2) Pour chaque $\omega \in \Omega$, la fonction $f(., \omega)$ est continue en x_0.

(3) L'intégrale

$$F(x) = \int_\Omega f(x, \omega) d\mu(\omega) \tag{17.72}$$

est uniformément convergente[9].

Alors la fonction F est continue en x_0.

Démonstration. Nous reprenons les notations du lemme 17.11. Les fonctions

$$F_k(x) = \int_{K_k} f(x, \omega) d\mu(\omega) \tag{17.73}$$

existent parce que les fonctions $f(x, .)$ sont dans $L^1(\Omega)$. Montrons que les fonctions F_k sont continues. Soit une suite $x_k \to x_0$ nous avons

$$\lim_{n\to\infty} F_k(x_n) = \lim_{n\to\infty} \int_{K_k} f(x_n, \omega) d\mu(\omega). \tag{17.74}$$

Nous pouvons inverser la limite et l'intégrale en utilisant le théorème de la convergence dominée. Pour cela, la fonction $f(x_n, \omega)$ étant continue sur le compact K_k, elle y est majorée par une constante. Le fait que les compacts soient de mesure finie (hypothèse) implique que les constantes soient intégrables sur K_k. Le théorème de la convergence dominée implique alors que

$$\lim_{n\to\infty} F_k(x_n) = \int_{K_k} \lim_{n\to\infty} f(x_n, \omega) d\mu(\omega) = \int_{K_k} f(x_0, \omega) d\mu(\omega) = F_k(x_0). \tag{17.75}$$

Nous avons utilisé le fait que $f(., \omega)$ était continue en x_0.

Le lemme 17.11 nous indique alors que la convergence $F_k \to F$ est uniforme. Les fonctions F_k étant continues, la fonction F est continue. □

Pour finir, citons ce résultat concernant les fonctions réelles.

ThoInDerrtCvUnifFContinue

Théorème 17.17.
Nous considérons $F(x) = \int_a^\infty f(x, t)dt$. Si f est continue sur $[\alpha, \beta] \times [a, \alpha[$ et l'intégrale converge uniformément, alors $F(x)$ est continue.

17.4.2 Le coup du compact

NORMooZWECooHvRgBw

17.18.
Nous avons vu des fonctions définies par toute une série de processus de limite (suites, séries, intégrales). Une des questions centrales est de savoir si la fonction limite est continue, dérivable, intégrable, etc. étant donné que les fonctions sont continues.

Pour cela, nous inventons le concept de *convergence uniforme*. Si la limite (série, intégrale) est uniforme, alors la fonction limite sera continue. Il arrive qu'une limite ne soit pas uniforme sur un intervalle ouvert $]0, 1]$, et que nous voulions quand même prouver la continuité sur cet intervalle. C'est à cela que sert la notion de convergence uniforme *sur tout compact*. En effet, la notion de continuité est une notion locale : savoir ce qu'il se passe dans un petit voisinage autour de x est suffisant pour savoir la continuité en x (idem pour sa dérivée).

9. Définition 17.10.

Si nous avons uniforme convergence sur tout compact de $]0,1]$, mais pas uniforme convergence sur cet intervalle, la limite sera quand même continue sur $]0,1]$. En effet, si $x \in]0,1]$, il existe un ouvert autour de x contenu dans un compact contenu dans $]0,1]$. L'uniforme convergence sur ce compact suffit à prouver la continuité en x.

Déduire la continuité sur un ouvert à partir de l'uniforme convergence sur tout compact de l'ouvert est appelé faire le **coup du compact**.

17.4.3 Dérivabilité sous l'intégrale

Nous traitons à présent de la dérivabilité de la fonction F définie comme intégrale de f. Dans le théorème 17.19 nous traitons de fonctions sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$. Pour les fonctions définies sur $\mathbb{C}$, voir le théorème 26.73.

ThoMWpRKYp

Théorème 17.19 (Dérivation sous le signe intégral, formule de Leibniz, thème 62[468, 500])**.**
Soit (Ω, μ) un espace mesuré, un intervalle ouvert $I \subset \mathbb{R}$, et une fonction $f : \mathbb{R} \times \Omega \to \mathbb{C}$. Nous supposons qu'il existe A mesurable de mesure nulle dans Ω tels que

ITEMooAFVMooAeCEco

(1) Pour tout $x \in I$, la fonction $\omega \mapsto f(x,\omega)$ soit dans $L^1(\Omega)$.

ITEMooXIZXooGPYFyT

(2) L'application $x \mapsto f(x,\omega)$ est dérivable[10] pour tout $x \in I$ et pour tout $\omega \in \Omega \backslash A$.

ITEMooDTTIooWkldfB

(3) Il existe une fonction $G : \Omega \to \mathbb{R}^+$ intégrable sur Ω telle que

EQooUHQYooHtwfML

$$\left|\frac{\partial f}{\partial x}(x,\omega)\right| \leqslant G(\omega) \tag{17.76}$$

pour tout $x \in I$ et pour tout $\omega \in \Omega \backslash A$.

Alors la fonction

$$\begin{aligned} F : \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \int_\Omega f(x,\omega) d\mu(\omega) \end{aligned} \tag{17.77}$$

est dérivable sur I et pour tout $a \in I$ nous avons

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega) d\mu(\omega), \tag{17.78}$$

Démonstration. Soit une suite (x_n) dans I telle que $x_n \neq a$ et $x_n \to a$. Si la limite

$$\lim_{n\to\infty} \frac{F(a) - F(x_n)}{a - x_n} \tag{17.79}$$

existe et ne dépend pas de la suite choisie, alors la fonction F est dérivable en a et sa dérivée vaut cette limite. Autrement dit, nous nous mettons en devoir d'étudier la limite

EqLIiralx

$$\lim_{n\to\infty} \int_\Omega \frac{f(a,\omega) - f(x_n,\omega)}{a - x_n} d\omega. \tag{17.80}$$

montrer qu'elle existe, ne dépend pas de la suite choisie et vaut $\int_\Omega \partial_x f(a,\omega) d\omega$. On y va.

(1) **La bonne suite de fonctions** D'abord nous posons

EqAEQUbQB

$$g_n(\omega) = \frac{f(x_n,\omega) - f(a,\omega)}{x_n - a}. \tag{17.81}$$

Nous montrons à présent que cette suite vérifie les hypothèses du théorème de la convergence dominée 14.207.

10. La dérivabilité pour une fonction à valeurs dans $\mathbb{C}$ n'a rien de mystérieux : c'est la dérivée composante par composante. Rien à voir avec la dérivée complexe.

— Chacune des fonctions g_n est dans $L^1(\Omega)$ parce que, a étant fixé, l'élément x_n est dans $I\backslash\{a\}$; le dénominateur n'a donc aucun rôle. L'hypothèse (1) montre que $\omega \mapsto f(x_n,\omega)$ et $\omega \mapsto f(a,\omega)$ sont dans L^1. La somme est donc dans L^1 (proposition 14.193).

— Par l'hypothèse (2), pour chaque ω nous avons une fonction dérivable. Nous pouvons donc passer à la limite :

$$\lim_{n\to\infty} g_n(\omega) = \frac{\partial f}{\partial x}(a,\omega). \tag{17.82}$$

— En ce qui concerne la majoration de g_n, nous utilisons le théorème des accroissements finis 12.192(2). Pour chaque ω, ce théorème peut être utilisé sur la fonction $x \mapsto f(x,\omega)$ qui est dérivable. Nous avons :

$$\Big|\frac{f(x_n,\omega)-f(a,\omega)}{x_n-a}\Big| \leqslant \sup_{x\in[a,x_n]} \Big|\frac{\partial f}{\partial x}(x,\omega)\Big| \leqslant G(\omega). \tag{17.83}$$

Nous avons utilisé l'hypothèse (3).

Les hypothèses de la convergence dominée sont satisfaites.

(2) **Convergence dominée** Le théorème de la convergence dominée de Lebesgue (théorème 14.207) nous permet alors de calculer la limite (17.80) :

$$\lim_{n\to\infty}\int_\Omega g_n(\omega)d\omega = \int_\Omega \lim_{n\to\infty} g_n(\omega)d\omega = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\omega. \tag{17.84}$$

Notons que l'existence de la dernière intégrale fait partie du théorème de la convergence dominée.

Nous avons donc prouvé que la limite de gauche existait et ne dépendait pas de la suite choisie. Donc F est dérivable en a et la dérivée vaut cette limite :

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\mu(\omega). \tag{17.85}$$

□

En ce qui concerne les fonctions dans $\mathbb{R}^n$, il y a les propositions 17.24 et 17.25 qui parlent de différentiabilité sous l'intégrale.

¡¡ Avertissement/question au lecteur !! 17.20

Attention : l'énoncé et la démonstration de la proposition 17.21 sont de moi. Écrivez-moi pour

— *me dire si ça vous semble correct (ou pas),*

— *me donner un lien vers une source qui énonce et démontre ce résultat.*

PROPooJKXJooLxgEGd

Proposition 17.21 ([1]).
Soient un espace mesuré (Ω,μ) ainsi qu'une fonction $f\colon \mathbb{R}^n\times\Omega\to\mathbb{C}$. Si $\omega\in\Omega$ est fixé, nous notons

$$\begin{aligned} f_\omega\colon \mathbb{R}^n &\to \mathbb{C}\\ x &\mapsto f(x,\omega) \end{aligned} \tag{17.86}$$

et si $x\in\mathbb{R}^n$ est fixé, nous notons

$$\begin{aligned} f_x\colon \Omega &\to \mathbb{C}\\ \omega &\mapsto f(x,\omega). \end{aligned} \tag{17.87}$$

Soient $\delta>0$, A de mesure nulle dans Ω et une liste d'indices[11] α tels que

(1) pour tout $x\in\mathbb{R}^n$, la fonction f_x est dans $L^1(\Omega)$,

(2) la dérivée partielle multiple $\partial^\alpha f_\omega(x)$ existe pour tout $x\in B(a,\delta)$ et pour tout $\omega\in A^c$.

11. Voir -1.1.

(3) pour toute queue β de la liste d'indice α, il existe une fonction $G_\beta \in L^1(\Omega)$ telle que

$$|(\partial^\beta f_\omega)(x)| \leqslant G_\beta(\omega) \tag{17.88}$$

pour tout $x \in B(a,\delta)$ et $\omega \in A^c$.

Enfin nous posons

$$F(x) = \int_\Omega f(x,\omega)d\mu(\omega). \tag{17.89}$$

Alors pour toute queue de suite β de α, nous avons

(1) $(\partial^\beta F)$ existe en a,

(2) nous avons la formule

$$(\partial^\beta F)(x) = \int_\Omega (\partial^\beta f_\omega)(a)d\mu(\omega). \tag{17.90}$$

EQooTDDWooAcLRwX

Démonstration. Nous allons opérer une récurrence sur α. Plus précisément, si $|\alpha| = p$, nous allons ajouter une dérivation à la fois, et dans l'ordre inverse de α. Donc nous commençons par ∂_{α_p}, puis par $(\partial_{p-1}, \partial_p)$, etc.

Nous commençons par prouver la formule (17.90) dans le cas de $\beta = (\alpha_p)$. Et pour alléger les notations nous notons $\alpha_p = i$. Nous posons

$$\begin{aligned}\varphi\colon \mathbb{R}\times\Omega &\to \mathbb{C}\\ (t,\omega) &\mapsto f(a+te_i,\omega).\end{aligned} \tag{17.91}$$

Posons $H(t) = F(a+te_i)$, c'est-à-dire

$$H(t) = \int_\Omega f(a+te_i,\omega)d\mu(\omega) = \int_\Omega \varphi(t,\omega)d\mu(\omega). \tag{17.92}$$

En utilisant le théorème 17.19 sur la fonction φ nous trouvons

$$H'(0) = \int_\Omega \varphi'(0,\omega)d\mu(\omega) = \int_\Omega \frac{\partial f}{\partial x_i}(a,\omega)d\mu(\omega), \tag{17.93}$$

ce qui est la formule demandée dans le cas $\alpha = (i)$.

Pour la récurrence, nous supposons que la formule est démontrée pour $\beta = (\alpha_{p-k},\dots,\alpha_p)$, et nous montrons qu'elle fonctionne encore pour $\sigma = (i,\beta)$.

Il s'agit simplement de remarquer que la fonction

$$g(x,\omega) = (\partial^\beta f_\omega)(x) \tag{17.94}$$

vérifie encore les conditions du théorème 17.19 [12]. □

17.4.4 Dérivation et intégrale

PROBooGVWHooOTNJFe

¡¡ Avertissement/question à la lectrice !! 17.22
Si je regarde bien, le théorème 17.23 n'est utilisé nulle part. Je vais donc le supprimer un jour. Si vous l'aimez, vous devriez m'en envoyer une démonstration.

ThoDerSousIntegrale

Théorème 17.23.
Soient A un intervalle ouvert de $\mathbb{R}$ et Ω un espace mesuré. Soient une fonction $f\colon A\times\Omega\to\mathbb{R}$ et

$$F(x) = \int_\Omega f(x,\omega)d\omega. \tag{17.95}$$

Nous supposons les points suivants.

(1) La fonction f est mesurable.

12. Je n'ai pas fait cette vérification. Écrivez-moi si vous l'avez faite.

(2) *Pour chaque $x \in A$, la fonction $\omega \mapsto f(x,\omega)$ est intégrable sur Ω.*
(3) *Pour presque tout $\omega \in \Omega$, la fonction $f(x,\omega)$ est une fonction absolument continue*[13] *de x.*
(4) *La fonction $\frac{\partial f}{\partial x}$ est localement intégrable, c'est-à-dire que pour tout $[a,b] \subset A$,*

$$\int_a^b \int_\Omega \left| \frac{\partial f}{\partial x}(x,\omega) \right| d\omega\, dx < \infty. \tag{17.96}$$

Alors la fonction F est absolument continue et pour presque tout $x \in A$, la dérivée est donné par

$$\frac{d}{dx} \int_\Omega f(x,\omega) d\omega = \int_\Omega \frac{\partial f}{\partial x}(x,\omega) d\omega. \tag{17.97}$$

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

17.4.5 Différentiabilité sous l'intégrale

Le théorème suivant est restrictif sur l'ensemble d'intégration (qui doit être compact), mais accepte des fonctions de plusieurs variables, ce qui est un premier pas vers la différentiabilité.

PropDerrSSIntegraleDSD

Proposition 17.24 (Dérivation sous l'intégrale)**.**
Supposons $A \subset \mathbb{R}^m$ ouvert et $B \subset \mathbb{R}^n$ compact. Nous considérons une fonction $f : A \times B \to \mathbb{R}$. Si pour un $i \in \{i, \ldots, n\}$, la dérivée partielle $\frac{\partial f}{\partial x_i}$ existe dans $A \times B$ et est continue, alors la fonction

$$F(x) = \int_B f(x,t) dt \tag{17.98}$$

admet une dérivée partielle dans la direction x_i sur A. Cette dérivée partielle y est continue et

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a,t) dt, \tag{17.99}$$

pour tout a dans l'ouvert A.

Démonstration. Nous procédons en plusieurs étapes.

(1) **F est dérivable** Nous voulons prouver que $\frac{\partial F}{\partial x_i}(a,t)$ existe. Pour cela nous posons

$$g_l(t) = \frac{f(a_1, \ldots, a_i + \epsilon_l, \ldots, a_n, t) - f(a_1, \ldots, a_i, \ldots, a_n, t)}{\epsilon_l} \tag{17.100}$$

où ϵ_l est une suite de nombres tendant vers zéro. La fonction f est dérivable dans la direction x_i si et seulement si $\lim_{l\to\infty} g_l(t)$ existe et ne dépend pas du choix de la suite. À ce moment, la valeur de la dérivée partielle sera cette limite. Dans notre cas, nous savons que f admet une dérivée partielle dans la direction x_i et donc nous avons

$$\frac{\partial f}{\partial x_i}(a,t) = \lim_{l\to\infty} g_l(t). \tag{17.101}$$

De la même façon pour F nous avons

$$\frac{\partial F}{\partial x_i} = \lim_{l\to\infty} \int_B g_l(t) dt. \tag{17.102}$$

Sous-entendu : si la limite de droite ne dépend pas de la suite choisie, alors $\frac{\partial F}{\partial x_i}$ existe et vaut cette limite.

Considérant la continuité de f, le seul point à vérifier pour le théorème de la convergence dominée de Lebesgue est l'existence d'une fonction intégrable de t majorant g_l. Pour cela le

13. Définition 14.313 et proposition 14.316.

théorème de accroissements finis (théorème 12.192) appliqué à la fonction $\epsilon \mapsto f(a_n, \ldots, a_i + \epsilon, \ldots, a_n)$ nous dit que

$$f(a_1, \ldots, a_i + \epsilon_l, \ldots, a_n, t) - f(a_1, \ldots, a_i, \ldots, a_n, t) = \epsilon_l \frac{\partial f}{\partial x_i}(a_1, \ldots, \theta, \ldots, a_n, t) \quad (17.103)$$

pour un certain $\theta \in B(a_i, \epsilon_l)$. Notons que ce θ dépend de t mais pas de l. Vu que $\partial_i f$ est continue par rapport à ses deux variables, si K est un voisinage compact autour de a, il existe $M > 0$ tel que

$$\left| \frac{\partial f}{\partial x_i}(x, t) \right| < M \quad (17.104)$$

EqMXqviPC

pour tout $x \in K$ et tout $t \in B$. La valeur de $\frac{\partial f}{\partial x_i}(a_1, \ldots, \theta, \ldots, a_n, t)$ est donc bien majorée par rapport à θ et par rapport à t en même temps par une constante qui n'a pas de mal à être intégrée sur le compact B.

Le théorème de la convergence dominée (théorème 14.207) s'applique donc bien et nous avons

$$\lim_{l \to \infty} \int_B g_l(t) dt = \int_B \lim_{l \to \infty} g_l(t) = \int_B \frac{\partial f}{\partial x_i}(a, t) dt. \quad (17.105)$$

Le membre de droite ne dépendant pas de la suite ϵ_l choisie, le membre de gauche est bien la dérivée de F par rapport à x_i et nous avons

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a, t) dt. \quad (17.106)$$

Cela prouve la première partie de la proposition.

(2) **La dérivée est continue** Soit K un voisinage compact autour de a et U' un ouvert tel que $a \in U' \subset K$. Nous avons encore la majoration (17.104) sur U' et donc le théorème de continuité sous l'intégrale 17.15 nous indique que la fonction

$$\begin{aligned} U' &\to \mathbb{R} \\ x &\mapsto \int_B \frac{\partial f}{\partial x_i}(x, t) dt \end{aligned} \quad (17.107)$$

est continue en a.

□

Une conséquence de la proposition 17.24 est que si elle fonctionne pour tous les i, alors F est différentiable et même de classe C^1, et la différentielle de F s'obtient comme intégrale de la différentielle de f.

PropAOZkDsh

Proposition 17.25.
Supposons $A \subset \mathbb{R}^m$ ouvert et $B \subset \mathbb{R}^n$ compact. Si pour tout $i \in \{i, \ldots, n\}$, la dérivée partielle $\frac{\partial f}{\partial x_i}$ existe dans $A \times B$ et est continue, alors F est de classe C^1 et

$$(dF)_a = \int_B (df_t)_a dt \quad (17.108)$$

où $f_t(x) = f(x, t)$.

Démonstration. En vertu de la proposition 17.24, toutes les dérivées partielles de F sont continues. Cela implique que F est de classe C^1 par le théorème 12.294 et que la différentielle s'écrive en termes

des dérivées partielles avec la formule usuelle. Nous avons alors

$$(dF)_a(u) = \sum_k \frac{\partial F}{\partial x_k}(a)u_k \tag{17.109a}$$
$$= \int_B \sum_k \frac{\partial f}{\partial x_k}(a,t)dt \tag{17.109b}$$
$$= \int_B \sum_k \frac{\partial f_t}{\partial x_k}(a)u_k dt \tag{17.109c}$$
$$= \int_B (df_t)_a(u)dt. \tag{17.109d}$$

Ce qui est la formule annoncée. □

Un autre théorème tourne autour du pot, et me semble inutile.

ThoOLAQyRL

Théorème 17.26.
Soit (Ω, μ) un espace mesuré, une fonction $f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ et $a \in \mathbb{R}^n$. Nous considérons la fonction

$$F(x) = \int_\Omega f(x,\omega)d\mu(\omega). \tag{17.110}$$

Pour chaque $k = 1, \ldots, n$ nous supposons avoir

$$\frac{\partial F}{\partial x_k}(a) = F'_{|_k}(a) = \int_\Omega \frac{\partial f_{|_k}}{\partial t}(a_k,\omega)d\mu(\omega) \tag{17.111}$$

où $F_{|_k}(t) = F(a_1, \ldots, t, \ldots, a_n)$ et $f_{|_k}$ est définie de façon similaire.
Nous supposons de plus que les fonctions $\partial_{x_k} F$ sont continues.
Alors F est de classe C^1 et sa différentielle est donnée par

$$dF_a = \int_\Omega (df_\omega)_a d\omega \tag{17.112}$$

où f_ω est définie par $f_\omega(x) = f(x,\omega)$.

Démonstration. Étant donné que les dérivées partielles de F en a existent et sont continues, le théorème 12.294 dit que F est différentiable et que

$$dF_a(u) = \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a)u_k. \tag{17.113}$$

La linéarité de l'intégrale et les hypothèses nous donnent alors

$$dF_a(u) = \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a)u_k \tag{17.114a}$$
$$= \int_\Omega \sum_k \frac{\partial f_{|_k}}{\partial t}(a_k;\omega)u_k d\mu(\omega) \tag{17.114b}$$
$$= \int_\Omega \sum_k \frac{\partial f}{\partial x_k}(a;\omega)u_k d\mu(\omega) \tag{17.114c}$$
$$= \int_\Omega (df_\omega)_a(u)d\mu(\omega), \tag{17.114d}$$

et donc $dF_a = \int_\Omega (df_\omega)_a d\mu(\omega)$. □

Notons qu'en passant aux composantes, ce théorème fonctionne tout aussi bien pour des fonctions à valeurs dans un espace vectoriel normé de dimension finie plutôt que dans $\mathbb{R}$.

17.5 Théorème d'Hadamard

LemWNBooGPlIwT

Lemme 17.27 (Hadamard[501]).
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^p avec $p \geqslant 1$. Pour tout $a \in \mathbb{R}^n$ il existe des fonctions $g_1,\dots, g_n$ de classe C^{p-1} telles que

$$f(x) = f(a) + \sum_{i=1}^{n}(x_i - a_i)g_i(x). \tag{17.115}$$

Démonstration. Puisque f est de classe C^1, le théorème fondamental de l'analyse 14.287 s'applique et

$$f(x) - f(a) = \int_0^1 \frac{d}{dt}\Big[f\big(a + t(x-a)\big)\Big]dt = \int_0^1 \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)(x_i - a_i). \tag{17.116}$$

EqZLTooVKmGlu

Plus de détails : la fonction $t \mapsto \frac{d}{dt}\Big[f\big(a + t(x-a)\big)\Big]$ possède comme primitive la fonction $F(t) = f\big(a + t(x-a)\big)$.

Nous posons

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)dt \tag{17.117}$$

L'intégrale existe parce qu'il s'agit d'une fonction continue sur un compact et donc majorée par une constante. Pour voir que g_i est de classe C^{p-1} nous pouvons calculer $\frac{\partial g_i}{\partial x_k}$ en permutant dérivée et intégrale par la proposition 17.24 :

$$\frac{\partial g_i}{\partial x_k}(x) = \int_0^1 \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_i}\big(a + t(x-a)\big)\right)dt = \int_0^1 t\frac{\partial^2 f}{\partial x_k \partial x_i}\big(a + t(x-a)\big). \tag{17.118}$$

Nous pouvons ainsi permuter $p-1$ dérivées tout en gardant une fonction continue dans l'intégrale. Le théorème 17.15 nous donne alors une fonction continue. Ainsi toutes les fonctions

$$\frac{\partial^{p-1} g_i}{\partial x_{i_1} \dots \partial x_{i_{p-1}}} \tag{17.119}$$

sont continues et g_i est de classe C^{p-1} par le théorème 12.330.

En repartant de (17.116) nous avons alors bien ce qui était annoncé :

$$f(x) = f(a) + \sum_{i=1}^{n} g_i(x)(x_i - a_i). \tag{17.120}$$

□

CorQBXHooZVKeNG

Corolaire 17.28.
Soit $\phi \in \mathscr{D}(\mathbb{R})$ tel que $\phi^{(k)}(x_0) = 0$ pour tout $k \leqslant n$. Alors il existe une fonction $\psi \in \mathscr{D}(\mathbb{R})$ telle que

$$\phi(x) = (x - x_0)^{n+1}\psi(x) \tag{17.121}$$

pour tout $x \in \mathbb{R}$.

Démonstration. En utilisant le lemme de Hadamard 17.27 avec $a = x_0$, $n = 1$ et $f(x_0) = 0$, nous avons une fonction g_1 à support compact telle que

$$\phi(x) = \phi(x_0) + (x - x_0)g_1(x). \tag{17.122}$$

EqTOJGooWZBBRJ

Alors $\phi'(x) = g_1(x) + (x - x_0)g_1'(x)$, ce qui donne immédiatement $g_1(x_0) = 0$ et donc une fonction g_2 telle que $g_1(x) = (x - x_0)g_2(x)$. En injectant dans (17.122) nous avons

$$\phi(x) = (x - x_0)^2 g_2(x). \tag{17.123}$$

Il suffit de continuer ainsi tant que les dérivées de ϕ s'annulent. □

17.6 Deux théorèmes de point fixe

Nous allons voir Picard. Les autres théorème de point fixe que sont Brouwer, Schauder et Markov-Kakutani sont plus bas [14] parce qu'ils utilisent de l'intégration. Voir le thème 65 pour les retrouver.

17.6.1 Points fixes attractifs et répulsifs

DEFooTMZUooMoBDGC

Définition 17.29 ([1]).
*Soit I un intervalle fermé de $\mathbb{R}$ et $\varphi\colon I \to I$ une application C^1. Soit a un point fixe de φ. Nous disons que a est **attractif** si il existe un voisinage V de a tel que pour tout $x_0 \in V$ la suite $x_{n+1} = \varphi(x_n)$ converge vers a.*

*Le point a est dit **répulsif** si il existe un voisinage [15] V de a tel que pour toute suite de la forme $x_{n+1} = \varphi(x_n)$, si $u_n \in V$, alors $|x_{n+1} - a| \geqslant |u_n - a|$.*

17.30.
Beaucoup de sources prennent le lemme 17.31 comme définition de point fixe attractif et répulsif. Celle en termes de suites ne demande au moins pas la dérivabilité de f. Elle est donc un tout petit peu plus générale.

LEMooTIRAooQIHLyt

Lemme 17.31 ([502, 503]).
Soient un intervalle I de $\mathbb{R}$, et une application $\varphi\colon I \to I$ de classe C^1. Soit a un point fixe de φ.

(1) Si $|\varphi'(a)| < 1$ alors a est attractif et la convergence est au moins exponentielle.

(2) Si $|\varphi'(a)| > 1$ alors a est répulsif et la divergence est au moins exponentielle.

Démonstration. Si $|\varphi'(a)| < 1$ alors il existe k tel que $|\varphi'(a)| < k < 1$ et par continuité il existe un voisinage V de a dans lequel $|\varphi'(x)| < k$ pour tout $x \in V$. En utilisant le théorème des accroissements finis [16] nous avons

$$|x_n - a| = \big|\varphi(x_n) - a\big| \leqslant k|x_{n-1} - a| \tag{17.124}$$

et par récurrence

$$|x_n - a| \leqslant k^n |x_0 - a|. \tag{17.125}$$

Pour la seconde partie, nous supposons que $|f'(a)| > 1$. Il existe un voisinage V sur lequel $f' > 1$. La formule des accroissements finis donne une application $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que

$$f(a + h) = f(a) + hf'(a) + \alpha(h) \tag{17.126}$$

et $\alpha(h)/h \to 0$. Nous restreignons V pour que tout $x \in V$ nous ayons $|\alpha(x-a)/|x-a| + |f'(a)| < 1$. Bref. Supposons que $x_n \in V$ avec $x_n \neq a$. Nous avons alors

$$\begin{aligned} x_{n+1} &= f(x_n) && \text{(17.127a)}\\ &= f\big(a + (x_n - a)\big) && \text{(17.127b)}\\ &= f(a) + (x_n - a)f'(a) + \alpha(x_n - a). && \text{(17.127c)} \end{aligned}$$

Nous avons donc

$$\frac{|x_{n+1} - f(a)|}{|x_n - a|} \leqslant |f'(a)| + \frac{|\alpha(x_n - a)|}{|x_n - a|}. \tag{17.128}$$

Avec toutes les hypothèses que nous avons prises, nous avons

$$\frac{|x_{n+1} - f(a)|}{|x_n - a|} \leqslant 1, \tag{17.129}$$

et donc $|x_{n+1} - f(a)| < |x_n - a|$. □

14. Dans la section 20.5.

15. Cette partie de la définition est de moi. Je suis loin d'être sûr que c'est une bonne définition. Si vous avez une idée, écrivez-moi. Sinon, tenez-vous en à prendre le lemme 17.31 comme définition.

16. Proposition 10.257.

Remarque 17.32.
Dans le cas $|\varphi'(a)| = 1$, nous ne pouvons rien conclure. Si $\varphi(x) = \sin(x)$ nous avons $\sin(x) < x$ et le point $a = 0$ est attractif. À contrario, si $\varphi(x) = \sinh(x)$ nous avons $|\sinh(x)| > |x|$ et le point $a = 0$ est répulsif.

17.6.2 Picard

DEFooRSLCooAsWisu

Définition 17.33.
*Une application $f\colon (X, d_X) \to (Y, d_Y)$ entre deux espaces métriques est une **contraction** si elle est k-**Lipschitz** pour un certain $0 \leqslant k < 1$, c'est-à-dire si pour tout $x, y \in X$ nous avons*

$$d_Y\big(f(x), f(y)\big) \leqslant k d_X(x, y). \tag{17.130}$$

ThoEPVkCL

Théorème 17.34 (Picard [504, 505] [17].)**.**
Soit X un espace métrique complet et $f\colon X \to X$ une application contractante, de constante de Lipschitz k. Alors f admet un unique point fixe, nommé ξ. Ce dernier est donné par la limite de la suite définie par récurrence

$$\begin{cases} x_0 \in X & \text{(17.131a)} \\ x_{n+1} = f(x_n). & \text{(17.131b)} \end{cases}$$

De plus nous pouvons majorer l'erreur par

$$\|x_n - x\| \leqslant \frac{k^n}{1-k}\|x_n - x_{n-1}\| \leqslant \frac{k^n}{1-k}\|x_1 - x_0\|. \tag{17.132}$$

EqKErdim

Soit $r > 0$, $a \in X$ tels que la fonction f laisse la boule $K = \overline{B(a, r)}$ invariante (c'est-à-dire que f se restreint à $f\colon K \to K$). Nous considérons les suites (u_n) et (v_n) définies par

$$\begin{cases} u_0 = v_0 \in K & \text{(17.133a)} \\ u_{n+1} = f(v_n), v_{n+1} \in B(u_n, \epsilon). & \text{(17.133b)} \end{cases}$$

Alors le point fixe ξ de f est dans K et la suite (v_n) satisfait l'estimation

$$\|v_n - \xi\| \leqslant \frac{k^n}{1-k}\|u_1 - u_0\| + \frac{\epsilon}{1-k}. \tag{17.134}$$

La première inégalité (17.132) donne une estimation de l'erreur calculable en cours de processus ; la seconde donne une estimation de l'erreur calculable avant de commencer.

Démonstration. Nous commençons par l'unicité du point fixe. Si a et b sont des points fixes, alors $f(a) = a$ et $f(b) = b$. Par conséquent

$$\|f(a) - f(b)\| = \|a - b\|, \tag{17.135}$$

ce qui contredit le fait que f soit une contraction.

En ce qui concerne l'existence, notons que si la suite des x_n converge dans X, alors la limite est un point fixe. En effet en prenant la limite des deux côtés de l'équation $x_{n+1} = f(x_n)$, nous obtenons $\xi = f(\xi)$, c'est-à-dire que ξ est un point fixe de f. Notons que nous avons utilisé ici la continuité de f, laquelle est une conséquence du fait qu'elle soit Lipschitz. Nous allons donc porter nos efforts à prouver que la suite est de Cauchy (et donc convergente parce que X est complet). Nous commençons par prouver que $\|x_{n+1} - x_n\| \leqslant k^n\|x_0 - x_1\|$. En effet pour tout n nous avons

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leqslant k\|x_n - x_{n-1}\|. \tag{17.136}$$

17. Il me semble qu'à la page 100 de [505], l'hypothèse H1 qui est prouvée ne prouve pas Hn dans le cas $n = 1$. Merci de m'écrire si vous pouvez confirmer ou infirmer. La preuve donnée ici ne contient pas cette « erreur ».

La relation cherchée s'obtient alors par récurrence. Soient $q > p$. En utilisant une somme télescopique,

$$\|x_q - x_p\| \leqslant \sum_{l=p}^{q-1} \|x_{l+1} - x_l\| \tag{17.137a}$$
$$\leqslant \left(\sum_{l=p}^{q-1} k^l \right) \|x_1 - x_0\| \tag{17.137b}$$
$$\leqslant \left(\sum_{l=p}^{\infty} k^l \right) \|x_1 - x_0\|. \tag{17.137c}$$

Étant donné que $k < 1$, la parenthèse est la queue d'une série qui converge, et donc tend vers zéro lorsque p tend vers l'infini.

En ce qui concerne les inégalités (17.132), nous refaisons une somme télescopique :

$$\|x_{n+p} - x_n\| \leqslant \|x_{n+p} - x_{n+p-1}\| + \cdots + \|x_{n+1} - x_n\| \tag{17.138a}$$
$$\leqslant k^p \|x_n - x_{n-1}\| + k^{p-1}\|x_n - x_{n-1}\| + \cdots + k\|x_n - x_{n-1}\| = k(1 + \cdots + k^{p-1})\|x_n - x_{n-1}\| \tag{17.138b}$$
$$\leqslant \frac{k}{1-k}\|x_n - x_{n-1}\|. \tag{17.138c}$$

En prenant la limite $p \to \infty$ nous trouvons

$$\|\xi - x_n\| \leqslant \frac{k}{1-k}\|x_n - x_{n-1}\| \leqslant \frac{k}{1-k}\|x_1 - x_0\|. \tag{17.139}$$ EqIUMVGW

Nous passons maintenant à la seconde partie du théorème en supposant que f se restreigne en une fonction $f\colon K \to K$. D'abord K est encore un espace métrique complet, donc la première partie du théorème s'y applique et f y a un unique point fixe.

Nous allons montrer la relation par récurrence. Tout d'abord pour $n = 1$ nous avons

$$\|v_1 - \xi\| \leqslant \|v_1 - u_1\| + \|u_1 - \xi\| \leqslant \epsilon + \frac{k}{1-k}\|u_1 - u_0\| \tag{17.140}$$

où nous avons utilisé l'estimation (17.139), qui reste valable en remplaçant x_1 par u_1 [18]. Nous pouvons maintenant faire la récurrence :

$$\|v_{n+1} - \xi\| \leqslant \|v_{n+1} - u_{n+1}\| + \|u_{n+1} - \xi\| \tag{17.141a}$$
$$\leqslant \epsilon + k\|v_n - \xi\| \tag{17.141b}$$
$$\leqslant \epsilon + k\left(\frac{k^n}{1-k}\|u_1 - u_0\| + \frac{\epsilon}{1-k} \right) = \frac{\epsilon}{1-k} + \frac{k^{n+1}}{1-k}\|u_1 - u_0\|. \tag{17.141c}$$

□

Remarque 17.35.
Ce théorème comporte deux parties d'intérêts différents. La première partie est un théorème de point fixe usuel, qui sera utilisé pour prouver l'existence de certaines équations différentielles.

La seconde partie est intéressante d'un point de vue numérique. En effet, ce qu'elle nous enseigne est que si à chaque pas de calcul de la récurrence $x_{n+1} = f(x_n)$ nous commettons une erreur d'ordre de grandeur ϵ, alors le procédé (la suite (v_n)) ne converge plus spécialement vers le point fixe, mais tend vers le point fixe avec une erreur majorée par $\epsilon/(1-k)$.

Remarque 17.36.
Au final l'erreur minimale qu'on peut atteindre est de l'ordre de ϵ. Évidemment si on commet une faute de calcul de l'ordre de ϵ à chaque pas, on ne peut pas espérer mieux.

18. Elle n'est cependant pas spécialement valable si on remplace x_n par u_n.

remIOHUJm

Remarque 17.37.
Si f elle-même n'est pas contractante, mais si f^p est contractante pour un certain $p \in \mathbb{N}$ alors la conclusion du théorème de Picard reste valide et f a le même unique point fixe que f^p. En effet nommons x le point fixe de f : $f^p(x) = x$. Nous avons alors

$$f^p\big(f(x)\big) = f\big(f^p(x)\big) = f(x), \tag{17.142}$$

ce qui prouve que $f(x)$ est un point fixe de f^p. Par unicité nous avons alors $f(x) = x$, c'est-à-dire que x est également un point fixe de f.

ThoagJPZJ

Théorème 17.38 (Équation de Fredholm)**.**
Soit $K\colon [a,b]\times[a,b] \to \mathbb{R}$ et $\varphi\colon [a,b] \to \mathbb{R}$, deux fonctions continues. Alors si λ est suffisamment petit, l'équation

$$f(x) = \lambda \int_a^b K(x,y)f(y)dy + \varphi(x) \tag{17.143}$$

admet une unique solution qui sera de plus continue sur $[a,b]$.

Démonstration. Nous considérons l'ensemble $\mathcal{F}$ des fonctions continues $[a,b] \to [a,b]$ muni de la norme uniforme. Le lemme 12.360 implique que $\mathcal{F}$ est complet. Nous considérons l'application $\Phi\colon \mathcal{F} \to \mathcal{F}$ donnée par

$$\Phi(f)(x) = \lambda \int_a^b K(x,y)f(y)dy + \varphi(x). \tag{17.144}$$

Nous montrons que Φ^p est une application contractante pour un certain p. Pour tout $x \in [a,b]$ nous avons

$$\begin{aligned}
\|\Phi(f) - \Phi(g)\|_\infty &\leqslant \|\Phi(f)(x) - \Phi(g)(x)\| && \text{(17.145a)}\\
&= |\lambda|\Big\| \int_a^b K(x,y)\big(f(y) - g(y)\big)dy\Big\| && \text{(17.145b)}\\
&\leqslant |\lambda|\|K\|_\infty|b-a|\|f-g\|_\infty && \text{(17.145c)}
\end{aligned}$$

Nous choisissons λ assez petit pour avoir $|\lambda|\|K\|_\infty|b-a| < 1$. Dans ce cas, l'application Φ est une contraction. Elle possède donc un unique point fixe par le théorème de Picard 17.34. □

17.7 Théorèmes de point fixes et équations différentielles

17.7.1 Théorème de Cauchy-Lipschitz

Nous démontrons ici deux théorèmes de Cauchy-Lipschitz. De nombreuses propriétés annexes seront démontrées dans le chapitre sur les équations différentielles, section 32.8.

Le théorème de Cauchy-Arzella 20.34 sera pour plus tard parce qu'il utilise Schauder 20.33.

ThokUUlgU

Théorème 17.39 (Cauchy-Lipschitz[506, 507])**.**
XtiXON
Nous considérons l'équation différentielle

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & \text{(17.146a)}\\ y(t_0) = y_0 & \text{(17.146b)} \end{cases}$$

avec $f\colon U = I\times\Omega \to \mathbb{R}^n$ où I est ouvert dans $\mathbb{R}$ et Ω ouvert dans $\mathbb{R}^n$. Nous supposons que f est continue sur U et localement Lipschitz[19] *par rapport à y.*

Alors il existe un intervalle $J \subset I$ sur lequel la solution au problème est unique. De plus toute solution du problème est une restriction de cette solution à une partie de J. La solution sur J (dite « solution maximale ») est de classe C^1.

19. Définition 12.319. Notons que nous ne supposons pas que f soit une contraction.

Démonstration. Nous divisions la preuve en plusieurs étapes (même pas toutes simples).

(1) **Cylindre de sécurité**
Précisons l'espace fonctionnel $\mathcal{F}$ adéquat. Soient V et W les voisinages de t_0 et y_0 sur lesquels f est localement Lipschitz. Nous considérons les quantités suivantes :

(1a) $M = \sup_{V\times W} f$;

(1b) $r > 0$ tel que $\overline{B(y_0, r)} \subset V$

(1c) $T > 0$ tel que $\overline{B(t_0, T)} \subset W$ et $T < r/M$.

Nous considérons alors l'ensemble

$$\mathcal{F} = C^0\big(\overline{B(t_0, T)}, \overline{B(y_0, r)}\big) \tag{17.147}$$

que nous munissons de la norme uniforme. Par le lemme 12.360 l'espace $\big(\mathcal{F}, \|.\|_\infty\big)$ est complet.

(2) **Une application $\Phi\colon \mathcal{F} \to \mathcal{F}$**
Si y est une solution de l'équation différentielle considérée, elle vérifie [20]

$$y(t) = y_0 + \int_{t_0}^{t} f\big(u, y(u)\big)du. \tag{17.148}$$

Ceci nous incite à considérer l'opérateur $\Phi\colon \mathcal{F} \to \mathcal{F}$ défini par

$$\Phi(y)(t) = y_0 + \int_{t_0}^{t} f\big(u, y(u)\big)du. \tag{17.149}$$

Pour que l'application Φ soit utile nous devons montrer que pour tout $y \in \mathcal{F}$,

— l'application $\Phi(y)$ est bien définie,

— pour tout $t \in \overline{B(t_0, T)}$ nous avons $\Phi(y)(t) \in \overline{B(y_0, r)}$,

— l'application $\Phi(y)\colon \overline{B(t_0, T)} \to \overline{B(y_0, r)}$ est continue.

Attention : nous ne prétendons pas que Φ elle-même soit continue. C'est parti.

(2a) $\Phi(y)$ **est bien définie**
Il faut montrer que l'intégrale converge. Le calcul de $\Phi(y)(t)$ ne se fait qu'avec $t \in \overline{B(t_0, T)}$. Vu que u prend ses valeurs dans $[t_0, t]$ et que $y \in \mathcal{F}$, le nombre $y(u)$ est toujours dans $\overline{B(y_0, r)}$. Ceci pour dire que dans l'intégrale, la fonction f n'est considérée que sur $[t_0, t] \times \overline{B(y_0, r)} \subset V \times W$. La fonction f est donc uniformément majorable, et l'intégrale ne pose pas de problèmes.

(2b) $\Phi(y)(t) \in \overline{B(y_0, r)}$
Prouvons que $\Phi(y)(t) \in \overline{B(y_0, r)}$. Pour cela, notons que

$$|\Phi(y)(t) - y_0| \leqslant \int_{t_0}^{t} |f\big(u, y(u)\big)|du \leqslant |t - t_0|\|f\|_\infty. \tag{17.150}$$

Étant donné que $t \in \overline{B(t_0, T)}$ nous avons $|t - t_0| \leqslant r/M$ et donc $|\Phi(y)(t) - y_0| \leqslant r$.

(2c) $\Phi(y)$ **est continue**
Nous pourrions invoquer le théorème 17.15, mais nous allons le faire à la main. Soit $s_0 \in B(t_0, T)$ et prouvons que $\Phi(y)$ est continue en s_0. Pour cela nous prenons $s \in B(s_0, \delta)$ et nous calculons :

$$|\Phi(y)(s) - \Phi(y)(s_0)| \leqslant \int_{s_0}^{s} |f\big(u, y(u)\big)|du \leqslant |s_0 - s|\|f\|_\infty. \tag{17.151}$$

C'est le fait que f soit bornée dans le cylindre de sécurité qui fait en sorte que cela tende vers zéro lorsque $s \to s_0$.

20. C'est le théorème fondamental du calcul intégral 14.287.

L'équation (17.148) signifie que y est un point fixe de Φ. L'espace $\mathcal{F}$ étant complet, le théorème de point fixe de Picard (théorème 17.34) s'applique. Nous allons montrer qu'il existe un $p \in \mathbb{N}$ tel que Φ^p soit contractante. Par conséquent Φ^p aura un unique point fixe qui sera également unique point fixe de Φ par la remarque 17.37.

(3) **Contractante**

Prouvons donc que Φ^p est contractante pour un certain p. Pour cela nous commençons par montrer la formule suivante par récurrence :

$$\|\Phi^p(x)(t) - \Phi^p(y)(t)\| \leqslant \frac{k^p|t-t_0|^p}{p!}\|x-y\|_\infty \tag{17.152}$$

pour tout $x, y \in \mathcal{F}$, et pour tout $t \in \overline{B(t_0,T)}$. Pour $p = 0$ la formule (17.152) est vérifiée parce que $\|x-y\|_\infty$ est le supremum de $\|x(t)-y(t)\|$ pour $t \in \overline{B(t_0,T)}$. Supposons que la formule soit vraie pour p et calculons pour $p+1$. Pour tout $t \in \overline{B(t_0,T)}$ nous avons

$$\begin{aligned}
\|\Phi^{p+1}(x)(t) - \Phi^{p+1}(y)(t)\| &\leqslant \left|\int_{t_0}^t \|f\big(u,\Phi^p(x)(u)\big) - f\big(u,\Phi^p(y)(u)\big)\|du\right| && (17.153a)\\
&\leqslant \left|\int_{t_0}^t k\|\Phi^p(x)(u) - \Phi^p(y)(u)\|du\right| && (17.153b)\\
&\leqslant \left|\int_{t_0}^t k\frac{k^p|t-t_0|}{p!}\|x-y\|_\infty\right| && (17.153c)\\
&= \frac{k^{p+1}|t-t_0|^{p+1}}{(p+1)!}\|x-y\|_\infty. && (17.153d)
\end{aligned}$$

Justifications :

— (17.153b) parce que f est Lipschitz.

— (17.153c) par hypothèse de récurrence.

La formule (17.152) est maintenant établie. Nous pouvons maintenant montrer que Φ^p est une contraction pour un certain p. Pour tout $t \in \overline{B(t_0,T)}$ nous avons

$$\|\Phi^p(x)(t) - \Phi^p(y)(t)\| \leqslant \frac{k^p}{p!}|t-t_0|^p\|x-y\|_\infty \leqslant \frac{k^pT^p}{p!}\|x-y\|_\infty \tag{17.154}$$

où nous avons utilisé le fait que $|t-t_0|^p < T^p$. En prenant le supremum sur t des deux côtés il vient

$$\|\Phi^p(x) - \Phi^p(y)\|_\infty \leqslant \frac{k^pT^p}{p!}\|x-y\|_\infty. \tag{17.155}$$

Le membre de droite tend vers zéro lorsque $p \to \infty$ parce que $k^pT^p/p! \to 0$ [21]. Nous concluons donc que Φ^p est une contraction pour un certain p.

(4) **Conclusion**

L'unique point fixe de Φ est alors l'unique solution continue de l'équation différentielle (17.146). Par ailleurs l'équation elle-même $y' = f(t,y)$ demande implicitement que y soit dérivable et donc continue. Nous concluons que l'unique point fixe de Φ est l'unique solution de l'équation différentielle donnée. Cette dernière est automatiquement C^1 parce que si y est continue alors $u \mapsto f(u, y(u))$ est continue, c'est-à-dire que y' est continue.

(5) **Unicité**

Nous passons maintenant à la partie « prolongement maximum » du théorème. Soient x_1 et x_2 deux solutions maximales du problème (17.146) sur des intervalles I_1 et I_2 respectivement. Les intervalles I_1 et I_2 contiennent $\overline{B(t_0,T)}$ sur lequel $x_1 = x_2$ par unicité.

Nous allons maintenant montrer que pour tout $t \geqslant t_0$ pour lequel x_1 ou x_2 est défini, $x_1(t)$ et $x_2(t)$ sont définis et sont égaux. Le raisonnement sur $t \leqslant t_0$ est similaire.

21. Parce que les factorielles vont plus vite que les puissances.

Supposons que l'ensemble des $t \geqslant t_0$ tels que $x_1 = x_2$ soit ouvert à droite, c'est-à-dire soit de la forme $[t_0, b[$. Dans ce cas, soit x_1, soit x_2 (soit les deux) cesse d'exister en b. En effet si nous avions les fonctions x_i sur $[t_0, b+\epsilon[$ alors l'équation $x_1 = x_2$ définirait un fermé dans $[t_0, b+\epsilon[$. Supposons pour fixer les idées que x_1 cesse d'exister : le domaine de x_1 (parmi les $t \geqslant 0$) est $[t_0, b[$ et sur ce domaine nous avons $x_1 = x_2$. Dans ce cas x_1 pourrait être prolongé en x_2 au-delà de b. Si x_1 et x_2 s'arrêtent d'exister en même temps en b, alors nous avons bien $x_1 = x_2$.

Nous devons donc traiter le cas où $x_1 = x_2$ sur $[t_0, b]$ alors que x_1 et x_2 existent sur $[t_0, b+\epsilon[$ pour un certain ϵ.

Nous pouvons appliquer le théorème d'existence locale au problème

$$\begin{cases} y' = f(t,y) & \text{(17.156a)} \\ y(b) = x_1(b). & \text{(17.156b)} \end{cases}$$

Il existe un voisinage de b sur lequel la solution est unique. Sur ce voisinage nous devons donc avoir $x_1 = x_2$, ce qui contredit le fait que $x_1 \neq x_2$ en dehors de $[t_0, b]$.

Donc x_1 et x_2 existent et sont égaux sur, au moins $I_1 \cup I_2$.

□

Le théorème de Cauchy-Lipschitz donne existence et unicité d'une solution maximale. Cependant cette solution peut ne pas exister partout où les hypothèses sur f sont remplies. En d'autres termes, il peut arriver que f soit Lipschitz jusqu'à t_1, mais que la solution maximale ne soit définie que jusqu'en $t_2 < t_1$. Ce cas fait l'objet du théorème d'explosion en temps fini 32.20.

Sous quelques hypothèses, nous pouvons nous assurer de l'existence d'une solution unique sur tout $\mathbb{R}$.

Ce théorème de Cauchy-Lipschitz global est utilisé pour faire le lien entre les représentations des algèbres de Lie et celles du groupe, voir la proposition 53.94.

THOooZIVRooPSWMxg

Théorème 17.40 (Cauchy-Lipschitz global[508, 104])**.**
Soit un intervalle I de $\mathbb{R}$, $y_0 \in \mathbb{R}^n$, $t_0 \in I$ et une fonction continue $f\colon I \times \mathbb{R}^n \to \mathbb{R}^n$ telle que pour tout compact K dans I, il existe $k > 0$ tel que

$$\|f(t,y_1) - f(t,y_2)\| \leqslant k\|y_1 - y_2\| \tag{17.157}$$

pour tout $t \in K$ et $y_1, y_2 \in \mathbb{R}^n$.

Alors le problème EQSooBNREooUTfbMH

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & \text{(17.158a)} \\ y(t_0) = y_0 & \text{(17.158b)} \end{cases}$$

possède une unique solution $y\colon I \to \mathbb{R}^n$ sur I.

Démonstration. Soit un intervalle compact K dans I et contenant t_0. Nous notons ℓ le diamètre de K. Sur l'espace $E = C^0(K, \mathbb{R}^n)$ nous considérons la topologie uniforme : $(E, \|.\|_\infty)$. C'est un espace complet par le lemme 12.360 (nous utilisons le fait que $\mathbb{R}^n$ soit complet, proposition 1.406). Nous allons utiliser l'application suivante :

$$\begin{aligned} \Phi\colon E &\to E \\ \Phi(y)(t) &= y_0 + \int_{t_0}^{t} f\big(s, y(s)\big)ds \end{aligned} \tag{17.159}$$ EQooJUTBooILBKoE

Démontrons quelques faits à propos de Φ.

(1) **Φ est bien définie** Nous devons commencer par prouver que cette application est bien définie. Si $y \in E$ alors f et y sont continues ; l'application $s \mapsto f\big(s, y(s)\big)$ est donc également

continue. L'intégrale de cette fonction sur le compact $[t_0, t]$ ne pose alors pas de problèmes. En ce qui concerne la continuité de $\Phi(y)$ sous l'hypothèse que y soit continue,

$$\|\Phi(y)(t) - \Phi(y)(t')\| \leqslant \int_t^{t'} \|f(s, y(s))\| ds \leqslant M|t - t'| \tag{17.160}$$

où M est une majoration de $\|s \mapsto f\big(s, y(s)\big)\|_{\infty,K}$.

(2) **Si y est solution alors $\Phi(y) = y$**
Supposons que y soit une solution de l'équation différentielle (17.158). Alors, vu que $y'(t) = f\big(t, y(t)\big)$ nous avons :

$$y(t) = y_0 + \int_{t_0}^t y'(s) ds = y_0 + \int_{t_0}^t f\big(s, y(s)\big) ds = \Phi(y)(t). \tag{17.161}$$

(3) **Si $\Phi(y) = y$ alors y est solution**
Nous avons, pour tout t :

$$y(t) = y_0 + \int_{t_0}^t f\big(s, y(s)\big) ds. \tag{17.162}$$

Le membre de droite est dérivable par rapport à t, et la dérivée est $f\big(t, y(t)\big)$. Donc le membre de gauche est également dérivable et nous avons bien

$$y'(t) = f\big(t, y(t)\big). \tag{17.163}$$

De plus $y(t_0) = y_0 + \int_{t_0}^{t_0} \ldots = y_0$.

Nous sommes encore avec K compact et $E = C^0(K, \mathbb{R}^n)$ muni de la norme uniforme. Nous allons montrer que Φ est une contraction de E pour une norme bien choisie.

(1) **Une norme sur E** Pour $y \in E$ nous posons

$$\|y\|_k = \max_{t \in K} \big(e^{-k|t-t_0|} \|y(t)\|\big). \tag{17.164}$$

Ce maximum est bien défini et fini, parce que dedans, la fonction de t est une fonction continue sur le compact K. C'est également une norme parce que si $\|y\|_k = 0$ alors $e^{-k|t-t_0|}\|y(t)\| = 0$ pour tout t. Étant donné que l'exponentielle ne s'annule pas, $\|y(t)\| = 0$ pour tout t.

(2) **Équivalence de norme** Nous montrons que les normes $\|.\|_k$ et $\|.\|_\infty$ sont équivalentes [22] :

$$\|y\|_\infty e^{-k\ell} \leqslant \|y\|_k \leqslant \|y\|_\infty \tag{17.165}$$

pour tout $y \in E$. Pour la première inégalité, $\ell \geqslant |t - t_0|$ pour tout $t \in K$, et $k > 0$, donc

$$\|y(t)\| e^{-k\ell} \leqslant e^{-k|t-t_0|} \|y(t)\|. \tag{17.166}$$

En prenant le maximum des deux côtés, $\|y\|_\infty e^{-k\ell} \leqslant \|y\|_k$.
En ce qui concerne la seconde inégalité dans (17.165), $k|t - t_0| \geqslant 0$ et donc $e^{-k|t-t_0|} < 1$.

Puisque les normes $\|.\|_\infty$ et $\|.\|_k$ sont équivalentes, l'espace $(E, \|.\|_k)$ est tout autant complet que $(E, \|.\|_\infty)$. Nous démontrons à présent que Φ est une contraction dans $(E, \|\|_k)$.
Soient $y, z \in E$. Si $t \geqslant t_0$ nous avons

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \int_{t_0}^t \|f\big(s, y(s)\big) - f\big(s, z(s)\big)\| ds \tag{17.167a}$$

$$\leqslant k \int_{t_0}^t \|y(s) - z(s)\| ds. \tag{17.167b}$$

22. Définition 10.46

Il convient maintenant de remarquer que

$$\|y(t)\| = e^{-k|t-t_0|}e^{k|t-t_0|}\|y(t)\| \leqslant \|y\|_k e^{k|t-t_0|}. \tag{17.168}$$

Nous pouvons avec ça prolonger les inégalités (17.167) par

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant k\|y-z\|_k \int_{t_0}^{t} e^{k|s-t_0|}ds = k\|y-z\|_k \int_{t_0}^{t} e^{k(s-t_0)}ds \tag{17.169}$$

où nous avons utilisé notre supposition $t \geqslant t_0$ pour éliminer les valeurs absolues. L'intégrale peut être calculée explicitement, mais nous en sommes arrivés à un niveau de fainéantise tellement inconcevable que

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: var('a,b,k')
7 (a, b, k)
8 sage: f(x)=exp(-k*x)
9 sage: f.integrate(x,a,b)
10 e^(-a*k)/k - e^(-b*k)/k
```

tex/sage/sageSnip014.sage

Au final, si $t \geqslant t_0$,

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k\big(e^{k(t-t_0)} - 1\big). \tag{17.170}$$

Si $t \leqslant t_0$, il faut retourner les bornes de l'intégrale avant d'y faire rentrer la norme parce que $\|\int_0^1 f\| \leqslant \int_0^1 \|f\|$, mais ça ne marche pas avec $\|\int_1^0 f\|$. Pour $t \leqslant t_0$ tout le calcul donne

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k\big(e^{k(t_0-t)} - 1\big). \tag{17.171}$$

Les deux inéquations sont valables a fortiori en mettant des valeurs absolues dans l'exponentielle, de telle sorte que pour tout $t \in K$ nous avons

$$e^{-k|t_0-t|}\|\phi(y)(t) - \Phi(z)(t)\| \leqslant \|y-z\|_k\big(1 - e^{-k|t_0-t|}\big). \tag{17.172}$$

En prenant le supremum sur t,

$$\|\Phi(y) - \Phi(z)\|_k \leqslant \|y-z\|_k(1 - e^{-k\ell}), \tag{17.173}$$

mais $0 < (1-e^{e-k\ell}) < 1$, donc Φ est contractante pour la norme $\|.\|_k$. Comme $(E, \|.\|_k)$ est complet, l'application Φ y a un unique point fixe par le théorème de Picard 17.34.

Ce point fixe est donc l'unique solution de l'équation différentielle de départ.

(1) **Existence et unicité sur I** Il nous reste à prouver que la solution que nous avons trouvée existe sur I : jusqu'à présent nous avons démontré l'existence et l'unicité sur n'importe quel compact dans I.

Soit une suite croissante de compacts K_n contenant t_0 (par exemple une suite exhaustive comme celle du lemme 7.306). Nous avons en particulier

$$I = \bigcup_{n=0}^{\infty} K_n. \tag{17.174}$$

(2) **Existence sur I** Soit y_n l'unique solution sur K_n. Il suffit de poser

$$y(t) = y_n(t) \tag{17.175}$$

pour n tel que $t \in K_n$. Cette définition fonctionne parce que si $t \in K_n \cap K_m$, il y a forcément un des deux qui est inclus dans l'autre et le résultat d'unicité sur le plus grand des deux donne $y_n(t) = y_m(t)$.

(3) **Unicité sur I**
Soient y et z des solutions sur I ; puisque I n'est pas spécialement compact, le travail fait plus haut ne permet pas de conclure que $y = z$.
Soit $t \in I$. Alors $t \in K_n$ pour un certain n et y et z sont des solutions sur K_n qui est compact. L'unicité sur K_n donne $y(t) = z(t)$.

□

17.41.
Il y a d'autres moyens de prouver qu'une solution existe globalement sur $\mathbb{R}$. Si f est globalement bornée, le théorème d'explosion en temps fini donne quelques garanties, voir 32.22.

Le théorème suivant donne une version du théorème de Cauchy-Lipschitz lorsque la fonction f dépend d'un paramètre. Ce théorème n'utilise rien de fondamentalement nouveau. Nous le donnons seulement pour montrer que l'on peut choisir l'espace $\mathcal{F}$ de façon un peu maligne pour élargir le résultat. Si vous voulez un théorème de Cauchy-Lipschitz avec paramètre vraiment intéressant, allez voir le théorème 32.37.

THOooDTCWooSPKeYu

Théorème 17.42 (Cauchy-Lipschitz avec paramètre[1, 509])**.**
Soit un intervalle ouvert I de $\mathbb{R}$, un connexe ouvert Ω de $\mathbb{R}^n$ et un intervalle ouvert Λ de $\mathbb{R}^d$. Soit une fonction $f\colon I \times \Omega \times \Lambda \to \mathbb{R}^n$ continue et localement Lipschitz en Ω. Soient $t_0 \in I$, $y_0 \in \Omega$ et $\lambda_0 \in \Lambda$. Il existe un voisinage compact de (t_0, y_0, λ_0) sur lequel le problème

$$\begin{cases} y'_\lambda(t) = f\big(t, y_\lambda(t), \lambda\big) & \text{(17.176a)} \\ y_\lambda(t_0) = y_0 & \text{(17.176b)} \end{cases}$$

possède une unique solution. De plus $(t, \lambda) \mapsto y_\lambda(t)$ est continue[23].

Idée rapide de la preuve.

¡¡ Avertissement/question au lecteur !! 17.43
Ceci est une idée de la preuve. Je n'ai pas vérifié toutes les étapes. Soyez prudent.

D'abord nous avons un voisinage compact $V \times \overline{B(y_0, r)} \times \Lambda_0$ de (t_0, y_0, λ_0) sur lequel f est bornée. Ensuite nous récrivons l'équation différentielle sous la forme

$$\begin{cases} \dfrac{\partial y}{\partial t}(t, \lambda) = f\big(t, y(t, \lambda), \lambda\big) & \text{(17.177a)} \\ y(t_0, \lambda) = y_0. & \text{(17.177b)} \end{cases}$$

pour une fonction $y\colon V \times \Lambda_0 \to \mathbb{R}^n$.

Nous posons $\mathcal{F} = C^0\big(V \times \Lambda_0, \mathbb{R}^n\big)$ et nous y définissons l'application

$$\begin{aligned} \Phi\colon \mathcal{F} &\to \mathcal{F} \\ \Phi(y)(t, \lambda) &= y_0 + \int_{t_0}^{t} f\big(s, y(s, \lambda), \lambda\big) ds. \end{aligned} \tag{17.178}$$

Il y a plein de vérifications à faire[509], mais je parie que Φ est bien définie, et qu'une de ses puissances est une contraction de $(\mathcal{F}, \|.\|_\infty)$. L'unique point fixe est une solution de notre problème et est dans C^0, donc $(t, \lambda) \mapsto y(t, \lambda) = y_\lambda(t)$ est de classe C^0, c'est-à-dire continue. □

23. Ici, la surprise est que ce soit continu par rapport à λ. Le fait qu'elle le soit par rapport à t est clair depuis le départ, parce que ce n'est finalement rien d'autre que le Cauchy-Lipschitz vieux et connu.

17.44.
Ce théorème marque un peu la limite de ce que l'on peut faire avec la méthode des points fixes dans le cadre de Cauchy-Lipschitz : nous sommes limités à la continuité de la solution parce que les espaces C^p ne sont pas complets [24]. Il n'y a donc pas d'espoir d'adapter la méthode pour prouver que si f est de classe C^p alors $(t,\lambda) \mapsto y_\lambda(t)$ est de classe C^p. On peut, à λ fixé, prouver que $t \mapsto y_\lambda(t)$ est de classe C^p (utiliser une récurrence), mais pas plus.

La régularité C^1 de y par rapport à la condition initiale sera l'objet du théorème 32.33. Ce résultat n'est vraiment pas facile et utilise des ingrédients bien autres qu'un point fixe. Ensuite la régularité C^p par rapport à la condition initiale et par rapport à un paramètre seront presque des cadeaux (proposition 32.35 et 32.37).

EXooJXIGooQtotMc

Exemple 17.45 ([510])**.**
Nous savons que le théorème de Picard permet de trouver le point fixe par itération de la contraction à partir d'un point quelconque. Tentons donc de résoudre

$$\begin{cases} y'(t) = y(t) & \text{(17.179a)} \\ y(0) = 1 & \text{(17.179b)} \end{cases}$$

dont nous savons depuis l'enfance que la solution est l'exponentielle [25]. Partons donc de la fonction constante $y_0 = 1$, et appliquons la contraction (17.159) :

$$u_1 = 1 + \int_0^1 u_0(s)ds = 1 + t. \tag{17.180}$$

Ensuite

$$u_2 = 1 + \int_0^t (1+s)ds = 1 + t + \frac{t^2}{2}. \tag{17.181}$$

Et on voit que les itérations suivantes vont donner l'exponentielle.

Nous sommes évidemment en droit de se dire que nous avons choisi un bon point de départ. Tentons le coup avec une fonction qui n'a rien à voir avec l'exponentielle : $u_0(x) = \sin(x)$.

Le programme suivant permet de faire de belles investigations numériques en partant d'à peu près n'importe quelle fonction :

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

x=var('x')

def Phi(f):
    prim=f.integrate()
    return 1+prim(x)-prim(0)

f=sin(x)

for i in range(1,30):
    print(i,f)
    f=Phi(f)

g=f(x)-exp(x)
```

24. Par exemple, le théorème de Stone-Weierstrass 12.420 nous dit que la limite uniforme de polynômes (de classe C^∞) peut n'être que continue. Voir aussi le thème 37.
25. Voir par exemple le théorème 15.76.

```
19 plot(g,(x,-10,10)).show()
```

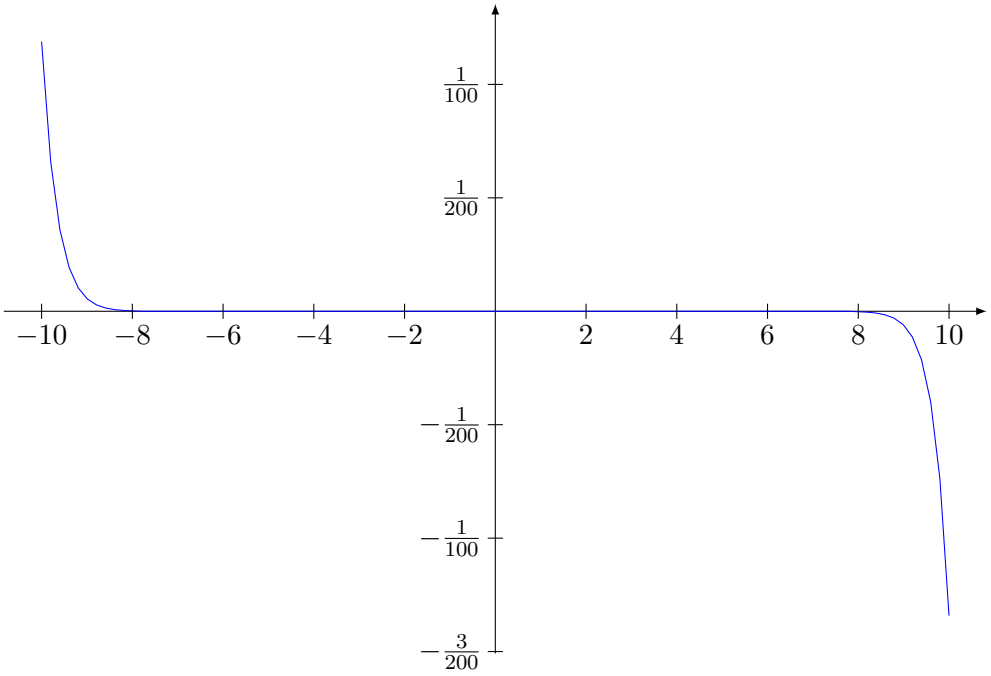
tex/sage/picard_exp.py

Ce programme fait 30 itérations depuis la fonction $\sin(x)$ pour tenter d'approximer $\exp(x)$. Pour donner une idée, après 7 itérations nous avons la fonction suivante :

$$\frac{1}{60}x^5 + \frac{1}{24}x^4 + \frac{1}{2}x^2 + 2x - \sin(x) + 1. \tag{17.182}$$

Nous voyons que les coefficients sont des factorielles, mais pas toujours celles correspondantes à la puissance, et qu'il manque certains termes par rapport au développement de l'exponentielle que nous connaissons. Bref, le polynôme qui se met en face de $\sin(x)$ s'adapte tout seul pour compenser.

Et après 30 itérations, ça donne quoi ? Voici un graphe de l'erreur entre $u_{30}(x)$ et $\exp(30)$:



Pour donner une idée, $\exp(10) \simeq 22000$. Donc il y a une faute de 0.01 sur 22000. Pas mal.

△

17.8 Théorèmes d'inversion locale et de la fonction implicite

17.8.1 Mise en situation

Dans un certain nombre de situation, il n'est pas possible de trouver des solutions explicites aux équations qui apparaissent. Néanmoins, l'existence « théorique » d'une telle solution est souvent déjà suffisante. C'est l'objet du théorème de la fonction implicite.

Prenons par exemple la fonction sur $\mathbb{R}^2$ donnée par

$$F(x, y) = x^2 + y^2 - 1. \tag{17.183}$$

Nous pouvons bien entendu regarder l'ensemble des points donnés par $F(x, y) = 0$. C'est le cercle dessiné à la figure 17.1.

Nous ne pouvons pas donner le cercle sous la forme $y = y(x)$ à cause du $\pm$ qui arrive quand on prend la racine carrée. Mais si on se donne le point P, nous pouvons dire que *autour de* P, le cercle est la fonction

$$y(x) = \sqrt{1 - x^2}. \tag{17.184}$$

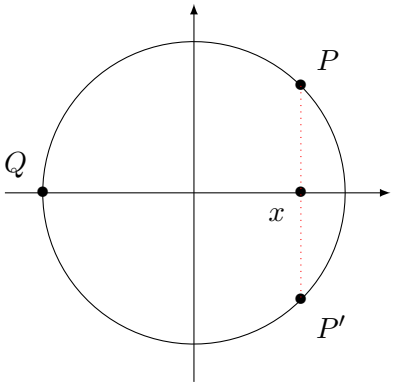


FIGURE 17.1: Un cercle pour montrer l'intérêt de la fonction implicite. Si on donne x, nous ne pouvons pas savoir si nous parlons de P ou de P'. LabelFigCercleImplicite

Tandis que autour du point P', le cercle est la fonction

$$y(x) = -\sqrt{1-x^2}. \tag{17.185}$$

Autour de ces deux points, donc, le cercle est donné par une fonction. Il n'est par contre pas possible de donner le cercle autour du point Q sous la forme d'une fonction.

Ce que nous voulons faire, en général, est de voir si l'ensemble des points tels que

$$F(x_1, \dots, x_n, y) = 0 \tag{17.186}$$

peut être donné par une fonction $y = y(x_1, \dots, x_n)$. En d'autre termes, est-ce qu'il existe une fonction $y(x_1, \dots, x_n)$ telle que

$$F\big(x_1, \dots, x_n, y(x_1, \dots, x_n)\big) = 0. \tag{17.187}$$

Plus généralement, soit une fonction

$$\begin{aligned} F\colon D \subset \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^m \\ (x, y) &\mapsto \big(F_1(x, y), \dots, F_m(x, y)\big) \end{aligned} \tag{17.188}$$

avec $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_m)$. Pour chaque x fixé, on s'intéresse aux solutions du système de m équations $F(x, y) = 0$ pour les inconnues y ; en particulier, on voudrait pouvoir écrire $y = \varphi(x)$ vérifiant $F(x, \varphi(x)) = 0$.

17.8.2 Théorème d'inversion locale

LemGZoqknC

Lemme 17.46 ([353]).
Soit E un espace de Banach (métrique complet) et $\mathcal{O}$ un ouvert de E. Nous considérons une λ-contraction $\varphi\colon \mathcal{O} \to E$. Alors l'application

$$f\colon x \mapsto x + \varphi(x) \tag{17.189}$$

est un homéomorphisme entre $\mathcal{O}$ et un ouvert de E. De plus f^{-1} est Lipschitz de constante plus petite ou égale à $(1-\lambda)^{-1}$.

Cette proposition utilise le théorème de point fixe de Picard 17.34, et sera utilisée pour démontrer le théorème d'inversion locale 17.47.

Démonstration. Soient $x_1, x_2 \in \mathcal{O}$. Nous posons $y_1 = f(x_1)$ et $y_2 = f(x_2)$. En vertu de l'inégalité de la proposition 7.159 nous avons subEqEBJsBfz

$$\|f(x_2) - f(x_1)\| = \|x_2 + \varphi(x_2) - x_1 - \varphi(x_1)\| \tag{17.190a}$$

$$\geqslant \Big|\|x_2 - x_1\| - \|\varphi(x_2) - \varphi(x_1)\|\Big| \tag{17.190b}$$

$$\geqslant (1-\lambda)\|x_2 - x_1\|. \tag{17.190c}$$

À la dernière ligne les valeurs absolues sont enlevées parce que nous savons que ce qui est à l'intérieur est positif. Cela nous dit d'abord que f est injective parce que $f(x_2) = f(x_1)$ implique $x_2 = x_1$. Donc f est inversible sur son image. Nous posons $A = f(\mathcal{O})$ et nous devons prouver que que $f^{-1}\colon A \to \mathcal{O}$ est continue, Lipschitz de constante majorée par $(1-\lambda)^{-1}$ et que A est ouvert.

Les inéquations (17.190) nous disent que

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| \leqslant \frac{\|y_1 - y_2\|}{1-\lambda}, \tag{17.191}$$

c'est-à-dire que

$$f^{-1}\big(B(y,r)\big) \subset B\big(f^{-1}(y), \frac{r}{1-\lambda}\big), \tag{17.192}$$

ce qui signifie que f^{-1} est Lipschitz de constante souhaitée et donc continue.

Il reste à prouver que $f(\mathcal{O})$ est ouvert. Pour cela nous prenons $y_0 = f(x_0)$ dans $f(\mathcal{O})$ est nous prouvons qu'il existe ϵ tel que $B(y_0, \epsilon)$ soit dans $f(\mathcal{O})$. Il faut donc que pour tout $y \in B(y_0, \epsilon)$, l'équation $f(x) = y$ ait une solution. Nous considérons l'application

$$L_y\colon x \mapsto y - \varphi(x). \tag{17.193}$$

Ce que nous cherchons est un point fixe de L_y parce que si $L_y(x) = x$ alors $y = x + \varphi(x) = f(x)$. Puisque

$$\|L_y(x) - L_y(x')\| = \|\varphi(x) - \varphi(x')\| \leqslant \lambda\|x - x'\|, \tag{17.194}$$

l'application L_y est une contraction de constante λ. Par ailleurs x_0 est un point fixe de L_{y_0}, donc en vertu de la caractérisation (12.856) des fonctions Lipschitziennes,

$$L_{y_0}\big(\overline{B(x_0,\delta)}\big) \subset \overline{B\big(L_{y_0}(x_0), \lambda\delta\big)} = \overline{B(x_0, \lambda\delta)}. \tag{17.195}$$

Comme pour tout y et x nous avons $L_y(x) = L_{y_0}(x) + y - y_0$,

$$L_y\big(\overline{B(x_0,\delta)}\big) = L_{y_0}\big(\overline{B(x_0,\delta)}\big) + (y - y_0) \subset \overline{B(x_0,\lambda\delta)} + (y - y_0) \subset \overline{B(x_0, \lambda\delta) + \|y - y_0\|}. \tag{17.196}$$

Si $\epsilon < (1-\lambda)\delta$ alors $\lambda\delta + \|y - y_0\| < \delta$. Un tel choix de $\epsilon > 0$ est possible parce que $\lambda < 1$. Pour une telle valeur de ϵ nous avons

$$L_y\big(\overline{B(x_0,\delta)}\big) \subset \overline{B(x_0,\delta)}. \tag{17.197}$$

Par conséquent L_y est une contraction sur l'espace métrique complet $\overline{B(x_0,\delta)}$, ce qui signifie que L_y y possède un point fixe par le théorème de Picard 17.34. □

Nous allons le démontrer dans le cas un peu plus général (mais pas plus cher [26]) des espaces de Banach en tant que conséquence du théorème de point fixe de Picard 17.34.

ThoXWpzqCn

Théorème 17.47 (Inversion locale dans un espace de Banach[511, 353])**.**
Soit une fonction $f \in C^p(E,F)$ avec $p \geqslant 1$ entre deux espaces de Banach. Soit $x_0 \in E$ tel que df_{x_0} soit une bijection bicontinue [27]. Alors il existe un voisinage ouvert V de x_0 et W de $f(x_0)$ tels que

(1) $f\colon V \to W$ soit une bijection,

(2) $f^{-1}\colon W \to V$ soit de classe C^p.

Démonstration. Nous commençons par simplifier un peu le problème. Pour cela, nous considérons la translation $T\colon x \mapsto x + x_0$ et l'application linéaire

$$\begin{aligned} L\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto (df_{x_0})^{-1}x \end{aligned} \tag{17.198}$$

26. Sauf la justification de la régularité de l'application $A \mapsto A^{-1}$

27. En dimension finie, une application linéaire est toujours continue et d'inverse continu.

qui sont tous deux des difféomorphismes (L en est un par hypothèse d'inversibilité). Quitte à travailler avec la fonction $k = L \circ f \circ T$, nous pouvons supposer que $x_0 = 0$ et que $df_{x_0} = \mathbb{1}$. Pour comprendre cela il faut utiliser deux fois la formule de différentielle de fonction composée de la proposition 12.724 :

$$dk_0(u) = dL_{(f\circ T)(0)}\Big(df_{T(0)}dT_0(u)\Big). \tag{17.199}$$

Puisque L est linéaire, sa différentielle est elle-même, c'est-à-dire $dL_{(f\circ T)(0)} = (df_{x_0})^{-1}$, et par ailleurs $dT_0 = \mathbb{1}$, donc

$$dk_0(u) = (df_{x_0})^{-1}\Big(df_{x_0}(u)\Big) = u, \tag{17.200}$$

ce qui signifie bien que $dk_0 = \mathbb{1}$. Pour tout cela nous avons utilisé en plein le fait que df_{x_0} était inversible.

Nous posons $g = f - \mathbb{1}$, c'est-à-dire $g(x) = f(x) - x$, qui a la propriété $dg_0 = 0$. Étant donné que g est de classe C^1, l'application [28]

$$\begin{aligned} dg\colon E &\to \mathrm{GL}(F) \\ x &\mapsto dg_x \end{aligned} \tag{17.201}$$

est continue. En conséquence, nous avons un voisinage U' de 0 pour lequel

$$\sup_{x\in U'} \|dg_x\| < \frac{1}{2}. \tag{17.202}$$ EqSGTOfyx

Maintenant le théorème des accroissements finis 10.292 (12.314 pour la dimension finie) nous indique que pour tout $x, x' \in U'$ nous avons [29]

$$\|g(x') - g(x)\| \leqslant \sup_{a\in[x,x']} \|dg_a\| \cdot \|x - x'\| \leqslant \frac{1}{2}\|x - x'\|, \tag{17.203}$$

ce qui prouve que g est une contraction au moins sur l'ouvert U'. Nous allons aussi donner une idée de la façon dont f fonctionne : si $x_1, x_2 \in U'$ alors

$$\|x_1 - x_2\| = \|g(x_1) - f(x_1) - g(x_2) + f(x_2)\| \tag{17.204a}$$
$$\leqslant \|g(x_1) - g(x_2)\| + \|f(x_1) - f(x_2)\| \tag{17.204b}$$
$$\leqslant \frac{1}{2}\|x_1 - x_2\| + \|f(x_1) - f(x_2)\|, \tag{17.204c}$$

ce qui montre que

$$\|x_1 - x_2\| \leqslant 2\|f(x_1) - f(x_2)\|. \tag{17.205}$$

Maintenant que nous savons que g est contractante de constante $\frac{1}{2}$ et que $f = g + \mathbb{1}$ nous pouvons utiliser la proposition 17.46 pour conclure que f est un homéomorphisme sur un ouvert U (partie de U') de E et f^{-1} a une constante de Lipschitz plus petite ou égale à $(1 - \frac{1}{2})^{-1} = 2$.

Nous allons maintenant prouver que f^{-1} est différentiable et que sa différentielle est donnée par $(df^{-1})_{f(x)} = (df_x)^{-1}$.

Soient $a, b \in U$ et $u = b - a$. Étant donné que f est différentiable en a, il existe une fonction $\alpha \in o(\|u\|)$ telle que

$$f(b) - f(a) - df_a(u) = \alpha(u). \tag{17.206}$$

En notant $y_a = f(a)$ et $y_b = f(b)$ et en appliquant $(df_a)^{-1}$ à cette dernière équation,

$$(df_a)^{-1}(y_b - y_a) - u = (df_a)^{-1}\big(\alpha(u)\big). \tag{17.207}$$

28. Ici $\mathrm{GL}(F)$ est l'ensemble des applications linéaires, inversibles et continues de F dans lui-même. Ce ne sont pas spécialement des matrices parce que nous n'avons pas d'hypothèses sur la dimension de F, finie ou non.

29. Ici nous supposons avoir choisi U' convexe afin que tous les $a \in [x, x']$ soient bien dans U' et donc soumis à l'inéquation (17.202), ce qui est toujours possible, il suffit de prendre une boule.

Puisque df_a est bornée (et son inverse aussi), le membre de droite est encore une fonction β ayant la propriété $\lim_{u\to 0}\beta(u)/\|u\|=0$; en réordonnant les termes,

$$b-a=(df_a)^{-1}(y_b-y_a)+\beta(u) \tag{17.208}$$

et donc

$$f^{-1}(y_b)-f^{-1}(y_a)-(df_a)^{-1}(y_b-y_a)=\beta(u), \tag{17.209}$$

ce qui prouve que f^{-1} est différentiable et que $(df^{-1})_{y_a}=(df_a)^{-1}$.

La différentielle df^{-1} est donc obtenue par la chaine

$$df^{-1}\colon f(U)\xrightarrow{f^{-1}}U'\xrightarrow{df}\mathrm{GL}(F)\xrightarrow{\mathrm{Inv}}\mathrm{GL}(F) \tag{17.210}$$

où l'application $\mathrm{Inv}\colon \mathrm{GL}(F)\to\mathrm{GL}(F)$ est l'application $X\mapsto X^{-1}$ qui est de classe C^∞ par le théorème 10.300. D'autre part, par hypothèse df est une application de classe C^{k-1} et donc au minimum C^0 parce que $k\geqslant 1$. Enfin, l'application $f^{-1}\colon f(U)\to U$ est continue (parce que la proposition 17.46 précise que f est un homéomorphisme). Donc toute la chaine est continue et df^{-1} est continue. Cela entraine immédiatement que f^{-1} est C^1 et donc que toute la chaine est C^1.

Par récurrence nous obtenons la chaine

$$df^{-1}\colon f(U)\xrightarrow[C^{k-1}]{f^{-1}}U'\xrightarrow[C^{k-1}]{df}\mathrm{GL}(F)\xrightarrow[C^\infty]{\mathrm{Inv}}\mathrm{GL}(F) \tag{17.211}$$

qui prouve que df^{-1} est C^{k-1} et donc que f^{-1} est C^k. La récurrence s'arrête ici parce que df n'est pas mieux que C^{k-1}. □

17.8.3 Théorème de la fonction implicite

Nous énonçons et démontrons le théorème de la fonction implicite dans le cas d'espaces de Banach.

ThoAcaWho

Théorème 17.48 (Théorème de la fonction implicite dans Banach[355])**.**
Soient E, F et G des espaces de Banach et des ouverts $U\subset E$, $V\subset F$. Nous considérons une fonction $f\colon U\times V\to G$ de classe C^r telle que[30]

$$d_yf_{(x_0,y_0)}\colon F\to G \tag{17.212}$$

soit un isomorphisme pour un certain $(x_0,y_0)\in U\times V$.

Alors nous avons des voisinages U_0 de x_0 dans E et W_0 de $f(x_0,y_0)$ dans G et une fonction de classe C^r

$$g\colon U_0\times W_0\to V \tag{17.213}$$

telle que

$$f\big(x,g(x,w)\big)=w \tag{17.214}$$

pour tout $(x,w)\in U_0\times W_0$.

Cette fonction g est unique au sens suivant : il existe un voisinage V_0 de y_0 tel que si $(x,y)\in U_0\times V_0$ et $w\in W_0$ satisfont à $f(x,y)=w$ alors $y=g(x,w)$. Autrement dit, la fonction $g\colon U_0\times W_0\to V_0$ est unique.

Démonstration. Nous commençons par considérer la fonction

$$\begin{aligned}\Phi\colon U\times V&\to E\times G\\ (x,y)&\mapsto\big(x,f(x,y)\big)\end{aligned} \tag{17.215}$$

30. La notation d_y est la différentielle partielle de la définition 10.297.

et sa différentielle

$$d\Phi_{(x_0,y_0)}(u,v) = \frac{d}{dt}\Big[\big(x_0+tu, f(x_0+tu,y_0+tv)\big)\Big]_{t=0} \tag{17.216a}$$
$$= \left(\frac{d}{dt}\Big[x_0+tu\Big]_{t=0}, \frac{d}{dt}\Big[f(x_0+tu,y_0+tv)\Big]_{t=0}\right) \tag{17.216b}$$
$$= \big(u, df_{(x_0,y_0)}(u,v)\big)\,. \tag{17.216c}$$

Nous utilisons alors la proposition 10.298 pour conclure que

$$d\Phi_{(x_0,y_0)}(u,v) = \big(u, (d_1f)_{(x_0,y_0)}(u) + (d_2f)_{(x_0,y_0)}(v)\big), \tag{17.217}$$

mais comme par hypothèse $(d_2f)_{(x_0,y_0)}\colon F \to G$ est un isomorphisme, l'application $d\Phi_{(x_0,y_0)}\colon E\times F \to E\times G$ est également un isomorphisme. Par conséquent le théorème d'inversion locale 17.47 nous indique qu'il existe un voisinage $\mathcal{O}$ de (x_0,y_0) et $\mathcal{P}$ de $\Phi(x_0,y_0)$ tels que $\Phi\colon \mathcal{O}\to\mathcal{P}$ soit une bijection et $\Phi^{-1}\colon \mathcal{P}\to\mathcal{O}$ soit de classe C^r. Comme $\mathcal{P}$ est un voisinage de

$$\Phi(x_0,y_0) = \big(x_0, f(x_0,y_0)\big), \tag{17.218}$$

nous pouvons par 7.230 le choisir un peu plus petit de telle sorte à avoir $\mathcal{P} = U_0\times W_0$ où U_0 est un voisinage de x_0 et W_0 un voisinage de $f(x_0,y_0)$. Dans ce cas nous devons obligatoirement aussi restreindre $\mathcal{O}$ à $U_0\times V_0$ pour un certain voisinage V_0 de y_0. L'application Φ^{-1} a obligatoirement la forme

$$\begin{aligned}\Phi^{-1}\colon U_0\times W_0 &\to U_0\times V_0\\ (x,w) &\mapsto \big(x, g(x,w)\big)\end{aligned} \tag{17.219}$$ EqMHT_QrHBn

pour une certaine fonction $g\colon U_0\times W_0\to V$. Cette fonction g est la fonction cherchée parce qu'en appliquant Φ à (17.219),

$$(x,w) = \Phi\big(x,g(x,w)\big) = \Big(x, f\big(x,g(x,w)\big)\Big), \tag{17.220}$$

qui nous dit que pour tout $x\in U_0$ et tout $w\in W_0$ nous avons

$$f\big(x,g(x,w)\big) = w. \tag{17.221}$$

Si vous avez bien suivi le sens de l'équation (17.219) alors vous avez compris l'unicité. Sinon, considérez $(x,y)\in U_0\times V_0$ et $w\in W_0$ tels que $f(x,y)=w$. Alors $\big(x,f(x,y)\big) = (x,w)$ et

$$\Phi(x,y) = (x,w). \tag{17.222}$$

Mais vu que $\Phi\colon U_0\times V_0\to U_0\times W_0$ est une bijection, cette relation définit de façon univoque l'élément (x,y) de $U_0\times V_0$, qui ne sera autre que $g(x,w)$. □

Le théorème de la fonction implicite s'énonce de la façon suivante pour des espaces de dimension finie.

ThoRYN_jvZrZ

Théorème 17.49 (Théorème de la fonction implicite en dimension finie)**.**
Soit une fonction $F\colon \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ de classe C^k et $(\alpha,\beta)\in\mathbb{R}^n\times\mathbb{R}^m$ tels que

(1) $F(\alpha,\beta)=0$,

(2) $\frac{\partial(F_1,\dots,F_m)}{\partial(y_1,\dots,y_m)}\neq 0$, *c'est-à-dire que $(d_yF)_{(\alpha,\beta)}$ est inversible.*

Alors il existe un voisinage ouvert V de α dans $\mathbb{R}^n$, un voisinage ouvert W de β dans $\mathbb{R}^m$ et une application $\varphi\colon V\to W$ de classe C^k telle que pour tout $x\in V$ on ait

$$F\big(x,\varphi(x)\big) = 0. \tag{17.223}$$

De plus si $(x,y)\in V\times W$ satisfait à $F(x,y)=0$, alors $y=\varphi(x)$.

RemPYA_pkTEx

Remarque 17.50.
Notons que cet énoncé est tourné un peu différemment en ce qui concerne le nombre de variables dont dépend la fonction implicite : comparez

$$f\big(x, g(x,w)\big) = w \tag{17.224a}$$
$$F\big(x, \varphi(x)\big) = 0. \tag{17.224b}$$

Le deuxième est un cas particulier du premier en posant

$$F(x,y) = f(x,y) - f(x_0,y_0) \tag{17.225}$$

et donc en considérant w comme valant la constante $f(x_0,y_0)$; dans ce cas la fonction g ne dépend plus que de la variable x.

Exemple 17.51.
La remarque 17.50 signifie entre autres que le théorème 17.48 est plus fort que 17.49 parce que le premier permet de choisir la valeur d'arrivée. Parlons de l'exemple classique du cercle et de la fonction $f(x,y) = x^2 + y^2$. Nous savons que

$$f(\alpha,\beta) = 1. \tag{17.226}$$

Alors le théorème 17.48 nous donne une fonction g telle que

$$f(x, g(x,r)) = r \tag{17.227}$$

tant que x est proche de α, que r est proche de 1 et que g donne des valeurs proches de β.

L'énoncé 17.49 nous oblige à travailler avec la fonction $F(x,y) = x^2 + y^2 - 1$, de telle sorte que

$$F(\alpha,\beta) = 0, \tag{17.228}$$

et que nous ayons une fonction φ telle que

$$F(x, \varphi(x)) = 0. \tag{17.229}$$

La fonction φ ne permet donc que de trouver des points sur le cercle de rayon 1. △

17.8.4 Exemple

Le théorème de la fonction implicite a pour objet de donner l'existence de la fonction φ. Maintenant nous pouvons dire beaucoup de choses sur les dérivées de φ en considérant la fonction

$$x \mapsto F\big(x, \varphi(x)\big). \tag{17.230}$$

Par définition de φ, cette fonction est toujours nulle. En particulier, nous pouvons dériver l'équation

$$F\big(x, \varphi(x)\big) = 0, \tag{17.231}$$

et nous trouvons plein de choses.

EXooTLNAooCJHPnq

Exemple 17.52.
Prenons par exemple la fonction[31]

$$F\big((x,y),z\big) = ze^z - x - y, \tag{17.232}$$

Le théorème de la fonction implicite 17.48 nous permet de considérer la fonction $z(x,y)$ vérifiant

$$F\big(x,y,z(x,y)\big) = 0, \tag{17.233}$$

31. Définition de l'exponentielle : 15.60.

c'est-à-dire telle que

$$z(x,y)e^{z(x,y)} - x - y = 0. \tag{17.234}$$ EqDefZImplExemple

pour tout x et $y \in \mathbb{R}$. Que pouvons dire de la fonction z ?

Nous pouvons facilement trouver $z(0,0)$ parce que

$$z(0,0)e^{z(0,0)} = 0, \tag{17.235}$$

donc $z(0,0) = 0$.

Nous pouvons dire des choses sur les dérivées de $z(x,y)$. Voyons par exemple $(\partial_x z)(x,y)$. Pour trouver cette dérivée, nous dérivons la relation (17.234) par rapport à x. Ce que nous trouvons est

$$(\partial_x z)e^z + ze^z(\partial_x z) - 1 = 0. \tag{17.236}$$

Cette équation peut être résolue par rapport à $\partial_x z$:

$$\frac{\partial z}{\partial x}(x,y) = \frac{1}{e^z(1+z)}. \tag{17.237}$$

Remarquez que cette équation ne donne pas tout à fait la dérivée de z en fonction de x et y, parce que z apparaît dans l'expression, alors que z est justement la fonction inconnue. En général, c'est la vie, nous ne pouvons pas faire mieux.

Dans certains cas, on peut aller plus loin. Par exemple, nous pouvons calculer cette dérivée au point $(x,y) = (0,0)$ parce que $z(0,0)$ est connu :

$$\frac{\partial z}{\partial x}(0,0) = 1. \tag{17.238}$$

Ceci est pratique pour calculer, par exemple, le développement en Taylor de z autour de $(0,0)$. △

Exemple 17.53.
Est-ce que l'équation $e^y + xy = 0$ définit au moins localement une fonction $y(x)$? Nous considérons la fonction

$$f(x,y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} \tag{17.239}$$

La différentielle de cette application est

$$df_{(0,0)}(u) = \frac{d}{dt}\Big[f(tu_1, tu_2)\Big]_{t=0} = \frac{d}{dt}\begin{pmatrix} tu_1 \\ e^{tu_2} + t^2u_1u_2 \end{pmatrix}_{t=0} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{17.240}$$

L'application f définit donc un difféomorphisme local autour des points (x_0,y_0) et $f(x_0,y_0)$. Soit $(u,0)$ un point dans le voisinage de $f(x_0,y_0)$. Alors il existe un unique (x,y) tel que

$$f(x,y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}. \tag{17.241}$$

Nous avons automatiquement $x = u$ et $e^y + xy = 0$. Notons toutefois que pour que ce procédé donne effectivement une fonction implicite $y(x)$ nous devons avoir des points de la forme $(u,0)$ dans le voisinage de $f(x_0,y_0)$. △

17.9 Décomposition polaire (régularité)

NomDJMUooTRUVkS

17.54.
Nous allons montrer que l'application

$$\begin{aligned} f\colon S^{++}(n,\mathbb{R}) &\to S^{++}(n,\mathbb{R}) \\ A &\mapsto \sqrt{A} \end{aligned} \tag{17.242}$$

est une difféomorphisme.

Cependant $S^{++}(n,\mathbb{R})$ n'est pas un ouvert de $\mathbb{M}(n,\mathbb{R})$ et nous ne savons pas ce qu'est la différentielle d'une application non définie sur un ouvert. Nous allons donc en réalité montrer que l'application racine carrée existe sur un voisinage de chacun des points de $S^{++}(n,\mathbb{R})$. Et comme une union quelconque d'ouverts est un ouvert, la fonction f sera bien définie sur un ouvert de $\mathbb{M}(n,\mathbb{R})$.

LemLBFOooDdNcgy

Lemme 17.55.
L'application

$$\begin{aligned} f\colon S^{++}(n,\mathbb{R}) &\to S^{++}(n,\mathbb{R}) \\ A &\mapsto A^2 \end{aligned} \tag{17.243}$$

est un C^∞-difféomorphisme.

Démonstration. Prouvons d'abord que f prend ses valeurs dans $S^{++}(n,\mathbb{R})$. Si $A\in S^{++}(n,\mathbb{R})$ alors par la diagonalisation 8.227 elle s'écrit $A=QDQ^{-1}$ où D est diagonale avec des nombres strictement positifs sur la diagonale. Avec cela, $A^2=QD^2Q^{-1}$ où D^2 contient encore des nombres strictement positifs sur la diagonale.

L'application f étant essentiellement des polynômes en les entrées de A, elle est de classe C^∞.

Passons à l'étude de la différentielle. Comme mentionné en 17.54 nous allons en réalité voir f sur un ouvert de $\mathbb{M}(n,\mathbb{R})$ autour de $A\in S^{++}(n,\mathbb{R})$. Par conséquent si $A\in S^{++}(n,\mathbb{R})$,

$$df\colon S^{++}(n,\mathbb{R}) \to \mathcal{L}\big(\mathbb{M}(n,\mathbb{R}),\mathbb{M}(n,\mathbb{R})\big) \tag{17.244a}$$
$$df_A\colon \mathbb{M}(n,\mathbb{R}) \to \mathbb{M}(n,\mathbb{R}). \tag{17.244b}$$

Le calcul de df_A est facile. Soit $u\in\mathbb{M}(n,\mathbb{R})$ et faisons le calcul en utilisant la formule du lemme (12.255) :

$$df_A(u) = \frac{d}{dt}\Big[f(A+tu)\Big]_{t=0} \tag{17.245a}$$
$$= \frac{d}{dt}\Big[A^2+tAu+tuA+t^2u^2\Big]_{t=0} \tag{17.245b}$$
$$= Au+uA. \tag{17.245c}$$

Nous allons utiliser le théorème d'inversion locale 17.47 à la fonction f. Dans la suite, A est une matrice de $S^{++}(n,\mathbb{R})$.

(1) **df_A est injective** Soit $M\in\mathbb{M}(n,\mathbb{R})$ dans le noyau de df_A. En posant $M'=Q^{-1}MQ$ nous avons $M=QM'Q^{-1}$ et on applique df_A à $QM'Q^{-1}$:

$$df_A(QM'Q^{-1}) = Q\big(DM+MD\big)Q^{-1}. \tag{17.246}$$

où $D=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$ avec $\lambda_i>0$. La matrice D est inversible. Nous avons $M'=-DM'D^{-1}$, et en coordonnées,

$$M'_{ij} = -\sum_{kl} D_{ik}M'_{kl}D^{-1}_{lj} \tag{17.247a}$$
$$= -\sum_{kl}\lambda_i\delta_{ik}M'_{kl}\frac{1}{\lambda_j}\delta_{lj} \tag{17.247b}$$
$$= -\frac{\lambda_i}{\lambda_j}M'_{ij}. \tag{17.247c}$$

C'est-à-dire que $M'_{ij}=-\frac{\lambda_i}{\lambda_j}M'_{ij}$ avec $-\frac{\lambda_i}{\lambda_j}<0$. Cela implique $M'=0$ et par conséquent $M=0$.

(2) $\underline{df_A \text{ est surjective}}$ Soit $N \in \mathbb{M}(n,\mathbb{R})$; nous cherchons $M \in \mathbb{M}(n,\mathbb{R})$ tel que $df_A(M) = N$. Nous posons $N' = Q^{-1}NQ$ et $M = QM'Q^{-1}$, ce qui nous donne à résoudre $df_D(M') = N'$. Passons en coordonnées :

$$(DM' + M'D)_{ij} = \sum_k (\delta_{ik}\lambda_i M'_{kj} + M'_{ik}\delta_{kj}\lambda_j) = M'_{ij}(\lambda_i + \lambda_j) \tag{17.248}$$

où $\lambda_i + \lambda_j \neq 0$. Il suffit donc de prendre la matrice M' donnée par

$$M'_{ij} = \frac{1}{\lambda_i + \lambda_j} N'_{ij} \tag{17.249}$$

pour que $df_A(M') = N'$.

Le théorème d'inversion locale donne un voisinage V de A dans $\mathbb{M}(n,\mathbb{R})$ et un voisinage W de A^2 dans $\mathbb{M}(n,\mathbb{R})$ tels que $f\colon V \to W$ soit une bijection et $f^{-1}\colon W \to V$ soit de même régularité, en l'occurrence C^∞. □

Remarque 17.56.
Oui, il y a des matrices non symétriques qui ont une unique racine carrée.

La proposition suivante, qui dépend du théorème d'inversion locale par le lemme 17.55, donne plus de régularité à la décomposition polaire donnée dans le théorème 13.32.

PropWCXAooDuFMjn

Proposition 17.57 (Décomposition polaire : cas réel (suite))**.**
L'application

$$\begin{aligned} f\colon \mathrm{O}(n,\mathbb{R}) \times S^{++}(n,\mathbb{R}) &\to \mathrm{GL}(n,\mathbb{R}) \\ (Q,S) &\mapsto SQ \end{aligned} \tag{17.250}$$

est un difféomorphisme de classe C^∞.

Démonstration. Si M est donnée dans $\mathrm{GL}(n,\mathbb{R})$ alors la décomposition polaire[32] $M = QS$ est donnée par $S = \sqrt{MM^t}$ et $Q = MS^{-1}$. Autrement dit, si nous considérons la fonction de décomposition polaire

$$f\colon \mathrm{O}(n,\mathbb{R}) \times S^{++}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R}) \tag{17.251}$$

alors

$$f^{-1}(M) = \big(M(\sqrt{MM^t})^{-1}, \sqrt{MM^t}\big). \tag{17.252}$$

Nous avons vu dans le lemme 17.55 que la racine carrée était un C^∞-difféomorphisme. Le reste n'étant que des produits de matrices, la régularité est de mise. □

17.10 Théorème de Von Neumann

Lemme 17.58 ([104])**.**
Soit G, un sous-groupe fermé de $\mathrm{GL}(n,\mathbb{R})$ et

$$\mathcal{L}_G = \{m \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } e^{tm} \in G\ \forall t \in \mathbb{R}\}. \tag{17.253}$$

Alors $\mathcal{L}_G$ est un sous-espace vectoriel de $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Si $m \in \mathcal{L}_G$, alors $\lambda m \in \mathcal{L}_G$ par construction. Le point délicat à prouver est le fait que si $a, b \in \mathcal{L}_G$, alors $a + b \in \mathcal{L}_G$. Soit $a \in \mathbb{M}(n,\mathbb{R})$; nous savons qu'il existe une fonction $\alpha_a\colon \mathbb{R} \to \mathbb{M}$ telle que

$$e^{ta} = \mathbb{1} + ta + \alpha_a(t) \tag{17.254}$$

et

$$\lim_{t\to 0} \frac{\alpha(t)}{t} = 0. \tag{17.255}$$

32. Proposition 13.32.

Si a et b sont dans $\mathcal{L}_G$, alors $e^{ta}e^{tb} \in G$, mais il n'est pas vrai en général que cela soit égal à $e^{t(a+b)}$. Pour tout $k \in \mathbb{N}$ nous avons

$$e^{a/k}e^{b/k} = \left(\mathbb{1} + \frac{a}{k} + \alpha_a(\frac{1}{k})\right)\left(\mathbb{1} + \frac{b}{k} + \alpha_b(\frac{1}{k})\right) = \mathbb{1} + \frac{a+b}{k} + \beta\left(\frac{1}{k}\right) \tag{17.256}$$

où $\beta\colon \mathbb{R} \to \mathbb{M}$ est encore une fonction vérifiant $\beta(t)/t \to 0$. Si k est assez grand, nous avons

$$\left\| \frac{a+b}{k} + \beta(\frac{1}{k}) \right\| < 1, \tag{17.257}$$

et nous pouvons profiter du lemme 15.152 pour écrire alors

$$\left(e^{a/k}e^{b/k}\right)^k = e^{k \ln\left(\mathbb{1} + \frac{a+b}{k} + \beta(\frac{1}{k})\right)}. \tag{17.258}$$

Ce qui se trouve dans l'exponentielle est

$$k\left[\frac{a+b}{k} + \alpha(\frac{1}{k}) + \sigma\left(\frac{a+b}{k} + \alpha(\frac{1}{k})\right)\right]. \tag{17.259}$$

Les diverses propriétés vues montrent que le tout tend vers $a+b$ lorsque $k \to \infty$. Par conséquent

$$\lim_{k\to\infty}\left(e^{a/k}e^{b/k}\right)^k = e^{a+b}. \tag{17.260}$$

Ce que nous avons prouvé est que pour tout t, $e^{t(a+b)}$ est une limite d'éléments dans G et est donc dans G parce que ce dernier est fermé. □

Comme $\mathcal{L}_G$ est un sous-espace vectoriel de $\mathbb{M}(n,\mathbb{R})$, nous pouvons considérer un supplémentaire M.

LemHOsbREC

Lemme 17.59.
Il n'existe pas de suite (m_k) dans $M\backslash\{0\}$ convergeant vers zéro et telle que $e^{m_k} \in G$ pour tout k.

Démonstration. Supposons que nous ayons $m_k \to 0$ dans $M\backslash\{0\}$ avec $e^{m_k} \in G$. Nous considérons les éléments $\epsilon_k = \frac{m_k}{\|m_k\|}$ qui sont sur la sphère unité de $\mathrm{GL}(n,\mathbb{R})$. Quitte à prendre une sous-suite, nous pouvons supposer que cette suite converge, et puisque M est fermé, ce sera vers $\epsilon \in M$ avec $\|\epsilon\| = 1$. Pour tout $t \in \mathbb{R}$ nous avons

$$e^{t\epsilon} = \lim_{k\to\infty} e^{t\epsilon_k}. \tag{17.261}$$

En vertu de la décomposition d'un réel en partie entière et décimale, pour tout k nous avons $\lambda_k \in \mathbb{Z}$ et $|\mu_k| \leqslant \frac{1}{2}$ tel que $t/\|m_k\| = \lambda_k + \mu_k$. Avec ça,

$$e^{t\epsilon} = \lim_{k\to\infty} \exp\left(\frac{t}{m_k}m_k\right) = \lim_{k\to\infty} e^{\lambda_k m_k}e^{\mu_k m_k}. \tag{17.262}$$

Pour tout k nous avons $e^{\lambda_k m_k} \in G$. De plus $|\mu_k|$ étant borné et m_k tendant vers zéro nous avons $e^{\mu_k m_k} \to 1$. Au final

$$e^{t\epsilon} = \lim_{k\to\infty} e^{t\epsilon_k} \in G \tag{17.263}$$

Cela signifie que $\epsilon \in \mathcal{L}_G$, ce qui est impossible parce que nous avions déjà dit que $\epsilon \in M\backslash\{0\}$. □

LemGGTtxdF

Lemme 17.60.
L'application

$$\begin{aligned} f\colon \mathcal{L}_G \times M &\to \mathrm{GL}(n,\mathbb{R}) \\ l, m &\mapsto e^l e^m \end{aligned} \tag{17.264}$$

est un difféomorphisme local entre un voisinage de $(0,0)$ dans $\mathbb{M}(n,\mathbb{R})$ et un voisinage de $\mathbb{1}$ dans $\exp\big(\mathbb{M}(n,\mathbb{R})\big)$.

Notons que nous ne disons rien de $e^{\mathbb{M}(n,\mathbb{R})}$. Nous n'allons pas nous embarquer à discuter si ce serait tout $\mathrm{GL}(n,\mathbb{R})$ [33] ou bien si ça contiendrait ne fut-ce que G.

Démonstration. Le fait que f prenne ses valeurs dans $\mathrm{GL}(n,\mathbb{R})$ est simplement dû au fait que les exponentielles sont toujours inversibles. Nous considérons ensuite la différentielle : si $u \in \mathcal{L}_G$ et $v \in M$ nous avons

$$df_{(0,0)}(u,v) = \frac{d}{dt}\Big[f\big(t(u,v)\big)\Big]_{t=0} = \frac{d}{dt}\Big[e^{tu}e^{tv}\Big]_{t=0} = u+v. \tag{17.265}$$

L'application df_0 est donc une bijection entre $\mathcal{L}_G \times M$ et $\mathbb{M}(n,\mathbb{R})$. Le théorème d'inversion locale 17.47 nous assure alors que f est une bijection entre un voisinage de $(0,0)$ dans $\mathcal{L}_G \times M$ et son image. Mais comme df_0 est une bijection avec $\mathbb{M}(n,\mathbb{R})$, l'image en question contient un ouvert autour de $\mathbb{1}$ dans $\exp\big(\mathbb{M}(n,\mathbb{R})\big)$. □

ThoOBriEoe

Théorème 17.61 (Von Neumann[104, 512, 513])**.**
Tout sous-groupe fermé de $\mathrm{GL}(n,\mathbb{R})$ *est une sous-variété de* $\mathrm{GL}(n,\mathbb{R})$.

Démonstration. Soit G un tel groupe ; nous devons prouver que c'est localement difféomorphe à un ouvert de $\mathbb{R}^n$. Et si on est pervers, on ne va pas faire localement difféomorphe à un ouvert de $\mathbb{R}^n$, mais à un ouvert d'un espace vectoriel de dimension finie. Nous allons être pervers.

Étant donné que pour tout $g \in G$, l'application

$$\begin{aligned} L_g\colon G &\to G \\ h &\mapsto gh \end{aligned} \tag{17.266}$$

est de classe C^∞ et d'inverse C^∞, il suffit de prouver le résultat pour un voisinage de $\mathbb{1}$.

Supposons d'abord que $\mathcal{L}_G = \{0\}$. Alors 0 est un point isolé de $\ln(G)$; en effet si ce n'était pas le cas nous aurions un élément m_k de $\ln(G)$ dans chaque boule $B(0,r_k)$. Nous aurions alors $m_k = \ln(a_k)$ avec $a_k \in G$ et donc

$$e^{m_k} = a_k \in G. \tag{17.267}$$

De plus m_k appartient forcément à M parce que $\mathcal{L}_G$ est réduit à zéro. Cela nous donnerait une suite $m_k \to 0$ dans M dont l'exponentielle reste dans G. Or cela est interdit par le lemme 17.59. Donc 0 est un point isolé de $\ln(G)$. L'application ln étant continue [34], nous en déduisons que $\mathbb{1}$ est isolé dans G. Par le difféomorphisme L_g, tous les points de G sont isolés ; ce groupe est donc discret et par voie de conséquence, une variété.

Nous supposons maintenant que $\mathcal{L}_G \neq \{0\}$. Nous savons par la proposition 15.151 que

$$\exp\colon \mathbb{M}(n,\mathbb{R}) \to \mathbb{M}(n,\mathbb{R}) \tag{17.268}$$

est une application C^∞ vérifiant $d\exp_0 = \mathrm{Id}$. Nous pouvons donc utiliser le théorème d'inversion locale 17.47 qui nous offre donc l'existence d'un voisinage U de 0 dans $\mathbb{M}(n,\mathbb{R})$ tel que $W = \exp(U)$ soit un ouvert de $\mathrm{GL}(n,\mathbb{R})$ et que $\exp\colon U \to W$ soit un difféomorphisme de classe C^∞.

Montrons que quitte à restreindre U (et donc W qui reste par définition l'image de U par exp), nous pouvons avoir $\exp\big(U \cap \mathcal{L}_G\big) = W \cap G$. D'abord $\exp(\mathcal{L}_G) \subset G$ par construction. Nous avons donc $\exp\big(U \cap \mathcal{L}_G\big) \subset W \cap G$. Pour trouver une restriction de U pour laquelle nous avons l'égalité, nous supposons que pour tout ouvert $\mathcal{O}$ dans U,

$$\exp\colon \mathcal{O} \cap \mathcal{L}_G \to \exp(\mathcal{O}) \cap G \tag{17.269}$$

ne soit pas surjective. Cela donnerait un élément de $\mathcal{O} \cap \complement\mathcal{L}_G$ dont l'image par exp n'est pas dans G. Nous construisons ainsi une suite en considérant une boule $B(0,\frac{1}{k})$ incluse à U et $x_k \in B(0,\frac{1}{k}) \cap \complement\mathcal{L}_G$ vérifiant $e^{x_k} \in G$. D'après le choix des boules, nous avons évidemment $x_k \to 0$.

33. Vu les dimensions y'a tout de même peu de chance.
34. Par le lemme 15.152.

L'élément e^{x_k} est dans $e^{\mathbb{M}(n,\mathbb{R})}$ et le difféomorphisme du lemme 17.60 [35] nous donne $(l_k, m_k) \in \mathcal{L}_G \times M$ tel que $e^{l_k}e^{m_k} = e^{x_k}$. À ce point nous considérons k suffisamment grand pour que e^{x_k} soit dans la partie de l'image de f sur lequel nous avons le difféomorphisme. Plus prosaïquement, nous posons

$$(l_k, m_k) = f^{-1}(e^{x_k}) \tag{17.270}$$

et nous profitons de la continuité pour permuter la limite avec f^{-1} :

$$\lim_{k\to\infty}(l_k, m_k) = f^{-1}\Big(\lim_{k\to\infty} e^{x_k}\Big) = f^{-1}(\mathbb{1}) = (0, 0). \tag{17.271}$$

En particulier $m_k \to 0$ alors que $e^{m_k} = e^{x_k}e^{-l_k} \in G$. La suite m_k viole le lemme 17.59. Nous pouvons donc restreindre U de telle façon à avoir

$$\exp\big(U \cap \mathcal{L}_G\big) = W \cap G. \tag{17.272}$$

Nous avons donc un ouvert de $\mathcal{L}_G$ (l'ouvert $U \cap \mathcal{L}_G$) qui est difféomorphe avec l'ouvert $W \cap G$ de G. Donc G est une variété et accepte $\mathcal{L}_G$ comme carte locale. □

Remarque 17.62.
En termes savants, nous avons surtout montré que si G est un groupe de Lie d'algèbre de Lie $\mathfrak{g}$, alors l'exponentielle donne un difféomorphisme local entre $\mathfrak{g}$ et G.

17.11 Recherche d'extrémums

DEFooJMMLooYMZehc

Définition 17.63 ([514])**.**
*Si $f\colon \mathbb{R}^n \to \mathbb{R}$ est une fonction, le point $a \in \mathbb{R}^n$ est un **maximum local** de f si il existe $\delta > 0$ tel que pour tout $x \in B(a, \delta)$, $f(x) \leqslant f(a)$.*

*Je vous laisse deviner la définition d'un **minimum local**.*

*Un **extrémum local** est un point qui est soit un minimum soit un maximum local.*

17.11.1 Extrema à une variable

Définition 17.64.
*Soit $f\colon A \subset \mathbb{R} \to \mathbb{R}$ et $a \in A$. Le point a est un **maximum local** de f si il existe un voisinage $\mathcal{U}$ de a tel que $f(a) \geqslant f(x)$ pour tout $x \in \mathcal{U} \cap A$. Le point a est un **maximum global** si $f(a) \geqslant g(x)$ pour tout $x \in A$.*

La proposition basique à utiliser lors de la recherche d'extrémums est la suivante :

PROPooNVKXooXtKkuz

Proposition 17.65.
Soit $f\colon A \subset \mathbb{R} \to \mathbb{R}$ et $a \in \mathrm{Int}(A)$. Supposons que f est dérivable en a. Si a est un extrémum local[36]*, alors $f'(a) = 0$.*

La réciproque n'est pas vraie, comme le montre l'exemple de la fonction $x \mapsto x^3$ en $x = 0$: sa dérivée est nulle et pourtant $x = 0$ n'est ni un maximum ni un minimum local.

Cette proposition ne sert donc qu'à sélectionner des *candidats* extrémum. Afin de savoir si ces candidats sont des extrémums, il y a la proposition suivante.

PROPooCXHPooSlRsEJ

Proposition 17.66.
Soit $f\colon I \subset \mathbb{R} \to \mathbb{R}$, une fonction de classe C^k au voisinage d'un point $a \in \mathrm{Int}\, I$. Supposons que

$$f'(a) = f''(a) = \ldots = f^{(k-1)}(a) = 0, \tag{17.273}$$

et que

$$f^{(k)}(a) \neq 0. \tag{17.274}$$

Dans ce cas,

35. Il me semble que l'utilisation de ce lemme manque à l'avant-dernière ligne de la preuve chez [104].
36. Définition 17.63.

(1) Si k est pair, alors a est un point d'extrémum local de f, c'est un minimum si $f^{(k)}(a) > 0$, et un maximum si $f^{(k)}(a) < 0$,

(2) Si k est impair, alors a n'est pas un extrémum local de f.

Note : jusqu'à présent nous n'avons rien dit des extrémums *globaux* de f. Il n'y a pas grand chose à en dire. Si un point d'extrémum global est situé dans l'intérieur du domaine de f, alors il sera extrémum local (a fortiori). Ou alors, le maximum global peut être sur le bord du domaine. C'est ce qui arrive à des fonctions strictement croissantes sur un domaine compact.

Une seule certitude : si une fonction est continue sur un compact, elle possède une minimum et un maximum global par le théorème 9.55.

Soit une fonction $f\colon I \to \mathbb{R}$, et soit $a \in I$. Si $f'(a) > 0$, alors la tangente au graphe de f au point $\big(a, f(a)\big)$ sera une droite croissante (coefficient directeur positif). Cela ne veut pas spécialement dire que la fonction elle-même sera croissante, mais en tout cas, cela est un bon indice.

Exemple 17.67.
Si $f(x) = x^2$, il est connu que $f'(x) = 2x$. Nous avons donc que f' est positive si $x \geqslant 0$ et f' est négative si $x < 0$. Cela correspond bien au fait que x^2 est décroissante sur $]-\infty, 0[$ et croissante sur $]0, \infty[$. △

Sur la figure 17.2, nous avons dessiné la fonction $f(x) = x\cos(x)$ et sa dérivée. Nous voyons que partout où la dérivée est négative, la fonction est décroissante tandis que, inversement, partout où la dérivée est positive, la fonction est croissante.

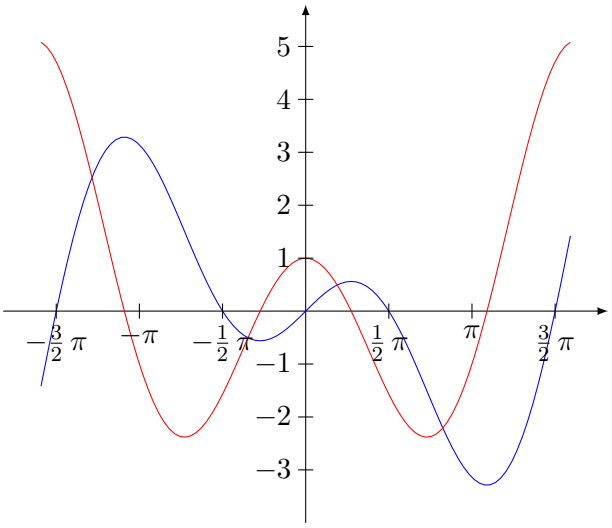


FIGURE 17.2: La fonction $f(x) = x\cos(x)$ en bleu et sa dérivée en rouge. LabelOfFigWIRAooTCcpOV

Les extrémums de la fonction f sont donc placés là où f' change de signe. En effet si $f'(x) < 0$ pour $x < a$ et $f'(x) > 0$ pour $x > a$, la fonction est décroissante jusqu'à a et est ensuite croissante. Cela signifie que la fonction connait un creux en a. Le point a est donc un minimum de la fonction.

Attention cependant. Le fait que $f'(a) = 0$ ne signifie pas automatiquement que f a un maximum ou un minimum en a. Nous avons par exemple tracé sur la figure 17.3 les fonctions x^3 et sa dérivée. Il est à noter que, conformément à ce que l'on pense, certes la dérivée s'annule en $x = 0$, mais elle ne change pas de signe.

17.11.2 Extrema libre

DEFooYJLZooLkEAYf

Définition 17.68.
*Un point a à l'intérieur du domaine d'une fonction $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$ est un **point critique** de f lorsque $df(a) = 0$.*

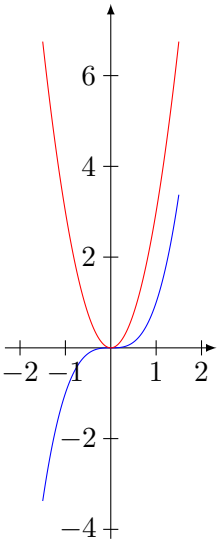


FIGURE 17.3: La dérivée de x^3 s'annule en $x = 0$, mais ce n'est ni un minimum ni un maximum. LabelFigVBOfooRHiKOfi

Ces points sont analogues aux points où la dérivée d'une fonction sur $\mathbb{R}$ s'annule. Les points critiques de f sont donc les candidats à être des points d'extrémum.

Dans le cas d'une fonction de deux variables, la proposition 12.348 nous permet de voir $(d^2 f)_a$ comme étant la matrice

$$d^2 f(a) = \begin{pmatrix} \frac{d^2 f}{dx^2}(a) & \frac{d^2 f}{dx\,dy}(a) \\ \frac{d^2 f}{dy\,dx}(a) & \frac{d^2 f}{dy^2}(a) \end{pmatrix}. \tag{17.275}$$

Dans le cas d'une fonction C^2, cette matrice est symétrique.

PropUQRooPgJsuz

Proposition 17.69 ([515]).
Soit un ouvert Ω de $\mathbb{R}^n$ et $a \in \Omega$. Soit une fonction $f\colon \Omega \to \mathbb{R}$ différentiable en a. Si a est un extrémum local de f, alors a est un point critique de f.

Démonstration. Nous supposons que a est un maximum local (ce sera la même chose si a est un minimum). Soit $r > 0$ tel que $f(x) \leqslant f(a)$ pour tout $x \in B(a, r)$ (et tel que cette boule reste dans Ω). Soit $u \in \mathbb{R}^n$ assez petit pour que $a \pm u \in B(a, r)$ de sorte que la définition suivante ait un sens :

$$\begin{aligned} g\colon [-1, 1] &\to \mathbb{R} \\ t &\mapsto f(a + tu) \end{aligned} \tag{17.276}$$

Cette fonction est différentiable en $t = 0$ (composée de fonctions différentiables, théorème 10.276) et a un maximum local en $t = 0$. Donc $g'(0) = 0$ par la proposition 17.65. Donc

$$0 = \frac{d}{dt}\Big[f(a + tu)\Big]_{t=0} = df_a(u). \tag{17.277}$$

□

17.11.3 Extremums et Hessienne

PropoExtreRn

Proposition 17.70 ([1, 516, 517]).
Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction $f\colon \Omega \to \mathbb{R}$ deux fois différentiable ainsi que $a \in \Omega$. ITEMooCBMYooQQMqQL

(1) Si a est un point critique de f et si il existe r tel que $(d^2 f_x)$ est semi-définie positive pour tout $x \in B(a, r)$ alors f possède un minimum local en a. ITEMooCVFVooWltGqI

(2) Si a est un point critique[37] *de f, et si $d^2 f_a$ est strictement définie positive*[38]*, alors a est un minimum local strict de f,*

37. Définition 17.68.
38. La fonction f est de classe C^2, donc les dérivées croisées sont égales et $d^2 f$ est symétrique. La définition 8.230 s'applique donc.

ItemPropoExtreRn

(3) *Si a est un minimum local, alors $(d^2f)_a$ est semi-définie positive.*

Démonstration. Nous subdivisons la preuve.

(1) **(1)**
Soit h tel que $a+h \in B(a,r)$. Nous allons montrer que $f(a) \leqslant f(a+h)$; cela montrera que $x=a$ est un minimum local. Pour cela nous utilisons un développement de Taylor [39] : il existe $c \in]a, a+h[$ tel que

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_c(h,h) \geqslant f(a) \tag{17.278}$$

parce que, par hypothèse, $(d^2f)_c$ est définie positive, et parce que $df_a = 0$.

(2) **(2)**
La forme bilinéaire d^2f_a est strictement définie positive, donc il existe $\alpha > 0$ tel que

$$d^2f_a(h,h) > \alpha\|h\|^2 \tag{17.279}$$

pour tout h. Nous utilisons encore Taylor : il existe une fonction ϵ telle que $\lim_{h\to 0}\epsilon(h) = 0$ et

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\epsilon(h). \tag{17.280}$$

En tenant compte du fait que $df_a = 0$,

$$f(a+h) > f(a) + \|h\|^2\big(\frac{1}{2}\alpha + \epsilon(h)\big). \tag{17.281}$$

La limite de ϵ nous dit qu'il existe $r > 0$ tel que $\|\epsilon(h)\| < \frac{1}{2}\alpha$ pour tout $h \in B(0,r)$. Pour ces valeurs de h nous avons

$$f(a+h) > f(a). \tag{17.282}$$

Donc a est un minimum local strict de f.

(3) **(3)** Si a est un minimum local, nous savons déjà $df_a = 0$ par la proposition 17.69. Nous écrivons le développement de Taylor de f à l'ordre 2 de la proposition 12.468 :

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(\|h\|). \tag{17.283}$$

En prenant h assez petit pour que $a+h$ ne sorte pas de la boule dans laquelle a est un minimum, nous avons $f(a+h) - f(a) > 0$. Donc

$$\frac{1}{2}(d^2f)_a(h,h) + \|h\|^2\alpha(\|h\|) > 0 \tag{17.284}$$

Nous divisons cela par $\|h\|^2$ et notons $e_h = h/\|h\|$:

$$\frac{1}{2}(d^2f)_a(e_h,e_h) + \alpha(\|h\|) > 0. \tag{17.285}$$

À la limite $h \to 0$, le premier terme est constant tandis que le deuxième tend vers zéro. À la limite,

$$(d^2f)_a(e_h,e_h) \geqslant 0. \tag{17.286}$$

La caractérisation du lemme 8.234(2) nous dit alors que $(d^2f)_a$ est semi-définie positive.

□

39. Proposition 12.471.

La partie (3) est tout à fait comparable au fait bien connu que, pour une fonction $f\colon \mathbb{R}\to\mathbb{R}$, si le point a est minimum local, alors $f'(a)=0$ et $f''(a)>0$.

Notons que le point (3) ne parle pas de minimum strict, et donc pas de matrice *strictement* définie positive.

Exemple 17.71 (Proposition 17.70(2) sans point critique)**.**
L'hypothèse de point critique pour l'utilisation de la stricte définition positive de d^2f_a est nécessaire. Soit en effet la fonction

$$f(x)=x^2+x. \tag{17.287}$$

Elle vérifie $f''(0)=2$, de telle sorte que sa différentielle seconde en zéro soit strictement définie positive. Le point $x=0$ n'est cependant même pas un minimum local. Entre autres parce que $f'(0)=1\neq 0$. △

La méthode pour chercher les extrémums de f est donc de suivre les points suivants :

(1) Trouver les candidats extrémums en résolvant $\nabla f=(0,0)$,
(2) écrire $d^2f(a)$ pour chacun des candidats
(3) calculer les valeurs propres de $d^2f(a)$, déterminer si la matrice est définie positive ou négative,
(4) conclure.

Une conséquence de la proposition 8.233(3) [40] est que si $\det M<0$, alors le point a n'est pas un extrémum dans le cas où $M=d^2f(a)$ par le point (3) de la proposition 17.70.

Exemple 17.72.
Soit la fonction $f(x,y)=x^4+y^4-4xy$. C'est une fonction différentiable sans problème. D'abord sa différentielle est

$$df=\big(4x^3-4y;4y^3-4x\big), \tag{17.288}$$

et la matrice des dérivées secondes est

$$M=d^2f(x,y)=\begin{pmatrix}12x^2 & -4\\ -4 & 12y^2\end{pmatrix}. \tag{17.289}$$

Nous avons $df=0$ pour les trois points $(0,0)$, $(1,1)$ et $(1,-1)$.

Pour le point $(0,0)$ nous avons

$$M=\begin{pmatrix}0 & -4\\ -4 & 0\end{pmatrix}, \tag{17.290}$$

dont les valeurs propres sont 4 et -4. Elle n'est donc ni définie ni semi-définie positive ou négative. La proposition 17.70(2) conclut donc $(0,0)$ n'est pas un extrémum local.

Au contraire pour les points $(1,1)$ et $(-1,-1)$ nous avons

$$M=\begin{pmatrix}12 & -4\\ -4 & 12\end{pmatrix}, \tag{17.291}$$

dont les valeurs propres sont 16 et 8. La matrice d^2f y est donc définie positive. Ces deux points sont donc extrémums locaux. △

17.11.4 Un peu de recettes de cuisine

(1) Rechercher les points critiques, càd les (x,y) tels que

$$\begin{cases}\frac{\partial f}{\partial x}(x,y)=0\\ \frac{\partial f}{\partial y}(x,y)=0\end{cases}$$

En effet, si (x_0,y_0) est un extrémum local de f, alors $\frac{\partial f}{\partial x}(x_0,y_0)=0=\frac{\partial f}{\partial y}(x_0,y_0)$.

40. La matrice $d^2f(a)$ est toujours symétrique quand f est de classe C^2.

(2) Déterminer la nature des points critiques : « test » des dérivées secondes :

$$\text{On pose } H(x_0,y_0)=\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\frac{\partial f^2}{\partial y^2}(x_0,y_0)-\left(\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)\right)^2$$

(2a) Si $H(x_0,y_0)>0$ et $\frac{\partial^2 f}{\partial x^2}(x_0,y_0)>0 \Longrightarrow (x_0,y_0)$ est un minimum local de f.

(2b) Si $H(x_0,y_0)>0$ et $\frac{\partial^2 f}{\partial x^2}(x_0,y_0)<0 \Longrightarrow (x_0,y_0)$ est un maximum local de f.

(2c) Si $H(x_0,y_0)<0 \Longrightarrow f$ a un point de selle en (x_0,y_0).

(2d) Si $H(x_0,y_0)=0 \Longrightarrow$, on ne peut rien conclure.

17.11.5 Extrema liés

Soit f, une fonction sur $\mathbb{R}^n$, et $M\subset\mathbb{R}^n$ une variété de dimension m. Nous voulons savoir quelles sont les plus grandes et plus petites valeurs atteintes par f sur M.

Pour ce faire, nous avons un théorème qui permet de trouver des extrémums *locaux* de f sur la variété. Pour rappel, $a\in M$ est une **extrémum local de f relativement** à l'ensemble M si il existe une boule $B(a,\epsilon)$ telle que $f(a)\leqslant f(x)$ pour tout $x\in B(a,\epsilon)\cap M$.

ThoRGJosS

Théorème 17.73 (Extremum lié [496]).
Soit A, un ouvert de $\mathbb{R}^n$ et

(1) une fonction (celle à minimiser) $f\in C^1(A,\mathbb{R})$,

(2) des fonctions (les contraintes) $G_1,\dots,G_r\in C^1(A,\mathbb{R})$,

(3) $M=\{x\in A \text{ tel que } G_i(x)=0\,\forall i\}$,

(4) un extrémum local $a\in M$ de f relativement à M.

Supposons que les gradients $\nabla G_1(a)$, ...,$\nabla G_r(a)$ soient linéairement indépendants. Alors $a=(x_1,\dots,x_n)$ est une solution de $\nabla L(a)=0$ où

$$L(x_1,\dots,x_n,\lambda_1,\dots,\lambda_r)=f(x_1,\dots,x_n)+\sum_{i=1}^{r}\lambda_i G_i(x_1,\dots,x_n). \tag{17.292}$$

Autrement dit, si a est un extrémum lié, alors $\nabla f(a)$ est une combinaisons des $\nabla G_i(a)$, ou encore il existe des λ_i tels que

EqRDsSXyZ

$$df(a)=\sum_i \lambda_i dG_i(a). \tag{17.293}$$

La fonction L est le **lagrangien** du problème et les variables λ_i sont les **multiplicateurs de Lagrange**.

Démonstration. Si $r=n$ alors les vecteurs linéairement indépendantes $\nabla G_i(a)$ forment une base de $\mathbb{R}^n$ et donc évidemment les λ_i existent. Nous supposons donc maintenant que $r<n$. Nous notons $(z_i)_{i=1\dots n}$ les coordonnées sur $\mathbb{R}^n$.

La matrice

$$\begin{pmatrix}\frac{\partial G_1}{\partial z_1}(a) & \cdots & \frac{\partial G_1}{\partial z_n}(a)\\ \vdots & \ddots & \vdots\\ \frac{\partial G_r}{\partial z_1}(a) & \cdots & \frac{\partial G_r}{\partial z_n}(a)\end{pmatrix} \tag{17.294}$$

est de rang r parce que les lignes sont par hypothèses linéairement indépendantes. Nous nommons $(y_i)_{i=1,\dots,r}$ un choix de r parmi les (z_i) tels que

$$\det\begin{pmatrix}\frac{\partial G_1}{\partial y_1} & \cdots & \frac{\partial G_1}{\partial y_r}\\ \vdots & \ddots & \vdots\\ \frac{\partial G_r}{\partial y_1} & \cdots & \frac{\partial G_r}{\partial y_r}\end{pmatrix}\neq 0. \tag{17.295}$$

Nous identifions $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$ dans lequel $\mathbb{R}^r$ est la partie générée par les $(y_i)_{i=1,\dots,r}$. Les coordonnées sur $\mathbb{R}^s$ seront nommées $(x_j)_{j=1,\dots,s}$, de telle sorte que les coordonnées sur $\mathbb{R}^n$ seront $x_1,\dots,x_s,y_1,\dots,y_r$. Dans ces coordonnées, nous nommons $a=(\alpha,\beta)$ avec $\alpha\in\mathbb{R}^s$ et $\beta\in\mathbb{R}^r$.

Si nous notons $G=(G_1,\dots,G_r)$, le théorème de la fonction implicite (théorème 17.48) nous dit qu'il existe un voisinage U' de $\alpha\in\mathbb{R}^n$, un voisinage V' de $\beta\in\mathbb{R}^r$ et une fonction $\varphi\colon U'\to V'$ de classe C^1 telle que si $(x,y)\in U'\times V'$, alors

$$G(x,y)=0 \tag{17.296}$$

si et seulement si $y=\varphi(x)$. Nous posons maintenant

$$\psi(x)=(x,\varphi(x)) \tag{17.297a}$$
$$h(x)=f\big(\psi(x)\big). \tag{17.297b}$$

Nous avons $\psi(\alpha)=a$ et $\psi(x)\in M$ pour tout $x\in U'$. La fonction h a donc un extrémum local en α et donc les dérivées partielles de h y sont nulles. Cela signifie que

$$0=\frac{\partial h}{\partial x_i}(\alpha)=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial x_i}+\sum_{k=1}^{r}\frac{\partial f}{\partial y_k}\frac{\partial\varphi_k}{\partial x_i}, \tag{17.298}$$

c'est-à-dire

$$\frac{\partial f}{\partial x_i}(\alpha)+\sum_{k=1}^{r}\frac{\partial f}{\partial y_k}(a)\frac{\partial\varphi_k}{\partial x_i}(\alpha)=0 \tag{17.299}$$

pour tout $i=1,\dots,s$. D'autre part pour tout k, la fonction $l_k(x)=G_k\big(x,\varphi(x)\big)$ est constante et vaut zéro ; ses dérivées partielles sont donc nulles :

$$\frac{\partial l}{\partial x_i}(\alpha)=\frac{\partial G_k}{\partial x_i}(\alpha)+\sum_{k=1}^{r}\frac{\partial G_k}{\partial y_k}(a)\frac{\partial\varphi_k}{\partial x_i}(\alpha)=0 \tag{17.300}$$

pour tout $i=1,\dots,s$ et $k=1,\dots,r$.

Les s premières colonnes de la matrice

$$\begin{pmatrix}
\frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_s} & \frac{\partial f}{\partial y_1} & \dots & \frac{\partial f}{\partial y_r}\\
\frac{\partial G_1}{\partial x_1} & \dots & \frac{\partial G_1}{\partial x_s} & \frac{\partial G_1}{\partial y_1} & \dots & \frac{\partial G_1}{\partial y_r}\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\frac{\partial G_r}{\partial x_1} & \dots & \frac{\partial G_r}{\partial x_s} & \frac{\partial G_r}{\partial y_1} & \dots & \frac{\partial G_r}{\partial y_r}
\end{pmatrix} \tag{17.301}$$

s'expriment en termes des r dernières. La matrice est donc au maximum de rang r. Notons que la première ligne est ∇f et les r suivantes sont les ∇G_i. Vu que ces lignes sont des vecteurs liés, il existe $\mu_0,\dots,\mu_r$ tels que

$$\mu_0\nabla f+\sum_{i=1}^{r}\mu_i\nabla G_i=0. \tag{17.302}$$

Par hypothèse les ∇G_i sont linéairement indépendants, ce qui nous dit que $\mu_0\neq 0$. Donc nous avons ce qu'il nous faut :

$$\nabla f(a)=\sum_i\frac{\mu_i}{\mu_0}\nabla G_i(a). \tag{17.303}$$

Notons qu'au vu de l'expression (17.293), le fait que les formes $\{dG_i(a)\}_{1\leqslant i\leqslant r}$ forment une partie libre dans $(\mathbb{R}^n)^*$ implique que les λ_i sont uniques. □

La proposition suivante est la même que 17.73.

PropfPPUxh

Proposition 17.74.
Soit U, un ouvert de $\mathbb{R}^n$ et des fonctions de classe C^1 $f, g_1, \dots, g_r \colon U \to \mathbb{R}$. Nous considérons

$$\Gamma = \{x \in U \text{ tel que } g_1(x) = \dots = g_r(x) = 0\}. \tag{17.304}$$

Soit a un extrémum de $f|_\Gamma$. Supposons que les formes $dg_1, \dots, dg_r$ soient linéairement indépendantes en a. Alors il existe $\lambda_1, \dots, \lambda_r$ dans $\mathbb{R}$ tel que

$$df_a = \sum_{i=1}^{r} \lambda_i (dg_i)_a. \tag{17.305}$$

En pratique les candidats extrémums locaux sont tous les points où les gradients ne sont pas linéairement indépendants, plus tous les points donnés par l'équation $\nabla L = 0$. Parmi ces candidats, il faut trouver lesquels sont maximums ou minima, locaux ou globaux.

L'existence d'extrémums locaux se prouve généralement en invoquant de la compacité, et en invoquant le lemme suivant qui permet de réduire le problème à un compact.

LemmeMinSCimpliqueS

Lemme 17.75.
Soit S, une partie de $\mathbb{R}^n$ et C, un ouvert de $\mathbb{R}^n$. Si $a \in \operatorname{Int} S$ est un minimum local relatif à $S \cap C$, alors il est un minimum local par rapport à S.

Démonstration. Vu que a est un minimum local relatif à $S \cap C$, il existe un $\epsilon_1 > 0$ tel que pour tout $x \in B(a, \epsilon_1) \cap S \cap C$, nous avons $f(x) \geqslant f(a)$.

Mais étant donné que C est ouvert, et que $a \in C$, il existe un ϵ_2 tel que $B(a, \epsilon_2) \subset C$. En prenant $\epsilon = \min\{\epsilon_1, \epsilon_2\}$, nous trouvons que $f(x) \geqslant f(a)$ pour tout $x \in B(a, \epsilon) \cap (S \cap C) = B(a, \epsilon) \cap S$. □

17.12 Fonctions convexes

SECooVZWeWoQXjRXQz

Définition 17.76 ([518]).
*Une fonction f d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$ est dite **convexe** lorsque, pour tous x_1 et x_2 de I et tout λ dans $[0, 1]$ nous avons*

EQooYNAPooFePQZy

$$f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) \leqslant \lambda\, f(x_1) + (1-\lambda)\, f(x_2) \tag{17.306}$$

Si pour tout $\lambda \in \,]0, 1[$ et pour tout $x \neq y$ dans I nous avons

EQooEGBSooCevoPR

$$f\big(\lambda\, x_1 + (1-\lambda)\, x_2\big) < \lambda\, f(x_1) + (1-\lambda)\, f(x_2) \tag{17.307}$$

*alors nous disons que la fonction f est **strictement convexe** sur I.*

*Une fonction est **concave** si son opposée est convexe.*

17.77 ([518]).
Les différents résultats pour les fonctions convexes s'adaptent généralement sans mal aux fonctions strictement convexes. Une nuance cependant : de même que les fonctions dérivables convexes sont celles qui ont une dérivée croissante, les fonctions dérivables strictement convexes sont celles qui ont une dérivée strictement croissante (proposition 17.83). En revanche, il ne faudrait pas croire que la dérivée seconde d'une fonction dérivable strictement convexe est nécessairement une fonction à valeurs strictement positives (voir théorème 17.84) : la dérivée d'une fonction strictement croissante peut s'annuler occasionnellement, ou plus exactement peut s'annuler sur un ensemble de points d'intérieur vide. Penser à $x \mapsto x^4$ pour un exemple de fonction strictement convexe dont la dérivée seconde s'annule.

17.78.
Demander un $\lambda \in [0, 1]$ est équivalent à demande des $\alpha, \beta \geqslant 0$ tels que $\alpha + \beta = 1$. En effet, si $\alpha + \beta = 1$, alors il suffit de se dire que α est le λ et que $\beta = 1 - \alpha$.

LEMooXUDSooBUMnaZ

Lemme 17.79.
soit une fonction strictement convexe f. Soient $\alpha, \beta > 0$ tels que $\alpha + \beta = 1$. Nous avons

EQooDKKUooSHLfwv
$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) \tag{17.308}$$

si et seulement si $u = v$.

Démonstration. Tant que $u \neq v$, la définition (17.307) est une inégalité stricte. Donc si (17.308) est valide, alors $u = v$.

Dans le sens inverse, si $u = v$, alors $\alpha u + \beta v = u$, et $\alpha f(u) + \beta f(v) = f(u)$. □

17.12.1 Inégalité des pentes

Dans l'étude des fonctions convexes nous allons souvent utiliser la fonction **taux d'accroissement** qui est, pour α dans le domaine de convexité de f définie par

EqRYBazWd
$$\begin{aligned} \tau_\alpha \colon I\backslash\{\alpha\} &\to \mathbb{R} \\ x &\mapsto \frac{f(x) - f(\alpha)}{x - \alpha}. \end{aligned} \tag{17.309}$$

PropMDMGjGO

Proposition 17.80 (Inégalité des pentes[519])**.**
Soit f une fonction convexe sur un intervalle $I \subset \mathbb{R}$. Alors pour tout $a < b < c$ dans I nous avons[41]

$$\frac{f(b) - f(a)}{b - a} \leqslant \frac{f(c) - f(a)}{c - a} \leqslant \frac{f(c) - f(b)}{c - b}. \tag{17.310}$$

En d'autres termes,

$$\tau_a(b) \leqslant \tau_a(c) \leqslant \tau_b(c), \tag{17.311}$$

c'est-à-dire que τ est croissante en ses deux arguments.

Démonstration. D'abord les inégalités $a < b < c$ impliquent $0 < b - a < c - a$ et donc

$$\lambda = \frac{b - a}{c - a} < 1. \tag{17.312}$$

L'astuce est de remarquer que $(1 - \lambda)a + \lambda c = b$. Donc λ a toutes les bonnes propriétés pour être utilisé dans la définition de la convexité :

$$f\big((1 - \lambda)a + \lambda c\big) \leqslant \lambda f(c) + (1 - \lambda) f(a), \tag{17.313}$$

c'est-à-dire

$$f(b) - f(a) \leqslant \lambda\big(f(c) - f(a)\big) \tag{17.314}$$

ou encore, en remplaçant λ par sa valeur :

$$\frac{f(b) - f(a)}{b - a} \leqslant \frac{f(c) - f(a)}{c - a}. \tag{17.315}$$

Cela fait déjà une des inégalités à savoir.

D'autre part en partant de $-a > -b < c$ nous posons

$$0 < \lambda = \frac{c - b}{c - a}. \tag{17.316}$$

Nous avons à nouveau $b = (1 - \lambda)c + \lambda a$ et nous pouvons obtenir la seconde inégalité

$$\frac{f(c) - f(a)}{c - a} \leqslant \frac{f(c) - f(b)}{c - b}. \tag{17.317}$$

□

41. Les inégalités sont strictes si la fonction f est strictement convexe.

Géométriquement, l'inégalité des pentes se comprend facilement : le coefficient angulaire de la corde du graphe augmente. Donc si $x < y < z$, le coefficient moyen entre x et y est plus petit que celui entre x et z qui est plus petit que celui entre y et z.

Donc si le coefficient angulaire moyen entre a et $b+u$ vaut celui entre a et b, ce coefficient ne peut qu'être constant entra a et b : sinon il serait plus grand entre b et $b+u$ et la moyenne sur $a \to b+u$ serait plus grande que sa moyenne sur $a \to b$. Mais avoir un coefficient angulaire constant signifie être une droite.

En résumé, si une fonction est convexe et non strictement convexe, alors son graphe est une droite. C'est en gros cela que la proposition 17.88 clarifiera.

17.12.2 Convexité et régularité

LemKLTsHIQ

Lemme 17.81 ([518]).
Une fonction convexe sur un ouvert de $\mathbb{R}$

(1) y admet des dérivées à gauche et à droite en chaque point,

(2) y est continue.

Démonstration. Si A est l'ouvert sur lequel la fonction f est convexe, pour chaque point x de A, il existe une boule $B(x,r) \subset A$. Pour $0 < r' < r$ nous avons aussi $\overline{B(0,r')} \subset A$. Bref, nous pouvons supposer que f est définie et convexe sur un intervalle $I = [a,b]$.

Soit $\alpha \in I$. Nous allons prouver que f est continue en α. Nous considérons τ_α le taux d'accroissement défini par (17.309) ; c'est une fonction croissante comme précisé dans l'inégalité des trois pentes 17.80 et de plus $\tau_\alpha(x)$ est bornée supérieurement par $\tau_\alpha(b)$ pour $x < \alpha$ et inférieurement par $\tau_\alpha(a)$ pour $x > \alpha$. Les limites existent donc et sont finies par la proposition 12.56. Autrement dit les limites

$$\lim_{x\to\alpha+} \frac{f(x)-f(\alpha)}{x-\alpha} = \lim_{x\to\alpha^+} \tau_\alpha(x) = \inf_{t>\alpha} \tau_\alpha(t) \tag{17.318a}$$

$$\lim_{x\to\alpha^-} \frac{f(x)-f(\alpha)}{x-\alpha} = \lim_{x\to\alpha^-} \tau_\alpha(x) = \sup_{t<\alpha} \tau_\alpha(t). \tag{17.318b}$$

existent et sont finies, c'est-à-dire que la fonction f admet une dérivée à gauche et à droite.

Pour tout x nous avons les inégalités

$$\tau_\alpha(a) \leqslant \frac{f(x)-f(\alpha)}{x-\alpha} \leqslant \tau_\alpha(b). \tag{17.319}$$

En posant $k = \max\{\tau_\alpha(a), \tau_\alpha(b)\}$ nous avons

$$\big|f(x)-f(\alpha)\big| \leqslant k|x-\alpha|. \tag{17.320}$$

La fonction est donc Lipschitzienne et par conséquent continue par la proposition 12.318. □

Remarque 17.82.
Les dérivées à gauche et à droite ne sont à priori pas égales. Penser par exemple à une fonction affine par morceaux dont les pentes augmentent à chaque morceau.

17.12.3 Dérivées d'une fonction convexe

PropYKwTDPX

Proposition 17.83 ([520, 521, 1]).
Une fonction dérivable sur un intervalle I de $\mathbb{R}$

ITEMooUTSAooJvhZNm

(1) est convexe si et seulement si sa dérivée est croissante sur I.

ITEMooLLSIooFwkxtV

(2) est strictement convexe si et seulement si sa dérivée est strictement croissante sur I

Démonstration. Pour la preuve de (1) et (2), nous allons démontrer les énoncés « non stricts » et indiquer ce qu'il faut changer pour obtenir les énoncés « stricts ».

(1) **Sens direct** Nous supposons que f est convexe. Soient $a < b$ dans I et $x \in]a,b[$. D'après l'inégalité des pentes 17.80,

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x}. \tag{17.321}$$

En faisant la limite $x \to a$ nous avons

$$f'(a) \leqslant \frac{f(b)-f(a)}{b-a} \tag{17.322}$$

et la limite $x \to b$ donne

$$\frac{f(b)-f(a)}{b-a} \leqslant f'(b). \tag{17.323}$$

Ici les inégalités sont à priori non-strictes, même si f est strictement convexe : même avec des inégalités strictes dans (17.321), le passage à la limite rend l'inégalité non stricte. Quoi qu'il en soit nous avons

$$f'(a) \leqslant f'(b). \tag{17.324}$$

(2) **Sens direct : strict** Nous savons déjà que f' est croissante. Si (17.324) était une égalité, alors f' serait constante sur $]a,b[$ parce qu'en prenant c entre a et b nous aurions $f'(a) \leqslant f'(c) \leqslant f'(b)$ avec $f'(a) = f'(b)$. Donc $f'(a) = f'(c)$. Avoir f' constante sur un intervalle est contraire à la stricte convexité.

(3) **Sens réciproque**
Nous supposons que f' est croissante et nous considérons $a < b$ dans I ainsi que $\lambda \in [0,1]$. Nous posons $x = \lambda a + (1-\lambda)b$, et nous savons que $a \leqslant x \leqslant b$. Le théorème des accroissements finis 12.192 donne $c_1 \in]a,x[$ et $c_2 \in]x,b[$ tels que

$$f'(c_1) = \frac{f(x)-f(a)}{x-a} \tag{17.325}$$

et

$$f'(c_2) = \frac{f(b)-f(x)}{b-x}. \tag{17.326}$$

Et en plus $c_1 < c_2$. Vu que f' est croissante nous avons $f'(c_1) \leqslant f'(c_2)$ et donc

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(x)}{b-x}. \tag{17.327}$$

En remplaçant x par sa valeur en termes de λ, a et b nous avons $x-a = (1-\lambda)(b-a)$ et $b-x = \lambda(b-a)$, et l'inégalité (17.327) nous donne

$$f(x) \leqslant \lambda f(a) + (1-\lambda) f(b). \tag{17.328}$$

(4) **Sens réciproque : strict** Si f' est strictement croissante, nous avons $f'(c_1) < f'(c_2)$ et les inégalité suivantes sont strictes, ce qui donne

$$f(x) < \lambda f(a) + (1-\lambda) f(b). \tag{17.329}$$

□

Théorème 17.84 ([520]).
Soit une fonction f de classe C^2.

(1) Est convexe si et seulement si f'' est positive.

(2) Si f'' est strictement positive, elle est strictement convexe.

Démonstration. En deux parties.

(1) **Pour (1)** La fonction est C^2, donc f'' est positive si et seulement si f' est croissante (proposition 12.185) alors que la proposition 17.83 nous jure que f sera convexe si et seulement si f' est croissante.

(2) **Pour (2)** Si f'' est strictement positive, f' sera strictement croissante et donc f strictement convexe (proposition 17.83).

□

REMooVRPQooIybxmp

Remarque 17.85.
Une fonction peut être strictement convexe sans que sa dérivée seconde ne soit toujours strictement positive. En exemple : $x \mapsto x^4$ est strictement convexe alors que sa dérivée seconde s'annule en zéro.

ExPDRooZCtkOz

Exemple 17.86.
Quelques exemples utilisant le théorème 17.84

(1) La fonction $x \mapsto x^2$ est convexe parce que sa dérivée seconde est la constante (positive) 2.

(2) La fonction $x \mapsto \frac{1}{x}$ est convexe sur $\mathbb{R}^+\backslash\{0\}$ (sa dérivée seconde est $2x^{-3}$).

ITEMooRXSBooDBerbx

(3) La fonction exponentielle est strictement convexe par le théorème 17.84.

(4) La fonction ln est concave parce que la dérivée seconde de $-\ln$ est $\frac{1}{x^2}$ qui est strictement positif.

△

Nous en démontrons une en détail ; elle sera utile en analyse fonctionnelle, lors de l'étude des espaces L^p. Voir par exemple le théorème de la projection 27.157.

LEMooSXTXooZOmtKq

Lemme 17.87.
Soient $p > 1$ et la fonction

$$\begin{aligned} f :]0, \infty[&\to \mathbb{R} \\ x &\mapsto x^p \end{aligned} \tag{17.330}$$

est strictement convexe.

Démonstration. La proposition 14.294 nous permet de dire que la fonction f est de classe C^∞ et que la dérivée seconde est donnée par

$$f''(x) = p(p-1)x^{p-2}. \tag{17.331}$$

Cela est strictement positif pour tous les x considérés, le théorème 17.84 conclut. □

17.12.4 Graphe d'une fonction convexe

L'idée principale du graphe d'une fonction convexe est qu'il est toujours au dessus du graphe de ses tangentes (lorsqu'elles existent). Lorsqu'elles n'existent pas, le lemme 17.81 donne des coefficients directeurs de droites qui vont rester en dessous du graphe de la fonction.

PROPooOCOEooEGybmS

Proposition 17.88 ([522]).
Une fonction convexe est strictement convexe si et seulement si il n'existe aucun intervalle de longueur non nulle sur lequel elle coïncide avec une fonction affine.

Démonstration. Si sur l'intervalle (non réduit à un point) $[x, y]$, la fonction convexe f coïncide avec une fonction affine, alors $f(t) = at + b$ et pour $\lambda \in]0, 1[$ nous avons

$$f\big(\lambda x + (1-\lambda)y\big) = a\lambda x + a(1-\lambda)y + b = \lambda f(x) + (1-\lambda)f(y) \tag{17.332}$$

où nous avons remplacé b par $\lambda b + (1-\lambda)b$. Par conséquent la fonction n'est pas strictement convexe.

Nous supposons maintenant que la fonction convexe f n'est pas strictement convexe sur l'intervalle I. Il existe $x \neq y \in I$ et $\lambda \in]0,1[$ tels que

$$f\big(\lambda x + (1-\lambda)y\big) = \lambda f(x) + (1-\lambda)f(y). \tag{17.333}$$

Nous posons $z = \lambda x + (1-\lambda)y$ et $u \in]x,z[$ pour écrire des inégalités des pentes entre $x < u < z < y$. Plus précisément si nous notons $a \to b$ la pente de a à b, c'est-à-dire $a \to b = \frac{f(b)-f(a)}{b-a}$, alors les inégalités des pentes pour $x < u < z$ puis $u < z < y$ donnent

$$x \to z \leqslant u \to z \leqslant z \to y. \tag{17.334}$$

EqooBMEEooMnoEzd

Voyons maintenant qu'en réalité $z \to y = x \to z$. En effet en replaçant

$$f(y) = \frac{f(z) - \lambda f(x)}{1-\lambda} \tag{17.335}$$

et

$$y = \frac{\lambda x}{1-\lambda} \tag{17.336}$$

dans l'expression $z \to y = \frac{f(y)-f(z)}{y-z}$ nous obtenons

$$z \to y = \frac{f(y)-f(z)}{y-z} = \frac{f(z)-f(x)}{z-x} = x \to z. \tag{17.337}$$

Les inégalités (17.334) sont donc des égalités :

$$\frac{f(z)-f(x)}{z-x} = \frac{f(z)-f(u)}{z-u} = \frac{f(y)-f(z)}{y-z}. \tag{17.338}$$

Nous avons donc montré que le nombre $a = \frac{f(z)-f(u)}{z-u}$ ne dépend pas de u. Nous avons alors

$$f(z) - f(u) = a(z-u) \tag{17.339}$$

ou encore :

$$f(u) = f(z) - a(z-u), \tag{17.340}$$

ce qui signifie que sur $]x,z[$, la fonction f est affine. □

PROPooQPOSooDZlUAJ

Proposition 17.89.
Une fonction dérivable sur un intervalle I de $\mathbb{R}$ est convexe si et seulement si son graphe est au dessus de chacune de ses tangentes.

Démonstration. En deux parties.

(1) **Sens direct** Soient $x, y \in I$. Nous voulons :

$$f(y) \geqslant f(x) + f'(x)(y-x). \tag{17.341}$$

Étant donné que nous aurons besoin, dans le quotient différentiel de quelque chose comme $f(x+t) - f(x)$ nous écrivons la définition (17.306) de la convexité en inversant les rôles de x et y et en manipulant un peu :

$$f\big(ty + (1-t)x\big) \leqslant tf(y) + (1-t)f(x) \tag{17.342a}$$
$$f\big(x + t(y-x)\big) \leqslant tf(y) + (1-t)f(x) \tag{17.342b}$$
$$f\big(x + t(y-x)\big) - f(x) \leqslant tf(y) - tf(x) \tag{17.342c}$$

Nous divisons par t :

$$\frac{f\big(x + t(y-x)\big) - f(x)}{t} \leqslant f(y) - f(x). \tag{17.343}$$

Le passage à la limite $t \to 0$ donne

$$(y-x)f'(x) \leqslant f(y) - f(x), \tag{17.344}$$

ce qu'il fallait.

(2) **Sens inverse** Pour tout $x, y \in I$ nous supposons avoir

$$f(y) \geqslant f(x) + f'(x)(y - x). \tag{17.345}$$

Si nous supposons $x \neq y$ et si nous posons $z = \lambda x + (1 - \lambda)y$ nous voulons prouver que

$$f(z) \leqslant \lambda f(x) + (1 - \lambda)f(y). \tag{17.346}$$

Pour cela nous écrivons l'inégalité (17.345) avec les couples (x, z) et (y, z) :

$$f(x) \geqslant f(z) + f'(z)'(x - z) \tag{17.347a}$$
$$f(y) \geqslant f(z) + f'(z)'(y - z) \tag{17.347b}$$

En multipliant la première par λ et la seconde par $(1 - \lambda)$ et en sommant,

$$\lambda f(x) + (1 - \lambda)f(y) \geqslant \lambda f(z) + \lambda f'(z)(x - z) + (1 - \lambda)f(z) + (1 - \lambda)f'(z)(y - z) \tag{17.348a}$$
$$= f(z) + f'(z)\big(\lambda(x - z) + (1 - \lambda)(y - z)\big) \tag{17.348b}$$
$$= f(z). \tag{17.348c}$$

□

Proposition 17.90 ([1])**.**
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction convexe et $a \in \mathbb{R}$. Il existe une constante $c_a \in \mathbb{R}$ telle que pour tout x nous ayons

$$f(x) - f(a) \geqslant c_a(x - a). \tag{17.349}$$

Autrement dit, le graphe de la fonction f est toujours au dessus de la droite d'équation

$$y = f(a) + c_a(x - a). \tag{17.350}$$

Démonstration. Les dérivées à gauche et à droite de f données par le lemme 17.81 sont les candidats tout désignés pour être coefficient directeur de la droite que l'on cherche. Nous allons prouver qu'en posant

$$c_a = \inf_{t>a} \tau_a(t), \tag{17.351}$$

la droite $y = f(a) + c_a(x - a)$ répond à la question [42].

Nous devons prouver que le nombre $\Delta_x = f(x) - \big(f(a) + c_a(x - a)\big)$ est positif pour tout x.

(1) **Si $x > a$**
Nous divisons par $x - a$ et nous devons prouver que $\frac{\Delta_x}{x-a}$ est positif :

$$\frac{\Delta_x}{x - a} = \frac{f(x) - f(a)}{x - a} - c_a \tag{17.352a}$$
$$= \tau_a(x) - \inf_{t>a} \tau_a(t) \tag{17.352b}$$
$$\geqslant 0 \tag{17.352c}$$

parce que $t \to \tau_a(t)$ est croissante et que $x > a$.

(2) **Si $x < a$**
Nous divisons par $x - a$ et nous devons prouver que $\frac{\Delta_x}{x-a}$ est négatif :

$$\frac{\Delta_x}{x - a} = \frac{f(x) - f(a)}{x - a} - c_a \tag{17.353a}$$
$$= \tau_a(x) - \inf_{t>a} \tau_a(t) \tag{17.353b}$$
$$\leqslant 0 \tag{17.353c}$$

parce que $t \to \tau_a(t)$ est croissante et que $x < a$.

42. En prenant l'autre, $c'_a = \sup_{t<a} \tau_a(t)$, ça fonctionne aussi. En pensant à une fonction affine par morceaux, on remarque qu'en choisissant un nombre entre les deux, nous avons plus facilement une inégalité stricte dans (17.349).

□
PropPEJCgCH

Proposition 17.91 ([1]).
Si g est une fonction convexe, il existe deux suites réelles (a_n) et (b_n) telles que

$$g(x) = \sup_{n\in\mathbb{N}}(a_n x + b_n). \tag{17.354}$$

Démonstration. Pour $u \in \mathbb{R}$ nous considérons $a(u)$ et $b(u)$ tels que la droite $y(x) = a(u)x + b(u)$ vérifie $y(u) = g(u)$ et $y(x) \leqslant g(x)$ pour tout x. Cela est possible par la proposition 17.90. Il s'agit d'une droite coupant le graphe de g en $x = u$ et restant en dessous. Nous considérons alors (u_n) une suite quelconque dense dans $\mathbb{R}$ (disons les rationnels pour fixer les idées) et nous posons

$$\begin{cases} a_n = a(u_n) & \text{(17.355a)}\\ b_n = b(u_n). & \text{(17.355b)} \end{cases}$$

Si $q \in \mathbb{Q}$ alors $a_n x + b_n \leqslant g(x)$ pour tout n et $g(q)$ est le supremum qui est atteint pour le n tel que $u_n = q$. Si maintenant x n'est pas dans $\mathbb{Q}$ il faut travailler plus.

Nous prenons $(\tilde{q}_n)$, une sous-suite de (q_n) convergeant vers x et N suffisamment grand pour que pour tout $n \geqslant N$ on ait $|\tilde{q}_n - x| \leqslant \epsilon$ et $|g(\tilde{q}_n) - g(x)| \leqslant \epsilon$; cela est possible grâce à la continuité de g (lemme 17.81). Ensuite les sous-suites $(\tilde{a}_n)$ et $(\tilde{b}_n)$ sont celles qui correspondent :

$$\tilde{a}_n\tilde{q}_n + \tilde{b}_n = g(\tilde{q}_n). \tag{17.356}$$

Nous considérons la majoration

$$\begin{aligned} |\tilde{a}_n x + \tilde{b}_n - g(x)| &\leqslant |\tilde{a}_n x + \tilde{b}_n - (\tilde{a}_n\tilde{q}_n + \tilde{b}_n)| + \underbrace{|\tilde{a}_n\tilde{q}_n + \tilde{b}_n - g(\tilde{q}_n)|}_{=0} + \underbrace{|g(\tilde{q}_n) - g(x)|}_{\leqslant\epsilon} && \text{(17.357a)}\\ &\leqslant |\tilde{a}_n||x - \tilde{q}_n| + \epsilon && \text{(17.357b)}\\ &= \epsilon\big(|\tilde{a}_n| + 1\big). && \text{(17.357c)} \end{aligned}$$

Il nous reste à montrer que $|\tilde{a}_n|$ est borné par un nombre ne dépendant pas de n (pour les $n > N$).

Comme la droite de coefficient directeur $\tilde{a}_n$ et passant par le point $\big(\tilde{q}_n, g(\tilde{q}_n)\big)$ reste en dessous du graphe de g, nous avons pour tout n et tout $y \in \mathbb{R}$ l'inégalité

$$g(y) \geqslant \tilde{a}_n(y - \tilde{q}_n) + g(\tilde{q}_n) \in \tilde{a}_n B(y - x, \epsilon) + B\big(g(x), \epsilon\big). \tag{17.358}$$

Si $\tilde{a}_n$ n'est pas borné vers le haut, nous prenons y tel que $B(y - x, \epsilon)$ soit minoré par un nombre k strictement positif et nous obtenons

$$g(y) \geqslant k\tilde{a}_n + l \tag{17.359}$$

avec k et l indépendants de n. Cela donne $g(y) = \infty$. Si au contraire $\tilde{a}_n$ n'est pas borné vers le bas, nous prenons y tel que $B(y - x, \epsilon)$ est majoré par un nombre k strictement négatif. Nous obtenons encore $g(y) = \infty$.

Nous concluons que $|\tilde{a}_n|$ est bornée. □
LemXOUooQsigHs

Lemme 17.92 ([104]).
L'application

$$\begin{aligned} \phi\colon S^{++}(n,\mathbb{R}) &\to \mathbb{R}\\ A &\mapsto \det(A) \end{aligned} \tag{17.360}$$

est ***log-convave****, c'est-à-dire que l'application $\ln\circ\phi$ est concave*[43]*. De façon équivalente, si $A, B \in S^{++}$ et si $\alpha + \beta = 1$, alors*

$$\det(\alpha A + \beta B) \geqslant \det(A)^\alpha \det(B)^\beta. \tag{17.361}$$
EqSPKooHFZymR

43. La définition 15.83 du logarithme ne fonctionne que pour les réels strictement positifs. C'est le cas du déterminant d'une matrice réelle symétrique strictement définie positive.

Ici S^{++} est l'ensemble des matrices symétriques strictement définies positives, définition 8.230.

Démonstration. En plusieurs étapes.

(1) **Pseudo-réduction** Le théorème de pseudo-réduction simultanée, corolaire 10.40, appliqué aux matrices A et B nous donne une matrice inversible Q telle que

$$\begin{cases} B = Q^tDQ & (17.362a)\\ A = Q^tQ & (17.362b)\end{cases}$$

avec

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \tag{17.363}$$

$\lambda_i > 0$. Nous avons alors

$$\det(A)^\alpha \det(B)^\beta = \det(Q)^{2\alpha} \det(Q)^{2\beta} \det(D)^\beta = \det(Q)^2 \det(D)^\beta \tag{17.364}$$

(parce que $\alpha + \beta = 1$) et

$$\begin{aligned} \det(\alpha A + \beta B) &= \det(\alpha Q^tQ + \beta Q^tDQ) & (17.365a)\\ &= \det\big(Q^t(\alpha\mathbb{1} + \beta D)Q\big) & (17.365b)\\ &= \det(Q)^2 \det(\alpha\mathbb{1} + \beta D). & (17.365c)\end{aligned}$$

(2) **Ré-expression** L'inégalité (17.361) qu'il nous faut prouver se réduit donc à

$$\det(\alpha\mathbb{1} + \beta D) \geqslant \det(D)^\beta. \tag{17.366}$$

Vue la forme de D nous avons

$$\det(\alpha\mathbb{1} + \beta D) = \prod_{i=1}^{n}(\alpha + \beta\lambda_i) \tag{17.367}$$

et

$$\det(D)^\beta = \big(\prod_{i=1}^{n}\lambda_i\big)^\beta. \tag{17.368}$$

Il faut donc prouver que

$$\prod_{i=1}^{n}(\alpha + \beta\lambda_i) \geqslant \big(\prod_{i=1}^{n}\lambda_i\big)^\beta. \tag{17.369}$$

EqGFLogQElciS

Cette dernière égalité de produit sera prouvée en passant au logarithme.

(3) **Logarithme** Puisque le logarithme est concave par l'exemple 17.86, nous avons pour chaque i que

$$\ln(\alpha + \beta\lambda_i) \geqslant \alpha\ln(1) + \beta\ln(\lambda_i) = \beta\ln(\lambda_i). \tag{17.370}$$

En sommant cela sur i et en utilisant les propriétés de croissance et de multiplicativité du logarithme nous obtenons successivement

$$\begin{aligned} \sum_{i=1}^{n}\ln(\alpha + \beta\lambda_i) &\geqslant \beta\sum_i\ln(\lambda_i) & (17.371a)\\ \ln\big(\prod_i(\alpha + \beta\lambda_i)\big) &\geqslant \ln\Big(\big(\prod_i\lambda_i\big)^\beta\Big) & (17.371b)\\ \prod_i(\alpha + \beta\lambda_i) &\geqslant \big(\prod_i\lambda_i\big)^\beta, & (17.371c)\end{aligned}$$

ce qui est bien (17.369).

□

Rappel de notations : $\mathbb{R}^+ = [0, \infty[$. Voir la remarque 1.428.

LEMooNUDOooVfVPkw

Lemme 17.93 ([1]).
Soit une fonction strictement convexe $g\colon \mathbb{R}^+ \to \mathbb{R}^+$. Soit une fonction $f\colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ vérifiant

(1) $f(0,0) = 0$,

(2) $f(tx, ty) = tf(x, y)$ *pour tout t (tant que ça ne déborde pas du domaine)*

(3) $f(1, y) = g(y)$ *pour tout y*

(4) $f(0, y) = f(y, 0)$.

Alors f est convexe.

Démonstration. Nous devons prouver que pour toute paire de points A, B sur le graphe de f, le segment $[A, B]$ est au-dessus du graphe de f. Ledit graphe étant d'ailleurs constitué de droites joignant $(0, 0, 0)$ et les points du graphe de g (situé en $x = 1$).

Nous notons $\mathcal{C}$ le graphe de f.

(1) **Une corde alignée à O** Soient deux point A et B alignés à l'origine O. Un point quelconque de $[A, B]$ (et même de toute la droite) s'écrit

$$\big(tA_x, tA_y, tf(A_x, A_y)\big) = \big(tA_x, tA_y, f(tA_x, tA_y)\big), \tag{17.372}$$

et donc est sur le graphe de f.

(2) **Autre corde** Nous prouvons que $[A, B] \cap \mathcal{C} = \{A, B\}$.
Si A et B sont dans le plan $x = 1$ alors c'est d'accord parce que le graphe de f dans le plan est le même que celui de la fonction strictement convexe g.
Si A et B ne sont pas alignés à O et si ils ne sont pas dans le plan $x = 1$ alors le plan AOB coupe le plan $x = 1$ en une droite.
Nous supposons l'existence d'un point $C \in]A, B[\cap \mathcal{C}$.
Nous considérons la droite (OA) qui est contenue dans ce plan et dans $\mathcal{C}$ (au moins la partie positive) et nous notons A' son intersection avec le plan $x = 1$. Même chose pour B et C qui donnent B' et C'.
Cela nous donne des points A', B' et C' qui sont alignés dans le graphe de f en $x = 1$. Or le graphe de f en $x = 1$ est le graphe de la fonction g qui est strictement convexe et qui ne contient donc pas de points alignés.
Nous en concluons que si $A, B \in \mathcal{C}$ alors $]A, B[$ est soit complètement strictement au-dessus de $\mathcal{C}$ soit complètement strictement en-dessous de $\mathcal{C}$.

Nous prouvons à présent que toutes les cordes sont au-dessus de $\mathcal{C}$. Pour cela, soient $A, B \in]0, \infty[\times]0, \infty[$, deux points non alignés à $O = (0, 0)$. Nous considérons les points A', B' qui sont les intersections entre les droites (AO) et (BO) et la droite $x = 1$ ainsi que le chemin σ qui parcourt le segment $[A, A']$ et le chemin γ qui parcourt le segment $[B, B']$:

$$\begin{aligned} \sigma(0) &= A, \quad \gamma(0) = B, \\ \sigma(1) &= A', \quad \gamma(1) = B'. \end{aligned} \tag{17.373}$$

Pour tout u, la seule droite passant par O et par $\sigma(u)$ passe également par A, et pas par B. En conséquence de quoi, pour tout $u_1, u_2 \in [0, 1]$, la droite $\big(\sigma(u_1)\gamma(u_2)\big)$ ne passe pas par $(0, 0)$.

Nous considérons à présent non seulement la corde joignant $\big(A, f(A)\big)$ à $\big(B, f(B)\big)$ et la corde joignant $\big(A', f(A')\big)$ à $\big(B', f(B')\big)$ mais également toutes les cordes intermédiaires (si vous aimez les grands mots, vous pouvez parler d'homotopie) :

$$c(u, t) = t\Big(\sigma(u), f\big(\sigma(u)\big)\Big) + (1 - t)\Big(\gamma(u), f\big(\gamma(u)\big)\Big) \tag{17.374}$$

Pour chaque $u \in [0,1]$, cela représente une corde entre deux points non alignés à $(0,0,0)$ et donc une corde qui est soit strictement au-dessus de $\mathcal{C}$ soit strictement en-dessous (à par les points correspondant à $t=0$ et $t=1$ qui, eux, sont sur $\mathcal{C}$).

Soit $t_0 \in]0,1[$. La courbe $c(u,t_0)$ avec $u \in [0,1]$ ne touche jamais $\mathcal{C}$. Or le point $c(1,t_0)$ est au-dessus de $\mathcal{C}$, donc le point $c(0,t_0)$ est également au-dessus de $\mathcal{C}$.

Nous en concluons que toutes les cordes entre $(A,f(A))$ et $(B,f(B))$ est située au-dessus de $\mathcal{C}$ et non en-dessous de $\mathcal{C}$. □

17.12.5 Convexité et hessienne

DEFooKCFPooLwKAsS

Définition 17.94.
Soit une partie convexe U de $\mathbb{R}^n$ et une fonction $f\colon U \to \mathbb{R}$.

*(1) La fonction f est **convexe** si pour tout $x,y \in U$ avec $x \neq y$ et pour tout $\theta \in]0,1[$ nous avons*

$$f\big(\theta x + (1-\theta)y\big) \leqslant \theta f(x) + (1-\theta)f(y). \tag{17.375}$$

*(2) Elle est **strictement convexe** si nous avons l'inégalité stricte.*

PROPooYNNHooSHLvHp

Proposition 17.95 ([517]).
Soit Ω ouvert dans $\mathbb{R}^n$ et U convexe dans Ω, et une fonction différentiable $f\colon U \to \mathbb{R}$.

ITEMooRVIVooIayuPS

(1) La fonction f est convexe sur U si et seulement si pour tout $x,y \in U$,

$$f(y) \geqslant f(x) + df_x(y-x). \tag{17.376}$$

ITEMooCWEWooFtNnKl

(2) La fonction f est strictement convexe sur U si et seulement si pour tout $x,y \in U$ avec $x \neq y$,

$$f(y) > f(x) + df_x(y-x). \tag{17.377}$$

Démonstration. Nous avons quatre petites choses à démontrer.

(1) **(1) sens direct** Soit une fonction convexe f. Nous avons :

$$f\big((1-\theta)x + \theta y\big) \leqslant (1-\theta)f(x) + \theta f(y), \tag{17.378}$$

donc

$$f\big(x + \theta(y-x)\big) - f(x) \leqslant \theta\big(f(y) - f(x)\big) \tag{17.379}$$

Puisque $\theta > 0$ nous pouvons diviser par θ sans changer le sens de l'inégalité :

$$\frac{f\big(x + \theta(y-x)\big) - f(x)}{\theta} \leqslant f(y) - f(x). \tag{17.380}$$

EQooAXXFooHWtiJh

Nous prenons la limite $\theta \to 0^+$. Cette limite est égale à a limite simple $\theta \to 0$ et vaut (parce que f est différentiable) :

$$\frac{\partial f}{\partial(y-x)}(x) \leqslant f(y) - f(x), \tag{17.381}$$

et aussi

$$df_x(y-x) \leqslant f(y) - f(x) \tag{17.382}$$

par le lemme 12.255.

(2) **(1) sens inverse** Pour tout $a \neq b$ dans U nous avons

$$f(b) \geqslant f(a) + df_a(b-a). \tag{17.383}$$

EQooEALSooJQszWr

Pour $x \neq y$ dans U et pour $\theta \in]0,1[$ nous écrivons (17.383) pour les couples $\big(\theta x + (1-\theta)y, y\big)$ et $\big(\theta x + (1-\theta)y, x\big)$. Ça donne :

$$f(y) \geqslant f\big(\theta x + (1-\theta)y\big) + df_{\theta x + (1-\theta)y}\big(\theta(y-x)\big), \tag{17.384}$$

et

$$f(x) \geqslant f\big(\theta x + (1-\theta)y\big) + df_{\theta x+(1-\theta)y}\big((1-\theta)(x-y)\big). \tag{17.385}$$

La différentielle est linéaire ; en multipliant la première par $(1-\theta)$ et la seconde par θ et en la somme, les termes en df se simplifient et nous trouvons

$$\theta f(x) + (1-\theta)f(y) \geqslant f\big(\theta x + (1-\theta)y\big). \tag{17.386}$$

(3) **(2) sens direct** Nous avons encore l'équation (17.380), avec une inégalité stricte. Par contre, ça ne va pas être suffisant parce que le passage à la limite ne conserve pas les inégalités strictes. Nous devons donc être plus malins.

Soient $0 < \theta < \omega < 1$. Nous avons $(1-\theta)x + \theta y \in [x, (1-\omega)x + \omega y]$, donc nous pouvons écrire $(1-\theta)x + \theta y$ sous la forme $(1-s)x + s\big((1-\omega)x + \omega y\big)$. Il se fait que c'est bon pour $s = \theta/\omega$ (et aussi que nous avons $\theta/\omega < 1$). Donc nous avons

$$f\big((1-\theta)x + \theta y\big) = f\Big((1-\frac{\theta}{\omega})x + \frac{\theta}{\omega}\big((1-\omega)x + \omega y\big)\Big) \tag{17.387a}$$

$$< (1-\frac{\theta}{\omega})f(x) + \frac{\theta}{\omega}f\big((1-\omega)x + \omega y\big). \tag{17.387b}$$

Cela nous permet d'écrire

$$\frac{f\big((1-\theta)x + \theta y\big) - f(x)}{\theta} < \frac{f\big((1-\omega)x + \omega y\big)}{\omega} < f(y) - f(x). \tag{17.388}$$

Le seconde inégalité est le pendant de (17.380). Maintenant en passant à la limite pour θ nous conservons une inégalité stricte par rapport à $f(y) - f(x)$:

$$df_x(y-x) < f(y) - f(x). \tag{17.389}$$

□

Avant de lire la proposition suivante, il faut relire la proposition 12.348 et ce qui s'y rapporte. Lire aussi la remarque 17.85 qui indique qu'il n'y a pas de réciproque dans l'énoncé (2).

REMooYCRKooEQNIkC

Remarque 17.96.
Notons que la condition (17.390) n'est pas équivalente à demander $(d^2f)_x(h,h) \geqslant 0$ pour tout h. En effet nous ne demandons la positivité que dans les directions atteignables comme différence de deux éléments de U. La partie U n'est pas spécialement ouverte ; elle pourrait n'être qu'une droite dans $\mathbb{R}^3$. Dans ce cas, demander que f (qui est C^2 sur l'ouvert Ω) soit convexe sur U ne demande que la positivité de $(d^2f)_x$ appliqué à des vecteurs situés sur la droite U.

PROPooBMIRooFkQSAb

Proposition 17.97 ([517])**.**
Soit une fonction $f\colon \Omega \to \mathbb{R}$ deux fois différentiable sur l'ouvert Ω de $\mathbb{R}^n$ et un convexe $U \subset \Omega$.

ITEMooZQCAooIFjHOn

(1) La fonction f est convexe sur U si et seulement si

$$(d^2f)_x(y-x, y-x) \geqslant 0 \tag{17.390}$$

EQooIBDCooJYdiBb

pour tout $x, y \in U$.

ITEMooHAGQooYZyhQk

(2) Si pour tout $x \neq y$ dans U nous avons

$$(d^2f)_x(y-x, y-x) > 0 \tag{17.391}$$

alors la fonction f est strictement convexe sur U.

Démonstration. Il y a trois parties à démontrer.

(1) **(1) sens direct**

Soit une fonction convexe f sur U. Soient aussi $x, y \in U$ et $h = y - x$. Nous utilisons ma version préférée de Taylor[44] : celle de la proposition 12.468 :

$$f(x + th) = f(x) + tdf_x(h) + \frac{t^2}{2}(d^2f_x)(h, h) + t^2\|h\|^2\alpha(th) \tag{17.392}$$

avec $\lim_{s\to 0} \alpha(s) = 0$. Le fait que f soit convexe donne

$$0 \leqslant f(x + th) - f(x) - tdf_x(h), \tag{17.393}$$

et donc

$$0 \leqslant \frac{t^2}{2}(d^2f)_x(h, h) + f^2\|h\|^2\alpha(th). \tag{17.394}$$

En multipliant par 2 et en divisant par t^2,

$$0 \leqslant (d^2f)_x(h, h) + 2\|h\|^2\alpha(th). \tag{17.395}$$

En prenant $t \to 0$ nous avons bien $(d^2f)_x(y - x, y - x) \geqslant 0$.

(2) **(1) sens inverse**

Soient $x, y \in U$. Nous écrivons Taylor en version de la proposition 12.469 :

$$f(y) = f(x) + df_x(y - x) + \frac{1}{2}(d^2f)_z(y - x, y - x) \tag{17.396}$$

pour un certain $z \in]x, y[$. En vertu de ce qui a été dit dans la remarque 17.96 nous ne pouvons pas évoquer l'hypothèse (17.390) pour conclure que $(d^2f)_z(y - x, y - x) \geqslant 0$. Il y a deux manières de nous sortir du problème :

— Trouver $s \in U$ tel que $y - x = s - z$.

— Trouver un $z \in U$ tel que $z - x$ soit un multiple de $y - x$.

La première approche ne fonctionne pas parce que $s = y - x + z$ n'est pas garanti d'être dans U ; par exemple avec $x = 1$, $z = 2$, $y = 3$ et $U = [0, 3]$. Dans ce cas $s = 4 \notin U$.

Pour suivre la seconde approche, nous posons $z = \theta x + (1 - \theta)y$, donc $z - x = (1 - \theta)(y - x)$. Dans ce cas la bilinéarité de $(d^2f)_z$ donne[45]

$$f(y) = f(x) + df_x(y - x) + \underbrace{\frac{1}{2}\frac{1}{(1 - \theta)^2}(d^2f)_z(z - x, z - x)}_{\geqslant 0}. \tag{17.397}$$

Nous en déduisons que f est convexe par la proposition 17.95(1).

(3) **(2)**

Le raisonnement que nous venons de faire pour le sens inverse de (1) tient encore, et nous avons

$$f(y) = f(x) + df_x(y - x) + \underbrace{\frac{1}{2}\frac{1}{(1 - \theta)^2}(d^2f)_z(z - x, z - x)}_{>0} \tag{17.398}$$

d'où nous déduisons la stricte convexité de f par la proposition 17.95(2).

□

CORooMBQMooWBAIIH

Corolaire 17.98.

Soit un ouvert Ω de $\mathbb{R}^n$ et une fonction deux fois différentiable f sur Ω.

44. Si vous présentez ceci au jury d'un concours, vous devriez être capable de raconter ce que signifie d^2f, et pourquoi nous l'utilisons comme une 2-forme.

45. Si vous avez bien suivi, la bilinéarité est contenue dans la proposition 12.348.

ITEMooUAFTooXfCviI

(1) *La fonction f est convexe si et seulement si pour tout x, la matrice hessienne d^2f_x est semi-définie positive.*

ITEMooDGISooPlRLOd

(2) *Si pour tout x de Ω, la matrice hessienne d^2f_x est strictement définie positive, alors f est strictement convexe.*

Démonstration. Nous pouvons voir ce résultat comme une conséquence directe de la proposition 17.97 en posant $U = \Omega$. Nous allons cependant en donner une démonstration directe.

Soit $a \in \Omega$ et posons la fonction

$$\begin{aligned} g\colon \Omega &\to \mathbb{R} \\ x &\mapsto f(x) - f(a) - (df)_a(x-a). \end{aligned} \tag{17.399}$$

Nous allons calculer des différentielles de f, et une chose importante à comprendre est que la différentielle de la fonction $x \mapsto df_a(x-a)$ ne fait pas intervenir la différentielle seconde de f ; c'est la différentielle de $a \mapsto df_a(x)$ qui demanderait la différentielle seconde de f. Ici le point a étant donné, df_a est une application linéaire sans histoires. En particulier, $df_a(x-a) = df_a(x) - df_a(a)$.

La fonction g vérifie :

(1) $g(a) = 0$,

(2) $dg_x = df_x - df_a$, parce que la différentielle de $x \mapsto df_a(x)$ est $x \mapsto df_a(x)$ en vertu du lemme 10.273.

(3) $dg_a = 0$. Le point a est un point critique de g.

(4) $d^2g_x = d^2f_x$ parce que la différentielle de $x \mapsto df_a$ est nulle.

Ceci étant dit, nous pouvons commencer avec la preuve.

(1) **(1) sens direct**
Nous supposons que f est convexe. Alors $g(x) \geqslant 0$ pour tout x par la caractérisation 17.95(1). Cela signifie que $x = 0$ est un minimum global de g. Par conséquent la proposition 17.70(3) nous dit que la Hessienne d^2f_a est semi-définie positive.

(2) **(1) sens inverse**
Nous sommes dans le cas de la proposition 17.70(1). Le point $x = a$ est un minimum local de g, ce qui signifie que $g(x) \geqslant 0$ pour tout x de Ω. Encore une fois la caractérisation 17.95(1) nous permet de conclure.

(3) **(2)**
La fonction g vérifie les conditions de 17.70(2), donc $x = 0$ est un minimum local strict de g. La caractérisation 17.95(2) nous fait conclure que f est strictement convexe.

□

17.99.
Nous rappelons que, avec $p > 1$, la fonction

$$\begin{aligned} f\colon [0,\infty[&\to \mathbb{R} \\ x &\mapsto x^p \end{aligned} \tag{17.400}$$

est strictement croissante. La proposition 12.403 est formelle sur ce point.

Lemme 17.100.
Soient un réel a, et $p > 1$. L'équation

$$|x|^p = ax \tag{17.401}$$

possède au plus deux solutions réelles.

Démonstration. En deux parties suivant le signe de a.

(1) **Si $a = 0$** Alors l'équation est $|x|^p = 0$, et l'unique solution est $x = 0$

(2) **Si $a > 0$** Si $x < 0$ alors $ax < 0$ et $|x|^p - ax > |x|^p > 0$. Cela prouve que notre équation n'a pas de solutions $x < 0$ lorsque $a > 0$.
Cherchons donc des solutions avec $x \geqslant 0$. D'abord $|x| = x$ et ensuite, en posant $p = 1 + \delta$ ($\delta > 0$), nous avons la factorisation

$$x^p - ax = x(x^\delta - a) = 0. \tag{17.402}$$

Poser $x = 0$ est clairement une solution. Si $x \neq 0$, nous avons le raisonnement suivant. Les réels formant un corps, c'est un anneau intègre qui vérifie alors la règle du produit nul [46]. Cela pour dire que si $x \neq 0$, alors

$$x^\delta = a. \tag{17.403}$$

La proposition 12.410 nous dit que cela a une unique solution dans les réels positifs.

(3) **Si $a < 0$** Ce cas se traite de façons similaire [47].

□
PROPooLIGIooPrHYlb

Proposition 17.101.
Soit $1 < p < \infty$. La fonction

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \|x\|^p \end{aligned} \tag{17.404}$$

est strictement convexe.

Démonstration. La preuve va se diviser en deux parties. D'abord nous allons utiliser la matrice Hessienne pour démontrer le résultat sur l'ouvert $\mathbb{R}^n\backslash\{0\}$, et ensuite nous allons un peu bricoler pour ajouter 0 au domaine de stricte convexité [48].

(1) **La Hessienne** Nous notons

$$f_p(x) = \|x\|^p = \big(\sum_i x_i^2\big)^{p/2}, \tag{17.405}$$

et nous dérivons :

$$\frac{\partial f_p}{\partial x_i}(x) = \frac{p}{2} 2x_i \big(\sum_i x_i^2\big)^{(p-2)/2} = p x_i f_{p-2}(x). \tag{17.406}$$
EQooNZZWooGeAlyj

Cela n'est déjà pas bien défini en $x = 0$ lorsque $p < 2$, mais qu'importe ? Nous dérivons encore en utilisant entre autres la formule (17.406) elle-même avec $p \to p - 2$:

$$\frac{\partial^2 f_p}{\partial x_j \partial x_i}(x) = p\delta_{ij}\|x\|^{p-2} + p(p-2)x_i x_j \|x\|^{p-4}. \tag{17.407}$$

Nous avons donc la matrice Hessienne

$$H_{ij}(x) = p\delta_{ij}\|x\|^{p-2} + p(p-2)x_i x_j \|x\|^{p-4}. \tag{17.408}$$

Pour prouver que cette matrice est strictement définie positive, nous avons le choix entre la proposition 8.233(1) ou le lemme 8.234(1). Nous utilisons le second. Nous avons [49] :

$$\begin{aligned} y \cdot H(x)y &= \sum_{kl} y_k H(x)_{kl} y_l && \text{(17.409a)} \\ &= p\sum_l y_l^2 \|x\|^{p-2} + p(p-2)\|x\|^{p-4} \sum_{kl} x_k y_k x_l y_l && \text{(17.409b)} \\ &= p\|y\|^2\|x\|^{p-2} + p(p-2)\|x\|^{p-4}(x \cdot y)^2 && \text{(17.409c)} \\ &= p\|x\|^{p-4}\big(\|y\|^2\|x\|^2 + (p-2)(x \cdot y)^2\big) && \text{(17.409d)} \\ &> p\|x\|^{p-4}\big(\|y\|^2\|x\|^2 - (x \cdot y)^2\big) && \text{(17.409e)} \\ &\geqslant 0. && \text{(17.409f)} \end{aligned}$$
SUBEQooUSZOooCqgWPE
SUBEQooBXQKooZcarVv

Justifications :

46. Voir le lemme 1.198 et la définition 1.197.
47. Vérifiez par vous-même et écrivez-moi si il y a un problème (personnellement, je n'ai pas vérifié.).
48. Si quelqu'un sait comment éviter ce bricolage en deux parties, je suis preneur.
49. Si vous ne savez pas où placer les indices, voyez la proposition 8.184.

— Pour 17.409e. Puisque $p > 1$ nous avons $p - 2 > -1$. Là, l'inégalité est stricte et c'est important.

— Pour 17.409f. C'est l'inégalité de Cauchy-Schwarz du théorème 10.1.

Voilà. La matrice Hessienne est strictement définie positive par le lemme 8.234(1) sur $\mathbb{R}^n\backslash\{0\}$. Le corolaire 17.98(1) nous indique que pour tout $x, y \in \mathbb{R}^n\backslash\{0\}$ et pour tout $\theta \in]0,1[$,

$$f_p\big(\theta x + (1-\theta)y\big) < \theta f_p(x) + (1-\theta) f_p(y) \tag{17.410}$$

EQooFXXZooBZhJYY

pourvu que $\theta x + (1-\theta)y \neq 0$ pour tout θ.

(2) **La suite**

Nous devons prouver que l'inéquation (17.410) tient également lorsque $\theta x + (1-\theta)y = 0$ pour une certaine valeur $\theta = \theta_0 \in]0,1[$.

(3) **Les cordes passant par zéro**

Pour ce faire, nous allons montrer que le segment de droite joignant $\big(a, f_p(a)\big)$ à $(b, f_p(b))$ est toujours au-dessus de la courbe $\big(x, f_p(x)\big)$. Nous commençons par $b = 0$. Vu que $p > 1$ et que $\theta \in]0,1[$ nous avons

$$\|\theta a\|^p = |\theta|^p \|a\|^p < \theta \|a\|^p = \theta f_p(a). \tag{17.411}$$

(4) **Les cordes passant au-dessus de zéro**

Dans le cas $b \neq 0$ nous considérons

$$\begin{aligned} l_1 \colon [a,b] &\to \mathbb{R} \\ x &\mapsto l_1(x) \end{aligned} \tag{17.412}$$

tel que $\big(x, l_1(x)\big)$ soit (le segment) la droite joignant $\big(a, \|a\|^p\big)$ à $\big(b, \|b\|^p\big)$.

Nous considérons aussi, pour $x \in [0,a]$ la fonction l_2 telle que $\big(x, l_2(x)\big)$ soit la droite joignant $(0,0)$ à $\big(a, \|a\|^p\big)$. Nous avons déjà vu que pour $x \in]0,a[$ nous avons $l_1(x) > \|x\|^p$.

Nous avons $l_1(a) = l_2(a) = \|a\|^p$. Donc $l_2(x) - l_1(x)$ ne change pas de signe sur $[0,a]$. Mais comme $l_1(0) > 0 = l_2(0)$ nous avons

$$l_2(x) > l_1(x) \tag{17.413}$$

pour tout $x \in [0,a]$.

Au final, $l_2(x) > l_1(x) > f_p(x)$.

Pour la partie $[b,0]$ nous faisons de même en considérant l_3 de telle sorte que $\big(x, l_3(x)\big)$ soit le segment joignant $(0,0)$ à $\big(b, \|b\|^p\big)$.

□

17.12.6 Quelques inégalités

17.12.6.1 Inégalité de Jensen

PropXIBooLxTkhU

Proposition 17.102 (Inégalité de Jensen)**.**
Soit $f \colon \mathbb{R} \to \mathbb{R}$ une fonction convexe et des réels $x_1, \ldots, x_n$. Soient des nombres positifs $\lambda_1, \ldots, \lambda_n$ formant une combinaison convexe[50]*. Alors*

$$f\big(\sum_i \lambda_i x_i\big) \leqslant \sum_i \lambda_i f(x_i). \tag{17.414}$$

50. Définition 11.31.

Démonstration. Nous procédons par récurrence sur n, en sachant que $n = 2$ est la définition de la convexité de f. Vu que

$$\sum_{k=1}^{n} \lambda_k x_k = \lambda_n x_n + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n}, \tag{17.415}$$

nous avons

$$f\Big(\sum_{k=1}^{n} \lambda_k x_k \Big) \leqslant \lambda_n f(x_n) + (1 - \lambda_n) f\Big(\sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n} \Big). \tag{17.416}$$

La chose à remarquer est que les nombres $\frac{\lambda_k}{1-\lambda_n}$ avec k allant de 1 à $n-1$ forment eux-mêmes une combinaison convexe. L'hypothèse de récurrence peut donc s'appliquer au second terme du membre de droite :

$$f\Big(\sum_{k=1}^{n} \lambda_k x_k \Big) \leqslant \lambda_n f(x_n) + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} f(x_k) = \lambda_n f(x_n) + \sum_{k=1}^{n-1} \lambda_k f(x_k). \tag{17.417}$$

□

17.12.6.2 Inégalité arithmético-géométrique

La proposition suivante dit que la moyenne arithmétique de nombres strictement positifs est supérieure ou égale à la moyenne géométrique.

PropWDPooBtHIAR

Proposition 17.103 (Inégalité arithmético-géométrique[523])**.**
Soient $x_1, \dots, x_n$ des nombres strictement positifs. Nous posons

$$m_a = \frac{1}{n}(x_1 + \cdots + x_n) \tag{17.418}$$

et

$$m_g = \sqrt[n]{x_1 \dots x_n} \tag{17.419}$$

Alors $m_g \leqslant m_a$ et $m_g = m_a$ si et seulement si $x_i = x_j$ pour tout i, j.

Démonstration. Par hypothèse les nombres m_a et m_g sont tous deux strictement positifs, de telle sorte qu'il est équivalent de prouver $\ln(m_g) \leqslant \ln(m_a)$ ou encore

$$\frac{1}{n}\big(\ln(x_1) + \cdots + \ln(x_n) \big) \leqslant \ln\left(\frac{x_1 + \cdots + x_n}{n} \right). \tag{17.420}$$

Cela n'est rien d'autre que l'inégalité de Jensen de la proposition 17.102 appliquée à la fonction ln et aux coefficients $\lambda_i = \frac{1}{n}$. □

17.12.6.3 Inégalité de Kantorovitch

PropMNUooFbYkug

Proposition 17.104 (Inégalité de Kantorovitch[524])**.**
Soit A une matrice symétrique strictement définie positive dont les plus grandes et plus petites valeurs propres sont λ_{min} et λ_{max}. Alors pour tout $x \in \mathbb{R}^n$ nous avons

$$\langle Ax, x \rangle \langle A^{-1}x, x \rangle \leqslant \frac{1}{4} \left(\frac{\lambda_{min}}{\lambda_{max}} + \frac{\lambda_{max}}{\lambda_{min}} \right)^2 \|x\|^4. \tag{17.421}$$

Démonstration. Sans perte de généralité nous pouvons supposer que $\|x\| = 1$. Nous diagonalisons [51] la matrice A par la matrice orthogonale $P \in \mathrm{O}(n, \mathbb{R})$: $A = PDP^{-1}$ et $A^{-1} = PD^{-1}P^{-1}$ où D est une matrice diagonale formée des valeurs propres de A.

51. Théorème spectral 8.227.

Nous posons $\alpha = \sqrt{\lambda_{min}\lambda_{max}}$ et nous regardons la matrice

$$\frac{1}{\alpha}A + \alpha A^{-1} \tag{17.422}$$

dont les valeurs propres sont

$$\frac{\lambda_i}{\alpha} + \frac{\alpha}{\lambda_i} \tag{17.423}$$

parce que les vecteurs propres de A et de A^{-1} sont les mêmes (ce sont les valeurs de la diagonale de D). Nous allons quelque peu étudier la fonction

$$\theta(x) = \frac{x}{\alpha} + \frac{\alpha}{x}. \tag{17.424}$$

Elle est convexe en tant que somme de deux fonctions convexes. Elle a son minimum en $x = \alpha$ et ce minimum vaut $\theta(\alpha) = 2$. De plus

$$\theta(\lambda_{max}) = \theta(\lambda_{min}) = \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{17.425}$$

Une fonction convexe passant deux fois par la même valeur doit forcément être plus petite que cette valeur entre les deux[52] : pour tout $x \in [\lambda_{min}, \lambda_{max}]$,

$$\theta(x) \leqslant \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{17.426}$$

Nous sommes maintenant en mesure de nous lancer dans l'inégalité de Kantorovitch.

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left(\frac{\langle Ax, x\rangle}{\alpha} + \alpha\langle A^{-1}x, x\rangle\right) \tag{17.427a}$$

$$= \frac{1}{2}\langle\left(\frac{A}{\alpha} + \alpha A^{-1}\right)x, x\rangle \tag{17.427b}$$

$$\leqslant \frac{1}{2}\left\|\left(\frac{A}{\alpha} + \alpha A^{-1}\right)x\right\|\|x\| \tag{17.427c}$$

$$\leqslant \frac{1}{2}\left\|\frac{A}{\alpha} + \alpha A^{-1}\right\| \tag{17.427d}$$

Justifications :

— 17.427a par l'inégalité arithmético-géométrique, proposition 17.103. Nous avons aussi inséré $\alpha\frac{1}{\alpha}$ dans le produit sous la racine.
— 17.427c par l'inégalité de Cauchy-Schwarz, théorème 10.1.
— 17.427d par la définition de la norme opérateur de la proposition 10.55

La norme opérateur est la plus grande des valeurs propres. Mais les valeurs propres de $A/\alpha + \alpha A^{-1}$ sont de la forme $\theta(\lambda_i)$, et tous les λ_i sont entre λ_{min} et λ_{max}. Donc la plus grande valeur propre de $A/\alpha + \alpha A^{-1}$ est $\theta(x)$ pour un certain $x \in [\lambda_{min}, \lambda_{max}]$. Par conséquent

$$\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} \leqslant \frac{1}{2}\left\|\frac{A}{\alpha} + \alpha A^{-1}\right\| \leqslant \sqrt{\frac{\lambda_{min}}{\lambda_{max}}} + \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{17.428}$$

□

17.12.7 Norme l^p

Si vous êtes seulement intéressés par le cas $p = 1, 2$ ou le cas $p = \infty$, allez voir la proposition 10.44.

PROPooUDFTooQyhAtq

Proposition 17.105.
Pour tout réel $p \geqslant 1$, la formule

$$\|x\|_{L^p} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \tag{17.429}$$

définit une norme[53] *sur $\mathbb{R}^n$.*

52. Je ne suis pas certain que cette phrase soit claire, non ?
53. Définition 7.158.

Démonstration. Le seul point un peu délicat parmi les propriétés à vérifier de la définition 7.158 est l'inégalité triangulaire. Grâce au lemme 7.160, nous allons seulement montrer que la partie $B = \{x \in \mathbb{R}^n \text{ tel que } \|x\|_p \leqslant 1\}$ est convexe.

Soient $x, y \in B$ ainsi que $\lambda \in [0,1]$. La fonction

$$\begin{aligned} \alpha\colon [0,\infty[&\to [0,\infty[\\ t &\mapsto t^p \end{aligned} \tag{17.430}$$

est convexe par le lemme 17.87. Nous avons donc le calcul suivant :

$$\begin{aligned}
\|\lambda x + (1-\lambda)y\|_p^p &= \sum_{i=1}^{n} |\lambda x_i + (1-\lambda)y_i|^p && \text{(17.431a)}\\
&\leqslant \sum_i \big(\lambda|x_i| + (1-\lambda)|y_i|\big)^p && \text{(17.431b)}\\
&\leqslant \sum_i \lambda^p|x_i|^p + (1-\lambda)^p|y_i|^p && \text{(17.431c)}\\
&\leqslant \sum_i \lambda|x_i|^p + \sum_i (1-\lambda)|y_i|^p && \text{(17.431d)}\\
&= \lambda\|x\|_p^p + (1-\lambda)\|y\|_p^p && \text{(17.431e)}\\
&\leqslant \lambda + (1-\lambda) && \text{(17.431f)}\\
&= 1. && \text{(17.431g)}
\end{aligned}$$

Justifications.

— Pour (17.431b). L'inégalité $|a+b| \leqslant |a| + |b|$.

— Pour (17.431c). La fonction α est convexe.

— Pour (17.431d). Nous avons $\lambda \in [0,1]$ et $p \geqslant 1$ et donc $\lambda^p \leqslant \lambda$ par 12.404.

— Pour (17.431f). Parce que $x \in B$ implique $\|x\|_p \leqslant 1$ qui implique $\|x\|_p^p \leqslant 1$.

□

17.12.8 Hölder

Notre étude de la fonction puissance permet de démontrer quelques inégalités de Hölder. Voir le thème 25.

Théorème 17.106 ([525]).
Si $1 \leqslant p < \infty$, alors pour tout $x \in \mathbb{R}^n$ nous avons[54]

$$\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p}\|x\|_\infty. \tag{17.432}$$

Démonstration. Vu que la fonction $t \mapsto |t|^p$ est croissante pour les t positifs[55], pour chaque i nous avons

$$|x_i| \leqslant \left(\sum_k |x_k|^p\right)^{1/p} = \|x\|_p. \tag{17.433}$$

Cela montre que

$$\|x\|_\infty = \max\{|x_i|\} \leqslant \|x\|_p. \tag{17.434}$$

D'autre part, pour chaque i nous avons $|x_i| \leqslant \|x\|_\infty$, donc

$$\|x\|_p \leqslant \big(n\|x\|_\infty\big)^{1/p} = n^{1/p}\|x\|_\infty. \tag{17.435}$$

□

54. Définition de la norme $\|.\|_p$ sur $\mathbb{R}^n$, proposition 17.105.
55. Parce que $p \geqslant 1$ et la proposition 12.403.

Le corolaire suivant donne une façon de majorer une norme ℓ^p par une norme ℓ^q moyennant un facteur. Notons cependant que l'inégalité de Hölder de la proposition 17.111 est plus précise. Ce corolaire est suffisant pour prouver l'équivalence des normes ℓ^p.

CORooEZGHooACHOiB

Corolaire 17.107.
Soient $p \geqslant 1$ et $q \leqslant \infty$. Pour tout $x \in \mathbb{R}^n$ nous avons

$$\|x\|_p \leqslant n^{1/p}\|x\|_q. \tag{17.436}$$

Démonstration. Il suffit d'utiliser les deux inégalités du théorème 17.106. D'abord la seconde avec p, et ensuite la première avec q. □

Si vous cherchez l'inégalité de Hölder dans L^p, c'est la proposition 27.33. Les normes ℓ^p sont définies dans 17.105.

LEMooLGGDooGLGFHj

Lemme 17.108 ([526]).
Soient $x, y > 0$ ainsi que $\alpha, \beta > 0$ tels que $\alpha + \beta = 1$. Alors

$$xy \leqslant \alpha x^{1/\alpha} + \beta y^{1/\beta}. \tag{17.437}$$

Nous avons

$$xy = \alpha x^{1/\alpha} + \beta y^{1/\beta} \tag{17.438}$$

si et seulement si $x^{1/\alpha} = y^{1/\beta}$.

Démonstration. Nous utilisons le logarithme[56] et ses propriétés (surtout la proposition 15.86). D'abord

$$xy = e^{\ln(xy)} = e^{\ln(x)+\ln(y)} = e^{\alpha\frac{\ln(x)}{\alpha}+\beta\frac{\ln(y)}{\beta}} \tag{17.439}$$

Comme l'exponentielle est strictement convexe (exemple 17.86(3)) et vu que $\alpha+\beta = 1$, nous avons

$$xy = e^{\alpha\frac{\ln(x)}{\alpha}+\beta\frac{\ln(y)}{\beta}} \leqslant \alpha e^{\ln(x)/\alpha} + \beta e^{\ln(y)/\beta} = \alpha x^{1/\alpha} + \beta y^{1/\beta}. \tag{17.440}$$

EQooNLQIooAYiEAQ

Cela prouve déjà l'inégalité demandée.

Nous passons au cas d'égalité. Puisque α et β ne sont pas nuls, l'inégalité (17.440) est une égalité si et seulement si[57]

$$\frac{\ln(x)}{\alpha} = \frac{\ln(y)}{\beta}. \tag{17.441}$$

Cela signifie $\ln(x^{1/\alpha}) = \ln(y^{1/\beta})$, qui implique $x^{1/\alpha} = y^{1/\beta}$ parce que le logarithme est une bijection. □

CORooTCBZooAcZxaC

Corolaire 17.109.
Si $p, q > 0$ vérifient $\frac{1}{p} + \frac{1}{q} = 1$ et si $p > 1$ alors nous avons

$$xy \leqslant \frac{1}{p}x^p + \frac{1}{q}y^q \tag{17.442}$$

EQooWKTSooQwRsLg

pour tout $x, y \geqslant 0$, avec une égalité si et seulement si $x^p = y^q$.

Démonstration. Il suffit de poser $\alpha = 1/p$ et $\beta = 1/q$ et appliquer le lemme 17.108. □

THOooYHMJooBlXfpl

Théorème 17.110 (Inégalité de Hölder[526]).
Soient $p, q > 0$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Pour tout $x, y \in \mathbb{R}^n$ nous avons[58]

$$|x \cdot y| = \sum_{i=1}^{n} |x_i y_i| \leqslant \|x\|_p \|y\|_q. \tag{17.443}$$

56. Définition 15.83
57. Par le lemme 17.79.
58. Définition de la norme $\|.\|_p$, proposition 17.105.

Il y a égalité si et seulement si x_iy_i est de signe constant[59] *et les vecteurs $\sum_i |x_i|^p e_i$ et $\sum_i |y_i|^q e_i$ sont proportionnels.*

Démonstration. En plusieurs parties.

(1) **Le cas des vecteurs nuls** Si x ou y est nul, les inégalités sont évidentes. Donc nous supposons que non.

(2) **Première inégalité** En ce qui concerne la première inégalité,

$$|x \cdot y| = |\sum_i x_iy_i| \leqslant \sum_i |x_i||y_i| \leqslant \sum_i \left(\frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q \right) \tag{17.444}$$

où nous avons utilisé le corolaire 17.109 dans chaque terme de la somme en tenant compte du fait que $|x_i|$ et $|y_i|$ sont positifs.

(3) **Seconde inégalité** Pour la seconde inégalité, nous commençons avec $\|x\|_p = \|y\|_q = 1$. Utilisant encore le corolaire 17.109 pour chaque terme, nous avons

$$\sum_i |x_iy_i| \leqslant \frac{1}{p} \underbrace{\sum_i |x_i|^p}_{=1} + \frac{1}{q} \underbrace{\sum_i |y_i|^q}_{=1} = \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q. \tag{17.445}$$

Si maintenant x et y sont arbitraires non nuls dans $\mathbb{R}^n$, nous posons $x' = x/\|x\|_p$ et $y' = y/\|y\|_q$; nous savons déjà que

$$\sum_i |x'_iy'_i| \leqslant 1. \tag{17.446}$$

En remplaçant x'_i par $x_i/\|x\|_p$ et y'_i par $y_i/\|y\|_q$, l'inégalité (17.446) devient

$$\sum_i \frac{|x_i||y_i|}{\|x\|_p\|y\|_q} \leqslant 1, \tag{17.447}$$

ce qui signifie

$$\sum_i |x_iy_i| \leqslant \|x\|_p\|y\|_q. \tag{17.448}$$

(4) **Cas d'égalité, dans un sens** Nous notons

$$\mathcal{D} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \text{ tel que } \sum_i |x_iy_i| = \|x\|_p\|y\|_q\}. \tag{17.449}$$

(4a) **Multiplications** D'abord si $(x, y) \in \mathcal{D}$, alors $(\mu x, \lambda y) \in \mathcal{D}$ pour tout $\mu, \lambda \in \mathbb{R}$. En effet,

$$\sum_i |(\mu x)_i(\lambda y)_i| = |\mu||\lambda| \sum_i |x_iy_i| = |\mu||\lambda|\|x\|_p\|y\|_q = \|\mu x\|_p\|\lambda y\|_q. \tag{17.450}$$

(4b) **Avec normes égales à 1** Soit $(x, y) \in \mathcal{D}$ tels que $\|x\|_p = \|y\|_q = 1$. Nous avons en particulier,

$$1 = \sum_i |x_iy_i| \leqslant \frac{1}{p} \sum_i |x_i|^p + \frac{1}{q} \sum_i |y_i|^q \tag{17.451}$$

grâce à l'inégalité (17.442) appliquée à chaque terme. Vu que $\|x\|_p = 1$, nous avons $\sum_i |x_i|^p = 1$, de telle sorte que le membre de droite de (17.451) se réduise à $\frac{1}{p} + \frac{1}{q} = 1$. Nous pouvons donc écrire

$$1 = \sum_i |x_iy_i| \leqslant \frac{1}{p} \sum_i |x_i|^p + \frac{1}{q} \sum_i |y_i|^q = 1. \tag{17.452}$$

59. Je n'utilie pas cette hypothèse de signe constant. Il doit y avoir une subtilité qui m'a échappée. Soyez prudente en lisant et écrivez-moi si vous trouvez une erreur.

L'inégalité est donc une égalité :

$$\sum_i |x_i y_i| = \sum_i \left(\frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q \right). \tag{17.453}$$

Mais chaque terme à gauche est en inégalité avec le terme correspondant à droite :

$$|x_i y_i| \leqslant \frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q. \tag{17.454}$$

EQooAGFKooUsYrWT

Pour que le tout soit une égalité, il faut que chaque inégalité (17.454) soit une égalité. Pour chaque i, nous avons

$$|x_i y_i| = \frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q. \tag{17.455}$$

La condition d'égalité du corolaire 17.109 nous dit alors que $|x_i|^p = |y_i|^q$.

(4c) **Avec normes arbitraires** Soit donc $(x, y) \in \mathcal{D}$. Nous savons qu'en posant $x' = x/\|x\|_p$ et $y' = y/\|y\|_q$ nous avons $(x', y') \in \mathcal{D}$ et donc

$$\left(\frac{|x_i|}{\|x\|_p} \right)^p = \left(\frac{|y_i|}{\|y\|_q} \right)^q. \tag{17.456}$$

Cela donne tout de suite

$$|x_i|^p = \frac{\|x\|_p^p}{\|y\|_q^q}|y_i|^q, \tag{17.457}$$

ce qui est bien ce que nous voulions : le vecteur $\sum_i |x_i|^p e_i$ est proportionnel au vecteur $\sum_i |y_i|^q e_i$.

(5) **Cas d'égalité dans l'autre sens[1]** Nous supposons que les vecteurs $\sum_i |x_i|^p e_i$ et $\sum_i |y_i|^q e_i$ sont proportionnels. Nous nommons c^q le facteur de proportionnalité, c'est-à-dire que nous posons

$$|x_i|^p = c^q |y_i|^q. \tag{17.458}$$

Dans ce cas, pour chaque i, les nombres $c|y_i|$ et $|x_i|$ sont dans le cas d'égalité du corolaire 17.109. Nous avons alors

SUBEQSooVULLooPGWUIP

$$\begin{aligned}
\sum_i |x_i y_i| &= \frac{1}{c}\sum_i c|x_i||y_i| && (17.459a)\\
&= \frac{1}{c}\sum_i \left(\frac{1}{p}|x_i|^p + \frac{1}{q}\big(c|y_i|\big)^q \right) && (17.459b)\\
&= \frac{1}{c}\sum_i \left(\frac{1}{p}|x_i|^p + \frac{1}{q}|x_i|^p \right) && (17.459c)\\
&= \frac{1}{c}\sum_i |x_i|^p. && (17.459d)
\end{aligned}$$

Et c'est maintenant que nous subdivisons.

(5a) **Si $\|x\|_p = \|y\|_q = 1$** Dans ce cas, l'égalité (17.459) se réduisent à

$$\sum_i |x_i y_i| = \frac{1}{c}. \tag{17.460}$$

Mais l'hypothèse sur les normes donne

$$1 = \sum_i |x_i|^p = \sum_i c^q |y_i|^q = c^q \sum_i |y_i|^q = c^q. \tag{17.461}$$

Donc $c = 1$ et nous avons bien

$$\sum_i |x_i y_i| = \frac{1}{c} = 1 = \|x\|_p \|y\|_q. \tag{17.462}$$

(5b) **Pour des normes arbitraires** Soit $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$ tels que $|x_i|^p=c^q|y_i|^q$. Nous posons comme d'habitude $x'=x/\|x\|_p$ et $y'=y/\|y\|_q$. En utilisant le cas « de norme 1 » nous avons

$$1=\sum_i|x'_iy'_i|=\frac{1}{\|x\|_p\|y\|_q}\sum_i|x_iy_i|. \tag{17.463}$$

Donc $\sum_i|x_iy_i|=\|x\|_p\|y\|_q$ comme nous le voulions.

□

La majoration de la proposition suivante sera utile pour les inégalités de Clarkson du lemme 27.148. Pour d'autres inégalités (plus simples) autour des normes $\|.\|_p$, voir le thème 25.

PROPooQZTNooGACMlQ

Proposition 17.111 ([526]).
Si $x\in\mathbb{R}^n$ et si $0<q<p$, alors

$$\|x\|_q\leqslant n^{\frac{1}{q}-\frac{1}{p}}\|x\|_p. \tag{17.464}$$

En particulier, si $0<q<p$, alors

$$\|x\|_q\leqslant\|x\|_p. \tag{17.465}$$

Démonstration. Nous posons $P=p/q$ et $Q=P/(P-1)$. Les nombres P et Q sont des exposants conjugués, parce que

$$\frac{1}{P}+\frac{1}{Q}=\frac{q}{p}+\frac{p-q}{p}=1. \tag{17.466}$$

Nous posons $y=(1,\dots,1)\in\mathbb{R}^n$ ainsi que

$$v=\sum_i|x_i|^qe_i, \tag{17.467}$$

et nous écrivons l'inégalité de Hölder de la proposition 17.110 sur les vecteurs v et y :

$$\sum_i|v_iy_i|\leqslant\|v\|_P\|y\|_Q. \tag{17.468}$$

En déballant,

$$\sum_i|x_i|^q\leqslant\left(\sum_i(|x_i|^q)^P\right)^{1/P}\Big(\underbrace{\sum_i1^Q}_{=n}\Big)^{1/Q} \tag{17.469a}$$

$$=\left(\sum_i|x_i|^p\right)^{q/p}n^{1/Q} \tag{17.469b}$$

$$=n^{1-q/p}\|x\|_p^q. \tag{17.469c}$$

Cela donne

$$\|x\|_q^q\leqslant n^{1-q/p}\|x\|_p^q. \tag{17.470}$$

En prenant la puissance $1/q$ des deux côtés,

$$\|x\|_q\leqslant n^{\frac{1}{q}-\frac{1}{p}}\|x\|_p\leqslant\|x\|_p \tag{17.471}$$

parce que $n\geqslant1$ et $p/q<1$. □

17.13 Algorithme du gradient à pas optimal

Une idée pour trouver un minimum à une fonction est de prendre un point p au hasard, calculer le gradient $\nabla f(p)$ et suivre la direction $-\nabla f(p)$ tant que ça descend. Une fois qu'on est « dans le creux », recalculer le gradient et continuer ainsi.

Nous allons détailler cet algorithme dans un cas très particulier d'une matrice A symétrique et strictement définie positive.

— Dans la proposition 17.113 nous montrons que résoudre le système linéaire $Ax = -b$ est équivalent à minimiser une certaine fonction.

— La proposition 17.114 donnera une méthode itérative pour trouver ce minimum.

DefQXPooYSygGP

Définition 17.112.
*Si X est un espace vectoriel normé et $f : X \to \mathbb{R} \cup \{\pm\infty\}$ nous disons que f est **coercive** sur le domaine non borné P de X si pour tout $M \in \mathbb{R}$, l'ensemble*

$$\{x \in P \text{ tel que } f(x) \leqslant M\} \tag{17.472}$$

est borné.

En langage imagé la coercivité de f s'exprime par la limite

$$\lim_{\substack{\|x\|\to\infty \\ x\in P}} f(x) = +\infty. \tag{17.473}$$

Nous rappelons que $S^{++}(n, \mathbb{R})$ est l'ensemble des matrices symétriques strictement définies positives définies en 8.231.

PROPooYRLDooTwzfWU

Proposition 17.113.
Soit $A \in S^{++}(n, \mathbb{R})$ et $b \in \mathbb{R}^n$. Nous considérons l'application

$$\begin{aligned} f : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle. \end{aligned} \tag{17.474}$$

Alors :

(1) Il existe un unique $\bar{x} \in \mathbb{R}^n$ tel que $A\bar{x} = -b$.

(2) Il existe un unique $x^* \in \mathbb{R}^n$ minimisant f.

(3) Ils sont égaux : $\bar{x} = x^*$.

Démonstration. Une matrice symétrique strictement définie positive est inversible, entre autres parce qu'elle se diagonalise par des matrices orthogonales (qui sont inversibles) et que la matrice diagonalisée est de déterminant non nul : tous les éléments diagonaux sont strictement positifs. Voir le théorème spectral symétrique 8.227.

D'où l'unicité du $\bar{x}$ résolvant le système $Ax = -b$ pour n'importe quel b.

(1) **f est strictement convexe**
La fonction f s'écrit

$$f(x) = \frac{1}{2}\sum_{kl} A_{kl}x_l x_k + \sum_k b_k x_k. \tag{17.475}$$

Elle est de classe C^2 sans problèmes, et il est vite vu que $\frac{\partial^2 f}{\partial x_i \partial x_j} = A_{ij}$, c'est-à-dire que A est la matrice hessienne de f. Cette matrice étant strictement définie positive par hypothèse, la fonction f est strictement convexe par le corolaire 17.98(2).

(2) **f est coercive** Montrons à présent que f est coercive. Nous avons :

$$|f(x)| = |\frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle| \tag{17.476a}$$
$$\geqslant \frac{1}{2}|\langle Ax, x\rangle| - |\langle b, x\rangle| \tag{17.476b}$$
$$\geqslant \frac{1}{2}\lambda_{min}\|x\|^2 - \|b\|\|x\| \tag{17.476c}$$

Pour la dernière ligne nous avons nommé λ_{min} la plus grande valeur propre de A et utilisé Cauchy-Schwarz pour le second terme. Nous avons donc bien $|f(x)| \to \infty$ lorsque $\|x\| \to \infty$ et la fonction f est coercive.

Soit M une valeur atteinte par f. L'ensemble

$$\{x \in \mathbb{R}^n \text{ tel que } f(x) \leqslant M\} \tag{17.477}$$

est fermé (parce que f est continue) et borné parce que f est coercive. Cela est donc compact [60] et f atteint un minimum qui sera forcément dedans. Cela est pour l'existence d'un minimum.

Pour l'unicité du minimum nous invoquons la convexité : si $\bar{x}_1$ et $\bar{x}_2$ sont deux points réalisant le minimum de f, alors

$$f\left(\frac{\bar{x}_1 + \bar{x}_2}{2}\right) < \frac{1}{2}f(\bar{x}_1) + \frac{1}{2}f(\bar{x}_2) = f(\bar{x}_1), \tag{17.478}$$

ce qui contredit la minimalité de $f(\bar{x}_1)$.

Nous devons maintenant prouver que $\bar{x}$ vérifie l'équation $A\bar{x} = -b$. Vu que $\bar{x}$ est minimum local de f qui est une fonction de classe C^2, le théorème des minimums locaux 17.69 nous indique que $\bar{x}$ est solution de $\nabla f(x) = 0$. Calculons un peu cela avec la formule

$$df_x(u) = \frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} = \frac{1}{2}\big(\langle Ax, u\rangle + \langle Au, x\rangle\big) + \langle b, u\rangle = \langle Ax, u\rangle + \langle b, u\rangle = \langle Ax+b, u\rangle. \tag{17.479}$$

Donc demander $df_x(u) = 0$ pour tout u demande $Ax + b = 0$. □

PropSOOooGoMOxG

Proposition 17.114 (Gradient à pas optimal)**.**
Soit $A \in S^{++}(n, \mathbb{R})$ (A est une matrice symétrique strictement définie positive) et $b \in \mathbb{R}^n$. Nous considérons l'application

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle. \end{aligned} \tag{17.480}$$

Soit $x_0 \in \mathbb{R}^n$. Nous définissons la suite (x_k) par

$$x_{k+1} = x_k + t_k d_k \tag{17.481}$$

où

— $d_k = -(\nabla f)(x_k)$

— *t_k est la valeur minimisant la fonction $t \mapsto f(x_k + td_k)$ sur $\mathbb{R}$.*

Alors pour tout $k \geqslant 0$ nous avons

$$\|x_k - \bar{x}\| \leqslant K\left(\frac{c_2(A) - 1}{c_2(A) + 1}\right)^k \tag{17.482}$$

où $c_2(A) = \frac{\lambda_{max}}{\lambda_{min}}$ est le rapport ente la plus grande et la plus petite valeur propre [61] de la matrice A et $\bar{x}$ est l'unique élément de $\mathbb{R}^n$ à minimiser f.

60. Théorème 9.24

61. Cela est certainement très lié au conditionnement de la matrice A, voir la proposition 34.110.

Démonstration. Décomposition en plusieurs points.

(1) **Existence de $\bar{x}$** Le fait que $\bar{x}$ existe et soit unique est la proposition 17.113.

(2) **Si $(\nabla f)(x_k) = 0$** D'abord si $\nabla f(x_k) = 0$, c'est que $x_{k+1} = x_k$ et l'algorithme est terminé : la suite est stationnaire. Pour dire que c'est gagné, nous devons prouver que $x_k = \bar{x}$. Pour cela nous écrivons (à partir de maintenant « x_k » est la k^{e}composante de x qui est une variable, et non le x_k de la suite)

$$f(x) = \frac{1}{2}\sum_{kl} A_{kl}x_l x_k + \sum_k b_k x_k \tag{17.483}$$

et nous calculons $\frac{\partial f}{\partial x_i}(a)$ en tenant compte du fait que $\frac{\partial x_k}{\partial x_i} = \delta_{ki}$. Le résultat est que $(\partial_i f)(a) = (Ax + b)_i$ et donc que

$$(\nabla f)(a) = Aa + b. \tag{17.484}$$

Vu que A est inversible (symétrique définie positive), il existe un unique $a \in \mathbb{R}^n$ qui vérifie cette relation. Par la proposition 17.113, cet élément est le minimum $\bar{x}$.
Cela pour dire que si $a \in \mathbb{R}^n$ vérifie $(\nabla f)(a) = 0$ alors $a = \bar{x}$. Nous supposons donc à partir de maintenant que $\nabla f(x_k) \neq 0$ pour tout k.

(3) **t_k est bien défini**
Pour $t \in \mathbb{R}$ nous avons

$$f(x_k+td_k) = f(x_k)+\frac{1}{2}t^2\langle Ad_k, d_k\rangle+t\langle \underbrace{Ax_k + b}_{=-d_k}, d_k\rangle = \frac{1}{2}t^2\langle Ad_k, d_k\rangle - t\|d_k\|^2+f(x_k). \tag{17.485}$$

qui est un polynôme du second degré en t. Le coefficient de t^2 est $\frac{1}{2}\langle Ad_k, d_k\rangle > 0$ parce que $d_k \neq 0$ et A est strictement définie positive. Par conséquent la fonction $t \mapsto f(x_k + td_k)$ admet bien un unique minimum. Nous pouvons même calculer t_k parce que l'on connaît par cœur le sommet d'une parabole :

$$t_k = -\frac{\langle Ax_k + b, d_k\rangle}{\langle Ad_k, d_k\rangle} = \frac{\|d_k\|^2}{\langle Ad_k, d_k\rangle} \tag{17.486}$$

parce que $d_k = -\nabla f(x_k) = -(Ax_k + b)$.

(4) **La valeur de d_{k+1}**
Par définition, $d_{k+1} = -\nabla f(x_{k+1}) = -(Ax_{k+1} + b)$. Mais $x_{k+1} = x_k + t_k d_k$, donc

$$d_{k+1} = -Ax_k - t_k Ad_k - b = d_k - t_k Ad_k \tag{17.487}$$

parce que $-Ax_k - b = d_k$.
Par ailleurs, $\langle d_{k+1}, d_k\rangle = 0$ parce que

$$\langle d_{k+1}, d_k\rangle = \langle d_k, d_k\rangle - t_k\langle d_k, Ad_k\rangle = \|d_k\|^2 - \frac{\|d_k\|^2}{\langle Ad_k, d_k\rangle}\langle d_k, Ad_k\rangle = 0 \tag{17.488}$$

où nous avons utilisé la valeur (17.486) de t_k.

(5) **Calcul de $f(x_{k+1})$**
Nous repartons de (17.485) où nous substituons la valeur (17.486) de t_k :

$$f(x_{k+1}) = f(x_k) + \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} = f(x_k) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle}. \tag{17.489}$$

(6) **Encore du calcul . . .**
Vu que le produit $\langle Ad_k, d_k\rangle$ arrive tout le temps, nous allons étudier $\langle A^{-1}d_k, d_k\rangle$. Le truc malin est d'essayer d'exprimer ça en termes de $\bar{x}$ et $\bar{f} = f(\bar{x})$. Pour cela nous calculons $f(\bar{x})$:

$$\bar{f} = f(\bar{x}) = f(-A^{-1}b) = -\frac{1}{2}\langle b, A^{-1}b\rangle. \tag{17.490}$$

Ayant cela en tête nous pouvons calculer :

$$\langle A^{-1}d_k, d_k\rangle = \langle A^{-1}(Ax_k+b), Ax_k+b\rangle \tag{17.491a}$$
$$= \langle x_k, Ax_k\rangle + \langle A^{-1}b, Ax_k\rangle + \langle b, x_k\rangle + \underbrace{\langle A^{-1}b, b\rangle}_{-2\bar{f}} \tag{17.491b}$$
$$= \langle x_k, Ax_k\rangle + 2\langle x_k, b\rangle - 2\bar{f} \tag{17.491c}$$
$$= 2\big(f(x_k)-\bar{f}\big) \tag{17.491d}$$

où nous avons utilisé le fait que $\langle x, Ay\rangle = \langle Ax, y\rangle$ parce que A est symétrique.

(7) **Erreur sur la valeur du minimum**

Nous voulons à présent estimer la différence $f(x_{k+1})-\bar{f}$. Pour cela nous mettons en facteur $f(x_k)-\bar{f}$ dans $f(x_{k+1}-\bar{f})$; et d'ailleurs c'est pour cela que nous avons calculé $\langle A^{-1}d_k, d_k\rangle$: parce que ça fait intervenir $f(x_k)-\bar{f}$.

$$f(x_{k+1})-\bar{f} = f(x_k) - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle} - \bar{f} \tag{17.492a}$$
$$= \big(f(x_k)-\bar{f}\big)\left(1 - \frac{1}{2}\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\big(f(x_k)-\bar{f}\big)}\right) \tag{17.492b}$$
$$= \big(f(x_k)-\bar{f}\big)\left(1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\langle A^{-1}d_k, d_k\rangle}\right). \tag{17.492c}$$

Nous traitons le dénominateur à l'aide de l'inégalité de Kantorovitch 17.104. Nous avons

$$\frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\langle A^{-1}d_k, d_k\rangle} \geqslant \frac{\|d_k\|^4}{\frac{1}{4}\left(\sqrt{c_2(A)} + \frac{1}{\sqrt{c_2(A)}}\right)^2\|d_k\|^4} = \frac{4c_2(A)}{(c_2(A)+1)^2}. \tag{17.493}$$

Mettre cela dans (17.492c) est un calcul d'addition de fractions :

$$f(x_{k+1})-\bar{f} \leqslant \big(f(x_k)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^2. \tag{17.494}$$

Par récurrence nous avons alors

$$f(x_k)-\bar{f} \leqslant \big(f(x_0)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^{2k}. \tag{17.495}$$

Notons qu'il n'y a pas de valeurs absolues parce que $\bar{f}$ étant le minimum de f, les deux côtés de l'inégalité sont automatiquement positifs.

(8) **Erreur sur la position du minimum**

Nous voulons à présent étudier la norme de $x_k-\bar{x}$. Pour cela nous l'écrivons directement avec la définition de f en nous souvenant que $b=-A\bar{x}$:

$$f(x_k)-\bar{f} = \frac{1}{2}\langle Ax_k, x_k\rangle - \langle A\bar{x}, x_k\rangle + \frac{1}{2}\langle A\bar{x}, \bar{x}\rangle \tag{17.496a}$$
$$= \frac{1}{2}\langle Ax_k, x_k\rangle - \frac{1}{2}\langle A\bar{x}, x_k\rangle - \frac{1}{2}\langle A\bar{x}, x_k\rangle + \frac{1}{2}\langle A\bar{x}, \bar{x}\rangle \tag{17.496b}$$
$$= \frac{1}{2}\Big(\langle A(x_k-\bar{x}), x_k\rangle + \langle A\bar{x}, \bar{x}-x_k\rangle\Big) \tag{17.496c}$$
$$= \frac{1}{2}\Big(\langle A(x_k-\bar{x}), (x_k-\bar{x})\rangle\Big) \tag{17.496d}$$

où à la dernière ligne nous avons fait $\langle A\bar{x}, \bar{x}-x_k\rangle = \langle \bar{x}, A(\bar{x}-x_k)\rangle$ en vertu de la symétrie de A.

Les produits de la forme $\langle Ay, y\rangle$ sont majorés par $\lambda_{min}\|y\|^2$ parce que λ_{min} est la plus petite valeur propre de A. Dans notre cas,

$$f(x_k)-\bar{f} \geqslant \frac{1}{2}\lambda_{min}\|x_k-\bar{x}\|^2 \tag{17.497}$$

(9) **Conclusion**

En combinant les inéquations (17.497) et (17.495) nous trouvons

$$\frac{1}{2}\lambda_{min}\|x_k-\bar{x}\|^2 \leqslant f(x_k)-\bar{f} \leqslant \big(f(x_0)-\bar{f}\big)\left(\frac{c_2(A)-1}{c_2(A)+1}\right)^{2k}, \tag{17.498}$$

c'est-à-dire

$$\|x_k-\bar{x}\| \leqslant \sqrt{\frac{2\big(f(x_0)-\bar{f}\big)}{\lambda_{min}+1}}^{2k}. \tag{17.499}$$

□

Notons que lorsque $c_2(A)$ est proche de 1 la méthode converge rapidement. Par contre si $c_2(A)$ est proche de zéro, la méthode converge lentement.

17.14 Formes quadratiques, signature, et lemme de Morse

NORMooHSWKooLtUbRl

17.115.
Soit $(E,\|.\|_E)$ un espace vectoriel réel normé de dimension finie n. L'ensemble des formes quadratiques réelles[62] sur E est vu comme l'ensemble des matrices symétriques $S_n(\mathbb{R})$; il sera noté $Q(E)$ et le sous-ensemble des formes quadratiques non dégénérées est $S_n(\mathbb{R})\cap \mathrm{GL}(n,\mathbb{R})$ qui sera noté $\Omega(E)$. Nous rappelons que la correspondance est donnée de la façon suivante.

Si $A\in S_n(\mathbb{R})$, la forme quadratique associée est q_A donnée par le produit scalaire $q_A(x)=x\cdot Ax$.

Pour information, le lemme de Morse est le lemme 20.198.

NORMooQZFLooYnILtn

17.116.
Nous noterons encore $Q^+(E)$ les formes quadratiques positives sur E et $Q^{++}(E)$ les formes quadratiques strictement définies positives sur E.

Sur $Q(E)$ nous mettons la norme

$$N(q)=\sup_{\|x\|_E=1}|q(x)|, \tag{17.500}$$

qui du point de vue de $S_n(\mathbb{R})$ est

$$N(A)=\sup_{\|x\|_E=1}|x^tAx|. \tag{17.501}$$

EqDOgBNAg

Notons que à droite, c'est la valeur absolue usuelle sur $\mathbb{R}$.

Nous savons par le théorème de Sylvester (théorème 8.251) que dans $\mathbb{M}(n,\mathbb{R})$, toute matrice symétrique de signature (p,q) est semblable à la matrice

$$\mathbb{1}_{p,q}=\begin{pmatrix}-\mathbb{1}_q & & \\ & \mathbb{1}_p & \\ & & 0_{n-p-q}\end{pmatrix}. \tag{17.502}$$

Donc deux matrices de S_n sont semblables si et seulement si elles ont la même signature (même si elles ne sont pas de rang maximum, cela soit dit au passage). Si nous notons $S_n^{p,q}(\mathbb{R})$ l'ensemble des matrices réelles symétriques de signature (p,q), alors

$$S_n^{p,q}(\mathbb{R})=\{P^tAP \text{ tel que } P\in \mathrm{GL}(n,\mathbb{R})\} \tag{17.503}$$

où A est une quelconque ce ces matrices.

62. Définition 8.125.

Nous voudrions en savoir plus sur ces ensembles. En particulier nous aimerions savoir si la signature est une notion « stable » au sens où ces ensembles seraient ouverts dans S_n. Pour cela nous considérons l'action de $\mathrm{GL}(n,\mathbb{R})$ sur S_n définie par

$$\begin{aligned}\alpha\colon \mathrm{GL}(n,\mathbb{R})\times S_n(\mathbb{R}) &\to S_n(\mathbb{R})\\ (P,A) &\mapsto P^tAP\end{aligned}\tag{17.504}$$

faite exprès pour que les orbites de cette action soient les ensembles $S_n^{p,q}(\mathbb{R})$.

La proposition suivante montre que lorsque $p+q=n$, c'est-à-dire lorsqu'on parle de matrices de rang maximum, les ensembles $S_n^{p,q}(\mathbb{R})$ sont ouverts, c'est-à-dire que la signature d'une forme quadratique est une propriété « stable » par petite variations des éléments de matrice. Notons tout de suite que si le rang n'est pas maximum, le théorème de Sylvester dit qu'elle est semblable à une matrice diagonale avec des zéros sur la diagonale ; en modifiant un peu ces zéros, on peut modifier évidemment la signature.

PropNPbnsMd

Proposition 17.117 ([104])**.**
Soit $(E,\|.\|_E)$ un espace vectoriel normé de dimension finie. Alors

(1) les formes quadratiques non dégénérées forment un ouvert dans l'ensemble des formes quadratiques,

(2) les ensembles $S_n^{p,q}(\mathbb{R})$ avec $p+q=n$ sont ouverts dans $S_n(\mathbb{R})$, ItemGOhRIiViii

(3) les composantes connexes de $\Omega(E)$ sont les $S_n^{p,q}(\mathbb{R})$ avec $p+q=n$, ItemGOhRIiViv

(4) les $S_n^{p,q}(\mathbb{R})$ non dégénérés sont connexes par arc.

Démonstration. Cette preuve est donnée du point de vue des matrices. La différence entre le point (3) et (4) est que dans le premier nous prouvons la connexité de $S_n^{p,q}(\mathbb{R})$ à partir de la connexité de $\mathrm{GL}^+(n,\mathbb{R})$, tandis que dans le second nous prouvons la connexité par arc de $S_n^{p,q}(\mathbb{R})$ à partir de la connexité par arc de $\mathrm{GL}^+(n,\mathbb{R})$. Bien entendu le second implique le premier.

(1) Il s'agit simplement de remarquer que $Q(E)=S_n(\mathbb{R})$, que $\Omega(E)=S_n(\mathbb{R})\cap\mathrm{GL}(n,\mathbb{R})$ et que le déterminant est une fonction continue sur $\mathbb{M}(n,\mathbb{R})$.

(2) Soit $A_0\in S_n^{p,q}(\mathbb{R})$. Le théorème de Sylvester 8.251 nous donne une matrice inversible P telle que $P^tA_0P=\mathbb{1}_{p,q}$. Nous allons montrer qu'il existe un voisinage $\mathcal{U}$ de $\mathbb{1}_{p,q}$ contenu dans $S_n^{p,q}(\mathbb{R})$. À partir de là, l'ensemble $(P^{-1})^t\mathcal{U}P^{-1}$ sera un voisinage de A_0 contenu dans $S_n^{p,q}(\mathbb{R})$.

Nous considérons les espaces vectoriels

$$F=\mathrm{Span}\{e_1,\dots,e_p\}\tag{17.505a}$$
$$G=\mathrm{Span}\{e_{p+1},\dots,e_n\}\tag{17.505b}$$

La norme euclidienne $\|.\|_p$ sur F est équivalente à la norme $|.|_E$ par le théorème 10.49. Donc il existe une constante $k_1>0$ telle que pour tout $x\in F$,

$$\|x\|_p\geqslant k_1\|x\|_E.\tag{17.506}$$ EqMVjCiJJ

De la même façon sur G, il existe une constante $k_2>0$ telle que

$$\|x\|_q\geqslant k_2\|x\|_E.\tag{17.507}$$ EqSEwOcDw

Si nous posons $k=\min\{k_1^2,k_2^2\}$, alors nous avons

$$\forall x\in F,\quad \|x\|_p^2\geqslant k_1^2\|x\|_E^2\geqslant k\|x\|_E^2\tag{17.508a}$$
$$\forall x\in G,\quad \|x\|_q^2\geqslant k_2^2\|x\|_E^2\geqslant k\|x\|_E^2.\tag{17.508b}$$

Soit une matrice $A \in S_n(\mathbb{R})$ telle que $N(A-\mathbb{1}_{p,q}) < k$, c'est-à-dire que A est dans un voisinage de $\mathbb{1}_{p,q}$ pour la norme sur $S_n(\mathbb{R})$ donné par (17.501). Si x est non nul dans E, nous avons

$$\left|x^t(A-\mathbb{1}_{p,q})x\right| \leqslant N(\mathbb{1}_{p,q}-A)\|x\|^2 \leqslant k\|x\|^2. \tag{17.509}$$

En déballant la valeur absolue, cela signifie que

$$-k\|x\|_E^2 \leqslant x^t(A-\mathbb{1}_{p,q})x \leqslant k\|x\|^2. \tag{17.510}$$

Si $x \in F$, alors la première inéquation et (17.506) donnent

$$x^tAx \geqslant \|x\|_p^2 - k\|x\|_E^2 > 0 \tag{17.511}$$

Si $x \in G$, alors la seconde inéquation et (17.507) donnent

$$x^tAx \leqslant k\|x\|_E^2 - \|x\|_q^2 < 0. \tag{17.512}$$

Nous avons donc montré que $x \mapsto x^tAx$ est positive sur F et négative sur G, ce qui prouve que A est bien de signature (p,q) et appartient donc à $S_n^{p,q}(\mathbb{R})$. Autrement dit nous avons

$$B(\mathbb{1}_{p,q},k) \subset S_n^{p,q}(\mathbb{R}). \tag{17.513}$$

(3) Cette partie de la preuve provient essentiellement de [527], et fonctionne pour tous les $S_n^{p,q}(\mathbb{R})$, même pour ceux qui ne sont pas de rang maximum.
Soit $A \in S_n^{p,q}(\mathbb{R})$. Nous savons que $\mathrm{GL}(n,\mathbb{R})$ contient exactement deux composantes connexes (proposition 13.20). Vu que l'application

$$\begin{aligned}\alpha\colon \mathrm{GL}(n,\mathbb{R}) &\to S_n\\ P &\mapsto P^tAP\end{aligned} \tag{17.514}$$

est continue, l'image d'un connexe de $\mathrm{GL}(n,\mathbb{R})$ par α est connexe (lemme 7.218). En particulier, $\alpha\big(\mathrm{GL}^{\pm}(n,\mathbb{R})\big)$ sont deux connexes et nous savons que $S_n^{p,q}(\mathbb{R})$ a au plus ces deux composantes connexes.
Notre but est maintenant de trouver une intersection entre les parties $\alpha\big(\mathrm{GL}^{+}(n,\mathbb{R})\big)$ et $\alpha\big(\mathrm{GL}^{-}(n,\mathbb{R})\big)$ [63]. Soit par le théorème de Sylvester, soit par le théorème de diagonalisation des matrices symétriques réelles 8.227, il existe une matrice $P \in \mathrm{GL}(n,\mathbb{R})$ diagonalisant A. En suivant la remarque 8.228, et en notant Q la matrice obtenue à partir de P en changeant le signe de sa première ligne, nous avons

$$\alpha(Q) = Q^tAQ = P^tAP = \alpha(P). \tag{17.515}$$

Or si $P \in \mathrm{GL}^+(n,\mathbb{R})$, alors $Q \in \mathrm{GL}^-(n,\mathbb{R})$ et inversement. Donc nous avons trouvé une intersection entre $\alpha\big(\mathrm{GL}^{+}(n,\mathbb{R})\big)$ et $\alpha\big(\mathrm{GL}^{-}(n,\mathbb{R})\big)$.

(4) Soient A et B dans $S_n^{p,q}(\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R})$. Par le théorème de Sylvester, il existe P et Q dans $\mathrm{GL}(n,\mathbb{R})$ telles que $A = P^t\mathbb{1}_{p,q}P$ et $B = Q^t\mathbb{1}_{p,q}Q$. Par la remarque 8.228 nous pouvons choisir P et Q dans $\mathrm{GL}^+(n,\mathbb{R})$. Ce dernier groupe étant connexe par arc, il existe un chemin

$$\gamma\colon [0,1] \to \mathrm{GL}^+(n,\mathbb{R}) \tag{17.516}$$

tel que $\gamma(0) = P$ et $\gamma(1) = Q$. Alors le chemin

$$s \mapsto \gamma(s)^t\mathbb{1}_{p,q}\gamma(s) \tag{17.517}$$

est un chemin continu dans $S_n^{p,q}(\mathbb{R})$ joignant A à B.

□

63. À ce point, il me semble que [527] fait erreur parce que la matrice $-\mathbb{1}_n$ est de déterminant 1 lorsque n est pair. L'argument donné ici provient de [104]

Nous savons déjà de la proposition 17.117 que les ensembles $S_n^{p,q}(\mathbb{R})$ (pas spécialement de rang maximum) sont ouverts dans $S_n(\mathbb{R})$. Le lemme suivant nous donne une précision à ce sujet, dans le cas des matrices de rang maximum, en disant que la matrice qui donne la similitude entre A_0 et A est localement un C^1-difféomorphisme de A.

LemWLCvLXe

Lemme 17.118.
Soit $A_0 \in \Omega(\mathbb{R}^n) = S_n \cap \mathrm{GL}(n,\mathbb{R})$, une matrice symétrique inversible. Alors il existe un voisinage V de A_0 dans S_n et une application $\varphi\colon V \to \mathrm{GL}(n,\mathbb{R})$ qui

(1) est de classe C^1,

(2) est telle que pour tout $A \in V$, $\varphi(A)^t A_0 \phi(A) = A$.

Démonstration. Nous considérons l'application

$$\begin{aligned}\varphi\colon \mathbb{M}(n,\mathbb{R}) &\to S_n\\ M &\mapsto M^t A_0 M.\end{aligned} \tag{17.518}$$

Étant donné que les composantes de $\varphi(M)$ sont des polynômes en les entrées de M, cette application est de classe C^1 – et même plus. Soit maintenant $H \in \mathbb{M}(n,\mathbb{R})$ et calculons $d\varphi_{\mathbb{1}}(H)$ par la formule (12.647) :

$$\begin{aligned}d\varphi_{\mathbb{1}}(H) &= \frac{d}{dt}\Big[\varphi(\mathbb{1}+tH)\Big]_{t=0} && (17.519a)\\ &= \frac{d}{dt}\Big[(\mathbb{1}+tH^t)A_0(\mathbb{1}+tH)\Big]_{t=0} && (17.519b)\\ &= \frac{d}{dt}\Big[A_0 + tA_0H + tH^tA_0 + t^2H^tA_0H\Big]_{t=0} && (17.519c)\\ &= A_0H + H^tA_0. && (17.519d)\end{aligned}$$

Donc

$$d\varphi_{\mathbb{1}}(H) = (A_0H) + (A_0H)^t. \tag{17.520}$$

Par conséquent

$$\ker(d\varphi_{\mathbb{1}}) = \{H \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } A_0H \text{ est antisymétrique}\}, \tag{17.521}$$

et si nous posons

$$F = \{H \in \mathbb{M}(n,\mathbb{R}) \text{ tel que } A_0H \text{ est symétrique}\} \tag{17.522}$$

nous avons

$$\mathbb{M}(n,\mathbb{R}) = F \oplus \ker(d\varphi_{\mathbb{1}}) \tag{17.523}$$

parce que toute matrice peur être décomposée de façon unique en partir symétrique et antisymétrique. De plus l'application

$$\begin{aligned}f\colon F &\to S_n\\ H &\mapsto A_0H\end{aligned} \tag{17.524}$$

EqGTBusDn

est une bijection linéaire. D'abord $A_0H = 0$ implique $H = 0$ parce que A_0 est inversible, et ensuite si $X \in S_n$, alors $X = A_0A_0^{-1}X$, ce qui prouve que X est l'image par f de $A_0^{-1}X$ et donc que f est surjective.

Maintenant nous considérons la restriction $\psi = \varphi_{|_F}$, $\psi\colon F \to S_n$. Remarquons que $\mathbb{1} \in F$ parce que $A_0 \in S_n$. L'application $d\psi_{\mathbb{1}}$ est une bijection. En effet d'abord

$$d(\varphi_{|_F})_{\mathbb{1}} = (d\varphi_{\mathbb{1}})_{|_F}, \tag{17.525}$$

ce qui prouve que

$$\ker(d\psi_{\mathbb{1}}) = \ker(d\varphi_{\mathbb{1}}) \cap F = \{0\}, \tag{17.526}$$

ce qui prouve que $d\psi_{\mathbb{1}}$ est injective. Pour montrer que $d\psi_{\mathbb{1}}$ est surjective, il suffit de mentionner le fait que $\dim F = \dim S_n$ du fait que l'application (17.524) est une bijection linéaire.

Nous pouvons utiliser le théorème d'inversion locale (théorème 17.47) et conclure qu'il existe un voisinage ouvert U de $\mathbb{1}$ dans F tel que ψ soit un difféomorphisme C^1 entre U et $V = \psi(U)$. Vu que $\mathrm{GL}(n,\mathbb{R})$ est ouvert dans $\mathbb{M}(n,\mathbb{R})$, nous pouvons prendre $U \cap \mathrm{GL}(n,\mathbb{R})$ et donc supposer que $U \subset \mathrm{GL}(n,\mathbb{R})$.

Pour tout $A \in V$, il existe une unique $M \in U$ telle que $\psi(M) = A$, c'est-à-dire telle que $A = M^t A_0 M$. Cette matrice M est $\psi^{-1}(A)$ et est une matrice inversible. Bref, nous posons

$$\begin{aligned} \phi\colon V &\to \mathrm{GL}(n,\mathbb{R}) \\ A &\mapsto \psi^{-1}(A), \end{aligned} \tag{17.527}$$

et ce ϕ est de classe C^1 sur V parce que c'est ce que dit le théorème d'inversion locale. Cette application répond à la question parce que V est un voisinage de $\varphi(\mathbb{1}) = A_0$ et pour tout $A \in V$ nous avons

$$\phi(A)^t A_0 \phi(A) = \varphi^{-1}(A)^t A_0 \varphi^{-1}(A) = A. \tag{17.528}$$

□

17.15 Ellipsoïde de John-Loewner

C'est le moment de relire les conventions de notations données en 17.115 et 17.116 ainsi que la définition 8.125 d'une forme quadratique.

Soit q une forme quadratique sur $\mathbb{R}^n$ ainsi que $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ dans laquelle la matrice de q est diagonale. Dans cette base, la forme q est donnée par la proposition 8.255 :

$$q(x) = \sum_i \lambda_i x_i^2 \tag{17.529}$$

où les λ_i sont les valeurs propres de q.

Plus généralement nous notons $mat_{\mathcal{B}}(q)$ la matrice de q dans la base $\mathcal{B}$ de $\mathbb{R}^n$.

PropOXWooYrDKpw

Proposition 17.119.
Soit $\mathcal{B}$ une base orthonormée de $\mathbb{R}^n$ et l'application[64]

$$\begin{aligned} D\colon Q(\mathbb{R}^n) &\to \mathbb{R} \\ q &\mapsto \det\big(mat_{\mathcal{B}}(q)\big). \end{aligned} \tag{17.530}$$

Alors :

1. *La valeur de D ne dépend pas du choix de la base orthonormée $\mathcal{B}$.*
2. *La fonction D est donnée par la formule $D(q) = \prod_i \lambda_i$ où les λ_i sont les valeurs propres de q.*
3. *La fonction D est continue.*

Démonstration. Soit q une forme quadratique sur $\mathbb{R}^n$. Nous considérons $\mathcal{B}$ une base de diagonalisation de q :

$$q(x) = \sum_i \lambda_i x_i^2 \tag{17.531}$$

où les x_i sont les composantes de x dans la base $\mathcal{B}$. Par définition, la matrice $mat_{\mathcal{B}}(q)$ est la matrice diagonale contenant les valeurs propres de q.

Nous considérons aussi $\mathcal{B}_1$, une autre base orthonormées de $\mathbb{R}^n$. Nous notons $S = mat_{\mathcal{B}_1}(q)$; étant symétrique, cette matrice se diagonalise par une matrice orthogonale : il existe $P \in \mathrm{O}(n,\mathbb{R})$ telle que

$$S = P mat_{\mathcal{B}}(q) P^t; \tag{17.532}$$

64. L'ensemble $Q(E)$ est l'ensemble des formes quadratiques sur E.

donc $\det(S) = \det(PP^t)\det\big(\operatorname{diag}(\lambda_1,\dots,\lambda_n)\big) = \lambda_1\dots\lambda_n$. Ceci prouve en même temps que D ne dépend pas du choix de la base et que sa valeur est le produit des valeurs propres.

Passons à la continuité. L'application déterminant $\det\colon S_n(\mathbb{R}^n) \to \mathbb{R}$ est continue car polynôme en les composantes. D'autre part l'application $mat_{\mathcal{B}}\colon Q(\mathbb{R}^n) \to S_n(\mathbb{R})$ est continue par la proposition 8.142. L'application D étant la composée de deux applications continues, elle est continue. □

LEMooLSTOooZiEOdx

Lemme 17.120 (Volume d'un ellipsoïde[104, 1])**.**
Soit une forme quadratique strictement définie positive q sur $\mathbb{R}^n$. Nous considérons l'ellipsoïde

$$\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant r^2\} \tag{17.533}$$

pour un certain $r > 0$.

Alors le volume de $\mathcal{E}$ est donné par

$$V_q = \frac{V_0 r^n}{\sqrt{D(q)}} \tag{17.534}$$

où V_0 est une constante[65] *et D est l'application de la proposition 17.119.*

Démonstration. Vu que la matrice de q est symétrique, elle est diagonalisable par une matrice orthogonale (théorème 8.227). Autrement dit, il existe une base orthonormée $\mathcal{B} = \{e_1,\dots,e_n\}$ de $\mathbb{R}^n$ telle que

$$q(x) = \sum_{i=1}^{n} a_i x_i^2 \tag{17.535}$$

EqELBooQLPQUj

où $x_i = \langle e_i, x\rangle$ et les a_i sont tous strictement positifs.

À priori nous devrions calculer

$$\int_{\mathcal{E}} dx, \tag{17.536}$$

mais nous effectuons le changement de variable[66] associé à la matrice qui diagonalise q et nous devons simplement calculer

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx \tag{17.537}$$

parce que le jacobien de ce changement de variable est 1 (déterminant d'une matrice orthogonale).

Tout cela pour dire que nous nommons $\mathcal{E}_q$ l'ellipsoïde associée à la forme quadratique q et V_q son volume que nous allons maintenant calculer[67] :

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx \tag{17.538}$$

Cette intégrale est écrite de façon plus simple en utilisant le C^1-difféomorphisme

$$\begin{aligned} \varphi\colon \mathcal{E}_q &\to B(0,1) \\ x &\mapsto \frac{1}{r}\Big(x_1\sqrt{a_1},\dots,x_n\sqrt{a_n}\Big). \end{aligned} \tag{17.539}$$

Le fait que φ prenne bien ses valeurs dans $B(0,1)$ est un simple calcul : si $x \in \mathcal{E}_q$, alors

$$\sum_i \varphi(x)_i^2 = \frac{1}{r^2}\sum_i a_i x_i^2 = \frac{1}{r^2} q(x) < 1. \tag{17.540}$$

65. C'est le volume de la boule unité dans $\mathbb{R}^n$ et ce n'est pas tout à fait évident à calculer [528].
66. Théorème 14.281.
67. Le volume ne change pas si nous écrivons l'inégalité stricte au lieu de large dans le domaine d'intégration ; nous le faisons pour avoir un domaine ouvert.

Cela nous permet d'utiliser le théorème de changement de variables 14.281 :

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx = \frac{r^n}{\sqrt{a_1 \dots a_n}} \int_{B(0,1)} dx. \tag{17.541}$$

La dernière intégrale est le volume de la boule unité dans $\mathbb{R}^n$ et nous la notons V_0. La proposition 17.119 nous permet d'écrire V_q sous la forme

$$V_q = \frac{V_0 r^n}{\sqrt{D(q)}}. \tag{17.542}$$

□

17.121.
Le théorème suivant dit en substance que si K est compact, alors il existe un unique ellipsoïde de volume minimal centré en l'origine et contenant K. Il faut se rendre compte que l'ellipsoïde n'est pas celui que l'on croirait intuitivement parce que la contrainte *centrée en l'origine* est forte. Si $K = \overline{B\big((4,0),1\big)}$, alors l'ellipsoïde donnée n'est pas du tout K lui-même comme on pourrait s'y attendre. Ce serait probablement quelque chose comme la boule centrée en $(0,0)$ et de rayon 5 [68].

Ce que l'on voudrait est un ellipsoïde qui soit centré où il faut pour que le volume soit minimal. Nous verrons que c'est possible en la proposition 17.123, mais qu'alors l'unicité est moins évidente (voir la remarque dans [529]). Si vous voulez en entendre parler, vous pouvez lire [530, 531].

PropJYVooRMaPok

Proposition 17.122 (Ellipsoïde de John-Loewner[104])**.**
Soit K compact dans $\mathbb{R}^n$ et d'intérieur non vide. Il existe une unique ellipsoïde [69] (pleine) centrée en l'origine de volume minimal contenant K.

Démonstration. Nous subdivisons la preuve en plusieurs parties.

(1) **Volume de l'ellipsoïde** Soit $\mathcal{E}$ un ellipsoïde centré en l'origine. La proposition 8.317 et son corolaire 8.318 nous indiquent que

$$\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant 1\} \tag{17.543}$$

pour une certaine forme quadratique strictement définie positive q. Le lemme 17.120 nous donne alors le volume de $\mathcal{E}$ par

$$V_q = \frac{V_0}{\sqrt{D(q)}} \tag{17.544}$$

où V_0 est une constante.

(2) **Existence de l'ellipsoïde**
Nous voulons trouver un ellipsoïde contenant K de volume minimal, c'est-à-dire une forme quadratique $q \in Q^{++}(\mathbb{R}^n)$ telle que

— $D(q)$ soit maximal
— $q(x) \leqslant 1$ pour tout $x \in K$.

Nous considérons l'ensemble des candidats semi-définis positifs.

$$A = \{q \in Q^+ \text{ tel que } q(x) \leqslant 1 \forall x \in K\}. \tag{17.545}$$

Nous allons montrer que A est convexe, compact et non vide dans $Q(\mathbb{R}^n)$; il aura ainsi un maximum de la fonction continue D définie sur $Q(\mathbb{R}^n)$. Nous montrerons ensuite que le maximum est dans Q^{++}. L'unicité sera prouvée à part.

68. Je ne sais pas très bien si il y a moyen de faire mieux. Ce serait sans doute un bon exercice ; faites-moi savoir si vous avez la réponse.
69. Définition 8.315.

(2a) **Non vide** L'ensemble K est compact et donc borné par $M > 0$. La forme quadratique $q\colon x \mapsto \|x\|^2/M^2$ est dans A parce que si $x \in K$ alors

$$q(x) = \frac{\|x\|^2}{M^2} \leqslant 1. \tag{17.546}$$

(2b) **Convexe** Soient $q, q' \in A$ et $\lambda \in [0,1]$. Nous avons encore $\lambda q + (1-\lambda)q' \in Q^+$ parce que

$$\lambda q(x) + (1-\lambda)q'(x) \geqslant 0 \tag{17.547}$$

dès que $q(x) \geqslant 0$ et $q'(x) \geqslant 0$. D'autre part si $x \in K$ nous avons

$$\lambda q(x) + (1-\lambda)q'(x) \leqslant \lambda + (1-\lambda) = 1. \tag{17.548}$$

Donc $\lambda q + (1-\lambda)q' \in A$.

(2c) **Fermé**

Pour rappel, la topologie de $Q(\mathbb{R}^n)$ est celle de la norme (8.292). Nous considérons une suite (q_n) dans A convergeant vers $q \in Q(\mathbb{R}^n)$ et nous allons prouver que $q \in A$, de sorte que la caractérisation séquentielle de la fermeture (proposition 7.252) conclue que A est fermé. En nommant e_x le vecteur unitaire dans la direction x nous avons

$$\big|q(x)\big| = \big|\|x\|^2 q(e_x)\big| \leqslant \|x\|^2 N(q), \tag{17.549}$$

de sorte que notre histoire de suite convergente donne pour tout x :

$$\big|q_n(x) - q(x)\big| \leqslant \|x\|^2 N(q_n - q) \to 0. \tag{17.550}$$

Vu que $q_n(x) \geqslant 0$ pour tout n, nous devons aussi avoir $q(x) \geqslant 0$ et donc $q \in Q^+$ (semi-définie positive). De la même manière si $x \in K$ alors $q_n(x) \leqslant 1$ pour tout n et donc $q(x) \leqslant 1$. Par conséquent $q \in A$ et A est fermé.

(2d) **Borné**

La partie K de $\mathbb{R}^n$ est borné et d'intérieur non vide, donc il existe $a \in K$ et $r > 0$ tel que $\overline{B(a,r)} \subset K$. Si par ailleurs $q \in A$ et $x \in \overline{B(0,r)}$ nous avons $a + x \in K$ et donc $q(a+x) \leqslant 1$. De plus $q(-a) = q(a) \leqslant 1$, donc

$$\sqrt{q(x)} = \sqrt{q\big(x + a - a\big)} \leqslant \sqrt{q(x+a)} + \sqrt{q(-a)} \leqslant 2 \tag{17.551}$$

par l'inégalité de Minkowski 10.12. Cela prouve que si $x \in \overline{B(0,r)}$ alors $q(x) \leqslant 4$. Si par contre $x \in \overline{B(0,1)}$ alors $rx \in \overline{B(0,r)}$ et

$$0 \leqslant q(x) = \frac{1}{r^2} q(rx) \leqslant \frac{4}{r^2}, \tag{17.552}$$

ce qui prouve que $N(q) \leqslant \frac{4}{r^2}$ et que A est borné.

L'ensemble A est compact parce que fermé et borné, théorème de Borel-Lebesgue 9.24. L'application continue $D\colon Q(\mathbb{R}^n) \to \mathbb{R}$ de la proposition 17.119 admet donc un maximum sur le compact A. Soit q_0 ce maximum.

Nous montrons que $q_0 \in Q^{++}(\mathbb{R}^d)$. Nous savons que l'application $f\colon x \mapsto \frac{\|x\|^2}{M^2}$ est dans A et que $D(f) > 0$. Vu que q_0 est maximale pour D, nous avons

$$D(q_0) \geqslant D(f) > 0. \tag{17.553}$$

Donc $q_0 \in Q^{++}$.

(3) **Unicité**

Si il existe une autre ellipsoïde de même volume que celle associée à la forme quadratique q_0, nous avons une forme quadratique $q \in Q^{++}$ telle que $q(x) \leqslant 1$ pour tout $x \in K$. C'est-à-dire que nous avons $q_0, q \in A$ tels que $D(q_0) = D(q)$.
Nous considérons la base canonique $\mathcal{B}_c$ de $\mathbb{R}^n$ et nous posons $S = mat_{\mathcal{B}_c}(q)$, $S_0 = mat_{\mathcal{B}_c}(q_0)$. Étant donné que A est convexe, $(q_0 + q)/2 \in A$ et nous allons prouver que cet élément de A contredit la maximalité de q_0. En effet

$$D\left(\frac{q+q_0}{2}\right) = \det\left(\frac{S+S_0}{2}\right) \tag{17.554}$$

Nous allons utiliser le lemme 17.92 qui dit que le déterminant est log-concave sous la forme de l'équation (17.361) avec $\alpha = \beta = \frac{1}{2}$:

$$D\left(\frac{q+q_0}{2}\right) = \det\left(\frac{S+S_0}{2}\right) > \sqrt{\det(S)}\sqrt{\det(S_0)} = \det(S_0) = D(q_0). \tag{17.555}$$ eqBHJooYEUDPC

Nous avons utilisé le fait que $D(q_0) = D(q)$ qui signifie que $\det(S_0) = \det(S)$. L'inéquation (17.555) contredit la maximalité de $D(q_0)$ et donne donc l'unicité.

□

Dans la proposition suivante nous oublions l'unicité, mais nous démontrons qu'il existe un ellipsoïde de volume minimal parmi les ellipsoïdes centrées où l'on veut et non seulement en zéro. La source de cette proposition est [1], et comme toujours avec cette source, vous devez regarder à la fois l'énoncé et la preuve avec un oeil encore plus prudent que d'habitude.

PROPooVIDPooOGrRJh

Proposition 17.123 ([1])**.**
Soit un compact d'intérieur non vide K dans $\mathbb{R}^n$. Il existe un ellipsoïde de volume minimal contenant K.

Démonstration. Au lieu de considérer le compact K et de chercher où centrer l'ellipsoïde afin qu'elle puisse contenir K en un volume minimal, nous allons chercher comment translater K pour qu'un ellipsoïde centré en zéro puisse contenir l'image translétée de K en un volume minimal.

Pour $a \in \mathbb{R}^n$, nous notons $\mathcal{E}_a$ la plus petite ellipsoïde centrée en 0 contenant $K + a$ (translation de K par le vecteur a). Elle est bien définie par la proposition 17.122. Notre jeu est maintenant d'étudier la fonction [70]

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R}^+ \\ a &\mapsto \mathrm{vol}(\mathcal{E}_a). \end{aligned} \tag{17.556}$$

Nous allons montrer que f est continue et qu'elle peut être restreinte à un compact sans risque de rater un minimum.

(1) **À propos de forme quadratique** Nous notons q la forme quadratique de l'ellipsoïde $\mathcal{E}_0$ et A sa matrice symétrique associée. Si $x \in K + h$ nous avons, pour un certain $k \in K$:

$$\begin{aligned} q(x) &= q(k+h) && \text{(17.557a)}\\ &= \langle k+h, A(k+h)\rangle && \text{(17.557b)}\\ &= \langle k, Ak\rangle + \langle k, Ahx\rangle + \langle h, Ak\rangle + \langle h, Ah\rangle. && \text{(17.557c)} \end{aligned}$$

En tenant compte du fait que $A = A^t$ nous avons aussi $\langle k, Ah\rangle = \langle h, Ak\rangle$ et donc

$$q(x) = 2\langle k, Ah\rangle + q(h). \tag{17.558}$$

En ce qui concerne la norme, pour tout $x \in K + h$ nous avons donc la majoration

$$\begin{aligned} |q(x)| &\leqslant 1 + 2\|k\|\|Ah\| + q(h) && \text{(17.559a)}\\ &\leqslant 1 + 2|K|\|Ah\| + q(h) && \text{(17.559b)} \end{aligned}$$

70. Notons que cette fonction est bien définie, même sans la partie unicité de 17.122 parce que même si $\mathcal{E}_a$ n'était pas bien définie, son volume, lui, est bien défini.

où nous avons noté $|K| = \sup_{k\in K} \|k\|$. Vu que $q(x)$ est en fait toujours positif nous pouvons oublier la valeur absolue à gauche et conclure que $K + h$ est contenu dans l'ellipsoïde

$$\mathcal{F}_1 = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant 1 + 2|K|\|Ah\| + q(h)\}. \tag{17.560}$$

(2) **Volumes**

Nous avons donc $\text{vol}(\mathcal{E}_h) \leqslant \text{vol}(\mathcal{F}_1)$. En reprenant la formule du lemme 17.120 pour le volume de l'ellipsoïde nous avons la majoration

$$\text{vol}(\mathcal{E}_h) \leqslant \text{vol}(\mathcal{E}_0)\big(1 + 2|K|\|Ah\| + q(h)\big)^n. \tag{17.561}$$

Autrement dit, en ce qui concerne notre fonction f :

$$f(h) \leqslant f(0)\big(1 + 2|K|\|Ah\| + q(h)\big)^n. \tag{17.562}$$ EQooLVZCooIVxVNx

Nous pouvons faire le même raisonnement en partant de $K+h$ et en voyant comment modifier le rayon de $\mathcal{E}_h$ pour que $\mathcal{E}_0$ y rentre. Autrement dit, nous refaisons le raisonnement en posant $K' = K + h$ et en étudiant $K' - h$. Si q_1 est la forme quadratique de $\mathcal{E}_h$ et si A_1 est sa matrice symétrique,

$$\text{vol}(\mathcal{E}_0) \leqslant \text{vol}(\mathcal{E}_h)\big(1 + 2|K + h|\|A_1h\| + q_1(h)\big)^n \tag{17.563}$$

En termes de f :

$$f(0) \leqslant f(h)\big(1 + 2|K + h|\|A_1h\| + q_1(h)\big)^n \tag{17.564}$$ EQooRHWNooDisTyn

(3) **Encadrement**

Les majoration (17.562) et (17.564) nous permettent de créer un encadrement pour $f(h)$. En effet, en écrivant (17.564) et en continuant la majoration en remplaçant $f(h)$ par (17.562), nous obtenons

$$\begin{aligned} f(0) &\leqslant f(h)\big(1 + 2|K + h|\|A_1h\| + q_1(h)\big)^n \\ &\leqslant f(0)\big(1 + 2|K|\|Ah\| + q(h)\big)^n\big(1 + 2|K + h|\|A_1h\| + q_1(h)\big)^n. \end{aligned} \tag{17.565}$$

Vu que nous avons dans l'idée de prendre des h petits (en norme) et que les parenthèses tendent vers 1 lorsque h est petit, nous pouvons supposer qu'elles sont non nulles et allègrement les passer au dénominateur. Nous divisons donc tout par le coefficient de $f(h)$:

$$\frac{f(0)}{\big(1 + 2|K + h|\|A_1h\| + q_1(h)\big)^n} \leqslant f(h) \leqslant f(0)\big(1 + 2|K|\|Ah\| + q(h)\big)^n \tag{17.566}$$

Nous utilisons la règle de l'étau [71] pour conclure que

$$\lim_{h\to 0} \text{vol}(\mathcal{E}_h) = \text{vol}(\mathcal{E}_0). \tag{17.567}$$ EQooCMAQooMkkEVR

(4) **Continuité** Si $a \in \mathbb{R}^n$ nous pouvons appliquer la limite (17.567) pour l'ellipsoïde $\mathcal{E}_a$ au lieu de $\mathcal{E}_0$, c'est-à-dire en partant du compact $K + a$ au lieu de K. Cela donne

$$\lim_{h\to 0} \text{vol}(\mathcal{E}_{a+h}) = \text{vol}(\mathcal{E}_a). \tag{17.568}$$

Donc f est continue sur $\mathbb{R}^n$.

(5) **Compact**

Nous avons en hypothèse que K est d'intérieur non vide. Il existe donc $a \in K$ et $r > 0$ tel que $B(a, r) \subset K$. Soit l'hyperplan affine H passant par a et perpendiculaire à a. Nous notons $S = H \cap B(a, r)$. Nous avons évidemment $S \subset K$ et donc $S \subset \mathcal{E}_0$. Vu que l'ellipsoïde $\mathcal{E}_0$ est convexe et contient K, tout le cône de base S et de hauteur a est dans $\mathcal{E}_0$.

71. Théorème 12.221.

Sans rentrer dans le détails du calcul du volume de ce cône, son volume tend vers l'infini si a tend vers l'infini [72]

Nous avons donc une majoration

$$\text{vol}(\mathcal{E}_h) \geqslant \text{volume du cône de base } S \text{ et de hauteur } a+h. \tag{17.569}$$

Dès que ce volume est plus grand que $\text{vol}(\mathcal{E}_0)$, il est impossible que $\text{vol}(\mathcal{E}_h) \leqslant \text{vol}(\mathcal{E}_0)$. Soit $R>0$ tel que pour tout $\|h\| \geqslant R$, le volume du cône soit plus grand que $\text{vol}(\mathcal{E}_0)$. Le minimum de f est certainement dans $B(0,R)$ parce que au moins

$$\text{vol}(\mathcal{E}_0) \in f\big(B(0,R)\big). \tag{17.570}$$

(6) **Minimum**

Nous considérons la fonction

$$\begin{aligned} f\colon \overline{B(0,R)} &\to \mathbb{R}^+ \\ h &\mapsto \text{vol}(\mathcal{E}_h). \end{aligned} \tag{17.571}$$

Elle est continue sur un compact, et le théorème de Weierstrass 7.146 nous dit alors qu'elle atteint son minimum. Il n'y a cependant pas unicité de la valeurs de h pour laquelle $f(h)$ est minimum.

(7) **Envoi**

Soit $s \in \mathbb{R}^n$ tel que $f(s)$ soit minimal. Cela signifie que parmi tous les ellipsoïdes centrés en zéro et contenant des ensembles de la forme $K+h$, le plus petit est celui qui contient $K+s$. Soit $\mathcal{E}_s$ cet ellipsoïde.

Je prétend que $\mathcal{E}_s - s$ est le plus petit ellipsoïde contenant K, tout centres confondus. En effet, si $\mathcal{F}$ est un ellipsoïde centré en a et contenant K, alors $\mathcal{F}-a$ est un ellipsoïde centré en zéro et contenant $K-a$. Nous avons alors

$$\text{vol}(\mathcal{F}-a) \geqslant \text{vol}(\mathcal{E}_s) = \text{vol}(\mathcal{E}_s - s). \tag{17.572}$$

L'invariance de la mesure par translation est le théorème 14.252.

□

17.16 Prolongement de fonctions

LEMooTUQIooEyTLBa

Lemme 17.124 ([1])**.**
Soient un espace topologique X et un ouvert $V \subset X$. Nous considérons une fonction $\phi \in C_c^0(V,\mathbb{C})$. Alors l'extension

$$\begin{aligned} h\colon X &\to \mathbb{C} \\ x &\mapsto \begin{cases} \phi(x) & \text{si } x \in V \\ 0 & \text{sinon} \end{cases} \end{aligned} \tag{17.573}$$

est continue sur X.

LemdCOMQM

Lemme 17.125.
Soit E, un espace vectoriel normé complet et (A_n) une suite emboîtée de fermés non vides dont le diamètre tend vers zéro. Alors l'intersection $\bigcap_{n\in\mathbb{N}} A_n$ contient exactement un point.

Démonstration. Si l'intersection contenait deux points distincts a et b, alors nous aurions pour tout n la majoration $\text{diam}(A_n) \geqslant \|a-b\|$ qui ne dépend pas de n. Cela contredirait la limite.

Soit une suite (x_n) avec $x_k \in A_k$ pour tout $k \in \mathbb{N}$. C'est une suite de Cauchy. En effet si $\epsilon > 0$, considérons N tel que $\text{diam}(A_N) < \epsilon$. Dans ce cas dès que $n,m > N$ nous avons $x_n, x_m \in A_N$ et donc $\|x_n - x_m\| \leqslant \epsilon$. La suite x_n converge donc vers un élément dans E.

72. Topologie du complété en un point, j'imagine que vous voyez de quoi il en retourne. Sinon c'est la définition 7.106, et il faut sans doute adapter le lemme 12.85 au cas de $\mathbb{R}^n$ pour tout faire dans les règles.

Nous devons montrer que $x \in A_k$ pour tout k. La queue de suite $(x_n)_{n\geqslant k}$ est une suite de Cauchy dans A_k qui converge donc vers un élément de A_k (ici nous utilisons le fait que A_k est fermé). Par unicité de la limite, cette dernière doit être x. Par conséquent $x \in \bigcap_{n\in\mathbb{N}} A_n$. □

ThoCaMpKO

Théorème 17.126 ([532]).
Soient X et Y des espaces vectoriels normés. Pour une application linéaire $f\colon X \to Y$, les assertions suivantes sont équivalentes :

(1) f est continue sur X,

(2) f est continue en un point de X,

(3) f est bornée.

PropTTiRgAq

Proposition 17.127 ([533]).
Soit un espace normé X, un espace de Banach F et une partie dense A de X. Si l'application linéaire

$$f\colon \big(A, \|.\|_X\big) \to F \tag{17.574}$$

est continue[73]*, alors il existe une unique application linéaire continue $\tilde{f}\colon X \to F$ prolongeant f. De plus $\|\tilde{f}\| = \|f\|$.*

Démonstration. Soient $x \in X$ et $(A_n)_{n\geqslant 0}$ la suite d'ensembles définie par

$$A_n = \{y \in A \text{ tel que } \|x - y\| \leqslant 2^{-n}\}. \tag{17.575}$$

Étant donné que A est dense, ces ensembles sont tous non vides. De plus $\operatorname{diam} A_n \to 0$ parce que si $y, y' \in A_n$ alors

$$\|y - y'\| \leqslant \|y - x\| + \|x - y'\| \leqslant 2^{-n+1}. \tag{17.576}$$

Vu que f est bornée, la suite d'ensembles $f(A_n)$ est une suite emboitée d'ensembles non vides de F. De plus leur diamètre tend vers zéro. En effet si $z, z' \in f(A_n)$, nous posons $z = f(y)$, $z' = f(y')$ et nous avons

$$\|z - z'\| \leqslant \|f(y) - f(x)\| + \|f(x) - f(y')\| \leqslant \|f\|\big(\|y - x\| + \|x - y'\|\big), \tag{17.577}$$

ce qui montre que $\operatorname{diam} f(A_n) \leqslant \|f\|2^{-n+1}$. Notons que nous avons utilisé la linéarité de f. Par le lemme 17.125, l'intersection $\bigcap_{n\in\mathbb{N}} \overline{f(A_n)}$ contient exactement un point. Nous posons

$$S(x) = \bigcap_{n\in\mathbb{N}} \overline{f(A_n)}. \tag{17.578}$$

Nous allons montrer que l'application $x \mapsto S(x)$ ainsi définie est l'application que nous cherchons.

Nous commençons par montrer que pour toute suite $y_k \to x$ avec $y_k \in A$ nous avons

$$f(y_k) \to S(x). \tag{17.579}$$

EqBnRZxW

Pour cela nous considérons $n_0 \in \mathbb{N}$ et k_0 tel que $y_{k_0} \in A_{n_0}$. Avec cela nous avons

$$\|f(y_k) - S(x)\| \leqslant \operatorname{diam}(A_{n_0}) \leqslant \|f\|2^{-n_0+1}. \tag{17.580}$$

Pour montrer que S est linéaire, nous considérons deux suites dans A : $y_k \to x$ et $y'_k \to x'$ ainsi que la somme $y_k + y'k \to x + x'$. Nous écrivons la relation (17.579) pour ces trois suites :

$$f(y_k) \to S(x) \tag{17.581a}$$

$$f(y'_k) \to S(x') \tag{17.581b}$$

$$f(y_k + y'_x) \to S(x + x'). \tag{17.581c}$$

73. Nous avons bien mis sur A la topologie induite de X. Notons que ce n'est pas toujours celle qui est la plus naturelle sur A.

Cependant, étant donné que f est linéaire, pour tout k nous avons $f(y_k + y'_k) = f(y_k) + f(y'_k)$ et par conséquent

$$f(y_k + y'_k) \to S(x) + S(x'). \tag{17.582}$$

Par unicité de la limite, $S(x + x') = S(x) + S(x')$. Le même genre de raisonnement montre que $S(\lambda x) = \lambda S(x)$. L'application S est donc linéaire.

En ce qui concerna la continuité, nous avons

$$\|S(x)\| = \lim \|f(y_k)\| \leqslant \|f\| \lim \|y_k\| = \|f\|\|x\|, \tag{17.583}$$

donc $\|S\| \leqslant \|f\|$, c'est-à-dire que S est borné et donc continue parce que linéaire (théorème 17.126).

Nous montrons maintenant que S prolonge f. Si $x \in A$, alors nous avons $\bigcap_{n\in\mathbb{N}} f(A_n) = f(x)$, et donc $S(x) = f(x)$. Cela montre du même coup que $\|f\| \leqslant \|S\|$ et que par conséquent $\|f\| = \|S\|$.

Passons à la partie sur l'unicité. Soient donc S et T deux prolongements continus de f sur X. Soient $x \in X$ et $x_n \to x$ une suite dans A. Par continuité nous avons $T(x_n) \to T(x)$ et $S(x_n) \to S(x)$. Étant donné que par ailleurs pour tout n nous avons $S(x_n) = T(x_n)$, l'unicité de la limite montre que $T(x) = S(x)$. □

17.16.1 Encore du prolongement

Dans la même veine que la proposition 17.127 nous avons ce résultat.

ThoPVFQMi

Théorème 17.128 ([534]).
Soient E et F, deux espaces métriques complets ainsi que A dense dans E. Si $u\colon A \to F$ est uniformément continue, alors elle se prolonge de façon unique en une fonction continue $\tilde{u}\colon E \to F$. De plus ce prolongement est uniformément continu.

Démonstration. Soit $x \in E\backslash A$ et une suite (x_n) contenue dans A et convergente vers x. Nous voulons définir

$$\tilde{u}(x) = \lim_{n\to\infty} u(x_n) \tag{17.584}$$

mais pour ce faire nous devons prouver que la suite $\big(u(x_n)\big)$ converge dans F et que la limite ne dépend pas de la suite choisie parmi les suites de A qui convergent (dans E) vers x.

Commençons par montrer que $\big(u(x_n)\big)$ est de Cauchy dans F. Pour cela nous prenons $\epsilon > 0$ et $\eta > 0$ telle que $d_E(a, b) < \eta$ implique $d_F\big(u(a), u(b)\big) < \epsilon$ (uniforme continuité de u). Après, il suffit de choisir N tel que pour tout $n, m > N$ nous ayons $d(x_m, x_n) < \eta$ (parce que x_n est de Cauchy). Avec tout ça nous avons

$$d_F\big(u(x_m), u(x_n)\big) < \epsilon, \tag{17.585}$$

ce qui signifie que $\big(u(x_n)\big)$ est de Cauchy et donc convergente dans F.

Nous voulons montrer maintenant que si (x_n) et (y_n) sont deux suites dans A convergentes vers x alors $\lim_{n\to\infty} u(x_n) = \lim_{n\to\infty} u(y_n)$. Pour cela nous considérons la suite $z = (x_1, y_1, x_2, y_2, \ldots)$. Nous avons évidemment $z_n \to x$, et donc $u(z_n)$ converge dans F par ce qui a été dit plus haut. Mais $u(x_n)$ et $u(y_n)$ en sont deux sous-suites convergentes. Donc leurs limites sont égales.

Il reste à montrer que ce $\tilde{u}$ est continue et uniformément continue. Pour cela nous utilisons le module de continuité et le lemme 10.312. Étant donné que $\tilde{u}$ prolonge u nous avons

EqERYoQN

$$\omega_{\tilde{u}}(h) \geqslant \omega_u(h). \tag{17.586}$$

Soient $h > 0$ et $\epsilon > 0$; soient aussi $x, y \in E$ tels que $d(x, y) < h$. Nous prenons des suites $(a_n) \to x$ et $(b_n) \to y$ tout en choisissant n assez grand pour avoir $d_E(a_n, b_n) < h$. Nous avons

$$d_F\big(\tilde{u}(x), \tilde{u}(y)\big) \leqslant d_F\big(\tilde{u}(x), u(a_n)\big) + d\big(u(a_n), u(b_n)\big) + d_F\big(u(b_n), \tilde{u}(y)\big). \tag{17.587}$$

Si n est assez grand, par construction de $\tilde{u}$, le premier et le dernier terme sont plus petits que ϵ. Par définition du module de continuité nous avons d'autre part $d_F\big(u(a_n), u(b_n)\big) \leqslant \omega_u(h)$. Du coup

$$d_F\big(\tilde{u}(x), \tilde{u}(y)\big) \leqslant \omega_u(h) + 2\epsilon. \tag{17.588}$$

Si nous prenons le supremum sur les x et y vérifiant $d_E(x,y)<h$, à gauche nous obtenons $\omega_{\tilde{u}}(h)$ tandis que le membre de droite ne dépend pas de x et$y$. Donc pour tout ϵ, nous avons

$$\omega_{\tilde{u}}(h)\leqslant\omega_u(h)+2\epsilon. \tag{17.589}$$

En comparaison avec (17.586), nous trouvons

$$\omega_{\tilde{u}(h)}\leqslant\omega_u(h). \tag{17.590}$$

Les fonctions u et $\tilde{u}$ ayant le même module de continuité, le lemme 10.312 nous enseigne que l'une est uniformément continue si et seulement si l'autre l'est. Vu que u est uniformément continue par hypothèse, le prolongement $\tilde{u}$ est uniformément continu. □

DEFooMHQMooEvLccV

Définition 17.129.
*Un **plongement** de l'espace topologique X dans Y est une application $f\colon X\to Y$ telle que $f\colon X\to f(X)$ soit un homéomorphisme.*

ThoPHllyoB

Théorème 17.130 (Extensiton des isométries)**.**
Soit $\tilde{M}$ un espace métrique complet et une application isométrique

$$f\colon A\to\tilde{M} \tag{17.591}$$

où A est une partie dense d'un espace métrique M (pas spécialement complet). Alors f accepte une unique extension isométrique

$$\tilde{f}\colon M\to\tilde{M} \tag{17.592}$$

Supposons de plus que M soit complet[74]*. Alors $\tilde{f}\colon M\to\tilde{M}$ est une bijection si et seulement si $f(A)$ est dense dans $\tilde{M}$.*

Démonstration. Nous commençons par prouver l'unicité. Soient $\tilde{f}_1$ et $\tilde{f}_2$ deux extensions de f et $x\in M$. Si (a_n) est une suite dans A convergeant vers x (possible parce que A est dense dans M), alors nous avons

$$\tilde{f}_1(a_n)=\tilde{f}_2(a_n) \tag{17.593}$$

et donc $\tilde{f}_1(x)=\tilde{f}_2(x)$ par continuité (une application isométrique est continue (proposition 7.246)).

Nous démontrons à présent l'existence.

(1) **Construction de $\tilde{f}$** Soient $x\in M$ et (a_n) une suite dans A qui converge vers x. Nous définissons

$$\tilde{f}(x)=\lim_{k\to\infty}f(a_k). \tag{17.594}$$

EqHEqmbqv

Note : nous pouvons prouver que cette définition ne dépend pas du choix de la suite (a_n) convergeant vers x, mais ce serait superflu parce que nous avons déjà prouvé l'unicité de $\tilde{f}$. Par contre nous devons expliquer pourquoi la limite du membre de droite de (17.594) existe dans $\tilde{M}$. D'abord la suite (a_n) est de Cauchy parce qu'elle est convergente (attention : M n'étant pas complet le fait d'être de Cauchy n'implique pas la convergence). Donc, étant donné que f est une isométrie, la suite $\big(f(a_n)\big)$ est de Cauchy dans $\tilde{M}$. Or ce dernier étant complet, la suite des images converge.

Montrons que cette application $\tilde{f}\colon M\to\tilde{M}$ répond à la question.

(2) **$\tilde{f}$ est isométrique**

74. Il me semble que cette hypothèse manque dans [535].

Soient $a, b \in M$ et des suites dans A convergeant vers eux : $a_n \to a$, $b_n \to b$. Nous avons, par continuité de l'application distance,

$$\begin{aligned} d\big(\tilde{f}(a), \tilde{f}(b)\big) &= \lim_{k\to\infty} d\big(\tilde{f}(a_k), \tilde{f}(b)\big) && (17.595a)\\ &= \lim_{k\to\infty}\lim_{l\to\infty} d\big(\tilde{f}(a_k), \tilde{f}(b_l)\big) && (17.595b)\\ &= \lim_{k\to\infty}\lim_{l\to\infty} d\big(f(a_k), f(b_l)\big) && (17.595c)\\ &= \lim_{k\to\infty}\lim_{l\to\infty} d\big(a_k, b_l\big) && (17.595d)\\ &= d(a,b). && (17.595e) \end{aligned}$$

Cela prouve que $\tilde{f}$ est une isométrie.

Pour la suite nous supposons que M est complet. Notons tout de suite que $\tilde{f}$ est injective parce qu'elle est isométrique.

(3) **Bijection implique dense**

Nous supposons que $\tilde{f}\colon M \to \tilde{M}$ est une bijection. Notons que $f\colon A \to \tilde{M}$ a son image dans $\tilde{M}$. Donc $\overline{f(A)} \subset \tilde{M}$.

Étant donné que $\tilde{f}$ a pour image des limites de suites dans $f(A)$, l'image de $\tilde{f}$ est contenue dans $\overline{f(A)}$, c'est-à-dire $\tilde{f}(M) \subset \overline{f(A)}$. Donc si $\tilde{f}$ est surjective, c'est que $\tilde{M} \subset \overline{f(A)}$. Au final nous avons l'égalité $\overline{f(A)} = \tilde{M}$. Cela prouve que si $\tilde{f}$ est bijective, alors $f(A)$ est dense dans $\tilde{M}$.

(4) **Dense implique bijective**

Nous supposons que $\overline{f(A)} = \tilde{M}$ et nous devons prouver que $\tilde{f}$ est surjective. Soient $x \in \tilde{M}$ et $f(a_n)$ une suite dans $f(A)$ qui converge vers x ; une telle suite existe parce que $f(A)$ est dense dans $\tilde{M}$. Cette suite est de Cauchy dans $\tilde{M}$ parce que dans un espace métrique, une suite convergente est de Cauchy. La suite (a_n) est elle-même également de Cauchy parce que

$$d(a_n, a_m) = d\big(f(a_n), f(a_m)\big). \tag{17.596}$$

Étant donné que (a_n) est de Cauchy dans M, elle converge vers un élément que nous nommons $a \in M$. Par continuité de f nous avons alors

$$\tilde{f}(a) = \lim_{k\to\infty} f(a_k) = x. \tag{17.597}$$

Cela prouve que x est bien dans l'image de $\tilde{f}$ et donc que $\tilde{f}$ est surjective.

□

17.17 Complétion d'un espace métrique

LEMooGJONooVwUgyv

Lemme 17.131 ([536]).
Soient des espaces métriques (X, d) et (Y, ρ). Soit une application isométrique $f\colon X \to Y$. ITEMooKJJHooAEDfIU

(1) Si (x_n) est de Cauchy dans X, alors $\big(f(x_n)\big)$ est de Cauchy dans Y. ITEMooBGWMooQPfTAz

(2) Si (a_n) et (b_n) sont deux suite des Cauchy dans X ayant la même limite, alors $\big(f(a_n)\big)$ et $\big(f(b_n)\big)$ ont même limite.

Démonstration. Soit $\epsilon > 0$. Soit N_ϵ tel que si $p, q \geqslant N_\epsilon$, nous avons $d(x_p, x_q) \leqslant \epsilon$. Pour ce même N_ϵ, vu que f est isométrique,

$$\rho\big(f(x_p), f(x_q)\big) = d(x_p, x_q) \leqslant \epsilon. \tag{17.598}$$

Cela prouve (1). Nous passons à (2).

Supposons $a_n \xrightarrow{X} x$ et $b_n \xrightarrow{X} x$, et nous considérons y_a et y_b les limites $f(a_n) \xrightarrow{Y} y_a$ et $f(b_n) \xrightarrow{Y} y_b$. Nous avons

$$\rho(y_a,y_b) \leqslant \rho\big(y_a,f(a_n)\big) + \underbrace{\rho\big(f(a_n),f(b_n)\big)}_{=d(a_n,b_n)} + \rho\big(f(b_n),y_b\big) \tag{17.599a}$$

$$\leqslant \rho\big(y_a,f(a_n)\big) + d(a_n,x) + d(x,b_n) + \rho\big(f(b_n),y_b\big). \tag{17.599b}$$

À droite, les quatre termes tendent vers zéro lorsque $n \to \infty$. Donc $\rho(y_a,y_b) = 0$, et donc $y_a = y_b$. □

LEMooNFIOooZUMhuA

Lemme 17.132 ([536]).
Soient des espaces métriques (X,d) et (Y,ρ) ainsi que D dense dans X. Nous supposons que Y est dense et nous considérons une application isométrique $f\colon D \to Y$.

Alors il existe une unique isométrie $F\colon X \to Y$ qui étend f.

Démonstration. En plusieurs points.

(1) **Construction de F** Nous commençons par l'existence. Soit $x \in X$. Étant donné que D est dense dans X, il existe une suite (x_n) dans D telle que $x_n \xrightarrow{X} x$. Nous déduisons du lemme 17.131 que $\big(f(x_n)\big)$ est de Cauchy dans Y, et même que la limite ne dépend pas de la suite (x_n) convergeant vers x. Bref, nous pouvons définir

$$\begin{aligned} F\colon X &\to Y \\ x &\mapsto \lim_{n\to\infty} f(x_n) \end{aligned} \tag{17.600}$$

où (x_n) est n'importe quelle suite dans D qui converge vers x. Si $x \in D$, alors il suffit de prendre la suite constante $x_n = x$ pour obtenir que $F(x) = f(x)$.

(2) **F est une isométrie**

Soient $x,y \in X$ ainsi que deux suites de Cauchy $x_n \xrightarrow{X} x$ et $y_n \xrightarrow{X} y$. Soit $\epsilon > 0$. Si n est assez grand,

$$d(x,y) \leqslant d(x,x_n) + d(x_n,y_n) + d(y,y_n) \tag{17.601a}$$

$$\leqslant 2\epsilon + \rho\big(f(x_n),f(y_n)\big) \qquad f \text{ est isom.} \tag{17.601b}$$

$$\leqslant 2\epsilon + \rho\big(f(x_n),F(x)\big) + \rho\big(F(x),F(y)\big) + \rho\big(F(y),f(y_n)\big) \tag{17.601c}$$

$$\leqslant 3\epsilon + \rho\big(F(x),F(y)\big). \tag{17.601d}$$

Donc nous avons l'inégalité

$$d(x,y) \leqslant \rho\big(F(x),F(y)\big). \tag{17.602}$$

En partant maintenant de $\rho\big(F(x),F(y)\big)$ et en effectuant les mêmes majorations, nous trouvons l'inégalité dans l'autre sens : $F\big(F(x),F(y)\big) \leqslant d(x,y)$, de telle sorte que

$$d(x,y) = \rho\big(F(x),F(y)\big), \tag{17.603}$$

ce qui signifie que F est une isométrie.

(3) **Unicité** Soient deux isométries $F\colon X \to Y$ et $F'\colon X \to Y$ telles que $F = F'$ sur D. Nous devons prouver que $F(x) = F'(x)$ pour tout $x \in X$.

Soient donc $x \in X$ ainsi qu'une suite (x_n) dans D telle que $x_n \xrightarrow{X} x$. Nous avons alors

$$\rho\big(F(x),F'(x)\big) \leqslant \rho\big(F(x),F(x_n)\big) + \underbrace{\rho\big(F(x_n),F'(x_n)\big)}_{=0} + \rho\big(F'(x_n),F'(x)\big) \tag{17.604a}$$

$$= d(x,x_n) + d(x_n,x). \tag{17.604b}$$

Donc pour tout n nous avons $\rho\big(F(x),F'(x)\big) \leqslant d(x,x_n) + d(x_n,x)$. En faisant $n \to \infty$ nous trouvons $\rho\big(F(x),F'(x)\big) \leqslant 0$ et donc $\rho\big(F(x),F'(x)\big) = 0$.

□

DEFooXHQFooNBIavy

Définition 17.133 ([536]).
*Soit un espace métrique (X,d). Une **complétion** de (X,d) est un triple $(\hat{X},\hat{d},j)$ où*

(1) L'espace $(\hat{X},\hat{d})$ est un espace métrique complet,

(2) L'application $j\colon X\to\hat{X}$ est une isométrie,

(3) La partie $j(X)$ est dense dans $\hat{X}$.

THOooRLYBooCVBjoP

Théorème 17.134 (Unicité de la complétion[536]).
Soient (X,d) un espace métrique ainsi que deux complétions[75] (X_1,d_1,j_1), (X_2,d_2,j_2). Alors les espaces (X_1,d_1) et (X_2,d_2) sont isométriques.

Démonstration. Nous considérons l'application $f=j_2\circ j_1^{-1}\colon j_1(X)\to X_2$. Cette application est isométrique (parce que j_1 et j_2 le sont), l'espace X_2 est complet et $j_1(X)$ est dense dans X_1. Le lemme 17.132 nous permet de considérer une isométrie $F\colon X_1\to X_2$. Nous devons prouver que F est bijective.

L'application F est injective parce qu'elle est isométrique. En ce qui concerne la surjectivité, soit $y\in X_2$. Étant donné que $j_2(X)$ est dense dans X_2, il existe une suite (y_n) dans X_2 telle que $y_n\xrightarrow{X_2}y$. Pour chaque n nous choisissons $x_n\in X$ tel que $j_2(x_n)=y_n$. Nous posons $z_n=j_1(x_n)\in X_1$.

Vu que (y_n) est de Cauchy et que j_1 est isométrique, la suite (z_n) est de Cauchy. Étant donné que X_1 est complet, la suite (z_n) a une limite dans X_1. Nous pouvons donc poser $z=\lim_{n\to\infty}z_n\in X_1$.

Nous prouvons à présent que $F(z)=y$. L'application $F\colon X_1\to X_2$ est isométrique et donc continue ; elle commute donc avec la limite et

$$\lim_{n\to\infty}d_2\big(F(z_n),y\big)=d_2\big(F(z),y\big).\tag{17.605}$$

Mais $j_1(x)\in j_1(X)$, et nous savons que $F=f$ sur $j_1(X)$. Donc

$$F\big(j_1(x_n)\big)=f\big(j_1(x_n)\big)=j_2(x_n).\tag{17.606}$$

Donc nous avons le calcul

$$\begin{aligned}
d_2\big(F(z),y\big)&=\lim_{n\to\infty}d_2\big(F(z_n),y\big)&(17.607a)\\
&=\lim_{n\to\infty}d_2\big(F(j_1(x_n)),y\big)&(17.607b)\\
&=\lim_{n\to\infty}d_2\big(j_2(x_n),y\big)&(17.607c)\\
&=\lim_{n\to\infty}d_2(y_n,y)&(17.607d)\\
&=0.&(17.607e)
\end{aligned}$$

Cela prouve que $F(z)=y$ et donc que F est surjective. □

Une conséquence du théorème de prolongement est le théorème suivant qui permet de compléter un espace métrique.

ThoKHTQJXZ

Théorème 17.135 (Complétion d'un espace métrique[537, 535, 536]).
Plus précisément, soit (M,d) un espace métrique. Il existe un espace métrique complet $\tilde{M}$ muni d'un plongement[76] isométrique $\varphi\colon M\to\tilde{M}$ tel que $\varphi(M)$ soit dense dans $\tilde{M}$.

Ce complété de M est unique au sens suivant. Si $\tilde{M}_1$ et $\tilde{M}_2$ sont deux espaces métriques complets munis de plongements isométriques $f_i\colon M\to\tilde{M}_i$ dont les images sont denses, alors il existe une bijection isométrique $\phi\colon\tilde{M}_1\to\tilde{M}_2$ telle que $\phi\circ f_1=f_2$.

Démonstration. L'unicité est déjà le théorème 17.134. Nous faisons l'existence.

75. Définition 17.133.
76. Définition 17.129.

(1) **Existence**

Soit C_M l'ensemble des suites de Cauchy de M. Nous définissons

$$\begin{aligned} f\colon C_M \times C_M &\to \mathbb{R} \\ u, v &\mapsto \lim_{n\to\infty} d(u_n, v_n). \end{aligned} \tag{17.608}$$

Notre première tâche est de nous assurer que cela est bien défini, c'est-à-dire que la limite existe toujours. En effet, si u et v sont des suites de Cauchy dans M, nous avons

$$\begin{aligned} |d(u_n, v_n) - d(u_m, v_m)| &\leqslant |d(u_n, v_n) + d(u_n, v_m)| + |d(u_n, v_m) - d(u_m, v_m)| && \text{(17.609a)} \\ &\leqslant d(v_n, v_m) + d(u_n, u_m) && \text{lem. 7.116} \\ & && \text{(17.609b)} \\ &\leqslant 2\epsilon && \text{(17.609c)} \end{aligned}$$

dès que m et n sont assez grand. Cela prouve que la suite $n \mapsto d(u_n, v_n)$ est de Cauchy dans $\mathbb{R}$. Par complétude de $\mathbb{R}$, elle converge [77].

Nous considérons la relation d'équivalence $u \sim v$ si et seulement si $f(u, v) = 0$. Nous posons $\tilde{M} = C_M/\sim$ et nous y mettons la distance

$$d([u], [v]) = f(u, v) \tag{17.610}$$

EqDDLNRNF

et nous devons encore vérifier que cela est bien défini. Prenons $u' \sim u$ et $v' \sim v$. Alors nous avons

$$d(u'_n, v'_n) \leqslant d(u'_n, u_n) + d(u_n, v_n) + d(v_n, v'_n), \tag{17.611}$$

et donc

$$d(u', v') = \lim_{n\to\infty} d(u'_n, v'n) \leqslant \lim_{n\to\infty} d(u_n, v_n) = d(u, v). \tag{17.612}$$

Le même argument en inversant les primes et les non primes montre l'inégalité inverse. Donc $d(u, v) = d(u', v')$ dans C_M, et donc la distance (17.610) est bien définie sur $\tilde{M}$.

Afin de s'assurer que $\tilde{M}$ répond bien à la question du théorème, il faut encore démontrer les points suivants :

— M se plonge isométriquement dans $\tilde{M}$.

— l'image de M par le plongement est dense dans $\tilde{M}$.

— $\tilde{M}$ est complet.

Nous allons maintenant considérer l'application

$$\begin{aligned} \varphi\colon M &\to \tilde{M} \\ x &\mapsto \text{la classe de la suite constante } x. \end{aligned} \tag{17.613}$$

(1a) **Plongement isométrique** Nous allons montrer que cela est une isométrie bijective et que $\varphi(M)$ est dense dans $\tilde{M}$. Le fait que φ soit bijective entre M et $\varphi(M)$ est évident. C'est une isométrie parce que

$$d\big(\varphi(x), \varphi(y)\big) = \lim_{n\to\infty} d\big(\varphi(x)_n, \varphi(y)_n\big) = d(x, y). \tag{17.614}$$

(1b) **Densité**

Soit $[u] \in \tilde{M}$. Tous les termes u_n sont des éléments de M. Nous considérons la suite dans $\varphi(M)$ donnée par

$$a_n = \varphi(u_n) \tag{17.615}$$

77. Ici nous utilisons la complétude de $\mathbb{R}$. Cette dernière doit donc être démontrée indépendamment, ce qui est fait dans le théorème 7.287. De plus nous ne pouvons pas définir $\mathbb{R}$ comme étant le complété de $\mathbb{Q}$ en utilisant ce théorème.

Chaque a_n est un élément[78] de $\tilde{M}$. Montrons que (a_n) converge dans $\tilde{M}$ vers u. Nous avons

$$\begin{aligned} d(a_n, u) &= \lim_{k\to\infty} d\big((a_n)_k, u_k\big) && (17.616a)\\ &= \lim_{k\to\infty} d(u_n, u_k) && (17.616b)\\ &= d(u_n, \ell) && (17.616c)\end{aligned}$$

en notant ℓ la limite de la suite (u_n). Ici nous avons utilisé le fait que la fonction distance était continue pour l'inverser avec la limite, par le théorème 12.219. Nous avons alors

$$\lim_{n\to\infty} d(a_n, [u]) = \lim_{n\to\infty} d(u_n, \ell) = 0. \tag{17.617}$$

(1c) **Complétude** Nous passons maintenant à la preuve du fait que $\tilde{M}$ est complet. Soit (y_n) une suite de Cauchy dans $\tilde{M}$. Soit $\epsilon > 0$; nous définissons $K(n)$ par

$$d\big((y_n)_k, (y_n)_l\big) < \epsilon \tag{17.618}$$

dès que $k, l \geqslant K(n)$. Cette définition fonctionne parce que pour chaque n, y_n est une suite de Cauchy dans M. Nous posons

$$x_n = (y_n)_{K(n)} \in M \tag{17.619}$$

et nous allons montrer que (x_n) est de Cauchy dans M –donc est un élément de $\tilde{M}$– et que $y_k \to (x_n)$ dans $\tilde{M}$.

Nous commençons par montrer que (x_n) est de Cauchy dans M. Nous avons

$$\begin{aligned} d(x_n, x_m) &= d\big((y_n)_{K(n)}, (y_m)_{K(m)}\big) && (17.620a)\\ &\leqslant d\big((y_n)_{K(n)}, (y_n)_l\big) + d\big((y_n)_l, (y_m)_l\big) + d\big((y_m)_l, (y_m)_{K(m)}\big) && (17.620b)\end{aligned}$$

Nous choisissons n, m tels que $d(y_n, y_m) < \epsilon$, ce qui nous permet de choisir l de telle façon à avoir $d\big((y_n)_k, (y_m)_k\big) < \epsilon$ pour tout $k \geqslant l$. De plus, quitte à encore augmenter l, nous supposons que $l > K(m)$ et $l > K(m)$. Avec ces choix nous voyons que $d(x_n, x_m) < 3\epsilon$, ce qui signifie que la suite (x_n) est de Cauchy dans M.

En ce qui concerne la convergence $y_n \to (x)$, on a

$$\begin{aligned} d\big(y_n, (x)\big) &= \lim_{k\to\infty} d\big((y_n)_k, (y_k)_{K(k)}\big) && (17.621a)\\ &\leqslant \lim_{k\to\infty} d\big((y_n)_k, (y_n)_l\big) + \lim_{k\to\infty} d\big((y_n)_l, (y_k)_l\big) + \lim_{k\to\infty} d\big((y_k)_l, (y_k)_{K(k)}\big) && (17.621b)\end{aligned}$$

EqABmqNwo

Nous devons trouver un n tel que si k est suffisamment grand, le tout est majoré par ϵ. Voici nos choix :

— n tel que $d(y_n, y_m) < \epsilon$ dès que $m \geqslant n$,
— $k > n$,
— $k > K(n)$,
— $l > k$,
— $l > K(k)$,
— l suffisamment grand pour que $d\big((y_n)_l, (y_k)_l\big) < \epsilon$.

Avec tous ces choix, les trois termes de (17.621b) sont plus petits que ϵ.

Ceci prouve que $\tilde{M}$ est complet.

Unicité Cette partie de la preuve est encore à écrire. Si vous savez comment faire, écrivez-moi.

□

78. À partir de maintenant nous n'écrivons plus explicitement la classe d'équivalence.

17.17.1 Principe des zéros isolés

ThoukDPBX

Théorème 17.136 (Principe des zéros isolés [538])**.**
Soit un ouvert $\Omega \subset \mathbb{C}$. Soient une fonction analytique[79] *$f : \Omega \to \mathbb{C}$ et $a \in \Omega$, un zéro non isolé de f. Alors f est nulle sur un voisinage de a.*

Démonstration. Nous écrivons f sous la forme d'une série entière[80] autour de a :

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \tag{17.622}$$

EqgryfVJ

valable sur une boule $B(a,r)$. Soit c_m le premier coefficient non nul (si il n'existe pas c'est que f est nulle sur tout $B(a,r)$ et alors le théorème est prouvé). Nous avons alors

$$f(z) = c_m (z-a)^m \big(1 + \sum_{k=1}^{\infty} d_k (z-a)^k\big) \tag{17.623}$$

avec $d_k = c_{k+m}/c_m$. Le rayon de convergence de la série $\sum_k d_k (z-a)^k$ est le même que celui de (17.622) parce que la suite $\frac{d_k r^{m+k}}{c_m}$ reste bornée (critère d'Abel, lemme 15.18). Si nous posons

$$g(z) = 1 + \sum_{k=1}^{\infty} d_k (z-a)^k, \tag{17.624}$$

alors g est une fonction continue et $g(a) = 1$. De plus

$$f(z) = c_m (z-a)^m g(z). \tag{17.625}$$

Soit une suite (z_n) de zéros de f qui converge vers a. Étant donné que g est continue, nous devrions avoir $\lim_{k\to\infty} g(z_k) = g(a) = 1$, mais si $f(z_k) = 0$ avec $z_k \neq a$, alors $g(z_k) = 0$. Cela est un paradoxe qui nous permet de conclure que si la suite z_n existe bien, alors f est identiquement nulle sur un voisinage, c'est-à-dire que tous les c_n sont nuls. □

CORooFBXXooZyfUQi

Corolaire 17.137.
Soit f une fonction analytique sur un ouvert connexe $\Omega \subset \mathbb{C}$. Si f s'annule sur un ouvert (non vide) de Ω, alors f s'annule sur tout Ω.

Démonstration. soit

$$N = \{z \in \Omega \text{ tel que } f = 0 \text{ sur un ouvert autour de } z\}. \tag{17.626}$$

Le fait que N soit ouvert est évident à partir de sa définition. Nous allons montrer que N est également fermé dans Ω, et donc conclure que $N = \Omega$. Soit (z_n) une suite dans N convergente vers $z \in \Omega$. Étant donné que $f(z_n) = 0$ et que f est continue, nous avons

$$f(z) = \lim_{n\to\infty} f(z_n) = 0, \tag{17.627}$$

ce qui fait de z un zéro non isolé de f. Par conséquent le principe des zéros isolés (théorème 17.136) nous enseigne que f s'annule dans un voisinage autour de z, c'est-à-dire que $z \in N$. L'ensemble N est donc fermé. □

ThoAVBCewB

Théorème 17.138 (Principe du prolongement analytique[1])**.**
Soit Ω un ouvert connexe de $\mathbb{C}$. Si deux fonctions analytiques coïncident sur un sous-ensemble D de Ω contenant un point d'accumulation dans Ω, alors elles sont égales sur Ω.

79. Définition 12.454.
80. Définition de fonction analytique 12.454.

Démonstration. Soient deux fonctions analytique $f, g: \Omega \to \mathbb{C}$ égales sur D. Nous considérons la fonction $h = f - g$. Il existe dans D une suite (z_n) qui converge vers $a \in \Omega$ et telle que $f(z_n) = g(z_n)$, c'est-à-dire $h(z_n) = 0$ pour tout n.

Vu que h est continue nous avons aussi $h(a) = 0$. Le point a est donc un zéro non isolé de h. Le théorème 17.136 conclut que $h = 0$ sur Ω. □

17.18 Un petit extra

Ceci provient de notes de TP de l'université libre de Bruxelles autour des années 2000-2001[539].

LEMooXRMAooRADhOM

Lemme 17.139 ([1]).
Soit une fonction $f: \mathbb{R} \to \mathbb{R}$ telle que

ItemExtrai

(1) $f(1) = 1$,

ItemExtraii

(2) $f(x + y) = f(x) + f(y)$ *pour tout réels x et y.*

Alors $f(q) = q$ pour tout $q \in \mathbb{Q}$.

Démonstration. Vu que $f(x) + f(y) = f(x + y)$, nous avons, pour $p \in \mathbb{N}$:

$$\sum_{i=1}^{p} f(\frac{1}{p}) = f(\sum_{i=1}^{p} \frac{1}{p}) = f(1) = 1. \tag{17.628}$$

Donc $f(1/p) = 1/p$. Nous en déduisons tout de suite que $p(p/q) = pf(1/q) = p/q$. □

Si nous ajoutons l'hypothèse de continuité, nous avons un résultat plus fort.

LEMooYKCUooUiIvPJ

Lemme 17.140 ([1]).
Soit une fonction $f: \mathbb{R} \to \mathbb{R}$ telle que

(1) $f(1) = 1$,

(2) $f(x + y) = f(x) + f(y)$ *pour tout réels x et y.*

(3) f *est continue.*

Alors f est la fonction identité : $f(x) = x$ pour tout $x \in \mathbb{R}$.

Démonstration. Nous savons déjà du lemme 17.139 que $f(q) = q$ pour tout $q \in \mathbb{Q}$. Soient $x \in \mathbb{R}$, ainsi qu'une suite q_k dans $\mathbb{Q}$ telle que $q_k \xrightarrow{\mathbb{R}} x$. Par continuité, nous avons

$$f(x) = \lim_{k\to\infty} f(q_k) = \lim_{k\to\infty} f(q_k) = \lim_{k\to\infty} q_k = x \tag{17.629}$$

parce que $f(q_k) = q_k$ pour tout k. □

Une question naturelle qu'on peut alors se poser est la suivante :

Est-il possible de définir une fonction réelle non continue ayant les propriétés (1) et (2) ?

Nous savons qu'une application linéaire sur un espace vectoriel de dimension finie est continue[81], mais qu'en dimension infinie, une application linéaire non continue est possible[82].

Donc une manière de trouver une réponse positive à la question posée plus haut, serait de voir $\mathbb{R}^n$ comme espace vectoriel de dimension infinie.

Nous nous souvenons que, grâce au lemme de Zorn[83], tout espace vectoriel admet une base[84]. En particulier, l'ensemble des réels vu comme espace vectoriel sur $\mathbb{Q}$ admet une base, i.e. $\exists (e_i)_{i\in I}$

81. Proposition 7.180.
82. Exemples 10.71 et 10.72.
83. Lemme 1.23.
84. Proposition 4.24.

des éléments de $\mathbb{R}$ tels que tout réel s'écrit comme combinaison linéaire à coefficients rationnels de ces e_i, i.e.

$$\forall r \in \mathbb{R}, \exists (\lambda_i)_{i\in I} \text{ des éléments de } \mathbb{Q} \text{ tels que } r = \sum_{i\in I} \lambda_i e_i. \tag{17.630}$$

C'est avec ça en main que nous allons faire le travail.

Proposition 17.141.
Soient $e_1 = 1$ ainsi qu'un irrationnel e_2. Soit une $\mathbb{Q}$-base complétée $(e_i)_{i\in I}$ avec[85] *de $\mathbb{R}$.*

Nous définissons $h\colon \mathbb{R} \to \mathbb{R}$ par

$$h(e_1) = e_1 \tag{17.631a}$$
$$h(e_2) = e_3 \tag{17.631b}$$
$$h(e_3) = e_2 \tag{17.631c}$$
$$h(e_i) = e_i \qquad \forall i \in I\backslash\{1,2,3\}, \tag{17.631d}$$

et par $\mathbb{Q}$-linéarité :

$$h(\sum \lambda_i e_i) = \sum \lambda_i \alpha_i \tag{17.632}$$

pour tout $(\lambda_i \in \mathbb{Q})_{i\in I}$.

La fonction h vérifie

ITEMooOXXMooJFSuuN

(1) $h(1) = 1$

(2) *h est $\mathbb{Q}$-linéaire, c'est-à-dire que*

ITEMooAVNMooEsodLR

(2a) $h(x+y) = h(x) + h(y)$ *pour tout $x, y \in \mathbb{R}$*

ITEMooDEJMooXtvrgi

(2b) $h(\lambda x) = \lambda h(x)$ *pour tout $\lambda \in \mathbb{Q}$.*

ITEMooMHNZooBHQywD

(3) *h n'est pas continue (pour la topologie usuelle sur $\mathbb{R}$).*

Démonstration. En plusieurs points

(1) **Pour (1)** C'est la condition $h(e_1) = e_1$ en nous souvenant que $e_1 = 1$.

(2) **Pour (2a)** Nous décomposons x et y dans la base $\{e_i\}_{i\in I}$. Il existe J_x et J_y finis dans I tels que[86] $x = \sum_{i\in J_x} x_i e_i$, $y = \sum_{i\in J_y} y_i e_i$. Vu que toutes les sommes sont finies, nous n'avons aucun problème à écrire

$$x + y = \sum_{i\in J_x \cup J_y} (x_i + y_i) e_i \tag{17.633}$$

et donc

$$h(x+y) = \sum_i h(x_i + y_i) = \sum_i h(x_i) + \sum_i h(y_i) = h(x) + h(y). \tag{17.634}$$

(3) **Pour (2b)** Même type de vérifications que pour (2a).

(4) **Pour (3)** Si h était continue, elle serait l'identité par le lemme 17.140. Or elle n'est pas l'identité parce que $h(e_2) = e_3$.

□

Donc nous avons trouvé une application $\mathbb{Q}$-linéaire $h\colon \mathbb{R} \to \mathbb{R}$ qui n'est pas continue.

85. Théorème de la base incomplète 4.25. Attention : nous ne prétendons pas que I est dénombrable. Par contre, par définition d'une base, les sommes sont toujours finies. Tout réel peut s'écrire comme une combinaison linéaire finie des e_i avec des coefficients rationnels.

86. Les vrais chasseurs de précisions diront que x_i est défini pour $i \in I$, mais uniquement non nul sur J_x.

Chapitre 18

Trigonométrie, isométries

18.1 Trigonométrie

18.1.1 Définitions, périodicité et quelques valeurs remarquables

PROPooZXPVooBjONka

Proposition-Définition 18.1 (Définition du cosinus et du sinus)**.**
La série

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \tag{18.1}$$

définit une fonction $\cos\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞*. Nous l'appelons* ***cosinus****.*

La série

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \tag{18.2}$$

EQooCMRFooCTtpge

définit une fonction $\sin\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞*. Nous l'appelons* ***sinus****.*

Démonstration. La série entière définissant $\cos(x)$ a pour coefficients

$$a_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{18.3}$$

Nous pouvons la majorer par la série entière donnée par les coefficients

$$b_n = \begin{cases} 1/n! & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{18.4}$$

Quelle que soit la parité de k nous avons toujours

$$\left| \frac{b_{k+1}}{b_k} \right| = \frac{1}{k+1}, \tag{18.5}$$

de telle sorte que la formule d'Hadamard (15.68) nous donne $R = \infty$ pour la série $\sum_{k=0}^{\infty} b_k x^k$. A fortiori [1] le rayon de convergence pour la série du cosinus est infini.

L'assertion concernant le sinus se démontre de même.

En ce qui concerne le fait que les fonctions sin et cos sont de classe C^∞ sur $\mathbb{R}$, il faut invoquer le corolaire 15.46. □

LEMooZMJTooJPnyfv

Lemme 18.2 ([1])**.**
Nous avons

SUBEQooTTNNooXzApSM

$$\cos(0) = 1 \tag{18.6a}$$
$$\sin(0) = 0. \tag{18.6b}$$

1. Remarque 15.16.

ainsi que SUBEQSooFOGNooQrBxYc

$$\cos(-x) = \cos(x) \tag{18.6ca}$$
$$\sin(-x) = -\sin(x) \tag{18.6cb}$$
$$\tag{18.6cc}$$

Démonstration. Par substitution directe dans les séries. □

LEMooBBCAooHLWmno

Lemme 18.3.
En ce qui concerne la dérivation, nous avons

$$\sin' = \cos \tag{18.4a}$$
$$\cos' = -\sin. \tag{18.4b}$$

Démonstration. Il s'agit de se permettre de dériver terme à terme (proposition 15.44) les séries qui définissent le sinus et le cosinus. □

LEMooAEFPooGSgOkF

Lemme 18.4.
Les fonctions sinus et cosinus vérifient

$$\cos^2(x) + \sin^2(x) = 1 \tag{18.5}$$

EQooNYCZooApyyBd

pour tout $x \in \mathbb{R}$.

Démonstration. Posons $f(x) = \sin^2(x) + \cos^2(x)$ et dérivons :

$$f'(x) = 2\sin(x)\cos(x) + 2\cos(x)(-)\sin(x) = 0. \tag{18.6}$$

La fonction f est donc constante par le corolaire 12.196. Nous avons donc pour tout x :

$$f(x) = f(0) = \sin^2(0) + \cos^2(0) = 1. \tag{18.7}$$

Le dernier calcul s'obtient en substituant directement x par zéro dans les séries : $\sin(0) = 0$ et $\cos(0) = 1$. □

18.1.2 Fonction puissance (pour les complexes)

La fonction puissance a déjà fait l'objet de nombreuses définitions et extensions. Voir le thème 52. Nous allons maintenant définir a^z pour $a > 0$ et $z \in \mathbb{C}$.

Soit $z = x + iy \in \mathbb{C}$. L'exponentielle $\exp(x + yi)$ est déjà définie en 15.60 ; il suffit donc maintenant de définir les notations e^z et a^z pour $z \in \mathbb{C}$.

DEFooRBTDooNLcWGj

Définition 18.5.
Pour le nombre $e \in \mathbb{R}$ et le nombre imaginaire pur iy ($y \in \mathbb{R}$), nous définissons

$$e^{iy} = \exp(iy) \tag{18.8}$$

où exp *est la série usuelle de la définition 15.60. Pour un nombre complexe général $x + yi$ nous définissons*

$$e^{x+iy} = e^x e^{iy}. \tag{18.9}$$

Et enfin, si $a > 0$ et si $z \in \mathbb{C}$ nous définissons

$$a^z = e^{z\ln(a)}, \tag{18.10}$$

la fonction logarithme ici étant celle $\ln :]0, \infty[\to \mathbb{R}$ *définie par la proposition 15.83.*

Si $z \in \mathbb{C}$ et si $n \in \mathbb{Z}$, la définition de z^n ne pose pas de problème, c'est la définition 1.206.

DefJilXoM

18.6.
Soit $z = x + iy \in \mathbb{C}$. L'exponentielle $\exp(x+yi)$ est déjà définie en 15.60 ; elle est la fonction donnée par

$$\begin{aligned} \exp\colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!}. \end{aligned} \tag{18.11}$$

PROPooXEYFooIEaPvU

Proposition 18.7.
Le rayon de convergence[2] *de la série exponentielle est infini.*

Démonstration. L'exponentielle est la série de puissance dont les coefficients sont donnés par la suite $(a_k) = 1/k!$. Nous utilisons la formule de Hadamard de la proposition 15.25 :

$$\frac{1}{R} = \lim_{n\to\infty} \left| \frac{1/(n+1)!}{1/n!} \right| = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{k} = 0. \tag{18.12}$$

Donc $R = \infty$ par la partie de Hadamard qui dit que si on tombe sur $1/R = 0$, alors c'est $R = \infty$. □

PROPooWSDKooJREQGk

Proposition 18.8.
Pour tout $z \in \mathbb{C}$ nous avons

$$\exp(z) = e^z. \tag{18.13}$$

PropdDjisy

Proposition 18.9 ([540])**.**
Quelques propriétés de l'exponentielle.

(1) Le fonction exp *est continue.*

ITEMooRLHCooJTuYKV

(2) Nous avons la formule $e^{z+w} = e^z e^w$ pour tout $z, w \in \mathbb{C}$.

(3) $(e^z)^{-1} = e^{-z}$

ITEMooIFYFooUniuKS

(4) $(\exp(z))^n = \exp(nz)$.

Démonstration. La proposition 18.7 nous enseigne que le rayon de convergence est infini. La fonction ainsi définie est alors continue par la proposition 12.369.

Les séries $\exp(z)$ et $\exp(w)$ ayant un rayon de convergence infini, nous pouvons utiliser le produit de Cauchy (théorème 15.32) :

$$e^z e^w = \sum_{n=0}^{\infty} \left(\sum_{i+j=n} \frac{z^i w^j}{i!j!} \right) \tag{18.14a}$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \frac{z^i w^{n-i}}{i!(n-i)!} \right) \tag{18.14b}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} z^i w^{n-i} \tag{18.14c}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \tag{18.14d}$$

$$= \exp(z+w). \tag{18.14e}$$

Nous avons utilisé la formule du binôme (proposition 3.43).

Les autres propriétés énoncées sont des corolaires :

$$e^z e^{-z} = e^0 = 1. \tag{18.15}$$

□

2. Définition 15.12.

D'autres propriétés de l'exponentielle sur $\mathbb{C}$, entre autres l'holomorphie, sont données dans le théorème 26.105.

LEMooTDGKooWdpUTD

Lemme 18.10 ([1])**.**
Soient $a > 0$, $z \in \mathbb{C}$ et $n \in \mathbb{Z}$. Alors

$$(a^z)^n = a^{nz}. \tag{18.16}$$

Démonstration. Il s'agit d'un calcul utilisant les propositions 18.9(4) et 18.8 :

$$\begin{aligned}(a^z)^n &= \big(e^{z\ln(a)}\big)^n && (18.17a)\\ &= \exp\big(z\ln(a)\big)^n && (18.17b)\\ &= \exp\big(nz\ln(a)\big) && (18.17c)\\ &= e^{nz\ln(a)} && (18.17d)\\ &= a^{nz}. && (18.17e)\end{aligned}$$

□

18.1.3 Formules de trigonométrie

Le lemme suivant est un premier pas pour le paramétrage du cercle dont nous parlerons dans la proposition 18.61(1).

LEMooHOYZooKQTsXW

Lemme 18.11.
Pour tout $x \in \mathbb{R}$ nous avons :

(1)

$$e^{ix} = \cos(x) + i\sin(x) \tag{18.18}$$

EQooRVPJooTMwNTU

(2) $|e^{ix}| = 1$.

Démonstration. La définition de l'exponentielle sur $\mathbb{C}$ est la définition 15.60. Cette définition fonctionne parce que $\mathbb{C}$ est une algèbre normée, et que $\mathbb{C}$ est un $\mathbb{C}$-module vérifiant l'inégalité $|zz'| \leqslant |z||z'|$ (en l'occurrence, une égalité).

Nous remarquons que i^k vaut 1, i, -1, $-i$. Donc un terme sur deux est imaginaire pur et parmi ceux-là, un sur deux est positif. À bien y regarder, les termes imaginaires purs forment la série du sinus et ceux réels la série du cosinus.

Si vous aimez les formules,

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!} = \cos(y) + i\sin(y). \tag{18.19}$$

Nous avons utilisé le fait que $i^{2n} = (-1)^n$ et $i^{2n+1} = i(-1)^n$. □

NORMooXOVSooEwjecU

18.12 (Formule d'Euler)**.**
En posant $x = \pi$ dans (18.18) nous trouvons la soi-disant géniale formule d'Euler :

$$e^{i\pi} = -1. \tag{18.20}$$

Cette formule fait intervenir les nombres π, e, -1 et i.

En fait non.

Le nombre e n'intervient pas dans la formule d'Euler. Vous vous souvenez de la définition (18.5) ? En fait e^z est juste un abus de notation pour $\exp(z)$.

Lorsque x est réel, nous avons la formule $\exp(x) = e^x$ qui est un théorème tout à fait non trivial liant la fonction exponentielle et la fonction puissance. Il n'y a pas de tel résultat pour z complexe. Dans le cas de z complexe, la formule $\exp(z) = e^z$ est juste une définition de la fonction puissance dans les complexes.

Bref, le nombre e n'intervient pas dans la formule d'Euler.

CORooWZFIooDTCoRo

Corolaire 18.13.
Le complexe conjugué[3] *de* e^{ix} *est* e^{-ix}.

Démonstration. Vu le lemme 18.11, le complexe conjugué de $z = e^{ix}$ est $\bar{z} = \cos(x) - i\sin(x)$. En utilisant (18.6c) nous avons également

$$\bar{z} = \cos(x) - i\sin(x) = \cos(-x) + i\sin(-x) = e^{-ix}. \tag{18.21}$$

□

LEMooJAWBooJGfZIL

Lemme 18.14.
Nous avons les formules d'addition d'angles[4] SUBEQSooFSSMooHcYwRc

EQooJYEMooQaQMib
$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \tag{18.22a}$$
EQooECAUooQzckDv
$$\sin(a+b) = \cos(a)\sin(b) + \sin(a)\cos(b) \tag{18.22b}$$
EQooCVZAooQfocya
$$\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \tag{18.22c}$$
$$\sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b). \tag{18.22d}$$

pour tout a, b *réels.*

Démonstration. Nous utilisons la formule d'addition dans l'exponentielle, proposition (15.291) et la formule (18.18) avant de séparer les parties réelles et imaginaires :

$$e^{i(a+b)} = e^{ia}e^{ib} = \cos(a)\cos(b) - \sin(a)\sin(b) + i\big(\cos(a)\sin(b) + \sin(a)\cos(b)\big). \tag{18.23}$$

Cela est également égal à

$$\cos(a+b) + i\sin(a+b). \tag{18.24}$$

En identifiant les parties réelle et imaginaires, nous obtenons les formules (18.22a) et (18.22c) annoncées.

Pour la formule (18.22c), il suffit de se souvenir que $\sin(-b) = -\sin(b)$ et $\cos(-b) = \cos(b)$ (ces deux égalités sont immédiatement visibles sur les développements en série : l'un a uniquement des puissances paires et l'autre impaires) et d'écrire (18.22a) avec $-b$ au lieu de b. □

CORooQZDQooWjMXTF

Corolaire 18.15.
Les formules suivantes pour les duplications d'angles s'ensuivent :

$$\cos(2a) = \cos^2(a) - \sin^2(a) \tag{18.25a}$$
SUBEQooLRJDooQuEvux
$$\sin(2a) = 2\cos(a)\sin(a). \tag{18.25b}$$

Démonstration. Poser $b = a$ dans les relations du lemme 18.14. □

LEMooPQWWooMdPWUT

Lemme 18.16.
Un sous-groupe de $(\mathbb{R}, +)$ *est soit dense dans* $\mathbb{R}$ *soit de la forme* $p\mathbb{Z}$ *pour un certain réel* $p \neq 0$.

Démonstration. Soit A, un sous-groupe de $(\mathbb{R}, +)$ qui ne soit pas dense. Soit un intervalle $]a, b[$ qui n'intersecte pas A (si vous voulez frimer, vous noterez ici que nous utilisons le fait que les intervalles ouverts forment une base de la topologie de $\mathbb{R}$[5]). Si $d = |b-a|$, l'ensemble A ne contient pas deux éléments séparés par strictement moins de d. Soit p, le plus petit élément strictement positif de A ; nous avons $p \geqslant d$ (parce que $0 \in A$ de toutes façons).

Puisque A est un groupe, nous avons $p\mathbb{Z} \subset A$.

Pour l'inclusion inverse, si $x \in A$ est hors de $p\mathbb{Z}$, il existe un $y \in p\mathbb{Z}$ avec $|x-y| < p$. Et donc le nombre $|x-y|$ est dans A tout en étant plus petit que p. Contradiction. □

3. Définition 1.486.
4. Rien ne nous empêche de donner ce nom à ces formules, mais seriez-vous capable de définir précisément le mot « angle » ?
5. Proposition 7.117.

PROPooFRVCooKSgYUM

Proposition-Définition 18.17 (Périodicité, le nombre π[541])**.**
Plusieurs choses à propos de la périodicité de la fonction cos*.*

(1) La fonction cos *est périodique*[6]*.*

ITEMooVPMWooBqidZG

(2) Un nombre $T > 0$ est une période si et seulement si $\cos(T) = 1$ et $\sin(T) = 0$.

Nous définissons le nombre $\pi > 0$ comme étant la moitié de la période de la fonction cos *:*

$$2\pi = \min\{T > 0 \text{ tel que } \cos(x+T) = \cos(x), \forall x \in \mathbb{R}\}. \tag{18.26}$$

Démonstration. Plusieurs étapes.

(1) **La fonction cosinus n'est pas toujours positive** Supposons d'abord que $\cos(x) > 0$ pour tout $x \in \mathbb{R}$. Dans ce cas, la fonction sin est strictement croissante. Mais les deux fonctions sont bornées par 1 du fait de la formule $\cos^2(x) + \sin^2(x) = 1$. La fonction sin étant croissante et bornée, elle est convergente vers un réel par la proposition 12.56 :

$$\lim_{x\to\infty} \sin(x) = \ell \tag{18.27}$$

pour un certain $\ell > 0$. Avec ça nous avons aussi (pour cause de dérivée) $\lim_{x\to\infty} \sin'(x) = 0$, c'est-à-dire $\lim_{x\to\infty} \cos(x) = 0$. Mais vu que $\cos^2(x) + \sin^2(x) = 1$, nous en déduisons que $\lim_{x\to\infty} \sin(x) = 1$. Mézalor $\lim_{x\to\infty} \cos'(x) = -1$, ce qui veut dire que la fonction cos n'est pas bornée. Cela est impossible. Nous en déduisons que $\cos(x)$ n'est pas toujours positive.

(2) **Il existe $T > 0$ tel que $\cos(T) = 1$ et $\sin(T) = 0$** Par ce que nous venons de faire, il existe $r > 0$ tel que $\cos(r) = 0$. Pour cette valeur, nous avons aussi obligatoirement $\sin(r) = \pm 1$. Nous avons aussi, en utilisant les formules (18.22),

$$\cos(2r) = \cos^2(r) - \sin^2(r) = -1 \tag{18.28a}$$
$$\sin(2r) = 2\cos(r)\sin(r) = 0. \tag{18.28b}$$

et par conséquent

$$\cos(4r) = \cos^2(2r) - \sin^2(2r) = 1 \tag{18.29a}$$
$$\sin(4r) = 2\cos(2r)\sin(2r) = 0. \tag{18.29b}$$

Donc $T = 4r$ fonctionne.

(3) **Si T est une période** Nous entrons dans le vif de la preuve. Soit un $T > 0$ tel que $\cos(x+T) = \cos(x)$ pour tout $x \in \mathbb{R}$. Avec la formule d'addition d'angle dans le cosinus nous cherchons un T tel que

$$\cos(x+T) = \cos(x)\cos(T) - \sin(x)\sin(T) = \cos(x) \tag{18.30}$$

et donc tel que

$$\cos(x)\big(\cos(T) - 1\big) = \sin(x)\sin(T). \tag{18.31}$$

EQooELSAooLNtBnm

Nous dérivons cette équation :

$$-\sin(x)\big(\cos(T) - 1\big) = \cos(x)\sin(T). \tag{18.32}$$

EQooCECFooLrxXaw

Nous multiplions chacune des deux équations (18.31) et (18.32) par $\sin(x)$ et $\cos(x)$ pour obtenir les quatre relations suivantes :

$$\cos^2(x)\big(\cos(T) - 1\big) - \sin(x)\cos(x)\sin(T) = 0 \tag{18.33a}$$
$$-\sin(x)\cos(x)\big(\cos(T) - 1\big) - \cos^2(x)\sin(T) = 0 \tag{18.33b}$$
$$\sin(x)\cos(x)\big(\cos(T) - 1\big) - \sin^2(x)\sin(T) = 0 \tag{18.33c}$$
$$-\sin^2(x)\big(\cos(T) - 1\big) - \sin(x)\cos(x)\sin(T) = 0 \tag{18.33d}$$

SUBEQooLGQXooIrLMLW SUBEQooCHTDooKwyyZF SUBEQooEWPTooTLCUMf SUBEQooGBXTooCFekGJ

En faisant (18.33a) moins (18.33d) nous trouvons $\cos(T) = 1$. Et en sommant (18.33b) avec (18.33c) nous avons $-\sin(T) = 0$.

6. Définition 12.182.

(4) **Si $T > 0$ est tel que $\sin(T) = 0$ et $\cos(T) = 1$** Alors les formules d'addition d'angle du lemme 18.14 donnent tout de suite

$$\cos(x + T) = \cos(x). \tag{18.34}$$

À ce niveau nous croyons avoir prouvé que cos était périodique et que la période est donnée par

$$\min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \tag{18.35}$$

Or rien n'est moins sûr parce qu'il pourrait arriver que ce minimum n'existe pas, c'est-à-dire que l'infimum soit zéro. Autrement dit, il peut arriver que l'ensemble des périodes soit dense. Plus précisément, soit $P \subset \mathbb{R}$ l'ensemble des périodes de cos. C'est un sous-groupe de $(\mathbb{R}, +)$ et le lemme 18.16 nous dit que P est soit dense dans $\mathbb{R}$, soit de la forme $p\mathbb{Z}$ pour un $p > 0$.

Si P est dense, soient $t \in \mathbb{R}$ et une suite (t_n) dans P telle que $t_n \to t$. Pour tout x et tout n nous avons

$$\cos(x + t_n) = \cos(x), \tag{18.36}$$

Comme la fonction cosinus est continue, nous pouvons passer à la limite et écrire $\cos(x+t) = \cos(x)$. Cela étant valable pour tout x et pour tout t, la fonction cosinus est constante. Or nous savons que ce n'est pas le cas, donc P n'est pas dense. Donc cosinus est périodique. □

PROPooKNLAooLwQHea

Proposition 18.18.
La fonction sin *est périodique de période* 2π *et*

$$2\pi = \min\{T > 0 \text{ } tel \text{ } que \text{ } \sin(T) = 0, \cos(T) = 1\}. \tag{18.37}$$

Démonstration. La proposition 18.17 dit que cos est périodique. Puisque $\sin = -\cos'$ par le lemme 18.3, la fonction sin est également périodique par le lemme 12.184. Si T est une période de cos, alors T est une période de sin.

Mais $\sin' = \cos$, de telle sorte que les périodes de sin sont périodes de cos. Bref, T est une période de sin si et seulement si T est une période de cos. □

18.19.
Notons que tout ceci ne nous donne pas la plus petite indication d'ordre de grandeur de la valeur de π. Cela peut encore être 0.1 autant que 500.

PROPooMWMDooJYIlis

Proposition 18.20 ([541, 1])**.**
Des propriétés à la chaine à propos des sinus, cosinus et de leurs périodes.

ITEMooRJZHooCXcKmM

(1) Nous avons

$$2\pi = \min\{x > 0 \text{ } tel \text{ } que \text{ } \cos(x) = 1, \sin(x) = 0\}. \tag{18.38}$$

ITEMooTNHMooUtOjNC

(2) Les fonctions sin *et* cos *sont périodiques de période* 2π*.*

ITEMooSPZBooIQLUXh

(3) Nous avons $\cos(\pi) = -1$ *et* $\sin(\pi) = 0$*.*

(4) Pour tout $a \in \mathbb{R}$ *nous avons*

$$\cos(a + \pi) = -\cos(a) \tag{18.39a}$$
$$\sin(a + \pi) = -\sin(a). \tag{18.39b}$$

ITEMooHDQNooYHVCkg

(5) Nous avons

$$\pi = \min\{x > 0 \text{ } tel \text{ } que \text{ } \cos(x) = -1, \sin(x) = 0\}. \tag{18.40}$$

ITEMooWFNUooYAybDB

(6) Nous avons SUBEQSooBTNPooSvCAHO

$$\begin{cases} \cos(\pi/2) = 0 & (18.41a) \\ \sin(\pi/2) = 1. & (18.41b) \end{cases}$$

ITEMooIRALooBMGOXP

(7) *Nous avons les formules* [EQSooRJZGooCFVqbZ]

$$\begin{cases} \cos(x+\pi/2) = -\sin(x) & (18.42a) \\ \sin(x+\pi/2) = \cos(x) & (18.42b) \end{cases}$$

pour tout $x \in \mathbb{R}$.

ITEMooMQQPooGwOdbt

(8) *Nous avons*

$$\frac{\pi}{2} = \min\{x > 0 \text{ tel que } \sin(x) = 1, \cos(x) = 0\}. \quad (18.43)$$

(9) *Nous avons les valeurs*

$$\begin{cases} \cos(\frac{3\pi}{2}) = 0 & (18.44a) \\ \sin(\frac{3\pi}{2}) = -1. & (18.44b) \end{cases}$$

ITEMooQKPKooEPeHER

(10) *Nous avons*

$$\frac{\pi}{2} = \min\{x > 0 \text{ tel que } \cos(x) = 0\}. \quad (18.45)$$

ITEMooMEXUooGfSInJ

(11) *Pour tout* $x \in]0, \frac{\pi}{2}[$, *nous avons* $\cos(x) > 0$ *et* $\sin(x) > 0$.

Démonstration. C'est parti.

(1) Le fond de la proposition 18.17 est que toutes les périodes $T > 0$ vérifient $\cos(T) = 1$ et $\sin(T) = 0$. La définition de 2π est que c'est la plus petite période.

(2) En utilisant le fait que l'une est la dérivée de l'autre, si T est une période de cos nous avons

$$\begin{aligned} \sin(x+T) &= -\cos'(x+T) & (18.46a) \\ &= -\lim_{\epsilon\to 0} \frac{\cos(x+T+\epsilon) - \cos(x+T)}{\epsilon} & (18.46b) \\ &= -\lim_{\epsilon\to 0} \frac{\cos(x+\epsilon) - \cos(x)}{\epsilon} & (18.46c) \\ &= -\cos'(x) & (18.46d) \\ &= \sin(x). & (18.46e) \end{aligned}$$

Nous déduisons que toute période de cos est une période de sin. De la même façon, nous pouvons prouver l'autre sens : toute période de sin est une période de cos.

(3) D'un côté nous avons

$$\cos(2\pi) = \cos^2(\pi) - \sin^2(\pi) = 1 \quad (18.47)$$

parce que $\cos(2\pi) = \cos(0) = 1$. Puisque $\cos(\pi)$ et $\sin(\pi)$ sont bornés par -1 et 1, nous devons avoir $\sin(\pi) = 0$ et $\cos(\pi) = \pm 1$.

Mais d'un autre côté, le nombre 2π est le plus petit T vérifiant $\cos(T) = 1$, $\sin(T) = 0$. Donc, avoir $\cos(\pi) = 1$ n'est pas possible. Nous concluons

$$\begin{cases} \cos(\pi) = -1 & (18.48a) \\ \sin(\pi) = 0. & (18.48b) \end{cases}$$

(4) Il s'agit d'utiliser les formules d'addition d'angles du lemme 18.14 pour calculer $\cos(a+\pi)$ et $\sin(a+\pi)$ en tenant compte du fait que $\cos(\pi) = -1$ et $\sin(\pi) = 0$.

(5) Soit $a \in]0, \pi[$ tel que $\cos(a) = -1$ et $\sin(a) = 0$. Alors nous avons

$$\begin{aligned} \cos(a+\pi) &= -\cos(\pi) = 1 & (18.49a) \\ \sin(a+\pi) &= -\sin(\pi) = 0, & (18.49b) \end{aligned}$$

ce qui donnerait $a + \pi \in]\pi, 2\pi[$ dont le cosinus est 1 et le sinus est zéro. Mais nous savons déjà que 2π est le minimum pour cette propriété.

(6) Nous avons
$$-1 = \cos(\pi) = \cos^2(\pi/2) - \sin^2(\pi/2), \tag{18.50}$$
donc $\cos(\pi/2) = 0$ et $\sin^2(\pi/2) = 1$, ce qui donne $\sin(\pi/2) = \pm 1$.
Nous devons départager le $\pm$. Pour cela nous savons que $\sin'(0) = \cos(0) = 1$ et que $\sin(0) = 0$, donc il existe $\epsilon > 0$ tel que pour tout $x \in]0, \epsilon[$ nous avons $0 < \cos(x) < 1$ et $0 < \sin(x) < 1$ (nous avons aussi utilisé le lien entre dérivation et croissance de la proposition 12.185). Nous choisissons ϵ plus petit que $\pi/2$.
Supposons que $\sin(\pi/2) = -1$. Le théorème des valeurs intermédiaires 9.93 dit qu'il existe $x_0 \in]\epsilon, \pi/2[$ tel que $\sin(x_0) = 0$. Pour cette valeur de x_0 nous devons aussi avoir $\cos(x_0) = \pm 1$. Mais puisque 2π est minimum pour avoir cos = 1 et sin = 0, nous devons avoir $\cos(x_0) = -1$. Alors nous avons aussi
$$\cos(x_0 + \pi) = \cos(x_0)\cos(\pi) - \sin(x_0)\sin(\pi) = -\cos(x_0) = 1 \tag{18.51a}$$
$$\sin(x_0 + \pi) = \cos(x_0)\sin(\pi) + \sin(x_0)\cos(\pi) = \sin(x_0) = 0. \tag{18.51b}$$
Encore une fois par minimalité de 2π, cela ne va pas. Conclusion : $\sin(\pi/2) = 1$.

(7) Il s'agit encore d'utiliser les formules d'addition d'angle en tenant compte des valeurs remarquables $\cos(\pi/2) = 0$ et $\sin(\pi/2) = 1$.

(8) Supposons $x_0 \in]0, \pi/2[$ tel que $\sin(x_0) = 1$ et $\cos(x_0) = 0$. En utilisant les formules (18.42) nous avons
$$\cos(x_0 + \pi/2) = -1 \tag{18.52a}$$
$$\sin(x_0 + \pi/2) = 0, \tag{18.52b}$$
avec $x_0 + \pi/2 < \pi$. Cela contredirait la minimalité de π.

(9) Il s'agit d'utiliser les formules (18.42) :
$$\cos(\frac{3\pi}{2}) = \cos(\pi + \pi/2) = -\sin(\pi) = 0 \tag{18.53a}$$
$$\sin(\frac{3\pi}{2}) = \sin(\pi + \pi/2) = \cos(\pi) = -1. \tag{18.53b}$$

(10) Si $\cos(x_0) = 0$ alors $\sin(x_0) = -1$ (parce que $\sin(x_0) = 1$ est déjà exclu). Alors $\cos(x_0 + \pi/2) = 1$ et $\sin(x_0 + \pi/2) = 0$, ce qui est également impossible.

(11) La fonction cosinus est continue (proposition 18.1) et $\cos(0) = 1$. Le théorème des valeurs intermédiaires implique que si $\cos(x) \leqslant 0$, alors il existe $t \in]0, x]$ avec $\cos(x) = 0$. Cela n'est pas possible pour $x < \pi/2$, par le point (8).
Le cosinus est positif sur l'intervalle considéré et $\sin'(x) = \cos(x)$. Donc $\sin(0) = 0$ et la dérivée est positive. La proposition 12.185 conclut que sin est strictement croissante et donc, strictement positive.

□
LEMooFESYooBoiuol

Lemme 18.21 (Positivité[1])**.**
À propos de positivité de la fonction cosinus.
ITEMooIXSDooJyCQyb

(1) $\cos(0) = 1$
ITEMooWJEVooGZykbO

(2) $\cos(x) > 0$ *pour* $x \in [0, \pi/2[$.
ITEMooANEPooLGmYtc

(3) $\cos(\pi/2) = 0$.
ITEMooRDWJooZXWyfv

(4) $\cos(x) < 0$ *pour* $x \in]\pi/2, 3\pi/2[$
ITEMooFKPAooBNlvPU

(5) $\cos(3\pi/2) = 0$.
ITEMooIDZGooBTDvDF

(6) $\cos(x) > 0$ *pour* $x \in]3\pi/2, 2\pi]$.

Démonstration. En plusieurs points.

(1) **Pour (1)** C'est déjà fait dans le lemme 18.2.

(2) **Pour (2)** C'est la proposition 18.20(11).

(3) **Pour (3)** C'est la proposition 18.20(6).

(4) **Pas d'annulation entre $\pi/2$ et π** Nous montrons à présent que cos ne s'annule pas entre $\pi/2$ et π. Supposons que $\cos(\frac{\pi}{2}+s)=0$ avec $s\in\,]0,\pi/2[$. Comme $\cos(x)^2+\sin(x)^2=1$ (lemme 18.4), nous avons

$$\begin{cases}\cos(\frac{\pi}{2}+s)=0 & (18.54a)\\ \sin(\frac{\pi}{2}+s)=\epsilon & (18.54b)\end{cases}$$

avec $\epsilon=\pm1$. Utilisant trois fois la proposition 18.20(7) nous trouvons

$$\begin{cases}\cos(x+\frac{3\pi}{2})=\sin(x) & (18.55a)\\ \sin(x+\frac{3\pi}{2})=-\cos(x) & (18.55b)\end{cases}$$

pour tout x. Nous appliquons cela à $x=\frac{\pi}{2}+s$, en nous souvenant que $\cos(x+2\pi)=\cos(x)$ et $\sin(x+2\pi)=\sin(x)$ (par 18.20(2)) :

$$\cos(s)=\cos(\frac{\pi}{2}+s+\frac{3\pi}{2})=\sin(\frac{\pi}{2}+s)=\epsilon \tag{18.56}$$

et

$$\sin(s)=\sin(\frac{\pi}{2}+s+\frac{3\pi}{2})=-\cos(\frac{\pi}{2}+s)=0. \tag{18.57}$$

Si $\epsilon=1$, nous avons une contradiction avec 18.20(1). Si $\epsilon=-1$, nous avons une contradiction avec 18.20(5).

Donc $\cos(x)\neq0$ pour $x\in\,]\frac{\pi}{2},\pi]$.

(5) **$\cos(x)<0$ sur $]\pi/2,\pi]$** Nous savons que $\cos(\pi)=-1$ (18.20(3)). Étant donné que la fonction cos est continue et qu'elle ne s'annule pas sur $]\pi/2,\pi]$, nous en déduisons qu'elle y est partout strictement négative.

(6) **Pour (4), (5), (6)** Il est directement visible sur le développement de définition que $\cos(-x)=\cos(x)$. Et comme $\cos(x+2\pi)=\cos(x)$, nous avons

$$\cos(\pi+s)=\cos(-\pi-s)=\cos(\pi-s). \tag{18.58}$$

Donc toutes les valeurs (et tous les signes) de $\cos(x)$ sur $[\pi,2\pi]$ peuvent être déduits de ceux sur $[0,\pi]$.

□

LEMooPARBooTXbbiB

Lemme 18.22.
Soient $x,y\in\mathbb{R}$.

(1) Nous avons $\cos(x)=\cos(y)$ si et seulement si

$$y\in\{x+2k\pi\}_{k\in\mathbb{Z}}\cup\{-x+2k\pi\}_{k\in\mathbb{Z}}. \tag{18.59}$$

(2) Nous avons $\sin(x)=\sin(y)$ si et seulement si

$$y\in\{x+2k\pi\}_{k\in\mathbb{Z}}\cup\{-x+(2k+1)\pi\}_{k\in\mathbb{Z}}. \tag{18.60}$$

PROPooZULQooBKWrcv

Proposition 18.23.
Les nombres $x, y \in \mathbb{R}$ vérifient SUBEQSooIHUGooAUchjn

$$\begin{cases} \cos(x) = \cos(y) & (18.61a) \\ \sin(x) = \sin(y) & (18.61b) \end{cases}$$

si et seulement si il existe $k \in \mathbb{Z}$ tel que $y = x + 2k\pi$.

Démonstration. En deux parties.

(1) $\Leftarrow$ Si $y = x + 2k\pi$, le résultat est correct parce que la proposition 18.20(2) dit que sin et cos sont périodiques de période 2π.

(2) $\Rightarrow$ Supposons que $x > y$. Nous calculons $\sin(x-y)$ et $\cos(x-y)$ en utilisant les formules du lemme 18.14 et en tenant compte de (18.61). Cela donne $\cos(x-y) = 1$ et $\sin(x-y) = 0$. La proposition 18.17(2) dit alors que $x - y$ est une période de la fonction cos.

Or *la* période de cos est 2π (proposition 18.20(2)). Donc toutes les périodes de cos sont les $2k\pi$ avec $k \in \mathbb{N}$ (lemme 12.183).

□

CORooTFMAooHDRrqi

Corolaire 18.24.
Des nombres $x, y \in \mathbb{R}$ vérifient $e^{ix} = e^{iy}$ si et seulement si il existe $k \in \mathbb{Z}$ tel que $y = x + 2k\pi$.

Démonstration. Le lemme 18.11 donne $e^{ix} = \cos(x) + i\sin(x)$. Donc l'équation $e^{ix} = e^{iy}$ revient au système (18.61) dont les solutions sont bien $y = x + 2k\pi$. □

LEMooBIPFooQNiTqZ

Lemme 18.25 ([1])**.**
À propos de croissance et décroissance des fonctions trigonométriques.

(1) Sur $]0, \pi[$, la fonction cos *est décroissante.*

(2) Sur $]-\pi, 0[$, la fonction cos *est croissante.*

Démonstration. Nous savons que $\cos' = -\sin$ par le lemme 18.3. La liaison entre dérivée et croissance est la proposition 12.185. Les signes de la fonction cosinus sont dans le lemme 18.21. Les signes de la fonction sinus peuvent être déduits de la proposition 18.20(7).

Vous avez tout en main. □

Tout cela nous permet de calculer quelques valeurs remarquables de cosinus et sinus ainsi que d'écrire le tableau de variations de sinus et cosinus.

LEMooIGNPooPEctJy

Lemme 18.26.
Nous avons les valeurs remarquables

$$\sin(\frac{\pi}{4}) = \cos(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}. \tag{18.62}$$

Démonstration. La relation (18.25b) donne

$$0 = \cos(\pi/2) = \cos^2(\pi/4) - \sin^2(\pi/4). \tag{18.63}$$

Donc $\cos^2(\pi/4) = \sin^2(\pi/4)$. Mais puisque $\sin(\pi/4)$ et $\cos(\pi/4)$ sont positifs, ils sont égaux.

Nous avons aussi $\sin^2(\pi/4) + \cos^2(\pi/4) = 1$. Donc le nombre $x = \cos(\pi/4) = \sin(\pi/4)$ vérifie l'équation $2x^2 = 1$, dont l'unique solution positive est $x = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. □

LEMooRMHAooDEAPMw

Lemme 18.27.
Nous avons la valeur remarquable

$$\cos(\frac{\pi}{3}) = \frac{1}{2}. \tag{18.64}$$

Démonstration. Il faut utiliser la formule (18.22a) avec $\cos(\pi) = \cos(2\pi/3 + \pi/3)$ en sachant que $\cos(\pi) = -1$. Ensuite $\cos(2\pi/3) = \cos(\pi/3 + \pi/3)$. En décomposant ainsi, nous exprimons $-1 = \cos(\pi)$ en termes de $\cos(\pi/3)$ et de $\sin(\pi/3)$. En substituant $\sin^2(\pi/3) = 1 - \cos^2(\pi/3)$ nous trouvons que le nombre $\cos(\pi/3)$ vérifie l'équation

$$4x^3 - 3x + 1 = 0. \tag{18.65}$$

Croyez-le ou non, les solutions de cette équation sont $x = -1$ et $x = 1/2$. Allez. Faisons comme si nous le savions pas. En tout cas, ces deux nombres sont des solutions, et nous avons la factorisation [7]

$$4x^3 - 3x + 1 = (2x - 1)^2(x + 1). \tag{18.66}$$

Donc $1/2$ est de multiplicité 2 et -1 de multiplicité 1. Le théorème 3.130 nous dit qu'il n'y a alors pas d'autres racines que ces deux-là [8].

Nous en déduisons que la valeur de $\cos(\pi/3)$ est soit $1/2$ soit -1. La proposition 18.20(5) nous dit qu'il est impossible que $\cos(\pi/3)$ soit égal à -1 parce que $\pi/3 < \pi$. Donc $\cos(\pi/3) = 1/2$ comme annoncé. □

Remarque 18.28.
Vous avez déjà sans doute vu la démonstration de $\cos(30^\circ) = 1/2$ à partir de la figure 18.4. Il n'est pas possible de l'utiliser parce que cela n'est en réalité pas loin d'être la définition de l'angle entre deux droites.

Si vous voulez savoir la définition de l'angle entre deux droites, il faut passer par la définition 18.152, laquelle se base sur le lemme 18.136 qui, elle-même, se base sur la proposition 18.58.

Bref, à notre niveau, nous sommes encore loin de pouvoir faire des raisonnements trigonométriques sur base de géométrie dans les triangles.

PROPooJFAGooYjRJcb

Proposition 18.29.
Pour tout $x \in [0, \pi/4[$ nous avons $\cos(x) > \sin(x)$.

Démonstration. Nous posons $f(x) = \cos(x) - \sin(x)$. Elle vérifie $f(0) = 1$. En utilisant les dérivées du lemme 18.3, nous trouvons

$$f'(x) = -\big(\sin(x) + \cos(x)\big). \tag{18.67}$$

Mais sur $]0, \pi/2[$ nous avons $\cos(x) > 0$ et $\sin(x) > 0$ (proposition 18.20(11)). Donc f est strictement décroissante. Elle ne peut donc passer qu'une seule fois par zéro. Le lemme 18.26 nous indique que $f(\pi/4) = 0$. Donc $f(x) > 0$ sur $[0, \pi/4[$. □

PROPooIBYEooClQttq

Proposition 18.30.
Quelques valeurs trigonométriques.

(1) *Pour le sinus :*
  (1a) $\sin(0) = 0$
  (1b) $\sin(\pi/6) = 1/2$
  (1c) $\sin(\pi/4) = \sqrt{2}/2$
  (1d) $\sin(\pi/3) = \sqrt{3}/2$
  (1e) $\sin(\pi/2) = 1$

(2) *Pour le cosinus :*
  (2a) $\cos(0) = 1$
  (2b) $\cos(\pi/6) = \sqrt{3}/2$
  (2c) $\cos(\pi/4) = \sqrt{2}/2$
  (2d) $\cos(\pi/3) = 1/2$
  (2e) $\cos(\pi/2) = 0$

(3) *Pour la tangente :*
  (3a) $\tan(0) = 0$
  (3b) $\tan(\pi/6) = \sqrt{3}/3$
  (3c) $\tan(\pi/4) = 1$
  (3d) $\tan(\pi/3) = \sqrt{3}$
  (3e) $\tan(\pi/2)$ *est non défini.*

Démonstration. Plusieurs ont déjà été faites. Les autres ne seront pas démontrées dans l'ordre énoncé.

7. Factorisation d'un polynôme en sachant des racines, proposition 3.123.
8. Nous attirons votre attention sur le fait que cela n'est en aucun cas une trivialité.

(1) $\sin(0) = 0$ Substitution dans la définition (18.2).

(2) $\sin(\pi/4) = \sqrt{2}/2$ C'est le lemme 18.26.

(3) $\sin(\pi/3) = 1/\sqrt{2}$ Nous utilisons la formule $\sin^2(x) + \cos^2(x) = 1$ avec $x = \pi/3$. Vu que $\cos(\pi/3) = 1/2$ (lemme 18.27), cela donne $\sin^2(\pi/3) = 3/4$. Nous en déduisons que $\sin(\pi/3)$ vaut $\pm\frac{\sqrt{3}}{2}$.
La proposition 18.20(5) nous dit que sin est positive sur $[0, \pi]$. Donc c'est bien la possibilité $\sqrt{3}/2$ qui est la bonne.

(4) $\sin(\pi/6) = 1/2$ **et** $\cos(\pi/6) = \sqrt{3}/2$ Nous partons de l'équation (18.25b) pour écrire

$$\sin(\pi/3) = 2\cos(\pi/6)\sin(\pi/6). \tag{18.68}$$

Nous avons déjà vu que $\sin(\pi/3) = \sqrt{3}/2$. En posant $x = \sin(\pi/6)$ nous avons également $\cos(\pi/6) = \sqrt{1 - x^2}$ parce que nous savons que la fonction cosinus est positive sur $[0, \pi/2]$ (proposition 18.20(11)). Nous avons donc l'équation

$$\frac{\sqrt{3}}{2} = 2x\sqrt{1 - x^2}. \tag{18.69}$$

Nous passons au carré et posons $y = x^2$. Après quelque manipulations,

$$16y^2 - 16y + 3 = 0. \tag{18.70}$$

Cela donne deux possibilités pour y : $\frac{3}{4}$ et $\frac{1}{4}$. Puisque $x > 0$, nous pouvons simplement passer à la racine carrée : $x = \sqrt{3}/2$ ou $x = 1/2$.
Notez que si nous avion posé $x = \cos(\pi/6)$ au lieu de $x = \sin(\pi/6)$, nous aurions obtenu le même résultat. Donc $\sin(\pi/6)$ et $\cos(\pi/6)$ peuvent tous deux avoir les valeurs $\sqrt{3}/2$ ou $1/2$. Cela fait 4 possibilités.
Étant donné que $\sin^2(\pi/6) + \cos^2(\pi/6) = 1$, les deux possibilités avec $\sin(\pi/6) = \cos(\pi/6)$ sont exclues.
La proposition 18.29 nous dit aussi que $\cos(\pi/6) > \sin(\pi/6)$. Donc $\cos(\pi/6) = \sqrt{3}/2$ et $\sin(\pi/6) = 1/2$.

(5) $\sin(\pi/2) = 1$ C'est dans (18.41).

(6) $\cos(0) = 1$ Substitution dans la définition.

(7) $\cos(\pi/6) = \sqrt{3}/2$ Déjà fait avec le sinus de $\pi/6$.

(8) $\cos(\pi/4) = \sqrt{2}/2$ Lemme 18.26.

(9) $\cos(\pi/3) = 1/2$ Lemme 18.27.

(10) $\cos(\pi/2) = 0$ Dans (18.41).

Toutes les valeurs pour la tangente s'obtiennent maintenant par la définition, en calculant $\tan(x) = \frac{\sin(x)}{\cos(x)}$. □

Voici un tableau qui rappelle les valeurs à retenir pour les fonctions sinus, cosinus et tangente.

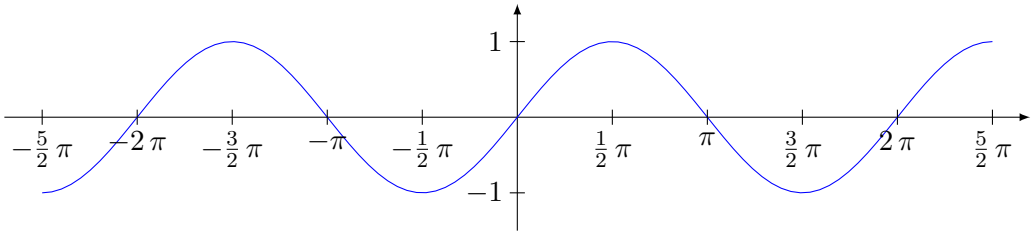
x	$\sin(x)$	$\cos(x)$	$\tan(x)$
0	0	1	0
$\pi/6$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$	1
$\pi/3$	$\sqrt{3}/2$	$1/2$	$\sqrt{3}$
$\pi/2$	1	0	N.D.

(18.71)

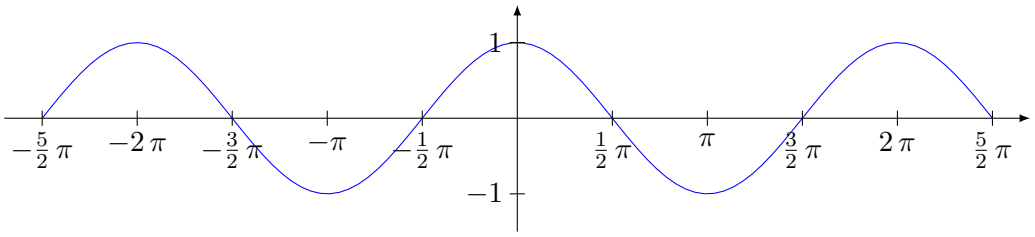
PGooIMQFooTnBdU

où « N.D. » signifie « non défini ».

Rappelons le graphe de la fonction sinus :



celui de la fonction cosinus :



developcosenpisur3

Exemple 18.31.
Développer la fonction cos autour de $x = \frac{\pi}{3}$. Utiliser la valeur remarquable du lemme 18.27. Nous développons autour de $h = 0$ la fonction $\cos(\frac{\pi}{3} + h)$:

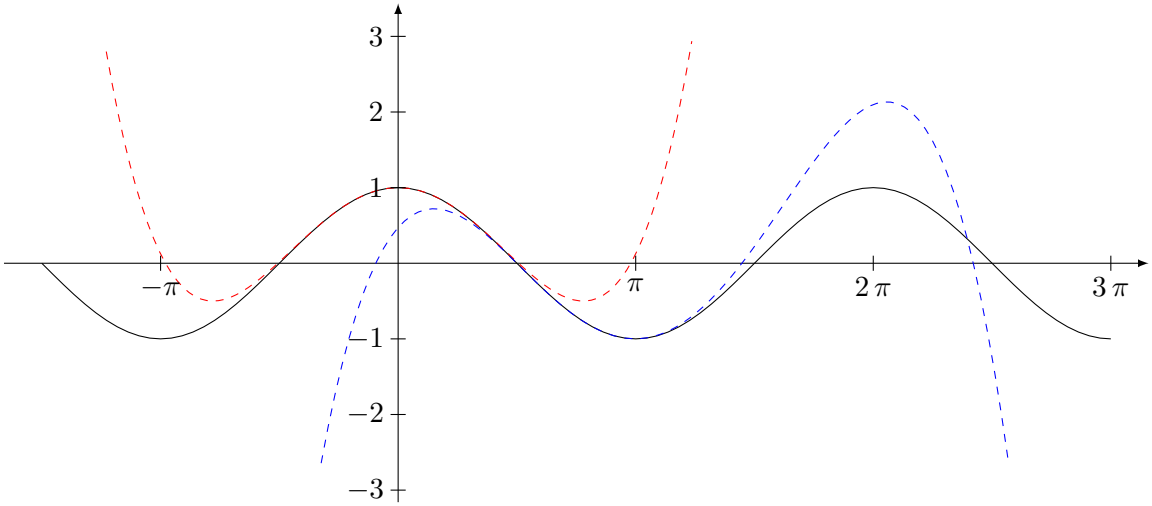
$$\cos\left(\frac{\pi}{3} + h\right) \sim \cos\left(\frac{\pi}{3}\right) + h\cos'\left(\frac{\pi}{3}\right) + \frac{h^2}{2}\cos''\left(\frac{\pi}{3}\right) = \frac{1}{2} - \frac{\sqrt{3}}{2}h - \frac{1}{4}h^2. \tag{18.72}$$

Il est aussi possible d'écrire cela en notant $x = x_0 + h$, c'est-à-dire en remplaçant h par $x - \frac{\pi}{3}$:

$$\cos(x) \sim \frac{1}{2} - \frac{\sqrt{3}}{2}(x - \frac{\pi}{3}) - \frac{1}{4}(x - \frac{\pi}{3})^2. \tag{18.73}$$

△

18.32.
Voici un petit dessin pour donner une idée.



(1) En noir le graphe de $\cos(x)$.
(2) En rouge, le développement de $\cos(x)$ à l'ordre 4 autour de $x = 0$.
(3) En bleu, le développement de $\cos(x)$ à l'ordre 4 autour de $x = 3\pi/4$.

18.2 Trucs et astuces de calcul d'intégrales

SECooKSOFooEVKDLh

Afin d'alléger le texte de calculs parfois un peu longs, nous regroupons ici les intégrales à une variable.

18.2.1 Quelques intégrales « usuelles »

ItemIntegrali

(1) L'intégrale

$$\boxed{I = \int x\ln(x)dx = \frac{x^2}{2}\big(\ln(x) - \frac{1}{2}\big)} \tag{18.74}$$

se fait par partie en posant

$$\begin{aligned} u &= \ln(x), \quad dv = x\,dx \\ du &= \frac{1}{x}\,dx, \qquad v = \frac{x^2}{2}, \end{aligned} \tag{18.75}$$

et ensuite

$$I = \ln(x)\frac{x^2}{2} - \int \frac{x}{2} = \frac{x^2}{2}\big(\ln(x) - \frac{1}{2}\big). \tag{18.76}$$

(2) L'intégrale

$$\boxed{I = \int x\ln(x^2)dx = x^2\ln(x) - \frac{x^2}{2}.} \tag{18.77}$$

En utilisant le fait que $\ln(u^2) = 2\ln(u)$, nous retombons sur une intégrale du type (1) :

$$I = x^2\ln(x) - \frac{x^2}{2}. \tag{18.78}$$

(3) L'intégrale

$$\boxed{I = \int x\ln(1+x^2)dx = \frac{1}{2}\ln(x^2+1)(x^2+1) - x^2 - \frac{1}{2}} \tag{18.79}$$

EqTrucIntxlnxsqpun

se traite en posant $v = 1 + x^2$ de telle sorte à avoir $dx = \frac{dv}{2x}$ et donc

$$I = \frac{1}{2}\ln(x^2+1)(x^2+1) - x^2 - \frac{1}{2}. \tag{18.80}$$

(4) L'intégrale

$$I = \int \cos(\theta)\sin(\theta)\ln\left(1 + \frac{1}{\cos^2(\theta)}\right)d\theta \tag{18.81}$$

demande le changement de variable $u = \cos(\theta)$, $d\theta = -\frac{du}{\sin(\theta)}$. Nous tombons sur l'intégrale

$$I = -\int u\ln\left(\frac{1+u^2}{u^2}\right) = -\int u\ln(1+u^2) + \int u\ln(u^2), \tag{18.82}$$

qui sont deux intégrales déjà faites. Nous trouvons

$$I = -\frac{1}{2}\ln\left(\frac{\sin^2(\theta)-1}{\sin^2(\theta)-2}\right)\sin^2(\theta) - \ln\big(\sin^2(\theta)-2\big) + \frac{1}{2}\ln\big(\sin^2(\theta)-1\big) \tag{18.83}$$

(5) L'intégrale

$$\boxed{\int \frac{r^3}{1+r^2}dr = \frac{r^2}{2} - \frac{1}{2}\ln(r^2+1).} \tag{18.84}$$

commence par faire la division euclidienne de r^3 par r^2+1 ; ce que nous trouvons est $r^3 = (r^2+1)r - r$. Il reste à intégrer

$$\int \frac{r^3}{1+r^2}dr = \int r\,dr - \int \frac{r}{1+r^2}dr. \tag{18.85}$$

La fonction dans la seconde intégrale est $\frac{r}{1+r^2} = \frac{1}{2}\frac{f'(r)}{f(r)}$ où $f(r) = 1+r^2$, et donc $\int \frac{r}{1+r^2} = \frac{1}{2}\ln(1+r^2)$. Au final,

$$I = \frac{1}{2}r^2 - \frac{1}{2}\ln(r^2+1). \tag{18.86}$$

(6) L'intégrale

$$\boxed{I = \int \cos(\theta)\sin(\theta)d\theta = \frac{\sin^2(\theta)}{2}} \tag{18.87}$$

EqTrucIntsxcosdx

se traite par le changement de variable $u = \sin(\theta)$, $du = \cos(\theta)d\theta$, et donc

$$\int \cos(\theta)\sin(\theta)d\theta = \int u du = \frac{u^2}{2} = \frac{\sin^2(\theta)}{2}. \tag{18.88}$$

(7) L'intégrale

$$\boxed{\int \sqrt{1+x^2}dx = \frac{x}{2}\sqrt{1+x^2} + \frac{1}{2}\operatorname{arcsinh}(x)} \tag{18.89}$$

EqTrucsIntsqrtAplusu

s'obtient en effectuant le changement de variable $u = \sinh(\xi)$.

(8) L'intégrale

$$\boxed{\int \cos^2(x)\sin^2(x)dx = \frac{x}{8} - \frac{\sin(4x)}{32}} \tag{18.90}$$

EqTrucIntcossqsinsq

s'obtient à coups de formules de trigonométrie. D'abord, la formule (18.25b) permet de récrire la fonction à intégrer sous la forme

$$f(x) = \frac{1}{4}\sin^2(x). \tag{18.91}$$

Ensuite nous utilisons le fait que $\sin^2(t) = (1-\cos(2t))/2$ pour transformer la formule à intégrer en

$$f(x) = \frac{1-\cos(4x)}{8}. \tag{18.92}$$

Cela s'intègre facilement en posant $u = 4x$, et le résultat est

$$\int f(x)dx = \frac{x}{8} - \frac{\sin(4x)}{32}. \tag{18.93}$$

(9) La fonction

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x} \tag{18.94}$$

est le **sinus cardinal** de x. Nous allons montrer que

$$\boxed{\int_0^\infty \big|\operatorname{sinc}(x)\big|dx = \infty}. \tag{18.95}$$

EqKNQmLEq

D'abord nous avons

$$\int_{(n-1)\pi}^{n\pi} \frac{\big|\sin(t)\big|}{t}dt \geqslant \int_{(n-1)\pi}^{n\pi} \frac{\big|\sin(t)\big|}{n\pi}dt, \tag{18.96}$$

mais par périodicité,

$$\int_{(n-1)\pi}^{n\pi} \big|\sin(t)\big|dt = \int_0^\pi \sin(t)dt = 2. \tag{18.97}$$

Par conséquent

$$\int_0^{n\pi} \big|\operatorname{sinc}(t)\big|dt \geqslant \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k}, \tag{18.98}$$

ce qui diverge lorsque $n \to \infty$.

(10) Les intégrales, pour $\epsilon > 0$,

$$\boxed{\int_0^\infty \cos(kx)e^{-\epsilon x}dx = \frac{\epsilon}{k^2+\epsilon^2}} \tag{18.99}$$

EQooNCVIooWqbbrH

et

$$\boxed{\int_0^\infty \sin(kx)e^{-\epsilon x}dx = \frac{k}{k^2+\epsilon^2}} \tag{18.100}$$

EQooSAYIooSatbGc

se calculent deux fois par partie. Nous posons

$$I = \int_0^\infty \cos(kx)e^{-\epsilon x}dx \tag{18.101a}$$

$$J = \int_0^\infty \sin(kx)e^{-\epsilon x}dx. \tag{18.101b}$$

L'intégrale I s'effectue par partie en posant $u = \cos(kx)$ et $v' = e^{-\epsilon x}$. Un peu de calcul montre que

$$I = \frac{1}{\epsilon} - \frac{k}{\epsilon}J. \tag{18.102}$$

Par ailleurs l'intégrale J se fait également par partie pour obtenir

$$J = \frac{k}{\epsilon}I. \tag{18.103}$$

En résolvant pour I et J les deux équations déduites, nous trouvons

$$I = \frac{\epsilon}{k^2+\epsilon^2} \tag{18.104a}$$

$$J = \frac{k}{k^2+\epsilon^2}. \tag{18.104b}$$

18.2.2 Reformer un carré au dénominateur

subsecCarreDenoPar

Lorsqu'on a un second degré au dénominateur, le bon plan est de reformer un carré parfait. Par exemple :

$$x^2 + 2x + 2 = (x+1)^2 + 1. \tag{18.105}$$

Ensuite, le changement de variable $t = x + 1$ est pratique parce que cela donne $t^2 + 1$ au dénominateur.

Cherchons

$$I = \int \frac{1-x}{x^2+2x+2}dx = \int \frac{1-x}{(x+1)^2+1}dx = \int \frac{1-(t-1)}{t^2+1} \tag{18.106}$$

où nous avons fait le changement de variable $t = x + 1$, $dt = dx$. L'intégrale se coupe maintenant en deux parties :

$$I = \int \frac{-t}{t^2+1} + \int \frac{2}{t^2+1}. \tag{18.107}$$

La seconde est dans les formulaires et vaut

$$2\arctan(t) = 2\arctan(x+1), \tag{18.108}$$

tandis que la première est presque de la forme f'/f :

$$\int \frac{t}{t^2+1} = \frac{1}{2}\int \frac{2t}{t^2+1} = \frac{1}{2}\ln(t^1+1) = \frac{1}{2}\ln(u^2+2u+2). \tag{18.109}$$

18.2.3 Décomposition en fractions simples

La décomposition en fractions simples décrite en 19.1.3 permet d'intégrer des fractions rationnelles. Elle peut parfois être évitée par la méthode de Rothstein-Trager que nous expliquerons dans 20.15.3.

18.3 Très modeste approximation de π

Nous sommes en droit de vouloir une valeur approchée de π.

LEMooJWSGooExmtDA

Lemme 18.33.
Nous avons l'encadrement

$$2\sqrt{2} < \pi < 4. \tag{18.110}$$

Démonstration. Grace au lemme 18.26 nous savons que la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivée majorée par 1. Par conséquent

$$\frac{\pi}{4} > \frac{\sqrt{2}}{2} \tag{18.111}$$

et donc [9]

$$\pi > 2\sqrt{2} \simeq 2.82 \tag{18.112}$$

De plus la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivée minorée par $\sqrt{2}/2$, donc

$$\frac{\pi}{4} < \frac{\sqrt{2}/2}{\sqrt{2}/2}, \tag{18.113}$$

ce qui donne

$$\pi < 4. \tag{18.114}$$

□

Pour avoir une meilleur approximation de π, nous pouvons remarquer que $\pi \in]2.82, 4[$, et que cet intervalle est suffisamment petit pour ne pas recouvrir l'intervalle correspondant pour 2π. L'équation $\cos(x) = -1$ possède donc une unique solution dans cet intervalle (et cette solution est π). Nous pouvons donc faire une dichotomie pour trouver la valeur de π, pourvu que nous ayons une façon d'évaluer des valeurs de $\cos(x)$ de façon pas trop ridicule.

LEMooIECUooQOGYyN

Lemme 18.34.
Pour toute valeur de $x \in \mathbb{R}$ on a $|\sin(x)| \leqslant |x|$.

Démonstration. Nous commençons par les valeurs $x > 0$. Considérons la fonction

$$f(x) = x - \sin(x). \tag{18.115}$$

La dérivée de f vaut

$$f'(x) = 1 - \cos(x) \geqslant 0, \tag{18.116}$$

De plus $f(0) = 0$. Vu que f est croissante, nous avons $f(x) \geqslant 0$ pour tout $x \geqslant 0$.

(1) **Pour $0 \leqslant x \leqslant \pi$** Nous avons $|x| = x$ et $|\sin(x)| = \sin(x)$, et ce que nous venons de dire à propos de f suffit pour conclure.

(2) **Pour $x \geqslant \pi$** En utilisant l'approximation de π du lemme 18.33, nous avons $\pi > 2\sqrt{2} > 2$ et donc [10]

$$|x| - |\sin(x)| > 2 - 1 = 1 > 0. \tag{18.117}$$

(3) **Pour $x < 0$** Si $x \leqslant 0$, ce que nous venons de faire nous assure que $\sin(-x) \leqslant -x$. Par ailleurs, nous savons déjà par (18.6c) que $\sin(-x) = -\sin(x)$. Nous avons donc

$$|\sin(x)| = \sin(-x) \leqslant -x = |x|. \tag{18.118}$$

□

9. Sérieusement, êtes vous capables de trouver une approximation de $\sqrt{2}$ en ne vous basant que sur des choses vues jusqu'ici ?

10. Ouais ; l'approximation $\pi > 2$ nous suffit. On est large.

18.3.1 Les fonctions tangente et arc tangente

Définition 18.35.
La fonction ***tangente*** *est :*

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \tag{18.119}$$

où sin *et* cos *sont de la définition 18.1.*

La fonction tangente n'est pas définie sur les points de la forme $x = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$. Une interprétation géométrique, qui justifie le nom, est donnée sur la figure 18.1.

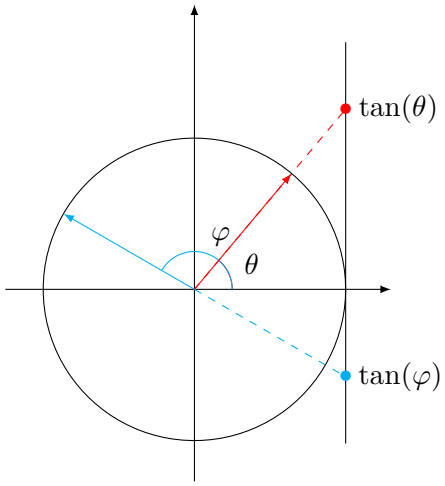


FIGURE 18.1: Interprétation géométrique de la fonction tangente. La tangente de l'angle θ est positive (et un peu plus grande que 1) tandis que celle de la tangente de l'angle φ est négative. labelFig TgCercle Trigono

LEMooYAMFooXCLXfr

Lemme 18.36.
La dérivée de la fonction tangente vaut

$$\tan'(x) = \tan^2(x) + 1, \tag{18.120}$$

Démonstration. Il suffit de suivre les règles de calcul comme la dérivation du quotient 12.172(5). □

PROPooTLHUooPpqupE

Proposition 18.37.
La fonction

$$\begin{aligned} \tan\colon]-\frac{\pi}{2}, \frac{\pi}{2}[&\to \mathbb{R} \\ x &\mapsto \tan(x) \end{aligned} \tag{18.121}$$

est une bijection.

Démonstration. Le cosinus ne s'annulant pas sur l'intervalle donné, la fonction est bien définie. Nous avons

$$\lim_{x \to \pi/2^-} \tan(x) = +\infty \tag{18.122}$$

parce que la limite du sinus est 1 est celle du cosinus est zéro par les valeurs positives. Le même raisonnement donne la limite en $-\pi/2$ qui vaut $-\infty$. Le théorème des valeurs intermédiaires [11] dit que la fonction tangente est alors surjective sur $\mathbb{R}$.

Par ailleurs le lemme 18.36 nous donne la dérivée de la tangente sous la forme $\tan'(x) = \tan^2(x) + 1$, ce qui nous donne une dérivée partout strictement positive, et donc une fonction strictement croissante et donc injective. □

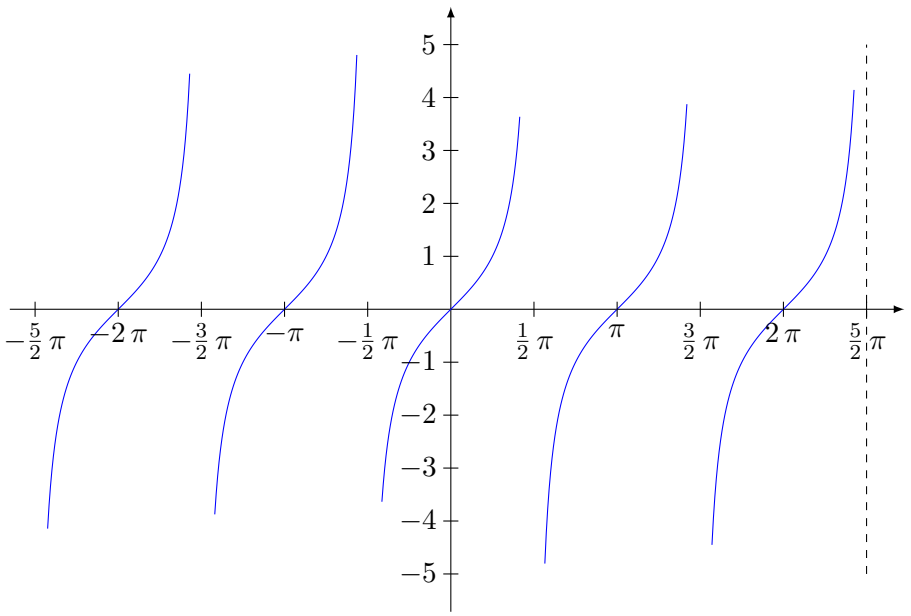


FIGURE 18.2: Le graphe de la fonction tangente. LabelFigPVJooJDyNAg

Le graphe de la fonction tangente est sur la figure 18.2.

En ce qui concerne la bijection réciproque nous avons le théorème suivant.

THOooUSVGooOAnCvC

Théorème 18.38.
La fonction inverse de la tangente,

$$\begin{aligned}\arctan\colon \mathbb{R} &\to \left]-\frac{\pi}{2},\frac{\pi}{2}\right[\\ x &\mapsto \arctan(x)\end{aligned} \tag{18.123}$$

nommée ***arc tangente*** *est*

(1) impaire et strictement croissante sur $\mathbb{R}$.

ITEMooMNHLooOVhIIb

(2) dérivable sur $\mathbb{R}$ *de dérivée*

$$\arctan'(x) = \frac{1}{1+x^2}. \tag{18.124}$$

EQooGCHGooPlwYWt

Démonstration. Il est immédiatement visible sur son développement de définition (18.2) que la fonction sinus est impaire. Une vérification similaire montre que la fonction cosinus est paire. La fonction tangente est alors impaire et sa réciproque l'est tout autant.

La fonction arc tangente est également dérivable (donc continue) par la proposition 12.178 parce que la fonction tangente l'est. Notons qu'ici nous nous sommes restreint à $]-\pi/2,\pi/2[$. Sinon, le résultat est faux.

La formule proposée pour la dérivée provient également de la proposition 12.178 et de la dérivée de la tangente du lemme 18.36. □

LEMooHRDCooGtnyeQ

Lemme 18.39.
Nous avons les limites

(1) $\lim_{x\to\infty}\arctan(x) = \frac{\pi}{2}$,

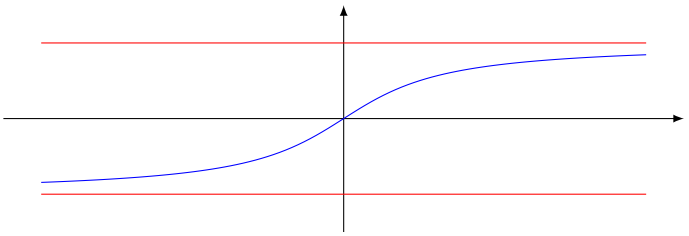
(2) $\lim_{x\to-\infty}\arctan(x) = -\frac{\pi}{2}$.

LEMooHJEKooARZMil

Lemme 18.40.
L'application $\arctan\colon \mathbb{R} \to]-\pi/2,\pi/2[$ *est un homéomorphisme*[12].

11. Théorème 9.93.
12. Bijection continue, inversion continu définition 7.41.

Exemple 18.41.
Un homéomorphisme $\phi\colon \mathbb{R} \to]-1,1[$ est donné par exemple par la fonction $\phi(t) = \frac{2}{\pi}\arctan(t)$ dont le graphique est donné ci-dessous :



△

LEMooJKIUooEMMOrs

Lemme 18.42.
Nous avons la valeur remarquable

$$\arctan(1/\sqrt{3}) = \frac{\pi}{6}. \tag{18.125}$$

Le nombre $\arctan(x_0)$ se calcule en cherchant l'angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ dont la tangente vaut x_0. Nous obtenons le tableau de valeurs suivant :

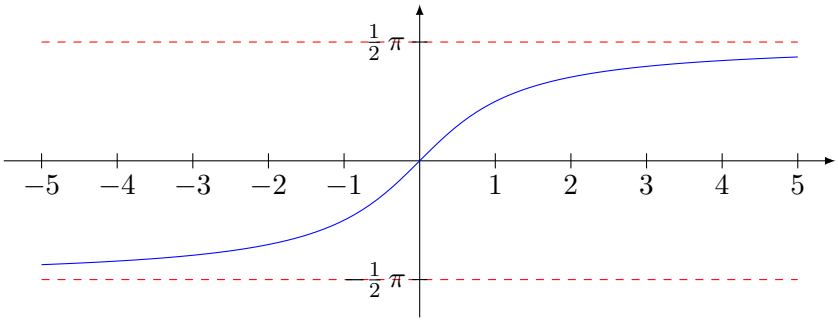
LEMooPQNCooDkEUyw

Lemme 18.43.
Quelques valeurs remarquables de l'arc tangente :

x	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$
$\arctan(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

(18.126)

En ce qui concerne la représentation graphique de la fonction $x \mapsto \arctan(x)$, elle s'obtient « en retournant » la partie entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ du graphique de la fonction tangente :



18.3.2 La fonction arc sinus

Nous voulons étudier la fonction

$$\begin{aligned} \sin\colon \mathbb{R} &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{18.127}$$

et sa réciproque éventuelle.

La fonction sinus est continue sur $\mathbb{R}$ mais n'est pas bijective : elle prend une infinité de fois chaque valeur de $J = [-1,1]$. Pour définir une bijection réciproque de la fonction sinus en utilisant le théorème 12.51, nous devons donc choisir un intervalle à partir duquel la fonction sinus est monotone. Nous choisissons l'intervalle

$$I = [-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{18.128}$$

La fonction

$$\begin{aligned} \sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{18.129}$$

est une bijection croissante et continue. Nous avons donc le résultat suivant.

Théorème 18.44 (Définition et propriétés de arc sinus)**.**
Nous nommons ***arc sinus*** *la bijection inverse de la fonction* $\sin\colon I\to J$. *La fonction*

$$\begin{aligned}\arcsin\colon [-1,1]&\to[-\frac{\pi}{2},\frac{\pi}{2}]\\ x&\mapsto\arcsin(x)\end{aligned}\tag{18.130}$$

ainsi définie est

(1) continue et strictement croissante ;

(2) impaire : pour tout $x\in[-1,1]$ *nous avons* $\arcsin(-x)=-\arcsin(x)$.

Démonstration. Nous prouvons le fait que arcsin est impaire. Un élément de l'ensemble de définition de arcsin est de la forme $y=\sin(x)$ avec $x\in[-\pi/2,\pi/2]$. La relation (12.99) s'écrit dans notre cas

$$x=\arcsin\big(\sin(x)\big).\tag{18.131}$$

Nous écrivons d'une part cette équation avec $-x$ au lieu de x :

$$-x=\arcsin\big(\sin(-x)\big)=\arcsin\big(-\sin(x)\big)=\arcsin(-y);\tag{18.132}$$

et d'autre part nous multiplions (18.131) par -1 :

$$-x=-\arcsin\big(\sin(x)\big)=-\arcsin(y).\tag{18.133}$$

En égalisant les valeurs (18.132) et (18.133) nous trouvons

$$\arcsin(-y)=-\arcsin(y),\tag{18.134}$$

ce qui signifie que arcsin est une fonction impaire. □

Notons que cette preuve repose sur le fait que tout élément de l'ensemble de définition de la fonction arc sinus peut être écrit sous la forme $\sin(x)$ pour un certain x.

Si $x_0\in[-1,1]$ est donné, calculer $\arcsin(x_0)$ revient à trouver un angle θ_0 dans $[-\frac{\pi}{2},\frac{\pi}{2}]$ pour lequel $\sin(\theta_0)=x_0$. Un tel angle sera forcément unique.

Remarque 18.45.
La définition de arc sinus découle du choix de l'intervalle I, qui est une convention. Il aurait été possible de faire un choix différent : pourriez-vous trouver la réciproque de la fonction sinus sur l'intervalle $[\pi/2,3\pi/2]$? Le mieux est de l'écrire comme une translatée de arc sinus, en utilisant le fait que sinus est une fonction périodique.

Lemme 18.46.
Nous avons $\arcsin(1)=\frac{\pi}{2}$.

Démonstration. Pour calculer $\arcsin(1)$, il faut chercher un angle entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ ayant 1 pour sinus : résoudre $\sin(\theta)=1$. La solution est $\theta=\frac{\pi}{2}$ et nous avons donc $\arcsin(1)=\frac{\pi}{2}$. □

À l'aide des valeurs remarquables de la fonction sinus nous obtenons le tableau suivant de valeurs remarquables pour l'arc sinus.

x	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arcsin(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$

Les autres valeurs remarquables peuvent être déduites du fait que l'arc sinus est une fonction impaire.

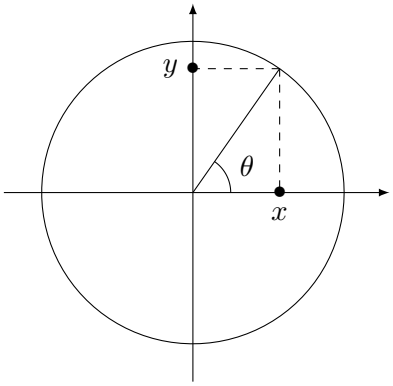
En ce qui concerne la dérivabilité de la fonction arc sinus, en application de la proposition 12.178 elle est dérivable en tout $y=\sin(x)$ tel que $\sin'(x)\neq0$, c'est-à-dire tel que $\cos(x)\neq0$. Or $\cos(x)=0$ pour $x=\pm\frac{\pi}{2}$, ce qui correspond à $y=\sin(\pm\frac{\pi}{2})=\pm1$. La fonction arc sinus est donc dérivable sur $]-1,1[$. Nous avons donc la propriété suivante pour la dérivabilité.

Proposition 18.47.
La fonction arc sinus est continue sur $[-1,1]$ *et dérivable sur* $]-1,1[$. *Pour tout* $y \in]-1,1[$, *la dérivée est donnée par la formule* (12.462), *qui dans ce cas s'écrit*

$$\arcsin'(y) = \frac{1}{\cos\big(\arcsin(y)\big)} = \frac{1}{\sqrt{1-y^2}}. \tag{18.135}$$

La dernière égalité viens du fait que si $x = \arcsin(y)$ alors $y = \sin(x)$ et $\cos(x) = \sqrt{1-\sin^2(x)} = \sqrt{1-y^2}$.

Pour comprendre la dernière égalité, remarquer que dans le dessin suivant, $\theta = \arcsin(y)$, donc $y = \sin(\theta)$, et $x = \cos(\theta)$.



Notons enfin que le graphe de la fonction arc sinus est donné à la figure 18.3.

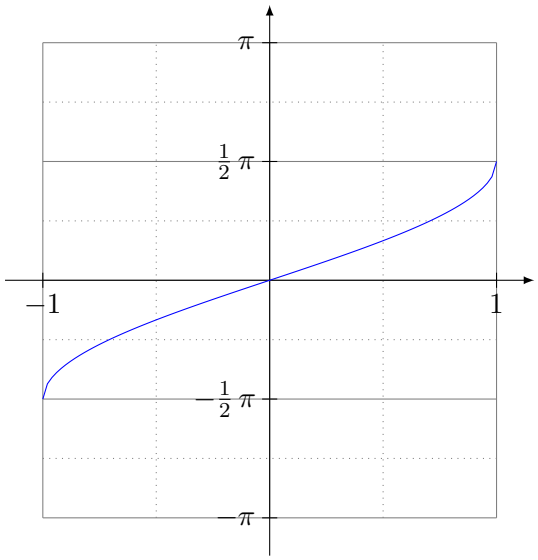


FIGURE 18.3: Le graphe de la fonction $x \mapsto \arcsin(x)$ LabelFigFGRooDhFkch

18.3.3 La fonction arc cosinus

Nous voulons étudier la fonction

$$\cos\colon \mathbb{R} \to [-1,1] \tag{18.136}$$

et son éventuelle réciproque. Encore une fois il n'est pas possible d'en prendre la réciproque globale parce que ce n'est pas une bijection ; ne fut-ce que parce qu'elle est périodique (proposition 18.17). Nous choisissons de considérer l'intervalle $[0,\pi]$ sur lequel la fonction cosinus est continue et strictement monotone décroissante.

Nous avons alors le résultat suivant :

PROPooZOZHooSMoYQD

Proposition-Définition 18.48.
Pour définir la fonction arc cosinus.

ITEMooMJWZooNHgQox

(1) La fonction

$$\cos\colon [0,\pi] \to [-1,1] \tag{18.137}$$

est une bijection continue strictement décroissante.

(2) Sa bijection réciproque est la fonction

$$\arccos\colon [-1,1] \to [0,\pi] \tag{18.138}$$

nommée ***arc cosinus****.*

(3) La fonction arc cosinus est continue, strictement décroissante.

(4) Elle est dérivable et pour tout $y \in]-1,1[$, sa dérivée est donnée par

$$\arccos'(y) = \frac{1}{-\sin\big(\arccos(y)\big)} = \frac{-1}{\sqrt{1-y^2}}. \tag{18.139}$$

Démonstration. La fonction cosinus est continue et même de classe C^∞ par la proposition 18.1. Elle est strictement décroissante parce que sa dérivée $(-\sin)$ y est strictement négative (strictement, à l'intérieur du domaine).

Le fait que arc cosinus soit une bijection continue strictement monotone est dans le théorème de la bijection 12.51. La dérivabilité et la formule sont de la proposition 12.178. □

Pour $y_0 \in [-1,1]$, trouver la valeur de $\arccos(y_0)$ revient à résoudre l'équation $\cos(x_0) = y_0$. Cela nous permet de construire une tableau de valeurs :

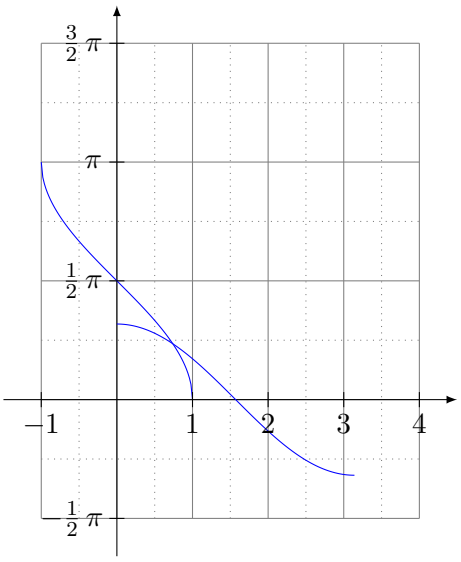
x	-1	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arccos(x)$	π	$\frac{5\pi}{6}$	$\frac{3}{4}\pi$	$\frac{2}{3}\pi$	$\frac{1}{2}\pi$	$\frac{\pi}{3}$	$\frac{1}{4}\pi$	$\frac{1}{6}\pi$	0

Remarque 18.49.
Certes la fonction cosinus est paire (vue sur $\mathbb{R}$), mais la fonction arc cosinus ne l'est pas car elle est une bijection entre $[-1,1]$ et $[0,\pi]$.

Exemple 18.50.
Cherchons $\arccos(\frac{1}{2})$. Il faut trouver un angle $\theta \in [0,\pi]$ tel que $\cos(\theta) = \frac{1}{2}$. La solution est $\theta = \frac{\pi}{3}$. Donc $\arccos(\frac{1}{2}) = \frac{\pi}{3}$.

Il n'est cependant pas immédiat d'en déduire la valeur de $\arccos(-\frac{1}{2})$. En effet $\theta = \arccos(-\frac{1}{2})$ si et seulement si $\cos(\theta) = -\frac{1}{2}$ avec $\theta \in [0,\pi]$. La solution est $\theta = \frac{2\pi}{3}$. △

En ce qui concerne la représentation graphique, il suffit de tracer la fonction cosinus entre 0 et π puis de prendre le symétrique par rapport à la droite $y = x$.



18.3.4 Une meilleure approximation de π

Nous avions laissé le nombre π avec l'approximation assez minable de $2\sqrt{2} < \pi < 4$ en le lemme 18.33. Nous pouvons maintenant faire nettement mieux.

Le lemme 18.42 donne

$$\arctan(1/\sqrt{3}) = \pi/6 \tag{18.140}$$

et l'idée est de donner un développement de arctan autour de zéro, de l'évaluer en $1/\sqrt{3}$ et d'égaliser le résultat à $\pi/6$. Tout cela donne lieu à des calculs peut-être fastidieux, mais comme un gars l'a fait dès l'an 1424[542] pour trouver 16 décimales correctes, nous faisons comme si c'était facile.

Pour trouver le développement en série de Taylor (théorème 12.461) de arc tangente autour de $x = 0$, il faut partir de la formule (18.124) et sans doute pas mal calculer et faire une récurrence [13]. Le résultat est :

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}, \tag{18.141}$$

valable pour $x \in]-1, 1[$. Avec cela nous avons

$$\arctan(\frac{1}{\sqrt{3}}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k} \times \frac{1}{\sqrt{3}} = \frac{\pi}{6}, \tag{18.142}$$

et donc

$$\pi = \frac{6}{\sqrt{3}} \sum_{k=0}^{\infty} \frac{(-1)k}{(2k+1)3^k}. \tag{18.143}$$

Pour donner une idée du fait que ça fonctionne pas mal, voici le calcul pour quelques termes :

```
1
2  SageMath version 7.3, Release Date: 2016-08-04
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6  sage: n(taylor(arctan(x),x,0,5)(1/sqrt(3)))*6
7  3.15618147156995
8  sage: n(taylor(arctan(x),x,0,10)(1/sqrt(3)))*6
9  3.14260474566308
10 sage: n(taylor(arctan(x),x,0,20)(1/sqrt(3))*6-pi)
11 -2.14265171338823e-6
12 sage: n(taylor(arctan(x),x,0,58)(1/sqrt(3))*6-pi)
13 8.88178419700125e-16
```

tex/sage/sageSnip012.sage

Calculer 5 termes donne déjà 3.15. Et on est à 10^{-6} de la bonne réponse avec 20 termes. Et avec 58 termes, on n'est à 10^{-16}.

¡¡ Avertissement/question à la lectrice !! 18.51
Pour bien faire, il faudrait étudier le reste et donner un encadrement.

18.3.5 Angle entre deux vecteurs

DEFooSVDZooPWHwFQ

Proposition-Définition 18.52.
Soient des vecteurs $X, Y \in \mathbb{R}^2$. Il existe un unique $\theta \in [0, \pi]$ tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{18.144}$$

eqDefAngleVect

Ce réel est appelé ***angle*** *entre X et Y.*

13. Je n'ai pas fait le calcul, merci de me faire savoir si il y a une astuce.

Démonstration. Si a et b sont des réels, l'inégalité $|a| \leqslant b$ peut se développer en une double inégalité

$$-b \leqslant a \leqslant b. \tag{18.145}$$

L'inégalité de Cauchy-Schwarz (10.2) devient alors

$$-\|X\|\|Y\| \leqslant X \cdot Y \leqslant \|X\|\|Y\|. \tag{18.146}$$

Si $X \neq 0$ et $Y \neq 0$, nous en déduisons

$$-1 \leqslant \frac{X \cdot Y}{\|X\|\|Y\|} \leqslant 1. \tag{18.147}$$

Il existe donc par la proposition 18.48 un angle $\theta \in [0, \pi]$ tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{18.148}$$

□

18.53.
Certains n'hésitent pas à écrire la formule

$$X \cdot Y = \|X\|\|Y\| \cos(\theta). \tag{18.149}$$ eqPropCosThet

comme une définition du produit scalaire. C'est ce qui arrive lorsqu'on définit les fonctions trigonométriques à partir de relations dans les triangles rectangles.

Nous pouvons prouver simplement que $\sin(30^\circ) = \frac{1}{2}$ et $\cos(30^\circ) = \frac{\sqrt{3}}{2}$ en s'inspirant de la figure 18.4.

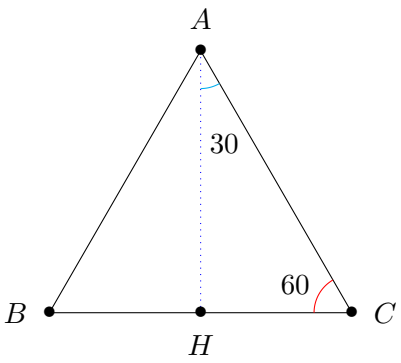


FIGURE 18.4: Un triangle équilatéral de côté 1. LabelFigGVDJooYzMxLW

18.3.6 Aire du parallélogramme

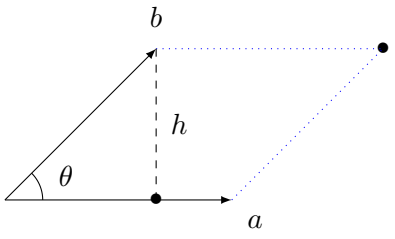


FIGURE 18.5: Calculer l'aire d'un parallélogramme. LabelFigBNHLooLDxdPA

RemaAireParalProdVect

Remarque 18.54.
Le nombre $\|a\|\|b\|\sin(\theta)$ est l'aire du parallélogramme [14] formé par les vecteurs a et b, comme cela se voit sur la figure 18.5. Un vrai calcul avec une intégrale sera effectué dans la proposition 20.26.

PropNormeProdVectoabsint

Proposition 18.55.
Nous avons

$$\|a \times b\| = \|a\|\|b\|\sin(\theta) \tag{18.150}$$

où $\theta \in [0, \pi]$ est l'angle formé par a et b.

Démonstration. En utilisant la décomposition du produit vectoriel [15], nous avons

$$\begin{aligned}
\|a \times b\|^2 &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 \\
&= (a_2b_3 - b_2a_3)^2 + (a_1b_3 - a_3b_1)^2 + (a_1b_2 - a_2b_1)^2 \\
&= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\
&= \|a\|^2\|b\|^2 - (a \cdot b)^2 \\
&= \|a\|^2\|b\|^2 - \|a\|^2\|b\|^2\cos^2(\theta) \\
&= \|a\|^2\|b\|^2\big(1 - \cos^2(\theta)\big) \\
&= \|a\|^2\|b\|^2\sin^2(\theta).
\end{aligned} \tag{18.151}$$

D'où le résultat. Nous avons utilisé la formule de la définition (18.52) donnant l'angle en fonction du produit scalaire. □

NORMooWWOKooWzScnZ

18.56.
Si les vecteurs a, b et c ne sont pas coplanaires, alors la valeur absolue du produit mixte (voir équation (10.99)) $a \cdot (b \times c)$ donne le volume du parallélépipède construit sur les vecteurs a, b et c.

En effet si φ est l'angle entre $b \times c$ et a, alors la hauteur du parallélépipède vaut $\|a\|\cos(\varphi)$ parce que la direction verticale est donnée par $b \times c$, et la hauteur est alors la « composante verticale » de a. Par conséquent, étant donné que $\|b \times c\|$ est l'aire de la base, le volume du parallélépipède vaut [16]

$$V = \|b \times c\|\|a\|\cos(\varphi). \tag{18.152}$$

Or cette formule est le produit scalaire de a par $b \times c$; ce dernier étant donné par le déterminant de la matrice formée des composantes de a, b et c grâce à la formule (10.99).

La valeur absolue du déterminant

EqDeratb

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \tag{18.153}$$

est l'aire du parallélogramme déterminé par les vecteurs $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ et $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. En effet, d'après la remarque 18.54, l'aire de ce parallélogramme est donnée par la norme du produit vectoriel

$$\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} = \begin{vmatrix} e_x & e_y & e_z \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} e_z, \tag{18.154}$$

donc la norme $\|a \times b\|$ est bien donnée par la valeur absolue du déterminant (18.153).

14. Définition de ce qu'est une aire : 20.21. Preuve dans le cas d'un parallélogramme : 20.26.
15. Directement de la définition 10.28.
16. Le calcul de ce volume mériterait une certaine réflexion, surtout à partir du moment où nous avons décidé de définir les fonctions trigonométriques à partir de leur développement (définition 18.1).

18.4 Paramétrisation du cercle

Nous allons parler de paramértisation du cercle. L'ensemble S^1 sera vu tantôt comme le cercle dans $\mathbb{R}^2$, tantôt comme le cercle dans $\mathbb{C}$. Nous n'allons pas pousser le vice jusqu'à écrire explicitement les isomorphismes lorsque nous passons d'une représentation à l'autre. Parmi les identifications que nous allons faire sans ménagement, il y a l'identification entre les applications

$$\begin{aligned}\gamma\colon [0,2\pi[&\to \mathbb{R}^2\\ t &\mapsto \big(\cos(t),\sin(t)\big)\end{aligned} \tag{18.155}$$

et

$$\begin{aligned}\varphi\colon [0,2\pi[&\to \mathbb{C}\\ t &\mapsto e^{it}.\end{aligned} \tag{18.156}$$

C'est évidemment la formule $e^{it}=\cos(t)+i\sin(t)$ (lemme 18.11) qui permet de transformer γ en φ et inversement. De plus $\mathbb{R}^2$ et $\mathbb{C}$ sont isomorphes en tant qu'espaces vectoriels normés (et aussi donc topologiques).

Nous allons prouver suffisamment de résultats à propos de ces deux applications pour pouvoir dire qu'elles paramétrisent le cercle et écrire des égalités du type

$$\int_{S^1} f=\int_0^{2\pi} f\circ\varphi. \tag{18.157}$$

18.5 Cercle trigonométriques

PROPooWZFGooMVLtFz

Proposition 18.57 ([543]).
Soient des fonctions $f,g\colon I\to\mathbb{R}$ de classe C^1 sur l'ouvert I de $\mathbb{R}$ telles que $f^2+g^2=1$. Soient $t_0\in I$ et θ_0 tel que $f(t_0)=\cos(\theta_0)$ et $g(t_0)=\sin(\theta_0)$.

Alors il existe une unique fonction continue $\theta\colon I\to\mathbb{R}$ telle que

$$\begin{cases}\theta(t_0)=\theta_0 & \text{(18.158a)}\\ f=\cos\circ\theta & \text{(18.158b)}\\ g=\sin\circ\theta. & \text{(18.158c)}\end{cases}$$

Démonstration. Nous commençons par l'existence, en passant par les nombres complexes. Soit $h\colon I\to\mathbb{C}$ définie par $h=f+ig$. Nous avons $h\bar{h}=1$ et nous définissons

$$\theta(t)=\theta_0-i\int_{t_0}^{t} h'(s)\overline{h(s)}ds. \tag{18.159}$$

Cette intégrale existe pour tout t parce que les fonctions f et g étant continues, elles sont bornées sur le compact $[t_0,t]$. De plus θ est une fonction continue parce que c'est une primitive (proposition 14.286) [17].

La dérivée de θ est la fonction $s\mapsto -ih'(s)\overline{h(s)}$.

Utilisant la formule du lemme 18.11 sur la forme trigonométrique des nombres complexes, nous calculons :

$$\frac{d}{dt}\Big[he^{-i\theta}\Big]_{t=0}=e^{-i\theta}(h'-h\theta')=e^{-i\theta}(h'-ih(-i)h'\bar{h})=0. \tag{18.160}$$

Par conséquent il existe $c\in\mathbb{C}$ tel que $he^{-i\theta}=c$. Mais $h(t_0)=f(t_0)+ig(t_0)=\cos(\theta_0)+i\sin(\theta_0)=e^{i\theta_0}$, du coup

$$h(t_0)e^{-i\theta(t_0)}=c \tag{18.161}$$

donne immédiatement $c=1$, ou encore $e^{i\theta(t)}=h(t)$, c'est-à-dire que

$$f+ig=\cos\circ\,\theta+i\sin\circ\,\theta, \tag{18.162}$$

17. En réalité nous appliquons le théorème 10.1 à chacune des parties réelles et imaginaires de la fonction $s\mapsto h'(s)\overline{h(s)}$.

ce qu'il fallait pour l'existence.

Pour l'unicité nous supposons avoir une autre fonction, α qui satisfait aux exigences. Pour tout $t \in I$ nous avons

$$e^{i\theta(t)} = e^{i\alpha(t)}. \tag{18.163}$$

Il existe donc une fonction $n: I \to \mathbb{N}$ telle que $\theta(t) = \alpha(t) + 2n(t)\pi$. Par continuité de θ et α, la fonction n doit être constante, mais vu que $\theta(t_0) = \alpha(t_0)$ nous avons $n = 1$. □

18.5.1 Bijection continue

PROPooKSGXooOqGyZj

Proposition 18.58.
L'application

$$\begin{aligned}\gamma: [0, 2\pi[&\to S^1 \subset \mathbb{R}^2 \\ t &\mapsto \big(\cos(t), \sin(t)\big)\end{aligned} \tag{18.164}$$

est une bijection continue.

Démonstration. En plusieurs points.

(1) **Image** Le fait que l'image de γ soit dans S^1 découle immédiatement du fait que $\sin^2 + \cos^2 = 1$.

(2) **Continuité** La continuité découle de la continuité des composantes.

(3) **Injectif** Soient $x_1 < x_2$ tels que $\sin(x_1) = \sin(x_2)$ et $\cos(x_1) = \cos(x_2)$. Supposons pour fixer les idées que $\sin(x_1) > 0$ et $\cos(x_1) > 0$: si ce n'est pas le cas, il faut traiter séparément les 4 possibilités de combinaisons de signes.

Nous avons obligatoirement $x_1, x_2 \in [0, \frac{\pi}{2}[$. Vu que nous avons supposé que $\sin(x_1) = \sin(x_2)$, le théorème de Rolle 12.189 nous permet de considérer un élément $c \in]x_1, x_2[$ tel que $\sin'(c) = 0$, c'est-à-dire $\cos(c) = 0$. Cela contredirait la proposition 18.20(10) à moins que $x_1 = x_2$.

(4) **Surjectif**

Soient x, y tels que $x^2 + y^2 = 1$. Supposons pour varier les plaisirs[18] que $x < 0$ et $y > 0$. Puisque la fonction cos va de 0 à -1 lorsque x va de $\pi/2$ à π, le théorème des valeurs intermédiaires donne $t \in [\pi/2, \pi]$ tel que $\cos(t) = x$.

En reportant cette valeur de x dans l'égalité $x^2 + y^2 = 1$, nous trouvons $y^2 = 1 - \cos^2(t) = \sin^2(t)$, et donc $\sin(t) = \pm y$. Mais pour $t \in [\pi/2, \pi]$ nous avons $\sin(t) > 0$. Par conséquent $\sin(t) = y$ parce que nous avons également supposé $y > 0$.

□

CORooAKMKooORqcrO

Corolaire 18.59.
Soient $R > 0$ ainsi que le cercle $S = \{x \in \mathbb{R}^2$ tel que $\|x\| = R\}$ de rayon R. L'application

$$\begin{aligned}\gamma: [0, 2\pi[&\to S \\ t &\mapsto \big(R\cos(t), R\sin(t)\big)\end{aligned} \tag{18.165}$$

est une bijection continue.

EXooJFDPooBZADKs

Exemple 18.60.
L'application

$$\begin{aligned}\varphi:]0, 2\pi[&\to S^1 \\ x &\mapsto \begin{pmatrix}\cos(x) \\ \sin(x)\end{pmatrix}\end{aligned} \tag{18.166}$$

est continue par la proposition 18.58. Comme $]0, 2\pi[$ est connexe (proposition 9.53) le lemme 7.218 implique que le cercle privé d'un point est connexe. △

18. Mais les autres cas sont à faire, pour être complet.

La proposition suivante ne va pas jusqu'à dire qu'il y a un C^∞-difféomorphisme entre $[0,2\pi[$ et S^1. La raison est que n'avons pas défini la notion de différentielle pour des applications définies sur des fermés.

Si vous voulez un C^∞ difféomorphisme, il y a deux possibilités. Soit vous acceptez d'élargir un peu le domaine et vous allez voir les coordonnées polaires du théorème 18.227, soit vous allez étudier la géométrie différentielle, et vous considérez S^1 comme une variété.

PROPooXELTooYKjDav

Proposition 18.61 ([1, 162])**.**
Dans cette proposition, S^1 désigne l'ensemble des nombres complexes de norme 1. Trois affirmations.

ITEMooOHRHooRXvxrL

(1) L'application

$$\begin{aligned} f\colon [0,2\pi[&\to S^1 \\ x &\mapsto e^{ix} \end{aligned} \tag{18.167}$$

est une bijection continue.

ITEMooPWUWooHsQBDg

(2) La réciproque de

$$\begin{aligned} f\colon [0,2\pi[&\to S^1 \\ x &\mapsto e^{ix} \end{aligned} \tag{18.168}$$

n'est pas continue en 1.

ITEMooCVAQooLsYspI

(3) L'application

$$\begin{aligned} f\colon]0,2\pi[&\to \mathbb{C} \\ x &\mapsto e^{ix} \end{aligned} \tag{18.169}$$

est de classe C^∞.

Démonstration. En plein de points.

(1) **Pour (1)**
Nous savons que

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{C} \\ (x,y) &\mapsto x+iy \end{aligned} \tag{18.170}$$

est une bijection isométrique. C'est pour cela que nous allons nous permettre de noter S^1 le cercle unité dans $\mathbb{R}^2$ aussi bien que l'ensemble des nombres complexes de norme 1.
Sur $\mathbb{R}^2$ nous avons l'application

$$\begin{aligned} \gamma\colon [0,2\pi[&\to S^1 \subset \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \end{aligned} \tag{18.171}$$

qui est une bijection continue (c'est la proposition 18.58). Et enfin le lemme 18.11 nous donne $e^{ix} = \cos(x) + i\sin(x)$.
Avec tout ça, l'application $\varphi^{-1}\circ f\colon [0,2\pi[\to S^1$ est une bijection continue. Et comme φ l'est également, f est une bijection continue.

(2) **Pour (2)** Nous vérifions que la définition 7.36(1) n'est pas satisfaite pour f^{-1}. Nous considérons le voisinage $W = [0,\epsilon[$ de 0 dans $[0,2\pi[$ avec $\epsilon < 1/2$. Nous allons montrer que tout voisinage V de 1 dans S^1 contient des points z_k tels que $f^{-1}(z_k) \notin W$. Tout voisinage de 1 dans S^1 contient une partie de la forme $B(1,r)\cap S^1$. Posons $z_k = \exp\big(i(2\pi - 1/k)\big)$. Nous avons

$$\begin{aligned} \|1 - \exp\big(i(2\pi-1/k)\big)\|^2 &= \|1 - \cos(2\pi - 1/k) - i\sin(2\pi - 1/k)\| \quad \text{lem 18.11} && \text{(18.172a)} \\ &= |1-\cos(2\pi-1/k)|^2 + |\sin(2\pi-1/k)|^2 && \text{(18.172b)} \\ &\to 0. && \text{(18.172c)} \end{aligned}$$

Donc pour tout r, nous avons $z_k \in B(1,r)\cap S^1$ dès que k est assez grand. Or $f^{-1}(z_k) = 2\pi - 1/k \notin [0,\epsilon[$ dès que k est assez grand.

(3) **Pour (3)** Nous avons $f(x) = \cos(x) + i\sin(x)$. Les fonctions sin et cos étant de classe C^∞ par la proposition 18.1, la fonction f l'est aussi.

□

La proposition suivante donne les coordonnées polaires sur $\mathbb{C}$. La régularité est l'objet du théorème 18.227 (à part le fait que ce dernier parle de $\mathbb{R}^2$ et non de $\mathbb{C}$).

PROPooRFMKooURhAQJ

Proposition 18.62 (Décomposition polaire des nombres complexes)**.**
Pour tout nombre complexe $z \neq 0$, il existe un unique $\theta \in [0, 2\pi[$ tel que

$$z = |z|e^{i\theta}. \tag{18.173}$$

Démonstration. Soit $z \in \mathbb{C}$. Nous considérons $z' = z/|z|$ qui est de norme 1. Donc il existe un unique $\theta \in [0, 2\pi[$ tel que $z' = e^{i\theta}$ (proposition 18.61(1)).

Pour ce θ nous avons $z = |z|e^{i\theta}$. □

Bien entendu, le θ est unique dans $[0, 2\pi[$, mais il n'est pas du tout unique dans $\mathbb{R}$.

LEMooOQKNooGZlJHf

Lemme 18.63 ([1])**.**
Pour tout $a, b \in \mathbb{C}$, il existe $t \in [0, 2\pi[$ tel que

$$|a + b|^2 = |a|^2 + |b|^2 + 2|a||b|\cos(t). \tag{18.174}$$

Plus précisément, si les formes trigonométriques[19] *de a et b sont $a = |a|e^{i\alpha}$ et $b = |b|e^{i\beta}$, alors*

$$|a + b|^2 = |a|^2 + |b|^2 + 2|a||b|\cos(\alpha - \beta). \tag{18.175}$$

Démonstration. Nous avons le calcul suivant :

$$\begin{aligned}
|a + b|^2 &= (a + b)(\bar{a} + \bar{b}) && \text{(18.176a)}\\
&= |a|^2 + a\bar{b} + b\bar{a} + |b|^2 && \text{(18.176b)}\\
&= |a|^2 + |b|^2 + |a||b|\big(e^{i(\alpha-\beta)} + e^{i(\beta-\alpha)}\big) && \text{(18.176c)}\\
&= |a^2| + |b|^2 + 2|a||b|\cos(\alpha - \beta). && \text{(18.176d)}
\end{aligned}$$

□

CORooONAVooKPhQuI

Corolaire 18.64.
La partie $B = \{z \in \mathbb{C} \text{ tel que } |z| > R\}$ est connexe et connexe par arcs[20]*.*

Démonstration. Nous prouvons que B est connexe par arcs ; la connexité découlera du lemme 9.66. Nous considérons les points $z_1 = r_1 e^{i\theta_1}$ et $z_2 = r_2 e^{i\theta_2}$ dans B. Une telle décomposition est toujours possible par la proposition 18.62.

Supposons $\theta_1 < \theta_2$ et $r_1 < r_2$; sinon il faut faire différents cas. Nous relions les points z_1 et z_2 par le chemin continue constitué des deux chemins suivants mis bout à bout. D'abord

$$\begin{aligned}
\gamma\colon [\theta_1, \theta_2] &\to B\\
t &\mapsto r_1 e^{it},
\end{aligned} \tag{18.177}$$

et ensuite

$$\begin{aligned}
\sigma\colon [r_1, r_2] &\to B\\
u &\mapsto u e^{i\theta_2}.
\end{aligned} \tag{18.178}$$

□

19. Proposition 18.62.
20. Définition 9.64.

18.5.2 Inverse

SUBSECooWFNMooOuZBRN

Nous pouvons écrire un inverse de la fonction φ grâce à la fonction arc tangente introduite au théorème 18.38. La fonction que nous écrivons à présent est la fonction $\arg_{0^-}$ définie par (26.127). Elle n'est pas exactement la fonction argument définie par (26.88).

Nous avons :

$$\varphi^{-1}\colon S^1 \to [0,2\pi[$$
$$x+iy \mapsto \begin{cases} \arctan(y/x) & \text{si } x>0, y\geqslant 0 \\ \frac{\pi}{2} & \text{si } (x,y)=(0,1) \\ \pi-\arctan(-y/x) & \text{si } x<0, y\geqslant 0 \\ \pi+\arctan(y/x) & \text{si } x<0, y<0 \\ 2\pi-\arctan(-y/x) & \text{si } x>0, y<0 \end{cases} \tag{18.179}$$

EQooSAYFooRFVSPc

Chacune des branches est continue parce que la fonction arc tangente l'est. Trois des raccords sont également continus grâce aux limites du lemme 18.39.

L'application φ^{-1} n'est cependant pas continue au point $(1,0)$[21]. C'est l'objet du lemme suivant.

LEMooEQVRooMAffCw

Lemme 18.65.
L'application $\varphi^{-1}\colon S^1 \to [0,2\pi[$ n'est pas continue en $(1,0)$. Mais elle est continue ailleurs. Autrement dit,

$$\varphi^{-1}\colon S^1\backslash\{(1,0)\} \to]0,2\pi[\tag{18.180}$$

est continue.

Démonstration. En effet, φ^{-1} serait continue si l'image de tout ouvert de $[0,2\pi[$ par φ était ouverte dans S^1 (topologie induite de $\mathbb{C}$). Prenons un petit ouvert $[0,\epsilon[$ (si vous êtes étonnés, c'est que vous n'avez pas bien la topologie induite en tête). Son image contient le point $(1,0)$, mais aucun point (x,y) avec $y<0$.

Montrons que tout voisinage de $(1,0)$ dans $\mathbb{C}$ contient des points $x+iy$ de S^1 avec $y<0$. Un point de S^1 est de la forme $\cos(t)+i\sin(t)$. Nous avons :

$$|\cos(t)+i\sin(t)-1|^2 = \big(\cos(t)-1\big)^2+\sin^2(t) = 2\big(1-\cos(t)\big). \tag{18.181}$$

Soit $\delta>0$, et montrons que $B\big((1,0),\delta\big)\cap S^1$ contient des points d'ordonnées négatives. D'abord il existe $\epsilon>0$ tel que pour $t=2\pi-\epsilon$,

$$2\big(1-\cos(t)\big)<\delta. \tag{18.182}$$

Ensuite pour de tels t, nous avons $\sin(t)<0$. Donc les points de S^1 correspondant à $2\pi-\epsilon$ sont dans $S^1\cap B\big((1,0),\delta\big)$.

Bref, l'image de $[0,\epsilon[$ n'est pas un ouvert de S^1. □

18.5.3 Cercle trigonométrique

Définition 18.66 ([544]).
Le cercle trigonométrique est le cercle dans $\mathbb{R}^2$ de rayon 1 centré en $(0,0)$ représenté à la figure 18.6. Nous n'hésiterons pas à parler de cercle trigonométrique dans $\mathbb{C}$.

Nous verrons plus tard (proposition 21.13) que la longueur de l'arc de cercle intercepté par un angle θ est égal à θ. Les radians sont donc l'unité d'angle la plus adaptée au calcul de longueurs sur le cercle.

21. Puisque nous avons considéré $S^1\subset\mathbb{C}$, nous aurions dû noter « 1 » ce point. Mais vous vous imaginez le clash de notation avec le $1\in[0,2\pi[\subset\mathbb{R}$?

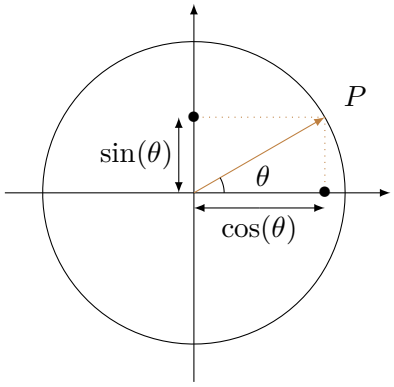


FIGURE 18.6: Le cercle trigonométrique. LabelFigCercleTrigono

LEMooXISFooRMWUEC

Lemme 18.67.
Soit $R > 0$. Soient le cercle $S = \{x \in \mathbb{R}^2 \text{ tel que } \|x\| = R\}$ de rayon R ainsi que $t_0 \in \mathbb{R}$. L'application

$$\begin{aligned}\gamma_{t_0} : [t_0, t_0 + 2\pi[&\to S \\ x &\mapsto \big(R\cos(x), R\sin(x)\big)\end{aligned} \tag{18.183}$$

est une bijection continue.

Démonstration. Le corolaire 18.59 nous dit que

$$\begin{aligned}\gamma : [0, 2\pi[&\to S \\ x &\mapsto \big(R\cos(x), R\sin(x)\big)\end{aligned} \tag{18.184}$$

est une bijection continue.

D'autre part, l'application

$$\begin{aligned}f : [t_0, t_0 + 2\pi[&\to [0, 2\pi[\\ x &\mapsto x - t_0\end{aligned} \tag{18.185}$$

est une bijection continue.

Comme $\gamma_{t_0} = \gamma \circ f$, nous déduisons que γ_{t_0} est une bijection continue. □

18.68.
La proposition 21.13 dira que la longueur d'arc de cercle de rayon R interceptée par un angle θ vaut $R\theta$.

18.5.4 Du point de vue de la tribu, mesure et co.

Nous avons considéré sur S^1 la topologie induite de $\mathbb{C}$. Nous allons y mettre la tribu induite de celle de Lebesgue de $\mathbb{C}$. Mais nous n'allons pas y mettre la *mesure* induite de $\mathbb{C}$; sinon tout serait toujours de mesure nulle.

PROPooQFYHooEajmbW

Proposition 18.69 ([1])**.**
L'application φ est borélienne d'inverse borélien, c'est-à-dire

$$\mathcal{B}or(S^1) = \varphi\big(\mathcal{B}or([0, 2\pi[)\big). \tag{18.186}$$

Démonstration. L'inclusion $\mathcal{B}\text{or}(S^1) \subset \varphi\big(\mathcal{B}\text{or}([0, 2\pi[)\big)$ est la plus simple : si $A \in \mathcal{B}\text{or}(S^1)$, alors $\varphi^{-1}(A) \in \mathcal{B}\text{or}\big([0, 2\pi[\big)$ parce que $\varphi : [0, 2\pi[\to S^1$ est continue et donc borélienne (théorème 14.56).

Pour l'autre inclusion, il faudra procéder par étapes.

(1) **Ouvert ne contenant pas zéro** Si A est un ouvert de $[0, 2\pi[$ ne contenant pas 0, il est un ouvert de $\mathbb{R}$ ou de $]0, 2\pi[$. Le lemme 18.65 nous indique que son image par φ est ouverte dans S^1. En particulier, $\varphi(A) \in \mathcal{B}\text{or}(S^1)$.

(2) **Ouvert de la forme $[0,\epsilon[$** Nous supposons que ϵ est petit. Disons pour fixer les idées, plus petit que $\pi/2$. Nous avons :

$$\varphi\big([0,\epsilon[\big) = \varphi\big(]0,\epsilon[\big) \cup \varphi\big(\{0\}\big). \tag{18.187}$$

Le premier élément de l'union est un ouvert, et le second, un unique point. L'union est un borélien.

(3) **Ouvert général** Si un ouvert de $[0,2\pi[$ ne contient pas 0, son image est ouverte. Nous nous penchons sur le cas d'un ouvert contenant 0.

Si un ouvert de $[0,2\pi[$ contient 0, alors il contient un ouvert de la forme $[0,\epsilon[$, parce qu'un ouvert contient une boule autour de chacun de ses points (théorème 7.8 couplé au fait que nous sommes dans la topologie induite de $\mathbb{R}$).

Si A est un ouvert contenant zéro, alors

$$A = [0,\epsilon[\cup \big(A\backslash[0,\frac{\epsilon}{2}]\big). \tag{18.188}$$

Nous avons déjà vu que l'image du premier élément de l'union est un borélien. Étant donné que $A\backslash[0,\frac{\epsilon}{2}]$ est un ouvert ne contenant pas zéro, son image est un ouvert. Donc l'image de A est un borélien.

(4) **Pause** Nous avons déjà vu que l'image par φ de tout ouvert de $[0,2\pi[$ était un borélien de S^1. Nous devons en déduire que l'image de tout borélien de $[0,2\pi[$ est un borélien de S^1.

C'est ce que nous faisons maintenant

(5) **Boréliens** Nous utilisons le lemme de transport 14.48 avec l'application φ^{-1} et l'ensemble des ouverts :

$$\varphi\big(\sigma(\mathcal{C})\big) = \sigma\big(\varphi(\mathcal{C})\big) \tag{18.189}$$

où $\mathcal{C}$ est la tribu des ouverts dans $[0,2\pi[$. L'ensemble $\sigma(\mathcal{C})$ est par définition l'ensemble $\mathcal{B}\text{or}\big([0,2\pi[\big)$. D'autre part nous avons vu que l'image d'un ouvert est un borélien : $\varphi(\mathcal{C}) \subset \mathcal{B}\text{or}(S^1)$. Nous avons donc

$$\varphi\big(\mathcal{B}\text{or}([0,2\pi[)\big) = \sigma\big(\varphi(\mathcal{C})\big) \subset \sigma\big(\mathcal{B}\text{or}(S^1)\big) \subset \mathcal{B}\text{or}(S^1). \tag{18.190}$$

La preuve est terminée. □

PROPooHMSCooRIjcJq

Proposition 18.70 (Boréliens sur S^1[1])**.**
Soit la structure usuelle d'espace mesurable $(\mathbb{C},\mathcal{B}or(\mathbb{C}))$. Nous considérons

- *la tribu $\mathcal{B}or(\mathbb{C})_{S^1}$ induite de la tribu des boréliens de $\mathbb{C}$ vers S^1,*
- *la tribu $\mathcal{B}or(S^1)$ des boréliens de S^1 construite à partir de la topologie induite de $\mathbb{C}$ vers S^1.*
- *la bijection $\varphi\colon [0,2\pi[\to S^1$,*
- *la mesure de Lebesgue sur $[0,2\pi[$ (induite de celle sur $\mathbb{R}$) et sur $\mathbb{C}$, que nous noterons toutes deux λ.*

Alors

ITEMooSUNEooRhAdep

(1) Nous avons les expressions

$$\mathcal{B}or(\mathbb{C})_{S^1} = \{A \in \mathcal{B}or(\mathbb{C}) \text{ tel que } A \subset S^1\} \tag{18.191a}$$
$$= \{A \cap S^1 \text{ tel que } A \in \mathcal{B}or(\mathbb{C})\} \tag{18.191b}$$

SUBEQooYZGCooDqXmft

ITEMooGYPNooRaZbNW

(2) Nous avons

$$\mathcal{B}or(S^1) = \mathcal{B}or(\mathbb{C})_{S^1} = \varphi\Big(\mathcal{B}or\big([0,2\pi[\big)\Big). \tag{18.192}$$

ITEMooFUXKooFQdoaw

(3) *En définissant* $\mu\colon \mathcal{B}or(S^1) \to \mathbb{R}$ *par*

$$\mu(A) = \frac{\lambda\big(\varphi^{-1}(A)\big)}{2\pi}, \tag{18.193}$$

EQooKHZBooSrFMdo

le triplet $\big(S^1, \mathcal{B}or(S^1), \mu\big)$ *est un espace mesuré.*

ITEMooBQLRooOsqesg

(4) *L'espace mesuré* $\big(S^1, \mathcal{B}or(S^1), \mu\big)$ *est fini et*

$$\mu(S^1) = 1. \tag{18.194}$$

Démonstration. Point par point.

(1) **Pour (1)** C'est la proposition 14.9.

(2) **Pour (2)** La première égalité est le lemme 14.54. Le fait que $\mathcal{B}\text{or}(S^1) = \varphi\Big(\mathcal{B}\text{or}\big([0,2\pi[\big)\Big)$ est déjà la proposition 18.69.

(3) **Pour (3)** Nous devons d'abord nous assurer que la formule ait un sens. Cela est chose aisée ; si $A \in \mathcal{B}\text{or}(S^1)$, le point (2) nous indique que $\varphi^{-1}(A) \in \mathcal{B}\text{or}\big([0,2\pi[\big)$. Ensuite, nous devons vérifier les deux conditions de la définition 14.18 pour avoir un espace mesuré.
En premier lieu,

$$\mu(\varnothing) = \frac{1}{2\pi}\lambda\big(\varphi^{-1}(\varnothing)\big) = \frac{1}{2\pi}\lambda(\varnothing) = 0. \tag{18.195}$$

Et en second lieu, si les $A_i \in \mathcal{B}\text{or}(S^1)$ sont disjoints, les $\varphi^{-1}(A_i)$ sont également disjoints parce que φ^{-1} est une bijection. Donc

$$\begin{aligned}
\mu\Big(\bigcup_i A_i\Big) &= \frac{1}{2\pi}\lambda\Big(\bigcup_i \varphi^{-1}(A_i)\Big) && \text{(18.196a)}\\
&= \frac{1}{2\pi}\sum_i \lambda\big(\varphi^{-1}(A_i)\big) && \text{(18.196b)}\\
&= \sum_i \frac{\lambda\big(\varphi^{-1}(A_i)\big)}{2\pi} && \text{(18.196c)}\\
&= \sum_i \mu(A_i). && \text{(18.196d)}
\end{aligned}$$

D'accord.

(4) **Pour (4)** En ce qui concerne la mesure de S^1 pour μ nous avons simplement

$$\mu(S^1) = \frac{\lambda\big([0,2\pi[\big)}{2\pi} = 1. \tag{18.197}$$

□

Maintenant que $(S^1, \mathcal{B}\text{or}(S^1), \mu)$ est un espace mesuré, nous pouvons compléter la tribu $\mathcal{B}\text{or}(S^1)$ pour la mesure μ.

Définition 18.71.
La ***tribu de Lebesgue*** *sur* S^1 *est la mesure complétée pour*

$$\big(S^1, \mathcal{B}or(S^1), \mu\big) \tag{18.198}$$

où μ *est la mesure définie par la proposition 18.70. Nous notons* $\mathcal{L}eb(S^1)$ *la tribu et encore* μ *la mesure.*

PROPooDLBCooUfQZOa

Proposition 18.72 (Lebesgue sur S^1[1])**.**
Soit la structure d'espace mesuré complet $\big(S^1, \mathcal{L}eb(S^1), \mu\big)$*. Nous considérons*

— *la tribu* $\mathcal{L}eb(\mathbb{C})_{S^1}$ *induite de la tribu des boréliens de* $\mathbb{C}$ *vers* S^1,
— *la bijection* $\varphi\colon [0,2\pi[\to S^1$,

Alors

ITEMooQMHDooHEThPf

(1) *La tribu* $\mathcal{L}eb(\mathbb{C})_{S^1}$ *est la tribu de toutes les parties de* S^1.

ITEMooNIRNooKSeyCa

(2) *La tribu* $\mathcal{L}eb(S^1)$ *est donnée par*

$$\mathcal{L}eb(S^1) = \varphi\big(\mathcal{L}eb(\mathbb{R})_{[0,2\pi[}\big) = \varphi\big(\mathcal{L}eb([0,2\pi[)\big) \tag{18.199}$$

où $\mathcal{L}eb([0,2\pi[)$ *est la tribu sur* $[0,2\pi[$ *obtenue par complétion de la tribu des boréliens de la topologie induite.*

ITEMooXDBTooYnauyi

(3) *Nous avons l'inclusion stricte*

$$\mathcal{L}eb(S^1) \subsetneq \mathcal{L}eb(\mathbb{C})_{S^1}. \tag{18.200}$$

Démonstration. Point par point.

(1) **Pour (1)** Si $A \subset S^1$, alors A est une partie de S^1 qui est mesurable et de mesure nulle pour $\mathbb{C}$. Donc A est λ-négligeable et par conséquent mesurable.

(2) **Pour (2)** Il s'agit de prouver que

$$\widehat{\mathcal{B}\text{or}(S^1)} = \varphi\big(\widehat{\mathcal{B}\text{or}([0,2\pi[)}\big). \tag{18.201}$$

Ce n'est rien d'autre que la proposition 14.83. La seconde partie de l'égalité est la proposition 14.78

(3) **Pour (3)** Comme indiqué au point (1), la tribu $\mathcal{L}\text{eb}(\mathbb{C})_{S^1}$ est la tribu de toutes les parties de S^1 ; l'inclusion est donc évidente. Le point pas tout à fait évident à prouver est l'existence de parties de S^1 à n'être pas dans $\mathcal{L}\text{eb}(S^1)$.
Soit V non mesurable dans $[0,2\pi[$ (prenez quelque chose comme l'ensemble de Vitali de l'exemple 14.157). Puisque, par le point (2),

$$\mathcal{L}\text{eb}(S^1) = \varphi\big(\mathcal{L}\text{eb}(\mathbb{R})_{[0,2\pi[}\big), \tag{18.202}$$

la partie $\varphi^{-1}(V)$ ne peut pas être dans $\mathcal{L}\text{eb}(S^1)$.

□

Si vous en voulez plus à propos de S^1 et la façon dont on passe la structure depuis $[0,2\pi[$, vous pouvez lire la proposition 27.94 qui donne la structure de

$$L^2\big(S^1, \mathcal{L}\text{eb}(S^1), \mu\big) \tag{18.203}$$

qui sera, sans surprise, la même que celle de

$$L^2\big([0,2\pi[, \mathcal{L}\text{eb}\big([0,2\pi[\big), \lambda\big). \tag{18.204}$$

18.6 Exemples trigonométriques

Nous mettons ici quelques exemples concernant les fonctions trigonométriques, qui n'ont pas pu être mis dans les chapitres les plus adaptés, parce que ces derniers sont plus haut dans la table des matières.

EXooSPFDooSluUGV

Exemple 18.73.
Prouvons que la fonction[22] $f(x) = x\sin(x)$ tend vers zéro lorsque x tend vers 0. D'abord, nous coinçons la fonction entre deux fonctions connues :

$$0 \leqslant |x\sin(x)| = |x||\sin(x)| \leqslant |x|. \tag{18.205}$$

Donc $|x\sin(x)|$ est coincé entre $g(x) = 0$ et $h(x) = |x|$. Ces deux fonctions tendent vers 0 lorsque $x \to 0$, et donc $f(x)$ tend vers zéro. △

22. La définition de la fonction sinus est 18.1.

18.6.1 Quelques équations trigonométriques

La proposition suivante se voit très facilement sur le cercle trigonométrique, mais il faut le démontrer.

PROPooTUUUooVrAGQo

Proposition 18.74 ([1]).
Si $\theta_0 \in [0, 2\pi[$ vérifie $\cos(\theta_0) = x_0$, alors l'ensemble des solutions de l'équation $\cos(\theta) = x_0$ (d'inconnue θ) est

$$\{\theta_0, 2\pi - \theta_0\}. \tag{18.206}$$

Cet ensemble est un singleton si et seulement si $x_0 = \pm 1$.

Démonstration. Commençons par prouver que θ_0 et $2\pi - \theta_0$ sont des solutions. Le nombre θ_0 est solution par hypothèse. En ce qui concerne $2\pi - \theta_0$, il est possible d'utiliser la formule d'addition d'angle (18.22c) :

EQooUCAOooTQsUUq
$$\cos(2\pi - \theta_0) = \cos(2\pi)\cos(\theta_0) + \sin(2\pi)\sin(\theta_0). \tag{18.207}$$

La proposition 18.20(1) nous indique que $\cos(2\pi) = 1$ et $\sin(2\pi) = 0$. Donc l'égalité (18.207) se réduit à $\cos(2\pi - \theta_0) = \cos(\theta_0)$.

Le lemme 18.4 dit que si $\cos(\theta) = x_0$, alors

$$\sin(\theta) = \pm\sqrt{1 - x_0^2}. \tag{18.208}$$

Nous avons donc soit

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} x_0 \\ \sqrt{1 - x_0^2} \end{pmatrix}, \tag{18.209}$$

soit

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} x_0 \\ -\sqrt{1 - x_0^2} \end{pmatrix}, \tag{18.210}$$

Comme $\theta \mapsto \big(\cos(\theta), \sin(\theta)\big)$ est une bijection avec S^1 (proposition 18.58), chacune de ces deux possibilités possède une unique solution. L'ensemble des solutions de $\cos(\theta) = x_0$ possède donc au maximum deux éléments.

L'ensemble des solutions possède exactement une solution lorsque les points $\big(x_0, \sqrt{1 - x_0^2}\big)$ et $\big(x_0, -\sqrt{1 - x_0^2}\big)$ sont identiques. C'est le cas, si et seulement si $\sqrt{1 - x_0^2} = 0$, c'est-à-dire, si et seulement si $x_0 = \pm 1$. □

18.6.2 Développements en série

PROPooNPYXooTuwAHP

Proposition 18.75 (Taylor pour cosinus).
Pour tout $n \in \mathbb{N}$, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

EQooGQOIooIkwbJV
$$\cos(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + \alpha(x)x^{2n}. \tag{18.211}$$

En ce qui concerne le sinus, pour tout n nous avons une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

EQooKYJAooRebHgc
$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + x^{2n+2}\alpha(x). \tag{18.212}$$

Démonstration. Il s'agit d'utiliser la proposition 15.51, en faisant attention à l'ordre. Le fait est que dans (18.211), nous avons écrit le polynôme de degré $2n+1$ (et non seulement $2n$), en sachant que le terme d'ordre $2n+1$ est nul.

C'est pour cela que nous avons pu écrire $\alpha(x)x^{2n+1}$ au lieu de $\alpha(x)x^{2n}$ qui aurait été attendu.

Même raisonnement pour le développement du sinus. □

Remarque 18.76.
Quelques remarques concernant l'ordre du polynôme.

(1) Notons que nous aurions aussi pu écrire le reste sous la forme $\alpha(x)x^{2n}$, mais ça aurait été avec une autre fonction α : celle correspondant au développement à l'ordre $2n$ au lieu de $2n+1$.
(2) Les développements de sinus et de cosinus ont un terme sur deux qui est nul. C'est pour cela qu'en ayant un polynôme de degré $2p$, nous avons le développement d'ordre $2p+1$.
(3) Nous aurions pu utiliser les dérivées données dans la proposition 18.3 et les valeurs spéciales (18.6).

Corolaire 18.77.
Il existe une fonction $\alpha\colon \mathbb{R}\to\mathbb{R}$ *telle que* $\lim_{t\to 0}\alpha(t)/t=0$ *et*

$$\sin(x)=x+\alpha(x). \tag{18.213}$$

EQooDLGIooXyfmtC

Nous avons la limite

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1. \tag{18.214}$$

Démonstration. Il s'agit de prendre la formule (18.212) avec $n=0$. Cela donne tout de suite (18.213). Pour la limite, on divise par x, ce qui donne (pour tout $x\neq 0$)

$$\frac{\sin(x)}{x}=1+\frac{\alpha(x)}{x}. \tag{18.215}$$

Et justement la fonction α a la propriété que $\lim_{x\to 0}\alpha(x)/x=0$. □

Exemple 18.78.
Cherchons le développement limité à l'ordre 5 de $\tan(x)=\frac{\sin(x)}{\cos(x)}$. Nous utilisons les développements de la proposition 18.75 :

$$\sin(x)=x-\frac{x^3}{6}+\frac{x^5}{120}+x^5\alpha_1(x) \tag{18.216a}$$

$$\cos(x)=1-\frac{x^2}{2}+\frac{x^4}{24}+x^5\alpha_2(x). \tag{18.216b}$$

Nous calculons alors la division des deux polynômes, en classant les puissances dans l'ordre croissant (c'est le sens inverse de ce qui est fait pour la divisions euclidienne !) :

$$\begin{array}{l|l}
x-\frac{1}{6}x^3+\frac{1}{120}x^5 & 1-\frac{1}{2}x^2+\frac{1}{24}x^4 \\
\quad -\left(x-\frac{1}{2}x^3+\frac{1}{24}x^5\right) & \overline{\phantom{1-\frac{1}{2}x^2+\frac{1}{24}x^4}} \\
\overline{\phantom{x-\frac{1}{6}x^3+\frac{1}{120}x^5}} & x+\frac{1}{3}x^3+\frac{2}{15}x^5 \\
\quad \frac{1}{3}x^3-\frac{1}{30}x^5 & \\
\quad -\left(\frac{1}{3}x^3-\frac{1}{6}x^5+\frac{1}{72}x^7\right) & \\
\quad \overline{\phantom{\frac{1}{3}x^3-\frac{1}{30}x^5}} & \\
\quad \frac{2}{15}x^5-\frac{1}{72}x^7 & \\
\quad -\left(\frac{2}{15}x^5-\frac{1}{15}x^7+\frac{1}{180}x^9\right) & \\
\quad \overline{\phantom{\frac{2}{15}x^5-\frac{1}{72}x^7}} & \\
\quad \frac{29}{360}x^7-\frac{1}{180}x^9 &
\end{array}$$

Nous avons continué la division jusqu'à obtenir un reste de degré plus grand que 5. Le développement à l'ordre 5 de la fonction tangente autour de zéro est alors (proposition 12.485)

$$\tan(x)=x+\frac{1}{3}x^3+\frac{2}{15}x^5+x^5\alpha(x). \tag{18.217}$$

Notons que, puisque le reste ne nous intéressait pas vraiment, nous aurions pu ne pas calculer les coefficients des termes en x^7 et x^8. La dernière soustraction était également inutile. △

18.7 Isométries de l'espace euclidien

Nous considérons l'espace affine euclidien $A=\mathcal{E}_n(\mathbb{R})$ modelé sur $\mathbb{R}^n$ avec sa métrique usuelle. Un premier grand résultat sera le théorème 8.157 qui dira que les isométries de cet espace sont des applications linéaires.

18.7.1 Structure du groupe $\mathrm{Isom}(\mathbb{R}^n)$

Si vous ne voulez pas savoir ce qu'est un produit semi-direct de groupes, vous pouvez lire seulement le point (1) du théorème suivant, et passer directement à la remarque 18.80.

THOooQJSRooMrqQct

Théorème 18.79.
Un peu de structure sur $\mathrm{Isom}(\mathbb{R}^n)$.

ITEMooLLUIooIGsknv

(1) *L'application*

$$\begin{aligned}\psi\colon T(n)\times \mathrm{O}(n) &\to \mathrm{Isom}(\mathbb{R}^n)\\ (v,\Lambda) &\mapsto \tau_v\circ\Lambda\end{aligned} \tag{18.218}$$

est une bijection. Ici, $T(n)$ *est le groupe des translations de* $\mathbb{R}^n$.

(2) *Un couple* $(v,\Lambda)\in T(n)\times \mathrm{SO}(n)$ *agit sur* $x\in\mathbb{R}^n$ *par*

$$(v,\Lambda)x=\Lambda x+v \tag{18.219}$$

au sens où $\psi(v,\Lambda)x=\Lambda x+v$.

ITEMooEWSIooNKzRxB

(3) *En tant que groupes,*

$$\mathrm{Isom}(\mathbb{R}^n)\simeq T(n)\times_\rho \mathrm{O}(n) \tag{18.220}$$

où ρ *représente l'action adjointe de* $\mathrm{O}(n)$ *sur* $T(n)$ *et* $\times_\rho$ *dénote le produit semi-direct de la définition 2.47.*

ITEMooSKUPooBDvNWX

(4) *Une isométrie de* $\mathbb{R}^n$ *est une application affine*[23].

ITEMooQLNPooSyHaps

(5) *La partie linéaire*[24] *d'une isométrie* f *est* $\psi^{-1}(f)_2$.

Démonstration. Point par point.

(1) Prouvons que l'application proposée est injective et surjective. Notons aussi que ce point ne parle pas de structure de groupe, mais seulement d'une bijection en tant qu'ensembles.

(1a) **Injection** Si $\psi(v,\Lambda)=\psi(w,\Lambda')$ alors en appliquant sur $x=0$ nous avons tout de suite $v=w$. Et ensuite $\Lambda=\Lambda'$ est immédiat.

(1b) **Surjection** Une isométrie $g\in\mathrm{Isom}(\mathbb{R}^n)$ est une application $g\colon\mathbb{R}^n\to\mathbb{R}^n$ telle que $d(x,y)=d\big(g(x),g(y)\big)$. Dans le cas de $\mathbb{R}^n$ cela se traduit par

$$\|x-y\|=\|g(x)-g(y)\|, \tag{18.221}$$

Comme $x\mapsto\|x\|^2$ est une forme quadratique, elle tombe sous le coup du théorème 8.157, ce qui nous permet de dire que g est affine. Or par définition une application est affine lorsqu'elle est la composée d'une translation et d'une application linéaire.

Donc $g=\tau_v\circ\Lambda$ pour une certaine application linéaire isométrique $\Lambda\colon\mathbb{R}^n\to\mathbb{R}^n$. L'application Λ est donc dans $\mathrm{O}(n)$ par la proposition 8.46(3).

(2) C'est seulement le fait que $(\tau_v\circ\Lambda)x=\tau_v\big(\Lambda x\big)=\Lambda(x)+v$.

(3) Nous allons étudier l'application

$$\psi\colon T(n)\times_\rho O(n)\to\mathrm{Isom}(\mathbb{R}^n). \tag{18.222}$$

(3a) **Le produit semi-direct est bien défini** Il faut montrer que

$$\begin{aligned}\rho\colon O(n)&\to \mathrm{Aut}\big(T(n)\big)\\ \Lambda&\mapsto \mathbf{Ad}(\Lambda)\end{aligned} \tag{18.223}$$

est correcte.

23. Définition 11.12.
24. Définition 11.13.

D'abord pour $\Lambda \in O(n)$, nous avons bien $\rho_\Lambda(\tau_v) \in T(n)$ parce qu'en appliquant à $x \in \mathbb{R}^n$,

$$(\Lambda\tau_v\Lambda^{-1})(x) = \Lambda\big(\tau_v(\Lambda^{-1}x)\big) = \Lambda\big(\Lambda^{-1}x + v\big) = x + \Lambda(v) = \tau_{\Lambda(v)}(x). \tag{18.224}$$

Donc $\rho_\Lambda(\tau_v) = \tau_{\Lambda(v)}$.
De plus, $\rho_\Lambda \in \mathrm{Aut}\big(T(n)\big)$ parce que

$$\rho_\Lambda\big(\tau_v \circ \tau_w\big) = \rho_\Lambda(\tau_v) \circ \rho_\Lambda(\tau_w), \tag{18.225}$$

comme on peut aisément vérifier que les deux membres sont égaux à $\tau_{\Lambda(v+w)}$.

(3b) ψ **est une bijection** Cela est déjà vérifié.

(3c) ψ **est un morphisme** Nous avons d'une part

$$\psi\big((v,g)(w,h)\big) = \psi\big(v\rho_g(w), gh\big) = \tau_v \circ g \circ \tau_w \circ g^{-1} \circ g \circ h = \tau_v \circ g \circ \tau_w \circ h. \tag{18.226}$$

Et d'autre part,

$$\psi(v,g) \circ \psi(w,h) = \tau_v \circ g \circ \tau_w \circ h, \tag{18.227}$$

ce qui est la même chose.

(4) Si f est une isométrie de $\mathbb{R}^n$, et si $\psi(f) = (v, \Lambda)$, nous avons

$$f(a+x) = \tau_v\big(\Lambda(a) + \Lambda(x)\big) = \Lambda(a) + \Lambda(x) + v = f(a) + \Lambda(x). \tag{18.228}$$

Donc f est affine et l'application qui serait notée u_M dans (11.27) est $u_M = \Lambda$ pour tout M.

(5) Si $f(x) = \psi(v,\Lambda)x = \Lambda x + v$, nous avons

$$f(a+x) = \psi(v,\Lambda)(a+x) = f(a) + \Lambda(x), \tag{18.229}$$

et donc Λ est bien la partie linéaire de f.

□

REMooLUEZooIwvTqu

Remarque 18.80.
Notons au passage la loi de groupe sur les couples qui est donnée, pour tout $v, v' \in \mathbb{R}^n$, $\Lambda, \Lambda' \in \mathrm{SO}(n)$, par

$$(v,\Lambda) \cdot (v',\Lambda') = (\Lambda v' + v, \Lambda\Lambda') \tag{18.230}$$

EqDiHcut

comme le montre le calcul suivant :

$$\begin{aligned}(v,\Lambda) \cdot (v',\Lambda')x &= (v,\Lambda)(\Lambda' x + v') && \text{(18.231a)}\\ &= \Lambda\Lambda' x + \Lambda v' + v && \text{(18.231b)}\\ &= (\Lambda v' + v, \Lambda\Lambda')x. && \text{(18.231c)}\end{aligned}$$

PROPooDHYWooXxEXvl

Proposition 18.81 ([545]).
Soient $n \geqslant 1$ et R un élément de $\mathrm{O}(n)$ de déterminant -1 tels que $R^2 = \mathrm{Id}$. En posant $C_2 = \{\mathrm{Id}, R\}$ nous avons

$$\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2 \tag{18.232}$$

Démonstration. Notons qu'un élément R comme décrit dans l'énoncé existe. Par exemple il y a l'application $(x_1, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$.

Cela étant dit, nous allons montrer que

$$\begin{aligned}\psi\colon \mathrm{SO}(n) \times C_2 &\to \mathrm{O}(n)\\ (A,h) &\mapsto Ah.\end{aligned} \tag{18.233}$$

est un isomorphisme.

(1) **Injectif** Soient $A, B \in \mathrm{SO}(n)$ et $h, k \in C_2$ tels que $\psi(A,h) = \psi(B,k)$, c'est-à-dire tels que $Ah = Bk$. Puisque $\det(A) = \det(B) = 1$, nous avons $\det(h) = \det(k)$. Mais comme C_2 contient un élément de déterminant 1 et un élément de déterminant -1, nous avons $h = k$. De là $A = B$.

(2) **Surjectif** Soit $X \in \mathrm{O}(n)$. Si $\det(X) = 1$ alors $X \in \mathrm{SO}(n)$ et $X = \psi(X, \mathbb{1})$. Si par contre $\det(X) = -1$, alors $XR \in \mathrm{SO}(n)$ parce que $\det(XR) = 1$, et nous avons

$$\psi(XR, R) = XR^2 = X. \tag{18.234}$$

(3) **Morphisme** Nous avons

$$\psi\Big((A,h)(B,k)\Big) = \psi\big(A\rho_h(B), hk\big) = A(hBh^{-1})hk = AhBk, \tag{18.235}$$

tandis que

$$\psi(A,h)\psi(B,k) = AhBk, \tag{18.236}$$

qui est la même chose.

□

18.8 Isométries dans $\mathbb{R}^n$

Définition 18.82.
*Un **hyperplan** de $\mathbb{R}^n$ est un sous-espace affine de dimension $n-1$.*

LEMooPXIOooSlMIIY

Lemme-Définition 18.83.
Si un hyperplan H de $\mathbb{R}^n$ est donné, et si $x \in \mathbb{R}^n$, il existe un unique point $y \in \mathbb{R}^n$ tel que

(1) $x - y \perp H$,

(2) Le segment $[x,y]$ coupe H en son milieu.

*La **réflexion** σ_H est l'application $\sigma_H \colon \mathbb{R}^n \to \mathbb{R}^n$ qui à x fait correspondre ce y.*

Démonstration. Il faut vérifier que les conditions données définissent effectivement un unique point de $\mathbb{R}^n$. Soit H_0 le sous-espace vectoriel parallèle à H et une base orthonormée $\{e_1, \dots, e_{n-1}\}$ de H_0. Nous complétons cette partie en une base orthonormée de $\mathbb{R}^n$ avec un vecteur e_n. Si $H = H_0 + v$, quitte à décomposer v en une partie parallèle et une partie perpendiculaire à H, nous avons

$$H = H_0 + \lambda e_n \tag{18.237}$$

pour un certain λ.

Une droite passant par x et perpendiculaire à H est de la forme $t \mapsto x + te_n$. Si $x = \sum_{i=1}^n x_i e_i$ alors l'unique point de cette droite à être dans H est le point tel que $x_n e_n + te_n = \lambda e_n$, c'est-à-dire $t = \lambda - x_n$. L'unique point y sur cette droite à être tel que $[x,y]$ coupe H en son milieu est celui qui correspond à $t = 2(\lambda - x_n)$. □

Lemme 18.84.
Soit un hyperplan vectoriel H_0 de $\mathbb{R}^n$. Nous considérons une base $\{e_1, \dots, e_n\}$ de $\mathbb{R}^n$ telle que $\{e_1, \dots, e_{n-1}\}$ soit une base de H_0 et $e_n \perp H_0$.

Soit l'hyperplan affine $H = H_0 + \lambda e_n$.

La réflexion de plan H est donnée par

$$\begin{aligned} \sigma \colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto x - 2(x_n - \lambda)e_n. \end{aligned} \tag{18.238}$$

EQooRTWLooLPsUpY

Démonstration. En posant $y = x - 2(x_n - \lambda)$, nous avons

$$x - y = 2(x_n - \lambda)e_n \perp H, \tag{18.239}$$

et

$$\frac{x+y}{2} = \frac{2x - 2(x_n - \lambda)e_n}{2} = x - (x_n - \lambda)e_n = \sum_{i=1}^{n-1} x_i e_i + \lambda e_n \in H_0 + \lambda e_n = H. \tag{18.240}$$

Les deux conditions de la définition 18.83 sont donc vérifiées. □

LEMooWYVRooQmWqvM

Lemme 18.85.
Si H_0 est un plan vectoriel, si $\{e_1, \dots, e_{n-1}\}$ est une base de H_0, si $e_n \perp H_0$ et si $H = H_0 + \lambda e_n$, alors

$$\sigma_H = \sigma_{H_0} + 2\lambda e_n. \tag{18.241}$$

Démonstration. En posant $\lambda = 0$ dans la formule (18.238), nous avons $\sigma_{H_0}(x) = x - 2x_n e_n$. Nous avons donc

$$\sigma_H(x) = x + 2(\lambda - x_n)e_n = x - 2x_n e_n + 2\lambda e_n = \sigma_{H_0}(x) + 2\lambda e_n. \tag{18.242}$$

□

Le lemme suivant est une généralisation du fait que tous les points de la médiatrice d'un segment sont à égale distance des deux extrémités du segment (très utile lorsqu'on étudie les triangles isocèles).

LEMooDPLYooJKZxiM

Lemme 18.86 ([545]).
Soient deux points distincts $x_0, y_0 \in \mathbb{R}^n$ l'ensemble $H \subset \mathbb{R}^n$ donné par

$$H = \{x \in \mathbb{R}^n \text{ tel que } d(x, x_0) = d(x, y_0)\}. \tag{18.243}$$

Alors

(1) H est hyperplan.

(2) $H \perp y_0 - x_0$

(3) H contient le milieu du segment $[x_0, y_0]$.

Démonstration. Nous savons que

$$d(x, x_0)^2 = \langle x - x_0, x - x_0 \rangle = \|x\|^2 + \|x_0\|^2 - 2\langle x, x_0 \rangle, \tag{18.244}$$

ou encore

$$\|x_0\|^2 - \|y_0\|^2 = 2\langle x, x_0 - y_0 \rangle. \tag{18.245}$$

En posant $v = y_0 - x_0$ et en considérant la forme linéaire

$$\begin{aligned} \beta \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \langle x, v \rangle, \end{aligned} \tag{18.246}$$

Nous avons $x \in H$ si et seulement si $\beta(x) = \frac{1}{2}\big(\|y_0\|^2 - \|x_0\|^2\big) = \lambda$. En d'autres termes, $H = \beta^{-1}(\lambda)$. Par la proposition 11.27 la partie H est un sous-espace affine. C'est même un translaté de $\ker(\beta)$, et comme $\ker(\beta)$ est l'espace vectoriel des vecteurs perpendiculaires à v, nous avons $\dim(H) = \dim\big(\ker(\beta)\big) = n - 1$.

Le fait que H contienne le milieu du segment $[x_0, y_0]$ est par définition. □

Pour le lemme suivant, et pour que la récurrence se passe bien nous disons que l'ensemble vide est un espace vectoriel de dimension -1.

LEMooJCDRooGAmlwp

Lemme 18.87 (Cartan-Dieudonné[546, 547]).
Soit un espace euclidien E de dimension n.

ITEMooFYEDooIJZBjP

(1) *Si f est une isométrie de E satisfaisant*

$$\dim\big(\mathrm{Fix}(f)\big) = n - k \tag{18.247}$$

alors f peut être écrit comme composition de k réflexions hyperplanes. ITEMooJTZVooWvyfDD

(2) *Une isométrie de E peut être écrite sous la forme de $\mathrm{rk}(f - \mathrm{Id})$ réflexions, mais pas moins.* ITEMooUCZWooSbyPwt

(3) *Toute isométrie de E peut être écrite comme composition de $n + 1$ réflexions.*

Démonstration. Les deux parties importantes à démontrer sont les points (1) et la partie « pas moins » de (2). Le reste proviendra de reformulations.

(1) **Pour (1)** Nous faisons une récurrence sur $k \geqslant 0$.

Pour l'initialisation, si $k = 0$ alors $\dim\big(\mathrm{Fix}(f)\big) = n$, c'est-à-dire que f fixe tout $\mathbb{R}^n$, autant dire que f est l'identité, une composition de zéro réflexions.

Pour la récurrence, nous supposons que le lemme est démontré jusqu'à $k \geqslant 0$. Soit donc $f \in \mathrm{Isom}(\mathbb{R}^n)$ tel que

$$\dim\big(\mathrm{Fix}(f)\big) = n - (k + 1). \tag{18.248}$$

Puisque $k \geqslant 0$, la dimension de $\mathrm{Fix}(f)$ est strictement plus petite que n, donc il existe un $x_0 \in \mathbb{R}^n$ tel que $f(x_0) \neq x_0$. Nous posons

$$H = \{x \in E \text{ tel que } d(x, x_0) = d\big(x, f(x_0)\big)\}. \tag{18.249}$$

Par le lemme 18.86, ce H est l'hyperplan orthogonal à $v = f(x_0) - x_0$ et passant par le milieu du segment $[x_0, f(x_0)]$.

Nous posons $g = \sigma_H \circ f$. Comme $g(x_0) = \sigma_H(f(x_0)) = x_0$, ce x_0 est un point fixe de g. Le fait que $\sigma_H\big(f(x_0)\big) = x_0$ est vraiment la définition de l'hyperplan H.

Nous avons donc

$$x_0 \in \mathrm{Fix}(g) \backslash \mathrm{Fix}(f). \tag{18.250}$$

Mais nous prouvons de plus que $\mathrm{Fix}(f) \subset \mathrm{Fix}(g)$. En effet si $y \in \mathrm{Fix}(f)$ alors $y \in H$ parce que

$$d(y, x_0) = d\big(f(y), f(x_0)\big) = d\big(y, f(x_0)\big). \tag{18.251}$$

Puisque $y \in H$ nous avons $y \in \mathrm{Fix}(g)$ parce que

$$g(y) = \sigma_H\big(f(y)\big) = \sigma_H(y) = y. \tag{18.252}$$

Tout cela pour dire que l'ensemble $\mathrm{Fix}(g)$ est *strictement* plus grand que $\mathrm{Fix}(f)$. Et comme ce sont des espaces affines, nous pouvons parler de dimension :

$$\dim\big(\mathrm{Fix}(g)\big) > \dim\big(\mathrm{Fix}(f)\big). \tag{18.253}$$

Par hypothèse de récurrence, l'application g peut être écrite comme composition de k réflexions. Donc l'application

$$f = \sigma_H \circ g \tag{18.254}$$

est une composition de $k + 1$ réflexions.

(2) **Pour (2), existence**

Ce point est une reformulation du point (1). Le fait est que $\mathrm{Fix}(f) = \ker(f - \mathrm{Id})$ parce que $x \in \mathrm{Fix}(f)$ si et seulement si $f(x) = x$ si et seulement si $(f - \mathrm{Id})x = 0$. Nous utilisons le théorème du rang 4.47 à l'endomorphisme $f - \mathrm{Id}$:

$$\dim\big(\mathrm{Fix}(f)\big) = \dim\big(\ker(f - \mathrm{Id})\big) = \dim(E) - \mathrm{rk}(f - \mathrm{Id}). \tag{18.255}$$

En remplaçant par les valeurs :

$$n - k = n - \mathrm{rk}(f - \mathrm{Id}). \tag{18.256}$$

Or le point (1) donnait f comme composée de $n - k$ réflexions. Donc f est composée de $\mathrm{rk}(f - \mathrm{Id})$ réflexions.

(3) **Pour (2), « pas moins »**
Supposons que $f = \sigma_1 \circ \ldots \circ \sigma_r$ où σ_i est la réflexion de l'hyperplan H_i. Nous devons prouver que $r \geqslant \mathrm{rk}(f - \mathrm{Id})$. Nous avons

$$\bigcap_{i=1}^{r} H_i \subset \ker(f - \mathrm{Id}). \tag{18.257}$$

D'autre part, la proposition 8.310 nous donne $\dim \bigcap_i H_i \geqslant n - r$. Donc

$$n - r \leqslant \dim\Big(\bigcap_{i=1}^{r} H_i\Big) \leqslant \dim\big(\ker(f - \mathrm{Id})\big) = n - \mathrm{rk}(f - \mathrm{Id}). \tag{18.258}$$

Donc $n - r \leqslant n - \mathrm{rk}(f - \mathrm{Id})$ ou encore

$$r \geqslant \mathrm{rk}(f - \mathrm{Id}). \tag{18.259}$$

(4) **Pour (3)** Le rang de $f - \mathrm{Id}$ vaut au maximum n. Donc f peut être écrite comme composition de $n + 1$ réflexions par le point (2).

□
PROPooUSKEooUbNVfs

Proposition 18.88.
Un élément de SO(3) *qui fixe deux vecteurs linéairement indépendants est l'identité.*

Démonstration. Soit un élément $A \in \mathrm{SO}(3)$ et deux vecteurs linéairement indépendants $v_1, v_2 \in \mathbb{R}^3$ tels que $Av_1 = v_1$ et $Av_2 = v_2$. Puisque v_1 et v_2 sont linéairement indépendants, le théorème de la base incomplète 4.14 nous permet de considérer $v_3 \in \mathbb{R}^3$ tel que $\{v_1, v_2, v_3\}$ soit une base. Dans cette base, la matrice de A est de la forme

$$A = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}. \tag{18.260}$$

Le déterminant de cette matrice est c. Or $\det(A) = 1$ parce qu'elle est dans SO(3). Donc $c = 1$. Le fait que A soit orthogonale implique que la troisième colonne doit être un vecteur de norme 1. Donc $a = b = 0$.

Donc $A = \mathrm{Id}$. □
CORooJCURooSRzSFb

Corolaire 18.89.
Tout élément de SO(3) *peut être écrit comme composée de deux réflexions.*

Démonstration. Un élément de SO(3) est une isométrie de $\mathbb{R}^3$ parce que si $A \in \mathrm{SO}(3)$ alors [25]

$$\langle Ax, Ay\rangle = \langle A^*Ax, y\rangle = \langle x, y\rangle. \tag{18.261}$$

Donc si le rang de A est k, alors A est la composée de $3 - k$ réflexions par le lemme 18.87.

Si $A = \mathrm{Id}$, c'est bon parce que l'identité est la composée de deux réflexions égales. Nous supposons que A n'est pas l'identité.

Comme discuté dans l'exemple 12.92, l'opérateur A possède trois valeurs propres dans $\mathbb{C}$ dont une réelle, et deux complexes conjuguées. Nous les notons $\lambda \in \mathbb{R}$ et $\alpha, \bar{\alpha} \in \mathbb{C}$. Le déterminant de A, qui vaut 1, est le produit de ces trois valeurs propres, c'est-à-dire $\lambda|\alpha|^2$. En particulier $\lambda > 0$.

Si v est un vecteur propre correspondant à la valeur propre λ, nous avons $\|v\| = \|Av\| = |\lambda|\|v\|$ parce que A est une isométrie. Donc $\lambda = \pm 1$.

Au final, $\lambda = 1$. Cela signifie que A laisse au moins un vecteur invariant. Vu que A n'est pas l'identité, la proposition 18.88 nous indique qu'il n'y a pas d'autres vecteurs de $\mathbb{R}^3$ à être fixé par A. Donc $\dim\big(\mathrm{Fix}(A)\big) = 1$ et le lemme 18.87(1) s'écrit avec $n = 3$, $k = 2$ et implique que A est la composée de deux réflexions. □

25. Opérateur orthogonal, définition 8.43.

LEMooMCVKooKzmlAg

Lemme 18.90.
Soit un hyperplan H et un vecteur v de $\mathbb{R}^n$. Nous avons

$$\tau_v \circ \sigma_H \circ \tau_v^{-1} = \sigma_{\tau_v(H)}. \tag{18.262}$$

Démonstration. Pour ce faire, nous considérons une base adaptée. Les vecteurs $\{e_1, \ldots, e_{n-1}\}$ forment une base orthonormée de H_0 et e_n complète en une base orthonormée de $\mathbb{R}^n$. Soit H_0 l'hyperplan parallèle à H et passant par l'origine ; nous avons, pour un certain $\lambda \in \mathbb{R}$,

$$H = H_0 + \lambda e_n \tag{18.263}$$

D'un autre côté, le vecteur v peut être décomposé en $v = v_1 + v_2$ où $v_1 \perp H$ et $v_2 \parallel H$. Alors

$$\tau_v(H) = H + v = H + v_2 = H_0 + \lambda e_n + v_2. \tag{18.264}$$

Nous pouvons maintenant utiliser le lemme 18.85 pour exprimer la transformation $\sigma_{\tau_v(H)}$:

$$\sigma_{\tau_v(H)}(x) = \sigma_{H_0}(x) + 2\lambda e_n + 2v_2 \tag{18.265}$$

EQooNYKFooXprXav

Mais d'autre part,

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_H(x - v) = v + \sigma_{H_0}(x - v) + 2\lambda e_n. \tag{18.266}$$

Vue la décomposition de $v = v_1 + v_2$ nous avons $\sigma_{H_0}(v) = -v_1 + v_2$ et donc

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_{H_0}(x) + v_1 - v_2 + 2\lambda e_n = \sigma_{H_0} + 2v_1 + 2\lambda e_n. \tag{18.267}$$

EQooGOHEooALPRER

Les expressions (18.265) et (18.267) coïncident, d'où l'égalité recherchée. □

PROPooLYCUooRQgGtF

Proposition 18.91.
Soit un hyperplan H de $\mathbb{R}^n$ passant par l'origine. Il existe une base orthonormée $\{e_1, \ldots, e_n\}$ de $\mathbb{R}^n$ telle que $\{e_1, \ldots, e_{n-1}\}$ soit une base orthonormée de H et $e_n \perp H$.

18.8.1 Préserver l'orientation

Le fait qu'une isométrie puisse être décomposé en réflexions est le lemme 18.87.

DEFooUZFHooXVVLBL

Proposition-Définition 18.92 (Isométrie positive et négative[545])**.**
Nous utilisons la bijection $\psi\colon T(n) \times \mathrm{O}(n) \to \mathrm{Isom}(\mathbb{R}^n)$ décrite par le théorème 18.79, et nous définissons

$$\begin{aligned} \epsilon\colon \mathrm{Isom}(\mathbb{R}^n) &\to \{\pm 1\} \\ f &\mapsto \det\big(\psi^{-1}(f)_2\big). \end{aligned} \tag{18.268}$$

Cette application ϵ est un morphisme de groupes.

*Nous disons que $f \in \mathrm{Isom}(\mathbb{R}^n)$ est **positive** ou **préserve l'orientation** si $\epsilon(f) = 1$, et nous notons*

$$\mathrm{Isom}^+(\mathbb{R}^n) = \epsilon^{-1}(1). \tag{18.269}$$

LEMooVRELooESIWQl

Lemme 18.93.
La partie $\mathrm{Isom}^+(\mathbb{R}^n)$ des isométries positives est un sous-groupe de $\mathrm{Isom}(\mathbb{R}^n)$.

LEMooJABDooOKHwWv

Lemme 18.94 ([545])**.**
Si H est un hyperplan et si σ_H est la réflexion de cet hyperplan, alors $\epsilon(\sigma_H) = -1$.

Démonstration. Nous commençons par supposer que H passe par l'origine, de telle sorte que σ_H soit linéaire, et que $\epsilon(\sigma_H) = \det(\sigma_H)$.

Nous choisissons une base orthonormée de $\mathbb{R}^n$ comme dans la proposition 18.91 : $\{e_1, \dots, e_{n-1}\}$ est une base de H et $e_n \perp H$. Nous avons alors

$$\sigma_H(e_i) = \begin{cases} e_i & \text{si } i = 1, \dots, n-1 \\ -e_n & \text{si } i = n. \end{cases} \tag{18.270}$$

Calculons ce déterminant à l'ancienne, en utilisant les définitions 8.15 et 8.10. Nous posons $\mu_i = 1$ pour $i = 1, \dots, n-1$ et $\mu_n = -1$, de telle sorte à avoir $\sigma_H(e_i) = \mu_i e_i$. Nous posons aussi $B = (e_1, \dots, e_n)$. Ensuite c'est parti pour le calcul :

$$\det_B\big(f(B)\big) = \sum_{s\in S_n} \epsilon(s) \prod_{i=1}^{n} e^*_{s(i)}\big(f(e_i)\big) \tag{18.271a}$$

$$= \sum_{s\in S_n} \epsilon(s) \prod_{i=1}^{n} \langle e_{s(i)}, \mu_i e_i\rangle \tag{18.271b}$$

$$= \sum_{s\in S_n} \epsilon(s) \prod_{i=1}^{n} \mu_i \langle e_{s(i)}, e_i\rangle. \tag{18.271c}$$

Vu que la base est orthogonale, $\langle e_{s(i)}, e_i\rangle = \delta_{s(i),i}$ et il ne reste, dans la somme sur S_n, que le terme $s = \mathrm{Id}$ dont la signature est $\epsilon(\mathrm{Id}) = 1$. Donc

$$\det_B\big(f(B)\big) = \prod_{i=1}^{n} \mu_i = -1. \tag{18.272}$$

Et si H ne passe pas par l'origine ? Soit $v \in \mathbb{R}^n$ tel que $\tau_v(H)$ passe par l'origine (prendre pour v l'opposé de n'importe que élément de H). Le lemme 18.90 nous indique que

$$\sigma_{\tau_v(H)} = \tau_v \circ \sigma_H \circ \tau_v^{-1}. \tag{18.273}$$

Comme ϵ est un morphisme,

$$\epsilon\big(\sigma_{\tau_v(H)}\big) = \epsilon(\tau_v \sigma_H \tau_v^{-1}) = \epsilon(\tau_v)\epsilon(\sigma_H)\epsilon(\tau_v^{-1}) = \epsilon(\sigma_H). \tag{18.274}$$

Nous avons utilisé le fait que $\epsilon(g^{-1}) = \epsilon(g)^{-1}$. Tout ça pour dire que

$$\epsilon(\sigma_H) = \epsilon(\sigma_{\tau_v(H)}) = -1 \tag{18.275}$$

parce que $\tau_v(H)$ passe par l'origine. □

THOooQEWRooYeOIfZ

Théorème 18.95 ([545]).
Une isométrie de $(\mathbb{R}^n, d)$ préserve l'orientation[26] *si et seulement si elle est composition d'un nombre pair de réflexions hyperplanes.*

Démonstration. Le lemme 18.87 nous indique qu'une isométrie f de $\mathbb{R}^n$ peut être décomposée en réflexions hyperplanes. Le lemme 18.94 nous dit que chacune de ces réflexions est négative. Donc si $f = \sigma_1 \circ \dots \circ \sigma_r$, alors

$$\epsilon(f) = \epsilon(\sigma_1) \dots \epsilon(\sigma_r) = (-1)^r. \tag{18.276}$$

□

26. Définition 18.92.

18.9 Groupes finis d'isométries

DEFooCUYLooAlbtzv

Définition 18.96.
Si X est une partie finie de $\mathbb{R}^n$, le ***barycentre*** *de X est le point*

$$B_X = \frac{1}{|X|}\sum_{x\in X} x \tag{18.277}$$

où $|X|$ est le cardinal de X.

Cela est à mettre en relation avec la définition dans le cadre affine 11.30.

LEMooSEZYooYceLIb

Lemme 18.97 ([545]).
Les applications affines de $\mathbb{R}^n$ préservent le barycentre[27] *des parties finies.*

Démonstration. Soit une partie finie X de $\mathbb{R}^n$ et une application affine $f \in \mathrm{Aff}(\mathbb{R}^n)$. Nous devons prouver que

$$f(B_X) = B_{f(X)}. \tag{18.278}$$

Nous savons que toute application affine est une composée de translation et d'une application linéaire : $f = \tau_v \circ g$ avec $v \in \mathbb{R}^n$ et $g \in \mathrm{End}(n, \mathbb{R})$. Nous vérifions le résultat séparément pour τ_v et pour g.

D'une part,

$$B_{\tau_v(X)} = \frac{1}{|\tau_v(X)|}\sum_{y\in\tau_v(X)} y = \frac{1}{|X|}\sum_{x\in X}(x+v) = B_x + \frac{1}{|X|}\sum_{x\in X} v = B_x + v = \tau_v(B_X). \tag{18.279}$$

Nous avons utilisé le fait que X et $\tau_v(X)$ possèdent le même nombre d'éléments, ainsi que le fait d'avoir une somme de $|X|$ termes tous égaux à v.

D'autre part,

$$B_{g(X)} = \frac{1}{|X|}\sum_{x\in X} g(x) = g\Big(\frac{1}{|X|}\sum_{x\in X} x\Big) = g(B_X) \tag{18.280}$$

où nous avons utilisé la linéarité de g dans tous ses retranchements. □

PROPooLAEBooWdcBoe

Proposition 18.98.
Points fixes d'un sous-groupe.

(1) Soit H un sous-groupe fini des isométries de $(\mathbb{R}^n, d)$. Alors il existe $v \in \mathbb{R}^n$ tel que $f(v) = v$ pour tout $f \in H$.

(2) Si H est un sous-groupe de $\mathrm{Isom}(\mathbb{R}^n, d)$ n'acceptant pas de point fixe, alors il est infini.

Démonstration. Le groupe H agit sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$ nous pouvons considérer son orbite

$$Hx = \{f(x) \text{ tel que } f \in H\}, \tag{18.281}$$

qui est une partie finie de $\mathbb{R}^n$. Considérons son barycentre $v = B_{Hx}$. Soit $f \in H$. Alors

$$\begin{aligned} f(v) &= f(B_{Hx}) && \text{(18.282a)}\\ &= B_{f(Hx)} && \text{(18.282b)}\\ &= B_{Hx} && \text{(18.282c)}\\ &= v, && \text{(18.282d)} \end{aligned}$$

SUBEQooOQBZooYIlbgN

SUBEQooXWEGooSoczYg

Justifications :

— Pour (18.282b), c'est le lemme 18.97.

27. Définition 18.96.

— Pour (18.282c), c'est le fait que, $f \in H$ étant donné, l'application $g \mapsto fg$ est une bijection de H, donc

$$f(Hx) = \{(fg)(x) \text{ tel que } h \in H\} = \{g(x) \text{ tel que } g \in H\} = Hx. \tag{18.283}$$

Bref, v est fixé par H.

La seconde affirmation n'est rien d'autre que la contraposée de la première. □

PROPooEUFIooDUIYzi

Proposition 18.99.
À propos de groupes finis d'isométries.

(1) Tout sous-groupe fini de $\mathrm{Isom}(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{O}(n)$.

(2) Tout sous-groupe fini de $\mathrm{Isom}^+(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{SO}(n)$.

Démonstration. Soit H un sous-groupe fini de $\mathrm{Isom}(\mathbb{R}^n)$ et $v \in \mathbb{R}^n$ un élément fixé par H (comme garanti par la proposition 18.98). Nous posons

$$\begin{aligned} \phi\colon H &\to \mathrm{Isom}(\mathbb{R}^n) \\ f &\mapsto \tau_v^{-1} \circ f \circ \tau_v. \end{aligned} \tag{18.284}$$

(1) **ϕ est un morphisme** Les opération du type $\phi = \mathbf{Ad}(\tau_v)$ sont toujours des morphismes.

(2) **ϕ consiste à extraire la partie linéaire** Si $f = \tau_w \circ g$ alors

$$\begin{aligned} \phi(f)(x) &= (\tau_{-v} \circ \tau_w \circ g \circ \tau_v)(x) && \text{(18.285a)} \\ &= \tau_{w-v}(g(x) + g(v)) && \text{(18.285b)} \\ &= g(x) + g(v) - v + w && \text{(18.285c)} \end{aligned}$$

Mais $g(v) + w = f(v)$ et nous savons que $f(v) = v$. Donc il ne reste que $\phi(f)(x) = g(x)$.

(3) **ϕ est injective** Si $f = \tau_w \circ g$ vérifie $\phi(f) = \mathrm{Id}$, il faut en particulier que $g = \mathrm{Id}$. Mais H est fini et ne peut donc pas contenir de translations non triviales. Donc $w = 0$ et $f = \mathrm{Id}$.

Donc ϕ est une injection à valeur dans les transformations linéaires de $\mathrm{Isom}(\mathbb{R}^n)$. Autrement dit, ϕ est un isomorphisme entre H et son image, laquelle image est dans $\mathrm{O}(n)$.

En ce qui concerne la seconde partie, si $f \in \mathrm{Isom}^+(\mathbb{R}^n)$, alors $\phi(f)$, qui est la partie linéaire de f (théorème 18.79(5)), est dans $\mathrm{O}(n)$ avec un déterminant égal à 1. Autrement dit, $\phi(f) \in \mathrm{SO}(n)$. □

L'extraction de la partie linéaire est injective ? Certe c'est prouvé, mais on peut se demander ce qu'il se passe si H contient deux éléments qui ont la même partie linéaire. Cela n'est pas possible parce si $f_1 = \tau_{w_1} \circ g$ et $f_2 = \tau_{w_2} \circ g$ sont dans H alors $f_1 f_2^{-1} = \tau_{w_1 - w_2}$ est également dans H. Vu que H est fini, cela n'est possible que si $w_1 = w_2$, c'est-à-dire si $f_1 = f_2$.

18.9.1 Points fixés par une affinité

LEMooGUEGooTUXRsQ

Lemme 18.100 ([547]).
Si $n \geqslant 3$, alors toute droite est intersection de deux plans non isotropes[28].

Démonstration. Soit une droite d dans E, ainsi que $x \in d$. Soit x est isotrope, soit il ne l'est pas.

(1) **Si x est isotrope**

Étant donné que f est non dégénérée, il existe $y \in E$ tel que $f(x, y) \neq 0$. Le plan $P = \mathrm{Span}\{x, y\}$ est alors non isotrope. En effet, considérons $z \in P \cap P^\perp$. Nous avons alors $f(z, x) = f(z, x) = 0$. Vu que $z \in P$ nous pouvons l'écrire sous la forme $z = \alpha x + \beta y$. La première condition sur z donne :

$$0 = f(\alpha x + \beta y, x) = \beta f(y, x) \tag{18.286}$$

28. Sous-espace isotrope, définition 8.143.

parce que $f(x,x)=0$. Étant donné que $f(x,y)\neq 0$ nous en déduisons que $\beta=0$. La seconde condition donne alors $f(\alpha x,y)=0$, ce qui donne $\alpha=0$.
Le plan P étant non isotrope, le plan $P^\perp$ est également non isotrope par le lemme 8.145.
Nous fixons maintenant $z\in P^\perp$.
L'élément z n'est pas dans P parce que $P\cap P^\perp=\{0\}$. L'élément $y+z$ n'est pas dans P non plus parce que P est vectoriel, $y\in P$ et $z\notin P$.
Nous considérons donc le plan $Q=\text{Span}\{x,y+z\}$.
Ce plan est non isotrope. En effet, soit $u\in Q\cap Q^\perp$. Nous avons $u=\alpha x+\beta(y+z)$, et cela doit vérifier $f(u,x)=f(y,y+z)=0$. Nous avons d'abord

$$0=f(u,x)=\alpha\underbrace{f(x,x)}_{=0}+\beta\underbrace{f(x,y)}_{\neq 0}+\beta\underbrace{f(x,z)}_{=0}. \tag{18.287}$$

Nous en déduisons que $\beta=0$. Il reste $u=\alpha x$, et nous imposons la seconde condition :

$$0=f(u,y+z)=\alpha f(x,y)+\alpha f(x,z). \tag{18.288}$$

Vu que $f(x,y)\neq 0$ et $f(x,z)=0$ nous déduisons $\alpha=0$ et donc $u=0$.
Au final, les plans P et Q sont non isotropes et $P\cap Q=d$.

(2) **Si x n'est pas isotrope** Nous supposons maintenant que x n'est pas isotrope. Le lemme 8.136 sur les dimensions, appliqué à l'espace de dimension 1 engendré par x donne $\dim(x^\perp)+1=\dim(E)$. Vu que $\dim(E)\geqslant 3$, nous déduisons que $\dim(x^\perp)\geqslant 2$.
Nous considérons une base f-orthogonale de $x^\perp$ (proposition 8.243). Les vecteurs de cette base ne sont pas isotropes. Soient y et z les vecteurs de cette base. Il nous reste à prouver que les plans $P=\text{Span}\{x,y\}$ et $Q=\text{Span}\{x,z\}$ font l'affaire.
Vu que x n'est pas isotrope, les vecteurs y et z ne sont pas dans d ; ce sont donc de vrais plans de dimension 2, et leur intersection est d. Montrons que P n'est pas isotrope, en considérant $v=\alpha x+\beta y\in P\cap P^\perp$. Vu que $v\in P^\perp$, il vérifie les deux conditions $f(v,x)=f(v,y)=0$.
La première condition donne

$$0=\alpha\underbrace{f(x,x)}_{\neq 0}+\beta\underbrace{f(y,x)}_{=0}=\alpha f(x,x) \tag{18.289}$$

parce que $y\in x^\perp$. Il nous reste $v=\beta y$. Appliquons la seconde condition :

$$0=f(v,y)=\alpha f(y,y). \tag{18.290}$$

Le scalaire $f(y,y)\in\mathbb{K}$ n'est pas nul parce que nous avons choisi une base orthogonale. Au final $v=0$, et P est non isotrope. Exactement le même raisonnement montre que Q n'est pas isotrope non plus.

□
PROPooVEEUooJQmmkN

Proposition 18.101 ([545])**.**
Si une isométrie de $\mathbb{R}^n$ fixe un ensemble F de points, alors elle fixe l'espace affine engendrée par F.

Démonstration. Soit $f\in\text{Isom}(\mathbb{R}^n)$ fixant F. Par le théorème 8.157, c'est une application affine et l'ensemble $\text{Fix}(f)$ des points fixés par f est un sous-espace affine de $\mathbb{R}^n$, grâce à la proposition 11.61.

Donc $\text{Fix}(f)$ est un espace affine contenant F. Puisque l'espace affine engendré par F est l'intersection de tous les espaces affines contenant F, il est en particulier contenu dans $\text{Fix}(f)$. □
CORooZHZZooDgTzsW

Corolaire 18.102.
Si f et g sont des isométries de $\mathbb{R}^n$ qui coïncident sur F, alors elles coïncident sur l'espace affine engendré par F.

Démonstration. Nous considérons $h=g^{-1}\circ f$ qui est une isométrie de $\mathbb{R}^n$ fixant F. Elle fixe donc, par la proposition 18.101, l'espace affine engendré par F. Or tout point fixé par h est un point sur lequel g et f coïncident. □

18.10 Classification des isométries dans $\mathbb{R}^2$

18.10.1 Projection orthogonale

Une version pour les espaces de Hilbert sera le théorème 25.7

DEFooHXJTooNPyDFQ

Proposition-Définition 18.103.
Soient un espace euclidien de dimension finie $(E, \cdot)$, un sous-espace vectoriel H_0 ainsi que $a, d \in E$. Nous considérons le sous-espace affine $H = H_0 + d$.

Il existe un unique $y \in H$ tel que

$$\|a - y\| = \inf\{\|a - x\| \text{ tel que } x \in H\}. \tag{18.291}$$

*Ce y est la **projection orthogonale** de a sur H et est notée $\texttt{proj}_H(a)$.*

LEMooXZMFooOPTjNx

Lemme 18.104.
Soient un espace vectoriel euclidien $(E, \cdot)$ ainsi que $a \in E$. Nous considérons un sous-espace affine H de E. Si $h \in H$, alors

$$a \cdot h = \texttt{proj}_H(a) \cdot h. \tag{18.292}$$

18.10.2 Réflexions

Soit un espace vectoriel E de dimension 2 muni d'un produit scalaire[29]. Cela pourrait très bien être $\mathbb{R}^2$, mais nous allons nous efforcer de l'appeler E pour rester un peu général.

DEFooLJKDooUaamen

Lemme-Définition 18.105 (Caractérisation des réflexions)**.**
Soit une droite ℓ de $\mathbb{R}^2$. Il existe une unique application $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que

(1) $f(x) = x$ pour tout $x \in \ell$.

(2) f échange les côtés de ℓ.

(3) f laisse invariants les droites perpendiculaires à ℓ et les cercles dont le centre est sur ℓ.

*Cette application est la **réflexion** d'axe ℓ.*

Démonstration. Soit x hors de ℓ et p la droite perpendiculaire à ℓ et passant par x. Nous avons $f(x) \in p$. En nommant P l'intersection entre ℓ et p, nous considérons le cercle $S(P, \|Px\|)$ qui est un cercle dont le centre est sur ℓ. Il contient x et donc $f(x) \in S(P, \|Px\|)$.

Donc $f(x) \in p \cap S(P, \|Px\|)$. L'intersection entre un cercle et une droite contient de façon générique deux points. L'un est x, mais $f(x) = x$ n'est pas possible parce que x est hors de ℓ et f doit inverser les côtés de ℓ. Donc $f(x)$ est l'autre.

Cela prouve l'unicité. En ce qui concerne l'existence, il suffit de noter que la réflexion σ_ℓ satisfait les contraintes. □

LEMooZSDRooUkNYer

Lemme 18.106.
Soit une droite ℓ et $A \in E$. Alors

$$\sigma_\ell(A) = 2\,\texttt{proj}_\ell(A) - A \tag{18.293}$$

EQooVUQDooKuwszl

où $\texttt{proj}_\ell$ est l'opération de projection orthogonale sur la droite ℓ.

Démonstration. Nous posons $f(X) = 2\,\texttt{proj}_\ell(X) - X$ et nous allons montrer que $f = \sigma_\ell$ en vérifiant les conditions de la définition 18.105. Nous nous gardons bien de faire un raisonnement du type « nous allons montrer que f et σ_ℓ coïncident sur deux points, et sont donc égales par le corolaire 18.102 » parce que nous ne savons pas encore que σ_ℓ est une application affine, ni même que c'est une isométrie.

Si $X \in \ell$ alors $\texttt{proj}_\ell(X) = X$ et nous avons $f(X) = 2X - X = X$. Donc ℓ est conservée.

29. Définition 8.168.

En ce qui concerne les deux côtés de ℓ, il existe une application linéaire $s\colon E\to\mathbb{R}$ et une constante $c\in\mathbb{R}$ telles qu'en posant $\ell(X)=s(X)+c$, la droite ℓ soit le lieux des points X tels que $\ell(X)=0$. Un côté de la droite est $\ell<0$ et l'autre côté est $\ell>0$. Nous avons :

$$\begin{aligned}\ell\big(f(A)\big)&=\ell\big(2\,\texttt{proj}_\ell(A)-A\big)&(18.294a)\\&=s(2\,\texttt{proj}_\ell(A)-A)+c&(18.294b)\\&=2s\big(\texttt{proj}_\ell(A)\big)-s(A)+c&(18.294c)\\&=s\big(\texttt{proj}_\ell(A)\big)-s(A)&(18.294d)\\&=-c-s(A)&(18.294e)\\&=-\ell(A)&(18.294f)\end{aligned}$$

où nous avons utilisé le fait que, $\texttt{proj}_\ell(A)$ étant sur ℓ, $s\big(\texttt{proj}_\ell(A)\big)+c=0$. Nous avons donc $\ell\big(f(A)\big)=-\ell(A)$, ce qui indique que A et $f(A)$ sont de part et d'autre de ℓ.

Si d est une droite perpendiculaire à ℓ et si $A\in d$ alors $f(A)=2\,\texttt{proj}_\ell(A)-A=\big(\texttt{proj}_\ell(A)-A\big)+A\in d$ parce que $\texttt{proj}_\ell(A)\in d$ du fait que d soit précisément perpendiculaire à ℓ. Nous avons aussi utilisé le fait que si $A,B,C\in d$ alors $(B-A)+C\in d$; pensez que $B-A$ est un vecteur directeur et que C est un point de d.

Enfin soit $K\in\ell$ et un cercle $S(K,r)$ centré en K. Soit $A\in S(K,r)$; nous devons vérifier que $f(A)=S(K,r)$. Le segment $[A,f(A)]$ est par définition perpendiculaire à ℓ. Soit M, le milieu, qui est sur la droite ℓ. Les triangles AMK et $f(A)MK$ sont rectangles en M, et $\|AM\|=\|Mf(A)\|$. Le théorème de Pythagore donne $\|AK\|=\|f(A)K\|$. Donc le cercle centré en K est alors préservé par f.

Nous en déduisons que $f=\sigma_\ell$. □

LEMooNYIWooXanBXh

Lemme 18.107.
Soit une droite ℓ dans $\mathbb{R}^2$ ainsi que $a\in\mathbb{R}^2$. Nous avons

$$\texttt{proj}_\ell\big(\sigma_\ell(a)\big)=\texttt{proj}_\ell(a).\tag{18.295}$$

PROPooFSVEooWmJsnv

Proposition 18.108 ([548])**.**
Une réflexion est une isométrie de (E,d) où $d(A,B)=\|A-B\|$.

Démonstration. Soient $A,B\in E$; il faut vérifier que $\|A-B\|=\|\sigma_\ell(A)-\sigma_\ell(B)\|$. Pour cela nous écrivons

$$B-A=\underbrace{B-\texttt{proj}_\ell(B)}_{=a}+\underbrace{\texttt{proj}_\ell(B)-\texttt{proj}_\ell(A)}_{=b}+\underbrace{\texttt{proj}_\ell(A)-A}_{=c}.\tag{18.296}$$

Vu que $b\perp a$ et $b\perp c$ (lemme 18.104) nous avons

$$\|B-A\|^2=\langle B-A,B-A\rangle=\|a\|^2+2\langle a,c\rangle+\|b\|^2+\|c\|^2.\tag{18.297}$$

Nous pouvons jouer le même jeu avec $\sigma_\ell(B)-\sigma_\ell(A)$ en tenant compte du lemme 18.107 et que

$$\sigma_\ell(A)-\texttt{proj}_\ell(A)=2\,\texttt{proj}_\ell(A)-A-\texttt{proj}_\ell(A)=-\big(A-\texttt{proj}_\ell(A)\big).\tag{18.298}$$

Là nous avons utilisé le lemme 18.106. Ce que nous trouvons est que

$$\sigma_\ell(B)-\sigma_\ell(A)=-a+b-c,\tag{18.299}$$

et donc encore une fois

$$\|\sigma_\ell(B)-\sigma_\ell(A)\|^2=\|a\|^2+2\langle a,c\rangle+\|b\|^2+\|c\|^2.\tag{18.300}$$

□

Remarque 18.109.
Il faut bien comprendre que si l'axe de la réflexion ne passe pas par 0 (le zéro de l'espace vectoriel normé $(E,\|.\|)$), la réflexion n'est pas une isométrie de $(E,\|.\|)$ au sens où nous n'avons pas $\|\sigma_\ell(x)\|=\|x\|$.

18.10.3 Segment, plan médiateur et équidistance

LEMooSZZWooPDHnGl

Lemme 18.110.
Un point M est sur la médiatrice du segment $[A, B]$ si et seulement si $\|M - A\| = \|M - B\|$.

Démonstration. Nous posons $I = (A + B)/2$. D'une part l'expression que M est sur la médiatrice est que

$$(M - I) \cdot (A - B) = 0. \tag{18.301}$$

D'autre part, l'expresson d'être équidistant de A est de B est

$$\|M - A\|^2 - \|M - B\|^2 = 0. \tag{18.302}$$

La preuve consiste à montrer que

$$2(M - I) \cdot (B - A) = \|M - A\|^2 - \|M - B\|^2. \tag{18.303}$$

Pour cela, il suffit d'injecter $I = (A + B)/2$ et d'effectuer les produits.

Nous avons par exemple

$$\begin{aligned} 2(M - I) \cdot (B - A) &= (2M - A - B) \cdot (B - A) && \text{(18.304a)}\\ &= 2M \cdot B - 2M \cdot A - A \cdot B + \|A\|^2 - \|B\|^2 + A \cdot B && \text{(18.304b)}\\ &= 2M \cdot (B - A) + \|A\|^2 - \|B\|^2. && \text{(18.304c)} \end{aligned}$$

□

LEMooVBVUooOTFFXT

Lemme 18.111.
Soient A et B deux points de $\mathbb{R}^3$. Alors le plan médiateur du segment $[A, B]$ est le lieu des points de $\mathbb{R}^3$ qui sont équidistants de A et B.

Démonstration. Nous nommons σ ce plan.

Soit X un point équidistant de A et B. Alors dans le plan (A, B, X), le triangle ABX est isocèle en X, et la hauteur issue de X coupe perpendiculairement $[A, B]$ en son milieu. Cela prouve que X est dans le plan médiateur du segment $[A, B]$ (lemme 18.110).

Réciproquement, supposons que X soit dans le plan médiateur de $[A, B]$. Soit M le point du segment $[A, B]$ équidistant de A et B. Nous avons $(X, M) \perp (A, B)$. Donc le triangle A, B, X est isocèle en X et donc X est équidistant de A et B. □

LEMooTCIEooXdyuHu

Lemme 18.112.
Si A' est l'image de A par σ_ℓ alors ℓ est la médiatrice du segment $[A, A']$.

Démonstration. Soit $M \in \ell$. Nous avons

$$\|A - M\|^2 = \|\mathtt{proj}_\ell(A) - A\|^2 + \|\mathtt{proj}_\ell(A) - M\|^2 \tag{18.305}$$

parce que $A - \mathtt{proj}_\ell(A) \perp M - \mathtt{proj}_\ell(A)$. Par ailleurs, vu que $\sigma_\ell(A) = 2\,\mathtt{proj}_\ell(A) - A$ et que $\mathtt{proj}_\ell(A) = \mathtt{proj}_\ell(A')$,

$$\|\mathtt{proj}_\ell(A) - A\| = \|\mathtt{proj}_\ell(A') - A'\|. \tag{18.306}$$

Nous avons donc

$$\|\sigma_\ell(A) - M\|^2 = \|A - M\|^2, \tag{18.307}$$

ce qui prouve que M est sur la médiatrice de $[A', A]$ par le lemme 18.110. □

18.113.
Si l est une droite dans $\mathbb{R}^2$, nous avons la réflexion $\sigma_l \in \mathrm{Isom}(\mathbb{R}^2)$ d'axe l. Cela est une isométrie et donc une application affine par le théorème 8.157. Le lemme suivant détermine comment la réflexion σ_ℓ se décompose en une translation et une application linéaire.

LEMooVOJLooCFgdNG

Lemme 18.114.
Soit une droite ℓ. Alors

$$\sigma_\ell = \tau_{2w} \circ \sigma_{\ell_0} \tag{18.308}$$

où ℓ_0 est la droite parallèle à ℓ passant par l'origine, et w est le vecteur perpendiculaire à ℓ tel que $\ell_0 = \ell + w$.

Démonstration. Il faut trouver trois points non alignés sur lesquels les deux applications coïncident ; cela suffira par le corolaire 18.102.

Pour tous les points de ℓ_0, l'égalité fonctionne parce que si $x \in \ell_0$,

$$\sigma_\ell(x) = x + 2w, \tag{18.309}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x + 2w \tag{18.310}$$

du fait que $\sigma_{\ell_0}(x) = x$.

Si $x \in \ell$, alors

$$\sigma_\ell(x) = x \tag{18.311}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x - 2w + 2w = x. \tag{18.312}$$

Donc les applications affines σ_ℓ et $x \mapsto \sigma_{\ell_0}(x) + 2w$ coïncident sur ℓ et ℓ_0. Elles coïncident donc partout. □

LEMooRVKEooSlIrIR

Lemme 18.115 ([545])**.**
Une isométrie de $(\mathbb{R}^2, d)$ fixant tous les points d'une droite est soit l'identité soit la réflexion de cette droite.

Démonstration. Soit l une droite fixée par f, et soient $x, y \in l$ et $z \notin l$ (avec $x \neq y$). Le fait que x et y soient des points fixes de f implique

$$\begin{cases} d\big(x, f(z)\big) = d(x, z) & \text{(18.313a)} \\ d\big(y, f(z)\big) = d(y, z) & \text{(18.313b)} \end{cases}$$

ce qui signifie que $f(z)$ est sur l'intersection des deux cercles[30] $S\big(x, d(x,z)\big)$ et $S\big(y, d(y,z)\big)$, et comme ce sont deux cercles centrés sur la droite l, les intersections sont liées par σ_l. Autrement dit, les intersections sont z et $\sigma_l(z)$.

Si $f(z) = z$ alors f fixe trois points non alignés et fixe donc $\mathbb{R}^2$, c'est-à-dire $f = \mathrm{Id}$.

Si par contre $f(z) = \sigma_l(z)$ alors les isométries f et σ_l coïncident sur trois points et coïncident donc partout par le corolaire 18.102 : $f = \sigma_l$. □

18.10.4 Translations et réflexions

LEMooMKTXooYKZcdQ

Lemme 18.116.
Si $A \in \mathbb{R}^2$, si ℓ est une droite de $\mathbb{R}^2$, alors nous avons

$$\texttt{proj}_{\tau_A(\ell)} = \tau_A \circ \texttt{proj}_\ell \circ \tau_A^{-1}. \tag{18.314}$$

Démonstration. Soit $x \in \mathbb{R}^2$. Soit v unitaire dans la direction[31] de ℓ. La condition $q = \texttt{proj}_\ell(x)$ est le système

$$\begin{cases} q \in \ell & \text{(18.315a)} \\ (q - x) \cdot v = 0. & \text{(18.315b)} \end{cases}$$

30. L'intersection existe parce que $d(x,z) + d(y,z) > d(x,y)$.
31. Cela signifie qu'il existe $a \in \mathbb{R}^2$ tel que $\ell = a + \mathbb{R}v$.

Nous voulons prouver que $\tau_A(q) = \mathtt{proj}_{\tau_A(\ell)}\big(\tau_A(x)\big)$, c'est-à-dire que

$$\begin{cases} \tau_A(q) \in \tau_A(\ell) & (18.316a)\\ \big(\tau_A(q) - \tau_A(x)\big) \cdot v = 0. & (18.316b) \end{cases}$$

Nous avons utilisé le fait que v est un vecteur unitaire dans la direction de $\tau_A(\ell)$ aussi bien que de ℓ.

Puisque $q \in \ell$, il est bien entendu que $\tau_A(q) \in \tau_A(\ell)$. D'autre part, $\tau_A(q) - \tau_A(x) = q - x$, donc

$$\big(\tau_A(q) - \tau_A(x)\big) \cdot v = (q - x) \cdot v = 0. \tag{18.317}$$

□

LEMooSMMMooAqsHWb

Lemme 18.117.
Si ℓ est une droite de $\mathbb{R}^2$, si $A \in \mathbb{R}^2$, alors

$$\sigma_{\tau_A(\ell)} = \tau_A \sigma_\ell \tau_A^{-1}. \tag{18.318}$$

Démonstration. Il s'agit d'un calcul mettant en scène les lemmes 18.106 et 18.116 :

$$\begin{aligned} \big(\sigma_{\tau_A(\ell)}\tau_A\big)(x) &= 2\,\mathtt{proj}_{\tau_A(\ell)}\big(\tau_A(x)\big) & (18.319a)\\ &= 2\tau_A\big(\mathtt{proj}_\ell(x)\big) - \tau_A(x) & (18.319b)\\ &= 2\,\mathtt{proj}_\ell(x) + 2A - x - A & (18.319c)\\ &= 2\,\mathtt{proj}_\ell(x) - x + A & (18.319d)\\ &= \tau_A\big(\sigma_\ell(x)\big). & (18.319e) \end{aligned}$$

□

18.10.5 Rotations

DEFooFUBYooHGXphm

Définition 18.118 (Rotation en dimension 2)**.**
*Une **rotation** d'un espace euclidien de dimension 2 est une composée de deux réflexions d'axes non parallèles. L'identité est une rotation.*

18.119.
Quelques remarques à propos de cette définition.

(1) Attention : nous ne parlons pas encore de rotations « vectorielles » : ici le centre de la rotation (que nous n'avons pas encore défini) peut ne pas être 0.

(2) Dans la même veine : plus tard, lorsque nous saurons que les rotations sont des isométries de (E, d) où $d(X, Y) = \|X - Y\|$, nous allons en réalité beaucoup plus souvent parler de rotations centrées en l'origine qu'en un point quelconque. C'est pourquoi à partir de 18.130 nous dirons le plus souvent « rotation » pour « rotation centrée en 0 ». D'où les énoncés comme « les rotations sont les matrice orthogonales » (corolaire 18.138), qui *stricto sensu* de la définition 18.118, sont faux.

(3) Une rotation est composée de deux réflexions d'axes non parallèles. Il est cependant trop tôt pour décréter que l'intersection de ces axes est le centre de la rotation. Rien ne dit en effet pour l'instant que deux décompositions différentes de la même rotation, avec des axes différents donnent le même point d'intersection.

(4) Pourquoi ajouter l'identité ? Pour avoir un groupe. Dans le cas vectoriel, il est suffisant de demander d'être une composée de deux réflexions, parce que toutes les réflexions vectorielles ont des axes qui s'intersectent en 0. Le cas des axes parallèles est seulement le cas des axes confondus et revient à l'identité.

Si nous voulons avoir un groupe même pour les rotations centrées ailleurs qu'en zéro, nous devons ajouter « à la main » l'identité.

Toutes ces remarques se résument par : « tout devient compliqué du fait que nous voulons considérer également les rotations centrées ailleurs qu'en zéro ». En se contentant du cas vectoriel, de nombreuses choses sont plus simples.

CORooNKKIooPGOUJl

Corolaire 18.120.
Si $A \neq B$ dans E alors il existe une unique réflexion envoyant A sur B.

Démonstration. En ce qui concerne l'existence, la réflexion dont l'axe est la médiatrice de $[A, B]$ fait l'affaire. En ce qui concerne l'unicité, le lemme 18.112 nous dit que si A est envoyé sur B, l'axe est forcément la médiatrice de $[A, B]$. □

LEMooIJELooLWqBfE

Lemme-Définition 18.121 ([549])**.**
Soit une rotation $r = \sigma_1 \circ \sigma_2$ différente de l'identité.

(1) Elle admet un unique point fixe.

(2) Ce point fixe est l'intersection des axes $\ell_1 \cap \ell_2$.

*Le **centre** d'une rotation (autre que l'identité) est cet unique point fixe.*

Démonstration. Nous nommons $O = \ell_1 \cap \ell_2$. Soit $A \in E$, et supposons que $r(A) = A$. Nous avons $\sigma_1 \circ r = \sigma_2$ et donc

$$\sigma_1(A) = (\sigma_1 \circ r)(A) = \sigma_2(A). \tag{18.320}$$

On pose $B = \sigma_1(A)$. Alors σ_1 et σ_2 envoient tous deux A sur B.

Si $A = B$ alors A est fixé par σ_1 et donc appartient à ℓ_1. Même chose pour A qui est fixé par σ_2, et donc $A \in \ell_2$. Cela donne $A = B = O$, et alors, le point fixé par r est O.

Si $A \neq B$ alors il existe une unique réflexion envoyant A sur B (corolaire 18.120). L'unicité signifie que $\sigma_1 = \sigma_2$. Dans ce cas, $r = \sigma_1 \circ \sigma_2 = \mathrm{Id}$. □

NORMooDPBOooKkRuTn

18.122.
La rotation $\sigma_1 \circ \sigma_2$ laisse évidemment fixé le point $\ell_1 \cap \ell_2$. Si $\sigma_1 \circ \sigma_2 = \sigma_a \circ \sigma_b$ alors rien n'oblige les axes de σ_1 et σ_2 d'être identiques à ceux de σ_a et σ_b. Mais l'intersection $\ell_1 \cap \ell_2$ doit être la même que l'intersection $\ell_a \cap \ell_b$ parce que c'est l'unique point fixé par la composée. Cela nous permet de poser la définition suivante.

LEMooTZNWooTVOklu

Lemme 18.123.
Les rotations sont des isométries pour la distance : $\|X - Y\| = \|r(X) - r(Y)\|$.

Démonstration. Si $r = \sigma_1 \circ \sigma_2$, en utilisant le fait que σ_1 et σ_2 sont des isométries de (E, d) (18.108) nous avons :

$$d(X, Y) = d\big(\sigma_2(X), \sigma_2(Y)\big) = d\big(\sigma_1\sigma_2(X), \sigma_1\sigma_2(Y)\big) = d\big(r(X), r(Y)\big). \tag{18.321}$$

□

Ce lemme nous dit qu'une rotation de centre O vérifie $\|OX\| = \|Or(X)\|$ pour tout X.

PROPooNXJKooEDOczh

Proposition 18.124 ([1, 549])**.**
Soient $A, B, O \in E$ tels que $\|AO\| = \|BO\| \neq 0$. Alors il existe une unique rotation r centrée en O telle que $r(A) = B$.

Démonstration. Existence et unicité séparément.

(1) **Existence** Si $A = B$, l'identité fait l'affaire. Sinon, $\|A - O\| = \|B - O\|$ implique que la médiatrice de $[A, B]$ contient O. Soit σ_m la réflexion selon cette médiatrice. La rotation $\sigma_m \circ \sigma_{(AO)}$ convient.

(2) **Unicité** Soit r une rotation de centre O et telle que $r(A) = B$. Si $A = B$ alors $r = \mathrm{Id}$ parce qu'une rotation autre que l'identité ne fixe que son centre par le lemme 18.121. Nous supposons que $A \neq B$.

Nous posons $g = \sigma_m \circ r$. Alors $g(A) = \sigma_m(B) = A$ parce que $\sigma_m(B) = A$ et $r(A) = B$. Cela signifie que g est une isométrie qui fixe A.

(2a) **Si A et O ne sont pas alignés** Attention : ici O est un point de E, pas le zéro de l'espace vectoriel E. Lorsqu'on dit que A et O ne sont pas alignés, nous parlons bien d'alignement avec le zéro de E.

Nous avons $g(A) = A$ et $g(O) = O$. Donc g coïncide avec $\sigma_{(AO)}$ en deux points non alignés, c'est-à-dire en deux points pour lesquels l'espace engendré est tout E. Nous en déduisons que $g = \sigma_{(AO)}$.

(2b) **Si A et O sont alignés** Soit maintenant un point C tel que $A - O \perp C - O$ et

$$\|OC\| = \|OA\| = \|OB\|. \tag{18.322}$$

Puisque g est une isométrie pour la distance sur E, pas pour la norme, nous ne pouvons pas écrire $g(C - O) \perp g(A - O)$ à partir de $C - O \perp A - O$. Nous décomposons $g(X) = s(X) + G$ où s est linéaire sur E. Il est vite vu que s est une isométrie de $(E, \|.\|)$:

$$\|X - Y\| = \|g(X) - g(Y)\| = \|s(X) + G - s(Y) - G\| = \|s(X) - s(Y)\| = \|s(X - Y)\| \tag{18.323}$$

pour tout $X, Y \in E$. Nous avons de plus $g(A) = A$ et $g(O) = O$, ce qui donne $O = s(O) + G$ et $A = s(A) + G$. En égalisant les valeurs de G nous avons

$$O - s(O) = A - s(A). \tag{18.324}$$

EQooPEWGooABHUvu

Comme s est une isométrie (une vraie) nous avons

$$s(A - O) \perp s(C - O), \tag{18.325}$$

mais $s(A - O) = s(A) - s(O) = A - O$ par (18.324). Donc

$$A - O \perp s(C - O). \tag{18.326}$$

Nous en concluons que $s(C - O) = \pm(C - O)$. Parce que les vecteurs $\pm(C - O)$ sont les deux seuls de norme $\|AO\| = \|CO\|$ à être perpendiculaire à $A - O$. Rappel : la définition de C, et le fait que nous soyons en dimension 2.

Est-il possible d'avoir $s(C - O) = C - O$? Cela donnerait

$$g(A - O) = s(A) - s(O) + G = s(A) - O + O - s(O) + G = A - O + G \tag{18.327a}$$
$$g(C - O) = C - O + G, \tag{18.327b}$$

ce qui signifierait que g et τ_G coïncideraient sur les points $A - O$ et $C - O$, et donc seraient égaux par le corolaire 18.102. Cela est cependant impossible parce que g fixe au moins les points A et O alors que la translation ne fixe aucun point. Nous en déduisons $s(C - O) = -(C - O)$.

Nous avons aussi, parce que (AO) est une droite passant par l'origine que

$$\sigma_{(AO)}(A - O) = A - O \tag{18.328}$$

et parce que $C - O$ est perpendiculaire à cette droite :

$$\sigma_{(AO)}(C - O) = -(C - O). \tag{18.329}$$

Nous avons donc quand même que g et $\sigma_{(AO)}$ coïncident sur deux points non alignés : $A - O$ et $C - O$.

Dans tous le cas, $g = \sigma_{(AO)}$. Nous avons donc

$$\sigma_{(OA)} = \sigma_m \circ r, \tag{18.330}$$

et donc r est fixé à

$$r = \sigma_m \circ \sigma_{(OA)}. \tag{18.331}$$

□

18.125.
Anticipons un peu et faisons semblant de déjà connaitre les matrices et les fonctions trigonométriques. La proposition 18.124 nous dit qu'il existe une seule rotation amenant A sur B. Vous pourriez objecter que le point $(1,0)$ peut être amené sur $(0,-1)$ soit par la rotation d'angle $3\pi/2$, soit par celle d'angle $-\pi/2$. Il n'en est rien parce que ces deux rotations sont les mêmes ! Pensez-y. En tant qu'application $\mathbb{R}^2 \to \mathbb{R}^2$, la rotation $R_{3\pi/2}$ est égale à $R_{-\pi/2}$.

Une rotation donnée peut être écrite de beaucoup de façons comme composée de deux réflexions. En fait d'autant de façons qu'il y a de réflexions.

PROPooKAZEooLTHWKe

Proposition 18.126 ([549])**.**
Soit une rotation r de E centrée en O. Pour toute réflexion σ_ℓ telle que le centre de r soit sur ℓ, il existe une réflexion σ_1 telle que $r = \sigma_1 \circ \sigma_\ell$. Il existe aussi une réflexion σ_2 telle que $r = \sigma_\ell \circ \sigma_2$.

Démonstration. Si $r = \mathrm{Id}$ c'est bon avec $\sigma_1 = \sigma_2 = \sigma_\ell$. Sinon nous considérons $A \neq O$ sur ℓ, et $B = r(A)$. Nous savons que $B \neq A$ parce que O est le seul point de E fixé par r (proposition 18.121). Il existe une réflexion (unique) σ_1 faisant $\sigma_1(A) = B$, et c'est la réflexion dont l'axe est la médiatrice de $[A,B]$. Le point O est sur cette médiatrice parce que les rotations sont des isométries de (E,d) (lemme 18.123).

La rotation $\sigma_1 \circ \sigma_\ell$ vérifie

$$(\sigma_1 \circ \sigma_\ell)(A) = \sigma_1(A) = B. \tag{18.332}$$

Or $\|OA\| = \|OB\|$, donc il y a unicité de la rotation centrée en O portant A sur B (proposition 18.124) ; nous avons donc $r = \sigma_1 \circ \sigma_\ell$.

En ce qui concerne $r = \sigma_\ell \circ \sigma_2$, il faut appliquer ce que nous venons de faire à la rotation r^{-1} : il existe σ_2 tel que $r^{-1} = \sigma_2 \circ \sigma_\ell$, ce qui donne

$$r = \sigma_\ell \circ \sigma_2. \tag{18.333}$$

□

18.10.6 Rotation d'un angle donné

DEFooADTDooKIZbrw

Lemme-Définition 18.127.
Soit $\theta \in \mathbb{R}$. Nous considérons l'application linéaire $R_0(\theta)\colon \mathbb{R}^2 \to \mathbb{R}^2$ dont la matrice dans la base canonique est[32]

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.334}$$

ITEMooIEKJooZfsAui

(1) Nous avons

EQooEVCTooBpTDDq

$$R_0(\theta) = \sigma_\ell \circ s \tag{18.335}$$

où s est la réflexion d'axe horizontal et ℓ est la droite

$$\ell = \mathbb{R} \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{18.336}$$

ITEMooBEYOooMHRRYk

(2) L'application $R_0(\theta)$ est une rotation autour de $(0,0)$.

32. Pour la définition des fonctions trigonométriques, définition 18.1.

ITEMooEQPAooQcsYfj

(3) *Si $A \in \mathbb{R}^2$, alors l'application*

$$R_A(\theta) = \tau_A \circ R_0(\theta) \circ \tau_A^{-1} \tag{18.337}$$

EQooKMJQooWGxWHe

est une rotation autour de A nommée ***rotation d'angle*** *θ.*

Démonstration. Nous allons prouver l'égalité (18.335) en calculant les deux membres sur les vecteurs $\begin{pmatrix}1\\0\end{pmatrix}$ et $\begin{pmatrix}0\\1\end{pmatrix}$ qui forment une base.

(1) **Pour $p = (1,0)$** D'abord la chose facile[33] : $s(1,0) = (1,0)$.
Pour calculer $\sigma_\ell(1,0)$, nous utilisons le lemme 18.106 ; nous commençons par chercher la projection orthogonale q de $p = (1,0)$ sur ℓ. Nous posons

$$q = \lambda \begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)\end{pmatrix} \tag{18.338}$$

EQooUFMWooWhwcHR

et nous cherchons λ satisfaisant $q \cdot (q-p) = 0$. Un peu de calculs passant par (18.5) nous donne

$$q \cdot (q-p) = \lambda\big(\lambda - \cos(\theta/2)\big). \tag{18.339}$$

Les deux solutions sont $\lambda = 0$ et $\lambda = \cos(\theta/2)$. Mais la solution $\lambda = 0$ revient à dire que la droite ℓ est verticale, c'est-à-dire $\cos(\theta/2) = 0$. Donc la solution est toujours donnée par

$$\lambda = \cos(\theta/2). \tag{18.340}$$

Nous introduisons cette valeur dans (18.338) pour fixer q, et nous utilisons la formule du lemme 18.106 :

$$\sigma_\ell \begin{pmatrix}1\\0\end{pmatrix} = 2q - p = \begin{pmatrix}2\cos^2(\theta/2) - 1\\ 2\cos(\theta/2)\sin(\theta/2)\end{pmatrix} = \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}. \tag{18.341}$$

Nous avons utilisé les formules trigonométriques de duplication d'angle (corolaire 18.15).

(2) **Pour $p = (0,1)$** Cette fois $s(p) = (0,-1)$ et l'équation pour déterminer λ est

$$0 = q \cdot (q-p) = \begin{pmatrix}\lambda\cos(\theta/2)\\ \lambda\sin(\theta/2)\end{pmatrix} \cdot \begin{pmatrix}\lambda\cos(\theta/2)\\ \lambda\sin(\theta/2) - 1\end{pmatrix}. \tag{18.342}$$

Nous trouvons $\lambda\big(\lambda + \sin(\theta/2)\big) = 0$. Le cas $\lambda = 0$ signifie que la droite ℓ est horizontale et donc que $\sin(\theta/2) = 0$. Donc la solution est dans tous les cas

$$\lambda = -\sin(\theta/2). \tag{18.343}$$

Nous trouvons

$$\sigma_\ell \begin{pmatrix}0\\-1\end{pmatrix} = 2q - p = \begin{pmatrix}-\sin(\theta)\\ \cos(\theta)\end{pmatrix}. \tag{18.344}$$

Le calcul de $\sigma_\ell \circ s$ étant effectué sur une base, il est facile de reconstituer la matrice

$$\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}. \tag{18.345}$$

Cette matrice étant celle, par définition, de $R_0(\theta)$, nous avons montré que $R_0(\theta)$ était bien une rotation. Nous avons prouvé les points (1) et (2).

33. Ici comme partout dans le Frido nous ne faisons aucune différence entre (a,b) et $\begin{pmatrix}a\\b\end{pmatrix}$; ce sont seulement deux façons différentes d'écrire le même élément de $\mathbb{R}^2$. Nous ne faisons pas semblant de croire que l'un ou l'autre serait un « covecteur » suivant que l'on tourne notre page dans un sens ou un autre.

Nous passons maintenant au point (3). Il est facile de voir que A est un point fixe de $R_A(\theta)$ parce que $\tau_A^{-1}(A) = (0,0)$.

Nommons ℓ' la droite horizontale $\mathbb{R}(1,0)$. Nous avons, par le point précédent $R_0(\theta) = \sigma_\ell \circ \sigma_{\ell'}$. En introduisant astucieusement $\tau_A^{-1}\tau_A$ dans l'expression définissant $R_A(\theta)$, nous avons

$$R_A(\theta) = \tau_A\sigma_\ell\sigma_{\ell'}\tau_A^{-1} = \underbrace{\tau_A\sigma_\ell\tau_A^{-1}}_{=\sigma_{\tau_A(\ell)}}\underbrace{\tau_A\sigma_{\ell'}\tau_A^{-1}}_{\sigma_{\tau_A(\ell')}} = \sigma_{\tau_A(\ell)}\sigma_{\tau_A(\ell')}. \tag{18.346}$$

Nous avons utilisé le lemme 18.117.

Nous voyons que $R_A(\theta)$ est une composée de deux réflexions se coupant en A. C'est donc une rotation centrée en A.

□

18.10.7 Rotations vectorielles

L'expression « rotation vectorielle » signifie rotation centrée en zéro. Elles sont « vectorielles » parce qu'elles sont linéaires comme nous le voyons à présent.

PROPooTFNSooFjiWHG

Proposition 18.128.
Quelques résultats à propos de rotations.

ITEMooONJOooRgycsQ

(1) *Toutes les rotations de $\mathbb{R}^2$ centrées en $(0,0)$ sont de la forme $R_0(\theta)$* [34] *pour un $\theta \in \mathbb{R}$.* ITEMooMQPRooAbPcrr

(2) *Les rotations de $\mathbb{R}^2$ centrées en $(0,0)$ sont des applications linéaires.* ITEMooSIHZooBEJhdu

(3) *Si r est une rotation dans $\mathbb{R}^2$, il existe $A \in \mathbb{R}^2$ et $\theta \in \mathbb{R}$*

$$r = \tau_A \circ R_0(\theta) \circ \tau_A^{-1}. \tag{18.347}$$

Démonstration. Soit r une rotation centrée en $(0,0)$. Par la proposition 18.126, il existe une droite ℓ passant par $(0,0)$ telle que $r = \sigma_\ell \circ s$ où s est la réflexion d'axe horizontal : $s(x,y) = (x,-y)$.

Soit un vecteur unitaire $v \in \ell$. Puisque $v \in S^1$, la proposition 18.58 nous donne $t \in [0, 2\pi[$ tel que

$$v = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}. \tag{18.348}$$

En posant $\theta = 2t$ nous avons

$$\ell = \mathbb{R}\begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}, \tag{18.349}$$

et donc $r = R_0(\theta)$ qui est l'application linéaire définie en 18.127. Ceci prouve les points (1) et (2).

Et enfin nous voyons le point (3). Soit A l'unique point fixe de la rotation r. Cette dernière s'écrit alors $r = \sigma_{\ell_1} \circ \sigma_{\ell_2}$ où ℓ_1 et ℓ_2 sont des droites telles que $\ell_1 \cap \ell_2 = \{A\}$.

En utilisant le lemme 18.117, nous avons $\sigma_{\ell_i} = \tau_A \circ \sigma_{\tau_A^{-1}(\ell_i)} \circ \tau_A^{-1}$. En substituant, et en nous rendant compte que $\tau_A^{-1}\tau_A = \mathrm{Id}$, nous avons

$$r = \sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_A r_0 \tau_1^{-1} \tag{18.350}$$

où r_0 est la rotation $\sigma_{\tau_A^{-1}(\ell_1)} \circ \sigma_{\tau_A^{-1}}(\ell_2)$. Cette dernière est une rotation autour de $(0,0)$ parce que $\tau_A^{-1}(\ell_1) \cap \tau_A^{-1}(\ell_2) = \{0\}$. Elle est donc, par le point (1), de la forme $R_0(\theta)$, pour une certaine valeur de θ. □

PROPooWMESooNJMdxf

Proposition 18.129 ([549]).
Les rotations basées en $(0,0)$ forment un groupe abélien.

Démonstration. L'identité est une rotation par définition. En ce qui concerne l'inverse, si $r = \sigma_1\sigma_2$ alors $r^{-1} = \sigma_2\sigma_1$. Nous commençons maintenant les choses pas tout à fait évidentes.

34. Définition 18.127.

(1) **Composition** Soient des rotations r, r' centrées en $(0,0)$. Soit également, une réflexion σ dont l'axe contient $(0,0)$. Alors la proposition 18.126 nous donne l'existence de σ_1 et σ_2 tels que $r = \sigma_1\sigma$ et $r' = \sigma\sigma_2$. Avec ça, la composition donne

$$rr' = \sigma_1\sigma\sigma\sigma_2 = \sigma_1\sigma_2, \tag{18.351}$$

qui est encore une rotation.

(2) **Commutativité** Soient deux rotations r et r' ainsi que des décompositions $r = \sigma_1\sigma$, $r' = \sigma\sigma_2$. Nous avons

$$rr' = \sigma_1\sigma_2 \tag{18.352a}$$
$$r'r = \sigma\sigma_2\sigma_1\sigma. \tag{18.352b}$$

Vu que $t = \sigma_2\sigma_1$ est une rotation, nous pouvons encore appliquer la proposition 18.126 pour avoir $t = \sigma_2\sigma_1 = \sigma\sigma_3$. Ainsi,

$$r'r = \sigma\sigma\sigma_3\sigma = \sigma_3\sigma. \tag{18.353}$$

Mais aussi $rr' = \sigma_1\sigma_2 = t^{-1} = \sigma_3\sigma$. Nous avons donc bien $rr' = r'r$, et le groupe est commutatif.

□

NORMooOUDJooRfbDEX

18.130.
Jusqu'à présent nous avons parlé de rotations « affines ». Parmi elles, les rotations centrées en 0 (zéro, l'origine de E comme espace vectoriel) sont de particulière importance. Ce sont des applications linéaires, et même des isométries. Dans la suite, nous allons souvent dire simplement « rotation » pour dire « rotation centrée en 0 ».

Nous allons maintenant prendre un point de vue plus vectoriel, et allons noter les points de E avec des lettres comme x, y, u, v et non plus, avec des majuscules, comme lorsqu'on avait un point de vue affine. En même temps, nous allons noter les applications $E \to E$ par des lettres comme A et ne plus écrire les parenthèses. Bref, nous écrivons Au au lieu de $r(A)$.

LEMooSYZYooWDFScw

Lemme 18.131.
En dimension 2*, les réflexions vectorielles (c'est-à-dire dont l'axe passe par* 0*) ont un déterminant qui vaut* -1*.*

Démonstration. Soit une réflexion d'axe ℓ. Prenons une base orthonormale de E constituée de e_1 sur ℓ et de $e_2 \perp \ell$. Alors $\sigma_\ell(e_1) = e_1$ et $\sigma_\ell(e_2) = -e_2$. La formule du déterminant donne

$$\det(\sigma_\ell) = e_1^*\big(\sigma_\ell(e_1)\big)e_2^*\big(\sigma_\ell(e_2)\big) - e_2^*\big(\sigma_\ell(e_1)\big)e_1^*\big(\sigma_\ell(e_2)\big) = 1 \times (-1) - 0 \times 0 = -1. \tag{18.354}$$

Nous utilisons de façon cruciale le fait que le calcul du déterminant ne dépende pas de la base choisie, lemme 8.15. □

PROPooTUJWooAjtEnQ

Proposition 18.132.
Les rotations[35] *sont*

(1) des applications linéaires orthogonales au sens de la définition 8.43,

(2) des applications de déterminant 1*,*

Démonstration. Le fait qu'elles soient linéaires est la proposition 18.128.

Nous avons, pour tout $u \in E$ l'égalité de la norme $\|u\|$ et $\|Au\|$ par le lemme 18.123 appliqué à $Y = 0$. En termes de produits scalaires nous avons alors $\langle Au, Au\rangle = \langle u, u\rangle$, et donc

$$\langle A^*Au, u\rangle = \|u\|^2. \tag{18.355}$$

35. Centrées en 0, nous ne le répéterons pas !

En particulier si $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de E nous avons

$$(A^*Ae_i)_i = \|e_i\|^2 = 1, \tag{18.356}$$

ce qui donne $\|A^*Ae_i\| \geqslant 1$, avec égalité si et seulement si $A^*Ae_i = e_i$. Ici nous avons utilisé le fait que $\langle x, e_i\rangle = x_i$, et le fait que pour tout i nous ayons $\|x\| \geqslant |x_i|$, avec égalité seulement si x est un multiple de e_i.

Par ailleurs l'inégalité de Cauchy-Schwarz 10.1 nous donne

$$\|u\|^2 = |\langle A^*Au, u\rangle| \leqslant \|A^*Au\|\|u\| \tag{18.357}$$

et donc

$$\|u\| \leqslant \|A^*Au\|. \tag{18.358}$$

Encore une fois, en appliquant cela à $u = e_i$ nous trouvons $1 \leqslant \|A^*Ae_i\|$. Puisque nous avions déjà l'inégalité dans l'autre sens, $\|A^*Ae_i\| = 1$. Et le cas d'égalité est uniquement possible avec $A^*Ae_i = e_i$.

Donc pour tout i de la base, nous avons $A^*Ae_i = e_i$. Nous avons donc $A^*A = \mathrm{Id}$ et l'application A est orthogonale.

En ce qui concerne le déterminant, les réflexions sont de déterminant -1 par le lemme 18.131, donc $A = \sigma_1 \circ \sigma_2$ est de déterminant 1. Nous avons utilisé le fait que le déterminant était un morphisme : proposition 8.16(1). □

Remarque 18.133.
Nous ne savons pas encore que les rotations forment tout le groupe SO(2) des endomorphismes orthogonaux de déterminant 1. Il faudra attendre le corolaire 18.138 pour le savoir.

LEMooMIJXooCjiQqP

Lemme 18.134.
L'application $-\mathrm{Id}$ *est une rotation de* $\mathbb{R}^2$.

Démonstration. Soit une base orthonormée $\{e_1, e_2\}$ de E et la rotation $r = \sigma_1\sigma_2$ où σ_i est la réflexion le long de l'axe $\ell_i = \{te_i\}_{t\in\mathbb{R}}$. Faut-il vous prouver que $r = -\mathrm{Id}$? La réflexion σ_2 retourne la composante y d'un vecteur écrit dans la base $\{e_1, e_2\}$ sans toucher à la composante x. La réflexion σ_1 fait le contraire. □

18.10.8 Matrice des transformations orthogonales

Nous donnons maintenant une forme générale (trigonométrique) pour les matrices de SO(2). Nous ne pouvons cependant pas invoquer les lemmes 18.135 ou 18.136 pour prétendre avoir une matrice des rotations, parce que nous n'avons pas encore prouvé que les rotations étaient des transformations orthogonales. Ce sera pour la proposition 18.139.

LEMooAJMAooXPSKtS

Lemme 18.135.
Si $A \in \mathrm{O}(2)$ *alors il existe un unique* $(\theta, \epsilon) \in [0, 2\pi[\times \{\pm 1\}$ *tel que*

$$A = \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{18.359}$$

Démonstration. Soit une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et imposons qu'elle soit dans O(2). Le fait que A soit orthogonale impose

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{18.360}$$

Nous avons alors le système

$$\begin{cases} a^2 + c^2 = 1 & \text{(18.361a)} \\ b^2 + d^2 = 1 & \text{(18.361b)} \\ ab + cd = 0 & \text{(18.361c)} \end{cases}$$

La proposition 18.58 nous permet de déduire qu'il existe un unique $\theta \in [0, 2\pi[$ tel que $a = \cos(\theta)$, $c = \sin(\theta)$, ainsi que plusieurs $\alpha \in \mathbb{R}$ tel que $b = \cos(\alpha)$, $d = \sin(\alpha)$.

Note : si nous voulons $\alpha \in [0, 2\pi[$, alors il y a unicité. Ici nous ne nous attachons pas à cette contrainte ; nous savons qu'il en existe plusieurs, et nous allons en fixer un en fonction de θ. Le α ainsi fixé ne sera peut-être pas dans $[0, 2\pi[$, mais ce ne sera pas grave.

Les angles θ et α sont alors liés par la contrainte

$$\cos(\theta)\cos(\alpha) + \sin(\theta)\sin(\alpha) = 0. \tag{18.362}$$

Utilisant l'identité (18.22c) cela signifie que $\cos(\theta - \alpha) = 0$. Donc

$$\alpha \in \{\theta + \frac{\pi}{2} + k\pi\}_{k\in\mathbb{Z}}. \tag{18.363}$$

Si k est pair, ça donne

$$\cos(\alpha) = -\sin(\theta) \tag{18.364a}$$
$$\sin(\alpha) = \cos(\theta) \tag{18.364b}$$

et alors

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.365}$$

EQooNAMKooKACIfd

Si au contraire k est impair, alors

$$\cos(\alpha) = \sin(\theta) \tag{18.366a}$$
$$\sin(\alpha) = -\cos(\theta), \tag{18.366b}$$

et

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{18.367}$$

EQooJMYEooGgAiMJ

Nous avons démontré qu'une matrice de O(2) était forcément d'une des deux formes (18.365) ou (18.367). Il est maintenant facile de vérifier que ces deux matrices sont effectivement dans O(2). □

LEMooHRESooQTrpMz

Lemme 18.136.
Si $g \in \mathrm{SO}(2)$, il existe un unique $\theta \in [0, 2\pi[$ tel que la matrice de g dans la base canonique soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.368}$$

Démonstration. Puisque SO(2) est la partie de O(2) constituée des matrices de déterminant 1, nous pouvons reprendre la forme donnée par le lemme 18.135 et fixer ϵ par la contrainte sur le déterminant. Nous avons, en utilisant la relation du lemme 18.4,

$$\det \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} = \epsilon. \tag{18.369}$$

Étant donné que $\det(g) = 1$, nous voyons que $\epsilon = 1$. □

COROOGGVUooLQYGET

Corolaire 18.137 ([1]).
Nous avons une bijection

$$\begin{aligned} \psi\colon \mathrm{SO}(2) &\to [0, 2\pi[\\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} &\mapsto \theta, \end{aligned} \tag{18.370}$$

et un isomorphisme de groupe

$$\begin{aligned} \varphi\colon \mathrm{SO}(2) &\to \mathrm{U}(1) = \{e^{i\theta}\}_{\theta\in\mathbb{R}} \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} &\mapsto e^{i\theta}. \end{aligned} \tag{18.371}$$

Démonstration. La première assertion est une paraphrase du lemme 18.136. Pour la seconde, il faut vérifier que c'est bien un morphisme et une bijection.

Pour le morphisme, ce sont les formules d'addition d'angle du lemme 18.14 qui jouent. En ce qui concerne la bijection...

(1) **Surjection** Comme $e^{i\theta+2ki\pi} = e^{i\theta}$, tout élément de U(1) est exponentielle de $i\theta$ pour un $\theta \in [0, 2\pi[$.

(2) **Injection** Remarquez que nous avons, pour $\theta \in [0, 2\pi[$, $\varphi\big(\psi^{-1}(\theta)\big) = e^{i\theta}$.
Soient $A, B \in \mathrm{SO}(2)$ tels que $\varphi(A) = \varphi(B)$. Soient $\theta_a, \theta_b \in [0, 2\pi[$ tels que $A = \psi^{-1}(\theta_a)$ et $B = \psi^{-1}(\theta_b)$. L'hypothèse $\varphi(A) = \varphi(B)$ donne

$$e^{i\theta_a} = e^{i\theta_b}, \tag{18.372}$$

et donc $\theta_a = \theta_b$ par la proposition 18.61(1).
Nous en déduisons que $A = B$, et donc que φ est injective.

□

18.10.9 Rotations, SO(2) et matrice de rotation

CORooVYUJooDbkIFY

Corolaire 18.138 ([1]).
Le groupe des rotations centrées en $(0,0)$ est le groupe SO(2).

Démonstration. Nous devons prouver deux choses :

— Toutes les rotations sont des éléments de SO(2).

— Tous les éléments de SO(2) sont des rotations.

La proposition 18.128 nous indique que toute rotation de $\mathbb{R}^2$ centrée en $(0,0)$ est de la forme $R_0(\theta)$, c'est-à-dire, a une matrice de la forme

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.373}$$

EQooSJNBooMPIBZS

Donc toute rotation est dans SO(2).

D'autre part, le lemme 18.136 indique que tout élément de SO(2) a, dans la base canonique, une matrice de la forme (18.373). Le lemme 18.127 nous indique alors que c'est une rotation. □

PROPooOTIVooZpvLnb

Proposition 18.139.
Si r est une rotation de $\mathbb{R}^2$ centrée en $(0,0)$, il existe un unique $\theta \in [0, 2\pi[$ tel que $r = R_0(\theta)$.

Démonstration. Soit une rotation r autour de $(0,0)$. Le corolaire 18.138 nous dit qu'il existe $\theta \in \mathbb{R}$ tel que

$$r = R_0(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.374}$$

Nous avons identifié l'application linéaire à sa matrice. L'élément $\big(\cos(\theta), \sin(\theta)\big)$ est dans S^1, et il existe donc, par la proposition 18.58, un unique $t \in [0, 2\pi[$ tel que $\cos(t) = \sin(\theta)$ et $\sin(t) = \sin(\theta)$. Pour ce t nous avons alors

$$r = R_0(\theta) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}. \tag{18.375}$$

□

PROPooISUCooRYJcwo

Proposition 18.140.
Nous avons la formule suivante pour la composition :

$$R_0(\alpha) \circ R_0(\beta) = R_0(\alpha + \beta). \tag{18.376}$$

Démonstration. Par définition de $R_0(\theta)$, dans la base canonique de $\mathbb{R}^2$, la composition se calcule avec le produit suivant, en utilisant les formules du lemme 18.14 :

$$\begin{pmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{pmatrix}\begin{pmatrix}\cos(\beta) & -\sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{pmatrix} \tag{18.377a}$$

$$= \begin{pmatrix}\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta)-\sin(\alpha)\cos(\beta)\\ \sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta)+\cos(\alpha)\cos(\beta)\end{pmatrix} \tag{18.377b}$$

$$= \begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix}. \tag{18.377c}$$

Donc dans la base canonique, la matrice de $R_0(\alpha)R_0(\beta)$ est celle de $R_0(\alpha+\beta)$. □

18.10.10 Rotation et application affine

Nous considérons à nouveau la définition 11.12 d'une application affine ainsi que sa décomposition en application linéaire et translation donnée par le lemme 11.13. Nous voyons maintenant comment ces propriétés se déclinent dans le cas d'une rotation non centrée en l'origine.

Exemple 18.141.
Soit $A \in \mathbb{R}^2$ ainsi qu'une rotation f autour de A, c'est-à-dire une composition de deux symétries dont les axes se coupent en A[36]. Nous allons extraire de f la partie linéaire définie en 11.13.

Il existe des axes ℓ_1 et ℓ_2 tels que $\ell_1 \cap \ell_2 = \{A\}$ et tels que

$$f = s_{\ell_1} \circ s_{\ell_2}. \tag{18.378}$$

En utilisant le lemme 18.117,

$$f = t_A \circ s_{t_A^{-1}(\ell_1)} \circ t_A^{-1} \circ t_A \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1} = t_A \circ s_{t_A^{-1}(\ell_1)} \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1}. \tag{18.379}$$

Comme $A \in \ell_i$, nous avons $O = (0,0) \in t_A^{-1}(\ell_i)$. Donc les axes $t_A^{-1}(\ell_1)$ et $t_A^{-1}(\ell_2)$ se coupent en O et nous pouvons écrire

$$f = t_A \circ R \circ t_A^{-1} \tag{18.380}$$

où R est une rotation centrée en O ; donc une application linéaire par la proposition 18.128.

Nous avons

$$f(x) = (t_A \circ R \circ t_A^{-1})(x) \tag{18.381a}$$
$$= (t_A \circ R)(x - A) \tag{18.381b}$$
$$= t_A\big(R(x) - R(A)\big) \tag{18.381c}$$
$$= R(x) + t_{R(A)+A} \tag{18.381d}$$
$$= (t_{R(A)+A} \circ R)(x). \tag{18.381e}$$

Donc

$$f = t_{R(A)+A} \circ R. \tag{18.382}$$

Il est maintenant aisé de montrer que R est la partie linéaire de f. Pour tout $M, x \in \mathbb{R}^2$ nous avons

$$f(M+x) = R(M+x) + R(A) + A \tag{18.383a}$$
$$= R(M) + R(x) + R(A) + A \tag{18.383b}$$
$$= R(x) + f(M). \tag{18.383c}$$

Donc nous confirmons la formule

$$f(M+x) = R(x) + f(M) \tag{18.384}$$

et R est la partie linéaire de f, voir la définition 11.13. △

36. Voir la définition 18.118.

LEMooUKEVooAEWvlM

Lemme 18.142 ([1]).
Tout sous-groupe fini de $\mathrm{SO}(2)$ *est cyclique*[37].

Démonstration. Soit un sous-groupe fini G de $\mathrm{SO}(2)$. Nous considérons les applications $\psi\colon \mathrm{SO}(2) \to [0,2\pi[$ et $\varphi\colon \mathrm{SO}(2) \to U(1)$ données par le corolaire 18.137.

Vu que $\psi(e) = 0$, la partie $\psi(G\backslash\{e\})$ est une partie finie de $[0,2\pi[$ qui ne contient pas 0. Nous en considérons le minimum $\theta_0 \neq 0$. Gardez en tête que ce minimum n'est pas nul : $\theta_0 \neq 0$.

Nous allons prouver que $\psi^{-1}(\theta_0)$ génère le groupe G. Pour cela nous devons montrer que pour tout $g \in G$, il existe un entier k tel que $g = \psi^{-1}(\theta_0)^k$.

Soient $g_1 \in \mathrm{SO}(2)$ et $\theta_1 = \psi(g_1)$. Si il existe $k \in \mathbb{N}$ tel que $\theta_1 = k\theta_0$, alors le résultat est prouvé. Supposons donc que $\theta_1 = \lambda\theta_0$ avec λ non entier. Nous allons en déduire une contradiction.

Soit k la partie entière de λ (définition 1.476). Vu que φ est un morphisme de groupes, en utilisant la formule de la proposition 18.9(2) nous avons :

$$\begin{aligned}
\varphi^{-1}\big(e^{i(\theta_1-k\theta_0)}\big) &= \varphi^{-1}\big(e^{i\theta_1}e^{-ik\theta_0}\big) && \text{(18.385a)}\\
&= \varphi^{-1}(e^{i\theta_1})\varphi^{-1}(e^{-ik\theta_0}) && \text{(18.385b)}\\
&= \varphi^{-1}(e^{i\theta_1})\varphi^{-1}(e^{t\theta_0})^k && \text{(18.385c)}\\
&\in G. && \text{(18.385d)}
\end{aligned}$$

Nous avons donc $\theta = \theta_1 - k\theta_0 \in \psi(G)$. Mais en utilisant le lemme 1.477,

$$\theta - \theta_0 = \theta_1 - (k+1)\theta_0 < \theta_1 - \lambda\theta_0 = 0. \tag{18.386}$$

Cela signifie que $\theta < \theta_0$, et contredit la minimalité de θ_0. □

18.10.11 Angle orienté

Avant d'aborder la classification des isométries, nous devons parler de l'angle entre deux droites. Si ℓ_1 et ℓ_2 sont deux droites, alors il est bien clair que deux angles peuvent prétendre être « l'angle entre ℓ_1 et ℓ_2 ». De plus chacun de ces deux angles sont doubles parce que si α peut prétendre être l'angle entre ℓ_1 et ℓ_2, alors $-\alpha$ peut également y prétendre.

Remarque 18.143.
Nous ne parlons pas de l'angle entre ℓ_1 et ℓ_2 mais bien de l'angle *de* ℓ_1 *à* ℓ_2. L'ordre des droites est important.

18.144.
Pour la suite, $R_O(\alpha)$ est la rotation d'angle α autour du point O tandis que $R(\alpha)$ est la rotation d'angle α autour de l'origine.

PROPooDWIMooQPkobw

Proposition 18.145 ([550]).
Si u *et* v *sont des vecteurs unitaires*[38] *de* $\mathbb{R}^2$ *alors il existe une unique rotation*[39] f *telle que* $f(u) = v$.

Démonstration. C'est la proposition 18.124 appliquée à $O = (0,0)$. □

Remarque 18.146.
Notons l'unicité. Nous ne faisons pas de différence entre R_θ et $R_{\theta+2\pi}$ et les autres $R_{\theta+2k\pi}$. En particulier si une rotation T est donnée, dire « $T = R_\theta$ » ne définit pas un nombre θ de façon univoque. Par contre ça définit une classe modulo 2π, c'est-à-dire un élément $\theta \in \mathbb{R}/2\pi$.

Nous avons déjà défini le groupe $\mathrm{SO}(2)$ en la définition 8.47 et nous avons déterminé ses matrices dans $\mathbb{R}^2$ en le lemme 18.136.

37. Définition 1.326.
38. De norme 1.
39. Définition 18.118.

La proposition 18.145 donne une application

$$T\colon S^1\times S^1\to \mathrm{SO}(2). \tag{18.387}$$

Et nous avons une relation d'équivalence sur $S^1\times S^1$ donnée par $(u,v)\sim(u',v')$ si et seulement si il existe $g\in\mathrm{SO}(2)$ telle que $g(u)=u'$ et $g(v)=v'$.

DEFooVBKIooWlHvod

Définition 18.147 (Angle orienté[550])**.**
*Les classes de $S^1\times S^1$ pour cette relation d'équivalence sont les **angles orientés de vecteurs**. Nous notons $[u,v]$ la classe de (u,v).*

PROPooIWJQooGQJBWR

Proposition 18.148.
Nous avons $T(u,v)=T(u',v')$ si et seulement si $(u,v)\sim(u',v')$.

Démonstration. En utilisant la commutativité du groupe $\mathrm{SO}(2)$ nous avons équivalence entre les affirmations suivantes :

— $(u,v)\sim(u',v')$
— $T(u,u')=T(v,v')$
— $T(u,u')\circ T(u',v)=T(v,v')\circ T(u',v)$
— $T(u,v)=T(u',v')$.

□

Proposition 18.149.
Nous avons une bijection

$$\begin{aligned} S\colon \frac{S^1\times S^1}{\sim}&\to \mathrm{SO}(2)\\ [u,v]&\mapsto T(u,v). \end{aligned} \tag{18.388}$$

Démonstration. En plusieurs points.

(1) **S est bien définie** En effet si $[u,v]=[z,t]$ alors $T(u,v)=T(z,t)$.
(2) **Injectif** Si $S[u,v]=S[z,t]$ alors $T(u,v)=T(z,t)$, qui implique $(u,v)\sim(z,t)$ par la proposition 18.148.
(3) **Surjectif** Nous avons $R_\theta=T(u,R_\theta u)$.

□

Définition 18.150 (Somme d'angles orientés[550])**.**
Si $[u,v]$ et $[z,t]$ sont des angles orientés, nous définissons la somme par

$$[u,v]+[z,t]=S^{-1}\Big(S[u,v]\circ S[z,t]\Big). \tag{18.389}$$

LEMooWISVooYsStJp

Lemme 18.151.
Quelques propriétés des angles plats liées à la somme.

(1) $(S^1\times S^1)/\sim$ est un groupe commutatif.

ITEMooBKTFooWbEvIU

(2) Relations de Chasles :

$$[u,v]+[v,w]=[u,w]. \tag{18.390}$$

(3) $-[u,v]=[v,u]$.

Démonstration. Pour la relation de Chasles, ça se base sur la propriété correspondante sur T :

$$\begin{aligned} [u,v]+[v,w]&=S^{-1}\Big(T(u,v)\circ T(v,w)\Big) && (18.391a)\\ &=S^{-1}\big(T(u,w)\big) && (18.391b)\\ &=[u,w]. && (18.391c) \end{aligned}$$

Pour l'inverse, la vérification est que

$$[u,v]+[v,u]=[u,u]=0. \tag{18.392}$$

□

DEFooFLGNooCZUkHY

Définition 18.152.
*La **mesure** de l'angle orienté $[u,v]$ est $[\theta]_{2\pi}$ si $T[u,v]=R_\theta$.*

Notons dans cette définition qu'écrire $T[u,v]=R_\theta$ dans $\mathrm{SO}(2)$ ne définit pas θ, mais seulement sa classe modulo 2π. C'est pour cela que la mesure de l'angle orienté n'est également définie que modulo 2π.

Pour la suite nous allons nous intéresser à des vecteurs qui ont, dans l'idée, un point de départ et un point d'arrivée. Si $A,B\in\mathbb{R}^2$ nous notons

$$\overrightarrow{AB}=\frac{B-A}{\|B-A\|}. \tag{18.393}$$

C'est le vecteur unitaire dans la direction « de A vers B ».

THOooQDNKooTlVmmj

Théorème 18.153 (Théorème de l'angle inscrit[551])**.**
Soit un cercle Γ de centre O et trois points distincts $A,B,M\in\Gamma$. Alors

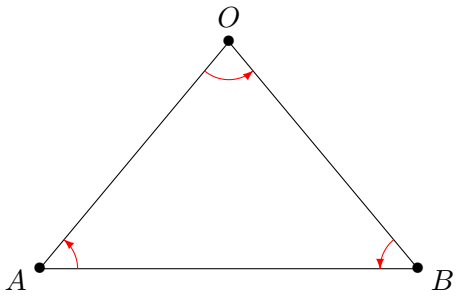
$$2(\overrightarrow{MA},\overrightarrow{MB})\in(\overrightarrow{OA},\overrightarrow{OB})_{2\pi} \tag{18.394}$$

où l'indice 2π indique la classe modulo 2π.

Démonstration. Le triangle MOA est isocèle en O, donc les angles à la base sont égaux. Et de plus la somme des angles est dans $[\pi]_{2\pi}$. Bon, entre nous, nous savons que la somme des angles est exactement π, mais comme nous n'avons pas défini les angles autrement que modulo π, nous ne pouvons pas dire mieux. Donc

$$2(\overrightarrow{AB},\overrightarrow{AO})+(\overrightarrow{OB},\overrightarrow{OA})\in[\pi]_{2\pi}. \tag{18.395}$$

Il faut être sûr de l'orientation de tout cela. Le nombre $(\overrightarrow{AB},\overrightarrow{AO})$ est l'angle qui sert à amener $\overrightarrow{AB}$ sur $\overrightarrow{AO}$. Vu que nous l'avons choisi dans le sens trigonométrique, il faut bien prendre les autres dans le sens trigonométrique, et utiliser $(\overrightarrow{OA},\overrightarrow{OB})$ et non $(\overrightarrow{OB},\overrightarrow{OA})$.



De la même manière sur le triangle MOB nous écrivons

$$2(\overrightarrow{MB},\overrightarrow{MO})+(\overrightarrow{OM},\overrightarrow{OB})\in[\pi]_{2\pi}. \tag{18.396}$$

Nous faisons la différence entre les deux équations en remarquant que la différence de deux représentants de $[\pi]_{2\pi}$ est un représentant de $[0]_{2\pi}$ et en nous souvenant que $-(\overrightarrow{MB},\overrightarrow{MO})=(\overrightarrow{MO},\overrightarrow{MB})$ et les relations de Chasles du lemme 18.151(2) nous avons :

$$2(\overrightarrow{MA},\overrightarrow{MB})+(\overrightarrow{OB},\overrightarrow{OA})\in[0]_{2\pi}. \tag{18.397}$$

□

18.154.
Comment exprimer le fait qu'un angle orienté soit égal à θ modulo π alors que les angles orientés sont des classes modulo 2π ? Nous ne pouvons certainement pas écrire

$$(u,v) = [\theta]_\pi \tag{18.398}$$

parce que (u,v) est un élément de $S^1 \times S^1$ alors que $[\theta]_\pi$ est un ensemble de nombres. Nous pouvons écrire

$$[u,v] \subset [\theta]_\pi. \tag{18.399}$$

C'est cohérent parce que nous avons des deux côtés des ensembles de nombres. Les opérations permises sont l'égalité ou l'inclusion. L'égalité entre les deux ensembles n'est pas possible parce que la différence minimale ente deux éléments dans $[u,v]$ est 2π alors que celle dans $[\theta]_\pi$ est π.

Si u et v forment un angle droit, nous avons

$$[u,v] = \{\frac{\pi}{2} + 2k\pi\}_{k\in\mathbb{Z}}. \tag{18.400}$$

Et cela est bien un sous-ensemble de $[\pi/2]_\pi$.

Pour exprimer que deux angles orientés diffèrent de π nous devrions écrire :

$$[u,v] \subset [a,b]_\pi \tag{18.401}$$

où le membre de droite signifie la classe modulo π d'un représentant de $[a,b]$.

Nous allons cependant nous permettre d'écrire

$$[u,v] = [a,b]_\pi \tag{18.402}$$

voire carrément

$$(u,v) = (a,b)_\pi. \tag{18.403}$$

Cette dernière égalité devant être comprise comme voulant dire que l'angle pour passer de u à v est, soit le même que celui pour aller de a à b, soit ce dernier, plus π.

THOooUDUGooTJKDpO

Théorème 18.155 ([551])**.**
Soient 4 points distincts du plan A,B,C,D. Ils sont alignés ou cocycliques[40] *si et seulement si*

$$(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{DA},\overrightarrow{DB})_\pi. \tag{18.404}$$

Nous allons seulement démontrer l'implication directe.

Démonstration. Si les quatre points sont alignés nous avons $[\overrightarrow{CA},\overrightarrow{CB}] = [0]_{2\pi}$ et $[\overrightarrow{DA},\overrightarrow{DB}] = [0]_{2\pi}$. En particulier nous avons

$$[\overrightarrow{CA},\overrightarrow{CB}] = [\overrightarrow{DA},\overrightarrow{DB}] \tag{18.405}$$

et a fortiori l'égalité modulo π au lieu de 2π.

Nous nous relâchons en termes de notations. Si les quatre points sont cocycliques, nous pouvons utiliser le théorème de l'angle inscrit 18.153 dans les triangles ABC et ADB :

$$2(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi} \tag{18.406a}$$
$$2(\overrightarrow{DA},\overrightarrow{DB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi}, \tag{18.406b}$$

ce qui donne $2(\overrightarrow{CA},\overrightarrow{CB}) = 2(\overrightarrow{DA},\overrightarrow{DB})_{2\pi}$ et donc

$$(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{DA},\overrightarrow{DB})_\pi. \tag{18.407}$$

Comme annoncé, nous ne faisons pas la preuve dans l'autre sens ; elle peut être trouvée dans [551]. □

40. C'est-à-dire sur un même cercle.

EXooOXAAooZMdDfP

Exemple 18.156.
À propos de groupe engendré et de générateur[41]. Soit G le groupe des rotations d'angle[42] $k\pi/5$ (avec k entier). Ce groupe est constitué des « dixièmes de tour », puisque $\frac{k\pi}{5} = \frac{2k\pi}{10}$.

La rotation d'angle $2\pi/5$ n'est pas génératrice parce qu'elle n'engendre que des « cinquièmes de tour » : $4\pi/5$, $6\pi/5$, $8\pi/5$ et l'identité.

Par contre, la rotation d'angle $\pi/5$ est génératrice. △

18.10.12 Angles et nombres complexes

SUBSECooKNUVooUBKaWm

Les nombres complexes peuvent être repérés par une norme et un angle, ce qui en fait un terrain propice à l'utilisation des angles orientés. Nous en ferons d'ailleurs usage dans $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ pour parler d'alignement, de cocyclicité et de birapport dans la proposition 23.88.

Soient deux éléments $z_1, z_2 \in \mathbb{C}$. Nous les écrivons sous la forme $z_1 = r_1 e^{i\theta_1}$ et $z_2 = r_2 e^{i\theta_2}$; remarquons que cela ne définit θ_i qu'à 2π près. Nous avons

$$[z_1, z_2] = [\theta_2 - \theta_1]_{2\pi}. \tag{18.408}$$

Soient maintenant $a, b, c, d \in \mathbb{C}$. Nous écrivons $\overrightarrow{ab}$ le vecteur unitaire dans le sens « de a vers b », c'est-à-dire un multiple positif bien choisi du nombre $b - a$. Nous notons θ_{ab} l'argument du nombre complexe $b - a$, et nous avons encore

$$[\overrightarrow{ab}, \overrightarrow{cd}] = [\theta_{ab} - \theta_{cd}]. \tag{18.409}$$

Avec toutes ces notations, ce qui est pas mal, c'est que les produits et quotients de nombres complexes se comportent très bien par rapport aux angles : l'argument de a/b est $\theta_a - \theta_b$ et en particulier l'argument de

$$\frac{a-b}{c-d} \tag{18.410}$$

est dans la classe de l'angle orienté

$$[\overrightarrow{ba}, \overrightarrow{dc}]. \tag{18.411}$$

DEFooUPUUooKAPFrh

Définition 18.157 (Angle avec 3 points)**.**
Soient trois points $A, O, B \in \mathbb{R}^2$. Voici comment nous définissons l'angle $\widehat{AOB}$; de façon informelle, c'est l'angle de la rotation qui permet d'aller de A vers B.

- *Nous nous mettons en l'origine : $A' = A - O$ et $B' = B - O$.*
- *Nous normalisons : $A'' = A'/\|A'\|$ et $B'' = B'/\|B'\|$.*
- *Soit f, l'unique rotation telle que $f(A'') = B''$ (proposition 18.145).*
- *Soit θ l'unique élément de $[0, 2\pi[$ tel que la matrice de f dans la base canonique soit*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{18.412}$$

  par la proposition 18.139.
- *L'angle $\widehat{AOB}$ est ce nombre θ.*

Nous noterons également l'angle $\widehat{AOB}$ par $\text{angle}(A, O, B)$.

Nous voyons que l'angle est toujours un nombre entre 0 et 2π. Par abus de notation, nous admettrons de temps en temps, de parler d'angle en dehors de cet intervalle.

41. Définition 1.325 et 1.326
42. Voir la définition 18.152.

DEFooMAKWooGRutOk

Définition 18.158 (Angle dans $\mathbb{C}$).
Si $a, b, c \in \mathbb{C}$, nous définissons l'angle $\text{angle}(a, b, c) = \text{angle}\big(s(a), s(b), s(c)\big)$ *où*

$$\begin{aligned} s\colon \mathbb{C} &\to \mathbb{R}^2 \\ x + iy &\mapsto (x, y), \end{aligned} \tag{18.413}$$

et le membre de droite étant défini par 18.157.

Avec cette définition, la multiplication dans $\mathbb{C}$ par un élément de S^1 revient à une rotation. C'est le lemme suivant.

LEMooUFOHooWdgIvO

Lemme 18.159 ([1]).
Soit

$$\begin{aligned} s\colon \mathbb{C} &\to \mathbb{R}^2 \\ x + iy &\mapsto (x, y). \end{aligned} \tag{18.414}$$

Nous avons

$$s(e^{i\xi}z) = R_0(\xi)s(z) \tag{18.415}$$

pour tout $\xi \in \mathbb{R}$ et $z \in \mathbb{C}$.

Démonstration. Nous écrivons $z = re^{i\theta}$ comme l'autorise la proposition 18.62. En d'autres termes, $z = r\cos(\theta) + ir\sin(\theta)$ et donc

$$s(z) = \big(r\cos(\theta), r\sin(\theta)\big). \tag{18.416}$$

D'autre part, la proposition 18.9 donne $e^{i\xi}z = re^{i(\theta+\xi)}$, et donc

$$s(e^{i\xi}z) = \begin{pmatrix} r\cos(\theta + \xi) \\ r\sin(\theta + \xi) \end{pmatrix}. \tag{18.417}$$

La conclusion est un calcul :

$$rR(\xi)\big(\cos(\theta), \sin(\theta)\big) = r\begin{pmatrix} \cos(\xi) & -\sin(\xi) \\ \sin(\xi) & \cos(\xi) \end{pmatrix}\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \qquad \text{prop. 18.139} \tag{18.418a}$$

$$= r\begin{pmatrix} \cos(\xi)\cos(\theta) - \sin(\xi)\sin(\theta) \\ \sin(\xi)\cos(\theta) + \cos(\xi)\sin(\theta) \end{pmatrix} \tag{18.418b}$$

$$= r\begin{pmatrix} \cos(\xi + \theta) \\ \sin(\xi + \theta) \end{pmatrix} \qquad \text{lemme 18.14} \tag{18.418c}$$

$$= s(e^{i\xi}z). \tag{18.418d}$$

□

LEMooXMYDooRrNblP

Lemme 18.160.
Pour tout $v \in S^1$, l'application

$$\begin{aligned} s\colon [0, 2\pi[&\to S^1 \\ \theta &\mapsto R(\theta)v \end{aligned} \tag{18.419}$$

est une bijection.

LEMooMFXWooGymSoF

Lemme 18.161 ([1]).
Si les points 0, $R(\theta_1)v$ et $R(\theta_2)v$ sont alignés, avec $\theta_1 < \theta_2$ dans $[0, 2\pi[$, alors $\theta_2 \in \{\theta_1, \theta_1 + \pi\}$.

Démonstration. Supposons dans un premier temps que $\|v\| = 1$. Étant donné que $R(\theta_2)v$ est sur la droite contenant 0 et $R(\theta_1)v$, il existe $\lambda \in \mathbb{R}$ tel que

$$R(\theta_2)v = \lambda R(\theta_1)v. \tag{18.420}$$

Vu que $R(\theta)$ est une isométrie, nous avons $\lambda = \pm 1$. Si $\lambda = 1$, alors nous avons $R(\theta_1)v = R(\theta_2)v$ et donc $\theta_1 = \theta_2$ par le lemme 18.160.

Si $\lambda = -1$, alors en remarquant que $R(\pi)v = -v$, nous avons $R(\theta_1 + \pi)v = R(\theta_1)R(\pi)v = -R(\theta_1)v$. Donc le lemme 18.160 nous dit que $\theta_1 + \pi = \theta_2$. □

DEFooEGKOooRPGOAs

Lemme-Définition 18.162 (Angle entre deux droites[1]).
Soient deux droites ℓ_1 et ℓ_2 de $\mathbb{R}^2$ sécantes au point $A \in \mathbb{R}^2$.
Il existe un unique $\theta_0 \in [0, \pi[$ tel que[43] *$R_A(\theta_0)x \in \ell_2$ pour tout $x \in \ell_1$.*
*Cet angle θ_0 est l'**angle** de ℓ_1 à ℓ_2.*

Démonstration. Nous considérons le vecteur v tel que $\ell_1 = \{A+tv\}_{t\in\mathbb{R}}$, ainsi qu'un point arbitraire $x_0 = A + \lambda v$ sur ℓ_1. Nous allons montrer qu'il existe un unique $\theta_\lambda \in]0, \pi[$ tel que $R_A(\theta_\lambda)x_0 \in \ell_2$. Ensuite nous allons montrer que $\theta_{\lambda_1} = \theta_{\lambda_2}$ pour tout λ_1, λ_2 dans $\mathbb{R}$.

Nous considérons une application affine

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ x &\mapsto u(x) + \alpha \end{aligned} \tag{18.421}$$

avec $u\colon \mathbb{R}^2 \to \mathbb{R}$ linéaire et $\alpha \in \mathbb{R}$ telle que $\ell_2 = \ker(f)$ (proposition 12.153). Nous supposons que $f(x_0) > 0$ pour fixer les idées.

Nous utilisons la formule (18.337) pour R_A ; pour tout θ nous avons

$$R_A(\theta)(A + \lambda v) = (\tau_A \circ R(\theta) \circ \tau_A^{-1})(A + \lambda v) = \tau_A\big(R(\theta)(\lambda)v\big) = A + \lambda R(\theta)v. \tag{18.422}$$

Nous avons d'une part

$$f(x_0) = f(A + \lambda v) = u(A + \lambda v) + \alpha = u(A) + \alpha + \lambda u(v) = f(A) + \lambda u(v) = \lambda u(v) \tag{18.423}$$

parce que $f(A) = 0$ du fait que $A \in \ell_2$.

De même, en tenant compte que $R(\pi)v = -v$ nous avons

$$R_A(\pi)x_0 = u(-\lambda v + A) + \alpha = -\lambda u(v). \tag{18.424}$$

Donc $f\big(R_A(\pi)x_0\big) = -f(x_0)$.

Nous avons prouvé qu'en posant

$$\begin{aligned} s\colon [0, \pi] &\to \mathbb{R}^2 \\ \theta &\mapsto R_A(\theta)(A + \lambda v), \end{aligned} \tag{18.425}$$

nous avons $s(0) = A + \lambda v$, $s(\pi) = A - \lambda v$ et que $(f \circ s)(0) > 0$ et $(f \circ s)(\pi) < 0$. Nous pouvons donc appliquer le théorème des valeurs intermédiaires 9.93 à la fonction continue $f \circ s$. Il existe $\theta_0 \in]0, \pi[$ tel que $(f \circ s)(\theta_0) = 0$.

Pour une telle valeur de θ_0, nous avons $s(\theta_0) \in \ker(f)$ et donc $R_A(\theta_0)x_0 \in \ell_2$.

(1) **Unicité** Nous prouvons que si $R_A(\theta_1)x_0$ et $R_A(\theta_2)x_0$ sont sur ℓ_2, alors $\theta_1 = \theta_2$. Supposons que ce les points A, $R_A(\theta_1)x_0$ et $R_A(\theta_2)x_0$ sont alignés sur ℓ_2. Alors les points 0, $R(\theta_1)v$ et $R(\theta_2)v$ sont alignés, et donc le lemme 18.161 nous assure que $\theta_2 = \theta_1 + \pi$ ou que $\theta_1 = \theta_2$. La possibilité $\theta_2 = \theta_1 + \pi$ est exclue parce que θ_1 et θ_2 sont dans $]0, \pi[$. Donc $\theta_1 = \theta_2$.
Nous avons vu qu'il existe une unique angle, que nous notons θ_λ tel que $R_A(\theta_\lambda)(A+\lambda v) \in \ell_2$.

(2) **Petite note au passage** La droite ℓ_2 est l'ensemble

$$\ell_2 = \{A + tR(\theta_\lambda)v\}_{t\in\mathbb{R}}. \tag{18.426}$$

(3) **Ce θ_λ fonctionne sur tout ℓ_1** Tout élément de ℓ_1 peut être écrit sous la forme $x_0 + \mu v$. Nous prouvons à présent que si $x_1 = x_0 + \mu v$, alors $R_A(\theta)x_1 \in \ell_2$. Nous avons

$$\begin{aligned} R_A(\theta_\lambda)v &= R_A(\theta_\lambda)(x_0 + \mu v) && \text{(18.427a)} \\ &= R_A(\theta_\lambda)\big(A + (\lambda + \mu)v\big) && \text{(18.427b)} \\ &= R(\theta_\lambda)\big((\lambda + \mu)v\big) + A && \text{(18.427c)} \\ &= (\lambda + \mu)R(\theta_\lambda)v + A. && \text{(18.427d)} \end{aligned}$$

Donc cela est dans ℓ_2.

43. Pour rappel, $R_A(\theta)$ est défini en 18.127.

□

PROPooKVSHooRODGWE

Proposition 18.163 ([1]).
Les angles sont invariants sous les translations.

Plus précisément, si $A, B, S, v \in \mathbb{R}^2$, alors

$$\widehat{T_v(A)T_v(S)T_v(B)} = \widehat{ASB} \tag{18.428}$$

où $T_v(X) = X + v$.

Démonstration. Nous notons $X_v = X + v$. Nous avons $A'_v = A_v - S_v = (A+v)-(S+v) = A-S = A''$. Donc les vecteurs A'' et B'' à partir desquels est calculé $\widehat{ASB}$ sont les mêmes que les vecteurs A''_v et B''_v qui servent à calculer $\widehat{T_v(A)T_v(S)T_v(B)}$. □

PROPooYWKJooRjybUJ

Proposition 18.164 ([1]).
Les angles sont invariants par rotations, c'est-à-dire que si $A, B, S \in \mathbb{R}^2$ et si R_θ est une rotation, alors

$$\widehat{ASB} = \widehat{R_\theta(A)R_\theta(S)R_\theta(B)}. \tag{18.429}$$

Démonstration. Pour être plus concis, nous écrivons A_θ for $R_\theta(A)$ et de même pour B et S. Afin de calculer l'angle $\widehat{R_\theta(A)R_\theta(S)R_\theta(B)}$, nous définissons

$$\begin{cases} A'_\theta = A_\theta - S_\theta & \text{(18.430a)} \\ S'_\theta = 0 & \text{(18.430b)} \\ B'_\theta = B_\theta - S_\theta. & \text{(18.430c)} \end{cases}$$

et

$$\begin{cases} A''_\theta = \dfrac{A_\theta - S_\theta}{\|A_\theta - S_\theta\|} & \text{(18.431a)} \\ B''_\theta = \dfrac{B_\theta - S_\theta}{\|B_\theta - S_\theta\|}. & \text{(18.431b)} \end{cases}$$

Par définition, l'angle est le α tel que $R_\alpha(A''_\theta) = B''_\theta$. Nous devons prouver que le même α vérifie $R_\alpha(A'') = B''$.

Le fait que R_θ soit une isométrie nous donne déjà

$$\|R_\theta(A) - R_\theta(B)\| = \|A - B\|. \tag{18.432}$$

Ensuite, la relation de définition de α s'écrit

$$\frac{R_\alpha R_\theta(A) - R_\alpha R_\theta(S)}{\|A_\theta - S_\theta\|} = \frac{R_\theta(B) - R_\theta(S)}{\|B_\theta - S_\theta\|}. \tag{18.433}$$

Vu que R_α et R_θ commutent, nous avons

$$R_\theta \frac{R_\alpha(A) - R_\alpha(B)}{\|A - S\|} = R_\theta \frac{B - S}{\|B - S\|}, \tag{18.434}$$

et comme R_θ est inversible, cela donne $R_\alpha(A'') = B''$. □

LEMooJLHGooQIpKIE

Lemme 18.165 ([1]).
Soit $A \in \mathbb{R}^2$ et une droite ℓ_1. Soit ℓ_2 une droite passant par A et intersectant ℓ_1 en O. Alors

$$\sigma_{\ell_1}(A) = R_O(-2\alpha)A \tag{18.435}$$

où α est l'angle de ℓ_1 à ℓ_2.

Démonstration. Nous allons utiliser des coordonnées autour de O. Il existe un vecteur v tel que

$$A = O + v \tag{18.436}$$

Par définition de l'angle α [44], la droite ℓ_2 s'obtient par rotation d'angle α depuis la droite ℓ_1. Donc le point

$$B = R_O(-\alpha)A \tag{18.437}$$

est sur ℓ_1.

Nous allons prouver que le point

$$D = R_O(-2\alpha)A \tag{18.438}$$

est $D = \sigma_{\ell_1} A$.

Nous commençons par montrer que la droite (DA) est perpendiculaire à ℓ_1, c'est-à-dire que

$$(D - A) \cdot (B - O) = 0. \tag{18.439}$$

En utilisant le fait que

$$R_O(\alpha)(O + X) = O + R(\alpha)X, \tag{18.440}$$

nous avons

$$D - A = R_O(-2\alpha)(O + v) - (O + v) = O + R(-2\alpha)v - O - v = R(-2\alpha)v - v \tag{18.441}$$

et de la même façon,

$$B - O = R(-\alpha)v. \tag{18.442}$$

Notons que tous les O se sont simplifiés et qu'il ne reste que des rotations usuelles. En utilisant le fait que $R(\alpha)$ est une isométrie, nous pouvons alors calculer

$$(D - A) \cdot (B - O) = \langle R(-2\alpha)v - v, R(-\alpha)v \rangle \tag{18.443a}$$
$$= \langle R(-\alpha)v - R(\alpha)v, v \rangle. \tag{18.443b}$$

En utilisant la matrice de rotation du lemme 18.136 nous trouvons

$$\big(R(-\alpha) - R(\alpha)\big)v = \begin{pmatrix} 2\sin(\alpha)v_2 \\ -2\sin(\alpha)v_1 \end{pmatrix} \tag{18.444}$$

et donc

$$\big\langle \big(R(-\alpha) - R(\alpha)\big)v, v \big\rangle = 0. \tag{18.445}$$

Le point D est bien sûr la droite perpendiculaire à ℓ_1 et passant par A. Mais vu que D est obtenu à partir de A par une rotation, le point D est également sur le cercle de rayon $\|OA\|$ et centré en O. Ce cercle possède exactement deux intersections avec cette droite. Le premier est A et le second est $\sigma_{\ell_1}(A)$. Comme D n'est pas A, nous avons $D = \sigma_\ell(A)$. □

18.10.13 Polygone convexe

DEFooBCSKooABKKtR

Définition 18.166 ([552]).
*Un **polygone** est une union de segments dans $\mathbb{R}^2$ de la forme*

$$[a_N, a_1] \cup \bigcup_{i=1}^{N} [a_i, a_{i+1}]. \tag{18.446}$$

La définition suivante dit ce qu'est un polygone convexe. À mon avis, le théorème de Jordan 27.215 doit permettre de donner une définition plus précise en disant que c'est la partie compacte qui est convexe.

44. Définition 18.162.

DEFooBCDPooXTpbFq

Définition 18.167.
Nous disons qu'un polygone est ***convexe*** *si il est la frontière d'une partie convexe de* $\mathbb{R}^2$.

PROPooFYRMooTqVDEm

Proposition 18.168 ([553])**.**
Un polygone est convexe[45] *si et seulement si il est contenu dans le demi-plan de chacun de ses côtés.*

PROPooUPPTooZBFvPg

Proposition 18.169.
Les racines de l'unité dans $\mathbb{C}$*, c'est-à-dire la partie*

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\}, \tag{18.447}$$

forment un polygone régulier convexe[46].

Démonstration. Nous posons $a_k = e^{2ik\pi/n}$ et $\xi = e^{2i\pi/n}$, de telle sorte que $a_{k+1} = \xi a_k$. Le polygone que nous considérons est celui formé par les segments $[a_i, a_{i+1}]$ avec $i = 0, \dots, n-1$. Nous savons que $a_n = a_0$; donc nous pouvons parler seulement de segments de la forme $[a_i, a_{i+1}]$ sans nous soucier du cas particulier $[a_{n-1}, a_0]$.

(1) **Longueurs** La longueur d'une arrête est donnée par

$$|a_k - a_{k+1}| = |a_k - \xi a_k| = |a_k||1 - \xi| = |1 - \xi| \tag{18.448}$$

parce que $|a_k| = 1$.

(2) **Angles** Nous avons un calcul utilisant pas mal de propriétés :

$$\begin{aligned}
\text{angle}(a_l, a_{l+1}, a_{l+2}) &= \text{angle}(\xi^l a_0, \xi^l a_1, \xi^l a_2) && && (18.449a)\\
&= \text{angle}\big(s(\xi^l a_0), s(\xi^l a_1), s(\xi^l a_2)\big) && \text{def. 18.158} && (18.449b)\\
&= \text{angle}\big(R_0(\xi^l)s(a_0), R_0(\xi^l)s(a_1), R_0(\xi^l)s(a_2)\big) && \text{lem. 18.159} && (18.449c)\\
&= \text{angle}\big(s(a_0), s(a_1), s(a_2)\big) && \text{prop. 18.164} && (18.449d)\\
&= \text{angle}(a_0, a_1, a_2). && && (18.449e)
\end{aligned}$$

(3) **Convexe** Voici un squelette de preuve[47].

- Utiliser la caractérisation 18.168.
- Tous les sommets sont sur le cercle unité.
- Considérer d, la droite passant par deux point successifs $e^{i\theta_1}$ et $e^{i\theta_2}$.
- Cette droite définit deux demi-plans.
- La proposition 12.161 devrait permettre de prouver que la partie du cercle entre θ_1 et θ_2 est dans un demi-plan, et le reste dans l'autre demi-plan.
- Il n'y a pas d'autres racines de l'unité entre θ_1 et θ_2. Donc elles sont toutes dans le même demi-plan.
- Les demi-plans étant convexes (lemme 12.155), les arrêtes sont dans le demi-plan.

□

45. Définition 18.167.
46. Polygone, définition 18.166, polygone convexe, définition 18.167. En ce qui concerne les angles, formellement c'est la définition 18.158.
47. Je ne n'ai pas vérifié. Écrivez-moi si vous voyez un problème.

18.10.14 Groupe diédral

18.10.14.1 Définition et générateurs : vue géométrique

DEFooIWZGooAinSOh

Définition 18.170.
*Le **groupe diédral** D_n ($n \geqslant 3$) est le groupe des isométries de $(\mathbb{C}, d)$ laissant invariant l'ensemble*

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\} \tag{18.450}$$

des racines de l'unité.

18.171.
La proposition 18.169 nous permet de dire que le groupe diédral est le groupe des isométries de $\mathbb{R}^2$ laissant invariant un polygone régulier à n côtés. C'est un peu pour cela que nous n'avons défini D_n que pour $n \geqslant 3$; et un peu aussi pour une raison technique qui arrivera quand on parlera des générateurs abstraits. Voir 18.180 et ce qui suit.

LEMooCUVPooMZKnzo

Lemme 18.172.
Nous avons

$$D_n \subset O(2, \mathbb{R}). \tag{18.451}$$

Démonstration. Si $f \in D_n$, alors $f(e^{2ik\pi/n})$ doit être l'un des $e^{2ik'\pi/n}$, et puisque f conserve les longueurs dans $\mathbb{C}$, nous devons avoir

$$1 = d(0, e^{2ik\pi/n}) = d\big(f(0), e^{2ik'\pi/n}\big). \tag{18.452}$$

Donc $f(0)$ est à l'intersection de tous les cercles de rayon 1 centrés en les $e^{2ik\pi/n}$, ce qui montre que $f(0) = 0$ (dès que $n \geqslant 3$). Par conséquent notre étude du groupe diédral ne doit prendre en compte que les isométries vectorielles de $\mathbb{R}^2$. En d'autres termes

$$D_n \subset O(2, \mathbb{R}). \tag{18.453}$$

□

PROPooELOIooVJtuZN

Proposition 18.173 ([554]).
Le groupe D_n contient un sous-groupe cyclique d'ordre 2 et un sous-groupe cyclique d'ordre n.

Démonstration. Nous notons s la conjugaison complexe [48]. C'est un élément d'ordre 2 qui est dans D_n parce que

$$s\big(e^{2ki\pi/n}\big) = e^{2(n-k)i\pi/n}. \tag{18.454}$$

EqSUshknP

De la même façon, la rotations d'angle $2\pi/n$, que l'on note r, agit sur les racines de l'unité et engendre le groupe d'ordre n des rotations d'angle $2k\pi/n$. □

Proposition 18.174 ([554]).
Si s est la conjugaison complexe et r la rotation d'angle $2\pi/n$, alors $(sr)^2 = \mathrm{Id}$.

Démonstration. Si $z^n = 1$, alors

$$(srsr)z = (srs)\big(e^{2i\pi/n}z\big) = (sr)\big(e^{-2i\pi/n}\bar{z}\big) = s(\bar{z}) = z. \tag{18.455}$$

□

PropLDIPoZ

Proposition 18.175 ([554]).
Nous notons s la conjugaison complexe et r la rotation d'angle $2\pi/n$.

ITEMooOEBHooULRmZk

(1) Le groupe diédral D_n est engendré par s et r.

(2) Tous les éléments de D_n s'écrivent sous la forme r^m ou $s \circ r^m$.

48. C'est une réflexion ; la réflexion d'axe $\mathbb{R}$ dans $\mathbb{C}$.

Démonstration. Nous considérons les points $A_0 = 1$ et $A_k = e^{2ki\pi/n}$ avec $k \in \{1, \dots, n-1\}$. Par convention, $A_n = A_0$. L'action des éléments s et r sur ces points est

$$r(A_k) = A_{k+1} \tag{18.456a}$$
$$s(A_k) = A_{n-k}. \tag{18.456b}$$

Cette dernière est l'équation (18.454).

Soit $f \in D_n$. Étant donné que c'est une isométrie de $\mathbb{R}^2$ avec un point fixe (le point 0), f est soit une rotation, soit une réflexion.

Supposons pour commencer qu'un des A_k soit fixé par f. Dans ce cas f a deux points fixes : O et A_k et est donc soit la réflexion d'axe (OA_k), soit l'identité (lemme 18.115). L'identité étant engendrée par s et r, nous supposons que f est la réflexion d'axe (OA_k). Dans ce cas, nous avons $f = s \circ r^{n-2k}$. En effet pour chaque l nous avons

$$(s \circ r^{n-2k})(A_l) = s(A_{l+n-2k}) = A_{n-l+n+2k} = A_{2k-l}. \tag{18.457}$$

Nous voyons que $s \circ r^{n-2k}$ fixe O et A_k, et donc la droite (OA_k). Elle est donc soit la réflexion d'axe (OA_k) soit l'identité. Mais elle n'est pas l'identité parce que elle ne fixe pas tous les A_l.

Nous passons à présent au cas où f ne fixe aucun des A_k.

(1) Supposons que f soit une rotation. Si $f(A_k) = A_m$, alors l'angle de la rotation est

$$\frac{2(m-k)\pi}{n}, \tag{18.458}$$

et donc $f = r^{m-k}$, qui est de la forme demandée.

(2) Supposons à présent que f soit une réflexion d'axe Δ. Cette fois, Δ ne passe par aucun des points A_k, par contre Δ passe par 0. Nous commençons par montrer que Δ doit être la médiatrice d'un des côtés $[A_p, A_{p+1}]$ du polygone. Comme Δ passe par O et n'est aucune des droites (OA_k), cette droite passe par l'intérieur d'un des triangles OA_pA_{p+1} et intersecte donc le côté correspondant.

Notre tâche est de montrer que Δ coupe $[A_p, A_{p+1}]$ en son milieu. Dans ce cas, Δ sera automatiquement perpendiculaire parce que le triangle OA_pA_{p+1} est isocèle en O. Nommons l la longueur des côtés du polygone $P = \Delta \cap [A_p, A_{p+1}]$, $x = d(A_p, P)$ et $\delta = d(A_p, \Delta)$. Vu que f est la symétrie d'axe Δ, nous avons aussi $d\big(f(A_p), \Delta\big) = \delta$ et $d\big(A_p, f(A_p)\big) = 2\delta$. D'autre part, par la définition de la distance, $\delta < x$. Si $x < \frac{l}{2}$, alors $\delta < \frac{l}{2}$ et donc $d\big(A_p, f(A_p)\big) < l$. Or cela est impossible parce que le polygone ne possède aucun sommet à distance plus courte que l de A_p.

De la même manière si $x > \frac{l}{2}$, nous raisonnons avec A_{p+1} pour obtenir une contradiction. Nous en concluons que la seule possibilité est $x = \frac{l}{2}$, et donc $f(A_p) = A_{p+1}$. Montrons alors que $f = s \circ r^{n-2p-1}$. Il faut montrer que c'est une réflexion qui envoie A_p sur A_{p+1}. D'abord c'est une réflexion parce que

$$\det(sr^{n-2p-1}) = \det(s)\det(r^{n-2p-1}) = -1 \tag{18.459}$$

parce que $\det(s) = -1$ alors que $\det(r^k) = 1$ parce que r est une rotation dans $\mathrm{SO}(2)$. Ensuite nous avons

$$s \circ r^{n-2p-1}(A_p) = s(A_{p+n-2p-1}) = s(A_{n-p-1}) = A_{n-(n-p-1)} = A_{p+1}. \tag{18.460}$$

Donc $s \circ r^{n-2p-1}$ est bien une réflexion qui envoie A_p sur A_{p+1}.

□

CorWYITsWW

Corolaire 18.176.
La liste des éléments de D_n est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{18.461}$$

et $|D_n| = 2n$.

Démonstration. Nous savons par la proposition 18.175 que tous les élément de D_n s'écrivent sous la forme r^k ou sr^k. Puisque r est d'ordre n, il ne faut considérer que $k \in \{1, \ldots, n-1\}$. Les éléments $1, r, \ldots, r^{n-1}$ sont tous différents, et sont (pour des raisons de déterminant) tous différents des sr^k. Les isométries sr^k sont toutes différentes entre elles pour essentiellement la même raison :

$$sr^k(A_p) = s(A_{p+k}) = A_{n-p+k} \tag{18.462}$$

donc si $k \neq k'$, $sr^k(A_p) \neq sr^{k'}(A_p)$. La liste des éléments de D_n est donc

$$D_n = \{1, r, \ldots, r^{n-1}, s, sr, \ldots, sr^{n-1}\} \tag{18.463}$$

et donc $|D_n| = 2n$. □

EXooHNYYooUDsKnm

Exemple 18.177.
Nous considérons le carré $ABCD$ dans $\mathbb{R}^2$ et nous cherchons les isométries de $\mathbb{R}^2$ qui laissent le carré invariant. Nous nommons les points comme sur la figure 18.7. La symétrie d'axe vertical est nommée s et la rotation de 90 degrés est notée r.

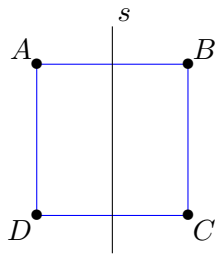


FIGURE 18.7: Le carré dont nous étudions le groupe diédral. LabelFigIsomCarre

Il est facile de vérifier que toutes les symétries axiales peuvent être écrites sous la forme $r^i s$. De plus le groupe engendré par s agit sur le groupe engendré par r parce que

$$(srs^{-1})(A, B, C, D) = sr(B, A, D, C) = s(A, D, C, B) = (B, C, D, A), \tag{18.464}$$

c'est-à-dire $srs^{-1} = r^{-1}$. Nous sommes alors dans le cadre du corolaire 2.49 et nous pouvons écrire que

$$D_4 = \mathrm{gr}(r) \times_\sigma \mathrm{gr}(s). \tag{18.465}$$

△

18.10.14.2 Table de multiplication

La proposition 18.175 nous indique que tous les éléments de D_n s'écrivent sous la forme $s^\epsilon \circ r^m$ avec $\epsilon \in \{0, 1\}$. Nous allons maintenant écrire la table de multiplication pour de telles transformations de $\mathbb{C}$.

LEMooBNJFooAbhsUa

Lemme 18.178.
Si R est une rotation autour de 0 (dans $\mathbb{C}$), et si s est la conjugaison complexe, alors

$$Rs = sR^{-1} \tag{18.466}$$

Démonstration. Il s'agit seulement d'un calcul en écrivant R comme la multiplication par $e^{i\alpha}$. Nous avons

$$(Rs)z = e^{i\alpha}\bar{z} = s\big(e^{-i\alpha}z\big) = sR^{-1}z. \tag{18.467}$$

□

PROPooPYDLooLgiUjk

Proposition 18.179.
Si $\epsilon_1, \epsilon_2 \in \{0, 1\}$ et si $k, l \in \mathbb{Z}$ nous avons

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1+\epsilon_2}r^{l+(-1)^{\epsilon_1}k}. \tag{18.468}$$

Démonstration. Si $\epsilon_2 = 1$ alors nous utilisons le lemme 18.178 pour trouver

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1}(r^k s^{\epsilon_2})r^l = s^{\epsilon_1}s^{\epsilon_2}r^{-k}r^l. \tag{18.469}$$

La proposition est déjà prouvée dans ce cas.

Passons à $\epsilon_2 = 0$. Dans ce cas nous avons

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1}r^{k+l}, \tag{18.470}$$

et c'est bon. □

18.10.14.3 Générateurs : vue abstraite

NORMooCCUEooRRENed

18.180.
Nous allons montrer que D_n peut être décrit de façon abstraite en ne parlant que de ses générateurs. Nous considérons un groupe G engendré par des éléments a et b tels que

(1) a est d'ordre 2,
(2) b est d'ordre n avec $n \geqslant 3$,
(3) $abab = e$.

Nous allons prouver que ce groupe doit avoir la même liste d'éléments que celle du corolaire 18.176.

Proposition 18.181 ([554])**.**
Le groupe G n'est pas abélien.

Démonstration. Nous savons que $abab = e$, donc $abab^{-1} = b^{-2}$, mais $b^{-2} \neq e$ parce que b est d'ordre $n > 2$. Donc $abab^{-1} \neq e$. En manipulant un peu :

$$e \neq abab^{-1} = (ab)(ba^{-1})^{-1} = (ab)(ba)^{-1} \tag{18.471}$$

parce que $a^{-1} = a$. Donc $ab \neq ba$. □

LemKKXdqdL

Lemme 18.182 ([554])**.**
Pour tout k entre 1 et $n-1$ nous avons

$$\boldsymbol{Ad}(a)b^k = ab^k a^{-1} = ab^k a = b^{-k}. \tag{18.472}$$

Démonstration. Nous faisons la démonstration par récurrence. D'abord pour $k = 1$, nous devons avoir $aba = b^{-1}$, ce qui est correct parce que par construction de G nous avons $abab = e$. Ensuite nous supposons que le lemme tient pour k et nous regardons ce qu'il se passe avec $k+1$:

$$ab^{k+1}a = ab^k ba = \underbrace{ab^k a}_{b^{-k}}\underbrace{aba}_{b^{-1}} = b^{-k}b^{-1} = b^{-(k+1)}. \tag{18.473}$$

□

PROPooVQARooWuKHMZ

Proposition 18.183.
L'élément a n'est pas une puissance de b.

Démonstration. Supposons le contraire : $a = b^k$. Dans ce cas nous aurions

$$e = (ab)(ab) = b^{k+1}b^{k+1} = b^{2k+2} = b^{2k}b^2 = a^2b^2 = b^2, \tag{18.474}$$

ce qui signifierait que b est d'ordre 2, ce qui est exclu par construction. □

PROPooEPVGooQjHRJp

Proposition 18.184 ([554])**.**
La liste des éléments de G est donnée par

$$G = \{1, b, \cdots, b^{n-1}, a, ab, \ldots, ab^{n-1}\} = \{a^\epsilon b^k\}_{\substack{\epsilon=0,1\\k=0,\ldots,n-1}} \tag{18.475}$$

Les éléments de ces listes sont distincts.

Démonstration. Étant donné que a n'est pas une puissance de b, les éléments 1, a, b,..., b^{n-1} sont distincts. De plus si k et $m = k + p$ sont deux éléments distincts de $\{1, \dots, n-1\}$, nous avons $ab^k \neq ab^m$ parce que si $ab^k = ab^{k+p}$, alors $a = ab^p$ avec $p < n$, ce qui est impossible. Pour la même raison, $ab^k \neq e$, et $ab^k \neq b^m$.

Au final les éléments $1, a, b, \dots, b^{n-1}, ab, \dots, ab^{n-1}$ sont tous différents. Nous devons encore voir qu'il n'y en a pas d'autres.

Par définition le groupe G est engendré par a et b, donc tout élément $x \in G$ s'écrit $x = a^{m_1}b^{k_1} \dots a^{m_r}b^{k_r}$ pour un certain r et avec pour tout i, $k_i \in \{1, \dots, n-1\}$ (sauf k_r qui peut être égal à zéro) et $m_i = 1$, sauf m_1 qui peut être égal à zéro. Donc

$$x = a^m b^{k_1} a b^{k_2} a \dots b^{k_{r-1}} a b^{k_r} \tag{18.476}$$

où m et k_r peuvent éventuellement être zéro. En utilisant le lemme 18.182 sous la forme $b^{k_i}a = ab^{-k_i}$, quitte à changer les valeurs des exposants, nous pouvons passer tous les a à gauche et tous les b à droite pour finir sous la forme $x = a^k b^m$.

Donc non, il n'existe pas d'autres éléments dans G que ceux déjà listés. □

LemooNFRIooPWuikH

Lemme 18.185 ([1]).
Tout élément de G s'écrit de façon unique sous la forme $a^\epsilon b^k$ ou $b^k a^\epsilon$ avec $\epsilon \in \{0, 1\}$ et $k \in \{0, \dots, n-1\}$.

Démonstration. Nous commençons par la forme $a^\epsilon b^k$. L'existence est la proposition 18.184. Pour l'unicité nous supposons $a^\epsilon b^k = a^\sigma b^l$ et nous décomposons en 4 cas distincts.

(1) $\epsilon = 0, \sigma = 0$ Alors $b^k = b^l$. Mais b étant d'ordre n et k, l étant égaux au maximum à $n-1$, cette égalité implique $k = l$.

(2) $\epsilon = 0, \sigma = 1$ Alors $b^k = ab^l$, ce qui donne $a = b^{k-l}$, ce qui est interdit par la proposition 18.183.

(3) $\epsilon = 1, \sigma = 0$ Même problème que ci-dessus.

(4) $\epsilon = 1, \sigma = 1$ Encore une fois $b^k = b^l$ implique $k = l$.

En ce qui concerne la forme $b^k a^\epsilon$, l'existence est à montrer. Soit l'élément $g = a^\epsilon b^k$; cherchons à le mettre sous la forme $b^l a^\sigma$. Si $\epsilon = 0$ c'est évident. Sinon $\epsilon = 1$ et nous avons par le lemme 18.182

$$ab^k = b^{-k}a^{-1} = b^{-k}b^n a = b^{-k}a. \tag{18.477}$$

En ce qui concerne l'unicité, nous distinguons 4 cas pour $b^k a^\epsilon = b^l a^\sigma$. Comme précédemment ils se traitement exactement comme précédemment. □

THOooYITHooTNTBuG

Théorème 18.186.
Les groupes G et D_n sont isomorphes.

Démonstration. Nous utilisons l'application

$$\begin{aligned} \psi\colon G &\to D_n \\ a^k b^m &\mapsto s^k r^m. \end{aligned} \tag{18.478}$$

C'est évidemment bien défini et bijectif, mais c'est également un morphisme parce que si nous calculons ψ sur un produit, nous devons comparer

$$\psi\big(a^{k_1}b^{m_1}a^{k_2}b^{m_2}\big) \tag{18.479}$$

EqBULPilp

avec

$$\psi\big(a^{k_1}b^{m_1}\big)\psi\big(a^{k_2}b^{m_2}\big) = s^{k_1}r^{m_1}s^{k_2}r^{m_2}. \tag{18.480}$$

EqIVFIphI

Vu que D_n et G ont les mêmes propriétés qui permettent de permuter a et b ou s et r, l'expression à l'intérieur du ψ dans (18.479) se simplifie en $a^k b^m$ avec les même k et m que l'expression à droite dans (18.480) et se simplifie en $s^k r^m$. □

Corolaire 18.187.
Toutes les propriétés démontrées pour G sont vraies pour D_n. En particulier, avec quelques redites :

(1) Le groupe D_n peut être défini comme étant le groupe engendré par un élément s d'ordre 2 et un élément r d'ordre $n-1$ assujettis à la relation $srsr = e$.

(2) Le groupe D_n n'est pas abélien.

(3) Pour tout $k \in \{1, \dots, n-1\}$ nous avons $sr^k s = r^{-k}$.

(4) L'élément s ne peut pas être obtenu comme une puissance de r.

(5) La liste des éléments de D_n est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{18.481}$$

(6) Le groupe diédral D_n est d'ordre $2n$.

Proposition 18.188.
En posant $C_n = \{r^k\}_{k=0,\dots,n-1}$ et $C_2 = \{a^\epsilon\}_{\epsilon=0,1}$, nous pouvons exprimer D_n comme le produit semi-direct

$$D_n = C_n \times_\rho C_2 \tag{18.482}$$

où ρ désigne l'action adjointe.

Démonstration. L'isomorphisme est :

$$\begin{aligned} \psi\colon C_n \times_\rho C_2 &\to D_n \\ (b^k, a^\epsilon) &\mapsto b^k a^\epsilon. \end{aligned} \tag{18.483}$$

(1) **Action adjointe** L'application $\rho_{a^\epsilon} = \mathbf{Ad}(a^\epsilon)$ est toujours un morphisme. Comme a^ϵ est, soit e, soit a, nous allons nous restreindre à a et oublier l'exposant ϵ. Il faut montrer que $\mathbf{Ad}(a) \in \mathrm{Aut}(C_n)$. En utilisant le lemme 18.182,

$$\mathbf{Ad}(a)b^k = ab^k a^{-1} = b^{-k} = b^{n-k}. \tag{18.484}$$

L'application $\mathbf{Ad}(a)\colon C_n \to C_n$ est donc bijective et homomorphique. Ergo isomorphisme.

(2) **Injectif** Si $\psi(b^k, a^\epsilon) = \psi(b^l, a^\sigma)$, alors par unicité du lemme 18.185 nous avons $k = l$ et $\epsilon = \sigma$.

(3) **Surjectif** Par la partie « existence » du lemme 18.185.

(4) **Morphisme** Lorqu'on prend deux sous-groupes d'un même groupe (ici le groupe des isométries de $\mathbb{R}^2$), et que l'on tente de faire un produit semi-direct en utilisant l'action adjointe, nous avons toujours un morphisme. Dans notre cas, le calcul est :

$$\psi\big((b^k, a^\epsilon)(b^l, a^\sigma)\big) = b^k \rho_{a^\epsilon}(b^l) a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^{-\epsilon} a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^\sigma = \psi(b^k, a^\epsilon)\psi(b^l, a^\sigma). \tag{18.485}$$

□

18.10.14.4 Classes de conjugaison

subsubsecZQnBcgo

Pour les classes de conjugaison du groupe diédral nous suivons [555].

D'abord pour des raisons de déterminants[49], les classes des éléments de la forme r^k et de la forme sr^k ne se mélangent pas. Nous notons $C(x)$ la classe de conjugaison de x, et $y \cdot x = yxy^{-1}$.

Les relations que nous allons utiliser sont

$$sr^k s = r^{-k} \tag{18.486a}$$

$$rs = sr^{-1} = sr^{n-1}. \tag{18.486b}$$

49. Vous notez qu'ici nous utilisons un argument qui utilise la définition de D_n comme isométries de $\mathbb{R}^2$. Si nous avions voulu à tout prix nous limiter à la définition « abstraite » en termes de générateurs, il aurait fallu trouver autre chose.

La classe de conjugaison qui ne rate jamais est bien entendu $C(1) = 1$. Nous commençons les vraies festivités avec $C(r^m)$. D'abord $r^k \cdot r^m = r^m$, ensuite

$$(sr^k) \cdot r^m = sr^k r^m r^{-k} s^{-1} = sr^m s^{-1} = r^{-m}. \tag{18.487}$$

Donc

$$C(r^m) = \{r^m, r^{-m}\}. \tag{18.488}$$ EqVFfExgi

À ce niveau il faut faire deux remarques. D'abord si $m > \frac{n}{2}$, alors $C(r^m)$ est la classe de C^{n-m} avec $n - m < \frac{n}{2}$. Donc les classes que nous avons trouvées sont uniquement à lister avec $m < \frac{n}{2}$. Ensuite si $m = \frac{n}{2}$ alors $r^m = r^{-m}$ et la classe est un singleton. Cela n'arrive que si n est pair.

Nous passons ensuite à $C(s)$. Nous avons

$$r^k \cdot s = r^k s r^{-k} = s s r^k s r^{-k} = s r^{-k} r^{-k} = s r^{n-2k}, \tag{18.489}$$

et

$$(sr^k) \cdot s = \underbrace{sr^k s}_{r^{-k}} r^{-k} s^{-1} = r^{-2k} s = r^{n-2k} s = s r^{(n-1)(n-2k)} = s r^{n^2-2kn-n+2k} = s r^{2k}. \tag{18.490}$$

donc

$$C(s) = \{sr^{n-2k}, sr^{2k}\}_{k=0,\dots,n-1}. \tag{18.491}$$

Ici aussi l'écriture n'est pas optimale : peut-être que pour certains k il y a des doublons. Nous reportons l'écriture exacte à la discussion plus bas qui distinguera n pair de n impair. Notons juste que si n est pair, l'élément sr n'est pas dans la classe $C(s)$.

Nous en faisons donc à présent le calcul en gardant en tête le fait qu'il n'a de sens que si n est pair. D'abord

$$s \cdot (sr) = ssrs = rs = sr^{n-1}. \tag{18.492}$$

Ensuite

$$(sr^k) \cdot (sr) = sr^k s r r^{-k} s = r^{-2k+1} s = s r^{2k-1}. \tag{18.493}$$

Avec $k = \frac{n}{2}$, cela rend $s \cdot (sr)$, donc pas besoin de le recopier. Nous avons

$$C(sr) = \{sr^{2k-1}\}_{k=1,\dots,n-1}. \tag{18.494}$$

18.10.14.5 Le compte pour n pair

SubsubsecROVmHuM

Si n est pair, nous avons les classes

$$C(1) = \{1\} \qquad \text{1 élément} \tag{18.495a}$$

$$C(r^m) = \{r^m, r^{-m}\} = \{r^m, r^{n-m}\} \qquad \text{pour } 0 < m < \frac{n}{2} \qquad \frac{n}{2} - 1 \text{ fois 2 éléments} \tag{18.495b}$$

$$C(r^{n/2}) = \{r^{n/2}\} \qquad \text{1 élément} \tag{18.495c}$$

$$C(s) = \{sr^{2k}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments} \tag{18.495d}$$

$$C(sr) = \{sr^{2k+1}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments.} \tag{18.495e}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n}{2} + 3$ classes différentes.

18.10.14.6 Le compte pour n impair

GJIzDEP

Si n est impair, nous avons les classes

$$C(1) = \{1\} \qquad \text{1 élément} \tag{18.496a}$$

$$C(r^m) = \{r^m, r^{-m}\} = \{r^m, r^{n-m}\} \qquad \text{pour } 0 < m < \frac{n-1}{2} \qquad \frac{n-1}{2} \text{ fois 2 éléments} \tag{18.496b}$$

$$C(s) = \{sr^k\}_{k=0,\dots,n-1} \qquad n \text{ éléments} \tag{18.496c}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n+3}{2}$ classes différentes.

18.10.15 Applications : du dénombrement

18.10.15.1 Le jeu de la roulette

pTqJLY

Soit une roulette à n secteurs que nous voulons colorier en q couleurs[556]. Nous voulons savoir le nombre de possibilités à rotations près. Soit d'abord E l'ensemble des coloriages possibles sans contrainte ; il y a naturellement q^n possibilités. Sur l'ensemble E, le groupe cyclique G des rotations d'angle $2\pi/n$ agit. Deux coloriages étant identiques si ils sont reliés par une rotation, la réponse à notre problème est donnée par le nombre d'orbites de l'action de G sur E qui sera donnée par la formule du théorème de Burnside 2.40.

Nous devons calculer $\text{Card}\big(\text{Fix}(g)\big)$ pour tout $g \in G$. Soit g, un élément d'ordre d dans G. Si g agit sur la roulette, chaque secteur a une orbite contenant d éléments. Autrement dit, g divise la roulette en n/d secteurs. Un élément de E appartenant à $\text{Fix}(g)$ doit colorier ces n/d secteurs de façon uniforme ; il y a $q^{n/d}$ possibilités.

Il reste à déterminer le nombre d'éléments d'ordre d dans G. Un élément de G est donné par un nombre complexe de la forme $e^{2ik\pi/n}$. Les éléments d'ordre d sont les racines primitives [50] d-ièmes de l'unité. Nous savons que –par définition– il y a $\varphi(d)$ telles racines primitives de l'unité. Bref il y a $\varphi(d)$ éléments d'ordre d dans G.

La formule de Burnside nous donne maintenant le nombre d'orbites :

$$\frac{1}{n}\sum_{d|n}\varphi(d)q^{n/d}. \tag{18.497}$$

Cela est le nombre de coloriage possibles de la roulette à n secteurs avec q couleurs.

18.10.15.2 L'affaire du collier

siOQlG

Nous avons maintenant des perles de q couleurs différentes et nous voulons en faire un collier à n perles. Cette fois non seulement les rotations donnent des colliers équivalents, mais en outre les symétries axiales (il est possible de retourner un collier, mais pas une roulette). Le groupe agissant sur E est maintenant le groupe diédral [51] D_n conservant un polygone à n sommets.

Nous devons séparer le cas n impair, du cas n pair.

Si n est impair, alors les axes de symétries passent par un sommet et par le milieu du côté opposé. Le groupe D_n contient n symétries axiales. Nous avons donc maintenant

$$|G| = 2n. \tag{18.498}$$

Nous écrivons la formule de Burnside

$$\text{Card}(\Omega) = \frac{1}{2n}\sum_{g\in G}\text{Card}\big(\text{Fix}(g)\big). \tag{18.499}$$

Si g est une rotation, le travail est déjà fait. Si g est une symétrie, nous avons le choix de la couleur du sommet par lequel passe l'axe et le choix de la couleur des $(n-1)/2$ paires de sommets. Cela fait

$$qq^{(n-1)/2} = q^{\frac{n+1}{2}} \tag{18.500}$$

possibilités. Nous avons donc

$$\text{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d|n}q^{n/d}\varphi(d) + nq^{\frac{n+1}{2}}\right). \tag{18.501}$$

Si n est pair, les choses se compliquent un tout petit peu. En plus de symétries axiales passant par les milieux de deux côtés opposés, il y a les axes passant par deux sommets opposés. Pour

50. Une racine non primitive 8ième de l'unité est par exemple i. Certes $i^8 = 1$, mais $i^4 = 1$ aussi. Le nombre i est d'ordre 4.

51. Définition 18.170.

colorier un collier en tenant compte d'une telle symétrie, nous pouvons choisir la couleur des deux perles par lesquelles passe l'axe ainsi que la couleur des $(n-2)/2$ paires de perles. Cela fait en tout

$$q^2 q^{\frac{n-2}{2}} = q^{\frac{n+2}{2}}. \tag{18.502}$$

Le groupe G contient $n/2$ tels axes.

Notons que cette fois G ne contient plus que $n/2$ symétries passant par un sommet et un côté. L'ordre de G est donc encore $2n$. La formule de Burnside donne

$$\mathrm{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d|n}\varphi(d)q^{n/d} + \frac{n}{2}q^{(n+2)/2} + \frac{n}{2}q^{n/2}\right). \tag{18.503}$$

18.10.16 Classification

THOooRORQooTDWFdv

Théorème 18.189 ([545]).
Toute isométrie du plan $(\mathbb{R}^2, d)$ est une composition d'au plus 3 réflexions.

Démonstration. Encore une fois nous décomposons la preuve en fonction du nombre de points fixes.

(1) **Si f n'a pas de point fixe** Soit $x \in \mathbb{R}^2$. Nous considérons le segment $[x, f(x)]$ et nous nommons l sa médiatrice. Par construction, $f(x) = \sigma_l(x)$. Nous posons $g = \sigma_l \circ f$, et nous avons

$$g(x) = x. \tag{18.504}$$

Donc nous avons $f = \sigma_l \circ g$ avec $x \in \mathrm{Fix}(g)$.

(2) **Si f a un unique point fixe** Soit x cet unique point fixe. Soit $y \neq x$ et l la médiatrice de $[y, f(y)]$. En posant $g = \sigma_l \circ f$ nous avons

$$g(y) = y \tag{18.505}$$

et $g(x) = x$ parce que

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y), \tag{18.506}$$

ce qui donne que x est à égale distance de y et de $f(y)$, c'est-à-dire que $x \in l$ et par conséquent $g(x) = (\sigma_l \circ f)(x) = \sigma_l(x) = x$.
Donc g fixe x et y. Par conséquent, g fixe toute la droite (xy).

(3) **Si f fixe une droite** Soit l une droite fixée par f. Alors f est soit σ_l soit l'identité par le lemme 18.115.

(4) **Conclusion**
Nous avons montré que si $\mathrm{Fix}(f)$ a pour dimension m, alors il existe une droite pour laquelle $f = \sigma_l \circ g$ avec $\dim\big(\mathrm{Fix}(g)\big) > m$. Donc il faut au maximum trois pas pour avoir $\dim\big(\mathrm{Fix}(g)\big) = 2$ c'est-à-dire pour avoir $g = \mathrm{Id}$.

□

DEFooJEOYooNwYtuQ

Définition 18.190.
*Une **réflexion glissée** est une transformation du plan de la forme $\tau_v \circ \sigma_\ell$ où le vecteur v est parallèle à la droite ℓ.*

THOooVRNOooAgaVRN

Théorème 18.191 ([545]).
Les isométries du plan $(\mathbb{R}^2, d)$ sont exactement

(1) l'identité (composée de 0 réflexions),

(2) les réflexions,

(3) les translations (composées de 2 réflexions d'axes parallèles),

(4) les rotations (composées de 2 réflexions d'axes non parallèles),

(5) les réflexions glissées (composées de 3 réflexions)

Démonstration. Nous savons déjà que $f \in \mathrm{Isom}(\mathbb{R}^2)$ est une composée de 0, 1, 2 ou 3 réflexions.

(1) **Zéro réflexion** Alors c'est l'identité. Ce n'est pas très profond.

(2) **Une réflexion** Alors f est une réflexion. Toujours pas très profond.

(3) **Deux réflexions** Soit $f = \sigma_{\ell_1} \circ \sigma_{\ell_2}$. Maintenant ça s'approfondit un bon coup.
Nous supposons d'abord que $\ell_1 \parallel \ell_2$. Dans ce cas nous allons prouver que $f = \tau_{2v}$ où v est le vecteur perpendiculaire à ℓ_1 tel que $\ell_1 + v = \ell_2$. Nous allons utiliser le lemme 18.114 pour montrer que $\sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_{2v}$. Nous avons

$$\ell_1 = \ell_0 + w \tag{18.507a}$$
$$\ell_2 = \ell_0 + w + v \tag{18.507b}$$

où w est un vecteur perpendiculaire à ℓ_1 et ℓ_0 est la droite passant par l'origine et parallèle à ℓ_1 et ℓ_2. Avec cela,

$$\begin{aligned}(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) &= \sigma_{\ell_1}\big(\sigma_{\ell_0}(x) + 2w\big) && (18.508a)\\ &= \sigma_{\ell_0}\big(\sigma_{\ell_0}(x) + 2w\big) + 2(v + w) && (18.508b)\\ &= x + \underbrace{\sigma_{\ell_0}(2w)}_{-2w} + 2v + 2w && (18.508c)\\ &= x + 2v. && (18.508d)\end{aligned}$$

Donc si f est composée de deux réflexions d'axes parallèles, alors f est une translation.
Toujours dans le cas où f est composée de deux réflexions, nous supposons que $f = \sigma_{\ell_2} \circ \sigma_{\ell_1}$ avec ℓ_1 et ℓ_2 non parallèles. Nous notons O le point d'intersection, et nous allons voir que $f = R_O(2\alpha)$, où α est l'angle de ℓ_1 à ℓ_2 donné par le lemme 18.162.
Soit $x \in \ell_1$. Alors

$$f(x) = \sigma_{\ell_2}(x), \tag{18.509}$$

et le lemme 18.165 nous donne un moyen de calculer $\sigma_{\ell_2}(x)$ parce que ℓ_1 est une droite passant par x et coupant ℓ_2 au point O. Le lemme dit que $\sigma_{\ell_2}(x) = R_O(2\alpha)$. Remarque : c'est bien 2α et non -2α parce qu'il s'agit de l'angle de ℓ_2 à ℓ_1 ; il y a inversion des numéros entre ici et l'énoncé du lemme.
Nous avons donc bien $f(x) = R_O(2\alpha)x$ pour $x \in \ell_1$.
Si $y \in \ell_2$ alors

$$f(y) = \sigma_{\ell_2}\big(R_O(-2\alpha)y\big) \tag{18.510}$$

Nous posons $z = \sigma_{\ell_1}(y) = R_O(-2\alpha)y$. Soit la droite ℓ_3 passant par O et z. Puisque $R_O(2\alpha)z = y \in \ell_2$, l'angle de ℓ_3 à ℓ_2 est 2α. Par conséquent

$$\sigma_{\ell_2}(z) = R_O\big(-2 \times (-2\alpha)\big)z = R_O(4\alpha)z = R_O(4\alpha)R_O(-2\alpha)y = R_O(2\alpha)y. \tag{18.511}$$

Donc les transformations f et $R_O(2\alpha)$ coïncident pour tous les points des droites ℓ_1 et ℓ_2, qui ne sont pas parallèles. Cela prouve que $f = R_O(2\alpha)$.

(4) **Trois réflexions** Nous écrivons $f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1}$. Nous allons transformer cela progressivement en une symétrie glissée en passant par plusieurs étapes : ITEMooHVYCooPhFMiv

(4a) $f = \sigma_\ell \circ \tau_v$, ITEMooUKGLooFlCcjt

(4b) $f = \tau_v \circ \sigma_\ell$, ITEMooWUCWooZSjofe

(4c) $f = \tau_v \circ \sigma_\ell$ avec $v \parallel \ell$.

À chacune de ces étapes, v et ℓ vont changer. La dernière est une réflexion glissée.
Nous commençons par supposer $\ell_2 \parallel \ell_3$. Dans ce cas, $\sigma_{\ell_3} \circ \sigma_{\ell_2}$ est une translation, comme nous l'avons déjà vu. Alors $f = \tau_v \circ \sigma_{\ell_1}$ et nous sommes déjà dans le cas (4b).

Nous supposons que ℓ_2 n'est pas parallèle à ℓ_3. Dans ce cas, si $O = \ell_2 \cap \ell_3$ nous avons

$$\sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha) \tag{18.512}$$

où α est l'angle de ℓ_2 à ℓ_3. En réalité tant que l'angle de ℓ_2' à ℓ_3' est α nous avons

$$\sigma_{\ell_3'} \circ \sigma_{\ell_2'} = \sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha). \tag{18.513}$$

Nous choisissons ℓ_2' parallèle à ℓ_1, de telle sorte que $\sigma_{\ell_2'} \circ \sigma_{\ell_1}$ soit une translation. Alors nous avons

$$f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1} = \sigma_{\ell_3'} \circ \sigma_{\ell_2'} \circ \sigma_{\ell_1} = \sigma_{\ell_3} \circ \tau_v. \tag{18.514}$$

où v est le vecteur de la translation en question.

Nous avons donc prouvé que toute composition de trois réflexions peut être écrite soit sous la forme (4a) soit sous la forme (4b).

Nous prouvons à présent que toute transformation de la forme (4a) peut être écrite sous la forme (4b). Plus précisément nous allons prouver que si ℓ est une droite, v un vecteur et ℓ_0 la droite parallèle à ℓ passant par l'origine, alors

$$\sigma_\ell \circ \tau_v = \tau_{\sigma_{\ell_0}(v)} \circ \sigma_l \tag{18.515}$$

D'abord nous savons que $\sigma_\ell(x) = \sigma_{\ell_0}(x) + 2w$ où w est le vecteur tel que $\ell = \ell_0 + w$. Ensuite c'est un simple calcul utilisant le fait que σ_{ℓ_0} est linéaire :

$$(\sigma_\ell \circ \tau_v)(x) = \sigma_l(x + v) = \sigma_{\ell_0}(x) + \sigma_{\ell_0}(v) + 2w, \tag{18.516}$$

et

$$(\tau_{\sigma_{\ell_0}(v)} \circ \sigma_\ell)(x) = \sigma_{\ell_0}(v) + \sigma_\ell(x) = \sigma_{\ell_0}(v) + \sigma_{\ell_0}(x) + 2w. \tag{18.517}$$

L'égalité est prouvée.

Nous montrons maintenant que toute transformation de la forme (4b) peut être mise sous la forme (4c). Soit donc $f = \tau_v \circ \sigma_\ell$ où v et ℓ ne sont pas spécialement parallèles.

Pour cela nous décomposons $v = v_1 + v_2$ avec $v_1 \perp \ell$ et $v_2 \parallel \ell$ et nous posons $\ell' = \ell + \frac{1}{2}v_1$. Nous montrons que

— $\tau_v \circ \sigma_\ell = \tau_{v_2} \circ \sigma_{\ell'}$

— $v_2 \parallel \ell'$.

Pour le deuxième point, $v_2 \parallel \ell$ et bien entendu $\ell' \parallel \ell$. Donc $v_2 \parallel \ell'$.

Soit ℓ_0 la droite parallèle à ℓ et ℓ' et passant par l'origine. Soit aussi le vecteur w tel que $\ell = \ell_0 + w$. Alors nous avons

$$\begin{cases} \sigma_\ell = \sigma_{\ell_0} + 2w & \text{(18.518a)} \\ \sigma_{\ell'} = \sigma_{\ell_0} + 2w + v_1 & \text{(18.518b)} \end{cases}$$

Nous avons

$$(\tau_v \circ \sigma_\ell)(x) = v + \sigma_{\ell_0}(x) + 2w \tag{18.519}$$

et

$$\begin{aligned} (\tau_{v_2} \circ \sigma_{\ell'})(x) &= v_2 + \sigma_{\ell_0}(x) + 2w + v_1 && \text{(18.520a)} \\ &= \sigma_{\ell_0}(x) + v + 2w && \text{(18.520b)} \end{aligned}$$

où dans la dernière ligne, nous avons regroupé $v_1 + v_2 = v$. Et voilà.

□

18.10.17 Classification des isométries de $\mathbb{R}$

Définition 18.192.
*Soit $x \in \mathbb{R}$; nous notons σ_x la **réflexion** par rapport à x, c'est-à-dire*

$$\sigma_x(y) = 2x - y. \tag{18.521}$$

Théorème 18.193 ([545]).
Toute isométrie de $\mathbb{R}$ est composée d'au plus 2 réflexions. Plus précisément toute isométrie de $\mathbb{R}$ est dans une des trois catégories suivantes :

- *l'identité (0 réflexion),*
- *les réflexions,*
- *les translations (2 réflexions)*

Démonstration. Nous divisons la preuve en fonction du nombre de points fixés par l'isométrie $f \in \mathrm{Isom}(\mathbb{R})$.

(1) **f fixe deux points distincts** Alors elle fixe l'espace affine engendré par ces deux points par la proposition 18.101. Donc f fixe tout $\mathbb{R}$, et est l'identité.

(2) **f fixe un unique point** Soit x l'unique point fixé par f et considérons $y \neq x$. Puisque $x = f(x)$ et que f est une isométrie,

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y). \tag{18.522}$$

Donc $f(y)$ est à égale distance de x que de y. Autrement dit, $f(y)$ est, soit y, soit $\sigma_x(y)$. Mais comme x est unique point fixe, $f(y) = \sigma_x(y)$. Ce raisonnement étant valable pour tout $y \neq x$, nous avons $f = \sigma_x$.

(3) **f n'a pas de point fixe** Soient $x \in \mathbb{R}$ et $y = \frac{x+f(x)}{2}$. Nous posons $g = \sigma_y \circ f$. Alors x est un point fixe de g parce que

$$g(x) = \sigma_y\big(f(x)\big) = 2y - f(x) = x. \tag{18.523}$$

Donc, soit g est l'identité, soit g est une réflexion (par les points précédents). La possibilité $g = \mathrm{Id}$ est exclue parce que cela donnerait $f = \sigma_y$ alors que f n'a pas de point fixe. Donc g est une réflexion ; et comme x est un point fixe de g nous avons $g = \sigma_x$. Au final

$$f = \sigma_y \circ \sigma_x. \tag{18.524}$$

Montrons que cela implique que f est une translation :

$$\sigma_y\sigma_x(z) = \sigma_y(2x - z) = 2y - 2x + z = z + 2(y - x). \tag{18.525}$$

Donc $\sigma_y \circ \sigma_x$ est la translation de vecteur $2(y - x)$.

□

18.10.18 Isométries du tétraèdre régulier

PROPooHTBIooTJQNZt

Proposition 18.194.
Soient un tétraèdre régulier[52] T ainsi qu'une application $f : \mathbb{R}^3 \to \mathbb{R}^3$ qui vérifie

(1) f est affine[53]

(2) f est bijective (sur $\mathbb{R}^3$)

(3) $f(T) \subset T$.

Alors

52. Définition 12.147
53. Définition 8.160.

ITEMooGAVPooVPESod

(1) *Il existe une bijection* $\alpha\colon \{1,2,3,4,5,6\} \to \{1,2,3,4,5,6\}$ *telle que* $f(d_i) = d_{\alpha(i)}$.

ITEMooWRINooTlmXyI

(2) *Pour la même bijection* $\alpha\colon \{1,2,3,4,5,6\} \to \{1,2,3,4,5,6\}$ *nous avons* $f(a_i) = a_{\alpha(i)}$.

ITEMooGPNPooZiEoLk

(3) *Il existe une bijection* $\beta\colon \{1,2,3,4\} \to \{1,2,3,4\}$ *telle que* $f(s_i) = s_{\beta(i)}$.

Autrement dit, f *agit comme une permutation*[54] *sur l'ensemble des sommets de* T.

Démonstration. Puisque d_i est une droite et que f est affine, l'ensemble $f(d_i)$ est une droite de $\mathbb{R}^3$ (proposition 12.137).

(1) **Pour (1)** Puisque $f(T) \subset T$, la partie $f(a_i)$ contient une infinité de points alignés dans T (parce que a_i contient une infinité de points et f est une bijection de $\mathbb{R}^3$).
Comme le tétraèdre n'a que 6 arrêtes, il y a forcément une des arrêtes qui contient au moins deux points de $f(a_i)$. Soit $a_{\beta(i)}$ une arrête qui contient deux points de $f(a_i)$. En particulier $d_{\alpha(i)}$ contient deux points de $f(a_i)$, et donc de $f(d_i)$. Donc $f(d_i) = d_{\alpha(i)}$.

(2) **Pour (2)** L'ensemble $f(a_i)$ est contenu dans $d_{\alpha(i)} \cap T = a_{\alpha(i)}$ (lemme 12.149). Nommons s_1 et s_2 les deux sommets de T sur a_i. Soit $k \in \{1,2\}$. Comme les sommets sont sur plusieurs arrêtes, il existe $j \neq i$ tel que $s_k = a_i \cap a_j$. Nous avons alors

EQooWJXRooAHVQpW
$$f(s_k) \in f(a_i) \cap f(a_j) \subset a_{\alpha(i)} \cap a_{\alpha(j)}. \tag{18.526}$$

Mais nous avons déjà vu que les seules intersections des segments sont les sommets (lemme 12.151). Donc $f(s_k)$ est un sommet de T.
Nous avons prouvé que $f(a_i)$ contient deux sommets sur $a_{\alpha(i)}$ donc il contient tout $a_{\alpha(i)}$. Cela prouve que $f(a_i) = f(a_{\alpha(i)})$.

(3) **Pour (3)** Nous avons déjà mentionné, juste en-dessous de (18.526) que l'image d'un sommet doit être un sommet. Nous avons donc une application $\beta\colon \{1,2,3,4\} \to \{1,2,3,4\}$ telle que $f(s_i) = s_{\beta(i)}$. Puisque f est injective, β est injective. Par principe des tiroirs, cette application doit également être surjective.

□

PROPooVNLKooOjQzCj

Proposition 18.195 (Isométries affines du tétraèdre régulier)**.**
Soient T *un tétraèdre régulier et* $\mathrm{Isom}(T)$ *son groupe d'isométries affines*[55]*. Alors*

$$\mathrm{Isom}(T) \simeq S_4 \tag{18.527}$$

où S_4 *est le groupe des permutations de quatre objets.*

Démonstration. Si $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ est une bijection affine (ce qui est le cas d'une isométrie affine), alors la proposition 18.194(3) donne une bijection $\beta\colon \{1,2,3,4\} \to \{1,2,3,4\}$ des sommets de T. Si nous numérotons les sommets $x_1,\dots,x_4$, nous obtenons un morphisme de groupe $\varphi\colon \mathrm{Isom}(T) \to S_4$ qui envoie g sur la permutation qui envoie 1 sur le numéro du sommet $g(x_1)$, 2 sur le numéro du sommet $g(x_2)$, etc.

(1) **Le morphisme φ est injectif** Supposons $\varphi(g_1) = \varphi(g_2)$. Alors $g_1^{-1} \circ g_2$ est une isométrie de $(\mathbb{R}^3, d)$ qui fixe les quatre sommets. Une application affine $\mathbb{R}^3 \to \mathbb{R}^3$ fixant 4 point est l'identité par la proposition 8.162. Donc $g_1^{-1} \circ g_2 = \mathrm{Id}$, ce qui prouve que $g_1 = g_2$. Vous noterez que nous utilisons l'unicité de l'inverse dans un groupe.

(2) **φ est surjectif**
Nous savons que S_4 est engendré par les transpositions (proposition 1.295). Or les transpositions sont dans l'image de φ. En effet, notons les sommets de notre tétraèdre par A, B, C et D et considérons la transposition $A \leftrightarrow B$. Elle est l'image par φ de la réflexion selon le plan σ, médiateur du segment $[A,B]$. Pour nous assurer de cela, nous devons nous assurer que C et D appartiennent à σ. C'est le contenu du lemme 18.111.

54. Une permutation est une bijection, définition 1.274.
55. Une isométrie est une isométrie, et une application est affine dans la définition 8.160.

(3) **Conclusion** L'application φ est un morphisme bijectif, c'est-à-dire un isomorphisme.

□

18.10.19 Représentation de S_4 via les isométries du tétraèdre

SUBSECooVEASooDUbsBh

18.196.
Lorsque le tétraèdre a son barycentre en l'origine de $\mathbb{R}^3$, l'isomorphisme $\varphi\colon \mathrm{Isom}(T) \to S_4$ donne une représentation de dimension 3 de S_4. Nous avons calculé les caractères de S_4 en la section 16.5 sans avoir besoin de savoir que l'une des représentations de dimension 3 est celle que nous venons de trouver via le groupe des isométries du tétraèdre. Nous allons cependant également y calculer les caractères de la représentation φ, pour le sport.

Une des représentations trouvées (la représentation ρ_s) peut être vue comme le groupe $\mathrm{Isom}(T)$ des isométries affine du tétraèdre grâce à la proposition 18.195 qui donne un isomorphisme de groupe $S_4 \simeq \mathrm{Iso}(T)$ lorsque T est un tétraèdre régulier de $\mathbb{R}^3$.

Si le barycentre de T est situé à l'origine de $\mathbb{R}^3$, alors les éléments de $\mathrm{Isom}(T)$ sont des applications linéaires parce que

— les affinités laissent invariantes les barycentres (proposition 11.43),

— les affinités qui laissent l'origine invariante sont linéaires (corolaire 11.60).

Nous allons à présent calculer la trace de cette représentation, en utilisant le fait que nous la connaissions explicitement. Nous savons que les caractères sont constants sur les classes de conjugaison ; nous allons donc écrire une matrice par classe de conjugaison (qui sont données dans l'exemple 1.292).

Pour tout cela nous allons considérer un tétraèdre dont l'isobarycentre est en $(0,0,0)$ et une base de $\mathbb{R}^3$ formée de trois sommets e_1, e_2 et e_3. Puisque l'isobarycentre des quatre sommets est en $(0,0,0)$, le quatrième sommet est forcément le point de coordonnées $e_4(-1,-1,-1)$, de telle sorte que $e_1 + e_2 + e_3 + e_4 = 0$.

Les transpositions Quelle isométrie de $\mathbb{R}^3$ permute deux sommets du tétraèdre sans bouger les autres ? Pour permuter les sommets e_1 et e_2 en laissant e_3 et e_4, c'est le symétrie par rapport au plan médiateur de $[e_1, e_2]$. Ce plan passe par les sommets e_3 et e_4, parce que le tétraèdre étant régulier, les points e_3 et e_4 sont équidistants de e_1 et e_2. Le lemme 18.111 dit qu'alors, ces points font partie du plan médiateur.

Dans notre base, la matrice de la transposition précédemment nommée (12) est

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{18.528}$$

dont la trace est 1. Donc $\chi_s(12) = 1$.

Les bitranspositions La bitransposition $(12)(34)$ est le produit des transpositions selon les plans médiateurs de $[e_1, e_2]$ et $[e_3, e_4]$. Ces deux plans sont perpendiculaires, et l'intersection est la droite qui passe par les milieux. Cette droite est perpendiculaire aux deux segments en même temps. La matrice est :

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \tag{18.529}$$

parce que $e_1 \mapsto e_2$, $e_2 \mapsto e_1$ et $e_3 \mapsto e_4$. Pour rappel, la matrice est formée des images des vecteurs de base. Cela donne

$$\chi_s\big((12)(34)\big) = -1. \tag{18.530}$$

Les 3-cycles La symétrie qui permute cycliquement les points e_1, e_2 et e_3 est la rotation d'angle[56] $2\pi/3$ dans le plan formé par les extrémités de ces trois vecteurs. Heureusement, la trace est

56. Angle d'une rotation, définition 18.127.

invariante par changement de base ; donc nous pouvons calculer la trace d'une rotation d'angle $2\pi/3$ dans n'importe quelle base. Par exemple :

$$\chi_s\big((12)(34)\big) = \operatorname{Tr}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & \sin(2\pi/3) \\ 0 & -\sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix} = 1 + 2\cos(2\pi/3) = 0. \tag{18.531}$$

Notons que, sans cette interprétation géométrique, nous y arrivons aussi facilement : dans notre base le 3-cycle est $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_1$, donc la matrice est :

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{18.532}$$

dont la trace est manifestement nulle : $\chi_s\big((123)\big) = 0$.

Le 4-cycle Il réalise $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto e_1$, dont la matrice est

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \tag{18.533}$$

EQooONDUooYlduup

et la trace est $\chi_s\big((1,2,3,4)\big) = -1$.

Nous avons retrouvé les caractères de la représentation ρ_s, et nous pouvons vérifier qu'elle est irréductible.

18.11 Transformations de Lorentz

Nous considérons dans cette section un nombre réel $c > 0$ ainsi que l'espace $\mathbb{R}^2$ muni du produit pseudo-scalaire[57] donné par la matrice

$$\eta = \begin{pmatrix} c^2 & 0 \\ 0 & -1 \end{pmatrix}. \tag{18.534}$$

Et pour faire plus vrai, nous notons (x_0, x_1) les coordonnées sur $\mathbb{R}^2$. Ainsi

$$x \cdot y = c^2 x_0 y_0 - x_1 y_1. \tag{18.535}$$

Nous insistons sur le fait que cela n'est pas un produit scalaire.

LEMooPZPZooVAdPVj

Lemme 18.197 ([1]).
Soit $c > 0$. L'application

$$\begin{aligned} \varphi\colon]-c, c[&\to \mathbb{R} \\ v &\mapsto \frac{-v/c}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned} \tag{18.536}$$

est une bijection.

Démonstration. Nous commençons par mentionner le fait que φ est continue du fait que le dénominateur ne s'annule pas. Une petite étude de fonction montre que

$$\lim_{v \to -c} \varphi(v) = +\infty, \tag{18.537}$$

et

$$\lim_{v \to c} \varphi(v) = -\infty, \tag{18.538}$$

57. Définition 8.169.

et

$$\varphi'(v) = -\frac{1}{c\sqrt{1-\frac{v^2}{c^2}}} - \frac{v^2/c^3}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} < 0. \tag{18.539}$$

Tout cela fait que φ est bijective (entre autres par le théorème des valeurs intermédiaires 9.93 et le théorème dérivée et croissance 12.185). □

LEMooUZFKooSIjery

Lemme 18.198.
La forme bilinéaire

$$\begin{aligned} b\colon \mathbb{R}^2 \times \mathbb{R}^2 &\to \mathbb{R} \\ x, y &\mapsto x \cdot y \end{aligned} \tag{18.540}$$

est non dégénérée[58].

Démonstration. Soit $(x_0, x_1) \in \mathbb{R}^2$ tel que

$$b\big((x_0, x_1), (y_0, y_1)\big) = 0 \tag{18.541}$$

pour tout $(y_0, y_1) \in \mathbb{R}^2$. Nous avons

$$c^2 x_0 y_0 - x_1 y_1 = 0. \tag{18.542}$$

En écrivant cela avec $(y_0, y_1) = (1, 0)$ puis $(0, 1)$ nous obtenons immédiatement que $(x_0, x_1) = (0, 0)$. □

THOooYHDWooWxVovH

Théorème 18.199.
Soit une bijection[59] $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ *telle que*

$$f(x) \cdot f(y) = x \cdot y \tag{18.543}$$

pour tout $x, y \in \mathbb{R}^2$. Alors :

(1) f est linéaire.

(2) Il existe un unique choix de $(\xi, \sigma_1, \sigma_2) \in \mathbb{R} \times \{\pm 1\} \times \{\pm 1\}$ tel que la matrice de f ait la forme

$$f = \begin{pmatrix} \sigma_1 \cosh(\xi) & \frac{\sigma_1\sigma_2}{c}\sinh(\xi) \\ c\sinh(\xi) & \sigma_2\cosh(\xi) \end{pmatrix}. \tag{18.544}$$

(3) Il existe un unique $v \in]-c, c[$ tel que la matrice de f ait la forme

$$f = \begin{pmatrix} \frac{\sigma_1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\sigma_1\sigma_2}{c^2}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{-v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{\sigma_2}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix}. \tag{18.545}$$

Démonstration. Puisque notre produit pseudo-scalaire est non dégénéré (lemme 18.198), le fait que f soit linéaire est la proposition 8.157. Nous posons

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{18.546}$$

et, conformément à la proposition 8.163, nous imposons $A^t \eta A = \eta$. Après un petit produit matriciel nous obtenons :

$$\begin{pmatrix} c^2\alpha^2 - \gamma^2 & c^2\alpha\beta - \gamma\delta \\ c^2\alpha\beta - \gamma\delta & c^2\beta^2 - \delta^2 \end{pmatrix} = \begin{pmatrix} c^2 & 0 \\ 0 & -1 \end{pmatrix}. \tag{18.547}$$

58. Définition 8.129.
59. À mon avis, il y a moyen d'affaiblir cette hypothèse. Écrivez-moi si vous avez une idée.

Voilà quatre équations à résoudre pour les quatre inconnues $\alpha, \beta, \gamma, \delta$. Déjà les équations des termes anti-diagonaux sont les mêmes. Nous recopions le reste :

$$\begin{cases} c^2\alpha^2 - \gamma^2 = c^2 & \text{(18.548a)} \\ c^2\alpha\beta - \gamma\delta = 0 & \text{(18.548b)} \\ c^2\beta^2 - \delta^2 = 1. & \text{(18.548c)} \end{cases}$$

C'est le moment d'utiliser la proposition 15.128. La relation (18.548a) donne

$$\alpha^2 - \left(\frac{\gamma}{c}\right)^2 = 1, \tag{18.549}$$

ce qui implique l'existence d'un unique [60] $\xi_1 \in \mathbb{R}$ et $\sigma_1 \in \{\pm 1\}$ tels que

$$\gamma = c\sinh(\xi_1) \tag{18.550a}$$
$$\alpha = \sigma_1 \cosh(\xi_1). \tag{18.550b}$$

La relation (18.548c) implique quant à elle l'existence de $\xi_2 \in \mathbb{R}$ et $\sigma_2 \in \{\pm 1\}$ tels que

$$\delta = \sigma_2 \cosh(\xi_2) \tag{18.551a}$$
$$\beta = \frac{1}{c}\sinh(\xi_2). \tag{18.551b}$$

Nous substituons maintenant toutes les valeurs (18.550) et (18.551) dans (18.548b). Cela donne

$$\sigma_1 \cosh(\xi_1)\sinh(\xi_2) = \sigma_2 \sinh(\xi_1)\cosh(\xi_2). \tag{18.552}$$

Nous mettons cette relation au carré et nous substituons $\cosh(\xi_1)^2 = 1 + \sinh^2(\xi_1)$. Ce que nous trouvons est

$$\sinh(\xi_1)^2 = \sinh(\xi_2)^2, \tag{18.553}$$

qui implique que $\xi_1 = \pm\xi_2$. Nous posons donc $\xi_2 = \sigma_3\xi_1$ pour un certain $\sigma_3 \in \{\pm 1\}$. Cela nous permet d'alléger la notation et d'écrire ξ au lieu de ξ_1.

Nous remettons la valeur $\xi = \xi_1 = \sigma_3\xi_2$ dans l'équation (18.552) en tenant compte du fait que sinh est impaire et cosh est paire :

$$\sigma_1\sigma_3 \cosh(\xi)\sinh(\xi) = \sigma_2 \sinh(\xi)\cosh(\xi). \tag{18.554}$$

Et cela nous enseigne que $\sigma_3 = \sigma_1\sigma_2$.

Jusqu'à présent nous avons prouvé qu'il existe un unique $\xi \in \mathbb{R}$ et $\sigma_1, \sigma_2 \in \{\pm 1\}$ tels que

$$A = \begin{pmatrix} \sigma_1\cosh(\xi) & \frac{\sigma_1\sigma_2}{c}\sinh(\xi) \\ c\sinh(\xi) & \sigma_2\cosh(\xi) \end{pmatrix}. \tag{18.555}$$

Nous utilisons à présent la bijection du lemme 18.197. Il existe un unique $v \in]-c, c[$ tel que $\sinh(\xi) = \varphi(v)$. En utilisant $\cosh(\xi)^2 = 1 + \varphi(v)^2$, nous trouvons

$$\cosh(\xi)^2 = \frac{1}{1 - \frac{v^2}{c^2}}. \tag{18.556}$$

Mais comme le cosinus hyperbolique est toujours strictement positif, nous pouvons prendre la racine carrée des deux côtés :

$$\cosh(\xi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{18.557}$$

En substituant dans (18.555), nous trouvons le résultat annoncé. □

60. Par 15.128.

18.11.1 Sous-groupe fini d'isométries du plan

THOooKDMUooUxQqbB

Théorème 18.200 ([557]).
Soit un groupe fini G d'isométries de $(\mathbb{R}^2, d)$ contenant n éléments.

ITEMooYEONooCOMpeb

(1) Il existe un point $C \in \mathbb{R}^2$ fixé par tous les éléments de G.

ITEMooGELWooFFAqkc

(2) Si G ne contient pas de réflexion, alors il est cyclique[61] *et engendré par la rotation d'angle $2\pi/n$ autour de C.*

ITEMooDHKEooFpCfmX

(3) Si G contient au moins une réflexion, et si C est un point fixe de G, alors

ITEMooGQZTooJIPPLtyf

(3a) toutes les réflexions ont un axe qui passe par C,

ITEMooKPQRooLquSiQ

(3b) n est pair,

ITEMooCHSWooHpDGHf

(3c) Si σ est une réflexion dans G, alors nous avons $G = \mathrm{gr}\big(\sigma, R_C(4\pi/n)\big)$ où $R_C(\theta)$ est la rotation d'angle θ autour de C,

ITEMooROUYooRghvMv

(3d) G est isomorphe au groupe diédral $D_{n/2}$.

Démonstration. Soit un groupe fini G constitué d'isométries de $(\mathbb{R}^2, d)$. Nous prouvons le théorème point par point.

(1) **Pour (1)** C'est la proposition 18.98.

(2) **Questions de réflexions** Le théorème 18.189 nous dit que les éléments de G sont des compositions d'au maximum 3 réflexions.

(3) **Exclure trois réflexions** Il n'est pas possible que G contienne un élément composé de trois réflexions. En effet, les composées de trois réflexions, par le théorème 18.191 sont des réflexions glissées[62], c'est-à-dire des transformations de la forme $g = \tau_v \circ \sigma_\ell$ où v est un vecteur parallèle à la droite ℓ. Si $x \in \ell$, alors

$$g(x) = \tau_v(x) = x + v, \tag{18.558}$$

de telle sorte que $g^k(x) = x + kv$, qui signifie que tous les g^k sont différents. Le groupe G ne peut pas être fini si il contient une réflexion glissée.

(4) **G^+ et G^-** Pour la même raison que celle qui exclut les réflexions glissées, G ne peut pas contenir de translation. Le théorème 18.191 nous donne la liste des possibilités. Après exclusion des translations et des réflexions glissées, il reste :

- l'identité,
- les rotations,
- les réflexions.

Nous notons G^+ la partie de G contenant l'identité et les rotations et G^- celle contenant les réflexions. Notons que G^+ n'est pas vide parce qu'il contient au moins l'identité, tandis que G^- peut être vide, mais n'est certainement pas un groupe.

(5) **Même nombre d'éléments** Nous prouvons à présent que si G^- est non vide, alors il a le même nombre d'éléments que G^+. Un élément de G^- est une réflexion. Soit $\sigma \in G^-$. Nous prouvons que

$$\begin{aligned} \varphi\colon G^+ &\to G^- \\ f &\mapsto \sigma \circ f \end{aligned} \tag{18.559}$$

EQooWRVVooBQCtPg

est une bijection.

(5a) **Surjective** Soit $s \in G^-$. Posons $f = \sigma^{-1} \circ s$. Puisque σ^{-1} et s sont des réflexions, f est une rotation. Donc $f \in G^+$ et $\varphi(f) = s$.

61. Définition 1.326.
62. Définition 18.190.

(5b) **Injective** La condition $\varphi(f) = \varphi(g)$ dit que $\sigma \circ f = \sigma \circ g$. En composant par σ^{-1} nous obtenons $f = g$.

(6) $G^+ = \text{gr}\big(R_C(2\pi/p)\big)$ Nous nommons p le nombre d'éléments de G^+. Si G^- est vide, $p = n$, et sinon $p = n/2$. Dans les deux cas, G^+ est un groupe de rotations à p éléments.

Le groupe G^+ contient seulement des rotations ; or le centre d'une rotation est l'unique point fixe. Donc tous les éléments de G^+ sont des rotations autour de C.

Le corolaire 2.14 au théorème de Lagrange nous indique que tous les éléments de G^+ vérifient $g^p = \text{Id}$. Seules les rotations d'angle $2k\pi/p$ autour de C satisfont la condition $g^p = \text{Id}$. Or il n'y a que p telles rotations. Donc elles sont toutes dans G^+. Nous en déduisons que

$$G^+ = \text{gr}\big(R_C(2\pi/p)\big). \tag{18.560}$$

EQooUWTVooEMgkVH

(7) **Pour (2)** Dans le cas où G ne contient pas de réflexion, G^- est vide et G contient n éléments. La relation (18.560) devient

$$G = G^+ = \text{gr}\big(R_C(2\pi/n)\big). \tag{18.561}$$

(8) **Pour (3)** Nous supposons maintenant que G contienne au moins une réflexion. De la sorte $G^- \neq \varnothing$.

(8a) **Pour (3a)** Les seuls points fixes d'une réflexion sont ceux de l'axe. Donc C doit être sur tous les axes des réflexions contenues dans G^-.

Notons au passage que deux réflexions d'axes qui se coupent forment une rotation. Donc G^- ne forme pas un groupe, mais même pas en rêve, quoi.

(8b) **Pour (3b)** Vu que l'union $G = G^+ \cup G^-$ est disjointe et que G^+ et G^- ont le même nombre d'éléments par la bijection 18.559, si G^- est non vide, G possède un nombre pair d'éléments.

(8c) **Pour (3c)** Si $\sigma \in G$ est une réflexion, nous savons que n est pair, que G^+ possède $p = n/2$ éléments et que

$$G^+ = \{R_C(2k\pi/p)\} = \{R_C(4k\pi/n)\}_{k=1,\dots,n/2}. \tag{18.562}$$

L'élément $\sigma \in G^-$ étant fixé, la bijection (18.559) nous indique que tous les éléments de G^- sont de la forme $\sigma \circ f$ avec $f \in G^+$. Donc

$$G^- \subset \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{18.563}$$

Nous avons aussi

$$G^+ \subset \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{18.564}$$

Et comme σ et $R_C(4\pi/n)$ sont dans G nous avons $\text{gr}\big(\sigma, R_C(4\pi/n)\big) \subset G$. Tout cela pour dire que

$$G = \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{18.565}$$

(8d) $R\sigma = \sigma R^{-1}$ Nous restons dans le cas où G^- n'est pas vide. Nous considérons R, la rotation d'angle θ autour de C. Si R_0 est la rotation d'angle θ autour de $(0,0)$, nous avons

$$R = \tau_C \circ R_0 \circ \tau_C^{-1}, \tag{18.566}$$

et si σ_0 est la symétrie d'axe parallèle à l'axe de σ, mais passant par $(0,0)$ nous avons :

$$\sigma = \tau_C \circ \sigma_0 \circ \tau_C^{-1}. \tag{18.567}$$

Si v est le vecteur directeur de la réflexion σ_0, nous considérons enfin α, la rotation qui réalise $\alpha(v) = (1,0)$. Nous avons alors

$$\sigma_0 = \alpha^{-1} \circ s \circ \alpha \tag{18.568}$$

où s est la symétrie autour de l'axe horizontal. En n'ayant pas peur d'identifier $\mathbb{R}^2$ à $\mathbb{C}$, l'application s est la conjugaison complexe. Avec tout ça nous avons

$$R\sigma = \tau_C R_0 \tau_C^{-1} \tau_C \sigma_0 \tau_C^{-1} = \tau_C R_0 \alpha^{-1} s \alpha \tau_C^{-1} = \tau_C \alpha^{-1} R_0 s \alpha \tau_C^{-1} \tag{18.569}$$

où nous avons utilisé le fait que les rotations autour de $(0,0)$ forment un groupe abélien pour commuter α^{-1} avec R_0. Nous utilisons à présent le lemme 18.178 pour commuter R avec s :

$$\begin{aligned} R\sigma &= \tau_C \alpha^{-1} s R_0^{-1} \alpha \tau_C^{-1} && (18.570a)\\ &= \tau_C \underbrace{\alpha^{-1} s \alpha}_{\sigma_0} R_0^{-1} \tau_C^{-1} && (18.570b)\\ &= \tau_C \sigma_0 \tau_C^{-1} \tau_C R_0^{-1} \tau_C^{-1} && (18.570c)\\ &= \sigma R^{-1}. && (18.570d) \end{aligned}$$

Nous avons utilisé le fait que $\tau_C R_0^{-1} \tau_C^{-1} = R^{-1}$ comme on peut s'en convaincre en calculant le produit.

(8e) **Table de multiplication**
Nous considérons une réflexion $\sigma \in G$. Les éléments de G^+ sont des rotations autour de C et ceux de G^- de la forme σR où R est une rotation autour de C. Pour connaître la table de multiplication de G, nous devons écrire

$$(\sigma^{\epsilon_1} R^k)(\sigma^{\epsilon_2} R^l) = \sigma^{\epsilon} R^m \tag{18.571}$$

où $\epsilon_1, \epsilon_2 \in \{0,1\}$, R est la rotation d'angle $4\pi/n$ autour de C et ϵ et m sont des constantes à exprimer en fonction de ϵ_1, ϵ_2, k et l.
Tous les éléments de G pouvant être écrits soit sous la forme R^m, soit sous la forme σR^m, nous avons les possibilités suivantes :

(8e.v) $R^m R^l = R^{m+l}$
(8e.v) $(R^m)(\sigma R^l) = \sigma R^{-m} R^l = \sigma R^{l-m}$
(8e.v) $(\sigma R^m) R^l = \sigma R^{m+l}$
(8e.v) $(\sigma R^m)(\sigma R^l) = \sigma\sigma R^{l-m} = R^{l-m}$.

(8f) **Pour (3d)** Récoltons quelque acquis.
— Nous venons de prouver que $R\sigma = \sigma R^{-1}$.
— Tout élément de G peut s'écrire soit sous la forme R^m, soit sous la forme σR^m, selon que l'élément est dans G^+ ou G^-.
— Tout élément du groupe diédral D_n s'écrit soit sous la forme r^m, soit sous la forme sr^m (proposition 18.175(2)).

L'application $\varphi\colon G \to D_n$ suivante est donc une bijection :

$$\begin{cases} \varphi(R^m) = r^m & (18.572a)\\ \varphi(\sigma R^m) = sr^m. & (18.572b) \end{cases}$$

Il nous reste à prouver que c'est un morphisme. Cela se fait en utilisant la table de multiplication du groupe diédral donnée dans la proposition 18.179 et celle du groupe G que nous venons d'écrire.

□

Définition 18.201 (Groupe de symétrie d'une partie de $\mathbb{R}^n$[545])**.**
Si Y est une partie de $\mathbb{R}^n$, nous définissons le ***groupe des symétries*** *de Y par*

$$\mathrm{Sym}(Y) = \{f \in \mathrm{Isom}(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \tag{18.573}$$

Nous définissons aussi le groupe des symétries propres de Y par

$$\mathrm{Sym}^+(Y) = \{f \in \mathrm{Isom}^+(\mathbb{R}^n) \text{ tel que } f(Y) = Y\} \tag{18.574}$$

où $\mathrm{Isom}^+(\mathbb{R}^n)$ est le groupe des isométries positives de $\mathbb{R}^n$, définition 18.92.

THOooAYZVooPmCiWI

Théorème 18.202 ([545]).
Soit $Y \subset \mathbb{R}^2$.

(1) Si le groupe $\mathrm{Sym}^+(Y)$ est fini d'ordre n, alors c'est un groupe cyclique d'ordre n.

(2) Si $\mathrm{Sym}^+(Y)$ est fini, alors $\mathrm{Sym}(Y)$ est soit cyclique[63] d'ordre n, soit isomorphe au groupe diédral[64] d'ordre $2n$.

Démonstration. Nous savons déjà par la proposition 18.99 que $\mathrm{Sym}^+(Y)$ est isomorphe à un sous-groupe H^+ d'ordre n de $\mathrm{SO}(2)$. Vérifions que ce groupe est cyclique. Si $n = 1$, c'est évident. Si $n \geqslant 2$ alors nous savons que H^+ est constitué de rotations d'angles dans $[0, 2\pi[$ et vu que c'est un ensemble fini, il possède une rotation d'angle minimal (à part zéro). Notons α_0 cet angle.

Nous montrons que H^+ est engendré par la rotation d'angle α_0. Soit une rotation d'angle α. Étant donné que $\alpha_0 < \alpha$ nous pouvons effectuer la division euclidienne[65] de α par α_0 et obtenir

$$\alpha = k\alpha_0 + \beta \tag{18.575}$$

avec $\beta < \alpha_0$. Mézalors $R(\beta) = R(\alpha)R(\alpha_0)^{-k}$ est également un élément du groupe. Cela contredit la minimalité dès que $\beta \neq 0$. Avoir $\beta = 0$ revient à dire que α est un multiple de α_0, ce qui signifie que le groupe H^+ est cyclique engendré par α_0.

Notons au passage que nous avons automatiquement $\alpha_0 = \frac{2\pi}{n}$ parce qu'il faut $R(\alpha_0)^n = \mathrm{Id}$. Nous avons prouvé que $\mathrm{Sym}^+(Y)$ est cyclique d'ordre n.

Nous étudions maintenant le groupe $\mathrm{Sym}(Y)$. Par la proposition 18.99 nous avons un morphisme injectif

$$\phi\colon \mathrm{Sym}(Y) \to \mathrm{O}(2), \tag{18.576}$$

et en posant $H = \phi\big(\mathrm{Sym}(Y)\big)$ nous avons un isomorphisme de groupes $\phi\colon \mathrm{Sym}(Y) \to H$. Nous savons aussi que ce ϕ se restreint en

$$\phi\colon \mathrm{Sym}^+(Y) \to H^+ \subset \mathrm{SO}(2) \tag{18.577}$$

où $H^+ = \phi\big(\mathrm{Sym}^+(Y)\big) = H \cap \mathrm{SO}(2)$. Le groupe H^+ est cyclique et est engendré par la rotation $R(2\pi/n)$.

Supposons un instant que $H \subset \mathrm{SO}(2)$. Alors nous avons $H = H^+$ et ϕ est un isomorphisme entre $\mathrm{Sym}(Y)$ et le groupe cyclique engendré par $R(2\pi/n)$.

Nous supposons à présent que H n'est pas un sous-ensemble de $\mathrm{SO}(2)$. Quelles sont les isométries de $\mathbb{R}^2$ qui ne sont pas de déterminant 1 ? Il faut regarder dans le théorème 18.191 quelles sont les isométries contenant un nombre impair de réflexions. Ce sont les réflexions et les réflexions glissées. Or il ne peut pas y avoir de réflexion glissée dans un groupe fini, parce que si f est une réflexion glissée, tous les f^k sont différents.

Nous en déduisons que si H n'est pas inclus dans $\mathrm{SO}(2)$, il contient une réflexion que nous nommons σ. Nous allons en déduire que $H \simeq H^+ \times_{\mathbf{Ad}} C_2$ où $C_2 = \{\mathrm{Id}, \sigma\}$. Si $h \in H$ nous pouvons écrire

$$h = (h\sigma^\epsilon)\sigma^\epsilon \tag{18.578}$$

pour n'importe quelle valeur de ϵ, et en particulier pour $\epsilon = \pm 1$.

Si $h \in \mathrm{SO}(2)$ alors nous écrivons $h = h\sigma^0$ et si $h \notin \mathrm{SO}(2)$ nous écrivons $h = (h\sigma)\sigma$. Vu que $h\sigma \in \mathrm{SO}(2)$, cette dernière écriture est encore de la forme $\mathrm{SO}(2) \times C_2$. Quoi qu'il en soit, tout élément de H s'écrit comme un produit

$$H = H^+C_2. \tag{18.579}$$

Cette décomposition est unique parce que si $h_1c_1 = h_2c_2$ alors $h_2^{-1}h_1 = c_2c_1^{-1}$, et comme $h_2^{-1}h_1 \in H^+$ nous avons $c_2c_1^{-1} \in H^+$ et donc $c_1 = c_2$. Partant, nous avons aussi $h_1 = h_2$. Pour avoir

63. Définition 1.326.
64. Définition 18.170.
65. Théorème 1.222.

le produit semi-direct, il faut encore montrer que $\mathbf{Ad}(C_2)H^+ \subset H^+$. Le seul cas à vérifier est $\mathbf{Ad}(\sigma)H^+ \subset H^+$. Comme les éléments de H^+ sont caractérisés par le fait d'avoir un déterminant positif, nous avons

$$\mathbf{Ad}(\sigma)R(\alpha) = \sigma R(\alpha)\sigma^{-1} \in H^+. \tag{18.580}$$

□

Remarque 18.203.
Tout ceci est cohérent avec le théorème de Burnside 8.313 parce que le sous-groupe fini de $\mathrm{SO}(n)$ engendré par la rotation $R(2\pi/n)$ est un groupe d'exposant fini, à savoir que si h est dans ce groupe, $h^n = \mathrm{Id}$.

18.11.2 Relations trigonométriques dans un triangle rectangle

Nous donnons maintenant quelques relations trigonométriques classiques dans un triangle rectangle. Le théorème de Pythagore est déjà le théorème 10.25 ; nous nous concentrons ici sur les angles.

PROPooCDZVooKOQzct

Proposition 18.204.
Soient $A, B, S \in \mathbb{R}^2$ des points distincts et non alignés formant un triangle rectangle en A :

$$(A - S) \cdot (B - A) = 0. \tag{18.581}$$

En posant $\theta = \widehat{ASB}$ nous avons

$$\cos(\theta) = \frac{\|A - S\|}{\|B - S\|} \tag{18.582}$$

et

$$\sin(\theta) = \pm\frac{\|B - A\|}{\|B - S\|}. \tag{18.583}$$

EQooEKZEooFeNImX

Démonstration. Nous posons $C = A - S$ et $D = B - S$. Comme $C \neq 0$, il existe une rotation R_α telle que

$$\begin{cases} (R_\alpha C)_x > 0 & \text{(18.584a)} \\ (R_\alpha C)_y = 0. & \text{(18.584b)} \end{cases}$$

Nous posons $X = R_\alpha C$ et $Y = R_\alpha D$.

Le triangle formé de O, X et Y est « posé » sur l'axe des abscisses et est rectangle en X, c'est-à-dire

$$X \cdot (Y - X) = 0. \tag{18.585}$$

De ce fait, le point Y satisfait à $Y_x = X_x$. Et enfin, grâce aux propositions 18.163 et 18.164 nous avons $\widehat{ASB} = \widehat{XOY}$.

Nous écrivons les relations qui définissent l'angle $\widehat{XOY}$. Pour cela nous posons $X' = X/\|X\|$ et $Y' = Y/\|Y\|$ et nous avons

SUBEQooVHNDooPOfbjC

$$\begin{cases} \cos(\theta) = X'_x Y'_x & \text{(18.586a)} \\ \sin(\theta) = X'_y Y'_y. & \text{(18.586b)} \end{cases}$$

Vu que $X = (X_x, 0)$, nous avons $X'_x = 1$. De plus

$$\|Y\| = \|R_\alpha(D)\| = \|D\| = \|B - S\|. \tag{18.587}$$

En substituant les valeurs dans (18.586),

$$\cos(\theta) = Y'_x = \frac{Y_x}{\|Y\|} = \frac{X_x}{\|B - S\|} = \frac{\|C\|}{\|B - S\|} = \frac{\|A - S\|}{\|B - S\|}. \tag{18.588}$$

Voilà déjà une chose de prouvée.

Pour la seconde, nous avons $\sin(\theta) = Y'_y$. Selon le signe de Y_y nous avons $Y_y = \pm\|Y - X\|$ et donc

$$\sin(\theta) = Y'_y = \frac{Y_y}{\|Y\|} = \frac{\pm\|Y - X\|}{\|Y\|} = \pm\frac{\|D - C\|}{\|Y\|} = \pm\frac{\|B - A\|}{\|B - S\|}. \tag{18.589}$$

□

Le signe sur la formule du sinus revient au fait que la définition de l'angle $\widehat{AOB}$ est de considérer la rotation qui fait aller A vers B. Donc suivant la position relative de A, O et B, il se peut que l'angle mesuré soit l'angle *extérieur* au triangle.

La proposition suivante est parfois prise comme définition de l'angle.

Proposition 18.205.
Soient trois points non alignés $A, S, B \in \mathbb{R}^2$. Nous avons

$$\cos\left(\widehat{ASB}\right) = \frac{(A - S) \cdot (B - S)}{\|A - S\|\|B - S\|}. \tag{18.590}$$

EQooOWULooVQntxE

Démonstration. Nous posons $C = A - S$, $D = B - S$, $X = C/\|C\|$ et $Y = D/\|D\|$. Avec cela, la définition 18.157 donne l'équation

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} X_x \\ X_y \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix} \tag{18.591}$$

que nous écrivons comme le système

$$\begin{cases} X_x \cos(\theta) - X_y \sin(\theta) = Y_x & \text{(18.592a)} \\ X_y \cos(\theta) + X_x \sin(\theta) = Y_y. & \text{(18.592b)} \end{cases}$$

Nous considérons maintenant cela comme un système pour $\big(\cos(\theta), \sin(\theta)\big)$:

$$\begin{pmatrix} X_x & -X_y \\ X_y & X_x \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \tag{18.593}$$

Le déterminant de la dernière matrice est $X_x + X_y^2 = \|X\|^2 = 1$ parce que X est unitaire. Cette matrice est donc inversible et son inverse est vite calculée. Nous avons

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} X_x & X_y \\ -X_y & X_x \end{pmatrix} \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \tag{18.594}$$

Cela donne ce que nous voulions :

$$\cos(\theta) = X \cdot Y = \frac{C \cdot D}{\|C\|\|D\|} = \frac{(A - S) \cdot (B - S)}{\|A - S\|\|B - S\|}. \tag{18.595}$$

□

Remarque 18.206.
Prendre la formule (18.590)

$$\cos\left(\widehat{ASB}\right) = \frac{(A - S) \cdot (B - S)}{\|A - S\|\|B - S\|}. \tag{18.596}$$

comme définition de l'angle $\widehat{ASB}$ est cependant trompeur parce que ça ne permet de définir les angles que sur une partie de $[0, 2\pi[$ sur laquelle le cosinus est injectif. Pour réellement définir tous les angles, il faut alors un peu bricoler.

Sans vouloir être méchant, je crois que ceux qui procèdent de cette manière sont ceux qui donnent un cours sur le produit scalaire sans avoir l'intention de lier la définition d'une rotation comme composée de réflexions aux matrices de $\mathrm{SO}(2)$ et aux fonctions trigonométriques.

18.11.3 Pavages du plan

DEFooHPKFooSIDhCM

Définition 18.207.
*Une application affine $f\colon \mathbb{R}^n \to \mathbb{R}^n$ est un **déplacement** lorsqu'elle est une isométrie de $(\mathbb{R}^n, d)$ qui préserve l'orientation*[66].

DEFooJPHKooRgCBJs

Définition 18.208 ([430]).
*Un **pavage** de $\mathbb{R}^2$ est une paire (G, K) où G est un groupe de déplacements*[67] *de $\mathbb{R}^2$ et K un compact de $\mathbb{R}^2$ d'intérieur non vide telle que*

(1) $G \cdot K = \mathbb{R}^2$,

ITEMooOIJZooZMKLUm

(2) Si $g_1, g_2 \in G$ satisfont $g_1 \cdot \mathrm{Int}(K) \cap g_2 \cdot \mathrm{Int}(K) \neq \emptyset$, alors $g_1 \cdot K = g_2 \cdot K$.

*Nous disons qu'un groupe G de déplacements de $\mathbb{R}^2$ est un **groupe de pavage** de $\mathbb{R}^2$ si il existe un compact K tel que la paire (G, K) soit un pavage.*

En termes de notations,

$$G \cdot K = \bigcup_{g\in G} g(K) = \bigcup_{g\in G} \bigcup_{k\in K} g(k). \tag{18.597}$$

LEMooWZSWooZYkICn

Lemme 18.209 ([1]).
Soient une bijection affine $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ ainsi qu'une droite d et un point A. Nous notons S la partie de $\mathbb{R}^2$ située du côté de d contenant A.

Alors $\varphi(S)$ est la partie de $\mathbb{R}^2$ située du côté de $\varphi(d)$ contenant $\varphi(A)$.

Démonstration. La droite d est donnée par une application affine $f\colon \mathbb{R}^2 \to \mathbb{R}$ et la définition

$$d = \{x \in \mathbb{R}^2 \text{ tel que } f(x) = 0\}. \tag{18.598}$$

Nous supposons que $f(A) > 0$; sinon, nous pouvons utiliser $-f$ au lieu de f. Donc

$$S = \{x \in \mathbb{R}^2 \text{ tel que } f(x) > 0\}. \tag{18.599}$$

La partie $\varphi(S)$ est alors donnée par

$$\varphi(S) = \{\varphi(x) \text{ tel que } f(x) > 0\}. \tag{18.600}$$

Comme φ est une bijection, cela s'écrit aussi bien

$$\varphi(S) = \{y \in \mathbb{R}^2 \text{ tel que } f\big(\varphi^{-1}(y)\big) > 0\}. \tag{18.601}$$

De même

$$\varphi(d) = \{s \in \mathbb{R}^2 \text{ tel que } f\big(\varphi^{-1}(s)\big) = 0\}. \tag{18.602}$$

Donc les deux côtés de la droite $\varphi(d)$ sont donnés par le signe de $f \circ \varphi^{-1}$. Nous avons

$$(f \circ \varphi^{-1})\big(\varphi(A)\big) = f(A) > 0. \tag{18.603}$$

Donc $\varphi(A) \in \varphi(S)$. □

LEMooZOXVooTJiLTF

Lemme 18.210.
Le groupe

$$G = \mathrm{gr}(\tau_{e_1}, \tau_{e_2}) \tag{18.604}$$

est un groupe de pavage du plan.

66. Définition 8.35.
67. Définition 18.207.

Démonstration. Il suffit de prendre le carré $K = [0,1] \times [0,1]$. En appliquant les translations, nous recouvrons tout le plan, sans intersection des intérieurs des carrés. Notons toutefois qu'il y a un recouvrement des bords. □

LEMooTMRGooChBzZg

Lemme 18.211 ([1, 558])**.**
Le groupe

$$G = \mathrm{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi)) \tag{18.605}$$

est un groupe de pavage du plan.

Démonstration. Le compact à considérer est $K = [0, \frac{1}{2}] \times [0,1]$. Le compact K et son image par $R_0(\pi)$ sont représentés sur la figure 18.8.

En agissant sur K avec les translations verticales et horizontales, nous recouvrons des bandes verticales de largeur 1/2. En agissant de même sur $R_0(\pi)(K)$, nous recouvrons les autres bandes verticales.

Donc $G \cdot K$ recouvre bien $\mathbb{R}^2$. Il serait cependant un peu présomptueux de croire en avoir fini. Il faut vérifier la condition (2) de la définition 18.208 d'un pavage.

Supposons que $g_1 \cdot \mathrm{Int}(K) \cap g_2 \cdot \mathrm{Int}(K) \neq \varnothing$. Cela signifie qu'il existe $k_1, k_2 \in \mathrm{Int}(K)$ tels que $g_1(k_1) = g_2(k_2)$, ou encore que

$$(g_2^{-1} g_1) k_1 = k_2 \in \mathrm{Int}(K). \tag{18.606}$$

EQooITJEooUkUKuu

Quelle est la forme d'un élément général de G ? Le lemme 1.323 nous indique qu'un élément général de G est un produit fini de τ_{e_1}, τ_{e_2} et $R_0(\pi)$. Mais nous savons que si α est linéaire,

$$\alpha \circ \tau_u = \tau_{\alpha(u)} \circ \alpha. \tag{18.607}$$

Dans notre cas, dans un produit général, nous pouvons déplacer tous les facteurs $R_0(\pi)$ à droite en changeant des τ_{e_i} en $\tau_{R_0(\pi)e_i} = \tau_{-e_i}$. Les translations par contre commutent sans faire d'histoires. Donc un élément général de G est de la forme

$$g = \tau_{e_1}^k \tau_{e_2}^l R_0(\pi)^m \tag{18.608}$$

Nous pouvons évidemment restreindre m à $\{0,1\}$. Supposons $k \in \mathrm{Int}(K)$ et $g(k) \in \mathrm{Int}(K)$. Nous avons $0 < k_x < 0.5$. Si $m = 1$, alors $g(k)_x \in]-1/2, 0[$ et aucun $\tau_{e_1}^k g(k)_x$ ne pourra plus être entre 0 et 1/2. Donc $m = 0$. À partir de là, pour avoir $g(k) \in \mathrm{Int}(K)$ nous devons avoir également $k = l = 0$. Donc $g = \mathrm{Id}$.

Deux éléments g_1 et g_2 vérifiant la condition (18.606) doivent donc vérifier $g_2^{-1} g_1 = \mathrm{Id}$, et donc $g_1 = g_2$. Par conséquent $g_1 \cdot K = g_2 \cdot K$. □

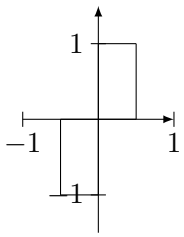


FIGURE 18.8: Le compact K et son image par $R_0(\pi)$ pour le lemme 18.211.

LabelFigATJSooefYkmCbP

LEMooJPNDooHDCLnY

Lemme 18.212 ([1, 558])**.**
Le groupe

$$\mathrm{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi/2)) \tag{18.609}$$

est un groupe de pavage.

Démonstration. Le pavé à considérer est

$$K = [-\frac{1}{2}, 0] \times [0, \frac{1}{2}]. \tag{18.610}$$

En lui appliquant trois fois la rotation $R_0(\pi/2)$, nous reconstituons le carré $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$. Ensuite, avec les translations, nous pavons tout le plan.

Pour la seconde condition, nous procédons comme dans la démonstration du lemme 18.211. D'abord $R_0(\pi/2)e_1 = e_2$ et $R_0(\pi/2)e_2 = -e_1$. Donc dans un produit général de τ_{e_1}, τ_{e_2} et $R_0(\pi/2)$ (et de leurs inverses), toutes les rotations peuvent être mises à droite ; nous avons donc un élément général de G sous la forme

$$g = \tau_a \circ R_0(\pi/2)^k \tag{18.611}$$

avec $a \in \mathbb{Z}e_1 + \mathbb{Z}e_2$ et $k \in \{0, 1, 2, 3\}$.

Vu que les translations se font par nombres entiers tandis que les différences de coordonnées entre les $R_0(\pi/2)^k K$ sont demi-entiers, si $k \in \text{Int}(K)$, alors aucun τ_a ne permet d'avoir $(\tau_a \circ R_0(\pi/2))k \in \text{Int}(K)$.

Bref, si $(g_2^{-1}g_1)k \in K$, alors encore une fois $g_2^{-1}g_1 = \text{Id}$. □

Et c'est maintenant que les choses compliquées commencent.

LEMooMWWEooEbZXtb

Lemme 18.213.
Le groupe

$$\text{gr}\left(\tau_{e_1}, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3)\right) \tag{18.612}$$

est un groupe de pavage.

Démonstration. Ici le compact K est le triangle de sommets $A = (0, 0)$, $B = (\frac{1}{2}, \frac{\sqrt{3}}{6})$ et $C = (1, 0)$. Plus précisément il s'agit de l'intersection des trois parties suivantes :

— le côté de la droite (AB) où est C,
— le côté de la droite (AC) où est B,
— le côté de la droite (BC) où est A.

Il est bon d'écrire ces trois conditions sous forme d'inéquations.

— La droite AB est donnée par l'équation $f_{AB}(x, y) = 0$ pour $f_{AB}(x, y) = x - \sqrt{3}y$. Puisque $f_{AB}(C) = 1$, la première inéquation pour K est

$$x - \sqrt{3}y \geqslant 0. \tag{18.613}$$

— Pour la droite (BC) nous avons $f_{BC}(x, y) = -x - \sqrt{3}y + 1$ et $f_{BC}(A) = 1$. Donc la seconde inéquation pour K est

$$-x - \sqrt{3}y + 1 \geqslant 0 \tag{18.614}$$

— La droite AC est donnée par l'application $f_{AC}(x, y) = y$. Vu que $f_{AC}(B) = \sqrt{3}/6$, nous avons la troisième inéquation pour K :

$$y \geqslant 0. \tag{18.615}$$

En résumé, la définition de K est le système SUBEQSooECKFooOdneOA

$$\begin{cases} x - \sqrt{3}y \geqslant 0 & \text{(18.616a)} \\ -x - \sqrt{3}y + 1 \geqslant 0 & \text{(18.616b)} \\ y \geqslant 0. & \text{(18.616c)} \end{cases}$$

La figure 18.9(a) nous montre ce triangle et l'action des puissances de $R_0(\pi/3)$ sur lui. Pour votre gouverne, la matrice de cette rotation est

$$R_0(\pi/3) = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}. \tag{18.617}$$ EQooMRBXooTebLJt

La figure 18.9(c) vous montre une partie de ce pavage dans toute sa splendeur.

Puisque les translations sont $u_1 = e_1$ et $u_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$, il est suffisant de montrer que notre pavage pave réellement le parallélogramme construit sur ces deux vecteurs. Nous vous avons très obligeamment dessiné ce parallélogramme pavé sur la figure 18.9(b).

Pour montrer que les triangles dessinés pavent effectivement le parallélogramme, nous allons procéder à une exhaustion de cas. Soit (x, y) dans le parallélogramme.

Nous commençons par couper le parallélogramme en deux parties suivant la diagonale allant de u_1 à u_2. Dans la vie mon p'ti gars, il y a deux types de points : ceux qui vérifient $x + \frac{1}{\sqrt{3}}y - 1 \leqslant 0$ et les autres.

(1) **Si** $x + y/\sqrt{3} - 1 \leqslant 0$ Ceci est le côté de $(0, 0)$. Nous subdivisons suivant la petite barre verticale (suivez le dessin), c'est-à-dire suivant les deux cas : $x \geqslant \frac{1}{2}$ et $x \leqslant \frac{1}{2}$.

(1a) **Si** $x \leqslant \frac{1}{2}$ Enfin nous coupons avec la droite diagonale partant de $(0, 0)$, c'est-à-dire selon que $x - \sqrt{3}y \leqslant 0$ ou $x - \sqrt{3}y \geqslant 0$.

(1a.i) **Si** $x - \sqrt{3}y \leqslant 0$ Les points dont nous parlons sont les $(x, y) \in \mathbb{R}^2$ vérifiant [SUBEQSooNYWDooYMNV]

$$\begin{cases} x + \dfrac{1}{\sqrt{3}}y - 1 \leqslant 0 & (18.618a) \\ x \leqslant \dfrac{1}{2} & (18.618b) \\ x - \sqrt{3}y \leqslant 0. & (18.618c) \end{cases}$$

En suivant le dessin, vous remarquerez que l'élément de G à considérer est

$$g = \tau_{e_1} \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})} \circ R_0(\pi/3)^4. \tag{18.619}$$

Nous avons [EQSooOJBFooCTaTtu]

$$g(A) = (\frac{1}{2}, \frac{\sqrt{3}}{2}) \tag{18.620a}$$
$$g(B) = (\frac{1}{2}, \frac{\sqrt{3}}{6}) \tag{18.620b}$$
$$g(C) = (0, 0). \tag{18.620c}$$

Ce que les équations (18.618) décrivent est l'intersection des trois parties suivanteas[68] :

— le côté de la droite $\big(g(A)g(B)\big)$ contenant $g(C)$,
— le côté de la droite $\big(g(A)g(C)\big)$ contenant $g(B)$,
— le côté de la droite $\big(g(B)g(C)\big)$ contenant $g(A)$,

Le lemme 18.209 nous permet d'exprimer ces trois parties en termes de K :

— l'image par g du côté de la droite (AB) contient C,
— l'image par g du côté de la droite (AC) contient B,
— l'image par g du côté de la droite (BC) contient A,

Comme g est une bijection, l'intersection des images par g est l'image par g de l'intersection. Bref, les équations (18.620) décrivent l'image par g de K.

Cette partie est donc pavée par (G, K).

(1a.i) **Si** $x - \sqrt{3}y \geqslant 0$ hop, même calculs avec des petites variations. Nous laissons la lectrice s'en occuper.

(1b) **Si** $x \geqslant \frac{1}{2}$ hop.

(2) **Si** $x + y/\sqrt{3} - 1 \geqslant 0$ hop.

68. Personnellement, je n'ai pas vérifié, mais ça m'étonnerait que ce soit faux. Vérifiez et écrivez-moi.

Les cas listés se traitent surement de la même façon. Nous tenons pour prouvé que (G, K) est bien surjectif sur $\mathbb{R}^2$.

Nous devons encore montrer la condition (2) de la définition 18.208. Commençons par déterminer la forme générale d'un élément de G sous la forme $\tau_a \circ r_0$ où r_0 est une application linéaire. Le lemme 1.323 nous indique qu'un élément général de $G = \text{gr}(\tau_{e_1}, \tau_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, R_0(\pi/3))$ est un produit arbitraire (mais fini) de τ_{e_1}, $\tau_{(\frac{1}{2},\frac{\sqrt{3}}{2})}$ et de $R_0(\pi/3)$ et de leurs inverses.

Comme toujours nous avons $\alpha \circ \tau_v = \tau_{\alpha(v)} \circ \alpha$. Donc nous pouvons passer toutes les rotations $R_0(\pi/3)$ et $R_0(\pi/3)^{-1}$ à droite du produit, quitte à produire des translations des formes suivantes :

$$R_0(\pi/3)^k e_1 \tag{18.621a}$$
$$R_0(\pi/3)^k (\frac{1}{2}, \frac{\sqrt{3}}{2}) \tag{18.621b}$$
$$R_0(\pi/3)^{-k} e_1 \tag{18.621c}$$
$$R_0(\pi/3)^{-k} (\frac{1}{2}, \frac{\sqrt{3}}{2}). \tag{18.621d}$$

En utilisant la matrice (18.617) nous trouvons assez vite que SUBEQooEMVIooNaaMqk

$$R_0(\pi/3)e_1 = \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix} \tag{18.622a}$$
$$R_0(\pi/3)^2 e_1 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix} \tag{18.622b}$$
$$R_0(\pi/3)^3 e_1 = -e_1. \tag{18.622c}$$

Les applications suivantes de $R_0(\pi/3)$ ne donnent rien de nouveau, si ce n'est le signe. Les puissances de $R_0(\pi/3)$ appliquées à $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ sont déjà parmi celles listées en (18.622). Quant aux inverses, $R_0(\pi/3)^{-1} = R_0(\pi/3)^5$; donc rien de nouveau non plus.

Un élément général de G est donc dans

$$\tau_v \circ R_0(\pi/3)^k \tag{18.623}$$

avec v dans

$$\mathbb{Z}e_1 + \mathbb{Z}\begin{pmatrix} 1/2 \\ \frac{\sqrt{3}}{2} \end{pmatrix} + \mathbb{Z}\begin{pmatrix} -1/2 \\ \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{18.624}$$

Remarquons que

$$\begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix} = -e_1 + \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}. \tag{18.625}$$

Donc un élément général de G est

$$\tau_{e_1}^k \circ \tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}^l \circ R_0(\pi/3)^m. \tag{18.626}$$

Nous pouvons maintenant prouver notre point. Pour cela nous allons successivement considérer les 5 rotations de K présentées dans la sous-figure 18.9(a). Pour chacune nous allons montrer qu'aucune translation ne permet d'obtenir une intersection avec K.

Nous allons en faire un seul en détail.

(1) **Pour $L = R_0(\pi/3)K$** hop. les détails sont laissés à notre aimable lecteur qui ne pourra pas ne pas compléter tous les cas.

(2) **Pour $L = R_0(\pi/3)^2 K$** Nous allons prouver que si $(x, y) \in \text{Int}(L)$, alors

$$\tau_{e_1}^k \circ \tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}^l (x, y) \tag{18.627}$$

ne peut pas être dans $\mathrm{Int}(K)$. Pour la simplicité des notations nous notons $r = R_0(\pi/3)^2$; nous avons

$$r = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}. \tag{18.628}$$

Nous considérons $g = \tau_{e_1}^k \circ \tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})} \circ r$. Un calcul nous donne l'image de (x, y) par g :

$$g(x,y) = \tau_{e_1}^k \circ \tau^l_{(-\frac{1}{2},\frac{\sqrt{3}}{2})} \begin{pmatrix} -\frac{1}{2}x - \frac{\sqrt{3}}{2}y \\ -\frac{\sqrt{3}}{2}x - \frac{1}{2}y \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}x - \frac{\sqrt{3}}{2}y + k - \frac{1}{2}l \\ \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l \end{pmatrix}. \tag{18.629}$$

Nous considérons $(x, y) \in \mathrm{Int}(K)$ tel que $g(x, y) \in \mathrm{Int}(K)$, et nous allons trouver une contradiction. Le fait que $(x, y) \in \mathrm{Int}(K)$ signifie que (x, y) satisfait les inéquations (18.616) mais avec des inégalités strictes. Le fait que $g(x, y) \in \mathrm{Int}(K)$ nous donne trois inéquations de plus. Assez rapide calcul :

$$f_{AB}\big(g(x,y)\big) = -2x + k - 2l, \tag{18.630a}$$
$$f_{BC}\big(g(x,y)\big) = -x - \sqrt{3}y - k - l + 1, \tag{18.630b}$$
$$f_{AC}\big(g(x,y)\big) = \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l. \tag{18.630c}$$

Au final, notre point (x, y) doit satisfaire le système suivant :

$$\begin{cases} x - \sqrt{3}y > 0 & \text{SUBEQooQJBRooSvcvfW} \quad (18.631a) \\ -x - \sqrt{3}y + 1 > 0 & \text{SUBEQooLWJQooAIQhCh} \quad (18.631b) \\ y > 0 & \text{SUBEQooYCVNooHJHVWt} \quad (18.631c) \\ -2x + k - 2l > 0 & \text{SUBEQooYYVUooORxLnp} \quad (18.631d) \\ -x - \sqrt{3}y - k - l + 1 > 0 & \text{SUBEQooRRGQooYtSxso} \quad (18.631e) \\ \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l > 0 & \text{SUBEQooSNVNooVrIVVy} \quad (18.631f) \end{cases}$$

Les inéquations 18.631c et 18.631a donnent déjà $x > 0$. De même avec 18.631b nous trouvons $x < 1$. Voilà déjà $x \in]0, 1[$ qui est directement visible sur le dessin du triangle K.
Puisque $x > 0$, l'inéquation (18.631d) donne

$$k - 2l > -2x + k - 2l > 0. \tag{18.632}$$

Donc $k - 2l > 0$.
Comme $x < 1$ et $y > 0$, l'inéquation (18.631f) donne

$$\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2}l > \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l > 0. \tag{18.633}$$

Donc $\frac{\sqrt{3}}{2}(1 + l) > 0$. Comme l est entier, cela donne $l \geqslant 0$.
Enfin, de (18.631e) nous tirons

$$-k + 1 > -x - \sqrt{3}y - k - l + 1 > 0, \tag{18.634}$$

Ce qui donne $k < 1$ et donc $k \leqslant 0$.
En résumé nous avons trouvé trois inéquations pour k et l :

$$\begin{cases} k \leqslant 0 & (18.635a) \\ l \geqslant 0 & (18.635b) \\ k - 2l > 0. & (18.635c) \end{cases}$$

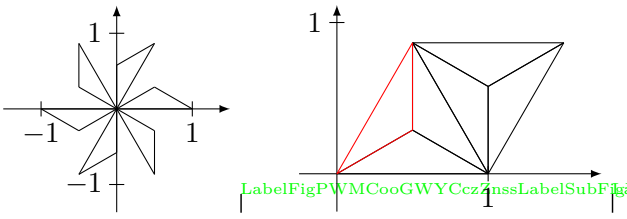
Ce système est impossible.
Il n'existe donc pas de translation qui, appliquée à $\mathrm{Int}(L)$, donne une intersection avec $\mathrm{Int}(K)$.

(3) **Pour** $L = R_0(\pi/3)^3K$ hop.

(4) **Pour** $L = R_0(\pi/3)^4K$ hop.

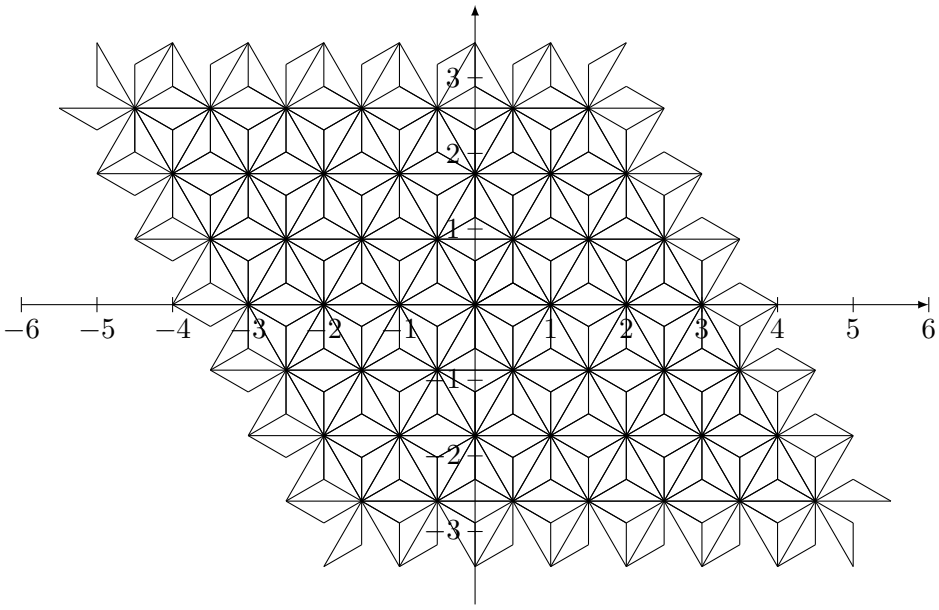
(5) **Pour** $L = R_0(\pi/3)^5K$ hop.

Voilà. J'espère que toutes les idées sont en place, et que les parties manquantes sont seulement des vérifications qui se font mécaniquement, de la même manière [69]. □



(a) Le compact K et ses rotations.

(b) Une maille du réseau des translations de G.



(c) Une partie du pavage complet.

FIGURE 18.9: Illustrations pour le pavage du lemme 18.213.

Lemme 18.214.
Le groupe

$$\mathrm{gr}\left(\tau_1, \tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}, R_0(\pi/3)\right) \tag{18.636}$$

est un groupe de pavage.

D'après [558], la démonstration du lemme 18.214 demande d'utiliser le losange de sommets $(0,0)$, $(1,0)$, $(\frac{1}{2},\frac{\sqrt{3}}{6})$ et $(\frac{1}{2},-\frac{\sqrt{3}}{6})$ [70].

Lemme 18.215 ([1])**.**
Soient $u_1, u_2 \in \mathbb{R}^2$ non colinéaires, ainsi que $\alpha, \beta \in [0,1]$. Nous considérons $v = \alpha u_1 + \beta u_2$.

Nous supposons pour fixer les idées, que $\|u_1\| \geqslant \|u_2\|$. Alors

$$\min\{\|v\|, \|u_1 + u_2 - v\|\} \leqslant \|u_1\|. \tag{18.637}$$

69. Je n'ai pas vérifié. Faites-le et écrivez-moi pour me dire ce qu'il en est.
70. Je n'ai pas vérifié, mais à mon avis une preuve doit prendre les mêmes idées que celles du lemme 18.213.

Autrement dit, tout point intérieur d'un parallélogramme est plus proche d'un angle que la longueur du plus long côté.

Démonstration. Les points $\alpha u_1 + \beta u_2$ ($\alpha, \beta \in [0,1]$) se divisent en deux parties : ceux avec $0 \leqslant \alpha + \beta \leqslant 1$ et ceux avec $1 \leqslant \alpha + \beta \leqslant 2$.

Si $\alpha + \beta \leqslant 1$ alors

$$\|\alpha u_1 + \beta u_2\| < \|\alpha u_1\| + \beta\|u_2\| = \alpha\|u_1\| + \beta\|u_2\| \leqslant (\alpha + \beta)\|u_1\| \leqslant \|u_1\|. \tag{18.638}$$

L'inégalité est stricte parce que u_1 et u_2 ne sont pas colinéaires.

Si au contraire $\alpha + \beta \geqslant 1$ nous avons

$$\|u_1 + u_2 - \alpha u_1 - \beta u_2\| < \|(1-\alpha)u_1\| + \|(1-\beta)u_2\| < (2 - \alpha - \beta)\|u_1\| \leqslant \|u_1\|. \tag{18.639}$$

□

LEMooWKTGooQlfuxm

Lemme 18.216.
Soient un ensemble E, et deux bijections $r, s : E \to E$ ayant chacune un unique point fixe. Si elles commutent, alors leurs points fixes sont égaux.

Démonstration. Nous nommons a le point fixe de r et b celui de s. Pour tout $x \in E$ nous avons $(rs)(x) = (sr)(x)$. En particulier pour $x = s^{-1}(a)$. D'une part

$$(rs)(x) = r(a) = a. \tag{18.640}$$

Et d'autre part,

$$(sr)(a) = (srs^{-1})(a) \tag{18.641}$$

Si nous imposons $(srs^{-1})(a) = a$, nous avons, en appliquant s^{-1} des deux côtés : $(rs^{-1})(a) = s^{-1}(a)$. Cela prouve que $s^{-1}(a)$ est un point fixe de r. Donc $s^{-1}(a) = a$.

Nous en déduisons que a est un point fixe de s et donc que $a = b$. □

LEMooDGSJooCiBhZz

Lemme 18.217.
Soit un sous-groupe T de $(\mathbb{R}^2, +)$ tel que $\|v\| > \delta$ pour tout $v \neq 0$ dans T.

Alors si u_1 et u_2 sont les plus petits éléments en norme de T, nous avons

$$T = \mathbb{Z}u_1 + \mathbb{Z}u_2. \tag{18.642}$$

Démonstration. Nous décomposons en plusieurs parties.

(1) $\underline{\mathbb{R}^+ v \cap T = \mathbb{N}v_m}$ Soit $v \in T$. L'ensemble $\{\lambda \in \mathbb{R}^+ \text{ tel que } \lambda v \in T\}$ a un minimum parce que tous les éléments de T sont en norme plus grands que $\delta > 0$. Soit λ_m ce minimum et $v_m = \lambda_m v$.

Nous prétendons à présent que $\mathbb{R}^+ v \cap T = \mathbb{N}v_m$. Nous ne faisons d'ailleurs pas que prétendre ; nous *prouvons*. En effet, soit $\lambda v \in T$. Nous devons prouver que $\lambda = l\lambda_m$ pour un certain $l \in \mathbb{N}$. Soit $k \in \mathbb{N}$ tel que $k \leqslant \lambda < k + 1$.

Nous avons

$$(k+1)\lambda_m - \lambda \leqslant (k+1)\lambda_m - k\lambda_m = \lambda_m. \tag{18.643}$$

Vu que T est un groupe pour l'addition, et que $\lambda_m v \in T$ et $\lambda v \in T$, cela implique que $\big((k+1)\lambda_m - \lambda\big)v \in T$. Mais

$$|\big((k+1)\lambda_m - \lambda\big)v| \leqslant \lambda_m. \tag{18.644}$$

Vue la propriété de minimalité de λ_m, nous avons forcément

$$(k+1)\lambda_m - \lambda = \lambda_m. \tag{18.645}$$

Cela prouve que $\lambda = k\lambda_m$.

Jusqu'ici nous avons prouvé que

$$\mathbb{R}^+ v \cap T = \mathbb{N}v_m \tag{18.646}$$

pour un certain multiple v_m de v.

(2) $\|u_1\| \leqslant \|v\|$ **pour tout** v

Nous montrons à présent qu'il existe $u_1 \in T$ tel que $\|u_1\| \leqslant \|v\|$ pour tout $v \in T$. Si tel n'était pas le cas, il existerait une suite $v_k \in T$ telle que $\|v_{k+1}\| < \|v_k\|$. Toute cette suite serait contenue dans la couronne (compacte) de rayons δ et $\|u_1\|$. Quitte à prendre une sous-suite, nous pouvons supposer que (v_k) converge[71]. Cette suite serait de Cauchy et pour tout ϵ (en particulier $\epsilon < \delta$), il existerait p, q tels que $\|v_p - v_q\| < \epsilon$. Puisque T est un groupe pour l'addition, nous aurions $v_p - v_q \in T$ avec $\|v_p - v_q\| < \epsilon \leqslant \delta$.

(3) **Première pause**

Si T est engendré seulement par u_1, nous avons fini. Autrement dit, si tout T est dans un sous-espace de dimension 1 de $\mathbb{R}^2$, nous avons terminé.

Dans la suite, nous supposons donc que T n'est pas contenu dans un sous-espace de dimension 1.

(4) $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$ Soient u_1 et u_2 les deux plus petits éléments de T en norme (peut-être ex-aequo). Ces deux éléments ne sont pas colinéaires, sinon leur différence serait plus petite. Ils forment donc une base de $\mathbb{R}^2$.

Soit $v \in T$. Comme $\{u_1, u_2\}$ est une base de $\mathbb{R}^2$, il existe $\alpha, \beta \in \mathbb{R}$ tels que

$$v = \alpha u_1 + \beta u_2. \tag{18.647}$$

Notre but est à présent de prouver que $\alpha, \beta \in \mathbb{Z}$.

Si ce n'était pas le cas, une simple translation nous mènerait dans les circonstances du lemme 18.215. Nous aurions alors que, soit $\|v\|$, soit $\|u_1 + u_2 - v\|$ serait strictement plus petit que le plus grand entre $\|u_1\|$ et $\|u_2\|$. Cela contredirait le fait que $\|u_1\|$ et $\|u_2\|$ étaient les deux plus petits.

□

PROPooPQYXooIDZlHy

Proposition 18.218 ([1])**.**
Si G est un groupe de pavage[72] *de $\mathbb{R}^2$ et si $f : \mathbb{R}^2 \to \mathbb{R}^2$ est une isométrie affine, alors le groupe*

$$G' = f \circ G \circ f^{-1} = \{f \circ g \circ f^{-1} \text{ tel que } g \in G\} \tag{18.648}$$

est un groupe de pavage.

Démonstration. Soit un compact K tel que (G, K) soit un pavage. Nous notons $K' = f(K)$. Nous devons prouver deux choses :

— $f \circ G \circ f^{-1}$ est un groupe de déplacements ;

— (G', K') est un pavage.

Ceci mène à prouver trois éléments.

(1) G' **est constitué de déplacements** Les éléments de G sont des isométries, ainsi que f et f^{-1}. Donc les éléments de G' sont des isométries.

Soit $g \in G$. Comme g est affine, il existe une décomposition $g = \tau_v \circ g_0$ où g_0 est linéaire. De même $f = \tau_w \circ f_0$. Les règles du produit et de l'inverse de la proposition 11.62(2)(3) nous indiquent que la partie linéaire de fgf^{-1} est $f_0 g_0 f_0^{-1}$.

En ce qui concerne le déterminant de $f_0 g_0 f_0^{-1}$, c'est la proposition 8.16 qui nous indique que

$$\det(f_0 g_0 f_0^{-1}) = \det(f_0)\det(g_0)\det(f_0^{-1}) = \det(f_0)\det(g_0)\det(f_0)^{-1} = \det(g_0) = 1. \tag{18.649}$$

71. Proposition 7.290.
72. Définition 18.208.

(2) $\underline{G' \cdot K' = \mathbb{R}^2}$ Comme f est affine, K' est encore compact [73] Nous avons :

$$\begin{aligned}
(f\circ G\circ f^{-1})\big(f(K)\big) &= \bigcup_{g\in G}\bigcup_{k\in K}(f\circ g\circ f^{-1})\big(f(k)\big) && (18.650a)\\
&= \bigcup_{g\in G}\bigcup_{k\in K}(f\circ g)(k) && (18.650b)\\
&= \bigcup_{g\in G} f\Big(\bigcup_{k\in K} g(k)\Big) && (18.650c)\\
&= \bigcup_{g\in G}(f\circ g)(K) && (18.650d)\\
&= f\Big(\bigcup_{g\in G} g(K)\Big) && (18.650e)\\
&= f\big(G\cdot K\big) && (18.650f)\\
&= \mathbb{R}^2. && (18.650g)
\end{aligned}$$

Pour la dernière égalité, nous avons utilisé le fait que f est bijective et que $G\cdot K=\mathbb{R}^2$.

(3) **L'autre condition** Deux éléments de G' s'écrivent fg_1f^{-1} et fg_2f^{-1}. Nous les supposons tels que

$$(fg_1f^{-1})\cdot \operatorname{Int}(K')\cap(fg_2f^{-1})\cdot\operatorname{Int}(K')\neq\varnothing. \tag{18.651}$$

Nous avons :

$$\begin{aligned}
(fg_1f^{-1})\cdot \operatorname{Int}(K')\cap(fg_2f^{-1})\cdot\operatorname{Int}(K') &= f\big(g_1\cdot\operatorname{Int}(K)\big)\cap f\big(g_2\cdot\operatorname{Int}(K)\big) && (18.652a)\\
&= f\Big(\underbrace{g_1\cdot\operatorname{Int}(K)\cap g_2\cdot\operatorname{Int}(K)}_{\neq\varnothing}\Big) && (18.652b)\\
&\neq\varnothing. && (18.652c)
\end{aligned}$$

SUBEQooWPMUooWxfdAw
SUBEQooTOGSooNrArAk

Justifications :

— Égalité (18.652a) parce qu'un peu de topologie nous enseigne que $f^{-1}\big(\operatorname{Int}(K')\big)=\operatorname{Int}\big(f^{-1}(K')\big)=\operatorname{Int}(K)$ parce que f est affine.

— Égalité (18.652b) parce que, f étant bijective, $f(A)\cap f(B)=f(A\cap B)$;

□

THOooUPHQooYfeHAy

Théorème 18.219 ([430, 558]).
Nous notons τ_v la translation de vecteur v, $r_{A,\theta}$ la rotation de centre A et d'angle θ ainsi que $\tau_i=\tau_{e_i}$.

Un groupe G est un groupe de pavage de $\mathbb{R}^2$ si et seulement si il existe une bijection affine $f\colon\mathbb{R}^2\to\mathbb{R}^2$ telle que $f\circ G\circ f^{-1}$ est un groupe de la liste suivante :

(1) $\operatorname{gr}\big(\tau_1,\tau_2\big)$
(2) $\operatorname{gr}\big(\tau_1,\tau_2,R_0(\pi)\big)$
(3) $\operatorname{gr}\big(\tau_1,\tau_{(\frac12,\frac{\sqrt3}{2})},R_0(2\pi/3)\big)$
(4) $\operatorname{gr}\big(\tau_1,\tau_2,R_0(\pi/2)\big)$
(5) $\operatorname{gr}\big(\tau_1,\tau_{(-\frac12,\frac{\sqrt3}{2})},R_0(\pi/3)\big)$.

Démonstration. Les lemmes 18.210, 18.211, 18.212, 18.213, et 18.214 montrent que les groupes listés sont des groupes de pavage. La proposition 18.218 nous montre alors que si $H=fGf^{-1}$ est dans la liste, alors G est un groupe de pavage. Il nous reste à montrer que si G est un groupe de pavage, alors il existe une application affine α telle que $\alpha G\alpha^{-1}$ est un groupe de la liste.

Soit (G,K) un pavage de $\mathbb{R}^2$. Nous notons T l'ensemble des translations dans G, plus précisément,

$$T=\{v\in\mathbb{R}^2 \text{ tel que } \tau_v\in G\}. \tag{18.653}$$

73. Il suffit de prouver que $f(K)$ est fermé et borné par le théorème de Borel-Lebesgue 9.24.

(1) **Une borne pour T** Nous prouvons qu'il existe $\delta > 0$ tel que $\|v\| > \delta$ pour tout $v \neq 0 \in T$. En effet, soit $m \in \text{Int}(K)$ ainsi que r tel que $B(m,r) \subset \text{Int}(K)$. Alors si $v \in T$ est tel que $\|v\| < r$ nous avons $\tau_v(m) \in \text{Int}(K)$, ce qui donnerait

$$m \in \tau_v\big(\text{Int}(K)\big) \cap \text{Int}(K) \tag{18.654}$$

par hypothèse, cette intersection est non vide seulement si $v = 0$.

Donc il existe δ tel que $\|v\| \geqslant \delta$ pour tout $v \in T$.

(2) **Utilisation du lemme**

La partie T est donc dans la position du lemme 18.217 et nous avons

$$T = \mathbb{Z}u_1 + \mathbb{Z}u_2 \tag{18.655}$$

pour les vecteurs u_1 et u_2 les plus petits en norme de T.

En réalité, il se peut que T soit plus petit que ça, parce que G peut par exemple ne contenir aucune translation. Nous avons trois possibilités :

— $T = \{0\}$,

— $T = \mathbb{Z}u$ pour un certain $u \neq 0$ dans $\mathbb{R}^2$,

— $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$ pour certains $u_1, u_2 \in \mathbb{R}^2\backslash\{0\}$.

(3) **Translation** Si $r, s \in G$, alors l'élément $rsr^{-1}s^{-1}$ est une translation. En effet, puisque les éléments de G sont des déplacements, ce sont des applications affines et donc il existe des applications linéaires A_r, A_s et des translations τ_r, τ_s telles que $r = A_r \circ \tau_r$ et $s = A_s \circ \tau_s$. La proposition 11.62 nous donne les inverses. Nous avons

$$rsr^{-1}s^{-1} = (A_r \circ \tau_r)(A_s \circ \tau_s)(A_r^{-1} \circ \tau_{-A_r v_r})(A_s^{-1} \circ \tau_{-A_s v_s}). \tag{18.656}$$

La partie linéaire de cela est

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1}. \tag{18.657}$$

C'est donc une composée de rotations centrées en $(0,0)$. Mais ces rotations forment un groupe abélien (proposition 18.129). Donc nous pouvons écrire

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1} = A_r \circ A_r^{-1} \circ A_s \circ A_s^{-1} = \text{Id}\,. \tag{18.658}$$

Tout ceci pour dire que dès que $r, s \in G$, l'élément $rsr^{-1}s^{-1}$ est une translation.

(4) **Les parties linéaires[1]** Nous savons de l'exemple 11.18 que les éléments de G s'écrivent sous la forme $f = \tau_v \circ \alpha$ où $v \in \mathbb{R}^2$ et $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ est linéaire.

De plus, f étant une isométrie de $(\mathbb{R}^2, d)$, l'application α est une isométrie de $(\mathbb{R}^2, \|.\|)$. Puisque α est une isométrie, $\det(\alpha) = \pm 1$. Mais les déplacements conservent l'orientation ; donc α doit conserver l'orientation, et la proposition 8.34 nous dit que $\det(\alpha) > 0$. Donc

$$\det(\alpha) = 1. \tag{18.659}$$

Le théorème 18.95 dit que α est la composition d'un nombre pair de réflexions. Mais comme il y en a au plus trois (théorème 18.189), l'application α est composée de zéro ou deux réflexions.

Donc les parties linéaires des éléments de G sont des rotations.

(5) **Les autres**

Les parties linéaires des éléments de G sont des rotations. Mais les éléments de G eux-mêmes ne sont pas tellement mystérieux. Puisque ce sont des isométries de $(\mathbb{R}^2, d)$, elles sont composées de 0, 1, 2 ou 3 réflexions.

Mais ce sont des déplacements, donc ils préservent l'orientation et le théorème 18.95 dit qu'ils sont des composées de zéro ou deux réflexions (nombre pair). Ce sont donc des rotations.

(6) **Hein ?** Les éléments linéaires de G sont des rotations. Et les autres aussi ? Les linéaires sont des rotations autour de $(0,0)$; les autres sont des rotations autour de points autres que $(0,0)$.
C'est pourquoi dans la suite, nous préciserons « rotation linéaire » pour une rotation autour de $(0,0)$ et nous dirons « rotation » pour une rotation en général. Dans le contexte affine, il faut toujours faire attention à ça : une rotation peut très bien n'être pas linéaire [74].

(7) **Les rotations linéaires stabilisent T** Nous prouvons maintenant que les rotations linéaires de G stabilisent T, c'est-à-dire que si $v \in T$ et si α est une rotation linéaire de G, alors $\alpha(v) \in T$. La transformation $\alpha\tau_v\alpha^{-1}$ est dans G. Mais pour tout $x \in \mathbb{R}^2$ nous avons

$$(\alpha\tau_v\alpha^{-1})(x) = \alpha\big(\alpha^{-1}(x) + v\big) = x + \alpha(v) = \tau_{\alpha(v)}(x). \tag{18.660}$$ EQooLLZVooUnabir

Donc $\alpha\tau\alpha^{-1} = \tau_{\alpha(v)}$ et $\alpha(v) \in T$.

(8) **Exclusion de $T = \{0\}$** Le fait que $T = \{0\}$ ne signifie pas que tous les éléments de G sont des rotations ; il peut encore y avoir des composées de rotations et de translations $A \circ \tau$. Cela étant dit, si $T = \{0\}$, il n'en reste pas moins que $rsr^{-1}s^{-1}$ est une translation, c'est-à-dire est égal à Id. Mais $rsr^{-1}s^{-1} = e$ implique $rs = sr$.
Donc G est abélien. Les éléments de G sont donc des rotations qui commutent deux à deux. Puisqu'une rotation a son centre comme unique point fixe, le lemme 18.216 nous dit que tous les éléments de G sont des rotations de même centre.
Soit c le centre commun de tous les éléments de G. Vu que K est compact dans $\mathbb{R}^2$, il existe $r > 0$ tel que $K \subset B(c,r)$. Puisque G stabilise toutes les boules centrées en c, nous avons

$$G \cdot K \subset B(c,r). \tag{18.661}$$

Donc nous n'avons pas un recouvrement de $\mathbb{R}^2$. Le cas $T = \{0\}$ est exclu.

(9) **Exclusion de $T = \mathbb{Z}u$** Nous supposons à présent que $T = \mathbb{Z}u$ pour un certain $u \in \mathbb{R}^2$.

(9a) $r_0 = \pm\,\mathrm{Id}$ Nous savons que tous les éléments de G sont des rotations ; soit un élément r de G. La proposition 18.128(3) nous indique qu'il existe un point $a \in \mathbb{R}^2$ ainsi qu'une rotation linéaire r_0 telle que $r = \tau_a^{-1}r_0\tau_a$. Nous allons prouver que r_0 est $\pm\,\mathrm{Id}$. D'abord,

$$r\tau_u r^{-1} = \tau_a^{-1}r_0\tau_a\tau_u\tau_a^{-1}r_0^{-1}\tau_a = \tau_a^{-1}r_0\tau_u r_0^{-1}\tau_a. \tag{18.662}$$

Ensuite, nous appliquons cela à $x \in \mathbb{R}^2$:

$$\begin{aligned}(\tau_a^{-1}r_0\tau_u r_0^{-1}\tau_a)(x) &= (\tau_a^{-1}r_0\tau_u)\big(r_0^{-1}(x+a)\big) && \text{(18.663a)}\\ &= (\tau_a^{-1}r_0)\big(r_0^{-1}(x+a) + u\big) && \text{(18.663b)}\\ &= \tau_a^{-1}\big(x + a + r_0(u)\big) && \text{(18.663c)}\\ &= x + r_0(u). && \text{(18.663d)}\end{aligned}$$

Donc $r \circ \tau_u \circ r^{-1} = \tau_{r_0(u)}$, ce qui prouve que $r_0(u) \in T$. Comme $\|r_0(u)\| = \|u\|$ nous avons forcément $r_0(u) = \pm u$.
Si $r_0(u) = u$, alors $r_0 = \mathrm{Id}$ (parce que r_0 est une rotation fixant plus qu'un seul point). Dans ce cas, $r = \mathrm{Id}$.
Si au contraire $r_0(u) = -u$, alors $r_0 = -\mathrm{Id}$.

(9b) **Forme générale** Donc si r est un élément non trivial de G nous avons $r = \tau_a^{-1} \circ (-\mathrm{Id}) \circ \tau_a$, et alors

$$\big(\tau_a^{-1} \circ (-\mathrm{Id}) \circ \tau_a\big)(x) = \tau_a^{-1}(-\mathrm{Id})(x+a) = \tau_a^{-1}(-x-a) = -x - 2a. \tag{18.664}$$

Donc pour tout $r \in G$, il existe $a \in \mathbb{R}^2$ tel que

$$r(x) = -x - 2a. \tag{18.665}$$ EQooQGNVooKyCCYW

Pour information, le centre de cette rotation est $-a$ (c'est le seul point fixe).

74. Lorsque, ailleurs dans le Frido, nous disons « rotation », souvent nous pensons « rotation linéaire ». Gardez cependant à l'esprit qu'une rotation peut très bien être centrée ailleurs qu'en l'origine, et soyez toujours capable de préciser le cas échéant.

(9c) **Les centres sont alignés** Soit r une rotation de centre $-a$ et s de centre $-b$. Alors

$$(rs)(x) = r(-x-2b) = x+2b-2a = x+2(b-a). \tag{18.666}$$

Donc $rs = \tau_{2(b-a)}$.
Cela prouve que $2(b-a) \in \mathbb{Z}u$.

(9d) **Une bande** Soit la droite $D = \mathbb{R}u$. Nous considérons la bande

$$B_r = \{x \in \mathbb{R}^2 \text{ tel que } d(x,D) < r\}. \tag{18.667}$$

(9e) **Une inclusion** Nous prouvons à présent que pour tout r, nous avons

$$G \cdot B_r \subset B_r. \tag{18.668}$$

Nous savons qu'un élément général de G est une rotation centrée en un point de D, et que l'action d'une telle rotation est donnée par (18.665). Nous avons

$$\begin{aligned} d\big(r(x),D\big) &= d(x+2a,D) && \text{EQooNCQGooGGAhCR} \quad (18.669a)\\ &= d\big(\tau_{2a}^{-1}(x+2a),\tau_{2a}^{-1}(D)\big) && \text{EQooNFJNooSAYtBD} \quad (18.669b)\\ &= d(x,D). && \text{SUBEQooMGTMooAlRwWT} \quad (18.669c) \end{aligned}$$

Justifications :

— Pour (18.669a), nous avons $D = -D$ et $d(x,y) = d(-x,-y)$.
— Pour (18.669b), invariance par translation de la distance dans $\mathbb{R}^2$.
— Pour (18.669c), les éléments de D sont les multiples de a ; donc cette droite est invariante par cette translation.

Bref, $d(x,D) = d\big(r(x),D\big)$. Donc, pour tout $r > 0$ [75] et pour tout $g \in G$, si $x \in B_r$, alors $g(x) \in B_r$.

(9f) **Exclusion**
Comme K est compact et que la fonction $x \mapsto d(x,D)$ est continue, il existe $r > 0$ tel que $K \subset B_r$. Avec ça, $G \cdot K \subset G \cdot B_r \subset B_r$. Donc $G \cdot K$ ne recouvre pas tout $\mathbb{R}^2$ et G n'est pas un groupe de pavage.

Et nous voilà avec seulement $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$ en lice.

(1) **Pause : quelques parties de G à ne pas confondre** Il convient de ne pas se perdre entre différentes parties de G. Je vous laisse méditer quelques temps sur la liste suivante :

(1a) G est le groupe que nous cherchons à déterminer ;
(1b) T est le groupe des translations de G ;
(1c) le groupe des rotations linéaires dans G ;
(1d) l'ensemble des $\tau_A^{-1} \circ r \circ \tau_A$ où r est une rotation de G centrée en A. ITEMooFWWMooNzLUGy
(1e) L est l'ensemble des parties linéaires des éléments de G. ITEMooIEJZooNaSKpc

En particulier les deux derniers points ne sont pas les mêmes.
Dans le cas (1d), il s'agit de dire que r est une rotation centrée en $A \in \mathbb{R}^2$ et écrire $r = \tau_A \circ r_0 \circ \tau_A^{-1}$ (proposition 18.128(3)) et considérer r_0. Dans ce cas, r_0 est une rotation, mais elle n'est pas ce que nous appelons la « partie linéaire » de r. Il n'y a pas de garantie que cela forme un groupe.
Si r est une rotation dans G, dans le cas (1e) il s'agit de décomposer $r = \tau_v \circ \alpha$ (lemme 11.13) et considérer α. Dans ce cas, α est linéaire, mais il n'y a pas de garantie que α soit une rotation.

75. Remarquez la notation malheureuse pour r qui est maintenant une distance alors que trois mots plus tôt, c'était un élément de G.

(2) **L est un groupe** En trois conditions.

— Si $\alpha, \beta \in L$, il existe $v, w \in \mathbb{R}^2$ tels que $\tau_v \circ \alpha \in G$ et $\tau_w \circ \beta \in G$. La loi de produit 11.62(2) dit que $\tau_v \circ \alpha \circ \tau_w \circ \beta = \tau_{\alpha(w)+v} \circ \alpha\beta$. Donc $\alpha\beta$ est la partie linéaire d'un élément de G.

— De la même façon, en utilisant l'inverse 11.62(3), $(\tau_v \circ \alpha)^{-1} = \tau_{-\alpha^{-1}(v)} \circ \alpha^{-1}$. Donc α^{-1} est la partie linéaire d'un élément de G.

— Et enfin $\mathrm{Id} = \tau_0 \circ \mathrm{Id}$. Donc l'identité est dans L.

Ok : L est un groupe.

(3) **Précision** L'ensemble L est un groupe, certes. Mais rien ne dit que L soit un sous-groupe de G.

(4) **L préserve le réseau** Soit $\alpha \in L$. Il existe $v \in \mathbb{R}^2$ tel que $g = \tau_v \circ \alpha \in G$. Soit $u \in T$. Nous allons montrer que $\alpha(u) \in T$. Vu que g et τ_u sont dans G, l'élément $g\tau_u g^{-1}$ est également dans G. Nous l'appliquons à $x \in \mathbb{R}^2$:

$$\begin{aligned}(\tau_v\alpha\tau_u\alpha^{-1}\tau_v^{-1})(x) &= (\tau_v\alpha\tau_u\alpha^{-1})(x-v) && (18.670a)\\ &= (\tau_v\alpha)\big(\alpha^{-1}(x-v)+u\big) && (18.670b)\\ &= \tau_v\big(x-v+\alpha(u)\big) && (18.670c)\\ &= x+\alpha(u) && (18.670d)\\ &= \tau_{\alpha(u)}(x). && (18.670e)\end{aligned}$$

Donc $g\tau_u g^{-1} = \tau_{\alpha(u)} \in G$.

(5) **Question de trace** Soit $\alpha \in L$; dans la base $\{u_1, u_2\}$ la matrice de α est

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{18.671}$$

avec $a, b, c, d \in \mathbb{Z}$. La trace de cette matrice est $a + d \in \mathbb{Z}$. Dans la base canonique de $\mathbb{R}^2$ par contre, la proposition 18.139 nous dit qu'il existe $\theta \in [0, 2\pi[$ tel que la matrice de α soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{18.672}$$

La trace est $2\cos(\theta)$. La trace est invariante par changement de base[76], donc $2\cos(\theta) = a + d \in \mathbb{Z}$. Les possibilités pour $\cos(\theta)$ sont donc -1, $-1/2$, 0, $1/2$ et 1.

(6) **Les angles possibles** Nous savons que $\cos(\frac{\pi}{2}) = 0$ et $\cos(\pi/3) = 1/2$ (proposition 18.20(6) et lemme 18.27). La proposition 18.74 nous dit alors que, dans notre cas, les valeurs possibles pour θ dans $[0, 2\pi[$ sont

$$\{0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{3\pi}{2}, \frac{5\pi}{3}\}. \tag{18.673}$$

EQooLMPIooPQoHUI

Donc les rotations possibles dans L sont les rotations de ces angles.

Nous devons trouver quels sont les groupes qui peuvent être formés seulement avec ces éléments.

(7) **Quelques combinaisons impossibles**

Puisque L est un groupe, il y a des combinaisons impossibles. Par exemple si $R_0(\pi/3)$ et $R_0(\pi/2)$ sont dans L, alors la composée[77] $R_0(\pi/2)R_0(\pi/3) = R_0(5\pi/6)$ est également dans L. Mais comme $5\pi/6$ n'est pas dans la liste (18.673), $R_0(5\pi/6)$ n'est pas dans L.

En raisonnant de la sorte, nous voyons que si $R_0(\pi/2) \in L$, alors $L = \mathrm{gr}\big(R_0(\pi/2)\big)$.

(8) **La liste** Plus généralement, les possibilités pour L sont

— $\{\mathrm{Id}\}$

76. Proposition 8.207.
77. Proposition 18.140 pour l'addition des angles.

— $\mathrm{gr}\big(R_0(\pi/2)\big)$
— $\mathrm{gr}\big(R_0(\pi/3)\big)$
— $\mathrm{gr}\big(R_0(\pi)\big)$
— $\mathrm{gr}\big(R_0(2\pi/3)\big)$

Une justification plus courte pour cette liste est d'invoquer le théorème 18.200(2) qui dit que L étant un sous-groupe fini des isométries de $(\mathbb{R}^2, d)$, il est cyclique et donc monogène. Notons pour cela que $\mathrm{gr}\big(R_0(4\pi/3)\big) = \mathrm{gr}\big(R_0(2\pi/3)\big)$ parce que $R_0(4\pi/3) = R_0(2\pi/3)^{-1}$. De la même manière $R_0(3\pi/2) = R_0(\pi/2)^{-1}$ et $R_0(5\pi/3) = R_0(\pi/3)^{-1}$.

(9) **Le cas un peu générique**

Nous supposons que $L = \mathrm{gr}\big(R_0(\theta)\big)$ pour un certain θ. Nous allons voir qu'à part dans les cas $\theta = 0$ et $\theta = \pi$, il est possible de trouver une application affine α telle que le groupe $\alpha G \alpha^{-1}$ soit alors dans la liste.

Nous considérons un élément de G de la forme $\tau_{v_0} \circ R_0(\theta)$. Pour être bien clair, il n'est absolument pas garanti que v_0 soit dans T. Nous allons chercher un élément $w_0 \in \mathbb{R}^2$ tel que le groupe de pavage[78] $G' = \tau_{w_0} G \tau_{w_0}^{-1}$ contienne $R_0(\theta)$. Le groupe G' contient l'élément $g = \tau_{w_0}\tau_{v_0}R_0(\theta)\tau_{w_0}^{-1}$; nous l'appliquons à $x \in \mathbb{R}^2$:

$$g(x) = R_0(\theta)x - R_0(\theta)w_0 + w_0 + v_0. \tag{18.674}$$

Nous avons $g(x) = R_0(\theta)x$ lorsque

$$w_0 = \big(R_0(\theta) - \mathbb{1}\big)^{-1} v_0. \tag{18.675}$$

Voilà pourquoi le cas $\theta = 0$ sera traité à part : dans le cas $\theta = 0$, l'opérateur $R_0(\theta) - \mathbb{1}$ n'est pas inversible. Dans les autres cas, nous avons un groupe $G' = \tau_{w_0} G \tau_{w_0}^{-1}$ qui contient $R_0(\theta)$. Puisqu'un élément général de G est de la forme $\tau_v \circ R_0(\theta)^k$, un élément général de G' est de la forme

$$\tau_{w_0} \circ \tau_v \circ R_0(\theta)^k \circ \tau_{w_0}^{-1} = \tau_{w_0 + v - R_0(\theta)^k w} \circ R_0(\theta)^k. \tag{18.676}$$

Donc tous les éléments de G' sont encore de la forme

$$\tau_v \circ R_0(\theta)^k. \tag{18.677}$$

Mais dans G' nous avons une information capitale : $R_0(\theta)^k$ lui-même est dans G'. Donc si $\tau_v \circ R_0(\theta)^k \in G'$, alors $\tau_v \in G'$.

Vu que G' est encore un groupe de pavage, tout ce qui a été dit précédemment tient, et le groupe des translations dans G' est un réseau $T' = \mathbb{Z}u_1 + \mathbb{Z}u_2$ où u_1 et u_2 peuvent être choisis arbitrairement parmi les deux plus petits vecteurs non colinéaires de T.

Tous les éléments de G' sont de la forme $\tau_v \circ R_0(\theta)^k$ avec $v \in T'$. Donc

$$G' = \mathrm{gr}\big(\tau_{u_1}, \tau_{u_2}, R_0(\theta)\big). \tag{18.678}$$

EQooUUDQooQnBcIi

Nous savons que $R_0(\theta)$ fixe T'. Or l'élément $R_0(\theta)u_1$ a la même norme que u_1 ; donc nous pouvons choisir $u_2 = R_0(\theta)u_1$ pour peu que $R_0(\theta)u_1$ soit non colinéaire à u_1. Et c'est ici que nous laissons le cas $\theta = \pi$ de côté.

Nous écrivons donc sans vergogne que

$$G' = \mathrm{gr}\big(\tau_{u_1}, \tau_{R_0(\theta)u_1}, R_0(\theta)\big). \tag{18.679}$$

Nous allons maintenant nous occuper de u_1. Pour cela nous considérons une rotation suivie d'une homothétie α telle que $\alpha(u_1) = e_1$. Une telle opération α d'une part, commute avec $R_0(\theta)$ et d'autre part, réalise $\alpha \circ \tau_v \circ \alpha^{-1} = \tau_{\alpha(v)}$. Donc le groupe $G'' = \alpha G' \alpha^{-1}$ est, par le lemme 1.324,

$$G'' = \mathrm{gr}\big(\tau_{e_1}, \tau_{\alpha R_0(\theta)u_1}, R_0(\theta)\big). \tag{18.680}$$

78. Proposition 18.218.

Cela s'écrit aussi bien

$$G'' = \mathrm{gr}\big(\tau_{e_1}, \tau_{R_0(\theta)e_1}, R_0(\theta)\big). \tag{18.681}$$

Et voilà, ce groupe G'' est un de ceux de la liste.

(10) **Le cas $L = \{\mathrm{Id}\}$** Dans ce cas, tous les éléments de G sont de la forme τ_v avec $v \in T$. Nous considérons l'application linéaire $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que $\alpha(u_1) = e_1$ et $\alpha(u_2) = e_2$. Nous avons

$$(\alpha \circ \tau_{u_i} \circ \alpha^{-1})(x) = (\alpha\tau_{u_i})\alpha^{-1}(x) = \alpha\big(\alpha^{-1}(x) + u_i\big) = x + \alpha(u_i) = \tau_i(x). \tag{18.682}$$

Donc $\alpha \circ \tau_{u_i} \circ \alpha^{-1} = \tau_i$. De la même façon, si $a \in \mathbb{Z}$ nous avons $\alpha\tau_{au_i}\alpha^{-1} = \tau_{ae_i}$. Et avec tout ça, si $v \in T$, alors $v = au_1 + bu_2$ et nous avons

$$\alpha\tau_v\alpha^{-1} = \alpha\tau_{au_1}\tau_{bu_2}\alpha^{-1} = \alpha\tau_{au_1}\alpha^{-1}\alpha\tau_{bu_2}\alpha^{-1} = \tau_{ae_1}\tau_{be_2}. \tag{18.683}$$

Nous avons donc

$$\alpha G\alpha^{-1} = \mathrm{gr}(\tau_1, \tau_2). \tag{18.684}$$

Cela est un des groupes de la liste.

Notez qu'à la place de ces calculs, nous pouvions aussi invoquer la proposition 1.324.

(11) **Le cas $L = \{-\mathrm{Id}\}$** Dans le cas $\theta = \pi$, nous pouvons aller jusqu'à (18.678) et écrire

$$G' = \mathrm{gr}\big(\tau_{u_1}, \tau_{u_2}, R_0(\pi)\big). \tag{18.685}$$

Puisque $R_0(\pi) = -\mathrm{Id}$ commute avec toutes les applications linéaires et que u_1 et u_2 ne sont pas colinéaires, nous pouvons considérer une application linéaire $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que $\alpha(u_1) = e_1$, $\alpha(u_2) = e_2$. Nous avons alors

$$G'' = \alpha G'\alpha^{-1} = \mathrm{gr}\big(\tau_{e_1}, \tau_{e_2}, -\mathrm{Id}\big), \tag{18.686}$$

qui est encore dans la liste.

□

18.12 Un peu de structure de O(n)

18.12.1 Valeurs propres dans O(n)

PROPooVEJGooWnqtMm

Proposition 18.220 ([559]).
Soit une matrice $A \in O(n)$. Si $\lambda \in \mathbb{C}$ est une valeur propre de A, alors $\bar{\lambda}$ est également une valeur propre de A, et de plus $|\lambda| = 1$.

Démonstration. Dire que $\lambda \in \mathbb{C}$ est une valeur propre de A signifie qu'il existe $x \in \mathbb{C}^n$ (non nul) tel que $Ax = \lambda x$. Comme les éléments de la matrice A sont réels,

$$A\bar{x} = \bar{A}\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}. \tag{18.687}$$

Donc $\bar{\lambda}$ est une valeur propre de A pour le vecteur propre $\bar{x}$.

Soit λ une valeur propre de A de vecteur propre x. Alors nous avons d'une part

$$\langle\overline{Ax}, Ax\rangle = \langle A^tA\bar{x}, x\rangle = \langle x, \bar{x}\rangle, \tag{18.688}$$

et d'autre part

$$\langle\overline{Ax}, Ax\rangle = \langle\bar{\lambda}\bar{x}, \lambda x\rangle = |\lambda|^2\langle\bar{x}, x\rangle. \tag{18.689}$$

Puisque $x \neq 0$ nous avons aussi $\langle\bar{x}, x\rangle \neq 0$. Par conséquent $|\lambda|^2 = 1$ et $|\lambda| = 1$. □

LEMooNEDQooNRmASH

Lemme 18.221 ([559]).
Soient un espace vectoriel euclidien E de dimension finie et une isométrie f de E. Soit F un sous-espace de E stable par f. Alors $F^\perp$ est stable par f.

Démonstration. La restriction $f_F\colon F\to F$ est encore une isométrie ; elle est donc inversible : pour tout $y\in F$, il existe $x\in F$ tel que $y=f(x)$. Soit $a\in F^\perp$; nous montrons que $f(a)\in F^\perp$. Soit donc $y\in F$ et calculons :

$$\langle y,f(a)\rangle=\langle f(x),f(a)\rangle=\langle x,a\rangle=0 \tag{18.690}$$

parce que $x\in F$ et $a\in F^\perp$. □

PROPooOMORooWzsrDB

Proposition 18.222 ([559]).
Soit une isométrie $f\colon\mathbb{R}^3\to\mathbb{R}^3$.

(1) L'application linéaire f possède au moins une valeur propre réelle qui vaut ±1.

(2) Il existe une base orthonormée de $\mathbb{R}^3$ dans laquelle la matrice de f est de la forme

$$\begin{pmatrix}\lambda&0&0\\0&\cos(\theta)&-\epsilon\sin(\theta)\\0&\sin(\theta)&\epsilon\cos(\theta)\end{pmatrix} \tag{18.691}$$

avec $\epsilon,\lambda=\pm1$ et $\theta\in[0,2\pi[$.

Démonstration. Le polynôme caractéristique de f, donné par $\det(f-\lambda\,\mathrm{Id})$, est à coefficients réels et de degré 3. Il possède donc au moins une solution réelle, par le corolaire 12.91. Soit donc une valeur propre réelle λ de χ_f ; par le lemme 18.220 nous avons $\lambda=\pm1$. Soit u_1 le vecteur propre correspondant. Nous notons F l'espace engendré par u_1.

Nous avons $f(F)=F$ et donc $f(F^\perp)=F^\perp$ par le lemme 18.221. Soit une base orthonormée $\{u_2,u_3\}$ de $F^\perp$ et la matrice B de la restriction f_p à $F^\perp$. Comme l'application f_p est une isométrie de $F^\perp$, la matrice B est, par le lemme 18.135, de la forme

$$B=\begin{pmatrix}\cos(\theta)&-\epsilon\sin(\theta)\\\sin(\theta)&\epsilon\cos(\theta)\end{pmatrix} \tag{18.692}$$

pour un certain $\theta\in[0,2\pi[$ et $\epsilon=\pm1$.

Dans la base $\{u_1,u_2,u_3\}$ de $\mathbb{R}^3$, la matrice de f est alors

$$\begin{pmatrix}\lambda&0\\0&B\end{pmatrix}, \tag{18.693}$$

comme annoncé. □

Pour classifier les isométries de $\mathbb{R}^3$, nous pouvons nous baser sur les possibilités de la matrice donnée dans le lemme 18.222. Il y a essentiellement quatre possibilités suivant les valeurs de $\lambda=\pm1$ et $\epsilon=\pm1$.

(1) **Si $\epsilon=\lambda=1$** Alors la matrice est

$$A=\begin{pmatrix}1&0&0\\0&\cos(\theta)&-\sin(\theta)\\0&\sin(\theta)&\cos(\theta)\end{pmatrix} \tag{18.694}$$

et l'isométrie correspondante est la rotation d'angle $-\theta$ autour de la droite de u_1.

(2) **Si $\epsilon=\lambda=-1$** Alors la matrice est

$$A=\begin{pmatrix}-1&0&0\\0&\cos(\theta)&\sin(\theta)\\0&\sin(\theta)&-\cos(\theta)\end{pmatrix} \tag{18.695}$$

Cette application est plus subtile, parce que même dans le plan $\mathrm{Span}(u_2,u_3)$, ce n'est pas une rotation. Nous allons montrer qu'il s'agit d'une réflexion autour de la droite d'angle $\theta/2$ dans

le plan $\mathrm{Span}(u_2, u_3)$. Nous nommons D cette droite. Dans la base (u_1, u_2, u_3), cette droite est donnée par [79]

$$t \mapsto t\big(0, \cos(\theta/2), \sin(\theta/2)\big). \tag{18.696}$$

L'image de u_1 par cette réflexion est $-u_1$, c'est clair.

Étudions en détail l'image de u_3. Nous devons démontrer que la droite D coupe le segment $[u_3, A(u_3)]$ perpendiculairement en son milieu.

Nous considérons la base (u_2, u_3) du plan $\mathrm{Span}(u_2, u_3)$. En utilisant les coordonnées dans cette base, nous avons $u_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et $A(u_3) = \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}$. Le milieu du segment $[u_3, A(u_3)]$ est le point

$$M = \left(\frac{\sin(\theta)}{2}, \frac{1 - \cos(\theta)}{2} \right). \tag{18.697}$$

Les formules de duplication d'angle du corolaire 18.15 nous permettent d'écrire $\sin(\theta)$ et $\cos(\theta)$ en fonction de $\sin(\theta/2)$ et $\cos(\theta/2)$, et donc d'exprimer le point M de la façon suivante :

$$\begin{aligned} M &= \left(\cos(\theta/2)\sin(\theta/2), \frac{1 - \big(\cos^2(\theta/2) - \sin^2(\theta/2)\big)}{2} \right) && \text{(18.698a)} \\ &= \big(\cos(\theta/2)\sin(\theta/2), \sin^2(\theta/2)\big) && \text{(18.698b)} \\ &= \sin(\theta/2)\big(\cos(\theta/2), \sin(\theta/2)\big). && \text{(18.698c)} \end{aligned}$$

Ce point fait donc partie de la droite D. La droite D coupe le segment $[u_3, A(u_3)]$ en son milieu.

En ce qui concerne l'orthogonalité, nous calculons le produit scalaire

$$\big(A(u_3) - u_3\big) \cdot \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = \sin(\theta)\cos(\theta/2) - \big(1 + \cos(\theta)\big)\sin(\theta/2) = 0 \tag{18.699}$$

où nous avons encore utilisé les duplications d'angles et le fait que $1 = \cos^2(\theta/2) + \sin^2(\theta/2)$ (lemme 18.4).

(3) **Si $\epsilon = -1$ et $\lambda = 1$** Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{18.700}$$

C'est la symétrie orthogonale par le plan engendré par u_1 et $v = \cos(\theta/2)u_2 + \sin(\theta/2)u_3$.

Le vecteur u_1 est bien évidemment préservé par A. En ce qui concerne le vecteur v,

$$A(v) = \cos(\theta/2) \begin{pmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{pmatrix} + \sin(\theta/2) \begin{pmatrix} 0 \\ -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = v. \tag{18.701}$$

Nous avons sauté quelques étapes de calcul mettant en scène les formules de duplication d'angle [80] : exprimer $\cos(\theta) = \cos^2(\theta/2) - \sin^2(\theta/2)$ et $\sin(\theta) = 2\cos(\theta/2)\sin(\theta/2)$.

Pour achever, nous devons trouver un vecteur w perpendiculaire au plan, et montrer qu'il est envoyé par A sur $-w$. Un vecteur $w = xu_1 + yu_2 + zu_3$ est perpendiculaire au plan si les deux égalités suivantes sont satisfaites :

$$\begin{aligned} \big(\cos(\theta/2)u_2 + \sin(\theta/2)u_3\big) \cdot (xu_1 + yu_2 + zu_3) = 0 && \text{(18.702a)} \\ u_1 \cdot (xu_1 + yu_2 + zu_3) = 0. && \text{(18.702b)} \end{aligned}$$

79. Les plus acharnées liront 4.6 pour comprendre la différence entre $\{u_1, u_2, u_3\}$ et (u_1, u_2, u_3).
80. Corolaire 18.15.

Nous avons immédiatement $x = 0$ et ensuite la relation

$$y\cos(\theta/2) + z\sin(\theta/2) = 0. \tag{18.703}$$

EQooXQMDooTvwrWk

Nous devons prouver que $A(w) = -w = -yu_2 - zu_3$, c'est-à-dire que

$$y\begin{pmatrix}0\\ \cos(\theta)\\ \sin(\theta)\end{pmatrix} + z\begin{pmatrix}0\\ \sin(\theta)\\ -\cos(\theta)\end{pmatrix} = \begin{pmatrix}0\\ -y\\ 0\end{pmatrix} + \begin{pmatrix}0\\ 0\\ -z\end{pmatrix}. \tag{18.704}$$

La première ligne est immédiate. Nous faisons la seconde, et nous laissons à la lectrice la troisième[81]. En utilisant les formules de duplication d'angle,

$$\begin{aligned} y\cos(\theta) + z\sin(\theta) &= y\big(\cos^2(\theta/2) - \sin^2(\theta/2)\big) + 2z\cos(\theta/2)\sin(\theta/2) && \text{(18.705a)}\\ &= y\big(2\cos^2(\theta/2) - \sin^2(\theta/2) - \cos^2(\theta/2)\big) + 2z\cos(\theta/2)\sin(\theta/2) && \text{(18.705b)}\\ &= -y + 2\cos(\theta/2)\big(y\cos(\theta/2) + z\sin(\theta/2)\big) && \text{(18.705c)}\\ &= -y && \text{(18.705d)} \end{aligned}$$

où pour la dernière ligne, nous avons utilisé la relation (18.703).

(4) **Si $\epsilon = 1$ et $\lambda = -1$** Alors la matrice est

$$A = \begin{pmatrix}-1 & 0 & 0\\ 0 & \cos(\theta) & -\sin(\theta)\\ 0 & \sin(\theta) & \cos(\theta)\end{pmatrix}. \tag{18.706}$$

Cela est la composition entre la symétrie de plan $\text{Span}(u_2, u_3)$ et la rotation d'angle θ dans ce plan.

18.12.2 Sous-groupes finis de SO(3)

LEMooWIMMooXOCfSt

Lemme 18.223 ([1]).
Points fixes pour SO(3).

(1) Tout élément de SO(3) *possède une droite de points fixes.*

(2) Tout élément non trivial de SO(3) *possède une seule droite de points fixes.*

Démonstration. Le polynôme caractéristique d'un élément de SO(3) est de degré trois et possède donc (en comptant les multiplicités), trois racines dont une réelle par le corolaire 12.91. Comme nous sommes en dimension impaire, le coefficient du terme de degré 3 est -1 et le polynôme caractéristique de $g \in \text{SO}(3)$ s'écrit

$$\chi_g(X) = -(X - \lambda_1)(X - \bar{\lambda}_1)(X - s) \tag{18.707}$$

avec $s = \pm 1$ que nous allons tout de suite fixer. Nous savons que $\det(g) = \chi_g(0)$ mais aussi que $\det(g) = 1$. Donc

$$1 = \det(g) = \chi_g(0) = \lambda_1\bar{\lambda}_1 s = s. \tag{18.708}$$

Tout cela pour dire que tout élément de SO(3) possède une valeur propre égale à 1, et donc une droite de points fixes.

Pour continuer, supposons que g possède deux droites distinctes de points fixes. En particulier g fixe un plan. Une base orthonormée de $\mathbb{R}^3$ peut être choisie en prenant deux vecteurs e_1, e_2 dans ce plan et un vecteur e_3 perpendiculaire au plan.

Puisque g est une isométrie, la base reste orthonormée sous l'action de g. Donc g a pour matrice

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \pm 1\end{pmatrix}. \tag{18.709}$$

Pour que le déterminant soit 1, il faut que la matrice soit l'identité. □

81. Je n'ai pas vérifié, alors faites attention.

PROPooBHPNooHPlgwH

Proposition 18.224 ([1, 560, 114, 561, 496]).
Les sous-groupes finis de SO(3) *sont :*

(1) *les groupes cycliques* $\mathbb{Z}/n\mathbb{Z}$,
(2) *les groupes diédraux* D_n,
(3) *le groupe alterné* A_4,
(4) *le groupe alterné* A_5
(5) *le groupe symétrique* S_4.

Démonstration. Soit G, un sous-groupe fini de SO(3). Par la proposition 8.46, les éléments de G sont des isométries de $\mathbb{R}^3$, et le lemme 18.223 dit que tout élément de G possède une droite de points fixes.

Un point de la sphère unité fixé par $g \in G$ est un **pôle** de g. Nous nommons Ω l'ensemble des pôles des éléments non triviaux de G.

(1) **Une action** Le groupe G agit sur Ω. En effet si $x \in \Omega$, alors x est fixé par un élément g. Montrons que $h(x)$ est également fixé par un élément de G. Pas dur : $(hgh^{-1})h(x) = h(x)$; donc $h(x)$ est un pôle de hgh^{-1}.

(2) **Les fixateurs sont cycliques** Nous montrons à présent que pour tout $x \in \Omega$, le sous-groupe $\mathrm{Fix}(x)$ est cyclique. Soit donc $x \in \Omega$, le plan orthogonal $\sigma = \mathrm{Span}(x)^\perp$ et $h \in \mathrm{Fix}(x)$. Nous avons $h(\sigma) = \sigma$. En effet si $y \in \sigma$ nous avons

$$0 = y \cdot x = h(y) \cdot h(x) = h(y) \cdot x, \tag{18.710}$$

donc $h(y)$ est perpendiculaire à x. L'inclusion inverse se démontre de même : si $y \in \sigma$ alors $y = h\big(h^{-1}(y)\big)$ alors que $h^{-1}(y) \in \sigma$.
La restriction de h à σ est une isométrie de σ. Prenant une isométrie $f : \sigma \to \mathbb{R}^2$, l'application

$$\begin{aligned}\varphi\colon \mathrm{Fix}(x) &\to \mathrm{SO}(2)\\ h &\mapsto f \circ h \circ f^{-1}.\end{aligned} \tag{18.711}$$

est un morphisme injectif de groupes. En effet nous avons d'une part

$$\varphi(hh') = f \circ h \circ h' \circ f^{-1} = fhf^{-1}fh'f^{-1} = \varphi(h)\varphi(h'), \tag{18.712}$$

d'où le morphisme. Et d'autre part, si $\varphi(h) = \varphi(h')$ alors $f \circ h \circ f^{-1} = f \circ h' \circ f^{-1}$, qui donne immédiatement $h = h'$.
Nous en déduisons que $\mathrm{Fix}(x)$ est isomorphe à un sous-groupe de $SO(2)$ (l'image de φ). Le lemme 18.142 en fait un groupe cyclique.

(3) **Taille des fixateurs** Soient Ω_i les orbites. Si $x, y \in \Omega_i$ alors nous montrons que $|\mathrm{Fix}(x)| = |\mathrm{Fix}(y)|$ avec la bijection

$$\begin{aligned}\varphi\colon \mathrm{Fix}(x) &\to \mathrm{Fix}(y)\\ h &\mapsto g^{-1}hg\end{aligned} \tag{18.713}$$

où g est choisie de façon à avoir $y = g(x)$ (possible parce que x et y sont dans la même orbite). Cela est surjectif parce que si $k \in \mathrm{Fix}(x)$ alors $k = \varphi(gkg^{-1})$ et l'on vérifie que $gkg^{-1} \in \mathrm{Fix}(y)$. L'application φ est également injective parce que si $ghg^{-1} = gh'g^{-1}$ alors $h = h'$.

(4) **Un peu de notations** Puisque tous les fixateurs des éléments d'une orbite ont la même taille (finie), nous pouvons noter

$$n_i = |\mathrm{Fix}(x_i)| \tag{18.714}$$

pour $x_i \in \Omega_i$. Nous notons également r le nombre d'orbites de G.
La formule de Burnside du théorème 2.40, avec les notations d'ici, donne

$$r = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Fix}(g)|. \tag{18.715}$$

(5) **Une belle formule** Soit l'ensemble

$$A = \{(g, x) \text{ tel que } g \in G\backslash\{e\}, x \in \mathrm{Fix}(g)\} \tag{18.716}$$

où par $\mathrm{Fix}(g)$ nous entendons les pôles de G fixés par g.
Il y a $|G|-1$ possibilités pour la composante g, mais chaque élément $g \neq e$ possède exactement deux pôles, donc l'ensemble A contient exactement $2(|G|-1)$ éléments.
Nous pouvons calculer le nombre d'éléments dans A d'une autre façon : pour chaque $x \in \Omega$ nous avons $|\mathrm{Fix}(x)| - 1$ éléments de $G\backslash\{e\}$ qui fixent x. Donc

$$|A| = \sum_{x\in\Omega} \big(|\mathrm{Fix}(x)| - 1\big). \tag{18.717}$$

Mais $|\mathrm{Fix}(x)|$ est constant sur les orbites. Nous coupons donc la somme sur Ω en plusieurs sommes sur les orbites Ω_i :

$$|A| = \sum_{i=1}^{r} \sum_{x\in\Omega_i} \big(|\mathrm{Fix}(x)| - 1\big) = \sum_i |\Omega_i|(n_i - 1). \tag{18.718}$$

En égalisant les deux façons de calculer $|A|$, nous déduisons la formule

$$2\big(|G| - 1\big) = \sum_i |\Omega_i|(n_i - 1). \tag{18.719}$$

Nous utilisons ensuite la relation orbite-stabilisateur, proposition 2.33 : $|\mathrm{Fix}(x_i)||\Omega_i| = |G|$; la formule (18.719) devient

$$2\big(|G| - 1\big) = \sum_i |G| - \sum_i \frac{|G|}{n_i} = r|G| + |G| \sum_i \frac{1}{n_i}, \tag{18.720}$$

ou encore, en simplifiant par $|G|$:

$$2 - \frac{2}{|G|} = r - \sum_i \frac{1}{n_i} = \sum_{i=1}^{r} \left(1 - \frac{1}{n_i}\right). \tag{18.721}$$

Nous pouvons aussi repartir de (18.719) et sommer de façon plus simple $\sum_i |\Omega_i| = |\Omega|$ pour obtenir

$$2\big(|G| - 1\big) = r|G| - |\Omega| \tag{18.722}$$

où Ω est l'ensemble des pôles de $G\backslash\{e\}$.

(6) **Quelles sont les possibilités ?** Les nombres $|G|$, r et n_i sont des entiers. Nous allons voir qu'il n'y a pas des centaines de possibilités pour satisfaire la relation (18.721). D'abord, pour toute valeur de $|G|$ (strictement plus grande que 1),

$$1 \leqslant 2 - \frac{2}{|G|} < 2. \tag{18.723}$$

Ensuite, si g fixe x alors g^{-1} fixe également x, de sorte que $n_i = |\mathrm{Fix}(x_i)| \geqslant 2$ pour tout i. Donc tous les termes dans la somme à droite de (18.721) sont dans $[\frac{1}{2}, 1[$. Nous avons donc au minimum deux termes, et au maximum trois. Autrement dit : $r = 2$ ou $r = 3$.

(7) **Si $r = 2$** Le plus simple est de repartir de (18.722). En posant $r = 2$ nous trouvons tout de suite $|\Omega| = 2$. Il y a donc exactement deux pôles pour l'action de G sur la sphère unité.
Tous les éléments de G laissent donc le même axe invariant et G est un sous-groupe des isométries du plan qui lui est perpendiculaire. Autrement dit, G est un sous-groupe fini de $\mathrm{SO}(2)$ et donc cyclique par le lemme 18.142.

Nous étudions à présent le cas $r = 3$. Vu que ça va être un peu long, nous sautons un niveau d'indentation.

(1) **Les possibilités pour $r = 3$** Nous devons voir les solutions entières $(n_1, n_2, n_3, |G|)$ de

$$2 - \frac{2}{|G|} = 3 - \frac{1}{n_1} - \frac{1}{n_2} - \frac{1}{n_3} < 2. \tag{18.724}$$

Il faut en particulier que

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} > 1, \tag{18.725}$$

ce qui signifie qu'au moins un des n_i doit être 1 ou 2, mais qu'il n'est pas possible que tous les n_i soient plus grands ou égaux à 3. Puisque $n_i = |\operatorname{Fix}(x_i)| \geqslant 2$, nous en déduisons qu'au moins un des n_i doit valoir 2. Nous posons donc $n_1 = 2$.

De plus, nous savons que les n_i doivent diviser $|G|$. Donc $|G|$ est pair.

(2) **Si $n_2 = 2$** Nous sommes dans le cas $r = 3$, $n_1 = 2$, $n_2 = 2$. Nous avons

$$\frac{1}{n_3} = \frac{2}{|G|}, \tag{18.726}$$

mais aussi $n_3 = |G|/|\Omega_3|$ d'où nous déduisons que $|\Omega_3| = 2$. Nous avons donc une orbite à deux éléments. Soit $\Omega_3 = \{x, y\}$ avec $x \neq y$.

Le groupe $\operatorname{Fix}(x)$ est un groupe à $|G|/2$ éléments. Il est donc normal par le lemme 3.28. Si $g \in G$ est tel que $g(x) = y$ alors nous avons $\operatorname{Fix}(y) = g\operatorname{Fix}(x)g^{-1}$, mais comme $\operatorname{Fix}(x)$ est normal nous avons $\operatorname{Fix}(x) = \operatorname{Fix}(y)$. Donc tous les éléments de $\operatorname{Fix}(x)$ fixent x et y. Le groupe $\operatorname{Fix}(x)$ est donc un sous-groupe de SO(2) et est cyclique, comme vu plus haut.

Mais de plus nous avons forcément $y = -x$ parce qu'un élément de G qui fixe un point fixe également l'opposé. Vu que $\Omega_3 = \{x, -x\}$, il existe $s \in G$ tel que $s(x) = -x$. Évidemment, s n'est pas dans $\operatorname{Fix}(x)$ et les points fixes de s ne sont pas parmi x et $-x$. Donc l'élément s^2 a au moins 4 points fixes : les deux de s ainsi que x et $-x$. Il a donc au moins deux droites de points fixes, et est donc l'identité : $s^2 = e$.

De plus, vu que $s(y)$ doit être égal soit à x soit à y, et que $s(x) = y$, l'injectivité de s donne $s(y) = x$.

Soit a, un générateur de $\operatorname{Fix}(x)$. Nous allons montrer que $G = \operatorname{gr}(s, sa)$. Nous avons déjà

$$(sa)(x) = s(x) = y \tag{18.727a}$$
$$(sa)(y) = s(y) = x. \tag{18.727b}$$

Donc sa inverse x et y. Mais sa a ses propres deux points fixes (qui ne sont ni x ni y). L'élément $(sa)^2$ a donc quatre points fixes sur la sphère unité : x, y et les deux de sa. Nous en déduisons que $(sa)^2 = e$.

Nous nous souvenons que a est un générateur $\operatorname{Fix}(x)$. Mais $a = s \cdot sa$, donc $a^k = (ssa)^k$. Nous en déduisons que $\operatorname{gr}(s, sa)$ contient au moins $\operatorname{Fix}(x)$.

D'autre part si h et h' sont des éléments distincts dans $\operatorname{Fix}(x)$, alors sh et sh' sont des éléments distincts de $\operatorname{gr}(s, sa)$ qui ne sont pas dans $\operatorname{Fix}(x)$. Autrement dit, la partie

$$A = \{sh \text{ tel que } h \in Fix(x)\} \tag{18.728}$$

est une partie de même cardinal que $\operatorname{Fix}(x)$ tout en n'ayant aucune intersection avec $\operatorname{Fix}(x)$ (note : l'identité n'est pas dans A). Mais $|\operatorname{Fix}(x)| = |G|/2$, donc $A \cup \operatorname{Fix}(x) = G$. Et justement $A \cup G \subset \operatorname{gr}(s, sa)$. Nous en déduisons que $\operatorname{gr}(s, sa) = G$.

Le théorème 18.186 nous assure que le groupe G est alors le groupe diédral parce que les éléments s et sa vérifient les relations données en 18.180.

(3) **Si $r = 3$, les autres cas possibles** Nous repartons de (18.721) en posant $r = 3$. Nous obtenons ceci :

$$1 + \frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}. \tag{18.729}$$

Nous avons déjà vu que $n_1 = 2$ était obligatoire, et que tous les cas où deux des n_i sont égaux à 2 sont déjà couverts. Donc n_2 et n_3 valent 3 ou plus.
Nous trions les n_i dans l'ordre croissant. Si $n_2 = 4$ ou plus, alors n_3 vaut 4 ou plus. Mais

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1 < 1 + \frac{2}{|G|}. \tag{18.730}$$

Donc $n_3 = 3$ est obligatoire. Nous avons alors l'inégalité suivante qui restreint n_3 :

$$\frac{1}{n_3} = \frac{1}{6} + \frac{3}{|G|} > \frac{1}{6}. \tag{18.731}$$

Donc n_3 est 3, 4 ou 5.
Les derniers cas à couvrir sont :

— $(n_1, n_2, n_3) = (2, 3, 3)$. Dans ce cas, $\frac{7}{6} = 1 + \frac{2}{|G|}$, donc $|G| = 12$.
— $(n_1, n_2, n_3) = (2, 3, 4)$. Dans ce cas, $|G| = 24$.
— $(n_1, n_2, n_3) = (2, 3, 5)$. Dans ce cas, $|G| = 60$.

(4) **Le cas** $(2, 3, 3)$
Nous utilisons les relations $n_i|\Omega_i| = |G|$ pour connaître la taille des orbites. Nous avons :
(4a) $2|\Omega_1| = 12$, donc $|\Omega_1| = 6$,
(4b) $3|\Omega_2| = 12$, donc $|\Omega_2| = 4$,
(4c) $3|\Omega_3| = 12$, donc $|\Omega_3| = 4$.
Nous avons $G \cdot \Omega_2 = \Omega_2$. D'une part parce que, par définition d'une orbite, $G \cdot \Omega_2 \subset \Omega_2$, et d'autre part parce que si $x \in \Omega_2$, alors $g^{-1}(x) \in \Omega_2$ et $g\big(g^{-1}(x)\big) = x$; donc Ω_2 est bien dans l'image de Ω_2 par G. Nous avons donc un morphisme $s\colon G \to S_{\Omega_2}$ que nous allons immédiatement voir comme

$$s\colon G \to S_4 \tag{18.732}$$

où S_4 est le groupe des permutations de $\{1, 2, 3, 4\}$.
Voyons que s est injective. Si $s(g) = s(h)$, alors $s(gh^{-1}) = \mathrm{Id}$. Autrement dit, l'élément sh^{-1} de G est l'identité sur Ω_2 qui contient 4 éléments. Fixant 4 points (au moins), l'élément sh^{-1} est l'identité. Par conséquent

$$s\colon G \to s(G) \subset S_4 \tag{18.733}$$

est un isomorphisme entre G et un sous-groupe de S_4. Mais $|G| = 12$ et $|S_4| = 24$, donc G est d'indice deux dans S_4 et est donc le groupe alterné A_4 par la proposition 5.45(3).

(5) **Le cas** $(2, 3, 4)$ Nous avons $|G = 24|$ et les orbites ont pour taille :
— $2|\Omega_1| = 24$, donc $|\Omega_1| = 12$,
— $3|\Omega_2| = 24$, donc $|\Omega_2| = 8$,
— $4|\Omega_3| = 24$, donc $|\Omega_3| = 6$.

(5a) Ω_2 **vient par paires** Soit $x \in \Omega$ tel que $|\operatorname{Fix}(x)| = 3$. Alors $x \in \Omega_2$ parce que x est forcément dans un des Ω_i et tout élément x_i de Ω_i vérifie $|\operatorname{Fix}(x_i)| = n_i$. Mais comme les éléments de $\mathrm{SO}(3)$ sont des applications linéaires, ceux qui fixent x fixent également $-x$. Cela pour dire que si $x \in \Omega_2$, alors $-x \in \Omega_2$. Nous avons donc quatre éléments distincts a_1, a_2, a_3 et a_4 tels que

$$\Omega_2 = \{\pm a_1, \pm a_2, \pm a_3, \pm a_4\}. \tag{18.734}$$

(5b) **Action sur les couples** Nous prétendons que G agit sur l'ensemble des couples $\{\pm a_i\}$. C'est encore la linéarité qui joue : l'élément $g(a_i)$ est forcément un des $\pm a_k$ (éventuellement $k = i$). Si $g(a_i) = a_k$, alors $g(-a_i) = -a_k$. Autrement dit, pour tout i, il existe un k tel que $g\big(\{a_i, -a_i\}\big) = \{a_k, -a_k\}$. Cette association $i \mapsto k$ est bijective (sinon g ne serait pas bijective), et fournit donc un morphisme de groupes

$$s\colon G \to S_4. \tag{18.735}$$

(5c) s **est injective** Nous prouvons à présent que $s(g) = \mathrm{Id}$ si et seulement si $g = e$. Dans un sens c'est évident : $g(e) = \mathrm{Id}$. Dans l'autre sens, nous devons prouver que si $g(a_i) \in \pm a_i$ pour tout i alors $g = e$.

Si $g(a_i) = a_i$ pour tout i, alors g stabilise 4 points et l'affaire est pliée. Nous supposons qu'au moins un des a_i n'est pas stabilisé par g. Pour fixer les idées nous disons que c'est a_1. Nous avons donc $g(a_1) = -a_1$. (oui : $g(a_1) = -a_1$ et non $\pm a_k$ pour un autre k parce que nous sommes sous l'hypothèse que g stabilise les couples)

L'élément g^2 fixe tout Ω_2 ; donc $g^2 = e$. Nommons $\pm b$ les points fixes de g. Si $b \in \Omega_2$ alors $|\mathrm{Fix}(b)| = 3$, c'est-à-dire que les éléments de G qui fixent b sont dans un groupe d'ordre 3, et le corolaire 2.14 nous indique que ces éléments ne peuvent être que d'ordre 1 ou 3, pas deux. Nous en déduisons que b n'est pas dans Ω_2 et donc que $g(a_i) = -a_i$ pour tout i.

Jusqu'à présent nous avons prouvé que si $g \in \ker(s)$ est non trivial, alors $g(a_i) = -a_i$ pour tout i.

Soit maintenant $h \in G$. Comme Ω_2 est une orbite, $h(a_i) \in \Omega_2$ et nous notons $h(a_i) = \epsilon a_k$ avec $\epsilon = \pm 1$ et éventuellement $k = i$ ou éventuellement pas. Nous avons :

$$(h^{-1}gh)(a_i) = \epsilon(h^{-1}g)(a_k) = -\epsilon h^{-1}(a_k) = -\epsilon^2 a_i = -a_i. \tag{18.736}$$

Donc g et $h^{-1}gh$ ont même restriction à Ω_2. En particulier $h^{-1}ghg^{-1}$ est l'identité sur Ω_2 et est donc l'identité.

Pour tout h nous avons $g = h^{-1}gh$. Les points fixes de $h^{-1}gh$ sont $\pm h^{-1}(b)$, mais aussi $\pm b$. Nous avons donc égalité d'ensemble $\{h(b), -h(b)\} = \{b, -b\}$ pour tout $h \in G$ (notez le changement de notation $h \to h^{-1}$). Cela signifie que $\{b, -b\}$ est une orbite de G. Maizon'a pas d'orbites de cardinal deux ; contradition. Nous en déduisons que e est l'unique élément de $\ker(s)$.

(5d) **Conclusion** La partie $s(G)$ est un sous-groupe de S_4 isomorphe à G. Mais au niveau des cardinaux, $|G| = 24$ en même temps que $|S_4| = 24$. Donc $G \simeq s(G) \simeq S_4$.

Nous passons au cas $(2, 3, 5)$, et comme ça va être long et douloureux [82], nous sautons un niveau d'indentation.

Au niveau du cardinal de G,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} = 1 + \frac{2}{|G|}, \tag{18.737}$$

donc $|G| = 60$. Et pour les orbites, $|\Omega_1| = 30$, $|\Omega_2| = 20$, $|\Omega_3| = 12$.

La proposition 5.55 nous indique que le seul groupe simple d'ordre 60 est le groupe A_5. Nous allons donc nous atteler à prouver que G est simple. Vous êtes prêts ?

(1) **Fixateurs et ordres** Tous les éléments de G sont dans un fixateur de type $\mathrm{Fix}(x)$, et comme l'ordre d'un élément divise l'ordre du groupe (corolaire 2.14), tous les éléments de G ont un ordre 2, 3 ou 5. Nous sommes dans un cas très particulier parce que

- Les trois nombres 2, 3 et 5 sont des nombres premiers distincts. Donc « diviser n_i » signifie pratiquement « être égal à n_i », surtout lorsqu'on parle de l'ordre d'un élément, qui ne peut pas être 1.
- Il existe une seule orbite de chaque taille.

Nous notons $G(n_i)$ l'ensemble des éléments de G d'ordre n_i. Les parties $G(n_i)$ ne contiennent pas l'identité.

(2) $g \in G(n_i)$ **implique** $\mathrm{Fix}(g) \subset \Omega_i$ Si $g \in G(n_i)$ et $x \in \mathrm{Fix}(g)$ alors $x \in \Omega_i$. En effet $x \in Fix(g)$ signifie $g(x) = x$ et donc aussi $g \in \mathrm{Fix}(x)$. Donc l'ordre de g divise $|Fix(x)|$, alors que l'ordre de g est n_i et que les possibilités pour $|\mathrm{Fix}(x)|$ sont exactement les n_i, lesquels sont premiers entre eux. Donc $|\mathrm{Fix}(x)| = n_i$ et $x \in \Omega_i$.

82. Mais pas autant que le théorème 32.33, cependant.

(3) $|\operatorname{Fix}(x)| = n_i$ **implique** $x \in \Omega_i$ Comme plus haut, $g \in \operatorname{Fix}(x)$ implique que l'ordre de g divise n_i et est donc égal à n_i. Autrement dit, $g \in G(n_i)$. De plus $g \in \operatorname{Fix}(x)$ implique $x \in \operatorname{Fix}(g)$. Par le cas juste au-dessus nous déduisons $x \in \Omega_i$.

(4) a **et** $-a$ **dans la même orbite**
Nous avons évidemment $\operatorname{Fix}(a) = \operatorname{Fix}(-a)$ parce que les éléments de G sont des applications linéaires. Si $|\operatorname{Fix}(a)| = n_i$ alors $a \in \Omega_i$ et aussi $|\operatorname{Fix}(-a)| = |\operatorname{Fix}(a)| = n_i$ et encore $-a \in \Omega_i$.

(5) **Nombre de** $\operatorname{Fix}(x_i)$ Soient $a, b \in \Omega_i$. Nous avons $|\operatorname{Fix}(a)| = |\operatorname{Fix}(b)| = n_i$ et $\operatorname{Fix}(a) = \operatorname{Fix}(b)$ si et seulement si $b = -a$ parce qu'un élément qui fixe a et b fixe automatiquement a, b, $-a$, et $-b$. Aucun élément non trivial ne peut fixer quatre points distincts. Autrement dit,

$$\operatorname{Fix}(a) \cap \operatorname{Fix}(b) = \begin{cases} \operatorname{Fix}(a) & \text{si } a = \pm b \\ \{e\} & \text{sinon.} \end{cases} \tag{18.738}$$

Chaque élément $x_i \in \Omega_i$ a son fixateur (il y en aurait $|\Omega_i| = 60/n_i$), mais ces fixateurs sont égaux deux à deux, donc il y a seulement $\frac{60}{2n_i}$ groupes distincts de la forme $|\operatorname{Fix}(x_i)|$ avec $x_i \in \Omega_i$.

(6) **Récapitulatif** En reprenant ce que nous venons de dire avec $i = 1, 2, 3$ nous trouvons :

(6a) $n_1 = 2$, avec $|\Omega_1| = 30$ et 15 groupes du type $\operatorname{Fix}(x_1)$ avec x_1 parcourant Ω_1.

(6b) $n_2 = 3$, avec $|\Omega_2| = 20$ et 10 groupes du type $\operatorname{Fix}(x_2)$ avec x_2 parcourant Ω_2.

(6c) $n_3 = 5$, avec $|\Omega_3| = 12$ et 6 groupes du type $\operatorname{Fix}(x_3)$ avec x_3 parcourant Ω_3.

Un élément non trivial de G se trouve forcément dans un et un seul de ces sous-groupes. Plus précisément, si $g \in G(n_i)$ alors g est dans un des $\operatorname{Fix}(x_i)$ avec $x_i \in \Omega_i$.
Comptons pour être sûr de ne pas s'être trompé. Chacune des lignes décrit 30 éléments de G ; par exemple, la seconde ligne donne 10 groupes de taille $|\operatorname{Fix}(x_2)| = n_2 = 3$. Mais tous ces groupes ont pour intersection exactement $\{e\}$. Donc le comptage des éléments se fait comme suit :

$$3 \times 30 - 15 - 10 - 6 + 1. \tag{18.739}$$

Le dernier $+1$ est parce que nous aurions décompté l'identité une fois de trop. Bref, on a bien 60 éléments comme il se doit.

(7) **Un ensemble à calculer deux fois** Soient les ensembles A_2, A_3 et A_5 définis par

$$A_i = \{(g, a) \in G(n_i) \times \Omega_i \text{ tel que } g(a) = a\} \tag{18.740}$$

où $G(n_i)$ est la partie de G des éléments d'ordre n_i.
Nous avons

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g) \cap \Omega_i|. \tag{18.741}$$

Mais les éléments de $G(n_i)$ sont d'ordre n_i, et par ce que nous avons dit plus haut, tous les éléments de $\operatorname{Fix}(g)$ sont dans Ω_i. Donc $\operatorname{Fix}(g) \cap \Omega_i = \operatorname{Fix}(g)$. Nous avons alors

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g)| = 2|G(n_i)| \tag{18.742}$$

parce que $|\operatorname{Fix}(g)| = 2$ pour tout g.
En comptant $|A_i|$ dans l'autre sens, nous avons

$$|A_i| = \sum_{x \in \Omega_i} |\operatorname{Fix}(x) \cap G(n_i)| \tag{18.743}$$

EQooBHIJooVcGgEd

Vu que $x \in \Omega_i$, les éléments de $\operatorname{Fix}(x)$ sont d'ordre n_i [83] (sauf e), et comme $G(n_i)$ est justement l'ensemble des éléments d'ordre n_i dans G nous avons $\operatorname{Fix}(x) \cap G(n_i) = \operatorname{Fix}(x) \backslash \{e\}$. Cela

83. Encore et toujours parce que les éléments de $\operatorname{Fix}(x)$ ont un ordre qui divise $|\operatorname{Fix}(x)| = n_i$ et que n_i est premier, et que nous avons exclu l'identité.

pour dire que

$$|A_i| = \sum_{x\in\Omega_i} \Big(|\operatorname{Fix}(x)| - 1\Big) \tag{18.744a}$$
$$= \sum_{x\in\Omega_i} |\operatorname{Fix}(x)| - \sum_{x\in\Omega_i} 1 \tag{18.744b}$$
$$= \sum_{x\in\Omega_i} n_i - |\Omega_i| \qquad |\operatorname{Fix}(x)| = n_i \text{ pcq } x \in \Omega_i \tag{18.744c}$$
$$= |\Omega_i|n_i - |\Omega_i| = |G| - |\Omega_i|. \tag{18.744d}$$

En égalisant cela à la valeur $2|G(n_i)|$ déjà trouvée, nous déduisons les valeurs des $|G(n_i)|$:

$$|G(n_i)| = \frac{|G| - |\Omega_i|}{2}. \tag{18.745}$$

Nous avons alors

(7a) $|G(2)| = 15$

(7b) $|G(3)| = 20$

(7c) $|G(5)| = 24$

(8) **Les Sylow de G** Les p-Sylow sont définis en 5.6, et le super théorème qui répond à toutes les questions est le théorème 5.11. Dans notre cas, les diviseurs premiers de $|G| = 60$ sont 2, 3 et 5. Il faut faire attention au 2 parce que sa plus haute puissance dans la décomposition de 60 est 4 et non 2. Nous avons :

(8a) Un 2-Sylow est un sous-groupe d'ordre 4.

(8b) Un 3-Sylow est un sous-groupe d'ordre 3.

(8c) Un 5-Sylow est un sous-groupe d'ordre 5.

Entre autres :

(8a) Les 10 sous-groupes $\operatorname{Fix}(x_2)$ avec $x_2 \in \Omega_2$ sont des 3-Sylow de G.

(8b) Les 6 sous-groupes $\operatorname{Fix}(x_3)$ avec $x_3 \in \Omega_3$ sont des 5-Sylow de G.

(8c) Les 15 sous-groupes $\operatorname{Fix}(x_1)$ avec $x_1 \in \Omega_1$ sont d'ordre 2 et ne sont donc pas des 2-Sylow de G.

(9) **Tous les 3-Sylow et les 5-Sylow** Nous avons déjà trouvé 10 3-Sylow et 6 5-Sylow. Nous montrons à présent qu'il n'y en a pas d'autres. Le théorème de Sylow 5.11(4) nous indique que le nombre n_3 de 3-Sylow est :

— diviseur de 60,

— dans $[1]_3$

— au moins 10.

Les diviseurs de 60 sont :

$$1, 5, 3, 15, 2, 10, 6, 30, 4, 20, 12, 60. \tag{18.746}$$

Le seul qui vérifie toutes les conditions est 10. Donc G possède seulement 10 3-Sylow et ils sont tous de la forme $\operatorname{Fix}(x_2)$ avec $x_2 \in \Omega_2$.

Même raisonnement pour les 5-Sylow : il faut

— diviseur de 60,

— dans $[1]_5$

— au moins 6.

La seule possibilité est 6.

(10) **Sous-groupe normal** Soit H, un sous-groupe normal de G. Notre but étant de prouver que G est simple, nous voulons prouver que H est soit $\{e\}$ soit G. Nous supposons que H est non trivial, et nous allons prouver que $H = G$.

Le théorème de Lagrange 2.13(1) nous dit que $|H|$ divise $|G|$. Le nombre $|H|$ ne peut donc avoir que 2, 3 et 5 comme facteurs premiers. Avec une mention spéciale pour le 2 : $|H|$ pourrait être divisible aussi par 4.

(11) **Diviseurs de $|H|$**
Soit un sous-groupe normal H de G. Puisque c'est un sous-groupe, son ordre divise celui de G (encore et toujours le théorème de Lagrange 2.13), et donc les facteurs premiers de $|H|$ ne peuvent être que 2, 3 et 5.

(12) **Si $|H|$ est divisible en 3** Alors H contient au moins un 3-Sylow. Mais nous avons vu que les 3-Sylow de H sont les 3-Sylow de G. Donc H contient tous les 3-Sylow de G, parce que les 3-Sylow sont conjugués et H est normal.

Soit E l'ensemble des sous-groupes de H. Puisqu'il est normal, H agit sur E par conjugaison, et les 3-Sylow forment une orbite. Si α est un 3-Sylow, la formule des classes (proposition 2.33(2)) nous donne

$$|H| = |\operatorname{Fix}(\alpha)||\mathcal{O}_\alpha|. \tag{18.747}$$

Mais l'orbite $\mathcal{O}_\alpha$ de α est l'ensemble des 3-Sylow, de sorte que $|\mathcal{O}_\alpha| = 10$. Donc $|H|$ est divisible en 10.

Mais il y a pire : H contient au moins les 10 sous-groupes $\operatorname{Fix}(x_2)$ pour $x_2 \in \Omega_2$. Ce sont 10 groupes de $|\operatorname{Fix}(x_2)| = 3$ éléments. En décomptant e qui est dans l'intersection, cela fait

$$10 \times |\operatorname{Fix}(x_2)| - 10 + 1 = 21 \tag{18.748}$$

éléments. Donc H contient au moins 21 éléments. Le nombre $|H|$ est donc :

— diviseur de 60
— multiple de 10
— au moins 21.

Donc c'est 30 ou 60.

(13) **Si $|H|$ est divisible en 5** Le même raisonnement tient et $|H|$ est 30 ou 60.

Nous restons avec les possibilités $|H|$ égal à 2, 4, 30 ou 60.

(1) **Si $|H| = 4$**
Alors H contient au moins un 2-Sylow. Un 2-Sylow de H est un sous-groupe contenant 4 éléments qui sont d'ordre 2^m. Le seul m possible dans G est $m = 1$. Vu qu'un 2-Sylow de H contient 4 éléments, nous sommes dans le cas où H est un 2-Sylow. Il est donc le seul 2-Sylow de H parce que H est normal, et que tous les 2-Sylow sont conjugués.

Mais tous les sous-groupes d'ordre 2 sont contenus dans un 2-Sylow. En particulier tous les 15 groupes $\operatorname{Fix}(x_1)$ sont dans l'unique 2-Sylow H qui est soi-disant d'ordre 4. IL y a là une belle impossibilité.

Donc, le cas $|H| = 4$ est hors-concours.

(2) **Si $|H| = 2$** Alors $H = \{e, g\}$ avec $g^2 = e$. Si $h \in G$, l'élément hgh^{-1} ne peut être que e ou g (parce que H est normal). Le premier cas est $g = e$, et le second donne $gh = hg$. Donc g est dans le centre de G : il commute avec tous les éléments de G.

Comme $g \in G(2)$, nous savons que les éléments $a \in \operatorname{Fix}(g)$ sont forcément dans Ω_1 parce que les points dont les fixateurs sont formés d'éléments d'ordre 2 sont dans Ω_1. Soit $h \in G$. Nous avons $g = hgh^{-1}$ et donc aussi

$$\big(hgh^{-1}\big)\big(h(a)\big) = hg(a) = h(a), \tag{18.749}$$

donc $h(a)$ et $-h(a)$ sont des points fixes de hgh^{-1}. Ce sont donc également des points fixes de g. Nous en déduisons que g a pour points fixes les points a, $-a$, $h(a)$ et $-h(a)$. Puisque

g n'est pas e, ces quatre points ne peuvent pas être distincts. Comme $h(a)$ ne peut pas être $-h(a)$, nous avons forcément $h(a) = \pm a$.

Donc l'orbite de a ne contiendrait que 2 éléments. Pas possible.

(3) **Si $|H| = 30$** À part $|H| = 60$, le dernier cas à traiter est $|H| = 30$. Nous rappelons obligeamment que

(3a) $|G(2)| = 15$

(3b) $|G(3)| = 20$

(3c) $|G(5)| = 24$.

Si H possède 30 éléments, le théorème de Sylow dit que H contient au moins un 3-Sylow et un 5-Sylow, et donc tous. Puisque pour 3 et 5, les Sylow de H et de G sont les mêmes et bien identifiés, nous allons nous baser dessus. Le sous-groupe H contient tous les 3 et 5-Sylow, donc le comptage des éléments est :

$$10 \times |\operatorname{Fix}(x_2)| + 6 \times |\operatorname{Fix}(x_3)| - 15 = 45. \tag{18.750}$$

Nous aurions aussi pu ajouter $+4 - 1$ pour compter au moins un 2-Sylow.

Donc dès que H compte 30 éléments, il en compte au moins 45. Autrement dit le cas $|H| = 30$ est impossible.

□

¡¡ Avertissement/question à la lectrice !! 18.225
La démonstration des groupes finis de SO(3) *est longue. Je me demande si il n'y a pas moyen de faire plus court. Par exemple [560] utilise le théorème de Cauchy 3.26 que je n'utilise pas. D'autre part, toutes les références me semblent utiliser plus ou moins implicitement le fait que si le sous-groupe normal H contient un élément de $G(n_i)$, alors il les contiennent tous. J'avoue ne pas trop comprendre pourquoi.*

18.13 Systèmes de coordonnées

SECooWTPRooZbOSzO

La trigonométrie nous offre de nouveaux systèmes de coordonnées qui peuvent se révéler pratiques dans certains cas : les coordonnées polaires sur $\mathbb{R}^2$, ainsi que les coordonnées cylindriques et sphériques sur $\mathbb{R}^3$.

18.13.1 Coordonnées polaires

18.13.1.1 Ce que ça signifie intuitivement

On a vu qu'un point M dans $\mathbb{R}^2$ peut être représenté par ses abscisses x et ses ordonnées y. Nous pouvons également déterminer le même point M en donnant un angle et une distance comme illustré sur la figure 18.10.

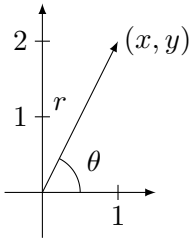


FIGURE 18.10: Un point en coordonnées polaires est donné par sa distance à l'origine et par l'angle qu'il fait avec l'horizontale.

LabelFigJWINooSfKCeA

Le même point M peut être décrit indifféremment avec les coordonnées (x, y) ou bien avec (r, θ).

Remarque 18.226.
L'angle θ d'un point n'étant à priori défini qu'à un multiple de 2π près, nous convenons de toujours choisir un angle $0 \leqslant \theta < 2\pi$. Par ailleurs l'angle θ n'est pas défini si $(x, y) = (0, 0)$.

La coordonnée r est toujours positive.

Nous avons dans l'idée de définir r et θ par les formules [EqrthetaxyPoal]

$$\begin{cases} x = r\cos(\theta) & (18.751a) \\ y = r\sin(\theta). & (18.751b) \end{cases}$$

18.13.1.2 Coordonnées polaires : le théorème

THOooBETSooXSQhdX

Théorème 18.227 (Coordonnées polaires[1])**.**
Soit l'application

$$\begin{aligned} T\colon [0,\infty[\times [0,2\pi[&\to \mathbb{R}^2 \\ (r,\theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{18.752}$$

ITEMooNGOKooFCXmwy

(1) L'application T est surjective.

ITEMooMCIOooJiBvug

(2) L'application

$$T\colon]0,\infty[\times [0,2\pi[\to \mathbb{R}^2\backslash\{(0,0)\} \tag{18.753}$$

est bijective.

ITEMooZFRGooQPDUtX

(3) En considérant la demi-droite $D = \{(x, 0)\}_{x\geqslant 0}$, l'application

$$T\colon]0,\infty[\times]0,2\pi[\to \mathbb{R}^2\backslash D \tag{18.754}$$

est un C^∞-difféomorphisme[84]*.*

Démonstration. Une bonne partie de ce théorème est une conséquence de 18.58. La vraie nouveauté de ce théorème sera la régularité. Nous démontrons point par point.

(1) Pour (1). Soit $a = (x, y) \in \mathbb{R}^2$. Nous avons $a/\|a\| \in S^1$. Par la proposition 18.58, il existe $\theta \in [0, 2\pi]$ tel que

$$\frac{a}{\|a\|} = \big(\cos(\theta), \sin(\theta)\big). \tag{18.755}$$

Alors $a = \|a\|\big(\cos(\theta), \sin(\theta)\big) = T(\|a\|, \theta)$. Voilà. L'application T est surjective.

(2) Pour (2). En ce qui concerne la surjectivité,

$$T\big(0, [0, 2\pi[\big) = \{(0, 0)\}. \tag{18.756}$$

Donc le point (1) donne le surjectif lorsque nous enlevons d'un côté les points avec $r = 0$ et de l'autre le point $(0, 0)$.

Pour l'injectivité, nous supposons $T(r_1, \theta_1) = T(r_2, \theta_2)$. Puisque $\|T(r, \theta)\| = r$, nous avons tout de suite $r_1 = r_2$. Nous restons donc avec l'égalité

$$\begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} = \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix}. \tag{18.757}$$

La proposition 18.58 dit alors que $\theta_1 = \theta_2$.

84. L'application est de classe C^∞ et son inverse est également de classe C^∞. Le plus souvent, vous voulez seulement utiliser ce théorème dans le but de faire un changement de variables dans une intégrale ; vous n'avez donc besoin que d'un C^1-difféomorphisme.

(3) Pour (3). L'application T est injective en tant que restriction d'une application injective. Pour le surjectif, soit $a \in \mathbb{R}^2 \backslash D$. Comme $a \notin D$, nous avons $\|a\| \neq 0$ et il est légitime de dire, comme plus haut, qu'il existe $\theta \in [0, 2\pi[$ tel que

$$\frac{a}{\|a\|} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \tag{18.758}$$

Ce θ n'est pas zéro parce que $\theta = 0$ donne le point $(1, 0)$ qui est sur D.

En ce qui concerne l'inverse, nous n'allons pas nous lancer dans une étude subtile de la fonction (18.179) ; nous avons déjà démontré la continuité dans le lemme 18.65, et monter dans les dérivées nous semble un peu compliqué. Au lieu de cela, nous allons procéder en deux étapes :

— Prouver que T est de classe C^p pour tout p en invoquant seulement des théorèmes à proposition de différentielle,

— En déduire que T^{-1} est également C^p pour tout p en invoquant le théorème d'inversion locale 17.47.

Les applications $(r, \theta) \mapsto r$, $(r, \theta) \mapsto \sin(\theta)$ et $(r, \theta) \mapsto \cos(\theta)$ sont de classe C^∞ grâce au lemme 12.248. Le lemme 12.270 sur la différentiabilité du produit montre alors que les fonctions T_1 et T_2 données par

$$T_1(r, \theta) = r\cos(\theta) \tag{18.759a}$$
$$T_2(r, \theta) = r\sin(\theta) \tag{18.759b}$$

sont différentiables [85]. Mieux, la proposition 12.273 montre que ces fonctions T_1 et T_2 sont de classe C^p pour tout p, c'est-à-dire qu'elles sont de classe C^∞. Cela montre que les coordonnées polaires sont de classe C^∞, et il faut encore parler de l'inverse.

En ce qui concerne la différentielle,

$$dT_{(r,\theta)}(u, v) = \begin{pmatrix} u\cos(\theta) - rv\sin(\theta) \\ u\sin(\theta) + rv\cos(\theta) \end{pmatrix}. \tag{18.760}$$

Donc la matrice de la différentielle est

$$dT_{(r,\theta)} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta), \end{pmatrix} \tag{18.761}$$

dont le déterminant est r (lemme 18.4 utilisé). Donc la différentielle en (r, θ) est une application linéaire inversible parce que $r \neq 0$ aux points que nous considérons. L'application $dT_{(r,\theta)}$ est bicontinue parce que nous sommes en dimension finie. Tout cela pour dire que le théorème d'inversion local 17.47 s'applique, et T^{-1} est C^p dès que T est C^p.

Puisque T est de classe C^p pour tout p, l'inverse T^{-1} est également C^p pour tout p, c'est-à-dire que T^{-1} est de classe C^∞.

□

Définition 18.228.
Ce que nous appelons ***les coordonnées polaires*** *est l'application*

$$\begin{aligned} T\colon [0, \infty[\times [0, 2\pi[&\to \mathbb{R}^2 \\ (r, \theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{18.762}$$

du théorème 18.227(3). Selon les circonstances, nous considérons l'une ou l'autre des restrictions pour avoir une bijection ou un difféomorphisme.

85. Si vous voulez seulement avoir un C^1-difféomorphisme, calculez explicitement la différentielle et montrez que c'est continu. Vous n'avez pas à utiliser la proposition 12.273 ni rien des produits tensoriels.

EXooSDHDooJzDioW

Exemple 18.229.
Soit à calculer

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{x-y}. \tag{18.763}$$

Nous introduisons la fonction

$$\begin{aligned} f\colon \mathbb{R}^2\backslash\{x=y\} &\to \mathbb{R} \\ (x,y) &\mapsto \frac{x^2+y^2}{x-y}. \end{aligned} \tag{18.764}$$

Une idée souvent fructueuse pour traiter ce genre de limite est de passer aux coordonnées polaires. Attention, si on veut faire les choses très explicitement, c'est un peu lourd en notations. Il s'agit de poser

$$\begin{aligned} g\colon \big]0,\infty[\times[0,2\pi[\big)\backslash\big\{\mathbb{R}\times\{\frac{\pi}{4}\}\cup\mathbb{R}\times\{\frac{5\pi}{4}\}\big\} &\to \mathbb{R} \\ (r,\theta) &\mapsto \frac{r^2}{r\big(\cos(\theta)-\sin(\theta)\big)}. \end{aligned} \tag{18.765}$$

Bon. À strictement parler, nous aurions pu dire que g est définie pour $r=0$, mais vu que nous voulons seulement calculer la limite pour $r\to 0$, on n'a pas besoin de la valeur en zéro. De plus les coordonnées polaires ne sont pas bijectives en l'origine. Donc bon . . .on s'en passe.

Quel est le lien entre f et g ? Du point de vue du calcul, le lien est qu'on a remplacé x par $r\cos(\theta)$ et y par $r\sin(\theta)$. Le vrai lien est l'égalité

$$g=f\circ T \tag{18.766}$$

où T est l'application de coordonnées polaires dont les principales propriétés sont données dans le théorème 18.227(2).

Soit un voisinage $B\big((0,0),R\big)$ de $(0,0)$ dans $\mathbb{R}^2$. Le but est de montrer que les valeurs $f(B)$ se regroupent autour d'une valeur ℓ lorsque $R\to 0$. Soyons plus précis et nommons ℓ le candidat limite. Soit $\epsilon>0$; nous devons trouver $R>0$ tel que $f\Big(B\big((0,0),R\big)\Big)\subset B(\ell,\epsilon)$.

Pour $R>0$, nous avons

$$B\big((0,0),R\big)=T\big([0,R[\times[0,2\pi[\big), \tag{18.767}$$

donc

$$f(B)=g\big([0,R[\times[0,2\pi[\big). \tag{18.768}$$

Soit $r<R$. Nous avons

$$\lim_{\theta\to\pi/4} g(r,\theta)=\infty. \tag{18.769}$$

Donc $f(B)$ contient des valeurs arbitrairement grandes, quelle que soit la valeur de R. Il n'y a donc pas de limite possible.

Si vous voulez un argument un peu plus imagé, en voici un [86] basé sur une combinaison entre la méthode des coordonnées polaires et la méthode des chemins.

Certes *pour chaque* θ nous avons $\lim_{r\to 0} g(r,\theta)=0$, mais il ne faut pas en déduire trop vite que la limite $\lim_{(x,y)\to(0,0)} g(x,y)$ vaut zéro parce que prendre la limite $r\to 0$ avec θ fixé revient à prendre la limite le long de la droite d'angle θ.

Il n'est pas possible de majorer $g(r,\theta)$ par une fonction ne dépendant pas de θ parce que cette fonction tend vers l'infini lorsque $\theta\to\pi/4$. Est-ce que cela veut dire que la limite n'existe pas ? Cela veut en tout cas dire que la méthode des coordonnées polaires ne parvient pas à résoudre l'exercice. Pour conclure, il faudra encore un peu travailler.

86. Qui satisfera tous vos professeurs, pourvu que vous ayez compris que ce qui se cache est une histoire de valeurs de f prises sur un voisinage de $(0,0)$.

Nous pouvons essayer de calculer le long d'un chemin plus général $(r(t), \theta(t))$. Choisissons $r(t) = t$ puis cherchons $\theta(t)$ de telle sorte à avoir

$$\cos\theta(t) - \sin\theta(t) = t^2. \tag{18.770}$$

Le mieux serait de résoudre cette équation pour trouver $\theta(t)$. Mais en réalité il n'est pas nécessaire de résoudre : montrer qu'il existe une solution suffit. Nous pouvons supposer que $t^2 < 1$. Pour $\theta = \pi/4$ nous avons $\cos(\theta) - \sin(\theta) = 0$ et pour $\theta = 0$ nous avons $\cos(\theta) - \sin(\theta) = 1$. Le théorème des valeurs intermédiaires nous enseigne alors qu'il existe une valeur de θ qui résout l'équation (18.770).

Pour être rigoureux, nous devons aussi montrer que la fonction $\theta(t)$ est continue. Pour cela il faudrait utiliser le théorème de la fonction implicite 17.49. Nous verrons dans l'exemple 18.245 comment s'en sortir sans théorème de la fonction implicite, au prix de plus de calculs. △

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0,\infty[\times]0,2\pi[&\to \mathbb{R}^2\backslash D \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{18.771}$$

où D est la demi-droite $y = 0$, $x \geqslant 0$. Le fait que les coordonnées polaires ne soient pas un difféomorphisme sur tout $\mathbb{R}^2$ n'est pas un problème pour l'intégration parce que le manque de difféomorphisme est de mesure nulle dans $\mathbb{R}^2$. Le jacobien est donné par

$$Jg = \det\begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{18.772}$$

La fonction qui donne les coordonnées polaires est

$$\begin{aligned} \varphi\colon \mathbb{R}^+ \times]0,2\pi[&\to \mathbb{R}^2 \\ (r,\theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{18.773}$$

Son Jacobien vaut

$$J_\varphi(r,\theta) = \det\begin{pmatrix} \frac{\partial x(r,\theta)}{\partial r} & \frac{\partial x(r,\theta)}{\partial \theta} \\ \frac{\partial y(r,\theta)}{\partial r} & \frac{\partial y(r,\theta)}{\partial \theta} \end{pmatrix} = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r. \tag{18.774}$$

Proposition 18.230.
Soit la fonction

$$\begin{aligned} T\colon]0,+\infty[\times \mathbb{R} &\to \mathbb{R}^2\backslash\{(0,0)\} \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big). \end{aligned} \tag{18.775}$$

(1) Elle est surjective.
(2) Pour tout $a \in \mathbb{R}$, l'application T est bijective sur la bande $]0,+\infty[\times [a-\pi, a+\pi[$.
(3) Si $a = 0$, la fonction inverse est donnée par

$$T^{-1}(x,y) = \big(\sqrt{x^2+y^2}, \arctan(y/x)\big). \tag{18.776}$$

Soit $P = (x,y)$ un élément dans $\mathbb{R}^2$, on dit que $r = \sqrt{x^2+y^2}$ est le rayon de P et que $\theta = \arctan(y/x)$ est son argument principal. L'origine ne peut pas être décrite en coordonnées polaires parce que si son rayon est manifestement zéro, on ne peut pas lui associer une valeur univoque de l'angle θ.

Exemple 18.231.
L'équation du cercle de rayon a et centre $(0,0)$ en coordonnées polaires est $r = a$. △

Exemple 18.232.
Une équation possible pour la demi-droite $x = y$, $x > 0$, est $\theta = \pi/4$. △

18.13.1.3 Polaires : calcul de limites

EXooBEROooPhPsSU

Exemple 18.233.
Calculer les limites suivantes :

(1) $\lim_{(x,y)\to(0,0)} \frac{(xy)^2}{(x+y)^2+(x-y)^2}$
(2) $\lim_{(x,y)\to(0,0)} \frac{xy^3}{x^2+y^2}$
(3) $\lim_{(x,y)\to(0,0)} \frac{x\sin(y)}{\sqrt{x^2+y^2}}$

Tentez de les faire par vous-même avant de regarder la solution qui suit.

(1) Si $(x,y) \in B(0,\delta)\backslash\{(0,0)\}$, alors le théorème 18.227 dit qu'il existe $0 < r < \delta$ et $\theta \in [0, 2\pi[$ tels que $x = r\cos(\theta)$ et $y = r\sin(\theta)$. Avec ça nous avons

$$f(x,y) = \frac{r^4\cos^2(\theta)\sin^2(\theta)}{r^2\big(\cos(\theta)+\sin(\theta)\big)^2 + r^2\big(\cos(\theta)-\sin(\theta)\big)^2}. \tag{18.777}$$

Vu que $r \neq 0$ nous pouvons simplifier par r^2. Ce qui reste au dénominateur est une fonction de θ continue sur le compact $[0, 2\pi]$. Je ne peux pas vous dire quel est le minimum, mais il est facile de voir que ce n'est pas zéro. Bref,

$$f(x,y) = r^2 f(\theta) \tag{18.778}$$

et $|f(x,y)| < ar^2$ où a est une constante que vous pouvez vous amuser à calculer. En posant $\delta < \sqrt{\epsilon/a}$ nous avons $|f(x,y)| < \epsilon$ pour tout $(x,y) \in B(0,\delta)$, ce qui signifie que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

(2) Regardons la technique des coordonnées polaires. Nous remplaçons x par $r\cos(\theta)$ et y par $r\sin(\theta)$:

$$f(r,\theta) = \frac{r^4\cos(\theta)\sin^3(\theta)}{r^2} = r^2\cos(\theta)\sin^3(\theta). \tag{18.779}$$

Cette fonction tend vers zéro quand $r \to 0$. Nous avons donc

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0. \tag{18.780}$$

Pour cet exercice nous pouvons aussi utiliser la règle de l'étau en écrivant d'abord

$$0 \leqslant |f(x,y)| \leqslant \frac{|x||y^3|}{|x^2+y^2|}. \tag{18.781}$$

Mais on a $|x| \leqslant \sqrt{x^2+y^2}$, $|y| \leqslant \sqrt{x^2+y^2}$ et $|x^2+y^2| = \big(\sqrt{x^2+y^2}\big)^2$, donc

$$0 \leqslant |f(x,y)| \leqslant \frac{\sqrt{x^2+y^2}\big(\sqrt{x^2+y^2}\big)^3}{\big(\sqrt{x^2+y^2}\big)^2} = \big(\sqrt{x^2+y^2}\big)^2 \to 0. \tag{18.782}$$

(3) En passant aux polaires, nous avons

$$f(r,\theta) = \frac{r\cos\theta\sin\big(r\sin\theta\big)}{r} = \cos(\theta)\sin\big(r\sin\theta\big). \tag{18.783}$$

La limite de cette dernière fonction lorsque $r \to 0$ vaut zéro.
Une autre façon de procéder consiste à multiplier et diviser par y de telle façon à faire apparaitre $\sin(y)/y$ dont nous connaissons la limite :

$$f(x,y) = \frac{\sin(y)}{y} \cdot \frac{xy}{\sqrt{x^2+y^2}}. \tag{18.784}$$

La limite du premier facteur est 1, tandis que le second peut être traité de façon classique en prenant la valeur absolue et en majorant $|x|$ par $\sqrt{x^2+y^2}$.

△

18.13.1.4 Transformation inverse : théorie

Voyons la question inverse : comment retrouver r et θ si on connait x et y ? Tout d'abord,

$$r = \sqrt{x^2 + y^2} \tag{18.785}$$

parce que la coordonnée r est la distance entre l'origine et (x, y). Comment trouver l'angle ? Nous supposons $(x, y) \neq (0, 0)$. Si $x = 0$, alors le point est sur l'axe vertical et nous avons

$$\theta = \begin{cases} \pi/2 & \text{si } y > 0 \\ 3\pi/2 & \text{si } y < 0 \end{cases} \tag{18.786}$$

Notez que si $y < 0$, conformément à notre convention $\theta \geqslant 0$, nous avons noté $\frac{3\pi}{2}$ et non $-\frac{\pi}{2}$.

Supposons maintenant le cas général avec $x \neq 0$. Les équations (18.751) montrent que

$$\tan(\theta) = \frac{y}{x}. \tag{18.787}$$

Nous avons donc

$$\theta = \tan^{-1}\left(\frac{y}{x}\right). \tag{18.788}$$

La fonction inverse de la fonction tangente est celle définie plus haut.

18.13.1.5 Transformation inverse : pratique

Le code suivant utilise Sage.

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

def PointToPolaire(x,y):
  x=SR(x)
  y=SR(y)
  r = sqrt(x**2+y**2)
  if x == 0:
    if y > 0:
      alpha = pi/2
    if y < 0:
      alpha = 3*pi/2
    if y == 0 :
      raise ValueError,"Pas d'angle pour le point (0,0) !!"
  else :
    alpha = atan(y/x)
  if (x < 0) and (y == 0) :
    alpha = pi
  if (x < 0) and (y > 0) :
    alpha = alpha + pi
  if (x < 0) and (y < 0 ) :
    alpha = alpha + pi
  alpha=alpha.simplify_trig()
  return (r,alpha)

print PointToPolaire(1,1)
print PointToPolaire(-2,1)
```

```
30 print PointToPolaire(6*sqrt(3)/2,3)
```

tex/frido/calculAngle.py

Son exécution retourne :

```
(sqrt(2), 1/4*pi)
(sqrt(5), pi - arctan(1/2))
(6, 1/6*pi)
```

Notez que ce sont des valeurs *exactes*. Ce ne sont pas des approximations, Sage travaille de façon symbolique.

18.13.1.6 Coordonnées polaires : dérivées partielles

Le changement de coordonnées pour les coordonnées polaires est la fonction

$$f\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}. \tag{18.789}$$

Considérons une fonction g sur $\mathbb{R}^2$, et définissons la fonction $\tilde{g}$ par

$$\tilde{g}(r,\theta) = g(r\cos\theta, r\sin\theta). \tag{18.790}$$

La formule (12.744) permet de trouver les dérivées partielles de g par rapport à r et θ en termes de celles par rapport à x et y de g.

Pour faire le lien avec les notations du point précédent, nous avons

$$\begin{aligned} f_1(r,\theta) &= r\cos(\theta) \\ f_2(r,\theta) &= r\sin(\theta) \\ (x_1,x_2) &\to (r,\theta) \\ (y_1,y_2) &\to (x,y). \end{aligned} \tag{18.791}$$

Nous avons donc

$$\begin{aligned} \frac{\partial\tilde{g}}{\partial r}(r,\theta) &= \sum_{i=1}^{2} \frac{\partial g}{\partial x_i}\big(f(r,\theta)\big)\frac{\partial f_i}{\partial r}(r,\theta) \\ &= \frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\cos\theta\big)}{\partial r}(r,\theta) \\ &\quad + \frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\sin\theta\big)}{\partial r}(r,\theta) \\ &= \cos\theta\frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta). \end{aligned} \tag{18.792}$$

Prenons par exemple $g(x,y) = \frac{1}{x^2+y^2}$. Étant donné que

$$\frac{\partial g}{\partial x} = \frac{-2x}{(x^2+y^2)^2}, \tag{18.793}$$

nous avons

$$\frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta) = \frac{-2\cos\theta}{r^3}. \tag{18.794}$$

En utilisant la formule,

$$\frac{\partial\tilde{g}}{\partial r}(r,\theta) = \cos(\theta)\left(\frac{-2\cos\theta}{r^3}\right) + \sin(\theta)\left(\frac{-2\sin\theta}{r^3}\right) = -\frac{2}{r^3}. \tag{18.795}$$

Nous pouvons vérifier directement que cela est correct. En effet

$$\tilde{g}(r,\theta) = g(r\cos\theta, r\sin\theta) = \frac{1}{r^2}, \tag{18.796}$$

dont la dérivée par rapport à r vaut $-2/r^3$.

En ce qui concerne la dérivée par rapport à θ, nous avons

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial\theta} &= \frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\cos(\theta)\big)}{\partial\theta} + \frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\sin(\theta)\big)}{\partial\theta}\\ &= \left(\frac{-2\cos\theta}{r^3}\right)(-r\sin\theta) + \left(\frac{-2\sin\theta}{r^3}\right)(r\cos\theta)\\ &= 0.\end{aligned} \tag{18.797}$$

En résumé et avec quelques abus de notation :

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial r} &= \cos(\theta)\frac{\partial g}{\partial x} + \sin(\theta)\frac{\partial g}{\partial y}\\ \frac{\partial\tilde{g}}{\partial\theta} &= -r\sin(\theta)\frac{\partial g}{\partial x} + r\cos(\theta)\frac{\partial g}{\partial y}\end{aligned} \tag{18.798}$$

18.13.2 Coordonnées cylindriques

Les **coordonnées cylindriques** sont un perfectionnement des coordonnées polaires. Il s'agit simplement de donner le point (x,y,z) en faisant la conversion $(x,y)\mapsto(r,\theta)$ et en gardant le z. Les formules de passage sont

$$\begin{cases} x = r\cos(\theta) & \text{(18.799a)}\\ y = r\sin(\theta) & \text{(18.799b)}\\ z = z. & \text{(18.799c)}\end{cases}$$

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\}\\ & (r,\theta,z) & \mapsto & (r\cos\theta, r\sin\theta, z),\end{array} \tag{18.800}$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme $]0,+\infty[\times[a-\pi,a+\pi[\times\mathbb{R}$, a dans $\mathbb{R}$. Il n'y a presque rien de nouveau par rapport aux coordonnées polaires. Les coordonnées cylindriques sont intéressantes si on décrit un objet invariant par rapport aux rotations autour de l'axe des z.

Exemple 18.234.
Il faut savoir ce que décrivent les équations les plus simples en coordonnées cylindriques,

- $r\leqslant a$, pour a constant dans $]0,+\infty[$, est le cylindre de hauteur infinie qui a pour axe l'axe des z et pour base le disque de rayon a centré à l'origine,
- $r=a$ est la surface du cylindre,
- $\theta=b$ est un demi-plan ouvert et sa fermeture contient l'axe des z,
- $z=c$ est un plan parallèle au plan x-y.

△

Exemple 18.235.
Un demi-cône qui a son sommet en l'origine et pour axe l'axe des z est décrit par $z=dr$. Si d est positif il s'agit de la moitié supérieure du cône, si $d<0$ de la moitié inférieure. △

Exemple 18.236.
De même, la sphère de rayon a et centrée à l'origine est l'assemblage des calottes $z=\sqrt{a^2-r^2}$ et $z=-\sqrt{a^2-r^2}$. △

En ce qui concerne les coordonnées cylindriques, le Jacobien est donné par

$$J(r,\theta,z)=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta} & \frac{\partial x}{\partial z}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta} & \frac{\partial y}{\partial z}\\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial\theta} & \frac{\partial z}{\partial z}\end{vmatrix}=\begin{vmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{vmatrix}=r. \tag{18.801}$$

Nous avons donc $dx\,dy\,dz=r\,dr\,d\theta\,dz$.

$$\begin{cases} x=r\cos\theta & \text{(18.802a)}\\ y=r\sin\theta & \text{(18.802b)}\\ z=z & \text{(18.802c)}\end{cases}$$

avec $r\in\,]0,\infty[$, $\theta\in[0,2\pi[$ et $z\in\mathbb{R}$. Le jacobien vaut r.

18.13.3 Coordonnées sphériques

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\}\\ & (\rho,\theta,\phi) & \mapsto & (\rho\cos\theta\sin\phi,\rho\sin\theta\sin\phi,\rho\cos\phi),\end{array} \tag{18.803}$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme $]0,+\infty[\times[a-\pi,a+\pi[\times[b-\pi/2,b+\pi/2[$, a et b dans $\mathbb{R}$. Si $a=0$ et $b=-\pi/2$ la fonction inverse T^{-1} est donnée par

$$\begin{array}{rccc} T^{-1}: & \mathbb{R}^3\backslash\{(0,0,0)\} & \to &]0,+\infty[\times[-\pi,\pi[\times[0,\pi[\\ & (x,y,z) & \mapsto & \left(\sqrt{x^2+y^2+z^2},\arctan\frac{y}{x},\arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)\right).\end{array} \tag{18.804}$$

Soit P un point dans $\mathbb{R}^3$. L'angle ϕ est l'angle entre le demi-axe positif des z et le vecteur $\overrightarrow{OP}$, ρ est la norme de $\overrightarrow{OP}$ et θ est l'argument en coordonnées polaires de la projection de $\overrightarrow{OP}$ sur le plan x-y.

Remarque 18.237.
Dans la littérature, les angles θ et ϕ sont parfois inversés (voire, changent de nom, par exemple φ au lieu de ϕ). Il faut donc être très prudent lorsqu'on veut utiliser dans un cours des formules données dans un autre cours.

Exemple 18.238.
Il faut connaitre le sens des équations plus simples,

- $\rho\leqslant a$, pour a constant dans $]0,+\infty[$, est la boule fermée de rayon a centrée à l'origine,
- $\rho=a$ est la sphère de rayon a centrée à l'origine,
- $\theta=b$ est un demi-plan ouvert et sa fermeture contient l'axe des z,
- $\phi=c$ est un demi-cône qui a son sommet à l'origine et pour axe l'axe des z. Si c est positif il s'agit de la moitié supérieure du cône, si $d<0$ de la moitié inférieure.

△

Les **coordonnées sphériques** sont ce qu'on appelle les « méridiens » et « longitudes » en géographie. Les formules de transformation sont [SubEqsCoordSphe]

$$\begin{cases} x=\rho\sin(\theta)\cos(\varphi) & \text{(18.805a)}\\ y=\rho\sin(\theta)\sin(\varphi) & \text{(18.805b)}\\ z=\rho\cos(\theta) & \text{(18.805c)}\end{cases}$$

avec $0\leqslant\theta\leqslant\pi$ et $0\leqslant\varphi<2\pi$.

Remarque 18.239.
Attention : d'un livre à l'autre les conventions sur les noms des angles changent. N'essayez donc pas d'étudier par cœur des formules concernant les coordonnées sphériques trouvées autre part. Par exemple sur le premier dessin de wikipédia, l'angle φ est noté θ et l'angle θ est noté Φ. Mais vous noterez que sur cette même page, les conventions de noms de ces angles changent plusieurs fois.

Les coordonnées sphériques sont données par

$$\begin{cases} x = r\cos\theta\sin\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\varphi \end{cases} \tag{18.806}$$ EqChmVarSpherique

avec $r \in]0,\infty[$, $\theta \in]0,2\pi[$ et $\varphi \in]0,\pi[$. Le jacobien associé est $Jg(r,\theta,\varphi) = -r^2\sin\varphi$. N'oubliez pas que lorsqu'on effectue un changement de variables dans une intégrale, la *valeur absolue* du jacobien apparaît.

Cependant notre convention de coordonnées sphériques fait venir $\sin(\phi)$ avec $\phi \in [0,\pi[$; vu que le signe de $\sin(\phi)$ y est toujours positif, cette histoire de valeur absolue est sans grandes conséquences. Ce n'est pas le cas de toutes les conventions possibles.

18.13.3.1 Coordonnées sphériques : inverse

Trouvons le changement inverse, c'est-à-dire trouvons ρ, θ et φ en termes de x, y et z. D'abord nous avons

$$\rho = \sqrt{x^2+y^2+z^2}. \tag{18.807}$$

Ensuite nous savons que

$$\cos(\theta) = \frac{z}{\rho} \tag{18.808}$$

détermine de façon unique[87] un angle $\theta \in [0,\pi]$. Dès que ρ et θ sont connus, nous pouvons poser $r = \rho\sin\theta$ et alors nous nous trouvons avec les équations

$$\begin{cases} x = r\cos(\varphi) & \text{(18.809a)} \\ y = r\sin(\varphi), & \text{(18.809b)} \end{cases}$$

qui sont similaires à celles déjà étudiées dans le cas des coordonnées polaires.

18.14 Calcul de limites

Beaucoup de techniques de calcul de limites fonctionnent bien avec les fonctions trigonométriques, entre autres grâce à l'utilisation des coordonnées polaires de la proposition 18.230. Nous en voyons quelques exemples à présent.

ExQWHooGddTLE

Exemple 18.240 (Limite et prolongement par continuité)**.**
La fonction

$$f(x) = \frac{\cos(x)-1}{x} \tag{18.810}$$

n'est pas définie en $x = 0$.

Nous avons vu dans l'équation (18.6) que $\cos(0) = 1$, donc la limite

$$\lim_{x\to 0}\frac{\cos(x)-1}{x} \tag{18.811}$$

est la limite définissant la dérivée de cosinus en 0 (ici, le x joue le rôle de ϵ). Le lemme 18.3 nous donne la dérivée du cosinus comme étant le sinus. Nous avons donc :

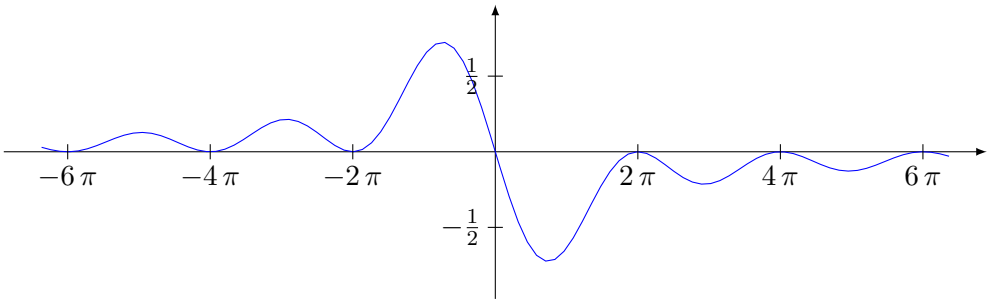
$$\lim_{x\to 0}\frac{\cos(x)-1}{x} = \sin(0) = 0, \tag{18.812}$$

87. Le problème $\rho = 0$ ne se pose pas ; pourquoi ?

et nous définissons le prolongement par continuité :

$$\tilde{f}(x) = \begin{cases} \frac{\cos(x)-1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases} \tag{18.813}$$

Encore une fois, le graphe de la fonction $\tilde{f}$ ne présente aucune particularité autour de $x = 0$.



△

EXooINLRooPzRWEA

Exemple 18.241 (Un calcul heuristique de limite)**.**
Soit à calculer la limite suivante :

$$\lim_{x\to 0} \frac{e^{-2\cos(x)+2}\sin(x)}{\sqrt{e^{2\cos(x)+2}-1}}. \tag{18.814}$$

La stratégie que nous allons suivre pour calculer cette limite est de développer certaines parties de l'expression en série de Taylor, afin de simplifier l'expression. La première chose à faire est de remplacer $e^{y(x)}$ par $1 + y(x)$ lorsque $y(x) \to 0$. La limite devient

$$\lim_{x\to 0} \frac{\big(-2\cos(x)+3\big)\sin(x)}{\sqrt{-2\cos(x)+2}}. \tag{18.815}$$

Nous allons maintenant remplacer $\cos(x)$ par 1 au numérateur et par $1 - x^2/2$ au dénominateur. Pourquoi ? Parce que le cosinus du dénominateur est dans une racine, donc nous nous attendons à ce que le terme de degré deux du cosinus donne un degré un en dehors de la racine, alors que du degré un est exactement ce que nous avons au numérateur : le développement du sinus commence par x.

Nous calculons donc

$$\lim_{x\to 0} \frac{\sin(x)}{\sqrt{-2\left(1-\frac{x^2}{2}\right)+2}} = \lim_{x\to 0} \frac{\sin(x)}{x} = 1. \tag{18.816}$$

Tout ceci n'est évidemment pas très rigoureux, mais en principe vous avez tous les éléments en main pour justifier les étapes. △

18.14.1 Méthode des coordonnées polaires

SUBSECooWCGMooPrXSpt

La proposition suivante exprime la définition de la limite en d'autres termes, et va être pratique dans le calcul de certaines limites.

PropMethodePolaire

Proposition 18.242.
Soit $f\colon D \subset \mathbb{R}^m \to \mathbb{R}^n$, a un point d'accumulation de D et $\ell \in \mathbb{R}^n$. Nous définissons

$$E_r = \{f(x) \text{ tel que } x \in B(a,r) \cap D\}, \tag{18.817}$$

et

$$s_r = \sup\{\|v-\ell\| \text{ tel que } v \in E_r\}. \tag{18.818}$$

Alors nous avons $\lim_{x\to a} f(x) = \ell$ si et seulement si $\lim_{r\to 0} s_r = 0$.

Dans cette proposition, E_r représente l'ensemble des valeurs atteintes par f dans un rayon r autour de a. Le nombre s_r sélectionne, parmi toutes ces valeurs, celle qui est la plus éloignée de ℓ et donne la distance. En d'autres termes, s_r est la distance maximale entre $f(x)$ et ℓ lorsque x est à une distance au maximum r de a.

Lorsque nous avons affaire à une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$, cette proposition nous permet de calculer facilement les limites en passant aux coordonnées polaires.

ExempleMethodeTrigigi

Exemple 18.243.
Reprenons la fonction de l'exemple 12.218 :

$$f(x,y) = \frac{xy}{x^2+y^2}. \tag{18.819}$$

Son domaine est $\mathbb{R}^2\backslash\{(0,0)\}$. Nous voulons calculer $\lim_{(x,y)\to(0,0)} f(x,y)$. Écrivons la définition de E_r :

$$E_r = \{f(x,y) \text{ tel que } (x,y) \in B\big((0,0),r\big)\}. \tag{18.820}$$

Les points de la boule sont, en coordonnées polaires, les points de la forme (ρ,θ) avec $\rho < r$. La chose intéressante est que $f(\rho,\theta)$ est relativement simple (plus simple que la fonction départ). En effet en remplaçant tous les x par $\rho\cos(\theta)$ et tous les y par $\rho\sin(\theta)$, et en utilisant le fait que $\cos^2(\theta)+\sin^2(\theta)=1$, nous trouvons

$$f(\rho,\theta) = \frac{\rho^2\cos(\theta)\sin(\theta)}{\rho^2} = \cos(\theta)\sin(\theta). \tag{18.821}$$

Eq2807fpolairerhodeuxcossin

Cela signifie que

$$E_r = \{\cos(\theta)\sin(\theta) \text{ tel que } \theta \in [0,2\pi[\}. \tag{18.822}$$

Prenons ℓ quelconque. Le nombre s_r est le supremum des

$$\|\ell - \cos(\theta)\sin(\theta)\| \tag{18.823}$$

lorsque θ parcours $[0,2\pi]$. Nous ne sommes pas obligés calculer la valeur exacte de s_r. Ce qui compte ici est que s_r ne vaut certainement pas zéro, et ne dépend pas de r. Donc il est impossible d'avoir $\lim_{r\to 0} s_r = 0$, et la fonction donnée n'a pas de limite en $(0,0)$. △

Nous pouvons retenir cette règle pour calculer les limites lorsque $(x,y) \to (0,0)$ de fonctions $f\colon \mathbb{R}^2 \to \mathbb{R}$:

(1) passer en coordonnées polaires, c'est-à-dire remplacer x par $\rho\cos(\theta)$ et y par $\rho\sin(\theta)$;

(2) nous obtenons une fonction g de ρ et θ. Si la limite $\lim_{r\to 0} g(r,\theta)$ n'existe pas ou dépend de θ, alors la fonction n'a pas de limite. Si on peut majorer g par une fonction ne dépendant pas de θ, et que cette fonction a une limite lorsque $r \to 0$, alors cette limite est la limite de la fonction.

La vraie difficulté de la technique des coordonnées polaires est de trouver le supremum de E_r, ou tout au moins de montrer qu'il est borné par une fonction qui a une limite qui ne dépend pas de θ. Une des situations classiques dans laquelle c'est facile est lorsque la fonction se présente comme une fonction de r multiplié par une fonction de θ.

Exemplexyxsqysq

Exemple 18.244.
Soit à calculer la limite

$$\lim_{(x,y)\to(0,0)} xy\left(\frac{x^2-y^2}{x^2+y^2}\right). \tag{18.824}$$

Le passage aux coordonnées polaires donne

$$f(r,\theta) = r^2\sin\theta\cos\theta(\cos^2\theta - \sin^2\theta). \tag{18.825}$$

Déterminer le supremum de cela est relativement difficile. Mais nous savons que de toutes façons, la quantité $\sin\theta\cos\theta(\cos^2\theta-\sin^2\theta)$ est bornée par 1. Donc

$$\|f(r,\theta)\|\leqslant r^2. \tag{18.826}$$

Maintenant la règle de l'étau montre que $\lim_{(x,y)\to(0,0)} f(x,y)$ est zéro.

La situation vraiment gênante serait celle avec une fonction de θ qui risque de s'annuler dans un dénominateur. △

L'exemple 18.229 donnera un cas où la méthode fonctionne plus difficilement. Entre autres parce qu'il utilisera en même temps la méthode des chemins et celle des coordonnées polaires.

ExmeASDLAf

Exemple 18.245.
Considérons fonction

$$f(x,y)=\frac{x^2+y^2}{x-y}. \tag{18.827}$$

Une mauvaise idée pour prouver que la limite n'existe pas pour $(x,y)\to(0,0)$ est de considérer le chemin (t,t). En effet, la fonction n'existe pas sur ce chemin. Or la méthode des chemins parle uniquement de chemins contenus dans le domaine de la fonction.

Nous prouvons que la limite n'existe pas en trouvant des chemins le long desquels les limites sont différentes. Si nous essayons le chemin (t,kt) avec k constant, nous trouvons

$$f(t,kt)=\frac{t(1+k^2)}{1-k}. \tag{18.828}$$

La limite $t\to 0$ est hélas toujours 0. Nous ne pouvons donc pas conclure.

Nous allons maintenant utiliser la même technique que celle utilisée en coordonnées polaires. Vous noterez que dans ce cas, travailler en cartésiennes donne lieu à des calculs plus longs. L'astuce consiste à prendre k non constant et à chercher par exemple $k(t)$ de façon à avoir

$$\frac{1+k(t)^2}{1-k(t)}=\frac{1}{t}. \tag{18.829}$$

Avec une telle fonction k, la fonction $t\mapsto f(t,tk(t))$ serait la constante 1. L'équation à résoudre pour k est

$$tk^2+k+(t-1)=0, \tag{18.830}$$

et les solutions sont

$$k(t)=\frac{-1\pm\sqrt{1-4t(t-1)}}{2t}. \tag{18.831}$$

Nous proposons donc les chemins

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}t\\ \frac{-1\pm\sqrt{1-4t(t-1)}}{2}\end{pmatrix} \tag{18.832}$$

Nous devons vérifier deux points. D'abord que ce chemin est bien défini, et ensuite que $tk(t)$ tend bien vers zéro lorsque $t\to 0$ (sinon $(t,k(t)t)$) n'est pas un chemin passant par $(0,0)$. Lorsque t est petit, ce qui se trouve sous la racine est proche de 1 et ne pose pas de problèmes. Ensuite,

$$\lim_{t\to 0} tk(t)=\frac{-1\pm 1}{2}. \tag{18.833}$$

En choisissant le signe +, nous trouvons un chemin qui nous convient.

Ce que nous avons prouvé est que

$$f\left(t,\frac{-1+\sqrt{1-4t(t-1)}}{2}\right)=1 \tag{18.834}$$

pour tout t. Le long de ce chemin, la limite de f est donc 1. Cette limite est différente des limites obtenues le long de chemins avec k constant. La limite $\lim_{(x,y)\to(0,0)} f(x,y)$ n'existe donc pas. △

seno

Exemple 18.246.
Considérons la fonction (figure 18.11)

$$f(x,y)=\begin{cases}\sqrt{x^2+y^2}\sin\frac{1}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0),\end{cases}\tag{18.835}$$

et cherchons la limite $(x,y)\to(0,0)$. Le passage en coordonnées polaires [88] donne

EqFoncRho2907
$$f(\rho,\theta)=\rho\sin\frac{1}{\rho}.\tag{18.836}$$

Pour calculer la limite de cela lorsque $\rho\to 0$, nous remarquons que

$$0\leqslant|\rho\sin\frac{1}{\rho}|\leqslant\rho\tag{18.837}$$

parce que $\sin(\frac{1}{\rho})\leqslant 1$ quel que soit ρ. Or évidemment $\lim_{\rho\to0}\rho=0$, donc la limite de la fonction (18.836) est zéro et ne dépend pas de θ. Nous en concluons que $\lim_{(x,y)\to(0,0)}f(x,y)=0$. △

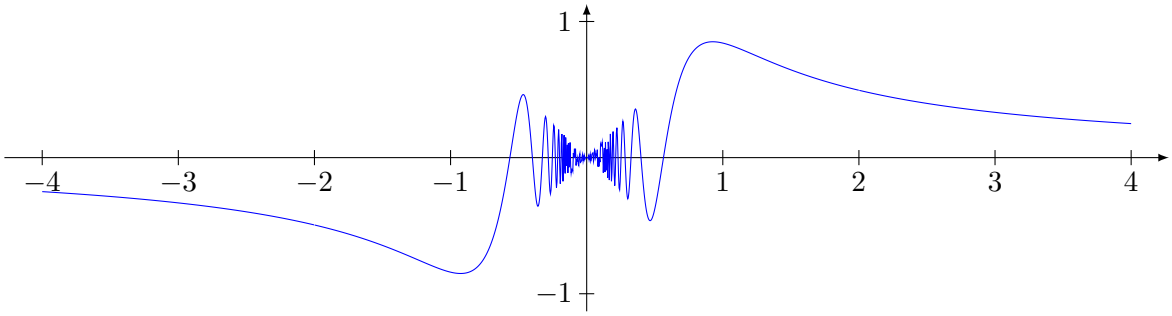


FIGURE 18.11: La fonction de l'exemple 18.246.

LabelFigsenotopologo

18.14.2 Méthode du développement asymptotique

SUBSECooRAKKooAnpvkE

Nous savons que nous pouvons développer certaines fonctions en série grâce au développement de Taylor (théorème 12.461). Lorsque nous avons une limite à calculer, nous pouvons remplacer certaines parties de la fonction à traiter par la formule (12.1298a). Cela est très utile pour comparer des fonctions trigonométrique à des polynômes.

LEMooZYNEooYkwsWD

Lemme 18.247.
Nous avons la limite

$$\lim_{x\to0}\frac{\sin(x)}{x}=1.\tag{18.838}$$

Démonstration. Une manière de prouver cela est d'écrire

$$\sin(x)=x+h(x)\tag{18.839}$$

avec $h\in o(x)$, c'est-à-dire $\lim_{x\to0}h(x)/x=0$. Alors nous avons

$$\lim_{x\to0}\frac{\sin(x)}{x}=\lim_{x\to0}\frac{x+h(x)}{x}=\lim_{x\to0}\frac{x}{x}+\lim_{x\to0}\frac{h(x)}{x}=1.\tag{18.840}$$

□

L'utilisation de la proposition 12.497 permet d'utiliser cette technique dans le cadre de limites à plusieurs variables. Reprenons le lemme 18.247 un tout petit peu modifié :

88. Proposition 18.230.

LEMooSFALooVRBdNb

Lemme 18.248.
Pour tout $x > 0$ nous avons $\sin(x) < x$.

Démonstration. Nous posons $f(x) = x - \sin(x)$. Cette fonction vérifie $f(0) = 0$ et

$$f'(x) = 1 - \cos(x). \tag{18.841}$$

Vu que $|\cos(x)| \leqslant 1$, nous avons toujours $f'(x) \geqslant 0$ et même $f'(x) > 0$ pour $x \in]0, \delta]$. Donc f est au moins strictement croissante sur $]0, \delta]$ et ensuite strictement croissante presque partout. □

Exemple 18.249.
Soit à calculer $\lim_{(x,y)\to(0,0)} f(x, y)$ où

$$f(x, y) = \frac{\sin(xy)}{xy}. \tag{18.842}$$

La première chose à faire est de voir f comme la composée de fonctions $f = f_1 \circ f_2$ avec

$$\begin{aligned} f_1 \colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \frac{\sin(t)}{t} \end{aligned} \tag{18.843}$$

et

$$\begin{aligned} f_2 \colon \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto xy. \end{aligned} \tag{18.844}$$

Étant donné que $\lim_{(x,y)\to(0,0)} f_2(x, y) = 0$, nous avons $\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{t\to 0} f_1(t) = 1$. △

EXooETZYooYsKPDJ

Exemple 18.250.
Les dérivées partielles de la fonction $f(x, y) = xy^3 + \sin y$ au point $(0, \pi)$ sont

$$\partial_x f(0, \pi) = \frac{\partial f}{\partial x}(0, \pi) = \lim_{\substack{t\to 0 \\ t\neq 0}} \frac{(t\pi^3 + \sin\pi) - (\sin\pi)}{t} = \pi^3,$$

$$\partial_y f(0, \pi) = \frac{\partial f}{\partial y}(0, \pi) = \lim_{\substack{t\to 0 \\ t\neq 0}} \frac{0(\pi + t)^3 + \sin(t + \pi) - 0 \cdot \pi^3}{t} = \cos\pi = -1,$$

△

18.15 Quelques intégrales avec de la trigonométrie

SECooOOPPooZLbaEH

Le théorème 14.281 manque un peu d'exemples. Nous allons en voir quelques-uns maintenant.

18.15.1 Changement de variables

Le domaine $E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 < 1\}$ s'écrit plus facilement $E = \{(r, \theta) \text{ tel que } r < 1\}$ en coordonnées polaires. Le passage aux coordonnées polaires permet de transformer une intégration sur un domaine rond à une intégration sur le domaine rectangulaire $]0, 2\pi[\times]0, 1[$. La question est évidemment de savoir si nous pouvons écrire

$$\int_E f = \int_0^{2\pi} \int_0^1 f(r\cos\theta, r\sin\theta) dr d\theta. \tag{18.845}$$

Hélas, non. La vie n'est pas aussi simple ; il faut tenir compte du fait que le changement de base dilate ou contracte certaines surfaces.

Soit $\varphi\colon D_1 \subset \mathbb{R}^2 \to D_2 \subset \mathbb{R}^2$ une fonction bijective de classe C^1 dont l'inverse est également de classe C^1. On désigne par x et y ses composantes, c'est-à-dire que

$$\varphi(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix} \tag{18.846}$$

avec $(u,v) \in D_1$.

ThoChamDeVarIntDDf

Théorème 18.251.
Soit une fonction continue $f\colon D_2 \to \mathbb{R}$. Alors

$$\int_{\varphi(D_1)} f(x,y)dxdy = \int_{D_1} f\big(x(u,v),y(u,v)\big)|J_\varphi(u,v)|dudv \tag{18.847}$$

où J_φ est le Jacobien de φ c'est-à-dire

$$J_\varphi(u,v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial u}{\partial v} \end{pmatrix}. \tag{18.848}$$

Ne pas oublier de prendre la valeur absolue lorsqu'on utilise le Jacobien dans un changement de variables.

18.15.2 Coordonnées polaires

Exemple 18.252.
Calculons la surface du disque D de rayon R. Nous devons calculer

$$\int_D dxdy. \tag{18.849}$$

Pour passer au polaires, nous savons que le disque est décrit par

$$D = \{(r,\theta) \text{ tel que } 0 \leqslant r \leqslant R, 0 \leqslant \theta \leqslant 2\pi\}. \tag{18.850}$$

Nous avons donc

$$\int_D dxdy = \int_D r\,drd\theta = \int_0^{2\pi}\int_0^R r\,drd\theta = 2\pi\int_0^R r\,dr = \pi R^2. \tag{18.851}$$

△

ExpmfDtAtV

Exemple 18.253.
Montrons comment intégrer la fonction $f(x,y) = \sqrt{1-x^2-y^2}$ sur le domaine délimité par la droite $y = x$ et le cercle $x^2 + y^2 = y$, représenté sur la figure 18.12. Pour trouver le centre et le rayon du cercle $x^2+y^2 = y$, nous commençons par écrire $x^2+y^2-y = 0$, et ensuite nous reformons le carré : $y^2 - y = (y - \frac{1}{2})^2 - \frac{1}{4}$.

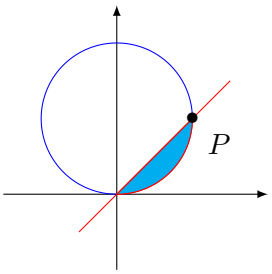


FIGURE 18.12: Passage en polaire pour intégrer sur un morceau de cercle.

LabelFigHFAYooOrfMAA

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned}\cos(\theta) &= \sin(\theta)\\ r^2 &= r\sin(\theta).\end{aligned} \tag{18.852}$$

L'angle θ parcours donc $]0,\pi/4[$, et le rayon, pour chacun de ces θ parcours $]0,\sin(\theta)[$. La fonction à intégrer se note maintenant $f(r,\theta)=\sqrt{1-r^2}$. Donc l'intégrale à calculer est

$$\int_0^{\pi/4}\left(\int_0^{\sin(\theta)}\sqrt{1-r^2}r\,dr\right). \tag{18.853}$$

PgOMRapIntMultFubiniBoutCercle

Remarquez la présence d'un r supplémentaire pour le jacobien.

Notez que les coordonnées du point P sont $(1,1)$. △

En pratique, lors du passage en coordonnées polaires, le « $dxdy$ » devient « $r\,drd\theta$ ».

Exemple 18.254.
On veut évaluer l'intégrale de la fonction $f(x,y)=x^2+y^2$ sur la région V suivante :

$$V=\{(x,y)\in\mathbb{R}^2\,|\,x^2+y^2\leqslant 1,\ x>0,\ y>0\}.$$

On peut faire le calcul directement,

$$\int_V f(x,y)\,dV=\int_0^1\int_0^{\sqrt{1-x^2}}x^2+y^2\,dy\,dx=\int_0^1\left(x^2\sqrt{1-x^2}+\frac{(1-x^2)^{3/2}}{3}\right)dx$$

mais c'est un peu ennuyeux. On peut simplifier beaucoup les calculs avec un changement de variables vers les coordonnées polaires. Dans ce cas, on sait bien que le difféomorphisme à utiliser est $\phi(r,\theta)=(r\cos\theta,r\sin\theta)$. Le jacobien J_ϕ est

$$J_\phi(r,\theta)=\begin{vmatrix}\cos\theta & \sin\theta\\ -r\sin\theta & r\cos\theta\end{vmatrix}=r, \tag{18.854}$$

qui est toujours positif. D'une part, la fonction f peut s'écrire sous la forme $f(\phi(r,\theta))=r^2$ et d'autre part, $\phi^{-1}(V)=]0,1]\times]0,\pi/2[$. Par conséquent, la formule du changement de variables nous donne

$$\int_V f(x,y)\,dV=\int_0^{\pi/2}\int_0^1 r^3dr\,d\theta=\int_0^{\pi/2}\frac{1}{4}\,d\theta=\frac{\pi}{8}.$$

△

18.15.3 Coordonnées cylindriques

Exemple 18.255.
On veut calculer le volume de la région A définie par l'intersection entre la boule unité et le cylindre qui a pour base un disque de rayon $1/2$ centré en $(0,1/2)$

$$A=\{(x,y,z)\in\mathbb{R}^3\,|\,x^2+y^2+z^1\leqslant 1\}\cap\{(x,y,z)\in\mathbb{R}^3\,|\,x^2+(y-1/2)^2\leqslant 1/4\}.$$

On peut décrire A en coordonnées cylindriques

$$\begin{aligned}A=\Big\{(r,\theta,z)\in{}&]0,+\infty[\times[-\pi,\pi[\times\mathbb{R}\,|\\ &-\pi/2<\theta<\pi,\ 0<r\leqslant\sin\theta,\ -\sqrt{1-r^2}\leqslant z\leqslant\sqrt{1-r^2}\Big\}.\end{aligned} \tag{18.855}$$

Le jacobien de ce changement de variables, J_{cyl}, est

$$J_{cyl}(r,\theta,z)=\begin{vmatrix}\cos\theta & \sin\theta & 0\\ -r\sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{vmatrix}=r, \tag{18.856}$$

qui est toujours positif. Le volume de A est donc

$$\int_{\mathbb{R}^3} \chi_A(x,y,z)\,dV = \int_{-\pi/2}^{\pi/2}\int_0^{\sin\theta}\int_{-\sqrt{1-r^2}}^{\sqrt{1-r^2}} r\,dz\,dr\,d\theta = \frac{2\pi}{8} + \frac{8}{9}.$$

△

Exemple 18.256 (Volume d'un solide de révolution)**.**
Soit $g : [a,b] \to \mathbb{R}_+$ une fonction continue et positive. On dit que le solide A décrit par

$$A = \left\{(x,y,z) \in \mathbb{R}^3 \,|\, z \in [a,b],\ \sqrt{x^2+y^2} \leqslant g^2(z)\right\}$$

est un solide de révolution. Afin de calculer son volume, on peut décrire A en coordonnées cylindriques,

$$A = \left\{(r,\theta,z) \in]0,+\infty[\times[-\pi,\pi[\times\mathbb{R} \,|\, a \leqslant z \leqslant b,\ 0 < r^2 \leqslant g^2(z)\right\}.$$

Le jacobien de ce changement de variables est $J_{cyl} = r$, comme dans l'exemple précédent. Le volume de A est donc

$$\int_{\mathbb{R}^3} \chi_A(x,y,z)\,dV = \int_a^b\int_{-\pi}^{\pi}\int_0^{g(z)} r\,dr\,d\theta\,dz = \int_a^b \pi g^2(z)\,dz.$$

Cette formule peut être utilisée pour tout solide de révolution. △

18.15.3.1 Coordonnées sphériques

Le calcul est un peu plus long :

$$\begin{aligned} J(\rho,\theta,\varphi) &= \begin{vmatrix} \frac{\partial x}{\partial\rho} & \frac{\partial x}{\partial\theta} & \frac{\partial x}{\partial\varphi} \\ \frac{\partial y}{\partial\rho} & \frac{\partial y}{\partial\theta} & \frac{\partial y}{\partial\varphi} \\ \frac{\partial z}{\partial\rho} & \frac{\partial z}{\partial\theta} & \frac{\partial z}{\partial\varphi} \end{vmatrix} \\ &= \begin{vmatrix} \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & -\rho\sin\theta\cos\varphi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} \\ &= \rho^2\sin\theta. \end{aligned} \tag{18.857}$$

Donc

$$dx\,dy\,dz = \rho^2\sin(\theta)\,d\rho\,d\theta\,d\varphi. \tag{18.858}$$

Exemple 18.257.
On veut calculer le volume du cornet de glace A

$$A = \left\{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in \mathbb{S}^2,\ \sqrt{x^2+y^2} \leqslant z \leqslant \sqrt{1-x^2-y^2}\right\}.$$

On peut décrire A en coordonnées sphériques.

$$A = \{(\rho,\theta,\phi) \in]0,+\infty[\times[-\pi,\pi[\times[0,\pi[\,|\,0 < \phi \leqslant \pi/4,\ 0 < \rho \leqslant 1\}.$$

Le jacobien de ce changement de variables J_{sph} est

$$J_{sph}(\rho,\theta,\phi) = \begin{vmatrix} \cos\theta\sin\phi & \sin\theta\sin\phi & \cos\phi \\ -\rho\sin\theta\sin\phi & \rho\cos\theta\sin\phi & 0 \\ \rho\cos\theta\cos\phi & \rho\sin\theta\cos\phi & -\rho\sin\phi \end{vmatrix} = \rho^2\sin\phi, \tag{18.859}$$

Le volume de A est donc

$$\int_{\mathbb{R}^3} \chi_A(x,y,z)\,dV = \int_{-\pi}^{\pi}\int_0^{\pi/4}\int_0^1 \rho^2\sin\phi\,d\rho\,d\phi\,d\theta = \frac{2\pi}{3}\left(1 - \frac{1}{\sqrt{2}}\right).$$

△

18.15.4 Un autre système utile

Un changement de variables que l'on voit assez souvent est

$$\begin{cases} u = x + y & (18.860a) \\ v = x - y. & (18.860b) \end{cases}$$

Afin de calculer son jacobien, il faut d'abord exprimer x et y en fonctions de u et v :

$$\begin{cases} x = (u + v)/2 & (18.861a) \\ y = (u - v)/2. & (18.861b) \end{cases}$$

La matrice jacobienne est

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}. \tag{18.862}$$

Le déterminant vaut $-\frac{1}{2}$. Nous avons donc

$$dxdy = \frac{1}{2}dudv. \tag{18.863}$$

Nous insistons sur le fait que c'est $\frac{1}{2}$ et non $-\frac{1}{2}$ qui intervient parce que que la formule du changement de variable demande d'introduire la *valeur absolue* du jacobien.

Exemple 18.258.
Calculer l'intégrale de la fonction $f(x, y) = x^2 - y^2$ sur le domaine représenté sur la figure 18.13.

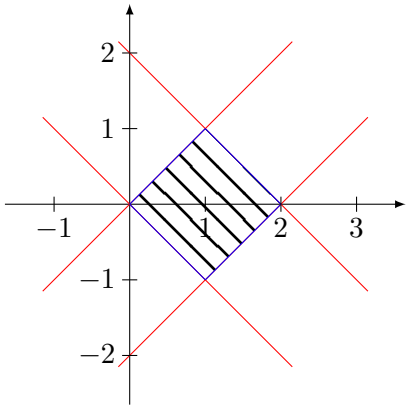


FIGURE 18.13: Un domaine qui s'écrit étonnament bien avec un bon changement de coordonnées.
LabelFigVWFLooFSrOqz

Les droites qui délimitent le domaine d'intégration sont

$$\begin{aligned} y &= -x + 2 \\ y &= x - 2 \\ y &= x \\ y &= -x \end{aligned} \tag{18.864}$$

Le domaine est donc donné par les équations

$$\begin{cases} y + x < 2 & (18.865a) \\ y - x > -2 & (18.865b) \\ y - x < 0 & (18.865c) \\ y + x > 0. & (18.865d) \end{cases}$$

En utilisant le changement de variables $u = x + y$, $v = x - y$ nous trouvons le domaine $0 < u < 2$, $0 < v < 2$. En ce qui concerne la fonction, $f(x, y) = (x + y)(x - y)$ et par conséquent

$$f(u, v) = uv. \tag{18.866}$$

L'intégrale à calculer est simplement

$$\int_0^2 \int_0^2 uv\, dudv = \int_0^2 u\, du \left[\frac{v^2}{2}\right]_0^2 = 2\int_0^2 u\, du = 4. \tag{18.867}$$

△

18.16 Aire d'une surface de révolution

¡¡ Avertissement/question à la lectrice !! 18.259
Vu que nous n'avons pas encore de définition d'une aire dans $\mathbb{R}^3$, ce qui suit n'est en réalité pas très rigoureux.

Soit γ une courbe dans le plan xy, paramétrée par

$$\gamma(u) = \begin{pmatrix} x(u) \\ y(u) \\ 0 \end{pmatrix} \tag{18.868}$$

avec $u \in [a, b]$. Nous supposons que la courbe est toujours positive, c'est-à-dire $y(u) > 0$ pour tout u.

Nous voulons considérer la surface obtenue en effectuant une rotation de cette ligne autour de l'axe X. Chaque point de la courbe va parcourir un cercle de rayon $y(u)$ dans le plan YX et centré en $(x(u), 0, 0)$. La surface est donc donnée par

$$\varphi(u, \theta) = \begin{pmatrix} x(u) \\ y(u)\cos\theta \\ y(u)\sin\theta \end{pmatrix} \tag{18.869}$$

avec $(u, \theta) \in [a, b] \times [0, 2\pi]$. Notez que la courbe de départ correspond à $\theta = 0$.

Les vecteurs tangents à la surface pour ce paramétrage sont

$$T_u = \frac{\partial\varphi}{\partial u} = \begin{pmatrix} x'(u) \\ y'(u)\cos\theta \\ y'(u)\sin\theta \end{pmatrix} \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} 0 \\ -y(u)\sin\theta \\ y(u)\cos\theta \end{pmatrix}. \tag{18.870}$$

Le produit vectoriel de ces deux vecteurs vaut

$$\begin{aligned} T_u \times T_\theta &= \begin{vmatrix} e_x & e_y & e_z \\ x' & y'\cos\theta & y'\sin\theta \\ 0 & -y\sin\theta & y\cos\theta \end{vmatrix} \\ &= y'(u)y(u)\, e_x - x'(u)y(u)\cos\theta\, e_y + x'(u)y(u)\sin\theta\, e_z. \end{aligned} \tag{18.871}$$

En ce qui concerne la norme :

$$dS = \|T_u \times T_\theta\| = \sqrt{(y'y)^2 + (x'y)^2} = |y(u)|\sqrt{y'(u)^2 + x'(u)^2}. \tag{18.872}$$

Étant donné que nous avons supposé que $y(u) > 0$, nous pouvons supprimer les valeurs absolues, et l'aire de la surface de révolution devient :

$$\begin{aligned} Aire(S) &= \int_0^{2\pi} d\theta \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du \\ &= 2\pi \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du. \end{aligned} \tag{18.873}$$

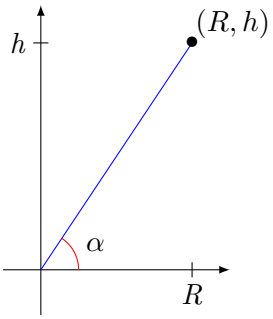


FIGURE 18.14: En faisant tourner cette droite autour de l'axe X, nous obtenons un cône. LabelFigTHJYooTEXLLn

EXooZCLXooVmXQgY

Exemple 18.260.
Calculons la surface du cône de révolution de rayon (à la base) R et de hauteur h. La courbe de départ est le segment droite qui part de $(0,0)$ et qui termine en (R,h) de la figure 18.14.

Ce segment peut être paramétré par

$$\gamma(u) = \begin{pmatrix} Ru \\ hu \\ 0 \end{pmatrix} \tag{18.874}$$

avec $u \in [0,1]$. Cela donne $x(u) = Ru$, $y(u) = hu$ et par conséquent

$$Aire = 2\pi \int_0^1 hu\sqrt{R^2+h^2} = \pi h\sqrt{R^2+h^2}. \tag{18.875}$$

Ce résultat peut aussi être exprimé en fonction de l'angle, grâce à la formule (18.583). En sachant que $h = \sqrt{h^2+R^2}\sin(\alpha)$, nous trouvons

$$Aire = \pi(R^2+h^2)\sin(\alpha). \tag{18.876}$$

△

Exemple 18.261.
Calculons la surface latérale du tore obtenu par révolution du cercle de la figure 18.15.

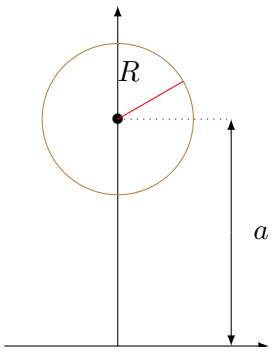


FIGURE 18.15: Si nous tournons ce cercle autour de l'axe X, nous obtenons un tore de rayon « externe » a et de rayon « interne » R. LabelFigROAOooPgUZIt

Le chemin qui détermine le cercle de départ est

$$\gamma(u) = \begin{pmatrix} R\cos(u) \\ a + R\sin(u) \\ 0 \end{pmatrix}, \tag{18.877}$$

c'est-à-dire $x(u) = R\cos(u)$, $y(u) = a + R\sin(u)$ avec $u \in [0, 2\pi]$. Nous avons donc l'aire

$$\begin{aligned} Aire &= 2\pi \int_0^{2\pi} \big(a + R\sin(u)\big) R\, du \\ &= 2\pi R\big(2\pi a + R[-\cos(u)]_0^{2\pi}\big) \\ &= 4\pi^2 aR. \end{aligned} \tag{18.878}$$

△

18.17 Table de caractères du groupe diédral

SecWMzheKf

Cette section vient de [104] ; nous avons comme but d'établir la table des caractères des représentations complexes du groupe diédral D_n.

18.17.1 Représentations de dimension un

Nous nous occupons des représentations de D_n sur $\mathbb{C}$. Les applications linéaires $\mathbb{C} \to \mathbb{C}$ sont seulement les multiplications par des nombres complexes. Nous cherchons donc $\psi\colon D_n \to \mathbb{C}^*$.

Nous savons que D_n est généré [89] par s et r. Vu que $s^2 = 1$, nous avons

$$\psi(s)^2 = \psi(s^2) = \psi(1) = 1, \tag{18.879}$$

donc $\psi(s) \in \{-1, 1\}$. Nous savons aussi que $srsr = 1$, donc

$$\psi(s)^2\psi(r)^2 = 1, \tag{18.880}$$

ce qui donne $\psi(r) \in \{-1, 1\}$.

Nous avons donc quatre représentations de dimension un données par

	$\psi(r) = 1$	$\psi(r) = -1$
$\psi(s) = 1$	ρ^{++}	ρ^{+-}
$\psi(s) = -1$	ρ^{-+}	ρ^{--}

Attention au fait que nous devons aussi avoir la relation $\psi(r)^n = \psi(r^n) = 1$. Donc $\psi(r)$ doit être une racine n^{e} de l'unité. Nous allons donc devoir avoir un compte différent selon la parité de n. Nous en reparlerons à la fin, au moment de faire les comptes. En ce qui concerne les caractères correspondants,

	r^k	sr^k
χ^{++}	1	1
χ^{+-}	$(-1)^k$	$(-1)^k$
χ^{-+}	1	-1
χ^{--}	$(-1)^k$	$(-1)^{k+1}$

Étant donné qu'ils sont tous différents, ce sont des représentations deux à deux non équivalentes, lemme 16.16.

18.17.2 Représentations de dimension deux

Nous cherchons maintenant les représentations $\rho\colon D_n \to \mathrm{End}(\mathbb{C}^2)$. Ici nous supposons connue la liste des éléments de D_n donnée par le corolaire 18.176. Soit $\omega = e^{2i\pi/n}$ et $h \in \mathbb{Z}$; nous considérons la représentation $\rho^{(h)}$ de D_n définie par

$$\rho^{(h)}(r^k) = \begin{pmatrix} \omega^{hk} & 0 \\ 0 & \omega^{-hk} \end{pmatrix} \tag{18.881a}$$

$$\rho^{(h)}(sr^k) = \begin{pmatrix} 0 & \omega^{-hk} \\ \omega^{hk} & 0 \end{pmatrix}. \tag{18.881b}$$

89. Voir proposition 18.175 et tout ce qui suit.

Cela donne bien $\rho^{(h)}$ sur tous les éléments de D_n par la proposition 18.175. Nous pouvons restreindre le domaine de h en remarquant d'abord que $\rho^{(h)}=\rho^{(h+n)}$, et ensuite que les représentations $\rho^{(h)}$ et $\rho^{(-h)}$ sont équivalentes. Un opérateur d'entrelacement est donné par $T=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, et il est facile de vérifier que $T\rho^{(h)}(x)=\rho^{-h}(x)T$ avec $x=r^k$ puis avec $x=sr^k$.

Donc $\rho^{(h)}\simeq\rho^{(-h)}\simeq\rho^{(n-h)}$ et nous pouvons restreindre notre étude à $0\leqslant h\leqslant\frac{n}{2}$.

Nous allons séparer les cas $n=0$, $h=n/2$ et les autres. En effet si nous notons par commodité $a=\omega^h$, alors un vecteur (x,y) est vecteur propre de $\rho^{(h)}(s)$ et de $\rho^{(h)}(r)$ si et seulement si il vérifie les systèmes d'équations SubEqsGXZoxLq

$$\begin{cases}ax=\lambda x & (18.882a)\\ \frac{1}{a}y=\lambda y & (18.882b)\end{cases}$$

et SubEqsFYZmzhT

$$\begin{cases}\frac{1}{a}y=\mu x & (18.883a)\\ ax=\mu y & (18.883b)\end{cases}$$

avec λ et μ des nombres non nuls. Une représentation sera réductible si et seulement si ces deux systèmes acceptent une solution non nulle commune. Il est vite vu que si $x\neq0$ et $y\neq0$, alors $a^2=1$, ce qui signifie $h=0$ ou $h=n/2$. Sinon, il n'y a pas de solutions, et la représentation associée est irréductible.

(1) $h=0$. Nous avons

$$\rho^{(0)}(r^k)=\begin{pmatrix}1&0\\0&1\end{pmatrix}\quad\rho^{(0)}(sr^k)=\begin{pmatrix}0&1\\1&0\end{pmatrix},\tag{18.884}$$

donc le caractère de cette représentation est $\chi^{(0)}(r^k)=2$ et $\chi^{(0)}(sr^k)=0$. Donc nous avons

$$\chi^{(0)}=\chi^{++}+\chi^{-+}.\tag{18.885}$$

Il y a maintenant (au moins) quatre façons de voir que la représentation $\rho^{(0)}$ est réductible.

Première méthode Trouver un opérateur d'entrelacement. Pour cela nous calculons les matrices :

$$S(r^k)=(\rho^{++}\oplus\rho^{-+})(r^k)=\begin{pmatrix}\rho^{++}(r^k)&0\\0&\rho^{-+}(r^k)\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\tag{18.886a}$$

$$S(sr^k)=(\rho^{++}\oplus\rho^{-+})(sr^k)=\begin{pmatrix}\rho^{++}(sr^k)&0\\0&\rho^{-+}(sr^k)\end{pmatrix}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\tag{18.886b}$$

$$\tag{18.886c}$$

Nous cherchons une matrice T telle que $TS(r^k)=\rho^{(0)}(r^k)T$ et $TS(sr^k)=\rho^{(0)}(sr^k)T$. Étant donné que $S(r^k)=\mathbb{1}=\rho^{(0)}(r^k)$, la première contrainte n'en est pas une. Nous pouvons vérifier qu'avec $T=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$, nous avons bien

$$T\begin{pmatrix}1&0\\0&-1\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}.\tag{18.887}$$

Donc ce T entrelace $\rho^{++}\oplus\rho^{-+}$ avec $\rho^{(0)}$ qui sont donc deux représentations équivalentes. Donc $\rho^{(0)}$ est réductible et ça ne nous intéresse pas de la lister.

Seconde méthode Invoquer le théorème 16.24(1) et dire que les représentations sont équivalentes parce que les caractères sont égaux.

Troisième méthode Utiliser le théorème 16.24(2) et calculer $\langle\chi^{(0)},\chi^{(0)}\rangle$:

$$\langle\chi^{(0)},\chi^{(0)}\rangle = \frac{1}{|D_n|}\sum_{g\in D_n}|\chi^{(0)}(g)|^2 \tag{18.888a}$$
$$= \frac{1}{2n}\big(4+0+4(n-1)\big) \tag{18.888b}$$
$$= 2. \tag{18.888c}$$

Ici le 4 est pour le 1, le zéro est pour les termes sr^k et $4(n-1)$ est pour les $n-1$ termes r^k. Vu que le résultat n'est pas 1, la représentation $\rho^{(0)}$ n'est pas irréductible.

Quatrième méthode Regarder les solutions des systèmes (18.882) et (18.883) dont nous avons parlé plus haut.

La première méthode a l'avantage d'être simple et ne demander aucune théorie particulière à part les définitions. La seconde méthode est la plus rapide, mais demande un théorème très puissant. La troisième utilise également un théorème assez avancé, mais a l'avantage sur les deux autres méthodes de ne pas avoir besoin de savoir à priori un candidat décomposition de $\rho^{(0)}$; cette méthode est applicable même sans faire la remarque que $\chi^{(0)}=\chi^{++}+\chi^{-+}$. Quoi qu'il en soit, nous ne listons pas $\chi^{(0)}$ dans notre table de caractères.

(2) $h=n/2$. Vu que $\omega^{n/2}=e^{i\pi}=-1$, nous avons

$$\rho^{(n/2)}(r^k)=\begin{pmatrix}(-1)^k & 0\\ 0 & (-1)^k\end{pmatrix}\quad \rho^{(n/2)}(sr^k)=\begin{pmatrix}0 & (-1)^k\\ (-1)^k & 0\end{pmatrix}\quad, \tag{18.889}$$

et donc

$$\chi^{(n/2)}(r^k)=2(-1)^k \tag{18.890a}$$
$$\chi^{(n/2)}(sr^k)=0. \tag{18.890b}$$

Il est vite vu que $\chi^{(n/2)}=\chi^{+-}+\chi^{-+}$. Ergo la représentation $\rho^{(n/2)}$ n'est pas irréductible.

(3) $0<h<\frac{n}{2}$. Dans ce cas nous avons $\omega^h\neq\omega^{-h}$, et en regardant les systèmes d'équations donnés plus haut, nous voyons que $\rho^{(h)}(s)$ et $\rho^{(h)}(r)$ n'ont pas de vecteurs propres communs. Donc ces représentations sont irréductibles.

Nous devons cependant encore vérifier si elles sont deux à deux non équivalentes. Supposons que pour $h\neq h'$ nous ayons une matrice $T\in\mathrm{GL}(2,\mathbb{C})$ telle que $T\rho^{(h)}(r)T^{-1}=\rho^{(h')}(r)$. Cela impliquerait en particulier que les matrices $\rho^{(h)}(r)$ et $\rho^{(h')}(r)$ aient même valeurs propres. Nous aurions donc $\{\omega^h,\omega^{-h}\}=\{\omega^{h'},\omega^{-h'}\}$. Mais cela est impossible avec $0<h<h'<\frac{n}{2}$. Donc toutes ces représentations sont distinctes.

Le caractère de la représentation $\rho^{(h)}$ est $\chi^{(h)}(r^k)=\omega^{hk}+\omega^{-hk}=2\cos\big(\frac{2\pi hk}{n}\big)$.
Nous ajoutons donc la ligne suivante à notre liste :

	r^k	sr^k
$\chi^{(h)}$	$2\cos\big(\frac{2\pi hk}{n}\big)$	0

18.17.3 Le compte pour n pair

Nous avons 4 représentations de dimension 1 puis $\frac{n}{2}-1$ représentations de dimension 2. En tout nous avons

$$\frac{n}{2}+3 \tag{18.891}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 18.10.14.5. Pour rappel, le nombre de représentations non équivalentes est égal au nombre de classes de conjugaison par le corolaire 16.27. Notons que c'est cela qui justifie le fait que nous ne devons pas chercher d'autres représentations. Nous sommes sûrs de les avoir toutes trouvées.

18.17.4 Le compte pour n impair

Nous avions fait mention plus haut du fait que si ψ est une représentation de dimension 1, le nombre $\psi(r)$ devait être une racine n^{e} de l'unité. Donc en dimension 1 nous avons seulement les représentations ρ^{++} et ρ^{-+}. Pour celles de dimension 2, nous en avons $\frac{n-1}{2}$. En tout nous avons donc

$$\frac{n+3}{2} \tag{18.892}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 18.10.14.6.

Chapitre 19

Corps finis, racines de l'unité

19.1 Le groupe des racines de l'unité dans les nombres complexes

SecGJOLooWdMYVl

19.1.1 Le groupe

DEFooDUWPooZaAByH

Définition 19.1.
Une ***racine*** n^e ***de l'unité*** *dans un anneau est une racine du polynôme* $X^n - 1$.

LEMooSXFBooYJmRTK

Lemme-Définition 19.2.
Soit

$$U_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\} \tag{19.1}$$

EqIEAXooIpvFPe

(1) L'ensemble U_n est un groupe pour la multiplication.
(2) L'ensemble U_n est l'ensemble des racines n^e de l'unité dans $\mathbb{C}$.

Démonstration. Il est vite vérifié que tous les éléments de U_n sont des racines de l'unité parce que

$$\left(e^{2i\pi k/n}\right)^n = e^{2i\pi k} = 1, \tag{19.2}$$

entre autres à cause du lemme 18.11.

Cela nous donne déjà n racines pour $X^n - 1$ dans $\mathbb{C}$. Le théorème 6.111 nous indique qu'il ne peut pas y en avoir plus. □

19.3.
En ce qui concerne les notations, dans U_n, le « U » signifie « unité ». Cela n'a à peu près rien à voir avec le « U » du groupe $\mathrm{SU}(n)$; dans ce dernier, le « U » est pour « unitaire ».

Un des intérêts du groupe des racines est qu'il permet de factoriser $X^n - 1$, comme nous le verrons via les polynômes cyclotomiques dans le lemme 19.18.

LEMooBKTNooTmtUNQ

Lemme 19.4 ([562])**.**
Un nombre complexe algébrique dont tous les conjugués[1] *sont de module* 1 *est une racine de l'unité*[2].

LemWHQGooXyeJiw

Lemme 19.5.
L'ensemble U_n est un groupe cyclique[3] *d'ordre n généré par $\xi = e^{2i\pi/n}$.*

Démonstration. Il y a trois propriétés à vérifier pour que ce soit un groupe.

(1) **Neutre** Le nombre 1 est une racine de l'unité.
(2) **Inverse** Si $\omega \in U_n$ alors $\omega^n = 1$ et donc $\omega\omega^{n-1} = 1$, ce qui signifie que ω^{n-1} est un inverse de ω. Il reste à voir que $\omega^{n-1} \in U_n$. En effet $\left(\omega^{n-1}\right)^n = (\omega^n)^{n-1} = 1^{n-1} = 1$.

1. Définition 6.68.
2. Définition 19.1.
3. Définition 1.326.

(3) **Associativité** Cas particulier de l'associativité dans $\mathbb{C}$.

Le fait que ce soit un groupe cyclique contenant n éléments fait partie de la définition. □

Le lemme suivant donne les autres générateurs.

LemcFTNMa

Lemme 19.6.
Le nombre ξ^a est un générateur de U_n si et seulement si $\operatorname{pgcd}(a,n)=1$.

Démonstration. Si $\operatorname{pgcd}(a,n)=1$ alors le théorème de Bézout 1.236 nous fournit des entiers u et v tels que $ua+vn=1$. Alors nous avons

$$e^{2i\pi/n}=e^{2(ua+vn)i\pi/n}=(e^{2ai\pi/n})^u, \tag{19.3}$$

ce qui signifie que ξ est dans le groupe engendré par ξ^a, et par conséquent tout U_n est engendré.

Pour l'implication inverse, nous utilisons le théorème de Bézout dans le sens inverse. Soit ξ^a un générateur de U_n. Alors il existe u tel que $(\xi^a)^u=\xi$, donc $\xi^{au-1}=1$, c'est-à-dire qu'il existe v tel que $au-1=vn$. Cette dernière égalité implique que $\operatorname{pgcd}(a,n)=1$. □

Exemple 19.7.
Une conséquence tout à fait extraordinaire de ce lemme est que le nombre 7 est générateur de $\mathbb{Z}/12\mathbb{Z}$ (parce que $\operatorname{pgcd}(7,12)=1$). Or en solfège, une quinte fait 7 demi-tons, et une gamme en fait 12. Le cycle des quintes est donc générateur de la gamme chromatique[563]. Ce fait est connu des musiciens [4] depuis des siècles. △

PROPooIOQEooGMcCJm

Proposition 19.8 (Intersection par deux)**.**
Les ensembles U_α et U_β ont une intersection réduite à $\{1\}$ si et seulement si α et β sont premiers entre eux.

Démonstration. Nous rappelons qu'une racine α^{e} de l'unité peut s'écrire sous la forme $e^{2i\pi k/\alpha}$ avec $0\leqslant k<\alpha$.

(1) **Sens direct** Par contraposée, nous supposons que α et β ne sont pas premiers entre eux, et nous notons d leur pgcd. Nous nommons $\alpha=d\alpha'$ et $\beta=d\beta'$. Pour trouver une intersection entre U_α et U_β nous devons trouver une valeur de $0<k<\alpha$ telle que

$$(e^{2i\pi k/\alpha})^\beta=e^{2i\pi k\beta/\alpha}=1, \tag{19.4}$$

c'est-à-dire une valeur de k telle que $k\beta/\alpha$ soit un entier. Mais $k\beta/\alpha=k\beta'/\alpha'$ et par conséquent prendre $k=\alpha'$ fonctionne. Surtout que par hypothèse $d>1$ et donc $k=\alpha'<\alpha$.

(2) **Sens réciproque** Supposons maintenant que α et β soient premiers entre eux. Soit $z\in U_\alpha\cap U_\beta$. Le fait que z soit une racine α^{e} de l'unité implique qu'il existe un $k<\alpha$ tel que $z=e^{2i\pi k/\alpha}$. Mais si z est également une racine β^{e} de l'unité, alors $z^\beta=1$, c'est-à-dire que $k\beta/\alpha$ doit être un entier, soit l cet entier. Nous avons

$$k\beta=l\alpha. \tag{19.5}$$

Si $k>0$, comme le nombre α divise $k\beta$, cela conduirait via le lemme de Gauss 3.12 à dire que α divise k. Mais α ne peut pas diviser k parce que nous avions supposé que k était strictement plus petit que α. Donc $k=0$ et $z=1$.

□

PropFDDHooEyYxBC

Proposition 19.9 (Intersection : le cas général[1])**.**
Soient des entiers positifs $\alpha_1,\dots,\alpha_p$. Nous avons

$$\bigcap_{i=1}^{p}U_{\alpha_i}=\{1\} \tag{19.6}$$

si et seulement si $\operatorname{pgcd}(\alpha_1,\dots,\alpha_p)=1$ *(c'est-à-dire que les α_i sont premiers dans leur ensemble).*

4. Même ceux qui ignorent le théorème de Bézout.

Démonstration. Nous décomposons les α_i en facteurs premiers[5] de la façon suivante : $\alpha_i = \prod_{k\in\mathbb{N}} p_k^{\alpha_i^{(k)}}$ où les p_k sont les nombres premiers.

(1) **Caractérisation par une décomposition en facteurs premiers** Les éléments z différents de 1 dans U_{α_1} s'écrivent sous la forme

$$z = e^{2i\pi k/\alpha_1} \tag{19.7}$$

avec $0 < k < \alpha_1$.
Pour tout $i \neq 1$, le fait que $z \in U_{\alpha_i} \cap U_{\alpha_1}$ se traduit par le fait que $\left(e^{2i\pi k/\alpha_1}\right)^{\alpha_i} = 1$, c'est-à-dire que $\alpha_i k/\alpha_1$ est entier, donc que α_1 divise $k\alpha_i$. Par conséquent il existera un élément différent de 1 dans l'intersection des U_{α_i} si et seulement si il existe un entier k strictement compris entre 0 et α_1 pour lequel α_1 divise tous les $k\alpha_i$.
Un entier $0 < k < \alpha_1$ convient si et seulement si pour tout l, la puissance de p_l dans la décomposition de k est au moins égale à

$$\alpha_1^{(l)} - \alpha_i^{(l)} \tag{19.8}$$

pour tout l.

(2) **Sens direct** L'hypothèse $\operatorname{pgcd}(\alpha_1, \dots, \alpha_p) \neq 1$ implique qu'il existe un l pour lequel tous les $\alpha_i^{(l)}$ sont non nuls. Nous construisons le k voulu en prenant pour tout p_i la même puissance que celle dans α_1, sauf pour p_l pour lequel nous prenons la puissance $\alpha_1^{(l)} - \min_i\{\alpha_i^{(l)}\}$. Le minimum en question est strictement positif, ce qui donne un k strictement inférieur à α_1.

(3) **Sens réciproque** Si $\operatorname{pgcd}(\alpha_1, \dots, \alpha_p) = 1$ alors pour tout l, il existe un i tel que $\alpha_i^{(l)} = 0$. Donc pour tout l, la puissance de p_l dans la décomposition de k est au moins $\alpha_1^{(l)}$. Cela implique que $k \geqslant \alpha_1$, ce qui est impossible.

□

19.1.2 Fonction indicatrice d'Euler

Voir le thème 11.

DefLYGTooFPOYGZ

Définition 19.10.
Les générateurs de U_n sont les ***racines primitives***[6] *de l'unité dans $\mathbb{C}$. Nous nommons Δ_n leur ensemble :*

$$\Delta_n = \{e^{2i\pi k/n} \text{ tel que } 1 \leqslant k \leqslant n \text{ et } \operatorname{pgcd}(k,n) = 1\}. \tag{19.9}$$

Nous avons par exemple

— Pour $n = 1$, $\operatorname{pgcd}(k,1) = 1$ et $1 \leqslant k \leqslant 1$ uniquement pour $k = 1$.
— Pour $n = 2$, $\operatorname{pgcd}(k,2) = 1$ et $1 \leqslant k \leqslant 2$ uniquement pour $k = 1$.
— Pour $n = 3$, $\operatorname{pgcd}(k,3) = 1$ et $1 \leqslant k \leqslant 3$ uniquement pour $k = 1$, $k = 2$.

$$\Delta_1 = \{1\} \tag{19.10a}$$
$$\Delta_2 = \{e^{i\pi}\} \tag{19.10b}$$
$$\Delta_4 = \{e^{i\pi/2}, e^{i3\pi/2}\}. \tag{19.10c}$$

Notons que $1 \in \Delta_d$ seulement avec $d = 1$.

5. Théorème 3.15.
6. parce qu'en prenant les puissances successives de l'une d'entre elles, nous retrouvons toutes les racines de l'unité, voir aussi la définition 19.54.

PROPooWILSooFHWLnt

Proposition 19.11.
En posant[7]

$$U_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\}, \tag{19.11}$$

et[8]

$$\Delta_d = \{e^{2i\pi k/d} \text{ tel que } 1 \leqslant k \leqslant d \text{ et } \operatorname{pgcd}(k,d) = 1\}, \tag{19.12}$$

nous avons

$$U_n = \bigcup_{d \mid n} \Delta_d \tag{19.13}$$

EqpZuIyL

et l'union est disjointe.

Démonstration. En plein de petites parties.

(1) **Union disjointe** Soient d_1, d_2 tels que $x \in \Delta_{d_1} \cap \Delta_{d_2}$. Il existe k, l tels que $e^{2i\pi k/d_1} = e^{2i\pi l/d_2}$. Le corolaire 18.24 dit alors qu'il existe $s \in \mathbb{N}$ tel que

$$\frac{k}{d_1} = \frac{l}{d_2} + s = \frac{l + sd_2}{d_2}. \tag{19.14}$$

Vu que k/d_2 est une fraction irréductible, nous avons $d_2 \leqslant d_1$ par la proposition 3.34. Nous faisons de même en permutant les rôles de d_1 et d_2 : il existe $t \in \mathbb{N}$ tel que

$$\frac{l}{d_1} = \frac{k + d_1 t}{d_1}, \tag{19.15}$$

de telle sorte que $d_1 \leqslant d_2$. Nous en déduisons que $d_1 = d_2$. Donc l'union à droite dans (19.13) est disjointe.

(2) **Première inclusion** Soit $x \in U_n$. Il existe $0 \leqslant k \leqslant n-1$ tel que $x = e^{2i\pi k/n}$. Si $x = 1$, alors $x \in \Delta_1$; nous supposons que ce n'est pas le cas, c'est-à-dire que $k \neq 0$. Nous écrivons la fraction k/n sous forme irréductible :

$$\frac{k}{n} = \frac{k'}{d} \tag{19.16}$$

avec $d = \operatorname{pgcd}(k,n)$, et $\operatorname{pgcd}(k',d) = 1$. Donc $x \in \Delta_d$.

(3) **Seconde inclusion** Nous supposons que $x \in \Delta_d$ pour un certain diviseur d de n. Il existe donc $1 \leqslant k \leqslant d$ tel que $\operatorname{pgcd}(k,d) = 1$ et $x = e^{2i\pi k/d}$. Si $d = 1$ alors $x = 1$ et $x \in U_n$. Nous excluons ce cas. Vu que $d \mid n$, nous considérons $s \in \mathbb{N}$ tel que $n = sd$. Nous avons alors

$$\frac{k}{d} = \frac{ks}{sd} = \frac{ks}{n}. \tag{19.17}$$

Multiplions par s les inéquations $1 \leqslant k \leqslant d$:

$$s \leqslant ks \leqslant sd = n. \tag{19.18}$$

Le cas $ks = n$ est exclu parce que cela demanderait $kn = ksd = nd$, et donc $k = d$, ce que nous avons exclu parce que $k = d$ est uniquement possible pour $d = 1$ à cause de la condition $\operatorname{pgcd}(k,d) = 1$.

□

PROPooFKCHooWdFicM

Proposition 19.12.
Nous avons

$$\varphi(n) = \operatorname{Card}(\Delta_n) \tag{19.19}$$

EqEulerGqPsvi

φ est l'indicatrice d'Euler[9].

7. Voir la définition 19.2.
8. Définition 19.10. On pourrait croire que la condition $\operatorname{pgcd}(k,d) = 1$ exclu $k = d$. On aurait alors envie d'écrire $1 \leqslant k < d$. Mais pour $d = 1$, ça fait une différence parce que tel qu'écrit $\Delta_1 = \{1\}$.
9. Définition 5.29.

Démonstration. Nous prouvons que l'application

$$\begin{aligned} f\colon \{k \in \mathbb{N}^* \text{ tel que } 1 \leqslant k \leqslant n, \operatorname{pgcd}(k,n) = 1\} &\to \{e^{2i\pi k/n} \text{ tel que } 1 \leqslant k \leqslant n, \operatorname{pgcd}(k,n) = 1\} \\ k &\mapsto e^{2i\pi k/n} \end{aligned} \tag{19.20}$$

est une bijection.

Le fait que ce soit surjectif est immédiat. Pour l'injectivité, nous considérons $1 \leqslant k \leqslant n$ et $1 \leqslant l \leqslant n$ tels que $f(k) = f(l)$. Nous avons donc $e^{2i\pi k/n} = e^{2i\pi l/n}$. Le corolaire 18.24 nous dit alors qu'il existe $s \in \mathbb{Z}$ tel que

$$2\pi k/n = 2\pi l/n + 2\pi s. \tag{19.21}$$

En simplifiant par 2π et en multipliant par n, nous trouvons $k = l + sn$. Étant donné que k et l sont tout deux entre 1 et n, nous avons obligatoirement $s = 0$ et donc $k = l$.

Donc f est bijective et $\operatorname{Card}(\Delta_n) = \varphi(n)$. □

19.1.3 Décomposition en éléments simples

SUBSECooSBXooDIHUpgD

Lemme 19.13 ([564, 1]).
Soient des nombres complexes distincts $a_1, \dots, a_N$. Alors pour tout $z \notin \{a_i\}_{i=1,\dots,N}$,

$$\prod_{i=1}^{N} \frac{1}{z - a_i} = \sum_{i=1}^{N} \frac{\lambda_i}{z - a_i} \tag{19.22}$$

où

$$\lambda_i = \prod_{j \neq i} \frac{1}{a_i - a_j}. \tag{19.23}$$

Démonstration. Nous posons $P(z) = \prod_{k=1}^{N}(z - a_k)$, et nous calculons :

$$\sum_{i=1}^{N} \frac{\lambda_i}{z - a_i} = \sum_{i=1}^{N} \frac{\lambda_i \prod_{k \neq i}(z - a_k)}{\prod_{k=1}^{N}(z - a_k)} \tag{19.24a}$$

$$= \frac{1}{P(z)} \sum_{i=1}^{N} \lambda_i \prod_{k \neq i}(z - a_k) \tag{19.24b}$$

Il s'agit maintenant de prouver que la somme de gauche vaut toujours 1. Nous posons

$$S(z) = \sum_{i=1}^{N} \frac{\prod_{k \neq i}(z - a_k)}{\prod_{k \neq i}(a_i - a_k)}. \tag{19.25}$$

Calculons $S(a_l)$:

$$S(a_l) = \sum_{i=1}^{N} \frac{\prod_{k \neq i}(a_l - a_k)}{\prod_{k \neq i}(a_i - a_k)}. \tag{19.26}$$

Pour les termes $i \neq l$, le numérateur est nul, car il contient le facteur $a_l - a_l = 0$. Donc la somme se réduit au seul terme $i = l$:

$$S(a_l) = \frac{\prod_{k \neq l}(a_l - a_k)}{\prod_{k \neq l}(a_l - a_k)} = 1. \tag{19.27}$$

Le polynôme $S - 1$ est donc un polynôme de degré $N - 1$ qui possède N racines distinctes. Le théorème 6.111 implique que $S - 1 = 0$ et donc que $S = 1$ comme nous le voulions. □

Il est possible de décomposer une fraction rationnelle en fractions dites « simples ». Si $|z| < 1$ nous avons par exemple la décomposition

EqDWYBooJIMBAt

$$\frac{1}{1 - z^r} = \sum_{\omega \in U_r} \frac{A_\omega}{\omega - z} \tag{19.28}$$

où U_r est le groupe des racines r^{e} de l'unité défini en (19.1). Les nombres A_ω peuvent alors être déterminés en effectuant la somme. Le dénominateur commun sera $1-z^r$ tandis que les A_ω sont déterminés en égalant le numérateur à 1.

Exemple 19.14.
Pour décomposer la fraction $\frac{1}{1-x^2}$ nous savons que les racines du dénominateur sont ± 1. Donc nous écrivons

$$\frac{1}{1-x^2}=\frac{A}{1-x}+\frac{B}{1+x}. \tag{19.29}$$

Nous trouvons les valeurs de A et B en effectuant la somme :

$$\frac{A(1+x)+B(1-x)}{1-x^2}=\frac{A+B+(A-B)x}{1-x^2}. \tag{19.30}$$

Les coefficients A et B doivent donc vérifier $A+B=1$ et $A-B=0$. Au final,

$$\frac{1}{1-x^2}=\frac{1}{2(1-x)}+\frac{1}{2(1+x)}. \tag{19.31}$$

△

19.2 Chiffrement RSA

SecEVaFYi

Ce passage sur RSA provient en bonne partie de la page de Wikipédia[565].

Alice veut envoyer un message à Bob. L'idée est que Bob va donner à Alice une clef publique qui va permettre de chiffrer le message, tandis que Bob va garder pour lui une clef privée qui permet de déchiffrer.

19.2.1 Mise en place par Bob

Bob se crée une paire de clef publique, clef privée de la façon suivante.

(1) Bob choisit deux nombres premiers distincts p, q.
(2) Il calcule $n=pq$.
(3) Par le corolaire 5.43, l'indicatrice d'Euler $\varphi(n)=(p-1)(q-1)$ est facile à calculer pour Bob.
(4) Bob choisit $e\in\mathbb{N}$ premier avec $\varphi(n)$, puis d tel que $ed\in[1]_{\varphi(n)}$.

Maintenant la paire est : clef publique (n,e) et clef privée (n,d) [10].

Bob envoie la paire (n,e) à Alice.

Remarque 19.15.
Ici nous ne supposons pas que la communication soit sure. Une tierce personne peut intercepter le message. D'ailleurs en principe, les gens publient leur clef publique sur leurs sites, voire sur des sites dédiés. Le problème de l'identification reste à résoudre à l'ancienne.

19.2.2 Chiffrement

Nous chiffrons en utilisant la clef publique (n,e). D'abord Alice se débrouille pour transformer son message en un nombre plus petit que n. Soit M ce message. Alice code M en

$$C=M^e \mod n. \tag{19.32}$$

Tout le truc est que nous allons voir que l'application $x\mapsto x^e$ est une bijection de $\mathbb{F}_n$, et que l'inverse est facile à calculer par Bob, et difficile pour les autres. Alice envoie C à Bob. Encore une fois, nous ne supposons pas que cette communication soit privée. Le nombre C peut être intercepté.

10. Le fait que e soit public et d soit privé est une convention. e comme *encryption* et d comme *decryption*.

19.2.3 Déchiffrement

Nous allons montrer que $M = C^d \mod n$, et donc que Bob, connaissant (n, d), peut déchiffrer. D'abord

$$C^d = (M^e)^d = M^{ed}, \tag{19.33}$$

mais nous savons qu'il existe k tel que

$$ed = 1 + k\varphi(n) = 1 + k(p-1)(q-1). \tag{19.34}$$

L'étape astucieuse est de remarquer que

$$M^{1+k(p-1)(q-1)} \in [M]_p \cap [M]_q. \tag{19.35}$$ EqreeHgn

Pour montrer cela nous utilisons le petit théorème de Fermat 6.15(4).

— Si M est premier avec p, alors $M^{p-1} \in [1]_p$.

— Si M n'est pas premier avec p, alors M est multiple de p et on sait que $M^{p-1} \in [0]_p = [M]_p$.

Dans les deux cas nous avons (19.35). Le nombre $M^{1+k\varphi(n)} - M$ est donc à la fois multiple de p et de q.

Le lemme chinois 6.28 nous dit immédiatement [11] qu'alors

$$M^{1+k\varphi(n)} - M \tag{19.36}$$

est un multiple de $pq = n$, c'est-à-dire que

$$C^d = M^{ed} \in [M]_n. \tag{19.37}$$

Si on ne croit pas au lemme chinois, on peut utiliser le lemme de Gauss. Posons

$$M^{1+k\varphi(n)} - M = ap = bq. \tag{19.38}$$

Dans ce cas p divise bq, mais q est premier avec p, donc le lemme de Gauss 3.74 nous enseigne [12] que p divise b.

19.2.4 Une imprudence à ne pas commettre

Nous avons pris deux cas selon que M soit ou non premier avec p. Une question qui se pose est la suivante : est-ce que c'est une bonne idée d'envoyer un message qui ne soit pas premier avec p ?

Si nous savons que M n'est pas premier avec p, alors nous avons $M^e = l^e p^e$ et $n = pq$ qui sont publics. Donc un calcul de PGCD permettrait de trouver p.

Il faut cependant savoir que PageAKTBooMDeQxY

— La probabilité que ça arrive est infime : comme M est entre 0 et $n = pq$, les multiples de p possibles sont $p, 2p, \cdots pq$. Il y a donc une chance sur p que cela arrive. Typiquement avec des p de l'ordre de 10^{120}, on peut utiliser RSA chaque milliseconde sur chaque atome de l'univers depuis le début des temps que ça ne se serait presque certainement pas encore produit.

— De toutes façons, Alice ne sait pas vérifier si son message est premier avec p, parce qu'elle ne connaît pas p.

— En conclusion, la partie de la preuve qui montre que $M^{1+\varphi(n)} \in [M]_p \cap [M]_q$ dans le cas M non premier avec p est, à toutes fins pratiques, inutile parce que ce cas de figure ne se présentera jamais dans toutes l'histoire de l'univers, même pas avec une civilisation intelligente autour de chaque étoile.

ProbGAYFooZATuYy

¡¡ Avertissement/question au lecteur !! 19.16
Est-ce que ces trois points sont corrects ?

11. C'est ici qu'il est important que p ne soit pas égal à q. Si $p = q$, alors le lemme chinois ne fonctionne pas.
12. Ici aussi, si $p = q$, ça ne marche pas.

19.2.5 Problèmes calculatoires

Pour implémenter RSA, il faut pouvoir faire (au moins) trois choses :

(1) Trouver de grands nombres premiers.

(2) Trouver des couples de Bézout.

(3) Calculer M^e lorsque e est très grand.

En ce qui concerne le problème de trouver des nombres premiers, c'est compliqué, mais il faut savoir qu'il y en a plein. À 120 chiffres, il y a environ autant de nombres premiers que d'atomes dans 10^{20} fois l'univers connu. Cela rend impossible toute tentative de factoriser un grand nombre en essayant toutes les possibilités. Même pas en science-fiction.

Trouver des nombres u et v tels que $Au + Bv = \text{pgcd}(A, B)$ est un problème expliqué en 3.2.1.

En ce qui concerne le calcul de M^e lorsque e est grand, il n'est évidemment pas pensable de faire $M \cdot M \cdot \ldots M$ avec e facteurs. Un truc pour calculer en moins d'étapes est l'**exponentiation rapide**. Si $e = 2k$ est pair, nous calculons

$$M^e = (M^k)^2; \tag{19.39}$$

si $e = 2k + 1$ alors nous calculons

$$M^e = M(M^k)^2. \tag{19.40}$$

Le calcul prend alors seulement environ $\log_2(e)$ étapes. Pour donner une idée,

$$\log_2(10^{120}) \simeq 400. \tag{19.41}$$

Très raisonnable, mais un ordinateur reste indispensable.

19.2.6 La solidité de RSA

La solidité de la méthode repose sur deux conjectures (non démontrées !!) :

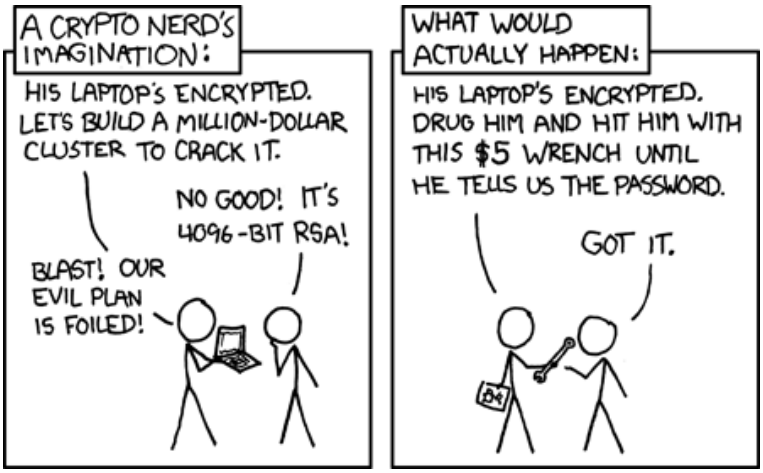
— Pour déchiffrer il *faut* connaitre p et q.

— La difficulté de trouver p et q en partant de $n = pq$ est exponentielle en n.

Dans la méthode de déchiffrage proposée ici, p et q sont utilisés pour calculer d qui est solution de $ed = [1]_{\varphi(n)}$. La seule formule connue pour calculer $\varphi(n)$ est $\varphi(n) = (p-1)(q-1)$. Si on trouve plus simple, alors RSA peut être craqué.

19.2.7 Note non mathématique pour doucher l'enthousiasme

Il est souvent dit[566] que différents systèmes de chiffrement peuvent aider à avoir des discussions « discrètes » dans les régimes totalitaires. La technologie au service de la démocratie, voilà qui enthousiasme la jeunesse [13]. La réalité est qu'il est souvent possible de craquer un système de chiffrement arbitrairement complexe, même sans connaitre le petit théorème de Fermat ...



13. Cela dit, le navigateur Tor[567], qui est un pur produit de RSA, permet effectivement d'accéder en France aux sites bloqués pour apologie du terrorisme (mars 2015).

http://xkcd.com/538/ Creative Commons Attribution-NonCommercial 2.5 License.

...tout dépend du contexte.

19.3 Polynômes cyclotomiques

19.3.1 Définitions et propriétés

DefXGHooRAXlpp

Définition 19.17.
*Le **polynôme cyclotomique** d'indice n est le polynôme*

$$\phi_n(X) = \prod_{z\in\Delta_n} (X - z) \tag{19.42}$$

EqLiGYKK

où Δ_n est l'ensemble des racines primitives de l'unité de la définition 19.10 :

$$\Delta_n = \{e^{2i\pi k/n} \text{ tel que } 1 \leqslant k \leqslant n \text{ et } \operatorname{pgcd}(k, n) = 1\}, \tag{19.43}$$

voir la définition 19.10.

Le polynôme ϕ_n est un polynôme unitaire de degré $\varphi(n)$ où φ est l'indicatrice d'Euler [14]. Nous avons par exemple

$$\Delta_1 = \{1\} \tag{19.44a}$$
$$\Delta_2 = \{-1\} \tag{19.44b}$$
$$\Delta_3 = \{e^{2i\pi/3}, e^{4i\pi/3}\} \tag{19.44c}$$

et les premiers polynômes cyclotomiques sont donnés par

$$\phi_1(X) = X - 1 \tag{19.45a}$$
$$\phi_2(X) = X + 1 \tag{19.45b}$$
$$\phi_3(X) = X^2 + X + 1. \tag{19.45c}$$

Pour le dernier nous avons utilisé le fait que $e^{6i\pi/3} = 1$ et $e^{2i\pi/3} + e^{4i\pi/3} = -1$.

LemKYGBooAwpOHD

Lemme 19.18.
Le polynôme $X^n - 1$ se factorise des diverses manières suivantes :

$$X^n - 1 = \prod_{z\in U_n} (X - z) = \prod_{d\,|\,n} \prod_{z\in\Delta_d} (X - z) = \prod_{d\,|\,n} \phi_d(X) \tag{19.46}$$

où U_n est défini en 19.2.

Démonstration. En ce qui concerne la première égalité, tous les éléments de U_n sont des racines simples de $X^n - 1$. Donc le théorème 3.130 dit qu'il existe un nombre k (polynôme de degré zéro) tel que $X^n - 1 = k\prod_{z\in U_n}(X - z)$. Vu le coefficient du terme de plus haut degré, ce k ne peut être que 1.

Pour la suite nous utilisons l'union disjointe $U_n = \bigcup_{d\,|\,n} \Delta_d$ de la proposition 19.11 et la définition (19.42) des polynômes cyclotomiques. □

Remarque 19.19.
Notons juste pour le plaisir que dans le produit $\prod_{d\,|\,n}\prod_{z\in\Delta_d}$, il y a bien n termes parce que $\operatorname{Card}(\Delta_d) = \varphi(d)$ et $\sum_{d\,|\,n}\varphi(d) = n$ (définition 19.10 et proposition 5.38).

Proposition 19.20.
Les polynômes cyclotomiques sont à coefficients entiers : $\phi_n \in \mathbb{Z}[X]$.

14. Définie par l'équation 19.19.

Démonstration. Nous devons démontrer que les coefficients de ϕ_n sont dans $\mathbb{Z}$ alors qu'ils sont à priori dans $\mathbb{C}$. Nous démontrons cela par récurrence. D'abord $\phi_1(X) = X - 1$, d'accord. Ensuite

$$X^{n+1} - 1 = \prod_{d|n+1} \phi_d(X) = \phi_{n+1}(X) \cdot \underbrace{\prod_{\substack{d|n+1\\ d\leqslant n}} \phi_d(X)}_{\in\mathbb{Z}[X] \text{ par récurrence}} \tag{19.47}$$

Le lemme 6.49 conclut que $\phi_{n+1} \in \mathbb{Z}[X]$. Nous avons considéré $\mathbb{Z}$ comme sous anneau du corps $\mathbb{C}$. □

PropUImYnL

Proposition 19.21.
Soient $1 \leqslant m \leqslant n$, deux entiers, et

$$T(X) = \frac{X^n - 1}{X^m - 1} \in \mathbb{Z}(X). \tag{19.48}$$

Alors :

ItemhpDPKE

(1) si $m \mid n$ alors $T \in \mathbb{Z}[X]$,
(2) si $m \mid n$ et si $m < n$ alors ϕ_n divise T dans $\mathbb{Z}[X]$.

Démonstration. Nous prouvons point par point.

(1) Si m divise n alors les diviseurs de n sont l'union des diviseurs de m et des diviseurs de n qui ne divisent pas m. Soit

$$Q = \{\text{diviseurs de } n \text{ ne divisant pas } m\}. \tag{19.49}$$

Nous avons alors

$$X^n - 1 = \prod_{d|n} \phi_d(X) = \prod_{d|m} \phi_d(X) \cdot \prod_{q\in Q} \phi_q(X) = (X^m - 1) \cdot \prod_{q\in Q} \phi_q(X). \tag{19.50}$$

Nous avons donc

$$T(X) = \frac{X^n - 1}{X^m - 1} = \prod_{q\in Q} \phi_q(X) \in \mathbb{Z}[X]. \tag{19.51}$$

(2) Nous venons de montrer que

$$T = \prod_{q\in Q} \phi_q \in \mathbb{Z}[X]. \tag{19.52}$$

Étant donné que $m < n$ nous avons $n \in Q$ et donc

$$T = \phi_n \cdot \prod_{q\in Q\setminus\{n\}} \phi_q. \tag{19.53}$$

Par conséquent ϕ_n divise T dans $\mathbb{Z}[X]$.

□

CorTVUooErJiAC

Corolaire 19.22.
Si p est premier alors le polynôme cyclotomique ϕ_p a une bonne tête :

$$\phi_p(X) = 1 + X + \cdots + X^{p-1}. \tag{19.54}$$

Démonstration. Nous utilisons la formule du lemme 19.18 en remarquant que seuls p et 1 divisent p :

$$X^p - 1 = \prod_{d|p} \phi_d(X) = \phi_1(X)\phi_p(X) = (X - 1)\phi_p(X). \tag{19.55}$$

Nous pouvons simplifier par $X - 1$ en utilisant la formule du lemme 3.136(2) :

$$1 + X + \cdots + X^{p-1} = \phi_p(X) \tag{19.56}$$

□

PropoIeOVh

Proposition 19.23 (Irréductibilité des polynômes cyclotomiques[568]).
Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}$.

Démonstration. Pour rappel, nous savons déjà que pour tout $n \in \mathbb{N}$, $\phi_n \in \mathbb{Z}[X]$. Puisque les racines de ϕ_n sont les racines primitives de l'unité, nous devons montrer que toutes les racines primitives de l'unité ont même polynôme minimal (qui sera alors ϕ_n) ; en effet comme ces polynômes divisent ϕ_n, s'ils sont distincts, la proposition 6.108 s'applique et le produit des polynômes minimaux diviserait ϕ_n. Dans le cas inverse, ϕ_n est polynôme minimal des racines primitives de l'unité et est donc irréductible. Soit donc ξ, une telle racine primitive. Une autre racine primitive est de la forme ξ^l où l est un nombre premier tel que $\operatorname{pgcd}(l, n) = 1$.

Soient f et g, les polynômes minimaux dans $\mathbb{Z}[X]$ de ξ et ξ^l. Nous allons montrer que $f = g$ et donc que $f = g = \phi_n$. Supposons par l'absurde que $f \neq g$. Dans ce cas ils seraient des facteurs irréductibles distincts de ϕ_n et il existerait un polynôme h tel que $\phi_n = fgh$. A priori, $h \in \mathbb{Q}[X]$ parce que nous sommes justement en train de prouver que ϕ_n est irréductible dans $\mathbb{Q}[X]$. Quoi qu'il en soit, le lemme de Gauss 6.51 nous montre que $h \in \mathbb{Z}[X]$ parce que ϕ_n, f et g ont des coefficients entiers. Nous avons

$$f(\xi) = g(\xi^l) = 0. \tag{19.57}$$

Considérons le polynôme $\psi(X) = g(X^l)$. Ce polynôme ψ est dans $\mathbb{Z}[X]$ et ψ est annulateur de ξ, donc f divise ψ en tant que polynôme minimal de ξ. Il y a un polynôme unitaire à coefficients entiers (lemme de Gauss forever) k tel que

$$\psi = fk \tag{19.58}$$

Nous considérons maintenant les projections sur $\mathbb{F}_l[X]$: étant donné que $\phi_n = fgh$, nous savons que $\bar{f}\bar{g}$ divise $\bar{\phi}_n$. En même temps, $\bar{f}$ divise $\bar{\psi}$. En utilisant le morphisme de Frobenius (c'est ici que la projection sur $\mathbb{F}_l$ joue), nous avons aussi

$$\bar{\psi}(X) = \bar{g}(X^l) = \bar{g}(X)^l. \tag{19.59}$$

Par conséquent dire que $\bar{f}$ divise $\bar{\psi}$ revient à dire que $\bar{f}(X)$ divise $\bar{g}(X)^l$. En particulier tout facteur irréductible de $\bar{f}$ divise $\bar{g}$. Un facteur irréductible de $\bar{f}$ serait donc à la fois dans $\bar{f}$ et dans $\bar{g}$ et donc deux fois (au moins) dans $\bar{\phi}_n$ parce que $\bar{f}\bar{g}$ divise ϕ_n. Dans un corps de décomposition de ce facteur, ϕ_n aurait une racine double, alors que ce n'est pas le cas. Contradiction. Nous concluons que $f = g$. □

Le corolaire suivant va être utilisé pour déterminer les polygones constructibles à la règle et au compas, théorème de Gauss-Wantzel 19.87.

CorKRTooTJtyvP

Corolaire 19.24.
Soit p un nombre premier et α un entier non nul. Nous posons $q = p^\alpha$. Alors le polynôme minimal de $e^{2i\pi/q}$ sur $\mathbb{Q}$ est le polynôme cyclotomique ϕ_q.

Démonstration. Le polynôme ϕ_q est irréductible par la proposition 19.23, il est unitaire par définition et contient le monôme $X - e^{2i\pi/q}$, donc il est annulateur. Annulateur, irréductible et unitaire, la proposition 6.82(2) en fait le polynôme minimal de $e^{2i\pi/q}$. □

ThojCJpFW

Théorème 19.25.
Soit $P \in \mathbb{Z}[X]$ un polynôme unitaire irréductible non constant tel que toutes les racines dans $\mathbb{C}$ soient de module $\leqslant 1$. Alors, soit $P = X$, soit P est un polynôme cyclotomique.

Démonstration. Le polynôme $P = X$ vérifie les conditions. Pour la suite, nous supposons que $P \neq X$.

Nous notons $P = \sum_i a_i X^i$. Étant donné que P est irréductible et différent de X, nous avons $a_0 \neq 0$ (sinon $x = 0$ serait une racine). Nous allons montrer que les racines de P sont toutes des racines N-ièmes de l'unité (avec le même N pour toutes).

Soient $\{\xi_i\}_{i=1,\dots,d}$ les racines de P ; on a

$$P = \prod_{i=1}^{d}(X - \xi_i) \tag{19.60}$$

avec $\prod_{i=1}^{d}\xi_i = a_0$. Par hypothèse, $|\xi_i| \leqslant 1$ et donc $0 < |a_0| \leqslant 1$. Puisque $P \in \mathbb{Z}[X]$ nous avons donc $a_0 = 1$ et $|\xi_i| = 1$, pour tout i.

Nous introduisons les polynômes

$$g_q(X) = \prod_{i=1}^{d}\big(X - (\xi_i)^q\big), \tag{19.61}$$

et en particulier $g_1 = P$, et nous développons

$$g_q(X) = X^n + C_{1,q}X^{n-1} + \cdots + C_{n,q} \tag{19.62}$$

où

$$C_{k,q} = (-1)^k \sum_{1\leqslant i_1<\dots<i_k\leqslant d} (\xi_{i_1}\dots\xi_{i_k})^q. \tag{19.63}$$

Nous introduisons aussi les polynômes

$$F_{k,q}(X_1,\dots,X_n) = (-1)^k \sum_{1\leqslant i_1<\dots<i_k\leqslant d} (X_{i_1}\dots X_{i_k})^q \tag{19.64}$$

qui sont des polynômes symétriques. Ils vérifient deux propriétés. La première est que

$$C_{r,q} = F_{r,q}(\xi_1,\dots,\xi_n), \tag{19.65}$$

et la seconde est que les polynômes $F_{r,1}$ sont les polynômes symétriques élémentaires à un coefficient près. Le théorème 6.181 nous donne alors des polynômes $G_{k,q} \in \mathbb{Z}[X_1,\dots,X_n]$ tels que

$$F_{k,q}(X_1,\dots,X_n) = G_{k,q}\big(F_{1,1}(X_1,\dots,X_n),\dots,F_{k,1}(X_1,\dots,X_n)\big). \tag{19.66}$$

Nous savons que

$$|C_{k,q}| \leqslant \sum_{1\leqslant i_1<\dots<i_k<d} 1 = \binom{d}{k}. \tag{19.67}$$

Donc g_q fait partie de l'ensemble fini des polynômes dans $\mathbb{Z}[q]$ dont tous les coefficients sont bornés en valeur absolue par

$$\max_{k=1,\dots,d}\binom{d}{k}. \tag{19.68}$$

Il existe un certain nombre d'ensembles $\{\xi_i\}$ qui sont racines de polynômes vérifiant les conditions du théorème. À chacun de ces ensembles est associé une suite de polynômes g_q et donc des coefficients $C_{k,q}$. Ce que nous avons vu est que l'ensemble de tous les coefficients $C_{k,q}$ possibles (pour un choix donné des $\{\xi_i\}$) est fini, en particulier, comme $C_{1,q} = \sum_i \xi_i^q$, pour chaque k, l'ensemble

$$\{\xi_k^q \text{ tel que } q \in \mathbb{N}\}. \tag{19.69}$$

Par le principe des tiroirs, il existe q_1 et q_2 tels que $\xi_k^{q_1} = \xi_k^{q_2}$. Ici, q_1 et q_2 dépendent de k et nous notons $N_k = q_1 - q_2$; nous avons donc $\xi_k^{N_k} = 1$.

En posant $N = \text{ppcm}(N_1,\dots,N_d)$, nous avons

$$\xi_k^N = 1 \tag{19.70}$$

pour tout k.

Mais P est irréductible dans $\mathbb{Z}[X]$; si il a ±1 comme racines, alors c'est que $P = X + 1$ ou $P = X - 1$ et ce sont des polynômes cyclotomiques. Si P n'a pas ±1 parmi ses racines, alors P n'a pas de racines dans $\mathbb{Q}$ parce que ±1 sont les seules racines de $X^N - 1$ dans $\mathbb{Q}$.

Par conséquent P est un facteur irréductible de $X^N - 1$ dans $\mathbb{Q}[X]$. Mais étant donné que

$$X^N - 1 = \prod_{d\,|\,N}\phi_d(X), \tag{19.71}$$

les polynômes cyclotomiques sont les seuls facteurs irréductibles de $X^N - 1$. Donc P est un polynôme cyclotomique. □

19.3.2 Nombres premiers

LemiAqLEn

Lemme 19.26 ([569]).
Soit $n \geqslant 1$. Il existe un nombre premier p et un entier a tels que

(1) p divise $\phi_n(a)$,

(2) p ne divise aucun de $\phi_d(a)$ avec $d \mid n$ et $d \neq n$.

De tels p et a vérifient automatiquement

(1) p divise $a^n - 1$,

(2) p ne divise aucun des $a^d - 1$ pour $d \mid n$, $d \neq n$.

Démonstration. Nous posons

$$B(X) = \prod_{\substack{d \mid n \\ d \neq n}} \phi_d(X), \tag{19.72}$$

et nous commençons par montrer que ϕ_n est premier avec B. Nous avons $X^n - 1 = B\phi_n$, donc B et ϕ_n n'ont pas de racine commune (même pas dans $\mathbb{C}$) parce que ce serait une racine double de $X^n - 1$. Notons que par définition 19.42, les polynômes cyclotomiques sont scindés (dans $\mathbb{C}$), donc en particulier les polynômes ϕ_n et B sont scindés, et donc premiers entre eux, dans $\mathbb{C}$ et a fortiori dans $\mathbb{Q}$. Par Bézout (corolaire 3.73), il existe $U, V \in \mathbb{Q}[X]$ tels que

$$U\phi_n + VB = 1. \tag{19.73}$$

Si nous prenons $a \in \mathbb{Z}$ tel que $U' = aU$ et $V' = aV$ soient tous deux dans $\mathbb{Z}[X]$, alors nous avons

EqCpNMEi

$$U'\phi_n + V'B = a, \tag{19.74}$$

égalité dans $\mathbb{Z}[X]$. Quitte à prendre un multiple assez grand de a, nous pouvons choisir a de telle sorte que $|\phi_n(a)| \geqslant 2$. Nous prenons alors un nombre premier p divisant $\phi_n(a)$.

Montrons que le a et le p ainsi construits satisfont aux exigences.

Puisque $X^n - 1 = B\phi_n$, si p divise $\phi_n(a)$, il divise automatiquement $a^n - 1$ et donc $[a^n]_p = 1$, ce qui signifie entre autres que a et p sont premiers entre eux. Évaluons l'équation (19.74) en a :

$$U'(a)\phi_n(a) + V'(a)B(a) = a. \tag{19.75}$$

Le nombre p ne divisant pas a, mais divisant $\phi_n(a)$, il ne peut pas diviser $B(a)$ [15]. Étant donné que p ne divise pas $B(a)$, il ne divise aucun des $\phi_d(a)$ avec $d \mid n$ et $d \neq n$.

Nous passons maintenant à la seconde partie de la preuve. Nous supposons avoir a et p tels que p soit un nombre premier divisant $\phi_n(a)$ et tels que p ne divise aucun des $\phi_d(a)$ avec $d \mid n$, $d \neq n$. Le fait de diviser $\phi_n(a)$ entraine le fait de diviser $a^n - 1$ parce que ϕ_n est un des facteurs de $X^n - 1$. Soit maintenant $d \neq n$ divisant n ; nous avons

EqwTWcCu

$$X^d - 1 = \prod_{d' \mid d} \phi_{d'}, \tag{19.76}$$

et cela est une partie du produit

$$\prod_{\substack{d \mid n \\ d \neq n}} \phi_d. \tag{19.77}$$

Puisque p ne divise aucun des $\phi_d(a)$ de ce dernier produit, a fortiori, il ne divise pas le produit 19.76, et donc pas $a^d - 1$. □

LemrZnmpG

Lemme 19.27.
Si $n \geqslant 1$, alors il existe un nombre premier dans $[1]_n$, c'est-à-dire un nombre premier de la forme $1 + kn$ avec $k \in \mathbb{N}\backslash\{0\}$.

15. C'est pour pouvoir dire ça que l'on a choisi $V' \in \mathbb{Z}[X]$ de telle sorte que $V'(a)$ soit dans $\mathbb{Z}$

Démonstration. Soient $n \geqslant 1$ et p, a les nombres donnés par le lemme 19.26. Puisque p divise $\phi_n(a)$, p divise $a^n - 1$ et donc $[a]_p$ a un ordre qui divise n dans $(\mathbb{Z}/p\mathbb{Z})\backslash\{0\}$ parce que $[a]_p^n = [1]_p$.

Prenons $d \neq n$ divisant n. Nous savons que

$$a^d - 1 = \prod_{d' \mid d} \phi_{d'}(a). \tag{19.78}$$

Par construction de a et p, nous avons

$$[\phi_{d'}(a)]_p \neq 0 \tag{19.79}$$

Comme $\mathbb{Z}/p\mathbb{Z}$ est intègre, le produit est également non nul, c'est-à-dire

$$\Big[\prod_{d' \mid d} \phi_{d'}(a)\Big]_p \neq 0, \tag{19.80}$$

et donc $[a]_p^a \neq 1$. Nous avons donc montré que si $d \neq n$ divise n, alors nous avons en même temps

$$[a]_p^n = 1 \tag{19.81}$$

et

$$[a]_p^d \neq 1. \tag{19.82}$$

Cela prouve que $[a]_p$ est d'ordre exactement n. Oui, mais l'ordre de $[a]_p$ doit diviser l'ordre du groupe $\mathbb{Z}/p\mathbb{Z}$ qui est $p-1$, donc n divise $p-1$ et nous écrivons $p = kn + 1$ avec k entier. □

ThoxwTjcl

Théorème 19.28 (Forme faible du théorème de Dirichlet [114])**.**
Pour tout $n \geqslant 1$, il existe une infinité de nombres premiers dans $[1]_n$.

Démonstration. Le lemme 19.27 nous donne déjà l'existence de nombres premiers dans $[1]_n$. Il faut maintenant voir qu'il y en a une infinité. Nous supposons qu'il y en ait seulement un nombre fini : $p_1, \dots, p_r$, et nous notons

$$N = np_1 \dots p_r. \tag{19.83}$$

Nous utilisons maintenant le lemme 19.27 avec ce N, c'est-à-dire qu'on a un nombre premier de la forme

$$p = 1 + kN = 1 + knp_1 \dots p_r. \tag{19.84}$$

C'est un nombre premier plus grand que tous les p_i, et de la forme $1 + \lambda n$. Cela contredit l'exhaustivité de la liste $p_1, \dots, p_r$. □

19.4 Corps finis

SecCorpsFinizkAcbS

Si vous cherchez des choses à propos de RSA, c'est à la section 19.2.

19.4.1 Théorème de Wedderburn

ThoMncIWA

Théorème 19.29 (Théorème de Wedderburn[165])**.**
Tout corps fini est commutatif.

Démonstration. Soit $\mathbb{K}$ un corps fini et Z, le centre de $\mathbb{K}$. Ce dernier est un corps fini et un sous-corps de $\mathbb{K}$. Si $q = \text{Card}(Z)$ alors par le lemme 6.58 nous avons

$$\text{Card}(\mathbb{K}) = q^n \tag{19.85}$$

pour un certain n.

Nous supposons maintenant que $\mathbb{K}$ est non commutatif. Dans ce cas $Z \neq \mathbb{K}$ et nous avons $n \geqslant 2$. Nous considérons aussi

$$Z_x = \{a \in \mathbb{K} \text{ tel que } ax = xa\}. \tag{19.86}$$

Le centre Z est un sous-corps de Z_x, donc il existe $d(x)$ tel que

$$\text{Card}(Z_x) = q^{d(x)}. \tag{19.87}$$

De la même manière, Z_x est un sous-corps de $\mathbb{K}$, donc il existe $m(x)$ tel que

$$\text{Card}(\mathbb{K}) = \text{Card}(Z_x)^{m(x)}. \tag{19.88}$$

En mettant bout à bout, nous avons

$$q^n = \text{Card}(Z_x)^{m(x)} = q^{d(x)m(x)}, \tag{19.89}$$

et par conséquent $n = d(x)m(x)$. Le point important à retenir est que $d(x)$ divise n pour tout $x \in \mathbb{K}$.

Nous considérons maintenant l'action adjointe du groupe $\mathbb{K}\backslash\{0\}$ sur lui-même :

$$\varphi(k)x = kxk^{-1}. \tag{19.90}$$

Nous notons $\mathcal{O}_x$ l'orbite de $x \in \mathbb{K}\backslash\{0\}$ pour cette action, et $\text{Fix}(x)$ son stabilisateur. Nous avons

$$Z_y = \text{Fix}(y) \cup \{0\} \tag{19.91}$$

parce que Z_y et $\text{Fix}(y)$ ont les mêmes définitions, sauf que $\text{Fix}(y)$ est dans $\mathbb{K}\backslash\{0\}$ alors que Z_y est dans $\mathbb{K}$. Nous avons donc

$$\text{Card}\big(\text{Fix}(y)\big) = \text{Card}(Z_y) - 1 = q^{d(y)} - 1. \tag{19.92}$$

Nous avons équivalence des affirmations suivantes :

— $\text{Card}(\mathcal{O}_x) = 1$
— $\mathcal{O}_x = \{x\}$
— $\text{Fix}(x) = \mathbb{K}\backslash\{0\}$
— $x \in Z\backslash\{0\}$.

Soient $z_0, \dots, z_{q-1}$ les éléments de Z avec $z_0 = 0$. Ce sont les éléments qui auront une orbite réduite à un point. Les orbites qui coupent $Z\backslash\{0\}$ sont

$$\{z_1\}, \dots, \{z_{q-1}\} \tag{19.93}$$

et il y en a $q-1$. Soient $\mathcal{O}_{y_1}, \dots, \mathcal{O}_{y_r}$, les autres orbites. Nous utilisons l'équation des classes (2.78) :

$$\text{Card}(\mathbb{K}^*) = \text{Card}(Z^*) + \sum_{i=1}^{r} \frac{\text{Card}(\mathbb{K}^*)}{\text{Card}(\text{Fix}(y_i))}, \tag{19.94}$$

mais $\text{Card}(Z^*) = q - 1$, $\text{Card}(\mathbb{K}^*) = q^n - 1$ et $\text{Card}\big(\text{Fix}(y_i)\big) = q^{d(y_i)} - 1$, donc

$$q^n - 1 = (q-1) + \sum_{i=1}^{r} \frac{q^n - 1}{q^{d(y_i)} - 1}. \tag{19.95}$$ EqBPBDzE

Nous considérons la fraction rationnelle

$$F(X) = (X^n - 1) - \sum_{i=1}^{r} \frac{X^n - 1}{X^{d(y_i)} - 1}. \tag{19.96}$$ EqATGciu

Étant donné que $d(y_i)$ divise n, nous avons, contrairement aux apparences, $F \in \mathbb{Z}[X]$, par la proposition 19.21(1).

Nous pouvons exploiter un peu mieux la proposition 19.21 en remarquant que $d(y_i) < n$ parce que sinon $\text{Card}(Z_{y_i}) = \text{Card}(\mathbb{K})$, ce qui signifierait que $y_i \in Z$, ce qui nous avions exclu. Par conséquent le polynôme cyclotomique ϕ_n divise

$$\frac{X^n - 1}{X^{d(y_i)} - 1} \tag{19.97}$$

dans $\mathbb{Z}[X]$. Le polynôme cyclotomique ϕ_n divise également $X^n - 1$ et par conséquent ϕ_n divise F. Il existe donc $Q \in \mathbb{Z}[X]$ tel que $F = Q\phi_n$. En particulier en évaluant en q :

$$F(q) = Q(q)\phi_n(q) = q - 1. \tag{19.98}$$ eqmoLdJy

En effet nous avons $F(q) = q - 1$ par construction : comparer (19.95) avec (19.96). Évidemment $q \neq 1$ parce que si $q = 1$ alors $\text{Card}(\mathbb{K}) = 1$ et le théorème est trivial. Par ailleurs $Q(q)$ est un entier (parce que $Q \in \mathbb{Z}[X]$ et $q \in \mathbb{N}$) et $Q(q) \neq 0$, parce qu'à droite de (19.98) nous avons $q - 1 \neq 0$. Nous avons donc $|Q(q)| \geqslant 1$ et donc

$$|\phi_n(q)| \leqslant q - 1. \tag{19.99}$$

Par définition du polynôme cyclotomique nous avons

$$|\phi_n(q)| = \prod_{z \in \Delta_n} |q - z|. \tag{19.100}$$

Étant donné que ce produit doit être inférieur à $q-1$, au moins un des termes doit l'être : il existe $z_0 \in \Delta_n$ tel que $|z_0 - q| \leqslant q - 1$. Étant donné que $n \geqslant 2$ nous avons $z_0 \neq 1$.

Mais d'autre part, comme indiqué sur la figure 19.1, la distance entre z_0 et q doit être strictement plus grande que $q-1$ parce que $q-1$ est le minimum de la distance entre le cercle trigonométrique et q, et n'est atteint qu'en $z = 1$.

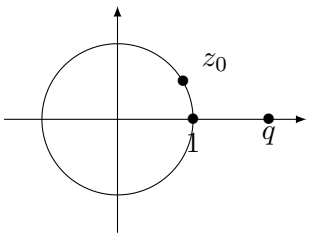


FIGURE 19.1: Nous devons avoir $|z_0 - q| > q - 1$. LabelFigtrigoWedd

Nous avons ainsi obtenu une contradiction, et nous concluons que le corps $\mathbb{K}$ est commutatif. □

19.4.2 Existence, unicité

Nous avons déjà défini le corps fini $\mathbb{F}_p$ lorsque p est un nombre premier dans la section 6.1.3. Le théorème suivant sert à définir $\mathbb{F}_{p^n}$ lorsque p est premier.

ThoOzgSfy

Théorème 19.30.
Soit p un nombre premier, soit $n \in \mathbb{N}\backslash\{0\}$ et $q = p^n$. Alors il existe un unique corps $\mathbb{K}$ de cardinal q. Ce corps est le corps de décomposition du polynôme $X^q - X$ sur $\mathbb{F}_p$.

Démonstration. Montrons l'unicité. Soit $\mathbb{K}$ un corps fini de cardinal $q = p^n$. Le groupe multiplicatif $\mathbb{K}^*$ est de cardinal $q - 1$, et par le corolaire 2.14 tous les éléments de $\mathbb{K}^*$ vérifient $g^{q-1} = e$, c'est-à-dire que dans $\mathbb{K}[X]$, les éléments de $\mathbb{K}^*$ sont des racines du polynôme

$$X^{q-1} - 1 \tag{19.101}$$

Par conséquent $\mathbb{K}$ est un corps de décomposition pour le polynôme $Q(X) = X^q - X = X(X^{q-1} - 1)$ parce que $Q(X) = 0$ dans $\mathbb{K}$. Il est unique par la proposition 6.146.

Montrons maintenant que le corps de décomposition de $P = X^q - X$ sur $\mathbb{F}_p$ est un corps de cardinal q. Pour ce faire nous considérons $\mathbb{K}$ ce corps de décomposition, et $\mathbb{E}$, l'ensemble des racines de P dans $\mathbb{K}$. Nous allons montrer que $\mathbb{E} = \mathbb{K}$ et que $\mathbb{E}$ est un corps contenant q éléments.

Montrons que $\mathbb{E}$ est un corps. Pour $\alpha, \beta \in \mathbb{E}$ nous avons

$$(\alpha\beta)^q = \alpha^q\beta^q = \alpha\beta \tag{19.102}$$

parce que $\alpha^q = \alpha$. Le produit $\alpha\beta$ est donc encore dans $\mathbb{E}$. Pour la somme,

$$(\alpha + \beta)^q = (\alpha + \beta)^{p^n} = \Big((\alpha + \beta)^p\Big)^{p^{n-1}} = (\alpha^p + \beta^p)^{p^{n-1}} = \ldots = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta. \tag{19.103}$$

En ce qui concerne l'inverse,

$$(\alpha^{-1})^q = (\alpha^q)^{-1} = \alpha^{-1}. \tag{19.104}$$

Donc $\mathbb{E}$ est un corps. Évidemment $\mathbb{E}$ est un corps de décomposition de P au sens où $\mathbb{E}$ est une extension de $\mathbb{F}_p$ sur lequel P est scindé (parce qu'il est scindé sur $\mathbb{K}$ et $\mathbb{E}$ est le sous-corps de $\mathbb{K}$ contenant les racines de P) et tel que $\mathbb{E} = \mathbb{F}_p(\{\alpha_i\})$ où les α_i sont les racines de P. Notons que $\mathbb{F}_p \subset \mathbb{E}$ parce que dans $\mathbb{F}_p$ on a $x^q = x$.

Par unicité, nous avons $\mathbb{K} = \mathbb{E}$. Nous devons montrer que P possède exactement q racines distinctes, afin d'avoir $\text{Card}(\mathbb{E}) = q$. Pour cela remarquons que

$$P'(X) = qX^{q-1} - 1 = -1 \tag{19.105}$$

dans $\mathbb{F}_p$. En effet $P \in \mathbb{F}_p$ et $q = 0$ dans $\mathbb{F}_p$. Par conséquent P' ne s'annule pas et P n'a pas de racine double. Toutes les racines étant simples, il y en a exactement q. □

Le théorème 19.30 ne permet pas de *construire* le corps à $q = p^n$ éléments. Nous allons maintenant voir un certain nombre de résultats donnant des façons de le construire. Ces résultats proviennent de [570, 571, 572] et de wikipedia

PropnfebjI

Proposition 19.31 ([572])**.**
Soit $\mathbb{K}$ un corps fini. Alors le groupe multiplicatif $\mathbb{K}\backslash\{0\}$ est cyclique.

Démonstration. Soit $\mathbb{K}$ un corps ayant q éléments. Le groupe $\mathbb{K}\backslash\{0\}$ en a $q - 1$; ergo l'ordre des éléments de $\mathbb{K}\backslash\{0\}$ sont des diviseurs de $q - 1$; c'est le corolaire 2.14. Soit d un diviseur de $q - 1$ et

$$H_d\backslash\{0\} = \{x \text{ d'ordre } d \text{ dans } \mathbb{K}\backslash\{0\}\} \tag{19.106a}$$

$$H_d = \{\text{racines de } X^d - 1 \text{ dans } \mathbb{K}\}. \tag{19.106b}$$

Ici le polynôme $X^d - 1$ est vu dans $\mathbb{K}[X]$. Notons que nous avons automatiquement $H_d^* \subset H_d$, mais l'inclusion inverse n'est pas assurée parce que les éléments d'ordre $d/2$ par exemple sont aussi dans H_d. Supposons $H_d^* \neq \varnothing$ et considérons $a \in H_d^*$. Alors l'application

$$\begin{aligned} \phi\colon \mathbb{Z}/d\mathbb{Z} &\to H_d \\ n &\mapsto a^n \end{aligned} \tag{19.107}$$

est un isomorphisme d'anneaux. En effet étant donné que $a \in H_d^* \subset H_d$, l'ensemble H_d contient le groupe cyclique engendré par a. Ce dernier contient, par construction, d éléments. Mais $\text{Card}(H_d) \leqslant d$ parce que H_d est l'ensemble des racines d'un polynôme de degré d. Par conséquent $\text{Card}(H_d) = d$ et l'ensemble H_d est bien engendré par a et ϕ est bien un isomorphisme. Par conséquent tous les éléments de H_d^* sont des générateurs de H_d.

Inversement soit x un générateur de H_d. L'ordre de H_d étant d, l'ordre de x doit être un diviseur de d. Supposons donc que x soit d'ordre d/k. Dans ce cas nous devrions avoir $\text{Card}(H_d) = d/k$, ce qui contredit l'isomorphisme ϕ.

En conclusion, H_d^* est l'ensemble des générateurs du groupe H_d. Le nombre de générateurs de $\mathbb{Z}/d\mathbb{Z}$ étant $\varphi(d)$ par la proposition 5.41, et H_d étant isomorphe à $\mathbb{Z}/d\mathbb{Z}$ nous avons

$$\text{Card}(H_d^*) = \varphi(d). \tag{19.108}$$

Par conséquent si H_d^* n'est pas vide, son cardinal est $\varphi(d)$. Nous avons

$$\begin{aligned} q-1 &= \text{Card}(\mathbb{K}^*) && \text{(19.109a)}\\ &= \text{Card}\Big(\bigcup_{d\,|\,q-1} H_d^*\Big) && \text{(19.109b)}\\ &= \sum_{d\,|\,q-1} \text{Card}(H_d^*) && \text{(19.109c)}\\ &\leqslant \sum_{d\,|\,q-1} \varphi(d) = q-1 && \text{(19.109d)} \end{aligned}$$

où nous avons utilisé la proposition 5.38. Par conséquent pour tout d divisant $q-1$ nous avons $\text{Card}(H_d^*) = \varphi(d)$ et il y a au moins un élément d'ordre $q-1$ dans $\mathbb{K}$. Cet élément engendre $\mathbb{K}^*$ parce que $\mathbb{K}^*$ contient exactement $q-1$ éléments. Par conséquent $\mathbb{K}^*$ est cyclique. □

CorpRUndR

Corolaire 19.32.
Si p est un nombre premier, alors

$$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}. \tag{19.110}$$

L'isomorphisme est un isomorphisme de groupes (abéliens). À gauche multiplicatif et à droite additif.

Démonstration. La proposition 19.31 nous enseigne que le groupe multiplicatif d'un corps fini est cyclique et donc isomorphe à un certain $\mathbb{Z}/n\mathbb{Z}$. Donc $(\mathbb{Z}/p\mathbb{Z})^*$ est un groupe cyclique d'ordre p, et donc isomorphe à $\mathbb{Z}/n\mathbb{Z}$ avec $n = p-1$. □

Lorsque $\mathbb{K}$ est un corps les éléments du groupe $\mathbb{K}^*$ sont les **éléments primitifs** de $\mathbb{K}$.

propQRcUlq

Proposition 19.33.
Soit $\mathbb{K}$ un corps contenant q éléments. Alors

(1) $x^q = x$ pour tout $x \in \mathbb{K}$,

(2) $X^q - X = \prod_{a\in\mathbb{K}}(X-a)$.

Démonstration. Le groupe $\mathbb{K}\backslash\{0\}$ ayant $q-1$ éléments, ses éléments vérifient $a^{q-1} = 1$ par le corolaire 2.14 et par conséquent $a^q = aa^{q-1} = a$.

Soit $a \in \mathbb{K}$. Étant donné que $a^q - a = 0$, le polynôme $(X-a)$ divise $X^q - X$ dans $\mathbb{K}[X]$. Par conséquent

$$\prod_{a\in\mathbb{K}}(X-a) \tag{19.111}$$

divise également $X^q - X$. Les polynômes $X^q - X$ et $\prod_{a\in\mathbb{K}}(X-a)$ étant deux polynômes unitaires de même degré, le fait que l'un divise l'autre montre qu'ils sont égaux. □

Une conséquence de $x^q = x$ est qu'il ne faut pas considérer le théorème 6.111 trop rapidement en disant « s'il s'annule partout, alors c'est le polynôme nul ». En effet dans un corps fini, « partout » n'est pas forcément très grand.

exVQBooBMPLkD

Exemple 19.34.
Si $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$ est le[16] corps à 3 éléments, alors le polynôme $P(X) = X^3 - X$ s'évalue à zéro pour tout $x \in \mathbb{F}_3$ (proposition 19.33.) mais il n'est pas le polynôme nul. △

16. Le singulier est justifié par le théorème 19.30, mais ça n'a pas d'importance ici.

19.4.3 Symboles de Legendre et carrés

Source : [573].

Nous disons que $a \in \mathbb{F}_p$ est un **carré** si il existe $b \in \mathbb{F}_p$ tel que $a = b^2$.

Définition 19.35.
*Soit $n \in \mathbb{N}$ et $p > 2$ un nombre premier. Le **symbole de Legendre** est défini par*

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{si } p \text{ divise } n \\ 1 & \text{si } n \text{ est un carré dans } \mathbb{F}_p \\ -1 & \text{sinon.} \end{cases} \tag{19.112}$$

Note : -1 peut être un carré, et pas que dans $\mathbb{C}$. Par exemple dans $\mathbb{F}_5$ nous avons $4 = -1$ et donc -1 est un carré.

PropcGsJjk

Proposition 19.36.
Soit un nombre premier $p > 2$. Le corps $\mathbb{F}_p^*$ contient autant de carrés que de non carrés. De plus pour tout $n \in \mathbb{N}$ nous avons

$$\left(\frac{n}{p}\right) = n^{(p-1)/2} \mod p. \tag{19.113}$$

Eqbcugos

Démonstration. Nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{F}_p^* &\to \mathbb{F}_p^* \\ x &\mapsto x^2. \end{aligned} \tag{19.114}$$

C'est un morphisme de groupes multiplicatifs et $\ker\psi = \{-1, 1\}$. Étant donné que $p > 2$, nous avons alors

$$\text{Card}(\ker\psi) = 2 \tag{19.115}$$

parce que $1 \neq -1$. Évidemment l'ensemble des carrés dans $\mathbb{F}_p^*$ est l'image de ψ. Le premier théorème d'isomorphisme 2.6(3) nous permet alors de conclure que

$$\text{Card}(\text{Image}(\psi)) = \frac{\text{Card}(\mathbb{F}_p^*)}{2}. \tag{19.116}$$

Ceci prouve la première assertion.

Par le petit théorème de Fermat (théorème 6.15), nous avons $x^{p-1} = 1$ pour tout $x \in \mathbb{F}_p^*$. Les $(p-1)$ éléments de $\mathbb{F}_p^*$ sont donc tous racines d'un des deux polynômes

$$X^{(p-1)/2} = \pm 1. \tag{19.117}$$

Mais chacun des deux ne peut avoir, au maximum, que $(p-1)/2$ solutions. Ils ont donc chacun exactement $(p-1)/2$ racines.

Nous pouvons maintenant prouver la formule (19.113). D'abord si $n = 0$, elle est évidente. Si n est un carré dans $\mathbb{F}_p$, nous posons $n = x^2$ et nous avons

$$n^{(p-1)/2} = x^{p-1} = 1 = \left(\frac{n}{p}\right). \tag{19.118}$$

Si n n'est pas un carré, c'est que n n'est pas une racine de $X^{(p-1)/2} = 1$. Le nombre n est alors une racine de $X^{(p-1)/2} = -1$. Nous avons alors

$$n^{(p-1)/2} = -1 = \left(\frac{n}{p}\right). \tag{19.119}$$

□

CoruJosNz

Corolaire 19.37.
Si $a, b \in \mathbb{N}$ et si $p > 2$ est un nombre premier, alors

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{19.120}$$

Démonstration. Par la formule (19.113),

$$\left(\frac{ab}{p}\right) = (ab)^{(p-1)/2} = a^{(p-1)/2}b^{(p-1)/2} = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{19.121}$$

□

Soit un nombre premier $q > 2$ et $\mathbb{A}$, un anneau de caractéristique p. Si $\alpha \in \mathbb{A}$ vérifie

$$1 + \alpha + \cdots + \alpha^{q-1} = 0, \tag{19.122}$$

nous définissons la **somme de Gauss** par

$$\tau = \sum_{x \in \mathbb{F}_q} \left(\frac{x}{q}\right)\alpha^x = \sum_{x=1}^{q-1} \left(\frac{x}{q}\right)\alpha^x. \tag{19.123}$$

Notons que la somme de Gauss dépend de q et du α choisis.

PropciRUov

Proposition 19.38.
Les sommes de Gauss vérifient les propriétés suivantes.

(1) $\tau^2 = \left(\frac{-1}{q}\right)q$. *Nous allons noter* $\epsilon(q) = \left(\frac{-1}{q}\right)$.

(2) Si $\mathbb{A}$ est de caractéristique $p \geqslant 3$ et si $p \neq q$ alors

$$\tau^p = \left(\frac{p}{q}\right)\tau. \tag{19.124}$$

EqxBNpJz

(3) Si $\mathbb{A}$ est de caractéristique p et si q est premier avec p, alors τ est inversible dans $\mathbb{A}$.

Démonstration. D'abord nous notons que

$$\alpha^q - 1 = (\alpha - 1)(1 + \alpha + \cdots + \alpha^{q-1}) = 0 \tag{19.125}$$

par définition de α. Nous calculons

$$\epsilon(q)\tau^2 = \epsilon(q)\sum_{x,y \in \mathbb{F}_q}\left(\frac{x}{q}\right)\left(\frac{y}{q}\right)\alpha^{x+y} \tag{19.126a}$$

$$= \sum_{x,y \in \mathbb{F}_q}\left(\frac{-xy}{q}\right)\alpha^{x+y}. \tag{19.126b}$$

EqlQbFeo

$$= \sum_{z \in \mathbb{F}_q}\sum_{y \in \mathbb{F}_q}\left(\frac{-(z-y)y}{q}\right)\alpha^z \tag{19.126c}$$

EqWxIhhk

$$= \sum_{z \in \mathbb{F}_q} s_z\alpha^z \tag{19.126d}$$

EqWoIszS

Justifications :

— Pour obtenir (19.126b) nous avons utilisé le corolaire 19.37.

— (19.126c) est un changement de variable $z = x + y$ dans la somme sur x.

— Pour (19.126d) nous avons posé

$$s_z = \sum_{y\in\mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right). \tag{19.127}$$

Nous avons

$$s_0 = \sum_{y\in\mathbb{F}_q} \left(\frac{y^2}{q}\right). \tag{19.128}$$

Dans cette somme, tous les termes sont égaux à 1, sauf celui avec $y = 0$ qui vaut zéro. Nous avons donc $s_0 = q - 1$. Voyons maintenant s_y avec $y \neq 0$. L'application

$$\begin{aligned} \mathbb{F}_q^* &\to \mathbb{F}_q\backslash\{1\} \\ k &\mapsto 1 - zy^{-1} \end{aligned} \tag{19.129}$$

étant une bijection nous pouvons effectuer le changement de variables $t = y^{-1}z - 1$ pour la somme sur y en notant y^{-1} l'inverse de y dans $\mathbb{F}_q^*$, nous trouvons alors

$$\sum_{y\in\mathbb{F}_q} \left(\frac{y(z-y)}{q}\right) = \sum_{y\in\mathbb{F}_q} \left(\frac{y^2(y^{-1}z-1)}{q}\right) \tag{19.130a}$$

$$= \sum_{y\in\mathbb{F}_q} \left(\frac{y^{-1}z-1}{q}\right) \tag{19.130b}$$

$$= \sum_{t\in\mathbb{F}_q\backslash\{1\}} \left(\frac{t}{q}\right) \tag{19.130c}$$

$$= \underbrace{\sum_{t\in\mathbb{F}_q} \left(\frac{y}{q}\right)}_{=0} - \left(\frac{1}{1}\right) \tag{19.130d}$$

$$= -1 \tag{19.130e}$$

parce qu'il y a autant de carrés que de non carrés dans $\mathbb{F}_q^*$ (proposition 19.36). En résumé nous avons

$$\epsilon(q)\tau^2 = \sum_{z\in\mathbb{F}_q} s_z\alpha^z \tag{19.131}$$

où

$$s_z = \begin{cases} q-1 & \text{si } z = 0 \\ -1 & \text{sinon.} \end{cases} \tag{19.132}$$

Cela donne

$$\epsilon(q)\tau^2 = (q-1) - \underbrace{(\alpha + \cdots + \alpha^{q-1})}_{=-1} = q \tag{19.133}$$

où nous avons utilisé l'hypothèse sur α. Donc $\epsilon(q)\tau^2 = q$, et étant donné que $\epsilon(q) = \pm 1$ nous concluons

$$\tau^2 = \epsilon(q)q. \tag{19.134}$$

Nous prouvons maintenant la seconde partie. Comme $\mathbb{A}$ est de caractéristique p, en utilisant le fait que le morphisme de Frobenius est un morphisme,

$$\tau^p = \left(\sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)\alpha^x\right)^p = \sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)^p \alpha^{px}. \tag{19.135}$$

Étant donné que $\left(\frac{x}{q}\right) = \pm 1$ et que p est impair, nous avons

$$\left(\frac{x}{q}\right)^p = \left(\frac{x}{q}\right). \tag{19.136}$$

Du coup nous avons

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{xp}{q}\right)\alpha^{px}. \tag{19.137}$$

Mais p étant inversible dans $\mathbb{F}_q$, l'application $x \mapsto px$ est une bijection et nous pouvons sommer sur px au lieu de x :

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{x}{q}\right)\alpha^{x} = \tau. \tag{19.138}$$

Nous trouvons alors que

$$\tau^p = \left(\frac{p}{q}\right)\tau. \tag{19.139}$$

Étant donné la formule du τ^2 que nous venons de démontrer, nous avons $\tau^2 = \pm q$. Les nombres p et q étant premiers entre eux, le théorème de Bézout (théorème 1.236) nous donne a et b tels que

$$ap + bq = 1. \tag{19.140}$$

Cela montre que b est un inverse de q modulo p. Donc τ^2 est inversible, et il en découle que τ lui-même est inversible. □

ThoMiEiUm

Théorème 19.39 (Loi de réciprocité quadratique)**.**
Soient deux nombres premiers distincts $p, q \geqslant 3$. Alors

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}\left(\frac{q}{p}\right). \tag{19.141}$$

Démonstration. Soit ϕ_q le polynôme $1 + X + \cdots + X^{q-1}$ et l'anneau

$$\mathbb{A} = \mathbb{F}_p[X]/(\phi_q). \tag{19.142}$$

C'est un anneau de caractéristique p parce que son unité est le polynôme constant 1. Nous nommons $\alpha = X/(\phi_q)$, c'est-à-dire que $\phi_q(\alpha) = 0$ dans $\mathbb{A}$, et nous pouvons considérer la somme de Gauss

$$\tau = \sum_{i\in\mathbb{F}_q}\left(\frac{i}{q}\right)\alpha^i. \tag{19.143}$$

Notons que ceci est un élément de $\mathbb{A}$ et plus précisément un polynôme de degré zéro dans $\mathbb{A}$, et encore plus précisément, une classe d'un tel polyôme. Donc les coefficients de α doivent être compris comme des éléments de $\mathbb{F}_p$. Nous savons (proposition 19.38) que

$$\tau^2 = \left(\frac{-1}{q}\right), \tag{19.144}$$

et en utilisant la formule (19.113) nous trouvons

$$\left(\frac{\tau^2}{p}\right) = (\tau^2)^{(p-1)/2} \mod p = \tau^{p-1} \mod p \tag{19.145}$$

En réalité sur cette dernière ligne, nous ne devrions pas préciser le « modulo p » parce que, comme mentionné plus haut, ce sont des éléments de $\mathbb{F}_p$. En utilisant cela, ainsi que (19.124), nous avons

$$\underbrace{\left(\frac{\tau^2}{p}\right)}_{\tau^{p-1}}\tau = \tau^p = \left(\frac{p}{q}\right)\tau \tag{19.146}$$

Puisque τ est inversible, nous écrivons

$$\left(\frac{\tau^2}{p}\right) = \left(\frac{p}{q}\right) \tag{19.147}$$

Nous utilisons maintenant la formule (19.113) sur le membre de gauche avec $n = \tau^2 = \left(\frac{-1}{q}\right)$:

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{q}\right)^{\frac{q-1}{2}} \left(\frac{q}{p}\right). \tag{19.148}$$

Toujours avec la même formule nous pouvons substituer $\left(\frac{-1}{q}\right)$ par $(-1)^{(q-1)/2}$ et obtenir

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(q-1)}{2}\frac{(p-1)}{2}}. \tag{19.149}$$

□

Lemoabzrn

Lemme 19.40.
Si p est un nombre premier $p \geqslant 3$, alors le symbole de Legendre $x \mapsto \left(\frac{x}{p}\right)$ est l'unique morphisme non trivial de $\mathbb{F}_p^*$ dans $\{-1, 1\}$.

Démonstration. Le fait que le symbole de Legendre soit non trivial est simplement le fait qu'il y ait des carrés et des non carrés dans $\mathbb{F}_p^*$; voir la proposition 19.36. Pour l'unicité, soit $\alpha\colon \mathbb{F}_p^* \to \{-1, 1\}$ un morphisme surjectif (c'est-à-dire non trivial). Étant donné que

$$\mathbb{F}_p^* = \ker(\alpha) \cup -\ker(\alpha), \tag{19.150}$$

le groupe $\mathbb{F}_p^*/\ker(\alpha)$ ne contient que deux éléments : $[1]$ et $[-1]$. Autrement dit, $\ker(\alpha)$ est d'indice 2 dans $\mathbb{F}_p^*$.

Or $\mathbb{F}_p\backslash\{0\}$ ne possède qu'un seul sous-groupe d'indice 2. En effet soit S un tel sous-groupe et a, un générateur de $\mathbb{F}_p\backslash\{0\}$ (qui est cyclique par la proposition 19.31), alors $a^2 \in S$ par le lemme 3.29. Par conséquent S contient le groupe des puissances paires de a. Le groupe S ne peut rien contenir de plus parce qu'il est d'indice 2 et que l'ordre de $\mathbb{F}_p\backslash\{0\}$ est pair.

Bref, le sous-groupe $\ker(\alpha)$ est l'unique sous-groupe d'indice 2 dans $\mathbb{F}_p\backslash\{0\}$. Mais la proposition 19.36 nous indique que $|(\mathbb{F}_p\backslash\{0\})^2| = \frac{p-1}{2}$, c'est-à-dire que le groupe des carrés est d'indice 2. Nous avons donc, par l'unicité,

$$\ker(\alpha) = (\mathbb{F}_p\backslash\{0\})^2. \tag{19.151}$$

Au final, pour $y \in \mathbb{F}_p\backslash\{0\}$,

$$\alpha(y) = \begin{cases} 1 & \text{si } y \text{ est un carré} \\ -1 & \text{sinon.} \end{cases} \tag{19.152}$$

Ce qui est bien la définition des symboles de Legendre. □

PROPooCLEUooHWnoXM

Proposition 19.41.
Si p est un nombre premier, nous avons

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \textit{si } p \in [1]_8 \textit{ ou } p \in [7]_8 \\ -1 & \textit{sinon.} \end{cases} \tag{19.153}$$

Démonstration. Soit le polynôme

$$X^4 + 1 \in \mathbb{F}_p[X] \tag{19.154}$$

et α, une racine dans une extension [17] de $\mathbb{F}_p$ [18]. Nous posons $\theta = \alpha + \alpha^{-1}$ et nous calculons

$$\theta^2 = (\alpha + \alpha^{-1})(\alpha + \alpha^{-1}) = \alpha^2 + 2 + (\alpha^2)^{-1} = \alpha^2 + 2 - \alpha^2 = 2 \tag{19.155}$$

parce que α^4 étant -1, nous avons $(\alpha^2)^{-1} = -\alpha^2$. Bref, $\theta^2 = 2$.

17. Dans la source que je suivais (je ne sais plus où), on parlait ici de « fermeture » de $\mathbb{F}_p$ et non d'extension. Il me semble que parler simplement d'extension suffit. Vous confirmez ?

18. Voir par exemple la proposition 6.143 pour l'existence d'une extension comme il faut.

Dire que 2 est un carré modulo p revient à dire que θ est dans $\mathbb{F}_p$. C'est-à-dire que pour calculer le symbole de Legendre $\left(\frac{2}{p}\right)$, nous étudions pour quels p, l'élément θ est vraiment dans $\mathbb{F}_p$ et non seulement dans l'extension $\mathbb{F}_p(\alpha)$. En tenant compte de l'exemple 6.97, il faut distinguer deux cas : $\alpha^p = \alpha$ et $\alpha^p \neq \alpha$. Autrement dit, si $\alpha^k = \alpha$ pour un certain nombre premier k, alors le cas $p = k$ est à traiter à part. La liste des puissances de α est :

$$1, \alpha, \alpha^2, \alpha^3, -1, -\alpha, -\alpha^2, -\alpha^3, 1, \alpha, \dots \tag{19.156}$$

Nous avons donc automatiquement $\alpha^{9k} = \alpha$, mais $p = 9k$ est exclu parce que p est premier. Nous devons donc vérifier si une des propriétés

$$\alpha^2 = \alpha \tag{19.157a}$$
$$\alpha^3 = \alpha \tag{19.157b}$$
$$-\alpha = \alpha \tag{19.157c}$$
$$-\alpha^3 = \alpha \tag{19.157d}$$

est possible. Il est aisément vérifiable, au cas par cas, que ces possibilités sont toutes incompatibles avec $\alpha^4 = -1$. Nous avons donc certainement $\alpha^p \neq \alpha$ et compte tenu de l'exemple 6.97, l'équation $x^p = x$ caractérise les éléments de $\mathbb{F}_p$ dans $\mathbb{F}_p(\alpha)$.

L'équation $X^2 = 2$ a exactement deux solutions qui sont $\pm\theta$. Nous avons donc $2 \in \mathbb{F}_p^2$ si et seulement si $\theta \in \mathbb{F}_p$ si et seulement si $\theta^p = \theta$. Nous avons réduit notre problème à déterminer pour quels p nous avons $\theta^p = \theta$. D'abord nous avons, par le morphisme de Frobenius,

$$\theta^p = (\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p}. \tag{19.158}$$

Nous pouvons maintenant conclure facilement. Un nombre premier étant impair (sauf $p = 2$ qui peut être traité à part), p est automatiquement dans un des ensembles $[1]_8$, $[3]_8$, $[5]_8$ ou $[7]_8$. Nous avons quatre petites vérifications à faire. Dans tous les cas $\alpha^{8k} = 1$. Si $p = 1 + 8k$, alors

$$\theta^p = \alpha^{1+8k} + (\alpha^{-1})^{1+8k} = \alpha + \alpha^{-1} = \theta, \tag{19.159}$$

donc 2 est un carré dans $\mathbb{F}_p$. Si $p \in [3]_8$, alors $\theta^p = \alpha^3 + \alpha^{-3}$. Si cela était égal à $\alpha + \alpha^{-1}$, alors nous aurions

$$\alpha^6 + 1 = \alpha^4 + \alpha^2, \tag{19.160}$$

et donc $\alpha^2 = 1$, ce qui est impossible. Les vérifications pour $p \in [5]_8$ et $p \in [7]_8$ sont du même style. □

19.4.4 Théorème de Chevalley-Warning

Lemme 19.42.
Soit $\mathbb{K}$ un corps de caractéristique p et de cardinal q. Pour $m \in \mathbb{N}$ nous définissons

$$S_m = \sum_{x \in \mathbb{K}} x^m. \tag{19.161}$$

Alors nous avons

$$S_m \mod p = \begin{cases} -1 & \text{si } m \geqslant 1 \text{ et } m \text{ divisible par } q-1 \\ 0 & \text{sinon.} \end{cases} \tag{19.162}$$

Démonstration. Si $m = 0$, alors $x^0 = 1$ et $S_m = q$. Par conséquent $S_m \mod p = 0$ parce que la caractéristique d'un corps divise son ordre (proposition 1.353).

Nous prenons maintenant $m \geqslant 1$ et nous voyons séparément les cas où $q - 1$ divise m ou non. Si $q - 1$ divise m, alors pour tout $x \neq 0$ nous avons

$$x^m = x^{k(q-1)} = 1 \tag{19.163}$$

parce que $\mathbb{K}\backslash\{0\}$ est cyclique et $x^{q-1}=1$ par le petit théorème de Fermat (théorème 6.15). Par conséquent nous avons

$$\sum_{x\in\mathbb{K}} x^m = \sum_{x\in\mathbb{K}\backslash\{0\}} 1 = q-1. \tag{19.164}$$

Si le nombre $m\geqslant 1$ n'est pas divisible par $q-1$ alors nous prenons un générateur y du groupe $\mathbb{K}\backslash\{0\}$. Un tel élément vérifie $y^m\neq 1$. En effet, si y vérifiait $y^m=1$ alors cela signifierait que l'ordre de $\mathbb{K}\backslash\{0\}$ est un diviseur de m, ce qui n'est pas le cas ici, parce que l'ordre de $\mathbb{K}\backslash\{0\}$ est $q-1$. Pour un tel y, l'application

$$\begin{aligned}\varphi\colon \mathbb{K}\backslash\{0\} &\to \mathbb{K}\backslash\{0\}\\ x &\mapsto yx\end{aligned} \tag{19.165}$$

est une bijection [19]. En ce qui concerne l'injectivité, $ya=yb$ implique $a=b$. En ce qui concerne la surjectivité, si a est un générateur, si $z=a^l$ et si $y=a^k$, alors

$$z=\varphi(a^{l-k}). \tag{19.166}$$

Nous pouvons maintenant poser le calcul.

$$S_m=\sum_{x\in\mathbb{K}\backslash\{0\}} x^n=\sum_{x\in\mathbb{K}\backslash\{0\}}(yx)^m=y^m\sum_{x\in\mathbb{K}\backslash\{0\}}x^m=y^mS_m. \tag{19.167}$$

Étant donné que $y^m\neq 1$, la seule solution est $S_m=0$. □

ThoLTcYKk

Théorème 19.43 (Chevalley-Warning[574])**.**
Soit $\mathbb{K}$ un corps fini de cardinal q et de caractéristique p. Soient $P_1,\dots,P_r$ des éléments de $\mathbb{K}[X_1,\dots,X_n]$ tels que $\sum_{i=1}^r \deg(P_i)<n$. Nous considérons l'ensemble des zéros communs à tous les polynômes :

$$V=\{x\in\mathbb{K}^n \text{ tel que } P_1(x)=\dots=P_r(x)=0\}. \tag{19.168}$$

Alors $\operatorname{Card}(V)=0 \mod p$.

Démonstration. Nous considérons le polynôme

$$P=\prod_{i=1}^r(1-P_i^{q-1}). \tag{19.169}$$

Montrons que

$$P(x)=\begin{cases}1 & \text{si } x\in V\\ 0 & \text{sinon.}\end{cases} \tag{19.170}$$

La première ligne est facile : étant donné que tous les $P_i(x)$ sont nuls pour $x\in V$, nous avons $P(x)=1$. Si x n'est pas dans V, alors nous avons un i tel que $P_i(x)\in\mathbb{K}\backslash\{0\}$. Mais dans ce cas (toujours la cyclicité de $\mathbb{K}\backslash\{0\}$) nous avons $P_i(x)^{q-1}=1$ et donc le produit est nul.

En utilisant l'hypothèse sur le degré des P_i, nous trouvons

$$\deg(P)=\sum_{i=1}^r(q-1)\deg(P_i)<n(q-1). \tag{19.171}$$

Pour un polynôme $Q\in\mathbb{K}[X_1,\dots,X_n]$, nous définissons

$$\int Q=\sum_{x\in\mathbb{K}^n}Q(x). \tag{19.172}$$

19. Notons que nous n'avons pas réellement besoin que y soit un générateur. Nous n'utilisons seulement le fait que $y^m\neq 1$ et $y\neq 0$.

Nous avons immédiatement

$$\int P = \sum_{x\in\mathbb{K}^n} P(x) = \sum_{x\in V} 1 = \text{Card}(V) \mod p. \tag{19.173}$$

Nous insistons sur le « modulo p » parce que dans la formule $P(x) = 1$, le membre de droite est le 1 de $\mathbb{K}$; il est donc automatiquement modulo la caractéristique de $\mathbb{K}$.

Il nous reste à prouver que $\int P = 0$. Pour cela nous décomposons

EqHnUVlM
$$P = \sum_m c_m X_1^{m_1} \dots X_n^{m_n} \tag{19.174}$$

où la somme s'étend sur les $m \in \mathbb{N}^n$ tels que $c_m \neq 0$. Nous avons

$$\begin{aligned}\int P &= \sum_{x\in\mathbb{K}^n} \sum_m c_m x_1^{m_1} \dots x_n^{m_n} && (19.175a)\\ &= \sum_m c_m \left(\sum_{x\in\mathbb{K}^n} x_1^{m_1} \dots x_n^{m_n} \right) && (19.175b)\\ &= \sum_m c_m S_{m_1} \dots S_{m_n}. && (19.175c)\end{aligned}$$

Le terme de plus haut degré dans la décomposition (19.174) est celui du m tel que $\sum_i m_i$ est le plus grand. Comme ce degré est plus petit que $n(q-1)$, pour chacun des m rentrant dans la somme, nous avons

$$\sum_{i=1}^{n} m_i < n(q-1). \tag{19.176}$$

En particulier pour tout $m \in \mathbb{N}^n$, il existe i tel que $m_i < q-1$, et dans ce cas $S_{m_i} = 0$. Donc tous les termes de la somme

$$\sum_{m\in\mathbb{N}^n} c_m S_{m_1} \dots S_{m_n} \tag{19.177}$$

ont un facteur nul. □

CorfuHNKz

Corolaire 19.44.
Soit P_i des polynômes à n variables avec $\sum_{i=1}^{r} \deg(P_i) < n$. Si les P_i n'ont pas de terme constant, alors ils ont un zéro commun non trivial.

Démonstration. Nous reprenons les notations du théorème 19.43. Étant donné que les P_i n'ont pas de terme constant, $0 \in V$, mais $\text{Card}(V) = 0 \mod p$. Par conséquent nous devons avoir $\text{Card}(V) > p$. □

Exemple 19.45.
Nous considérons les polynômes

$$\begin{aligned}P_1(x,y,t,u) &= xy + x + ux && (19.178a)\\ P_2(x,y,t,u) &= x + y - 3t. && (19.178b)\end{aligned}$$

La somme de leurs degrés est 3 et ce sont des polynômes à 4 variables. Nous devons donc avoir, en vertu du corolaire 19.44, d'autres racines que la racine triviale $(x,y,t,u) = (0,0,0,0)$.

Le corolaire nous donne aussi une borne inférieure du nombre de racines à chercher : plus que la caractéristique du corps sur lequel nous travaillons. Nous pouvons dire cela sans avoir la moindre idée de la façon dont on pourrait résoudre le système $P_1 = P_2 = 0$. △

19.4.5 Contenu d'un polynôme

LemHULrVaF

Lemme 19.46 (de Gauss[104, 575]).
Soient $P, Q \in \mathbb{Z}[X]$. Alors

$$c(PQ) = c(P)c(Q) \tag{19.179}$$

où c est le contenu, définition 3.113.

Démonstration. Afin de fixer les notations, nous posons $P = \sum_i a_i X^i$ et $Q = \sum_j b_j Y^j$.

(1) **Pour les polynômes primitifs**
Nous commençons par supposer que $c(P) = c(Q) = 1$. Dans ce cas si $c(PQ) \neq 1$, nous considérons un nombre premier p divisant $c(PQ)$. Puisque le contenu de P et de Q vaut 1, le nombre p ne peut pas diviser tous leurs coefficients. Nous définissons i_0 de façon que a_{i_0} soit le premier à ne pas être divisible par p, et j_0 de telle façon que b_{j_0} soit le premier à ne pas être divisible par p. Autrement dit :

$$p \mid a_0, p \mid a_1, \ldots, p \mid a_{i_0-1}, p \nmid a_{i_0} \tag{19.180}$$

et de façon similaire pour j_0. Donc p ne divise ni a_{i_0}, ni b_{j_0}. Nous nous demandons alors avec malice quel est le coefficient de $X^{i_0+j_0}$ dans PQ. La réponse est :

$$a_{i_0}b_{j_0} + \sum_{\substack{i+j=i_0+j_0 \\ i<i_0 \text{ ou } j<j_0}} a_i b_j. \tag{19.181}$$

Par définition p divise soit a_i soit b_j pour chacun des termes de la grande somme. Comme p ne divise pas $a_{i_0}b_{j_0}$, il ne divise pas le coefficient de $X^{i_0+j_0}$ dans PQ, alors que nous étions partis en disant que p divisait tous les coefficients de PQ.
Nous concluons donc que $c(PQ) = 1$.

(2) **Cas général**
Si P et Q sont maintenant des polynômes sans condition particulière dans $\mathbb{Z}[X]$, nous considérons $P_1 = \frac{P}{c(P)}$ et $Q_1 = \frac{Q}{c(Q)}$; ces deux polynômes sont primitifs et nous avons alors, en utilisant la première partie :

$$c(P_1Q_1) = 1. \tag{19.182}$$

Étant donné que

$$P_1Q_1 = \frac{1}{c(P)c(Q)}PQ, \tag{19.183}$$

nous avons

$$c(PQ) = c(P)c(Q)c(P_1Q_1) = c(P)c(Q). \tag{19.184}$$

□

19.4.6 Théorème de l'élément primitif

Définition 19.47.
*Soit $\mathbb{K}$ un corps. Une extension $\mathbb{L}$ de $\mathbb{K}$ est dite **finie** si $\mathbb{L}$ est un espace vectoriel de dimension finie sur $\mathbb{K}$.*

Notez que la définition d'extension finie ne suppose ni que $\mathbb{K}$, ni que $\mathbb{L}$, soient finis en tant qu'ensembles.

THOooYFLZooHYENKp

Théorème 19.48 (de l'élément primitif). *Si $\mathbb{K}$ est un corps fini, toute extension finie de $\mathbb{K}$ est simple*[20].

Si $\mathbb{K}$ est un corps quelconque alors toute extension séparable finie est simple.

20. Définition 6.93.

Démonstration. Nous ne donnons la preuve que dans le cas où $\mathbb{K}$ est fini. Dans ce cas nous savons par la proposition 19.31 que le groupe $\mathbb{K}\backslash\{0\}$ est cyclique. Si de plus $\mathbb{L}$ est une extension finie alors $\mathbb{L}$ est fini en tant qu'ensemble. Par conséquent $\mathbb{L}\backslash\{0\}$ est un groupe cyclique. Si α est un générateur de $\mathbb{L}$ alors $\mathbb{L} = \mathbb{K}(\alpha)$ et l'extension est donc simple.

Une preuve de l'assertion dans le cas où $\mathbb{K}$ est infini peut être trouvée sur wikipédia. □

PROPooWNHYooSzMJqF

Proposition 19.49.
L'ordre d'un polynôme P vérifie les propriétés suivantes :

(1) L'ordre de P est l'ordre multiplicatif de ses racines

(2) L'ordre de P divise $p^n - 1$.

LemZrUUOz

Lemme 19.50.
Soit p un nombre premier et P un polynôme irréductible unitaire de degré n. Si $\alpha, \beta \in \mathbb{F}_p[X]/P$, alors $(\alpha + \beta)^p = \alpha^p + \beta^p$.

Démonstration. La preuve est exactement la preuve classique[21] :

$$(\alpha + \beta)^p = \sum_k \binom{p}{k} \alpha^k \beta^{p-k} \tag{19.185}$$

où les coefficients binomiaux sont dans $\mathbb{F}_p$ et donc nuls pour les k différents de p et de 0. □

Cette proposition est encore vraie avec $\alpha, \beta \in \mathbb{F}_{p^n}$ et $(\alpha + \beta)^{p^n}$.

Lemme 19.51.
Si $\alpha \in \mathbb{F}_q$ est une racine d'ordre k de P (de degré n) alors les racines de $X^k - 1$ sont $\{\alpha^i$ tel que $i = 0, \dots, k-1\}$.

Nous serions donc intéressés à construire $\mathbb{F}_q$ comme quotient de $\mathbb{F}_p[X]$ par un polynôme primitif. Le théorème suivant donne une description abstraite de $\mathbb{F}_q$ qui va nous servir de point de départ pour la construction.

ThoqSludu

Théorème 19.52 (Théorème de l'élément primitif)**.**
Soit p un nombre premier, $n \in \mathbb{N}$ et $q = p^n$. Soit $\mathbb{K}$ un corps à q éléments. Alors

(1) Il existe $\alpha \in \mathbb{K}$ tel que $\mathbb{K} = \mathbb{F}_p[\alpha]$.

(2) Il existe un polynôme irréductible $P \in \mathbb{F}_p[X]$ de degré n tel que

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K} \\ \overline{X} &\mapsto \alpha \end{aligned} \tag{19.186}$$

EqWlMhhm

soit un isomorphisme de corps.

Soient α et P choisis pour avoir les propriétés citées plus haut. Alors nous avons les propriétés suivantes.

(1) P est primitif[22]*.*

(2) P est scindé dans $\mathbb{K}$.

(3) L'ensemble des racines de P est $\{\alpha, \alpha^p, \dots, \alpha^{p^{n-1}}\}$.

(4) Le polynôme P divise $X^q - X$ dans $\mathbb{F}_p[X]$.

Démonstration. Le corps $\mathbb{K}$ étant fini, il est cyclique par la proposition 19.31. Soit α un générateur de $\mathbb{K}\backslash\{0\}$ alors

$$\mathbb{K} = \mathbb{F}_p[\alpha]. \tag{19.187}$$

21. Voir le coefficients binomiaux de la proposition 3.43.
22. Définition 3.115.

Soit ℓ le plus grand entier tel que l'ensemble

$$\{1, \alpha, \cdots, \alpha^{\ell-1}\} \subset \mathbb{K} \tag{19.188}$$

soit libre. Pour rappel, $\mathbb{K}$ est un espace vectoriel sur $\mathbb{F}_p$. Il existe des $a_i \in \mathbb{F}_p$ tels que

$$\alpha^\ell + a_{\ell-1}\alpha^{\ell-1} + \cdots + a_0 = 0. \tag{19.189}$$

De façon équivalente, il existe un polynôme unitaire $P \in \mathbb{F}_p[X]$ de degré ℓ tel que $P(\alpha) = 0$. Étant donné que α est générateur de $\mathbb{K}$,

$$\mathbb{K} = \text{Span}\{1, \alpha, \ldots, \alpha^{\ell-1}\} \tag{19.190}$$

parce que $\mathbb{K}$ est généré par les puissances de α alors que les puissances de α plus hautes que $\ell - 1$ peuvent être générées par $1, \alpha, \ldots, \alpha^{\ell-1}$. L'espace $\mathbb{K}$ est donc un $\mathbb{F}_p$-espace vectoriel de dimension ℓ ; par conséquent

$$\text{Card}(\mathbb{K}) = p^n = q \tag{19.191}$$

et $\ell = n$.

Montrons que P est irréductible dans $\mathbb{F}_p$. Si P était réductible dans $\mathbb{F}_p$, l'élément $\alpha \in \mathbb{K}$ serait une racine d'un des facteurs, c'est-à-dire qu'il serait racine d'un polynôme de degré inférieur à n, ce qui contredirait le fait que

$$\{\alpha^{\ell-1}, \ldots, 1\} \tag{19.192}$$

soit libre.

Montrons que l'application

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K} \\ \overline{X} &\mapsto \alpha \end{aligned} \tag{19.193}$$

est un isomorphisme. Pour l'injectivité, deux éléments $Q_1, Q_2 \in \mathbb{F}_p[X]/(P)$ s'écrivent

$$Q_1 = \sum_{k=0}^{n-1} a_k \overline{X}^k \tag{19.194a}$$

$$Q_2 = \sum_{k=0}^{n-1} b_k \overline{X}^k. \tag{19.194b}$$

Dans ce cas si $\phi(Q_1) = \phi(Q_2)$ alors

$$\phi(Q_1) = \sum_{k=0}^{n-1} a_k \alpha^k = \phi(Q_2) = \sum_{k=0}^{n-1} b_k \alpha^k. \tag{19.195}$$

Mais l'ensemble $\{1, \alpha, \ldots, \alpha^{n-1}\}$ étant libre sur $\mathbb{F}_p$, cela implique $a_k = b_k$. La surjectivité de ϕ provient du fait que α génère $\mathbb{K}$.

Nous passons maintenant à la seconde partie de la démonstration. Soient $\alpha \in \mathbb{K}$ tel que $\mathbb{K} = \mathbb{F}_p[\alpha]$ et $P \in \mathbb{F}_p[X]$ un polynôme irréductible de degré n tel que $\alpha \mapsto \overline{X}$ soit un isomorphisme entre $\mathbb{K}$ et $\mathbb{F}_p[X]/(P)$.

Le polynôme P est primitif parce que α est d'ordre p^n dans $\mathbb{K}$ alors que $\overline{X} \mapsto \alpha$ est un isomorphisme. Par conséquent $\overline{X}$ est d'ordre p^n dans $\mathbb{F}_p[X]/P$.

Nous commençons par prouver que l'ensemble

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{19.196}$$

est l'ensemble des racines distinctes de P. Pour cela nous posons

$$P(X) = \sum_{k=0}^{n} a_k X^k \tag{19.197}$$

avec $a_k \in \mathbb{F}_p$. D'abord α est une racine de P. En effet

$$P(\overline{X}) = \sum_k a_k \overline{X}^k = 0 \tag{19.198}$$

parce que cette somme est calculée dans $\mathbb{F}_p[X]/(P)$. En appliquant l'isomorphisme ϕ à l'égalité (19.198) nous trouvons

$$0 = \phi\big(P(\overline{X})\big) = \sum_k a_k \phi(\overline{X}^k) = \sum_k a_k \alpha^k. \tag{19.199}$$

Donc α est bien une racine de P dans $\mathbb{F}_p[X]$. Nous devons montrer qu'il en est de même pour les autres puissances dans l'ensemble (19.196). Étant donné que pour tout x dans $\mathbb{F}_p$ nous avons $x^p = x$, nous avons aussi

$$P(X^p) = \sum_k a_k (X^p)^k = \sum_k a_k^p (X^p)^k = \sum_k (a_k X^k)^p \tag{19.200}$$

alors que nous savons que $x \mapsto x^p$ est un automorphisme de $\mathbb{F}_p$ par la proposition 1.355. Par conséquent

$$P(X^p) = \sum_k (a_k X^k)^p = \left(\sum_k a_k X^k \right)^p = P(X)^p. \tag{19.201}$$

Nous avons montré que si β est une racine de P, alors β^p est également une racine de P. Nous savons déjà que α est une racine de P, et que α est également générateur de $\mathbb{K}$, c'est-à-dire que α est d'ordre $q-1$. Les puissances

$$\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}} \tag{19.202}$$

sont donc distinctes ($\alpha^{p^n} = \alpha^q = 1$) et sont toutes des racines de P. Étant donné que P est de degré n il ne peut pas y avoir d'autres racines. Nous concluons que l'ensemble

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{19.203}$$

est l'ensemble des racines distinctes de P dans $\mathbb{K}$. Le polynôme P est alors scindé dans $\mathbb{K}[X]$.

Le dernier point du théorème est de montrer que P divise $X^q - X$. Pour cela nous allons montrer que toutes les racines de P sont des racines de $X^q - X$. Soit β une racine de P ; il s'écrit $\beta = \alpha^k$ pour un certain k. Étant donné que $\alpha^{q-1} = e = \alpha^{p^n - 1}$,

$$\begin{aligned}
\beta^q &= (\alpha^{p^n})^k && \text{(19.204a)}\\
&= \big(\alpha^{p^n - 1}\alpha\big)^k && \text{(19.204b)}\\
&= \big(\alpha^{q-1}\alpha\big)^k && \text{(19.204c)}\\
&= \alpha^k && \text{(19.204d)}\\
&= \beta. && \text{(19.204e)}
\end{aligned}$$

Cela signifie que $\beta^q = \beta$ et donc que β est racine de $X^q - X$. □

Corolaire 19.53.
Le corps fini à $q = p^n$ éléments est de caractéristique p.

Démonstration. Nous considérons le corps fini $\mathbb{K}$ à q éléments sous la forme $\mathbb{K} = \mathbb{F}_p[X]/P$ comme indiqué par l'équation (19.186). Soit 1_q la classe du polynôme 1 modulo P, nous considérons le morphisme

$$\begin{aligned}
\mu\colon \mathbb{Z} &\to \mathbb{F}_q\\
n &\mapsto n1_q.
\end{aligned} \tag{19.205}$$

Le noyau de cette application est $\ker\mu = \mathbb{Z}_p$ parce que $p1_q = 0$, les coefficients étant à comprendre dans $\mathbb{F}_p$. □

DefnPNCFO

Définition 19.54.
Soient P, un polynôme de degré n, et p, un nombre premier. Un élément $\alpha \in \mathbb{F}_p[X]/(P)$ est une ***racine primitive*** *si les puissances de α parcourent tout le groupe multiplicatif $(\mathbb{F}_p[X]/P)\backslash\{0\}$.*

Lembcerei

Lemme 19.55.
Soit p un nombre premier et P, un polynôme de degré n. Si $\alpha \in \mathbb{F}_p[X]/P$ est une racine primitive de P alors les autres racines de P sont également primitives.

Démonstration. Soit $\alpha \in \mathbb{F}_p[X]/P$ une racine primitive de P. L'élément α^p est également une racine parce que si $P = \sum_k a_k X^k$,

$$P(\alpha^p) = \sum_k (a_k \alpha^k)^p = \big(\sum_k a_k \alpha^k\big)^p = 0 \tag{19.206}$$

où nous avons utilisé le fait que $a_k^p = a_k$ étant donné que $a_k \in \mathbb{F}_p$. Par hypothèse α est une racine primitive ; cela implique que les éléments $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^n-1}$ sont distincts dans $\mathbb{F}_p[X]/P$. Ces éléments constituent donc *toutes* les racines de P.

Soit $\beta = \alpha^{p^k}$ une racine de P. Montrons que α est une puissance de β. Étant donné que $(\mathbb{F}_p[X]/P)\backslash\{0\}$ est un groupe à p^n-1 éléments, le corolaire 2.14 indique que $\alpha^{p^n} = \alpha$. En particulier avec $r = p^{n-k}$ nous avons

$$\beta^r = \alpha^{rp^k} = \alpha^{p^n} = \alpha. \tag{19.207}$$

Par suite toutes les puissances de α sont des puissances de β, ce qui implique que β est générateur du groupe cyclique $(\mathbb{F}_p[X]/P)\backslash\{0\}$. □

LemkzWjse

Lemme 19.56.
Soit p un nombre premier et n, un entier. Un polynôme de degré d, irréductible dans $\mathbb{F}_p[X]$, divise $X^{p^n} - X$ si et seulement si, d divise n.

Théorème 19.57.
Soient P et Q deux polynômes irréductibles de degré n dans $\mathbb{F}_p[X]$. Alors les quotients $\mathbb{F}_p[X]/P$ et $\mathbb{F}_p[X]/Q$ sont isomorphes en tant que corps.

En guise de démonstration de ce théorème, nous allons démontrer la proposition suivante.

PropCRPjZsp

Proposition 19.58.
Si $\mathbb{K}$ et $\mathbb{L}$ sont deux corps à $q = p^n$ éléments, alors ils sont isomorphes.

Démonstration. Soit a un élément primitif de $\mathbb{K}$ et P son polynôme minimal. Nous savons que $\mathbb{K} \simeq \mathbb{F}_p[X]/P$ par le théorème de l'élément primitif 19.52. L'élément a est en particulier une racine de $X^q - X$. Par ailleurs P divise $X^q - X$ par le lemme 19.56.

Nous avons aussi

$$X^q - X = \prod_{b \in \mathbb{L}} (X - b) \tag{19.208}$$

par la proposition 19.33. Étant donné que P divise $X^q - X$, un des éléments de $\mathbb{L}$ annule P. Soit $b \in \mathbb{L}$ tel que $P(b) = 0$. Soit Q le polynôme minimal de b. Par définition nous savons que Q divise P, mais P étant irréductible et unitaire, nous avons immédiatement $P = Q$. En particulier

$$\mathbb{F}_p[X]/P \simeq \mathbb{F}_p[X]/Q \simeq \mathbb{K}. \tag{19.209}$$

Nous montrons maintenant que $\mathbb{F}_p[X]/Q \simeq \mathbb{L}$ par l'application

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/Q &\to \mathbb{L} \\ \overline{X} &\mapsto b \end{aligned} \tag{19.210}$$

qui se prolonge en $R(\overline{X}) \mapsto R(b)$ pour tout $R \in \mathbb{F}_p[X]$. Cette application est bien définie parce que $Q(b) = 0$. Elle est injective parce que $R(b) = 0$ ne peut pas avoir lieu avec $R \in \mathbb{F}_p[X]/Q$ parce

que Q est le polynôme minimal de b. La surjectivité vient alors du fait que les deux corps ont le même nombre d'éléments. □

19.4.7 Construction de $\mathbb{F}_{p^n}$

Le théorème 19.30 nous indique que, pour tout $n \in \mathbb{N}$ et $q = p^n$, il existe un unique corps possédant q éléments. Ce corps est noté $\mathbb{F}_q$.

Le théorème 19.52 nous incite à chercher à écrire $\mathbb{F}_q$ sous la forme

$$\mathbb{F}_q = \mathbb{F}_p[X]/(P) \tag{19.211}$$

pour un certain polynôme irréductible $P \in \mathbb{F}_p[X]$.

19.4.7.1 La version du faignant

Nous pouvons construire le corps à $q = p^n$ éléments en prenant le quotient de $\mathbb{F}_p[X]$ par n'importe quel polynôme irréductible de degré n. Le résultat est le suivant.

PropHfrNCB

Proposition 19.59.
Soit P un polynôme unitaire irréductible dans $\mathbb{F}_p[X]$. Nous posons $\mathbb{K} = \mathbb{F}_p[X]/(P)$. Alors

(1) $\mathbb{K}$ est un corps à q éléments.

(2) $\alpha = \overline{X}$ est une racine de P dans $\mathbb{K}$.

ItemiEFRTg

(3) $\mathbb{K} = \mathbb{F}_p[\alpha]$.

Démonstration. (1) En vertu du corolaire 6.44, $\mathbb{K}$ est un corps. Il est aussi un espace vectoriel de dimension n sur $\mathbb{F}_p$, et contient donc $p^n = q$ éléments.

(2) Nous avons $P(\overline{X}) = 0$ par construction de $\mathbb{K} = \mathbb{F}_p[X]/(P)$.

(3) En tant que quotient de $\mathbb{F}_p[X]$, les éléments de $\mathbb{K}$ sont des polynômes en $\overline{X}$.

□

19.4.7.2 La version plus élaborée

Construire $\mathbb{F}_q$ comme quotient de $\mathbb{F}_p[X]$ par un polynôme irréductible quelconque ne donne pas d'information sur les générateurs de $\mathbb{F}_q\backslash\{0\}$, et en particulier il n'est pas toujours vrai que $\overline{X}$ est générateur.

Exemple 19.60.
Construisons $\mathbb{F}_4$. Le polynôme $X^2 + X + 1$ est irréductible dans $\mathbb{F}_2$ parce qu'il n'a pas de racine (c'est vite vu : dans $\mathbb{F}_2$ il n'y a que deux candidats). Donc $\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2 + X + 1)$. △

Remarque 19.61.
Le corps $\mathbb{F}_2$ n'est pas un sous-corps de $\mathbb{C}$ parce que leurs caractéristiques ne sont pas les mêmes. Une conséquence est que les racines de polynômes peuvent être très différentes. Par exemple le polynôme $X^2 + 1$ accepte $x = 1$ comme racine dans $\mathbb{F}_2$ tandis qu'il a pour racines $\pm i$ dans $\mathbb{C}$.

En changeant de corps, les racines peuvent donc complètement changer. Ce n'est pas juste qu'il y a des racines dans l'un et pas dans l'autre.

ExemWUdrcs

Exemple 19.62.
Cherchons à construire $\mathbb{F}_{16}$ comme quotient de $\mathbb{F}_2$ par un polynôme de degré 4.

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: x=polygen(GF(2))
```

```
sage: -x-1
x + 1
sage: Q=x**15-1
sage: Q.factor()
(x + 1) * (x^2 + x + 1) * (x^4 + x + 1) * (x^4 + x^3 + 1)
           * (x^4 + x^3 + x^2 + x + 1)
```

Les polynômes candidats à avoir des racines génératrices sont donc au nombre de 3 :

$$P_1 = X^4 + X + 1 \tag{19.212a}$$
$$P_2 = X^4 + X^3 + 1 \tag{19.212b}$$
$$P_3 = X^4 + X^3 + X^2 + X + 1. \tag{19.212c}$$

Dans le quotient $\mathbb{F}_2[X]/P_3$, l'élément $\overline{X}$ n'est pas générateur. En effet nous avons $X^4 = X^3 + X^2 + X + 1$ et par conséquent les puissances successives de X sont

$$X \tag{19.213a}$$
$$X^2 \tag{19.213b}$$
$$X^3 \tag{19.213c}$$
$$X^4 = X^3 + X^2 + X + 1 \tag{19.213d}$$
$$1. \tag{19.213e}$$

La classe de X dans $\mathbb{F}_2[X]/P_3$ n'est donc pas génératrice du groupe $(\mathbb{F}_2[X]/P_3)\backslash\{0\}$.

Le polynôme $P_1 = X^4 + X + 1$ par contre est primitif parce que les puissances de X dans $\mathbb{F}_2[X]/P_1$ sont

$$X \tag{19.214a}$$
$$X^2 \tag{19.214b}$$
$$X^3 \tag{19.214c}$$
$$X + 1 \tag{19.214d}$$
$$X^2 + X \tag{19.214e}$$
$$X^3 + X^2 \tag{19.214f}$$
$$X + 1 + X^3 \tag{19.214g}$$
$$X^2 + 1 \tag{19.214h}$$
$$X^3 + X \tag{19.214i}$$
$$X + 1 + X^2 \tag{19.214j}$$
$$X^2 + X + X^3 \tag{19.214k}$$
$$X^3 + X^2 + X + 1 \tag{19.214l}$$
$$1 + X^2 + X^2 \tag{19.214m}$$
$$1 + X^3 \tag{19.214n}$$
$$1 \tag{19.214o}$$

Cela fait 15 puissances distinctes, ce qui prouve que P_1 est primitif. Nous verrons plus loin comment alléger un peu la vérification de la primitivité de P_1. △

PropNsLqWb

Proposition 19.63 ([1])**.**
Soient un nombre premier p, un entier non nul $n \in \mathbb{N}\backslash\{0\}$ ainsi qu'un polynôme P irréductible unitaire primitif dans $\mathbb{F}_p[X]$. Nous considérons $\mathbb{K} = \mathbb{F}_p[X]/P$ et $\alpha = \overline{X} \in \mathbb{K}$. En notant $q = p^n$ nous avons

(1) Les racines de P sont $\{\alpha, \alpha^p, \dots, \alpha^{p^{n-1}}\}$ et $\alpha^q = \alpha$.

(2) P est le polynôme minimal de α.

(3) P est scindé dans $\mathbb{K}$.

(4) P divise $X^q - X$ dans $\mathbb{K}$.

(5) La famille $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ est une base de $\mathbb{K}$ en tant qu'espace vectoriel sur $\mathbb{F}_p$.

(6) En tant qu'ensemble,

$$\mathbb{F}_q = \{0, \alpha, \alpha^2, \alpha^3, \dots, \alpha^{q-1}\}, \tag{19.215}$$

et les α^k sont distincts pour $k = 1, \dots, q-1$.

Démonstration. La plupart des assertions sont des corolaires ou des paraphrases de résultats contenus dans les propositions précédentes.

(1) L'assertion à propos des racines de P est contenue dans le lemme 19.55. D'autre part le groupe $(\mathbb{F}_p[X]/P)\backslash\{0\}$ est cyclique d'ordre $q-1$. Par conséquent le corolaire 2.14 indique que $\alpha^{q-1} = 1$ et donc $\alpha^q = \alpha$.

(2) Soit $\tilde{P}$ un polynôme annulateur de α. Nous voyons que si β est racine de $\tilde{P}$ alors β^p est également racine de $\tilde{P}$ en utilisant les techniques habituelles. Par conséquent toutes les racines de P sont racines de $\tilde{P}$, ce qui implique que $\tilde{P}$ est de degré au moins égal à celui de P.

(3) Possédant n racines distinctes dans $\mathbb{K}$, le polynôme P est scindé.

(4) D'après le lemme 19.33 un polynôme irréductible de degré n divise le polynôme $X^{p^n} - X$. Une autre façon de montrer ce point est de remarquer que le polynôme P est scindé et que toutes ses racines sont également racines de $X^q - X$.

(5) Une combinaison linéaire nulle entre les éléments de $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ serait un polynôme annulateur de degré $n-1$ de α. Cet ensemble est donc libre. Par ailleurs un ensemble libre de n éléments dans un espace vectoriel de dimension n est générateur.

(6) Si $\alpha^l = \alpha^k$ avec $k < l$ et $k, l \leqslant q$ alors nous avons $\alpha^r = 1$ avec $r = l - k < q$, ce qui contredirait la primitivité de P. Les éléments $0, \alpha, \dots, \alpha^{q-1}$ étant distincts et au nombre de q, ils forment tout l'ensemble $\mathbb{F}_q$.

□

19.4.8 Exemple : étude de $\mathbb{F}_{16}$

Dans cette section nous voulons construire $\mathbb{F}_{16}$. Nous considérons donc $p = 2$ et $n = 4$. Des polynôme irréductibles de degré 4 dans $\mathbb{F}_2[X]$ ne sont pas très difficiles à trouver. Par exemple $X^4 + X^3 + X^2 + X + 1$.

Si vous en voulez d'autres, en voici.

¡¡ Avertissement/question à la lectrice !! 19.64
Le lemme suivant me semble douteux. Écrivez-moi si vous avez une preuve ou un contre-exemple.
LEMooTBROooANstIL

Lemme 19.65 ([1]).
Soit un polynôme de degré 4 dans $\mathbb{F}_2[X]$. Si il vérifie

(1) le terme constant est non nul,

(2) il y a un nombre impair de termes non nuls,

alors il est irréductible.

Les polynômes primitifs par contre, doivent être trouvés parmi les diviseurs irréductibles de $X^{15} - 1$. Montrons que

$$P = X^4 + X^3 + 1 \tag{19.216}$$

est primitif. Nous posons $\omega = \overline{X} \in \mathbb{F}_2[X]/P$. L'ordre de ω dans le groupe $(\mathbb{F}_2[X]/P)\backslash\{0\}$ doit être un diviseur de 15 et donc peut être seulement 1, 3, 5 ou 15. Le fait que l'ordre ne soit ni 1 ni 3 est trivial parce que le degré de P est 4. Montrons que l'ordre de ω n'est pas 5 non plus :

$$\omega^5 = \omega^4\omega = (\omega^3 + 1)\omega = \omega^4 + \omega = \omega^3 + \omega + 1 \neq 1. \tag{19.217}$$

Dans ce calcul nous avons abondamment utilisé le fait que $-1 = 1$.

À partir de maintenant nous posons $\mathbb{K} = \mathbb{F}_2[X]/P$. Les racines de P sont $\omega, \omega^2, \omega^4$ et ω^8. En effet si β est une racine de P, alors β^2 est une racine en vertu de

$$P(\beta^2) = (\beta^2)^4 + (\beta^2)^3 + 1 = (\beta^4)^2 + (\beta^3)^2 + 1^2 = (\beta^4 + \beta^3 + 1)^2 = 0. \tag{19.218}$$

Ici nous avons implicitement utilisé le lemme 19.50. D'autre part P ne peut pas avoir plus de 4 racines.

Proposition 19.66.
L'ensemble $\{\omega, \omega^2, \omega^4, \omega^8\}$ est une base de $\mathbb{F}_{16}$ sur $\mathbb{F}_2$.

Démonstration. Nous savons que $\{1, \omega, \omega^2, \omega^3\}$ est une base. En effet cet ensemble est libre (sinon ω aurait un polynôme annulateur de degré 3) et générateur parce que l'espace engendré par 4 vecteurs indépendants sur $\mathbb{F}_2$ contient $2^4 = 16$ éléments.

Nous posons $e_0 = 1$, $e_1 = \omega$, $e_2 = \omega^2$, $e_3 = \omega^3$ et $f_1 = \omega$, $f_2 = \omega^2$, $f_3 = \omega^4$, $f_4 = \omega^8$. En utilisant le calcul modulo $\omega^4 + \omega^3 + 1 = 0$ et $2 = 0$ nous trouvons

$$f_1 = \omega \tag{19.219a}$$
$$f_2 = \omega^2 \tag{19.219b}$$
$$f_3 = \omega^3 + 1 \tag{19.219c}$$
$$f_4 = \omega^3 + \omega^2 + \omega. \tag{19.219d}$$

Ensuite nous montrons que les vecteurs e_i peuvent être construits comme combinaisons linéaires des vecteurs f_j :

$$f_1 + f_2 + f_3 + f_4 = e_0 \tag{19.220a}$$
$$f_1 = e_1 \tag{19.220b}$$
$$f_2 = e_2 \tag{19.220c}$$
$$f_1 + f_2 + f_4 = e_3. \tag{19.220d}$$

Les quatre vecteurs f_j forment donc bien une base parce qu'ils sont générateurs d'un espace de dimension 4. □

Exemple 19.67. (1) Résoudre dans $\mathbb{F}_{16}$ l'équation $x^5 = a$ en discutant éventuellement en fonction de la valeur de a.

(2) Montrer qu'il existe quatre éléments $\gamma \in \mathbb{F}_{16}$ tels que pour chacun d'eux l'ensemble $B_\gamma = \{\gamma, \gamma^2, \gamma^4, \gamma^8\}$ est une base de $\mathbb{F}_{16}$ sur $\mathbb{F}_2$ telle que le produit de deux éléments de B_γ est, soit un élement de B_γ, soit 1.

C'est parti !

(1) Si $a = 0$, alors $x = 0$ est la seule solution. Si $a \neq 0$ alors a est une puissance de ω ; nous posons $a = \omega^l$. Nous cherchons x sous la forme $x = \omega^k$. L'équation à résoudre pour k est

$$\omega^{5k} = \omega^l \tag{19.221}$$

où l est donné. Cette équation revient à

$$5k = l \mod 15. \tag{19.222}$$

Si l n'est pas un multiple de 5, alors il n'y a pas de solution. Il n'y a des solutions que pour $l = 0, 5, 10$ et elles sont :

$$k = \begin{cases} 3, 6, 9, 12 & \text{si } l{=}0 \\ 1 & \text{si } l{=}5 \\ 2 & \text{si } l{=}10 \end{cases} \tag{19.223}$$

(2) Nous cherchons γ sous la forme $\gamma = \omega^k$. Parmi les nombreuses contraintes liées à l'énoncé, nous devons avoir

$$\gamma^5 = 1, \gamma, \gamma^2, \gamma^4, \gamma^8. \tag{19.224}$$

Les possibilités $\gamma^5 = \gamma, \gamma^2, \gamma^4, \gamma^5$ ne sont pas bonnes parce qu'elles impliqueraient que B_γ n'est pas une base. Reste à explorer $\gamma^5 = 1$.
Étant donné le premier point, nous restons avec les possibilités

$$\gamma = 1, \omega^3, \omega^6, \omega^9, \omega^{12}. \tag{19.225}$$

Évidemment $\gamma = 1$ ne produit pas une base. Avec $\gamma = \omega^3$ nous trouvons

$$B_\gamma = \{\omega^3, \omega^6, \omega^{12}, \omega^{24}\} = \{\omega^3, \omega^6, \omega^{12}, \omega^9\} \tag{19.226}$$

où nous avons utilisé le fait que $\omega^k = \omega^{k \bmod 15}$. En utilisant le fait que $\omega^4 = \omega^3 + 1$ nous trouvons

$$\omega^5 = \omega^3 + \omega + 1 \tag{19.227a}$$
$$\omega^6 = \omega^3 + \omega^2 + \omega + 1 \tag{19.227b}$$
$$\omega^9 = \omega^2 + 1 \tag{19.227c}$$
$$\omega^{12} = \omega + 1. \tag{19.227d}$$

L'ensemble B_γ est alors formé des éléments

$$f_1 = \omega^3 \tag{19.228a}$$
$$f_2 = \omega^3 + \omega^2 + \omega + 1 \tag{19.228b}$$
$$f_3 = \omega + 1 \tag{19.228c}$$
$$f_4 = \omega^2 + 1. \tag{19.228d}$$

Il est assez simple de vérifier que c'est une base en remarquant que $f_1 + f_2 + f_2 + f_4 = 1$.
Les possibilités $\gamma = \omega^6, \omega^9, \omega^{12}$ produisent les mêmes ensembles B_γ.

△

19.4.9 Polynômes irréductibles sur $\mathbb{F}_q$

DefWXBkOxg

Définition 19.68.
*La **fonction de Möbius** est la fonction $\mu\colon \mathbb{N}\backslash\{0\} \to \{-1, 0, 1\}$ définie par*

$$\mu(n) = \begin{cases} 0 & \text{si } n \text{ est divisible par un carré différent de } 1, \\ 1 & \text{si } n \text{ est le produit d'un nombre pair de nombres premiers distincts,} \\ -1 & \text{si } n \text{ est le produit d'un nombre impair de nombres premiers distincts,} \end{cases} \tag{19.229}$$

PROPooOVYJooFvmxyj

Proposition 19.69 ([576])**.**
Si m et n sont strictement positifs et premiers entre eux, alors

$$\mu(mn) = \mu(m)\mu(n). \tag{19.230}$$

De plus nous avons

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases} \tag{19.231}$$

PropLBZoIoO

Proposition 19.70 (Formule d'inversion de Möbius[576]).
Soient $f, g : \mathbb{N} \to \mathbb{C}$ telles que pour tout $n \geqslant 1$,

$$g(n) = \sum_{d|n} f(d). \tag{19.232}$$

Alors

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) \tag{19.233}$$

où μ est la fonction de Möbius pour tout $n \geqslant 1$.

LemRGuWqNu

Lemme 19.71 ([577]).
Soient $P, Q \in \mathbb{K}[X]$ ayant une racine commune dans une extension $\mathbb{L}$ de $\mathbb{K}$. Si P est irréductible, alors $P \mid Q$.

Démonstration. Si P ne divise pas Q, alors P et Q sont premiers entre eux parce que dans la décomposition en irréductibles de Q, il n'y a pas de P tandis que dans celle de P, il n'y a que P. Par conséquent, il existe $a, b \in \mathbb{K}[X] \subset \mathbb{L}[X]$ tels que[23] $aP + bQ = 1$. Cette dernière égalité est encore valable dans $\mathbb{L}$ et donc rend impossible l'existence d'une racine commune. □

PropVFNOvzZ

Proposition 19.72 ([577, 104]).
Soit p un nombre premier, $n \geqslant 1$ et $r \in \mathbb{N}\backslash\{0\}$. Nous notons $q = p^r$, $A(n,q)$, l'ensemble des polynômes unitaires irréductibles de degré n sur $\mathbb{F}_q$. Nous notons aussi $I(n,q) = \text{Card}\big(A(n,q)\big)$. Alors :

(1) Le polynôme $X^{q^n} - X$ se décompose en irréductibles de la façon suivante :

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in A(d,q)} P. \tag{19.234}$$

(2) Le nombre d'irréductibles est donné par

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \tag{19.235}$$

où μ est la fonction de Möbius (définition 19.68).

(3) Nous avons l'équivalence de suite

$$I(n,q) \sim_{n\to\infty} \frac{q^n}{n}. \tag{19.236}$$

Démonstration. (1) Soit un diviseur d de n et $P \in A(d,q)$. Montrons que P divise $X^{q^n} - X$. Nous considérons le corps $\mathbb{K} = \mathbb{F}_q[X]/(P)$, qui est une extension de degré $\deg(P)$ de $\mathbb{F}_q$ parce qu'il s'agit des polynômes de degré au maximum $\deg(P)$ à coefficients dans $\mathbb{F}_q$. Ce corps possède donc q^d éléments et est isomorphe à $\mathbb{F}_{q^d}$ par la proposition 19.58. Par construction dans $\mathbb{K}$, l'élément $\alpha = [X]$ (la classe de X dans le quotient par P) est une racine de P. Cet élément est également une racine de $X^{q^d} - X$ parce que tout élément de $\mathbb{F}_{q^d}$ est une racine de ce polynôme. Ce dernier point est la proposition 19.33.

Nous sommes donc dans la situation où P et $X^{q^d} - X$ ont une racine commune dans l'extension $\mathbb{F}_q[X]/(P)$. Nous en déduisons que α est aussi une racine de $X^{q^n} - X$. En effet en utilisant le fait que $\alpha^{q^d} = \alpha$, nous avons

$$\alpha^{q^n} = \alpha^{q^{kd}} = \alpha^{q^d q^{(k-1)d}} = \left(\alpha^{q^d}\right)^{q^{(k-1)d}} = \alpha^{q^{(k-1)d}}, \tag{19.237}$$

23. Théorème de Bézout, 6.47.

donc par récurrence, on a encore $\alpha^{q^n}=\alpha$, et α est racine de $X^{q^n}-X$. Puisque P est irréductible, le lemme 19.71 nous indique alors, que P divise $X^{q^n}-X$.
Étant donné que tous les éléments de $A(d,q)$ divisent $X^{q^n}-X$ et sont irréductibles, leur produit divise encore $X^{q^n}-X$:

$$\prod_{d\,|\,n}\prod_{P\in A(d,q)}P\mid X^{q^n}-X. \tag{19.238}$$

Nous devons à présent montrer que tous les facteurs irréductibles de $X^{q^n}-X$ sont dans un $A(d,q)$ avec $d\mid n$. Soit donc P un facteur irréductible de $X^{q^n}-X$ de degré $d\geqslant 1$. Nous posons encore $\mathbb{K}=\mathbb{F}_q[X]/(P)$ et nous utilisons la propriété de multiplication sur les degrés (proposition 6.63) :

$$[\mathbb{F}_{q^n}:\mathbb{K}][\mathbb{K}:\mathbb{F}_q]=[\mathbb{F}_{q^n}:\mathbb{F}_q]=n, \tag{19.239}$$

donc $[\mathbb{K}:\mathbb{F}_q]$, qui vaut $\deg(P)$ est un diviseur de n.
Étant donné que $X^{q^n}-X$ n'a que des racines simples sur $\mathbb{F}_{q^n}$ (à nouveau la proposition 19.33), dans sa décomposition en irréductibles sur $\mathbb{F}_q$, il n'a pas de facteur carré ; il n'a donc qu'une fois chacun des $P\in A(d,q)$ avec $d\mid n$. Autrement dit, tous les facteurs irréductibles de $X^{q^n}-X$ sont dans le produit $\prod_{d\,|\,n}\prod_{P\in A(d,q)}P$ et donc $X^{q^n}-X$ divise ce gros produit :

$$X^{q^n}-X\mid\prod_{d\,|\,n}\prod_{P\in A(d,q)}P. \tag{19.240}$$

Ayant déjà obtenu la divisibilité inverse et les polynômes étant unitaires, nous avons égalité.

(2) Nous passons au degré dans l'expression que nous venons de démontrer :

$$q^n=\sum_{d\,|\,n}d\,\mathrm{Card}\left(A(d,q)\right)=\sum_{d\,|\,n}dI(d,q). \tag{19.241}$$

Nous pouvons utiliser la formule d'inversion de Möbius (proposition 19.70) pour les fonctions $g(n)=q^n$ et $f(n)=dI(n,q)$. Nous écrivons alors

$$f(n)=\sum_{d\,|\,n}\mu\left(\frac{n}{d}\right)q^d, \tag{19.242}$$

ou encore

$$I(n,q)=\frac{1}{n}\sum_{d\,|\,n}\mu\left(\frac{n}{d}\right)q^d, \tag{19.243}$$

ce qu'il fallait.

(3) Nous posons

$$r_n=\sum_{\substack{d\,|\,n\\ d<n}}\mu\left(\frac{n}{d}\right)q^d, \tag{19.244}$$

mais sachant que les diviseurs de n, outre n lui-même, sont tous plus petits ou égaux à $n/2$ et qu'en valeur absolue, la fonction de Möbius est toujours plus petite ou égale à [24] 1,

$$|r_n|\leqslant\sum_{d=1}^{\lfloor n/2\rfloor}q^d=\frac{q-q^{\lfloor n/2\rfloor}}{1-q}=q\frac{q^{\lfloor n/2\rfloor}-1}{q-1}\leqslant\frac{q^{\lfloor n/2\rfloor+1}}{q-1}. \tag{19.245}$$

D'autre part en reprenant la formule déjà prouvée,

$$I(n,q)=\frac{1}{n}\sum_{d\,|\,n}\mu\left(\frac{n}{d}\right)q^d=\frac{1}{n}\left(r_n+\mu\left(\frac{n}{n}\right)q^n\right)=\frac{r_n+q^n}{n}. \tag{19.246}$$

Au numérateur, le plus haut degré en n est q^n parce que r_n est en $q^{\lfloor n/2\rfloor}$. Donc nous avons bien l'équivalence de suite pour $n\to\infty$:

$$\frac{q^n+r_n}{n}\sim_{n\to\infty}\frac{q^n}{n}. \tag{19.247}$$

□

24. Dans [104], ma dernière inégalité arrive comme une égalité.

19.4.10 Matrices

Proposition 19.73.
Nous avons

$$|\mathrm{GL}(n,\mathbb{F}_p)| = (p^n-1)(p^n-p)\dots(p^n-p^{n-1}). \tag{19.248}$$

Démonstration. Par construction il existe une bijection entre $\mathrm{GL}(n,\mathbb{F}_p)$ et l'ensemble des bases de $\mathbb{F}_p^n$. Nous devons donc seulement compter le nombre de bases. Pour le premier vecteur de base nous avons le choix entre les p^n-1 éléments non nuls de $\mathbb{F}_p^n$. Pour le second nous avons le choix entre p^n-p éléments, et ainsi de suite. □

LemcDOTzM

Lemme 19.74.
Soit $\mathbb{K}$ un corps fini autre que $\mathbb{F}_2$ [25]*, soit un groupe abélien M et un morphisme $\varphi\colon \mathrm{GL}(n,\mathbb{K})\to M$. Alors il existe un unique morphisme $\delta\colon \mathbb{K}\backslash\{0\}\to M$ tel que $\varphi=\delta\circ\det$.*

Démonstration. D'abord le groupe dérivé de $\mathrm{GL}(n,\mathbb{K})$ est $\mathrm{SL}(n,\mathbb{K})$ parce que les éléments de $D\big(\mathrm{GL}(n,\mathbb{K})\big)$ sont de la forme $ghg^{-1}h^{-1}$ dont le déterminant est 1.

De plus le groupe $\mathrm{SL}(n,\mathbb{K})$ est normal dans $\mathrm{GL}(n,\mathbb{K})$. Par conséquent $\mathrm{GL}(n,\mathbb{K})/\mathrm{SL}(n,\mathbb{K})$ est un groupe et nous pouvons définir l'application relevée

$$\tilde{\varphi}\colon \frac{\mathrm{GL}(n,\mathbb{K})}{\mathrm{SL}(n,\mathbb{K})}\to M \tag{19.249}$$

vérifiant $\varphi=\tilde{\varphi}\circ\pi$ où π est la projection.

Nous pouvons faire la même chose avec l'application

$$\det\colon \mathrm{GL}(n,\mathbb{K})\to\mathbb{K}\backslash\{0\} \tag{19.250}$$

qui est un morphisme de groupes dont le noyau est $\mathrm{SL}(n,\mathbb{K})$. Cela nous donne une application

$$\tilde{\det}\colon \frac{\mathrm{GL}(n,\mathbb{K})}{\mathrm{SL}(n,\mathbb{K})}\to\mathbb{K}\backslash\{0\} \tag{19.251}$$

telle que $\det=\tilde{\det}\circ\pi$. Cette application $\tilde{\det}$ est un isomorphisme. En effet, elle est surjective parce que le déterminant l'est, et elle est injective parce que son noyau est précisément ce par quoi on prend le quotient. Par conséquent $\tilde{\det}$ possède un inverse et nous pouvons écrire

$$\varphi=\tilde{\varphi}\circ\tilde{\det}^{-1}\circ\tilde{\det}\circ\pi. \tag{19.252}$$

Étant donné que $\tilde{\det}\circ\pi=\det$, nous avons alors $\varphi=\delta\circ\det$ avec $\delta=\tilde{\varphi}\circ\tilde{\det}^{-1}$. Ceci conclut la partie existence de la preuve.

En ce qui concerne l'unicité, nous considérons $\delta'\colon\mathbb{K}\backslash\{0\}\to M$ telle que $\varphi=\delta'\circ\det$. Pour tout $u\in\mathrm{GL}(n,\mathbb{K})$ nous avons $\delta'(\det(u))=\varphi(u)=\delta(\det(u))$. L'application det étant surjective depuis $\mathrm{GL}(n,\mathbb{K})$ vers $\mathbb{K}\backslash\{0\}$, nous avons $\delta'=\delta$. □

Théorème 19.75.
Soient $p\geqslant 3$ un nombre premier et E, un $\mathbb{F}_p$-espace vectoriel de dimension finie n. Pour tout $u\in\mathrm{GL}(E)$ nous avons

$$\epsilon(u)=\left(\frac{\det(u)}{p}\right). \tag{19.253}$$

Ici ϵ est la signature de u vue comme une permutation des éléments de $\mathbb{F}_p$.

Démonstration. Commençons par prouver que

$$\epsilon\colon \mathrm{GL}(E)\to\{-1,1\}. \tag{19.254}$$

25. Je ne comprends pas très bien à quel moment joue cette hypothèse.

est un morphisme. Si nous notons $\overline{u} \in S(E)$ l'élément du groupe symétrique correspondant à la matrice $u \in \mathrm{GL}(E)$, alors nous avons $\overline{uv} = \overline{u} \circ \overline{v}$, et la signature étant un morphisme (proposition 1.300),

$$\epsilon(uv) = \epsilon(\overline{u} \circ \overline{v}) = \epsilon(\overline{u})\epsilon(\overline{v}). \tag{19.255}$$

Par ailleurs $\{-1, 1\}$ est abélien, donc le lemme 19.74 s'applique et nous pouvons considérer un morphisme $\delta\colon \mathbb{F}_p\backslash\{0\} \to \{-1, 1\}$ tel que $\epsilon = \delta \circ \det$.

Nous allons utiliser le lemme 19.40 pour montrer que δ est le symbole de Legendre. Pour cela il nous faudrait trouver un $x \in \mathbb{F}_p\backslash\{0\}$ tel que $\delta(x) = -1$. Étant donné que det est surjective, nous cherchons ce x sous la forme $x = \det(u)$. Par conséquent nous aurions

$$\delta(x) = (\delta \circ \det)(u) = \epsilon(u), \tag{19.256}$$

et notre problème revient à trouver une matrice $u \in \mathrm{GL}(E)$ dont la permutation associée soit de signature -1.

Soit $n = \dim E$; en conséquence de la proposition 19.59(3), l'espace $\mathbb{E}_q = \mathbb{F}_{p^n}$ est un $\mathbb{F}_p$-espace vectoriel de dimension n et est donc isomorphe en tant qu'espace vectoriel à E. Étant donné que $\mathbb{F}_q$ est un corps fini, nous savons que $\mathbb{F}_q\backslash\{0\}$ est un groupe cyclique à $q-1$ éléments. Soit y, un générateur de $\mathbb{F}_q\backslash\{0\}$ et l'application

$$\begin{aligned} \beta\colon \mathbb{F}_q &\to \mathbb{F}_q \\ x &\mapsto yx. \end{aligned} \tag{19.257}$$

Cela est manifestement $\mathbb{F}_p$-linéaire (ici y et x sont des classes de polynômes et $\mathbb{F}_p$ est le corps des coefficients). L'application β fixe zéro et à part zéro, agit comme le cycle

$$(1, y, y^2, \dots, y^{q-2}). \tag{19.258}$$

Nous savons qu'un cycle de longueur n est de signature $(-1)^{n+1}$. Ici le cycle est de longueur $q-1$ qui est pair (parce que $p \geqslant 3$) et par conséquent, l'application β est de signature -1. □

19.5 Constructions à la règle et au compas

Définition 19.76 ([578]).
Soit E une partie de $\mathbb{R}^2$. Un point de $\mathbb{R}^2$ est constructible en une étape à partir de E si il est un point de E ou une intersection de deux objets parmi

- *les droites passant par deux points distincts de E ;*
- *les cercles centrés en un point de E et dont le rayon est la distance entre deux points de E.*

Nous notons $C_1(E)$ l'ensemble des points constructibles en une étape à partir de E.

*Les points constructibles en n étapes à partir de E sont définis par récurrence : $C_{n+1}(E) = C_1\big(C_n(E)\big)$. Enfin un point de $\mathbb{R}^2$ est **constructible** à partir de E si il appartient à*

$$C(E) = \bigcup_{n=1}^{\infty} C_n(E). \tag{19.259}$$

*Un réel est **constructible** si il est l'abscisse d'un point constructible.*

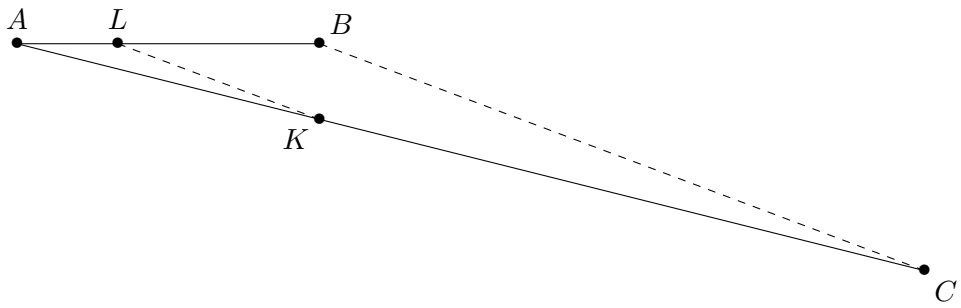
Pour toute la suite, nous allons considérer les points et réels constructibles à partir de l'ensemble $E = \{(0,0), (0,1)\}$.

19.5.1 Quelques constructions

PropIMFooDWAyoH

Proposition 19.77 ([579]).
Les nombres rationnels sont tous constructibles.

Démonstration. Si le réel r est constructible, alors kr est également constructible pour tout $k \in \mathbb{Z}$. Nous devons donc seulement pouvoir construire le nombre $1/n$ pour tout $n \in \mathbb{N}\backslash\{0\}$.

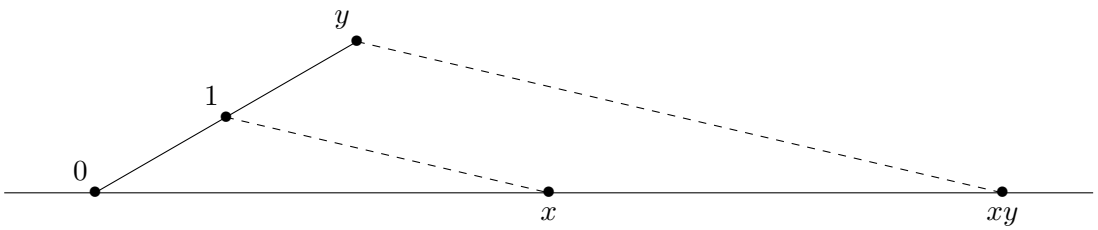


La méthode pour construire le nombre $1/n$ est la suivante. Soit $[AB]$ un segment de longueur 1 (par exemple $A = (0,0)$ et $B = (1,0)$) et un point C, non aligné avec A et B, et tel que $[AC]$ ait une longueur n. Nous plaçons sur $[AC]$ le point K situé à une distance 1 de A en pointant le compas en A et en traçant le cercle de rayon $[AB]$.

La droite passant par K et parallèle à (BC) coupe $[AB]$ en un point L. Maintenant, le segment $[AL]$ a une longueur de $1/n$, par le théorème de Thalès [26]. □

ExGROooIosiBt

Exemple 19.78 (Multiplication à la règle et au compas[578])**.**
Soient x et y deux nombres constructibles. Montrons qu'il est possible de construire le nombre xy. La construction est la suivante :



— On trace deux droites sécantes en A.

— Sur la première nous plaçons le point Y à distance y de A et le point P à distance 1 de A.

— Sur la seconde on place le point X à distance x de A.

— On trace la droite (PX)

— Puis la parallèle à (PX) passant par Y.

— Le point d'intersection entre cette dernière droite et (AX) est le point B.

La longueur AB est égale à xy. △

ExTYMooSMCvSr

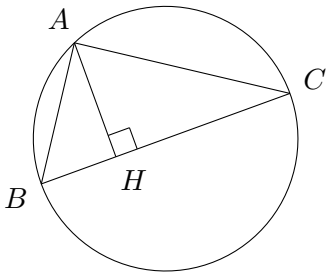
Exemple 19.79 (Racine carrée à la règle et au compas[578])**.**
Nous supposons que le nombre x est constructible, et nous voulons une construction qui donne un segment de longueur $\sqrt{x}$. Nous traçons un segment $[BC]$ dont la longueur correspond à la plus grande des valeurs entre x et 1, puis le cercle de diamètre BC, ensuite le point H sur $[BC]$ tel que BH corresponde à la plus petite des valeurs entre x et 1, enfin la perpendiculaire à (BC) menée par H, qui rencontre le cercle en un point A. D'après le théorème de Thalès sur le cercle [27], le triangle ABC est rectangle en A.

Les triangles ABC et ABH sont donc semblables parce qu'ils sont rectangles avec un angle (autre que l'angle droit) égal. Nous avons donc proportionnalité des longueurs des côtés :

$$\frac{AB}{BH} = \frac{BC}{AB}, \tag{19.260}$$

ce qui donne $AB^2 = BC \times BH = x$ (BC et BH valent respectivement 1 et x ou le contraire).

26. Théorème 12.162.
27. Théorème 12.163.

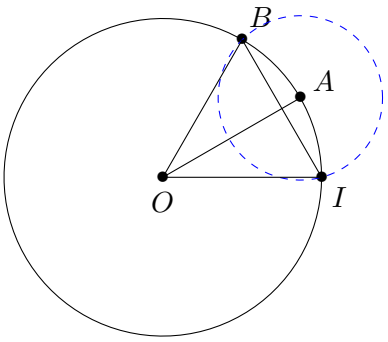


△

ExAHCooELGGPa

Exemple 19.80 (Duplication d'un angle)**.**
Si un angle α est constructible, nous allons construire les angles 2α, 3α, etc. Pour cela nous considérons un cercle de centre O et les points A et I sur le cercle tels que $\widehat{AOI} = \alpha$. Le cercle de centre A et de rayon AI intersecte le cercle de départ en les points I et B.

Le point A est à égale distance de B et I ; le point O également. Donc la droite (OA) est médiatrice du segment $[BI]$. Par conséquent elle est la hauteur du triangle isocèle OBI. L'angle $\widehat{BOA}$ est alors le même que $\widehat{AOI}$; par conséquent $\widehat{BOI} = 2\alpha$.



En traçant le cercle de centre B et de rayon BA, nous continuons et nous construisons 3α. △

L'exemple suivant qui permet d'additionner des angles repose sur le fait que deux cordes de mêmes longueurs sous-tendent des angles égaux, et est une adaptation simple de la duplication d'angle.

ExOVDooXnWPDl

Exemple 19.81 (Addition d'angles)**.**
Quitte à soustraire ou additionner un certain nombre de fois $90°$, nous supposons que les deux angles donnés sont entre $0°$ et $90°$.

Soient A, B, I sur un cercle de centre O, et nous notons $\alpha = \widehat{OAI}$, $\beta = \widehat{BOI}$. Nous traçons le cercle de centre B et de rayon AI ; il intersecte le cercle en des points K_1 et K_2. Les angles $\widehat{K_1OB}$ et $\widehat{K_2OB}$ sont tous deux égaux à α.

Les angles $\widehat{K_1OA}$ et $\widehat{K_2OA}$ sont égaux à $\beta - \alpha$ et $\beta + \alpha$. △

19.82.
Notez que l'exemple 19.81 donne un moyen de construire $\alpha + \beta$ et $\alpha - \beta$ ensemble. Il ne permet pas de savoir lequel est $\alpha + \beta$ et lequel est $\alpha - \beta$.

19.5.2 Nombres constructibles

ThoRHFooZsLbqd

Théorème 19.83 (Wantzel[580])**.**
Le réel a est constructible si et seulement si il existe une suite finie de corps $\mathbb{L}_i$ tels que

(1) $\mathbb{L}_0 = \mathbb{Q}$,

(2) $\mathbb{L}_{i+1}$ *est un extension quadratique*[28] *de* $\mathbb{L}_i$

28. C'est-à-dire une extension finie de degré 2.

(3) $a \in \mathbb{L}_n$.

Démonstration. Soit $\mathbb{K}$ un corps de nombres constructibles (par exemple $\mathbb{Q}$) ; nous notons $E_{\mathbb{K}}$ l'ensemble des points de $\mathbb{R}^2$ dont les coordonnées sont dans $\mathbb{K}$. Ce sont des points forcément constructibles.

(1) **Intersection de droites** Si $A, B \in E_{\mathbb{K}}$ alors la droite (AB) a pour équation $ax+by+c=0$ avec $a, b, c \in \mathbb{K}$, et le point d'intersection entre deux droites est donné par la solution du système

$$\begin{cases} ax + by + c = 0 & (19.261a) \\ a'x + b'y + c' = 0, & (19.261b) \end{cases}$$

dont les solutions sont encore dans $\mathbb{K}$.

(2) **Intersection droite-cerle** L'équation d'un cercle est de la forme

$$(x-u)^2 + (y-v)^2 = r^2 \tag{19.262}$$

où $(u, v) \in E_{\mathbb{K}}$ et r est la distance entre deux points de $E_{\mathbb{K}}$; donc $r^2 \in \mathbb{K}$. En développant et en redéfinissant u, v nous voyons que tous les cercles à considérer ont une équation de la forme

$$x^2 + y^2 + ux + vy + t = 0 \tag{19.263}$$

avec $u, v, t \in \mathbb{K}$. Il s'agit de voir où sont les solutions du système

$$\begin{cases} ax + by + c = 0 & (19.264a) \\ x^2 + y^2 + ux + vy + t = 0. & (19.264b) \end{cases}$$

Si $a \neq 0$ alors nous pouvons opérer la substitution $x = -(c + by)/a$ et obtenir l'équation suivante pour y :

$$\left(\frac{-c-by}{a}\right)^2 + y^2 + u\left(\frac{-c-by}{a}\right) + vy + t = 0. \tag{19.265}$$

Cette équation est de la forme $P(y) = 0$ où P est un polynôme du second degré à coefficients dans $\mathbb{K}$.

Si $a = 0$ nous substituons y au lieu de x et le résultat est le même.

C'est le moment de relire la proposition 6.103 qui nous assure que si le réel α est une solution de $P(y) = 0$ hors de $\mathbb{K}$, alors l'extension $\mathbb{K}(\alpha)$ est de degré 2 parce que le polynôme minimal de α est de degré 2 :

$$\big[\mathbb{K}[\alpha] : \mathbb{K}\big] = 2. \tag{19.266}$$

De plus $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$.

(3) **Intersection cercle-cercle** Le système

$$\begin{cases} x^2 + y^2 + ux + vy + t = 0 & (19.267a) \\ x^2 + y^2 + ax + by + c = 0 & (19.267b) \end{cases}$$

est équivalent au système (en substituant la seconde équation par la différence entre les deux)

$$\begin{cases} x^2 + y^2 + ux + vy + t = 0 & (19.268a) \\ (u-a)x + (v-b)y + t - c = 0 & (19.268b) \end{cases}$$

qui est à nouveau une intersection entre un cercle et une droite.

Passons à la conclusion. Si α est un nombre constructible, alors il apparaît dans les coordonnées d'un point de $C_m(E_{\mathbb{Q}})$ pour un certain m. Nous supposons (pour la récurrence) que chaque réel constructible en $m-1$ étapes possède sa pile d'extension quadratiques en partant de $\mathbb{Q}$. Le nombre α vérifie donc

$$a\alpha^2 + b\alpha + c = 0 \tag{19.269}$$

avec a, b, c trois réels constructibles en $m-1$ étapes. L'hypothèse de récurrence donne donc des piles d'extensions

$$\begin{array}{ccc} \mathbb{L}_0 = \mathbb{Q} & \mathbb{L}'_0 = \mathbb{Q} & \mathbb{L}''_0 = \mathbb{Q} \\ L_{i+1} = \mathbb{L}_i(a_i) & \mathbb{L}'_{i+1} = \mathbb{L}'_i(b_i) & \mathbb{L}''_{i+1} = \mathbb{L}''_i(c_i) \\ a \in \mathbb{L}_n & b \in \mathbb{L}'_{n'} & c \in \mathbb{L}''_{n''}. \end{array} \tag{19.270}$$

Nous considérons donc la suite d'extensions de $\mathbb{Q}$ qui consiste à étendre successivement par les nombres $a_1, \dots, a_n, b_1, \dots, b_{n'}, c_1, \dots, c_{n''}$ tout en excluant les doublons : il est possible que par exemple b_3 soit déjà dans $\mathbb{Q}(a_1, \dots, a_n)$. Cela nous fournit une suite d'extensions

$$\begin{aligned} \mathbb{M}_0 &= \mathbb{Q} \\ \mathbb{M}_{i+1} &= \mathbb{M}_i(\alpha_i) \\ a, b, c &\in \mathbb{M}_n. \end{aligned} \tag{19.271}$$

Ici le n n'est pas spécialement le même que celui plus haut. Maintenant, α est dans une extension quadratique de $\mathbb{M}_n$.

Pour la réciproque, nous supposons avoir une tour d'extensions quadratiques $\mathbb{L}_0 = \mathbb{Q}$, $\mathbb{L}_i$, et $\alpha \in \mathbb{L}_n$, et nous voulons prouver que α est constructible. Nous y allons par récurrence : si $n = 0$ alors α est rationnel, et il est constructible par la proposition 19.77.

Supposons que tous les points à coordonnées dans $\mathbb{L}_i$ sont constructibles. Alors, nous allons prouver que les éléments de $\mathbb{L}_{i+1}$ le sont également. Puisque α est solution d'une équation de degré 2 à coefficients dans $\mathbb{L}_i$, nous avons

$$a\alpha^2 + b\alpha + c = 0 \tag{19.272}$$

pour certains $a, b, c \in \mathbb{L}_i$ et donc

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{19.273}$$

Les exemples 19.78 et 19.79 montrent que les produits et les racines carrées de nombres constructibles sont constructibles. Donc α est constructible. □

19.5.3 Polygones constructibles

Définition 19.84.
*Un angle α est **constructible** si le nombre $\cos(\alpha)$ est constructible.*

La raison est qu'il suffit de prendre la perpendiculaire à l'axe horizontal.

LemMAHooXcOCpr

Lemme 19.85 ([104]).
Soient m et n, deux nombres premiers entre eux. L'angle $\frac{2\pi}{mn}$ est constructible si et seulement si les angles $\frac{2\pi}{m}$ et $\frac{2\pi}{n}$ sont constructibles.

Démonstration. (1) **Sens direct** Il suffit de pouvoir multiplier un angle par un entier, ce qui est fait dans l'exemple 19.80.

(2) **Sens réciproque** Le théorème de Bézout 1.236 nous donne $a, b \in \mathbb{Z}$ tels que $an + bm = 1$. Cela donne immédiatement

$$\frac{1}{mn} = \frac{a}{m} + \frac{b}{n}, \tag{19.274}$$

et donc

$$\frac{2\pi}{mn} = a\frac{2\pi}{m} + b\frac{2\pi}{n}, \tag{19.275}$$

ce qui fait que l'angle $2\pi/mn$ est une combinaison entière d'angles constructibles. Il est donc constructible par l'exemple 19.81.

□

LemUKNooSBzDyY

Lemme 19.86.
Soit un entier $n \geqslant 3$ ayant

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{19.276}$$

comme décomposition en facteurs premiers. Le polynôme régulier à n côtés est constructible si et seulement si les angles $\frac{2\pi}{p_i^{\alpha_i}}$ sont constructibles.

Démonstration. Le polygone régulier à n côtés est constructible si et seulement si l'angle $\frac{2\pi}{n}$ est constructible, c'est-à-dire si l'angle

$$\frac{2\pi}{\prod_{i=1}^{k} p_i^{\alpha_i}} \tag{19.277}$$

est constructible. Par récurrence sur le lemme 19.85, cet angle est constructible si et seulement si les angles $\frac{2\pi}{p_i^{\alpha_i}}$ sont constructibles. □

ThoTWAooEsLjJu

Théorème 19.87 (Gauss-Wantzel[104])**.**
Soit $\alpha \in \mathbb{N}\backslash\{0\}$.

ItemFSEooONDFrSi

(1) L'angle $\frac{2\pi}{2^\alpha}$ est constructible.

ItemFSEooONDFrSii

*(2) Si p est premier et $p \neq 2$, alors l'angle $\frac{2\pi}{p^\alpha}$ est constructible si et seulement si $\alpha = 1$ et p est un **nombre de Fermat**, c'est-à-dire de la forme $1 + 2^{(2^\beta)}$ pour $\beta \in \mathbb{N}$.*

ItemFSEooONDFrSiii

(3) Le polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de 2 et d'un nombre fini de nombres premiers de Fermat distincts.

Démonstration. Le point (1) est une construction de bissectrice, et le point (3) consistera à remettre en place différents morceaux. Le gros de la preuve est donc consacré à (2).

(1) **Sens direct** Nous supposons que l'angle $\frac{2\pi}{p^\alpha}$ est constructible ; alors le nombre $\cos\big(2\pi/p^\alpha\big)$ l'est également et le théorème de Wantzel 19.83 nous indique que

$$\Big[\mathbb{Q}\Big(\cos(2\pi/p^\alpha)\Big) : \mathbb{Q}\Big] = 2^m \tag{19.278}$$

pour un certain m. Posons $q = p^\alpha$ et $\omega = e^{2i\pi/q}$. Grâce au corolaire 19.24, nous savons que le polynôme minimal de ω est le polynôme cyclotomique ϕ_q dont le degré est [29]

$$\varphi(q) = \varphi(p^\alpha) = p^{\alpha-1}(p-1). \tag{19.279}$$

Par conséquent

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big] = p^{\alpha-1}(p-1). \tag{19.280}$$

Mais par ailleurs

$$\omega + \omega^{-1} = 2\cos(2\pi/q), \tag{19.281}$$

EqJDBooHURUQa

donc $\cos(2\pi/q) \in \mathbb{Q}(\omega)$. Et en multipliant (19.281) par ω nous trouvons le polynôme annulateur suivant pour ω :

$$\omega^2 - 2\cos\left(\frac{2\pi}{q}\right)\omega + 1 = 0. \tag{19.282}$$

Cela signifie que ω est de degré 2 dans $\mathbb{Q}\big(\cos(2\pi/q)\big)$, c'est-à-dire

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/q)\big)\big] = 2. \tag{19.283}$$

EqSJXooRCLJyt

29. Voir juste en dessous de la définition 19.17.

En remettant bout à bout et en utilisant la propriété multiplicative des degrés des extensions [30],

$$\underbrace{\big[\mathbb{Q}(\omega):\mathbb{Q}\big]}_{p^{\alpha-1}(p-1)}=\underbrace{\Big[\mathbb{Q}(\omega):\mathbb{Q}\big(\cos(2\pi/p^\alpha)\big)\Big]}_{2}\underbrace{\Big[\mathbb{Q}\big(\cos(2\pi/p^\alpha)\big):\mathbb{Q}\Big]}_{2^m},\tag{19.284}$$

donc $p^{\alpha-1}(p-1)=2^{m+1}$, mais comme p est premier et impair, $\alpha=1$ et $p=2^{m+1}+1$. Par ailleurs $m+1$ est un entier et nous nous proposons de mettre en facteur la puissance de 2 dans son développement en facteurs premiers : $m+1=\lambda 2^\beta$ avec $\beta\in\mathbb{N}$ et un certain nombre impair $\lambda\in\mathbb{N}^*$. Nous avons :

$$p=2^{\lambda 2^\beta}+1=\Big(2^{2^\beta}\Big)^\lambda+1,\tag{19.285}$$

mais le lemme 3.136(1) nous indique alors que $1+2^{2^\beta}$ divise $1+\Big(2^{2^\beta}\Big)^\lambda=p$. Mais vu que p est premier, il ne peut être divisé que par p lui-même et donc λ soit être égal à 1 et nous avons

$$p=1+2^{2^\beta},\tag{19.286}$$

ce qui signifie que p est un nombre de Fermat premier.

(2) **Sens réciproque** Nous supposons que p est un nombre premier de la forme $p=1+2^{2^\beta}$ et nous devons prouver que l'angle $\frac{2\pi}{p}$ est constructible. Pour cela nous posons tout de suite $n=2^\beta$ et $\omega=e^{2i\pi/p}$. Comme dans la première partie nous nous souvenons que le polynôme minimal de ω est le polynôme cyclotomique ϕ_p de degré $p-1$; donc

$$[\mathbb{Q}(\omega):\mathbb{Q}]=p-1.\tag{19.287}$$

(2a) **Un groupe d'automorphismes** Nous considérons le groupe $G=\mathrm{Aut}_{\mathbb{Q}}\big(\mathbb{Q}(\omega)\big)$ des automorphismes du corps $\mathbb{Q}(\omega)$ agissant sur $\mathbb{Q}$ comme l'identité. Tous les éléments de $\mathbb{Q}(\omega)$ étant des polynômes en ω (proposition 6.103), un élément $g\in G$ est uniquement déterminé par $g(\omega)$, et de plus $g(0)=0$ ainsi que $\phi_p(\omega)=0$ et que ϕ_p commute avec g, donc

$$\phi_p\big(g(\omega)\big)=g\big(\phi_p(\omega)\big)=g(0)=0,\tag{19.288}$$

ce qui signifie que $g(\omega)$ est une racine du polynôme cyclotomique ϕ_p. Par définition 19.17, les racines sont $\{\omega,\omega^2,\dots,\omega^{p-1}\}$, ce qui signifie que l'action de g consiste à élever ω à une certaine puissance entre 1 et $p-1$. Le groupe G est donc d'ordre $|G|=p-1$ et a pour éléments

$$g_k:\omega\mapsto\omega^k\tag{19.289}$$

avec $k=1,\dots,p-1$. L'application [31]

$$\begin{aligned}\psi:G&\to(\mathbb{Z}/p\mathbb{Z})^*\\ g_k&\mapsto[k]_p.\end{aligned}\tag{19.290}$$

est un morphisme surjectif entre deux groupes finis de même cardinal, donc c'est un isomorphisme. Le corolaire 19.32 nous donne de plus l'isomorphisme $(\mathbb{Z}/p\mathbb{Z})^*\simeq\mathbb{Z}/(p-1)\mathbb{Z}$, ce qui fait que G a un élément d'ordre $p-1$ (parce que $\mathbb{Z}/(p-1)\mathbb{Z}$ en a un). Nous notons g_0 cet élément.

(2b) **La tour d'extensions** À partir de cet élément $g_0\in G$ nous définissons avec $0\leqslant i\leqslant n$:

$$\mathbb{K}_i=\{z\in\mathbb{Q}(\omega)\text{ tel que }g_0^{2^i}(z)=z\}.\tag{19.291}$$

Ces ensembles sont bien des corps parce que g_0 est un morphisme : $g_0(zz')=g_0(z)g_0(z')$; on en déduit immédiatement que $g_0^2(zz')=g_0^2(z)g_0^2(z')$.

30. Proposition 6.63.
31. Pour rappel, la notation $[k]_p$ est la classe de l'entier k modulo p, qui est un élément de $\mathbb{Z}/p\mathbb{Z}$.

(2c) $\mathbb{K}_n = \mathbb{Q}(\omega)$ Cela est dû au fait que g_0 est, par définition, d'ordre 2^n, ce qui signifie que $g_0^{2^n}(z) = z$ pour tout $z \in \mathbb{Q}(\omega)$.

(2d) $\mathbb{K}_0 = \mathbb{Q}$ Puisque les éléments de G laissent $\mathbb{Q}$ invariant nous avons forcément $\mathbb{Q} \subset \mathbb{K}_0$. Nous nous attelons maintenant à prouver l'inclusion inverse. Nous savons par la proposition 6.103(2) que $\{\omega^i\}_{i=0,\dots,p-2}$ est une base [32] de $\mathbb{Q}(\omega)$. Vu que $g_0 : \mathbb{Q}(\omega) \to \mathbb{Q}(\omega)$ est un isomorphisme, l'image d'une base est une base :

$$\{g_0(\omega)^i\}_{i=0,\dots,p-2} \tag{19.292}$$

est également une base de $\mathbb{Q}(\omega)$. Soit $z \in \mathbb{K}_0$ et décomposons-le dans cette base [33] :

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^i; \tag{19.293}$$

nous appliquons g_0 à cette égalité en tenant compte du fait que $g_0(z) = z$:

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^{i+1}. \tag{19.294}$$

En identifiant les coefficients (et en remarquant que le dernier terme dans (19.293) est le premier dans (19.294)) nous voyons que tous les coefficients sont égaux :

$$z = \lambda_0 g_0(\omega + \cdots + \omega^{p-1}) = \lambda_0 g_0\big(\phi_p(\omega) - 1\big) = -\lambda_0 \in \mathbb{Q}. \tag{19.295}$$

Dans cette chaine d'égalités nous avons utilisé le fait que $\phi_p(X) = 1 + X + \cdots + X^{p-1}$ (corolaire 19.22) et que $\phi_p(\omega) = 0$ par définition.

(2e) $\mathbb{K}_i \subset \mathbb{K}_{i+1}$ **strictement** Nous avons une inclusion pour la simple raison que si $z \in \mathbb{K}_i$, alors $g_0^{2^i}(z) = z$. Par conséquent :

$$g_0^{2^{i+1}}(z) = g_0^{2\times 2^i}(z) = (g_0^{2^i})^2(z) = z. \tag{19.296}$$

Afin de voir que l'inclusion est stricte nous montrons que l'élément

$$x = \sum_{k=0}^{2^{n-i}-1} g_0^{k2^{i+1}}(\omega) \tag{19.297}$$

est dans $\mathbb{K}_{i+1}\backslash\mathbb{K}_i$. Dans la base $\{1, \omega, \dots, \omega^{p-1}\}$ l'élément x a une composante ω avec le terme $k = 0$, mais si on lui applique $g_0^{2^i}$ nous obtenons

$$g_0^{2^i}(x) = \sum_{k=0}^{2^{n-i-1}-1} g_0^{2^i(1+2k)}(\omega). \tag{19.298}$$

Une composante ω pour cela demanderait d'avoir $2^i(1 + 2k) = \lambda 2^n$ avec $\lambda \in \mathbb{N}$; cela demanderait

$$k = \lambda 2^{n-i-1} - \frac{1}{2} \tag{19.299}$$

ce qui est impossible. Donc $x \notin \mathbb{K}_i$.

Nous montrons que $x \in \mathbb{K}_{i+1}$ de la même manière :

$$g_0^{2^{i+1}}(x) = \sum_{k=0}^{2^{n-i-1}-1} g^{2^{i+1}(k+1)}(\omega). \tag{19.300}$$

Soit k_0 entre 0 et $2^{n-i-1} - 1$; nous voulons trouver un k entre 0 et $2^{n-i-1} - 1$ tel que $k_0 2^{i+1} = 2^{i+1}(k + 1)$; cela est très simple : il suffit de prendre $k = k_0 - 1$ tant que $k_0 \neq 0$. Si $k_0 = 0$ alors cela correspond au terme ω dans (19.297), et il se trouve dans (19.300) avec $k = 2^{n-i-1} - 1$. Donc $g_0^{2^{i+1}}(x) = x$ et $x \in \mathbb{K}_{i+1}$.

32. Ici [104] parle de $\mathbb{Q}(\omega)/\mathbb{Q}$. Dans ce cas il me semble qu'il faille faire partir les valeurs de i de 0 et non de 1.
33. Il s'agit d'une base en tant qu'espace vectoriel sur $\mathbb{Q}$, donc les λ_i sont dans $\mathbb{Q}$.

(2f) **Les degrés dans la tour** Par définition de n et de ω, et par la propriété multiplicative des degrés[34] nous avons :

$$2^n = p - 1 = \big[\mathbb{Q}(\omega) : \mathbb{Q}\big] = \underbrace{\big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] \dots \big[\mathbb{K}_1 : \mathbb{Q}\big]}_{n \text{ facteurs}}. \tag{19.301}$$

Un produit de n facteurs entiers tous strictement plus grands que zéro doit valoir 2^n. Ils doivent donc tous valoir 2 et nous avons en particulier

$$\big[\mathbb{K}_{n-1} : \mathbb{Q}\big] = 2^{n-1}. \tag{19.302}$$

(2g) $\cos(2\pi/p) \in \mathbb{K}_{n-1}$ Soit $f = g_0^{2^n-1}$; puisque l'action de g_0 (comme de tous les éléments de G) est de décaler les puissances de ω, il existe $\lambda \in \mathbb{N}$ tel que $f(\omega) = \omega^\lambda$. Mais $f^2 = \text{Id}$ donc

$$\omega = f^2(\omega) = \omega^{2\lambda}, \tag{19.303}$$

ce qui donne $1 = \omega^{\lambda^2-1}$ ou encore que p divise λ^2-1 parce que ω est une racine primitive p^{e}de l'unité. Par conséquent $[\lambda^2-1]_p = 0$ ce qui donne $[\lambda]_p = \pm[1]_p$, mais comme $f \neq \text{Id}$ nous avons

$$[\lambda]_p = -[1]_p. \tag{19.304}$$

Ce que donne $f(\omega) = \omega^{-1}$.
Par ailleurs $\mathbb{Q}(\omega)$ est un corps, donc il contient

$$\cos\left(\frac{2\pi}{p}\right) = \frac{1}{2}(\omega + \omega^{-1}). \tag{19.305}$$

Nous en déduisons que $\cos(2\pi/p)$ est un point fixe de $f = g_0^{2^{-1}}$:

$$f\big(\cos(2\pi/p)\big) = \frac{1}{2}\big(f(\omega) + f(\omega^{-1})\big) = \cos(2\pi/p). \tag{19.306}$$

Être un point fixe de $g_0^{2^{n-1}}$ signifie être un élément de $\mathbb{K}_{n-1}$:

$$\cos\left(\frac{2\pi}{p}\right) \in \mathbb{K}_{n-1}. \tag{19.307}$$

(2h) **Questions de degrés** Nous avons alors les (non)inclusions suivantes :

$$\mathbb{Q}\left(\cos(2\pi/p)\right) \subset \mathbb{K}_{n-1} \subsetneq \mathbb{K}_n = \mathbb{Q}(\omega), \tag{19.308}$$

ce qui fait que

$$1 < \big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] \leqslant \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2 \tag{19.309}$$

EqGEKooJBAQbg

La première inégalité est le fait que $\mathbb{Q}(\omega)$ n'est pas égal à $\mathbb{K}_{n-1}$. La dernière égalité se démontre de la même manière que (19.283) (ici $\alpha = 1$). Les inégalités de (19.309) sont donc en réalité des égalités :

$$\big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] = \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2. \tag{19.310}$$

Cela, combiné au fait que $\mathbb{Q}\big(\cos(2\pi/p)\big) \subset \mathbb{K}_{n-1}$ donne

$$\mathbb{K}_{n-1} = \mathbb{Q}\big(\cos(2\pi/p)\big). \tag{19.311}$$

Mais nous savons déjà que $\big[\mathbb{K}_{n-1} : \mathbb{Q}\big] = 2^{n-1}$, donc

$$\big[\mathbb{Q}\big(\cos(2\pi/p)\big) : \mathbb{Q}\big] = 2^{n-1} \tag{19.312}$$

et le nombre $\cos(2\pi/p)$ est bien sûr le sommet d'une tour d'extensions quadratiques partant de $\mathbb{Q}$. Il est donc constructible par le théorème de Wantzel 19.83.

34. Proposition 6.63

Il est maintenant l'heure de conclure en prouvant le point (3). D'abord si n est un produit de nombres premiers de Fermat distincts, alors $n = p_1 \dots p_m$ et l'angle $\frac{2\pi}{n}$ est constructible si et seulement si chacun des angles $\frac{2\pi}{p_k}$ est constructible (lemme 19.86), ce qui est le cas d'après la partie (2). Si n est le produit d'une puissance de 2 avec un produit de nombres premiers de Fermat, le polygone à n côtés est tout autant constructible : il suffit de bissecter les angles.

À l'inverse si le polynôme à n côtés (avec n impair) est constructible, alors $n = \prod_i p_i^{\alpha_i}$, dont l'angle est constructible si et seulement si $\alpha = 1$ et p_i est un nombre premier de Fermat. Et si le polygone à n côtés (n pair) est constructible, alors $n = 2^\lambda \prod_i p_i^{\alpha_i}$. Ce polygone est constructible si et seulement si le polygone à $\prod_i p_i^{\alpha_i}$ côtés est constructible : il suffit de regrouper les côtés par deux. □

Remarque 19.88.
D'après Wikipédia[581], les seuls nombres de Fermat connus pour être premiers sont

$$F_0 = 3 \tag{19.313a}$$
$$F_1 = 5 \tag{19.313b}$$
$$F_2 = 17 \tag{19.313c}$$
$$F_3 = 257 \tag{19.313d}$$
$$F_4 = 65537. \tag{19.313e}$$

Pour les autres, essentiellement on ne sait pas. Il n'est même pas sûr qu'il y en ait d'autres. Le problème des polygones constructibles n'est donc pas encore tout à fait terminé.

Remarque 19.89.
Les angles $\frac{2\pi}{p^\alpha}$ avec p premier et $\alpha > 1$ ne sont pas constructibles. Il n'est donc pas possible de trisecter l'angle $\frac{2\pi}{3}$. Voilà qui règle un des vieux problèmes de l'antiquité.

Chapitre 20

Intégration sur des variétés

20.1 Variétés

20.1.1 Introduction

Soit $f : S^2 \to \mathbb{R}$ une fonction définie sur la sphère usuelle $S^2 \subset \mathbb{R}^3$. Une question naturelle est d'estimer la régularité de f ; est-elle continue, dérivable, différentiable ? Il n'existe pas de dérivée directionnelle étant donné que le quotient différentiel

$$\frac{f(x + \epsilon u_1, y + \epsilon u_2) - f(x, y)}{\epsilon}$$

n'a pas de sens pour un point $(x + \epsilon u_1, y + \epsilon u_2)$ qui n'est pas –sauf valeurs particulières– dans la surface. Pour la même raison il n'est pas possible de parler de différentiabilité de cette manière. Comment faire, sans devoir étendre le domaine de définition de f à un voisinage de la sphère ? Une solution possible est de parler de la notion de variété.

Une variété est un objet qui ressemble, vu de près, à $\mathbb{R}^m$ pour un certain m. En d'autres termes, on imagine une variété comme un recollement de morceaux de $\mathbb{R}^m$ vivant dans un espace plus grand $\mathbb{R}^n$. Ces morceaux sont appelés des ouverts de carte, et l'application qui exprime la ressemblance à $\mathbb{R}^m$ est l'application de carte.

20.1.2 Définition, carte

DEFooJOAHooAxcLHe

Définition 20.1.
*Une **variété** de dimension m de classe C^1 est une partie M de $\mathbb{R}^n$ ($n \geqslant m$) munie d'un ensemble de paires $\{U_\alpha, \varphi_\alpha\}$ où U_α est un ouvert de $\mathbb{R}^m$ et $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ est une application vérifiant*

(1) φ_α est une bijection entre U_α et $\varphi_\alpha(U_\alpha)$;

(2) pour tout $a \in M$, il existe un α tel que $a \in \varphi_\alpha(U_\alpha)$;

(3) pour tout α, β, la partie $\varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big)$ est un ouvert de U_α ;

(4) L'application

$$\varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \to U_\beta \tag{20.1}$$

est un C^1-difféomorphisme vers son image.

Notez que cette définition n'utilise pas du tout la structure de $\mathbb{R}^n$ dans lequel se trouve M. Seulement la structure différentielle du $\mathbb{R}^m$ depuis lequel les cartes partent. En particulier, nous n'avons pas dit si $\varphi_\alpha : U_\alpha \to M$ était de classe C^1 pour une topologie induite de $\mathbb{R}^n$ vers M. Nous pouvons en réalité définir plus généralement une variété en remplaçant « une partie de $\mathbb{R}^n$ » par « un ensemble » sans rien y changer.

La difficulté qui apparaît lorsque nous voulons dire que M est un ensemble quelconque au lieu d'une partie de $\mathbb{R}^n$ est pour la mesure.

Nous verrons plus loin comment les cartes φ_α peuvent transporter la mesure de U_α (celle de Lebesgue dans $\mathbb{R}^m$) vers une partie de $\mathbb{R}^n$ en utilisant le déterminant de $d\varphi_\alpha$. Cela n'est possible

que parce que $\mathbb{R}^n$ a une structure différentielle qui permet de définir $d\varphi_\alpha$. Si M est un ensemble quelconque, cette façon de faire n'est plus possible. Dans ce cas, il faut définir des fibrés tangents et cotangents puis en extraire une « forme volume ». Le choix de la forme volume n'étant pas canonique, il n'y a pas d'intégrale canonique sur une variété générale comme il y en a une sur une variété « dans $\mathbb{R}^n$ ».

DEFooWABZooGIEDEV

Définition 20.2.
*Soit une variété M de dimension m. Une **carte** pour M est un couple (U, φ) où U est un ouvert de $\mathbb{R}^m$ et $\varphi\colon U \to M$ est une application*

(1) qui est une bijection entre U et son image,

(2) pour tout α,

$$\varphi_\alpha^{-1} \circ \varphi\colon \varphi^{-1}\big(\varphi(U) \cap \varphi_\alpha(U_\alpha)\big) \to \varphi_\alpha^{-1}\big(\varphi(U) \cap \varphi_\alpha(U_\alpha)\big) \tag{20.2}$$

est un C^1-difféomorphisme.

Proposition 20.3 ([1]).
Si (U_1, φ_1) et (U_2, φ_2) sont des cartes pour la variété M, alors en posant $S = \varphi_1(U_1) \cap \varphi_2(U_2)$, l'application

$$\varphi_2^{-1} \circ \varphi_1\colon \varphi_1^{-1}(S) \to \varphi_2^{-1}(S) \tag{20.3}$$

est un C^1-difféomorphisme.

Démonstration. Le fait que $\varphi_2^{-1} \circ \varphi_1$ soit une bijection est dû au fait que nous ayons choisi les espaces de départ et d'arrivée pour que ce soit une bijection.

Soit $a \in \varphi_1(U_1) \cap \varphi_2(U_2)$. Soient α tel que $a \in \varphi_\alpha(U_\alpha)$ et un ouvert A autour de $\varphi_\alpha^{-1}(a)$ tel que $\varphi_\alpha(A) \subset \varphi_1(U_1) \cap \varphi_2(U_2)$. Nous allons montrer que

$$\varphi_2^{-1} \circ \varphi_1\colon \varphi_1\big(\varphi_\alpha(A)\big) \to \varphi_2^{-1}\big(\varphi_\alpha(A)\big) \tag{20.4}$$

est un C^1-difféomorphisme. Nous avons

$$\varphi_2^{-1} \circ \varphi_1 = \varphi_2^{-1} \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_1, \tag{20.5}$$

mais vu que φ_1 et φ_2 sont des cartes, les applications $\varphi_2^{-1} \circ \varphi_\alpha$ et $\varphi_\alpha \circ \varphi_1$ sont de classe C^1 et d'inverses C^1. La composition l'est encore. □

Nous avons demandé qu'une variété n'admette que des cartes partant d'ouverts de $\mathbb{R}^m$. Aurions-nous pu admettre, dans la définition 20.2 que la carte parte d'un ouvert de $\mathbb{R}^p$ avec $p \neq m$? Non. Voici une résultat qui dit que si une carte part de $\mathbb{R}^m$, alors toutes les cartes doivent partir de $\mathbb{R}^m$.

Proposition 20.4.
Soit une variété M de dimension m et une carte $\varphi_\alpha\colon U_\alpha \to M$. Si U est un ouvert de $\mathbb{R}^p$ et si $\varphi\colon U \to M$ est telle que $\varphi_\alpha^{-1} \circ \varphi$ soit un difféomorphisme, alors $p = m$.

Démonstration. Soit $A = \varphi_\alpha(U_\alpha) \cap \varphi(U)$. Les parties $\varphi_\alpha^{-1}(A)$ et $\varphi^{-1}(U)$ sont des ouverts de $\mathbb{R}^m$ et de $\mathbb{R}^p$ respectivement. Par hypothèse, l'application

$$\varphi_\alpha^{-1} \circ \varphi\colon \varphi^{-1}(A) \to \varphi_\alpha^{-1}(A) \tag{20.6}$$

est un difféomorphisme entre l'ouvert $\varphi^{-1}(A) \subset \mathbb{R}^p$ et l'ouvert $\varphi_\alpha^{-1}(A) \subset \mathbb{R}^m$. La proposition 10.285 implique que $m = p$. □

20.1.3 Ancienne définition

Cette section est à recycler.

Définition 20.5.
Soit $\varnothing \neq M \subset \mathbb{R}^n$, $1 \leqslant m < n$ et $k \geqslant 1$. M est une variété de classe C^k de dimension m *si pour tout $a \in M$, il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, et un ouvert V de $\mathbb{R}^m$ tel que $U \cap M$ soit le graphe d'une fonction $f : V \subset \mathbb{R}^m \to \mathbb{R}^{n-m}$ de classe C^1, c'est-à-dire qu'il existe un réagencement des coordonnées $(x_{i_1}, \ldots, x_{i_m}, x_{i_{m+1}}, \ldots, x_{i_n})$ avec*

$$M \cap U = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ tel que } (x_{i_1}, \ldots, x_{i_m}) \in V \quad \left\{ \begin{array}{rcl} x_{i_{m+1}} & = & f_1(x_{i_1}, \ldots, x_{i_m}) \\ \vdots & = & \vdots \\ x_{i_n} & = & f_{n-m}(x_{i_1}, \ldots, x_{i_m}) \end{array} \right. \right\}$$

où V est un voisinage ouvert de $(a_{i_1}, \ldots, a_{i_m}) \in \mathbb{R}^m$.

La littérature regorge de théorèmes qui proposent des conditions équivalentes à la définition d'une variété. Celle que nous allons le plus utiliser est la suivante

Proposition 20.6.
Soit $M \subset \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. L'ensemble M est une variété si et seulement si $\forall a \in M$, il existe un voisinage ouvert $\mathcal{U}$ de a dans $\mathbb{R}^n$ et une application $F : W \subset \mathbb{R}^m \to \mathbb{R}^n$ où W est un ouvert tels que

(1) F est un homéomorphisme de W vers $M \cap \mathcal{U}$,

(2) $F \in C^1(W, \mathbb{R}^n)$,

(3) Le rang de $dF(w) \in L(\mathbb{R}^m, \mathbb{R}^n)$ est de rang maximum (c'est-à-dire m) en tout point $w \in W$.

Pour rappel, si $T : \mathbb{R}^m \to \mathbb{R}^n$ est une application linéaire, son rang[1] est la dimension de son image. Si A est la matrice d'une application linéaire, alors le rang de cette application linéaire est égal à la taille de la plus grande matrice carrée de déterminant non nul contenue dans A[2].

La condition de rang maximum sert à éviter le genre de cas de la figure 20.1 qui représente l'image de l'ouvert $[-1, 1]$ par l'application $F(t) = (t^2, t^3)$.

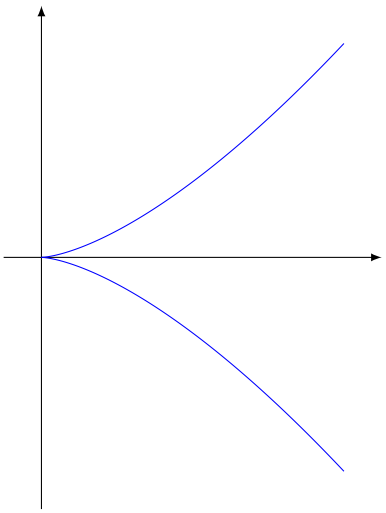


FIGURE 20.1: Quelque chose qui n'est pas de rang maximum et qui n'est pas une variété. LabelFigExemplePasRang

La différentielle a pour matrice

$$dF(t) = (2t, 3t^2). \tag{20.7}$$

Le rang maximum est 1, mais en $t = 0$, la matrice vaut $(0, 0)$ et son rang est zéro. Pour toute autre valeur de t, c'est bon.

Une autre caractérisation des variétés est donnée par la proposition suivante

1. Définition 4.46.
2. Proposition 4.125

PropCarVarZerFonc

Proposition 20.7.
Soit $M \subset \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. L'ensemble M est une variété si et seulement si $\forall a \in M$, il existe un voisinage ouvert $\mathcal{U}$ de a dans $\mathbb{R}^n$ tel et une application $G \in C^1(\mathcal{U}, \mathbb{R}^{n-m})$ tel que

(1) le rang de $dG_a \in L(\mathbb{R}^n, \mathbb{R}^{n-m})$ soit maximum (c'est-à-dire $n-m$) en tout $a \in M$,

(2) $M \cap \mathcal{U} = \{x \in \mathcal{U} \text{ tel que } G(x) = 0\}$.

20.1.4 Espace tangent

DEFooMCEYooHuCAdg

Proposition-Définition 20.8 ([582])**.**
Soit une variété M de dimension m dans $\mathbb{R}^n$. Soit $a \in M$. L'ensemble

$$T_aM = \left\{ \gamma'(0) \text{ tel que } \begin{cases} \gamma \in C^1(]-1,1[, M) \\ \gamma(0) = a \end{cases} \right\} \tag{20.8}$$

*est un sous-espace vectoriel de dimension m, appelé **espace tangent** de M en a.*

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

20.2 Intégration

20.2.1 Le problème pour une intégration globale

Soient une variété M et une partie A de M. Si nous avons une fonction $f\colon A \to \mathbb{R}$, nous voudrions définir une intégrale de f sur A en utilisant les cartes.

Il y a une petite complication. Supposons que A soit dans $\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$. Pour la carte φ_α nous aurions tendance à vouloir écrire

$$\int_A f = \int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha \tag{20.9}$$

EQooUEVHooQXgonw

pour la mesure de Lebesgue induite de $\mathbb{R}^m$ vers son ouvert U_α. Pas de problème à ce que (20.9) soit bien définie. Le problème est que α n'est pas à priori spécial par rapport à β et qu'il faudrait également que

$$\int_A f = \int_{\varphi_\beta^{-1}(A)} f \circ \varphi_\beta \tag{20.10}$$

EQooEOJIooJGLQHs

Mais en utilisant le changement de variable (théorème 14.281(3)) pour le C^1-difféomorphisme $\phi = \varphi_\alpha^{-1} \circ \varphi_\beta$, nous avons

$$\int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha = \int_{\varphi_\beta^{-1}(A)} (f \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_\beta)|J_\phi| = \int_{\varphi_\beta^{-1}(A)} (f \circ \varphi_\beta)|J_\phi|. \tag{20.11}$$

À moins d'une coïncidence extraordinaire sur les valeurs du jacobien, il n'y aura pas égalité entre (20.9) et (20.10).

Pour faire mieux, il faudra ajouter quelque chose qui compense l'arrivée du jacobien. C'est ce que nous allons faire maintenant.

20.2.2 Intégrale sur une carte

Dans un premier temps, nous allons définir l'intégrale sur une carte. Nous verrons plus tard comment combiner les cartes pour faire une intégrale sur une variété entière.

PROPooAHWooAfxvyv

Proposition-Définition 20.9 ([1])**.**
Soient deux cartes (U_1, φ_1) et (U_2, φ_2) pour la partie A d'une variété de dimension m dans $\mathbb{R}^n$. Soit une fonction mesurable $f\colon A \to [0, +\infty]$ [3].

3. Si f est seulement mesurable, allez voir les hypothèses du théorème 14.281(3).

Alors[4]

$$\int_{U_1}(f\circ\varphi_1)(x)\sqrt{\det\big((d\varphi_1)_x^*(d\varphi_1)_x\big)}dx=\int_{U_2}(f\circ\varphi_2)(y)\sqrt{\det\big((d\varphi_2)_y^*(d\varphi_2)_y\big)}dy \tag{20.12}$$

où les deux intégrales sont au sens de la mesure de Lebesgue sur des ouverts de $\mathbb{R}^m$.

Ce nombre est ***l'intégrale*** *de* f *sur* A*, et est noté*

$$\int_A f. \tag{20.13}$$

Démonstration. Vu que nous sommes en présence de cartes pour une variété, les applications $\varphi_1\colon U_1\to A$ et $\varphi_2\colon U_2\to A$ sont des bijections telles que $\varphi_1^{-1}\circ\varphi_2\colon U_2\to U_1$ soit un C^1-difféomorphisme.

Nous considérons le C^1-difféomorphisme $\phi=\varphi_1^{-1}\circ\varphi_2$. Notre but est d'utiliser le théorème de changement de variables 14.281(2).

Vu que φ_1 est une bijection, $(\varphi_1\circ\varphi_1^{-1})(A)=A$ et nous avons

$$\phi^{-1}\big(\varphi_1^{-1}(A)\big)=\varphi_2^{-1}(A). \tag{20.14}$$

Ensuite, $f\circ\varphi_1\circ\phi=f\circ\varphi_2$.

Jusqu'ici, rien de drôle me diriez-vous. Ok. Alors voyons un peu comment se passe le jacobien et ce qui se trouve sous la racine carrée.

L'application $x\mapsto\det\big((d\varphi_1)_x^*\circ(d\varphi_1)_x\big)$ doit être composée avec ϕ pour donner

$$y\mapsto\det\big((d\varphi_1)_{\phi(y)}^*\circ(d\varphi_1)_{\phi(y)}\big). \tag{20.15}$$

Attardons-nous un peu sur $(d\varphi_1)_{\phi(y)}$. Nous savons par la règle de différentiation en chaine 10.277 que

$$d(\varphi_1\circ\phi)_y=(d\varphi_1)_{\phi(y)}\circ(d\phi)_y. \tag{20.16}$$

EQooCBGAooCneiJv

Le lemme 10.284 nous dit que l'application linéaire $d\phi_y\colon\mathbb{R}^m\to\mathbb{R}^m$ est inversible et que $d\phi_y^{-1}=(d\phi^{-1})_{\phi(y)}$. En composant les deux côtés de (20.16) par cela nous trouvons

$$(d\varphi_1)_{\phi(y)}=d(\varphi_1\circ\phi)_y\circ(d\phi)_y^{-1}=(d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}. \tag{20.17}$$

C'est me moment d'utiliser la proposition 8.40 à propos de la composition des applications adjointes. Ce qui est dans le déterminant est :

$$\big((d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}\big)^*\circ\big((d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}\big)=\underbrace{(d\phi^{-1})_{\phi(y)}^*}_{\mathbb{R}^m\to\mathbb{R}^m}\circ\underbrace{(d\varphi_2)_y^*\circ(d\varphi_2)_y}_{\mathbb{R}^m\to\mathbb{R}^m}\circ\underbrace{(d\phi^{-1})_{\phi(y)}}_{\mathbb{R}^m\to\mathbb{R}^m}. \tag{20.18}$$

Nous utilisons à présent un combo entre les propriétés des déterminants et de l'adjoint : les propositions 8.16(2) et 8.39 font sortir $|\det\big((d\phi^{-1})_{\phi(y)}\big)|$ de la racine carré. Ne pas oublier la valeur absolue parce que $\sqrt{x^2}=|x|$ et non $\sqrt{x^2}=x$. Cela pour dire que

$$\sqrt{\det\big((d\varphi_1)_{\phi(y)}^*\circ(d\varphi_1)_{\phi(y)}\big)}=|\det\big((d\phi^{-1})_{\phi(y)}\big)|\sqrt{\det\big((d\varphi_2)_y^*\circ(d\varphi_2)_y\big)}. \tag{20.19}$$

En ce qui concerne l'intégrale que nous voulons calculer, en mettant les bouts ensemble,

$$\begin{aligned}\int_{\varphi_1^{-1}(A)}&(f\circ\varphi_1)(x)\sqrt{\det(\cdots)}dx\\&=\int_{\varphi_2^{-1}(A)}(f\circ\varphi_2)(y)|\det\big((d\phi^{-1})_{\phi(y)}\big)|\sqrt{\det\big((d\varphi_2)_y^*\circ(d\varphi_2)_y\big)}|\det(d\phi_y)|dy.\end{aligned} \tag{20.20}$$

Enfin, les déterminants se suppriment : dans la valeur absolue nous avons :

$$\det\big((d\phi^{-1})_{\phi(y)}\big)\det(d\phi_y)=\det\big((d\phi^{-1})_{\phi(y)}\circ d\phi_y\big)=\det(\mathrm{Id})=1. \tag{20.21}$$

□

4. Voir la définition de l'adjoint 8.37.

20.10.
Notez que si A est un ouvert de $\mathbb{R}^m$ lui-même, le nombre défini dans la définition 20.9 est l'intégrale usuelle. Pour le voir, choisir $U = A$ ainsi que la carte

$$\begin{aligned} \varphi\colon U &\to A \\ x &\mapsto x. \end{aligned} \tag{20.22}$$

La différentielle de φ est l'identité, de telle sorte que $\det\big((d\varphi)^*_x(d\varphi)_x\big) = 1$. De même, $f \circ \varphi = f$. Il reste seulement $\int_U f(x)dx$.

Si nous voulons que la formule de la proposition 20.9 fournisse une définition raisonnable de $\int_A f$, il faut que si (U, φ) est une carte,

$$\det\big(d\varphi^*_x \circ d\varphi_x\big) \geqslant 0 \tag{20.23}$$

pour tout x, et ne soit nul sur aucun ouvert de U.

Lemme 20.11 ([1])**.**
Soit une carte (U, φ) d'une variété M de dimension m. Soit $x \in U$. Alors

(1) L'opérateur $d\varphi^*_x \circ d\varphi_x$ est autoadjoint semi-défini positif.

(2) Il existe une base de $\mathbb{R}^m$ formée de vecteurs propres de $d\varphi^*_x \circ d\varphi_x$. Les valeurs propres sont réelles et positives.

(3) Pour tout $x \in U$ nous avons

$$\det(d\varphi^*_x \circ d\varphi_x) > 0. \tag{20.24}$$

Démonstration. Point par point.

(1) Soit $x \in U$. Nous notons $A = d\varphi_x$. Par la proposition 8.40 nous avons $(AA^*)^* = AA^*$ du fait que $(A^*)^* = A$. Donc A^*A est autoadjoint. Sa matrice est donc symétrique par la proposition 8.37(2).
De plus, pour tout $u \in \mathbb{R}^m$ nous avons

$$\langle A^*Au, u\rangle = \langle Au, Au\rangle \geqslant 0. \tag{20.25}$$

Cela implique que la matrice de A^*A est semi-définie positive par le lemme 8.234(2). L'opérateur A^*A est également semi-défini positif par la proposition 8.235.

(2) L'existence de la base de vecteurs propres est le théorème 8.227. Ce théorème dit de plus que les valeurs propres sont réelles. Le fait qu'elles soient positives est le fait déjà prouvé que A^*A est semi-défini positif.

(3) Dans la base de vecteurs propres, la matrice est diagonale avec des nombres positifs sur la diagonale. Le déterminant est alors le produit des valeurs propres. Voilà pourquoi le déterminant est positif.
Il nous reste à montrer que ce déterminant ne peut pas être nul. Si $\det(d\varphi^*_a \circ d\varphi_a) = 0$, alors l'application $d\varphi^*_a \circ d\varphi$ a un noyau (proposition 8.16(2)), c'est-à-dire un vecteur u tel que $d\varphi^*_a d\varphi_a(u) = 0$ en particulier,

$$0 = \langle d\varphi^*_a d\varphi_a u, u\rangle = \langle d\varphi_a u, d\varphi_a u, \rangle = \|d\varphi_a u\|^2. \tag{20.26}$$

Donc $d\varphi_a(u) = 0$.
Mais $\varphi\colon U \to \varphi(U)$ est bijective, donc l'application $\varphi^{-1}\colon \varphi(U) \to U$ l'est aussi et vu que φ est une carte, l'application $\varphi^{-1} \circ \varphi\colon U \to U$ est égale à l'identité. L'identité est une application linéaire dont la différentielle est elle-même, c'est-à-dire que

$$d(\varphi^{-1} \circ \varphi)_a = (d\varphi^{-1})_{\varphi(a)} \circ d\varphi_a = \mathrm{Id}\,. \tag{20.27}$$

Si donc $u \neq 0$, nous devons avoir $d\varphi_a(u) \neq 0$.

□

20.2.3 Quelques expressions pour l'élément de volume

Ce que nous appelons « élément de volume » (pour des raisons qui apparaîtront plus tard) est le coefficient $\sqrt{\det(d\varphi_x^* \circ d\varphi_x)}$, introduit pour le besoin de compenser l'arrivée du jacobien en cas de changement de variables. Mais à quoi ressemble ce coefficient pour les variétés de petite dimension ?

C'est ce que nous allons voir maintenant.

20.2.3.1 En dimension un

Nous supposons une variété de dimension un dans $\mathbb{R}^n$. Une carte est donc une application $\varphi\colon \mathbb{R} \to \mathbb{R}^n$ dont nous notons φ_i les composantes. Nous avons, pour $u \in \mathbb{R}$,

$$d\varphi_a(u) = \frac{d}{dt}\Big[\varphi(a+tu)\Big]_{t=0} = \frac{d}{dt}\begin{pmatrix}\varphi_1(a+tu)\\ \vdots \\ \varphi_n(a+tu)\end{pmatrix} = u\begin{pmatrix}\varphi_1'(a)\\ \vdots \\ \varphi_n'(a)\end{pmatrix} = u\varphi'(a). \tag{20.28}$$

Là dedans, vous vous souviendrez que $a \in \mathbb{R}$ et que $\varphi'(a) \in \mathbb{R}^n$.

Nous devons savoir ce que vaut $d\varphi_a^* x$ lorsque $a \in \mathbb{R}$ et $x \in \mathbb{R}^n$. Pour tout $u \in \mathbb{R}$ nous avons

$$\langle d\varphi_a x, u\rangle_{\mathbb{R}} = \langle x, d\varphi_a(u)\rangle_{\mathbb{R}^n} = u\langle x, \varphi'(a)\rangle_{\mathbb{R}^n} = u\varphi'(a) \cdot x = \langle \varphi'(a) \cdot x, u\rangle_{\mathbb{R}}. \tag{20.29}$$

Donc $d\varphi_a^*(x) = \varphi'(a) \cdot x$.

Question notation, nous avons noté $\langle a, b\rangle_{\mathbb{R}} = ab$ et indifféremment $\langle x, y\rangle_{\mathbb{R}^n} = x \cdot y$.

L'application $d\varphi_a^* \circ d\varphi_a$ est une application linéaire bijective entre $\mathbb{R}$ et $\mathbb{R}$. Sa matrice est donc 1×1 et elle se calcule en appliquant l'application au vecteur de base 1 de $\mathbb{R}$:

$$(d\varphi_a^* \circ d\varphi_a)(1) = d\varphi_a^*\big(\varphi'(a)\big) = \varphi'(a) \cdot \varphi'(a) = \|\varphi'(a)\|^2. \tag{20.30}$$

Nous avons donc

$$\det\big(d\varphi_a^* \circ d\varphi_a\big) = \|\varphi'(a)\|^2. \tag{20.31}$$

20.2.3.2 En dimension quelconque

Soit une carte $\varphi\colon \mathbb{R}^m \to \mathbb{R}^n$ pour une variété de dimension m. Les éléments de matrice de la différentielle d'une application sont donnés par la proposition 12.247 :

$$(d\varphi_a)_{ij} = \frac{\partial \varphi_i}{\partial x_j}(a). \tag{20.32}$$

Les éléments de la matrice de $d\varphi_a^*$ sont ceux de la matrice transposée. En ce qui concerne la composition, c'est le produit des matrices :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \sum_k (d\varphi_a^*)_{ik}(d\varphi_a)_{kj} = \sum_k (d\varphi_a)_{ki}(d\varphi_a)_{kj} = \sum_k \frac{\partial \varphi_k}{\partial x_i}(a)\frac{\partial \varphi_k}{\partial x_j}(a). \tag{20.33}$$

Cela pour écrire cette bonne formule :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \frac{\partial \varphi}{\partial x_i}(a) \cdot \frac{\partial \varphi}{\partial x_j}(a). \tag{20.34}$$

EQooQRQKooJVJRsv

20.2.3.3 En dimension deux

Nous considérons maintenant une carte $\varphi\colon U \to \mathbb{R}^3$ avec $U \subset \mathbb{R}^2$. En notant $v_i = \partial_i \varphi(a)$ nous notons la formule (20.34) sous la forme

$$\det\big(d\varphi_a^* \circ d\varphi_a\big) = \det\begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 \\ v_2 \cdot v_1 & v_2 \cdot v_2\end{pmatrix} = \|v_1\|^2\|v_2\|^2 - (v_1 \cdot v_2)^2. \tag{20.35}$$

L'identité de Lagrange de la proposition 10.35 nous donne alors

$$\sqrt{\det(d\varphi_a^* \circ d\varphi_a)} = \|\frac{\partial \varphi}{\partial x} \times \frac{\partial \varphi}{\partial y}\|. \tag{20.36}$$

20.2.3.4 En dimension trois

Nous considérons maintenant une carte $\varphi\colon U \to \mathbb{R}^3$ avec $U \subset \mathbb{R}^3$. Nous notons $u = \frac{\partial\varphi}{\partial x}(a)$, $v = \frac{\partial\varphi}{\partial y}(a)$ et $w = \frac{\partial\varphi}{\partial z}(a)$.

En utilisant l'expression du lemme 10.33 à propos du produit mixte, la formule 20.34 donne

$$\det(d\varphi_a^* \circ d\varphi_a) = \det\begin{pmatrix} \|u\|^2 & u\cdot v & u\cdot w \\ v\cdot u & \|v\|^2 & v\cdot w \\ w\cdot u & w\cdot v & \|w\|^2 \end{pmatrix} = \big|(u\times v)\cdot w\big|. \tag{20.37}$$

Donc

$$\det(d\varphi_a^* \circ d\varphi_a) = \big|(\frac{\partial\varphi}{\partial x}(a)\times\frac{\partial\varphi}{\partial y}(a))\cdot\frac{\partial\varphi}{\partial z}(a)\big|. \tag{20.38}$$

EQooYIJSooHtkXfu

20.3 Intégrale sur une variété

20.3.1 Mesure sur une carte

Nous considérons dans cette section uniquement des variétés M de dimension 2 dans $\mathbb{R}^3$. Une particularité de $\mathbb{R}^3$ (par rapport aux autres $\mathbb{R}^n$) est qu'il existe le produit vectoriel.

Si v, $w \in \mathbb{R}^3$, alors le vecteur $v \times w$ est un vecteur normal au plan décrit par v et w qui jouit de l'importante propriété suivante :

$$\text{aire du parallélogramme} = \|v\times w\|. \tag{20.39}$$

L'aire du parallélogramme construit sur v et w est donnée par la norme du produit vectoriel. Afin de donner une mesure infinitésimale en un point $p \in M$, nous voudrions prendre deux vecteurs tangents à M en p, et puis considérer la norme de leur produit vectoriel. Cette idée se heurte à la question du choix des vecteurs tangents à considérer.

Dans $\mathbb{R}^2$, le choix est évident : nous choisissons e_x et e_y, et nous avons $\|e_x \times e_y\| = 1$. L'idée est donc de choisir une carte $F\colon W \to F(W)$ autour du point $p = F(w)$, et de choisir les vecteurs tangents qui correspondent à e_x et e_y via la carte, c'est-à-dire les vecteurs

$$\frac{\partial F}{\partial x}(w), \quad \text{et} \quad \frac{\partial F}{\partial y}(w). \tag{20.40}$$

L'**élément infinitésimal de surface** sur M au point $p = F(w)$ est alors défini par

$$d\sigma_F = \|\frac{\partial F}{\partial x}(w)\times\frac{\partial F}{\partial y}(w)\|dw, \tag{20.41}$$

et si la partie $A \subset M$ est entièrement contenue dans $F(W)$, nous définissons la **mesure** de A par

$$\mu_2(A) = \int_{F^{-1}(A)} d\sigma_F = \int_{F^{-1}(A)} \|\frac{\partial F}{\partial x}(w)\times\frac{\partial F}{\partial y}(w)\|dw. \tag{20.42}$$

EqDefMuDeuxDF

Remarque 20.12.
Afin que cette définition ait un sens, nous devons prouver qu'elle ne dépend pas du choix de la carte F. En effet, les vecteurs $\partial_x F$ et $\partial_y F$ dépendent de la carte F, donc leur produit vectoriel (et sa norme) dépendent également de la carte F choisie. Il faudrait donc un petit miracle pour que le nombre $\mu_2(A)$ donné par (20.42) soit indépendant du choix de F. Nous allons bientôt voir comme cas particulier du théorème 20.16 que c'est en fait le cas. C'est-à-dire que si F et $\tilde{F}$ sont deux cartes qui contiennent A, alors

$$\int_{F^{-1}(A)} d\sigma_F = \int_{\tilde{F}^{-1}(A)} d\sigma_{\tilde{F}}. \tag{20.43}$$

20.3.1.1 Exemple : la mesure de la sphère

PROPooAFGSooJBZiuo

Proposition 20.13 ([1]).
L'aire de la sphère $S^2 \equiv x^2 + y^2 + z^2 + R^2$ est $4\pi R^2$.

Démonstration. Nous allons calculer l'aire de la demi-sphère. L'application $F\colon B((0,0),R) \to \mathbb{R}^3$ donnée par

$$F(x,y) = \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} \tag{20.44}$$

est une carte pour une demi-sphère. Ses dérivées partielles sont

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -\frac{x}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -\frac{y}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}. \tag{20.45}$$

Le produit vectoriel de ces deux vecteurs tangents donne

$$\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) = \frac{x}{\alpha}e_1 + \frac{y}{\alpha}e_2 + e_3 \tag{20.46}$$

où $\alpha = \sqrt{R^2 - x^2 - y^2}$. En calculant la norme, nous trouvons

$$\| \frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) \| = \sqrt{\frac{R^2}{R^2 - x^2 - y^2}}, \tag{20.47}$$

et en passant aux coordonnées polaires, nous écrivons l'intégrale (20.42) sous la forme

$$\int_B \|\partial_x F \times \partial_y F\| = \int_0^{2\pi} d\theta \int_0^R r\sqrt{\frac{R^2}{R^2 - x^2 - y^2}} dr = 2\pi R^2, \tag{20.48}$$

EQooYGRFooKwEYfV

qui est bien la mesure de la demi-sphère. □

20.3.2 Intégrale sur une carte

Nous pouvons maintenant définir l'intégrale d'une fonction sur une carte de la variété M.

DEFooZNFOooZPiBWY

Définition 20.14.
*Soit $F\colon W \subset \mathbb{R}^2 \to \mathbb{R}^3$, une carte pour une variété M. Soit A, une partie de $F(W)$ telle que $A = F(B)$ où $B \subset W$ est mesurable. Soit encore $f\colon A \to \mathbb{R}$, une fonction continue. L'**intégrale** de f sur A est le nombre*

$$\int_A f = \int_A f d\sigma_F = \int_{F^{-1}(A)} (f \circ F)(w) \| \frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w) \| dw \tag{20.49}$$

EqDefIntDeuxDF

Remarque 20.15.
L'intégrale (20.49) n'est pas toujours bien définie. Étant donné que F est C^1 et que f est continue, l'intégrande est continue. L'intégrale sera donc bien définie par exemple lorsque B est borné et si la fermeture $\bar{A}$ est un compact contenu dans $F(W)$.

Le théorème suivant montre que le travail que nous avons fait jusqu'à présent ne dépend en fait pas du choix de carte F effectué.

ThoIntIndepF

Théorème 20.16.
Soient $F\colon W \to F(W)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, deux cartes de la variété M. Soit une partie $A \subset F(W) \cap \tilde{F}(\tilde{W})$ telle que $A = F(B)$ avec $B \subset W$ mesurable. Alors $A = \tilde{F}(\tilde{B})$ avec $\tilde{B} \subset \tilde{W}$ mesurable.

Si f est une fonction continue, et si $\int_A f d\sigma_F$ existe, alors $\int_A f d\sigma_{\tilde{F}}$ existe et

$$\int_A f d\sigma_F = \int_A f d\sigma_{\tilde{F}}. \tag{20.50}$$

20.3.3 Exemples

Intégrons la fonction $f(x,y,z)$ sur le carré $K =]0,1[\times]0,2[\times \{1\}$. La première carte que nous pouvons utiliser est

$$\begin{aligned} F\colon]0,1[\times]0,2[&\to K \\ (x,y) &\mapsto (x,y,1). \end{aligned} \tag{20.51}$$

Nous trouvons aisément les vecteurs tangents qui forment l'élément de surface :

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{20.52}$$

donc $d\sigma_F = 1 \cdot dxdy$, et

$$\int_K f d\sigma_F = \int_{]0,1[\times]0,2[} f(x,y,1) \cdot 1 \cdot dxdy. \tag{20.53}$$

IntKSurcarUn

Nous pouvons également utiliser la carte

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,6[&\to K \\ (\tilde{x},\tilde{y}) &\mapsto (2\tilde{x},\frac{\tilde{y}}{3},1). \end{aligned} \tag{20.54}$$

Les vecteurs tangents sont maintenant

$$\frac{\partial \tilde{F}}{\partial \tilde{x}} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial \tilde{F}}{\partial \tilde{y}} = \begin{pmatrix} 0 \\ 1/3 \\ 0 \end{pmatrix}, \tag{20.55}$$

et nous avons donc $d\sigma_{\tilde{F}} = \|\frac{2}{3}e_3\| = \frac{2}{3}$. Cette fois, l'intégrale de f sur K s'écrit

$$\int_K f d\sigma_{\tilde{F}} = \int_{]0,\frac{1}{2}[\times]0,6[} f\big(2\tilde{x},\frac{\tilde{y}}{3},1\big) \cdot \frac{2}{3} \cdot d\tilde{x}d\tilde{y}. \tag{20.56}$$

Conformément au théorème 20.16, cette dernière intégrale est égale à l'intégrale (20.53) parce qu'il s'agit juste d'un changement de variable.

20.3.4 Orientation

Soient $F\colon W \to F(W)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, deux cartes de la variété M. Nous pouvons considérer la fonction $h = \tilde{F}^{-1} \circ F$, définie uniquement sur l'intersection des cartes :

$$h\colon F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big) \to \tilde{F}^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big). \tag{20.57}$$

Nous disons que F et $\tilde{F}$ ont même **orientation** si

$$J_h(w) > 0 \tag{20.58}$$

pour tout $w \in F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big)$.

Considérons les deux cartes suivantes pour le même carré :

$$\begin{aligned} F\colon]0,1[\times]0,1[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (x,y,0) \end{aligned} \tag{20.59}$$

et

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,\frac{1}{3}[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (2x,3y,0) \end{aligned} \tag{20.60}$$

Ici, $h(x,y) = \left(\frac{x}{2}, \frac{y}{3}\right)$ et nous avons $J_h = \frac{1}{6} > 0$. Ces deux cartes ont même orientation. Notez que

$$\frac{\partial F}{\partial x} \times \frac{\partial F}{\partial y} = e_3, \tag{20.61}$$

tandis que

$$\frac{\partial \tilde{F}}{\partial x} \times \frac{\partial \tilde{F}}{\partial y} = 6e_3. \tag{20.62}$$

Les vecteurs normaux à le paramétrage pointent dans le même sens.

Si par contre nous prenons le paramétrage

$$\begin{aligned} G\colon]0,1[\times]0,1[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (x,(1-y),0), \end{aligned} \tag{20.63}$$

nous avons

$$\frac{\partial G}{\partial x} \times \frac{\partial G}{\partial y} = -e_3, \tag{20.64}$$

et si $g = G^{-1} \circ F$, alors $J_g = -1$.

L'orientation d'une carte montre donc si le vecteur normal à la surface pointe d'un côté ou de l'autre de la surface.

DEFooSWREooNdQpdA

Définition 20.17 (Variété orientable)**.**
*Une variété M est **orientable** si il existe un atlas de M tel que deux cartes quelconques ont toujours même orientation. Une variété est **orientée** lorsque qu'un tel choix d'atlas est fait.*

Proposition 20.18.
Soit M, une variété orientable et un atlas orienté $\{F_i\colon W_i \to \mathbb{R}^3\}$. Alors le vecteur unitaire

$$\frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\left\| \frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) \right\|} \tag{20.65}$$

ne dépend pas du choix de F parmi les F_i.

Démonstration. Considérons deux cartes F_1 et F_2, ainsi que l'application $h = F_2^{-1} \circ F_1$. Écrivons le vecteur $\partial_x F_1 \times \partial_y F_1$ en utilisant $F_1 = F_2 \circ h$. D'abord, par la règle de dérivation de fonctions composées,

$$\frac{\partial (F_2 \circ h)}{\partial x} = \frac{\partial F_2}{\partial x}\frac{\partial h_1}{\partial x} + \frac{\partial F_2}{\partial y}\frac{\partial h_2}{\partial x}. \tag{20.66}$$

Après avoir fait le même calcul pour $\frac{\partial (F_2 \circ h)}{\partial y}$, nous pouvons écrire

$$\partial_x(F_2 \circ h) \times \partial_y(F_2 \circ h) = (\partial_x h_1 \partial_x F_2 + \partial_x h_2 \partial_y F_2) \times (\partial_y h_1 \partial_x F_2 + \partial_y h_2 \partial_y F_2). \tag{20.67}$$

Dans cette expression, les facteurs $\partial_i h_j$ sont des nombres, donc ils se factorisent dans les produits vectoriels. En tenant compte du fait que $\partial_x F_2 \times \partial_x F_2 = 0$ et $\partial_y F_2 \times \partial_y F_2 = 0$, ainsi que de l'antisymétrie du produit vectoriel, l'expression se réduit à

$$\left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y} \right) (\partial_x h_1 \partial_y h_2 - \partial_x h_2 \partial_y h_2). \tag{20.68}$$

Par conséquent,

$$\frac{\partial F_1}{\partial x} \times \frac{\partial F_1}{\partial y} = \frac{\partial (F_2 \circ h)}{\partial x} \times \frac{\partial (F_2 \circ h)}{\partial y} = \left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y} \right) \det J_h. \tag{20.69}$$

Donc, tant que J_h est positif, les vecteurs unitaires correspondants au membre de gauche et de droite sont égaux. □

Corolaire 20.19.
Si nous avons choisi un atlas orienté pour la variété M, nous avons une fonction continue $G\colon M \to \mathbb{R}^3$ telle que $\|G(p)\| = 1$ pour tout $p \in M$. Cette fonction est donnée par

$$G(F(x,y)) = \frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\|\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)\|} \tag{20.70}$$

DefCarteGOritn

sur l'image de la carte F.

Démonstration. La fonction G est construite indépendamment sur chaque carte $F(W)$ en utilisant la formule (20.70). Cette fonction est une fonction bien définie sur tout M parce que nous venons de démontrer que sur $F_1(W_1) \cap F_2(W_2)$, les fonctions construites à partir de F_1 et à partir de F_2 sont égales. □

Il est possible que prouver, bien que cela soit plus compliqué, que la réciproque est également vraie.

PROPooGHXTooPrAESF

Proposition 20.20.
Une variété M de dimension 2 dans $\mathbb{R}^3$ est orientable si et seulement si il existe une fonction continue $G\colon M \to \mathbb{R}^3$ telle que pour tout $p \in M$, le vecteur $G(p)$ soit de norme 1 et normal à M au point p.

20.4 Longueur, aire, volumes etc.

Grâce à la mesure de Lebesgue (définition 14.249), nous avons la définition d'aires dans $\mathbb{R}^2$ et de volumes dans $\mathbb{R}^3$. Dans tout ce qui suit, nous considérons toujours la tribu de Lebesgue, la mesure de Lebesgue, et l'intégrale de Lebesgue correspondante.

DEFooPZRDooWbbBXy

Définition 20.21.
*L'**aire** de la partie mesurable S de $\mathbb{R}^2$ est le nombre*

$$\int_{\mathbb{R}^2} \mathbb{1}_S \tag{20.71}$$

au sens de l'intégrale pour la mesure de Lebesgue de la fonction caractéristique de la partie S.

Définition 20.22.
*Le **volume** de la partie mesurable V de $\mathbb{R}^3$ est le nombre*

$$\int_{\mathbb{R}^3} \mathbb{1}_V \tag{20.72}$$

au sens de l'intégrale pour la mesure de Lebesgue de la fonction caractéristique de la partie V.

Ceci est bien beau, mais ne permet pas de définir l'aire d'une surface dans $\mathbb{R}^3$, ni une longueur dans $\mathbb{R}^2$. Nous n'avons pas encore défini ce que nous appelons une surface dans $\mathbb{R}^3$, mais selon toute définition raisonnable, si S en est une, elle sera négligeable au sens de la mesure de Lebesgue dans $\mathbb{R}^3$ et nous aurons toujours

$$\int_{\mathbb{R}^3} \mathbb{1}_S = 0. \tag{20.73}$$

L'objet de ce chapitre sera de donner un sens aux notions de longueurs dans $\mathbb{R}^2$, de surfaces dans $\mathbb{R}^3$ et d'y définir des intégrales permettant de définir longueurs et aires.

20.4.1 Quelques aires faciles

Nous nous souvenons de la proposition 10.37 qui donnait une bonne propriété du produit vectoriel dans $\mathbb{R}^3$. Nous en donnons une autre.

LEMooVHGKooDjcfOL

Lemme 20.23.
L'aire d'une droite dans $\mathbb{R}^2$ est nulle.

Démonstration. Nous considérons la mesure le Lebesgue λ sur $\mathbb{R}$ et une droite A dans $\mathbb{R}^2$.

La mesure dans $\mathbb{R}^2$ est donné par la définition 14.249 qui demande de calculer les intégrales (14.754) :

$$(\lambda\otimes\lambda)(A)=\int_{\mathbb{R}}\lambda\big(A_2(x)\big)d\lambda(x)=\int_{\mathbb{R}}\lambda\big(A_1(y)\big)d\lambda(y) \tag{20.74}$$

EQooUTAHooUyESco

où $A_2(x)=\{y\in\mathbb{R}\text{ tel que }(x,y)\in A\}$ et $A_1(y)=\{x\in\mathbb{R}\text{ tel que }(x,y)\in A\}$.

En vertu du lemme 12.152, cette droite est soit d'équation $y=ax+b$, soit d'équation $x=a$.

(1) **Droite $y=ax+b$** Dans ce cas, $A_2(x)=\{ax+b\}$. C'est un singleton et nous avons $\lambda\big(A_2(x)\big)=0$ pour tout x. Nous avons donc $(\lambda\otimes\lambda)(A)=0$.

(2) **Droite $x=a$** Nous utilisons la seconde possibilité laissée par l'égalité (20.74). L'ensemble A_1 est facile à déterminer : $A_1(y)=\{a\}$ pour tout y. Donc $\lambda\big(A_1(y)\big)=0$ pour tout y, et l'intégrale est nulle.

□

Remarque 20.24.
Dans la preuve de 20.23, nous aurions pu faire le cas $x=a$ en utilisant la première formule. Dans ce cas nous serions partis de

$$A_2(x)=\begin{cases}\mathbb{R} & \text{si } x=a\\ \varnothing & \text{sinon,}\end{cases} \tag{20.75}$$

et donc de

$$\lambda\big(A_2(x)\big)=\begin{cases}\infty & \text{si } x=a\\ 0 & \text{sinon.}\end{cases} \tag{20.76}$$

Le lemme 14.172 nous dit que l'intégrale de A_2 sur $\mathbb{R}$ est nulle.

DEFooDTFCooCTdaDO

Définition 20.25.
*Soient a, u_1 et u_2 dans $\mathbb{R}^2$. Le **parallélogramme** basé en a formé sur u_1 et u_2 est l'ensemble*

$$\{a+su_1+tu_2\text{ tel que }0\leqslant t\leqslant 1,0\leqslant s\leqslant 1\}. \tag{20.77}$$

Assez souvent, nous supposeront que u_1 et u_2 ne sont pas colinéaires.

PROPooAVVNooOOlSzr

Proposition 20.26.
L'aire[5] *du parallélogramme*[6] *basé en a et formé sur u et v (que nous supposons n'être pas colinéaires) est donnée par*

$$\|u\times v\|=|u_1v_2-v_1u_2|. \tag{20.78}$$

Démonstration. En vertu de la définition 20.21 d'une aire, le nombre à calculer est $\int_D 1d\lambda$ où

$$D_0=\{a+xu+yv\text{ tel que }x\in[0,1],y\in[0,1]\}. \tag{20.79}$$

En posant

$$D=\{a+xu+yv\text{ tel que }x\in\,]0,1[,y\in\,]0,1[\}, \tag{20.80}$$

nous avons

$$\int_{D_0}d\lambda=\int_D d\lambda, \tag{20.81}$$

parce que ce que nous avons enlevé sont des segments de droites alors que les droites sont de mesure nulle (lemme 20.23).

5. Définition 20.21.
6. Définition 20.25.

Nous considérons la paramétrisation

$$\begin{aligned}\varphi\colon]0,1[\times]0,1[&\to D\\ (x,y) &\mapsto a+xu+yv.\end{aligned}\tag{20.82}$$

Pour le théorème 14.281, nous devons montrer que ϕ est un C^1-difféomorphisme. Comme nous sommes un peu fatigués, nous allons seulement prouver ϕ est injective [7]. Supposons que $\phi(x_1,x_2)=\phi(y_1,y_2)$. Alors

$$(x_1-x_2)u+(y_1-y_2)v=0.\tag{20.83}$$

Si $y_1-y_2=0$, alors $x_1-x_2=0$ et on est bon. Sinon,

$$v=-\frac{x_1-x_2}{y_1-y_2}u\tag{20.84}$$

Vu que les vecteurs u et v ne sont pas colinéaires, nous en déduisons que $x_1-x_2=0$ et on est encore bon.

Nous utilisons la formule (14.855) de changement de variable avec la fonction ϕ. Le jacobien à calculer est

$$J=\det\begin{pmatrix}(\partial_x\phi_1)(x,y) & (\partial_x\phi_2)(x,y)\\ (\partial_y\phi_1)(x,y) & (\partial_y\phi_2)(x,y)\end{pmatrix}=\det\begin{pmatrix}u_1 & u_2\\ v_1 & v_2\end{pmatrix}=|u_1v_2-v_1u_2|.\tag{20.85}$$

Le lien avec le produit vectoriel est un petit abus de notation : il s'agit du produit vectoriel entre $(u1,u2,0)$ et $(v_1,v_2,0)$ qui peut être obtenu en utilisant directement la formule de définition (10.78). □

20.5 Autres théorèmes de points fixes

SECooDWMPooWZgzRZ

En termes de théorème de points fixes nous avons déjà vu le théorème de Picard 17.34. Voir aussi le thème 65.

20.5.1 Brouwer

subSecZCCmMnQ

Proposition 20.27.
Soit $f\colon[a,b]\to[a,b]$ une fonction continue. Alors f accepte un point fixe.

Démonstration. En effet si nous considérons $g(x)=f(x)-x$ alors nous avons $g(a)=f(a)-a\geqslant 0$ et $g(b)=f(b)-b\leqslant 0$. Si $g(a)$ ou $g(b)$ est nul, la proposition est démontrée ; nous supposons donc que $g(a)>0$ et $g(b)<0$. La proposition découle à présent du théorème des valeurs intermédiaires 9.93. □

Exemple 20.28.
La fonction $x\mapsto\cos(x)$ est continue entre $[-1,1]$ et $[-1,1]$. Elle admet donc un point fixe. Par conséquent il existe (au moins) une solution à l'équation $\cos(x)=x$. △

PROPooBRZCooLvscEf

Proposition 20.29 (Théorème de rétraction[583])**.**
Il n'existe pas d'applications $g\colon\overline{B(0,1)B(0,1)}\to$ qui soit de classe C^1 et qui soit l'identité sur $\partial B(0,1)$.

Démonstration. En plein d'étapes.

(1) **Définition de l'application F** Nous définissons la fonction

$$\begin{aligned}F\colon\overline{B(0,1)}\times[0,1] &\to \overline{B(0,1)}\\ (x,t) &\mapsto (1-t)x+tg(x).\end{aligned}\tag{20.86}$$

7. Pour le reste, écrivez l'inverse explicitement, et prouvez que c'est C^1. Si ça pose problème, écrivez-moi parce que je n'ai pas vérifié.

Lorsque t est fixé, nous notons

$$\begin{aligned} F_t\colon \overline{B(0,1)} &\to \overline{B(0,1)} \\ x &\mapsto F(x,t). \end{aligned} \tag{20.87}$$

Nous avons $F_0 = \mathrm{Id}|_{\overline{B(0,1)}}$ et $F_1 = g$. De plus, nous avons $F(x,t) = x$ pour tout $x \in \partial B(0,1)$ et pour tout $t \in [0,1]$.

(2) **définition de V** Nous posons[8]

$$\begin{aligned} V\colon [0,1] &\to \mathbb{R} \\ t &\mapsto \int_{B(0,1)} \det(dF_t)_x dx. \end{aligned} \tag{20.88}$$

En ce qui concerne la différentielle de F_t, nous avons $F_t(x) = (1-t)x + tg(x)$. La différentielle de $x \mapsto x$ est $v \mapsto v$ (lemma 10.269), et donc

$$(dF_t)_x = (1-t)\,\mathrm{Id} + t dg_x. \tag{20.89}$$

Le déterminant de cela est un polynôme en t.

(3) $V(1) = 0$ Vu que $g(x) \in \partial B(0,1)$, nous avons $g(x) \cdot g(x) = \|g(x)\|^2 = 1$. En utilisant la proposition 10.283, pour tout $h \in \mathbb{R}^n$ nous avons :

$$2g(x) \cdot (dg_x)(h) = 0. \tag{20.90}$$

Autrement dit, l'image de dg_x est dans $g(x)^\perp$, qui est de dimension plus petite que n (parce que $g(x) \neq 0$). Donc dg_x n'est pas de rang maximum ; bref : son déterminant est nul.
Vu que $F_1(x) = F(1,x) = g(x)$, nous avons bien $V(1) = 0$.

(4) $\det\Big((dF_t)_x\Big) > 0$ **si t est assez petit**

Supposons le contraire : pour tout $n \in \mathbb{N}$, il existe $t_n \in]0, 1/n]$ et $x_n \in \overline{B(0,1)}$ tels que $\det\big((dF_{t_n})_{x_n}\big) \leqslant 0$. La suite (x_n) étant dans le compact $\overline{B(0,1)}$, elle admet une sous-suite convergente : il existe une application strictement croissante $\phi\colon \mathbb{N} \to \mathbb{N}$ et un $x \in \overline{B(0,1)}$ tels que $x_{\phi(n)} \to x$. Notez au passage que $t_{\phi(n)} \to 0$.
Pour tout n nous avons

$$\det\Big(\big(dF_{t_{\phi(n)}}\big)_{x_{\phi(n)}}\Big) \leqslant 0. \tag{20.91}$$

EqooABSWooUErEXU

Vu que l'application $x \mapsto \det\big((dF_t)_x\big)$ est continue en t et en x, nous pouvons passer à la limite dans (20.91) :

$$\det\Big((dF_0)_x\Big) \leqslant 0. \tag{20.92}$$

Mais comme $dF_0 = \mathrm{Id}$, nous avons en réalité $\det\big((dF_0)_x\big) > 0$. Contradiction. Nous en déduisons que il existe $m_1 > 0$ tel que pour tout $t < m_1$ et pour tout $x \in \overline{B(0,1)}$, nous avons

$$\det\Big((dF_t)_x\Big) > 0. \tag{20.93}$$

(5) **F_t est injective si t est assez petit**

Soient $t < 1/N$ et $x, y \in \overline{B(0,1)}$ tels que $F(x,t) = F(y,t)$. Nous avons

$$0 = F(x,t) - F(y,t) = (1-t)(x-y) + t\big(g(x) - g(y)\big). \tag{20.94}$$

En passant aux normes et en utilisant le théorème des accroissements finis 10.292,

$$(1-t)\|x-y\| = t\|g(x) - g(y)\| \leqslant tM\|x-y\| \tag{20.95}$$

8. L'intégrale est définie en (14.539), et la mesure de Lebesgue en 14.249. Mais en principe, ça fait longtemps que vous êtes à l'aise avec ça.

où $M = \sup_{x\in\overline{B(0,1)}} \|dg_x\|$. En simplifiant par $\|x-y\|$, nous trouvons $tM/(1-t) \geqslant 1$ ou encore

$$t \geqslant \frac{1}{M+1}. \tag{20.96}$$

Autrement dit, dès que $t < 1/(M+1)$, l'application F_t est injective.

(6) t **est assez petit**
Nous posons

$$\alpha = \min\{m_1, \frac{1}{M+1}\}. \tag{20.97}$$

Dès que $t < \alpha$,

— F_t est injective,

— $\det\Big((dF_t)_x\Big) > 0$.

Cela étant dit, nous continuons la preuve.

(7) **Un peu de relations ensemblistes**
Nous considérons $t < \alpha$ et $x \in \partial B(0,1)$. Nous avons alors $F(x,t) = x$. Nous avons alors

$$F\big(\partial B(0,1),t\big) \cup B(0,1) = \overline{B(0,1)}, \tag{20.98}$$

qui est fermé dans $\mathbb{R}^n$.
Vu que F_t est injective, nous pouvons utiliser la formule ensembliste (1.8) : SUBEQSooLWJLooBEVkro

$$\begin{aligned} F\big(B(0,1),t\big) &= F\big(\overline{B(0,1)},t\big)\backslash F\big(\partial B(0,1),t\big) && (20.99a)\\ &= F\big(\overline{B(0,1)},t\big)\backslash \partial B(0,1) && (20.99b)\\ &= F\big(\overline{B(0,1)},t\big) \cap B(0,1), && (20.99c)\end{aligned}$$

parce que $X\backslash\partial B(0,1) = X \cap B(0,1)$.

(8) $F\big(B(0,1),t\big)$ **est ouvert**
L'application F est de classe C^1 et injective. Donc $(dF_t)_x$ est continue et injective (proposition 10.256). Soit $x \in B(0,1)$. Par le théorème d'inversion locale 17.47 il existe un ouvert $V_x \subset B(0,1)$ et un ouvert W_x de $\mathbb{R}^n$ tels que $F_t\colon V_x \to W_x$ soit une bijection. Nous avons donc

$$F_t\big(B(0,1)\big) = \bigcup_{x\in B(0,1)} F_t(V_x) = \bigcup_{x\in B(0,1)} W_x. \tag{20.100}$$ EQooDCMAooVJnhxy

Une union d'ouverts étant ouverte, nous en déduisons que $F\big(B(0,1),t\big)$ est ouvert.

(9) **Un peu de topologie induite** Nous considérons l'espace $B(0,1)$ avec la topologie induite de $\mathbb{R}^n$.

(9a) **C'est fermé**
La partie $F\big(\overline{B(0,1)},t\big)$ est l'image d'un compact par une fonction continue. C'est donc un compact (et donc fermé) par le théorème 7.220. Du coup la proposition 7.29 dit que la partie

$$F\big(B(0,1),t\big) = F\big(\overline{B(0,1)},t\big) \cap B(0,1) \tag{20.101}$$

est fermée dans $B(0,1)$.

(9b) **C'est ouvert**
La partie $F\big(B(0,1),t\big)$ est ouverte (voir (20.100)) et contenue dans $B(0,1)$ (voir (20.99)). Elle est donc un ouvert pour la topologie de $B(0,1)$.

(9c) **Et donc c'est ...**
La partie $F\big(B(0,1),t\big)$ est ouvert et fermée dans $B(0,1)$. Par connexité, la proposition 7.74(4) conclu que

$$F\big(B(0,1),t\big) = B(0,1). \tag{20.102}$$

(10) **Volume de la boule**

Nous sommes toujours avec $t < \alpha$. Nous avons donc $\det\big((dF_t)_x\big) > 0$ pour tout x. Nous calculons le volume de la boule $B(0,1)$ en introduisant la fonction constante 1. En utilisant le changement de variables[9] $F_t\colon B(0,1) \to B(0,1)$:

$$\int_{B(0,1)} 1dx = \int_{F_t^{-1}\left(B(0,1)\right)} \underbrace{(1\circ F_t)(x)}_{=1} |\det\big((dF_t)_x\big)|dx \tag{20.103a}$$

$$= \int_{B(0,1)} \det\big((dF_t)_x\big)dx \tag{20.103b}$$

$$= V(t). \tag{20.103c}$$

Nous venons de prouver que sur $t \in [0,\alpha[$, la fonction $V(t)$ est en réalité une constante égale au volume de la boule unité.

(11) **Et enfin...**

Vu que V est un polynôme et que V est constant sur un intervalle, V est constante partout. En particulier, $V(1)$ est encore égal au volume de le boule. Ah ben non parce que nous avions démontré plus haut que $V(1) = 0$.

Contradiction, fin de l'histoire.

□

PropDRpYwv

Proposition 20.30 (Brouwer dans $\mathbb{R}^n$ version C^∞)**.**
Toute application $f\colon \overline{B(0,1)} \to \overline{B(0,1)}$ de classe C^∞ admet un point fixe.

Démonstration. Supposons que f ne possède pas de points fixes. Alors pour tout $x \in \overline{B(0,1)}$ nous considérons la demi-droite issue de $f(x)$ et passant par x (cette demi-droite existe parce que x et $f(x)$ sont supposés distincts). Cette demi-droite intersecte $\partial B(0,1)$ en un point que nous appelons $g(x)$. Prouvons que cette fonction est C^k dès que f est C^k (y compris avec $k = \infty$).

L'application

$$g\colon \overline{B(0,1)} \to \partial B(0,1) \tag{20.104}$$

est construite comme solution du système

$$\begin{cases} g(x) - f(x) = \lambda\big(x - f(x)\big) & \text{(20.105a)}\\ \|g(x)\|^2 = 1 & \text{(20.105b)}\\ \lambda \geqslant 0. & \text{(20.105c)} \end{cases}$$

En substituant nous obtenons l'équation

$$P_x(\lambda) = \|\lambda\big(x - f(x)\big) + f(x)\|^2 - 1 = 0, \tag{20.106}$$

ou encore

$$\lambda^2\|x - f(x)\|^2 + 2\lambda\big(x - f(x)\big) \cdot f(x) + \|f(x)\|^2 - 1 = 0. \tag{20.107}$$

En tenant compte du fait que $\|f(x)\| < 1$ (parce que les images de f sont dans $\mathcal{B}$), nous trouvons que $P_x(0) \leqslant 0$ et $P_x(1) \leqslant 0$. De même $\lim_{\lambda\to\infty} P_x(\lambda) = +\infty$. Par conséquent le polynôme de second degré P_x a exactement deux racines distinctes $\lambda_1 \leqslant 0$ et $\lambda_2 \geqslant 1$. La racine que nous cherchons est la seconde. Le discriminant est strictement positif, donc pas besoin d'avoir peur de la racine dans

$$\lambda(x) = \frac{-\big(x - f(x)\big) \cdot f(x) + \sqrt{\Delta_x}}{\|x - f(x)\|^2} \tag{20.108}$$

où

$$\Delta_x = 4\Big(\big(x - f(x)\big) \cdot f(x)\Big)^2 - 4\|x - f(x)\|^2\big(\|f(x)\|^2 - 1\big). \tag{20.109}$$

Notons que la fonction $\lambda(x)$ est C^k dès que f est C^k ; et en particulier elle est C^∞ si f l'est.

En résumé la fonction g ainsi définie vérifie deux propriétés :

9. Théorème 14.281.

(1) $g\colon \overline{B(0,1)} \to \overline{B(0,1)}$
(2) elle est C^∞ ;
(3) elle est l'identité sur $\partial B(0,1)$.

Cela viole le théorème de contraction 20.29. Contradiction, fin de la preuve. □

20.31.
Voici comment on démontre le théorème de rétraction dans le cas C^∞ en utilisant Stokes.

La suite de la preuve consiste à montrer qu'une telle application $g\colon B \to \partial B$ ne peut pas exister [10]. Nous considérons une forme de volume ω sur ∂B : l'intégrale de ω sur ∂B est l'aire de ∂B qui est non nulle. Nous avons alors la contradiction suivante :

$$0 < \int_{\partial B} \omega = \int_{\partial B} g^*\omega = \int_B d(g^*\omega) = \int_B g^*(d\omega) = 0 \tag{20.110}$$

Justifications :

— L'intégrale $\int_{\partial B} \omega$ est l'aire de ∂B et est donc strictement positive.
— La fonction g est l'identité sur ∂B. Nous avons donc $\omega = g^*\omega$.
— Le lemme 27.82.
— La forme ω est de volume, par conséquent de degré maximum et $d\omega = 0$.

Un des points délicats pour démontrer le théorème de Brouwer en version continue est de se ramener au cas de fonctions C^∞. Pour la régularisation par convolution, voir [584] ; pour celle utilisant le théorème de Weierstrass, voir [585].

ThoRGjGdO

Théorème 20.32 (Brouwer dans $\mathbb{R}^n$, version continue)**.**
Toute application continue [11] $\overline{B(0,1)} \to \overline{B(0,1)}$ *admet un point fixe.*

Démonstration. Soit une application continue $f\colon \overline{B(0,1)} \to \overline{B(0,1)}$. Nous commençons par définir une suite de fonctions

$$f_k(x) = \frac{f(x)}{1 + \frac{1}{k}}. \tag{20.111}$$

Nous avons $\|f_k - f\|_\infty \leqslant \frac{1}{1+k}$ où la norme est la norme uniforme sur B. Par le théorème de Weierstrass 12.419 il existe une suite de fonctions $C^\infty(B, \mathbb{R})$ que nous nommons g_k telles que

$$\|g_k - f_k\|_\infty \leqslant \frac{1}{1+k}. \tag{20.112}$$

Vérifions que cette fonction g_k soit bien une fonction qui prend ses valeurs dans B :

$$\|g_k(x)\| \leqslant \|g_k(x) - f_k(x)\| + \|f_k(x)\| \tag{20.113a}$$
$$\leqslant \frac{1}{1+k} + \frac{\|f(x)\|}{1 + \frac{1}{k}} \tag{20.113b}$$
$$\leqslant \frac{1}{1+k} + \frac{1}{1 + \frac{1}{k}} \tag{20.113c}$$
$$= 1. \tag{20.113d}$$

Par la version C^∞ du théorème (proposition 20.30), g_k admet un point fixe que l'on nomme x_k.

10. Notons qu'il n'existe pas non plus de rétractions continues sur B, mais pour le montrer il faut utiliser d'autres méthodes que Stokes, ou alors présenter les choses dans un autre ordre.
11. Une fonction continue sur un fermé de $\mathbb{R}^n$ est à comprendre pour la topologie induite.

Étant donné que x_k est dans le compact B, quitte à prendre une sous-suite nous supposons que la suite (x_k) converge vers un élément $x \in B$. Nous montrons maintenant que x est un point fixe de f :

$$\|f(x) - x\| = \|f(x) - g_k(x) + g_k(x) - x_k + x_k - x\| \tag{20.114a}$$
$$\leqslant \|f(x) - g_k(x)\| + \underbrace{\|g_k(x) - x_k\|}_{=0} + \|x_k - x\| \tag{20.114b}$$
$$\leqslant \frac{1}{1+k} + \|x_k - x\|. \tag{20.114c}$$

En prenant le limite $k \to \infty$ le membre de droite tend vers zéro et nous obtenons $f(x) = x$. □

20.5.2 Théorème de Schauder

Une conséquence du théorème de Brouwer est le théorème de Schauder qui est valide en dimension infinie.

ThovHJXIU

Théorème 20.33 (Théorème de Schauder[586])**.**
Soit E, un espace vectoriel normé, K un convexe compact de E et $f\colon K \to K$ une fonction continue. Alors f admet un point fixe.

Démonstration. Étant donné que $f\colon K \to K$ est continue, elle y est uniformément continue. Si nous choisissons ϵ alors il existe $\delta > 0$ tel que

$$\|f(x) - f(y)\| \leqslant \epsilon \tag{20.115}$$

dès que $\|x - y\| \leqslant \delta$. La compacité de K permet de choisir un recouvrement fini par des ouverts de la forme

$$K \subset \bigcup_{1 \leqslant i \leqslant p} B(x_j, \delta) \tag{20.116}$$

EqKNPUVR

où $\{x_1, \ldots, x_p\} \subset K$. Nous considérons maintenant $L = \text{Span}\{f(x_j) \text{ tel que } 1 \leqslant j \leqslant p\}$ et

$$K^* = K \cap L. \tag{20.117}$$

Le fait que K et L soient convexes implique que K^* est convexe. L'ensemble K^* est également compact parce qu'il s'agit d'une partie fermée de K qui est compact (lemme 7.96). Notons en particulier que K^* est contenu dans un espace vectoriel de dimension finie, ce qui n'est pas le cas de K.

Nous allons à présent construire une sorte de partition de l'unité subordonnée au recouvrement (20.116) sur K (voir le théorème 27.75). Nous commençons par définir

$$\psi_j(x) = \begin{cases} 0 & \text{si } \|x - x_j\| \geqslant \delta \\ 1 - \frac{\|x - x_j\|}{\delta} & \text{sinon.} \end{cases} \tag{20.118}$$

pour chaque $1 \leqslant j \leqslant p$. Notons que ψ_j est une fonction positive, nulle en-dehors de $B(x_j, \delta)$. En particulier la fonction suivante est bien définie :

$$\varphi_j(x) = \frac{\psi_j(x)}{\sum_{k=1}^{p} \psi_k(x)} \tag{20.119}$$

et nous avons $\sum_{j=1}^{p} \varphi_j(x) = 1$. Les fonctions φ_j sont continues sur K et nous définissons finalement

$$g(x) = \sum_{j=1}^{p} \varphi_j(x) f(x_j). \tag{20.120}$$

Pour chaque $x \in K$, l'élément $g(x)$ est une combinaison des éléments $f(x_j) \in K^*$. Étant donné que K^* est convexe et que la somme des coefficients $\varphi_j(x)$ vaut un, nous avons que g prend ses valeurs dans K^* par la proposition 11.29.

Nous considérons seulement la restriction $g\colon K^* \to K^*$ qui est continue sur un compact contenu dans un espace vectoriel de dimension finie. Le théorème de Brouwer 20.32 nous enseigne alors que g a un point fixe. Nous nommons y ce point fixe. Notons que y est fonction du ϵ choisi au début de la construction, via le δ qui avait conditionné la partition de l'unité.

Nous avons [EqoXuTzE]

$$\begin{aligned} f(y) - y &= f(y) - g(y) && \text{(20.121a)}\\ &= \sum_{j=1}^{p} \varphi_j(y) f(y) - \sum_{j=1}^{p} \varphi_j(y) f(x_j) && \text{(20.121b)}\\ &= \sum_{j=1}^{p} \varphi_j(y)\big(f(y) - f(x_j)\big). && \text{(20.121c)} \end{aligned}$$

Par construction, $\varphi_j(y) \neq 0$ seulement si $\|y - x_j\| \leqslant \delta$ et par conséquent seulement si $\|f(y) - f(x_j)\| \leqslant \epsilon$. D'autre part nous avons $\varphi_j(y) \geqslant 0$; en prenant la norme de (20.121) nous trouvons

$$\|f(y) - y\| \leqslant \sum_{j=1}^{p} \|\varphi_j(y)\big(f(y) - f(x_j)\big)\| \leqslant \sum_{j=1}^{p} \varphi_j(y)\epsilon = \epsilon. \tag{20.122}$$

Nous nous souvenons maintenant que y était fonction de ϵ. Soit y_m le y qui correspond à $\epsilon = 2^{-m}$. Nous avons alors

$$\|f(y_m) - y_m\| \leqslant 2^{-m}. \tag{20.123}$$

L'élément y_m est dans K^* qui est compact, donc quitte à choisir une sous-suite nous pouvons supposer que y_m est une suite qui converge vers $y^* \in K$ [12]. Nous avons les majorations

$$\|f(y^*) - y^*\| \leqslant \|f(y^*) - f(y_m)\| + \|f(y_m) - y_m\| + \|y_m - y^*\|. \tag{20.124}$$

Si m est assez grand, les trois termes du membre de droite peuvent être rendus arbitrairement petits, d'où nous concluons que

$$f(y^*) = y^* \tag{20.125}$$

et donc que f possède un point fixe. □

20.5.3 Théorème de Cauchy-Arzella

[ThoHNBooUipgPX]

Théorème 20.34 (Cauchy-Arzela[504]).
Nous considérons le système d'équation différentielles [EqTXlJdH]

$$\begin{cases} y' = f(t, y) & \text{(20.126a)}\\ y(t_0) = y_0. & \text{(20.126b)} \end{cases}$$

avec $f\colon U \to \mathbb{R}^n$, continue où U est ouvert dans $\mathbb{R} \times \mathbb{R}^n$. Alors il existe un voisinage fermé V de t_0 sur lequel une solution C^1 du problème (20.126) existe.

Idée de la démonstration. Nous considérons $M = \|f\|_\infty$ et K, l'ensemble des fonctions M-Lipschitz sur U. Nous prouvons que $(K, \|.\|_\infty)$ est compact. Ensuite nous considérons l'application

$$\begin{aligned} &\Phi\colon K \to K\\ &\Phi(f)(t) = x_0 + \int_{t_0}^{t} f\big(u, f(u)\big)du. \end{aligned} \tag{20.127}$$

Après avoir prouvé que Φ était continue, nous concluons qu'elle a un point fixe par le théorème de Schauder 20.33. □

12. Notons que même dans la sous-suite nous avons $\|f(y_m) - y_m\| \leqslant 2^{-m}$, avec le même « m » des deux côtés de l'inégalité.

Remarque 20.35.
Quelques remarques.

(1) Les théorème de Cauchy-Lipschitz et Cauchy-Arzella donnent des existences pour des équations différentielles du type $y' = f(t, y)$. Et si nous avons une équation du second ordre ? Alors il y a la méthode de la réduction de l'ordre qui permet de transformer une équation différentielle d'ordre élevé en un système d'ordre 1.

(2) Ces théorèmes posent des *conditions initiales* : la valeur de y est donnée en un point, et la méthode de la réduction de l'ordre permet de donner l'existence de solutions d'un problème d'ordre k en donnant les valeurs de $y(0)$, $y'(0)$, $\dots y^{(k-1)}(0)$. C'est-à-dire de la fonction et de ses dérivées en un point. Rien n'est dit sur l'existence de *conditions aux bords*.

Ces deux points sont illustrés dans les exemples 32.17 et 32.18.

20.5.4 Théorème de Markov-Kakutani

Le théorème de Markov-Kakutani, nous donne l'existence d'une mesure de Haar sur un groupe compact. Voir la définition 48.81 et la construction de la mesure de Haar, théorème 48.89.

ThoeJCdMP

Théorème 20.36 (Markov-Katutani[587])**.**
Soit E un espace vectoriel normé et L, une partie non vide, convexe, fermée et bornée de E'. Soit $T : L \to L$ une application continue. Alors T a un point fixe.

Démonstration. Nous considérons un point $x_0 \in L$ et la suite

$$x_n = \frac{1}{n+1} \sum_{i=0}^{n} T^i x_0. \tag{20.128}$$

La somme des coefficients devant les $T^i(x_0)$ étant 1, la convexité de L montre que $x_n \in L$. Nous considérons l'ensemble

$$C = \bigcap_{n \in \mathbb{N}} \overline{\{x_m \text{ tel que } m \geqslant n\}}. \tag{20.129}$$

Le lemme 7.301 indique que C n'est pas vide, et de plus il existe une sous-suite de (x_n) qui converge vers un élément $x \in C$. Nous avons

$$\lim_{n \to \infty} x_{\sigma(n)}(v) = x(v) \tag{20.130}$$

pour tout $v \in E$. Montrons que x est un point fixe de T. Nous avons

$$\|(Tx_{\sigma(k)} - x_{\sigma(k)})v\| = \Big\| T \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v) - \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v) \Big\| \tag{20.131a}$$

$$= \Big\| \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^{i+1} x_0(v) - T^i x_0(v) \Big\| \tag{20.131b}$$

$$= \frac{1}{1+\sigma(k)} \big\| T^{\sigma(k)+1} x_0(v) - x_0(v) \big\| \tag{20.131c}$$

$$\leqslant \frac{2M}{\sigma(k)+1} \tag{20.131d}$$

où $M = \sum_{y \in L} \|y(v)\| < \infty$ parce que L est borné. En prenant $k \to \infty$ nous trouvons

$$\lim_{k \to \infty} \big(Tx_{\sigma(k)} - x_{\sigma(k)} \big) v = 0, \tag{20.132}$$

ce qui signifie que $Tx = x$ parce que T est continue. □

20.6 Intégrales curvilignes

secintcurvi

20.6.1 Chemins de classe C^1

DEFooQTAWooOCmSJo

Définition 20.37 ([588]).
*Soit $p, q \in \mathbb{R}^n$. Un **chemin** C^1 **par morceaux** joignant p à q est une application continue*

$$\gamma : [a, b] \to \mathbb{R}^n \tag{20.133}$$

vérifiant :

(1) L'application $\gamma\colon [a, b] \to \mathbb{R}^n$ est continue.

(2) Les extrémités : $\gamma(a) = p$, $\gamma(b) = q$

(3) Il existe une subdivision $a = t_0 < t_1 < \ldots < t_{r-1} < t_r = b$ telle que la restriction de γ sur chaque fermé $[t_i, t_{i+1}]$ est de classe C^1.

Pour être d'accord sur les définitions, quand on dit qu'une application est C^1 sur un fermé, c'est qu'elle est C^1 sur ce fermé, avec la topologie de ce fermé (induite[13] de $\mathbb{R}$). Autrement dit, toutes les limites et tous les ouverts qui interviennent dans la notion de « être C^1 » sont à prendre dans la topologie du fermé $[a_i, b_i]$.

Remarque 20.38.
Si a et b sont des points de $\mathbb{R}^n$, on peut créer le chemin particulier

$$\gamma : [0, 1] \to \mathbb{R}^n : t \mapsto (1 - t)a + tb \tag{20.134}$$

qui relie ces points par un segment de droite.

Il existe des fonctions C^1 par morceaux qui ne sont pas C^1.

Proposition 20.39.
L'application

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} x & si\ x \leqslant 0 \\ 2x & si\ x > 0 \end{cases} \end{aligned} \tag{20.135}$$

est C^1 par morceaux.

Elle n'est pas de classe C^1.

Démonstration. Le fait que f ne soit pas de classe C^1 est immédiat parce que sa dérivée n'est pas définie en zéro.

Pour montrer que f est C^1 par morceaux, nous définissons les deux fonctions

$$\begin{aligned} f_1\colon]-\infty, 0] &\to \mathbb{R} \\ x &\mapsto x \end{aligned} \tag{20.136}$$

et

$$\begin{aligned} f_2\colon [0, \infty[&\to \mathbb{R} \\ x &\mapsto 2x. \end{aligned} \tag{20.137}$$

Nous calculons la dérivée de f_1 en $x = 0$. C'est par définition

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} \frac{f(x)}{x}. \tag{20.138}$$

Attention : cette limite est à comprendre au sens de la topologie de $]-\infty, 0]$. Nous allons montrer que cette limite vaut 1, en remontant aux définition pour le sport.

13. Topologie induite, définition 7.26.

Posons $g(x) = f(x)/x$. Si V est un voisinage de 1 dans $\mathbb{R}$, il existe une boule $B(1,r)$ qui contient V. Nous allons trouver un $\epsilon > 0$ tel que $g\big(B(0,\epsilon)\backslash\{0\}\big) \subset B(1,r)$. Attention : à gauche c'est une boule au sens de la topologie induite. Au sens de la topologie normale, nous cherchons $\epsilon > 0$ tel que

$$g\Big(\big(B(0,\epsilon) \cap]-\infty,0]\big)\backslash\{0\}\Big) \tag{20.139}$$

Bref, pour tout x dans cette boule, nous devons avoir $|g(x)-1| < r$. Vu que $x \in]-\infty,0]\backslash\{0\}$, nous avons $g(x) = 1$ pour tout x. Et voila, n'importe que ϵ fait l'affaire.

Nous avons prouvé que la dérivée de f_1 est donnée par

$$\begin{aligned} f_1' \colon]-\infty,0] &\to \mathbb{R} \\ x &\mapsto 1. \end{aligned} \tag{20.140}$$

Elle est continue. Nous laissons la démonstration de la continuité de f_1' à la lectrice ; attention : continuité pour la topologie induite, toujours. Bref, f_1 est de classe C^1.

Nous démontrons de même que f_2 est de classe C^2 et que $f_2'(x) = 2$ pour tout x. □

20.6.2 Intégrer une fonction

DEFooFAYUooCaUdyo

Définition 20.40.
Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}$ une fonction continue, et $\gamma : [a,b] \to D$ un chemin C^1. On définit ***l'intégrale de f sur γ*** *par*

$$\int_\gamma f ds = \int_\gamma f = \int_a^b f(\gamma(t)) \big\|\gamma'(t)\big\| \, dt. \tag{20.141}$$

EqhJGRcb

Note : dans le cadre de l'analyse complexe, ce n'est pas exactement cette définition. Voir 26.5.

Exemple 20.41.
Soit l'hélice

$$\begin{aligned} \sigma \colon [0,2\pi] &\to \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}, \end{aligned} \tag{20.142}$$

et la fonction $f(x,y,z) = x^2 + y^2 + z^2$. L'intégrale de f sur σ est

$$\begin{aligned} \int_\sigma f &= \int_0^{2\pi} (\cos^2 t + \sin^2 t + t^2) \|\sigma'(t)\| dt \\ &= \int_0^{2\pi} (1+t^2)\sqrt{2} dt \\ &= \sqrt{2}\left[t + \frac{t^3}{3}\right]_0^{2\pi} \\ &= \sqrt{2}\left(2\pi + \frac{8\pi^3}{3}\right). \end{aligned} \tag{20.143}$$

△

Remarque 20.42.
Si $f = 1$, alors nous tombons sur

$$\int_\gamma ds = \int_a^b \|\gamma'(t)\| dt, \tag{20.144}$$

Nous verrons par le théorème 21.9 que cette dernière intégrale est la longueur de la courbe. Il est un fait général que l'intégrale de la fonction 1 sur un ensemble en donne la « mesure ». Cela est à mettre en rapport avec le lemme 14.175 en gardant en tête que $\int_\gamma 1$ n'est pas la mesure de l'image de γ dans $\mathbb{R}^2$.

Proposition 20.43 (Indépendence en le paramétrage)**.**
La valeur de l'intégrale de f sur γ ne dépend pas du paramétrage (équivalent ou pas) choisi.

Démonstration. Soit donc un chemin $\gamma\colon [c,d]\to\mathbb{R}^3$ ainsi que $\varphi\colon [c,d]\to[a,b]$, un reparamétrage de classe C^1, strictement monotone et le chemin σ définit par $\gamma(s)=\sigma\big(\varphi(s)\big)$ avec $s\in[c,d]$. En supposant que $\varphi'(s)\geqslant 0$, nous avons

$$\begin{aligned} I=\int_\gamma f&=\int_c^d f\big(\gamma(s)\big)\|\gamma'(s)\|ds\\ &=\int_c^d f\Big(\sigma\big(\varphi(s)\big)\Big)\|\sigma'\big(\varphi(s)\big)\||\varphi'(s)|ds. \end{aligned} \tag{20.145}$$

Pour cette intégrale, nous posons $t=\varphi(s)$, et par conséquent $dt=\varphi'(s)ds$. Étant donné que $\varphi'(s)\geqslant 0$, nous pouvons supprimer les valeurs absolues, et obtenir

$$\begin{aligned} I&=\int_{\varphi(c)}^{\varphi(d)} f\big(\sigma(t)\big)\|\sigma'(t)\|dt\\ &=\int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\|dt\\ &=\int_\sigma f. \end{aligned} \tag{20.146}$$

Essayez de faire le cas $\varphi'(s)\leqslant 0$. □

RemiqswPd

Remarque 20.44.
Attention : les intégrales sur des chemins dans $\mathbb{C}$ ne sont la même chose. En effet $\mathbb{C}$ doit être souvent plutôt traité comme $\mathbb{R}$ que comme $\mathbb{R}^2$. Si γ est un chemin dans $\mathbb{C}$, l'intégrale

$$\int_\gamma f \tag{20.147}$$

doit être comprise comme une généralisation de $\int_a^b f(x)dx$ et non comme l'intégrale sur un chemin. La différence est qu'en retournant les bornes d'une intégrale usuelle sur $\mathbb{R}$ on change le signe, alors qu'en retournant un chemin dans $\mathbb{R}^2$, on ne change pas. Bref, la définition est que si $\gamma\colon[a,b]\to\mathbb{C}$ est un chemin, alors

$$\int_\gamma f=\int_\gamma f(z)dz=\int_a^b f\big(\gamma(t)\big)\gamma'(t)dt, \tag{20.148}$$

sans valeur absolue autour de $\gamma'(t)$.

20.6.3 Intégrer un champ de vecteurs

DEFooSHHFooVdsxMf

Définition 20.45.
*Un **champ de vecteur** est une application $G:\mathbb{R}^n\to\mathbb{R}^n$. On définit l'intégrale de G sur un chemin $\gamma:[a,b]\to\mathbb{R}^n$ par*

$$\int_\gamma G\stackrel{def}{=}\int_a^b \big\langle G(\gamma(t)),\gamma'(t)\big\rangle\,dt.$$

Remarque 20.46.
Cette définition ne dépend pas du paramétrage choisi, mais le signe change selon le sens du chemin.

20.6.4 Intégrer une forme différentielle sur un chemin

La formule d'intégration d'un champ de vecteur [14],

$$\int_\gamma G=\int_{[a,b]}\langle G(\gamma(t)),\gamma'(t)\rangle dt, \tag{20.149}$$

14. Définition 20.45.

contient quelque chose d'intéressant : la combinaison $\langle G(\gamma(t)), \gamma'(t)\rangle$. Cette combinaison sert à transformer le vecteur tangent $\gamma'(t)$ en un nombre en utilisant le produit scalaire avec le vecteur $G(\gamma(t))$.

Si G est un champ de vecteur sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$, nous pouvons utiliser l'isomorphisme musical (définition 12.608)

$$\begin{aligned} G_x^\flat \colon \mathbb{R}^n &\to \mathbb{R} \\ v &\mapsto \langle G(x), v\rangle \end{aligned} \tag{20.150}$$

pour écrire de façon plus compacte :

$$\int_\gamma G = \int_{[a,b]} G^\flat_{\gamma(t)}\big(\gamma'(t)\big)dt. \tag{20.151}$$

DEFooRMHGooFtMEPB

Définition 20.47.
*Soient une forme différentielle ω sur $\mathbb{R}^n$ et un chemin de classe C^1 $\gamma\colon [a,b] \to \mathbb{R}^n$. L'**intégrale** de ω sur γ est définie par*

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}\gamma'(t)dt \tag{20.152}$$

FqEFIZyEe

Remarque 20.48.
Cette définition ne dépend pas du paramétrage choisi, mais le signe change selon le sens du chemin.

20.6.5 Intégration d'une forme différentielle sur un chemin

Les formes intégrales que nous avons déjà vues sont celles de fonctions et de champs de vecteur sur des chemins. Si $\gamma\colon [a,b] \to \mathbb{R}^n$ est le chemin, les formules sont

$$\begin{aligned} \int_\gamma f &= \int_{[a,b]} f\big(\gamma(t)\big)\|\gamma'(t)\|dt \\ \int_\gamma G &= \int_{[a,b]} \langle G\big(\gamma(t)\big), \gamma'(t)\rangle dt. \end{aligned} \tag{20.153}$$

Dans les deux cas, le principe est que nous disposons de quelque chose (la fonction f ou le vecteur G), et du vecteur tangent $\gamma'(t)$, et nous essayons d'en tirer un nombre que nous intégrons. Lorsque nous avons une 1-forme, la façon de l'utiliser pour produire un nombre avec le vecteur tangent est évidemment d'appliquer la 1-forme au vecteur tangent. La définition suivante est donc naturelle.

Définition 20.49.
*Soit $\gamma\colon [a,b] \to \mathbb{R}^n$, un chemin de classe C^1 tel que son image est contenue dans le domaine D. Si ω est une 1-forme différentielle sur D, nous définissons l'**intégrale de ω le long de** γ le nombre*

$$\begin{aligned} \int_\gamma \omega &= \int_a^b \omega_{\gamma(t)}\big(\gamma'(t)\big)dt \\ &= \int_a^b \Big[a_1\big(\gamma(t)\big)\gamma_1'(t) + \cdots + a_n\big(\gamma(t)\big)\gamma_n'(t)\Big]dt. \end{aligned} \tag{20.154}$$

Cette définition est une bonne définition parce que si on change le paramétrage du chemin, on ne change pas la valeur de l'intégrale, c'est la proposition suivante.

Proposition 20.50.
Si γ et β sont des chemins équivalents, alors

$$\int_\gamma \omega = \int_\beta \omega, \tag{20.155}$$

c'est-à-dire que l'intégrale est invariante sous les reparamétrages du chemin.

Démonstration. Deux chemins sont équivalents quand il existe un difféomorphisme C^1 $h\colon [a,b]\to[c,d]$ tel que $\gamma(t)=(\beta\circ h)(t)$. En remplaçant γ par $(\beta\circ h)$ dans la définition de $\int_\gamma\omega$, nous trouvons

$$\int_a^b \omega_{\gamma(t)}\big(\gamma'(t)\big)dt=\int_a^b \omega_{(\beta\circ h)(t)}\big((\beta\circ h)'(t)\big)dt. \tag{20.156}$$

Un changement de variable $u=h(t)$ transforme cette dernière intégrale en $\int_\beta\omega$, ce qui prouve la proposition. □

Remarque 20.51.
Si γ est une somme de chemins, $\gamma=\gamma^{(1)}+\cdots+\gamma^{(n)}$, où chacun des $\gamma^{(i)}$ est un chemin, alors

$$\int_\gamma\omega=\sum_{i=1}^n\int_{\gamma_i}\omega \tag{20.157}$$

parce que ω est linéaire.

Remarque 20.52.
Si $-\gamma$ est le chemin

$$\begin{aligned}-\gamma\colon [a,b]&\to\mathbb{R}^n\\ t&\mapsto\gamma\big(b-(t-a)\big),\end{aligned} \tag{20.158}$$

alors

$$\int_{-\gamma}\omega=-\int_\gamma\omega, \tag{20.159}$$

c'est-à-dire que si l'on parcours le chemin en sens inverse, alors on change le signe de l'intégrale.

L'intégrale d'une forme différentielle sur un chemin est compatible avec l'intégrale déjà connue d'un champ de vecteur sur le chemin parce que si G est un champ de vecteurs,

$$\int_\gamma G^\flat=\int_\gamma G. \tag{20.160}$$

En effet,

$$\begin{aligned}\int_\gamma G^\flat&=\int_a^b G^\flat_{\gamma(t)}(\gamma'(t))\\ &=\int_a^b\big[G_1(\gamma(t))dx_1+\dots G_n(\gamma(t))dx_n\big]\big(\gamma_1'(t),\dots,\gamma_n'(t)\big)\\ &=\int_a^b\langle G(\gamma(t)),\gamma'(t)\rangle\\ &=\int_\gamma G.\end{aligned} \tag{20.161}$$

Proposition 20.53.
Soit $\omega=df$, une 1-forme exacte et continue sur le domaine D. Alors la valeur de $\int_\gamma df$ ne dépend que des valeurs de f aux extrémités de γ.

Démonstration. Nous avons

$$\begin{aligned}\int_\gamma\omega=\int_\gamma df&=\int_a^b\sum_{i=1}^n\frac{\partial f}{\partial x_i}\big(\gamma(t)\big)\gamma_i'(t)dt\\ &=\int_a^b\frac{d}{dt}\Big((f\circ\gamma)(t)\Big)dt\\ &=(f\circ\gamma)(b)-(f\circ\gamma(a)).\end{aligned} \tag{20.162}$$

□

20.6.6 Interprétation physique : travail

Définition 20.54 ([589]).
*Une force $F\colon D \subset \mathbb{R}^n \to \mathbb{R}^n$ est **conservative** si elle dérive d'un potentiel, c'est-à-dire si il existe une fonction $V \in C^1(D, \mathbb{R})$ telle que*

$$F(x) = (\nabla V)(x). \tag{20.163}$$

Étant donné que F est un champ de vecteurs, nous avons une forme différentielle associée $F^\flat$,

$$F_x^\flat\colon v \mapsto \langle F(x), v\rangle. \tag{20.164}$$

Lemme 20.55.
Le champ F est conservatif si et seulement si la 1-forme différentielle $F^\flat$ est exacte.

Démonstration. Supposons que la force F soit conservative, c'est-à-dire qu'il existe une fonction V telle que $F = \nabla V$. Dans ce cas, il est facile de prouver que $F^\flat$ est exacte et est donnée par $F_x^\flat = dV(x)$. En effet,

$$\begin{aligned} F_x^\flat(v) &= \langle F(x), v\rangle \\ &= F_1(x)v_1 + \cdots + F_n(x)v_n \\ &= \frac{\partial V}{\partial x_1}(x)v_1 + \ldots \frac{\partial V}{\partial x_n}(x)v_n \\ &= dV(x)v. \end{aligned} \tag{20.165}$$

Pour le sens inverse, supposons que $F^\flat$ soit exacte. Dans ce cas, nous avons une fonction V telle que $F^\flat = dV$. Il est facile de prouver qu'alors, $F = \nabla V$. □

En résumé, nous avons deux façons équivalentes d'exprimer que la force F dérive du potentiel V : soit nous disons $F = \nabla V$, soit nous disons $F^\flat = dV$.

Proposition 20.56.
Si F est une force conservative, alors le travail[15] *de F lors d'un déplacement ne dépend pas du chemin suivi.*

Démonstration. Le travail d'une force le long d'un chemin n'est autre que l'intégrale de la force le long du chemin, et le calcul est facile :

$$W_\gamma(F) = \int_\gamma F = \int_\gamma dV = V\big(\gamma(b)\big) - V\big(\gamma(a)\big). \tag{20.166}$$

Donc si β est un autre chemin tel que $\beta(a) = \gamma(a)$ et $\beta(b) = \gamma(b)$, nous avons $W_\beta(F) = W_\gamma(F)$. □

20.6.7 Intégrer un champ de vecteurs sur un bord en $2D$

Si $D \subset \mathbb{R}^2$ est tel que ∂D est une variété de dimension 1 et tel que D accepte un champ de vecteur normal extérieur unitaire ν. Si nous voulons définir

$$\int_{\partial D} G, \tag{20.167}$$

le mieux est de prendre un paramétrage $\gamma\colon [0,1] \to \mathbb{R}^2$ et de calculer

$$\int_0^1 \langle G_{\gamma(t)}, \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}\rangle dt. \tag{20.168}$$

Hélas, cette définition ne fonctionne pas parce que son signe dépend du sens du paramétrage γ. Si le paramétrage tourne dans l'autre sens, il y a un signe de différence.

15. Voir [590].

Nous allons définir

$$\int_{\partial D} G = \int_0^1 \langle G_{\gamma(t)}, T(t)\rangle dt \tag{20.169}$$

où $T(t) = \dot{\gamma}(t)/\|\dot{\gamma}(t)\|$ et où γ est choisi de telle façon que la rotation d'angle $\frac{\pi}{2}$ amène ν sur T. Cela fixe le choix de sens.

Ce choix de sens aura des répercussions dans l'application de la formule de Green et du théorème de Stokes.

20.6.8 Intégrer une forme différentielle sur un bord en $2D$

Nous n'allons pas chercher très loin :

$$\int_{\partial D} \omega = \int_{\partial D} \omega^\sharp, \tag{20.170}$$

c'est-à-dire que l'intégrale de la forme différentielle est celle du champ de vecteur associé. Le membre de droite est défini par (20.169), avec le choix d'orientation qui va avec.

20.6.9 Intégrer une forme différentielle sur un bord en $3D$

Nous allons maintenant intégrer une forme différentielle sur certains chemins fermés dans $\mathbb{R}^3$. Soit $F(D) \subset \mathbb{R}^3$, une variété de dimension 2 dans $\mathbb{R}^3$ où $F\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$ est la carte. Nous supposons que D vérifie les hypothèses de la formule de Green. Alors nous définissons

$$\int_{F(\partial D)} \omega = \int_{\partial D} F^*\omega \tag{20.171}$$

où $F^*\omega$ est la forme différentielle définie sur ∂D par $(F^*\omega)(v) = \omega\big(dF(v)\big)$.

Cette définition est très abstraite, mais nous n'allons, en pratique, jamais l'utiliser, grâce au théorème de Stokes.

20.6.10 Intégrer un champ de vecteurs sur un bord en $3D$

Encore une fois, nous n'allons pas chercher bien loin :

$$\int_{F(\partial D)} G = \int_{F(\partial D)} G^\flat \tag{20.172}$$

où $G^\flat$ est la forme différentielle associée au champ de vecteur. Le membre de droite est défini par l'équation (20.171).

20.6.11 Dérivées croisées et forme différentielle exacte

Nous considérons le problème suivant : trouver une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\begin{cases} \dfrac{\partial f}{\partial x} = a(x,y) & \text{(20.173a)} \\ \dfrac{\partial f}{\partial y} = b(x,y) & \text{(20.173b)} \end{cases}$$

où a et b sont des fonctions supposées suffisamment régulières. Nous savons que ce problème n'a pas de solutions lorsque

$$\frac{\partial a}{\partial y} \neq \frac{\partial b}{\partial x} \tag{20.174}$$

parce que cela impliquerait $\partial^2_{xy} f \neq \partial^2_{yx} f$. Nous sommes en droit de nous demander si la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x} \tag{20.175}$$

impliquerait qu'il existe une solution au problème (20.173). La réponse est oui, et nous allons brièvement la justifier. Pour plus de détails nous vous demandons de chercher un peu. La référence [591] peut être utile.

Proposition 20.57.
Si a et b sont des fonctions qui satisfont à la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}, \tag{20.176}$$

alors la fonction

$$f(x,y) = \int_0^x a(t,0)dt + \int_0^y b(x,t)dt \tag{20.177}$$

répond au problème

$$\begin{cases} \frac{\partial f}{\partial x} = a(x,y) & (20.178a) \\ \frac{\partial f}{\partial y} = b(x,y) & (20.178b) \end{cases}$$

La preuve qui suit n'en est pas complètement une parce qu'il manque des justifications, notamment au moment de permuter la dérivée et l'intégrale.

Démonstration. La clef de la preuve est le théorème fondamental de l'analyse :

$$\int_0^x \frac{\partial f}{\partial x}(t,y)dt = f(x,y) \tag{20.179}$$

et son pendant par rapport à y :

$$\int_0^y \frac{\partial f}{\partial y}(x,t)dt = f(x,y). \tag{20.180}$$

En appliquant ces version du théorème fondamental, nous obtenons immédiatement.

$$\frac{\partial f}{\partial y} = b(x,y). \tag{20.181}$$

En ce qui concerne la dérivée par rapport à y,

$$\begin{aligned} \frac{\partial f}{\partial x} &= a(x,0) + \int_0^y \frac{\partial b}{\partial x}(x,t)dt & (20.182a) \\ &= a(x,0) + \int_0^y \frac{\partial a}{\partial y}(x,t)dt & (20.182b) \\ &= a(x,0) + [a(x,t)]_{t=0}^{t=y} & (20.182c) \\ &= a(x,y). & (20.182d) \end{aligned}$$

□

En ce qui concerne l'unicité, supposons que f et g soient deux solutions au problème. L'équation

$$\frac{\partial f}{\partial x} = a(x,y) = \frac{\partial g}{\partial x} \tag{20.183}$$

implique que

$$f(x,y) = g(x,y) + C(y) \tag{20.184}$$

où C est une fonction seulement de y. L'autre équation implique

$$f(x,y) = g(x,y) + D(x) \tag{20.185}$$

où D est seulement une fonction de x. En égalisant nous voyons que les fonctions C et D doivent être des constantes.

Par conséquent la fonction f est donnée à une constante près et en réalité la fonction (20.177) est suffisante pour répondre au problème de trouver toutes les fonctions dont les dérivées partielles sont données par les fonctions a et b.

La fonction f ainsi créée est un **potentiel** pour le champ de force

$$F(x,y) = \begin{pmatrix} a(x,y) \\ b(x,y) \end{pmatrix}. \tag{20.186}$$

Notez que ce champ de vecteurs est le gradient de f. La question initiale aurait donc pu être posée en les termes suivants : trouver une fonction f dont le gradient est donné par

$$\nabla f = \begin{pmatrix} a(x,y) \\ b(x,y) \end{pmatrix}. \tag{20.187}$$

20.7 Surfaces paramétrées

De la même façon qu'un chemin dans $\mathbb{R}^3$ est décrit comme une application $\sigma\colon \mathbb{R} \to \mathbb{R}^3$, une surface dans $\mathbb{R}^3$ sera vue comme une application $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^3$. Une **surface paramétrée** dans $\mathbb{R}^3$ est une application

$$\begin{aligned} \varphi\colon D \subset \mathbb{R}^2 &\to \mathbb{R}^3 \\ (u,v) &\mapsto \varphi(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,z) \end{pmatrix}. \end{aligned} \tag{20.188}$$

Nous allons parler de la « surface φ » pour désigner l'image de φ dans $\mathbb{R}^3$.

Si on fixe le paramètre $u = u_0$, alors l'application

$$v \mapsto \varphi(u_0, v) \tag{20.189}$$

est un chemin dans la surface. Un vecteur tangent à ce chemin sera tangent à la courbe :

$$\frac{\partial\varphi}{\partial v}(u_0,v_0) = \begin{pmatrix} \frac{\partial x}{\partial v}(u_0,v_0) \\ \frac{\partial y}{\partial v}(u_0,v_0) \\ \frac{\partial z}{\partial v}(u_0,v_0) \end{pmatrix}. \tag{20.190}$$

De même, en fixant v_0, on considère le chemin

$$u \mapsto \varphi(u, v_0). \tag{20.191}$$

Le vecteur tangent à ce chemin est égalent tangent à la surface :

$$\frac{\partial\varphi}{\partial u} = \begin{pmatrix} \frac{\partial x}{\partial u}(u_0,v_0) \\ \frac{\partial y}{\partial u}(u_0,v_0) \\ \frac{\partial z}{\partial u}(u_0,v_0) \end{pmatrix} \tag{20.192}$$

DefSurfReguliere

Définition 20.58.
Nous disons que la surface est ***régulière*** *si les vecteurs $\partial_u\varphi(u_0,v_0)$ et $\partial_v\varphi(u_0,v_0)$ sont non nuls et non colinéaires.*

Si la surface est régulière, les vecteurs tangents au paramétrage forment le plan tangent à la surface au point $\varphi(u_0,v_0)$.

Un vecteur orthogonal à la surface (et donc au plan tangent) est donc donné par le produit vectoriel :

$$n(u_0,v_0) = \frac{\partial\varphi}{\partial u}(u_0,v_0) \times \frac{\partial\varphi}{\partial v}(u_0,v_0). \tag{20.193}$$

L'équation du plan tangent est alors obtenue par

$$\begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \cdot n(u_0,v_0) = 0 \tag{20.194}$$ EqPlanTgSurfaceParm

où $x_0 = x(u_0,v_0)$, $y_0 = y(u_0,v_0)$, $z_0 = z(u_0,v_0)$.

20.7.1 Graphe d'une fonction

Soit la fonction $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$. Le graphe de f est l'ensemble des points de la forme

$$\big(x,y,f(x,y)\big) \tag{20.195}$$

tels que $(x,y) \in D$. Cela est une surface paramétrée par

$$\begin{aligned} \varphi\colon D &\to \mathbb{R}^3 \\ (x,y) &\mapsto \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}. \end{aligned} \tag{20.196}$$

Les vecteurs tangents sont

$$\frac{\partial\varphi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix}, \quad \frac{\partial f}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix}. \tag{20.197}$$

La surface est donc partout régulière parce que ces deux vecteurs ne sont jamais nuls ou colinéaires. Un vecteur normal à cette surface au point $(x_0,y_0,f(x_0,y_0))$ est donné par le produit vectoriel

$$n = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \partial_x f(x_0,y_0) \\ 0 & 1 & \partial_y f(x_0,y_0) \end{vmatrix} = -\frac{\partial f}{\partial x}(x_0,y_0)e_x - \frac{\partial f}{\partial y}(x_0,y_0)e_y + e_z. \tag{20.198}$$

En suivant l'équation (20.194), nous avons l'équation suivante pour le plan :

$$\begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \cdot \begin{pmatrix} -\frac{\partial f}{\partial x}(x_0,y_0) \\ -\frac{\partial f}{\partial y}(x_0,y_0) \\ 1 \end{pmatrix} = 0, \tag{20.199}$$

c'est-à-dire

$$-(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) - (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0) + z - f(x_0,y_0) = 0, \tag{20.200}$$

ce qui revient à

$$z - f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0). \tag{20.201}$$

Nous retrouvons donc l'équation du plan tangent à un graphe.

Exemple 20.59.
La sphère de rayon R peut être paramétrée par les angles sphériques :

$$\phi(\theta,\varphi) = \begin{pmatrix} R\sin\theta\cos\varphi \\ R\sin\theta\sin\varphi \\ R\cos\theta \end{pmatrix} \tag{20.202}$$

avec $(\theta,\varphi) \in [0,\pi] \times [0,2\pi]$.

Tentons d'en trouver le plan tangent au point $(x,y,z)=(R,0,0)$. Un petit dessin nous montre que c'est un plan vertical d'équation $x=R$. Montrons cela en utilisant la théorie que nous venons de découvrir. D'abord le point $(R,0,0)$ correspond à $\theta_0=\frac{\pi}{2}$ et $\varphi=0$. Les vecteurs tangents sont

$$T_\theta=\frac{\partial\phi}{\partial\theta}(R,\frac{\pi}{2},0)=\begin{pmatrix}R\cos\theta\cos\varphi\\R\cos\theta\sin\varphi\\-R\sin\theta\end{pmatrix}=\begin{pmatrix}0\\0\\-R\end{pmatrix}, \tag{20.203}$$

et

$$T_\varphi=\frac{\partial\phi}{\partial\varphi}(R,\frac{\pi}{2},0)=\begin{pmatrix}-R\sin\theta\sin\varphi\\R\sin\theta\cos\varphi\\0\end{pmatrix}=\begin{pmatrix}0\\R\\0\end{pmatrix}. \tag{20.204}$$

Cela sont de toute évidence bien les deux vecteurs tangents à la sphère au point $(x,y,z)=(R,0,0)$. Le vecteur normal est

$$\begin{vmatrix}e_x&e_y&e_z\\0&0&-R\\0&R&0\end{vmatrix}=R^2e_x. \tag{20.205}$$

Ici encore, nous avons le vecteur que nous attendions sur un dessin. L'équation du plan tangent est maintenant

$$\begin{pmatrix}x-R\\y\\z\end{pmatrix}\cdot\begin{pmatrix}R^2\\0\\0\end{pmatrix}=0, \tag{20.206}$$

c'est-à-dire $R^2(x-R)=0$ et donc $x=R$. △

20.7.2 Intégrale sur une partie de $\mathbb{R}^m$

Soit M une variété de dimension n dans $\mathbb{R}^m$. Soit $F:W\subset\mathbb{R}^n\to M$ un paramétrage d'un ouvert relatif de M.

Si f est une fonction définie sur un sous-ensemble $A\subset F(W)$ tel que $F^{-1}(A)$ est mesurable, l'*intégrale de f sur A* est définie par

$$\int_A f=\int_{F^{-1}(A)}f(F(w))\sqrt{\det(J_F(w)^tJ_F(w))}dw$$

où l'intégrale est l'intégration usuelle (de Lebesgue) sur $F^{-1}(A)\subset\mathbb{R}^n$. On écrit parfois cette intégrale $\int_{F^{-1}(A)}f(F(w))d\sigma$ où

$$d\sigma=\sqrt{\det(J_F(w)^tJ_F(w))}dw \tag{20.207}$$

est l'*élément infinitésimal de volume* de la variété.

Si $m=3$ et $n=2$, l'élément infinitésimal de volume vaut

$$d\sigma=\left\|\frac{\partial F}{\partial w_1}\times\frac{\partial F}{\partial w_2}\right\|dw$$

où $\times$ représente le produit vectoriel dans $\mathbb{R}^3$, et (w_1,w_2) sont les coordonnées sur $W\subset\mathbb{R}^2$. Dans la suite, nous ne regarderons plus que ce cas.

20.8 Intégrales de surface

20.8.1 Intégrale d'un champ de vecteurs

Dans l'intégration curviligne, on a noté que si l'intégrale d'une fonction ne dépendait pas de l'orientation du chemin, l'intégrale d'un champ de vecteurs ou d'une forme différentielle en dépendait. Ce problème d'orientation apparait également dans l'intégration sur des surfaces de l'espace.

DEFooFTQLooXXbtOQ

Définition 20.60.
*Une **orientation** sur une surface $S \subset \mathbb{R}^3$ est le choix d'un champ de vecteurs continu $\nu : S \to \mathbb{R}^3$ dont la norme en tout point de S vaut* 1.

On remarque qu'ayant fait un tel choix d'orientation $\nu(x)$ en un point x, le seul autre choix possible en x est $-\nu(x)$. Si S est le bord d'un ouvert $D \subset \mathbb{R}^3$, l'*orientation induite par D sur S* est, si elle existe, l'orientation qui pointe hors de D en tout point de S. Plus précisément, il faut que pour tout $x \in D$ il existe $\epsilon > 0$ vérifiant, pour tout $0 < t < \epsilon$, la relation $t\nu(x) \notin D$. Dans ce cas, le champ de vecteurs ν est appelé le *vecteur normal unitaire extérieur* à D et il est forcément unique.

Soit G un champ de vecteurs défini sur une surface orientée par un champ ν. L'intégrale de G sur S, aussi appelée le *flux de G à travers S*, est

$$\iint_S G \cdot dS \stackrel{\text{def}}{=} \iint_S \langle G, \nu \rangle \, d\sigma. \tag{20.208}$$

eqflux-star

Si on suppose que la surface est paramétrée par une application

$$F : W \subset \mathbb{R}^2 \to \mathbb{R}^3 : (u, v) \mapsto (F_1(u, v), F_2(u, v), F_3(u, v))$$

alors un vecteur unitaire ν peut s'écrire sous la forme

$$\nu = \frac{\frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\|}$$

et grâce à ce paramétrage l'intégrale (20.208) devient

$$\iint_S G \cdot dS = \iint_W \left\langle G(F(u, v)), \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\rangle du dv.$$

où on utilise l'expression de $d\sigma$ obtenue précédemment dans le cas qui nous intéresse (surface dans l'espace).

20.9 Aires et intégrales

20.9.1 Aire d'une surface paramétrée

Lorsque nous avions vu la longueur d'une courbe paramétrée, nous avions pris comme « élément de longueur » la norme du vecteur tangent. Il est donc naturel de prendre comme « élément de surface » une petite surface que l'on peut construire à partir des deux vecteurs tangents à la surface.

Au point $\varphi(u_0, v_0)$, nous avons les deux vecteurs tangents

$$T_u = \frac{\partial \varphi}{\partial u}(u_0, v_0) \quad T_v = \frac{\partial \varphi}{\partial v}(u_0, v_0). \tag{20.209}$$

L'élément de surface que nous pouvons construire à partir de ces deux vecteurs est la surface du parallélogramme, donnée par la norme du produit vectoriel :

$$dS = \|T_u \times T_v\|. \tag{20.210}$$

EQooNYWSooZuvcPe

L'aire de la surface donnée par φ: $D \subset \mathbb{R}^2 \to \mathbb{R}^3$ sera donc donnée par

$$Aire\big(\varphi(D)\big) = \iint_D \|T_u \times T_v\| du \, dv. \tag{20.211}$$

Exemple 20.61.
Calculons l'aire de la sphère. Les vecteurs tangents ont déjà été calculés aux équations (20.203) et (20.204) :

$$T_\theta = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix}, \quad T_\varphi = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix}. \tag{20.212}$$

Le produit vectoriel vaut

$$\begin{aligned} T_\theta \times T_\varphi &= \begin{vmatrix} e_x & e_y & e_z \\ R\cos\theta\cos\varphi & R\cos\theta\sin\varphi & -R\sin\theta \\ -R\sin\theta\sin\varphi & R\sin\theta\cos\varphi & 0 \end{vmatrix} \\ &= (R^2\sin^2\theta\cos\varphi)e_x + (R^2\sin^2\theta\sin\varphi)e_y \\ &\quad + (R^2\cos\theta\sin\theta\cos^2\varphi + R^2\sin\theta\cos\theta\sin^2\varphi)e_z. \end{aligned} \tag{20.213}$$

La norme demande quelques calculs et mises en évidences. Le résultat est :

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{20.214}$$ EqProdVectTTSPh

L'aire de la sphère est donc donnée par

$$Aire = \int_0^{2\pi} d\varphi \int_0^\pi R^2\sin\theta d\theta = 2\pi R^2[-\cos\theta]_0^\pi = 4\pi R^2. \tag{20.215}$$

Il est bon de se souvenir que, en coordonnées sphériques,

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{20.216}$$

Or nous savons que ce vecteur est dirigé dans le sens de e_r parce que ce dernier est le vecteur qui est constamment dirigé radialement. En coordonnées sphériques nous avons donc

$$T_\theta \times T_\varphi = R^2\sin(\theta)e_r. \tag{20.217}$$ EqNormalEnSph

△

Remarque 20.62.
L'équation (20.214) donne l'élément de surface pour la sphère. Notez que cela est justement l'expression du jacobien des coordonnées sphériques. Cela n'est évidemment pas une coïncidence.

Exemple 20.63.
Nous pouvons donner l'aire du graphe d'une fonction quelconque. La surface est paramétrée par

$$\varphi(x,y) = \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}. \tag{20.218}$$

Les vecteurs tangents sont

$$T_x = \begin{pmatrix} 1 \\ 0 \\ \partial_x f \end{pmatrix}, \quad T_y = \begin{pmatrix} 0 \\ 1 \\ \partial_y f \end{pmatrix}. \tag{20.219}$$

Le produit vectoriel est donné par

$$T_x \times T_y = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \partial_x f \\ 0 & 1 & \partial_y f \end{vmatrix} = (-\partial_x f)e_x - (\partial_y f)e_y + e_z. \tag{20.220}$$

L'élément de surface est par conséquent

$$dS = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}, \tag{20.221}$$

et la surface du graphe sera

$$Aire = \iint_D \sqrt{\left(\frac{\partial f}{\partial x}(x,y)\right)^2 + \left(\frac{\partial f}{\partial y}(x,y)\right)^2 + 1}\, dx\, dy \tag{20.222}$$

△

20.9.2 Intégrale d'une fonction sur une surface

Si S est une surface dans $\mathbb{R}^3$ paramétrée par

$$\begin{aligned} &\varphi\colon D \to \mathbb{R}^3 \\ &(u,v) \mapsto \varphi(u,v) \in S, \end{aligned} \tag{20.223}$$

et si f est une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$ définie au moins sur S, l'intégrale de f sur S est logiquement définie par

$$\int_S f\, dS = \iint_D f\big(\varphi(u,v)\big) \|T_u(u,v) \times T_v(u,v)\| du dv \tag{20.224}$$

où $T_u = \frac{\partial \varphi}{\partial u}$ et $T_v = \frac{\partial \varphi}{\partial v}$. La quantité

$$\|T_u(u,v) \times T_v(u,v)\| du dv \tag{20.225}$$

est appelé **élément de surface**.

Encore une fois, si on prend $f = 1$, alors on retrouve la surface de S :

$$\int_S dS = Aire(S). \tag{20.226}$$

Remarque 20.64.
Le nombre $\int_S f dS$ ne dépend pas de le paramétrage choisie pour S.

20.9.3 Intégrale d'une 2-forme

Nous considérons ω, une 2-forme différentielle sur $\mathbb{R}^2$.

Définition 20.65.
Si $\omega_{(x,y)} = u(x,y)dx \wedge dy$ et si D est un ouvert de $\mathbb{R}^2$ alors nous définissons

$$\int_D \omega = \int_D u(x,y) dx\, dy. \tag{20.227}$$

Nous voulons maintenant intégrer une 2-forme sur une surface dans $\mathbb{R}^3$. Soit $S \subset \mathbb{R}^3$, une surface orientée (c'est-à-dire que nous avons un choix continu d'un vecteur normal unitaire n). Nous supposons de plus avoir un paramétrage $\phi\colon D \to S$ de S avec D ouvert dans $\mathbb{R}^2$ compatible avec l'orientation, c'est-à-dire que pour tout $(t,s) \in D$,

$$n\big(\phi(t,s)\big) = \frac{\partial \phi}{\partial t}(t,s) \times \frac{\partial \phi}{\partial s}(t,s). \tag{20.228}$$

Définition 20.66.
Pour intégrer ω sur S nous faisons

$$\int_S \omega = \int_D \phi^* \omega \tag{20.229}$$

où $\phi^*\omega$ est de la forme $F(t,s)dt \wedge ds$.

Montrons ce que cela fait. Soient u, v des vecteurs de D et calculons

$$(\phi^*\omega)(u,v) = \omega\big(d\phi(u), d\phi(v)\big) \tag{20.230a}$$
$$= \omega\left(u_1\frac{\partial\phi}{\partial x_1} + u_2\frac{\partial\infty}{\partial x_2}, v_1\frac{\partial\phi}{\partial x_1} + v_2\frac{\partial\infty}{\partial x_2}\right). \tag{20.230b}$$

Les termes en u_1v_1 et u_2v_2 sont nuls ; par exemple :

$$\omega\left(u_1\frac{\partial\phi}{\partial x_1}, v_1\frac{\partial\phi}{\partial x_1}\right) = u_1v_1\omega\left(\frac{\partial\infty}{\partial x_1}, \frac{\partial\phi}{\partial x_1}\right) = 0 \tag{20.231}$$

parce que ω est antisymétrique. Il nous reste donc

$$(\phi^*\omega)(u,v) = (u_1v_2 - u_2v_1)\omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right) \tag{20.232a}$$
$$= \omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right)(dt\wedge ds)(u,v). \tag{20.232b}$$

Cette dernière ligne est bien de la forme $\phi^*\omega = F(t,s)dt\wedge ds$.

20.10 Flux d'un champ de vecteurs à travers une surface

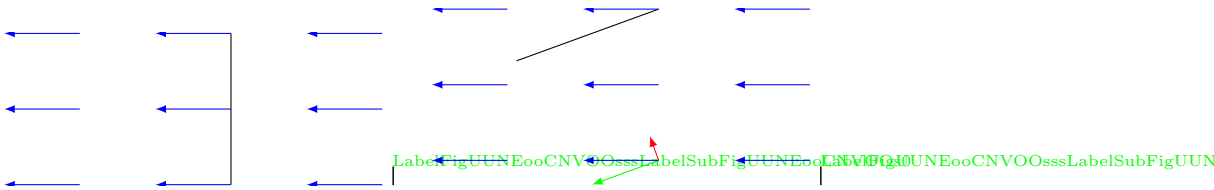
Nous voulons construire un moulin à eau. Comment placer les pales pour maximiser le travail de la pression de l'eau ? On n'a pas attendu l'invention du calcul intégral pour répondre à cette question. Trois paramètres rentrent en ligne de compte :

(1) plus il y a d'eau, plus ça pousse ;

(2) plus la surface de la pale est grande, plus on va utiliser d'eau ;

(3) plus la pale est perpendiculaire au courant, plus on va en profiter.

Nous voyons sur la figure 20.2 que lorsque la pale du moulin est inclinée, non seulement elle prend moins d'eau sur elle, mais qu'en plus elle la prend avec un moins bon angle : une partie de la force ne sert pas à la faire tourner.



(a) L'eau pousse bien perpendiculairement à la palle du moulin.

(b) Lorsque l'eau ne pousse pas perpendiculairement, une partie de la force est perdue. Ici la partie normale (en rouge) est très petite.

FIGURE 20.2: La partie rouge de la force est perdue si l'eau ne pousse pas perpendiculairement. De plus lorsque la pale est inclinée, elle prend moins d'eau sur elle. LabelFigUUNEooCNVOOs

L'idée du flux d'un champ de vecteurs à travers une surface est de savoir quelle est la quantité « utile » de vecteurs qui traverse la surface. Ce sera simplement l'intégrale sur la surface de la composante du champ de vecteurs normale à la surface. Il reste deux problèmes à régler : le premier est de savoir quel est le vecteur normal à la surface, et le second est de savoir comment « sélectionner » la composante normale d'un champ de vecteurs F.

Le problème de trouver un vecteur normal est résolu par le produit vectoriel des vecteurs tangents. Si la surface est donnée par $\varphi\colon D\subset\mathbb{R}^2\to\mathbb{R}^3$, les vecteurs tangents sont $T_u = \partial_u\varphi(u,v)$ et $T_v = \partial_v\varphi(u,v)$. Le normal de norme 1 est donné par :

$$n(u,v) = \frac{T_u\times T_v}{\|T_u\times T_v\|}. \tag{20.233}$$

Si p est un point de la surface $\varphi(D)$, la composante de $F(p)$ qui est normale à la surface au point p est donnée par le produit scalaire

$$F(p)_\perp = F(p) \cdot n(p). \tag{20.234}$$

C'est ce nombre là que nous intégrons sur la surface.

Définition 20.67.
*Le **flux du champ de vecteurs** à travers la surface $S = \varphi(D)$ est*

$$\int F \cdot dS = \int F \cdot n \, dudv. \tag{20.235}$$

Une petite simplification se produit lorsqu'on veut calculer effectivement cette intégrale. En effet $F \cdot n$ est, en soi, une fonction sur S. Pour l'intégrer, il faut donc la multiplier par $\|T_u \times T_v\|$ (c'est la définition de l'intégrale d'une fonction sur une surface). Donc, étant donné que $n = (T_u \times T_v)/\|T_u \times T_v\|$, nous avons

$$\int F \cdot dS = \iint_D F\big(\varphi(u,v)\big) \cdot (T_u \times T_v) \, dudv \tag{20.236}$$

où $T_u = \frac{\partial \phi}{\partial u}$ et $T_v = \frac{\partial \varphi}{\partial v}$.

Exemple 20.68.
Soit le champ de vecteurs

$$F = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}. \tag{20.237}$$

Calculons son flux au travers de la sphère de rayon R.

Nous choisissons de paramétrer la sphère en coordonnées sphériques avec $\phi(\theta, \varphi)$. Nous pouvons reprendre le résultat (20.217) :

$$T_\theta \times T_\varphi = R^2 \sin(\theta). \tag{20.238}$$

Nous savons aussi que

$$F\big(\phi(\theta,\varphi)\big) = Re_r. \tag{20.239}$$

L'intégrale à calculer est donc

$$I = \int_0^\pi d\theta \int_0^{2\pi} d\varphi \, 2e_r \cdot \big(R^2 \sin(\theta) e_r\big). \tag{20.240}$$

Vu que le produit scalaire $e_r \cdot e_r$ vaut 1, nous calculons

$$I = 4\pi R^2 \int_0^\pi \sin(\theta) d\theta = 8\pi R^2. \tag{20.241}$$

△

Exemple 20.69.
Calculons le flux du champ de force de gravitation d'une masse au travers de la sphère de centre R centrée autour la masse. À un coefficient constant près, le champ vaut

$$G(r,\theta,\varphi) = \frac{1}{r^2} e_r. \tag{20.242}$$

Sur la sphère de rayon R, nous avons

$$G\big(\phi(\theta,\varphi)\big) = \frac{1}{R^2} e_r. \tag{20.243}$$

L'intégrale est donc

$$\int_0^\pi d\theta \int_0^{2\pi} \frac{1}{R^2} e_r \cdot \big(R^2 \sin(\theta) e_r\big) d\varphi = 8\pi. \tag{20.244}$$

Ce flux ne dépend pas de R. △

Exemple 20.70.
Soit S le disque de rayon 5 placé horizontalement à la hauteur 12. Calculer le flux du champ de vecteurs

$$F(x,y,z) = xe_x + ye_y + ze_z. \tag{20.245}$$

Les équations de la surface sont $z = 12$, $x^2+y^2 \leqslant 25$. Nous prenons le paramétrage en coordonnées cylindriques :

$$\varphi(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 12 \end{pmatrix}. \tag{20.246}$$

Les vecteurs tangents sont

$$T_r = \frac{\partial\varphi}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}. \tag{20.247}$$

Le vecteur normal est alors

$$T_r \times T_\theta = re_z. \tag{20.248}$$

Sur la surface, le champ de vecteurs s'écrit

$$F\big(\varphi(r,\theta)\big) = r\cos(\theta)e_x + r\sin(\theta)e_y + 12e_z. \tag{20.249}$$

Par conséquent

$$F \cdot (T_r \times T_\theta) = 12r. \tag{20.250}$$

L'intégrale à calculer est

$$\begin{aligned} \int_0^5 dr \int_0^{2\pi} 12r\,d\theta &= 12 \cdot 2\pi \int_0^5 r\,dr \\ &= \frac{25}{2}24\pi \\ &= 300\pi. \end{aligned} \tag{20.251}$$

△

20.11 Divergence, Green, Stokes

Le théorème de Stokes (et ses variations) peut se voir comme une généralisation du théorème fondamental du calcul différentiel et intégral qui stipule que

$$\int_a^b f'(x)dx = f(b) - f(a)$$

c'est-à-dire qui relie l'intégrale de f' sur $I = [a,b]$ aux valeurs de f sur le bord $\partial I = \{a,b\}$. Le signe $-$ qui apparait vient de l'orientation ; celle-ci requiert de la prudence dans l'utilisation des théorèmes.

Voici, pour votre culture générale, un énoncé général :

ThoATsPuzF

Théorème 20.71.
Si M est une variété orientable de dimension n avec un bord noté ∂M, alors pour toute forme différentielle ω de degré $n-1$ on a

$$\int_M d\omega = \int_{\partial M} \omega.$$

où $d\omega$ désigne la différentielle extérieure de ω.

Nous allons maintenant voir quelques cas particuliers.

Une des nombreuses formes du théorème de Stokes (théorème 20.71) est que si la forme différentielle ω est exacte alors son intégrale est facile.

ThoUJMhFwU

Théorème 20.72 ([1]).
Si γ est une chemin de classe C^1 dans un ouvert Ω et si ω est la forme différentielle exacte $\omega = df$, alors

$$\int_\gamma df = f\big(\gamma(1)\big) - f\big(\gamma(0)\big). \tag{20.252}$$

Démonstration. C'est une application du lemme 12.280 et du théorème fondamental du calcul intégral 14.287. □

20.11.1 Théorème de la divergence

Si nous considérons une surface dans $\mathbb{R}^n$ et un champ de vecteurs, il est bon de se demander quelle « quantité de vecteurs » traverse la surface. Soit D, un ouvert borné de $\mathbb{R}^n$ telle que ∂D soit une variété de dimension $n-1$, et G, un champ de vecteurs défini sur $\bar{D}$. Afin de compter combien de G traverse ∂D, il faudra faire en sorte de ne considérer que la composante de G normale à ∂D : pas question d'intégrer par exemple la norme de G sur ∂D.

Comme nous le savons, la composante du vecteur v dans la direction w est le produit scalaire $v \cdot 1_w$ où 1_w est le vecteur de norme 1 dans la direction w. Nous allons donc introduire le concept de vecteur normal extérieur.

Définition 20.73.
*Soit $x \in \partial D$ et $\nu \in \mathbb{R}^n$, nous disons que ν est un **vecteur normal extérieur** de ∂D si*

(1) $\langle \nu, v\rangle = 0$ pour tout vecteur tangent v à ∂D au point x. Pour rappel, ∂D étant une variété de dimension $n-1$, il y a $n-1$ tels vecteurs v linéairement indépendants.
(2) Il existe un $\delta > 0$ tel que $\forall t \in]0, \delta[$, nous avons $x + t\nu \notin \bar{D}$ et $x - t\nu \in D$.

Nous pouvons maintenant définir le concept de flux. Soit $D \subset \mathbb{R}^n$ tel que ∂D soit une variété de dimension $n-1$ qui admette un vecteur normal extérieur $\nu(x)$ en chaque point. Soit aussi $G\colon \bar{D} \to \mathbb{R}^n$, un champ de vecteur de classe C^1. Le **flux** de G au travers de ∂D est le nombre

$$\int_{\partial D} \langle G(x), \nu(x)\rangle d\sigma(x). \tag{20.253}$$

Cette intégrale est en général très compliquée à calculer parce qu'il faut trouver le champ de vecteur normal, puis un paramétrage de la surface ∂D et ensuite appliquer la méthode décrite au point 20.8.

Heureusement, il y a un théorème qui nous permet de calculer plus facilement : sans devoir trouver de vecteurs normaux.

Il n'est pas plus contraignant d'énoncer ce théorème dans le cadre d'une hypersurface de $\mathbb{R}^n$, ce que nous faisons donc :

Théorème 20.74 (Formule de la divergence).
Soit D un ouvert borné de $\mathbb{R}^n$ dont le bord est « assez régulier par morceaux », c'est-à-dire :

$$\partial D = A_1 \cup \dots A_p \cup N \tag{20.254}$$

où

(1) $A_1, \dots, A_p, N$ sont deux à deux disjoints,
(2) pour tout $i \leqslant p$, A_i est un ouvert relatif d'une certaine variété M_i de dimension $(n-1)$
(3) $\bar{A}_i \subset M_i$
(4) N est un compact contenu dans une réunion finie de variétés de dimensions $(n-2)$.

Supposons également qu'en chaque point de $A_1 \cup \dots \cup A_p$ il existe un vecteur normal extérieur ν.

Si G est un champ de vecteurs de classe C^1 sur $\bar{D}$ alors

$$\int_D \nabla \cdot G = \sum_{i=1}^{p} \int_{A_i} \langle G, \nu\rangle. \tag{20.255}$$

L'intégrale du membre de gauche est l'intégrale sur un ouvert d'une simple fonction.

20.11.2 Lacets et homotopie

DEFooQZMSooYYkGDv

Définition 20.75 (chemin, lacet[205, 592])**.**
Plusieurs notions autour de chemins dans un espace topologique X.

(1) *Un* ***chemin*** *dans un espace vectoriel X est une application continue $\gamma\colon [a,b] \to X$ avec $a < b$.*
(2) *Un* ***lacet*** *dans X est un chemin $\gamma\colon [a,b] \to X$ tel que $\gamma(a) = \gamma(b)$.*
(3) *Un chemin $\gamma\colon [0,1] \to \mathbb{R}^n$ est* ***régulier*** *si il est C^1 et si $\gamma'(t) \neq 0$ pour tout t*
(4) *Le chemin $\gamma\colon [a,b] \to X$ est un* ***chemin de Jordan*** *si il est injectif.*
(5) *Un lacet $\gamma\colon [a,b] \to X$ est un* ***lacet de Jordan*** *si $\gamma\colon [a,b[\to X$ est injective.*
(6) *Une* ***courbe de Jordan*** *est l'image d'un lacet de Jordan.*
(7) *Une* ***courbe simple*** *est l'image d'un chemin de Jordan.*

L'homotopie sera définie en 20.77.

20.76.
Un chemin de Jordan peut évidemment être vu comme une application $\gamma\colon [0,2\pi] \to \mathbb{R}^2$ telle que $\gamma(0) = \gamma(2\pi)$. En particulier il n'est jamais mauvais de se rappeler qu'on peut choisir un paramétrage normal par la proposition 21.46.

DEFooHJQTooYUFcee

Définition 20.77 (homotopie de lacets)**.**
Soit un espace topologique X. Les lacets[16] *$\gamma_0, \gamma_1\colon [a,b] \to X$ sont* ***homotopes*** *dans X si il existe une application continue*

$$H\colon [0,1] \times [a,b] \to X \tag{20.256}$$

telle que

(1) *Pour tout $s \in [0,1]$, l'application*

$$\begin{aligned} \Gamma_s\colon [a,b] &\to X \\ t &\mapsto H(s,t) \end{aligned} \tag{20.257}$$

est un lacet dans X.
(2) *Nous avons $\Gamma_0 = \gamma_0$ et $\Gamma_1 = \gamma_1$.*

L'application H est l'homotopie entre γ_0 et γ_1.

DEFooLXDTooDPgxqL

Définition 20.78 (Homotopie à extrémités fixées[205])**.**
Soit un espace topologique X. Soient des chemins $\gamma_0, \gamma_1\colon [a,b] \to X$. Nous supposons que leurs extrémités soient égales, et nous les notons p, q : $\gamma_0(a) = \gamma_1(a) = p$ et $\gamma_0(b) = \gamma_1(b) = q$. Nous disons que γ_0 et γ_1 sont ***homotopes avec les extrémités fixées*** *si il existe une application continue $H\colon [0,1] \times [a,b] \to X$ telle que*

(1) *les extrémités sont fixées : $H(s,a) = p$ et $H(s,b) = q$ pour tout $s \in [0,1]$.*
(2) *pour tout $t \in [a,b]$, nous avons $H(0,t) = \gamma_0(t)$ et $H(1,t) = \gamma_1(t)$.*

20.11.3 Formule de Green

La formule de Green est un cas particulier du théorème de la divergence dans le cas $n = 2$, légèrement reformulé.

THOooQSWMooAZasTl

Théorème 20.79.
Soit $D \subset \mathbb{R}^2$ ouvert borné tel que son bord est la réunion finie d'un certain nombre de chemins

16. Définition 20.75.

de classe C^1 de Jordan réguliers. Supposons qu'en chaque point de son bord, D possède un vecteur normal unitaire extérieur ν. Soient P et Q deux fonctions réelles de classe C^1 sur $\bar{D}$. Alors

$$\iint_D (\partial_x Q - \partial_y P) dx\, dy = \oint_{\partial D} P dx + Q dy \tag{20.258}$$

où chaque chemin γ formant le bord de D est orienté de sorte que $T\nu = \frac{\dot{\gamma}}{\|\dot{\gamma}\|}$ où T représente la rotation d'angle $+\frac{\pi}{2}$.

Justifions le fait que cela soit un cas particulier de la formule de Stokes du théorème 20.71. Nous considérons la forme différentielle

$$\omega = P dx + Q dy, \tag{20.259}$$

et sa différentielle

$$\begin{aligned} d\omega &= \sum_i d\omega_i \wedge dx_i && \text{(20.260a)}\\ &= \left(\frac{\partial P}{\partial x} dx + \frac{\partial P}{\partial y} dy\right) \wedge dx + \left(\frac{\partial Q}{\partial x} dx + \frac{\partial Q}{\partial y} dy\right) \wedge dy && \text{(20.260b)}\\ &= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy. && \text{(20.260c)} \end{aligned}$$

Intégrons cette forme $d\omega$ sur le domaine ouvert D que nous paramétrons de façon triviale par

$$\begin{aligned} \varphi\colon D &\to \mathbb{R}^2 \\ (u, v) &\mapsto (u, v). \end{aligned} \tag{20.261}$$

Ce que nous avons est

$$\iint_D d\omega = \iint_D d\omega_{(u,v)} \left(\frac{\partial \varphi}{\partial u}, \frac{\partial \varphi}{\partial v}\right) du dv \tag{20.262}$$

Nous avons aussi $T_u = \frac{\partial \varphi}{\partial u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $T_v = \frac{\partial \varphi}{\partial v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et donc

$$(dx \wedge dy)(T_u, T_v) = dx(T_u)dy(T_v) - dx(T_v)dy(T_u) = 1 - 0 = 1. \tag{20.263}$$

L'intégrale (20.262) se développe donc en

$$\iint_D d\omega = \iint_D \left(\frac{\partial Q}{\partial x}(u, v) - \frac{\partial P}{\partial y}(u, v)\right) (dx \wedge dy)(T_u, T_v) du dv = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) du dv. \tag{20.264}$$

Par conséquent la formule de Stokes nous donne la formule (20.258).

La formule de Green nous permet de calculer l'aire de la surface délimitée par une courbe fermée en termes de l'intégrale d'une forme bien choisie le long du contour. Pour cela nous prenons la forme

$$\omega = -\frac{y}{2} dx + \frac{x}{2} dy, \tag{20.265}$$

de telle sorte que $\partial_x Q - \partial_y Q = 1$ et que

$$\iint_D d\omega = \iint_D du dv = S, \tag{20.266}$$

et au final l'aire est donnée par

$$S = \int_{\partial D} \left(-\frac{y}{2} dx + \frac{x}{2} dy\right). \tag{20.267}$$

Lorsque le bord de D est paramétré par

$$\begin{aligned}\gamma\colon [a,b] &\to \mathbb{R}^2\\ u &\mapsto \begin{pmatrix} x(u)\\ y(u)\end{pmatrix},\end{aligned} \tag{20.268}$$

nous avons

$$(Pdx+Qdy)\gamma'(u) = Px'+Qy', \tag{20.269}$$

et alors

$$\int_{\partial D} Pdx+Qdy = \int_a^b P\big(x(u),y(u)\big)x'(u)+Q\big(x(u),y(u)\big)y'(u)du. \tag{20.270}$$

En ce qui concerne l'aire de la surface, nous prenons les P et Q de la forme 20.265 :

$$S = \frac{1}{2}\int_a^b \Big(-y(u)x'(u)+x(u)y'(u)\Big)du. \tag{20.271}$$ EqAJGrtOk

20.11.4 Formule de Stokes

secstokesusuel

La formule de Stokes est la version classique, qui permet d'exprimer la circulation d'un champ de vecteur le long d'une courbe de $\mathbb{R}^3$ comme le flux de son rotationnel à travers n'importe quelle surface dont le bord est la courbe. La version présentée ici suppose que la surface peut se paramétrer en un seul morceau :

THOooDCPKooMqNOMU

Théorème 20.80 (Formule de Stokes)**.**
Soit $F : W \subset \mathbb{R}^2 \to \mathbb{R}^3$ un paramétrage (carte) d'une surface dans $\mathbb{R}^3$, supposée de classe C^2. Soit D un ouvert de $\mathbb{R}^2$ vérifiant les hypothèses de la formule de Green, et tel que $\bar{D} \subset W$. Soit G un champ de vecteurs de classe C^1 défini sur $F(\bar{D})$, et soit N le champ normal unitaire donné par le paramétrage

$$N = \frac{\frac{\partial F}{\partial u}\wedge\frac{\partial F}{\partial v}}{\left\|\frac{\partial F}{\partial u}\wedge\frac{\partial F}{\partial v}\right\|} \tag{20.272}$$

alors

$$\iint_{F(D)} \langle \nabla\times G, N\rangle\, d\sigma_F = \int_{F(\partial D)} G \tag{20.273}$$ EqStokesTho

où les chemins formant le bord ∂D sont orientés comme dans le théorème de Green.

Notons, juste pour avoir une bonne nouvelle de temps en temps, que

$$d\sigma_F = \left\|\frac{\partial F}{\partial u}\times\frac{\partial F}{\partial v}\right\| dudv, \tag{20.274}$$

mais cette norme apparaît exactement au dénominateur de N. Il ne faut donc pas la calculer parce qu'elle se simplifie.

Sous forme un peu plus physicienne [17], la formule (20.273) s'écrit

$$\int_{F(D)} \langle \nabla\times G, N(x)\rangle\, d\sigma_F(x) = \int_{F(\gamma)} \langle G, T\rangle\, ds \tag{20.275}$$

où T est le vecteur unitaire tangent à $F(\gamma)$.

17. et surtout plus explicite.

20.11.4.1 Quelle est la bonne orientation ?

Le signe du vecteur normal N dépend du choix de l'ordre des coordonnées dans la carte. Supposons que je veuille paramétrer la surface $x^2 + y^2 = 1$, $z = 1$. Nous prenons naturellement comme carte le cercle C de rayon 1 dans $\mathbb{R}^2$ et la carte

$$F(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{20.276}$$

Mais nous aurions aussi pu mettre les coordonnées r et θ dans l'autre ordre :

$$\tilde{F}(\theta, r) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{20.277}$$

Les vecteurs normaux ne sont pas les même : la carte F donnera $\partial_r F \times \partial_\theta F$, tandis que l'autre donnera $\partial_\theta \tilde{F} \times \partial_r \tilde{F}$. Le signe change !

Il faut savoir laquelle choisir. Le cercle $C \subset \mathbb{R}^2$ a une orientation donnée par le théorème de Green. Nous choisissons l'ordre des coordonnées pour que 1_θ et 1_r soient dans la même orientation que les vecteurs ν et T tels que donnés par le théorème de Green, et tels que dessinés sur la figure 20.3.

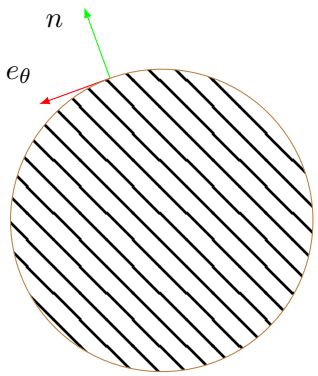


FIGURE 20.3: L'orientation sur le cercle. Si nous les prenons dans l'ordre, les vecteurs $(1_r, 1_\theta)$ ont la même orientation que celle donnée par les vecteurs (ν, T) donnés par la convention de Green. LabelFigCercleInu

Plus généralement, nous choisissons l'ordre des coordonnées u et v pour que la base $(1_u, 1_v)$ ait la même orientation que (ν, T) où T a le sens convenu dans le théorème de Green.

20.12 Résumé des intégrales vues

Nous sommes maintenant capables de revoir tous les types d'intégrales vues jusqu'ici de façon très cohérentes. Nous commencerons par les intégrales de fonctions et nous ferons ensuite les intégrales de champs de vecteurs.

20.12.1 L'intégrale d'une fonction sur les réels

Si $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ est une fonction usuelle, sont intégrale est

$$\int_a^b f(x)dx = F(b) - F(a) \tag{20.278}$$

où F est une primitive de f.

20.12.2 Intégrale d'une fonction sur un chemin

Si f est une fonction sur $\mathbb{R}^3$ et si $\sigma\colon [a,b] \to \mathbb{R}^3$ est un chemin dans $\mathbb{R}^3$, l'intégrale de f sur σ est, par définition,

$$\int f\, d\sigma = \int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\|dt. \tag{20.279}$$

20.12.3 Intégrale d'une fonction sur une surface

Nous devons paramétrer la surface S par une application $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$. À partir d'un tel paramétrage, nous construisons un élément de surface en prenant le produit vectoriel des deux vecteurs tangents :

$$dS = \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} dudv. \tag{20.280}$$

L'intégrale est

$$\int f\, dS = \iint_D f\big(\varphi(u,v)\big) \left\| \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} \right\| dudv. \tag{20.281}$$

EqDefIntSurffS

Il ne faut pas rajouter de jacobien : la norme du produit vectoriel *est* le jacobien.

Remarque 20.81.
La formule (20.281) est autant valable pour des surfaces dans $\mathbb{R}^2$ que dans $\mathbb{R}^3$. Si nous considérons une surface dans $\mathbb{R}^2$, nous la voyons dans $\mathbb{R}^3$ en ajoutant un zéro comme troisième composante.

Exemple 20.82.
Les coordonnées polaires sont données par

$$\varphi(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix}. \tag{20.282}$$

Les vecteurs tangents à ce paramétrage sont

$$T_r = \frac{\partial\varphi}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}. \tag{20.283}$$

Le vecteur normal est

$$\frac{\partial\varphi}{\partial r} \times \frac{\partial\varphi}{\partial\theta} = \begin{vmatrix} e_x & e_y & e_z \\ \cos\theta & \sin\theta & 0 \\ -r\sin\theta & r\cos\theta & 0 \end{vmatrix} = re_z. \tag{20.284}$$

Nous trouvons donc que l'élément de surface est la norme de re_z, c'est-à-dire r, le jacobien connu. △

20.12.4 Intégrale d'une fonction sur un volume

Si V est un volume dans $\mathbb{R}^3$, nous effectuons la même procédure : nous trouvons un paramétrage, et nous formons un élément de volume avec les vecteurs tangents de le paramétrage. Nous avons donc un volume déterminé par l'application

$$\varphi\colon D \subset \mathbb{R}^3 \to \mathbb{R}^3, \tag{20.285}$$

et ses trois vecteurs tangents

$$\begin{aligned} T_u &= \frac{\partial\varphi}{\partial u} \\ T_v &= \frac{\partial\varphi}{\partial v} \\ T_w &= \frac{\partial\varphi}{\partial w}. \end{aligned} \tag{20.286}$$

Comment former un volume avec trois vecteurs ? Réponse : le produit mixte. L'intégrale de f sur V sera

$$\int f\,dV = \iiint_D f\big(\varphi(u,v,w)\big)\left\|\frac{\partial\varphi}{\partial u}\cdot\left(\frac{\partial\varphi}{\partial v}\times\frac{\partial\varphi}{\partial w}\right)\right\|dudvdw. \tag{20.287}$$

Encore une fois, le produit mixte *est* le jacobien. Prenons les coordonnées sphériques :

$$\begin{aligned} x(r,\theta,\varphi) &= r\sin(\theta)\cos(\varphi)\\ y(r,\theta,\varphi) &= r\sin(\theta)\sin(\varphi)\\ z(r,\theta,\varphi) &= r\cos(\theta) \end{aligned} \tag{20.288}$$

Les trois vecteurs tangents seront

$$T_r = \begin{pmatrix}\frac{\partial x}{\partial r}\\ \frac{\partial y}{\partial r}\\ \frac{\partial z}{\partial r}\end{pmatrix} = \begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix} \tag{20.289a}$$

$$T_\theta = \begin{pmatrix}\frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial \theta}\\ \frac{\partial z}{\partial \theta}\end{pmatrix} = \begin{pmatrix}r\cos(\theta)\cos(\varphi)\\ r\cos(\theta)\sin(\varphi)\\ -r\sin(\theta)\end{pmatrix} \tag{20.289b}$$

$$T_\varphi = \begin{pmatrix}\frac{\partial x}{\partial \varphi}\\ \frac{\partial y}{\partial \varphi}\\ \frac{\partial z}{\partial \varphi}\end{pmatrix} = \begin{pmatrix}-r\sin(\theta)\sin(\varphi)\\ r\sin(\theta)\cos(\varphi)\\ 0\end{pmatrix} \tag{20.289c}$$

Nous avons vu que le produit mixte revient à mettre toutes les composantes dans une matrice. Ici nous avons donc

$$\frac{\partial\phi}{\partial r}\cdot\left(\frac{\partial\phi}{\partial\theta}\times\frac{\partial\phi}{\partial\varphi}\right) = \begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r}\\ \frac{\partial x}{\partial\theta} & \frac{\partial y}{\partial\theta} & \frac{\partial z}{\partial\theta}\\ \frac{\partial x}{\partial\varphi} & \frac{\partial y}{\partial\varphi} & \frac{\partial z}{\partial\varphi}\end{vmatrix} \tag{20.290}$$

Cela est précisément le jacobien dont nous parlions plus haut.

20.12.5 Conclusion pour les fonctions

Lorsque nous intégrons une fonction sur un chemin, une surface ou un volume, la technique est toujours la même :

(1) Trouver un paramétrage à une, deux ou trois variables.

(2) Dériver le paramétrage par rapport à ses variables.

(3) Construire un élément de longueur, surface ou volume à partir des vecteurs que l'on a. Cela se fait en prenant la norme, le produit vectoriel ou le produit mixte.

20.12.6 Circulation d'un champ de vecteurs

Pour les champs de vecteurs, nous faisons la même chose, mais au lieu de *multiplier* par l'élément de longueur ou de surface, nous prenons le produit scalaire. Si nous considérons la courbe paramétrée $\sigma\colon [a,b]\to\mathbb{R}^3$ et le champ de vecteurs F, nous avons donc

$$\int_\sigma F = \int F\cdot d\sigma = \int_a^b F\big(\sigma(t)\big)\cdot\sigma'(t)dt. \tag{20.291}$$

20.12.7 Flux d'un champ de vecteurs

Si la surface $S\subset\mathbb{R}^3$ est paramétrée par

$$\begin{aligned} \phi\colon D\subset\mathbb{R}^2 &\to \mathbb{R}^3\\ (u,v) &\mapsto \phi(u,v), \end{aligned} \tag{20.292}$$

et si F est un champ de vecteurs, alors on a

$$\int_S F = \int_S F \cdot dS = \iint_D F\big(\phi(u,v)\big) \cdot \left(\frac{\partial\phi}{\partial u} \times \frac{\partial\phi}{\partial v}\right) dudv. \tag{20.293}$$

EqResIntFluxPhi

20.12.8 Conclusion pour les champs de vecteurs

La circulation et le flux ne représentent pas tout à fait la même chose. En effet pour la circulation, nous sélectionnons la composante *tangente* à la courbe, c'est-à-dire la partie du vecteurs qui « circule » le long de la courbe. Une force perpendiculaire au mouvement ne travaille pas.

La situation est exactement le contraire pour le flux. Étant donné que le vecteur

$$\frac{\partial\phi}{\partial u} \times \frac{\partial\phi}{\partial v} \tag{20.294}$$

est normal à la surface, le fait de prendre le produit scalaire du champ de vecteurs avec lui sélectionne la composante *normale* à la surface, c'est-à-dire la partie du vecteur qui traverse la surface.

20.12.9 Attention pour les surfaces fermées !

Si nous considérons une surface fermée, il faut faire attention à choisir une *orientation*. Les vecteurs normaux doivent soit tous pointer vers l'intérieur soit tous vers l'extérieur. En effet, en tant que vecteur normal, nous avons choisi de prendre

$$T_u \times T_v. \tag{20.295}$$

Mais le vecteur $T_v \times T_u$ est tout aussi normal ! Il n'y a pas à priori de façon standard pour choisir l'un ou l'autre. Il faut juste être cohérent : il faut que si on divise la surface en plusieurs morceaux, tous les vecteurs pointent dans le même sens.

Notez que si vous faites un choix et que votre voisin fait le choix inverse, vous obtiendrez des réponses qui diffèrent d'un signe. Sans plus de précisions [18], les deux réponses sont correctes.

Un exemple de ce problème est donné dans l'exemple 20.83.

EXooAJRLooSTPChN

Exemple 20.83.
Calculer le flux du champ de vecteurs

$$F(x,y,z) = e_x \tag{20.296}$$

au travers du cylindre de rayon R et de hauteur h autour de l'axe z.

Même question si le cylindre est autour de l'axe x.

Remarque : ces cylindres sont considérés *avec* leur « couvercles ».

Un paramétrage du cylindre autour de l'axe z est

$$\phi(\theta,z) = \begin{pmatrix} R\cos\theta \\ R\sin\theta \\ z \end{pmatrix}. \tag{20.297}$$

Les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \\ 0 \end{pmatrix}, T_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{20.298}$$

Le vecteur normal est donc

$$T_\theta \times T_z = R\cos(\theta)e_x + R\sin(\theta)e_y. \tag{20.299}$$

18. Il faudrait définir ce qu'est une surface *orientable* et choisir une orientation.

C'est un vecteur dirigé vers l'extérieur.

Le champ de vecteurs considéré est constant : $F(\theta, z) = e_x$. Nous avons donc

$$F(\theta, z) \cdot (T_\theta \times T_z) = R\cos(\theta) \tag{20.300}$$

et le flux vaut

$$\Phi = \int_0^{2\pi} d\theta \int_0^h R\cos(\theta) dz = 0. \tag{20.301}$$

En ce qui concerne les couvercles haut au bas, ils sont paramétrés par

$$\phi_1(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ h \end{pmatrix}, \phi_2(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ 0 \end{pmatrix}. \tag{20.302}$$

Les vecteurs normaux correspondants sont dans la direction de e_z, de façon que le produit scalaire avec $F(r, \theta)$ soit nul. Le flux total est donc nul.

Regardons maintenant le cylindre le long de l'axe x. Un paramétrage est

$$\phi(\theta, x) = \begin{pmatrix} x \\ R\cos(\theta) \\ R\sin(\theta) \end{pmatrix}, \tag{20.303}$$

et le vecteurs tangents sont

$$T_\theta = \begin{pmatrix} 0 \\ -R\sin\theta \\ R\cos\theta \end{pmatrix}, T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{20.304}$$

Le vecteur normal est alors donné par

$$T_\theta \times T_x = R\cos(\theta)e_y + R\sin(\theta)e_z. \tag{20.305}$$

Nous avons par conséquent $F(\theta, x) \cdot (T_\theta \times T_x) = 0$. Pas de flux par le côté du cylindre.

Regardons les « couvercles ». Le premier est donné par le paramétrage

$$\phi_1(r, \theta) = \begin{pmatrix} 0 \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{20.306}$$

Le vecteur normal serait $T_r \times T_\theta = re_x$, et le flux

$$\Phi = \int_0^{2\pi} d\theta \int_0^R r\, dr = \pi R^2. \tag{20.307}$$

Le second couvercle est donné par

$$\phi_2(r, \theta) = \begin{pmatrix} h \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{20.308}$$

Le vecteur normal est encore re_x, et le flux est à nouveau πR^2.

Le flux total serait donc $2\pi R^2$.

Cela n'est pas possible parce que tous les vecteurs qui « rentrent » d'un côté doivent « sortir » de l'autre côté. L'erreur est le le premier vecteur normal est un vecteur qui pointe vers l'intérieur du cylindre, tandis que le second pointe vers l'extérieur. Si nous choisissons, par convention, de prendre uniquement les vecteurs extérieurs, il faut changer le vecteur normal du premier couvercle en $-re_x$. Le premier flux vaudra donc

$$-\pi R^2, \tag{20.309}$$

de telle sorte que le flux total sera nul. △

20.13 Formes différentielles exactes et fermées

Nous avons déjà parlé de formes différentielles et de leurs intégrales sur un chemin dans la section 12.21.

DefEFKQmPs

Définition 20.84.
*La forme différentielle ω est **exacte** si il existe une fonction f telle que $\omega = df$; elle est dite **fermée** si $d\omega = 0$.*

Dire que la forme différentielle $\omega = f dx + g dy$ est fermée, c'est dire que

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}. \tag{20.310}$$

Il est naturel de se demander si toutes les formes différentielles sont des différentielles de fonctions. Une réponse complète est délicate à établir, mais a d'innombrables conséquences en physique, notamment en ce qui concerne l'existence d'un potentiel vecteur pour le champ magnétique dans les équations de Maxwell.

Le fait qu'une forme exacte soit fermée est relativement facile à établir ; c'est la proposition suivante. La question plus délicate est la réciproque : sous quelles conditions une forme fermée est-elle exacte ?

Proposition 20.85.
Si ω est une 1-forme exacte de classe C^1, alors ω est fermée.

Démonstration. Le fait que ω soit exacte implique l'existence d'une fonction f telle que $\omega = df$, c'est-à-dire

$$\omega_x = \sum_i a_i(x) dx_i = \sum_i \frac{\partial f}{\partial x_i}(x) dx_i, \tag{20.311}$$

c'est-à-dire que $a_i(x) = \frac{\partial f}{\partial x_i}(x)$. L'hypothèse que ω est C^1 implique que f est C^2, et donc que nous pouvons inverser l'ordre de dérivation pour les dérivées secondes $\partial^2_{ij} f = \partial^2_{ji} f$. Nous pouvons donc faire le calcul suivant :

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j} = \frac{\partial a_j}{\partial x_i}, \tag{20.312}$$

ce qu'il fallait démontrer. □

La réciproque est vraie sur un ouvert simplement connexe.

ThoFermeExactFormRappel

Théorème 20.86.
Supposons que $D \subset \mathbb{R}^n$ soit un ouvert simplement connexe. Alors toute forme différentielle de degré 1 et de classe C^1 sur D qui est fermée est exacte.

Nous allons prouver ce théorème dans un cas un peu moins général : celui d'un domaine étoilé de $\mathbb{R}^2$ plutôt que simplement connexe de $\mathbb{R}^n$.

ThoMSofFxL

Théorème 20.87.
Soit $D \subset \mathbb{R}^2$, une ouvert étoilé, et ω, une 1-forme fermée de classe C^1. Alors ω est exacte.

Démonstration. Soit $D \subset \mathbb{R}^2$, un ouvert étoilé par rapport à l'origine. Soient $f\colon D \to \mathbb{R}$, $g\colon D \to \mathbb{R}$, des fonctions de classe C^1 telles que

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \tag{20.313}$$

sur D, et

$$F(x,y) = \int_0^1 \big[f(tx,ty)x + g(tx,ty)y\big]dt \tag{20.314}$$

EqIMDefFformI33

pour tout $(x,y) \in D$.

Étant donné que nous ne définissons $F(x,y)$ que pour des $(x,y)\in D$, la fonction $t\mapsto f(tx,ty)$ est C^1 sur tout le compact $[0,1]$ et aucune divergence de l'intégrale n'est à craindre. Nous sommes donc dans le cadre de la proposition 17.24, et nous pouvons dériver sous le signe intégral.

Nous calculons, en utilisant la règle de dérivation de fonctions composées

$$\begin{aligned}\frac{\partial F}{\partial x}(x,y)&=\int_0^1\left[f\frac{\partial f}{\partial x}(tx,ty)x+f(tx,ty)+t\frac{\partial g}{\partial x}(tx,ty)y\right]dt\\&=\int_0^1\left[t\Big(x\frac{\partial f}{\partial x}(tx,ty)+y\frac{\partial f}{\partial y}(tx,ty)\Big)+f(tx,ty)\right]dt\end{aligned}\tag{20.315}$$

où nous avons utilisé l'hypothèse $\partial_y f=\partial_x g$. Ce qui se trouve dans la parenthèse n'est autre que $\partial_t\big(f(tx,ty)\big)$, plus précisément, si nous posons $\mathcal{F}(x,y,t)=f(tx,ty)$, nous avons

$$\frac{\partial\mathcal{F}}{\partial t}(x,y,t)=x\frac{\partial f}{\partial x}(tx,ty)+y\frac{\partial f}{\partial y}(tx,ty).\tag{20.316}$$

En recopiant le résultat (20.315) en termes de $\mathcal{F}$, nous avons

$$\begin{aligned}\frac{\partial F}{\partial x}(x,y)&=\int_0^1\left(t\frac{\partial\mathcal{F}}{\partial t}(x,y,t)+\mathcal{F}(x,y,t)\right)dt\\&=\int_0^1\partial_t\big(t\mathcal{F}(x,y,t)\big)dt\\&=\big[t\mathcal{F}(x,y,t)\big]_0^1\\&=\mathcal{F}(x,y,1)\\&=f(x,y).\end{aligned}\tag{20.317}$$

Le résultat correspondant pour $\frac{\partial F}{\partial y}(x,y)=g(x,y)$ s'obtient de la même manière. Nous avons donc obtenu que

$$\frac{\partial F}{\partial x}=f,\quad\text{et}\quad\frac{\partial F}{\partial y}=g.\tag{20.318}$$

En ayant prouvé cela, nous avons prouvé que si $\omega=fdx+gdy$ avec $\partial_y f=\partial_x g$, alors $\omega=dF$ où F est définie par (20.314). □

Démonstration alternative du théorème 20.87. Nous posons $u=tx$ et $v=ty$, ainsi que $\mathcal{F}(x,y,t)=f(u,v)$ et $\mathcal{G}(x,y,t)=g(u,v)$. Avec cette notation, nous avons

$$F(x,y)=\int_0^1\big(x\mathcal{F}(x,y,t)+y\mathcal{G}(x,y,t)\big)dt,\tag{20.319}$$

et

$$\begin{aligned}\frac{\partial\mathcal{F}}{\partial x}&=\frac{\partial f}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial x}=t\frac{\partial f}{\partial u},\\\frac{\partial\mathcal{G}}{\partial x}&=t\frac{\partial g}{\partial u}.\end{aligned}\tag{20.320}$$

Ainsi,

$$\begin{aligned}\frac{\partial F}{\partial x}&=\int_0^1\left(x\frac{\partial\mathcal{F}}{\partial x}+\mathcal{F}+y\frac{\partial G}{\partial x}\right)dt\\&=\int_0^1\left(xt\frac{\partial f}{\partial u}+\mathcal{F}+yt\frac{\partial g}{\partial u}\right)dt\\&=\int_0^1\left[t\left(x\frac{\partial f}{\partial u}+y\frac{\partial f}{\partial v}\right)+\mathcal{F}\right]dt.\end{aligned}\tag{20.321}$$

où nous avons utilisé le fait que, par hypothèse, $\frac{\partial g}{\partial u}=\frac{\partial f}{\partial v}$. Nous calculons par ailleurs que

$$\frac{\partial F}{\partial t}=\frac{\partial f}{\partial u}\frac{\partial u}{\partial t}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial t}=x\frac{\partial f}{\partial u}+y\frac{\partial f}{\partial v}.\tag{20.322}$$

Donc, nous avons

$$\frac{\partial F}{\partial x} = \int_0^1 \left(t\frac{\partial \mathcal{F}}{\partial t} + \mathcal{F} \right) dt = \int_0^1 \frac{\partial}{\partial t}(t\mathcal{F})dt. \tag{20.323}$$

Par conséquent,

$$\frac{\partial F}{\partial x} = [t\mathcal{F}]_0^1 = \mathcal{F}(x, y, 1) = f(x, y). \tag{20.324}$$

Le même genre de calculs fournit $\frac{\partial F}{\partial y} = g(x, y)$. □

20.14 Théorème d'Abel angulaire

ThoTGjmeen

Théorème 20.88 (Abel angulaire[104])**.**
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence plus grand ou égal à 1 et de somme f. Soit $\theta_0 \in [0, \frac{\pi}{2}[$. Nous posons

$$\Delta_{\theta_0} = \{z = 1 - \rho e^{i\varphi} \text{ tel que } \rho > 0, \varphi \in [-\theta_0, \theta_0], |z| < 1\}. \tag{20.325}$$

Nous supposons de plus que $\sum_n a_n$ converge. Alors

$$\lim_{\substack{z\to 1 \\ z\in\Delta_0}} f(z) = \sum_{k=0}^{\infty} a_k. \tag{20.326}$$

Démonstration. Le résultat de ce théorème est que l'on peut calculer la limite $z \to 1$ avec des chemins contenus dans un domaine de la forme de celui dessiné à la figure 20.4.

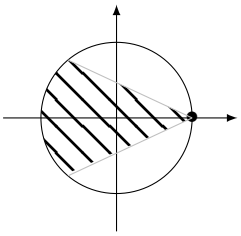


FIGURE 20.4: La zone dans laquelle peut être le chemin qui va vers $z = 1$.LabelFigJGuKEjH

De façon très classique nous posons

$$S = \sum_{k=0}^{\infty} a_k \quad S_n = \sum_{k=0}^{n} a_k, \tag{20.327}$$

et $R_n = S - S_n$. En particulier $a_n = R_{n-1} - R_n$.

Le but du théorème est de montrer que $\sum a_n z^n$ converge vers S lorsque z converge vers 1 à l'intérieur de Δ_{θ_0}. Pour cela nous calculons pour un N donné la différence $\sum_{n=0}^{N} a_n z^n - S_N$ en triant les termes par ordre de R_n, en isolant le terme R_0 et le terme R_N :

$$\sum_{n=0}^{N} a_n z^n - S_N = \sum_{n=1}^{N} a_n(z^n - 1) \tag{20.328a}$$

$$= \sum_{n=1}^{N} (R_{n-1} - R_n)(z^n - 1) \tag{20.328b}$$

$$= R_0(z-1) + \sum_{n=1}^{N-1} R_n(z^{n+1} - 1 - z^n + 1) + R_N(z^N - 1) \tag{20.328c}$$

$$= R_0(z-1) + \sum_{n=1}^{N-1} R_n z^n(z-1) + R_N(z^N - 1) \tag{20.328d}$$

$$= (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). \tag{20.328e}$$

Cela est valable pour tout N et $|z|<1$. Nous avons donc

$$\sum_{n=0}^{N} a_n z^n - S_N = (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N-1). \tag{20.329}$$

Par hypothèse nous avons $\lim_{N\to\infty} R_N = 0$. Et de plus le membre de gauche converge parce que chacun des deux termes converge séparément. En passant à la limite nous avons pour tout $|z|<1$:

$$f(z) - S = (z-1)\sum_{n=0}^{\infty} R_n z^n. \tag{20.330}$$

Nous voudrions étudier le comportement de la différence $f(z)-S$ lorsque z tend vers 1. Pour cela nous nous fixons $\epsilon>0$ et $N\geqslant 1$ tel que $|R_n|<\epsilon$ dès que $n\geqslant N$. Alors pour tout $|z|<1$ nous avons

$$|f(z)-S| \leqslant |z-1|\left(\sum_{n=0}^{N}|R_n|\underbrace{|z^n|}_{\leqslant 1} + \sum_{n=N+1}^{\infty}\underbrace{|R_n|}_{\leqslant\epsilon}|z^n|\right) \tag{20.331a}$$

$$\leqslant |z-1|\sum_{n=0}^{N}|R_n| + \epsilon\frac{|z-1|}{1-|z|} \tag{20.331b}$$

où nous avons utilisé la somme de la série géométrique (1.687) et l'égalité $|z^n|=|z|^n$. Avant de nous particulariser à $z\in\Delta_{\theta_0}$ nous devons anticiper un problème au dénominateur en multipliant par le binôme conjugué :

$$\frac{|z-1|}{1-|z|} = \frac{|z-1|(1+|z|)}{1-|z|^2}. \tag{20.332}$$

C'est maintenant que nous nous particularisons à $z\in\Delta_{\theta_0}$ en posant $z=\rho e^{i\varphi}$ et en remarquant que $|z|^2 = 1-2\rho\cos(\varphi)+\rho^2$. Nous avons le calcul suivant :

$$\frac{|z-1|}{1-|z|} = \frac{\rho(1+|z|)}{2\rho\cos(\varphi)-\rho^2} \tag{20.333a}$$

$$= \frac{1+|z|}{2\cos(\varphi)-\rho} \tag{20.333b}$$

$$\leqslant \frac{2}{2\cos(\varphi)-\rho} \tag{20.333c}$$

$$\leqslant \frac{2}{2\cos(\varphi)-\cos(\theta_0)} \tag{20.333d}$$

$$\leqslant \frac{2}{2\cos(\theta_0)-\cos(\theta_0)} \tag{20.333e}$$

$$= \frac{2}{\cos(\theta_0)}. \tag{20.333f}$$

Quelques justifications.

— Vu que nous avons dans l'idée de faire $\rho\to 0$ nous supposons que $\rho<\cos(\theta_0)$.

— Nous avons $\cos(\varphi)>\cos(\theta_0)$ parce que z est dans Δ_{θ_0}.

Nous avons donc, pour tout $z\in\Delta_{\theta_0}$ que

$$|f(z)-S| \leqslant |z-1|\sum_{n=0}^{N}|R_n| + \epsilon\frac{2}{\cos(\theta_0)}. \tag{20.334}$$

Il suffit de prendre ρ assez petit pour que

$$|z-1|\sum_{n=0}^{N}|R_n| < \epsilon \tag{20.335}$$

et nous avons

$$|f(z)-S|\leqslant\epsilon\left(1+\frac{2}{\cos(\theta_0)}\right). \tag{20.336}$$

Nous avons donc bien $\lim_{\substack{z\to 1\\ z\in\Delta_0}} f(z)=S$, comme nous le voulions. □

La réciproque du théorème d'Abel angulaire est que si $f(z)=\sum_n a_n z^n$ sur $B(0,1)$ se prolonge par continuité en $z=1$ alors cette prolongation se fait par $f(1)=\sum_n a_n$. Cela est faux comme le montre l'exemple suivant.

Exemple 20.89.
Nous considérons la série entière $\sum_{n=0}^{\infty}(-1)^n z^n$ qui converge[19] vers

$$f(z)=\frac{1}{1+z} \tag{20.337}$$

sur $B(0,1)$. De plus nous avons

$$\lim_{\substack{z\to 1\\ |z|<1}}\frac{1}{1+z}=\frac{1}{2}. \tag{20.338}$$

Donc la fonction converge bien vers quelque chose lorsque z tend vers 1. La fonction f se prolonge par continuité en 1. Pourtant la série ses coefficients $\sum_n(-1)^n$ ne converge pas. △

Le théorème suivant donne une espèce d'inverse au théorème d'Abel angulaire. En effet il dit que si la série converge en allant vers 1 le long de l'axe réel, alors ça converge vers la somme des coefficients. Il faut cependant une hypothèse en plus sur les a_n.

Théorème 20.90 (Théorème taubérien faible[104])**.**
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1 et de somme f. Nous supposons

(1) Il existe $S\in\mathbb{C}$ tel que $\lim_{\substack{x\to 1\\ x\in]-1,1[}} f(x)=S$.

(2) $\lim_{n\to\infty} na_n=0$.

Alors la série $\sum_{n=0}^{\infty}a_n$ converge et vaut S.

Démonstration. Nous notons $S_n=\sum_{k=0}^{n}a_k$ et $M=\sup_{k\geqslant 1}k|a_k|$, qui est fini par hypothèse. Pour $x\in]0,1[$ et $n\geqslant 0$ nous avons

$$S_n-f(x)=\sum_{k=1}^{n}a_k-\sum_{k=1}^{n}a_kx^k-\sum_{k=n+1}^{\infty}a_kx^k=\sum_{k=1}^{n}a_k(1-x^k)-\sum_{k=n+1}^{\infty}a_kx^k. \tag{20.339}$$

Nous utilisons la série géométrique sous la forme $1-x^k=(1-x)\sum_{i=0}^{n}x^i$ pour écrire

$$S_n-f(x)=\sum_{k=1}^{n}a_k(1-x)\underbrace{\sum_{i=0}^{k-1}x^i}_{\leqslant k}-\sum_{k=n+1}^{\infty}a_kx^k \tag{20.340a}$$

$$\leqslant\sum_{k=1}^{n}ka_k(1-x)-\sum_{k=n+1}^{\infty}a_kx^k, \tag{20.340b}$$

19. C'est la série géométrique de raison $-z$.

donc en passant à la norme

$$\big|S_n - f(x)\big| \leqslant (1-x)Mn + \sum_{k=n+1} |a_k| x^k \tag{20.341a}$$

$$\leqslant (1-x)Mn + \sum_{k=n+1}^{\infty} \underbrace{\frac{k}{n}|a_k|}_{\leqslant M/n} x^k \tag{20.341b}$$

$$\leqslant (1-x)Mn + \frac{M}{n} \sum_{k=n+1}^{\infty} x^k \tag{20.341c}$$

$$\leqslant (1-x)Mn + \frac{M}{n} \frac{1}{1-x}. \tag{20.341d}$$

Ce que nous cherchons à étudier est le comportement $x \to 1$ et montrer que $S_n \to S$, ce qui nous incite à calculer $|S_n - f(1 - \frac{\epsilon}{n})|$ avec $0 < \epsilon < 1$:

$$\big|S_n - f\big(1 - \frac{\epsilon}{n}\big)\big| \leqslant \epsilon M + \epsilon. \tag{20.342}$$

Nous choisissons N_1 tel que $\frac{M}{n} \leqslant \epsilon^2$ dès que $n \geqslant N_1$. En sus nous savons que

$$\lim_{\epsilon \to 0} f(1-\epsilon) = S. \tag{20.343}$$

Nous choisissons N_2 de telle sorte à avoir

$$\left|f\left(1 - \frac{\epsilon}{n}\right) - S\right| < \epsilon, \tag{20.344}$$

et en prenant $n \geqslant \max(N_1, N_2)$ nous avons

$$|S_n - S| \leqslant \left|S_n - f\left(1 - \frac{\epsilon}{n}\right)\right| + \left|f\left(1 - \frac{\epsilon}{n}\right) - S\right| \leqslant \epsilon M + 2\epsilon. \tag{20.345}$$

Il suffit de choisir ϵ suffisamment petit (en particulier pour que ϵM soit petit) pour montrer que $|S_n - S|$ est borné par un nombre arbitrairement petit. □

20.15 Passage à la limite sous le signe intégral

Un autre résultat très important pour l'étude de l'intégrabilité est le théorème de la **convergence dominée de Lebesgue** :

Théorème 20.91.
Soit $E \subset \mathbb{R}^n$ un ensemble mesurable et $\{f_k\}$, une suite de fonctions intégrables sur E qui converge simplement vers une fonction $f\colon E \to \overline{\mathbb{R}}$. Supposons qu'il existe une fonction g intégrable sur E telle que pour tout k,

$$|f(x)| \leqslant g(x) \tag{20.346}$$

pour tout $x \in E$. Alors f est intégrable sur E et

$$\int_E f = \lim_{k \to \infty} \int_E f_k. \tag{20.347}$$

20.15.1 Intégrale en dimension un

Proposition 20.92 (Critère de comparaison)**.**
Soit f mesurable sur $]a, \infty[$ et bornée sur tout $]a, b]$, et supposons qu'il existe un $X_0 \geqslant a$, tel que sur $]X_0, \infty[$,

$$|f(x)| \leqslant g(x) \tag{20.348}$$

où $g(x)$ est intégrable. Alors $f(x)$ est intégrable sur $]a, \infty[$.

Corolaire 20.93 (Critère d'équivalence)**.**
Soient f et g des fonctions mesurables et positives ou nulles sur $]a,\infty[$, bornées sur tout $]a,b]$, telles que

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=L \tag{20.349}$$

existe dans $\bar{\mathbb{R}}$.

(1) Si $L\neq\infty$ et $\int_a^\infty g(x)dx$ existe, alors $\int_a^\infty f(x)dx$ existe,
(2) Si $L\neq 0$ et si $\int_a^\infty f(x)dx$ existe, alors $\int_a^\infty g(x)dx$ existe,

CorCritFonsTest

Corolaire 20.94 (Critère des fonctions test)**.**
Soit $f(x)$ une fonction mesurable et positive ou nulle sur $]a,\infty[$ et bornée pour tout $]a,b]$. Nous posons

$$L(\alpha)=\lim_{x\to\infty}x^\alpha f(x), \tag{20.350}$$

et nous supposons qu'elle existe.

(1) Si il existe $\alpha>1$ tel que $L(\alpha)\neq\infty$, alors $\int_a^\infty f(x)dx$ existe,
(2) Si il existe $\alpha\leqslant 1$ et $L(\alpha)\neq 0$, alors $\int_a^\infty f(x)dx$ n'existe pas.

CorAlphaLCasInteabf

Corolaire 20.95.
Soit $f\colon]a,b]\to\mathbb{R}$ une fonction mesurable, positive ou nulle, et bornée sur $[a+\epsilon,b]$ $\forall\epsilon>0$. Si $\lim_{x\to a}(x-a)^\alpha f(x)=L$ existe, alors

(1) Si $\alpha<1$ et $L\neq\infty$, alors $\int_a^b f(x)dx$ existe,
(2) Si $\alpha\geqslant 1$ et $L\neq 0$, alors $\int_a^b f(x)dx$ n'existe pas.

20.15.2 Intégrales convergentes

Définition 20.96.
Soit f, une fonction mesurable sur $[a,\infty[$, bornée sur tout intervalle $[a,b]$. On dit que l'intégrale

$$\int_a^\infty f(x)dx \tag{20.351}$$

converge *si la limite*

$$\lim_{X\to\infty}\int_a^X f \tag{20.352}$$

EqDEfConvergeZeroInftX

existe et est finie.

20.15.3 La méthode de Rothstein-Trager

subSecBCRYooRVjFpS

Mes sources pour parler d'intégration de fractions rationnelles : [593].

ThoXJFatfu

Théorème 20.97 (Rothstein-Trager[594])**.**
Soient $P,Q\in\mathbb{Q}[X]$ premiers entre eux avec $\operatorname{pgcd}(P,Q)=1$ et $\deg(P)<\deg(Q)$. Nous supposons que Q est unitaire et sans facteurs carrés. Supposons que nous puissions écrire, dans un extension $\mathbb{K}$ de $\mathbb{Q}$ la primitive de P/Q de la façon suivante :

$$\int\frac{P}{Q}=\sum_{i=1}^n c_i\ln(P_i) \tag{20.353}$$

EqCHVaDay

où les c_i sont des constantes non nulles et deux à deux distinctes et où les P_i sont des polynômes unitaires non constants sans facteurs carrés et premiers deux à deux entre eux dans $\mathbb{K}[X]$.

Alors les c_i sont les racines distinctes du polynôme

$$R(Y)=\operatorname{res}_X(P-YQ',Q)\in\mathbb{K}[Y] \tag{20.354}$$

et

$$P_i=\operatorname{pgcd}(P-c_iQ',Q). \tag{20.355}$$

Démonstration. Nous posons

$$U_i = \prod_{j\neq i} P_j. \tag{20.356}$$

(1) **Question de division** Ensuite nous dérivons formellement l'équation (20.353) et nous multiplions les deux côtés du résultat par $\prod_{j=1}^{n} P_j$:

$$P\prod_{j=1}^{n} P_j = Q\sum_{i=1}^{n} c_i \frac{P'i}{P_i}\prod_{j=1}^{n} P_j = Q\sum_{i=1}^{n} c_i P_i' U_i. \tag{20.357}$$

Une première chose que nous en tirons est que Q divise le produit $P\prod_{j=1}^{n} P_j$; mais P et Q étant premiers entre eux,

$$Q \mid \prod_{j=1}^{n} P_j \tag{20.358}$$

par le théorème de Gauss 6.50.

Une seconde chose que nous tirons de (20.357) est que P_j divise $Q\sum_{i=1}^{n} c_i P_i' U_i$. De cette somme, à cause du U_i qui est divisé par P_j pour tout i sauf $i = j$, le polynôme P_j divise tous les termes sauf peut-être un. Donc il les divise tous et en particulier

$$P_j \mid Qc_j P_J' U_j \tag{20.359}$$

En nous souvenant que les P_k sont premiers entre eux, P_j ne divise pas U_j. De plus P_j étant sans facteurs carrés, les polynômes P_j et P_j' sont premiers entre eux. Il ne reste que Q. Nous en déduisons que

$$P_j \mid Q \tag{20.360}$$

pour tout $1 \leqslant j \leqslant n$. Et vu que les P_i sont premiers entre eux, le fait que chacun divise Q implique que leur produit divise Q, c'est-à-dire

$$\prod_{j=1}^{n} P_j \mid Q. \tag{20.361}$$

Or nous avions déjà prouvé la division contraire. Du fait que les deux polynômes sont unitaires nous en déduisons qu'ils sont en réalité égaux :

$$Q = \prod_{j=1}^{n} P_j. \tag{20.362}$$

Nous pouvons simplifier les deux membres de (20.357) par cela :

$$P = \sum_{i=1}^{n} c_i P_i' U_i. \tag{20.363}$$

(2) **Encore un peu de division**
En dérivant (20.362) nous trouvons

$$Q' = \sum_{j=1}^{n} P_j' U_j, \tag{20.364}$$

et en écrivant P sous sa forme (20.363),

$$P - c_i Q' = \sum_{j=1}^{n} c_j P_j' U_j - \sum_{j=1}^{n} c_i P_j' U_j = \sum_{j=1}^{n} (c_j - c_i) P_j' U_j. \tag{20.365}$$

Le terme $i = j$ de la somme est nul ; en ce qui concerne les autres termes, ils sont divisés par P_i parce que $P_i \mid U_j$. Donc P_i divise tous les termes de la somme et nous avons

$$P_i \mid P - c_i Q'. \tag{20.366}$$

(3) **Un pgcd pour continuer**

Nous montrons à présent que $P_i = \text{pgcd}(P - c_iQ', Q)$. Pour cela nous utilisons la multiplicativité du PGCD lorsque les facteurs sont premiers entre eux :

$$\text{pgcd}(P - c_iQ', Q) = \text{pgcd}(P - c_iQ', \prod_{j=1}^{n} P_j) = \prod_{j=1}^{n} \text{pgcd}(P - c_iQ', P_j). \tag{20.367}$$

Nous remplaçons $P - c_iQ'$ par son expression (20.365) et nous écrivons un des facteurs du produit :

$$\text{pgcd}(P - c_iQ', P_j) = \text{pgcd}(\sum_{k=1}^{n} (c_k - c_i)P'_kU_k, P_j) \tag{20.368}$$

Le polynôme P_j divise tous les U_k sauf celui avec $k = j$. Donc le lemme 6.57(1) nous permet de dire

$$\text{pgcd}(P - c_iQ', P_j) = \text{pgcd}\big((c_j - c_i)P'_jU_j, P_j\big) = \begin{cases} 1 & \text{si } i \neq j \\ P_j & \text{si } i = j. \end{cases} \tag{20.369}$$

La seconde ligne provient du fait que nous ayons déjà montré que $P_j \mid P - c_jQ'$. En fin de compte,

$$\text{pgcd}(P - c_iQ', Q) = P_i. \tag{20.370}$$

(4) **Une histoire de résultant**

Les nombres c_i sont tels que les polynômes $P - c_iQ'$ et Q ne sont pas premiers entre eux. Vu que les P_i sont non nuls, la proposition 8.25 nous dit que le résultant

$$\text{res}_X(P - c_iQ', Q) = 0. \tag{20.371}$$

Donc les c_i sont des racines du polynôme (en Y)

$$R(Y) = \text{res}_X(P - YQ', Q). \tag{20.372}$$

EqQOjnwJj

Nous n'avons pas prouvé qu'ils étaient *toutes* les racines [20].

(5) **Toutes les racines**

Nous allons maintenant montrer que les c_i étaient toutes les racines imaginables du polynôme (20.372) dans toutes les extensions de $\mathbb{Q}$. Soit donc c une racine de R dans une extension $\hat{\mathbb{K}}$ de $\mathbb{K}$ qui ne soit pas parmi les c_i de la formule (20.353). Étant donné que c est racine du résultat, les polynômes $P - cQ'$ et Q ont un PGCD non trivial, c'est-à-dire non constant. Donc

$$\text{pgcd}(P - cQ', Q) = S \in \hat{\mathbb{K}}[X] \tag{20.373}$$

est un polynôme non constant. Si T un facteur irréductible de S, alors T divise $P - cQ'$ et Q, mais $Q = \prod_{i=1}^{n} P_i$ avec les P_i premiers entre eux. Donc T ne peut diviser que l'un (et exactement un) d'entre eux [21]. Soit P_{i_0} celui qui est divisé par T. La relation (20.365) dans ce contexte donne :

$$P - cQ' = \sum_{j=1}^{n} (c_j - c)P'_jU_j \tag{20.374}$$

Le polynôme T divise tous les U_j avec $j \neq i_0$, mais comme en plus il divise $P - cQ'$, il divise aussi le dernier terme de la somme :

$$T \mid (c_{i_0} - c)P'_{i_0}U_{i_0}. \tag{20.375}$$

Le polynôme T ne divisant pas U_{i_0} et $(c_{i_0} - c)$ étant non nul, nous concluons que T divise P'_{i_0}. Mais cela n'est pas possible parce que nous avons supposé que P_{i_0} était sans facteur carré, ce qui voulait entre autres dire que P_{i_0} et P'_{i_0} n'ont pas de facteurs communs.

20. De plus, nous n'avons pas de garanties que ces racines soient dans $\mathbb{Q}$, et en fait il y a des cas dans lesquels les c_i n'y sont pas.

21. On ne peut pas diviser deux trucs qui sont premiers entre eux ; c'est une question de cohérence, madame !

□

Ce théorème suggère la méthode suivante pour trouver la primitive de la fraction rationnelle P/Q (si elle vérifie les hypothèses)

(1) Écrire le résultant $R(y) = \text{res}_X(P - yQ', Q)$ et en trouver les racines $\{c_i\}_{i=1,\dots,n}$.

(2) Calculer les polynômes $P_i = \text{pgcd}(P - c_iQ', Q)$.

(3) Écrire la réponse :

$$\int \frac{P}{Q} = \sum_{i=1}^{n} c_i \ln(P_i). \tag{20.376}$$

Notons que le polynôme $R(Y)$ est de degré $\deg(Q)$ (pour le voir, faire un peu de comptage de lignes et colonnes dans la matrice de Sylvester), donc il n'est à priori pas pire à factoriser que Q lui-même [22]. Mais il se peut que nous ayons de la chance et que R soit plus facile que Q.

À part qu'on a peut-être plus de chance avec R qu'avec Q, l'avantage de la méthode est qu'elle permet d'éviter de passer par des extensions de $\mathbb{Q}$ non nécessaires [23].

Exemple 20.98 ([594])**.**
Prenons la fraction rationnelle $\frac{x}{x^2-3}$. L'intégration via les fractions simples est :

$$\int \frac{x}{x^2-3}dx = \frac{1}{2}\int \frac{1}{x-\sqrt{3}} + \frac{1}{2}\int \frac{1}{x+\sqrt{3}} = \frac{1}{2}\ln(x-\sqrt{3}) + \frac{1}{2}\ln(x+\sqrt{3}) = \frac{1}{2}\ln(x^2-3). \tag{20.377}$$

Nous voyons que dans la réponse, il n'y a pas de racines. Passer par l'extension $\mathbb{Q}[\sqrt{3}]$ est par conséquent peut-être un effort inutile. Voyons comment les choses se mettent avec la méthode Rothstein-Trager.

D'abord

$$\begin{aligned}
R(Y) &= \text{res}_X(X - 2YX, X^2 - 3) && \text{(20.378a)}\\
&= \text{res}_X\big((1-2Y)X, X^2-3\big) && \text{(20.378b)}\\
&= \det\begin{pmatrix} 1-2Y & 0 & 0 \\ 0 & 1-2Y & 0 \\ 1 & 0 & -3 \end{pmatrix} && \text{(20.378c)}\\
&= (1-2Y)\big(-3(1-2Y)\big) && \text{(20.378d)}\\
&= -3(1-2Y)^2, && \text{(20.378e)}
\end{aligned}$$

dont les solutions sont faciles : il n'y a que la racine double $y = \frac{1}{2}$. La somme (20.353) sera donc réduite à un seul terme avec $c_1 = \frac{1}{2}$. Nous calculons P_1 :

$$P_1 = \text{pgcd}(X - \frac{1}{2}2X, X^2 - 3) = \text{pgcd}(0, X^2 - 3) = X^2 - 3, \tag{20.379}$$

et par conséquent

$$\int \frac{X}{X^2-3} = \frac{1}{2}\ln(X^2-3). \tag{20.380}$$

À aucun moment nous ne sommes sortis de $\mathbb{Q}$. △

Comme vu sur cet exemple, l'intérêt du théorème de Rothstein-Trager est de permettre, lorsqu'on a de la chance, d'en profiter, et non de nous en rendre compte à la fin en remarquant bêtement que la réponse pouvait s'écrire dans $\mathbb{Q}[X]$.

22. C'est de la factorisation de Q qu'on a besoin pour utiliser la méthode de décomposition en fractions simples.
23. J'imagine que pour un ordinateur, c'est plus facile d'éviter les extensions.

Remarque 20.99.
Afin d'utiliser cette méthode, il faut s'assurer que Q soit sans facteurs carrés. Si nous devons intégrer un $\frac{P}{Q}$ quelconque, nous devons commencer par écrire

$$Q = Q_1 Q_2^2 Q_3^3 \dots Q_r^r, \tag{20.381}$$

et ensuite il y a moyen de ramener l'intégrale de P/Q à des intégrales de $\frac{P}{Q_1 \dots Q_r}$. Cela ne demande pas de factoriser complètement Q, mais seulement de trouver ses facteurs irréductibles Q_i dans $\mathbb{Q}[X]$.

Dans l'exemple donné plus haut, $Q = X^2 - 3$ a des facteurs irréductibles autres que Q lui-même dans $\mathbb{R}[X]$, mais nous n'en avons pas besoin.

Voici une exemple où nous évitons de passer par les complexes.

EXooIPEQooGKDjea

Exemple 20.100 ([595]).
À calculer : $\int \frac{1}{x^3+x}$. La décomposition en fractions simples donne :

$$\frac{1}{x^3+x} = \frac{1}{x} - \frac{1/2}{x-i} - \frac{1/2}{x+i}. \tag{20.382}$$

Déjà cette décomposition passe par l'extension $\mathbb{Q}[i]$, et le calcul de la primitive de $\frac{1}{x+i}$ demande le logarithme complexe qui ne sera vu que dans la proposition 26.42 au prix d'un peu de sang. Nous verrons dans l'exemple 26.43 comment ça se passe en passant par les complexes.

En ce qui concerne la méthode de Rothstein-Trager, nous commençons par calculer le résultant (qui est tout de même un peu de calcul) :

$$P(y) = \operatorname{res}_X\big(-3yX^2 - y + 1, X^3 - X\big) \tag{20.383a}$$

$$= \det \begin{pmatrix} -3y & 0 & 1-y & 0 & 0 \\ 0 & -3y & 0 & 1-y & 0 \\ 0 & 0 & -3y & 0 & 1-y \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \tag{20.383b}$$

$$= -(y-1)^2(2y+1)^2 \tag{20.383c}$$

```
----------------------------------------------------------------------
| Sage Version 5.7, Release Date: 2013-02-19                         |
| Type "notebook()" for the browser-based notebook interface.        |
| Type "help()" for help.                                            |
----------------------------------------------------------------------
sage: y=var('y')
sage: R=matrix(5,5,[-3*y,0,1-y,0,0,0,-3*y,0,1-y,0,0,0,
                    -3*y,0,1-y,1,0,1,0,0,0,1,0,1,0])
sage: R.determinant().factor()
-(y - 1)*(2*y + 1)^2
```

Les solutions sont $c_1 = 1$ et $c_2 = -\frac{1}{2}$. Nous pouvons alors calculer les P_i :

$$P_1 = \operatorname{pgcd}(-3X^2, X^3 + X) = X \tag{20.384}$$

et

$$P_2 = \operatorname{pgcd}(\frac{3}{2}X^2 + \frac{3}{2}, X^3 + X) = X^2 + 1, \tag{20.385}$$

et finalement

$$\int \frac{1}{X^3 + X} = \ln(X) - \frac{1}{2}\ln(X^2 + 1). \tag{20.386}$$

△

Notons qu'il n'y a pas de miracles : lorsque la réponse contient des racines, nous ne pouvons pas couper à passer par des extensions et factoriser un peu à la dure.

ExYQODuyU

Exemple 20.101 ([595]).
Nous voulons calculer

$$\int \frac{1}{X^2+1}. \tag{20.387}$$

Nous posons donc $P=1$ et $Q=X^2+1$. Le résultant à calculer est

$$P(y)=\operatorname{res}_X(-2yX+1,X^2+1)=\det\begin{pmatrix}-2y & 1 & 0\\ 0 & -2y & 1\\ 1 & 0 & 1\end{pmatrix}=4y^2+1. \tag{20.388a}$$

Les racines de cela sont complexes et il n'y a donc pas d'échappatoires : $c_1=\frac{i}{2}$, $c_2=-\frac{i}{2}$. Ensuite, étant donné que $X^2+1=(X+i)(X-i)=i(-iX+1)(X-i)$ nous avons

$$P_1=\operatorname{pgcd}(-iX+1,X^2+1)=X+i. \tag{20.389}$$

Notons que de façon naturelle, nous aurions écrit $P_1=1-iX$, mais par convention nous considérons le PGCD unitaire. Cela ne change rien à la réponse parce que changer P_i en kP_i ne fait que rajouter une constante $\ln(k)$ à la primitive trouvée.

De la même façon,

$$P_2=\operatorname{pgcd}(1+iX,X^2+1)=X-i. \tag{20.390}$$

Au final nous écrivons

$$\int \frac{1}{X^2+1}=\frac{i}{2}\ln(X+i)-\frac{i}{2}\ln(X-i). \tag{20.391}$$

△

Remarque 20.102.
Tout cela est si nous voulons absolument écrire la primitive avec des logarithmes de polynômes. Pour celui de l'exemple 20.101, nous avons trouvé

$$\int \frac{1}{X^2+1}=\frac{i}{2}\ln(X+i)-\frac{i}{2}\ln(X-i). \tag{20.392}$$

EqBCjCCbs

Mais

```
sage: f(x)=1/(x**2+1)
sage: f.integrate(x)
x |--> arctan(x)
```

Si nous acceptons de passer aux fonctions trigonométriques (inverses), la primitive prend un tour très différent et bien réel. Ces deux visions de l'univers sont bien entendu [24] compatibles. En effet, afin de tomber juste, nous allons prendre la primitive

$$f(x)=\frac{i}{2}\ln(ix-1)-\frac{i}{2}\ln(ix+1) \tag{20.393}$$

au lieu de (20.392). Il s'agit seulement de multiplier l'intérieur des logarithmes, ce qui ne donne qu'une constante de différence. Ensuite nous passons à la forme trigonométrique des nombres complexes : $ix-1=\sqrt{x^2+1}e^{i\arctan(-x)}$ et $ix+1=\sqrt{x^2+1}e^{i\arctan(x)}$. Avec un peu de calcul,

$$f(x)=-\frac{1}{2}\Big(\arctan(-x)-\arctan(x)\Big)=\arctan(x). \tag{20.394}$$

24. Si on croit que la mathématique est cohérente.

20.16 Rappel sur les intégrales usuelles

Soit une fonction

$$\begin{aligned} f\colon [a,b] \subset \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto f(x). \end{aligned} \tag{20.395}$$

L'intégrale de f sur le segment $[a,b]$, notée $\int_a^b f(x)dx$ est le nombre égal à l'aire de la surface située entre le graphe de f et l'axe des x, comme indiqué à la figure 20.5.

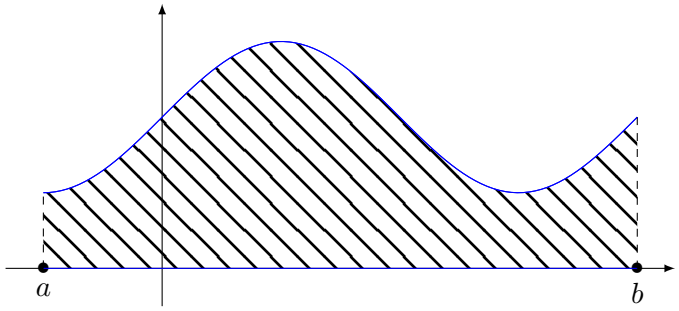


FIGURE 20.5: L'intégrale de f entre a et b représente la surface sous la fonction.
Label: fgKKLooMbjxdI

Définition 20.103.
*Si f est une fonction de une variable à valeurs réelles, une **primitive** de f est une fonction F telle que $F' = f$.*

Toute fonction continue admet une primitive.

Théorème 20.104 (Théorème fondamental du caclul intégral)**.**
Si f est une fonction positive et continue, et si F est une primitive de f, alors

$$\int_a^b f(x)dx = F(b) - F(a). \tag{20.396}$$

Remarque 20.105.
Si f est une fonction continue par morceaux, l'intégrale de f se calcule comme la somme des intégrales de ses morceaux. Plus précisément si nous avons $a = x_0 < x_1 < \ldots < x_n = b$ et si f est continue sur $]x_i, x_{i+1}[$ pour tout i, alors nous posons

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)dx. \tag{20.397}$$

Sur chacun des morceaux, l'intégrale se calcule normalement en passant par une primitive.

20.17 Intégrales le long de chemins

20.17.1 Circulation d'un champ de vecteur

Définition 20.106.
*Soit $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$. On appelle **circulation** de F le long du chemin σ le scalaire*

$$\int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{20.398}$$
EqDeffvkZwh

Il existe de nombreuses notations pour cela ; entre autres :

$$\int_\sigma F = \int_\sigma F \cdot ds. \tag{20.399}$$

En physique, la circulation de la force le long d'un chemin est le travail de la force.

Exemple 20.107.
À la surface de la Terre, le champ de gravitation est donné par

$$G(x,y,z) = -mg\begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{20.400}$$

Si nous considérons un mobile qui monte à vitesse constante jusqu'à la hauteur h, c'est-à-dire le chemin

$$\sigma(t) = \begin{pmatrix}0\\0\\t\end{pmatrix} \tag{20.401}$$

avec $t \in [0,h]$. Le travail de la gravitation est alors donné par

$$W = \int_0^h G\big(\sigma(t)\big) \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = -mg\int_0^h \begin{pmatrix}0\\0\\1\end{pmatrix} \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = -mgh. \tag{20.402}$$

Cela est bien le résultat usuel de l'énergie potentielle. Nous allons voir bientôt que nous nommons la fonction mgh énergie *potentielle* précisément parce que la force dérive de ce potentiel. △

Exemple 20.108.
Soit le chemin

$$\begin{aligned}\sigma\colon [0,2\pi] &\to \mathbb{R}^3\\ t &\mapsto \begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}.\end{aligned} \tag{20.403}$$

et le champ de vecteurs

$$F\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}. \tag{20.404}$$

La circulation de ce champ de vecteur le long de l'hélice σ est

$$\begin{aligned}\int_\sigma F\cdot ds &= \int_0^{2\pi} (F\circ\sigma)(t)\cdot\sigma'(t)dt\\ &= \int_0^{2\pi}\begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}\cdot\begin{pmatrix}\cos(t)\\ -\sin(t)\\ 1\end{pmatrix}dt\\ &= \int_0^{2\pi} t dt\\ &= \left[\frac{t^2}{2}\right]_0^{2\pi}\\ &= 2\pi^2.\end{aligned} \tag{20.405}$$

△

Proposition 20.109.
La circulation d'un champ de vecteurs le long d'un chemin ne dépend pas du paramétrage. En d'autres termes, si σ_1 et σ_2 sont deux chemins équivalents, alors

$$\int_{\sigma_1} F = \int_{\sigma_2} F. \tag{20.406}$$

Démonstration. Soient deux chemins $\sigma_1\colon [a,b] \to \mathbb{R}^3$ et $\sigma_2\colon [c,d] \to \mathbb{R}^3$ équivalents, c'est-à-dire tels que

$$\sigma_1(t) = \sigma_2\big(\varphi(t)\big) \tag{20.407}$$

où $\varphi\colon [a,b] \to [c,d]$ strictement croissante. En utilisant le fait que $\sigma_1'(t) = \varphi'(t)\sigma_2'\big(\varphi(t)\big)$, nous avons

$$\begin{aligned}
\int_{\sigma_1} F \cdot ds &= \int_a^b F\big(\sigma_1(t)\big) \cdot \sigma_1'(t)dt \\
&= \int_a^b F\Big(\sigma_2\big(\varphi(t)\big)\Big) \cdot \sigma_2'\big(\varphi(t)\big)\varphi'(t)dt \\
&= \int_{\varphi(a)}^{\varphi(b)} F\big(\sigma_2(s)\big) \cdot \sigma_2(s)ds \\
&= \int_c^d F\big(\sigma_2(s)\big) \cdot \sigma_2'(s)ds \\
&= \int_{\sigma_2} F \cdot ds.
\end{aligned} \tag{20.408}$$

où nous avons effectué le changement de variables $s = \varphi(t)$, $ds = \varphi'(t)dt$. □

Remarque 20.110.
Si σ_2 est le chemin opposé de σ, alors

$$\int_{\sigma_2} F = -\int_{\sigma_1} F. \tag{20.409}$$

20.18 Circulation d'un champ conservatif

Si nous avons une fonction scalaire $V\colon \mathbb{R}^3 \to \mathbb{R}$, nous pouvons construire un champ de vecteur en prenant le gradient :

$$F(x) = \nabla V(x). \tag{20.410}$$

On dit que le champ de vecteur F **dérive** de V, et on dit que V est le **potentiel** de F. Nous posons la définition suivante :

Définition 20.111.
*Un champ de vecteurs $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ est un champ **conservatif** si il existe une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ telle que*

$$F(x) = \nabla V(x). \tag{20.411}$$

Nous disons aussi parfois que le champ F dérive d'un potentiel *ou bien qu'il s'agit d'un* champ de gradient*.*

Les champs de vecteurs conservatifs sont particulièrement importants parce que presque toutes les forces connues en physiques dérivent d'un potentiel. Nous verrons que la terminologie « conservatif » provient du fait que les forces de ce type conservent l'énergie associée.

Proposition 20.112.
Considérons une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ (que nous appellerons potentiel*) et le champ de vecteur qui en dérive :*

$$F = \nabla V. \tag{20.412}$$

Alors

$$\int_\sigma F \cdot ds = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{20.413}$$

Autrement dit, le travail nécessaire pour déplacer un objet d'un point à un autre dans un champ de force conservatif vaut la différence de potentiel entre le point de départ et le point d'arrivée.

Démonstration. Par définition,

$$\int_\sigma F \cdot ds = \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{20.414}$$

Eqintparddeftray

Nous pouvons transformer l'intégrante de la façon suivante :

$$\begin{aligned} F\big(\sigma(t)\big) \cdot \sigma'(t) &= \nabla V\big(\sigma(t)\big) \cdot \sigma'(t) \\ &= \frac{\partial V}{\partial x}\big(\sigma(t)\big)\sigma'_x(t) + \frac{\partial V}{\partial y}\big(\sigma(t)\big)\sigma'_y(t) + \frac{\partial V}{\partial z}\big(\sigma(t)\big)\sigma'_z(t) \\ &= \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] \end{aligned} \tag{20.415}$$

où nous avons posé

$$\sigma(t) = \begin{pmatrix} \sigma_x(t) \\ \sigma_y(t) \\ \sigma_z(t) \end{pmatrix} \tag{20.416}$$

et utilisé à l'envers la formule de dérivation de fonction composée pour

$$\frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] = \Big((V \circ \sigma)(t)\Big)'. \tag{20.417}$$

En remettant ces expressions dans l'intégrale (20.414),

$$\int_\sigma F \cdot ds = \int_a^b \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big]dt = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{20.418}$$

□

Exemple 20.113.
Nous savons que le champ de gravitation dérive d'un potentiel. À la surface de la Terre, le potentiel de gravitation vu par une masse m est donné par la fonction $V(x, y, z) = mgz$. Si nous voulons soulever cette masse d'une hauteur h, cela demandera toujours une énergie mgh, quel que soit le chemin suivi : en ligne droite vertical, en diagonal, en hélice, ... △

Exemple 20.114.
À plus grande échelle, le champ de gravitation est encore un champ qui dérive d'un potentiel. En coordonnées sphériques,

$$V(\rho, \theta, \varphi) = k\frac{m}{\rho} \tag{20.419}$$

Lorsqu'un satellite a une orbite de rayon R autour la Terre, il reste sur la sphère $\rho = R$. Donc il reste sur une surface sur laquelle V est constante. Il n'y a donc pas de travail de la force de gravitation ! C'est pour cela qu'un satellite peut tourner pendant des siècles sans apport énergétique. △

Exemple 20.115.
Soit le champ de vecteurs

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} \tag{20.420}$$

et le chemin

$$\sigma(t) = \begin{pmatrix} t^4/4 \\ \sin^3(t\frac{\pi}{2}). \end{pmatrix} \tag{20.421}$$

Nous voulons calculer la circulation de F le long du chemin σ entre $t = 0$ et $t = 1$.

La première chose à voir est que $F = \nabla V$ avec $V(x, y) = xy$. Donc la circulation sera donnée par

$$\int_\sigma F \cdot ds = V\big(\sigma(1)\big) - V\big(\sigma(0)\big) = V\big(\frac{1}{4}, 1\big) - V(0, 0) = \frac{1}{4} - 0 = \frac{1}{4}. \tag{20.422}$$

Nous n'avons pas réellement calculé l'intégrale. △

20.19 Intégration de fonction à deux variables

20.19.1 Intégration sur un domaine rectangulaire

PgRapIntMultFubiniRect

Soit une fonction positive

$$\begin{aligned} f\colon [a,b]\times[c,d] &\to \mathbb{R}^+ \\ (x,y) &\mapsto f(x,y). \end{aligned} \tag{20.423}$$

L'intégrale de f sur le rectangle $[a,b]\times[c,d]$ est le volume sous le graphe de la fonction. C'est-à-dire le volume de l'ensemble

$$\{(x,y,z) \text{ tel que } (x,y)\in[a,b]\times[c,d], z\leqslant f(x,y)\}. \tag{20.424}$$

Théorème 20.116 (Théorème de Fubini)**.**
Soit une fonction $f\colon \mathbb{R}^2\to\mathbb{R}$ une fonction continue par morceaux sur $\mathcal{R}=[a,b]\times[c,d]$. Alors

$$\int_{\mathcal{R}} f(x,y)dxdy = \int_a^b \left[\int_c^d f(x,y)dy\right]dx = \int_c^d \left[\int_a^b f(x,y)dx\right]dy. \tag{20.425}$$

En pratique, nous utilisons le théorème de Fubini pour calculer les intégrales sur des rectangles.

Exemple 20.117.
Nous voudrions intégrer la fonction $f(x,y)=4+x^2+y^2$ sur le rectangle de la figure 20.6.

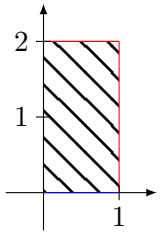


FIGURE 20.6: Intégration sur un rectangle

LabelFigVNBGooSqMsGU

L'ensemble sur lequel nous intégrons est donné par le produit cartésien d'intervalles $E=[0,1]\times[0,2]$. Le théorème de Fubini montre que nous pouvons intégrer séparément sur l'intervalle horizontal et vertical :

$$\int_{E=[0,1]\times[0,2]} f = \int_{[0,1]}\left(\int_{[0,2]}(4-x^2-y^2)dy\right)dx. \tag{20.426}$$

Ces intégrales sont maintenant des intégrales usuelles qui s'effectuent en calculant des primitives :

$$\begin{aligned} \int_0^1\int_0^2(4-x^2-y^2)dy\,dx &= \int_0^1\left[4y-x^2y-\frac{y^3}{3}\right]_0^2 dx \\ &= \int_0^1(8-2x^2-\frac{8}{3})dx \\ &= \left[\frac{16x}{3}-\frac{2x^3}{3}\right]_0^1 \\ &= \frac{14}{3}. \end{aligned} \tag{20.427}$$

Avec Sage, on peut faire comme ceci :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
```

```
------------------------------------------------------------------------
sage: f(x,y)=4-x**2-y**2
sage: f.integrate(y,0,2).integrate(x,0,1)
(x, y) |--> 14/3
```

△

20.19.2 Intégration sur un domaine non rectangulaire

PgRapIntMultFubiniTri

Nous voulons maintenant intégrer la fonction $f(x,y) = x^2 + y^2$ sur le triangle de la figure 20.7.

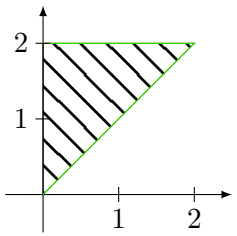


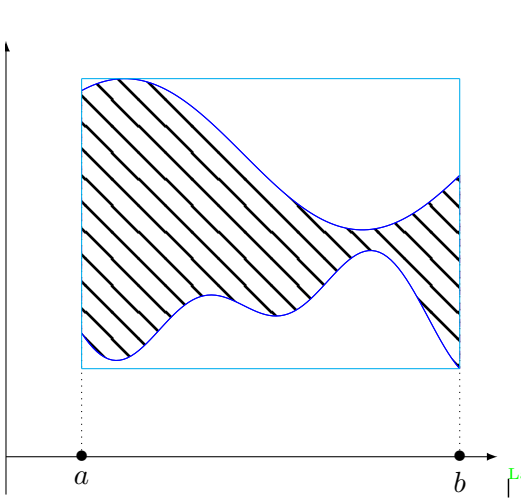
FIGURE 20.7: Intégration sur un triangle

LabelFigCURGooXvruWV

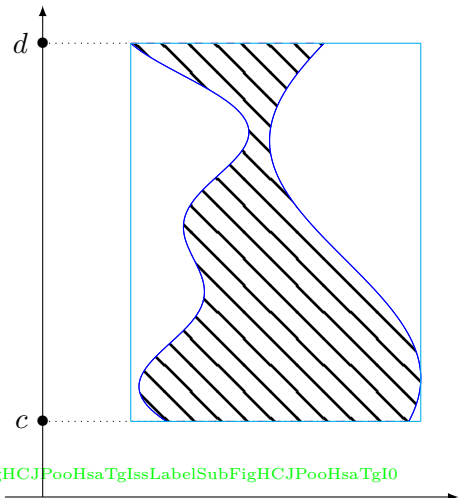
Étant donné que y varie de 0 à 2 et que *pour chaque* y, la variable x varie de 0 à y, nous écrivons l'intégrale sur le triangle sous la forme :

$$\int_{\text{triangle}} (x^2 + y^2) dxdy = \int_0^2 \left(\int_0^y (x^2 + y^2) dx \right) dy. \tag{20.428}$$

Il existe principalement deux types de domaines non rectangulaires : les « horizontaux » et les « verticaux », voir figure 20.8.



(a) Un domaine horizontal.



(b) Un domaine vertical.

FIGURE 20.8: Deux types de surfaces. Nous avons tracé un rectangle qui contient chacune des deux surfaces. L'intégrale sur un domaine sera l'intégrale sur le rectangle de la fonction qui vaut zéro en dehors du domaine.

LabelFigHCJPooHsaTgI

Les surfaces horizontales sont de la forme

$$D = \{(x,y) \text{ tel que } x \in [a,b], \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\} \tag{20.429}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Pour un tel domaine nous avons

$$\int_D f(x,y)dxdy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy. \tag{20.430}$$

Les surfaces verticales sont de la forme

$$D = \{(x,y) \text{ tel que } y \in [c,d], \psi_1(y) \leqslant x \leqslant \psi_2(y)\} \tag{20.431}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Dans ces cas nous avons

$$\int_D f = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx. \tag{20.432}$$

Proposition 20.118.
L'aire du domaine D vaut l'intégrale de la fonction $f(x,y) = 1$ sur D :

$$Aire(D) = \int_D dxdy. \tag{20.433}$$

Démonstration. Supposons que le domaine soit du type « horizontal ». En utilisant le théorème de Fubini avec $f(x,y) = 1$ nous avons

$$\int_D dxdy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} dy \right] dx = \int_a^b \big[\varphi_2(x) - \varphi_1(x)\big]. \tag{20.434}$$

Cela représente l'aire sous φ_2 moins l'aire sous φ_1, et par conséquent l'aire contenue entre les deux. □

Exemple 20.119.
Cherchons la surface du disque de centre $(0,0)$ et de rayon 1 dessinée à la figure 20.9.

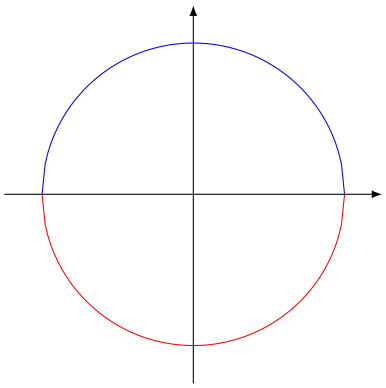


FIGURE 20.9: En bleu, la fonction $\sqrt{r^2-x^2}$ et en rouge, la fonction $-\sqrt{r^2-x^2}$. LabelFigCMMAooQegASg

Le domaine est donné par $\varphi_1(x) \leqslant y \leqslant \varphi_2(x)$ et $x \in [-r,r]$ où $\varphi_1(x) = -\sqrt{r^2-x^2}$ et $\varphi_2(x) = \sqrt{r^2-x^2}$. L'aire est donc donnée par

$$A = \int_{-r}^{r} \big[\varphi_2(x) - \varphi_1(x)\big]dx = 2\int_{-r}^{r} \sqrt{r^2-x^2}dx = 4\int_0^r \sqrt{r^2-x^2}. \tag{20.435}$$

Nous effectuons le premier changement de variables $x = ru$, donc $dx = rdu$. En ce qui concerne les bornes, si $x = 0$, alors $u = 0$ et si $x = r$, alors $u = 1$. L'intégrale à calculer devient

$$A = 4\int_0^1 \sqrt{r^2 - r^2u^2}rdu = 4r^2\int_0^1 \sqrt{1-u^2}du. \tag{20.436}$$

Cette dernière intégrale se calcule en posant

$$\begin{aligned} u &= \sin(t) \quad du = \cos(t)dt \\ u &= 0 \qquad\quad t = 0 \\ u &= 1 \qquad\quad t = \pi/2. \end{aligned} \tag{20.437}$$

Nous avons

$$A = 4r^2 \int_0^{\pi/2} \sqrt{1 - \sin^2(t)} \cos(t)dt = 4r^2 \int_0^{\pi/2} \cos^2(t)dt. \tag{20.438}$$

En utilisant la formule $2\cos^2(x) = 1 + \cos(2x)$, nous avons

$$A = 4r^2 \int_0^{\pi/2} \frac{1 + \cos(2t)}{2} dt = \pi r^2. \tag{20.439}$$

△

20.19.3 Changement de variables

Nous n'allons pas parler de changements de variables maintenant parce que les principaux exemples sont les coordonnées polaires, cylindriques et sphériques qui requièrent les fonctions trigonométriques. Ce sera pour la section 18.15.

20.20 Les intégrales triples

Les intégrales triples fonctionnent exactement de la même manière que les intégrales doubles. Il s'agit de déterminer sur quelle domaine les variables varient et d'intégrer successivement par rapport à x, y et z. Il est autorisé de permuter l'ordre d'intégration [25] à condition d'adapter les domaines d'intégration.

Exemple 20.120.
Soit le domaine parallélépipédique rectangle

$$R = [0, 1] \times [1, 2] \times [0, 4]. \tag{20.440}$$

Pour intégrer la fonction $f(x, y, z) = x^2 y \sin(z)$ sur R, nous faisons

$$\begin{aligned} I &= \int_R x^2 y \sin(z)\, dxdydz \\ &= \int_0^1 dx \int_1^2 dy \int_0^4 x^2 y \sin(z)dz \\ &= \int_0^1 dx \int_1^2 x^2 y(1 - \cos(4))dy \\ &= \int_0^1 \frac{3}{2}(1 - \cos(4))x^2 dx \\ &= \frac{1}{2}\big(1 - \cos(4)\big). \end{aligned} \tag{20.441}$$

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=x**2*y*sin(z)
sage: f.integrate(x,0,1).integrate(y,1,2).integrate(z,0,4)
(x, y, z) |--> -1/2*cos(4) + 1/2
```

25. En toute rigueur, cela n'est pas vrai, mais nous ne considérons seulement des cas où cela est autorisé.

△

Exemple 20.121.
Soit D la région délimitée par le plan $x = 0$, $y = 0$, $z = 2$ et la surface d'équation

$$z = x^2 + y^2. \tag{20.442}$$

Cherchons à calculer $\int_D x\,dx\,dy\,dz$. Ici, un dessin indique que le volume considéré est $z \geqslant x^2 + y^2$. Il y a plusieurs façons de décrire cet ensemble. Une est celle-ci :

$$\begin{aligned} &z\colon 0 \to 2\\ &x\colon 0 \to \sqrt{z}\\ &y\colon 0 \to \sqrt{z - x^2}. \end{aligned} \tag{20.443}$$

Cela revient à dire que z peut prendre toutes les valeurs de 0 à 2, puis que pour chaque z, la variable x peut aller de 0 à $\sqrt{z}$, mais que pour chaque z et x fixés, la variable y ne peut pas dépasser $\sqrt{z - x^2}$. En suivant cette méthode, l'intégrale à calculer est

$$\int_0^2 dz \int_0^{\sqrt{z}} dx \int_0^{\sqrt{z-x^2}} f(x, y, z)dy. \tag{20.444}$$

```
--------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                     |
| Type notebook() for the GUI, and license() for information.      |
--------------------------------------------------------------------
sage: f(x,y,z)=x
sage: assume(z>0)
sage: assume(z-x**2>0)
sage: f.integrate(y,0,sqrt(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Notez qu'il a fallu aider Sage en lui indiquant que $z > 0$ et $z - x^2 > 0$.

Un autre paramétrage serait

$$\begin{aligned} &x\colon 0 \to \sqrt{2}\\ &y\colon 0 \to \sqrt{2 - x^2}\\ &z\colon x^2 + y^2 \to 2. \end{aligned} \tag{20.445}$$

```
--------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                     |
| Type notebook() for the GUI, and license() for information.      |
--------------------------------------------------------------------
sage: f(x,y,z)=x
sage: assume(2-x**2>0)
sage: f.integrate(y,0,sqrt(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Écrivons le détail de cette dernière intégrale :

$$\begin{aligned} I &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} dy \int_{x^2+y^2}^2 x dz\\ &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} x(2 - x^2 - y^2)dy\\ &= \int_0^{\sqrt{2}} dx\, x \left[(2 - x^2)y - \frac{y^3}{3}\right]_0^{\sqrt{2-x^2}}\\ &= \int_0^{\sqrt{2}} \frac{2}{3} x(2 - x^2)^{3/2} dx. \end{aligned} \tag{20.446}$$

Ici nous effectuons le changement de variable $u = x^2$, $du = 2xdx$. Ne pas oublier de changer les bornes de l'intégrale :

$$I = \frac{1}{3}\int_0^2 (2-u)^{3/2}du. \tag{20.447}$$

Le changement de variable $t = 2 - u$, $dt = -du$ fait venir (attention aux bornes!!)

$$I = -\frac{1}{3}\int_2^0 t^{3/2}dt = \frac{1}{3}\left[\frac{t^{5/2}}{5/2}\right]_0^2 = \frac{8}{15}\sqrt{2}. \tag{20.448}$$

△

20.20.1 Volume

Parmi le nombreuses interprétations géométriques de l'intégrale triple, notons celle-ci :

Proposition 20.122.
Soit $D \subset \mathbb{R}^3$. Le volume de D est donné par

$$Vol(D) = \int_D dxdydz. \tag{20.449}$$

C'est-à-dire l'intégrale de la fonction $f(x, y, z) = 1$ sur D.

Suivant les points de vue, cette proposition peut être considérée comme une *définition* du volume.

ExemVolSphCart

Exemple 20.123.
Calculons le volume de la sphère de rayon R. Le domaine de variation des variables x, y et z pour la sphère est

$$\begin{aligned} x&: -R \to R \\ y&: -\sqrt{R^2-x^2} \to \sqrt{R^2-x^2} \\ z&: -\sqrt{R^2-x^2-y^2} \to \sqrt{R^2-x^2-y^2}. \end{aligned} \tag{20.450}$$

Par conséquent nous devons calculer l'intégrale

$$V = \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} dz. \tag{20.451}$$

La première intégrale est simple :

$$V = 2\int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2-x^2-y^2}dy. \tag{20.452}$$

Afin de simplifier la notation, nous posons $a = R^2 - x^2$. Ceci n'est pas un changement de variables : juste une notation provisoire le temps d'effectuer l'intégration sur y. Étudions donc

$$I = \int_{-\sqrt{a}}^{\sqrt{a}} \sqrt{a-y^2}dy, \tag{20.453}$$

ce qui est la surface du demi-disque de rayon $\sqrt{a}$. Nous avons donc

$$I = \frac{\pi a}{2} = \frac{\pi}{2}(R^2 - x^2), \tag{20.454}$$

et

$$V = 2\int_{-R}^{R} \frac{\pi}{2}(R^2-x^2)dx = \pi\left[R^2x - \frac{x^3}{3}\right]_{-R}^{R} = \frac{4}{3}\pi R^3. \tag{20.455}$$

△

Exemple 20.124.
Nous pouvons calculer le volume de la sphère en utilisant les coordonnées sphériques. Les bornes des variables pour la sphère de rayon R sont

$$\begin{aligned}\rho &: 0 \to R\\ \theta &: 0 \to \pi\\ \varphi &: 0 \to 2\pi.\end{aligned} \tag{20.456}$$

En n'oubliant pas le jacobien $\rho^2 \sin(\theta)$, l'intégrale à calculer est

$$V = \int_0^R d\rho \int_0^{2\pi} d\varphi \int_0^{\pi} \rho^2 \sin(\theta) d\theta \tag{20.457}$$

L'intégrale sur φ fait juste une multiplication par 2π. Celle sur ρ vaut

$$\int_0^R \rho^2 d\rho = \frac{R^3}{3}. \tag{20.458}$$

L'intégrale sur θ donne

$$\int_0^{\pi} \sin(\theta) d\theta = [-\cos(\theta)]_0^{\pi} = 2. \tag{20.459}$$

Le tout fait par conséquent

$$V = \frac{4}{3}\pi R^3. \tag{20.460}$$

Sans contestes, le passage aux coordonnées sphériques a considérablement simplifié le calcul par rapport à celui de l'exemple 20.123. △

20.21 Un petit peu plus formel

20.21.1 Intégration sur un domaine non rectangulaire

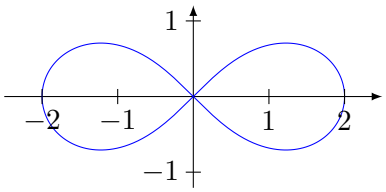


FIGURE 20.10: Intégrer sur des domaines plus complexes. LabelFigPONXooXYjEot

La méthode de Fubini ne fonctionne plus sur un domaine non rectangulaire tel que celui de la figure 20.10. Nous allons donc utiliser une astuce. Considérons le domaine

$$E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } a < x < b \text{ et } \alpha(x) < y < \beta(x)\} \tag{20.461}$$

représenté sur la figure 20.10. Nous considérons la fonction

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in E\\ 0 & \text{sinon.}\end{cases} \tag{20.462}$$

Ensuite intégrons $\tilde{f}$ sur un rectangle qui englobe la surface à intégrer à l'aide de Fubini. Étant donné que $\tilde{f} = f$ sur la surface et que $\tilde{f}$ est nulle en dehors, nous avons

$$\int_E f = \int_E \tilde{f} = \int_{\text{rectangle}} \tilde{f} = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right) dx. \tag{20.463}$$

Dans le cas de l'intégrale de $f(x, y) = x^2 + y^2$ sur le triangle de la figure 20.7, nous avons

$$\int_{\text{triangle}} (x^2 + y^2) dx dy = \int_0^2 \left(\int_0^y (x^2 + y^2) dx \right) dy. \tag{20.464}$$

Remarque 20.125.
Le nombre $\int_D f(x,y)dxdy$ ne dépend pas du choix du rectangle englobant D.

En pratique, nous calculons l'intégrale en utilisant une extension du théorème de Fubini :

Théorème 20.126.
Soit $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction continue où D est un domaine de type vertical ou horizontal.

(1) Si D est vertical, alors

$$\int_D f = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy \right] dx. \tag{20.465}$$

(2) Si D est horizontal, alors

$$\int_D f = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx \right] dy. \tag{20.466}$$

20.21.1.1 Coordonnées polaires

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0,\infty[\times]0,2\pi[&\to \mathbb{R}^2\backslash D \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{20.467}$$

où D est la demi-droite $y = 0$, $x \geqslant 0$. Le fait que les coordonnées polaires ne soient pas un difféomorphisme sur tout $\mathbb{R}^2$ n'est pas un problème pour l'intégration parce que le manque de difféomorphisme est de mesure nulle dans $\mathbb{R}^2$. Le jacobien est donné par

$$Jg = \det \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{20.468}$$

20.21.1.2 Coordonnées sphériques

SubSubCoordSpJxhMwm

Les coordonnées sphériques sont données par

$$\left\{ \begin{array}{ll} x = r\cos\theta\sin\varphi & r \in]0,\infty[\\ y = r\sin\theta\sin\varphi \quad \text{avec} & \theta \in]0,2\pi[\\ z = r\cos\varphi & \phi \in]0,\pi[. \end{array} \right. \tag{20.469}$$

OMEqChmVarSpherique

Le jacobien associé est $Jg(r,\theta,\varphi) = -r^2\sin\varphi$. Rappelons que ce qui rentre dans l'intégrale est la valeur absolue du jacobien.

Si nous voulons calculer le volume de la sphère de rayon R, nous écrivons donc

$$\int_0^R dr \int_0^{2\pi} d\theta \int_0^{\pi} r^2\sin(\phi)d\phi = 4\pi R = \frac{4}{3}\pi R^3. \tag{20.470}$$

Ici, la valeur absolue n'est pas importante parce que lorsque $\phi \in]0,\pi,[$, le sinus de ϕ est positif.

Des petits malins pourraient remarquer que le changement de variable (20.469) est encore un paramétrage de $\mathbb{R}^3$ si on intervertit le domaine des angles :

$$\begin{aligned} &\theta\colon 0 \to \pi \\ &\phi\colon 0 \to 2\pi, \end{aligned} \tag{20.471}$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr \int_0^{\pi} d\theta \int_0^{2\pi} r^2\sin(\phi)d\phi = 0. \tag{20.472}$$

EqOMVolumeIncorrectSphere

Pourquoi ces « nouvelles » coordonnées sphériques sont-elles mauvaises ? Il y a que quand l'angle ϕ parcours $]0, 2\pi[$, son sinus n'est plus toujours positif, donc la *valeur absolue* du jacobien n'est plus $r^2 \sin(\phi)$, mais $r^2 \sin(\phi)$ pour les ϕ entre 0 et π, puis $-r^2 \sin(\phi)$ pour ϕ entre π et 2π. Donc l'intégrale (20.472) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr \int_0^\pi d\theta \int_0^\pi r^2 \sin(\phi) d\phi - \int_0^R dr \int_0^\pi d\theta \int_\pi^{2\pi} r^2 \sin(\phi) d\phi = \frac{4}{3}\pi R^3 \tag{20.473}$$

20.22 Aire et primitive

Soient $f : \mathbb{R} \to \mathbb{R}$ une fonction continue et $x \in \mathbb{R}$. Pour chaque $a \in \mathbb{R}$, nous pouvons considérer le nombre $F(x)$ défini par

$$F(x) = \int_a^x f(t)dt. \tag{20.474}$$

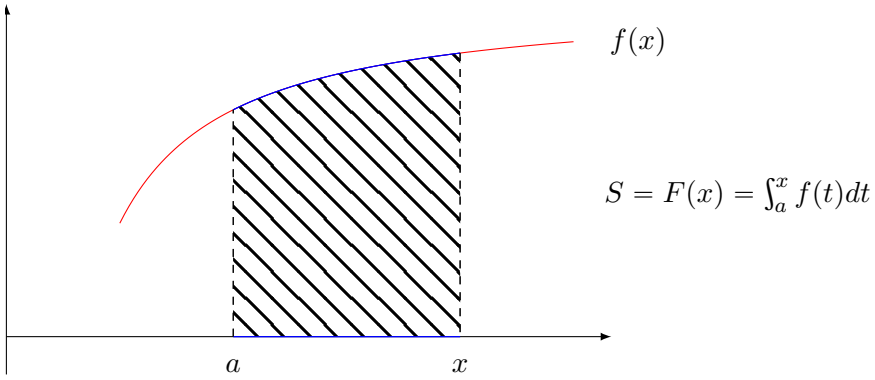


FIGURE 20.11: Surface sous une courbe LabelFigVSZRooRWgUGu

La fonction F ainsi définie a deux importantes propriétés :

(1) C'est une primitive de f,
(2) Elle donne la surface en dessous de f entre les points a et x, voir la figure 20.11.

Notons que tant que f est positive, la surface est croissante.

La manière de calculer la surface comprise entre deux fonctions est dessinée à la figure 20.12.

La surface entre les deux fonctions $y_1(x)$ et $y_2(x)$ se calcule comme suit.

(1) On calcule les intersections entre $y1$ et y_2. Notons a et b les ordonnées obtenues.
(2) La surface demandée est la différence entre la surface sous la fonction y_1 (la plus grande) et la surface sous la fonction y_2 (la plus petite), donc

$$S = \int_a^b y_1 - \int_a^b y_1. \tag{20.475}$$

20.22.1 Longueur d'arc de courbe

La longueur de l'arc de courbe de la fonction $y = f(x)$ entre les abscisses x_0 et x_1 est donné par la formule

$$l(x_0, x_1) = \int_{x_0}^{x_1} \sqrt{1 + y'(t)^2} dt. \tag{20.476}$$ EqLongArcCourbe

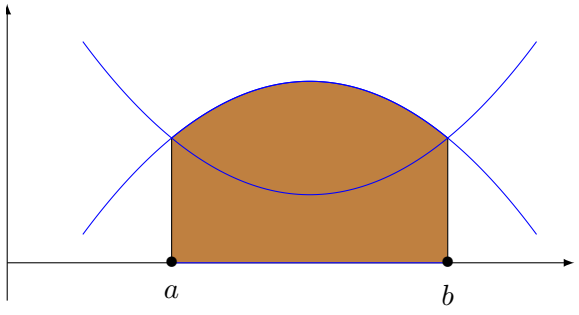
Lorsque la courbe est donnée sous forme paramétrique

$$\begin{cases} x = x(t) & \text{(20.477a)} \\ y = y(t), & \text{(20.477b)} \end{cases}$$

alors la formule devient

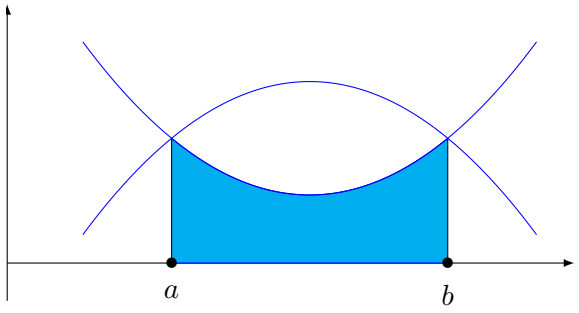
$$l(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt, \tag{20.478}$$ EqLongArcParam

où $\dot{x}(t) = x'(t)$.



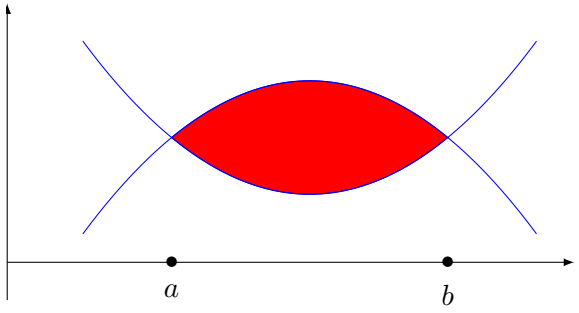
LabelFigQOBAooZZZOrlssLabelSubFigQOBAooZZZ

(a) La grande surface.



LabelFigQOBAooZZZOrlssLabelSubFigQOBAooZZZ

(b) La surface à soustraire.



LabelFigQOBAooZZZOrlssLabelSubFigQOBAooZZZ

(c) La surface entre les deux fonctions.

FIGURE 20.12: Le calcul de la surface comprise entre deux fonctions.

LabelFigQOBAooZZZOrl

20.22.2 Aire de révolution

Pour savoir l'aire engendrée par la ligne $y = f(x)$ entre a et b autour de l'axe Ox, on utilise la formule

$$S = 2\pi \int_a^b \sqrt{1 + f'(x)^2} f(x) dx. \tag{20.479}$$

20.23 L'aire en dessous d'une courbe

Soit f une fonction à valeurs dans $\mathbb{R}^+$.

Nous voudrions pouvoir calculer l'aire au-dessous du graphe de la fonction f. Nous notons $S_f(x)$ l'aire là-dessous de la fonction f entre l'abscisse 0 et x, c'est-à-dire l'aire bleue de la figure 20.13.

Si la fonction f est continue et que Δx est assez petit, la fonction ne varie pas beaucoup entre x et $x + \Delta x$. L'augmentation de surface entre x et $x + \Delta x$ peut donc être approximée par le rectangle de surface $f(x)\Delta x$. Ce que nous avons donc, c'est que quand Δx est très petit,

$$S_f(x + \Delta x) - S_f(x) = f(x)\Delta x, \tag{20.480}$$

ou encore

$$f(x) = \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{20.481}$$

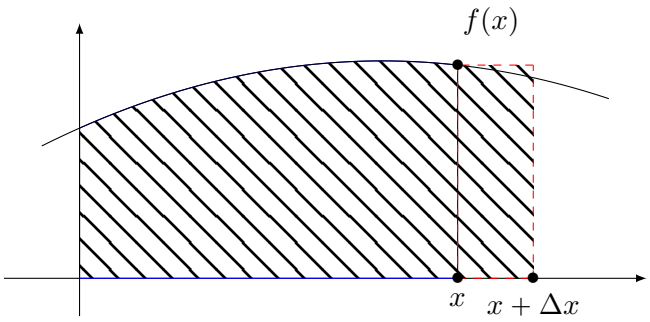


FIGURE 20.13: L'aire en dessous d'une courbe. Le rectangle rouge d'aire $f(x)\Delta x$ approxime de combien la surface augmente lorsqu'on passe de x à $x + \Delta x$. LabelFigKKRooHseDzC

Nous formalisons la notion de « lorsque Δx est très petit » par une limite :

$$f(x) = \lim_{\Delta x \to 0} \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{20.482}$$

Donc, la fonction f est la dérivée de la fonction qui représente l'aire là-dessous de f. Calculer des surfaces revient donc au travail inverse de calculer des dérivées.

20.24 Propriétés des intégrales

LemIneqnormeintintnorm

Lemme 20.127.
Pour toute fonction $u\colon [a,b] \to \mathbb{R}^n$, nous avons

$$\| \int_a^b u(t)dt \| \leqslant \int_a^b \|u(t)\| dt \tag{20.483}$$

pourvu que le membre de gauche ait un sens.

Démonstration. Étant donné que $\int_a^b u(t)dt$ est un élément de $\mathbb{R}^n$, par la proposition 10.11, il existe un $\xi \in \mathbb{R}^n$ de norme 1 tel que

$$\| \int_a^b u(t)dt \| = \xi \cdot \int_a^b u(t)dt = \int_a^b u(t) \cdot \xi dt \leqslant \int_a^b \|u(t)\|\|\xi\| = \int_a^b \|u(t)\|dt. \tag{20.484}$$

□

Proposition 20.128 (Relations de Chasles)**.**
Soit f une fonction continue sur l'intervalle I. Si $a, b, c \in I$ nous avons

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx. \tag{20.485}$$

Sur la figure 20.14, la surface de a à c est évidemment égale à la somme des surfaces de a à b et de b à c.

Corolaire 20.129.

$$\int_a^b f(x)dx = -\int_b^a f(x)dx. \tag{20.486}$$

linearitéintegrale

Proposition 20.130 (Linéarité de l'intégrale)**.**
Si f et g sont deux fonctions continues sur $I \subset \mathbb{R}$, a, $b \in I$ et $\lambda \in \mathbb{R}$ nous avons

$$\int_a^b \big(f(x) + g(x)\big)dx = \int_a^b f(x)dx + \int_a^b g(x)dx, \tag{20.487}$$

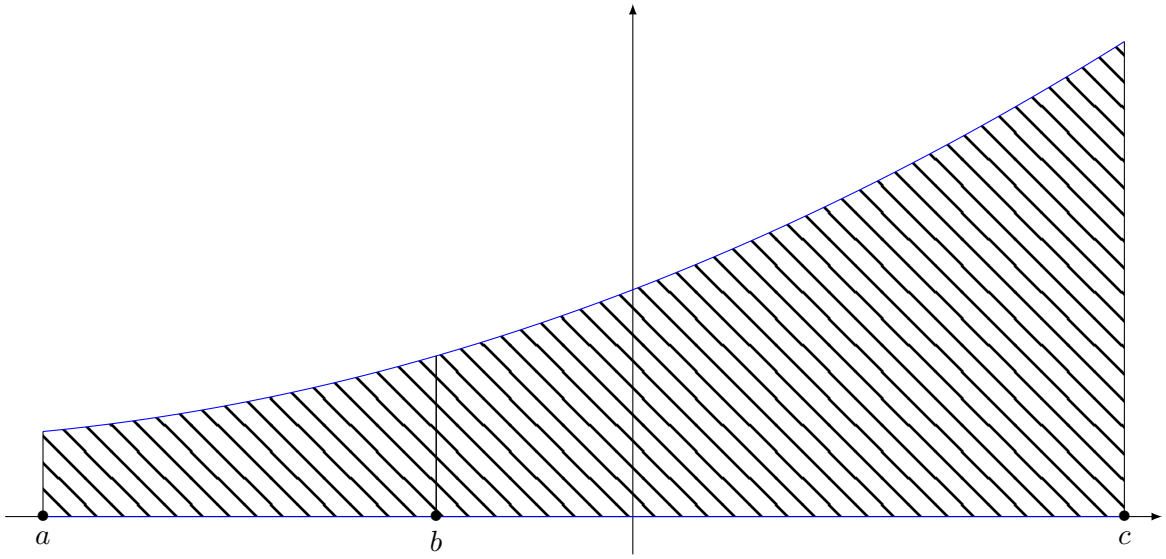


FIGURE 20.14: Illustration pour les relations de Chasles. LabelFigNWDooOObSHB

et

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx. \tag{20.488}$$

PropCJIooHqECbq

Proposition 20.131 (L'intégrale est monotone)**.**
Soient $a, b \in I$ avec $a < b$. Si $f \geqslant g$ sur $[a, b]$ alors

$$\int_a^b f(x)dx \geqslant \int_a^b g(x)dx. \tag{20.489}$$

PropHVWooBDRhCX

Corolaire 20.132 (Positivité)**.**
Si $a < b$ et $f \geqslant 0$ sur $[a, b]$ alors

$$\int_a^b f(x)dx \geqslant 0. \tag{20.490}$$

Ce résultat n'est qu'une application de la proposition 20.131 car il consiste à prendre comme fonction g la fonction nulle.

20.25 Techniques d'intégration

Par le théorème 14.287, la calcul d'une intégrale consiste essentiellement à trouver une primitive de la fonction à intégrer. Il est donc indispensable de bien connaitre les dérivées des fonctions usuelles.

Voici un tableau des primitives à connaitre. PageLCHooMbWjOj

Fonction $f(x)$	Primitive $\int f(x)\,dx$	Ensemble de définition de f	Remarques
x^α	$\frac{x^{\alpha+1}}{\alpha+1}+C$	dépend de α	$\alpha\in\mathbb{R}\backslash\{-1\}$
$\frac{1}{x}$	$\ln\big(\lvert x\rvert\big)+C$	$x\neq 0$	
$\frac{1}{1+x^2}$	$\arctan(x)+C$	$\mathbb{R}$	
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)+C$	$]-1,1[$	
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos(x)+C$	$]-1,1[$	
e^x	e^x+C	$\mathbb{R}$	
$\sin(x)$	$-\cos(x)+C$	$\mathbb{R}$	
$\cos(x)$	$\sin(x)+C$	$\mathbb{R}$	
$1+\tan^2(x)$	$\tan(x)+C$	un intervalle de la forme $]-\frac{\pi}{2},\frac{\pi}{2}[+k\pi$	

Notez que au signe près, les fonctions arcsin et arccos ont la même dérivée.

Si la fonction à intégrer est une combinaison linéaire de fonctions usuelles alors sa primitive peut être calculée en utilisant la proposition 20.130. Dans les sections suivantes on abordera deux autres cas où la fonction à intégrer peut s'écrire en termes de fonctions dont on connaît une primitive.

20.25.1 Intégration par parties

PROPooRLFIooQHnyJY

Proposition 20.133 ([1]).
Si u et v sont deux fonctions dérivables de dérivées continues sur l'intervalle $[a,b]$ alors

$$\int_a^b u(x)v'(x)dx=\big[u(x)v(x)\big]_a^b-\int_a^b u'(x)v(x)dx. \tag{20.491}$$

EQooKISBooQyGMQT

Si de plus les fonctions uv' et $u'v$ sont intégrables sur $[a,\infty]$, alors

$$\int_a^\infty u(x)v'(x)dx=\big[u(x)v(x)\big]_a^\infty-\int_a^\infty u'(x)v(x)dx. \tag{20.492}$$

Démonstration. Il s'agit d'utiliser à l'envers la formule de dérivation d'un produit :

$$uv'=(uv)'-u'v. \tag{20.493}$$

Les fonctions à gauche et à droite étant égales, elles ont même intégrale sur $[a,b]$ et par linéarité, voir proposition 20.130, on a :

$$\int_a^b u(x)v'(x)dx=\int_a^b (uv)'(x)-\int_a^b u'(x)v(x)dx. \tag{20.494}$$

La fonction uv est évidemment une primitive de $(uv)'$, de telle sorte que l'on puisse un peu simplifier cette expression :

$$\int_a^b u(x)v'(x)dx=\Big[u(x)v(x)\Big]_a^b-\int_a^b u'(x)v(x)dx, \tag{20.495}$$

ce qu'il fallait démontrer.

En ce qui concerne l'affirmation avec $b=\infty$, le lemme 14.261 est applicable et il suffit de passer à la limite dans (20.491). □

ExWIEooVUgvSp

Exemple 20.134.
Un cas typique d'utilisation de l'intégrale par parties est le suivant. Soit à calculer

$$\int_0^\pi x\cos(x)dx. \tag{20.496}$$

Nous devons écrire $x\cos(x)$ comme un produit $u(x)v'(x)$. Il y a (au moins) deux moyens de le faire :

$$\begin{cases} u = x & \text{(20.497a)} \\ v' = \cos(x). & \text{(20.497b)} \end{cases}$$

ou

$$\begin{cases} u = \cos(x) & \text{(20.498a)} \\ v' = x. & \text{(20.498b)} \end{cases}$$

Nous allons choisir le premier [26]. Nous avons donc

$$\begin{aligned} u = x, \quad v' = \cos(x) \\ u' = 1 \qquad v = \sin(x). \end{aligned} \tag{20.499}$$

En utilisant la formule d'intégration par parties,

$$\int_0^\pi x\cos(x)dx = \Big[x\sin(x)\Big]_0^\pi - \int_0^\pi 1\times\sin(x)dx = \pi\sin(\pi) - \Big[-\cos(x)\Big]_0^\pi = -2. \tag{20.500}$$

△

Le plus souvent, pour alléger les notations, il est plus pratique d'utiliser l'intégration par parties pour déterminer une primitive. Nous utilisons pour cela la formule (sans doute plus simple à retenir)

$$\int uv' = uv - \int u'v. \tag{20.501}$$

ExLTJooDZIYWP

Exemple 20.135.
Nous reprenons l'exemple 20.134 en déterminant cette fois une primitive de $x\cos(x)$:

$$\int x\cos(x)dx = x\sin(x) - \int \sin(x)dx = x\sin(x) + \cos(x) + C, \qquad C\in\mathbb{R}. \tag{20.502}$$

EqTQNooVTYkZX

Nous retrouvons le résultat numérique de l'exemple précédent en ajoutant les extrêmes d'intégration

$$\int_0^\pi x\cos(x)dx = \big[x\sin(x) + \cos(x)\big]_0^\pi = -2. \tag{20.503}$$

△

Remarque 20.136.
Lorsqu'on calcule des intégrales, il est bon de passer par la primitive (c'est-à-dire en suivant l'exemple 20.135 et non 20.134) parce qu'il est alors facile de vérifier le résultat en calculant la dérivée de la primitive trouvée.

Par exemple pour vérifier si (20.502) est correct, il suffit de dériver $x\sin(x) + \cos(x)$:

$$\big(x\sin(x) + \cos(x)\big)' = \sin(x) + x\cos(x) - \sin(x) = x\cos(x). \tag{20.504}$$

La fonction $x\sin(x) + \cos(x)$ est donc bien une primitive de $x\cos(x)$.

primln

Exemple 20.137 (Primitive du logarithme)**.**
La primitive de la fonction logarithme définie en 15.83 nous offre un bon moment d'intégration par partie.

Trouver la primitive de la fonction $x\mapsto\ln(x)$. Pour calculer

$$\int \ln(x)dx \tag{20.505}$$

26. Mais nous conseillons vivement au lecteur d'essayer le deuxième pour se rendre compte qu'il ne fonctionne pas.

nous écrivons $\ln(x) = 1 \times \ln(x)$ et nous posons $u' = 1$ et $v = \ln(x)$, c'est-à-dire

$$\begin{aligned} u' &= 1 \quad & v &= \ln(x) \\ u &= x & v' &= \frac{1}{x}. \end{aligned} \tag{20.506}$$

La formule d'intégration par parties (20.491) donne donc

$$\int \ln(x) = x\ln(x) - \int x \times \frac{1}{x} = x\ln(x) - \int 1 = x\ln(x) - x + C, \qquad C \in \mathbb{R}. \tag{20.507}$$

Il est facile de vérifier par un petit calcul que

$$\big(x\ln(x) - x\big)' = \ln(x). \tag{20.508}$$

△

20.25.2 Changement de variables – pour trouver des primitives

De la même manière que l'utilisation « à l'envers » de la formule de dérivation du produit avait donné la méthode d'intégration par parties, nous allons voir que que l'utilisation « à l'envers » de la formule de dérivation d'une fonction composée donne lieu à la méthode d'intégration par changement de variables.

PROPooMVIUooZmvHxS

Proposition 20.138.
Soient I et J des intervalles de $\mathbb{R}$, $u\colon I \to J$ une fonction qui est dérivable de dérivée continue et $f\colon J \to \mathbb{R}$ une fonction admettant une primitive F. Alors la fonction

$$x \mapsto F\big(u(x)\big) \tag{20.509}$$

est une primitive de

$$f\big(u(x)\big)u'(x). \tag{20.510}$$

changvar

Démonstration. Cela est une utilisation immédiate de la formule de dérivée des fonctions composées. □

Exemple 20.139.
Soit à calculer

$$\int x\sqrt{1-x^2}dx. \tag{20.511}$$

La fonction $g(x) = x\sqrt{1-x^2}$ est le produit de x et de $\sqrt{1-x^2}$. On remarque que la dérivée de $1-x^2$ est $-2x$: nous avons alors, à un facteur -2 près, une expression de la forme (20.510) où la racine carrée joue le rôle de f, $f(t) = \sqrt{t}$, et $1-x^2$ le rôle de u. Une primitive de la fonction $f(t) = \sqrt{t}$ est $F(t) = 2t^{3/2}/3$.

Donc la fonction $\frac{2u(x)^{3/2}}{3} = \frac{2}{3}(1-x^2)^{3/2}$ est primitive de $-2x\sqrt{1-x^2} = -2g(x)$. Autrement dit,

$$\int -2x\sqrt{1-x^2}\,dx = \frac{2(1-x^2)^{3/2}}{3} + C, \tag{20.512}$$

et en divisant par -2 nous trouvons la primitive demandée :

$$\int x\sqrt{1-x^2}\,dx = -\frac{(1-x^2)^{3/2}}{3} + C. \tag{20.513}$$

△

L'exemple suivant donne une façon plus économe de retenir la méthode du changement de variables.

exempleprimitivechangvar

Exemple 20.140.
Soit à calculer

$$\int \cos(x)e^{\sin(x)}dx. \tag{20.514}$$

Vu qu'il y a beaucoup de fonctions trigonométriques dans la fonction à intégrer, nous allons poser $u(x) = \sin(x)$, et remplacer élément par élément tout ce qui contient du « x » dans l'intégrale demandée par la quantité correspondante en termes de u.

La difficulté est de savoir ce que nous allons faire du « dx » dans l'intégrale. Ce dx marque une variation (infinitésimale) de x. La formule des accroissements finis dit que si x augmente de la valeur dx, alors $u(x)$ augmente de $u'(x)dx$, c'est-à-dire que

$$du = \cos(x)dx. \tag{20.515}$$

Nous avons donc les substitutions suivantes à faire :

$$\sin(x) = u \tag{20.516a}$$
$$du = \cos(x)dx \tag{20.516b}$$
$$dx = \frac{du}{\cos(x)}. \tag{20.516c}$$

La chose « magique » est que le $\cos(x)$ se trouvant dans la fonction se simplifie avec le cosinus qui arrive lorsqu'on remplace dx par $\frac{du}{\cos(x)}$. Les substitutions faites nous restons avec

$$\int \cos(x)e^{\sin(x)}dx = \int e^u du = e^u + C, \qquad \text{où } u = \sin(x). \tag{20.517}$$

Attention : la réponse doit être impérativement donnée en termes de x et non de u. Nous écrivons donc

$$\int \cos(x)e^{\sin(x)} = e^{\sin(x)} + C. \tag{20.518}$$

△

20.25.3 Changement de variables – pour calculer des intégrales

Le théorème 14.287 fixe la relation entre la recherche des primitives de f et la calcul de l'intégrale de f sur l'intervalle d'extrêmes a et b. On a vu dans la section précédente comment utiliser le changement de variable pour trouver une primitive de f. Il faut maintenant comprendre comment appliquer ce qu'on a vu dans le calcul d'une intégrale.

En effet nous avons le choix entre

— trouver une primitive de f comme dans la section précédente et appliquer ensuite la formule du corolaire 14.287 ;

— écrire une intégrale pour la nouvelle variable $u = u(x)$ sur l'intervalle entre $u(a)$ et $u(b)$.

Nous allons voir ces deux méthodes dans des exemples.

Exemple 20.141.
Soit à calculer

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{20.519}$$

Les primitives $\int x\sqrt{1-x^2}dx$ ont été trouvé dans l'exemple 20.140. Une primitive est

$$F(x) = \int x\sqrt{1-x^2}dx = -\frac{(1-x^2)^{3/2}}{3}. \tag{20.520}$$

Nous pouvons maintenant calculer l'intégrale de $x\sqrt{1-x^2}$ sur l'intervalle $[1/3, 1/2]$ par la définition

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = F\left(\frac{1}{2}\right) - F\left(\frac{1}{3}\right) = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{20.521}$$

△

Remarque 20.142.
Pour que le calcul d'intégrale donne quelque chose de sensé il faut absolument que la primitive soit écrite en tant que fonction de x et non comme fonction de u. La méthode que nous allons voir dans l'exemple suivant réduit grandement la probabilité d'oublier ce détail, d'où le fait qu'elle soit de loin la plus utilisée.

Exemple 20.143.
Calculons à nouveau

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{20.522}$$

Cette fois nous allons toucher à l'intervalle d'intégration en même temps que faire le changement de variables. Nous savons déjà les substitutions

$$\begin{cases} u = 1 - x^2 & \text{(20.523a)} \\ du = -2xdx & \text{(20.523b)} \\ dx = \dfrac{du}{-2x}. & \text{(20.523c)} \end{cases}$$

En ce qui concerne les extrêmes d'intégration, si $x = 1/3$ alors $u = 1 - \frac{1}{9} = \frac{8}{9}$ et si $x = \frac{1}{2}$ alors $u = \frac{3}{4}$. Nous avons donc encore les substitutions suivantes :

$$\begin{cases} x = 1/3 \to u = 8/9 & \text{(20.524a)} \\ x = 1/2 \to u = 3/4 & \text{(20.524b)} \end{cases}$$

Le calcul est alors

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = -\frac{1}{2}\int_{8/9}^{3/4} \sqrt{u}du = -\frac{1}{2}\left[\frac{u^{3/2}}{3/2}\right]_{8/9}^{3/4} = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{20.525}$$

Attention : la dernière égalité n'est pas immédiate ; elle demande quelques calculs et une bonne utilisation des règles de puissances. △

La deuxième méthode est plus utilisée et, avec un peu d'exercice, plus rapide à mettre en place que la première.

Jusqu'à présent nous avons utilisé des changements de variables dans lesquels nous exprimions u en termes de x. Comme le montre l'exemple suivant, il est parfois fructueux d'utiliser le changement de variable dans le sens inverse : avec x exprimé en termes d'un paramètre.

exemplepassagepolaires

Exemple 20.144.
À calculer :

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx. \tag{20.526}$$

Nous posons $x = \sin(\theta)$ parce que nous savons que $1 - \sin^2(\theta) = \cos^2(\theta)$; nous espérons que le changement de variables simplifie l'expression [27]. Les substitutions à faire dans l'intégrale sont :

$$\begin{cases} x = \sin(\theta) & \text{(20.527a)} \\ dx = \cos(\theta)d\theta, & \text{(20.527b)} \end{cases}$$

27. Lorsqu'on fait un changement de variables, il s'agit toujours d'*espérer* que l'expression se simplifie. Il n'y a pas moyen de savoir à priori si tel changement de variable va être utile. Il faut essayer.

et en ce qui concerne les bornes, si $x = 1/2$ alors $\sin(\theta) = \frac{1}{2}$, c'est-à-dire $\theta = \frac{\pi}{6}$. Si $x = \sqrt{3}/2$ alors $\theta = \frac{\pi}{3}$. Donc

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \int_{\pi/6}^{\pi/3} \sqrt{1-\sin^2(\theta)}\cos(\theta)d\theta. \tag{20.528}$$

Nous avons $1-\sin^2(\theta) = \cos^2(\theta)$ et vu que $\theta \in [\frac{\pi}{6}, \frac{\pi}{3}]$ nous avons toujours $\cos(\theta) > 0$, ce qui donne $\sqrt{\cos^2(\theta)} = \cos(\theta)$. Nous devons donc calculer

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta. \tag{20.529}$$

Pour celle-là, il faut utiliser une formule de trigonométrie [28] :

$$\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}. \tag{20.530}$$

Donc

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \int_{\pi/6}^{\pi/3} \frac{1+\cos(2\theta)}{2}d\theta = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} + \int_{\pi/6}^{\pi/3} \frac{\cos(2\theta)}{2}d\theta, \tag{20.531}$$

Pour calculer proprement la dernière intégrale nous effectuons un autre changement de variable (facile) en posant $t = 2\theta$, $dt = 2d\theta$, $t(\pi/6) = \pi/3$ et $t(\pi/3) = 2\pi/3$, nous avons alors

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} + \int_{\pi/3}^{2\pi/3} \frac{\cos(t)}{4}dt = \left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3} = \frac{\pi}{6} - \frac{\pi}{12} = \frac{\pi}{12}, \tag{20.532}$$

parce que $\sin\left(\frac{2\pi}{3}\right) = \sin\left(\frac{\pi}{3}\right)$. Au final,

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \frac{\pi}{12}. \tag{20.533}$$

△

20.25.4 Intégrations des fractions rationnelles réduites

Définition 20.145.
*Une **fraction rationnelle** est un quotient de deux polynômes à coefficients réels ou complexes.*

Par exemple

$$\frac{x^5 + 7x^4 - \frac{x^3}{2} + x}{x^2 - 1} \tag{20.534}$$

est une fraction rationnelle.

Il sera expliqué dans le cours d'algèbre que toute fraction rationnelle peut être écrite sous forme d'une somme d'éléments simples, c'est-à-dire de fractions rationnelles d'un des deux types suivants :

$$\frac{\alpha}{(x-a)^m}, \qquad \alpha, a \in \mathbb{R}, m \in \mathbb{N} \tag{20.535a}$$

CasMMIooZnZpUWi

$$\frac{\alpha x + \beta}{(x^2+ax+b)^m}; \qquad \alpha, \beta, a, b \in \mathbb{R}, m \in \mathbb{N}, a^2 - 4b < 0. \tag{20.535b}$$

CasMMIooZnZpUWi

Nous allons nous contenter de donner un exemple de chaque type.

(1) En ce qui concerne le cas (20.535a) avec $m = 1$, nous avons par exemple

$$\int \frac{1}{x-3}dx = \ln\big(|x-3|\big) + C. \tag{20.536}$$

Si vous voulez en être tout à fait sûr, effectuez d'abord le changement de variables $u = x - 3$ qui donne $dx = du$.

28. En fait, il y a moyen de terminer le calcul en intégrant deux fois par parties, mais c'est plus compliqué.

(2) En ce qui concerne le cas (20.535a) avec $m \neq 1$, nous avons par exemple

$$\int \frac{1}{(x-1)^4}dx = -\frac{1}{3(x-1)^3} + C. \tag{20.537}$$

Encore une fois, pour s'en convaincre, utiliser le changement de variables $u = x-1$, $dx = du$:

$$\int \frac{1}{(x-1)^4}dx = \int \frac{1}{u^4}du = \int u^{-4}du = -\frac{u^{-3}}{3} + C = -\frac{1}{3}\frac{1}{(x-1)^3} + C. \tag{20.538}$$

(3) En ce qui concerne le cas (20.535b) avec $\alpha \neq 0$, nous avons par exemple

$$\int \frac{x}{x^2+4}dx = \frac{1}{2}\ln(x^2+4) + C. \tag{20.539}$$

Pour ce faire, il faut faire le changement de variables $u = x^2+4$, $du = 2xdx$, $dx = \frac{du}{2x}$ qui donne

$$\int \frac{x}{x^2+4}dx = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln(|u|) + C = \frac{1}{2}\ln(|x^2+4|) + C. \tag{20.540}$$

Dans ce cas nous pouvons oublier d'écrire la valeur absolue dans le logarithme parce que de toutes façons, x^2+4 est toujours positif.

(4) En ce qui concerne le cas (20.535b) avec $\alpha = 0$, nous avons par exemple

$$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{(\frac{x}{2})^2+1} = \frac{1}{2}\arctan(\frac{x}{2}) + C. \tag{20.541}$$

où nous avons utilisé la primitive $\int \frac{dx}{x^2+1}dx = \arctan(x)$ du tableau de la page 1911. Pour vous en convaincre vous pouvez faire la dernière étape avec le changement de variables $u = x/2$, $dx = 2du$.

20.25.5 Quelques formules à connaitre

À retenir 20.146

$$\int (\alpha f(x) + \beta g(x))\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx. \tag{20.542a}$$

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx. \tag{20.542b}$$

$$\int f'(u(x))u'(x)\, dx = \int f(t)\, dt, \qquad \text{avec } t = u(x). \tag{20.542c}$$

$$\int \frac{f'(x)}{f(x)}\, dx = \log|f(x)| + C, \qquad \text{c'est un cas particulier de la formule précédente.} \tag{20.542d}$$

20.25.6 Taylor avec reste intégral

THOooDGCJooXKmFTT

Théorème 20.147 (Taylor, reste intégral)**.**
Soit une fonction $f\colon [a,b] \to \mathbb{R}$ de classe C^{n+1}. Alors pour tout $N \leqslant n$ nous avons

$$f(b) = f(a) + \sum_{k=1}^{N} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{1}{N!}\int_a^b (b-t)^N f^{(N+1)}(t)dt. \tag{20.543}$$

EQooSCKCooXcKzCc

Démonstration. Notons que dans l'énoncé, n est fixé ; nous faisons une récurrence sur N. Ça ne change pas grand chose, mais il faut être conscient de ce qui est exactement dans l'hypothèse du théorème et ce qui est dans l'hypothèse de récurrence.

Bref, n est fixé, la fonction f est de classe C^{n+1} et nous vérifions d'abord la formule avec $N = 1$. À droite dans (20.543) nous avons

$$f(a) + f'(a)(b-a) + \int_a^b (b-t)f''(t)dt. \tag{20.544}$$

EQooETPRooJHcQXh

Nous évaluons l'intégrale à part en faisant une intégration par parties [29]. Il s'agit de poser

$$u = b - t \tag{20.545a}$$
$$v' = f'', \tag{20.545b}$$

de déduire

$$u' = -1 \tag{20.546a}$$
$$v = f' \tag{20.546b}$$

et d'écrire

$$\int_a^b (b-t)f''(t)dt = \big[(b-t)f'(t)\big]_a^b - \int_a^b (-)f'(t)dt \tag{20.547a}$$
$$= -(b-a)f'(a) + \int_a^b f'(t)dt \tag{20.547b}$$
$$= -(b-a)f'(a) + f(b) - f(a). \tag{20.547c}$$

Dans le calcul nous avons utilisé le théorème fondamental du calcul intégral 14.287. En remettant ça dans (20.544) nous trouvons $f(b)$ comme il se doit.

En ce qui concerne la récurrence, nous devons calculer

$$f(a) + \sum_{k=1}^{N+1} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{1}{(N+1)!}\int_a^b (b-t)^{N+1} f^{(N+2)}(t)dt. \tag{20.548}$$

EQooKQWZooGBytlZ

Ici encore, il s'agit de faire une intégration par partie, et sortir de la somme le terme $k = N+1$. L'intégration par partie donne

$$\int_a^b (b-t)^{N+1} f^{(N+2)}(t)dt = -(b-a)^{N+1} f^{(N+1)}(a) + (N+1)\int_a^b (b-t)^N f^{(N+1)}(t)dt. \tag{20.549}$$

En remettant tout ensemble, il y a encore deux termes qui se simplifient, et des termes qui se remettent pour former la formule de récurrence. Bref, on obtient que (20.548) se réduit bien à $f(b)$. □

Cette formule avec reste intégral sert par exemple à prouver un encadrement pour $\ln(2)$, voir la proposition 20.151.

PROPooWGWKooYydpFg

Proposition 20.148 (Formule de Taylor avec reste intégral[596])**.**
Soit une fonction $f\colon \mathbb{R} \to \mathbb{R}$ de classe C^{m+1} sur un ouvert U contenant $a \in \mathbb{R}$. Alors pour tout $k \leqslant m$ et pour tout $x \in U$ nous avons

$$f(x) = f(a) + \sum_{l=1}^{k} \frac{f^{(l)}(a)}{l!}(x-a)^l + \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k dt \tag{20.550a}$$
$$= \sum_{l=0}^{k} \frac{f^{(l)}(a)}{l!}(x-a)^l + \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k dt \tag{20.550b}$$
$$\tag{20.550c}$$

29. Proposition 20.133.

Démonstration. Nous faisons par récurrence sur k. Nous commençons par voir la formule pour $k = 0$. La formule se réduit à

$$f(x) = f(a) + \int_a^x f'(t)dt, \tag{20.551}$$

qui est correcte par le théorème 14.287 sous la forme (14.906).

La vraie récurrence maintenant. Nous supposons donc avoir la formule

$$f(x) = f(a) + \sum_{l=1}^{k} \frac{f^{(l)}(a)}{l!}(x-a)^l + \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k dt \tag{20.552}$$

EQooJGVGooZBKDyM

valide pour un certain $k \leqslant m$, et nous allons prouver la formule pour $k+1$. L'idée est d'effectuer l'intégrale dans (20.552) par partie (proposition 20.133) en posant

$$\begin{aligned} u &= f^{(k+1)} & dv &= (x-t)^k \\ du &= f^{(k+2)}dt & v &= -\frac{(x-t)^{k+1}}{k+1}. \end{aligned} \tag{20.553}$$

Nous avons

$$\int_a^x f^{(k+1)}(t)(x-t)^k = \left[-f^{(k+1)}(t)\frac{(x-t)^{k+1}}{k+1}\right]_a^x - \int_a^x -\frac{(x-t)^{k+1}}{k+1}f^{(k+2)}(t)dt \tag{20.554a}$$

$$= f^{(k+1)}(a)\frac{(x-a)^{k+1}}{k+1} + \int_a^x \frac{f^{(k+2)}(t)}{k+1}(x-t)^{k+1}dt. \tag{20.554b}$$

En remettant dans (20.552), nous trouvons la réponse. □

THOooFKZZooAgecfp

Théorème 20.149 (Formule de Taylor, reste intégral[107])**.**
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^{m+1} sur un ouvert convexe U. Soient $a \in U$ et $h \in \mathbb{R}^n$ tel que $a + h \in U$. Alors

$$f(a+h) = \sum_{\substack{\alpha\in\mathbb{N}^n\\ |\alpha|\leqslant m}} \frac{h^\alpha}{\alpha!}(\partial^\alpha f)(a) + (m+1)\sum_{|\alpha|=m+1}\frac{h^\alpha}{\alpha!}\int_0^1 (1-t)^m(\partial^\alpha f)(a+th)dt. \tag{20.555}$$

Démonstration. Étant donné que U est connexe et que a et $a+h$ sont dans U, nous pouvons poser

$$\begin{aligned} g\colon [0,1] &\to \mathbb{R} \\ t &\mapsto f(a+th). \end{aligned} \tag{20.556}$$

Cela est une fonction de classe C^{m+1} parce que f l'est (théorème 10.277). Nous appliquons à la fonction g la formule du corolaire 12.302 :

$$g'(t) = h \cdot (\nabla f)(a+th). \tag{20.557}$$

EQooQDLSooGbrePQ

(1) **Une récurrence pour $g^{(k)}$** Par $h \cdot \nabla$ nous entendons l'opérateur $h \cdot \nabla = h_1\partial_1 + \ldots + h_n\partial_n$. Nous prouvons maintenant par récurrence que

$$g^{(k)}(t) = \big[(h\cdot\nabla)^k f\big](a+th). \tag{20.558}$$

En ce qui concerne le cas $k = 1$, c'est une réécriture de la formule (20.557) :

$$g'(t) = h\cdot(\nabla f)(a+th) = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \cdot \begin{pmatrix} (\partial_1 f)(a+th) \\ \vdots \\ (\partial_n f)(a+th) \end{pmatrix} \tag{20.559a}$$

$$= \sum_i h_i(\partial_i f)(a+th) \tag{20.559b}$$

$$= \sum_i (h_i\partial_i f)(a+th) \qquad \text{cf. justif} \tag{20.559c}$$

SUBEQooTRIUooPHkKvE

$$= \big(\sum_i h_i\partial_i f\big)(a+th) \qquad \text{cf. justif.} \tag{20.559d}$$

SUBEQooHXSJooRZUTPT

$$= \big((h\cdot\nabla)f\big)(a+th). \tag{20.559e}$$

Justifications.

— Pour (20.559c). Dans un cas il s'agit de calculer $(\partial_i f)(a+th)$ et de multiplier le résultat par h_i, et dans l'autre cas, de considérer la fonction $(h_i\partial_i f)$ et de l'appliquer à $a+th$.

— Pour (20.559d). Avant, on calculait $(h_i\partial_i f)(a+th)$ et on sommait sur i, et maintenant on applique la fonction $\sum_i h_i\partial_i f$ à $a+th$.

Nous y allons maintenant pour la récurrence. Nous posons

$$s(t) = \big((h\cdot\nabla)^k f\big)(a+th), \tag{20.560}$$

et nous essayons de calculer $s'(t)$. Nous avons

$$\begin{aligned}
s'(t) &= h\cdot\Big(\nabla(h\cdot\nabla)^k f\Big)(a+th) && (20.561a)\\
&= \sum_i h_i\,\mathtt{proj}_i\Big[\big(\nabla(h\cdot\nabla)^k f\big)(a+th)\Big] && (20.561b)\\
&= \sum_i h_i\big(\partial_i(h\cdot\nabla)^k f\big)(a+th) && (20.561c)\\
&= \Big((h\cdot\nabla)^{k+1} f\Big)(a+th). && (20.561d)
\end{aligned}$$

(2) **Expression pour $g^{(k)}$** En mettant les bouts ensemble, nous avons

$$g^{(k)}(t) = (h\cdot\nabla)^k f(a+th). \tag{20.562}$$

Nous prouvons à présent que

$$(h\cdot\nabla)^k = \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=k}} \frac{k!}{\alpha!}h^\alpha\partial^\alpha. \tag{20.563}$$

EQooTBDVooBggqP

Pour cela nous nous souvenons du théorème multinomial 3.48. Il s'applique parce que f est de classe C^{m+1}, ce qui permet de permuter les dérivées [30]. Les « variables » $h_i\partial_i$ sont donc commutatives. D'abord nous avons

$$(h\cdot\nabla)^k = \big(\sum_i h_i\partial_i\big)^k = (h_1\partial_1+\ldots+h_n\partial_n)^k, \tag{20.564}$$

et ensuite nous appliquons le théorème multinomial avec $x_i = h_i\partial_i$.

En ce qui concerne g, nous avons

$$g^{(k)}(t) = (h\cdot\nabla)^k f(a+th) = \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=k}} \frac{k!}{\alpha!}(h^\alpha\partial^\alpha f)(a+th). \tag{20.565}$$

Nous savons que $g(1) = f(a+h)$ et $g(0) = f(a)$. Nous écrivons donc un développement de Taylor pour $g(1)$ autour de $t=0$. Il s'agit d'utiliser la proposition 20.148.

$$\begin{aligned}
g(1) &= \sum_{l=0}^{m}\frac{g^l(0)}{l!}1^l + \frac{1}{m!}\int_0^1 g^{(m+1)}t(t)(1-t)^m dt && (20.566a)\\
&= \sum_{l=0}^{m}\frac{1}{l!}\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=l}}\frac{l!}{\alpha!}(h^\alpha\partial^\alpha f)(a) + \frac{1}{m!}\int_0^1\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=m+1}}\frac{(m+1)!}{\alpha!}(h^\alpha\partial^\alpha f)(a+th)(1-t)^{m+1}dt && (20.566b)\\
&= \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|\leqslant m}}\frac{1}{\alpha!}h^\alpha(\partial^\alpha f)(a) + (m+1)\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=m+1}}\int_0^1\frac{h^\alpha}{\alpha!}(1-t)^m(\partial^\alpha f)(a+th). && (20.566c)
\end{aligned}$$

□

30. Proposition 12.352.

PROPooJFSRooGJcLyv

Proposition 20.150 (Taylor à l'ordre $m = 1$[1])**.**
Soit une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ de classe C^2 sur un ouvert convexe U. Soient $a \in U$ et $h \in \mathbb{R}^n$ tel que $a + h \in U$. Alors

$$f(a+h) = h \cdot (\nabla f)(a) + h \cdot Q(a,h)h \tag{20.567}$$

où $Q(a,h)\colon \mathbb{R}^n \to \mathbb{R}^n$ est l'opérateur linéaire dont les éléments de matrice sont donnés par

$$Q(a,h)_{kl} = \int_0^1 (1-t)(\partial_{kl} f)(a+th)dt. \tag{20.568}$$

Démonstration. Nous commençons par écrire la formule Taylor avec reste intégral du théorème 20.149 avec $m = 1$:

$$f(a+h) = \sum_k h_k (\partial_k f)(a) + 2 \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=2}} \frac{h^\alpha}{\alpha!} \int_0^1 (1-t)(\partial^\alpha f)(a+th)dt \tag{20.569a}$$

$$= h \cdot (\nabla f)(a) + \frac{1}{2} \sum_{k,l} 2 h_k h_l \int_0^1 (1-t)(\partial_{kl} f)(a+th)dt \qquad \text{cf. justif} \tag{20.569b}$$

SUBEQooJLEYooPbiCjj

$$= h \cdot (\nabla f)(a) + \sum_l h_l Q(a,h)h \tag{20.569c}$$

$$= h \cdot (\nabla f)(a) + h \cdot Q(a,h)h \tag{20.569d}$$

$$\tag{20.569e}$$

Justification pour (20.569b). Nous avons utilisé la formule du lemme 3.49 ainsi que le fait que $h^{e_k+e_l} = h_k h_l$. □

20.25.7 Approximation de ln(2)

PROPooHOMYooFclkCU

Proposition 20.151 (Approximation de ln(2)[597])**.**
Pour tout n nous avons

$$\left| \ln(2) - \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} \right| \leqslant \frac{1}{n+1}. \tag{20.570}$$

Démonstration. Nous écrivons la formule de Taylor avec reste intégral du théorème 20.147 pour la fonction $f = \ln$ et pour $a = 1$ et $b = x$. Cela donne :

$$\ln(x) = \ln(1) + \sum_{k=1}^{N} \frac{\ln^{(k)}(1)}{k!}(x-1)^k + \frac{1}{N!} \int_1^x (x-t)^N \ln^{(N+1)}(t)dt. \tag{20.571}$$

Sachant que la dérivée du logarithme[31] est $1/x$ et faisant une petite récurrence, pour $k \geqslant 1$ nous avons

$$\ln^{(k)}(x) = \frac{(-1)^{k+1}(k-1)!}{x^k}. \tag{20.572}$$

En remplaçant,

$$\ln(x) = \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k}(x-1)^k + \int_1^x \frac{(-1)^N (x-t)^N}{t^{N+1}} dt. \tag{20.573}$$

C'est le moment de poser $x = 2$ et de faire les simplifications qui s'imposent,

$$\ln(2) = \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k} + \int_1^2 \frac{(-1)^N (2-t)^N}{t^{N+1}} dt. \tag{20.574}$$

31. Voir la proposition 15.93.

Nous déplaçons la somme à gauche, et nous prenons la valeur absolue des deux côtés :

$$|\ln(2)-\sum_{k=1}^{N}\frac{(-1)^{k+1}}{k}| = |\int_1^2 \frac{(-1)^N(2-t)^N}{t^{N+1}}dt| \tag{20.575a}$$

$$\leqslant \int_1^2 \frac{(2-t)^N}{t^{N+1}}dt \tag{20.575b}$$

SUBEQooZSXEooVcbIpd

$$\leqslant \int_1^2 (2-t)^N dt. \tag{20.575c}$$

SUBEQooHGZLooGIhoVt

Justifications :

— Pour 20.575b. Majoration en rentrant la valeur absolue dans l'intégrale, suppression de $(-1)^N$, et le fait que pour $t\in[1,2]$, $2-t\geqslant 0$.

— Pour 20.575c. Majoration en supprimant purement et simplement le dénominateur $t^N+1\geqslant 1$.

Ais-je vraiment besoin de vous dire que la dernière intégrale se calcule en posant le changement de variables[32] $u=2-t$? Le résultat est que

$$\int_1^2 (2-t)^N dt = \frac{1}{N+1}. \tag{20.576}$$

□

EXooYMEEooMGpUNM

Exemple 20.152 ([1])**.**
La convergence de l'encadrement (20.151) n'est pas terrible. Pour avoir une erreur de $\frac{1}{10}$, il faut

$$\frac{1}{10}=\frac{1}{n+1}, \tag{20.577}$$

ce qui demande $n=9$. Ça reste jouable, même pour les jeunes d'aujourd'hui. Écrivons 9 termes :

$$|\ln(2)-1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\frac{1}{6}-\frac{1}{7}+\frac{1}{8}-\frac{1}{9}|\leqslant\frac{1}{10}. \tag{20.578}$$

En calculant[33],

$$|\ln(2)-\frac{1879}{2520}|\leqslant\frac{1}{10}. \tag{20.579}$$

Voici donc un bel encadrement

$$\frac{1879}{2520}-\frac{1}{10}\leqslant\ln(2)\leqslant\frac{1879}{2520}+\frac{1}{10}. \tag{20.580}$$

Pour avoir quelque chose avec des virgules, d'abord un peu de calcul mental donne

$$\frac{1879}{2520}\simeq 0.745634920634921. \tag{20.581}$$

Donc en majorant et minorant, disons, la troisième décimale[34], on n'est pas moins précis que le $\frac{1}{10}$. On a

$$0.744-\frac{1}{10}\leqslant\ln(2)\leqslant 0.746+\frac{1}{10}. \tag{20.582}$$

Bref, on retient l'approximation

$$0.644\leqslant\ln(2)\leqslant 0.846. \tag{20.583}$$

Pour la quantité de travail, avouez que ce n'est pas terrible comme résultat. Eh oui ; le calcul numérique c'est tout un métier ; il existe des méthodes nettement plus efficaces que ce que nous venons de faire. △

32. Proposition 20.138.
33. Moi j'ai utilisé Sage, mais si tu es au tableau, débrouille-toi.
34. Notez ici que nous utilisons le fait que la division euclidienne, elle, donne un encadrement pour les fractions. Pensez-y.

20.26 Constructions plus naïves de l'intégrale dans le cas réel

Les sections 14.2 et 14.8 ont donné une construction très complète de la mesure de Lebesgue, et nous avons définit la théorie de l'intégration sur un espace mesuré quelconque dans la définition 14.167.

Dans cette section nous allons donner différentes choses plus rapides qui servent souvent de définition dans les cours moins avancés.

20.26.1 Mesure de Lebesgue, version rapide

Nous construisons à présent la mesure de Lebesgue sur $\mathbb{R}^n$. Un **pavé** dans $\mathbb{R}^n$ est un ensemble de la forme

$$B = \prod_{i=1}^{n} [a_i, b_i]; \tag{20.584}$$

le volume d'un tel pavé est défini par $\text{Vol}(B) = \prod_i (b_i - a_i)$. Soit maintenant $A \subset \mathbb{R}^n$. La **mesure externe** de A est le nombre

$$m^*(A) = \inf\{ \sum_{B \in \mathcal{F}} \text{Vol}(B) \text{ où } \mathcal{F} \text{ est un ensemble dénombrable de pavés dont l'union recouvre } A.\} \tag{20.585}$$

DefKTzOlyH

Définition 20.153.
*Nous disons que A est **mesurable** au sens de Lebesgue si pour tout ensemble $S \subset \mathbb{R}^n$ nous avons l'égalité*

$$m^*(S) = m^*(A \cap S) + m^*(S \backslash A). \tag{20.586}$$

Dans ce cas nous disons que la mesure de Lebesgue de A est $m(A) = m^*(A)$.

PropNCMToWI

Proposition 20.154.
Deux fonctions continues égales presque partout pour la mesure de Lebesgue[35] *sont égales.*

Démonstration. Soient f et g deux fonctions continues telles que $f(x) = g(x)$ pour presque tout $x \in D$. La fonction $h = f - g$ est alors presque partout nulle et nous devons prouver qu'elle est nulle sur tout D. La fonction h est continue ; si $h(a) \neq 0$ pour un certain $a \in D$ alors h est non nulle sur un ouvert autour de a par continuité et donc est non nulle sur un ensemble de mesure non nulle. □

20.26.2 Pavés et subdivisions

Définition 20.155.
*Nous appelons **pavé** de $\mathbb{R}^p$ toute partie de $\mathbb{R}^p$ obtenue comme produit de p intervalles de $\mathbb{R}$. Plus explicitement, une partie R est un pavé de $\mathbb{R}^p$ si il s'écrit sous la forme*

$$R = \left\{ (x_1, \ldots, x_p) \in \mathbb{R}^p \,\middle|\, x_i \in \mathcal{I}_i, i = 1, \ldots, p \right\},$$

où $\mathcal{I}_i$ est un intervalle de $\mathbb{R}$ pour tout $i = 1, \ldots, p$.

On appelle pavé fermé de $\mathbb{R}^p$ le produit de p intervalles fermés

$$R = \prod_{i=1}^{p} [a_i, b_i].$$

On définit de même le pavé ouvert

$$S = \prod_{i=1}^{p}]a_i, b_i[.$$

35. Définition 20.153.

Un pavé $R = \prod_{i=1}^{p} \mathcal{I}_i$ est dit borné si tous les intervalles $\mathcal{I}_i$ sont bornés dans $\mathbb{R}$. Les pavés non bornés sont des produits d'intervalles où un (ou plusieurs) des intervalles n'est pas borné. Par exemple,

$$N =]-\infty, 5] \times [0, 13].$$

L'espace $\mathbb{R}^p$, lui-même, est un pavé de $\mathbb{R}^p$.

Définition 20.156.
Une partie A de $\mathbb{R}^p$ est dite ***pavable*** *si il existe une famille finie de pavés R_j, $j = 1, \ldots, n$, et deux à deux disjoints tels que*

$$A = \bigcup_{j=1}^{n} R_j.$$

Un exemple d'ensemble pavable dans $\mathbb{R}^2$ est donné à la figure 20.15. Il existe beaucoup d'ensembles dans $\mathbb{R}^2$ qui ne sont pas pavables, par exemple les ellipses.

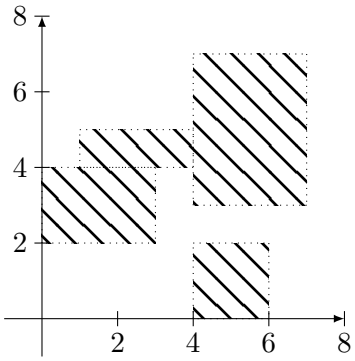


FIGURE 20.15: Un ensemble pavable. LabelFigPolirettangolo

Le complémentaire d'un pavé est un ensemble pavable et, en particulier, tout complémentaire d'un pavé borné est une réunion de pavés non bornés. Toute union finie et toute intersection d'ensemble pavables est pavable.

Définition 20.157.
Soit R un pavé borné de $\mathbb{R}^p$, pour fixer les idées on peut penser $R = \prod_{i=1}^{p} [a_i, b_i]$. On appelle ***longueur*** *de l'i-ème arrête de R le nombre $b_i - a_i$. La* ***mesure p-dimensionnelle de*** *R, $m(R)$, est le produit des longueurs*

$$m(R) = \prod_{i=1}^{p} (b_i - a_i).$$

Exemple 20.158.
Dans $\mathbb{R}^3$, l'ensemble $R = [-1, 1] \times [3, 4] \times [0, 2]$ est un pavé fermé de mesure

$$m(R) = (1 + 1) \cdot (4 - 3) \cdot (2 - 0) = 4.$$

Il s'agit du volume usuel du parallélépipède rectangle. △

Exemple 20.159.
L'ensemble $R =]-1, 1[\times [3, 4] \times [0, 2]$ est un pavé de $\mathbb{R}^3$. Il n'est ni fermé ni ouvert, sa mesure est encore 4. △

Si R est un pavé non borné on peut encore définir sa mesure. La notion de mesure se généralise en deux étapes. D'abord on dit que la longueur d'une arête non bornée est ∞. Ensuite, on adopte la convention $0 \cdot \infty = 0$. Il faut remarquer que avec cette généralisation tout point et toute droite dans $\mathbb{R}^2$ ont mesure nulle.

Afin de définir les intégrales, nous allons intensivement faire appel à la notion de subdivision d'intervalles, voir définition 21.2 et la discussion qui suit.

Lorsqu'on considère un pavé borné $R = \prod_{i=1}^{p} \mathcal{I}_i$ de $\mathbb{R}^p$, on note $\mathcal{S}_i$ l'ensemble des subdivisions de l'intervalle $\mathcal{I}_i$. La notion de subdivision de généralise au cas des pavés.

Définition 20.160.
*Soit R un pavé fermé borné de $\mathbb{R}^p$, pour fixer les idées on peut penser à $R = \prod_{i=1}^{p}[a_i, b_i]$. On appelle **subdivision** finie de R les éléments de l'ensemble $\mathcal{S} = \prod_{i=1}^{p} \mathcal{S}_i$,*

$$\mathcal{S} = \left\{ (Y_1, \dots, Y_p) \,\middle|\, Y_i = (y_{i,j})_{j=1}^{n_i} \in \mathcal{S}_i,\ i = 1, \dots, p \right\}.$$

On peut définir de même l'ensemble des subdivisions d'un pavé non borné.

Souvent, une subdivision d'un pavé $R = \prod_{i=1}^{p} \mathcal{I}_i$ sera noté $\sigma = (y_{i,j})_{j=1}^{n_i}$. Dans cette notation, on sous-entend que pour chaque i fixé, les nombres $y_{i,j}$ (il y en a n_i) forment une subdivision de l'intervalle $\mathcal{I}_i$. Afin de vous familiariser avec ces notations, repérez bien tous les éléments de la figure 20.16.

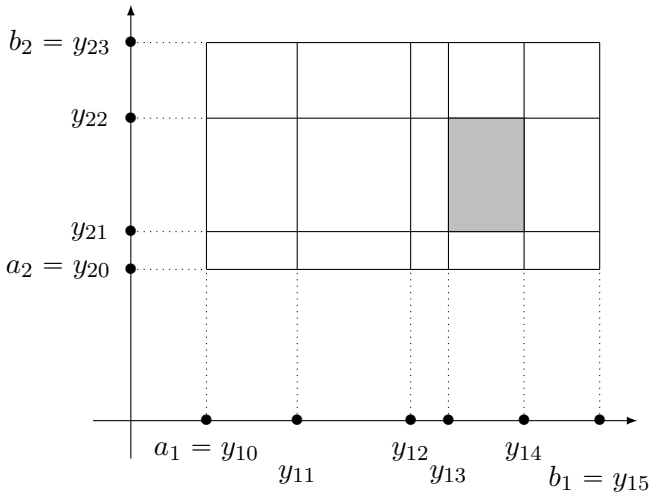


FIGURE 20.16: Une cellule d'une subdivision d'un pavé de $\mathbb{R}^2$. La cellule grisée est $R_{(4,2)}$. LabelFigCelluleCellule

Définition 20.161.
*Si σ est une subdivision d'un pavé R, un **raffinement** de σ est une subdivision de R obtenue en fixant plus de points dans chaque intervalle.*

La subdivision σ de R détermine $n_1 n_2 \dots n_p$ pavés fermés de la forme

$$R_{(k_1,\dots,k_p)} = \{(x_1, \dots, x_p) \in \mathbb{R}^p \,|\, y_{i,k_{i-1}} \leqslant x_i \leqslant y_{i,k_i}\},$$

où k_i est dans $\{1, \dots, n_i\}$ et i dans $\{1, \dots, p\}$. On les appelles **cellules** de σ. On remarque que les cellules de σ sont toujours deux à deux disjointes (sauf au plus sur leurs bords).

meas_sous

Lemme 20.162.
Soit R un pavé borné de $\mathbb{R}^p$ et soit $\sigma = (y_{i,j})_{j=1}^{n_i}$ une subdivision de R. On a

$$m(R) = \sum_{(k_1,\dots,k_p)\in K} m(R_{(k_1,\dots,k_p)}),$$

où $K = \{1, \dots, n_1\} \times \{1, \dots, n_2\} \times \dots \times \{1, \dots, n_p\}$.

Le lemme 20.162 suggère de définir la mesure d'un ensemble borné pavable $P = \bigcup_{j=1}^{n} R_j$ comme la somme des mesures des pavés disjoints R_j, $j = 1, \dots, n$.

Définition 20.163.
*Une application $f : \mathbb{R}^p \to \mathbb{R}$ est dite **application en escalier** sur $\mathbb{R}^p$ si*

— *f est une application bornée,*

— *il existe une subdivision σ de $\mathbb{R}^p$ telle que la restriction de f est une application constante sur toute cellule R_k de σ*

$$f_{|_{R_k}} = C_k, \qquad C_k \in \mathbb{R},$$

*Une telle subdivision σ est dite **associée** à f.*

Exemple 20.164.
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in [0,3] \times [-1,2], \\ 2 & \text{sinon.} \end{cases} \tag{20.587}$$

est une application en escalier. Exercice : donner une subdivision de $\mathbb{R}^2$ associée à cette fonction. △

Exemple 20.165.
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{1}{m^2+n^2}, & \text{si } (x,y) \in [m,m+1] \times [n,n+1], \quad m,\, n \in \mathbb{N}_0, \\ 0, & \text{sinon} \end{cases} \tag{20.588}$$

est une application en escalier. Observez que, dans ce cas, il n'existe pas une subdivision finie de $\mathbb{R}^2$ associée à f. △

Remarque 20.166.
Si la subdivision σ est associée à f alors tout raffinement de σ (c'est-à-dire, toute subdivision obtenue en fixant plus de points dans chaque intervalle) a la même propriété.

Si f et g sont deux applications en escalier sur R et σ_f et σ_g sont des subdivisions de R associées respectivement à f et g, alors on peut construire une troisième subdivision de R qui est associée à f et à g en même temps. Soient $\sigma_f = (Y_1, \ldots, Y_p)$ et $\sigma_g = (Z_1, \ldots, Z_p)$, où $Y_i = (y_{i,j})_{j=1}^{m_i}$ et $Z_i = (z_{i,j})_{j=1}^{n_i}$ sont des subdivisions de l'intervalle $[a_i, b_i]$, pour $i = 1, \ldots, p$. La subdivision de $[a_i, b_i]$ obtenue par l'union de Y_i et Z_i est encore une subdivision finie, qu'on appellera $\bar{Y}_i$. La subdivision $\bar{\sigma} = (\bar{Y}_1, \ldots, \bar{Y}_p)$ de R est un raffinement de σ_f et de σ_g, donc elle est associée à la fois à f et à g.

Cela nous permet de prouver que si f et g sont des applications en escalier, alors $f+g$, fg, $\min\{f,g\}$, $\max\{f,g\}$ et $|f|$ sont des applications en escalier.

20.26.3 Intégrale d'une fonction en escalier

Définition 20.167.
*Soit f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Le **support** de f est la fermeture de l'ensemble des points x tels que $f(x) \neq 0$.*

Définition 20.168.
*Une application en escalier f est dite **intégrable** si son support est compact.*

Soit f une application en escalier sur $\mathbb{R}^p$. Soit σ une subdivision de $\mathbb{R}^p$ associée à f et appelons R_k les cellules de σ, avec $k = (k_1, \ldots, k_p)$ dans $K = \{1, \ldots, n_1\} \times \{1, \ldots, n_2\} \times \ldots \times \{1, \ldots, n_p\}$. Alors

$$f_{|_{R_k}} = C_k, \qquad C_k \in \mathbb{R}.$$

Définition 20.169.
*On définit l'**intégrale** de f sur $\mathbb{R}^p$ par*

$$\int_{\mathbb{R}^p} f\, dV = \sum_{k \in K} C_k m(R_k).$$

L'intégrale ainsi définie est un nombre réel. La proposition suivante nous dit que l'intégrale est « bien définie », au sens que sa valeur ne dépend pas de la subdivision associée à f qu'on utilise dans le calcul.

Proposition 20.170.
Soit f une application en escalier intégrable sur $\mathbb{R}^p$. Soient σ_1 et σ_2 deux subdivisions de $\mathbb{R}^p$ associées à f. L'intégrale de f ne dépend pas de la subdivision choisie.

On ne donne pas une preuve complète de cette proposition. En fait elle est une conséquence de la formule de réduction introduite dans la suite de ce chapitre.

20.26.4 Intégrales partielles

Soit f de $\mathbb{R}^p$ dans $\mathbb{R}$ une fonction continue, nulle hors du pavé borné R. Posons $R = \prod_{i=1}^{p}[a_i, b_i]$, pour fixer les idées. Pour chaque i dans $\{1, \dots, p\}$ fixé, on peut associer à f la fonction F_i de $p-1$ variables définie par

$$F_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) = \int_{a_i}^{b_i} f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_p)\, dy.$$

La fonction F_i est l'intégrale partielle de f par rapport à la i-ème variable. En particulier, si $f(x_1, \dots, x_p) = g(x_i)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ on obtient

$$F_i = \int_{a_i}^{b_i} g(y)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)\, dy = h \cdot \int_{a_i}^{b_i} g\, dy.$$

La fonction d'une seule variable qu'on obtient à partir de f en fixant $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p$ et qui associe à x_i la valeur $f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_p)$, est appelée x_i-ème section de f en $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p$.

Exemple 20.171.
Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x, y) = \begin{cases} x + 3y & \text{si } (x, y) \in [9, 10] \times]\pi, 5] \\ 0 & \text{sinon.} \end{cases} \tag{20.589}$$

Les intégrales partielles de f sont

$$F_1(y) = \int_9^{10} x + 3y\, dx = \left[\frac{x^2}{2} + 3xy\right]_{x=9}^{x=10} = \frac{19}{2} + 3y,$$

$$F_2(x) = \int_\pi^5 x + 3y\, dy = \left[xy + \frac{3y^2}{2}\right]_{y=\pi}^{y=5} = x(5 - \pi) + \frac{3}{2}(25 - \pi^2).$$

△

20.26.5 Réduction d'une intégrale multiple

Soit $R = [a, b] \times [c, d]$ un pavé fermé et borné de $\mathbb{R}^2$ et soit f une application en escalier intégrable sur $\mathbb{R}^2$ telle que le support de f soit contenu dans R. On considère la subdivision σ de R définie par les subdivisions

$$a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_m = b,$$

$$c = y_0 \leqslant y_1 \leqslant \ldots \leqslant y_n = d.$$

Les cellules de σ sont

$$R_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], \qquad i = 0, \dots, m-1, \quad j = 0, \dots, n-1.$$

La mesure de R est la somme des mesures des $R_{i,j}$

$$\begin{aligned} m(R) &= \sum_{(i,j)\in\{0,\dots,m-1\}\times\{0,\dots,n-1\}} m(R_{i,j}) = \\ &= \sum_{j=0}^{n-1}\sum_{i=0}^{m-1}(x_{i+1}-x_i)\cdot(y_{i+1}-y_i) = \\ &= \sum_{i=0}^{m-1}(x_{i+1}-x_i)\cdot\sum_{j=0}^{n-1}(y_{i+1}-y_i) = \\ &= (b-a)\cdot(d-c). \end{aligned} \tag{20.590}$$

Si f est constante sur chaque cellule de σ on peut écrire f de la forme suivante

$$f(x,y) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{R_{i,j}}$$

où les $C_{i,j}$ sont des constantes réelles et $\chi_{R_{i,j}}$ est la **fonction caractéristique** de $R_{i,j}$

$$\chi_{R_{i,j}}(x,y) = \begin{cases} 1, & \text{si } (x,y)\in R_{i,j}, \\ 0, & \text{sinon.} \end{cases} \tag{20.591}$$

Comme (x,y) est dans $R_{i,j}$ si et seulement si $x\in[x_i,x_{i+1}]$ et $y\in[y_j,y_{j+1}]$, on vérifie que la fonction $\chi_{R_{i,j}}$ est égal au produit des fonctions caractéristiques des intervalles $[x_i,x_{i+1}]$ et $[y_j,y_{j+1}]$

$$\chi_{R_{i,j}}(x,y) = \chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y).$$

On peut donc écrire la fonction f de la façon suivante

$$f(x,y) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y).$$

Comme on suppose que le support de f est une partie de R, l'intégrale de f sur $\mathbb{R}^2$ est

$$\int_{\mathbb{R}^2} f\,dV = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,m(R_{i,j}) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,(x_{i+1}-x_i)\cdot(y_{j+1}-y_j). \tag{20.592}$$

Cette intégrale peut être réduite à la composition de deux intégrales partielles. Il suffit de remarquer que la valeur de l'intégrale de la fonction caractéristique d'un intervalle est la longueur de l'intervalle,

$$\begin{aligned} C_{i,j}(x_{i+1}-x_i)\cdot(y_{j+1}-y_j) &= \\ &= C_{i,j}\left(\int_{x_i}^{x_{i+1}}\chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_{y_j}^{y_{j+1}}\chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= C_{i,j}\left(\int_a^b\chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_c^d\chi_{[y_j,y_{j+1}]}(y)\,dy\right), \end{aligned} \tag{20.593}$$

et utiliser les propriétés de linéarité de l'intégrale

$$\begin{aligned} \int_{\mathbb{R}^2} f\,dV &= \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\left(\int_a^b\chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_c^d\chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= \int_c^d\int_a^b\sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_c^d\int_a^b f\,dxdy. \end{aligned} \tag{20.594}$$

De même on obtient

$$\begin{aligned}\int_{\mathbb{R}^2} f\,dV &= \int_a^b \int_c^d \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_a^b\int_c^d f\,dxdy.\end{aligned} \tag{20.595}$$

En général, on prouve la proposition suivante

Proposition 20.172.
Soit f une application en escalier intégrable sur $\mathbb{R}^p$ et soit R un pavé borné dans $\mathbb{R}^p$ qui contient le support de f. Comme d'habitude, pour fixer les idées nous écrivons $R = \prod_{i=1}^p [a_i, b_i]$. Alors

$$\begin{aligned}\int_{\mathbb{R}^p} f(x_1,\dots,x_p)\,dV &= \int_{a_p}^{b_p}\int_{a_{p-1}}^{b_{p-1}}\cdots\int_{a_1}^{b_1} f(x_1,\dots,x_p)\,dx_1\cdots dx_p = \\ &= \int_{a_{s_p}}^{b_{s_p}}\int_{a_{s_{p-1}}}^{b_{s_{p-1}}}\cdots\int_{a_{s_1}}^{b_{s_1}} f(x_1,\dots,x_p)\,dx_1\cdots dx_p,\end{aligned} \tag{20.596}$$

pour toute permutation $(s_1,\dots,s_p)$ de l'ensemble $\{1,\dots p\}$.

20.26.6 Propriétés de l'intégrale

Soient f et g deux fonctions en escalier intégrables de $\mathbb{R}^p$ dans $\mathbb{R}$, et soient a et b dans $\mathbb{R}$.

Linéarité de l'intégrale :

— Additivité : $f+g$ est intégrable et

$$\int_{\mathbb{R}^p}(f+g)\,dV = \int_{\mathbb{R}^p} f\,dV + \int_{\mathbb{R}^p} g\,dV,$$

— Homogénéité : λf est intégrable pour tout réel λ

$$\int_{\mathbb{R}^p}\lambda f\,dV = \lambda\int_{\mathbb{R}^p} f\,dV,$$

Monotonie Si $f \leqslant g$ alors

$$\int_{\mathbb{R}^p} f\,dV \leqslant \int_{\mathbb{R}^p} g\,dV,$$

Inégalité fondamentale

$$\left|\int_{\mathbb{R}^p} f\,dV\right| \leqslant \int_{\mathbb{R}^p}|f|\,dV.$$

Cette dernière inégalité s'obtient de la façon suivante :

$$\left|\int_{\mathbb{R}^p} f\,dV\right| = \left|\sum_{k\in K} C_k m(R_k)\right| \leqslant \sum_{k\in K}|C_k| m(R_k) = \int_{\mathbb{R}^p}|f|\,dV.$$

Inégalité de Čebičeff Si f est une application en escalier alors pour tout $a>0$ dans $\mathbb{R}$ l'ensemble $\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}$ est pavable et borné, et l'inégalité suivante est satisfaite

$$m\left(\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}\right) \leqslant \frac{1}{a}\int_{\mathbb{R}^p}|f|\,dV.$$

20.26.7 Intégrales multiples, cas général

Nous voulons généraliser la définition d'intégrale multiple au cas des domaines non pavables et de fonctions qui ne sont pas en escalier. Il y a plusieurs méthodes de le faire et ici on ne considère qu'une seule, introduite par Riemann.

Définition 20.173.
Soit $f : \mathbb{R}^p \to \mathbb{R}$ une fonction.

- *Pour toute application en escalier intégrable f_* telle que $f_* \leqslant f$, l'intégrale de f_* est dit une* ***somme inférieure*** *de f.*
- *Pour toute application en escalier intégrable f^* telle que $f^* \geqslant f$, l'intégrale de f^* est dit une* ***somme supérieure*** *de f.*

Soient $\sum_* f$ et $\sum^* f$ les ensembles des sommes inférieures et supérieures de f. Grâce à la propriété de monotonie de l'intégrale on sait que si a est dans $\sum_* f$ et b est dans $\sum^* f$ alors $a \leqslant b$.

Définition 20.174.
La fonction f est intégrable (au sens de Riemann) si $\sum_* f$ et $\sum^* f$ ne sont pas vides et

$$\inf \Sigma^* f = I = \sup \Sigma_* f.$$

Dans ce cas, la valeur I est appelée intégrale de f sur $\mathbb{R}^p$.

Remarque 20.175.
Toute fonction intégrable est bornée et à support compact. En effet, si le support de la fonction n'est pas compact alors soit $\sum_* f$ soit $\sum^* f$ doit être vide !

L'intégrale qu'on vient de définir possède toutes les propriétés de l'intégrale pour les fonctions en escalier. Le produit de deux fonctions intégrables est intégrable.

Il y a des cas où l'intégrabilité d'une fonction n'est pas évidente. Cependant, dans la plupart des exercices et des exemples de ce cours, nous nous aidons avec le critère suivant

Proposition 20.176.
Toute fonction continue à support compact est intégrable.

Cette proposition n'est à priori pas étonnante, vu qu'une fonction continue sur un support compact est bornée (théorème de Weierstrass 7.146).

20.26.8 Réduction d'une intégrale multiple

On n'utilise jamais la définition pour calculer la valeur d'une intégrale multiple. La méthode plus efficace, en pratique, est de réduire l'intégrale à la composition de plusieurs intégrales d'une variable.

fub

Théorème 20.177 (de Fubini)**.**
Soit f une fonction intégrable de $\mathbb{R}^2$ dans $\mathbb{R}$. Si pour tout x dans $\mathbb{R}$ la section $f(x, \cdot\,)$ est intégrable par rapport à y, alors

$$\int_{\mathbb{R}^2} f(x,y)\,dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,dx \right) dy.$$

De même, si pour tout y dans $\mathbb{R}$ la section $f(\,\cdot\,, y)$ est intégrable par rapport à x, alors

$$\int_{\mathbb{R}^2} f(x,y)\,dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,dy \right) dx.$$

ThoSectionINte

En général, on ne peut pas dire que les sections d'une fonction intégrable sont intégrables, donc il faut vraiment se souvenir des hypothèses du théorème 20.177. En dimension plus haute, on a le même résultat

Théorème 20.178.
Soit f une fonction intégrable de $\mathbb{R}^p$ dans $\mathbb{R}$. Si pour tout $(p-1)$-uple $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ dans $\mathbb{R}^{p-1}$ la section $f(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_p)$ est intégrable par rapport à x_i, alors

$$\int_{\mathbb{R}^p} f\, dV = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{p-1}} f\, dV \right) dx_i.$$

Si f est une fonction positive et intégrable de $\mathbb{R}^2$ dans $\mathbb{R}$ on peut interpréter l'intégrale de f comme le volume du solide au-dessous du graphe de f. Avec cette interprétation, l'intégrale partielle par rapport à x pour $y = y_0$ fixé est l'aire de la tranche qu'on obtient en coupant le solide par le plan $y = y_0$.

Exemple 20.179.
Le premier exemple à faire est celui d'une fonction en escalier intégrable et positive. Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ la fonction

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in R_1 =]-1,3] \times [4,5] \\ 3 & \text{si } (x,y) \in R_2 =]13,15[\times [0,2[\\ 0 & \text{dans les autres cas.} \end{cases} \tag{20.597}$$

L'intégrale de f sur $\mathbb{R}^2$ est $1 \cdot m(R_1) + 3 \cdot m(R_2) = 16$. On voit tout de suite qu'il s'agit de la somme du volume des deux parallélépipèdes de hauteurs respectives 1 et 3 et bases R_1 et R_2. △

Exemple 20.180.
On veut calculer le volume du solide S, borné par le paraboloïde elliptique $x^2 + 2y^2 + z = 16$ et le plan $x = 2$, $x = 0$, $y = 2$ $y = 0$, $z = 0$. On observe que la portion de paraboloïde elliptique qui nous intéresse est le graphe de la fonction $f(x,y) = 16 - x^2 - 2y^2$ pour (x,y) dans $R = [0,2] \times [0,2]$. La fonction f est continue ainsi que ses sections, donc on peut appliquer le théorème 20.177 et décomposer l'intégrale double en deux intégrales simples :

$$\begin{aligned} &\int_R 16 - x^2 - 2y^2\, dV = \int_0^2 \int_0^2 f(x,y)\, dxdy = \\ &= \int_0^2 \left[(16 - 2y^2)x - \frac{x^3}{3} \right]_{x=0}^{x=2} dy = \\ &= \left[\left(32 - \frac{8}{3} \right) y - \frac{4y^3}{3} \right]_{x=0}^{x=2} = 64 - \frac{16+32}{3} = 48. \end{aligned} \tag{20.598}$$

Vérifiez, comme exercice, qu'on obtient le même résultat en intégrant d'abord par rapport à y et puis par rapport à x. △

Exemple 20.181.
Dans les hypothèses du théorème 20.177 l'ordre des intégrations partielles ne change pas la valeur de l'intégrale. En fait, si les calculs sont faits par des êtres humains l'ordre d'intégration peut faire une certaine différence comme dans cet exemple. On veut évaluer la valeur de l'intégrale

$$\int_{\mathbb{R}^2} f(x,y)\, dV$$

où

$$f(x,y) = \begin{cases} y \sin(xy) & \text{si } (x,y) \in [1,2] \times [0,\pi] \\ 0 & \text{sinon.} \end{cases} \tag{20.599}$$

Les deux sections de $f(x,y) = y \sin(xy)$ sont continues. Si on intègre d'abord par rapport à y on obtient

$$-\int_1^2 \frac{\pi \cos(\pi x)}{x} dx + \int_1^2 \frac{\sin(\pi x)}{x^2} dx,$$

qui n'est pas du tout immédiat, alors que, si on intègre d'abord par rapport à x on obtient

$$\int_0^\pi \cos y - \cos(2y)\, dy.$$

△

20.26.9 Intégrales sur des parties de $\mathbb{R}^2$

On veut évaluer l'intégrale de la fonction $f(x,y) = \sqrt{1-x^2}$ sur son domaine, la boule unité $B((0,0),1)$. La théorie introduite jusqu'ici n'est pas suffisante pour résoudre ce problème, parce que $B((0,0),1)$ n'est pas pavable. Les parties bornées de $\mathbb{R}^p$ sur lesquelles on peut intégrer des fonctions sont dites mesurables (au sens de Riemann) parce que, comme on verra dans la suite, la mesure d'une partie de $\mathbb{R}^p$ est l'intégrale (si il existe) de sa fonction caractéristique.

On peut dire qu'une partie de $\mathbb{R}^p$ est mesurable si son bord est «assez régulier». Dans $\mathbb{R}^2$ il est suffisant que le bord de A soit une réunion finie de courbes paramétrées continues. En particulier, on est très souvent dans un des deux cas suivants

Régions du premier type A est borné et contenu entre les graphes de deux fonctions continues de x

$$A = \{(x,y) \in \mathbb{R}^2 \,:\, a \leqslant x \leqslant b,\, g_1(x) \leqslant y \leqslant g_2(x)\},$$

avec g_1 et g_2 continues.

Régions du deuxième type A est borné et contenu entre les graphes de deux fonctions continues de y

$$A = \{(x,y) \in \mathbb{R}^2 \,:\, c \leqslant y \leqslant d,\, h_1(y) \leqslant x \leqslant h_2(y)\},$$

avec h_1 et h_2 continues.

Exemple 20.182.
Il y a des régions qui sont des deux types au même temps, comme les boules centrées à l'origine, le triangle de sommets $(0,0)$, $(0,a)$ et $(b,0)$, ou la région C délimité par les courbes $y = 2x$ et $y = x^2$. Cette dernière admet les représentations suivantes

$$C = \{(x,y) \in \mathbb{R}^2 \,:\, 0 \leqslant x \leqslant 1,\, x^2 \leqslant y \leqslant 2x\},$$

et

$$C = \{(x,y) \in \mathbb{R}^2 \,:\, 0 \leqslant y \leqslant 1,\, y/2 \leqslant x \leqslant \sqrt{y}\}.$$

△

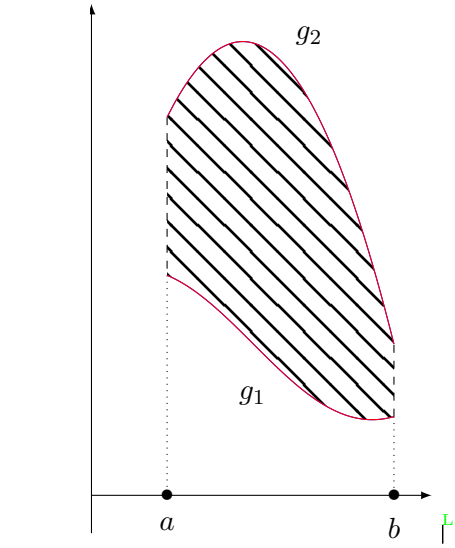
Définition 20.183.
Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ dont le support A est une région du premier ou du deuxième type. On définit la fonction $\bar{f}$ comme

$$\bar{f}(x,y) = \begin{cases} f(x,y), & \text{si } (x,y) \in A, \\ 0, & \text{sinon.} \end{cases} \tag{20.600}$$

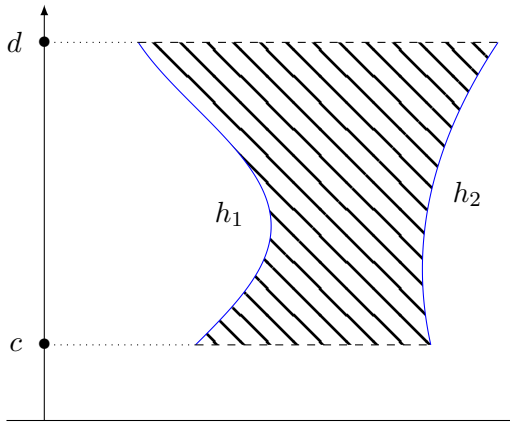
*La fonction f est dite **intégrable** si $\bar{f}$ est intégrable, et la valeur de son intégrale est*

$$\int_A f\, dV = \int_{\mathbb{R}^2} \bar{f}\, dV.$$

Une fonction continue définie sur une région du premier ou du deuxième type est toujours intégrable.



(a) Une region du premier type



(b) Une region du deuxième type

FIGURE 20.17: Régions du premier et du deuxième type. LabelFigRegioniPrimoeSecondoTipo

Pour fixer les idées on suppose ici que A est du premier type et contenue dans le pavé borné $R = [a, b] \times [c, d]$. En suivant la définition on obtient

$$\begin{aligned}\int_A f\,dV &= \int_{\mathbb{R}^2} \bar{f}\,dV \\ &= \int_a^b \int_c^d \bar{f}\,dydx \\ &= \int_a^b \left(\int_c^{g_1(x)} \bar{f}\,dy + \int_{g_1(x)}^{g_2(x)} \bar{f}\,dy + \int_{g_2(x)}^d \bar{f}\,dy \right) dx \\ &= \int_a^b \int_{g_1(x)}^{g_2(x)} f\,dydx.\end{aligned} \tag{20.601}$$

De même, si A est du deuxième type on obtient

$$\int_A f\,dV = \int_c^d \int_{h_1(y)}^{h_2(y)} f\,dxdy. \tag{20.602}$$

Exemple 20.184.
On peut maintenant résoudre notre problème de départ, évaluer l'intégrale de la fonction $f(x, y) = \sqrt{1 - x^2}$ sur $B((0, 0), 1)$. Nous choisissons de décrire la boule unité de $\mathbb{R}^2$ comme une région du

premier type : $B((0,0),1) = \{(x,y) \,:\, x \in [-1,1],\, -\sqrt{1-x^2} \leqslant y \leqslant \sqrt{1-x^2}\}$.

$$I = \int_B \sqrt{1-x^2}\, dV = \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2} dy dx \tag{20.603}$$

La première intégrale à effectuer, par rapport à y, est l'intégrale d'une fonction constante. Ne pas oublier que l'on intègre $\sqrt{1-x^2}$ par rapport à y ; c'est bien une constante et l'intégrale consiste seulement à multiplier par y :

$$I = \int_{-1}^{1} \left[y\sqrt{1-x^2} \right]_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} dx = 2\int_{-1}^{1} (1-x^2) dx. \tag{20.604}$$

Cela est à nouveau une intégrale simple à effectuer. Le résultat est

$$2\int_{-1}^{1} (1-x^2) dx = 2\left[x - \frac{x^3}{3} \right]_{x=-1}^{x=1} = \frac{8}{3}. \tag{20.605}$$

△

Remarque 20.185.
Toutes les techniques d'intégration à une variable restent valables. Par exemple, lorsqu'une des intégrales est l'intégrale d'une fonction impaire sur un intervalle symétrique par rapport à zéro, l'intégrale vaut zéro.

NORMooDSNXooFhyHkx

20.186.
Par le lemme 14.175 nous savons que la mesure d'une région bornée de $\mathbb{R}^2$ est l'intégrale de sa fonction caractéristique, si elle existe.

La mesure d'une région bornée de $\mathbb{R}^2$ est dite son **aire**, et celle d'une région bornée de $\mathbb{R}^3$ est son **volume**. Voir aussi la remarque 21.10.

exint

Exemple 20.187.
On veut calculer l'aire de la région de la figure 20.18 définie par

$$A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, x^3 - 1 \leqslant y \leqslant x\}.$$

On considère l'intégrale

$$\int_{\mathbb{R}^2} \chi_A \, dV = \int_0^1 \int_{x^3+1}^{x} 1 \, dy \, dx = \int_0^1 -x^3 + x + 1 \, dx = -\frac{1}{4} + \frac{1}{2} + 1 = \frac{5}{4}.$$

△

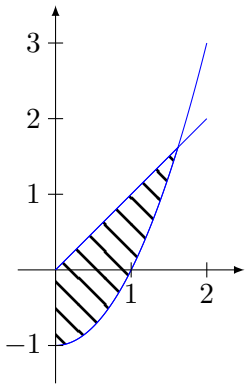


FIGURE 20.18: La région de l'exemple 20.187. Dans l'exemple, l'image est coupée en $x = 1$.

LabelFigExampleIntegration

Exemple 20.188.
Parfois la région sur laquelle on veut intégrer peut être décrite indifféremment de deux façons, mais la fonction à intégrer nous force a choisir un ordre particulier. Vérifiez que la fonction $f(x,y) = \sin(y^2)$ sur la région triangulaire de sommets $(0,0)$, $(0,2)$, $(2,2)$ doit être intégrée d'abord par rapport à x. △

Si une région bornée n'est pas de premier ou de deuxième type on peut normalement la découper en morceaux plus faciles à décrire. On utilise alors la propriété suivante.

Lemme 20.189.
Soit A un sous-ensemble borné de $\mathbb{R}^2$ et soient B_1 et B_2 deux parties de A telles que $B_1 \cap B_2 = \varnothing$ et $B_1 \cup B_2 = A$. Alors, pour toute fonction f intégrable sur A (et en particulier pour sa fonction caractéristique) on a

$$\int_A f\,dV = \int_{B_1} f\,dV + \int_{B_2} f\,dV.$$

exint2

Exemple 20.190.
La région D que nous voyons sur la figure 20.19 est bornée par la parabole $y^2 = 2x+6$ et la droite $y = x-1$. La région D est une région du deuxième type. Nous pouvons aussi la décrire comme l'union de deux régions du premier type D_1 et D_2,

$$D_1 = \{(x,y) \,:\, -3 \leqslant x \leqslant -1,\, -\sqrt{2x+6} \leqslant y \leqslant \sqrt{2x+6}\},$$

et

$$D_2 = \{(x,y) \,:\, -3 \leqslant x \leqslant -1,\, x-1 \leqslant y \leqslant \sqrt{2x+6}\}.$$

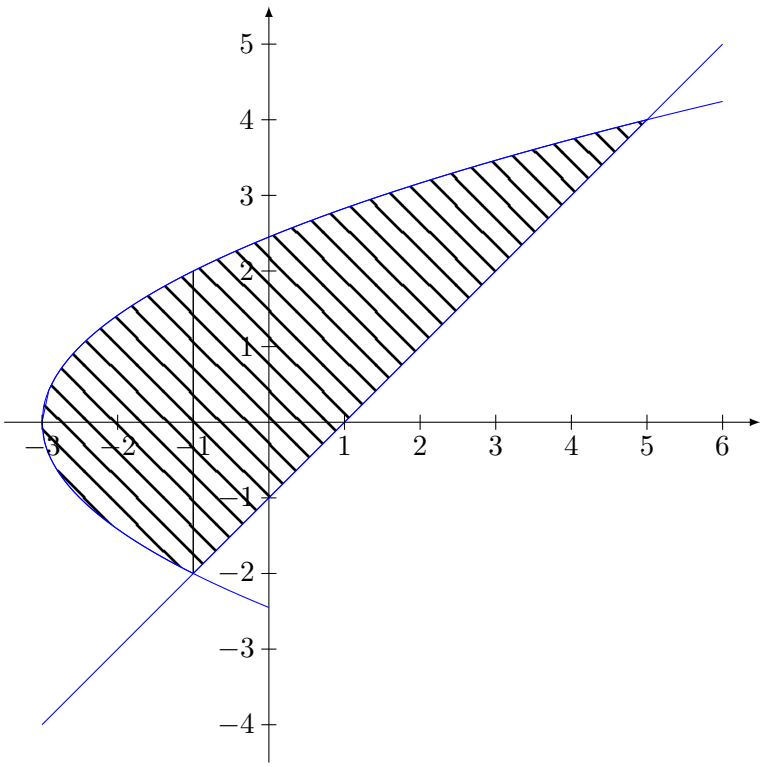


FIGURE 20.19: La région D de l'exemple 20.190 LabelFigExampleIntegrationdeux

△

20.26.10 Intégrales sur des parties de $\mathbb{R}^3$

Dans ces notes nous n'avons pas l'ambition de traiter d'une façon rigoureuse l'étude des ensembles mesurables de $\mathbb{R}^3$. Comme dans la section précédente on se limitera à considérer des cas particuliers.

primotipo_solida

Définition 20.191.
Soit E une région de $\mathbb{R}^3$. On dit que E est une ***région solide de premier type*** *si E est contenue entre les graphes de deux fonctions continues de x et y.*

$$E = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in A \subset \mathbb{R}^2, u_1(x,y) \leqslant z \leqslant u_2(x,y)\}.$$

Le sous-ensemble de A de $\mathbb{R}^2$ qui apparaît dans la définition 20.191 est la projection (ou l'ombre) de E sur le plan x-y.

cornet

Exemple 20.192.
La région E donnée par une portion de sphère collée à un cône est une région solide de premier type

$$E = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in \overline{B\big((0,0),1\big)}, \sqrt{x^2+y^2} \leqslant z \leqslant \sqrt{1-x^2-y^2}\}. \tag{20.606}$$

L'ombre de E est la boule unité de $\mathbb{R}^2$. L'ensemble $\sqrt{x^2+y^2} \leqslant z$ est un cône posé sur sa pointe tandis que l'ensemble $z \leqslant \sqrt{1-x^2-y^2}$ est la demi-sphère. L'ensemble E contient les points entre les deux, voir la figure 20.20.

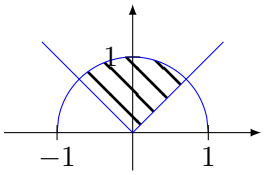


FIGURE 20.20: Il faut voir ça en trois dimensions.

LabelFigCornetGlace

△

Si la fonction f, à intégrer sur E, et ses sections sont intégrables alors on peut réduire l'intégrale

$$\begin{aligned}\int_E f(x,y,z)\,dV &= \int_A \left(\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz \right) dV = \\ &= \int_A \left(F(x,y,u_2(x,y)) - F(x,y,u_1(x,y))\right)\,dV,\end{aligned} \tag{20.607}$$

où F est une primitive de f par rapport à la variable z, c'est-à-dire en considérant x et y comme des constantes. Il faut ensuite évaluer la partie qui reste comme dans la section précédente. Comme le calcul des aires dans $\mathbb{R}^2$, le calcul des volumes dans $\mathbb{R}^3$ est fait par des intégrales. En fait le **volume** d'une région solide dans $\mathbb{R}^3$ est sa mesure.

Définition 20.193.
La mesure d'une région de $\mathbb{R}^3$ est l'intégrale de sa fonction caractéristique.

Soit E une région solide du premier type, nous pouvons évaluer son volume par l'intégrale

$$\int_A \left(u_2(x,y) - u_1(x,y)\right)\,dV.$$

Parfois c'est plus intéressant de calculer le volume avec la formule de réduction contraire : l'intégrale double d'abord et puis l'intégrale simple par rapport à z. On parle alors de calcul de volume « par tranche ».

Exemple 20.194.
On veut calculer le volume de la boule de rayon a, centrée à l'origine $B = \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 \leqslant a^2\}$. On peut décrire B par

$$B = \left\{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in D_a, -\sqrt{a^2 - x^2 - y^2} \leqslant z \leqslant \sqrt{a^2 - x^2 - y^2}\right\},$$

où D_a est le disque de rayon a centré en $(0,0)$, donc le volume B sera

$$2\int_{D_a} \sqrt{a^2 - x^2 - y^2} dV.$$

Cette intégrale est un peu ennuyeuse à calculer. On peut simplifier le calcul en observant que pour $\bar{z}$ fixé dans l'intervalle $[-a,a]$ la section de la boule au niveau $\bar{z}$ est un disque de rayon $\sqrt{a^2 - z^2}$. L'aire d'un tel disque est $\pi(a^2 + z^2)$. Si on réduit l'intégrale de volume de la façon

$$\int_B 1\, dV = \int_{-a}^{a} \sqrt{a^2 - z^2}\, dz,$$

on obtient tout de suite la valeur cherchée : le volume de B est $4/3\pi a^3$. △

Exemple 20.195.
On calcule l'intégrale de $f(x,y,z) = z$ sur la pyramide P bornée par les plans $x = 0$, $y = 0$, $x + y + z = 1$, $x + y + z/2 = 1$. On remarque tout de suite que les plans $x + y + z = 1$, $x + y + z/2 = 1$ se coupent en la droite $x + y = 1$, $z = 0$ (on se souvient qu'*une* droite dans $\mathbb{R}^3$, c'est *deux* équations). Cela veut dire que la projection de P sur le plan x-y est le triangle T borné par les droites $x = z = 0$, $y = z = 0$ et $x + y = 1$, $z = 0$. On décrit donc P par

$$P = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in T,\, 1 - 2x - 2y \leqslant z \leqslant 1 - x - y\}$$

et T par

$$T = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1,\, 0 \leqslant y \leqslant 1 - x\},$$

donc l'intégrale de f sur P est

$$\int_P f(x,y,z)\, dV = \int_0^1 \int_0^{1-x} \int_{1-2x-2y}^{1-x-y} z\, dz\, dy\, dx = -\frac{1}{24}.$$

Notez que lorsque x et y sont entre 0 et 1, nous avons bien $1 - 2x - 2y < 1 - x - y$, d'où le fait que nous mettons $1 - 2x - 2y$ dans la borne inférieure de l'intégrale. △

De façon analogue on définit les régions solides du deuxième et du troisième type.

20.26.11 Intégrales de fonctions non bornées sur des ensembles non bornés

Soit $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, une fonction positive. Nous notons $E_r = E \cap B(0,r)$. On dit qu'elle est **intégrable** sur $E \subset \mathbb{R}^n$ si

(1) Pour chaque $r > 0$, la fonction $f_r(x) = f(x)\mathbb{1}_{f<r}$ est intégrable sur E_r ;
(2) la limite $\lim_{r\to\infty} \int_{E_r} f_r$ est finie.

Dans ce cas, on pose

$$\int_E f = \lim_{r\to\infty} \int_{E_r} f_r. \tag{20.608}$$

ThoFnTestIntnnBorn

Théorème 20.196.
Soit E mesurable dans $\mathbb{R}^n$ et $f\colon E \to \overline{\mathbb{R}}$. Si f est mesurable et si il existe $g\colon E \to \overline{\mathbb{R}}$ intégrable sur E telle que $|f(x)| \leqslant g(x)$ pour tout $x \in E$, alors f est intégrable sur E.

Réciproquement, si f est intégrable sur E, alors f est mesurable.

LemTHBSEs

Lemme 20.197.
Si f est une fonction sur $[a,\infty[$, alors nous avons la formule

$$\lim_{b\to\infty}\int_a^b f(x)dx = \int_a^\infty f(x)dx \tag{20.609}$$

au sens où si un des deux membres existe, alors l'autre existe et est égal.

Démonstration. Supposons que le membre de gauche existe. Cela signifie que la fonction

$$\psi(x) = \int_a^x f \tag{20.610}$$

est bornée. Soit M, un majorant. Pour toute fonction simple φ dominée f, on a $\int\varphi \leqslant M$, donc l'ensemble sur lequel on prend le supremum pour calculer $\int_a^\infty f$ est majoré par M et possède donc un supremum. Nous avons donc

$$\int_a^\infty f \leqslant \lim_{b\to\infty}\int_a^b f. \tag{20.611}$$

□

20.26.12 Lemme de Morse

LemNQAmCLo

Lemme 20.198 (Lemme de Morse)**.**
Soit $f\in C^3(\mathcal{U},\mathbb{R})$ où $\mathcal{U}$ est un ouvert de $\mathbb{R}^n$ contenant 0. Nous supposons que

(1) $df_0 = 0$

(2) La matrice $H_{kl} = (\partial_{kl}f)(0)$ est non dégénérée et de signature $(p, n-p)$.

Alors il existe un C^1-difféomorphisme φ entre deux voisinages de 0 dans $\mathbb{R}^n$ tel que

(1) $\varphi(0) = 0$,

(2) si $\varphi(x) = u$ alors

$$f(x) - f(0) = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2. \tag{20.612}$$

Une autre façon de dire est qu'il existe un C^1-difféomorphisme local ψ tel que

$$(f\circ\psi)(x) - f(0) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \ldots - x_n^2. \tag{20.613}$$

Démonstration. Nous utilisons le théorème de Taylor avec reste intégral sous la forme de la proposition 20.150 avec $a = 0$:

$$f(x) = x\cdot(\nabla f)(0) + x\cdot Q(x)x \tag{20.614}$$

avec

$$Q(x)_{kl} = \int_0^1 (1-t)(\partial_{kl}f)(tx)dt. \tag{20.615}$$

Par hypothèse sur $df_0 = 0$, le premier terme est nul. En ce qui concerne Q, nous avons

$$Q(0)_{kl} = \int_0^1 (1-t)(\partial_{kl}f)(0) = \frac{1}{2}H_{kl}. \tag{20.616}$$

Étant donné que f est de classe C^2, le théorème de Schwaz 12.343 dit que H est une matrice symétrique. Elle est également inversible par hypothèse.

À partir de là, le lemme 17.118 donne un voisinage V de $Q(0)$ dans l'ensemble S_n des matrices symétriques et une application ϕ de classe C^1

$$\phi\colon V\to \mathrm{GL}(n,\mathbb{R}) \tag{20.617}$$

telle que pour tout $A \in V$,

$$\phi(A)^t Q(0)\phi(A) = A. \tag{20.618}$$

Si on pose $M = \phi \circ Q$, et si x est dans un voisinage de zéro, Q étant continue nous avons $Q(x) \in V$ et donc

$$Q(x) = M(x)^t Q(0) M(x). \tag{20.619}$$

Notons que l'application $M\colon \mathbb{R} \to \mathrm{GL}(n,\mathbb{R})$ est de classe C^1 parce que Q et ϕ le sont.

Nous avons

$$\begin{aligned}
f(x) - f(0) &= x \cdot Q(x)x && && \text{(20.620a)}\\
&= x \cdot M(x)^t Q(0) M(x)x && && \text{(20.620b)}\\
&= M(x)x \cdot Q(0)M(x)x && \text{cf. justif} && \text{(20.620c)}\\
&= y(x) \cdot Q(0)y(x) && \text{cf. justif} && \text{(20.620d)}\\
& && && \text{(20.620e)}
\end{aligned}$$

SUBEQooPDUKooHDpkHu

SUBEQooEJGYooRHWmzV

Justifications.

— Pour 20.620c. Voir la proposition 8.190 et le blabla 8.191.

— Pour 20.620d. Nous notons $y(x) = M(x)x = (\phi \circ Q)(x)x$, qui est encore une fonction de classe C^1 parce que la multiplication est une application C^∞.

D'un autre côté le théorème de Sylvester 8.251 nous donne une matrice inversible P telle que

$$Q(0) = P^t \begin{pmatrix} \mathbb{1}_p & \\ & -\mathbb{1}_{n-p} \end{pmatrix} P. \tag{20.621}$$

Et nous posons enfin $u = \varphi(x) = Py(x)$ qui est toujours de classe C^1 et qui donne

$$\begin{aligned}
f(x) - f(0) &= y \cdot Q(0)y && \text{(20.622a)}\\
&= y \cdot P^t \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} Py && \text{(20.622b)}\\
&= u \cdot \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} u && \text{(20.622c)}\\
&= u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2. && \text{(20.622d)}
\end{aligned}$$

Nous devons maintenant montrer que, quitte à réduire son domaine à un ouvert plus petit, φ est un C^1-difféomorphisme. Dans la chaine qui donne φ, seule l'application

$$\begin{aligned} g\colon U \subset \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto M(x)x \end{aligned} \tag{20.623}$$

est sujette à caution. Nous allons appliquer le théorème d'inversion locale. Nous savons que g est de classe C^1 et donc différentiable ; calculons la différentielle en utilisant la formule (12.647) :

$$dg_0(x) = \frac{d}{dt}\Big[g(tx)\Big]_{t=0} = \frac{d}{dt}\Big[tM(tx)x\Big]_{t=0} = M(0)x. \tag{20.624}$$

Note que nous avons utilisé la règle de Leibniz pour la dérivée d'un produit, mais le second terme s'est annulé. Donc $dg_0 = M(0) \in \mathrm{GL}(n,\mathbb{R})$ et g est localement un C^1-difféomorphisme.

Il suffit de restreindre φ au domaine sur lequel g est un C^1-difféomorphisme pour que φ devienne lui-même un C^1-difféomorphisme.

□

Définition 20.199.
*Un point a est un **point critique** de la fonction différentiable f si $df_a = 0$.*

Corolaire 20.200 ([598]).
Les points critiques non dégénérés d'une fonction C^3 sont isolés.

Démonstration. Soit a un point critique non dégénéré. Par le lemme de Morse 20.198, il existe un C^1-difféomorphisme ψ et un entier p tel que

$$(f\circ\psi)(x)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\ldots-x_n^2+f(a) \tag{20.625}$$

sur un voisinage $\mathcal{U}$ de a. Vue la formule générale $df_x(u)=\nabla f(x)\cdot u$, si x est un point critique de f, alors $\nabla f(x)=0$. Dans notre cas, les points critiques de $f\circ\psi$ dans $\mathcal{U}$ doivent vérifier $x_i=0$ pour tout i, et donc $x=a$.

Nous devons nous assurer que la fonction f elle-même n'a pas de points critiques dans $\mathcal{U}$. Pour cela nous utilisons la formule générale de dérivation de fonction composée :

$$\nabla(f\circ\psi)(x)=\sum_k\frac{\partial f}{\partial y_k}\big(g(x)\big)\nabla g_k(x). \tag{20.626}$$

Si $\psi(x)$ est un point critique de f, alors le membre de droite est le vecteur nul parce que tous les $\partial_k f\big(\psi(x)\big)$ sont nuls. Par conséquent le membre de gauche est également nul, et x est un point critique de $f\circ\psi$. Or nous venons de voir que $f\circ\psi$ n'a pas de points critiques dans $\mathcal{U}$.

Donc f n'a pas de points critiques dans un voisinage d'un point critique non dégénéré. □

20.27 Autres intégrales sympathiques

20.27.1 Intégrale de Wallis

LEMooMGUVooIIQSmC

Lemme 20.201 ([599, 1]).
Nous définissons les fonctions I_n par récurrence de la façon suivante :

$$\begin{cases} I_0(x)=x & \text{(20.627a)}\\ I_1(x)=\cos(x) & \text{(20.627b)}\\ I_n(x)=\dfrac{\cos(x)\sin^{n-1}(x)}{n}-\dfrac{n-1}{n}I_{n-2}(x). & \text{(20.627c)}\end{cases}$$

Pour chaque n, la fonction I_n est une primitive de $\sin^n(x)$.

EQooWJMIooSgBbJx

De même en définissant

$$\begin{cases} I_0(x)=x & \text{(20.628a)}\\ I_1(x)=\sin(x) & \text{(20.628b)}\\ I_n(x)=\dfrac{\cos^{n-1}(x)\sin(x)}{n}+\dfrac{n-1}{n}I_{n-2}(x), & \text{(20.628c)}\end{cases}$$

nous avons des primitives de $\cos^n(x)$.

Démonstration. Nous faisons par récurrence. Les cas $n=0$ et $n=1$ sont immédiats. Pour I_n c'est pas trop compliqué non plus : il suffit de dériver en supposant que $I'_{n-2}(x)=\sin^{n-2}(x)$. Voici le calcul :

$$\begin{aligned}
I'_n(x)&=\frac{1}{n}\big(-\sin^n(x)+\cos(x)(n-1)\sin^{n-2}(x)\cos(x)\big)-\frac{n-1}{n}\sin^{n-2}(x) && \text{(20.629a)}\\
&=\frac{1}{n}\sin^{n-2}(x)\big(-\sin^2(x)+(n-1)\cos^2(x)\big)-\frac{n-1}{n}\sin^{n-2}(x) && \text{(20.629b)}\\
&=\frac{1}{n}\sin^{n-2}(x)\big(-\sin^2(x)+(n-1)\cos^2(x)-n+1\big) && \text{(20.629c)}\\
&=\frac{1}{n}\sin^{n-2}(x)\big(n\cos^2(x)-n\big) && \text{(20.629d)}\\
&=\frac{1}{n}\sin^{n-2}(x)n\sin^2(x) && \text{(20.629e)}\\
&=\sin^n(x). && \text{(20.629f)}
\end{aligned}$$

Nous avons prouvé que la dérivée de $I_n(x)$ est bien $\sin^n(x)$.

La formule (20.628) se démontre de la même façon. □

LEMooUOIBooLyMDft

Lemme 20.202 ([600, 601])**.**
Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(t)dt. \tag{20.630}$$

Alors :

(1) une formule de récurrence :

$$W_n = \frac{n-1}{n} W_{n-2}, \tag{20.631}$$

EQooILMZooBUgJpk

(2) et une formule un peu explicite :

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2}. \tag{20.632}$$

EQooUYIDooEpHCnP

Démonstration. Le nombre W_n est seulement la seconde intégrale du lemme 20.201, évaluée entre 0 et $\pi/2$. En partant donc de (20.628), nous avons

$$W_n = \int_0^{\pi/2} \cos^n(x)dx = \left[\frac{\cos^{n-1}(x)\sin(x)}{n} \right]_0^{\pi/2} + \frac{n-1}{n} \int_0^{\pi/2} \cos^{n-2}(x)dx. \tag{20.633}$$

Le terme aux bords disparaît grâce aux valeurs trigonométriques remarquables [36]. Il reste immédiatement

$$W_n = \frac{n-1}{n} W_{n-2}. \tag{20.634}$$

À partir de là, nous démontrons (20.632) par récurrence. D'abord pour $n = 0$, c'est l'égalité $W_0 = \pi/2$ qui est correcte parce que

$$\int_0^{\pi/2} \cos^0(x)dx = \int_0^{\pi/2} 1dx. \tag{20.635}$$

Pour la récurrence elle-même,

$$\begin{aligned} W_{2(n+1)} &= \frac{2n+1}{2(n+1)} W_{2n} && \text{(20.636a)} \\ &= \frac{2n+1}{2n+2} \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2} && \text{(20.636b)} \\ &\quad \ldots\text{pas mal de petits calculs } \ldots && \text{(20.636c)} \\ &= \frac{(2n+2)!}{\big((n+1)!2^{n+1}\big)^2} \frac{\pi}{2}. && \text{(20.636d)} \end{aligned}$$

Voilà. □

Maintenant, la suite (W_n) se divise en ses termes pairs et ses termes impairs. Pour les pairs, nous avons une formule assez explicite donnée par le lemme 20.202. Pour les termes impairs, nous n'avons rien. Dans tous les cas, nous avons la formule de récurrence

$$W_n = \frac{n-1}{n} W_{n-2} \tag{20.637}$$

qui ne sert à rien pour déduire des choses sur les termes impairs à partir de ce que l'on sait des termes pairs.

Sommes-nous perdus ? Non. La situation se débloque grâce au lemme suivant.

36. Par exemple la proposition 18.20(10).

LEMooZFBVooQsOuOx

Lemme 20.203.
La suite donnée par

$$W_n = \int_0^{\pi/2} \cos^n(x)dx \tag{20.638}$$

est décroissante.

Démonstration. Vu que l'intégrale est sur $[0, \pi/2]$, le nombre $\cos(x)$ prend ses valeurs dans $[0, 1]$. Nous avons donc

$$\cos^{n+1}(x) \leqslant \cos^n(x). \tag{20.639}$$

Les intégrales suivent les mêmes inégalités. □

Ce lemme permet de relancer le jeu parce que les termes impairs sont coincés entre les termes pairs, qui décroissent. Les termes impairs doivent donc décroître à la même vitesse. Le lemme suivant met cela en musique.

LEMooAXTEooLBXQuM

Lemme 20.204.
Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(x)dx \tag{20.640}$$

La fonction $\alpha\colon \mathbb{N} \to \mathbb{R}$ *définie par* $W_n = \alpha(n)W_{n-1}$ *vérifie* $\lim_{n\to\infty} \alpha(n) = 1$.

Démonstration. Parce que nous en aurons besoin, nous triturons d'abord un peu la formule de récurrence (20.631). D'abord nous l'inversons un peu pour avoir

$$W_n = \frac{n+2}{n} W_{n+2}, \tag{20.641}$$

et ensuite nous écrivons

$$W_{n-1} = \frac{n+1}{n} W_{n+1}. \tag{20.642}$$

EQooYINNooBUKUYB

Ne vous posez pas de questions, ça va être utile. Le fait que la suite soit décroissante (lemme 20.203) nous permet d'écrire

$$W_{n+1} \leqslant W_n \leqslant W_{n-1}. \tag{20.643}$$

En y remplaçant W_n par $\alpha(n)W_{n+1}$ et W_{n-1} par (20.642),

$$W_{n+1} \leqslant \alpha(n)W_{n+1} \leqslant \frac{n+1}{n} W_{n+1}. \tag{20.644}$$

Nous simplifions par W_{n+1} et nous trouvons l'encadrement, valable pour tout n :

$$1 \leqslant \alpha(n) \leqslant \frac{n+1}{n}. \tag{20.645}$$

Nous en déduisons par la règle de l'étau que $\alpha(n) \to 1$. □

LEMooWQZAooOXAPQO

Lemme 20.205.
Soit

$$W_n = \int_0^{\pi/2} \cos^n(x)dx. \tag{20.646}$$

Encore plusieurs choses à dire.

(1) Pour tout n nous avons

$$nW_nW_{n-1} = \frac{\pi}{2}. \tag{20.647}$$

EQooLOLFooMIwMXN

(2) Nous avons l'équivalence de suites[37]

$$W_n \sim \sqrt{\frac{\pi}{2n}}. \tag{20.648}$$

Démonstration. En deux parties

(1) **La suite constante** Nous posons $K_n = nW_nW_{n-1}$. Grâce aux formules de récurrence (20.631) que nous écrivons sous la forme

$$W_n = \frac{n-1}{n}W_{n-2} \tag{20.649a}$$
$$W_{n+1} = \frac{n}{n+1}W_{n-1}, \tag{20.649b}$$

nous avons

$$K_{n+1} = (n+1)W_{n+1}W_n = (n+1)\frac{n}{n+1}\frac{n-1}{n}W_{n-1}W_{n-2} = (n-1)W_{n-1}W_{n-2} = K_{n-1}. \tag{20.650}$$

Nous avons montré que $K_{n+1} = K_{n-1}$.
Il nous reste à prouver que $K_1 = K_2 = \frac{\pi}{2}$. Pour cela nous avons immédiatement $W_0 = \frac{\pi}{2}$ ainsi que

$$W_1 = \int_0^{\pi/2} \cos(t)dt = 1. \tag{20.651}$$

Pour W_2, il ne faut pas calculer d'intégrales, mais seulement utiliser la formule (20.632). Nous trouvons vite $W_2 = \frac{\pi}{4}$. Donc

$$K_1 = W_1W_0 = \frac{\pi}{2} \tag{20.652}$$

et

$$K_2 = 2W_2W_1 = \frac{\pi}{2}. \tag{20.653}$$

(2) **L'équivalence de suites** Soit la fonction α définie par $W_n = \alpha(n)W_{n-1}$. Le lemme 20.204 nous dit que $\alpha(n) \to 1$. Nous l'utilisons dans (20.647) :

$$\frac{\pi}{2} = nW_nW_{n-1} = n\alpha(n)W_{n-1}^2. \tag{20.654}$$

Donc

$$W_{n-1}^2 = \frac{\pi}{2n\alpha(n)}, \tag{20.655}$$

et

$$W_{n-1}\sqrt{\frac{1}{\alpha(n)}}\sqrt{\frac{\pi}{2n}}. \tag{20.656}$$

Vu que le coefficient $\sqrt{1/\alpha(n)}$ tend vers 1 pour $n \to 1$, nous avons l'équivalence demandée.

□

20.27.2 Formule de Stirling

LEMooDXJOooOGFcIv

Lemme 20.206 ([602]).
Si $n \in \mathbb{N}$ nous avons la formule

$$\ln(1+\frac{1}{n}) = 2\sum_{k=0}^{\infty}\frac{1}{2k+1}\left(\frac{1}{2n+1}\right)^{2k+1} \tag{20.657}$$

37. Définition 9.32.

Démonstration. Il s'agit d'utiliser astucieusement le développement de la proposition 15.102. Nous avons d'une part

$$\ln(1+x)=\sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \tag{20.658}$$

et d'autre part,

$$-\ln(1-x)=\sum_{k=1}^{\infty}\frac{x^k}{k}. \tag{20.659}$$

Cela permet de calculer, en utilisant l'associativité de la série [38] EQooYVXHooNDetVx

$$\ln\left(\frac{1+t}{1-t}\right)=\ln(1+t)-\ln(1-t) \tag{20.660a}$$

$$=\sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}x^k}{k}+\frac{x^k}{k}\right) \tag{20.660b}$$

$$=2\sum_{k=0}^{\infty}\frac{x^{2k+1}}{2k+1} \tag{20.660c}$$ SUBEQooIIOBooXYoHkK

parce que tous les termes pairs s'annulent, tandis que les termes impairs sont doublés. Notez que la somme dans (20.660c) commence à $k=0$, contrairement aux autres qui commencent à 1.

Et là c'est l'astuce : on écrit l'égalité (20.660) avec le t qu'il faut pour que

$$\frac{1+t}{1-t}=1+\frac{1}{n}. \tag{20.661}$$

Nous posons donc $t=\frac{1}{2n+1}$ nous avons

$$\ln\left(1+\frac{1}{n}\right)=2\sum_{k=0}^{\infty}\frac{1}{2k+1}\left(\frac{1}{2n+1}\right)^{2k+1}, \tag{20.662}$$

ce qu'il fallait. □

LemCEoBqrP

Lemme 20.207 (Formule de Stirling[603, 604, 600, 602, 605])**.**
Nous avons l'équivalence de suites [39]

$$n!\sim\left(\frac{n}{e}\right)^n\sqrt{2\pi n}. \tag{20.663}$$

Démonstration. Nous posons

$$a_n=\frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n} \tag{20.664}$$

et $b_n=\ln(a_n)$.

(1) **Une formule pour b_n-b_{n+1}** Nous faisons un beau calcul qui utilise les formules de la proposition 15.86 ainsi que $\ln(e)=1$: SUBEQSooZTABooBukFGM

$$b_n-b_{n+1}=\ln\left(\frac{a_n}{a_{n+1}}\right) \tag{20.665a}$$

$$=\ln\left(\frac{(n+1)^{n+1/2}}{n^{n+1/2}}\frac{1}{e}\right) \tag{20.665b}$$

$$=\ln\left((\frac{n+1}{n})^{n+1/2}\right)-1 \tag{20.665c}$$

$$=\left(n+\frac{1}{2}\right)\ln(1+\frac{1}{n})-1. \tag{20.665d}$$

38. Proposition 10.102.
39. Définition 9.32.

(2) **La suite (b_n) est décroissante** Nous écrivons l'égalité (20.665) en utilisant le lemme 20.206 :

$$b_n - b_{n+1} = (n + \frac{1}{2}) \ln(1 + \frac{1}{n}) - 1 \tag{20.666a}$$
$$= \frac{1}{2}(2n+1)2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k+1} - 1 \tag{20.666b}$$
$$= (2n+1) \sum_{k=1}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k+1} \tag{20.666c}$$
$$= \sum_{k=1}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k}. \tag{20.666d}$$

Notez que le terme $k = 0$ s'est simplifié avec le -1. Vu que le tout est une somme de termes positifs, nous avons

$$b_n - b_{n+1} > 0 \tag{20.667}$$

et la suite est décroissante.

(3) **Majoration pour $b_n - b_{n+1}$** Vu que $\frac{1}{2k+1} < 1$, nous pouvons majorer :

$$b_n - b_{n+1} \leqslant \sum_{k=1}^{\infty} \left(\frac{1}{(2n+1)^2} \right)^k. \tag{20.668}$$

Nous remarquons que cela est une série géométrique déjà traitée dans la proposition 10.135. Nous faisons un peu de calcul en partant de

$$\sum_{k=1}^{\infty} q^n = \frac{q}{1-q} \tag{20.669}$$

avec $q = \frac{1}{(2n+1)^2}$. Après quelques simplifications,

$$b_n - b_{n+1} \leqslant \frac{1}{4} \frac{1}{n(n+1)} \tag{20.670}$$

(4) **Le coup de la somme télescopique** Nous avons

$$b_1 - b_n = (b_1 - b_2) + (b_2 - b_3) + \ldots + (b_{n-1} - b_n). \tag{20.671}$$

Chacun de ces termes est majoré ; nous avons donc

$$b_1 - b_n < \frac{1}{4} \sum_{m=1}^{n-1} \frac{1}{m(m+1)} \leqslant \frac{1}{4} \sum_{m=1}^{\infty} \frac{1}{m(m+1)} = \frac{1}{4} \tag{20.672}$$

grâce au lemme 10.139 pour la dernière somme.

(5) **La suite b_n est bornée vers le bas**
Vu que $b_1 = 1 - \frac{\ln(2)}{2}$, nous avons

$$b_n > b_1 - \frac{1}{4} = \frac{3}{4} - \frac{\ln(2)}{2}. \tag{20.673}$$

En utilisant la majoration de l'exemple 20.152 nous trouvons

$$0.327 \leqslant b_n \leqslant 0.427. \tag{20.674}$$

Cet encadrement n'est pas très important. Le point est que la suite (b_n) soit bornée vers le bas ; savoir que la borne est strictement positive n'est pas indispensable.

(6) **Une limite pour (a_n)** La suite (b_n) est décroissante et bornée vers le bas, donc elle est convergente par le lemme 9.35. Vu que l'exponentielle est une fonction continue [40], la suite $a_n = e^{b_n}$ est également convergente.
Vu que $\lim_{n\to\infty} b_n > 0$, nous avons $\lim_{n\to\infty} a_n = e^{\lim_{n\to\infty} b_n} > 1$.
L'important est que nous sachions que (a_n) est une suite convergente. Nous notons L sa limite :

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n} = L. \tag{20.675}$$

Ou encore :

$$\lim_{n\to\infty} \frac{n!}{L\sqrt{2n}\left(\frac{n}{e}\right)^n} = 1. \tag{20.676}$$

Autrement dit, en posant

$$n! = \alpha(n)L\sqrt{2n}\left(\frac{n}{e}\right)^n, \tag{20.677}$$

EQyiBooFVCCooIbGpsW

nous avons $\lim_{n\to\infty} \alpha(n) = 1$.

(7) **Introduction de Wallis** Nous avons déjà parlé dans le lemme 20.202 du nombre

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2}\frac{\pi}{2}. \tag{20.678}$$

Le lemme 20.205 implique que la suite $I_n = W_{2n}$ est équivalente à $\sqrt{\pi/4n}$.
Cela étant dit, nous faisons un gros calcul en remplaçant les factorielles dans W_{2n} par la formule (20.677). Après pas mal de calculs [41], nous trouvons

$$I_n = \frac{\alpha(2n)}{\alpha(n)^2}\frac{1}{\sqrt{n}}\frac{\pi}{2n}. \tag{20.679}$$

Nous avons donc

$$\sqrt{\frac{\pi}{4n}} \sim \frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{\sqrt{n}}\frac{1}{2L}. \tag{20.680}$$

Il existe donc une fonction $\beta(n)$ avec $\beta(n) \to 1$ telle que

$$\frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{n}} = \beta(n)\frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{\sqrt{n}}\frac{1}{2L}. \tag{20.681}$$

En simplifiant par $1/\sqrt{n}$,

$$\frac{\sqrt{\pi}}{2} = \beta(n)\frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{2L}. \tag{20.682}$$

Nous prenons à présent la limite $n \to \infty$ en nous rappelant que α et β donnent 1. Après simplifications, nous trouvons

$$L = \sqrt{\pi}. \tag{20.683}$$

(8) **La fin** Nous introduisons la valeur $L = \sqrt{\pi}$ dans l'expression (20.677) de la factorielle :

$$n! = \alpha(n)\sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{20.684}$$

La dernière équation est exactement ce qui signifie l'équivalence de suite demandée. □

20.208.
J'ai déjà dit du mal de la formule d'Euler $e^{i\pi} = -1$ dans (18.12). La formule de Stirling est plus intéressante parce qu'elle lie (réellement) les nombres e et π. Mais elle fait aussi intervenir le nombre $\sqrt{2}$, et la fonction factorielle.

La formule de Stirling est également nettement moins triviale à démontrer : la formule d'Euler ne fait intervenir de non trivial que les définition et propriétés des fonctions trigonométriques. Toutes ces propriétés sont également utilisées pour prouver la formule de Stirling.

40. Parce qu'elle est dérivable, voir par exemple le théorème 15.76.
41. Si vous êtes en manque de papier de brouillon, c'est le moment de vous inquiéter.

20.27.3 La fonction sinus cardinal, intégrale de Dirichlet

Définition 20.209.
La fonction ***sinus cardinal*** *est*

$$f(t) = \begin{cases} 1 & si\ t = 0 \\ \frac{\sin(t)}{t} & sinon. \end{cases} \tag{20.685}$$

Elle sert à plein de choses. Entre autres, le lemme 20.212 montrera que la fonction $x \mapsto |\sin(x)/x|$ a une intégrale sur $\mathbb{R}$ qui vaut ∞. Cela nous permettra de donner un exemple d'une fonction dans $L^1(\mathbb{R})$ dont la transformée de Fourier n'est pas dans $L^1(\mathbb{R})$ (lemme 29.8).

20.210.
Le but que nous nous fixons maintenant est de prouver que

$$\int_0^\infty \frac{\sin(t)}{t} dt = \frac{\pi}{2}. \tag{20.686}$$

Un adage dit que si un théorème est trop long, c'est qu'il n'a pas assez de lemmes. Nous allons faire plein de lemmes.

LEMooMJFBooAjtNjV

Lemme 20.211.
La fonction sinus cardinal est continue.

Démonstration. Elle est continue en zéro parce que le lemme 18.247 nous donne

$$\lim_{t\to 0} \frac{\sin(t)}{t} = 1. \tag{20.687}$$

□

Nous commençons par une mauvaise nouvelle.

LEMooEEWSooZwLSAP

Lemme 20.212 ([1])**.**
Nous avons

$$\int_{\mathbb{R}} |\frac{\sin(x)}{x}| dx = \infty. \tag{20.688}$$

Démonstration. Soit $\delta > 0$ tel que sur l'intervalle $[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta]$, nous ayons $\sin(x) > 0.9$ [42] .

Les intervalles $I_k = [\frac{\pi}{2} - \delta + 2k\pi, \frac{\pi}{2} + \delta + 2k\pi]$ sont disjoints et la fonction que nous intégrons est partout positive. Nous découpons

$$\mathbb{R} = C + \bigcup_{k=0}^{\infty} [\frac{\pi}{2} - \delta + 2k\pi, \frac{\pi}{2} + \delta + 2k\pi] \tag{20.689}$$

où C est le complémentaire qu'il faut pour faire $\mathbb{R}$.

La σ-additivité de l'intégrale de Lebesgue (proposition 14.211) nous indique que

$$\int_{\mathbb{R}} |\frac{\sin(x)}{x}| = \int_C |\frac{\sin(x)}{x}| + \sum_{k=0}^{\infty} \int_{I_k} |\frac{\sin(x)}{x}| dx \tag{20.690}$$

Vu que tous les termes sont positifs, nous obtenons une majoration en en supprimant un. Allons-y :

42. Ça existe par une astucieuse combinaison du théorème 9.93 des valeurs intermédiaires, de la valeur remarquable $\sin(\pi/2) = 1$ (de (18.41)) et du fait que sin est continue (proposition 18.1).

SUBEQSooSRAYooEOBwiC

$$\int_{\mathbb{R}} |\frac{\sin(x)}{x}|dx \geqslant \sum_{k=0}^{\infty} \int_{I_k} |\frac{\sin(x)}{x}|dx \tag{20.691a}$$
$$\geqslant \sum_{k=0}^{\infty} 0.9 \int_{I_k} \frac{1}{x}dx \tag{20.691b}$$
$$\geqslant 0.9 \sum_{k=0}^{\infty} 2\delta \frac{1}{\frac{\pi}{2}+\delta+2k\pi} \tag{20.691c}$$
$$\geqslant 1.8\delta \sum_{k=0}^{\infty} \frac{1}{2\pi(k+1)} \tag{20.691d}$$
$$= \frac{1.8\delta}{2\pi} \sum_{k=0}^{\infty} \frac{1}{k+1}. \tag{20.691e}$$

Justifications :

— Pour (20.691c), nous avons majoré $\frac{1}{x}$ par $\frac{1}{\frac{\pi}{2}+\delta+2k\pi}$ sur I_k.

— Pour (20.691d), nous avons dit que $\frac{\pi}{2}+\delta < 2\pi$.

La dernière somme dans (20.691) diverge.

Donc la fonction sinus cardinal n'est pas dans $L^1(\mathbb{R})$. □

La mauvaise nouvelle suivante en est un corolaire immédiat.

LEMooBEQRooHaugKj

Lemme 20.213.
La fonction $t \mapsto \frac{\sin(t)}{t}$ n'est pas intégrable sur $[0,\infty[$ au sens de Lebesgue.

Démonstration. Le lemme 20.212 nous dit que $\int_0^\infty |f| = \infty$. Dans ce cas, $\int_0^\infty f$ n'existe pas par le lemme 14.188. □

Donc l'intégrale $\int_0^\infty \frac{\sin(t)}{t}dt$ n'existe pas parce que la définition de l'intégrale de Lebesgue ne permet pas de profiter des compensations qui arrivent entre les valeurs positives et négatives.

Nous définissons donc

$$\int_0^\infty \frac{\sin(t)}{t}dt = \lim_{b\to\infty} \int_0^b \frac{\sin(t)}{t}dt. \tag{20.692}$$

EQooWAQLooTFQPbl

Pour chaque b, l'intégrale existe sans problèmes (fonction continue sur le compact $[0,b]$), et les compensations se font. Il n'est pas pas sans espoir que la limite (20.692) existe et vaille un nombre fini.

LEMooTFVZooRAmjUN

Lemme 20.214 ([606])**.**
La limite

$$\lim_{b\to\infty} \int_0^b \frac{\sin(t)}{t}dt \tag{20.693}$$

existe dans $\mathbb{R}$.

Démonstration. En plusieurs parties.

(1) **Découpage**
Nous découpons l'intervalle $[0,b]$ en morceaux du type $[k\pi,(k+1)\pi]$ et un morceau restant lorsque b n'est pas un multiple de π :

$$[0,b] = \bigcup_{k=0}^{N(b)-1} [k\pi,(k+1)\pi] \cup [N(b)\pi,b] \tag{20.694}$$

où $N(b)$ est un entier bien choisi[43]. En tout cas $\lim_{b\to\infty} N(b) = \infty$.

43. Il me semble que le traitement de ce terme manque dans [606].

(2) **Majoration 1**
Pour chaque $b \in \mathbb{R}^+$ nous avons

$$\int_0^b \frac{\sin(t)}{t}dt = \sum_{k=0}^{N(b)} \int_{[k\pi,(k+1)\pi]} \frac{\sin(t)}{t}dt + \int_{[(N(b)+1)\pi,b]} \frac{\sin(t)}{t} \tag{20.695}$$

Dans le dernier terme, nous majorons $|\sin(t)| \leqslant 1$ et $\frac{1}{t} \leqslant \frac{1}{(N(b)+1)\pi}$. Cela donne

$$\left| \int_{[(N(b)+1)\pi,b]} \frac{\sin(t)}{t} \right| \leqslant \frac{b-(N(b)+1)\pi}{(N(b)+1)\pi} \leqslant \frac{1}{N(b)+1} \tag{20.696}$$

où nous avons encore majoré $b \leqslant (N(b)+2)\pi$.

(3) **Majoration 2**
En ce qui concerne les autres termes, sur l'intervalle $[k\pi,(k+1)\pi]$, nous avons $\sin(t) = (-1)^k|\sin(t)|$. Nous avons alors

$$\int_0^b \frac{\sin(t)}{t}dt = \sum_{k=0}^{N(b)} (-1)^k \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t}dt + \alpha(b) \tag{20.697}$$

EQooHZPRooFuwRWQ

où $|\alpha(b)| \leqslant \frac{1}{N(b)+1}$; l'important est que $\lim_{b\to\infty} \alpha(b) = 0$.

(4) **Une suite alternée**
L'inégalité (20.697) nous incite à étudier la série $\sum_{k=0}^{\infty}(-1)^k a_k$ en ayant posé

$$a_k = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t}dt. \tag{20.698}$$

Nous montrons à présent que la suite (a_k) vérifie les conditions du critère des séries alternées 10.140.
D'abord, $a_{k+1} \leqslant a_k$. En effet en utilisant le changement de variables [44] $u = t - \pi$,

$$a_{k+1} = \int_{(k+1)\pi}^{(k+2)\pi} \frac{|\sin(t)|}{t}dt = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(u+\pi)|}{u+\pi}du = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(u)|}{u+\pi}du < a_k. \tag{20.699}$$

Nous avons utilisé le fait que $|\sin(u+\pi)| = |\sin(u)|$ pour tout u.
De plus, vu que $|\sin(t)| \leqslant 1$ et que $t \in [k\pi,(k+1)\pi]$, nous avons

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t} \leqslant \int_{k\pi}^{(k+1)\pi} \frac{1}{k\pi} = \frac{(k+1)\pi - k\pi}{k\pi} = \frac{1}{k}. \tag{20.700}$$

Donc $a_k \leqslant \frac{1}{k} \to 0$.
Le critère des séries alternées 10.140 nous dit que

$$\sum_{k=0}^{\infty}(-1)^k a_k < \infty. \tag{20.701}$$

(5) **Conclusion** Nous repartons de (20.697) :

$$\int_0^b \frac{\sin(t)}{t}dt = \sum_{k=0}^{N(b)} (-1)^k \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t}dt + \alpha(b). \tag{20.702}$$

Cette égalité est valable pour tout $b \in \mathbb{R}^+$. Le passage à la limite $b \to 0$ à droite donne un nombre fini ; donc à gauche aussi, et nous avons prouvé que

$$\lim_{b\to\infty} \int_0^b \frac{\sin(t)}{t}dt < \infty. \tag{20.703}$$

44. Le théorème 14.281 est toujours bon à citer.

□

Notre tache n'est donc pas sans espoir. Au moins l'intégrale que nous cherchons à évaluer est finie.

LEMooARPIooDPSGwR

Lemme 20.215.
Soit $x \in]0,\infty[$. L'intégrale

$$F(x) = \int_0^\infty e^{-tx} \frac{\sin(t)}{t} dt \tag{20.704}$$

existe au sens de Lebesgue usuel.

Démonstration. Vu qu'en $t = 0$ nous avons $\frac{\sin(t)}{t} = 1$, il n'y a pas de problèmes de ce côté. Lorsque $t > 1$ nous avons la majoration

$$|e^{-tx} \frac{\sin(t)}{t}| \leqslant |e^{-tx}|. \tag{20.705}$$

Lorsque t est assez grand, le lemme 15.107 nous donne aussi la majoration

$$|e^{-tx}| \leqslant \frac{1}{t^2}. \tag{20.706}$$

La proposition 14.300(2) implique que $\int_1^\infty \frac{1}{t^2} < \infty$. Et les majorations font que la proposition 14.176 nous donne le résultat. □

LEMooRDCSooBrWmep

Lemme 20.216 ([606])**.**
Il existe une constante $C \in \mathbb{R}$ telle que

$$I(x) = \int_0^\infty e^{-tx} \frac{\sin(t)}{t} dt = -\arctan(x) + C \tag{20.707}$$

pour tout $x > 0$.

Démonstration. En permutant dérivée et intégrale, nous allons prouver que $I'(x) = -\frac{1}{1+x^2}$.

(1) **Permuter** Nous posons

$$f(x,t) = e^{-tx} \frac{\sin(t)}{t}, \tag{20.708}$$

et nous vérifions les hypothèses du théorème 17.19.

(1a) Pour chaque $x > 0$ fixé, la fonction $t \mapsto f(x,t)$ est intégrable sur $[0,\infty[$, c'est le lemme 20.215.

(1b) Pour $t > 0$ fixé, la fonction $x \mapsto f(x,t)$ est dérivable.

(1c) Nous avons la dérivée partielle

$$\frac{\partial f}{\partial x}(x,t) = -e^{-tx} \sin(t) \tag{20.709}$$

qui vérifie

$$|\frac{\partial f}{\partial x}(x,t)| \leqslant e^{-tx}, \tag{20.710}$$

alors que la fonction $t \mapsto e^{-tx}$ est intégrable sur $[0,\infty[$.

Nous pouvons donc dériver sous l'intégrale et obtenir

$$F'(x) = -\int_0^\infty e^{-xt} \sin(t) dt. \tag{20.711}$$

(2) **Quelques intégrations par partie** Nous posons

$$J(x) = \int_0^\infty e^{-xt} \sin(t) dt. \tag{20.712}$$

Cette intégrale est une intégrale de Lebesgue toute simple. En effet, posons provisoirement $s(t) = e^{-xt}\sin(t)$. La partie positive de s vérifie $s_+(t) \leqslant e^{-xt}$ et la partie négative vérifie $s_-(t) > -e^{-xt}$. Les deux ont une intégrale qui converge. La définition 14.186 dit alors que J est une intégrale tout à fait propre. Elle serait d'ailleurs mieux écrite sous la forme

$$J(x) = \int_{[0,\infty]} e^{-xt} \sin(t) dt. \tag{20.713}$$

Bref. Tout ça pour dire que nous avons tout à fait le droit de la faire par parties [45], en deux fois.

D'abord en posant $u = e^{-xt}$ et $v' = \sin(t)$ nous avons

$$J(x) = \big[-e^{-xt}\cos(t)\big]_{t=0}^{t=\infty} - \int_0^\infty (-)xe^{-xt}(x)\cos(t)dt = 1 - x\int_0^\infty e^{-xt}\cos(t)dt. \tag{20.714}$$

Nous faisons l'intégrale encore par parties en posant $u = e^{-xt}$ et $v' = \cos(t)$:

$$\int_0^\infty e^{-xt}\cos(t)dt = \big[e^{-xt}\sin(t)\big]_0^\infty - \int_0^\infty (-x)e^{-xt}\sin(t)dt = xJ(x). \tag{20.715}$$

Donc

$$\begin{aligned} J(x) &= 1 - x\int_0^\infty e^{-xt}\cos(t)dt && \text{(20.716a)}\\ &= 1 - x\int_0^\infty e^{-xt}\cos(t)dt && \text{(20.716b)}\\ &= 1 - x\Big(\underbrace{x\int_0^\infty e^{-xt}\sin(t)dt}_{J(x)}\Big) && \text{(20.716c)}\\ &= 1 - x^2J(x). && \text{(20.716d)} \end{aligned}$$

Voilà qui prouve que $J(x) = \frac{1}{1+x^2}$, et donc que

$$F'(x) = -J(x) = -\frac{1}{1+x^2}. \tag{20.717}$$

(3) **Et enfin**

Le théorème 18.38(2) nous dit que la dérivée de la fonction arctan est précisément $1/(1+x)$. Donc F et arctan ont la même dérivée (au signe près). Donc il existe $C \in \mathbb{R}$ tel que

$$F(x) = -\arctan(x) + C \tag{20.718}$$

pour tout $x > 0$.

□

LEMooEOYHooVIMCCa

Lemme 20.217.
Nous avons

$$F(x) = \int_0^\infty e^{-xt}\frac{\sin(t)}{t}dt = -\arctan(x) + \frac{\pi}{2} \tag{20.719}$$

pour tout $x > 0$.

45. Proposition 20.133.

Démonstration. Le but de ce lemme est de fixer la constante laissée arbitraire dans le lemme 20.216. Nous savons qu'il existe $C \in \mathbb{R}$ tel que

$$F(x) = -\arctan(x) + C. \tag{20.720}$$ EQooTZGXooUxfAiT

Le but est de prendre la limite $x \to \infty$ des deux côtés.

Par le lemme 18.248, nous avons $|\frac{\sin(t)}{t}| \leqslant 1$ sur $[0, \infty[$. Donc

$$F(x) \leqslant \int_0^\infty e^{-xt} dt = \left[-\frac{1}{x} e^{-xt} \right]_{t=0}^{t=\infty} = \frac{1}{x}. \tag{20.721}$$

Vu que $F(x) \leqslant \frac{1}{x}$ pour tout x, nous avons certainement $\lim_{x\to\infty} F(x) = 0$.

D'autre part,

$$\lim_{x\to\infty} \arctan(x) = \frac{\pi}{2}. \tag{20.722}$$

En passant à la limite dans (20.720), nous avons

$$0 = -\frac{\pi}{2} + C, \tag{20.723}$$

et donc $C = \pi/2$. □

Nous avons maintenant la formule

$$F(x) = \int_0^\infty e^{-tx} \frac{\sin(t)}{t} dt = -\arctan(x) + \frac{\pi}{2} \tag{20.724}$$

qui est valable pour tout $x > 0$.

Notre but sera de prendre la limite $x \to 0$ des deux côtés. Vu que arctan est continue, le membre de droite ne pose pas de problèmes et donne $\pi/2$. Pour le membre de gauche, il faut encore permuter une limite et une intégrale.

Pour la suite, nous allons étudier[606]

$$\int_0^\infty (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \tag{20.725}$$

Cette intégrale n'existe pas au sens de Lebesgue et est définie par

$$L(x) = \lim_{b\to\infty} \int_0^b (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \tag{20.726}$$

Rien n'indique cependant pour l'instant que cette limite existe.

LEMooZGODooLaBuHo

Lemme 20.218 ([606])**.**
Soit $x > 0$. Nous posons

$$L_k(x) = \int_{k\pi}^{(k+1)\pi} (1 - e^{-tx}) \frac{|\sin(t)|}{t} dt. \tag{20.727}$$ EQooJXWMooRbbCtt

La suite $(L_k(x))_{k\in\mathbb{N}}$ satisfait le critère des séries alternées[46]*, c'est-à-dire que cette suite est positive, décroissante à limite nulle.*

Démonstration. Notons que chacune des intégrales $L_k(x)$ est sans problèmes : fonction continue sur un compact. Trois éléments à prouver.

(1) **Positive** Vu que dans toute notre histoire, $x, t > 0$, nous avons $1 - e^{-tx} > 0$ et donc toute la fonction intégrée est positive.

46. Théorème 10.140.

(2) **Tend vers zéro** Vu que $1-e^{-tx}<1$, nous avons

$$L_k(x)\leqslant\int_{k\pi}^{(k+1)\pi}\frac{1}{t}dt\leqslant\pi\frac{1}{k\pi}=\frac{1}{k}.\tag{20.728}$$

Donc $\lim_{k\to\infty}L_k(x)=0$.

(3) **Décroissante** Nous devons à présent prouver que $L_k(x)$ est décroissante en k lorsque x est fixé.

Nous avons $L_{k+1}(0)=L_k(0)$ pour tout k. Nous allons montrer que $L'_{k+1}(x)<L'_k(x)$ pour tout $x>0$. De cette façon nous aurons bien $L_{k+1}(x)<L_k(x)$ pour tout k et x.

En permutant (encore) intégrale et dérivée,

$$\begin{aligned}L'_{k+1}(x)&=\int_{(k+1)\pi}^{(k+2)\pi}e^{-tx}|\sin(t)|dt&(20.729a)\\&=\int_{k\pi}^{(k+1)\pi}e^{-(u+\pi)x}|\sin(u+\pi)|&(20.729b)\\&=e^{-\pi x}\int_{k\pi}^{(k+1)\pi}e^{-ux}|\sin(u)|du&(20.729c)\\&=e^{-\pi x}L'_k(x).&(20.729d)\end{aligned}$$

Justifications :

— Pour (20.729a), permuter dérivée et intégrale ; je ne donne pas tout le détail. Ça a déjà été fait.

— Pour (20.729b), nous avons fait le changement de variables $u=t-\pi$.

— Pour (20.729c), nous avons utilisé le fait que $|\sin(u+\pi)|=|\sin(u)|$ ainsi que $e^{-(u+\pi)x}=e^{-ux}e^{-\pi x}$ par (12.1096).

Nous avons donc prouvé que

$$L'_{k+1}(x)=e^{-\pi x}L'_k(x)<L'_k(x).\tag{20.730}$$

□

Lemme 20.219.
Pour chaque $x>0$, nous avons la limite

$$\lim_{b\to\infty}\int_0^b(1-e^{-tx})\frac{\sin(t)}{t}dt=\sum_{k=0}^{\infty}(-1)^kL_k(x)<\infty.\tag{20.731}$$

Démonstration. Nous fixons (provisoirement) b et nous découpons l'intervalle d'intégration comme

$$[0,b]=\bigcup_{k=1}^{N}[k\pi,(k+1)\pi]\cup\big[(N+1)\pi,b\big]\tag{20.732}$$

où N est une fonction de b ; quelque chose comme $N(b)$ est le plus grand entier tel que $\big(N(b)+1\big)\pi\leqslant b$. Sur chacun des intervalles nous avons $\sin(t)=(-1)^k|\sin(t)|$. Nous avons donc

$$\begin{aligned}\int_0^b(1-e^{-tx})\frac{\sin(t)}{t}dt=&\sum_{k=0}^{N(b)}(-1)^k\int_{k\pi}^{(k+1)\pi}(1-e^{-tx})\frac{|\sin(t)|}{t}dt\\&+\int_{(N+1)\pi}^{b}(1-e^{-tx})\frac{\sin(t)}{t}dt\end{aligned}\tag{20.733}$$

Le premier terme est $\sum_{k=0}^{N(b)}(-1)^k L_k(x)$, dont nous savons que la limite $b\to\infty$ existe parce que $L_k(x)$ vérifie le critère des séries alternées (lemme 20.218). En ce qui concerne le second terme,

$$\Big|\int_{(N+1)\pi}^{b}(1-e^{-tx})\frac{\sin(t)}{t}\Big|<\frac{b-(N+1)\pi}{(N+1)\pi}<\frac{1}{N(b)+1}. \tag{20.734}$$

La dernière inégalité est le fait que $N(b)$ est choisi pour avoir $b-\big(N(b)+1\big)\pi<\pi$.

Les deux termes de (20.733) ont donc une limite lorsque $b\to\infty$. Nous pouvons donc passer à la limite en sommant les deux limites :

$$\int_0^\infty(1-e^{-tx})\frac{\sin(t)}{t}dt=\sum_{k=0}^{\infty}(-1)^k L_k(x). \tag{20.735}$$

Cette égalité est valable pour chaque $x>0$.

Le fait que la limite soit finie est dans le critère des séries alternées. Pour chaque x, la suite $L_k(x)$ vérifie ce critère par le lemme 20.218. □

LEMooNZVSooDbZCZx

Lemme 20.220.
Nous avons

$$\lim_{x\to0^+}\int_0^\infty(1-e^{-tx})\frac{\sin(t)}{t}dt=0. \tag{20.736}$$

Démonstration. La définition de l'intégrale ainsi que le lemme 20.219 nous ont déjà donné

$$\int_0^\infty(1-e^{-tx})\frac{\sin(t)}{t}dt=\lim_{b\to\infty}\int_0^b(1-e^{-tx})\frac{\sin(t)}{t}dt=\sum_{k=0}^{\infty}(-1)^k L_k(x) \tag{20.737}$$

ainsi que l'assurance que le tout est un nombre réel fini [47].

(1) **Majoration pour la série alternée**
Nous majorons un peu. Pour $x>0$ et $N\in\mathbb{N}$ nous avons SUBEQSQooLIGNooNAzpmi

$$\begin{aligned}
\Big|\sum_{k=0}^{N}(-1)^k L_k(x)\Big|&\leqslant\sum_{k=0}^{N}|L_k(x)| && (20.738a)\\
&=\sum_{k=0}^{N}\Big|\int_{k\pi}^{(k+1)\pi}(1-e^{-tx})\frac{|\sin(t)|}{t}dt\Big| && (20.738b)\\
&\leqslant\sum_{k=0}^{N}\int_{k\pi}^{(N+1)\pi}(1-e^{-tx})\frac{|\sin(t)|}{t}dt && (20.738c)\\
&=\int_0^{(N+1)\pi}(1-e^{-tx})\frac{|\sin(t)|}{t}dt && (20.738d)\\
&\leqslant\int_0^{(N+1)\pi}tx\frac{|\sin(t)|}{t}dt && (20.738e)\\
&\leqslant\int_0^{(N+1)\pi}xdt && (20.738f)\\
&=x(N+1)\pi. && (20.738g)
\end{aligned}$$

EQooIIVVooLvHFbS
SUBEQooZPVVooTBZzdl

Justifications :

— Pour (20.738b) c'est la définition (20.727).

— Pour (20.738e), c'est le fait que $0\leqslant1-e^{-u}\leqslant u$ pour tout $u\geqslant0$ ainsi que la sous-additivité de l'intégrale de la proposition 14.194.

47. De toutes façons, il n'existe pas de nombres réels infinis, mais vous voyez ce que je veux dire.

(2) **Majoration pour l'intégrale**
Nous fixons $N \in \mathbb{N}$, et nous avons : SUBEQSooSBSJooMAkJPh

$$\Big|\lim_{b\to\infty}\int_0^b (1-e^{-tx})\frac{\sin(t)}{t}dt\Big| = \Big|\sum_{k=0}^{\infty}(-1)^k L_k(x)\Big| \tag{20.739a}$$
$$= \Big|\sum_{k=0}^{N}(-1)^k L_k(x) + \sum_{k=N+1}^{\infty}(-1)^k L_k(x)\Big| \tag{20.739b}$$
$$\leqslant \Big|\sum_{k=0}^{N}(-1)^k L_k(x)\Big| + \Big|\sum_{k=N+1}^{\infty}(-1)^k L_k(x)\Big| \tag{20.739c}$$
$$\leqslant \Big|\sum_{k=0}^{N}(-1)^k L_k(x)\Big| + L_{N+1}(x) \tag{20.739d}$$ SUBEQooYFWDooYKhYtd
$$\leqslant \Big|\sum_{k=0}^{N}(-1)^k L_k(x)\Big| + \frac{1}{N+1} \tag{20.739e}$$ SUBEQooDDBGooNDfxoQ
(20.739f)

Justifications :

— Pour (20.739d) c'est le reste du critère des séries alternées, théorème 10.140(3).

— Pour (20.739e) c'est la majoration (20.728) déjà faite.

(3) **Les deux ensemble** Pour chaque N et pour chaque $x > 0$ nous avons, en mettant (20.738) au bout de (20.739) :

$$\Big|\int_0^{\infty}(1-e^{-tx})\frac{\sin(t)}{t}dt\Big| \leqslant x(N+1)\pi + \frac{1}{N+1}. \tag{20.740}$$

En prenant la limite $x \to 0$ nous trouvons

$$\lim_{x\to 0}\int_0^{\infty}(1-e^{-tx})\frac{\sin(t)}{t}dt \leqslant \frac{1}{N+1} \tag{20.741}$$

pour tout N. Donc cette limite est nulle :

$$\lim_{x\to 0}\int_0^{\infty}(1-e^{-tx})\frac{\sin(t)}{t}dt = 0. \tag{20.742}$$

□

Maintenant que nous avons fait plein de lemmes, nous pouvons énoncer notre résultat principal, et le démontrer facilement.

Théorème 20.221 (Intégrale de Dirichlet[606])**.**
Nous avons

$$\lim_{b\to\infty}\int_0^b \frac{\sin(t)}{t}dt = \frac{\pi}{2}. \tag{20.743}$$

Nous avons écrit $\lim_{b\to\infty}\int_0^b \frac{\sin(t)}{t}dt$ et non $\int_0^{\infty}\frac{\sin(t)}{t}dt$ parce que cette dernière intégrale n'existe pas vraiment au sens de Lebesgue, voir le lemme 20.213. Dans la suite nous écrirons cependant $\int_0^{\infty}\frac{\sin(t)}{t}dt$, en gardant en tête que cela n'est défini que via la limite.

Démonstration. Nous nommons D la valeur que nous cherchons. Le lemme 20.214 nous assure que

$$D = \lim_{b\to\infty}\int_0^b \frac{\sin(t)}{t} < \infty. \tag{20.744}$$

Le lemme 20.217 nous donne, quant à lui,

$$\lim_{b\to\infty}\int_0^b e^{-tx}\frac{\sin(t)}{t}dt=-\arctan(x)+\frac{\pi}{2}. \tag{20.745}$$

Vu que les deux limites existent, on peut permuter somme et limite [48] :

$$D+\arctan(x)-\frac{\pi}{2}=\lim_{b\to\infty}\Big(\int_0^b\frac{\sin(t)}{t}+\int_0^b e^{-tx}\frac{\sin(t)}{t}dt\Big) \tag{20.746a}$$

$$=\lim_{b\to\infty}\int_0^b(1-e^{-xt})\frac{\sin(t)}{t}dt. \tag{20.746b}$$

SUBEQooXIUNooPZnCPb

Pour (20.746b), nous avons des fonctions bornées sur un intervalle borné $]0,b[$, donc il n'y a pas de mal à sommer les intégrales.

Donc pour tout $x>0$, nous avons

$$D+\arctan(x)-\frac{\pi}{2}=\int_0^\infty(1-e^{-xt})\frac{\sin(t)}{t}dt. \tag{20.747}$$

Nous passons à la limite $x\to 0$ en utilisant le lemme 20.220 et le fait que $\arctan(0)=0$ (lemme 18.43) :

$$D-\frac{\pi}{2}=0, \tag{20.748}$$

c'est-à-dire que résultat annoncé. □

48. C'est une phrase un peu grandiloquente pour dire que $\lim_{b\to a}f(b)-\lim_{b\to a}g(b)=\lim_{b\to a}(f(b)-g(b))$. Ici nous avons $a=\infty$ et les fonctions f et g sont celles définies par les intégrales.

Chapitre 21

Arcs paramétrés

La structure de ce chapitre, comme beaucoup de choses dans le Frido, est fortement liée au choix de présenter toutes les matières dans l'ordre mathématiquement logique. Nous devons donc le placer après la trigonométrie ; les propriétés principales des fonctions trigonométriques étant dans la proposition 18.20, et c'est la proposition 18.58 qui nous permet de dire que $\big(\cos(t), \sin(t)\big)$ décrit le cercle.

Et enfin nous n'avons pas encore calculé la circonférence du cercle, et pour cause : nous n'avons pas encore donné de définition à la longueur d'un chemin dans $\mathbb{R}^2$. C'est pourquoi ce chapitre va aller droit à la longueur avant de donner des exemples.

21.1 Définitions

SecDeExCPar

Définition 21.1.
Un ***arc paramétré*** *dans $\mathbb{R}^p$ est un couple (I, γ) où I est un intervalle de $\mathbb{R}$ et γ est une application continue de I dans $\mathbb{R}^p$. Nous disons que (I, γ) est un arc paramétré* ***compact*** *(ou un* ***chemin*** *dans $\mathbb{R}^p$) lorsque I est compact dans $\mathbb{R}$.*

L'intervalle I d'un arc paramétré compact est toujours de la forme $[a, b]$, étant donné que tous les intervalles compacts de $\mathbb{R}$ sont de cette forme. Un **sous arc** de (I, γ) est un arc de la forme (I_0, γ) avec $I_0 \subset I$.

La définition 20.75 a déjà dit qu'un chemin dans $\mathbb{R}^n$ est une application continue $\sigma\colon [a, b] \to \mathbb{R}^n$.

Si le chemin σ est dérivable, la fonction σ' est la **vitesse** du chemin σ. Si le chemin est deux fois dérivable, la fonction σ'' l'**accélération**.

L'image de σ dans $\mathbb{R}^3$ est ce qu'on appellera la **courbe**.

21.2 Longueur d'arc

SecLongArc

Nous voulons définir et étudier la notion de longueur d'un arc paramétré. Pour cela, le plus raisonnable est d'approcher l'arc par des petits segments de droites (dont les longueurs sont évidentes), et d'extraire la « meilleure » approximation.

Une des notions clefs pour la suite est celle de subdivision d'intervalles. Cette notion sera encore utilisée par la suite à propos des intégrales.

DefSubdivisionIntervalle

Définition 21.2.
Si I est un intervalle d'extrêmes a et b avec $a < b$, nous appelons ***subdivision finie*** *de I un choix de nombres t_i tels que*

$$a = t_0 < t_1 < \ldots < t_n = b. \tag{21.1}$$

Nous disons qu'une subdivision σ' est ***plus fine*** *que la subdivision σ si l'ensemble des points de σ est inclus dans celui des points de σ'. Dans ce cas, la subdivision σ' est un* ***raffinement*** *de σ. Nous désignons par $\mathcal{S}(I)$ l'ensemble des subdivisions finies de l'intervalle I.*

Dans la suite, toutes les subdivisions que nous considérons seront des subdivisions finies. Aussi nous parlerons simplement de *subdivisions* sans préciser. Nous allons souvent noter $\sigma = (t_i)_{i=1}^n$ pour désigner la subdivision formée par les nombres t_i. Il faut garder en tête que dans une subdivision, les nombres *sont ordonnés*.

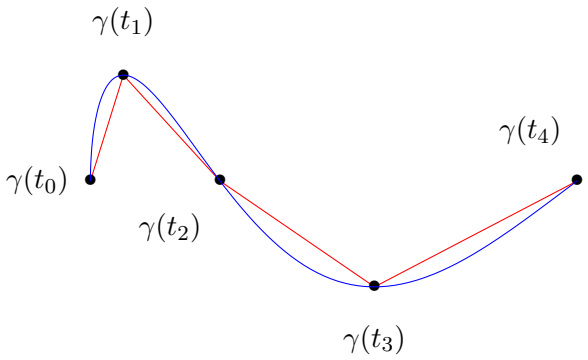


FIGURE 21.1: La longueur d'un découpage. La somme des longueurs des segments droits est facile à calculer. LabelFigCourbeRectifiable

DEFooDNZWooXmxhsU

Définition 21.3.
Soit un arc paramétré compact (I, γ) et une subdivision $\sigma = (t_i)_{i=0}^n$ de $I = [a, b]$. À partir de γ et du découpage σ nous définissons le nombre (voir figure 21.1)

$$l_\sigma(\gamma) = \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\|. \tag{21.2}$$ Eqlsigmagammasss

On appelle ***longueur*** *de l'arc γ le nombre*

$$l(\gamma) = \sup_\sigma l_\sigma(\gamma) \in [0, \infty]. \tag{21.3}$$

Nous disons que γ est ***rectifiable*** *lorsque $l(\gamma) < \infty$.*

Lorsque nous voulons spécifier sur quel intervalle nous considérons l'arc, nous noterons $l(I, \gamma)$ au lieu de $l(\gamma)$ pour être plus précis.

Par l'inégalité triangulaire, si σ_1 est plus fine que σ, nous avons

$$l_\sigma(\gamma) \leqslant l_{\sigma_1}(\gamma), \tag{21.4}$$

Comme cela peut être vu sur la figure 21.2.

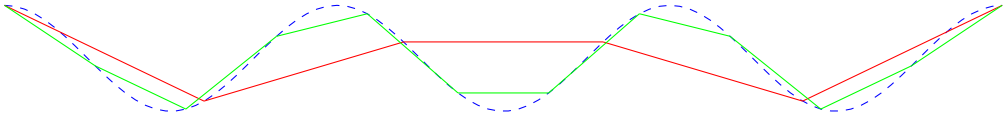


FIGURE 21.2: Il est visible que la longueur donnée par l'approximation par des petits segments (verts) est plus longue et plus précise que celle donnée par les longs segments (rouge). LabelFigArcLongueurFinesse

PROPooCXLYooRpKDMs

Proposition 21.4.
Si P et Q sont des points de $\mathbb{R}^2$, alors le segment de droite joignant P à Q est le plus court des arcs paramétrés passant par P et Q.

Démonstration. Si γ est un arc paramétré joignant P et Q, la longueur de γ est donné par un supremum dont un des éléments est la longueur du segment de droite. □

Dans la vie réelle, il est souvent difficile et peu pratique de calculer le supremum « à la main ». C'est pourquoi nous allons travailler à exprimer la longueur d'un arc à l'aide d'une intégrale (théorème 21.9).

Lemme 21.5.
Nous avons $l(\gamma) = 0$ si et seulement si $\gamma(t)$ est un vecteur constant.

Démonstration. Si l'application $\gamma(t)$ est constante, le résultat est évident. Supposons maintenant que γ ne soit pas constante. Cela signifie qu'il existe t_1 et t_2 dans I tels que $\gamma(t_1) \neq \gamma(t_2)$. Dans ce cas, si nous prenons le découpage $\sigma = \{a, t_1, t_2, b\}$, la somme (21.2) contient au moins le terme non nul $\|\gamma(t_2) - \gamma(t_1)\|$, et donc $l_\sigma(\gamma) > 0$. Par définition du supremum, nous avons alors $l(\gamma) \geqslant l_\sigma(\gamma) > 0$. □

Propletautredecop

Proposition 21.6.
Soit (I, γ) un arc paramétré compact.

(1) Si $\gamma' = (I', \gamma)$ avec $I' \subset I$, alors $l(\gamma') \leqslant l(\gamma)$.

(2) Soit $c \in [a, b]$, et considérons les arcs $\gamma_1 = \big([a, c], \gamma\big)$ et $\gamma_2 = \big([c, b], \gamma\big)$. Alors

$$l(\gamma) = l(\gamma_1) + l(\gamma_2). \tag{21.5}$$

En particulier, γ est rectifiable si et seulement si γ_1 et γ_2 le sont.

Démonstration. (1) Nous notons $I = [a, b]$ et $I' = [a', b']$. Étant donné que $I' \subset I$, nous avons

$$a \leqslant a' < b' \leqslant b. \tag{21.6}$$

Pour chaque subdivision $\sigma_0 : a' = t_0 < t_1 < \ldots < t_n = b'$ de I', nous pouvons construire une subdivision de I en « ajoutant » les points a et b, c'est-à-dire

$$\sigma : a \leqslant t_0 < \ldots < t_n \leqslant b. \tag{21.7}$$

Si nous calculons $l_\sigma(\gamma)$, nous avons tous les termes qui arrivent dans $l_{\sigma_0}(\gamma')$ plus le premier et dernier terme : $\|\gamma(t_0) - \gamma(a)\|$ et $\|\gamma(b) - \gamma(t_n)\|$. Nous avons donc

$$l_{\sigma_0}(\gamma') \leqslant l_\sigma(\gamma) \leqslant \sup_\sigma l_\sigma(\gamma) = l(\gamma). \tag{21.8}$$

Étant donné que pour toute subdivision σ_0 nous avons $l_{\sigma_0}(\gamma') \leqslant l(\gamma)$, en prenant le supremum sur les subdivisions σ_0 de I', nous avons comme annoncé

$$l(\gamma') \leqslant l(\gamma). \tag{21.9}$$

(2) Soit $\sigma = \{t_i\}$ une subdivision de $[a, b]$. Nous considérons les subdivisions σ_1 et σ_2 définies comme suit :

$$\begin{aligned} \sigma_1 &: \{t_i \text{ tel que } t_i < c\} \cup \{c\}, \\ \sigma_2 &: \{t_i \text{ tel que } t_i > c\} \cup \{c\}. \end{aligned} \tag{21.10}$$

L'inégalité triangulaire implique que

$$l_\sigma(\gamma) \leqslant l_{\sigma\cup\{c\}}(\gamma) = l_{\sigma_1}(\gamma_1) + l_{\sigma_2}(\gamma_2) \leqslant l(\gamma_1) + l(\gamma_2). \tag{21.11}$$

Nous avons donc

$$l(\gamma) \leqslant l(\gamma_1) + l(\gamma_2). \tag{21.12}$$

EqIneglglglgud

Nous prouvons maintenant l'inégalité inverse. Soit $\varepsilon > 0$. Étant donné que $l(\gamma_1)$ est le supremum des quantités $l_{\sigma_1}(\gamma_1)$ lorsque σ_1 parcours toutes les subdivisions possibles, il existe une partition σ_1^ε telle que (idem pour γ_2)

$$\begin{aligned} l_{\sigma_1^\varepsilon}(\gamma_1) + \frac{\varepsilon}{2} &> l(\gamma_1), \\ l_{\sigma_2^\varepsilon}(\gamma_2) + \frac{\varepsilon}{2} &> l(\gamma_2), \end{aligned} \tag{21.13}$$

EqAllsigmaepsgammaufd

où σ_1^ε est une subdivision de $[a,c]$ et σ_2^ε en est une de $[c,b]$. En faisant la somme des deux équations (21.13), nous trouvons

$$l(\gamma_1)+l(\gamma_2)<l_{\sigma_1^\varepsilon}(\gamma_1)+l_{\sigma_2^\varepsilon}(\gamma_2)+\varepsilon=l_{\sigma_1^\varepsilon\cup\sigma_2^\varepsilon}(\gamma)\leqslant l(\gamma)+\varepsilon. \tag{21.14}$$

L'inégalité $l(\gamma_1)+l(\gamma_2)<l(\gamma)+\varepsilon$ étant valable pour tout ε, nous avons

$$l(\gamma_1)+l(\gamma_2)\leqslant l(\gamma). \tag{21.15}$$

Cette inégalité, combinée avec l'inégalité (21.12), donne bien $l(\gamma)=l(\gamma_1)+l(\gamma_2)$.

□

21.3 Abscisse curviligne

Définition 21.7.
Soit (I,γ) un arc rectifiable compact avec $I=[a,b]$. L'application

$$\begin{aligned}\varphi\colon [a,b]&\to\mathbb{R}^+\\ t&\mapsto l\big([a,t],\gamma\big)\end{aligned} \tag{21.16}$$

*est la **longueur d'arc** de γ.*

Cette fonction nous permet de calculer la distance (suivant la courbe) entre deux points arbitraires parce que si $a\leqslant t<u\leqslant b$, nous avons

$$l\big([t,u],\gamma\big)=\varphi(u)-\varphi(t). \tag{21.17}$$

En effet,

$$\varphi(u)-\varphi(t)=l\big([a,u],\gamma\big)-l\big([a,t],\gamma\big), \tag{21.18}$$

mais en utilisant la proposition 21.6, nous avons

$$l\big([a,u],\gamma\big)=l\big([a,t],\gamma\big)+l\big([t,u],\gamma\big). \tag{21.19}$$

Proposition 21.8.
La longueur d'arc d'un arc rectifiable compact est une fonction continue et croissante.

Démonstration. Soit (I,γ) un arc paramétré rectifiable compact avec $I=[a,b]$. Afin de montrer que φ est croissante, prenons $t\in I$ ainsi que $h>0$ et montrons que $\varphi(t+h)\geqslant\varphi(t)$. La proposition 21.6 implique que

$$l\big([a,t+h],\gamma\big)=l\big([a,t],\gamma\big)+l\big([t,t+h],\gamma\big), \tag{21.20}$$

c'est-à-dire

$$\varphi(t+h)=\varphi(t)+l\big([t,t+h],\gamma\big)\geqslant\varphi(t). \tag{21.21}$$

Pour la continuité, soit t fixé dans $[a,b]$ et $\varepsilon>0$. Il nous faut démontrer qu'il existe $\eta>0$ tel que si s est dans $[0,\eta]$ alors

$$|\varphi(t+s)-\varphi(t)|\leqslant\varepsilon,\qquad\forall t\in[a,b].$$

Étant donné que $l\big([t,b],\gamma\big)$ est le supremum des $l_\sigma\big([t,b],\gamma\big)$, il existe une subdivision σ donnée par les points $t,t_1,\cdots,t_{n-1},b$ telle que

$$l_\sigma\big([t,b],\gamma\big)>l\big([t,b],\gamma\big)-\frac{\varepsilon}{2}=\varphi(b)-\varphi(t)-\frac{\varepsilon}{2}. \tag{21.22}$$

La continuité de γ implique qu'il existe un η tel que

$$s\in[0,\eta]\Rightarrow\|\gamma(t+s)-\gamma(t)\|<\frac{\varepsilon}{2} \tag{21.23}$$

Quitte à prendre η encore plus petit, nous supposons que $t+\eta<t_1$. Soit $s\in[0,\eta]$ et considérons la subdivision de $[t,b]$ donnée par $\sigma'=\sigma\cup\{t+s\}$. Étant donné que σ' est plus fine que σ, le nombre $l_\sigma\big([t,b],\gamma\big)$ est inférieur ou égal à $l_{\sigma'}\big([t,b],\gamma\big)$. Nous avons donc les inégalités

$$\begin{aligned}\varphi(b)-\varphi(t)-\frac{\varepsilon}{2}&\leqslant l_\sigma\big([t,b],\gamma\big)\\&\leqslant l_{\sigma'}\big([t,b],\gamma\big)\\&=\big\|\gamma(t+s)-\gamma(t)\big\|+l_{\sigma'\setminus\{t\}}\big([t+s,b]\gamma\big)\\&\leqslant\|\gamma(t+s)-\gamma(t)\|+\varphi(b)-\varphi(t+s)\\&\leqslant\frac{\varepsilon}{2}+\varphi(b)-\varphi(t+s).\end{aligned}\tag{21.24}$$

Au final, nous avons trouvé que

$$\varphi(t+s)-\varphi(t)\leqslant\varepsilon,\tag{21.25}$$

ce qui prouve que φ est continue au point t. □

En guise de paramètre sur un arc, nous pouvons utiliser la longueur d'arc elle-même. En effet si (I,γ) est un arc de longueur l, nous pouvons donner le même arc avec le couple $\big([0,l],g\big)$ où g est la fonction qui au réel s fait correspondre l'élément $\gamma\big(\varphi^{-1}(s)\big)$ de $\mathbb{R}^n$. Dire

$$P=(\gamma\circ\varphi^{-1})(s)\tag{21.26}$$

revient à dire que le point P est le point sur la courbe sur lequel on tombe après avoir marché une distance s sur la courbe.

Nous allons revenir sur ce « changement de paramètre » plus tard, en particulier dans la section 21.7.

21.3.1 Formule intégrale de la longueur

Nous pouvons voir un chemin γ comme étant la trajectoire d'une particule en fonction du temps. Sa vitesse à l'instant t est le vecteur $\gamma'(t)$, tandis que sa vitesse *scalaire* est le nombre $\|\gamma'(t)\|$. Une question naturelle est de savoir quelle est la longueur de la trajectoire parcourue entre $t=a$ et $t=b$.

Si nous prenons un petit intervalle de temps dt, nous pouvons supposer que le mobile avance à la vitesse constante $\|\gamma'(t)\|$. Cela ferait un trajet parcouru de longueur $\|\gamma'(t)\|dt$. Nous nous attendons donc à une formule de la forme suivante pour la longueur de γ :

$$l(\gamma)=\int_a^b\|\gamma'(t)\|dt.\tag{21.27}$$ EqDefLongueurChemin

Plus explicitement, si $\gamma(t)=\big(x(t),y(t),z(t)\big)$, alors nous aurions la formule

$$l(\gamma)=\int_a^b\sqrt{x'(t)^2+y'(t)^2+z'(t)^2}dt.\tag{21.28}$$

ThoLongueurIntegrale

Théorème 21.9.
Soit (I,γ) un arc paramétré compact de classe $\mathcal{C}^1$. Alors γ est rectifiable et

$$l(\gamma)=\int_a^b\|\gamma'(t)\|dt=\int_\gamma 1,\tag{21.29}$$ EqLongGammalInt

où $I=[a,b]$.

Démonstration. L'égalité avec l'intégrale le long de γ de la fonction 1 est simplement la définition 20.40 de l'intégrale curviligne.

Si $\sigma = \{t_i\}$ est une subdivision de l'intervalle $[a,b]$, alors

$$\begin{aligned} l_\sigma(\gamma) &= \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| \\ &= \sum_{i=1}^{n} \| \int_{t_{i-1}}^{t_i} \gamma'(t)dt\| \\ &\leqslant \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \|\gamma'(t)\|dt \\ &= \int_a^b \|\gamma'(t)\|dt. \end{aligned} \tag{21.30}$$

Cela prouve déjà que

$$l(\gamma) = \sup_\sigma l_\sigma(\gamma) \leqslant \int_a^b \|\gamma'(t)\|dt. \tag{21.31}$$ Eq_0208lsigsigmmintifp

Nous devons maintenant prouver l'inégalité inverse.

Notons φ l'abscisse curviligne $\varphi(t) = l\big([a,t],\gamma\big)$. Cette dernière vérifie

$$\varphi(t+h) - \varphi(t) = l\big([t,t+h],\gamma\big) \geqslant \|\gamma(t+h) - \gamma(t)\|, \tag{21.32}$$

et en particulier

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leqslant \frac{\varphi(t+h) - \varphi(t)}{h}. \tag{21.33}$$ Eq_0208intervpvpintfrach

D'autre part, en utilisant (21.31) sur le segment $[t,t+h]$, nous avons

$$\varphi(t+h) - \varphi(t) = l\big([t,t+h],\gamma\big) \leqslant \int_t^{t+h} \|\gamma'(u)\|du. \tag{21.34}$$

Cela nous permet de continuer l'inéquation (21.33) en

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leqslant \frac{\varphi(t+h) - \varphi(t)}{h} \leqslant \frac{1}{h} \int_t^{t+h} \|\gamma'(u)\|du. \tag{21.35}$$

Prenons la limite $h \to 0$. À gauche nous reconnaissons la formule de la dérivée, et nous obtenons $\|\gamma'(t)\|$; au centre nous avons $\varphi'(t)$ et à droite, si $n(u)$ représente une primitive de la fonction $u \mapsto \|\gamma'(u)\|$,

$$\lim_{h\to 0} \frac{n(t+h) - n(t)}{h} = n'(t) = \|\gamma'(t)\|. \tag{21.36}$$

Au final,

$$\|\gamma'(t)\| \leqslant \varphi'(t) \leqslant \|\gamma'(t)\|, \tag{21.37}$$

c'est-à-dire $\varphi'(t) = \|\gamma'(t)\|$ et donc par le théorème fondamental du calcul intégral 14.287,

$$\varphi(t) - \varphi(a) = \int_a^t \|\gamma'(u)\|du. \tag{21.38}$$

Par construction de la longueur d'arc, $\varphi(a) = 0$ et en posant $t = b$ nous obtenons la relation recherchée :

$$l(\gamma) = \varphi(b) = \int_a^b \|\gamma'(u)\|du. \tag{21.39}$$

□ RemLongIntUn

Remarque 21.10.
Cela est cohérent avec 20.186, mais il faut garder en tête que $l(\gamma)$ n'est pas la mesure de Lebesgue de l'image de γ dans $\mathbb{R}^2$. Cette dernière est nulle.

Exemple 21.11.
Soient donc a et b deux points de $\mathbb{R}^m$, et γ la droite joignant a à b, c'est-à-dire

$$\gamma(t) = (1-t)a + tb \tag{21.40}$$

avec $t \in [0,1]$. Le théorème 21.9 nous enseigne que la longueur de ce chemin est

$$l\big([0,1],\gamma\big) = \int_0^1 \|\gamma'(t)\|dt = \int_0^1 \| - a + b\| = \|b-a\|, \tag{21.41}$$

qui est bien la distance entre a et b. △

Exemple 21.12 (Circonférence du cercle)**.**
Nous savons que l'image de

$$\begin{aligned}\gamma\colon [0,2\pi[&\to \mathbb{R}^2\\ t &\mapsto \big(R\cos(t), R\sin(t)\big)\end{aligned} \tag{21.42}$$

est le cercle de centre $(0,0)$ et de rayon $R>0$. Et de plus cet arc est de classe C^1 (et même C^∞) par la proposition 18.1. La longueur sera, d'après la formule (21.9)

$$l_\gamma = \int_0^{2\pi} \|\gamma'(t)\|dt = 2\pi R \tag{21.43}$$

grâce à la formule $\sin^2 + \cos^2 = 1$ du lemme 18.4.

Mais tout cela n'est pas satisfaisant parce que nous n'avons pas encore de valeur numérique de π.

Il y a une autre façon de faire en considérant le quart de cercle dont la longueur en fonction de π est vite calculée par

$$l_\gamma = \int_0^{\pi/2} \|\gamma'(t)\|dt = \frac{\pi R}{2}. \tag{21.44}$$

Cette même longueur est calculée en termes de fonctions plus courantes avec le chemin

$$\begin{aligned}\sigma\colon]0,1[&\to \mathbb{R}^2\\ t &\mapsto \left(\frac{t}{\sqrt{R^2-t^2}}\right)\end{aligned} \tag{21.45}$$

La longueur s'exprime avec

$$l_\sigma = \int_0^1 \sqrt{\frac{R^2}{R^2+t^2}}dt. \tag{21.46}$$

EQooIIKSooBQMgWY

Notons que le changement de variables $t = R\sin(u)$ permet de retrouver l'expression $l_\sigma = \pi R/2$.

Pour avoir une approximation de π, il est loisible de calculer une approximation numérique de l'intégrale (21.46) (avec $R=1$) et de l'égaler à $\pi/2$. △

PROPooDMSTooEOFExj

Proposition 21.13.
Soit le cercle $S = \{x \in \mathbb{R}^2$ tel que $\|x\| = 1\}$. Nous considérons la bijection[1] *$\varphi\colon [0,2\pi[\to S$. Soient deux points $a, b \in S$ ainsi que $\theta_a = \varphi^{-1}(a)$, $\theta_b = \varphi^{-1}(b)$. Nous supposons que $\theta_a \leqslant \theta_b$ et que $\widehat{a(0,0)b} = \theta$*[2]*.*

Nous considérons le chemin

$$\begin{aligned}\gamma\colon [\theta_a,\theta_b] &\to S\\ t &\mapsto \big(R\cos(t), R\sin(t)\big).\end{aligned} \tag{21.47}$$

Alors $l(\gamma) = R\theta$[3]*.*

1. Corolaire 18.59.
2. Angle, définition 18.157
3. Longueur d'arc, définition 21.3.

Démonstration. Vu que $\|a\| = \|b\| = R$, la définition de l'angle $\widehat{aob}$ dit que

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \frac{a}{R} = \frac{b}{R}. \tag{21.48}$$

En substituant $a = R\big(\cos(\theta_a), \sin(\theta_b)\big)$ et $b = R\big(\cos(\theta_b), \sin(\theta_b)\big)$ et en faisant une petite multiplication matrice par vecteur :

$$\begin{pmatrix} \cos(\theta)\cos(\theta_a) - \sin(\theta)\sin(\theta_a) \\ \sin(\theta)\cos(\theta_a) + \cos(\theta)\sin(\theta_a) \end{pmatrix} = \begin{pmatrix} cos(\theta_b) \\ \sin(\theta_b). \end{pmatrix} \tag{21.49}$$

En utilisant les formules de trigonométrie du lemme 18.14, il vient

$$\begin{cases} \cos(\theta + \theta_a) = \cos(\theta_b) & \text{(21.50a)} \\ \sin(\theta + \theta_a) = \sin(\theta_b). & \text{(21.50b)} \end{cases}$$

La proposition 18.23 nous enseigne alors que $\theta_b = \theta + \theta_a + 2k\pi$ pour un certain $k \in \mathbb{Z}$. Vu que θ_a et θ_b sont entre 0 et 2π, il est obligatoire que $k = 0$ et donc que $\theta_b = \theta_a + \theta$.

Enfin nous calculons la longueur de γ avec le théorème 21.9 qui permet de réduire la longueur à une intégrale : $l(\gamma) = \int_{\theta_a}^{\theta_a+\theta} \|\gamma'(t)\|dt$.

Dans notre cas, $\gamma'(t) = R\big(-\sin(t), \cos(t)\big)$ et donc $\|\gamma'(t)\| = R$. Donc

$$l(\gamma) = \int_{\theta_a}^{\theta_a+\theta} \|\gamma'(t)\|dt = R\theta, \tag{21.51}$$

comme annoncé. □

Exemple 21.14.
Considérons l'arc de cercle de rayon R interceptée par l'angle θ présenté sur la figure 21.3.

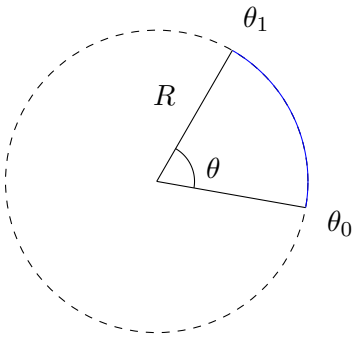


FIGURE 21.3: Quelle est la longueur de la partie bleue de ce cercle de rayon R ?
LabelFigAMDUooZZUOqa

Par définition, cette longueur sera

$$\int_{\theta_0}^{\theta_1} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}dt = R(\theta_1 - \theta_0). \tag{21.52}$$

Le radian comme unité de mesure d'angle est donc l'unité parfaite : elle est la longueur d'arc interceptée (si le rayon est $R = 1$). △

Une conséquence à peine indirecte de ce que nous venons de voir à propos de longueur d'arc de cercle est la proposition suivante [4].

PROPooYMMKooSUBtoo

Proposition 21.15.
Pour tout $x, y \in \mathbb{R}$, nous avons

$$|e^{ix} - e^{iy}| \leqslant |x - y|. \tag{21.53}$$

4. À mon avis il y a moyen de prouver ça avec un développement limité, mais je ne sais pas trop comment majorer l'erreur sans accepter que x soit arbitrairement proche de y. Si vous savez comment faire, écrivez-moi.

Démonstration. Évacuons tout de suite la différence entre $\mathbb{R}^2$ et $\mathbb{C}$: ils sont isométriques. Si vous n'êtes pas convaincu que tout se passe bien, vous pouvez récrire toute la démonstration en écrivant systématiquement $\big(\cos(x), \sin(x)\big)$ au lieu de e^{ix}. Cela serait au passage un bon exercice pour voir que les formules de dérivation fonctionnent bien.

Nous considérons les points e^{ix} et e^{iy} dans $\mathbb{C}$ et deux chemins différents les joignant. Le premier est le segment de droite

$$\begin{aligned}\sigma_1 : [0,1] &\to \mathbb{C}\\ t &\mapsto te^{ix} + (1-t)e^{iy}.\end{aligned} \tag{21.54}$$

Le second est l'arc de cercle

$$\begin{aligned}\sigma_1 : [0,1] &\to \mathbb{C}\\ t &\mapsto e^{i\big(tx+(1-t)y\big)}.\end{aligned} \tag{21.55}$$

Nous avons $\sigma_1'(t) = e^{ix} - e^{iy}$ qui ne dépend pas de t, et donc la longueur est facile à calculer à partir de la formule intégrale du théorème 21.9 :

$$l(\sigma_1) = \int_0^1 |\sigma_1'(t)| = |e^{ix} - e^{iy}|. \tag{21.56}$$

En ce qui concerne le second chemin,

$$\sigma_2'(t) = (x-y)e^{i\big(tx+(1-t)y\big)}. \tag{21.57}$$

Nous avons [5] $|\sigma_2'(t)| = |x-y|$ qui ne dépend pas non plus de t. Donc

$$l(\sigma_2) = |x-y|. \tag{21.58}$$

Étant donné la proposition 21.4 qui dit que le chemin le plus court est le segment de droite,

$$l(\sigma_1) < l(\sigma_2) \tag{21.59}$$

et donc le résultat annoncé. □

21.16.
Si on veut savoir la longueur d'une courbe donnée sous la forme d'une fonction $y = y(x)$, un chemin qui trace la courbe est évidemment donné par

$$\gamma(t) = (t, y(t)), \tag{21.60}$$

et le vecteur tangent au chemin est $\gamma'(t) = (1, y'(t))$. Donc

$$\|\gamma'(t)\| = \sqrt{1 + y'(t)^2}, \tag{21.61}$$

et

$$L = \int_a^b \sqrt{1 + y'(t)^2}. \tag{21.62}$$

EqLongFonction

Exemple 21.17.
La longueur de l'hélice

$$\sigma(t) = \begin{pmatrix}\cos(2t)\\ \sin(2t)\\ \sqrt{5}t\end{pmatrix} \tag{21.63}$$

pour $t \in [0, 4\pi]$ est donnée par

$$l(\sigma) = \int_0^{4\pi} \sqrt{4\sin^2(2t) + 4\cos^2(2t) + 5}dt = \int_0^{4\pi} \sqrt{9} = 12\pi. \tag{21.64}$$

△

5. Si vous voulez citer des résultats, lemme 18.11 et proposition 9.106.

Définition 21.18.
Soit $\sigma_1\colon [a,b]\to\mathbb{R}^3$, un chemin et $\sigma_2\colon [c,d]\to\mathbb{R}^3$, un autre chemin. On dit que ces chemins sont ***équivalents*** *si il existe une fonction $\varphi\colon [a,b]\to[c,d]$ strictement croissante telle que $\sigma_1(t)=\sigma_2\big(\varphi(t)\big)$.*

Deux chemins équivalents parcourent la même courbe dans le même sens. Ils ne le parcourent toutefois pas à la même vitesse. On dit que les chemins sont **opposés** si la fonction φ de la définition est strictement décroissante. Dans ce cas, ils ont la même image, mais parcourue dans le sens opposés. Nous disons que deux chemins équivalents sont un **changement de paramétrage** pour la même courbe.

Dans le cas d'un paramétrage équivalente, nous avons $\varphi(a)=c$ et $\varphi(b)=d$. Les points de départ et d'arrivée des deux paramètres coïncident. Dans le cas d'un paramètre qui va dans le sens opposé par contre nous avons automatiquement $\varphi(a)=d$ et $\varphi(b)=c$.

Proposition 21.19.
La longueur d'une courbe ne dépend pas du paramètre (équivalent ou opposé) choisi.

Démonstration. Soient $\sigma_1\colon [a,b]\to\mathbb{R}^3$ et $\sigma_2\colon [c,d]\to\mathbb{R}^3$ tels que

$$\sigma_1(t)=\sigma_2\big(\varphi(t)\big) \tag{21.65}$$

où $\varphi\colon [a,b]\to[c,d]$ est une bijection strictement monotone. Par définition on a

$$l(\sigma_1)=\int_a^b\|\sigma_1'(t)\|dt. \tag{21.66}$$

Nous pouvons exprimer la dérivée de σ_1 en termes de celle de σ_2 en dérivant la relation (21.65) :

$$\sigma_1'(t)=\varphi'(t)\sigma_2'\big(\varphi(t)\big). \tag{21.67}$$

En ce qui concerne la norme,

$$\|\sigma_1'(t)\|=|\varphi'(t)|\|\sigma_2'(t)\|. \tag{21.68}$$

Notez dans cette relation que $\varphi'(t)$ est un nombre (et non un vecteur). Étant donné que nous avons supposé que φ était monotone, soit elle est monotone croissante et $\|\varphi'(t)\|=\varphi'(t)$ pour tout t, soit elle est monotone décroissante et $\|\varphi'(t)\|=' \varphi(t)$ pour tout t.

Considérons d'abord le premier cas, c'est-à-dire $\|\varphi'(t)\|=\varphi'(t)$. Nous posons $s=\varphi(t)$, $ds=\varphi'(t)dt$. En remplaçant cela dans la formule de la longueur est

$$\begin{aligned}l(\sigma_1)&=\int_a^b\varphi'(t)\|\sigma_2'\big(\varphi(t)\big)\|dt\\&=\int_{\varphi(a)}^{\varphi(b)}\|\sigma_2'(s)\|ds\\&=\int_c^d\|\sigma_2'(s)\|ds\\&=l(\sigma_2).\end{aligned} \tag{21.69}$$

Si nous considérons maintenant un paramétrage strictement décroissante. Dans ce cas, $\varphi'(t)\leqslant 0$ et $\|\varphi'(t)\|=-\varphi'(t)$. Nous posons encore une fois $s=\varphi(t)$, $ds=\varphi'(t)ds$. Ici il ne faut pas oublier

que $\varphi(a) = d$ et $\varphi(b) = c$. Le calcul est à part cela le même en faisant attention au singe :

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t)\|\sigma_2'\big(\varphi(t)\big)\|dt \\ &= -\int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\|ds \\ &= -\int_d^c \|\sigma_2'(s)\|ds \\ &= \int_c^d \|\sigma_2'(s)\|ds \\ &= l(\sigma_2). \end{aligned} \tag{21.70}$$

Nous avons changé le signe en changeant l'ordre des bornes. □

21.4 Suite du chapitre

Le grand avantage des arcs paramétrés par rapports aux graphes de fonctions est qu'un arc paramétré peut « faire des retours en arrière », ou bien des auto intersections. Outre les deux exemples typiques de la la figure 21.4, un exemple classique est la droite verticale. Les fonctions $y = ax + b$ permettent de décrire toutes les droites, sauf les droites verticales. Dans le cadre des courbes paramétrées, les droites verticales et horizontales sont sur pied d'égalité. Quelques exemples classiques :

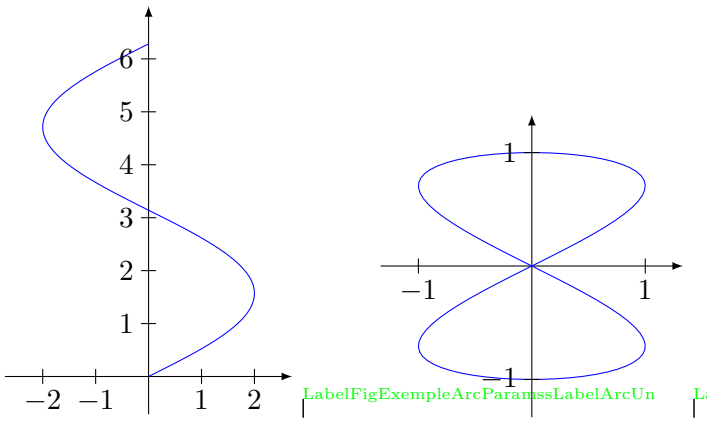
Droite horizontale Une droite horizontale à la hauteur a est donnée par la courbe paramétrée $\gamma(t) = (t, a)$, avec $t \in I = \mathbb{R}$.

Droite verticale Une droite verticale à la distance b de l'origine est donnée par la courbe paramétrée $\gamma(t) = (b, t)$, avec $t \in I = \mathbb{R}$.

PgGrqFnGamma

Graphe d'une fonction Le graphe d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est donné par l'arc $\gamma(t) = \big(t, f(t)\big)$.

Un cercle Le cercle de rayon R est donné par l'arc $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$.



LabelFigExempleArcParamssLabelArcUn LabelFigExempleArcParamssLabelArcI

(a) $\gamma(t) = (2\sin(t), t)$, avec $0 \leqslant t \leqslant 2\pi$ (b) $\gamma(t) = (sin(2t), sin(t))$ avec $-\pi \leqslant t \leqslant \pi$

FIGURE 21.4: Des exemples d'arcs paramétrées. Ceux ne sont pas des graphes.

LabelFigExempleArcParam

Remarque 21.20.
Afin d'alléger la notation, nous allons le plus souvent désigner l'arc (I, γ) simplement par la fonction γ. Il est cependant toujours *très* important de savoir sur quel intervalle nous considérons le chemin. Cela dépendra le plus souvent du contexte, et nous indiquerons l'intervalle I explicitement lorsqu'une ambigüité est à craindre.

Par exemple, lorsque nous considérons le cercle $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$, le plus souvent l'intervalle de variation de t sera $I = [0, 2\pi]$. Par contre, si nous considérons la droite $\gamma(t) = (t, 2t)$, l'intervalle de variation de t sera naturellement $I = \mathbb{R}$.

21.5 Autres exemples

Exemple 21.21.
Soit $v \in \mathbb{R}^3$ et $x_0 \in \mathbb{R}^3$. Le chemin

$$\sigma(t) = x_0 + tv \tag{21.71}$$

est une droite. Sa vitesse est $\sigma'(t) = v$. △

Exemple 21.22.
La courbe

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix} \in \mathbb{R}^2 \tag{21.72}$$

avec $t \in [0, 2\pi[$ est le cercle unité parcouru une fois dans le sens trigonométrique.

Notez que si on prend $t \in [0, 4\pi[$, nous avons un *autre* chemin ; c'est le même cercle unité, mais parcouru *deux* fois. Même si le « dessin » (le graphe) des deux est le même, le chemin n'est pas le même.

Le chemin

$$\gamma(t) = \begin{pmatrix}\cos(2\pi - t)\\ \sin(2\pi - t)\end{pmatrix} \tag{21.73}$$

est le cercle unité parcouru une fois dans le sens inverse. Encore une fois le « dessin » est le même, mais le chemin n'est pas le même. △

Exemple 21.23.
Le chemin

$$\sigma(t) = \begin{pmatrix}t\\ t^2\end{pmatrix} \tag{21.74}$$

est un chemin dont l'image est la parabole d'équation $y = x^2$. △

L'importance de la dérivée du chemin réside en le fait qu'elle donne la tangente. En effet le vecteur $\sigma'(t)$ est tangent au graphe de σ au point $\sigma(t)$.

CorKBEMooRvYAcJ

Corolaire 21.24.
La tangente à un cercle est perpendiculaire au rayon.

Démonstration. Nous savons que pour un cercle,

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \tag{21.75}$$

Un point général du cercle a pour abscisse $x = R\cos(\theta)$. En remplaçant nous trouvons le coefficient directeur suivant pour la tangente :

$$y'\big(R\cos(\theta)\big) = -\frac{1}{\tan(\theta)}. \tag{21.76}$$

Par conséquent une droite perpendiculaire à la tangente aurait comme coefficient directeur le nombre $\tan(\theta)$. Or cela est bien le coefficient directeur du rayon qui joint le point $(0, 0)$ au point $\big(R\cos(\theta), R\sin(\theta)\big)$.

□

Exemple 21.25.
Pour le cercle,

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}, \tag{21.77}$$

la dérivée est donnée par

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t). \end{pmatrix} \tag{21.78}$$

Le produit scalaire $\sigma(t) \cdot \sigma'(t)$ est nul. Le vecteur $\sigma'(t)$ est donc bien tangent (corolaire 21.24). △

Exemple 21.26.
Le courbe donnée par le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix} \tag{21.79}$$

est une hélice. Sa vitesse est

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix}. \tag{21.80}$$

Notez que pour tout $t \in \mathbb{R}$, nous avons $\|\sigma'(t)\| = \sqrt{2}$. △

Remarque 21.27.
Lorsqu'on parle d'une courbe dans l'espace, l'intervalle sur lequel on considère la variation du paramètre est une donné fondamentale. Elle fait partie intégrante de la définition de la courbe.

21.6 Élément de longueur

21.6.1 Élément de longueur : cartésiennes

Étant donné que la longueur d'arc d'une courbe paramétrée (I, γ) est donnée par l'intégrale de $\|\gamma'(t)\|$, il est naturel d'appeler le nombre $\|\gamma'(t)\|\,dt$, **l'élément de longueur** de la courbe γ au point $\gamma(t)$.

En coordonnées cartésiennes dans le plan, une courbe paramétrée est donnée par

$$\gamma(t) = \big(x_1(t), x_2(t)\big), \tag{21.81}$$

et l'élément de longueur est

$$\|x'(t)\|\,dt = \sqrt{(x_1')^2 + (x_2')^2}\,dt. \tag{21.82}$$ EqElLongCart

21.6.2 Élément de longueur : polaires (1)

En coordonnées polaires, une courbe est donnée par

$$\gamma(t) = \big(\rho(t), \theta(t)\big), \tag{21.83}$$

et le passage aux cartésiennes se fait via les formules

$$\begin{cases} x(t) = \rho(t)\cos\big(\theta(t)\big) & \text{(21.84a)} \\ y(t) = \rho(t)\sin\big(\theta(t)\big). & \text{(21.84b)} \end{cases}$$

L'élément de longueur se trouve directement en remplaçant $x(t)$ et $y(t)$ dans la formule (21.82). Les dérivées sont données par

$$\begin{aligned} x'(t) &= \rho'(t)\cos\theta(t) - \rho(t)\theta'(t)\sin\theta(t) \\ y'(t) &= \rho'(t)\sin\theta(t) + \rho(t)\theta'(t)\cos\theta(t), \end{aligned} \tag{21.85}$$

et un calcul montre que

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \big(\rho'(t)\big)^2 + \big(\rho(t)\big)^2\big(\theta'(t)\big)^2. \tag{21.86}$$ EqElLongEnPolaires

Nous reviendrons plus en détail sur le concept de changement de paramétrage (ici, les polaires) à la section 21.7.

21.6.3 Élément de longueur : polaires (2)

Parfois on utilise θ comme paramètre. L'équation de la courbe est alors donnée en coordonnées polaires sous la forme

$$\rho(\theta) = f(\theta), \tag{21.87}$$

où f est une fonction réelle et il faut comprendre que nous parlons de la courbe $\big(\rho(\theta), \theta\big)$ en coordonnées polaires. En coordonnées cartésiennes, cette courbe est donnée par

$$\begin{cases} x(t) = \rho(t)\cos(t) & (21.88a)\\ y(t) = \rho(t)\sin(t) & (21.88b) \end{cases}$$

avec t qui parcourt le plus souvent l'intervalle $[0, 2\pi]$. Notez qu'il se peut que le domaine ne soit pas toujours $[0, 2\pi]$; cela peut dépendre des circonstances. Quoi qu'il en soit, la donnée du domaine fait partie de la donnée d'une courbe, et il ne peut donc pas y avoir d'équivoques à ce niveau.

Nous utilisons à nouveau la formule (21.82) en y mettant les valeurs (21.88) :

$$\begin{cases} x'(t) = \rho'(t)\cos(t) - \rho(t)\sin(t) & (21.89a)\\ y'(t) = \rho'(t)\sin(t) + \rho(t)\cos(t), & (21.89b) \end{cases}$$

et

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \rho'(t)^2 + \rho(t)^2. \tag{21.90}$$

Remarque 21.28.
N'oubliez pas, en utilisant ces formules, que ce qui rentre dans l'intégrale est la *racine carré* de $(x')^2 + (y')^2$.

Exemple 21.29.
Calculons la circonférence du cercle. En coordonnées polaires, le graphe du cercle correspond à l'équation

$$\big(\rho(t), \theta(t)\big) = (R, t) \tag{21.91}$$

où R est constante (le rayon du cercle) et t va de 0 à 2π. En substituant dans l'équation (21.86), l'élément de longueur à intégrer est seulement

$$\sqrt{R^2} = R \tag{21.92}$$

parce que $\rho'(t) = 0$ et $\theta'(t) = 1$. La longueur du cercle est alors directement donnée par

$$l = \int_0^{2\pi} R dt = 2\pi R. \tag{21.93}$$

Nous pouvions aussi faire le calcul en coordonnées cartésiennes. Alors la courbe est donnée par les équations

$$\begin{aligned} x(t) &= R\cos(t)\\ y(t) &= R\sin(t) \end{aligned} \tag{21.94}$$

et $t \in [0, 2\pi]$. La circonférence du cercle est alors

$$l = \int_0^{2\pi} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}\, dt = \int_0^{2\pi} R\, dt = 2\pi R. \tag{21.95}$$

△

Remarque 21.30.
Il faut bien comprendre que quand on parle de courbes paramétrées en coordonnées cartésiennes on pense à une courbe dont le paramètre est, par exemple, t et les équations de la courbe sont $(x(t), y(t))$. Cela ne veut pas dire que x ou y soit le paramètre. Le cas où x ou y est le paramètre est un cas particulier qui est possible seulement pour certaines courbes et notamment pour les graphes. Le cercle de rayon 1 n'est pas un graphe, donc si on veut utiliser x ou y comme paramètre il faut d'abord découper la courbe en deux morceaux, par exemple, la moitié inférieure ($y < 0$) et la moitié supérieure ($y > 0$).

ExCycloLong

Exemple 21.31.
Une **cycloïde** est une courbe paramétrée par

$$\begin{cases} x(t) = a(t - \sin(t)) & \text{(21.96a)} \\ y(t) = a(1 - \cos(t)) & \text{(21.96b)} \end{cases}$$

avec $a > 0$ et $t \in \mathbb{R}$. Comme montré sur la figure 21.5, la cycloïde donne lieu à un graphe périodique. Il est possible de montrer (le faire) que le premier arc correspond à $t \in [0, 2\pi]$. Nous voulons donc calculer la longueur de l'arc sur cet intervalle.

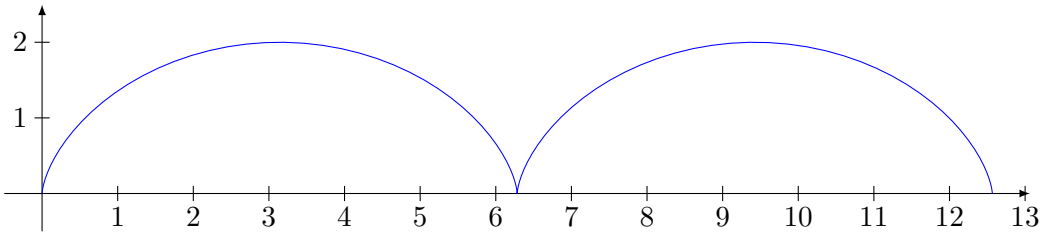


FIGURE 21.5: La cycloïde de paramètre $a = 1$ entre 0 et 4π.

LabelFigCycloideA

Nous avons $x'(t) = a(1 - \cos(t))$ et $y'(t) = a\sin(t)$, de telle façon que

$$\sqrt{(x')^2 + (y')^2} = a\sqrt{2 - 2\cos(t)} = a\sqrt{4\sin^2\left(\frac{t}{2}\right)} = 2a\left|\sin\frac{t}{2}\right|. \quad \text{(21.97)}$$

Eq_0508dlcycloide

La longueur est donc donnée par

$$\int_0^{2\pi} 2a|\sin\frac{t}{2}|dt = 4a\int_0^{\pi}\sin(t)dt = 8a. \quad \text{(21.98)}$$

△

Exemple 21.32.
La **cardioïde** est la courbe donnée par

$$\rho(\theta) = a(1 + \cos(\theta)). \quad \text{(21.99)}$$

EqCardioide

avec $\theta \in [-\pi, \pi]$. Le nom de cette courbe provient de son graphe illustré à la figure 21.6.

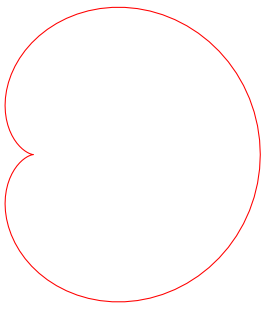


FIGURE 21.6: Une cardioïde, $\rho = 1 + \cos(\theta)$.

LabelFigCardioid

L'équation (21.99) est donnée sous la forme (21.87), c'est-à-dire que $\theta(t) = t$ et $\theta'(t) = 1$, et par conséquent l'élément de longueur est donné par

$$\begin{aligned}(\rho')^2 + (\rho)^2 &= \big(-a\sin(\theta)\big)^2 + a^2\big(1 + \cos(\theta)\big)^2 \\ &= a^2\sin^2(\theta) + a^2\big(1 + 2\cos(\theta) + \cos^2(\theta)\big) \\ &= a^2\big(1 + 1 + 2\cos(\theta)\big) \\ &= 2a^2\big(1 + \cos(\theta)\big) \\ &= 4a^2\cos^2\frac{\theta}{2}.\end{aligned} \quad \text{(21.100)}$$

La longueur d'arc est donc donnée par

$$l = \int_{-\pi}^{\pi} 2a\cos\frac{\theta}{2}d\theta = 2a\int_{-\pi/2}^{\pi/2}\cos(t)2dt = 8a. \tag{21.101}$$

△

21.6.4 Approximation de la longueur par des cordes

Définition 21.33 (Point régulier, birégulier[607])**.**
Soit un arc paramétré (I,γ). Un point $t \in I$ est dit ***régulier*** *si $\gamma'(t) \neq 0$, et il est dit* ***critique*** *si $\gamma'(t) = 0$. Le point $t \in I$ est dit* ***birégulier*** *si les vecteurs $\gamma'(t)$ et $\gamma''(t)$ sont linéairement indépendants et non nuls.*

Par extension, nous dirons également que le point $\gamma(t)$ lui-même est régulier, critique ou birégulier. Un arc est dit régulier *lorsque tous ses points sont réguliers.*

Note : dans le lemme 21.71 et ses dépendances, nous utilisons effectivement que l'arc γ est de classe C^2.

Nous savons que la longueur d'une courbe est donnée par le supremum sur toutes les subdivisions de la longueur des cordes correspondantes. De plus l'inégalité triangulaire nous enseigne que plus la subdivision est fine, plus la longueur sera grande. Il est donc naturel de penser que sur un petit intervalle, la longueur de la courbe ne doit pas être très différente de la longueur de la corde correspondante.

La proposition suivante est un énoncé précis et quantitatif de ce fait.

Proposition 21.34.
Soit (I,γ) un arc de classe $\mathcal{C}^1$ et $t_0 \in I$ un point régulier (c'est-à-dire $\gamma'(t_0) \neq 0$). Alors pour tout $\varepsilon > 0$, il y a un $\delta > 0$ tel que on trouve $t, t' \in I \cap (t_0, \delta)$ tels que

$$\left|\int_t^{t'} \|\gamma'(u)\|du - \|\gamma(t) - \gamma(t')\|\right| \leqslant 2\varepsilon|t - t'|. \tag{21.102}$$

Intuitivement, cette proposition signifie qu'au voisinage de t_0, la longueur d'arc est équivalente à celle de la corde.

Démonstration. Par la continuité de γ' (parce que γ est $\mathcal{C}^1$), pour tout ε, il existe un δ tel que

$$|t - t_0| < \delta \Rightarrow \|\gamma'(t) - \gamma'(t_0)\| \leqslant \varepsilon. \tag{21.103}$$

Nous considérons la fonction

$$u \mapsto \gamma(u) - \gamma(t_0) - (u - t_0)\gamma'(t_0), \tag{21.104}$$

dont la dérivée (par rapport à u) est

$$\gamma'(u) - \gamma'(t_0). \tag{21.105}$$

Nous y appliquons la formule des accroissements finis entre t et t' choisis dans $I \cap]t_0 - \delta, t_0 + \delta[$. Il existe un u entre t et t' tel que

$$\begin{aligned}\|\gamma(t) - \gamma(t_0) - (t - t_0)\gamma'(t_0) - \gamma(t') + \gamma(t_0) + (t' - t_0)\gamma'(t_0)\| \\ = |t - t'|\|\gamma'(u) - \gamma'(t_0)\| \\ \leqslant \varepsilon|t - t'|.\end{aligned} \tag{21.106}$$

En simplifiant ce qui peut être simplifié dans le membre de gauche, nous trouvons

$$\|\gamma(t) - \gamma(t') - (t - t')\gamma'(t_0)\| \leqslant \varepsilon|t - t'|. \tag{21.107}$$

Le membre de gauche peut être minoré en utilisant la proposition 7.159 :

$$\Big|\|\gamma(t)-\gamma(t')\|-\|(t-t')\gamma'(t_0)\|\Big|\leqslant\varepsilon|t-t'|. \tag{21.108}$$

D'autre part, les inégalités (7.140) montrent que

$$-\|\gamma'(u)-\gamma'(t_0)\|\leqslant\|\gamma'(u)\|-\|\gamma'(t_0)\|\leqslant\|\gamma'(u)-\gamma'(t_0)\|. \tag{21.109}$$

Si de plus u est compris entre t et t', ces inégalités sont encore coincées entre $-\varepsilon$ et ε. En intégrant (21.109) par rapport à u entre t et t', nous obtenons

$$\left|\int_t^{t'}\|\gamma'(u)\|-(t-t')\|\gamma'(t_0)\|\right|\leqslant\varepsilon|t-t'|. \tag{21.110}$$

Afin d'alléger les notations pour la ligne suivante, nous notons A le nombre positif $\int_t^{t'}\|\gamma'(u)\|du$. Nous avons

$$\begin{aligned}\Big|A-\|\gamma(t)-\gamma(t')\|\Big|&=\Big|A-|t-t'|\,\|\gamma'(t_0)\|+|t-t'|\,\|\gamma'(t_0)\|-\|\gamma(t)-\gamma(t')\|\Big|\\&\leqslant\Big|A-|t-t'|\,\|\gamma'(t_0)\|\Big|+\Big||t-t'|\,\|\gamma'(t_0)\|-\|\gamma(t)-\gamma(t')\|\Big|.\end{aligned} \tag{21.111}$$

L'équation (21.108) montre que le second terme est plus petit ou égal à $\varepsilon|t-t'|$. En ce qui concerne le premier terme, étant donné que A est positif,

$$\Big|A-|t-t'|\,\|\gamma'(t_0)\|\Big|\leqslant\Big|A-(t-t')\,\|\gamma'(t_0)\|\Big|\leqslant\varepsilon|t-t'|. \tag{21.112}$$

Au final, l'inéquation (21.111) donne

$$\Big|A-\|\gamma(t)-\gamma(t')\|\Big|\leqslant2\varepsilon\,|t-t'|, \tag{21.113}$$

ce qu'il fallait démontrer. □

21.7 Arc géométrique

Définition 21.35.
Soient (I,γ) et (J,g) deux arcs paramétrés de classe $\mathcal{C}^k$. On dit qu'il sont ***équivalents*** *si il existe une bijection $\theta\colon I\to J$ de classe $\mathcal{C}^k$, d'inverse de classe $\mathcal{C}^k$ telle que $g=\gamma\circ\theta$. Nous notons $\gamma\sim g$ lorsque γ et g sont équivalents (les ensembles I et J sont sous-entendus).*

Le passage d'un paramétrage (I,γ) à une autre (J,g) se fait selon le diagramme suivant :

$$\begin{array}{ccc} I & \xrightarrow{\gamma} & \mathbb{R}^n \\ {\scriptstyle\theta}\uparrow & \nearrow_{g} & \\ J & & \end{array} \tag{21.114}$$

Proposition 21.36.
La relation donnée dans la définition 21.35 est une relation d'équivalence.

Démonstration. Les trois points d'une relation d'équivalence se vérifient en utilisant le fait que θ est inversible, et que l'inverse θ^{-1} jouit des mêmes propriétés de continuité ($\mathcal{C}^k$) que θ.

Réflexivité Nous avons $\gamma\sim\gamma$ avec $\theta=\mathrm{Id}$.

Symétrie Si $\gamma\sim g$, alors nous avons une application θ telle que $g=\gamma\circ\theta$, et donc $\gamma=g\circ\theta^{-1}$, ce qui montre que $g\sim\gamma$.

Transitivité Si $\gamma \sim g$ et $g \sim h$ avec $g = \gamma \circ \theta$ et $h = g \circ \omega$, alors $h = \gamma \circ (\theta \circ \omega)$, ce qui montre que $\gamma \sim h$.

□

Si les arcs (I, γ) et (J, g) sont équivalents, les images dans $\mathbb{R}^n$ sont identiques, et décrivent donc « le même dessin ». Nous allons préciser cette notion plus loin.

Définition 21.37.
Pour cette raison les classes d'équivalences sont appelées des ***arcs géométriques*** *(de classe $\mathcal{C}^k$).*

Si Γ est un arc géométrique, ses représentants sont dits des **paramétrages admissibles** ou, plus simplement *paramétrage*. On dit que l'application $\theta\colon J \to I$ est un **changement de variable**. Nous disons que un arc géométrique est *compact* quand ses représentants sont compacts.

LemChamVarsStriMomnot

Lemme 21.38.
Dans le cas d'un arc $\mathcal{C}^1$, les changements de variables sont strictement monotones (croissants ou décroissants).

Démonstration. Nous considérons (I, γ) et (J, g), deux paramétrages différents du même arc géométrique, et $\theta \in \mathcal{C}^1(J, I)$ le changement de variable. Nous allons noter t la variable sur I et s la variable sur J. Par définition, $\theta\big(\theta^{-1}(t)\big) = t$, et par conséquent,

$$\theta'\big(\theta^{-1}(t)\big)(\theta^{-1})'(t) = 1. \tag{21.115}$$

En particulier $\theta'\big(\theta^{-1}(t)\big)$ ne s'annule pas pour aucune valeur de t. Mais $\theta^{-1}(t)$ peut prendre n'importe quelle valeur dans J, donc nous avons $\theta'(s) \neq 0$ pour tout $s \in J$. Cela signifie bien que θ est strictement monotone. En effet, θ' étant continue, elle ne peut pas changer de signe sans passer par zéro (théorème 9.93 des valeurs intermédiaires). □

ThoLongArcGeom

Théorème 21.39.
La longueur d'un arc est indépendante de son paramétrage, c'est-à-dire que les représentants d'un arc géométrique compact de classe $\mathcal{C}^1$ ont même longueur.

Démonstration. Nous utilisons les mêmes notations que celles du lemme 21.38. Nous savons déjà que le changement de variable $\theta\colon J \to I$ est strictement monotone. Supposons que θ soit croissante. En effectuant un changement de variable dans l'intégrale qui donne la longueur [6] nous avons

$$\begin{aligned}
l(\gamma) &= \int_I \|\gamma'(t)\| dt \\
&= \int_J \|\gamma'\big(\theta(s)\big)\|\theta'(s) ds \\
&= \int_J \|\gamma'\big(\theta(s)\big)\theta'(s)\| ds \\
&= \int_J \|\frac{d}{ds}(\gamma \circ \theta)(s)\| ds \\
&= \int_J \|g'(s)\| ds \\
&= l(J, g).
\end{aligned} \tag{21.116}$$

□

Définition 21.40.
Nous nommons ***longueur*** *d'un arc géométrique la longueur commune de tous ses représentants. On dit que l'arc géométrique est* ***rectifiable*** *si sa longueur est $< \infty$.*

6. Théorème 21.9.

21.7.1 Abscisse curviligne et paramétrage normal

SubSecAbsCurv

Définition 21.41.
*Soit (I,γ) un arc paramétré continu rectifiable. Nous appelons **abscisse curviligne** de γ toute application $\phi\colon I\to\mathbb{R}$ telle que pour tout $t,t'\in I$ avec $t<t'$, nous ayons*

$$l\big([t,t'],\gamma\big)=\big|\phi(t')-\phi(t)\big|. \tag{21.117}$$

*Si il existe un $t_0\in I$ tel que $\phi(t_0)=0$, alors nous disons que t_0 est l'**origine** de l'abscisse ϕ.*

DEFooJJQFooEITCvG

Définition 21.42.
*Un arc paramétré (I,γ_N) continu rectifiable est dit **normal** si l'identité est une abscisse curviligne.*

LEMooLADUooBlHjuT

Lemme 21.43.
Si γ est de classe C^1 et est un paramétrage normal, alors

(1) pour tout choix de t et t' dans I avec $t<t'$, nous avons

$$l\big([t,t'],\gamma_N\big)=t'-t. \tag{21.118}$$

(2) $\|\gamma'(t)\|=1$ pour tout t.

Démonstration. Pour tout x_1,x_2 dans le domaine nous avons

$$l\big([x_1,x_2],\gamma\big)=\int_{x_1}^{x_2}\|\gamma'(t)\|dt=x_2-x_1. \tag{21.119}$$

Cela implique $\|\gamma'(t)\|=1$ pour tout t. En effet, pour fixer les idées, supposons que $\|\gamma'(t)\|>1$ en un point, par continuité, cela reste strictement supérieur à 1 sur un intervalle[7]. L'intégrale sur cet intervalle ne peut alors pas être la taille de l'intervalle. □

ExCerlceRadNorm

Exemple 21.44.
Le cercle unitaire est donné par l'arc

$$\gamma(t)=\big(\cos(t),\sin(t)\big) \tag{21.120}$$

et $t\in[0,2\pi]$. Pour tout choix de t et t' dans $[0,2\pi]$, nous avons

$$l\big([t,t'],\gamma\big)=\int_t^{t'}\sqrt{\sin^2(u)+\cos^2(u)}du=t'-t. \tag{21.121}$$

Les angles exprimés en radians forment donc un paramétrage normal du cercle de rayon 1. △

Lemme 21.45.
Pour un arc paramétré compact, la longueur d'arc est une abscisse curviligne.

Démonstration. Par définition de la longueur d'arc φ, nous avons

$$\varphi(t')-\varphi(t)=l\big([a,t'],\gamma\big)-l\big([a,t],\gamma\big)=\diamondsuit. \tag{21.122}$$

Supposons pour fixer les idées que $t'>t$. En utilisant la proposition 21.6, nous avons

$$l\big([a,t'],\gamma\big)=l\big([a,t],\gamma\big)+l\big([t,t'],\gamma\big), \tag{21.123}$$

et donc après simplification de deux termes,

$$\diamondsuit=l\big([t,t'],\gamma\big), \tag{21.124}$$

ce qui est précisément la propriété demandée pour être une abscisse curviligne. □

7. Corolaire 9.90.

PropExisteChmNorm

Proposition 21.46.
Pour tout arc paramétré C^1 sans points critiques, il existe un changement de coordonnées qui rend l'arc normal.

Démonstration. Soit (I,γ) un arc de classe $\mathcal{C}^1$. Nous devons montrer qu'il existe un intervalle J et une application $\theta\colon J\to I$ de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$ tel que l'arc (J,γ_N) soit $\mathcal{C}^1$ où $\gamma_N=\gamma\circ\theta$.

Si $I=[a,b]$, nous considérons la fonction

$$\begin{aligned}\phi\colon I&\to\mathbb{R}^+\\ t&\mapsto\int_a^t\|\gamma'(u)\|du.\end{aligned}\tag{21.125}$$

EqDevVarPhi

Étant définie par l'intégrale d'une fonction $\mathcal{C}^0$, la fonction ϕ est $\mathcal{C}^1$, et nous avons $\phi'(t)=\|\gamma'(t)\|>0$ pour tout $t\in I$. Vue comme application $\phi\colon[a,b]\to[0,l(\gamma)]$, l'application ϕ est bijective et d'inverse $\mathcal{C}^1$. Voyons cela point par point.

(1) La fonction ϕ est injective parce que strictement croissante.

(2) Elle est surjective parce que $\phi(a)=0$ et $\phi(b)=l(\gamma)$.

(3) La continuité de l'inverse est plus délicate. Soit $s\in[0,l(\gamma)]$ et $\varepsilon>0$. Pour prouver la continuité de ϕ^{-1} en s, nous devons trouver un δ tel que

$$|s-s'|<\delta\Rightarrow\left|\phi^{-1}(s)-\phi^{-1}(s')\right|<\varepsilon.\tag{21.126}$$

Étant donné que s et s' sont dans l'image de ϕ, nous considérons les uniques t et t' tels que $s=\phi(t)$ et $s'=\phi(t')$. La quantité $\phi(t)-\phi(t')$ devient

$$\int_a^t\|\gamma'(u)\|du-\int_a^{t'}\|\gamma'(u)\|du=\int_t^{t'}\|\gamma'(u)\|du.\tag{21.127}$$

EqCondypemuCont

D'autre part, $\phi^{-1}(s)=t$ et $\phi^{-1}(s')=t'$, donc la condition (21.127) devient

$$\left|\int_{t'}^t\|\gamma'(u)\|du\right|\leqslant\delta\Rightarrow|t-t'|<\varepsilon.\tag{21.128}$$

Cela revient à la continuité des fonctions définies par une intégrale.

(4) La dérivée de son inverse est donnée par [8]

$$(\phi^{-1})'(s)=\frac{1}{\phi'\big(\phi^{-1}(s)\big)}.\tag{21.129}$$

Nous avons vu que ϕ^{-1} et ϕ' étaient continues. La fonction $(\phi^{-1})'$ étant exprimée en termes de ces deux fonctions elle est également continue.

Nous considérons l'arc paramétré (J,γ_N) avec $J=[0,l(\gamma)]$ et

$$\gamma_N(s)=(\gamma\circ\phi^{-1})(s).\tag{21.130}$$

8. Pour obtenir cette formule, dérivez les deux membres de l'équation $\phi\big(\phi^{-1}(s)\big)=s$.

Nous montrons maintenant que ce nouveau paramétrage est normal. Soient $0 \leqslant s \leqslant s' \leqslant l(\gamma)$,

$$\begin{aligned}
l\big([s,s'],g\big) &= \int_s^{s'} \|\gamma'_N(u)\|du \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma'_N\circ\phi)(t)\|\phi'(t)dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma_N\circ\phi)'(t)\|dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|\gamma'(t)\|dt \\
&= \int_0^{\phi^{-1}(s')} \|\gamma'(t)\|\,dt - \int_0^{\phi^{-1}(s)} \|\gamma'(t)\|\,dt \\
&= \phi\big(\phi^{-1}(s')\big) - \phi\big(\phi^{-1}(s)\big) \\
&= s' - s,
\end{aligned} \tag{21.131}$$

ce qui prouve que le paramétrage (J,γ_N) est normal. □

Nous retenons que le paramétrage normal de γ est donné par (J,γ_N) avec $J=[0,l(\gamma)]$ et

$$\gamma_N(s) = (\gamma\circ\phi^{-1})(s) \tag{21.132}$$

où

$$\begin{aligned}
\phi\colon I &\to \mathbb{R}^+ \\
t &\mapsto \int_a^t \|\gamma'(u)\|du.
\end{aligned} \tag{21.133}$$

Notons aussi que ϕ est une fonction croissante, étant l'intégrale d'une fonction positive.

Exemple 21.47.
Trouvons les coordonnées normales pour la cycloïde donnée par

$$\begin{cases} x(t) = a(t-\sin(t)), & \text{(21.134a)} \\ y(t) = a(1-\cos(t)) & \text{(21.134b)} \end{cases}$$

et $t\in\,]0,2\pi[$. Relire l'exemple 21.31.

D'abord nous trouvons ϕ avec la formule (21.133) avec $a=0$. En utilisant le bout de calcul (21.97), nous avons

$$\phi(t) = 2a\int_0^t \sin\frac{u}{2}du = 4a\left(1-\cos\frac{t}{2}\right). \tag{21.135}$$

Pour trouver $\phi^{-1}(s)$, nous résolvons l'équation

$$s = \phi\big(\phi^{-1}(s)\big) \tag{21.136}$$

par rapport à $\phi^{-1}(s)$. Dans un premier temps, nous trouvons

$$1-\frac{s}{4a} = \cos\frac{\phi^{-1}(s)}{2}, \tag{21.137}$$

donc $\frac{\phi^{-1}(s)}{2} = \arccos(\frac{4a-s}{4a})$, et finalement

$$\phi^{-1}(s) = 2\arccos\left(\frac{4a-s}{4a}\right). \tag{21.138}$$

Il nous reste à injecter cela dans les expressions de $x(t)$ et $y(t)$ pour trouver $(\gamma_N)_x(s)$ et $(\gamma_N)_y(s)$. D'abord,

$$(\gamma_N)_x(s) = a\big[\phi^{-1}(s) - \sin\big(\phi^{-1}(s)\big)\big]. \tag{21.139}$$

Nous utilisons maintenant la formule trigonométrique $\sin(x) = 2\sin\frac{x}{2}\cos\frac{x}{2}$ afin de simplifier les expressions :

$$\begin{aligned}(\gamma_N)_x &= a\Big[2\arccos\left(\frac{4a-s}{4a}\right) - 2\sin\big(\arccos\left(\frac{4a-s}{4a}\right)\big)\cos\big(\arccos\left(\frac{4a-s}{4a}\right)\big)\Big]\\ &= a\Big[2\arccos\left(\frac{4a-s}{4a}\right) - \frac{4a-s}{2a}\sqrt{1-\left(\frac{4a-s}{4a}\right)^2}\Big]\\ &= 2a\arccos\left(\frac{4a-s}{4a}\right) - \sqrt{8as-s^2}\,\frac{4a-s}{8a}\end{aligned} \tag{21.140}$$

où nous avons utilisé la formule $\sin\big(\arccos(x)\big) = \sqrt{1-x^2}$. Ensuite, pour obtenir $(\gamma_N)_y$ nous devons calculer

$$(\gamma_N)_y(s) = a\big[1-\cos\big(\phi^{-1}(s)\big)\big]. \tag{21.141}$$

Encore une fois, il est intéressant d'exprimer le cosinus en termes des angles divisés par deux : $\cos(x) = \cos^2\frac{x}{2} - \sin^2\frac{x}{2}$.

$$\begin{aligned}(\gamma_N)_y &= a\Big[1-\cos^2\frac{\phi^{-1}(s)}{2} + \sin^2\frac{\phi^{-1}(s)}{2}\Big]\\ &= a\Big[2-2\cos^2\frac{\phi^{-1}(s)}{2}\Big]\\ &= 2a\Big[1-\left(\frac{4a-s}{4a}\right)^2\Big].\end{aligned} \tag{21.142}$$

Dans ce paramétrage, $s \in]0, 8a[$. △

Exemple 21.48.
La cardioïde $\rho(\theta) = a\big(1+\cos(\theta)\big)$ avec θ entre $-\pi$ et π. Avant d'utiliser la formule (21.133), nous devons trouver l'élément de longueur de la cardioïde. Étant donné la façon dont l'équation de la cardioïde nous est donnée, l'élément de longueur est donné par [9] (21.90) :

$$\begin{aligned}\|\gamma'(u)\|^2 &= a^2\sin^2(u) + a^2(1+\cos(u))^2\\ &= 2a^2\big(1+\cos(u)\big),\end{aligned} \tag{21.143}$$

et par conséquent [10]

$$\begin{aligned}\phi(t) &= \int_0^t \sqrt{2a^2\big(1+\cos(u)\big)}du\\ &= \int_0^t \sqrt{2a^2\left(1+\cos^2\frac{u}{2}-\sin^2\frac{u}{2}\right)}du\\ &= 2a\int_0^t \cos\frac{u}{2}du\\ &= 4a\sin\frac{t}{2}.\end{aligned} \tag{21.144}$$

Pour trouver l'inverse, nous résolvons $\phi\big(\phi^{-1}(s)\big) = s$ par rapport à $\phi^{-1}(s)$:

$$\begin{aligned}4a\sin\left(\frac{\phi^{-1}(s)}{2}\right) &= s,\\ \phi^{-1}(s) &= 2\arcsin\left(\frac{s}{4a}\right).\end{aligned} \tag{21.145}$$

9. Nous vous déconseillons d'étudier cette formule par cœur. Sachez cependant la retrouver assez vite.
10. L'utilisation stricte de la formule (21.133) demanderait d'intégrer à partir de $-\pi$. Pour plus de simplicité, nous intégrons à partir de zéro, et nous verrons plus tard comment adapter l'intervalle du nouveau paramètre.

Avant d'écrire trop brutalement $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$, il faut comprendre comment est γ. Nous avons reçu la courbe sous forme polaire, c'est-à-dire

$$\gamma(t) = \big(\gamma_r(t), \gamma_\theta(t)\big) = \Big(a\big(1 + \cos(t)\big), t\Big). \tag{21.146}$$

C'est comme cela qu'il faut comprendre la donnée $\rho(\theta) = a\big(1 + \cos(\theta)\big)$. Maintenant la formule $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$ devient

$$\begin{cases} (\gamma_N)_r(s) = \gamma_r\big(\phi^{-1}(s)\big) & \text{(21.147a)} \\ (\gamma_N)_\theta(s) = \gamma_\theta\big(\phi^{-1}(s)\big). & \text{(21.147b)} \end{cases}$$

Étant donné que $\gamma_\theta(t) = t$, la seconde est facile :

$$(\gamma_N)_\theta(s) = 2\arcsin\Big(\frac{s}{4a}\Big). \tag{21.148}$$

Pour la première,

$$(\gamma_N)_r(s) = a\big[1 + \cos\big(2\arcsin\frac{s}{4a}\big)\big] = \frac{16a^2 - s^2}{8a}. \tag{21.149}$$

Nous écrivons donc le nouveau paramétrage en coordonnées polaires sous la forme

$$\Big(\frac{16a^2 - s^2}{8a}, 2\arcsin\frac{s}{4a}\Big). \tag{21.150}$$

La question qui arrive maintenant est de savoir quel intervalle parcours la nouvelle variable s. D'après le résultat de l'exemple 21.99, la longueur de la cardioïde est de $8a$ et nous avons donc $s \in [0, 8a]$. Cependant, la condition d'existence de arcsin nous interdit d'avoir s plus grand que $4a$ en valeur absolue. Où est le problème ?

Le problème est que nous avons changé l'origine de notre paramètre en donnant $\phi(t)$ comme une intégrale à partir de 0 au lieu de $-\pi$. Cela se voit en regardant de quel point nous partons : en $s = 0$ nous sommes sur le point $(2a, 0)$ tandis qu'avec le paramètre original, c'est-à-dire $\theta \in [-\pi, \pi]$, nous avons pour $\theta = -\pi$ le point $(0, -\pi)$.

Il se passe donc que si nous commençons à parcourir la cardioïde avec $s = 0$, nous partons du milieu, et nous ne parcourons donc pas tout. Étant donné que le « premier » point de la cardioïde est le point $(0, -\pi)$, le paramètre s commence en $s = -4a$, et nous avons comme intervalle :

$$s \in [-4a, 4a], \tag{21.151}$$

ce qui est en accord avec la conditions d'existence. △

Quel enseignement tirer de cet exemple ? Lorsqu'on calcule $\phi(t)$ pour trouver les coordonnées normales, il y a deux solutions.

(1) Utiliser strictement la formule $\phi(t) = \int_a^t \|\gamma'(u)\|du$, en prenant bien comme borne de départ le point de départ de le paramétrage de γ. À ce moment la coordonnée normale construite aura $[0, l(\gamma)]$ comme intervalle de variation.

(2) Faire commencer l'intervalle d'intégration en zéro (ou ailleurs). Un bon choix peut simplifier quelques calculs, mais alors il faudra bien choisir la valeur de départ de la nouvelle coordonnées pour que le « premier » point de la courbe soit correct. Dans ce cas, la longueur de l'intervalle sera quand même $l(\gamma)$. Il n'y a donc pas de problèmes pour trouver la valeur du bout de l'intervalle de variation du paramètre normal.

Dans tous les cas, il faut bien préciser l'intervalle de variation du paramètre lorsqu'on donne une courbe paramétrée.

21.7.2 Tangente à une courbe paramétrée

Définition 21.49.
*Soit (I,γ) un arc paramétré de classe $\mathcal{C}^k$ avec $k \geqslant 1$. Nous disons que la courbe admet une **tangente** en $\gamma(t_0) \in \mathbb{R}^n$ lorsque les deux conditions suivantes sont remplies*

(1) $\gamma(t) \neq \gamma(t_0)$ pour tout t dans un voisinage de t_0 ;

(2) la direction de la droite qui passe par $\gamma(t)$ et $\gamma(t_0)$ admet une limite lorsque $t \to t_0$.

Dans ce cas, la tangente sera la droite passant par le point $\gamma(t_0)$ et dont la direction est donnée par la limite.

Dans cette définition, par **direction** d'une droite, nous entendons le vecteur de norme 1 parallèle à celle-ci sans tenir compte du signe. La tangente sera donc la droite passant par $\gamma(t_0)$ et parallèle au vecteur

$$\lim_{t\to t_0} \frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}. \tag{21.152}$$

Évidemment si nous avions écrit $\gamma(t_0)-\gamma(t)$, ça n'aurait pas changé la droite. Par abus de langage, nous parlerons souvent de « la direction u » même lorsque u n'est pas de norme 1.

Formellement, une direction est une classe d'équivalence de vecteurs pour la relation $u \sim v$ si il existe $\lambda \neq 0$ tel que $u = \lambda v$, mais nous n'aurons pas besoin de cette précision ici.

Sans surprises, la tangente est à peu près toujours donnée par la dérivée lorsqu'elle existe. Plus précisément nous avons le

Théorème 21.50.
Soit (I,γ), un arc paramétré de classe $\mathcal{C}^k$ ($k \geqslant 1$) et $t_0 \in I$ tel que

$$\gamma'(t_0) = \gamma''(t_0) = \ldots = \gamma^{(q-1)}(t_0) = 0 \tag{21.153}$$

et

$$\gamma^{(q)}(t_0) \neq 0 \tag{21.154}$$

pour un entier $1 \leqslant q \leqslant k$. Alors γ admet une tangente en $\gamma(t_0)$ de direction $\gamma^{(q)}(t_0)$.

Démonstration. Le développement de $\gamma(t_0)$ en série de Taylor autour de t jusqu'à l'ordre q est

$$\begin{aligned}\gamma(t) = \gamma(t_0) + \gamma'(t_0)|t-t_0| + \frac{\gamma' t(t_0)}{2}|t-t_0|^2 + \cdots + \frac{\gamma^{(q)}(t_0)}{q!}|t-t_0|^q \\ + \varepsilon(t)|t-t_0|^q\end{aligned} \tag{21.155}$$

où ε est une application $\varepsilon\colon \mathbb{R} \to \mathbb{R}^n$ telle que $\lim_{t\to t_0}\varepsilon(t) = 0$. En utilisant les hypothèses, nous éliminons la majorité des termes dans le développement (21.155) :

$$\gamma(t)-\gamma(t_0) = \frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q + \varepsilon(t)|t-t_0|^q. \tag{21.156}$$

La direction de la droite qui joint $\gamma(t)$ à $\gamma(t_0)$ est donc donnée par

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|} = \frac{\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q + \varepsilon(t)|t-t_0|^q}{\|\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q + \varepsilon(t)|t-t_0|^q\|} \tag{21.157}$$

et la limite lorsque $t \to t_0$ donne $\gamma^{(q)}(t_0)$ comme direction de la tangente.

□

Lorsque le théorème s'applique, le vecteur

$$\tau = \frac{\gamma^{(q)}(t_0)}{\|\gamma^{(q)}(t_0)\|} \tag{21.158}$$

est appelé le **vecteur unitaire tangent** en $\gamma(t_0)$ à l'arc paramétré γ.

CorTgSoCun

Corolaire 21.51.
Si (I,γ) est un arc paramétré de classe $\mathcal{C}^1$ régulier (c'est-à-dire $\gamma'(t)\neq 0$ pour tout t) alors l'arc admet une tangente en tout point et le vecteur unitaire de la tangente est donné par

$$\tau(t)=\frac{\gamma'(t)}{\|\gamma'(t)\|},\tag{21.159}$$

pour tout t dans I.

CorUnitTgtaugpnorma

Corolaire 21.52.
Si $\gamma=(J,\gamma_N)$ est un arc paramétré de classe $\mathcal{C}^1$, normal, alors le vecteur unitaire de la tangente au point $\gamma_N(s)$ est donné par $\tau(s)=\gamma'_N(s)$.

Démonstration. Nous devons démontrer que dans le cas d'un paramétrage normal nous avons $\|\gamma'_N(s)\|=1$ pour tout s. Par définition,

$$l\big([s,s'],g\big)=\int_s^{s'}\|\gamma'_N(u)\|du=s'-s.\tag{21.160}$$

Par conséquent,

$$\lim_{h\to 0}\frac{1}{h}\int_s^{s+h}\|\gamma'_N(u)\|du=\lim_{y\to 0}\frac{s+h-s}{h}=1.\tag{21.161}$$

Cela implique que $\|\gamma'_N(s)\|=1$, et donc en particulier que (J,γ_N) est un arc régulier. Le corolaire précédent montre alors que $\tau(s)=\gamma'_N(s)/\|\gamma'_N(s)\|=\gamma'_N(s)$. □

Exemple 21.53.
Considérons la courbe $\gamma(t)=(t^2,t^3)$, et cherchons la tangente en $t_0=0$. En dérivant nous avons successivement

$$\begin{aligned}\gamma(t)&=(t^2,t^3)\\ \gamma'(t)&=(2t,3t^2)\\ \gamma''(t)&=(2,6t).\end{aligned}\tag{21.162}$$

En posant $t=0$, nous trouvons que $\gamma'(0)=0$ mais $\gamma''(0)=(2,0)\neq 0$. Le théorème nous dit donc que la direction de la tangente est horizontale. Nous pouvons faire le calcul directement :

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}=\frac{(t^2,t^3)}{\sqrt{t^4+t^6}}=\frac{(t^2,t^3)}{t^2\sqrt{1+t^2}}=\frac{(1,t)}{\sqrt{1+t^2}},\tag{21.163}$$

dont la limite $t\to 0$ est bien le vecteur horizontal $(1,0)$.

La figure 21.7 montre quelques tangentes, c'est-à-dire quelques vecteurs dans la direction $\gamma'(t)$ (pour les $t\neq 0$, il ne faut pas aller à la dérivée seconde). Nous remarquons que de part et d'autres du sommet, les vecteurs ne sont pas dirigés dans le même sens. *En tant que vecteurs* de norme 1, ces vecteurs n'ont pas de limites quand $t\to 0$. Ce sont bien les *directions* qui ont une limite, parce que la direction ne tient pas compte du sens.

△

21.8 Un peu de topologie

La proposition 21.54 donne une sorte de théorème des valeurs intermédiaires pour le cas d'une application à valeurs dans un chemin.

PROPooJYGVooShNewy

Proposition 21.54 ([1]).
Soit une application continue et injective $\gamma\colon[0,1]\to\mathbb{R}^n$. Nous posons $\Gamma=\gamma\big([0,1]\big)$ sur lequel nous considérons la topologie induite[11] *de $\mathbb{R}^n$.*

11. Définition 7.26.

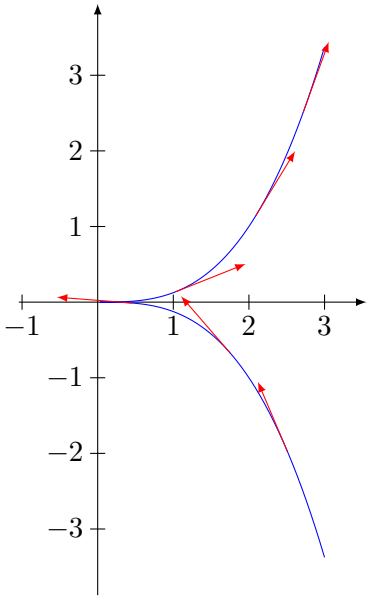


FIGURE 21.7: Quelques tangentes de la courbe $\gamma(t) = (t^2, t^3)$. LabelFigParamTangente

Nous supposons que $\gamma^{-1}\colon \Gamma \to [0,1]$ est continue[12]
Nous considérons un chemin $\alpha\colon [0,1] \to \mathbb{R}^n$ tel que

(1) α *est continu,*
(2) $\alpha(0) = \gamma(0)$
(3) $\alpha\big([0,1]\big) \subset \Gamma$
(4) $\alpha(1) = \gamma(t_0)$ *pour un certain* $t_0 \in [0,1]$.

Alors pour tout $t \in [0, t_0]$, il existe $u \in [0,1]$ tel que $\alpha(u) = \gamma(t)$.

Démonstration. Nous commençons par montrer que l'application $\alpha\colon [0,1] \to \Gamma$ est encore continue lorsque nous voyons bien l'espace d'arrivée comme Γ muni de sa propre topologie et non comme $\mathbb{R}^n$ muni de sa topologie usuelle.

Soit un ouvert $\mathcal{O}$ de Γ ; il existe un ouvert $\mathcal{O}'$ de $\mathbb{R}^n$ tel que $\mathcal{O} = \mathcal{O}' \cap \Gamma$. Vu que α ne prend ses valeurs que dans Γ, nous avons $\alpha^{-1}(\Gamma \cap \mathcal{O}') = \alpha^{-1}(\mathcal{O}')$ et comme α est continue pour la topologie de $\mathbb{R}^n$, la partie $\alpha^{-1}(\mathcal{O}')$ est un ouvert de $[0,1]$.

Nous considérons donc l'application

$$\gamma^{-1} \circ \alpha\colon [0,1] \to [0,1] \tag{21.164}$$

qui est continue et vérifie donc le théorème des valeurs intermédiaires 9.93. Les hypothèses $\alpha(0) = \gamma(0)$ et $\alpha(1) = \gamma(t_0)$ donnent

$$(\gamma^{-1} \circ \alpha)(0) = 0 \tag{21.165a}$$
$$(\gamma^{-1} \circ \alpha)(1) = t_0. \tag{21.165b}$$

Donc pour tout $t \in [0, t_0]$, il existe $u \in [0,1]$ tel que $(\gamma^{-1} \circ \alpha)(u) = t$. En appliquant γ des deux côtés, nous voyons que ce u vérifie $\alpha(u) = \gamma(t)$ comme demandé. □

La proposition suivante dit essentiellement que la longueur d'un chemin est minoré par la longueur de son graphe.

PROPooXENVooMvkTZW

Proposition 21.55 ([1])**.**
Soient $a, b \in \mathbb{R}^n$. Soit une application continue et injective $\gamma\colon [0,1] \to \mathbb{R}^n$ telle que $\gamma(0) = a$ et $\gamma(1) = b$.

12. Je ne suis pas certain que cette hypothèse soit indispensable.

Nous considérons maintenant un second chemin continu $\alpha\colon [0,1] \to \mathbb{R}^n$ *tel que* $\alpha(0) = a$, $\alpha(1) = b$ *et* $\alpha([0,1]) \subset \gamma([0,1])$.

Alors la longueur de γ *minore celle de* α *:* $l(\gamma) \leqslant l(\alpha)$.

Démonstration. Pour rappel, la longueur d'un chemin est donné par le supremum des longueurs des lignes brisées reliant des points du chemin (définition 21.3). Soit une subdivision σ de $[0,1]$ et la « longueur » correspondante pour γ :

$$l_\sigma(\gamma) = \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\|. \tag{21.166}$$

Nous allons construite une subdivision de $[0,1]$ pour laquelle la longueur de α sera la même. De cette façon, le supremum pour α sera un supremum sur un ensemble plus grand que celui du supremum pour γ.

Vu que le graphe de α est contenu dans celui de γ, et que les points de départ et d'arrivée sont les mêmes, la proposition 21.54 nous donne que pour tout $t \in [0,1]$, il existe $u \in [0,1]$ tel que $\alpha(u) = \gamma(t)$.

Cela ne suffit pas à considérer une subdivision $(u_i)_{i=1,\dots,n}$ de $[0,1]$ pour laquelle $\alpha(u_i) = \gamma(t_i)$ parce qu'il faut encore que les u_i soient ordonnés. Vu que α n'est pas injective, il n'y a pas de garanties de ce côté. Nous posons :

$$u_i = \min\{x \in [0,1] \text{ tel que } \alpha(x) = \gamma(t_i)\}. \tag{21.167}$$

Montrons que ce minimum existe. Ce minimum existe parce que α étant continue, l'ensemble des x sur lesquels $\alpha(x)$ a une valeur donnée est fermé (le complémentaire est ouvert par le théorème des valeurs intermédiaires). De plus cet ensemble est borné parce qu'il est inclus dans $[0,1]$. Il est donc compact et possède un minimum.

Nous prouvons à présent que $u_{i+1} \geqslant u_i$. Nous avons $\alpha(u_{i+1}) = \gamma(t_{i+1})$ et il existe u dans $[0, u_{i+1}[$ tel que$\alpha(u) = \gamma(t_i)$ (proposition 21.54). Vu que u_i est le minimum de tels u, nous avons $u_i \in [0, u_{i+1}[$. Cela prouve bien que $u_i < u_{i+1}$.

Ces $(u_i)_{i=1,\dots,n}$ forment une subdivision de $[0,1]$ telle que $\alpha(u_i) = \gamma(t_i)$. La longueur associée à la subdivision (u_i) pour α est la même que celle associée à (t_i) pour γ.

Donc $l(\alpha) \geqslant l(\gamma)$. □

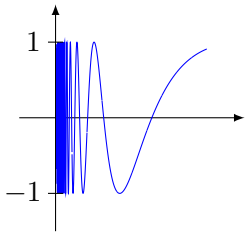


FIGURE 21.8: La figure de la proposition 21.56.

LabelFigYWxOAkh

PROPooVXDNooPZYKPr

Proposition 21.56 ([1])**.**
Nous considérons la partie suivante de $\mathbb{R}^2$ *:*

$$A_2 = \big\{\big(t, \sin(1/t)\big)\big\}_{t\in]0,1]} \tag{21.168}$$

qui est dessinée est sur la figure 21.8. Ensuite nous posons

$$A = \big\{(0,0)\big\} \cup A_2. \tag{21.169}$$

La partie A *est connexe, mais pas connexe par arcs.*

Démonstration. En plusieurs points.

(1) **A_2 est connexe** La partie A_2 est l'image de la fonction

$$\begin{aligned} f\colon]0,1] &\to \mathbb{R}^2 \\ t &\mapsto \big(t, \sin(1/t)\big). \end{aligned} \tag{21.170}$$

Vu que f est continue et que son ensemble de départ est connexe, A_2 est connexe (lemme 7.218).

(2) **A est connexe** Soient deux ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ dont l'union contient A. Nous supposons que $(0,0) \in \mathcal{O}_1$. Si $\mathcal{O}_1$ contient $B\big((0,0),r\big)$, alors il contient tous les points de la forme $\big(\frac{1}{2k\pi},0\big)$ pour k assez grand. Ces points sont dans A_2.
Vu que $\mathcal{O}_1$ contient des points de A_2, il doit contenir tous les points de A_2 ; sinon les ouverts $\mathcal{O}_1$ et $\mathcal{O}_2$ contrediraient la connexité de A_2. Finalement, $A \subset \mathcal{O}_1$ et A est connexe.

(3) **Connexité par arcs : le chemin** Pour faire aller le chemin dans le bon sens, nous considérons

$$\begin{aligned} \gamma\colon [0,1[&\to \mathbb{R}^2 \\ t &\mapsto f(1-t). \end{aligned} \tag{21.171}$$

Si A était connexe par arcs, il existerait une application continue $\alpha\colon [0,1] \to A$ telle que $\alpha(0) = \big(1,\sin(1)\big) = \gamma(0)$ et $\alpha(1) = (0,0)$.

(4) **Minoration de la longueur de α** Vu que α est continue, ses composantes le sont et vérifient le théorème des valeurs intermédiaires. Pour tout $t \in]0,1[$, il existe $u \in]0,1[$ tel que $\alpha(u) = \gamma(t)$.
Par la proposition 21.55, la longueur de α est minorée par la longueur du chemin γ entre 0 et t_0 pour tout t.

(5) **La longueur de γ** Vu que γ passe une infinité de fois par zéro et par 1, il est possible de construire une subdivision de $]0,1[$ par rapport à la quelle la longueur de γ est arbitrairement grande.

(6) **Conclusion** Donc la longueur de α est minoré par tous les nombres, arbitrairement grand. Autrement dit, la longueur de α est infinie. Mais α étant un chemin continu depuis un compact, cela est impossible.
Nous déduisons qu'un chemin continu liant $\big(1,\sin(1)\big)$ à $(0,0)$ en restant sur A est impossible.

□

21.9 Repère de Frenet

SecFrenet

Dans cette section, nous ne considérons que des courbes dans $\mathbb{R}^3$.

Proptausclataupzero

Proposition 21.57.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$. Alors pour toute valeur de s dans J, nous avons

$$\tau(s) \cdot \tau'(s) = 0 \tag{21.172}$$

où $\tau(s) = \gamma_N'(s)$. C'est-à-dire que la dérivée seconde est perpendiculaire à la dérivée première.

Démonstration. Le paramétrage étant normal, nous avons

$$\|\gamma_N'(s)\|^2 = \sum_{i=1}^{n} x_i'(s)^2 = 1; \tag{21.173}$$

ce qui implique, en dérivant les deux membres, que

$$0 = 2\sum_{i=1}^{n} x_i'(s)x_i''(s), \tag{21.174}$$

c'est-à-dire exactement $\gamma_N'(s) \cdot \gamma_N''(s) = 0$; d'où la thèse. □

Remarque 21.58.
Si nous n'utilisons pas des coordonnées normales, la proposition 21.57 n'est pas spécialement vraie. Prenons par exemple la courbe qui donne la parabole :

$$\gamma(t) = (t, t^2) \tag{21.175a}$$
$$\gamma'(t) = (1, 2t) \tag{21.175b}$$
$$\gamma''(t) = (0, 2) \tag{21.175c}$$

Nous avons $\gamma'(t) \cdot \gamma''(t) = 4t$. Par conséquent, la dérivée seconde n'est la normale à la courbe que en $t = 0$. Cela est une propriété très intéressante des coordonnées normales : la dérivée seconde d'une coordonnées normale donne un vecteur normal à la courbe, c'est-à-dire perpendiculaire à la tangente.

DefCourbureNormleUnit

Définition 21.59.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$.

*(1) Le **vecteur unitaire tangent** est donné par le corolaire 21.51 : $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$.*

*(2) La **normale principale** est le vecteur $\tau'(s)$. Le **vecteur unitaire normal** est le vecteur*

$$\nu(s) = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}. \tag{21.176}$$

Nous déduirons une formule plus pratique en dehors des coordonnées normales en (21.204).

*(3) La **courbure** au point $\gamma_N(s)$ est le réel*

$$c(s) = \|\tau'(s)\| = \|\gamma_N''(s)\|. \tag{21.177}$$

Note : il y a une notion de courbure signée qui sera donnée dans la définition 21.73.

*(4) Le **rayon de courbure** est le réel*

$$R(s) = \frac{1}{c(s)} = \frac{1}{\|\gamma_N''(s)\|}. \tag{21.178}$$

Par la proposition 21.57, nous avons $\nu(s) \cdot \tau(s) = 0$. En combinant toutes les formules, nous avons les différentes expressions suivantes pour le vecteur normal unitaire :

$$\nu(s) = \frac{\gamma_N''(s)}{c(s)} = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\tau'(s)}{c(s)} = R(s)\tau'(s) = R(s)\gamma_N''(s). \tag{21.179}$$

Eq0908mfractauRc

Proposition 21.60.
La fonction courbure s'écrit $c = \|\gamma_N' \times \gamma_N''\|$.

Démonstration. Par la proposition 10.35 nous avons :

$$\langle \gamma_N', \gamma_N'' \rangle^2 + \|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N'\|^2 \|\gamma_N''\|^2 = \|\gamma_N''\|^2 \tag{21.180}$$

parce que, le paramétrage étant normal, $\|\gamma_N'\| = 1$. Mais $\langle \gamma_N', \gamma_N'' \rangle = 0$, donc il reste $\|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N''\|^2$, d'où

$$c(s) = \|\gamma_N''(s)\| = \|\gamma_N'(s) \times \gamma_N''(s)\| \tag{21.181}$$

Eqcsnormgpgpps

pour chaque s dans J. □

Définition 21.61.
*Soit s un point birégulier (c'est-à-dire $\gamma_N'(s) \neq 0$ et $\gamma_N''(s) \neq 0$) de l'arc normal $\gamma = (J, \gamma_N)$. Le **vecteur unitaire de la binormale** est le vecteur*

$$\beta(s) = \tau(s) \times \nu(s) \tag{21.182}$$

Par leurs définitions, τ et ν sont unitaires, tandis que la proposition 21.57 montre qu'ils sont également orthogonaux. Les propriétés du produit vectoriel font que β est également unitaire, et simultanément orthogonal à τ et à ν.

Définition 21.62.
Le repère orthonormal $\{\tau(s), \nu(s), \beta(s)\}$ est le ***repère de Frenet*** *au point $\gamma_N(s)$.*

Lemme 21.63.
Le vecteur unitaire normal est donné par $\nu(s) = \beta(s) \times \tau(s)$.

Démonstration. Ceci est une application de la formule d'expulsion (10.95) et de l'orthonormalité de la base de Frenet :

$$\beta \times \tau = (\tau \times \nu) \times \tau = -(\nu \cdot \tau)\tau + (\tau \cdot \tau)\nu = \nu. \tag{21.183}$$

□

21.9.1 Torsion

Décomposons le vecteur $\beta'(s)$ dans la base de Frenet. Pour cela nous allons utiliser la proposition 8.174 et montrer que $\beta'(s) \cdot \tau(s) = \beta'(s) \cdot \beta(s) = 0$, ce qui voudra dire que, dans la base de Frenet, les composantes de β' le long de τ et β sont nulles. Le vecteur β' sera donc colinéaire à ν.

D'abord, étant donné que la norme de $\beta(s)$ est constante par rapport à s, nous avons

$$0 = \frac{d}{ds}\|\beta(s)\|^2 = 2\beta'(s) \cdot \beta(s). \tag{21.184}$$

Ensuite, nous dérivons la définition $\beta(s) = \tau(s) \times \nu(s)$ en utilisant la formule de Leibniz (12.425) :

$$\beta'(s) = \tau'(s) \times \nu(s) + \tau(s) \times \nu'(s). \tag{21.185}$$

Mais $\tau'(s) = \gamma_N''(s)$ tandis que $\nu(s) = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}$, de telle sorte que $\tau'(s) \times \nu(s) = 0$. Nous restons donc avec $\beta'(s) = \tau(s) \times \nu'(s)$, ce qui prouve que $\beta'(s)$ est perpendiculaire à $\tau(s)$ et donc que $\beta'(s) \cdot \tau(s) = 0$.

Le vecteur $\beta'(s)$ est donc un multiple de $\nu(s)$. Nous notons $t(s)$ le facteur de proportionnalité :

$$\beta'(s) = t(s)\nu(s). \tag{21.186}$$

DefTorsion

Définition 21.64.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^3$. La ***torsion*** *de γ au point $\gamma_N(s)$ est le réel*

$$t(s) = \|\beta'(s)\| = \|\tau(s) \times \nu'(s)\|. \tag{21.187}$$

Lorsque $t(s) \neq 0$, le réel $T(s) = \frac{1}{t(s)}$ est le ***rayon de torsion*** *de γ en $\gamma_N(s)$.*

Étant donné que pour chaque s, l'ensemble $\{\tau(s), \nu(s), \beta(s)\}$ est une base, il est naturel de vouloir décomposer leurs dérivées dans cette base. D'abord, par définition de c et de t, nous avons

$$\begin{aligned} \tau'(s) &= c(s)\nu(s) \\ \beta'(s) &= t(s)\nu(s). \end{aligned} \tag{21.188}$$

Il reste à décomposer $\nu'(s)$. Définissons α_τ, α_ν et α_β (qui peuvent dépendre de s) par

$$\nu'(s) = \alpha_\tau\tau(s) + \alpha_\nu\nu(s) + \alpha_\beta\beta(s). \tag{21.189}$$

En vertu de la proposition 8.174, nous avons

$$\begin{aligned} \alpha_\tau &= \langle \nu'(s), \tau(s) \rangle = -\langle \nu(s), \tau'(s) \rangle = -\langle \nu(s), c(s)\nu(s) \rangle = -c(s), \\ \alpha_\nu &= \langle \nu'(s), \nu(s) \rangle = 0, \\ \alpha_\beta &= \langle \nu'(s), \beta(s) \rangle = -\langle \nu(s), \beta'(s) \rangle = -t(s), \end{aligned} \tag{21.190}$$

où nous avons utilisé le fait que $\langle \nu(s), \nu(s) \rangle = \|\nu(s)\|^2 = 1$. Si nous mettons ces résultats sous forme matricielle, nous avons les **formules de Frenet** :

$$\begin{pmatrix} \tau'(s) \\ \nu'(s) \\ \beta'(s) \end{pmatrix} = \begin{pmatrix} 0 & c(s) & 0 \\ -c(s) & 0 & -t(s) \\ 0 & t(s) & 0 \end{pmatrix} \begin{pmatrix} \tau(s) \\ \nu(s) \\ \beta(s) \end{pmatrix}. \tag{21.191}$$

Proposition 21.65.
Si s est un point birégulier, alors la torsion est donnée par

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \times \gamma_N'''}{\|\gamma_N'(s) \times \gamma_N''(s)\|^2}. \tag{21.192}$$

Démonstration. Par l'équation (21.179), nous avons $\gamma_N''(s) = c(s)\nu(s)$, et par conséquent

$$\gamma_N'''(s) = c'(s)\nu(s) + c(s)\nu'(s) = c'(s)\nu(s) + c(s)\big[-c(s)\tau(s) - t(s)\beta(s)\big], \tag{21.193}$$

où nous avons utilisé la formule de Frenet pour $\nu'(s)$. Par ailleurs, sachant le corolaire 21.52 et la formule de Frenet pour τ', nous avons

$$\gamma_N' \times \gamma_N'' = \tau(s) \times \tau'(s) = \tau(s) \times c(s)\nu(s) = c(s)\beta(s). \tag{21.194}$$

En combinant les deux dernières équations, et en se souvenant que la base de Frenet et orthonormale,

$$(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''(s) = -c(s)^2 t(s), \tag{21.195}$$

et donc, en remplaçant $c(s)$ par la formule (21.181),

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''}{\|\gamma_N' \times \gamma_N''\|^2}. \tag{21.196}$$

□

21.10 Hors des coordonnées normales

Remfougnormoupad

Remarque 21.66.
Notons que la définition de τ est donnée pour tout arc $\mathcal{C}^1$ régulier (I, γ) par $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$. La propriété $\tau = \gamma_N'$ n'est valable que lorsque le paramétrage est normal. Les autres définitions ont toutes été données dans le cas d'un paramétrage normal.

La remarque 21.66 nous incite à exprimer toute la base de Frenet en termes de γ lorsque le paramétrage n'est pas normal. Étant donné que nous pouvons toujours faire le changement de variable $\gamma(t) = \gamma_N\big(\phi(t)\big)$ (proposition 21.46), il est possible d'exprimer les vecteurs τ, ν et β ainsi que les réels c et t en fonction de γ et de ses dérivées.

Nous allons maintenant travailler à écrire les formules.

Pour plus de facilité, nous collectons les définitions. Afin d'alléger la notation, nous n'exprimons pas explicitement les dépendances en s :

Vecteur unitaire tangent Par le corolaire 21.52, τ est donné par $\tau = \gamma_N'$.

Vecteur unitaire normal Par la définition 21.59, ν est donné par $\nu = \frac{\tau'}{\|\tau'\|}$.

Vecteur unitaire de la binormale Par la définition 21.59, β est donné par $\beta = \tau \times \nu$.

Courbure Par la définition 21.59, c est donné par $c = \|\tau'\|$.

Torsion Par la définition 21.64, t est donné par $t = \|\beta'\|$.

Le schéma du changement de variable est

$$\begin{array}{ccc} t \in I & \xrightarrow{\gamma} & \mathbb{R}^3 \\ \phi \downarrow & \nearrow_{\gamma_N} & \\ s \in J & & \end{array} \tag{21.197}$$

La difficulté ne sera pas d'éliminer γ_N de toutes les formules, mais bien de se débarrasser des fonctions ϕ qui arrivent quand nous exprimons γ_N en termes de γ, et en particulier lorsque nous voulons exprimer les dérivées de γ_N en termes de γ et de ses dérivées.

Regardons d'abord comment les dérivées de γ_N s'expriment en termes de γ. En utilisant le fait que $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$ et que $\|\gamma_N'(s)\| = 1$, nous avons

$$\gamma_N'(s) = \frac{\gamma_N'(s)}{\|\gamma_N'(s)\|} = \frac{(\gamma \circ \phi^{-1})'(s)}{\|(\gamma \circ \phi^{-1})'(s)\|} = \frac{\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\||(\phi^{-1})'(s)|} = \frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{21.198}$$

où nous avons utilisé le fait que ϕ^{-1} étant croissante (parce que l'inverse d'une fonction croissante est croissante), $(\phi^{-1})'(s) = |(\phi^{-1})'(s)|$. Pourquoi écrivons nous $|\phi^{-1}(s)|$ et non $\|\phi^{-1}(s)\|$?

Pour la dérivée seconde, nous dérivons la relation (21.198) :

$$\gamma_N''(s) = \frac{\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\|} + \gamma'\big(\phi^{-1}(s)\big)\frac{d}{ds}\Big[\|\gamma'\big(\phi^{-1}(s)\big)\|\Big]. \tag{21.199}$$

Le petit calcul suivant va nous permettre de simplifier cette expression :

$$(\phi^{-1})'(s) = (\phi^{-1})'\big(\phi(t)\big) = \frac{1}{\phi'(t)} = \frac{1}{\|\gamma'(t)\|}. \tag{21.200}$$

Donc

$$\gamma_N''(s) = \frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big] \tag{21.201}$$

où il est entendu que $t = \phi^{-1}(s)$. Avec cette expression, nous ne nous sommes pas encore débarrassés de la fonction ϕ, mais nous allons voir que cela nous sera suffisant.

Pour le vecteur unitaire tangent $\tau(s)$, nous avons donc immédiatement

$$\tau(s) = \gamma_N'(s) = \frac{\gamma'(t)}{\|\gamma'(t)\|}. \tag{21.202}$$

Ici encore il est sous-entendu que le t dans le membre de droite est lié au s du membre de gauche par $t = \phi^{-1}(s)$. Il est donc naturel de nous demander si nous avons gagné quelque chose, étant donné que la formule (21.202) contient encore la fonction ϕ.

Géométriquement, le vecteur $\tau(s)$ est le vecteur normal unitaire de la courbe au point $\gamma_N(s)$. En utilisant les relations du diagramme (21.197), nous avons en réalité $\gamma_N(s) = \gamma_N\big(\phi(t)\big) = \gamma(t)$. Le vecteur $\frac{\gamma'(t)}{\|\gamma'(t)\|}$ représente donc le vecteur normal tangent au point $\gamma(t)$.

Pour calculer la courbure, nous devons d'abord calculer le produit vectoriel

$$\begin{aligned} \gamma_N'(s) \times \gamma_N''(s) &= \frac{\gamma'(t)}{\|\gamma'(t)\|} \times \left(\frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big] \right) \\ &= \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \end{aligned} \tag{21.203}$$

parce que le deuxième terme dans la parenthèse est un multiple de $\gamma'(t)$, de telle sorte que son produit vectoriel avec $\gamma'(t)/\|\gamma'(t)\|$ soit nul. En prenant la norme,

$$c(s) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}. \tag{21.204}$$

Encore une fois, cette équation nous enseigne que la courbure au point $\gamma(t) \in \mathbb{R}^3$ est donnée par le membre de droite, qui ne dépend que de t.

Le vecteur unitaire binormal est donné par $\beta(s) = \tau(s) \times \nu(s)$. En utilisant (21.202) et (21.179),

$$\beta(s) = \tau(s) \times \nu(s) = \gamma_N'(s) \times \frac{\gamma_N''(s)}{c(s)}. \tag{21.205}$$

Les formules (21.203) pour le produit vectoriel et (21.204) pour la courbure donnent ensuite

$$\beta(s) = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \cdot \frac{1}{c(s)} = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{21.206}$$

Cela donne le vecteur unitaire binormal au point $\gamma(t)$ en termes de $\gamma'(t)$ et $\gamma''(t)$.

La torsion demande d'utiliser la dérivée troisième de γ_N. Nous avons

$$\begin{aligned}
\gamma_N'''(s) &= (\gamma \circ \phi^{-1})'''(s) \\
&= \Big(\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)\Big)'' \\
&= \Big(\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^2 + \gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})''(s)\Big)' \\
&= \gamma'''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^3 + v \\
&= \frac{\gamma'''\big(\phi^{-1}(s)\big)}{\|\gamma'(t)\|^3} + v \qquad \text{par (21.200)}
\end{aligned} \tag{21.207}$$

où v est un élément de $\operatorname{Span}\{\gamma''\big(\phi^{-1}(s)\big), \gamma'\big(\phi^{-1}(s)\big)\}$. Le vecteur v est donc perpendiculaire à $\gamma' \times \gamma''$ et donc à $\gamma_N' \times \gamma_N''$ à cause de la relation (21.203) qui montre que $\gamma' \times \gamma''$ est parallèle à $\gamma_N' \times \gamma_N''$. De ce fait, lorsque nous calculons $(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''$, la partie v de γ_N''' n'entre pas en ligne de compte.

Nous avons donc le calcul suivant, en remplaçant les diverses occurrences de $\gamma_N' \times \gamma_N''$ par sa valeur (21.203) en termes de γ,

$$\begin{aligned}
t(s) &= -\frac{(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''}{\|\gamma_N' \times \gamma_N''\|^2} \\
&= -\frac{(\gamma_N' \times \gamma_N'') \cdot \gamma'''(t)}{\|\gamma_N' \times \gamma_N''\|^2 \|\gamma'(t)\|^2} \\
&= -\frac{(\gamma' \times \gamma'') \cdot \gamma'''}{\|\gamma' \times \gamma''\|^2}.
\end{aligned} \tag{21.208}$$

Dans cette expression, il est sous-entendu que tous les γ_N sont fonctions de s et tous les γ sont fonction de t où s et t sont liés par $s = \phi(t)$.

Voici un théorème qui résume ce que nous venons de faire.

THOooXOVRooRXtEuU

Théorème 21.67.
Pour tout représentant (I, γ), les éléments métriques (τ, ν, β, c, t) au point $\gamma(t)$ s'expriment en fonction de $\gamma(t)$, $\gamma'(t)$, $\gamma''(t)$ et $\gamma'''(t)$.

Lemme 21.68.
Si γ est le graphe de la fonction y alors la courbure de γ est donnée par la formule

$$c\big(\gamma(t)\big) = \frac{|y''(t)|}{\big(1 + y'(t)^2\big)^{3/2}} \tag{21.209}$$

Démonstration. Nous avons :

$$\gamma(t) = \big(t, y(t)\big) \tag{21.210a}$$
$$\gamma'(t) = \big(1, y'(t)\big) \tag{21.210b}$$
$$\gamma''(t) = \big(0, y''(t)\big). \tag{21.210c}$$

Il s'agit maintenant seulement d'utiliser la formule (21.204) en se souvenant comment on calcule un produit vectoriel [13] :

$$\gamma' \times \gamma'' = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & y' & 0 \\ 0 & y'' & 0 \end{vmatrix} = y''. \tag{21.211}$$

□

21.11 Tracer des courbes paramétriques dans $\mathbb{R}^2$

SecTracerParmCourbe

Nous allons maintenant voir comment les concepts introduits nous aident à effectivement tracer des courbes dans le plan. Les courbes que nous regardons sont de la forme $\gamma(t) = \big(x(t), y(t)\big)$, et nous supposons que ces fonctions soient suffisamment régulières (disons trois fois continument dérivables). Nous ne supposons pas que la courbe soit donnée en coordonnées normales, en particulier, $\gamma''(t)$ n'est pas le vecteur normal en $\gamma(t)$.

La notion clef qui va jouer est le **cercle osculateur** de la courbe γ au point $\gamma(t)$. Sans rentrer dans les détails, disons que c'est le cercle qui « colle » le mieux possible la courbe. Le rayon de ce cercle est le rayon de courbure :

$$R(t) = \frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{21.212}$$

En pratique, le produit vectoriel se calcule comme ceci :

$$\gamma'(t) \times \gamma''(t) = \begin{vmatrix} e_x & e_y & e_z \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{vmatrix} = (x'y'' - x''y')e_z. \tag{21.213}$$

Le centre du cercle osculateur va se trouver quelque part sur la normale. Le vecteur normal est donné par

$$n(t) = J\frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{21.214}$$

où J est la rotation d'angle $\frac{\pi}{2}$:

$$J\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix}. \tag{21.215}$$

Cela nous laisse deux possibilités pour le centre du cercle osculateur : $\gamma(t) + R(t)n(t)$ ou bien $\gamma(t) - R(t)n(t)$. Il faut savoir de quel côté de la courbe est situé le centre du cercle osculateur. Il faut choisir le côté de la concavité, c'est-à-dire le côté de la dérivée seconde.

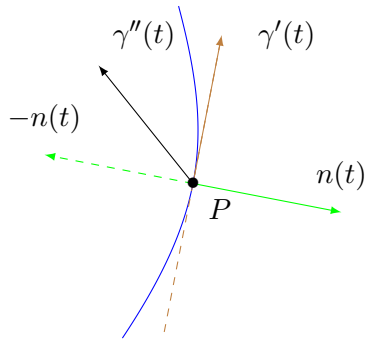


FIGURE 21.9: De quel côté de $\gamma'(t)$ se trouvent $n(t)$ et $-n(t)$? LabelFigQuelCote

La difficulté maintenant est de savoir qui de $n(t)$ ou $-n(t)$ est du côté de $\gamma''(t)$. Il faut savoir si $n(t)$ est du même côté de la droite tangente que $\gamma''(t)$ ou non. Par construction, si nous regardons

13. Définition 10.28.

la figure 21.9, le vecteur $n(t)$ sera toujours à gauche de $\gamma'(t)$. Le fait que $\gamma''(t)$ soit à gauche ou à droite de $\gamma'(t)$ est donné par le signe du produit vectoriel $\gamma'(t) \times \gamma''(t)$. Si ce produit vectoriel est positif, il faut choisir $-n(t)$ et si il est négatif, il faut choisir $n'(t)$.

Le truc pour obtenir le signe de $x'y'' - x''y'$ est de faire

$$\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|}. \tag{21.216}$$

Le centre de courbure sera donc situé à la position

$$\Omega(t) = \gamma(t) - n(t)\frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|^2}(\gamma' \times \gamma'') \cdot e_z \tag{21.217}$$

Nous pouvons écrire cela plus explicitement en nous souvenant que $\gamma' \times \gamma'' = (x'y'' - x''y')e_z$, par conséquent $\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|^2} = \frac{1}{x'y'' - x''y'}$. Nous avons

$$\Omega_x(t) = x(t) - y'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'} \tag{21.218a}$$

$$\Omega_y(t) = y(t) + x'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'}. \tag{21.218b}$$

Quelques exemples de cercles osculateurs sont sur la figure 21.10.

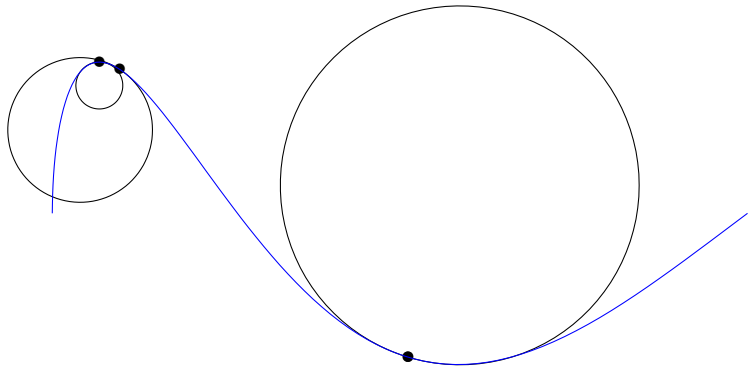


FIGURE 21.10: Exemple de cercles osculateurs. LabelFigOsculateur

21.12 Courbes planes

Définition 21.69.
*Une courbe $\gamma\colon [a,b] \to \mathbb{R}^n$ est **fermée** si $\gamma(a) = \gamma(b)$. Elle est **simple** si $\gamma(t) \neq \gamma(t')$ dès que $t, t' \in]a,b[$ et $t \neq t'$.*

DEFooSAZTooZGQrQG

Définition 21.70.
Nous disons qu'une courbe fermée est continue, de classe C^1, de classe C^2 ou autre condition de régularité si son extension périodique comme application $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ a cette régularité.

21.12.1 Angle

LEMooUECMooNBDGiR

Lemme 21.71 ([543]).
Soient des courbes régulières γ et σ de classe C^2 de l'intervalle ouvert I vers $\mathbb{R}^2$. Soit $\theta_0 \in \mathbb{R}$ tel que

$$\frac{\gamma'(t_0) \cdot \sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \cos(\theta_0) \tag{21.219a}$$

$$\frac{\gamma'(t_0) \cdot J\sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \sin(\theta_0). \tag{21.219b}$$

Alors il existe une unique fonction différentiable $\theta\colon I\to\mathbb{R}$

$$\frac{\gamma'(t)\cdot\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|}=\cos\big(\theta(t)\big) \tag{21.220a}$$
$$\frac{\gamma'(t)\cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|}=\sin\big(\theta(t)\big). \tag{21.220b}$$

Démonstration. Il suffit de prendre

$$f(t)=\frac{\gamma'(t)\cdot\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{21.221}$$

et

$$g(t)=\frac{\gamma'(t)\cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{21.222}$$

dans la proposition 18.57. Ces courbes sont de classe C^1 parce que γ et σ sont de classe C^2. □

21.12.2 Courbure signée

Nous avons déjà défini la courbure d'une courbe en la définition 21.59. Nous introduisons maintenant la courbure signée qui est propre à la dimension deux.

DEFooTSJXooTIyRXf

Définition 21.72.
*La **structure complexe** sur $\mathbb{R}^2$ est l'application*

$$\begin{aligned}J\colon\mathbb{R}^2&\to\mathbb{R}^2\\(x,y)&\mapsto(-y,x).\end{aligned} \tag{21.223}$$

Cette définition est un cas très particulier des structures complexes sur les variétés, voir la définition 66.62.

DEFooJFWEooXcIVUs

Définition 21.73.
*La **courbure signée** de la courbe $\gamma\colon I\to\mathbb{R}^2$ (I est un intervalle dans $\mathbb{R}$) est la fonction*

$$\kappa(t)=\frac{\gamma''(t)\cdot J\gamma'(t)}{\|\gamma'(t)\|^3} \tag{21.224}$$

EQooWOUQooXrVzGx

où J est la structure complexe de la définition 21.72.

Cette définition est motivée par le fait qu'en identifiant $\mathbb{R}^2$ à $\mathbb{C}$, l'application J revient à l'application $z\mapsto iz$.

Si $v,w\in\mathbb{R}^2$ nous avons formellement

$$v\times w=-(v\cdot Jw)e_3. \tag{21.225}$$

En particulier pour tout $v\in\mathbb{R}^2$ nous avons

$$v\cdot Jv=0. \tag{21.226}$$

Lemme 21.74.
Soir une courbe régulière $\gamma\colon[a,b]\to\mathbb{R}^2$ et un difféomorphisme $h\colon[c,d]\to[a,b]$. Si nous posons $\sigma=\gamma\circ h$ alors

$$\kappa_\sigma(u)=\operatorname{sgn}\big(h'(u)\big)\kappa_\gamma\big(h(u)\big). \tag{21.227}$$

EQooSQNMooUKGhPd

Démonstration. Nous utilisons la définition (21.224) de la courbure signée. La règle de dérivation en chaine donne :

$$\sigma'(u)=\gamma'\big(h(u)\big)h'(u) \tag{21.228a}$$
$$\sigma''(u)=\gamma''\big(h(u)\big)h'(u)^2+\gamma'\big(h(u)\big)h''(u). \tag{21.228b}$$

La numérateur de $\kappa_\sigma(u)$ est :

$$(\gamma'' \circ h)h'^2 \cdot J(\gamma' \circ h)h' + h''(\gamma' \circ h) \cdot J(\gamma' \cdot h)h' \tag{21.229}$$

dont le second terme est nul parce que $v \cdot Jv = 0$. Il nous reste donc

$$\kappa_\sigma(u) = \frac{(h')^3}{|h'|^3} \frac{(\gamma'' \circ h) \cdot J(\gamma' \circ h)}{\|\gamma' \circ h\|^3} = \operatorname{sgn}(h')\kappa_\gamma\big(h(u)\big). \tag{21.230}$$

□
LEMooKPORooEGJCRm

Lemme 21.75 ([543]).
Si γ_N est un arc paramétré normal, alors

$$\gamma_N''(s) = \kappa(s)J\gamma_N'(s). \tag{21.231}$$

Démonstration. Vu que le paramétrage est normal, $\gamma_N' \cdot \gamma_N' = 1$, et en dérivant, $\gamma_N'' \cdot \gamma_N' = 0$. Donc γ_N'' est un multiple de $J\gamma'$. En tenant compte du fait que le paramétrage est normal, la courbure est

$$\kappa(s) = \gamma_N''(s) \cdot J\gamma'(s). \tag{21.232}$$

En y injectant $\gamma''(s) = \lambda(s)J\gamma'(s)$ nous trouvons

$$\kappa(s) = \lambda(s)J\gamma'(s) \cdot J\gamma'(s) = \lambda(s). \tag{21.233}$$

Donc le facteur de proportionnalité est $\kappa(s)$. □
THOooDLDVooFQnLWn

Théorème 21.76.
Soit une courbe $\gamma\colon I \to \mathbb{R}^2$ de classe C^2.

(1) γ est une partie de droite si et seulement si $\kappa(t) = 0$ pour tout t.

(2) γ est une partie d'un cercle de rayon $r > 0$ si et seulement si $|\kappa(s)| = \frac{1}{r}$.

Démonstration. Si γ est une droite, la dérivée seconde est nulle et la courbure est nulle. Supposons pour la réciproque que $\kappa(t) = 0$ pour tout t. Nous utilisons un paramétrage normal de γ, ce qui ne change pas que la courbure reste nulle. Nous avons par le lemme 21.75 que $\gamma''(s) = 0$ et donc l'existence de $a, b \in \mathbb{R}^2$ tels que $\gamma(t) = at + b$.

Passons au cas du cercle. Si γ est un cercle, le paramétrage normal est

$$\gamma(t) = R\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix} \tag{21.234a}$$

$$\gamma'(t) = \begin{pmatrix}-\sin(\frac{t}{R})\\ \cos(\frac{t}{R})\end{pmatrix} \tag{21.234b}$$

$$\gamma''(t) = -\frac{1}{R}\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix}. \tag{21.234c}$$

Avec tout cela nous avons $\kappa(s) = \gamma''(s) \cdot J\gamma'(s) = \frac{1}{R}$.

Nous supposons enfin que $\kappa(t) = 1/R$ et que le paramétrage soit normal (encore une fois, un reparamétrage ne change pas la courbure lorsqu'elle est constante). Nous définissons la courbe

$$\begin{aligned}\beta\colon [b, c] &\to \mathbb{R}^2\\ t &\mapsto \gamma(t) + RJ\gamma'(t).\end{aligned} \tag{21.235}$$

Nous avons $\beta'(t) = \gamma'(t) + RJ\gamma''(t)$. Mais par le lemme 21.75 nous avons $\gamma'' = kJ\gamma' = \frac{1}{R}J\gamma'$. Donc

$$\beta'(t) = \gamma'(t) - \gamma'(t) = 0. \tag{21.236}$$

Du coup β est constante : $\beta(t) = a$. Alors $a = \gamma(t) + RJ\gamma'(t)$ et en particulier

$$\|\gamma(t) - a\| = \|RJ\gamma'(t)\| = r. \tag{21.237}$$

Donc effectivement γ reste sur un cercle de rayon r et de centre a. □

DEFooFAENooXWgBmh

Définition 21.77 ([543]).
*La **courbure totale** de la courbe $\gamma\colon [a,b]\to\mathbb{R}^2$ est le nombre*

$$K=\int_a^b \kappa(t)\|\gamma'(t)\|dt \tag{21.238}$$

EQooTIFWooQfIOfd

où κ est la courbure signée définie en 21.73.

*Nous nommons aussi **nombre de tours** de γ le nombre $\mathrm{Turn}(\gamma)=K/2\pi$.*

Lemme 21.78.
La courbure signée ne change pas sous reparamétrage positif, et change de signe sous reparamétrage négatif.

Démonstration. Soit la courbe $\gamma\colon [a,b]\to\mathbb{R}^2$ et un difféomorphisme $h\colon [a,b]\to[c,d]$. Il s'agit d'intégrer la relation (21.227) en effectuant le changement de variables $t=h(u)$:

$$K_\gamma=\int_a^b \kappa_\gamma(t)\|\gamma'(t)\|dt \tag{21.239a}$$

$$=\int_c^d \underbrace{\kappa_\gamma\big(h(u)\big)}_{\kappa_\sigma(u)}\|\gamma'\big(h(u)\big)\|h'(u)du. \tag{21.239b}$$

En utilisant le fait que $\sigma'(u)=\gamma'\big(h(u)\big)h'(u)$ nous avons alors

$$K_\gamma=\int_c^d \kappa_\sigma(u)\left\|\frac{\sigma'(u)}{h'(u)}\right\|h'(u)du \tag{21.240a}$$

$$=\int_c^d \kappa_\sigma(u)\|\sigma'(u)\|\frac{h'(u)}{|h'(u)|}du \tag{21.240b}$$

$$=\mathrm{sgn}(h')K_\sigma. \tag{21.240c}$$

□

LEMDEFooLPWJooAnWZjb

Lemme-Définition 21.79 ([543]).
Soit une courbe régulière $\gamma\colon I\to\mathbb{R}^2$ de classe C^2 et t_0 dans l'intérieur de I. Soit $\theta_0\in\mathbb{R}$ tel que

$$\frac{\gamma'(t_0)}{\|\gamma'(t_0)\|}=\big(\cos(\theta_0),\sin(\theta_0)\big). \tag{21.241}$$

Alors il existe une unique fonction différentiable $\theta\colon I\to\mathbb{R}$ telle que $\theta(t_0)=\theta_0$ et

$$\frac{\gamma'(t)}{\|\gamma'(t)\|}=\begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix} \tag{21.242}$$

pour tout $t\in I$.

*Cette fonction est l'**angle** de γ déterminé par θ_0.*

Démonstration. Soit $\beta(t)=(t,0)$; alors $\beta'(t)=(1,0)$ et nous avons

$$\gamma'\cdot\beta'=\gamma'_x \tag{21.243a}$$

$$\gamma'\cdot J\beta'=\gamma'_y. \tag{21.243b}$$

Par la proposition 21.71 il existe une unique fonction $\theta\colon I\to\mathbb{R}$ telle que $\theta(t_0)=\theta_0$ et

$$\begin{cases}\cos\big(\theta(t)\big)=\dfrac{\gamma'\cdot\beta'}{\|\gamma'\|\|\beta'\|}=\dfrac{\gamma'_x}{\|\gamma'\|} & (21.244a)\\ \sin\big(\theta(t)\big)=\dfrac{\gamma'\cdot J\beta'}{\|\gamma'\|\|\beta'\|}=\dfrac{\gamma'_y}{\|\gamma'\|}. & (21.244b)\end{cases}$$

Une telle fonction est bien celle que l'on demande ici. □

LEMooWLAUooKetUiW

Lemme 21.80.
Si γ est une courbe régulière de classe C^2, alors sa courbure et son angle vérifient la relation

$$\theta'(t) = \|\gamma'(t)\|\kappa(t). \tag{21.245}$$

Démonstration. Par définition de l'angle (lemme 21.79) nous avons

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \big(\cos\theta(t), \sin\theta(t)\big). \tag{21.246}$$

Dérivant cela,

$$\frac{\gamma''(t)}{\|\gamma'(t)\|} + \gamma'(t)\frac{d}{dt}\left(\frac{1}{\|\gamma'(t)\|}\right) = \theta'(t)\begin{pmatrix}-\sin\big(\theta(t)\big)\\ \cos\big(\theta(t)\big)\end{pmatrix} \tag{21.247a}$$

$$= \theta'(t)J\begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix} \tag{21.247b}$$

$$= \theta'(t)\frac{J\gamma'(t)}{\|\gamma'(t)\|}. \tag{21.247c}$$

Nous prenons le produit scalaire de cette égalité avec $J\gamma'$ en tenant compte du fait que $\gamma' \cdot J\gamma' = 0$:

$$\frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|} = \frac{\theta'}{\|\gamma'\|}J\gamma' \cdot J\gamma'. \tag{21.248}$$

En remarquant que $Jv \cdot Jv = \|v\|^2$ nous trouvons $\theta'\|\gamma'\| = \|\gamma'\|^2\kappa_\gamma$ et donc

$$\theta'(t) = \|\gamma'(t)\|\kappa_\gamma(t), \tag{21.249}$$

ce qu'il fallait prouver. □

Ce lemme nous fournit la formule attendue pour la courbure totale.

Lemme 21.81 ([543])**.**
Soit une courbe régulière de classe C^2 $\gamma\colon [a,b] \to \mathbb{R}^2$. Sa courbure totale[14] *est donnée par*

$$K = \theta(b) - \theta(a). \tag{21.250}$$

Démonstration. Il suffit de remplacer dans la définition (21.238) de la courbure totale l'intégrande par son expression du lemme 21.80 :

$$K = \int_a^b \kappa(t)\|\gamma'(t)\|dt = \int_a^b \theta'(t)dt = \theta(b) - \theta(a) \tag{21.251}$$

par le théorème 14.287. □

21.12.3 Degré, indice et homotopie

La définition de l'indice d'un chemin par rapport à un point est 26.50.

DEFooTKBUooNVcheO

Lemme-Définition 21.82.
Soit une application continue $\phi\colon S^1 \to S^1$. Un ***relèvement*** *de ϕ est une application continue $\varphi\colon \mathbb{R} \to \mathbb{R}$ une application continue telle que*

$$\phi\big(\cos(t), \sin(t)\big) = \begin{pmatrix}\cos\big(\varphi(t)\big)\\ \sin\big(\varphi(t)\big)\end{pmatrix}. \tag{21.252}$$

Le ***degré*** *est l'entier $\deg(\phi)$ tel que*

$$\varphi(2\pi) - \varphi(0) = 2\deg(\phi)\pi. \tag{21.253}$$

Ce nombre ne dépend pas du choix du relèvement φ.

14. Définition 21.77.

Démonstration. Soient φ_1 et φ_2, deux applications qui satisfont les contraintes. Nous avons une fonction continue $n\colon S^1 \to \mathbb{R}$ telle que

$$\varphi_1(t) - \varphi_2(t) = 2\pi n(t). \tag{21.254}$$

La fonction n ne pouvant prendre que des valeurs entières et étant continue, elle est constante. Par conséquent

$$\varphi_1(2\pi) - \varphi_1(0) = \varphi_2(2\pi) - \varphi_2(0), \tag{21.255}$$

ce que nous voulions. □

DEFooOCUQooUAlbLo

Définition 21.83.
Soit une courbe fermée $\gamma\colon [0,L] \to \mathbb{R}^2$ de classe C^1. Nous posons

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} \tag{21.256}$$

*où $\tilde{\gamma}(u) = \gamma\big(\frac{L}{2\pi}u\big)$. L'**indice de rotation** de γ est le degré de Φ_γ, c'est-à-dire*

$$\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma). \tag{21.257}$$

Notons que dans cette définition, Φ_γ n'est rien d'autre que le vecteur unitaire tangent à γ, ramené à $[0, 2\pi]$.

PROPooXHSDooDDnlJQ

Proposition 21.84.
Pour une courbe fermée de classe[15] C^2, l'indice de rotation est égal au nombre de tours.

Démonstration. Soit $\gamma\colon [0,L] \to \mathbb{R}^2$ une courbe vérifiant les hypothèses. Par définition, $\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma)$ où

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \begin{pmatrix} \cos\tilde{\theta}(t) \\ \sin\tilde{\theta}(t) \end{pmatrix} \tag{21.258}$$

EQooGTAPooDRCFPG

où nous avons noté $\tilde{\theta}$ l'angle tournant[16] de $\tilde{\gamma}$ et θ celui de γ. Vu que $\tilde{\theta}$ vérifie les hypothèses de la définition 21.82 nous pouvons calculer le degré de Φ_γ par

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\big(\tilde{\theta}(2\pi) - \tilde{\theta}(0)\big) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{\theta}'(s)ds. \tag{21.259}$$

EQooKNCJooTwTBMQ

Il faut trouver le lien entre θ et $\tilde{\theta}$. Pour cela nous notons que

$$\frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}\right)\|}. \tag{21.260}$$

En comparant avec (21.258) il vient

$$\tilde{\theta}(t) = \theta\big(\frac{L}{2\pi}t\big) \tag{21.261}$$

et

$$\tilde{\theta}'(s) = \frac{L}{2\pi}\theta'\left(\frac{L}{2\pi}s\right). \tag{21.262}$$

Le changement de variables $t = \frac{L}{2\pi}s$ est donc tout vu dans l'intégrale (21.259) :

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\int_0^{2\pi} \frac{L}{2\pi}\theta'\big(\frac{L}{2\pi}s\big)ds = \frac{1}{2\pi}\int_0^{L} \theta'(t)dt = \frac{1}{2\pi}\int_0^{L} \kappa_\gamma(t)\|\gamma'(t)\|dt = \mathrm{Turn}(\gamma) \tag{21.263}$$

où nous avons aussi utilisé le lemme 21.80 qui donne le lien entre θ' et κ. □

15. La dérivée seconde arrive dans la définition de la courbure ; il faudrait donc supposer au moins C^2 pour avoir la continuité de la courbure.
16. Définition 21.79.

Les définitions à propos d'homotopie sont 20.77 et 20.78.

LEMooTDTAooCRSYzI

Lemme 21.85 ([1]).
Soit un ouvert Ω de $\mathbb{C}$. Soit un chemin continu $\gamma\colon [a,b]\to\Omega$. Soit une boule B contenant $\gamma\big([a,b]\big)$ ainsi que le segment joignant $\gamma(a)$ et $\gamma(b)$.

Alors γ est homotope[17] *au segment joignant $\gamma(a)$ et $\gamma(b)$.*

Démonstration. Notons $d\colon [a,b]\to\Omega$ le segment de droite joignant $\gamma(a)$ à $\gamma(b)$. Nous notons aussi $l_{x,y}$ le segment joignant x à y et paramétré sur $[0,1]$:

$$\begin{aligned} l_{x,y}\colon [0,1]&\to\Omega\\ t&\mapsto (1-t)x+ty. \end{aligned} \tag{21.264}$$

L'homotopie entre γ et d est alors donnée par

$$\begin{aligned} H\colon [0,1]\times[a,b]&\to\Omega\\ (s,t)&\mapsto l_{\gamma(t),d(t)}(s). \end{aligned} \tag{21.265}$$

□

LEMooQNFCooWorCxN

Lemme 21.86 ([205]).
Soit un ouvert Ω de $\mathbb{C}$. Tout chemin dans Ω est homotope à extrémités fixées[18] *dans Ω à un chemin droit par morceaux.*

Démonstration. Nous allons utiliser le lemme 9.23 en considérant $S\subset[a,b]$ définie en disant que $t\in S$ soit si $t=a$ soit si le chemin $\gamma\colon [a,t]\to\Omega$ est homotope à un chemin droit par morceaux à extrémités fixes.

Nous prouvons que S est inductif (définition 9.22).

(1) $\underline{a\in S}$ Le point a est dans S par définition.

(2) **Si $\underline{t\in S}$** Supposons que $t\in[a,b[\cap S$, et cherchons $s>t$ tel que $[t,s]\subset S$. Considérons une boule $B=B\big(\gamma(t),\delta\big)$ incluse dans Ω. Vu que γ est continue, l'application $u\mapsto|\gamma(t)-\gamma(t+s)|$ est continue. Pour $s=0$, cette fonction est nulle. Il existe $\epsilon>0$ tel que cette fonction reste plus petite que δ, et donc tel que

$$\gamma\big([t,t+\epsilon]\big)\subset B\big(\gamma(t),\delta\big)\subset\Omega. \tag{21.266}$$

Le segment de droite $l_{\gamma(t),\gamma(t+\epsilon)}$ est donc dans Ω et la partie $\gamma\colon [t,t+\epsilon]\to\Omega$ est homotope à extrémités fixée à ce segment (lemme 21.85).

En joignant ce segment aux segments de droites homotopes à $\gamma\big([a,t]\big)$, nous avons des segments de droites homotopes à $\gamma\big([a,s]\big)$. Donc $[a,s]\subset S$.

(3) **Si $\underline{[a,t[\subset S}$** Nous prouvons que $t\in S$. Nous considérons $\delta>0$ tel que $B\big(\gamma(t),\delta\big)\subset\Omega$. Nous considérons également $\epsilon>0$ tel que $t-\epsilon>a$. Par hypothèse, $t-\epsilon\in S$, et il existe une homotopie à extrémités fixées entre des segments de droites et $\gamma\colon [a,t+\epsilon]\to\Omega$. Nous complétons ces segments par le segment $l_{\gamma(t-\epsilon),\gamma(t)}$ et voila.

Donc S est inductif, et nous en déduisons que $S=[a,b]$. □

PROPooZIAKooHqtnZj

Proposition 21.87 (Homotopie, degré et indice[543]).
Il y a deux résultats à ne pas confondre.

ITEMooLEHFooXEyTHY

(1) Si $\phi_i\colon S^1\to S^1$ sont homotopes, alors $\deg(\phi_1)=\deg(\phi_2)$.

(2) Si $\gamma_i\colon [0,L]\to\mathbb{R}^2$ sont homotopes, alors $\mathrm{Ind}(\gamma_1)=\mathrm{Ind}(\gamma_2)$.

Démonstration. Nous allons décomposer la preuve en deux parties.

17. Définition 20.78.
18. Définition 20.78.

(1) **Le degré pour les applications $S^1 \to S^1$** Soient deux applications homotopes $\gamma_i : S^1 \to S^1$. Si F est l'homotopie [19] entre γ_0 et γ_1, nous posons $\gamma_u(t) = F(u,t)$ qui est encore une courbe fermée $\gamma_u : S^1 \to S^1$. Nous pouvons donc considérer le degré de γ_u. C'est un entier n_u qui vérifie

$$\tilde{\gamma}_u(2\pi) - \tilde{\gamma}_u(0) = 2\pi n_u. \tag{21.267}$$

Le membre de gauche est une fonction continue de u ; dans le membre de droite n_u ne pouvant prendre que des valeurs entières, elle est alors constante.

(2) **Indice de rotation pour des courbes dans $\mathbb{R}^2$** Soient maintenant γ_0 et γ_1 deux courbes homotopes de $[0,L]$ dans $\mathbb{R}^2$. Par définition, $\text{Ind}(\gamma_i) = \deg(\Phi_{\gamma_i})$. Prouvons alors que Φ_{γ_0} et homotope à Φ_{γ_1}. De cette façon, la première partie de la preuve conclura à

$$\text{Ind}(\gamma_0) = \deg(\Phi_{\gamma_0}) = \deg(\Phi_{\gamma_1}) = \text{Ind}(\gamma_1). \tag{21.268}$$

Nous savons que

$$\Phi_\gamma = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}t\right)\|}, \tag{21.269}$$

donc en posant

$$\Phi_F(u,t) = \frac{\gamma'_u\left(\frac{L}{2\pi}t\right)}{\|\gamma'_u\left(\frac{L}{2\pi}t\right)\|} \tag{21.270}$$

nous avons une homotopie entre Φ_{γ_0} et Φ_{γ_1}.

□

THOooEQWOooBCRMMZ

Théorème 21.88 ([543])**.**
Pour une courbe simple fermée de classe C^2 dans $\mathbb{R}^2$,

(1) le nombre de tours [20] vaut ± 1,

(2) la courbure totale vaut $\pm 2\pi$.

Démonstration. Soit une telle courbe γ et son image Γ. Notons tout de suite que les deux points sont équivalents parce que $K = 2\pi \,\text{Turn}(\gamma)$.

(1) **Choix d'un point et d'une tangente**

Si ℓ est une droite dans le plan, soit p le point de Γ le plus proche de ℓ. Il n'est peut-être pas unique, mais la parallèle à ℓ passant par p est une tangente à Γ telle que tout Γ se trouve d'un seul côté de ℓ_p. Pour voir cela, il suffit de choisir un système d'axes pour lequel ℓ est l'axe $y = 0$. La distance entre ℓ et les points de Γ est donnée par γ_y, et donc les extrémums sont atteints là où $\gamma'_y = 0$, c'est-à-dire pour les points sur lesquels la tangente est parallèle à ℓ.

Étant donné qu'il y a (au moins) un maximum et un minimum distincts, nous pouvons choisir le point p de telle sorte que pour tout point $q \in \Gamma$, l'angle du vecteur $q - p$ soit entre α_0 et $\alpha_0 + \pi$ et non entre $\alpha_0 - \pi$ et α_0. Ce choix revient à choisir p de telle sorte que Γ soit d'un côté ou de l'autre de ℓ_p.

Pour simplifier les notations plus tard nous choisissons ℓ_p horizontale, de telle sorte que $\alpha_0 = 0$, et que Γ est au dessus de ℓ_p. Les angles des vecteurs $q - p$ pour $q \in \Gamma$ sont donc tous entre 0 et π.

(2) **Définition de Σ**

Soit L la longueur [21] de γ et β, un paramétrage de γ telle que $\beta(0) = p$. Nous considérons le triangle

$$\mathcal{T} = \{(t_1,t_2) \in \mathbb{R}^2 \text{ tel que } 0 \leqslant t_1 \leqslant t_2 \leqslant L\} \tag{21.271}$$

19. Définition 20.77.
20. Nombre de tours et courbure totale, définition 21.77.
21. Définition 21.3.

et l'application sécante $\Sigma\colon \mathcal{T} \to S^1$ définie par

$$\Sigma(t_1,t_2) = \begin{cases} \dfrac{\beta'(t)}{\|\beta'(t)\|} & \text{si } t_1 = t_2 = t \\ -\dfrac{\beta'(0)}{\|\beta'(0)\|} & \text{si } t_1 = 0 \text{ et } t_2 = L \\ \dfrac{\beta(t_2)-\beta(t_1)}{\|\beta(t_2)-\beta(t_1)\|} & \text{sinon.} \end{cases} \tag{21.272}$$

(3) **Continuité de Σ** Nous devons prouver les limites[22] suivantes :

$$\lim_{\substack{(t_1,t_2)\to(t,t)\\ 0\leqslant t_1\leqslant t\leqslant t_2\leqslant L}} \frac{\beta(t_1)-\beta(t_2)}{\|\beta(t_1)-\beta(t_2)\|} = \frac{\beta'(t)}{\|\beta'(t)\|} \tag{21.273a}$$

$$\lim_{\substack{(t_1,t_2)\to(0,L)\\ 0\leqslant t_1\leqslant t_2\leqslant L}} \frac{\beta(t_1)-\beta(t_2)}{\|\beta(t_1)-\beta(t_2)\|} = -\frac{\beta'(0)}{\|\beta'(0)\|} \tag{21.273b}$$

Nous commençons par (21.273a) en multipliant et divisant par $t_2 - t_1$:

$$\frac{\beta(t_2)-\beta(t_1)}{\|\beta(t_2)-\beta(t_1)\|} = \frac{\beta(t_2)-\beta(t_1)}{t_2-t_1}\frac{t_2-t_1}{\|\beta(t_2)-\beta(t_1)\|}. \tag{21.274}$$

Si chacune des limites des deux facteurs existent dans $\mathbb{R}$, la limite du produit sera le produit des limites. Pour le premier facteur nous développons $\beta(t_2)$ autour de $t = t_1$ via la formulation (12.1300) :

$$\beta(t_2) = \beta(t_1) + (t_2-t_1)\beta'(t_1) + \alpha(t_2-t_1) \tag{21.275}$$

où α est une fonction vérifiant $\lim_{t\to 0}\alpha(t)/t = 0$. Nous avons à calculer la limite de

$$\frac{(t_2-t_1)\beta'(t_1)+\alpha(t_2-t_1)}{t_2-t_1} = \beta'(t_1) + \frac{\alpha(t_2-t_1)}{t_2-t_1}. \tag{21.276}$$

Prendre la limite $(t_1,t_2)\to(t,t)$ donne bien $\beta'(t)$ parce que β est de classe C^1. En ce qui concerne la limite de la deuxième partie, nous allons la faire plus en détail pour la limite (21.273b).

La limite (21.273b). Nous prenons la prolongation périodique de β. Alors si $t_2 = L - \epsilon$ nous pouvons écrire $\beta(-\epsilon)$ au lieu de $\beta(t_2)$. Nous développons $\beta(t_1)$ autour de $t = -\epsilon$ (parce que t_1 est petit) :

$$\beta(t_1) = \beta(-\epsilon) + (t_1+\epsilon)\beta'(-\epsilon) + \alpha(t_1+\epsilon). \tag{21.277}$$

Après multiplication et division par $t_1 + \epsilon$, la première limite à calculer est celle de

$$\frac{\beta(-\epsilon)-\beta(t_1)}{t_1+\epsilon} = -\beta'(-\epsilon) + \frac{\alpha(t_1+\epsilon)}{t_1+\epsilon}, \tag{21.278}$$

pour $(\epsilon,t_1)\to(0,0)$. Cela donne bien $-\beta'(0)$. La seconde limite à calculer est celle de

$$\frac{t_1+\epsilon}{\|\beta(-\epsilon)-\beta(t_1)\|} = \frac{t_1+\epsilon}{\|-(t_1+\epsilon)\beta'(-\epsilon)-\alpha(t_1+\epsilon)\|} = \frac{t_1+\epsilon}{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}. \tag{21.279}$$

Nous calculons la limite de l'inverse (qui, si elle est non nulle donnera la réponse en inversant à nouveau) en nous souvenant de la formule

$$\big|a-|b|\big| \leqslant |a+b| \leqslant \big|a+|b|\big|. \tag{21.280}$$

Nous avons l'encadrement

$$\frac{\big|(t_1+\epsilon)\beta'(-\epsilon)-|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} \leqslant \frac{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}{t_1+\epsilon} \tag{21.281a}$$

$$\leqslant \frac{\big|(t_1+\epsilon)\beta'(-\epsilon)+|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} \tag{21.281b}$$

Les limites des deux extrêmes existent et valent $\beta'(0)$; la règle de l'étau 12.221 conclu.

22. Limite au sens de la définition 7.110, en sachant qu'elle est unique par la proposition 7.113.

(4) **Deux chemins homotopes** Nous considérons dans $\mathcal{T}$ les points $A = (0,0)$, $B = (0,L)$ et $C = (L,L)$. Nous allons considérer le chemin direct de A à C et celui passant via B. Et comme ces chemins doivent être paramétrés de 0 à 2π, il faut faire un peu attention. Nous définissons les chemins

$$\sigma_i \colon S^1 \to S^1 \tag{21.282}$$

par

$$\sigma_1(t) = \Sigma\left(\frac{L}{2\pi}t, \frac{L}{2\pi}t\right) = \frac{\beta'(tL/2\pi)}{\|\beta'(tL/2\pi)\|} = \frac{\tilde{\beta}'(t)}{\|\tilde{\beta}'(t)\|} = \Phi_\beta \tag{21.283a}$$

$$\sigma_2(t) = \begin{cases} \Sigma(0, \frac{L}{\pi}t) & \text{si } 0 \leqslant t \leqslant \pi \\ \Sigma\big(\frac{L}{\pi}(t-\pi), L\big) & \text{si } \pi \leqslant t \leqslant 2\pi \end{cases} \tag{21.283b}$$

où nous avons repris les notations de la définition 21.83. Notons que pour σ_2, en $t = \pi$ les deux expressions donnent

$$\sigma_2(\pi) = \Sigma(0,L) = -\frac{\beta'(0)}{\|\beta'(0)\|}. \tag{21.284}$$

Ce sont des chemins fermés parce que

$$\sigma_1(2\pi) = \Sigma(L,L) = \frac{\beta'(L)}{\|\beta'(L)\|} = \frac{\beta'(0)}{\|\beta'(0)\|} = \Sigma(0,0) = \sigma_1(0). \tag{21.285}$$

Notons que dans toutes ces définitions et calculs, nous avons utilisé de façon assez cruciale la définition 21.70 pour définir la dérivée de β en $t = 0$.

Les chemins σ_1 et σ_2 sont homotopes par construction.

(5) **Indices et degrés** La proposition 21.87(1) nous donne $\deg(\sigma_1) = \deg(\sigma_2)$. Nous avons alors la chaine d'égalités

$$\deg(\sigma_2) = \deg(\sigma_1) = \deg(\Phi_\beta) = \mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma) \tag{21.286}$$

EQooSXOAooZVOVxc

où les justifications sont :

(5a) $\deg(\sigma_2) = \deg(\sigma_1)$ par homotopie : proposition 21.87.

(5b) $\deg(\sigma_1) = \deg(\Phi_\beta)$ parce que $\sigma_1 = \Phi_\beta$.

(5c) $\deg(\Phi_\beta = \mathrm{Ind}(\beta)$ par définition 21.83 de l'indice.

(5d) $\mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma)$ par invariance de l'indice sous reparamétrage.

(5e) $\mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma)$ par la proposition 21.84.

Il nous reste à montrer que $\deg(\sigma_2) = \pm 1$.

(6) **La géométrie de σ_2** Notons que par définition les valeurs de $\sigma_2(t)$ sont les vecteurs (unitaires) joignant p aux points de Γ lorsque $0 \leqslant t \leqslant \pi$, et les vecteurs inverses pour $\pi \leqslant t \leqslant 2\pi$. Plus précisément nous avons, si $0 < a < \pi$:

$$\sigma_2(a) = \Sigma(0, \frac{L}{\pi}a) = \frac{\beta(\frac{L}{\pi}a) - \beta(L)}{\|\beta(\frac{L}{\pi}a) - \beta(L)\|} \tag{21.287}$$

et

$$\sigma_2(\pi + a) = \Sigma(\frac{L}{\pi}a, L) = \frac{\beta(L) - \beta(\frac{L}{\pi}a)}{\|\beta(L) - \beta(\frac{L}{\pi}a)\|}. \tag{21.288}$$

En sachant que $\beta(0) = \beta(L)$ nous avons alors

$$\sigma_2(\pi + a) = -\sigma_2(a). \tag{21.289}$$

EQooKMSRooEzWkyL

Notons aussi que le fait que γ soit une courbe simple assure que le numérateur et le dénominateur de σ_2 ne s'annulent pas autrement que pour $t = L$ ou $t = 0$.

(7) **Un relèvement pour σ_2** Ces propriétés motivent cette idée pour le relèvement de σ_2 :

$$\begin{aligned}\theta\colon [0,2\pi[&\to [0,2\pi[\\ t &\mapsto \text{l'angle du vecteur } \sigma_2(t),\end{aligned} \tag{21.290}$$

avec $\theta(2\pi)$ définit par continuité. Nous allons cependant voir, en étant plus prudent, que cette définition n'assure pas la continuité (surtout en $t=\pi$).

Soyons donc plus prudent et construisons θ petit à petit.

Étant donné que ℓ_p est tangente à Γ au point p, la droite ℓ_p est parallèle à $\beta'(0)$, et l'angle entre ℓ_p et $\beta'(0)$ est soit 0 soit π. Donc $\theta(0)$ devrait valoir soit 0 soit π.

Nous commençons par définir ceci :

$$\begin{aligned}\theta\colon [0,\pi[&\to [0,\pi]\\ t &\mapsto \text{angle de } \sigma_2(t) = \arccos\big(\sigma_2(t)_x\big).\end{aligned} \tag{21.291}$$

C'est le choix d'avoir ℓ_p horizontale et Γ au dessus de ℓ_p qui nous assure que pour tout $t\in[0,\pi]$, l'angle de $\sigma_2(t)$ peut être choisi entre 0 et π. De plus cette fonction est continue en tant que partie de la fonction $\arccos\colon [-1,1]\to[0,\pi]$ qui est elle-même continue par le lemme 7.222.

Le fait que θ soit continue est assuré par le fait que σ_2 est C^∞.

(7a) **Si $\theta(0)=0$**

Cela correspond à la situation des vecteurs rouges sur la figure 21.11.

Vu que $\sigma_2(\pi)=-\sigma_2(0)$, l'angle de σ_2 en $t=\pi$ est le supplémentaire de celui en $t=0$. Mais pour $0\leqslant t<\pi$, $\theta(t)$ prend ses valeurs entre 0 et π, le seul supplémentaire de 0 à être disponible est $\theta(\pi)=\pi$ (et non $\theta(\pi)=-\pi$ par exemple). Nous définissons donc $\theta(\pi)=\pi$ pour la continuité.

En ce qui concerne les $\pi<t<2\pi$ nous savons que $\sigma_2(\pi+a)=-\sigma_2(a)$. Les angles sont donc les supplémentaires de ceux pour $0\leqslant t\leqslant\pi$. Pour assurer la continuité en $t=\pi$ nous sélectionnons la place $]\pi,2\pi]$; et nous définissons

$$\begin{aligned}\theta\colon]\pi,2\pi[&\to]\pi,2\pi]\\ t &\mapsto \text{angle de } \sigma_2(t).\end{aligned} \tag{21.292}$$

Le nombre $\theta(2\pi)$ est défini par continuité. Il doit valoir $\theta(2\pi)=2\pi$.

La fonction θ ainsi définie est un relèvement pour σ_2, et le degré peut être calculé :

$$\deg(\sigma_2)=\frac{1}{2\pi}\big(\theta(2\pi)-\theta(0)\big)=1. \tag{21.293}$$

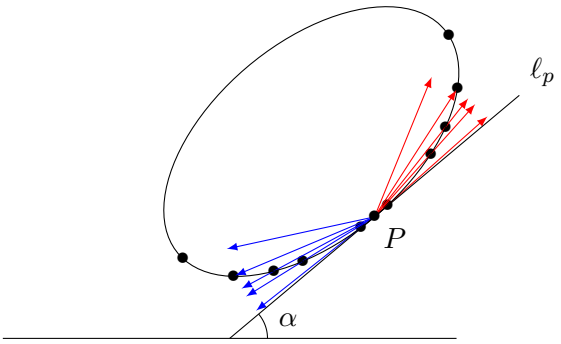


FIGURE 21.11: Les vecteurs représenant σ_2 dans le cas où $\beta'(0)$ est dans le sens de ℓ_p ou dans le sens inverse. Pour le sport nous avons dessiné la situation avec une droite ℓ quelconque plutôt que horizontale. LabelFigERPMooZibfNOiU

(7b) **Si $\theta(0) = \pi$** Cela correspond à la situation des vecteurs bleus sur la figure 21.11. Alors $\theta(\pi) = 0$ est obligatoire parce qu'il doit être supplémentaire à $\theta(0)$. Les angles atteints par $\sigma_2(t)$ pour $t \in]\pi, 2\pi[$ sont encore les complémentaires, mais cette fois la continuité en $t = \pi$ nous impose de les chercher dans $]-\pi, 0]$ et nous définissons

$$\begin{aligned} \theta\colon]\pi, 2\pi[&\to]-\pi, 0] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \tag{21.294}$$

Par continuité nous devons avoir $\theta(2\pi) = \theta(\pi) - \pi$ et donc $\theta(2\pi) = -\pi$.
Le degré de σ_2 est alors

$$\deg(\sigma_2) = \frac{1}{2\pi}\big(\theta(2\pi) - \theta(0)\big) = -1. \tag{21.295}$$

(8) **Conclusion**
Nous avons montré que le degré de σ_2 est 1 ou -1, et en remontant les égalités (21.286) nous déduisons que $\text{Turn}(\gamma) = \pm 1$.

□

21.13 Courbes fermées planes

21.13.1 Cercle circonscrit

La proposition suivante est dans le même esprit que l'ellipse de John-Loewer [23].

PROPDEFooCWESooVbDven

Proposition-Définition 21.89 (Cercle circonscrit[608, 1])**.**
Soit une courbe fermée simple et continue $\gamma\colon [0,1] \to \mathbb{R}^2$. Soit Γ son image. Il existe un unique cercle de rayon minimum contenant Γ. Ce cercle est le ***cercle circonscrit*** *à γ.*

Nous nommons C ce cercle circonscrit à γ. Il a les propriétés suivantes :

(1) L'intersection $C \cap \Gamma$ contient au moins deux points distincts.

(2) Tout arc de C plus long que sa moitié intersecte Γ.

Démonstration. Division de la preuve.

(1) **Existence** L'application γ étant continue, l'ensemble Γ est compact (théorème 7.220). Nous considérons l'ensemble Q des formes quadratiques de la forme

$$q_{a,r}(x) = \|a - x\|^2 - r^2 \tag{21.296}$$

avec $a \in \mathbb{R}^2$ et $r \in \mathbb{R}$. Nous mettons sur cet ensemble la topologie de $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$. Le nombre $q_{a,r}(x)$ est continu en a, r et x. Soit A l'ensemble des formes quadratiques de cette forme et vérifiant

$$q\big(\gamma(t)\big) \leqslant 0 \tag{21.297}$$

pour tout $t \in [0,1]$.
Cet ensemble A est non vide parce que Γ est compact et donc borné ; il existe donc une boule qui contient Γ en son intérieur.
Montrons que A est fermé dans Q. Si $q_{a,r} \notin A$ alors il existe t_0 tel que $q_{a,r}\big(\gamma(t_0)\big) > 0$. Par continuité, il existe un voisinage de (a,r) dans $\mathbb{R}^3$ et donc de $q_{a,r}$ dans Q tel que $q_{a',r'}\big(\gamma(t_0)\big)$ reste strictement positif pour tout (a',r') dans ce voisinage. Le complémentaire de A dans Q est ouvert. Donc A est fermé.
L'ensemble

$$B = \{r \geqslant 0 \text{ tel que } \exists\, a \text{ tel que } q_{a,r} \in A\} \tag{21.298}$$

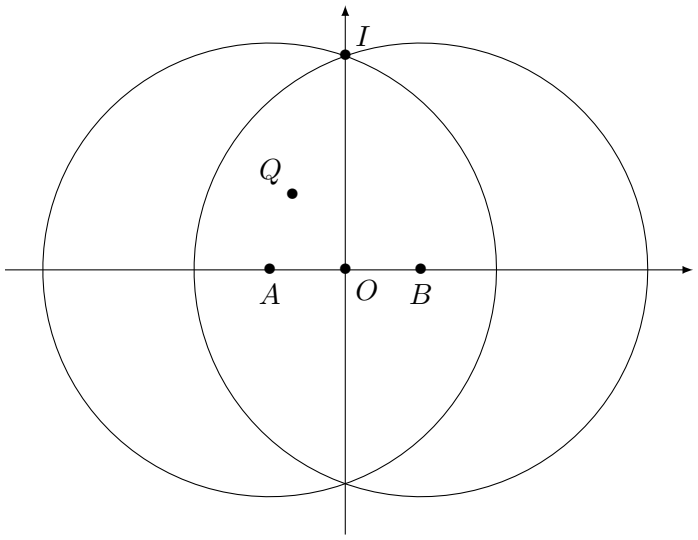
est fermé [24]. La partie B est évidemment borné vers le bas. De plus $r = 0$ n'est pas dans cet ensemble. Il possède donc un minimum strictement positif. Le cercle correspondant donne l'existence.

23. Proposition 17.122.
24. Lemme 7.322.

(2) **Unicité**

En ce qui concerne l'unicité, si Γ est contenu dans les boules $B(a,R)$ et $B(b,R)$ alors

$$\Gamma \subset B(a,R) \cap B(b,R). \tag{21.299}$$



Nous choisissons les axes comme indiqué sur le dessin et nous montrons que l'intersection est dans le cercle de centre O et de rayon $\|OI\| < R$. Soit Q un point de l'intersection ; par symétrie il est suffisant de supposer $Q_x < 0$ et $Q_y > 0$. Vu que Q est dans le cercle de centre B et de rayon R, il doit satisfaire

$$(B_x - Q_x)^2 + Q_y^2 < R. \tag{21.300}$$

D'autre part le point I est d'abscisse $I_x = 0$ et d'ordonnée donnée par $I_y^2 = R^2 - B_x^2$.
Nous devons prouver que $Q_x^2 + Q_y^2 \leqslant I_y^2$. Il s'agit simplement de calculer

$$Q_x^2 + Q_y^2 \leqslant Q_x^2 + R^2 - B_x^2 + 2B_xQ_x - Q_x^2 = R^2 - B_x^2 + 2B_xQ_x \leqslant R^2 - B_x^2 = I_y^2 \tag{21.301}$$

parce que $B_x > 0$ et $Q_x < 0$.
Nous concluons que Γ est inclus dans un cercle de rayon plus petit que R et donc que R n'est pas minimum. D'où l'unicité.

(3) **Au moins deux intersections**

Nous nommons C le cercle circonscrit à γ, et nous écrivons, pour $p \in \Gamma$

$$r(p) = d(p,C) \tag{21.302}$$

la distance entre p et C. Cela est une fonction continue sur le compact Γ. Elle atteint donc ses bornes sur Γ.

(3a) **Impossible d'avoir zéro intersection**

Si $C \cap \Gamma = \varnothing$ alors $r(p) > 0$ pour tout p et le minimum est également strictement positif. Soit r_0 ce minimum :

$$r_0 = \min\{d(p,C) \text{ tel que } p \in \Gamma\}. \tag{21.303}$$

Alors le cercle C_2 même centre que C mais de rayon $r_0/2$ n'intersecte pas non plus Γ. En effet soient $p \in \Gamma$ et $s \in C_2$. Nous notons s' le point de C situé sur la demi-droite ps. Nous avons

$$d(p,s) \geqslant d(p,s') - \frac{r_0}{2} \geqslant d(p,C) - \frac{r_0}{2} \tag{21.304a}$$

$$\geqslant r_0 - \frac{r_0}{2} = \frac{r_0}{2} \tag{21.304b}$$

$$> 0. \tag{21.304c}$$

Cela étant valable pour tout $s \in C_2$ (et d étant continue en s sur le compact C_2), nous en déduisons que $d(p,C_2) > 0$ pour tout $p \in \Gamma$. Donc $\Gamma \cap C_2 = \varnothing$.

(3b) **Impossible d'avoir une seule intersection**

Si $C \cap \Gamma = \{q\}$ alors un choix d'axe place le centre de C en $(0,0)$, le point q en $(1,0)$ et fixe le rayon de C à 1. Vu que Γ ne coupe C qu'un un seul point, situé tout à droite, il devrait être possible de décaller un peu C pour le « décoller » de q tout en continuant à ne pas toucher Γ en les autres points.

C'est ce que nous allons faire maintenant.

Pour $p \in \Gamma$ nous notons $r(p)$ la distance entre p et le point de C situé sur sa gauche, c'est-à-dire, si $p = (p_x, p_y)$,

$$r(p) = \sqrt{1 - p_y^2} + p_x. \tag{21.305}$$

EQooCWRDooRzwqKj

Cela est encore une fonction continue sur Γ qui atteint son minimum valant r_0. Nous considérons maintenant le cercle C_2 de centre $(\frac{r_0}{2}, 1)$, et nous prouvons que Γ est dans C_2 (au sens dans la boule dont la frontière est C_2) et que $C_2 \cap \Gamma = \varnothing$. Donc C_2 retombe dans un cas que nous avons prouvé être impossible.

Soit $p \in \Gamma$. Si $p \in C_2$, alors $p - (r_0/2, 0) \in C$. Cela est impossible parce que la distance entre p et le point à sa gauche sur C est au moins r_0. Donc $C_2 \cap \Gamma = \varnothing$.

Le point q est dans C_2 parce que $(1,0) \in B\big((r_0/2, 0), 1\big)$. Par continuité de C_2 et de Γ, tous les points de Γ sont dans C_2. Sinon il y aurait une intersection.

(3c) **Si on en veut plus** Nous avons fini de démontrer qu'il y a au moins deux intersections.

Pour le sport, notez que tout nombre de points d'intersection entre C et Γ est possible à partir de 2. Pour en avoir deux, prendre une ellipse, et pour en avoir plus, prendre des polynômes dont les angles sont un peu modifiés de façon à rester C^1.

(4) **Intersection avec les demi-arcs**

Supposons, en fixant encore les axes, que le cercle circonscrit soit encore centré en $(0,0)$ et que Γ n'intersecte pas le demi-cercle $x < 0$. Alors pour tout $p \in \Gamma$ la distance entre p et ce demi-cercle est strictement positive. Il y a un minimum r_0. En décalant le centre du cercle de $r_0/2$ vers la droite, nous obtenons un nouveau cercle contenant Γ mais ne l'intersectant pas.

□

21.13.2 Description locale

DEFooVQODooJSNYLw

Définition 21.90.
Une courbe plane différentiable est ***convexe*** *si son graphe est en tout point d'un seul côté de sa tangente.*

LEMooGEVEooHxPTMO

Lemme 21.91 ([1])**.**
Soit une courbe simple, convexe $\gamma\colon [0,1] \to \mathbb{R}^2$ que nous supposons être de classe C^1. Nous notons Γ l'image de γ. Alors Γ est localement le graphe d'une fonction convexe (définition 17.76).

Démonstration. Soit $p \in \Gamma$ et l_p la tangente à Γ en p. Nous considérons un système d'axe centré en p de telle sorte que l_p soit l'axe des abscisses et l'axe des ordonnées soit dirigé de tel manière que Γ se trouve dans la partie $y > 0$. De plus nous paramétrons γ de telle sorte à avoir $\gamma(0) = p = (0,0)$.

Vu que $l_p \equiv y = 0$ est la tangente à Γ nous avons $\gamma_y'(0) = 0$ et $\gamma_x'(0) > 0$. Nous en déduisons que $\gamma_x'(t) > 0$ pour tout $t \in B(0,\delta)$ pour δ suffisamment petit. Nous posons alors

$$g(x) = \gamma_y\big(\gamma_x^{-1}(x)\big) \tag{21.306}$$

qui est bien définie parce que γ_x est une bijection entre $B(0,\delta)$ et son image. La fonction g est continue par le théorème de la bijection 12.51, et même dérivable par la proposition 12.187. De plus, vu la formule (12.461), la fonction g^{-1} est de classe C^1 parce que $(g^{-1})'$ est une composée d'applications continues.

(1) **Si g est C^2** Dans ce cas, g'' ne peut pas changer de signe, sinon la tangente coupe le graphe. Par positivité de g (et le fait que $g(0) = g'(0) = 0$), il n'est pas possible d'avoir $g'' < 0$ partout. Donc $g'' \geqslant 0$ partout. Cela prouve que g est convexe par la caractérisation 17.84.

(2) **Si g est seulement de classe C^1** Le graphe de g correspond au graphe de Γ. Nous montrons que g est convexe en utilisant la caractérisation de la proposition 17.89.
La tangente au graphe de g en $x = x_0$, que nous notons l_0, est la tangente à Γ en $t = \gamma_x^{-1}(x_0)$. Le graphe de g, qui est une partie de Γ se trouve donc d'un seule côté de l_0.
Nous nous restreignons g à un compact I et nous considérons la fonction

$$d_a(x) = g(x) - l_a(x) \tag{21.307}$$

qui donne la distance entre le graphe de g et la tangente à g en $x = a$. Cela est une fonction continue en x et en a. Le graphe de g est au dessus de la tangente en $x = 0$ (par construction des axes). Supposons que le graphe de g soit en dessous de la tangente en $x = x_2$. Alors nous avons, pour tout x :

$$\begin{cases} d_0(x) \geqslant 0 & \text{(21.308a)} \\ d_{x_1}(x) \leqslant 0. & \text{(21.308b)} \end{cases}$$

Nous posons

$$s(a) = \sup_{x \in I} d_a(x), \tag{21.309}$$

qui est une fonction continue par la proposition 12.33. L'application s vérifie $s(0) \geqslant 0$ et $s(x_2) \leqslant 0$. Il existe donc $m \in [0, x_2]$ tel que $s(m) = 0$. À ce moment nous avons $g(x) = l_m(x)$ pour tout $x \in I$ et donc g est une droite, et en réalité toutes les inégalités sont des égalités. La fonction g est alors bien convexe (mais pas strictement).

□

21.13.3 Enveloppe convexe

PROPooWZITooTFiWsi

Proposition 21.92 ([609, 1])**.**
Soit une courbe simple, fermée et convexe $\gamma\colon [0,1] \to \mathbb{R}^2$ que nous supposons être de classe C^2. Nous notons Γ l'image de γ. Alors il existe un convexe D tel que $\partial D = \Gamma$.

Démonstration. Pour $p \in \Gamma$ nous notons l_p la tangente à Γ en p et H_p le demi-plan (fermé) contenant Γ. Nous posons

$$D = \bigcap_{p \in \Gamma} H_p. \tag{21.310}$$

EQooDYFTooCHRbsD

Cet ensemble est convexe comme intersection de convexes [25] et fermé comme intersection de fermés [26]. Nous prouvons que $\Gamma = \partial D$.

L'inclusion $\Gamma \subset \partial D$ est la plus facile. Si $p \in \Gamma$ alors p est dans chacun des H_q et donc dans D. De plus tout voisinage de p contient des points en dehors de H_p, donc p n'est pas dans l'intérieur de D. Ce dernier étant fermé, un point hors de l'intérieur est sur le bord. Ergo $p \in \partial D$.

Pour l'inclusion inverse, soit $p \in \partial D$.

(1) **Il existe q tel que $p \in l_q$** Vu que D est fermé, le point p est dans D, et donc dans tous les H_q. Supposons qu'il soit dans l'intérieur de tous les H_q. Alors nous considérons la fonction

$$r(q) = \frac{d(p, l_q)}{2} \tag{21.311}$$

définie sur Γ. C'est une fonction continue [27] strictement positive définie sur le compact Γ qui possède donc un minimum strictement positif. Si r_0 est ce minimum, alors $B(p, r_0)$ est incluse à tous les H_q, ce qui ferait que p est à l'intérieur de D. Nous concluons à l'existence de $q \in \Gamma$ tel que $p \in l_q$.

25. Proposition 7.155.
26. Lemme 7.6(1).
27. L'équation de la droite l_q a des coefficients continus parce que γ est de classe C^1.

(2) **Le point où ça décolle** Nous supposons que $p \notin \Gamma$, sinon ce serait trop facile. Nous paramétrons γ de telle sorte à avoir $q = \gamma(0)$ et nous posons

$$T = \{t \in \mathbb{R}^+ \text{ tel que } l_{\gamma(t)} = l_q\}. \tag{21.312}$$

Cela est un fermé dans $\mathbb{R}$ parce que γ est de classe C^1. Nous posons $t_0 = \inf(T^c)$ et pour tout ϵ suffisamment petit,

$$\begin{cases} l_{\gamma(t_0+\epsilon)} \neq l_q, & \text{(21.313a)} \\ l_{\gamma(t_0-\epsilon)} = l_q. & \text{(21.313b)} \end{cases}$$

La seconde est parce que si $l_{\gamma(t_0-\epsilon)} \neq l_q$ nous aurions $t_0 - \epsilon \in T^c$. Soit $r = \gamma(t_0)$; nous avons $l_r = l_q$ parce que si $l_r \neq l_q$ alors par continuité de γ' nous aurions $l_{r'} \neq l_q$ pour tout $r' \in \gamma\big(B(t_0, \delta)\big)$.

(3) **Graphe d'une fonction strictement convexe** En suivant le lemme 21.91, l'ensemble Γ est localement (autour de r) le graphe d'une fonction convexe au-dessus de l_r. Soit $g\colon B(0,\delta) \to \mathbb{R}^2$ cette fonction convexe.

Soit $\epsilon > 0$. Si $g''(x) = 0$ sur $]0, \epsilon]$ alors g' y est constante. Mais $g'(0) = 0$, ce qui signifierait que sur $[0, \epsilon]$ nous ayons $g'(x) = 0$ et donc $g(x) = 0$. Cela ferait que $l_{r'} \equiv y = 0$ pour tout r' de la forme $(x, 0)$ avec $x \in [0, \epsilon]$ (qui sont des points de Γ). Cela est en contradiction avec la définition de r. Donc il existe un point $x \in [0, \epsilon[$ tel que $g''(x) > 0$.

Rappelons que $p \in l_q$, ce qui fait que p a pour coordonnées $(p_x, 0)$. Nous restreignons δ et ϵ de telle sorte que p_x soit plus grand à la fois que ϵ et δ.

Il existe donc un intervalle $[a, b]$ avec $a, b \geqslant 0$ et $a, b < p_x$ sur lequel g est strictement convexe.

(4) **La tangente qui tue**

En particulier $g(b) > 0$, de telle sorte que le théorème des accroissements finis 12.192(1) nous donne l'existence de $m \in [0, b]$ tel que la tangente à g en $x = m$ est parallèle au segment joignant $(0, 0)$ à $(b, g(b))$. Cette tangente, que nous nommons l_m, est en dessous de la corde, par stricte convexité. En particulier, son point d'intersection avec $y = 0$ est strictement entre 0 et m.

L'ensemble D est d'un seul côté de l_m. Ce côté est forcément celui de $q = (0, 0)$ (parce que $q \in \Gamma \subset D$), et donc le points de coordonnées $(x, 0)$ avec $x > m$ ne sont pas dans D. Pas de chance, p est un point de ce type.

(5) **La contradiction** Nous avons prouvé que si $p \in \partial D \backslash \Gamma$ alors p n'est pas dans D, ce qui est impossible parce que, l'ensemble D étant fermé, nous avons $\partial D \subset D$.

□

Remarque 21.93.
Bien que cela puisse paraitre évident dès le début, nous ne démontrerons que dans la proposition 21.95 que D est l'enveloppe convexe de Γ.

CORooSXDGooJEmVcf

Corolaire 21.94.
Si $p \in \mathrm{Int}(D)$ alors toute droite passant par p intersecte Γ en exactement 2 points.

Démonstration. Vu que la partie D est bornée, toute droite passant par son intérieur coupe ∂D en au moins deux points (un dans chaque sens, et en utilisant le lemme de passage de douane 7.130).

Soit ℓ une droite passant par p et supposons qu'elle coupe Γ en trois points distincts. Alors au moins deux d'entre eux sont du même côté de p. Soient q_1 et q_2 ces points. Nous avons donc dans l'ordre $p \in \mathrm{Int}(D)$, $q_1 \in \partial D$ et $q_2 \in \partial D$.

Vu que tout Γ est d'un seul côté de ses tangentes, lesdites tangentes ne passent pas par l'intérieur de D. Ni p ni q_2 ne sont sur ℓ_{q_1}, parce que si $q_2 \in \ell_{q_1}$ alors $\ell_{q_1} = \ell$, ce qui est impossible parce que ℓ passe par $p \in \mathrm{Int}(D)$.

Or p et q_2 sont de part et d'autres de ℓ_{q_1}, ce qui est impossible parce que Γ est d'un seul côté de cette droite. □

PROPooORPooCXrIQi

Proposition 21.95.
Soit D l'ensemble défini en (21.310).

(1) $D = \mathcal{C}\text{onv}(\Gamma)$ *($\mathcal{C}\text{onv}(\Gamma)$ désigne l'enveloppe convexe de Γ)*

(2) $\partial\mathcal{C}\text{onv}(\Gamma) = \Gamma$.

Pour la définition d'enveloppe convexe, voir la définition 11.38.

Démonstration. L'ensemble D est un convexe contenant Γ. Donc $\mathcal{C}\text{onv}(\Gamma) \subset D$. L'inclusion inverse est à prouver.

Soit $x \in D$. Si $x \in \partial D$ alors $x \in \Gamma$ (proposition 21.92) et donc $x \in \mathcal{C}\text{onv}(\Gamma)$. Nous ne devons donc traiter que le cas $x \in \text{Int}(D)$.

Le corolaire 21.94 nous dit que toute droite passant par x coupe Γ en exactement deux points. Soient p et q ces points. Alors $p, q \in \Gamma \subset \mathcal{C}\text{onv}(\Gamma)$. Vu que $\mathcal{C}\text{onv}(\Gamma)$ est convexe, tout le segment $[p, q]$ est dans $\mathcal{C}\text{onv}(\Gamma)$, et en particulier $x \in \mathcal{C}\text{onv}(\Gamma)$.

Nous passons à la seconde affirmation. Nous savons que $D = \mathcal{C}\text{onv}(\Gamma)$. En prenant le bord des deux côtés, $\partial D = \partial\mathcal{C}\text{onv}(\Gamma)$, donc $\Gamma = \partial\mathcal{C}\text{onv}(\Gamma)$. □

LEMooUEKQooWhGyKn

Lemme 21.96 (Des tangentes parallèles[610])**.**
Une courbe fermée γ de classe C^1 est convexe si et seulement si elle ne possède pas 3 tangentes parallèles distinctes.

Démonstration. Si le graphe Γ possédait trois tangentes parallèles distinctes, une serait entre les deux autres et le graphe Γ serait de part et d'autres de cette tangente. Dans ce cas, γ n'est pas convexe.

Nous montrons maintenant que si γ n'est pas convexe, alors elle possède trois tangentes parallèles distinctes. Pour $p \in \Gamma$ tel que Γ soit des deux côtés de l_p.

Soient Γ_1 et Γ_2 les parties de Γ délimitées par l_p. Nous notons q_i le point de Γ_i le plus éloigné de la droite l_p, il existe parce que Γ est compact et que la fonction distance à l_p est continue sur Γ. Les points p, q_1 et q_2 sont distincts (sinon l_p ne couperait pas Γ en deux parties).

Montrons que $l_{q_i} \parallel l_p$. Pour cela nous choisissons un système d'axe dans lequel $l_p \equiv y = 0$. Dans ce système, la distance entre $\gamma(t)$ et l_p est $\gamma_y(t)$ et les extrémums de cette fonction ont lieu aux points t tels que $\gamma_y'(t) = 0$, c'est-à-dire aux points sur lesquels la tangente est parallèle à l_p.

À quel moment avons nous utilisé le fait que la courbe soit fermée ? Au moment de dire que le point le plus éloigné devait vérifier $\gamma_y'(t) = 0$. En effet un extrémum peut ne pas vérifier cette condition si il n'est pas à l'intérieur du domaine. Dans notre cas, nous avons $\gamma\colon [0, 1] \to \mathbb{R}^2$ qui est fermée : $\gamma(0) = \gamma(1)$. Donc en réalité nous pouvons considérer $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ et tous les points du domaine sont intérieurs au domaine. L'extrémum doit donc vérifier la condition d'annulation de la dérivée. □

Exemple 21.97.
Si la courbe n'est pas fermée, alors le lemme 21.96 ne tient pas comme le montre le contre-exemple du graphe de $f(x) = x^3 - x$. Il est non convexe et pourtant ne présente que 2 tangentes parallèles (dans chaque direction). △

LEMooCSXCooIDPiKW

Lemme 21.98.
Soit γ une courbe convexe et ℓ une droite qui intersecte Γ mais qui n'est pas tangente. Alors l'intersection entre ℓ et Γ comprend au maximum 2 points.

Démonstration. Supposons que ℓ coupe Γ en trois points p, q, r (dans cet ordre). Vu que ℓ n'est pas tangente à Γ, la tangente ℓ_q est distincte de ℓ (et intersecte ℓ en l'unique point q). Les points p et r sont dans Γ et sont pourtant de deux côtés différents de ℓ_q. Contradiction avec la convexité de γ. □

La proposition suivante nous dit que si deux points de Γ ont la même tangente, alors entre ces deux points, Γ est le segment de droite les joignant.

PROPooCKTZooIPcUca

Proposition 21.99 ([1]).
Soit $\gamma\colon [0,L]\to\mathbb{R}^2$ une courbe fermée simple et convexe de classe C^1. Si la tangente en $p=\gamma(s_p)$ et la tangente en $q=\gamma(s_q)$ sont identiques (pas seulement parallèles), alors soit

$$\gamma\big([s_p,s_q]\big)=[p,q] \tag{21.314}$$

soit

$$\gamma\big([s_q,L]\big)=[p,q] \tag{21.315}$$

Démonstration. Nous considérons un système d'axe dans lequel $p=(0,0)$, $\ell_p=\ell_q\equiv y=0$ et tel que $\gamma_y(t)\geqslant 0$ pour tout t. Nous choisissons enfin un paramétrage de γ tel que $p=\gamma(0)$.

Supposons un instant que $\gamma_y\big([0,s_q]\big)=\{0\}$. Nous avons $\gamma_x(0)=p_x=0$ et $\gamma_x(s_q)=q_x$. Donc le théorème des valeurs intermédiaires 9.93 implique que toutes les valeurs entre 0 et q_x sont atteintes par γ_x entre 0 et s_q. Donc tous les points du type $(t,0)$ avec $0\leqslant t\leqslant q_x$ sont atteints par $\gamma(t)$ avec $0\leqslant t\leqslant s_q$. De plus aucun autre point ne peut être atteint parce que γ étant simple, elle ne peut pas faire de retour en arrière.

Nous supposons donc que γ_y n'est pas identiquement nulle sur $[0,s_q]$; il existe donc un s_M avec $0<s_M<s_q$ qui maximise γ_y sur $[0,s_q]$. La tangente à Γ en $\gamma(s_M)$ est horizontale. Les droites $y=0$ et $y=\gamma_y(s_M)$ sont donc deux tangentes parallèles à Γ. Par le lemme des tangentes parallèles 21.96, il n'y a pas d'autres tangentes horizontales. Donc pour tout $n\in\mathbb{N}$ la droite $y=\frac{1}{n}$ n'est tangente nulle part [28] à Γ.

Par le lemme 21.98, la droite $y=\frac{1}{n}$ ne peut intersecter Γ qu'en seulement deux points. Or le théorème des valeurs intermédiaires appliqué à γ_y sachant que $\gamma_y(0)=\gamma_y(s_q)=0$ et $\gamma_y(s_M)>0$ nous donne a_n,b_n tels que

$$0<a_n<s_M \tag{21.316a}$$
$$s_M<b_n<s_q \tag{21.316b}$$

et $\gamma_y(a_n)=\gamma_y(b_n)=\frac{1}{n}$. Donc la droite $y=\frac{1}{n}$ intersecte γ deux fois dans $[0,s_q]$. En conséquence de quoi $\gamma_y(t)<\frac{1}{n}$ pour tout $t\in[s_q,L]$. Cela étant valable pour tout n nous avons $\gamma_y\big([s_q,L]\big)=\{0\}$ et nous sommes ramenés essentiellement au premier cas. □

PROPooKHUQooIOUxFw

Proposition 21.100.
Soit une courbe fermée simple γ de classe C^1 et une droite ℓ. Alors il existe au moins deux points distincts q_1,q_2 tels que $\ell_{q_1}\parallel\ell_{q_2}\parallel\ell$ avec $\ell_{q_1}\neq\ell_{q_2}$.

Démonstration. Nous considérons un système d'axes dans lequel la droite ℓ est la droite $y=0$ [29]. Nous considérons la fonction

$$\begin{aligned} r\colon \gamma &\to \mathbb{R}\\ p &\mapsto p_y. \end{aligned} \tag{21.317}$$

En tant que fonction continue sur un compact, elle possède un minimum et un maximum. Les extrémums arrivent en $\gamma(s)$ avec $\gamma_y'(s)=0$, ce qui signifie que les tangentes aux extrémums sont parallèles à ℓ. Vu que le maximum et le minimum ne peuvent pas être égaux (sinon la courbe serait horizontale et pas simple), les tangentes en ces points sont distinctes. □

Corolaire 21.101.
Soit une courbe convexe fermée simple γ de classe C^2. L'ensemble $\mathcal{C}\mathrm{onv}(\Gamma)$ est compact.

Démonstration. Nous savons que $\mathcal{C}\mathrm{onv}(\Gamma)$ n'est autre que D par la proposition 21.95. Nous savons déjà que D est fermé. Il nous suffit donc de prouver qu'il est borné (théorème de Borel-Lebesgue 9.24). Nous considérons deux droites perpendiculaires et les 4 tangentes correspondantes par la proposition 21.100. Vu que Γ est d'un seul côté de chacune de ces tangentes, elle est contenue dans le rectangle délimité par ces 4 droites. □

28. À part $n=0$ et si par manque de chance, $\gamma_y(s_M)$ est un nombre de la forme $1/n$.
29. Pour bien faire, il faut faire la démonstration en deux parties : une qui montre le cas particulier $\ell\equiv y=0$, et un second qui en déduit rigoureusement le cas général.

21.13.4 Courbure et convexité

SUBSECooNJOLooYuGRjA

À propos de paramétrage normal.

LEMooHMFSooFlhanD

Lemme 21.102.
Soit $\gamma\colon [0,L]\to\mathbb{R}^2$ une courbe simple, fermée de classe C^2 en paramétrage normal. Nous considérons sa fonction d'angle $\theta\colon [0,L]\to\mathbb{R}$.

Alors l'application

$$\begin{aligned}\sigma\colon [0,L]&\to S^1\\ t&\mapsto \begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix}\end{aligned}\tag{21.318}$$

est surjective.

Démonstration. Le théorème 21.88 nous dit que si $\theta(0)=a$ alors $\theta(2\pi)$ ne peut valoir que $a+2\pi$ ou $a-2\pi$. Le théorème des valeurs intermédiaires nous dit alors que θ prend toutes les valeurs entre a et $a+2\pi$ ou $a-2\pi$. □

PROPooWXUKooPOtPdj

Proposition 21.103 ([610]).
Une courbe fermée simple de classe C^2 est convexe si et seulement si sa courbure est soit toujours positive, soit toujours négative.

Démonstration. Nous considérons la courbe γ munie d'un paramétrage de vitesse 1, c'est-à-dire avec $\|\gamma'(t)\|=1$ pour tout t. Si θ est sa fonction d'angle[30], alors nous avons $\theta'=\kappa$ par le lemme 21.80. Donc la fonction θ est monotone si et seulement si la courbure ne change pas de signe. Nous allons montrer que θ est monotone si et seulement si γ est convexe.

(1) $\Rightarrow$

Nous supposons que θ est monotone et γ non convexe. Soient $p\in\Gamma$ tel que Γ soit des deux côtés de ℓ_p, et soient les points q_1, q_2 donnés par le lemme des tangentes parallèles 21.96 tels que $\ell_p\parallel\ell_{q_1}\parallel\ell_{q_2}$. Parmi les vecteurs tangents en p, q_1 et q_2, deux au moins ont la même direction ; supposons que ce sont q_1 et q_2. C'est-à-dire que si $p=\gamma(s_0)$, $q_1=\gamma(s_1)$ et $q_2=\gamma(s_2)$ alors nous avons $\gamma'(s_1)=\gamma'(s_2)$ et donc aussi

$$\theta(s_1)=\theta(s_2)+2n\pi\tag{21.319}$$

pour un certain n. Mais θ est monotone et la différence entre sa première et sa dernière valeur doit valoir 2π ou -2π par le théorème 21.88. Donc n ne peut valoir que -1, 0 et 1.

Si $n=0$ alors θ est constante sur $[s_1,s_2]$. Si $n=1$ alors $\theta(s_1)=\theta(s_2)+2\pi$ alors que sur toute la courbe, θ ne peut faire que 2π. Donc θ est constant sur $[0,s_1]$ et sur $[s_2,L]$ (où L est le bord de le paramétrage). Si $n=-1$, même conclusion.

Dans tous les cas, Γ contient une ligne droite, soit de q_1 à q_2, soit de q_2 à q_1. Et dans ces cas nous avons $\ell_{q_1}=\ell_{q_2}$, ce qui est contraire à la construction de q_i.

Nous concluons que γ est convexe.

(2) $\Leftarrow$

Nous supposons que γ est convexe, mais que θ n'est pas monotone. Il existe donc $s_1<s_0<s_2$ tels que

$$\theta(s_1)=\theta(s_2)\neq\theta(s_0).\tag{21.320}$$

Et vu le lemme 21.102, il existe s_3 tel que $\gamma'(s_3)=-\gamma'(s_1)$.

Donc en s_1, s_2 et s_3 nous avons trois tangentes parallèles. La proposition 21.96 est alors formelle, γ étant convexe, deux de ces tangentes doivent être identiques.

La proposition 21.99 dit qu'entre deux points dont les tangentes sont identiques, la courbe doit être un segment de droite. Or sur un segment de droite, $\kappa=0$ et θ est constante.

30. Définition 21.79.

— La partie $\gamma\big([s_1,s_2]\big)$ ne peut pas être droite parce que nous avions supposé l'existence d'un $s_0\in]s_1,s_2[$ tel que $\theta(s_1)=\theta(s_2)\neq\theta(s_0)$.

— La partie $\gamma\big([s_1,s_3]\big)$ ne peut pas être droite parce que $\theta(s_3)\neq\theta(s_1)$.

— La partie $\gamma\big([s_2,s_3]\big)$ ne peut pas être droite parce que $\theta(s_2)\neq\theta(s_1)$.

Nous sommes donc devant une contradiction.

Nous en concluons que θ doit être monotone.

□

21.13.5 Théorème des quatre sommets

LEMooELIRooNDVXPh

Lemme 21.104.
Soit une droite ℓ du plan. Il existe $a,c\in\mathbb{R}^2$ avec $c\neq 0$ tels que $z\in\ell$ si et seulement si $(z-a)\cdot c=0$.

Démonstration. Une droite est paramétrée par $\gamma(t)=p+tq$. En posant $a=p$ et $c=Jq$ nous avons la réponse. En effet nous allons montrer qu'avec ces valeurs de a et c, nous avons $z\in\Gamma$ si et seulement si $(z-a)\cdot c=0$.

D'abord un point de γ est de la forme $z=\gamma(t)=p+tq$. Nous avons :

$$\big(\gamma(t)-a\big)\cdot c=\big(\gamma(t)-p\big)\cdot Jq=tq\cdot Jq=0. \tag{21.321}$$

Et dans l'autre sens, si $(z-a)\cdot c=0$ nous devons prouver que $z\in\Gamma$. Nous avons : $(z-p)\cdot Jq=0$, ce qui fait que $z-p$ est un multiple de q. Autrement dit : $z-p=\lambda q$ ou encore $z=\alpha q+p$, qui est sur la droite Γ. □

Définition 21.105.
*Un **sommet** d'une courbe est un point d'extrémum local de la courbure.*

THOooFRBBooWKZcfY

Théorème 21.106 (Théorème des quatre sommets[104, 543])**.**
Soit un arc paramétrique $\gamma\colon\mathbb{R}\to\mathbb{R}^2$ fermé, simple et convexe[31] *de classe C^3 et T-périodique.*

Alors γ possède au moins 4 points critiques sur chaque période.

Démonstration. Nous supposons que le paramétrage de γ soit normal.

Si la courbure κ est constante sur une partie ouverte de la (du paramétrage de la) courbe, alors tous les points de cette partie sont des sommets et le théorème est fait. Nous supposons que κ n'est pas constante et en particulier que Γ ne contient ni bouts de droites ni bouts de cercles (théorème 21.76).

La fonction κ étant de classe C^1 sur le compact Γ, elle admet au moins un maximum et un minimum distincts. Vu que ces points sont intérieurs, ils correspondent au changement de signe de κ'. Soient p et q ces points. Pour la simplicité nous supposons que γ est paramétré de telle sorte que $\gamma(0)=p$, et $q=\gamma(s_q)$ avec $0<s_q<T$.

Nous supposons que p et q sont les seuls points de changement de signe de κ'.

Soit ℓ la droite passant par p et q. Tous les points du segment $[p,q]$ (qui sont dans $\mathcal{C}\text{onv}(\Gamma)$) ne peuvent pas être sur Γ (sinon nous aurions un morceau de droite). Donc certains points sont dans l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$. Donc la droite ℓ passe par l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$ et le corolaire 21.94 nous dit que la droite ℓ ne coupe Γ en seulement deux points.

Par conséquent, les ensembles $\gamma\big([0,s_q]\big)$ et $\gamma\big([s_q,T]\big)$ sont de part et d'autre de Γ. Vu qu'en ces points, κ' change de signe et qu'il ne change de signe en aucun autre points, la fonction κ' est positive d'un côté de ℓ et négative de l'autre.

D'autre part par le lemme 21.104, il existe $a\in\mathbb{R}^2$ et $c\neq 0$ tels que $z\in\ell$ si et seulement si $(z-a)\cdot c=0$. La fonction $z\mapsto(z-a)\cdot c$ est donc positive d'un côté de ℓ et négative de l'autre.

31. Par la proposition 21.103 nous pouvons aussi bien demander à la courbure d'être toujours strictement positive, comme le fait [104].

En résumé les fonctions

$$s \mapsto \kappa'(s) \tag{21.322a}$$
$$s \mapsto \big(\gamma(s) - a\big) \cdot c \tag{21.322b}$$

changent de signe en même temps et le produit a donc un signe constant. Ce produit n'est de plus pas nul parce que κ' n'est nul sur aucun intervalle (sinon κ y serait constant et Γ un segment de droite) et $\big(\gamma(s) - a\big) \cdot c$ ne s'annule pour aucun s sauf ceux qui correspondent à p et q.

Nous avons donc

$$0 \neq \int_0^T \kappa'(s)\big(\gamma(s) - a\big) \cdot c\, ds \tag{21.323a}$$
$$= \underbrace{\Big[\big(\gamma(s) - a\big) \cdot c\kappa(s)\Big]_0^T}_{A=0} - \int_0^T \kappa(s)\gamma'(s) \cdot c\, ds \tag{21.323b}$$
$$= -\int_0^T \kappa(s)\big(\gamma'(s) \cdot c\big) ds \tag{21.323c}$$
$$= \int_0^T J\gamma''(s) \cdot c\, ds \tag{21.323d}$$
$$= J\int_0^T \gamma''(s) \cdot c\, ds \tag{21.323e}$$
$$= J\big[\gamma'(s) \cdot c\big]_0^T \tag{21.323f}$$
$$= 0. \tag{21.323g}$$

Justifications :

— L'expression A est nulle parce que les valeurs en 0 et en T sont identiques.

— Nous utilisons le lemme 21.75 pour faire $-\kappa(s)\gamma'(s) = J\gamma''(s)$.

Le tout est une contradiction de la forme $0 \neq a = 0$.

Nous avons donc au moins un troisième point de changement de signe de κ'. Vu que la courbe est périodique, il en faut un nombre pair et donc un quatrième. □

L'exemple de l'ellipse montre qu'il n'y a pas lieu de chercher d'autres extrémums de κ à part les 4 déjà trouvés.

Exemple 21.107.
Nous trouvons les sommets de l'ellipse.

$$\gamma(t) = \big(a\cos(t), b\sin(t)\big) \tag{21.324a}$$
$$\gamma'(t) = \big(-a\sin(t), b\cos(t)\big) \tag{21.324b}$$
$$\gamma''(t) = -\big(a\cos(t), b\sin(t)\big) \tag{21.324c}$$
$$\tag{21.324d}$$

La courbure est

$$\kappa(t) = \frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|^3} \tag{21.325a}$$
$$= \frac{-1}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}} \begin{pmatrix} a\cos(t) \\ b\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -b\cos(t) \\ -a\sin(t) \end{pmatrix} \tag{21.325b}$$
$$= \frac{ab}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}}. \tag{21.325c}$$

Vu que $ab > 0$, les extrémums de cela sont ceux du dénominateur et il suffit donc d'étudier les extrémums de

$$f(t) = a^2\sin^2(t) + b^2\cos^2(t). \tag{21.326}$$

Nous avons

$$f'(t) = 2(a^2 - b^2)\cos(t)\sin(t), \tag{21.327}$$

fonction qui s'annule effectivement 4 fois sur une période. Deux maximums et deux minima. △

21.13.6 Espace topologique normal

DEFooNNKVooLtzImT

Définition 21.108.
*Un **espace topologique normal** est un espace topologique séparé dont les ouverts séparent strictement les fermés disjoints.*

Autrement dit, si F_1 et F_2 sont des fermés disjoints de X, alors il existe des ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ tels que $F_1 \subset \mathcal{O}_1$ et $F_2 \subset \mathcal{O}_2$.

PROPooWSKZooHKvuob

Proposition 21.109.
L'espace topologique $\mathbb{R}^n$ est normal.

¡¡ Avertissement/question à la lectrice !! 21.110
J'ai écrit le lemme 21.111 de façon adhoc pour une utilisation dans le théorème d'Urysohn. J'avoue n'être pas certain qu'il soit vrai. Si vous en trouvez une preuve, écrivez-moi.

LEMooVVASooEXYKOZ

Lemme 21.111 ([1]).
Soit un espace topologique normal X. Si $\mathcal{O}$ est ouvert dans X et si F est un fermé contenu dans $\mathcal{O}$, alors il existe un ouvert U tel que

$$F \subset U \subset \bar{U} \subset \mathcal{O}. \tag{21.328}$$

21.13.7 Théorème d'Urysohn

LEMooWRLRooCdubfZ

Lemme 21.112.
Soit un espace topologique X. Soit une partie D dense de $[0,1]$. Nous considérons des ouverts $\{U_s\}_{s\in D}$ de X indexés par D. Nous supposons que si $s_1 < s_2$ alors $\bar{U}_{s_1} \subset U_{s_2}$. Enfin pour tout $x \in X$ nous posons

$$A_x = \{s \in D \text{ tel que } x \in U_s\}. \tag{21.329}$$

Alors l'application

$$\begin{aligned} f\colon X &\to [0,1] \\ x &\mapsto \begin{cases} \inf(A_x) & \text{si } x \in \bigcup_{s\in D} U_s \\ 1 & \text{sinon.} \end{cases} \end{aligned} \tag{21.330}$$

est continue.

Démonstration. Dans la suite nous prenons les conventions suivantes :

$$\bigcup_{s\in\emptyset} U_s = \emptyset \tag{21.331a}$$

$$\bigcap_{s\in\emptyset} U_s = X \tag{21.331b}$$

(21.331c)

Soit $t \in [0,1]$.

(1) $\{x \text{ tel que } f(x) < t\} = \bigcup_{s<t} U_s$ D'abord nous considérons $x \in X$ tel que $f(x) < t$, et nous prouvons que $x \in \bigcup_{\substack{s\in D\\ s<t}} U_s$. Vu que $f(x) = \inf(A_x) < t \leqslant 1$, il existe $s_0 < t$ tel que $s_0 \in A_x$. Pour ce s_0 nous avons $x \in A_{s_0}$, et donc

$$x \in U_{s_0} \subset \bigcup_{\substack{s\in D\\ s<t}} U_s. \tag{21.332}$$

Dans l'autre sens, soient $s_0 < t$ et $x \in U_{s_0}$. Nous avons $s_0 \in A_x$ et donc $f(x) \leqslant s_0 < t$.

(2) $\{x$ **tel que** $f(x) \leqslant t\} = \bigcup_{s>t} \bar{U}_s$ Si $t = 1$, alors $\{x$ tel que $f(x) \leqslant t\} = X$, et $\bigcap_{s>t} \bar{U}_s = \bigcup_{s\in\emptyset} \bar{U}_s = X$. Supposons que $t < 1$.
Soit $x \in X$ tel que $f(x) \leqslant t$. Vu que $t < 1$ et que D est dense dans $[0, 1]$, il existe $s \in D$ tel que

$$s > t \geqslant f(x) = \inf(A_x). \tag{21.333}$$

Donc il existe $a \in A_x$ tel que $a \in]\inf(A_x), s[$. Vu que $a \in A_x$, nous avons $x \in U_a$, et vu que $a < s$, nous avons $\bar{U}_a \subset U_s$. Donc

$$x \in U_a \subset \bar{U}_a \subset U_s. \tag{21.334}$$

Donc $x \in U_s$ pour tout $s > t$ (dans D), et donc $x \in \bigcap_{s>t} \bar{U}_s$. Nous avons donc déjà prouvé que

$$\{x \text{ tel que } f(x) \leqslant t\} \subset \bigcap_{\substack{s\in D\\ s>t}} U_s. \tag{21.335}$$

Nous faisons à présent l'inclusion dans l'autre sens. Soit $x \in \bigcap_{\substack{s\in D\\ s>t}} \bar{U}_s$, et montrons que $f(x) \leqslant t$. Nous avons déjà traité le cas $t = 1$; nous continuous à supposer que $t < 1$.
Soit $\epsilon > 0$ assez petit pour avoir $t + \epsilon < 1$. Vu que D est dense dans $[0, 1]$, il existe $s_1, s_2 \in D$ tels que

$$t < s_1 < s_2 < t + \epsilon. \tag{21.336}$$

Par hypothèse $x \in \bar{U}_{s_1}$ et donc

$$x \in \bar{U}_{s_1} \subset U_{s_2}. \tag{21.337}$$

Cela prouve que $s_2 \in A_x$, et nous avons alors

$$f(x) = \inf(A_x) \leqslant s_2 < t + \epsilon. \tag{21.338}$$

Étant donné que ϵ est arbitrairement petit, nous avons $f(x) \leqslant t$.

(3) $\{x$ **tel que** $f(x) < t\}$ **est ouvert pour tout** $t \in \mathbb{R}$ Nous commençons par $t \in [0, 1]$. Dans ce cas nous avons vu que

$$\{x \text{ tel que } f(x) < t\} = \bigcup_{\substack{s\in D\\ s<t}} U_s \tag{21.339}$$

qui est une union d'ouverts, et donc qui est ouvert.
Si $t < 0$, alors

$$\bigcup_{s<t} U_s = \bigcup_{s\in\emptyset} U_s = \emptyset, \tag{21.340}$$

est ouvert. Si $t > 1$, alors $\{x$ tel que $f(x) < t\} = X$ parce que f prend ses valeurs dans $[0, 1]$.

(4) $\{x$ **tel que** $f(x) \leqslant t\}$ **est fermé pour tout** $t \in \mathbb{R}$ Si $t \in [0, 1]$, alors nous savons déjà que

$$\{x \text{ tel que } f(x) \leqslant t\} = \bigcup_{s>t} \bar{U}_s \tag{21.341}$$

qui est fermé comme intersection de fermés (lemme 7.6(1)). Vu que f prend ses valeurs dans $[0, 1]$, nous avons $\{x$ tel que $f(x) < t\} = \emptyset$ lorsque $t < 0$ et $\{x$ tel que $f(x) < t\} = X$ lorsque $t > 1$.

f est contiue Soient $a, b \in \mathbb{R}$ avec $a < b$. Nous avons,

$$f^{-1}\big(]a, b[\big) = \{f < b\}\backslash\{f \leqslant a\} = \{f < b\} \cap \{f \leqslant a\}^c. \tag{21.342}$$

Mais nous savons que $\{f < b\}$ est ouvert et que $\{f \leqslant a\}$ est fermé. Donc $f^{-1}\big(]a, b[\big)$ est ouvert comme intersection d'ouverts.
La proposition 7.138 dit (entre autres) que les $]a, b[$ forment une base de la topologie de $\mathbb{R}$. La proposition 7.38 dit que cela suffit pour conclure à la continuité de f.

□
THOooKYYEooLFcNpg

Théorème 21.113 (Théorème d'Urysohn[205]).
Soit un espace topologique normal X. Soient A et B, des fermés disjoints de X. Il existe une application continue $f\colon X \to [0,1]$ telle que $f|_A = 0$ et $f|_B = 1$.

Démonstration. Pour chaque $m \in \mathbb{N}$ nous considérons $D_m = \{k/2^m\}_{k=0,\dots,2^m}$, et posons $D = \bigcup_{m=0}^{\infty} D_m$.

(1) **D est dense dans $[0,1]$** Soient $x \in [0,1]$ et $\epsilon > 0$. Considérons m tel que $1/2^m < \epsilon$. Dans ce cas, il existe $s \in D_m$ tel que $|s - x| < \frac{1}{2^m} < \epsilon$.

(2) **Construction des U_s** Nous allons maintenant construire un ouvert U_s pour chaque $s \in D$. Et nous allons le faire de telle sorte d'avoir $s_1 < s_2 \Rightarrow \bar{U}_{s_1} \subset U_{s_2}$.

(3) **Pour $s \in D_0$** Nous commençons avec les éléments de $D_0 = \{0,1\}$. D'abord nous posons

$$U_1 = B^c. \tag{21.343}$$

Vu que B est fermé, c'est un ouvert.
Pour construire U_0, nous remarquons que A est un fermé disjoint de B dans l'espace normal X. Il existe donc des ouverts U_0 et U_0' qui séparent A et B, c'est-à-dire tels que $A \subset U_0$, $B \subset U_0'$ et $U_0 \cap U_0' = \varnothing$.
Nous avons $\bar{U}_0 \subset B^c$. En effet si $a \in \bar{U}_0 \cap B$, alors $a \in B$ et donc U_0' est un voisinage ouvert de a. Ce voisinage n'intersectant pas U_0, a n'est pas dans la fermeture de U_0.
Nous avons donc bien la relation

$$\bar{U}_0 \subset B^c = U_1. \tag{21.344}$$

(4) **Récurrence** Nous avons défini U_s pour $s \in D_0$. Pour la récurrence, nous supposons que U_s est défini pour $s \in D_{m-1}$, et qu'ils vérifient la condition. Nous allons construire les U_s pour $s \in D_m$.
Remarquons que $D_{m-1} \subset D_m$. Plus précisément, les éléments de D_m qui ne sont pas dans D_{m-1} sont ceux de la forme $k/2^m$ avec k impair.
Soit donc k impair et considérons $s = k/2^m$. Par hypothèse de récurrence nous avons déjà

$$\bar{U}_{(k-1)/2^m} \subset U_{(k+1)/2^m}. \tag{21.345}$$

Le lemme 21.111 nous permet de considérer un ouvert U tel que

$$\bar{U}_{(k-1)/2^m} \subset U \subset \bar{U} \subset U_{(k+1)/2^m}. \tag{21.346}$$

Nous posons $U_{k/2^m} = U$ et nous prouvons que $\bar{U}_{s_1} \subset U_{s_2}$ dès que $s_1 < s_2$ dans D_m.
Soient $s_1 = k/2^m$ et $s_2 = l/2^m$ avec $k < l$. Il y a quatre possibilités suivant la parité de k et l.

(4a) **Si k et l sont pairs** Alors s_1 et s_2 sont dans D_{m-1} et l'hypothèse de récurrence dit $\bar{U}_{s_1} \subset U_{s_2}$.

(4b) **Si k est impair et l est pair** Nous notons $k = 2k' + 1$ et $l = 2l'$. Vu que $k < l$, nous avons $k' + \frac{1}{2} < l'$. Mais comme k' et l' sont entiers, nous avons $k' + 1 \leqslant l'$. Et nous sommes partis pour le calcul :

$$\begin{aligned}
\bar{U}_{s_1} &= \bar{U}_{k/2^m} && && \text{(21.347a)}\\
&\subset U_{(k+1)/2^m} && \text{construction pour } k \text{ pair} && \text{(21.347b)}\\
&= U_{(k'+1)/2^{m-1}} && && \text{(21.347c)}\\
&\subset U_{l'/2^{m-1}} && \text{hyp. rec. dans } D_{m-1} && \text{(21.347d)}\\
&= U_{l/2^m} && && \text{(21.347e)}\\
&= U_{s_2}. && && \text{(21.347f)}
\end{aligned}$$

(4c) **Si k est pair et l est impair** Laissé aux bons soins du lecteur. J'ai pas vérifié, mais [205] dit que c'est similaire. Écrivez-moi si vous voyez un problème.

(4d) **Si k et l sont impairs** Laissé aux bons soins du lecteur. J'ai pas vérifié, mais [205] dit que c'est similaire. Écrivez-moi si vous voyez un problème.

(5) **Résumé** Nous avons posé $D_m = \{k/2^m\}_{k=0,\dots,2^m}$, et $D = \bigcup_{m=0}^{\infty} D_m$. Nous avons prouvé que D est dense dans $[0,1]$, et nous avons construit, pour chaque $s \in D$ un ouvert U_s de X, de telle façon que

$$\bar{U}_{s_1} \subset U_{s_2} \tag{21.348}$$

pour tout $s_1 < s_2$. Bref, le lemme 21.112 est applicable. Nous posons donc

$$\begin{aligned} f\colon X &\to [0,1] \\ x &\mapsto \begin{cases} \inf(A_x) & \text{si } x \in \bigcup_{s\in D} U_s \\ 1 & \text{sinon.} \end{cases} \end{aligned} \tag{21.349}$$

avec $A_x = \{s \in D \text{ tel que } x \in U_s\}$.

(6) **Ce f est ce qu'il nous fait** Si $x \in B$, alors x n'est pas dans U_1 et donc n'est dans aucun des U_s. Donc $f(x) = 1$.

Si $x \in A$, alors nous avons $x \in U_0$ (parce que $A \subset U_0$). En particulier $0 \in A_x$, et donc $f(x) = \inf(A_x) = 0$.

□

21.114.
Les fractions de la forme $a/2^b$ sont nommées **fractions dyadiques** et nous avons déjà vu qu'elles sont denses dans $\mathbb{R}$ dans le corolaire 1.446.

Chapitre 22

Géométrie hyperbolique

22.1 Inversion

LEMooNLYKooKBMdDk

Lemme-Définition 22.1.
Soit un cercle C et un point A du plan $\mathbb{R}^2$. Soit une droite passant par A et coupant C en deux points P et P' (pas spécialement distincts). Alors le nombre

$$\|AP\|\|AP'\| \tag{22.1}$$

*ne dépend pas du choix de la droite et est nommé la **puissance** du point A par rapport au cercle C.*

22.1.1 Cercles perpendiculaires

Définition 22.2.
*Deux cercles sont **perpendiculaires** lorsque leurs tangentes aux points d'intersection sont perpendiculaires.*

LEMooWMGOooAieUjD

Lemme 22.3 ([1])**.**
Soient deux cercles perpendiculaires C_1 et C_2. Alors

ITEMooJVYYooSrlSdA

(1) le centre de C_1 est hors de C_2.

ITEMooTQATooNWvllF

(2) Si ℓ est une droite passant par ce le centre de C_1 (nommé O) et si ℓ coupe C_2 en les deux points P et P', alors P et P' sont situés du même côté de O.

Démonstration. Nous nommons O_1 le centre de C_1 ainsi que Q et Q' les points d'intersection de C_1 avec C_2. Si ℓ_Q et $\ell_{Q'}$ sont les tangentes à C_1 en Q et Q', alors ce sont des rayons de C_2 (parce que les cercles sont perpendiculaires). Par conséquent le centre O_2 de C_2 est le point d'intersection $O_2 = \ell_Q \cap \ell_{Q'}$.

Le triangle O_1O_2Q est rectangle en Q et donc $\|O_1O_2\| > \|QO_2\|$. Or le nombre $\|QO_2\|$ est le rayon de C_2, donc O_1 est en dehors de C_2.

Ceci achève de prouver le point (1) ; nous démontrons le point (2). Les points Q et Q' sont sur le cercle C_2, donc tous les points du segment $[QQ']$ sont dans le cercle. Or le centre de C_1 doit être en dehors de C_2 ; il ne peut donc pas être dans le segment $[QQ']$, ce qui prouve que Q et Q' ne sont pas de part et d'autre de O sur la droite (OQQ'). □

PROPooYSVYooOFKxib

Proposition 22.4 ([1])**.**
Soit un cercle C_1 de centre O et de rayon R.

ITEMooWYIJooAAmXUl

(1) Un cercle C_2 est perpendiculaire à C_1 si et seulement si il existe une droite ℓ passant par O telle que les points d'intersection $\{P, P'\} = \ell \cap C_2$ soient situés du même côté de O et vérifient

$$\|OP\|\|OP'\| = R^2. \tag{22.2}$$

ITEMooYKZOooYZKyhF

(2) *Dans ce cas, toutes les droites passant par O et coupant C_2 en deux points P, P' vérifient le fait que P et P' soient du même côté de O et $\|OP\|\|OP'\| = R^2$.*

Démonstration. Nous commençons par prouver le point (2). Le fait que les points P et P' soient du même côté de O est le lemme 22.3(2). Pour la relation sur les distances, soit $Q \in C_1 \cap C_2$. Vu que C_1 et C_2 sont perpendiculaires, la droite (OQ) ne coupe C_2 qu'en ce point, et la puissance de O par rapport au cercle C_2 est $\|OQ\|^2 = R^2$.

La même puissance peut être calculée via la droite ℓ :

$$\|OP\|\|OP'\|. \tag{22.3}$$

Donc $\|OP\|\|OP'\| = R^2$.

Soit un cercle C_2 passant par P et P'. Notons que P et P' ne sont pas sur C_1 parce qu'ils ne pourraient pas être alignés avec O. De plus l'un est à l'intérieur de C_1 et l'autre à l'extérieur de C_1. Les cercles C_1 et C_2 possèdent donc deux points distincts d'intersections. La puissance de O par rapport à C_2 est :

$$\|OP\|\|OP'\| = R^2 \tag{22.4}$$

parce que (OP) est une droite coupant C_2 en les points P et P'.

Soit Q un point d'intersection de C_1 et C_2, et Q' l'autre point d'intersection de C_2 avec la droite (OQ). La puissance de O par rapport à C_2 peut également être calculée à partir de cette droite (lemme 22.1) et nous avons

$$\|OQ\|\|OQ'\| = R^2, \tag{22.5}$$

mais $Q \in C_1$, donc $\|OQ\| = R$ et partant $\|OQ'\| = R$. Nous en déduisons que $Q' \in C_1$ également. Or Q' ne peut pas être l'autre point d'intersection de C_1 avec C_2 (sinon O, Q, Q' ne seraient pas alignés). Donc $Q = Q'$ et nous déduisons que la droite (OQ) est tangente à C_2.

Le rayon de C_1 est tangent à C_2. Cela signifie que C_1 est perpendiculaire à C_2. □

22.1.2 Inversion

PROPDEFooVLIWooQgpLQa

Proposition-Définition 22.5 ([545]).
Soit un cercle C de centre O dans $\mathbb{R}^2$. Il existe une unique application

$$i_C : \mathbb{R}^2\backslash\{O\} \to \mathbb{R}^2\backslash\{O\} \tag{22.6}$$

telle que

(1) $i_C(x) = x$ *pour tout* $x \in C$ ITEMooXLZCooEGAxHu

(2) i_C *échange l'intérieur et l'extérieur de* C. ITEMooCPPUooDJIzSk

(3) i_C *laisse invariants les droites et les cercles orthogonales à* C.

*Cette application est l'**inversion** de cercle C.*

Démonstration. Soit P à l'intérieur de C, mais différent de O. Nous notons ℓ la droite (OP) et nous considérons un cercle C_2 passant par P et perpendiculaire à C (existence par la proposition 22.4). Nous avons $\ell \perp C$ (parce que ℓ est un rayon) et $C_2 \perp C$. Donc $i_C(\ell) = \ell$ et $i_C(C_2) = C_2$ par l'exigence (3). Mais comme $P \in \ell \cap C_2$ nous avons aussi

$$i_C(P) \in \ell \cap C_2. \tag{22.7}$$

Mais ℓ et C_2 se coupent en exactement deux points. Vu que $i_C(P)$ doit être hors de C (exigence (2)), avoir $i_C(P) = P$ est impossible. Nous en concluons que $i_C(P)$ doit être l'autre intersection.

Nous avons prouvé que les conditions (2) et (3) fixent l'image d'un point situé dans l'intérieur de C.

Si P est extérieur au cercle C, la même procédure fonctionne : nous considérons la droite $\ell = (OP)$ et un cercle C_2 perpendiculaire à C et passant par P. Encore une fois, ces deux objets sont fixés par i_C, et vu que $i_C(P)$ doit être à l'intérieur de C, il est fixé.

L'unicité est montrée.

En ce qui concerne l'existence, si $P \neq O$, la procédure suivante donne P' :

— Soit la droite $\ell = (OP)$.

— Soit un cercle C_2 perpendiculaire à C et passant par P.

— Le point P' est le point de l'intersection $C_2 \cap \ell$ qui n'est pas P.

Il est aisé de vérifier que poser $i_C(P) = P'$ donne une application qui vérifie toutes les propriétés. □

NORMooUBFQooCzXRJj

22.6.

Nous recopions la construction de l'inversion d'un point par rapport à un cercle. Si C_1 est un cercle de centre O et P est un point différent de O, alors la procédure suivante construit $P' = i_{C_1}(P)$:

— Soit la droite $\ell = (OP)$.

— Soit un cercle C_2 perpendiculaire à C_1 et passant par P.

— Le point P' est le point de l'intersection $C_2 \cap \ell$ qui n'est pas P.

Que se passe-t-il si ℓ et C_2 n'ont qu'une seule intersection ? Alors la droite $\ell = (OP)$ est tangente à C_2. Or de O il n'existent que deux droites tangentes à C_2, et ce sont les rayons passant par les intersections parce que les cercles sont perpendiculaires. En d'autres mots, cette situation se présente lorsque P est sur le cercle C_1. Dans ce cas, $i_{C_1}(P) = P$.

Remarque 22.7.

Lorsque nous disons qu'une inversion « conserve les droites passant par O », il y a pour sous-entendu que nous considérons la droite privée du point O, parce que de toutes façons, l'inversion n'est pas définie sur O.

Nous allons résoudre cet intéressant problème en 23.6.1, en ajoutant le point ∞ à toutes les droites.

Corolaire 22.8.

L'inversion i_C est une involution ($i_C^2 = \mathrm{Id}$).

Démonstration. Soit un cercle C de centre O et un point P. Si C_2 est un cercle perpendiculaire à C passant par P, alors nous avons vu en 22.6 que $P' = i_C(P)$ est l'autre intersection entre C_2 et la droite (OP).

Pour construire l'image de P', il faut un cercle perpendiculaire à C passant par P'. Le cercle C_2 déjà utilisé fait l'affaire. Ensuite, la droite (OP') est la même que la droite (OP). Donc l'image de P' est P. □

Soit C_1 le cercle dans $\mathbb{C}$ de centre 0 et de rayon 1. Nous notons $\alpha\colon \mathbb{C} \to \mathbb{C}$ la dilatation de rapport R.

PROPooYHQVooVFamhr

Proposition 22.9 ([1])**.**

Soit C le cercle de rayon R centré en 0 ($C = \alpha(C_1)$). Alors

$$i_{\alpha(C_1)} \circ \alpha = \alpha \circ i_{C_1}. \tag{22.8}$$

Démonstration. Soit $z = re^{i\theta}$; nous devons prouver que

$$i_C\big(\alpha(z)\big) = \alpha\big(i_{C_1}(z)\big). \tag{22.9}$$

Nous avons :

$$i_C\big(\alpha(z)\big) = i_C(Rre^{i\theta}) = R^2\frac{1}{Rr}e^{i\theta} = R\frac{1}{r}e^{i\theta} = \alpha\big(\frac{1}{r}e^{i\theta}\big) = \alpha\big(i_{C_1}(z)\big). \tag{22.10}$$

□

La définition 23.70 pour l'inversion sur $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ sera basée sur la proposition suivante.

PROPooEWXNooNshvHq

Proposition 22.10 ([1]).
Soit le cercle de rayon R et de centre $a \in \mathbb{C}$. Alors

$$i_C(z) = \frac{R^2}{\bar{z} - \bar{a}} + a. \tag{22.11}$$

EQooEQEEooHtJuje

De plus si t_a est la translation de vecteur a, nous avons la décomposition

$$i_{t_a(C_R)} = t_a \circ i_{C_R} \circ t_{-a} \tag{22.12}$$

où C_R est le cercle de rayon R centré en 0.

Démonstration. Si $z \in \mathbb{C}$ et z' est son image par i_C, alors non seulement

$$\|a - z'\|\|a - z\| = R^2, \tag{22.13}$$

mais en plus $a - z' = \lambda(a - z)$ pour un certain $\lambda > 0$. Cela est l'expression du fait que z' est sur la demi-droite qui joint a à z. Nous avons donc

$$\lambda\|a - z\|^2 = R^2 \tag{22.14}$$

et alors

$$\lambda = \frac{R^2}{(a - z)(\bar{a} - \bar{z})}. \tag{22.15}$$

En récrivant $a - z' = \alpha(a - z)$ avec cette valeur de λ nous trouvons

$$a - z' = \frac{R^2}{\overline{a - z}}, \tag{22.16}$$

ce qu'il fallait démontrer.

La décomposition demandée est une simple vérification en utilisant $i_{C_R}(z) = \frac{R^2}{\bar{z}}$ qui découle de la proposition 22.9. □

Avant d'aller plus loin, donnons l'équation d'un cercle dans $\mathbb{C}$. Si C est un cercle de centre ω et de rayon r, alors $z \in C$ si et seulement si $d(z, \omega) = r$. En développant, et en passant au carré sans perte d'information (les deux membres sont positifs), $z \in C$ si et seulement si

$$(z - \omega)(\bar{z} - \bar{\omega}) = r^2. \tag{22.17}$$

En développant nous pouvons aussi écrire l'équation du cercle sous la forme

$$z\bar{z} - z\omega - \bar{z}\omega = r^2 - \omega\bar{\omega}. \tag{22.18}$$

EQooWZVRooXciOsA

PROPooMIMRooTbQRVI

Proposition 22.11 ([545]).
Soit un cercle C de centre O. L'inversion

$$i_C : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}^2\backslash\{0\} \tag{22.19}$$

transforme

ITEMooNOXMooQYNPnv

(1) *les droites passant par O en elles-mêmes ;*
(2) *les cercles passant par O en des droites ne passant pas par O ;*
(3) *les droites ne passant pas par O en des cercles passant par O ;*
(4) *les cercles ne passant pas par O en des cercles ne passant pas par O.*

ITEMooRFPSooGdJdHD

Démonstration. Point par point.

(1) Une droite passant par O est une droite perpendiculaire à C. Par le point (3) de la définition 22.5, elle est invariante.

(2) Nous construisons successivement :
- Un cercle C_1 de centre O_1 et passant par O. Le but est de déterminer l'image de ce cercle.
- Le point P de C_1 tel que $[OP]$ en soit un diamètre.
- Le point $P' = i_C(P)$.
- La droite ℓ perpendiculaire à (OP) et passant par P'.

Nous montrons maintenant que $i_C(C_1\backslash\{O\}) = \ell$. Soit $Q \in C_1\backslash\{P, O\}$. Nous posons

$$Q' = (QO) \cap \ell. \tag{22.20}$$

Vu que $[OP]$ est un diamètre de C_1 et que $Q \in C_1$, le triangle OPQ est rectangle en Q. Et étant donné que Q' est sur ℓ nous savons que $OP'Q'$ est rectangle en P'.

De plus les angles en O de ces deux triangles sont identiques (parce que c'est l'angle formé par les droites (OQ) et (OP)) ; les triangles OPQ et $QP'Q'$ sont donc semblables et nous pouvons utiliser le théorème de Thalès [1] 12.162 :

$$\frac{OP}{OQ} = \frac{OQ'}{OP'}. \tag{22.21}$$

Donc

$$\|OP\|\|OP'\| = \|OQ\|\|OQ'\|, \tag{22.22}$$

mais P' est l'image de P par l'inversion du cercle C, c'est-à-dire $\|OP\|\|OP'\| = R^2$. Nous en déduisons que

$$\|OQ\|\|OQ'\| = R^2, \tag{22.23}$$

c'est-à-dire que Q' est l'image de Q par i_C, et donc que

$$i_C\big(C_1\backslash\{O\}\big) \subset \ell. \tag{22.24}$$

EQooPJBGooGeIVQQ

Pour avoir l'inclusion inverse, il faut remarquer que ℓ est parallèle à la tangente à C_1 en O. Donc si $Q \in \ell$, la droite (OQ) intersecte le cercle C_1 en un point Q'. En refaisant le cheminement du résultat (22.24) à l'envers, il est loisible de prouver que $i_C(Q') = Q$ et donc que ℓ est bien incluse à l'image de C_1 par i_C.

(3) Nous commençons par prouver que toutes les droites ne passant pas par O sont des images de cercles passant par O.

Nous considérons :
- une droite ℓ ne passant pas par O.
- la droite d, perpendiculaire à ℓ passant par O
- le point $P' = \ell \cap d$,
- le point $P = i_C(P')$,
- le cercle C_1 dont $[OP]$ est un diamètre.

Par tout ce que nous avons fait jusqu'à présent, la droite ℓ est l'image du cercle C_1. Or si $\|OP\| = r$ alors

$$\|OP'\| = \frac{R^2}{r}. \tag{22.25}$$

Donc quelle que soit la valeur de $\|OP'\|$ dans $]0, \infty[$, il existera un point P tel que le cercle passant par O et P ait pour image la droite perpendiculaire à (OP) passant par $i_C(P)$.

Étant donné que i_C est une involution surjective des cercles passant par O vers les droites ne passant pas par O, elle transforme également toutes les droites ne passant pas par O en un cercle passant par O.

1. Faites bien le dessin : ce n'est pas une situation de Thalès ultra-standard de collège.

(4) Pour cette partie, nous allons utiliser un peu de géométrie analytique dans $\mathbb{C}$[2].

(4a) **Le cas centré** Nous supposons que C est centré en 0 et de rayon, 1, ce telle sorte que $i_C(z) = 1/\bar{z}$ (proposition 22.10). Soit C_1 un cercle de centre ω et de rayon r, ne passant pas par 0, en particulier tel que $|\omega| \neq r$.
Si $z \in i_C(C_1)$ alors $i_C(z) \in C_1$ et nous avons

$$\big(i_C(z) - \omega\big)\overline{\big(i_C(z) - \omega\big)} = r^2. \tag{22.26}$$

En développant et en multipliant par $z\bar{z}$ nous trouvons

$$\bar{\omega}z + \omega\bar{z} + z\bar{z}(r^2 - \omega\bar{\omega}) = 1. \tag{22.27}$$

EQooCWXSooUzjRNr

ou, en posant $s = r^2 - \omega\bar{\omega}$,

$$z\bar{\omega} + \bar{z}\omega + sz\bar{z} = 1. \tag{22.28}$$

En comparant avec l'équation (22.18) d'un cercle, nous voyons qu'il y a de l'espoir de montrer que z est sur un cercle centré en $a = -\omega/s$. En effet, en divisant par s et en substituant a comme il faut, nous trouvons

$$-z\bar{a} - \bar{z}a + a\bar{z} = \frac{1}{s}. \tag{22.29}$$

En ce qui concerne le rayon R, il doit vérifier

$$\frac{1}{s} = R^2 - a\bar{a}, \tag{22.30}$$

ou encore

$$R = \frac{1}{s} + \frac{\omega\bar{\omega}}{s^2}. \tag{22.31}$$

Nous dévons vérifier que $R > 0$. En multipliant par s^2, il s'agit de vérifier que $s + \omega\bar{\omega} \geqslant 0$, ce qui est correct parce que $s + \omega\bar{\omega} = r^2$.
En résumé, si $z \in i(C_1)$ alors z est dans le cercle C_2 de centre $-\frac{\omega}{s}$ et de rayon R. Nous voysons que ce cercle ne passe pas par 0 par exemple directement sur l'équation (22.27) qui, pour $z = 0$ donne $0 = 1$.
Nous avons prouvé que $i(C_1) \subset C_2$. Pour prouver l'inclusion inverse, vu que i est une involution, il faut prouver $i(C_2) \subset C_1$. Pour cela nous écrivons l'équation qui donne $i(z) \in C_2$ et en développant nous devons conclure que $z \in C_1$. Nous ne le faisons pas ici.

(4b) **Le cas de rayon non unité**
Si C est un cercle quelconque, nous écrivons l'inversion du cercle C via la formule (22.11). Si $z \in i_C(C_1)$ alors $i(z) \in C_1$ et nous pouvons écrire

$$\big(i(z) - a\big)\overline{\big(i(z)} - \bar{a}\big) = R^2. \tag{22.32}$$

De là il faut déduire que z est sur un cercle ne passant pas par 0. Bons calculs...

□

2. Principalement parce que je ne comprends pas le raisonnement fait dans [545].

Chapitre 23

Espaces projectifs

Sur les espaces projectifs : [611].

Définition 23.1.
*Soit E un espace vectoriel de dimension finie sur le corps commutatif $\mathbb{K}$. Nous définissons sur $E\backslash\{0\}$ la relation d'équivalence $u \sim v$ si et seulement si $u = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Cette relation est la relation de **colinéarité**. L'ensemble des classes d'équivalence de $\sim$ est l'**espace projectif** de E et sera noté $P(E)$.*

DEFooTPPMooTDxNpg

Définition 23.2.
*Si $\dim E = 2$, l'ensemble $P(E)$ est la **droite projective**, et si $\dim E = 3$ nous parlons du **plan projectif**.*

Étant donné que tous les $\mathbb{K}$-espaces vectoriels de dimensions $n+1$ sont isomorphes à $\mathbb{K}^{n+1}$, nous noterons $P_n(\mathbb{K})$ ou P_n l'espace projectif $P(\mathbb{K}^{n+1})$. PgNotimesjNtMoW

Exemple 23.3.
Si $n = 1$ et $\mathbb{K} = \mathbb{R}$, l'espace projectif est l'ensemble des droites vectorielles dans le plan usuel. Il y en a une pour chaque point du type $(x, 1)$ avec $x \in \mathbb{R}$ et ensuite une horizontale, passant par le point $(1, 0)$. Nous avons donc

$$P_1(\mathbb{R}) = \{(1,0)\} \cup \{(x,1) \text{ tel que } x \in \mathbb{R}\}. \tag{23.1}$$

Le point $(1, 0)$ est dit « point à l'infini ». △

23.1 Sous espaces projectifs

Un **sous-espace projectif** de $P(E)$ est une partie de la forme $P(F)$ où F est un sous-espace vectoriel de E.

PropuqpWVx

Proposition 23.4.
Si F et G sont des sous-espaces vectoriels de E, alors

$$P(F) \cap P(G) = P(F \cap G) \tag{23.2}$$

et nous avons

$$\dim P(F) + \dim P(G) = \dim P(F+G) + \dim P(F \cap G). \tag{23.3}$$ EqNAdWfN

Démonstration. Nous avons

$$P(F) = \{[v] \text{ tel que } v \in F\} \tag{23.4}$$

où les crochets signifient la classe par rapport à la relation de colinéarité. Nous avons alors

$$P(F) \cap P(G) = \{[v] \text{ tel que } v \in F \cap G\} = P(F \cap G). \tag{23.5}$$

Cela prouve le premier point.

En ce qui concerne l'équation (23.3), en considérant $\dim P(E) = \dim E - 1$ nous devons prouver l'égalité

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G) \tag{23.6}$$

concernant les dimensions des espaces vectoriels usuelles. Si nous considérons une base de E telle que $B_1 = \{e_1, \dots, e_{k_1}\}$ est une base de $F \cap G$, $B_2 = \{e_{k_1+1}, \dots, e_{k_2}\}$ complète B_1 en une base de F et $B_3 = \{e_{k_2+1}, \dots, e_n\}$ complète $B_1 \cup B_2$ en une base de G.

Nous avons alors

$$\dim F + \dim G = 2\operatorname{Card}(B_1) + \operatorname{Card}(B_2) + \operatorname{Card}(b_3) \tag{23.7a}$$
$$\dim(F + G) = \operatorname{Card}(B_1) + \operatorname{Card}(b_2) + \operatorname{Card}(B_3) \tag{23.7b}$$
$$\dim(F \cap G) = \operatorname{Card}(B_1). \tag{23.7c}$$

De là la relation (23.3) se déduit immédiatement. □

Théorème 23.5 (incidence)**.**
Soient F et G deux sous-espaces vectoriels de E tels que

$$\dim P(F) + \dim P(G) \geqslant \dim P(E). \tag{23.8}$$

Alors $P(F) \cap P(G) \neq \varnothing$.

Démonstration. En utilisant les hypothèses et la proposition 23.4 nous avons

$$\dim P(E) + \dim P(G) = \dim P(F + G) + \dim P(F \cap G) \geqslant \dim P(E). \tag{23.9}$$

En passant aux espaces vectoriels correspondants,

$$\dim(F + G) + \dim(F \cap G) \geqslant \dim(E) + 1. \tag{23.10}$$

Mais nous avons aussi $\dim(F + G) \leqslant \dim(E)$ et par conséquent $\dim(F \cap G) \geqslant 1$. Au final, $\dim P(F \cap G) \geqslant 0$. Cela prouve que $P(F \cap G)$ contient au moins un élément (nous rappelons que lorsqu'un espace projectif contient un seul élément, sa dimension est zéro). □

Exemple 23.6.
Soient les plans $\Pi_1 \equiv x = 0$ et $\Pi_2 \equiv y = 0$. Nous avons

$$P(\Pi_1) = \{[0, y, 1]\} \cup \{[0, 1, 0]\} \tag{23.11a}$$
$$P(\Pi_2) = \{[x, 0, 1]\} \cup \{[1, 0, 0]\} \tag{23.11b}$$

où le crochet signifie la classe pour la colinéarité. Ces deux droites projectives ont comme point d'intersection le point $[0, 0, 1]$. △

Définition 23.7.
*Un **hyperplan projectif** est un sous-espace projectif de $P(E)$ de la forme $P(V)$ où V est un hyperplan de E.*

DEFooBBMBooSVgTnn

Définition 23.8.
*Soit E un espace vectoriel de dimension au moins 3. Nous disons que $d \subset P(E)$ est une **droite projective** de $P(E)$ si $d = P(D)$ pour un plan vectoriel $D \subset E$.*

*Nous disons que trois points de $P(E)$ sont **alignés** lorsqu'il existe une droite projective les contenant.*

23.9.
Dans la définition 23.8 nous voyons $P(D)$ comme inclus dans $P(E)$ dès que D est un sous-espace vectoriel de E. Cela est possible parce que si la direction de $v \in D$, c'est-à-dire la classe $[v]$ est également une direction dans E.

Le lemme suivant peut paraitre idiot, mais ce qui serait surement idiot est de l'utiliser sans s'en rendre compte.

Lemme 23.10.
Deux points dans $P(E)$ sont toujours alignés.

Démonstration. Soient deux points $A, B \in P(E)$. Si $A = \pi(a)$ et $B = \pi(b)$ alors le plan D passant par a, b et 0 est vectoriel et $P(D)$ contient A et B.

Note : si a, b et 0 sont trois points alignés, alors $A = B$. Il suffit de prendre les points a, c et 0 où $c \in E$ est un point quelconque non aligné avec 0 et a. Nous avons de toutes façons $A = B = \pi(a)$. □

Lemme 23.11.
Trois points distincts A, B, C dans $P(E)$ sont alignés si et seulement si il existe trois points non alignés $a, b, c \in E$ tels que

(1) le plan passant par a, b et c est vectoriel (c'est-à-dire passe par 0),

(2) $A = \pi(a)$, $B = \pi(b)$, et $C = \pi(c)$.

Démonstration. Deux implications à montrer.

(1) **Sens direct** Soient A, B, C distincts et alignés dans $P(E)$. Alors il existe un plan vectoriel D tel que $A, B, C \in P(D)$.

La condition $A \in P(D)$ implique qu'il existe $a \in D$ tel que $A = \pi(A)$. Idem pour B et C. Les points a, b et c ainsi construits sont distincts parce que A, B et C sont distincts. Si par malheur ces trois points étaient alignés, ce n'est pas grave : il suffit de remplacer a par λa avec $\lambda \neq 0$ pour qu'ils ne le soient plus (cette manipulation ne change pas le fait que le nouveau choix de point a reste dans D parce que D est vectoriel). Nous avons donc trois points non alignés a, b et c tous contenus dans D. Le plan D répond à la question.

(2) **Sens réciproque**

Soient a, b et c non alignés dans E tels que $A = \pi(a)$, $B = \pi(b)$ et $C = \pi(c)$. Le plan D les contenant tous trois est vectoriel par hypothèse. Nous avons $A, B, C \in P(D)$ et donc A, B et C sont alignés dans $P(E)$.

□

Proposition 23.12.
Soit $H = P(V)$ un hyperplan projectif de $P(E)$ et soit m hors de H. Alors toute droite projective passant par m coupe H en un et un seul point.

Démonstration. Si $\dim E = n$ nous avons $\dim V = n - 1$. Soit $d = P(D)$ une droite projective passant par m, c'est-à-dire que D est de dimension 2 dans E. Si $D \subset V$ alors $m \in P(D) \subset P(V)$; or nous avons demandé que m soit hors de $P(V)$. Par conséquent D n'est pas inclus dans V et en particulier $\dim(D + V) = \dim(E)$.

Nous recopions la formule (23.3) pour notre cas :

$$\underbrace{\dim d}_{=1} + \underbrace{\dim H}_{=n-2} = \underbrace{\dim P(D + V)}_{=n-1} + \dim P(D \cap V). \tag{23.12}$$

Nous avons donc $\dim P(D \cap V) = 0$, ce qui signifie que l'ensemble $P(D \cap V) = P(D) \cap P(V) = d \cap H$ contient un et un seul point. □

23.2 Espace projectifs comme « complétés » d'espaces affines

Soit E un espace vectoriel de dimension 2 et $P(E)$ la droite projective correspondante, et soit $\{e_1, e_2\}$ une base de E. Nous considérons la droite affine $d \equiv y = 1$. Nous avons la bijection

$$\begin{aligned}\phi\colon d \cup \{\infty\} &\to P(E)\\ (x,1) &\mapsto \text{la droite vectorielle passant par } (x,1)\\ \infty &\mapsto \text{la droite vectorielle passant par } (1,0).\end{aligned} \tag{23.13}$$

Lemme 23.13.
Si nous munissons l'ensemble $d \cup \{\infty\}$ de la topologie compactifiée d'Alexandrov[1]*, la bijection* (23.13) *est un homéomorphisme.*

Soient maintenant les plans affines dans l'espace vectoriel E de dimension 3

$$\Pi_1 \equiv z = 0 \tag{23.14a}$$
$$\Pi_2 \equiv z = 1. \tag{23.14b}$$

Une droite (vectorielle) de E coupe Π_2 en un et un seul point, sauf si elle est contenue dans Π_1. Nous avons donc une bijection

$$\begin{aligned}\phi\colon P(E) &\to \Pi_2 \cup P(\Pi_1)\\ d &\mapsto \begin{cases}\Pi_2 \cap d & \text{si cette intersection est non vide}\\ d & \text{sinon.}\end{cases}\end{aligned} \tag{23.15}$$

La droite projective $P(\Pi_1)$ est la droite à l'infini du plan projectif $P(E)$. Nous voyons que le plan projectif $P(E)$ peut être vu comme un plan affine (Π_2) « complété » par une droite affine $P(\Pi_1)$. Cette dernière droite est elle-même une droite affine complétée par un point à l'infini.

Nous pouvons généraliser cette démarche en considérant un espace affine $\mathcal{E}$ de direction E sur le corps $\mathbb{K}$. Nous construisons $F = E \times \mathbb{K}$ et nous considérons un repère affine sur F tel que $E \equiv x_{n+1} = 0$. Nous pouvons donc identifier $\mathcal{E}$ à l'hyperplan affine d'équation $x_{n+1} = 1$ dans F.

Une droite vectorielle de F non contenue dans E coupe $\mathcal{E}$ en un unique point ; nous avons donc une bijection

$$\mathcal{E} \cup P(E) \to P(F). \tag{23.16}$$

Dans ce cadre, $P(E)$ est l'hyperplan à l'infini et nous disons que $P(E)$ est la **complétion projective** de $\mathcal{E}$.

Exemple 23.14.
Nous considérons les plans affines

$$\Pi_1 \equiv z = 0 \tag{23.17a}$$
$$\Pi_2 \equiv z = 1 \tag{23.17b}$$

et nous avons la bijection

$$P(E) = \Pi_2 \cup P(\Pi_1). \tag{23.18}$$

Un plan vectoriel D a deux possibilités : soit il coupe Π_2 en une droite, soit il est égal à Π_1. Si $D \cap \Pi_2 = d$ (d est une droite affine), alors nous avons

$$P(D) = d \cup \{\infty_D\}, \tag{23.19}$$

ce qui justifie la terminologie comme quoi $P(D)$ est une droite dans $P(E)$. △

Soit E un espace vectoriel de dimension 3 et le plan projectif $P(E)$. Nous avons deux types de droites projectives :

1. Définition 7.106.

(1) D'abord nous avons la droite à l'infini, donnée[2] par $P(z = 0)$.

(2) Ensuite nous avons toutes les droites affines du plan $z = 1$. Chacune de ces droites est complétée par un point à l'infini.

ExempMyTmFp

Exemple 23.15.
Étudions un peu le second type de droites. D'abord si deux droites sont parallèles, leurs points à l'infini sont identiques. Prenons par exemple les droites $d = \{z = 1, x = 1\}$ et $d' = \{z = 1, x = 2\}$. Elles décrivent les directions des vecteurs

$$\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}. \tag{23.20}$$

En normalisant, ce sont les vecteurs

$$\frac{1}{\sqrt{2+y^2}} \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \frac{1}{\sqrt{5+y^2}} \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}, \tag{23.21}$$

et toutes deux tendent vers le vecteur $(0, 1, 0)$ pour $y \to \infty$. △

Lemme 23.16.
Deux droites d'un plan projectif ont toujours une intersection.

Démonstration. Si les deux droites sont des droites affines non parallèles, le résultat est évident. Si elles sont parallèles, alors l'intersection est donnée par le point à l'infini comme indiqué dans l'exemple 23.15.

Supposons que d est la droite à l'infini tandis que d' est une droite affine. Dans notre représentation usuelle du plan affine, la droite à l'infini d a contient les vecteurs $(1, y, 0)$ et le point à l'infini $(0, 1, 0)$. La droite affine d' a pour équation paramétriques

$$\begin{cases} x = at + c & \text{(23.22a)} \\ y = bt + d & \text{(23.22b)} \\ z = 1. & \text{(23.22c)} \end{cases}$$

Les directions données par la droite d' sont donc

$$\frac{1}{a^2t^2 + b^2t^2 + c^2 + d^2} \begin{pmatrix} at + c \\ bt + d \\ 1 \end{pmatrix} \tag{23.23}$$

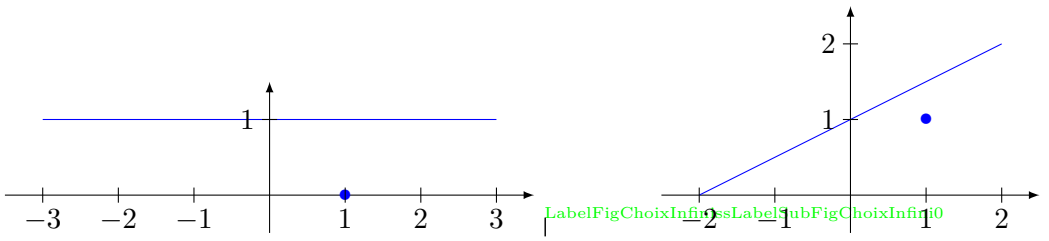
Son point à l'infini est la direction du vecteur $(a, b, 0)$, qui est bien un point de la droite à l'infini (éventuellement son point à l'infini[3]). □

La plupart du temps nous considérons le plan projectif comme étant le plan affine $z = 1$ de l'espace affine de dimension 3 complété par la droite affine $x = 1, z = 0$, elle-même complétée par le point $(0, 1, 0)$. Ce n'est évidemment pas la seule manière. Tout plan peut être considéré comme le plan à l'infini et pour une droite projective, tout point peut être considéré comme point à l'infini.

Sur la figure 23.1(a), le point à l'infini est la direction $(1, 0)$ tandis que la direction $(1, 1)$ n'a rien de spécial. À l'inverse sur la figure 23.1(b), la direction à l'infini est $(1, 1)$ tandis que la direction $(1, 0)$ est une direction usuelle.

2. Dans notre représentation usuelle du plan projectif $z = 1$.

3. D'accord, aller chercher le point à l'infini de la droite à l'infini, c'est chercher loin, mais n'empêche que ça existe.



(a) Ici le point à l'infini est la direction $(1,0)$.

(b) Ici le point à l'infini est la direction $(1,1)$.

FIGURE 23.1: Deux façons de voir la droite projective. Étant donné que les points de la droite projective doivent être interprétés comme des directions (des classes d'équivallence), en réalité les deux dessins représentent les mêmes ensembles. LabelFigChoixInfini

Remarque 23.17.
Du point de vue de la topologie, si nous mettons celle de la compactification d'Alexandrov, tous les points de la droite projective sont équivalents.

Du point de vue de la géométrie différentielle, c'est la même chose. En effet nous pouvons mettre sur la droite projective un système de deux cartes en pensant aux angles. La première sur $]-a,a[$ avec par exemple $a<\pi/4$. La seconde carte serait $]a/2,\pi[$. Dans ce cas la direction $\theta=0$ semble jouer un rôle spécial, mais il n'en est rien.

Nous pouvons également considérer les cartes $]\pi/4-a,\pi/4+a[$ et $]\pi/4+a/2,5\pi/4[$. Dans ces cartes, c'est plutôt le point $\theta=\pi/4$ qui semble différent (encore qu'il soit bien centré dans une carte).

23.3 Théorème de Pappus

Théorème 23.18.
Soient deux droites d et d' dans un plan affine. Soient $A,B,C\in d$ et $A',B',C'\in d'$ tels que $AB'\parallel BA'$ et $BC'\parallel B'C$. Alors $AC'\parallel A'C$.

Démonstration. Si d et d' ne sont pas parallèles nous considérons o, le point d'intersection. Les relations de parallélisme des hypothèses impliquent qu'il existe λ_1 et λ_2 tels que

$$\begin{cases} A=\lambda_1 B & (23.24a)\\ B'=\lambda_1 A' & (23.24b)\end{cases}$$

et

$$\begin{cases} B'=\lambda_2 C' & (23.25a)\\ C=\lambda_2 B. & (23.25b)\end{cases}$$

En substituant nous trouvons

$$\begin{cases} C=\dfrac{\lambda_2}{\lambda_1}A & (23.26a)\\ A'=\dfrac{\lambda_2}{\lambda_1}C', & (23.26b)\end{cases}$$

ce qui implique que $A'C\parallel AC'$.

Si les droites d et d' sont parallèles, alors nous avons les translations

$$\begin{cases} B=A+x & (23.27a)\\ A'=B'+x & (23.27b)\end{cases}$$

et

$$\begin{cases} B=C+y & (23.28a)\\ C'=B'+y, & (23.28b)\end{cases}$$

ce qui montre que

$$\begin{cases} C = A + x - y & (23.29a) \\ A' = C' + x - y, & (23.29b) \end{cases}$$

et donc que $A'C \parallel AC'$. □

Le théorème suivant est une version projective.

Théorème 23.19.
Soient d et d' deux droites projectives d'un plan projectif. Soient $A, B, C \in d$ et $A', B', C' \in d'$. Alors les points $B'C \cap C'B$, $C'A \cap A'C$ et $A'B \cap B'A$ sont alignés.

Démonstration. Soient $E = B'C \cap C'B$ et $E' = C'A \cap A'C$. Ces deux points existent parce que deux droites projectives distinctes ont toujours un unique point d'intersection. Nous allons prendre EE' comme droite à l'infini et prouver que le point $A'B \cap B'A$ est dessus. Étant donné que le point d'intersection de $B'C$ et $C'B$ est à l'infini nous avons $B'C \parallel C'B$ (cela est un exemple de la flexibilité de la notion de parallélisme en géométrie projective). De la même façon nous avons $C'A \parallel A'C$.

Par le théorème de Pappus affine nous avons alors $A'B \parallel B'A$ et par conséquent le point d'intersection est sur la droite à l'infini, c'est-à-dire sur la droite EE'. □

23.4 Homographies

23.4.1 Homographies

DEFooKWSMooXvOeEP

Définition 23.20.
Soient E et F deux espaces vectoriels avec leurs projections naturelles

$$\pi_E \colon E\backslash\{0\} \to P(E) \quad (23.30a)$$
$$\pi_F \colon F\backslash\{0\} \to P(F). \quad (23.30b)$$

*Une application $g \colon P(E) \to P(F)$ est une **homographie** si il existe un isomorphisme d'espaces vectoriels $\bar{g} \colon E \to F$ tel que le diagramme*

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{\bar{g}} & F\backslash\{0\} \\ \downarrow{\scriptstyle \pi_E} & & \downarrow{\scriptstyle \pi_F} \\ P(E) & \xrightarrow[g]{} & P(F) \end{array} \quad (23.31)$$

commute, c'est-à-dire si il existe $\bar{g} \colon E \to F$ telle que

$$\pi_F\big(\bar{g}(v)\big) = g\big(\pi_E(v)\big) \quad (23.32)$$

EQooSEFWooRpiLxt

pour tout $v \in E$.

Lemme 23.21.
Si $\bar{g} \colon E \to F$ est linéaire et si $\ker \bar{g} = \{0\}$ alors l'application g définie par

$$g\big(\pi_E(v)\big) = \pi_F\big(\bar{g}(v)\big) \quad (23.33)$$

EqRlGLlW

est une homographie.

Démonstration. Nous devons simplement vérifier que l'équation (23.33) définit bien une application. Soient $v, w \in E$ tels que $\pi_E v = \pi_E w$; nous devons montrer que

$$\pi_F \bar{g} v = \pi_F \bar{g} w. \quad (23.34)$$

EqmoIUkH

L'équation (23.34) sera vérifiée si et seulement si il existe $\lambda \in \mathbb{R}$ tel que $\bar{g}v = \lambda \bar{g}w$, c'est-à-dire si et seulement si $\bar{g}(v - \lambda w) = 0$. Étant donné que nous supposons que le noyau de $\bar{g}$ est réduit à $\{0\}$, l'équation (23.34) sera vérifiée si et seulement si $v = \lambda w$, ce qui signifie exactement $\pi_E(v) = \pi_E(w)$. □

La proposition suivante donne les premières propriétés des homographies.

PROPooGVYXooDIiIbW

Proposition 23.22.
Quelques propriétés des homographies.

ITEMooTIONooSKjfny

(1) *Une homographie est bijective.*
(2) *Si deux espaces projectifs sont homographes, alors ils ont même dimension.* ITEMooIZAPooNxEigb
(3) *L'ensemble des homographies $P(E) \to P(E)$ est un groupe (pour la composition).*
(4) *Une homographie conserve l'alignement des points.*

Démonstration. Nous considérons une homographie $g\colon P(E) \to P(F)$, et $\bar{g}$ l'isomorphisme d'espaces vectoriels correspondant.

(1) Pour l'injectivité, si $g\big([v]\big) = g\big([w]\big)$ alors en utilisant la définition d'une homographie, $\pi_F \bar{g} v = \pi_F \bar{g} w$, ce qui implique que $\bar{g}v = \lambda \bar{g} w$, et donc $v = \lambda w$, ce qui signifie $[v] = [w]$.
Pour la surjectivité, un élément général de $P(F)$ prend la forme $\pi_F \bar{g} v$ pour un certain $v \in E$. Nous avons $g\big(\pi_E v\big) = \pi_F \bar{g} v$. Par conséquent l'élément $\pi_F \bar{g} v$ est bien dans l'image de g.
(2) Une homographie $P(E) \to P(F)$ n'existe que si il existe un isomorphisme $E \to F$. Les dimensions sont donc automatiquement égales.
(3) Il suffit de vérifier que l'application

$$\begin{aligned} \varphi\colon P(E) &\to P(E) \\ \pi_F \bar{g} v &\mapsto \pi_E v \end{aligned} \tag{23.35}$$

est bien définie et donne l'inverse de g.
(4) Soient les points A, B, C alignés dans $P(E)$; ils correspondent à des directions de E qui sont données par des vecteurs situés sur la même droite affine. Autrement dit, il existe trois points $a, b, c \in E$ situés sur la même droite affine tels que $A, B, C = \pi_E(a, b, c)$. Les images par g sont données par $\pi_F \bar{g} a$, $\pi_F \bar{g} b$, et $\pi_F \bar{g} c$.
Étant donné qu'un isomorphisme d'espaces vectoriels conserve l'alignement affin, les points $\bar{g}a$, $\bar{g}b$ et $\bar{g}c$ sont alignés dans F. Cela implique que les projections par π_F sont alignés dans $P(F)$.

□

23.4.2 Le groupe projectif

DEFooWUSDooSLVKwV

Définition 23.23.
Le groupe des homographies de l'espace $P(E)$ est le ***groupe projectif****, noté* $\mathrm{PGL}(E)$.

Nous avons une surjection naturelle

$$\begin{aligned} \mathrm{GL}(E) &\to \mathrm{PGL}(E) \\ \bar{g} &\mapsto g \end{aligned} \tag{23.36}$$

EqnoNEfe

qui s'avère être un morphisme de groupes.

Proposition 23.24.
Nous avons l'isomorphisme de groupes

$$\frac{\mathrm{GL}(E)}{\{\textit{homothéties}\}} \simeq \mathrm{PGL}(E). \tag{23.37}$$

Démonstration. Nous devons prouver que le noyau de l'application (23.36) est constitué des homothéties. Considérons un automorphisme d'espace vectoriel $f: E \to E$ dont l'homographie associée est l'identité, et prouvons que f est une homothétie. Nous avons le diagramme commutatif suivant :

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{f} & E\backslash\{0\} \\ {\scriptstyle\pi_E}\downarrow & & \downarrow{\scriptstyle\pi_E} \\ P(E) & \xrightarrow[\mathrm{Id}]{} & P(E). \end{array} \tag{23.38}$$

Pour tout vecteur $v \in E$ nous avons $\pi_E(v) = \pi_E\big(f(v)\big)$. Cela implique qu'il existe $\lambda \in \mathbb{R}$ tel que $f(v) = \lambda v$. Tous les vecteurs de E sont donc des vecteurs propres de f. Cela n'est possible que si toutes les valeurs propres sont identiques, c'est-à-dire que f est une homothétie. □

23.4.3 Repères projectifs

Définition 23.25 ([612])**.**
*Des éléments $\{P_i\}_{i\in I}$ sont **projectivement independents** si en choisissant $v_i \in \pi^{-1}(P_i)$ nous obtenons des vecteurs $\{v_i\}_{i\in I}$ linéairement indépendants.*

DEFooPZKFooDBXtEn

Définition 23.26.
*Soit E un espace vectoriel de dimension $n+1$. Un **repère projectif** de $P(E)$ est la donnée de $n+2$ points $m_0, \dots, m_{n+1}$ tels que*

(1) les vecteurs m_i, $i \neq 0$, sont les images d'une base $\{e_i\}$ de E

(2) $m_0 = \pi_E(e_1 + e_2 + \dots + e_{n+1})$.

Note que si $m_k = \pi_E(v_k)$ $(k = 0, \dots, n+1)$, alors tout choix de $n+1$ vecteurs parmi les v_k est une base de E.

Exemple 23.27.
Un repère projectif de l'espace $P(\mathbb{R}^3)$ est par exemple les éléments $\{m_i\}_{i=1,\dots,3}$ donnés par

$$m_1 = \pi(e_1) \tag{23.39a}$$
$$m_2 = \pi(e_2) \tag{23.39b}$$
$$m_3 = \pi(e_3) \tag{23.39c}$$
$$m_0 = \pi(e_1 + e_2 + e_3). \tag{23.39d}$$

△

Exemple 23.28.
Pour $P(\mathbb{C}^2)$, un repère projectif possible est $m_1 = [1, 0]$, $m_2 = [0, 1]$, $m_0 = [1, 1]$. △

23.29.
Pourquoi voulons nous des repères projectifs ? Pourquoi demander un quatrième élément alors que trois devraient suffire ? Le fait est que si E est de dimension 3, nous voudrions pouvoir identifier E et $P(\mathbb{R}^3)$.

Plus précisément, si E est de dimension $n+1$ et possède une base $\{f_i\}_{i=1,\dots,n+1}$, il existe un unique isomorphisme d'espaces vectoriels $E \to \mathbb{R}^{n+1}$ qui envoie cette base sur la base canonique de $\mathbb{R}^{n+1}$. La base de E étant fixée, nous pouvons donner à un point de E les coordonnées de son image dans $\mathbb{R}^{n+1}$ par cet isomorphisme *qui est unique*.

Dans le cas des espaces projectifs, nous voudrions avoir une unique homographie $\phi: P(E) \to P(\mathbb{R}^{n+1})$ qui permet de donner à un point $A \in P(E)$ les coordonnées de $\pi^{-1}\big(\phi(A)\big)$. Bien entendu ce dernier n'est pas un élément bien défini de $\mathbb{R}^{n+1}$ parce qu'il y a toute une droite d'éléments de $\mathbb{R}^n$ qui se projettent sur $\phi(A)$.

L'idée d'imposer un point de plus est la bonne. Si nous imposons un point de plus, nous pouvons dire que les coordonnées de $A \in P(E)$ sont celles dans $\mathbb{R}^{n+1}$ de l'élément de $\pi^{-1}\big(\phi(A)\big)$ dont la dernière coordonnée est par exemple 1.

Nous allons maintenant mettre ça en musique.

D'abord nous donnons un exemple de non unicité.

Exemple 23.30.
Soit un espace vectoriel E de dimension 2 et une base $\{b_1, b_2\}$ de E. Nous considérons également l'espace $\mathbb{R}^2$ muni de sa base canonique $\{e_1, e_2\}$.

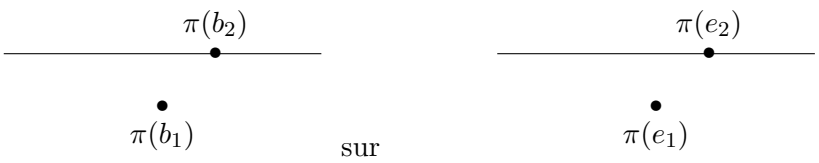
Soit une homographie $\phi\colon P(E) \to P(\mathbb{R}^2)$ telle que

$$\phi\big(\pi(b_i)\big) = \pi(e_i) \tag{23.40}$$

pour $i = 1, 2$. Nous allons facilement construire une autre homographie qui vérifie les mêmes conditions.

L'idée est la suivante. L'espace $P(E)$ peut être vu comme la droite complétée $\{(x,1)\}_{x\in\mathbb{R}} \cup \{(1,0)\}$ et l'espace $P(\mathbb{R}^2)$ également. Une homographie respectant (23.40) doit envoyer le $(1,0)$ de E vers le $(1,0)$ de $\mathbb{R}^2$ et le $(1,1)$ de E vers le $(1,1)$ de $\mathbb{R}^2$. Mais en ce qui concerne le reste de la droite, l'homographie peut la parcourir à la vitesse qu'elle veut.

Il faut envoyer



Soit donc une homographie $\phi\colon P(E) \to P(\mathbb{R}^2)$, et nous définissons

$$\begin{aligned} \phi'\colon P(E) &\to P(\mathbb{R}^2) \\ \pi(xb_1 + yb2) &\mapsto \phi\big(\pi(xb_1 + \lambda y b_2)\big) \end{aligned} \tag{23.41}$$

pour un certain $\lambda \neq 1$. En ce qui concerne le relèvement, l'application $\bar{\phi}'\colon E \to \mathbb{R}^2$ donnée par

$$\bar{\phi}'(xb_1 + yb_2) = \bar{\phi}(xb_1 + \lambda y b_2) \tag{23.42}$$

est bien définie et vérifie

$$\pi_{\mathbb{R}^2} \circ \bar{\phi}' = \phi' \circ \pi_E. \tag{23.43}$$

Donc ϕ' est une homographie. De plus

$$\phi'\big(\pi(b_1)\big) = \phi\big(\pi(b_1)\big) \tag{23.44a}$$
$$\phi'\big(\pi(b_2)\big) = \phi\big(\pi(\lambda b_2)\big) = \phi\big(\pi(b_2)\big) \tag{23.44b}$$

parce que $\pi(\lambda b_2) = \pi(b_2)$.

Nous n'avons donc pas l'unicité. △

C'est pour rétablir cette unicité que nous demandons d'avoir un point de plus pour avoir un repère projectif. De cette façon nous aurons une unique homographie $\phi\colon P(E) \to P(\mathbb{R}^{n+1})$ vérifiant $\phi\big(\pi_E(b_i)\big) = \pi_{\mathbb{R}^{n+1}}(e_i)$ pour tout $i = 0, \dots, n+1$.

Lemme 23.31.
Soit un espace vectoriel E de dimension $n+1$ muni de deux bases $\{e_i\}_{i=1,\dots,n+1}$ et $\{f_i\}_{i=1,\dots,n+1}$. Soit un repère projectif $\{m_0, m_i\}_{i=1,\dots,n+1}$ de $P(E)$.

Si $\pi(e_i) = \pi(f_i) = m_i$ pour tout $i = 1, \dots, n+1$ et si

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}) \tag{23.45}$$

alors les deux bases sont proportionnelles : il existe λ tel que $f_i = \lambda e_i$ pour $i = 1, \dots, n+1$.

Démonstration. Nous avons $\pi(e_i) = \pi(f_i)$ pour tout $i = 1, \dots, n+1$. Donc pour chaque $i = 1, \dots, n+1$ il existe $\lambda_i \in \mathbb{K}$ tel que $e_i = \lambda f_i$. Nous devons voir que les λ_i sont en réalité tous égaux.

Pour cela nous avons aussi l'égalité pour $i = 0$:

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}), \tag{23.46}$$

ce qui donne un $\mu \in \mathbb{K}$ tel que $e_1 + \dots + e_{n+1} = \mu(f_1 + \dots + f_{n+1})$, c'est-à-dire

$$\lambda_1 f_1 + \dots + \lambda_{n+1} f_{n+1} = \mu f_1 + \dots + \mu f_{n+1}. \tag{23.47}$$

Du fait que les f_i forment une base, cette égalité impose à tous les λ_i d'être égal à μ. □

THOooTXPVooJGigne

Théorème 23.32 ([613]).
Soient $P(E)$ et $P(F)$ deux espaces projectifs de dimensions n.

ITEMooRSIWooXbEnlT

(1) Une homographie $P(E) \to P(F)$ envoie un repère projectif sur un repère projectif.

ITEMooQXQXooDyIsxsh

(2) Si $(m_0, \dots, m_{n+1})$ est un repère projectif de $P(E)$, si $(m'_0, \dots, m'_{n+1})$ est un repère projectif de $P(F)$ alors il existe une unique homographie $g\colon P(E) \to P(F)$ telle que $g(m_i) = m'_i$ pour tout $i = 0, 1, \dots, n+1$

Démonstration. Un point à la fois.

(1) **(1)**
Soit une homographie $\phi\colon P(E) \to P(F)$ et un repère projectif $\{m_0, m_1, \dots, m_{n+1}\}$ de $P(E)$. Nous posons $m'_i = \phi(m_i)$ pour tout $i = 0, \dots, n+1$. Nous devons prouver que ces m'_i forment un repère projectif de $P(F)$.
D'abord pour $i = 1, \dots, n+1$ nous avons $m'_i = \phi\big(\pi_E(e_i)\big) = \pi_F\big(\bar{\phi}(e_i)\big)$, mais $\{\bar{\phi}(e_i)\}_{i=1,\dots,n+1}$ est une base de F parce que $\bar{\phi}$ est un isomorphisme d'espaces vectoriels. Donc oui : les m'_i $(i = 1, \dots, n+1)$ sont les projetés d'une base de F.
Nous posons au passage $f_i = \bar{\phi}(e_i)$. En ce qui concerne m_0 nous savons que $m_0 = \pi_E(e_1 + \dots + e_{n+1})$ et

$$m'_0 = \phi\big(\pi_E(e_1 + \dots + e_{n+1})\big) = \pi_F\big(\bar{\phi}(e_1 + \dots + e_{n+1})\big) = \pi_F(f_1 + \dots + f_{n+1}), \tag{23.48}$$

ce qui termine de montrer que $\{m'_i\}_{i=0,\dots,n+1}$ est un repère projectif de $P(F)$.

(2) **(2)**
Soient un repère projectif $(m_0, \dots, m_{n+1})$ de $P(E)$ et un repère projectif $(m'_0, \dots, m'_{n+1})$ de $P(F)$. Nous choisissons des bases $\{e_i\}$ de E et $\{f_i\}$ de F telles que

$$m_i = \pi_E(e_i) \tag{23.49a}$$
$$m'_i = \pi_F(f_i) \tag{23.49b}$$

pour $i = 1, \dots, n+1$ et

$$m_0 = \pi_E(e_1 + \dots + e_{n+1}) \tag{23.50a}$$
$$m'_0 = \pi_F(f_1 + \dots + f_{n+1}). \tag{23.50b}$$

Nous considérons un isomorphisme d'espace vectoriel $\bar{\phi}\colon E \to F$ tel que $\bar{\phi}(e_i) = f_i$ pour tout i, et nous voulons définir $\phi\colon P(E) \to P(F)$ par

$$\phi\big(\pi_E(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \tag{23.51}$$

EQooRMYIooKcPZwD

Cela est bien défini parce que si $\pi_E(v) = \pi_E(w)$ alors $w = \lambda v$ et

$$\pi_F\big(\bar{\phi}(\lambda v)\big) = \pi_F\big(\lambda\bar{\phi}(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \tag{23.52}$$

L'application définie par (23.51) est une homographie qui envoie m_i sur m'_i pour tout $i = 0, \dots, n+1$. Ceci prouve la partie « existence » du point (2).

Pour l'unicité, soient des homographies

$$\phi_1 \colon P(E) \to P(F) \tag{23.53a}$$
$$\phi_2 \colon P(E) \to P(F) \tag{23.53b}$$

telles que $\phi_1(m_i) = \phi_2(m_i)$ pour tout $i = 0, \dots, n+1$. Soit aussi une base $\{e_i\}_{i=1,\dots,n+1}$ de E adaptée au repère projectif, c'est-à-dire $m_i = \pi_E(e_i)$ pour $i = 1, \dots, n+1$ et $\pi_E(e_1 + \dots + e_{n+1}) = m_0$. Nous considérons aussi les isomorphismes d'espaces vectoriels $\bar{\phi}_1$ et $\bar{\phi}_2$. Avec tout ce beau monde nous avons

$$\phi_1(m_i) = \pi_E\big(\bar{\phi}_1(e_i)\big) \tag{23.54a}$$
$$\phi_2(m_i) = \pi_E\big(\bar{\phi}_2(e_i)\big). \tag{23.54b}$$

Mais nous savons que $\phi_1(m_i) = \phi_2(m_i)$, donc nous savons que $\pi_E\big(\bar{\phi}_1(e_i)\big) = \pi_E\big(\bar{\phi}_2(e_i)\big)$, ce qui nous fait conclure que

$$\bar{\phi}_1(e_i) = \lambda_i \bar{\phi}_2(e_i) \tag{23.55}$$

pour certaines constantes $\lambda_i \in \mathbb{K}$. Le même raisonnement appliqué à m_0 nous donne un $\mu \in \mathbb{K}$ tel que

$$\bar{\phi}_1(e_1) + \dots + \bar{\phi}_1(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \dots + \bar{\phi}_2(e_{n+1})\big). \tag{23.56}$$

En mettant l'un dans l'autre :

$$\lambda_1 \bar{\phi}_2(e_1) + \dots + \lambda_{n+1} \bar{\phi}_2(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \dots + \bar{\phi}_2(e_{n+1})\big). \tag{23.57}$$

Sachant que $\{\bar{\phi}_2(e_i)\}_{i=1,\dots,n+1}$ est une base de F et nous souvenant de l'unicité de la décomposition d'un élément dans une base[4], nous en déduisons que tous les λ_i doivent être égaux à μ. Donc pour tout $v \in E$ nous avons $\bar{\phi}_1(v) = \lambda \bar{\phi}_2(v)$.
Cela a pour conséquence que $\phi_1 = \phi_2$.

□

23.33.
Si nous avons une droite projective, trois points sont nécessaires pour créer un repère et donc pour construire une homographie de la droite sur elle-même. Soit E un espace vectoriel de dimension 2 et $P(E)$ la droite projective qui lui est associée. Soit une homographie $f \colon P(E) \to P(E)$ et $\bar{f} \colon E \to E$, l'isomorphisme d'espaces vectoriels associé (par $f \circ \pi_E = \pi_E \circ \bar{f}$). Si $\{e_1, e_2\}$ est une base de E alors l'application $\bar{f}$ a une matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{M}(2, \mathbb{K}) \tag{23.58}$$

avec $\det A \neq 0$ parce que $\bar{f}$ est un isomorphisme.

La plupart des points de $P(E)$ sont représentés par des points de la forme $(z, 1)$. Nous voudrions savoir quelle est la direction représentée par le point $\bar{f}(z, 1)$; c'est-à-dire que nous voudrions savoir $f([z, 1])$ sous la forme $[z', 1]$ (si possible). Nous avons

$$\bar{f}(z, 1) = (a_{11}z + a_{12}, a_{21}z + a_{22}). \tag{23.59}$$

Nous posons $\lambda = a_{21}z + a_{22}$ et nous avons

$$\bar{f}(z, 1) = \lambda \left(\frac{a_{11}z + a_{12}}{\lambda}, 1 \right). \tag{23.60}$$

Il y a plusieurs possibilités suivant les valeurs de λ et de z.

(1) Si $\lambda = 0$ c'est que nous avons $\bar{f}(z, 1) = (a_{11}z + a_{12}, 0)$. L'application f envoie donc le point $(z : 1)$ sur le point à l'infini.

4. Proposition 4.7.

(2) Si $\lambda \neq 0$, alors f envoie le point $(z:1)$ vers un autre point « normal ».
(3) Si le point de départ est le point à l'infini alors $\bar{f}(1,0) = (a_{11}, a_{21})$. Cela peut être le point à l'infini ou non selon les valeurs des a_{ij}.

Dans tous les cas si nous posons

$$\begin{cases} \varphi_f(z) = \dfrac{a_{11}z + a_{12}}{a_{21}z + a_{22}} & \text{(23.61a)} \\ \varphi_f(\infty) = \dfrac{a_{11}}{a_{21}} & \text{(23.61b)} \end{cases}$$

alors nous avons

$$\bar{f}(z,1) = \big(\varphi_f(z), 1\big). \tag{23.62}$$

Si nous prenons la convention que $\frac{1}{0} = \infty$ et que $(\infty, 0)$ est le point à l'infini, alors cette application φ_f donne bien toutes les valeurs de f, y compris les cas à l'infini.

LEMooXNKOooBKhzyt

Lemme 23.34 ([614]).
Trois points distincts d'une droite projective forment un repère projectif.

Démonstration. Soit une droite projective $d = P(E)$ où E est un espace vectoriel de dimension 2 sur le corps $\mathbb{K}$. Soient trois points distincts A, B et C de d. Nous avons $a, b, c \in E$ tels que $A = \pi(a)$, $B = \pi(b)$ et $C = \pi(c)$. Vu que $A \neq B$, les vecteurs a et b ne sont pas proportionnels et la partie $\{a, b\}$ est libre dans E. Autrement dit, c'est une base [5].

Il existe donc $\alpha, \beta \in \mathbb{K}$ tels que $c = \alpha a + \beta b$. De plus α et β ne sont pas nuls parce que $C \neq A$ et $C \neq B$. En prenant $a' = \alpha a$, $b' = \beta b$ et $c' = c$ nous avons : $A = \pi(a')$, $B = \pi(b')$, $C = \pi(c')$ en même temps que $\{a', b'\}$ est une base de E et $c' = a' + b'$. Donc A, B, C est un repère projectif de $d = P(E)$. □

CORooRFCZooGZiQBJ

Corolaire 23.35 ([615]).
Soient E et F des espaces vectoriels de dimension 2. Soient A_i $(i = 1, 2, 3)$ des points distincts sur $P(E)$ et B_i distincts sur $P(F)$. Alors il existe une unique homographie $P(E) \to P(F)$ portant A_i sur B_i pour tout $i = 1, 2, 3$.

Démonstration. Il s'agit de mettre en conjonction le lemme 23.34 qui dit que les A_i forment un repère projectif de $P(E)$ (idem : les B_i forment un repère projectif de $P(F)$) et le théorème 23.32(2) qui dit que l'une va sur l'autre part une unique homographie. □

23.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$

Pour rappel, une droite projective est l'espace projectif modelé sur un espace vectoriel de dimension deux (définition 23.8).

NORMooUQRUooOMIzJD

23.36.
Nous allons faire un usage assez intense de bijections entre $P(\mathbb{K}^2)$ et $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$. Une possible est

$$\begin{aligned} \varphi_0 \colon P(\mathbb{K}^2) &\to \mathbb{K} \cup \{\infty\} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_2 \neq 0 \\ \infty & \text{si } k_2 = 0. \end{cases} \end{aligned} \tag{23.63}$$

EQooSIJDooTHPYMb

Notons que nous utilisons ici le fait que $\mathbb{K}$ soit commutatif, sinon il aurait fallu choisir $k_1 k_2^{-1}$ ou $k_2^{-1} k_1$ au lieu d'écrire gentiment k_1/k_2.

CORooFJSCooNOeAel

Corolaire 23.37.
Soit une bijection $\varphi \colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. Les points $\{\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)\}$ forment un repère projectif de $P(\mathbb{K}^2)$.

5. Il convient de citer ici la proposition 4.19.

Démonstration. Il s'agit seulement d'une application du lemme 23.34.

Juste pour l'amusement, nous allons le prouver explicitement pour la bijection $\varphi = \varphi_0$ donnée en (23.63). Un repère projectif est la définition 23.26. Nous avons

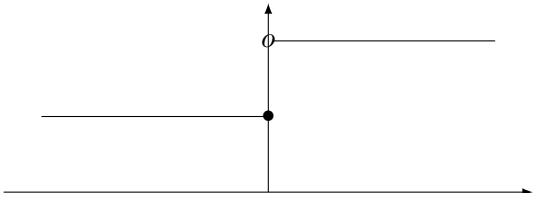
$$\varphi_0^{-1}(\infty) = [1,0] = \pi_{\mathbb{K}^2}\big((1,0)\big) \tag{23.64a}$$
$$\varphi_0^{-1}(0) = [0,1] = \pi_{\mathbb{K}^2}\big((0,1)\big) \tag{23.64b}$$
$$\varphi_0^{-1}(1) = [1,1] = \pi_{\mathbb{K}^2}\big((1,1)\big) \tag{23.64c}$$

Les points $(1,0)$ et $(0,1)$ forment un base de $\mathbb{K}^2$ et nous avons bien $(1,1) = (1,0) + (0,1)$. Donc le tout vérifie bien la définition d'un repère projectif. □

D'autre part, comme il est plus agréable de travailler avec $\hat{\mathbb{K}}$ qu'avec $P(\mathbb{K}^2)$ nous avons envie de voir $\hat{\mathbb{K}}$ comme un espace projectif (qu'il n'est pas). Il y a cependant nombre d'autres identifications possibles. En voici un autre :

$$\begin{aligned}\varphi_1\colon P(\mathbb{K}^2) &\to \hat{\mathbb{K}}\\ [k_1,k_2] &\mapsto \begin{cases} \frac{k_2}{k_1} & \text{si } k_1 \neq 0\\ \infty & \text{si } k_1 = 0. \end{cases}\end{aligned} \tag{23.65}$$

Vous en voulez une plus compliquée ? En voici une pour $\mathbb{K} = \mathbb{R}$, basée sur le dessin suivant :



La bijection associée est :

$$\begin{aligned}\varphi_d\colon P(\mathbb{R}^2) &\to \hat{\mathbb{R}}\\ [k_1,k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_1k_2 \leqslant 0\\ 2\frac{k_1}{k_2} & \text{si } k_1k_2 > 0\\ \infty & \text{si } k_2 = 0. \end{cases}\end{aligned} \tag{23.66}$$

EQooMIGCooAtXPaS

Pour mettre un peu d'ordre dans toutes ces identifications possibles, nous introduisons une classe.

DEFooMLQUooGwvQMh

Définition 23.38.
Pour une bijection $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ nous définissons

$$A(\varphi) = \big\{\varphi_a\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}} \text{ tel que } \varphi_a^{-1} \circ \varphi \text{ soit une homographie } \big\} \tag{23.67}$$

Les classes sont assez larges parce que pour toute homographie $\phi\colon P(\mathbb{K}^2) \to P(\mathbb{K}^2)$, nous avons $\varphi \circ \phi^{-1} \in A(\varphi)$. Mieux, nous avons le lemme suivant.

Lemme 23.39.
L'application

$$\begin{aligned}\psi\colon A(\varphi) &\to \mathrm{PGL}(\mathbb{K}^2)\\ \varphi_a &\mapsto \varphi_a^{-1} \circ \varphi\end{aligned} \tag{23.68}$$

est une bijection.

Démonstration. Pour rappel, $\mathrm{PGL}(E)$ est le groupe des homographies de E, voir la définition 23.23.

(1) **Surjectif** Si $\phi \in \mathrm{PGL}(\mathbb{K}^2)$ nous avons $\phi = \psi(\varphi \circ \phi^{-1})$.

(2) **Injectif** Si $\psi(\varphi_a) = \psi(\varphi_b)$ alors

$$\varphi_a^{-1} \circ \varphi = \varphi_b^{-1} \circ \varphi, \tag{23.69}$$

d'où nous déduisons $\varphi_a^{-1} = \varphi_b^{-1}$ parce que φ est une bijection.

□

23.4.5 Birapport

23.40.
Tout le monde semble définir le birapport en identifiant $P(\mathbb{K}^2)$ à $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$. Bien entendu, personne ne semble s'être attribué la mission d'expliciter la dépendance du birapport en le choix de l'identification. Je le fais à la définition 23.42.

Mais cette définition dépend du choix d'identification $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$, comme le montre l'exemple 23.44. J'ai donc défini des classes d'identifications possibles $A(\varphi)$ en 23.38. Et je démontre la proposition 23.45 que si $\varphi_a \in A(\varphi)$ alors les birapports construits à partir de φ et φ_a sont identiques.

Question : pourquoi personne ne semble faire ce travail ? En quoi l'identification φ_0 que tout le monde utilise est plus canonique qu'une autre ? Est-ce que l'on peut décrire simplement les classes $A(\varphi)$? Le groupe qui conserve le birapport associé à φ est-il isomorphe au groupe qui conserve le birapport associé à φ' ? Quels que soient φ et φ' ?

Suis-je la seule personne au monde à m'être demandé si le birapport était un objet canonique ?

23.41.
Une utilisation très intéressante du birapport dans la vidéo « The Cross-Ratio » de numberphile : https://www.youtube.com/watch?v=ffvojZONF_A

DEFooBFSKooDwzwmO

Proposition-Définition 23.42.
Soit une droite projective $d = P(E)$ et trois points distincts A, B et C sur cette droite. Soit une bijection $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. Si X est un point de d alors nous nommons le ***birapport*** *de X par rapport à A, B et C l'élément de $\hat{\mathbb{K}}$ donné par*

$$[A, B, C, X]_\varphi = (\varphi \circ \phi)(X) \tag{23.70}$$

où $\phi\colon d \to P(\mathbb{K}^2)$ est l'unique homographie telle que SUBEQooYWMSooYlFKQv

$$\phi(A) = \varphi^{-1}(\infty) \tag{23.71a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{23.71b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{23.71c}$$

Démonstration. Nous devons prouver qu'il existe effectivement une unique homographie vérifiant les conditions (23.71).

(1) A, B, C **est un repère projectif de** $P(E)$ Voir le lemme 23.34.

(2) $\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)$ **est un repère projectif de** $P(\mathbb{K}^2)$ Lemme 23.34 ou corolaire 23.37 au choix.

(3) **Conclusion** Le théorème 23.32(2) nous donne existence et unicité d'une homographie $P(E) \to P(\mathbb{K}^2)$ envoyant le premier repère sur le second.

□

Remarque 23.43.
La majorité des sources ne parlent pas de la dependence du birapport en le choix de φ parce que tout le monde ne semble ne considérer que $\varphi = \varphi_0$ définie en (23.63). Il est cependant naturel de se demander si la définition dépend effectivement du choix de φ. La réponse est oui : ça dépend du choix.

EXooYCOYooWFSfUv

Exemple 23.44 (Une autre identification qui ne va pas bien)**.**
Nous montrons que l'identification $\varphi_d\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ donnée en (23.66) ne donne pas lieu au même birapport que celui de φ_0.

Nous travaillons le birapport sur la droite projective la plus simple : $P(\mathbb{K}^2)$ avec $\mathbb{K} = \mathbb{R}$. Prenons pour la simplicité $A = [1, 0]$, $B = [0, 1]$ et $C = [1, 1]$. Alors l'homographie demandée dans la définition de $[.,.,.,.]_{\varphi_0}$ est $\phi_0 = \mathrm{Id}$. Par conséquent,

$$\big[A, B, C, [k_1, k_2]\big]_{\varphi_0} = (\varphi_0 \circ \phi_0)[k_1, k_2] = \varphi_0[k_1, k_2] = \frac{k_1}{k_2}. \tag{23.72}$$

En ce qui concerne le birapport définit par φ_d nous avons

$$\varphi_d^{-1}(\infty) = [1,0] \tag{23.73a}$$
$$\varphi_d^{-1}(0) = [0,1] \tag{23.73b}$$
$$\varphi_d^{-1}(1) = [1,2], \tag{23.73c}$$

de telle sorte que nous cherchons une homographie $\phi_d\colon P(\mathbb{K}^2) \to P(\mathbb{K}^2)$ telle que

$$\phi_d[1,0] = [1,0] \tag{23.74a}$$
$$\phi_d[0,1] = [0,1] \tag{23.74b}$$
$$\phi_d[1,1] = [1,2] \tag{23.74c}$$

L'homographie $\phi_d[k_1,k_2] = [k_1,2k_2]$ convient et nous avons

$$\big[A,B,C,D,[k_1,k_2]\big]_{\varphi_d} = (\varphi_d \circ \phi_d)[k_1,k_2] = \varphi_d[k_1,2k_2] = \frac{k_1}{2k_2} \tag{23.75}$$

dès que $k_1k_2 < 0$. Nous avons donc bien trouvé

$$[A,B,C,X]_{\varphi_0} \neq [A,B,C,X]_{\varphi_d}. \tag{23.76}$$

△

Le birapport n'est pas un objet tout à fait canonique parce qu'il dépend effectivement du choix de l'identification entre $P(\mathbb{K}^2)$ et $\hat{\mathbb{K}}$.

PROPooTFMQooIOQGvs

Proposition 23.45.
Soit une bijection $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. Si $\varphi_a \in A(\varphi)$[6] *alors les birapports construits sur φ et φ_a coïncident.*

Démonstration. Soient trois points distincts $A,B,C \in P(E)$, et $X \in P(E)$. Nous avons

$$[A,B,C,X]_{\varphi} = (\varphi \circ \phi)(X) \tag{23.77}$$

où $\phi\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi(A) = \varphi^{-1}(\infty) \tag{23.78a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{23.78b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{23.78c}$$

Et

$$[A,B,C,X]_{\varphi_a} = (\varphi_a \circ \phi_a)(X) \tag{23.79}$$

où $\phi_a\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi_a(A) = \varphi_a^{-1}(\infty) \tag{23.80a}$$
$$\phi_a(B) = \varphi_a^{-1}(0) \tag{23.80b}$$
$$\phi_a(C) = \varphi_a^{-1}(1). \tag{23.80c}$$

Il est facile de voir que $\phi_a = \varphi_a^{-1} \circ \varphi\phi$. En effet, cela est une homographie parce que $\varphi_a \in A(\varphi)$, et parce que la composée d'homographies est une homographie. De plus,

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(A) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(\infty) = \varphi_a(\infty) \tag{23.81a}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(B) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(0) = \varphi_a(0) \tag{23.81b}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(C) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(1) = \varphi_a(1). \tag{23.81c}$$

Au final nous avons :

$$[A,B,C,X]_{\varphi_a} = (\varphi_a \circ \varphi_a^{-1} \circ \varphi \circ \phi)(X) = (\varphi \circ \phi)(X) = [A,B,C,X]_{\varphi}. \tag{23.82}$$

□

6. $A(\varphi)$ définie en 23.38.

Remarque 23.46.
Tout le monde semble ne considérer que l'identification usuelle $\varphi_0 : P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ donnée par $\varphi_0[k_1, k_2] = k_1/k_2$. Toute la discussion concernant la dépendance du birapport en le choix de l'identification (y compris la définition des classes $A(\varphi)$) peut être sautée en disant qu'on ne considéra que φ_0.

Et c'est ce que nous allons faire : sauf avis contraire, nous utiliserons le birapport associé à l'identification φ_0.

LEMooCOFTooVGKdVO

Lemme 23.47 ([613])**.**
Nous avons

$$[A, B, C, X]_\varphi = \begin{cases} \infty & \text{si et seulement si } X = A \\ 0 & \text{si et seulement si } X = B \\ 1 & \text{si et seulement si } X = C. \end{cases} \tag{23.83}$$

Démonstration. Par définition $[A, B, C, X]_\varphi = \varphi \circ \phi(X)$. Nous avons donc équivalence entre les affirmations suivantes :

— $[A, B, C, X]_\varphi = \infty$
— $(\varphi \circ \phi)(X) = \infty$
— $\phi(X) = \varphi^{-1}(\infty)$
— $\phi(X) = \phi(A)$
— $A = X$

parce que ϕ et φ sont des bijections.

Le même raisonnement tient pour les deux autres. □

PROPooKQZRooVCXPLW

Proposition 23.48.
Autres petites propriétés faciles ... Soit une droite projective $d = P(E)$ et trois points distincts $A, B, C \in d$.

ITEMooOIPZooQFFYIn

(1) Les points A, B, C et X sont distincts si et seulement si $[A, B, C, X] \in \mathbb{K}\backslash\{0, 1\}$.

ITEMooBEBEooVfiJXY

(2) Pour tout $k \in \hat{\mathbb{K}}$, il existe un unique $X \in d$ tel que $[A, B, C, X] = k$.

Démonstration. Notons pour le point (1) que l'énoncé demande déjà que A, B et C soient distincts. Sinon le birapport n'est pas défini.

(1) **(1)**
Les points A, B et C sont distincts par hypothèse. Vu le lemme 23.47, pour que X soit distincts de A, B et C il faut et il suffit que le birapport ne soit ni ∞ ni 1 ni 0. Donc $\mathbb{K}\backslash\{0, 1\}$.

(2) **(2)**
Nous avons $[A, B, C, X] = \phi(X)$ où $\phi : P(E) \to P(\mathbb{K}^2)$ est une homographie et donc une bijection par la proposition 23.22(1). Donc oui, pour tout éléments de $P(\mathbb{K}^2)$ il existe un unique élément de $P(E)$ dont le birapport par rapport à A, B et C soit cet élément.

Notons encore une fois que nous avons identifié $P(\mathbb{K}^2)$ à $\hat{\mathbb{K}}$ par la bijection (23.63). □

PROPooMGYDooHqSoJs

Proposition 23.49 ([613])**.**
Nous considérons deux droites projectives d et d' ainsi que 4 points sur chacune. Nous les nommons $A_1, A_2, A_3, A_4 \in d$ et $A'_1, A'_2, A'_3, A'_4 \in d'$. Nous supposons que A_1, A_2, A_3 sont distincts et que A'_1, A'_2, A'_3 également. Alors il y a équivalence entre

ITEMooIDKBooXHnNDi

(1) Il existe une homographie $\phi : d \to d'$ telle que $\phi(A_i) = A'_i$ pour $i = 1, 2, 3, 4$,

ITEMooCDWAooIckJwT

(2) égalité des birapports :

$$[A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4]. \tag{23.84}$$

Dans ce cas, l'homographie est unique.

Démonstration. Nous divisons la preuve en trois parties évidentes.

(1) **(1) implique (2)**

Nous avons une homographie $\mu'\colon d' \to \hat{\mathbb{K}}$ telle que $\mu'(A'_1) = \infty$, $\mu'(A'_2) = 0$ et $\mu'(A'_3) = 1$. En composant [7] avec l'homographie $\phi\colon d \to d'$ de l'hypothèse nous avons une homographie $\mu' \circ \phi\colon d \to \hat{\mathbb{K}}$ qui vérifie

$$(\mu' \circ \phi)(A_1) = \mu'(A'_1) = \infty \tag{23.85a}$$
$$(\mu' \circ \phi)(A_2) = \mu'(A'_2) = 0 \tag{23.85b}$$
$$(\mu' \circ \phi)(A_3) = \mu'(A'_3) = 1, \tag{23.85c}$$

ce qui signifie que $\mu' \circ \phi$ est l'homographie qui définie le birapport sur d. Par conséquent

$$[A_1, A_2, A_3, A_4] = (\mu' \circ \phi)(A_4) = \mu'(A'_4) = [A'_1, A'_2, A'_3, A'_4]. \tag{23.86}$$

(2) **(2) implique (1)**

La partie $\{A_1, A_2, A_3\}$ est un repère projectif de d par le lemme 23.34. Idem pour $\{A'_1, A'_2, A'_3\}$. Nous considérons les homographies $\mu\colon d \to \hat{\mathbb{K}}$ et $\mu'\colon d' \to \hat{\mathbb{K}}$ définissant les birapports par rapport à ces repères. Ces homographies vérifient, en utilisant l'hypothèse :

$$\mu(A_4) = [A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4] = \mu'(A'_4). \tag{23.87}$$

Et de plus

$$\mu(A_1) = [A_1, A_2, A_3, A_1] = \infty = \mu'(A'_1) \tag{23.88a}$$
$$\mu(A_2) = [A_1, A_2, A_3, A_2] = 0 = \mu'(A'_2) \tag{23.88b}$$
$$\mu(A_3) = [A_1, A_2, A_3, A_3] = 1 = \mu'(A'_3). \tag{23.88c}$$

Autrement dit : $\mu(A_i) = \mu'(A'_i)$ pour tout $i = 1, 2, 3, 4$. Nous nous inspirons de ce diagramme :

$$\begin{array}{ccc} d & \xrightarrow{\ \phi\ } & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{23.89}$$

La composée $\phi = \mu'^{-1} \circ \mu$ vérifie

$$(\mu'^{-1} \circ \mu)(A_i) = A'_i \tag{23.90}$$

pour tout i, et est une homographie.

(3) **Unicité**

Le fait qu'une homographie vérifiant $\phi(A_i) = A'_i$ pour $i = 1, 2, 3$ soit unique découle du fait qu'il existe une unique homographie portant un repère projectif sur un autre. A fortiori la condition $\phi(A_4) = A'_4$ ne retire rien à l'unicité.

□

Théorème 23.50 ([613])**.**
Une bijection entre deux droites projectives est une homographie si et seulement si elle conserve le birapport.

Démonstration. Chacun des deux sens séparément.

7. Proposition 23.22(3), la composition est encore une homographie.

(1) $\Rightarrow$

Soit une homographie $\phi\colon d \to d'$ entre deux droites projectives. Nous devons prouver que pour tout choix 4 points A, B, C, X dans d (dont A, B et C sont distincts) nous avons

$$[A, B, C, X] = [\phi(A), \phi(B), \phi(C), \phi(X)]. \tag{23.91}$$

EQooVFIOooCdumAe

Nous nommons $\mu\colon d \to \hat{\mathbb{K}}$ l'homographie qui donne le birapport sur d par rapport à A, B et C, et $\mu'\colon d' \to \hat{\mathbb{K}}$ celle qui donne le birapport sur d' par rapport à $\phi(A)$, $\phi(B)$, $\phi(C)$. Voici un diagramme de la situation :

$$\begin{array}{ccc} d & \xrightarrow{\quad\phi\quad} & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{23.92}$$

Nous prouvons maintenant que $\mu' = \mu \circ \phi^{-1}$. En effet :

$$(\mu \circ \phi^{-1})\big(\phi(A)\big) = \mu(A) = \infty \tag{23.93a}$$
$$(\mu \circ \phi^{-1})\big(\phi(B)\big) = \mu(B) = 0 \tag{23.93b}$$
$$(\mu \circ \phi^{-1})\big(\phi(C)\big) = \mu(C) = 1. \tag{23.93c}$$

Par conséquent le birapport à droite dans (23.91) peut se calculer à l'aide de $\mu \circ \phi^{-1}$:

$$[\phi(A), \phi(B), \phi(C), \phi(X)] = \mu'\big(\phi(X)\big) = (\mu \circ \phi^{-1})\big(\phi(X)\big) = \mu(X) = [A, B, C, X]. \tag{23.94}$$

La première implication est prouvée.

(2) $\Leftarrow$

Soit une bijection $f\colon d \to d'$ conservant le birapport, ainsi que trois points distincts A, B et C dans d. Vu que f est une bijection les points $f(A)$, $f(B)$ et $f(C)$ sont distincts dans d'. Par le lemme 23.34 et le théorème 23.32(2), il existe une unique homographie $\phi\colon d \to d'$ telle que $\phi(A) = f(a)$, $\phi(B) = f(B)$ et $\phi(C) = f(C)$. Pour tout $X \in d$ nous avons

SUBEQooVCSBooWvQLih

$$[f(A), f(B), f(C), f(X)] = [A, B, C, X] \tag{23.95a}$$
SUBEQooIYQKooJEnyo
$$= [\phi(A), \phi(B), \phi(C), \phi(X)] \tag{23.95b}$$
SUBEQooWYUPooQMZrrU
$$= [f(A), f(B), f(C), \phi(X)]. \tag{23.95c}$$

Justifications :

— (23.95a) parce que f conserve le birapport par hypothèse.

— (23.95b) parce que ϕ conserve le birapport étant une homographie (c'est le premier sens du présent théorème)

Nous nommons $\mu'\colon d' \to \hat{\mathbb{K}}$ l'homographie donnant le birapport par rapport aux points $f(A)$, $f(B)$, $f(C)$. Alors le résultat (23.95) se lit

$$\mu'\big(f(X)\big) = \mu'\big(\phi(X)\big). \tag{23.96}$$

Mais comme μ' est une bijection (proposition 23.22(1)) cela implique $f(X) = \phi(X)$. Vu que nous avons fait ce raisonnement pour un X quelconque dans d nous avons $f = \phi$, ce qui prouve que f est une homographie.

□

Lemme 23.51.
Soient a, b, c distincts sur la droite projective $D = P(E)$. Soient $x, y \in E$ tels que $\pi_E(x) = a$, $\pi_E(y) = b$, $\pi_E(x + y) = c$. Alors

$$d = \pi_E(\lambda x + \mu y) \tag{23.97}$$

si et seulement si

$$[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{23.98}$$

Démonstration. Étant donné que a et b sont distincts, les vecteurs x et y forment une base de E. Soit $f\colon E \to \mathbb{K}^2$ un isomorphisme qui envoie (x,y) sur e_1, e_2 où e_i sont les vecteurs de base de $\mathbb{K}^2$. Ensuite nous considérons $g\colon P(E) \to P(\mathbb{K}^2)$, l'homographie associée à f. Par définition $g\big(\pi_E z\big) = \pi_{\mathbb{K}^2}\big(f(z)\big)$. Par f nous avons

$$a \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad b \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{23.99}$$

Donc par g nous avons

$$a \mapsto \infty \quad b \mapsto 0. \tag{23.100}$$

Nous avons aussi $f(\lambda x + \mu y) = (\lambda, \mu)$ et

$$\begin{aligned} g(c) &= g\big(\pi_E(x+y)\big) && \text{(23.101a)}\\ &= \pi_F f(x+y) && \text{(23.101b)}\\ &= \pi_F(f(x)+f(y)) && \text{(23.101c)}\\ &= \pi_F \begin{pmatrix} 1 \\ 1 \end{pmatrix} && \text{(23.101d)}\\ &= 1. && \text{(23.101e)} \end{aligned}$$

La dernière égalité est le fait que la direction $(1,1)$ dans $\mathbb{R}^2$ est représentée par le point $x = 1$ sur la droite $y = 1$ qui est notre « représentation » de la droite affine. L'application g a donc toutes les propriétés qu'il faut pour être l'application qui définit le birapport. Nous avons donc bien $g(d) = [a,b,c,d]$.

D'une part si $d = \pi_E(\lambda x + \mu y)$ alors

$$g(d) = \pi_{\mathbb{K}^2} f(\lambda x + \mu y) = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{23.102}$$

Dans l'autre sens si $[a,b,c,d] = \pi_{\mathbb{K}^2}(\lambda,\mu)$ alors supposons que $g(d) = \pi_{\mathbb{K}^2}(\lambda,\mu)$ avec $d = \pi_E(v)$ alors

$$g\pi_E v = \pi_{\mathbb{K}^2} f(v), \tag{23.103}$$

ce qui implique $f(v) = \alpha(\lambda,\mu)$ pour un certain $\alpha \in \mathbb{K}$. Par conséquent $v = \alpha(\lambda x + \mu y)$ et $d = \pi_E(\lambda x + \mu y)$. □

23.5 Coordonnées homogènes

Soit E un espace vectoriel de dimension $n+1$ et une base $\{e_0, \dots, e_n\}$ de E. Soit $M \in P(E)$ et $u \in E$ un élément engendrant M. Au point M nous voudrions associer les coordonnées $(x_0, \dots, x_n)$ de u dans E. Notons que toutes les coordonnées de u ne sont jamais nulles en même temps parce que u doit indiquer une direction. Nous savons par ailleurs que les coordonnées $(x_0, \dots, x_n)$ indiquent le même point de $P(E)$ que les coordonnées $(x'_0, \dots, x'_n)$ si et seulement si $x_i = \lambda x_i$.

DEFooLWMHooMWxAFq

Définition 23.52.
La classe d'équivalence de $(x_0, \dots, x_n)$ est la ***coordonnée homogène*** *de M. Nous la notons $(x_0 : \dots : x_n)$.*

Si nous avons une base $\{e_i\}$ de $\mathbb{R}^n$ nous associons à $M \in P(E)$ les coordonnées $(X : Y : T)$. Mais si on prend la base $\{2e_1, e_2, \dots, e_n\}$, les coordonnées du même point deviennent $(X/2 : Y : T)$ alors que du point de vue de l'espace projectif, rien n'a été changé : la classe de e_1 est la même que celle de $2e_1$. Les coordonnées homogènes[8] ne sont donc pas intrinsèques.

8. Définition 23.52.

23.5.1 Curiosité : matrice de translation

Si E est un espace vectoriel, l'espace projectif $P(E)$ est l'ensemble des classes d'équivalence dans E pour la relation $v \sim \lambda v$ pour tout $\lambda \neq 0$.

Il se fait que l'étude de $\mathbb{R}^3$ peut être fait à partir de $\mathbb{R}^4$ en considérant les coordonnées homogènes sur $P(\mathbb{R}^4)$. Plus précisément, nous considérons

$$\begin{aligned}\varphi\colon \mathbb{R}^3 &\to P(\mathbb{R}^4)\\ (x,y,z) &\mapsto \big[(x,y,z,1)\big].\end{aligned} \tag{23.104}$$

Cela est injectif mais pas surjectif parce que les éléments de la forme $[(x,y,z,0)]$ ne sont pas atteints. Ces éléments sont alors dits « à l'infini » .

Nous aurions pu placer $\mathbb{R}^3$ dans $\mathbb{R}^4$ de nombreuses autres manières ; chacune aurait donné une notion différente de « point à l'infini ».

Nous allons maintenant montrer une petite curiosité qui a une grande importance en informatique, lors de la manipulation d'objets 3D. Nous considérons la bijection

$$\begin{aligned}\varphi\colon \mathbb{R}^3 &\to \mathbb{R}^4\\ (x,y,z) &\mapsto (x,y,z,1).\end{aligned} \tag{23.105}$$

Soit l'opérateur de translation $T_a\colon \mathbb{R}^3 \to \mathbb{R}^3$. En considérant la matrice

$$T_a^h = \begin{pmatrix} 1 & 0 & 0 & a_x \\ 0 & 1 & 0 & a_y \\ 0 & 0 & 1 & a_z \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{23.106}$$

nous avons

$$T_a = \varphi^{-1} \circ T_a^h \circ \varphi. \tag{23.107}$$

Autrement dit, ce passage de $\mathbb{R}^3$ à $\mathbb{R}^4$ permet de voir les translations comme des matrices, et c'est bien pratique.

Si $R\colon \mathbb{R}^3 \to \mathbb{R}^3$ est une rotation, la matrice correspondante sur $\mathbb{R}^4$ est

$$R^h = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}. \tag{23.108}$$

Elle se combine assez bien avec une translation parce que le produit donne

$$T_a^h R^h = \begin{pmatrix} & R & & \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} & & & 1 \end{pmatrix}. \tag{23.109}$$

C'est-à-dire que la translation et la rotation restent assez visible dans la matrice composée.

Note : pour la composition $R^h T_a^h$, c'est beaucoup moins vrai.

23.5.2 Dualité

Soit E un espace vectoriel de dimension $n+1$. Une forme linéaire non nulle est un élément de E^*, mais aussi un représentant d'un élément de $P(E^*)$.

Le noyau d'une forme linéaire ω est un hyperplan. Le noyau de la forme linéaire $\lambda\omega$ étant le même hyperplan, l'hyperplan est donné par toute la classe de ω dans $P(E^*)$. Nous avons donc une bijection

$$P(E^*) \leftrightarrow \{\text{hyperplans vectoriels de } E\}. \tag{23.110}$$

Soit un espace vectoriel E de dimension 3 muni d'une base $\{e_1, e_2, e_3\}$ à partir de laquelle nous construisons la base duale $\{e_1^*, e_2^*, e_3^*\}$ de l'espace dual E^*. À un élément $m \in P(E^*)$ nous associons la droite

$$H_m = \{(X:Y:T) \text{ tel que } m(X,Y,T) = 0\} \tag{23.111}$$

dans $P(E)$. Si les coordonnées homogènes de m étaient $(u : v : w)$ alors l'équation de la droite H_m est

$$uX + vY + wT = 0. \tag{23.112}$$

Eqezgpmk

En effet si $\omega \in E^*$ est un représentant de m alors $\omega = \lambda(ue_1^* + ve_2^* + we_3^*)$ et l'équation (23.112) est indépendante de λ ainsi que du choix du représentant dans E du point $(X : Y : T)$ dans $P(E)$.

Si les points m_1 et m_2 sont distincts dans $P(E^*)$, ils donnent deux droites $m_1(X, Y, T) = 0$ et $m_2(X, Y, T) = 0$. Les points de la droite qui joint m_1 à m_2 dans $P(E^*)$ sont de la forme $\lambda m_1 + \mu m_2$ et ils sont associés à l'équation

$$\lambda m_1(X, Y, T) + \mu m_2(X, Y, T) = 0 \tag{23.113}$$

qui sont encore des droites dans $P(E)$. Toutes ces droites passent par le point d'intersection des droits associées à m_1 et m_2. Nous avons donc

$$\bigcap_{\lambda,\mu} H_{\lambda m_1 + \mu m_2} = H_{m_1} \cap H_{m_2}. \tag{23.114}$$

Lemme 23.53.
L'application

$$\begin{aligned} P(E^*) &\to \{\textit{droites dans } P(E)\} \\ m &\mapsto H_m \end{aligned} \tag{23.115}$$

est une bijection.

Démonstration. Une droite dans $P(E)$ est donnée en coordonnées homogènes par une équation $aX + bY + cT = 0$. Cette droite est décrite par le point $(a : b : c)$ dans $P(E^*)$. Ce dernier correspond à la direction de la forme $ae_1^* + be_2^* + ce_3^*$. Cela prouve que l'application est surjective.

Pour l'injectivité, si $m_1 \neq m_2$ dans $P(E^*)$, les formes ω_1 et ω_2 associées dans E^* ne sont pas multiples l'une de l'autre. Donc les équations

$$a_1X + b_1Y + z_1T = 0 \tag{23.116}$$

et

$$a_2X + b_2Y + z_2T = 0 \tag{23.117}$$

n'ont pas de solutions communes et décrivent donc des droites distinctes. □

LemjXywjH

Lemme 23.54.
Trois points distincts m_1, m_2 et m_3 dans $P(E^*)$ sont alignés si et seulement si les droites H_{m_1}, H_{m_2} et H_{m_3} sont distinctes et concourantes.

Démonstration. Supposons avoir trois points alignés, c'est-à-dire

$$m_3 = m_1 + \mu(m_2 - m_1). \tag{23.118}$$

EqXyfbmF

Soit $X : Y : T$ le point d'intersection de H_{m_1} avec H_{m_2}. Alors $m_1(X, Y, T) = m_2(X, Y, T) = 0$. En tenant compte de (23.118) nous avons alors évidemment $m_3(X, Y, T) = 0$.

Supposons maintenant que les trois droites H_{m_i} soient concourantes. Nous avons donc un point $(X : Y : T)$ dans $P(E)$ tel que $m_i(X, Y, T) = 0$. Si m_i est la classe de $a_ie_1^* + b_ie_2^* + c_ie_3^*$ alors nous avons le système

$$\begin{cases} a_1X + b_1Y + c_1T = 0 & \text{(23.119a)} \\ a_2X + b_2Y + c_2T = 0 & \text{(23.119b)} \\ a_3X + b_3Y + c_3T = 0. & \text{(23.119c)} \end{cases}$$

Afin que cela ait une solution non triviale nous devons avoir

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = 0, \tag{23.120}$$

c'est-à-dire que les points (a_i, b_i, c_i) soient alignés. □

En tenant compte de ce qui a été dit, une droite dans $P(E^*)$ est constituée de points qui fournissent des droites concourantes dans $P(E)$. Donc une droite de $P(E^*)$ se caractérise par un point de $P(E)$ (l'intersection) de la façon suivante. Un point $M_d \in P(E)$ donne lieu à un **faisceau de droites** passant par M_d. Chacune de ces droites donne lieu à un point de $P(E^*)$ et tous ces points sont alignés. Nous avons ainsi construit la droite d dans $P(E^*)$ correspondante au point M_d de $P(E)$.

23.5.3 Polynômes

Soit l'espace projectif de dimension n avec ses coordonnées homogènes $(X_0 : \ldots : X_n)$. Nous considérons l'espace affine $H \equiv X_n = 1$ dans l'espace vectoriel E de dimension $n+1$. Nous considérons pour H un repère affine ayant pour origine le point $(0, \ldots, 0, 1)$. Considérons un polynôme homogène P sur le corps $\mathbb{K}$. L'équation

$$P(X_0, \ldots, X_n) = 0 \tag{23.121}$$

sur l'espace vectoriel E descend immédiatement à l'espace projectif : étant donné que P est homogène nous avons $P(u) = 0$ si et seulement si $P(\lambda u) = 0$.

Nous essayons de décrire l'ensemble A des points de $P(E)$ satisfaisant $P(X_0, \ldots, X_n) = 0$. Nous savons que les éléments de $P(E)$ ont chacun un représentant soit dans H soit sur la droite à l'infini. Ceux de A ayant un représentant dans H sont d'équation

$$Q(x_0, \ldots, x_{n-1}) = 0 \tag{23.122}$$

où Q est le polynôme donné par $Q(x_0, \ldots, x_{n-1}) = P(x_0, \ldots, x_{n-1}, 1)$. Les points de A ayant un représentant sur la droite à l'infini s'obtiennent par l'équation

$$R(x_0, \ldots, x_{n-1}) = 0 \tag{23.123}$$

où R est le polynôme donné par $R(x_0, \ldots, x_{n-1}) = P(x_0, \ldots, x_{n-1}, 0)$.

Exemple 23.55.
Nous considérons la conique projective

$$X^2 - XT - Y^2 - T^2 = 0. \tag{23.124}$$

Elle est décomposée en deux parties : une dans l'espace affine « normale » et une à l'infini. La première s'obtient en posant $T = 1$ dans (23.124) :

$$x^2 - x - y^2 - 1 = 0. \tag{23.125}$$

L'autre est obtenue en posant $T = 0$:

$$x^2 - y^2 = 0. \tag{23.126}$$

La partie à l'infini est donc composée de deux points : $(1 : 1 : 0)$ et $(1 : -1 : 0)$.

Le graphique de l'équation (23.125) est donné à la figure 23.2. Nous y voyons que les asymptotes sont effectivement données par les directions $(1, 1)$ et $(1, -1)$ dans le plan.

△

Nous pouvons tenter de faire l'exercice inverse : considérer une conique dans $\mathbb{R}^2$, la voir comme une partie d'une conique dans l'espace projectif et trouver les points à l'infini qui la complètent.

Exemple 23.56.
La droite projective usuelle est donnée par la droite affine $y - 1 = 0$. L'homogénéisation donne $y - z = 0$ et par conséquent la partie à l'infini est donnée par $y = 0$, c'est-à-dire la direction $(1, 0)$ comme il se doit. △

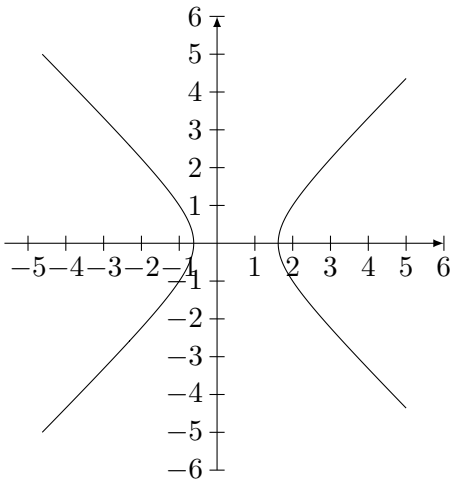


FIGURE 23.2: Le graphique de $x^2 - x - y^2 - 1 = 0$. LabelFigProjPoly

Exemple 23.57.
Prenons la conique

$$x^2 + xy + y^3 - 2 = 0. \tag{23.127}$$

D'abord nous homogénéisons cette équation pour la voir dans $\mathbb{R}^3$:

$$x^2z + xyz + y^3 - 2z^3 = 0. \tag{23.128}$$

Les points à l'infini sont ceux qui correspondent à $z = 0$, c'est-à-dire la droite donnée en coordonnées homogènes par $(1 : 0 : 0)$. △

23.6 La sphère de Riemann $P_1(\mathbb{C})$

DEFooSZGNooTzFYbh

Définition 23.58.
*La **sphère de Riemann** est l'espace projectif modelé sur $\mathbb{C}^2$: en vertu des notations données à la page 2025, c'est*

$$P_1(\mathbb{C}) = P(\mathbb{C}^2). \tag{23.129}$$

Nous parlerons aussi de sphère de Riemann en tant que compactification à un point de $\mathbb{C}$ dans la sous-section 48.5.1.

L'ensemble $P_1(\mathbb{C})$ est le quotient $\mathbb{C}^2\backslash\{(0,0)\}/\sim$ où $\sim$ est la relation d'équivalence de $\mathbb{C}$-colinéarité dans $\mathbb{C}^2$.

LEMooKWZDooEIraSJ

Lemme 23.59.
L'application

$$\begin{aligned}\varphi_0 \colon P_1(\mathbb{C}) &\to \mathbb{C} \cup \{\infty\} \\ [z_1, z_2] &\mapsto \begin{cases} \frac{z_1}{z_2} & si\ z_2 \neq 0 \\ \infty & si\ z_2 = 0 \end{cases}\end{aligned} \tag{23.130}$$

EQooKJIZooZihzuU

est une bijection qui respecte la conjugaison complexe : $\varphi_0\big([z_1, z_2]^*\big) = \varphi_0\big([z_1, z_2]\big)^*$.

Démonstration. Notons d'abord que la définition a un sens parce que si un représentant que $[z_1, z_2]$ est de la forme $(z, 0)$ alors ils sont tous de cette forme. L'affirmation « $z_1 \neq 0$ dans $[z_1, z_2]$ » a donc un sens.

(1) **Injectif** Supposons $\varphi_0\big([z_1, z_2]\big) = \varphi_0\big([t_1, t_2]\big)$.
Si les deux membres sont égaux à ∞ alors nous avons $z_2 = t_2 = 0$, et alors avec $\lambda = z_1/t_1$ nous avons $(z_1, z_2) = \lambda(t_1, t_2)$, ce qui prouve que $[z_1, z_2] = [t_1, t_2]$.
Si les deux membres sont égaux à zéro alors $z_1 = t_1 = 0$ et le même raisonnement tient.

Sinon nous avons $z_1/z_2 = t_1/t_2$ où tous les nombres sont non nuls. Cela donne

$$z_2 = \frac{t_2}{t_1} z_1, \tag{23.131}$$

et donc

$$\frac{t_1}{z_1}(z_1, z_2) = (t_1, t_2), \tag{23.132}$$

qui montre qu'au niveau des classes, $[z_1, z_2] = [t_1, t_2]$.

(2) **Surjectif**
Nous avons

$$\infty = \varphi_0\big([1, 0]\big) \tag{23.133}$$

et si $z \neq \infty$ nous avons $z = \varphi_0\big([z, 1]\big)$.

□

23.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$

SUBSECooQPRLooAjMNqp

Étant donné que nous sommes partis pour faire de la géométrie dans $\mathbb{C}$ et même dans $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, autant nous armer des équations de cercles et de droites dans $\mathbb{C}$, ainsi que de quelques notions adjacentes.

Remarque 23.60.
La définition 23.2 parle de plan et de droites projectives. Ici nous ne sommes pas dans ce cadre parce que nous travaillons sur $P_1(\mathbb{C})$ où $\mathbb{C}$ n'est certainement pas un espace de dimension 3. Les droites dont nous allons parler ne sont pas des droites projectives avec leur point à l'infini.

23.6.1.1 Équation complexe d'une droite

L'équation d'une droite dans $\mathbb{R}^2$ est $d \equiv ax + by = c$ avec $a, b, c \in \mathbb{R}$ et a, b non nuls en même temps. En posant $z = x + iy$ nous voulons exprimer l'équation en termes de z au lieu de x et y. Nous avons[616]

$$x = \frac{z + \bar{z}}{2}, \quad y = \frac{z - \bar{z}}{2i}, \tag{23.134}$$

et nous pouvons écrire $d \equiv a(z + \bar{z}) - ib(z - \bar{z}) = 2c$, ou encore $d \equiv (a - bi)z + (a + ib)\bar{z} = 2c$. En posant $\omega = a + bi \in \mathbb{C}^*$ et $k = 2c \in \mathbb{R}$ nous avons l'équation

$$\bar{\omega} z + \omega z = k. \tag{23.135}$$

EQooPRCPooVvrHME
DEFooAQSMooWNOzAI

Définition 23.61.
*Une **droite** est une partie de $\hat{\mathbb{C}}$ de la forme*

$$d(\omega, k) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega} z + \omega \bar{z} = k\} \cup \{\infty\} \tag{23.136}$$

avec $\omega \in \mathbb{C}^*$ et $k \in \mathbb{R}$.

Dans $\hat{\mathbb{C}}$, toutes les droites contiennent le point ∞.

PROBooZHHTooIFNwxR

¡¡ Avertissement/question au lecteur !! 23.62
La proposition 22.11 montre que toute inversion transforme un cercle-droite en un cercle-droite, nonobstant d'accepter de prolonger toute droite par ∞.

Est-ce que l'on peut dire que toutes les droites contiennent le point ∞ ?

En donnant ∞ à toutes les droites et à aucun cercle, la proposition 22.11 fonctionne partout en posant $i_C(O) = \infty$ et $i_C(\infty) = O$.

De plus en pensant à la projection stéréographique, ce serait logique : quelle que soit la direction dans laquelle un point s'éloigne de $z = 0$, son image par l'inverse de la projection stéréographique s'approche du pôle nord.

23.6.1.2 Équation complexe d'un cercle

Un cercle de centre $\omega \in \mathbb{C}$ et de rayon r a pour équation $|z-\omega| = r$, et nous avons les équivalences suivantes :

$$|z-\omega| = r \Leftrightarrow |z-\omega|^2 = r^2 \Leftrightarrow (z-\omega)(\bar{z}-\bar{\omega}) = r^2 \Leftrightarrow r\bar{r} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{23.137}$$

Donc un cercle de centre $\omega \in \mathbb{C}$ et de rayon $r \in \mathbb{R}$ a pour équation

$$z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{23.138}$$

EQooDIFBooKRZZoi
DEFooAUDJooVqLDhe

Définition 23.63.
*Un **cercle** dans $\hat{\mathbb{C}}$ est une partie de la forme*

$$C(\omega, r) = \{z \in \mathbb{C} \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2\} \tag{23.139}$$

avec $\omega \in \mathbb{C}$ et $r \in \mathbb{R}$.

Dans la sphère de Riemann, aucun cercle ne contient le point ∞.

Exemple 23.64.
Trouvons le centre et le rayon du cercle d'équation

$$\bar{\omega}z + \omega\bar{z} = kz\bar{z} \tag{23.140}$$

avec $k \neq 0$. En divisant par k et en posant $\sigma = \omega/k$ nous avons :

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = 0. \tag{23.141}$$

Cela est un cercle de centre σ et de rayon $|\sigma|$. En effet si $z \in \mathbb{C}$ vérifie cette équation,

$$|z-\sigma|^2 = (z-\sigma)(\bar{z}-\bar{\sigma}) = \underbrace{z\bar{z} - \bar{\sigma}z - \sigma\bar{z}}_{=0} + |\sigma|^2 = |\sigma|^2, \tag{23.142}$$

c'est-à-dire que tous les points de $\mathbb{C}$ qui vérifient l'équation donnée sont à la distance $|\sigma|$ de σ. En particulier $z = 0$ est sur le cercle. △

23.6.1.3 Cercle-droite

Une chose de bien avec les équations complexes, c'est que nous pouvons écrire les droites et les cercles avec le même type d'équations.

LEMooHKHOooHpBuBZ

Lemme-Définition 23.65 ([616]).
*Un **cercle-droite** est l'ensemble des points $z \in \mathbb{C}$ tels que*

$$az\bar{z} - \bar{\omega}z - \omega\bar{z} = k \tag{23.143}$$

EQooUJAKooEVQNqe

avec $a, k \in \mathbb{R}$ et $\omega \in \mathbb{C}$.

(1) Si $a = 0$, cela est une droite ;
(2) si $a \neq 0$, cela est un cercle.
(3) Un cercle-droite peut être l'ensemble vide.

Démonstration. Si $a = 0$ alors nous tombons tout de suite sur l'équation (23.135). Si $a \neq 0$ alors nous pouvons diviser par a, poser $\sigma = \omega/a$ et $l = k/a$ pour obtenir

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = l, \tag{23.144}$$

EQooNBILooArHPCG

qui est l'équation (23.138) d'un cercle . . .ou pas tout à fait. En effet, (23.144) serait l'équation du cercle de centre σ et de rayon r donné par $l = r^2 - |\sigma|^2$, c'est) dire

$$r^2 = l + |\sigma|^2, \tag{23.145}$$

alors que rien n'assure que le nombre $l + |\sigma|^2$ soit positif. Dans le cas où c'est positif, nous avons bien un cercle. Sinon c'est l'ensemble vide. □

REMooBMAEooHDvNID

Remarque 23.66.
Lorsque nous parlons de cercle-droite, nous parlons de partie de $\mathbb{C}$ et non de $\hat{\mathbb{C}}$ parce que l'équation (23.143) a du mal à traiter le cas $z = \infty$. À cause du fait que nous avons décider de donner le point ∞ à toutes les droites, la fusion des notions de droites et de cercles n'est pas totale ; en tout cas pas en une seule équation.

EXooKFBIooOJKjGL

Exemple 23.67 ([616])**.**
Soit le cercle de centre $\omega = ir$ et de rayon r. Quelle que soit la valeur de $r > 0$, ce cercle passe par le point 0 et l'axe réel lui est tangent. L'équation de ce cercle est :

$$z\bar{r} + irz - ir\bar{z} = 0. \tag{23.146}$$

Vu que $ir \neq 0$ nous pouvons diviser et obtenir

$$\frac{z\bar{z}}{ir} + z - \bar{z} = 0. \tag{23.147}$$

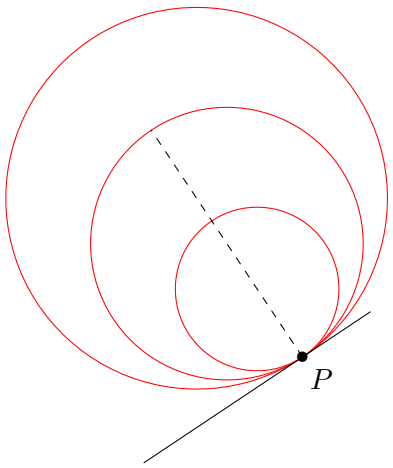
En faisant tendre r vers ∞ nous obtenons $z - \bar{z} = 0$, c'est-à-dire l'équation de la droite réelle.

Cela explique pourquoi il est souvent dit qu'une droite est un cercle dont le rayon est à l'infini. △

NORMooCXVJooMTMqEU

23.68.
Notons que l'exemple 23.67 est générique : prenez une droite ℓ, un point P sur ℓ, et considérez un cercle dont le centre est situé sur la perpendiculaire à ℓ passant par P, et dont le rayon est tel que le cercle passe par P. En prenant $|\omega - P| \to \infty$, l'équation du cercle devient celle de la droite ℓ.

Cela est particulièrement pratique lorsque nous travaillons dans $\hat{\mathbb{C}}$ parce nous y avons une notion précise du point à l'infini. Notons que (peut-être contre-intuitivement), il existe un seul point à l'infini dans $\hat{\mathbb{C}}$. Et ce point est le centre de tous les cercles que l'on veut transformer en droites. Cela pose évidemment la question de savoir comment on définit précisément un cercle dont le centre est réellement ∞.



23.6.1.4 Rotation-homothétie

Définition 23.69.
Une ***rotation-homothétie*** *est une application $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ de la forme $z \mapsto \lambda z$ avec $\lambda \in \mathbb{C}$.*

Le nom provient du fait que si λ est réel, alors $z \mapsto \lambda z$ est une vraie homothétie, et si $\lambda = e^{i\theta}$ alors $z \mapsto e^{i\theta}z$ est une vraie rotation. Pour une valeur $\lambda \in \mathbb{C}$ générique, l'application $z \mapsto \lambda z$ est une composée des deux.

23.6.1.5 Application linéaire

SSUBSooRBCWooSCIQEL

Nous nous en voudrions de ne pas parler d'applications linéaires lorsque nous parlons de géométrie sur $\hat{\mathbb{C}}$. Soit $\alpha \in \mathbb{C}^*$ et $\beta \in \mathbb{C}$. Lorsque nous parlons de l'application linéaire

$$\begin{aligned} f\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \alpha z + \beta, \end{aligned} \tag{23.148}$$

nous entendons implicitement que $f(\infty) = \infty$.

23.6.1.6 Inversion

SSUBSooPOUNooTPilbE

L'inversion d'un cercle de $\mathbb{R}^2$ est définie par la proposition 22.5. De nombreuses propriétés y sont décrites, y compris son écriture complexe dans la proposition 22.10. Tout cela était du temps de $\mathbb{R}^2$ ou de $\mathbb{C}$, mais maintenant nous sommes dans $\hat{\mathbb{C}}$ et nous voulons plus.

DEFooIUTZooWRaXts

Définition 23.70 ([616])**.**
*Soient $\omega \in \mathbb{C}$ et $R \in \mathbb{R}^*$. L'**inversion** de centre ω et de **puissance** R^2 est l'application*

$$\begin{aligned} i\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \begin{cases} \frac{R^2}{\bar{z} - \bar{\omega}} + \omega & si\ z \in \mathbb{C}\backslash\{\omega\} \\ \infty & si\ z = \omega \\ \omega & si\ z = \infty. \end{cases} \end{aligned} \tag{23.149}$$

Notons que grâce aux conventions type $1/0 = \infty$ et $1/\infty = 0$, nous pouvons nous contenter de la première formule pour tout $z \in \hat{\mathbb{C}}$, et nous n'avons en réalité pas besoin de décrire $i(\infty)$ et $i(\omega)$ séparément.

Exemple 23.71.
L'inversion de cercle de centre 0 et de rayon 1 est l'application $z \mapsto \frac{1}{\bar{z}}$, que l'on prolonge avec $0 \mapsto \infty$ et $\infty \mapsto 0$. △

23.6.2 Homographies

La notion d'homographie est la définition 23.20. Pour une homographie $\phi\colon P_1(\mathbb{C}) \to P_1(\mathbb{C})$ nous avons un isomorphisme d'espace vectoriel $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$. Vue la bijection (23.130), nous voulons plutôt travailler avec $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ qui est un ensemble avec lequel nous sommes plus familier. Nous allons donc travailler avec

$$\begin{aligned} &\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}} \\ &\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi_0^{-1}. \end{aligned} \tag{23.150}$$

PROPooTZJBooPpowOo

Proposition 23.72.
L'application $\phi = \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$ est une homographie de $P(\mathbb{C}^2)$ si et seulement si l'application $\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ est de la forme

$$\tilde{\phi}(z) = \frac{az + b}{cz + d} \tag{23.151}$$

avec $a, b, c, d \in \mathbb{C}$ tels que $ad - bc \neq 0$.

Par convention nous posons $z/0 = \infty$ dès que $z \neq 0$; en particulier $\tilde{\phi}(\infty) = a/c$ et $\tilde{\phi}(-d/c) = \infty$.

Démonstration. Nous séparons la condition suffisante de la condition nécessaire.

(1) $\Rightarrow$ La condition $\pi \circ \bar{\phi} = \phi \circ \pi$ (de (23.32)) nous dit que

$$\big[\bar{\phi}(z_1, z_2)\big] = \phi\big([z_1, z_2]\big), \tag{23.152}$$

et comme $\bar{\phi}$ est un isomorphisme d'espace vectoriel nous avons $a, b, c, d \in \mathbb{C}$ vérifiant $ad-cb \neq 0$ pour lesquels

$$\bar{\phi}(z_1, z_2) = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}. \tag{23.153}$$

Soit $z \in \mathbb{C}$. Alors nous avons

$$\tilde{\phi}(z) = (\varphi_0 \circ \phi)[z, 1] = \varphi_0\Big(\big[\bar{\phi}(z, 1)\big]\Big) = \varphi_0\big([az + b, cz + d]\big) = \frac{az + b}{cz + d}. \tag{23.154}$$

EQooNWVZooUClOSd

Il est important de comprendre que cette formule fonctionne pour tout $z \in \mathbb{C}$. En effet nous pourrions avoir un doute sur $z = -d/c$. D'abord si $c = 0$ alors $d \neq 0$ et ce problème n'existe pas : le dénominateur est toujours non nul. Nous avons donc seulement un doute lorsque $c \neq 0$. Dans ce cas,

$$\tilde{\phi}(-d/c) = \frac{-\frac{ad}{c} + b}{0}. \tag{23.155}$$

Mais $c \neq 0$, donc le numérateur est non nul. Or lorsque $z \neq 0$ nous avons posé $z/0 = \infty$, donc dans notre cas,

$$\tilde{\phi}(-d/c) = \infty \tag{23.156}$$

automatiquement, et cela est encodé dans la formule (23.154).
Il nous reste à déterminer $\tilde{\phi}(\infty)$. Nous avons :

$$\tilde{\phi}(\infty) = (\varphi_0 \circ \phi)\big([1, 0]\big) = \varphi_0\big[\bar{\phi}(1, 1)\big] = \varphi_0\big([a, c]\big) = \begin{cases} a/c & \text{si } c \neq 0 \\ \infty & \text{si } c = 0 \end{cases} \tag{23.157}$$

où la distinction entre les deux cas n'est pas fondamentale parce que si $c = 0$ alors $a \neq 0$ et $a/c = \infty$.

(2) $\Leftarrow$
En notant $Z = [z_1, z_2]$ (pour $z_1, z_2 \in \mathbb{C}$) nous avons

$$\phi\big([z_1, z_2]\big) = \varphi^{-1}\left(\frac{a\varphi_0(Z) + b}{c\varphi_0(Z) + d}\right) = \big[a\varphi_0(Z) + b, c\varphi_0(Z) + d\big]. \tag{23.158}$$

Nous définissons $\bar{\phi}$ par son action sur les vecteurs de base : $\bar{\phi}(1, 0) = (a, c)$ et $\bar{\phi}(0, 1) = (b, d)$. Nous avons bien l'isomorphisme d'espace vectoriel $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$ dont nous avons besoin pour la définition 23.20. D'abord le fait que $ad - cb$ soit non nul assure que le $\bar{\phi}$ ainsi défini est bien bijectif[9]. Et de plus ce $\bar{\phi}$ vérifie la condition

$$\begin{aligned} \pi\big(\bar{\phi}(z_1, z_2)\big) &= \pi\big((az_1, cz_1) + (bz_2, dz_2)\big) && \text{(23.159a)} \\ &= [az_1 + bz_2, cz_1 + dz_2] && \text{(23.159b)} \\ &= \big[a\varphi_0(Z) + b, c\varphi_0(Z) + d\big] && \text{(23.159c)} \\ &= \phi\big(\pi(z_1, z_2)\big). && \text{(23.159d)} \end{aligned}$$

Nous avons utilisé la notion de classe pour diviser par z_2 et faire apparaitre $\varphi_0(Z)$.

□

Remarque 23.73.
En prenant les conventions relativement claires $\infty \times a = \infty$ (pour $a \neq 0$) et $\infty \pm a = \infty$ (avec $a \neq \infty$), alors tout est dans la formule

$$\bar{\phi}(z) = \frac{az + b}{cz + d} \tag{23.160}$$

avec $ad - cb \neq 0$. Il n'y a pas besoin de traiter séparément le cas $z = \infty$ ou $z = -d/c$.

9. Si vous n'en êtes pas convaincu, écrivez la matrice de l'application qui envoie $(1, 0)$ sur (a, c) et $(0, 1)$ sur (b, d), et demandez-vous sous quelle condition elle est inversible.

DEFooAMQHooPFUgIa

Définition 23.74.
*Nous aimons tellement l'identification $\varphi_0\colon P(\mathbb{C}^2) \to \hat{\mathbb{C}}$ que nous allons parler d'**homographie** sur $\hat{\mathbb{C}}$ pour les applications $\tilde{\phi}$ de la forme*

$$\tilde{\phi}(z) = \frac{az+b}{cz+d} \tag{23.161}$$

avec $ad - cb \neq 0$.

NORMooCVYKooYvjIeE

23.75.
La proposition 23.72 nous indique que les homographies de $\hat{\mathbb{C}}$ sont de la forme $\phi = \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$ pour une homographie $\phi\colon P(\mathbb{C}^2) \to P(\mathbb{C}^2)$.

PROPooSQFOooRginjJ

Proposition 23.76 ([616])**.**
L'application $h\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ associée à une homographie est soit linéaire, soit de la forme $h = l_1 \circ \iota \circ l_2$ où l_i sont linéaires et ι est l'application $z \mapsto 1/z$.

Démonstration. Commençons par une remarque : lorsque nous parlons d'une application linéaire, c'est au sens de la note 23.6.1.5 qui explique qu'une application linéaire sur $\mathbb{C}$ est automatiquement prolongée à $\hat{\mathbb{C}}$ par $f(\infty) = \infty$.

Soit donc l'application

$$h(z) = \frac{az+b}{cz+d}. \tag{23.162}$$

Si $c = 0$, alors c'est une application linéaire et la preuve est terminée. Nous supposons que $c \neq 0$. Nous posons

$$l_2(z) = cz + d, \tag{23.163}$$

et ensuite $l_1(z) = \alpha z + \beta$ avec α et β à déterminer. Un peu de calcul :

$$(l_1 \circ \iota \circ l_2)(z) = l_1\left(\frac{1}{cz+d}\right) = \frac{\alpha + \beta cz + \beta d}{cz+d}, \tag{23.164}$$

et en imposant que cela soit égal à $\frac{az+b}{cz+d}$ nous trouvons $\beta = a/c$ et $\alpha = b - ad/c$. Il est vite vérifié que ces choix donnent le bon résultat. □

NORMooMMKOooQlzjqJ

23.77.
Vu que les applications linéaires sont des composées d'une translation et d'une rotation-homothétie, et que l'application i est une composée d'une inversion $z \mapsto 1/\bar{z}$ et d'une réflexion $z \mapsto \bar{z}$, toutes les homographies sont des composées des éléments suivants :

— inversion[10] $z \mapsto 1/\bar{z}$, prolongée par $i(\infty) = 0$ et $i(0) = \infty$;

— réflexion $z \mapsto \bar{z}$;

— translation $z \mapsto z + \alpha$ avec $\alpha \in \mathbb{C}$;

— rotation-homothétie $z \mapsto \lambda z$ avec $\lambda \in \mathbb{C}$.

Toutes ces opérations sont prolongées à $\hat{\mathbb{C}}$ par $1/\infty = 0$, $\lambda \cdot \infty = \omega$ (si $\lambda \neq 0$) et $\infty + \lambda = \infty$. Nous ne définissons pas $0 \cdot \infty$ et $\infty - \infty$.

Certes nous pouvons construire des homographies à partir d'ingrédients dont la conjugaison complexe. Il ne faudrait cependant pas déduire que cette conjugaison est une homographie.

LEMooGDDJooBpJlUf

Lemme 23.78.
La conjugaison complexe n'est pas une homographie.

10. Oui, c'est l'inversion de la géométrie hyperbolique, voir 23.6.1.6.

Démonstration. Si elle l'était nous aurions des nombres $a, b, c, d \in \mathbb{C}$ tels que $ad - bc \neq 0$ et

$$\frac{az + b}{cz + d} = \bar{z} \tag{23.165}$$

EQooJMAZooVcJIAP

pour tout $z \in \mathbb{C}$.

En posant $z = 0$ nous avons déjà $b/d = 0$, c'est-à-dire $b = 0$. Avec $z = 1$ nous trouvons alors $a/(c + d) = 1$, c'est-à-dire

$$a = c + d. \tag{23.166}$$

(1) **Si** $ci + d \neq 0$

Dans ce cas nous pouvons évaluer (23.165) en $z = i$ et avoir $a = -ci + d$. Mais comme nous avions déjà $a = c + d$ nous déduisons $c = 0$. Nous restons donc avec

$$\frac{a}{d}z = \bar{z} \tag{23.167}$$

pour tout z. En prenant $z = 1$ puis $z = i$, il est vite remarqué que cela n'est pas possible.

(2) **Si** $ci + d = 0$

Nous rappelons que $ad \neq 0$. Nous écrivons l'équation avec $z = -i$ pour trouver

$$\frac{-ai}{-ci + d} = i, \tag{23.168}$$

qui donne immédiatement $a = ci - d$. Nous avons donc les trois équations

$$\begin{cases} c = id & \text{(23.169a)} \\ a = ci - d & \text{(23.169b)} \\ a = c + d. & \text{(23.169c)} \end{cases}$$

Une tentative de résolution tombe rapidement sur une impossibilité (en substituant la première dans les deux autres et en comparant les deux valeurs de a par exemple).

□

La proposition suivante ressemble à s'y méprendre à la proposition 22.11, mais elle diffère en deux points. D'abord elle ne traite que de l'inversion par rapport à l'origine, mais surtout, elle traite le point $z = \infty$. C'est un avantage de travailler sur $\hat{\mathbb{C}}$ plutôt que sur $\mathbb{R}^2$.

PROPooEAKXooUIqWEv

Proposition 23.79 (Inversion de cercles et de droites)**.**
L'inversion dans $\hat{\mathbb{C}}$ envoie

(1) une droite passant par 0 sur elle-même

(2) une droite ne passant pas par 0 sur un cercle passant par 0.

(3) un cercle ne passant pas par 0 en un cercle ne passant pas par 0.

(4) un cercle passant par 0 en une droite ne passant pas par 0.

Démonstration. Décomposition en tous les cas possibles.

(1) **Droite passant par** 0

La façon la plus simple de traiter la droite passant par 0 est de l'écrire sous forme paramétrique :

$$z(t) = te^{i\theta} \tag{23.170}$$

pour θ fixé et $t \in \mathbb{R} \cup \{\infty\}$. En appliquant l'inversion :

$$\iota\big(z(t)\big) = \frac{1}{(te^{i\theta})^*} = \frac{1}{t}e^{i\theta}. \tag{23.171}$$

Notons que les cas particuliers fonctionnent : pour $t = 0$ nous avons le point ∞ et pour $t = \infty$ nous avons 0.

(2) **Droite ne passant pas par 0**
Une droite ne passant par par $z = 0$ est un ensemble de la forme

$$d(\omega, k) = \{z \in \mathbb{C}^* \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \tag{23.172}$$

avec $k \neq 0$. Étant donné que ι est une bijection et même une involution nous avons $z \in \iota\big(d(\omega, k)\big)$ si et seulement si $\tau(z) \in d(\omega, k)$. L'équation est donc, pour $z \neq 0$:

$$\frac{\bar{\omega}}{z} + \frac{\omega}{\bar{z}} = k. \tag{23.173}$$

Et comme $z \neq 0$ nous pouvons multiplier par $z\bar{z}$ pour trouver $\bar{z}\bar{\omega} + z\omega = kz\bar{z}$. Donc

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C}^* \text{ tel que } \bar{z}\bar{\omega} + z\omega = kz\bar{z}\} \cup \{0\}. \tag{23.174}$$

Dans l'ensemble, nous pouvons renommer z et $\bar{z}$ pour avoir une forme plus symétrique. De plus il se fait que $z = 0$ vérifie l'équation donnée ; nous pouvons donc lever la condition $z \in \mathbb{C}^*$ et ne plus ajouter $\{0\}$ à côté :

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = kz\bar{z}\}. \tag{23.175}$$

Cela est l'équation d'un cercle passant par l'origine (définition 23.63).

(3) **Cercle ne passant pas par 0**
Nous considérons le cercle $C(\omega, r)$ avec $r^2 \neq |\omega|^2$. Il ne contient ni ∞ ni 0 et nous avons alors

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C}^* \text{ tel que } \frac{1}{z\bar{z}} - \bar{\omega}\frac{1}{\bar{z}} - \omega\frac{1}{z} = r^2 - |\omega|^2\}. \tag{23.176}$$

Vu que z n'est jamais nul nous pouvons multiplier l'équation par $z\bar{z}$:

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C} \text{ tel que } \big(|\omega|^2 - r^2\big)z\bar{z} - \bar{\omega}z - \omega\bar{z} = -1\}. \tag{23.177}$$

Le coefficient $|\omega|^2 - r^2$ est non nul par hypothèse et cet ensemble est un cercle par le lemme 23.65. Il ne passe manifestement pas par $z = 0$.

(4) **Cercle passant par 0**
Le cercle passe par 0, et donc son image par ∞. Nous écrivons alors

$$\iota\big(C(\omega, |\omega|)\big) = \iota\big(C(\omega, |\omega|)\backslash\{0\}\big) \cup \{\infty\}. \tag{23.178}$$

Nous avons

$$C(\omega, |\omega|)\backslash\{0\} = \{z \in \mathbb{C}^* \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{23.179}$$

et un calcul usuel donne

$$\iota\big(C(\omega, |\omega|)\backslash\{0\}\big) = \{z \in \mathbb{C}^* \text{ tel que } 1 - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{23.180}$$

et donc

$$\iota\big(C(\omega, |\omega|)\big) = d(\omega, 1), \tag{23.181}$$

en nous souvenant que le point ∞ est contenu dans $d(\omega, 1)$.

□

PROPooYFJBooAWxFIs

Proposition 23.80.
Une homographie conserve l'ensemble des droites et cercles de $\hat{\mathbb{C}}$.

Attention : cela ne veut pas dire qu'une homographie transforme une droite en une droite et un cercle en un cercle. Ça veut dire qu'une homographie transforme une droite en une droite ou un cercle et un cercle en une droite ou un cercle.

Démonstration. Nous savons par la proposition 22.11 et 23.77 que les homographies se décomposent en inversion, réflexion, translation et rotation-homothétie.

À part pour l'inversion, tout est clair comment ça fonctionne hein. En ce qui concerne l'inversion, nous avons la proposition 23.79 qui donne déjà toutes les réponses.

□

23.81.
Les homographies préservent les angles, c'est l'objet du théorème suivant. Il ne faudrait cependant pas croire que si A, B et C sont trois points distincts, l'angle entre $\overline{AC}$ et $\overline{BC}$ est le même que celui entre $\overline{f(A)f(C)}$ et $\overline{f(B)f(C)}$ dès que f est une homographie. Cela serait préserver les angles globalement, c'est-à-dire préserver les angles lorsque les points sont déplacés par f.

Nous allons regarder les angles locaux, c'est-à-dire lorsque les points sont déplacés par df.

Définition 23.82.
Nous disons qu'une application $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ préserve localement les angles non orientés lorsque

$$\cos\big(df_a(u), df_a(v)\big) = \cos(u, v) \tag{23.182}$$

pour tout $a \in \mathbb{R}^2$ et $u, v \in \mathbb{R}^2$. Ici il est mieux de penser à $u, v \in T_a\mathbb{R}^2$ pour qui sait les espaces tangents en géométrie différentielle.

Voir la définition de l'angle 18.52.

Théorème 23.83.
Les homographies de $P(\mathbb{C}^2)$ préservent localement les angles non orientés.

Démonstration. En ce qui concerne les translations, dilatations et rotations, les choses sont claires. Vérifions pour l'inversion, qu'il faut interpréter dans $\mathbb{R}^2$. Pour $z = x + iy$ nous avons

$$\iota(z) = \frac{1}{z} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}. \tag{23.183}$$

Nous devons donc étudier la fonction

$$\begin{aligned} F\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto \Big(\frac{x}{r^2}, -\frac{y}{r^2}\Big) \end{aligned} \tag{23.184}$$

où nous avons posé $r^2 = x^2 + y^2$ pour simplifier les notations.

Soient deux vecteurs $u, v \in \mathbb{R}^2$ et un point $a \in \mathbb{R}^2$. Nous devons prouver que

$$\frac{u \cdot v}{\|u\|\|v\|} = \frac{dF_a(u) \cdot dF_a(v)}{\|dF_a(u)\|\|dF_a(v)\|}. \tag{23.185}$$ EQooRXSJooFnLVLC

Pour cela, nous pourrions calculer dF_a et passer en coordonnées polaires[613] mais nous préférons faire les calculs à la dure parce que nous avons Sage avec nous.

Nous notons A la matrice de dF en $a = (x, y)$ et nous avons

$$Au \cdot Av = A^tAu \cdot v \tag{23.186}$$

ainsi que $\|u\| = \sqrt{u \cdot u} = \sqrt{A^tAu \cdot u}$, de telle sorte qu'il devienne urgent de calculer A^tA. Voici le calcul :

```
var('y')

# les fonctions coordonnées
F1(x,y)=x/(x**2+y**2)
F2(x,y)=-y/(x**2+y**2)
```

```
6
7  # La matrice différentielle :
8  A=matrix( [  [F1.diff(x).simplify_full(),F1.diff(y).simplify_full()],[F2.diff(x).simplify_full(),F2.diff(y).simplify_full()] ] )
9
10 # Quelqu'un peut expliquer pourquoi ceci ne fonctionne pas ?
11 # A=matrix( [  [F1.diff(x),F1.diff(y)],[F2.diff(x),F2.diff(y)]  ]).simplify_full()
12
13 # A^tA
14 S=A.transpose()*A
15 S=S.simplify_full()      # Mais ça, ça marche !!
16 print(S)
```

tex/sage/sageSnip009.sage

Le résultat est que

$$A^tA = \begin{pmatrix} \frac{1}{r^4} & 0 \\ 0 & \frac{1}{r^4} \end{pmatrix} = \frac{1}{r^4}\mathrm{Id}\,. \tag{23.187}$$

La vérification de (23.185) est alors immédiate. □

23.6.3 Birapport

Nous introduisons maintenant quelque chose qui s'appelle le « birapport » et qui n'est à priori pas du tout lié au birapport défini en 23.42.

DEFooQYHVooMZwQMB

Définition 23.84 (Birapport dans $\hat{\mathbb{C}}$[545])**.**
*Soient $a, b, c, x \in \hat{\mathbb{C}}$ où a, b et c sont distincts. Le **birapport** de ces quatre nombres est l'élément de $\hat{\mathbb{C}}$ donné par, si $a, b, c \neq \infty$:*

$$[a, b, c, x] = \frac{(a-c)(b-x)}{(b-c)(a-x)}, \tag{23.188}$$

EQooQJWZooQXKslh

et

$$[\infty, b, c, x] = \frac{b-x}{b-c} \tag{23.189a}$$

SUBEQooNSQNooYUhuzB

$$[a, \infty, c, x] = \frac{a-c}{a-x} \tag{23.189b}$$

$$[a, b, \infty, x] = \frac{b-x}{a-x} \tag{23.189c}$$

23.85.
Notons la « logique » des cas particuliers. Pour le premier [11], si $a \to \infty$ tandis que les autres restent dans $\mathbb{C}$ alors $a - c$ et $a - x$ deviennent du même ordre de grandeur et se simplifient. Il reste les deux autres parties de la fraction.

C'est cette même logique qui, partant de $[a, b, \infty, x] = \frac{b-x}{a-x}$ donne

$$[a, b, \infty, \infty] = 1 \tag{23.190}$$

comme il se doit si nous avons l'intention de ressembler au lemme 23.47.

L'objet « birapport » introduit ici est évidemment lié au birapport sur $P(\mathbb{C}^2)$ défini plus haut. Le lien est la proposition suivante.

11. Pour avoir une notion de topologie sur $\hat{\mathbb{C}}$ vous pouvez aller voir la définition 48.57 et ce qui l'entoure. En particulier la compactification en un point 48.55.

PROPooLKQQooEOrjwC

Proposition 23.86.
Soit l'application $\varphi_0 : P(\mathbb{C}^2) \to \hat{\mathbb{C}}$ définie en 23.36. Si $A, B, C, X \in P(\mathbb{C}^2)$ alors

$$[A, B, C, X]_{\varphi_0} = [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(X)]. \tag{23.191}$$

EQooEOZZooMRHJfH

Cela est une égalité dans $\hat{\mathbb{C}}$.

Démonstration. Nous écrivons $A = [a_1, a_2]$, $B = [b_1, b_2]$, $C = [c_1, c_2]$ avec $a_1, a_2, b_2, b_2, c_1, c_2 \in \mathbb{C}$. Par définition $[A, B, C, X]_{\varphi_0} = (\varphi_0 \circ \phi)(X)$ où $\phi : P(\mathbb{C}^2) \to P(\mathbb{C}^2)$ est l'unique homographie telle que

SUBEQSooTFWTooRnijrY

$$\phi[a_1, a_2] = [1, 0] \tag{23.192a}$$
$$\phi[b_1, b_2] = [0, 1] \tag{23.192b}$$
$$\phi[c_1, c_2] = [1, 1] \tag{23.192c}$$

Une des difficultés de cette preuve va être de calculer ce ϕ. D'abord nous pouvons introduire $\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi_0^{-1}$ qui est obligatoirement (proposition 23.72) de la forme

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma z + \delta}. \tag{23.193}$$

Nous allons imposer les relations (23.192) pour déterminer les coefficients α, β, γ et δ.

D'abord

$$(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)\big([a_1, a_2]\big) = (\varphi_0^{-1} \circ \tilde{\phi})(a_1/a_2) \tag{23.194a}$$
$$= \varphi_0^{-1}\left(\frac{\alpha\frac{a_1}{a_2} + \beta}{\gamma\frac{a_1}{a_2} + \delta}\right) \tag{23.194b}$$
$$= \big[\alpha\frac{a_1}{a_2} + \beta, \gamma\frac{a_1}{a_2} + \delta\big]. \tag{23.194c}$$

Égaler cela à $[1, 0]$ donne

$$\begin{cases} \alpha\frac{a_1}{a_2} + \beta \neq 0 & \text{(23.195a)} \\ \gamma\frac{a_1}{a_2} + \delta = 0. & \text{(23.195b)} \end{cases}$$

Donc nous avons déjà

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma(z - \frac{a_1}{a_2})} = \frac{\alpha z + \beta}{\gamma\big(z - \varphi_0(A)\big)}. \tag{23.196}$$

En y imposant la contrainte $(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)([b_1, b_2]) = [0, 1]$ nous trouvons les contraintes

$$\begin{cases} \gamma\big(\varphi_0(B) - \varphi_0(A)\big) \neq 0 & \text{(23.197a)} \\ \beta = -\alpha\varphi_0(B). & \text{(23.197b)} \end{cases}$$

Nous avons décidé d'écrire $\varphi_0(A)$ au lieu de a_1/a_2 à la fois pour un soucis de simplification d'écriture et dans le but de ressembler à (23.191). En substituant :

$$\tilde{\phi}(z) = \frac{\alpha\big(z - \varphi_0(B)\big)}{\gamma\big(z - \varphi_0(A)\big)}. \tag{23.198}$$

La condition pour $[c_1, c_2]$ donne

$$\big[\alpha\big(\varphi_0(C) - \varphi_0(B)\big), \gamma\big(\varphi_0(C) - \varphi_0(A)\big)\big] = [1, 1], \tag{23.199}$$

ce qui donne

$$\alpha\big(\varphi_0(C) - \varphi_0(B)\big) = \gamma\big(\varphi_0(C) - \varphi_0(A)\big). \tag{23.200}$$

Nous avons alors

$$\gamma = \alpha \frac{\varphi_0(C) - \varphi_0(B)}{\varphi_0(C) - \varphi_0(A)}, \tag{23.201}$$

et les α se simplifient dans la formule pour $\tilde{\phi}$:

$$\tilde{\phi}(z) = \frac{\big(z - \varphi_0(B)\big)\big(\varphi_0(C) - \varphi_0(A)\big)}{\big(z - \varphi_0(A)\big)\big(\varphi_0(C) - \varphi_0(B)\big)}. \tag{23.202}$$

Par la proposition 23.72, l'application $\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$ est une homographie. Nous pouvons donc calmement calculer le birapport $[A, B, C, X]_{\varphi_0}$ de la façon suivante :

$$\begin{aligned}
[A, B, C, X]_{\varphi_0} &= (\varphi_0 \circ \phi)(X) && \text{(23.203a)}\\
&= (\varphi_0 \circ \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)(X) && \text{(23.203b)}\\
&= \tilde{\phi}\big(\varphi_0(X)\big) && \text{(23.203c)}\\
&= \frac{\big(\varphi_0(X) - \varphi_0(B)\big)\big(\varphi_0(C) - \varphi_0(A)\big)}{\big(\varphi_0(X) - \varphi_0(A)\big)\big(\varphi_0(C) - \varphi_0(B)\big)} && \text{(23.203d)}\\
&= [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(A)]. && \text{(23.203e)}
\end{aligned}$$

□

PROPooQGPFooReNaGq

Proposition 23.87.
Les homographies de $\hat{\mathbb{C}}$ conservent le birapport.

Démonstration. Ici le mot « homographie » réfère à la définition 23.74 et le birapport à 23.84. Soient $a, b, c, x \in \hat{\mathbb{C}}$ et une homographie $\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$. Il existe une homographie $\phi\colon P(\mathbb{C}^2) \to P(\mathbb{C}^2)$ telle que $\tilde{\phi} = \varphi_0 \circ \phi \varphi_0^{-1}$. Alors

$$\begin{aligned}
\big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)\big] &= \big[(\phi \circ \varphi_0^{-1})(a), (\phi \circ \varphi_0^{-1})(b), (\phi \circ \varphi_0^{-1})(c), (\phi \circ \varphi_0^{-1})(x), \big]_{\varphi_0} && \text{(23.204a)}\\
&= \big[\varphi_0^{-1}(a), \varphi_0^{-1}(b), \varphi_0^{-1}(c), \varphi_0^{-1}(x), \big]_{\varphi_0} && \text{(23.204b)}\\
&= [a, b, c, x]. && \text{(23.204c)}
\end{aligned}$$

Justifications :

— Identification des birapports sur $P(\mathbb{C}^2)$ et sur $\hat{\mathbb{C}}$, proposition 23.86.

— Invariance du birapport sour les homographies (dans $P(\mathbb{C}^2)$), proposition 23.49.

□

PROPooSGCJooLnOLCx

Proposition 23.88.
Soient des points a, b, c, x dans $\hat{\mathbb{C}}$ avec a, b, c distincts. Ils sont alignés ou cocycliques si et seulement si $[a, b, c, x] \in \hat{\mathbb{R}}$.

Démonstration. Nous allons faire plusieurs cas. Mais dans tous les cas vous pouvez relire la définition des angles orientés 18.147 et la partie sur les angles dans les nombres complexes 18.10.12.

(1) **Tous les points sont distincts et dans $\mathbb{C}$**

D'une part, nous savons que le nombre complexe $re^{i\theta}$ est réel si et seulement si $\theta \in [0]_\pi$, et d'autre part l'argument du birapport (23.188) est

$$[\overrightarrow{ca}, \overrightarrow{cb}] + [\overrightarrow{xb}, \overrightarrow{xa}]. \tag{23.205}$$

Le birapport est réel si et seulement si

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \in [0]_\pi. \tag{23.206}$$

À gauche nous avons une classe modulo 2π et à droite une classe modulo π. L'égalité signifie qu'il y a un représentant du membre de gauche qui appartient au membre de droite. Si vous

aimez faire très attention à ce que signifient les notations, voici trois manières d'écrire la condition, par ordre croissant de précision :

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] = [0]_\pi, \tag{23.207a}$$

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \in [0]_\pi, \tag{23.207b}$$

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \subset [0]_\pi. \tag{23.207c}$$

SUBEQooEAGIooIKdPQV

Nous avons donc que le birapport est réel si et seulement si la condition (23.207c) est vérifiée. D'après le théorème 18.155, cette dernière condition est équivalente à dire que les points a, b, c et x sont alignés.

(2) **Pas quatre points distincts, dans $\mathbb{C}$**

Nous supposons encore que a, b, c et x sont dans $\mathbb{C}$. Mais nous supposons que x est un de a, b ou c. Vu que par hypothèse a, b et c sont distincts, c'est le seul cas à considérer dans la catégorie des 4 points non distincts.

Trois points sont toujours alignés ou cocycliques[12]. Donc nous devons seulement montrer que dans ce cas le birapport est toujours dans $\hat{\mathbb{R}}$. Par définition,

— Si $x = a$ alors $[a, b, c, x] = \infty$,
— Si $x = b$ alors $[a, b, c, x] = 0$,
— Si $a = c$ alors $[a, b, c, x] = 1$.

Dans tous les cas de figure le birapport est dans $\hat{\mathbb{R}}$.

À ce niveau de la preuve nous devons encore vérifier les cas où a, b, c ou x valent ∞. Si l'un de a, b ou c est ∞ et si x l'est aussi, alors, comme ∞ est aligné avec tout, nous avons seulement une droite passant par deux points. Il nous faut donc seulement regarder les cas où un seul des 4 points est ∞.

(1) **Si $a = \infty$** Le birapport est alors (par (23.189a))

$$[\infty, b, c, x] = \frac{b - x}{a - x}, \tag{23.208}$$

qui est un nombre à priori complexe dont le dénominateur est supposé non nul parce que le cas $a = x$ est déjà traité. L'argument de ce nombre est dans la classe de l'angle orienté

$$\arg\left(\frac{b - x}{a - x}\right) \in [\overrightarrow{xa}, \overrightarrow{xb}]. \tag{23.209}$$

Le birapport est réel si et seulement si le membre de gauche est dans $[0]_\pi$. Et cela est justement le cas où le membre de droite donne des points alignés.

Les cas $b = \infty$, $c = \infty$ et $x = \infty$ se traitent de la même manière. □

23.6.4 Division harmonique

Définition 23.89.
*Nous disons que les éléments a, b, c et x de $\hat{\mathbb{C}}$ sont en **division harmonique** lorsque $[a, b, c, x] = -1$.*

NORMooUWYDooAZTTWu

23.90.
Une chose qui sera utile par la suite est de remarquer que $[a, b, c, \infty] = -1$ lorsque $c = \frac{a+b}{2}$.

Nous allons maintenant voir comment, pour $a, b, c \in \mathbb{C}$ donnés nous pouvons construire x tels que a, b, c, x soient en division harmonique. Vu que trois points sont soit cocycliques soit alignés nous divisons la construction en deux parties.

Notons que si nous trouvons une construction qui donne un point x vérifiant $[a, b, c, x] = -1$ alors nous prouvons au passage que la construction ne dépend pas des choix intermédiaires parce que il n'existe qu'un unique x tel que $[a, b, c, x] = -1$ lorsque a, b, c sont donnés.

12. Si ils ne sont pas alignés, prendre la médiatrice du segment $[a, b]$ et celle de $[b, c]$, et l'intersection vous donnera le centre d'un cercle passant par a, b et c.

LEMooAEDTooKsUoPw

Lemme 23.91 ([617, 618])**.**
Soient a, b, c cocycliques dans $\mathbb{C}$. Nous nommons $\mathcal{C}$ le cercle contenant a, b et c ainsi que T_a et T_b les tangentes à $\mathcal{C}$ en a et b. Soit $m = T_a \cap T_b$ et la droite $L = (mc)$. Alors le point

$$x = (mc) \cap \mathcal{C} \tag{23.210}$$

vérifie $[a, b, c, x] = -1$.

Si $m = \infty$ (arrive lorsque $T_a \parallel T_b$) alors en guise de L nous prenons la parallèle à T_a passant par c.

Démonstration. Nous séparons les cas suivant que $m = \infty$ ou non.

(1) $\underline{m = \infty}$

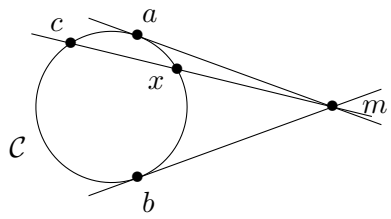
Les tangentes à $\mathcal{C}$ en a et en b sont parallèles, c'est-à-dire que ces points sont diamétralement opposés sur $\mathcal{C}$. Les homographies préservent le birapport (proposition 23.87), et les rotations, dilatations et translations sont des homographies (voir 23.77).

Nous pouvons donc nous ramener au cas où $\mathcal{C}$ est centré en 0 et de rayon 1 avec $a = i$ et $b = -i$. Dans ce cas, $c = e^{i\theta}$. Vu que x est donné par l'intersection entre le cercle et la droite horizontale passant par c nous avons $x = e^{i(\pi-\theta)}$. Le birapport se calcule explicitement :

$$[a, b, c, x] = \frac{(i - e^{i\theta})(-i - e^{i(\pi-\theta)})}{(-i - e^{i\theta})(i - e^{i(\pi-\theta)})} = -1. \tag{23.211}$$

(2) $\underline{m \neq \infty}$

Nous sommes dans la situation suivante où à une translation près nous supposons $x = 0$:



Nous allons prouver que dans ce cas, $[a, b, c, x] = -1$. Pour cela nous considérons l'inversion de centre x (qui est $x = 0$ par translation). Soit $\tilde{\phi}$ cette homographie. Elle conserve le birapport, il nous allons voir que calculer $[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)]$ se révèle être plus facile[13].

Nous nommons $A = (am)$, $B = (bm)$, $C = (cm)$ et $\mathcal{C}$, le cercle. Nous allons maintenant faire intensément usage de la proposition 23.79. Nous avons :

— $\tilde{\phi}(A)$ est un cercle passant par 0.
— $\tilde{\phi}(B)$ est un cercle passant par 0.
— $\tilde{\phi}(C)$ est la droite C.
— $\tilde{\phi}(\mathcal{C})$ est une droite ne passant pas par 0.

Les droites A et B se coupent en m et en ∞ (qui sont des points distincts). Donc les cercles $\tilde{\phi}(A)$ et $\tilde{\phi}(B)$ se coupent en 0 et $\tilde{\phi}(m)$, aucun de ces deux points n'est sur la droite $\tilde{\phi}(\mathcal{C})$.

Par tangence, la droite A et le cercle $\mathcal{C}$ se coupent en un seul point (a). Donc $\tilde{\phi}(A)$ coupe $\tilde{\phi}(\mathcal{C})$ en un seul point, $\tilde{\phi}(a)$. Idem pour le cercle $\tilde{\phi}(B)$.

Nous avons donc que les cercles $\tilde{\phi}(A)$ et $\tilde{\phi}(B)$ sont tangents à la droite $\tilde{\phi}(\mathcal{C})$ et se coupent en exactement deux points distincts (qui sont donc du même côté de la droite).

Nous nous intéressons à la droite $\tilde{\phi}(C)$. C'est une droite parce que c'est l'image d'une droite passant par 0. Elle passe par 0, par $\tilde{\phi}(c)$ et $\tilde{\phi}(m)$. Le fait qu'elle passe par 0 et $\tilde{\phi}(m)$ fait

13. Si vous n'avez peur d'aucun calculs, il suffit de poser $a = e^{i\theta}$, $b = e^{-i\theta}$ et $c = e^{i\sigma}$ et vous êtes théoriquement capable de calculer les coordonnées de tous les points, y compris de x en termes de θ et σ. Ensuite le calcul du birapport est explicite.

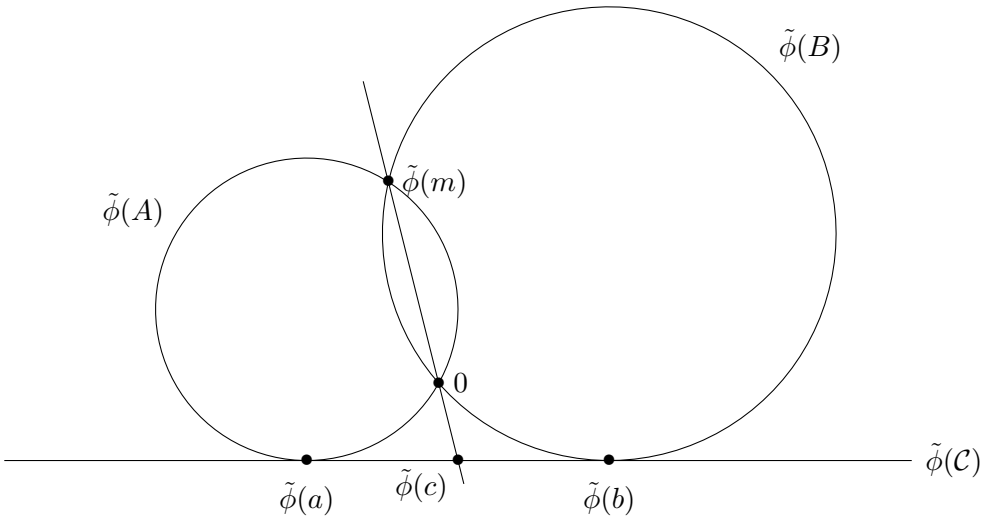
que c'est la droite passant par les deux intersections des cercles. Vu que $c \in \mathcal{C} \cap C$, le point d'intersection $\tilde{\phi}(C) \cap \tilde{\phi}(\mathcal{C})$ est $\tilde{\phi}(c)$.

Quelle est la puissance du point $\tilde{\phi}(c)$ par rapport au cercle $\tilde{\phi}(A)$? En la calculant avec la droite $\tilde{\phi}(C)$, qui intersecte les deux cercles aux points déjà étudiés, la puissance est :

$$k = d\big(\tilde{\phi}(c), 0\big) \times d\big(\tilde{\phi}(c), \tilde{\phi}(m)\big). \tag{23.212}$$

Vu que ces points sont également sur le cercle $\tilde{\phi}(B)$, la puissance de $\tilde{\phi}(c)$ par rapport à ce second point est la même.

Tout cela justifie le dessin suivant [14] :



Mais la droite passant par $\tilde{\phi}(a)$ et $\tilde{\phi}(c)$ (qui est tangente au cercle) permet également de calculer cette puissance :

$$k = d\big(\tilde{\phi}(a), \tilde{\phi}(c)\big). \tag{23.213}$$

Idem pour la puissance par rapport à l'autre cercle :

$$k = d\big(\tilde{\phi}(b), \tilde{\phi}(c)\big). \tag{23.214}$$

Nous en déduisons que $\tilde{\phi}(c)$ est le milieu entre $\tilde{\phi}(a)$ et $\tilde{\phi}(b)$.

Du coup

$$\big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)\big] = \big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \infty\big] = -1 \tag{23.215}$$

en vertu de ce que nous avons raconté en 23.90.

□

LEMooYBTHooABWkeo

Lemme 23.92 ([618]).
Soit $a, b, c \in \mathbb{C}$ colinéaires. Soit m un point hors de cette droite. Nous considérons une droite issue de c coupant $[ma]$ en p et $[mb]$ en q.

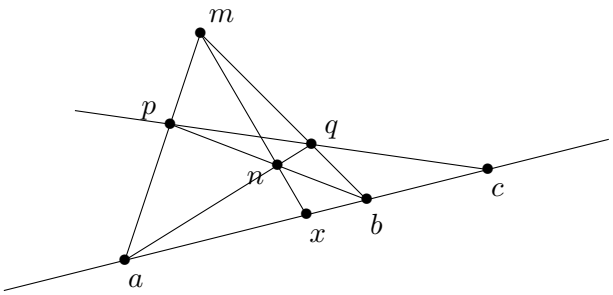
Nous construisons $n = (aq) \cap (pb)$ et finalement $x = (mn) \cap (ab)$.

À la fin nous avons

$$[a, b, c, x] = -1. \tag{23.216}$$

Démonstration. Commençons par un dessin de la situation :

14. Bien que ce ne soit pas strictement nécessaire à la preuve, est-ce que vous savez si les deux cercles ont le même rayon ? Et si par hasard la droite $\big(\tilde{\phi}(m)\tilde{\phi}(c)\big)$ n'arrive pas perpendiculairement à $\tilde{\phi}(\mathcal{C})$?



Les points a, b et m ne sont pas alignés, et nous pouvons les utiliser comme repère barycentrique (voir 11.51 pour savoir en deux mots ce que c'est). Nous nommons (α,β,γ) les coordonnées de n dans ce système, c'est-à-dire que

$$\alpha\overrightarrow{na}+\beta\overrightarrow{nb}+\gamma\overrightarrow{nm}=0. \tag{23.217}$$

Dans notre contexte, nous pouvons voir le vecteur $\overrightarrow{st}$ comme une façon d'écrire le nombre $t-s$. Par la proposition 11.58 nous savons les coordonnées barycentriques de p, x et q en regardant le triangle acb. Voici les coordonnées et les relations qu'elles signifient : SUBEQSooKKIXooZbWyHe

$$\begin{aligned}
n&=(\alpha,\beta,\gamma), & \alpha\overrightarrow{na}+\beta\overrightarrow{nb}+\gamma\overrightarrow{nm}&=0 && (23.218a)\\
p&=(\alpha,0,\gamma), & \alpha\overrightarrow{pa}+\gamma\overrightarrow{pm}&=0 && (23.218b)\\
q&=(0,\beta,\gamma), & \beta\overrightarrow{qb}+\gamma\overrightarrow{qm}&=0 && (23.218c)\\
x&=(\alpha,\beta,0). & \alpha\overrightarrow{xa}+\beta\overrightarrow{xb}&=0. && (23.218d)
\end{aligned}$$ SUBEQooRUUDooDGsMoq

Nous voudrions maintenant voir les coordonnées de c. Nous posons $c=(\lambda,\mu,\sigma)$:

$$\lambda\overrightarrow{ca}+\mu\overrightarrow{cb}+\sigma\overrightarrow{cm}=0. \tag{23.219}$$ EQooZKBSooZQzqrV

Mais a, b et c sont alignés, donc $\overrightarrow{ca}$ et $\overrightarrow{cb}$ sont colinéaires, alors que $\overrightarrow{cm}$ n'est pas aligné avec les deux autres. L'annulation (23.219) demande donc l'annulation séparément

$$\begin{cases}
\sigma\overrightarrow{cm}=0 & (23.220a)\\
\lambda\overrightarrow{ca}+\mu\overrightarrow{cb}=0. & (23.220b)
\end{cases}$$ SUBEQooOTAOooQLndsd

Nous en déduisons que $\sigma=0$ et aussi que λ et μ ne sont pas nuls. Nous posons arbitrairement $\lambda=1$ parce que les coordonnées barycentriques sont définies à coefficient multiplicatif près.

Nous imposons à présent le fait que p, q et c sont alignés. Pour cela nous devons faire apparaitre les vecteurs $\overrightarrow{pq}$, $\overrightarrow{pc}$, $\overrightarrow{pc}$. Vu le dessin et les relations disponibles (23.218) le mieux est d'utiliser les relations de Chasles (proposition 11.5) pour faire $\overrightarrow{ca}=\overrightarrow{cp}+\overrightarrow{pa}$ et $\overrightarrow{cb}=\overrightarrow{cq}+\overrightarrow{qb}$. La relation (23.220b) devient :

$$\overrightarrow{cp}+\overrightarrow{pa}+\mu(\overrightarrow{cq}+\overrightarrow{qb})=0. \tag{23.221}$$

Les vecteurs $\overrightarrow{cp}$ et $\overrightarrow{cq}$ sont alignés, donc nous les écrivons ensemble. Les vecteurs $\overrightarrow{pa}$ et $\overrightarrow{qb}$ se transforment en utilisant les relations (23.218) :

$$\overrightarrow{cp}+\mu\overrightarrow{cq}-\frac{\gamma}{\alpha}\overrightarrow{pm}-\mu\frac{\gamma}{\beta}\overrightarrow{qm}. \tag{23.222}$$

Enfin nous voulons faire la somme du terme $\overrightarrow{pm}$ avec le terme $\overrightarrow{qm}$. D'abord on change le signe :

$$\overrightarrow{cp}+\mu\overrightarrow{cq}-\frac{\gamma}{\alpha}\overrightarrow{pm}+\frac{\mu\gamma}{\beta}\overrightarrow{mq} \tag{23.223}$$

ensuite nous écrivons

$$\frac{\mu\gamma}{\beta}=\frac{\mu\gamma}{\beta}+\frac{\gamma}{\alpha}-\frac{\gamma}{\alpha}, \tag{23.224}$$

et

$$\overrightarrow{cp}+\mu\overrightarrow{cq}-\frac{\gamma}{\alpha}(\overrightarrow{pm}+\overrightarrow{mq})+\left(\frac{\mu\gamma}{\beta}+\frac{\gamma}{\alpha}\right)\overrightarrow{mq}=0. \tag{23.225}$$

Tout cela pour

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pq} + \left(\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha}\right)\overrightarrow{mq} = 0. \tag{23.226}$$

Le dernier terme n'est pas colinéaire aux deux premiers et s'annule donc séparément :

$$\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} = 0. \tag{23.227}$$

Cela donne $\mu = -\beta/\alpha$.

Au final nous avons

$$\overrightarrow{ca} - \frac{\beta}{\alpha}\overrightarrow{cb} = 0 \tag{23.228}$$

et donc

$$\alpha\overrightarrow{ca} - \beta\overrightarrow{cb} = 0, \tag{23.229}$$

EQooWKBWooEWhDiM

ce qui donne les coordonnées $(\alpha, -\beta, 0)$ pour le point c.

Vu que nous somme dans l'espace vectoriel $\mathbb{C}$, ce que nous notons $\overrightarrow{AB}$ n'est rien d'autre que la différence $B - A$ dans $\mathbb{C}$ [15]. La relation (23.229) signifie donc

$$\alpha(a - c) = \beta(b - c). \tag{23.230}$$

Nous avons alors :

$$\frac{a - c}{b - c} = \frac{\beta}{\alpha}. \tag{23.231}$$

Par ailleurs, la relation (23.218d) à propos des coordonnées de x donne

$$\frac{a - x}{b - x} = -\frac{\beta}{\alpha}. \tag{23.232}$$

En égalisant les deux valeurs de β/α nous trouvons :

$$\frac{a - c}{b - c} = \frac{x - a}{b - x}, \tag{23.233}$$

ce qui donne (via un petit jeu de signes)

$$\frac{(c - a)(x - b)}{(c - b)(x - a)} = -1. \tag{23.234}$$

C'est cela que nous voulions. □

23.6.5 Groupe circulaire

Nous avons vu que les homographies préservent l'ensemble des cercles et droites. Nous pouvons nous demander quel est le groupe maximum préservant l'ensemble des cercles et droites.

Définition 23.93.
*Le **groupe circulaire** de $\mathbb{C}$ est le groupe de transformations de $\hat{\mathbb{C}}$ engendré*[16] *par les homographies*[17] *et la conjugaison complexe. Le groupe circulaire de l'espace projectif est l'ensemble des applications de la forme $\varphi_0^{-1} \circ f \circ \varphi_0$ où f est un élément du groupe circulaire de $\mathbb{C}$.*

Vu le lemme 23.78, la conjugaison complexe n'est pas une homographie. Donc cette définition n'est pas stupide : le groupe circulaire est strictement plus grand que le groupe des homographies.

15. Les mauvaise langues diront que tout le chapitre sur les espaces affines, et surtout la partie sur les barycentres ne sont rien d'autres que le snobisme d'écrire $\overrightarrow{xy}$ au lieu de $y - x$. C'est aussi une facilité d'écriture.
16. Définition 1.318.
17. Homographie de $\hat{\mathbb{C}}$: définition 23.74.

LEMooOPOMooWZLSeH

Lemme 23.94.
Soit une application $\alpha\colon \mathbb{C} \to \mathbb{C}$ fixant 1 et 0 et préservant les divisions harmoniques (c'est-à-dire tel que son prolongement à $\hat{\mathbb{C}}$ donné par $\alpha(\infty) = \infty$ préserve les divisions harmoniques). Alors α est un automorphisme de corps[18].

Démonstration. Nous savons que si $a, b, c \in \mathbb{C}$ nous avons $c = (a + b)/2$ si et seulement si $[a, b, c, \infty] = -1$. Vu que α préserve les divisions harmoniques nous avons équivalence entre les affirmations suivantes :

$$c = \frac{a+b}{2} \tag{23.235a}$$
$$[a, b, c, \infty] = -1 \tag{23.235b}$$
$$[\alpha(a), \alpha(b), \alpha(c), \alpha(\infty)] = -1 \tag{23.235c}$$
$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha(c). \tag{23.235d}$$

Donc α préserve les milieux : pour tout $a, b \in \mathbb{C}$ nous avons

$$\alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a) + \alpha(b)}{2}. \tag{23.236}$$

EQooZCWFooZkwYVW

En particulier, cette relation avec $b = 0$ donne (parce que $\alpha(0) = 0$) : $\alpha(a/2) = \alpha(a)/2$. Nous avons au final, en utilisant cela en conjonction avec (23.236) :

$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a+b)}{2}. \tag{23.237}$$

Cela démontre déjà que

$$\alpha(a + b) = \alpha(a) + \alpha(b). \tag{23.238}$$

En particulier $\alpha(-a) = \alpha(0 - a) = \alpha(0) - \alpha(a) = -\alpha(a)$.

Nous passons maintenant à la démonstration du fait que $\alpha(ab) = \alpha(a)\alpha(b)$. Pour tout a différent de 0 et ± 1 nous avons

$$[a, -a, a^2, 1] = \frac{(a - a^2)(-a - 1)}{(-a - a^2)(a - 1)} = -1. \tag{23.239}$$

EQooUPTOooOsEXjp

Et en prenant $\alpha(a)$ en guise de a nous avons aussi

$$[\alpha(a), -\alpha(a), \alpha(a)^2, \alpha(1)] = -1. \tag{23.240}$$

EQooXYKYooQJAiMB

Vu que α préserve les divisions harmoniques, l'équation (23.239) donne aussi

$$\big[\alpha(a), \alpha(-a), \alpha(a^2), \alpha(1)\big] = -1, \tag{23.241}$$

c'est-à-dire

$$\big[\alpha(a), -\alpha(a), \alpha(a^2), 1\big] = -1. \tag{23.242}$$

EQooYYHLooSELB8

Comparant (23.240) avec (23.242) et en tenant compte de l'unicité du birapport[19] nous avons

$$\alpha(a^2) = \alpha(a)^2. \tag{23.243}$$

Avec cela nous pouvons y aller en remarquant que

$$ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2. \tag{23.244}$$

18. Définition 1.43.
19. C'est-à-dire que si trois éléments du birapport sont donnés, le quatrième est fixé. C'est une variation sur la thème de la proposition 23.48(2).

Nous appliquons α à cette dernière équations en tenant compte de ce que nous savons déjà

$$\alpha(ab) = \left(\frac{\alpha(a)+\alpha(b)}{2}\right)^2 - \left(\frac{\alpha(a)-\alpha(b)}{2}\right)^2 = \alpha(a)\alpha(b). \tag{23.245}$$

□

THOooKMKWooZPIDaK

Théorème 23.95 ([618, 619])**.**
Le groupe circulaire de $\mathbb{C}$ est le groupe des bijections $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ préservant l'ensemble des cercles-droites.

Démonstration. L'inclusion dans un sens est facile : les homographies conservent l'ensemble des cercles et droites par la proposition 23.80. Et la conjugaison complexe aussi.

Soit une bijection $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ préservant les cercles-droites. Nous supposons dans un premier temps que $f(0) = 0$, $f(1) = 1$ et $f(\infty) = \infty$.

(1) **Pour f vérifiant $f(0, 1, \infty) = 0, 1, \infty$**
Si $\mathcal{C}$ est un cercle alors $f(\mathcal{C})$ est un cercle ou une droite, mais vu que $\mathcal{C}$ ne contient pas ∞, l'ensemble $f(\mathcal{C})$ ne le contient pas non plus. Donc f transforme un cercle en un cercle et une droite en une droite.

(1a) **f préserve les divisions harmoniques**
Soient a, b, c, x dans $\hat{\mathbb{C}}$ tels que $[a, b, c, x] = -1$. Nous allons prouver que $[f(a), f(b), f(c), f(x)] = -1$.

Si a, b et c sont colinéaires, nous suivons la construction du lemme 23.92. Soit m hors de la droite (ab) et une droite D passant par c et coupant $[ma]$ en p et $[mb]$ en q. Nous posons $n = (pb) \cap (qa)$. Alors $x = (mn) \cap (ac)$.

L'application f est une bijection qui respecte les intersectons, les tangences, les cercles et les droites). Le point $f(m)$ est hors de la droite $\big(f(a)f(b)\big)$. La droite $f(D)$ passe par $f(c)$ et coupe les segments $[f(m)f(a)]$ en $f(p)$ et $[f(m)f(b)]$ en $f(q)$. Alors nous avons

$$f(n) = \big(f(p)f(b)\big) \cap \big(f(q)f(a)\big) \tag{23.246}$$

et aussi

$$f(x) = \big(f(m)f(n)\big) \cap \big(f(a)f(c)\big) \tag{23.247}$$

Donc $f(x)$ se construit à partir de $f(a)$, $f(b)$ et $f(c)$ en suivant la même construction que x à partir de a, b et c. Nous en concluons que $[f(a), f(b), f(c), f(x)] = -1$.

Si a, b et c sont cocycliques, le même raisonnement, en suivant le lemme 23.91 nous donne le même résultat.

(1b) **f est un automorphisme du corps $\mathbb{C}$**
C'est le lemme 23.94.

(1c) **Et enfin ...**
Notre application f est un automorphisme du corps $\mathbb{C}$ qui fixe $\mathbb{R}$ parce qu'elle laisse invariante les droites dans $\mathbb{C}$. Donc la proposition 6.6 nous dit que f est soit l'identité soit la conjugaison complexe. Dans les deux cas, f est dans le groupe circulaire.

(2) **Pour f plus générale**
Nous ne supposons plus que f fixe 0, 1 et ∞. En tout cas les nombres $f^{-1}(1)$, $f^{-1}(0)$ et $f^{-1}(\infty)$ sont distincts parce que f est une bijection. Nous pouvons considérer une homographie[20] $\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ telle que $\phi(1) = f^{-1}(1)$, $\phi(0) = f^{-1}(0)$ et $\phi(\infty) = f^{-1}(\infty)$. Dans ce cas l'application

$$g = f \circ \phi \tag{23.248}$$

20. Attention : ici nous parlons d'homographies de $\hat{\mathbb{C}}$, pas de $P(\mathbb{C}^2)$. L'existence d'une telle application demande de composer le corolaire 23.35 avec l'application φ_0 et la définition 23.74 et 23.75.

vérifie $g(1) = 1$, $g(0) = 1$ et $g(\infty) = \infty$ tout en continuant à transformer un cercle-droite en un cercle-droite. Donc $f \circ \phi$ est soit l'identité soit la conjugaison complexe. Avec ça, l'application

$$f = g \circ \phi^{-1} \tag{23.249}$$

est la composée d'une homographie avec soit l'identité soit la conjugaison complexe. Elle est donc dans le groupe circulaire.

□

23.6.6 Action du groupe modulaire

Le **demi-plan de Poincaré** est l'ensemble

$$P = \{z \in \mathbb{C} \text{ tel que } \Im(z) > 0\}. \tag{23.250}$$

Le **groupe modulaire** est le quotient de groupes

$$\mathrm{PSL}(2, \mathbb{Z}) = \frac{\mathrm{SL}(2, \mathbb{Z})}{\mathbb{Z}_2}. \tag{23.251}$$

Ce sont donc les matrices au signe près de la forme

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{23.252}$$

où a, b, c et d sont entiers tels que $ad - cb = 1$.

ThoItqXCm

Théorème 23.96 ([620]).
Le groupe modulaire agit fidèlement (définition 2.30) sur le demi-plan de Poincaré par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} * z = \frac{az + b}{cz + d}. \tag{23.253}$$

EqVXywlB

L'ensemble $D = D_1 \cup D_2$ avec

$$D_1 = \{z \in P \text{ tel que } |z| > 1, -\frac{1}{2} \leqslant \Re(z) < \frac{1}{2}\} \tag{23.254a}$$

$$D_2 = \{z \in P \text{ tel que } |z| = 1, -\frac{1}{2} \leqslant \Re(z) \leqslant 0\} \tag{23.254b}$$

est un domaine fondamental (définition 2.37) de cette action.

De plus si nous notons

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{23.255}$$

alors pour tout $z \in P$, il existe $A \in \mathrm{gr}(S, T)$ telle que $A * z \in D$.

Démonstration. Nous divisions la preuve en plusieurs étapes.

(1) **Bien définie**
D'abord il faut remarquer que l'action (23.253) est bien définie par rapport au quotient : $A * z = (-A) * z$. La vérification est immédiate.

(2) **Interne**
Montrons que si $A \in \mathrm{PSL}(2, \mathbb{Z})$ et $z \in P$ alors $A * z \in P$. Nous avons

$$A * z = \frac{az + b}{cz + d} = \frac{(az + b)(c\bar{z} + d)}{|cz + d|^2} = \frac{a|z|c + azd + bc\bar{z} + bd}{|cz + d|^2}, \tag{23.256}$$

et donc en décomposant $z = \Re(z) + i\Im(z)$,

$$\Im(A * z) = \Im\left(\frac{azd + bc\bar{z}}{|cz + d|^2}\right) = \frac{ad - bc}{|cz + d|^2}\Im(z) = \frac{\Im(z)}{|cz + d|^2} \tag{23.257}$$

où nous avons tenu compte de $ad - bc = 1$. Donc l'action respecte la (stricte) positivité de la partie imaginaire.

(3) **Action**

Nous vérifions maintenant que la formule donne bien une action : $A * (B * z) = (AB) * z$. Cela est un bon calcul :

$$A * (B * z) = A * \left(\frac{a'z + b'}{c'z + d'} \right) \tag{23.258a}$$

$$= \frac{a\left(\frac{a'z+b'}{c'z+d'}\right) + b}{c\left(\frac{a'z+b'}{c'z+d'}\right) + d} \tag{23.258b}$$

$$= \frac{a(a'z + b') + b(c'z + d')}{c(a'z + b') + d(c'z + d')} \tag{23.258c}$$

$$= \frac{(aa' + bc')z + (ab' + bd')}{(ca' + dc')z + (cb' + dd')} \tag{23.258d}$$

$$= \begin{pmatrix} aa' + bc' & ab' + bd' \\ a'c + dc' & cb' + dd' \end{pmatrix} * z \tag{23.258e}$$

$$= (AB) * z. \tag{23.258f}$$

(4) **Fidèle**

Soit $A \in \mathrm{PSL}(2, \mathbb{Z})$ tel que pour tout $z \in P$ nous ayons

$$\frac{az + b}{cz + d} = z. \tag{23.259}$$

Alors nous avons

$$cz^2 + (d - a)z + b = 0. \tag{23.260}$$

Cela est donc un polynôme en z qui s'annule sur un ouvert [21] (le demi-plan de Poincaré). Il doit donc être identiquement nul, donc $c = b = a - d = 0$. Si vous n'y croyez pas, écrivez pour $z = \epsilon i$ (avec $\epsilon > 0$) :

$$-c\epsilon^2 + \epsilon(d - a)i + b = 0 \tag{23.261}$$

pour tout ϵ. Le fait d'avoir $c\epsilon^2 = b$ pour tout ϵ implique que $c = b = 0$. Donc A est de la forme

$$A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \tag{23.262}$$

avec la contrainte supplémentaire que $ad = 1$, les nombres a et b étant entiers. Nous avons donc soit $a = d = 1$ soit $a = d = -1$. Étant donné le quotient par $\mathbb{Z}_2$, ces deux possibilités donnent le même élément de $\mathrm{PSL}(2, \mathbb{Z})$.

(5) **Les orbites intersectent D**

Soit $z \in P$. Nous devons trouver $A \in \mathrm{PSL}(2, \mathbb{Z})$ tel que $A * z \in D$. Nous savons déjà que

$$\Im(A * z) = \frac{\Im(z)}{|cz + d|^2}. \tag{23.263}$$

Nous notons $\mathcal{O}_z$ l'orbite de z sous le groupe modulaire et nous posons

$$I_z = \{\Im(u) \text{ tel que } u \in \mathcal{O}_z\} = \{\Im(A * z) \text{ tel que } A \in \mathrm{PSL}(2, \mathbb{Z})\}, \tag{23.264}$$

l'ensemble des parties imaginaires des éléments de l'orbite de z. Nous allons montrer que cet ensemble est borné vers le haut en montrant que la quantité $|cz + d|$ ne peut, à z donné, prendre qu'un nombre fini de valeurs plus grandes que $\Im(z)$ [22]. Nous cherchons donc les couples $(c, d) \in \mathbb{Z}^2$ tels que $|cz + d| < 1$.

21. On ne peut pas dire que $b = 0$ simplement en justifiant qu'on l'obtient en posant $z = 0$ parce que $z = 0$ n'est pas dans le demi-plan de Poincaré.

22. Bien que cela ne soit pas indispensable pour la preuve, remarquons que I_z ne comprend qu'une quantité au plus dénombrable de valeurs. Le fait que, à z donné, la quantité $|cz + d|^2$ puisse être rendue aussi grande que l'on veut est évident. Donc I_z est borné vers le bas par zéro (qui n'est pas atteint, mais qui est une valeur d'adhérence).

Nous avons $\Im(cz+d)=c\Im(z)$, donc $|cz+d|\geqslant|c\Im(z)|$, mais il n'y a qu'un nombre fini de $c\in\mathbb{Z}$ tels que $|c\Im(z)|<1$. De la même façon, pour la partie réelle nous avons

$$\Re(cz+d)=c\Re(z)+d, \tag{23.265}$$

et pour chaque c, il n'y a qu'un nombre fini de $d\in\mathbb{Z}$ qui laissent cette quantité plus petite que 1 (en valeur absolue).

Donc I_z possède un maximum. Soit $A_1\in\mathrm{PSL}(2,\mathbb{Z})$ tel que $\Im(A_1*z)=\max I_z$. Nous notons $z_1=A_1*z$, et que nous n'avons à priori pas l'unicité. Nous allons maintenant agir sur z_1 avec l'élément

$$T=\begin{pmatrix}1&1\\0&1\end{pmatrix} \tag{23.266}$$

pour ramener z_1 dans le domaine D. Si $u\in P$ nous avons $T*u=u+1$ et donc

$$T^n*u=u+n. \tag{23.267}$$

Vu que D est de largeur 1, il existe un n (éventuellement négatif) tel que

$$\Re(T^n*z_1)\in[-\frac{1}{2},\frac{1}{2}[. \tag{23.268}$$

Notons qu'ici le fait d'être ouvert d'un côté et fermé de l'autre joue de façon essentielle (pour l'unicité aussi). Nous notons $z_2=T^n*z_1$.

Supposons un instant que $|z_2|<1$. Nous considérons l'élément

$$S=\begin{pmatrix}0&-1\\1&0\end{pmatrix} \tag{23.269}$$

qui fait

$$\Im(S*z)=\frac{\Im z}{|z|^2}. \tag{23.270}$$

Donc si $|z_2|<1$ alors $\Im(S*z_2)>\Im(z_2)$, ce qui contredit la maximalité de $\Im(z_2)$ dans I_z. Nous en déduisons que $|z_2|\geqslant 1$.

Si $|z_2|>1$, alors $z_2\in D_1$ et c'est bon. Si $|z_2|=1$, alors il faut encore un peu travailler. Si $z_2\pm1$ est à l'intérieur du disque, alors en agissant avec T ou T^{-1} nous retrouvons la même contradiction que précédemment. En écrivant $z_2=e^{i\theta}$, nous devons donc avoir $2\cos(\theta)\leqslant 1$ ou encore $|\Re(z_2)|\leqslant\frac{1}{2}$. Donc si $\Re(z_2)\leqslant 0$ alors $z_2\in D_2$.

Le dernier cas à traiter est $\Re(z_2)\in\,]0,\frac{1}{2}]$, c'est-à-dire $\theta\in[\frac{\pi}{3},\frac{\pi}{2}[$. Dans ce cas l'action avec S ramène l'angle dans la bonne zone parce que $S*z=-\frac{1}{z}$ et donc $S*(\rho e^{-i\theta})=-\frac{1}{\rho}e^{-i\theta}$.

(6) **Unicité**

Nous voulons à présent montrer que si $z\in D$, alors $A*z$ n'est plus dans D (sauf si $A=\pm\mathbb{1}$). Nous supposons que $z\in D$ et $A\in\mathrm{PSL}(2,\mathbb{Z})$ soient tels que $A*z\in D$, et nous prouvons qu'alors soit nous arrivons à une contradiction soit nous arrivons à $A=\mathbb{1}$. Pour cela nous allons décomposer en de nombreux cas.

(6a) Nous commençons par $\Im(A*z)\geqslant\Im(z)$. Dans ce cas nous avons $|cz+d|\leqslant 1$ et en particulier $|c||\Im(z)|\leqslant 1$. Étant donné que le point de D qui a la partie imaginaire la plus petite est $-\frac{1}{2}+\frac{2}{\sqrt{3}}i$, nous trouvons $|c|\leqslant 2/\sqrt{3}$. Vu que c doit être entier, nous avons trois cas : $c=-1,0,1$.

(6a.i) Soit $c=0$. Alors $A=\begin{pmatrix}a&b\\0&d\end{pmatrix}$ et la condition de déterminant est $ad=1$, ce qui signifie $a=d=1$ (la possibilité $a=b=-1$ est « éliminée » le quotient par $\mathbb{Z}_2$ définissant $\mathrm{PSL}(2,\mathbb{Z})$). La matrice A doit alors être de la forme

$$A=\begin{pmatrix}1&b\\0&1\end{pmatrix} \tag{23.271}$$

et $A*z=z+b$. Si $z\in D$, alors le seul $z+b$ à être (peut-être) encore dans D est $b=0$, mais alors A est l'identité.

(6a.i) Soit $c = 1$. Alors la condition $|cz + d| \leqslant 1$ nous donne trois possibilités[23] : $d = -1, 0, 1$.

A. Si $d = -1$, alors nous devons avoir $|z-1| \leqslant 1$. Il est instructif de faire un dessin, mais le point d'intersection entre les cercles $|z| = 1$ et $|z - 1| = 1$ est le point $\frac{1}{2} + \frac{\sqrt{3}}{2}i$, qui n'est pas dans D. Bref, il n'y a pas de points dans D vérifiant $|z - 1| \leqslant 1$.

B. Si $d = 1$, alors (et c'est maintenant que la dissymétrie de D intervient) nous avons le point

$$z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \tag{23.272}$$

qui est dans D et qui vérifie $|z + 1| \leqslant 1$. Voyons à quoi ressemble la matrice A dans ce cas. Son déterminant est $a - b = 1$. Nous écrivons donc

$$A = \begin{pmatrix} b+1 & b \\ 1 & 1 \end{pmatrix}, \tag{23.273}$$

et en tenant compte du fait que $z\bar{z} = |z + 1| = 1$, nous calculons

$$A * z = \frac{(b+1)z + b}{z+1} \tag{23.274a}$$

$$= \frac{(bz + z + b)(\bar{z} + 1)}{|z+1|^2} \tag{23.274b}$$

$$= z + b + 1. \tag{23.274c}$$

La seule façon de ne pas quitter D est d'avoir $b = -1$, mais alors nous avons

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \tag{23.275}$$

et $A * z = z$. Donc au final z est quand même le seul de son orbite à être dans D.

Notons au passage cette très intéressante propriété du point

$$z_0 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i. \tag{23.276}$$

C'est un point de qui vérifie $z_0 = A * z_0$ pour un élément non trivial A de $\mathrm{PSL}(2, \mathbb{Z})$. L'existence d'un tel élément est ce qui va nous coûter un peu de sueur pour prouver que $PSL(2, \mathbb{Z})$ est engendré par S et T.

C. Le cas $d = 0$ nous fait écrire $1 = \det A = -b$, donc $b = -1$ et

$$A = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}. \tag{23.277}$$

Nous avons alors $A * z = a - \frac{1}{z}$. De plus la condition $|z| \leqslant 1$ revient à $|z = 1|$. Pour les nombres complexes de module 1, l'opération $z \to -1/z$ est la symétrie autour de l'axe des imaginaires purs. Le seul à ne pas sortir de D est le fameux $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, qui revient sur lui-même avec $a = -1$.

Nous passons à la possibilité $c = -1$. Dans ce cas la matrice est de la forme

$$A = \begin{pmatrix} a & b \\ -1 & d \end{pmatrix}, \tag{23.278}$$

et nous revenons au cas $c = 1$ en prenant $-A$ au lieu de A.

23. Je ne rigolais pas quand je disais qu'on allait avoir de nombreux cas.

(6b) Nous passons au cas $\Im(A*z)<\Im(z)$. Nous récrivons cette condition avec

$$\Im(A*z)<\Im\big(A^{-1}*(A*z)\big). \tag{23.279}$$

Si nous supposons que z et A sont tels que z et $A*z$ soient tous deux dans D, alors $z'=A*z$ est un élément de D tel que

$$\Im(z')<\Im(A^{-1}*z'). \tag{23.280}$$

Or nous avons vu qu'aucun élément de D vérifiant cette condition n'existait sans être trivial (celui qui ne bouge pas). Pour cela il suffit d'appliquer tout ce que nous venons de dire avec A^{-1} au lieu de A.

(7) **Quelques conclusions**

Après avoir passé tous les cas en revue, le fameux point $z_0=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$ est l'unique point de D à accepter une matrice non triviale $A\in\mathrm{PSL}(2,\mathbb{Z})$ telle que $z_0=A*z_0$.

Nous remarquons aussi que tous les points de P sont ramenés dans D par une matrice obtenue comme produit de T, S, T^{-1} et S^{-1}.

□

CorJQwgNp

Corolaire 23.97 ([621]).
Les matrices S et T génèrent le groupe modulaire au sens où toute matrice de $\mathrm{PSL}(2,\mathbb{Z})$ *s'écrit comme*

$$T^{m_1}S^{p_1}\dots T^{m_k}S^{p_k} \tag{23.281}$$

pour un certain k et des nombres $m_i,p_i\in\mathbb{Z}$. Autrement dit, $\mathrm{PSL}(2,\mathbb{Z})=\mathrm{gr}(S,T)$.

Démonstration. Soit z, un point de D autre que z_0. Alors si $A\in\mathrm{PSL}(2,\mathbb{Z})$ est non trivial nous avons $A*z$ hors de D. Du coup, comme vu dans la démonstration du théorème 23.96, il existe $B\in\mathrm{gr}(S,T)$ tel que $B*(A*z)\in D$. Vu que D ne contient qu'un seul point de chaque orbite, nous avons

$$B*A*z=z, \tag{23.282}$$

et donc $BA=\pm\mathbb{1}$, ce qui prouve que[24] $A=B^{-1}$, c'est-à-dire que $A\in\mathrm{gr}(S,T)$. □

24. Dans $\mathrm{PSL}(2,\mathbb{Z})$, nous n'avons pas besoin de mettre $\pm$ parce qu'il est compris dans la définition.

Chapitre 24

Analyse vectorielle

24.1 Le théorème de Green

Soit un champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_2(x,y,z) \end{pmatrix} \tag{24.1}$$

et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$ donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{24.2}$$

Nous avons défini la circulation de F le long de σ par

$$\begin{aligned} \int_\sigma F \cdot d\sigma &= \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt \\ &= \int_a^b \Big[F_1\big(\sigma(t)\big)x'(t) + F_2\big(\sigma(t)\big)y'(y) + F_3\big(\sigma(t)\big)z'(t)\Big]dt \\ &= \int_\sigma F_1dx + F_2dy + F_3dz. \end{aligned} \tag{24.3}$$

La dernière ligne est juste une notation compacte [1]. Elle sert à se souvenir qu'on va mettre x' à côté de F_1, y' à côté de F_2 et z' à côté de F_3. L'avantage de cette notation est qu'on peut écrire d'autres combinaisons.

Si f et g sont deux fonctions sur $\mathbb{R}^3$, nous pouvons écrire

$$\int_\sigma f dy + g dz. \tag{24.4}$$

Cela signifie

$$\int_a^b \Big[f\big(\sigma(t)\big)y'(t) + g\big(\sigma(t)\big)z'(t)\Big]dt. \tag{24.5}$$

Soit D une région du plan et σ, son contour que nous prenons, par convention [2], dans l'orientation trigonométrique, comme indiqué sur la figure 24.1. Nous supposons également que le domaine D n'a pas de trous intérieurs.

Nous notons par $\sigma = \partial D$ le bord de D, c'est-à-dire le contour dont nous venons de parler.

1. Il y aurait beaucoup de choses à dire là-dessus, mais la vie est trop courte pour parler de formes différentielles, et c'est dommage.

2. Il y aurait beaucoup de choses à dire sur ça aussi, mais...

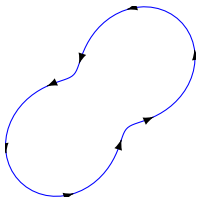


FIGURE 24.1: Un contour avec son orientation. LabelFigVDFMooHMmFZr

Théorème 24.1 (Théorème de Green).
Soient $P, Q : D \to \mathbb{R}$ deux fonctions de classe C^1. Alors

$$\int_{\partial D} P dx + Q dy = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy. \tag{24.6}$$ EqThoGreen

Pour rappel, l'intégrale du membre de gauche signifie

$$\int_a^b \Big[P\big(\sigma(t)\big)\sigma'_x(t) + Q\big(\sigma(t)\big)\sigma'_y(t) \Big] dt. \tag{24.7}$$

Ce n'est d'ailleurs rien d'autre que l'intégrale du champ de vecteurs $\begin{pmatrix} P \\ Q \end{pmatrix}$.

Corolaire 24.2.
L'aire du domaine D est donnée par

$$A = \frac{1}{2} \int_{\partial D} (xdy - ydx). \tag{24.8}$$

Démonstration. L'intégrale $\int_{\partial D}(xdy - ydx)$ se traite avec le théorème de Green où l'on pose $P = -y$ et $Q = x$. Nous avons donc

$$\begin{aligned} \int_{\partial D} -ydx + xdy &= \int_D \left(\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y} \right) dxdy \\ &= \int_D 2\, dxdy. \end{aligned} \tag{24.9}$$

La dernière ligne est bien le double de la surface. □

Exemple 24.3.
Calculons (encore une fois) l'aire du disque de rayon R. Il s'agit de calculer l'intégrale

$$I = \frac{1}{2} \int_\sigma (xdt - ydx) \tag{24.10}$$

où σ est le cercle donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R\cos(t) \\ R\sin(t) \end{pmatrix} \tag{24.11}$$

Le calcul est

$$\begin{aligned} I &= \frac{1}{2} \int_0^{2\pi} \underbrace{R\cos(\theta)}_{x} \underbrace{R\cos(\theta)}_{y'} - \underbrace{R\sin(\theta)}_{y} \underbrace{(-R\sin(\theta))}_{x'}\, d\theta \\ &= \frac{R^2}{2} \int_0^{2\pi} d\theta \\ &= \pi R^2. \end{aligned} \tag{24.12}$$

△

Exemple 24.4.
Calculons l'aire de l'ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant 1 \tag{24.13}$$

dont le bord est donné par

$$\begin{cases} x(t)=a\cos(t) & (24.14a)\\ y(t)=b\sin(t). & (24.14b)\end{cases}$$

Le terme xdy devient $a\cos(t)b\cos(t)=ab\cos^2(t)$ et le terme ydx devient $b\sin(t)(-a\sin(t))=-ab\sin^2(t)$. L'intégrale qui donne la surface est donc

$$\frac{1}{2}\int_{\partial D}(xdy-ydx)=\frac{1}{2}\int_0^{2\pi}ab=\pi ab. \tag{24.15}$$

△

Le théorème de Green peut être mis sous une autre forme.

ThoGreenVecto

Théorème 24.5 (Théorème de Green, forme vectorielle)**.**
Si G est un champ de vecteurs sur D, nous avons

$$\int_{\partial D}G\cdot d\sigma=\int_D(\nabla\times G)\cdot dS \tag{24.16}$$

EqGreenVecto

où le second membre est le flux de $\nabla\times G$ sur la surface D.

Démonstration. Analysons le membre de droite. Nous savons que D est une surface dans le plan $\mathbb{R}^2$. Le vecteur normal à la surface est donc simplement le vecteur (constant) e_z. Le produit scalaire $(\nabla\times F)\cdot dS$ est donc $(\nabla\times F)\cdot e_z$ et se réduit à la troisième composante du rotationnel, c'est-à-dire

$$\frac{\partial F_2}{\partial x}-\frac{\partial F_1}{\partial y}. \tag{24.17}$$

Cela est bien le membre de droite de l'équation (24.6). Le membre de gauche de cette dernière est bien le membre de gauche de (24.16). □

ExempleGreenSqL

Exemple 24.6.
Soit le champ de vecteurs $F(x,y)=\begin{pmatrix}xy^2\\ y+x\end{pmatrix}$, et soit à calculer

$$\int_D\nabla\times F\cdot dS \tag{24.18}$$

où D est la région comprise entre les courbes $y=x^2$ et $y=x$ pour $x\geqslant 0$ (voir la figure 24.2).

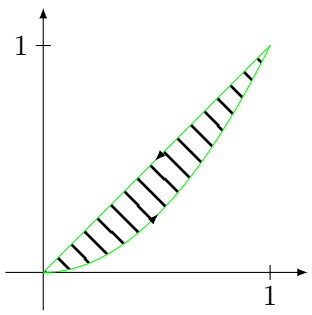


FIGURE 24.2: Le contour d'intégration pour l'exemple 24.6. LabelFigLLVMooWOkvAB

Nous pouvons calculer cette intégrale directement en calculant le rotationnel de F :

$$\nabla \times F = \begin{pmatrix} 0 \\ 0 \\ 1-2xy \end{pmatrix}. \tag{24.19}$$

Par conséquent l'intégrale à effectuer est

$$I = \int_0^1 dx \int_{x^2}^x (1-2xy)dy = \frac{1}{12}. \tag{24.20}$$

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=1-2*x*y
sage: f.integrate(y,x**2,x).integrate(x,0,1)
(x, y) |--> 1/12
```

L'autre façon de calculer l'intégrale est d'utiliser le théorème de Green et de calculer la circulation de F le long de ∂D :

$$I = \int_{\partial D} F \cdot \sigma. \tag{24.21}$$

Le chemin $\sigma = \partial D$ est composé de la parabole $y = x^2$ et du segment de droite $x = y$. Attention : il faut respecter l'orientation. Nous avons

$$\sigma_1(t) = (t, t^2) \tag{24.22}$$

et

$$\sigma_2(t) = (1-t, 1-t). \tag{24.23}$$

Notez bien que le second chemin est $(1-t, 1-t)$ et non (t, t) parce qu'il faut le parcourir dans le bon sens (voir le dessin).

Commençons par le premier chemin :

$$\begin{aligned} \sigma_1(t) &= (t, t^2) \\ \sigma_1'(t) &= (1, 2t) \\ F\big(\sigma_1(t)\big) &= \begin{pmatrix} t^5 \\ t+t^2 \end{pmatrix}, \end{aligned} \tag{24.24}$$

et par conséquent

$$F\big(\sigma_1(t)\big) \cdot \sigma_1'(t) = t^5 + 2t^2 + 2t^3, \tag{24.25}$$

et le premier morceau de la circulation vaut

$$\int_{\sigma_1} F \cdot d\sigma_1 = \int_0^1 t^5 + 2t^2 + 2t^3 = \frac{4}{3}. \tag{24.26}$$

Pour le second chemin :

$$\begin{aligned} \sigma_2(t) &= (1-t, 1-t) \\ \sigma_2'(t) &= (-1, -1) \\ F\big(\sigma_2(t)\big) &= \begin{pmatrix} (1-t)^3 \\ 2(1-t) \end{pmatrix}. \end{aligned} \tag{24.27}$$

Par conséquent

$$F\big(\sigma_2(t)\big) \cdot \sigma_2(t) = -(1-t)^2 - 2(1-t). \tag{24.28}$$

Le second morceau de la circulation est par conséquent

$$\int_0^1 -(1-t)^2 - 2(1-t)dt = -\frac{5}{4}. \tag{24.29}$$

La circulation de F le long de σ est donc égale à

$$\frac{4}{3} - \frac{5}{4} = \frac{1}{12}. \tag{24.30}$$

Comme prévu, nous obtenons le même résultat. △

24.2 Théorème de la divergence dans le plan

24.2.1 La convention de sens de parcours

Soient D, un domaine dans le plan et un paramétrage

$$\begin{aligned} \sigma\colon [a,b] &\to \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \end{aligned} \tag{24.31}$$

un paramétrage du bord ∂D de D. La normale à σ est perpendiculaire à la tangente, donc la normale extérieure de norme 1 vaut

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}} \quad \text{ou} \quad n = -\frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{24.32}$$

Comment faire le choix ?

Nous prenons comme convention que le sens *du chemin* doit être tel que le vecteur normal extérieur soit

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{24.33}$$

Donc si le chemin σ donne lieu à un vecteur n pointant vers l'intérieur, il faut utiliser le chemin qui va dans le sens contraire : $\tilde{\sigma}(t) = \sigma(1-t)$.

Les vecteurs tangents et normaux d'un contour sont dessinés sur la figure 24.3.

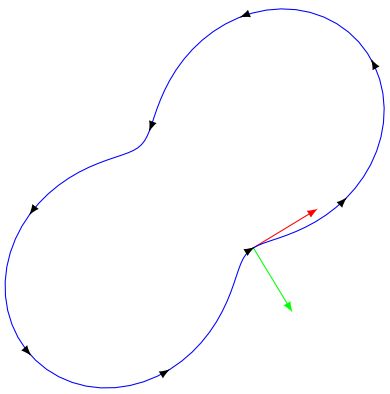


FIGURE 24.3: Le champ de vecteurs tangents est dessiné en rouge tandis qu'en vert nous avons le champ de vecteurs normaux extérieurs. LabelFigDDCTooYscVzA

24.2.2 Théorème de la divergence

Théorème 24.7 (Théorème de la divergence)**.**
Soit F un champ de vecteurs sur $\mathbb{R}^2$. Le flux de F à travers le bord de D est égal à l'intégrale de la divergence de F sur D. En formule :

$$\int_{\partial D} F \cdot n \, d\sigma = \int_D \nabla \cdot F \, dxdy. \tag{24.34}$$

Démonstration. Tant $F \cdot n$ que $\nabla \cdot F$ sont des fonctions. Le membre de gauche est donc l'intégrale d'une fonction sur un chemin et le membre de droite est l'intégrale d'une fonction sur une surface. Notre convention de sens de parcours du chemin permet d'écrire le produit scalaire $F \cdot n$ sous la forme suivante :

$$\begin{aligned} F \cdot n &= \frac{1}{\|\sigma'\|} \begin{pmatrix} F_x \\ F_y \end{pmatrix} \cdot \begin{pmatrix} y' \\ -x' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} (F_x y' - F_y x') \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \end{aligned} \tag{24.35}$$

Par conséquent, la *fonction*

$$F \cdot n \tag{24.36}$$

est la même que la *fonction*

$$\frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \tag{24.37}$$

L'intégrale de cette dernière fonction sur le chemin σ est

$$\begin{aligned} I &= \int_\sigma F \cdot n \\ &= \int_\sigma \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma' \\ &= \int_a^b \frac{1}{\|\sigma'(t)\|} \begin{pmatrix} -F_y\big(\sigma(t)\big) \\ F_x\big(\sigma(t)\big) \end{pmatrix} \cdot \sigma'(t) \|\sigma'(t)\| dt \\ &= \int_a^b \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'(t) dt. \end{aligned} \tag{24.38}$$

Cette dernière intégrale est la circulation du champ de vecteurs $\begin{pmatrix} -F_y \\ F_x \end{pmatrix}$ sur le chemin σ. Le théorème de Green 24.5 nous enseigne que la circulation le long d'un chemin est égale au flux du rotationnel à travers la surface. Par conséquent,

$$I = \int_D \left(\nabla \times \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \right) \cdot dS = \int_D \nabla \cdot F \, dxdy \tag{24.39}$$

□

24.3 Théorème de Stokes

Nous nous mettons maintenant dans $\mathbb{R}^3$, et nous y considérons une surface paramétrée S donc le bord est ∂S.

THOooIRYTooFEyxif

Théorème 24.8 (Théorème de Stokes)**.**
Alors le flux du rotationnel de F à travers S est égal à la circulation de F le long du bord. En formule :

$$\int_S \nabla \times F \cdot dS = \int_{\partial S} F \cdot d\sigma. \tag{24.40}$$

Nous pouvons nous donner une idée du pourquoi ce théorème est vrai. D'abord, si la surface est plate, cela est exactement le théorème de Green 24.5. Supposons maintenant que le bord reste plat, mais que la surface se déforme un petit peu. Le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 0 \end{pmatrix} \tag{24.41}$$

est tout autant le bord du disque plat de rayon 1 que celui de la demi-sphère

$$\phi(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{1 - x^2 - y^2} \end{pmatrix}. \tag{24.42}$$

Le champ de vecteur que nous considérons est $G = \nabla \times F$. Il a un certain flux à travers le disque plat, et ce plus est égal à la circulation de F sur σ. Quel est le flux de G à travers la demi-sphère ? Étant donné que $\nabla \cdot G = \nabla \cdot (\nabla \times F) = 0$, le champ de vecteurs G est incompressible, de telle façon que tout ce qui rentre dans la demi-sphère doit en sortir. Le flux de G à travers la demi-sphère doit par conséquent être égal à celui à travers le disque plat.

Exemple 24.9.
Soit C l'intersection entre le cylindre $x^2 + y^2 = 1$ et le plan $x + y + z = 1$. Calculer la circulation de

$$F(x, y, z) = \begin{pmatrix} -y^3 \\ x^3 \\ -z^3 \end{pmatrix} \tag{24.43}$$

le long de C.

Au lieu de calculer directement

$$\int_C F \cdot d\sigma, \tag{24.44}$$

nous allons calculer

$$\int_S \nabla \times F \cdot dS \tag{24.45}$$

où S est une surface dont C est le bord. Cette intégrale est à calculer avec la formule (20.293).

La première chose à faire est de trouver une surface dont le bord est C et en trouver un paramétrage ϕ. Le plus simple est de prendre le graphe du plan sur le cercle $x^2 + y^2 + 1$. Un paramétrage de cette surface est simplement

$$\begin{aligned} \phi\colon D &\to \mathbb{R}^3 \\ (x, y) &\mapsto \begin{pmatrix} x \\ y \\ 1 - x - y \end{pmatrix} \end{aligned} \tag{24.46}$$

où D est le disque de rayon 1. Étant donné que cela paramètre le plan $x + y + z - 1 = 0$, le vecteur normal est $n = e_x + e_y + z_z$. Nous pouvons cependant calculer ce vecteur normal en suivant la recette usuelle. D'abord les vecteurs tangents sont

$$\frac{\partial \phi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{\partial \phi}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{24.47}$$

Et le vecteur normal est donné par le produit vectoriel :

$$\begin{aligned} n &= \frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y} \\ &= \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} \\ &= e_x + e_y + e_z. \end{aligned} \tag{24.48}$$

Ensuite, le rotationnel de F est donné par

$$\nabla \times F = 3(x^2 + y^2)e_z. \tag{24.49}$$

Par conséquent,

$$\nabla \times F \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) = 3(x^2 + y^2). \tag{24.50}$$

L'intégrale à calculer est donc

$$\begin{aligned} \int_S \nabla \times F \cdot dS &= \int_D (\nabla \times F)\big(\phi(x,y)\big) \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) dxdy \\ &= 3\int_D (x^2 + y^2)dxdy. \end{aligned} \tag{24.51}$$

Cette dernière intégrale est l'intégrale d'une fonction sur le disque de rayon 1. Elle s'effectue en passant aux coordonnées polaires :

$$3\int_D (x^2 + y^2)dxdy = 3\int_0^{2\pi} d\theta \int_0^1 (r^2)r\,dr = \frac{3\pi}{2}. \tag{24.52}$$

△

24.4 Théorème de Gauss

Soit V une partie de $\mathbb{R}^3$ délimitée par une surface S sur laquelle nous considérons la normale extérieure. Soit F un champ de vecteurs sur $\mathbb{R}^3$.

Théorème 24.10 (Théorème de la divergence ou de Gauss)**.**
Le flux d'un champ de vecteur F à travers une surface fermée est égal à l'intégrale de la divergence sur le volume correspondant :

$$\int_{\partial V} F \cdot dS = \int_V \nabla \cdot F\,dxdydz. \tag{24.53}$$

Ce théorème signifie que la quantité de fluide qui s'accumule dans le volume (le flux est ce qui rentre moins ce qui sort) est égale à l'intégrale de $\nabla \cdot F$ sur le volume, alors que nous savons que, localement, la quantité $\nabla \cdot F(x,y,z)$ est la quantité de fluide qui s'accumule au point (x,y,z).

Remarque 24.11.
Ce théorème ne fonctionne qu'avec des surfaces fermées. Essayer de l'appliquer au calcul de flux à travers des surfaces ouvertes n'a pas de sens parce qu'une surface ouverte ne délimite pas un volume.

24.12.
La formule de la divergence peut être utilisée comme intégration par partie. Si u est une fonction et F un champ de vecteurs, $\nabla(uF) = \nabla(u) \cdot F + u\nabla \cdot F$ et alors

$$\int_{\partial V} uF \cdot n = \int_V \nabla(uF) = \int_V u\nabla \cdot F + \int_V F \cdot \nabla u \tag{24.54}$$

où n est le champ de vecteurs normal extérieur à V. En remettant les termes dans un ordre qui ressemble plus à l'intégration par partie :

$$\int_V F \cdot \nabla u = \int_{\partial V} uF \cdot n - \int_V u \nabla F. \tag{24.55}$$

EQooRUCKooUUrxxJ

Exemple 24.13.
Calculer le flux du champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} 2x \\ y^2 \\ z^2 \end{pmatrix} \tag{24.56}$$

à travers la sphère de rayon 1 centrée à l'origine. Nous utilisons le théorème de la divergence

$$\int_S F \cdot n \, dS = \int_B \nabla \cdot F \, dxdydz \tag{24.57}$$

où S est la sphère et B est la boule (la sphère pleine). La divergence de F se calcule :

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 2 + 2x + 2y. \tag{24.58}$$

L'intégrale est donc en trois termes :

$$\begin{aligned} \int_B 2 = 2\text{Volume(B)} = \frac{8\pi}{3} \\ \int_B y \, dxdydz = 0 \\ \int_B z \, dxdydz = 0. \end{aligned} \tag{24.59}$$

△

Dans certains cas le théorème de Gauss permet de simplifier le calcul de l'intégrale d'une fonction sur une surface.

Exemple 24.14.
Soit à calculer l'intégrale

$$I = \int_{\partial B} (x^2 + y + z) dS, \tag{24.60}$$

c'est-à-dire l'intégrale de la fonction $x^2 + y + z$ sur la sphère. Le vecteur normal à la sphère est

$$n = xe_x + ye_y + ze_z. \tag{24.61}$$

Étant donné que nous sommes sur la sphère de rayon 1, ce vecteur est même normé. La fonction que nous regardons n'est rien d'autre que $F \cdot n$ avec

$$F = \begin{pmatrix} x \\ 1 \\ 1 \end{pmatrix}. \tag{24.62}$$

Nous pouvons donc simplement intégrer $\nabla \cdot F$ sur toute la boule :

$$I = \int_B \nabla \cdot F \, dxdydz = \int_B 1 \, dxdudz = \frac{4\pi}{3}. \tag{24.63}$$

△

24.5 Coordonnées curvilignes

24.5.1 Base locale

Nous connaissons déjà les coordonnées sphériques et cylindriques sur $\mathbb{R}^3$. Ce sont des systèmes « un peu courbes ». Il en existe bien entendu de nombreux autres (sinon ce ne serait pas drôle), et nous allons faire une étude un peu générale de ces systèmes de **coordonnées curvilignes**. Des coordonnées curvilignes sur $\mathbb{R}^3$ est n'importe quel [3] système qui permet de repérer un point de $\mathbb{R}^3$ à partir de trois nombres.

Il s'agit donc d'un ensemble de trois applications

$$x_i : \mathbb{R}^3 \to \mathbb{R}. \tag{24.64}$$

Les coordonnées cylindriques sont

$$\begin{cases} x_1(r,\theta,z) = r\cos\theta & \text{(24.65a)} \\ x_2(r,\theta,z) = r\sin\theta & \text{(24.65b)} \\ x_3(r,\theta,z) = z & \text{(24.65c)} \end{cases}$$

Soit donc un système général $q = (q_1, q_2, q_3)$ et

$$M(q) = \begin{pmatrix} x_1(q) \\ x_2(q) \\ x_3(q) \end{pmatrix}. \tag{24.66}$$

Si nous fixons q_2 et q_3 et que nous laissons varier q_1, nous obtenons une courbe [4] dont nous pouvons considérer le vecteur vitesse, c'est-à-dire le vecteur tangent. En chaque point nous avons ainsi trois vecteurs

$$\frac{\partial M}{\partial q_i}(q). \tag{24.67}$$

Nous disons que le système de coordonnées curviligne est **orthogonal** si ces trois vecteurs sont orthogonaux. Dans la suite nous supposerons que c'est toujours le cas.

Nous posons

$$h_i = \left\| \frac{\partial M}{\partial q_i} \right\| \tag{24.68}$$

et nous considérons les trois vecteurs normés

$$e_i = h_i^{-1} \frac{\partial M}{\partial q_i}. \tag{24.69}$$ EqDefeihMq

Les trois vecteurs $\{e_1, e_2, e_3\}$ forment une base orthonormée dite **base locale**. Ce sont des vecteurs liés [5] au point M.

24.5.2 Importance de l'orthogonalité

Nous avons dit que nous nous restreignons au cas où les vecteurs e_i sont orthogonaux. En termes de produits scalaires, cela signifie

$$e_i \cdot e_j = \delta_{ij}. \tag{24.70}$$

Nous en étudions maintenant quelques conséquences. L'équation (24.69) peut s'écrire plus explicitement sous la forme

$$e_i = \sum_k h_i^{-1} \frac{\partial x_k}{\partial q_i} 1_k. \tag{24.71}$$

3. Nous n'entrons pas dans les détails de régularité.
4. Dans le cas des sphériques, c'est une demi-droite horizontale d'angle q_2 et de hauteur q_3.
5. En géométrie différentielle on dira que ce sont des élément de l'espace tangent, mais c'est une toute autre histoire.

Notez que pour chaque k et i, la quantité $h_i^{-1}\frac{\partial x_k}{\partial q_i}$ est un simple nombre. Nous allons les mettre dans une matrice :

$$A_{ki} = h_i^{-1}\frac{\partial x_k}{\partial q_i}. \tag{24.72}$$

Cela nous donne le changement de base

$$e_i = \sum_k A_{ki} 1_k. \tag{24.73}$$ EqChmBasecisAkiAk

Le produit $e_i \cdot e_j$ s'écrit alors

$$\begin{aligned} e_i \cdot e_j &= \sum_{kl} A_{ki}A_{lj}\underbrace{1_k \cdot 1_l}_{=\delta_{kl}} \\ &= \sum_{kl} A_{ki}A_{lj}\delta_{kl} \\ &= \sum_k A_{ki}A_{kj} \\ &= \sum_k (A^T)_{ik}A_{kj}. \end{aligned} \tag{24.74}$$

Or cela doit valoir δ_{ij}. Par conséquent

$$A^T = A^{-1}. \tag{24.75}$$

Le fait que les coordonnées curvilignes considérées soient orthogonales s'exprime donc par la fait que la matrice de changement de base est une matrice orthogonale.

Cette circonstance nous permet d'inverser le changement de base (24.73) en multipliant cette équation par $(A^{-1})_{il}$ des deux côtés et en faisant la somme sur i :

$$\sum_i (A^{-1})_{il}e_i = \sum_{kl} \underbrace{A_{ki}(A^{-1})_{il}}_{=\delta_{kl}} 1_k, \tag{24.76}$$

par conséquent

$$\sum_i (A^T)_{il}e_i = 1_l, \tag{24.77}$$

et

$$1_l = \sum_i A_{li}e_i = \sum_i h_i^{-1}\frac{\partial x_l}{\partial q_i}e_i. \tag{24.78}$$ EqChamvarunAei

Armés de cette importante formule, nous pouvons exprimer les quantités que nous connaissons dans la base canonique en termes de la base locale.

Une autre conséquence du fait que e_1, e_2 et e_3 est une base orthonormée est que, éventuellement en réordonnant les vecteurs, on a

$$\begin{aligned} e_1 \times e_2 &= e_3 \\ e_2 \times e_3 &= e_1 \\ e_3 \times e_1 &= e_2 \end{aligned} \tag{24.79}$$

Ces trois relations s'écrivent en une seule avec

$$e_i \times e_j = \sum_k \epsilon_{ijk}e_k \tag{24.80}$$

où

$$\epsilon_{ijk} = \begin{cases} 0 & \text{si } i,j,k \text{ ne sont pas tous différents} \\ 1 & \text{si } ijk \text{ se ramène à 123 par un nombre pair de transpositions} \\ -1 & \text{si } ijk \text{ se ramène à 123 par un nombre impair de transpositions} \end{cases} \tag{24.81}$$

est le **symbole de Levi-Civita**. La formule du produit vectoriel peut également être utilisée à l'envers sous la forme

$$e_k = \frac{1}{2}\sum_{ij} \epsilon_{ijk}\, e_i \times e_j. \tag{24.82}$$

Eqekeitimesej

Le symbole de Levi-Civita possède de nombreuses formules. En voici certaines, facilement démontrables en considérant tous les cas :

$$\epsilon_{ijk}\epsilon_{ijl} = \delta_{kl}|\epsilon_{ijk}|. \tag{24.83}$$

Grâce aux symboles de Levi-Civita, le produit mixte des vecteurs de base a une belle forme :

$$e_l \cdot (e_i \times e_j) = \sum_k \epsilon_{ijk} e_l \times e_k = \sum_k \epsilon_{ijk}\delta_{lk} = \epsilon_{ijl}. \tag{24.84}$$

EqProdMixteepsilonCicivr

24.5.3 Coordonnées polaires

Les coordonnées curvilignes polaires sont données par

$$M(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}, \tag{24.85}$$

et par conséquent

$$\frac{\partial M}{\partial r} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \frac{\partial M}{\partial \theta} = \begin{pmatrix} -r\sin(\theta) \\ r\cos(\theta) \end{pmatrix}. \tag{24.86}$$

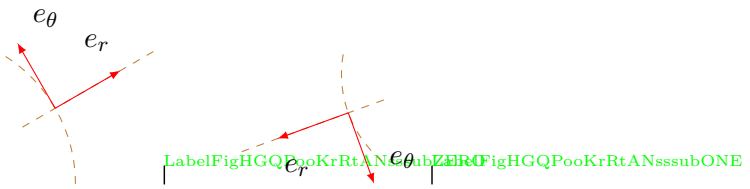
Nous avons les normes $h_r = 1$ et $h_\theta = r$, et donc les vecteurs de la base locale en (r,θ) sont

$$e_r = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \cos(\theta)e_x + \sin(\theta)e_y \tag{24.87}$$

ainsi que

$$e_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = -\sin(\theta)e_x + \cos(\theta)e_y. \tag{24.88}$$

Ces vecteurs sont représentés à la figure 24.4. Notez qu'il y en a une paire différente en chaque point.



(a) Base locale. (b) Base locale.

LabelFigHGQPooKrRtANsubZERO LabelFigHGQPooKrRtANsubONE

FIGURE 24.4: En brun, les lignes que le point suivrait si on ne variait qu'une coordonnées polaires à la fois. Les vecteurs rouges sont les vecteurs e_r et e_θ.

LabelFigHGQPooKrRtAN

24.5.4 Coordonnées cylindriques

Les coordonnées cylindriques sont les mêmes que les coordonnées polaires à part qu'il faut écrire

$$M(r,\theta,z) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ z \end{pmatrix}, \tag{24.89}$$

et nous avons le vecteur de base supplémentaire

$$e_z = \frac{\partial M}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{24.90}$$

parce que $h_z = 1$.

24.5.5 Coordonnées sphériques

Les coordonnées curvilignes sphériques sont données par

$$M(\rho,\theta,\varphi)=\begin{pmatrix}\rho\sin(\theta)\cos(\varphi)\\ \rho\sin(\theta)\sin(\varphi)\\ \rho\cos(\theta)\end{pmatrix}, \tag{24.91}$$

dont les dérivées sont données par

$$\begin{aligned}\frac{\partial M}{\partial\rho}&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\qquad \frac{\partial M}{\partial\theta}=\begin{pmatrix}\rho\cos(\theta)\cos(\varphi)\\ \rho\cos(\theta)\sin(\varphi)\\ -\rho\sin(\theta)\end{pmatrix},\\ \frac{\partial M}{\partial\varphi}&=\begin{pmatrix}-\rho\sin(\theta)\sin(\varphi)\\ \rho\sin(\theta)\cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{24.92}$$

Les normes de ces vecteurs sont $h_\rho=1$, $h_\theta=\rho$ et $h_\varphi=\rho\sin(\theta)$. Les vecteurs de la base locale en (ρ,θ,φ) sont donc

$$\begin{aligned}e_\rho&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\quad e_\theta=\begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix},\\ e_\varphi&=\begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{24.93}$$

24.5.6 Gradient en coordonnées curvilignes

Soit $(x,y,z)\mapsto f(x,y,z)$ une fonction sur $\mathbb{R}^3$. Nous pouvons la composer avec les coordonnées curvilignes q pour obtenir la fonction

$$\tilde{f}(q_1,q_2,q_3)=f\big(x_1(q),x_2(q),x_3(q)\big). \tag{24.94}$$

Nous disons que $\tilde{f}$ est l'expression de f dans les coordonnées q. Nous savons déjà comment calculer le gradient de f en coordonnées cartésiennes :

$$F(x,y,z)=\nabla f(x,y,z)=\begin{pmatrix}\partial_x f(x,y,z)\\ \partial_y f(x,y,z)\\ \partial_z f(x,y,z)\end{pmatrix}. \tag{24.95}$$

Cela est un vecteur lié au point (x,y,z). Notre objectif de bonheur dans la vie serait d'exprimer les coordonnées de ce vecteur dans la base $\{e_1,e_2,e_3\}$. En d'autres termes, nous voudrions trouver les nombres $\tilde{F}_1$, $\tilde{F}_2$ et $\tilde{F}_3$ tels que

$$F(x,y,z)=F\big(x(q),y(q),z(q)\big)=\tilde{F}_1e_1+\tilde{F}_2e_2+\tilde{F}_3e_3. \tag{24.96}$$

Ces nombres seront des fonctions de (q_1,q_2,q_3).

Par définition,

$$\nabla f=\sum_l\frac{\partial f}{\partial x_l}1_l. \tag{24.97}$$

En remplaçant 1_l par sa valeur en termes des e_i par la formule (24.78),

$$\begin{aligned}\nabla f &= \sum_l \frac{\partial f}{\partial x_l} 1_l \\ &= \sum_l \frac{\partial f}{\partial x_l} \sum_i h_i^{-1} \frac{\partial x_l}{\partial q_i} e_i \\ &= \sum_{il} \frac{1}{h_i} \frac{\partial f}{\partial x_l} \frac{\partial x_l}{\partial q_i} e_i \\ &= \sum_i \frac{1}{h_i} \frac{\partial \tilde{f}}{\partial q_i} e_i.\end{aligned} \tag{24.98}$$

Plus explicitement,

$$\nabla f\big(x(q), y(q), z(q)\big) = \sum_i \frac{1}{h_i(q)} \frac{\partial \tilde{f}}{\partial q_i}(q) e_i \tag{24.99}$$ EqGradientenCurviligne

où

$$h_i(q) = \left\| \frac{\partial M}{\partial q_i}(q) \right\|. \tag{24.100}$$

Le plus souvent nous n'allons pas noter explicitement la dépendance de h_i en q.

24.5.6.1 Coordonnées sphériques

Nous pouvons exprimer le gradient d'une fonction en coordonnées sphériques en utilisant la formule (24.99) :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho} \frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)} \frac{\partial \tilde{f}}{\partial \varphi} e_\varphi. \tag{24.101}$$ EqGradientSpherique

Cette expression peut paraitre peu pratique parce que les vecteurs e_ρ, e_θ et e_φ eux-mêmes changent en chaque point. Elle est effectivement peu adaptée au dessin, mais elle est très pratique pour des fonctions ayant des symétries.

Exemple 24.15.
Le potentiel de la gravitation est la fonction

$$V(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}. \tag{24.102}$$

En coordonnées sphériques elle s'écrit

$$\tilde{V}(\rho, \theta, \varphi) = \frac{1}{\rho}. \tag{24.103}$$

En voilà une fonction qu'elle est facile à dériver, contrairement à V ! En suivant la formule (24.101), nous avons immédiatement

$$\nabla \tilde{V} = -\frac{1}{\rho^2} e_\rho. \tag{24.104}$$

Nous voyons immédiatement que cela est un champ de vecteurs dont la norme diminue comme le carré de la distance à l'origine et qui est en permanence dirigé vers l'origine. △

24.5.7 Divergence en coordonnées curvilignes

Nous savons que

$$\nabla \tilde{f} = \sum_j \frac{1}{h_j} \frac{\partial \tilde{f}}{\partial q_j} e_j. \tag{24.105}$$

Nous pouvons en particulier considérer la fonction $f(q) = q_i$. De la même manière que nous avions noté x_i la fonction $x \mapsto x_i$, nous notons q_i la fonction $q \mapsto q_i$. Le gradient de cette fonction est donné par

$$\nabla q_i = \sum_j \frac{1}{h_j} \frac{\partial q_i}{\partial q_j} e_j, \tag{24.106}$$

mais $\frac{\partial q_i}{\partial q_j} = \delta_{ij}$, donc

$$\nabla q_i = \frac{e_i}{h_i}, \tag{24.107}$$

ou encore

$$e_i = h_i \nabla q_i. \tag{24.108}$$

Cela n'est pas étonnant : la direction dans laquelle la coordonnées q_i varie le plus est le vecteur e_i qui donne la tangente à la courbe obtenue lorsque *seul* q_i varie.

Commençons par calculer la divergence de e_i. En utilisant la formule (24.82),

$$\nabla \cdot e_k = \frac{1}{2} \sum_{ij} \epsilon_{ijk} \nabla \cdot (e_i \times e_j). \tag{24.109}$$

Nous avons, en utilisant les règles de Leibniz de la proposition 12.510,

$$\begin{aligned} \nabla \cdot (e_i \times e_j) &= \nabla \cdot (h_i \nabla q_i \times h_j \nabla q_j) \\ &= \nabla(h_i h_j) \cdot \big(\nabla q_i \times \nabla q_j\big) + h_i h_j \nabla \cdot \big(\nabla q_i \times \nabla q_j\big) \\ &= \nabla(h_i h_j) \cdot \big(\nabla q_i \times \nabla q_j\big) \\ &\quad + h_i h_j \nabla q_j \cdot \big(\underbrace{\nabla \times \nabla q_i}_{=0}\big) \\ &\quad + h_i h_j \nabla q_i \cdot \big(\underbrace{\nabla \times \nabla q_j}_{=0}\big) \end{aligned} \tag{24.110}$$

Cela nous fait

$$\nabla \cdot e_k = \sum_{ij} \epsilon_{ijk} \frac{\nabla(h_i h_j)}{h_i h_j} \cdot (e_i \times e_j). \tag{24.111}$$

parce que $\nabla q_i = h_i^{-1} e_i$. Nous pouvons développer le gradient qui intervient :

$$\nabla(h_i h_j) = \sum_l \frac{1}{h_l} \frac{\partial}{\partial q_l}(h_i h_j) e_l. \tag{24.112}$$

Nous voyons donc arriver le produit mixte $e_l \cdot (e_i \times e_j)$. En utilisant la formule (24.84), cela s'exprime directement sous la forme ϵ_{ijl}.

Nous avons alors

$$\begin{aligned} \nabla \cdot e_k &= \frac{1}{2} \sum_{ijl} \frac{1}{h_i h_j h_l} \frac{\partial}{\partial q_l}(h_i h_j) \epsilon_{ijk} \epsilon_{ijl} \\ &= \frac{1}{2} \sum_{ijl} \delta_{kl} |\epsilon_{ijk}| \frac{\partial}{\partial q_l}(h_i h_j) \\ &= \frac{1}{2} \sum_{ij} \frac{|\epsilon_{ijk}|}{h_i h_j h_k} \frac{\partial}{\partial q_k}(h_i h_j). \end{aligned} \tag{24.113}$$

EqFragradekdvj

Par exemple,

$$\nabla \cdot e_1 = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_1}(h_2 h_3). \tag{24.114}$$

Nous devons maintenant chercher le gradient d'un champ général

$$F(q) = \sum_k F_k(q) e_k. \tag{24.115}$$

La première chose à faire est d'utiliser la formule de Leibniz :

$$\nabla \cdot F = \sum_k \nabla F_k(q) \cdot e_k + \sum_k F_k(q) \nabla \cdot e_k. \tag{24.116}$$

Afin d'alléger les notations, nous allons nous concentrer sur le terme numéro k et ne pas écrire la somme. Si i et j sont les nombres tels que $\epsilon_{ijk} = 1$, alors ce que la formule (24.113) signifie, c'est que

$$\nabla \cdot e_k = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_k}(h_i h_j). \tag{24.117}$$

Nous savons déjà par la formule (24.99) que

$$\nabla F_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} e_l, \tag{24.118}$$

par conséquent

$$\nabla F_k \cdot e_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} \delta_{kl} = \frac{1}{h_k} \frac{\partial F_k}{\partial q_k}. \tag{24.119}$$

Pour obtenir cela nous avons utilisé le fait que $e_l \cdot e_k = \delta_{lk}$. Le terme numéro k de la somme (24.116) est donc

$$\frac{1}{h_k} \frac{\partial F_k}{\partial q_k} + \frac{F_k}{h_k h_i h_j} \frac{\partial (h_i h_j)}{\partial q_k} = \frac{1}{h_i h_j h_k} \frac{\partial (F_k h_i h_j)}{\partial q_k} \tag{24.120}$$

où il est entendu que i et j représentent les nombres tels que $\epsilon_{ijk} = 1$.

Au final, nous avons

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \sum_{ijk} |\epsilon_{ijk}| \frac{\partial (F_k h_i h_j)}{\partial q_k}. \tag{24.121}$$

Ici, la somme sur i et j consiste seulement à sélectionner les termes tels que i et j ne sont pas k. En écrivant la somme explicitement,

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}(F_1 h_2 h_3) + \frac{\partial}{\partial q_2}(F_2 h_1 h_3) + \frac{\partial}{\partial q_3}(F_3 h_1 h_2) \right]. \tag{24.122}$$

24.5.7.1 Coordonnées cylindriques

En coordonnées cylindriques, nous avons déjà vu que $h_r = 1$, $h_\theta = r$ et $h_z = 1$. La divergence est donc donnée par

$$\nabla \cdot F = \frac{1}{r} \left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(r F_z) \right]. \tag{24.123}$$

Par exemple si

$$F(r, \theta, z) = r e_\theta + e_z, \tag{24.124}$$

nous avons

$$(\nabla \cdot F)(r, \theta, z) = \frac{1}{r} \left[\frac{\partial}{\partial \theta}(r) + \frac{\partial}{\partial z}(r) \right] = 0. \tag{24.125}$$

Cela est logique parce que $r e_\theta$ est à peu près le champ dont nous avons parlé dans l'exemple (12.507), qui était à divergence nulle. En réalité, le champ dont on parlait dans cet exemple était exactement $-e_\theta$. Le champ e_z est également à divergence nulle parce qu'il est constant.

24.5.7.2 Coordonnées sphériques

En coordonnées sphériques, nous avons $h_\rho = 1$, $h_\theta = r$ et $h_\varphi = r \sin\theta$, donc

$$\nabla \cdot F = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi) \right]. \tag{24.126}$$

si $F(\rho, \theta, \varphi) = F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi$.

24.5.8 Laplacien en coordonnées curvilignes orthogonales

Soit une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$. Le **laplacien** de f est donné par

$$\Delta f = \nabla \cdot (\nabla f). \tag{24.127}$$

En utilisant les formules données, nous avons

$$\Delta f = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_1h_3}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right]. \tag{24.128}$$

Dans cette expression, la fonction f est donnée comme fonction de q_1, q_2 et q_3.

En coordonnées cylindriques, cela s'écrit

$$\begin{aligned}\Delta f &= \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{1}{r}\frac{\partial f}{\partial \theta}\right) + \frac{\partial}{\partial z}\left(r\frac{\partial f}{\partial z}\right)\right]\\ &= \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}.\end{aligned} \tag{24.129}$$

Dans cette expression, f est fonction de r, θ et z.

En coordonnées sphériques, cela devient

$$\Delta f = \frac{1}{\rho^2 \sin\theta}\left[\frac{\partial}{\partial \rho}\left(\rho^2 \sin\theta \frac{\partial f}{\partial \rho}\right) + \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial f}{\partial \theta}\right) + \frac{\partial}{\partial \varphi}\left(\frac{1}{\sin\theta}\frac{\partial f}{\partial \varphi}\right)\right]. \tag{24.130}$$ EqLaplaceSphe

Dans cette expression, f est fonction de ρ, θ et φ.

24.5.9 Rotationnel en coordonnées curvilignes orthogonales

Nous voulons calculer le rotationnel de $F(q) = \sum_k F_k(q)e_k$. Pour cela nous commençons par écrire $e_k = h_k \nabla q_k$ et nous utilisons la formule de la proposition 12.510(3) avec $F_k h_k$ en guise de f :

$$\begin{aligned}\nabla \times F_k e_k &= \nabla \times (F_k h_k \nabla q_k)\\ &= F_k h_k \underbrace{\nabla \times (\nabla q_k)}_{=0} + \nabla(F_k h_k) \times \nabla q_k\\ &= \frac{1}{h_k}\nabla(F_k h_k) \times e_k.\end{aligned} \tag{24.131}$$

Nous utilisons à présent la formule (24.99) du gradient et le formule $e_j \times e_k = \sum_l \epsilon_{jkl} e_l$:

$$\begin{aligned}\nabla \times (F_k e_k) &= \sum_j \frac{1}{h_j h_k}\frac{\partial}{\partial q_j}(F_k h_k) e_j \times e_k\\ &= \sum_{jl} \frac{1}{h_j h_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_k h_k) e_l.\end{aligned} \tag{24.132}$$

Le rotationnel s'écrit donc

$$\nabla \times F = \sum_{jkl} \frac{1}{h_j h_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_k h_k) e_l. \tag{24.133}$$

Devant e_1 par exemple nous avons seulement les termes $j = 2$, $k = 3$ et $j = 3$, $k = 2$. Étant donné que $\epsilon_{231} = 1$ et $\epsilon_{321} = -1$, le coefficient de e_1 sera simplement

$$\frac{1}{h_2 h_3}\left(\frac{\partial}{\partial q_2}(F_3 h_3) - \frac{\partial}{\partial q_3}(F_2 h_2)\right). \tag{24.134}$$

La formule complète devient

$$\begin{aligned}\nabla \times \sum_k F_k e_k = &\frac{1}{h_2 h_3}\left(\frac{\partial}{\partial q_2}(F_3 h_3) - \frac{\partial}{\partial q_3}(F_2 h_2)\right) e_1\\ &+ \frac{1}{h_1 h_3}\left(\frac{\partial}{\partial q_3}(F_1 h_1) - \frac{\partial}{\partial q_1}(F_3 h_3)\right) e_2\\ &+ \frac{1}{h_2 h_1}\left(\frac{\partial}{\partial q_1}(F_2 h_2) - \frac{\partial}{\partial q_2}(F_1 h_1)\right) e_3.\end{aligned} \tag{24.135}$$

24.5.9.1 Coordonnées cylindriques

En utilisant $h_r = 1$, $h_\theta = r$ et $h_z = 1$, nous trouvons

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{24.136}$$

EqRotationnelCylin

24.5.9.2 Coordonnées sphériques

En utilisant $h_\rho = 1$, $h_\theta = \rho$ et $h_\varphi = \rho \sin\theta$, nous trouvons

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi) \sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{24.137}$$

Note : dans le premier terme, il y a une simplification par ρ.

24.6 Les formules

24.6.1 Coordonnées polaires

Les vecteurs de base :

$$e_r = \begin{pmatrix}\cos(\theta) \\ \sin(\theta)\end{pmatrix} = \cos(\theta) e_x + \sin(\theta) e_y \tag{24.138a}$$

$$e_\theta = \begin{pmatrix}-\sin(\theta) \\ \cos(\theta)\end{pmatrix} = -\sin(\theta) e_x + \cos(\theta) e_y. \tag{24.138b}$$

Le gradient :

$$\nabla \tilde{f}(r,\theta) = \frac{\partial \tilde{f}}{\partial r}(r,\theta) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta) e_\theta. \tag{24.139}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta)\right]. \tag{24.140}$$

EqqRxJKd

Le rotationnel :

$$\nabla \times (F_r e_r + F_\theta e_\theta) = \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z. \tag{24.141}$$

EqtBnqCw

Notons que le rotationnel n'existe pas vraiment en deux dimensions. Ici nous avons vu le champ $F(r,\theta)$ comme un champ dans $\mathbb{R}^3$ ne dépendant pas de z et n'ayant pas de composante z. Le résultat est un rotationnel qui est dirigé selon l'axe z.

24.6.2 Coordonnées cylindriques

Les vecteurs de base : idem qu'en coordonnées polaires, et on ajoute e_z sans modifications.
Le gradient :

$$\nabla \tilde{f}(r,\theta,z) = \frac{\partial \tilde{f}}{\partial r}(r,\theta,z) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta,z) e_\theta + \frac{\partial \tilde{f}}{\partial z}(r,\theta,z) e_z. \tag{24.142}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(rF_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(rF_z)\right]. \tag{24.143}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{24.144}$$

Note : les formules concernant les coordonnées polaires se réduisent de celles-ci en enlevant toutes les références à z.

24.6.3 Coordonnées sphériques

Les vecteurs de base :

$$\begin{aligned} e_r &= \begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix}, \quad e_\theta = \begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix}, \\ e_\varphi &= \begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{24.145}$$

Le gradient :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)}\frac{\partial \tilde{f}}{\partial \varphi} r_\varphi. \tag{24.146}$$

La divergence :

$$\nabla \cdot F = \frac{1}{\rho^2 \sin\theta}\left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi)\right]. \tag{24.147}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi)\sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_\rho}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{24.148}$$

Chapitre 25

Espaces de Hilbert

25.1 Espaces de Hilbert

DEFooKSDFooGIBtrG

Définition 25.1.
*Soit un espace vectoriel E sur le corps $\mathbb{K}$. Son **dual topologique**, noté E', est l'ensemble des formes linéaires continues de E vers $\mathbb{K}$.*

PROPooQFTSooPFfbCc

Proposition 25.2 ([622]).
Soient des espaces vectoriels normés X et Y ainsi qu'une forme sesquilinéaire $\phi\colon X \times Y \to \mathbb{C}$. L'espace $X \times Y$ est muni de la topologie produit[1]*. Il y a équivalence des faits suivants.*

ITEMooWMFBooOWPzgP

(1) ϕ *est continue.*

ITEMooSFQZooHNcWwH

(2) ϕ *est continue en* $(0,0)$

ITEMooWSWQooSYzOol

(3) $\sup\{|\phi(x,y)| \text{ tel que } \|x\|, \|y\| \leqslant 1\} < \infty$.

ITEMooRKHQooMVXlYD

(4) Il existe $C \geqslant 0$ *telle que*

$$|\phi(x,y)| \leqslant C\|x\|\|y\| \tag{25.1}$$

pour tout $(x,y) \in X \times Y$.

ITEMooLWDCooLIGVDs

(5) ϕ *est bornée*

Dans le cas où ϕ vérifie ces propriétés, alors nous posons

$$\mathcal{A} = \{|\phi(x,y)| \text{ tel que } \|(x,y)\| \leqslant 1\} \tag{25.2}$$

et

$$\mathcal{B} = \{\beta \geqslant 0 \text{ tel que } |\phi(x,y)| \leqslant \beta\|x\|\|y\|, \forall x, y\} \tag{25.3}$$

et nous avons

$$\|\phi\| = \sup(\mathcal{A}) = \min(\mathcal{B}). \tag{25.4}$$

Démonstration. En plusieurs parties.

(1) **(1) ⇒ (2)** Une application est continue si et seulement si elle est continue en tout point (théorème 7.205).

(2) **(2) ⇒ (3)** Pour rappel, nous avons mis sur $X \times Y$ la topologie produit. Le lemme 7.227 nous indique que cette topologie est la même que la topologie de la norme produit.
Bref, nous supposons que ϕ est continue en $(0,0)$ et nous prouvons que

$$\sup\{|\phi(x,y)| \text{ tel que } \|x\|, \|y\| \leqslant 1\} < \infty. \tag{25.5}$$

Nous y allons par contradiction. Supposons que pour n dans $\mathbb{N}$ il existe $(x_n, y_n) \in X \times Y$ tels que $\|x_n\|, \|y_n\| \leqslant 1$ et $|\phi(x_n, y_n) > n^2|$. En posant $v_n = x_n/n$ et $w_n = y_n/n$, nous avons

$$\|v_n\| \leqslant \frac{\|x_n\|}{n} \leqslant \frac{1}{n}, \tag{25.6}$$

1. Quand on ne précise pas, un produit est toujours muni de la topologie produit 7.15.

et la même chose pour $\|w_n\|$. Nous avons $(v_n, w_n) \xrightarrow{X\times Y} (0,0)$ et donc, par continuité de ϕ en $(0,0)$ nous avons $\phi(v_n, w_n) \xrightarrow{\mathbb{C}} \phi(0,0) = 0$. Mais

$$|\phi(v_n, w_n)| = \frac{1}{n^2}|\phi(x_n, y_n)| \geqslant \frac{1}{n^2}n^2 = 1. \tag{25.7}$$

Donc nous n'avons pas $|\phi(v_n, w_n)| \to 0$.

(3) **(3) ⇒ (4)** Nous notons

$$M = \sup\{|\phi(x,y)| \text{ tel que } \|x\|, \|y\| \leqslant 1\} < \infty, \tag{25.8}$$

et nous allons prouver que $|\phi(x,y)| \leqslant M\|x\|\|y\|$.
Si $x = 0$ ou si $y = 0$, c'est évident ; nous supposons donc que x et y sont tous deux non nuls. Nous posons $v = x/\|x\|$ et $w = y/\|y\|$. Nous avons alors

$$|\phi(x,y)| = \|x\|\|y\||\phi(v,w)| \leqslant \|x\|\|y\|M. \tag{25.9}$$

(4) **(4) ⇒ (5)** La norme de ϕ est le nombre

$$\|\phi\| = \sup_{\substack{(x,y)\in X\times Y\\ \|(x,y)\|=1}} |\phi(x,y)|, \tag{25.10}$$

que nous devons prouver être fini. Les éléments (x,y) sur lesquels porte le supremum vérifient

$$1 = \|(x,y)\| = \max\{\|x\|, \|y\|\}, \tag{25.11}$$

et donc vérifient $\|x\|, \|y\| \leqslant 1$. En utilisant (4), nous avons donc

$$\|\phi\| = \sup_{\|(x,y)\|=1} |\phi(x,y)| \leqslant \sup_{\|(x,y)\|=1} C\|x\|\|y\| \leqslant C. \tag{25.12}$$

(5) **(5) ⇒ (1)** Il nous reste à prouver que si ϕ est bornée, alors elle est continue. Soit une suite convergente $(x_n, y_n) \xrightarrow{X\times Y} (x,y)$. Nous allons prouver que $\phi(x_n, y_n) \xrightarrow{\mathbb{C}} \phi(x,y)$. Ce sont les majorations usuelles :

$$\begin{aligned}
|\phi(x_n, y_n) - \phi(x,y)| &\leqslant |\phi(x_n, y_n) - \phi(x_n, y)| + |\phi(x_n, y) - \phi(x,y)| && \text{(25.13a)}\\
&= |\phi(x_n, y_n - y)| + |\phi(x_n - x, y)| && \text{(25.13b)}\\
&\leqslant \|\phi\|\|x_n\|\|y_n - y\| + \|\phi\|\|x_n - x\|\|y\| && \text{(25.13c)}\\
&\to 0. && \text{(25.13d)}
\end{aligned}$$

Enfin nous supposons que toutes ces conditions équivalentes, et nous prouvons que $\|\phi\| = \sup(\mathcal{A}) = \min(\mathcal{B})$.

(1) $\|\phi\| = \sup(\mathcal{A})$ En posant

$$\mathcal{C} = \{|\phi(x,y)| \text{ tel que } \|(x,y)\| = 1\}, \tag{25.14}$$

nous avons, par définition, $\|\phi\| = \sup(\mathcal{C})$. Étant donné que $\mathcal{C} \subset \mathcal{A}$, nous avons $\sup(\mathcal{C}) \leqslant \sup(\mathcal{A})$. Faisons l'inégalité dans l'autre sens. Soit $|\phi(x,y)| \in \mathcal{A}$, c'est-à-dire $\|(x,y)\| \leqslant 1$. En posant $v = x/\|x\|$ et $w = y/\|y\|$ nous avons $|\phi(v,w)| \in \mathcal{C}$ et surtout $\|x\|, \|y\| \leqslant 1$. Donc

$$|\phi(v,w)| = \frac{1}{\|x\|\|y\|}|\phi(x,y)| \geqslant |\phi(x,y)|. \tag{25.15}$$

Donc tout élément de $\mathcal{C}$ est majoré par un élément de $\mathcal{A}$. Donc $\sup(\mathcal{A}) \geqslant \sup(\mathcal{C})$.

(2) $\|\phi\| \in \mathcal{B}$ Soit $(x,y) \in X \times Y$. Étant donné que $\|x/\|x\|\| = 1$ nous avons

$$\left|\phi\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right)\right| \leqslant \sup_{\|(v,w)\|=1} |\phi(v,w)| = \|\phi\|. \tag{25.16}$$

En multipliant les deux côtés par $\|x\|\|y\|$ nous trouvons

$$|\phi(x,y)| \leqslant \|\phi\|\|x\|\|y\|, \tag{25.17}$$

et donc $\|\phi\| \in \mathcal{B}$.

(3) $\|\phi\|$ **minore** $\mathcal{B}$
Nous montrons que pour tout $\alpha \in \mathcal{A}$ et pour tout $\beta \in \mathcal{B}$, nous avons $\alpha \leqslant \beta$. En effet si $\alpha \in \mathcal{A}$, alors $\alpha = |\phi(x,y)|$ pour un certain $(x,y) \in X \times Y$ vérifiant $\|(x,y)\| = 1$. Mais $\beta \in \mathcal{B}$ signifie que $|\phi(x,y)| \leqslant \beta\|x\|\|y\|$. Nous avons donc

$$\alpha = |\phi(x,y)| \leqslant \beta\|x\|\|y\| \leqslant \beta \tag{25.18}$$

parce que $\|x\|, \|y\| \leqslant 1$.

□

DefORuBdBN

Définition 25.3.
*Un espace **préhilbertien** est*

- *soit un espace vectoriel réel muni d'un produit scalaire*[2],
- *soit un espace vectoriel complexe muni d'un produit scalaire hermitien*[3].

*Un **espace de Hilbert** est un espace préhilbertien qui est complet*[4] *pour la norme induite par son produit (scalaire ou hermitien), c'est-à-dire tel que*

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{25.19}$$

pour tout élément x.

Dans les deux cas nous considérons la topologie métrique dérivant du produit.

Dans les cas de dimension finie, les espaces vectoriels normés sont automatiquement complets par la proposition 7.293.

La différence entre un espace de Hilbert et un espace de Banach est que dans le cas d'un espace de Hilbert, nous demandons que la norme dérive d'un produit scalaire.

PropTdupIG

Proposition 25.4.
Si $\mathscr{H}$ est un espace de Hilbert réel, alors

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle. \tag{25.20}$$

Si $\mathscr{H}$ est un espace de Hilbert complexe alors

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle. \tag{25.21}$$

EqrbBlkK

Dans les deux cas nous avons l'inégalité de Cauchy-Schwarz :

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|. \tag{25.22}$$

Dans un espace vectoriel de dimension infinie, tous les opérateurs linéaires ne sont pas continus.

2. Définition 8.168.
3. Définition 8.176.
4. Définition 7.261.

Exemple 25.5.
Soit un espace vectoriel V engendré par la base $\{e_k\}_{k\in\mathbb{N}}$ et l'application linéaire $T: V \to V$ donnée par

$$Te_k = ke_k. \tag{25.23}$$

Nous allons montrer que l'image inverse de la boule unité ouverte $\mathcal{O}$ n'est pas ouverte. En effet si (ϵ_k) est une suite de réels strictement positifs tendant vers zéro, les vecteurs

$$a_k = \left(\frac{1}{k} + \epsilon_k\right) e_k \tag{25.24}$$

sont hors de $T^{-1}\mathcal{O}$ parce que

$$Ta_k = (1 + k\epsilon_k)e_k. \tag{25.25}$$

Mais la suite (a_k) converge vers 0 qui fait partie de $T^{-1}\mathcal{O}$. Donc le complémentaire de $T^{-1}\mathcal{O}$ n'est pas fermé, ce qui prouve que $T^{-1}\mathcal{O}$ n'est pas ouvert. △

PROPooMSAYooONHLYq

Proposition 25.6 ([1]).
Soit un espace de Hilbert $\mathscr{H}$ et un espace vectoriel normé $(E, \|.\|)$ tel qu'il existe une bijection linéaire isométrique $\Phi: \mathscr{H} \to E$. Alors en posant

$$\langle \alpha, \beta \rangle_E = \langle \Phi^{-1}(\alpha), \Phi^{-1}(\beta) \rangle_{\mathscr{H}}, \tag{25.26}$$

l'espace E devient un espace de Hilbert.

Démonstration. Le fait que $\langle ., . \rangle_E$ soit un produit scalaire (hermitien) lorsque $\langle ., . \rangle_{\mathscr{H}}$ l'est est une conséquence de la linéarité de Φ. Il donne la norme parce que

$$\sqrt{\langle \alpha, \alpha \rangle} = \sqrt{\langle \Phi^{-1}(\alpha), \Phi^{-1}(\alpha) \rangle} = \|\Phi^{-1}(\alpha)\| = \|\alpha\| \tag{25.27}$$

parce que Φ^{-1} est une isométrie.

Pour voir que E est complet, soit une suite (α_k) de τ-Cauchy dans E. Nous devons prouver qu'elle converge. Nous posons $x_k = \Phi^{-1}(\alpha_k)$.

Commençons par voir que (x_k) est τ-Cauchy dans $\mathscr{H}$. Soit un voisinage V de 0 dans $\mathscr{H}$. Nous posons $U = \Phi(V)$. Vu que (α_k) est τ-Cauchy, il existe $N > 0$ tel que $\alpha_k - \alpha_l \in U$ dès que $k, l > N$. Du coup, si $k, l > N$ nous avons aussi $x_k - x_l = \Phi^{-1}(\alpha_k - \alpha_l) \subset \Phi^{-1}(U) = V$. Cela prouve que (x_k) est τ-Cauchy.

Vu que $\mathscr{H}$ est complet, il existe $x \in \mathscr{H}$ tel que $x_k \xrightarrow{\mathscr{H}} x$. Montrons que $\alpha_k \xrightarrow{E} \Phi(x)$.

Soit $\epsilon > 0$ et $N > 0$ tel que $\|x_k - x\| < \epsilon$ dès que $k \geqslant N$. Alors nous avons aussi

$$\|\alpha_k - \Phi(x)\| = \|\Phi(x_k) - \Phi(x)\| = \|\Phi(x_k - x)\| = \|x_k - x\| < \epsilon \tag{25.28}$$

où nous avons utilisé le fait que Φ était linéaire et une isométrie. □

25.1.1 Sous-espace vectoriel fermé ? ? ?

Nous verrons que beaucoup de résultats demandent un sous-espace vectoriel *fermé*. Une question légitime est : est-ce qu'il existe des sous-espaces vectoriels qui ne soient pas fermés ? En dimension finie, tous les sous-espaces vectoriels sont fermés, mais cela n'est pas vrai en dimension infinie.

Soit en effet une partie libre infinie $A = \{v_i\}_{i\in\mathbb{N}}$ dans un espace de Hilbert $\mathscr{H}$. L'ensemble $\mathrm{Span}(A)$ des combinaisons linéaires d'éléments de A est un sous-espace vectoriel de $\mathscr{H}$, mais il n'est pas fermé.

En effet, supposons pour simplifier les notations que $\|v_i\| = 1$ pour tout i. Alors nous considérons la combinaison

$$a = \sum_{k=1}^{\infty} \alpha_k v_k \tag{25.29}$$

où les α_k sont suffisamment décroissants pour assurer les convergences[5]. Le vecteur a n'est pas dans $\text{Span}(A)$, mais la suite $a_n = \sum_{k=1}^{n} \alpha_k v_k$ est dans $\text{Span}(A)$ et converge vers a : $a_n \xrightarrow{\mathscr{H}} a$. En effet,

$$\|a_n - a\| = \| \sum_{k=n+1}^{\infty} \alpha_k v_k \| \leqslant \sum_{k=n+1}^{\infty} |\alpha_k| \xrightarrow{n\to\infty} 0. \tag{25.30}$$

Vous voulez des détails sur la dernière limite ? Vu que la somme $\sum_k |\alpha_k|$ converge, la suite des sommes de queues de suites[6] converge vers zéro :

$$\lim_{n\to\infty} \sum_{k=n+1}^{\infty} |\alpha_k| = 0. \tag{25.31}$$

25.2 Théorème de la projection

Voir la version plus simple dans 18.103.

ThoProjOrthuzcYkz

Théorème 25.7 (Projection sur partie fermée convexe[623, 624])**.**
Soit $\mathscr{H}$ un espace de Hilbert, $x \in \mathscr{H}$, et C un sous-ensemble fermé convexe de $\mathscr{H}$.

(1) Les deux conditions suivantes sur $y \in \mathscr{H}$ sont équivalentes : ETsfYCSItemi

(1a) $\|x - y\| = \inf\{\|x - z\| \text{ tel que } z \in C\}$, ETsfYCSItemii

(1b) pour tout $z \in C$, $\text{Re}\langle x - y, z - y\rangle \leqslant 0$.

(2) Il existe un unique $y \in \mathscr{H}$, noté $y = \texttt{proj}_C(x)$ vérifiant ces conditions.

Démonstration. Nous commençons par prouver l'existence et l'unicité d'un élément dans C vérifiant la première condition. Ensuite nous verrons l'équivalence.

Nous nommons d l'infimum en question de la première condition.

Existence Soit (y_n) une suite dans C telle que

$$\lim_{n\to\infty} \|x - y_n\| = \inf\{\|z - y\| \text{ tel que } z \in C\} = d. \tag{25.32}$$

Nous allons montrer que cette suite peut être choisie de Cauchy. Elle convergera donc dans $\mathscr{H}$ parce que ce dernier est complet. Mais C étant supposé fermé dans $\mathscr{H}$, la limite appartiendra à C. Soient $r, s \in \mathbb{N}$. D'abord nous avons

$$\|y_r - y_s\|^2 = \langle y_r - y_s + x - x, y_r - y_s + x - x\rangle \tag{25.33a}$$
$$= \|y_r - x\|^2 + \|y_s - x\|^2 - 2\langle y_r - x, y_s - x\rangle. \tag{25.33b}$$

Ensuite,

$$4\left\|\frac{y_r + y_s}{2} - x\right\|^2 = \langle y_r + y_s - 2x, y_r - y_s - 2x\rangle \tag{25.34a}$$
$$= \|y_r - x\|^2 + \|y_s - x\|^2 + 2\langle y_r - x, y_s - x\rangle. \tag{25.34b}$$

Si nous égalisons les valeurs de $2\langle y_r - x, y_s - x\rangle$ nous trouvons

$$\|y_r - y_s\|^2 = -4\left\|\frac{y_r + y_s}{2} - x\right\|^2 + 2\|y_r - x\|^2 + 2\|y_s - x\|^2. \tag{25.35}$$ EqiqCxUa

La distance infimum étant d, nous pouvons choisir y_n de telle façon à avoir

$$\|y_n - x\| \leqslant d + \frac{1}{n}. \tag{25.36}$$

5. Par exemple $\alpha_k = 1/k^2$ si les v_k sont orthonormaux.
6. On se comprend hein.

D'autre part étant donné que C est convexe, $(y_r + y_s)/2$ est dans C et nous avons

$$\left\| \frac{y_r + y_s}{2} - x \right\| \geqslant d. \tag{25.37}$$

En mettant ces majorations dans (25.35) nous trouvons

$$\|y_r - y_s\|^2 \leqslant -4d + 2\left(d + \frac{1}{r}\right) + 2\left(d + \frac{1}{s}\right) = \frac{1}{r} + \frac{1}{s}. \tag{25.38}$$

La suite (y_n) est donc de Cauchy et la limite est un élément de C. Prouvons que cet élément y réalise l'infimum. Pour cela nous avons les inégalités

$$d \leqslant \|x - y\| \leqslant \|x - y_n\| + \|y_n - y\|. \tag{25.39}$$

En prenant le limite $n \to \infty$ nous trouvons

$$d \leqslant \|x - y\| \leqslant d. \tag{25.40}$$

Unicité Même preuve que pour le théorème en dimension finie 12.142.

(1a)⇒ (1b) Même preuve que pour le théorème en dimension finie 12.142.

(1b)⇒ (1a) Même preuve que pour le théorème en dimension finie 12.142.

□

PropAXJpCe

Proposition 25.8.
Soit C une partie convexe et fermée de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x_1, x_2 \in \mathscr{H}$ nous avons

$$\| \texttt{proj}_C(x_1) - \texttt{proj}_C(x_2)\| \leqslant \|x_1 - x_2\|. \tag{25.41}$$

En particulier la projection est une application continue.

Démonstration. Nous posons $y_1 = \texttt{proj}_C(x_1)$ et $y_2 = \texttt{proj}_C(x_2)$. Par la partie (1b) du théorème 25.7, nous avons, pour tout $z, z' \in C$ les inégalités

$$\mathrm{Re}\langle x_1 - y_1, z - y_1\rangle \leqslant 0 \tag{25.42a}$$
$$\mathrm{Re}\langle x_2 - y_2, z' - y_2\rangle \leqslant 0. \tag{25.42b}$$

En prenant $z = y_2$ et $z' = y_1$ et en sommant nous trouvons

$$\mathrm{Re}\langle (x_1 - y_1) + (y_2 - x_2), y_2 - y_1\rangle \leqslant 0. \tag{25.43}$$

Nous pouvons maintenant calculer

$$\begin{aligned}
\|y_1 - y_2\|^2 &= \mathrm{Re}\, \|y_1 - y_2\| \\
&= \mathrm{Re}\langle y_1 - y_2, (y_1 - x_1) + x_1 - x_2 + (x_2 - y_2)\rangle \\
&= \mathrm{Re}\langle y_1 - y_2, x_1 - x_2\rangle + \mathrm{Re}\langle y_2 - y_1, (x_1 - y_1) + (y_2 - x_2)\rangle \\
&\leqslant \langle y_1 - y_2, x_1 - x_2\rangle \\
&\leqslant \|\langle y_1 - y_2, x_1 - x_2\rangle\| \\
&\leqslant \|y_1 - y_2\|\|x_1 - x_2\|.
\end{aligned} \tag{25.44}$$

En simplifiant par $\|y_1 - y_2\|$ [7] nous trouvons le résultat

$$\|y_1 - y_2\| \leqslant \|x_1 - x_2\|. \tag{25.45}$$

□

7. Si c'est nul, alors la preuve est évidente.

ThoMXwOjb

Théorème 25.9 (Projection orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert et K, un sous-espace vectoriel fermé non réduit à $\{0\}$ et $x \in \mathscr{H}$. L'élément $y = \texttt{proj}_K(x)$ est l'unique élément de K tel que

$$x - y \in K^\perp. \tag{25.46}$$

De plus l'application $x \mapsto \texttt{proj}_K(x)$ est linéaire, continue et de norme 1.

L'élément y ainsi défini est la **projection orthogonale** de x sur K et sera noté $\texttt{proj}_K(f)$.

Démonstration. La continuité de la projection est donnée par la proposition 25.8.

Soit z un élément de K tel que $\langle z - x, a\rangle = 0$ pour tout $a \in K$. Nous avons

$$\|x - a\|^2 = \|z - x\|^2 + \|a - z\|^2 + 2\underbrace{\langle z - x, a - z\rangle}_{=0} \tag{25.47a}$$

$$\geqslant \|z - x\|^2. \tag{25.47b}$$

Le produit scalaire est nul parce que $a - z \in K$. La distance $\|z - x\|$ est donc bien la plus petite distance entre x et les éléments de K.

Dans l'autre sens, nous supposons que $y \in K$ minimise la distance à x dans K. Par hypothèse pour tout a et pour tout $\lambda \in \mathbb{R}$, la différence

$$\|(y + \lambda a) - x\|^2 - \|y - x\|^2 \tag{25.48}$$

est positive. En développant les produits scalaires nous trouvons la conditions suivante

$$\lambda^2\|a\|^2 + 2\lambda\langle a, y - x\rangle \geqslant 0 \tag{25.49}$$

qui doit être vraie pour tout $\lambda \in \mathbb{R}$. En tant que polynôme du second degré en λ, cela n'aura pas deux racines réelles distinctes uniquement si $\langle a, y - x\rangle = 0$.

Nous montrons maintenant la linéarité de la projection orthogonale. Soient $x_1, x_2 \in \mathscr{H}$. L'élément $y = \texttt{proj}_K x_1 + \texttt{proj}_K x_2$ satisfait à la condition d'orthogonalité : pour tout $z \in K$,

$$\langle x_1 + x_2 - \texttt{proj}_K x_1 - \texttt{proj}_K x_2, z\rangle = \langle x_1 - \texttt{proj}_K x_1, z\rangle + \langle x_2 - \texttt{proj}_K x_2, z\rangle = 0. \tag{25.50}$$

Étant donné que K est un sous-espace vectoriel, la condition de minimalité est automatiquement vérifiée (seconde partie du théorème 25.7).

En ce qui concerne la norme opérateur de $\texttt{proj}_K$, la décomposition de $x \in \mathscr{H}$ en composantes dans K et $K^\perp$ est

$$x = x + (x - \texttt{proj}_K x). \tag{25.51}$$

Étant deux parties orthogonales nous avons

$$\|\texttt{proj}_K x\|^2 = \|x\|^2 - \|x - \texttt{proj}_K x\|^2. \tag{25.52}$$

En prenant $\|x\| = 1$ nous trouvons $\|\texttt{proj}_K x\|^2 \leqslant 1$ et par conséquent $\|\texttt{proj}_K\| \leqslant 1$. Mais d'autre part en prenant $x \in K$ nous avons automatiquement $\|\texttt{proj}_K\| \geqslant 1$. □

Proposition 25.10.
Soit $\mathscr{H} = L^2(\Omega, \mathcal{A}, \mu)$, $\mathcal{F}$ une sous tribu de $\mathcal{A}$ et K l'ensemble de fonctions $\mathcal{F}$-mesurables dans $L^2(\Omega, \mathcal{A}, \mu)$. Si $f \in L^2(\Omega, \mathcal{A}, \mu)$ est positive, alors $\texttt{proj}_K f$ est positive (presque partout).

Démonstration. L'ensemble $A = \{\texttt{proj}_K f < 0\}$ est dans $\mathcal{F}$. En effet

$$A = (\texttt{proj}_K f)^{-1}\big(]-\infty, 0[\big) \tag{25.53}$$

alors que, par construction, $\texttt{proj}_K f$ est $\mathcal{F}$-mesurable. La fonction indicatrice $\mathbb{1}_A$ est alors $\mathcal{F}$-mesurable (c'est-à-dire $1_A \in K$) et nous avons

$$0 \leqslant \int_\Omega f\mathbb{1}_A = \int_\Omega \texttt{proj}_K f\mathbb{1}_A \leqslant 0. \tag{25.54}$$

Étant donné que nous avons supposé $f \geqslant 0$ nous avons alors $\mu(A) = 0$. D'où le fait que $\texttt{proj}_K f$ est presque partout positive. □

25.3 Systèmes orthogonaux et bases

Dans cette partie nous noterons $\mathbb{K}$ le corps de base de l'espace $\mathscr{H}$. Seuls deux cas sont envisagés : $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$.

Pour chaque $x \in \mathscr{H}$ nous considérons l'application

$$\begin{aligned} \Phi_y \colon \mathscr{H} &\to \mathbb{K} \\ x &\mapsto \langle x, y \rangle. \end{aligned} \tag{25.55}$$

EQooMWZAooZUFKOs

Ce sont des applications continues.

25.3.1 Orthogonal d'une partie

DEFooXUXQooMmDnhW

Définition 25.11.
*Soit une partie A de l'espace de Hilbert $\mathscr{H}$. L'**orthogonal** de A est l'ensemble*

$$A^\perp = \{v \in \mathscr{H} \text{ tel que } \langle v, x \rangle = 0 \forall x \in A\}. \tag{25.56}$$

PropdpaMpH

Proposition 25.12.
Si $\mathscr{H}$ est un préhilbert et si $A \subset \mathscr{H}$, alors l'ensemble $A^\perp$ est un sous-espace fermé de $\mathscr{H}$.

Démonstration. L'application Φ_x définie en (25.55) est continue pour chaque $x \in \mathscr{H}$ et par conséquent l'ensemble $\ker \Phi_x = \Phi_x^{-1}(\{0\})$ est fermé. L'ensemble

$$A^\perp = \bigcap_{x \in A} \ker \Phi_x \tag{25.57}$$

est donc fermé. □

ThowZyaiz

Théorème 25.13.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace fermé de $\mathscr{H}$. Alors

(1) Nous avons la décomposition

$$\mathscr{H} = F \oplus F^\perp. \tag{25.58}$$

EqxXIvrG

ItemThowZyaizii

(2) La projection sur F par rapport à la somme directe (25.58) est la projection $\texttt{proj}_F$ du théorème de projection.

(3) La décomposition (25.58) est topologique.

Proposition 25.14.
Si F est un sous-espace de l'espace de Hilbert $\mathscr{H}$ alors on a $F^{\perp\perp} = \bar{F}$.

Démonstration. Nous savons par la proposition 25.12 que $F^\perp$ est fermé, par conséquent le théorème 25.13 donne la somme directe

$$\mathscr{H} = F^\perp \oplus F^{\perp\perp}. \tag{25.59}$$

Mais $\bar{F}$ étant également fermé nous avons la somme directe

$$\mathscr{H} = \bar{F} \oplus (\bar{F})^\perp. \tag{25.60}$$

Montrons que $(\bar{F})^\perp = F^\perp$. En effet si $x \in F^\perp$ et si $y \in \bar{F}$, alors il existe une suite y_n dans F qui converge vers y. Pour chaque n nous avons $\langle x, y_n \rangle = 0$ et donc $\langle x, y \rangle = 0$ par continuité du produit scalaire.

Nous avons donc

$$\mathscr{H} = F^\perp \oplus F^{\perp\perp} = \bar{F} \oplus F^\perp. \tag{25.61}$$

Mais $\bar{F} \subset F^{\perp\perp}$ (prendre une suite). Les espaces $\bar{F}$ et $F^{\perp\perp}$ étant tous deux des supplémentaires de $F^\perp$, nous déduisons qu'ils doivent être égaux. □

PropqiWonByiBmc

Proposition 25.15.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace vectoriel de $\mathscr{H}$. Alors F est dense si et seulement si $F^\perp = \{0\}$.

Démonstration. Nous savons que

$$\mathscr{H} = F^\perp \oplus \bar{F}. \tag{25.62}$$

Donc nous avons $\mathscr{H} = \bar{F}$ si et seulement si $F^\perp = \{0\}$. □

Pour vérifier si un sous-espace vectoriel d'un espace de Hilbert est dense, il suffit donc de montrer que son orthogonal est réduit à zéro.

25.3.2 Dual, théorème de représentation de Riesz

LemjYVcHE

Lemme 25.16.
L'application linéaire

$$\begin{aligned}\Phi\colon \mathscr{H} &\to \mathcal{L}(\mathscr{H}, \mathbb{K})\\ y &\mapsto \Phi_y\end{aligned} \tag{25.63}$$

est une isométrie : nous avons $\|\Phi_y\| = \|y\|$.

De plus pour chaque y, l'application Φ_y est continue.

Démonstration. En utilisant la définition de la norme opérateur et l'inégalité de Cauchy-Schwarz,

$$\|\Phi_y\| = \sup_{\|x\|=1} \|\Phi_y(x)\| = \sup |\langle x, y\rangle| \leqslant \sup \|x\|\|y\| = \|y\|. \tag{25.64}$$

Par conséquent $\|\Phi_y\| \leqslant \|y\|$. Mais d'autre part le fait que $\Phi_y(y) = \|y\|^2$ montre que $\|\Phi_y\| \geqslant \|y\|$.

En ce qui concerne la continuité de Φ_y, elle est garantie par le fait que c'est une application linéaire bornée via la proposition 10.70. □

Définition 25.17.
*Le **dual** de l'espace de Hilbert $\mathscr{H}$ est l'ensemble*

$$\mathscr{H}' = \{f\colon \mathscr{H} \to \mathbb{K} \text{ linéaire et continue}\}. \tag{25.65}$$

Notons que dans le contexte des espaces de Hilbert nous demandons la continuité des éléments du dual parce qu'elle n'est pas automatique par la linéarité dans les cas de dimension infinie. En principe nous devrions préciser dual *topologique*, mais nous ne le ferons pas systématiquement lorsque le contexte parle clairement de topologie (ce qui est le cas lorsqu'on parle d'espaces de Hilbert). De temps en temps le dual *algébrique* d'un espace est noté E^* ; dans ce cas la continuité n'est pas demandée.

ThoQgTovL

Théorème 25.18 (Théorème de représentation de Riesz, thème 21)**.**
Soit un espace de Hilbert $\mathscr{H}$ sur le corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). L'application

$$\begin{aligned}\Phi\colon \mathscr{H} &\to \mathscr{H}'\\ y &\mapsto \Phi_y\end{aligned} \tag{25.66}$$

est une bijection isométrique.

Démonstration. Nous savons du lemme 25.16 que Φ est une isométrie. Nous devons seulement montrer que Φ est surjective. L'application nulle est dans l'image de Φ. Soit $f \in \mathscr{H}'$ non nulle. Par continuité nous savons que $F = \ker(f)$ est fermé, donc

$$\mathscr{H} = \ker(f) \oplus (\ker f)^\perp \tag{25.67}$$

EQooFUCWooUnYcZG

par le théorème 25.13.

(1) **Une base adaptée**
Nous considérons une base de $\mathscr{H}$ adaptée à la décomposition 25.67, c'est-à-dire
— $(u_s)_{s\in S}$ une base de $\ker f$,
— $(v_t)_{t\in T}$ une base de $(\ker f)^\perp$.

(2) $(\ker f)^\perp \cap \ker(f) = \{0\}$
Si $v \in (\ker f)^\perp$ alors pour tout $w \in \ker(f)$ nous avons $\langle v, w\rangle = 0$. En particulier si v lui-même est dans $\ker(f)$ alors $\langle v, v\rangle = 0$ et $v = 0$.

(3) **Une base encore plus adaptée** Si v est non nul dans $(\ker f)^\perp$, alors $f(v) \neq 0$ (sinon il serait dans l'intersection entre $\ker(f)$ et $(\ker f)^\perp$). Mais $f(v) \neq 0$ dans $\mathbb{K}$ implique que $f(\mathbb{K}v) = \mathbb{K}$, c'est-à-dire que $\{f(v)\}$ est une base de $\text{Image}(f)$ (l'image de f est $\mathbb{K}$).
Le théorème du rang (théorème 4.47) assure alors que

$$\{v\} \cup \{u_s\}_{s\in S} \tag{25.68}$$

est une base de $\mathscr{H}$ avec $v \in (\ker f)^\perp$ et $u_s \in \ker f$. Nous choisissons v pour avoir $\|v\| = 1$.

(4) **Existence**
Nous pouvons maintenant prouver l'existence de y tel que $\Phi_y = f$. Prouvons qu'en posant $y = \overline{f(v)}v$ nous avons $\Phi_y = f$. Pour cela, un peu de calcul :

$$\Phi_y(v) = \langle v, y\rangle = f(v)\langle v, v\rangle = f(v) \tag{25.69a}$$
$$\Phi_y(u_s) = \langle u_s, y\rangle = 0. \tag{25.69b}$$

Par conséquent Φ_y et f coïncident sur une base de $\mathscr{H}$.

(5) **Unicité**
En ce qui concerne l'unicité, d'abord si $\Phi_y = f$ alors nous devons avoir $y \in (\ker f)^\perp$ et par conséquent $y = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Nous avons alors

$$\Phi_y(v) = \bar{\lambda}\langle v, v\rangle = \bar{\lambda}. \tag{25.70}$$

Pour que cela soit égal à $f(v)$, nous fixons $\lambda = \overline{f(v)}$.

□

Notons que nous avons réellement utilisé le théorème du rang pour l'unicité. Si nous ne demandions pas l'unicité, alors nous n'avions pas besoin du fait que $\dim(\ker f)^\perp = 1$, et nous aurons donc pu parler de formes plus générales à valeurs dans $\mathbb{K}^n$.

25.3.3 Séparabilité

DEFooSFOJooGICSbT

Définition 25.19.
*Un espace topologique est **séparable** si il possède une partie dénombrable dense.*

DEFooQVPHooJaSWyF

Définition 25.20 ([625]).
*Si E est un espace vectoriel normé nous disons que Δ est une partie **totale** de E si* $\text{Span}(\Delta)$ *est dense dans E. Attention : nous rappelons que* $\text{Span}(\Delta)$ *est l'ensemble des combinaisons linéaires* finies *d'éléments de Δ.*

PROPooZMWHooVwvNBY

Proposition 25.21.
Un espace vectoriel normé est séparable si et seulement si il possède une partie totale dénombrable.

Démonstration. Si Δ est une partie dénombrable de E, alors le $\mathbb{Q}$-espace vectoriel $\text{Span}_{\mathbb{Q}}\Delta$ est dénombrable, et sa fermeture est la même que celle de $\text{Span}_{\mathbb{K}}\Delta$. □

Définition 25.22.
*Une famille $(u_i)_{i\in I}$ d'éléments de $\mathscr{H}$ indicée par un ensemble quelconque I est un **système orthonormé** si*

(1) $\|u_i\| = 1$ *pour tout* $i \in I$,

(2) $u_i \perp u_j$ *pour tout* $i \neq j$.

Notons que si $(u_n)_{n\in\mathbb{N}}$ est un système orthonormé dénombrable alors en utilisant les formules de la proposition 25.4, nous avons

$$\| \sum_{k=1}^{n} \xi_k u_k \|^2 = \sum_{k=1}^{n} |\xi_k|^2. \tag{25.71}$$

Exemple 25.23.
Dans l'ensemble $L^2(a,b)$ avec $b-a=L$ l'ensemble des fonctions

$$e_n(t) = \frac{1}{\sqrt{b-a}} \exp(\frac{2\pi}{L} int) \tag{25.72}$$

avec $n \in \mathbb{Z}$ forme un système orthonormé. En effet

$$\langle e_n, e_n \rangle = \frac{1}{b-a} \int_a^b \exp(\frac{2\pi}{L} int) \exp(-\frac{2\pi}{L} int) = 1 \tag{25.73}$$

et

$$\langle e_n, e_m \rangle = \frac{1}{b-a} \int_a^b e^{\frac{2\pi}{L}(n-m)t} dt = 0. \tag{25.74}$$

Dans cette intégrale nous utilisons le fait que $b = a + (b-a)$ pour simplifier les expressions en cours de calcul.

La famille (25.72) est le **système trigonométrique** de $L^2(a,b)$. On en parle aussi dans [585]. △

Proposition 25.24.
Une partie orthonormée est libre.

Démonstration. Soit $(u_i)_{i\in I}$ une famille orthonormée de $\mathscr{H}$ et une combinaison linéaire *finie* nulle :

$$\sum_{k=1}^{n} a_k u_k = 0. \tag{25.75}$$

Nous développons la somme en utilisant les formules de la proposition 25.4 :

$$0 = \| \sum_k a_k u_k \|^2 = \sum_k |a_k|^2 \|u_k\|^2 + 2 \sum_{k<l} \langle a_k u_k, a_l u_l \rangle. \tag{25.76}$$

La famille étant orthonormée, les choses se simplifient en

$$\sum_k |a_k|^2 = 0, \tag{25.77}$$

ce qui signifie que $a_k = 0$ pour tout k. □

Avant de continuer nous devons définir comment nous calculons des sommes sur des ensembles quelconques. Si I est un ensemble et si pour chaque $i \in I$ nous avons un nombre réel positif a_i, alors nous définissons

$$\sum_{i\in I} a_i = \sup_{\substack{J\subset I \\ J \text{ fini}}} \sum_{j\in J} a_j. \tag{25.78}$$

Cela est discuté dans la section 10.8.

Proposition 25.25 (Inégalités de Bessel)**.**
Soit $\mathscr{H}$ *un préhilbert. Si* $(u_i)_{i\in I}$ *est un système orthonormé et si* $x \in \mathscr{H}$, *alors*

$$\sum_{i\in I} |\langle x, u_i \rangle|^2 \leqslant \|x\|^2. \tag{25.79}$$

Démonstration. Les éléments de la somme étant des réels positifs, la proposition 10.116 fonctionne pour décrire les sommes.

Posons $c_i(x) = \langle x, u_i \rangle$. Pour toute partie finie $J \subset I$ nous avons

$$0 \leqslant \Big\| x - \sum_{j\in J} c_j(x)u_j \Big\|^2 = \|x\|^2 - 2\operatorname{Re}\sum_j \langle x, c_j(x)u_j\rangle + \sum_j |c_j(x)|^2. \tag{25.80}$$

Mais en tenant compte du fait que

$$\langle x, c_j(x)u_j\rangle = \overline{c_j(x)}\langle x, u_j\rangle = |c_j(x)|^2, \tag{25.81}$$

nous restons avec

$$\Big\| x - \sum_{j\in J} c_j(x)u_j \Big\| = \|x\|^2 - \sum_{j\in J} |c_j(x)|^2. \tag{25.82}$$

Finalement,

$$\sum_{j\in J} |c_j(x)|^2 \leqslant \|x\|^2. \tag{25.83}$$

Ayant cette inégalité pour toute partie finie de I, nous l'avons encore pour le supremum. □

Proposition 25.26.
Soit $\mathscr{H}$ un préhilbert et une famille orthonormée $(u_i)_{i\in I}$. Si

$$x = \sum_{i\in I} \xi_i u_i \tag{25.84}$$

alors $\xi_i = \langle x, u_i\rangle$.

Démonstration. Nous appliquons l'application Φ_{u_k} du théorème de représentation de Riesz[8] à l'équation (25.84).

Si I est dénombrable, alors permuter Φ_{u_k} avec la somme consiste à invoquer la continuité, et permuter la limite des sommes partielles avec Φ_{u_k} (l'application Φ_{u_k} est continue parce qu'isométrique).

Sinon, il faut utiliser la proposition 10.127. Il faut donc montrer que la famille $\Phi_{u_k}\big(\xi_i u_i\big)$ est sommable. Cela est fort vrai parce que cette famille ne contient en réalité qu'en seul élément non nul, celui avec $i = k$, qui vaut ξ_k. Au final nous avons :

$$\langle x, u_k\rangle = \sum_{i\in I} \Phi_{u_k}(\xi_i u_i) = \xi_k. \tag{25.85}$$

□

25.3.4 Base hilbertienne

Définition 25.27.
*Une **base orthonormée** est une famille dénombrable orthonormée et totale*[9]*. Cela sera souvent aussi appelé une **base hilbertienne**.*

25.28.
Cette définition demande quelques remarques.

(1) La notion de base hilbertienne n'est pas la même notion de base qu'en algèbre. En effet pour avoir une base algébrique d'un espace vectoriel, nous demandons que les éléments soient des combinaisons linéaires *finies* des éléments de la base, tandis qu'ici en demandant que la partie soit totale nous demandons simplement que les combinaisons linéaires finies soient denses.

(2) Nous allons voir qu'un espace de Hilbert est généré par les sommes *infinies* de vecteurs d'une base hilbertienne avec des coefficients qui forment une suite dans ℓ^2.

8. Théorème 25.18.
9. Définition 25.20

(3) Nous ne demandons pas que la famille soit libre ? La belle affaire. Une famille orthogonale est toujours libre, proposition 25.24.

(4) Nous verrons dans la proposition 25.37 que toute base hilbertienne est maximale (définition 25.33), c'est-à-dire que si $\langle x, b\rangle = 0$ pour tout b dans la base, alors $x = 0$.

LEMooHWOBooQJKdTD

Lemme 25.29 ([626]).
Si $\mathcal{E}$ est une base hilbertienne[10] *de l'espace de Hilbert $\mathscr{H}$ et si $x \in \mathscr{H}$ alors l'ensemble*

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} \tag{25.86}$$

est au plus dénombrable.

Démonstration. Notons qu'ici, $\mathscr{H}$ n'est pas supposé séparable. Nous savons par l'inégalité de Bessel (25.79) que

$$\sum_{e\in\mathcal{E}} |\langle x, e\rangle|^2 \leqslant \|x\|^2. \tag{25.87}$$

Donc si $\epsilon > 0$ est donné, l'ensemble $\{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \epsilon\}$ est fini. Or

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} = \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > 0\} = \bigcup_{n=1}^{\infty} \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \frac{1}{n}\}. \tag{25.88}$$

Bref, cet ensemble est une union dénombrable d'ensembles finis. Il est donc dénombrable. □

Remarque 25.30.
Le lemme 25.29 ne signifie pas que la base $\mathcal{E}$ doive être dénombrable. Il signifie seulement que pour chaque x séparément, seule une partie dénombrable de $\mathcal{E}$ est nécessaire.

CORooFROTooNupAQs

Corolaire 25.31 ([626]).
Si un espace de Hilbert possède une base hilbertienne dénombrable, alors toutes ses bases hilbertiennes sont dénombrables.

Démonstration. Soient $\mathcal{E}$ et $\mathcal{F}$ des bases hilbertiennes de l'espace de Hilbert $\mathscr{H}$, en supposant que $\mathcal{E}$ soit dénombrable. Pour chaque $f \in \mathcal{F}$ nous avons $\langle f, e\rangle \neq 0$ pour au moins un $e \in \mathcal{E}$, sinon en vertu de la décomposition (25.84) de f dans la base $\mathcal{E}$, nous aurions $f = 0$. Nous avons donc

$$\mathcal{F} \subset \bigcup_{e\in\mathcal{E}} \{f \in \mathcal{F} \text{ tel que } \langle f, e\rangle \neq 0\}, \tag{25.89}$$

EQooKRTFooNABXnH

alors que le lemme 25.29 indique que chacun des ensembles de l'union est au plus dénombrable. La partie $\mathcal{F}$ est donc une union dénombrable d'ensemble dénombrables. Elle est dénombrable. □

Remarque 25.32.
En travaillant un peu plus sur la notion de cardinalité, le corolaire 25.31 indique que toutes les bases hilbertiennes ont même cardinalité. En effet en laissant tomber l'hypothèse de dénombrabilité sur $\mathcal{E}$, l'inclusion (25.89) donne que $\mathcal{F}$ est une union de $\mathrm{Card}(\mathcal{E})$ ensembles dénombrables et est alors de cardinalité $\mathrm{Card}(\mathcal{E})$.

DEFooRFATooDRKWoJ

Définition 25.33.
*Une partie orthonormale B est **maximale** si le seul $x \in \mathscr{H}$ vérifiant $\langle x, b\rangle = 0$ pour tout $b \in B$ est $x = 0$.*

LEMooXIECooCAQeJN

Lemme 25.34 ([1, 627]).
Tout espace de Hilbert possède une partie orthonormale maximale.

10. Définition 25.27.

Démonstration. Soit $\mathscr{H}$ un espace de Hilbert et $\mathcal{O}$, l'ensemble de parties orthonormales de $\mathscr{H}$, ordonné [11] par l'inclusion. Cet ensemble est inductif. En effet soit une partie totalement ordonnée $\{A_i\}_{i\in I}$ de $\mathcal{O}$: chaque A_i est une partie orthonormale de $\mathscr{H}$, et de plus pour $i,j\in I$ nous avons soit $A_i\subset A_j$ soit $A_j\subset A_i$. Nous pouvons considérer

$$A=\bigcup_{i\in I}A_i. \tag{25.90}$$

Cela est encore une partie orthonormée de $\mathscr{H}$ parce que si $x,y\in A$, alors il existe $i,j\in I$ tels que $x\in A_i$ et $y\in A_j$. Vu que $\{A_i\}_{i\in I}$ est totalement ordonné nous supposons pour fixer les idées que $A_i\subset A_j$. Alors x et y sont dans A_j qui est une partie orthonormée ; nous en déduisons que $\langle x,y\rangle=0$ et donc que A est une partie orthonormée. C'est-à-dire : $A\in\mathcal{O}$. Par ailleurs, A est un majorant de $\{A_i\}_{i\in I}$ pour l'inclusion parce que $A_i\subset A$ pour tout i.

Nous avons prouvé que $\mathcal{O}$ est un ensemble inductif. Le lemme de Zorn 1.23 nous dit alors que $\mathcal{O}$ possède un élément maximum. Cet élément est une partie orthonormale incluse dans aucune autre partie orthonormale. C'est-à-dire qu'il est une partie orthonormale maximale au sens de la définition 25.33. □

PROPooENTIooIplRAS

Proposition 25.35 ([627])**.**
Si un espace de Hilbert possède une partie orthonormale maximale dénombrable, alors toutes les parties orthonormales sont dénombrables (ou finies).

Démonstration. Soit B une partie orthonormale maximale, et A une partie orthonormale. Pour chaque $b\in B$ nous notons

$$A(b)=\{a\in A\text{ tel que }\langle a,b\rangle\neq 0\}\cup\{0\}. \tag{25.91}$$

Un élément de A qui ne serait dans aucun des $A(b)$ serait perpendiculaire à tous les éléments de B, et serait donc l'élément nul. Mais l'élément nul est dans tous les $A(b)$; donc nous avons

$$A\subset\bigcup_{b\in B}A(b). \tag{25.92}$$

Or par l'inégalité de Bessel, l'ensemble $A(b)$ est dénombrable. Par conséquent A est inclus dans une union dénombrable d'ensembles dénombrables ; A est dénombrable. □

Le fait que tout espace de Hilbert possède une base hilbertienne est vrai. Nous allons démontrer ce résultat d'abord pour les espaces séparables et ensuite, indépendamment, en général. Si vous vous la sentez de maîtriser la proposition 25.37, vous pouvez sauter la 25.36.

PROPooDRFBooWTcunC

Proposition 25.36.
Tout espace de Hilbert séparable possède une base hilbertienne.

Démonstration. Vu que nous supposons avoir un espace de Hilbert séparable, il possède une partie totale dénombrable par la proposition 25.21. Soit $(v_n)_{n\in\mathbb{N}}$ une telle partie. Quitte à supprimer les v_i qui sont combinaisons linéaires des précédents, nous pouvons supposer que cette partie est libre. Nous considérons l'espace vectoriel

$$F_n=\text{Span}\{v_1,\dots,v_n\}. \tag{25.93}$$

Sur F_n nous pouvons appliquer un procédé de Gram-Schmidt pour construire une base orthonormée $\{u_1,\dots,u_n\}$ de F_n au sens usuel. En considérant F_{n+1} et en recommençant, les vecteurs $u_1,\dots,u_n$ ne changent pas, mais nous obtenons un vecteur u_{n+1}.

Nous construisons ainsi une suite (u_n) qui est alors orthonormée au sens des espaces de Hilbert. Nous devons encore prouver qu'il s'agit d'un ensemble total. Cela est simplement dû au fait que tout élément de $\text{Span}\{v_n\}$ est contenu dans $\text{Span}\{u_n\}$ parce que Span ne considère que des combinaisons linéaires finies. □

11. Définition 1.11.

PROPooLDXFooRaxBsI

Proposition 25.37 ([627]).
À propos de parties orthonormales maximales.

ITEMooVUFXooDrVwum

(1) *Une partie d'un espace de Hilbert est orthonormale maximale si et seulement si elle est une base hilbertienne.*

ITEMooZFENooQnSlrv

(2) *Tout espace de Hilbert possède une base hilbertienne.*

Démonstration. Soit une partie orthonormale maximale B et la fermeture de son espace engendré : $F = \overline{\text{Span}(B)}$. Pour que B soit une base, nous devons démontrer que $F = \mathscr{H}$. Pour cela nous considérons $x \in \mathscr{H}$ et nous utilisons le théorème de projection orthogonale 25.9 pour mentionner le fait que

$$x - \texttt{proj}_F(x) \perp F \tag{25.94}$$

Cela dit que $x - \texttt{proj}_F(x)$ est un vecteur orthogonal en particulier à tous les éléments de B ; par maximalité nous avons $x - \texttt{proj}_F(x) = 0$, c'est-à-dire $x = \texttt{proj}_F(x)$ ou encore $x \in F$. Cela prouve que $F = \mathscr{H}$.

Note : pour être pointilleux, nous aurions dû travailler non avec $x - \texttt{proj}_F(x)$, mais avec le vecteur normalisé à 1.

En ce qui concerne le second point, nous invoquons le lemme 25.34 pour dire que tout espace de Hilbert possède une partie orthonormale maximale. Ensuite la première partie de cette proposition nous dit que cette dernière est une base hilbertienne. □

PROPooEYRSooCRPjdK

Proposition 25.38 (Unicité de la décomposition dans une base hilbertienne[1]).
Soit un espace de Hilbert $\mathscr{H}$ muni d'une base hilbertienne $\{u_i\}_{i\in I}$. Alors :

(1) *Si*

$$\sum_{i\in I} a_i u_i = 0, \tag{25.95}$$

alors $a_i = 0$ pour tout $i \in I$.

(2) *Si $\sum_{i\in I} a_i u_i = \sum_{i\in I} b_i u_i$, alors $a_i = b_i$ pour tout i dans I.*

Démonstration. Supposons que $\sum_i a_i u_i = 0$. La proposition 25.26 nous dit alors que

$$a_j = \langle \sum_i a_i u_i, u_j \rangle = \langle 0, u_j \rangle = 0. \tag{25.96}$$

Pour la même raison, si

$$x = \sum_{i\in I} a_i u_i = \sum_{i\in I} b_i u_i, \tag{25.97}$$

alors, pour chaque $i \in I$, nous avons $a_i = \langle x, u_i \rangle$ et $b_i = \langle x, u_i \rangle$, c'est-à-dire $a_i = b_i$. □

Voici un petit résumé de ce que nous avons vu en termes de dénombrabilité et séparabilité.

THOooMKNFooVrCNGA

Théorème 25.39.
Pour un espace de Hilbert, les choses suivantes sont équivalentes.

ITEMooSJKVooFIIbwg

(1) *L'espace est séparable*[12].

ITEMooQIZLooYdtYqF

(2) *L'espace possède au moins une base hilbertienne*[13] *dénombrable.*

ITEMooHYSXooOubwUy

(3) *Toutes les bases hilbertiennes sont dénombrables.*

ITEMooMZICooNBAVum

(4) *Toute partie libre est dénombrable.*

Démonstration. Plein de résultats à citer ...

12. Définition 25.19.
13. Définition 25.27.

(1) **(1) implique (2)**
est la proposition 25.36.

(2) **(2) implique (3)**
est la proposition 25.31.

(3) **(3) implique (4)**
Le procédé de Gram-Schmidt met en bijection une partie libre avec une partie orthonormale (qui engendre le même espace, mais c'est une autre affaire). Or lorsque l'espace de Hilbert possède une base dénombrable, toutes les parties orthonormales sont dénombrables par la proposition 25.35.

(4) **(4) implique (1)** La proposition 25.37(2) dit que tout espace de Hilbert possède des bases hilbertiennes. Une telle base est forcément une partie libre parce que toute famille orthonormale est libre (proposition 25.24), et donc dénombrable par hypothèse. À ce point nous avons montré que notre espace de Hilbert possédait une base hilbertienne dénombrable. Cela implique qu'il est séparable par la proposition 25.21.

□
RemfdJcQF

Remarque 25.40.
À mon avis il doit exister un théorème de complétion de base hilbertienne disant que si on a une famille orthonormée, alors elle se prolonge en base. Utilisant cela, nous trouvons une nouvelle démonstration de la proposition 25.25 en disant que la somme sur la « partie de base » est plus petite que la somme sur la « base complète ».

25.41.
Vu que nous n'avons l'intention de ne travailler qu'avec des espaces de Hilbert séparables et que toutes leurs bases sont dénombrables, nous n'allons travailler qu'avec des bases dénombrables, et donc des systèmes orthonormés dénombrables. Nous allons conventionnellement les indicer par $\mathbb{N}$.

La proposition suivante explique que la notion de projection est compatible avec la décomposition d'un vecteur dans un système orthonormé.

Proposition 25.42.
Soit $(u_k)_{k\geqslant 1}$ un système orthonormé d'un préhilbert $\mathscr{H}$. Soient

$$x = \sum_{k=1}^{\infty} \xi_k u_k \tag{25.98}$$

et

$$F = \operatorname{Span}\{u_1, \dots, u_n\}. \tag{25.99}$$

Alors

$$\texttt{proj}_F(x) = \sum_{k=1}^{n} \xi_k u_k. \tag{25.100}$$

Démonstration. Nous allons dans un premier temps montrer que

$$y = x - \sum_{k=1}^{n} \xi_k u_k \tag{25.101}$$

est dans $F^\perp$. Pour cela nous calculons

$$\langle x - \sum_{k=1}^{n} \xi_k u_k, u_j \rangle = \langle \sum_{k=n+1}^{\infty} \xi_k u_k, u_j \rangle = 0 \tag{25.102}$$

où nous avons utilisé la continuité du produit scalaire pour permuter la somme (infinie) et le produit. Étant donné que $y \in F^\perp$ nous avons $\texttt{proj}_F y = 0$ par le point (2) du théorème 25.13.

D'autre part $\mathtt{proj}_F y$ peut être calculé selon

$$\mathtt{proj}_F y = \mathtt{proj}_F x - \sum_{k=1}^{n} \xi_k \mathtt{proj}_F u_k \tag{25.103}$$

tandis que $\mathtt{proj}_F u_k = u_k$ lorsque $1 \leqslant k \leqslant n$. Par conséquent l'annulation de $\mathtt{proj}_F y$ donne

$$\mathtt{proj}_F x = \sum_{k=1}^{n} \xi_k u_k, \tag{25.104}$$

donc le résultat. □

ThooRArDp

Théorème 25.43 (Meilleure approximation)**.**
Soit $\{u_i\}_{i\in I}$ une famille orthonormée de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x \in \mathscr{H}$ et pour tout J fini dans I et pour toute famille de nombres complexes $(a_j)_{j\in J}$ nous avons

$$\| \sum_{j\in J} \langle x, u_j \rangle u_j - x \| \leqslant \| \sum_{j\in J} a_j u_j - x \|. \tag{25.105}$$

Ce théorème exprime le fait que les nombres $\langle x, u_i \rangle$ sont les meilleurs coefficients à mettre devant les u_i pour approximer x.

Corolaire 25.44.
Soit $\{u_i\}_{i\in I}$ une famille orthonormée de $\mathscr{H}$. Pour tout $x \in \mathscr{H}$ et pour toutes parties finies J, K de I avec $J \subset K$ nous avons

$$\| \sum_{j\in K} \langle x, u_j \rangle u_j - x \| \leqslant \| \sum_{j\in J} \langle x, u_j \rangle u_j - x \|. \tag{25.106}$$

Ce corolaire exprime le fait que plus on prend de termes de la forme $\langle , u_i \rangle u_i$, mieux c'est.

PropzaKXHq

Proposition 25.45.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) un système orthonormé dans $\mathscr{H}$. Si $(\xi_n)_{n\in\mathbb{N}}$ est une suite dans ℓ^2 alors la série

$$\sum_{n=1}^{\infty} \xi_n u_n \tag{25.107}$$

converge dans $\mathscr{H}$.

Autrement dit l'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n \rangle\big)_{n\geqslant 1} \end{aligned} \tag{25.108}$$

est surjective.

Démonstration. Nous allons montrer que la série $\sum_{n=1}^{\infty} \xi_n u_n$ est de Cauchy, c'est-à-dire que la limite

$$\lim_{n\to\infty} \| \sum_{k=n}^{n+p} \xi_k u_k \| = 0 \tag{25.109}$$

est uniforme en p. Cela est un corolaire de la formule (25.71) parce que

$$\| \sum_{k=n}^{n+p} \xi_k u_k \|^2 = \sum_{k=n}^{n+p} |\xi_k|^2. \tag{25.110}$$

Mais si (ξ_n) est dans ℓ^2, pour tout ϵ, il existe N tel que si $n > N$ alors le membre de droite est inférieur à ϵ indépendamment de p. □

25.3.5 Décomposition dans une base hilbertienne

Étant donné une base hilbertienne de $\mathscr{H}$, nous notons

$$c_k(x) = \langle x, u_k \rangle. \tag{25.111}$$

Dans le théorème suivant (et d'ailleurs partout), les sommes sur I sont prises au sens de la définition 10.112.

ThoyAjoqP

Théorème 25.46 (Décomposition dans une base orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert séparable $\{u_i\}_{i\in I}$ une base orthonormée (I est un ensemble dénombrable quelconque).

ItemQGwoIxi

(1) Pour tout $x \in \mathscr{H}$ nous avons

EqPUPTXJ
$$x = \sum_{i\in I} \langle x, u_i \rangle u_i \tag{25.112}$$

où la somme est prise au sens de la définition 10.112. En particulier, la somme converge de façon commutative.

(2) Si $\{e_i\}_{i\in I}$ est une famille orthonormée qui satisfait la décomposition (25.112) pour tout $x \in \mathscr{H}$ alors $\{e_i\}$ est une base hilbertienne.

(3) Nous avons l'identité de ***Plancherel***

$$\|x\|^2 = \sum_{i\in I} |\langle x, u_i \rangle|^2. \tag{25.113}$$

Le point (5) nous indiquera que cette égalité est en fait suffisante pour dire que nous avons une base.

(4) Nous avons l'identité de ***Parseval***

EqHZxitKt
$$\langle x, y \rangle = \sum_{i\in I} \langle x, u_i \rangle \overline{\langle y, u_i \rangle}. \tag{25.114}$$

ItemQGwoIx

(5) Si $\{u_i\}$ est une famille de vecteurs unitaires vérifiant l'identité de Plancherel, alors c'est une base hilbertienne.

(6) Si $\{u_i\}_{i\in I}$ est une base hilbertienne, la suite $n \mapsto |\langle x, u_n \rangle|$ appartient à $\ell^2(I)$.

Démonstration. (1) Étant donné que le système $\{u_i\}_{i\in I}$ est total, nous pouvons considérer une suite de combinaisons linéaires finies des u_i qui converge vers x. Nous écrivons

$$x_n = \sum_{x\in J_n} a_{n,k} u_j \tag{25.115}$$

et $x_n \to x$ dans $\mathscr{H}$. Les ensembles J_n sont des sous-ensembles finis de I. Nous pouvons les choisir de telle sorte que $J_n \subset J_{n+1}$ et $\bigcup_{n\in\mathbb{N}} J_n = I$. Ce choix correspond à éventuellement prendre $a_{n,j} = 0$ pour toutes les valeurs de j « en trop ».
Soit $\epsilon > 0$ et N tel que $\|x_n - x\| < \epsilon$ pour tout $n \geqslant N$. Nous allons montrer que pour tout J fini tel que $J_N \subset J$ nous avons $\| \sum_{j\in J} \langle x, u_j \rangle u_j - x \| < \epsilon$. Étant donné que

$$\sum_{j\in J_N} \langle x, u_j \rangle u_j = \sum_{j\in J} a_j u_j \tag{25.116}$$

avec

$$a_j = \begin{cases} \langle x, u_j \rangle & \text{si } j \in J_N \\ 0 & \text{sinon,} \end{cases} \tag{25.117}$$

le théorème de meilleure approximation 25.43 nous enseigne que

$$\| \sum_{j\in J} \langle x, u_j \rangle u_j - x \| \leqslant \| \sum_{j\in J_N} \langle x, u_j \rangle u_j - x \| < \epsilon \tag{25.118}$$

par conséquent la somme $\sum_{i\in I}\langle x,u_i\rangle u_i$ converge vers x au sens général, et en particulier commutativement.

Les sommes étant commutatives (en particulier $x=\sum_{i\in I}\langle x,u_i\rangle u_i$), et les bases hilbertiennes étant dénombrables, nous ne perdons aucune généralité en ne considérant que des bases indexées par $\mathbb{N}$.

(2) Nous devons montrer que l'ensemble des combinaisons linéaires finies est dense dans $\mathscr{H}$. Par hypothèse, pour tout ϵ, il existe un ensemble fini J tel que

$$\|\sum_{j\in J}\langle x,u_j\rangle u_j-x\|<\epsilon. \tag{25.119}$$

Cela prouve la densité dont nous avions besoin.

(3) La norme étant une fonction continue, elle commute avec les sommes infinies, de telle sorte que l'égalité de Plancherel donne

$$\|x\|^2=\|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2. \tag{25.120}$$

Le système des u_i étant orthonormé,

$$\|x\|^2=\|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2=\sum_{i\in I}|\langle x,u_i\rangle|^2. \tag{25.121}$$

(4) Au tour de Parseval. Nous commençons par prouver que la somme du membre de droite converge. En utilisant l'inégalité $|zz'|\leqslant|z|^2+|z'|^2$ (valable pour $z,z'\in\mathbb{C}$) et Plancherel, nous avons

$$\begin{aligned}\sum_{i\in I}|\langle x,u_i\rangle\overline{\langle y,u_i\rangle}|&\leqslant\sum_{i\in I}|\langle x,u_i\rangle|^2+|\langle y,u_i\rangle|^2\\&\leqslant\|x\|^2+\|y\|^2.\end{aligned} \tag{25.122}$$

Nous en déduisons que la famille $\langle x,u_i\rangle\overline{\langle y,u_i\rangle}$ est (commutativement) sommable en utilisant la proposition 10.123. Par ailleurs nous savons que

$$x=\sum_{i\in I}c_i(x)u_i \tag{25.123a}$$

$$y=\sum_{i\in I}c_i(y)u_i, \tag{25.123b}$$

et le produit scalaire étant une forme bilinéaire continue,

$$\langle x,y\rangle=\sum_{i\in I}\sum_{j\in I}\langle c_i(x)u_i,\overline{c_j(y)u_j}\rangle \tag{25.124a}$$

$$=\sum_{i\in I}\sum_{j\in J}c_i(x)\overline{c_j(y)}\delta_{ij} \tag{25.124b}$$

$$=\sum_{i\in I}c_i(x)\overline{c_i(y)}. \tag{25.124c}$$

(5) Nous utilisons l'égalité de Plancherel avec $x=u_j$:

$$\|u_j\|^2=\|u_j\|^2+\sum_{i\in I\setminus\{j\}}|\langle u_j,u_i\rangle|^2. \tag{25.125}$$

Par conséquent $\langle u_j,u_i\rangle=0$ dès que $i\neq j$. Cela prouve que le système $\{u_i\}_{i\in I}$ est orthonormé. Nous devons encore prouver que le système est total. Pour cela nous repartons de l'équation (25.82) que nous avions déduites dans la démonstration de l'inégalité de Bessel :

$$\|x-\sum_{j\in J}c_j(x)u_j\|=\|x\|^2-\sum_{j\in J}|c_j(x)|^2. \tag{25.126}$$

Par hypothèse le membre de droite peut être rendu aussi petit que l'on veut en prenant J grand (mais fini) dans I. Le membre de gauche indique alors que le système $\{u_i\}_{i\in I}$ est total.

(6) L'identité de Plancherel signifie entre autres que si $x \in \mathscr{H}$ alors $\sum_{i\in I} |\langle x, u_i\rangle|^2$ converge. Du coup la suite $(\langle x, u_i\rangle)_{i\in I}$ est dans $\ell^2(I)$.

□

Remarque 25.47.
Nous avons décidé d'indexer les bases hilbertiennes par $\mathbb{N}$; cela est légitime parce que les sommes sont commutatives. Il ne faut cependant pas perdre de vue qu'en pratique l'ensemble naturel avec lequel on indexe une base est parfois $\mathbb{Z}$. Un tel cas est donné par la base trigonométrique de L^2. Indexer cette dernière par $\mathbb{N}$ plutôt que par $\mathbb{Z}$ serait une contorsion inutile.

Remarque 25.48.
L'égalité de Parseval est la raison pour laquelle les physiciens écrivent souvent

$$\mathrm{Id} = \sum_{n=1}^{\infty} |u_n\rangle\langle u_n| \tag{25.127}$$

dans les livres de mécanique quantique par exemple. Dans certains[628], nous lisons même

$$\int_{-\infty}^{+\infty} dq|q\rangle\langle q| = \hat{I}. \tag{25.128}$$

Notons que ces personnes travaillent avec un espace de Hilbert dont la base n'est pas dénombrable. Pour dire que la physique, ça n'utilise pas des mathématiques pour rire !

Remarque 25.49.
Par définition une base orthonormée est donc une partie dénombrable dont l'espace vectoriel engendré est dense. Un espace de Hilbert possédant une base orthonormée est donc séparable. C'est ce fait qui nous pousse à ne considérer que des espaces de Hilbert séparables ; nous n'allons donc pas étudier ce qu'il se passerait par exemple en considérant l'espace vectoriel librement engendré par les éléments de $\mathbb{R}$.

Exemple 25.50.
L'identité de Parseval (25.114), dans le cas de l'espace des fonctions continues périodiques de période 2π signifie qu'en posant

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(s)e^{-ins}, \tag{25.129}$$

nous avons

$$\frac{1}{2\pi}\int_0^{2\pi} |f(s)|^2 = \sum_{n=-\infty}^{\infty} |c_n(f)|^2. \tag{25.130}$$ EqMIuCSfz

△ CorQETwUdF

Corolaire 25.51.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) une base orthonormée. L'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n\rangle\big)_{n\in\mathbb{N}} \end{aligned} \tag{25.131}$$

est un isomorphisme d'espaces de Hilbert.

De plus l'isomorphisme réciproque est

$$\begin{aligned} S^{-1}\colon \ell^2 &\to \mathscr{H} \\ (\xi_n) &\mapsto \sum_{n=1}^{\infty} \xi_n u_n. \end{aligned} \tag{25.132}$$

Démonstration. Nous devons prouver que l'application est bijective et qu'elle vérifie

$$S(x) \cdot S(y) = \langle x, y \rangle \tag{25.133}$$

où le point dénote le produit dans ℓ^2.

Pour la surjectivité, si $(\xi_n) \in \ell^2$ alors nous savons que la somme $\sum_n \xi_n u_n$ converge par la proposition 25.45 et par conséquent (ξ_n) est l'image par S de ce vecteur de $\mathscr{H}$.

Pour l'injectivité, si $S(x) = S(y)$ alors

$$x = \sum_n \langle x, u_n \rangle u_n = \sum_n \langle y, u_n \rangle u_n = y \tag{25.134}$$

en utilisant la décomposition (25.112).

Le fait que S soit une isométrie est contenu dans Parseval. □

PROPooPVQIooPcEFSe

Proposition 25.52 ([1]).
Soit un espace de Hilbert séparable $\mathscr{H}$, un sous-espace vectoriel fermé V et une base orthonormée $\{b_i\}_{i\in I}$ de $\mathscr{H}$. En posant $v_i = \texttt{proj}_V(b_i)$ alors

$$C = \overline{\text{Span}\{b_i - v_i\}} \tag{25.135}$$

est l'orthogonal de V.

Démonstration. Juste pour rappel, lorsque nous écrivons $v_i = \texttt{proj}_V(b_i)$, nous parlons de la projection orthogonale du théorème 25.9. Donc tous les vecteurs $b_i - v_i$ sont dans $V^\perp$. En passant aux limites, $C \subset V^\perp$.

Soit $x \in V^\perp$ que nous décomposons dans la base $\{b_i\}_{i\in I}$ comme $x = \sum_{i\in I} x_i b_i$. Posons $c_i = b_i - v_i = b_i - \texttt{proj}_V(b_i)$. Alors d'une part

$$0 = \texttt{proj}_V(x) = \sum_i x_i \texttt{proj}_V(b_i) = \sum_i x_i(b_i - c_i). \tag{25.136}$$

Nous avons utilisé la continuité de $\texttt{proj}_V$ pour permuter avec la somme. D'autre part,

$$x = \sum_i x_i b_i = \sum_i x_i(b_i - c_i + c_i) = \underbrace{\sum_i x_i(b_i - c_i)}_{=0} + \sum_i x_i c_i \in C. \tag{25.137}$$

Notons que pour la dernière appartenance, il est important de prendre la fermeture pour définir C. □

Proposition 25.53 ([1, 629]).
Toute partie orthonormée d'un espace de Hilbert séparable se prolonge en une base hilbertienne.

Démonstration. Soit une partie orthonormée $\{u_i\}_{i\in I}$ de l'espace de Hilbert $\mathscr{H}$. Nous mentionnons que cette partie est libre et donc, par le théorème 25.39, dénombrable [14]. Montrons pour commencer que la partie $V = \overline{\text{Span}\{u_i\}_{i\in I}}$ est un sous-espace vectoriel fermé de $\mathscr{H}$.

(1) **Vectoriel** Soit $v, w \in V$ et $\epsilon > 0$. Il existe $a \in \text{Span}\{u_i\}$ tel que $\|a - v\| < \epsilon$ et $b \in \text{Span}\{u_i\}$ tel que $\|b - w\| < \epsilon$. Dans ce cas,

$$\|v + w - (a + b)\| \leqslant 2\epsilon. \tag{25.138}$$

Cela prouve que $v + w \in V$. Nous procédons de même pour λv.

(2) **Fermé** Par construction.

14. Avec un peu de mauvaise foi, vous pouvez quand même dire que cela n'implique pas que I lui-même soit dénombrable, mais vous pouvez le supposer pour fixer les idées.

L'orthogonal[15] de V est un sous-espace vectoriel de $\mathscr{H}$, et nous pouvons donc en considérer une base hilbertienne $C = \{c_\alpha\}_{\alpha\in A}$. Nous prétendons que $C \cup U$ est une base hilbertienne[16] de $\mathscr{H}$.

(1) $C \cup U$ **est libre** Bing ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

(2) $C \cup U$ **est générateur** Bang ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

(3) $C \cup U$ **est orthogonal** Ah, voilà quelque chose à démontrer. Nous devons vérifier que les produits sont nuls. Soient $u, v \in U$ et $a, b \in C$. Nous avons :

— $\langle u, v\rangle = 0$ par hypothèse.

— $\langle u, a\rangle = 0$ parce que les éléments de C sont orthogonaux à V et que $u_i \in V$.

— $\langle a, b\rangle = 0$ parce que C est une base hilbertienne de $V^\perp$.

(4) $C \cup U$ **est dénombrable**
L'ensemble C est dénombrable parce que c'est une base hilbertienne. Quant à $\{u_i\}_{i\in I}$, nous avons déjà mentionné le fait qu'il doive être dénombrable. L'union deux parties dénombrables est dénombrable.

(5) $C \cup U$ **est total**
Nous devons prouver que $\overline{\text{Span}(C \cup U)} = \mathscr{H}$ parce qu'il y a bien la fermeture qui intervient dans la définition 25.20. Pour cela nous utilisons la proposition 25.15. Si $x \in \mathscr{H}$ alors nous avons

$$x = \texttt{proj}_V(x) + \big(x - \texttt{proj}_V(x)\big) = \sum_{i\in I} x_i u_i + \sum_{\alpha\in A} x_\alpha c_\alpha \tag{25.139}$$

pour des coefficients x_i et x_α. Notons que I et A sont deux ensembles différents. Aucun des x_α n'est un des x_i, ni inversement. Supposons que $x \in \text{Span}(C \cup U)^\perp$; alors

$$0 = \langle x, u_j\rangle = \sum_{i\in I} x_i \underbrace{\langle u_i, u_j\rangle}_{=\delta_{ij}} + \sum_{\alpha\in A} \underbrace{\langle c_\alpha, u_j\rangle}_{=0} = x_j. \tag{25.140}$$

Donc $x_j = 0$. En faisant de même avec $0 = \langle x, c_\beta\rangle = x_\beta$ nous déduisons $x = 0$.

□

PROPooZQIAooUEWLtg

Proposition 25.54 ([1]).
Si $\mathscr{H}$ est un espace de Hilbert séparable de base hilbertienne $\{e_i\}_{i\in I}$, alors $\mathscr{H}'$ est un espace de Hilbert séparable dont une base hilbertienne est donnée par les formes

$$\begin{aligned} \alpha_i \colon \mathscr{H} &\to \mathbb{C} \\ e_j &\mapsto \delta_{ij}. \end{aligned} \tag{25.141}$$

Démonstration. C'est la proposition 25.6 qui fait tout. □

25.3.6 Digression sur les normes opérateurs

subsecaeSywF

Le théorème 25.46 nous indique que si $\{u_i\}_{i\in I}$ est une base hilbertienne, alors pour tout $x \in \mathscr{H}$ nous avons

$$\texttt{proj}_{u_i}(x) = \langle x, u_i\rangle u_i, \tag{25.142}$$

et donc

$$\sum_{i\in I} \texttt{proj}_{u_i} x = x. \tag{25.143}$$

15. Voir la définition 25.11 et la proposition 25.52.
16. Définition 25.27.

Nous ne pouvons cependant pas conclure que

$$\sum_{i\in I} \mathtt{proj}_{u_i} = \mathrm{Id} \tag{25.144}$$

au sens de la norme opérateur de la définition 10.55. En effet en prenant $I = \mathbb{N}$, l'égalité (25.144) demanderait d'avoir

$$\lim_{N\to\infty} \left\| \sum_{i=1}^{N} \mathtt{proj}_{u_i} - \mathrm{Id} \right\|_\infty = 0, \tag{25.145}$$

or pour tout N, le vecteur u_{N+1} réalise

$$\sum_{i=1}^{N} \mathtt{proj}_{u_i}(u_{N+1}) - u_{N+1} = -u_{N+1}. \tag{25.146}$$

Par conséquent pour tout N nous avons

$$\sup_{\|x\|=1} \left\| \sum_{i=1}^{N} \mathtt{proj}_{u_i} x - x \right\| \geqslant 1. \tag{25.147}$$

Nous ne pouvons donc pas dire

$$\sum_{n=1}^{\infty} \mathtt{proj}_{u_i} = \mathrm{Id} \tag{25.148}$$

au sens de la norme opérateur.

Nous avons cependant la convergence au sens faible.

Proposition 25.55.
Soit $\mathscr{H}$ un espace de Hilbert et $\{u_i\}_{i\in\mathbb{N}}$ une base hilbertienne de $\mathscr{H}$. Au sens de la topologie faible sur l'espace des opérateurs nous avons

$$\sum_{i=1}^{\infty} \mathit{proj}_{u_i} = \mathrm{Id}\,. \tag{25.149}$$

Démonstration. Pour chaque $N \in \mathbb{N}$ et $x \in \mathscr{H}$, en vertu de la décomposition (25.112) nous avons

$$\sum_{i=1}^{N} \mathtt{proj}_{u_i}(x) - x = \sum_{i=N+1}^{\infty} \langle x, u_i\rangle u_i. \tag{25.150}$$

Par l'orthonormalité de la base nous avons

$$\sum_{i=N+1}^{\infty} \|\langle x, u_i\rangle u_i\| = \sum_{i=N+1}^{\infty} |\langle x, u_i\rangle|, \tag{25.151}$$

dont la limite $N \to \infty$ est zéro étant donné que la suite $i \mapsto |\langle x, u_i\rangle|$ est dans $\ell^2(\mathbb{R})$ par le théorème 25.46. □

25.3.7 Applications linéaires et continuité

Nous avons déjà vu dans l'exemple 10.71 que la fonction

$$\begin{aligned} f\colon H &\to H \\ e_k &\mapsto k e_k \end{aligned} \tag{25.152}$$

n'était pas continue en zéro alors qu'elle est linéaire. Nous allons maintenant voir qu'elle est un contre-exemple à la proposition 12.295. Calculons les dérivées partielles :

$$\frac{\partial f}{\partial e_j}(x) = \frac{d}{dt}\Big[f(x + te_j)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + t\delta_{jk})e_k\Big]_{t=0} = je_j. \tag{25.153}$$

où nous avons permuté la somme et la dérivée en considérant la suite de fonctions $f_k(t) = k(x_k + tu_k)e_k$. Donc $\frac{\partial f}{\partial e_j}(x)$ existe et est continue sur un voisinage de $x = 0$ (c'est même constant). Nous savons pourtant que la fonction f n'est pas différentiable en zéro parce que non continue.

L'endroit qui coince dans la preuve de la proposition 12.295 est l'introduction des « contres-termes » dans l'équation (12.799). En effet les contre-termes à ajouter seraient

$$l(x) = \frac{d}{dt}\Big[f\big(a + s(x-a)\big)\Big]_{t=0} \tag{25.154}$$

qui ici serait

$$l\big(\sum_k x_k e_k\big) = \frac{d}{dt}\Big[f\big(t\sum_k x_k e_k\big)\Big]_{t=0} = \sum_k x_k e_k, \tag{25.155}$$

dont la convergence est plus que douteuse.

Notons que les dérivées directionnelles n'existent pas toutes, loin s'en faut : si $u \in H$ nous avons

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + tu_k)e_k\Big]_{t=0} = \sum_k k u_k e_k \tag{25.156}$$

EQooWNLOooJNRUMQ

La convergence de la dernière somme n'est pas garantie pour tout u.

Exemple 25.56.
Soit un espace de Hilbert $\mathscr{H}$. L'application

$$\begin{aligned} \phi\colon \mathscr{H} &\to \mathbb{C} \\ e_i &\mapsto 1 \end{aligned} \tag{25.157}$$

n'est pas continue.

Nous considérons les vecteurs $y_n = \sum_{i=1}^n e_i$ et nous posons $x_n = y_n/\|y_n\|$. Nous avons

$$\|y_n\|^2 = \langle y_n, y_n\rangle = \sum_{i,j}^n \langle e_i, e_j\rangle = n, \tag{25.158}$$

donc $\|y_n\| = \sqrt{n}$, et

$$\phi(x_n) = \frac{n}{\sqrt{n}} = \sqrt{n}. \tag{25.159}$$

L'application ϕ n'est donc pas bornée et pas continue non plus. △

PROPooWHZKooEXEIrV

Proposition 25.57.
Soient deux espaces de Hilbert séparables[17] $\mathscr{H}_1$ *et* $\mathscr{H}_2$. *Soit une base orthonormée* $\{e_i\}_{i\in I}$ *où* I *est dénombrable. Si* $\Phi\colon \mathscr{H}_1 \to \mathscr{H}_2$ *est une bijection linéaire continue, alors*

$$\Phi\big(\sum_{i\in I} x_i e_i\big) = \sum_{i\in I} x_i \Phi(e_i). \tag{25.160}$$

Démonstration. Soit $\epsilon > 0$. Vu que $\sum_{i\in I} x_i e_i$ est une somme convergente[18], il existe une partie finie J de I telle que pour tout K fini contenant J,

$$\|\sum_{j\in K} x_j e_j - x\| < \epsilon. \tag{25.161}$$

Vu que Φ est une isométrie linéaire, nous avons

$$\|\sum_{j\in K} x_j \Phi(e_j) - \Phi(x)\| < \epsilon, \tag{25.162}$$

et donc bien l'égalité

$$\sum_{i\in I} x_i \Phi(e_i) = \Phi(x). \tag{25.163}$$

□

17. Donc à base dénombrable, voir 25.39.
18. Définition 10.112.

La proposition suivante donne une formule pour l'inverse du Φ du théorème de représentation de Riesz 25.18 lorsque nous avons une base hilbertienne.

Proposition 25.58.
Soit un espace de Hilbert séparable $\mathscr{H}$. Si $\phi \in \mathscr{H}'$, alors

$$\Phi^{-1}(\phi) = \sum_{i\in I} \phi(e_i)e_i \tag{25.164}$$

EQooRLJEooTaUtWy

dès que $\{e_i\}_{i\in I}$ est une base hilbertienne de $\mathscr{H}$.

Démonstration. Nous considérons la base des $\{\alpha_i\}_{i\in I}$ de $\mathscr{H}'$. Remarquez au passage que $\alpha_i = \Phi_{e_i}$. Nous avons $\phi = \sum_{i\in I} \phi_i\alpha_i$ pour certains $\phi_i \in \mathbb{C}$ (qui valent $\phi_i = \phi(e_i)$, ça a son importance pour la suite). En utilisant la proposition 25.57 pour permuter Φ et la somme nous avons le calcul suivant :

$$\phi = \sum_{i\in I} \phi_i\alpha_i = \sum_{i\in I} \phi_i\Phi(e_i) = \sum_{i\in I} \Phi(\phi_i e_i) = \Phi\big(\sum_{i\in I} \phi_i e_i\big) = \Phi(\sum_{i\in I} \phi(e_i)e_i). \tag{25.165}$$

En résulté :

$$\phi = \Phi(\sum_{i\in I} \phi(e_i)e_i), \tag{25.166}$$

et donc la formule (25.164). □

25.4 Théorème de Kochen-Specker

Le théorème suivant est central en mécanique quantique. La démonstration provient de [630] et de Wikipédia. Nous allons démontrer complètement le théorème seulement pour les espaces de Hilbert de dimension plus grande ou égale à 4.

THOooRHHPooQbcDAe

Théorème 25.59 (Kochen-Specker[630])**.**
Soit $\mathscr{H}$ un espace de Hilbert de dimension plus grande ou égale à 3. Une fonction v sur l'ensemble des opérateurs de $\mathscr{H}$ ne peut pas satisfaire aux conditions suivantes :

(1) Si A et B sont compatibles, alors $v(A + B) = v(A) + v(B)$,

(2) Si A et B sont compatibles, alors $v(AB) = v(A)v(B)$.

*Ici nous disons que deux opérateurs sont **compatibles** lorsqu'ils possèdent une base hilbertienne commune de vecteurs propres.*

Démonstration. Soit $\{u_n\}_{n\in\mathbb{N}}$ une base hilbertienne de $\mathscr{H}$. Nous notons $\texttt{proj}_i$ l'opérateur de projection sur l'espace (fermé) engendré par u_i. Ce sont des opérateurs compatibles deux à deux parce que la base $\{u_n\}_{n\in\mathbb{N}}$ est une base commune de vecteurs propres [19].

D'abord nous devons avoir $v(\mathbb{1}) = 1$. En effet pour tout opérateur A, nous avons

$$v(A) = v(A\mathbb{1}) = v(A)v(\mathbb{1}). \tag{25.167}$$

Pour peu que $v(A) \neq 0$, cela nous fait $v(\mathbb{1}) = 1$.

En vertu du théorème 25.46, un vecteur $x \in \mathscr{H}$ se décompose en $x = \sum_n \langle x, u_n\rangle u_n$, et nous avons

$$\texttt{proj}_i\, x = \langle x, u_i\rangle u_i. \tag{25.168}$$

EqFFuXyq

En effet le théorème de la projection orthogonale 25.9 nous enseigne que $\texttt{proj}_i\, x$ serait l'unique vecteur de la forme λu_i tel que $\lambda u_i - x \perp u_i$. Il est facile de vérifier que le vecteur proposé par (25.168) vérifie cette propriété.

Une conséquence est que

$$\left(\sum_{i=0}^{\infty} \texttt{proj}_i\right)(x) = x. \tag{25.169}$$

19. Pour les besoins de la physique, nous remarquons que ces opérateurs sont des opérateurs hermitiens qui commutent, mais ça ne joue pas ici.

Par conséquent, par hypothèse du théorème nous devons avoir

$$\sum_i v(\mathtt{proj}_i) = v(\mathbb{1}) = 1. \tag{25.170}$$

EqUeZHJH

Étant donné que les projections sont idempotentes,

$$v(\mathtt{proj}_i) = v(\mathtt{proj}_i^2) = v(\mathtt{proj}_i)^2 \tag{25.171}$$

et donc $v(\mathtt{proj}_i)$ doit valoir zéro ou un. Mais la relation (25.170) donne une forte contrainte sur le choix de 0 et de 1. En effet, parmi les $v(\mathtt{proj}_i)$, un et un seul doit valoir 1, les autres doivent valoir 0.

Refaisant le raisonnement pour une autre base orthonormale hilbertienne, nous trouvons que les valeurs de v sur les opérateurs de projection sur les différentes directions doivent être choisies de telle façon que tout choix de base hilbertienne orthogonale contienne exactement un 1 et le reste de zéros.

Nous voudrions maintenant insister sur un point. Le problème de déterminer de façon cohérente les valeurs 0 ou 1 pour tous les $v(\mathtt{proj}_i)$ revient à attacher 0 ou 1 à tous les *rayons* de $\mathscr{H}$ de façon que toute base orthogonale de $\mathscr{H}$ contienne exactement un 1. Un rayon est une direction, c'est-à-dire une classe d'équivalence $x \sim \lambda x$. Si nous décidons de nommer « blanc » les rayons attachés à la valeur 0 et « noirs » ceux attachés à la valeur 1, le problème se réduit à colorer la boule unité de façon compatible.

Soit $\{u_i\}_{i\in\mathbb{N}}$ une base orthogonale de $\mathscr{H}$ numérotée de telle sorte que $v(u_0) = 1$ et $v(u_k) = 0$ pour $k \neq 0$. Nous allons maintenant nous particulariser au cas de dimension supérieure ou égale à 4. Si R est une rotation dans le plan $\mathrm{Span}\{u_0, u_1, u_2, u_3\}$, alors l'ensemble

$$\{Ru_0, Ru_1, Ru_2, Ru_3, u_k\}_{k\geqslant 4} \tag{25.172}$$

est encore une base orthogonale de $\mathscr{H}$ et nous avons encore $v(u_k) = 0$ pour $k \geqslant 4$. Par conséquent un et un seul des vecteurs Ru_0, Ru_1, Ru_2 ou Ru_3 est colorié en noir ; les trois autres étant blancs. Le problème est maintenant complètement réduit à la dimension 4. Note : pour réduire à la dimension 3, on procède de même, mais pour conclure, il faut travailler plus.

Nous allons construire 9 bases orthogonales de $\mathbb{R}^4$ à partir de 18 vecteurs, chacun arrivant dans exactement deux des bases. Ils sont donnés dans le tableau suivant :

0	1	0	0	0	0	1	1	1	−1	1	−1	−1	1	1	1	1	1
0	0	1	0	0	1	1	1	1	−1	−1	1	1	1	−1	1	1	−1
0	0	0	1	0	0	1	−1	1	−1	1	1	1	1	−1	1	1	1
0	1	0	0	1	0	1	−1	−1	1	1	1	1	−1	1	1	−1	1
1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	1	−1
1	0	0	1	0	0	−1	0	0	0	0	−1	0	1	1	0	0	0
1	0	1	0	1	−1	0	1	0	0	0	1	0	1	0	1	0	0
−1	0	0	−1	0	0	0	−1	1	1	−1	0	−1	0	0	−1	1	1

Chaque case de ce table représente un rayon de $\mathbb{R}^4$; il y en a 18 différents, chacun écrit deux fois. Une simple vérification montre que chaque colonne est un système orthogonal. La preuve du théorème de Kochen-Specker revient à montrer que nous ne pouvons pas colorier ce tableau de façon cohérente. En effet, étant donné que chaque vecteur est écrit deux fois, le tableau doit contenir un nombre pair de cases blanches et un nombre pair de cases noires.

Par ailleurs chaque colonne étant un système orthogonal, chaque colonne contient exactement une case noire ; il y a donc exactement neuf cases noires dans le tableau, ce qui est impossible.

□

25.5 Théorème de Lax-Milgram

DEFooUNOKooCitMjL

Définition 25.60.
Une forme bilinéaire $a\colon V \times V \to \mathbb{R}$ sur un espace vectoriel normé V est ***coercitive*** *si il existe $\alpha > 0$ tel que $a(u,u) \geqslant \alpha\|u\|^2$ pour tout $u \in V$.*

THOooLLUXooHyqmVL

Théorème 25.61 (Lax-Milgram[631]).
Soit un espace de Hilbert réel V muni de différentes choses.

(1) *L'application linéaire $L: V \to \mathbb{R}$ qui est bornée sur V. Nous notons $\|L\|$ sa norme.*
(2) *La forme bilinéaire continue a sur $V \times V$.*
(3) *La forme a est coercitive*[20].

Alors

(1) *Il existe $M > 0$ tel que $|a(u,v)| \leqslant M\|u\|\|v\|$ pour tout $u, v \in V$.*
(2) *Le problème qui consiste à chercher $u \in V$ tel que $a(u,v) = L(v)$ pour tout $v \in V$ admet une unique solution. De plus cette solution vérifie l'inégalité*

$$\|u\| \leqslant \frac{\|L\|}{\alpha} \tag{25.173}$$

EQooUAYSooKYyQBU

Démonstration. Évacuons deux points faciles avant de commencer les choses sérieuses. D'abord L est continue par la proposition 10.70. Ensuite, l'existence du M est donnée par la proposition 25.2.

Maintenant nous commençons.

(1) **Reformulation en équation linéaire** La forme L est continue et donc dans le dual V' ; le théorème de Riesz 25.18 nous donne donc $f \in V$ tel que

$$L(v) = \langle f, v \rangle \tag{25.174}$$

pour tout $v \in V$. De plus si $w \in V$ est fixé, l'application $b_w: v \mapsto a(w,v)$ est linéaire et bornée parce que

$$\|b_w\| = \sup_{\|v\|=1} |b_w(v)| = \sup_{\|v\|=1} |a(w,v)| \leqslant M\|w\|\|v\| = M\|w\|. \tag{25.175}$$

Encore une fois, b_w étant continue et linéaire, elle est dans V' et Riesz nous fournit un élément $A(w) \in V$ tel que

$$b_w(v) = \langle A(w), v \rangle \tag{25.176}$$

pour tout $v \in V$.

Le problème variationnel $a(u,v) = L(v)$ est équivalent à $\langle A(u), v\rangle = \langle f, v\rangle$. L'ensemble des solutions de cette dernière est égal à l'ensemble des solutions de l'équation

$$A(u) = f. \tag{25.177}$$

EQooLPMPooMVuYUX

(2) **A est linéaire**
Soient $\alpha, \beta \in \mathbb{R}$ et $w, z \in V$. Nous avons pour tout $v \in V$:

$$\begin{aligned}
\langle A(\alpha w + \beta z), v\rangle &= a(\alpha w + \beta z, v) && (25.178a)\\
&= \alpha a(w,v) + \beta a(z,v) && (25.178b)\\
&= \alpha\langle A(w), v\rangle + \beta\langle A(z), v\rangle && (25.178c)\\
&= \langle \alpha A(w) + \beta A(z), v\rangle. && (25.178d)
\end{aligned}$$

Étant donné que nous avons égalité pour tout $v \in V$ nous en déduisons que $A(\alpha w + \beta z) = \alpha A(w) + \beta A(z)$, ce qui signifie que A est linéaire.

(3) **Une autre propriété de A** Nous déduisons une majoration de $\|A(v)\|^2$ lorsque ce n'est pas nul. Pour ce faire,

$$\begin{aligned}
\|A(v)\|^2 &= \langle A(v), A(v)\rangle && (25.179a)\\
&= a(v, A(v)) && (25.179b)\\
&\leqslant M\|v\|\|A(v)\|. && (25.179c)
\end{aligned}$$

En simplifiant, $\|A(v)\| \leqslant M\|v\|$. Et donc

$$\|A(v)\|^2 \leqslant M^2\|v\|^2. \tag{25.180}$$

EQooBQAHooAZRdAW

20. Définition 25.60.

(4) **Une contraction**
Nous allons choisir une valeur de $\rho > 0$ telle que l'application

$$T\colon w \mapsto w - \rho\big(A(w) - f\big) \tag{25.181}$$

soit une contraction [21]. Nous avons $T(w) - T(w') = w - w' - \rho\big(A(w - w')\big)$ et donc

$$\|T(w) - T(w')\|^2 = \|w - w'\|^2 + \rho^2\|A(w - w')\|^2 - 2\rho\langle A(w - w'), w - w'\rangle \tag{25.182a}$$
$$= \|w - w'\|^2 + \rho^2\|A(w - w')\|^2 - 2\rho a(w - w', w - w'). \tag{25.182b}$$

Vu que le dernier terme arrive avec un signe moins, pour majorer l'expression, il faut minorer ce terme, c'est-à-dire utiliser $a(w - w', w - w') \geqslant \alpha\|w - w - w'\|$. Et en même temps nous utilisons (25.180) pour le second terme. Au final pour pouvons factoriser $\|w - w'\|$ et

$$\|T(w) - T(w')\| \leqslant \|w - w'\|\big(1 + \rho^2 M^2 - 2\rho\alpha\big). \tag{25.183}$$

Pout que T soit contractante, il faut $0 < P(x) < 1$ avec $P(x) = M^2x^2 - 2\alpha x + 1$. Le minimum de ce polynôme est obtenu en $x = \frac{\alpha}{M^2}$ (la formule du $x_{min} = -b/2a$) et vaut $1 - \frac{\alpha^2}{M^2} < 1$. Vu que par ailleurs $\lim_{x\to\infty} P(x) = +\infty$, et que ce polynôme passe par au moins une valeur strictement inférieure à 1, nous savons qu'il existe un x tel que $0 < P(x) < 1$. En donnant à ρ cette valeur, l'application T est une contraction.

(5) **Point fixe et conclusion** L'ensemble des solutions du problème (25.177) est égal à l'ensemble des points fixes de $T(v) = v - \rho\big(A(v) - f\big)$.
L'application $T\colon V \to V$ est contractante et V est métrique et complet. Ergo le théorème de point fixe de Picard 17.34 s'applique et il existe un unique point fixe $u \in V$ pour l'application T. Ce point fixe est l'unique solution de notre problème initial.

(6) **La majoration**
Nous savons que pour tout $v \in V$, la relation $a(u, v) = L(v)$ est vérifiée. En particulier pour $v = u$ nous avons

$$a(u, u) = L(u). \tag{25.184}$$

D'un côté nous utilisons $a(u, u) \geqslant \alpha\|u\|^2$ et de l'autre, $L(u) \leqslant \|L\|\|u\|$:

$$\alpha\|u\|^2 \leqslant \|L\|\|u\| \tag{25.185}$$

et donc

$$\|u\| \leqslant \frac{\|L\|}{\alpha}. \tag{25.186}$$

Notons que $L(u)$ et $a(u, u)$ sont positifs.

□
EXooTTBDooUNhBOc

Exemple 25.62 ([632]).
La borne $\|u\| \leqslant \|L\|/\alpha$ est optimale au sens où il existe des cas d'égalité. En effet nous pouvons considérer

$$a(u, v) = s\langle u, v\rangle \tag{25.187}$$

et

$$L(v) = \langle f, v\rangle \tag{25.188}$$

pour un certain $s > 0$ et $f \in V$. L'application L n'est autre que Φ_f dont nous avons abondamment parlé autour du théorème de représentation de Riesz 25.18. Le lemme 25.16 nous apprend que L est une application bornée (et donc continue) de norme $\|L\|_{V'} = \|f\|_V$.

L'application $a\colon V \times V \to \mathbb{R}$ est continue par la proposition 10.79. Elle est coercive parce que

$$a(u, u) = s\|u\|^2. \tag{25.189}$$

21. Définition 17.33.

Les hypothèses du théorème de Lax-Milgram 25.61 sont réunies et il existe un unique élément $u \in V$ tel que $a(u,v) = L(v)$ pour tout v. Cet élément n'est autre que

$$u = \frac{1}{s} f. \tag{25.190}$$

Nous avons alors $\|u\| = \frac{1}{s}\|f\|$.

Mais s est la constante de coercivité de a et $\|f\|$ est la norme de L. Cette dernière affirmation est contenue dans le fait que Φ est une isométrique dans le lemme 25.16. En reprenant les notations du théorème, nous avons l'égalité

$$\|u\| = \frac{\|L\|}{\alpha}, \tag{25.191}$$

ce qui signifie que l'inégalité du théorème est dans un certain sens optimale. △

PROPooFEOZooTNPcBJ

Proposition 25.63.
Soit un espace de Hilbert réel V et une application $a\colon V \times V \to \mathbb{R}$ coercive et continue.

L'application $\phi\colon V' \to V$ qui à L fait correspondre l'unique solution u de $a(u,v) = L(v)$ est

(1) bien définie,

(2) linéaire,

(3) continue.

Démonstration. Le fait que ϕ soit bien définie est le théorème de Lax-Milgram 25.61.

L'application ϕ dont nous parlons ici est une application entre deux espaces normés. Soient $L_1, L_2 \in V'$; nous définissons $u_1 = \phi(L_1)$ et $u_2 = \phi(L_2)$. Nous avons

$$a(u_1, v) = L_1(v) \tag{25.192a}$$
$$a(u_2, v) = L_2(v) \tag{25.192b}$$

pour tout $v \in V$. Par bilinéarité de a nous avons aussi

$$a(u_1 + u_2, v) = (L_1 + L_2)(v). \tag{25.193}$$

Vu que $L_1 + L_2$ est également bornée, $u_1 + u_2 = \phi(L_1 + L_2)$. Pour le même type de raisons, si $\lambda \in \mathbb{R}$,

$$a(\lambda u, v) = (\lambda L)(v), \tag{25.194}$$

c'est-à-dire que $\phi(\lambda L) = \lambda\phi(L)$.

Donc ϕ est linéaire. Pour la continuité, il suffit de s'assurer que ϕ est bornée. Pour cela nous utilisons la majoration (25.173) et la définition de la norme d'une application linéaire :

$$\|\phi\| = \sup_{\|L\|=1} \|\phi(L)\| \leqslant \sup_{\|L\|=1} \frac{1}{\alpha}\|L\| = \frac{1}{\alpha}. \tag{25.195}$$

Donc ϕ est bornée et par voie de conséquence continue (proposition 10.70). □

Un problème est bien posé au sens de Hadamard si la solution est unique et dépend continument des données ; nous ne rentrerons pas plus dans les détails pour l'instant : nous en reparlerons dans la définition 33.14 mais sachez qu'il existe en fin de compte autant de Hadamard qu'il y a de contextes dans lequel le mot « problème » a un sens.

Ce que dit la proposition 25.63 est que, a étant donné, le problème qui consiste à résoudre $a(u,v) = L(v)$ est bien posé au sens de Hadamard, au moins par rapport à L.

Théorème 25.64 (Lax-Milgram version symétrique[631])**.**
Nous considérons les mêmes hypothèses que celles du théorème de Lax-Milgram, c'est-à-dire un espace de Hilbert réel V muni de différentes choses.

(1) L'application linéaire $L\colon V \to \mathbb{R}$ qui est bornée sur V. Nous notons C sa norme.

(2) *La forme bilinéaire continue a sur $V\times V$. Nous considérons $M>0$ tel que $|a(u,v)|\leqslant M\|u\|\|v\|$ pour tout $u,v\in V$.*

(3) *La forme a est coercitive.*

Nous supposons de plus que a est symétrique.

Alors un élément $u\in V$ est tel que $a(u,w)=l(w)$ pour tout $w\in V$ si et seulement si il minimise la ***fonctionnelle d'énergie***

$$J(w)=\frac{1}{2}a(w,w)-l(w). \tag{25.196}$$

Démonstration. Nous séparons les deux sens.

(1) $\Rightarrow$ Soit u, un élément vérifiant $a(u,w)=l(w)$ pour tout $w\in V$. Soit aussi un élément quelconque $w\in V$, et montrons que $J(u+w)\geqslant J(u)$ (tout élément de V peut être écrit sous la forme $u+w$). Nous avons :

$$J(u+w)=\frac{1}{2}\big(a(u,u)+a(u,w)+a(w,u)+a(w,w)\big)-l(u)-l(w) \tag{25.197a}$$

$$=J(u)+\underbrace{a(u,w)-l(w)}_{=0}+\frac{1}{2}a(w,w) \tag{25.197b}$$

$$=J(u)+\frac{1}{2}a(w,w) \tag{25.197c}$$

Vu que a est coercive, le second terme est positif et nous avons

$$J(u+w)\geqslant J(u), \tag{25.198}$$

ce qu'il fallait.

(2) $\Leftarrow$

Soit $u\in V$, un élément minimisant la fonctionnelle J. Nous fixons $w\in V$ et considérons la fonction $g\colon\mathbb{R}\to\mathbb{R}$ définie par

$$g(\epsilon)=J(u+\epsilon w). \tag{25.199}$$

En développant un peu et en regroupant les termes,

$$g(\epsilon)=J(u)+\epsilon\big(a(u,w)-l(w)\big)+\frac{\epsilon^2}{2}a(w,w). \tag{25.200}$$

EQooHGDOooUMCRda

Cela est une fonction éminemment continue et dérivable ; en réalité c'est un bête polynôme de degré deux. Vu que u minimise J, pour tout $\epsilon\neq 0$ nous avons $g(\epsilon)\geqslant g(0)$ ou encore : $\epsilon=0$ est minimum local (et même global) de g. Le polynôme (25.200) prend son minimum en $\epsilon=0$ si et seulement si

$$a(u,w)-l(w)=0. \tag{25.201}$$

Vous ne me croyez pas ? Faites $g'(\epsilon)=0$ ou bien reprenez la formule du $-b/2a$ pour le sommet d'une parabole, en tenant compte que $a(w,w)\neq 0$. Notons qu'ici encore le fait que a soit coercive joue parce que c'est cela qui nous permet d'affirmer que la parabole a un minimum et non un maximum.

□

Chapitre 26

Analyse complexe

ChapICHIooXbLccl

26.1 Fonctions holomorphes

La dérivée complexe est discutée à la section 12.28, et la définition d'une fonction holomorphe est 12.305.

26.1.1 Équations de Cauchy-Riemann

Notons que la formule (12.844) donne un **développement limité** pour les fonctions holomorphes. Si f est holomorphe en z_0 alors si z est dans un voisinage de z_0, il existe une fonction $s\colon \mathbb{R} \to \mathbb{C}$ telle que $\lim_{t\to 0} s(t)/t = 0$ et

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|). \tag{26.1}$$ EqptwBEG

Nous introduisons les opérateurs

$$\frac{\partial}{\partial z} = \partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \tag{26.2a}$$

$$\frac{\partial}{\partial \bar{z}} = \bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \tag{26.2b}$$

Si f est une fonction $\mathbb{C}$-dérivable représentée par la fonction $F = P + iQ$, les équations de Cauchy-Riemann signifient que $\Delta P = \Delta Q = 0$, c'est-à-dire que les composantes de la fonction f sont harmoniques [1].

THOooHXKMooKYkOjt

Théorème 26.1.
Si $f \in C^1(\Omega)$ alors nous avons équivalence des faits suivants :

(1) f est holomorphe sur Ω,

(2) f vérifie $\partial_{\bar{z}} f = 0$.

PropkwIQwg

Proposition 26.2.
Une application $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-dérivable sur Ω si et seulement si elle est différentiable et EqmblExI

$$\begin{cases} \dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y} & (26.3a) \\ \dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x} & (26.3b) \end{cases}$$

où $f(x + iy) = u(x, y) + iv(x, y)$. Ces équations se notent de façon plus compacte

$$\frac{\partial f}{\partial \bar{z}} = 0. \tag{26.4}$$

*Ces équations sont les équations de **Cauchy-Riemann**.*

1. Une fonction u est harmonique si $\Delta u = 0$.

Démonstration. La différentielle de $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ est donnée par la matrice

$$T = \begin{pmatrix} \partial_x u(a) & \partial_y u(a) \\ \partial_x v(a) & \partial_y v(a) \end{pmatrix}. \tag{26.5}$$

EQwtagsz

Cette matrice est une similitude si et seulement si les équations de Cauchy-Riemann sont satisfaites. En effet si $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, la matrice T est une similitude (écrivons $\alpha + i\beta$ son coefficient) si

$$\begin{cases} T(1) = \alpha + i\beta & \text{(26.6a)} \\ T(i) = -\beta + i\alpha, & \text{(26.6b)} \end{cases}$$

c'est-à-dire

$$T = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{26.7}$$

Identifier cette matrice à (26.5) fournit le résultat annoncé. □

PROPooAGGMooIVQFQB

Proposition 26.3 (Cauchy-Riemann en coordonnées polaires[1, 633])**.**
Soit une fonction $f\colon \mathbb{C} \to \mathbb{C}$ que nous supposons être $\mathbb{C}$-dérivable dans un voisinage de $r_0 e^{i\theta_0}$ (r_0 et θ_0 sont des réels). Nous posons

$$\begin{aligned} \tilde{f}\colon \mathbb{R} \times \mathbb{R} &\to \mathbb{C} \\ r, \theta &\mapsto f(re^{i\theta}). \end{aligned} \tag{26.8}$$

Alors

ITEMooRTYYooSTgTAQ

(1) La fonction $\tilde{f}$ admet des dérivées partielles dans les deux directions.

ITEMooDHXTooBjxwjY

(2) Les dérivées partielles de $\tilde{f}$ sont liées par la relation

$$\frac{\partial \tilde{f}}{\partial \theta}(r_0, \theta_0) = i r_0 \frac{\partial \tilde{f}}{\partial r}(r_0, \theta_0). \tag{26.9}$$

ITEMooUUXTooZoDMHI

(3) La fonction $\tilde{f}$ est de classe C^∞.

Démonstration. Nous considérons les fonctions réelles u et v donnant les parties réelles et imaginaires de f :

$$f(x + iy) = u(x, y) + iv(x, y). \tag{26.10}$$

Nous avons :

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial r}(r_0, \theta_0) = {} & \frac{\partial u}{\partial x}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) \cos(\theta_0) \\ & + \frac{\partial u}{\partial y}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) \sin(\theta_0) \\ & + i\frac{\partial v}{\partial x}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) \cos(\theta_0) \\ & + i\frac{\partial v}{\partial y}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) \sin(\theta_0) \end{aligned} \tag{26.11}$$

et

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial \theta}(r_0, \theta_0) = {} & -\frac{\partial u}{\partial x}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) r_0 \sin(\theta_0) \\ & + \frac{\partial u}{\partial y}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) r_0 \cos(\theta_0) \\ & - \frac{\partial v}{\partial x}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) r_0 \sin(\theta_0) \\ & + i\frac{\partial v}{\partial y}\big(r_0 \cos(\theta_0), r_0 \sin(\theta_0)\big) r_0 \cos(\theta_0). \end{aligned} \tag{26.12}$$

Nous exprimons tout en termes de $\partial_y v$ et $\partial_x v$ en utilisant les équations de Cauchy-Riemann de la proposition 26.2 ainsi que la formule $\cos(\theta)+i\sin(\theta)=e^{i\theta}$ du lemme 18.11. Pour simplifier les notations, nous notons $a_0=\big(r_0\cos(\theta_0),r_0\sin(\theta_0)\big)\in\mathbb{R}^2$. Après quelques calculs :

$$(\partial_r\tilde{f})(r_0,\theta_0)=e^{i\theta_0}\frac{\partial v}{\partial y}\big(r_0\cos(\theta_0),r_0\sin(\theta_0)\big)+ie^{i\theta_0}\frac{\partial v}{\partial x}\big(r_0\cos(\theta_0),r_0\sin(\theta_0)\big) \tag{26.13a}$$

$$=e^{i\theta_0}\left(\frac{\partial v}{\partial y}(a_0)+i\frac{\partial v}{\partial x}(a_0)\right), \tag{26.13b}$$

et

$$(\partial_\theta\tilde{f})(r_0,\theta_0)=r_0e^{i\theta_0}\left(i\frac{\partial v}{\partial y}(a_0)-\frac{\partial v}{\partial x}(a_0)\right) \tag{26.14}$$

En comparant les deux, nous trouvons

$$(\partial_\theta\tilde{f})(r_0,\theta_0)=r_0i(\partial_r\tilde{f})(r_0,\theta_0). \tag{26.15}$$

□

PROPooCHUEooYsGcQK

Proposition 26.4.
Si $f\colon\mathbb{C}\to\mathbb{C}$ est holomorphe, alors nous avons

$$df_{z_0}=(\partial_z f)(z_0) \tag{26.16}$$

au sens où l'opérateur linéaire $df_{z_0}\colon\mathbb{C}\to\mathbb{C}$ est l'opération de multiplication par le nombre complexe $(\partial_z f)(z_0)$.

Démonstration. Soit $f(x+iy)=f_1(x,y)+if_2(x,y)$ une fonction holomorphe[2]. Les fonctions réelles f_1 et f_2 sont assujetties aux équations de Cauchy-Riemann de la proposition 26.2 :

$$\begin{cases}\partial_x f_1=\partial_y f_2 & \text{(26.17a)}\\ \partial_x f_2=-\partial_y f_1. & \text{(26.17b)}\end{cases}$$

Nous avons, en recourant à un petit abus de notation entre $f_i\colon\mathbb{R}^2\to\mathbb{R}$ et $f_i\colon\mathbb{C}\to\mathbb{R}$:

$$df_{z_0}(u)=\frac{d}{dt}\Big[f(z_0+tu)\Big]_{t=0} \tag{26.18a}$$

$$=\frac{d}{dt}\Big[f_1(z_0+tu)+if_2(z_0+tu)\Big]_{t=0} \tag{26.18b}$$

$$=\partial_x f_1u_1+\partial_y f_1u_2+i\big(\partial_x f_2u_1+\partial_y f_2u_2\big) \tag{26.18c}$$

$$=\begin{pmatrix}\partial_x f_1 & \partial_y f_1\\ \partial_x f_2 & \partial_y f_2\end{pmatrix}\begin{pmatrix}u_1\\u_2\end{pmatrix} \tag{26.18d}$$

$$=\begin{pmatrix}\partial_x f_1 & \partial_y f_1\\ -\partial_y f_1 & \partial_x f_1\end{pmatrix}\begin{pmatrix}u_1\\u_2\end{pmatrix}. \tag{26.18e}$$

En utilisant le lemme 12.308 nous reconnaissons la matrice de multiplication par le nombre $\partial_x f_1-i\partial_y f_1$. Or justement,

$$\partial_z f=\frac{1}{2}\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)f=\frac{1}{2}\big(\partial_x f_1+i\partial_x f_2-i\partial_y f_1+\partial_y f_2\big), \tag{26.19}$$

qui se réduit à $\partial_x f_1-i\partial_y f_1$ lorsque nous y appliquons les équations de Cauchy-Riemann. □

2. Définition 12.305.

26.1.2 Intégrale sur un chemin dans $\mathbb{C}$

DEFooBPLJooZwsmxi

Définition 26.5.
Si nous avons une application $\gamma\colon [a,b]\to\mathbb{C}$ et une fonction $f\colon \mathbb{C}\to\mathbb{C}$, nous définissons

$$\int_\gamma f(z)dz=\int_a^b (f\circ\gamma)(t)\gamma'(t)dt \tag{26.20}$$

pour tous les couples (f,γ) pour lesquels le membre de droite a un sens.

26.6.
Vous noterez que cette définition n'est pas exactement la même que celle 20.40 d'une intégrale curviligne en analyse réelle. Cette dernière demande de prendre la norme de γ', alors qu'ici nous la gardons telle quelle. D'ailleurs, gardez en tête que γ est une fonction à valeurs dans $\mathbb{C}$. Donc $\gamma'(t)$ peut très bien être un nombre complexe.

PROPooCUBTooZDcdHX

Proposition-Définition 26.7 (Intégrale sur un chemin C^1 par morceaux)**.**
Soit un chemin C^1 par morceaux $\gamma\colon [a,b]\to\mathbb{C}$. Nous considérons deux subdivisions $a=t_1<t_2<\dots t_n=b$ et $a=u_1<u_2<\dots<u_m=b$ telles que γ soit de classe C^1 sur les intervalles $]t_i,t_{i+1}[$ et $]u_j,u_{j+1}[$. Alors

$$\sum_{i=1}^{n-1}\int_{\gamma\colon]t_i,t_{i+1}[\to\mathbb{C}} f=\sum_{j=1}^{m-1}\int_{\gamma\colon]u_j,u_{j+1}[\to\mathbb{C}} f. \tag{26.21}$$

Cette valeur est notée

$$\int_\gamma f. \tag{26.22}$$

26.1.3 Intégrales sur des chemins fermés

LemtpEOmi

Lemme 26.8.
Si g est une fonction continue dans un ouvert $\Omega\subset\mathbb{C}$ et si g admet une primitive complexe sur Ω alors

$$\int_\gamma g(z)dz=0 \tag{26.23}$$

pour tout chemin fermé γ de classe C^1 contenu dans Ω.

Démonstration. Nommons G une primitive de g. Par définition,

$$\begin{aligned}\int_\gamma g&=\int_\gamma G' && (26.24a)\\ &=\int_0^1 G'\big(\gamma(t)\big)\gamma'(t)dt && (26.24b)\\ &=\int_0^1 (G\circ\gamma)'(t)dt && (26.24c)\\ &=G(\gamma(1))-G\big(\gamma(0)\big) && (26.24d)\\ &=0 && (26.24e)\end{aligned}$$

parce que le chemin est fermé : $\gamma(0)=\gamma(1)$. □

LemwbwbUR

Lemme 26.9 (Goursat[538])**.**
Soit Ω un ouvert dans $\mathbb{C}$ et f une fonction continue sur Ω, holomorphe sur Ω moins éventuellement un point (nommé $z_1\in\Omega$). Soit T, un triangle[3] fermé inclus dans Ω. Alors nous avons

$$\int_{\partial T} f(z)dz=0. \tag{26.25}$$

3. Nous considérons ici le triangle « plein ».

Démonstration. Nous notons $\gamma = \partial T$. Dans la suite nous allons définir une suite de triangles $T^{(n)}$ et nous noterons $\gamma_n = \partial T^{(n)}$ avec une orientation que nous allons expliquer. Pour commencer nous posons $T^{(0)} = T$ et $\gamma_0 = \partial T^{(0)}$.

Nous considérons le cas $z_1 \notin T$, et nous posons

$$c = l(\gamma)^{-2} \left| \int_\gamma f \right|. \tag{26.26}$$

Notre objectif est de montrer que $c = 0$. Soit A, B, C les trois sommets du triangle ; nous divisons le triangle de la façon suivante. D'abord nous considérons les points A', B', C' respectivement milieux de BC, AC et AB. En traçant le triangle $A'B'C'$, nous construisons quatre triangles que nous nommons $T_i^{(0)}$. Le théorème de Thalès [4] assure que le périmètre de chacun des quatre triangles est la moitié du périmètre du grand triangle T.

Sur T nous choisissons l'orientation ABC. De façon à être « compatible », nous choisissons les orientations $AC'B'$, $BA'C'$ et $A'CB'$. La somme de ces trois triangles donne T plus le triangle $A'C'B'$. Par conséquent nous choisissons sur le triangle central l'orientation (inverse) $A'B'C'$ de façon à avoir

$$\int_\gamma f = \sum_{i=1}^{4} \int_{\partial T_i^{(0)}} f. \tag{26.27}$$

Cela implique que pour au moins un des quatre triangles (disons $T_k^{(0)}$ pour fixer les idées) nous ayons

$$\int_{\partial T_k^{(0)}} f \geqslant \frac{1}{4} \int_{\partial T^{(0)}} f \tag{26.28}$$

Nous notons $T^{(1)}$ ce triangle. Comme noté précédemment nous avons

$$l(\partial T^{(1)}) = \frac{1}{2} l(\partial T^{(0)}), \tag{26.29}$$

et donc

$$l(\gamma_1)^{-2} \left| \int_{\gamma_1} f \right| = 4 l(\gamma_0)^{-2} \left| \int_{\gamma_1} f \right| \geqslant 4 l(\gamma_0)^{-2} \frac{1}{4} \left| \int_{\gamma_0} f \right| = c. \tag{26.30}$$

En répétant le procédé nous construisons une suite de triangles $T^{(n)}$ qui satisfont toujours

$$l(\partial T^{(n)}) = \frac{1}{2^n} l(\partial T^{(0)}). \tag{26.31}$$

Ces triangles forment une suite de fermés emboités dont le diamètre tend vers zéro. Leur intersection contient donc exactement un point (lemme 17.125) que nous nommons z_0 (et qui appartient évidemment à Ω). Étant donné que f est holomorphe nous utilisons le développement limité (26.1) autour de z_0 :

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|)(z - z_0) \tag{26.32}$$

avec $\lim_{t \to 0} s(t) = 0$. Nous posons $g(z) = f(z_0) + f'(z_0)(z - z_0)$ et nous considérons $\epsilon > 0$. Soit $\alpha > 0$ tel que

$$|f(z) - g(z)| < \epsilon |z - z_0| \tag{26.33}$$

pour tout $|z - z_0| < \alpha$. Le α à choisir pour obtenir cet effet est celui qui donne $s(|z - z_0|) < \epsilon$. Soit $N \in \mathbb{N}$ tel que $l(\gamma_n) < \alpha$ pour tout $n > N$. D'autre part, deux points dans un triangle sont toujours à distance moindre que la longueur d'un côté, donc pour tout $z \in T^{(n)}$ nous avons $|z - z_0| < \alpha$ et par conséquent pour tout z dans $T^{(n)}$ nous avons

$$|f(z) - g(z)| < \epsilon |z - z_0|. \tag{26.34}$$

Notons que la fonction g est une dérivée : c'est la dérivée de la fonction

$$G(z) = z f(z_0) + \frac{1}{2} f'(z_0)(z - z_0)^2. \tag{26.35}$$

4. Théorème 12.162.

Par conséquent nous avons

$$\int_{\gamma_n} g = 0 \tag{26.36}$$

par le lemme 26.8. Nous avons donc

$$\begin{aligned} |\int_{\gamma_n} f| &= |\int_{\gamma_n} (f-g)| && (26.37a)\\ &\leqslant l(\gamma_n) \max\{|f(z)-g(z)| \text{ tel que } z \in T^{(n)}\} && (26.37b)\\ &\leqslant \epsilon l(\gamma_n)^2, && (26.37c) \end{aligned}$$

et par conséquent

$$c \leqslant l(\gamma_n)^{-2} |\int_{\gamma_n} f| \leqslant \epsilon, \tag{26.38}$$

ce qui signifie que $c = 0$ parce que ϵ est arbitraire. Nous avons donc prouvé le lemme de Goursat dans le cas où le point de non holomorphie z_1 est en dehors de T.

Si z_1 est sur un côté, disons sur le côté AB, alors nous considérons un vecteur $v \in \mathbb{C}$ tel que $T_\epsilon = T + \epsilon v$ ne contienne z_1 pour aucun ϵ. Le vecteur $v = z_1 - C$ fait par exemple l'affaire. En vertu du point précédent nous avons

$$\int_{\partial T_\epsilon} f = 0 \tag{26.39}$$

pour tout $\epsilon > 0$. Étant donné que la fonction f est continue (y compris en z_1), l'intégrale sur ∂T est également nulle.

Si maintenant le point z_1 est à l'intérieur de T nous décomposons T en trois triangles ayant z_1 comme sommet commun. Si nous considérons les orientations Az_1C, ABz_1 et BCz_1, alors nous avons

$$\int_T f = \int_{Az_1C} f + \int_{ABz_1} f + \int_{BCz_1} f, \tag{26.40}$$

alors que par le point précédent les trois intégrales du membre de droite sont nulles. □

PrpopwQSbJg

Proposition 26.10 ([538])**.**
Soient Ω un ouvert étoilé et f une fonction holomorphe sur Ω sauf éventuellement en un point z_1 où f est seulement continue. Alors si γ est un chemin fermé dans Ω, nous avons

$$\int_\gamma f = 0. \tag{26.41}$$

PropRZCKeO

Proposition 26.11.
Si $f(z) = \sum_n a_n z^n$ a pour rayon de convergence R, alors f est $\mathbb{C}$-dérivable et nous pouvons dériver terme à terme dans la boule ouverte $B(0,R)$.

Démonstration. Cela est exactement la proposition 15.44. □

26.1.4 Homotopie entre applications

DEFooPJKLooCvgxsu

Définition 26.12 (homotopie entre applications)**.**
*Soient des espaces topologiques X et Y. Deux applications continues $f_1, f_2 \colon X \to Y$ sont **homotopes** si il existe une application continue $H \colon [0,1] \times X \to Y$ telle que pour tout $x \in X$ nous avons*

$$\begin{aligned} H(0,x) &= f_1(x) && (26.42a)\\ H(1,x) &= f_2(x). && (26.42b) \end{aligned}$$

LEMooMGFZooGOaGYl

Lemme 26.13.
La relation « être homotope à[5] » est une relation d'équivalence sur $C(X,Y)$.

5. Définition 26.12.

Démonstration. Pour obtenir $f \sim f$, il suffit de prendre $H(s,x) = f(x)$.

Si $f \sim g$, nous considérons l'homotopie $H\colon [0,1] \times X \to Y$. Alors l'application

$$\begin{aligned} M\colon [0,1] \times X &\to Y \\ (t,x) &\mapsto H(1-t,x) \end{aligned} \tag{26.43}$$

est une homotopie pour $g \sim f$.

Si $f \sim g$ et $g \sim f$, nous avons les applications continues H et M telles que

$$H(0,x) = f(x) \qquad H(1,x) = g(x) \tag{26.44a}$$

$$M(0,x) = g(x) \qquad M(1,x) = h(x). \tag{26.44b}$$

L'application

$$\begin{aligned} S\colon [0,1] \times X &\to Y \\ (t,x) &\mapsto \begin{cases} H(2t,x) & \text{si } t \in [0, \frac{1}{2}[\\ M(2t-1,x) & \text{si } t \in [\frac{1}{2}, 1]. \end{cases} \end{aligned} \tag{26.45}$$

Nous vérifions que S est continue en vérifiant la valeur en $t = 1/2$. De plus

$$S(0,x) = H(0,x) = f(x) \tag{26.46}$$

et

$$S(1,x) = M(2-1,x) = M(1,x) = h(x). \tag{26.47}$$

□

LEMooMJKEooCaVhjD

Lemme 26.14 ([1])**.**
Soient des espaces topologiques X et Y. Si Y est connexe par arcs, alors toutes les applications constantes $X \to Y$ sont homotopes.

Démonstration. Soient les applications constantes $u(x) = u_0$ et $v(x) = v_0$. Étant donné que Y est connexe par arcs, il existe une application continue $\gamma\colon [0,1] \to Y$ telle que $\gamma(0) = u_0$ et $\gamma(1) = v_0$. Alors l'application

$$\begin{aligned} H\colon [0,1] \times X &\to Y \\ (t,x) &\mapsto \gamma(t) \end{aligned} \tag{26.48}$$

est une homotopie entre les applications u et v. □

PROPooNABDooFtKukO

Proposition 26.15 ([205])**.**
Soit un compact K de $\mathbb{R}^n$. Deux applications $f, g\colon K \to \mathbb{C}^*$ sont homotopes dans $\mathbb{C}^*$ si et seulement si f/g est homotope à la fonction

$$\begin{aligned} u\colon K &\to \mathbb{C} \\ x &\mapsto 1. \end{aligned} \tag{26.49}$$

Démonstration. Supposons que $H\colon [0,1] \times K \to \mathbb{C}^*$ est une homotopie entre f et g. Dans ce cas, l'application

$$\begin{aligned} M\colon [0,1] \times K &\to \mathbb{C}^* \\ (t,x) &\mapsto \frac{H(t,x)}{g(x)} \end{aligned} \tag{26.50}$$

est une homotopie entre f/g et u. En effet

$$M(0,x) = \frac{H(0,x)}{g(x)} = \frac{f(x)}{g(x)} \tag{26.51}$$

et

$$M(1,x) = \frac{H(1,x)}{g(x)} = 1 = u(x). \tag{26.52}$$

Dans l'autre sens, si H est une homotopie entre f/g et u, alors l'application $M(t,x) = H(t,x)g(x)$ est une homotopie entre f et g. □

26.1.5 Intégrale et homotopie

LEMooOFCEooIsuchR

Lemme 26.16 ([205]).
Soit un rectangle fermé $R \subset \mathbb{R}^2$. Si l'application $H : R \to \Omega$ est continue, alors pour toute fonction continue $f : \Omega \to \mathbb{C}$ nous avons

$$\int_{\partial H} f = 0 \tag{26.53}$$

où ∂H désigne la frontière de $H(R)$ dans $\mathbb{C}$, et l'intégrale est celle de la définition 26.7.

Démonstration. La frontière ∂H peut être décomposée en 4 parties de classe C^1 :

(1)
$$\begin{aligned} \alpha_1 : [0,1] &\to \Omega \\ t &\mapsto H(t,a), \end{aligned} \tag{26.54}$$

(2)
$$\begin{aligned} \alpha_2 : [a,b] &\to \Omega \\ t &\mapsto H(1,t) \end{aligned} \tag{26.55}$$

(3)
$$\begin{aligned} \alpha_3 : [0,1] &\to \Omega \\ t &\mapsto H(1-t,b) \end{aligned} \tag{26.56}$$

(4)
$$\begin{aligned} \alpha_4 : [a,b] &\to \Omega \\ t &\mapsto H\big(0,b+a-t\big). \end{aligned} \tag{26.57}$$

Étudions les diverses parties α_i.

(1) **Pour α_2** Première observation : $\alpha_2 = \gamma_1$.

(2) **Pour α_4** Nous avons $\alpha_4 = \gamma_0(b+a-t)$ et donc

$$\int_{\alpha_4} f = \int_a^b (f \circ \alpha_4)(t)\alpha_4'(t) = -\int_a^b (f \circ \gamma_0)(b+a-t)\gamma_0'(b+a-t). \tag{26.58}$$ EQooFWIQooJALDpQ

Nous utilisons le changement de variables [6]

$$\begin{aligned} \phi : [a,b] &\to [a,b] \\ t &\mapsto b+a-t. \end{aligned} \tag{26.59}$$

Le jacobien $|J_\phi|$ est égal à 1 et donc nous continuons (26.58) de la façon suivante :

$$\begin{aligned} \int_{\alpha_4} f &= -\int_a^b (f \circ \gamma_0)\big(\phi(t)\big)\gamma'\big(\phi(t)\big)dt && \text{(26.60a)} \\ &= -\int_a^b (f \circ \gamma_0)(t)\gamma_0'(t)dt && \text{(26.60b)} \\ &= -\int_{\gamma_0} f. && \text{(26.60c)} \end{aligned}$$

(3) **Lien entre α_1 et α_3** Il y a deux possibilités : soit γ_1 et γ_2 sont des lacets, soit ce sont des chemins normaux.

6. Théorème 14.281(3).

(3a) **Si ils sont des lacets** Alors pour tout s, l'application

$$\begin{aligned}\Gamma_s\colon [a,b] &\to C\\ t &\mapsto H(s,t)\end{aligned} \tag{26.61}$$

est un lacet. En particulier $H(s,a)=H(s,b)$. Donc $\alpha_3(t)=H(1-t,b)=\alpha_1(1-t)$. Dans ce cas nous avons

$$\begin{aligned}\int_{\alpha_3} f &= \int_a^b f\big(\alpha_3 t\big)\alpha_3'(t) && \text{(26.62a)}\\ &= \int_a^b f\big(\alpha_1(1-t)\big)(-)\alpha_1'(1-t) && \text{(26.62b)}\\ &= -\int_a^b (f\circ\alpha_1)(t)dt && \text{chm. var. } \phi(t)=1-t \quad \text{(26.62c)}\\ &= -\int_{\alpha_1} f. && \text{(26.62d)}\end{aligned}$$

Nous avons alors le calcul

$$0=\int_{\partial H} f=\int_{\alpha_1} f+\int_{\gamma_1} f+\underbrace{\int_{\alpha_3} f}_{-\int_{\alpha_1} f}-\int_{\gamma_0} f=\int_{\gamma_1} f-\int_{\gamma_0} f. \tag{26.63}$$

Le lemme est prouvé.

(3b) **Si γ_0 et γ_1 sont des chemins** Dans ce cas nous avons une homotopie à extrémités fixées. Donc α_1 et α_3 sont des chemins constants, et les intégrales dessus sont nulles et

$$0=\int_{\partial H} f=\underbrace{\int_{\alpha_1} f}_{=0}+\int_{\gamma_1} f+\underbrace{\int_{\alpha_3} f}_{=0}-\int_{\gamma_0} f=\int_{\gamma_1} f-\int_{\gamma_0} f. \tag{26.64}$$

Et le lemme est prouvé.

□

THOooVTFXooBgvVyD

Théorème 26.17 (Cauchy, version homotopique[205, 634])**.**
Soit un ouvert Ω dans $\mathbb{C}$. Soit une fonction holomorphe $f\colon \Omega\to\mathbb{C}$.

Supposons que l'une des deux conditions suivantes soit respectée :

(1) Les chemins $\gamma_0,\gamma_1\colon [a,b]\to\Omega$ sont homotopes à extrémités fixées.

ITEMooESDVooFgVarr

(2) Les lacets $\gamma_0,\gamma_1\colon [a,b]\to\Omega$ sont homotopes.

Alors

$$\int_{\gamma_0} f=\int_{\gamma_1} f. \tag{26.65}$$

Démonstration. Considérons l'homotopie $H\colon [0,1]\times[a,b]\to\Omega$. Étant donné que $[0,1]\times[a,b]$ est un rectangle dans $\mathbb{R}^2$, le lemme 26.16 nous indique que

$$\int_{\partial H} f=0. \tag{26.66}$$

□

CorGZXzuZR

Corolaire 26.18 ([634])**.**
Soient $a\in\mathbb{C}$ ainsi que deux chemins γ_1 et γ_2 homotopes dans $\mathbb{C}\backslash\{a\}$. Alors $\mathrm{Ind}(\gamma_1,a)=\mathrm{Ind}(\gamma_2,a)$.

26.1.6 Théorème de Tietze (espace normal)

LEMooCLVAooTaNGJk

Lemme 26.19 ([205]).
Soit un espace topologique normal[7] X. *Soient* $M \in \mathbb{R}^+$ *et* A *fermé dans* X, *et une application continue* $f\colon A \to [-M, M]$.

Il existe une application continue $g\colon X \to [-M/3, M/3]$ *telle que pour tout* $x \in A$,

$$|f(x) - g(x)| < \frac{2M}{3}, \tag{26.67}$$

Démonstration. Nous divisons A en trois parties :

$$A_- = f^{-1}\big([-M, -M/3]\big), \tag{26.68a}$$
$$A_+ = f^{-1}\big([M/3, M/3]\big), \tag{26.68b}$$
$$A_0 = f^{-1}\big([-M/3, M/3]\big). \tag{26.68c}$$

Les parties A_+ et A_- sont fermées parce que f est continue. D'autre part, X est normal, de telle sorte que le théorème d'Urysohn 21.113 s'applique.

Nous considérons donc une application continue $g_1\colon X \to [0, 1]$ telle que $g_1^{-1}(A_-) = \{0\}$ et $g_1^{-1}(A_+) = \{1\}$. Nous posons alors

$$g(x) = \left(\frac{2M}{3}\right) g_1(x) - \frac{M}{3}. \tag{26.69}$$

Vérification des propriétés de g.

(1) g **est continue** Parce que g_1 est continue.

(2) g **prend ses valeurs dans** $[-M/3, M/3]$ Parce que g_1 prend ses valeurs dans $[0, 1]$ et que $t \mapsto (2M/3)t - M/3$ est croissante.

(3) $|f(x) - g(x)|$ **sur** A Si $x \in A_-$, alors $g_1(x) = 0$ et donc $g(x) = -M/3$. Donc

$$|f(x) - g(x)| = |f(x) - M/3| \leqslant 2M/3 \tag{26.70}$$

EQooEUOKooFLRJiz

parce que $f(x) \in [-M/3, M/3]$.

Si $x \in A_+$, alors $g_1(x) = 1$ et donc $g(x) = M/3$. Même fin de raisonnement qu'en (26.70).

Si $x \in A_0$, alors $f(x) \in [-M/3, M/3]$ et $g(x) \in [-M/3, M/3]$ et donc encore $|f(x) - g(x)| \leqslant 2M/3$.

□

LEMooSKSNooEdgFcR

Lemme 26.20 ([205]).
Soient un espace topologique normal X *ainsi qu'un fermé* A *dans* X. *Nous considérons une fonction continue* $f\colon A \to [-M, M]$.

Il existe une application $g\colon X \to [-M, M]$ *continue prolongeant* f.

Démonstration. Nous allons commencer par construire une suite d'applications $g_i\colon X \to \mathbb{R}$ telles que

ITEMooGAIVooQYBCZj

(1) Pour tout $x \in A$ et pour tout $n \in \mathbb{N}$,

$$\Big|f(x) - \sum_{i=1}^{r} n g_i(x)\Big| \leqslant \left(\frac{2}{3}\right)^n M \tag{26.71}$$

ITEMooOLNAooEJPdbV

(2) Pour tout $x \in X$ et pour tout i,

$$|g_i(x)| \leqslant \left(\frac{2}{3}\right)^i \frac{M}{2}. \tag{26.72}$$

7. Définition 21.108.

Nous commençons par construire g_1 à partir du lemme 26.19 appliqué à la fonction f. Nous avons donc une application $g_1 : X \to [-M/3, M/3]$ telle que

$$|f(x) - g_1(x)| \leqslant \frac{2M}{3} \tag{26.73}$$

EQooGXUIooWnATPw

pour tout $x \in A$. Vu que g_1 prend ses valeurs dans $[-M/3, M/3]$, elle vérifie la condition (2). De plus (26.73) montre que la condition (1) est vérifiée pour $n = 1$.

Et c'est parti pour la récurrence. Nous supposons avoir des applications g_i pour $i = 1, \dots, k$ qui vérifient la condition (2), et telles que la condition (1) est satisfaite pour $n = 1, \dots, k$. Nous allons maintenant construire g_{k+1}.

Nous posons

$$\begin{aligned} h_k : A &\to \mathbb{R} \\ x &\mapsto f(x) - \sum_{i=1}^{k} g_i(x). \end{aligned} \tag{26.74}$$

Par hypothèse de récurrence, la fonction h_k ne prend pas n'importe quelles valeurs dans $\mathbb{R}$, mais

$$h_k : A \to \Big[-\Big(\frac{2}{3}\Big)^k M, \Big(\frac{2}{3}\Big)^k M\Big]. \tag{26.75}$$

Nous construisons g_{k+1} à partir de ce h_k et du lemme 26.19. Nous avons donc

$$g_{k+1} : X \to \Big[-\frac{1}{3}\Big(\frac{2}{3}\Big)^k M, \frac{1}{3}\Big(\frac{2}{3}\Big)^k M\Big] \tag{26.76}$$

vérifiant

$$|h_k(x) - g_{k+1}(x)| \leqslant \Big(\frac{2}{3}\Big)^{k+1} M. \tag{26.77}$$

La condition (1) est donc maintenant vérifiée jusqu'à $n = k + 1$. Nous vérifions la condition (2) pour $i = k + 1$. Simple calcul :

$$|g_{k+1}(x)| \leqslant \frac{1}{3}\Big(\frac{2}{3}\Big)^k M = \Big(\frac{2}{3}\Big)^{k+1} \frac{M}{2}. \tag{26.78}$$

Et voilà pour la définition des applications g_i.

Vu que $2/3 < 1$, la série $\sum_{i=1}^{\infty} \|g_i\|_\infty$ converge normalement (définition 10.93). Notons g la somme. Le lemme 10.95 donne alors la convergence uniforme $g_i \xrightarrow{unif} g$, et le théorème 12.370 nous assure que g est continue sur X.

En ce qui concerne la norme de g, nous avons, en utilisant la formule (10.344) avec $q = 2/3$,

$$\|g\|_\infty \leqslant \sum_{i=1}^{\infty} \Big(\frac{2}{3}\Big)^i \frac{M}{2} \leqslant M. \tag{26.79}$$

Donc $|g(x)| \leqslant M$ pour tout $x \in X$.

Enfin nous vérifions que g prolonge f. Soit $x \in A$. Prenez la limite $n \to \infty$ dans l'inégalité

$$|f(x) - \sum_{i=1}^{n} g_i(x)| \leqslant \Big(\frac{2}{3}\Big)^n M. \tag{26.80}$$

Nous trouvons que $|f(x) - g(x)| = 0$. □

THOooXKGWooFUYlux

Théorème 26.21 (Théorème de Tietze)**.**
Soit une partie fermée A de l'espace normal X. Si la fonction $f : A \to \mathbb{R}$ est continue, alors elle se prolonge en une fonction continue $g : X \to \mathbb{R}$.

Démonstration. Nous supposons dans un premier temps que f prenne ses valeurs dans $]-M,M[$. À fortiori, elle prend ses valeurs dans $[-M,M]$ et le lemme 26.20 dit qu'il existe un prolongement continu $g\colon X\to[-M,M]$. Nous allons construire à partir de là un prolongement continu $h\colon X\to]-M,M[$ de f.

Nous posons

$$B=g^{-1}\big(\{-M,M\}\big). \tag{26.81}$$

Étant donné que $f(A)\subset]-M,M[$, nous avons $B\cap A=\varnothing$. De plus A est fermé par hypothèse et B est fermé en tant qu'image réciproque du fermé $\{-M,M\}$ par l'application continue g. Nous pouvons donc appliquer le théorème d'Urysohn 21.113.

Nous considérons donc une application $g_1\colon X\to[0,1]$ telle que $g_1=0$ sur B et $g_1=1$ sur A. Enfin nous posons $h=gg_1$. Cette application prend ses valeurs dans $]-M,M[$, est continue et si $x\in A$ nous avons

$$h(x)=g(x)g_1(x)=f(x)\times 1=f(x). \tag{26.82}$$

Ceci règle la question si f prend ses valeurs dans $]-M,M[$.

Nous considérons à présent le cas général $f\colon A\to\mathbb{R}$. Soit un homéomorphisme $\phi\colon\mathbb{R}\to]-M,M[$ (par exemple via la proposition 12.179). Nous considérons

$$\begin{aligned}\tilde{f}\colon A&\to]-M,M[\\ x&\mapsto(\phi\circ f)(x).\end{aligned} \tag{26.83}$$

Nous lui appliquons le premier cas pour avoir une fonction $\tilde{g}\colon X\to]-M,M[$ qui prolonge $\tilde{f}$. Il suffit maintenant de poser

$$\begin{aligned}g\colon X&\to\mathbb{R}\\ x&\mapsto(\phi^{-1}\circ\tilde{g})(x).\end{aligned} \tag{26.84}$$

Cela est une application continue et si $x\in A$, nous avons

$$g(x)=(\phi^{-1}\circ\tilde{g})(x)=(\phi^{-1}\circ\tilde{f})(x)=(\phi^{-1}\circ\phi\circ f)(x)=f(x). \tag{26.85}$$

□

Corolaire 26.22.
Soit un fermé A dans un espace normal X. Si $f\colon A\to\mathbb{C}$ est continue, alors elle se prolonge en une fonction continue $g\colon X\to\mathbb{C}$.

Démonstration. Les parties réelles et imaginaires de f sont continues. Il suffit de leur appliquer le théorème de Tietze 26.21. □

26.2 Logarithme complexe

26.2.1 La fonction argument

Nous savons la définition 18.6 de l'exponentielle complexe.

Définition 26.23.
*Un **logarithme** de $\alpha\in\mathbb{C}$ est une solution de l'équation $e^z=\alpha$.*

Notons bien que cela définit *un* logarithme, et non *le* logarithme.

LEMooUMESooJVzeDb

Lemme 26.24.
Si z_1 et z_2 sont des logarithmes de α alors il existe $k\in\mathbb{Z}$ tel que $z_1=z_2+2ik\pi$.

Démonstration. Nous commençons par déterminer les logarithmes de $\alpha=1$. Nous avons besoin de $e^{a+bi}=1$ $(a,b\in\mathbb{R})$. Nous avons

$$e^ae^{bi}=1, \tag{26.86}$$

et en prenant la norme nous trouvons $|e^a| = 1$, ce qui donne $a = 0$. Ensuite $e^{bi} = 1$, qui signifie $b = 2k\pi$. Les logarithmes de 1 sont donc les nombres de la forme $2ik\pi$.

Soient maintenant z_1 et z_2 des logarithmes de α. Alors $e^{z_1} = e^{z_2}$, donc [8] $e^{z_1 - z_2} = 1$, ce qui signifie que $z_1 - z_2$ est un logarithme de 1. Donc il existe un $k \in \mathbb{Z}$ tel que $z_1 - z_2 = 2ik\pi$. □

Remarque 26.25.
Jusqu'ici nous n'avons pas donné de conditions donnant l'existence d'un logarithme. Nous avons seulement supposé des existences et donné des propriétés sur ces hypothétiques objets.

Définition 26.26 ([635])**.**
Si $z \in \mathbb{C}^*$ nous définissons la ***valeur principale*** *de son argument le nombre $\theta \in]-\pi, \pi]$ tel que*

$$z = |z|e^{i\theta} \tag{26.87}$$

Nous le notons $\arg(z)$.

NORMooOGHNooYriCBH

26.27.
Il ne faut pas se ruer sur $\arg(x + iy) = \arctan(y/x)$. Pour rappel, la fonction arctan a été définie dans le théorème 18.38, et elle prend ses valeurs dans $]-\pi/2, \pi/2[$. La formule $v(x, y) = \arctan(y/x)$ n'est donc valable que pour $x > 0$. Les valeurs sont :

$$\arg(x + iy) = \begin{cases} \arctan(y/x) & \text{si } x > 0 \\ \pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y \geqslant 0 \\ -\pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y < 0 \\ \frac{\pi}{2} & \text{si } x = 0 \text{ et } y > 0 \\ \frac{-\pi}{2} & \text{si } x = 0 \text{ et } y < 0. \end{cases} \tag{26.88}$$

EQooPJVFooSEKTny

Pour $x > 0$ nous avons $\arg(x+iy) = \arctan(y/x)$ parce que justement la fonction arctan prend ses valeurs en particulier entre $-\pi$ et π. Pour $x < 0$ et $y > 0$ nous avons $\arg(x+iy) = \pi+\arctan(y/x)$ (dans ce cas, $\arctan(y/x) < 0$) et si $x < 0$, $y < 0$ nous avons $\arg(x + iy) = -\pi + \arctan(y/x)$.

NORMooMRBEooVtTcIA

26.28 (Les dérivées partielles de la fonction argument)**.**
Vu que nous en aurons besoin plusieurs fois, nous calculons maintenant les dérivées partielles de la fonction

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto \arg(x + iy). \end{aligned} \tag{26.89}$$

Nous commençons par la dérivée $\partial_x\varphi(x, y)$. Et il y a de nombreux cas à séparer.

(1) $\underline{x > 0}$
Nous avons

$$\frac{\partial\varphi}{\partial x}(x, y) = \lim_{\epsilon\to 0} \frac{\arctan(y/(x + \epsilon)) - \arctan(y/x)}{\epsilon}, \tag{26.90}$$

qui n'est autre que la dérivée de la fonction $x \mapsto \arctan(y/x)$. Nous pouvons la calculer facilement avec le théorème 18.38(2) :

$$\frac{\partial\varphi}{\partial x}(x, y) = -\frac{y}{x^2 + y^2}. \tag{26.91}$$

(2) $\underline{x < 0}$
Nous avons

$$\frac{\partial\varphi}{\partial x}(x, y) = \lim_{\epsilon\to 0} \frac{\pm\pi + \arctan(y/(x + \epsilon)) - \big(\pm\pi + \arctan(y/x)\big)}{\epsilon} \tag{26.92}$$

où les signes $\pm$ dépendent du signe de y. De toutes façons, les termes en π se simplifient et le calcul est le même que celui du cas $x > 0$. Encore une fois nous avons

$$\frac{\partial\varphi}{\partial x}(x, y) = -\frac{y}{x^2 + y^2}. \tag{26.93}$$

8. C'est facile de dire « donc ». Il faut surtout citer la proposition 18.9(2).

(3) $\underline{x = 0}$

Nous devons calculer

$$\frac{\partial\varphi}{\partial x}(0,y) = \lim_{\epsilon\to 0} \frac{\arg(\epsilon + iy) - \arg(iy)}{\epsilon}. \tag{26.94}$$

Il y a quatre cas d'après les signes de ϵ (séparer limite à gauche et à droite) et y.

Si $\epsilon > 0$ et $y > 0$ alors nous avons à faire le calcul

$$\lim_{\epsilon\to 0^+} \frac{\arctan(y/\epsilon) - \pi/2}{\epsilon} \tag{26.95}$$

qui se traite par la règle de l'Hospital. Cela donne $-1/y$.

Les trois autres cas ne se distinguent que par des constantes au numérateur, lesquelles disparaissent en appliquant la règle de l'Hospital [9]. Au final,

$$\frac{\partial\varphi}{\partial x}(0,y) = -\frac{1}{y}. \tag{26.96}$$

Nous avons calculé jusqu'ici :

$$\frac{\partial\varphi}{\partial x}(x,y) = \frac{-y}{x^2+y^2} \tag{26.97}$$

EQooAOJPooOrvUBR

pour tout $(x,y) \in \mathbb{R}^2\backslash\{(0,0)\}$. En particulier vous avez noté que cette dérivée partielle est continue sur $\mathbb{R}^2\backslash\{(0,0)\}$.

Nous calculons à présent la dérivée partielle par rapport à y :

$$\frac{\partial\varphi}{\partial y}(x,y) = \lim_{\epsilon\to 0} \frac{\arg(x + iy + i\epsilon) - \arg(x + iy)}{\epsilon}. \tag{26.98}$$

(1) $\underline{x > 0}$

Nous avons à calculer

$$\lim_{\epsilon\to 0} \frac{\arctan\frac{y+\epsilon}{x} - \arctan\frac{y}{x}}{\epsilon}, \tag{26.99}$$

qui n'est autre que la dérivée de la fonction $t \mapsto \arctan\frac{t}{x}$ en $t = y$. Résultat :

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2}. \tag{26.100}$$

(2) $\underline{x < 0 \text{ **et** } y \neq 0}$

Le calcul à faire est :

$$\lim_{\epsilon\to 0} \frac{\pm\pi + \arctan\frac{y+\epsilon}{x} - \left(\pm\pi + \arctan\frac{y}{x}\right)}{\epsilon} \tag{26.101}$$

Une chose importante à remarquer est que dans le calcul de la limite nous pouvons supposer que y et $y+\epsilon$ aient le même signe, quelle que soit la valeur et le signe de ϵ (assez petit). C'est pour cela que les deux termes $\pm\pi$ arrivent avec le même signe des deux côtés de la différence, et se simplifient. Nous tombons sur une limite déjà faite et

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2} \tag{26.102}$$

(3) $\underline{x < 0 \text{ **et** } y = 0}$

Vu que $x < 0$ nous avons $\arg(x) = \pi$ et nous devons calculer

$$\lim_{\epsilon\to 0} \frac{\arg(x + i\epsilon) - \pi}{\epsilon}. \tag{26.103}$$

La limite $\epsilon \to 0^+$ est classique et donne $1/x$.

9. Nonobstant le fait que ces constantes se mettent bien pour avoir un vrai cas d'indétermination 0/0, sinon la règle de l'Hospital ne s'applique pas.

Mais la limite $\epsilon \to 0^-$ n'existe pas :

$$\lim_{\epsilon\to 0^-} \frac{-\pi + \arctan(\epsilon/x) - \pi}{\epsilon} \tag{26.104}$$

n'existe pas.
Donc

$$\frac{\partial\varphi}{\partial y}(x,0) \tag{26.105}$$

n'existe pas pour $x < 0$.

(4) $\underline{x = 0 \textbf{ et } y \neq 0}$
Le calcul est immédiat

$$\lim_{\epsilon\to 0} \frac{\arg(iy + i\epsilon) - \arg(iy)}{\epsilon} = 0, \tag{26.106}$$

donc

$$\frac{\partial\varphi}{\partial y}(0,y) = 0. \tag{26.107}$$

En ce qui concerne la continuité, nous avons que $\partial_y\varphi$ est continue partout sauf sur la demi-droite $\{(x,0) \text{ tel que } x \leqslant 0\}$ où elle n'existe pas.

26.2.2 Une définition possible du logarithme

DEFooWDYNooYIXVMC

Définition 26.29.
Nous définissons la fonction ***logarithme*** *par*

$$\begin{aligned} \ln\colon \mathbb{C}^* &\to \mathbb{C} \\ z &\mapsto \ln\big(|z|\big) + i\arg(z) \end{aligned} \tag{26.108}$$

où le ln *à droite est le logarithme usuel sur* $\mathbb{R}^+$.

Remarque 26.30.
Cette fonction généralise le logarithme déjà vu sur $]0,\infty[\subset \mathbb{R}$. En effet pour des valeurs de z dans cette partie nous avons $\arg(z) = 0$ et $|z| = z$.

Lemme 26.31.
Le nombre $\ln(z)$ *est un logarithme de* z.

Démonstration. Nous avons

$$e^{\ln(z)} = e^{\ln|z|}e^{i\arg(z)} = |z|e^{i\arg(z)} = z. \tag{26.109}$$

Nous avons utilisé le fait que $e^{\ln(x)} = x$ pour $x \in \mathbb{R}^+$ et $|z|e^{i\arg(z)} = z$ par définition de la fonction arg. □

Notons que si on avait pris d'autres conventions pour définir arg, nous aurions eu d'autres définitions possibles de ln.

Exemple 26.32.
Nous avons

$$\ln(-1) = \ln(1) + i\arg(-1). \tag{26.110}$$

Mais $\ln(1) = 0$ et $\arg(-1) = \pi$ (et non $-\pi$), donc

$$\ln(-1) = i\pi. \tag{26.111}$$

C'est cette définition du logarithme qui est prise par Sage, et c'est cela qui lui permet de donner la primitive de $1/x$ comme $\ln(x)$ et non $\ln(|x|)$, parce que Sage connaît les logarithmes de nombres réels négatifs :

```
 SageMath version 7.3, Release Date: 2016-08-04
 Type "notebook()" for the browser-based notebook interface.
 Type "help()" for help.

sage: ln(-1)
I*pi
sage: f(x)=1/x
sage: f.integrate(x)
x |--> log(x)
```

tex/sage/sageSnip010.sage

△

Nous avons jusqu'ici défini une fonction sur $\mathbb{C}^*$ qui fait correspondre à chaque nombre complexe un de ses logarithmes. Il reste quelques questions à régler :

— Est-ce que cette fonction est continue ? Holomorphe ? (réponses : non et non)

— Si non, est-ce qu'il y avait moyen de trouver une définition plus efficace ? (réponse : non)

LEMooMUOIooCnoWwq

Lemme 26.33.
La fonction ln *n'est pas continue sur* $]-\infty,0]$.

Démonstration. Attention à bien comprendre l'énoncé. La fonction

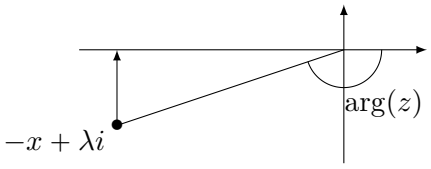
$$\begin{aligned} f\colon]-\infty,0[&\to \mathbb{C} \\ x &\mapsto \ln(x) \end{aligned} \tag{26.112}$$

est continue. D'ailleurs c'est $\ln(x) = \ln(|x|) + i\pi$. Ce dont il est question dans l'énoncé, c'est de la fonction ln vue comme fonction sur $\mathbb{C}^*$.

Soit $x > 0$ dans $\mathbb{R}$; nous avons

$$\ln(-x) = \ln(x) + i\pi. \tag{26.113}$$

Cependant $\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i)$ va valoir $\ln(|x|) - i\pi$. En effet lorsque $\lambda < 0$ est petit, l'argument de $-x+\lambda i$ se rapproche de $-\pi$ (et non de π).



Donc

$$\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i) = \lim \ln(|x+\lambda i|) + i\arg(-x+\lambda i) = \ln(|x|) - i\pi. \tag{26.114}$$

Nous n'avons donc pas continuité de la fonction logarithme comme fonction sur $\mathbb{C}^*$. □

THOooWUXOooYKvLbJ

Théorème 26.34.
La restriction

$$\ln\colon \mathbb{C}\backslash]-\infty,0] \to \mathbb{C} \tag{26.115}$$

est holomorphe.

Démonstration. Nous allons utiliser la proposition 12.311 et considérer la fonction

$$\begin{aligned} F\colon S &\to \mathbb{R}^2 \\ (x,y) &\mapsto \big(\ln(|x+iy|), \arg(x+iy)\big) \end{aligned} \tag{26.116}$$

où $S = \mathbb{R}^2\backslash\{(x,0) \text{ tel que } x \leqslant 0\}$. Nous devons vérifier que F est différentiable et que sa différentielle en un point de S est une similitude.

Nous posons

$$u(x,y) = \ln\big(\sqrt{x^2+y^2}\big) \tag{26.117}$$

et

$$v(x,y) = \arg(x+iy). \tag{26.118}$$

Les dérivées partielles de u ne sont pas très compliquées :

```
sage: var('x,y')
(x, y)
sage: u(x,y)=ln(sqrt(x**2+y**2))
sage: u.diff(x)
(x, y) |--> x/(x^2 + y^2)
```

tex/sage/sageSnip011.sage

c'est-à-dire

$$\frac{\partial u}{\partial x} = \frac{x}{x^2+y^2} \tag{26.119a}$$

$$\frac{\partial u}{\partial y} = \frac{y}{x^2+y^2}. \tag{26.119b}$$

Pour celles de v par contre, il faut se poser des questions, par exemples résister à la tentation d'écrire $v(x,y) = \arctan(y/x)$ et lire 26.27.

Nous avons déjà calculé les dérivées partielles de v dans 26.28, et nous avons vu qu'elles étaient continues sur $\mathbb{R}^2$ privé de la demi-droite.

Vu que les dérivées partielles sont continues, le théorème 12.294 nous dit que F est différentiable. La matrice de la différentielle est alors la matrice des dérivées partielles

$$\begin{pmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}, \tag{26.120}$$

qui a la forme requise (12.837) pour que la proposition 12.311 nous assure que ln soit $\mathbb{C}$-dérivable, c'est-à-dire holomorphe. □

26.2.3 Pas plus de continuité

Bon. La fonction logarithme que nous avons définie est holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite

$$U = \{z \in \mathbb{C} \text{ tel que } \operatorname{Im}(z) = 0, \operatorname{Re}(z) \leqslant 0\}. \tag{26.121}$$

EQooDOELooBQpPvG

Et elle n'est pas continue sur U ; elle y est cependant continue « par le haut ». Pouvons-nous faire mieux ? Nous allons maintenant prouver quelques résultats d'impossibilité de faire mieux que holomorphe partout sauf une partie pas si petite que ça.

Proposition 26.35.
Il n'existe pas de fonctions continues $f\colon \mathbb{C}^* \to \mathbb{C}$ telle que $e^{f(z)} = z$ pour tout $z \in \mathbb{C}^*$.

Démonstration. Pour tout z, le nombre $f(z)$ est un logarithme de z. Or $\ln(z)$ en est également un. Donc par le lemme 26.24

$$f(z) = \ln(z) + 2ik(z)\pi \tag{26.122}$$

pour une certaine fonction $k\colon \mathbb{C}^* \to \mathbb{Z}$. Sur le domaine d'holomorphie de ln, les fonctions ln et f étant continues, la fonction k l'est aussi. Mais une fonction continue à valeurs dans $\mathbb{Z}$ est constante (son domaine est connexe).

Il existe donc $k \in \mathbb{Z}$ tel que

$$f(z) = \ln(z) + 2ik\pi \tag{26.123}$$

au moins pour tout $z \in \mathbb{C}^*\backslash U$. Une telle fonction ne peut pas être continue sur U parce que ln ne l'est pas. □

Ok. Pas continue sur tout $\mathbb{C}$. Mais continue sur un peu plus que $\mathbb{C}$ privé de toute une demi-droite ? La proposition suivante répond que bof.

Proposition 26.36.
Soit Ω un ouvert de $\mathbb{C}$ contenant $S(0,r)$ (le cercle centré en 0 et de rayon $r > 0$). Il n'existe pas de fonction continue $f\colon \Omega \to \mathbb{C}$ telle que $e^{f(z)} = z$ pour tout $z \in \Omega$.

Démonstration. Encore une fois, pour tout $z \in \Omega$ nous avons

$$f(z) = \ln(z) + 2i\pi k(z) \tag{26.124}$$

pour une certaine fonction $k\colon \Omega \to \mathbb{Z}$. Nous considérons la demi-droite U de (26.121). Sur $\Omega\backslash U$, la fonction ln est continue et k doit également l'être. Donc k est constante sur les composantes connexes de $\Omega\backslash U$.

Vu que $S(0,r)$ est compact, on peut le recouvrir par un nombre fini de boules centrées en des points de $S(0,r)$. En prenant le minimum des rayons de ces boules, nous voyons que Ω contient une couronne

$$\{z \in \mathbb{C} \text{ tel que } r - \delta \leqslant |z| \leqslant r + \delta\}. \tag{26.125}$$

Soit le point $x_0 = -r$. C'est un point de Ω contenu dans U. Nous allons prouver que $B(x_0,\delta)\backslash U$ est dans une seule composante connexe de Ω.

Soit un point $z_1 \in B(x_0,\delta)$ situé au-dessus de U, et z_2 un point de $B(x_0,\delta)$ situé en dessous de U. Le cercle $S(0,r)$ coupe $B(x_0,\delta)$ en deux points : un au-dessus et un en-dessous de U. On peut lier z_1 au point de « sortie » supérieur de $S(0,r)$ en restant dans $B(x_0,\delta)$; ce point est ensuite relié en suivant le cercle au point d'entrée inférieur du cercle dans $B(x_0,\delta)$. Ce dernier point est lié à z_2 par un chemin restant dans la boule.

Tout cela pour dire que z_1 et z_2 sont dans la même composante connexe de Ω et que $k(z_1) = k(z_2)$. Il existe donc $k \in \mathbb{Z}$ tel que

$$f(z) = \ln(z) + 2ik\pi \tag{26.126}$$

sur $B(x_0,\delta)\backslash U$. Une telle fonction f ne peut pas être continue. □

26.2.4 Pas d'unicité : autres déterminations de l'argument

NORMooFCDOooFDzAjp

26.37.
Nous avons pris la fonction d'argument $\arg\colon \mathbb{C} \to]-\pi,\pi]$. Il y en a évidemment beaucoup d'autres de possibles. Par exemple pour $\alpha \in \mathbb{R}$ nous pouvons considérer

EQooNKKDooOuJxXe

$$\arg_{\alpha^+}\colon \mathbb{C} \to]\alpha,\alpha+2\pi] \tag{26.127}$$

ou

$$\arg_{\alpha^-}\colon \mathbb{C} \to [\alpha,\alpha+2\pi[. \tag{26.128}$$

En posant

$$\ln_{\alpha^\pm}(z) = \ln(|z|) + i\arg_{\alpha^\pm}(z) \tag{26.129}$$

nous avons une fonction réciproque de l'exponentielle définie sur $\mathbb{C}^*$ et holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite D_α (dépendante de la valeur de α).

La différence entre $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ est seulement la valeur sur la demi-droite de non-holomorphie. L'une sera semicontinue d'un côté et l'autre, de l'autre côté.

Remarque 26.38.
La fonction $\arg_{0^-}$ a déjà été utilisée en 18.5.2 pour écrire un inverse de la fonction

$$\begin{aligned}\varphi\colon [0,2\pi[&\to S^1\\ t &\mapsto e^{it}.\end{aligned} \tag{26.130}$$

Définition 26.39 ([636]).
Soit un ouvert $\Omega\subset\mathbb{C}^*$. Nous disons que la fonction $f\colon\Omega\to\mathbb{C}$ est une ***détermination*** *sur Ω si elle est continue et vérifie*

$$e^{f(z)}=z \tag{26.131}$$

pour tout $z\in\Omega$.

Les différents résultats vus jusqu'ici montrent qu'il n'existe pas de détermination du logarithme sur $\mathbb{C}^*$.

Définition 26.40.
La ***détermination principale*** *du logarithme est la restriction de notre logarithme 26.29*

$$\begin{aligned}\ln\colon \mathbb{C}^* &\to \mathbb{C}\\ z &\mapsto \ln(|z|)+i\arg(z)\end{aligned} \tag{26.132}$$

à l'ouvert $\mathbb{C}^*\backslash U$ où U est la demi-droite $\mathrm{Re}(z)\leqslant 0$, $\mathrm{Im}(z)=0$ de $\mathbb{C}$.

REMooFBLLooDnkmjR

Remarque 26.41.
Beaucoup de sources[637] ne définissent pas $\ln_{\alpha^\pm}$ sur la droite D_α. C'est-à-dire qu'ils notent $\ln_\alpha$ notre fonction $\ln_{\alpha^+}$ restreinte à $\mathbb{C}^*\backslash D_\alpha$. Dans ce cas, les fonctions $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ sont identiques[10].

Cette remarque est importante parce que certains vont vous dire « le logarithme n'est pas défini sur la demi-droite » ; de leur point de vue, la fonction que nous avons définie est une prolongation (non continue) à U du logarithme, qui est continu.

(1) Certaines personnes pourraient vous dire que notre logarithme « n'est pas bien défini parce que si on fait le tour dans un sens ou dans l'autre nous n'obtenons pas la même valeur pour $\ln(z)$ lorsque z est sur U ». Et cela avec des arguments aussi forts que « 2π et 0, c'est le même point ».
Nous préférons être bien clairs[11] sur ce point : notre fonction ln est parfaitement définie sur $\mathbb{C}^*$ et 2π n'est pas la même chose que zéro. En particulier $\arg(e^{2i\pi})=0$ et $\arg(e^{-i\pi})=\pi$ et non $-\pi$.

(2) Il n'en reste pas moins que Sage donne $\ln(-1)=i\pi$ et que nous avons choisi de faire de même, parce que le Frido n'est pas un cours d'agrégation, mais un texte qui donne quelques éléments de mathématique dans le but d'utiliser Sage efficacement.

(3) Tout ceci pour dire que si vous utilisez ce livre pour l'agrégation, vous devriez sérieusement considérer l'option de ne pas donner du logarithme la définition donnée ici, mais bien sa restriction.

En fait notre logarithme est maximum pour la propriété « être une réciproque de l'exponentielle » alors que beaucoup de monde préfère avoir une fonction maximale pour la propriété « être réciproque de l'exponentielle tout en étant continue ».

De toutes les fonctions ayant le droit de vouloir être appelée « logarithme », celle que nous avons choisie (un peu arbitrairement) pour s'appeler « logarithme » et accaparer de la notation « ln » est $\ln_{\pi^+}$. Elle est d'une certaine manière celle qui arrive le plus naturellement.

10. Cela n'est pas tout à fait évident ; vous devriez y penser.
11. Est-ce qu'il faut vraiment un pluriel ici ?

En effet si nous pensons au logarithme népérien $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ que nous voulons prolonger sur $\mathbb{R}$, nous devons poser

$$\ln(-x) = \ln(-1) + \ln(x) \tag{26.133}$$

pour $x > 0$. Que peut valoir $\ln(-1)$? Il doit vérifier $e^{\ln(-1)} = -1$. La première valeur qui nous tombe sous la main est $\ln(-1) = \pi$. Bien entendu, d'autres possibilités existent, comme $\ln(-1) = 2017\pi$ par exemple.

26.2.5 Pas d'unicité : développement en série

Pour $z_0 \in \mathbb{C}^*$ nous pouvons écrire un développement en série de la réciproque de l'exponentielle autour de z_0. La fonction ainsi définie est holomorphe sur la boule $B(z_0, |z_0|)$ et diverge en dehors de cette boule.

Voilà encore une fonction « logarithme » pour chaque point de $\mathbb{C}^*$. Nous nommons $\ln_{z_0}$ la fonction

$$\ln_{z_0}\colon B(z_0, |z_0|) \to \mathbb{C} \tag{26.134}$$

donnée par la série.

En général nous n'avons pas $\ln_{z_1} = \ln_{z_2}$ sur l'intersection des disques de convergence. Si c'était le cas, de proche en proche nous pourrions construire une fonction continue réciproque du logarithme sur $\mathbb{C}^*$, ce qui est impossible.

26.2.6 Pas d'unicité : laquelle choisir ?

Bon. Pour chaque demi-droite D nous avons une détermination du logarithme sur $\mathbb{C}^*\backslash D$. Et pour tout $z_0 \in \mathbb{C}^*$ nous en avons une sur $B(z_0, |z_0|)$.

En pratique, quel logarithme choisir ? Cela dépend du problème.

Si vous avez besoin ou envie de travailler avec des série entières, le mieux est de choisir une détermination donnée par un développement autour d'un point bien choisi par rapport à votre problème.

Si vous avez surtout besoin d'holomorphie, et que vous en avez besoin sur un grand domaine, vous devriez choisir une détermination sur un des ensembles $\mathbb{C}^*\backslash D_\alpha$ en choisissant α de telle sorte que la demi-droite maudite ne passe pas par la zone sur laquelle vous travaillez.

Dans tous les cas, vous devez préciser très explicitement la détermination choisie. Dans ce texte, sauf mention du contraire, nous utiliserons la détermination principale, et même son extension (non continue) à $\mathbb{C}^*$. Lorsque nous aurions besoin d'holomorphie, nous préciserons que nous considérons la restriction.

26.2.7 Logarithme comme primitive

Tout le monde sait [12] que le logarithme $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ est une primitive de la fonction $x \mapsto 1/x$. Qu'en est-il dans le cas complexe ? Tout d'abord précisons que nous ne comptons pas encore parler d'intégrale sur $\mathbb{C}$, mais seulement d'intégrales sur $\mathbb{R}$ d'une fonction à valeur complexes.

PROPooNIJVooKueuYJ

Proposition 26.42.
Si $z \in \mathbb{C}$ alors

$$\int \frac{1}{x+z}dx = \ln(x+z) \tag{26.135}$$

EQooAHYXooTPGXDS

Démonstration. Il est important de comprendre que la formule (26.135) est un abus de notation pour dire que si nous considérons la fonction

$$\begin{aligned}\varphi\colon \mathbb{R} &\to \mathbb{C}\\ x &\mapsto \ln(x+z)\end{aligned} \tag{26.136}$$

12. Proposition 15.85.

alors nous avons $\varphi'(x) = \frac{1}{x+z}$. Ici la dérivation est une dérivation sur $\mathbb{R}$ et l'intégrale est une intégrale sur $\mathbb{R}$, c'est-à-dire « composante par composantes ». La fonction φ se décompose en partie réelle et imaginaire qui sont à dériver séparément :

$$\varphi(x) = \ln(|x+z|) + i\arg(x+z). \tag{26.137}$$

(1) **Si z est imaginaire pur**
Nous posons $z = \lambda i$ avec $\lambda \in \mathbb{R}^*$. D'abord nous avons

$$\frac{1}{x+\lambda i} = \frac{x}{x^2+\lambda^2} - i\frac{\lambda}{x^2+\lambda^2}. \tag{26.138}$$

La partie réelle de $\varphi(x)$ est

$$\varphi_1(x) = \ln\left(\sqrt{x^2+\lambda^2}\right), \tag{26.139}$$

dont la dérivée est

$$\varphi_1'(x) = \frac{x}{x^2+\lambda^2}, \tag{26.140}$$

qui correspond bien à la partie réelle de $\frac{1}{x+\lambda i}$.
En ce qui concerne la partie imaginaire, $\varphi_2(x) = \arg(x+\lambda i)$, et sa dérivée n'est rien d'autre que la dérivée partielle par rapport à x de la fonction argument, déjà calculée en (26.97) :

$$\varphi_2'(x) = \frac{-\lambda}{x^2+\lambda}. \tag{26.141}$$

Cela est bien la partie imaginaire de $\frac{1}{x+\lambda i}$.
Notons que nous n'avons pas de problèmes sur la demi-droite des réels négatifs parce que nous ne considérons au final que la dérivée partielle par rapport à x de la fonction argument, laquelle existe et est continue, même sur cette partie.

(2) **Pour z quelconque**
Soit $z = s + \lambda i$ avec $s, \lambda \in \mathbb{R}$. En posant $\varphi_0(x) = \ln(x+\lambda i)$ nous avons $\varphi(x) = \varphi_0(x+s)$ et donc

$$\varphi'(x) = \varphi_0'(x+s) = \frac{1}{x+s+\lambda i} = \frac{1}{x+z}. \tag{26.142}$$

Tout va bien.

□
EXooAKEDooZgjocX

Exemple 26.43.
Un petit calcul d'intégrale, que nous avions déjà faite dans l'exemple 20.100 (avec la méthode de Rothstein-Trager). En passant par une décomposition en fractions simples :

$$\int \frac{1}{x^3+x} = \int \left(\frac{1}{x} - \frac{1/2}{x-i} - \frac{1/2}{x+i}\right) \tag{26.143a}$$

$$= \ln(x) - \frac{1}{2}\ln(x-i) - \frac{1}{2}\ln(x+i) \tag{26.143b}$$

$$= \ln(x) - \frac{1}{2}\ln(x^2+1). \tag{26.143c}$$

SUBEQooRNQLooScfSlG

Attention aux justifications. Il n'est pas vrai en général dans le cas de nombres complexes a et b que $\ln(ab) = \ln(a) + \ln(b)$. En effet, pour la partie réelle, ça passe parce que $|ab| = |a||b|$. Mais en ce qui concerne la partie imaginaire,

$$\arg(ab) \neq \arg(a) + \arg(b) \tag{26.144}$$

lorsque la somme dépasse les bornes de $]-\pi, \pi]$. Le passage à (26.143c) fonctionne parce que dans le cas particulier des nombres $x+i$ et $x-i$, les arguments se somment à zéro : $\arg(x+i)+\arg(x-i) = 0$. △

26.2.8 Logarithme sur un chemin

DEFooBBGFooCEdsFR

Définition 26.44 (logarithme continu)**.**
*Soient un espace topologique X, et une application $f\colon X \to \mathbb{C}^*$. Nous disons que $g\colon X \to \mathbb{C}^*$ est un **logarithme** de f si pour tout $x \in X$ nous avons*

$$f(x) = \exp\big(g(x)\big). \tag{26.145}$$

DEFooOCDGooGyvvWi

Définition 26.45 (Détermination du logarithme)**.**
Soit un chemin[13] *$\gamma\colon [a,b] \to \mathbb{C}^*$. Nous disons qu'une application $g\colon \gamma\big([a,b]\big) \to \mathbb{C}^*$ est une **détermination du logarithme** sur γ si*

$$\exp\big((g\circ\gamma)(t)\big) = \gamma(t) \tag{26.146}$$

pour tout $t \in [a,b]$.

THOooUPANooMiECqe

Théorème 26.46.
Tout chemin dans $\mathbb{C}^*$ admet une détermination continue du logarithme[14]*, et si l est une détermination sur le chemin γ, nous avons*

$$\int_\gamma \frac{dz}{z} = l\big(\gamma(b)\big) - l\big(\gamma(a)\big). \tag{26.147}$$

Le théorème de Borsuk parle de lien entre logarithme continu et homotopie de fonctions à la fonction constante. Il s'applique donc bien de concert avec la proposition 26.15 qui dit que si f et g sont homotopes, alors f/g est homotope à une constante.

THOooTCUMooEByCKg

Théorème 26.47 (Théorème de Borsuk[205])**.**
Soient un compact K de $\mathbb{R}^n$ ainsi qu'une application continue $f\colon K \to \mathbb{C}^*$. Les propriétés suivantes sont équivalentes :

ITEMooKZYDooKoEEbl

(1) f admet un logarithme continu[15]*.*

ITEMooXVNXooVAHklr

(2) f est homotope à l'application constante $u\colon K \to \mathbb{C}^*$, $u(z) = 1$.

ITEMooQDHXooObjxLA

(3) f admet une extension continue $\tilde{f}\colon \mathbb{R}^n \to \mathbb{C}^*$.

Démonstration. En plusieurs points.

(1) **(1) ⇒ (3)** Soit un logarithme continu $g\colon K \to \mathbb{C}$ de f. Vu que K est fermé, le théorème de Tietze 26.21 dit que g possède une extension continue $\tilde{g}\colon \mathbb{R}^n \to \mathbb{C}$. Nous posons

$$\begin{aligned}\tilde{f}\colon \mathbb{R}^n &\to \mathbb{C}^*\\ x &\mapsto \exp\big(\tilde{g}(x)\big).\end{aligned} \tag{26.148}$$

L'application $\tilde{f}$ est continue parce que exp et $\tilde{g}$ le sont. Elle est une extension de f parce que si $x \in K$, nous avons

$$\tilde{f}(x) = \exp\big(\tilde{g}(x)\big) = \exp\big(g(x)\big) = f(x) \tag{26.149}$$

parce que $\tilde{g}(x) = g(x)$ et g est un inverse de exp.

(2) **(3) ⇒ (2)** Soit une extension continue $\tilde{f}\colon \mathbb{R}^n \to \mathbb{C}^*$ de f. Soit $x_0 \in \mathbb{R}^n$. Nous posons

$$\begin{aligned}H\colon [0,1]\times K &\to \mathbb{C}^*\\ (t,x) &\mapsto \tilde{f}\big((1-t)x + tx_0\big).\end{aligned} \tag{26.150}$$

L'application H est continue, et pour $x \in K$ elle vérifie $H(0,x) = \tilde{f}(x) = f(x)$ ainsi que $H(1,x) = \tilde{f}(x_0)$.

13. Définition 20.75.
14. Définition 26.45.
15. Définition 26.44.

Donc f est homotope à l'application constante $\tilde{f}(x_0)$. Vu que $\mathbb{C}^*$ est connexe par arcs, toutes les applications constantes sont homotopes [16]. Donc f est homotope à $\tilde{f}(x_0)$ qui est homotope à u. L'homotopie étant une relation d'équivalence [17], f est homotope à u.

(3) **(2) ⇒ (1)** Nous considérons l'homotopie $H\colon [0,1]\times K\to\mathbb{C}^*$ entre f et u. En particulier pour tout $x\in K$ nous avons $H(0,x)=f(x)$ et $H(1,x)=1$.
La partie $[0,1]\times K$ est compacte [18] et $|H|$ y est continue. Donc elle atteint ses bornes. Mais elle prend ses valeurs dans $\mathbb{C}^*$; donc

$$\inf_{x,t}|H(x,t)|>0. \tag{26.151}$$

Et d'ailleurs cet infimum est un minimum. Nous considérons la norme suivante sur $\mathbb{R}\times\mathbb{R}^n$:

$$\|(t,x)\|_\infty=\max\big(|t|,|x|\big). \tag{26.152}$$

Toutes les normes étant équivalentes [19] sur $\mathbb{R}\times\mathbb{R}^n$, nous pouvons caractériser la compacité, la continuité et tout ça en termes de cette norme. L'application H est uniformément continue [20]. Soit $\epsilon>0$, il existe $\eta>0$ tel que si $\|(t,x)-(s,y)\|_\infty<\eta$ alors $\|H(t,x)-H(s,y)\|<\epsilon$.
Soient un entier $n>\frac{1}{\eta}+1$, $x\in K$ et $k\in\{0,\dots,n\}$. Nous avons

$$\|(\frac{k}{n},x)-(\frac{k+1}{n},x)\|_\infty=\|(-\frac{1}{n},0)\|=\frac{1}{n}<\eta. \tag{26.153}$$

Nous avons donc également

$$\big|H(\frac{k}{n},x)-H(\frac{k+1}{n},x)\big|<\epsilon. \tag{26.154}$$

Nous posons

$$\begin{aligned}F_k\colon K&\to\mathbb{C}^*\\ x&\mapsto H\big(\frac{k}{n},x\big).\end{aligned} \tag{26.155}$$

Cette application est continue sur K et pour tout $x\in K$ nous avons [SUBEQSooLJIJooYkxyMO]

$$\big|\frac{F_k(x)}{F_{k+1}(x)}-1\big|=\big|\frac{F_k(x)-F_{k+1}(x)}{F_{k+1}(x)}\big| \tag{26.156a}$$

$$=\big|\frac{H(\frac{k}{n},x)-H(\frac{k+1}{n},x)}{H(\frac{k+1}{n},x)}\big| \tag{26.156b}$$

$$<\frac{\epsilon}{|H(\frac{k+1}{n},x)|}. \tag{26.156c}$$

Pour rappel, pour tout $\epsilon>0$, il existe un $\eta>0$ tel qu'en posant $n>\frac{1}{\eta}+1$ nous avons (26.156). Nous choisissons ϵ tel que $0<\epsilon<\inf_{(t,x)}|H(t,x)|$, de telle sorte à avoir

$$\big|\frac{F_k(x)}{F_{k+1}(x)}-1\big|<\frac{\epsilon}{|H(\frac{k+1}{n},x)|}<1. \tag{26.157}$$

La fonction $\ln\colon\mathbb{C}^*\to\mathbb{C}^*$ est continue sur $B(1,1)$ [21]. Donc la fonction

$$\begin{aligned}h_k\colon K&\to\mathbb{C}^*\\ x&\mapsto\ln\left(\frac{F_k(x)}{F_{k+1}(x)}\right)\end{aligned} \tag{26.158}$$

16. Lemme 26.14.
17. Lemme 26.13.
18. Théorème 7.314.
19. Équivalence de normes, théorème 10.49.
20. Théorème de Heine 12.78.
21. Le logarithme est celui défini en 26.29 et sa continuité est le théorème 26.34.

est continue. Nous posons

$$g = \sum_{k=0}^{n-1} \ln\left(\frac{F_k}{F_{k+1}}\right), \tag{26.159}$$

qui est également continue sur K. Cette fonction g est un logarithme continu de f sur K parce que[22]

$$\begin{aligned}
\exp\big(g(x)\big) &= \exp\left(\sum_{k=0}^{n-1} \ln\left(\frac{f_k(x)}{F_{k+1}(x)}\right)\right) && (26.160a)\\
&= \prod_{k=0}^{n-1} \exp\left(\ln\left(\frac{F_k(x)}{F_{k+1}(x)}\right)\right) && (26.160b)\\
&= \prod_{k=0}^{n-1} \frac{F_k(x)}{F_{k+1}(x)} && (26.160c)\\
&= \frac{F_0(x)}{F_n(x)} && (26.160d)\\
&= \frac{H(0,x)}{H(1,x)} && (26.160e)\\
&= \frac{f(x)}{1} && (26.160f)\\
&= f(x). && (26.160g)
\end{aligned}$$

□

CORooXOOZooUJMKxu

Corolaire 26.48 ([205]).
Soit un compact convexe K de $\mathbb{R}^n$. Toute application continue $K \to \mathbb{C}^*$ admet un logarithme continu.

Démonstration. Soit une application continue $f\colon K \to \mathbb{C}^*$. Soit $x_0 \in K$. Vu que K est convexe, f est homotope à la fonction constante $f(x_0)$. Par arc-connexité de $\mathbb{C}^*$, l'application f est alors homotope à la fonction constante 1.

Le théorème de Borsuk 26.47 nous dit alors que f admet un logarithme continu. □

LEMooJNPTooScfSvA

Lemme 26.49.
Si K est compact dans $\mathbb{C}$, alors la partie $\mathbb{C}\backslash K$ possède exactement une composante connexe non bornée.

Démonstration. La partie K étant compacte, elle est bornée. Il existe donc $r > 0$ tel que $K \subset B(0,r)$. Soit $A = \mathbb{C}\backslash B(0,r)$. Cela est une partie connexe de $\mathbb{C}$ et est donc contenue dans une composante connexe de $\mathbb{C}\backslash K$. Il y a donc existence d'une composante connexe non bornée de $\mathbb{C}\backslash K$.

Pour l'unicité, les autres composantes connexes de $\mathbb{C}\backslash K$ sont contenues dans $B(0,r)$ et sont donc bornées. □

26.2.9 Lacets, indice et homotopie

DEFooLFBNooGlvJmp

Proposition-Définition 26.50 ([638]).
Soit γ un chemin fermé[23] *dans $\mathbb{C}$ que nous supposons continu et C^1 par morceaux*[24]*. L'**indice** de*

22. Dans le calcul suivant, nous utilisons entre autres la formule $\exp(a+b) = \exp(a)\exp(b)$ de la proposition 18.9(2).
23. Par abus de langage, nous désignerons par γ à la fois le chemin et son image.
24. Définition 20.37

la courbe γ est la fonction

$$\begin{aligned}\mathrm{Ind}_\gamma\colon \mathbb{C}\backslash\gamma &\to \mathbb{Z}\\ z &\mapsto \frac{1}{2\pi i}\int_\gamma \frac{d\omega}{\omega - z}.\end{aligned} \tag{26.161}$$

Alors :

ITEMooHYHMooWcaJxP

(1) La fonction Ind_γ prend ses valeurs dans $\mathbb{Z}$.

(2) Elle est continue sur chaque composante connexe de $\mathbb{C}\backslash\gamma$.

(3) La fonction indice est constante sur chaque composante connexe[25] *de $\mathbb{C}\backslash\gamma$.*

(4) Elle est nulle sur les composantes non bornées de $\mathbb{C}\backslash\gamma$[26].

Démonstration. Point par point.

(1) **Prend ses valeurs dans $\mathbb{Z}$**

Explicitons l'intégrale sur $\gamma\colon [a,b]\to\mathbb{C}$:

$$\mathrm{Ind}_\gamma(z_0) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z - z_0} = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t) - z_0}dt. \tag{26.162}$$

Note. Comme γ n'est peut-être pas dérivable en certains points, en réalité nous n'avons pas une intégrale entre a et b, mais une somme d'intégrales entre a_i et b_i avec $a_0 = a$ et $b_n = b$. Nous nommons S l'ensemble de ces points.

Nous allons calculer l'exponentielle de ce qui est dans l'intégrale en posant

$$\begin{aligned}\varphi\colon [a,b] &\to \mathbb{C}\\ u &\mapsto \exp\left(\int_a^u \frac{\gamma'(t)}{\gamma(t) - z_0}dt\right).\end{aligned} \tag{26.163}$$

Nous savons par le lien entre primitive et intégrale (proposition 14.286) que la dérivée de $x\mapsto\int_0^x f$ est f elle-même. Nous pouvons donc facilement dériver φ :

$$\varphi'(u) = \frac{\gamma'(u)}{\gamma(u) - z_0}\exp\left(\int_a^u \frac{\gamma'(t)}{\gamma(t) - z_0}dt\right) = \frac{\gamma'(u)\varphi(u)}{\gamma(u) - z_0}. \tag{26.164}$$

EQooHEAWooCWnNLZ

Cette égalité est vraie pour tout $u\in[a,b]$ sauf en un nombre fini de points (les points de S). Nous considérons maintenant

$$\begin{aligned}f\colon [a,b] &\to \mathbb{C}\\ t &\mapsto \frac{\varphi(t)}{\gamma(t) - z_0},\end{aligned} \tag{26.165}$$

et nous dérivons :

$$f'(t) = \frac{\varphi'(t)\big(\gamma(t) - z_0\big) - \varphi(t)\gamma'(t)}{\big(\gamma(t) - z_0\big)^2} \tag{26.166a}$$

$$= \frac{\varphi'(t)}{\gamma(t) - z_0} - \frac{\varphi(t)\gamma'(t)}{\big(\gamma(t) - z_0\big)^2}. \tag{26.166b}$$

Nous y injectons l'expression (26.164) de φ' :

$$f'(t) = \frac{\gamma'(t)\varphi(t)}{\big(\gamma(t) - z_0\big)^2} - \frac{\varphi(t)\gamma'(t)}{\big(\gamma(t) - z_0\big)^2} = 0. \tag{26.167}$$

25. Définition 7.73.

26. Elle a une seule telle composante par le lemme 26.49, mais ce n'est pas important ici.

Comme f est continue[27] sur $[a,b]$ et de dérivée nulle partout sauf sur S, elle est constante sur $[a,b]$. Nous avons donc $f(t)=f(a)$ pour tout t, c'est à dire

$$\frac{\varphi(t)}{\gamma(t)-z_0}=\frac{1}{\gamma(a)-z_0}. \tag{26.168}$$

Nous pouvons isoler de là une valeur pour $\varphi(t)$:

$$\varphi(t)=\frac{\gamma(t)-z_0}{\gamma(a)-z_0}. \tag{26.169}$$

C'est le moment de calculer $\varphi(b)$. Dans le calcul suivant, nous utilisons $\gamma(a)=\gamma(b)$:

$$\varphi(b)=\frac{\gamma(b)-z_0}{\gamma(a)-z_0}=\frac{\gamma(b)-z_0}{\gamma(b)-z_0}=1. \tag{26.170}$$

Nous avons prouvé que

$$\exp\left(\int_a^b\frac{\gamma'(t)}{\gamma(t)-z_0}dt\right)=1. \tag{26.171}$$

Pour qu'une exponentielle soit égale à 1, il faut que ce qui se trouve dedans soit égal à $2ik\pi$ pour un certain k. Il existe $k\in\mathbb{Z}$ tel que

$$\int_a^b\frac{\gamma'(t)}{\gamma(t)-z_0}dt=2i\pi k. \tag{26.172}$$

Pour cette valeur de k nous avons $\mathrm{Ind}_\gamma(z_0)=k$.

(2) **Une borne pour $|\gamma'|$** L'application γ est C^1 par morceaux[28]. La restriction $\gamma'\colon[a_i,b_i]\to\mathbb{C}$ est continue sur un compact et est donc bornée. Pour chaque i nous avons donc une borne M_i pour $|\gamma' j|$. En nommant M le maximum des ces bornes M_i nous avons $\gamma'(t)<M$ pour tout $t\in[a,b]$.

(3) **L'indice est ponctuellement continu** Nous utilisons le théorème 17.15 avec l'application

$$\begin{aligned}f\colon[a,b]\times\mathbb{C}\backslash\gamma&\to\mathbb{C}\\(t,z)&\mapsto\frac{\gamma'(t)}{\gamma(t)-z}.\end{aligned} \tag{26.173}$$

Sauf que non. Il ne va pas être possible de majorer cette fonction uniformément en z, comme il le faudrait pour remplir la condition (3). La raison est que quand z s'approche de γ, le dénominateur n'est pas contrôlé.

Nous allons donc restreindre le domaine de f. Soit $z_0\in\mathbb{C}\backslash\gamma$. Vu que γ est continue et que $[a,b]$ est compact, la fonction $t\mapsto 1/(\gamma(t)-z_0)$ atteint une borne inférieure. Disons que cette borne est atteinte pour $t=t_0$. Nous considérons $r<d\big(z_0,\gamma(t_0)\big)$ et nous prenons une borne M pour $|\gamma'|$. Nous avons alors, pour tout t que

$$\frac{|\gamma'(t)|}{|\gamma(t)-z_0|}<\frac{M}{r}. \tag{26.174}$$

Le z_0 étant fixé, la fonction constante M/r majore f uniformément en t. La théorème de continuité sous l'intégrale 17.15 dit alors que l'indice est continue en z_0.

(4) **Continuité globale** La fonction indice étant continue en chaque point de $\mathbb{C}\backslash\gamma$, elle est continue sur $\mathbb{C}\backslash\gamma$ (théorème 7.205).

(5) **Sur les composantes connexes** Soit une composante connexe C de $\mathbb{C}\backslash\gamma$. L'application Ind_γ est continue et à valeurs entières sur C. La proposition 7.216 dit alors que l'indice est constant sur C.

27. Fait important : nous utilisons la continuité de f, pas de l'indice. Nous n'avons pas encore démontré que l'indice est une fonction continue.

28. Nous utilisons ici crucialement le fait que les morceaux sont fermés.

(6) **Nulle sur la non bornée** En majorant $|\gamma'|$ par M, nous avons

$$|\operatorname{Ind}_\gamma(z_0)| \leqslant \frac{1}{2\pi}\int_a^b \frac{|\gamma'(t)|}{|\gamma(t)-z_0|} \leqslant \frac{M}{2\pi}\int_a^b \frac{dt}{|\gamma(t)-z_0|}. \tag{26.175}$$

Soit $\epsilon > 0$ assez petit pour que $\epsilon M(b-a)/2\pi < 1$. Vu que γ est bornée, il existe $R > 0$ tel que $\gamma(t) \in B(0,R)$ pour tout t. Nous considérons r assez grand pour que $r - R > 1/\epsilon$.

Comme nous étudions $\operatorname{Ind}_\gamma$ sur une composante non bornée, nous considérons $|z_0| > r$. Avec ça nous avons $|\gamma(t) - z_0| > r - R$ et donc

$$|\operatorname{Ind}_\gamma(z_0)| \leqslant \frac{M}{2\pi}\int_a^b \epsilon dt = \frac{M(b-a)}{2\pi}\epsilon < 1. \tag{26.176}$$

La dernière inégalité est un choix de ϵ fait au départ.

Étant donné que $\operatorname{Ind}_\gamma(z_0) \in \mathbb{Z}$, nous avons forcément $\operatorname{Ind}_\gamma(z_0) = 0$. Et comme l'indice est constant sur les composantes connexe, il est nul partout.

□

PROPooXWULooFZgHfL

Proposition 26.51.
Soit $p \in \mathbb{C}$. Soit un chemin $\gamma\colon [0,1] \to \mathbb{C}\backslash\{p\}$. Quelques hypothèses :

(1) Il existe une demi-droite D d'origine p et tel que $\gamma^{-1}(D)$ est fini.

(2) Le chemin γ traverse D en tout point de $\gamma^{-1}(D)$.

(3) Soit n^+ le nombre de points de $\gamma^{-1}(D)$ où γ traverse positivement D et n^- idem négativement.

Alors

$$\operatorname{Ind}(\gamma, p) = n^+ - n^-. \tag{26.177}$$

PROPooEKFHooOWcIMk

Proposition 26.52 ([205])**.**
Soient $p \in \mathbb{C}$ ainsi que deux chemins $\gamma_1, \gamma_2\colon [a,b] \to \mathbb{C}\backslash\{p\}$. Si γ_1 et γ_2 sont homotopes dans $\mathbb{C}\backslash\{p\}$. Alors

$$\operatorname{Ind}(\gamma_1, p) = \operatorname{Ind}(\gamma_2, p). \tag{26.178}$$

Démonstration. Posons

$$\begin{aligned} f\colon \mathbb{C}\backslash\{p\} &\to \mathbb{C} \\ z &\mapsto \frac{1}{z-p}. \end{aligned} \tag{26.179}$$

La partie $\mathbb{C}\backslash\{p\}$ est ouverte. Le théorème 26.17 d'invariance de l'intégrale par homotopie nous indique que

$$\operatorname{Ind}(\gamma_1, p) = \frac{1}{2\pi i}\int_{\gamma_1} f = \frac{1}{2\pi i}\int_{\gamma_2} f = \operatorname{Ind}(\gamma_2, p). \tag{26.180}$$

□

DefECnFJQp

Définition 26.53.
*Si γ_1 et γ_2 sont deux lacets en $x_0 \in X$ (un espace topologique), une **équivalence d'homotopie** est une application $f\colon [0,1]\times[0,1] \to X$ telle que*

(1) $f(0,t) = \gamma_1(t)$ pour tout t ;

(2) $f(1,t) = \gamma_2(t)$ pour tout t ;

(3) pour chaque $t \in [0,1]$, l'application $s \mapsto f(s,t)$ est continue ;

(4) pour chaque $s \in [0,1]$, l'application $t \mapsto f(s,t)$ est un lacet basé en x_0.

LEMooMEICooOOGzdf

Lemme 26.54.
Si γ est un cercle de centre $z_0 \in \mathbb{C}$ et de rayon r, alors

$$\operatorname{Ind}_\gamma(z) = \begin{cases} 1i & \text{si } z \in B(z_0, r) \\ 0 & \text{sinon.} \end{cases} \tag{26.181}$$

Démonstration. La seconde ligne provient directement de la proposition 26.50. Pour la première, le cercle γ se paramètre par

$$\gamma(\theta) = z_0 + re^{i\theta}. \tag{26.182}$$

Nous commençons par calculer $\text{Ind}_\gamma(z_0)$. Il vaut

$$\frac{1}{2\pi i}\int_\gamma \frac{d\omega}{\omega - z_0} = \frac{1}{2\pi i}\int_0^{2\pi} \frac{1}{re^{i\theta}} ire^{i\theta} d\theta = 1. \tag{26.183}$$

Nous considérons maintenant un point $p \in B(z_0, r)$ et nous calculons $\text{Ind}_\gamma(p)$. D'abord si σ est le cercle $\sigma(t) = \gamma(t) + p - z_0 = re^{it} + p$ de rayon r centré en p, nous avons

$$\text{Ind}_\sigma(p) = 1 \tag{26.184}$$

par ce que nous venons de voir.

Ensuite les chemins σ et γ sont homotopes dans $\mathbb{C}\backslash\{p\}$. En effet nous posons

$$H(u,t) = (1-u)\gamma(t) + u\sigma(t). \tag{26.185}$$

Nous avons

$$H(u,t) - p = (1-u)z_0 + re^{it} + (u-1)p = (1-u)(z_0 - p) + re^{it}. \tag{26.186}$$

En termes de modules, nous avons d'une part

$$|(1-u)(z_0-p)| \leqslant |z_0 - p| < r \tag{26.187}$$

et $|re^{it}| = r$. Donc la somme des deux ne peut pas faire zéro. Nous avons donc bien $H(u,t) \neq p$ pour tout $u \in [0,1]$ et pour tout $t \in [0, 2\pi]$.

Les chemins γ et σ étant homotopes dans $\mathbb{C}\backslash\{p\}$, la proposition 26.52 dit que

$$\text{Ind}_\gamma(p) = \text{Ind}_\sigma(p) = 1. \tag{26.188}$$

□

PROPooGAOIooFTOuli

Proposition 26.55 ([205])**.**
Soit un intervalle $J \subset \mathbb{R}$. Nous considérons l'application

$$\begin{aligned} \psi\colon \mathcal{L}(J, \mathbb{C}^*) &\to \mathbb{Z} \\ \gamma &\mapsto \text{Ind}(\gamma, 0) \end{aligned} \tag{26.189}$$

où $\mathcal{L}(J, \mathbb{C}^*)$ est l'ensemble des lacets $J \to \mathbb{C}^*$ sur lequel nous considérons la loi de groupe multiplicatif tandis que nous considérons la loi additive sur $\mathbb{Z}$. Nous avons :

ITEMooKCRIooSEyhlp

(1) ψ est effectivement à valeurs dans $\mathbb{Z}$,

(2) ψ est un morphisme de groupes surjectif.

ITEMooZFXTooDCXTVU

(3) Le noyau de ψ est

$$\ker(\psi) = \{\textit{lacets homotopes dans } \mathbb{C}^* \textit{ à un lacet constant}\}. \tag{26.190}$$

Démonstration. Pour (1), c'est la proposition 26.50(1).

Soient deux lacets γ_1 et γ_2 dans $\mathbb{C}^*$. Le théorème 26.46 nous permet de considérer des logarithmes continus $l_i\colon [a,b] \to \mathbb{C}$ le long de γ_i. L'application $l_1 + l_2$ est une détermination du logarithme le long de $\gamma_1\gamma_2$ parce que [29]

$$\exp\big(l_1(t) + l_2(t)\big) = \exp\big(l_1(t)\big)\exp\big(l_2(t)\big) = \gamma_1(t)\gamma_2(t). \tag{26.191}$$

29. Nous ne nous lasserons jamais de citer la proposition 18.9(2).

(1) **Morphisme** Nous avons

$$\begin{aligned}\mathrm{Ind}(\gamma_1\gamma_2,0) &= \frac{1}{2\pi i}\int_a^b f\big((\gamma_1\gamma_2)(t)\big)(\gamma_1\gamma_2)'(t)dt && (26.192a)\\ &= \frac{1}{2\pi i}\int_{\gamma_1\gamma_2}\frac{dz}{z} && (26.192b)\\ &= (l_1+l_2)(b)-(l_1+l_2)(a) && (26.192c)\\ &= \big(l_1(b)-l_1(a)\big)+\big(l_2(b)-l_1(a)\big) && (26.192d)\\ &= \mathrm{Ind}(\gamma_1,0)+\mathrm{Ind}(\gamma_2,0). && (26.192e)\end{aligned}$$

Nous avons prouvé que $\psi(\gamma_1\gamma_2)=\psi(\gamma_1)+\psi(\gamma_2)$, et donc que ψ est un morphisme.

(2) **Surjection** Nous considérons le lacet

$$\begin{aligned}\gamma\colon [a,b]&\to\mathbb{C}^*\\ t&\mapsto \exp\big(2i\pi kt(b-a)\big),\end{aligned} \tag{26.193}$$

et nous montrons que $\psi(\gamma)=k$. D'abord nous remarquons que

$$\gamma'(t)=\gamma(t)\frac{2i\pi k}{(b-a)}, \tag{26.194}$$

et ensuite nous calculons :

$$\begin{aligned}\psi(\gamma) &= \frac{1}{2\pi i}\int_a^b\frac{1}{\gamma(t)}\gamma'(t)dt && (26.195a)\\ &= \frac{1}{2\pi i}\frac{1}{\int_a^b}\gamma(t)\frac{2\pi ik}{(b-a)}\gamma(t)dt && (26.195b)\\ &= \frac{1}{2\pi i}\int_a^b\frac{2i\pi k}{(b-a)}dt && (26.195c)\\ &= k. && (26.195d)\end{aligned}$$

(3) **Noyau** Notons pour la simplicité

$$\mathcal{H}=\{\text{lacets homotopes dans } \mathbb{C}^* \text{ à un lacet constant}\}. \tag{26.196}$$

Nous notons u le chemin constant $u(t)=1$. Nous avons $\psi(u)=0$ parce que, en notant $f(z)=1/z$,

$$\mathrm{Ind}(u,0)=\frac{1}{2\pi i}\int_u f(z)dz=\frac{1}{2\pi i}\int_0^1 f\big(u(t)\big)u'(t)dt=0 \tag{26.197}$$

parce que $u'(t)=0$ pour tout t. Nous avons donc $u\in\ker(\psi)$. Si $\gamma\in\mathcal{H}$, alors, utilisant le théorème de Cauchy homotopique 26.17(2), nous avons $\mathrm{Ind}(\gamma,0)=\mathrm{Ind}(u,0)=0$. Donc $\mathcal{H}\subset\ker(\psi)$.

Soient $\gamma\in\ker(\psi)$ et une détermination continue l du logarithme le long de γ. Nous avons

$$0=\psi(\gamma)=l(b)-l(a). \tag{26.198}$$

Nous posons

$$\begin{aligned}H\colon [0,1]\times[a,b]&\to\mathbb{C}^*\\ (s,t)&\mapsto\exp\big(sl(a)+(1-s)l(t)\big).\end{aligned} \tag{26.199}$$

Cette application continue vérifie $H(0,t)=e^{l(t)}=\gamma(t)$ et $H(1,t)=e^{l(a)}=\gamma(a)$. Donc elle est une homotopie entre γ et le chemin constant $\gamma(a)$. Encore par Cauchy homotopique, $\psi(\gamma)=0$ parce que ψ est nul sur les chemins constants.

□

LEMooCTIBooNRAyZH

Lemme 26.56 ([1]).
Si γ est un lacet dans $\mathbb{C}\backslash\{p\}$, alors

$$\mathrm{Ind}(\gamma, p) = \mathrm{Ind}(\gamma - p, 0) \tag{26.200}$$

PROPooPRIIooKWCHBZ

Proposition 26.57 ([205]).
Soient $p \in \mathbb{C}$ ainsi que deux lacets γ_1 et γ_2 dans $\mathbb{C}\backslash\{p\}$. Si

$$\mathrm{Ind}(\gamma_1, p) = \mathrm{Ind}(\gamma_2, p), \tag{26.201}$$

alors γ_1 et γ_2 sont homotopes dans $\mathbb{C}\backslash\{p\}$.

Démonstration. Vu que γ_1 ne passe pas par p, nous pouvons considérer le chemin

$$\begin{aligned} \gamma\colon [a,b] &\to \mathbb{C}^* \\ t &\mapsto \frac{\gamma_0(t) - p}{\gamma_1(t) - p}. \end{aligned} \tag{26.202}$$

En posant $\sigma_i(t) = \gamma_i(t) - p$, nous avons

$$\begin{aligned} \mathrm{Ind}(\gamma, 0) &= \mathrm{Ind}(\sigma_0\sigma_1^{-1}, 0) && && (26.203a)\\ &= \mathrm{Ind}(\sigma_0, 0) + \mathrm{Ind}(\sigma_1^{-1}, 0) && \text{prop. 26.55} && (26.203b)\\ &= \mathrm{Ind}(\sigma_0, 0) - \mathrm{Ind}(\sigma_1, 0) && && (26.203c)\\ &= \mathrm{Ind}(\gamma_0, p) - \mathrm{Ind}(\gamma_1, p) && \text{lem. 26.56} && (26.203d)\\ &= 0. && && (26.203e) \end{aligned}$$

Vu que l'indice de γ est nul, ce chemin est homotope dans $\mathbb{C}^*$ au chemin constant u par la proposition 26.55(3). Soit une homotopie $H\colon [0,1] \times [a,b] \to \mathbb{C}^*$ entre γ et u. Et nous posons

$$\begin{aligned} S\colon [0,1] \times [a,b] &\to \mathbb{C}^* \\ (s,t) &\mapsto \big(\gamma_1(t) - p\big)H(s,t). \end{aligned} \tag{26.204}$$

Cette application est une homotopie entre

$$t \mapsto S(0,t) = \big(\gamma_2(t) - p\big)H(0,t) = \big(\gamma_2(t) - p\big)\gamma(t) = \gamma_1(t) - p \tag{26.205}$$

et

$$t \mapsto S(1,t) = \big(\gamma_2(t) - p\big)H(1,t) = \big(\gamma_2(t) - p\big)u(t) = \gamma_2(t) - p. \tag{26.206}$$

Donc $S + p$ est une homotopie entre γ_1 et γ_2 dans $\mathbb{C}\backslash\{p\}$. □

PROPooCFMFooXjlhfV

Proposition 26.58 ([205]).
Soit un espace topologique X localement connexe par arcs. Soit une application continue $f\colon X \to \mathbb{C}^*$. Il y a équivalence entre

(1) f admet un logarithme continu.

(2) Pour tout lacet γ dans X, nous avons $\mathrm{Ind}(f \circ \gamma, 0) = 0$.

26.2.10 Théorème de Cauchy et analycité

Cette sous-section veut prouver le théorème de Cauchy.

ThoUHztQe

Théorème 26.59 (Formule de Cauchy).
Soient Ω ouvert dans $\mathbb{C}$, $z_0 \in \Omega$ et f une fonction holomorphe sur Ω. Soit $r > 0$ tel que $B(z_0, r) \subset \Omega$. Alors pour tout $z \in B(z_0, r)$ nous avons

$$f(z) = \frac{1}{2\pi i}\int_{\partial B(z_0,r)} \frac{f(\omega)}{\omega - z} d\omega. \tag{26.207}$$

EqPzUABM

Démonstration. Soit $z \in B(z_0, r)$. Considérons la fonction

$$g(\omega) = \begin{cases} \frac{f(\omega)-f(z)}{\omega-z} & \text{si } \omega \neq z \\ f'(z) & \text{si } \omega = z. \end{cases} \tag{26.208}$$

Cette fonction est holomorphe sur $B(z_0, r)\backslash\{z\}$ et continue en z. Elle vérifie donc la proposition 26.10 et nous avons

$$\int_\gamma g = 0 \tag{26.209}$$

où γ est le cercle de centre z_0 et de rayon r. Nous avons donc

$$0 = \int_\gamma \frac{f(\omega)}{\omega - z} - \int_\gamma \frac{f(z)}{\omega - z}, \tag{26.210}$$

et ayant déjà calculé la seconde intégrale dans l'exemple 26.54 nous en déduisons

$$\int_\gamma \frac{f(\omega)}{\omega - z} d\omega = 2\pi i f(z), \tag{26.211}$$

ce qu'il fallait. □

ThomcPOdd

Théorème 26.60.
Soient Ω ouvert dans $\mathbb{C}$ et f holomorphe sur Ω. Soient encore $z_0 \in \Omega$ et r_0 tels que $B(z_0, r_0) \subset \Omega$. Alors :

ITEMooYWSOooHJtxGr

(1) Sur $B(z_0, r_0)$, la fonction f s'écrit

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \tag{26.212}$$

(2) Nous avons

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{(\omega - z_0)^{n+1}} d\omega \tag{26.213}$$

où $\gamma = \partial B(z_0, r)$ avec $|z - z_0| < r < r_0$.

ItemMRRTooMChmuZ

(3) En particulier f est infiniment dérivable.

Démonstration. Soit $r > 0$ tel que $|z - z_0| < r < r_0$. La formule de Cauchy (théorème 26.59) nous dit que

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{\omega - z} d\omega \tag{26.214}$$

où $\gamma = \partial B(z_0, r)$. Nous pouvons paramétrer ce chemin par $\omega = z_0 + re^{i\theta}$ et $\theta \in [0, 2\pi]$. Nous avons

$$f(z) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{z_0 + re^{i\theta} - z} rie^{i\theta} d\theta \tag{26.215a}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{1 - e^{-i\theta}(z - z_0)/r} d\theta. \tag{26.215b}$$

Nous pouvons développer l'intégrante en puissance de $(z - z_0)$ en utilisant la formule 15.462. Ici le rôle de x est tenu par

$$e^{-i\theta}(z - z_0)/r \tag{26.216}$$

dont le module est bien plus petit que 1, par hypothèse sur r. Nous avons donc

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \sum_{n=0}^{\infty} f(z_0 + re^{i\theta}) e^{-in\theta} r^{-n} (z - z_0)^n d\theta. \tag{26.217}$$

L'art est maintenant de permuter la somme et l'intégrale. Pour cela nous remarquons que ce qui se trouve dans la somme est majoré en module par

$$M\left|\frac{z-z_0}{r}\right|^n \tag{26.218}$$

FqbykTLD

où M est le maximum de $|f|$ sur γ. La borne (26.218) ne dépend pas de θ ; par conséquent la convergence de la somme est uniforme en θ par le critère de Weierstrass (théorème 12.372). Le théorème 15.1 s'applique[30] et nous pouvons permuter la somme avec l'intégrale.

Ce que nous trouvons est que

$$f(z)=\sum_{n=0}^{\infty}a_n(z-z_0)^n \tag{26.219}$$

où

$$a_n=\frac{1}{2\pi}\int_0^{2\pi}f(z_0+re^{i\theta})e^{-in\theta}r^{-n}d\theta=\frac{1}{2\pi i}\int_\gamma\frac{f(\omega)}{(\omega-z_0)^{n+1}}. \tag{26.220}$$

Cette formule est valable pour $|z-z_0|<r$. Sur cette boule, la fonction est donc une série entière. Le théorème de Taylor 15.137 nous permet donc d'affirmer que f est partout infiniment continument dérivable (parce que en chaque point on a un voisinage sur lequel c'est vrai), et d'identifier les coefficients (qui, eux, ne sont valables que localement) sous la forme

$$a_n=\frac{f^{(n)}(z_0)}{n!}. \tag{26.221}$$

□

¡¡ Avertissement/question au lecteur !! 26.61
L'énoncé du corolaire 26.62 n'est peut-être pas précis.

CORooJISDooFgwOPh

Corolaire 26.62.
Si le développement de f autour de 0 est

$$f(z)=\sum_{n\in\mathbb{N}}a_nz^n, \tag{26.222}$$

alors $a_0=f(0)$ et $a_1=f'(0)$.

CorwfHtJu

Corolaire 26.63.
Soit f une fonction continue sur un ouvert Ω telle que pour toute boule $B(a,r)$ contenue dans Ω, nous ayons

$$f(a)=\frac{1}{2\pi i}\int_{\partial B(a,r)}\frac{f(\xi)}{\xi-a}d\xi. \tag{26.223}$$

Alors f est holomorphe.

Démonstration. Il suffit de recopier la démonstration du théorème 26.60 pour savoir que f se développe en série de puissances et est donc en particulier dérivable. □

Le fait qu'une fonction holomorphe soit C^∞ comme dit dans la proposition 26.60 permet de démonter un résultat de dérivation sous l'intégrale, qui dépend de pouvoir majorer la différentielle.

PROPooZCLYooUaSMWA

Proposition 26.64.
Soit une fonction continue $g\colon\mathbb{R}\times\mathbb{C}\to\mathbb{C}$. Nous supposons que pour tout t, la fonction $z\mapsto g(t,z)$ est $\mathbb{C}$-dérivable (définition 12.304) et différentiable. Soit B compact dans $\mathbb{R}$ et la fonction

$$G(z)=\int_B g(t,z)dt. \tag{26.224}$$

30. Étant donné que nous savions déjà que la somme était une fonction intégrable, nous sommes loin d'avoir utilisé toute la puissance du théorème.

que nous supposons exister pour tout z.

Alors

$$G'(z) = \int_B g'(t,z)dt \tag{26.225}$$

où le prime réfère à la $\mathbb{C}$-dérivée par rapport à la variable z à t fixé.

Démonstration. Nous fixons $z \in \mathbb{C}$ et nous considérons la suite de fonctions

$$g_i(t) = \frac{g(t, z+\epsilon_i) - g(t,z)}{\epsilon_i} \tag{26.226}$$

où ϵ_i est une suite de nombres complexes tendant vers zéro ($\epsilon_i \xrightarrow{\mathbb{C}} 0$). Si la limite existe et ne dépend pas de la suite choisie, alors $\lim_{i\to\infty} g_i(t) = g'(t,z)$. Et vu que g est supposée dérivable, c'est le cas.

Nous avons aussi, par linéarité de l'intégrale :

$$G'(z) = \lim_{i\to\infty} \int_B g_i(t)dt. \tag{26.227}$$

La difficulté est de permuter la limite et l'intégrale. Pour cela nous allons utiliser la convergence dominée de Lebesgue (théorème 14.207). Afin de majorer $|g_i(t)|$ par une fonction intégrable en t (uniformément en i), nous exploitons le théorème des accroissements finis, théorème 10.292. En notant dg la différentielle de g par rapport à z à t fixé, pour chaque t et chaque i nous avons

$$|g(t,z+\epsilon_i) - g(t,z)| \leqslant \sup_{\xi\in[z,z+\epsilon_i]} \|dg_\xi\|\|\epsilon_i\|. \tag{26.228}$$

Vu que z est fixé et que ξ est dans le compact $[z, z+\epsilon_i]$ et que dg est continue (parce que la $\mathbb{C}$-dérivabilité implique la continuité de la différentielle parce que nous avons l'analycité par le théorème 26.60), nous pouvons majorer $\|dg_\xi\|$ par une constante $M_i(z)$ qui dépend à priori de i et de z.

Heureusement, nous pouvons prendre a fortiori le supremum sur $\overline{B(z,|\epsilon_i|)}$ (qui est tout autant compact) et supposer que $|\epsilon_i|$ est strictement décroissante ; de toutes façons, il y a un maximum parce que $|\epsilon_i| \to 0$. Dans ce cas, il suffit de prendre le supremum de $\|dg_\xi\|$ pour $\xi \in \overline{B(z,|\epsilon_1|)}$ et ça contente tout le monde.

Quoi qu'il en soit nous avons une constante $M(z)$ telle que

$$|g(t,z+\epsilon_i) - g(t,z)| \leqslant M(z)\|\epsilon_i\| \tag{26.229}$$

et donc $|g_i(t)| \leqslant M(z)$. La constante (par rapport à t) $M(z)$ est évidemment intégrable sur le compact B et nous pouvons permuter la limite avec l'intégrale :

$$G'(z) = \lim_{i\to\infty} \int_B g_i(t)dt = \int_B \lim_{i\to\infty} g_i(t)dt = \int_B g'(t,z)dt. \tag{26.230}$$

□

PropZOkfmO

Proposition 26.65.
Une fonction continue f est holomorphe si et seulement si la 1-forme différentielle $f(z)dz$ est localement exacte.

Démonstration. Si f est holomorphe, alors nous avons vu que f était différentiable et que $df_z = f(z)dz$ par la formule 12.850.

Dans le sens inverse, supposons que $f(z)dz$ est localement exacte, et soit F telle que $dF = f(z)dz$. Ce que nous allons faire est montrer que la dérivée de F existe et vaut f. En effet, la définition de la différentielle nous dit que

$$\lim_{h\to 0} \left| \frac{F(z+h) - F(z) - dF_z(h)}{h} \right| = 0. \tag{26.231}$$

La limite vaut évidemment encore zéro si nous enlevons les modules :

$$0 = \lim_{h\to 0} \frac{F(z+h) - F(z) - f(z)h}{h} \tag{26.232a}$$

$$= \lim_{h\to 0} \frac{F(z+h) - F(z)}{h} - f(z). \tag{26.232b}$$

Donc $F' = f$. Cela montre que F est $\mathbb{C}$-dérivable et donc holomorphe. En conséquence du théorème 26.60, la fonction F est infiniment dérivable et f l'est alors aussi. La fonction f est donc holomorphe[31]. □

26.2.11 Théorème de Brouwer en dimension 2

Pour d'autres versions du théorème de Brouwer, voir la sous-section 20.5.1.

ThoLVViheK

Théorème 26.66 (Brouwer en dimension 2[104])**.**
Soit $\mathcal{B} = \overline{B(0,1)}$ la boule unité fermée de $\mathbb{R}^2$. Alors toute application continue de $\mathcal{B}$ dans elle-même admet un point fixe.

Démonstration. Supposons que la fonction $f \in C^0(\mathcal{B}, \mathcal{B})$ n'admette pas de points fixes sur $\mathcal{B} = \overline{B(0,1)}$. Pour $x \in \mathcal{B}$ nous notons $g(x)$ l'intersection entre $\partial\mathcal{B}$ et la demi-droite allant de $f(x)$ vers x. C'est bien parce que f n'a pas de points fixes que g est bien définie.

En reprenant le même début de la preuve de la proposition 20.30 nous savons que la fonction

$$\begin{aligned} g\colon \overline{B(0,1)} &\to \partial B(0,1) \\ x &\mapsto \lambda(x)\big(x - f(x)\big) + f(x) \end{aligned} \tag{26.233}$$

est continue. De plus $g(x) = x$ sur $\partial B(0,1)$. Nous allons montrer qu'une telle fonction[32] ne peut pas exister.

Pour $s \in [0,1]$ nous paramétrons le cercle $\partial B(0,s)$ par

$$\begin{aligned} x_s\colon [0,1] &\to \partial B(0,s) \\ t &\mapsto \big(s\cos(2\pi t), s\sin(2\pi t)\big). \end{aligned} \tag{26.234}$$

Ensuite nous considérons les chemins

$$\begin{aligned} \gamma_s\colon [0,1] &\to \partial B(0,1) \\ t &\mapsto g \circ x_s(t). \end{aligned} \tag{26.235}$$

L'application γ_s est continue et $\gamma_s(0) = \gamma_s(1)$. Les chemins γ_s sont des lacets ; nous nous intéressons maintenant à l'indice au point 0 de γ_0 et γ_1. D'une part $\gamma_0(t) = g(0)$ (lacet constant) et $\gamma_1(t) = e^{2i\pi t}$ (parce que $g(x) = x$ sur le bord). Nous avons donc, en utilisant l'indice de la définition 26.50,

$$\mathrm{Ind}_{\gamma_0}(0) = \frac{1}{2\pi i}\int_{\gamma_0} \frac{d\omega}{\omega} = \frac{1}{2\pi i}\int_0^1 \frac{\gamma_0'(t)}{\gamma_0(t)}dt = 0, \tag{26.236}$$

alors que

$$\mathrm{Ind}_{\gamma_1}(0) = \frac{1}{2\pi i}\int_0^1 \frac{2i\pi e^{2i\pi t}}{e^{2i\pi t}}dt = 1. \tag{26.237}$$

Nous considérons l'homotopie

$$\begin{aligned} \gamma\colon [0,1]\times[0,1] &\to \overline{B(0,1)} \\ (s,t) &\mapsto \gamma_s(t) = (g\circ x_s)(t). \end{aligned} \tag{26.238}$$

Nous avons $g(0) \neq 0$ parce que g prend ses valeurs sur le bord. Vu que c'est une équivalence d'homotopie[33] entre γ_1 et γ_2, les indices devraient être égaux par le corolaire 26.18. □

31. Dire que la dérivée d'une fonction holomorphe est holomorphe est un raisonnement classique.
32. Qui est nommée *rétraction* de la sphère sur elle-même.
33. Définition 26.53

26.2.12 Principe des zéros isolés

LEMooYYZQooClmOgG

Lemme 26.67.
Si f est une fonction holomorphe[34] *sur le compact K, alors il existe un polynôme P_f et une fonction holomorphe h_f ne s'annulant pas sur K telles que $f = h_f P_f$.*

Démonstration. Soit une fonction f vérifiant les conditions. Si f est identiquement nulle, alors il suffit de prendre $P_f = 0$ et c'est fait. Nous supposons donc que f n'est pas identiquement nulle.

(1) **Quantité finie de racines**
D'abord f ne peut s'annuler qu'un nombre fini de fois sur K. Sinon, on pourrait considérer une suite des racines[35] de f dans K. Vu qu'une suite dans un compact contient une sous-suite convergente (théorème 7.144), la fonction f aurait un point d'accumulation de racines. Alors le principe des zéros isolés (théorème 17.136) nous donne un ouvert sur lequel f est nulle et donc le corolaire 17.137 nous dit que f est identiquement nulle.

(2) **Autour d'une racine**
Bref, la fonction f possède un nombre fini de racines sur K. Soit z_0 l'une d'elles.
Par le théorème 26.60(1), nous avons, sur un voisinage de z_0 :

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \tag{26.239}$$

En particulier, $0 = f(z_0) = a_0$. Donc $a_0 = 0$. Soit k, le plus petit naturel pour lequel $a_k \neq 0$. Nous avons

$$f(z) = (z - z_0)^k g(z) \tag{26.240}$$

avec $g(z) = \sum_{n=0}^{\infty} a_{k+n}(z - z_0)^n$.. Vu que $a_k \neq 0$ nous avons $g(z_0) \neq 0$. Montrons à présent que g est holomorphe sur un voisinage de z_0. Vu que la série définissant g est une sous-série d'une série convergente sur un voisinage, elle converge sur un voisinage et la proposition 26.11 nous dit que g est $\mathbb{C}$-dérivable. C'est-à-dire holomorphe par définition.

(3) **Autour de toutes les racines**
Soient (z_i) les racines (en nombre fini). Pour chaque i nous avons une boule $B(z_i, r_i)$ sur laquelle $f = P_i g_i$ où P_i est un polynôme de la forme $(z - z_i)^k$ et g_i est holomorphe sur $B(z_i, r_i)$. Nous définissons la fonction suivante :

$$h(z) = \begin{cases} \dfrac{f(z)}{\prod_k P_k(z)} & \text{si } z \neq z_i \\ \dfrac{g_i}{\prod_{k \neq i} P_k(z)} & \text{si } z = z_i. \end{cases} \tag{26.241}$$

Cette fonction ne s'annule jamais. Mais est-elle holomorphe ?
Si $z \neq z_i$ (sous-entendu : pour tout i), alors sur un voisinage, $h = f / \prod P_k$ qui est un quotient de fonctions holomorphes dont le dénominateur ne s'annule pas. Elle est donc holomorphe sur ce voisinage par le lemme 12.313.
Pour les autres notons que pour tout $z \in B(z_i, r_i)$,

$$h = \frac{g_i}{\prod_{k \neq i} P_k}. \tag{26.242}$$

Cela est encore un quotient dont le dénominateur ne s'annule pas[36].

(4) **La réponse**
Nous avons, pour tout $z \in K$:

$$f(z) = h(z) \prod_k P_k(z). \tag{26.243}$$

34. Définition 12.306.
35. Notez l'utilisation de la proposition 1.121 que je vous invite à ne pas considérer comme une trivialité absolue.
36. Nous avons choisi les r_i de telle sorte que les boules ne s'intersectent pas.

□

Afin de détendre l'atmosphère, nous allons laisser tomber l'analyse quelques instants et prouver un résultat d'algèbre.

PROPooVWRPooGQMenV

Proposition 26.68 ([1, 639])**.**
L'anneau des fonctions holomorphes sur un compact[37] *donné de* $\mathbb{C}$ *est principal*[38]*.*

Démonstration. Nous nommons A l'ensemble des fonctions holomorphes sur le compact K, et J un idéal de A.

(1) **A est un anneau**
Le point délicat de la définition 1.42 est le fait que la somme et le produit d'éléments de A sont des éléments de A parce que les résultats type « la somme de deux fonctions holomorphes est holomorphes sont valides sur des ouverts alors que nous sommes ici sur un compact. Soient f et g dans A ; nous nommons Ω_f et Ω_g des ouverts contenant K tels que f est holomorphe sur Ω_f et g sur Ω_g.
L'ensemble $\Omega_f \cap \Omega_g$ est un ouvert (intersection d'ouverts) contenant K et sur lequel f et g sont holomorphes. Donc $f+g$ et fg y sont holomorphes.

(2) **Engendré par des polynômes**
Pour chaque $f \in J$ nous écrivons $f = P_f h_f$ en vertu de la décomposition donnée par le lemme 26.67. Vu que h_f ne s'annule pas, $1/h_f$ est encore holomorphe sur K et nous déduisons que $P_f = f/h_f$ est dans J. La partie

$$S = \{P_f \text{ tel que } f \in J\} \tag{26.244}$$

est génératrice de J parce que, par construction, tous les éléments de J sont des produits d'éléments de S par des fonctions holomorphes sur K (donc, des éléments de A). Mais tous les éléments de S sont dans J, donc $(S) = J$.

(3) **Un polynôme pour tous les engendrer**
Soit M, l'idéal de $\mathbb{C}[X]$ engendré par S. Attention : J est l'idéal de A engendré par S. Mais l'idéal de $\mathbb{C}[X]$ engendré par S est peut-être autre chose. Vu que $\mathbb{C}$ est un corps, le lemme 3.110 dit que $\mathbb{C}[X]$ est principal. Donc M est un idéal principal de $\mathbb{C}[X]$ et nous avons un polynôme $p \in \mathbb{C}[X]$ tel que

$$M = \mathbb{C}[X]p. \tag{26.245}$$

Si vous avez compris le chausse trappe, vous saurez pourquoi il faut écrire $M = \mathbb{C}[X]p$ et non utiliser l'écriture plus simple « $M = (p)$ ».

(4) $A\mathbb{C}[X] = A$
L'inclusion $A \subset A\mathbb{C}[X]$ est dûe au fait que $1 \in \mathbb{C}[X]$, et l'autre inclusion est le fait que $\mathbb{C}[X] \subset A$ alors que A est un anneau.

(5) **Suite des opérations**
Nous avons :

$$J = AS \subset A\mathbb{C}[X]p. \tag{26.246}$$

Voilà une inclusion de montrée. Reste à faire l'autre.
Vu que $p \in J$ nous avons aussi $Ap \subset J$. Et donc

$$A\mathbb{C}[X]p = Ap \subset J. \tag{26.247}$$

Avec ces deux inclusions, $J = A\mathbb{C}[X]p = Ap$. Donc J est engendré par un seul élément et est principal.

□

37. Être holomorphe sur un compact signifie qu'il existe une extension holomorphe à un ouvert contenant le compact.
38. Définition 1.228

26.2.13 Prolongement de fonctions holomorphes

PropDRnYkKP

Proposition 26.69.
Soient Ω un ouvert de $\mathbb{C}$ et $f\colon \Omega \to \mathbb{C}$ une fonction holomorphe sur $\Omega\backslash\{a\}$ ($a \in \Omega$). Nous supposons qu'il existe $r > 0$ tel que f est bornée sur $B(a,r) \cap \Omega$. Alors f se prolonge en une fonction holomorphe sur Ω.

Le théorème de prolongement de Riemann 26.89 donnera plus d'informations.

Démonstration. Nous définissons la fonction $g\colon \Omega \to \mathbb{C}$ par

$$g(z) = \begin{cases} (z-a)f(z) & \text{si } z \neq a \\ 0 & \text{si } z = a. \end{cases} \tag{26.248}$$

Sur $\Omega\backslash\{a\}$, la fonction g est holomorphe (produit de fonctions holomorphes), et elle est continue en a. Par conséquent elle est holomorphe sur Ω. Nous la développons en série entière sur une boule $B(a,r)$:

$$g(z) = \sum_{n=0}^{\infty} c_n (z-a)^n. \tag{26.249}$$

Nous avons $g(a) = c_0 = 0$. Nous posons

$$\varphi(z) = \sum_{n=0}^{\infty} c_{n+1}(z-a)^n. \tag{26.250}$$

Si $z \neq a$, alors $\varphi(z) = f(a)$ parce que $\varphi(z) = g(z)/(z-a)$. Mais φ est continue en a, et donc holomorphe en a.

La fonction φ est par conséquent un prolongement holomorphe de f en a. □

26.2.14 Théorème de Runge

Le théorème que nous allons prouver n'est en réalité qu'une partie de ce qui est usuellement appelle le théorème de Runge.

ThoMvMCci

Théorème 26.70 (Théorème de Runge)**.**
Soit K, un compact de $\mathbb{C}$ tel que $\complement K$ soit connexe. Si $a \in \complement K$ alors la fonction

$$\varphi_a(z) = \frac{1}{z-a} \tag{26.251}$$

est limite uniforme de polynômes sur K.

Démonstration. Nous considérons $P(K)$, l'adhérence des polynômes sur K pour la norme uniforme (sur K). Nous devons montrer que pour tout $a \in \complement K$, la fonction φ_a est dans $P(K)$. Pour cela nous considérons l'ensemble

$$A = \{a \in \complement K \text{ tel que } \varphi_a \in P(K)\} \tag{26.252}$$

et nous allons montrer qu'il est à la fois non vide, ouvert et fermé dans le connexe $\complement K$.

Je répète : nous allons prouver l'ouverture et la fermeture *pour la topologie de* $\complement K$. Nous n'allons pas prouver que A est un ouvert de $\mathbb{C}$. Ce qui sera par conséquent prouvé est que $A = \complement K$.

(1) **Non vide** Soit $R = \sup_{z\in K} |z|$ et $a \in \complement K$ tel que $|a| > R$. Nous avons

$$\varphi_a(z) = \frac{1}{a}\frac{1}{\frac{z}{a}-1} = -\frac{1}{a}\frac{1}{1-\frac{z}{a}} = -\frac{1}{a}\sum_{k=0}^{\infty}\left(\frac{z}{a}\right)^k = -\sum_{k=0}^{\infty}\frac{z^k}{a^{k+1}}. \tag{26.253}$$

Ici la convergence de la série et sa limite sont assurées par le fait que $|z/a| < 1$ par choix de R et a. La suite de polynômes

$$P_n(z) = \sum_{k=0}^{n}\frac{z^k}{a^{k+1}} \tag{26.254}$$

converge uniformément sur $B(0,R)$ et en particulier sur K. Donc $P_n \to \varphi_a$.

(2) **Fermé**

Nous allons montrer que la fermeture de A (dans $\complement K$) est incluse dans A, et donc qu'elle est égale à A et donc que A est fermé. Par le lemme 7.30, la fermeture de A dans $\complement K$ est l'ensemble $\bar{A} \cap \complement K$ où $\bar{A}$ est la fermeture de A au sens usuel.

Bref, soit $a \in \bar{A} \cap \complement K$, et montrons que $\varphi_a \in \overline{P(K)}$. Vu que $P(K)$ est déjà une fermeture, nous aurons en fait $\varphi_a \in P(K)$ et donc $a \in A$, ce qui signifierait que $\bar{A} \cap \complement A = A$ et donc que A est fermé.

Au travail.

Soit $(a_n) \in A$ une suite convergente vers a. Soit aussi $d = d(a, K)$; on a $d > 0$ parce que K est compact et a est hors de K alors le complémentaire de K est ouvert. Nous choisissons en plus la suite a_n pour avoir $|a_n - a| < \frac{d}{2}$; au pire on prend la queue de suite. Soit $z \in K$; nous avons

$$|\varphi_{a_n}(z) - \varphi_a(z)| = \left| \frac{1}{z - a_n} - \frac{1}{z - a} \right| = \left| \frac{a_n - a}{(z - a_n)(z - a)} \right|. \tag{26.255}$$

Vu que $a_n \in B(a, \frac{d}{2})$ et que $z \in K$ et $d = d(a, K)$ nous avons $|a_n - z| \geqslant \frac{d}{2}$; et aussi $|a - z| \geqslant \frac{d}{2}$. Nous pouvons donc majorer (26.255) par

$$|\varphi_{a_n}(z) - \varphi_a(z)| \leqslant 2 \frac{|a_n - a|}{d^2}. \tag{26.256}$$

Donc nous avons

$$\| \varphi_a - \varphi_{a_n} \|_K \leqslant 2 \frac{|a_n - a|}{d^2} \to 0 \tag{26.257}$$

où la norme $\|.\|_K$ est la norme supremum sur K. Donc $a \in \overline{P(K)} = P(K)$ et A est fermé.

(3) **Ouvert** Vu que K est compact, il est fermé et donc $\complement K$ est ouvert. Par conséquent, ainsi que précisé dans l'exemple 7.32, les ouverts de $\complement K$ sont les ouverts de $\mathbb{C}$ contenus dans $\complement K$. Afin de prouver que A est ouvert, nous prenons $a \in A$ et nous cherchons une boule (au sens de $\mathbb{C}$) autour de a qui serait incluse dans A.

Soit donc $h \in \mathbb{C}$ « petit » dans un sens que nous allons préciser plus tard. Encore une fois nous posons $d = d(a, K)$. Nous avons

$$\varphi_{a+h}(z) = \frac{1}{z - a - h} = \frac{1}{z - a} \frac{1}{1 - \frac{h}{z-a}} = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}}. \tag{26.258}$$

Déjà ici nous demandons $h < \sup_{z \in K} |z - a|$. Puisque $|z - a| > d$, nous avons alors

$$|\varphi_{a+h}(z)| \leqslant \sum_{k=0}^{\infty} \frac{h^k}{d^{k+1}} < \infty. \tag{26.259}$$

Cela pour dire que la somme à droite de (26.258) converge bien pourvu que h soit bien petit. Nous pouvons donc poursuivre :

$$\varphi_{a+h}(z) = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}} = \sum_{k=0}^{\infty} h^k \varphi_a(z)^{k+1}. \tag{26.260}$$

Nous montrons maintenant que la convergence de la somme (26.260) est en réalité uniforme en z. En effet

$$\left| \varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1} \right| = \left| \sum_{k=N+1}^{\infty} h^k \varphi_a(z)^{k+1} \right| \tag{26.261a}$$

$$\leqslant \sum_{k=N+1}^{\infty} |h|^k |\varphi_a(z)|^{k+1}. \tag{26.261b}$$

Étant donné que φ_a est continue sur le compact K, elle y est majorée en module ; on peut même être plus précis :

$$|\varphi_a(z)| = \frac{1}{|z-a|} \leqslant \frac{1}{d}. \tag{26.262}$$

Nous pouvons donc écrire

$$\left|\varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1}\right| \leqslant \frac{1}{d} \sum_{k=N+1}^{\infty} \left|\frac{h}{d}\right|^k. \tag{26.263}$$

Étant donné que la somme $\sum_{k=0}^{\infty} |h/d|^k$ converge, la limite $N \to \infty$ est nulle et nous avons

$$\lim_{N\to\infty} \|\varphi_{a+h} - \sum_{k=0}^{N} h^k \varphi_a^{k+1}\|_K = 0. \tag{26.264}$$

Pour avoir $\varphi_{a+h} \in P(K)$, il faut encore savoir si les fonctions φ_a^k sont dans $P(K)$ pour tout k. Dans ce cas pour chaque N la somme sera encore dans $P(K)$ et φ_{a+h} sera limite uniforme d'éléments de $P(K)$.

Par hypothèse, $\varphi_a \in P(K)$; soit P_n une suite de polynômes qui converge uniformément vers φ_a. Nous allons montrer qu'alors la suite de polynômes P_n^k converge uniformément vers φ_a^k. Soit n tel que $\|P_n - \varphi_a\|_K \leqslant \epsilon$ et utilisons le produit remarquable

$$a^k - b^k = (a-b) \sum_{i=0}^{k-1} a^i b^{k-1-i} \tag{26.265}$$

pour obtenir

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant |P_n(z) - \varphi_a(z)| \sum_{i=0}^{k-1} |P_n(z)^i \varphi_a(z)^{k-1-i}|. \tag{26.266}$$

Vu que P_n et φ_a sont continues sur le compact K, on peut majorer la somme par une constante M, et il restera

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant M|P_n(z) - \varphi_a(z)|, \tag{26.267}$$

ou encore

$$\|P_n^k - \varphi_a^k\| \leqslant M\epsilon. \tag{26.268}$$

Cela prouve que $\varphi_a^k \in P(K)$ et donc que φ_{a+h} est limite uniforme (sur K) d'éléments de $P(K)$ et donc fait partie de $P(K)$ lui aussi.

Ceci achève de prouver que A est ouvert dans $\complement K$.

(4) **Conclusion**

L'ensemble A est non vide, ouvert et fermé dans $\complement K$, donc il est égal à $\complement K$. Le théorème est ainsi démontré.

□

26.3 Intégrales de fonctions holomorphes

Nous commençons par le lemme technique.

LemNAnweA

Lemme 26.71 ([538]).
Soit f une fonction holomorphe sur $B(z_0, r_0)$. Pour tout $z \in B(z_0, r)$ (avec $r < r_0$) nous avons

$$|f'(z)| \leqslant \frac{r}{\big(r - |z-z_0|\big)^2} \max\big\{f(z_0 + re^{i\theta})\big\}_{\theta\in\mathbb{R}}. \tag{26.269}$$

Démonstration. Par translation nous pouvons supposer que $z_0 = 0$. Étant donné que f est holomorphe, elle admet un développement en séries entières

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{26.270}$$

et nous notons $M = \max\{f(z) \text{ tel que } z \in \overline{B(0,r)}\}$. Nous avons $r^n|a_n| \leqslant M$. Par conséquent

$$|f'(z)| = \left| \sum_{n=1}^{\infty} n a_n z^{n-1} \right| \tag{26.271a}$$

$$\leqslant \frac{1}{r} \sum r^n |a_n| n \left(\frac{|z|}{r} \right)^{n-1} \tag{26.271b}$$

$$< \frac{M}{r} \sum n \left(\frac{|z|}{r} \right)^{n-1}. \tag{26.271c}$$

À ce point nous devons utiliser la série de l'exemple 15.145. Nous avons alors

$$|f'(z)| \leqslant \frac{M}{r} \frac{1}{\left(1 - \frac{|z|}{r}\right)^2} = \frac{Mr}{(r - |z|)^2}. \tag{26.272}$$

□

26.3.1 Holomorphie sous l'intégrale

26.72.
Notez une grande différence entre les théorèmes 26.73 et 17.19 : la condition (17.76) demande de contrôler l'intégrabilité de la dérivée de f, alors que la condition (26.274) demande de contrôler l'intégrabilité de f elle-même. Oh oui, on voudrait faire de l'analyse, mais ces fonctions holomorphes . . .tellement déloyal !!

ThopCLOVN

Théorème 26.73 (Holomorphie sous l'intégrale[538])**.**
Soit un espace mesuré (Ω, μ), un ouvert A dans $\mathbb{C}$ et une fonction $f\colon A \times \Omega \to \mathbb{C}$. Nous voulons étudier la fonction

$$F(z) = \int_{\Omega} f(z, \omega) d\mu(\omega) \tag{26.273}$$

pour tout $z \in A$. Nous supposons que

1. *la fonction $f(., \omega)$ est holomorphe sur A pour chaque ω.*
2. *La fonction $f(z, .)$ est mesurable sur (Ω, μ).*
3. *Pour tout compact $K \subset A$, il existe une fonction $g_K\colon \Omega \to \mathbb{R}$ telle que $|f(z, \omega)| \leqslant g_K(\omega)$ et telle que*

$$\int_{\Omega} g_K(\omega) d\mu(\omega) \tag{26.274}$$

EQooYILMooAIBQof

existe.

Alors la fonction F est holomorphe et

$$F'(z) = \int_{\Omega} \frac{\partial f}{\partial z}(z, \omega) d\mu(\omega). \tag{26.275}$$

Démonstration. Soient $z_0 \in A$ et $r > 0$ tels que $K = \overline{B(z_0, r)} \subset A$. Pour chaque $\omega \in \Omega$ nous considérons la fonction

$$\begin{aligned} f_\omega\colon \overline{B(z_0, r)} &\to \mathbb{C} \\ z &\mapsto f(z, \omega). \end{aligned} \tag{26.276}$$

Étant donné que $\overline{B(z_0,r)}$ est compacte, la fonction $|f_\omega|$ est majorée par un nombre que nous notons $f_K(\omega)$ qui est indépendant de z (pour autant que $z\in K$). Nous désignons par $S(z_0,r)$ la frontière de la boule $B(z_0,r)$. Étant donné que la majoration est valable sur $\overline{B(z_0,r)}$, nous avons en particulier

$$|f_\omega(z)|\leqslant f_K(\omega) \tag{26.277}$$

pour tout $z\in S$. En utilisant la lemme 26.71 nous avons

$$|f'_\omega(z)|\leqslant \frac{r}{(r-|z-z_0|)^2}\max\{f(z_0+re^{i\theta})\}_{\theta\in\mathbb{R}} \tag{26.278a}$$

$$\leqslant \frac{rf_K(\omega)}{(r-|z-z_0|)^2}. \tag{26.278b}$$

Cette majoration est valable pour tout $z\in B(z_0,r)$. Si nous supposons de plus que $z\in B(z_0,r/2)$ nous avons

$$|f'(z)|\leqslant \frac{rf_K(\omega)}{\left(r-\frac{r}{2}\right)^2}=\frac{4}{r}f_K(\omega). \tag{26.279}$$

Étant donné que la boule $B(z_0,r/2)$ est convexe, la fonction f_ω est Lipschitz et pour tout $h\in\mathbb{C}$ tel que $|h|<r/2$ nous avons

$$\left|\frac{f_\omega(z_0+h)-f_\omega(z_0)}{h}\right|\leqslant\frac{4f_K(\omega)}{r}. \tag{26.280}$$

Soit maintenant une suite (h_n) qui converge vers 0 dans $\mathbb{C}$. Nous considérons la suite de fonctions correspondantes

$$g_n(\omega)=\frac{f(z_0+h_n,\omega)-f(z_0,\omega)}{h_n}. \tag{26.281}$$

Cette suite de fonctions vérifie la convergence ponctuelle

$$g_n(\omega)\to\frac{\partial f}{\partial z}(z_0,\omega). \tag{26.282}$$

De plus g_n est une fonction (de ω) dominée par $\frac{4f_K}{r}$ qui est intégrable. Par conséquent le théorème de la convergence dominée[39] nous indique que

$$\int_\Omega g_n(\omega)d\mu(\omega)\to\int_\Omega\frac{\partial f}{\partial z}(z_0,\omega)d\mu(\omega), \tag{26.283}$$

tandis que

$$F'(z)=\lim_{n\to\infty}\frac{F(z_0+h_n)-F(z_0)}{h_n}=\lim_{n\to\infty}\int_\Omega g_n(\omega)d\mu(\omega). \tag{26.284}$$

□

26.74.
Dans le corolaire suivant, l'intégrale sur $\partial B(z,r)$ est une intégrale sur un chemin.

CorNxTjEj

Corolaire 26.75.
Si f est une fonction holomorphe sur un ouvert Ω contenant la fermeture de la boule $B(z_0,r)$, alors pour tout z dans $B(z_0,\rho)$ ($\rho<r$) les dérivées de f s'expriment par la formule suivante :

$$f^{(k)}(z)=\frac{k!}{2\pi i}\int_{\partial B(z_0,r)}\frac{f(\xi)}{(\xi-z)^{k+1}}d\xi. \tag{26.285}$$

EQooBPIQooNhQTtB

Démonstration. Nous faisons par récurrence.

39. Lebesgue, théorème 14.207.

(1) **Pour la dérivée première** Soit $\rho < r$. Nous utilisons le théorème 26.59 pour écrire

$$f(z) = \frac{1}{2\pi i}\int_{\partial B(z_0,r)} \frac{f(\xi)}{\xi - z}d\xi. \tag{26.286}$$

Ensuite nous dérivons sous l'intégrale en appliquant le théorème 26.73 à la fonction

$$\begin{aligned} g\colon B(z_0,\rho)\times\partial B(z_0,r) &\to \mathbb{C} \\ (z,\xi) &\mapsto \frac{f(\xi)}{\xi - z}. \end{aligned} \tag{26.287}$$

Étant donné que f est holomorphe, la fonction g est continue et donc bornée sur tout compact $K \subset A$ par une constante M (qui dépend du compact choisi). D'autre part, nous avons toujours $|\xi - z| > r - \rho$ et donc

$$|g(z,\xi)| \leqslant \frac{M}{r-\rho}. \tag{26.288}$$

La fonction constante $g_K = \frac{M}{r-\rho}$ est évidemment intégrable. Le théorème conclut que f est holomorphe (cela, nous le savions déjà [40]), et

$$f'(z) = \frac{1}{2i\pi}\int_{\partial B} \frac{f(\xi)}{(\xi - z)^2}d\xi. \tag{26.289}$$

(2) **Les dérivées suivantes** Pour la récurrence[640] nous supposons que

$$f^{(k)}(z) = \frac{k!}{2i\pi}\int_{\partial B(z_0,r)} \frac{f(\xi)}{(\xi - z)^{k+1}}d\xi, \tag{26.290}$$

et nous tentons de calculer $f^{(k+1)}(z)$. Pour cela nous paramétrons l'intégrale de façon très usuelle :

$$f^{(k)}(z) = \frac{k!}{2\pi i}\int_0^{2\pi} \frac{f(re^{it})}{(re^{it} - z)}ire^{it}dt. \tag{26.291}$$

Nous permettons de permuter la $\mathbb{C}$-dérivation (par rapport à z) et l'intégrale en vertu de la proposition 26.64 appliquée à la fonction

$$g(t,z) = \frac{f(re^{it})}{(re^{it} - z)^{k+1}}ire^{it}. \tag{26.292}$$

Cela donne

$$f^{(k+1)}(z) = \frac{k!}{2i\pi}\int_0^{2\pi} f(re^{it})ire^{it}\frac{k+1}{(re^{it} - z)^{k+1}}dt = \frac{(k+1)!}{2i\pi}\int_{\partial B}\frac{f(\xi)}{(\xi - z)^{k+2}}d\xi. \tag{26.293}$$

□

THOooSULFooHTLRPE

Théorème 26.76.
Si f est une fonction holomorphe sur le disque ouvert $B(z_0, R)$ alors

$$f(z) = \sum_{n=0}^{\infty}\frac{f^{(n)}(z_0)}{n!}(z - z_0)^n \tag{26.294}$$

et cette série converge uniformément sur tout compact.

40. Et cela fournit une preuve alternative à la réciproque du théorème de Cauchy : une fonction continue qui vérifie la formule de Cauchy est holomorphe.

Démonstration. Sans perte de généralité nous supposons que $z_0 = 0$. La formule de Cauchy (théorème 26.59) fournit, pour $z \in B(0,R)$,

$$f(z) = \frac{1}{2\pi i}\int_{\partial B} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i}\int_{\partial B} \frac{f(\xi)}{1-(z/\xi)}\frac{d\xi}{\xi}. \tag{26.295}$$

En particulier notons que $z \in B(0,R)$ alors que ξ est sur le bord de cette boule ouverte. Donc $|\xi| > |z|$ pour tous les ξ et z qui interviennent. Nous utilisons la série géométrique

$$\frac{1}{1-(z/\xi)} = \sum_{n=0}^{\infty} \left(\frac{z}{\xi}\right)^n. \tag{26.296}$$

(1) **Permuter une intégrale et une somme** En utilisant la mesure de comptage[41] sur $\mathbb{N}$ (qui est σ-finie), nous pouvons écrire

$$\int_{\partial B} \sum_{n=0}^{\infty} \frac{z^n f(\xi)}{\xi^{n+1}} d\xi = \int_{\partial B} \left(\int_{\mathbb{N}} g(\xi, n) dm(n) \right) d\xi \tag{26.297}$$

où

$$\begin{aligned} g\colon \partial B \times \mathbb{N} &\to \mathbb{C} \\ (\xi, n) &\mapsto \frac{z^n f(\xi)}{\xi^{n+1}}. \end{aligned} \tag{26.298}$$

Nous allons permuter les intégrales en utilisant le théorème de Fubini, selon la procédure décrite en 14.305. Nous commençons par l'intégrale sur $\mathbb{N}$:

$$\int_{\mathbb{N}} |g(n,\xi)| = |\frac{f(\xi)}{\xi}| \sum_{n\in\mathbb{N}} |\frac{z}{\xi}|^n = \frac{1}{R}|f(\xi)|\frac{1}{1-|z|/R}. \tag{26.299}$$

Ici nous avons utilisé $|\xi| = R$. Notons que z est fixé depuis longtemps à l'intérieur de la boule de rayon R de telle sorte que $|z/\xi|$ est une constante strictement inférieure à 1.

L'intégrale sur $\xi \in \partial B$ n'a pas à être effectuée explicitement : nous nous contentons de prouver qu'elle est finie. La fonction f est continue sur le compact ∂B. Cela parce que B est une boule fermée dans l'ouvert Ω sur lequel f est continue. Au final l'expression à droite de (26.299) est bornée sur le compact ∂B et son intégrale donne un nombre fini.

Tout ceci pour invoquer le corolaire 14.303 qui nous indique que $g \in L^1(\mathbb{N} \times \partial B)$.

Une fois g intégrable sur l'espace produit $\mathbb{N} \times \partial B$, nous pouvons utiliser Fubini 14.304 pour permuter les intégrales.

Une fois la somme et l'intégrale permutées, nous avons

$$f(z) = \frac{1}{2\pi i}\sum_{n=0}^{\infty} \int_{\partial B} \frac{z^n f(\xi)}{\xi^{n+1}} = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi^{n+1}} \right) z^n. \tag{26.300}$$

Nous devons maintenant montrer que ce qui se trouve dans la grande parenthèse vaut $f^{(n)}(0)/n!$. Cela est immédiat en comparant avec la formule (26.285).

□

Proposition 26.77 (Morera [641]).
Soit Ω ouvert dans $\mathbb{C}$ et f continue. Si

$$\int_{\partial T} f = 0 \tag{26.301}$$

pour tout triangle (plein) T contenu dans Ω, alors f est holomorphe sur Ω.

41. Définition 14.263.

Démonstration. Il est suffisant de prouver que f est holomorphe sur toute boule ouverte $B(a,r)$ incluse dans Ω. Nous posons, pour tout $z \in B(a,r)$,

$$F(z) = \int_{[a,z]} f, \tag{26.302}$$

et nous considérons le chemin triangulaire $a \to z \to z+h \to a$ où $h \in \mathbb{C}$ est choisi assez petit pour que $z+h \in B(a,r)$. L'intégrale sur le triangle étant nulle, nous avons

$$0 = \int_{a\to z} f + \int_{z\to z+h} f + \int_{z+h\to a} f, \tag{26.303}$$

c'est-à-dire

$$F(z+h) - F(z) = \int_{z\to z+h} f. \tag{26.304}$$

En paramétrant le chemin par $z+th$ avec $t \in [0,1]$, et en tenant compte de la remarque 20.44,

$$F'(z) = \lim_{h\to 0} \frac{F(z+h)-F(z)}{h} \tag{26.305a}$$

$$= \lim_{h\to 0} \frac{1}{h} \int_0^1 f(z+th)h dt, \tag{26.305b}$$

ce qui prouve que F est dérivable et $F' = f$. Par définition (12.305), F est holomorphe, et donc C^∞ par le théorème 26.60. Du coup f est également C^∞ et donc en particulier holomorphe. □

26.3.2 Mesure de Radon

Soit un compact K de $\mathbb{C}$. Le lemme 12.360 nous dit que $\big(C(K), \|.\|_\infty\big)$ est complet. Mais cela n'est pas vraiment utile pour l'instant. Je dis ça juste pour le garder dans un coin de la tête.

DEFooTLRKooGbKkAj

Définition 26.78.
*Une **mesure de Radon** sur un compact K de $\mathbb{C}$ est une forme linéaire continue sur $\big(C(K), \|.\|_\infty\big)$. Si μ est une mesure de Radon, on définit la **transformée de Cauchy** de μ par*

$$\begin{aligned} \hat{\mu}\colon \mathbb{C}\backslash K &\to \mathbb{C} \\ z &\mapsto -\frac{1}{\pi}\mu(f_z) \end{aligned} \tag{26.306}$$

où, pour tout $z \in \mathbb{C}\backslash K$ nous avons défini la fonction f_z par

$$\begin{aligned} f_z\colon K &\to \mathbb{C} \\ \xi &\mapsto \frac{1}{\xi - z}. \end{aligned} \tag{26.307}$$

EQooNHTLooWkIofe

26.79.
Les fonctions f_z sont définies pour $z \in \mathbb{C}\backslash K$, et elles sont définies sur K. Donc il n'y a pas de danger que $\xi = z$ dans le dénominateur de (26.307). De plus si f_z est définie (c'est à dire si $z \in \mathbb{C}\backslash K$), alors f_{z+h} est également définie tant que $h \in \mathbb{C}$ n'est pas trop grand parce que $\mathbb{C}\backslash K$ est un ouvert de $\mathbb{C}$.

ThoJVNTzn

Théorème 26.80.
Si μ est une mesure de Radon[42] sur K alors $\hat{\mu}$ est infiniment $\mathbb{C}$-dérivable[43] sur $\Omega = \mathbb{C}\backslash K$ et nous avons

$$\hat{\mu}^{(n)}(z) = -\frac{n!}{\pi}\mu\left(\frac{1}{(\xi - z)^{n+1}}\right). \tag{26.308}$$

42. Définition 26.78.
43. Définition 12.304.

Démonstration. Calculons la dérivée complexe de $\hat{\mu}$. Si la limite existe, nous avons

$$\hat{\mu}'(z) = \lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{\hat{\mu}(z+h)-\hat{\mu}(z)}{h}. \tag{26.309}$$

Nous allons nous concentrer sur ce qui est dans la limite pour montrer qu'elle existe, et pour calculer la valeur. Nous avons

$$\frac{\hat{\mu}(z+h)-\hat{\mu}(z)}{h} = -\frac{1}{\pi}\frac{\mu(f_{z+h})-\mu(f_z)}{h} \tag{26.310a}$$

$$= -\frac{1}{\pi}\mu\left(\frac{f_{z+h}-f_z}{h}\right) \qquad \mu \text{ est linéaire} \tag{26.310b}$$

$$= -\frac{1}{\pi}\mu\left(\frac{1}{(\xi-z-h)(\xi-z)}\right) \tag{26.310c}$$

où la dernière ligne est un abus de notations pour dire que μ est appliquée à la fonction $\xi \mapsto 1/(\xi-z-h)(\xi-z)$.

Histoire d'éviter de regretables abus de notations, nous introduisons la fonction

$$\begin{aligned}\phi_{z,h}\colon K &\to \mathbb{C}\\ \xi &\mapsto \frac{1}{(\xi-z-h)(\xi-z)},\end{aligned} \tag{26.311}$$

ainsi que la famille

$$\begin{aligned}f_{z,n}\colon K &\to \mathbb{C}\\ \xi &\mapsto \frac{1}{(\xi-z)^n}.\end{aligned} \tag{26.312}$$

Ce que nous avons montré est que

$$\frac{\hat{\mu}(z+h)-\hat{\mu}(z)}{h} = -\frac{1}{\pi}\mu(\phi_{z,h}). \tag{26.313}$$

De plus nous avons la convergence

$$\phi_{z,h} \xrightarrow{\left(C(K),\|.\|_\infty\right)} f_{z,2} \tag{26.314}$$

lorsque $h\to 0$. Vu que μ est continue, nous pouvons permuter la limite avec μ et avoir :

$$\hat{\mu}'(z) = \lim_{h\to 0}\mu(\phi_{z,h}) = -\frac{1}{\pi}\mu(f_{z,2}). \tag{26.315}$$

Cela prouve le résultat pour $n=1$.

Pour le reste, vous faites un récurrence [44]. □

Cette théorie permet de fournir une démonstration plus technologique du corolaire 26.75.

Lemme 26.81.
Si f est holomorphe sur Ω et si B est une boule fermée dans Ω alors pour tout $z\in \mathrm{Int}(B)$ nous avons

$$f^{(k)}(z) = \frac{k!}{2i\pi}\int_{\partial B}\frac{f(\xi)}{(\xi-z)^{k+1}}d\xi. \tag{26.316}$$

Démonstration. Appliquer le théorème 26.80 à la mesure de Radon

$$\mu(\phi) = \int_{\partial B}\phi(\xi)d\xi. \tag{26.317}$$

□

44. Je n'ai pas vérifié. Faites-le et écrivez-moi pour me dire si c'est bon.

Tout ce petit monde à propos de la mesure de Radon permet également de redémontrer que

$$\left(\frac{1}{2\pi i}\int_{\partial B}\frac{f(\xi)}{\xi^{n+1}}\right)=f^{(n)}(0)/n!, \tag{26.318}$$

comme nous l'avons déjà fait autour de l'équation (26.300). Nous utilisons le théorème de Radon 26.80 à la mesure

$$\mu(\phi)=\int_{\partial B}\phi(\xi)d\xi. \tag{26.319}$$

La transformée de Cauchy est

$$\hat{\mu}(z)=-\frac{1}{\pi}\mu\left(\frac{1}{\xi-z}\right)=-\frac{1}{\pi}\int_{\partial B}\frac{1}{\xi-z}d\xi, \tag{26.320}$$

EqTzkmeL

et le théorème assure que

$$\hat{\mu}^{(n)}(z)=-\frac{n!}{\pi}\mu\left(\frac{1}{(\xi-z)^{n+1}}\right)=-\frac{n!}{\pi}\int_{\partial B}\frac{1}{(\xi-z)^{n+1}}d\xi. \tag{26.321}$$

En comparant la formule (26.320) avec la formule de Cauchy nous voyons que $\hat{\mu}(z)=-2if(z)$. Par conséquent

$$f^{(n)}(z)=-\frac{1}{2i}\hat{\mu}^{(n)}(z)=\frac{n!}{2\pi i}\int_{\partial B}\frac{1}{(\xi-z)^{n+1}}d\xi, \tag{26.322}$$

et

$$f^{(n)}(0)=\frac{n!}{2\pi i}\int_{\partial B}\frac{1}{\xi^{n+1}}d\xi. \tag{26.323}$$

26.4 Conditions équivalentes à l'holomorphie

Nous nous proposons de lister les conditions que nous avons vues être équivalentes à l'holomorphie.

THOooOGOCooUalFaG

Théorème 26.82.
Soient Ω un ouvert de $\mathbb{C}$ et $f\colon\Omega\to\mathbb{C}$ une fonction continue. Les conditions suivantes sont équivalentes.

ItemOtPcTb

(1) f est holomorphe.

ItemHWRnxx

(2) Pour tout triangle (plein) T contenu dans Ω, $\int_T f=0$.

ItempBBPVv

(3) f est $\mathbb{C}$-dérivable.

ItemmLhzbB

(4) f est C^∞

ItemCCrSrLj

(5) $\partial_{\bar{z}}f=0$; ce sont les équations de Cauchy-Riemann.

ItemEvxRSn

(6) La 1-forme différentielle $f(z)dz$ est localement exacte.

ItemVSCHtY

(7) Pour toute boule $B(a,r)$ contenue dans Ω nous avons

$$f(a)=\frac{1}{2\pi i}\int_{\partial B(a,r)}\frac{f(z)}{z-a}dz. \tag{26.324}$$

La fonction f est holomorphe en z_0 si et seulement si il existe un voisinage $B(z_0,r)$ de z_0 et des nombres a_k tels que sur la boule,

$$f(z)=\sum_{n=0}^{\infty}a_n(z-z_0)^n. \tag{26.325}$$

Dans ce cas, f est holomorphe sur toute la boule.

Démonstration. (1) implique (2) est le lemme de Goursat 26.9. (2) implique (1) est le théorème de Morera 26.77.

(3) est la définition de l'holomorphie, définition 12.305.

(4) implique (1) est un a fortiori sur la définition. (1) implique (4) est contenu dans le théorème de développement en série entière 26.60.

L'équivalence entre (5) et l'holomorphie est le théorème 26.2.

L'équivalence entre (6) et (1) est la proposition 26.65.

L'équivalence entre (1) et (7) est d'une part le théorème 26.60 et d'autre part le corolaire 26.63.

En ce qui concerne la dernière affirmation, si f est holomorphe en z_0, alors le théorème 26.60(1) donne la série. Si au contraire nous avons la série, la proposition 26.11 nous donne le résultat. □

26.5 Singularités, pôles et méromorphe

DEFooKWDUooVPvtpy

Définition 26.83.
Si f est holomorphe[45] *sur un ouvert Ω, alors une **singularité** de f est un point isolé du bord de Ω.*

(1) *La singularité est **effaçable** si la fonction f s'y prolonge en une fonction holomorphe.*

(2) *La singularité a est **isolée** si f est holomorphe sur $B(a,r)\backslash\{a\}$.*

DEFooUIJTooUJPiDG

Définition 26.84 (pôle d'une fonction[1, 642])**.**
*Soient un ouvert Ω de $\mathbb{C}$ ainsi que $a \in \Omega$. La fonction $f\colon \Omega\backslash\{a\} \to \mathbb{C}$ a un **pôle d'ordre** n en a si il existe $r > 0$ et une fonction holomorphe $g\colon B(a,r) \to \mathbb{C}$ telle que*

(1) $g(a) \neq 0$

(2) *pour tout $z \in B(a,r)\backslash\{a\}$ nous avons*

$$f(z) = \frac{g(z)}{(z-a)^n}. \tag{26.326}$$

Lemme 26.85.
Soit $n \in \mathbb{N}$. Nous notons U_n l'ensemble des racines n^e de l'unité[46] *dans $\mathbb{C}$. La fonction*

$$\begin{aligned} f\colon \mathbb{C}\backslash \mathrm{U}_n &\to \mathbb{C} \\ z &\mapsto \frac{1}{z^n - 1} \end{aligned} \tag{26.327}$$

est holomorphe et possède un pôle d'ordre 1 en chaque point de U_n.

Démonstration. Le fait que f soit holomorphe est simplement le fait que sur le domaine, le dénominateur est un bête polynôme qui ne s'annule pas.

Nous énumérons $\mathrm{U}_n = \{\xi_i\}_{i=1,\dots,n}$. Prouvons que ξ_k est un pôle d'ordre 1 de f. La première égalité du lemme 19.18 donne $z^n - 1 = \prod_i (z - \xi_i)$. D'abord nous considérons $r > 0$ tel que $B(\xi_k, r) \cap \mathrm{U}_n = \{\xi_k\}$.

Nous posons

$$\begin{aligned} g_k\colon B(\xi_k,r)\backslash\{\xi_k\} &\to \mathbb{C} \\ z &\mapsto \frac{1}{\prod_{i\neq k}(z-\xi_i)}. \end{aligned} \tag{26.328}$$

Cela est bien une fonction holomorphe et nous avons

$$f(z) = \frac{g_k(z)}{(z-\xi_k)}. \tag{26.329}$$

□

45. Définition 12.305.
46. Voir la définition 19.1 et le lemme 19.2.

¡¡ Avertissement/question au lecteur !! 26.86
Je ne suis pas certain que la proposition suivante soit vraie. Écrivez-moi si vous avez un avis.

PROPooKXXSooGzbmeo

Proposition 26.87.
La fonction

$$z \mapsto \begin{cases} \frac{\sin(z)}{z} & si\ z \neq 0 \\ 1 & si\ z = 0 \end{cases} \tag{26.330}$$

est continue sur un voisinage de zéro dans $\mathbb{C}$.

PROPooQJDCooYwmkRS

Proposition 26.88.
Une singularité de f est un pôle si et seulement si

$$\lim_{z \to Z} f(z) = \infty. \tag{26.331}$$

Le théorème suivant complète la proposition 26.69.

ThoTLQOEwW

Théorème 26.89 (Prolongement de Riemann[643])**.**
Soient un ouvert $\Omega \subset \mathbb{C}$, un point $a \in \Omega$ et une fonction holomorphe $f\colon \Omega\backslash\{a\} \to \mathbb{C}$. Nous supposons que a est une singularité[47] *de f. Les points suivants sont équivalents.* ITEMooMLXJooMfuifN

(1) la singularité a est effaçable[48] *;* ITEMooBWPEooEltHAa

(2) f possède un prolongement continu en a ; ITEMooEAUOooIWcxHS

(3) il existe un voisinage épointé de a sur lequel f est bornée ; ITEMooETRWooDTTpxs

(4) $\lim_{z\to a}(z-a)f(z) = 0$.

Démonstration. En plusieurs implications.

(1) **(1) implique (2)** La fonction f admet même un prolongement holomorphe.

(2) **(2) implique (3)** Soit un prolongement continu $\tilde{f}\colon B(a,r) \to \mathbb{C}$ de f. La restriction $\tilde{f}\colon \overline{B(a,r/2)} \to \mathbb{C}$ est continue sur un compact et donc bornée[49] tout en étant égale à f sur $B(a,r/2)\backslash\{a\}$.

(3) **(3) implique (4)** Nous supposons que $f\colon B(a,r)\backslash\{a\} \to \mathbb{C}$ est bornée. Disons $|f(z)| < A$. Alors pour tout $z \in B(a,r)\backslash\{a\}$ nous avons

$$|(z-a)f(z)| \leqslant A|z-a| \tag{26.332}$$

Or $\lim_{z\to a} A(z-a)$ existe et vaut zéro. Donc $\lim_{z\to a}|(z-a)f(z)|$ existe et vaut également zéro.

(4) **(4) implique (1) si $a = 0$**
Nous commençons par supposer que $a = 0$, et nous posons $D = B(0,r)\backslash\{0\}$. La fonction $f\colon D \to \mathbb{C}$ est holomorphe et vérifie $\lim_{z\to 0} zf(z) = 0$.
Nous considérons la fonction suivante :

$$\begin{aligned} g\colon B(0,r) &\to \mathbb{C} \\ z &\mapsto \begin{cases} 0 & \text{si } z = 0 \\ z^2 f(z) & \text{sinon.} \end{cases} \end{aligned} \tag{26.333}$$

La fonction g est holomorphe sur D parce que f l'est. Voyons que g est dérivable en zéro. Pour tout z sur un voisinage,

$$\frac{g(z)-g(0)}{z} = \frac{z^2 f(z)}{z} = zf(z). \tag{26.334}$$

47. Singularité et singularité effaçable : définition 26.83.
48. Définition 26.83.
49. Théorème 7.146.

Or par hypothèse $\lim_{z\to 0} zf(z) = 0$ donc $g'(0) = 0$, et g est holomorphe en 0 (c'est la définition 12.305 d'une fonction holomorphe). Bref, g est holomorphe sur $B(0,r)$.
Nous pouvons donc développer g en série entière[50] dans un voisinage $B(0,r)$:

$$g(z) = \sum_{n=0}^{\infty} a_n z^n. \tag{26.335}$$

En utilisant le 26.62, $a_0 = g(0) = 0$ et $a_1 = g'(0) = 0$. Donc en réalité

$$g(z) = \sum_{n=2}^{\infty} a_n z^n. \tag{26.336}$$

Considérons la série entière

$$\sum_{n=0}^{\infty} b_n z^n \tag{26.337}$$

avec $b_n = a_{n+2}$. Le lemme 15.17 dit que son rayon de convergence est le même que celui de g. Donc la fonction

$$\begin{aligned} h\colon B(0,r) &\to \mathbb{C} \\ z &\mapsto \sum_{n=2}^{\infty} a_n z^{n-2} \end{aligned} \tag{26.338}$$

est holomorphe.
Par ailleurs, sur la partie D (qui ne contient pas $z = 0$) nous pouvons écrire

$$f(z) = \frac{g(z)}{z^2} \tag{26.339}$$

et donc

$$f(z) = \sum_{n=2}^{\infty} a_n z^{n-2}. \tag{26.340}$$

Autrement dit $f = h$ sur D, et h en est un prolongement holomorphe.

(5) **(3) implique (4)**
Nous posons $g(z) = f(z+a)$. Nous avons

$$\lim_{z\to 0} zg(z) = \lim_{z\to 0} zf(z+a) = \lim_{z\to a}(z-a)f(z) = 0. \tag{26.341}$$

Le changement de variable dans la limite est le lemme 7.177. Donc le premier cas s'applique à g et nous avons un prolongement holomorphe $\tilde{g}\colon B(0,r) \to \mathbb{C}$ de g. La fonction donnée par $\tilde{f}(z) = \tilde{g}(z-a)$ prolonge f.

□

Définition 26.90 (Fonction méromorphe[644])**.**
Soient $\mathcal{U}$ un ouvert de $\mathbb{C}$ et $\{p_i\}$ une suite de points dans $\mathcal{U}$ sans points d'accumulation (éventuellement il y a un nombre fini de p_i). Si la fonction f est holomorphe sur $\mathcal{U}\backslash\{p_i\}$ et si chaque p_i est un point régulier ou un pôle de f, alors nous disons que f est ***méromorphe*** *sur $\mathcal{U}$.*

LEMooCSAFooTYasYM

Lemme 26.91.
Soient $0 < a < b < \infty$. Nous considérons l'équation différentielle

$$\begin{cases} y'(t) = \frac{n}{t} y(t) & \text{(26.342a)} \\ y(a) = y_a & \text{(26.342b)} \end{cases}$$

pour la fonction $y\colon]a,b[\to \mathbb{R}$.
L'unique solution est

$$y(t) = \frac{y_a}{a^n} t^n. \tag{26.343}$$

EQooKPYIooMHPIBP

Note : $a \neq 0$ de toutes façons, donc pas de problèmes.

50. Théorème 26.60.

Démonstration. En termes du théorème de Cauchy-Lipschitz 17.39, nous avons $y'(t) = f\big(t, y(t)\big)$ avec

$$\begin{aligned} f\colon]a,b[\times \mathbb{R} &\to \mathbb{R} \\ (t,y) &\mapsto \frac{n}{t}y. \end{aligned} \tag{26.344}$$

Cette fonction f est Lipschitz et tout ce qu'on veut parce que $t = 0$ est hors de son domaine ; la régularité de f peut être étudiée sur le compact $[a/2, b]$.

Un calcul direct vérifie que la solution proposée (26.343) est bien une solution. Le théorème de Cauchy-Lipschitz dit qu'elle est unique. □

26.92.
Vous voulez savoir comment on trouve la solution $y(t) = Kt^n$? Allez, on vous la fait un peu détendue, sans trop regarder les détails. D'abord,

$$\frac{y'(t)}{y(t)} = \frac{n}{t}. \tag{26.345}$$

Nous intégrons des deux côtés :

$$\int_a^x \frac{y'(t)}{y(t)}dt = n[\ln(t)]_a^x. \tag{26.346}$$ EQooWCQRooKqjfVj

À gauche nous posons $u(t) = \ln\big(y(t)\big)$ et nous avons

$$\int_a^x \frac{y'(t)}{y(t)}dt = \int_a^x u'(t)dt = [u(t)]_a^x = \ln\big(y(x)\big) - \ln\big(y(a)\big). \tag{26.347}$$

Nous mettons ça à la place du membre de gauche de (26.346), nous mettons toutes les constantes dans un L (en ne nous posant aucune question sur le fait que ce soit positif, nul, que ça peut rentrer dans un logarithme ou non) et :

$$\ln\big(y(x)\big) = n\ln(x) + L = \ln\big(Kx^n\big). \tag{26.348}$$

où $K = e^L$. Et voila.

PropPUZTQKl

Proposition 26.93.
Soient Ω un ouvert de $\mathbb{C}$ et $f_n\colon \Omega \to \mathbb{C}$ une suite de fonctions telles que pour tout compact K de Ω il existe $N_K \geqslant 0$ tel que

(1) f_n n'a pas de pôle dans K dès que $n \geqslant N_K$;

(2) la série $\sum_{n\geqslant N_K} f_n$ converge uniformément sur K.

Alors

(1) La fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \tag{26.349}$$

est méromorphe sur Ω et ses pôles sont l'union de ceux des f_n.

(2) Nous pouvons permuter la somme et la dérivée :

$$f'(z) = \sum_{n=0}^{\infty} f_n'(z). \tag{26.350}$$

26.6 Dénombrement des solutions d'une équation diophantienne

Le théorème 26.99 peut être vu soit comme un dénombrement de solutions d'une certaine équation diophantienne, soit comme partition d'un entier en parts fixées. Avant de nous lancer dans sa démonstration, nous prouvons un certain nombre de lemmes qui vont traiter des aspects combinatoires de la preuve.

Soit $n, N \in \mathbb{N}$. Soit $a \in \mathbb{N}^N$. Nous considérons les ensembles suivants :

$$V_n(N) = \{x \in \mathbb{N}^N \text{ tel que } \sum_{i=1}^{N} x_i = n\}, \tag{26.351a}$$

$$W_n(a, N) = \{y \in \mathbb{N}^N \text{ tel que } y \cdot a = n\}, \tag{26.351b}$$

$$V_n(N)_a = \{x \in V_n(N) \text{ tel que } a_i \mid x_i \, \forall i = 1, \dots, N\}. \tag{26.351c}$$

L'ensemble $V_n(N)$ avait déjà été rencontré en (15.98).

LEMooLKCAooCeDnSj

Lemme 26.94.
L'application

$$\begin{aligned} \psi \colon W_n(a, N) &\to V_n(N)_a \\ (y_1, \dots, y_N) &\mapsto (y_1 a_1, \dots, y_N a_N) \end{aligned} \tag{26.352}$$

(1) est bien définie, c'est-à-dire qu'elle prend effectivement ses valeurs dans $V_n(N)_a$,

(2) est une bijection.

Démonstration. En plusieurs parties.

(1) **Bien définie** Soit $y \in W_n(a, N)$. Nous avons $\psi(y)_i = a_i y_i$. Le nombre a_i divise donc bien $\psi(y)_i$ et $\sum_{i=1}^{N} \psi(y)_i = n$.

(2) **Injective** Si $\psi(y) = \psi(y')$, alors pour tout i nous avons $y_i a_i = y'_i a_i$, et donc $y_i = y'_i$. La fonction ψ est donc bien injective.

(3) **Surjective** Soit $x \in V_n(N)_a$. Vu que $a_i \mid x_i$, il existe $y_i \in \mathbb{N}$ tel que $x_i = a_i y_i$. On vérifie que $y \in W_n(a, N)$ et que $\psi(y) = x$.

□

LEMooOPXHooHzoHrm

Lemme 26.95.
Pour $i = 1, \dots, N$, nous posons

$$\begin{aligned} b_i \colon \mathbb{N}^N &\to \{0, 1\} \\ x &\mapsto \begin{cases} 1 & \text{si } a_i \mid x_i \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{26.353}$$

Nous avons

$$\sum_{x \in V_n(N)} \prod_{i=1}^{N} b_i(x) = \operatorname{Card}\big(W_n(a, N)\big). \tag{26.354}$$

Démonstration. Nous décomposons la somme en $V_n(N)_a$ et son complémentaire dans $V_n(N)$:

$$\sum_{x \in V_n(N)} \prod_{i=1}^{N} b_i(x) = \sum_{x \in V_n(N)_a} \prod_{i=1}^{N} b_i(x) + \sum_{x \in V_n(N) \setminus V_n(N)_a} \prod_{i=1}^{N} b_i(x) \tag{26.355a}$$

$$= \sum_{x \in V_n(N)_a} 1 + \sum_{x \in V_n(N) \setminus V_n(N)_a} 0 \tag{26.355b}$$

SUBEQooLMBGooIfxjgy

$$= \operatorname{Card}\big(V_n(N)_a\big) \tag{26.355c}$$

$$= \operatorname{Card}\big(W_n(N, a)\big) \tag{26.355d}$$

SUBEQooBBOIooBHDgYF

Justifications :

— Pour (26.355b). Si $x \in V_n(N)_a$, alors $a_i \mid x_i$ pour tout i, et donc $b_i(x) = 1$ pour tout i. Si au contraire $x \in V_n(N)\backslash V_n(N)_a$, il existe un i tel que a_i ne divise pas x_i et donc tel que $b_i(x) = 0$.

— Pour (26.355d). Les deux ensembles sont en bijection par le lemme 26.94.

□

LEMooRJOKooPJGVTr

Lemme 26.96.
Soit $s \geqslant 1$. La série entière

$$\sum_{n=0}^{\infty} z^{ns} \tag{26.356}$$

a un rayon de convergence égal à 1. Pour $z \in B(0,1)$ nous avons

$$\sum_{n=0}^{\infty} z^{ns} = \frac{1}{1-z^s}. \tag{26.357}$$

EQooIRAZooKQoZnp

Démonstration. Pour un $z \in \mathbb{C}$ fixé, nous avons $z^{ns} = (z^s)^n$ [51]. Nous appelons donc la proposition 10.135 avec $q = z^s$.

Si $|z| < 1$, alors $|z^s| < 1$ et la proposition 10.135 nous dit que la série converge. Si au contraire $|z| > 1$, alors $|z^s| > 1$ et la série diverge.

Le corolaire 15.19 conclut que le rayon de convergence est bien 1.

La valeur (26.357) est également une partie de la proposition 10.135. □

LEMooVMLEooCzPuKy

Lemme 26.97.
Si $a, b \in \mathbb{C}$, si $N \in \mathbb{N}$ et si $p \in \mathbb{N}$ avec $p < N$, nous posons

$$a_n = a\frac{(n+N-1)!}{n!} \tag{26.358}$$

et

$$b_n = b\frac{(n+p-1)!}{n!}. \tag{26.359}$$

Nous avons $a_n \sim a_n + b_n$.

Démonstration. Nous posons $\alpha(n) = (a_n + b_n)/a_n$ et nous prouvons que $\lim_{n\to\infty} \alpha(n) = 1$. Pour ce faire,

$$\frac{b_n}{a_n} = \frac{b}{a}\frac{(n+p-1)!}{n!}\frac{n!}{(n+N-1)!} \tag{26.360a}$$

$$= \frac{b}{a}\frac{(n+p-1)!}{(n+N-1)!} \tag{26.360b}$$

$$= \frac{b}{a}\frac{(n+p-1)!}{(n+p-1)!\prod_{k=n+p}^{n+N-1} k} \tag{26.360c}$$

$$= \frac{b}{a}\prod_{k=n+p}^{n+N-1}\frac{1}{k} \tag{26.360d}$$

$$\leqslant \frac{b}{a}\frac{1}{n+p}, \tag{26.360e}$$

et nous avons

$$\lim_{n\to\infty}\frac{b_n}{a_n} = 0, \tag{26.361}$$

de telle sorte que $\lim_{n\to\infty} \alpha(n) = 1$. □

51. Vu qu'ici n et s sont entiers, c'est pas profond ça. Il ne faut pas invoquer la proposition générale 12.408.

LEMooTGHHooZHZsgE

Lemme 26.98.
Si $N \in \mathbb{N}$, nous avons équivalence des suites

$$\frac{(n+N-1)!}{n!} \sim n^{N-1}. \tag{26.362}$$

Démonstration. Sachez que dans $\prod_{i=a}^{b}$, il y a $b-a+1$ facteurs, et non $b-a$ comme on pourrait naïvement le croire. Cela dit, nous avons le calcul

$$\frac{(n+N-1)!}{n!n^{N-1}} = \left(\prod_{k=n+1}^{n+N-1} k \right) \frac{1}{n^{N-1}} \tag{26.363a}$$

$$= \prod_{k=n+1}^{n+N-1} \frac{k}{n} \tag{26.363b}$$

$$= \prod_{k=1}^{N-1} \frac{n+k}{n}. \tag{26.363c}$$

Donc la limite

$$\lim_{n\to\infty} \frac{(n+N-1)!}{n!n^{N-1}} = 1. \tag{26.364}$$

□

THOooQDYWooCOiUMb

Théorème 26.99 ([1, 114, 430, 645, 104])**.**
Soient $N \in \mathbb{N}$ et $a \in \mathbb{N}^N$ tel que $\operatorname{pgcd}(a_1, \dots, a_N) = 1$. *Nous posons*

$$W_n(a, N) = \{y \in \mathbb{N}^N \text{ tel que } y \cdot a = n\}. \tag{26.365}$$

Nous avons alors[52]

$$\operatorname{Card}\big(W_n(a,N)\big) \sim \frac{1}{\prod_{k=1}^{N} a_k} \frac{n^{N-1}}{(N-1)!}. \tag{26.366}$$

Démonstration. Pour chaque $i = 1, \dots, N$, nous considérons la série entière

$$s_i(z) = \sum_{k=0}^{\infty} z^{ka_i}. \tag{26.367}$$

dont le rayon de convergence vaut 1 par le lemme 26.96. Nous nous apprêtons à faire le produit de Cauchy multiple de la proposition 15.36 ; nous posons donc

$$b_{ik} = \begin{cases} 1 & \text{si } a_i | k \\ 0 & \text{sinon ,} \end{cases} \tag{26.368}$$

et nous écrivons toutes les séries s_i sous la forme

$$s_i(z) = \sum_{k=0}^{\infty} b_{ik} z^k. \tag{26.369}$$

La proposition 15.36 nous assure que si $|z| < 1$, le produit $\prod_{i=1}^{N} s_i(z)$ peut être écrit sous la forme de la série entière

SUBEQooYESHooChEKGm

$$\prod_{i=1}^{N} s_i(z) = \sum_{s=0}^{\infty} \left(\sum_{x \in V_s(N)} \prod_{i=1}^{N} b_{ix_i} \right) z^s \tag{26.370a}$$

$$= \sum_{s=0}^{\infty} \left(\sum_{x \in V_s(N)} \prod_{i=1}^{N} b_i(x) \right) z^s \tag{26.370b}$$

SUBEQooEEHKooPSrpiT

$$= \sum_{s=0}^{\infty} \operatorname{Card}\big(W_s(a,N)\big) z^s \tag{26.370c}$$

SUBEQooULBCooSTuHvy

52. Équivalence de suites, définition 9.32.

Justifications :

— Pour (26.370b). Notation b_i du lemme 26.95.

— Pour (26.370c). Utilisation du lemme 26.95

D'autre part, le lemme 26.96 nous permet d'écrire

$$f(z)=\prod_{i=1}^{N}s_i(z)=\prod_{i=1}^{N}\frac{1}{1-z^{a_i}}. \tag{26.371}$$

Notre but sera d'écrire ce produit sous forme de série entière est d'identifier les coefficients avec ceux que l'on trouve dans (26.370c).

(1) $m_\omega=N$ **si et seulement si** $\omega^{a_i}=1$ Nous montrons à présent que ω est un pôle d'ordre N de f si et seulement si $\omega^{a_1}=\ldots=\omega^{a_N}=1$.

(1a) **Un polynôme** Nous considérons le polynôme

$$P(X)=\prod_{i=1}^{N}(1-X^{a_i})=(-1)^N\prod_{i=1}^{N}\prod_{\omega\in\mathrm{U}_{a_i}}(X-\omega) \tag{26.372}$$

où, pour la seconde égalité, nous avons utilisé le lemme 19.18. En posant $\mathrm{U}=\bigcup_{i=1}^{N}\mathrm{U}_{a_i}$ nous écrivons encore

$$P(X)=(-1)^N\prod_{\omega\in\mathrm{U}}(X-\omega)^{m_\omega} \tag{26.373}$$

où $m_\omega=\mathrm{Card}\{i \text{ tel que } \omega\in\mathrm{U}_{a_i}\}$. Cela pour dire que, pour $|z|<1$,

$$f(z)=\frac{1}{P(z)}=\frac{(-1)^N}{\prod_{\omega\in\mathrm{U}}(z-\omega)^{m_\omega}}. \tag{26.374}$$

(1b) **Sens $\Rightarrow$** Si ω est un pôle d'ordre N, alors $N=m_\omega=\{i \text{ tel que } \omega\in\mathrm{U}_{a_i}\}$. Donc $\omega\in\mathrm{U}_{a_i}$ pour tout i, c'est-à-dire que $\omega^{a_i}=1$ pour tout i.

(1c) **Sens $\Leftarrow$** Dans l'autre sens, si $\omega^{a_i}=1$ pour tout i, alors $m_\omega=N$ et ω est un pôle d'ordre N.

Nous pouvons continuer.

(2) $m_\omega=N$ **si et seulement si** $\omega=1$ Nous savons que $\omega=1$ est un pôle d'ordre N parce que $1\in\mathrm{U}_{a_i}$ pour tout i. Dans l'autre sens, si ω est d'ordre N, alors nous venons de voir que $\omega^{a_i}=1$ pour tout i.

Vu que les a_i sont premiers entre eux, le théorème de Bézout 1.237 nous donne des entiers u_i tels que $\sum_{i=1}^{N}u_ia_i=1$. Nous avons alors

$$\omega=\omega^{u\cdot a}=\prod_{i=1}^{N}(\omega^{a_i})^{u_i}=1. \tag{26.375}$$

Donc les pôles de f différents de 1 sont d'ordre strictement inférieur à N.

(3) **Décomposition en éléments simples** Décomposons un peu l'expression de $f(z)$ en repartant de (26.372) :

$$f(z)=\prod_{i=1}^{N}\frac{1}{1-z^{a_i}} \tag{26.376a}$$

$$=\prod_{i=1}^{N}\frac{1}{-\prod_{\omega\in\mathrm{U}_{a_i}}(z-\omega)} \tag{26.376b}$$

$$=(-1)^N\prod_{i=1}^{N}\prod_{\omega\in\mathrm{U}_{a_i}}\frac{1}{z-\omega} \tag{26.376c}$$

$$=(-1)^N\prod_{i=1}^{N}\sum_{\omega\in\mathrm{U}_{a_i}}\frac{\lambda_{\omega,i}}{z-\omega} \tag{26.376d}$$

Justifications :

— Pour (26.376b). Lemme 19.18 ; vous noterez le signe de différence.

— Pour (26.376d). Lemme 19.13 pour la décomposition en éléments simples.

(4) **Isoler le terme $\omega = 1$** Nous notons $\mathrm{U} = \bigcup_{i=1}^{N} \mathrm{U}_{a_i}$. Chaque U_{a_i} contient $\omega = 1$. L'expression (26.376d) contient donc un terme en $\frac{1}{(z-1)^N}$. Tous les autres ω de U ne sont présents que dans au maximum $N-1$ des U_{a_i}. Nous avons donc

$$f(z) = \frac{A}{(z-1)^N} + \sum_{p=1}^{N-1} \sum_{\omega \in \mathrm{U}} \frac{B_{p,\omega}}{(z-\omega)^p} \tag{26.377}$$

avec $B_{p,\omega} \in \mathbb{C}$.

(5) **Une belle lampée de factorielles** Le lemme 15.47(3) permet d'écrire f avec des séries entières :

$$\begin{aligned} f(z) = {} & \frac{A}{(N-1)!} \sum_{s=0}^{\infty} \frac{(s+N-1)!}{s!} z^s \\ & + \sum_{p=1}^{N-1} \sum_{\omega \in \mathrm{U}} \frac{(-1)^p B_{p,\omega}}{(p-1)!} \sum_{s=0}^{\infty} \frac{(s+p-1)!}{s!} \frac{z^s}{\omega^{s+p+1}} \end{aligned} \tag{26.378}$$

qui est valable pour $z \in B(0,1)$.

(6) **Ce qu'on en fait** Pour rappel, l'équation (26.370) nous dit que

$$f(z) = \sum_{s=0}^{\infty} \mathrm{Card}\big(W_s(a,N)\big) z^s. \tag{26.379}$$

Nous allons donc identifier le coefficient de z^n dans (26.378) avec $\mathrm{Card}\big(W_n(a,N)\big)$:

$$\begin{aligned} \mathrm{Card}\big(W_n(a,N)\big) = {} & \frac{A}{(N-1)!} \frac{(n+N-1)!}{n!} \\ & + \sum_{p=1}^{N-1} \sum_{\omega \in \mathrm{U}} \frac{(-1)^p B_{p,\omega}}{(p-1)!} \frac{(n+p-1)!}{n!} \frac{1}{\omega^{n+p+1}}. \end{aligned} \tag{26.380}$$

Voici une belle suite (par rapport à n) dont nous devons étudier le comportement asymptotique.

(7) **Des équivalences** Le lemme 26.97 nous permet de supprimer tous les termes autres que celui qui contient A :

$$\mathrm{Card}\big(W_n(a,N)\big) \sim \frac{A}{(N-1)!} \frac{(n+N-1)!}{n!}. \tag{26.381}$$

Notez que, à gauche, nous avons une suite dans $\mathbb{N}$ et à droite, une suite dans $\mathbb{C}$ (il n'y a pas de raisons à priori que A soit entier ou réel). Cela n'a pas d'importance ; ça n'empêche pas les suites d'être équivalentes.

Le lemme 26.98 donne maintenant

$$\mathrm{Card}\big(W_n(a,N)\big) \sim \frac{A}{(N-1)!} n^{N-1}. \tag{26.382}$$

Cela est déjà très bien parce que ça donne la vitesse de croissance en fonction de N et n. Mais puisque nous sommes perfectionistes, nous allons encore déterminer la valeur de A.

(8) **La valeur de A** Pour déterminer la valeur de A, l'astuce est de considérer la fonction $z \mapsto f(z)(1-z)^N$:

$$f(z)(1-z)^N = (1-z)^N \prod_{i=1}^{N} \frac{1}{1-z^{a_i}} \tag{26.383a}$$

$$= \prod_{i=1}^{N} \frac{1-z}{1-z^{a_i}} \tag{26.383b}$$

$$= \prod_{i=1}^{N} \frac{1}{1+\ldots+z^{a_i-1}}. \tag{26.383c}$$

SUBEQooUZTCooYgHaES

Justifications :

— Pour (26.383c). C'est le lemme 3.136(2).

La dernière expression montre qu'il n'y a pas de mal à prendre la limite $z \to 1$; elle vaut

$$\lim_{z\to 1} f(z)(1-z)^N = \prod_{i=1}^{N} \frac{1}{a_i}. \tag{26.384}$$

Mais en partant d'autre part de (26.377), nous avons

$$f(z)(1-z)^N = A + \sum_{p=1}^{N-1} \sum_{\omega\in\mathrm{U}} \frac{B_{p,\omega}(1-z)^N}{(z-\omega)^p}. \tag{26.385}$$

Vu que $N > p$, la limite $z \to 0$ existe et vaut zéro dans tous les éléments de la somme, y compris les éléments avec $\omega = 1$. Donc

$$\lim_{z\to 1} f(z)(1-z)^N = A. \tag{26.386}$$

Nous savons donc que

$$A = \prod_{i=1}^{N} \frac{1}{a_i}. \tag{26.387}$$

EQooJMALooUrXJZc

En remettant la valeur (26.387) dans l'équivalence (26.382), nous trouvons le résultat demandé. □

Exemple 26.100.
Pour $p = 1$, l'équation est $\alpha x = n$, qui possède au maximum une solution, quel que soit n. Et de plus pour avoir une solution il faut et suffit que α divise n, c'est-à-dire que n soit un multiple de α. Il n'y a que un nombre sur α à être multiple de α. D'où le comportement en $\frac{1}{\alpha}$.

Pour $p = 2$, c'est l'équation (3.58) déjà étudiée. Il y a une famille à un paramètre de solutions dont seulement un certain nombre sont positives. À priori, le nombre de solutions positives croît linéairement en n. △

26.101.
Si vous aimez les séries génératrices. Si vous aimez l'idée de mettre toute l'information d'un problème dans les coefficients d'une série puis de trouver des réponses en les manipulant, vous pouvez regarder *introduction à la théorie analytique des nombres*[646].

Cette vidéo explique comment payer n euros avec des pièces et des billets de valeur données. On pourrait croire que cela est exactement le résultat du théorème 26.99. Il n'en est rien parce que l'hypothèse de pgcd du théorème n'est pas du tout réalisée par les pièces et billets actuellement en circulation.

Du coup, je ne sais pas si ce théorème est intéressant au sens de la définition 0.2.

26.7 Fonctions d'Euler

ThoZJYooWKfbVz

Théorème 26.102 (Prolongement méromorphe de la fonction Γ d'Euler[104])**.**
Nous considérons la formule

$$\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt. \tag{26.388}$$

Alors

(1) Cette formule définit une fonction holomorphe sur

$$\mathcal{P} = \{z \in \mathbb{C} \text{ tel que } \Re(z) > 0\}. \tag{26.389}$$

(2) La fonction $\Gamma : \mathcal{P} \to \mathbb{C}$ admet un unique prolongement méromorphe sur $\mathbb{C}$, lequel a des pôles sur les entiers négatifs.

Démonstration. (1) **Holomorphie sous l'intégrale**

Pour étudier l'holomorphie de la fonction Γ sur $\mathcal{P}$ nous utilisons le théorème 26.73.

Nous considérons la fonction

$$\begin{aligned} g : \mathcal{P} \times \mathbb{R}^+ &\to \mathbb{C} \\ (z,t) &\mapsto e^{-t}t^{z-1} \end{aligned} \tag{26.390}$$

et nous commençons par montrer que c'est holomorphe en z pour chaque $t > 0$ fixé. Nous le vérifions par le critère de $\partial_{\bar{z}} f = 0$ [53] et en nous souvenant que $t^i = e^{\ln(t^i)} = e^{i\ln(t)}$. Nous obtenons rapidement que

$$\frac{\partial g}{\partial \bar{z}} = 0. \tag{26.391}$$

Le fait que la fonction $t \mapsto g(z,t)$ soit mesurable pour tout z est d'accord.

Et enfin soit K compact dans $\mathcal{P}$. Il faut trouver une fonction $g_K(t)$ intégrable sur $[0,\infty[$ telle que pour tout $z \in K$ et $t \in [0,\infty[$ nous ayons $|g(z,t) \leqslant g_K(t)|$. Pour cela nous majorons séparément les parties $t \in]0,1[$ et $t \geqslant 1$.

Soit donc K compact dans $\mathcal{P}$; nous posons $M = \max_{z\in K} \Re(z)$ et $\epsilon = \min_{z\in K} \Re(z)$.

Si $t \in]0,1[$ alors nous avons

$$e^{-t}t^{z-1} = e^{-t}e^{(z-1)\ln(t)}, \tag{26.392}$$

de telle façon à que que

$$\begin{aligned} |e^{-t}t^{z-1}| &\leqslant |e^{(x-1+iy)\ln(t)}| && \text{(26.393a)} \\ &= |e^{(\Re(z)-1)\ln(t)}| && \text{(26.393b)} \\ &= |t^{\Re(z)-1}| && \text{(26.393c)} \\ &\leqslant |t^{\epsilon-1}| && \text{(26.393d)} \\ &= \frac{1}{t^{1-\epsilon}}. && \text{(26.393e)} \end{aligned}$$

Cette dernière fonction est intégrable sur $]0,1[$.

Nous considérons maintenant $t \geqslant 1$. Dans ce cas nous avons

$$|e^{-t}t^{z-1}| = e^{-t}t^{\Re(z)-1} \leqslant e^{-t}t^{M-1}. \tag{26.394}$$

Cette dernière fonction est un produit d'une exponentielle décroissante avec un polynôme. C'est donc intégrable entre 1 et l'infini.

La fonction g_K que nous considérons est donc

$$g_K(t) = \begin{cases} \frac{1}{t^{1-\epsilon}} & \text{si } t < 1 \\ \text{borné} & \text{si } 1 \leqslant t \leqslant b \\ e^{-t}t^{M-1} & \text{si } t > b. \end{cases} \tag{26.395}$$

53. Théorème 26.2.

Cela est une fonction intégrable sur $]0,\infty[$ et qui majore g uniformément en z sur le compact K de $\mathcal{P}$. Le théorème 26.73 nous permet donc de conclure que

$$\Gamma(z) = \int_0^\infty g(z,t)dt \tag{26.396}$$

est holomorphe en z sur $\mathcal{P}$ et que

$$\Gamma'(z) = \int_0^\infty \frac{\partial g}{\partial z}(z,t)dt. \tag{26.397}$$

(2) **En deux morceaux** Nous passons maintenant à la seconde partie du théorème. Pour $z \in \mathcal{P}$ nous coupons l'intégrale en deux :

$$\Gamma(z) = \int_0^1 e^{-t}t^{z-1}dt + \int_1^\infty e^{-t}t^{z-1}dt \tag{26.398}$$

(3) **Première partie** Nous commençons par parler de la première partie : $\int_0^1 e^{-t}t^{z-1}dt$ dans laquelle nous voulons utiliser le développement en série de l'exponentielle e^{-t}. Nous devons donc traiter

$$\int_0^1 \sum_{n=0}^\infty \frac{(-1)^n}{n!}t^{n+z-1}dt. \tag{26.399}$$

Nous allons permuter la somme avec l'intégrale à l'aide du théorème de Fubini 14.304 en posant la fonction

$$g(n,t) = \frac{(-1)^n}{n!}t^{n+z-1} \tag{26.400}$$

et en considérant le produit entre la mesure de Lebesgue sur $\mathbb{C}$ et la mesure de comptage sur $\mathbb{N}$, c'est-à-dire que nous étudions

$$\int_0^1 \int_{\mathbb{N}} g(n,t)dndt. \tag{26.401}$$

Pour permuter il suffit de prouver que $|g|$ est intégrable pour la mesure produit, c'est-à-dire que

$$\int_0^1 \int_{\mathbb{N}} \left|\frac{(-1)^n}{n!}t^{n+z-1}\right| < \infty. \tag{26.402}$$

Nous avons $|t^z| = t^{\Re(z)}$, donc

$$\sum_{n=0}^\infty \left|\frac{t^{n+z-1}}{n!}\right| = t^{\Re(z)-1}\sum_{n=0}^\infty \frac{t^n}{n!} = t^{\Re(z)-1}e^t. \tag{26.403}$$

Étant donné que nous avons fixé $z \in \mathcal{P}$, nous avons $\Re(z) - 1 > -1$ et donc $t^{\Re(z)-1}$ est intégrable entre 0 et 1 (proposition 14.300). La partie e^t se majore sur $[0,1]$ par une constante quelconque. Nous avons donc payé le droit d'inverser la somme et l'intégrale :

$$\int_0^1 e^{-t}t^{z-1}dt = \sum_{n=0}^\infty \int_0^1 \frac{(-1)^n}{n!}t^{n+z-1}dt = \sum_{n=0}^\infty \frac{(-1)^n}{n!}[t^{n+z}]_0^1 = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+z)}. \tag{26.404}$$

Nous avons donc l'intéressante formule suivante, valable pour tout $z \in \mathcal{P}$:

$$\Gamma(z) = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+z)} + \int_1^\infty e^{-t}t^{z-1}dt. \tag{26.405}$$

(4) **Prolongation de la première partie** Nous voudrions montrer maintenant que la fonction

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n-z)} \tag{26.406}$$

est méromorphe sur $\mathbb{C}$ avec des pôles en les entiers négatifs. Pour cela nous considérons la suite de fonctions

$$f_n(z) = \frac{(-1)^n}{n!(z+n)} \tag{26.407}$$

et nous allons utiliser la proposition 26.93. Si $n \geqslant 0$, la fonction f_n est méromorphe sur $\mathbb{C}$ avec un pôle simple en $z = -n$. Soit K compact de $\mathbb{C}$ et N_K tel que $K \subset \overline{B(0, N_K)}$. Pour $n \geqslant N_K + 1$, la fonction f_n n'a pas de pôle dans K et de plus pour tout $z \in K$ nous avons

$$|z+n| = |z-(-n)| \geqslant \big|n - |z|\big| \geqslant n - |z| \geqslant n - N_K, \tag{26.408}$$

et par conséquent

$$|f_n(z)| \leqslant \frac{1}{n!(n-N)}, \tag{26.409}$$

ou pour le dire de façon plus snob :

$$\|f_n\|_{\infty,K} \leqslant \frac{1}{n!(n-N)}, \tag{26.410}$$

dont la série converge. Cela signifie que la série $\sum_{n\geqslant N} f_n$ converge normalement [54] sur K, donc la fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \tag{26.411}$$

est une fonction méromorphe dont les pôles sont ceux des f_n, c'est-à-dire les entiers négatifs (proposition 26.93).

(5) **La seconde partie**
Nous allons à présent prouver que la fonction

$$h(z) = \int_1^{\infty} e^{-t} t^{z-1} dt \tag{26.412}$$

est holomorphe sur $\mathbb{C}$. Pour cela nous considérons la fonction de deux variables $f(z,t) = e^{-t}t^{z-1}$ et nous utilisons le théorème d'holomorphie sous l'intégrale 26.73. D'abord pour z_0 fixé dans $\mathbb{C}$ nous avons

$$\int_1^{\infty} |e^{-t}t^{z_0-1}| \leqslant \int_1^{\infty} e^{-t}t^{\Re(z_0)-1}dt, \tag{26.413}$$

donc l'intégrale converge parce que c'est polynôme contre exponentielle. Par ailleurs pour chaque t_0 fixé sur $[0,\infty[$, la fonction $z \mapsto e^{-t_0}t_0^{z-1}$ est holomorphe sur $\mathbb{C}$ comme en témoigne le calcul suivant :

$$\frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) t_0^{x+iy-1} = 0. \tag{26.414}$$

Et enfin si K est compact dans $\mathbb{C}$ nous avons

$$|f(z,t)| = |e^{-t}t^{z-1}| = e^{-t}|t^{\Re(z)-1}| \leqslant e^{-t}t^{M-1} \tag{26.415}$$

où $M = \max_{z\in K} \Re(z)$. Nous en déduisons que la fonction

$$z \mapsto \int_1^{\infty} e^{-t}t^{z-1}dt \tag{26.416}$$

est une fonction holomorphe sur $\mathbb{C}$.

54. Définition 10.93.

(6) **Conclusion**
Au final nous avons prouvé que la fonction Γ d'Euler admet le prolongement méromorphe sur $\mathbb{C}$ donné par

$$\Gamma(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(z+n)} + \int_1^{\infty} e^{-t}t^{z-1}dt. \tag{26.417}$$

□

26.7.1 Euler et factorielle

Proposition 26.103.
Nous avons la formule $\Gamma(n) = (n-1)!$ *pour tout* $n \in \mathbb{N}$.

Démonstration. Nous partons de la formule

$$\Gamma(n) = \int_0^{\infty} e^{-t}t^{n-1}dt \tag{26.418}$$

que nous intégrons par partie en posant

$$\begin{aligned} u &= t^{n-1} & u' &= (n-1)t^{n-2} \\ v &= e^{-t} & v' &= -e^{-t}. \end{aligned} \tag{26.419}$$

Les termes au bord s'annulent (ici il y a un passage à la limite qui n'est pas écrit) et nous trouvons

$$\Gamma(n) = \int_0^{\infty} (n-1)e^{-t}t^{n-2}dt = (n-1)\Gamma(n-1). \tag{26.420}$$

Pour conclure il suffit de remarquer que

$$\Gamma(1) = \int_0^{\infty} e^{-t} = -[e^{-t}]_0^{\infty} = 1. \tag{26.421}$$

□

26.8 Exponentielle et logarithme complexe

26.8.1 Propriétés de l'exponentielle

Proposition 26.104.
Soit $z \in \mathbb{C}$ *fixé. La fonction*

$$\begin{aligned} E\colon \mathbb{R} &\to \mathbb{C} \\ t &\mapsto e^{tz} \end{aligned} \tag{26.422}$$

est C^{∞}*, sa dérivée est*

$$E'(t) = ze^{tz}. \tag{26.423}$$

La fonction E *est développable en série entière (voir définition 15.135) sur* $\mathbb{R}$ *en* $t = 0$ *et*

$$e^{tz} = \sum_{n=0}^{\infty} \frac{z^n}{n!}t^n. \tag{26.424}$$

Démonstration. Nous fixons $z \in \mathbb{C}$. Par définition 18.6, la série suivante est e^{tz} :

$$f(t) = \sum_{n=0}^{\infty} \frac{z^n}{n!}t^n. \tag{26.425}$$

Cette série a un rayon de convergence infini et la fonction f est donc C^{∞} sur $\mathbb{R}$. Nous pouvons la dériver terme à terme :

$$f'(t) = \sum_{n=1}^{\infty} \frac{z^n}{n!}nt^{n-1} = z\sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!}t^{n-1} = ze^{tz}. \tag{26.426}$$

□

THOooNGOIooEECfAv

Théorème 26.105.
La fonction exponentielle vérifie les propriétés suivantes.

(1) exp *est holomorphe*[55].

(2) $(e^z)' = e^z$.

(3) L'exponentielle est développable en série entière,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{26.427}$$

et la série converge normalement sur tout compact de $\mathbb{C}$.

Démonstration. En tant que application $E\colon \mathbb{R}^2 \to \mathbb{C}$, la fonction

$$E(x,y) = e^x(\cos y + i \sin y) \tag{26.428}$$

est C^∞. De plus nous avons

$$\frac{\partial E}{\partial x}(x,y) = e^{x+iy} = E(x,y) \tag{26.429a}$$

$$\frac{\partial E}{\partial y}(x,y) = iE(x,y), \tag{26.429b}$$

et par conséquent la fonction E vérifie les équations de Cauchy-Riemann.

Si r est fixé, par le critère d'Abel appliqué à la suite $r^n/n!$ nous savons que la série $\sum z^n/n!$ converge normalement sur le compact $B(0,r)$. □

26.8.2 Intégrale de Fresnel

Nous allons calculer l'**intégrale de Fresnel**

$$\int_0^\infty e^{-ix^2} dx = \frac{\sqrt{\pi}}{2} e^{-i\pi/4} \tag{26.430}$$

en suivant la démarche présentée par Wikipédia[647]. Nous commençons par prouver que l'intégrale est convergente en nous contentant de justifier la convergence de

$$\int_0^\infty \sin(x^2) dx. \tag{26.431}$$

Pour chaque $a > 0$ fixé, l'intégrale $\int_0^a \sin(x^2)dx$ ne pose pas de problèmes. Le lemme 20.197 nous permet de passer à la limite ; nous devons donc seulement calculer

$$\lim_{b\to\infty} \int_a^b \sin(x^2) dx \tag{26.432}$$

où a est une constante strictement positive. Nous effectuons une intégration par partie en posant

$$u = \frac{1}{x} \qquad u' = -\frac{1}{x^2} \tag{26.433a}$$

$$v' = x\sin(x^2) \qquad v = \frac{1-\cos(x^2)}{2}. \tag{26.433b}$$

Notons que la primitive v a été choisie pour avoir $v(0) = 0$. Nous avons

$$\int_a^b \sin(x^2) dx = \left[\frac{1-\cos(x^2)}{2x}\right]_a^b - \int_a^b \frac{\cos(x^2)-1}{2x^2} dx \tag{26.434}$$

FqOdeKre

55. Définition 12.305.

Pour le premier terme nous avons

$$\lim_{b\to\infty}\left[\frac{1-\cos(x^2)}{2x}\right]_a^b=\lim_{b\to\infty}\frac{1-\cos(b^2)}{2b}-\frac{1-\cos(a^2)}{2a}=-\frac{1-\cos(a^2)}{2a}. \tag{26.435}$$

C'est borné. Pour le second terme de (26.434), la fonction

$$\frac{\cos(x^2)-1}{2x^2} \tag{26.436}$$

est majorée par la fonction $1/x^2$ qui est intégrable entre a et ∞.

Nous allons calculer l'intégrale demandée en passant par la fonction

$$f(x)=e^{-x^2} \tag{26.437}$$

définie sur le plan complexe. Nous l'intégrons sur le chemin $\gamma=\gamma_1+\gamma_2-\gamma_3$ indiqué à la figure 26.1.

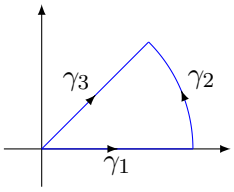


FIGURE 26.1: Chemin d'intégration pour l'intégrale de Fresnel LabelFigCheminFresnel

Ces chemins sont donnés par

$$\begin{aligned}\gamma_1\colon[0,R]&\to\mathbb{C}\\ t&\mapsto t,\end{aligned} \tag{26.438}$$

$$\begin{aligned}\gamma_2\colon[0,\frac{\pi}{4}]&\to\mathbb{C}\\ t&\mapsto Re^{it},\end{aligned} \tag{26.439}$$

$$\begin{aligned}\gamma_3\colon[0,R]&\to\mathbb{C}\\ t&\mapsto te^{i\pi/4}.\end{aligned} \tag{26.440}$$

Tout d'abord la fonction f est bien holomorphe par le critère du théorème 26.2. Le calcul de $\frac{\partial f}{\partial\bar{z}}$ se fait simplement en posant $f(x,y)=e^{-(x+iy)^2}$. Le calcul est usuel :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=exp(-(x+I*y)**2)
sage: A=f.diff(x)+I*f.diff(y)
sage: A.simplify_full()
(x, y) |--> 0
```

Nous avons donc

$$0=\int_\gamma f=\underbrace{\int_0^R e^{-t^2}dt}_{I_1(R)}+\underbrace{\int_0^{\pi/4}e^{-R^2e^{2it}}Rie^{it}dt}_{I_2(R)}+\underbrace{\int_0^R e^{-t^2e^{i\pi/2}}e^{i\pi/4}dt}_{I_3(R)}. \tag{26.441}$$ EqfaoRgU

L'intégrale est nulle pour tout R en vertu de la proposition 26.10. L'intégrale I_1 est une gaussienne et nous avons

$$\lim_{R\to\infty}I_1(R)=\frac{\sqrt{\pi}}{2} \tag{26.442}$$

par l'exemple 14.307. Nous montrons maintenant que $\lim_{R\to\infty}|I_2(R)|=0$ [56]. D'abord nous majorons en prenant la norme puis nous effectuons le changement de variables $u=2t$:

$$|I_2(R)|\leqslant\int_0^{\pi/4}Re^{-R^2\cos(2t)}dt \tag{26.443a}$$

$$=\frac{R}{2}\int_0^{\pi/2}e^{-R^2\cos(u)}du. \tag{26.443b}$$

Nous savons que le graphe du cosinus est concave : il reste au dessus de la droite que joint $(0,1)$ à $(\frac{\pi}{2},0)$. Du coup $\cos(u)\geqslant 1-\frac{2}{\pi}u$ et par conséquent

$$e^{-R^2\cos(u)}\leqslant e^{-R^2(1-\frac{2}{\pi}u)}=e^{R^2(\frac{2}{\pi}u-1)}. \tag{26.444}$$

Nous effectuons l'intégrale

$$|I_2(R)|\leqslant\frac{R}{2}\int_0^{\pi/2}e^{-R^2}e^{\frac{2R^2}{\pi}u}du \tag{26.445a}$$

$$=\frac{R}{2}e^{-R^2}\left[\frac{\pi}{2R^2}e^{2R^2u/\pi}\right]_0^{\pi/2} \tag{26.445b}$$

$$=\frac{\pi}{4R}-\frac{\pi e^{-R^2}}{4R}, \tag{26.445c}$$

et nous avons bien $\lim_{R\to\infty}|I_2(R)|=0$. Nous passons à la troisième intégrale. En tenant compte que $e^{i\pi/2}=i$, nous avons

$$I_3(R)=-\int_0^R e^{-\gamma_3(t)^2}e^{i\pi/4}dt \tag{26.446a}$$

$$=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-t^2}e^{2i\pi/4} \tag{26.446b}$$

$$=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-it^2}. \tag{26.446c}$$

En passant à la limite $R\to 0$, de l'équation (26.441) il ne reste que

$$0=\frac{\sqrt{2}}{2}-\frac{1+i}{\sqrt{2}}\int_0^\infty e^{-it^2}dt, \tag{26.447}$$

ce qui signifie que

$$\int_0^\infty e^{-it^2}dt=\frac{\sqrt{2\pi}}{2(1+i)}=\frac{\sqrt{\pi}}{2}e^{-i\pi/4}. \tag{26.448}$$

26.9 Théorème de Weierstrass

ThoArYtQO

Théorème 26.106 (Théorème de Weierstrass[649])**.**
Soit (f_n) une suite de fonctions holomorphes sur un ouvert Ω de $\mathbb{C}$ que nous supposons converger uniformément sur tout compact vers f. Alors f est holomorphe sur Ω et pour tout k nous avons

$$f_n^{(k)}\to f^{(k)} \tag{26.449}$$

uniformément sur tout compact.

Dit en peu de mots, la limite uniforme d'une suite de fonctions holomorphes est holomorphe, et on peut permuter la limite avec la dérivation.

56. Il y a moyen de démontrer cela via le lemme de Jordan[648]. Nous donnons ici une démonstration moins technologique.

Démonstration. Chacune des fonctions f_n étant holomorphes, si $a \in \Omega$ et r est tel que $B(a,r) \subset \Omega$, nous avons par la formule de Cauchy 26.59 :

$$f_n(z) = \frac{1}{2\pi i} \int_{\partial B(a,r)} \frac{f_n(\xi)}{\xi - z} d\xi \tag{26.450}$$

pour tout z dans un boule $B(a,\rho)$ incluse dans $B(a,r)$. Étant donné que le cercle ∂B est compact, elle y est majorée par une constante M. Montrons que de plus nous pouvons choisir M de telle façon à avoir $|f_n(\xi)| \leqslant M$ pour tout n et tout ξ en même temps. D'abord nous utilisons la continuité de la limite f sur le compact ∂B pour poser $A = \max_{z\in\partial B} |f(z)|$. Ensuite nous considérons un $\epsilon > 0$ et N tel que $\|f_n - f\|_{\partial B} \leqslant \epsilon$ pour tout $n \geqslant N$. Nous savons maintenant que

$$\{|f_n(\xi)| \text{ tel que } n \geqslant N, \xi \in \partial B\} \tag{26.451}$$

est majoré par $A + \epsilon$. Nous posons enfin

$$B = \max_{n\leqslant N} \max_{\xi\in\partial B} |f_n(z)|, \tag{26.452}$$

et alors le nombre $M = \max\{A + \epsilon, B\}$ majore $|f_n(\xi)|$ pour tout n et tout $\xi \in \partial B$.

De plus pour tout $\xi \in \partial B$ et pour tout z dans la petite boule, nous avons $|\xi - z| > r - \rho$, donc la fonction dans l'intégrale est majorée par une constante ne dépendant ni de n ni de ξ. Nous pouvons donc permuter l'intégrale et la limite sur n :

$$f(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{\xi - z}. \tag{26.453}$$

Cela implique que la fonction f est holomorphe par le corolaire 26.63.

Nous voudrions maintenant parler des dérivées des f_n et de f. Pour cela nous voulons permuter l'intégrale et les dérivées, ce qui est fait au corolaire 26.75 :

$$f_n^{(k)} = \frac{1}{2\pi i} \int_{\partial B(z_0,r)} \frac{f(\omega)}{(\omega - z)^{k+1}} d\omega. \tag{26.454}$$

Nous voulons la convergence sur tout compact contenu dans l'ouvert Ω. Pour ce faire, nous allons considérer un compact $K \subset \Omega$ et prouver la convergence uniforme dans toute boule de la forme $B(z_0,r)$ avec $z_0 \in K$ et $B(z_0,r) \subset \Omega$. Pour chaque tel couple (z_0,r), nous aurons un $N_{(z_0,r)} \in \mathbb{N}$ tel que si $n \geqslant N_{(z_0,r)}$,

$$\|f_n^{(k)} - f^{(k)}\|_{B(z_0,r)} \leqslant \epsilon. \tag{26.455}$$

Vu que ces boules $B(z_0,r)$ forment un recouvrement de K par des ouverts, nous pouvons en retirer un sous-recouvrement fini et prendre, comme N, le maximum des $N_{(z_0,r)}$ correspondants. Pour ce N nous aurons

$$\|f_n^{(k)} - f^{(k)}\|_K \leqslant \epsilon. \tag{26.456}$$

Au travail !

Pour $z \in B(z_0,r)$ nous considérons $r' > r$ tel que $B(z_0,r') \subset \Omega$ et nous avons

$$|f_n^{(k)}(z) - f^{(k)}(z)| = \left| \frac{1}{2\pi i} \int_{\partial B(z_0,r')} \frac{f_n(\xi) - f(\xi)}{(\xi - z)^{k+1}} d\xi \right| \tag{26.457a}$$

$$\leqslant \frac{1}{2\pi} \int_{\partial B(z_0,r')} \frac{|f_n(\xi) - f(\xi)|}{|r - r'|^{k+1}} d\xi. \tag{26.457b}$$

Nous avons pris ce r' de telle manière que $|\xi - z|$ soit borné par le bas par $|r - r'|$; sinon la majoration que nous venons de faire ne marche pas. Étant donné que $f_n \to f$ uniformément, nous pouvons considérer n assez grand pour que le numérateur soit plus petit que ϵ indépendamment de ξ et de z. Donc pour un n assez grand,

$$|f_n^{(k)}(z) - f^{(k)}(z)| \leqslant \frac{\epsilon}{2\pi} \frac{2\pi r'}{|r - r'|^{k+1}} \tag{26.458}$$

pour tout $z \in B(z_0,r)$. Donc nous avons convergence uniforme $f_n^{(k)} \to f^{(k)}$ sur cette boule. Par l'argument de compacité donné plus haut, nous avons la convergence uniforme sur tout compact. □

26.10 Analyse complexe en plusieurs variables

Définition 26.107 ([650]).
Nous définissons le ***polydisque*** *centré en $a \in \mathbb{C}^n$ et de « rayons » $r \in \mathbb{C}^n$ par*

$$D(a,r) = \{z \in \mathbb{C}^n \text{ tel que } |z_j - a_j| < r_j \, \forall j = 1, \dots, n\}. \tag{26.459}$$

Définition 26.108.
Pour une fonction $f : \mathbb{C}^n \to \mathbb{C}^n$, nous notons

$$\frac{\partial f}{\partial z_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} - i\frac{\partial}{\partial y_j}\right) f. \tag{26.460}$$

Définition 26.109 ([650]).
Soit un ouvert[57] *D dans $\mathbb{C}^n$. Nous disons que $f : D \to \mathbb{C}$ est* ***holomorphe*** *si*

(1) *f est continue sur D,*
(2) *pour tout $z \in D$ et pour tout $1 \leqslant j \leqslant n$, le nombre $\frac{\partial f}{\partial z_j}(z)$ existe et est fini.*

Définition 26.110 ([651]).
Soient des ouverts B_1 et B_2 dans $\mathbb{C}^n$. Nous disons que $f : B_1 \to B_2$ est ***biholomorphe*** *si*

(1) *f est holomorphe,*
(2) *f est une bijection entre B_1 et B_2,*
(3) *f^{-1} est holomorphe.*

THOooNBGZooHuGtxW

Théorème 26.111 (Inversion locale, version holomorphe)**.**
Soient des ouverts B_1 et B_2 dans $\mathbb{C}^n$. Nous considérons une fonction holomorphe $f : B_1 \to B_2$. Soit $z_0 \in B_1$ et $w_0 = f(z_0)$. Il y a équivalence entre

(1) $\det(df_{z_0}) \neq 0$
(2) *Il existe un voisinage U de z_0 dans B_1 et V de w_0 dans B_2 tels que la restriction $f : U \to V$ est biholomorphe.*

26.10.1 Inverse de fonctions analytiques

THOooXWCSooLSPzhO

Théorème 26.112 ([652]).
Une fonction analytique admet un inverse local analytique.

57. La topologie sur $\mathbb{C}^n$ est celle de la norme définie en 9.114.

Chapitre 27

Analyse fonctionnelle

27.1 Théorème d'isomorphisme de Banach

ThofQShsw

Théorème 27.1 (théorème d'isomorphisme de Banach[653]).
Une application linéaire continue et bijective entre deux espaces de Banach est un homéomorphisme[1].

Démonstration. Soit une application linéaire bijective et continue $f : E \to F$ entre deux espaces de Banach. En particulier elle est surjective, et le théorème de l'application ouverte 10.161 s'applique : f est une application ouverte.

Vu que f est bijective et ouverte, la proposition 10.157 implique que f^{-1} est continue. L'application f est donc continue d'inverse continue. Elle est donc un homéomorphisme. □

¡¡ Avertissement/question au lecteur !! 27.2
Il est donné dans [654] un exemple d'application linéaire et bijective dont l'inverse n'est pas continue. Si un jour vous le mettez au propre, faites-m'en profiter en m'envoyant une photo de votre feuille.

27.2 Théorème d'Ascoli

PropDGsPtpU

Proposition-Définition 27.3 (Application compacte[413]).
Soient E et F deux espaces vectoriels normés sur $\mathbb{R}$ ou $\mathbb{C}$ et une application $f \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes.

ITEMooAFNXooUGbZsh

(1) L'image d'un borné de E par f est relativement compacte[2] *dans F.*

ItemJIkpUbLii

(2) L'image par f de la boule unité fermée est relativement compacte dans F.

ITEMooQKISooBpyyee

(3) Si (x_n) est une suite bornée dans E, alors nous pouvons en extraire une sous-suite $(x_{\varphi(n)})$ telle que $y_n = f\big(x_{\varphi(n)}\big)$ converge dans F.

Une application vérifiant les conditions équivalentes de la proposition 27.3 est dite ***compacte****.*

Démonstration. En trois parties.

(1) **(1) ⇒ (2)** La boule unité fermée est bornée, et donc son image est relativement compacte.

(2) **(2) ⇒ (3)** Soient (x_n) bornée dans E, et soit M, une borne. Nous avons donc

$$z_n = f\left(\frac{x_n}{M}\right) \in f\big(B(0,1)\big) \subset \overline{f\big(B(0,1)\big)}. \tag{27.1}$$

1. Définition 7.41.
2. Relativement compact, définition 7.103.

Par hypothèse, la partie $\overline{f\big(B(0,1)\big)}$ est compacte ; la suite (z_n) est contenue dans ce compact. Nous en déduisons que (z_n) admet une sous-suite convergente. Soit donc $\varphi\colon \mathbb{N}\to\mathbb{N}$ telle que

$$f\left(\frac{x_{\varphi(n)}}{M}\right)\to y. \tag{27.2}$$

Vu que f est linéaire, nous avons aussi $f\big(x_{\varphi(n)}\big)\to My$.

(3) **(3) ⇒ (1)** Soit une partie bornée A de F. Nous devons prouver que $\overline{f(A)}$ est compact. Pour cela nous considérons une suite (y_n) dans $\overline{f(A)}$. Soit $\epsilon_k\to 0$. Pour chaque n, l'élément y_n est dans la fermeture de $f(A)$, c'est-à-dire qu'il existe un élément de $f(A)$ dont la distance à y_n est plus petite que ϵ_n. Bref, il existe une suite (x_n) dans A telle que

$$\|f(x_n)-y_n\|\leqslant\epsilon_k. \tag{27.3}$$

La suite (x_n) est bornée parce qu'elle est contenue dans A. Donc il existe une sous-suite $\varphi\colon\mathbb{N}\to\mathbb{N}$ telle que

$$f\big(x_{\varphi(n)}\big)\to y \tag{27.4}$$

pour un certain $y\in\overline{f(A)}$. Le fait que y soit dans $\overline{f(A)}$ est dû au fait que la suite $n\mapsto f(x_n)$ est une suite dans $f(A)$; donc la limite est dans la fermeture de $f(A)$.
Nous avons

$$\begin{aligned}\|y_{\varphi(n)}-y\|&\leqslant\|y_{\varphi(n)}-f\big(x_{\varphi(n)}\big)\|+\|f\big(x_{\varphi(n)}\big)-y\| &\text{(27.5a)}\\ &\leqslant\epsilon_k+\|f\big(x_{\varphi(n)}\big)-y\|\to 0. &\text{(27.5b)}\end{aligned}$$

Donc la suite $(y_{\varphi(n)})$ converge vers $y\in\overline{f(A)}$. Nous avons montré que toute dans $\overline{f(A)}$ possédait une sous-suite convergente dans $\overline{f(A)}$. Le théorème de Bolzano-Weierstrass 7.144 conclut que $\overline{f(A)}$ est compact et donc que $f(A)$ est relativement compact.

□
ThoKRbtpah

Théorème 27.4 (Théorème d'Ascoli[655]).
Soient un espace topologique compact K et un espace métrique (E,d). Nous considérons la topologie uniforme sur $C^0(K,E)$. Une partie A de $C^0(K,E)$ est relativement compacte si et seulement si les deux conditions suivantes sont remplies :

(1) A est équicontinu[3]*,*

(2) $\forall x\in K$, l'ensemble $\{f(x)$ tel que $f\in A\}$ est relativement compact dans E.

La version suivante du théorème de Banach-Steinhaus est énoncée de façon *ad hoc* pour fonctionner avec l'espace $\mathscr{D}(K)$ des fonctions de classe C^∞ à support dans le compact K. Un énoncé un peu plus fort est donné dans le cadre des espaces de Fréchet dans [216].

ThoNBrmGIg

Théorème 27.5 (Banach-Steinhaus avec des seminormes[1]).
Soit (E,d) un espace vectoriel métrique complet dont la topologie est également[4] *donnée par une famille $\mathcal{P}=\{p_i\}_{i\in I}$ de seminormes. Soit $\{T_\alpha\}_{\alpha\in A}$ une famille d'applications linéaires continues $T_\alpha\colon E\to\mathbb{R}$ telles que pour tout $x\in E$ nous ayons*

$$\sup_{\alpha\in A}\big|T_\alpha(x)\big|<\infty. \tag{27.6}$$

Alors il existe une constante $K>0$ et un sous-ensemble fini $J\subset I$ tels que pour tout $x\in E$ et pour tout $\alpha\in A$ nous ayons

$$\big|T_\alpha(x)\big|\leqslant K\max_{j\in J}p_j(x). \tag{27.7}$$
EqIFNGbtr

3. Définition 7.317.
4. Au sens où les ouverts sont les mêmes.

Démonstration. Pour chaque $k \in \mathbb{N}\backslash\{0\}$ nous posons

$$\Omega_k = \{x \in E \text{ tel que } \sup_{\alpha\in A} \left|T_\alpha(x)\right| > k\}. \tag{27.8}$$

Ces ensembles sont des ouverts (pour la même raison que dans la preuve du théorème 10.152) et leur union est E en entier parce que par hypothèse $\sup_{\alpha\in A} \left|T_\alpha(x)\right| < \infty$.

(1) **Les Ω_k ne sont pas tous denses** Le théorème 7.366 nous dit que (E, d) est un espace de Baire. Supposons que les Ω_k sont tous denses. Dans ce cas, l'intersection des Ω_k est également dense (c'est la définition 7.364 d'un espace de Baire), et en particulier non vide.
Soit $x_0 \in \bigcap_{k\in\mathbb{N}} \Omega_k$. Nous avons

$$\sup_{\alpha\in A} \left|T_\alpha(x_0)\right| = \infty, \tag{27.9}$$

ce qui contredirait l'hypothèse. Donc les Ω_k ne sont pas tous denses.

(2) **Un non dense** Soit $k_0 \in \mathbb{N}^*$ tel que Ω_{k_0} n'est pas dense dans E. Il existe donc $x_0 \in E$ et un ouvert $\mathcal{O}$ autour de x_0 n'intersectant pas Ω_{k_0}.

(3) **Topologie**
Parlons de topologie sur E. L'ouvert $\mathcal{O}$ est un d-ouvert et donc un $\mathcal{P}$-ouvert, lequel contient une $\mathcal{P}$-boule ouverte [5]. Cette dernière boule n'est pas spécialement une d-boule, mais c'est un d-ouvert.

(4) **Une majoration**
Il existe dont J fini dans I et $\rho > 0$ tels que $B_J(x_0, \rho) \subset \mathcal{O}$ et donc tels que

$$B_J(x_0, \rho) \cap \Omega_{k_0} = \varnothing. \tag{27.10}$$

Donc pour tout $z \in B_J(x_0, \rho)$ nous avons

$$\sup_{\alpha\in A} \left|T_\alpha(z)\right| \leqslant k_0. \tag{27.11}$$

Si maintenant $y \in B_J(0, \rho)$, nous avons $y = (x_0 + y) - x_0$ et donc

$$\begin{aligned} \left|T_\alpha(y)\right| &= \left|T_\alpha(x_0 + y) - T_\alpha(x_0)\right| && (27.12a)\\ &\leqslant \left|T_\alpha(x_0 + y)\right| + \left|T_\alpha(x_0)\right| && (27.12b)\\ &\leqslant k_0 + C. && (27.12c) \end{aligned}$$

Justifications.
— $x_0 + y \in B_J(x_0, \rho)$, donc $\left|T_\alpha(x_0 + y)\right| \leqslant k_0$.
— nous avons posé $C = \sup_{\alpha\in A} \left|T_\alpha(x_0)\right|$

(5) **Sur la boule unité** En divisant par ρ, pour $y \in B_J(0, 1)$ nous avons

$$\left|T_\alpha(y)\right| \leqslant \rho^{-1}(k_0 + C). \tag{27.13}$$

EQooOVHIooSonnig

(6) **Un bon élément dans la boule unité** Pour n'importe quel $x \in E$ nous avons

$$\frac{x}{\max_{j\in J} p_j(x)} \in B_J(0, 1). \tag{27.14}$$

Nous pouvons donc appliquer (27.13) à cet élément :

$$\left|T_\alpha\left(\frac{x}{\max_{j\in J} p_j(x)}\right)\right| \leqslant \rho^{-1}(k_0 + C). \tag{27.15}$$

Utilisant encore la linéarité de T_α nous trouvons ce que nous devions trouver :

$$\left|T_\alpha(x)\right| \leqslant \max_{j\in J} p_j(x)\rho^{-1}(k_0 + C) = K \max_{j\in J} p_j(x) \tag{27.16}$$

en posant $K = \rho^{-1}(k_0 + C)$.

5. Définition 7.342.

□
CorPGwLluz

Corolaire 27.6 ([1, 195]).
Soit $\big(X, \{p_i\}_{i\in I}\big)$, un espace vectoriel sur $\mathbb{C}$ seminormé qui est

(1) localement compact[6]*,*

(2) métrisable et complet.

Soit $(T_j)_{j\in\mathbb{N}}$ une suite d'application linéaires continues $X \to \mathbb{C}$ telles que pour tout $x \in X$ il existe[7] *$\alpha_x \in \mathbb{C}$ tel que*

$$T_j x \xrightarrow{\mathbb{C}} \alpha_x. \tag{27.17}$$

Si nous posons $Tx = \alpha_x$ alors
ItemAEOtOMLi

(1) L'application T est linéaire et continue.
ITEMooEVIXooBpaWOc

(2) Si $x_k \xrightarrow{X} x$, alors

$$|T_k(x_k) - T(x_k)| \xrightarrow{\mathbb{R}} 0. \tag{27.18}$$
ITEMooJYOVooPIkHBo

(3) Pour tout compact K dans X nous avons

$$\sup_{x\in K} |T_k x - Tx| \xrightarrow{\mathbb{R}} 0. \tag{27.19}$$
ItemAEOtOMLiii

(4) si $x_k \to x$ dans X alors

$$T_k x_k \xrightarrow{\mathbb{C}} Tx. \tag{27.20}$$

Démonstration. En plusieurs parties.

(1) **Linéaire** Vu que X est seminormé et que $\mathbb{C}$ est un corps valué, la proposition 7.347 nous indique que X est un espace vectoriel topologique et la proposition 7.202 nous dit que la limite de suite est une opération linéaire. Pour $x, y \in X$ nous pouvons donc faire

$$\begin{aligned}
T(x+y) &= \lim_{j\to\infty} \big(T_j(x+y)\big) && (27.21a)\\
&= \lim_{j\to\infty} \big(T_j(x) + T_j(y)\big) && (27.21b)\\
&= \lim_{j\to\infty} T_j(x) + \lim_{j\to\infty} T_j(y) && (27.21c)\\
&= T(x) + T(y). && (27.21d)
\end{aligned}$$

De même nous avons

$$\begin{aligned}
T(\lambda x) &= \lim_{k\to\infty} \big(T_j(\lambda x)\big) && (27.22a)\\
&= \lim_{j\to\infty} \big(\lambda T_j(x)\big) && (27.22b)\\
&= \lambda \lim_{j\to\infty} T_j(x) && (27.22c)\\
&= \lambda T(x). && (27.22d)
\end{aligned}$$

(2) **Séquentiellement continue** Nous prouvons que T est séquentiellement continue[8]. Soit une suite $x_k \xrightarrow{X} x$. Nous voulons utiliser le théorème de Banach-Steinhaus 27.5. L'ensemble $\{T_k\}_{k\in\mathbb{N}}$ est un ensemble d'applications linéaires continues. De plus, si $x \in X$, la suite $k \mapsto T_k(x)$ est une suite convergente dans $\mathbb{C}$; elle est donc bornée :

$$\sup_{k\in\mathbb{N}} |T_k(x)| < \infty. \tag{27.23}$$

6. Définition 7.86.
7. L'unicité des α_x et donc le fait que tout le reste ait un sens provient de l'hypothèse de séparabilité et la proposition 7.62.
8. Définition 7.208

Le théorème de Banach-Steinhaus nous dit qu'il existe un $C > 0$ et J fini dans I tels que pour tout k et pour tout x nous ayons

$$|T_k(x)| \leqslant C \max_{j\in J} p_j(x). \tag{27.24}$$

Un seul $i \in J$ nous suffira. Soient donc $C > 0$ et $i \in I$ tels que $|T_k(x)| \leqslant Cp_i(x)$ pour tout x. Vu que le module est une opération continue sur $\mathbb{C}$, elle commute avec la limite et nous pouvons faire le calcul suivant :

$$\begin{aligned}
|T(x)| &= |\lim_{k\to\infty} T_k(x)| && \text{(27.25a)}\\
&= \lim_{k\to\infty} |T_k(x)| && \text{(27.25b)}\\
&\leqslant \lim_{k\to\infty} Cp_i(x) && \text{(27.25c)}\\
&= Cp_i(x). && \text{(27.25d)}
\end{aligned}$$

Avec ça nous pouvons prouver que T est séquentiellement continue. Soit $x_k \xrightarrow{X} 0$. En utilisant la proposition 7.344 nous avons :

$$|T(x_k)| \leqslant Cp_i(x_k) \xrightarrow{\mathbb{R}} 0 \tag{27.26}$$

Donc $|T(x_k)| \to 0$ et T est séquentiellement continue en zéro. Vu que T est linéaire, elle est séquentiellement continue partout.

(3) **Continue** Vu que l'espace X est métrisable, la proposition 7.255 conclut que T est continue.

(4) **Point (2)**

Nous passons à la preuve de (2). Supposons que $x_k \xrightarrow{X} x$. Voici des majorations avec justifications juste en-dessous :

$$\begin{aligned}
|T_k(x_k) - T(x_k)| &\leqslant |T_k(x_k) - T_k(x)| + |T_k(x) - T(x_k)| && \text{(27.27a)}\\
&= |T_k(x_k - x)| + |T_k(x) - T(x_k)| && \text{(27.27b)}\\
&\leqslant Cp_i\big(x_k - x\big) + |T_k(x) - T(x_k)| && \text{(27.27c)}\\
&\leqslant C\epsilon + |T_k(x) - T(x_k)| && \text{(27.27d)}\\
&\leqslant C\epsilon + |T_k(x) - T(x)| + |T(x) - T(x_k)| && \text{(27.27e)}\\
&\leqslant (C + 2)\epsilon. && \text{(27.27f)}
\end{aligned}$$

SUBEQooQLXIooSyvNcx
SUBEQooVLNVooOVsQno
SUBEQooHSTIooJgfChZ

Justifications :

— Pour (27.27c). C'est le théorème de Banach-Steinhaus 27.5 qui nous assure l'existence d'un $i \in I$ et d'une $C \in \mathbb{R}^+$ qui donne la majoration.

— Pour (27.27d). Nous nous sommes fixé un $\epsilon > 0$ et nous avons pris k assez grand. Pour la topologie des seminormes, la convergence $y_k \xrightarrow{X} y$ implique la limite $p_i(y_k - y) \xrightarrow{\mathbb{R}} 0$ pour toutes les seminormes, et en particulier pour celle donnée par le théorème de Banach-Steinhaus.

— Pour (27.27f). D'une part, par hypothèse et par définition de T, nous avons $T_k(x) \xrightarrow{\mathbb{C}} T(x)$ et d'autre part par continuité séquentielle de T et par convergence de $x_k \xrightarrow{X} x$ nous avons $T(x_k) \xrightarrow{\mathbb{C}} T(x)$.
En prenant k encore plus grand, nous avons toutes les majorations en même temps.

Nous avons donc bien

$$|T_k(x_k) - T(x_k)| \xrightarrow{\mathbb{R}} 0. \tag{27.28}$$

(5) **Point (3)** Nous nous lançons dans la preuve du point (3). En vertu du point (1), l'application

$$\begin{aligned}
S_k\colon K &\to \mathbb{R}\\
x &\mapsto |T_k(x) - T(x)|
\end{aligned} \tag{27.29}$$

est continue. Vu que K est un compact, le théorème 7.146, nous indique qu'il existe x_k tel que

$$|T_k(x_k) - T(x_k)| = \sup_{x\in K} |T_k(x) - T(x)|. \tag{27.30}$$

Nous considérons la suite

$$a_k = |T_k(x_k) - T(x_k)| \tag{27.31}$$

et nous allons montrer que

— toutes les sous-suites convergentes de (a_k) convergent vers 0.

— la suite (a_k) est contenue dans un compact de $\mathbb{R}$.

Ensuite le corolaire 7.291 terminera.

(5a) **Toutes les sous-suites convergentes** Soit $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que la suite $b_k = a_{\varphi(k)}$ soit convergente. Nous allons trouver une sous-suite qui converge vers 0. La suite $k \mapsto a_{\varphi(k)}$ est une suite dans le compact K ; elle contient donc une sous-suite convergente. Soit donc $\psi\colon \mathbb{N} \to \mathbb{N}$ telle que $k \mapsto a_{(\varphi\circ\psi)(k)}$ soit convergente. Nous étudions la sous-suite $c_k = b_{\psi(k)}$ et nous allons prouver que $c_k \to 0$.

Pour ne pas mourir sous les notations, nous posons $\sigma = \varphi\circ\psi$. Nous avons donc $x_{\sigma(k)} \xrightarrow{X} x$, et donc $|T_{\sigma(k)}(x_{\sigma(k)}) - T(x_{\sigma(k)})| \xrightarrow{\mathbb{R}} 0$ par le point (2).
Donc la suite $a_{\varphi(k)}$ qui est convergente contient une sous-suite qui converge vers 0. La suite $a_{\varphi(k)}$ converge donc vers 0.

(5b) **Dans un compact** Nous allons prouver que $|T_k(x_k) - T(x_k)|$ est bornée. Supposons que ce ne soit pas le cas. Pour chaque $n \in \mathbb{N}$, il existe un k tel que $|T_k(x_k) - T(x_k)| > n$. En numérotant bien, il existe une application $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que

$$|T_{\varphi(k)}(x_{\varphi(k)}) - T(x_{\varphi(k)})| > k. \tag{27.32}$$ EQooRGVOooXJDkQx

La suite $k \mapsto x_{\varphi(k)}$ est une suite dans le compact K. Elle admet une sous-suite convergente. En nommant $x_{\sigma(k)}$ cette sous-suite, nous avons $x_{\sigma(k)} \xrightarrow{X} x$, et nous nous avons encore $|T_{\sigma(k)}(x_{\sigma(k)}) - T(x_{\sigma(k)})| \xrightarrow{\mathbb{R}} 0$.
Ouais mais bon. Une suite qui vérifie (27.32) ne peut pas avoir une sous-suite convergente. Contradiction.

(6) **Point (4)** Nous avons la majoration

$$|T_k x_k - Tx| \leqslant |T_k x_k - Tx_k| + |Tx_k - T_x|. \tag{27.33}$$

Par le point (2), le premier terme tend vers zéro. La continuité de T (point (1)) s'occupe du deuxième terme.

□

27.3 Espaces de Lebesgue L^p

SecVKiVIQK

27.3.1 Généralités

DEFooKMJQooXeaUtp

Définition 27.7.
*Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré. Deux fonctions à valeurs complexes f et g sur cet espaces sont dites **équivalentes** et nous notons $f \sim g$ si elles sont μ-presque partout égales. Nous notons $[f]$ la classe de f pour cette relation.*

Nous allons souvent noter f indifféremment pour la fonction f et un des représentants de la classe de f. Toutefois, lorsque la distinction sera importante, nous essayerons de faire faire l'effort de distinguer la fonction f de sa classe $[f]$.

Lemme 27.8.
Une classe contient au maximum une seule fonction continue.

Démonstration. Soient deux fonctions continues f_1 et f_2 avec $f_1(a) \neq f_2(a)$. Si $|f_1(a) - f_2(a)| > \delta$ alors il existe un ϵ tel que $|f_1(x) - f_1(a)| > \delta$ pour tout $x \in B(a, \epsilon)$. En particulier $f_1 \neq f_2$ sur $B(a, \epsilon)$. Cette dernière boule est de mesure de Lebesgue non nulle ; ergo f_1 et f_2 ne sont pas dans la même classe. □

DEFooTHIDooWYzBtn

Définition 27.9.
Pour $p \geqslant 1$, nous introduisons l'opération

$$\|f\|_p = \left(\int_\Omega |f(x)|^p d\mu(x) \right)^{1/p} \tag{27.34}$$

EQooBDBXooCHRmpo

et nous notons $\mathcal{L}^p(\Omega, \mu)$ l'ensemble des fonctions mesurables sur Ω telles que $\|f\|_p < \infty$.

27.10.
Le fait que f doive être mesurable pour être dans L^p est bien dans la définition de L^p, et non une propriété qui pourrait être déduite de la finitude de l'intégrale (27.34).

Soit $\Omega = \mathbb{R}$ muni de ses boréliens ou même de sa tribu de Lebesgue et de la mesure de Lebesgue. Si A est une partie bornée non mesurable (exemple 14.157), alors nous considérons un borné mesurable K contenant A et la fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in K \backslash A \\ 0 & \text{sinon.} \end{cases} \tag{27.35}$$

Nous avons alors $|f| = \mathbb{1}_K$. Vu que K est borné, $\int_{\mathbb{R}} |f| = \mu(K)$ ne présente pas de problèmes. Donc f serait un élément de $L^1(\mathbb{R})$ sans être mesurable.

Pour éviter cela, nous incluons la mesurabilité dans la définition de $\mathcal{L}^p$.

LEMooHVZGooRwHXMk

Lemme 27.11 ([1]).
L'ensemble $\mathcal{L}^p$ est un espace vectoriel sur $\mathbb{C}$.

Démonstration. Pour rappel, nous ne considérons les choses que pour $p \geqslant 1$. Le fait que si $f \in L^p$, alors $\lambda f \in L^p$ est évident. Ce qui est moins immédiat, c'est le fait que $f + g \in L^p$ lorsque f et g sont dans L^p. Cela découle d'une part du fait que la fonction $\varphi : x \mapsto x^p$ est convexe sur les positifs (lemme 17.87), de telle sorte que

$$\varphi\left(\frac{|a| + |b|}{2} \right) \leqslant \frac{\varphi(|a|) + \varphi(|b|)}{2}, \tag{27.36}$$

ou encore

$$(|a| + |b|)^p \leqslant 2^{p-1}(|a|^p + |b|^p) \tag{27.37}$$

EqZESduEa

Et d'autre part, nous savons que pour $z_1, z_2 \in \mathbb{C}$, $|z_1 + z_2| \leqslant |z_1| + |z_2|$ (proposition 9.102(7)). Donc

EQooKRMEooSLHUUc

$$\begin{aligned}
\int_\Omega |f + g|^p d\mu &= \int_\Omega |f(\omega) + g(\omega)|^p d\mu(\omega) && \text{(27.38a)} \\
&\leqslant \int_\Omega \big(|f(\omega)| + |g(\omega)| \big)^p d\mu(\omega) && \text{(27.38b)} \\
&\leqslant 2^{p-1} \int_\Omega \big(|f(\omega)|^p + |g(\omega)|^p \big) d\mu(\omega) && \text{(27.38c)} \\
&= 2^{p-1} \big(\|f\|_p^p + |g|_p^p \big). && \text{(27.38d)}
\end{aligned}$$

Donc si f et g sont dans $\mathcal{L}^p$, alors $f + g$ est dans $\mathcal{L}^p$. □

27.12.
Il est à noter que nous ne considérons que des valeurs $p \geqslant 1$, précisément parce que la fonction $x \mapsto |x|^p$ n'est pas convexe lorsque $p < 1$.

Dans le même ordre d'idées, si $p \geqslant 1$, alors le $q \in \mathbb{R}$ tel que

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{27.39}$$

est également $q \geqslant 1$. Cela est important pour un certain nombres de théorèmes qui vont venir, en particulier l'inégalité de Hölder (27.81).

Si vous en voulez à propos de $0 < p < 1$, vous pouvez lire [656].

27.13.
L'opération $f \mapsto \|f\|_p$ n'est pas une norme sur $\mathcal{L}^p$ parce que pour f presque partout nulle, nous avons $\|f\|_p = 0$. Il y a donc des fonctions non nulles sur lesquelles $\|.\|_p$ s'annule.

27.14.
Soit un espace mesuré non complet[9] $(\Omega, \mathcal{A}, \mu)$. Il existe une partie N de mesure nulle et $A \subset N$ non mesurable. Considérons

$$f(x) = \begin{cases} 2 & \text{si } x \in A \\ 1 & \text{si } x \in N \backslash A \\ 0 & \text{sinon.} \end{cases} \tag{27.40}$$

Cette fonction est non nulle exactement sur N. Donc $f \sim 0$. Mais f n'est pas mesurable parce que $f^{-1}(2) = A$ n'est pas mesurable.

Il est donc possible d'être dans la classe d'équivalence d'une fonction mesurable sans être mesurable. Ceci est cependant un détail (presque) sans importance pour deux raisons.

— La mesure de Lebesgue est complète par définition (définition 14.143) si nous considérons bien la tribu de Lebesgue et non seulement les boréliens.

— Dans le cas des espaces de Lebesgue $L^p(\Omega, \mathcal{A}, \mu)$, il s'agit d'un quotient de $\mathcal{L}^p$ qui ne contient que des fonctions mesurables. Donc dans l'étude de L^p, tous les représentants sont mesurables, même si $(\Omega, \mathcal{A}, \mu)$ n'est pas complet.

LemKZVHVAR

Lemme 27.15.
Si $f \in \mathcal{L}^p(\Omega)$ et $f \sim g$, alors $g \in \mathcal{L}^p(\Omega)$ et $\|f\|_p = \|g\|_p$.

Démonstration. Soit $h(x) = |g(x)|^p - |f(x)|^p$; c'est une fonction par hypothèse presque partout nulle et donc intégrable sur Ω ; son intégrale y vaut zéro. Nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = \int_\Omega \big(|f(x)|^p + h(x)\big) d\mu(x) = \int_\Omega |g(x)|^p d\mu(x). \tag{27.41}$$

Cela prouve que la dernière intégrale existe et vaut la même chose que la première. □

Nous pouvons donc considérer la norme $|.|_p$ comme une norme sur l'ensemble des classes plutôt que sur l'ensemble des fonctions. Nous notons L^p l'ensemble des classes des fonctions de $\mathcal{L}^p$. Cet espace est muni de la norme

$$\|[f]\|_p = \|f\|_p, \tag{27.42}$$

formule qui ne dépend pas du représentant par le lemme 27.15.

PROPooTYCYooAKJWOX

Proposition 27.16.
Soient un espace mesuré $(\Omega, \mathcal{A}, \mu)$, et $p \geqslant 1$.

(1) L'ensemble des classes $L^p(\Omega, \mathcal{A}, \mu)$ est un espace vectoriel.

9. Définition 14.68.

ITEMooZNXPooGzFnIK

(2) *La formule*

$$\|[f]\|_p = \left(\int_\Omega |f(x)|^p d\mu(x) \right)^{1/p} \tag{27.43}$$

définit une norme[10] *sur* $L^p(\Omega, \mathcal{A}, \mu)$.

Démonstration. En deux parties.

(1) **Espace vectoriel** Le lemme 27.11 dit que $\mathcal{L}^p$ est un espace vectoriel. Une structure d'espace vectoriel sur $L^p(\Omega, \mathcal{A}, \mu)$ est donnée en posant

$$[f] + [g] = [f + g] \tag{27.44}$$

et

$$\lambda[f] = [\lambda f] \tag{27.45}$$

qui sont deux définitions correctes parce qu'elles ne dépendent pas du choix du représentant. De plus le lemme 27.11 dit que $f + g \in \mathcal{L}^p$ dès que $f, g \in \mathcal{L}^p$. Donc $[f + g] \in L^p$ dès que $[f], [g] \in L^p$.

(2) **Norme**

Pour être une norme, il faut vérifier les trois propriétés de la définition 7.158.

D'abord, si $\|[f]\|_p = 0$, nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = 0, \tag{27.46}$$

ce qui par le lemme 14.199 implique que $|f(x)|^p = 0$ pour presque tout x. Ou encore $f \sim 0$, c'est-à-dire $[f] = [0]$ au niveau des classes.

Ensuite pour $\lambda \in \mathbb{C}$ et $f \in L^p$ nous avons

$$\|\lambda f\|_p = \left(\int_\Omega |\lambda f(\omega)|^p d\mu(\omega) \right)^{1/p} \tag{27.47a}$$

$$= \left(\int_\Omega |\lambda|^p |f(\omega)|^p d\mu(\omega) \right)^{1/p} \tag{27.47b}$$

$$= |\lambda| \|f\|_p. \tag{27.47c}$$

Et enfin, en suivant le calcul (27.38) nous avons

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p. \tag{27.48}$$

□

27.17.

À partir de maintenant $\big(L^p(\Omega, \mu), \|.\|_p\big)$ est un espace métrique avec toute la topologie qui va avec.

Dans la suite nous n'allons pas toujours écrire $[f]$ pour la classe de f. Par abus de notations nous allons souvent parler de $f \in L^p$ comme si c'était une fonction.

De même nous notons $L^p(\Omega)$ ou $L^p(\Omega, \mu)$ ou $L^p(\Omega, \mathcal{A}, \mu)$ d'après ce sur quoi nous voulons insister. Mais seule la dernière notation est parfaitement correcte.

10. Définition 7.158.

27.3.2 Un peu de convergence de suites

PropWoywYG

Proposition 27.18 ([657]).
Soit $1 \leqslant p \leqslant \infty$ et supposons que la suite $[f_n]$ dans $L^p(\Omega, \mathcal{F}, \mu)$ converge vers $[f]$ au sens L^p. Alors il existe une sous-suite (h_n) qui converge ponctuellement μ-presque partout vers f.

Démonstration. Si $p = \infty$ nous sommes en train de parler de la convergence uniforme et il ne faut même pas prendre ni de sous-suite ni de « presque partout ».

Supposons que $1 \leqslant p < \infty$. Nous considérons une sous-suite $[h_n]$ de $[f_n]$ telle que

$$\|[h_j] - [f]\|_p < 2^{-j}, \tag{27.49}$$

puis nous posons $u_k(x) = |h_k(x) - f(x)|^p$. Notons que ce u_k est une vraie fonction, pas une classe. Et en plus c'est une fonction positive. Nous avons

$$\int_\Omega u_k d\mu = \int_\Omega |h_k(x) - f(x)|^p d\mu(x) = \|h_k - f\|_p^p \leqslant 2^{-kp}. \tag{27.50}$$

Vu que u_k est une fonction positive la suite des sommes partielles de $\sum_k u_k$ est croissante et vérifie donc le théorème de la convergence monotone 14.178 :

$$\int_\Omega \left(\sum_{k=0}^{\infty} u_k(x) \right) d\mu(x) = \sum_{k=0}^{\infty} \int_\Omega u_k(x) d\mu(x) \leqslant \sum_{k=0}^{\infty} 2^{-kp} < \infty. \tag{27.51}$$

Le fait que l'intégrale de la fonction $\sum_k u_k$ est finie implique que cette fonction est finie μ-presque partout. Donc le terme général tend vers zéro presque partout, c'est-à-dire

$$|h_k(x) - f(x)|^p \to 0. \tag{27.52}$$

Cela signifie que $h_k \to f$ presque partout ponctuellement. □

Est-ce qu'on peut faire mieux que la convergence ponctuelle presque partout d'une sous-suite ? En tout cas on ne peut pas espérer grand chose comme convergence pour la suite elle-même, comme le montre l'exemple suivant.

ExPOmxICc

Exemple 27.19 ([658]).
Nous allons montrer une suite de fonctions qui converge vers zéro dans $L^p[0,1]$ (avec $p < \infty$) mais qui ne converge ponctuellement pour *aucun* point.

Nous construisons la suite de fonctions par paquets. Le premier paquet est formé de la fonction constante 1.

Le second paquet est formé de deux fonctions. La première est $\mathbb{1}_{[0,1/2]}$ et la seconde $\mathbb{1}_{[1/2,1]}$.

Plus généralement le paquet numéro k est constitué des k fonctions $\mathbb{1}_{[i/k,(i+1)/k]}$ avec $i = 0, \ldots, k-1$.

Vu que les fonctions du paquet numéro k ont pour norme $\|f\|_p = \frac{1}{k}$, nous avons évidemment $f_n \to 0$ dans L^p. Il est par contre visible que chaque paquet passe en revue tous les points de $[0,1]$. Donc pour tout x et pour tout N, il existe (même une infinité) $n > N$ tel que $f_n(x) = 1$. Il n'y a donc convergence ponctuelle nulle part. △

La proposition suivante est une espèce de convergence dominée de Lebesgue pour L^p.

PropBVHXycL

Proposition 27.20.
Soit $f \in L^p(\Omega)$ avec $1 \leqslant p < \infty$ et (f_n) une suite de fonctions convergeant ponctuellement vers f et telle que $|f_n| \leqslant |f|$. Alors $f_n \xrightarrow{L^p} f$.

Démonstration. Nous avons immédiatement $|f_n(x)|^p \leqslant |f(x)|^p$, de telle sorte que le théorème de la convergence dominée implique que $f_n \in L^p$. La convergence dominée donne aussi que $\|f_n\|_p \to \|f\|_p$, mais cela ne nous intéresse pas ici.

Nous posons $h_n(x) = |f_n(x) - f(x)|$. En reprenant la formule de majoration (27.37) et en tenant compte du fait que $|f_n(x)| \leqslant |f(x)|$, nous avons

$$h_n(x) \leqslant 2^{p-1}\big(|f_n(x)|^p + |f(x)|^p\big) \leqslant 2^p |f(x)|^p, \tag{27.53}$$

ce qui prouve que $|h_n|$ est uniformément (en n) majorée par une fonction intégrable, donc h_n est intégrable et on peut permuter la limite et l'intégrale (théorème de la convergence dominée 14.207) :

$$\lim_{n\to\infty} \|f_n - f\|_p^p = \lim_{n\to\infty} \int_{\mathbb{R}^d} |f_n(x) - f(x)|^p dx = \int_{\mathbb{R}^d} \lim_{n\to\infty} h_n(x) dx = 0. \tag{27.54}$$

□

PropRERZooYcEchc

Proposition 27.21 ([1])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et une fonction mesurable $f : \Omega \to \mathbb{C}$ telle que

$$\int_\Omega f \mathbb{1}_A d\mu = 0 \tag{27.55}$$

pour toute partie mesurable A de mesure finie. Alors $[f] = 0$, c'est-à-dire que f est non nulle uniquement sur une partie de mesure nulle.

Démonstration. Nous rappelons que dire que f est intégrable signifie que $\mathrm{Re}(f)^+$, $\mathrm{Re}(f)^-$, $\mathrm{Im}(f)^+$ et $\mathrm{Im}(f)^-$ sont intégrables.

Les parties

$$\{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \in [k, k+1[\}_{k=1,\dots} \tag{27.56}$$

et

$$\{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \in [\frac{1}{k+1}, \frac{1}{k}[\}_{k=1,\dots} \tag{27.57}$$

sont mesurables en tant qu'images inverses de mesurables par la fonction mesurable $\mathrm{Re}(f)^+$.

Nous notons $\{A_i\}_{i\in\mathbb{N}}$ une énumération quelconque de ces parties. L'important est que

$$\bigcup_{i\in\mathbb{N}} A_i = \{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \neq 0\} \tag{27.58}$$

et que pour chaque i, il existe $\alpha_i > 0$ tel que $\mathrm{Re}(f)^+(x) > \alpha_i$ pour tout x dans A_i.

Par hypothèse nous avons $\int_\Omega f \mathbb{1}_{A_i} d\mu = 0$. En particulier,

$$0 = \int_\Omega \mathrm{Re}(f\mathbb{1}_{A_i})^+ = \int_\Omega \mathrm{Re}(f)^+(x) \mathbb{1}_{A_i}(x) d\mu(x). \tag{27.59}$$

Mais pour tout $x \in A_i$ nous avons $\mathrm{Re}(f)^+(x) > \alpha_i > 0$, donc

$$0 = \int_\Omega \mathrm{Re}(f)^+(x) \mathbb{1}_{A_i}(x) d\mu(x) \geqslant \alpha_i \int_\Omega \mathbb{1}_{A_i} = \alpha_i \mu(A_i) \geqslant 0. \tag{27.60}$$

Vu cet encadrement par zéro, nous avons $\alpha_i \mu(A_i) = 0$ et donc $\mu(A_i) = 0$ pour tout i.

Nous en déduisons, par union dénombrable de parties de mesures nulles, que $\mathrm{Re}(f)^+$ est non nulle seulement sur une partie de mesure nulle.

Le même raisonnement pour $\mathrm{Re}(f)^-$, $\mathrm{Im}(f)^+$ et $\mathrm{Im}(f)^-$ donne que f est non nulle sur une partie de mesure nulle. Donc $f = 0$ au sens des classes dans L^p. □

27.3.3 L'espace L^∞

SUBSECooYFJTooBqrLXv

Il n'est pas possible de définir le supremum d'une fonction définie à ensemble de mesure nulle près parce que toute classe contient des fonctions qui peuvent être arbitrairement grandes en n'importe que point. Nous cherchons alors à définir une notion de supremum qui ne tient pas compte des ensembles de mesure nulle.

DEFooIQOOooLpJBqi

Définition 27.22.
*Soit $f\colon \Omega \to \mathbb{C}$. Un nombre M est un **majorant essentiel** de f si*

$$\mu\big(|f(x)| \geqslant M\big) = 0. \tag{27.61}$$

Nous posons alors

$$N_\infty(f) = \inf\{M \text{ tel que } |f(x)| \leqslant M \text{ presque partout}\}. \tag{27.62}$$

Cela revient à prendre le supremum à ensemble de mesure nulle près.

DEFooXUKHooXYrlYq

Définition 27.23.
Nous définissons alors les espaces de Lebesgue correspondants :

$$\mathcal{L}^\infty(\Omega) = \{f\colon \Omega \to \mathbb{C} \text{ tel que } N_\infty(f) < \infty\}, \tag{27.63}$$

et L^∞ en est le quotient usuel.

27.24.
Le point sur les notations telles qu'elles devraient être respectées :

(1) Si f est une fonction, $N_\infty(f)$ est son supremum essentiel.
(2) Si f est une fonction, $\|f\|_\infty$ est sa norme supremum. Ce n'est pas la même chose que $N_\infty(f)$.
(3) Si $[f]$ est une classe de fonctions (pour l'égalité presque partout), alors $\|[f]\|_{L^\infty}$ est la norme de cette classe dans L^∞, c'est-à-dire $\|[f]\|_{L^\infty} = N_\infty(f)$ où f est un représentant.

Le point sur les abus tolérables :

(1) Si f est une fonction, on peut écrire $\|f\|_{L^\infty}$ pour $N_\infty(f)$. Attention toutefois que $N_\infty(f)$ peut valoir ∞, alors que les éléments de L^∞ sont sélectionnés pour être les classes des fonctions telles que $N_\infty(f) < \infty$. Donc il est parfois possible, lorsque f est une fonction, de parler de $\|f\|_{L^\infty}$ alors que $[f]$ n'est pas un élément de L^∞.
(2) Si $[f]$ est une classe, on peut écrire $N_\infty([f])$ pour $\|[f]\|_{L^\infty}$.
(3) Noter f la classe de la fonction f. Attention qu'alors, écrire $\|f\|_\infty$ n'a pas de sens.

Le point sur les abus intolérables :

(1) Si f est une fonction, noter $\|f\|_\infty$ pour $N_\infty(f)$.
(2) Si $[f]$ est une classe de fonctions, noter $\|[f]\|_\infty$ pour $\|[f]\|_{L^\infty}$.

27.25.
Tout ceci pour dire que si f_k sont des fonctions et si f est une fonction, nous avons la convergence

$$f_k \xrightarrow{L^\infty} f \tag{27.64}$$

si et seulement si $N_\infty(f_k - f) \to 0$. Cette convergence (qui se sert des abus tolérables de notations) est équivalente à la convergence

$$[f_k] \xrightarrow{L^\infty} [f]. \tag{27.65}$$

Mais ce n'est pas du tout équivalent à $\|f_k - f\|_\infty \to 0$. Tout au plus, il est vrai que si $\alpha \in L^\infty$ (donc α est une classe), il existe un représentant $f \in \alpha$ tel que $\|f\|_\infty = \|\alpha\|_{L^\infty}$.

Lemme 27.26 ([1])**.**
Si f_k est une suite de fonctions telle que $\|f_k - f\|_\infty \to 0$, alors

$$f_k \xrightarrow{L^\infty} f \tag{27.66}$$

où la convergence signifie $N_\infty(f_k - f) \to 0$.

Démonstration. Si g est une fonction nous avons toujours $N_\infty(g) \leqslant \|g\|_\infty$. Donc

$$N_\infty(f_k - f) \leqslant \|f_k - f\|_\infty \to 0. \tag{27.67}$$

□

27.27.
Attention toutefois que ce lemme ne signifie pas que si $\|f_k - f\|_\infty \to 0$, alors $[f_k] \xrightarrow{L^\infty} [f]$ parce que nous pourrions avoir $N_\infty(f_k) = \infty$ et alors $[f_k]$ n'est pas un élément de L^∞.

Cela arrive par exemple pour $f_k(x) = x$ et $f(x) = x$. Nous avons $\|f_k - f\|_\infty = 0$ pour tout k, alors que ni f_k ni f ne donnent lieu à un élément de L^∞.

Exemple 27.28.
Même si g est bornée, nous n'avons pas spécialement $\|g\|_\infty = N_\infty(g)$. Par exemple

$$\begin{aligned} g\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon .} \end{cases} \end{aligned} \tag{27.68}$$

Cette fonction vérifie $\|g\|_\infty = 1$ mais $N_\infty(g) = 0$. △

27.3.4 Quelques identifications

Il est intuitivement clair que ce qui peut arriver à une fonction en un seul point ne va pas influencer la fonction lorsqu'elle est vue dans L^p. En tout cas lorsqu'on considère des mesures pour lesquelles les singletons sont de mesure nulle, et c'est bien le cas de la mesure de Lebesgue. Il est peut-être intuitivement moins clair que l'on peut non seulement modifier le comportement d'une fonction en un point, mais également modifier l'ensemble de base. En voici un exemple.

Proposition 27.29.
Nous avons les égalités suivantes d'espaces

$$L^p\big(]0, 2\pi[\big) = L^p\big([0, 2\pi]\big) = L^p(S^1) \tag{27.69}$$

au sens où il existe des bijections isométriques de l'un à l'autre. Ici nous sous-entendons la mesure de Lebesgue partout[11].

Démonstration. Voici une application où le crochet dénote la prise de classe :

$$\begin{aligned} \psi\colon L^p\big(]0, 2\pi[\big) &\to L^p\big([0, 2\pi]\big) \\ [f] &\mapsto \text{la classe de } f_e(x) = \begin{cases} f(x) & \text{si } x \in]0, 2\pi[\\ 0 & \text{si } x = 0 \text{ ou } x = 2\pi. \end{cases} \end{aligned} \tag{27.70}$$

(1) **Bien définie** Si $[f] = [g]$ dans $L^p\big(]0, 2\pi[\big)$ alors $f_e(x) = g_e(x)$ pour tout $x \in]0, 2\pi[$ sauf une partie de mesure nulle. L'union de cette partie avec $\{0, 2\pi\}$ est encore de mesure nulle dans $[0, 2\pi]$. Les images par ψ sont donc égales dans $L^p\big([0, 2\pi]\big)$.

(2) **Injective** Supposons que $[f], [g] \in L^p\big(]0, 2\pi[\big)$ soient tels que $\psi\big([f]\big) = \psi\big([g]\big)$. En particulier $[f_e] = [g_e]$ et donc $f_e = g_e$ presque partout sur $[0, 2\pi]$. Vu que $\{0\} \cup \{2\pi\}$ est de mesure nulle, nous avons $f_e = g_e$ presque partout sur $]0, 2\pi[$. Mais pour $x \in]0, 2\pi[$, nous avons $f_e(x) = f(x)$ et $g_e(x) = g(x)$. Bref nous avons $f(x) = g(x)$ pour presque tout $x \in]0, 2\pi[$.

(3) **Surjective** Un élément de $L^p\big([0, 2\pi]\big)$ est l'image de sa restriction ... ou plutôt l'image de la classe de la restriction d'un quelconque de ses représentants.

11. Vu que la mesure de Lebesgue est définie pour $\mathbb{R}^d$ munie de sa tribu des boréliens (complétée), vous êtes en droit de vous demander quelle est la tribu et la mesure que nous considérons sur le cercle S^1.

(4) **Isométrie** L'intégrale qui donne la norme sur L^p ne change pas selon que nous ajoutions ou non les bornes au domaine d'intégration.

De la même manière nous avons

$$L^p\big([0,2\pi[\big) = L^p\big([0,2\pi]\big). \tag{27.71}$$

En ce qui concerne l'identification avec $L^p(S^1)$, il faut passer par l'isométrie $\varphi\colon [0,2\pi[\to S^1$ donnée par $\varphi(t) = e^{it}$, et être heureux que ce soit bien une isométrie parce qu'il faudra l'utiliser pour un changement de variables pour montrer que

$$\int_0^{2\pi} f(t)dt = \int_{S^1} (f\circ\varphi^{-1})(z)dz. \tag{27.72}$$

□

27.3.5 Inégalité de Young, Jensen, Hölder et de Minkowski

PROPooCQUBooCvtMSi

Proposition 27.30 (Inégalité de Young[659])**.**
Soient $a,b \geqslant 0$ ainsi que $p,q>0$ tels que

$$\frac{1}{p}+\frac{1}{q}=1. \tag{27.73}$$

Alors

$$ab \leqslant \frac{a^p}{p}+\frac{b^q}{q}. \tag{27.74}$$

Il y a égalité si et seulement si $a^p = b^q$.

PropXISooBxdaLk

Proposition 27.31 (Inégalité de Jensen[468])**.**
Soit un espace mesuré de probabilité[12] *$(\Omega,\mathcal{A},\mu)$ ainsi qu'une fonction convexe $f\colon\mathbb{R}\to\mathbb{R}$ et une application $\alpha\colon\Omega\to\mathbb{R}$ tels que α et $f\circ\alpha$ soient intégrables sur Ω. Alors*

$$f\Big(\int_\Omega \alpha\,d\mu\Big) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{27.75}$$

Démonstration. Soient $a\in\mathbb{R}$ et c_a le nombre donné par la proposition 17.90 : pour tout $\omega\in\Omega$ nous avons

$$f\big(\alpha(\omega)\big) - f(a) \geqslant c_a\big(\alpha(\omega)-a\big). \tag{27.76}$$

EqOUMooIvknIP

Cela est en particulier vrai pour $a = \int_\Omega \alpha\,d\mu$. Nous intégrons l'inégalité (27.76) sur Ω en nous souvenant que $\int d\mu = 1$:

$$\int_\Omega (f\circ\alpha)d\mu - \int_\Omega f(a)d\mu \geqslant c_a\Big(\int_\Omega \alpha - \int_\Omega a\Big) \tag{27.77a}$$

$$\int_\Omega (f\circ\alpha)d\mu - f(a) \geqslant c_a(a-a) \tag{27.77b}$$

$$f(a) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{27.77c}$$

Cette dernière inégalité est celle que nous devions prouver. □

Corolaire 27.32.
Soit un espace mesuré de probabilité $(\Omega,\mathcal{A},\mu)$ et une application $\alpha\in L^1(\Omega,\mu)$ et $\alpha\in L^p(\Omega)$ avec $1\leqslant p<+\infty$. Alors

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big| \leqslant \|\alpha\|_p. \tag{27.78}$$

12. C'est-à-dire que $\int_\Omega d\mu = 1$.

Démonstration. Il suffit d'utiliser l'inégalité de Jensen sur la fonction convexe $f(x) = |x|^p$. Nous avons alors

$$|\int_\Omega \alpha(s)d\mu(s)|^p \leqslant \int_\Omega |\alpha(s)|^p d\mu(s), \tag{27.79}$$

c'est-à-dire

$$|\int_\Omega \alpha(s)d\mu(s)| \leqslant \left[\int_\Omega |\alpha(s)|^p d\mu(s)\right]^{1/p} = \|\alpha\|_p \tag{27.80}$$

où ma norme $\|.\|_p$ est prise au sens de la mesure μ. □

ProptYqspT

Proposition 27.33 (Inégalité de Hölder[660])**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $1 \leqslant p, q \leqslant \infty$ satisfaisant $\frac{1}{p} + \frac{1}{q} = 1$. Si $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, alors nous avons les choses suivantes :

ITEMooNDKPooRKdmgS

(1) Le produit fg est dans $L^1(\Omega)$ et nous avons

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{27.81}$$

EqLPKooPBCQYN
ITEMooQHLPooRWWMOP

(2) Si $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ alors

$$\|fg\|_r \leqslant \|f\|_p \|g\|_q \tag{27.82}$$

EqAVZooFNyzmT
ITEMooBOYJooRkiAqJ

(3) Nous avons $\|fg\|_1 = \|f\|_p\|g\|_q$ si et seulement si nous sommes dans un des trois cas suivants :
- *$f = 0$ presque partout,*
- *$g = 0$ presque partout,*
- *Il existe $\lambda > 0$ tel que $|f|^p = \lambda|g|^q$ presque partout.*

Ce qui est appelé « inégalité de Hölder » est généralement l'inéquation (27.81).

Démonstration. En plusieurs points.

(1) **Pour (1)** Nous allons le voir comme cas particulier de (2).

(2) **Pour (2)** D'abord nous supposons $\|g\|_q = 1$ et nous posons

$$A = \{x \in \Omega \text{ tel que } |g(x)| > 0\}. \tag{27.83}$$

Hors de A, les intégrales que nous allons écrire sont nulles. Nous avons

$$\|fg\|_r^p = \Big|\int_A |f|^r |g|^{r-q} |g|^q\Big|^{p/r}, \tag{27.84}$$

et le coup tordu est de considérer cette intégrale comme étant une intégrale par rapport à la mesure $\nu = |g|^q d\mu$ qui a la propriété d'être une mesure de probabilité par hypothèse sur g. Nous pouvons alors utiliser l'inégalité de Jensen [13] parce que $p/r > 1$, ce qui fait de $x \mapsto |x|^{p/r}$ une fonction convexe. Nous avons alors

$$\|fg\|_r^p \leqslant \int_A \left(|f|^r |g|^{r-q}\right)^{p/r} |g|^q d\mu \tag{27.85a}$$

$$= \int_A |f|^p |g|^{p(r-q)/r} |g|^q d\mu \tag{27.85b}$$

La puissance de $|g|$ dans cette expression est : $q + \frac{p(r-q)}{r} = 0$ parce que $p(q-r) = rq$. Nous avons alors montré que

$$\|fg\|_r^p \leqslant \int_A |f|^p d\mu \leqslant \|f\|_p^p. \tag{27.86}$$

La dernière inégalité est le fait que le domaine A n'est pas tout le domaine Ω.
Si maintenant $\|g\|_q \neq 1$ alors nous calculons

$$\|fg\|_r = \|g\|_q \|f \frac{g}{\|g\|_q}\|_r \leqslant \|g\|_q \|f\|_p \tag{27.87}$$

en appliquant la première partie à la fonction $\frac{g}{\|g\|_q}$ qui est de norme 1.

13. Proposition 27.31.

(3) **Pour (3)** Si f ou g est nulle presque partout, il y a immédiatement égalité. Nous supposons donc que f et g ne sont pas nulles presque partout et donc que $\|f\|_p$ et $\|g\|_q$ sont non nuls.

(3a) **Deux fonctions intermédiaires** Nous posons

$$\hat{f} = \frac{|f|}{\|f\|_p}, \quad \hat{g} = \frac{|g|}{\|g\|_q}. \tag{27.88}$$

(3b) **Égalité préliminaire** Nous avons

$$\frac{1}{p}\int \hat{f}^p = \frac{1}{p}\int \frac{|f|^p}{\|f\|_p^p} = \frac{1}{p\|f\|_p^p}\int |f|^p = \frac{1}{p}, \tag{27.89}$$

et de même avec g et q au lieu de f et p. Nous avons donc

$$\frac{1}{p}\int \hat{f}^p + \frac{1}{q}\int \hat{g}^q = 1. \tag{27.90}$$

(3c) **Les équivalences** Les choses suivantes sont équivalentes.

(3c.iii) $\int |fg| = \|f\|_p\|g\|_q$

(3c.iii) $\int \hat{f}\hat{g} = 1$

(3c.iii)

$$\int \hat{f}\hat{g} = \frac{1}{p}\int \hat{f}^p + \frac{1}{q}\int \hat{g}^q \tag{27.91}$$

(3c.iii)

$$\int \left(\hat{g}\hat{f} - \frac{1}{p}\hat{f}^p - \frac{1}{q}\hat{g}^q\right) = 0. \tag{27.92}$$

(3c.iii)

$$\hat{g}\hat{f} - \frac{1}{p}\hat{f}^p - \frac{1}{q}\hat{g}^q = 0 \tag{27.93}$$

presque partout.

En effet l'inégalité de Young [14] dit que l'intégrante est positive partout. Pour que l'intégrale soit nulle, il faut que l'intégrante soit nulle ; c'est le lemme 14.199.

(3c.iii) $\hat{f}(x)^p = \hat{g}(x)^q$ pour presque tout x. C'est le cas d'égalité dans l'inégalité de Young.

(3c.iii)

$$\frac{|f|^p}{\|f\|_p^p} = \frac{|g|^q}{\|g\|_q^q}. \tag{27.94}$$

(3c.iii)

$$|f|^p = \lambda|g|^q \tag{27.95}$$

avec $\lambda = \|f\|_p^p/\|g\|_q^q$.

(3d) **Conclusion** En lisant les implications de haut en bas, nous avons la condition nécessaire au cas d'égalité. Pour traiter la condition suffisante, nous supposons qu'il existe $\lambda > 0$ tel que $|f|^p = \lambda|g|^q$. Alors nous avons

$$\|f\|_p^p = \int |f|^p = \lambda\int |g|^q = \lambda\|g\|_q^q, \tag{27.96}$$

ce qui donne immédiatement $\lambda = \|f\|_p^p/\|g\|_q^q$. Nous pouvons donc remonter les équivalences.

□

14. Proposition 27.30.

RemNormuptNird

Remarque 27.34.
Dans le cas d'un espace de probabilité, la fonction constante $g = 1$ appartient à $L^p(\Omega)$. En prenant $p = q = 2$ nous obtenons

$$\|f\|_1 \leqslant \|f\|_2. \tag{27.97}$$

PROPooDVCCooKqbAwo

Proposition 27.35.
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ vérifiant $\mu(\Omega) = 1$. Soient des nombres $0 < r \leqslant s$ ainsi qu'une application $f : \Omega \to \mathbb{R}$.

Si $\int_\Omega |f|^2 < \infty$, alors $\int_\Omega |f|^r < \infty$.

Démonstration. Si $s = r$, la conclusions est immédiate. Nous supposons donc $r \neq s$. Nous considérons q tel que $\frac{1}{s} + \frac{1}{q} = \frac{1}{r}$, c'est-à-dire $q = rs/(s - r)$. Vu que $s > r$, nous avons $q > 0$, pas de problèmer.

Par hypothèse, nous avons $f \in L^s(\Omega)$. Et comme Ω est de mesure finie, la fonction constante 1_Ω est dans $L^q(\Omega)$. Nous povons donc écrire l'inégalité de Hölder 27.82 :

$$\|f 1_\Omega\|_r \leqslant \|f\|_s \|1_\Omega\|_q. \tag{27.98}$$

Cela donne $\|f\|_r \leqslant \|f\|_s$, et en déballant,

$$\left(\int_\Omega |f|^r \right)^{1/r} \leqslant \left(\int_\Omega |f|^s \right)^{1/s}. \tag{27.99}$$

Vu que $r > 0$, l'application $t \mapsto t^r$ est croissante, et nous pouvons écrire

$$\int_\Omega |f|^r \leqslant \left(\int_\Omega |f|^s \right)^{r/s} < \infty. \tag{27.100}$$

Notez l'utilisation de la proposition 12.408. □

LemTLHwYzD

Lemme 27.36.
Lorsque I est borné nous avons $L^2(I) \subset L^1(I)$. Si I n'est pas borné alors $L^2(I) \subset L^1_{loc}(I)$.

Démonstration. En effet si I est borné, alors la fonction constante 1 est dans $L^2(I)$ et l'inégalité de Hölder 27.33 nous dit que le produit $1u$ est dans $L^1(I)$.

Si I n'est pas borné, nous refaisons le même raisonnement sur un compact K de I. □

CORooIIEAooNmbkTo

Corolaire 27.37 ([661]).
Soit l'espace $L^2(I)$ avec $I =]0, 1[$ avec la mesure de Lebesgue. Si $u_n \in L^2$ converge vers u dans L^2 alors nous pouvons permuter l'intégrale et la limite :

$$\lim_{n\to\infty} \int_I u_n = \int_I u. \tag{27.101}$$

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} T : L^2(I) &\to \mathbb{R} \\ u &\mapsto \int_I u. \end{aligned} \tag{27.102}$$

Elle est bien définie par l'inégalité de Hölder $\|fg\|_1 \leqslant \|f\|_2 \|g\|_2$ appliqué à $g(x) = 1$ qui vérifie $\|g\|_2 = 1$. Nous avons aussi

$$T(u) \leqslant \int_I |u| \leqslant \|u\|_1 \leqslant \|u\|_2 \tag{27.103}$$

où la dernière inégalité est celle de Hölder 27.33. Bref, T est continue. Cela signifie que si $u_n \xrightarrow{L^2(I)} u$ alors $T(u_n) = T(u)$. Cela est l'égalité demandée. □

¡¡ Avertissement/question au lecteur !! 27.38
Dans la proposition suivante, la partie « égalité » est très personnelle. Je n'en n'ai pas trouvé de preuve complète. Donc soyez doublement vigilant et écrivez-moi si vous avez quelque chose à dire.
PropInegMinkKUpRHg

Proposition 27.39 (Inégalité de Minkowski[662, 663, 664, 665, 666, 1]).
Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors
ItemDHukLJi

(1) $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$
ITEMooGRXBooMLRMww

(2) *Si $p = 1$, il y a égalité si et seulement si $f(x)\overline{g(x)} \geqslant 0$ pour presque tout x.*
ITEMooQCSHooNUDwtM

(3) *Si $p > 1$, il y a égalité si et seulement si il existe des réels positifs α, β pas tous deux nuls tels que*

$$\alpha f(x) = \beta g(x) \tag{27.104}$$

pour presque tout x.

Démonstration. En plusieurs points.

(1) **Pour (1)** En utilisant l'inégalité $|z_1 + z_2| \leqslant |z_1| + |z_2|$ (proposition 9.102(7)) pour chaque x dans l'intégrale, de façon à pouvoir majorer

$$|f(x)+g(x)|^p = |f(x)+g(x)||f(x)+g(x)|^{p-1} \tag{27.105a}$$
$$\leqslant \big(|f(x)|+|g(x)|\big)|f(x)+g(x)|^{p-1}. \tag{27.105b}$$

Nous mettons ça dans une intégrale et nous calculons un peu :
SUBEQSooGWMTooDBXSgL

$$\|f+g\|_p^p = \int |f+g|^p d\mu \tag{27.106a}$$
$$\leqslant \int \big(|f|+|g|\big)|f+g|^{p-1} d\mu \tag{27.106b}$$
SUBEQooYCTWooQpHNqa
$$= \int |f||f+g|^{p-1} + \int |g||f+g|^{p-1} \tag{27.106c}$$
SUBEQooNQKBooQFtTIJ

Lorsque $p = 1$, nous nous arrêtons ici parce que (27.106c) s'écrit

$$\|f+g\|_1 \leqslant \int |f| + \int |g| = \|f\|_1 + \|g\|_1. \tag{27.107}$$

Lorsque $p > 1$, nous devons continuer et utiliser Hölder. Attardons nous sur le premier terme. Nous posons $q = p/(p-1)$, $a = f$ et $b = |f+g|^{p-1}$, et nous utilisons l'inégalité de Hölder [15] :
SUBEQSooFINUooQfIdMS

$$\int |f||f+g|^{p-1} = \|ab\|_1 \tag{27.108a}$$
$$\leqslant \|a\|_p \|b\|_q \tag{27.108b}$$
$$= \|f\|_p \left[\int \big(|f+g|^{p-1}\big)^q \right]^{1/q} \tag{27.108c}$$
$$= \|f\|_p \left(\int |f+g|^p \right)^{1-\frac{1}{p}} \tag{27.108d}$$
$$= \|f\|_p \left(\int |f+g|^p \right)^{(p-1)/p} \tag{27.108e}$$
$$= \|f\|_p \|f+g\|_p^{p-1} \tag{27.108f}$$

Nous avons utilisé la règle de produit d'exposants et de somme d'exposants [16]

15. Proposition 27.33.
16. Propositions 12.408 et 12.398(3).

Nous utilisons cette inégalité dans les deux termes de (27.106) : SUBEQSooKHKOooOCfndf

$$\|f+g\|_p^p \leqslant \int |f||f+g|^{p-1} + \int |g||f+g|^{p-1} \tag{27.109a}$$
$$\leqslant \|f\|_p \|f+g\|_p^{p-1} + \|g\|_p \|f+g\|^{p-1} \tag{27.109b}$$

Nous obtenons le résultat en divisant le tout par $\|f+g\|_p^{p-1}$.

(2) **Pour (2)** Toutes les inégalités de (27.106) sont des égalités. En particulier nous avons celle-ci :

$$\int |f+g| d\mu = \int (|f|+|g|) d\mu \tag{27.110}$$

Cela donne

$$\int \big(|f+g| - |f| - |g|\big) d\mu = 0. \tag{27.111}$$

Vue l'inégalité de la proposition 9.102(7), la fonction intégrée est toujours négative. Pour que l'intégrale soit nulle, il faut que la fonction intégrée soit presque partout nulle :

$$|f(x)+g(x)| = |f(x)| + |g(x)| \tag{27.112}$$

pour presque tout x. La partie « égalité » de la proposition 9.102(7) donne alors le résultat.

———— Cas d'égalité, préliminaire ————

Nous commençons par prouver que si $f+g=0$ alors $f=g=0$. Pour cela nous faisons le calcul SUBEQSooGHMTooWSTtOn

$$\|f\|_p + \|g\|_p = \|f+g\| \tag{27.113a}$$ SUEQooPUQGooAJaFFa
$$\leqslant \big\||f|+|g|\big\| \tag{27.113b}$$ SUBEQooFXHBooYXMhDf
$$\leqslant \big\||f|\big\|_p + \big\||g|\big\| \tag{27.113c}$$ SUBEQooQZFXooPVgQrI
$$= \|f\|_p + \|g\|_p. \tag{27.113d}$$

Justifications :

— Pour (27.113a). L'hypothèse d'égalité dans l'inégalité de Minkowski.

— Pour (27.113b). Intégrale de l'inégalité de la proposition 9.102(6).

— Pour (27.113c). Inégalité de Minkowski.

Étant donné que le premier et le dernier membre de (27.113) sont égaux, toutes les inégalités sont des égalités. En particulier,

$$\|f+g\|_p = \big\||f|+|g|\big\|. \tag{27.114}$$

En termes d'intégrales, cette égalité donne

$$\int |f+g|^p - \big(|f|+|g|\big)^p = 0. \tag{27.115}$$ EQooRYQBooGdgtLN

L'inégalité déjà mentionnée $|z_1+z_2|^p \leqslant \big(|z_1|+|z_2|\big)^p$ dit que la fonction à intégrer dans (27.115) est partout négative. Pour que l'intégrale soit nulle, il faut donc que la fonction soit presque partout nulle. Autrement dit, pour presque tout x nous avons

$$|f(x)+g(x)| = |f(x)| + |g(x)|. \tag{27.116}$$ EQooULHAooWXViEH

Cette égalité nous sera utile dans la suite. En particulier, retenez que si $f+g=0$ presque partout, alors $f=g=0$ presque partout. Dans ce cas, l'égalité $\alpha f + \beta g = 0$ fonctionne pour tout α et β.

———— Cas d'égalité (2), $p=1$. ————

(1) $\Rightarrow$

L'inégalité (27.106b) est une égalité :

$$\int |f+g|^{p-1}\big(|f+g|-|f|-|g|\big)=0. \tag{27.117}$$

Par la proposition 9.102(6), la fonction intégrée est presque partout nulle ou positive. Pour que l'intégrale soit nulle, il faut que la fonction soit presque partout nulle (lemme 14.199). Presque chaque x vérifie sont une des deux conditions suivantes :

(1a) $|f(x)+g(x)|=0$

(1b) $|f(x)+g(x)|-|f(x)|-|g(x)|=0$.

Pour les x vérifiant la première condition, nous avons déjà vu que $f=g=0$, juste en dessous de (27.116). En particulier, nous avons $f(x)\overline{g(x)}\geqslant 0$ pour tout x dans ce cas.

En ce qui concerne les x vérifiant la seconde condition, la proposition 9.102(7) montre qu'ils vérifient $f(x)\overline{g(x)}\geqslant 0$.

Pour les x dans le second cas, la proposition 9.102(7) donne $f(x)\overline{g(x)}\geqslant 0$.

Bref, dans tous les cas nous avons $f(x)\overline{g(x)}\geqslant 0$ pour presque tout x.

(2) $\Leftarrow$

Nous supposons que $f(x)\overline{g(x)}\geqslant 0$ pour presque tout x. La proposition 9.102(7) (utilisée dans l'autre sens) dit que $|f(x)+g(x)|-|f(x)|-|g(x)|=0$ pour presque tout x. L'inégalité (27.106b) est alors une égalité.

———— Cas d'égalité (3), $p>1$. ————

Il y a deux sens.

(1) **$\Rightarrow$, les inégalités** Nous supposons que $f,g\in L^p(\Omega,\mathcal{A},\mu)$ vérifient $\|f+g\|_p=\|f\|_p+\|g\|_p$, et nous allons voir ce que cela implique. Regroupons les différentes inégalités qui mènent à Minkowski.

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p d\mu && && (27.118a)\\
&\leqslant \int \big(|f|+|g|\big)|f+g|^{p-1}d\mu && \text{eq. (27.106)} && (27.118b)\\
&\leqslant \|f\|_p\|f+g\|_p^{p-1}+\int |g||f+g|^{p-1} && \text{eq. (27.109)} && (27.118c)\\
&\leqslant \|f\|_p\|f+g\|_p^{p-1}+\|g\|_p\|f+g\|^{p-1} && \text{eq. (27.109)} && (27.118d)\\
&= \big(\|f\|_p+\|g\|_p\big)\|f+g\|_p^{p-1} && && (27.118e)\\
&= \|f+g\|_p^p && \text{hyp. égalité.} && (27.118f)
\end{aligned}$$

La dernière ligne est l'hypothèse d'égalité $\|f\|_p+\|g\|_p=\|f+g\|_p$. Comme le premier et le dernier membre de ces inégalités sont égaux, toutes les inégalités sont des égalités.

(2) **$\Rightarrow$, première inégalité**

L'inégalité (27.118b) est une égalité :

$$\int |f+g|^{p-1}\big(|f+g|-|f|-|g|\big)=0. \tag{27.119}$$

Par la proposition 9.102(6), la fonction intégrée est presque partout nulle ou positive. Pour que l'intégrale soit nulle, il faut que la fonction soit presque partout nulle (lemme 14.199). Presque chaque x vérifie sont une des deux conditions suivantes :

(2a) $|f(x)+g(x)|=0$

(2b) $|f(x)+g(x)|-|f(x)|-|g(x)|=0$.

Pour les x vérifiant la première condition, nous avons déjà vu que $\alpha f+\beta g=0$ pour n'importe quels α et β, juste en dessous de (27.116).

En ce qui concerne les x vérifiant la seconde condition, la proposition 9.102(7) montre qu'ils vérifient $f(x)\overline{g(x)}\geqslant 0$.

Nous avons donc déjà montré que si f et g vérifient une égalité, alors pour presque tout x, nous avons

$$f(x)\overline{g(x)}\geqslant 0. \tag{27.120}$$

(3) **⇒, seconde inégalité (Hölder)**

La seconde inégalité à être une égalité est (27.118c) :

$$\int |f||f+g|^{p-1}=\|f\|_p\|f+g\|_p^{p-1}. \tag{27.121}$$

Il s'agit du cas d'égalité de Hölder avec $a=f$, $b=|f+g|^{p-1}$ traité dans la proposition 27.33(3). Nous sommes donc dans un des trois cas suivants :

— $f=0$ presque partout,

— $f+g=0$ presque partout,

— il existe $\lambda>0$ tel que $|a|^p=\lambda|b|^q$.

Dans le premier cas, nous avons $\alpha f=\beta g$ avec $\alpha=1$ et $\beta=0$. Le second cas donne $f=g=0$ et donc $\alpha f+\beta g=0$ pour n'importe quels α et β (voir autour de (27.116)).

Nous supposons donc être dans le troisième cas. Étant donné que $q=p/(p-1)$, nous avons

$$|b|^q=\big(|f+g|^{p-1}\big)^q=|f+g|^p, \tag{27.122}$$

et donc la contrainte est l'existence de $\lambda>0$ tel que $|f|^p=\lambda|f+g|^p$.

(4) **⇒, troisième inégalité (Hölder)**

La seconde application de Hölder est avec $a=g$ et $b=|f+g|^{p-1}$. Nous sommes dans un cas d'égalité si nous sommes dans un des trois cas suivants :

— $g=0$

— $f+g=0$

— il existe $\sigma>0$ tel que $|a|^p=\sigma|b|^q$.

Encore une fois nous supposons être dans le troisième cas. Il existe donc $\sigma>0$ tel que $|g|^p=\sigma|f+g|^p$.

(5) **Petit résumé**

Nous avons prouvé l'existence de $\lambda,\sigma>0$ tels que

$$\begin{cases} f(x)\overline{g(x)}\geqslant 0 \quad \text{pour presque tout } x & \text{(27.123a)}\\ |f|^p=\lambda|f+g|^p & \text{(27.123b)}\\ |g|^p=\sigma|f+g|^p & \text{(27.123c)} \end{cases}$$

(6) **Des calculs et des cas** En passant à la racine p^{e} :

$$\begin{cases} f(x)\overline{g(x)}\geqslant 0 \quad \text{pour presque tout } x & \text{(27.124a)}\\ |f|=s|f+g| & \text{(27.124b)}\\ |g|=t|f+g|. & \text{(27.124c)} \end{cases}$$

Les deux dernières égalités du système permettent d'écrire

$$t|f|=st|f+g|=s|g|. \tag{27.125}$$

Nous considérons donc le système, valide pour presque tout x : SUBEQSooLKGDooQouvhW

$$\begin{cases} f(x)\overline{g(x)} \geqslant 0 & (27.126a)\\ t|f(x)| = s|g(x)| & (27.126b) \end{cases}$$

Avec la contrainte $(s,t) \neq (0,0)$.
Le lemme 9.104 appliqué à chaque couple $f(x)$, $g(x)$ donne l'existence d'une fonction réelle a telle que pour chaque x nous ayons un des deux cas suivants :

— $g(x) = 0$
— $f(x) = a(x)g(x)$.

De plus $a(x) \in \mathbb{R}^+$ parce que $f(x)\overline{g(x)}$ est non seulement réel, mais également positif.

(6a) $\underline{s = 0}$ Alors nous avons $t|f(x)| = 0$ et donc $f(x) = 0$ pour presque tout x parce que $t \neq 0$. Nous avons alors $\alpha f(x) + \beta g(x) = 0$ avec $\alpha = 1$ et $\beta = 0$.

(6b) $\underline{s \neq 0}$ Les points de Ω se séparent en deux parties : $g(x) = 0$ et les autres.

(6b.ii) $\underline{g(x) \neq 0}$ Dans ce cas nous avons le système

$$\begin{cases} f(x) = a(x)g(x) & (27.127a)\\ t|f(x)| = s|g(x)|. & (27.127b) \end{cases}$$

Vu que $s \neq 0$ et $g(x) \neq 0$, la dernière équation donne $t|f(x)| \neq 0$ et donc $t \neq 0$ et $f(x) \neq 0$.
La première équation donne $a(x) = f(x)/g(x)$ que nous pouvons mette dans la secondes pour obternir :

$$|a(x)| = \frac{s}{t}. \tag{27.128}$$

Vu que a est à valeurs dans $\mathbb{R}^+$, cela donne

$$a(x) = \frac{s}{t} \tag{27.129}$$

pour tout x tel que $g(x) \neq 0$.

(6b.ii) $\underline{g(x) = 0}$ Dans ce cas, le système (27.126) devient

$$t|f(x)| = 0. \tag{27.130}$$

Vu que $s \neq 0$, nous avons $f(x) = 0$. Donc n'importe quel choix de α et β fait l'affaire pour ces points. Il n'y a pas de contraintes.

(6c) $\underline{\Leftarrow}$ Nous supposons $\alpha f(x) = \beta g(x)$ pour presque tout x. Pour fixer les idées, nous supposons que $\beta \neq 0$ (sinon, refaire le raisonnement en inversant les rôles de f et g). En posant $\lambda = \alpha/\beta$ nous avons

$$g(x) = \lambda f(x) \tag{27.131}$$

pour presque tout x. En passant par l'intégrale,

$$\|f + g\|_p = \|(1+\lambda)f\|_p = (1+\lambda)\|f\|_p = \|f\|_p + \lambda\|f\|_p = \|f\|_p + \|\lambda f\|_p = \|f\|_p + \|g\|_p. \tag{27.132}$$

Le λ et $(1+\lambda)$ rentrent et sortent des normes parce qu'ils sont positifs.

□

Dans le cas où nous n'avons pas une somme de deux fonctions mais d'une infinité paramétrée par $y \in Y$, nous pouvons convertir la somme de l'inégalité de Minkowski en une intégrale :

$$\left[\int_X \Big(\int_Y |f(x,y)|d\nu(y)\Big)^p d\mu(x)\right]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y). \tag{27.133}$$

Ce sera la proposition 27.41.

LEMooPSBWooGLggTe

Lemme 27.40 ([662]).
Si (X,μ) est un espace mesuré σ-fini non nul[17]*, il existe une fonction mesurable strictement positive $k\colon X\to]0,\infty[$ telle que*

$$\int_X k d\mu = 1. \tag{27.134}$$

PROPooGZJZooXfZdqn

Proposition 27.41 (Forme intégrale de Minkowski[662]).
Soient deux espaces σ-finis (X,μ) et (Y,ν). Si $f\colon X\times Y\to\mathbb{C}$ est mesurable[18] *et si $p\geqslant 1$, alors*

$$\Big[\int_X \Big(\int_Y |f(x,y)|d\nu(y)\Big)^p d\mu(x)\Big]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y). \tag{27.135}$$

EQooAEXWooYJtGGR

De façon plus compacte :

$$\Big\| x\mapsto \int_Y f(x,y)d\nu(y)\Big\|_p \leqslant \int_Y \|f_y\|_p d\nu(y) \tag{27.136}$$

où $f_y(x)=f(x,y)$.

Démonstration. En plusieurs parties.

(1) **La fonction M** Pour toute application mesurable $g\colon X\times Y\to\mathbb{C}$ nous posons

$$M(g)=\int_Y\Big[\int_X |g(x,y)|^p d\mu(x)\Big]^{1/p} d\nu(y). \tag{27.137}$$

Cette fonction a une propriété d'homogénéité : $M(\lambda g)=\lambda M(g)$ pour tout $\lambda\in[0,\infty[$.

(2) **Si $M(g)<1$** Nous considérons une fonction mesurable $g\colon X\times Y\to\mathbb{C}$ telle que $M(g)<1$. Soit $k\colon Y\to]0,\infty[$ une fonction donnée par le lemme 27.40, c'est-à-dire telle que $\int_Y k=1$. Nous posons

$$\begin{aligned} s\colon Y&\to[0,\infty]\\ y&\mapsto\Big(\int_X |g(x,y)|^p d\mu(x)\Big)^{1/p}, \end{aligned} \tag{27.138}$$

ainsi que

$$h(y)=s(y)+\big(1-M(g)\big)k(y). \tag{27.139}$$

Vu que $M(g)<1$ et que $k>0$, nous avons l'inégalité stricte $h(y)>s(y)\geqslant 0$ pour tout $y\in Y$. Nous avons de plus

$$\int_Y h(y)d\nu(y)=\underbrace{\int_Y s}_{=M(g)}+\big(1-M(g)\big)\underbrace{\int_Y k}_{=1}=1. \tag{27.140}$$

Nous considérons maintenant la mesure $\rho=h\nu$ (produite d'une mesure par une fonction, définition 14.212). Nous venons de montrer que c'est une mesure de probablilité. En posant $\alpha(x)=|x|^p$, nous pouvons faire quelques calculs avec les justifications en-dessous :

SUBEQSooPPZIooSEDcpS

$$\begin{aligned} \alpha\Big(\int_Y |g(x,y)|d\nu(y)\Big)&=\alpha\Big(\int_Y h(y)^{-1}|g(x,y)|h(y)d\nu(y)\Big) && \text{(27.141a)}\\ &=\alpha\left(\int_Y [|g(x,y)|h(y)^{-1}]d\rho(y)\right) && \text{(27.141b)}\\ &\leqslant\int_Y \alpha\big(|g(x,y)|h(y)^{-1}\big)d\rho(y) && \text{(27.141c)}\\ &=\int_Y |g(x,y)|^p|h(y)|^{1-p}d\nu(y) && \text{(27.141d)}\\ & && \text{(27.141e)} \end{aligned}$$

SUBEQooDAWRooOOfQBE

SUBEQooLDFHopJelfxU

17. C'est-à-dire que μ n'est pas la mesure identiquement nulle.
18. Pour l'espace mesuré produit $X\times Y,\mu\otimes\nu$.

Justifications.

— Pour (27.141c). Inégalité de Jensen, proposition 27.31

— Pour (27.141d). La fonction h est à valeurs positives ; donc $h = |h|$ et nous pouvons coller le h qui ressort de ρ avec le $|h^{-1}|^p$.

Nous continuons en intégrant l'inégalité (27.141) sur X :

$$\int_X \left(\int_Y |g(x,y)| d\nu(y) \right)^p d\mu(x) \leqslant \int_X \left[\int_Y |g(x,y)|^p h(y)^{1-p} d\nu(y) \right] d\mu(x) \tag{27.142a}$$

$$= \int_Y \int_X \left[|g(x,y)|^p h(y)^{1-p} d\mu(x) \right] d\nu(y) \tag{27.142b}$$

$$= \int_Y \left(\int_X \left[|g(x,y)|^p d\mu(x) \right] \right) h(y)^{1-p} d\nu(y) \tag{27.142c}$$

$$= \int_Y s(y)^p h(y)^{1-p} d\nu(y) \tag{27.142d}$$

$$\leqslant \int_Y h(y) d\nu(y) \tag{27.142e}$$

$$= 1. \tag{27.142f}$$

Justifications.

— Pour (27.142b). Les intégrales sont permutées grâce au théorème de Fibini-Tonelli 14.301.

— Pour (27.142e). Nous savons que $h > s \geqslant 0$, donc $s(y)^p / h(y)^p < 1$.

Au final nous avons prouvé que si $M(g) < 1$, alors nous avons

$$\int_X \left(\int_Y |g(x,y)| d\nu(y) \right)^p d\mu(x) \leqslant 1. \tag{27.143}$$

(3) **La preuve proprement dite** Nous considérons une fonction mesurable $f : X \times Y \to \mathbb{C}$ comme dans les hypothèses. Astuce : considérons $0 < t < 1$ et appliquons l'inégalité (27.143) à $g = \frac{t}{M(f)} f$. Cela est possible parce que

$$M(g) = \frac{t}{M(f)} M(f) = t < 1. \tag{27.144}$$

Nous avons :

$$\int_X \left(\int_Y \left| \frac{t}{M(f)} \right| |f(x,y)|^p d\nu(y) \right)^p d\mu(x) \leqslant 1, \tag{27.145}$$

et donc, vu que $M(f)$ et t sont positifs,

$$\int_X \left(\int_Y |f(x,y)|^p d\nu(y) \right)^p d\mu(x) \leqslant \frac{M(f)}{t}. \tag{27.146}$$

Cette inégalité est valable pour tout $t < 1$ et, ô incroyable !, ce t n'est qu'à droite. Nous pouvons prendre la limite $t \to 1$ dans cette inégalité [19] :

$$\int_X \left(\int_Y |f(x,y)|^p d\nu(y) \right)^p d\mu(x) \leqslant M(f). \tag{27.147}$$

Et ça, c'est ce que nous devons démontrer.

□

19. Nous ne pouvons pas simplement poursuivre les inégalités en majorant t par 1 parce que le t est au dénominateur : l'inégalité irait dans le mauvais sens.

27.3.6 Ni inclusions ni inégalités

Aucun espace $L^p(\mathbb{R})$ n'est inclus dans aucun autre ni aucune norme n'est plus grande qu'une autre (sur les intersections). Nous verrons cependant en la proposition 27.42 que de telles inclusions et inégalités sont possibles pour $L^p\big([0,1]\big)$.

Nous allons donner des exemples de tout ça en supposant $p<q$ et en nous appuyant lourdement sur les intégrales de $\frac{1}{x^\alpha}$ étudiées par la proposition 14.300.

(1) $\underline{L^p \not\subset L^q}$
La fonction

$$f(x)=\begin{cases}\frac{1}{x^{1/q}} & \text{si } 0<x<1\\ 0 & \text{sinon}\end{cases} \tag{27.148}$$

est dans L^p mais pas dans L^q. En effet

$$\|f\|_p^p=\int_0^1\frac{1}{x^{p/q}}dx<\infty \tag{27.149}$$

parce que $p<q$ et $p/q<1$. Par contre

$$\|f\|_q^q=\int_0^1\frac{1}{x}dx=\infty. \tag{27.150}$$

(2) $\underline{L^q \not\subset L^p}$
La fonction

$$f(x)=\begin{cases}\frac{1}{x^{1/p}} & \text{si } x>1\\ 0 & \text{sinon}\end{cases} \tag{27.151}$$

est dans L^q mais pas dans L^p. En effet

$$\|f\|_p^p=\int_1^\infty\frac{1}{x}=\infty \tag{27.152}$$

alors que

$$\|f\|_q^q=\int_1^\infty\frac{1}{x^{q/p}}dx<\infty. \tag{27.153}$$

(3) $\underline{\textbf{Exemple de } \|f\|_p>\|f\|_q}$
La fonction

$$f(x)=\begin{cases}1 & \text{si } x\in[0,2]\\ 0 & \text{sinon.}\end{cases} \tag{27.154}$$

Nous avons

$$\|f\|_p=2^{1/p} \tag{27.155a}$$
$$\|f\|_q=2^{1/q}. \tag{27.155b}$$

Mais comme $p<q$ donc $\|f\|_p>\|f\|_q$.

(4) $\underline{\textbf{Exemple de } \|f\|_p<\|f\|_q}$
La fonction

$$f(x)=\begin{cases}1 & \text{si } x\in[0,\frac{1}{2}]\\ 0 & \text{sinon.}\end{cases} \tag{27.156}$$

Alors

$$\|f\|_p=\frac{1}{2^{1/p}} \tag{27.157a}$$
$$\|f\|_q=\frac{1}{2^{1/q}} \tag{27.157b}$$

et donc $\|f\|_p<\|f\|_q$.

Ces exemples donnent un exemple de fonction f telle que $\|f\|_p < \|f\|_q$ pour tout espace $L^p(I)$ et $L^q(I)$ avec $I \subset \mathbb{R}$. Par contre l'exemple $\|f\|_p > \|f\|_q$ ne fonctionne que si la taille de I est plus grande que 1. Et pour cause : il y a des inclusions si I est borné.

PropIRDooFSWORl

Proposition 27.42 ([11]).
Inclusions et inégalités dans le cas d'un ensemble de mesure finie.

(1) Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré fini et $1 \leqslant p \leqslant +\infty$. Alors $L^q(\Omega) \subset L^p(\Omega)$ dès que $p \leqslant q$. ItemWS4ooLcpOvXii

(2) Si $1 < p < 2$ et si $f \in L^2\big([0,1]\big)$ alors $\|f\|_p \leqslant \|f\|_2$.

Démonstration. Pour la simplicité des notations nous allons noter L^p pour $L^p(\Omega)$, et pareillement pour L^q. Soit $f \in L^q$. Nous posons

$$A = \{x \in \Omega \text{ tel que } |f(x)| \geqslant 1\}. \tag{27.158}$$

Étant donné que $p \leqslant q$ nous avons $|f|^p \leqslant |f|^q$ sur A ; par conséquent $\int_A |f|^p$ converge parce que $\int_A |f|^q$ converge.

L'ensemble A^c est évidemment borné (complémentaire dans Ω) et sur A^c nous avons $|f(x)| \leqslant 1$ et donc $|f|^p \leqslant 1$. L'intégrale $\int_{A^c} |f|^p$ converge donc également.

Au final $\int_\Omega |f|^p$ converge et $f \in L^p$.

Soit à présent $f \in L^2$; par le premier point nous avons immédiatement $f \in L^2 \cap L^p$. Soit aussi $r \in \mathbb{R}$ tel que $\frac{1}{2/p} + \frac{1}{r} = 1$. Nous avons $|f|^p \in L^{2/p}$, et vu que nous sommes sur un domaine borné, $1 \in L^r$. Nous écrivons l'inégalité de Hölder (27.81) avec ces fonctions. D'une part

$$\|f\|_1 = \||f|^p\|_1 = \|f\|_p^p. \tag{27.159}$$

D'autre part

$$\||f|^p\|_{2/p} = \left(\int |f|^2\right)^{p/2} = \|f\|_2^p. \tag{27.160}$$

Donc $\|f\|_p^p \leqslant \|f\|_2^p$, ce qui prouve l'assertion (2) parce que $p > 1$. □

Remarque 27.43.
Nous n'avons cependant pas $L^2\big([0,1]\big) = L^p\big([0,1]\big)$ parce que l'exemple (27.148) fonctionne encore :

$$f(x) = \frac{1}{\sqrt{x}} \tag{27.161}$$

pour $x \in [0,1]$ donne bien

$$\|f\|_2 = \int_0^1 \frac{1}{x} = \infty \tag{27.162}$$

et $\|f\|_p = \int_0^1 \frac{1}{x^{p/2}} < \infty$ parce que $1 < p < 2$.

27.3.7 Complétude

ThoUYBDWQX

Théorème 27.44 ([667, 668]).
Pour $1 \leqslant p < \infty$, l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est complet.

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans L^p. Pour tout i, il existe $N_i \in \mathbb{N}$ tel que $\|f_k - f_l\|_p \leqslant 2^{-i}$ pour tout $k, l \geqslant N_i$. Nous considérons la sous-suite $g_i = f_{N_i}$, de telle sorte qu'en particulier

$$\|g_i - g_{i-1}\|_p \leqslant 2^{-i+1}. \tag{27.163}$$ EqJLoDID

Pour chaque j nous considérons la somme télescopique

$$g_j = g_0 + \sum_{i=1}^{j} (g_i - g_{i-1}) \tag{27.164}$$

et l'inégalité

$$|g_j| \leqslant |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{27.165}$$

Nous allons noter

$$h_j = |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{27.166}$$ EqSomPaFPQOWC

La suite de fonctions (h_j) ainsi définie est une suite croissante de fonctions positives qui converge donc (ponctuellement) vers une fonction h qui peut éventuellement valoir l'infini en certains points. Par continuité de la fonction $x \mapsto x^p$ nous avons

$$\lim_{j\to\infty} h_j^p = h^p, \tag{27.167}$$

puis par le théorème de la convergence monotone (théorème 14.178) nous avons

$$\lim_{j\to\infty} \int_\Omega h_j^p d\mu = \int_\Omega h^p d\mu. \tag{27.168}$$

Utilisant à présent la continuité de la fonction $x \mapsto x^{1/p}$ nous trouvons

$$\lim_{j\to\infty} \left(\int h_j^p \right)^{1/p} = \left(\int |h|^p \right)^{1/p}. \tag{27.169}$$

Nous avons donc déjà montré que

$$\lim_{j\to\infty} \|h_j\|_p = \left(\int |h|^p \right)^{1/p} \tag{27.170}$$

où, encore une fois, rien ne garantit à ce stade que l'intégrale à droite soit un nombre fini. En utilisant l'inégalité de Minkowski (proposition 27.39) et l'inégalité (27.163) nous trouvons

$$\|h_j\|_p \leqslant \|g_0\|_p + \sum_{i=1}^{j} \|g_i - g_{i-1}\|_p \leqslant \|g_0\|_p + 1. \tag{27.171}$$

En passant à la limite,

$$\left(\int |h|^p \right)^{1/p} = \lim_{j\to\infty} \|h_j\|_p \leqslant \|g_0\|_p + 1 < \infty. \tag{27.172}$$

Par conséquent $\int |h|^p$ est finie et

$$h \in L^p(\Omega, \mathcal{A}, \mu). \tag{27.173}$$ EqgLpdUPQBP

En particulier, l'intégrale $\int h$ est finie (parce que $p \geqslant 1$) et donc que $h(x) < \infty$ pour presque tout $x \in \Omega$.

Nous savons que $h(x)$ est la limite des sommes partielles (27.166), en particulier la série

$$\sum_{j=1}^{\infty} |g_i - g_{i-1}| \tag{27.174}$$

converge ponctuellement. En vertu du corolaire 14.209, la série de terme général $g_i - g_{i-1}$ converge ponctuellement. La suite g_i converge donc vers une fonction que nous notons g. Par ailleurs la suite g_i est dominée par $h \in L^p$, le théorème de la convergence dominée (théorème 14.207) implique que

$$\lim_{j\to\infty} \|g_j - g\|_p = 0. \tag{27.175}$$

Nous allons maintenant prouver que $\lim_{n\to\infty} \|f_n - g\|_p = 0$. Soit $\epsilon > 0$. Pour tout n et i nous avons

$$\|f_n - g\|_p = \|f_n - f_{N_i} + f_{N_i} - g\|_p \leqslant \|f_n - f_{N_i}\|_p + \|f_{N_i} - g\|_p. \tag{27.176}$$

Pour rappel, $f_{N_i} = g_i$. Si i et n sont suffisamment grands nous pouvons obtenir que chacun des deux termes est plus petit que $\epsilon/2$.

Il nous reste à prouver que $g \in L^p(\Omega, \mathcal{A}, \mu)$. Nous avons déjà vu (équation (27.173)) que $h \in L^p$, mais $|g_i| \leqslant h^p$, par conséquent $g \in L^p$.

Nous avons donc montré que la suite de Cauchy (f_n) converge vers une fonction de L^p, ce qui signifie que L^p est complet. □

ThoGVmqOro

Théorème 27.45 (Riesz-Fischer[104, 669]).
Soient un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et $p \in [1, \infty]$. Alors

ItemPDnjOJzi

(1) Toute suite convergente dans $L^p(\Omega, \mathcal{A}, \mu)$ admet une sous-suite convergente presque partout sur Ω.

ItemPDnjOJzii

(2) La sous-suite donnée en (1) est dominée par un élément de $L^p(\Omega, \mathcal{A}, \mu)$.

ItemPDnjOJziii

(3) L'espace $L^p(\Omega, \mathcal{A}, \mu)$ est de Banach.

Démonstration. Le cas $p = \infty$ est à séparer des autres valeurs de p parce qu'on y parle de norme uniforme, et aucune sous-suite n'est à considérer.

(1) **Cas $p = \infty$.** Nous commençons par prouver dans le cas $p = \infty$. Soit (f_n) une suite de Cauchy dans $L^\infty(\Omega)$, ou plus précisément une suite de représentants d'éléments de L^p. Pour tout $k \geqslant 1$, il existe $N_k \geqslant 0$ tel que si $m, n \geqslant N_k$, on a

$$\|f_m - f_n\|_\infty \leqslant \frac{1}{k}. \tag{27.177}$$

En particulier, l'ensemble E_k sur lequel

$$|f_m(x) - f_n(x)| \geqslant \frac{1}{k} \tag{27.178}$$

est de mesure nulle. D'après le lemme 14.28, la partie $E = \bigcup_{k\in\mathbb{N}} E_k$, est encore de mesure nulle. En résumé, nous avons un N_k tel que si $m, n \geqslant N_k$, alors

$$|f_n(x) - f_m(x)| \leqslant \frac{1}{k} \tag{27.179}$$

EqKAWSptG

pour tout x hors de E. Donc pour chaque $x \in \Omega\backslash E$, la suite $n \mapsto f_n(x)$ est de Cauchy dans $\mathbb{R}$ et converge donc. Cela définit donc une fonction

$$\begin{aligned} f : \Omega\backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_n(x). \end{aligned} \tag{27.180}$$

Cela prouve le point (1) : la convergence ponctuelle.

En passant à la limite $n \to \infty$ dans l'équation (27.179) et tenant compte que cette majoration tient pour presque tout x dans Ω, nous trouvons

$$\|f - f_n\|_\infty \leqslant \frac{1}{k}. \tag{27.181}$$

Donc non seulement f est dans L^∞, mais en plus la suite (f_n) converge vers f au sens L^∞, c'est-à-dire uniformément. Cela prouve le point (3). En ce qui concerne le point (2), la suite f_n est entièrement (à partir d'un certain point) dominée par la fonction $1 + |f|$ qui est dans L^∞.

(2) **Cas $p < \infty$.**
Toute suite convergente étant de Cauchy, nous considérons une suite de Cauchy (f_n) dans $L^p(\Omega)$ et ce sera suffisant pour travailler sur le premier point. Pour montrer qu'une suite de

Cauchy converge, il est suffisant de montrer qu'une sous-suite converge. Soit $\varphi\colon \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante telle que pour tout $n \geqslant 1$ nous ayons

$$\|f_{\varphi(n+1)} - f_{\varphi(n)}\|_p \leqslant \frac{1}{2^n}. \tag{27.182}$$

Pour créer la fonction φ, il est suffisant de prendre le N_k donné par la condition de Cauchy pour $\epsilon = 1/2^k$ et de considérer la fonction définie par récurrence [20] par $\varphi(1) = N_1$ et $\varphi(n+1) > \max\{N_n, \varphi(n)\}$. Ensuite nous considérons la fonction

$$g_n(x) = \sum_{k=1}^{n} |f_{\varphi(k+1)}(x) - f_{\varphi(k)}(x)|. \tag{27.183}$$

Notons que pour écrire cela nous avons considéré des représentants f_k qui sont alors des fonctions à l'ancienne. Étant donné que g_n est une somme de fonctions dans L^p, c'est une fonction L^p, comme nous pouvons le constater en calculant sa norme :

$$\|g_n\|_p \leqslant \sum_{k=1}^{n} \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p \leqslant \sum_{k=1}^{n} \frac{1}{2^k} \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} = 1. \tag{27.184}$$

Étant donné que tous les termes de la somme définissant g_n sont positifs, la suite (g_n) est croissante. Mais elle est bornée en norme L^p et donc sujette à obéir au théorème de Beppo-Levi 14.178 sur la convergence monotone. Il existe donc une fonction $g \in L^p(\Omega)$ telle que $g_n \to g$ presque partout.
Soit un $x \in \Omega$ pour lequel $g_n(x) \to g(x)$; alors pour tout $n \geqslant 2$ et $\forall q \geqslant 0$, [EqWTHojCq]

$$\begin{aligned}
|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| &= \Big| \underbrace{f_{\varphi(n+q)}(x) + \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x)}_{=\sum_{k=1}^{q}} \underbrace{\sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - f_{\varphi(n)}(x)}_{=\sum_{k=1}^{q}} \Big| && \text{(27.185a)}\\
&= \left| \sum_{k=1}^{q} f_{\varphi(n+k)}(x) - \sum_{k=1}^{q} f_{\varphi(n+k-1)}(x) \right| && \text{(27.185b)}\\
&\leqslant \sum_{k=1}^{q} \left| f_{\varphi(n+k)}(x) - f_{\varphi(n+k-1)}(x) \right| && \text{(27.185c)}\\
&= g_{n+q+1}(x) - g_{n+1}(x) && \text{(27.185d)}\\
&\leqslant g(x) - g_{n+1}(x). && \text{(27.185e)}
\end{aligned}$$

Nous prenons la limite $n \to \infty$; la dernière expression tend vers zéro et donc

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| \to 0 \tag{27.186}$$

pour tout q. Donc pour presque tout $x \in \Omega$, la suite $n \mapsto f_{\varphi(n)}(x)$ est de Cauchy dans $\mathbb{R}$ et donc y converge vers un nombre que nous nommons $f(x)$. Cela définit une fonction

$$\begin{aligned} f\colon \Omega\backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_{\varphi(n)}(x) \end{aligned} \tag{27.187}$$

où E est de mesure nulle. Montrons que f est bien dans $L^p(\Omega)$; pour cela nous complétons la série d'inégalités (27.185) en

$$\left|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)\right| \leqslant g(x) - g_{n-1}(x) \leqslant g(x). \tag{27.188}$$

20. Utilisation du théorème 1.48. Vous n'êtes pas obligé de le citer à chaque fois, mais c'est bien de garder en tête que la définition de fonctions par récurrence n'est pas quelque chose de complètement trivial.

En prenant la limite $q \to \infty$ nous avons l'inégalité

$$|f(x) - f_{\varphi(n)}(x)| \leqslant g(x) \tag{27.189}$$

pour presque tout $x \in \Omega$, c'est-à-dire pour tout $x \in \Omega \backslash E$. Cette inégalité implique deux choses valables pour presque tout x dans Ω :

$$f(x) \in B\big(g(x), f_{\varphi(n)}(x)\big) \tag{27.190a}$$

$$f_{\varphi(n)}(x) \leqslant |f(x)| + |g(x)|. \tag{27.190b}$$

La première inégalité assure que $|f|^p$ est intégrable sur $\Omega\backslash E$ parce que $|f|$ est majorée par $|g| + |f_{\varphi(n)}|$. Elle prouve par conséquent le point (1) parce que $n \mapsto f_{\varphi(n)}$ est une sous-suite convergente presque partout. La seconde montre le point (2).

Attention : à ce point nous avons prouvé que $n \mapsto f_{\varphi(n)}$ est une suite de fonctions qui converge *ponctuellement presque partout* vers une fonction f qui s'avère être dans L^p. Nous n'avons pas montré que cette suite convergeait au sens de L^p vers f. Ce que nous devons montrer est que

$$\|f - f_{\varphi(n)}\|_p \to 0. \tag{27.191}$$

L'inégalité (27.189) nous donne aussi, toujours pour presque tout $x \in \Omega$:

$$\big|f(x) - f_{\varphi(n)}(x)\big|^p \leqslant g(x)^p \tag{27.192}$$

ce qui signifie que la suite [21] $|f - f_{\varphi(n)}|^p$ est dominée par la fonction $|g|^p$ qui est intégrable sur $\Omega\backslash E$ et tout autant sur Ω parce que E est négligeable ; cela prouve au passage le point (2), et le théorème de la convergence dominée de Lebesgue (14.207) nous dit que

$$\lim_{n\to\infty} \int_\Omega \big|f(x) - f_{\varphi(n)}(x)\big|^p dx = \int_\Omega \lim_{n\to\infty} \big|f(x) - f_{\varphi(n)}(x)\big| dx = 0. \tag{27.193}$$

Cette dernière suite d'égalités se lit de la façon suivante :

$$\lim_{n\to\infty} \|f - f_{\varphi(n)}\|_p = \Big\| \lim_{n\to\infty} |f - f_{\varphi(n)}| \Big\|_p = 0. \tag{27.194}$$

Nous en déduisons que la suite $n \mapsto f_{\varphi(n)}$ est convergente vers f au sens de la norme $L^p(\Omega)$. Or la suite de départ (f_n) était de Cauchy (pour la norme L^p) ; donc l'existence d'une sous-suite convergente implique la convergence de la suite entière vers f, ce qu'il fallait démontrer.

□

Le théorème suivant est souvent cité en disant que L^p est un espace de Hilbert si et seulement si $p = 2$. Comme vous le voyez, il faut un peu plus d'hypothèses.

Je précise que je suis le seul à nommer ce théorème par le nom de Weinersmith. Je ne sais pas si il a déjà un nom ; alors pourquoi pas celui-là plutôt qu'un autre ? La raison de ce choix est dans la constante de Weiner, définition 44.1.

Théorème 27.46 (Théorème de Weinersmith[670, 671]).
Nous considérons un espace mesuré $(\Omega, \mathcal{A}, \mu)$ ainsi qu'un nombre $p \in [1, \infty]$. Nous supposons

(1) $L^p(\Omega, \mathcal{A}, \mu)$ est un espace de Hilbert,

(2) Il existe des parties $A, B \subset L^p(\Omega, \mathcal{A}, \mu)$ telles que $A \cap B = \varnothing$ et $0 < \mu(A) < \infty$ et $0 < \mu(B) < \infty$.

Alors $p = 2$.

Démonstration. Vu que L^p est un espace de Hilbert (hypothèse), il vérifie l'identité du parallélogramme de la proposition 10.19, c'est-à-dire

$$\|f + g\|_p^2 + \|f - g\|_p^2 = 2\|f\|_p^2 + 2\|g\|_p^2. \tag{27.195}$$

21. À ce point, [104] se contente de majorer $|f_{\varphi(n)}(x)|$ par $|f(x)| + |g(x)|$, mais je ne comprends pas comment cette majoration nous permet d'utiliser la convergence dominée de Lebesgue pour montrer (27.191).

(1) **Pour $1 \leqslant p < \infty$**
Soient donc A, B comme dans l'hypothèse. Nous considérons les fonctions

$$f = \frac{1}{\mu(A)^{1/p}} \mathbb{1}_A \tag{27.196a}$$

$$g = \frac{1}{\mu(B)^{1/p}} \mathbb{1}_B. \tag{27.196b}$$

En ce qui concerne les normes L^p de f et g, c'est un calcul simple :

$$\|f\|_p^2 = \left(\int_\Omega |f(\omega)|^p d\mu \right)^{2/p} = \left(\int_A \frac{1}{\mu(A)} d\mu \right)^{2/p} = 1. \tag{27.197}$$

De même pour g : $\|f\|_p^2 = \|g\|_p^2 = 1$. Donc

$$2\|f\|_p^2 + 2\|g\|_p^2 = 4 \tag{27.198}$$

En ce qui concerne la somme,

$$\|f + g\|_p^2 = \left(\int_A \frac{d\mu}{\mu(A)} + \int_B \frac{d\mu}{\mu(B)} \right)^{2/p} = 2^{2/p}. \tag{27.199}$$

Pour la différence, la seule subtilité à voir est que

$$\int_\Omega |\mathbb{1}_A - \mathbb{1}_B|^p = \int_A |\mathbb{1}_A - \mathbb{1}_B|^p + \int_B |\mathbb{1}_A - \mathbb{1}_B|^p = \int_A |\mathbb{1}_A| + \int_B |-\mathbb{1}_B| = \int_A \mathbb{1}_A + \int_B \mathbb{1}_B. \tag{27.200}$$

Ce n'est pas de la magie que le moins se change en plus. Bref, pour la différence nous avons

$$\|f - g\|^2 = \left(\int_\Omega |f(\omega) - g(\omega)|^p d\mu \right)^{2/p} \tag{27.201a}$$

$$= \left(\int_\Omega |\frac{1}{\mu(A)^{1/p}} \mathbb{1}_A(\omega) - \frac{1}{\mu(B)^{1/p}} \mathbb{1}_B(\omega)|^p d\mu \right)^{2/p} \tag{27.201b}$$

$$= \left(\int_A |\frac{\mathbb{1}_A(\omega)}{\mu(A)^{1/p}}|^p + \int_B | - \frac{\mathbb{1}_B(\omega)}{\mu(B)^{1/p}}|^p \right)^{2/p} \tag{27.201c}$$

$$= 2^{2/p}. \tag{27.201d}$$

Donc $\|f + g\|_p^2 + \|f - g\|_p^2 = 2 \times 2^{2/p}$.
Vu que L^p est un espace de Hilbert, nous avons finalement

$$4 = 2 \times 2^{2/p}. \tag{27.202}$$

Cela est uniquement valable pour $p = 2$.

(2) **Pour $p = \infty$** Il suffit de prendre $f = \mathbb{1}_A$ et $g = \mathbb{1}_B$. Nous avons alors $\|f\|_{L^\infty}^2 = \|g\|_{L^\infty}^2 = \|f + g\|_{L^\infty}^2 = \|f - g\|_{L^\infty}^2 = 1$.
L'égalité (27.195) devient $2 = 4$, ce qui est faux.

□

Si $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré, est-ce que $L^p(\Omega, \mathcal{A}, \mu)$ est assuré de n'être pas de Hilbert ? Non.

Exemple 27.47 ([671])**.**
Soit la mesure de Dirac sur $\mathbb{R}$, c'est-à-dire

$$\delta(A) = \begin{cases} 1 & \text{si } 0 \in A \\ 0 & \text{sinon.} \end{cases} \tag{27.203}$$

Nous allons prouver que pour tout $p \in [0,\infty]$, l'espace $L^p(\mathbb{R}, \mathcal{B}or(\mathbb{R}), \delta)$ est un espace de Hilbert. Pour cela nous introduisons le produit hermitien

$$\langle f, g\rangle = f(0)\overline{g(0)}. \tag{27.204}$$

Nous avons par ailleurs la norme

$$\|f\|^p = \int_{\mathbb{R}} |f(x)|^p d\delta = f(0)^p. \tag{27.205}$$

Donc oui, $\|f\|_p = \sqrt{\langle f, f\rangle}$. △

27.3.8 Théorèmes d'approximation

PROPooUQUBooAWgNhm

Proposition 27.48 ([451], thème 24)**.**
Soient $1 \leqslant p \leqslant \infty$ et un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Alors les fonctions étagées[22] *dans $L^p(\Omega)$ sont denses dans $L^p(\Omega)$.*

Démonstration. Soit $f \in L^1(\Omega, \mathcal{A}, \mu)$. Nous supposons dans un premier temps que $f\colon \Omega \to [0,\infty[$. Pour les fonctions à valeurs dans $\mathbb{C}$, nous verrons plus bas.

Notons que la partie $\{x \in \Omega \text{ tel que } f(x) = \infty\}$ est de mesure nulle, donc nous pouvons vraiment choisir un représentant à valeurs dans $[0,\infty[$ et non à valeurs dans $[0,\infty]$ comme le serait un représentant un peu quelconque.

Par le théorème 14.118, il existe une suite croissante de fonctions étagées $\phi_n\colon \Omega \to [0,\infty[$ telles que $\phi_n \to f$ ponctuellement. Notons que ce théorème fonctionne parce que les fonctions L^p (en tout cas leurs représentants) sont mesurables parce que c'est dans la définition 27.9. Notre devoir est maintenant de prouver que sous l'hypothèse que f est dans L^p, alors la convergence $\phi_n \to f$ est une convergence dans L^p.

(1) $\underline{1 \leqslant p < \infty}$ Vu que f et ϕ_n sont à valeurs positives nous avons $|f - \phi_n|^p \leqslant |f|^p$. Mais par hypothèse $|f|^p \in L^1(\Omega)$. Donc la suite $g_n = |f - \phi_n|^p$ est majorée (uniformément en n) par $|f|^p$ qui est dans L^1. Le théorème de la convergence dominée permet de permuter

$$\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n = 0. \tag{27.206}$$

Cela revient à dire que

$$\lim_{n\to\infty} \|f - \phi_n\|_p^p = \lim_{n\to\infty} \int |f - \phi_n|^p = \int \lim_{n\to\infty} |f - \phi_n|^p = 0, \tag{27.207}$$

ce qui signifie que $\phi_n \xrightarrow{L^p(\Omega)} f$.

(2) $\underline{\textbf{Pour } p = \infty}$ Si $f \in L^\infty(\Omega)$, alors nous pouvons prendre un représentant borné. Avec lui, nous avons $\phi_n \xrightarrow{\|.\|_\infty} f$. Avec cela nous avons, pour ϵ donné, un n assez grand pour avoir

$$N_\infty(\phi_n - f) \leqslant \|\phi_n - f\|_\infty < \epsilon. \tag{27.208}$$

Si f est à valeurs dans $\mathbb{C}$ au lieu de $[0,\infty[$, il suffit de faire valoir le travail que nous venons de faire quatre fois, pour les valeurs réelles, imaginaires, positives et négatives. □

22. Définition 14.112. Pour rappel, une fonction est simple lorsqu'elle prend un nombre fini de valeurs, et elle est étagée lorsqu'elle est en outre mesurable.

27.3.9 Densité des fonctions infiniment dérivables à support compact

Définition 27.49.
*Une fonction est **étagée par rapport à** L^p si elle est de la forme*

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{27.209}$$

où les B_k sont des mesurables disjoints et $\mathbb{1}_{B_k} \in L^p$ pour tout k.

LemWHIRdaX

Lemme 27.50.
Si f est une fonction étagée en même temps qu'être dans L^p, alors elle est étagée par rapport à L^p.

Démonstration. Nous pouvons écrire

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{27.210}$$

où les B_k sont disjoints. Par hypothèse $\|f\|_p$ existe. Donc chacune des intégrales $\int_\Omega |\mathbb{1}_{B_k}|^p$ doit exister parce que les B_k étant disjoints, nous pouvons inverser la norme et la somme ainsi que la somme et l'intégrale :

$$\int_\Omega |f|^p = \int_\Omega \sum_{k=1}^{N} |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} \int |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} |c_k|^p \int_\Omega |\mathbb{1}_{B_k}(x)|^p dx. \tag{27.211}$$

□

L'ensemble $C_c^\infty(\mathbb{R}^d)$ des fonctions de classe C^∞ et à support compact sur $\mathbb{R}^d$ est souvent également noté $\mathscr{D}(\mathbb{R}^d)$.

ThoILGYXhX

Théorème 27.51 ([662]).
Nous avons des densités emboitées. Ici D est un borélien borné de $\mathbb{R}^d$ contenu dans $B(0,r)$ et K est un compact contenant $B(0,r+2)$.

(1) Les fonctions étagées par rapport à L^p sur $\mathbb{R}^d$ sont denses dans $L^p(\mathbb{R}^d)$. A fortiori les fonctions étagées sont denses dans L^p, mais nous n'en aurons pas besoin ici.
ItemYVFVrOIii

(2) Il existe une suite f_n dans $C(K,\mathbb{C})$ telle que

$$f_n \xrightarrow{L^p} \mathbb{1}_D. \tag{27.212}$$

ItemYVFVrOIiii

(3) Si A est un borélien tel que $\mathbb{1}_A \in L^p(\mathbb{R}^d)$[23], alors il existe une suite de boréliens bornée $(D_n)_{n\in\mathbb{N}}$ tels que

$$\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A. \tag{27.213}$$

ItemYVFVrOIiv

(4) Il existe une suite φ_n dans $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ telle que

$$\varphi_n \xrightarrow{L^p} \mathbb{1}_D. \tag{27.214}$$

ItemYVFVrOIv

(5) L'ensemble $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^d)$ pour tout $1 \leqslant p < \infty$.

Démonstration. Nous allons montrer les choses point par point.

23. Je pense que cette hypothèse manque dans [662]. En tout cas je vois mal comment je pourrais justifier les différentes étapes de la preuve en prenant par exemple $A = \mathbb{R}^d$.

(1) Si $f \in L^1(\mathbb{R}^d)$, nous savons par le théorème 14.118 qu'il existe une suite f_n de fonctions étagées convergeant ponctuellement vers f telle que $|f_n| \leqslant |f|$. La proposition 27.20 nous dit qu'alors $f_n \xrightarrow{L^p} f$.
La fonction f_n étant étagée et dans L^p en même temps, elle est automatiquement étagée par rapport à L^p par le lemme 27.50.
ItemYVFVrOIi

(2) C'est le théorème d'approximation 14.242 appliqué au borélien D contenu dans l'espace mesuré K.

(3) En vertu du point (2), il existe $f \in C^0(K, \mathbb{R})$ telle que

$$\|f - \mathbb{1}_D\|_p \leqslant \epsilon. \tag{27.215}$$

Ensuite, par le théorème de Weierstrass, il existe $\varphi \in C^\infty(K, \mathbb{R})$ telle que $\|f - \varphi\|_\infty \leqslant \epsilon$. Nous avons aussi

$$\|\varphi - f\|_p^p = \int_K |\varphi(x) - f(x)|^p dx \leqslant \mu(X)\|\varphi - f\|_\infty^p \leqslant \epsilon^p \mu(K). \tag{27.216}$$

Quitte à prendre un φ correspondant à un ϵ plus petit, nous avons

$$\|\varphi - f\| \leqslant \epsilon. \tag{27.217}$$

En combinant et en passant à $\epsilon/2$ nous avons trouvé une fonction $\varphi \in C^\infty(K, \mathbb{R})$ telle que

$$\|\varphi - \mathbb{1}_D\| \leqslant \epsilon. \tag{27.218}$$

(4) Nous considérons les boréliens fermés $D_n = A \cap B(0, n)$. Alors $\mathbb{1}_{D_n} \in L^p$ et nous avons pour n assez grand :

$$\int_{\mathbb{R}^d} |\mathbb{1}_{D_n}(x) - \mathbb{1}_A(x)|^p dx = \int_{\mathbb{R}^d \setminus B(0,n)} |\mathbb{1}_A(x)|^p < \epsilon, \tag{27.219}$$

c'est-à-dire que $\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A$.

(5) Il suffit de remettre tout ensemble. Si $f \in L^p(\mathbb{R}^d)$, par le point (2) nous commençons par prendre σ étagée par rapport à L^p telle que

$$\|\sigma - f\|_p \leqslant \epsilon. \tag{27.220}$$

Ensuite nous écrivons σ sous la forme

$$\sigma = \sum_{k=1}^N c_k \mathbb{1}_{B_k} \tag{27.221}$$

et nous appliquons le point (3) à chacune des $\mathbb{1}_{B_k}$ pour trouver des boréliens bornés D_k tels que

$$\|\mathbb{1}_{D_k} - \mathbb{1}_{B_k}\|_p \leqslant \epsilon. \tag{27.222}$$

Enfin nous appliquons le point (4) pour trouver des fonctions $\varphi_k \in C_c^\infty(\mathbb{R}^d)$ telles que

$$\|\varphi_k - \mathbb{1}_{D_k}\|_p \leqslant \epsilon. \tag{27.223}$$

Il n'est pas compliqué de calculer que

$$\Big\| \sum_{k=1}^N c_k \varphi_k - f \Big\|_p \leqslant 2\epsilon \sum_k c_k + \epsilon. \tag{27.224}$$

□

CorFZWooYNbtPz

Corolaire 27.52.
Si $1 < p < \infty$ alors la partie[24] $L^2\big([0,1]\big) \cap L^p\big([0,1]\big)$ *est dense dans* $L^p\big([0,1]\big)$.

Démonstration. Nous savons du théorème 27.51(5) que $C_c^\infty\big([0,1]\big)$ est dense dans L^p. Mais nous avons évidemment $C_c^\infty \subset L^2 \cap L^p$, donc $L^2 \cap L^p$ est dense dans L^p. □

LemCUlJzkA

Lemme 27.53 ([662, 629])**.**
Soit $1 \leqslant p < \infty$ et $f \in L^p(\Omega)$. Nous notons τ_v l'opérateur de translation par v :

$$\begin{aligned}\tau_v \colon L^p(\Omega) &\to L^p(\Omega)\\ f &\mapsto \Big[x \mapsto f(x-v)\Big].\end{aligned} \tag{27.225}$$

Pour chaque $f \in L^p(\Omega)$, l'application

$$\begin{aligned}\tau(f) \colon \mathbb{R}^d &\to L^p(\Omega)\\ v &\mapsto \tau_v(f)\end{aligned} \tag{27.226}$$

est continue en $v = 0$, c'est-à-dire

$$\lim_{v\to 0} \|\tau_v(f) - f\|_p = 0. \tag{27.227}$$

Démonstration. Nous commençons par supposer que f est dans $\mathscr{D}(\Omega)$, et nous verrons ensuite comment généraliser.

(1) **Si $f \in \mathscr{D}(\Omega)$**

Soit une suite $v_i \xrightarrow{\mathbb{R}^d} 0$, et posons $f_i = \tau_{v_i}(f)$; le but est de montrer que $f_i \xrightarrow{L^p} f$. Pour cela, la fonction $f - f_i$ est également à support compact, et qui plus est, si $\text{supp}(f) \subset B(0,r)$, alors $\text{supp}(f - f_i) \subset B(0, r + |v_i|)$, et l'ensemble

$$S = \overline{B\big(0, r + \max_i |v_i|\big)} \tag{27.228}$$

est un compact contenant les supports de tous les $f - f_i$. Le maximum existe parce que $v_i \to 0$. Voilà qui « majore » le domaine de $f - f_i$ uniformément en i.
Majorons maintenant $|f - f_i|^p$ de façon uniforme en i. Soit le nombre

$$M = 2 \max_{x\in\mathbb{R}^d}\{f(x)\}. \tag{27.229}$$

La fonction qui vaut M^p sur S et zéro ailleurs est une fonction intégrable qui majore $|f - f_i|^p$. Nous pouvons donc utiliser la convergence dominée de Lebesgue (théorème 14.207) pour écrire

$$\lim_{i\to\infty} \|f - f_i\|_p^p = \lim_{i\to\infty} \int_\Omega |f(x) - f(x - v_i)|^p dx = \int_\Omega \lim_{i\to\infty} |f(x) - f(x - v_i)| dx = 0. \tag{27.230}$$

(2) **Pour $f \in L^p(\Omega)$**
Soit $\epsilon > 0$, $f \in L^p(\Omega)$ et $\varphi \in \mathscr{D}(\Omega)$ tel que $\|f - \varphi\|_p \leqslant \epsilon$. Cela est possible par la densité de $\mathscr{D}(\Omega)$ dans $L^p(\Omega)$ vue en 27.51(5). Nous choisissons de plus $|v|$ assez petit pour avoir $\|\tau_v(\varphi) - \varphi\|_p < \epsilon$, qui est possible en vertu de ce que nous venons de démontrer à propos des fonctions à support compact. De plus τ_v étant une isométrie de L^p nous avons $\|\tau_v(\varphi) - \tau_v(f)\| = \|\varphi - f\| < \epsilon$. Nous avons tout pour majorer :

$$\|f - \tau_v(f)\| \leqslant \|f - \varphi\| + \|\varphi - \tau_v(\varphi)\| + \|\tau_v(\varphi) - \tau_v(f)\| \leqslant 3\epsilon. \tag{27.231}$$

Nous avons donc bien $\lim_{v\to 0} \|f - \tau_v(f)\| = 0$.

□

24. Nous parlons bien ici de l'*ensemble* L^2 parce que nous le considérons sans norme ou topologie particulière. La densité dont nous parlons ici est celle pour la topologique de L^p.

27.3.10 Approximation

LempTBaUw

Lemme 27.54 (Théorème fondamental d'approximation [433])**.**
Soit Ω un espace mesurable et $f\colon \Omega \to [0,\infty]$ une application mesurable. Alors il existe une suite croissante d'applications étagées $\varphi_n\colon \Omega \to \mathbb{R}^+$ dont la limite est f.

De plus si f est bornée, la convergence est uniforme.

ThoJsBKir

Théorème 27.55 ([320])**.**
Soit I un intervalle de $\mathbb{R}$. L'espace $\mathscr{D}(I)$ des fonctions continues à support compact sur I est dense dans $L^2(I)$.

Ce théorème sera généralisé à tous les $L^p(\mathbb{R}^d)$ par le théorème 27.51. Cependant L^p n'étant pas un Hilbert, il faudra travailler sans produit scalaire.

Démonstration. Soit $g \in L^2(I)$ une fonction telle que $g \perp f$ pour toute fonction $f \in C_c(I)$. Nous avons donc

$$\langle f, g\rangle = \int_I f\bar{g} = 0. \tag{27.232}$$

En passant éventuellement aux composantes réelles et imaginaires nous pouvons supposer que les fonctions sont toutes réelles. Nous décomposons g en parties positives et négatives : $g = g^+ - g^-$. Notre but est de montrer que $g^+ = g^-$, c'est-à-dire que g est nulle. La proposition 25.15 conclura que $C_c(I)$ est dense dans $L^2(I)$.

Soit un intervalle $[a,b] \subset I$ et une suite croissante de fonctions $f_n \in C_c(I)$ qui converge vers $\mathbb{1}_{[a,b]}$. Par hypothèse pour chaque n nous avons

$$\int_I f_n g^+ = \int_I f_n g^-. \tag{27.233}$$

La suite étant croissante, le théorème de la convergence monotone (théorème 14.178) s'applique et nous avons

$$\lim_{n\to\infty} \int_I f_n g^+ = \int_a^b g^+, \tag{27.234}$$

de telle sorte que nous ayons, pour tout intervalle $[a,b] \subset I$ l'égalité

$$\int_a^b g^+ = \int_a^b g^-. \tag{27.235}$$

EqYlErAM

De plus ces intégrales sont finies parce que

$$\int_a^b g^+ \leqslant \int_a^b |g| = \int_I |g|\mathbb{1}_{[a,b]} = \langle |g|, \mathbb{1}_{[a,b]}\rangle \leqslant \|g\|_{L^2}\sqrt{b-a} < \infty \tag{27.236}$$

par l'inégalité de Cauchy-Schwarz.

Soit maintenant un ensemble mesurable $A \subset I$. La fonction caractéristique $\mathbb{1}_A$ est mesurable et il existe une suite croissante de fonctions étagées (φ_n) convergente vers $\mathbb{1}_A$ par le lemme 27.54. À multiples près, les fonctions φ_n sont des sommes de fonctions caractéristiques du type $\mathbb{1}_{[a,b]}$, par conséquent, en vertu de (27.235) nous avons

$$\int_I \varphi_n g^+ = \int_I \varphi_n g^-. \tag{27.237}$$

Une fois de plus nous pouvons utiliser le théorème de la convergence monotone et obtenir

$$\int_A g^+ = \int_A g^- \tag{27.238}$$

pour tout ensemble mesurable $A \subset I$. Si nous notons dx la mesure de Lebesgue, les mesures g^+dx et g^-dx sont par conséquent égales et dominées par dx. Par le corolaire 14.236 du théorème de Radon Nikodym, les fonctions g^+ et g^- sont égales. □

27.3.11 Partie uniformément intégrable

DefOZlZnse

Définition 27.56.
*Une partie $H \subset L^1(\Omega, \mu)$ est **uniformément intégrable** si*

$$\lim_{a\to\infty} \left(\sup_{f\in H} \int_{|f|>a} |f(x)| d\mu(x) \right) = 0. \tag{27.239}$$

Notons dans cette définition que vu que $f \in L^1$ nous avons toujours

$$\lim_{a\to\infty} \int_{|f|>a} |f(x)| d\mu(x) = 0. \tag{27.240}$$

LEMooUOQDooIkVyWm

Lemme 27.57 ([1]).
La partie $H = \{f_n\}_{n\in\mathbb{N}}$ de $L^1(\Omega, \mu)$ est uniformément intégrable[25] *si et seulement si*

$$\lim_{a\to\infty} \left(\limsup_n \int_{|f_n|>a} |f_n| \right) = 0. \tag{27.241}$$

Cette formule est prise comme définition de l'uniforme intégrabilité par [672].

Démonstration. Quand on a une suite (x_n) de réels positifs, nous avons toujours $\limsup_n x_n \leqslant \sup_n x_n$. □

¡¡ Avertissement/question au lecteur !! 27.58
Je n'ai pas vérifié l'énoncé de 27.59. Si vous pensez que c'est faux, écrivez-moi. Écrivez-moi également si vous avez une démonstration.

PROPooPZRAooSiqgKT

Proposition 27.59 ([1]).
Soit une partie uniformément intégrable $\{f_n\}_{n\in\mathbb{N}}$. Soit une fonction intégrable g.

1. *La partie $\{f_n + g\}$ est uniformément intégrable.*
2. *Pour tout $\lambda \in \mathbb{R}$, la partie $\{\lambda f_n\}$ est uniformément intégrable.*
3. *Si pour tout n nous avons $g_n \leqslant f_n$, alors la partie $\{g_n\}$ est uniformément intégrable.*

PROPooPFZJooLySrgp

Proposition 27.60 ([1]).
Soit $0 < r < \infty$. Si $f \in L^r(\Omega, \mathcal{A}, \mu)$, alors

$$\int_{|f|^r\geqslant\lambda} |f|^r d\mu \xrightarrow{\lambda\to\infty} 0. \tag{27.242}$$

THOooGLYJooCbNysA

Théorème 27.61 ([673]).
Soient un espace mesuré fini $(\Omega, \mathcal{A}, \mu)$ et une partie uniformément intégrable[26] *$H \subset L^1(\Omega)$. Nous avons*

ITEMooCAIFooHXOKQb

1. *$\sup_{f\in H} \|f\|_{L^1(\Omega)} < \infty$.*
2. *Pour tout $\epsilon > 0$, il existe $\delta_\epsilon > 0$ tel que*

$$\int_A |f| d\mu < \epsilon \tag{27.243}$$

pour tout $f \in H$ et pour tout $A \in \mathcal{A}$ tel que $\mu(A) < \epsilon$.

25. Définition 27.56.
26. Définition 27.56.

27.4 Autres résultats de convergence

PROPooGJBIooAYsZQL

Proposition 27.62 ([1]).
Soit $0 < p < \infty$. Si $|f_n| \xrightarrow{L^p} 0$, alors $\liminf_n |f_n| = 0$ presque partout.

Démonstration. Nous y allons par contradiction en supposant qu'il existe une partie mesurable A de mesure non nulle telle que $\liminf_n |f_n(\omega)| > 0$ pour tout $\omega \in A$.

Pour chaque $\omega \in A$, nous posons $n_\omega \in \mathbb{N}$ et $v_\omega > 0$ tels que pour tout $n \geqslant n_\omega$, $|f_n(\omega)| > v_\omega$. Ensuite nous posons

$$A_{n,k} = \{\omega \in A \text{ tel que } n_\omega \leqslant n \text{ et } v_\omega > 1/k\}. \tag{27.244}$$

Nous avons l'union dénombrable $A = \bigcup_{n,k\in\mathbb{N}^2} A_{n,k}$. Vu qu'une union dénombrable d'ensembles de mesure nulle est de mesure nulle, il existe (n_0, k_0) tel que $\mu(A_{n_0,k_0}) > 0$. Pour tout $\omega \in A_{n_0,k_0}$ et pour tout $n > n_0$, nous avons

$$|f_\omega(\omega)| > \frac{1}{k_0}. \tag{27.245}$$

Nous avons donc la minoration, pour tout $n \geqslant n_0$,

$$\begin{aligned}
\|f_n\|_p^p &= \int_\Omega |f_n(\omega)|^p d\mu(\omega) && \text{(27.246a)}\\
&\geqslant \int_{A_{n_0,k_0}} |f_n(\omega)|^p d\mu(\omega) && \text{(27.246b)}\\
&\geqslant \int_{A_{n_0,k_0}} \frac{1}{k_0^p} d\mu(\omega) && \text{(27.246c)}\\
&= \frac{1}{k_0^p}\mu(A_{n_0,k_0}) > 0. && \text{(27.246d)}
\end{aligned}$$

Nous ne pouvons donc pas avoir $|f_n| \xrightarrow{L^p} 0$. □

PROPooUMBQooAhUgHA

Proposition 27.63 ([674]).
Si $0 < p \leqslant 1$, alors la formule

$$d_r(f, g) = \int_\Omega |f - g|^r d\mu \tag{27.247}$$

est une métrique invariante par translation.

Parfois il faut admettre que la théorie des probabilités a des notations pratiques. Par exemple si $X_n \xrightarrow{L^p} X$ alors $E(X_n) \to E(X)$. C'est la proposition suivante.

PROPooXXOYooCUtDZf

Proposition 27.64 ([675]).
Soient un espace mesuré $(\Omega, \mathcal{A}, \mu)$ ainsi que $0 < r < \infty$. Nous supposons que $f_n \xrightarrow{L^r} f$. Alors

$$\lim_{n\to\infty} \int_\Omega |f_n|^r = \int_\Omega |f|^r. \tag{27.248}$$

Démonstration. Si $r \geqslant 1$, c'est facile. L'application $f \mapsto \|f\|_r$ est une norme[27] et la proposition 7.159 s'applique. Donc

$$\big|\|f_n\|_r - \|f\|_r\big| \leqslant \|f - f_n\|_r. \tag{27.249}$$

Nous considérons $0 < r \leqslant 1$. Nous considérons la métrique d_r donnée par la proposition 27.63. Notre hypothèse de convergence signifie que $d_r(f_n, f) \to 0$. Nous utilisons l'inégalité triangulaire :

$$d_r(f_n, 0) \leqslant \underbrace{d_r(f_n, f)}_{\to 0} + d_r(f, 0) \to d(f, 0). \tag{27.250}$$

27. Proposition 27.16(2).

Nous avons donc [28]

$$\limsup_n d_r(f_n,0) \leqslant d_r(f,0). \tag{27.251}$$

EQooZATHooVuHffk

D'autre part la proposition 27.62 nous dit que $\liminf_n |f_n(\omega)-f(\omega)|=0$ pour tout $\omega\in\Omega\backslash Z$ avec $\mu(Z)=0$. Nous avons donc aussi $\liminf|f_n|=|f|$ (hors de Z) et $\liminf|f_n|^r=|f|^r$ (hors de Z). Vu que Z est de mesure nulle, nous pouvons l'oublier quand nous faisons des intégrales sur Ω :

$$\begin{aligned}
d_r(f,0) &= \int_\Omega |f|^r && (27.252a)\\
&= \int_\Omega \liminf_n |f_n|^r && (27.252b)\\
&= \liminf \int_\Omega |f_n|^r && \text{lem Fatou 14.182} \quad (27.252c)\\
&= \liminf d_r(f_n,0) && (27.252d)\\
&\leqslant \liminf \Big(d_r(f_n,f)+d_r(f,0)\Big) && (27.252e)\\
&= d_r(f,0). && (27.252f)
\end{aligned}$$

Vu que les expressions tout à gauche et tout à droite sont égales, toutes les inégalités sont des égalités. En particulier

$$\liminf_n d_r(f_n,0) = d_r(f,0). \tag{27.253}$$

En combinant avec (27.251), et en nous souvenant qu'une limite sup est toujours plus grande qu'une limite inf,

$$\limsup d_r(f_n,0) \leqslant d_r(f,0) \leqslant \liminf d_r(f_n,0) \leqslant \liminf d_r(f_n,0). \tag{27.254}$$

Encore une fois toutes les inégalités sont des égalités et nous avons

$$\liminf d_r(f_n,0) = \limsup d_r(f_n,0) = d_r(f,0). \tag{27.255}$$

Cela prouve que $\lim_{n\to\infty} d_r(f_n,0)$ existe et vaut $d_r(f,0)$ (lemme 9.45). Enfin nous pouvons faire le calcul

$$\int_\Omega |f|^r = d_r(f,0) = \lim_{n\to\infty} d_r(f_n,0) = \lim_{n\to\infty}\int_\Omega |f_n|^r. \tag{27.256a}$$

□

27.5 Convolution

DEFooHHCMooHzfStu

Définition 27.65.
Pour toutes fonctions $f,g\colon\mathbb{R}^n\to\mathbb{C}$ et pour tout $x\in\mathbb{C}$ tels que l'intégrale de droite ait un sens [29]*, nous définissons*

$$(f*g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy. \tag{27.257}$$

L'éventuelle fonction $f*g$ ainsi définie est le ***produit de convolution*** *de f et g.*

Le théorème qui permet de dire que le produit de convolution n'est pas tout à fait ridicule est le suivant.

THOooMLNMooQfksn

Théorème 27.66 ([468, 676])**.**
Soient $f,g\in L^1(\mathbb{R}^d)$.

28. Notons que $x_n\leqslant y_n\to\lambda$ n'implique pas que x_n converge. Par contre ça implique que $\limsup_n(x_n)\leqslant\lambda$.
29. Attention divulgâchis : ce sera le cas pour $f,g\in L^1(\mathbb{R}^n)$ par le théorème 27.66.

(1) *Pour presque tout $x \in \mathbb{R}^d$, la fonction*

$$\begin{aligned} h_x\colon \mathbb{R}^d &\to \mathbb{C} \\ y &\mapsto g(x-y)f(y) \end{aligned} \tag{27.258}$$

est dans $L^1(\mathbb{R}^d)$.

(2) $f * g \in L^1(\mathbb{R}^d)$.

(3) $\|f * g\|_1 \leqslant \|f\|_1\|g\|_1$.

Démonstration. Nous considérons l'application

$$\begin{aligned} \phi\colon \mathbb{R}^d \times \mathbb{R}^d &\to \mathbb{R}^d \times \mathbb{R}^d \\ (x,y) &\mapsto (x-y,y). \end{aligned} \tag{27.259}$$

Cela est un C^1-difféomorphisme dont le jacobien vaut

$$J_\phi(x,y) = \det\begin{pmatrix} (\partial_1\phi_1)(x,y) & (\partial_2\phi_1)(x,y) \\ (\partial_1\phi_2)(x,y) & (\partial_2\phi_2)(x,y) \end{pmatrix} = \det\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1. \tag{27.260}$$

C'est une première bonne chose.

Ensuite nous considérons la fonction

$$\begin{aligned} \alpha\colon \mathbb{R}^d \times \mathbb{R}^d &\to \mathbb{C} \\ (x,y) &\mapsto f(y)g(x). \end{aligned} \tag{27.261}$$

Par hypothèse, pour chaque y, la fonction $x \mapsto |\alpha(x,y)|$ est dans $L^1(\mathbb{R}^d)$. Bref, le calcul suivant a un sens :

$$\begin{aligned} \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|\alpha|dx\right]dy &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|f(y)g(x)|dx\right]dy && \text{(27.262a)} \\ &= \int_{\mathbb{R}^d}\left[|f(y)|\int_{\mathbb{R}^d}|g(x)|dx\right]dy && \text{(27.262b)} \\ &= \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)} && \text{(27.262c)} \\ &< \infty. && \text{(27.262d)} \end{aligned}$$

SUBEQooMYLWooPeNLlk

Le corolaire 14.303 nous dit alors que $\alpha \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Et notons au passage que

$$\|\alpha\|_{L^1(\mathbb{R}^d\times\mathbb{R}^d)} = \int_{\mathbb{R}^d\times\mathbb{R}^d}|\alpha| = \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)} \tag{27.263}$$

EQooLCWNooHLLdRd

parce que le théorème de Fubini 14.304 permet de scinder l'intégrale et de retomber sur (27.262c).

Vu que $\alpha \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$, nous pouvons utiliser le changement de variable 14.281 avec l'application ϕ ci-dessus. En notant λ la mesure de Lebesgue,

$$\begin{aligned} \infty > \int_{\mathbb{R}^d\times\mathbb{R}^d}|\alpha|d\lambda &= \int_{\mathbb{R}^d\times\mathbb{R}^d}\big(|\alpha|\circ\phi\big)|J_\phi|d\lambda && \text{(27.264a)} \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d}|\alpha|(x-y,y)d\lambda(x,y) && \text{(27.264b)} \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d}|f(y)g(x-y)|d\lambda(x,y) && \text{(27.264c)} \\ &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|f(y)g(x-y)|dy\right]dx. && \text{(27.264d)} \end{aligned}$$

Nous avons scindé l'intégrale avec le théorème de Fubini pour la dernière étape.

Passons à l'intégrabilité de $f * g$. Nous avons

$$|(f * g)(x)| \leqslant \int_{\mathbb{R}^d} |f(y)g(x-y)|dy. \tag{27.265}$$

EQooXMFYooKPIGfU

Or nous venons de voir que (27.265) était, en tant que fonction de x, intégrable sur $\mathbb{R}^d$ et que

$$\int_{\mathbb{R}^d} |(f * g)(x)|dx \leqslant \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} |f(y)g(x-y)|dy\right] dx = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha|d\lambda < \infty. \tag{27.266}$$

Cela prouve que $f * g \in L^1(\mathbb{R}^d)$, mais en nous souvenant de (27.263), cela prouve aussi que

$$\int_{\mathbb{R}^d} |(f * g)(x)|dx \leqslant \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha|d\lambda = \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)}, \tag{27.267}$$

c'est-à-dire que $\|f * g\|_{L^2(\mathbb{R}^d)} \leqslant \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^1(\mathbb{R}^d)}$. □

LEMooMRWZooHjrnHD

Lemme 27.67.
Le produit de convolution est commutatif : pour tout $f, g \in L^1(\mathbb{R}^d)$ nous avons $f * g = g * f$.

Démonstration. Le théorème de Fubini (théorème 14.304) permet d'écrire

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy = \int_{-\infty}^{\infty} dy_1 \dots \int_{-\infty}^{\infty} dy_n f(y)g(x-y). \tag{27.268}$$

En effectuant le changement de variable $z_i = x_i - y_i$ dans chacune des intégrales nous obtenons

$$(f * g)(x) = \int_{\mathbb{R}^n} g(z)f(x-z)dz = (g * f)(x). \tag{27.269}$$

Attention : on pourrait croire qu'un signe apparaît du fait que $z = x - y$ donne $dz = -dy$. Mais en réalité, l'intégrale $\int_{-\infty}^{+\infty}$ devient par le même changement de variables $\int_{+\infty}^{-\infty}$ qui redonne un nouveau signe au moment de remettre dans l'ordre. □

LEMooTUMSooSmnlHc

Lemme 27.68 ([677, 1])**.**
Le produit de convolution est associatif sur $L^1(\mathbb{R}^d)$.

Démonstration. Soient $f, g, h \in L^1(\mathbb{R}^d)$. L'existence de $(f * g) * h$ ne fait pas de doute grâce au théorème 27.66. Nous avons d'abord

$$\begin{aligned}
\big((f * g) * h\big)(u) &= \int_{\mathbb{R}^d} (f * g)(x)h(u-x)dx && \text{(27.270a)}\\
&= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} f(y)g(x-y)dy\right] h(u-x)dx && \text{(27.270b)}\\
&= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} f(y)g(x-y)h(u-x)dy\right] dx. && \text{(27.270c)}
\end{aligned}$$

SUBEQooJKRHooDaScXV

Nous permutons les intégrales en suivant la procédure 14.305. Pour cela nous commençons par poser

$$\begin{aligned}
s\colon \mathbb{R}^d \times \mathbb{R}^d &\to \mathbb{C}\\
(x, y) &\mapsto f(y)g(x-y)h(u-x),
\end{aligned} \tag{27.271}$$

et nous vérifions que $|s|$ peut être successivement intégrée par rapport à y puis x. D'abord l'intégrale par rapport à y est

$$\int_{\mathbb{R}^d} |f(y)||g(x-y)|dy, \tag{27.272}$$

qui existe et qui vaut $(|f| * |g|)(x)$ parce que $|f|$ et $|g|$ sont dans $L^1(\mathbb{R}^d)$. D'après le théorème 27.66, la fonction $|f| * |g|$ est encore dans $L^1(\mathbb{R}^d)$. En ce qui concerne l'intégrale du résultat par rapport à x, nous avons

$$\int_{\mathbb{R}^d} (|f| * |g|)(x)|h(u-x)|dx, \tag{27.273}$$

qui existe et qui vaut $\big((|f|*|g|)*|h|\big)(u)$. Le corolaire 14.303 nous assure donc que $s \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$.

Nous permutons donc les intégrales dans (27.270c) pour obtenir

$$\big((f*g)*h\big)(u) = \int_{\mathbb{R}^d} f(y) \left[\int_{\mathbb{R}^d} g(x-y)h(u-x)dx \right] dy. \tag{27.274}$$

EQooHXTBooQpmXcR

Attardons-nous un instant sur l'intégrale interne, et utilisons l'invariance par translation de l'intégrale (lemme 14.257). Nous effectuons la translation $x \to x+y$:

$$\int_{\mathbb{R}^d} g(x-y)h(u-x) = \int_{\mathbb{R}^d} g(x)h(u-x-y)dx \qquad = \int_{\mathbb{R}^d} g(x)h\big((u-y)-x\big)dx \tag{27.275a}$$

$$= (g*h)(u-y). \tag{27.275b}$$

Nous pouvons reprendre notre calcul en (27.274) :

$$\big((f*g)*h\big)(u) = \int_{\mathbb{R}^d} f(y) \left[\int_{\mathbb{R}^d} g(x-y)h(u-x)dx \right] dy \tag{27.276a}$$

$$= \int_{\mathbb{R}^d} f(y)(g*h)(u-y) \tag{27.276b}$$

$$= \big(f*(g*h)\big)(u). \tag{27.276c}$$

C'est ce que nous voulions. □

PROPooNBHNooInwoar

Proposition 27.69.
Le couple $\big(L^1(\mathbb{R}^d), *\big)$ est une algèbre de Banach[30].

Démonstration. Point par point.

(1) **Algèbre** La définition d'une algèbre est 1.347. Les différents points sont dans la linéarité de l'intégrale.

(2) **Commutative** C'est la proposition 27.67.

(3) **Associative** C'est le lemme 27.68.

(4) **Normé** L'espace L^1 a une norme ; c'est la norme $\|\|_{L^1}$.

(5) **Complet** C'est le théorème de Riesz-Fischer 27.45.

□

La proposition suivante est une conséquence de l'inégalité de Minkowski sous forme intégrale de la proposition 27.41.

PROPooDMMCooPTuQuS

Proposition 27.70.
Si $1 \leqslant p \leqslant \infty$ et si $f \in L^p(\mathbb{R}^d)$ et $g \in L^1(\mathbb{R}^d)$ alors

(1) $f*g \in L^p$

(2) $\|f*g\|_p \leqslant \|f\|_p \|g\|_1$.

PropHNbdMQe

Proposition 27.71 ([678])**.**
Si $f \in L^1(\mathbb{R})$ et si g est dérivable avec $g' \in L^\infty$, alors $f*g$ est dérivable et $(f*g)' = f*g'$.

Démonstration. La fonction qu'il faut intégrer pour obtenir $f*g$ est $f(t)g(x-t)$, dont la dérivée par rapport à x est $f(t)g'(x-t)$. La norme de cette dernière est majorée (uniformément en x) par $G(t) = |f(t)|\|g'\|_\infty$. La fonction f étant dans $L^1(\mathbb{R})$, la fonction G est intégrable et le théorème de dérivation sous l'intégrale (théorème 17.19) nous dit que $f*g$ est dérivable et

$$(f*g)'(x) = \frac{d}{dx}\int_{\mathbb{R}} f(t)g(x-t)dt = \int_{\mathbb{R}} f(t)g'(x-t)dt = (f*g')(x). \tag{27.277}$$

□

30. Algèbre de Banach, définition 7.265.

CORooBSPNooFwYQrc

Corolaire 27.72.
Si $f \in L^1(\mathbb{R}^d)$ et si g est de classe C^∞, alors $f * g$ est de classe C^∞.

Démonstration. Il s'agit d'itérer la proposition 27.71. □

27.5.1 Fonction plateau, lemme d'Urysohn

PROPooQWNNooHmHRbk

Proposition 27.73 ([1]).
Soient deux fonction f et g sur $\mathbb{R}^n$ telles que le produit de convolution $f * g$ ait un sens. Si le support de f est dans le compact K et le support de g est dans $B(0,r)$, alors

$$\operatorname{supp}(f * g) \subset K + \overline{B(0,r)} \tag{27.278}$$

PROPooBOZIooAhKbPs

Proposition 27.74 (Urysohn, Fonctions plateau[243, 679]).
Soit un ouvert $\Omega \subset \mathbb{R}^d$ ainsi qu'un compact K dans Ω. Il existe une fonction $f \in C^\infty(\mathbb{R}^d)$ telle que

(1) $\operatorname{supp}(f)$ *est compact dans* Ω,

(2) $0 \leqslant f \leqslant 1$

(3) $f = 1$ *sur un voisinage de* K.

Démonstration. Nous posons $\delta = d(K, \Omega^c)/3$.

(1) **Les applications ϕ et ϕ_ϵ** Nous nous souvenons de l'application ξ de la proposition 15.162. Elle est de classe C^∞ à support compact dans $B(0,1)$. Nous considérons [31]

$$\phi(x) = \frac{\xi(x)}{\int_{\mathbb{R}^n} \xi(x)dx}, \tag{27.279}$$

de telle sorte que $\int_{\mathbb{R}^n} \phi = 1$. Pour chaque $\epsilon > 0$ nous posons également

$$\begin{aligned} \phi_\epsilon \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{1}{\epsilon^n}\phi(x/\epsilon). \end{aligned} \tag{27.280}$$

Par le lemme 15.167, nous avons

$$\int_{B(0,\epsilon)} \phi_\epsilon = \int_{\mathbb{R}^n} \phi_\epsilon = 1. \tag{27.281}$$ EQooDUOCooLKPmQZ

(2) **La partie K_1** Nous posons

$$K_1 = \{x \in \mathbb{R}^n \text{ tel que } \phi(x, K) < \delta\}. \tag{27.282}$$ EQooLXRMooKxKHSF

La partie K_1 est bornée parce que si $K \subset B(0,r)$, alors $K_1 \subset B(0, r+\delta)$.
La partie K_1 est fermée. Pour ce voir, nous prouvons que K_1^c est ouvert. Soit $y \in K_1^c$, et prouvons qu'il existe une boule centrée en y et contenue dans K_1^c. Soit $k \in K_1$ réalisant $d(z, K_1) = d(z,k)$ (d est une fonction continue sur le compact K_1, elle atteint ses bornes). Soit $z \in B(y,\alpha)$ où α va être précisé de telle sorte à avoir $B(y,\alpha) \subset K_1^c$. Nous avons

$$\begin{aligned} d(z,k) &\geqslant |d(z,y) - d(y,k)| && \text{cf. justif} && \text{SUBEQooAKWEooGpnxaX (27.283a)} \\ &= d(y,k) - d(z,y) && \text{cf. justi.} && \text{SUBEQooGTKYooEPMiZR (27.283b)} \\ &> \delta + s - \alpha && \text{cf. justif.} && \text{SUBSETooGVBEooIUCcKE (27.283c)} \\ & && && \text{(27.283d)} \end{aligned}$$

Justifications.

31. L'intégrale de ξ ne pose pas de problèmes parce qu'elle est continue sur le compact de son support, de telle sorte qu'elle y soit bornée.

— Pour (27.283a). Lemme 7.116.

— Pour (27.283b). Nous choisissons $\alpha < d(y,k)$.

— Pour 27.283c. Vu que $d(y,k) > \delta$, nous posons $s > 0$ tel que $d(y,k) = \delta + s$.

Nous avons prouvé qu'il existe $s > 0$ tel que pour tout $\alpha > 0$ suffisamment petit, nous ayons $d(z,k) > \delta + s - \alpha$. Nous en déduisons que $d(z,k) > \delta$ et donc que $z \in K_1^c$.

Nous avons prouvé que $K - 1$ est borné et fermé. Il est donc compact (théorème de Borel-Lebesgue 9.24).

(3) **La fonction f** Nous posons $f = \phi_\epsilon * \mathbb{1}_{K_1}$ où $\mathbb{1}_{K_1}$ est la fonction indicatrice de K_1 et $*$ dénote le produit de convolution [32]. Vu que ϕ_ϵ est de classe C^∞, la fonction f est également de classe C^∞ par le corolaire 27.72.

(4) **À propos de support** La proposition 27.73 dit que le support de f est

$$\text{supp}(f) \subset K_1 + \overline{B(0,\epsilon)}. \tag{27.284}$$

Vue la définition 27.282, nous avons $d(K_1, \Omega^c) \geqslant \delta$, de telle sorte que, en prenant $\epsilon < \delta$ nous avons

$$\text{supp}(f) \subset \Omega. \tag{27.285}$$

Le support de f est également borné parce que contenu dans le borné $K_1 + \overline{B(0,\epsilon)}$. Nous en déduisons que f est une application C^∞ à support compact dans Ω.

(5) **$f(x)$ si $x \in K$** Soit $x \in K$. Nous avons, par définition,

$$\begin{aligned}
f(x) &= \int_{\mathbb{R}^n} \phi_\epsilon(y)\mathbb{1}_{K_1}(x-y)dy && && (27.286a)\\
&= \int_{\overline{B(0,\epsilon)}} \phi_\epsilon(y)\mathbb{1}_{K_1}(x-y)dy && \text{cf. justif.} && (27.286b)\\
&= \int_{\overline{B(0,\epsilon)}} \phi_\epsilon(y)dy && \text{cv. justif.} && (27.286c)\\
&= 1 && \text{par (27.281).} && (27.286d)
\end{aligned}$$

Justifications.

— Pour 27.286b. Le support de ϕ_ϵ est dans $\overline{B(0,\epsilon)}$.

— Pour 27.286c. Vu que $x \in K$ et que $|y| < \epsilon < \delta$, nous avons

$$d(x-y, K) \leqslant d(x-y, x) \leqslant \delta, \tag{27.287}$$

et donc $x - y \in K_1$.

(6) **$0 \leqslant f \leqslant 1$** Nous avons

$$\begin{aligned}
|f(x)| &\leqslant \int_{\mathbb{R}^n} \big|\phi_\epsilon(y)\mathbb{1}_{K_1}(x-y)\big|dy && (27.288a)\\
&\leqslant \int_{\mathbb{R}^n} |\phi_\epsilon(y)|dy && (27.288b)\\
&= 1. && (27.288c)
\end{aligned}$$

Note : nous avons majoré $\mathbb{1}_{K_1}(x-y)$ par 1.

□

32. Définition 27.65.

27.5.2 Partition de l'unité

THOooQFCQooSlgLpz

Théorème 27.75 (Partition de l'unité[243])**.**
Soient un compact $K \subset \mathbb{R}^d$ et des ouverts $\{\Omega_i\}_{i=1,\dots,n}$ recouvrant K. Alors il existe des fonctions $\phi_i \in C^\infty(\mathbb{R}^d)$ $(i = 1, \dots, n)$ telles que

(1) $0 \leqslant \phi_k \leqslant 1$

(2) $\text{supp}(\phi_k) \subset \Omega_k$,

(3) $\sum_{k=1}^n \phi_k = 1$ *sur un voisinage de* K.

Ces fonctions ϕ_i sont une ***partition de l'unité*** *subordonnée aux ouverts Ω_i.*

Démonstration. Nous considérons des compacts $\{K_i\}_{i=1,\dots,n}$ comme dans le lemme 7.307. Pour chaque i nous avons $K_i \subset \Omega_i$, de telle sorte que nous pouvions utiliser la proposition 27.74. Nous avons donc des fonctions $\psi_j \in C^\infty(\mathbb{R}^d)$ telles que $\text{supp}(\psi_j)$ est compact dans Ω_j, $\psi_j \geqslant 0$ et $\psi_j = 1$ sur V_j qui est un voisinage ouvert de K_j.

Nous posons $V = V_1 \cup \dots \cup V_n$. C'est un ouvert qui contient K parce que

$$K \subset \bigcup_j K_j \subset \bigcup_j \Omega_j \subset \bigcup_j V_j. \tag{27.289}$$

Nous faisons de même pour K lui-même. Il existe une fonction $\theta \in C^\infty(\mathbb{R}^d)$ telle que

— $\text{supp}(\theta)$ est compact dans Vn

— $\theta = 1$ sur un voisinage de K,

— $0 \leqslant \theta \leqslant 1$.

Vu que $\psi_j = 1$ sur V_j, nous avons $\sum_{j=1}^n \psi_j > 0$ sur V.

Nous prouvons à présent que

$$1 - \theta + \sum_{k=1} \psi_k > 0 \tag{27.290}$$

EQooGABMooZZmlyQ

sur $\mathbb{R}^d$.

(1) **Si $x \in V$** Alors $1 - \theta(x) \geqslant 0$ et donc

$$1 - \theta(x) + \sum_{k=1}^n \psi_k(x) \geqslant \sum_{k=1}^n \psi_k(x) > 0. \tag{27.291}$$

(2) **Si $x \notin V$** Vu que le support de θ est dans V, nous avons $\theta(x) = 0$. Quant aux fonctions ψ_k, elles font un peu ce qu'elles veulent, mais elles sont positives [33] et donc

$$1 - \theta(x) + \sum_{k=1}^n \psi_k(x) \geqslant 1 - \theta(x) = 1. \tag{27.292}$$

Vu que (27.290) est prouvé, nous pouvons poser

$$\phi_j = \frac{\psi_j}{1 - \theta + \sum_{k=1}^n \psi_k} \tag{27.293}$$

sans peur pour le dénominateur. Nous avons $\phi_j \in C^\infty(\mathbb{R}^d)$. Nous savons que $\theta = 1$ sur un voisinage de K. Su ce voisinage nous avons

$$\sum_{j=1}^n \phi_j(x) = \frac{\sum_j \psi_j(x)}{1 - \theta(x) + \sum_k \psi_k(x)} = 1. \tag{27.294}$$

De plus $\phi_j \geqslant 0$ parce que chacun des ψ_j l'est et parce que nous avons montré que le dénominateur était toujours strictement positif.

En enfin,

$$\text{supp}(\phi_j) = \text{supp}(\psi_j) \subset \Omega_j. \tag{27.295}$$

Donc les fonctions ϕ_j sont celles que dont nous avions besoin. □

33. Dans le Frido, « positif » signifie dans $[0, \infty]$.

27.5.3 Approximation de l'unité

LemDQEKNNf

Lemme 27.76.
Soit $f \in L^2(I)$ telle que

$$\int_I f\varphi = 0 \tag{27.296}$$

pour toute fonction $\varphi \in C_c^\infty(I)$. Alors $f = 0$ presque partout sur I.

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} \phi\colon L^2(I) &\to \mathbb{C} \\ g &\mapsto \langle f, g\rangle = \int_I f\bar{g}. \end{aligned} \tag{27.297}$$

Par densité[34] nous pouvons aussi considérer une suite (φ_n) dans $C_c^\infty(I)$ convergeant dans L^2 vers f. Alors nous avons pour tout n :

$$\langle f, \varphi_n\rangle = 0. \tag{27.298}$$

En passant à la limite, $\langle f, f\rangle = 0$, ce qui implique $f = 0$ dans L^2 et donc $f = 0$ presque partout en tant que bonne fonction. □

Ce résultat est encore valable dans les espaces L^p (proposition 27.189), mais il demande le théorème de représentation de Riesz[35].

DEFooEFGNooOREmBb

Définition 27.77 ([1, 662]).
*Nous considérons $\Omega = \mathbb{R}^d$ ou $(S^1)^d$. Une **approximation de l'unité** sur Ω autour de $a \in \Omega$ est une suite (φ_n) de fonctions à valeurs réelles dans $L^1(\Omega)$ telle que*

ITEMooGVRQooHDbrcf

(1) $\sup_k \|\varphi_k\|_1 < \infty$,

(2) pour chaque n nous avons $\int_\Omega \varphi_n = 1$,

(3) si V est un voisinage de a, alors

$$\lim_{k\to\infty} \int_{\Omega\setminus V} |\varphi_k| = 0. \tag{27.299}$$

En pratique, nous allons, sur $\mathbb{R}^d$ toujours considérer des approximations de l'unité autour de 0, même si nous ne le préciserons pas. Vous noterez que dans le cas de S^1, le choix du « point de base » est plus arbitraire.

Ce sont des fonctions dont la masse vient s'accumuler autour de zéro. En effet quel que soit le voisinage $B(0,\alpha)$, si k est assez grand, il n'y a presque plus rien en dehors.

Pour le point ((3)), si Ω est S^1, la mesure que nous considérons est $\frac{dx}{2\pi}$.

Exemple 27.78.
Une façon de construire une approximation de l'unité sur $\mathbb{R}$ est de considérer une fonction $\varphi \in L^1(\Omega)$ telle que $\int\varphi = 1$ puis de poser

$$\varphi_k(x) = k^d\varphi(kx). \tag{27.300}$$

Ici, Ω peut être $\mathbb{R}$ ou S^1. △

Le lemme suivant permet de construire des approximations de l'unité intéressantes. Nous aurons une version pour S^1 dans le lemme 27.113.

34. Théorème 27.51(5).
35. Théorème 27.186.

LemCNjIYhv

Lemme 27.79 ([662]).
Soit une fonction φ continue et positive à support compact sur $\mathbb{R}^d$ telle que $\varphi(x) > \varphi(0)$ pour tout $x \neq 0$. Si nous posons

$$\varphi_n(x) = \left(\int \varphi(y)^n \right)^{-1} \varphi(x)^n, \tag{27.301}$$

alors la suite (φ_n) est une approximation de l'unité.

Voici un théorème qui donne les propriétés à propos du produit de convolution avec une approximation de l'unité dans $\mathbb{R}^d$. Une version pour S^1 sera le théorème 27.114.

ThoYQbqEez

Théorème 27.80 ([662]).
Soit (φ_k) une approximation de l'unité sur $\mathbb{R}^d$.

(1) Si g est mesurable et bornée sur $\mathbb{R}^d$ et si g est continue en x_0 alors

$$(\varphi_k * g)(x_0) \to g(x_0). \tag{27.302}$$

(2) Si $g \in L^p(\mathbb{R}^d)$ $(1 \leqslant p < \infty)$ alors

$$\varphi_k * g \xrightarrow{L^p} g. \tag{27.303}$$

(3) Si g est uniformément continue et bornée, alors

$$\varphi_k * g \xrightarrow{L^\infty} g \tag{27.304}$$

Démonstration. En plusieurs points.

(1) Nous notons $d_k = (\varphi_k * g)(x_0) - g(x_0)$ et nous devons prouver que $d_k \to 0$. Vu que φ_k est d'intégrale 1 sur $\mathbb{R}^d$ nous pouvons écrire

$$d_k = \int_{\mathbb{R}^d} \varphi_k(y) g(x_0 - y) dy - \int_{\mathbb{R}^d} g(x_0)\varphi_k(y) dy, \tag{27.305}$$

et donc

$$|d_k| = \Big| \int_{\mathbb{R}^d} \big(g(x_0 - y) - g(x_0)\big)\varphi_k(y) dy \Big| \leqslant \int_{\mathbb{R}^d} \big|g(x_0 - y) - g(x_0)\big||\varphi_k(y)| dy. \tag{27.306}$$

Nous notons $M = \sup_k \|\varphi_k\|_1$, et nous considérons $\alpha > 0$ tel que

$$\big|g(x_0 - y) - g(x_0)\big| \leqslant \epsilon \tag{27.307}$$

pour tout $y \in B(0, \alpha)$. Nous nous restreignons maintenant aux k suffisamment grands pour que $\int_{\complement B(0,\alpha)} |\varphi_k(y)| dy \leqslant \epsilon$. Alors en découpant l'intégrale en $B(0, \alpha)$ et son complémentaire dans $\mathbb{R}^d$,

$$|d_k| \leqslant \epsilon M + \int_{\complement B(0,\alpha)} 2\|g\|_\infty |\varphi_k(y)| dy \leqslant \epsilon M + 2\|g\|_\infty \epsilon \leqslant \epsilon C. \tag{27.308}$$

Donc oui, nous avons $|d_k| \to 0$, et donc le premier point du théorème.

(2) Cette fois $g \in L^p(\mathbb{R}^d)$ et nous cherchons à montrer que $\|d_k\|_p \to 0$. Encore qu'ici d_k soit défini à partir d'un représentant dans la classe de g et que d'ailleurs, nous allons travailler avec ce représentant.

D'abord nous développons un peu ce d_k :

$$\|d_k\|_p = \left[\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \big(g(x - y) - g(x)\big)\varphi_k(y) dy \right|^p dx \right]^{1/p} \tag{27.309a}$$

$$\leqslant \left[\int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |g(x - y) - g(x)| \cdot |\varphi_k(y)| dy \Big)^p dx \right]^{1/p}. \tag{27.309b}$$

À cette dernière expression nous appliquons l'inégalité de Minkowski (théorème 27.39) sous la forme (27.135) pour la mesure $d\nu(y) = |\varphi_k(y)|dy$ et $f(x,y) = g(x-y) - g(x)$:

$$\|d_k\|_p \leqslant \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |g(x-y) - g(x)|^p dx \Big)^{1/p} |\varphi_k(y)|dy = \int_{\mathbb{R}^d} \|\tau_y g - g\|_p |\varphi_k(y)|dy. \quad (27.310)$$

Par le lemme 27.53 nous pouvons trouver $\alpha > 0$ tel que $\|\tau_y g - g\|_p \leqslant \epsilon$ pour tout $y \in B(0,\alpha)$. Avec cela nous découpons encore le domaine d'intégration :

$$\|d_k\|_p \leqslant \int_{B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant\epsilon} |\varphi_k(y)|dy + \int_{\complement B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant 2\|g\|_p} |\varphi_k(y)|dy \leqslant \epsilon M + 2\epsilon\|g\|_p. \quad (27.311)$$

(3) Nous posons $d_k(x) = (\varphi_k * g)(x) - g(x)$ et nous voulons prouver que $\|d_k\|_\infty \to 0$, c'est-à-dire que $d_k(x)$ converge vers zéro uniformément en x. Nous posons aussi

$$\tau_y(g)\colon x \mapsto g(x-y). \quad (27.312)$$

En récrivant le produit de convolution, une petite majoration donne

$$|d_k(x)| \leqslant \int_{\mathbb{R}^d} \|\tau_y(g) - g\|_\infty |\varphi_k(y)|dy. \quad (27.313)$$

L'uniforme continuité de g signifie que pour tout ϵ, il existe un α tel que pour tout $y \in B(0,\alpha)$,

$$\|\tau_y(g) - g\|_\infty \leqslant \epsilon. \quad (27.314)$$

Encore une fois nous découpons le domaine d'intégration en $B = B(0,\alpha)$ et son complémentaire :

$$\|d_k\|_\infty \leqslant \int_B \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant\epsilon} |\varphi_k(y)|dy + \int_{\complement B} \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant 2\|g\|_\infty} |\varphi_k(y)| \quad (27.315a)$$

$$\leqslant \epsilon M + 2\|g\|_\infty \epsilon \quad (27.315b)$$

où la seconde ligne est justifiée par le choix d'un k assez grand pour que $\int_{\complement B} |\varphi_k(y)|dy \leqslant \epsilon$. Nous avons donc bien $\|d_k\|_\infty \to 0$.

□

Exemple 27.81.
Une petite remarque en passant : aussi triste que cela en ait l'air, la convergence uniforme n'implique pas la convergence $L^p(\Omega)$ si Ω n'est pas borné. En effet si $f \in L^p$, la suite donnée par

$$f_n(x) = f(x) + \frac{1}{n} \quad (27.316)$$

converge uniformément vers f, mais

$$\|f_n - f\|_p = \int_\Omega \frac{1}{n} \quad (27.317)$$

n'existe même pas si le domaine Ω n'est pas borné. △

27.6 Intégrale sur des variétés

27.6.1 Formes différentielles

Nous allons donner une toute petite introduction aux formes différentielles sur des variétés compactes.

LemdwLGFG

Lemme 27.82 ([680]).
Soit ω une k-forme sur $\mathbb{R}^n$ et f, une fonction C^∞ sur $\mathbb{R}^n$. Alors $d(f^*\omega) = f^*d\omega$.

Démonstration. Nous effectuons la preuve par récurrence sur le degré de la forme. Soit d'abord une 0-forme, c'est-à-dire une fonction $g\colon \mathbb{R}^n \to \mathbb{R}$. Nous avons

$$d(d^*g)X = d(g\circ f)X = (dg\circ df)X = dg\big(df X\big) = (f^*dg)(X). \tag{27.318}$$

Supposons maintenant que le résultat soit exact pour toutes les $p-1$-formes et montrons qu'il reste valable pour les p-formes. Par linéarité de la différentielle nous pouvons nous contenter de considérer la forme différentielle

$$\omega = g\,dx^1\wedge\ldots dx^p \tag{27.319}$$

où g est une fonction C^∞. Pour soulager les notations nous allons noter $dx^I = dx^1\wedge\ldots dx^{p-1}$. Nous avons

$$\begin{aligned}
d(f^*\omega) &= d\big(f^*(gdx^I\wedge dx^p)\big) && \text{(27.320a)}\\
&= d\big(f^*(gdx^I)\wedge f^*dx^p\big) && \text{(27.320b)}\\
&= d\big(f^*(gdx^I)\big)\wedge f^*dx^p + (-1)^{p-1}f^*(gdx^I)\wedge(f^*dx^p) && \text{(27.320c)}\\
&= f^*\big(d(gdx^I)\big)\wedge f^*dx^p && \text{(27.320d)}\\
&= f^*\big(d(gdx^I)\wedge dx^p\big) && \text{(27.320e)}\\
&= f^*d\omega && \text{(27.320f)}
\end{aligned}$$

Justifications : (27.320c) est la formule de Leibniz. (27.320d) est parce que le second terme est nul : $d(f^*dx^p) = f^*(d^2x^p) = 0$. Nous avons utilisé l'hypothèse de récurrence et le fait que $d^2 = 0$. L'étape (27.320f) est une utilisation à l'envers de la règle de Leibniz en tenant compte que $d^2x^p = 0$. □

Soit M une variété de dimension n et ω une n-forme différentielle

$$\omega_p = f(p)dx_1\wedge\ldots\wedge dx_n. \tag{27.321}$$

Si (U,φ) est une carte ($U\subset\mathbb{R}^n$ et $\varphi\colon U\to M$) alors nous définissons

$$\int_{\varphi(U)}\omega = \int_U f\big(\varphi(x)\big)dx_1\ldots dx_n. \tag{27.322}$$

Lorsque nous voulons intégrer sur une partie plus grande qu'une carte nous utilisons une partition de l'unité du théorème 27.75.

27.6.2 Intégrale d'une fonction sur une sous-variété

L'exemple typique est l'intégrale sur une surface dans $\mathbb{R}^3$, ou de volumes.

Nous supposons à présent que M est une variété compacte de dimension 2 dans $\mathbb{R}^3$. La compacité fait que M possède un atlas contenant un nombre fini de cartes $F_i\colon W_i\to F_i(W_i)$.

Si $A\subset M$ est tel que pour chaque i, $A\cap F_i(W_i) = F_i(V_i)$ pour un ensemble V_i mesurable dans $\mathbb{R}^2$, alors nous considérons

$$A_1 = A\cap F_1(W_1) = F_1(V_1). \tag{27.323}$$

Ensuite, nous construisons A_2 en considérant $F_2(W_2)$ et en lui retranchant A_1 :

$$A_2 = \big(A\cap F_2(W_2)\big)\backslash F_1(V_1). \tag{27.324}$$

En continuant de la sorte, nous construisons la décomposition

$$A = A_1\cup\ldots\cup A_p \tag{27.325}$$

de A en ouverts disjoints, chacun de ouverts A_p étant compris dans une carte.

Il est possible de prouver que dans ce cas, la définition suivante a un sens et ne dépend pas du choix de l'atlas effectué.

Définition 27.83.
*Si $f : A \to \mathbb{R}$ est une fonction continue, alors l'**intégrale** est le nombre*

$$\int_A f = \sum_{i=1}^{p} \int_{A_i} f d\sigma_{F_i}. \tag{27.326}$$

27.6.3 Densité des polynômes trigonométriques

DEFooGCZAooFecAHB

Définition 27.84.
*Le **système trigonométrique** donné par $\{e_n\}_{n\in\mathbb{Z}}$ est*

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}. \tag{27.327}$$

Une bonne partie de la douleur qu'évoque mot « densité » consiste à montrer que ce système est total dans $L^2(S^1) = L^2([0, 2\pi])$, et donc en est une base hilbertienne.

Définition 27.85.
*Un **polynôme trigonométrique** est une fonction de la forme*

$$P(t) = \sum_{n=-N}^{N} c_n e_n(t). \tag{27.328}$$

DEFooZDUKooRnYhhF

Définition 27.86 (Coefficients de Fourier)**.**
Pour toute fonction pour laquelle ça a un sens (que ce soit des fonctions L^2 ou non), nous posons

$$c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt. \tag{27.329}$$

EqhIPoPH

*Ces nombres sont les **coefficients de Fourier** de f.*

Ces trois définitions n'ont à priori aucun rapport entre elles, et rien en particulier ne devrait vous faire penser à une égalité du type

$$f(x) = \sum_{n=-\infty}^{\infty} c_n(f) e_n(x). \tag{27.330}$$

Nous avons toutefois quelques liens.

LemZVfZlms

Lemme 27.87.
Deux petits résultats simples mais utiles à propos des polynômes trigonométriques.

(1) Si $f \in L^1(S^1)$, alors nous avons la formule

$$f * e_n = c_n(f) e_n. \tag{27.331}$$

(2) Si P est un polynôme trigonométrique et si $f \in L^1(S^1)$ alors $f * P$ est encore un polynôme trigonométrique.

Démonstration. Le premier point est un simple calcul :

$$(f * e_n)(x) = \int_0^{2\pi} f(x - t) e_n(t) \tag{27.332a}$$

En ce qui concerne le second point, nous notons $P = \sum_{k=-N}^{N} P_k e_k$, et par linéarité de la convolution,

$$f * P = \sum_{k=-N}^{N} P_k f * e_k = \sum_{k=-N}^{n} P_k c_k(f) e_k, \tag{27.333}$$

qui est encore un polynôme trigonométrique. □

ExDMnVSWF

Exemple 27.88.
Sur S^1 nous construisons alors l'approximation de l'unité basée sur la fonction $1+\cos(x)$ et le lemme 27.79. Cette fonction est évidemment un polynôme trigonométrique parce que

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}. \tag{27.334}$$

Ensuite les puissances le sont aussi à cause de la formule du binôme :

$$\big(1+\cos(x)\big)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^n(x), \tag{27.335}$$

dans laquelle nous pouvons remettre $\cos(x)$ comme un polynôme trigonométrique et développer à nouveau la puissance avec (encore) la formule du binôme. La chose importante est qu'il existe une approximation de l'unité (φ_n) formée de polynômes trigonométrique.

Ce qui fait la spécificité des polynômes trigonométriques est qu'ils sont à la fois stables par convolution (lemme 27.87) et qu'ils permettent de créer une approximation de l'unité sur $[0, 2\pi]$. Ce sont ces deux choses qui permettent de prouver l'important théorème suivant. △

ThoQGPSSJq

Théorème 27.89.
Les polynômes trigonométriques sont dense dans $L^p(S^1)$ pour $1 \leqslant p < \infty$.

Démonstration.

$$\varphi_k * f \xrightarrow{L^p} f \tag{27.336}$$

par le théorème 27.80. Nous avons donc convergence L^p d'une suite de polynômes trigonométrique, ce qui prouve que l'espace de polynômes trigonométriques est dense dans $L^p(S^1)$. □

Remarque 27.90.
Deux remarques.

— Il n'est pas possible que les polynômes trigonométriques soient dense dans L^∞ parce qu'une limite uniforme de fonctions continues est continue (c'est le théorème 12.357). Donc les polynômes trigonométriques ne peuvent engendrer que des fonctions continues.

— Nous donnerons au théorème 28.8 une démonstration indépendante de la densité des polynômes trigonométriques dans $L^p(S^1)$.

27.7 Espaces L^2, généralités

SECooEVZSooLtLhUm

L'espace L^2 est l'espace L^p défini en 27.7 avec $p = 2$. Cependant il possède une propriété extraordinaire par rapport aux autres L^p, c'est que la norme $|.|_2$ dérive d'un produit scalaire. Il sera donc un espace de Hilbert.

NORMooUEIEooYtlFse

27.91.
Nous en rappelons la construction. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Nous considérons l'opération

$$\langle f, g \rangle = \int_\Omega f(\omega)\overline{g(\omega)} d\mu(\omega) \tag{27.337}$$

DefProdScalLubrgTi

et la norme associée

$$\|f\|_2 = \sqrt{\langle f, f \rangle}. \tag{27.338}$$

Nous considérons l'ensemble

$$\mathcal{L}^2(\Omega, \mu) = \{f : \Omega \to \mathbb{C} \text{ tel que } \|f\|_2 < \infty\} \tag{27.339}$$

et la relation d'équivalence $f \sim g$ si et seulement si $f(x) = g(x)$ pour μ-presque tout x.

Et enfin, nous considérons le quotient

$$L^2(\Omega, \mu) = \mathcal{L}^2(\Omega, \mu)/\sim . \tag{27.340}$$

LemIVWooZyWodb

Lemme 27.92.
Soit un espace mesuré[36] $(\Omega, \mathcal{A}, \mu)$.

(1) Pour tout $f, g \in L^2(\Omega, \mathcal{A}, \mu)$, le produit

$$\langle f, g \rangle = \int_\Omega f \bar{g} \, d\mu \tag{27.341}$$

EQooGLVUooQbPmaX

est bien défini et est un nombre complexe[37].

(2) L'opération $(f, g) \mapsto \langle f, g \rangle$ est un produit hermitien[38].

(3) Le couple $\big(L^2(\Omega, \mathcal{A}, \mu), \langle ., . \rangle\big)$ est un espace de Hilbert[39].

Démonstration. Que $L^2(\Omega)$ soit un espace vectoriel est un cas particulier de la proposition 27.16. Voyons cette histoire de produit scalaire.

(1) **Pour de vraies fonctions** Nous commençons par analyser l'intégrale (27.341) dans le cas où f et g sont des fonctions, c'est-à-dire des représentants d'éléments de L^2.
Dans ce cas, l'inégalité de Hölder (proposition 27.33) avec $p = q = 2$ nous indique que le produit $f\bar{g}$ est un élément de L^1. Par conséquent la formule a un sens.

(2) **Passage aux classes**
Ensuite nous montrons que la formule passe au quotient. Pour cela, nous considérons des fonctions α et β nulles presque partout et nous regardons le produit de $f_1 = f + \alpha$ par $g_1 = g + \beta$:

$$\langle f_1, g_1 \rangle = \int f\bar{g} + \bar{\beta} f + \alpha \bar{g} + \alpha\bar{\beta}. \tag{27.342}$$

Les fonctions $\bar{\beta} f$, $\alpha\bar{g}$ et $\alpha\bar{\beta}$ étant nulles presque partout, leur intégrale est nulle et nous avons bien $\langle f_1, g_1 \rangle = \langle f, g \rangle$. Nous pouvons donc considérer le produit sur l'ensemble des classes.

(3) **Produit hermitien** Pour vérifier que la formule est un produit hermitien, le seul point non évidement est de prouver que $\langle f, f \rangle = 0$ implique $f = 0$. Cela découle du fait que

$$\langle f, f \rangle = \int_\Omega |f|^2. \tag{27.343}$$

La fonction $x \mapsto |f(x)|^2$ vérifie les hypothèses du lemme 14.199. Par conséquent $|f(x)|^2$ est presque partout nulle.

(4) **Espace de Hilbert** En ce qui concerne le fait que $L^2(\Omega)$ soit un espace de Hilbert, il s'agit simplement de se remémorer que c'est un espace complet (théorème 27.44) et dont la norme dérive d'un produit scalaire ou hermitien. Nous sommes donc bien dans la définition 25.3.

□

27.93.
Ces espaces seront utilisés pour de nombreuses applications. Nous en aurons besoin pour plusieurs combinaisons d'ensembles Ω et de mesures μ.

— Pour $\mathbb{R}^d$
— Pour S^1
— Pour $[a, b]$
— Pour $[0, 2\pi[$

36. Est-ce qu'il ne faudrait pas un peu plus d'hypothèses, comme σ-fini par exemple ? Vérifiez et écrivez-moi quand vous avez la réponse.
37. Par opposition au fait que ce serait l'infini.
38. Définition 8.176. Pour rappel, nous considérons des fonctions à valeurs complexes. Si au contraire nous avions considéré seulement des fonctions à valeurs réelles, nous aurions eu un produit scalaire.
39. Définition 25.3.

— Pour $[-T, T[$

Le premier est non compact et il est raisonnable de penser qu'il sera fondamentalement différent des autres. À isomorphismes assez triviaux près, les espaces des fonctions sur les trois autres sont identiques. Nous nous attendons donc à ce qu'ils aient les mêmes propriétés. Notons que du point de vue de L^2, étant donné qu'il y a un quotient par les parties de mesures nulles, prendre $]0, 2\pi[$ ou $[0, 2\pi]$ ou n'importe quelle autre possibilité de ce genre revient au même.

Afin de pouvoir utiliser ces espaces de façon optimale, et entre autres y définir les séries de Fourier, nous avons besoin, pour chacun d'entre eux de définir les éléments suivants :

— mesure

— produit de convolution

— le système trigonométrique (que nous allons montrer être une base hilbertienne)

— coefficients de Fourier

Ça fait pas mal de choses à définir. Il n'est pas besoin de définir un produit scalaire parce que le lemme 27.92 nous en donne un générique.

Les définitions qui viennent sont à prendre « tant que les formules ont un sens ». Nous parlons donc de fonctions dans $\text{Fun}(\Omega, \mathbb{C})$, l'ensemble de toutes les fonctions sur Ω à valeurs dans $\mathbb{C}$. Nous verrons plus tard les espaces de fonctions sur lesquels tout a un sens.

27.8 L'espace $L^2(\mathbb{R}^d)$

La mesure est celle de Lebesgue. Le produit de convolution est donné, pour $f, g \in \text{Fun}(\mathbb{R}^d, \mathbb{C})$, par

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x - y)dy \tag{27.344}$$

Certaines de ses propriétés ont déjà été vues dans le théorème 27.66.

En ce qui concerne le système trigonométrique, pour tout $\xi \in \mathbb{R}^d$ nous définirions bien

$$e_\xi(x) = e^{i\xi \cdot x}, \tag{27.345}$$

genre pour faire que les transformations de Fourier sont des séries continues ...mais bon. Nous n'allons pas tenter le diable plus que ça, et nous ne définissons

— pas de système trigonométrique,

— pas de coefficients de Fourier non plus,

— pas de théorie des séries de Fourier sur $\mathbb{R}^d$.

Quand je disais que la non-compacité de $\mathbb{R}^d$ allait un peu changer les choses par rapport aux autres, je ne rigolais pas.

27.9 L'espace $L^2(S^1)$

L'espace S^1 sera fait avec forces détails, parce qu'il va servir de base pour les espaces $L^2([0, 2\pi[)$, $L^2([-T, T[)$ ainsi que pour l'étude des fonctions périodiques sur $\mathbb{R}$.

En tant qu'ensemble,

$$S^1 = \{e^{it}\}_{t\in\mathbb{R}}, \tag{27.346}$$

sans garanties que ce paramétrage soit une bijection.

Il y a essentiellement deux façons de définir une intégrale sur S^1.

(1) Voir S^1 comme une sous-variété de $\mathbb{R}^2$ et utiliser la définition 20.9. Cette façon a cependant deux inconvénients :

— Elle ne donne pas la tribu des mesurables sur S^1, c'est-à-dire que cette méthode ne donne pas de façon évidente une théorie de la mesure sur S^1.

— Il faut au moins deux cartes pour paramétrer le cercle. La fainéantise nous prévient que ça va être technique.

(2) Rapporter la structure d'espace mesuré de $[0,2\pi[$ vers S^1, de force via le premier difféomorphisme qui nous passe par la tête, à savoir $t \mapsto e^{it}$.

Nous allons choisir la seconde possibilité, en gardant en tête qu'elle fonctionne de façon très simple un peu par coup de chance, voir la remarque 14.284(5).

27.9.1 Espace mesuré

Plusieurs choses sont déjà faites.

— Les boréliens de S^1 sont décrits dans la proposition 18.70,

— la tribu de Lebesgue de S^1 est décrite dans la proposition 18.72. Non, ce n'est pas la tribu induite de la tribu de Lebesgue de $\mathbb{C}$.

PROPooDJERooYirMru

Proposition 27.94 (Espaces de fonctions sur S^1[1])**.**
Soit l'espace mesuré $\big(S^1, \mathcal{L}eb(S^1), \mu\big)$.

(1) La formule

$$\langle f, g\rangle = \int_{S^1} f\bar{g}d\mu \tag{27.347}$$

EQooHPFQooEavjfZ

est un produit hermitien[40] *sur* $L^2(S^1, \mathcal{L}eb(S^1), \mu)$.

(2) L'espace $L^2(S^1)$ *est un espace de Hilbert.*

ITEMooQZAPooKEeQBW

(3) L'application[41]

$$\begin{aligned}\phi\colon L^2(S^1) &\to L^2\big([0,2\pi[\big)\\ f &\mapsto \frac{1}{\sqrt{2\pi}} f\circ\varphi.\end{aligned} \tag{27.348}$$

est une bijection isométrique (isomorphisme d'espaces de Hilbert)

Tout cela se résume par l'égalité

$$L^2(S^1, \mathcal{L}eb(S^1), \mu) = L^2\big(]0,2\pi[, \mathcal{L}eb(\mathbb{R}), \lambda\big) \tag{27.349}$$

où nous avons fait un minuscule abus de notations : ici $\mathcal{L}eb(\mathbb{R})$ *est en réalité la tribu induite sur* $]0,2\pi[$.

Démonstration. Le fait que la formule (27.347) donne bien un produit hermitien est le lemme 27.92. Ce même lemme assure que le tout donne un espace de Hilbert.

Il nous reste à prouver le point (3). En ce qui concerne l'isométrie, nous posons[42]

$$\begin{aligned}\phi\colon L^2(S^1) &\to L^2\big([0,2\pi[\big)\\ f &\mapsto \frac{1}{\sqrt{2\pi}} f\circ\varphi\end{aligned} \tag{27.350}$$

(1) **Injection** Si $\phi(f) = \phi(g)$, alors pour tout $x \in [0,2\pi[$ nous avons $f\big(\varphi(x)\big) = g\big(\varphi(x)\big)$. Vu que $\varphi\colon [0,2\pi[\to S^1$ est une bijection nous avons alors $f(s) = g(s)$ pour tout $s \in S^1$.

(2) **Surjection** Si $f \in L^2\big(]0,2\pi]\big)$, nous posons $g\colon S^1 \to \mathbb{C}$ par

$$g(s) = \sqrt{2\pi} f\big(\varphi^{-1}(s)\big). \tag{27.351}$$

Nous avons alors bien $\phi(g)(x) = f(x)$.

40. Définition 8.176.
41. Ici $\varphi\colon [0,2\pi[\to S^1$ est la bijection usuelle de la proposition 18.61(1).
42. Notez que cette définition passe aux classes. Nous le répéterons pas.

(3) **Isométrie** Nous montrons que ϕ préserve le produit scalaire : SUBEQSooRYYHooPcLXHN

$$\langle \phi(f), \phi(g) \rangle = \int_0^{2\pi} \phi(f)(x)\overline{\phi(g)(x)} d\lambda(x) \tag{27.352a}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} (f \circ \varphi)(x)\overline{(g \circ \varphi)(x)}\, d\lambda(x) \tag{27.352b}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} (f\bar{g}) \circ \varphi \, d\lambda \tag{27.352c}$$

Pour la suite nous devons invoquer la proposition 14.283 qui nous permet de changer une intégrale sur $\big([0, 2\pi[, \mathcal{L}\text{eb}\big([0, 2\pi[\big), \lambda\big)$ en une intégrale sur $\big(S^1, \mathcal{L}\text{eb}(S^1), \mu\big)$. La première condition de cette proposition est que $\mathcal{L}\text{eb}(S^1) = \varphi\big(\mathcal{L}\text{eb}([0, 2\pi[)\big)$. Cela est la proposition 18.72(2). La condition sur la mesure dans la proposition 14.283 n'est vraie ici qu'à un facteur 2π près. Nous avons :

$$\int_{[0,2\pi[} f d\lambda = 2\pi \int_{S^1} (f \circ \varphi^{-1}) d\mu. \tag{27.353}$$

Nous continuons le calcul (27.352) :

$$\langle \phi(f), \phi(g) \rangle = \frac{1}{2\pi} \int_0^{2\pi} (f\bar{g}) \circ \varphi \, d\lambda = \int_{S^1} f\bar{g} d\mu = \langle f, g \rangle. \tag{27.354}$$

□

27.9.2 Topologie

Nous considérons sur S^1 la topologie induite de $\mathbb{C}$. Vu que S^1 est fermé et borné dans $\mathbb{C}$, il en est une partie compacte. Par le lemme 7.84, l'espace S^1 muni de sa topologie est un espace topologique compact.

Nous pouvons donc sans crainte affirmer que toute fonction continue $f : S^1 \to \mathbb{K}$ est bornée et atteint ses bornes.

PROPooEQDBooDfOrTZ

Proposition-Définition 27.95.
Soit la formule

$$d(e^{ix}, e^{iy}) = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi|. \tag{27.355}$$

(1) Elle est bien définie (ne dépend pas des choix de x et y donnant les mêmes points dans S^1)

(2) L'infimum est en réalité un minimum : il est atteint par un certain $k \in \mathbb{Z}$ (qui, lui, dépend des choix).

(3) La formule définit une distance[43] *sur S^1.*

Nous considérons sur S^1 la topologie τ_d découlant de cette distance.

Démonstration. Point par point.

(1) Soient $x', y' \in \mathbb{R}$ tels que $e^{ix'} = e^{ix}$ et $e^{iy'} = e^{iy}$. Alors $x' = x + 2l\pi$ et $y' = y + 2l'\pi$ pour certains entiers $l, l' \in \mathbb{Z}$ (corolaire 18.24). Nous avons alors $|x'-y'+2k\pi| = |x-y+2\pi(k+l-l')|$ et

$$\inf_{k\in\mathbb{Z}} |x' - y' + 2k\pi| = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi|. \tag{27.356}$$

(2) Quels que soient x et y fixés, nous avons

$$\lim_{k\to\pm\infty} |x - y + 2k\pi| = \infty. \tag{27.357}$$

Donc l'infimum est forcément atteint par un $k \in \mathbb{Z}$.

(3) Pour la distance, il y a plusieurs points à prouver.

43. Définition 7.115.

— Pour tout $z, z' \in S^1$ nous avons $d(z, z') \geqslant 0$ parce que la distance est donnée par une valeur absolue.

— Si $d(z, z') = 0$, alors il existe k tel que $x = y + 2k\pi$. Alors $e^{ix} = e^{i(y+2k\pi)} = e^{iy}e^{2ki\pi} = e^{iy}$. C'est-à-dire $z = z'$.

— Pour la symétrie, nous avons

$$|x - y + 2k\pi| = |y - x - 2k\pi| = |y - x + 2k'\pi| \tag{27.358}$$

en posant $k' = -k$. L'infimum étant pris sur $k \in \mathbb{Z}$, nous avons al symétrique $d(e^{ix}, e^{iy}) = d(e^{iy}, e^{ix})$.

— Pour attaquer l'inégalité triangulaire, nous considérons $z_1 = e^{ix_1}$, $z_2 = e^{ix_2}$ et $z_3 = e^{ix_3}$. Nous posons également $k_1, k_2, k_3 \in \mathbb{Z}$ tels que $d(z_1, z_3) = |x_1 - x_3 + 2k_1\pi|$, $d(z_1, z_2) = |x_1 - x_2 + 2k_2\pi|$ et $d(z_2, z_3) = |x_2 - x_3 + 2k_3\pi|$. Nous avons alors

$$\begin{aligned}
d(z_1, z_3) &= \inf_{k\in\mathbb{Z}} |x_1 - x_3 + 2k\pi| && (27.359a)\\
&= \inf_{k\in\mathbb{Z}} |x_1 - x_2 + x_2 - x_3 + 2k\pi| && (27.359b)\\
&= \inf_{k\in\mathbb{Z}} \big|(x_1 - x_2 + 2k_2\pi) + (x_2 - x_3 + 2k_3\pi) + 2k\pi\big| && (27.359c)\\
&\leqslant \big|(x_1 - x_2 + 2k_2\pi) + (x_2 - x_3 + 2k_3\pi)\big| && (27.359d)\\
&\leqslant |x_1 - x_2 + 2k_2\pi| + |x_2 - x_3 + 2k_3\pi|. && (27.359e)
\end{aligned}$$

SUBEQooXOGCooCZxCsf

Notez que pour (27.359d), en posant $k = 0$ nous avons bien une inégalité. En y pensant, $k = 0$ réalise effectivement l'infimum, mais ce n'est pas indispensable.

□

Le cercle est bien connu pour être symétrique et en particulier avoir une symétrie sous les rotations. Nous allons voir quelques résultats qui vont dans le sens de dire que la distance définie sur S^1 respecte cette symétrie.

LEMooCQCAooAEctbe

Lemme 27.96.
Plusieurs points à propos de l'invariance de la topologie sous les rotations.

(1) La distance est invariante sous les rotations, c'est-à-dire que si $a, b \in S^1$ et si $s \in \mathbb{R}$, alors

$$d(e^{is}a, e^{is}b) = d(a, b). \tag{27.360}$$

ITEMooCIPYooTyPQLj

(2) Les boules sont préservées sous les rotations[44]*, c'est-à-dire que*

$$e^{is}B_d(a, r) = B_d(e^{is}a, r). \tag{27.361}$$

(3) La topologie est invariante sous les rotations : $e^{is}\tau_d = \tau_d$.

Démonstration. Point par point.

(1) Si $a = e^{ix}$ et $b = e^{iy}$, nous avons

$$d(a, b) = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi| = \inf_{k\in\mathbb{Z}} |(x - s) - (y - s) + 2k\pi| = d(e^{is}a, e^{is}b). \tag{27.362}$$

Et de un.

(2) Il faut une inclusion dans chaque sens.

(2a) $\underline{e^{is}B_d(a, r) \subset B_d(e^{is}a, r)}$ Soit $b \in e^{is}B_d(a, r)$. Alors $b = e^{is}b'$ pour un certain $b' \in B_d(a, r)$. Nous avons alors, en utilisant le premier point,

$$d(b, ae^{is}) = d(e^{-is}b, a) = d(b', a) < r. \tag{27.363}$$

Donc $b \in B_d(e^{is}a, r)$.

44. Pas chaque boule séparément, mais l'ensemble des boules

(2b) $B_d(e^{is}a, r) \subset e^{is}B_d(a, r)$ Soit $b \in B_d(e^{is}a, r)$. Nous devons prouver que $b \in e^{is}B_d(a, r)$, c'est-à-dire que $b = e^{is}b'$ pour un certain $b' \in B_d(a, r)$ ou encore que $e^{-is}b \in B_d(a, r)$. En utilisant encore le premier point,

$$d(e^{-is}b, a) = d(b, e^{is}a) < r. \tag{27.364}$$

Donc oui, $e^{-is}b \in B(a, r)$.

(3) Soit $A \in \tau_d$. Si $a \in e^{is}A$, alors $a = e^{is}a'$ pour un certain $a' \in A$. Notre but est de prouver que $e^{is}A$ contient un voisinage de a.
Vu que $a' \in A$, il existe $r > 0$ tel que $B_d(a', r) \subset A$. Nous avons alors

$$e^{is}aB_d(a', r) \subset e^{is}A, \tag{27.365}$$

et comme $e^{is}B_d(a', r) = B_d(e^{is}a', r) = B_d(a, r)$ nous avons bien

$$B_d(a, r) \subset e^{is}A. \tag{27.366}$$

□

Nous allons voir maintenant quelques résultats à propos de $B_d(1, r)$ qui a la bonne figure d'être un ouvert qui s'étale symétriquement en partant de 1 (le point le plus à droite du cercle). Par rapport à la figure 27.1, il s'agit ni plus ni moins que de voir qu'une boule de rayon r autour de 1 est bien la partie indiquée (symétrique par rapport à 1 et de longueur d'arc r des deux côtés). De plus, ce voisinage n'est autre que la partie du cercle située à droite de la ligne en pointillés.

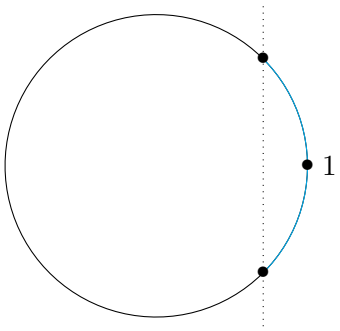


FIGURE 27.1: Un voisinage de 1 dans S^1. LabelFigJOQVoolPTsYPZK

Ce lemme-ci montre que $B_d(1, r)$ est une partie de S^1 qui s'étale symétriquement autour de 1. LEMooMYNVooIWWsiV

Lemme 27.97.
Soit l'application

$$\begin{aligned}\varphi\colon \mathbb{R} &\to S^1\\ x &\mapsto e^{ix}.\end{aligned} \tag{27.367}$$

Nous avons $B_d(1, r) = \varphi\big(]-r, r[\big)$.

Démonstration. Soit $b \in B_d(1, r)$ de la forme $b = e^{iy}$ avec y choisi de telle sorte que $d(1, b) = |y|$. Vu que $d(1, b) < r$, nous avons $|y| < r$ et donc $b \in \varphi\big(]-r, r[\big)$.

Dans l'autre sens, si $y \in]-r, r[$, alors

$$d(1, e^{iy}) = \inf_{k\in\mathbb{Z}} |y + 2k\pi| \leqslant |y| < r. \tag{27.368}$$ EQooRWASooVZnQCJ

Nous avons utilisé le fait que l'infimum sur $k \in \mathbb{Z}$ est plus petit ou égal à la valeur pour $k = 0$. Les inégalités (27.368) montrent que $e^{iy} \in B_d(1, r)$. □

Le lemme suivant montre que que les boules autour de 1 sont délimitées par la droite en pointillé de la figure 27.1.

LEMooLINCooHJmJWx

Lemme 27.98.
Soit $r \in [0, \pi]$. Nous avons

$$B_d(1, r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \tag{27.369}$$

Démonstration. En deux inclusions.

(1) **Inclusion $\subset$**
Si $r = \pi$, alors $B(1, r) = S^1 \backslash \{-1\}$, alors que $\cos(\pi) = -1$.
Si $r < \pi$, alors nous partons de la formule (18.18) qui dit que $e^{ir} = \cos(r) + i \sin(r)$. D'après le lemme 27.97, un élément de $B_d(1, r)$ est de la forme e^{iy} avec $y \in]-r, r[$. Nous voudrions donc prouver que $\cos(y) > \cos(r)$ dès que $y \in]-r, r[$ et $r < \pi$.
Sur $]-r, 0[$, la fonction cos est croissante [45], donc si $y < 0$ alors

$$\cos(y) > \cos(-r) = \cos(r). \tag{27.370}$$

De la même façon, sur $]0, r[$, la fonction cos est décroissante, de telle sorte que si $y > 0$, alors $\cos(y) > \cos(r)$.
Nous avons prouvé que $B_d(1, r) \subset S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}$.

(2) **Inclusion inverse**
Soit $z = e^{is}$ avec $s \in [-\pi, \pi[$, et vérifiant $\cos(s) > \cos(r)$. Nous avons

$$d(1, z) = \inf_{k \in \mathbb{Z}} |s + 2k\pi| = |s| \tag{27.371}$$

parce que l'infimum est réalisé par $k = 0$. Nous avons donc

$$\begin{cases} \cos(s) > \cos(r) & \text{(27.372a)} \\ |s| \in [0, \pi] & \text{(27.372b)} \\ r \in [0, \pi], & \text{(27.372c)} \end{cases}$$

et nous devons en déduire que $|s| < r$. Si $s > 0$, alors la décroissance de cosinus (proposition 18.48(1)) dit que $s < r$.
Si $s < 0$, alors nous savons que $\cos(s) = \cos(-s)$ et donc, en posant $s' = -s$, le cas "$s > 0$" nous dit que $s' < r$, et donc que $|s| < r$.
Dans les deux cas nous avons $d(1, z) = |s| < r$.

□

Proposition 27.99.
La topologie τ_d sur S^1 [46] est la topologique induite depuis $\mathbb{C}$.

Démonstration. Nous notons τ_i la topologie induite (c'est-à-dire l'ensemble des ouverts) et τ_d la topologie de la distance fraichement définie. Nous allons également noter $B_{\mathbb{C}}(a, r)$ la boule dans $\mathbb{C}$ de centre a et de rayon r, et $B_d(z, r)$ celle dans S^1, de centre $z \in S^1$ et de rayon r pour notre distance d.

(1) $\tau_i \subset \tau_d$ Un élément général de τ_i est de la forme $\mathcal{O} \cap S^1$ où $\mathcal{O}$ est un ouvert de $\mathbb{C}$. Soit $a \in \mathcal{O} \cap S^1$ et prouvons qu'il existe un ouvert de τ_d contenant a et contenu dans $\mathcal{O} \cap S^1$; cela prouvera que $\mathcal{O} \cap S^1$ est ouvert de τ_d par le théorème 7.8.
Soient $a = e^{ix}$ et r tel que $B_{\mathbb{C}}(a, r) \subset \mathcal{O}$. Nous allons montrer que $B_d(a, r) \subset B_{\mathbb{C}}(a, r)$. Un élément général de $B_d(a, r)$ est $b = e^{iy}$ tel que

$$d(a, b) = \inf_{k \in \mathbb{Z}} |x - y + 2k\pi| \leqslant r. \tag{27.373}$$

45. Lemme 18.25.
46. Définition 27.95.

Quitte à redéfinir x ou y nous pouvons supposer que l'infimum est atteint en $k = 0$. En utilisant la proposition 21.15 nous majorons :

$$|a - b| = |e^{ix} - e^{iy}| \leqslant |x - y| = d(a, b) \leqslant r. \tag{27.374}$$

Donc nous avons bien $b \in B_{\mathbb{C}}(a, r)$ dès que $b \in B_d(a, r)$.

(2) $\tau_d \subset \tau_i$

Ce sens est plus délicat parce que, si nous voulons suivre les mêmes pas que le premier sens, nous devrons nous appuyer sur la continuité de l'application $\ln\colon \mathbb{C}^* \to \mathbb{C}$, laquelle n'est pas vraie en -1 (voir par exemple le lemme 26.33).

Soient $A \in \tau_d$ et $a \in A$. Nous devons prouver l'existence d'un ouvert $\mathcal{O}$ de $\mathbb{C}$ tel que $S^1 \cap \mathcal{O}$ soit inclus dans A et contienne a. Nous allons prouver cela dans le cas $a = 1$ et ensuite propager le résultat en utilisant la symétrie de S^1.

(2a) **Si $a = 1$** Vu que A est ouvert pour la topologie de la distance d, et vu que $1 \in A$, il existe $r > 0$ tel que $B_d(1, r) \subset A$. Pour ce r le lemme 27.98 donne

$$B_d(1, r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \tag{27.375}$$

Nous montrons que $\mathcal{O} = B_{\mathbb{C}}(1, \delta)$ avec $\delta < 1 - \cos(r)$ fait l'affaire. Si $x + iy \in B_{\mathbb{C}}(1, \delta)$, alors $x > 1 - \delta$ et donc

$$1 - \delta > 1 - (1 - \cos(r)) = \cos(r), \tag{27.376}$$

ce qui prouve que la partie de $\mathcal{O}$ qui est dans S^1 est bien dans $B_d(1, r)$.

(2b) **Si $a \neq 1$**

Soient un ouvert quelconque $A \in \tau_d$ ainsi que $a = e^{ix} \in A$. Nous considérons $r > 0$ tel que $B_d(a, r) \subset A$; nous avons $e^{-ix}B_d(a, r) \subset e^{-ix}A$ et donc, en tenant compte du lemme 27.96(2) :

$$B_d(1, r) \subset e^{-ix}A. \tag{27.377}$$

Par le premier point, il existe un ouvert $\mathcal{O}$ de $\mathbb{C}$ tel que $1 \in \mathcal{O}$ et

$$\mathcal{O} \cap S^1 \subset B_d(1, r) \subset e^{-ix}A. \tag{27.378}$$

Nous avons évidemment que $a \in e^{ix}\mathcal{O}$ et

$$e^{ix}(\mathcal{O} \cap S^1) \subset e^{ix}B_d(1, r) \subset A. \tag{27.379}$$

Donc $e^{ix}\mathcal{O} \cap S^1 \subset A$. Vu que $e^{ix}\mathcal{O}$ est un ouvert de $\mathbb{C}$, l'ensemble $e^{ix}\mathcal{O} \cap S^1$ est un ouvert de τ_i.

□

LEMooTKFHooJaeMyc

Lemme 27.100 ([1]).
Deux résultats de limites dans S^1.

ITEMooEUDIooDuynRg

(1) Pour tout $a_0 \in S^1$, nous avons

$$\lim_{s \to 1} d(a, as) = 0. \tag{27.380}$$

ITEMooXCBUooUxQldB

(2) Si $f\colon S^1 \to \mathbb{C}$ est continue en $a \in S^1$, alors

$$\lim_{s \to 1} f(as) = f(a). \tag{27.381}$$

Démonstration. Point par point.

(1) Soient $a, s \in S^1$. Donnons une formule pour $d(a, as)$. Si $a = e^{ix}$ et $s = e^{iy}$ nous avons $as = e^{i(x+y)}$ et donc

$$d(a, as) = \inf_{k\in\mathbb{Z}} |x - (x + y) + 2k\pi| = \inf_{k\in\mathbb{Z}} |y + 2k\pi|. \tag{27.382}$$

Voilà pour la formule. Maintenant la preuve de notre point.
Soit $\epsilon > 0$. Si $\delta < \epsilon$ et si $s \in B(1, \delta)$, alors il existe $y \in]-\delta, \delta[$ tel que $s = e^{iy}$ par le lemme 27.97. Pour un tel s nous avons

$$d(a, sa) = \inf_{k\in\mathbb{Z}} |y + 2k\pi| \leqslant |y| < \delta < \epsilon. \tag{27.383}$$

Nous avons trouvé $\delta > 0$ tel que $s \in B(1, \delta)$ implique $d(a, as) < \epsilon$. Cela est la limite que nous devions prouver.

(2) Soit $\epsilon > 0$. Soit $r > 0$ tel que si $b \in B(a, r)$, alors $|g(b) - g(a)| < \epsilon$; l'existence d'un tel r est la continuité de g en a. Nous considérons $\delta > 0$ tel que $s \in B(1, \delta)$ implique $sa \in B(a, r)$; l'existence d'un tel δ est le point (1) de ce lemme.
Avec tout cela nous avons $|g(as) - g(a)| < \epsilon$ dès que $s \in B(1, \delta)$. Nous avons donc, comme nous le voulions, la limite $\lim_{s\to 1} g(as) = g(a)$.

□

Lemme 27.101.
Pour $s \in S^1$, nous considérons l'application

$$\begin{aligned} \alpha_s\colon \operatorname{Fun}(S^1) &\to \operatorname{Fun}(S^1) \\ \alpha_s(g)(u) &= g(u\bar{s}) - g(u). \end{aligned} \tag{27.384}$$

Quelques propriétés avec $1 \leqslant p < \infty$:

(1) Si $f \in L^p(S^1)$, alors $\alpha_s(f) \in L^p(S^1)$.

(2) Si f est continue dans $L^p(S^1)$ nous avons la limite

$$\lim_{s\to 1} \alpha_s(f) = 0 \tag{27.385}$$

dans $L^p(S^1)$.

Démonstration. D'abord un calcul de norme :

$$\|\alpha_s(f)\|_p^p = \int_{S^1} |f(u\bar{s}) - f(u)|^p du \leqslant \int_{S^1} |f(u\bar{s})|^p du + \int_{S^1} |f(u)|^p du = 2\|f\|_p^p. \tag{27.386}$$

Donc oui pour que $\alpha_s(f) \in L^p(S^1)$.

La fonction f étant supposée continue sur le compact S^1, elle est majorée. Nous savons qu'en posant $M = \|f\|_\infty$ nous avons $|\alpha_s(f)| \leqslant 2M$. Donc la fonction constante

$$\begin{aligned} g\colon S^1 &\to \mathbb{C} \\ u &\mapsto 2M \end{aligned} \tag{27.387}$$

est une fonction intégrable sur S^1 qui majore $|\alpha_s(f)|$ uniformément en s. Soit une suite $s_i \to 1$ dans S^1, et posons $f_i = \alpha_{s_i}(f)$. Alors nous avons

$$\|f_i\|_p^p = \int_{S^1} |f_i(u)|^p du \tag{27.388}$$

et aussi $|f_i|^p \leqslant (2M)^p$. Le théorème de la convergence dominée de Lebesgue 14.207 nous permet de permuter limite et intégrale :

$$\lim_{i\to\infty} \|f_i\|_p^p = \int_{S^1} \lim_{i\to\infty} |f_i(u)|^p du. \tag{27.389}$$

Mais

$$\lim_{i\to\infty} f_i(u) = \lim_{i\to\infty} \alpha_{s_i}(f)(u) = \lim_{i\to\infty} \big(f(u\bar{s}_i) - f(u)\big) = 0. \tag{27.390}$$

La dernière limite est due au fait que $\lim_{s\to 1} g(us) = g(u)$ (lemme 27.100(2)). □

27.9.3 Système trigonométrique

Définition 27.102.
*La **famille trigonométrique** sur S^1 est l'ensemble de fonctions $\{e_n\}_{n\in\mathbb{Z}}$ données par*

$$\begin{aligned} e_n\colon S^1 &\to \mathbb{C} \\ z &\mapsto z^n \end{aligned} \tag{27.391}$$

*avec $n\in\mathbb{Z}$. Un **polynôme trigonométrique** est une application $S^1\to\mathbb{C}$ de la forme*

$$\sum_{k=-n}^{n} a_k e_k \tag{27.392}$$

pour des nombres $a_k\in\mathbb{C}$, peut-être pas tous non-nuls (autrement dit, il n'est pas forcé d'avoir autant de termes négatifs que positifs).

Le but de $z\mapsto z^n$ dans cette définition est d'être lu $t\mapsto e^{int}$ lorsqu'on considère les fonctions sur $[0,2\pi[$.

PROPooOMGFooROFFFr

Proposition 27.103.
La famille trigonométrique est une famille orthonormale pour le produit scalaire $L^2(S^1,\mathcal{A},\mu)$.

Démonstration. En utilisant la proposition 27.94(3) nous avons :

$$\langle e_n, e_n\rangle = \int_{S^1} e_n\overline{e_n} = \frac{1}{2\pi}\int_{[0,2\pi[} e_n\big(\varphi(x)\big)\overline{e_n\big(\varphi(x)\big)} \tag{27.393a}$$

$$= \frac{1}{2\pi}\int_{[0,2\pi[} e^{inx}e^{-inx}dx \tag{27.393b}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} 1dx \tag{27.393c}$$

$$= 1. \tag{27.393d}$$

Et nous avons également, pour $m\neq n$:

$$\langle e_n, e_m\rangle = \frac{1}{2\pi}\int_{[0,2\pi[} e^{i(n-m)x}dx = \frac{1}{2\pi}\left[\frac{1}{i(n-m)e^{i(n-m)x}}\right]_0^{2\pi} = 0. \tag{27.394}$$

□

Remarque 27.104.
Vous aurez noté que le facteur $\frac{1}{2\pi}$ qui permet d'avoir $\langle e_n, e_n\rangle = 1$ ne provient ni de la définition du produit scalaire ni de celle de la famille trigonométrique, mais bien de la mesure, voir la définition 18.193.

REMooUCANooVyXPxj

Remarque 27.105.
Notez aussi que nous avons bien $\langle e_n, e_{-n}\rangle = 0$. Il faut donc bien prendre tous les e_n avec $n\in\mathbb{Z}$ et non seulement $n\in\mathbb{N}$.

PROPooTGBHooXGhdPR

Proposition 27.106.
Les polynômes trigonométriques forment une partie dense dans $\big(C(S^1,\mathbb{C}),\|.\|_\infty\big)$.

Démonstration. Pour préciser les notations, $C(S^1,\mathbb{C})$ est l'ensemble des fonctions continues de S^1 vers $\mathbb{C}$, et l'espace topologique que nous considérons est cet ensemble sur lequel nous considérons la distance supremum.

Nous utilisons le théorème de Stone-Weierstrass 12.418.

Le système contient une fonction constante non nulle, à savoir e_0.

Il sépare les points grâce à la fonction e_1 qui n'est autre que la fonction identité $z \mapsto z$. De plus l'ensemble des polynômes trigonométriques est stable par conjugaison parce que si

$$P = \sum_{k=-n}^{n} a_k e_k, \tag{27.395}$$

alors $\bar{P} = \sum_{k=-n}^{n} \overline{a_k e_k} = \sum_{k=-n}^{n} \overline{a_k} e_{-k}$ qui est encore un polynôme trigonométrique. □

Définition 27.107.
*Si nous avons une fonction $f : S^1 \to \mathbb{C}$, nous définissons ses **coefficients de Fourier** par*

$$c_n(f) = \langle f, e_n \rangle \tag{27.396}$$

pourvu que l'intégrale existe.

27.9.4 Convolution

La convolution sur $\mathbb{R}^n$ est donnée par la définition 27.65. Nous voyons maintenant comment cela s'adapte à S^1.

DEFooSKWOooEdIHoH

Définition 27.108.
*Si f et g sont des fonctions sur S^1 à valeurs dans $\mathbb{C}$, nous définissons la **convolution** de f et g comme étant la fonction sur S^1 définie par*

$$(f * g)(z) = \int_{S^1} f(s) g(z\bar{s}) d\mu(s). \tag{27.397}$$

EQooILQNooBKtSRj

Cette définition appelle plusieurs remarques.

— Dès que $z, s \in S^1$, nous avons $z\bar{s} \in S^1$, de telle sorte qu'au moins l'intégrande ait un sens.
— Nous ne prétendons pas que l'intégrale (27.397) converge pour toutes les fonctions f et g. Cela est une définition « pour tous les couples f, g pour lesquels l'intégrale fonctionne ».
— Le lemme 27.109 nous dira que $L^1(S^1)$ est stable par convolution : si f et g sont dans L^1, alors $f * g$ y est aussi.
— Dans la formule (27.397), la variable s est vraiment une variable muette. Cette formule aurait également pu être écrite

$$(f * g)(z) = \int_{S^1} \big[s \mapsto f(s) g(z\bar{s}) \big] d\mu. \tag{27.398}$$

LEMooTYSSooItOiYE

Lemme 27.109 ([681]).
Si $f, g \in L^1(S^1)$, alors pour presque tout $z \in S^1$, la fonction $s \mapsto \int f(s) g(z\bar{s})$ est dans $L^1(S^1)$.

Démonstration. Nous considérons la fonction

$$\begin{aligned} \psi : S^1 \times S^1 &\to \mathbb{C} \\ (z, s) &\mapsto f(s) g(z\bar{s}). \end{aligned} \tag{27.399}$$

(1) $\psi \in L^1(S^1 \times S^1)$ en chaine[47] : Nous utilisons le corolaire 14.303, et pour cela nous calculons les intégrales

$$\int_{S^1} \left[\int_{S^1} |f(s) g(z\bar{s})| dz \right] ds = \int_{S^1} |f(s)| \underbrace{\left[\int_{S^1} |g(z\bar{s})| dz \right]}_{=A<\infty} ds \tag{27.400a}$$

$$= A \int_{S^1} |f(s) ds| \tag{27.400b}$$

$$< \infty. \tag{27.400c}$$

47. Dans les expressions suivantes, les symboles « ds » et « dz » n'ont pas d'autres valeurs que purement de notation pour indiquer le nom de la variable d'intégration.

Le fait que $A < \infty$ provient directement de l'hypothèse $g \in L^1(S^1)$ [48].

Par le corolaire sus-cité nous avons bien $\psi \in L^1(S^1 \times S^1)$.

(2) **Et par Fubini** Le théorème de Fubini 14.304(1) nous renseigne que pour presque tout $z \in S^1$, l'application

$$s \mapsto \psi(z, s) \tag{27.401}$$

est dans $L^1(S^1)$. Et la partie 14.304(2) ajoute que l'application

$$z \mapsto \int_{S^1} \psi(s, z) ds \tag{27.402}$$

EQooPLLBooJZsZzu

est également $L^1(S^1)$.

(3) **Conclusion** L'application donnée en (27.402) est précisément $(f * g)$. Donc $f * g \in L^1(S^1)$.

□

Lemme 27.110 ([1]).
Si $f \in L^1(S^1)$ et si g est continue sur S^1, alors $f * g$ existe et est continue sur S^1.

Démonstration. Vu que S^1 est compact, la continuité de g implique que g est bornée et donc dans $L^1(S^1)$. Le lemme 27.109 dit alors que $f * g$ est bien définie sur S^1.

Soit $z_0 \in S^1$. Nous montrons que $f * g$ est continue en z_0 ; pour cela nous considérons $\epsilon > 0$ et ensuite nous réfléchissons un peu.

Vu que g est continue sur S^1 qui est compact, g y est uniformément continue par le théorème de Heine12.78. Il existe donc un $\delta > 0$ tel que pour tout $z_0 \in S^1$, si $z \in B(z_0, \delta)$, alors $|g(z_0) - g(z)| < \epsilon$.

Soit $s \in S^1$. Si $z \in B(z_0, \delta)$, alors $\bar{s}z \in B(\bar{s}z_0, \delta)$ par le lemme 27.96(2). Dans ce cas nous avons aussi

$$|g(\bar{s}z_0) - g(\bar{s}z)| < \epsilon. \tag{27.403}$$

Un peu de calcul maintenant. D'une part

$$(f * g)(z_0) - (f * g)(z) = \int_{S^1} f(s)\big(g(z_0\bar{s}) - g(z\bar{s})\big) ds, \tag{27.404}$$

et donc

$$|(f * g)(z_0) - (f * g)(z)| \leqslant \int_{S^1} |f(s)| \underbrace{\big|g(z_0\bar{s}) - g(z\bar{s})\big|}_{<\epsilon} ds \tag{27.405a}$$

$$\leqslant \epsilon \int_{S^1} |f| \tag{27.405b}$$

$$= A\epsilon \tag{27.405c}$$

pour une certaine constante A ne dépendant pas de z_0.

Nous avons prouvé que pour tout $\epsilon > 0$, il existe un δ tel que $z \in B(z_0, \alpha)$ implique

$$|(f * g)(z_0) - (f * g)(z)| \leqslant A\epsilon, \tag{27.406}$$

ce qui signifie que $f * g$ est continue en z_0. □

Notez que dans cette démonstration, l'uniforme continuité de g a été utilisée pour effectuer d'un seul coup la majoration pour tout s dans l'intégrale.

PROPooCSRNooDyClBY

Proposition 27.111.
Si $f \in L^1(S^1)$, nous avons

$$f * e_n = c_n(f) e_n. \tag{27.407}$$

48. Avec un changement de variables $z \mapsto z\bar{s}$ que je vous conseille d'être capable de justifier.

Démonstration. Il s'agit d'un bon calcul. En considérant $z = e^{i\theta}$ nous avons

$$\begin{aligned}(f * e_n)(z) &= \int_{S^1} f(s)e_n(z\bar{s})ds && (27.408a)\\ &= \frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})e_n(e^{i(\theta-x)})dx && (27.408b)\\ &= e^{in\theta}\frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})e^{-inx}dx && (27.408c)\\ &= e_n(e^{i\theta})\frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})\overline{e_n(e^{ix})}dx && (27.408d)\\ &= e_n(z)\int_{S^1} f(s)\overline{e_n(s)}ds && (27.408e)\\ &= e_n(z)\langle f, e_n\rangle. && (27.408f)\end{aligned}$$

Donc $(f * e_n)(z) = \langle f, e_n\rangle e_n(z)$, c'est-à-dire que

$$f * e_n = c_n(f)e_n. \tag{27.409}$$

□

LEMooDGHJooRAnwpy

Lemme 27.112.
Si P est un polynôme trigonométrique et si $f \in L^1(S^1)$, alors $f * P$ est également un polynôme trigonométrique.

Démonstration. Soit $P = \sum_{k=-n}^{n} a_k e_k$. Par la linéarité du produit de convolution,

$$f * P = \sum_{k=-n}^{n} a_k f * e_k = \sum_k a_k c_k(f)e_k \tag{27.410}$$

où nous avons également utilisé la proposition 27.111. Nous avons donc un polynôme trigonométrique dont les coefficients sont $a_k c_k(f)$ au lieu de a_k. □

27.9.5 Approximation de l'unité

LEMooUNFBooRCzwIn

Lemme 27.113 ([662])**.**
Soient une fonction continue $f\colon S^1 \to [0,\infty[$ et $a \in S^1$ telle que $f(z) < f(a)$ pour tout $z \in S^1\backslash\{a\}$. Alors la suite de fonctions $f_n\colon S^1 \to \mathbb{R}$ donnée par

$$f_n(z) = \left(\int_{S^1} f^n\right)^{-1} f(z)^n \tag{27.411}$$

EQooTQYBooLZprJj

est une approximation de l'unité[49] *autour de a.*

Démonstration. En plusieurs points, dont d'abord une série de vérifications pour voir que la formule a un sens.

(1) **Strictement positive** D'abord, vu que f prend ses valeurs dans $[0,\infty[$ et vu que $f(z) < f(a)$, nous avons $f(a) > 0$ (strict). Peut-être que f s'annule à certains endroits de S^1, mais pas a.

(2) **f^n est intégrable sur S^1** La fonction f^n est dans les hypothèses de la proposition 14.218 parce que S^1 est compact, f^n y est continue et la mesure sur S^1 est compatible avec la topologie (voir les hypothèses précises).

49. Définition 27.77.

(3) **L'intégrale n'est pas nulle** Vu que $f(a) > 0$, il existe un ouvert A contenant a sur lequel $f > 0$. Nous avons alors

$$\int_K f^n \geqslant \int_A f^n > 0. \tag{27.412}$$

Cela pour dire que l'inverse dans (27.411) ne pose pas de problèmes.

(4) **Norme** Vu que toutes les fonctions tant f que f_n sont positives, les valeurs absolues ne jouent aucun rôle et nous avons

$$\|f_n\|_1 = \int_{S^1} f_n d\mu = \left(\int_{S^1} f^n\right)^{-1} \int_{S^1} |f(z)|^n d\mu = 1. \tag{27.413}$$

Ce calcul donne d'un seul coup les deux conditions

— $\sup_k \|f_k\| = 1$

— $\int_{S^1} f_n = 1$ pour tout n.

Nous passons maintenant au vrai travail. Soit un voisinage V de a dans S^1. Soit une suite croissante (t_k) qui converge vers $f(a)$, c'est-à-dire $0 < t_k < f(a)$. Nous posons

$$A_k = \{x \in S^1 \backslash V \text{ tel que } f(x) \geqslant t_k\}. \tag{27.414}$$

Cet ensemble est contenu dans S^1 et est donc borné (pour la métrique de S^1).

(1) **A_k est fermé** Attention : ici nous démontrons que A_k est fermé dans S^1, et les complémentaires sont pris dans S^1.

Nous montrons que le complémentaire est ouvert en prenant $y \in A^c$ et en montrant que y admet un voisinage contenu dans A^c (le fameux théorème 7.8 que nous ne nous lasserons jamais de citer). Si $y \in A^c$, il y a deux possibilités (non exclusives) : soit $y \in V$ soit $f(y) < t_k$. Si $y \in V$, alors le voisinage V lui-même est encore dans A^c. Si par contre $f(y) < t_k$, alors par continuité, il existe un voisinage de y sur lequel $f < t_k$.

(2) **A_k est compact** L'espace S^1 est compact, par exemple grâce au lemme 7.84. La partie A_k est fermée dans le compact S^1, donc elle est compacte par le lemme 7.96.

(3) **$A_{k+1} \subset A_k$** Si $x \in A_{k+1}$, alors $f(x) \geqslant t_{k+1} > t_k$. Donc $f(x) > t_k$ et $x \in A_k$.

(4) **Intersection vide** Si $x \in \bigcap_{k\in\mathbb{N}} A_k$, alors $f(x) \geqslant t_k$ pour tout k. En passant à la limite et en sachant que $\lim_{k\to\infty} t_k = f(a)$, nous avons $f(x) \geqslant f(a)$. Par hypothèse, cela n'est pas. Donc

$$\bigcap_{k\in\mathbb{N}} A_k = \varnothing. \tag{27.415}$$

Nous avons, dans un compact, des fermés emboîtés dont l'intersection est vide. Le corolaire 7.93 nous dit qu'il existe un indice à partir duquel tous les A_k sont vides.

Soit $\delta = t_k$ pour un k tel que A_k est vide. Nous avons

$$\{x \in S^1 \backslash V \text{ tel que } f(x) \geqslant \delta\} = \varnothing, \tag{27.416}$$

c'est-à-dire que sur $S^1\backslash V$, nous avons $f < \delta$ et donc

$$\int_{S^1\backslash V} f(s)^n < \int_{S^1\backslash V} \delta^n = \text{Vol}(S^1\backslash V)\delta^n \tag{27.417}$$

où $\text{Vol}(S^1\backslash V) = \int_{S^1\backslash V} 1 = \mu'(S^1\backslash V)$ est une constante réelle strictement positive.

Nous avons aussi $\delta < f(a)$ parce que δ est un des t_k (et que cette suite croissante converge vers $f(a)$ sans l'atteindre par hypothèse). Soit δ_1 tel que $\delta < \delta_1 < f(a)$ [50].

Nous posons

$$W = \{x \in S^1 \text{ tel que } f(x) > \delta_1\}. \tag{27.418}$$

50. Dans [662], il prend $\delta < \delta_1 < 1$ et je crois qu'il aurait dû écrire $\varphi(0)$ au lieu de 1.

Cet ensemble n'est pas vide parce qu'il contient a et est ouvert parce que f est continue. Nous avons

$$\int_{S^1} f(s)^n ds \geqslant \int_W f(s)^n ds \geqslant \delta_1^n \operatorname{Vol}(W). \tag{27.419}$$

Nous avons donc déjà ces deux inégalités :

$$\int_{S^1\backslash V} f(s)^n \leqslant \operatorname{Vol}(S^1\backslash V)\delta^n \tag{27.420}$$

et

$$\int_{S^1} f(s)^n ds \geqslant \delta_1^n \operatorname{Vol}(W). \tag{27.421}$$

En ce qui concerne les fonctions f_n que nous voulions étudier,

$$f_n(z) = \left(\int_{S^1} f(s)^n ds\right)^{-1} f(z)^n \leqslant \big(\operatorname{Vol}(W)\delta_1^n\big)^{-1} f(z)^n, \tag{27.422}$$

et donc

$$\int_{S^1\backslash V} f_n \leqslant \big(\operatorname{Vol}(W)\delta_1^n\big)^{-1} \operatorname{Vol}(S^1\backslash V)\sigma^n = \frac{\operatorname{Vol}(S^1\backslash V)}{\operatorname{Vol}(W)} \left(\frac{\delta}{\delta_1}\right)^n. \tag{27.423}$$

Étant donné que $\delta < \delta_1$, nous avons $(\delta/\delta_1)^n \to 0$. Donc aussi

$$\lim_{n\to\infty} \int_{S^1\backslash V} f_n(z) dz = 0. \tag{27.424}$$

□

Le théorème suivant est une version pour S^1 du théorème 27.80. Le produit de convolution dans S^1 est la définition 27.108.

THOooIAOPooELSNxq

Théorème 27.114 ([662, 1])**.**
Soient (φ_k) une approximation de l'unité sur $\Omega = S^1$ ainsi qu'une fonction $g\colon S^1 \to \mathbb{C}$. Si g est mesurable et bornée sur Ω et si g est continue en a_0 alors

$$(\varphi_k * g)(a_0) \to g(a_0). \tag{27.425}$$

Démonstration. Pour chaque $k \in \mathbb{N}$ nous posons

$$d_k = \varphi_k * g - g. \tag{27.426}$$

Le but de ce théorème est de montrer que $d_k \to 0$ pour diverses notions de convergence.

Soit $a_0 \in S^1$. Par définition de l'approximation de l'unité, $\int_{S^1} \varphi_k = 1$ et donc on peut écrire $g(a_0) = \int_{S^1} g(a_0)\varphi_k(s)ds$. En ce qui concerne $d_k(a_0)$ nous avons alors

$$d_k(a_0) = \int_{S^1} \varphi_k(s) g(a_0\bar{s}) ds - \int_{S^1} g(a_0)\varphi_k(s) ds \tag{27.427a}$$

$$= \int_{S^1} \varphi_k(s)\big(g(a_0\bar{s}) - g(a_0)\big). \tag{27.427b}$$

Nous pouvons passer à la norme (et non la valeur absolue parce que d_k prend ses valeurs dans $\mathbb{C}$) :

$$|d_k(a_0)| \leqslant \int_{S^1} |\varphi_k(s)|\big|g(a_0\bar{s}) - g(a_0)\big| ds. \tag{27.428}$$

La définition d'une approximation de l'unité nous permet de considérer $M = \sup_k \|\varphi_k\|_1 < \infty$. Le lemme 27.100(2) nous permet, lui, de considérer $\alpha > 0$ tel que

$$|g(a_0\bar{s}) - g(a_0)| < \epsilon \tag{27.429}$$

dès que $s \in B(1,\alpha)$ [51]. Vu que la suite (φ_k) est une approximation de l'unité, nous avons

$$\lim_{k\to\infty} \int_{S^1\setminus B(1,\alpha)} |\varphi_k| = 0. \tag{27.430}$$

Soit k suffisamment grand pour avoir $\int_{S^1\setminus B(1,\alpha)} |\varphi_k| < \epsilon$. Avec tout cela nous avons les majorations

$$\begin{aligned}
|d_k(a_0)| &\leqslant \int_{S^1} |\varphi_k(s)|\big|g(a_0\bar{s}) - g(a_0)\big| ds && (27.431a)\\
&= \int_{S^1\setminus B(1,\alpha)} |\varphi_k(s)| \underbrace{\big|g(a_0\bar{s}) - g(a_0)\big|}_{\leqslant 2\|g\|_\infty} ds + \int_{B(1,\alpha)} |\varphi_k(s)| \underbrace{\big|g(a_0\bar{s}) - g(a_0)\big|}_{<\epsilon} ds && (27.431b)\\
&\leqslant 2\|g\|_\infty \int_{S^1\setminus B(1,\alpha)} |\varphi_k(s)| ds + \epsilon \int_{S^1} |\varphi_k(s)| ds && (27.431c)\\
&\leqslant \epsilon\big(1 + 2\|g\|_\infty\big). && (27.431d)
\end{aligned}$$

Nous avons donc bien $\lim_{k\to\infty} |d_k(a_0)| = 0$ et donc la continuité de $\varphi_k * g$ en a_0.

□

Voici une version un peu forte sous l'hypothèse de continuité. Vu que S^1 est compact, la continuité est en réalité une hypothèse assez forte : ça implique l'uniforme continuité et l'existence d'un maximum et d'un minimum.

Proposition 27.115.
Soient (φ_k) une approximation de l'unité sur $\Omega = S^1$ ainsi qu'une fonction continue $g\colon S^1 \to \mathbb{C}$.
ITEMooPHBJooOHDVoW

(1) Si $g \in L^p(\Omega)$ $(0 \leqslant p < \infty)$ et si g est continue, alors [52]

$$\varphi_k * g \xrightarrow{L^p} g. \tag{27.432}$$

ITEMooLOSVooDtaugF

(2) Si g est continue sur S^1, alors

$$\varphi_k * g \xrightarrow{L^\infty} g \tag{27.433}$$

Démonstration. En plusieurs points

(1) $\lim_{u\to 1} \|\tau_u(g) - g\| = 0$ Ceci est un petit point intermédiaire. Pour des besoins de notations, nous posons

$$\begin{aligned}
\tau_u(g)\colon S^1 &\to \mathbb{C}\\
s &\mapsto g(s\bar{u})
\end{aligned} \tag{27.434}$$

pour $u \in S^1$.
La fonction g est continue sur le compact S^1, et y est donc uniformément continue [53]. Nous allons en déduire que $\lim_{u\to 1} \|\tau_u(g) - g\|_\infty = 0$.
Soit $\epsilon > 0$. L'uniforme continuité de g signifie qu'il existe $\delta > 0$ tel que pour tout $a \in S^1$ si $s \in B(a,\delta)$, alors $\big|g(s) - g(a)\big| < \epsilon$. Si $u \in B(1,\delta)$ nous avons aussi $\bar{u} \in B(1,\delta)$ et donc $s\bar{u} \in B(s,\delta)$; ça c'est le lemme 27.96(2).
Pour tout $a \in S^1$ nous avons la chaine

$$u \in B(1,\delta) \Rightarrow a\bar{u} \in B(a,\delta) \Rightarrow \big|g(a\bar{u}) - g(a)\big| < \epsilon. \tag{27.435}$$

Cela étant valable pour tout a, c'est encore valable en passant au supremum [54] :

$$u \in B(1,\delta) \Rightarrow \sup_{a\in S^1} \big|g(a\bar{u}) - g(a)\big| \leqslant \epsilon \tag{27.436}$$

51. Notez que $s \in B(1,\alpha)$ si et seulement si $\bar{s} \in B(1,\alpha)$. Il n'y a donc pas d'incohérence entre l'hypothèse sur s et notre condition sur $g(a_0\bar{s})$
52. Vous noterez les $p \in]0,1[$ en bonus par rapport au cas de $\mathbb{R}^n$.
53. Théorème de Heine 12.78. C'est fondamentalement ce fait qui unifie les parties (1) et (2) de cette preuve.
54. Notez l'inégalité qui n'est plus stricte.

et donc d'accord pour

$$\lim_{u\to 1}\|\tau_u(g)-g\|=0. \tag{27.437}$$

(2) $\|d_k\|_\infty\to 0$ Nous prouvons la convergence uniforme sur S^1 de d_k vers zéro. Ensuite nous verrons que la compacité de S^1 permet d'en déduire les points (1) et (2).
En utilisant la notation τ_u, nous pouvons écrire

$$d_k(s)=\int_{S^1}\varphi_k(u)g(s\bar{u})du-g(s)=\int_{S^1}\varphi_k(u)\big(\underbrace{g(s\bar{u})}_{=\tau_u(g)(s)}-g(s)\big)du, \tag{27.438}$$

et donc

$$|d_k(s)|\leqslant\int_{S^1}|\varphi_k(u)|\|\tau_u(g)-g\|_\infty du. \tag{27.439}$$

Nous posons $M=\sup_k\|\varphi_k\|_1$, et nous considérons δ tel que $\|\tau_u(g)-g\|_\infty<\epsilon$ pour tout $u\in B(1,\delta)$. Ensuite nous subdivisons S^1 en $B(1,\delta)$ et $B(1,\delta)^c$:

$$\|d_k\|_\infty\leqslant\int_{B(1,\delta)}|\varphi_k(u)|\underbrace{\|\tau_u(g)-g\|_\infty}_{\leqslant\epsilon}du+\int_{B(1,\delta)^c}|\varphi_k(u)|\|\tau_u(g)-g\|_\infty du \tag{27.440a}$$

$$\leqslant\epsilon M+2\|g\|_\infty\int_{B(1,\delta)^c}|\varphi_k(u)|du \tag{27.440b}$$

$$\leqslant\epsilon\big(M+2\|g\|_\infty\big) \tag{27.440c}$$

parce que pour chaque $s\in S^1$ nous avons $\tau_u(g)(s)-g(s)$ et donc $\|\tau_u(g)-g\|_\infty\leqslant 2\|g\|_\infty$.

Tout cela montre que $d_k\xrightarrow{\|.\|_\infty}0$.

(3) **Convergence L^p, $0<p<\infty$**
Soit $\epsilon>0$. Nous avons

$$\|d_k\|_p^p\leqslant\int_{S^1}\big|(\varphi_k*g)(s)-g(s)\big|^p ds. \tag{27.441}$$

Il existe un k à partir duquel $\|\varphi_k*g-g\|_\infty<\epsilon$. Pour de tels k nous avons

$$\|d_k\|_p^p<\epsilon^p. \tag{27.442}$$

Ce passage est très possible dans le cas de S^1 parce que $\int_{S^1}1=1$. Dans le cas de $\mathbb{R}^d$, c'est pas du tout bon ; c'est pour cela que nous avons un résultat un peu plus fort dans S^1. La croissance de la fonction puissance (proposition 12.403) nous permet de conclure que $\|d_k\|_p<\epsilon$.
Nous avons donc la convergence L^p pour $0<p<\infty$.

(4) **Convergence L^∞**
Non, la convergence L^∞ n'est pas la convergence pour la norme $\|.\|_\infty$. Voir la sous-section 27.3.3. Il n'en reste pas moins que si $\epsilon>0$ et si k est assez grand pour que $\|f_k-f\|_\infty<\epsilon$, nous aurons

$$N_\infty(f_k-f)\leqslant\|f_k-f\|<\epsilon. \tag{27.443}$$

□

27.9.6 Base hilbertienne (suite des polynômes trigonométriques)

Voici le plan pour la suite :

— Construire un polynôme trigonométrique qui vérifie les hypothèse du lemme 27.113.
— En déduire une approximation de l'unité constituée de polynômes trigonométriques.
— Dire que si $f\in L^2(S^1)$, alors $f*\varphi_k$ est un polynôme trigonométrique dès que φ_k en est un.

— Invoquer le théorème 27.114(1) pour déduire que $\varphi_k * f$ est une suite de polynômes trigonométriques dans $L^2(S^1)$ qui converge $\varphi_k * f \xrightarrow{L^2} f$.

LEMooQQILooWlhntZ

Lemme 27.116.
La fonction $P = e_1 + e_{-1}$ est continue à valeurs réelles sur S^1.

Démonstration. Nous avons $e_1(z) = z$ et $e_{-1}(z) = z^{-1}$, c'est-à-dire que pour $z = e^{ix}$ ($x \in \mathbb{R}$), nous avons $e_{-1}(e^{ix}) = e^{-ix}$, de telle sorte que, en utilisant le lemme 18.11 qui donne e^{ix} en termes des fonctions trigonométriques usuelles :

$$(e_1 + e_{-1})(e^{ix}) = e^{ix} + e^{-ix} = \cos(x) + i\sin(x) + \cos(x) - i\sin(x) = 2\cos(x). \tag{27.444}$$

Nous avons donc la continuité et les valeurs réelles. □

LEMooIDTVooYTpfEm

Lemme 27.117.
Il existe un polynôme trigonométrique à valeurs dans $[0, \infty[$ et tel que $f(a) < f(1)$ pour tout $a \neq 1$ dans S^1.

Démonstration. Le lemme 27.116 nous dit déjà que $P = e_1 + e_{-1}$ est continue à valeurs réelles. Or qui est continue sur un compact (ici S^1), atteint donc ses bornes. Il est donc facile de considérer [55] $M > 0$ tel que $Q = M + e_1 + e_{-1}$ est à valeurs dans $[0, \infty[$.

Une forme explicite de Q est que

$$Q(e^{ix}) = M + 2\cos(x). \tag{27.445}$$

Le maximum de $\cos(x)$ est obtenu en $x = 0$ et vaut 1. Le maximum de Q est alors $Q(1) = 2 + M$. Il n'est atteint qu'une seule fois sur S^1 parce que pour avoir $Q(e^{ix}) = 2 + M$, il faut avoir $2\cos(x) = 2$, c'est-à-dire $x = 2k\pi$. Mais $e^{i2k\pi} = 1$.

Donc $Q(a) < M + 2 = Q(1)$ pour tout $a \neq 1$ dans S^1. □

Proposition 27.118.
Les polynômes trigonométriques $\{e_n\}_{n\in\mathbb{Z}}$ forment une base hilbertienne de $L^2(S^1)$.

Démonstration. Le fait que les e_n soient orthonormée est la proposition 27.103. Il reste à prouver que ce soit un système total.

Soit $f \in L^2(S^1)$. Soit un polynôme Q vérifiant le lemme 27.117 ; nous posons

$$\varphi_k(z) = \left(\int_{S^1} Q^n \right)^{-1} Q(z)^n. \tag{27.446}$$

Cela est une approximation de l'unité par la proposition 27.113. Les φ_k sont des polynômes trigonométriques parce que les Q^n le sont et que que $\int_{S^1} Q^n$ est seulement un nombre.

Le lemme 27.112 nous dit alors que pour tout k, la fonction

$$\varphi_k * f \tag{27.447}$$

est un polynôme trigonométrique. □

Nous nous permettons de confirmer la remarque 27.105 comme quoi il faut bien tous les e_n avec $n \in \mathbb{Z}$, parce que le polynôme trigonométrique Q est bien construit à partir de $e_1 + e_{-1}$.

55. Par exemple, M est le maximum de $|P|$.

27.9.7 Convolution, bis

LEMooLUBQooWLMFrN

Lemme 27.119.
Nous considérons l'application

$$\begin{aligned}\varphi\colon \mathbb{R} &\to S^1\\ t &\mapsto e^{it}.\end{aligned} \tag{27.448}$$

Soient $t,u\in\mathbb{R}$ tels que $\varphi(t)=\varphi(u)$. Alors pour toutes fonctions pour lesquelles les intégrales convergent,

$$\int_0^{2\pi}(f\circ\varphi)(\theta)(g\circ\varphi)(t-\theta)\frac{d\theta}{2\pi}=\int_0^{2\pi}(f\circ\varphi)(\theta)(g\circ\varphi)(u-\theta)\frac{d\theta}{2\pi}. \tag{27.449}$$

Démonstration. Si $\varphi(u)=\varphi(t)$, alors $u=t+2k\pi$ pour un certain $k\in\mathbb{Z}$. Cette condition implique que $\varphi(t-\theta)=\varphi(u-\theta)$, et donc l'égalité

$$\int_0^{2\pi}(f\circ\varphi)(\theta)(g\circ\varphi)(t-\theta)\frac{d\theta}{2\pi}=\int_0^{2\pi}(f\circ\varphi)(\theta)(g\circ\varphi)(u-\theta)\frac{d\theta}{2\pi}. \tag{27.450}$$

□

Définition 27.120 (Convolution sur S^1)**.**
Le lemme 27.119 permet de définir

$$(f*g)\big(\varphi(t)\big)=\int_0^{2\pi}(f\circ\varphi)(\theta)(g\circ\varphi)(t-\theta)\frac{d\theta}{2\pi} \tag{27.451}$$

pour toutes les paires de fonctions $f,g\in\mathrm{Fun}(S^1,\mathbb{C})$ pour lesquelles l'intégrale converge.

27.10 L'espace de lebesgue $L^2\big([a,b]\big)$

L'espace $L^2\big([a,b]\big)$ est l'espace générique sur lequel nous allons construire les espaces L^2 sur $[-T,T]$ et $[0,2\pi]$. Pour fixer les idées, nous considérons $b>a$.

Si f et g sont dans $L^2\big([a,b]\big)$, il n'est pas possible de définir $f*g$ par la formule intégrale usuelle parce que $f(x_0+t)$ n'existe pas pour tout x_0 et t dans $[a,b]$. Donc soit nous utilisons un truc pas très net comme étendre les fonctions sur $[a,b]$ en fonctions périodiques sur $\mathbb{R}$, soit nous intégrons vraiment seulement sur $[a,b]$.

Nous n'allons suivre aucune de ces deux voies ou plutôt les deux en même temps. Nous allons seulement tout ramener de S^1 que nous venons de travailler.

Proposition-Définition 27.121.
Sur $[a,b]$ nous considérons la mesure de Lebesgue dx usuelle. Si $f,g\in L^2\big([a,b]\big)$, alors

(1) $f\bar{g}\in L^1\big([a,b]\big)$,

(2) La formule

$$\langle f,g\rangle=\int_a^b f(x)\overline{g(x)}dx. \tag{27.452}$$

EQooCRSXooPEopzm

définit un produit hermitien[56]*.*

Démonstration. Pour le premier point, d'abord $\bar{g}\in L^2$, et ensuite l'inégalité de Hölder 27.33(1) dit que $f\bar{g}$ est dans L^1.

Le fait que la formule donne une forme sesquilinéaire découle des propriétés de l'intégrale. Le fait que ce soit hermitien découle du fait que $\overline{\int f}=\int\bar{f}$.

Et enfin,

$$\langle f,f\rangle=\int_a^b|f(x)|^2dx\geqslant 0. \tag{27.453}$$

56. Définition 8.176.

Si il existe une partie de mesure non nulle A sur laquelle $f \neq 0$, alors

$$\int_a^b |f|^2 = \int_A |f|^2 + \int_{[a,b]\setminus A} |f|^2. \tag{27.454}$$

Le premier terme est strictement positif, alors que le second est positif ou nul. Donc le tout est strictement positif. □

27.122.
Il y a (au moins) deux conventions possibles pour le produit scalaire :

$$\langle f, g\rangle = \int_a^b f(x)\bar{g}(x)\,dx \tag{27.455}$$

EQooAJLHooTKraYR

et

$$\langle f, g\rangle = \frac{1}{b-a}\int_a^b f(x)\bar{g}(x)\,dx \tag{27.456}$$

EQooSJJEooQLGzDG

L'argument en faveur de (27.455). Il est plus facile d'être cohérent avec les espaces $L^p(\Omega, \mathcal{A}, \mu)$. En effet, pour de telles espaces, on a vite $\mu(\Omega) = \infty$ et donc du mal à mettre un coefficient $\frac{1}{\mu(\Omega)}$ dans la définition de la norme. Voir la définition 27.9.

L'argument en faveur de (27.456). Le facteur dx a les mêmes unités que $b-a$. En mettant donc le facteur $b-a$, le tout a les unités de fg, comme il se doit pour le produit scalaire.

PROPooLNALooSVNMfe

Proposition 27.123.
Nous considérons

$$\begin{aligned} s\colon [a,b] &\to [0,2\pi] \\ x &\mapsto 2\pi\frac{x-a}{b-a} \end{aligned} \tag{27.457}$$

ainsi que l'application usuelle

$$\begin{aligned} \varphi\colon [0,2\pi[&\to S^1 \\ t &\mapsto e^{it}. \end{aligned} \tag{27.458}$$

L'application

$$\begin{aligned} \phi\colon L^2\big([a,b]\big) &\to L^2(S^1) \\ \phi(f)(z) &= f\big((s^{-1}\circ\varphi^{-1})(z)\big) \end{aligned} \tag{27.459}$$

est une bijection isométrique.

Démonstration. La preuve du fait que ϕ est isométrique suffira pour prouver qu'elle prend bien ses valeurs dans $L^2(S^1)$.

(1) **Isométrique** C'est un calcul :

$$\begin{aligned} \|\phi(f)\|^2 &= \langle\phi(f),\phi(f)\rangle && \text{(27.460a)}\\ &= \int_{S^1} |\phi(f)|^2 && \text{(27.460b)}\\ &= \frac{1}{2\pi}\int_{[0,2\pi[} |\phi(f)\big(\varphi(u)\big)|^2 du && \text{(27.460c)}\\ &= \frac{1}{2\pi}\int_0^{2\pi} |f\big((s^{-1}\circ\varphi^{-1}\circ\varphi)(u)\big)|^2 du && \text{(27.460d)}\\ &= \frac{1}{2\pi}\int_0^{2\pi} |f\big(s^{-1}(u)\big)|^2 du. && \text{(27.460e)} \end{aligned}$$

Il est temps de faire le changement de variables[57] $y = s^{-1}(u)$, c'est-à-dire

$$y = \frac{b-a}{2\pi}u + a. \tag{27.461}$$

57. Nous le faisons de façon un peu informelle ; soyez capable de bien justifier.

En ce qui concerne la différentielle,

$$dy = \frac{b-a}{2\pi}du \tag{27.462}$$

et pour les bornes, si $u = 0$ alors $y = a$ et si $u = 2\pi$, $y = b$. Donc

$$\|\phi(f)\|^2 = \frac{1}{2\pi}\int_a^b |f(y)|^2 \frac{2\pi}{b-a}dy \tag{27.463a}$$

$$= \frac{1}{b-a}\int_a^b |f|^2 \tag{27.463b}$$

$$= \|f\|^2. \tag{27.463c}$$

(2) **Injectif** Soit f telle que $\phi(f) = 0$. Alors pour tout $z \in S^1$ nous avons

$$f\big((s^{-1} \circ \varphi^{-1})(z)\big) = 0. \tag{27.464}$$

Vu que $s^{-1} \circ \varphi^{-1} \colon S^1 \to [a, b[$ est une bijection, pour tout $u \in [a, b[$ nous avons $f(u) = 0$. Donc $f = 0$ dans $L^2\big([a, b]\big)$ parce que du point de vue de L^2, que l'on prenne ou non les bornes, ce n'est pas important.

(3) **Surjectif** Si $g \in L^2(S^1)$, alors en posant

$$f(u) = g\big((\varphi \circ s)(u)\big) \tag{27.465}$$

nous avons $g = \phi(f)$.

□

Définition 27.124.
En ce qui concerne le produit de convolution, si f et g sont des fonctions sur $[a, b]$ nous définissons

$$f * g = \phi^{-1}\big(\phi(f) * \phi(g)\big) \tag{27.466}$$

tant que les formules ont un sens.

Définition 27.125.
*Le **système trigonométrique** sur $[a, b]$ est l'ensemble de fonctions*

$$\begin{aligned} e_k \colon [a, b] &\to \mathbb{C} \\ t &\mapsto \frac{1}{\sqrt{b-a}} e^{2\pi i k t/(b-a)} \end{aligned} \tag{27.467}$$

pour $k \in \mathbb{Z}$.

27.126.
Pour prouver que ce système est une base hilbertienne, il faut prouver que c'est orthonormal et total. Pour prouver que le système est total, il y a (au moins) trois moyens.

(1) Prouver que le système est orthonormal maximal et invoquer la proposition 25.37(1). Cela est fait dans [682].

(2) Prouver que le système trigonométrique sépare les points pour la densité dans les fonctions continues. Ensuite travailler comme dans [683].

(3) Adapter le théorème 27.89 pour prouver directement la densité des polynômes trigonométriques dans $L^2\big([a, b]\big)$.

THOooAVWIooDhnjpN

Théorème 27.127 ([682]).
Le système trigonométrique $\{e_k\}_{k\in\mathbb{Z}}$ est une base hilbertienne de $L^2\big([a, b]\big)$.

Démonstration. En vertu de la proposition 25.37(1), ils nous suffit de prouver que $\{e_k\}_{k\in\mathbb{Z}}$ est une famille orthonormale maximale.

(1) **Orthonormale** Nous calculons le produit :

$$\langle e_k, e_l\rangle = \frac{1}{b-a}\int_a^b e^{2\pi ikt/(b-a)}e^{-2\pi ilt/(b-a)}dt = \frac{1}{b-a}\int_a^b e^{2\pi it(k-l)/(b-a)}dt. \tag{27.468}$$

Justifications.

— Le complexe conjugué de e^{it} est e^{-it} par le corolaire 18.13.

— Les exponentielles sont « fusionnées » avec la proposition 18.9(2).

Si $k = l$ nous avons

$$\langle e_k, e_k\rangle = \frac{1}{b-a}\int_a^b 1\,dt = 1. \tag{27.469}$$

Si $k \neq l$ nous pouvons continuer avec une primitive. Une primitive de e^{at} est $\frac{1}{a}e^{at}$. Dans notre cas, en regroupant toutes les constantes sous le nom C nous avons :

$$\langle e_k, e_l\rangle = C\left[e^{2\pi it(k-l)/(b-a)}\right]_a^b = C\left(e^{2\pi ia(k-l)/(b-a)} - e^{2\pi ib(k-l)/(b-a)}\right). \tag{27.470}$$

Cela vaut zéro. Vous n'y croyez pas ? Faites un effort, relisez le corolaire 18.24, et remarquez que

$$\frac{2\pi a(k-l)}{b-a} - \frac{2\pi b(k-l)}{b-a} = 2\pi(k-l) \in 2\pi\mathbb{Z}. \tag{27.471}$$

(2) **Maximale** Nous devons prouver que pour tout $f \in L^2\big([a,b]\big)$ (autre que $f = 0$), il existe $k \in \mathbb{Z}$ tel que $\langle e_k, f\rangle = 0$. Autrement dit, nous supposons que $\langle e_k, f\rangle = 0$ pour tout k, et nous allons montrer que $f = 0$. Nous allons largement confondre $f \in L^2$ et une fonction f qui représente la classe.
Pour le reste, voir [682].

□

27.11 Sur $[0, 2\pi[$

Le produit de convolution est un peut subtil parce que $f(t-x)$ n'est pas défini à priori pour tout $t, x \in [0, 2\pi[$, vu que f n'est définie que sur $[0, 2\pi[$. Au moins trois solutions s'offrent à nous :

— considérer implicitement la fonction prolongée par périodicité.

— considérer les fonctions sur $\mathbb{R}/2\pi$, et définir un peu toutes les opérations modulo 2π (fastidieux)

— utiliser une bijection ayant les bonnes propriétés avec S^1 sur lequel tout est déjà fait.

Nous sélectionnons la troisième voie. Pour cela nous considérons la fonction (attention, elle n'est pas tout à fait la même que celle plus haut)

$$\begin{aligned}\varphi\colon [0, 2\pi[&\to S^1\\ t &\mapsto e^{it}\end{aligned} \tag{27.472}$$

qui est une bijection par la proposition 18.61(1). Pour le produit de convolution,

$$(f * g)(x) = (f\circ\varphi^{-1}) * (g\circ\varphi^{-1})\big(\varphi(x)\big) \tag{27.473}$$

pour toutes les fonctions $f, g\colon [0, 2\pi[\to \mathbb{C}$ pour lesquelles les intégrales en jeu ont un sens.

27.12 Sur $[-T,T[$

Pour rappel, les éléments de L^2 sont des classes de fonctions à valeurs dans $\mathbb{C}$.

PROPooHNJZooGfRCfU

Proposition 27.128.
Les fonctions

$$\begin{aligned} e_n\colon [-T,T] &\to \mathbb{C} \\ t &\mapsto \frac{1}{\sqrt{2T}} e^{\pi i n t/T}. \end{aligned} \tag{27.474}$$

forment une base hilbertienne[58] *de* $L^2\big([-T,T[\big)$.

Démonstration. C'est un cas particulier du théorème 27.127. □

27.12.1 Le cas dans $[0,2\pi]$

En pratique, nous n'allons pas souvent travailler avec des fonctions sur intervalle symétrique $[-T,T]$, mais le plus souvent nous serons sur $[0,2\pi]$.

Nous notons ici une conséquence du théorème 27.45 dans le cas de l'espace L^2. La proposition suivante est une petite partie du corolaire 25.51, qui sera d'ailleurs démontré de façon indépendante.

Proposition 27.129.
Si nous avons une suite de réels (a_k) *telle que* $\sum_{k=0}^{\infty} |a_k|^2 < \infty$ *alors la suite*

$$f_n(x) = \sum_{k=0}^{n} a_k e^{ikx} \tag{27.475}$$

converge dans $L^2\big(]0,2\pi[\big)$.

Démonstration. Quitte à séparer les parties réelles et imaginaires, nous pouvons faire abstraction du fait que nous parlons d'une série de fonctions à valeurs dans $\mathbb{C}$ au lieu de $\mathbb{R}$.

Un simple calcul est :

$$\|f_n - f_m\|^2 \leqslant \int_0^{2\pi} \sum_{k=n}^{m} |a_k|^2 dx \leqslant 2\pi \sum_{k=n}^{m} |a_k|^2. \tag{27.476}$$

EqHVdJxZT

Par hypothèse le membre de droite est $|s_m - s_n|$ où s_k dénote la suite des sommes partielles de la série des $|a_k|^2$. Cette dernière est de Cauchy (parce que convergente dans $\mathbb{R}$) et donc la limite $n \to \infty$ (en gardant $m > n$) est zéro. Donc la suite des f_n est de Cauchy dans L^2 et donc converge dans L^2. □

27.130.
Adaptons tout cela pour l'espace $L^2\big([0,2\pi]\big)$. Nous posons

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\overline{g(t)} dt \tag{27.477}$$

EQooBFKDooMkCZQt

et

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}. \tag{27.478}$$

EQooKMYOooLZCNap

27.131.
Attention que $e_n(x)$ n'est pas exactement e^{inx} : il y a un coefficient. Lorsque ça a un sens, la théorie de Fourier permet d'écrire

$$f(x) = \sum_{n\in\mathbb{Z}} c_n(f) e^{inx}. \tag{27.479}$$

58. Définition 25.27.

Ici les $c_n(f)$ sont les coefficients de Fourier de f. Ce développement n'est pas le même que

$$f(x) = \sum_{n\in\mathbb{Z}} a_n(f)e_n(x). \tag{27.480}$$

Dans cette dernière égalité, les $a_n(f)$ ne sont pas les coefficients de Fourier, mais $a_n = \langle f, e_n\rangle$. Le lien entre les deux est fondamentalement l'objet du corolaire 28.18.

L'importance du système trigonométrique défini en 27.84 est d'être une base de $L^2\big([0,2\pi]\big)$, comme précisé dans le lemme suivant.

LEMooBJDQooLVPczR

Lemme 27.132.
Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est une base hilbertienne[59] *de $L^2\big([0,2\pi]\big)$.*

Démonstration. Cas particulier du théorème 27.127. □

Note : le théorème 28.8 donne aussi la densité, mais sera démontré plus tard, indépendamment. Voir aussi les thèmes 34 et 46.

Pour un élément donné $f \in L^2\big([0,2\pi]\big)$, nous définissons

$$S_n f = \sum_{k=-n}^{n} \langle f, e_k\rangle e_k \tag{27.481}$$

et nous avons le théorème suivant, qui récompense les efforts consentis à propos de la densité des polynômes trigonométriques dans L^2.

ThoYDKZLyv

Théorème 27.133.
Soit $f \in L^2\big([0,2\pi]\big)$. Nous avons égalité[60]

$$f = \sum_{n\in\mathbb{Z}} c_n(f)e_n \tag{27.482}$$

EqXMMRpSN

dans L^2.

Nous avons aussi la convergence

$$S_n f \xrightarrow{L^2} f. \tag{27.483}$$

EqRBWKsYP

Démonstration. Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est total pour l'espace de Hilbert $L^2\big([0,2\pi]\big)$ (sans périodicité particulière). Donc le point (1) du théorème 25.46 nous donne l'égalité demandée.

La convergence (27.483) est une reformulation de l'égalité (27.482). □

27.134.
Obtenir la convergence L^2 ne demande pas d'hypothèses de périodicité : la convergence (27.483) est automatique du fait que le système trigonométrique soit total. Ce n'est cependant pas plus qu'une convergence L^2 et elle ne demande pas $f(0) = f(2\pi)$, même si pour chacun des e_k nous avons $e_k(0) = e_k(2\pi)$.

Si $f(2\pi) \neq f(0)$, alors il existe tout de même une suite (f_n) convergente vers f au sens L^2 telle que $f_n(0) = f_n(2\pi)$. Cela ne contredit en rien le fait que $e_k(0) = e_k(2\pi)$ parce que dans L^2, la valeur d'un point seul n'a pas d'importance.

Si nous voulons une vraie convergence ponctuelle ou uniforme $(S_nf)(x) \to f(x)$, alors il faut ajouter des hypothèses sur la continuité de f, sa périodicité ou le comportement des coefficients c_n. Voir aussi le thème 73.

EXooQDWUooLtuIOm

Exemple 27.135.
Si $f \in L^2\big([0,2\pi]\big)$ est (la classe de) une fonction à valeurs réelles, alors on peut la développer avec nettement moins de termes. D'abord nous savons que $e_{-n} = \overline{e_n}$, et donc

$$\langle f, e_n\rangle = \overline{\langle f, e_{-n}\rangle}, \tag{27.484}$$

59. Définition 25.27.
60. Notons que la somme sur $\mathbb{Z}$ dans (27.482) est commutative ; il n'est donc pas besoin d'être plus précis.

ce qui donne

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n = \sum_{n>0} \langle f, e_n\rangle e_n + \sum_{n<0} \overline{\langle f, e_n\rangle e_n} + \langle f, e_0\rangle e_0 = \sum_{n\in\mathbb{N}} \Re\big(\langle f, e_n\rangle e_n\big). \tag{27.485}$$

Notez que f étant supposée réelle et e_0 étant la fonction constante (réelle) $1/\sqrt{2\pi}$, le terme $n = 0$ est bien réel.

Or

$$\Re\big(\langle f, e_n\rangle e_n\big) = \frac{1}{(2\pi)^{3/2}} \cos(nx) \int_0^{2\pi} f(t)\cos(nt)dt - \frac{1}{(2\pi)^{3/2}} \sin(nx) \int_0^{2\pi} f(t)\sin(nt)dt. \tag{27.486}$$

EQooMWJNooSjPCpR

Considérons la fonction impaire $\tilde{f} \in L^2\big([-2\pi, 2\pi]\big)$ créée à partir de f. Elle se développe de même et nous avons la même formule (27.486) à part quelques coefficients et le fait que les intégrales sont entre -2π et 2π. Vu que $\tilde{f}$ est impaire, l'intégrale avec $\cos(nt)$ s'annule et

$$\tilde{f}(x) = \sum_{n\in\mathbb{N}} c_n \sin(nx) \tag{27.487}$$

pour certains coefficients réels c_n. Cette égalité est à considérer dans L^2, c'est-à-dire presque partout et en particulier presque partout sur $[0, 2\pi]$.

Donc les fonctions réelles sur $[0, 2\pi]$ peuvent être écrites sous la forme d'une série de seulement des sinus.

Note : en choisissant $\tilde{f}$ paire, nous aurions eu une série de cosinus. △

27.13 Des inégalités

Avant d'entrer dans le vif du sujet, nous nous fendons d'une petite étude de fonction. Soit

$$\begin{aligned} \phi\colon [0,1] &\to \mathbb{R} \\ x &\mapsto \frac{(1+x)^r}{1+x^r}. \end{aligned} \tag{27.488}$$

Un peu de calcul montre que

$$\frac{\phi'(x)}{\phi(x)} = \frac{r(1-x^{r-1})}{(1+x^r)(1+x)}. \tag{27.489}$$

27.13.1 c-r inégalité

LEMooFKKEooDTypUd

Lemme 27.136.
Soient $a, b > 0$ et $r \geqslant 1$. Nous avons les inégalités

$$a^r + b^r \leqslant (a+b)^r \leqslant 2^{r-1}(a^r + b^r). \tag{27.490}$$

Démonstration. Notons d'abord que si $r = 1$, nous avons juste de égalités ; pas de problèmes. Nous supposons donc que $r > 1$. Pour la première inégalité, nous posons $f(x) = a^r + x^r$ et $g(x) = (a+x)^r$. Nous avons $f(0) = g(0) = a^r$, et, en utilisant la fonction f_α définite par $f_\alpha(x) = x^\alpha$, nous avons

$$f'(x) = rx^{r-1} = rf_{r-1}(x) \tag{27.491a}$$
$$g'(x) = r(a+x)^{r-1} = rf_{r-1}(a+x). \tag{27.491b}$$

Vu que $r > 1$, la fonction f_{r-1} est strictement croissante sur les positifs par la proposition 12.403. Donc pour $x \geqslant 0$ nous avons

$$f'(x) = rf_{r-1}(x) \leqslant rf_{r-1}(a+x) = g'(x). \tag{27.492}$$

Nous avons donc $f(b) \leqslant g(b)$ comme souhaité.

Nous passons à la seconde inégalité. Le lemme 17.87 nous dit que la fonction $f : x \mapsto x^r$ est convexe. Donc

$$f\left(\frac{a}{2}+\frac{b}{2}\right) \leqslant \frac{1}{2}f(a)+\frac{1}{2}f(b). \tag{27.493}$$

De là nous déduisons

$$\frac{(a+b)^r}{2^r} \leqslant \frac{1}{2}(a^r+b^r), \tag{27.494}$$

c'est-à-dire la seconde inégalité. □

Comme semble indiquer le code suivant, la seconde inégalité du lemme 27.136 semble vraie pour tous les complexes.

```python
import random


def f1(z1: complex, z2: complex, r: float) -> float:
    return abs((z1+z2)/2)**r


def f2(z1: complex, z2: complex, r: float) -> float:
    return (abs(z1)**r + abs(z2)**r)/2


def test(z1: complex, z2: complex, r: float) -> float:
    """Should be positive."""
    return f2(z1, z2, r) - f1(z1, z2, r)


for _ in range(1, 10):
    r = random.uniform(1, 1.3)
    x1 = random.uniform(-10, 10)
    y1 = random.uniform(-10, 10)
    x2 = random.uniform(-10, 10)
    y2 = random.uniform(-10, 10)
    z1 = complex(x1, y1)
    z2 = complex(x2, y2)
    result = test(z1, z2, r)
    print("-----")
    print(f"z1={z1}")
    print(f"z2={z2}")
    print(f"r={r}")
    print("---> ", result)
    if result < 0:
        print("!!!!!!!!!!!!!!!!!!!!!!!!!!!!")
```

tex/frido/codeSnip_5.py

PROPooPMIGooRAjROX

Proposition 27.137 ([684]).
Soient $\alpha, \beta \in \mathbb{C}$ et $r \geqslant 1$. Nous avons

$$\left|\frac{\alpha+\beta}{2}\right|^r \leqslant \frac{|\alpha|^r+|\beta|^r}{2}. \tag{27.495}$$

Démonstration. Nous prouvons en montant progressivement en généralité.

(1) $\alpha = 1$ **et** $\beta = t \geqslant 1$ SPITEMooUDCCooSxNxhv Nous commençons avec $\alpha = 1$ et en supposant que β est un réel plus grand ou égal à 1. Nous devons prouver que

$$\left(\frac{1+t}{2}\right)^r \leqslant \frac{1+t^r}{2}. \tag{27.496}$$

Nous posons

$$\begin{aligned} f\colon [1,\infty[&\to \mathbb{R} \\ t &\mapsto \frac{1+t^r}{2} - \left(\frac{1+t}{2}\right)^r, \end{aligned} \tag{27.497}$$

et nous prouvons que $f(t) \geqslant 0$ pour tout $t \geqslant 1$. Pour cela nous calculons la dérivée :

$$f'(t) = \frac{r}{2}\left(t^{r-1} - \left(\frac{1+t}{2}\right)^{r-1}\right) \geqslant 0. \tag{27.498}$$

Nous avons utilisé le fait que $x \mapsto x^{r-1}$ est croissante et que $t \geqslant 1$. Vu que $f(1) = 1 - 1 = 0$, la proposition 12.186(2) donne

$$f(t) \geqslant f(1) = 0 \tag{27.499}$$

pour tout $t \geqslant 1$

(2) $\alpha = z \in \mathbb{C}$ **avec** $|z| \geqslant 1$ **et** $\beta = 1$ SPITEMooLSGLooXiItES
Nous avons

$$\left|\frac{z+1}{2}\right|^r \leqslant \left(\frac{|z|+1}{2}\right)^r \qquad \text{pcq.} |z+1| \leqslant |z|+1 \tag{27.500a}$$

$$\leqslant \frac{1+|z|^2}{2} \qquad \text{par point (1).} \tag{27.500b}$$

(3) **Cas général**
Notons que si $\beta = 0$ c'est facile. Nous supposons que $\beta \neq 0$ et que $|\alpha| \geqslant |\beta|$. En posant $z = \alpha/\beta$ nous avons $|z| \geqslant 1$ et nous pouvons calculer un peu :

$$\left|\frac{\alpha+\beta}{2}\right|^r = \left|\frac{\beta\frac{\alpha+\beta}{\beta}}{2}\right| \tag{27.501a}$$

$$= |\beta|^r \left|\frac{z+1}{2}\right| \tag{27.501b}$$

$$\leqslant |\beta|^r \frac{1+|z|^r}{2} \qquad \text{par le point (2)} \tag{27.501c}$$

$$= \frac{|\beta|^r + |\alpha|^r}{2}. \tag{27.501d}$$

□

PROPooHEWFooJpfzEL

Proposition 27.138 ([672]).
Si $a, b \geqslant 0$ et si $0 < r \leqslant 1$, alors

$$(a+b)^r \leqslant a^r + b^r. \tag{27.502}$$

Démonstration. Nous avons un calcul :

$$(a+b)^r - a^r = \int_a^{a+b} rt^{r-1}dt \qquad \text{tho. 14.287} \tag{27.503a}$$

$$= \int_0^b r(s+a)^{r-1}ds \qquad \text{ch.var. } s = t - a \tag{27.503b}$$

$$\leqslant \int_0^b rs^{r-1}ds \qquad \text{cf. justif.} \tag{27.503c}$$

SUBEQooUCTQooRULeqL

$$= b^r. \tag{27.503d}$$

Pour (27.503c), nous savons que $r < 1$ (le cas $r = 0$ est facile), de telle sorte que $r - 1 < 0$. La fonction $s \mapsto (s+a)^{r-1}$ est donc décroissante. □

PROPooTIJYooJGxphQ

Proposition 27.139 (c-r inégalité[672]).
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Soient $f, g \in L^r(\Omega, \mathcal{A}, \mu)$ des fonctions à valeurs réelles[61]*. En posant $c_r = 1$ pour $0 < r \leqslant 1$ et $c_r = 2^{r-1}$ si $r \geqslant 1$, nous avons*

(1) Pour tout $\omega \in \Omega$,

$$\big|f(\omega) + g(\omega)\big|^r \leqslant c_r |f(\omega)|^r + c_r |g(\omega)|^r. \tag{27.504}$$

EQooWIQQooLWPioQ

(2) Nous avons

$$\int_\Omega |f + g|^r d\mu \leqslant c_r \int_\Omega |f|^r d\mu + c_r \int_\Omega |g|^r d\mu. \tag{27.505}$$

EQooAJTKooGNxHot

Démonstration. L'inégalité (27.504) provient de propositions 27.138 et 27.137 (suivant la valeur de r). L'inégalité (27.505) demande maintenant d'intégrer sur Ω. □

27.13.2 Inégalité de Hanner

Nous allons démontrer les inégalités de Hanner dans le théorème 27.144. Vu que ce sera un peu longuet, nous faisons un lemme.

LEMooDHRCooQiSpyC

Lemme 27.140.
Soient $z_1, z_2 \in \mathbb{C}$. Nous avons

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geqslant \big(|z_1| + |z_2|\big)^p + \big||z_1| - |z_2|\big|^p. \tag{27.506}$$

EQooMUXVooSpGSyG

Démonstration. Soient $z_1, z_2 \in \mathbb{C}$. Nous posons

$$d = |z_1 + z_2|^p + |z_1 - z_2|^p. \tag{27.507}$$

EQooJKYZooEzbETG

Pour $|z_1|$ et $|z_2|$ fixés, nous nous demandons quel est le minimum possible de d.

Si $|z_1| = 0$, alors le minimum est $2|z_2|^p$ et si $|z_2| = 0$ alors il est $2|z_1|^p$. Pour les autres cas, nous posons $|z_1| = a > 0$ ainsi que $b \in \mathbb{R}$ et $\theta \in \mathbb{R}$ tels que[62]

$$z_2 = z_1 a^{-1} b e^{i\theta}. \tag{27.508}$$

Nous avons déjà que $z_1 + z_2 = z_1(1 + a^{-1} b e^{i\theta})$ et donc

$$|z_1 + z_2| = a|1 + a^{-1} b e^{i\theta}| = |a + b e^{i\theta}| \tag{27.509}$$

parce que $a > 0$. De plus,

$$|a + b e^{i\theta}|^2 = (a + b e^{i\theta})(a + b e^{-i\theta}) = a^2 + b^2 + 2ab\cos(\theta) \tag{27.510}$$

parce que $e^{i\theta} + e^{-i\theta} = \cos(\theta)$. Nous posons

$$d(\theta) = |a + b e^{i\theta}|^p + |a - b e^{i\theta}|^p. \tag{27.511}$$

En développant,

$$d(\theta) = \big(a^2 + b^2 + 2ab\cos(\theta)\big)^{p/2} + \big(a^2 + b^2 - 2ab\cos(\theta)\big)^{p/2}. \tag{27.512}$$

Trouvons le minimum de cette fonction de θ. D'abord sa dérivée :

$$\begin{aligned} d'(\theta) &= pab\sin(\theta)\big[\big(a^2 + b^2 - 2ab\cos(\theta)\big)^{p/2-1} - \big(a^2 + b^2 + 2ab\cos(\theta)\big)^{p/2-1}\big] && \text{(27.513a)} \\ &= pab\sin(\theta) s(\theta). && \text{(27.513b)} \end{aligned}$$

61. Par défaut, les éléments de L^p sont à valeurs complexe. Et oui, je sais que ce sont des classes de fonctions, et donc elles peuvent être à valeurs complexes sur une partie de mesure nulle.

62. Proposition 18.62

Nous avons $s(\theta)=0$ pour $\theta=\pi/2$ et $\theta=3\pi/2$. Il faut surtout remarquer que $1<p<2$, ce qui donne $\frac{p}{2}-1<0$. La fonction $x\mapsto x^{p/2-1}$ est donc décroissante. Cela pour dire que

$$s(0)=\big(|a-b|^2\big)^{p/2-1}-\big(|a+b|^2\big)^{p/2-1}>0. \tag{27.514}$$

De la même façon, $s(\pi)=-s(0)<0$. Cela permet d'écrire un petit tableau de signe de d', et de conclure que $d(\theta)$ a un minimum en 0 et en π. Calcul fait, nous avons

$$d(0)=d(\pi)=|a+b|^p+|a-b|^p. \tag{27.515}$$

En reliant à (27.507) nous avons l'inégalité

$$|z_1+z_2|^p+|z_1-z_2|^p\geqslant(a+b)^p-|a-b|^p. \tag{27.516}$$ EQooVHQOooJcbsCR

Nous rappelons que $a=|z_1|$ et que $z_2=z_1a^{-1}be^{i\theta}$. Notons au passage que $|z_2|=b$, donc que ce que nous dit l'équation (27.516) est que

$$|z_1+z_2|^p+|z_1-z_2|^p\geqslant\big(|z_1|+|z_2|\big)^p+\big||z_1|-|z_2|\big|^p. \tag{27.517}$$

□

Encore dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

LEMooTCNEooADpNai

Lemme 27.141 ([1, 685])**.**
La fonction

$$\begin{aligned}\eta\colon\ ]0,\infty[&\to\mathbb{R}\\ a&\mapsto(a^{1/p}+1)^p+|a^{1/p}-1|^p\end{aligned} \tag{27.518}$$

est strictement convexe.

Démonstration. La fonction η est une fonction de classe C^∞ sur $]0,\infty[\backslash\{1\}$. Quelle est sa régularité en $a=1$? À cause de la valeur absolue, il n'est pas clair qu'elle y soit dérivable. En tout cas, la fonction $x\mapsto|x-1|$ n'est pas dérivable en $x=1$, mais peut-être que les exposants aident à lisser. Nous y reviendrons.

Afin de suivre les calculs nous introduisons quelques fonctions :

$$so(x)=1+x^{1/p} \tag{27.519a}$$
$$di(x)=1-x^{1/p} \tag{27.519b}$$
$$dj(x)=x^{1/p}-1 \tag{27.519c}$$

Pour les dérivées, nous avons

$$so'(x)=\frac{1}{p}x^{1/p-1} \tag{27.520a}$$
$$di'(x)=-so'(x) \tag{27.520b}$$
$$dj'(x)=so'(x). \tag{27.520c}$$

Nous divisons les cas selon $a<1$ ou $a>1$.

(1) **Pour $a<1$** Nous avons

$$\eta(a)=so(a)^p+di(a)^p, \tag{27.521}$$

et la première dérivée donne :

$$\eta'(a)=p\,so'(a)\big(so(a)^{p-1}-di(a)^{p-1}\big). \tag{27.522}$$ EQooCLXZooXCJQwd

Pour la seconde dérivée nous trouvons d'abord

$$\begin{aligned}\eta''(a)=&\left(\frac{1-p}{p}\right)a^{\frac{1}{p}-2}\big(so(a)^{p-1}-di(a)^{p-1}\big)\\ &+\frac{p-1}{p}a^{\frac{2}{p}-2}\big(so(a)^{p-2}+di(a)^{p-2}\big).\end{aligned} \tag{27.523}$$

À partir de là, le truc est de substituer les expressions suivantes :

$$so(a)^{p-1} = so(a)^{p-2}so(a) = so(a)^{p-2} + so(a)^{p-2}a^{1/p} \tag{27.524a}$$
$$di(a)^{p-1} = di(a)^{p-2} - a^{1/p}di(a)^{p-2}. \tag{27.524b}$$

Plein de trucs se simplifient et nous obtenons

$$\eta''(a) = \frac{p-1}{p}a^{\frac{1}{p}-2}\big(di(a)^{p-1} - so(a)^{p-2}\big). \tag{27.525}$$

(2) **Pour $a > 1$** Les calculs sont essentiellement les mêmes, en partant de

$$\eta(a) = so(a)^p + dj(a)^p. \tag{27.526}$$

Les résultats sont :

$$\eta'(a) = p\,so'(a)\big(so(a)^{p-1} + dj(a)^{p-1}\big), \tag{27.527}$$

EQooAJLHooGWiPlz

et

$$\eta''(a) = \frac{p-1}{p}a^{\frac{1}{p}-2}\big(dj(a)^{p-2} - so(a)^{p-2}\big). \tag{27.528}$$

(3) **Résumé pour $a \neq 1$**
Au final, nous avons pour tout $a \neq 1$:

$$\eta''(a) = \frac{p-1}{p}a^{\frac{1}{p}-2}\big(|1 - a^{1/p}|^{p-2} - (1 + a^{1/p})^{p-2}\big). \tag{27.529}$$

Ce qu'il se passe en $a = 1$ est encore une question ouverte que nous traitons maintenant.

(4) **Pour $a = 1$**
Les limites des expressions (27.522) et (27.527) en $a = 1$ sont vite calculées et c'est 2^{p-1} dans les deux cas. Donc la dérivée admet une prolongation continue en $a = 1$. Nous allons prouver que la fonction η est en réalité dérivable en $a = 1$ et que la dérivée vaut 2^{p-1}.
Nous nous concentrons sur la partie difficile donnée par $f(x) = |x^{1/p} - 1|^p$. Elle est donnée par

$$f(x) = \begin{cases} di(x)^p & \text{si } x < 1 \\ dj(x)^p & \text{si } x > 1 \\ 0 & \text{si } x = 1. \end{cases} \tag{27.530}$$

Si $f'(1)$ existe, alors elle est égale à la limite

$$f'(1) = \lim_{\epsilon\to 0}\frac{f(1) - f(1-\epsilon)}{\epsilon}. \tag{27.531}$$

Les deux limites à calculer sont :

$$\lim_{\epsilon\to 0^+}\frac{\big((1+\epsilon)^{1/p} - 1\big)^p}{\epsilon} \tag{27.532}$$

et

$$\lim_{\epsilon\to 0^-}\frac{\big(1 - (1+\epsilon)^{1/p}\big)^p}{\epsilon}. \tag{27.533}$$

La première se traite par la règle de l'Hospital [63], et le résultat est zéro. Pour la seconde, il faut juste transformer

$$\lim_{\epsilon\to 0^+}\frac{\big((1+\epsilon)^{1/p} - 1\big)^p}{\epsilon} = \lim_{h\to 0^+}\frac{\big(1 - (1-h)^{1/p}\big)^p}{-h}, \tag{27.534}$$

qui se traite également par la règle de l'Hospital. Le résultat est également zéro.
Donc η est dérivable en $a = 1$ et la dérivée vaut $\eta'(1) = 2^{p-1}$.

63. Proposition 12.194

En récapitulant, nous avons $\eta'' > 0$ sur $]0,\infty[\backslash\{1\}$, donc η' est croissante sur cette partie (proposition 12.185). Vu que η' est continue sur $]0,\infty[$, elle est même croissante (strictement) sur tout $]0,\infty[$.

La proposition 17.83 conclut que η est strictement convexe sur $]0,\infty[$. □

Toujours dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

Lemme 27.142 ([685]).
Soit $1 < p < 2$. La fonction

$$\begin{aligned}\xi\colon \mathbb{R}^+ \times \mathbb{R}^+ &\to \mathbb{R}\\ (a,b) &\mapsto \big(a^{1/p} + b^{1/p}\big)^p + |a^{1/p} - b^{1/p}|^p\end{aligned} \tag{27.535}$$

est convexe.

Pour rappel, les conventions de données en 1.428 donnent $\mathbb{R}^+ = [0,\infty[$.

Démonstration. La fonction ξ vérifie facilement les conditions suivante :

— $\xi(a,b) = \xi(b,a)$,

— $\xi(0,0) = 0$,

— $\xi(ta,tb) = t\xi(a,b)$ pour tout $t \geqslant 0$.

Nous posons

$$\begin{aligned}\eta\colon \mathbb{R}^+ &\to \mathbb{R}\\ a &\mapsto \xi(a,1).\end{aligned} \tag{27.536}$$

Le lemme 27.141 dit que η est strictement convexe, et le lemme 17.93 conclut que ξ est convexe. □

LEMooWIPYooMZqjbn

Lemme 27.143 ([686]).
Soit $1 < p < 2$. Nous considérons les fonctions

$$\begin{aligned}\alpha\colon [0,1] &\to \mathbb{R}\\ x &\mapsto (1+x)^{p-1} + (1-x)^{p-1}\end{aligned} \tag{27.537}$$

et

$$\begin{aligned}\beta\colon [0,1] &\to \mathbb{R}\\ x &\mapsto x^{1-p}\big((1+x)^{p-1} - (1-x)^{p-1}\big).\end{aligned} \tag{27.538}$$

Soient $A, B \in \mathbb{R}$. Nous avons

$$\alpha(x)|A|^p + \beta(x)|B|^p \leqslant |A+B|^p + |A-B|^p. \tag{27.539}$$

Démonstration. Plusieurs étapes.

(1) $\beta(x) \leqslant \alpha(x)$ Nous avons $\alpha(1) = \beta(1) = 2^{p-1}$. Pour les autres valeurs de x, nous allons raisonner avec la dérivée. La valeur de $\alpha'(x)$ est facile à calculer

$$\alpha'(x) = (p-1)(x+1)^{p-2} - (p-1)(1-x)^{p-2}. \tag{27.540}$$

Pour $\beta'(x)$ c'est un peu plus lourd. En substituant $(1+x)^{p-1} = (1+x)^{p-2}(1+x)$ et $(1-x)^{p-1} = (1-x)^{p-2}(1-x)$ nous pouvons regrouper les termes en $(1+x)^{p-2}$ et $(1-x)^{p-2}$. Après un peu de travail,

$$\beta'(x) = \frac{p-1}{x^p}\big((1-x)^{p-2} - (1+x)^{p-2}\big). \tag{27.541}$$

Cela nous permet de calculer $\alpha' - \beta'$:

$$\alpha'(x) - \beta'(x) = (p-1)\big(1 + \frac{1}{x^p}\big)\big((1+x)^{p-2} - (1-x)^{p-2}\big). \tag{27.542}$$

Vu que $1<p<2$, le nombre $p-2$ est strictement négatif ; afin de travailler avec des exposants positifs, nous écrivons

$$\alpha'(x)-\beta'(x)=\underbrace{(p-1)}_{>0}\underbrace{\Big(1+\frac{1}{x^p}\Big)}_{>0}\underbrace{\left(\frac{1}{(1+x)^{2-p}}-\frac{1}{(1-x)^{2-p}}\right)}_{<0}. \tag{27.543}$$

Nous avons $\alpha'(x)-\beta'(x)<0$ pour tout $x\in]0,1]$. Du fait qu'en plus nous ayons $\alpha(1)=\beta(1)$, nous déduisons que $\alpha(x)\geqslant\beta(x)$.

(2) **Une petite étude de fonction** Soit $R\in[0,1]$. Nous considérons la fonction

$$\begin{aligned}F\colon[0,1]&\to\mathbb{R}\\ x&\mapsto\alpha(x)+R^p\beta(x).\end{aligned} \tag{27.544}$$

Nous montrons maintenant que cette fonction a un maximum global pour $x=R$. D'abord sa dérivée :

$$F'(x)=\underbrace{(p-1)}_{>0}\underbrace{\Big((1-x)^{p-1}-(1+x)^{p-2}\Big)}_{<0}\left(1-\left(\frac{R}{x}\right)^p\right) \tag{27.545}$$

Nous avons

- $F'(x)=0$ pour $x=R$,
- $F'(x)<0$ pour $x>R$,
- $F'(x)>0$ pour $x<R$.

Donc $x=R$ est bien un maximum global.

(3) **Pause** Nous avons les petits résultats utiles pour commencer à prouver. Petite pause avant de commencer ; pas de panique, ça ne va pas être trop violent.

(4) **Pour $0<B<A$** Nous devons prouver que

$$\alpha(x)A^p+\beta(x)B^p\leqslant(A+B)^p+(A-B)^p. \tag{27.546}$$

EQooEPKRooBYJDSF

En divisant par A^p et en posant $R=B/A$, l'inéquation (27.546) est équivalente à

$$\alpha(x)+\beta(x)R^p\leqslant(1+R)^p+(1-R)^p \tag{27.547}$$

où $R\in]0,1[$ parce que nous avons supposé $0<B<A$. Nous avons (il y a un petit calcul pour $F(R)$)

$$(1+R)^p+(1-R)^p=F(R)\geqslant F(x)=\alpha(x)+\beta(x)R^p. \tag{27.548}$$

ok.

(5) **Pour $0<A<B$** Lorsque $0<A<B$ nous avons

$$\begin{aligned}\alpha(x)|A|^p+\beta(x)|B|^p&=\alpha(x)A^p+\beta(x)B^p &(27.549a)\\ &\leqslant\alpha(x)B^p+\beta(x)A^p &(27.549b)\\ &\leqslant(B+A)^p+(B-A)^p &(27.549c)\\ &=|A+B|^p+|A-B|^p. &(27.549d)\end{aligned}$$

SUBEQooSHNUooCoWMFR
SUBEQooBPYVooPsAjbq

Justification :

- Pour (27.549b), c'est parce que $\alpha(x)>\beta(x)$; alors en mettant le plus grand de A et B devant le α au lieu du β, nous majorons.
- Pour (27.549c), c'est l'inégalité dans le cas $0<B<A$, mais en inversant les noms de A et B.

(6) **Pour $0<A=B$** Toutes les expressions sont continues par rapport à B (fixons x et A). Nous avons prouvé pour $B<A$ et pour $B>A$. Par continuité, l'inégalité est encore valide pour $A=B$.

(7) **Pour $A<0$, $B>0$** En posant $A'=-A$ nous avons $A'>0$ et nous pouvons écrire

$$|A+B|^p+|A-B|^p=|-A'+B|^p+|-A'-B|^p=|B-A'|^p+|B+A'|^p\geqslant\alpha(x)|A'|^p+\beta(x)|B|^p. \tag{27.550}$$

Nous avons utilisé, avec A' et B le cas déjà prouvé $A',B>0$.

(8) **Pour $A>0$, $B<0$** Celui-là, je vous le laisse.

(9) **Pour $A<0$, $B<0$** Posez $A'=-A$ et $B'=-B$ et hop.

□

THOooZRRYooBTBQKW

Théorème 27.144 (Inégalités de Hanner[685, 686]).
Soit un espace mesuré $(\Omega,\mathcal{A},\mu)$. Soit $1<p<2$ et $f,g\in L^p(\Omega,\mathcal{A},\mu)$; nous avons

$$\big(\|f\|_p+\|g\|_p\big)^p+\Big|\|f\|_p-\|g\|_p\Big|^p\leqslant\|f+g\|_p^p+\|f-g\|_p^p \tag{27.551}$$

Il y a égalité si et seulement si $f(t)$ et $g(t)$ sont colinéaires pour presque tout t.

Démonstration. Nous supposons que $\|f\|_p\geqslant\|g\|_p$ pour fixer les idées. De toutes façons, la symétrie des formules nous fait passer de ce cas à l'autre sans difficulté.

Soit $x\in[0,1]$. Nous écrivons l'inégalité du lemme 27.143 pour $A=|f(\omega)|$ et $B=|g(\omega)|$:

$$\alpha(x)|f(\omega)|^p+\beta(x)|g(\omega)|^p\leqslant\big|f(\omega)+g(\omega)\big|^p+\big|f(\omega)-g(\omega)\big|^p. \tag{27.552}$$

Nous intégrons cela par rapport à ω sur Ω :

$$\alpha(x)\|f\|_p^p+\beta(x)\|g\|_p^p\leqslant\|f+g\|_p^p+\|f-g\|_p^p. \tag{27.553}$$

Et là vient l'idée qu'on se demande ce qui est passé par l'esprit du mec qui a tout combiné : nous évaluons cela pour $x=\frac{\|g\|_p}{\|f\|_p}$, ce qui est permis parce que nous avons supposé $\|f\|_p\geqslant\|g\|_p$. Faites le calcul, collectez les termes identiques, vous obtiendrez

$$\big(\|f\|_p+\|g\|_p\big)^p+\big(\|f\|_p-\|g\|_p\big)^p\leqslant\|f+g\|_p^p+\|f-g\|_p^p. \tag{27.554}$$

Et vu que $\|f\|_p\geqslant\|g\|_p$, nous pouvons gratuitement faire

$$\|f\|_p-\|g\|_p=\big|\|f\|_p-\|g\|_p\big|. \tag{27.555}$$

Fini pour Hanner. □

27.13.3 Inégalités de Clarkson

LEMooWEODooLHeVrP

Lemme 27.145 ([687]).
Si $p\geqslant2$ et si $a,b\in\mathbb{C}$, alors

$$\left|\frac{a+b}{2}\right|^p+\left|\frac{a-b}{2}\right|^p\leqslant\frac{1}{2}\big(|a|^p+|b|^p\big). \tag{27.556}$$

Démonstration. Nous prouvons l'inégalité en montant petit à petit en généralité.

(1) **Avec $x>0$** Soit $x\geqslant0$. Nous montrons dans cette partie l'inégalité

$$x^p+1\leqslant(x+1)^{p/2}. \tag{27.557}$$

EQooDJBNooEyfNtq

Pour cela nous considérons la fonction

$$\begin{aligned}f\colon[0,\infty[&\to\mathbb{R}\\ t&\mapsto(t^2+1)^{p/2}-t^p-1.\end{aligned} \tag{27.558}$$

Nous avons $f(0)=0$, mais aussi, en utilisant les règle de dérivation[64] nous trouvons vite

$$f'(t)=p(t^2+1)^{p/2-1}t-pt^{p-1}. \tag{27.559}$$

Vu que $(t^2+1)^{p/2-1}\geqslant t^{p-2}$, le signe de $f'(t)$ est toujours strictement positif pour $t>0$. La proposition 12.185 fait que f est strictement croissante et que $f(t)>0$ pour tout $t>0$.

64. Par exemple celle de la proposition 14.294.

(2) **Avec $x, y \geqslant 0$** Soient $x, y \geqslant 0$ dans $\mathbb{R}$. Nous prouvons dans cette partie que

$$(x^2 + y^2)^{p/2} \geqslant x^p + y^p. \tag{27.560}$$

EQooGFGMooSiDfKX

Il s'agit d'appliquer l'inégalité (27.557) à x/y :

$$\left(\left(\frac{x}{y}\right)^2 + 1\right)^{p/2} \geqslant \left(\frac{x}{y}\right)^p + 1. \tag{27.561}$$

En multipliant par y^p et en simplifiant un peu, nous trouvons le résultat (27.560).

(3) **Avec $a, b \in \mathbb{C}$** Nous appliquons l'inégalité (27.560) à $x = |\frac{a+b}{2}|$ et $y = |\frac{a-b}{2}|$. Cela donne :

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leqslant \left(\left|\frac{a+b}{2}\right|^2 + \left|\frac{a-b}{2}\right|^2\right)^{p/2} \tag{27.562a}$$

$$= \left(\frac{2|a|^2 + 2|b|^2}{4}\right)^{p/2} \tag{27.562b}$$

$$\leqslant \frac{1}{2}|a|^p + \frac{1}{2}|b|^p. \tag{27.562c}$$

La dernière ligne est la convexité de la fonction $t \mapsto t^{p/2}$ (lemme 17.87).

□

LEMooFGKXooZCHNln

Lemme 27.146.
Si $1 < p < 2$, alors l'exposant conjugué q vérifie $q > 2$.

Démonstration. Nous considérons q en fonction de p, sur le domaine $1 < p < 2$:

$$q(p) = \frac{p}{p-1}. \tag{27.563}$$

Donc [65] $q(1) = \infty$ et $q(2) = 2$. Nous étudions ensuite la dérivée :

$$q'(p) = -\frac{1}{(p-1)^2} < 0. \tag{27.564}$$

C'est donc une fonction strictement décroissante. Vues les valeurs aux bornes, nous voyons que $q(p) > 2$ sur tout son domaine. □

LEMooMKIXooVOYaxI

Lemme 27.147 ([688])**.**
Soit $1 < p < 2$. Pour $x, y \in \mathbb{R}$ nous avons

$$|x + y|^q + |x - y|^q \leqslant 2\big(|x|^p + |y|^p\big)^{q-1}. \tag{27.565}$$

Démonstration. Nous considérons l'exposant conjugué q et p, c'est-à-dire q tel que $\frac{1}{p} + \frac{1}{q} = 1$. Nous considérons la fonction

$$\begin{aligned} f : \mathbb{R} \times [0,1] &\to \mathbb{R} \\ (\alpha, z) &\mapsto (1 + \alpha^{1-q}z)(1 + \alpha z)^{q-1} + (1 - \alpha^{1-q}z)(1 - \alpha z)^{q-1}. \end{aligned} \tag{27.566}$$

Cette fonction vérifie

$$f(1, z) = (1 + z)^q + (1 - z)^q, \tag{27.567}$$

EQooRFZQooJvdocT

ainsi que

EQooISBRooHMiPRE

$$f(z^{p-1}, z) = \big(1 + z^{(p-1)(1-q)}\big)(1 + z^p)^{q-1} + \big(1 - z^{(p-1)(1-q)}\big)(1 + z^p)^{q-1} \tag{27.568a}$$

$$= 2(1 + z^p)^{q-1}. \tag{27.568b}$$

65. Dire que $q(1) = \infty$ est un abus de notations pour parler de la limite $p \to 1$ avec $p > 1$.

Nous montrons maintenant que $(\partial_\alpha f)(\alpha,z)\leqslant 0$ pour tout $\alpha\in]0,1[$ et pour tout $z\in]0,1[$. C'est du calcul :

$$\frac{\partial f}{\partial\alpha}(\alpha,z)=(1-q)z\big[\alpha^{-q}(1+\alpha z)^{q-1}-(1+\alpha^{1-q}z)(1+\alpha z)^{q-1}-\alpha^{-q}(1-\alpha z)^{q-1}+(1-\alpha^{1-q}z)(1-\alpha z)^{q-2}\big].\tag{27.569}$$

Maintenant nous factorisons $(1+\alpha z)^{q-2}$ grâce à la décomposition $(1+\alpha z)^{q-1}=(1+\alpha z)(1+\alpha z)^{q-2}$. Notez que le lemme 27.146 donne $q>2$, et donc pas de problèmes avec la puissance $q-2$. Nous continuons le calcul

$$\begin{aligned}\frac{\partial f}{\partial\alpha}(\alpha,z)&=(1-q)z(1+\alpha z)^{q-2}\big[\alpha^{-q}(1+\alpha z)-(1+\alpha^{1-q}z)\big]&(27.570a)\\&\quad+(1-q)z(1-\alpha z)^{q-2}\big[-\alpha^{-q}(1-\alpha z)+(1-\alpha^{1-q}z)\big]\\&=(1-q)z(1+\alpha z)^{q-2}\big[\alpha^{-q}+\alpha^{-q+1}z-1-\alpha^{1-q}z\big]&(27.570b)\\&\quad+(1-q)z(1-\alpha z)^{q-2}\big[-\alpha^{-q}+\alpha^{-q+1}+1-\alpha^{1-q}z\big]\\&=(1-q)z(1+\alpha z)^{q-2}[\alpha^{-q}-1]+(1-q)z(1-\alpha z)^{q-2}[1-\alpha^{-q}]&(27.570c)\\&=(1-q)z(\alpha^{-q}-1)\big[(1+\alpha z)^{q-2}-(1-\alpha z)^{q-2}\big].&(27.570d)\end{aligned}$$

Vu que $q>2$, la fonction $x\mapsto x^{q-2}$ est strictement croissante sur les positifs [66]. Et vu que $\alpha z<1$, le crochet est strictement positif. Par ailleurs, $z>0$, $(1-q)<0$ et $(\alpha^{-q}-1)>0$ donc nous avons prouvé que $(\partial_\alpha f)(\alpha,z)\leqslant 0$.

Donc f est décroissante par rapport à α. Vu que $z\in[0,1]$ et que $p>1$, nous avons $z^{p-1}<1$ et donc $f(1,z)\leqslant f(z^{p-1},z)$. Nous y substituons les valeurs calculées en (27.567) et (27.568) :

$$(1+z)^q+(1-z)^q\leqslant 2(1+z^p)^{q-1}.\tag{27.571}$$

EQooFQJAooPCYtMG

Nous pouvons maintenant facilement prouver notre inégalité.

(1) **Pour $0<x<y$** Si $0<x<y$, nous avons $x/y\in[0,1]$ et nous pouvons appliquer l'inégalité (27.571) à $z=x/y$. Nous avons successivement :

$$\begin{aligned}\left(1+\frac{x}{y}\right)^q+\left(1-\frac{x}{y}\right)^q&\leqslant\left(1+\frac{x^p}{y^p}\right)^{q-1}&(27.572a)\\\left(\frac{x+y}{y}\right)^q+\left(\frac{y-x}{y}\right)^q&\leqslant 2\left(\frac{y^p+x^p}{y^p}\right)^{q-1}&(27.572b)\\y^{-q}(x+y)^q+y^{-q}(y-x)^q&\leqslant 2y^{-p(q-1)}(y^p+x^p)&(27.572c)\\(x+y)^q+(y-x)^q&\leqslant 2y^{-p(q-1)+q}(y^p+x^p).&(27.572d)\end{aligned}$$

Nous avons utilisé la proposition 12.408 sur la composition de puissances. Maintenant il suffit de remarquer que $q=p(q-1)$ pour avoir le résultat.

(2) **$x<0$ et $y>0$** En posant $x'=-x$ nous avons

$$\begin{aligned}|x+y|^q+|x-y|^q&=|-x'+y|^q+|-x'-y|^q&(27.573a)\\&=|x'-y|^q+|x'+y|^q&(27.573b)\\&\leqslant 2\big(|x'|^p+|y|^p\big)&(27.573c)\\&\leqslant 2\big(|x|^p+|y|^p\big).&(27.573d)\end{aligned}$$

(3) **Les autres cas** Je vous prie de faire la liste, et d'adapter.

□

Pour d'autres preuves du lemme suivant, voir [689].

66. Proposition 12.403.

LEMooLTROooVusGte

Lemme 27.148 ([690]).
Soient $a, b \in \mathbb{C}$ ainsi que $1 < p < 2$. Nous notons q l'exposant conjugué de p. Nous avons l'inégalité

$$|a+b|^q + |a-b|^q \leqslant 2\big(|a|^p + |b|^p\big)^{q-1}. \tag{27.574}$$

Démonstration. Commençons doucement avec le cas $b = 0$. À gauche nous gardons $2|a|^q$, et pour le membre de droite nous remarquons que

$$q - 1 = \frac{1}{p-1}, \tag{27.575}$$

de telle sorte que

$$(|a|^p)^{q-1} = |a|^{p/(p-1)} = |a|^q. \tag{27.576}$$

Gardez en tête que, par le lemme 27.146, $q > 2$; ce sera utile.

Nous commençons le vrai combat. Vu que $|a-b| = |b-a|$ nous pouvons supposer $|a| \geqslant |b|$ pour fixer les idées. En utilisant le lemme 18.63, il existe $t_0 \in [0, \frac{\pi}{2}]$ tel que

$$\begin{aligned} |a+b|^2 &= |a|^2 + |b|^2 + 2|a||b|\cos(t_0) \\ |a-b|^2 &= |a|^2 + |b|^2 - 2|a||b|\cos(t_0) \end{aligned} \tag{27.577}$$

Nous considérons la fonction

$$\begin{aligned} f\colon [0, 2\pi] &\to \mathbb{R} \\ t &\mapsto \Big(|a|^2 + |b|^2 + 2|a||b|\cos(t)\Big)^{q/2} + \Big(|a|^2 + |b|^2 - 2|a||b|\cos(t)\Big)^{q/2}. \end{aligned} \tag{27.578}$$

Elle vérifie

$$f(t_0) = (|a+b|^2)^{q/2} + (|a-b|^2)^{q/2}. \tag{27.579}$$

Vu que $|a+b|$ et $|a-b|$ sont positifs, nous pouvons « simplifier » le carré et la racine carré, de telle sorte que

$$f(t_0) = |a+b|^q + |a-b|^q. \tag{27.580}$$

Nous cherchons un maximum pour f sur $[0, \frac{\pi}{2}]$. Pour cela, nous prenons d'abord la dérivée :

$$f'(t) = -q|a||b|\sin(t)\Big[\big(|a|^2 + |b|^2 + 2|a||b|\cos(t)\big)^{q/2-1} - \big(|a|^2 + |b|^2 - 2|a||b|\cos(t)\big)^{q/2-1}\Big]. \tag{27.581}$$

Notez que $q > 2$, donc la fonction $x \mapsto x^{q/2-1}$ est croissante.

Pour $t \in [0, \pi/2]$, nous avons $\cos(t) \geqslant -\cos(t)$ ainsi que $\sin(t) \geqslant 0$. Donc $f'(t) \leqslant 0$. De la même manière, nous avons $f'(t) \geqslant 0$ pour $t \in [\pi/2, \pi]$.

Par le lien entre dérivée et croissance (proposition 12.186), nous savons que le maximum de f sur $[0, \pi]$ est atteint en 0 ou en π.

Nous avons, en utilisant la supposition $|a| \geqslant |b|$:

$$f(0) = f(\pi) = \big(|a|^2 + |b|^2 + 2|a||b|\big)^{q/2} + \big(|a|^2 + |b|^2 - 2|a||b|\big)^{q/2} \tag{27.582a}$$

$$= \big(\big||a| + |b|\big|^2\big)^{q/2} + \Big(\big||a| - |b|\big|^2\Big) \tag{27.582b}$$

$$= \big||a| + |b|\big|^q + \big||a| - |b|\big|^q. \tag{27.582c}$$

En particulier $f(t_0) \leqslant f(0)$ et donc

$$|a+b|^q + |a-b|^q \leqslant \big||a| + |b|\big|^q + \big||a| - |b|\big|^q \leqslant 2\big(|a|^p + |b|^p\big)^{q-1}. \tag{27.583}$$

La dernière inégalité est le lemme 27.147 appliqué aux réels $|a|$ et $|b|$. □

PROPooJDOQooWsGlkr

Proposition 27.149 (Inégalité de Clarkson[688]).
Soient $f, g \in L^p(\Omega, \mathcal{A}, \mu)$.

(1) Si $p \geqslant 2$, alors

$$\| \frac{f+g}{2} \|_p^p + \| \frac{f-g}{2} \|_p^p \leqslant \frac{1}{2}\Big(\|f\|_p^p + \|g\|_p^p \Big). \tag{27.584}$$

(2) Si $1 < p < 2$ et si q est l'exposant conjugué de p, alors

$$\|f+g\|_p^q + \|f-g\|_p^q \leqslant 2\Big(\|f\|_p^p + \|g\|_p^p \Big)^{q-1}, \tag{27.585}$$

ou

$$\| \frac{f+g}{2} \|_p^q + \| \frac{f-g}{2} \|_p^q \leqslant 2^{1-q} \big(\|f\|_p^p + \|g\|_p^p \big)^{q-1}. \tag{27.586}$$

Démonstration. En deux parties.

(1) **Pour $p \geqslant 2$** Soient $f, g \in L^p(\Omega, \mathcal{A}, \mu)$; ce sont des fonctions à valeurs dans $\mathbb{C}$. Pour chaque $\omega \in \Omega$ nous considérons les nombres complexes $f(\omega)$ et $g(\omega)$; nous pouvons écrire l'inégalité du lemme 27.145 :

$$\left| \frac{f(\omega)+g(\omega)}{2} \right|^p + \left| \frac{f(\omega)-g(\omega)}{2} \right|^p \leqslant \frac{1}{2}\big(|f(\omega)|^p + |g(\omega)|^p \big). \tag{27.587}$$

Nous avons les substitutions évidentes $f(\omega)+g(\omega) = (f+g)(\omega)$ et $f(\omega)-g(\omega) = (f-g)(\omega)$. En intégrant alors (27.587) sur Ω nous trouvons l'inégalité demandée.

(2) **Pour $1 < p < 2$** Il s'agit de faire la même chose, en utilisant l'inégalité de Clarkson du lemme 27.148.

Pour obtenir (27.586), il s'agit simplement de multiplier et diviser le member de gauche de (27.585) par 2^q.

□

27.14 Théorème de la projection normale

27.14.1 Espace uniformément convexe

DEFooOPQBooBhufew

Définition 27.150 (Espace uniformément convexe[691])**.**
*Un espace de Banach B est **uniformément convexe** si il existe une fonction $\delta :]0, \infty[\to \mathbb{R}^+$ telle que si*

(1) $\|x\| \leqslant \|y\| \leqslant 1$,

(2) $\|x-y\| \geqslant \epsilon$,

alors

$$\| \frac{x+y}{2} \| \leqslant \|y\| - \delta(\epsilon). \tag{27.588}$$

Lemme 27.151 ([1])**.**
Si B est un espace de Banach uniformément convexe, alors pour tout $k > 0$, il existe une fonction $\delta_k :]0, \infty[\to \mathbb{R}^2$ telle que si

(1) $\|x\| \leqslant \|y\| \leqslant k$,

(2) $\|x-y\| \geqslant \epsilon$,

alors

$$\| \frac{x+y}{2} \| \leqslant \|y\| - \delta_k(\epsilon). \tag{27.589}$$

Démonstration. Nous posons $x' = x/k$ et $y' = y/k$. Nous avons alors

$$\|x'-y'\| = \frac{\|x-y\|}{k} > \frac{\epsilon}{k}. \tag{27.590}$$

L'uniforme convexité de B dit alors que

$$\| \frac{x'+y'}{2} \| > \|y'\| - \delta(\epsilon/k). \tag{27.591}$$

En multipliant cette inégalité par k nous trouvons

$$\|\frac{x+y}{2}\| > \|y\| - k\delta(\epsilon/k). \tag{27.592}$$

Donc en posant $\delta_k(\epsilon) = k\delta(\epsilon/k)$, nous avons le résultat escompté. □

DEFooMYYLooJyACPL

Définition 27.152 (Projection normale[691])**.**
Soient un espace de Banach B ainsi que $V \subset B$. Soit $a \in B$. La fonction

$$\begin{aligned} f : V &\to \mathbb{R} \\ x &\mapsto d(x,a) \end{aligned} \tag{27.593}$$

possède un infimum[67] *m. Si $x \in V$ est tel que $d(x,a) = m$, alors x est une **projection normale** de a sur V.*

PROPooDKXVooUoYPgz

Proposition 27.153 ([691])**.**
Soient un espace de Banach B et un sous-espace vectoriel $V \subset B$. Si une projection normale de $a \in B$ sur V existe, alors elle est unique.

Démonstration. Soient deux projections normales b, b' de a sur V.

Si $m = 0$, alors $\|a - b\| = 0$ et $\|a - b'\| = 0$, ce qui donne $a = b$ et $a = b'$. Donc d'accord pour $b = b'$.

Si $m > 0$ alors nous utilisons l'inégalité $\|x + y\| \leqslant \|x\| + \|y\|$ sous la forme

$$\|a - \frac{b+b'}{2}\| = \|\frac{a-b}{2} + \frac{a-b'}{2}\| \leqslant \|\frac{a-b}{2}\| + \|\frac{a-b'}{2}\| = \frac{m}{2} + \frac{m}{2} = m. \tag{27.594}$$

EQooQWJWooVaWMCL

Mais $\frac{b+b'}{2} \in V$, donc

$$\|a - \frac{b+b'}{2}\| \geqslant m. \tag{27.595}$$

Nous en déduisons que dans (27.594), toutes les inégalités sont des égalités et en particulier

$$\|\frac{b+b'}{2} - a\| = m. \tag{27.596}$$

Nous avons donc les deux égalités suivantes :

$$2m = \|a - b\| + \|a - b'\| \tag{27.597}$$

et

$$2m = \|b + b' - 2a\|. \tag{27.598}$$

Cela donne

$$\|a - b\| + \|a - b'\| = \|(a - b) + (a - b')\|. \tag{27.599}$$

Vu que B est strictement convexe, cela n'est possible que si $a - b = a - b'$, ce qui signifie que $b = b'$. □

THOooOOVVooMhzHqd

Théorème 27.154 ([688, 691])**.**
Si B est un espace de Banach uniformément convexe, si $V \subset B$ est un sous-espace vectoriel fermé et si $a \in B$, alors a admet une unique projection normale[68] *sur V.*

Démonstration. En deux parties.

67. Toute fonction à valeurs positives possède un infimum, c'est la proposition 1.461.
68. Définition 27.152.

(1) **Unicité** Soient deux projections normales b et b' de a sur V. Nous avons $\|a-b\| = \|a-b'\| = m$. Si $b \neq b'$, il existe $\epsilon > 0$ tel que

$$\|b - b'\| > \epsilon > 0. \tag{27.600}$$

En posant

$$x = \frac{1}{2}\frac{b-a}{m}, \quad y = \frac{1}{2}\frac{b'-a}{m}, \tag{27.601}$$

nous avons $\|x\| = \|y\| = \frac{1}{2} < 1$. L'uniforme convexité de B donne alors

$$\|\frac{x+y}{2}\| \leqslant \|y\| - \delta(\epsilon). \tag{27.602}$$

Mais

$$x + y = \frac{\frac{1}{2}(b+b') - a}{m} \tag{27.603}$$

et

$$\|x + y\| \leqslant 2\|y\| - 2\delta(\epsilon) = 1 - 2\delta(\epsilon) < 1. \tag{27.604}$$

Nous avons donc prouvé que

$$\|\frac{1}{2}(b+b') - a\| < m, \tag{27.605}$$

ce qui est impossible parce que cela dirait que $\frac{b+b'}{2}$ est une « meilleure » projection normale que b et b'.

(2) **Existence** Soient b_k dans V tels que $\|a - b_k\| \to m$. Nous supposons (quitte à passer à une sous-suite) que

$$\|a - b_{k+1}\| \leqslant \|a - b_k\|. \tag{27.606}$$

(2a) **La suite (b_k) converge**
Nous supposons qu'elle ne converge pas. Elle n'est donc pas de Cauchy parce que B est de Banach [69] et donc complet. Il existe $\epsilon > 0$ tel que pour tout $N \in \mathbb{N}$ il existe $p, q > N$ tels que

$$\|b_p - b_q\| > \epsilon. \tag{27.607}$$

Nous effectuons quelque choix.

(2a.i) nous choisissons $q > p$ de telle sorte que $\|a - b_p\| \leqslant \|a - b_q\|$,

(2a.i) nous choisissons N assez grand pour avoir

$$\|a - b_p\| \leqslant \|a - b_q\| < m. \tag{27.608}$$

Nous avons $\|(a - b_p) - (a - b_q)\| = \|b_q - b_p\| > \epsilon$, ce qui avec l'uniforme convexité donne

$$\frac{\|(a - b_p) + (a - b_q)\|}{2} \leqslant \|a - b_q\| - \delta(\epsilon). \tag{27.609}$$

Donc

$$m \leqslant \|a - \frac{b_p - b_q}{2}\| = \|\frac{(a - b_p) + (a - b_q)}{2}\| \leqslant \|a - b_q\| - \delta(\epsilon) < m - \delta(\epsilon) < m. \tag{27.610}$$

Cela signifie que $m < m$, ce qui est impossible.

(2b) **Conclusion** La suite (b_k) converge dans B. Vu que V est fermé, la limite est dans V. Cette limite, que nous nommons b, vérifie

$$\|a - b\| \leqslant \|a - b_k\| \tag{27.611}$$

pour tout k. Mais comme nous avons $m \leqslant \|a - b_k\| \to m$, nous avons $\|a - b\| = m$, c'est-à-dire que b est une projection normale de a sur V.

□

69. Définition 7.265.

27.14.2 Uniforme convexité des espaces de Lebesgue

PROPooFNLJooDlyIKV

Proposition 27.155 ([692]).
Si $1 < p < \infty$, l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est uniformément convexe[70].

Démonstration. En deux parties.

(1) $\underline{1 < p \leqslant 2}$ Nous montrons que la fonction $\delta(\epsilon) = 2^{-q}\epsilon^q$ fonctionne.
Soient $f, g \in L^p$ telles que $\|f\|_p \leqslant \|g\|_p \leqslant 1$ et $\|f - g\|_p \geqslant \epsilon$. Nous commençons par écrire l'inégalité de Clarkson (27.585) :

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leqslant 2^{1-q}\big(\|f\|_p^p + \|g\|_p^p\big)^{q-1}. \tag{27.612}$$

EQooOWVEooGGfCpy

Par hypothèse, $\|f\|_p$ et $\|g\|_p$ sont plus petites que 1. Vu que $p > 1$, nous avons

$$\|f\|_p^p + \|g\|_p^p \leqslant 1 + 1 = 2. \tag{27.613}$$

En remplaçant dans le membre de droite de (27.612) nous avons

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leqslant 2^{1-q}2^{q-1} = 1, \tag{27.614}$$

et donc

$$\|\frac{f+g}{2}\|_p^q \leqslant 1 - \|\frac{f-g}{2}\|_p^q. \tag{27.615}$$

EQooKARVooDrQuJI

Par ailleurs nous avons supposé $\|f - g\|_p \geqslant \epsilon$. Donc aussi[71]

$$\|\frac{f-g}{2}\|_p^q \geqslant 2^{-q}\epsilon^q. \tag{27.616}$$

EQooCGDDooWtDokf

Et par un autre ailleurs,

$$\|\frac{f+g}{2}\|_p = \frac{1}{2}\|f+g\|_p \leqslant \frac{1}{2}\big(\|f\|_p + \|g\|_p\big) \leqslant 1. \tag{27.617}$$

EQooOFWYooLVvNDc

Vu que nous avons $q \geqslant 2$, cela donne aussi

$$\|\frac{f+g}{2}\|_p \leqslant \|\frac{f+g}{2}\|^q. \tag{27.618}$$

EQooGMPRooGiLSss

Avec les inégalités (27.616) et 27.618 nous finissons l'inégalité (27.615) :

$$\|\frac{f+g}{2}\|_p \leqslant \|\frac{f+g}{2}\|_p^q \leqslant 1 - 2^{-q}\epsilon^q \leqslant \|g\|_p - \delta(\epsilon). \tag{27.619}$$

Okay, c'est bon.

(2) $\underline{2 \leqslant p < \infty}$ Il s'agit de faire la même chose en partant de Clarkson (27.584). Le résultat est que la fonction $\delta(\epsilon) = (\epsilon/2)^p$, ça fonctionne.

□

27.14.3 Théorème de la projection normale

PROPooTZMRooCvQtGg

Proposition 27.156.
Si $1 < p < \infty$, et si V est un sous-espace vectoriel fermé de $L^p(\Omega, \mathcal{A}, \mu)$, alors la projection normale[72] *de $a \in L^p$ sur V existe et est unique.*

70. Définition 27.150.
71. Ici j'ai un coefficient un peu différent que celui de [692]. Écrivez-moi pour confirmer ou infirmer mes calculs.
72. Définition 27.152.

Démonstration. La proposition 27.155 nous indique que l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est uniformément convexe. Or le théorème 27.154 nous indique que les espaces uniformément convexes vérifient la présente proposition. □

Nous pouvons donner une preuve directe, sans passer par l'uniforme convexité, dans les cas $p \geqslant 2$.

THOooRJFUooQivDKm

Théorème 27.157 (Théorème de la projection normale[693])**.**
Nous considérons $p \geqslant 2$. Soit un sous-espace vectoriel fermé $W \subset L^p(\Omega, \mathcal{A}, \mu)$ et $u_0 \in L^p$. Nous notons

$$d(u_0, W) = \inf_{w \in W} d(u_0, W). \tag{27.620}$$

Alors il existe $w_0 \in W$ tel que $\|u_0 - w_0\| = d(u_0, W)$.

Démonstration. Nous allons séparer trois cas : $p = 2$ et $p > 2$.

(1) $p = 2$ Pour $p = 2$, nous savons que L^2 est un espace de Hilbert [73], et nous avons déjà le théorème de la projection 25.7.

(2) $p > 2$ Pour chaque $x \in \Omega$ nous avons $f(x), g(x) \in \mathbb{C}$ et donc l'identité du parallélogramme [74] :

$$\big|f(x) - g(x)\big|^2 + \big|f(x) + g(x)\big|^2 = 2|f(x)|^2 + 2|g(x)|^2. \tag{27.621}$$

EQooUBFEooDUjLnb

Vu que $p > 2$, la fonction $s\colon x \mapsto x^{p/2}$ est convexe (lemme 17.87). Calcul :

$$\begin{aligned}
|f(x) - g(x)|^p + |f(x) + g(x)|^p &= \big(|f(x) - g(x)|^2\big)^{p/2} + \big(|f(x) + g(x)|^2\big)^{p/2} && \text{(27.622a)}\\
&= s\big(|\dots|^2\big) + s\big(|\dots|^2\big) && \text{(27.622b)}\\
&\leqslant \big(|f(x) - g(x)|^2 + |f(x) + g(x)|^2\big)^{p/2} && \text{(27.622c)}\\
&= \big(2|f(x)|^2 + 2|g(x)|^2\big)^{p/2} && \text{(27.622d)}\\
&= 2^{p/2}\big(|f(x)|^2 + |g(x)|^2\big)^{p/2} && \text{(27.622e)}\\
&\leqslant 2^{p/2}2^{p/2-1}\big(|f(x)|^p + |g(x)|^p\big) && \text{(27.622f)}\\
&= 2^{p-1}\big(|f(x)|^p + |g(x)|^p\big) && \text{(27.622g)}
\end{aligned}$$

SUBEQooRHAEooHkYNLH
SUBEQooQFSLooJkoecN
SUBEQooQSUHooXKaWwQ

Justifications :

— Pour (27.622c) : la convexité de s.
— Pour (27.622d) : la relation (27.621).
— Pour (27.622f) : la seconde inégalité du lemme 27.622f.

Nous isolons $|f(x) - g(x)|^p$:

$$\begin{aligned}
|f(x) - g(x)|^p &\leqslant 2^{p-1}\big(|f(x)|^p + |g(x)|^p\big) - |f(x) + g(x)|^p && \text{(27.623a)}\\
&= 2^p\left(\frac{|f(x)|^p + |g(x)|^p}{2} - \left|\frac{|f(x)| + |g(x)|}{2}\right|^p\right) && \text{(27.623b)}
\end{aligned}$$

Cette inégalité étant valable pour tout x, nous pouvons intégrer sur Ω et découper l'intégrale en petits morceaux :

$$\|f - g\|_p^p \leqslant 2^p\left(\frac{\|f\|_p^p + \|g\|_p^p}{2} - \|\frac{f+g}{2}\|_p^p\right). \tag{27.624}$$

EQooVNHSooPXiFNC

Voilà une bonne chose de prouvée. Nous pouvons maintenant passer au vif du sujet.

Soit une suite w_j dans W telle que $\|u_0 - w_j\| \to d(u_0, W)$. Trois choses à savoir sur cette suite :

(2a) Une telle suite existe parce que $d(u_0, W)$ est défini comme un infimum.

73. Lemme 27.92.
74. Théorème 10.1 en remarquant que $(z_1, z_2) \mapsto z_1\bar{z}_2$ est un produit scalaire hermitien sur $\mathbb{C}$.

(2b) Rien ne garantit qu'elle converge.

(2c) Même si elle convergeait, rien ne garantirait que la limite soit encore dans W.

Le troisième point est facile à régler : vu que W est fermé par hypothèse, une suite convergente contenue dans W a sa limite dans W. Nous allons régler la convergence de w_j en prouvant qu'elle est de Cauchy.

Remarquons que W est vectoriel, donc $(w_j + w_k)/2$ est dans W pour tout j et k ; donc

$$\| \frac{w_j + w_k}{2} - u_0 \| \geqslant d(u_0, W). \tag{27.625}$$

En tenant compte de cela, nous écrivons l'inégalité (27.624) avec $f = w_j - u_0$ et $g = w_k - u_0$:

$$\|f - g\|_p^p = \|w_j - w_k\|_p^p \leqslant 2^p \left(\frac{\|w_j - u_0\|^p + \|w_k - u_0\|^p}{2} - d(u_0, W) \right). \tag{27.626}$$

Soit $\epsilon > 0$ et $0 < \epsilon_1, \epsilon_2 < \epsilon$ tels que $\epsilon_1 + \epsilon_2 < \epsilon$. Il existe un N tel que si $j, k > N$ alors $\|w_j - u_0\|^p \leqslant d(u_0, W)^p + \epsilon_1$ et $\|w_k - u_0\|^p \leqslant d(u_0, W)^p + \epsilon_2$. Pour de telles valeurs de j et k, nous avons

$$\|w_j - w_k\|_p \leqslant 2 \left(\frac{\epsilon_1 + \epsilon_2}{2} \right) < 2\epsilon^{1/p}. \tag{27.627}$$

Donc la suite (w_j) est de Cauchy.

L'espace L^p étant complet par le théorème 27.44, nous en déduisons que (w_j) converge dans L^p. Mais comme W est fermé, nous avons $w_j \xrightarrow{L^p} w \in W$.

En termes de normes, nous avons

$$\|w - u_0\| = \lim_j \|w_j - u_0\| = d(W, u_0). \tag{27.628}$$

□

27.15 Jauge

DEFooRNZBooLSdowc

Définition 27.158 ([216]).
*Soit un ouvert connexe C dans un espace vectoriel topologique E. Nous supposons que $0 \in C$. La **jauge** de C est l'application*

$$\begin{aligned} j_C \colon E &\to \mathbb{R} \\ x &\mapsto \inf\{t > 0 \text{ tel que } \frac{x}{t} \in C\}. \end{aligned} \tag{27.629}$$

LEMooEUETooMAAnNO

Lemme 27.159 ([1]).
Soit un ouvert connexe C dans un espace vectoriel topologique E. Nous supposons que $0 \in C$, et nous considérons la jauge $j_C \colon E \to \mathbb{R}$. Nous posons

$$A_x = \{t > 0 \text{ tel que } \frac{x}{t} \in C\}. \tag{27.630}$$

Nous avons :

ITEMooORQGooUkHWKJ

(1) $]j_C(x), \infty[\subset A_x$.

ITEMooFZSGooLuyzcw

(2) Si $x_0 \notin C$, alors $j_C(x_0) \geqslant 1$.

Démonstration. Soit $t \in A_x$. Vu que C est connexe et contient 0, vu que $x/t \in C$ nous avons

$$(1 - \lambda)0 + \lambda \frac{x}{t} \in C \tag{27.631}$$

pour tout $\lambda \in [0, 1]$. C'est à dire que $(\lambda/t)x \in C$ pour tout $\lambda \in [0, 1]$. Donc $(t/\lambda) \in A_x$, et donc $[t, \infty[\subset A_x$. Soit maintenant $t > j_C(x)$. Alors il existe $t_0 \in [j_C(x), t[\cap A_x$. Dans ce cas, $t \in [t_0, \infty[\subset A_x$. Cela prouve le point (1).

Nous passons au point (2). Soit $x_0 \notin C$. Si $j_C(x_0) < 1$, alors nous avons $1 \in A_{x_0}$ par le point (1). Et si $1 \in A_{x_0}$, alors $x_0 \in C$, ce qui est contraire à l'hypothèse. Donc $j_C(x_0) \leqslant 1$. □

PROPooEBPPooDWUwLx

Proposition 27.160 ([216]).
Soient un espace vectoriel topologique E, un ouvert connexe $C \subset E$ tel que $0 \in C$. Nous la jauge $j_C : E \to \mathbb{R}$ a les propriétés suivantes.

ITEMooAOUUooGhrpk

(1) Pour tout $x \in E$ et $\lambda > 0$ nous avons $f_C(\lambda x) = \lambda j_C(x)$.

ITEMooQWIQooMRramc

(2) Pour tout $x, y \in E$ nous avons $j_C(x + y) \leqslant j_C(x) + j_C(y)$.

ITEMooKGNFooKAEUPH

(3) Si C est symétrique[75], alors j_C est une seminorme[76]

Démonstration. Nous posons

$$A_x = \{t > 0 \text{ tel que } \frac{x}{t} \in C\} \tag{27.632}$$

et $A = A_x$, $B = A_{\lambda x}$.

(1) $\underline{B = \lambda A}$ Soit $t \in B$. Alors $x/(t/\lambda) \in C$. Et donc $t/\lambda \in A$ et donc $t \in \lambda A$. De même si $t \in A$, alors $x/t \in C$ et donc $(\lambda x)/(\lambda t) \in C$ de telle sorte que $\lambda t \in B$.

(2) $\underline{\textbf{infimum, point (1)}}$ La proposition 1.463 dit que $\inf(B) = \inf(\lambda A) = \lambda \inf(A)$. Cela prouve le point (1).

(3) $\underline{\textbf{Pour (3)}}$ Nous savons par le lemme 27.159(1) $]j_C(x), \infty[\subset A_x$. Si $\epsilon > 0$, nous avons donc $j_C(x) + \epsilon \in A_x$ et donc

$$\frac{x}{j_C(x) + \epsilon} \in C, \quad \frac{y}{j_C(y) + \epsilon} \in C. \tag{27.633}$$

Comme C est convexe, nous pouvons prendre une combinaison convexe de ces deux éléments. L'astuce est d'écrire

$$\lambda x + (1 - \lambda)y \in C \tag{27.634}$$

avec

$$\lambda = \frac{j_C(x) + \epsilon}{j_C(x) + j_C(y) + 2\epsilon}, \tag{27.635}$$

et

$$1 - \lambda = \frac{j_C(y) + \epsilon}{j_C(x) + j_C(y) + 2\epsilon}. \tag{27.636}$$

Un simple calcul montre que

$$\frac{x + y}{j_C(x) + j_C(y) + 2\epsilon} \in C, \tag{27.637}$$

c'est à dire $j_C(x) + j_C(y) + 2\epsilon \in A_{x+y}$ ou encore que

$$j_C(x + y) = \inf(A_{x+y}) \leqslant j_C(x) + j_C(y) + 2\epsilon. \tag{27.638}$$

Vu que cette inégalité est vraie pour tout $\epsilon > 0$, nous avons $j_C(x + y) \leqslant j_C(x) + j_C(y)$.

(4) $\underline{\textbf{Pour (3)}}$ Nous devons prouver que $j_C(\lambda x) = |\lambda| j_C(x)$ pour tour λ. Nous l'avons déjà prouvé pour $\lambda > 0$. Vu que $0/t = 0 \in C$ pour tout $t > 0$, nous avons $j_C(0) = 0$ et donc $j_C(\lambda x) = |\lambda| j_C(x)$ pour $\lambda = 0$. Reste à voir cette égalité pour $-\lambda$ lorsque $\lambda > 0$.
Vu que C est symétrique, $A_x = A_{-x}$ et alors

$$j_C(-\lambda x) = \inf\{t > 0 \text{ tel que } \frac{-\lambda x}{t} \in C\} \tag{27.639a}$$

$$= \inf\{t > 0 \text{ tel que } \frac{\lambda x}{t} \in C\} \tag{27.639b}$$

$$= \lambda j_C(x). \tag{27.639c}$$

□

75. C'est à dire que $-x \in C$ si et seulement si $x \in C$.
76. Définition 7.326.

27.16 Théorèmes de Hahn-Banach

27.16.1 Applications $\mathbb{R}$-linéaires et $\mathbb{C}$-linéaires

LEMooBZHIooSQJSnM

Lemme 27.161 ([694, 695]).
Soit un espace vectoriel X sur $\mathbb{C}$. Nous considérons[77] *l'application*

$$\begin{aligned}\psi\colon \mathcal{L}_{\mathbb{R}}(X,\mathbb{R}) &\to \mathcal{L}_{\mathbb{C}}(X,\mathbb{C})\\ f &\mapsto \big[x\mapsto f(x)-if(ix)\big].\end{aligned} \tag{27.640}$$

EQooLYYGooJfKIfu

(1) Cette application est bien définie.

(2) Elle est une bijection.

ITEMooNVKBooIGhzWM

(3) Elle vérifie $\|\psi(f)\|_{\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})}=\|f\|_{\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})}$.

ITEMooTFWOooIIhcnZ

(4) Si $g\in\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})$ se décompose en $g(x)=u(x)+iv(x)$ où u et v sont à valeurs réelles, alors $g=\psi(u)$.

Démonstration. En plusieurs parties.

(1) **Bien définie** Nous devons avant tout prouver que (27.640) a un sens : si $f\in\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})$, alors nous devons prouver que $\psi(f)\in\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})$, c'est-à-dire que $\psi(f)$ est $\mathbb{C}$-linéaire. Soient donc $x,y\in X$ et $\alpha,\beta\in\mathbb{R}$ de telle sorte que $\alpha+\beta i$ soit un élément générique de $\mathbb{C}$. Prouver que

$$\psi(f)(x+y)=\psi(f)(x)+\psi(f)(y) \tag{27.641}$$

est un simple calcul. Ensuite

$$\begin{aligned}\psi(f)\big((\alpha+i\beta)x\big)&=f\big((\alpha+i\beta)x\big)-if\big(i(\alpha+i\beta)x\big) &(27.642a)\\ &=\alpha f(x)+\beta f(ix)-i\alpha f(ix)-i(-\beta)f(x) &(27.642b)\\ &=(\alpha+i\beta)\big(f(x)-if(ix)\big) &(27.642c)\\ &=(\alpha+i\beta)\psi(f)(x). &(27.642d)\end{aligned}$$

(2) **ψ est injective** Soient $f,g\in\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})$ telles que $\psi(f)=\psi(g)$. Nous avons l'égalité

$$f(x)-if(ix)=g(x)-ig(ix), \tag{27.643}$$

EQooOFWHooIJutnV

qui est une égalité dans $\mathbb{C}$. En sachant que les nombres $f(x)$, $f(ix)$, $g(x)$ et $g(ix)$ sont des réels, nous séparons les parties réelles et imaginaires dans (27.643) et nous trouvons $f(x)=g(x)$ et $f(ix)=g(ix)$. Chacune de ces deux égalités nous assurent que $f=g$.

(3) **ψ est surjective** Soit $g\in\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})$. Nous séparons ses parties réelles et imaginaires :

$$g(x)=g_1(x)+ig_2(x). \tag{27.644}$$

Avec $g_1,g_2\in\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})$. Nous allons prouver que $g=\psi(g_1)$.
Ensuite, utilisant la $\mathbb{C}$-linéarité de g pour calculer $g(ix)$ de deux façon différentes. D'une part

$$g(ix)=g_1(ix)+ig_2(ix), \tag{27.645}$$

EQooRMJAooQlTJLY

et d'autre part

$$g(ix)=i\big(g_1(x)+ig_2(x)\big). \tag{27.646}$$

EQooCJQHooFVJXGa

En égalisant les parties réelles et imaginaires de (27.645) et de (27.646),

$$\begin{aligned}g_1(x)&=g_2(ix) &(27.647a)\\ g_2(x)&=-g_1(ix), &(27.647b)\end{aligned}$$

77. Voir la définition 4.29 pour les notations.

et en particulier $g_1(ix) = g_2(-x)$.

Nous pouvons maintenant calculer

$$\begin{aligned}\psi(g_1)(x) &= g_1(x) - ig_1(ix) && (27.648a)\\ &= g_1(x) - ig_2(-x) && (27.648b)\\ &= g_1(x) + ig_2(x) && (27.648c)\\ &= g(x). && (27.648d)\end{aligned}$$

(4) **Norme** Nous prouvons maintenant le point (3) qui prétend que $\|\psi(f)\|_{\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})} = \|f\|_{\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})}$. Nous allons montrer les inégalités dans les deux sens.

(4a) $\|f\| \leqslant \|\psi(f)\|$

D'abord nous prouvons que $\|f\| \leqslant \|\psi(f)\|$. Pour cela, nous considérons $x \in X$, et nous écrivons

$$|f(x)| \leqslant |f(x) - if(ix)| = |\psi(f)x| \tag{27.649}$$

dont la première inégalité est vraie parce que $f(x)$ et $f(ix)$ sont réels. Avec cette inégalité nous avons tout de suite

$$\sup_{\|x\|=1} |f(x)| \leqslant \sup_{\|x\|=1} |\psi(f)x|. \tag{27.650}$$

(4b) $\|f\| \geqslant \|\psi(f)\|$

Nous passons maintenant à l'autre sens. Soit $x \in X$. Nous commençons par trouver $\alpha \in S^1$ (c'est-à-dire $|\alpha| = 1$) tel que $f(\alpha ix) = 0$. Si $f(ix) = 0$, alors $\alpha = 1$ fait l'affaire. Sinon, supposons $f(ix) \neq 0$, et considérons l'application

$$\begin{aligned}s\colon]0, 2\pi[&\to \mathbb{R}\\ t &\mapsto f\big(\varphi(t)ix\big).\end{aligned} \tag{27.651}$$

où φ est la bijection continue de la proposition 18.61(1). D'une part, pour $t = \pi$, nous avons

$$s(\pi) = f\big(\varphi(\pi)ix\big) = f(-ix) = -f(ix), \tag{27.652}$$

et d'autre part, pour $t = 0$ nous avons

$$s(0) = f\big(\varphi(0)ix\big) = f(ix) = f(ix). \tag{27.653}$$

Comme nous avons supposé que $f(ix) \neq 0$ le théorème des valeurs intermédiaires 9.93 appliqué à s donne un $t \in]0, \pi[$ tel que $s(t_0) = 0$. En posant $\alpha = \varphi(t_0)$ nous avons le α qu'il nous fallait.

Armés de ce α, nous montrons que $\|\psi(f)\| \leqslant \|f\|$. Nous avons

$$\begin{aligned}|\psi(f)x| &= |\alpha||\psi(f)x| && && (27.654a)\\ &= |\psi(f)(\alpha x)| && \psi(f) \text{ est } \mathbb{C}\text{-linéaire} && (27.654b)\\ &= |f(\alpha x) - if(i\alpha x)| && && (27.654c)\\ &= |f(\alpha x)|. && && (27.654d)\end{aligned}$$

Nous avons prouvé que pour tout $x \in X$ tel que $\|x\| = 1$, il existe $y \in X$ tel que $\|y\| = 1$ et $|\psi(f)x| = |f(y)|$. En l'occurrence, $y = \alpha x$. Notez que α est une fonction de x. Nous avons donc

$$\sup_{\|y\|=1} |f(y)| \geqslant \sup_{\|x\|=1} |\psi(f)x|, \tag{27.655}$$

autrement dit $\|f\| \geqslant \|\psi(f)\|$.

(5) **Décomposition** Nous prouvons le point (4). Supposons $g(x) = u(x) + iv(x)$. Nous avons $g(ix) = ig(x)$ par $\mathbb{C}$-linéarité de g. Donc

$$u(ix) + iv(ix) = i\big(u(x) + iv(x)\big), \tag{27.656}$$

c'est-à-dire $u(ix) + iv(ix) = iu(x) - v(x)$. En égalisant les partie imaginaires, $v(x) = -u(ix)$. En remettant dans l'expression de g,

$$g(x) = u(x) + iv(x) = u(x) + i\big(-u(ix)\big) = u(x) - iu(ix) = \psi(u)x, \tag{27.657}$$

ce qu'il fallait démontrer.

□

LEMooUFMFooEXecXE

Lemme 27.162 ([694]).
Soient un espace vectoriel X sur $\mathbb{C}$ ainsi qu'une seminorme[78] *p. Soit $f \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$. Nous avons*

$$|\psi(f)(x)| \leqslant p(x) \quad \forall x \in X \tag{27.658}$$

si et seulement si

$$|f(x)| \leqslant p(x) \quad \forall x \in X \tag{27.659}$$

Démonstration. En deux parties.

(1) $\Rightarrow$ Nous supposons que $f \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ vérifie $|\psi(f)(x)| \leqslant p(x)$ pour tout x. L'observation à faire est $f(x) = \operatorname{Re}\big(\psi(f)(x)\big)$. Nous avons alors[79]

$$|f(x)| = |\operatorname{Re}\big(\psi(f)(x)\big)| \leqslant |\psi(f)(x)| \leqslant p(x). \tag{27.660}$$

(2) $\Leftarrow$ Nous supposons que $|f(x)| \leqslant p(x)$ pour tout x. Soit $x \in X$. Si $\psi(f)(x) = 0$, alors nous avons bien $|\psi(f)(x)| \leqslant p(x)$ parce qu'une seminorme est toujours positive. Nous supposons donc que $\psi(f)(x) \neq 0$.
La décomposition polaire (proposition 18.62) du nombre complexe $\psi(f)x$ donne un $\theta \in \mathbb{R}$ tel que $\psi(f)x = e^{i\theta}|\psi(f)x|$. Donc

$$|\psi(f)(x)| = e^{-i\theta}\psi(f)(x) = \psi(f)(e^{-i\theta}x). \tag{27.661}$$

EQooFYVFooQvQTHz

Ces égalités montrent entre autres que $\psi(f)(e^{-i\theta}x) \in \mathbb{R}^+$; et en particulier, il est égal à sa partie réelle. Vu la définition (27.640), ça nous dit que

$$\psi(f)(e^{-i\theta}x) = f(e^{-i\theta}x). \tag{27.662}$$

Nous pouvons alors continuer les égalités (27.661) en mettant des normes partout :

$$|\psi(f)(x)| = |\psi(f)(e^{-i\theta}x)| = |f(e^{-i\theta}x)| \leqslant p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x). \tag{27.663}$$

□

27.16.2 Hahn-Banach, théorème d'extension dominée

THOooXALCooFrkvDo

Théorème 27.163 (Hahn-Banach, extension dominée cas réel[696, 216]).
Soit E, un espace vectoriel réel et une application $p\colon E \to \mathbb{R}$ satisfaisant[80]

(1) $p(\lambda x) = \lambda p(x)$ pour tout $x \in E$ et pour tout $\lambda > 0$,

(2) $p(x + y) \leqslant p(x) + p(y)$ pour tout $x, y \in E$.

78. Définition 7.326.

79. Dans [694], l'auteur sépare le calcul en deux parties : une majoration pour $f(x)$ et une pour $-f(x)$. Je ne vois pas où est le mal à la faire d'un seul coup.

80. Ce n'est pas tout à fait une seminorme (définition 7.326) à cause de la valeur absolue. Le théorème 27.165 parlera pour les vraies seminormes, et pour les espaces sur $\mathbb{C}$.

Soit de plus $M \subset E$ un sous-espace vectoriel muni d'une application $g\colon M \to \mathbb{R}$ vérifiant $g(m) \leqslant p(m)$ pour tout $m \in M$. Alors il existe $f \in \mathcal{L}(E, \mathbb{R})$ telle que

(1) L'application f étend g : $f(m) = g(m)$ pour tout $m \in M$,

(2) L'application f reste dominée par p : $f(x) \leqslant p(x)$ pour tout $x \in E$.

Démonstration. Si h est une application linéaire définie sur un sous-espace de E, nous notons D_h ledit sous-espace.

(1) **Un ensemble inductif**
Nous considérons P, l'ensemble des fonctions linéaires suivant

$$P = \left\{ h\colon D_h \to \mathbb{R} \text{ tel que } \begin{cases} M \subset D_h \\ h(m) = g(m) & \forall m \in M \\ h(x) \leqslant p(x) & \forall x \in D_h \end{cases} \right\} \tag{27.664}$$

Cet ensemble est non vide parce que g est dedans. Nous le munissons de la relation d'ordre $h_1 \leqslant h_2$ si et seulement si $D_{h_1} \subset D_{h_2}$ et h_2 prolonge h_1. Nous montrons à présent que P est un ensemble inductif. Soit un sous-ensemble totalement ordonné $Q \subset P$; nous définissons une fonction h de la façon suivante. D'abord $D_h = \sup_{l \in Q} D_l$ et ensuite

$$\begin{aligned} h\colon D_h &\to \mathbb{R} \\ x &\mapsto l(x) \quad \text{si } x \in D_l \end{aligned} \tag{27.665}$$

Cela est bien défini parce que si $x \in D_l \cap D_{l'}$ alors, vu que Q est totalement ordonné (i.e. $l \leqslant l'$ ou $l' \leqslant l$), on a obligatoirement $D_l \subset D_{l'}$ et l' qui prolonge l (ou le contraire). Donc h est un majorant de Q dans P parce que $h \geqslant l$ pour tout $l \in Q$. Cela montre que P est inductif (définition 1.22). Le lemme de Zorn 1.23 nous dit alors que P possède un élément maximal f qui va être la réponse à notre théorème.

(2) **Le support de f**
La fonction f est dans P ; donc $f(x) \leqslant p(x)$ pour tout $x \in D_h$ et $f(m) = g(m)$ pour tout $m \in M$. Pour terminer nous devons montrer que $D_f = E$. Supposons donc que $D_f \neq E$ et prenons $x_0 \notin D_f$. Nous allons contredire la maximalité de f en considérant la fonction h donnée par $D_h = D_f + \mathbb{R}x_0$ et

$$h(x + tx_0) = f(x) + t\alpha \tag{27.666}$$

où α est une constante que nous allons fixer plus tard.
Nous commençons par prouver que h est dans P. Nous devons prouver que

$$h(x + tx_0) = f(x) + t\alpha \leqslant p(x + tx_0) \tag{27.667}$$

Pour cela nous allons commencer par fixer α pour avoir les relations suivantes :

$$\begin{cases} f(x) + \alpha \leqslant p(x + x_0) & \text{(27.668a)} \\ f(x) - \alpha \leqslant p(x - x_0) & \text{(27.668b)} \end{cases}$$

pour tout $x \in D_f$. Ces relations sont équivalentes à demander α tel que

$$\begin{cases} \alpha \leqslant p(x + x_0) - f(x) & \text{(27.669a)} \\ \alpha \geqslant f(x) - p(x - x_0) & \text{(27.669b)} \end{cases}$$

Nous nous demandons donc si il existe un α qui satisfasse

$$\sup_{y \in D_f} \big(f(y) - p(y - x_0)\big) \leqslant \alpha \leqslant \inf_{z \in D_f} \big(p(z + x_0) - f(z)\big). \tag{27.670}$$

Ou encore nous devons prouver que pour tout $y, z \in D_f$,

$$p(z + x_0) - f(x) \geqslant f(y) - p(y - x_0) \geqslant 0. \tag{27.671}$$

Par les propriétés de p et de f,

$$p(z + x_0) + p(y - x_0) - f(z) - f(y) \geqslant p(z + y) - f(z + y) \geqslant 0. \tag{27.672}$$

La dernière inégalité est le fait que $f \in P$. Un choix de α donnant les inéquations (27.668) est donc possible.

À partir des inéquations (27.668) nous obtenons la relation (27.667) de la façon suivante. Si $t > 0$ nous multiplions l'équation (27.668a) par t :

$$tf(x) + t\alpha \leqslant tp(x + x_0). \tag{27.673}$$

Et nous écrivons cette relation avec x/t au lieu de x en tenant compte de la linéarité de f :

$$f(x) + t\alpha \leqslant tp\big(\frac{x}{t} + x_0\big) = p(x + tx_0). \tag{27.674}$$

Avec $t < 0$, c'est similaire, en faisant attention au sens des inégalités.

Nous avons donc construit $h\colon D_h \to \mathbb{R}$ avec $h \in P$, $D_f \subset D_h$ et $h(x) = f(x)$ pour tout $x \in D_f$. Cela pour dire que $h > f$, ce qui contredit la maximalité de f. Le domaine de f est donc E tout entier.

La fonction f est donc une fonction qui remplit les conditions.

□

Lemme 27.164 ([251])**.**
Soient un espace vectoriel complexe X, et une application linéaire $f\colon X \to \mathbb{C}$. Il existe une application $u \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ telle que

(1) Pour tout $x \in X$, nous avons

$$f(x) = u(x) - iu(ix) \tag{27.675}$$

(2) Nous avons égalité des normes[81]

$$\|f\|_{\mathcal{L}_{\mathbb{C}}(X,\mathbb{C})} = \|u\|_{\mathcal{L}_{\mathbb{R}}(X,\mathbb{R})}. \tag{27.676}$$

Démonstration. Nous décomposons, pour chaque $x \in X$, l'élément $f(x)$ en parties réelles et imaginaires :

$$f(x) = u(x) + iv(x). \tag{27.677}$$ EQooZJZIooBQxffg

Il ne reste plus qu'à trouver le lien entre u et v. Pour cela, nous écrions $f(ix) = if(x)$ en substituant $f(x)$ et $f(ix)$ par leurs valeurs en termes de u et v :

$$u(ix) + iv(ix) = iu(x) - v(x). \tag{27.678}$$

En égalisant les parties réelles, nous avons $u(ix) = -v(x)$ et donc bien

$$f(x) = u(x) + iv(x) = u(x) - iu(ix). \tag{27.679}$$ EQooZUTLooVzwRDM

Nous prouvons maintenant que u est $\mathbb{R}$-linéaire. Il suffit de considérer $\lambda \in \mathbb{R}$ et d'écrire $f(\lambda x) = \lambda f(x)$ en substituant (27.677) pour $f(\lambda x)$ et $f(x)$:

$$u(\lambda x) + iv(\lambda x) = \lambda\big(u(x) + iv(x)\big). \tag{27.680}$$

En égalisant encore les parties réelles, nous avons $u(\lambda x) = \lambda u(x)$. Nous faisons de même pour prouver que $u(x + y) = u(x) + u(y)$. □

81. Pas encore démontrée...

THOooVQLJooWuBMoZ

Théorème 27.165 (Hahn-Banach complexe[251])**.**
Soient un espace vectoriel X sur $\mathbb{K}$ (qui est $\mathbb{R}$ ou $\mathbb{C}$), et une seminorme $p\colon X \to \mathbb{R}$. Soient un sous-espace vectoriel M de X ainsi qu'une application linéaire $\ell \in \mathcal{L}(M, \mathbb{K})$ telle que

$$|\ell(m)| \leqslant p(m) \ \forall m \in M. \tag{27.681}$$

Alors il existe une extension linéaire $f \in \mathcal{L}(X, \mathbb{K})$ telle que

(1) $f(m) = \ell(m)$ pour tout $m \in M$,

(2) $|f(x)| \leqslant p(x)$ pour tout $x \in X$.

Démonstration. Nous considérons X comme espace vectoriel réel (de dimension double de celle de X comme espace vectoriel réel). Là-dedans, l'application p vérifie $p(\lambda x) = \lambda p(x)$ pour tout $\lambda > 0$. D'autre part nous décomposons l'application $\ell\colon M \to \mathbb{C}$ en parties réelles et imaginaires :

$$\ell(x) = u(x) + iv(x), \tag{27.682}$$

où $u\colon M \to \mathbb{R}$ est une application vérifiant, pour tout $m \in M$,

$$u(m) \leqslant |u(m)| \leqslant |\ell(m)| \leqslant p(m). \tag{27.683}$$

Donc le théorème en version réelle 27.163 s'applique et il existe une extension $R\colon X \to \mathbb{R}$ de u qui vérifie $R \leqslant p$ sur X. Nous posons alors

$$f(x) = R(x) - iR(ix). \tag{27.684}$$

En vertu du lemme 27.161, cela est une application $\mathbb{C}$-linéaire sur X, et nous avons $f = \ell$ sur M parce que leurs parties réelles sont égales.

De plus nous avons $R \leqslant p$ et donc $|f| \leqslant p$ par le lemme 7.330. □

27.16.3 Hyperplan séparateur

DEFooQUDEooMrgcCV

Définition 27.166 (Hyperplan qui sépare)**.**
*Soit E un espace vectoriel topologique ainsi que A, B des sous-ensembles de E. Nous disons que l'hyperplan d'équation $f = \alpha$ **sépare au sens large** les parties A et B si $f(x) \leqslant \alpha$ pour tout $x \in A$ et $f(x) \geqslant \alpha$ pour tout $x \in B$.*

*La séparation est **au sens strict** si il existe $\epsilon > 0$ tel que*

$$f(x) \leqslant \alpha - \epsilon \qquad \text{pour tout } x \in A \tag{27.685a}$$

$$f(x) \geqslant \alpha + \epsilon \qquad \text{pour tout } x \in B. \tag{27.685b}$$

PROPooSKBXooQDOmcL

Proposition 27.167 ([216])**.**
Soit un espace vectoriel topologique E. Soit un ouvert connexe non vide $C \subset E$. Soit $x_0 \notin C$. Alors il existe $f \in E'$ [82] tel que $f(x) < f(x_0)$ pour tout $x \in C$.

Démonstration. En deux parties.

(1) **Si $O \in C$** Nous commençons par supposer que $0 \in C$, de telle sorte à pouvoir considérer la jauge [83] $j_C\colon E \to \mathbb{R}$. Nous posons $M = \mathbb{R}x_0$ et

$$\begin{aligned} g\colon M &\to \mathbb{R} \\ tx_0 &\mapsto t. \end{aligned} \tag{27.686}$$

82. E' est l'ensemble des applications linéaires continues sur E à valeurs dans $\mathbb{R}$.
83. Définition 27.158.

Vu que $x_0 \notin C$, le lemme 27.159(2) dit que $j_C(x_0) \geqslant 1 = g(x_0)$. Nous multiplions cette inégalité par $t \geqslant 0$ pour obtenir $tj_C(x_0) \geqslant tg(x_0)$. Vu que $tj_C(x_0) = j_C(tx_0)$ (proposition 27.160(1)), et que $tg(x_0) = g(tx_0)$, nous avons

$$j_C(tx_0) \geqslant g(tx_0) \tag{27.687}$$

EQooYSVTooFKJabg

pour tout $t \geqslant 0$. Si $t < 0$, alors nous avons $j_C(tx_0) \geqslant 0$ et $g(tx_0) = t < 0$. Donc l'inégalité (27.687) est valide pour tout $t \in \mathbb{R}$. Nous avons donc $j_C \geqslant g$ sur M. Le théorème 27.163 dit qu'il existe $f \in \mathcal{L}(E, \mathbb{R})$ tel que

(1a) $f(x) = g(x)$ pour tout $x \in M$.

(1b) $f(x) \leqslant j_C(x)$ pour tout $x \in E$.

Pour $x \in C$ en particulier nous avons $f(x) \leqslant j_C(x) < 1 = f(x_0)$. Nous devons encore prouver que f est continue. Pour cela nous considérons l'ouvert $U = C \cap (-C)$, qui est un voisinage ouvert de 0. Soit $x \in U$. Nous avons $f(x) \leqslant j_C(x) < 1$ parce que $x \in C$, mais aussi $f(-x) \leqslant j_C(-x) < 1$ parce que $-x \in C$. Vu que f est linéaire, l'inégalité $f(-x) < 1$ donne $f(x) > -1$. Au final nous avons prouvé que

$$f(x) \in]-1, 1[\tag{27.688}$$

pour tout $x \in U$. La proposition 7.224 dit qu'alors f est continue.

(2) **Si $0 \notin C$** Enfin nous prouvons le cas où 0 n'est pas dans C. Soit $c \in C$. Nous posons $C' = C - c$ et $x_0' = x_0 - c$ de telle sorte que $0 \in C'$ et $x_0' \in C'$. Par la première partie, il existe $f \in E'$ telle que $f(x) < f(x_0')$ pour tout $x \in C'$. Nous posons ensuite

$$f_1(x) = f(x - c). \tag{27.689}$$

Nous avons alors, pour tout $x \in C$ que

$$f_1(x) = f(x - c) < f(x_0') = f(x_0 - c) = f_1(x_0). \tag{27.690}$$

parce que $x - c \in C'$.

L'application f_1 est encore continue et linéaire. Elle répond donc à la question.

□

ThoSAJjdZc

Théorème 27.168 (Hahn-Banach, première forme géométrique[216])**.**
Soit E un espace vectoriel topologique et A, B deux convexes non vides disjoints de E. Si A est ouvert, il existe un hyperplan fermé qui sépare A et B au sens large[84]*.*

Démonstration. Vu que A et B sont convexes, la proposition 7.156 dit que $A - B$ est convexe. De plus $A - B$ est ouvert et ne contient pas zéro. Nous pouvons donc lui appliquer la proposition 27.167 : il existe $f \in E'$ tel que $f(x) < f(0) = 0$ pour tout $x \in A - B$. Autrement dit, pour tout $a, b \in A \times B$ nous avons $f(a) < f(b)$. Cela implique

$$\sup\{f(a) \text{ tel que } a \in A\} \leqslant \inf\{f(b) \text{ tel que } b \in B\}. \tag{27.691}$$

En posant $\alpha \in \mathbb{R}$ tel que

$$\sup\{f|_A\} \leqslant \alpha \leqslant \inf\{f|_B\}, \tag{27.692}$$

l'hyperplan $\{f = \alpha\}$ sépare bien A et B au sens large. □

ThoACuKgtW

Théorème 27.169 (Hahn-Banach, seconde forme géométrique)**.**
Soient un espace vectoriel topologique localement convexe[85] *E ainsi que des parties A, B qui sont*

(1) des convexes

84. Définition 27.166.
85. Définition 7.197.

(2) *non vides*

(3) *disjoints*

(4) *A compact et B soit fermé.*

Alors il existe un hyperplan fermé qui sépare strictement A et B.

Démonstration. Vu que B est fermé, A est dans l'ouvert $E\backslash B$. Donc si $a\in A$, il existe un voisinage ouvert convexe de a inclus dans $E\backslash B$. Soit U_a un voisinage ouvert et convexe de 0 tel que $(a+U_a)\cap B=\varnothing$.

Vu que la fonction $(x,y)\mapsto x+y$ est continue, nous pouvons trouver un ouvert convexe V_a tel que $V_a+V_a\subset U_a$. L'ensemble $a+V_a$ est alors un voisinage ouvert de a et bien entendu $\bigcup_a(a+V_a)$ recouvre A qui est compact. Nous en extrayons un sous-recouvrement fini, c'est-à-dire que nous considérons $a_1,\dots,a_n\in A$ tels que

$$A\subset\bigcup_{i=1}^{n}(a_i+V_{a_i}).\tag{27.693}$$

Nous posons alors

$$V=\bigcap_{i=1}^{n}V_{a_i}.\tag{27.694}$$

Cet ensemble est non vide parce et il contient un voisinage de zéro parce que c'est une intersection finie de voisinages de zéro. Soit $x\in A+V$. Il existe i tel que

$$x\in a_i+U_{a_i}+V\subset a_i+V_{a_i}+V_{a_i}\subset a_i+U_{a_i}\subset E\backslash B.\tag{27.695}$$

Donc $(A+V)\cap B=\varnothing$. L'ensemble $A+V$ est alors un ouvert convexe disjoint de B. Par la première forme géométrique du théorème de Hahn-Banach 27.168 nous avons un hyperplan qui sépare $A+V$ de B au sens large : il existe $f\in E'\backslash\{0\}$ tel que $f(a)+f(v)\leqslant f(b)$ pour tout $a\in A$, $v\in V$ et $b\in B$.

Il suffit donc de trouver un $v\in V$ tel que $f(v)\neq 0$ pour avoir la séparation au sens strict. Cela est facile : V étant un voisinage de zéro et f étant linéaire, si elle était nulle sur V, elle serait nulle sur E. □

CORooHTZVooFhgrSN

Corolaire 27.170 ([216]).
Soit un espace vectoriel localement convexe. Soit un sous-espace vectoriel F tel que $\bar{F}\neq E$. Alors il existe une application $f\in E'\backslash\{0\}$ telle que $f=0$ sur F.

Démonstration. Soit $a\in E$ avec $a\notin\bar{F}$. Vu que le singleton $\{a\}$ est compact et que $\bar{F}$ est fermé, le théorème de Hahn-Banach 27.169 dit qu'il existe $f\in E'\backslash\{0\}$ et $\alpha\neq 0$ tel que $f<\alpha$ sur $\bar{F}$ et $f(a)>\alpha$.

Étant donné que F est vectoriel et que f est linéaire, avoir $f(x)<\alpha$ pour tout x dans F implique d'avoir $f(x)=0$ pour tout $x\in F$. En effet si $f(x)\neq 0$, il existe un λ tel que $f(\lambda x)>\alpha$ ($\mathbb{R}$ est archimédien, théorème 1.440). □

27.16.4 Prolongement de fonctionnelles (dimension finie)

Nous allons prouver quelques lemmes qui permettent de prolonger des fonctionnelles d'un sous-espace vers un sous-espace contenant un nombre fini de dimensions en plus.

LEMooHWSJooGVmIPV

Lemme 27.171 ([697]).
Soient un espace vectoriel réel normé X ainsi qu'un sous-espace M. Soient $\ell\in\mathcal{L}(M,\mathbb{R})$ de norme finie, et $x_1\in X\backslash M$. On pose $M_1=\operatorname{Span}\{M,x_1\}$. Alors il existe $\ell_1\in\mathcal{L}(M_1,\mathbb{R})$ tel que

(1) *$\ell_1(x)=\ell(x)$ pour tout $x\in M$,*

(2) *$\|\ell_1\|_{\mathcal{L}(M_1,\mathbb{R})}=\|\ell\|_{\mathcal{L}(M,\mathbb{R})}$*

Démonstration. Si $\ell=0$, alors c'est facile : on prend $\ell_1=0$. Nous commençons par supposer que $\|\ell\|=1$; nous ferons le cas général ensuite.

(1) **Deux fonctions** Nous considérons les fonctions

$$\begin{aligned} f_+ \colon M &\to \mathbb{R} \\ z &\mapsto \|x_1 + z\| - \ell(z) \end{aligned} \tag{27.696}$$

et

$$\begin{aligned} f_- \colon M &\to \mathbb{R} \\ z &\mapsto -\|x_1 + z\| - \ell(z). \end{aligned} \tag{27.697}$$

Pour tout $z_1, z_2 \in M$ nous avons

$$\begin{aligned} f_+(z_1) - f_-(z_2) &= \|x_1 + z_1\| - \ell(z_1) + \|x_1 + z_2\| + \ell(z_2) && \text{(27.698a)} \\ &= \|x_1 + z_1\| + \|x_1 + z_2\| - \ell(z_1 - z_2) && \text{(27.698b)} \\ &\geqslant \|(x_1 + z_1) - (x_1 + z_2)\| - \ell(z_1 - z_2) && \text{(27.698c)} \\ &= \|z_1 - z_2\| - \ell(z_1 - z_2) && \text{(27.698d)} \\ &\geqslant 0 && \text{(27.698e)} \end{aligned}$$

parce que $\|\ell\| = 1$. Nous en déduisons que $f_+(z_1) \geqslant f_-(z_2)$.

(2) **Un entre les deux**
En posant

$$c_- = \sup_{z \in M} \big(-\|x_1 + z\| - \ell(z) \big) \tag{27.699a}$$

$$c_+ = \inf_{z \in M} \big(\|x_1 + z\| - \ell(z) \big), \tag{27.699b}$$

nous avons $c_- \leqslant c_+$. Nous choisissons $c_1 \in [c_-, c_+]$.

(3) **La définition** Si $x \in M_1 = \operatorname{Span} M, x_1$, alors il existe $y \in M$ et $\lambda \in \mathbb{R}$ tels que

$$x = \lambda x_1 + y. \tag{27.700}$$

Nous posons alors

$$\ell_1(x) = \lambda c_1 + \ell(y). \tag{27.701}$$

Voilà qui définit notre ℓ_1. Nous devons prouver qu'elle satisfait $\|\ell_1\| = 1$ et $\ell_1(x) = \ell(x)$ pour tout $x \in M$. La seconde condition est facile : si $x \in M$, alors $\lambda = 0$ dans (27.701) et nous avons bien $\ell_1(x) = \ell(x)$.
Pour la condition sur la norme, nous allons devoir un peu travailler.

(4) $|\ell_1(x)| \leqslant \|x\|$ **pour tout** $x \in M_1$ Soit $x \in M_1$. Nous avons $x = \lambda x_1 + y$ avec $\lambda \in \mathbb{R}$ et $y \in M$.
i

(4a) **Si $\lambda = 0$** Alors $x \in M$ et $|\ell_1(x)| = |\ell(x)| \leqslant \|x\|$ parce que $\|\ell\| = 1$.

(4b) **Si $\lambda \neq 0$** Soient $z_1, z_2 \in M$. Par définition de c_1, c_+ et c_- nous avons les inégalités

$$-\|x_1 + z_1\| - \ell(z_1) \leqslant c_- \leqslant c_1 \leqslant c_+ \leqslant \|x_1 + z_2\| - \ell(z_2). \tag{27.702}$$

Nous écrivons ces inégalités pour $z_1 = z_2 = y/\lambda$:

$$-\|x_1 + \frac{y}{\lambda}\| - \ell(y/\lambda) \leqslant c_1 \leqslant \|x_1 + \frac{y}{\lambda}\| - \ell(y/\lambda). \tag{27.703}$$

(4c) **Si $\lambda > 0$** Nous multiplions (27.703) par λ et nous profitons de la linéarité de ℓ :

$$-\|\lambda x_1 + y\| - \ell(y) \leqslant \lambda c_1 \leqslant \|\lambda x_1 + y\| + \ell(y), \tag{27.704}$$

donc

$$-\|\lambda x_1 + y\| \leqslant \underbrace{\lambda c_1 + \ell(y)}_{=\ell_1(\lambda x_1 + y)} \leqslant \|\lambda x_1 + y\| \tag{27.705}$$

Nous en déduisons que

$$|\ell(\lambda x_1 + y)| \leqslant \|\lambda x_1 + y\|, \tag{27.706}$$

ce qu'il fallait.

(4d) **Si $\lambda < 0$** Le calcul est le même, mais il faut faire attention à bien reverser les inégalités au bon moment, et en manipulant bien les valeur absolues. Nous avons par exemple

$$\lambda\|x_1 + \frac{y}{\lambda}\| = -\big\||\lambda|x_1 + \frac{|\lambda|y}{\lambda}\big\| = -\| - \lambda x_1 - y\| = -\|\lambda x_1 + y\|. \tag{27.707}$$

En multipliant encore (27.703) par λ, nous trouvons

$$-\lambda\|x_1 + \frac{y}{\lambda}\| - \ell(y) \geqslant \lambda c_1 \geqslant \lambda\|x_1 + \frac{y}{\lambda}\| - \ell(y). \tag{27.708}$$

qui devient

$$\|\lambda x_1 + y\| \geqslant \lambda c_1 + \ell(y) \geqslant -\|\lambda x_1 + y\|, \tag{27.709}$$

qui revient au même que (27.705).

(5) **Première conclusion** Nous avons prouvé que $|\ell_1(x)| \leqslant \|x\|$ pour tout x dans M_1. Cela signifie que $\|\ell_1\|_{M_1'} \leqslant 1$. Pour prouver que $\|\ell_1\| = 1$ nous prouvons l'inégalité inverse :

$$\|\ell_1\|_{M_1'} = \sup_{x\in M_1} \frac{|\ell_1(x)|}{\|x\|} \geqslant \sup_{x\in M} \frac{|\ell_1(x)|}{\|x\|} = \sup_{x\in M} \frac{|\ell(x)|}{\|x\|} = \|\ell\|_M = 1. \tag{27.710}$$

Nous en déduisons que $\|\ell_1\| = 1$ et cela termine la preuve dans le cas où $\|\ell\| = 1$.

Maintenant, si $\|\ell\| = a \neq 1$, nous considérons la forme linéaire $f = \ell/a$ qui satisfait $\|f\| = 1$. Par la partie déjà prouvée, nous définissons une extension $f_1 : M_1 \to \mathbb{R}$ telle que $\|f_1\| = 1$.

Il suffit alors de poser $\ell_1 = af_1$, et nous avons le résultat. □

Ce lemme est également valable pour les complexes.

LEMooBYEGooRswAmh

Lemme 27.172 ([1])**.**
Soit un espace vectoriel normé X sur $\mathbb{C}$. Soit $\ell \in \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$ de norme finie. Soient un sous-espace M de X ainsi qu'un élément $a \in X\backslash M$. Il existe $\ell_1 \in \mathcal{L}_{\mathbb{C}}\big(\mathrm{Span}_{\mathbb{C}}(M, a), \mathbb{C}\big)$ tel que

(1) $\ell_1(x) = \ell(x)$ pour tout $x \in M$,

(2) $\|\ell_1\| = \|\ell\|$.

Démonstration. Nous considérons l'application ψ du lemme 27.161. En posant $f = \psi^{-1}(\ell)$, nous avons $f \in \mathcal{L}_{\mathbb{R}}(M, \mathbb{R})$ vérifiant $\|\ell\| = \|f\|$.

L'espace X peut être vu comme vectoriel sur $\mathbb{R}$. Le lemme 27.171 permet de prolonger f à

$$f_1 \in \mathcal{L}_{\mathbb{R}}\big(\mathrm{Span}_{\mathbb{R}}(M, a), \mathbb{R}\big) \tag{27.711}$$

vérifiant $\|\ell\| = \|f\| = \|f_1\|$.

Vu que M est déjà un espace vectoriel sur $\mathbb{C}$, l'espace $\mathrm{Span}_{\mathbb{C}}(M, a)$ qui nous intéresse est donné par

$$\mathrm{Span}_{\mathbb{C}}(M, a) = \mathrm{Span}_{\mathbb{R}}(M, a, ia). \tag{27.712}$$

Nous pouvons donc utiliser une deuxième fois le lemme 27.171 avec le vecteur ia, et définir une extension

$$f_2 \in \mathcal{L}_{\mathbb{R}}\big(\mathrm{Span}_{\mathbb{R}}(M, a, ia), \mathbb{R}\big) = \mathcal{L}_{\mathbb{R}}\big(\mathrm{Span}_{\mathbb{C}}(M, a), \mathbb{R}\big) \tag{27.713}$$

vérifiant $\|\ell\| = \|f\| = \|f_1\| = \|f_2\|$.

En utilisant à nouveau le lemme 27.161, nous avons encore une extension

$$\psi(f_2) \in \mathcal{L}_{\mathbb{C}}\big(\mathrm{Span}_{\mathbb{C}}(M, a), \mathbb{C}\big) \tag{27.714}$$

vérifiant $\|\ell\| = \|f\| = \|f_1\| = \|f_2\| = \|\psi(f_2)\|$.

Et voilà ! La fonctionnelle $\psi(f_2)$ est celle que nous voulions. □

27.16.5 Prolongement de fonctionnelles (dimension infinie)

Les lemmes 27.171 et 27.172 permettent de prolonger une forme linéaire une dimension réelle ou complexe à la fois. Rien ne nous permet de prolonger d'une infinité de dimensions d'un seul coup. Le théorème de Hahn-Banach va nous permettre de faire une infinité de dimension d'un coup à l'aide du lemme de Zorn.

THOooTZSSooBKfxXE

Théorème 27.173 (Hahn-Banach[697])**.**
Soit X, un espace vectoriel normé sur $\mathbb{K}$ ($=\mathbb{R}$ ou $\mathbb{C}$). Soient un sous-espace vectoriel M, et une fonctionnelle linéaire $\ell \in \mathcal{L}_{\mathbb{K}}(M,\mathbb{K})$ de norme finie. Il existe $f \in \mathcal{L}_{\mathbb{K}}(X,\mathbb{K})$ telle que

(1) $f(x) = \ell(x)$ *pour tout* $x \in M$,

(2) $\|f\|_{\mathcal{L}(X,\mathbb{K})} = \|\ell\|_{\mathcal{L}(M,\mathbb{K})}$.

Démonstration. Nous considérons l'ensemble $\mathcal{E}$ des paires (N,f) telles que

(1) N est un sous-espace de X contenant M,

(2) $f \in \mathcal{L}(N,\mathbb{K})$

(3) $\|f\| = \|\ell\|$

(4) $f(x) = \ell(x)$ pour tout $x \in M$.

Cet ensemble n'est pas vide parce que $(M,\ell) \in \mathcal{E}$. Nous mettons un ordre partiel sur $\mathcal{E}$ en posant $(N,f) \leqslant (N',f')$ si et seulement si $N \subset N'$ et $f'|_N = f$.

(1) **$\mathcal{E}$ est inductif** Nous commençons par prouver que $(\mathcal{E},\leqslant)$ est un ensemble inductif[86]. Soit une partie $\mathcal{F}$ totalement ordonnée dans $\mathcal{E}$.

(1a) **L'espace vectoriel**
Nous commençons par poser

$$Y = \bigcup_{(N,f)\in\mathcal{F}} N. \tag{27.715}$$

Et nous prouvons que c'est un espace vectoriel. Soient $x,y \in Y$. Supposons $x \in N_1$ et $y \in N_2$. Vu que $\mathcal{F}$ est totalement ordonné, nous avons $N_1 \subset N_2$ (ou le contraire). Donc $x+y \in N_2 \subset Y$. De même pour λx avec $\lambda \in \mathbb{K}$ et $x \in Y$.

(1b) **La fonctionnelle** Nous devons trouver une fonctionnelle g sur Y. Soit $y \in Y$. Commençons par prouver que l'ensemble

$$\{f(y) \text{ tel que } (N,f) \in \mathcal{F}, y \in N\} \tag{27.716}$$

est un singleton. Soient $(N_1,f_1),(N_2,f_2) \in \mathcal{F}$ avec $y \in N_1 \cap N_2$. Nous supposons que $(N_1,f_1) \leqslant (N_2,f_2)$ (sinon c'est le contraire). Alors $f_2(y) = f_2|_{N_1}(y) = f_1(y)$. Nous pouvons donc définir

$$\begin{aligned} g\colon Y &\to \mathbb{K} \\ y &\mapsto f(y) \end{aligned} \tag{27.717}$$

où $(N,f) \in \mathcal{F}$ est tel que $y \in N$.

(1c) **g est linéaire** Soient $x,y \in Y$. Nous supposons que $x \in N_1$ et $y \in N_2$ avec $N_1 \subset N_2$. Donc $x,y,x+y \in N_2$ et nous avons $g(x) = f_2(x)$, $g(y) = f_2(y)$ et $g(x+y) = f_2(x+y)$. La linéarité de f_2 fait alors le boulot. Même raisonnement pour $g(\lambda x) = \lambda g(x)$.

(1d) **g se restreint à ℓ** Soit $x \in M$. Nous avons $g(x) = f(x)$ pour un couple $(N,f) \in \mathcal{F}$ vérifiant $x \in N$. Vu que f prolonge ℓ nous avons $g(x) = f(x) = \ell(x)$.

86. Définition 1.22.

(1e) **Norme de g** Nous devons voir que $\|g\|_{\mathcal{L}(Y,\mathbb{K})} = \|\ell\|_{\mathcal{L}(M,\mathbb{K})}$. L'inégalité dans un sens est facile pour qui comprend la norme opérateur[87]. Étant donné que $M \subset Y$ nous avons

$$\|g\|_{\mathcal{L}(Y,\mathbb{K})} = \sup_{y\in Y} \frac{|g(y)|}{\|y\|} \geqslant \sup_{y\in M} \frac{|g(y)|}{\|y\|} = \sup_{y\in M} \frac{|\ell(y)|}{\|y\|} = \|\ell\|_{\mathcal{L}(M,\mathbb{K})}. \tag{27.718}$$

L'inégalité dans l'autre sens n'est pas trop compliquée non plus. Prenons $x \in Y$ vérifiant $\|x\| = 1$. Nous considérons $(N, f) \in \mathcal{E}$ tel que $x \in N$. Alors

$$|g(x)| = |f(x)| \leqslant \|f\| = \|\ell\|. \tag{27.719}$$

Donc $\|g\| \leqslant \|\ell\|$.

(1f) **Conclusion pour le moment** Nous avons prouvé que (Y, g) est un majorant de $\mathcal{F}$. Donc $(\mathcal{E}, \leqslant)$ est un ensemble inductif.

(2) **Lemme de Zorn** Vu que $(\mathcal{E}, \leqslant)$ est un ensemble inductif non vide, il possède un élément maximal par lemme de Zorn 1.23. Nous nommons (Y, f) un tel élément maximal.

(3) **Fin de la preuve** Nous devons prouver que $Y = X$, de telle sorte que $f \in \mathcal{L}(Y, \mathbb{K})$ soit définie sur X. Supposons que $Y \neq X$. Dans ce cas nous considérons $x_1 \in X\backslash Y$. Suivant que $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$, nous utilisons le lemme 27.171 ou 27.172 pour construire la paire

$$\big(\mathrm{Span}_{\mathbb{K}}(Y, x_1), f_1\big) \tag{27.720}$$

qui majore (Y, f). Contradiction avec la maximalité de (Y, f). Donc $Y = X$.

□

PROPooFJPXooWrjbuH

Proposition 27.174 ([1]).
Soit un espace de Banach E sur $\mathbb{K}(= \mathbb{R}$ ou $\mathbb{C})$. Soit $a \in E$. Nous avons

$$\|a\| = \max_{\substack{\varphi\in E'\\ \|\varphi\|=1}} |\varphi(a)|. \tag{27.721}$$

Démonstration. Soit $\varphi \in E'$ tel que $\|\varphi\| = 1$. Si $|\varphi(a)| > \|a\|$ alors

$$\|\varphi\| = \sup_{x\in E} \frac{|\varphi(x)|}{\|x\|} \geqslant \frac{|\varphi(a)|}{\|a\|} > 1 \tag{27.722}$$

Nous en déduisons que pour tout $\varphi \in B_{E'}(0, 1)$, $|\varphi(a)| \leqslant \|a\|$.

Maintenant nous construisons une application linéaire $\varphi \in E'$ telle que $|\varphi(a)| = \|a\|$ et $\|\varphi\| = 1$. Pour cela nous considérons le sous-espace $M = \mathrm{Span}\{a\}$ et l'application

$$\begin{aligned} f\colon M &\to \mathbb{K} \\ \lambda a &\mapsto \lambda\|a\|. \end{aligned} \tag{27.723}$$

Cette application vérifie $f(a) = \|a\|$ et $\|f\| = 1$ parce que

$$\|f\| = \sup_{x\in M} \frac{|f(x)|}{\|x\|} = \sup_{\lambda\in\mathbb{K}} \frac{|f(\lambda a)|}{\|\lambda a\|} = \frac{|\lambda|\|a\|}{|\lambda|\|a\|} = 1. \tag{27.724}$$

Le théorème de Hahn-Banach 27.173 prolonge f en un élément $\varphi \in E'$ vérifiant $\|\varphi\| = \|f\| = 1$. Cette fonctionnelle vérifie donc aussi $\varphi(a) = f(a) = \|a\|$. □

CORooOBDHooJpiBrs

Corolaire 27.175 ([1]).
Voici trois façons différentes de dire la même chose, par ordre décroissant de frime.

(1) Le dual d'un espace vectoriel normé sépare les points.

87. Si ce n'est pas votre cas, vous ne devriez franchement pas être en train de lire ces lignes. Ce n'est que mon avis ; après tout, vous faites comme vous le sentez.

(2) *Les fonctionnelles bornées d'un espace vectoriel normé séparent les points.*

(3) *Si X est un espace vectoriel normé sur $\mathbb{K}$ ($=\mathbb{R}$ ou $\mathbb{C}$), et si $x,y\in X$, alors il existe une application $f\in\mathcal{L}(X,\mathbb{K})$ telle que $f(x)\neq f(y)$.*

Démonstration. Le théorème de la base incomplète 4.25 nous permet de considérer une base $\{e_i\}_{i\in I}$ de X telle que $e_0=x$ et $e_1=y$.

Évacuons quelque objections.

— Si x et y sont colinéaires, on complète seulement $\{x\}$ et ça ne changera rien pour la suite[88].

— En écrivant « e_0 » et « e_1 » ne prétends pas que I soit un ensemble de nombres. C'est juste une facilité d'écriture pour éviter de dire « il existe $\alpha,\beta\in I$ tels que $x=e_\alpha$ et $y=e_\beta$ ».

Nous considérons l'application linéaire

$$\begin{aligned}\ell\colon \operatorname{Span}_{\mathbb{K}}(e_0,e_1)&\to\mathbb{K}\\ \alpha e_1+\beta e_2&\mapsto\alpha.\end{aligned}\tag{27.725}$$

Cette application est parfaitement bornée. Le théorème de Hahn-Banach 27.173 nous permet de considérer une extension f sur X. Cette extension a la même norme et est donc bornée (donc continue par la proposition 10.70). C'est donc un élément du dual de X.

Comme $f=\ell$ sur $\operatorname{Span}(e_0,e_1)$, nous avons $f(x)\neq f(y)$. □

27.17 Théorème de Tietze

Définition 27.176.
*Si E et F sont des espaces normés, une application $f\colon E\to F$ est **presque surjective** si il existe $\alpha\in]0,1[$ et $C>0$ tels que pour tout $y\in\overline{B_F(0,1)}$, il existe $x\in\overline{B_E(0,C)}$ tel que $\|y-f(x)\|\leqslant\alpha$.*
LemBQLooRXhJzK

Lemme 27.177 ([104]).
Soient E et F des espaces de Banach et $f\in\mathcal{L}^0(E,F)$[89]. Si f est presque surjective, alors
ItemTSOooYkxvBui

(1) *f est surjective*
ItemTSOooYkxvBuii

(2) *pour tout $y\in\overline{B_F(0,1)}$, il existe $x\in\overline{B_E(0,\frac{C}{1-\alpha})}$ tel que $y=f(x)$.*

Le point (2) est une précision du point (1) : il dit quelle est la taille de la boule de E nécessaire à obtenir la boule unité dans F.

Démonstration. Soit $y\in\overline{B_F(0,1)}$. Nous allons construire $x\in B\big(0,\frac{C}{1-\alpha}\big)$ qui donne $f(x)=y$. Ce x sera la limite d'une série que nous allons construire par récurrence. Pour $n=1$ nous utilisons la presque surjectivité pour considérer $x_1\in\overline{B_E(0,C)}$ tel que $\|y-f(x_1)\|\leqslant\alpha$. Ensuite nous considérons la récurrence

$$x_n\in\overline{B_E(0,C)}\tag{27.726}$$

tel que

$$\Big\|y-\sum_{i=1}^{n}\alpha^{i-1}f(x_i)\Big\|\leqslant\alpha^n\tag{27.727}$$

Pour montrer que cela existe nous supposons que la série est déjà construite jusqu'à $n>1$:

$$\frac{1}{\alpha^n}\Big(y-\sum_{i=1}^{n}\alpha^{i-1}f(x_i)\Big)\in\overline{B_F(0,1)}\tag{27.728}$$

À partir de là, par presque surjectivité il existe un $x_{n+1}\in\overline{B_E(0,C)}$ tel que

$$\Big\|\frac{y-\sum_{i=1}^{n}\alpha^{i-1}f(x_i)}{\alpha^n}-f(x_{n+1})\Big\|\leqslant\alpha.\tag{27.729}$$

88. Soyez quand même attentives à ne pas vous laisser enfumer.
89. L'ensemble des applications linéaires continues

En multipliant par α^n, le terme $\alpha^n f(x_{n+1})$ s'intègre bien dans la somme :

$$\|y - \sum_{i=1}^{n+1} \alpha^{i-1} f(x_i)\| \leqslant \alpha^{n+1}. \tag{27.730}$$

Nous nous intéressons à une éventuelle limite à la somme des $\alpha^{n-1}x_n$. D'abord nous avons la majoration $\|\alpha^{n-1}x_n\| \leqslant \alpha^{n-1}C$, et vu que par la définition de la presque surjectivité $0 < \alpha < 1$, la série

$$\sum_{n=1}^{\infty} \alpha^{n-1}x_n \tag{27.731}$$

converge absolument [90] parce que la suite des normes est une suite géométrique de raison α. Vu que E est de Banach, la convergence absolue implique la convergence simple (la suite des sommes partielles est de Cauchy et Banach est complet). Nous posons

$$x = \sum_{n=1}^{\infty} \alpha^{n-1}x_n \in E, \tag{27.732}$$

et en termes de normes, ça vérifie

$$\|x\| \leqslant \sum_{n=1}^{\infty} \alpha^{n-1}\|x_n\| \leqslant C \sum_{n=1}^{\infty} \alpha^{n-1} = \frac{C}{1-\alpha}. \tag{27.733}$$

Donc c'est bon pour avoir $x \in B\big(0, \frac{C}{1-\alpha}\big)$. Nous devons encore vérifier que $y = f(x)$. Pour cela nous remarquons que

$$\|y - f\Big(\sum_{n=1}^{N} \alpha^{n-1}x_n\Big)\| \leqslant \alpha^N. \tag{27.734}$$

Nous pouvons prendre la limite $N \to \infty$ et permuter f avec la limite (par continuité de f). Vu que $0 < \alpha < 1$ nous avons

$$\|y - f(x)\| = 0. \tag{27.735}$$

□

ThoFFQooGvcLzJ

Théorème 27.178 (Tietze[104, 496, 698])**.**
Soit un espace métrique (X, d) et un fermé $Y \subset X$. Soit $g_0 \in C^0(Y, \mathbb{R})$. Alors g_0 admet un prolongement continu sur X.

Démonstration. Soit l'opération de restriction

$$\begin{aligned} T\colon (C_b^0(X, \mathbb{R}), \|.\|_\infty) &\to (C_b^0(Y, \mathbb{R}), \|.\|_\infty) \\ f &\mapsto f|_Y. \end{aligned} \tag{27.736}$$

L'application T est évidemment linéaire. Elle est de plus bornée pour la norme opérateur usuelle donnée par la proposition 10.55 parce que $\|T(f)\| \leqslant \|f\| < \infty$. L'application T est alors continue par la proposition 10.70.

(1) **Presque surjection**
Soit $g \in C_b^0(Y, \mathbb{R})$ avec $\|g\|_\infty \leqslant 1$. Nous posons

$$Y^+ = \{x \in Y \text{ tel que } \frac{1}{3} \leqslant g(x) \leqslant 1\} \tag{27.737a}$$

$$Y^- = \{x \in Y \text{ tel que } -1 \leqslant g(x) \leqslant -\frac{1}{3}\}. \tag{27.737b}$$

90. Définition 10.92.

Nous considérons alors

$$\begin{aligned} f\colon X &\to \mathbb{R} \\ x &\mapsto \frac{1}{3}\frac{d(x,Y^-)-d(x,Y^+)}{d(x,Y^-)+d(x,Y^+)} \end{aligned} \tag{27.738}$$

Vu qu'en valeur absolue le dénominateur est plus grand que le numérateur nous avons $\|f\|_\infty \leqslant \frac{1}{3}$. Notons que

— Si $x \in Y^+$ alors $f(x) = \frac{1}{3}$ et $g(x) \in [\frac{1}{3}, 1]$;

— Si $x \in Y^-$ alors $f(x) = -\frac{1}{3}$ et $g(x) \in [-1, -\frac{1}{3}]$;

— Si x n'est ni dans Y^+ ni dans Y^- alors nous avons[91] $g(x) \in [-\frac{1}{3}, \frac{1}{3}]$ et donc $\big|f(x) - g(x)\big| \leqslant \big|f(x)\big| + \big|g(x)\big| \leqslant \frac{2}{3}$.

Dans les trois cas nous avons $\big|f(x) - g(x)\big| \in [0, \frac{2}{3}]$ pour tout $x \in X$. Cela prouve que

$$\|T(f) - g\|_{Y,\infty} \leqslant \frac{2}{3}. \tag{27.739}$$

En résumé nous avons pris g dans la boule $\overline{B(0,1)}$ de $\big(C_b^0(Y,\mathbb{R}), \|.\|_\infty\big)$ et nous avons construit une fonction f dans la boule $\overline{B(0,\frac{1}{3})}$ de $\big(C_b^0(X,\mathbb{R}), \|.\|_\infty\big)$ telle que $\|T(f) - g\|_\infty \leqslant \frac{2}{3}$. L'application T est donc une presque surjection avec $\alpha = \frac{2}{3}$ et $C = \frac{1}{3}$.

(2) **Extension avec $\|.\|_\infty \leqslant 1$** SPITEMooSTXMooIdockk Nous montrons que si $g \in C^0(Y,\mathbb{R})$ vérifie $\|g\|_\infty \leqslant 1$, alors il existe une extension $f \in C^0(X,\mathbb{R})$ vérifiant $\|f\|_\infty \leqslant 1$.

La proposition 12.359 nous assure que les espaces $C_b^0(X,\mathbb{R})$ et $C_b^0(Y,\mathbb{R})$ sont de Banach (complets), et le lemme 27.177 nous dit alors que T est surjective et que pour tout $g \in \overline{B(0,1)}$, il existe

$$f \in \overline{B\left(0, \frac{1/3}{1-\frac{2}{3}}\right)} = \overline{B(0,1)}. \tag{27.740}$$

telle que $g = T(f)$.

Pour le redire avec des mots plus simples, nous avons prouvé que si $g\colon Y \to \mathbb{R}$ vérifie $\|g\|_\infty \leqslant 1$, il existe une extension $f\colon X \to \mathbb{R}$ telle que $\|f\|_\infty \leqslant 1$.

(3) **Extension avec $|.| < 1$** SPITEMooZBGTooKbyuch Nous montrons que si $g \in C^0(Y,\mathbb{R})$ vérifie $|g| < 1$, alors il existe une extension $f \in C^0(X,\mathbb{R})$ vérifiant $|f| < 1$.

Soit $g \in B_{C_b^0(Y)}(0,1)$ et son prolongement $h \in \overline{B_{C_b^0(X)}(0,1)}$ du point (2). Si $|h| < 1$ alors nous avons le résultat en posant $f = g$.

Si au contraire il existe $z \in X$ tel que $|h(z)| = 1$, nous posons

$$Z = \{x \in X \text{ tel que } |h(x)| = 1\}. \tag{27.741}$$ EQooIVXAooZQZxlq

Nous avons $Y \cap Z = \varnothing$ parce que nous avons $h = g$ sur Y et nous avons choisi $\|g\|_\infty < 1$. Par ailleurs Y est fermé par hypothèse et Z est fermé parce que h est continue ; par conséquent $Y \cap Z$ est fermé, donc[92]

$$\bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing. \tag{27.742}$$

Nous posons

$$\begin{aligned} u\colon X &\to \mathbb{R}^+ \\ x &\mapsto \frac{d(x,Z)}{d(x,Y)+d(x,Z)} \end{aligned} \tag{27.743}$$

Le dénominateur n'est pas nul parce qu'il faudrait $d(x,Y) = d(x,Z) = 0$, ce qui demanderait $x \in \bar{Y} \cap \bar{Z}$, ce qui n'est pas possible. Nous posons $f = uh$. Si $x \in Y$ alors $u(x) = 1$, donc f est encore un prolongement de g. De plus f est encore continue, et donc encore un bon

91. Nous rappelons que $\|g\| = 1$, donc $g(x)$ est forcément ente -1 et 1.

92. Si vous avez l'intention de dire que $\overline{Y \cap Z} = \bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing$, allez d'abord voir l'exemple 7.133. Ici c'est correct parce que Y et Z sont fermés.

candidat. Enfin si x est hors de Y alors $d(x,Y)>0$ (strictement parce que Y est fermé) et donc $0<u(x)<1$, ce qui donne $|f(x)|<|h(x)|\leqslant 1$.

Et voilà.

(4) **Démonstration du théorème**

Nous sommes maintenant prêts à prouver le théorème. Soit $g_0\in C^0(Y,\mathbb{R})$. Nous allons nous ramener au cas de la boule unité ouverte en utilisant un homéomorphisme[93] $\phi\colon\mathbb{R}\to]-1,1[$.

L'application $g=\phi\circ g_0$ vérifie $|g(y)|<1$ pour tout $y\in Y$. Par le point (3), il existe une extension $f\in C^0(X,\mathbb{R})$ vérifiant $|f(x)|<1$ pour tout $x\in X$. Attention : cette application f prolonge g et non g_0.

Nous posons finalement $f_0=\phi^{-1}\circ f$. Cela est encore une application dans $C^0(X,\mathbb{R})$, mais cette fois elle prolonge g_0. En effet si $y\in Y$, en tenant compte de $f(y)=g(y)$, nous avons

$$f_0(y)=(\phi^{-1}\circ f)(y)=\phi^{-1}\big(f(y)\big)=\phi^{-1}\big(g(y)\big)=\phi^{-1}\big((\phi\circ g_0)(y)\big)=g_0(y). \tag{27.744}$$

□

27.17.1 Pas de bicontinues entre dimensions différentes

THOooLGJMooIYzOBD

Théorème 27.179 ([699]).
Soit une application continue et injective $f\colon\overline{B(0,1)}\to\mathbb{R}^n$. Alors

$$f(0)\in\operatorname{Int}\Big(f\big(\overline{B(0,1)}\big)\Big). \tag{27.745}$$

Démonstration. Soit une application continue et injective $f_0\colon\overline{B(0,1)}\to\mathbb{R}^n$. Soit $\epsilon<0$. Si nous supposons que $f_0(0)$ n'est pas à l'intérieur de $f_0\big(\overline{B(0,1)}\big)$, alors il existe un $c\in\mathbb{R}^n$ hors de $f_0\big(\overline{B(0,1)}\big)$ et tel que $\|c-f_0(0)\|<\epsilon$.

Nous posons maintenant $f(x)=f_0(x)-c$. Nous avons

$$\begin{cases}\|f(0)\|<\epsilon & \text{(27.746a)}\\ 0\notin f\big(\overline{B(0,1)}\big). & \text{(27.746b)}\end{cases}$$

Vu que f est continue et que $\overline{B(0,1)}$ est compact, la partie $f\big(\overline{B(0,1)}\big)$ est compacte (théorème 7.220). L'application restreinte $f\colon\overline{B(0,1)}\to f_0\big(\overline{B(0,1)}\big)$ est une bijection continue entre compacts. Son inverse est alors également continue (lemme 7.222).

Mais comme $f\big(\overline{B(0,1)}\big)$ est un fermé de $\mathbb{R}^n$, le théorème de Tietze (27.178) prolonge son inverse en une application continue

$$G\colon\mathbb{R}^n\to\mathbb{R}^n. \tag{27.747}$$

Nous avons $G\big(f(0)\big)=0$. Nous choisissons ϵ assez petit pour que $\|G(y)\|<0.1$ dès que $\|y-f(0)\|\leqslant 2\epsilon$. Comme $\|f(0)\|<\epsilon$, nous avons aussi

$$\|y\|\leqslant\epsilon\Rightarrow\|y-f(0)\|\leqslant 2\epsilon\Rightarrow\|G(y)\|<0.1. \tag{27.748}$$

Autrement dit,

$$G\Big(\overline{B(0,\epsilon)}\Big)\subset B(0,\frac{1}{10}). \tag{27.749}$$

Nous posons

$$\Sigma_1=\{y\in f\big(\overline{B(0,1)}\big)\text{ tel que }\|y\|\geqslant\epsilon\}, \tag{27.750}$$

ainsi que

$$\Sigma_2=\{y\in\mathbb{R}^n\text{ tel que }\|y\|=\epsilon\}, \tag{27.751}$$

et $\Sigma=\Sigma_1\cup\Sigma_2$.

93. Existence dans le lemme 18.40.

(1) La partie Σ_2 est compacte en tant que fermée et bornée[94]. La partie Σ_1 est compacte en tant que intersection entre le compact $f\big(\overline{B(0,1)}\big)$ et le fermé $\|y\| \geqslant \epsilon$ (corolaire 7.97). Finalement la partie Σ est également compacte[95]

(2) La partie Σ ne contient pas $f(0)$ parce que $\|f(0) < \epsilon\|$ tandis que les éléments $y \in \Sigma$ vérifient $\|y\| \geqslant \epsilon$.

(3) La partie Σ_1 ne contient pas 0. En effet, vu que G étend f^{-1}, si $x \in G(\Sigma_1)$, nous avons $f(x) \in \Sigma_1$ et en particulier $\|f(x)\| \geqslant \epsilon$. Vu que $\|f(0)\| < \epsilon$, nous avons $x \neq 0$.

(4) La partie Σ_1 est compacte, G est continue dessus et ne prend pas la valeur 0. Il existe donc $\delta > 0$ tel que

$$G(\Sigma_1) > \delta. \tag{27.752}$$

Nous considérons un tel δ qui vérifie de plus $\delta < 1/10$.

Nous introduisons l'application

$$\begin{aligned} \phi\colon f\big(\overline{B(0,1)}\big) &\to \Sigma \\ y &\mapsto \max\big(\epsilon/\|y\|, 1\big)y. \end{aligned} \tag{27.753}$$

Prouvons que ϕ prend effectivement ses valeurs dans Σ. Il y a deux possibilités : soit $\|y\| < \epsilon$, soit $\|y\| \geqslant \epsilon$. Si $\|y\| < \epsilon$, alors $\phi(y) = \frac{\epsilon}{\|y\|}y$, ce qui donne $\|\phi(y)\| = \epsilon$, et donc $\phi(y) \in \Sigma_2$. Si $\|y\| \geqslant \epsilon$, alors le max est 1 et nous avons $\phi(y) = y$. Dans ce cas, nous avons d'une part $\|\phi(y)\| = \|y\| \geqslant \epsilon$ et d'autre part $\phi(y) = y \in f\big(\overline{B(0,1)}\big)$; donc $\phi(y) \in \Sigma_1$.

Le théorème de Stone-Weierstrass 12.421 nous permet de considérer un polynôme (à n variables) $P\colon \mathbb{R}^n \to \mathbb{R}^n$ vérifiant

$$\|P - G\|_\Sigma < \delta. \tag{27.754}$$

Un tel polynôme ne s'annule pas sur Σ_1 parce que $\|G\|_{\Sigma_1} > \delta$.

Le polynôme P est de classe C^∞ (proposition 12.353). Vu que Σ_2 est de mesure nulle, la proposition 14.279 dit que $P(\Sigma_2)$ est de mesure nulle dans $\mathbb{R}^n$.

À l'aide de la proposition 14.280, nous choisissons $s \in \mathbb{R}^n$ tel que $P_2(x) = P(x+s)$ ne s'annule ni sur Σ_1 ni sur Σ_2. Vu que P est uniformément continue sur Σ (théorème de Heine 12.78), il existe M (majorant de la norme de la dérivée sur Σ) tel que

$$\|P(x) - P_2(x)\| = \|P(x+s) - P(x)\|_\Sigma \leqslant sM. \tag{27.755}$$

En choisissant $s < \epsilon'/M$, nous avons

$$\|P - P_2\|_\Sigma < \epsilon'. \tag{27.756}$$

En prenant ϵ' assez petit[96], nous avons encore

$$\|P_2 - G\|_\Sigma < \delta. \tag{27.757}$$

EQooIVQLooViVzsA

Nous introduisons maintenant

$$\begin{aligned} \tilde{G}\colon f\big(\overline{B(0,1)}\big) &\to \mathbb{R}^n \\ \tilde{G} &= P_2 \circ \phi. \end{aligned} \tag{27.758}$$

Nous allons prouver que $\|G(y) - \tilde{G}(y)\| \leqslant 0.3$ pour tout $y \in f\big(\overline{B(0,1)}\big)$.

(1) **Si $\|y\| > \epsilon$**

Si $\|y\| > \epsilon$ et $y \in b\big(\overline{B(0,1)}\big)$, alors

$$\tilde{G}(y) = (P_2 \circ \phi)(y) = P\big(\min(\epsilon/\|y\|, 1)y\big). \tag{27.759}$$

Vu que $\|y\| > \epsilon$, le max est 1 et nous avons $\tilde{G}(y) = P_2(y)$, et donc, en utilisant (27.757),

$$\|G(y) - \tilde{G}(y)\| = \|G(y) - P_2(y)\| < \delta. \tag{27.760}$$

Quitte à choisir δ plus petit, nous pouvons supposer $\delta < 0.3$.

94. Borel-Lebesgue 9.24.
95. Union de compacts, lemme 7.98.
96. Plus petit que la différence entre δ et $\|P - G\|_\Sigma$.

(2) **Si $\|y\| \leqslant \epsilon$** Nous considérons maintenant le cas $\|y\| \leqslant \epsilon$ et $y \in b\big(\overline{B(0,1)}\big)$.
Montrons que $\|G(y) - \tilde{G}(y)\| < \delta$ pour tout $y \in \big(\overline{B(0,1)}\big)$ tel que $\|y\| > \epsilon$. Nous avons

$$\begin{aligned}\|G(y) - \tilde{G}(y)\| &= \|G(y) - P_2\big(\phi(y)\big)\| && (27.761a)\\ &\leqslant \|G(y) - G\big(\phi(y)\big)\| + \|G\big(\phi(y)\big) - P_2\big(\phi(y)\big)\| && (27.761b)\\ &\leqslant \|G(y)\| + \|G\big(\phi(y)\big)\| + \|(G - P_2)\big(\phi(y)\big)\|. && (27.761c)\end{aligned}$$

Vu que $\|y\| \leqslant \epsilon$, nous avons $\phi(y) = \frac{\epsilon}{\|y\|}y$, et donc $\|\phi(y)\| = \epsilon$. De ce fait, $\phi(y) \in \Sigma_2$ et nous avons $\|(G - P_2)\big(\phi(y)\big)\| \leqslant \delta$.
Si nous choisissons $\delta < 0.1$ nous avons donc encore

$$\|G(y) - \tilde{G}(y)\| \leqslant 2 \times 0.1 + \delta \leqslant 0.3. \tag{27.762}$$

Bref, pour tout $y \in f\big(\overline{B(0,1)}\big)$ nous avons

$$\|G(y) - \tilde{G}(y)\| \leqslant 0.3. \tag{27.763}$$ EQooMIVFooPwKeSR

Ça va être le moment de demander au théorème de Brouwer d'intervenir.

(1) **Définition de g**
Vu que $\tilde{G}$ est définie sur $f\big(\overline{B(0,1)}\big)$, nous pouvons considérer l'application

$$\begin{aligned}g\colon \overline{B(0,1)} &\to \mathbb{R}^n\\ x &\mapsto x - \tilde{g}\big(f(x)\big).\end{aligned} \tag{27.764}$$

Notons pour $x \in \overline{B(0,1)}$, nous avons $x = G\big(f(x)\big)$, et donc nous pouvons aussi écrire

$$g(x) = (G \circ f)(x) - (\tilde{G} \circ f)(x). \tag{27.765}$$ EQooJWNRooKIfex

(2) $g(x) \in \overline{B(0,1)}$
Nous avons déjà vu en (27.763) que $\|G - \tilde{G}\|_{f\left(\overline{B(0,1)}\right)} < 0.3 < 1$. L'équation (27.765) dit alors que $g(x) \in \overline{B(0,1)}$.

(3) **Brouwer** Nous pouvons appliquer le théorème de Brouwer (20.32) à la fonction g. Il existe $x_0 \in \overline{B(0,1))}$ tel que $g(x_0) = x_0$, c'est-à-dire $x_0 - \tilde{G}\big(f(x_0)\big) = x_0$. Nous en déduisons que

$$\tilde{G}\big(f(x_0)\big) = 0. \tag{27.766}$$

(4) **La contradiction** Rappel : $\tilde{G} = P_2 \circ \phi$. Si $y \in \overline{B(0,1)}$, alors $\phi(y) \in \Sigma$, et nous avons tout fait pour que P_2 ne s'annule pas sur Σ.
Et voilà la contradiction.

□

Remarque 27.180.
Il n'existe pas de bijection bicontinues d'un ouvert de $\mathbb{R}^m$ vers un ouvert de $\mathbb{R}^n$ si $m \neq n$. Il n'y a donc pas de notion de difféomorphismes entre ouverts de dimensions différentes.

Je crois que c'est une conséquence facile de 27.179. Écrivez-moi si vous pouvez rédiger une preuve.

27.18 Dualité, réflexivité et théorème de représentation de Riesz

Dans la suite E' est le dual topologique, c'est-à-dire l'espace des formes linéaires et continues sur E. Nous notons également V'' le dual de $(V', \|.\|)$. Certes en tant qu'ensembles, $(V', *)$ et $(V', \|.\|)$ sont identiques, mais comme ils n'ont pas la même topologie, les duaux ne sont pas les mêmes.

Bref, V'' est l'ensemble des applications linéaires continues $(V', \|.\|) \to \mathbb{C}$. Et lorsque nous disons $\mathbb{C}$ ici, ça peut aussi bien être $\mathbb{R}$ selon le contexte.

De plus nous considérons que V'' la norme opérateur qui dérive de la norme de V', laquelle dérive de la norme vectorielle sur V.

PROPooMAQSooCGFBBM

Proposition-Définition 27.181 (Espace réflexif)**.**
Soit un espace vectoriel normé V sur $\mathbb{R}$ ou $\mathbb{C}$. Nous considérons l'application

$$\begin{aligned} J &: V \to V'' \\ J(x)\varphi &= \varphi(x). \end{aligned} \tag{27.767}$$

ITEMooNVVSooNFXgnE

(1) L'application J est bien définie : $J(x) : V' \to \mathbb{C}$ est continue.

ITEMooKURHooZZWpbu

(2) L'application J est continue.

ITEMooTFYVooKhMOjp

(3) Elle est injective.

*Lorsque J est bijective, l'espace V est dit **réflexif**.*

Démonstration. Point par point.

(1) **(1)** Nous commençons par montrer que $J(x) : (V', \|.\|) \to \mathbb{C}$ est continue pour chaque $x \in V$. Soit une suite $\varphi_k \xrightarrow{\|.\|} 0$. Nous avons :

$$J(x)\varphi_k = \varphi_k(x) \leqslant \|\varphi_k\|\|x\| \to 0 \tag{27.768}$$

où vous aurez noté l'utilisation du lemme 10.66. Cela prouve que $J(x)$ est continue et donc que J est bien à valeurs dans V''.

(2) **(2)**
Soit une suite $x_k \xrightarrow{V} 0$, et étudions $\|J(x_k)\|$ pour la norme dans V''. Nous posons $x'_k = x_k/\|x_k\|$ et nous calculons (encore une fois, nous écrivons « $\mathbb{C}$ », mais ça pourrait être $\mathbb{R}$)

$$\|J(x_k)\| = \sup_{\|\varphi\|=1} |J(x_k)\varphi|_{\mathbb{C}} = \sup_{\|\varphi\|=1} |\varphi(x_k)| = \|x_k\| \sup_{\|\varphi\|=1} |\varphi(x'_k)| \leqslant \|x_k\| \to 0. \tag{27.769}$$

La dernière inégalité pourrait être sans doute une égalité [97], mais nous n'en avons pas besoin ici.

(3) **(3)** Soient $x \neq y$ dans V. Le corolaire 27.175 nous permet de considérer un élément $\varphi \in V'$ tel que $\varphi(x) \neq \varphi(y)$. Nous avons alors

$$J(x)\varphi \neq J(y)\varphi, \tag{27.770}$$

et donc $J(x) \neq J(y)$. Cela prouve que J est injective.

□

LEMooWEMFooEHIaxY

Lemme 27.182 ([1])**.**
Soient des espaces de Banach V et W. Si $\alpha : V \to W$ est une bijection linéaire et isométrique, alors l'application définie par

$$\begin{aligned} \beta &: V' \to W' \\ \beta(\varphi)y &= \varphi\big(\alpha^{-1}(y)\big) \end{aligned} \tag{27.771}$$

EQooTSVHooEQCuNw

pour tout $\varphi \in V'$ et $y \in W$ est une bijection linéaire isométrique.

97. Écrivez-moi si vous en êtes certain.

Démonstration. En plusieurs points.

(1) **α^{-1} est continue** Le théorème d'isomorphisme de Banach 27.1 implique que α^{-1} est continue parce que α est bijective, continue et linéaire entre deux espaces de Banach.

(2) **β prend ses valeurs dans W'** Nous devons prouver que, si $\varphi \in V'$, alors $\beta(\varphi)$ est bien un élément de W'. Autrement dit, nous devons prouver que $\beta(\varphi)\colon W \to \mathbb{C}$ est continue. Considérons une suite $y_k \xrightarrow{W} 0$. Alors, vu que α est continue, nous avons aussi $\alpha^{-1}(y_k) \xrightarrow{V} 0$, et donc

$$\varphi\big(\alpha^{-1}(y_k)\big) \xrightarrow{\mathbb{C}} 0 \tag{27.772}$$

parce que φ est continue. Donc $\beta(\varphi)$ est continue.

(3) **β est linéaire** Parce que α l'est.

(4) **β est isométrique** Nous devons prouver que la norme de l'application linéaire $\beta(\varphi)\colon W \to \mathbb{C}$ est la même que celle de φ. Vu que $\alpha\colon V \to W$ est une bijection isométrique, nous avons

$$\{\alpha^{-1}(y) \text{ tel que } \|y\| = 1, y \in W\} = \{x \in V \text{ tel que } \|x\| = 1\}. \tag{27.773}$$

Donc nous pouvons faire le calcul suivant :

$$\|\beta(\varphi)\| = \sup_{\substack{y\in W\\ \|y\|=1}} \|\beta(\varphi)y\| = \sup_{\substack{y\in W\\ \|y\|=1}} |\varphi\big(\alpha^{-1}(y)\big)| = \sup_{\substack{x\in V\\ \|x\|=1}} |\varphi(x)| = \|\varphi\|. \tag{27.774}$$

□

La proposition suivante dit que la notion d'être réflexif passe aux isomorphismes d'espaces vectoriels normés.

PROPooVRQKooLdmajh

Proposition 27.183 ([1]).
Soient des espaces vectoriels normés V et W ainsi qu'une bijection linéaire isométrique $\alpha\colon V \to W$. L'espace W est réflexif[98] *si et seulement si V l'est.*

Démonstration. Nous allons prouver que si W est réflexif, alors V est réflexif. Pour l'implication inverse, il suffira de noter que $\alpha^{-1}\colon W \to V$ est une bijection linéaire isométrique.

En plusieurs points.

(1) **Quelques applications** Nous notons J_V l'application de la proposition 27.181 :

$$\begin{aligned} J_V &\colon V \to V'' \\ J_V(v)\varphi &= \varphi(v), \end{aligned} \tag{27.775}$$

et J_W l'application correspondante pour W. Notre hypothèse est que J_W est bijective.

Nous posons

$$\begin{aligned} \beta &\colon V' \to W' \\ \beta(\varphi)w &= \varphi\big(\alpha^{-1}(w)\big), \end{aligned} \tag{27.776}$$

et

$$\begin{aligned} \gamma &\colon V'' \to W'' \\ \gamma(\sigma)\varphi &= \sigma\big(\beta^{-1}(\varphi)\big). \end{aligned} \tag{27.777}$$

L'application β est une bijection linéaire isométrique par le lemme 27.182 ; l'application γ l'est aussi, par le même lemme appliqué à $\beta\colon V' \to W'$.

(2) **Le S** Nous posons

$$\begin{aligned} S &\colon V \to V'' \\ S(v)\varphi &= J_W\big(\alpha(v)\big)\beta(\varphi). \end{aligned} \tag{27.778}$$

Nous allons montrer que $S = J_V$ et que S est bijectif.

98. Définition 27.181.

(3) $\underline{S = J_V}$ Cela est un calcul :

$$\begin{aligned} J_W\big(\alpha(v)\big)\beta(\varphi) &= \beta(\varphi)\big(\alpha(v)\big) && \text{(27.779a)}\\ &= \varphi\Big(\alpha^{-1}\big(\alpha(v)\big)\Big) && \text{(27.779b)}\\ &= \varphi(v) && \text{(27.779c)}\\ &= J_V(v)\varphi. && \text{(27.779d)} \end{aligned}$$

(4) $\underline{S \text{ est injective}}$ L'application J_V est toujours injective par la proposition 27.181.

(5) $\underline{S \text{ est surjective}}$ Soi $\sigma \in V''$. Nous devons prouver l'existence de $v \in V$ tel que $S(v) = \sigma$. Vu que $J_W : W \to W''$ est surjective et que $\gamma(\sigma) \in W''$, il existe $w \in W$ tel que

$$J_W(w) = \gamma(\sigma). \tag{27.780}$$

Un simple calcul montre alors que $S\big(\alpha^{-1}(w)\big) = \sigma$:

$$\begin{aligned} S\big(\alpha^{-1}(w)\big)\varphi &= J_W(w)\beta(\varphi) && \text{(27.781a)}\\ &= \gamma(\sigma)\big(\beta(\varphi)\big) && \text{(27.781b)}\\ &= \sigma\Big(\beta^{-1}\big(\beta(\varphi)\big)\Big) && \text{(27.781c)}\\ &= \sigma(\varphi). && \text{(27.781d)} \end{aligned}$$

Nous avons donc bien que $S(v) = \sigma$ en posant $v = \alpha^{-1}(w)$.

(6) **Conclusion** L'application $S = J_V$ est surjective ; l'espace V est donc réflexif par la proposition 27.181.

□

Voici déjà un bel énoncé. Pour des espaces mesurés $(\Omega, \mathcal{A}, \mu)$ plus généraux, voir l'arme totale en le théorème 27.186.

PropOAVooYZSodR

Proposition 27.184 ([380], thème 21)**.**
Soit $1 < p < 2$ et q tel que $\frac{1}{p} + \frac{1}{q} = 1$. L'application

$$\begin{aligned} \Phi &: L^q\big([0,1]\big) \to L^p\big([0,1]\big)'\\ \Phi_g(f) &= \int_{[0,1]} f\bar{g}. \end{aligned} \tag{27.782}$$

est une isométrie linéaire surjective.

Démonstration. Pour la simplicité des notations nous allons noter L^2 pour $L^2\big([0,1]\big)$, et pareillement pour L^p.

(1) $\underline{\Phi_g \text{ est un élément de } (L^p)'}$
Si $f \in L^p$ et $g \in L^q$ nous devons prouver que $\Phi_g(f)$ est bien définie. Pour cela nous utilisons l'inégalité de Hölder [99] qui dit que $fg \in L^1$; par conséquent la fonction $f\bar{g}$ est également dans L^1 et nous avons

$$|\Phi_g(f)| \leqslant \int_{[0,1]} |f\bar{g}| = \|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{27.783}$$

En ce qui concerne la norme de l'application Φ_g nous avons tout de suite

$$\|\Phi_g\| = \sup_{\|f\|_p = 1} \big|\Phi_g(f)\big| \leqslant \|g\|_q. \tag{27.784}$$

Cela signifie que l'application Φ_g est bornée et donc continue par la proposition 10.70. Nous avons donc bien $\Phi_g \in (L^p)'$.

99. Proposition 27.33.

(2) **Isométrie**

Afin de prouver que $\|\Phi_g\| = \|g\|_q$ nous allons trouver une fonction $f \in L^p$ telle que $\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q$. De cette façon nous aurons prouvé que $|\Phi_g| \geqslant \|g\|_q$, ce qui conclurait que $|\Phi_g| = \|g\|_q$. Nous posons $f = g|g|^{q-2}$, de telle sorte que $|f| = |g|^{q-1}$ et

$$\|f\|_p = \left(\int |g|^{p(q-1)} \right)^{1/p} = \left(\int |g|^q \right)^{1/p} = \|g\|_q^{q/p} \tag{27.785}$$

où nous avons utilisé le fait que $p(q-1) = q$. La fonction f est donc bien dans L^p. D'autre part,

$$\Phi_g(f) = \int f\bar{g} = \int g|g|^{q-2}\bar{g} = \int |g|^q = \|g\|_q^q. \tag{27.786}$$

Donc

$$\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q^{q-\frac{q}{p}} = \|g\|_q \tag{27.787}$$

où nous avons encore utilisé le fait que $q - \frac{q}{p} = \frac{q(p-1)}{p} = 1$.

(3) **Surjectif**

Soit $\ell \in (L^p)'$; c'est une application $\ell\colon L^p \to \mathbb{C}$ dont nous pouvons prendre la restriction à L^2 parce que la proposition 27.42 nous indique que $L^2 \subset L^p$. Nous nommons $\phi\colon L^2 \to \mathbb{C}$ cette restriction.

(3a) $\phi \in (L^2)'$

Nous devons montrer que ϕ est continue pour la norme sur L^2. Pour cela nous montrons que sa norme opérateur (subordonnée à la norme de L^2 et non de L^p) est finie :

$$\sup_{f \in L^2} \frac{|\phi(f)|}{\|f\|_2} \leqslant \sup_{f \in L^2} \frac{|\ell(f)|}{\|f\|_p} < \infty. \tag{27.788}$$

Nous avons utilisé l'inégalité de norme $\|f\|_p \leqslant \|f\|_2$ de la proposition 27.42(2).

(3b) **Utilisation du dual de L^2**

Étant donné que L^2 est un espace de Hilbert (lemme 27.92) et que $\phi \in (L^2)'$, le théorème 25.18 nous donne un élément $g \in L^2$ tel que $\phi(f) = \Phi_g(f)$ pour tout $f \in L^2$.

Nous devons prouver que $g \in L^q$ et que pour tout $f \in L^p$ nous avons $\ell(f) = \Phi_g(f)$.

(3c) $g \in L^q$

Nous posons $f_n = g|g|^{q-2}\mathbb{1}_{|g|<n}$. Nous avons d'une part

$$\Phi_g(f_n) = \int_0^1 f_n\bar{g} = \int_{|g|<n} |g|^q. \tag{27.789}$$

Et d'autre part comme $f_n \in L^2$ nous avons aussi $\phi(f_n) = \Phi_g(f_n)$ et donc

$$\begin{aligned} 0 \leqslant \Phi(f_n) = \phi(f_n) &\leqslant \|\ell\|\|f_n\|_p && (27.790a)\\ &= \|\ell\| \left(\int_{|g|<n} |g|^{(q-1)p} \right)^{1/p} && (27.790b)\\ &= \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}. && (27.790c) \end{aligned}$$

où nous avons à nouveau tenu compte du fait que $p(q-1) = q$. En combinant avec (27.789) nous trouvons

$$\int_{|g|<n} |g|^q \leqslant \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}, \tag{27.791}$$

et donc

$$\left(\int_{|g|<n} |g|^q\right)^{1-\frac{1}{p}} \leqslant \|\ell\|, \tag{27.792}$$

c'est-à-dire

$$\Big(\int_{|g|<n} |g|^q\Big)^{1/q} \leqslant \|\ell\|. \tag{27.793}$$

Si ce n'était pas encore fait nous nous fixons un représentant de la classe g (qui est dans L^2), et nous nommons également g ce représentant. Nous posons alors

$$g_n = |g|^q \mathbb{1}_{|g|<n} \tag{27.794}$$

qui est une suite croissante de fonctions convergeant ponctuellement vers $|g|^q$. Le théorème de Beppo-Levi 14.178 nous permet alors d'écrire

$$\lim_{n\to\infty} \int_{|q|<n} |g|^q = \int_0^1 |g|^q. \tag{27.795}$$

Mais comme pour chaque n nous avons $\int_{|g|<n} |g|^q \leqslant \|\ell\|^q$, nous conservons l'inégalité à la limite et

$$\int_0^1 |g|^q \leqslant \|\ell\|^q. \tag{27.796}$$

Cela prouve que $g \in L^p$.

(3d) $\ell(f) = \Phi_g(f)$

Soit $f \in L^p$. En vertu de la densité de L^2 dans L^p prouvée dans le corolaire 27.52 nous pouvons considérer une suite (f_n) dans L^2 telle que $f_n \xrightarrow{L^p} f$. Pour tout n nous avons

$$\ell(f_n) = \Phi_g(f_n). \tag{27.797}$$

Mais ℓ et Φ_g étant continues sur L^p nous pouvons prendre la limite et obtenir

$$\ell(f) = \Phi_g(f). \tag{27.798}$$

□

LemHNPEooHtMOGY

Lemme 27.185 ([11]).
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré fini. Soit $g \in L^1(\Omega)$ et S fermé dans $\mathbb{C}$. Si pour tout $E \in \mathcal{A}$ nous avons

$$\frac{1}{\mu(E)} \int_E g d\mu \in S, \tag{27.799}$$

alors $g(x) \in S$ pour presque tout $x \in \Omega$.

Démonstration. Soit $D = \overline{B(a,r)}$ un disque fermé dans le complémentaire de S (ce dernier étant fermé, le complémentaire est ouvert). Posons $E = g^{-1}(D)$. Prouvons que $\mu(E) = 0$ parce que cela prouverait que $g(x) \in D$ pour seulement un ensemble de mesure nulle. Mais S^c pouvant être écrit comme une union dénombrable de disques fermés [100], nous aurions $g(x) \in S^c$ presque nulle part.

Vu que $\frac{1}{\mu(E)} \int_E a = a$ nous avons

$$\Big|\frac{1}{\mu(E)} g d\mu - a\Big| = \Big|\frac{1}{\mu(E)} \int_E (g-a)\Big| \leqslant \frac{1}{\mu(E)} \int_E |g-a| \leqslant \frac{1}{\mu(E)} \mu(E) r = r. \tag{27.800a}$$

Donc

$$\frac{1}{\mu(E)} \int_E g d\mu \in D, \tag{27.801}$$

ce qui est une contradiction avec le fait que $D \subset S^c$. □

100. Tout ouvert peut être écrit comme union dénombrable d'éléments d'une base de topologie par la proposition 7.2 et $\mathbb{C}$ a une base dénombrable de topologie par la proposition 7.138.

Dans toute la partie d'analyse fonctionnelle, sauf mention du contraire, nous considérons dans L^p des fonctions à valeurs complexes, et donc les éléments du dual sont des applications linéaires continues à valeurs dans $\mathbb{C}$. La raison est que nous allons utiliser les résultats concernant L^p dans la partie sur les transformations de Fourier, tandis que les transformations de Fourier demandent naturellement de travailler sur les complexes.

ThoLPQPooPWBXuv

Théorème 27.186 (Théorème de représentation de Riesz, thème 21, [11, 466, 700, 451])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Soit q tel que $\frac{1}{p} + \frac{1}{q} = 1$ avec la convention que $q = \infty$ si $p = 1$. Alors l'application

$$\begin{aligned} \Phi\colon L^q &\to (L^p)' \\ \Phi_g(f) &= \int_\Omega f\bar{g}d\mu \end{aligned} \tag{27.802}$$

est une bijection isométrique[101] *dans les cas suivants :* ITEMooSQQBooWSFBmX

(1) si $1 < p < \infty$ et $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré quelconque, ITEMooCQGJooOWzjoV

(2) si $p = 1$ et $(\Omega, \mathcal{A}, \mu)$ est σ-fini.

Démonstration. Par petits bouts.

(1) **Φ est injective** Nous commençons par prouver que Φ est injectif. Soient $g, g' \in L^q$ tels que $\Phi_g = \Phi_{g'}$. Alors pour tout $f \in L^p$ nous avons

$$\int_\Omega f(\bar{g} - \bar{g}')d\mu = 0. \tag{27.803}$$

Soient des parties A_i de mesures finies telles que $\Omega = \bigcup_{i=1}^{\infty} A_i$. Étant donné que $\mu(A_i)$ est fini, nous avons $\mathbb{1}_{A_i} \in L^p(\Omega)$ et donc

$$\int_{A_i} (\bar{g} - \bar{g}')d\mu = \int_\Omega \mathbb{1}_{A_i}(x)(\bar{g} - \bar{g}')(x)d\mu(x) = 0. \tag{27.804}$$

La proposition 27.21 nous dit alors que $g - g' = 0$ dans $L^q(A_i)$. Pour chaque i, la partie $N_i = \{x \in A_i \text{ tel que } (g - g')(x) \neq 0\}$ est de mesure nulle.

Vu que Ω est l'union de tous les A_i, la partie de Ω sur laquelle $g - g'$ est non nulle est l'union des N_i et donc de mesure nulle parce que une réunion dénombrable de parties de mesure nulle est de mesure nulle. Donc $g - g' = 0$ presque partout dans Ω, ce qui signifie $g - g' = 0$ dans $L^q(\Omega)$.

(2) **La suite**

La partie difficile est de montrer que Φ est surjective.

Soit $\phi \in L^p(\Omega)'$. Si $\phi = 0$, c'est bien dans l'image de Φ ; nous supposons donc que non. Nous allons commencer par prouver qu'il existe une (classe de) fonction $g \in L^1(\Omega)$ telle que $\Phi_g(f) = \phi(f)$ pour tout $f \in L^\infty(\Omega, \mu)$; nous montrerons ensuite que $g \in L^q$ et que le tout est une isométrie.

(3) **Une mesure complexe**

Si $E \in \mathcal{A}$ nous notons $\nu(E) = \phi(\mathbb{1}_E)$. Nous prouvons maintenant que ν est une mesure complexe[102] sur $(\Omega, \mathcal{A})$. Le seul point pas facile est de démontrer est l'additivité dénombrable. Il est déjà facile de voir que A et B sont disjoints, $\nu(A \cup B) = \nu(A) + \nu(B)$. Soient ensuite des ensembles A_n deux à deux disjoints et posons $E_k = \bigcup_{i \leqslant k} A_i$ pour avoir $\bigcup_k A_k = \bigcup_k E_k$ avec l'avantage que les E_k soient emboîtés. Cela donne

$$\|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = \mu(E \backslash E_k)^{1/p}, \tag{27.805}$$

101. Pour rappel, la norme sur le dual est la norme opérateur 10.55.
102. Définition 14.239.

mais vu que $1 \leqslant p < \infty$, avoir $x_k \to 0$ implique d'avoir $x_k^{1/p} \to 0$. Prouvons que $\mu(E\backslash E_k) \to 0$. En vertu du lemme 14.23 nous avons pour chaque k :

$$\mu(E\backslash E_k) = \mu(E) - \mu(E_k), \tag{27.806}$$

et vu que $E_k \to E$ est une suite croissante, le lemme 14.24(1), sachant que μ est une mesure « normale », donne

$$\lim_{n\to\infty} \mu(E_k) = \mu\big(\bigcup_k E_k\big). \tag{27.807}$$

Donc effectivement $\mu(E_k) \to \mu(E)$ et donc oui : $\mu(E\backslash E_k) \to 0$. Jusqu'à présent nous avons

$$\lim_{k\to\infty} \|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = 0, \tag{27.808}$$

c'est-à-dire $\mathbb{1}_{E_k} \xrightarrow{L^p} \mathbb{1}_E$. La continuité de ϕ sur L^p donne alors

$$\lim_{k\to\infty} \nu(E_k) = \lim_{k\to\infty} \phi(\mathbb{1}_{E_k}) = \phi(\lim_{k\to\infty} \mathbb{1}_{E_k}) = \phi(\mathbb{1}_E) = \nu(E). \tag{27.809}$$

Par additivité finie de ν nous avons

$$\nu(E_k) = \sum_{i\leqslant k} \nu(A_i) \tag{27.810}$$

et en passant à la limite, $\sum_{i=1}^{\infty} \nu(A_i) = \nu(\bigcup_i A_i)$. L'application ν est donc une mesure complexe.

(4) **Mesure absolument continue**
En prime, si $\mu(E) = 0$ alors $\nu(E) = 0$ parce que

$$\mu(E) = 0 \Rightarrow \|\mathbb{1}_E\|_p = 0 \Rightarrow \mathbb{1}_E = 0 \text{ (dans } L^p) \Rightarrow \phi(\mathbb{1}_E) = 0 \tag{27.811}$$

(5) **Utilisation de Radon-Nikodym**
Nous sommes donc dans un cas où $\nu \ll \mu$ et nous utilisons le théorème de Radon-Nikodym 14.240 sous la forme de la remarque 14.241 : il existe une fonction intégrable $g\colon \Omega \to \mathbb{C}$ [103] telle que pour tout $A \in \mathcal{A}$,

$$\nu(A) = \int_A \bar{g} d\mu. \tag{27.812}$$

C'est-à-dire que

$$\phi(\mathbb{1}_A) = \int_A \bar{g} d\mu = \int_\Omega \bar{g}\mathbb{1}_A d\mu. \tag{27.813}$$

Nous avons donc exprimé ϕ comme une intégrale pour les fonctions caractéristiques d'ensembles.

(6) **Pour les fonctions étagées**
Par linéarité si f est mesurable et étagée nous avons aussi

$$\phi(f) = \int f\bar{g} d\mu = \Phi_g(f). \tag{27.814}$$

(7) **Pour $f \in L^\infty(\Omega)$**
Une fonction $f \in L^\infty$ est une fonction presque partout bornée. Nous supposons que f est presque partout bornée par M. Par ailleurs cette f est limite uniforme de fonctions étagées : $\|f_k - f\|_\infty \to 0$ en posant $f_k = f\mathbb{1}_{|f|\leqslant k}$. Pour chaque k nous avons l'égalité

$$\Phi_g(f_k) = \phi(f_k). \tag{27.815}$$

EqPDCJooGNjuAQ

103. On peut écrire, pour utiliser de la notation compacte que $g \in L^1(\Omega, \mathbb{C})$.

Par ailleurs la fonction $f_k\bar{g}$ est majorée par la fonction intégrable $M\bar{g}$ et le théorème de la convergence dominée 14.207 nous donne

$$\lim_{k\to\infty}\Phi_g(f_k)=\lim_{k\to\infty}\int f_k\bar{g}=\int f\bar{g}=\Phi_g(f). \tag{27.816}$$

Et la continuité de ϕ sur L^p couplée à la convergence $f_k\xrightarrow{L^p}f$ donne $\lim_{k\to\infty}\phi(f_k)=\phi(f)$. Bref prendre la limite dans (27.815) donne

$$\Phi_g(f)=\phi(f) \tag{27.817}$$

pour tout $f\in L^\infty(\Omega)$.

(8) **La suite ...**

Voici les prochaines étapes.

— Nous avons $\int f\bar{g}=\phi(f)$ tant que $f\in L^\infty$. Nous allons étendre cette formule à $f\in L^p$ par densité. Cela terminera de prouver que notre application est une bijection.

— Ensuite nous allons prouver que $\|\phi\|=\|\Phi_g\|$, c'est-à-dire que la bijection est une isométrie.

(9) **De L^∞ à L^p**

Soit $f\in L^p$. Si nous avions une suite (f_n) dans L^∞ telle que $f_n\xrightarrow{L^p}f$ alors $\lim\phi(f_n)=\phi(f)$ par continuité de ϕ. La difficulté est de trouver une telle suite de façon à pouvoir permuter l'intégrale et la limite :

$$\lim_{n\to\infty}\int_\Omega f_n\bar{g}=\int_\Omega\lim_{n\to\infty}f_n\bar{g}=\int_\Omega f\bar{g}=\Phi_g(f). \tag{27.818}$$

EqLYYAooUQnbfV

Nous allons donc maintenant nous atteler à la tâche de trouver $f_n\in L^\infty$ avec $f_n\xrightarrow{L^p}f$ et telle que (27.818) soit valide.

Nous allons d'abord supposer que $f\in L^p$ est positive à valeurs réelles. Nous avons alors par le théorème 14.118 qu'il existe une suite croissante de fonction étagées (et donc L^∞) telles que $f_n\to f$ ponctuellement. De plus étant donné que $|f_n|\leqslant|f|$, la proposition 27.20 nous dit que $f_n\xrightarrow{L^p}f$. Pour chaque n nous avons

$$\int_\Omega f_n\bar{g}=\phi(f_n). \tag{27.819}$$

Soit g^+ la partie réelle positive de $\bar{g}$. Alors nous avons la limite croissante ponctuelle $f_ng^+\to fg^+$ et le théorème de la convergence monotone 14.178 nous permet d'écrire

$$\lim_{n\to\infty}\int f_ng^+=\int fg^+. \tag{27.820}$$

Faisant cela pour les trois autres parties de $\bar{g}$ nous avons prouvé que si $f\in L^p$ est réelle et positive,

$$\int f\bar{g}=\phi(f), \tag{27.821}$$

c'est-à-dire que $\Phi_g(f)=\phi(f)$.

Refaisant le tout pour les trois autres parties de f nous montrons que

$$\Phi_g(f)=\phi(f) \tag{27.822}$$

pour tout $f\in L^p(\Omega)$. Nous avons donc égalité de ϕ et Φ_g dans $(L^p)'$ et donc bijection entre $(L^p)'$ et L^q.

(10) **Isométrie : mise en place**

Nous devons prouver que cette bijection est isométrique. Soit $\phi \in (L^p)'$ et $g \in L^q$ telle que $\Phi_g = \phi$. Il faut prouver que

$$\|g\|_q = \|\phi\|_{(L^p)'}. \tag{27.823}$$

(11) $\|\phi\| \leqslant \|g\|_q$

Nous savons que $\phi(f) = \int f\bar{g}$, et nous allons écrire la définition de la norme dans $(L^p)'$:

$$\|\phi\|_{(L^p)'} = \sup_{\|f\|_p=1} \big|\phi(f)\big| \tag{27.824a}$$

$$= \sup \Big| \int f\bar{g} \Big| \tag{27.824b}$$

$$\leqslant \sup \underbrace{\int |f\bar{g}|}_{=\|f\bar{g}\|_1}. \tag{27.824c}$$

Il s'agit maintenant d'utiliser l'inégalité de Hölder 27.33 :

$$\|\phi\| \leqslant \sup_{\|f\|_p=1} \|f\|_p \|\bar{g}\|_q = \|g\|_q. \tag{27.825}$$

L'inégalité dans l'autre sens sera démontrée en séparant les cas $p = 1$ et $1 < p < \infty$.

(12) **Si $p = 1$, une formule** Si E est un ensemble mesurable de mesure finie, alors

$$\Big| \int_E g d\mu \Big| = \big|\phi(\mathbb{1}_E)\big|. \tag{27.826}$$

Mais le fait que $\mu(E) < \infty$ donne que $\mathbb{1}_E \in L^1(\Omega)$. Donc $\mathbb{1}_E \in L^\infty \cap L^1$; nous pouvons alors écrire $\phi(\mathbb{1}_E) = \int_\Omega \mathbb{1}_E \bar{g} d\mu$ et donc

$$\Big| \int_\Omega \mathbb{1}_E \bar{g} d\mu \Big| = \Big| \int_E \bar{g} d\mu \Big| = \big|\phi(\mathbb{1}_E)\big| \leqslant \|\phi\|_{(L^1)'} \|\mathbb{1}_E\|_1 = \|\phi\|\mu(E). \tag{27.827}$$

EqUPCTooJvoKKI

Nous écrivons cela dans l'autre sens :

$$\|\phi\| \geqslant \frac{1}{\mu(E)} \Big| \int_\Omega \mathbb{1}_E \bar{g} d\mu \Big| = \Big| \frac{1}{\mu(E)} \int_E \bar{g} d\mu \Big|. \tag{27.828}$$

Si nous prenons $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\}$, c'est un fermé vérifiant que

$$\frac{1}{\mu(E)} \int_E \bar{g} d\mu \in S. \tag{27.829}$$

EQooMRLCooYPEiUo

Voilà une petite formule qui va nous aider à utiliser le lemme 27.185. Nous ne pouvons cependant pas l'utiliser immédiatement parce que l'appartenance (27.829) n'est vraie que pour les parties de mesure finie.

(13) **Si $p = 1$, conclusion[1]**

Pour utiliser le lemme 27.185, nous utilisons l'hypothèse que Ω est σ-fini. Soient des mesurables A_i de mesure fine tels que $\bigcup_{i\in\mathbb{N}} A_i = \Omega$.

Pour chaque i nous considérons la restriction $g_i \colon A_i \to \mathbb{C}$ de g à A_i. Par le point précédent, elle vérifie

$$\frac{1}{\mu(A_i)} \int_{A_i} \bar{g}_i d\mu = \frac{1}{\mu(A_i)} \int_{A_i} \bar{g} d\mu \in S. \tag{27.830}$$

En appliquant le lemme 27.185 à l'espace restreint $(A_i, \mathcal{A}_i, \mu_i)$, nous concluons $\bar{g}_i \in S$ presque partout, ce qui signifie que $\|g_i\|_\infty \in S$. Nous en concluons que

$$\|g_i\|_\infty \leqslant \|\phi\| \tag{27.831}$$

où, dans ce contexte, $\|g_i\|_\infty$ signifie $\sup_{x\in A_i}|g_i(x)|$.
Nous avons alors

$$\|g\|_\infty = \sup_{x\in\Omega}|g(x)| = \sup_{i\in\mathbb{N}}\|g_i\|_\infty \leqslant \|\phi\|. \tag{27.832}$$

Une petite justification pour cela ? Prenons une suite x_k telle que $|g(x_k)| \to \|g\|_\infty$. Vu que les A_i recouvrent Ω, existe un naturel $i(k)$ tel que $x_k \in A_{i(k)}$. Nous avons alors

$$|g(x_k)| \leqslant \|g_{i(k)}\|_\infty \leqslant \|\phi\|. \tag{27.833}$$

Cela pour conclure que $g \in L^\infty$.
Notons que cet argument ne tient pas avec $p > 1$ parce que l'équation (27.827) terminerait sur $\|\phi\|\mu(E)^{1/p}$. Du coup l'ensemble S à prendre serait $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\mu(E)^{1/p-1}\}$ et nous sommes en dehors des hypothèses du lemme parce qu'il n'y a pas d'ensemble *indépendant* de E dans lequel l'intégrale $\frac{1}{\mu(E)}\int_E \bar{g}d\mu$ prend ses valeurs.

(14) $1 < p < \infty$
La fonction

$$\alpha(x) = \begin{cases} \frac{g(x)}{|g(x)|} & \text{si } g(x) \neq 0 \\ 1 & \text{si } g(x) = 0 \end{cases} \tag{27.834}$$

a la propriété de faire $\alpha g = |g|$ en même temps que $|\alpha(x)| = 1$ pour tout x. Nous définissons

$$E_n = \{x \text{ tel que } |g(x)| \leqslant n\} \tag{27.835}$$

et

$$f_n = \mathbb{1}_{E_n}|g^{q-1}|\alpha. \tag{27.836}$$

Ce qui est bien avec ces fonctions c'est que [104]

$$|f_n|^p = |g^{p(q-1)}||\alpha|^p = |g|^q \tag{27.837}$$

sur E_n. Dans E_n nous avons $|f_n| = |g^{q-1}| \leqslant n^{q-1}$ et dans E_n nous avons $f_n = 0$. Au final, $f_n \in L^\infty$. Par ce que nous avons vu plus haut, nous avons alors

$$\phi(f_n) = \Phi_g(f_n). \tag{27.838}$$

Par ailleurs,

$$f_n\bar{g} = \mathbb{1}_{E_n}|g^{q-1}|\frac{g}{|g|}\bar{g}, \tag{27.839}$$

donc [105]

$$\begin{aligned}
\left|\int_{E_n}|g|^q d\mu\right| &= |\int_\Omega f_n\bar{g}d\mu| && \text{(27.840a)}\\
&= |\phi(f_n)| && \text{(27.840b)}\\
&\leqslant \|\phi\|\|f_n\|_p && \text{(27.840c)}\\
&= \|\phi\|\left(\int_{E_n}|f_n|^p\right)^{1/p} && \text{(27.840d)}\\
&= \|\phi\|\left(\int_{E_n}|g|^q\right)^{1/p}. && \text{(27.840e)}
\end{aligned}$$

Nous avons de ce fait une inégalité de la forme $A \leqslant \|\phi\|A^{1/p}$ et donc aussi $A^{1/p} \leqslant \|\phi\|^{1/p}A^{1/p^2}$, et donc $A \leqslant \|\phi\|\|\phi\|^{1/p}A^{1/p^2}$. Continuant ainsi à injecter l'inégalité dans elle-même, pour tout $k \in \mathbb{N}$ nous avons :

$$\left|\int_{E_n}|g|^q d\mu\right| \leqslant \|\phi\|^{1+\frac{1}{p}+\cdots+\frac{1}{p^k}}\left(\int_{E_n}|g|^q d\mu\right)^{1/p^k}. \tag{27.841}$$

104. C'est ici que nous utilisons le lien entre p et q. En l'occurrence, de $1/p + 1/q = 1$ nous déduisons $q(p-1) = p$.
105. Dans [11], cette équation arrive sans modules, ce qui me laisse entendre que $\phi(f_n)$ est réel et positif pour pouvoir écrire que $\phi(f_n) \leqslant \|\phi\|\|f_n\|_p$, mais je ne comprends pas pourquoi.

Nous pouvons passer à la limite $k \to \infty$. Sachant que $p > 1$ nous savons $A^{1/k} \to 1$ et

$$1 + \frac{1}{p} + \cdots + \frac{1}{p^k} \to \frac{p}{p-1} = q. \tag{27.842}$$

Nous avons alors

$$\int_{E_n} |g|^q d\mu \leqslant \|\phi\|^q. \tag{27.843}$$

L'intégrale s'écrit tout aussi bien sous la forme $\int_\Omega |g|^q \mathbb{1}_{E_n}$. La fonction dans l'intégrale est une suite croissante de fonctions mesurables à valeurs dans $[0, \infty]$. Nous pouvons alors permuter l'intégrale et la limite $n \to \infty$ en utilisant la convergence monotone (théorème 14.178) qui donne alors $\int_\Omega |g|^q \leqslant \|\phi\|^q$ ou encore

$$\|g\|_q \leqslant \|\phi\|. \tag{27.844}$$

Ceci achève de prouver que l'application $\phi \mapsto \Phi_g$ est une isométrie, et donc le théorème.

□

THOooXMVTooBAbyvr

Théorème 27.187.
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Pour $f, g \in \mathrm{Fun}(\Omega, \mathbb{C})$, nous écrivons

$$\Phi_g(f) = \int_\Omega f\bar{g} \tag{27.845}$$

pour toutes les combinaisons de f et g pour lesquelles l'intégrale a un sens.

ITEMooNCVEooTyNsoJ
(1) Si $1 < p < \infty$, alors $L^p(\Omega, \mathcal{A}, \mu)$ est réflexif[106].

ITEMooTQDJooFShTiA
(2) Si $(\Omega, \mathcal{A}, \mu)$ est σ-finie, alors

ITEMooHMMZooMQxWgB
(2a) L'application $\Phi\colon L^\infty(\Omega, \mathcal{A}, \mu) \to L^1(\Omega, \mathcal{A}, \mu)'$ est une bijection isométrique.

ITEMooBFFZooNxoHER
(2b) L'application $\Phi\colon L^1(\Omega, \mathcal{A}, \mu) \to L^\infty(\Omega, \mathcal{A}, \mu)'$ est une injection isométrique.

Note : nous verrons dans 27.191 que l'application $\Phi\colon L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)'$ n'est pas surjective.

Démonstration. En plusieurs parties, en notant toujours p et q les exposants conjugués, c'est-à-dire $\frac{1}{p} + \frac{1}{q} = 1$.

(1) **Pour (1)**
Nous considérons les applications

$$\begin{aligned} \alpha_1 &\colon L^p \to (L^q)' \\ \alpha_1(g)f &= \int_\Omega f\bar{g} \end{aligned} \tag{27.846}$$

et

$$\begin{aligned} \alpha_2 &\colon L^q \to (L^p)' \\ \alpha_2(g)f &= \int_\Omega g\bar{f}. \end{aligned} \tag{27.847}$$

Ce sont des bijections linéaires par le théorème de représentation de Riesz 27.186 [107]. L'espace L^p est de Banach par le théorème de Riesz-Fischer 27.45 ; le lemme 27.182 s'applique donc. Nous considérons alors les applications correspondantes

$$\beta_1 \colon (L^p)' \to (L^q)'' \tag{27.848}$$

106. Définition 27.181.
107. Notez que la conjugaison complexe dans α_2 n'est pas à la même place que dans α_1. L'application α_1 est exactement la même que le Φ de représentation de Riesz alors que α_2 est un tout petit peu modifiée. La raison de ce changement n'est pas très profonde : c'est seulement pour que les choses tombent juste à la fin.

et

$$\beta_2 \colon (L^q)' \to (L^p)'', \tag{27.849}$$

qui sont également des bijections linéaires. Nous allons montrer que la bijection

$$\beta_2 \circ \alpha_1 \colon L^p \to (L^p)'' \tag{27.850}$$

est l'application J de la définition 27.181.

Soient $f \in L^p$ et $\varphi \in (L^p)'$. Il existe $g \in L^q$ tel que $\varphi = \alpha_2(g)$. Nous pouvons donc calculer d'une part, en développant β_2 par sa définition (27.771),

$$(\beta_2 \circ \alpha_1)(f)\varphi = \beta_2\big(\alpha_1(f)\big)\varphi = \alpha_1(f)\big(\alpha_2^{-1}(\varphi)\big) = \alpha_1(f)(g) = \int_\Omega g\bar{f}, \tag{27.851}$$

et d'autre part

$$J(f)\varphi = \varphi(f) = \alpha_2(g)f = \int_\Omega g\bar{f}. \tag{27.852}$$

(2) **Pour (2a)** Il s'agit du théorème 27.186(2).

(3) **Pour (2b)[1]** Il nous reste à couvrir le cas de $(L^\infty)'$. Pour $g \in L^1$ nous prouvons que $\Phi_g \in (L^\infty)'$.

(3a) $\Phi_g(f)$ **est bien définie** Nous prouvons d'abord que si $f \in L^\infty$, alors l'intégrale $\int_\Omega f\bar{g}$ est bien définie. Par définition du supremum essentiel [108], il existe $M > 0$ tel que $|f(x)| < M$ pour tout x hors d'une partie A de mesure nulle. Nous avons alors

$$\int_\Omega |f\bar{g}| = \int_{\Omega\setminus A} |f\bar{g}| \leqslant M \int_{\Omega\setminus A} |f| = M \int_\Omega |f| < \infty. \tag{27.853}$$

(3b) Φ_g **est continue** Soit une suite $f_k \xrightarrow{L^\infty} f$ ainsi que $g \in L^1$. Pour chaque k, il existe une partie de mesure nulle A_k et un nombre $M_k = \|f_k\|_{L^\infty}$ tel que $|f_k(x)| < \|f_k\|_{L^\infty}$ pour tout x hors de A_k. Nous avons alors

$$|\Phi_g(f_k)| \leqslant \int_{\Omega\setminus A_k} |f_k\bar{g}|d\mu \leqslant \|f_k\|_{L^\infty} \int_{\Omega\setminus A_k} |g| \leqslant \|f_k\|_{L^\infty}\|g\|_1. \tag{27.854}$$

Vu que par hypothèse $f_k \to 0$ dans L^∞, nous avons $\|f_k\|_{L^\infty} \to 0$, et donc aussi

$$|\Phi_g(f_k)| \to 0. \tag{27.855}$$

□

27.188.

Les gens qui n'ont peur d'aucun abus de notations écrivent le théorème 27.187(2a) en disant simplement que $L^1 = L^\infty$ et le démontrent de la façon suivante. Le théorème 27.186(1) nous indique que

$$(L^p)' = L^q. \tag{27.856}$$

EQooJVDEooNGYEtg

Vu que $1 < p < \infty$, nous avons aussi $1 < q < \infty$ et donc $(L^q)' = L^p$. En prenant le dual des deux côtés de (27.856),

$$(L^p)'' = (L^q)' = L^p. \tag{27.857}$$

À ce moment, un second abus vient en aide et nous disons qu'un espace V est réflexif quand $V = V''$. Et voilà.

108. Voir les définitions 27.22 et 27.23.

PropUKLZZZh

Proposition 27.189.
Soit $f \in L^p(\Omega)$ telle que

$$\int_\Omega f\varphi = 0 \tag{27.858}$$

pour tout $\varphi \in \mathscr{D}(\Omega)$. Alors $f = 0$ presque partout.

Démonstration. Nous considérons la forme linéaire $\Phi_f \in (L^q)'$ donnée par

$$\begin{aligned} \Phi_f \colon L^p &\to \mathbb{C} \\ u &\mapsto \int_\Omega fu \end{aligned} \tag{27.859}$$

Par hypothèse cette forme est nulle sur la partie dense $C_c^\infty(\Omega)$. Si (φ_n) est une suite dans $C_c^\infty(\Omega)$ convergente vers u dans L^p, nous avons pour tout n que

$$0 = \Phi_f(\varphi_n) \tag{27.860}$$

En passant à la limite, nous voyons que Φ_f est la forme nulle. Elle est donc égale à Φ_0. La partie « unicité » du théorème de représentation de Riesz 27.186 nous indique alors que $f = 0$ dans L^p et donc $f = 0$ presque partout. □

PropLGoLtcS

Proposition 27.190.
Si $f \in L^1_{loc}(I)$ est telle que

$$\int_I f\varphi' = 0 \tag{27.861}$$

pour tout $\varphi \in C_c^\infty(I)$, alors il existe une constante C telle que $f = C$ presque partout.

Démonstration. Soit $\psi \in C_c^\infty(I)$ une fonction d'intégrale 1 sur I. Si $w \in C_c^\infty(I)$ alors nous considérons la fonction

$$h = w - \psi \int_I w, \tag{27.862}$$

qui est dans $C_c^\infty(I)$ et dont l'intégrale sur I est nulle. Par la proposition 17.2, la fonction h admet une primitive dans $C_c^\infty(I)$; et nous notons φ cette primitive. L'hypothèse appliquée à φ donne

$$0 = \int_I f\varphi' = \int_I f\left(w - \psi\int_I w\right) = \int_I fw - \underbrace{\left(\int_I f(x)\psi(x)dx\right)}_{C}\left(\int_I w(y)dy\right) = \int_I w(f - C). \tag{27.863}$$

L'annulation de la dernière intégrale implique par la proposition 27.189 que $f - C = 0$ dans L^2, c'est-à-dire $f = C$ presque partout. □

PROPooXNRRooUdgFPr

Proposition 27.191 ([195])**.**
L'application

$$\begin{aligned} \Phi \colon L^1(\mathbb{R}^d) &\to L^\infty(\mathbb{R}^d)' \\ \Phi_g(f) &= \int_\Omega f\bar{g}d\lambda \end{aligned} \tag{27.864}$$

n'est pas surjective.

Démonstration. Nous allons construire un élément de $(L^\infty)'$ qui n'est pas dans l'image de L^1.

(1) **Mise en place du décor** Nous considérons l'espace vectoriel normé $\left(L^\infty(\mathbb{R}^d), N_\infty\right)$ défini en 27.22 et 27.23. Dedans, nous considérons le sous-espace D des classes des fonctions dans $\mathscr{D}(\mathbb{R}^d)$; nous aurions très envie de le noter $\left(\mathscr{D}(\mathbb{R}^d), N_\infty\right)$, mais nous allons le noter $\left(D, N_\infty\right)$ parce que $\mathscr{D}$ est un espace de fonctions tandis que nous considérons un espace de classes de fonctions.

Si $f \in D$, alors f est une classe de fonctions contenant un unique représentant continu [109] ; nous le notons $\tilde{f}$.

(2) **Une première fonctionnelle**
Avec ça nous posons [110]

$$\begin{aligned}\phi_0 \colon D &\to \mathbb{C}\\ f &\mapsto \tilde{f}(0).\end{aligned} \tag{27.865}$$

(3) **ϕ_0 est linéaire** Ça ne devrait pas poser de problèmes.

(4) **ϕ_0 est continue** Soit $f_k \xrightarrow{N_\infty} 0$. Nous devons prouver que $\tilde{f}_k(0) \xrightarrow{\mathbb{C}} 0$. Supposons le contraire et considérons $\epsilon > 0$ ainsi que k tels que $\tilde{f}_k(0) > \epsilon$.
Par continuité de $\tilde{f}_k$, il existe un $\delta > 0$ tel que $\tilde{f}_k(x) > \epsilon/2$ pour tout $x \in B(0, \delta)$. Avec cela nous avons $N_\infty(f_k) > \epsilon/2$, et une impossibilité d'avoir [111] $f_k \xrightarrow{N_\infty} 0$.

(5) **ϕ_0 est de norme finie** C'est parce qu'elle est continue.

(6) **Utilisation de Hahn-Banach** Le théorème de Hahn-Banach 27.173 donne une extension continue

$$\phi \colon L^\infty(\mathbb{R}^d) \to \mathbb{C}. \tag{27.866}$$

Cela est un élément de $(L^\infty)'$ et nous allons montrer qu'il n'est pas dans $\Phi(L^1)$.

(7) **Par l'absurde** Supposons qu'il existe $u \in L^1(\mathbb{R}^d)$ telle que $\phi = \Phi_u$. Pour tout $f \in L^\infty$, nous avons

$$\phi(f) = \Phi_u(f) = \int_{\mathbb{R}^d} f\bar{u}. \tag{27.867}$$

Si f est la classe d'une fonction de $\mathscr{D}(\mathbb{R}^d)$ s'annulant en 0, alors $\phi(f) = 0$. De telles fonctions non identiquement nulles existent par le lemme 15.159 [112].
Soient $\varphi \in \mathscr{D}(\mathbb{R}^d)$, un compact K ne contenant pas 0 et χ_K sa fonction indicatrice. Nous avons $\chi_K\varphi \in \mathcal{L}^\infty(\mathbb{R}^d)$, et $(\chi_K\varphi)(0) = 0$. En passant aux classes, $\phi\big([\chi_K f]\big) = 0$. Nous avons :

$$0 = \int_{\mathbb{R}^d} u\overline{(\chi_K\varphi)} = \int_{\mathbb{R}^d} \overline{\varphi}(u\overline{\chi_K}). \tag{27.868}$$

Vu que cela est valable pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$, la proposition 27.189 dit que $u\chi_K = 0$ presque partout.

(8) **Le coup du compact**
Soit une suite de compacts K_n recouvrant $\mathbb{R}^d\backslash\{0\}$. Par exemple

$$K_n = \overline{B(0,n)}\backslash B(0, \frac{1}{n}). \tag{27.869}$$

Pour chacun des K_n, nous avons $u\chi_{K_n} = 0$ presque partout. Il existe donc une partie de mesure nulle N_k telle que $u\chi_{K_n}$ est nulle à part sur N_k. Au total, u est non nulle seulement sur $\bigcup_k N_k$ et peut-être en 0.
Bref, u est non nulle sur une partie de mesure nulle par le lemme 14.28.

(9) **Conclusion**
La fonction u est nulle presque partout. Donc $\phi = \Phi_u = 0$. Nous savons pourtant que ϕ n'est pas nulle parce qu'il existe des fonctions dans $\mathscr{D}(\mathbb{R}^d)$ qui ne s'annulent pas en 0.
Cela est une contradiction. Donc ϕ n'est pas dans l'image de Φ tout en étant dans $L^\infty(\mathbb{R}^d)'$.

□

109. Existence parce que les éléments de D sont des classes d'éléments de $\mathscr{D}$ qui sont C^∞. Unicité par la proposition 20.154.
110. Dans [195], l'auteur ne définit pas L^∞ comme un espace de classes de fonctions, et ces complications disparaissent.
111. Je vous laisse emballer ce raisonnement dans « si pour tout N, il existe $k > N$ tel que ».
112. Ce lemme donne un exemple f sur $\mathbb{R}$. Si vous voulez vraiment un exemple dans $\mathbb{R}^d$, prenez $g(x) = f(\|x\|)$.

Dans [701], il est dit que « la preuve [du lemme suivant], un peu fastidieuse mais en rien ingénieuse, est laissée en exercice ». La preuve est donc de moi ; elle est un tout petit peu ingénieuse mais en rien fastidieuse. J'espère ne pas m'être trompé et me demande bien ce que l'auteur avait en tête. Ma preuve s'appuie sur la proposition 17.101 dont la preuve ne me paraît pas non plus « fastidieuse mais en rien ingénieuse ».

LEMooLDQRooEGWDlm

Lemme 27.192 ([701, 1])**.**
Soit $r > 0$. Il existe $\delta > 0$ tel que pour tout $s, t \in \mathbb{C}$ vérifiant $|s| \leqslant 1$, $|t| \leqslant 1$ et $|s - t| \geqslant r$ nous ayons

$$\left|\frac{s+t}{2}\right|^p \leqslant (1-\delta)\frac{|s|^p + |t|^p}{2}. \tag{27.870}$$

Démonstration. Soit $r > 0$. La partie de $\mathbb{C}^2$ donnée par

$$D = \{(s,t) \in \mathbb{C}^2 \text{ tel que } |s| \leqslant 1, |t| \leqslant 1, |s-t| \geqslant r\} \tag{27.871}$$

est compacte. En effet elle est bornée (par la sphère de rayon $\sqrt{2}$) et fermée comme intersection de fermés [113]. Nous considérons la fonction $\Delta\colon D \to \mathbb{R}$ donnée par

$$\left|\frac{s+t}{2}\right|^p = \Delta(s,t)\frac{|s|^p + |t|^p}{2}. \tag{27.872}$$

Si vous voulez une expression explicite,

$$\Delta(s,t) = \frac{2^{p-1}|s+t|^p}{|s|^p + |t|^p}. \tag{27.873}$$

Cela est bien défini et continu sur D parce que le complémentaire D^c (qui est ouvert) contient $(0,0)$ et donc aussi un voisinage de $(0,0)$.

La proposition 17.101 nous dit que la fonction $z \mapsto |z|^p$ est strictement convexe. En prenant la définition 17.94 de la stricte convexité avec $\theta = \frac{1}{2}$, nous trouvons que

$$\Delta(s,t) < 1 \tag{27.874}$$

pour tout $(s,t) \in D$. Vu que par ailleurs Δ est une fonction continue sur le compact D, elle atteint un minimum dans D. Soit Δ_0 ce minimum qui vérifie forcément $\Delta_0 < 1$.

En posant $1 - \delta = \Delta_0$ nous avons le résultat. □

27.18.1 Le théorème de Jordan

Les définitions de chemins, de lacets et de courbes de Jordan sont dans 20.75.

LEMooZPRLooJPvrOE

Lemme-Définition 27.193 ([205])**.**
Soit un lacet $\gamma\colon [a,b] \to X$. Nous considérons l'application

$$\begin{aligned} q\colon [a,b] &\to S^1 \\ t &\mapsto e^{2i\pi t/(b-a)} \end{aligned} \tag{27.875}$$

Il existe une unique application continue $\tilde{\gamma}\colon S^1 \to \text{Image}(\gamma)$ telle que $\gamma = \tilde{\gamma} \circ q$.

*Cette application $\tilde{\gamma}$ est l'application **quotient** associée à γ.*

Démonstration. La proposition 18.61(1) dit que q est une bijection continue. Donc l'existence et l'unicité d'une application $\tilde{\gamma} = \gamma \circ q^{-1}$. Cette application est continue comme composée d'applications continues. □

LEMooCGVOooVPlSRD

Lemme 27.194 ([205])**.**
Soit $\gamma\colon [a,b] \to X$ un chemin d'image Γ.

113. Lemme 7.6 suivit du théorème de Borel-Lebesgue 9.24.

ITEMooWKVAooCQDvpL

(1) *Si γ est un chemin de Jordan*[114]*, alors $\gamma : [a,b] \to \Gamma$ est un isomorphisme d'espaces topologiques.*

ITEMooVYMXooEtgPJT

(2) *Si γ est un lacet de Jordan, alors le quotient*[115] *$\tilde{\gamma} : S^1 \to \Gamma$ est un homéomorphisme*[116]*.*

Démonstration. En deux parties.

(1) **Pour (1)** Un chemin de Jordan est injectif, et donc bijectif sur son image. Il est donc une bijection continue depuis $[a,b]$ qui est compact. Il est donc un homéomorphisme par le lemme 7.221.

(2) **Pour (2)** L'application $\gamma : [a,b[\to \Gamma$ est une bijection continue parce que γ est un lacet de Jordan. Son quotient $\tilde{\gamma} : S^1 \to \Gamma$ est donc une bijection continue depuis le compact S^1. Le lemme 7.221 conclut que c'est un homéomorphisme.

□

CORooPGFLooUTVZMi

Corolaire 27.195.
À propos de courbes simples et de Jordan.

ITEMooIUAXooOfvNov

(1) *Toute courbe simple est homéomorphe à $[0,1]$.*

(2) *Toute courbe de Jordan est homéomorphe à S^1.*

Démonstration. Une courbe simple est l'image d'un lacet de Jordan. Donc $\gamma : [a,b] \to \Gamma$ est un homéomorphisme par le lemme 27.194(1). Le lemme 10.270 fournit un homéomorphisme entre $[a,b]$ et $[0,1]$. Une composition d'homéomorphismes est un homéomorphisme.

Une courbe de Jordan est l'image d'un lacet de Jordan $\gamma : [a,b] \to \Gamma$. Le lemme 27.194(2) dit alors que Γ est homéomorphe à S^1. □

DEFooURFMooXIaRkl

Définition 27.196 ([205])**.**
Soit un compact K dans $\mathbb{C}$. Nous notons

(1) *$\mathcal{G}(K)$ le groupe multiplicatif des fonctions continues $f : K \to \mathbb{C}^*$.*

(2) *$\mathcal{E}(K)$ le sous-groupe des fonctions admettant un logarithme continu*[117]*.*

(3) *$G(K)$ le quotient*[118] *$G(K) = \mathcal{G}(K)/\mathcal{E}(K)$.*

LEMooHEOWooHTtHsJ

Lemme 27.197 ([205])**.**
Si K_1 et K_2 sont des compacts homéomorphes, alors les groupes $G(K_1)$ et $G(K_2)$ sont isomorphes.

Démonstration. Soit un homéomorphisme $\theta : K_1 \to K_2$. Nous posons

$$\begin{aligned} \psi : \mathcal{G}(K_1) &\to \mathcal{G}(K_2) \\ f &\mapsto f \circ \theta^{-1}. \end{aligned} \tag{27.876}$$

(1) **Injective** Suppose $\psi(f) = \psi(g)$. Vu que θ est une bijection, l'égalité $f \circ \theta^{-1} = g \circ \theta^{-1}$ implique $f = g$. Donc ψ est injective.

(2) **Surjective** Soit $f \in \mathcal{G}(K_2)$. Nous avons $f \circ \theta \in \mathcal{G}(K_1)$, et évidemment $\psi(f \circ \theta) = f$.

(3) **Morphisme** Nous avons

$$\psi(fg) = fg \circ \theta^{-1} = (f \circ \theta^{-1})(g \circ \theta^{-1}) = \psi(f)\psi(g). \tag{27.877}$$

114. Définition 20.75.
115. Définition 27.193.
116. Bijection continue d'inverse continu.
117. Logarithme continu, définition 26.44.
118. Le groupe $\mathcal{G}(K)$ est commutatif ; donc pas de problèmes pour que $\mathcal{E}(K)$ soit normal.

Nous avons prouvé que ψ est un isomorphisme. Nous devons maintenant voir qu'il passe au quotient, c'est-à-dire que $\psi\big(\mathcal{E}(K_1)\big) \subset \mathcal{E}(K_2)$. Soit $f \in \mathcal{E}(K_1)$, soit un logarithme continu g de f, c'est-à-dire $\exp\big(g(x)\big) = f(x)$. Pour $y \in K_2$, il existe $x \in K_1$ tel que $y = \theta(x)$. Dans ce cas nous avons

$$\exp\big((g \circ \theta^{-1})(y)\big) = \exp\big(g(x)\big) = f(x) = (f \circ \theta^{-1})(y). \tag{27.878}$$

Autrement dit,

$$\exp\big(\psi(g)(y)\big) = \psi(f). \tag{27.879}$$

Donc $\psi(g)$ est un logarithme continu de $\psi(f)$. □

DEFooKXHUooTsKuOe

Définition 27.198.
Petite notation. Si un compact K est donné, pour $p \in \mathbb{C}\backslash K$, nous notons

$$\begin{aligned} f_p\colon K &\to \mathbb{C}^* \\ z &\mapsto z - p. \end{aligned} \tag{27.880}$$

LEMooSULYooWiEoyf

Lemme 27.199.
Nous considérons l'application

$$\begin{aligned} \phi\colon \mathbb{C} &\to \mathbb{C} \\ x &\mapsto \lambda x + \alpha \end{aligned} \tag{27.881}$$

avec $\lambda \neq 0$ ($\lambda \in \mathbb{C}$) et $\alpha \in \mathbb{C}$.
Soit $q \in \mathbb{C}$. Nous avons[119]

$$f_q\big(\phi(x)\big) = \lambda f_{\phi^{-1}(q)}(x). \tag{27.882}$$

Démonstration. En ce qui concerne la multiplication nous avons $f_q(\lambda x) = \lambda x - q$, et donc

$$\frac{1}{\lambda} f_q(\lambda x) = x - \frac{q}{\lambda} = f_{q/\lambda}(x). \tag{27.883}$$

Nous retenons : $f_q(\lambda x) = \lambda f_{q/\lambda}(x)$. En ce qui concerne la translation, c'est plus facile :

$$f_q(x + \alpha) = x + \alpha - q = f_{q-\alpha}(x). \tag{27.884}$$

Lorsque nous faisons les deux en même temps,

$$f_q\big(\phi(x)\big) = f_{q-\alpha}(\lambda x) = \lambda f_{(q-\alpha)/\lambda}(x). \tag{27.885}$$

Il suffit maintenant de remarquer que $\phi^{-1}(q) = \frac{q-\alpha}{\lambda}$, ce qui est facilement vérifié :

$$\phi\left(\frac{q-\alpha}{\lambda}\right) = \lambda\left(\frac{q-\alpha}{\lambda}\right) + \alpha = q. \tag{27.886}$$

□

LEMooBZUCooHWfolf

Lemme 27.200 ([205])**.**
Soit un lacet de Jordan[120] *$\gamma\colon [a,b] \to \mathbb{C}$. Nous notons Γ son image.*

(1) Γ est compact.

(2) L'application[121]

$$\begin{aligned} I_\gamma\colon \mathcal{G}(\Gamma) &\to \mathbb{Z} \\ f &\mapsto \mathrm{Ind}(f \circ \gamma, 0) \end{aligned} \tag{27.887}$$

est un morphisme surjectif de noyau $\mathcal{E}(\Gamma)$.

119. L'application f_q est celle de la définition 27.198.
120. Définition 20.75.
121. Rappel définitions 27.196 pour $\mathcal{G}(\Gamma)$ et $\mathcal{E}(\Gamma)$.

Démonstration. Nous notons $\mathcal{L}(J,\mathbb{C}^*)$ le groupe multiplicatif des lacets $J\to\mathbb{C}^*$. Le fait que Γ soit compact est parce qu'il est l'image du compact $[a,b]$ par l'application continue γ (théorème 7.220).

Nous considérons l'application

$$\begin{aligned}\xi\colon \mathcal{G}(\Gamma)&\to\mathcal{L}(J,\mathbb{C}^*)\\ f&\mapsto f\circ\gamma.\end{aligned}\tag{27.888}$$

(1) **ξ est un morphisme** Nous avons

$$\xi(fg)=(fg)\circ\gamma=(f\circ\gamma)(g\circ\gamma)=\xi(f)\xi(g).\tag{27.889}$$

(2) **ξ est injective** Nous nous souvenons de l'application quotient $\tilde{\gamma}\colon S^1\to\Gamma$, qui est un homéomorphisme par le lemme 27.194(1). Nous nous souvenons également de l'application

$$\begin{aligned}q\colon [a,b]&\to\mathbb{C}\\ t&\mapsto\exp\big(2i\pi t/(b-a)\big)\end{aligned}\tag{27.890}$$

qui vérifie $\gamma=\tilde{\gamma}\circ q$. Nous montrons que ceci est l'inverse de ξ :

$$\begin{aligned}\theta\colon \mathcal{L}(J,\mathbb{C}^*)&\to\mathcal{G}(\Gamma)\\ \lambda&\mapsto\tilde{\lambda}\circ\tilde{\gamma}^{-1}.\end{aligned}\tag{27.891}$$

Soit $\sigma\in\mathcal{L}(J,\mathbb{C}^*)$. Nous avons

$$\xi\big(\theta(\sigma)\big)=\theta(\sigma)\circ\gamma=\tilde{\sigma}\circ\tilde{\gamma}^{-1}\circ\gamma=\tilde{\sigma}\circ\tilde{\gamma}^{-1}\circ\tilde{\gamma}\circ q=\tilde{\sigma}\circ q=\sigma.\tag{27.892}$$

(3) **Conclusion 1** La proposition 26.55 donne un morphisme surjectif $\psi\colon\mathcal{L}(J,\mathbb{C}^*)\to\mathbb{Z}$. L'application I_γ dont nous parlons ici est exactement $I_\gamma=\psi\circ\xi$. Donc I_γ est un morphisme surjectif.

(4) **Noyau** Nous devons encore montrer que le noyau de I_γ est $\mathcal{E}(\Gamma)$.

(4a) **Dans un sens** Soit $k\in\ker(I_\gamma)$. Alors $(\psi\circ\xi)(k)=0$, c'est-à-dire $\psi(k\circ\gamma)=0$. La proposition 26.55(3) dit qu'alors $k\circ\gamma$ est un lacet homotope dans $\mathbb{C}^*$ à un lacet constant. Donc il existe une application continue $H\colon[0,1]\times[a,b]\to\mathbb{C}^*$ telle que

$$\begin{cases}H(0,t)=(k\circ\gamma)(t) & \text{(27.893a)}\\ H(1,t)=1 & \text{(27.893b)}\end{cases}$$

pour tout t. Vu que l'application $\gamma\colon[a,b[\to\mathbb{C}^*$ est injective, nous pouvons considérer son inverse ; pour la suite, quand nous écrirons γ^{-1}, c'est bien de $\gamma^{-1}\colon\Gamma\to[a,b[$ que nous parlerons. C'est une application continue par le lemme 7.223. Nous pouvons donc considérer

$$\begin{aligned}L\colon[0,1]\times\Gamma&\to\mathbb{C}^*\\ (u,z)&\mapsto H\big(u,\gamma^{-1}(z)\big).\end{aligned}\tag{27.894}$$

Cela est une application continue parce que H et γ le sont. Nous avons

$$L(0,z)=H\big(0,\gamma^{-1}(z)\big)=(k\circ\gamma)\big(\gamma^{-1}(z)\big)=k(z),\tag{27.895}$$

et

$$L(1,z)=1.\tag{27.896}$$

Donc k est homotope à une application constante sur Γ. Le théorème de Borsuk 26.47 dit alors que k admet un logarithme continu sur Γ, c'est-à-dire que $k\in\mathcal{E}(\Gamma)$.

(4b) **Dans l'autre sens** C'est le même raisonnement que le premier, mais un peu plus simple parce qu'il n'y a pas de débats sur la continuité. Soit $k \in \mathcal{E}(\Gamma)$. Le théorème de Boruk indique que k est homotope dans $\mathbb{C}^*$ à une application constante. Il existe une application continue $H\colon [0,1]\times\Gamma\to\mathbb{C}^*$ telle que

$$\begin{cases} H(0,z)=k(z) & (27.897a)\\ H(1,z)=1 & (27.897b)\end{cases}$$

pour tout $z\in\Gamma$. Le chemin $k\circ\gamma$ est homotope à un lacet constant. En effet nous avons l'application continue

$$\begin{aligned} L\colon [0,1]\times[a,b]&\to\mathbb{C}^*\\ (u,t)&\mapsto H\big(u,\gamma(t)\big)\end{aligned} \tag{27.898}$$

qui vérifie $L(0,t)=(k\circ\gamma)(t)$ et $L(1,t)=1$. Nous avons donc

$$0=\psi(k\circ\gamma)=I_\gamma(k). \tag{27.899}$$

Nous avons prouvé que $k\in\ker(I_\gamma)$.

□

LEMooODIPooBZJPAW

Lemme 27.201 ([205]).
Soit un lacet de Jordan γ dont l'image est Γ. Des fonctions $f_0,f_1\in\mathcal{G}(\Gamma)$ sont homotopes dans $\mathbb{C}^*$ si et seulement si les lacets $f_0\circ\gamma$ et $f_1\circ\gamma$ le sont.

Démonstration. Dans les deux sens.

(1) $\Rightarrow$ Nous supposons que f_0 et f_1 sont homotopes : il existe une application continue $H\colon[0,1]\times\Gamma\to\mathbb{C}^*$ telle que

$$\begin{cases} H(0,z)=f_0(z) & (27.900a)\\ H(1,z)=f_1(z) & (27.900b)\end{cases}$$

pour tout $z\in\Gamma$. L'homotopie entre $f_0\circ\gamma$ et $f_1\circ\gamma$ est donnée par l'application continue

$$\begin{aligned} L\colon [0,1]\times[a,b]&\to\mathbb{C}^*\\ (u,t)&\mapsto H\big(u,\gamma(t)\big).\end{aligned} \tag{27.901}$$

Elle vérifie $L(0,t)=(f_0\circ\gamma)(t)$ et $L(1,t)=(f_1\circ\gamma)(t)$.

(2) $\Leftarrow$ Nous supposons l'existence d'une application continue $H\colon[0,1]\times[a,b]\to\mathbb{C}^*$ telle que $H(0,t)=(f_0\circ\gamma)(t)$ et $H(1,t)=(f_1\circ\gamma)(t)$.
L'application $\gamma\colon[a,b[\to\Gamma$ est continue et bijective. Sa réciproque γ^{-1} est donc également continue (lemme 7.223). Nous pouvons donc considérer l'application continue

$$\begin{aligned} L\colon [0,1]\times\Gamma&\to\mathbb{C}^*\\ (u,z)&\mapsto H\big(u,\gamma^{-1}(z)\big).\end{aligned} \tag{27.902}$$

Elle vérifie

$$L(0,z)=H\big(0,\gamma^{-1}(z)\big)=(f_0\circ\gamma)\big(\gamma^{-1}(z)\big)=f_0(z) \tag{27.903}$$

et $L(1,z)=f_1(z)$. Donc L est une homotopie dans C^* entre f_0 et f_1.

□

CORooDOMAooHtKPTe

Corolaire 27.202 ([205]).
Si Γ est une courbe de Jordan, alors[122] $G(\Gamma)\simeq\mathbb{Z}$.

122. Rappel : $G(\Gamma)$ est défini en 27.196.

Démonstration. Nous allons utiliser le premier théorème d'isomorphisme 2.6. Si γ est un lacet de Jordan dont la courbe est Γ, le lemme 27.200 dit que $I_\gamma\colon \mathcal{G}(\Gamma)\to\mathbb{Z}$ est un morphisme surjectif tel que $\ker(I_\gamma)=\mathcal{E}(\Gamma)$. Le théorème d'isomorphisme dit alors que

$$\frac{\mathcal{G}(\Gamma)}{\mathcal{E}(\Gamma)}=I_\gamma\big(\mathcal{G}(\Gamma)\big)=\mathbb{Z}. \tag{27.904}$$

□

Rappel, si K est compact dans $\mathbb{C}$, et si $p\in\mathbb{C}$, nous considérons la fonction $f_p\in\mathcal{G}(K)$ par $f_p(z)=z-p$.

LEMooAKVFooFMHaOZ

Lemme 27.203 ([1]).
Soient une courbe de Jordan γ ainsi que $p\in\mathbb{C}$. Nous avons

$$\mathrm{Ind}(f_p\circ\gamma,0)=\mathrm{Ind}(\gamma,p). \tag{27.905}$$

PROPooUPPEooEQOOkh

Proposition 27.204 ([205]).
Soit $p\in\mathbb{C}$. Si V est un disque ou un rectangle de centre p, alors $\mathcal{G}(\partial V)$ est engendré par $\mathcal{E}(\partial V)$ et par f_p.

Démonstration. Ici ∂V peut être vu comme un chemin que nous allons noter aussi γ. Utilisant le lemme 27.203, nous avons

$$I_{\partial V}(f_p)=\mathrm{Ind}(f_p\circ\partial V,0)=\mathrm{Ind}(\partial V,p). \tag{27.906}$$

EQooRSXRooTJwDNR

(1) **Si V est un disque** Alors ∂V est le chemin

$$\begin{aligned}\gamma\colon[0,2\pi]&\to\mathbb{C}\\ t&\mapsto p+Re^{it}.\end{aligned} \tag{27.907}$$

Nous avons alors le calcul suivant :

$$\begin{aligned}\mathrm{Ind}(\gamma,p)&=\frac{1}{2\pi i}\int_0^{2\pi}\frac{\gamma'(t)}{\gamma(t)-p}dt &(27.908a)\\ &=\frac{1}{2\pi i}\int_0^{2\pi}\frac{Rie^{it}}{Re^{it}}dt &(27.908b)\\ &=\frac{1}{2\pi i}\int_0^{2\pi}idt &(27.908c)\\ &=1. &(27.908d)\end{aligned}$$

(2) **Si V est un rectangle** Supposons que V ait une hauteur de h et une largeur de l tout en étant centré en p. Le chemin ∂V se décompose en quatre partie dont la première est

$$\begin{aligned}\gamma_1\colon[0,l]&\to\mathbb{C}\backslash\{p\}\\ t&\mapsto p-\frac{hi}{2}-\frac{l}{2}+t\end{aligned} \tag{27.909}$$

Je vous laisse voir les autres parties et calculer les intégrales [123]

Dans les deux cas nous avons $\mathrm{Ind}(\partial V,p)=1$ et donc, en continuant (27.906),

$$I_{\partial V}(f_p)=1. \tag{27.910}$$

Nous savons par le lemme 27.200 que

$$\begin{aligned}I_{\partial V}\colon\mathcal{G}(\partial V)&\to\mathbb{Z}\\ f&\mapsto\mathrm{Ind}(f\circ\gamma,0)\end{aligned} \tag{27.911}$$

123. Je n'ai pas vérifié ; si ça pose un problème, écrivez-moi.

est un morphisme surjectif de noyau $\mathcal{E}(\Gamma)$. Le fait que ce soit un morphisme implique que

$$I_{\partial V}(f_p^n) = nI_{\partial V}(f_p) = n. \tag{27.912}$$

Mais $G(\partial V) = \mathcal{G}(\partial V)/\mathcal{E}(\partial V)$. Donc nous avons déjà prouvé que $\{f_p^k\}_{k\in\mathbb{N}}$ contient un élément de chaque classe dans $\mathcal{G}(\partial V)/\mathcal{E}(\partial V)$.

Soit $f \in \mathcal{G}(\partial V)$. En notant [.] la classe par rapport à $\mathcal{E}(\partial V)$, il existe $k \in \mathbb{N}$ tel que

$$f \in \big[I_{\partial V}(f_p^k)\big]. \tag{27.913}$$

Autrement dit, en posant $\alpha = I_{\partial V}(f_p^k)$, il existe $h \in \mathcal{E}(\partial V)$ tel que $f = \alpha h$. Nous avons donc bien prouvé que $\{f_p, \mathcal{E}(\partial V)\}$ engendre $\mathcal{G}(\partial V)$. □

PROPooJNZQooLWDKww

Proposition 27.205 ([205]).
Soit un compact K de $\mathbb{C}$ ainsi que p, q deux points dans la même composante connexe de $\mathbb{C}\backslash K$. Alors

ITEMooQMKOooQHhUGn

(1) $f_p/f_q \in \mathcal{E}(K)$.

ITEMooUJLPooApgXIF

(2) Si p est dans la composante non bornée[124] *de $\mathbb{C}\backslash K$, alors $f_p \in \mathcal{E}(K)$.*

Démonstration. Soient p et q dans la même composante connexe de $\mathbb{C}\backslash K$. La proposition 9.67 nous autorise à considérer un chemin $\gamma\colon [0,1] \to \mathbb{C}\backslash K$ tel que $\gamma(0) = p$ et $\gamma(1) = q$. Nous posons

$$\begin{aligned} F\colon [0,1]\times K &\to \mathbb{C}^* \\ (t,z) &\mapsto f_{\gamma(t)}(z) = z - \gamma(t). \end{aligned} \tag{27.914}$$

C'est une application continue vérifiant $F(0,z) = f_p(z)$ et $F(1,z) = f_q(z)$. Donc les applications f_p et f_q sont homotopes dans $\mathbb{C}^*$. La proposition 26.15 nous indique qu'alors l'application f_p/f_q est homotope à 1. Le théorème de Borsuk 26.47 conclut que $f_p/f_q \in \mathcal{E}(K)$.

Soit p dans la composante non bornée de $\mathbb{C}\backslash K$. Vu que K est borné, nous pouvons prendre $r > 0$ tel que $K \subset B(0,r)$. Soit $p_0 \notin \overline{B(0,r)}$; ce point est également dans la composante non bornée de K. Par le point (1), nous avons $f_p/f_{p_0} \in \mathcal{E}(K)$.

Nous allons prouver que $f_{p_0} \in \mathcal{E}(K)$. Pour cela nous posons

$$\begin{aligned} F\colon [0,1]\times K &\to \mathbb{C}^* \\ (t,z) &\mapsto tz - p_0. \end{aligned} \tag{27.915}$$

Vu que $t \in [0,1]$ et que $z \in K \subset B(0,r)$, nous avons toujours $tz \in B(0,r)$. Étant donné que p_0 n'est pas dans $B(0,r)$, l'application F prend bien ses valeurs dans $\mathbb{C}^*$.

L'application F vérifie également $F(0,z) = -p_0$ et $F(1,z) = z - p_0 = f_{p_0}(z)$. Donc f_{p_0} est homotope dans $\mathbb{C}^*$ à l'application constante $-p_0$. Nous en déduisons que $f_{p_0} \in \mathcal{E}(K)$. Au final,

$$f_p = f_{p_0}\frac{f_p}{f_{p_0}} \in \mathcal{E}(K) \tag{27.916}$$

parce que f_{p_0} et f_p/f_{p_0} sont dans $\mathcal{E}(K)$. □

LEMooOTMGooZeUoCy

Lemme 27.206 ([205]).
Soient des composantes connexes bornées $\mathcal{O}_1$, ..., $\mathcal{O}_N$ de $\mathbb{C}\backslash K$ que nous supposons deux à deux disjointes. Soient $p_i \in \mathcal{O}_i$. Si $n_1, \dots, n_N \in \mathbb{Z}$ sont des entiers non tous nuls, alors

$$\prod_{i=1}^{N} f_{p_i}^{n_i} \notin \mathcal{E}(K). \tag{27.917}$$

Démonstration. En plusieurs points.

124. Il y en a une seule, c'est le lemme 26.49.

(1) **Par l'absurde** Nous notons $f = \prod_{i=1}^{N} f_{p_i}^{n_i}$ et nous supposons par l'absurde que $f \in \mathcal{E}(K)$. Par le théorème de Borsuk 26.47, nous avons une extension continue $\tilde{f}\colon \mathbb{C} \to \mathbb{C}^*$.

(2) $\partial\mathcal{O}_i \subset K$ C'est le lemme 7.78.

(3) p_j **n'est pas dans** $\bar{\mathcal{O}}_i$ Nous savons que $p_j \in \mathcal{O}_j$. Mais $\partial\mathcal{O}_i \subset K$ et $K \cap \mathcal{O}_j = \varnothing$. Donc p_j n'est pas dans $\partial\mathcal{O}_i$. Comme les $\mathcal{O}_i$ sont disjoints, nous avons aussi $p_j \notin \mathcal{O}_i$. Bref,

$$p_j \notin \bar{\mathcal{O}}_i. \tag{27.918}$$

(4) **Quelques fonctions** Nous posons

$$\begin{aligned} G_i\colon \bar{\mathcal{O}}_i &\to \mathbb{C}^* \\ z &\mapsto \prod_{j\neq i}(z-p_j)^{n_j}. \end{aligned} \tag{27.919}$$

Le fait que G_i prenne ses valeurs dans $\mathbb{C}^*$ est parce que $p_j \notin \bar{\mathcal{O}}_i$ pour tout $i \neq j$.

Nous reprenons notre extension continue $\tilde{f}\colon \mathbb{C} \to \mathbb{C}^*$ et nous posons

$$\begin{aligned} F_i\colon \bar{\mathcal{O}}_i &\to \mathbb{C}^* \\ z &\mapsto \frac{\tilde{f}(z)}{G_i(z)}. \end{aligned} \tag{27.920}$$

Vu que G_i ne s'annule pas sur $\bar{\mathcal{O}}_i$, cette fonction F_i est bien définie et continue.

Nous nous fixons i, et nous allons prouver que $n_i = 0$. Nous posons $D = B(0,R)$ avec R assez grand pour contenir $\bar{\mathcal{O}}_i$. Posons

$$\begin{aligned} H_i\colon \overline{B(0,R)} &\to \mathbb{C}^* \\ z &\mapsto \begin{cases} F_i(z) & \text{si } z \in \bar{\mathcal{O}}_i \\ (z-p_i)^{n_i} & \text{sinon.} \end{cases} \end{aligned} \tag{27.921}$$

(5) H_i **est continue** L'application F_i est continue sur $\bar{\mathcal{O}}_i$. L'application $z \mapsto (z-p_i)^{n_i}$ est continue sur $\mathbb{C}$.

Soit $z \in \partial\mathcal{O}_i$. Nous avons $z \in \partial\mathcal{O}_i \subset K$ et donc $\tilde{f}(z) = f(z)$. Donc

$$F_i(z) = \frac{f(z)}{G_i(z)} = f_{p_i}(z)^{n_i}. \tag{27.922}$$

Les deux applications F_i et $z \mapsto (z-p_i)^{n_i}$ sont égales sur le bord $\partial\mathcal{O}_i$. Donc H_i est continue sur $\overline{B(0,R)}$.

(6) **Logarithme continu** La partie $\overline{B(0,R)}$ est convexe et compacte. Le corolaire 26.48 dit que H_i admet un logarithme continu. La proposition 26.58 nous affirme que pour tout lacet $\gamma\colon [0,1] \to \overline{B(0,R)}$, nous avons $\operatorname{Ind}(H_i \circ \gamma, 0) = 0$. En particulier le chemin $\partial B(0,R)$ est dans $\overline{B(0,R)}$. Avant de calculer $\operatorname{Ind}(H_i \circ \gamma, 0)$, remarquons que $\gamma(t) \in \overline{B(0,R)}$ et reste donc en-dehors de $\bar{\mathcal{O}}_i$. Donc

$$H_i\big(\gamma(t)\big) = \big(\gamma(t) - p_i\big)^{n_i} \tag{27.923}$$

et

$$(H_i \circ \gamma)'(t) = n_i\gamma'(t)\big(\gamma(t)-p_i\big)^{n_i-1}. \tag{27.924}$$

De plus si vous voulez une forme explicite de γ, c'est pas compliqué : c'est un cercle de rayon

R. Nous avons un calcul :

$$\begin{aligned}
0 = \operatorname{Ind}(H_i \circ \gamma, 0) &= \frac{1}{2\pi i} \int_{H_i \circ \gamma} \frac{dz}{z} && \text{(27.925a)}\\
&= \frac{1}{2\pi i} \int_0^1 \frac{(H_i \circ \gamma)'(t)}{(H_i \circ \gamma)(t)} && \text{(27.925b)}\\
&= \frac{1}{2\pi i} \int_0^1 \frac{n_i \gamma'(t) \big(\gamma(t) - p_i\big)^{n_i - 1}}{\big(\gamma(t) - p_i\big)^{n_i}} && \text{(27.925c)}\\
&= \frac{1}{2\pi i} \int_0^1 \frac{n_i \gamma'(t)}{\gamma(t) - p_i} dt && \text{(27.925d)}\\
&= n_i \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - p_i} && \text{(27.925e)}\\
&= n_i \operatorname{Ind}(\gamma, p_i). && \text{(27.925f)}
\end{aligned}$$

Vu que $p_i \in B(0, R)$ et que γ est un cercle autour, nous avons $\operatorname{Ind}(\gamma, p_i) = 1$. Donc

$$0 = \operatorname{Ind}(H_i \circ \gamma, 0) = n_i. \tag{27.926}$$

□

LEMooEJRMooNJhMov

Lemme 27.207 ([205, 1])**.**
Nous notons $(\mathcal{O}_i)_{i \in I}$ les composantes connexes bornées de $\mathbb{C}\backslash K$. Pour chaque $i \in I$ nous considérons $p_i \in \mathcal{O}_i$. Alors

$$\mathcal{G}(K) = \operatorname{gr}\big(\mathcal{E}(K), \{f_{p_i}\}_{i \in I}\big). \tag{27.927}$$

Démonstration. Nous commençons par prouver que si $q \in \mathbb{C}\backslash K$, alors

$$f_q \in \operatorname{gr}\big(\mathcal{E}(K), \{f_{p_i}\}_{i \in I}\big). \tag{27.928}$$

Il y a deux possibilités : soit q est dans la composante non bornée, soit il est dans un des $\mathcal{O}_i$.

(1) **Si q est dans la composante non bornée** Alors la proposition 27.205(2) dit que $f_q \in \mathcal{E}(K)$.

(2) **Si $q \in \mathcal{O}_i$** Alors la proposition 27.205(1) dit que $f_q / f_{p_i} \in \mathcal{E}(K)$. Dans ce cas nous avons

$$f_q = f_{p_i} \frac{f_q}{f_{p_i}} \in \operatorname{gr}\big(f_{p_i}, \mathcal{E}(K)\big). \tag{27.929}$$

Soit $f \in \mathcal{G}(K)$. Nous devons prouver l'existence de $g \in \mathcal{E}(K)$, $q_1, \dots, q_l \in \mathbb{C}\backslash K$ et $m_1, \dots, m_l \in \mathbb{Z}$ tels que

$$f = g \prod_{i=1}^{l} f_{q_i}^{m_i}. \tag{27.930}$$

Nous commençons par supposer que $K \subset Q = [0,1] \times [0,1]$.

(1) **Les petits carrés** Soit $n \in \mathbb{N}$. Nous subdivisons le carré Q en carrés fermés de taille $1/n$; il y a n^2 tels carrés. Nous notons $Q_1, \dots, Q_l$ ceux qui n'ont pas d'intersection avec K, et nous notons A l'union des autres.

Grâce au lemme 10.27, nous savons que pour tout $z \in A$, nous avons $d(z, K) \leqslant \sqrt{2}/n$.

(2) **Prolongement** La partie K est fermée, f est continue dessus. Le théorème de Tietze 27.178 donne un prolongement continu $F \colon \mathbb{C} \to \mathbb{C}$. Nous avons

$$K \subset S = \{z \in \mathbb{C} \text{ tel que } F(z) \neq 0\}. \tag{27.931}$$

Vu que F est continue, S est ouverte.

(3) **Grand n** Nous prouvons que n peut être choisi assez grand pour que F ne s'annule pas sur A. La fonction F est continue et ne s'annule pas sur l'ouvert S qui contient K. Si $S = \mathbb{C}$, alors n'importe quel n convient.
Sinon nous considérons

$$\begin{aligned}\varphi\colon K &\to [0,\infty[\\ z &\mapsto d(z,S^c).\end{aligned} \tag{27.932}$$

L'application φ est continue sur le compact K. Elle a donc un minimum global que nous nommons s. Pour tout $z \in K$ nous avons $d(z,S^c) \geqslant s$. Donc pour tout $z \in K$ nous avons $B(z,s/2) \subset S$. Nous prenons n assez grand pour avoir

$$\frac{\sqrt{2}}{n} < \frac{s}{2}. \tag{27.933}$$ EQooEWQYooWYsImg

Avec ça nous avons $A \subset S$. En effet, soit $z \in A$. Nous avons $d(z,K) < \sqrt{2}/n < s/2$. Donc il existe $k \in K$ tel que

$$z \in B(k, \frac{s}{2}). \tag{27.934}$$

Comme $k \in K$, nous avons $d(k,S^c) = \varphi(k) \geqslant s$ et donc $B(k,s/2) \subset S$. Nous avons prouvé que $z \in S$.
Bref, nous prenons n assez grand pour vérifier (27.933), et nous avons alors

$$A \subset \{z \in \mathbb{C} \text{ tel que } F(z) \neq 0\}. \tag{27.935}$$

(4) **Une propriété à prouver** SPooVQHTooUNukEE Nous allons prouver l'énoncé suivant. Soit $k = 0,\dots,l$, soit une application continue $h\colon K \to \mathbb{C}^*$ admettant un prolongement continu $F\colon A \cup (Q_1 \cup \dots \cup Q_k) \to \mathbb{C}^*$. Alors il existe $m \in \mathbb{N}$ tel que l'application $h/f_{q_{k+1}}^m$ admette un prolongement continu $\tilde{F}\colon A \cup (Q_1 \cup \dots \cup Q_{k+1}) \to \mathbb{C}^*$.

(5) **Note pour f**
Jusqu'ici, nous avons prouvé que f vérifie les hypothèses de la récurrence pour $k = 0$.
De plus, cette propriété à prouver n'est pas à prouver par récurrence sur k. Le k est fixé, et nous prouvons la propriété de façon directe. Une récurrence va venir un peu plus loin.

(6) **Preuve de la propriété** Soit $h\colon K \to \mathbb{C}^*$ continue et admettant un prolongement continu $F\colon A \cup (Q_1 \cup \dots Q_k) \to \mathbb{C}^*$. Notons $B_k = A \cup (Q_1 \cup \dots \cup Q_k)$ et posons

$$\begin{aligned}g\colon \partial Q_{k+1} &\to \mathbb{C}^*\\ z &\mapsto \begin{cases} F(z) & \text{si } z \in Q_{k+1} \cap B_k\\ s(z) & \text{sinon.}\end{cases}\end{aligned} \tag{27.936}$$

Voyons ce que s peut être pour que g soit continue. La partie ∂Q_{k+1} est le bord d'un carré et est donc constitué de 4 côtés. La valeur de g sur certains de ces côtés (ceux qui sont aussi dans B_k) est prescrite par F. Sur les autres côtés, nous pouvons mettre n'importe quoi pourvu que ce soit égal à F sur les coins et que $s(z) \neq 0$. Étant donné que les valeurs de F sur les coins est non nulle et que $\mathbb{C}^*$ est connexe par arcs, c'est possible.
Nous notons q_{k+1} le centre de Q_{k+1}. La proposition 27.204 s'applique au compact ∂Q_{k+1}.

$$g \in \mathcal{G}(\partial Q_{k+1}) = \operatorname{gr}\big(\mathcal{E}(\partial Q_{k+1}), f_{q_{k+1}}\big). \tag{27.937}$$

Il existe donc $\xi \in \mathcal{E}(\partial Q_{k+1})$ et $m \in \mathbb{Z}$ tels que

$$g(z) = \xi(z) f_{q_{k+1}}^m(z). \tag{27.938}$$

Prouvons que $Q_{k+1} \cap B_k$ est une union de côtés de Q_{k+1}. Nous avons

$$Q_{k+1} \cap B_k = (Q_{k+1} \cap A) \cup (Q_{k+1} \cap Q_1) \cup \dots \cup (Q_{k+1} \cap Q_k). \tag{27.939}$$ EQooSPSQooVErzCI

Nous nous souvenons que $A = \bigcup_{i=1}^{s} R_i$ où R_i sont les carrés qui intersectent K. Donc $R_i \cap Q_j$ est au maximum les côtés de Q_j. Bref, chacun des termes de l'union à droite de (27.939) est une union de côtés de Q_{k+1}.

L'application $\xi \in \mathcal{E}(\partial Q_{k+1})$ admet un logarithme continu et admet donc une extension continue $\mathbb{C} \to \mathbb{C}^*$ (théorème de Borsuk 26.47). Nous considérons cette extension seulement sur Q_{k+1} et nous l'appelons $G\colon Q_{k+1} \to \mathbb{C}^*$. Nous posons

$$\begin{aligned}\tilde{F}\colon A \cup (Q_1 \cup \ldots \cup Q_{k+1}) &\to \mathbb{C}^* \\ z &\mapsto \begin{cases} G(z) & \text{si } z \in Q_{k+1} \\ \frac{F(z)}{(z-q_{k+1})^m} & \text{si } z \in A \cup (Q_1 \cup \ldots \cup Q_k). \end{cases}\end{aligned} \tag{27.940}$$

Voyons que cette fonction est continue en montrant que les deux morceaux ont les mêmes valeurs sur l'intersection. Soit $z \in Q_{k+1} \cap (A \cup Q_1 \cup \ldots \cup Q_k)$. Alors $z \in \partial Q_{k+1}$ parce que nous avons dit que (27.939) est dans le bord dans Q_{k+1}. Dans ce cas nous avons

$$G(z) = \xi(z) = \frac{g(z)}{f^m_{q_{k+1}}(z)} \tag{27.941a}$$

$$= \frac{F(z)}{f^m_{q_{k+1}}(z)}, \tag{27.941b}$$

$$\tag{27.941c}$$

et donc bien la continuité de $\tilde{F}$.

Si $z \in K$, étant donné que $K \subset A$ nous avons

$$\tilde{F}(z) = \frac{F(z)}{(z - q_{q_{k+1}})^m} = \frac{h(z)}{f^m_{q_{k+1}}(z)}, \tag{27.942}$$

donc $\tilde{F}$ prolonge bien $h/f^m_{q_{k+1}}$.

(7) **Récurrence** En appliquant la propriété (4) l fois, nous construisons une suite d'entiers $(m_i)_{i=1,\ldots,l}$ et une application continue $F_l\colon A \cup (Q_1 \cup \ldots \cup Q_l) = Q \to \mathbb{C}^*$ qui prolonge $f/\prod_{i=1}^{l} f^{m_i}_{q_i}$.

Sur K nous avons alors

$$f = F_l \prod_{i=1}^{l} f^{m_i}_{q_i}. \tag{27.943}$$

C'est une application continue sur le convexe Q. Donc $F_l \in \mathcal{E}(Q)$ et donc $F_l \in \mathcal{E}(K)$.

Maintenant vous croyez avoir fini mais non. Vous devez encore vérifier que tout est encore valable si K n'est pas contenu dans $[0,1] \times [0,1]$.

Soit un compact général $K \subset \mathbb{C}$. Pour $\lambda > 0$ et $\alpha \in \mathbb{C}$, nous posons $\phi(x) = \lambda x + \alpha$. Il existe forcément un choix de λ et de α tels que $\phi(K) \subset [0,1] \times i[0,1]$.

Soit $f \in \mathcal{G}(K)$. Nous appliquons le résultat déjà prouvé à l'application $f \circ \phi^{-1} \in \mathcal{G}\big(\phi(K)\big)$. Nous avons $g \in \mathcal{E}\big(\phi(K)\big)$, $q_1, \ldots, q_l \in \mathbb{C}\backslash\phi(K)$ et $m_1, \ldots, m_l \in \mathbb{Z}$ tels que, pour tout $y \in \phi(K)$,

$$(f \circ \phi^{-1})(y) = g(y) \prod_{i=1}^{l} f^{m_i}_{q_i}(y). \tag{27.944}$$

Pour $x \in K$ nous avons donc

$$f(x) = (g \circ \phi)(x) \prod_{i=1}^{l} f^{m_i}_{q_i}\big(\phi(x)\big). \tag{27.945}$$

C'est le moment de sortir le lemme 27.199 :

$$f(x) = (g \circ \phi)(x) \prod_{i=1}^{l} f^{m_i}_{q_i}\big(\phi(x)\big) = (g \circ \phi)(x) \prod_{i=1}^{l} \lambda^{m_i} f^{m_i}_{\phi^{-1}(q_i)}(x). \tag{27.946}$$

En posant

$$\begin{cases} \tilde{g}(x) = \left(\prod_{i=1}^{l} \lambda^{m_i} \right) (g \circ \phi)(x) & (27.947a) \\ \tilde{m}_i = m_i & (27.947b) \\ \tilde{q}_i = \phi^{-1}(q_i), & (27.947c) \end{cases}$$

nous avons $\tilde{q}_i \in \mathbb{C}\backslash K$, $\tilde{g} \in \mathcal{E}(K)$ et $f = \tilde{g} \prod_{i=1}^{l} f_{\tilde{q}_i}$. □

LEMooBHOGooXxBYGA

Lemme 27.208.
Une partie ouverte de $\mathbb{C}$ a au plus un nombre dénombrable de composantes connexes.

Démonstration. Soit un ouvert A de $\mathbb{C}$. Nous notons $\mathcal{B}$ une base dénombrable d'ouverts de $\mathbb{C}$ (proposition 7.138). Une composante connexe de A est ouverte [125]. Les composantes connexes de A sont disjointes, et chacune est l'union d'une quantité (au plus) dénombrable d'éléments de $\mathcal{B}$. Si il y avait une quantité non dénombrable de composantes connexes, cela ferait une quantité non dénombrable d'éléments différents de $\mathcal{B}$. Donc non. □

THOooDGXSooMMHTds

Théorème 27.209.
Soit un compact K dans $\mathbb{C}$. Alors

(1) Le nombre de composantes connexes bornées de $\mathbb{C}\backslash K$ est fini ou dénombrable.

ITEMooWLGDooZGORrP

(2) En notant N le nombre de composantes connexes bornées de $\mathbb{C}\backslash K$, nous avons

$$G(K) = \mathcal{G}(K)/\mathcal{E}(K) \simeq \mathbb{Z}^N \tag{27.948}$$

où, si N est infini, $\mathbb{Z}^N$ désigne l'ensemble des suites finies dans $\mathbb{Z}$.

Démonstration. Le fait que le nombre de composantes connexes bornées de $\mathbb{C}\backslash K$ est au maximum dénombrable est le lemme 27.208.

Soient $\{\mathcal{O}_i\}_{i=0,\dots,N}$ les composantes connexe bornées de $\mathbb{C}\backslash K$ (avec $i \in \mathbb{N}$ dans le cas où N il y a une infinité dénombrable de composants connexes). Pour chaque i nous prenons un élément $p_i \in \mathcal{O}_i$. Nous notons aussi $\pi\colon \mathcal{G}(K) \to G(K)$ le morphisme canonique. Nous posons

$$\begin{aligned} \phi\colon \mathbb{Z}^N &\to G(K) \\ n &\mapsto \prod_{i=1}^{N} \pi(f_{p_i})^{n_i}. \end{aligned} \tag{27.949}$$

Notez que le produit a toujours un sens parce que $n \in \mathbb{Z}^N$ est toujours une suite finie (de taille N ou arbitraire) d'entiers. Pas de soucis de convergence.

(1) **ϕ est un morphisme** Sur $\mathbb{Z}^n$, nous mettons l'addition $(a+b)_i = a_i + b_i$. Si $N = \infty$ et si a et b ont des longueurs différentes, on prend pour N le maximum des deux longueurs. Nous avons :

$$\begin{aligned} \phi(a)\phi(b) &= \prod_{i=1}^{N} \pi(f_{p_i})^{a_i} \prod_{j=1}^{N} \pi(f_{p_j})^{b_j} & (27.950a) \\ &= \prod_{i=1}^{N} \pi(f_{p_i})^{a_i+b_i} & (27.950b) \\ &= \phi(a+b). & (27.950c) \end{aligned}$$

(2) **ϕ est injective** Soit $a \in \mathbb{Z}^N$. Si $a \neq 0$, alors le lemme 27.206 nous dit que $\prod_{i=1}^{N} f_{p_i}^{a_i} \notin \mathcal{E}(K)$. Vu que dans $G(K)$, le zéro est $\mathcal{E}(K)$, nous avons $\prod_{i=1}^{N} \pi(f_{p_i})^{a_i} \neq 0$ dès qu'un des a_i est non nul.

125. Proposition 7.77 et le fait que $\mathbb{C}$ est localement convexe.

(3) ϕ **est surjective** Soit $F \in G(K)$. Vu que π est surjective, il existe $f \in \mathcal{G}(K)$ tel que $\pi(f) = F$. Le lemme 27.207 nous dit qu'il existe $g \in \mathcal{E}(K)$ et $a \in \mathbb{Z}^N$ tels que

$$f = g \prod_{i=1}^{N} f_{p_i}^{a_i}. \tag{27.951}$$

Nous avons alors

$$F = \pi(f) = \pi(g)\pi\left(\prod_{i=1}^{N} f_{p_i}^{a_i}\right) = \prod_{i=1}^{N} \pi(f_{p_i})^{a_i} = \phi(a). \tag{27.952}$$

□

DEFooWQSAooDQYaip

Définition 27.210 ([205, 1]).
*Si $(G, +)$ est un groupe abélien, la **pseudo-dimension** de G est le cardinal maximum des familles $\mathbb{Z}$-indépendantes de G.*

Dans ce contexte, une famille $\mathbb{Z}$-indépendante est une partie $\{a_i\}_{i\in I}$ de G telle que si $\sum_i z_i a_i = 0$ avec $z_i \in \mathbb{Z}$, alors $z_i = 0$ pour tout i.

Si il existe une partie libre idempotente à un ensemble A, mais pas de parties libres surpotentes[126] *à A, nous disons que la pseudo-dimension de G est « la cardialité de A ».*

Dans tous les cas, nous notons $\operatorname{pdim}(G)$ *la pseudo-dimension de G.*

LEMooPASKooFGJgBz

Lemme 27.211 ([205, 1]).
À propos de pseudo-dimension.

(1) La pseudo-dimension de $\mathbb{Z}^n$ est n.

(2) La pseudo-dimension de $\mathbb{Z}^\infty$[127] *est la cardinalité de $\mathbb{N}$.*

Démonstration. Nous commençons par le cas de $\mathbb{Z}^n$ et nous ferons $\mathbb{Z}^\infty$ après.

(1) $\geqslant n$ Prenons la base canonique $(e_i)_{i=1,\dots,n}$ de $\mathbb{Z}^n$. C'est une partie $\mathbb{Z}$-indépendante de $\mathbb{Z}^n$. Donc la pseudo-dimension de $\mathbb{Z}^n$ est plus grande ou égale à n.

(2) $\leqslant n$ Dans l'autre sens, supposons que $\{a_i\}_{i\in I}$ soit $\mathbb{Z}$-indépendant dans $\mathbb{Z}^n$. Pour rappel, la définition 4.1 de partie libre ne parle que de sommes finies.

Prouvons que $\{a_i\}_{i\in I}$ est $\mathbb{Q}$-libre dans $\mathbb{Q}^n$. Soient une partie finie J de I ainsi que des rationnels $\{q_j\}_{j\in J}$ que $\sum_{j\in J} q_j a_j = 0$. Nous pouvons utiliser le lemme 1.400 pour considérer $k \in \mathbb{N}\backslash\{0\}$ tel que $kq_j \in \mathbb{Z}$ pour tout j (ici il est crucial que J soit fini, sinon on ne peut pas prendre un maximum). Dans ce cas nous avons $\sum_{j\in J} kq_j a_j = 0$, ce qui montre que $\{a_i\}$ n'est pas $\mathbb{Z}$-indépendante.

L'intérêt de $\mathbb{Q}$ est d'être un corps (alors que $\mathbb{Z}$ ne l'est pas). Nous savons que $\mathbb{Q}^n$ est un $\mathbb{Q}$-espace vectoriel de dimension n. Il ne contient donc pas de parties $\mathbb{Q}$-libres de cardinal plus grand que n.

Nous passons maintenant à la preuve pour $\mathbb{Z}^\infty$[1]. L'existence d'une partie dénombrable et $\mathbb{Z}$-libre dans $\mathbb{Z}^\infty$ est facile : il faut juste prendre e_i étant la suite nulle partout sauf en i. La partie $\{e_i\}_{i\in\mathbb{N}}$ est $\mathbb{Z}$-libre dans $\mathbb{Z}^\infty$.

Nous devons maintenant montrer que toute partie libre de $\mathbb{Z}^\infty$ est au maximum dénombrable. Soit une telle partie A. En posant

$$A_n = \{x \in A \text{ tel que } x_n \neq 0 \text{ et } x_k = 0 \text{ si } k > n\}. \tag{27.953}$$

Étant donné que A ne contient que des suites finies, $A = \bigcup_{n\in\mathbb{N}} A_n$.

D'abord A_n est $\mathbb{Z}$-libre dans $\mathbb{Z}^n$. En effet considérons une partie finie $\{a_i\}_{i\in I}$ de A_n ainsi que des entiers $\{z_i\}_{i\in I}$ tels que $\sum_{i\in I} z_i a_i = 0$. Vu que $\{a_i\}_{i\in I}$ est également une partie finie de A, nous en déduisons que $z_i = 0$ pour tout i.

La partie A_n étant libre dans $\mathbb{Z}^n$, elle est finie (de cardinal au maximum n). La partie A étant une union dénombrable de parties finies, elle est au maximum dénombrable. □

126. Définition 1.115.
127. Pour rappel, $\mathbb{Z}^\infty$ est l'ensemble des suites finies d'éléments de $\mathbb{Z}$.

LEMooEYDNooEUKUpn

Lemme 27.212.
Deux groupes abéliens isomorphes ont même pseudo-dimension.

Démonstration. Soient des groupes abéliens G et H ainsi qu'un isomorphisme $\phi\colon G\to H$. Soit une partie $\mathbb{Z}$-indépendante $\{a_i\}_{i\in I}$ dans G. Nous montrons que $\{\phi(a_i)\}_{i\in I}$ est $\mathbb{Z}$-indépendante dans H.

En effet si $\sum_{i\in I} z_i\phi(a_i)=0$, alors $0=\sum_i\phi(z_ia_i)=\phi\big(\sum_i z_ia_i\big)$, et donc $\sum_i z_ia_i=0$, ce qui entraine $z_i=0$.

Cela prouve que $\operatorname{pdim}(H)\geqslant\operatorname{pdim}(G)$. En inversant les rôles de G et H (et en utilisant ϕ^{-1}), nous prouvons que $\operatorname{pdim}(G)\geqslant\operatorname{pdim}(H)$.

Si une des pseudo-dimensions est infinie, le raisonnement donné doit être récrit en disant que G possède une partie $\mathbb{Z}$-libre de telle cardinalité si et seulement si H en possède une. □

CORooQNUIooKLtWVD

Corolaire 27.213 ([205]).
Soient des compacts K_1 et K_2 homéomorphes dans $\mathbb{C}$. Alors les parties $\mathbb{C}\backslash K_1$ et $\mathbb{C}\backslash K_2$ ont le même nombre de composantes connexes.

Démonstration. Vu que K_1 est homéomorphe à K_2, les groupes $G(K_1)$ et $G(K_2)$ sont isomorphes par le lemme 27.197. Le théorème 27.209(2) nous enseigne que, si on désigne par N_i le nombre de composantes connexes de $\mathbb{C}\backslash K_i$, alors $G(K_i)=\mathbb{Z}^{N_i}$. Le fait que $G(K_1)$ soit isomorphe à $G(K_2)$ implique

$$\mathbb{Z}^{N_1}\simeq\mathbb{Z}^{N_2}. \tag{27.954}$$

Le lemme 27.212 indique alors que $\operatorname{pdim}(\mathbb{Z}^{N_1})=\operatorname{pdim}(\mathbb{Z}^{N_2})$. Et enfin le lemme 27.211 conclut que $N_1=N_2$. □

Dans le lemme suivant, $S^1(a,r)$ est la sphère centrée en a et de rayon r : $S^1(a,r)=\{z\in\mathbb{C}$ tel que $|a-z|=r\}$.

LEMooUSJZooFVAlTa

Lemme 27.214.
La partie $\mathbb{C}\backslash S^1(0,1)$ a exactement 2 composantes connexes.

Démonstration. Posons $A=\{z\in\mathbb{C}$ tel que $|z|<1\}$ et $B=\{z$ tel que $|z|>1\}$. Nous avons $A\cup B=\mathbb{C}\backslash S^1$. La partie A est connexe et même convexe. En ce qui concerne la connexité de B, c'est le corolaire 18.64.

De plus A et B sont des ouverts disjoints.

Supposons que C soit une composante connexe de $\mathbb{C}\backslash S^1(0,1)$ contenant un point de A et un point de B. Soit $a\in C\cap A$. Par définition 7.73(1) d'une composante connexe, C contient tous les connexes contenant a. En particulier $A\subset C$. De même si $b\in C\cap B$, alors $B\subset C$.

Dans ce cas nous aurions $C=A\cup B=\mathbb{C}\backslash S^1(0,1)$. Mais cela est impossible parce que A et B forment un recouvrement de $\mathbb{C}\backslash S^1$ par deux ouverts disjoints.

Donc A et B sont des composantes connexes de $\mathbb{C}\backslash S^1$. Et ce sont les deux seules parce que $A\cup B=\mathbb{C}\backslash S^1$. □

27.18.2 Théorème de Jordan

ThoHSPWBuh

Théorème 27.215 (Théorème de Jordan[205, 702]).
Soit une courbe de Jordan Γ fermée dans $\mathbb{C}$.

ITEMooOAKXooMirWiD

(1) La partie $\mathbb{C}\backslash\Gamma$ a exactement deux composantes connexes.

ITEMooICBGooXHMALK

(2) Une composante connexe est bornée, l'autre non.

(3) Les deux composantes connexes ont Γ comme frontière.

Démonstration. En plusieurs parties.

(1) **Pour (1)** Nous savons par le corolaire 27.195(2) que Γ est homéomorphe à S^1. Le corolaire 27.213 nous dit alors que $\mathbb{C}\backslash\Gamma$ a le même nombre de composantes connexes que $\mathbb{C}\backslash S^1$, c'est-à-dire deux par le lemme 27.214.

(2) **Pour (2)** Le lemme 26.49 dit que $\mathbb{C}\backslash\Gamma$ a exactement une composante connexe non bornée. Vu que nous venons de voir qu'elle a exactement deux composants connexes, c'est que l'autre est bornée.

□

27.19 Topologie faible

DEFooZGLDooRRarRj

Définition 27.216 ([195, 1]).
Soit un espace de Banach[128] *sur le corps* $\mathbb{F}$ *(=* $\mathbb{R}$ *ou* $\mathbb{C}$*)* E*, et son dual* E'*. La* ***topologie faible*** *sur* E*, est la plus petite topologie* τ *pour laquelle*

$$\big(E, \|.\|\big)' = \big(E, \tau\big)'. \tag{27.955}$$

Autrement dit, c'est la topologie de la proposition 7.44 rendant continues toutes les applications de E'*.*

Elle sera notée τ_w *ou* $\sigma(E, E')$*.*

Vu que, pour rendre continue une application, il suffit que toutes les images inverses des ouverts de $\mathbb{F}$ soient des ouverts, la topologie faible sur E est également la topologie engendrée[129] par les parties $\varphi^{-1}(V)$ avec $\varphi \in E'$ et $V \in \tau_{\mathbb{F}}$.

Remarque 27.217.
Il faut noter que la topologie faible n'est pas une topologie métrique. Cela même si la condition $A_i x \to Ax$, elle, est métrique vu qu'elle est écrite dans E.

Dans le cas où E est de dimension infinie, la topologie faible est réellement différente de la topologie forte. Nous verrons à la sous-section 25.3.6 que dans le cas des projections sur un espace de Hilbert, l'égalité

$$\sum_{i=1}^{\infty} \mathtt{proj}_{u_i} = \mathrm{Id} \tag{27.956}$$

est vraie pour la topologie faible, mais pas pour la topologie forte.

Lemme 27.218 ([195]).
La topologie faible est Hausdorff[130].

Démonstration. Soient un espace de Banach E ainsi que $x_1 \neq x_2$ dans E. Les parties $\{x_1\}$ et $\{x_2\}$ vérifient les hypothèses de Hahn-Banach (seconde forme géométrique, 27.169). Il existe donc une fonctionnelle $\varphi\colon E \to \mathbb{F}$ telle que $\varphi(x_1) \neq \varphi(x_2)$. Avec un peu de bonne volonté, nous supposons que $\mathrm{Re}\big(\varphi(x_1)\big) \neq \mathrm{Re}\big(\varphi(x_2)\big)$. Soit $\alpha \in \mathbb{R}$ tel que

$$\mathrm{Re}\big(\varphi(x_1)\big) < \alpha < \mathrm{Re}\big(\varphi(x_2)\big). \tag{27.957}$$

Posons

$$\mathcal{O}_1 = \{x \in E \text{ tel que } \mathrm{Re}\big(\varphi(x)\big) < \alpha\}, \tag{27.958a}$$
$$\mathcal{O}_2 = \{x \in E \text{ tel que } \mathrm{Re}\big(\varphi(x)\big) > \alpha\}. \tag{27.958b}$$

Nous avons $\mathcal{O}_i = \varphi^{-1}(V_i)$ avec[131]

$$V_1 =]-\infty, \alpha[+ i\mathbb{R} \tag{27.959}$$

128. Définition 7.265.
129. Proposition 7.12.
130. Espace topologique Hausdorff, définition 7.60.
131. Adaptez si $\mathbb{F} = \mathbb{R}$ au lieu de $\mathbb{C}$.

et quelque chose du même genre pour V_2. Vu que V_1 et V_2 sont ouverts dans $\mathbb{F}$, et que φ est continue pour la topologie faible, les parties $\mathcal{O}_i$ sont ouvertes dans (E, τ_w).

Nous avons de plus $\mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$ et $x_i \in \mathcal{O}_i$, de telle sorte que x_1 et x_2 sont correctement séparés. □

Lemme 27.219 ([195]).
Soit un espace de Banach E. Soient $x_0 \in E$, $\varphi_1, \dots, \varphi_k \in E'$. La partie

$$V(\varphi_1, \dots, \varphi_k, \epsilon) = \{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon\} \tag{27.960}$$

est faiblement ouverte.

Démonstration. Remarquez que

$$V(\varphi_1, \dots, \varphi_k, \epsilon) = \bigcap_{i=1}^{k} \varphi_i^{-1}\{z \in \mathbb{C} \text{ tel que } |z - \varphi_i(x_0)| < \epsilon\}. \tag{27.961}$$

Vu que φ_i^{-1} transforme un ouvert en un ouvert, $V(\varphi_1, \dots, \varphi_k, \epsilon)$ est une intersection d'ouverts, et donc un ouvert. De plus x_0 est dedans parce que $\varphi_i(x_0 - x_0) = \varphi_i(0) = 0 < \epsilon$. □

Lemme 27.220 ([195]).
Soit $x_0 \in E$. Les ensembles

$$V(\varphi_1, \dots, \varphi_k, \epsilon) = \{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon \, \forall i\} \tag{27.962}$$

forment une base de topologie en x_0.

Démonstration. Soit un ouvert faible U contenant x_0. Nous devons prouver que U contient une partie de la forme $V(\varphi_1, \dots, \varphi_k, \epsilon)$. C'est le moment d'avoir bien en tête la construction de la topologie engendrée donnée en la proposition 7.12.

La topologie faible est engendrée par les parties de la forme $\varphi^{-1}(V)$ avec $\varphi \in E'$ et $V \in \tau_{\mathbb{F}}$. La partie U est une union de parties de la forme

$$\bigcap_{j=1}^{k} \varphi_j^{-1}(V_j) \tag{27.963}$$

où V_j est ouvert dans $\mathbb{F}$ et $\varphi_j \in E'$. Vu que $x_0 \in U$, il existe une partie de la forme (27.963) contenant x_0. Nous la notons W :

$$x_0 \in W = \bigcap_{j=1}^{k} \varphi_j^{-1}(V_j) \subset U. \tag{27.964}$$

Vu que V_j est un ouvert de $\mathbb{F}$ contenant $\varphi_j(x_0)$, il existe $\delta_j > 0$ tel que $B\big(\varphi_j(x_0), \delta_j\big) \subset V_j$. En prenant $\epsilon = \min_{j=1,\dots,k}(\delta_j)$, nous avons $B\big(\varphi_j(x_0), \epsilon\big) \subset V_j$ pour tout j. En particulier,

$$\bigcap_{j=1}^{k} \varphi_j^{-1}\Big(B\big(\varphi_j(x_0), \epsilon\big)\Big) \subset W. \tag{27.965}$$

Nous montrons à présent que $V(\varphi_1, \dots, \varphi_k, \epsilon) \subset W$. Pour cela nous considérons $x \in V(\varphi_1, \dots, \varphi_k, \epsilon)$ et nous montrons que $x \in \varphi_j^{-1}\Big(B\big(\varphi_j(x_0), \epsilon\big)\Big)$ pour tout $j = 1, \dots, k$. Nous avons

$$|\varphi_j(x) - \varphi_j(x_0)| = |\varphi_j(x - x_0)| < \epsilon, \tag{27.966}$$

et donc bien $\varphi_j(x) \in B\big(\varphi_k(x_0), \epsilon\big)$. Au final nous avons

$$V(\varphi_1, \dots, \varphi_k, \epsilon) \subset \bigcap_{j=1}^{k} \varphi_j^{-1}\Big(B\big(\varphi_j(x_0), \epsilon\big)\Big) \subset W \subset U. \tag{27.967}$$

□

PROPooFJBBooKkRwIp

Proposition 27.221 ([195, 703]).
Soient un espace de Banach E ainsi qu'une suite (x_n) dans E.

ITEMooDMMTooSBINKN

(1) $x_n \xrightarrow{w} x$ *si et seulement si* $\varphi(x_n) \to \varphi(x)$ *pour tout* $\varphi \in E'$.

ITEMooFZKXooNqFGUb

(2) $x_n \xrightarrow{\|.\|} x$ *implique* $x_n \xrightarrow{w} x$.

ITEMooXPTSooYPwNgU

(3) Si $x_n \xrightarrow{w} x$ *alors* $\{\|x_n\|\}$ *est borné et*

$$\|x\| \leqslant \liminf \|x_n\|. \tag{27.968}$$

ITEMooAFRFooOXdsBy

(4) Si $x_n \xrightarrow{w} x$ *et* $\varphi_n \xrightarrow{E'} \varphi$ *alors* $\varphi_n(x_n) \to \varphi(x)$.

Démonstration. Point par point.

(1) **Pour (1)** C'est le lemme 7.47.

(2) **Pour (2)** Nous prouvons que $\varphi(x_n) \to \varphi(x)$ pour tout $\varphi \in E'$:

$$|\varphi(x_n) - \varphi(x)| = |\varphi(x_n - x)| \leqslant \|\varphi\|_{E'}\|x_n - x\|_E \to 0 \tag{27.969}$$

parce que par hypothèse $\|x_n - x\| \to 0$.

(3) **Pour (3)**

Vu que $\varphi(x_n) \xrightarrow{\mathbb{F}} \varphi(x)$, l'ensemble $\{\varphi(x_n)\}$ est borné dans $\mathbb{F}$. Considérons les opérateur suivants :

$$\begin{aligned} T_n\colon E' &\to \mathbb{F} \\ \varphi &\mapsto \varphi(x_n). \end{aligned} \tag{27.970}$$

Cet opérateur a la propriété que $\|T_n\|_{(E')'} = \|x_n\|_E$; en effet, en utilisant la proposition 27.174,

$$\|T_n\| = \sup_{\|\varphi\|=1} |T_n(\varphi)| = \sup_{\|\varphi\|=1} |\varphi(x_n)| = \|x_n\|. \tag{27.971}$$

Chacun des T_n est donc un opérateur linéaire borné. Nous vérifions que la famille $\{T_n\}_{n\in\mathbb{N}}$ satisfait aux hypothèses du théorème de Banach-Steinhaus 10.151. D'abord E' est un Banach par la proposition 7.267. Ensuite nous venons de voir que chacun de T_n est borné dans E'. Et enfin, pour chaque $\varphi \in E'$ nous avons

$$\sup_{n\in\mathbb{N}} \|T_n(\varphi)\| = \sup_{n\in\mathbb{N}} |\varphi(x_n)| < \infty. \tag{27.972}$$

Le théorème de Banach-Steinhaus nous assure donc que

$$\sup_{n\in\mathbb{N}} \|x_n\| = \sup_{n\in\mathbb{N}} \|T_n\| < \infty. \tag{27.973}$$

Donc l'ensemble $\{\|x_n\|\}$ est borné.
Pour chaque $\varphi \in E'$ tel que $\|\varphi\| \leqslant 1$ nous avons

$$|\varphi(x_n)|_{\mathbb{F}} \leqslant \|\varphi\|_{E'}\|x_n\|_E \leqslant \|x_n\| \tag{27.974}$$

Étant donnée l'inégalité $|\varphi(x_n)| \leqslant \|x_n\|$ pour tout n et étant donnée la convergence $|\varphi(x_n)| \to |\varphi(x)|$, nous avons pour tout n :

$$|\varphi(x)| \leqslant \|x_n\|. \tag{27.975}$$

Cette inégalité valable pour tout n donne la conclusion :

$$|\varphi(x)| \leqslant \liminf \|x_n\|. \tag{27.976}$$

(4) **Pour (4)**
Nous avons [SUBEQooPHWLooZUiQLO]

$$|\varphi_n(x_n)-\varphi(x)| \leqslant |\varphi_n(x_n)-\varphi(x_n)|+|\varphi(x_n)-\varphi(x)| \tag{27.977a}$$
$$\leqslant \|x_n\|\|\varphi_n-\varphi\|+|\varphi(x_n-x)| \tag{27.977b}$$

Mais par hypothèse $\|\varphi_n-\varphi\|\to 0$, par le point (3), $\|x_n\|$ est borné, et par le point (1), nous avons $|\varphi(x_n)-\varphi(x)|\to 0$. Tout ça mis ensemble nous permet de prendre la limite dans (27.977) et de voir que $\|\varphi_n(x_n)-\varphi(x)\|\to 0$.

□

Proposition 27.222.
Si E est un espace de Banach de dimension finie, alors sa topologie normée est la même que sa topologie faible : $\tau_w=\tau_{\|.\|}$.

Démonstration. Vu que E' est défini comme étant l'ensemble des formes continues pour la topologie $\|.\|$, et que τ_w est la plus petite topologie pour laquelle tous les éléments de E' sont continus, nous avons $\tau_w\subset\tau_{\|.\|}$.

Pour prouver l'inclusion inverse, nous considérons $U\in\tau_{\|.\|}$ et nous prouvons que U est également un ouvert faible en montrant que tout élément de U est inclus dans un ouvert faible contenu dans U.

Soit $x_0\in U$ ainsi que $r>0$ tel que $B(x_0,r)\subset U$. Nous considérons une base $\{e_i\}_{i=1,\dots,n}$ de E telle que $\|e_i\|=1$ pour tout i. Vu que tout élément de E peut être décomposé de façon unique en $x=\sum_i x_ie_i$, nous considérons les fonctionnelles linéaires

$$\begin{aligned}\varphi_j\colon E&\to\mathbb{F}\\ x&\mapsto x_j.\end{aligned} \tag{27.978}$$

C'est le moment de ressortir notre ouvert préféré [132] autour de x_0 :

$$V(x_0,\varphi_1,\dots,\varphi_n,\epsilon)=\bigcap_{i=1}^{n}\varphi_i^{-1}\Big(\{\|z\in\mathbb{C}\text{ tel que }|z-\varphi_i(x_0)|\|<\epsilon\}\Big). \tag{27.979}$$

Supposons que $x\in V(x_0,\varphi_1,\dots,\varphi_n,\epsilon)$. Il vérifie $|\varphi_j(x)-\varphi_j(x_0)|\leqslant\epsilon$ et donc

$$\|x-x_0\|=\|\sum_{j=1}^{n}\varphi_j(x-x_0)e_j\| \tag{27.980a}$$
$$\leqslant\sum_j|\varphi_j(x-x_0)|\|e_j\| \tag{27.980b}$$
$$=\sum_j|\varphi_j(x-x_0)| \tag{27.980c}$$
$$\leqslant\sum_j|\varphi_j(x)-\varphi_j(x_0)| \qquad \leqslant n\epsilon. \tag{27.980d}$$

En choisissant $\epsilon<\frac{r}{n}$, nous avons $x\in B(x_0,r)$ et donc

$$V(x_0,\dots,)\subset B(x_0,r). \tag{27.981}$$

□
[LEMooMCYAooGMzbbs]

Lemme 27.223.
Soit un espace de Banach E sur le corps $\mathbb{F}$. Nous notons $\mathrm{Fun}(E,\mathbb{F})$ l'ensemble de toutes les applications de E vers $\mathbb{F}$. Pour chaque $x\in E$ nous considérons l'application

$$\begin{aligned}f_x\colon \mathrm{Fun}(E,\mathbb{F})&\to\mathbb{F}\\ \omega&\mapsto\omega(x).\end{aligned} \tag{27.982}$$

132. Nous avons un . . .faible pour lui !

Nous considérons sur $\mathrm{Fun}(E,\mathbb{F})$ *la plus petite topologie telle que tous les* f_x *soient continues.*

La suite (ω_n) *dans* $\mathrm{Fun}(E,\mathbb{F})$ *converge vers* $\omega\in\mathrm{Fun}(E,\mathbb{F})$ *si et seulement si* $\omega_n(x)\to\omega(x)$ *pour tout* $x\in E$.

LEMooWXBVooSjafZr

Lemme 27.224.
Soit un espace de Banach E *sur le corps* $\mathbb{F}$ *(*$=\mathbb{R}$ *ou* $\mathbb{C}$*). Nous considérons l'ensemble* $\mathrm{Fun}(E,\mathbb{F})$ *sur lequel nous mettons la topologie minimale qui rend continue les applications*

$$\begin{aligned} f_x\colon \mathrm{Fun}(E,\mathbb{F})&\to\mathbb{F}\\ \omega&\mapsto\omega(x). \end{aligned} \tag{27.983}$$

Nous notons $I=\{f_x \text{ tel que } x\in E\}$ *et* τ_I *la topologie de* $\mathrm{Fun}(E,\mathbb{F})$.

Si (ω_n) *est une suite dans* $\mathrm{Fun}(E,\mathbb{F})$ *nous avons* $\omega_n\xrightarrow{\tau_I}\omega$ *si et seulement si* $\omega_n(x)\to\omega(x)$ *pour tout* $x\in E$.

THOooRECTooEVLHSq

Théorème 27.225 (Banach-Alaogly-Bourbaki[195])**.**
Soit un espace de Banach E. *La boule unité*

$$B=\{\varphi\in E' \text{ tel que } \|\varphi\|=1\} \tag{27.984}$$

est compacte pour la topologie faible.

Démonstration. Nous considérons à nouveau $\mathrm{Fun}(E,\mathbb{F})$ muni de la topologie du lemme 27.224. Nous considérons l'inclusion

$$\begin{aligned} \Phi\colon \big(E',\|.\|\big)&\to\big(\mathrm{Fun}(E,\mathbb{F}),\tau_I\big)\\ \omega&\mapsto\omega. \end{aligned} \tag{27.985}$$

En particulier nous notons $S=\Phi(E')$ et nous allons prouver que $\Phi\colon E'\to S$ bijective et continue et que $\Phi^{-1}\colon S\to E'$ est également continue. □

PROPooYARHooOpmztY

Proposition 27.226 ([195])**.**
Un convexe dans un espace de Banach est fermé si et seulement si il est faiblement fermé.

27.19.1 Espace de Banach réflexif

PROPooPVVYooMZjQSq

Proposition 27.227 ([195])**.**
Une suite bornée dans un espace de Banach réflexif[133] *contient une sous-suite faiblement convergente.*

THOooTFIHooPQjVAr

Théorème 27.228 (Kakutami[195])**.**
Un espace de Banach est réflexif si et seulement si la boule fermée $\overline{B(0,1)}$ *est compacte pour la topologie faible*[134].

PROPooBBNBooGcXDRH

Proposition 27.229 ([195])**.**
Un espace de Banach est réflexif si et seulement si son dual est réflexif.

27.19.2 Espaces L^∞

LEMooMSYAooGEMgoc

Lemme 27.230 ([195])**.**
Si $(\Omega,\mathcal{A},\mu)$ *est un espace mesuré* σ*-fini, l'espace* $L^1(\Omega,\mathcal{A},\mu)$ *n'est pas réflexif.*

LEMooUSXTooFvpsVd

Lemme 27.231 ([195])**.**
Si $(\Omega,\mathcal{A},\mu)$ *est un espace mesuré* σ*-fini, alors* $L^\infty(\Omega,\mathcal{A},\mu)$ *n'est pas réflexif*[135].

133. Définition 27.181.
134. Définition 27.216.
135. Définition 27.181.

Démonstration. Supposons que $L^\infty(\Omega,\mathcal{A},\mu)$ est réflexif. Le théorème de représentation de Riesz 27.186(2) dit que $(L^1)'$ est en bijection linéaire isométrique avec L^∞ ; le lemme 27.183 dit alors que $(L^1)'$ est réflexif [136]. La proposition 27.229 dit alors que L^1 est réflexif.

Or le lemme 27.230 dit que L^1 n'est pas réflexif. □

La proposition suivante est souvent présentée en disant que l'inclusion $L^1 \subset (L^\infty)'$ est stricte, ou, pire, en disant que $(L^\infty)'$ est strictement plus grand que L^1. Cette façon de dire est un gros abus de langage. D'abord L^1 n'est même pas inclus dans L^∞ ; ce sont deux ensembles qui n'ont rien à voir. Ensuite, ce que signifie réellement cette proposition est seulement que la première injection $L^1 \to (L^\infty)'$ qui nous tombe sous la main (celle du théorème de représentation de Riesz) n'est pas surjective [137].

PROPooXXRQooNSBZOi

Proposition 27.232.
Soit $g \in L^1(\mathbb{R}^d)$. L'application

$$\begin{aligned}\Phi_g\colon L^\infty(\mathbb{R}^d) &\to \mathbb{C}\\ f &\mapsto \int_{\mathbb{R}^d} f\bar{g}\end{aligned} \tag{27.986}$$

est linéaire et bien définie.

L'application

$$\begin{aligned}\Phi\colon L^1(\mathbb{R}^d) &\to L^\infty(\mathbb{R}^d)'\\ g &\mapsto \Phi_g\end{aligned} \tag{27.987}$$

n'est pas surjective.

Démonstration. Prouvons d'abord que Φ_g est bien définie. Vu que $f \in L^\infty$, il existe $M \in \mathbb{R}$ tel que $|f(x)| < M$ sur $\mathbb{R}^d\backslash A$ où A est de mesure nulle dans $\mathbb{R}^d$. La fonction $x \mapsto |f(x)\overline{g(x)}|$ est donc majorée par la fonction intégrable $x \mapsto M|g(x)|$ qui est intégrable (sur $\mathbb{R}^d\backslash A$). L'intégrabilité de $f\bar{g}$ n'est donc pas un problème.

Le fait que Φ prenne ses valeurs dans $(L^\infty)'$ est le théorème 27.187(2b).

Le fait qu'elle ne soit pas surjective est la proposition 27.191. □

27.20 Espace de Schwartz

Pour un multiindice $\alpha = (\alpha_1,\dots,\alpha_d) \in \mathbb{N}^d$, nous notons

$$\partial^\alpha\varphi = \partial_{x_1}^{\alpha_1}\dots\partial_{x_d}^{\alpha_d}\varphi \tag{27.988}$$

pour peu que la fonction φ soit $|\alpha| = \alpha_1 + \cdots + \alpha_d$ fois dérivable.

DefHHyQooK

Définition 27.233.
*Soit $\Omega \subset \mathbb{R}^d$. L'**espace de Schwartz** $\mathscr{S}(\Omega)$ est le sous-ensemble de $C^\infty(\Omega)$ des fonctions dont toutes les dérivées décroissent plus vite que tout polynôme :*

$$\mathscr{S}(\Omega) = \big\{\varphi \in C^\infty(\Omega) \text{ tel que } \forall \alpha,\beta \in \mathbb{N}^d, p_{\alpha,\beta}(\varphi) < \infty\big\} \tag{27.989}$$

où nous avons considéré

$$p_{\alpha,\beta}(\varphi) = \sup_{x\in\Omega} |x^\beta(\partial^\alpha\varphi)(x)| = \|x^\beta\partial^\alpha\varphi\|_\infty. \tag{27.990}$$

EqQWdChCu

Pour simplifier les notations (surtout du côté de Fourier), nous allons parfois écrire $M_i\varphi$ pour la fonction $x \mapsto x_i\varphi(x)$.

136. Dans de nombreuses références, par exemple[195], il est simplement dit que $(L^1)' = L^\infty$. C'est un abus de notation qui permet de se passer du lemme 27.183.

137. Si vous savez comment prouver qu'il n'existe pas de surject de L^1 vers $(L^\infty)'$, écrivez-moi.

Exemple 27.234.
La fonction e^{-x^2} est une fonction à décroissance rapide sur $\mathbb{R}$. △

Définition 27.235.
*Une fonction $f\colon \mathbb{R}^d \to \mathbb{C}$ est dite à **décroissance rapide** si elle décroît plus vite que n'importe quel polynôme. Plus précisément, si pour tout polynôme Q, il existe un $r > 0$ tel que $|f(x)| < \frac{1}{|Q(x)|}$ pour tout $\|x\| \geqslant r$.*

PropCSmzwGv

Proposition 27.236.
Une fonction Schwartz est à décroissance rapide.

Démonstration. Nous commençons par considérer un polynôme P donné par

$$P(x) = \sum_k c_k x^{\beta_k} \tag{27.991}$$

où les β_k sont des multiindices, les c_k sont des constantes et la somme est finie. Nous avons la majoration

$$\sup_{x\in\mathbb{R}^d} |\varphi(x)P(x)| \leqslant \sum_k \sup_x \left|c_k\varphi(x)x^{\beta_k}\right| \leqslant \sum_k |c_k| p_{0,\beta_k}(\varphi) < \infty. \tag{27.992}$$

Nous allons noter M_P la constante $\sum_k |c_k| p_{0,\beta_k}(\varphi)$, de sorte que pour tout $x \in \mathbb{R}^d$ nous ayons $|\varphi(x)P(x)| \leqslant M_P$ et donc

$$|\varphi(x)| \leqslant \frac{M_P}{|P(x)|} = \frac{1}{|\frac{1}{M_P}P(x)|}. \tag{27.993}$$

Notons que cette inégalité est a fortiori correcte pour les x sur lesquels P s'annule.

Soit maintenant un polynôme Q. Nous considérons le polynôme $P(x) = \|x\|Q(x)$. Étant de plus haut degré, pour toute constante C il existe un rayon r_C tel que $|P(x)| \geqslant C|Q(x)|$ pour tout $|x| \geqslant r_C$. En particulier pour $|x| \geqslant r_{M_P}$ nous avons

$$|P(x)| \geqslant M_P|Q(x)| \tag{27.994}$$

et donc, pour ces x,

$$|\varphi(x)| \leqslant \frac{1}{|\frac{1}{M_P}P(x)|} \leqslant \frac{1}{|Q(x)|}. \tag{27.995}$$

La première inégalité est valable pour tout x, et la seconde pour $\|x\| \geqslant r_{M_P}$. □

CORooZFPSooHCFUSH

Corolaire 27.237 ([1])**.**
Soit φ une fonction Schwartz sur $\mathbb{R}^m \times \mathbb{R}^n$. Alors la fonction

$$y \mapsto \sup_{x\in\mathbb{R}^n} |\varphi(x,y)| \tag{27.996}$$

est intégrable.

Démonstration. Soit un polynôme Q en la variable y. Par la proposition 27.236, il existe $r > 0$ tel que

$$|\varphi(x,y)| < \frac{1}{Q(y)} \tag{27.997}$$

pour tout $\|(x,y)\| > r$. A fortiori l'inégalité tient pour tout $|y| > r$. Donc

$$\int_{\mathbb{R}^m} \sup_{x\in\mathbb{R}^n} |\varphi(x,y)|dy = \int_{\|y\|\leqslant r} \sup_x |\varphi(x,y)|dy + \int_{\|y\|>r} \sup_x |\varphi(x,y)|dy. \tag{27.998}$$

La première intégrale est bornée par $\mathrm{Vol}\big(B(0,r)\big)\|\varphi\|_\infty$ tandis que la seconde est bornée par l'intégrale de $\frac{1}{Q(y)}$. En prenant Q de degré suffisamment élevé en toutes les composantes de y nous avons intégrabilité. □

27.20.1 Topologie

LEMDEFooZEFVooMMmiBr

Lemme-Définition 27.238.
Les $p_{\alpha,\beta}$ donnés par l'équation (27.990) ci-dessus sont des seminormes[138]*. La topologie considérée sur $\mathscr{S}(\mathbb{R}^d)$ est celle des seminormes $p_{\alpha,\beta}$.*

NORMooVQESooRwJShl

27.239.
Nous avons un enchainement de résultats qui nous aident à prouver la continuité d'une application $T\colon \mathscr{S}(\mathbb{R}^d) \to X$.

(1) La topologie de $\mathscr{S}(\mathbb{R}^d)$ est donnée par une famille dénombrable de seminormes. Donc la proposition 7.350 nous dit que $\mathscr{S}(\mathbb{R}^d)$ est métrisable.

(2) La proposition 7.297 nous dit alors que si X est métrique, toute application séquentiellement continue $T\colon \mathscr{S}(\mathbb{R}^d) \to X$ est continue.

(3) Donc si X est métrique, il suffit de prouver que pour $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ nous avons $T(f_n) \xrightarrow{X} 0$ où $f_n\colon \mathscr{S}(\mathbb{R}^d) \to X$. Dans les cas usuels, T sera une distribution et $X = \mathbb{C}$.

(4) En vertu de la proposition 7.344, la convergence $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ signifie que pour tout choix de multiindice α et β, $p_{\alpha,\beta}(f_n) \to 0$, c'est-à-dire

$$\| x^\beta \partial^\alpha f_n \|_\infty \to 0. \tag{27.999}$$

EQooPUJPooNbtNFh

(5) Et enfin, la technique pour montrer que $T\colon \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ est continue est de montrer que sous l'hypothèse d'avoir (27.999) pour tout choix de α et β, nous avons $T(f_n) \to 0$ dans $\mathbb{C}$.

LemRJhCbkO

Lemme 27.240 ([704])**.**
La topologie sur $\mathscr{S}(\mathbb{R}^d)$ est donnée aussi par les seminormes

$$q_{n,m} = \max_{|\alpha|\leqslant n} \sup_{x\in\mathbb{R}^d} \big(1 + \|x\|\big)^m \big|\partial^\alpha \varphi(x)\big|. \tag{27.1000}$$

Autrement dit, une suite $\varphi_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ si et seulement si $q_{n,m}(\varphi) \to 0$ pour tout n et m.

Le fait que les $q_{n,m}(\varphi)$ restent bornés est la proposition 27.236. Cependant ce lemme est plus précis parce qu'en disant seulement que φ est majoré par des polynôme, nous ne disons pas que les polynômes correspondants aux φ_n tendent vers zéro si $\varphi_n \xrightarrow{\mathscr{S}} 0$. Et d'ailleurs on ne sait pas très bien ce que signifierait $P_n \to 0$ pour une suite de polynômes.

PropGNXBeME

Proposition 27.241.
Pour $p \in [1,\infty]$, l'espace $\mathscr{S}(\mathbb{R}^d)$ s'injecte continument dans $L^p(\mathbb{R}^d)$.

Démonstration. L'injection dont nous parlons est l'identité ou plus précisément l'identité suivie de la prise de classe. Il faut vérifier que cela est correct et continu, c'est-à-dire d'abord qu'une fonction à décroissance rapide est bien dans L^p et ensuite que si $f_n \xrightarrow{\mathscr{S}} 0$, alors $f_n \xrightarrow{L^p} 0$.

Commençons par $p = \infty$. Alors $\|f_n\|_\infty = p_{0,0}(f_n) \to 0$ parce que si $f_n \xrightarrow{\mathscr{S}} 0$, alors en particulier $p_{0,0}(f_n) \to 0$.

Au tour de $p < \infty$ maintenant. Nous savons qu'en dimension d, la fonction

$$x \mapsto \frac{1}{(1 + \|x\|)^s} \tag{27.1001}$$

est intégrable dès que $s > d$. Pour toute valeur de m nous avons

$$\|\varphi\|_p^p = \int_{\mathbb{R}^d} |\varphi(x)|^p dx = \int_{\mathbb{R}^d} \frac{\big|(1 + \|x\|)^m \varphi(x)\big|^p}{\big(1 + \|x\|\big)^{mp}} \leqslant \int_{\mathbb{R}^d} \frac{q_{0,m}(\varphi)^p}{\big(1 + \|x\|\big)^{mp}}. \tag{27.1002}$$

138. Définition 7.326.

En choisissant m de telle sorte que $mp > d$, nous avons convergence de l'intégrale et donc $\|\varphi\|_p < \infty$. Nous retenons que

$$\|\varphi\|_p^p \leqslant C q_{0,m}(\varphi)^p \tag{27.1003}$$

pour une certaine constante C et un bon choix de m.

Ceci prouve que $\mathscr{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$. Nous devons encore vérifier que l'inclusion est continue. Si $\varphi_n \xrightarrow{\mathscr{S}} 0$, alors en particulier nous avons $q_{0,m}(\varphi_n) \to 0$ par le lemme 27.240. Par conséquent la majoration (27.1003) nous dit que $\|\varphi_n\|_p \to 0$ également.

□

En résumé, si $\varphi_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} \varphi$ alors $\varphi_n \xrightarrow{L^p} \varphi$.

Théorème 27.242 ([468]).
Soit μ une mesure sur les boréliens de $\mathbb{R}^n$ finie sur les compacts. Alors $\mathscr{D}(\mathbb{R}^n)$ est dense dans $L^1(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \mu)$.

Proposition 27.243 ([705]).
La partie $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}(\mathbb{R}^d)$.

Démonstration. Soit $f \in \mathscr{S}(\mathbb{R}^d)$, et ϕ, une fonction de $\mathscr{D}(\mathbb{R}^d)$ telle que $\phi(x) = 1$ pour $|x| \leqslant 1$ (l'existence de telles fonctions est discutée en 15.14.1). Soit aussi $\phi_k(x) = \phi(x/k)$. Nous posons

$$f_k(x) = \phi_k(x) f(x), \tag{27.1004}$$

et nous allons prouver que pour tout multiindice α et γ,

$$p_{\alpha,\gamma}(f_k - f) = \|x^\gamma \partial^\alpha (f_k - f)\|_\infty \to 0. \tag{27.1005}$$

Pour cela nous allons noter $\beta \leqslant \alpha$ lorsque β est un multiindice contenu dans α. En utilisant la dérivée du produit nous avons

$$\begin{aligned}
(\partial^\alpha f_k)(x) &= \sum_{\beta \leqslant \alpha} (\partial^{\alpha-\beta}\phi_k)(x) \partial^\beta f(x) && \text{(27.1006a)}\\
&= \sum_{\beta \leqslant \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) && \text{(27.1006b)}\\
&= \sum_{\beta < \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) + \phi(x/k)(\partial^\alpha f)(x). && \text{(27.1006c)}
\end{aligned}$$

Nous devons donc étudier et majorer

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d} |x^\gamma \partial^\alpha (f_k - f)| \leqslant \sup \Big| x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) \Big| \\
+ \sup \big| x^\gamma \big(\phi(x/k) - 1\big)(\partial^\alpha f)(x) \big|
\end{aligned} \tag{27.1007}$$

En ce qui concerne le second terme, soit $\epsilon > 0$, vu que f est Schwartz, il existe R tel que

$$|x^\gamma (\partial^\alpha f)(x)| < \epsilon \tag{27.1008}$$

dès que $\|x\| > R$. En prenant $k > R$,

$$|x^\gamma (\partial^\alpha f)(x)| \begin{cases} = 0 & \text{si } \|x\| < R \\ \leqslant \epsilon & \text{si } \|x\| > R. \end{cases} \tag{27.1009}$$

En ce qui concerne le premier terme,

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d} \Big| x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) \Big| && \text{(27.1010a)}\\
&\leqslant \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k)(x^\gamma \partial^\beta f)(x) \Big| && \text{(27.1010b)}\\
&= \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k) p_{\beta,\gamma}(f) \Big| && \text{(27.1010c)}
\end{aligned}$$

La somme ne contient qu'un nombre fini de β différents, donc nous pouvons considérer un nombre K qui majore tous les $p_{\beta,\gamma}(f)$ en même temps. La partie avec ϕ peut être majorée par $\|\partial^{\alpha-\beta}\phi\|_\infty$ (qui est fini) dont nous pouvons prendre le maximum sur $\beta<\alpha$. Toute l'expression dans la somme est donc majorée par un nombre qui ne dépend ni de x ni de β. Vu que la somme est finie, elle est majorée par ce nombre multiplié par le nombre de termes dans la somme et au final

$$\sup_{x\in\mathbb{R}^d}\left|x^\gamma\sum_{\beta<\alpha}k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\right|\leqslant\frac{K'}{k}. \tag{27.1011}$$

La limite $k\to\infty$ ne fait alors plus de doutes. □

Remarque 27.244.

Vu la topologie de $\mathscr{S}(\mathbb{R}^d)$ (définition 27.238), la convergence $f_k\xrightarrow{\mathscr{S}(\mathbb{R}^d)}f$ peut être exprimée par le fait que pour tout k,l,

$$t^k f_n^{(l)}\xrightarrow{unif}t^k f^{(l)}. \tag{27.1012}$$

C'est-à-dire convergence uniforme de toutes les dérivées multipliées par n'importe quel polynôme.

27.20.2 Produit de convolution

PROPooUNFYooYdbSbJ

Proposition 27.245 (Stabilité de Schwartz par convolution [139] [678]).
Si $\varphi\in L^1(\mathbb{R}^d)$ et $\psi\in\mathscr{S}(\mathbb{R}^d)$, alors $\varphi*\psi\in\mathscr{S}(\mathbb{R}^d)$.

Démonstration. Nous devons prouver que

$$p_{\alpha,\beta}(\varphi*\psi)=\sup_{x\in\mathbb{R}^d}|x^\beta(\partial^\alpha(\varphi*\psi))(x)| \tag{27.1013}$$

est borné pour tout multiindices α et β. En appliquant $|\alpha|$ fois la proposition 27.71, nous mettons toutes les dérivées sur ψ : $\partial^\alpha(\varphi*\psi)=(\varphi*\partial^\alpha\psi)$. Cela étant fait, nous majorons

$$\left|x^\beta(\varphi*\partial^\alpha\psi)(x)\right|\leqslant|x^\beta|\int_{\mathbb{R}^d}|\varphi(y)|\underbrace{\left|(\partial^\alpha\psi)(x-y)\right|}_{\leqslant\|\partial^\alpha\psi\|_\infty}dy\Big| \tag{27.1014a}$$

$$\leqslant|x^\beta|\|\partial^\alpha\psi\|_\infty\int_{\mathbb{R}^d}|\varphi(y)|dy \tag{27.1014b}$$

$$\leqslant p_{\alpha,\beta}(\psi)\|\varphi\|_{L^1}. \tag{27.1014c}$$

Par conséquent, $p_{\alpha,\beta}(\varphi*\psi)\leqslant\|\varphi\|_{L^1}p_{\alpha,\beta}(\psi)<\infty$. □

27.21 Théorème de Montel

ThoXLyCzol

Théorème 27.246 (Montel[104]).
Soient Ω un ouvert de $\mathbb{C}$ et $\mathcal{F}$ une famille de fonctions holomorphes sur Ω, uniformément bornée sur tout compact de Ω. Alors de toute suite dans $\mathcal{F}$ nous pouvons extraire une sous-suite convergeant uniformément sur tout compact de Ω.

Démonstration. (1) **Un ensemble équicontinu**

Nous commençons par prendre une suite de compacts dans Ω comme dans le lemme 7.306, et une suite δ_n de réels strictement positifs tels que

$$B(z,2\delta_n)\subset K_{n+1} \tag{27.1015}$$

139. Définition 27.65.

pour tout $z \in K_n$. Soient $x, y \in K_n$ tels que $|x-y| < \delta_n$; nous notons $\partial B(x, 2\delta_n)$ le cercle de rayon $2\delta_n$ autour de x, parcouru dans le sens positif. La formule de Cauchy 26.207 nous donne

$$f(x) - f(y) = \frac{1}{2\pi i} \int_{\partial B} \left(\frac{f(\xi)}{\xi - x} - \frac{f(\xi)}{\xi - y} \right) d\xi = \frac{x-y}{2\pi i} \int_{\partial B} \frac{f(\xi)}{(\xi - x)(\xi - y)} d\xi \tag{27.1016}$$

Nous majorons ça par

$$\big|f(x) - f(y)\big| \leqslant \frac{|x-y|}{2\pi} \int_{\partial B} \frac{|f(\xi)|}{2\delta_n^2} d\xi \leqslant \frac{|x-y|}{\delta_n} M_n. \tag{27.1017}$$

Justifications :

— $|\xi - x| = 2\delta_n$ et $|\xi - y| \geqslant \delta_n$ parce que ξ est au mieux sur le rayon passant par x et y.
— $|f(\xi)| \leqslant M_n$ où M_n est la borne uniforme de $\mathcal{F}$ sur le compact K_n.
— Nous avons aussi fini par calculer l'intégrale dans laquelle il ne restait plus rien, ça a donné la circonférence du cercle de rayon $2\delta_n$.

Jusqu'à présent nous avons prouvé que l'ensemble

$$\mathcal{F}_n = \{f|_{K_n} \text{ tel que } f \in \mathcal{F}\} \tag{27.1018}$$

est équicontinu. Il est aussi équiborné par hypothèse.

(2) **Application du théorème d'Ascoli**
L'ensemble $\mathcal{F}_n$ vérifie les hypothèses du théorème d'Ascoli 27.4. Donc l'ensemble $\mathcal{F}_n$ est relativement compact dans $C(K_n, \mathbb{C})$ pour la norme uniforme. Autrement dit l'ensemble $\bar{\mathcal{F}}$ est compact et si nous avons une suite de fonctions dans $\mathcal{F}_n$, il existe une sous-suite convergeant dans $\bar{\mathcal{F}}_n$, c'est-à-dire uniformément. Autrement dit il existe une fonction strictement croissante $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que la suite $k \mapsto f_{\varphi(k)}$ converge uniformément sur K_n. La limite n'est cependant pas spécialement dans $\mathcal{F}_n$.

(3) **L'argument diagonal**
La suite $k \mapsto f_{\varphi_1 \circ \ldots \varphi_k(k)}$ converge uniformément sur tous les K_n. Si K est un compact de Ω, alors les petites propriétés sympas du lemme 7.306 nous disent que $K \subset \text{Int}(K_m)$ pour un certain m. Ladite suite convergeant uniformément sur K_m, elle converge uniformément sur K et nous avons montré la convergence uniforme sur tout compact de Ω.

□

Corolaire 27.247 ([104]).
Soient Ω un ouvert connexe borné de $\mathbb{C}$ et $a \in \Omega$. Soit f holomorphe sur Ω telle que $f(a) = a$ et $|f'(a)| < 1$.

Alors de (f^n) on peut extraire une sous-suite convergeant uniformément sur tout compact de Ω vers la fonction constante a.

Démonstration. Nous considérons un voisinage de a inclus dans Ω ; sachant que $|f(a)| < 1$, nous trouvons un voisinage encore plus petit de a sur lequel $|f'(z)| < 1$. Soit donc r tel que $\overline{B(a,r)} \subset \Omega$ et tel que $|f'(z)| < 1$ sur $\overline{B(a,r)}$. Étant donné que $f'(z)$ est continue sur le compact $\overline{B(a,r)}$, nous en prenons le maximum λ (qui est strictement inférieur à 1) et nous avons au final

$$|f'(z)| \leqslant \lambda < 1 \tag{27.1019}$$

pour tout $z \in \overline{B(a,r)}$. Le théorème des accroissements finis 12.314 nous dit que

$$\big|f(z) - a\big| \leqslant \lambda |z - a| \tag{27.1020}$$

pour tout $z \in \overline{B(a,r)}$. C'est ici que nous utilisons l'hypothèse de convexité de Ω. Nous montrons alors par récurrence que

$$\big|f^n(z) - a\big| \leqslant \lambda^n |z - a| \leqslant \lambda^n r \leqslant r. \tag{27.1021}$$

EqJQUzKpg

L'ensemble $A = \{f^n \text{ tel que } n \geqslant 1\}$ est donc uniformément borné sur $\overline{B(a,r)}$ par $a + r$. Autre manière de le dire : pour tout $z \in \overline{B(a,r)}$ nous avons

$$f^n(z) \in \overline{B(a,r)}. \tag{27.1022}$$

La suite (f^n) est donc uniformément bornée sur tout compact de $B(a,r)$. Le théorème de Montel 27.246 nous indique que l'on peut extraire une sous-suite convergente uniformément sur tout compact. Au vu de (27.1021) cette convergence ne peut avoir lieu que vers une fonction g qui vaut la constante a sur $B(a,r)$.

D'autre part la fonction g est holomorphe en tant que limite uniforme de fonctions holomorphes, théorème 26.106. Or une fonction holomorphe constante sur un ouvert est constante sur tout son domaine d'holomorphie (principe d'extension analytique, théorème 17.138). □

27.22 Espaces de Bergman

Source : [496].

Soit Ω un borné dans $\mathbb{C}$ et D le disque unité ouvert de $\mathbb{C}$.

Définition 27.248.
*L'**espace de Bergman** sur Ω, noté $A^2(\Omega)$ est l'espace des fonctions holomorphes sur Ω qui sont en même temps dans $L^2(\Omega)$.*

Nous mettons sur $A^2(\Omega)$ le produit scalaire usuel hérité de L^2 :

$$\langle f, g\rangle = \int_\Omega f(z)\overline{g(z)}dz. \tag{27.1023}$$

LemIZxKfB

Lemme 27.249.
Soient un compact $K \subset \Omega$ et une fonction $f \in A^2(\Omega)$. Alors

$$\max_{z\in K} |f(z)| \leqslant \frac{1}{\sqrt{\pi}} \frac{1}{d(K, \partial\Omega)} \|f\|_2. \tag{27.1024}$$

Démonstration. Soient $a \in \Omega$ et $r > 0$ tels que $B(a,r) \subset \Omega$. Nous considérons aussi $\rho \leqslant r$. La formule de Cauchy (26.207) nous donne

$$f(a) = \frac{1}{2\pi i} \int_{B(a,\rho)} \frac{f(\xi)}{\xi - a} d\xi = \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta}) d\theta \tag{27.1025}$$

où nous avons utilisé le chemin $\gamma(\theta) = a + \rho e^{i\theta}$, $\gamma'(\theta) = i\rho e^{i\theta}$ et $\rho = |\xi - a|$. Maintenant une astuce est d'écrire

$$\frac{r^2}{2} f(a) = \int_0^r f(a)\rho d\rho, \tag{27.1026}$$

et d'y substituer la valeur de $f(a)$ que nous venons de calculer :

$$\begin{aligned}
\frac{r^2}{2} f(a) &= \int_0^r \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta}) d\theta \rho d\rho && (27.1027a)\\
&= \frac{1}{2\pi} \int_{B(a,r)} f(z) dz && \text{passage aux polaires} \quad (27.1027b)\\
&= \frac{1}{2\pi} \langle 1, f\rangle_B && \text{produit scalaire sur } B(a,r) \quad (27.1027c)\\
&\leqslant \frac{1}{2\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B} && (27.1027d)
\end{aligned}$$

Nous avons donc

$$r^2 f(a) \leqslant \frac{1}{\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B}, \tag{27.1028}$$

et donc

$$\pi r^2 f(a) \leqslant \sqrt{\pi r^2}\|f\|_2, \tag{27.1029}$$

parce que $\langle f, f\rangle_B \leqslant \|f\|_2^2$. En effet le produit scalaire $\|.\|_2$ est donné par une intégrale sur Ω alors que $B(a,r) \subset \Omega$ et que la fonction qu'on y intègre est positive (c'est $|f(z)|^2$). En simplifiant,

$$f(a) \leqslant \frac{1}{\sqrt{\pi} r}\|f\|_2. \tag{27.1030}$$

Mais r a été choisi pour avoir $B(a,r) \subset \Omega$, donc $r \leqslant d(a, \partial\Omega)$ et

$$|f(a)| \leqslant \frac{1}{d(a,\partial\Omega)\sqrt{\pi}}\|f\|_2. \tag{27.1031}$$

Maintenant si nous prenons $a \in K$, nous avons encore la minoration $d(a,\partial K) \leqslant d(a,\partial\Omega)$ et donc

$$|f(a)| \leqslant \frac{1}{d(a,\partial K)\sqrt{\pi}}\|f\|_2. \tag{27.1032}$$

□

Théorème 27.250.
Soit Ω un ouvert de $\mathbb{C}$.

(1) L'espace $A^2(\Omega)$ est un espace de Hilbert.

(2) Si D est la boule unité dans $\mathbb{C}$, une base hilbertienne de $A^2(D)$ est donnée par les fonctions

$$e_n(z) = \sqrt{\frac{n+1}{\pi}} z^n \tag{27.1033}$$

pour $n \geqslant 0$.

Démonstration. Nous commençons par montrer que $A^2(\Omega)$ est complet. Pour cela nous considérons une suite de Cauchy (f_n) dans $A^2(\Omega)$ et un compact $K \subset \Omega$. Nous savons par le lemme 27.249 que

$$\max_{z\in K} |f_n(z) - f_m(z)| \leqslant \frac{1}{\sqrt{\pi} d(K,\partial\Omega)}\|f_n - f_m\|_2. \tag{27.1034}$$

Donc f_n converge uniformément sur K. Par le théorème de Weierstrass 26.106, la fonction f est holomorphe. Il existe donc une fonction holomorphe f qui est limite uniforme sur tout compact de Ω de la suite (f_n).

Mais $L^2(\Omega)$ étant complet, la suite (f_n) a une limite $g \in L^2(\Omega)$. Ce que nous voudrions faire est prouver que $f = g$. Notons que tel quel, ce n'est pas vrai parce que f est une vraie fonction alors que g est une classe. Ce que nous enseigne la proposition 27.18 est qu'il existe une sous-suite (qu'on note (g_n)) qui converge vers g presque partout. Dans cette dernière phrase, g_n et g sont de vraies fonctions, des représentants des classes dans L^2.

Nous déduisons que $f = g$ presque partout (ici f et g sont les fonctions) parce que la sous-suite converge uniformément vers f en même temps que presque partout vers g. Donc $f = g$ dans $L^2(\Omega)$ (ici f et g sont les classes). Donc $f \in L^2(\Omega)$ et l'espace $A^2(\Omega)$ est de Hilbert.

Il nous faut encore prouver que $(e_n)_{n\geqslant 0}$ est une base orthonormale. En ce qui concerne les produits scalaires,

$$\langle e_m, e_n\rangle = \sqrt{\frac{(m+1)(n+1)}{\pi}} \int_D z^n \overline{z^m} dz \tag{27.1035a}$$

$$= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \int_0^1 \rho\, d\rho \int_0^{2\pi} d\theta \rho^{m+n} e^{i\theta(n-m)} \tag{27.1035b}$$

$$= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \frac{1}{m+n+2} \underbrace{\int_0^{2\pi} e^{i\theta(n-m)} d\theta}_{2\pi\delta_{mn}} \tag{27.1035c}$$

$$= \sqrt{\frac{(n+1)^2}{\pi^2}} \frac{1}{2n+2} 2\pi\delta_{nm} \tag{27.1035d}$$

$$= \delta_{nm}. \tag{27.1035e}$$

Donc les fonctions données sont bien orthonormales. Nous devons montrer qu'elles sont denses dans $A^2(D)$. Soit $f \in A^2(D)$ et $c_n(f) = \langle f, e_n \rangle$; nous allons montrer que

$$\|f\|_2^2 = \sum_{n=0}^{\infty} |\langle f, e_n \rangle|^2, \tag{27.1036}$$

parce que le point (5) du théorème 25.46 nous indique que ce sera suffisant pour avoir une base hilbertienne.

Étant donné que f est holomorphe sur D, le théorème 26.59 nous développe f en série entière :

$$f(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{27.1037}$$

En permutant la somme avec le produit scalaire,

$$c_n(f) = \int_D f(z)\bar{e}_n(z) = \sqrt{\frac{n+1}{\pi}} \int_D f(z)\bar{z}^n dz. \tag{27.1038}$$

Afin de profiter de la convergence uniforme de la série (27.1037) à l'intérieur de D, nous allons exprimer l'intégrale sur D comme une intégrale sur $|z| < r$ en faisant tendre r vers 1 (par le bas). Pour ce faire nous considérons les fonctions

$$g_k(z) = \begin{cases} f(z)\bar{z}^n & \text{si } |z| < 1 - 1/k \\ 0 & \text{sinon.} \end{cases} \tag{27.1039}$$

Ces fonctions sont intégrables sur D et dominées par $f(z)\bar{z}^n$ qui est intégrable sans dépendre de k. Mais nous avons évidemment $g_k(z) \to f(z)\bar{z}^n$. Le théorème de la convergence dominée permet alors de permuter l'intégrale et la limite $k \to \infty$. Cela nous permet d'écrire

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \int_{|z|<r} \bar{z}^n f(z) dz = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \int_{|z|<r} \sum_{k=0}^{\infty} a_k z^k \bar{z}^n. \tag{27.1040}$$

Par la convergence uniforme de la série entière *à l'intérieur* du disque D nous pouvons permuter l'intégrale et la somme (proposition 15.49) :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \sum_{k=0}^{\infty} a_k \int_{|z|<r} z^k \bar{z}^n dz. \tag{27.1041}$$

L'intégrale proprement dite est vite calculée et vaut

$$\int_{|z|<1} \bar{z}^n z^k dz = \frac{\pi r^{2n+2}}{n+1} \delta_{kn}. \tag{27.1042}$$

Nous pouvons donc continuer le calcul de $c_n(f)$ en effectuant la somme sur k qui se réduit à changer k en n puis en effectuant la limite :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r\to 1^-} \sum_k a_k \frac{\pi r^{2n+2}}{n+1} \delta_{kn} = \sqrt{\frac{\pi}{n+1}} a_n. \tag{27.1043}$$

Nous effectuons le même genre de calculs pour évaluer $\|f\|_2^2$:

$$\|f\|_2^2 = \int_D |f(z)|^2 dz \tag{27.1044a}$$

$$= \lim_{r\to 1^-} \int_{|z|<r} f(z) \sum_{k=0}^{\infty} \bar{a}_k \bar{z}_k dz \tag{27.1044b}$$

$$= \lim_{r\to 1^-} \sum_{k=0}^{\infty} \bar{a}_k \int_{|z|<r} f(z)\bar{z}^k dz \qquad \text{permuter } \sum \text{ et } \int \tag{27.1044c}$$

$$= \lim_{r\to 1^-} \sum_{k=0}^{\infty} \bar{a}_k a_k \frac{\pi r^{2k+2}}{k+1} \qquad \text{intégrale déjà faite.} \tag{27.1044d}$$

Mais nous savons déjà que $c_n(f) = \sqrt{\pi/(n+1)}$, donc ce qui est dans la somme est $\pi\bar{a}_k a_k/(n+1) = |c_k(f)|^2$. Nous avons donc

$$\|f\|_2^2 = \lim_{r\to 1^-} \sum_{k=0}^{\infty} |c_k(f)|^2 r^{2k+2}. \tag{27.1045}$$

La fonction (de r) constante $|c_k(f)|^2$ domine $|c_k(f)r^{2k+2}|$ tout en ayant une somme (sur k) qui converge ; en effet la proposition 25.25 nous indique que $\sum_k |c_k(f)|^2 \leqslant \|f\|_2^2$. Le théorème de la convergence dominée nous permet d'inverser la limite et la somme pour obtenir le résultat attendu :

$$\|f\|_2^2 = \sum_{k=0}^{\infty} |c_k(f)|^2. \tag{27.1046}$$

□

Chapitre 28

Séries de Fourier

28.1 Densité des polynômes trigonométriques

28.1.1 Convergence pour les fonctions continues (via Weierstrass)

Le résultat fondamental qui nous permet d'utiliser les polynômes trigonométriques comme base pour les fonctions *continues* périodiques est le suivant. Notons que pour les fonctions non continues, il y a encore du travail.

LemXGYaRlC

Lemme 28.1.
Si $f\colon \mathbb{R} \to \mathbb{C}$ est une fonction continue 2π-périodique et si $\epsilon > 0$, alors il existe un polynôme trigonométrique P tel que $\|f - P\|_\infty \leqslant \epsilon$.

Démonstration. Nous allons utiliser le théorème de Stone-Weierstrass 12.418. Soit le compact de Hausdorff [1]

$$S^1 = \{z \in \mathbb{C} \text{ tel que } |z| = 1\}, \tag{28.1}$$

et $C(S^1, \mathbb{C})$ l'algèbre des fonctions continues de S^1 vers $\mathbb{C}$. Il suffit de vérifier que les polynômes trigonométriques vérifient les hypothèses du théorème de Stone-Weierstrass. Un polynôme trigonométrique est un polynôme en z et $\bar{z}$ défini sur S^1.

(1) Le polynôme constant est dans l'algèbre, ok.

(2) Pour la séparation des points, considérons le polynôme trigonométrique $x \mapsto e^{ix}$.

(3) Si P est un polynôme en z et $\bar{z}$, alors $\bar{P}$ l'est aussi.

Donc si $\epsilon > 0$ et $\tilde{f} \in C(S^1, \mathbb{C})$ sont donnés, il existe un polynôme trigonométrique P tel que

$$\sum_t |\tilde{f}(e^{it}) - P(t)| < \epsilon. \tag{28.2}$$

Soit $f\colon \mathbb{R} \to \mathbb{C}$ une fonction continue 2π-périodique. Nous considérons $\tilde{f} \in C(S^1, \mathbb{C})$ donnée par $\tilde{f}(e^{it}) = f(t)$. Alors $\sup_t |f(t) - P(t)| \leqslant \epsilon$. □

28.1.2 Convergence pour les fonctions continues (via Fejér)

Si nous ne voulons pas passer par le gros théorème de Stone-Weierstrass pour prouver la densité des polynômes trigonométriques dans $\left(C^0_{2\pi}, \|.\|_\infty\right)$, nous pouvons passer par le gros théorème de Fejér. C'est ce que nous faisons maintenant.

Si vous vous intéressez seulement au théorème sur les séries de Fourier, vous cherchez probablement le théorème 28.17.

PROPooUOKAooGzGZWc

1. Définition 7.60.

Proposition-Définition 28.2.
Pour chaque $n \in \mathbb{N}$, la fonction

$$\begin{aligned} D_n : \mathbb{R} &\to \mathbb{C} \\ t &\mapsto \sum_{k=-n}^{n} e^{ikt} \end{aligned} \tag{28.3}$$

est bien définie. Elle s'appelle ***noyau de Dirichlet****.*

Définition 28.3.
Le ***noyau de Fejér*** *est la moyenne de Cesàro*[2] *des noyaux de Dirichlet :*

$$F_n(t) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(t). \tag{28.4}$$

LemHPoIkwu

Lemme 28.4.
Le noyau de Dirichlet s'exprime sous la forme

$$D_n(t) = \sum_{k=-n}^{n} e^{-ikt} = \frac{\sin\left(\frac{2n+1}{2}t\right)}{\sin(t/2)} \tag{28.5}$$

Démonstration. Nous commençons par mettre en facteur le premier terme :

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = e^{-int} \sum_{k=0}^{2n} e^{ikt}. \tag{28.6}$$

En utilisant la formule de la somme géométrique,

$$D_n(t) = e^{-int} \frac{1 - (e^{it})^{2n+1}}{1 - e^{it}} \tag{28.7a}$$

$$= e^{-int} \frac{1 - e^{(2n+1)it}}{1 - e^{it}} \tag{28.7b}$$

$$= e^{-int} \frac{e^{(2n+1)it/2}}{e^{i\frac{t}{2}}} \frac{e^{-(2n+1)it/2} - e^{(2n+1)it/2}}{e^{-it/2} - e^{it/2}} \tag{28.7c}$$

$$= \frac{(-2i)\sin\left(\frac{2n+1}{2}t\right)}{(-2i)\sin\left(\frac{t}{2}\right)}. \tag{28.7d}$$

□

Théorème 28.5 (Théorème de Dirichlet)**.**
Soit f une fonction 2π-périodique et C^1 par morceaux. Pour tout $x \in \mathbb{R}$ nous posons

$$s_n(x) = \sum_{k=-n}^{n} c_k(f) e^{ikx}. \tag{28.8}$$

Alors nous avons

$$\lim_{n\to\infty} s_n(x) = \frac{f(x^+) + f(x^-)}{2}. \tag{28.9}$$

LemtCAjJz

Lemme 28.6.
Le noyau de Fejér s'exprime sous la forme

$$F_n(t) = \frac{1}{n} \left(\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right)^2. \tag{28.10}$$

EqLOtzCf

2. Définition 10.141.

Note : ce noyau est positif. C'est important parce qu'on s'en sert dans la preuve du théorème de Fejér.

Démonstration. L'astuce est de noter $\sin(x) = \Im(e^{ix})$ et de repartir du résultat à propos du noyau de Dirichlet. En utilisant encore la formule de la série géométrique partielle [3],

$$F_n(t) = \frac{1}{n\sin(t/2)} \Im \sum_{k=0}^{n-1} e^{(2k+1)it/2} \tag{28.11a}$$

$$= \frac{1}{n\sin(t/2)} \Im e^{\frac{it}{2}} \sum_{k=0}^{n-1} e^{kit} \qquad \text{lem. 1.487} \tag{28.11b}$$

$$= \frac{1}{n\sin(t/2)} \Im e^{\frac{it}{2}} \left(\frac{1-e^{nit}}{1-e^{it}} \right) \tag{28.11c}$$

$$= \frac{1}{n\sin(t/2)} \Im e^{it/2} \frac{e^{\frac{nit}{2}} \left(e^{-\frac{int}{2}} - e^{\frac{nit}{2}} \right)}{e^{\frac{it}{2}} \left(e^{-it/2} - e^{it/2} \right)} \tag{28.11d}$$

$$= \frac{1}{n\sin(t/2)} \underbrace{\Im e^{\frac{nit}{2}}}_{\sin(nt/2)} \frac{\sin\left(\frac{nt}{2}\right)}{\sin\left(\frac{t}{2}\right)} \tag{28.11e}$$

$$= \frac{1}{n} \left(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}} \right)^2. \tag{28.11f}$$

□

Le théorème de Fejér donne la convergence au sens de Cesàro [4] de la série de Fourier dans le cas continu et périodique. Pour avoir une convergence plus forte que Cesàro, il faut plus d'hypothèses, comme le montre le contre-exemple de la proposition 28.21. Voir la discussion 28.22.

ThoJFqczow

Théorème 28.7 (Fejér)**.**
Soit $f : \mathbb{R} \to \mathbb{C}$ une fonction continue et 2π-périodique. Pour tout $k \in \mathbb{Z}$ nous notons

$$\begin{aligned} e_k : \mathbb{R} &\to \mathbb{C} \\ x &\mapsto e^{ikx}. \end{aligned} \tag{28.12}$$

Pour chaque $n \in \mathbb{N}$ nous posons

$$D_n = \sum_{k=-n}^{n} e_k \qquad S_n(f) = \sum_{k=-n}^{n} c_k(f) e_k \tag{28.13a}$$

$$F_n = \frac{D_0 + \cdots + D_{n-1}}{n} \qquad \tilde{F}_n = \sigma_n\big(S(f)\big) = \frac{1}{n} \sum_{k=0}^{n-1} S_k(f), \tag{28.13b}$$

où

$$c_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt. \tag{28.14}$$

Alors

(1) $\frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(t)dt = 1$.
(2) *Pour tout $\alpha \in]0, \pi[$, F_n converge uniformément vers 0 sur $[-\pi, \pi] \backslash [-\alpha, \alpha]$.*
(3) *La suite $\tilde{F}_n$ converge uniformément sur $\mathbb{R}$ vers f.*
ItemUNQSPmyiv
(4) *Le système trigonométrique $\{e_k\}_{k\in\mathbb{Z}}$ est total pour l'espace $\left(C^0_{2\pi}(\mathbb{R}), \|.\|_\infty\right)$ des fonctions continues 2π-périodiques.*

3. Proposition 10.135.
4. Convergence au sens de Cesàro, définition 10.141.

Démonstration. Un calcul usuel montre que

$$\int_{-\pi}^{\pi} e_l(t)dt = \begin{cases} 0 & \text{si } l \neq 0 \\ 2\pi & \text{si } l = 0 \end{cases} \tag{28.15}$$

Nous avons alors

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(t)dt = \frac{1}{2\pi}\frac{1}{n}\sum_{k=0}^{n-1}\sum_{l=-k}^{k}\underbrace{\int_{-\pi}^{\pi} e_l(t)dt}_{2\pi\delta_{l,0}} = \frac{1}{n}\sum_{k=0}^{n-1} 1 = 1. \tag{28.16}$$

Cela prouve déjà le premier point.

Pour le second point, en partant de l'expression (28.10) et en considérant $x \in [-\pi, \pi,]\backslash[-\alpha, \alpha]$ (ce qui nous évite l'annulation du dénominateur),

$$|F_n(x)| \leqslant \frac{1}{(n+1)\sin^2(\alpha/2)}, \tag{28.17}$$

et donc $F_n \to 0$ uniformément sur l'ensemble considéré.

Nous passons maintenant à cette histoire de convergence uniforme de la moyenne de Cesàro vers f. Pour tout $n \in \mathbb{N}$ nous avons

$$S_n(x) = \frac{1}{2\pi}\sum_{k=-n}^{n}\left(\int_{-\pi}^{\pi} f(t)e^{-ikt}dt\right)e^{ikx} \tag{28.18a}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{k=-n}^{n} e_k(x-t)dt \tag{28.18b}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_k(x-t)dt. \tag{28.18c}$$

Par conséquent, en effectuant le changement de variable $u = x - t$ et en utilisant la périodicité, [EqkDsyAc]

$$\tilde{F}_n(x) = \int_{-\pi}^{\pi} f(t)F_n(x-t)dt \tag{28.19a}$$

$$= -\int_{x+\pi}^{x-\pi} f(x-u)F_n(u)du \tag{28.19b}$$

$$= \int_{-\pi}^{\pi} f(x-u)F_n(u)du. \tag{28.19c}$$

Nous prouvons à présent l'uniforme convergence. Soit $\epsilon > 0$; étant donné que f est continue et 2π-périodique, elle est uniformément continue et nous considérons $\delta > 0$ tel que $|x-y| < \delta$ implique $\big|f(x) - f(y)\big| < \epsilon$. Soit M un majorant de $|f|$ sur $\mathbb{R}$. L'équation (28.19) nous donne

$$\big|f(x) - \tilde{F}_n(x)\big| = \left\|\frac{1}{2\pi}\int_{-\pi}^{\pi}\big(f(x-t) - f(x)\big)F_n(t)dt\right\| \tag{28.20a}$$ [ykuGGh]

$$\leqslant \frac{1}{2\pi}\int_{\delta\leqslant|t|\leqslant\pi}|2MF_n(t)|dt + \frac{1}{2\pi}\int_{-\delta}^{\delta}\epsilon|F_n(t)|dt \tag{28.20b}$$

$$\leqslant \frac{2M}{2\pi}\int_{\delta\leqslant|t|\leqslant\pi} F_n(t)dt + \epsilon' \tag{28.20c}$$ [uRAMyq]

Pour obtenir (28.20a) nous avons pu rentrer $f(x)$ dans l'intégrale en utilisant le premier point. Pour obtenir (28.20c) nous avons d'abord utilisé la positivité de F_n (lemme 28.6) pour enlever les valeurs absolues, et nous avons ensuite utilisé le fait que son intégrale valait 2π.

Étant donné que $F_n \to 0$ uniformément sur $[-\pi, \pi,]\backslash[-\alpha, \alpha]$, il existe un N tel que

$$\int_{\delta\leqslant|t|\leqslant\pi} F_n(t)dt \leqslant \epsilon \tag{28.21}$$

dès que $n > N$. Le résultat en découle.

Pour le point (4), il suffit de remarquer que chacun des $\tilde{F}_n$ est une combinaison finie d'éléments du système trigonométrique. □

28.1.3 Densité dans L^p

Nous venons de voir (de deux façons différentes) que les polynômes trigonométriques étaient dense dans $\left(C^0_{2\pi}(\mathbb{R}), \|.\|_\infty\right)$. Nous avons aussi déjà vu par le théorème 27.89 que ces polynômes trigonométriques étaient denses dans $L^p(S^1)$. Nous présentons à présent une autre façon de prouver cette dernière densité.

ThoDPTwimI

Théorème 28.8.
Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$.

Démonstration. Par les théorèmes 28.1 ou 28.7 (au choix), nous savons que les polynômes trigonométriques sont denses dans $\left(C^0_{2\pi}(S^1), \|.\|_\infty\right)$. Comme S^1 est compact, la densité est également au sens L^p. En effet si $\|f_n - f\|_\infty \leqslant \epsilon$, alors

$$\|f_n - f\|_p^p = \int_0^{2\pi} |f_n - f|^p \leqslant \int_0^{2\pi} \epsilon^p = 2\pi\epsilon^p. \tag{28.22}$$

Donc les polynômes trigonométriques sont denses dans $\left(C^0_{2\pi}(S^1), \|.\|_p\right)$. Mais nous savons par le théorème 27.51(5) que les fonctions continues sont denses dans $L^p(S^1)$.

Par composition de densités, les polynômes trigonométriques sont denses dans $L^p(S^1)$. □

28.1.4 Suite équirépartie, critère de Weyl

Définition 28.9.
Soit u une suite dans $[0,1]$. Pour $0 \leqslant a \leqslant b \leqslant 1$ nous posons

$$X_n(a,b) = \text{Card}\left\{k \in \{1,\dots,n\} \text{ tel que } u_k \in [a,b]\right\}. \tag{28.23}$$

*Nous disons que la suite u est **équirépartie** si pour tout $0 \leqslant a < b < 1$, on a*

$$\lim_{n\to\infty} \frac{X_n(a,b)}{n} = b - a. \tag{28.24}$$

Voir aussi la remarque 36.184 sur les nombres normaux.

PropDMvPDc

Proposition 28.10 (Critère de Weyl[496, 104])**.**
Soit (x_n) une suite dans $[0,1[$. Les conditions suivantes sont équivalentes.

ItemKWcZTHqi

(1) La suite (x_n) est équirépartie.

ItemKWcZTHqii

(2) Pour toute fonction continue à valeurs réelles sur $[0,1]$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(x_k) = \int_0^1 f(x)dx. \tag{28.25}$$

EqBSqdjpa

ItemKWcZTHqiii

(3) Pour tout $p \in \mathbb{N}\backslash\{0\}$ nous avons

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{2i\pi p x_k} = 0. \tag{28.26}$$

Démonstration. On pose

$$S_n(f) = \frac{1}{n} \sum_{k=1}^{n} f(x_k). \tag{28.27}$$

(1) **Une espèce de lemme**

Supposons connaitre un ensemble de fonctions A dense dans $C^0([0,1])$ pour toutes les fonctions desquelles nous avons la limite (28.25). Alors la limite a lieu pour toute fonction de $C^0([0,1])$. En effet, soit $f\in C^0([0,1])$ et $g\in A$ tel que $\|f-g\|_\infty<\epsilon$. Alors

$$\left\|\frac{1}{n}\sum_{k=1}^{n}f(x_k)-\int_0^1 f(t)dt\right\|\leqslant\left\|\frac{1}{n}\sum_{k=1}^{n}\big(f(x_k)-g(x_k)\big)\right\| \tag{28.28a}$$

$$+\left\|\frac{1}{n}\sum_{k=1}^{n}g(x_k)-\int_0^1 g(t)dt\right\| \tag{28.28b}$$

$$+\left\|\int_0^1 g(t)dt-\int_0^1 f(t)dt\right\|. \tag{28.28c}$$

Le premier terme est majoré par ϵ. Le troisième a la même majoration : $\int_0^1\big(f(t)-g(t)\big)dt\leqslant\|f-g\|_\infty=\epsilon$. Par hypothèse sur l'espace A, le second terme se majore par ϵ lorsque n est grand.

(2) **(1)⇒(2)** Nous supposons que la suite est équirépartie et nous commençons par montrer le résultat pour les fonctions en escalier. Soit donc la fonction en escalier $\eta(x)=c_j$ sur $a_{j-1}<x<a_j$. Sur le point a_j lui-même, la fonction η vaut soit c_j soit c_{j+1}. Nous avons

$$\frac{1}{n}\sum_{k=1}^{n}\eta(x_k)=\frac{1}{n}\left[\sum_{j=1}^{n}c_jX_n(a_j,a_{j+1})-\sum_{j=1}^{n}c_jX_n(a_j,a_j)+\sum_{j=1}^{n}\eta(a_j)X_n(a_j,a_j)\right]. \tag{28.29}$$

EgohMuel

À la limite $n\to\infty$, les deux derniers termes tombent [5] et il reste

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\eta(x_k)=\sum_{j=1}^{n}c_j(a_{j-1}-a_j). \tag{28.30}$$

Or par construction, pour une fonction en escalier,

$$\sum_{j=1}^{n}c_j(a_{j-1}-a_j)=\int_0^1\eta. \tag{28.31}$$

Étant donné que les fonctions en escalier sont denses dans les fonctions continues, l'espèce de lemme plus haut conclut.

(3) **(2)⇒(1)** Nous prouvons maintenant le sens inverse. C'est-à-dire que pour toute fonction continue sur $[0,1]$, nous avons

$$\int_0^1 f(x)dx=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}f(x_k). \tag{28.32}$$

Nous devons en déduire que (x_n) est équirépartie. Pour ce faire, soit $x\in[0,1[$ et $\epsilon>0$ tel que $x+\epsilon<1$. Nous considérons $\varphi=\mathbb{1}_{[x,1[}$ et

$$\varphi_\epsilon(t)=\begin{cases}0 & \text{si } t\in[0,x[\\ \frac{t-x}{\epsilon} & \text{si } t\in[x,x+\epsilon[\\ 1 & \text{si } t\geqslant x+\epsilon.\end{cases} \tag{28.33}$$

C'est une fonction continue, donc

$$\lim_{n\to\infty}S_n\big(\varphi_\epsilon(t)\big)=\int_0^1\varphi_\epsilon(t)dt=\int_x^{x+\epsilon}\frac{t-x}{\epsilon}dt+\int_{x+\epsilon}^1 1dt=1-x-\frac{\epsilon}{2}. \tag{28.34}$$

5. J'en profite pour mentionner que mon équation (28.29) n'est pas la même que celle de [496] dans laquelle il me semble voir une faute ; quoi qu'il en soit, les termes litigieux tombent.

Mais $\varphi_\epsilon \leqslant \varphi$, donc $S_n(\varphi_\epsilon) \leqslant S_n(\varphi)$ et donc

$$\liminf_{n\to\infty} S_n(\varphi) \geqslant 1 - x. \tag{28.35}$$

Notons que nous ne savons pas si la *vraie* limite de gauche existe ; c'est pourquoi nous prenons la limite inférieure, qui existe toujours.

Nous définissons aussi

$$\psi_\epsilon(t) = \begin{cases} 0 & \text{si } t \in [0, x-\epsilon[\\ \frac{t-x+\epsilon}{\epsilon} & \text{si } t \in [x-\epsilon, x[\\ 1 & \text{si } t > x. \end{cases} \tag{28.36}$$

C'est encore une fonction continue et nous trouvons [6]

$$\int_0^1 \psi_\epsilon(t)dt = 1 - x + \frac{\epsilon}{2}. \tag{28.37}$$

Puisque $\psi_\epsilon \geqslant \varphi$, nous avons $S_n(\psi_\epsilon) \geqslant S_n(\varphi)$ et donc

$$\limsup_n S_n(\varphi) \leqslant 1 - x. \tag{28.38}$$

Nous avons déjà obtenu que

$$1 - x \leqslant \liminf S_n(\varphi) \leqslant \limsup S_n(\varphi) \leqslant 1 - x, \tag{28.39}$$

donc la limite existe et vaut

$$\lim_{n\to\infty} S_n(\varphi) = 1 - x. \tag{28.40}$$

Le résultat est maintenant démontré dans le cas très particulier de la fonction caractéristique $\varphi = \mathbb{1}_{[x,1[}$.

Si nous prenons une fonction caractéristique $\mathbb{1}_{[a,b]}$, nous avons le même genre de preuve parce que $\mathbb{1}_{[a,b[}$ est une combinaisons linéaire de fonctions du type $\mathbb{1}_{[x,1[}$.

Nous avons donc

$$\lim_{n\to\infty} S_n\big(\mathbb{1}_{[a,b]}\big) = b - a, \tag{28.41}$$

alors que le membre de gauche n'est autre que

$$S_n\big(\mathbb{1}_{[a,b]}\big) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[a,b]}(x_k) = \frac{1}{n}X_n(a,b). \tag{28.42}$$

(4) **(2)⇒(3)** Vu que [7] $e^{2i\pi p x_k} = \cos(2\pi p x_k) + i\sin(2\pi p x_k)$. est une fonction périodique, c'est immédiat.

(5) **(3)⇒(2)** Par linéarité, le point (2) montre que si f est un polynôme trigonométrique, alors

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(x_k) = \int_0^1 f(t)dt. \tag{28.43}$$

(6) **Densité des polynômes trigonométriques** Il nous reste à prouver que les polynômes trigonométriques sont denses dans les fonctions continues sur $[0,1]$. Soit une fonction continue sur $[0,1]$ avec $f(0) = f(1)$. Alors le théorème de Stone-Weierstrass dans sa version trigonométrique (lemme 28.1) nous donne la densité.

Si $f(1) \neq f(0)$ c'est pas très grave : on peut trouver une fonction g vérifiant $g(0) = g(1)$ et $\|f - g\|_\infty \leqslant \epsilon$. Ensuite un polynôme trigonométrique approxime très bien g.

□

6. Je recommande chaudement de dessiner les fonctions φ_ϵ et ψ_ϵ pour avoir une idée de la situation.
7. Lemme 18.11.

28.2 Fonctions de Dirichlet

Définition 28.11.
*Une fonction $f : \mathbb{R} \to \mathbb{C}$ est une **fonction de Dirichlet** si*

(1) elle est 2π-périodique,

(2) elle est continue par morceaux,

(3) pour tout $x \in \mathbb{R}$ nous avons

$$f(x) = \frac{f(x^+) + f(x^-)}{2}. \tag{28.44}$$

Nous notons $\mathcal{D}$ l'ensemble des fonctions de Dirichlet.

LemVIwMsTC

Lemme 28.12 ([706]).
L'ensemble $C^0(S^1)$ est dense dans l'ensemble des fonctions de Dirichlet $\big(\mathcal{D}, \|.\|_2\big)$.

Démonstration. Nous commençons par supposer que $f \in \mathcal{D}$ n'a qu'un seul point de discontinuité, x_0. Alors nous considérons la fonction

$$f_n(x) = \begin{cases} f(x) & \text{si } x \in S^1 \backslash B(x_0, \frac{1}{n}) \\ d(x) & \text{si } x \in B(x_0, \frac{1}{n}) \end{cases} \tag{28.45}$$

où d est la droite joignant $f(x_0 - \frac{1}{n})$ et $f(x_0 + \frac{1}{n})$. La fonction f_n est continue et vérifie

$$|f_n(x)| \leqslant \|f\|_\infty \tag{28.46}$$

pour tout x. En effet si x est en dehors de $B(x_0, \frac{1}{n})$ c'est évident, et si $x \in B(x_0, \frac{1}{n})$, alors $|f_n(x)|$ est majoré soit par $f(x_0 - \frac{1}{n})$ soit par $f(x_0 + \frac{1}{n})$ suivant que d soit croissant ou décroissant. Avec ça nous avons

$$\|f_n - f\|_2^2 = \int_{x_0-1/n}^{x_0+1/n} |f_n(x) - f(x)|^2 dx \leqslant \int_{x_0-1/n}^{x_0+1/n} 4\|f\|_\infty = \frac{8\|f\|_\infty}{n}. \tag{28.47}$$

Et nous voyons que $\|f_n - f\|_2 \to 0$.

Si f contient plusieurs points de discontinuité, on fait le même coup autour de chaque point, en prenant n assez grand pour que si x_0 est un point de discontinuité, $B(x_0, \frac{1}{n})$ n'en contienne pas d'autres. □

Notons que la densité de $C^0(S^1)$ dans $\big(\mathcal{D}, \|.\|_\infty\big)$ est impossible, parce qu'une limite uniforme de fonctions continues est continue.

Théorème 28.13.
Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est total[8] *dans $\big(\mathcal{D}, \|.\|_2\big)$.*

Démonstration. Soit $f \in \mathcal{D}$. Si elle est continue, le théorème de Fejér 28.7 nous donne convergence uniforme sur S^1 d'une suite de polynômes trigonométriques vers f. Cette convergence est également une convergence L^2 parce que S^1 est compact.

Prenons donc $f \in \mathcal{D}$ non continue et $\epsilon > 0$[9]. Par le lemme 28.12, il existe une fonction $g \in C^0(S^1)$ telle que

$$\|g - f\|_2 \leqslant \epsilon. \tag{28.48}$$

Le théorème de Fejér donne aussi un polynôme trigonométrique P tel que $\|P - g\|_2 < \epsilon$; nous avons alors

$$\|P - f\|_2 \leqslant \|P - g\|_2 + \|g - f\|_2 \leqslant 2\epsilon. \tag{28.49}$$

□

8. Définition 25.20.

9. Par exemple $\epsilon = 0.4$, mais ce n'est qu'un exemple hein. Si vous en voulez un autre, prenez p, un nombre premier puis calculez $\epsilon = 1/p$.

Notons que cette histoire de fonctions de Dirichlet n'a pas attaqué le vrai fond du problème de la densité des polynômes trigonométriques dans $L^2(S^1)$ parce que nous restons avec une hypothèse de continuité, alors que les représentants des éléments de $L^2(S^1)$ n'ont strictement aucune régularité à priori.

28.3 Coefficients et série de Fourier

DEFooJUUIooNMdCtN

Définition 28.14.
Pour toutes les fonctions f définie sur $[0, 2\pi[$ ou périodique de période 2π, pour lesquelles les expressions ont un sens, nous définissons

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt, \tag{28.50}$$

EqNDBaXRL

*et nous nommons **série de Fourier** associée à f la série*

$$S(f)(x) = \sum_{k=-\infty}^{\infty} c_k(f) e^{ikx}. \tag{28.51}$$

Nous considérons aussi la suite (nous ne précisons pas dans quel espace)

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f) e^{ikx}. \tag{28.52}$$

Si la fonction f est de période T, nous définissons

$$c_n(f) = \frac{1}{T} \int_0^{T} f(t) e^{-2i\pi nt/T} dt. \tag{28.53}$$

EQooBOFSooFCJXzu

Le sport de la théorie des séries de Fourier est de donner des conditions sous lesquelles :

— les coefficients de Fourier et la série de Fourier ont un sens,

— la série de f est égale à f.

PropmrLfGt

Proposition 28.15 ([707]).
Soit f une fonction continue et 2π-périodique telle que sa série de Fourier converge uniformément. Alors la convergence est vers f.

Démonstration. Notons d'abord que f étant continue sur $[0, 2\pi]$, elle y est bornée et L^2. Par conséquent Parseval nous enseigne que

$$\|S_N(f) - f\|_{L^2} \to 0. \tag{28.54}$$

Cela signifie que

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_0^{2\pi} |f(t) - S_N(t)|^2 dt = 0. \tag{28.55}$$

L'hypothèse de convergence uniforme nous dit que la fonction $|f(t) - S_N(t)|^2$ converge uniformément vers la fonction $|f(t) - S(t)|^2$ où nous avons écrit S la limite de S_N. En permutant la limite et l'intégrale,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t) - S(t)|^2 dt = 0, \tag{28.56}$$

ce qui signifie que la fonction $t \mapsto |f(t) - S(t)|^2$ est la fonction nulle. Nous en déduisons que $f = S$. □

PropSgvPab

Proposition 28.16.
Soit f une fonction 2π-périodique. Si $\sum_{n\in\mathbb{Z}} |c_n(f)| < \infty$, alors pour tout $x \in \mathbb{R}$ nous avons

$$f(x) = \sum_{n\in\mathbb{Z}} c_n(f)e^{inx}. \tag{28.57}$$

De plus, la suite $(S_n(f))$ converge uniformément vers f.

Démonstration. Nous posons

$$g(x) = \sum_{n\in\mathbb{Z}} c_n(f)e^{inx}. \tag{28.58}$$

Étant donné les hypothèses, la série de droite converge absolument, la fonction g est continue sur $\mathbb{R}$. Nous avons

$$\big|g(x) - (S_n(f))(x)\big| \leqslant \sum_{|k|>n} |c_k(f)|, \tag{28.59}$$

mais le terme de droite tend vers zéro lorsque $n \to \infty$ parce que c'est le reste d'une série convergente. Cela signifie que $S_n(f)$ converge uniformément vers g.

Par ailleurs nous savons que dans L^2 nous avons la convergence $S_n(f) \to f$ (parce que f est continue sur le compact $[0, 2\pi]$ et donc y est bornée et L^2), ce qui signifie que $g = f$ presque partout. Ces deux fonctions étant continues, elles sont égales partout. □

ThozHXraQ

Théorème 28.17 ([1])**.**
Soit $f\colon \mathbb{R} \to \mathbb{C}$, une fonction C^1 et 2π-périodique. Pour $n \in \mathbb{Z}$ nous posons

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt. \tag{28.60}$$

Alors

ITEMooIDVEooJdMEmU

(1) Les coefficients de Fourier sont sommables : $(c_n) \in \ell^1(\mathbb{Z})$

ITEMooGIEUooKLyXej

(2) Pour tout $x \in \mathbb{R}$ nous avons

$$f(x) = \sum_{n\in\mathbb{Z}} c_n(f)e^{inx}. \tag{28.61}$$

ITEMooAUCTooTgJEPv

(3) La convergence est uniforme. C'est à dire que si nous posons

$$S_N(x) = \sum_{k=-N}^{N} c_k(f)e^{ikx}, \tag{28.62}$$

alors

$$\|S_N - f\|_\infty \to 0. \tag{28.63}$$

Démonstration. Point par point.

(1) **Pour (1)** Soit $n \in \mathbb{Z}$. Nous posons $g(t) = f(t)e^{-int}$. Nous avons

$$0 = g(2\pi) - g(0) = \int_0^{2\pi} g'(t)dt = \int_0^{2\pi} \big[f'(t)e^{-int} - inf(t)e^{-int}\big]dt. \tag{28.64}$$

Du coup, $c_n(f') = inc_n(f)$. La fonction f' étant bornée (parce que continue sur $[0, 2\pi]$), elle est de carré intégrable sur $[0, 2\pi]$ et par les inégalités de Parseval (théorème 25.46) nous avons

$$\sum_{n\in\mathbb{Z}} |c_n(f')|^2 < \infty. \tag{28.65}$$

Par conséquent $(c_n(f')) \in \ell^2(\mathbb{Z})$ et a fortiori $(c_n(f'))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$. L'inégalité de Cauchy-Schwarz nous indique alors

$$\sum_{n=1}^{\infty} |c_n(f)| = \sum_{n\in\mathbb{N}} \frac{1}{n}|c_n(f')| \leqslant \left(\sum_n \frac{1}{n^2}\right)^{1/2} \left(\sum_n |c_n(f')|^2\right)^{1/2} < \infty. \tag{28.66}$$

Nous procédons de même pour $n < 0$. Cela prouve que

$$\sum_{n\in\mathbb{Z}} |c_n(f)| = |c_0(f)| + \sum_{n<0} |c_n(f)| + \sum_{n>0} |c_n(f)| < \infty. \tag{28.67}$$

(2) **Pour (2)**

(3) **Pour (3)**

□

CordgtXlC

Corolaire 28.18 (Unicité des coefficients de Fourier[1])**.**
Soient f, g deux fonctions continues et 2π-périodiques.

ITEMooPLTIooSDykYF

(1) Si $c_n(f) = c_n(g)$ alors $f = g$.

ITEMooQMMSooEpIFbt

(2) Si $f(x) = \sum_{n\in\mathbb{Z}} a_n e^{inx}$, alors $a_n = c_n(f)$.

Démonstration. En deux points.

(1) **Pour (1)** Dans le cas de fonctions continues, le théorème de Fejér 28.7 nous enseigne que si nous posons

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx} \tag{28.68}$$

alors nous avons la convergence

$$\frac{1}{N+1}\sum_{n=0}^{N} S_n(f)(x) \to f(x). \tag{28.69}$$

Donc en supposant que $c_k(f) = c_k(g)$, nous avons $S_n(f)(x) = S_n(g)(x)$ et

$$f(x) = \lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} S_n(f)(x) = \lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} S_n(g)(x) = g(x). \tag{28.70}$$

(2) **Pour (2)** Nous considérons la restriction

$$\begin{aligned} \tilde{f}\colon [0,2\pi] &\to \mathbb{C} \\ x &\mapsto f(x). \end{aligned} \tag{28.71}$$

C'est une fonction bornée parce qu'elle est la restriction de f qui est continue sur, disons, le compact $[-\delta, 2\pi+\delta]$. Elle est donc dans l'espace de Hilbert $L^2\big([0,2\pi[\big)$.

En utilisant la base trigonométrique (27.478) (qui est une base par le lemme 27.132), nous écrivons l'hypothèse sous la forme

$$\tilde{f}(x) = \sum_{n\in\mathbb{Z}} \sqrt{2\pi}a_n e_n(x). \tag{28.72}$$

Autrement dit, $\tilde{f} = \sum_{n\in\mathbb{Z}} \sqrt{2\pi}a_n e_n$. La proposition 25.26 permet d'identifier les coefficients :

$$\sqrt{2\pi}a_n = \langle \tilde{f}, e_n\rangle. \tag{28.73}$$

Nous avons donc

$$a_n = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \tilde{f}(t)\overline{e_n(t)}dt \tag{28.74a}$$
$$= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \tilde{f}(t)\frac{1}{\sqrt{2\pi}}e^{-int}dt \tag{28.74b}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \tilde{f}(t)e^{-int}dt \tag{28.74c}$$
$$= \frac{1}{2\pi} \int_0^{2\pi} f(t)e^{-int}dt \tag{28.74d}$$
$$= c_n(f), \tag{28.74e}$$

ce qu'il fallait démontrer.

□

28.19.
La proposition 28.21 dit que les hypothèses de continuité et de périodicité ne sont pas suffisantes pour assurer la convergence de la série de Fourier. En particulier, pour 28.18(2), l'hypothèse de la convergence de la série est une vraie hypothèse.

Exemple 28.20.
Considérons la fonction

$$f(x) = 1 - \frac{x^2}{\pi^2} \tag{28.75}$$

sur $[-\pi, \pi]$. Nous la développons en série trigonométrique, et étant paire il n'y a pas de sinus. Un calcul montre que

$$a_0 = \frac{4}{3} \tag{28.76}$$

et

$$a_n = (-1)^{n+1}\frac{4}{n^2\pi^2}, \tag{28.77}$$

de telle sorte que

$$f(x) = \frac{2}{3} - \frac{4}{\pi^2}\sum_{n=1}^{\infty}(-1)^n\frac{\cos(nx)}{n^2}. \tag{28.78}$$

Nous avons $f(\pi) = 0$, mais avec le développement,

$$f(\pi) = \frac{2}{3} - \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}, \tag{28.79}$$

donc

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}. \tag{28.80}$$

△

28.3.1 Le contre-exemple que nous attendions tous

Nous montrons maintenant que la continuité et la périodicité ne sont pas suffisantes pour avoir convergence de la série de Fourier.

PropREkHdol

Proposition 28.21 ([104]).
Soit $C^0_{2\pi}(\mathbb{R})$ l'ensemble des fonctions périodiques continues muni de la norme uniforme. Nous définissons

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx}. \tag{28.81}$$

Alors il existe $f \in C^0_{2\pi}$ tel que la suite $n \mapsto S_n(f)(0)$ soit divergente. En particulier f n'est pas la somme de sa série de Fourier.

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} l_n \colon C^0_{2\pi} &\to \mathbb{C} \\ f &\mapsto S_n(f)(0) = \sum_{k=-n}^{n} c_k(f). \end{aligned} \tag{28.82}$$

(1) **La forme est continue** Nous montrons d'abord que l_n est continue en montrant que $\|l_n\| < \infty$ et en utilisant la proposition 10.70. Pour cela nous calculons un peu :

$$l_n(f) = \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e^{-ikt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_n(t) dt \tag{28.83}$$

où $D_n(t)$ est le noyau de Dirichlet dont nous connaissons une formule par le lemme 28.4. Nous avons donc

$$|l_n(f)| \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| \|f\|_\infty dt. \tag{28.84}$$

En prenant $\|f\|_\infty = 1$ nous avons la borne suivante pour la norme de l_n :

$$\|l_n\| \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| dt < \infty. \tag{28.85}$$

Notons que la convergence de l'intégrale vient de la continuité de la fonction

$$t \mapsto \frac{\sin\left(\frac{2n+1}{2}t\right)}{\sin\left(\frac{t}{2}\right)} \tag{28.86}$$

qui, elle même, se prouve avec une règle de l'Hospital :

$$\lim_{t\to 0} \frac{\sin(at)}{\sin(t)} = \lim_{t\to 0} \frac{a\cos(at)}{\cos(t)} = a. \tag{28.87}$$

Donc $D_n(t)$ a une limite bien définie pour $t \to 0$ et est alors une fonction continue sur le compact $[-\pi, \pi]$.

(2) **La norme de l_n (début)**
Nous avons prouvé que $\|l_n\| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)| dt$. Nous allons à présent prouver que ceci est effectivement la norme de l_n. Pour $\epsilon > 0$ nous considérons la fonction

$$\begin{aligned} f_\epsilon \colon \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \frac{D_n(x)}{|D_n(x)| + \epsilon}. \end{aligned} \tag{28.88}$$

C'est une fonction continue et 2π-périodique satisfaisant $\|f_\epsilon\| \leqslant 1$ parce que le dénominateur est toujours plus grand que le numérateur. Nous nous proposons de calculer

$$l_n(f_\epsilon) = \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f_\epsilon(t) e^{-ikt} dt. \tag{28.89}$$

Puisque $f_\epsilon(t)e^{-ikt}$ vaut en norme $|f_\epsilon(t)|$, qui est une fonction intégrable (ne dépendant pas de k) sur $[-\pi, \pi]$, le théorème de la convergence dominée 14.207 nous permet de permuter la somme et l'intégrale :

$$l_n(f_\epsilon) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{D_n(t)}{|D_n(t)| + \epsilon} \underbrace{\sum_{k=-n}^{n} e^{-ikt}}_{=D_n(t)} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|D_n(t)|^2}{|D_n(t)| + \epsilon} dt. \tag{28.90}$$

Nous avons donc

$$\lim_{\epsilon\to 0} l_n(f_\epsilon) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|dt. \tag{28.91}$$

Mais vue l'inégalité (28.85) nous avons

$$\|l_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|dt. \tag{28.92}$$

Notre tâche est maintenant de donner une valeur à cette intégrale.

(3) **Norme de l_n tend vers ∞** D'abord nous écrivons

$$\|l_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|\sin\left(\frac{2n+1}{2}t\right)\right|}{\left|\sin(t/2)\right|}dt, \tag{28.93}$$

ensuite nous nous souvenons que $|\sin(x)| \leqslant |x|$ pour tout x, ce qui nous permet de changer le dénominateur :

$$\|l_n\| \geqslant \frac{2}{\pi}\int_{0}^{\pi} \frac{\left|\sin\left(\frac{2n+1}{2}t\right)\right|}{|t|}dt \tag{28.94}$$

Nous y effectuons le changement de variable $u = \frac{2n+1}{2}t$ qui donne

$$\|l_n\| \geqslant \frac{2}{\pi}\int_{0}^{(n+\frac{1}{2})\pi} \frac{\left|\sin(u)\right|}{|u|}du. \tag{28.95}$$

Nous y reconnaissons l'intégrale (18.95) du sinus cardinal que nous savons diverger. Cela donne

$$\lim_{n\to\infty} \|l_n\| = \infty. \tag{28.96}$$

(4) **La conclusion**

L'espace $\left(C^0_{2\pi}, \|.\|_\infty\right)$ est complet [10], donc le théorème de Banach-Steinhaus 10.152 s'applique. Par rapport aux notations de l'énoncé de Banch-Steinhaus, nous posons

$$E = \left(C^0_{2\pi}, \|.\|_\infty\right) \tag{28.97a}$$
$$F = \mathbb{R} \tag{28.97b}$$
$$H = \{l_n\}_{n\in\mathbb{N}}. \tag{28.97c}$$

Comme la suite $(\|l_n\|)$ n'est pas bornée, il existe $f \in C^0_{2\pi}$ tel que

$$\sup_n \|l_n(f)\| = \infty. \tag{28.98}$$

Pour cette fonction nous avons

$$\sup_{n\geqslant 0} S_n(f)(0) = \infty, \tag{28.99}$$

et donc la série de Fourier de f ne converge pas en zéro.

□

NORMooGKKWooFmOBeE

28.22.

La proposition 28.21 ne contredit pas Fejér 28.7(4). Alors là c'est subtil, donc soyez bien attentive.

Le système trigonométrique est total dans l'espace des fonctions continues périodiques sur $\mathbb{R}$: $\left(C^0_{2\pi}(\mathbb{R}), \|.\|_\infty\right)$. Cela signifie que si $f \in C^0_{2\pi}(\mathbb{R})$, il existe une suite de polynômes trigonométriques P_k tels que $P_k \xrightarrow{unif} f$.

Cela ne signifie pas que cette suite soit la suite des sommes partielles de la série de Fourier. Et en effet, le théorème de Fejér ne donne pas la convergence de la suite des sommes partielles de

10. Parce qu'une limite uniforme de fonctions continues est continue, théorème 12.357.

Fourier, mais la convergence au sens de Cesàro de la somme des $c_k(f)e_k$. Ce n'est pas la même chose.

Notez que le coefficient de e_1 dans F_2 est $c_1(f)/2$ alors que dans F_3, il est $2c_1(f)/3$.

Il y a donc bien une suite de polynômes trigonométriques qui converge vers f, mais ce n'est pas la suite des sommes partielles de la série de Fourier.

De plus, $C^0_{2\pi}(\mathbb{R})$ n'est pas contenu dans $L^2(\mathbb{R})$, donc nous ne pouvons pas invoquer la théorie de Hilbert pour dire que le système trigonométrique serait quelque chose comme une base. Si $f \in C^0_{2\pi}(\mathbb{R})$, il n'est pas garanti qu'il existe des nombres $(a_i)_{i\in\mathbb{Z}}$ tels que $f(x) = \sum_{n\in\mathbb{Z}} a_n e^{inx}$.

Enfin, me diriez-vous, les fonctions continues et périodiques sur $\mathbb{R}$ sont les mêmes que les fonctions définies sur $[0, 2\pi[$ avec une petite condition de $\lim_{x\to 2\pi} f(x) = f(0)$. Les fonctions qui vérifient cela sont une partie de $L^2\big([0, 2\pi[\big)$, qui est un espace de Hilbert, lui. Or le système trigonométrique est une base hilbertienne (lemme 27.132). Alors oui, la série de Fourier de f converge vers f lorsque f est continue sur $[0, 2\pi[$. Cela n'est cependant pas un contre-argument pour deux raisons :

— La convergence $S_n(f) \to f$ qu'on a dans $L^2\big([0, 2\pi[\big)$ est seulement une convergence pour la norme L^2, et non une convergence uniforme.

— Une convergence L^2 sur $[0, 2\pi[$ ne se prolonge pas spécialement en une convergence L^2 sur $\mathbb{R}$, et encore moins en une convergence uniforme sur $\mathbb{R}$.

28.3.2 Inégalité isopérimétrique

Le théorème suivant dit que parmi les courbes C^1, le cercle a la plus grande surface possible à périmètre donné.

ThoIXyctPo

Théorème 28.23 (Inégalité isopérimétrique[104])**.**
Soit $f\colon S^1 \to \mathbb{C}$ une courbe de Jordan[11] de classe C^1. Nous notons L sa longueur et S l'aire contenue de la surface délimitée[12] par f. Alors

*(1) Nous avons l'**inégalité isopérimétrique** : $L^2 \geqslant 4\pi S$.*

(2) Nous avons l'égalité $L^2 = 4\pi S$ si et seulement si la courbe donnée par f est un cercle.

Démonstration. Nous commençons par considérer un chemin dont la longueur est 2π et nous en considérons son paramétrage normal. Nous allons exprimer l'aire S en utilisant le théorème de Green, et plus particulièrement la formule de surface (20.271).

Si $f(s) = x(s) + iy(s)$, nous devons intégrer $y'x - x'y$, qui n'est rien d'autre que la partie imaginaire de $f'(s)\overline{f(s)}$. Donc

$$S = \frac{1}{2}\operatorname{Im}\int_0^{2\pi} f'(s)\overline{f(s)}ds \tag{28.100}$$

EqCSWKbPX

Nous considérons les coefficients de Fourier de f donnés par la formule (28.50) :

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(s)e^{-ins}ds. \tag{28.101}$$

Ceux de f' (qui est aussi continue sur le compact S^1 et donc tout autant L^2) sont donnés par

$$c_n(f') = inc_n(f). \tag{28.102}$$

D'autre part en vertu du théorème 21.9, la longueur de γ s'exprime en termes de l'intégrale de la norme de sa dérivée :

$$2\pi = L = \int_0^{2\pi} |f'(s)|ds = \int_0^{2\pi} |f'(s)|^2 ds \tag{28.103}$$

11. Définition 20.75

12. C'est la partie connexe bornée de $\mathbb{C}\backslash\gamma$ dont l'existence est donnée par le théorème de Jordan 27.215.

parce que nous avons choisi un paramétrage normal qui vérifie automatiquement $|f'(s)| = 1$ pour tout s. L'identité de Parseval sous sa forme (25.130) appliquée à f' nous enseigne que

$$L = 2\pi = \int_0^{2\pi} |f'(s)|^2 ds = 2\pi \sum_{n=-\infty}^{\infty} |c_n(f')|^2 = 2\pi \sum_{n\in\mathbb{Z}} n^2 |c_n(f)|^2, \tag{28.104}$$

et donc que

$$\sum_{n\in\mathbb{Z}} n^2 |c_n(f)|^2 = 1. \tag{28.105}$$

Par ailleurs le système trigonométrique étant une base hilbertienne, et les fonctions f et f' étant dans $L^2\big([0,2\pi]\big)$ (parce que continues sur un compact), elles sont égales à leurs séries de Fourier (au sens L^2), c'est-à-dire que nous avons l'égalité (27.482). Nous avons alors

$$\begin{aligned}
\langle f', f\rangle_{L^2} &= \langle \sum_{n\in\mathbb{Z}} c_n(f')e_n, \sum_{m\in\mathbb{Z}} c_m(f)e_m\rangle && (28.106a)\\
&= \sum_m \sum_n c_n(f')\overline{c_m(f)} \underbrace{\langle e_n, e_m\rangle}_{\delta_{n,m}} && (28.106b)\\
&= \sum_{n\in\mathbb{Z}} c_n(f')\overline{c_n(f)} && (28.106c)\\
&= \sum_n in|c_n(f)|^2 && (28.106d)
\end{aligned}$$

où nous avons utilisé la continuité du produit scalaire pour sortir les sommes. Avec cela nous pouvons exprimer l'aire (28.100) en termes de coefficients de Fourier :

$$S = \frac{1}{2} \operatorname{Im} 2\pi\langle f', f\rangle = \pi \sum_{n\in\mathbb{Z}} n|c_n(f)|^2. \tag{28.107}$$

En utilisant les expressions (28.104) et (28.107) pour L et S, et en écrivant $L = 2\pi$, nous avons

$$\begin{aligned}
L^2 - 4\pi S &= 4\pi^2 \left(\sum_{n\in\mathbb{Z}} n^2 |c_n(f)|^2 \right)^2 - 4\pi^2 \sum_{n\in\mathbb{Z}} n|c_n(f)|^2 && (28.108a)\\
&\geqslant 4\pi^2 \sum_{n\in\mathbb{Z}} n^2 |c_n(f)|^2 - 4\pi^2 \sum_{n\in\mathbb{Z}} n|c_n(f)|^2 && (28.108b)\\
&= 4\pi^2 \sum_{n\in\mathbb{Z}} |c_n(f)|^2 (n^2 - n) && (28.108c)\\
&\geqslant 0. && (28.108d)
\end{aligned}$$

Justifications.

— Pour (28.108a). Expression (28.104) pour L et (28.107) pour S.

— Pour (28.108b). La somme dans le premier terme valant 1 par (28.105), nous pouvons supprimer le carré.

Cela prouve l'inégalité demandée dans le cas où $L = 2\pi$.

Si γ n'est pas de longueur 2π mais L, alors nous considérons le chemin $\sigma(t) = \frac{2\pi\gamma(t)}{L}$. Sa longueur est 2π et son aire, au vu de la formule de Green (28.100), est de $4\pi^2 \frac{S}{L^2}$. L'inégalité isopérimétrique appliquée au chemin σ donne alors $L^2 \geqslant 4\pi S$.

Le cas d'égalité s'obtient uniquement si $c_n = 0$ pour tout n différent de 0 ou 1. Dans ce cas nous avons

$$f(s) = c_0(f) + c_1(f)e^{is}, \tag{28.109}$$

qui est un cercle de centre $c_0(f)$ et de rayon $|c_1(f)|$. □

28.3.3 À propos des coefficients

Pour la suite, nous avons besoin d'une notation pour désigner l'ensemble des suites dans $\mathbb{C}$ à index dans $\mathbb{Z}$, c'est-à-dire l'ensemble $\text{Fun}(\mathbb{Z}, \mathbb{C})$. Pour alléger les notations, nous allons l'écrire $\mathbb{C}^{\mathbb{Z}}$, conformément à des notations déjà introduites par exemple en 1.335.

Nous considérons l'application

$$\begin{aligned} c\colon \big(L^1_{2\pi}, \|.\|_1\big) &\to \big(\mathbb{C}^{\mathbb{Z}}, \|.\|_\infty\big) \\ f &\mapsto (c_n(f))_{n\in\mathbb{Z}} \end{aligned} \tag{28.110}$$

qui à une fonction 2π-périodique fait correspondre la suite (bornée) de ses coefficients de Fourier. Nous rappelons la définition

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}dt. \tag{28.111}$$

Nous allons montrer que cette application est linéaire, continue, injective et non surjective. Pour la continuité, par la linéarité il suffit de la montrer en 0. Nous devons donc montrer que si nous avons une suite de fonctions f_k qui tend vers 0 au sens L^1, alors $c(f_k) \to 0$ au sens de la norme $\|.\|_\infty$ sur l'ensemble des suites.

Si nous posons $r_k = \int_0^{2\pi} |f_k(t)|dt$, alors $r_k = \|f_k\|_1$ et nous avons $r_k \to 0$. Mais par définition

$$|c_n(f_k)| \leqslant r_k, \tag{28.112}$$

et donc $\|c(f_k)\|_\infty \leqslant r_k$. L'application c est donc continue. L'injectivité est donnée par le corolaire 28.18.

Si nous supposons que l'application c est continue, alors le théorème d'isomorphisme de Banach (27.1) nous dit que cela devrait être un homéomorphisme, c'est-à-dire que c^{-1} serait également continue. Nous allons montrer qu'il n'en est rien.

Nous considérons la suite de suite

$$(c_n)_k = \begin{cases} 0 & \text{si } n < 0 \\ 1 & \text{si } k < n \\ 0 & \text{sinon.} \end{cases} \tag{28.113}$$

EqdMtbQB

Ici $(c_n)_k$ est le terme numéro k de la suite (c_n). Par exemple $c_0 = (0, 0, \ldots)$ et $c_2 = (1, 1, 0, \ldots)$.

Par injectivité de l'application qui à une fonction fait correspondre la suite de ses coefficients de Fourier, l'unique fonction qui possède ces coefficients est

$$f_n(t) = \sum_{k\in\mathbb{N}} c_{n,k}e^{ikt}. \tag{28.114}$$

En ce qui concerne la norme de f_n, nous avons

$$\|f_n\|_1 = \frac{1}{2\pi}\int_0^{2\pi} \sum_{k\in\mathbb{N}} (c_n)_k|e^{ikt}|dt = \sum_{k\in\mathbb{N}} (c_n)_k = n. \tag{28.115}$$

Étant donné que $\|f_n\|_1 = n$, la suite $(\|f_n\|_1)$ n'est pas bornée alors que la suite de suites (28.113) est bornée dans l'ensemble des suites parce que $\|c_n\|_\infty = 1$.

LEMooPUJDooKRBTaU

Lemme 28.24.
Soit une fonction $f\colon \mathbb{R} \to \mathbb{C}$ qui est T-périodique et de classe C^1. Alors

$$c_n(f') = \frac{2\pi n}{T} i c_n(f). \tag{28.116}$$

Démonstration. Nous rappelons la définition (28.53) des coefficients de Fourier :

$$c_n(f) = \frac{1}{T}\int_0^T f(t)e^{-2i\pi nt/T}dt. \tag{28.117}$$

Le coefficient pour f' ne pose pas de problème d'existence parce que f' est continue sur le compact $[0,T]$. Il vaut

$$c_n(f') = \frac{1}{T}\int_0^T f'(t)e^{-2i\pi nt/T}dt \tag{28.118a}$$
$$= \frac{1}{T}\left[f(t)e^{-2i\pi nt/T}\right]_0^T - \frac{1}{T}\int_0^T f(t)\left(\frac{-2i\pi n}{T}\right)e^{-2i\pi nt/T}dt \tag{28.118b}$$
SUBEQooXYOVooGmoXbZ
$$= \frac{2i\pi n}{T}\frac{1}{T}\int_0^T f(t)e^{-2i\pi nt/T}dt \tag{28.118c}$$
SUBEQooXSCEooUXExT
$$= \frac{2i\pi n}{T}c_n(f). \tag{28.118d}$$

Justifications.

— Pour (28.118b). C'est une intégration par partie avec $u' = f'$ et $v = e^{-2i\pi nt/T}$.

— Pour (28.118c). Comme $f(T) = f(0)$, et que $t \mapsto e^{-2i\pi nt/T}$ est périodique de période T, le terme au bord est nul : $f(T)e^{-2i\pi n} - f(0)e^{i0} = 0$.

□

LEMooYJQWooDVvSyj

Lemme 28.25 ([633]).
Soit une fonction $f\colon \mathbb{R} \to \mathbb{C}$ de classe C^2 et T-périodique. Alors

$$|c_n(f)| \leqslant \left(\frac{T}{2\pi}\right)^2 \frac{\|f''\|_\infty}{n^2}. \tag{28.119}$$

Démonstration. En utilisant la définition (28.53) des coefficients de Fourier,

$$|c_n(f)| \leqslant \frac{1}{T}\int_0^T |f(t)|dt \leqslant \frac{1}{T}\|f\|_\infty \int_0^T 1dt = \|f\|_\infty. \tag{28.120}$$

En appliquant le lemme 28.24 à f' nous avons

$$c_n(f'') = \left(\frac{2i\pi n}{T}\right)^2 c_n(f). \tag{28.121}$$

Donc

$$|c_n(f)| = \left(\frac{T}{2\pi n}\right)^2 |c_n(f'')| \leqslant \left(\frac{T}{2\pi n}\right)^2 \|f''\|_\infty. \tag{28.122}$$

□

28.4 Série de Laurent

THOooMKJOooVghZyG

Théorème 28.26 (Série de Laurent[633]).
Soient la couronne

$$C(r_1, r_2) = \{z \in \mathbb{C} \text{ tel que } r_1 < |z| < r_2\} \tag{28.123}$$

et une fonction holomorphe $f\colon C(r_1, r_2) \to \mathbb{C}$. Alors :

(1) Il existe une suite (a_n) dans $\mathbb{C}$ telle que

$$f(z) = \sum_{n\in\mathbb{Z}} a_n z^n. \tag{28.124}$$

ITEMooUOPHooSJRGKs

(2) Cette suite est unique : si

$$f(z) = \sum_{n\in\mathbb{Z}} a_n z^n = \sum_{n\in\mathbb{Z}} b_n z^n, \tag{28.125}$$

alors $a_n = b_n$, pour tout n.

ITEMooDGGZooJkDSxC

(3) *Si on pose, pour* $r \in]r_1, r_2[$,

$$\begin{aligned} f_r \colon \mathbb{R} &\to \mathbb{C} \\ \theta &\mapsto f(re^{i\theta}), \end{aligned} \tag{28.126}$$

les valeurs a_n *sont liés aux coefficients de Fourier de* f_r *par*

$$a_n = \frac{c_n(f_r)}{r^n}. \tag{28.127}$$

ITEMooOYCPooZZAyKs

(4) *Cette série converge uniformément sur tout compact contenu dans* $C(r_1, r_2)$.

(5) *Pour tout* $r_1 < s < r_2$, *les coefficients sont donnés par*[13]

$$a_n = \frac{1}{2\pi i} \int_{C_s} \frac{f(z)}{z^{n+1}} dz \tag{28.128}$$

où C_s *est un cercle centré en* 0, *et de rayon* s.

La série ainsi définie est la ***série de Laurent*** *de la fonction* f.

Démonstration. Pour $r \in]r_1, r_2[$ nous posons

$$\begin{aligned} f_r \colon \mathbb{R} &\to \mathbb{C} \\ \theta &\mapsto f(re^{i\theta}). \end{aligned} \tag{28.129}$$

(1) **Coefficients de Fourier** La fonction f_r est de classe C^1 et périodique. Le théorème 28.17 sur les séries de Fourier nous indique que

$$f_r(t) = \sum_{n \in \mathbb{Z}} c_n(r) e^{int} \tag{28.130}$$

EQooIHQRooZWqJKL

avec[14]

$$c_n(r) = \frac{1}{2\pi} \int_0^{2\pi} f_r(t) e^{-int} dt. \tag{28.131}$$

La fonction $c_n \colon]r_1, r_2[\to \mathbb{C}$ est une fonction définie par une intégrale que nous voudrions dériver en r_0.

(2) **Digression** Deux voies s'offrent à nous.

— Le plus immédiatement disponible est le théorème 17.19, mais il demande de travailler avec la dérivée (réelle) de $r \mapsto f(re^{i\theta})$ et de se poser des questions quant à son lien avec la dérivée (complexe) de f.

— Une façon plus indirecte est de considérer une extension

$$\begin{aligned} h \colon B(r_0, \delta)_{\mathbb{C}} \times [0, 2\pi[&\to \mathbb{C} \\ z, \theta &\mapsto f(ze^{i\theta}) e^{-in\theta} \end{aligned} \tag{28.132}$$

où δ est assez petit pour que le tout reste dans le domaine de f. Alors nous pouvons utiliser le théorème 26.73 qui a l'avantage de ne pas devoir majorer la dérivée. Mais cette voie demande de réellement faire le lien entre la dérivée complexe de h et la dérivée réelle de $r \mapsto f(re^{i\theta})$.

Nous choisissons la première voie parce qu'en réalité, elle évite complètement de parler de dérivée complexe.

13. Pour le dire clairement, ces a_n ne dépendent pas de s, même si s entre dans le membre de droite.

14. Oui, on devrait écrire $c_n(f_r)$ pour suivre scrupuleusement les notation introduites plus haut. Mais comme toute la suite de la démonstration sera de voir le tout comme fonction de r, je vous laisse juger.

(3) **Permuter dérivée et intégrale** Nous allons essayer de la dériver en $r_0 \in]r_1, r_2[$ en utilisant le théorème 17.19. Pour y voir plus clair, ce qui joue le rôle de f dans l'énoncé de 17.19 est l'application

$$\begin{aligned} h\colon]r_1, r_2[\times [0, 2\pi[&\to \mathbb{C} \\ (r, \theta) &\mapsto f_r(\theta) e^{-in\theta}. \end{aligned} \tag{28.133}$$

Nous considérons un intervalle $I = B(r_0, \delta)$ assez petit pour être dans $]r_1, r_2[$. Passons en revue les conditions.

(3a) **Pour (1)**
Nous introduisons la fonction

$$\begin{aligned} \tilde{f}\colon \mathbb{R} \times [0, 2\pi] &\to \mathbb{C} \\ (r, \theta) &\mapsto f(re^{i\theta}) \end{aligned} \tag{28.134}$$

Pour tout $r \in I$, la fonction $\theta \mapsto f(re^{i\theta})e^{-in\theta}$ est dans $L^1\big([0, 2\pi[\big)$ parce que

$$|f(re^{i\theta})e^{-in\theta}| = |f(re^{i\theta})| = |\tilde{f}(r, \theta)|. \tag{28.135}$$

La fonction $\tilde{f}$ étant continue, elle est bornée sur le compact $\overline{I} \times [0, 2\pi]$.

(3b) **Pour (2)** Pour chaque θ, la fonction $r \mapsto f(re^{i\theta})e^{-in\theta}$ est dérivable par la proposition 26.3(1).

(3c) **Pour (3)** En utilisant la proposition 26.3(3), nous savons que la fonction

$$\frac{\partial h}{\partial r}(r, \theta) = \frac{\partial \tilde{f}}{\partial r}(r, \theta) e^{-in\theta} \tag{28.136}$$

est continue et donc bornée sur le compact $\overline{I} \times [0, 2\pi]$. Une fonction constante majorant de $\partial_r h$ est intégrable sur le compact $[0, 2\pi]$.

En permutant nous avons donc

$$c_n'(r_0) = \frac{1}{2\pi} \int_0^{2\pi} (\partial_r \tilde{f})(r_0, \theta) e^{-in\theta} d\theta. \tag{28.137}$$

(4) **Cauchy-Riemann** C'est le moment d'utiliser Cauchy-Riemann en coordonnées polaires sous la forme de la proposition 26.3(2). Et tant que nous y sommes, nous notons $g(\theta) = \tilde{f}(r_0, \theta)$ pour avoir moins de choses à écrire :

$$c_n'(r_0) = \int_0^{2\pi} (\partial_r \tilde{f})(r_0, \theta) e^{-in\theta} d\theta \tag{28.138a}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{ir_0} (\partial_\theta \tilde{f})(r_0, \theta) e^{-in\theta} d\theta \tag{28.138b}$$

$$= \frac{1}{2\pi i r_0} \int_0^{2\pi} g'(\theta) e^{-in\theta} d\theta. \tag{28.138c}$$

La dernière expression a manifestement envie de se soumettre à une intégration par partie.

(5) **Une intégration par partie** Nous posons $u(\theta) = e^{-in\theta}$ et $v = g$, de telle sorte que

$$c_n'(r_0) = \frac{1}{2\pi i r_0} \left(\big[e^{-in\theta} g(\theta) \big]_0^{2\pi} - \int_0^{2\pi} (-in) e^{-in\theta} g(\theta) d\theta \right) \tag{28.139a}$$

$$= \frac{n}{2\pi r_0} \int_0^{2\pi} g(\theta) e^{-in\theta} d\theta. \tag{28.139b}$$

$$= \frac{n}{r_0} c_n(r_0). \tag{28.139c}$$

Justification pour (28.139b). Pour rappel, $g(\theta) = \tilde{f}(r_0, \theta) = f(r_0 e^{i\theta})$; donc $g(0) = g(2\pi)$ et le « terme au bord » est nul.

(6) **Équation différentielle** L'équation (28.139) dit que c_n satisfait à l'équation différentielle

$$c_n'(r) = \frac{n}{r} c_n(r) \tag{28.140}$$

pour tout $r \in]r_1, r_2[$. La fonction c_n est une fonction à valeurs complexes dont les parties réelles et imaginaires vérifient toutes deux l'équation du lemme 26.91. Il existe donc $a_n \in \mathbb{C}$ tel que

$$c_n(r) = a_n r^n. \tag{28.141}$$

Cela prouve au passage le point (3) parce que r n'est jamais nul.

(7) **La valeur de a_n** Nous avons

$$\begin{aligned} a_n &= \frac{c_n(r)}{r^n} \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\theta}) e^{-in\theta} r^{-n} d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{f(re^{i\theta})}{(e^{i\theta} r)^n} d\theta \\ &= \frac{1}{2\pi i} \int_{C_s} \frac{f(z)}{z^{n+1}} dz. \end{aligned} \tag{28.142}$$

EQooFNUHooZbbNAT

Une justification pour l'intégrale curviligne s'impose. La définition est 26.5. Dans le cas du cercle, nous considérons

$$\begin{aligned} C_r : [0, 2\pi[&\to \mathbb{C} \\ \theta &\mapsto re^{i\theta}, \end{aligned} \tag{28.143}$$

et donc

$$\int_{C_r} \frac{f(z)}{z^{n+1}} dz = \int_0^{2\pi} \frac{f(re^{i\theta})}{(re^{i\theta})^{n+1}} rie^{i\theta} d\theta = i \int_0^{2\pi} \frac{f(re^{i\theta})}{(re^{i\theta})^n} d\theta. \tag{28.144}$$

Le fait que le tout soit égal à a_n prouve que l'intégrale est en réalité indépendante de r [15].

(8) **Conclusion** Reprenons la formule (28.130) :

$$f(re^{i\theta}) = f_r(\theta) = \sum_n c_n(r) e^{in\theta} = \sum_n a_n r^n e^{in\theta} = \sum_n a_n (re^{i\theta})^n. \tag{28.145}$$

Autrement dit,

$$f(z) = \sum_n a_n z^n, \tag{28.146}$$

avec les a_n donnés par la formule (28.142), comme nous devions le prouver.

(9) **Point (2) (unicité)** Supposons que $f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$, fixons $r \in]r_1, r_2[$, et posons

$$\begin{aligned} f_r : \mathbb{R} &\to \mathbb{C} \\ \theta &\mapsto f(re^{i\theta}). \end{aligned} \tag{28.147}$$

C'est une fonction continue et périodique. Elle peut s'écrire sous la forme

$$f_r(\theta) = f(re^{i\theta}) = \sum_{n\in\mathbb{Z}} a_n (r^n e^{in\theta}) = \sum_{n\in\mathbb{Z}} (a_n r^n) e^{in\theta}. \tag{28.148}$$

Le corolaire 28.18(2) à propos de l'unicité des coefficients de Fourier implique que $a_n r^n = c_n(f)$ et donc que

$$a_n = \frac{c_n(f)}{r^n}. \tag{28.149}$$

EQooBNSMooGLIBqU

Donc les coefficients a_n sont déterminés par f. Si nous avions fait le calcul en partant de $f(z) = \sum_{n\in\mathbb{Z}} b_n z^n$, nous aurions eu $b_n = \frac{c_n(f)}{r^n}$, et donc bien $a_n = b_n$.

15. Voir ma question 28.27.

(10) **Point (4)** La fonction

$$\begin{aligned} f_r\colon \mathbb{R} &\to \mathbb{C} \\ \theta &\mapsto f(re^{i\theta}) \end{aligned} \tag{28.150}$$

est continue et périodique de période 2π. Le lemme 28.25 nous indique que, pour tout $r \in]r_1, r_2[$, nous avons

$$|c_n(f_r)| \leqslant \frac{\|f_r''\|_\infty}{n^2}. \tag{28.151}$$

(10a) **Sur une couronne** Soient deux rayons intermédiaires $r_1 < s_1 < s_2 < r_2$. La couronne fermée $\overline{C(s_1, s_2)}$ est compacte. Nous considérons la fonction

$$\begin{aligned} g\colon [s_1, s_2] \times [0, 2\pi] &\to \mathbb{C} \\ (u, \theta) &\mapsto f_u''(\theta) \end{aligned} \tag{28.152}$$

Nous notons g_u l'application $\theta \mapsto g(u, \theta)$; c'est une application définie sur $[s_1, s_2]$. Par le lemme 12.79, l'application $u \mapsto \|g_u\|_\infty$ est continue et donc majorée : nous pouvons considérer $M \in \mathbb{R}$ tel que $\|g_u\|_\infty < M$ pour tout $u \in [s_1, s_2]$.
Puisque, pour $r \in [s_1, s_2]$ nous avons $g_u = f_u''$, en combinant avec (28.151), nous voyons qu'il existe M tel que

$$|c_n(f_r)| \leqslant \frac{M}{n^2}. \tag{28.153}$$

En prenant les notations de la définition 10.93 de la convergence normale, nous posons

$$\begin{aligned} u_n\colon \overline{C(s_1, s_2)} &\to \mathbb{C} \\ z &\mapsto a_n z^n \end{aligned} \tag{28.154}$$

En ce qui concerne sa norme[16], nous avons

$$\begin{aligned} \|u_n\|_\infty &= |a_n| \sup_{z \in \overline{C(s_1,s_2)}} |z^n| && \text{(28.155a)} \\ &= |a_n| s_2^n && \text{(28.155b)} \\ &= \left| \frac{c_n(f_s)}{s^n} \right| s_2^n && \text{(28.155c)} \\ &\leqslant \frac{|c_n(f_s)|}{s_2^n} s_2^n && \text{(28.155d)} \\ &\leqslant \frac{M}{n^2} && \text{(28.155e)} \end{aligned}$$

Justifications :

— Pour (28.155c). Prendre n'importe quel $s \in [s_1, s_2]$ et c'est bon par (28.149).

— Pour (28.155e). Équation (28.153).

Comme la somme $\sum_{n\in\mathbb{Z}} M/n^2$ converge, nous avons la convergence normale de $\sum_{n\in\mathbb{Z}} a_n z^n$ sur $\overline{C(s_1, s_2)}$.

(10b) **Sur un compact quelconque** Soit K, un compact dans $C(r_1, r_2)$. La fonction $z \mapsto |z|$ est continue sur K ; donc elle a un minimum et un maximum. Nous posons $s_1 = \min_{z\in K} |z|$ et $s_2 = \max_{z\in K} |z|$.
Ah ah ! non. En fait nous ne définissons pas s_1 et s_2 de cette manière parce qu'il y a un risque que $s_1 = s_2$ et qu'alors $\overline{C(s_1, s_2)}$ soit vide et ne contienne donc pas K – pour rappel, $C(s_1, s_2)$ est la couronne ouverte.
Nous choisissons donc s_1 et s_2 de telle sorte que

$$r_1 < s_1 < \min_{z\in K} |z| \leqslant \max_{z\in K} |z| < s_2 < r_2. \tag{28.156}$$

16. Faites bien attention que dans cette partie, $\|.\|_\infty$ est la norme uniforme sur $\overline{C(s_1, s_2)}$ et non sur $C(r_1, r2)$ ou sur $\mathbb{C}$.

Tout ça pour dire que $K \subset \overline{C(s_1, s_2)}$. La convergence normale sur $\overline{C(s_1, s_2)}$ déjà prouvée implique la convergence normale sur K.

□

PROPBooYWDNooMXVPLJ

¡¡ Avertissement/question à la lectrice !! 28.27

Sur Wikipédia[633], le fait que a_n ne dépende pas de r est prouvé en disant que $a_n(s) = c_n(s)/s^n$ et $a_n(t) = c_n(t)/t^n$ et que

$$\frac{c_n(s)}{c_n(t)} = \frac{s^n}{t^n}. \tag{28.157}$$

EQooCPDTooBDxIKn

En mettant tout cela bout à bout,

$$\frac{a_n(s)}{a_n(t)} = 1. \tag{28.158}$$

Je ne comprends pas très bien pourquoi cette justification est nécessaire. À mon avis, la rédactrice de la démonstration sur Wikipédia parvient à déduire la relation (28.157) directement depuis l'équation différentielle, sans réellement avoir besoin de la résoudre.

Si vous comprenez n'hésitez pas à m'écrire, parce que j'ai l'impression d'avoir manqué quelque chose.

¡¡ Avertissement/question au lecteur !! 28.28

L'énoncé de la proposition 28.29 n'est peut-être pas précis. Si vous avez un énoncé correct sous le coude, écrivez-moi.

PROPooBMZGooLoaGLK

Proposition 28.29.

Si f est holomorphe sur $B(a, r)$, alors sa série de Laurent est de la forme

$$f(z) = \sum_{n=0}^{\infty} a_n (z - a)^n. \tag{28.159}$$

avec

$$a_n = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{(z - a)^{n+1}} dz. \tag{28.160}$$

où

$$\begin{aligned} \gamma \colon [0, 2\pi[&\to \mathbb{C} \\ t &\mapsto a + re^{it} \end{aligned} \tag{28.161}$$

est le cercle de centre a et de rayon r.

Chapitre 29

Transformation de Fourier

DEFooRIXGooECoIbx

Définition 29.1.
*Soit une fonction f sur $\mathbb{R}^d$, dont nous ne précisons pas la régularité. Sa **transformée de fourier** est la fonction $\hat{f}$ définie par*

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dx \tag{29.1}$$

EQooCHMUooFbSmaz

si elle existe.

29.2.
Ce qui est bien avec cette définition est que si la formule (29.1) ne définit pas $\hat{f}$ (parce que l'intégrale n'existe pas, par exemple), nous nous réservons le droit de définir tout de même $\hat{f}$ par d'autres biais. Ce sera d'ailleurs l'objet du théorème 29.32 qui définira $\hat{f}$ pour tout $f \in L^2(\mathbb{R}^d)$ alors que la formule (29.1) ne fonctionne pas sur toutes ces fonctions.

Une bonne partie de ce qui va suivre aura pour objet de déterminer des espaces de fonctions sur lesquels la transformée est bien définie, et sur lesquels elle a de bonnes propriétés.

Nous allons par ailleurs utiliser indifféremment les notations $\mathcal{F}(f)$ ou $\hat{f}$ pour la transformée de Fourier de f. La notation $\mathcal{F}$ est pratique pour les transformées de looooooongues expressions ainsi que pour parler de l'application « transformée de Fourier » d'un espace de fonction vers un autre.

29.3.
Nous verrons dans le théorème 29.32 que la transformée de Fourier n'est pas une isométrie de L^2. Pour avoir une isométrie, il aurait fallu choisir des coefficients moins simples.

Proposition 29.4.
La transformée de Fourier est $\mathbb{C}$-linéaire au sens suivant. Soient des fonctions $f, g \colon \mathbb{R}^d \to \mathbb{C}$ telles que $\mathcal{F}(f)$ et $\mathcal{F}(g)$ existent. Alors

(1) La transformée de Fourier de $f+g$ existe et vaut $\mathcal{F}(f+g) = \mathcal{F}(f) + \mathcal{F}(g)$

(2) Pour tout $\lambda \in \mathbb{C}$, la transformée de Fourier de λf existe et vaut $\mathcal{F}(\lambda f) = \lambda \mathcal{F}(f)$.

Démonstration. C'est la proposition 14.193 qui nous permet de séparer l'intégrale dans le calcul suivant :

$$\mathcal{F}(f+g)(\xi) = \int_{\mathbb{R}^d} (f+g)(x) e^{-i\xi \cdot x} dx \tag{29.2a}$$

$$= \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dx + \int_{\mathbb{R}^d} g(x) e^{-i\xi \cdot x} dx \tag{29.2b}$$

$$= \mathcal{F}(f)(\xi) + \mathcal{F}(g)(\xi). \tag{29.2c}$$

En ce qui concerne $\mathcal{F}(\lambda f)$, c'est la même chose, en utilisant la linéarité de l'intégrale. □

29.1 Transformée de Fourier sur $L^1(\mathbb{R}^d)$

Nous rappelons que les espaces L^p sont des ensembles de classes de fonctions, définition 27.7. La transformée de Fourier, comme presque tout ce qui a trait aux intégrales, passe aux classes.

Lemme 29.5.
Soit une fonction $f : \mathbb{R}^d \to \mathbb{C}$ telle que $\mathcal{F}(f)$ existe. Alors pour toute fonction $g \in [f]$ la transformée $\mathcal{F}(g)$ existe et $\hat{g} = \hat{f}$.

Démonstration. Par définition des classes, il existe une fonction $s : \mathbb{R}^d \to \mathbb{C}$ presque partout nulle telle que $g = f + s$. Soit $\xi \in \mathbb{R}^d$ fixé. La fonction $x \mapsto s(x)e^{-i\xi x}$ est presque partout nulle et donc intégrable d'intégrale nulle. La proposition 14.193 nous permet alors d'affirmer que $f + s$ est intégrable et que

$$\mathcal{F}(f+s)(\xi) = \int_{\mathbb{R}^d} (f+s)(x)e^{-i\xi x}dx = \int_{\mathbb{R}^d} f(x)e^{-i\xi x}dx + \int_{\mathbb{R}^d} s(x)e^{-i\xi x}dx = \mathcal{F}(f)(\xi). \tag{29.3}$$

□

À partir de maintenant, lorsque nous parlons de transformée de Fourier d'une fonction dans L^p, nous parlons indifféremment d'une vraie fonction ou d'une classe.

Lemme 29.6.
Si $f \in L^1(\mathbb{R}^d)$, alors $\hat{f}$ existe.

Démonstration. Par définition, si $f \in L^1(\mathbb{R}^d)$, alors l'intégrale $\int_{\mathbb{R}^d} |f|$ existe et est finie. Alors la fonction qui arrive dans la transformée de Fourier en ξ, la fonction $s : x \to f(x)e^{-i\xi x}$ est également dans $L^1(\mathbb{R}^d)$ parce que $|s| = |f|$. □

29.7.
Si $f \in L^1(\mathbb{R})$, la transformée de Fourier $\hat{f}$ existe, mais il n'est pas garanti qu'elle soit elle-même dans L^1, même si il est vrai que $\hat{f}$ est continue (proposition 29.11) et bornée (proposition 29.14). Nous allons immédiatement montrer un exemple de fonction L^1 dont la transformée de Fourier n'est pas dans L^1.

LEMooROPHooOSguhN

Lemme 29.8.
La transformée de Fourier de

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & si\ x \in [-\frac{1}{2}, \frac{1}{2}] \\ 0 & sinon \end{cases} \end{aligned} \tag{29.4}$$

est

$$\hat{f}(\xi) = -\frac{\sin(\xi/2)}{\xi/2}. \tag{29.5}$$

La fonction $\hat{f}$ n'est pas dans $L^1(\mathbb{R})$.

Démonstration. Pour $\xi \in \mathbb{R}$ fixé, la fonction $x \mapsto f(x)e^{i\xi x}$ est continue à support compact, donc

nous n'avons aucun problèmes avec l'intégrale

$$\begin{aligned}\hat{f}(\xi) &= \int_{\mathbb{R}} f(x)e^{-i\xi x}dx && (29.6a)\\ &= \int_{[-1/2,1/2]} e^{-i\xi x} && (29.6b)\\ &= \big[-\frac{1}{i\xi}e^{-i\xi x}\big]_{x=-1/2}^{x=1/2} && (29.6c)\\ &= -\frac{1}{i\xi}\big(e^{-i\xi/2}-e^{i\xi/2}\big) && (29.6d)\\ &= -\frac{2i\sin(\xi/2)}{i\xi} && (29.6e)\\ &= \frac{\sin(\xi/2)}{\xi/2}. && (29.6f)\end{aligned}$$

Le fait que cette fonction ne soit pas dans $L^1(\mathbb{R})$ est le lemme 20.212. □

LEMooKGDKooVXSMCn

Lemme 29.9.
Si $f \in L^1(\mathbb{R}^d)$ et si $g(x) = f(\lambda x)$ alors

$$\hat{g}(\xi) = \lambda^{-d}\hat{f}(\xi/\lambda). \tag{29.7}$$

Démonstration. Il s'agit de faire le changement de variable $y = \lambda x$ dans l'intégrale

$$\hat{g}(\xi) = \int_{\mathbb{R}^d} f(\lambda x)e^{-i\xi x}dx. \tag{29.8}$$

Dans le changement de variables, vient le coefficient $dx = \lambda^{-d}dy$. □

PropfqvLOl

Proposition 29.10.
La transformée de Fourier est un morphisme vis-à-vis de la convolution sur $L^1(\mathbb{R}^n)$:

$$\widehat{f * g} = \hat{f}\hat{g}. \tag{29.9}$$

Démonstration. Nous devons étudier l'intégrale

$$\widehat{f * g}(\xi) = \int_{\mathbb{R}}\left[\int_{\mathbb{R}} f(y)g(t-y)dy\right]e^{-it\xi}dt. \tag{29.10}$$

Ici nous avons choisi des représentants f et g dans les classes de L^1. Montrons que f est borélienne. D'abord $f(x) = f_+(x) - f_-(x)$ où f_+ et f_- sont des fonctions positives. Afin d'alléger les notations nous supposons un instant que f est positive et nous posons

$$f_n(x) = \sum_{k=1}^{2^n}\frac{k}{n}\mathbb{1}_{f(x)\in[\frac{k}{n},\frac{k+1}{n}[}. \tag{29.11}$$

Le fait que f soit dans L^1 implique que chacune des fonctions f_n est borélienne [1] et donc que f l'est aussi en tant que limite ponctuelle de fonctions boréliennes [2].

Nous allons appliquer le théorème de Fubini 14.303 à la fonction

$$\phi(x,y) = f(x)g(y)e^{-i\xi(x+y)} \tag{29.12}$$

1. Ceci demanderait plus de justification. Dites moi si vous savez comment justifier que les f_n soient boréliennes.
2. Le fait que f soit borélienne est une conséquence du théorème 27.242.

qui est borélienne en tant que produit et composée de fonctions boréliennes. Nous avons

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x)e^{-i\xi x}||g(y)e^{-i\xi y}|dy \right) dx = \int_{\mathbb{R}} \left(|f(x)| \int_{\mathbb{R}} |g(y)|dy \right) dx \tag{29.13a}$$

$$= \int_{\mathbb{R}} |f(x)|\|g\|_1 \tag{29.13b}$$

$$= \|f\|_1\|g\|_1 < \infty. \tag{29.13c}$$

Le théorème est donc applicable. D'abord nous avons :

$$\hat{f}(\xi)\hat{g}(\xi) = \left(\int_{\mathbb{R}} f(x)e^{-i\xi x}dx \right) \left(\int_{\mathbb{R}} g(y)e^{-i\xi y}dy \right) \tag{29.14a}$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(y)e^{-i\xi(x+y)}dy \right) dx \tag{29.14b}$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(t-x)e^{-i\xi t}dt \right) dx. \tag{29.14c}$$

Jusqu'ici nous n'avons pas utilisé Fubini. Nous avons seulement introduit le nombre $\int_{\mathbb{R}} g(y)e^{-i\xi y}dy$ dans l'intégrale par rapport à x et effectué le changement de variables $y \mapsto t = x + y$. Maintenant nous appliquons le théorème de Fubini pour inverser l'ordre des intégrales :

$$\hat{f}(\xi)\hat{g}(\xi) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(t-x)e^{-it\xi}dt \right) dx \tag{29.15a}$$

$$= \int_{\mathbb{R}} e^{-it\xi} \left(\int_{\mathbb{R}} f(x)g(t-x)dx \right) dt \tag{29.15b}$$

$$= \int_{\mathbb{R}} e^{-it\xi}(f * g)(t)dt \tag{29.15c}$$

$$= \widehat{f * g}(\xi). \tag{29.15d}$$

□

PropJvNfj

Proposition 29.11.
Soit une fonction $f \in L^1(\mathbb{R}^d)$. Alors sa transformée de Fourier est continue.

Démonstration. Nous considérons une fonction f définie sur $\mathbb{R}^d$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Sa transformée de Fourier est donnée par

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi x}f(x)dx. \tag{29.16}$$

Pour montrer que cette fonction $\hat{f}$ est continue en ξ_0 nous considérons une suite $(\xi_n) \to \xi_0$ et nous voulons montrer que $\hat{f}(\xi_n) \to \hat{f}(\xi_0)$. Pour cela nous considérons les fonctions

$$g_n(x) = e^{-i\xi_n x}f(x) \tag{29.17}$$

qui convergent simplement vers $g(x) = e^{-i\xi_0 x}f(x)$. Étant donné que

$$|g_n(x)| < |f(x)|, \tag{29.18}$$

le théorème de la convergence dominée donne alors

$$\lim_{n\to\infty} \int g_n(x) = \int \lim_{n\to\infty} g_n(x), \tag{29.19}$$

c'est-à-dire $\lim_{n\to\infty} \hat{f}(\xi_n) = \hat{f}(\xi_0)$. La fonction $\hat{f}$ est donc continue. □

LEMooCBPTooYlcbrR

Lemme 29.12.
Pour tout $f \in L^1(\mathbb{R}^n)$ nous avons $\|\hat{f}\|_\infty \leqslant \|f\|_1$.

Démonstration. Cela est une simple vérification :

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-ix\xi}dx, \tag{29.20}$$

nous avons, pour tout ξ,

$$|\hat{f}(\xi)| \leqslant \int_{\mathbb{R}} |f(x)|dx, \tag{29.21}$$

ce qui signifie exactement $\|\hat{f}\|_\infty \leqslant \|f\|_1$. □

LesmRLaxXkQV

Lemme 29.13 (Lemme de Riemann-Lebesgue[708])**.**
Si f est une fonction $L^1(\mathbb{R})$ alors $\lim_{\xi\to\pm\infty} \hat{f}(\xi) = 0$.

Démonstration. Nous commençons par prouver le résultat dans le cas d'une fonction g en escalier, et plus précisément par une fonction caractéristique d'un compact $K = [a, b]$. Au niveau de la transformée de Fourier nous avons

$$\hat{\mathbb{1}}_K(\xi) = \int_a^b e^{-i\xi x}dx = -\frac{1}{i\xi}(e^{-ib\xi} - e^{-ia\xi}). \tag{29.22}$$

Par conséquent

$$|\hat{\mathbb{1}}_K(\xi)| \leqslant \frac{2}{|\xi|}. \tag{29.23}$$

Plus généralement si $g = \sum_{i=1}^N c_i \mathbb{1}_{K_i}$, alors

$$|\hat{g}(\xi)| \leqslant \frac{2}{|\xi|} \sum_{i=1}^N |c_i|, \tag{29.24}$$

et donc nous avons effectivement $\lim_{\xi\to\pm\infty} |\hat{g}(\xi)| = 0$.

Nous passons maintenant au cas général $f \in L^1(\mathbb{R})$. Étant donné que les fonctions L^1 en escalier sont denses dans L^1, nous considérons une fonction $g \in L^1(\mathbb{R})$ en escalier telle que $\|f - g\|_1 < \epsilon$. Nous avons donc

$$\|\hat{f} - \hat{g}\|_\infty \leqslant \|f - g\|_1 < \epsilon. \tag{29.25}$$

Donc dans

$$\|\hat{f}(\xi)\| \leqslant \|\hat{f}(\xi) - \hat{g}(\xi)\| + |\hat{g}(\xi)|, \tag{29.26}$$

le premier terme est plus petit que ϵ. Il nous reste à voir que

$$\lim_{\xi\to\infty} |\hat{g}(\xi)| = 0, \tag{29.27}$$

mais cela est le résultat de la première partie de la preuve. □

CORooHSNYooZlZoyV

Corolaire 29.14.
La transformée de Fourier d'une fonction $L^1(\mathbb{R})$ est bornée.

Démonstration. Par le corolaire 29.11, la transformée de Fourier d'une fonction L^1 est continue. Le lemme de Riemann-Lebesgue 29.13 impliquant qu'elle tend vers zéro en $\pm\infty$, elle doit être bornée. □

29.1.1 Formule sommatoire de Poisson

ProprPbkoQ

Proposition 29.15 (Formule sommatoire de Poisson)**.**
Soit $f : \mathbb{R} \to \mathbb{C}$ une fonction continue et $L^1(\mathbb{R})$. Nous supposons que

(1) il existe $M > 0$ et $\alpha > 1$ tels que

$$|f(x)| \leqslant \frac{M}{(1 + |x|)^\alpha}, \tag{29.28}$$

(2) $\sum_{n=-\infty}^{\infty} |\hat{f}(2\pi n)| < \infty$.

Alors nous avons

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n). \tag{29.29}$$

Démonstration. (1) **Convergence normale**

Nous commençons par montrer qu'il y a convergence normale sur tout compact séparément des séries sur les $n \geqslant 0$ et sur les $n < 0$.

Soit K un compact de $\mathbb{R}$ contenu dans $[-A, A]$ et $n \in \mathbb{Z}$ tel que $|n| \geqslant 2A$. Pour $x \in K$ nous avons

$$|x + n| \geqslant |n| - |x| \geqslant |n| - A \geqslant \frac{|n|}{2}. \tag{29.30}$$

Du coup nous avons un $\alpha > 1$ tel que

$$|f(x + n)| \leqslant \frac{M}{\big(1 + |x + n|\big)^{\alpha}} \leqslant \frac{M}{\Big(1 + \frac{|n|}{2}\Big)^{\alpha}}. \tag{29.31}$$

Lorsque n est grand, cela a le comportement de $M/|n|^{\alpha}$ et donc la série

$$\sum_{n=0}^{\infty} f(x + n) \tag{29.32}$$

est une série convergeant normalement. Les deux séries (usuelles)

$$a_+ = \sum_{n \geqslant 0} f(x + n) \tag{29.33a}$$

$$a_- = \sum_{n < 0} f(x + n) \tag{29.33b}$$

convergent normalement.

(2) **Convergence commutative** Au sens de la définition 10.112 nous avons

$$\sum_{n \in \mathbb{Z}} f(x + n) = a_+ + a_-. \tag{29.34}$$

En effet si nous prenons $J_0' \subset \mathbb{N}$ fini tel que $|\sum_{\mathbb{N} \setminus J_0'} f(x + n) - a_+| \leqslant \epsilon$ et $J_1' \in -\mathbb{N}$ tel que $|\sum_{n \in -\mathbb{N} \setminus J_1'} f(x + n) - a_-| < \epsilon$, et si nous posons $J_0 = J_0' \cup J_1'$ alors si K est un ensemble fini de $\mathbb{Z}$ contenant J_0 nous avons

$$|\sum_{n \in K} f(n + x) - (a_+ + a_-)| \leqslant |\sum_{n \in K^+} f(n + x) - a_+| + |\sum_{n \in K^-} f(n + x) - a_-| \leqslant 2\epsilon \tag{29.35}$$

où K^+ sont les éléments positifs de K et K^- sont les *strictement* négatifs. Maintenant que la famille $\{f(n + x)\}_{n \in \mathbb{Z}}$ est une famille sommable, nous savons qu'elle est commutativement sommable et que la proposition 10.126 nous permet de sommer dans l'ordre que l'on veut. Nous pouvons donc écrire sans ambigüité l'expression $\sum_{n \in \mathbb{Z}} f(x + n)$ ou $\sum_{n=-\infty}^{\infty} f(x + n)$.

(3) **re-convergence normale**

Nous posons donc sans complexes la série

$$F(x) = \sum_{n \in \mathbb{Z}} f(x + n) \tag{29.36}$$

qui converge tant commutativement que normalement. Notons que nous pouvons maintenant dire que la série sur $\mathbb{Z}$ converge normalement ; pas seulement les deux séries séparément.

(4) **Continuité, périodicité** Étant donné que chacune des fonctions $f(x+n)$ est continue, la convergence normale nous assure que F est continue.
De plus F est périodique de période 1 parce que

$$F(x+1) = \sum_{n=-\infty}^{\infty} f(x+1+n) = \sum_{p=-\infty}^{\infty} f(x+p) = F(x) \tag{29.37}$$

où nous avons posé $p = 1+n$.
Notons que nous n'avons pas spécialement prouvé que F n'était pas périodique avec des périodes plus petites que 1. Mais cela n'a pas d'importance ici.

(5) **Coefficients de Fourier**
En vertu de la définition (27.329) et de la périodicité de F,

$$\begin{aligned}
c_n(F) &= \int_{-1/2}^{1/2} F(t)e^{-2\pi int}dt && (29.38a)\\
&= \int_0^1 F(t)e^{-2\pi int}dt && (29.38b)\\
&= \int_0^1 \sum_{n\in\mathbb{Z}} f(t+n)e^{-2i\pi nt}dt && (29.38c)\\
&= \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(u)e^{-2\pi i(u-n)n}du && (29.38d)\\
&= \int_{-\infty}^{\infty} f(u)e^{-2\pi inu}du && (29.38e)\\
&= \hat{f}(2\pi n). && (29.38f)
\end{aligned}$$

Justifications :

— Changement de variables $u = t+n$.
— permuté l'intégrale et la somme en vertu du fait que la somme converge normalement.
— Égalité $e^{-2\pi i(u-n)n} = e^{-2\pi iun}e^{-2\pi in^2} = e^{-2\pi iun}$.

(6) **Conclusion**
Étant donné l'hypothèse $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$ la proposition 28.16 nous dit que

$$F(x) = \sum_{n\in\mathbb{Z}} c_n(F)e^{2\pi inx}, \tag{29.39}$$

c'est-à-dire que

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n)e^{2\pi inx}. \tag{29.40}$$

En écrivant cette égalité en $x = 0$ nous trouvons le résultat :

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(2\pi n). \tag{29.41}$$

□
ExDLjesf

Exemple 29.16.
La formule sommatoire de Poisson peut être utilisée pour calculer des sommes dans l'espace de Fourier plutôt que dans l'espace direct. Nous allons montrer dans cet exemple l'égalité

$$\sum_{n=-\infty}^{\infty} e^{-\alpha n^2} = \sum_{n=-\infty}^{\infty} \sqrt{\frac{\pi}{\alpha}} e^{-\pi^2 n^2/\alpha}. \tag{29.42}$$

Si α est grand, alors la somme de gauche est plus rapide, tandis que si α est petit, c'est le contraire.

Nous appliquons la formule sommatoire de Poisson à la fonction

$$f(x) = e^{-\alpha x^2}. \tag{29.43}$$

Nous avons [EqCDeLht]

$$\hat{f}(x) = \int_{\mathbb{R}} e^{-\alpha t^2 - ixt} dt \tag{29.44a}$$

$$= e^{-x^2/4\alpha} \int_{\mathbb{R}} e^{-(\sqrt{\alpha}t + \frac{ix}{2\sqrt{\alpha}})^2} \tag{29.44b}$$

$$= e^{-x^2/4\alpha} \frac{1}{\sqrt{\alpha}} \int_{\mathbb{R}+\frac{ix}{2\sqrt{\alpha}}} e^{-u^2} du. \tag{29.44c}$$

Pour traiter cette intégrale nous utilisons la proposition 26.10 en considérant le chemin rectangulaire fermé qui joint les points $-R$, R, $R+ai$, $-R+ai$ et $f(z) = e^{-z^2}$. Calculons l'intégrale sur les deux côtés verticaux. Nous posons

$$\gamma_R(t) = R + tia \tag{29.45}$$

avec $t\colon 0 \to 1$. Nous avons

$$\int_{\gamma_R} f = \int_0^1 f\big(\gamma_R(t)\big)\|\gamma_R'(t)\|dt \tag{29.46a}$$

$$= ae^{-R^2} \int_0^1 e^{-2tRia + at^2} dt, \tag{29.46b}$$

donc en module nous avons

$$\Big|\int_{\gamma_R} f\Big| \leqslant ae^{-R^2} \int_0^1 e^{at^2} dt \leqslant Me^{-R^2}, \tag{29.47}$$

où M est une constante ne dépendant pas de R. Lorsque $R \to \infty$, la contribution des chemins verticaux s'annule et nous trouvons que

$$\int_{\mathbb{R}+ai} e^{-u^2} du = \int_{\mathbb{R}} e^{-u^2} du, \tag{29.48}$$ [EqirNxLr]

que nous pouvons utiliser pour continuer le calcul (29.44). Nous avons

$$\hat{f}(x) = \frac{e^{-x^2/4\alpha}}{\sqrt{\alpha}} \int_{\mathbb{R}} e^{-u^2} du = \sqrt{\frac{\pi}{\alpha}} e^{-x^2/4\alpha} \tag{29.49}$$

où nous avons utilisé la formule (14.984). Par conséquent ce qui rentre dans la formule sommatoire de Poisson est

$$\hat{f}(2\pi n) = \sqrt{\frac{\pi}{\alpha}} e^{-\pi^2 n^2/\alpha}. \tag{29.50}$$

△

29.2 Suite régularisante

[DEFooRIFYooUUUoha]

Définition 29.17.
*Une **suite régularisante** est une suite (ρ_n) dans $L^1(\mathbb{R}^d)$ telle que* [ITEMooEYXYooAkKeXX]

(1) pour tout n, $\rho_n \geqslant 0$ et $\int_{\mathbb{R}^d} \rho_n = 1$;

(2) pour tout $\alpha > 0$,

$$\lim_{n\to\infty} \int_{|t|>\alpha} \rho_n = 0. \tag{29.51}$$

Une telle suite est régularisante parce que souvent $\rho_n \in \mathscr{D}(\mathbb{R}^d)$, ce qui donne $f * \rho_n \in C^\infty$ par le corolaire 27.72.

PROPooYUVUooMiOktf

Proposition 29.18 ([709, 710])**.**
Soit une suite régularisante $\rho_n \in L^1(\mathbb{R}^d)$. Alors :

ITEMooLWMIooFFamdf

(1) Si f est continue à support compact, nous avons la convergence uniforme sur $\mathbb{R}^d$:

$$f * \rho_n \xrightarrow{unif} f. \tag{29.52}$$

ITEMooEJKKooChcgyM

(2) Si $g \in L^p$ $(1 \leqslant p < \infty)$ alors

$$g * \rho_n \xrightarrow{L^p} g. \tag{29.53}$$

Démonstration. Si f est continue à support compact, elle est uniformément continue[3], et elle est bornée. Soit $\epsilon > 0$ et $\alpha > 0$ tel que pour tout x, y tels que $\|x-y\| < \alpha$ nous ayons $|f(x)-f(y)| < \epsilon$. Nous prenons de plus n suffisamment grand pour avoir $\int_{B(0,\alpha)^c} \rho_n < \epsilon$. Nous avons alors

$$\begin{aligned}
|f(x) - (f*\rho_n)(x)| &= |\int_{\mathbb{R}^d} \big(f(x)-f(y)\big)\rho_n(x-y)dy| && \text{(29.54a)}\\
&\leqslant \int_{B(x,\alpha)} \underbrace{|f(x)-f(y)|}_{\leqslant\epsilon} \rho_n(x-y)dy + \int_{B(x,\alpha)^c} \underbrace{|f(x)-f(y)|}_{\leqslant 2\|f\|_\infty} \rho_n(x-y)dy && \text{(29.54b)}\\
&\leqslant \epsilon(1+2\|f\|_\infty). && \text{(29.54c)}
\end{aligned}$$

Nous avons prouvé que pour tout $\epsilon > 0$, il existe N tel que $n > N$ implique $|f(x)-(f*\rho_n)(x)| \leqslant \epsilon$. Cela prouve l'uniforme convergence sur $\mathbb{R}^d$ de $f * \rho_n$ vers f.

Pour le point (2) nous considérons $g \in L^p(\mathbb{R}^d)$ et $\phi \in \mathscr{D}(\mathbb{R}^d)$. Nous avons la majoration

$$\|g*\rho_n - g\|_p \leqslant \|g*\rho_n - \phi*\rho_n\|_p + \|\phi*\rho_n - \phi\|_p + \|\phi - g\|_p \tag{29.55}$$

En ce qui concerne le premier terme ;

$$\|(g-\phi)*\rho_n\|_p \leqslant \|g-\phi\|_p \tag{29.56}$$

par la proposition 27.70. Donc

EQooXWZYooVtsROK

$$\|g*\rho_n - g\|_p \leqslant 2\|g-\phi\|_p + \|\phi*\rho_n - \phi\|_p. \tag{29.57}$$

Par la densité de $\mathscr{D}$ dans L^p (théorème 27.51(5)) nous pouvons considérer une suite $\phi_i \xrightarrow{L^p} g$ dans $\mathscr{D}(\mathbb{R}^d)$. L'inégalité (29.57) est valable pour chaque i :

$$\|g*\rho_n - g\|_p \leqslant 2\|g-\phi_i\|_p + \|\phi_i*\rho_n - \phi_i\|_p. \tag{29.58}$$

Nous effectuons la limite sur $n \to \infty$:

$$\lim_{n\to\infty} \|g*\rho_n - g\|_p \leqslant 2\|g-\phi_i\| + \underbrace{\lim_{n\to\infty} \|\phi_i*\rho_n - \phi_i\|_p}_{=0} \tag{29.59}$$

parce que le point (1) s'applique à ϕ_i. Nous effectuons ensuite la limite sur $i \to \infty$ dans

$$\lim_{n\to\infty} \|g*\rho_n - g\| \leqslant 2\|g-\phi_i\| \to 0. \tag{29.60}$$

□

3. Théorème de Heine 12.76.

29.3 Transformée de Fourier dans l'espace de Schwartz

La définition de la transformée de Fourier de $\varphi \in \mathscr{S}(\mathbb{R}^d)$ est

$$\hat{\varphi}(\xi) = \int_{\mathbb{R}^n} \varphi(x) e^{-ix \cdot \xi} dx. \tag{29.61}$$

Si α est un multiindice de taille m, nous notons

$$(M_\alpha f)(x) = x_{\alpha_1} \dots x_{\alpha_m} f(x). \tag{29.62}$$

LemQPVQjCx

Lemme 29.19 (Lemme de transfert)**.**
Si $\varphi \in \mathscr{S}(\mathbb{R}^d)$ et si α est un multiindice, alors

$$\partial^\alpha \hat{\varphi} = (-i)^{|\alpha|} \widehat{M_\alpha \varphi}. \tag{29.63}$$

et

$$\widehat{\partial^\alpha \varphi}(\xi) = (-i)^{|\alpha|} \xi^\alpha \hat{\varphi}(\xi). \tag{29.64}$$

Démonstration. Nous considérons la fonction $h(x, \xi) = \varphi(x) e^{-ix \cdot \xi}$ dont la dérivée par rapport à ξ_i est donnée par $-i(M_i\varphi)(x)e^{x \cdot \xi}$. Cette fonction est majorée en norme par

$$G(x) = M_i \varphi(x), \tag{29.65}$$

qui est encore une fonction à décroissance rapide et donc parfaitement intégrable sur $\mathbb{R}^d$. Le théorème 17.19 nous dit donc que la dérivée de $\hat{\varphi}$ par rapport à ξ_j existe et vaut

$$\frac{\partial \hat{\varphi}}{\partial \xi_j}(\xi) = -i \int_{\mathbb{R}^n} x_j \varphi(x) e^{-i\xi \cdot x} = -i \widehat{M_j \varphi}(\xi). \tag{29.66}$$

En appliquant ce résultat en chaine, nous trouvons la première formule annoncée.

Nous passons à la seconde formule annoncée. Étant donné que $\varphi \in \mathscr{S}$, ses dérivées le sont aussi et par conséquent, il n'y a pas de problèmes pour écrire

$$\widehat{\partial_{x_k} \varphi}(\xi) = \int_{\mathbb{R}^d} \frac{\partial \varphi}{\partial x_k}(x) e^{-ix \cdot \xi} dx. \tag{29.67}$$

EqTYizlnia

Étant donné que

$$\frac{\partial}{\partial x_k} \left(\varphi(x) e^{-ix \cdot \xi} \right) = \frac{\partial \varphi}{\partial x_k}(x) e^{-ix \cdot \xi} - i\xi_k \varphi(x) e^{-ix \cdot \xi}, \tag{29.68}$$

EqZAsYaCR

notre tâche sera de prouver que

$$\int_{\mathbb{R}^d} \frac{\partial}{\partial x_k} \left(\varphi(x) e^{-ix \cdot \xi} \right) dx = 0. \tag{29.69}$$

EqVGvYBNK

Autrement dit, nous voulons montrer que le terme au bord d'une intégration par partie s'annule. D'abord le fait que φ soit à décroissance rapide nous assure que l'intégrale (29.69) converge. Pour chaque ξ, la fonction

$$f(x, \xi) = \frac{\partial}{\partial x_k} \left(\varphi(x) e^{-ix \cdot \xi} \right) \tag{29.70}$$

est intégrable par rapport à x. De plus, f est dans $\mathscr{S}(\mathbb{R})$ pour chacune de ses variables (les autres étant fixées). Le théorème de Fubini 14.304 nous permet alors de décomposer l'intégrale en

$$\int_{\mathbb{R}^d} f(x, \xi) dx = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_d) dx_1 \dots dx_d. \tag{29.71}$$

De plus nous pouvons intégrer dans l'ordre de notre choix et nous choisissons évidemment d'intégrer d'abord par rapport à x_k. Étudions donc l'intégrale

$$\int_{\mathbb{R}} \frac{\partial}{\partial x}\Big(\varphi(x)e^{-ix\xi}\Big)\,dx = \lim_{A\to\infty}\int_{-A}^{A}\frac{\partial}{\partial x}\Big(\varphi(x)e^{-ix\xi}\Big)\,dx \tag{29.72}$$

dans laquelle nous avons un peu allégé les notations. Une primitive de ce qui est intégré est toute trouvée : c'est $\varphi(x)e^{-ix\xi}$, et nous pouvons utiliser le théorème fondamental du calcul intégral pour écrire que

$$\int_{-A}^{A}\Big(\varphi(x)e^{-ix\xi}\Big)'\,dx = \Big[\varphi(x)e^{-ix\xi}\Big]_{x=-A}^{x=A}. \tag{29.73}$$

Vu que φ est dans $\mathscr{S}$, la limite $A\to\infty$ donne zéro.

En substituant maintenant (29.68) dans (29.67) et en tenant compte du terme que nous venons de montrer s'annuler, nous avons

$$\widehat{\partial_k\varphi}(\xi) = -i\xi_k\int_{\mathbb{R}^d}\varphi(x)e^{-ix\cdot\xi} = -i\xi_k\hat{\varphi}(\xi). \tag{29.74}$$

En recommençant la procédure $|\alpha|$ fois nous trouvons la seconde formule annoncée. □

PropKPsjyzT

Proposition 29.20 ([468]).
À propos de transformée de Fourier dans l'espace de Schwartz.

(1) L'espace de Schwartz[4] *est stable par transformée de Fourier, c'est-à-dire que* $\mathcal{F}\big(\mathscr{S}(\mathbb{R}^d)\big)\subset\mathscr{S}(\mathbb{R}^d)$[5].

(2) L'application $\mathcal{F}\colon\mathscr{S}(\mathbb{R}^d)\to\mathscr{S}(\mathbb{R}^d)$ *est continue.*

Démonstration. La linéarité découle de celle de l'intégrale. La difficulté est de prouver que pour $\varphi\in\mathscr{S}(\mathbb{R}^d)$ nous avons bien que $\hat{\varphi}\in\mathscr{S}(\mathbb{R}^d)$ et que cette association est continue[6].

(1) **Première inclusion : $\mathcal{F}\big(\mathscr{S}(\mathbb{R}^d)\big)\subset\mathscr{S}(\mathbb{R}^d)$** Nous devons prouver que pour tout multiindices α et β, nous avons $p_{\alpha,\beta}(\hat{\varphi})<\infty$. Nous avons

$$\xi^\beta\partial^\alpha\hat{\varphi}(\xi) = \xi^\beta(-i)^{|\alpha|}\widehat{M_\alpha\varphi}(\xi) = (-i)^{|\alpha|+|\beta|}\widehat{\partial^\beta M_\alpha\varphi}(\xi). \tag{29.75}$$

Ensuite nous nous souvenons que $\|\hat{f}\|_\infty\leqslant\|f\|_1$ parce que

$$|\hat{f}(\xi)|\leqslant\int_{\mathbb{R}^d}\big|f(x)e^{-ix\cdot\xi}\big| = \int_{\mathbb{R}^d}|f(x)|dx = \|f\|_1. \tag{29.76}$$

Donc

$$p_{\alpha,\beta}(\hat{\varphi}) = \|\widehat{\partial^\beta M_\alpha\varphi}\|_\infty\leqslant\|\partial^\beta M_\alpha\varphi\|_1. \tag{29.77}$$

Du fait que φ soit dans $\mathscr{S}$, la dernière expression est finie. Cela prouve déjà que

$$\mathcal{F}\big(\mathscr{S}(\mathbb{R}^d)\big)\subset\mathscr{S}(\mathbb{R}^d). \tag{29.78}$$

(2) **Continuité**

Nous supposons avoir une suite $\varphi_n\xrightarrow{\mathscr{S}}\varphi$, et nous devons prouver que $\hat{\varphi}_n\xrightarrow{\mathscr{S}}\hat{\varphi}$. Pour alléger les notations, nous posons $f_n=\varphi_n-\varphi$. Nous avons

subEqsSGsGGih

$$\begin{aligned}\|\hat{f}\|_{\alpha,\beta} &= \|\xi^\beta\partial^\alpha\hat{f}\|_\infty && (29.79a)\\ &= \|\widehat{\partial^\beta M_\alpha f}\|_\infty \text{ lemme 29.19.} && (29.79b)\\ &\leqslant \|\partial^\beta M_\alpha f\|_1 && (29.79c)\end{aligned}$$

4. Définition 27.233, avec rappel que $p_{\alpha\beta}(\varphi)=\|x\mapsto x^\beta(\partial^\alpha\varphi)(x)\|_\infty$.
5. Nous verrons, avec la formule d'inversion dans la proposition 29.26 que c'est même une bijection.
6. Pour rappel, en dimension infinie, il n'est pas garanti qu'une application linéaire soit continue.

La convergence $f_n \xrightarrow{\mathscr{S}} 0$ nous dit entre autres que $\partial^\beta M_\alpha f_n \xrightarrow{\mathscr{S}} 0$; en particulier la proposition 27.241 nous dit que $\partial^\beta M_\alpha f_n \xrightarrow{L^1} 0$, ce qui signifie, par les majorations (29.79) que

$$\|\hat{f}_n\|_{\alpha,\beta} \leqslant \|\partial^\beta M_\alpha f_n\|_1 \to 0, \tag{29.80}$$

ce qui prouve la continuité de transformée de Fourier dans $\mathscr{S}(\mathbb{R}^d)$.

(3) **Bijection** Une preuve peut être trouvée dans [711].

□

PROPooMVQMooGYAzSX

Proposition 29.21 ([1]).
Soit $\varphi \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$ et la transformée de Fourier partielle

$$\tilde{\varphi}(x,k) = \int_{\mathbb{R}^m} e^{-iky}\varphi(x,y)dy. \tag{29.81}$$

Alors $\tilde{\varphi} \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$.

Démonstration. Il s'agit de reprendre les étapes de la partie correspondante de la preuve de la proposition 29.20. Soient des multiindices α, α', β et β' où α et β se réfèrent à la variable x tandis que α' et β' se réfèrent à la variable k.

Vu que la multiplication par $k^{\beta'}$ commute avec ∂^α nous avons

$$x^\beta k^{\beta'}\partial^\alpha\partial^{\alpha'}\tilde{\varphi}(x,k) = x^\beta k^{\beta'}\partial^\alpha(-i)^{|\alpha'|}\widetilde{M_{\alpha'}\varphi}(x,k) = (-i)^{|\alpha'|+|\beta'|}x^\beta\partial^\alpha\partial^{\beta'}\widetilde{M_{\alpha'}\varphi}(x,k). \tag{29.82}$$

D'autre part nous avons $\partial^\alpha\tilde{\varphi} = \widetilde{\partial^\alpha\varphi}$ parce que la fonction $\partial_x\varphi$ étant Schwartz, la fonction

$$G(y) = \sup_{x\in\mathbb{R}^n} |(\partial_x\varphi)(x,y)| \tag{29.83}$$

est dans $L^1(\mathbb{R}^m)$ par le corolaire 27.237. Par conséquent le théorème 17.19 permet de permuter la dérivée et l'intégrale dans

$$\frac{\partial}{\partial x}\tilde{\varphi}(x,k) = \frac{\partial}{\partial x}\int_{\mathbb{R}^m} e^{-iky}\varphi(x,y)dy. \tag{29.84}$$

Dans le même ordre d'esprit des difficultés de permutation de limites nous avons $M_\beta\tilde{\varphi} = \widetilde{M_\beta\varphi}$.

D'autre part nous avons encore $\|\tilde{\varphi}\|_\alpha < \infty$ parce que

$$|\tilde{\varphi}(x,k)| \leqslant \int_{\mathbb{R}^m} |\varphi(x,y)|dy \leqslant \sup_x \int_{\mathbb{R}^m} |\varphi(x,y)|dy \leqslant \int_{\mathbb{R}^m} |\sup_x \varphi(x,y)|dy < \infty \tag{29.85}$$

parce que φ est Schwartz et le corolaire 27.237 donne l'intégrabilité.

Donc nous avons

$$p_{(\alpha\alpha'),(\beta\beta')}(\tilde{\varphi}) = \|\partial^{\beta'}\widetilde{M_{\alpha'}M_\beta}\partial^\alpha\varphi\|_\infty < \infty. \tag{29.86}$$

Cela prouve que $\tilde{\varphi}$ est Schwartz. □

29.3.1 Quelques transformées de Fourier

LEMooPAAJooCsoyAJ

Lemme 29.22 (Transformée de Fourier de la Gaussienne [712]).
La transformée de Fourier de

$$\begin{aligned} g_\epsilon\colon \mathbb{R}^d &\to \mathbb{R} \\ x &\mapsto e^{-\epsilon\|x\|^2} \end{aligned} \tag{29.87}$$

est donnée par

$$\hat{g}_\epsilon(\xi) = \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-\|\xi\|^2/4\epsilon} \tag{29.88}$$

Démonstration. Nous commençons par la fonction $g(x) = e^{-\|x\|^2/2}$ et nous prouvons que sa transformée de Fourier est $\hat{g}(\xi) = (2\pi)^{d/2}g(\xi)$.

(1) **Réduction à la dimension 1** La fonction g est dans l'espace de Schwartz. Par le théorème de Fubini,

$$\hat{g}(\xi) = \int_{\mathbb{R}^d} \prod_{k=1}^{d} e^{-x_k^2} e^{-i\xi_k x_k} dx = \prod_{k=1}^{d} \int_{\mathbb{R}} e^{-t^2/2} e^{-\xi_k x} dt = \prod_{k=1}^{d} \hat{f}(\xi_k) \tag{29.89}$$

où f est la fonction d'une variable

$$f(x) = e^{-x^2/2}. \tag{29.90}$$

Notons que $f \in \mathscr{D}(\mathbb{R})$.

(2) **Une équation différentielle**

Voyons l'équation différentielle satisfaite par la transformée de Fourier $\hat{f}$ de la fonction (29.90). Grâce au lemme 29.19 nous trouvons l'équation différentielle [7]

$$\xi \hat{f}(\xi) + (\hat{f})'(\xi) = 0. \tag{29.91}$$

C'est le moment d'utiliser le théorème de Cauchy-Lipschitz (17.39), appliqué à la fonction $f(t, y) = -ty$ qui est Lipschitz et continue au problème

$$\begin{cases} y' + ty = 0 & \text{(29.92a)} \\ y(0) = y_0 & \text{(29.92b)} \end{cases}$$

possède une unique solution maximale, en l'occurrence $y(x) = y_0 e^{-x^2/2}$. En ce qui concerne la condition initiale nous avons

$$\hat{f}(0) = \int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}. \tag{29.93}$$

par l'exemple 14.307. Donc

$$\hat{f}(\xi) = \sqrt{2\pi} e^{-\xi^2/2}. \tag{29.94}$$

En reformant le produit (29.89) nous concluons.

Nous passons maintenant à la fonction g_ϵ. Nous pouvons écrire g_ϵ sous la forme

$$g_\epsilon(x) = g(\sqrt{2\epsilon} x). \tag{29.95}$$

Utilisant successivement la transformée de Fourier de g que nous venons de calculer et 29.9 (facteur d'échelle) nous trouvons

$$\begin{aligned} \hat{g}(\xi) &= (2\pi)^{d/2} g(\xi) && \text{(29.96a)} \\ \hat{g}_\epsilon(\xi) &= (2\epsilon)^{-d/2} \hat{g}\big(\xi/\sqrt{2\epsilon}\big) && \text{(29.96b)} \\ &= \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-|\xi|^2/4\epsilon}.. && \text{(29.96c)} \end{aligned}$$

Nous voyons que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ (c'était gagné d'avance par la proposition 29.20). □

Lemme 29.23.
Si $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ alors la suite

$$\rho_n = \frac{1}{(2\pi)^d} \hat{g}_{1/n} \tag{29.97}$$

est une suite régularisante (définition 29.17).

7. Une façon alternative, plus directe de déduire cette équation différentielle sera donnée dans l'exemple 29.25.

Démonstration. Nous savons déjà la transformée de Fourier de g_ϵ par le lemme 29.22. Nous montrons que la suite ρ_n est régularisante. Nous avons $\hat{g}_\epsilon \in L^1(\mathbb{R}^d)$ et $\hat{g}_\epsilon \geqslant 0$ ainsi que $\lim_{\epsilon\to 0}\int_{B(0,\alpha)}\hat{g}_\epsilon = 0$ pour tout α. Il y a seulement un couac avec la norme. Nous calculons $\int_{\mathbb{R}^d}\hat{g}_\epsilon(\xi)d\xi$ avec la forme (29.96c). En utilisant sauvagement Fubini [8] pour séparer les intégrales et en effectuant le changement de variable $u = t/(2\sqrt{\epsilon})$ nous calculons :

$$\int_{\mathbb{R}^d} e^{-|\xi|^2/4\epsilon}d\xi = \prod_{k=1}^{d}\int_{\mathbb{R}} e^{-t^2/4\epsilon}dt \tag{29.98a}$$

$$= \prod_{k=1}^{d} 2\sqrt{\epsilon}\int_{\mathbb{R}} e^{-u^2}du \tag{29.98b}$$

$$= \prod_{k=1}^{d} 2\sqrt{\epsilon}\sqrt{\pi} \tag{29.98c}$$

$$= 2^d(\pi\epsilon)^{d/2}. \tag{29.98d}$$

Nous avons utilisé l'exemple 14.307 pour le calcul de l'intégrale gaussienne. Avec tout cela nous avons

$$\int_{\mathbb{R}^d}\hat{g}_\epsilon = (2\pi)^d. \tag{29.99}$$

Donc $\frac{1}{(2\pi)^d}\hat{g}_{1/n}$ est une suite régularisante. □

Le corolaire suivant regroupe les résultats à propos des suites régularisantes, leur utilité et leur existence.

CORooQLELooUjzIoM

Corolaire 29.24.
Si la suite régularisante ρ_n est dans $L^1(\mathbb{R}^d)\cap C^\infty(\mathbb{R}^d)$ alors pour $f \in L^p(\mathbb{R}^d)$ en posant $f_n = \rho_n * f$ nous avons

(1) $f_n \in C^\infty(\mathbb{R}^d)\cap L^p(\mathbb{R}^d)$

(2) $f_n \xrightarrow{L^p} f$

De plus, de telles suites existent.

Démonstration. Le fait que f_n soit de classe C^∞ est le corolaire 27.72, et la convergence est la proposition 29.18(2).

De telles suites existent, par exemple celle donnée par le lemme 29.23. □

EXooLMXKooFcAZGR

Exemple 29.25 ([104])**.**
Soit la fonction $g_\epsilon(x) = e^{-\epsilon x^2}$. Sa transformée de Fourier a été vue dans le lemme 29.22 en utilisant le lemme de transfert 29.19. Nous nous proposons ici de déduire de façon directe l'équation différentielle vérifiée par la transformée de Fourier de g_ϵ.

Nous posons

$$I(k) = \int_{\mathbb{R}} e^{-ikx}e^{-\epsilon x^2}dx. \tag{29.100}$$

et nous considérons la fonction

$$f(k,x) = e^{-ikx}e^{-\epsilon x^2}. \tag{29.101}$$

Elle est de classe C^1 par rapport à k, et intégrable en x pour chaque k. De plus sa dérivée

$$(\partial_k f)(k,x) = -ixe^{-ikx}e^{-\epsilon x^2} \tag{29.102}$$

8. Le pauvre !

vérifie $|\partial_k f| \leqslant xe^{-\epsilon x^2}$. La dérivée est donc majorée (uniformément en k) par une fonction intégrable. Le théorème 17.19 permet de permuter la dérivée et l'intégrale :

$$I'(k) = \int_{\mathbb{R}} -ixe^{-ikx}e^{-\epsilon x^2}dx \tag{29.103a}$$
$$= i\int_{\mathbb{R}} e^{-ikx}\frac{1}{2\epsilon}\frac{d}{dx}\left(e^{-\epsilon x^2}\right)dx \tag{29.103b}$$
$$= \frac{-i}{2\epsilon}\int_{\mathbb{R}} \frac{d}{dx}\left(e^{-ikx}\right)e^{-\epsilon x^2}dx \qquad \text{par partie} \tag{29.103c}$$
$$= \frac{-k}{2\epsilon}\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon x^2}dx \tag{29.103d}$$
$$= \frac{-k}{2\epsilon}I(k). \tag{29.103e}$$

D'où l'équation différentielle $I'(k) = -\frac{k}{2\epsilon}I(k)$. △

29.3.2 Formule d'inversion

PROPooLWTJooReGlaN

Proposition 29.26 (Formule d'inversion[468, 104])**.**
À propos d'inversion de la transformée de Fourier.

ITEMooLVOTooUDJSWt

(1) Si $f \in \mathscr{S}(\mathbb{R})$, alors nous avons la formule d'inversion

EQooHIDAooHARdNZ

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ikx}\hat{f}(k)dk. \tag{29.104}$$

ITEMooWINLooJWcDIX

(2) Nous avons, pour $f \in \mathscr{S}(\mathbb{R})$,

$$f(x) = \frac{1}{2\pi}\mathcal{F}(\hat{f})(-x). \tag{29.105}$$

ITEMooCZYMooRvKTfS

(3) L'application

$$\mathcal{F}\colon \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R}) \tag{29.106}$$

est une bijection continue.

(4) Cette formule peut d'écrire de plusieurs autres façons : pour tout $f \in \mathscr{S}(\mathbb{R}^d)$ nous avons

EQooWBZTooPeBNeb

$$f(x) = \frac{1}{2\pi}\mathcal{F}(\hat{f})(-x). \tag{29.107a}$$
$$\mathcal{F}^{-1}(f)(\xi) = -\frac{1}{2\pi}\mathcal{F}(f)(-\xi). \tag{29.107b}$$

Démonstration. Le fait que la transformée de Fourier sur $\mathscr{S}(\mathbb{R}^d)$ prenne ses valeurs dans $\mathscr{S}(\mathbb{R}^d)$ est déjà prouvé dans 29.20. Nous commençons maintenant la preuve de (1).

Soit $f \in \mathscr{S}(\mathbb{R}^d)$. Pour $\epsilon > 0$ nous posons

$$f_\epsilon(k) = e^{-\epsilon k^2}e^{ikx}\hat{f}(k). \tag{29.108}$$

Nous allons calculer

$$\lim_{\epsilon\to 0}\int_{\mathbb{R}} e^{-\epsilon k^2}e^{ikx}\hat{f}(k)dk \tag{29.109}$$

de deux façons ; d'abord avec la convergence dominée, et ensuite avec Fubini.

(1) **Premier calcul : convergence dominée** D'abord en utilisant directement le théorème de la convergence dominée 14.207. La fonction $\hat{f}$ est dans $\mathscr{S}(\mathbb{R})$ (théorème 29.20) et par conséquent $f_\epsilon \in L^1(\mathbb{R})$ parce que le facteur $e^{-\epsilon k^2}$ ne va certainement pas empêcher de converger. De plus $|f_\epsilon| \leqslant |\hat{f}|$ et $\hat{f} \in L^1$. Le théorème est de la convergence dominée est applicable et

EQooYIYGooXYubbW

$$\lim_{\epsilon\to 0}\int_{\mathbb{R}} e^{-\epsilon k^2}e^{ikx}\hat{f}(k)dk = \int_{\mathbb{R}} e^{ikx}\hat{f}(k)dk. \tag{29.110}$$

(2) **Deuxième calcul : Fubini**

Pour le deuxième calcul nous allons faire appel à Fubini [9] pour la fonction

$$\begin{aligned} u\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (k,y) &\mapsto e^{ik(x-y)}e^{-\epsilon k^2}f(y). \end{aligned} \tag{29.111}$$

D'abord nous nous assurons que $u\in L^1(\mathbb{R}\times\mathbb{R})$ par le corolaire 14.303, et ensuite nous utilisons le théorème de Fubini 14.304 pour manipuler les intégrales (et en particulier les permuter).

(2a) $u\in L^1(\mathbb{R}\times\mathbb{R})$ Dans un premier temps nous avons :

$$\int_{\mathbb{R}}\int_{\mathbb{R}}|e^{ik(x-y)}e^{-\epsilon k^2}f(y)|dy\,dk \leqslant \int_{\mathbb{R}}e^{-\epsilon k^2}\Big[\int_{\mathbb{R}}|f(y)|dy\Big]dk<\infty \tag{29.112}$$

parce que f étant dans $\mathscr{S}(\mathbb{R})$, l'intégrale intérieure se réduit à un nombre. Nous savons maintenant que $u\in L^1(\mathbb{R}\times\mathbb{R})$ grâce au corolaire 14.303.

(2b) **Calcul**

Nous pouvons alors calculer un peu ...

$$\begin{aligned}
\int_{\mathbb{R}}e^{ikx}e^{-\epsilon k^2}\hat{f}(k)dk &= \int_{\mathbb{R}}\int_{\mathbb{R}}e^{ikx}e^{-\epsilon k^2}e^{-iky}f(y)dy\,dk && (29.113a)\\
&= \int_{\mathbb{R}}\Big[\int_{\mathbb{R}}e^{ik(x-y)}e^{-\epsilon k^2}f(y)dk\Big]dy && (29.113b)\\
&= \int_{\mathbb{R}}f(y)\Big[\int_{\mathbb{R}}e^{ik(x-y)}e^{-\epsilon k^2}dk\Big]dy && (29.113c)\\
&= \int_{\mathbb{R}}f(y)\hat{g}_\epsilon(y-x)dy && (29.113d)\\
&= \sqrt{\frac{\pi}{\epsilon}}\int_{\mathbb{R}}f(y)e^{-(y-x)^2/4\epsilon}dy && (29.113e)\\
&= 2\sqrt{\epsilon}\sqrt{\frac{\pi}{\epsilon}}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}dt && (29.113f)\\
&= 2\sqrt{\pi}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}dt && (29.113g)\\
& && (29.113h)
\end{aligned}$$

Justifications :

- Pour (29.113b), c'est Fubini 14.304.
- Pour (29.113d), nous avons reconnu dans le crochet la transformée de Fourier de la fonction $g_\epsilon\colon x\mapsto e^{-\epsilon x^2}$.
- Pour (29.113e), nous utilisons la transformée de Fourier de g_ϵ donnée dans le lemme 29.22.
- Pour (29.113f), nous avons effectué le changement de variables $t=(y-x)/(2\sqrt{\epsilon})$ qui donne $dt=dy/2\sqrt{\epsilon}$.

(3) **Second passage à la limite**

Nous avons obtenu l'égalité

$$\int_{\mathbb{R}}e^{ikx}e^{-\epsilon k^2}\hat{f}(k)dk = 2\sqrt{\pi}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}dt, \tag{29.114}$$

et nous voudrions passer à la limite $\epsilon\to 0$. Le membre de gauche est déjà fait en (29.110).

9. Parce qu'il est toujours plus simple de refiler le boulot aux autres que de le faire soi-même...pauvre Fubini !

Pour le membre de droite, la fonction f étant Schwartz (en particulier bornée), nous pouvons effectuer la majoration

$$f(x+2\sqrt{\epsilon}t)e^{-t^2}\leqslant\|f\|_\infty e^{-t^2},\tag{29.115}$$

qui est une fonction intégrable de t. Nous avons donc le droit de permuter la limite $\epsilon\to 0$ et l'intégrale dans le calcul suivant :

$$\lim_{\epsilon\to0}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}=\int_{\mathbb{R}}f(x)e^{-t^2}dt.\tag{29.116}$$

(4) **Fin**

Nous avons maintenant les limites des deux membres de (29.114). Récrivons :

$$\lim_{\epsilon\to0}\int_{\mathbb{R}}e^{ikx}e^{-\epsilon k^2}\hat{f}(k)dk=\lim_{\epsilon\to0}2\sqrt{\pi}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}dt,\tag{29.117}$$

À gauche nous avons déjà la limite depuis (29.110), et à droite nous obtenons

$$\lim_{\epsilon\to0}2\sqrt{\pi}\int_{\mathbb{R}}f(x+2\sqrt{\epsilon}t)e^{-t^2}dt=2\sqrt{\pi}\int_{\mathbb{R}}f(x)e^{-t^2}dt=2\sqrt{\pi}f(x)\sqrt{\pi}=2\pi f(x)\tag{29.118}$$

où nous avons utilisé l'intégrale gaussienne faite dans l'exemple 14.307.

En remettant tout ensemble,

$$2\pi f(x)=\lim_{\epsilon\to0}\int_{\mathbb{R}}e^{-\epsilon k^2}e^{ikx}\hat{f}(k)dk=\int_{\mathbb{R}}e^{ikx}\hat{f}(k)dk,\tag{29.119}$$

ce qu'il fallait prouver.

Le plus gros est fait ; nous avons prouvé

$$f(x)=\frac{1}{2\pi}\int_{\mathbb{R}}e^{ikx}\hat{f}(k)dk.\tag{29.120}$$

EQooYVUZooZXQoEH

Pour (2), c'est simplement une reformulation de cela.

Nous prouvons maintenant (3).

(1) **La transformée de Fourier est injective** Vu qu'elle est linéaire, il suffit de démontrer que si $\mathcal{F}(f)=0$, alors $f=0$. Si $\hat{f}=0$, alors la formule (29.120) donne immédiatement $f=0$.

(2) **La transformée de Fourier est surjective** Soit $f\in\mathscr{S}(\mathbb{R})$. Nous devons trouver $g\in\mathscr{S}(\mathbb{R})$ tel que $\hat{g}=f$. La formule $\mathcal{F}(\hat{f})(-x)=2\pi f(x)$ nous incite à essayer

$$g(\xi)=-\frac{1}{2\pi}\hat{f}(-\xi).\tag{29.121}$$

Calculons $\mathcal{F}(g)$:

$$\begin{aligned}\mathcal{F}(g)(x)&=-\int_{\mathbb{R}}e^{-i\xi x}g(\xi)d\xi&(29.122a)\\&=-\int_{\mathbb{R}}e^{-i\xi x}\frac{1}{2\pi}\hat{f}(-\xi)d\xi&(29.122b)\\&=\frac{1}{2\pi}\int_{\mathbb{R}}e^{itx}\hat{f}(t)dt&(29.122c)\\&=\frac{1}{2\pi}\mathcal{F}(\hat{f})(-x)&(29.122d)\\&=f(x).&(29.122e)\end{aligned}$$

Pour la dernière ligne, nous avons utilisé $\mathcal{F}(\hat{f})(-x)=2\pi f(x)$.

□

CORooAZLZooSviTej

Corolaire 29.27.
Nous avons la formule

$$\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-ikx}f(x)dx\,dk = 2\pi f(0). \tag{29.123}$$

EQooRJXRooElEMAa

Démonstration. Poser $x = 0$ dans l'équation (29.104). □

29.28.
Les physiciens qui n'ont que rarement peur écrivent souvent la formule (29.123) sous la forme

$$\int_{\mathbb{R}} e^{-ikx}dk = \delta(x) \tag{29.124}$$

où δ serait la fonction de Dirac qui vaut zéro partout sauf en $x = 0$ où elle vaudrait l'infini, mais pas n'importe quel infini ; juste celui qu'il faut pour que sont intégrale vaille 1.

LemYYjFZSa

Lemme 29.29.
Si $\phi \in \mathscr{S}(\mathbb{R}\times\mathbb{R}^n)$, alors

$$\partial_t\hat{\phi} = \widehat{\partial_t\phi} \tag{29.125}$$

où le chapeau dénote la transformée de Fourier par rapport à la variable dans $\mathbb{R}^n$ et non par rapport à celle dans $\mathbb{R}$. Le t par contre est la variable dans $\mathbb{R}$.

Démonstration. Par définition de la transformée de Fourier nous avons

$$(\partial_t\hat{\phi})(t,\xi) = \frac{\partial}{\partial t}\int_{\mathbb{R}^n}\phi(t,x)e^{-ix\xi}dx. \tag{29.126}$$

Notre but est de permuter l'intégrale et la dérivée en utilisant le théorème 17.19. Il nous faut une fonction $G\colon \mathbb{R}^n \to \mathbb{R}$ qui soit intégrable sur $\mathbb{R}^n$ et telle que

$$\left|\frac{\partial\phi}{\partial t}\phi(t,x)\right| \leqslant G(x) \tag{29.127}$$

pour tout $t \in B(t_0,\delta)$. Étant donné que la fonction $\partial_t\phi$ est tout autant Schwartz que ϕ elle-même nous pouvons alléger les notations et chercher une fonction G qui convient pour toute fonction $\varphi \in \mathscr{S}(\mathbb{R}\times\mathbb{R}^n)$. Soit la fonction

$$G(x) = \sup_{t\in B(t_0,\delta)}|\varphi(t,x)|. \tag{29.128}$$

Pour tout multiindice α nous avons alors

$$\sup_{x\in\mathbb{R}^n}\left|x^\alpha G(x)\right| \leqslant \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^n}\left|x^\alpha\varphi(t,x)\right| \leqslant M_\alpha \in \mathbb{R}. \tag{29.129}$$

Grâce à la proposition 27.236, cela signifie que φ décroît plus vite que n'importe quel polynôme ; G est donc intégrable sur $\mathbb{R}^n$. □

29.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$

La théorie des transformées de Fourier est intéressante sur $L^2(\mathbb{R}^d)$ parce qu'elle y donne une isométrie. Nous allons la donner avec des fonctions à valeurs dans $\mathbb{C}$.

Remarque 29.30.
Une remarque qui vaut ce qu'elle vaut, mais si u est une classe de fonction pour la relation $u \sim v$ si et seulement si $u(x) = v(x)$ pour presque tout v alors l'intégrale

$$\hat{u}(\xi) = \int_{\mathbb{R}^d} u(x)e^{ix\xi}dx \tag{29.130}$$

ne dépend pas du choix du représentant. Nous pouvons donc parfaitement parler de transformée de Fourier d'une classe de fonctions.

29.4.1 Le problème

Nous avons défini en général la transformée de Fourier d'une fonction $f\colon \mathbb{R} \to \mathbb{C}$ par la formule

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{i\xi x} f(x) dx \tag{29.131}$$

EQooOPLJooEkqRuC

tant que cette intégrale existe.

Il se fait que cette intégrale n'existe pas toujours pour des fonctions dans $L^2(\mathbb{R})$. Donc nous devons faire mieux pour définir la transformée de Fourier sur L^2.

EXooSWCHooTdHTsl

Exemple 29.31 ([1])**.**
Prenons la fonction

$$f(x) = \begin{cases} 0 & \text{si } x < 1 \\ \frac{1}{x} & \text{si } x \geqslant 1. \end{cases} \tag{29.132}$$

Vu que l'intégrale $\int_1^\infty \frac{1}{x^2} dx$ existe et est finie (proposition 14.300(2)), la fonction f est dans $L^2(\mathbb{R})$.

Cependant l'intégrale (29.131) n'existe pas. Pour nous convaincre de cela, nous pouvons simplement nous souvenir de la définition d'une intégrale à valeurs dans un espace vectoriel (définition 14.197). Nous fixons $\xi \in \mathbb{R}$ et nous posons $g(x) = f(x)e^{i\xi x}$.

Bien évidemment, $|g(x)| = \frac{1}{x}$ sur $]1, \infty[$. Donc $\int_{\mathbb{R}} |g| = \infty$, et la fonction g n'est pas intégrable. Fin de l'histoire.

Nous pouvons toujours essayer de comprendre mieux. Vu que $\int_{\mathbb{R}} |g| = \infty$, la proposition 14.198 nous dit qu'au moins une des intégrales parmi

$$\int g_{re}^+, \int g_{im}^+, \int g_{re}^-, \int g_{im}^- \tag{29.133}$$

est égale à $+\infty$.

Note qu'en travaillant un peu, on se convainc qu'en réalité, elles divergent toutes les quatre. △

29.4.2 Extension de $L^1 \cap L^2$ vers L^2

THOooJLCDooAjTvJf

Théorème 29.32 (Extention de la transformée de Fourier vers $L^2(\mathbb{R}^d)$[712])**.**
Soit $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Alors

(1) Nous avons $\mathcal{F}(f) \in L^2$ et $\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2}$.

(2) L'application $\mathcal{F}\colon L^1 \cap L^2 \to L^2$ peut être étendue en une application $\mathcal{F}\colon L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ vérifiant

$$\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2} \tag{29.134}$$

pour tout $f \in L^2(\mathbb{R}^d)$.

Démonstration. Le fait que $f \in L^1$ implique $\|\mathcal{F}(f)\|_\infty \leqslant \|f\|_1$ (c'est le lemme 29.12). En particulier, $|\mathcal{F}(f)(\xi)|^2$ est majoré et l'intégrale

$$\clubsuit = \int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi \tag{29.135}$$

existe et est finie.

(1) **Découper l'intégrale** Dans un premier temps nous développons les intégrales. Dans les égalités suivantes, $x\xi$ est le produit scalaire $x \cdot \xi$ dans $\mathbb{R}^d$.

$$\clubsuit = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \overline{f(x)} e^{ix\xi} dx \right) \left(\int_{\mathbb{R}^d} f(y) e^{-y\xi} \right) e^{-\epsilon|\xi|^2} d\xi \tag{29.136a}$$

$$= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) e^{i\xi(x-y)} dx dy \right] e^{-\epsilon|\xi|^2} d\xi. \tag{29.136b}$$

Nous avons utilisé le théorème de Fubini pour regrouper les intégrales [10]. Vu que $f \in L^1(\mathbb{R}^d)$, la fonction $(x,y,\xi) \mapsto f(x)f(y)e^{-\epsilon|\xi|^2}$ est dans $L^1(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ et le théorème de Fubini 14.304 avec $\Omega_1 = \mathbb{R}^d \times \mathbb{R}^d$ et $\Omega_2 = \mathbb{R}^d$ nous permet de permuter les intégrales pour avoir

$$\clubsuit = \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) \left[\int_{\mathbb{R}^d} e^{i\xi(x-y)} e^{-\epsilon|\xi|^2} d\xi \right] dxdy. \tag{29.137}$$

EQooSUYWooCmtFeF

(2) **Discuter de cette gaussienne** En posant

$$g(x) = e^{-|x|^2/2} \tag{29.138a}$$

$$g_\epsilon(x) = g(\sqrt{2\epsilon}x) = e^{-\epsilon|x|^2} \tag{29.138b}$$

nous avons $g_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ et le lemme 29.23 nous autorise à écrire

$$\hat{g}(\xi) = (2\pi)^{d/2} g(\xi) \tag{29.139a}$$

$$\hat{g}_\epsilon(\xi) = \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-|\xi|^2/4\epsilon} \tag{29.139b}$$

Nous voyons que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$ (c'était gagné d'avance par la proposition 29.20) et que $\hat{g}_\epsilon$ est une fonction paire (encore une fois, c'était gagné d'avance parce que la transformée de Fourier d'une fonction paire est paire).

Tout cela pour dire que l'intégrale entre crochet dans (29.137) est $\hat{g}_\epsilon(y-x) = \hat{g}_\epsilon(x-y)$, et donc

$$\clubsuit = \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{f(x)} f(y) \hat{g}_\epsilon(x-y) dxdy. \tag{29.140}$$

Encore une fois le théorème de Fubini permet de séparer les intégrales et de calculer l'intégrale sur y en premier. Vu que $f \in L^1$ et que $\hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d)$, le produit de convolution $f * \hat{g}_\epsilon$ est un élément de $\mathscr{S}(\mathbb{R}^d)$ par la proposition 27.245. Nous avons donc

$$\clubsuit = \int_{\mathbb{R}^d} \overline{f(x)} (f * \hat{g}_\epsilon)(x) dx. \tag{29.141}$$

Là, nous reconnaissons un produit scalaire dans $L^2(\mathbb{R}^d)$, et donc

$$\int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi = \langle f, f * \hat{g}_\epsilon \rangle_{L^2(\mathbb{R}^d)}. \tag{29.142}$$

EQooWIHNooHutHlS

Notons que tout a un sens : $f \in L^2(\mathbb{R}^d)$ et $f * \hat{g}_\epsilon \in \mathscr{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$.

(3) **Suite régularisante**

Nous prenons la suite régularisante du lemme 29.23 donnée par

$$\rho_n = \frac{1}{(2\pi)^d} \hat{g}_{1/n}. \tag{29.143}$$

(4) **Première conclusion** Nous reprenons (29.142)

$$\int_{\mathbb{R}^d} |\hat{f}|^2 e^{-|\xi^2|/n} d\xi = \langle f, f * \hat{g}_{1/n} \rangle_{L^2(\mathbb{R}^d)} = (2\pi)^d \langle f, f * \rho_n \rangle. \tag{29.144}$$

En prenant la limite $n \to \infty$ nous trouvons

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} |\hat{f}|^2 e^{-\epsilon\xi^2} d\xi = (2\pi)^d \|f\|^2. \tag{29.145}$$

Pour effectuer la limite du membre de gauche nous devons remarquer qu'en posant

$$g_n(\xi) = |\hat{f}(\xi)| e^{-|\xi|^2/n}, \tag{29.146}$$

10. Dans la suite nous allons encore utiliser Fubini quelques fois pour regrouper et dégrouper des intégrales.

nous avons une suite décroissante de fonctions (c'est-à-dire que à ξ fixé, $g_n(\xi)$ est décroissant en n). Par ailleurs ces fonctions sont toujours à valeurs dans $[0,\infty]$ et nous pouvons utiliser le théorème de la convergence monotone 14.178 pour permuter la limite et l'intégrale. Au final :

$$\|\hat{f}\|_{L^2} = (2\pi)^d \|f\|_{L^2}. \tag{29.147}$$

En ce qui concerne l'extension, soit $f \in L^2(\mathbb{R}^d)$ et une suite (f_n) dans $L^1 \cap L^2$ telle que $f_n \xrightarrow{L^2} f$.

(1) **Existence d'une telle suite** Si $f \in L^2(\mathbb{R}^d)$, alors nous pouvons poser

$$f_n(x) = f(x)e^{-|x|^2/n^2}. \tag{29.148}$$ EQooHGJYooJsmxoX

Par l'inégalité de Hölder (27.81) nous avons $f_n \in L^1(\mathbb{R}^d)$; de plus $f_n \in L^2(\mathbb{R}^d)$ parce que pour tout x nous avons $|f_n(x)| \leqslant |f(x)|$. Montrons que $f_n \xrightarrow{L^2} f$. Nous avons

$$\|f_n - f\|_{L^2}^2 = \int_{\mathbb{R}^d} |f(x)(1 - e^{-|x^2|/n^2})|^2 dx. \tag{29.149}$$

Nous voulons prendre la limite $n \to \infty$. Pour ce faire à droite nous remarquons que $e^{-|x|^2/n^2}$ est majoré par 1 ; ce qui se trouve dans l'intégrale est donc majoré (uniformément en n) par $|f(x)|^2$, qui est une fonction L^1 parce que f est L^2. Le théorème de la convergence dominée 14.207 nous permet alors de permuter la limite et l'intégrale, ce qui donne

$$\lim_{n\to\infty} \|f_n - f\|_2^2 = \int_{\mathbb{R}^d} \lim_{n\to\infty} |f(x)(1 - e^{-|x|^2/n^2})|^2 dx = 0. \tag{29.150}$$

(2) **Définition de $\mathcal{F}: L^2 \to L^2$**
La suite (f_n) est une suite convergence dans L^2, et elle est donc de Cauchy. De plus pour chaque n, m nous avons

$$\|\hat{f}_n - \hat{f}_m\| = (2\pi)^d \|f_n - f_m\|. \tag{29.151}$$

Nous voyons donc que la suite $(\hat{f}_n)$ est également de Cauchy, dans l'espace $L^2(\mathbb{R}^d)$ qui est complet (lemme 27.92). Nous posons

$$\hat{f} = \lim_{n\to\infty} \hat{f}_n. \tag{29.152}$$

(3) **Indépendance aux choix** Nous devons montrer que la définition de $\hat{f}$ ne dépend pas de la suite approximant f dans $L^1 \cap L^2$. Soient dans deux suites $f_n \xrightarrow{L^2} f$ et $g_n \xrightarrow{L^2} f$ telles que $\hat{f}_n \xrightarrow{L^2} F$ et $\hat{g}_n \xrightarrow{L^2} G$. Alors

$$\|\hat{f}_n - \hat{g}_n\| = (2\pi)^d \|f_n - g_n\| \leqslant (2\pi)^d \|f_n - f\| + (2\pi)^d \|g_n - f\| \to 0. \tag{29.153}$$

Par conséquent $(\hat{f}_n - \hat{g}_n)_n$ est une suite qui converge vers zéro. Par unicité de la limite, $F = G$.

□

Remarque 29.33.
Une autre suite possible, à la place de (29.148), est

$$f_n(x) = f(x)\mathbb{1}_{|x|<n}. \tag{29.154}$$

C'est-à-dire la fonction f limitée à une boule de rayon n autour de 0.

29.4.3 Une formule de Leibniz

PROPooBTXLooFhuYSs

¡¡ Avertissement/question au lecteur !! 29.34
Si f et g sont des fonctions comme il faut et si P est un polynôme, je suis quasiment certain que nous avons la formule

$$P * (fg) = (Pf) * g + f * (Pg). \tag{29.155}$$

La raison est que, à travers la transformée de Fourier, P est seulement constituée de dérivation. Cela est donc une formule de Leibniz.

Si vous avez un énoncé exact et une preuve, je suis preneur.

Chapitre 30

Distributions

Nous donnons ici une partie de la théorie sur les distributions. L'utilisation des distributions dans le cadre des équations différentielles est mise dans le chapitre sur les équations différentielles, section 32.11.

PropAAjSURG

Proposition 30.1 ([713]).
Soient un ouvert Ω de $\mathbb{R}$ et une fonction intégrable $f\colon \big(\Omega, \mathcal{B}or(\Omega), \lambda\big) \to \mathbb{C}$ telle que

$$\int_\Omega f\varphi = 0 \tag{30.1}$$

pour toute fonction $\varphi \in \mathscr{D}(\Omega)$. Alors $f = 0$ presque partout sur Ω.

Démonstration. Nous commençons par prouver que f est nulle sur tout compact de Ω. Soit un compact K de Ω. Le lemme d'Urysohn 15.165 nous donne une fonction θ à support compact qui vaut 1 sur K.

Nous considérons une suite régularisante (ϕ_k) de fonctions toujours strictement positives (par exemple celle du lemme 29.23). Vu que $f\theta$ est à support compact, elle est dans $L^p(\Omega)$ et le corolaire 29.24 s'applique :

$$\phi_k * (\theta f) \to \theta f. \tag{30.2}$$

Mais, x et k étant fixés, nous avons

$$\big(\phi_k * (\theta f)\big)(x) = \int_{\mathbb{R}} \phi_k(x-t)\theta(t)f(t)dt. \tag{30.3}$$

EQooIUIMooMTAHCY

La fonction

$$t \mapsto \phi_k(x-t)\theta(t) \tag{30.4}$$

étant à support compact, l'hypothèse à propos de f fait que l'intégrale (30.3) est nulle :

$$\phi_k * (\theta f) = 0 \tag{30.5}$$

pour tout k. En prenant la limite $k \to \infty$,

$$\theta f = 0. \tag{30.6}$$

Vu que $\theta(x) = 1$ pour tout $x \in K$, nous avons $f(x) = 0$ pour tout $x \in K$.

Nous avons démontré que f était nulle sur tout compact de Ω.

Nous considérons maintenant une suite exhaustive (K_n) de compacts (lemme 7.306). La fonction f est nulle sur chaque K_n, et comme $\Omega = \bigcup_{n=0}^{\infty} K_n$, la fonction f est nulle sur Ω. □

30.1 Dérivée faible

30.1.1 Dérivée partielle au sens faible

DEFooIRJQooMVNopl

Lemme-Définition 30.2.
Soit $f \in L^p(I)$ où I est l'intervalle ouvert $]a,b[$. Il existe au maximum[1] *une fonction g telle que*

$$\int_I f\varphi' = -\int_I g\varphi \tag{30.7}$$

*pour tout $\varphi \in \mathscr{D}(I)$. Lorsqu'une telle fonction existe, nous la nommons **dérivée faible** de f.*

Démonstration. Soient $g, h \in L^2$ tels que

$$\int_I u\varphi' = -\int_I g\varphi = -\int_I h\varphi \tag{30.8}$$

pour tout $\varphi \in C_c^\infty(I)$. Nous avons alors

$$\int_I (g-h)\varphi = 0. \tag{30.9}$$

Cela implique que $g - h = 0$ presque partout par la proposition 27.189[2]. □

Exemple 30.3 (Dérivée faible de $\mathbb{1}_{\mathbb{Q}}$)**.**
Vu que $\mathbb{Q}$ est de mesure nulle dans $\mathbb{R}$, nous avons

$$\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}}\varphi' = 0 \tag{30.10}$$

pour tout $\varphi \in \mathscr{D}(\mathbb{R})$. Pour $g = 0$ nous avons aussi $\int_{\mathbb{R}} g\varphi = 0$. Donc $g = 0$ est la dérivée faible de $\mathbb{1}_{\mathbb{Q}}$.

Cela n'est pas étonnant du fait qu'en théorie de l'intégration, les parties de mesure nulle ne comptent pas. De ce point de vue, $\mathbb{1}_{\mathbb{Q}} = 0$. D'ailleurs cette égalité est vraie dans L^p (les classes et tout ça). △

EXooRVGHooTWOCtF

Exemple 30.4 (La fonction de Heaveside n'a pas de dérivée faible)**.**
Nous montrons que la fonction

$$H(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geqslant 0 \end{cases} \tag{30.11}$$

n'a pas dérivée faible. Nous nommons g une hypothétique fonction vérifiant les conditions pour être la dérivée faible de H.

Soit une fonction $\varphi \in \mathscr{D}(\mathbb{R})$, dont le support est contenu dans $]0,\infty[$. Sur le support de φ, et donc aussi de φ', nous avons $H(x) = 1$ et donc

$$\int_{\mathbb{R}} \varphi' = -\int_{\mathbb{R}} g\varphi. \tag{30.12}$$

Vu que φ est à support compact, $\int_{\mathbb{R}} \varphi' = 0$. En effet, si le support de φ est contenu dans $[-M, M]$, alors en utilisant le théorème fondamental de l'analyse 14.287, nous trouvons $\int_{\mathbb{R}} \varphi = \int_{-M}^{M} \varphi' = \varphi(M) - \varphi(-M) = 0 - 0 = 0$.

Donc g doit satisfaire

$$\int_{\mathbb{R}} g(x)\varphi(x) = 0 \tag{30.13}$$

pour tout $\varphi \in \mathscr{D}(x > 0)$. La proposition 30.1 nous dit que $g = 0$ presque partout sur $]0,\infty[$.

1. En réalité, c'est une classe au sens de l'égalité presque partout.
2. Ou alors par le lemme 27.76 qui est moins général mais tout aussi bien pour ici.

Le même raisonnement dit que $g = 0$ presque partout sur les négatifs. Que g soit maintenant nulle ou non en $x = 0$ ne change pas le fait que $g = 0$ presque partout sur $\mathbb{R}$.

Par conséquent, $\int_{\mathbb{R}} g\varphi = 0$ pour toute $\varphi \in \mathscr{D}(\mathbb{R})$. Hélas, nous avons d'autre part

$$\int_{\mathbb{R}} H(x)\varphi'(x)dx = \int_0^\infty \varphi'(x)dx, \tag{30.14}$$

qui n'est pas forcément nul. Notons que pour avoir un exemple de φ qui donne $\int_{\mathbb{R}} H\varphi \neq 0$, il faut chercher des fonctions dont le support contient des négatifs et des positifs.

La fonction H n'a donc pas de dérivée faible. Notons cependant que cela ne présume en rien la possibilité d'accepter une dérivée au sens des distributions. △

30.5.
Nous verrons dans la proposition 30.39 que la dérivée de Heaveside au sens des distributions est le delta de Dirac.

Exemple 30.6 (Dérivée faible de la valeur absolue)**.**
L'exemple de base de fonction continue qui n'est pas dérivable est la valeur absolue $f(x) = |x|$ prise en $x = 0$. Nous allons montrer ici que la fonction

$$H(x) = \begin{cases} -1 & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{cases} \tag{30.15}$$

est la dérivée faible de f.

Commençons par noter que H peut valoir la valeur qu'on veut en zéro ; de toutes façons la dérivée faible n'est définie qu'à partie de mesure nulle près.

Soit $\varphi \in \mathscr{D}(\mathbb{R})$ et $M > 0$ tel que le support de φ soit contenu dans $[-M, M]$. Nous avons d'une part

$$\int_{\mathbb{R}} |x|\varphi'(x)dx = -\int_{-M}^0 x\varphi'(x)dx + \int_0^M x\varphi'(x)dx \tag{30.16a}$$

$$= -[\varphi x]_{-M}^0 + \int_{-M}^0 \varphi + [x\varphi]_0^M - \int_0^M \varphi \tag{30.16b}$$

où nous avons utilisé l'intégration par partie de la proposition 20.133 en posant

$$u = x \qquad v' = \varphi' \tag{30.17a}$$

$$u' = 1 \qquad v = \varphi. \tag{30.17b}$$

Tout cela pour dire que

$$\int_{\mathbb{R}} |x|\varphi'(x)dx = \int_{-M}^0 \varphi - \int_0^M \varphi. \tag{30.18}$$

D'autre part, l'égalité

$$\int_{\mathbb{R}} H(x)\varphi(x)dx = -\int_{-M}^0 \varphi + \int_0^M \varphi \tag{30.19}$$

est immédiate.

Nous en déduisons que H est bien la dérivée faible de $x \mapsto |x|$. △

REMooBGJFooPBkFqm

Remarque 30.7.
La dérivée faible ne doit pas être confondue avec la dérivée au sens des distributions qui sera définie en 30.37. Nous avons donc trois notions distinctes de dérivation pour une fonction :

— la dérivée usuelle,

— la dérivée au sens des distributions,

— la dérivée faible.

La dérivée faible d'une fonction reste une fonction, tandis que la dérivée distributionnelle d'une fonction est une distribution.

Je vous mets en garde contre l'idée que l'existence de l'une impliquerait trop facilement l'existence d'une autre [3].

30.1.2 Dérivée faible partielle

La notion de dérivée partielle faible est la même que l'autre. Histoire de nous mettre dans le bain, nous écrivons la définition avec les notation du produit scalaire au lieu de l'intégrale.

DEFooBRFCooPncSCE

Définition 30.8.
*Si $i = 1, \dots, n$, la **dérivée faible** de v dans la direction e_i est l'application* [4] *notée $\partial_i v$ définie par*

$$\langle \partial_i v, \phi \rangle = -\langle v, \partial_i \phi \rangle \tag{30.20}$$

EQooMRZUooFoqPqv

pour tout $\phi \in C_c^\infty(\Omega)$.

Lemme 30.9.
Si $v \in L^2$ admet une dérivée faible, alors cette dernière est unique.

Démonstration. Supposons f, g telles que $\langle g, \phi \rangle$ et $\langle f, \phi \rangle$ soient tous deux égaux à $-\langle v, \partial_i \phi \rangle$. En particulier pour tout $\phi \in \mathscr{D}(\Omega)$ nous avons $\langle (f - g), \phi \rangle = 0$.

Cela donne $f - g = 0$ par la proposition 27.189. □

Soit Ω un ouvert de $\mathbb{R}^d$. Le but de notre histoire est de définir une distribution comme étant un élément de l'espace dual (topologique, voir définition 25.1) de l'espace $\mathscr{D}(\Omega)$ des fonctions C^∞ à support compact dans Ω. Pour ce faire nous devons voir un peu de topologie sur différents espaces de fonctions. Notons que l'espace $\mathscr{D}(\Omega)$ n'est pas réduit à la fonction nulle comme en témoigne l'exemple donné par l'équation (15.570).

PROPooVZFHooKfSpfO

Proposition-Définition 30.10.
Pour chaque K compact dans Ω et chaque entier m, l'application

$$\begin{aligned} p_{K,m} \colon\ & C^\infty(\Omega) \to \mathbb{R} \\ & f \mapsto \sum_{|\mu| \leqslant m} \| \partial^\mu f \|_{K,\infty}. \end{aligned} \tag{30.21}$$

EQooZSQUooAJRIFe

est une seminorme.

Démonstration. Nous devons vérifier les trois conditions de la définition 7.326. Soient $f, g \in C^\infty(\Omega)$. Nous avons d'abord $p_{K,m}(f) \geqslant 0$ parce qu'une norme est toujours positive. Ensuite $p_{K,m}(\lambda f) = |\lambda| p_{K,m}(f)$ parce que $\|\lambda f\| = |\lambda| \|f\|$. Pour la troisième condition,

$$p_{K,m}(f + g) = \sum_{|\mu| \leqslant m} \| \partial^\mu f + \partial^\mu g \|_K \tag{30.22a}$$

$$= \sum_{|\mu| \leqslant m} \sup_{x \in K} \Big(\| (\partial^\mu f)(x) + (\partial^\mu g)(x) \| \Big) \tag{30.22b}$$

$$\leqslant \sum_{|\mu| \leqslant m} \sup_{x \in K} \Big(\| (\partial^\mu f)(x) \| + \| (\partial^\mu g)(x) \| \Big) \tag{30.22c}$$

$$\leqslant \sum_{|\mu| \leqslant m} \sup_{x \in K} \| (\partial^\mu f)(x) \| + \sup_{x \in K} \| (\partial^\mu g)(x) \| \tag{30.22d}$$

$$= p_{K,m}(f) + p_{K,m}(g). \tag{30.22e}$$

□

3. Wikipédia cite l'exemple de la fonction de Cantor qui est dérivable presque partout au sens usuel, mais qui n'est pas faiblement dérivable. Écrivez-moi si vous connaissez des théorèmes qui lient les trois notions de dérivée.

4. En fait c'est une classe au sens de l'égalité presque partout.

En particulier,

$$p_{K,0}(f) = \sup_{x\in K} |f(x)| = \|f\|_{\infty,K}. \tag{30.23}$$

30.1.3 Formules de Leibniz

PROPooTEACooKXUXKl

Proposition 30.11 ([714]).
Soient des fonctions $f, g\colon \mathbb{R} \to \mathbb{R}^n$ chacune n fois dérivable. Alors nous avons la formule

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}. \tag{30.24}$$

Démonstration. Nous faisons par récurrence sur n. Le premier pas étant la règle de Leibniz usuelle de la proposition 12.172(3). Nous avons

$$\begin{aligned}
(fg)^{(n+1)} &= \left(\sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)} \right)' && \text{(30.25a)}\\
&= \sum_{k=0}^{n} \binom{n}{k} f^{(n+1-k)} g^{(k)} + \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k+1)} && \text{(30.25b)}\\
&= \sum_{k=0}^{n} \binom{n}{k} f^{(n+1-k)} g^{(k)} + \sum_{k=1}^{n} \binom{n}{k-1} f^{(n+1-k)} g^{(k)} && \text{(30.25c)}\\
&= \binom{n}{0} f^{(n+1)} g + \binom{n}{n} f g^{(n+1)} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k+1} \right] f^{(n+1-k)} g^{(k)} && \text{cf. justif.}\\
&&& \text{SUBEQooVBEMooPmPWIS} \quad \text{(30.25d)}\\
&= \binom{n+1}{0} f^{(n+1)} g + \binom{n+1}{n+1} f g^{(n+1)} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k+1} \right] f^{(n+1-k)} g^{(k)} && \text{cf. justif.}\\
&&& \text{SUBEQooKDNYooDmqUvB} \quad \text{(30.25e)}\\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(n+1-k)} g^{(k)} && \text{lem. 3.41.}\\
&&& \text{(30.25f)}
\end{aligned}$$

Justifications.

(1) Pour (30.25d). Nous avons séparé le termes $k = 0$ de la première somme et le terme $k = n+1$ de la seconde.

(2) Pour (30.25e). Quelques cas particuliers de la définition (3.73) :

$$\binom{n}{0} = \binom{n+1}{0} = \binom{n}{n} = \binom{n+1}{n+1} = 1. \tag{30.26}$$

□

LEMooOLQTooEHJuBc

Lemme 30.12 (Formule de Leibniz[4, 715]).
Soient des fonctions f, g chacune dérivable p fois sur $\mathbb{R}^d$. Alors pour tout multiindices[5] α de taille plus petite ou égale à p, nous avons

$$\partial^\alpha (fg) = \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} (\partial^\beta f)(\partial^{\alpha-\beta} g). \tag{30.27}$$

5. Voir les définitions 3.47 et -1.1.

Démonstration. Nous faisons une récurrence sur la taille de α. Nous supposons que la formule est valable lorsque $|\alpha| \leqslant k$, et nous prouvons la formule avec $\alpha = (\alpha_1, , \alpha_n)$ avec $|\alpha| = k + 1$.

Soit j tel que $\alpha_j \neq 0$. Nous définissons $\tilde{\alpha}$ par $\tilde{\alpha} = \alpha_j e_j + \tilde{\alpha}$.

Si $\tilde{\alpha} = 0$, alors α contient seulement une composant selon e_j. Dans ce cas, le résultat se réduit au cas à une dimension de la proposition 30.11. Si $\tilde{\alpha} \neq 0$, alors les multiindices $\alpha_j e_j$ et $\tilde{\alpha}$ sont tout deux de longueur au maximum k. Il reste à calculer :

$$\partial^{\alpha}(fg) = \partial^{\alpha_j e_j}\partial^{\tilde{\alpha}}(fg) \tag{30.28a}$$
$$= \partial^{\alpha_j e_j} \sum_{\beta \leqslant \tilde{\alpha}} \binom{\tilde{\alpha}}{\beta} (\partial^{\beta} f)(\partial^{\tilde{\alpha}-\beta} g) \tag{30.28b}$$
$$= \sum_{\beta \leqslant \tilde{\alpha}} \binom{\tilde{\alpha}}{\beta} \sum_{\gamma=0}^{\alpha_j} \binom{\alpha_j}{\gamma} (\partial^{\gamma e_j}\partial^{\beta} f)(\partial^{(\alpha_j-\gamma)e_j}\partial^{\tilde{\alpha}-\beta} g) \qquad \text{prop. 30.11.} \tag{30.28c}$$

Maintenant nous faisons quelque manipulations. Remarquons que dans la somme $\beta_j \leqslant \tilde{\alpha}_j = 0$. Donc nous avons égalité des ensembles

$$\{\beta \text{ tel que } \beta \leqslant \tilde{\alpha}\} = \{\beta \text{ tel que } \beta \leqslant \alpha, \beta_j = 0\}. \tag{30.29}$$

Et comme il n'y a pas de composantes j dans β, la définition 3.90 des coefficients binomiaux donne

$$\binom{\tilde{\alpha}}{\beta}\binom{\alpha_j e_j}{\gamma e_j} = \binom{\tilde{\alpha} + \alpha_j e_j}{\beta + \gamma e_j}. \tag{30.30}$$

Enfin, pour y voir plus clair nous posons

$$u(\delta) = \binom{\alpha}{\delta}(\partial^{\delta} f)(\partial^{\alpha-\delta} g). \tag{30.31}$$

Nous pouvons continuer le calcul :

$$\partial^{\alpha}(fg) = \sum_{\substack{\beta \leqslant \alpha \\ \beta_j = 0}} \sum_{\gamma=0}^{\alpha_j} \binom{\alpha}{\beta + \gamma e_j} (\partial^{\beta+\gamma e_j} f)(\partial^{\alpha-(\beta+\gamma e_j)} g) \tag{30.32a}$$
$$= \sum_{\substack{\beta \leqslant \alpha \\ \beta_j = 0}} \sum_{\gamma=0}^{\alpha_j} u(\beta + \gamma e_j) \tag{30.32b}$$
$$= \sum_{\delta \leqslant \alpha} u(\delta), \tag{30.32c}$$

ce qu'il fallait. □

30.2 Topologie et convergence sur des espaces de fonctions

30.2.1 Limite inductive

DEFooFCLUooSGJIKJ

Définition 30.13 (limite inductive[216])**.**
*Soit un espace vectoriel E ainsi que des sous-espaces vectoriels $(E_i)_{i\in I]}$ tels que $E = \bigcup_{i\in I} E_i$. Nous supposons avoir les seminormes $\{p_j^{(i)}\}_{j\in J_i}$ sur E_i. La topologie **limite inductive** sur E est celle définie par l'ensemble des seminormes p sur E telles que $p|_{E_i}$ soit continue pour tout $i \in I$.*

30.2.2 Les espaces classiques

30.14.
Ici, Ω est un ouvert de $\mathbb{R}^d$, et K est un compact de $\mathbb{R}^d$. Toutes les fonctions sont sur $\mathbb{R}^d$ à

valeurs dans $\mathbb{C}$. Nous rappelons que les seminormes $p_{K,m}$ sont définies en 30.10 et que la topologie engendrée par une seminorme est définie en 7.342.

En tant qu'ensemble, $\mathscr{D}(\Omega)$ est l'ensemble des fonctions de classe C^∞ sur $\mathbb{R}^d$ et dont le support est un compact dans Ω. L'ensemble $\mathscr{D}(K)$ est l'ensemble des fonctions de classe C^∞ sur $\mathbb{R}^d$ et dont le support est un compact dans K.

DefFGGCooTYgmYf

Définition 30.15 (Topologie sur $C^\infty(\Omega)$)**.**
Sur $C^\infty(\Omega)$, la topologie des seminormes $p_{K,m}$ de la définition 30.10 où K parcourt les compacts de Ω et m parcourt $\mathbb{N}$.

30.16.
Pour l'ensemble des fonctions C^∞ à support compact dans A, la notation devrait être $C_c^\infty(A)$, et non $\mathscr{D}(A)$. Hélas je ne suis pas consistant dans les notations.

DEFooJJJYooIrekSp

Définition 30.17 (Toplogie sur $C_c^\infty(K)$)**.**
Soient un ouvert Ω dans $\mathbb{R}^n$, ainsi qu'un compact $K \subset \Omega$. Nous notons

$$C^\infty(K) = \{f \in C^\infty(\Omega) \text{ tel que } \operatorname{supp}(f) \subset K\}. \tag{30.33}$$

La topologie dessus est celle des seminormes $(p_{K,m})_{m\in\mathbb{N}}$

EQooXGINooYJqmtD

$$p_{K,m}(f) = \max_{|\alpha|\leq m} \|\partial^\alpha f\|_K \tag{30.34}$$

30.18.
Quelque remarques.

(1) Il n'est dit que les fonctions dans $C^\infty(K)$ ont un support égal à K.
(2) L'espace $C^\infty(K)$ est égal à $C_c^\infty(K)$ parce que si f a un support compact dans K, elle a un support compact (un fermé dans un compact est compact).

LEMooFKWMooSoqITJ

Lemme 30.19.
Si K est compact dans l'ouvert Ω dans $\mathbb{R}^n$, l'espace $C^\infty(K)$ est métrisable.

Démonstration. C'est la proposition 7.350 : la topologie sur $C_c^\infty(K)$ est donnée par la famille dénombrable de seminormes (30.34). □

DEFooVSCRooLyYBzT

Définition 30.20 (Bonne seminorme, topologie sur $C_c^\infty(\Omega)$)**.**
Soit un ouvert Ω dans $\mathbb{R}^n$. Une seminorme[6] *p sur $C_c^\infty(\Omega)$ est **bonne** si pour tout compact $K \subset \Omega$, la restriction $p|_{C_c^\infty(K)}$ est continue.*

Sur $C_c^\infty(\Omega)$, nous considérons la topologie des bonnes seminormes.

Cela n'est pas très explicite, mais heureusement nous n'aurons souvent pas besoin de plus que de la notion de convergence dans $\mathscr{D}'(\Omega)$. Rappelons que la topologie d'un espace donne la notion de convergence par la définition 7.13.

PROPooRAZSooDttIbK

Proposition 30.21 ([716, 1])**.**
Nous avons

$$\phi_k \xrightarrow{C_c^\infty(\Omega)} \phi \tag{30.35}$$

si et seulement si il existe un compact $K \subset \Omega$ tel que

(1) $\operatorname{supp}(\phi_k) \subset K$ *pour tout k.*
(2) $\operatorname{supp}(\phi) \subset K$.
(3) $\phi_k \xrightarrow{C_c^\infty(K)} \phi$.

6. Seminorme, définition 7.326.

Autrement dit nous avons convergence dans $C_c^\infty(\Omega)$ si et seulement si il existe un compact dans lequel tout se passe y compris la convergence.

Démonstration. En deux parties.

Sens direct, $\Rightarrow$.

Nous supposons que $\phi_k \xrightarrow{C_c^\infty(\Omega)} \phi$. Nous supposons, par l'absurde qu'il n'existe pas de compacts K dans Ω tel que $\text{supp}(\phi_k) \subset K$ et $\text{supp}(\phi) \subset K$.

(1) **Une sous-suite exhaustive**
Soit (K_i) une suite exhaustive de compacts pour Ω (lemme 7.306). La partie $\text{supp}(\phi)$ est un compact dans Ω, et le lemme 7.306(2) dit qu'il existe un n tel que $\text{supp}(\phi) \subset K_n$. Nous considérons une nouvelle suite exhaustive $K'_i = K_{i+n}$. Autrement dit, $\text{supp}(\phi) \subset K'_i$ pour tout i.
Nous oublions la suite (K_i) et nous ne travaillons plus qu'avec (K'_i) que nous allons écrire (K_i) en oubliant le prime.

(2) **Les k_i**
Par hypothèse d'absurdité, nous avons alors qu'il n'existe aucun i contenant tous les $\text{supp}(\phi_k)$: pour tout k, il existe $k_i \in \mathbb{N}$ tel que $\text{supp}(\phi_{k_i})$ n'est pas inclus dans K_i.

(3) **$\{k_i\}$ n'est pas borné** Nous prouvons que l'ensemble des K_i n'est pas borné. Par l'absurde, nous supposons que $k_i < n$ pour tout i. Nous posons alors

$$A = \bigcup_{k<n} \text{supp}(\phi_k). \tag{30.36}$$

La partie A est un compact en tant que union finie de compacts[7]. Elle est donc contenue dans K_s pour un certain $s \in \mathbb{N}$. De plus $\text{supp}(\phi_{k_i}) \subset A$ pour tout i. Nous avons donc, pour tout i :

$$\text{supp}(\phi_{k_i}) \subset A \subset K_s. \tag{30.37}$$

En posant $i = s$, nous trouvons que $\text{supp}(\phi_{k_s}) \subset K_s$, ce qui est faux. Nous en déduisons que $(k_i)_{i\in\mathbb{N}}$ n'est pas bornée[8].

(4) **Sous-suite des ϕ_k**
Vu que (k_i) est non bornée, nous pouvons en extraire une sous-suite $(k'_i)_{i\in\mathbb{N}}$ telle que $k'_i \to \infty$. Nous considérons alors la sous-suite des ϕ_k donnée par

$$\tilde{\phi}_i = \phi_{k'_i}. \tag{30.38}$$

Et nous oublions immédiatement la suite initiale ainsi que le tilde. Donc pour tout i, nous avons

$$\text{supp}(\phi_i) \not\subset K_i. \tag{30.39}$$

Pour tout n, il existe $x_n \notin K_n$ tel que $\phi_n(x_n) \neq 0$.

(5) **Une nouvelle seminorme**
Si $f \in C_c^\infty(\Omega)$, nous avons $\text{supp}(f) \subset K_r$ pour un certain r. Si $n > r$, nous avons $K_r \subset K_n$ et $x_n \notin K_n$, donc $x_n \notin K_r$. Bref, nous avons alors $f(x_n) = 0$ pour tout $n > r$.
Tout cela pour dire que $f(x_n) \neq 0$ pour seulement un nombre fini de n. Il nous est donc loisible de poser

$$\begin{aligned} p\colon C_c^\infty(\Omega) &\to \mathbb{R} \\ f &\mapsto \max_{n\in\mathbb{N}} \Big|\frac{1}{\phi_n(x_n)} f(x_n)\Big|. \end{aligned} \tag{30.40}$$

Il s'agit bien d'un max parce qu'il n'y a qu'un nombre fini de termes non nuls. Vérifions que c'est une seminorme en vérifiant les conditions de la définition 7.326.

7. Lemme 7.98.

8. À mon avis, et à l'avis de [716], il y a moyen d'adapter pour prouver que non seulement (k_i) est non bornée, mais en plus $k_i \to \infty$. Prouver cela permettrait d'éviter la sous-suite ci-dessous. Écrivez-moi si vous savez prouver que $k_i \to \infty$.

(5a) En tant que maximum de choses dans des valeurs absolues, nous avons $p(f) \geqslant 0$.

(5b) Si $\lambda \in \mathbb{R}$,

$$p(\lambda f) = \max_n \left|\frac{1}{\phi_n(x_n)}(\lambda f)(x_n)\right| = |\lambda| \max_n |\dots| = |\lambda| p(f). \tag{30.41}$$

(5c) Nous avons

$$\begin{aligned} p(f+g) &= \max_n \left|\frac{1}{\phi_n(x_n)}\big(f(x_n)+g(x_n)\big)\right| && (30.42a)\\ &\leqslant \max_n \left[\left|\frac{1}{\phi_n(x_n)}f(x_n)\right| + \left|\frac{1}{\phi_n(x_n)}g(x_n)\right|\right] && (30.42b)\\ &\leqslant \max_n |\dots| + \max_n |\dots| && (30.42c)\\ &= p(f) + g(f). && (30.42d) \end{aligned}$$

Donc oui, p est une seminorme.

(6) **p est une bonne seminorme** Soit un compact K. Nous devons prouver que $p|_{C_c^\infty(K)}$ est continue. Soit $N \in \mathbb{N}$ tel que $x_n \notin K$ pour tout $n \geqslant N$. Si $f \in C_c^\infty(K)$, nous avons SUBEQSooTBSDooExlzPQ

$$\begin{aligned} p(f) &= \max_n \left|\frac{1}{\phi_n(x_n)}f(x_n)\right| && (30.43a)\\ &= \max_{n\leqslant N} \left|\frac{1}{\phi_n(x_n)}f(x_n)\right| && (30.43b)\\ &\leqslant \max_{n\leqslant N} \left|\frac{1}{\phi_n(x_n)}\|f\|_K\right| && (30.43c)\\ &= \|f\|_K \max_{n\leqslant N} \left|\frac{1}{\phi_n(x_n)}\right| && (30.43d)\\ &= c_K \|f\|_K. && (30.43e) \end{aligned}$$

où c_K est une constante qui dépend de K.

Supposons que $f_k \xrightarrow{C_c^\infty(K)} f$. Pour prouver la continuité de p, nous devons prouver que $p(f) - p(f_k) \to 0$. Nous avons

$$\begin{aligned} |p(f) - p(f_k)| &\leqslant p(f - f_k) && \text{lem. 7.328} && (30.44a)\\ &\leqslant c\|f_k - f\|_K && \text{eq. (30.43).} && (30.44b) \end{aligned}$$

La définition 30.17 donne la topologie sur $C_c^\infty(K)$ par les seminormes

$$p_{K,m}(f - f_k) = \sup_{|\alpha|\leqslant m} \|\partial^\alpha(f - f_k)\|_K. \tag{30.45}$$

La proposition 7.344 nous dit que $p_{K,m}(f - f_k) \to 0$ pour tout m. En particulier pour $m = 0$ nous avons $p_{K,0}(f - f_k) = \|f - f_k\|_K \to 0$.

Au final nous avons

$$|p(f) - p(f_k)| \leqslant c_K \|f - f_k\|_K \to 0, \tag{30.46}$$

et donc la continuité de p en f.

(7) **La contradiction**

Nous avons prouvé que p était une bonne seminorme. Nous devrions donc avoir $p(\phi_k - \phi) \to 0$. Vérifions ça .... Nous avons

$$\begin{aligned} p(\phi_k - \phi) &= \max_n \left|\frac{1}{\phi_n(x_n)}\big(\phi_k(x_n) - \phi(x_n)\big)\right| && (30.47a)\\ &= \max_n \left|\frac{1}{\phi_n(x_n)}\phi_k(x_n)\right| && \text{cf. justif.} && (30.47b)\\ &\geqslant \left|\frac{1}{\phi_k(x_k)}\phi_k(x_k)\right| && (30.47c)\\ &= 1. && (30.47d) \end{aligned}$$

SUBEQooBPBEooWiXATm

Justification. Pour (30.47b). Nous avons fait plein de choix pour avoir $\phi(x_n) = 0$ pour tout n.

Cela contredit la possibilité d'avoir $p(\phi_k - \phi) \to 0$.

(8) **Précision**

Nous avons trouvé une sous-suite non convergente des ϕ_k originaux. Donc les ϕ_k de départ ne convergent pas.

(9) **Première conclusion**

Nous avons prouvé par l'absurde qu'il existe un compact $K \subset \Omega$ tel que $\text{supp}(\phi) \subset K$ et $\text{supp}(\phi_k) \subset K$ pour tout k.

(10) $\phi_k \xrightarrow{C_c^\infty(K)} \phi$ Nous devons encore prouver que $\phi_k \xrightarrow{C_c^\infty(K)} \phi$. Par la proposition 7.344, nous devons prouver que $q(\phi_k - \phi) \to 0$ pour toutes les seminormes $p_{k,m}$ de la définition 30.17.

Par hypothèse nous avons $\phi_k \xrightarrow{C_c^\infty(\Omega)}$, et donc convergence $p(\phi_k - \phi) \to 0$ pour toute bonne seminorme sur Ω. Oh, mais justement pour tout m, l'application $p_{K,m}$ est continue, et donc est une bonne seminorme. Nous avons donc bien

$$p_{K,m}(\phi_k - \phi) \to 0. \tag{30.48}$$

Sens réciproque, $\Leftarrow$.

Nous supposons qu'il existe un compact $K \subset \Omega$ tel que

(1) $\text{supp}(\phi_k) \subset K$ pour tout k,

(2) $\text{supp}(\phi) \subset K$,

(3) $\phi_k \xrightarrow{C_c^\infty(K)} \phi$,

et nous devons prouver que $\phi_k \xrightarrow{C_c^\infty(\Omega)} \phi$, c'est-à-dire que $q(\phi_k - \phi) \to 0$ pour toute bonne seminorme q sur Ω.

Soit une bonne seminorme q sur Ω. En particulier, sa restriction à $C_c^\infty(K)$ est continue. Par hypothèse, $\phi_k \xrightarrow{C_c^\infty(K)} \phi$; par continuité de q, nous avons $q(\phi_k) \to q(\phi)$. □

THOooDFNAooPnDSNN

Théorème 30.22 ([716]).
Soit T, une application linéaire sur $C_c^\infty(\Omega)$, c'est-à-dire $T \in \mathcal{L}\big(C_c^\infty(\Omega), \mathbb{C}\big)$. Elle est continue si et seulement si pour tout compact $K \subset \Omega$, il existe $c > 0$ et $n \in \mathbb{N}$ tels que

$$|T(\phi)| \leqslant c \max_{|\alpha| \leqslant n} \|\partial^\alpha \phi\|_K. \tag{30.49}$$

pour tout $\phi \in C_c^\infty(K)$.

Démonstration. En deux parties

Sens direct, $\Rightarrow$.

Nous supposons que T est continue sur $C_c^\infty(\Omega)$. Nous utilisons le corolaire 7.354 qui dit qu'il existe une seminorme p (parmi celles qui définissent la topologie de $C_c^\infty(\Omega)$, c'est-à-dire une bonne seminorme) et $M > 0$ tels que

$$|T(\phi)| < Mp(\phi) \tag{30.50}$$

pour tout $\phi \in C_c^\infty(\Omega)$. Soit un compact K. Étant donné que p est bonne, elle est continue sur $C_c^\infty(K)$; nous pouvons donc appliquer le théorème 7.352 à la seminorme continue $q = p|_{C_c^\infty(K)}$. Il existe une seminorme de $C_c^\infty(K)$ (c'est-à-dire un $p_{K,m}$) et $c > 0$ tels que $q(\phi) \leqslant cp_{K,m}(\phi)$.

Pour tout $\phi \in C_c^\infty(K)$ nous avons donc

$$\begin{aligned}
|T(\phi)| &< Mp(\phi) && \text{(30.51a)}\\
&= Mq(\phi) && \text{(30.51b)}\\
&\leqslant Mcp_{K,m}(\phi) && \text{(30.51c)}\\
&\leqslant Mc \max_{|\alpha| \leqslant m} \|\partial^\alpha \phi\|_K. && \text{(30.51d)}
\end{aligned}$$

La constante recherchée est donc Mc.

Sens réciproque, $\Leftarrow$.

Soit une application linéaire T sur $C_c^\infty(\Omega)$. Nous supposons que pour tout compact K, il existe $c_K > 0$ et n_K tels que $|T(\phi)| \leqslant c_K \max_{|\alpha|<n_K} \|\partial^\alpha \phi\|_K$.

Soit une suite exhaustive de compacts (K_i) dans Ω. Nous avons des c_i et n_i tels que

$$|T(\phi)| \leqslant c_i \max_{|\alpha|\leqslant n_i} \|\partial^\alpha \phi\|_{K_i} \tag{30.52}$$

pour tout i.

Pour chaque i nous considérons $\phi_i \in C_c^\infty(K_i)$ tel que $\phi_i = 1$ sur K_{i-1}. Cela existe par le lemme d'Urysohn [9] parce que K_{i-1} est inclus dans l'ouvert $\text{Int}(K_i)$.

(1) **Les fonctions ϕ_i et ψ_i** Le lemme d'Urysohn [10], appliqué à K_{i-1} inclus dans l'ouvert $\text{Int}(K_i)$ nous permet de considérer des fonctions ϕ_i vérifiant :

(1a) $\text{supp}(\phi_i) \subset \text{Int}(K_i)$

(1b) $0 \leqslant \phi_i \leqslant 1$

(1c) $\phi_i = 1$ sur K_{i-1}.

Nous posons aussi $\phi_0 = 0$ et $\psi_i = \phi_i - \phi_{i-1}$.

(2) $\sum_n \psi_n(x) = 1$ Soit $x \in \Omega$. Soit $i_0 = \min\{i \in \mathbb{N} \text{ tel que } x \in K_i\}$. Nous n'avons pas de garanties sur la valeur de $\phi_{i_0}(x)$. Par contre, si $i < i_0$, alors $\text{supp}(\phi_i) \subset K_i$ et par minimalité de i_0, nous avons $\phi_i(x) = 0$.

Si par contre $i > i_0$, alors $x \in K_{i-1}$, mais nous savons que $\phi_i = 1)$ sur K_{i-1}. Donc pour $i > i_0$ nous avons $\phi_i(x) = 1$.

Nous avons donc :

(2a) $\psi_{i_0-k}(x) = \phi_{i_0-k}(x) - \phi_{i_0-k-1}(x) = 0 - 0 = 0$.

(2b) $\psi_{i_0}(x) = \phi_{i_0}(x) - \phi_{i_0-1}(x) = \phi_{i_0}(x) - 0 = \phi_{i_0}(x)$

(2c) $\psi_{i_0+1}(x) = \phi_{i_0+1}(x) - \phi_{i_0}(x) = 1 - \phi_{i_0}(x)$

(2d) $\psi_{i_0+k}(x) = 1 - 1 = 0$ si $k > 1$.

Tout ça pour dire que

$$\sum_n \psi_n(x) = \psi_{i_0}(x) + \psi_{i_0+1}(x) = 1. \tag{30.53}$$

Il y a seulement deux termes non nuls dans la somme.

(3) $\text{supp}(\psi_i) \subset K_i \backslash K_{i-2}$ Dans un premier temps nous prouvons que $\text{supp}(\psi_i) \subset K_i$. Si x est hors de K_i, alors il est également hors de K_{i-1}. Du coup$\phi_i(x) = 0$ et $\phi_{i-1}(x) = 0$. Nous avons alors $\psi_i(x) = 0$.

Si par contre $x \in K_{i-2}$, alors $x \in K_{i-2} \subset K_{i-1}$. Vu que $\phi_i = 1$ sur K_{i-1} et $\phi_{i-1} = 1$ sur K_{i-2}, nous avons $\psi_i(x) = 0$.

9. Proposition 27.74.
10. Proposition 27.74.

(4) **Quelques majorations** Un peu de calcul. Soit $\phi \in C_c^\infty(\Omega)$. Nous avons

$$|T(\phi)| = |\sum_n T(\phi\psi_n)| \qquad T \text{ est lin. et (30.53)} \tag{30.54a}$$

$$\leqslant \sum_n |T(\phi\psi_n)| \tag{30.54b}$$

$$\leqslant \sum_n c_n \max_{|\alpha|\leqslant k_n} \|\partial^\alpha(\phi\psi_n)\| \qquad \text{hyp. (30.52)} \tag{30.54c}$$

$$\leqslant \sum_n c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|(\partial^\beta\phi)(\partial^{\alpha-\beta}(\psi_n))\|_{K_n} \qquad \text{lem. 30.12} \tag{30.54d}$$

$$= \sum_n c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|(\partial^\beta\phi)(\partial^{\alpha-\beta}(\psi_n))\|_{K_n\setminus K_{n-2}} \qquad \text{point (3)} \tag{30.54e}$$

$$\leqslant \sum_n c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|\partial^\beta\phi\|_{K_n\setminus K_{n-2}} \|\partial^{\alpha-\beta}\psi\|_{K_n\setminus K_{n-2}}. \tag{30.54f}$$

(5) **Une bonne seminorme** Nous considérons l'application

$$\begin{aligned} p\colon C_c^\infty(\Omega) &\to \mathbb{R} \\ \phi &\mapsto \sum_n c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|\partial^{\alpha-\beta}\psi_n\|_{K_n} \|\partial^\beta\phi\|_{K_n\setminus K_{n-2}}. \end{aligned} \tag{30.55}$$

Pour vérifier que ce p est une seminorme, il suffit de vérifier les conditions de la définition 7.326, et c'est facile. Pour prouver que c'est une bonne [11] seminorme, c'est un peu plus chaud, et il nous faudra utiliser le théorème 7.352(2).

Soit un compact K dans Ω, et considérons la restriction $p\colon C_c^\infty(K) \to \mathbb{R}$. Il existe un $n_0 \in \mathbb{N}$ tel que $K \subset K_{n_0}$. Pour $\phi \in C_c^\infty(K)$, nous avons alors $\text{supp}(\phi) \subset K \subset K_{n_0}$, et donc

$$\|\partial^\beta\phi\|_{K_{n_0+2}\setminus K_{n_0}} = 0. \tag{30.56}$$

Nous avons donc

$$p(\phi) = \sum_{n\leqslant n_0+2} c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|\partial^{\alpha-\beta}\psi_n\|_{K_n} \|\partial^\beta\phi\|_{K_n\setminus K_{n-2}} \tag{30.57a}$$

$$\leqslant \max_{|\beta|\leqslant k_{n_0+2}} \|\partial^\beta\phi\|_K \sum_{n\leqslant n_0+2} c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|\partial^{\alpha-\beta}\psi_n\| \tag{30.57b}$$

$$= c' \max_{|\beta|\leqslant k_{n_0+2}} \|\partial^\beta\phi\|_K \tag{30.57c}$$

$$= c' p_{K,k_{n_0+2}}(\phi) \qquad \text{eq. (30.34)} \tag{30.57d}$$

où

$$c' = \sum_{n\leqslant n_0+2} c_n \max_{|\alpha|\leqslant k_n} \sum_{\beta\leqslant\alpha} \binom{\alpha}{\beta} \|\partial^{\alpha-\beta}\psi_n\|. \tag{30.58}$$

Par le théorème 7.352(2), nous déduisons que p est continue.

(6) **Conclusion** Nous avons montré que $|T(\phi)| \leqslant p(\phi)$ pour une bonne seminorme sur Ω. Le corolaire 7.354 conclut que T est continue.

□

11. Bonne seminorme, définition 30.20

30.2.3 Espaces de fonctions et topologie inductive

PROBooQXUDooJwnzpn

¡¡ Avertissement/question au lecteur !! 30.23
Je ne suis pas tout à fait sûr que la proposition suivante 30.24 soit vraie. Il y a peut-être une preuve dans [216].

PROPooUFJJooCoIdoI

Proposition 30.24.
La topologie sur $\mathscr{D}(\Omega)$ est la limite inductive[12] *des $\mathscr{D}(K_i)$ où (K_i) est une suite exhaustive de compacts dans Ω.*

30.2.4 Convergence dans les fonctions test

LemXXwDjui

Lemme 30.25 (Convergence dans $\mathscr{D}(K)$)**.**
Si α est un multiindice et si $\varphi_n \xrightarrow{\mathscr{D}(K)} \varphi$, alors nous avons

$$\partial^\alpha \varphi_n \xrightarrow{unif} \partial^\alpha \varphi. \tag{30.59}$$

Démonstration. Quitte à considérer la suite $\varphi_n - \varphi$ nous pouvons supposer $\varphi_n \xrightarrow{\mathscr{D}(K)} 0$. Nous avons

$$\|\partial^\alpha \varphi_n\| \leqslant \sum_{\mu \leqslant \alpha} \|\partial^\mu \varphi_n\|_{K,\infty}. \tag{30.60}$$

Vu que le membre de droite tend vers zéro, nous avons

$$\lim_{n\to\infty} \|\partial^\alpha \varphi_n\|_{K,\infty} \to 0, \tag{30.61}$$

ce qui revient à dire que $\partial^\alpha \varphi_n$ converge uniformément sur K vers $\partial^\alpha \varphi$. □

LemWEGpemo

Lemme 30.26.
Si une fonction $f\colon \mathscr{D}(\Omega) \to \mathbb{R}$ est continue sur chacun des $\mathscr{D}(K)$ pour tout K compact dans Ω alors est continue sur $\mathscr{D}(\Omega)$.

Démonstration. Soit I ouvert dans $\mathbb{R}$; nous devons trouver un ouvert $\mathcal{O}$ dans $C^\infty(\Omega)$ tel que $f^{-1}(I) = \mathscr{D}(\Omega) \cap \mathcal{O}$. Vu que f est continue sur chacun des $\mathscr{D}(K)$ avec K compact dans Ω, pour tout tel compact nous avons un ouvert $\mathcal{O}_K$ dans $\mathscr{D}(K)$ tel que $f^{-1}(I) \cap \mathscr{D}(K) = \mathcal{O}_K$. En tant qu'union d'ouverts[13], l'ensemble

$$\mathcal{O} = \bigcup_{K \text{ compact de } \Omega} \mathcal{O}_K \tag{30.62}$$

est ouvert dans $C^\infty(\Omega)$. Si $\phi \in f^{-1}(I)$, nous avons $\phi \in \mathscr{D}(K)$ pour un certain K compact de Ω, donc $f^{-1}(I) \subset \mathcal{O}$. À fortiori nous avons $f^{-1}(I) \subset \mathcal{O} \cap \mathscr{D}(\Omega)$.

Dans l'autre sens, si $\phi \in \mathcal{O}$, alors ϕ est dans un des $\mathcal{O}_K$ et donc dans $f^{-1}(I)$. Nous avons donc bien $f^{-1}(I) = \mathscr{D}(\Omega) \cap \mathcal{O}$. □

ThoXYADBZr

Théorème 30.27 (Convergence dans $\mathscr{D}(\Omega)$[216])**.**
Soit $(\varphi_n)_{n\in\mathbb{N}}$ une suite dans $\mathscr{D}(\Omega)$ et $\varphi \in \mathscr{D}(\Omega)$. Nous avons $\varphi_n \xrightarrow{\mathscr{D}(\Omega)} \varphi$ si et seulement si il existe K compact dans Ω tel que $\varphi_n \in \mathscr{D}(K)$ pour tout n et $\varphi_n \xrightarrow{\mathscr{D}(K)} \varphi$.

Démonstration. Supposons que $\varphi_n \xrightarrow{\mathscr{D}(\Omega)} \varphi$ et qu'il n'existe pas de compacts contenant tous les supports des φ_n. Alors pour tout compact de Ω il existe un n tel que le support de φ_n ne soit pas dans K. Nous considérons une suite de compacts (K_i) tels que $\text{Int}(K_n) \subset K_{n+1}$ et $\Omega = \bigcup_n K_n$. Une telle suite existe par le lemme 7.306. Ensuite nous construisons des sous-suites de la façon suivante. D'abord $L_1 = K_1$ et $n_1 \in \mathbb{N}$ est choisi de telle sorte que φ_{n_1} ait un support non contenu dans L_1. Ensuite L_i est un compact de la suite (K_n) choisi plus loin que L_{i-1} et tel que $\varphi_{n_{i-1}} \in \mathscr{D}(L_i)$. Le

12. Définition 30.13.
13. Voir définition 7.1.

nombre n_i est alors choisi plus grand que n_{i-1} de telle sorte que $\varphi_{n_i} \notin \mathscr{D}(L_i)$. Ce faisant, en posant $\phi_i = \varphi_{n_i}$ nous avons

$$\phi_i \in \mathscr{D}(L_{i+1}) \backslash \mathscr{D}(L_i) \tag{30.63}$$

et $\text{Int}(L_n) \subset L_{n+1}$ et $\Omega = \bigcup_n L_n$. Étant donné que (ϕ_i) et une sous-suite de (φ_i) nous avons encore $\phi_i \xrightarrow{\mathscr{D}(\Omega)} \varphi$.

Soit $i \in \mathbb{N}$. Nous allons utiliser le corolaire 27.170 de la seconde forme géométrique du théorème de Hahn-Banach pour séparer les parties $\{\phi_i\}$ (compact) et $\mathscr{D}(L_i)$ (sous-espace vectoriel fermé de $\mathscr{D}(\Omega)$) dans $\mathscr{D}(\Omega)$. Nous avons $f_i \in \mathscr{D}'(\Omega)$ telle que

$$\begin{cases} f_i(\phi_i) > \alpha & \text{(30.64a)} \\ f_i\big(\mathscr{D}(L_i)\big) = 0. & \text{(30.64b)} \end{cases}$$

Nous introduisons la fonction définie sur $\mathscr{D}(\Omega)$ par

$$p(\phi) = \sum_{i=1}^{\infty} i \frac{f_i(\phi)}{|f_i(\phi_i)|}. \tag{30.65}$$

EqJCqeXti

Si $\phi \in \mathscr{D}(L_k)$, alors $f_k(\phi) = 0$ et même $f_l(\phi) = 0$ pour tout $l \geqslant k$. Donc pour chaque k, la somme définissant p est finie sur $\mathscr{D}(L_k)$. Nous en déduisons que p est continue sur chacun des $\mathscr{D}(L_k)$ et donc sur $\mathscr{D}(\Omega)$ par le lemme 30.26.

L'image de la suite convergente $\phi_k \xrightarrow{\mathscr{D}(\Omega)} \varphi$ par p doit être bornée parce que p est continue. Mais dans la somme (30.65), tous les termes sont positifs et en particulier le terme $i = k$ vaut k, donc $p(\phi_k) \geqslant k$, ce qui contredit le fait que l'image de la suite soit bornée. Nous en déduisons donc l'existence d'un compact K tel que $\varphi_n \in \mathscr{D}(K)$ pour tout n.

Nous devons encore prouver que $\varphi_n \xrightarrow{\mathscr{D}(K)} \varphi$ pour ce choix de K. Vu que $\varphi_n \xrightarrow{\mathscr{D}(\Omega)} \varphi$, le lemme 7.58 nous dit que nous avons aussi $\varphi_n \xrightarrow{C^\infty(\Omega)} \varphi$, ce qui signifie que pour tout K et m nous avons

$$p_{K,m}(\varphi_n - \varphi) \to 0. \tag{30.66}$$

En particulier pour le K fixé plus haut nous avons $p_m(\varphi_n - \varphi) \to 0$, c'est-à-dire que $\varphi_n \xrightarrow{\mathscr{D}(K)} \varphi$. □

PropQAEVcTi

Proposition 30.28.
Soit un compact K de Ω.

(1) L'espace $\mathscr{D}(K)$ est complet.

(2) L'espace $\mathscr{D}(K)$ est métrique.

Démonstration. Nous allons d'abord montrer que $\mathscr{D}(K)$ est complet. Ensuite nous allons montrer que sa topologie peut être donnée par une distance.

(1) **Complet** Nous considérons une suite de Cauchy (φ_n) dans $\mathscr{D}(K)$ au sens de la définition 7.260. Soient $\epsilon > 0$ et $i \in \mathbb{N}$; si k et l sont assez grands nous avons

$$\varphi_k - \varphi_l \in B_i(0, \epsilon). \tag{30.67}$$

En particulier pour $i = 0$ nous avons l'inégalité

$$\|\varphi_k - \varphi_l\|_\infty \leqslant \epsilon, \tag{30.68}$$

La suite (φ_n) est donc de Cauchy dans $\big(C(K), \|.\|_\infty\big)$ et y converge donc par complétude, proposition 12.359. Il existe donc une fonction $\varphi \in C(K)$ telle que

$$\varphi_n \xrightarrow{unif} \varphi. \tag{30.69}$$

Notre jeu à présent est de prouver que $\varphi \in \mathscr{D}(K)$, c'est-à-dire qu'elle est de classe C^∞.

Soit un multiindice $\alpha = \mu_1, \dots, \mu_n, i$. Si k et l sont assez grands nous avons

$$\|\partial^\alpha(\varphi_k - \varphi_l)\|_\infty \leqslant \epsilon, \tag{30.70}$$

c'est-à-dire que

$$\|\partial_i(\partial^\mu\varphi_k) - \partial_i(\partial^\mu\varphi_l)\|_\infty \leqslant \epsilon. \tag{30.71}$$

Si nous notons $\psi_k = \partial^\mu\varphi_k$ cela signifie que $(\partial_i\psi_n)$ est une suite de Cauchy dans $\big(C(K), \|.\|_\infty\big)$. Elle y converge donc et il existe une fonction $g_i \in C(K)$ telle que

$$\partial_i\psi_n \xrightarrow{unif} g_i. \tag{30.72}$$

Dans ce cas le théorème 12.377 nous indique que g_i est de classe C^n, c'est-à-dire que $\varphi \in C^{n+1}(K)$.

(2) **Métrique**

La proposition 7.351 nous dit que la topologie donnée par l'écart

$$d(\varphi_1, \varphi_2) = \sup_{k\geqslant 1} \min\{\frac{1}{k}, p_{k-1}(\varphi_1 - \varphi_2)\} \tag{30.73}$$

est la même que celle de $\mathscr{D}(K)$. Il reste à montrer que cette formule est bien une distance au sens de la définition 7.115.

(2a) Nous avons bien $d(\varphi_1, \varphi_2) \geqslant 0$ parce que tous les éléments du supremum et du minimum sont positifs.

(2b) Si $d(\varphi_1, \varphi_2) = 0$ alors pour tout k nous devons avoir $p_{k-1}(\varphi_1 - \varphi_2) = 0$; en particulier pour $k = 1$ cela donne $\varphi_1 = \varphi_2$.

(2c) Nous avons

$$p_k(\varphi_1 - \varphi_2) = p_k\big(-(\varphi_2 - \varphi_1)\big) = p_k(\varphi_2 - \varphi_1) \tag{30.74}$$

en utilisant la propriété (1) de la définition 7.326 de seminorme.

(2d) Nous avons

$$p_k(\varphi_1 - \varphi_2) = p_k(\varphi_1 - \varphi_3 + \varphi_3 - \varphi_2) \leqslant p_k(\varphi_1 - \varphi_3) + p_k(\varphi_3 - \varphi_2) \tag{30.75}$$

en utilisant la propriété (2) de la définition 7.326.

□

Notons que la proposition 7.351 nous dit que $\mathscr{D}(K)$ est complet tout autant pour la topologie des seminormes que pour celle de la distance que nous venons de décrire. Ces deux topologies sont les mêmes. Étant métrique et complet, l'espace $\mathscr{D}(\Omega)$ et donc de Baire par le théorème 7.366. Ce qui est bien avec ces deux topologies identiques c'est qu'on peut utiliser la propriété de Baire même en ne parlant que des seminormes.

30.3 Distributions

Si Ω est un ouvert de $\mathbb{R}^d$, alors l'ensemble $\mathscr{D}(\Omega)$ est contenu dans $C^\infty(\Omega)$. Nous allons commencer par définir une topologie sur $C^\infty(\Omega)$ et ensuite donner à $\mathscr{D}(\Omega)$ la topologie induite[14].

DefPZDtWVP

Définition 30.29 (Distribution)**.**
Une ***distribution*** *sur un ouvert Ω de $\mathbb{R}^d$ est une forme linéaire continue sur $\mathscr{D}(\Omega) = C_c^\infty(\Omega)$. C'est donc un élément de $\mathscr{D}'(\Omega)$.*

Le théorème suivant donne quelques façons de vérifier qu'une forme linéaire soit continue. En particulier il nous dit que pour prouver qu'une forme linéaire est une distribution il suffit de prouver la continuité séquentielle.

14. Définition 7.26.

ThoVDDBnVn

Théorème 30.30 ([216, 717]).
Soit T une forme linéaire sur $C_c^\infty(\Omega)$. Nous avons équivalence entre les points suivants.

ITEMooQXXDooDleaQD

(1) T est continue.

ITMooMDODooPULDzj

(2) Pour tout compact $K \subset \Omega$, il existe $c_K > 0$ et $n_K \in \mathbb{N}$ tels que

$$|T(\varphi)| \leqslant c_K \max_{|\alpha| \leqslant n_K} \|\partial^\alpha \varphi\|_K. \tag{30.76}$$

pour tout $\varphi \in C_c^\infty(K)$.

ItemSPvoijoii

(3) Pour tout compact $K \subset \Omega$ il existe $m_K \in \mathbb{N}$ et $C_K \geqslant 0$ tels que pour tout $\varphi \in C_c^\infty(K)$ nous ayons

$$\big|T(\varphi)\big| \leqslant C_K p_{K,m_K}(\varphi) = C_K \sum_{|\mu| \leqslant m_K} \|\partial^\mu \varphi\|_K \tag{30.77}$$

voir la définition de la la seminorme donnée en (30.21).

ITEMooBXFSooYtAXjy

(4) T est séquentiellement continue sur $C_c^\infty(\Omega)$.

ITEMooXZJPooYyQQod

(5) T est séquentiellement continue en 0.

ITEMooPTDMooHBrnXk

(6) Pour tout compact $K \subset \Omega$, la restriction de T à $C_c^\infty(K)$ est continue.

Un certain nombre d'ouvrages prennent le point (3) comme la définition d'une distribution.

Démonstration. En plusieurs parties.

(1) **(1) si et seulement si (2)** C'est le théorème 30.22.

(2) **(2) ⇒ (3)** Notons s le nombre de termes dans la somme $|\mu| < n_K$. Nous avons

$$\begin{aligned}
|T(\varphi)| &\leqslant c_K \max_{|\alpha| \leqslant n_K} \|\partial^\alpha \varphi\|_K && (30.78a)\\
&< c_K s \max_{|\alpha| \leqslant n_K} \|\partial^\alpha \varphi\|_K && (30.78b)\\
&= c_K \sum_{|\mu| \leqslant n_K} \Big(\max_{|\alpha| \leqslant n_K} \|\partial^\alpha \varphi\|_K \Big) && (30.78c)\\
&\leqslant c_K \sum_{|\mu| \leqslant n_K} \|\partial^\mu \varphi\|_K. && (30.78d)
\end{aligned}$$

(3) **(3) ⇒ (2)** Même principe : chaque terme de la somme est majoré par le max. On adapte la constante en multipliant par le nombre de termes.

(4) **(3) ⇒ (4)** Nous considérons une suite convergente $\varphi_k \xrightarrow{C_c^\infty(\Omega)} \varphi$. En vertu de la proposition 30.21, nous avons un compact K dans Ω tel que $\varphi_k \xrightarrow{C_c^\infty(K)} \varphi$. Pour un tel compact nous avons

$$\begin{aligned}
|T(\varphi_k) - T(\varphi)| &= |T(\varphi_k - \varphi)| && (30.79a)\\
&\leqslant c_k \max_{|\alpha| \leqslant n_k} \|\partial^\alpha(\varphi_k - \varphi)\|_K \qquad \text{pcq. (3)⇒(2).} && (30.79b)\\
&= c_K p_{n_K}(\varphi_k - \varphi) && (30.79c)
\end{aligned}$$

où p_{n_K} est une des seminormes (30.34) définissant la topologie sur $C_c^\infty(K)$. Par la propositon 7.344, nous avons alors $p_{n_K}(\varphi_k - \varphi) \to 0$.

(5) **(4)⇒ (5)** À fortiori.

(6) **(5)⇒ (6)** Soit un compact K dans Ω. Nous faisons plusieurs étapes.

(6a) $T|_{C_c^\infty(K)}$ **est séquentiellement continue** Soient φ_k et φ dans $C_c^\infty(K)$ tels que $\varphi_k \xrightarrow{C_c^\infty(K)} \varphi$. Par la proposition 30.21, nous avons $\varphi_k \xrightarrow{C_c^\infty(\Omega)} \varphi$. Par hypothèse sur la séquentielle continuité de T [15] nous avons alors

$$T(\varphi_k) \to T(\varphi), \tag{30.80}$$

ce qui prouve que la restriction de T à $C_c^\infty(K)$ est séquentiellement continue.

(6b) **Topologie sur $C_c^\infty(K)$** Elle est métrisable par le lemme 30.19.

(6c) **Conclusion** Vu que $C_c^\infty(K)$ est métrisable, la proposition 7.255 dit que continuité séquentielle y est équivalente à la continuité. L'application T étant séquentiellement continue, elle est continue.

(7) **(6)⇒ (1)** C'est le lemme 30.26.

□

DefASmjVaT

Définition 30.31 (Topologie sur $\mathscr{D}'(\Omega)$)**.**
*Nous munissons l'espace $\mathscr{D}'(\Omega)$ de la **topologie $*$-faible**, c'est-à-dire celle de la famille de seminormes*

$$\begin{aligned} p_\varphi\colon \mathscr{D}'(\Omega) &\to \mathbb{R} \\ T &\mapsto \big|T(\varphi)\big| \end{aligned} \tag{30.81}$$

avec $\varphi \in \mathscr{D}(\Omega)$.

Oui, c'est bien une famille de seminormes indicée par l'ensemble $\mathscr{D}(\Omega)$. Il n'y en a donc à priori pas du tout une quantité dénombrable.

PropEUIsNhD

Proposition 30.32 (Convergence au sens des distributions)**.**
Nous avons $T_n \xrightarrow{\mathscr{D}'(\Omega)} T$ si et seulement si $T_n(\varphi) \to T(\varphi)$ pour tout $\varphi \in \mathscr{D}(\Omega)$.

Démonstration. La convergence $T_n \xrightarrow{\mathscr{D}'(\Omega)} T$ signifie que l'on ait $p_\varphi(T_n - T) \to 0$ pour tout $\varphi \in \mathscr{D}(\Omega)$, ce qui en retour signifie que

$$\big|(T_n - T)(\varphi)\big| \to 0. \tag{30.82}$$

□

Cette proposition suppose que l'on ait une distribution T qui vérifie $T_n(\varphi) \to T(\varphi)$ et conclut qu'on a une convergence dans les distributions. Le théorème suivant est plus fort : il va seulement supposer que $T_n(\varphi)$ converge dans $\mathbb{C}$ et va conclure que $T\colon \varphi \mapsto \lim_{n\to\infty} T_n(\varphi)$ est une distribution.

Définition 30.33.
Si f est une fonction sur $\mathbb{R}^d$ telle que $f\varphi \in L^1(\mathbb{R}^d)$ pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$, alors nous définissons la distribution $T_f \in \mathscr{D}'(\mathbb{R}^d)$ par

$$\langle T_f, \varphi \rangle = \int_{\mathbb{R}^d} f(x)\varphi(x)dx. \tag{30.83}$$

Théorème 30.34 ([718])**.**
Soit (T_n) une suite dans $\mathscr{D}'(\Omega)$ et nous supposons que pour tout $\varphi \in \mathscr{D}(\Omega)$ la suite $\big(T_n(\varphi)\big)$ converge dans $\mathbb{C}$. Alors il existe $T \in \mathscr{D}'(\Omega)$ telle que $T_n \xrightarrow{\mathscr{D}'(\Omega)} T$.

PROPooYAJSooMSwVOm

Proposition 30.35.
L'application

$$\begin{aligned} i\colon L^2(\Omega) &\to \mathscr{D}'(\Omega) \\ f &\mapsto T_f \end{aligned} \tag{30.84}$$

est une injection continue.

15. L'hypothèse est juste en 0, mais comme T est linéaire, elle est séquentiellement continue partout.

Démonstration. Le fait que ce soit une injection est le fait que si $T_f = T_g$ alors pour tout $\phi \in \mathscr{D}(\Omega)$ nous avons $\langle f-g, \phi\rangle = 0$, et cela implique que $f-g$ est nulle presque partout en tant que fonction et est simplement nulle en tant que classe de fonction dans L^2.

En ce qui concerne la continuité, il suffit de la prouver en zéro (par linéarité). Soit donc $f_n \xrightarrow{L^2} 0$ et montrons que $T_{f_n} \xrightarrow{\mathscr{D}'(\Omega)} T_0$. Pour prouver cela, la proposition 30.32 nous indique qu'il est suffisant de tester $T_n(\phi) \to 0$ pour tout $\phi \in \mathscr{D}(\Omega)$.

Notons que si $\phi \in \mathscr{D}$ a fortiori $\phi \in L^2$. Nous avons

$$T_{f_n}(\phi) = \int_\Omega f_n\phi \leqslant \|f_n\phi\|_{L^1} \leqslant \|f_n\|_2\|\phi\|_2 \to 0 \tag{30.85}$$

où nous avons utilisé l'inégalité de Hölder de la proposition 27.33. □

Cette proposition permet de donner un sens à des phrases du type « Soit une distribution T. Si $T \in L^2$, alors … ». Cela signifie qu'il existe $u \in L^2$ tel que $T = T_u$. Notons que dans ce cas, la distribution est définie sur L^2 et non seulement sur $\mathscr{D}$.

30.3.1 Multiplication d'une distribution par une fonction

DefZVRNooDXAoTU

Définition 30.36.
Si $T \in \mathscr{D}'(\Omega)$ et si $f \in C^\infty(\Omega)$ nous définissons la distribution fT par

$$(fT)(\varphi) = T(f\varphi). \tag{30.86}$$

DefTDkrgkA

Souvent écrit sous la forme plus compacte $\langle fu, \phi\rangle = \langle u, f\phi\rangle$.

Cela a un sens parce que si $\varphi \in \mathscr{D}(\Omega)$ alors $f\varphi$ est aussi dans $\mathscr{D}(\Omega)$.

Cette définition est motivée par ce que l'on ferait pour une distribution à densité. Si T est une distribution de densité notée également T, nous avons $T(\phi) = \int T(x)\phi(x)$ et donc

$$(fT)(\phi) = \int (fT)(x)\phi(x) = \int T(x)f(x)\phi(x) = \int T(x)(f\phi)(x) = T(f\phi). \tag{30.87}$$

En ce qui concerne les distributions tempérées, nous pouvons définir le produit avec une fonction $f \in \mathscr{S}(\Omega)$ par la même formule : si $f, \varphi \in \mathscr{S}(\Omega)$ alors le produit $f\varphi$ est encore Schwartz. Notons toutefois que nous ne pouvons pas définir fT dans $\mathscr{S}'(\Omega)$ si f est seulement dans $C^\infty(\Omega)$.

30.3.2 Dérivée de distribution

PropKJLrfSX

Proposition-Définition 30.37.
Soit T une distribution sur Ω et $\alpha \in \mathbb{N}^d$. Alors la formule

$$(\partial^\alpha T)(\varphi) = (-1)^{|\alpha|}T(\partial^\alpha\varphi) \tag{30.88}$$

définit une distribution $\partial^\alpha T$.

*Cette distribution $\partial^\alpha T$ sera la **dérivée distributionnelle** de T. Notons que le même résultat est encore valide pour des distributions tempérées, et la démonstration est la même.*

Démonstration. La forme linéaire $\partial^\alpha T$ sera continue si elle est séquentiellement continue par le théorème 30.30. Nous considérons donc une suite $\varphi_n \xrightarrow{\mathscr{D}(\Omega)} \varphi$ et nous vérifions que

$$\lim_{n\to\infty}(\partial^\alpha T)(\varphi_n) = (\partial^\alpha T)(\varphi). \tag{30.89}$$

D'abord T étant une distribution (et donc continue) nous pouvons la permuter avec la limite :

$$\lim_{n\to\infty}(\partial^\alpha T)(\varphi_n) = \lim_{n\to\infty}(-1)^{|\alpha|}T(\partial^\alpha\varphi_n) = (-1)^{|\alpha|}T\big(\lim_{n\to\infty}\partial^\alpha\varphi_n\big). \tag{30.90}$$

Notons qu'à gauche la limite est une limite dans $\mathbb{R}$ tandis qu'à droite c'est une limite dans $\mathscr{D}(\Omega)$. Ensuite le lemme 30.25 nous dit que l'hypothèse $\varphi_n \xrightarrow{\mathscr{D}(\Omega)} \varphi$ signifie en particulier que nous avons un compact $K \subset \Omega$ contenant tous les supports des φ_n et que $\partial^\alpha \varphi_n$ converge uniformément (sur K et donc sur Ω) vers $\partial^\alpha \varphi$. Donc

$$\lim_{n\to\infty} (\partial^\alpha T)(\varphi_n) = (-1)^{|\alpha|} T\big(\lim_{n\to\infty} \partial^\alpha \varphi_n\big) = (-1)^{|\alpha|} T\big(\partial^\alpha \varphi\big) = (\partial^\alpha T)(\varphi), \tag{30.91}$$

ce qui est la relation demandée. □

Le lemme suivant montre une compatibilité entre la dérivée des distributions, la dérivée faible et l'injection de L^2 dans l'espace des distributions.

LEMooQRUOooWVjCAV

Lemme 30.38.
Soit Ω un ouvert bornée de $\mathbb{R}^n$ et $f \in L^2(\Omega)$. Alors nous avons

$$\partial_i(T_f) = T_{\partial_i f} \tag{30.92}$$

où la dérivée à droite est la dérivée faible définie en 30.8.

Démonstration. En utilisant la définition de la dérivation de distribution, pour tout $\phi \in \mathscr{D}$ nous avons

$$\partial_i(T_f)\phi = -T_f(\partial_i\phi) = -\langle f, \partial_i\phi\rangle = \langle \partial_i f, \phi\rangle = T_{\partial_i f}(\phi). \tag{30.93}$$

Nous avons utilisé la définition (30.20) de la dérivée faible. □

Nous avons déjà vu dans l'exemple 30.4 que la fonction de Heaveside n'a pas de dérivée faible. Nous allons à présent voir que cette fonction a une dérivée au sens des distributions. Intuitivement, la fonction de Heaveside a une dérivée nulle partout sauf en $x = 0$ où sa dérivée serait infinie ; nous nous attendons à un delta de Dirac.

PROPooVUDVooAlwZzB

Proposition 30.39.
La dérivée de la fonction de Heaveside

$$\begin{aligned} H\colon \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto \begin{cases} 0 & si\ x \leqslant 0 \\ 1 & si\ x > 0 \end{cases} \end{aligned} \tag{30.94}$$

est le delta de Dirac.

Démonstration. Par définition, la dérivée de H au sens des distributions est la distribution H' qui fait

$$H'(\varphi) = -\langle H, \varphi'\rangle \tag{30.95}$$

pour tout élément $\varphi \in \mathscr{D}(\mathbb{R})$. Un petit calcul :

$$\begin{aligned} -\langle H, \varphi'\rangle &= -\int_{\mathbb{R}} H(t)\varphi'(t)dt && \text{(30.96a)} \\ &= -\int_0^\infty \varphi'(t)dt && \text{(30.96b)} \\ &= -\lim_{x\to\infty} \int_0^x \varphi'(t)dt && \text{(30.96c)} \\ &= -\lim_{x\to\infty} \big(\varphi(x) - \varphi(0)\big) && \text{(30.96d)} \\ &= \varphi(0). && \text{(30.96e)} \end{aligned}$$

SUBEQooMRMBooEGeDra
SUBEQooZDSDooTbzgWD
SUBEQooSUAHooOmjggr

Justifications :

— Pour (30.96b), c'est que $H(t) = 0$ pour $t \in]-\infty, 0[$.
— Pour (30.96c), c'est le lemme 14.260.
— Pour (30.96d), c'est le corolaire 14.299.

□

30.3.3 Ordre et support d'une distribution

DefVILMooBIYerO

Définition 30.40 (support d'une distribution[704])**.**
*Soit T une distribution. Le **support** de T est le complémentaire de l'union des ouverts $\mathcal{O}$ tels que $T(\varphi) = 0$ pour tout φ à support dans $\mathcal{O}$.*

DefXAHIooFeiRMB

Définition 30.41.
*Si T est une distribution sur Ω, nous disons que T est d'**ordre** inférieur ou égal à $p \in \mathbb{N}$ si pour tout compact K de Ω, il existe $C_K \in \mathbb{R}$ tel que pour tout $\varphi \in \mathscr{D}(K)$,*

$$\left|\langle T, \varphi\rangle\right| \leqslant C_K \max_{|\alpha|\leqslant p} \|\partial^\alpha \varphi\|_\infty. \tag{30.97}$$

Ici α est un multiindice.

La distribution T est d'ordre p si elle est d'ordre inférieur ou égal à p mais pas à $p-1$.

Pour la proposition suivante, on peut se remémorer la définition 30.15 de la topologie sur $C^\infty(\Omega)$.

PROPooSQPHooDPOrqK

Proposition 30.42 ([719])**.**
Restriction entre C^∞ et $\mathscr{D}$.

(1) Si $T \in C^\infty(\Omega)'$, alors la restriction de T à $\mathscr{D}(\Omega)$ est une distribution à support[16] compact.

(2) Si T est une distribution à support compact alors elle se prolonge de façon unique en une forme linéaire continue sur $C^\infty(\Omega)$.

PropZLUEooHcVxQj

Proposition 30.43 ([216])**.**
Une distribution à support compact est d'ordre fini.

LemYHRDooOdSnnK

Lemme 30.44 ([380])**.**
Soit $u \in \mathscr{D}'(\mathbb{R})$ et $\phi \in \mathscr{D}(\mathbb{R})$ tels que $\operatorname{supp}(u) \cap \operatorname{supp}(\phi) = \varnothing$. Alors $\langle u, \phi\rangle = 0$.

Démonstration. Soit $x \notin \operatorname{supp}(u)$. Alors il existe un voisinage V_x de x tel que $\langle u, \psi\rangle = 0$ pour tout $\psi \in \mathscr{D}(V_x)$. En particulier, si $x \in \operatorname{supp}(\phi)$, alors x n'est pas dans le support de u et les ensembles $\{V_x \text{ tel que } x \in \operatorname{supp}(\phi)\}$ recouvrent $\operatorname{supp}(\phi)$. Cependant ϕ est à support compact et nous pouvons extraire un sous-recouvrement fini de $\operatorname{supp}(\phi)$: il existe $x_1, \dots, x_p$ tels que

$$\operatorname{supp}(\phi) \subset \bigcup_{i=1}^{p} V_{x_i}. \tag{30.98}$$

Nous prenons une partition de l'unité[17] subordonnée à ce recouvrement. C'est-à-dire des fonctions $\chi_i \in \mathscr{D}(V_{x_i})$ telles que pour tout $x \in \operatorname{supp}(\phi)$,

$$\sum_{i=1}^{p} \chi_i(x) = 1. \tag{30.99}$$

En particulier nous avons $\sum_i \chi_i(x)\phi(x) = \phi(x)$, et donc

$$\langle u, \phi\rangle = \langle u, \sum \chi_i \phi\rangle = \sum \langle u, \chi_i \phi\rangle = 0 \tag{30.100}$$

parce que $\operatorname{supp}(\chi_i\phi) \subset V_{x_i}$. □

Lemme 30.45 ([380])**.**
Si u est une distribution d'ordre fini N sur $\mathbb{R}$, si $\operatorname{supp}(u) = \{x_0\}$ et si

$$\phi(x_0) = \dots = \phi^{(N)}(x_0) = 0 \tag{30.101}$$

alors $\langle u, \phi\rangle = 0$.

16. Définition 30.40.
17. Théorème 27.75.

Démonstration. Les fonctions plateaux dont nous avons parlé dans la section 15.14.1 nous permettent de considérer une fonction $\chi \in \mathscr{D}(\mathbb{R})$ vérifiant

$$\chi(x) = \begin{cases} 1 & \text{si } x \in \overline{B(0,1)} \\ 0 & \text{si } |x| > 2 \end{cases} \tag{30.102}$$

Ensuite nous posons $\chi_n(x) = \chi\big(n(x-x_0)\big)$. Par conséquent $\chi_n(x_0) = \chi(0) = 1$ et même $\chi_n(x_0+\epsilon) = \chi(\epsilon) = 1$ tant que ϵ est plus petit que disons $\frac{1}{2n}$ pour être sûr. Nous en déduisons que la fonction $1-\chi_n$ s'annule sur un voisinage de x_0 et que donc x_0 n'est pas dans le support de $1-\chi_n$. Donc $\text{supp}(1-\chi_n) \cap \text{supp}(u) = \varnothing$ et le lemme 30.44 est utilisable : $\langle u, (1-\chi_n\phi)\rangle = 0$, ou encore :

$$\langle u, \phi\rangle = \langle u, \chi_n\phi\rangle \tag{30.103}$$

pour tout n. Vu que le but est de prouver que $\langle u, \phi\rangle = 0$, nous allons prouver que

$$|\langle u, \chi_n\phi\rangle| \overset{n\to\infty}{\longrightarrow} 0. \tag{30.104}$$

Dans ce dessein nous posons

$$\|f\|_n = \sup_{x\in\overline{B(0,\frac{2}{n})}} \|f(x)\| \tag{30.105}$$

et

$$\|f\|_{(p)} = \sup_{i\leqslant p} \|\partial^i f\|_\infty. \tag{30.106}$$

La distribution u est d'ordre fini N, et nous en écrivons la définition 30.41 en prenant $\text{supp}(\chi_n\phi)$ en tant que K :

$$\big|\langle u, \chi_n\phi\rangle\big| \leqslant C \max_{k\leqslant N} \|\partial^k(\chi_n\phi)\|_\infty. \tag{30.107}$$

En remplaçant le maximum par une somme de $k=0$ à $k=N$, nous majorons. De plus le support de χ_n étant contenu dans $B_n = B(x_0, 2/n)$ nous ne changeons rien en utilisant $\|.\|_n$ au lieu de $\|.\|_\infty$. Donc

$$|\langle u, \chi_n\phi\rangle| \leqslant C \sum_{k=0}^{N} \|\partial^k(\chi_n\phi)\|_n \leqslant C \sum_{k=0}^{N}\sum_{i=0}^{k} \binom{k}{i} \|\partial^i\chi_n\|_n \|\partial^{k-i}\phi\|_n. \tag{30.108}$$

Notons que la seconde inégalité est une inégalité du type $\|fg\| \leqslant \|f\|\|g\|$. En dérivant un petit peu nous trouvons que

$$(\partial^i\chi_n)(x) = n^i(\partial^i\chi)\big(n(x-x_0)\big). \tag{30.109}$$

Donc [18]

$$\|\partial^i\chi_n\|_n = \sup_{x\in B_n} n^i\big|(\partial^i\chi)\big(n(x-x_0)\big)\big| = n^i \sup_{y\in[-2,2]} \big|(\partial^i\chi)(y)\big| = n^i\|\partial^i\chi\|_\infty. \tag{30.110}$$ EqFKWTooFfZoSM

Nous pouvons donc remplacer $\|\partial^i\chi_n\|_n$ par $n^i\|\partial^i\chi\|_\infty$.

D'autre part nous voulons majorer $\|\partial^{k-i}\phi\|_n$ par quelque chose ne dépendant ni de k ni de i. Nous faisons le théorème des accroissements finis 10.292 : $\|\partial^l\phi\|_n \leqslant \frac{2}{n}\|\partial^{l+1}\phi\|_n$. Ce n au dénominateur est salutaire parce que nous avions un n^i apparu à cause du remplacement (30.110). Nous faisons donc $i+1$ fois le théorème des accroissements finis :

$$\|\partial^{k-i}\phi\|_n \leqslant \left(\frac{2}{n}\right)^{i+1} \|\partial^{k+1}\phi\|_n. \tag{30.111}$$

Toutes ces majorations donnent

$$\big|\langle u, \chi_n\phi\rangle\big| \leqslant C \sum_{k=0}^{N}\sum_{i=0}^{k} \binom{k}{i} n^i \underbrace{\|\partial^i\chi\|_\infty}_{\leqslant\|\chi\|_{(N)}} \left(\frac{2}{n}\right)^{i+1} \underbrace{\|\partial^{k+1}\phi\|_n}_{\leqslant\|\phi\|_{(N+1)}} \tag{30.112a}$$

$$\leqslant C\|\chi\|_{(N)}\|\phi\|_{(N+1)} \frac{1}{n} \sum_{k=0}^{N}\sum_{i=0}^{k} \binom{k}{i} 2^{i+1} \tag{30.112b}$$

$$= \frac{C'}{n} \tag{30.112c}$$

18. Dans [380], la dernière égalité vient avec une inégalité, et je comprends pas pourquoi.

où C' est une constante qui dépend de χ, de ϕ et de N, mais pas de n. Vu que $\frac{C'}{n} \to 0$ nous avons bien

$$\langle u, \phi\chi_n \rangle = 0, \tag{30.113}$$

ce qu'il fallait démontrer. □

PropXXPLooSkgxOz

Proposition 30.46 ([380]).
Soit $u \in \mathscr{D}(\mathbb{R})'$ avec $\text{supp}(u) = \{x_0\}$. Alors $u = \sum_{i=0}^{N} a_i \partial^i \delta_{x_0}$ où N est l'ordre de u.

Démonstration. D'abord il faut préciser que l'ordre de u est fini parce que son support est compact (proposition 30.43) ; nous notons N cet ordre.

Soit $\phi \in \mathscr{D}(\mathbb{R})$. Nous considérons $\chi \in \mathscr{D}(\mathbb{R})$ telle que

$$\chi(x) = \begin{cases} 1 & \text{si } x \in \overline{B(x_0, 1)} \\ 0 & \text{si } |x - x_0| > 2. \end{cases} \tag{30.114}$$

Encore une fois, $1 - \chi$ s'annule sur un voisinage autour de x_0, ce qui fait que

$$\text{supp}(u) \cap \text{supp}\big((1 - \chi)\phi\big) = \varnothing, \tag{30.115}$$

et donc $\langle u, (1 - \chi)\phi \rangle = 0$. Au final,

$$\langle u, \phi \rangle = \langle u, \chi\phi \rangle. \tag{30.116}$$

C'est le moment de poser

$$\psi(x) = \chi(x)\big[\phi(x) - \sum_{k=1}^{N} \frac{1}{k!}(\partial^k \phi)(x_0)(x - x_0)^k\big] \tag{30.117}$$

La fonction ψ ayant un support disjoint de celui de u, nous avons aussi $\langle u, \psi \rangle = 0$, ce qui donne

$$\langle u, \phi \rangle = \langle u, \chi\phi \rangle = \langle u, \chi \sum_{k=0}^{N} \frac{1}{k!}(\partial^k \phi)(x_0)(x - x_0)^k \rangle. \tag{30.118}$$

En posant $a_k = \frac{1}{k!}\langle u, x \mapsto \chi(x)(x - x_0)^k \rangle$ nous avons alors

$$\langle u, \phi \rangle = \sum_{k=0}^{N} a_k (\partial^k \phi)(x_0) = \sum_k (-1)^k a_k (\partial^k \delta_{x_0})(\phi). \tag{30.119}$$

□

30.4 L'espace $C^\infty(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d))$

SecTEgDVWO

D'abord parlons un peu de continuité en recopiant la proposition 7.362 dans notre contexte.

PropIPlKQBa

Proposition 30.47.
Soient I un intervalle ouvert de $\mathbb{R}$ et $u\colon I \to \mathscr{D}'(\mathbb{R}^d)$ une fonction continue. Alors

ItemYAhnNhBi

(1) Pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$, l'application $t \mapsto u_t(\varphi)$ est continue.

(2) Pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$, nous avons la limite

$$\lim_{t \to t_0} u_t(\varphi) = u_{t_0}(\varphi). \tag{30.120}$$

(3) Nous avons la limite dans $\mathscr{D}'(\mathbb{R}^d)$

$$\lim_{t \to t_0} u_t = u_{t_0}. \tag{30.121}$$

En ce qui concerne la définition de l'espace $C^\infty(I, \mathscr{D}'(\mathbb{R}^d))$, c'est la définition 7.363. Grâce au point (1) de la proposition 30.47, nous retenons que la propriété fondamentale d'une application $T \in C^k\big(I, \mathscr{D}'(\Omega)\big)$ est que pour tout $\varphi \in \mathscr{D}(\Omega)$, l'application

$$\begin{aligned} I &\to \mathbb{C} \\ t &\mapsto T_t(\varphi) \end{aligned} \tag{30.122}$$

est de classe C^k.

PropOTlWzog

Proposition 30.48.
Soit I, un intervalle ouvert de $\mathbb{R}$. Soit $T \in C^0\big(I, \mathscr{D}'(\Omega)\big)$ et $\psi \in \mathscr{D}(I \times \Omega)$. Alors l'application

$$t \mapsto T_t\big(\psi(t, .)\big) \tag{30.123}$$

est continue sur I.

Démonstration. La fonction dont nous voulons prouver la continuité est une fonction $\mathbb{R} \to \mathbb{R}$; il est donc loisible de se contenter de la continuité séquentielle. Soient $t_0 \in I$ et (t_j) une suite dans I convergeant vers t_0. Nous posons $U_j = T_{t_j}$ et $\psi_j = \psi\big(t_j, .\big)$. Par hypothèse de continuité de (T_t) nous avons $U_j \xrightarrow{\mathscr{D}'(\Omega)} T_{t_0}$. D'autre part le support de ψ étant compact nous avons $\operatorname{supp}(\psi) \subset [c, d] \times K$ où $[c, d] \subset I$ et K est compact dans Ω. Par conséquent nous avons aussi $\operatorname{supp}(\psi_j) \subset K$.

Afin d'alléger les notations notons $\tilde{\psi}(x) = \psi(t_0, x)$. Pour tout multiindice α et pour tout j nous avons

$$p_\alpha(\psi_i - \tilde{\psi}) = \Big|\partial^\alpha \psi(t_j, x) - \partial^\alpha \psi(t_0, x)\Big| \leqslant |t_j - t_0| \sup_{\substack{t \in [c,d] \\ x \in K}} |\partial_t \partial^\alpha \psi(t, x)| \to 0. \tag{30.124}$$

Nous avons donc la convergence

$$\psi_j \xrightarrow{\mathscr{D}(K)} \psi(t_0, .). \tag{30.125}$$

Étant donné que $U_j \xrightarrow{\mathscr{D}'(\Omega)} T_{t_0}$ et $\psi_j \xrightarrow{\mathscr{D}(\Omega)} \tilde{\psi}$, le corolaire 27.6(4) nous donne la convergence

$$U_j(\psi_j) \xrightarrow{\mathbb{C}} T_{t_0}(\tilde{\psi}) \tag{30.126}$$

Cela est bien la continuité de la fonction $t \mapsto T_t\big(\psi(t, .)\big)$. □

PropLKtBsVi

Proposition 30.49 ([718])**.**
Soit $(T_t) \in C^0\big(I, \mathscr{D}'(\Omega)\big)$. Nous définissons l'application $T \colon \mathscr{D}(\Omega) \to \mathbb{C}$ par la formule

$$T(\psi) = \int_I T_t\big(\psi(t, .)\big)\, dt \tag{30.127}$$

pour tout $\psi \in \mathscr{D}(I \times \Omega)$. Alors $T \in \mathscr{D}'(I \times \Omega)$.

Démonstration. La proposition 30.48 nous indique que la fonction $t \mapsto T_t\big(\psi(t, .)\big)$ est continue. Étant donné qu'elle est seulement non nulle sur un compact, l'intégrale

$$\int_I T_t\big(\psi(t, .)\big) dt \tag{30.128}$$

a un sens et est finie. L'application $T \colon \mathscr{D}(I \times \Omega) \to \mathbb{C}$ ainsi définie est linéaire. Il reste à voir qu'elle est continue. Pour cela nous allons utiliser le théorème 30.30(3) qui nous dit que nous pouvons nous fixer un compact $[c, d] \times K \subset I \times \Omega$ et considérer $\psi \in \mathscr{D}\big([c, d] \times K\big)$.

Soit, pour commencer, donnée une application $\varphi \in \mathscr{D}(K)$. L'application $t \mapsto T_t(\varphi)$ est continue et non nulle sur le compact $[c, d]$ et il existe donc $C_\varphi > 0$ tel que

$$|T_t(\varphi)| \leqslant C_\varphi \tag{30.129}$$

pour tout $t \in [c,d]$.

Nous voulons utiliser le théorème de Banach-Steinhaus dans sa version 27.5 sur la famille d'applications paramétrée par $u \in [c,d]$:

$$\begin{aligned} U_u \colon \mathscr{D}\big([c,d]\times K\big) &\to \mathbb{R} \\ \psi &\mapsto T_u\big(\psi(u,.)\big). \end{aligned} \tag{30.130}$$

Commençons par prouver que cela est une application continue pour chaque u. Ce sera le cas si la projection

$$\begin{aligned} \texttt{proj} \colon \mathscr{D}\big([c,d]\times K\big) &\to \mathscr{D}(K) \\ \psi &\mapsto \psi(u,.) \end{aligned} \tag{30.131}$$

est continue. Pour cela nous notons P_{kl} la seminorme sur $\mathscr{D}\big([c,d]\times K\big)$ donnée par

$$P_{k,l}(\psi) = \sum_{n\leqslant k} \sum_{|\alpha|\leqslant l} \sup_{\substack{t\in[c,d]\\ x\in K}} \big|\partial_t^n \partial^\alpha \psi(t,x)\big|. \tag{30.132}$$

Nous montrons que proj est séquentiellement continue ; étant donné que les topologies sur $\mathscr{D}(K)$ et $\mathscr{D}\big([c,d]\times K\big)$ sont données par des métriques (proposition 30.28), cela suffit pour assurer la continuité grâce à la proposition 7.255. Montrons que si $\psi_n \xrightarrow{\mathscr{D}([c,d]\times K)} 0$, alors $\texttt{proj}(\psi_n) \xrightarrow{\mathscr{D}(K)} 0$. Pour cela nous remarquons que

$$\begin{aligned} p_j\big(\texttt{proj}(\psi)\big) &= \sum_{|\alpha|\leqslant j} \sup_{x\in K} |\partial^\alpha \psi(u,x)| && \text{(30.133a)}\\ &\leqslant \sum_{|\alpha|\leqslant j} \sup_{t\in[c,d]} \sup_{x\in K} |\partial^\alpha \psi(t,x)| && \text{(30.133b)}\\ &= P_{0,j}(\psi). && \text{(30.133c)} \end{aligned}$$

Par conséquent

$$p_j\big(\texttt{proj}(\psi_n)\big) \leqslant P_{0,j}(\psi_n) \to 0 \tag{30.134}$$

où nous avons utilisé la proposition 7.344. Utilisant cette même proposition à l'envers, nous déduisons que $\texttt{proj}(\psi_n) \xrightarrow{\mathscr{D}(K)} 0$. Les applications U_u sont donc continues ; elles sont également bornées parce que si $\psi \in \mathscr{D}\big([c,d]\times K\big)$ nous avons

$$\sup_{u\in[c,d]} \big|U_u(\psi)\big| = \sup_u \big|T_u\big(\psi(u,.)\big)\big|, \tag{30.135}$$

et la continuité déjà évoquée, sur le compact $[c,d]$, nous dit que cette quantité est finie. Le théorème de Banach-Steinhaus peut maintenant être appliqué et il existe $C>0$ et $k,l\in\mathbb{N}$ tels que pour tout $\psi \in \mathscr{D}\big([c,d]\times K\big)$,

$$\big|U_u(\psi)\big| \leqslant C P_{k,l}(\psi) = C \sum_{|\alpha|\leqslant k} \sum_{n\leqslant l} \sup_{t,x} \big|\partial_t^n \partial^\alpha \psi(t,x)\big| \leqslant C \sum_{|\alpha|+n\leqslant k+l} \sup_{t,x} \big|\partial_t^n \partial^\alpha \psi(t,x)\big|. \tag{30.136}$$

Quelques remarques

— Nous n'avons pas mis de maximum devant le supremum (alors que la conclusion (27.7) en demande) parce que dans le cas des seminormes P_{kl}, c'est toujours celle avec k et l le plus grand possible qui sont les plus grandes parce qu'elles sont des sommes emboitées les unes les autres.

— La fusion de deux sommes est bien une majoration parce qu'il y a plus de termes dans la seconde que dans la première.

— La quantité la plus à droite est (à part le C) ce que nous pouvons noter $P_{k+l}(\psi)$: c'est bien une des seminormes associées à l'espace de dimension $d+1$.

Nous majorons maintenant $T(\psi)$ par

$$|T(\psi)| \leqslant \int_c^d |T_t\big(\psi(t,.)\big)|dt = \int_c^d |U_t(\psi)|dt \leqslant C|d-c|P_{k+l}(\psi). \tag{30.137}$$

Maintenant le théorème 30.30(3) appliqué à l'ouvert $I \times \Omega$ et avec ψ au lieu de φ nous informe que $T \in \mathscr{D}(I \times K)$. □

30.4.1 Dérivation

Quelques propriétés de dérivation des fonctions $I \to \mathscr{D}(\Omega)$ seront directement énoncées et démontrées dans le cas des distributions tempérées. Les résultats 30.84 et 30.85 seront a fortiori valables si nous remplaçons $\mathscr{S}$ par $\mathscr{D}$.

30.5 Une équation de distribution

Nous allons étudier l'équation

EqLLTPooJHUVyU
$$(x-x_0)^\alpha u = 0 \tag{30.138}$$

pour $u \in \mathscr{D}'(\mathbb{R})$ et $\alpha \in \mathbb{N}$ est donné fixé. Notons tout de suite que (30.138) est un petit abus de notation pour dire qu'en vertu de la définition 30.36 du produit d'une distribution par une fonction, pour tout $\phi \in \mathscr{D}(\mathbb{R})$, nous avons $u\Big(x \mapsto (x-x_0)^\alpha \phi(x)\Big) = 0$.

LemWIGKooQpGXoI

Lemme 30.50 ([380]).
Soit $\alpha \in \mathbb{N}$. Une solution à l'équation

EqKVNEooJNwsPc
$$(x-x_0)^\alpha u = 0 \tag{30.139}$$

est une distribution à support dans $\{x_0\}$ et d'ordre fini.

Démonstration. Nous commençons par prouver que u est une solution de (30.139) si et seulement si [19] $\langle u, \phi\rangle = 0$ pour tout $\phi \in \mathscr{D}$ telle que

EqYLIPooYBvzwC
$$\phi(x_0) = \ldots = \partial^{\alpha-1}\phi(x_0) = 0. \tag{30.140}$$

(1) **Condition nécéssaire** Supposons que u soit une solution. Alors le corolaire 17.28 du théorème de Hadamard donne $\psi \in \mathscr{D}(\mathbb{R})$ telle que $\phi(x) = (x-x_0)^\alpha\psi(x)$. Dans ce cas, si u est solution de (30.139), alors

$$0 = \langle (x-x_0)^\alpha u, \psi\rangle = \langle u, (x-x_0)^\alpha\psi(x)\rangle = \langle u, \phi\rangle. \tag{30.141}$$

Nous avons vu que si u est solution, alors $\langle u, \phi\rangle = 0$ pour tout ϕ satisfaisant la condition (30.140).

(2) **Condition suffisante** Supposons maintenant l'inverse : u est une distribution s'annulant sur toute fonction $\phi \in \mathscr{D}'$ satisfaisant (30.140). Nous allons alors prouver que u est une solution. Soit donc $\psi \in \mathscr{D}$ et calculons

$$\langle (x-x_0)u, \psi\rangle = \langle u, (x-x_0)\psi\rangle = 0 \tag{30.142}$$

parce que la fonction $(x-x_0)\psi(x)$ vérifie la condition (30.140).

Nous passons maintenant au cœur de la preuve : nous supposons que u est une solution. Si le support de ϕ est contenu dans $\mathbb{R}\backslash\{x_0\}$ alors ϕ est nulle dans un voisinage de x_0 (et donc $\partial^k\phi = 0$ pour tout k) et $\langle u, \phi\rangle = 0$. Autrement dit, pour tout $\phi \in \mathscr{D}\big(\mathbb{R}\backslash\{x_0\}\big)$ nous avons $\langle u, \phi\rangle = 0$, ce qui signifie que $\text{supp}(u) \cap \big(\mathbb{R}\backslash\{x_0\}\big) = \varnothing$ ou encore que $\text{supp}(u) = \{x_0\}$.

Maintenant que u a un support compact, la proposition 30.43 nous indique qu'elle est d'ordre fini. □

19. En réalité nous n'aurons besoin que de la condition nécessaire, en particulier pour le théorème 30.51.

ThoRDUXooQBlLNb

Théorème 30.51 ([380]).
Soit $\alpha \in \mathbb{N}$ et l'équation

$$(x - x_0)^\alpha u = 0 \tag{30.143}$$

EqDONTooKPfDWU

pour $u \in \mathscr{D}'(\mathbb{R})$. Les solutions sont les combinaisons linéaires des dérivées de δ_{x_0} jusqu'à la α^eexclue.

Démonstration. D'abord montrons que les $\partial^i\delta_{x_0}$ sont des solutions. Avec les définition 30.36 et 30.37 des dérivées de distributions et de leur produits avec des fonctions [20],

$$(x - x_0)^\alpha \partial^i\delta_{x_0}(\phi) = \delta_{x_0}\Big(\partial^i\big((x - x_0)^\alpha\phi(x)\big)\Big) \tag{30.144}$$

Si $i < \alpha$ alors dans chaque terme de Leibniz [21], il y aura un facteur $(x - x_0)$, et la prise de δ_{x_0} annulera. Si par contre $i \geqslant \alpha$ alors il y aura le terme

$$\binom{i}{\alpha}\partial^\alpha\big((x - x_0)^\alpha\big)\partial^{i-\alpha}\phi(x_0) = \binom{i}{\alpha}\alpha!(\partial^{i-\alpha}\phi)(x_0) \tag{30.145}$$

qui est le seul terme contenant $(\partial^{i-\alpha}\phi)(x_0)$. Il suffit alors de choisir $\phi \in \mathscr{D}(\mathbb{R})$ de sorte que

$$(\partial^k\phi)(x_0) = \begin{cases} 0 & \text{si } k \neq i - \alpha \\ 1 & \text{si } k = i - \alpha \end{cases} \tag{30.146}$$

et alors on est certain que le tout n'est pas nul, et donc que $(x - x_0)^\alpha(\partial^i\delta_{x_0}) \neq 0$.

Jusqu'ici nous avons prouvé que $\partial^i\delta_{x_0}$ est solution si et seulement si $0 \leqslant i < \alpha$.

Il faut encore prouver que les solutions sont toutes des combinaisons linéaires de dérivées de delta de Dirac centrées en x_0. Pour cela nous invoquons d'abord le lemme 30.50 qui nous assure que u est d'ordre fini et de support $\{x_0\}$. Ensuite la proposition 30.46 nous indique que u doit alors être une combinaisons linéaire de dérivées de Dirac. □

30.6 Localisation, principe de recollement

PROPooNCHIooNOPfBt

Proposition-Définition 30.52 ([243, 1]).
Soient un ouvert Ω de $\mathbb{R}^d$, un ouvert A dans Ω, et $T \in \mathscr{D}'(\Omega)$. Nous définissons

$$\begin{aligned} T|_A \colon \mathscr{D}(A) &\to \mathbb{C} \\ \phi &\mapsto T(\bar\phi) \end{aligned} \tag{30.147}$$

où $\bar\phi \colon \Omega \to \mathbb{C}$ est définie par

$$\bar\phi(x) = \begin{cases} \phi(x) & \text{si } x \in A \\ 0 & \text{sinon.} \end{cases} \tag{30.148}$$

Alors $T|_A \in \mathscr{D}'(A)$.

*Cette distribution $T|_A$ est la **restriction** de T à A. Elle sera aussi souvent notée $r_A(T)$.*

Démonstration. Nous allons prouver la séquentielle continuité de $T|_A$. Ce sera suffisant par le théorème 30.30(4). Soit donc une suite $\phi_n \xrightarrow{\mathscr{D}(A)} 0$.

(1) **Débroussailler** Nous devons prouver que $T|_A(\phi_n) \xrightarrow{\mathbb{C}} 0$, c'est-à-dire que

$$T(\bar\phi_n) \xrightarrow{\mathbb{C}} 0. \tag{30.149}$$

20. Comme souvent, dans l'expression suivante, il y a un abus de notation parce que x est une variable muette : il faudrait écrire « $x \mapsto$ » au début de la grande parenthèse.
21. Lemme 30.12.

(2) **Notations** La topologie sur $\mathscr{D}(A)$ est celle de la définition 30.20. Fixons un peu de notations. Si K est compact dans A et si $m \in \mathbb{N}$, nous posons

$$\begin{aligned} p_{K,m} \colon C^\infty(A) &\to \mathbb{R} \\ f &\mapsto \sum_{|\mu|\leq m} \|\partial^\mu f\|_K. \end{aligned} \tag{30.150}$$

EQooYZWHooVRmlU

De même, si K est compact dans Ω et si $m \in \mathbb{N}$ nous posons

$$\begin{aligned} q_{K,m} \colon C^\infty(\Omega) &\to \mathbb{R} \\ f &\mapsto \sum_{|\mu|\leq m} \|\partial^\mu f\|_K. \end{aligned} \tag{30.151}$$

EQooCNPJooSiqAzt

Les familles de seminormes $(p_{K,m})$ et $(q_{K,m})$ ne sont pas indexées par les mêmes ensembles et n'ont pas le même domaine. Bien que les formules (30.150) et (30.151) se ressemblent, elles n'ont rien à voir l'une avec l'autre.

(3) **La notion de convergence**

La proposition 30.21 dit que[22] $f_n \xrightarrow{\mathscr{D}(A)} 0$ si et seulement si $p_{K_0,m}(f_n) \to 0$ pour tout m pour un certain compact K_0 contenant tous les supports de f_n.

(4) $\bar{\phi}_n \xrightarrow{\mathscr{D}(\Omega)} 0$

Nous utilisons le théorème 30.27 dans les deux sens. D'abord, vu que $\phi_n \xrightarrow{\mathscr{D}(A)} 0$, il existe un compact K_0 tel que $\phi_n \in \mathscr{D}(K_0)$ pour tout n et $\phi_n \xrightarrow{\mathscr{D}(K_0)} 0$, c'est-à-dire tel que

$$p_{K_0,m}(\phi_n) \to 0 \tag{30.152}$$

EQooRRMUooIznPsE

pour tout $m \in \mathbb{N}$.
Le compact K_0 vérifie également $\bar{\phi}_n \in \mathscr{D}(K_0)$ pour tout n, vu que le support de $\bar{\phi}_n$ est le même que celui de ϕ_n. De plus sur K_0 nous avons $\phi_n = \bar{\phi}_n$, donc

$$q_{K_0,m}(\bar{\phi}_n) = \sum_{|\mu|\leq m} \|\partial^\mu \bar{\phi}_n\|_{K_0} = \sum_{|\mu|\leq m} \|\partial^\mu \phi_n\|_{K_0} = p_{K_0,m}(\phi_n). \tag{30.153}$$

En vertu de (30.152) nous avons donc $q_{K_0,m}(\bar{\phi}_n) \to 0$, c'est-à-dire que $\bar{\phi}_n \xrightarrow{\mathscr{D}(K_0)} 0$. Donc le théorème 30.27 (dans l'autre sens, cette fois) nous indique que $\bar{\phi}_n \xrightarrow{\mathscr{D}(\Omega)} 0$.

(5) **Conclusion**

Vu que $T \in \mathscr{D}'(\Omega)$, nous avons $T(\bar{\phi}_n) \xrightarrow{\mathbb{C}} 0$. Donc

$$T|_A(\phi_n) = T(\bar{\phi}_n) \xrightarrow{\mathbb{C}} 0. \tag{30.154}$$

Cela prouve que $T|_A \in \mathscr{D}'(A)$ parce que ϕ_n est une suite quelconque tendant vers 0 dans $\mathscr{D}(A)$.

□

LEMooCXIZooAbeMpF

Lemme 30.53 ([1]).
Quelques propriétés de la restriction d'une distribution. Nous considérons un ouvert Ω dans $\mathbb{R}^d$ ainsi qu'un ouvert $A \subset \Omega$.

(1) L'application de restriction $r_A \colon \mathscr{D}'(\Omega) \to \mathscr{D}'(A)$ est linéaire.

ITEMooGXSKooYomqpg

(2) Si $\phi \in C^\infty(\mathbb{R}^d)$ vérifie $\operatorname{supp}(\phi) \subset A$, alors

$$T(\phi) = r_A(T)\phi. \tag{30.155}$$

22. Pour rappel, $\mathscr{D}$ signifie C_c^∞.

PROPooDVYNooKquEUl

Proposition 30.54 (Principe de recollement[243, 1]).
Soient un ouvert Ω de $\mathbb{R}^d$ ainsi que des ouverts $\{A_i\}_{i\in I}$ de $\mathbb{R}^d$ tels que $\bigcup_{i\in I} A_i = \Omega$.

Pour chaque $i \in I$ nous supposons avoir un élément $T_i \in \mathscr{D}'(A_i)$ tel que

$$r_{A_i\cap A_j}(T_i) = r_{A_i\cap A_j}(T_j) \tag{30.156}$$

pour tout $i, j \in I$ tels que $A_i \cap A_j \neq \varnothing$.

Alors il existe une unique distribution $T \in \mathscr{D}'(\Omega)$ telle que $r_{A_i}(T) = T_i$ pour tout i dans I.

Démonstration. Nous prouvons l'unicité et l'existence séparément [23].

(1) **Unicité** Soient T et U des distributions qui satisfont à la demande. Nous posons $S = T - U$ et nous allons prouver que $S = 0$. Grâce à la linéarité de r_{A_i} (lemme 30.53),

$$r_{A_i}(S) = 0 \tag{30.157}$$

pour tout i dans I.

Soit $\phi \in \mathscr{D}(\Omega)$; nous notons $K = \text{supp}(\phi) \subset \Omega$. Vu que les A_i recouvrent Ω, ils recouvrent K. Il existe dont une partie finie J de I telle que $K \subset \bigcup_{j\in J} A_j$. Et si nous prenions une partition de l'unité [24] subordonnée à ces A_j ?

Soit $\{\psi_j\}_{j\in J}$ une telle partition de l'unité. Vu que le support de ϕ est borné, le support de $\psi_k\phi$ est également borné. Et comme un support est toujours fermé, l'application $\psi_k\phi$ est dans $\mathscr{D}(A_k)$. En utilisant consciencieusement le lemme 30.53(2) nous avons

$$S(\phi) = S\Big(\sum_{k=1}^{n} \psi_k\phi\Big) = \sum_k S(\psi_k\phi) = \sum_{k=1}^{n} (r_{A_k}S)(\psi_k\phi) = 0 \tag{30.158}$$

parce que $r_{A_k}(S) = 0$.

(2) **Existence, début** La preuve de l'existence va se faire en plusieurs étapes.

(3) **Définition de T_K** Soit un compact K dans Ω. Vu que les A_j forment un recouvrement de K, il existe une partie finie J dans I telle que $K \subset \bigcup_{j\in J} A_j$. Nous considérons une partition de l'unité [25] $\{\psi_j\}_{j\in J}$ sur K subordonnée au recouvrement $\{A_j\}_{j\in J}$.

Nous définissons alors

$$\begin{aligned} T_K\colon \mathscr{D}(K) &\to \mathbb{C} \\ \phi &\mapsto \sum_{j\in J} T_j(\psi_j\phi). \end{aligned} \tag{30.159}$$

Nous supposons avoir fait, pour chaque compact K de Ω, un choix de sous-recouvrement fini (c'est-à-dire de partie J) et un choix de partition de l'unité.

(4) **$T_K = T_{K'}$ sur les intersection** Soient deux compacts K, M dans Ω. Nous prouvons que si $\phi \in \mathscr{D}(K) \cap \mathscr{D}(M)$, alors $T_K(\phi) = T_M(\phi)$. Nous avons donc en main les objets suivants :

$$\begin{aligned} &\big(K, J, \{A_j\}_{j\in J}, \{\chi_j\}_{j\in J}\big) \\ &\big(M, L, \{A_l\}_{l\in L}, \{\varphi_l\}_{l\in L}\big). \end{aligned} \tag{30.160}$$

Nous considérons $\phi \in \mathscr{D}(K) \cap \mathscr{D}(M)$. Vu que $\{\varphi_l\}_{l\in L}$ est une partition de l'unité sur M et que ϕ a son support contenu dans M, nous avons $\phi = \sum_{l\in L} \varphi_l\phi$, et nous pouvons écrire

$$T_K(\phi) = \sum_{j\in J} T_j(\chi_j\phi) = \sum_{j\in J}\sum_{l\in L} T_j(\chi_j\varphi_l\phi). \tag{30.161}$$ EQooXKYQooFUUoyL

23. Ce genre phrase semble ne servir à rien, mais elle sert à éviter qu'en l'environnement `description` qui suit ne soit tout moche.

24. Théorème 27.75.

25. Théorème 27.75.

Nous savons que $\operatorname{supp}(\chi_j\varphi_l\phi)$ est dans $A_j\cap A_l$, donc

$$T_j(\chi_j\varphi_l\phi)=r_{A_j\cap A_l}(T_j)(\chi_j\varphi_l\phi). \tag{30.162}$$

Mais l'hypothèse sur les T_i est que $r_{A_j\cap A_l}(T_j)=r_{A_j\cap A_l}(T_l)$. Cela nous permet de continuer le calcul de (30.161) :

$$T_K(\phi)=\sum_{j\in J}\sum_{l\in L}r_{A_j\cap A_l}(T_j)(\chi_j\varphi_l\phi)=\sum_{jl}r_{A_j\cap A_l}(T_l)(\chi_j\varphi_l\phi)=\sum_{jl}T_l(\chi_j\varphi_l\phi) \tag{30.163a}$$

$$=\sum_{l\in L}T_l\big(\varphi_l\sum_{j\in J}\chi_j\phi\big)=\sum_{l\in L}T_l(\varphi_l\phi)=T_M(\phi). \tag{30.163b}$$

Bien.

(5) **Ce qu'on pose** Si $\phi\in\mathscr{D}(\Omega)$, la valeur de $T_K(\phi)$ ne dépend pas du choix du compact K dans Ω contenant le support de ϕ. Nous pouvons donc poser

$$\begin{aligned}T\colon\mathscr{D}(\Omega)&\to\mathbb{C}\\ \phi&\mapsto T_K(\phi)\end{aligned} \tag{30.164}$$

où K est un compact quelconque de Ω contenant le support de ϕ.

(6) **T est une distribution** Nous devons prouver que $T\colon\mathscr{D}(\Omega)\to\mathbb{C}$ est une distribution, c'est-à-dire qu'elle est linéaire et continue. Pour la linéarité nous disons que c'est bon, et nous nous concentrons sur la continuité. Pour cela nous allons nous baser sur le critère du théorème 30.30(3).

Soit un compact K dans Ω. Nous devons trouver des constantes m et C telles que $|T(\phi)|\leqslant Cp_{m,K}(\phi)$ pour tout $\phi\in\mathscr{D}(K)$.

Vu que $\phi\in\mathscr{D}(K)$ nous avons $T(\phi)=T_K(\phi)=\sum_{j\in J}T_j(\chi_j\phi)$. Étant donné que T_j est continue sur $\mathscr{D}(A_j)$ et que $\chi_j\phi$ est une fonction de classe C^∞ à support dans A_j, en posant $M_j=\operatorname{supp}(\chi_j)$, il existe des constantes $m_j\in\mathbb{N}$ et $C_j\geqslant 0$ telles que

$$|T_j(\chi_j\phi)|\leqslant C_jp_{m_j,M_j}(\chi_j\phi). \tag{30.165}$$

Donc

$$|T(\phi)|\leqslant\sum_{j\in J}C_j|p_{m_j,M_j}(\chi_j\phi)|=\sum_{j\in J}C_j\sum_{|\alpha|\leqslant m_j}\|\partial^\alpha(\chi_j\phi)\|_{M_j}. \tag{30.166}$$

Nous utilisons la formule de Leibniz [26] pour décomposer $\partial^\alpha(\chi_j\phi)$:

$$|T(\phi)|\leqslant\sum_{j\in J}C_j\sum_{|\alpha|\leqslant m_j}\|\sum_{\beta\leqslant\alpha}\binom{\alpha}{\beta}\partial^{\alpha-\beta}\chi_j\partial^\beta\phi\|_{M_j} \tag{30.167a}$$

$$\leqslant\sum_jC_j\sum_{|\alpha|\leqslant m_j}\sum_{\beta\leqslant\alpha}\binom{\alpha}{\beta}\|\partial^{\alpha-\beta}\chi_j\|_{M_j}\|\partial^\beta\phi\|_{M_j}. \tag{30.167b}$$

Pour chaque j, il y a un nombre fini de multiindices α en jeu, et pour chacun d'entre eux, un nombre fini de β. Donc le coefficient $\binom{\alpha}{\beta}$ se majore par une constante dépendante de j ; nous l'incluons dans C_j. De même pour le coefficient $\|\partial^{\alpha-\beta}\chi_j\|$; il se majore et s'absorbe dans C_j. Nous avons donc

$$|T(\phi)|\leqslant\sum_{j\in J}C'_j\sum_{|\alpha|\leqslant m_j}\sum_{\beta\leqslant\alpha}\|\partial^\beta\phi\|_{M_j}. \tag{30.168}$$

Dans ces sommes, chaque nombre de la forme $\|\partial^\mu\phi\|_{M_j}$ avec $|\mu|\leqslant m_j$ est pris un certain nombre de fois. On majore ce « certain nombre [27] » et on l'inclut dans la constante C'_j. Cela donne

$$|T(\phi)|\leqslant\sum_{j\in J}C''_j\sum_{|\alpha|\leqslant m_j}\|\partial^\alpha\phi\|_{M_j}. \tag{30.169}$$

26. Lemme 30.12.

27. Il n'est pas nécessaire de savoir exactement combien, mais si vous y tenez vous pouvez vous embarquer dans un petit peu de combinatoire.

L'ensemble $M = \bigcup_{j\in J} M_j$ est un compact dans Ω et bien entendu $\|\partial^\alpha\phi\|_{M_j} \leqslant \|\partial^\alpha\phi\|_M$. Nous avons donc encore plein de majorations :

$$|T(\phi)| \leqslant \sum_{j\in J} C''_j \sum_{|\alpha|\leqslant m_j} \|\partial^\alpha\phi\|_M \tag{30.170a}$$
$$\leqslant C \sum_{j\in J} \sum_{|\alpha|\leqslant m_j} \|\partial^\alpha\phi\|_M \tag{30.170b}$$
SUBEQooOALIooTIhSXJ
$$\leqslant C \sum_{j\in J} \sum_{|\alpha|\leqslant m} \|\partial^\alpha\phi\|_M \tag{30.170c}$$
SUBEQooMNSNooVcTHXP
$$\leqslant C' \sum_{|\alpha|\leqslant m} \|\partial^\alpha\phi\|_M \tag{30.170d}$$
SUBEQooKJYOooHNwQAe
$$= C' p_{m,M}(\phi). \tag{30.170e}$$

Justifications :

— Pour 30.170b. On définit $C = \max_{j\in J} C''_j$.

— Pour 30.170c. On définit $m = \max_{j\in J} m_j$.

— Pour 30.170d. Comme ce qui est dans la somme ne dépend plus de j, nous supprimons la somme sur j et nous incluons le nombre de termes de la somme dans le changement $C \to C'$.

Et maintenant on admire le fait qu'en ayant fait autant de majorations dans tous les sens, nous obtenons encore une inégalité qui ait un sens. Bref, nous avons trouvé un compact M, un nombre m ainsi qu'une constante C' tels que

$$|T(\phi)| \leqslant C' p_{M,m}(\phi) \tag{30.171}$$

pour tout $\phi \in \mathscr{D}(K)$. Notons que M, m et C' dépendent fortement de K, mais pas de ϕ. Cela prouve que T est continue par 30.30(3).

(7) **T répond à la question** Il nous reste à prouver que T est bien la distribution recollée, c'est-à-dire que $r_{A_i}(T) = T_i$ pour tout $i \in I$. Le domaine de T_i et de $r_{A_i}(T)$ est $\mathscr{D}(A_i)$. Pour être clair, les fonctions dans $\mathscr{D}(A_i)$ sont des fonctions définies sur $\mathbb{R}^d$, mais dont le support est compact dans A_i.

Utilisant la définition de la restriction, pour $\phi \in \mathscr{D}(A_i)$,

$$r_{A_i}(T)\phi = T(\bar{\phi}) \tag{30.172}$$

où

$$\bar{\phi}(x) = \begin{cases} \phi(x) & \text{si } x \in A_i \\ 0 & \text{sinon.} \end{cases} \tag{30.173}$$

Mais vu que $\phi \in \mathscr{D}(A_i)$ nous avons $\phi = \bar{\phi}$ et donc pas de pièges de ce côté.

Nous posons $M = \operatorname{supp}(\phi)$ et nous avons

$$r_{A_i}(T)(\phi) = T(\bar{\phi}) = T(\phi) = T_M(\phi). \tag{30.174}$$

Nous nommons $\big(L, \{\varphi_l\}_{l\in L}\big)$ les choix faits pour le compact M dans la construction de T_M. Nous avons alors

$$T_M(\phi) = \sum_{l\in L} T_l(\varphi_p\phi). \tag{30.175}$$

Mais le support de $\varphi_k\phi$ est dans $A_i \cap A_l$, de telle sorte que l'hypothèse faite sur les T_i aux intersections donne $T_l(\varphi_l\phi) = T_i(\varphi_l\phi)$, et donc que

$$T_M(\phi) = \sum_{l\in L} T_l(\varphi_l\phi) = T_i\Big(\sum_{l\in L} \varphi_l\phi\Big) = T_i(\phi). \tag{30.176}$$

La dernière égalité vient du fait que $\{\varphi_l\}_{l\in L}$ est une partition de l'unité sur (au moins) $\operatorname{supp}(\phi)$.

□

30.7 Permuter distributions, dérivées et intégrales

PROPooCNYTooWCKHpV

Proposition 30.55 ([243]).
Soient un ouvert $\Omega \subset \mathbb{R}^d$, une distribution $T \in \mathscr{D}'(\Omega)$ ainsi qu'une fonction $\phi \in C^\infty(\Omega \times \mathbb{R}^n, \mathbb{C})$ à support dans $K \times \mathbb{R}^n$ où K est compact dans Ω.

Alors

ITEMooBIVOooHwGglM

(1) La fonction

$$y \mapsto T\big(\phi(.,y)\big) \tag{30.177}$$

est de classe C^∞ sur $\mathbb{R}^n$.

(2) Pour toute liste d'indices α et pour tout $y_0 \in \mathbb{R}^n$, nous avons

EQooYMXXooYkceTv

$$\partial_y^\alpha \big(T\big(x \mapsto \phi(x,y)\big)\big)_{y=y_0} = T\big(x \mapsto (\partial_y^\alpha \phi)(x,y_0)\big). \tag{30.178}$$

Démonstration. Nous y allons par récurrence sur les dérivations, en commençant par $\alpha = (i)$. Pourvu que le membre de gauche de (30.178) existe, il est donné par

EQooBHSSooXDrYQh

$$\partial_y^{(i)}\big[T\big(\phi(x,y)\big)\big]_{y=y_0} = \lim_{t\to 0} \frac{T\big(\phi(x,y_0+te_i)\big) - T\big(\phi(x,y)\big)}{t}. \tag{30.179}$$

Vu que T est linéaire, nous allons travailler sur

EQooXEFAooNzLBZx

$$T\Big(x \mapsto \frac{\phi(x,y_0+te_i) - \phi(x,y_0)}{t}\Big). \tag{30.180}$$

(1) **Un développement** Nous commençons par fixer y_0, x et t et écrire un développement de Taylor à l'ordre 1 avec reste intégral[28] pour la fonction $v(y) = \phi(x,y)$ autour de $y = y_0$ et dans la direction $h = te_i$. La formule est

EQooIAXAooWFhNVv

$$\phi(x,y_0+te_i) = \phi(x,y_0) + \sum_i h_i(\partial_{y_i}\phi)(x,y_0) + \text{reste} \tag{30.181}$$

où $\partial_{y_i}\phi$ signifie la dérivée partielle par rappport à la i^{e} composante de la seconde variable de ϕ. Avec la notations en multiindices, ça peut aussi bien s'écrire $\partial_y^{e_i}\phi$. En ce qui concerne le reste, il faut considérer la somme $\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=2}}$. Cette somme contient, entre autres,

$$h^\alpha = h_1^{\alpha_1}\dots h_n^{\alpha_n} \tag{30.182}$$

où

$$h_j = \begin{cases} 0 & \text{si } j \neq i \\ t & \text{si } j = i. \end{cases} \tag{30.183}$$

Autrement dit $h^\alpha = 0$ dès que $\alpha \neq \alpha_i e_i$. Et comme nous devons avoir $|\alpha| = 2$, le seul terme non nul dans la somme est $\alpha = 2e_i$. Le reste est donc

$$\frac{t^2}{2}\int_0^1 (1-t)^2(\partial_y^{2e_i}\phi)(x,y_0+te_i) \tag{30.184}$$

parce que $(te_i)^{2e_i} = t^2$. Au final,

EQooENWWooYftPch

$$\phi(x,y_0+te_i) = \phi(x,y_0) + \sum_i h_i(\partial_{y_i}\phi)(x,y_0) + \frac{t^2}{2}\int_0^1 (1-t)(\partial_y^{2e_i}\phi)(x,y_0+te_i). \tag{30.185}$$

28. Donc $m = 1$ dans la proposition 20.149. N'hésitez pas à revoir les notations 3.47 pour les multiindices.

(2) **Régularité du reste** Nous regardons le développement (30.185) en disant que t et y_0 sont fixés. Les fonctions

$$x \mapsto \phi(x, y_0 + te_i) \tag{30.186a}$$
$$x \mapsto \phi(x, y_0) \tag{30.186b}$$
$$x \mapsto (\partial_{y_i}\phi)(x, y_0) \tag{30.186c}$$

sont de classe C^∞ à support compact dans K. Donc $x \mapsto r(x, y_0, te_i)$ est également C^∞ à support dans K.
Nous allons d'ailleurs tellement fixer y_0 et t que nous allons définir

$$\begin{aligned} s\colon \mathbb{R}^d &\to \mathbb{C} \\ x &\mapsto \frac{t^2}{2}\int_0^1 (1-u)(\partial_y^{(ii)}\phi)(x, y_0 + ute_i)du \end{aligned} \tag{30.187}$$

et voir $s \in \mathscr{D}(K)$. Nous pouvons donc parler de $T(s)$ sans peurs.

(3) **Beaucoup de majorations** Vu que K est un compact dans Ω et que T est dans $\mathscr{D}'(\Omega)$, le théorème 30.30(3) dit qu'il existe des constantes m et C telles que $|T(s)| \leqslant Cp_{m,K}(s)$. Nous avons alors les majorations suivantes : SUBEQSooYPLGooOFMSyJ

$$|T(s)| \leqslant Cp_{m,K}(s) \tag{30.188a}$$
$$= C\sum_{|\mu|\leqslant m} \|\partial^\mu s\|_K \tag{30.188b}$$
$$= \sum_{|\mu|\leqslant m} \frac{t^2}{2}\int_0^1 (1-u)(\partial_x^\mu\partial_y^{ii}\phi)(x, y_0 + tue_i)du \tag{30.188c}$$ SUBEQooKIETooZIpBBF
$$\leqslant \sum_{|\mu|\leqslant m} \frac{t^2}{2}\int_0^1 (1-u)\|(\partial_x^\mu\partial_y^{ii}\phi)\|_{K\times\overline{B(y_0,1)}}du \tag{30.188d}$$ SUBEQooDYJCooFAsAxE
$$= \frac{Ct^2}{2}\sum_{|\mu|\leqslant m} \|\partial_x^\mu\partial_y^{ii}\phi\|_{K\times\overline{B(y_0,1)}} \tag{30.188e}$$ SUBEQooULTUooYUHNQ
$$\leqslant C't^2 \max_{|\mu|\leqslant m} \|\partial_x^\mu\partial_y^{ii}\phi\|_{K\times\overline{B(y_0,1)}} \tag{30.188f}$$ SUBEQooQENJooOgJuxi

Justifications :

— Pour (30.188c). Pour évaluer $\partial^\mu s$, nous utilisons la proposition 17.21 pour permuter l'intégrale sur u avec la dérivée sur x.
— Pour (30.188d). Nous supposons que $|t| < 1$, de telle sorte que $|ut| < 1$ et que $y_0 + tue_i$ reste dans $\overline{B(y_0, 1)}$. Notez au passage qu'il n'est pas nécessaire de prendre la fermeture ; c'est juste pour le plaisir de rester sur un compact.
— Pour (30.188e). Évaluer l'intégrale : $\int_0^1(1-u)du = 1/2$.
— Pour (30.188f). Le changement de constante $C \to C'$ intègre le nombre de termes dans la somme qui a été remplacée par son maximum et le $1/2$.

(4) **La dérivée (enfin)** Nous reprenons le calcul de la dérivée laissé en (30.179) et (30.180). Nous avons

$$T\left(x \mapsto \frac{\phi(x, y_0 + te_i) - \phi(x, y_0)}{t}\right) = T\left(\frac{t(\partial_y^{(i)}\phi)(x, y_0) + s(x)}{t}\right) \tag{30.189a}$$
$$= T\Big(x \mapsto (\partial_y^{(i)}\phi)(x, y_0)\Big) + T\left(\frac{s(x)}{t}\right). \tag{30.189b}$$

Le premier terme n'a plus de dépendance en t. La limite $t \to 0$ est donc sans problèmes. Pour le second terme, en prenant la majoration de (30.188) nous avons

$$\left|\frac{s(x)}{t}\right| \leqslant C't \max_{|\mu|\leqslant m} \|\partial_x^\mu\partial_y^{(ii)}\phi\|_{K\times\overline{B(y_0,1)}}. \tag{30.190}$$

La limite $t \to 0$ de cela est zéro.

(5) **Conclusion** La limite du membre de droite de (30.179) existe et vaut

$$T\Big[x \mapsto (\partial_y^{(i)}\phi)(x, y_0)\Big]. \tag{30.191}$$

Nous avons prouvé que

$$\partial_y^{(i)}\Big[T\big(x \mapsto \phi(x, y)\big)\Big]_{y=y_0} = T\Big[x \mapsto (\partial_y^{(i)}\phi)(x, y_0)\Big]. \tag{30.192}$$

Une récurrence permet de passer de la liste d'indices (i) à une liste générale pour terminer la preuve du résultat (30.178). □

Théorème 30.56 (Permuter distribution et intégrale[243])**.**
Soit un ouvert Ω dans $\mathbb{R}^n$. Soient $T \in \mathscr{D}'(\Omega)$ et $\phi \in \mathscr{D}(\Omega \times \mathbb{R}^n)$. Alors

$$\int_{\mathbb{R}^n} T\big(\phi(., y)\big)dy = T\big(\int_{\mathbb{R}^n} \phi(., y)dy\big). \tag{30.193}$$

Démonstration. Nous faisons une récurrence sur la valeur de n en partant de ...

(1) $n = 1$ Vu que le support de ϕ est compact dans $\Omega \times \mathbb{R}$, la proposition 7.101 dit qu'il existe un compact K dans Ω et $R > 0$ tel que le support de ϕ soit dans $K \times \overline{B(0, R)}$.
Nous introduisons plein de fonctions intermédiaires.

(1a) **La fonction θ**
Considérons une seconde la fonction

$$\begin{aligned}\theta\colon \Omega &\to \mathbb{C}\\ x &\mapsto \int_{\mathbb{R}} \phi(x, z)dz.\end{aligned} \tag{30.194}$$

Son support est dans K. Est-elle de classe C^∞ ? Oui. En effet, la proposition 17.21 dit que l'on peut permuter les dérivées partielles et l'intégrale. Donc toutes les dérivées partielles de θ existent (et sont donc continues). Le théorème 12.330 nous indique alors que θ est de classe C^∞.

(1b) **La fonction ξ** Le corolaire 15.163 nous permet de considérer $\xi \in \mathscr{D}\big(B(0, R)\big)$ telle que

$$\int_{\mathbb{R}} \xi(x)dx = 1. \tag{30.195}$$

(1c) **La fonction ψ** Nous posons alors

$$\begin{aligned}\psi\colon \mathbb{R}^d \times \mathbb{R} &\to \mathbb{C}\\ (x, y) &\mapsto \phi(x, y) - \xi(y)\int_{\mathbb{R}} \phi(x, z)dz.\end{aligned} \tag{30.196}$$

C'est une fonction $\psi \in \mathscr{D}(\Omega \times \mathbb{R})$ telle que

$$\operatorname{supp}(\psi) \subset K \times \overline{B(0, R)}. \tag{30.197}$$

Prouvons que pour tout $x \in \Omega$ nous avons

$$\int_{\mathbb{R}} \psi(x, y)dy = \int_{-R}^{R} \psi(x, y)dy = 0. \tag{30.198}$$

La première égalité est simplement le fait que $\psi(x, y) = 0$ lorsque $y \notin \overline{B(0, R)}$. Pour le fait que ce soit nul,

$$\int_{\mathbb{R}} \psi(x, y)dy = \int_{\mathbb{R}} \phi(x, y)dy - \int_{\mathbb{R}} \xi(y)dy \int_{\mathbb{R}} \phi(x, z)dz = 0. \tag{30.199}$$

(1d) **La fonction σ** Nous posons enfin

$$\begin{aligned}\sigma\colon \Omega \times \mathbb{R} &\to \mathbb{C}\\ (x,y) &\mapsto \int_{-\infty}^{y} \psi(x,z)dz.\end{aligned} \tag{30.200}$$

Nous avons $\sigma \in C^\infty(\Omega \times \mathbb{R})$ et $\text{supp}(\sigma) \subset K \times \overline{B(0,R)}$.

Nous pouvons maintenant entrer dans le dur.

(1a) **Une première égalité** Nous montrons que

$$T\big(\sigma(.,y)\big) = \int_{-\infty}^{y} T\big(\psi(.,z)\big)dz. \tag{30.201}$$

EQooDRISooWsWXnj

En voyant les deux côtés de cette égalité à prouver comme des fonctions de y, nous allons montrer qu'elles sont dérivables, de même dérivée et égale en $y < -R$.

À gauche nous permutons T avec la dérivée par la proposition 30.55 :

$$\partial_y\big(T\big(\sigma(.,y)\big)\big)_{y=y_0} = T\big((\partial_y\sigma)(.,y_0)\big) = T\big(\psi(.,y_0)\big) \tag{30.202}$$

parce que $(\partial_y\sigma)(x,y_0) = \psi(x,y_0)$.

À droite, la proposition 14.328 donne la première dérivée :

$$\partial_y\Big(\int_{-\infty}^{y} T\big(\psi(.,z)\big)dz\Big)_{y=y_0} = T\big(\psi(.,y_0)\big). \tag{30.203}$$

Notez que la proposition 30.55(1) dit alors que les deux membres de (30.201) sont de classe C^∞.

Pour $y < -R$, d'une part $\sigma(.,y) = 0$ et d'autre part, $\int_{-\infty}^{y} T\big(\psi(.,z)\big)dz = 0$ parce que l'intégrale est entièrement sur un domaine où $\psi(.,y) = 0$.

Donc l'égalité (30.201) est établie.

(1b) **Une seconde égalité** Nous prouvons maintenant que

$$\int_{\mathbb{R}} T\big(\psi(.,z)\big)dz = 0. \tag{30.204}$$

EQooTKHVooEqhpHH

Nous nous souvenons que $\psi(.,r) = 0$ si $r > R$. Donc

$$\begin{aligned}\int_{\mathbb{R}} T\big(\psi(.,z)\big)dz &= \int_{-\infty}^{\infty} T\big(\psi(.,z)\big)dz && \text{(30.205a)}\\ &= \int_{-\infty}^{R} T\big(\psi(.,z)\big)dz && \text{(30.205b)}\\ &= T\big(\sigma(.,R)\big). && \text{(30.205c)}\end{aligned}$$

La dernière égalité est (30.201).

(1c) **On ressort la définition de ψ** C'est le moment de nous souvenir la définition de ψ et de la substituer dans (30.204) en tenant compte de la linéarité de T. Nous avons :

$$\begin{aligned}0 &= \int_{\mathbb{R}} T\big(\psi(.,z)\big)dz && \text{(30.206a)}\\ &= \int_{\mathbb{R}} T\big(\phi(.,z)\big)dz - \int_{\mathbb{R}} T\Big(\xi(z)\int_{\mathbb{R}}\phi(.,s)ds\Big)dz && \text{(30.206b)}\\ &= \int_{\mathbb{R}} T\big(\phi(.,z)\big)dz - \int_{\mathbb{R}} \xi(z)dz\, T\Big(\int_{\mathbb{R}}\phi(.,s)ds\Big) && \text{(30.206c)}\\ &= \int_{\mathbb{R}} T\big(\phi(.,z)\big)dz - T\Big(\int_{\mathbb{R}}\phi(.,s)ds\Big). && \text{(30.206d)}\end{aligned}$$

SUBEQooXSQMooEuUZUe

Justifications :

— Pour (30.206c). Le nombre $\xi(z)$ sort de T. Et $T\big(\int_{\mathbb{R}} \phi(.,s)ds\big)$ sort de l'intégrale en z.

Le cas $n=1$ est prouvé.

(2) **La récurrence** Nous supposons que le théorème est prouvé jusqu'à n et nous le prouvons pour $n+1$.
Soit donc $\phi \in \mathscr{D}(\Omega \times \mathbb{R}^{n+1})$. La fonction $y \mapsto T\big(\phi(.,y)\big)$ est continue et à support compact. Donc Fubini 14.301 s'applique et nous allons permuter les intégrales une à une avec T. Avant cela nous définissons quelques fonctions intermédiaires.
D'abord, pour $t \in \mathbb{R}$ nous posons

$$\begin{aligned}\tilde{\phi}_t\colon \Omega \times \mathbb{R}^n &\to \mathbb{R}\\ (\omega,y) &\mapsto \phi\big(\omega,(y,t)\big).\end{aligned} \tag{30.207}$$

Ensuite, nous posons

$$\begin{aligned}\bar{\phi}\colon \Omega \times \mathbb{R} &\to \mathbb{R}\\ (\omega,t) &\mapsto \int_{\mathbb{R}^n} \tilde{\phi}_t(\omega,y_1,\dots,y_n)dy_1\dots dy_n.\end{aligned} \tag{30.208}$$

Maintenant nous calculons :

$$\int_{\mathbb{R}^{n+1}} T\big(\phi(.,y)\big)dy = \int_{\mathbb{R}}\Big[\int_{\mathbb{R}^n} T\big(\phi(.,y)\big)dy_1\dots dy_n\Big]dy_{n+1} \tag{30.209a}$$
$$= \int_{\mathbb{R}}\Big[\int_{\mathbb{R}^n} T\big(\tilde{\phi}_{y_{n+1}}(.,y_1,\dots,y_n)\big)dy_1\dots dy_n\Big]dy_{n+1} \tag{30.209b}$$
$$= \int_{\mathbb{R}}\Big[T\Big(\int_{\mathbb{R}^n} \tilde{\phi}_{y_{n+1}}(.,y_1,\dots,y_n)\Big)dy_1\dots dy_n\Big]dy_{n+1} \tag{30.209c}$$
$$= \int_{\mathbb{R}} T\big(\bar{\phi}(.,y_{n+1})\big)dy_{n+1} \tag{30.209d}$$
$$= T\Big(\int_{\mathbb{R}} \bar{\phi}(.,y_{n+1})dy_{n+1}\Big) \tag{30.209e}$$
$$= T\Big(\int_{\mathbb{R}^{n+1}} \phi(.,y)dy\Big) \tag{30.209f}$$

Justifications :

— Pour (30.209a). Théorème de Fubini 14.301.

— Pour (30.209c). Hypothèse de récurrence.

— Pour (30.209e). Encore la récurrence, avec $n=1$.

— Pour (30.209f). Encore Fubini, mais dans l'autre sens.

□

30.8 Distributions tempérées

L'espace de Schwartz[29] $\mathscr{S}(\Omega)$ est défini dans la définition 27.233 ; sa topologie y est discutée.

Définition 30.57.
*Une **distribution tempérée** est une forme linéaire continue sur $\mathscr{S}(\mathbb{R}^d)$. L'ensemble des distributions tempérées est noté $\mathscr{S}'(\mathbb{R}^d)$. Si T est une telle distribution, nous notons $\langle T,\varphi\rangle$ l'image de φ par T.*

29. Attention : ce Schwartz (avec un t) est le Schwartz des distributions dont le prénom est Laurent. À ne pas confondre avec Schwarz (sans t) dont le prénom est Cauchy.

30.58.
Deux rappels.

(1) La topologique sur $\mathscr{S}(\mathbb{R}^d)$ est celle de la définition 27.238.
(2) La topologie sur $\mathscr{D}(\mathbb{R}^d)$ est donnée par la définition 30.15.

DEFooUSTNooYEZfPN

Lemme-Définition 30.59.
L'application

$$\begin{aligned}\delta\colon \operatorname{Fun}(\mathbb{R}^d) &\to \mathbb{C}\\ \varphi &\mapsto \varphi(0).\end{aligned} \tag{30.210}$$

est

(1) une distribution,
(2) une distribution tempérée.

Cette distribution est nommée ***distribution de Dirac****.*

Démonstration. Juste pour rappel, $\operatorname{Fun}(X)$ est l'ensemble de toutes les fonctions sur X. Pour prouver que δ est une distribution, nous devons démontrer que $\delta\colon \mathscr{D}(\mathbb{R}) \to \mathbb{C}$ est continue. Et pour qu'elle soit une distribution tempérée, il faut démontrer que $\delta\colon \mathscr{S}(\mathbb{R}) \to \mathbb{C}$ est continue.

Nous utilisons le théorème 30.30(3) pour démontrer la continuité de δ sur $\mathscr{D}(\mathbb{R})$. Soit un compact $K \subset \mathbb{R}$ et $\varphi \in \mathscr{D}(K)$. En prenant $m = 0$ nous devons avons la majoration

$$|\delta(\varphi)| = \varphi(0) \leqslant \|\varphi\|_{K,\infty} = p_{0,K}(\varphi). \tag{30.211}$$

Pour la continuité de δ sur $\mathscr{S}(\mathbb{R}^d)$, nous utilisons les résultats de 27.239. Soit une suite $\varphi_n \xrightarrow{\mathscr{S}} 0$. En particulier, $p_{0,0}(\varphi_n) = \sup_x |\varphi_n(x)| \to 0$. Donc $\varphi_n(0) \to 0$ comme il le faut. □

Exemple 30.60.
La **valeur principale** de la fonction $x \mapsto \frac{1}{x}$ est la distribution

$$\begin{aligned}T\colon \mathscr{S}(\mathbb{R}) &\to \mathbb{R}\\ \varphi &\mapsto \lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \int_{|x|>\epsilon} \frac{\varphi(x)}{x}.\end{aligned} \tag{30.212}$$

Montrons que cela définit bien une distribution tempérée.

D'abord l'intégrale existe pour tout ϵ, par exemple parce que pour les grands $|x|$ nous avons par exemple $|\varphi(x)| \leqslant \frac{1}{x}$ et donc $\varphi(x)/x \leqslant 1/x^2$ dont l'intégrale converge. Nous devons maintenant regarder la limite.

Nous considérons une suite $\epsilon_n \to 0$ et la suite

$$a_n = \int_{|x|\geqslant\epsilon_n} \frac{\varphi(x)}{x} dx. \tag{30.213}$$

Nous montrons que cette suite converge dans $\mathbb{R}$ en montrant qu'elle est de Cauchy. Pour cela nous travaillons un peu la forme de φ :

$$\varphi(x) = \varphi(0) + \int_0^x \varphi'(t)dt = \varphi(0) + \int_0^1 x\varphi'(x\theta)d\theta. \tag{30.214}$$

Ce qui est dans l'intégrale est borné par $K = \|M_x\varphi'\|_\infty$ qui est parfaitement fini parce que φ est à décroissance rapide. Lorsque nous calculons $|a_m - a_n|$, le terme $\varphi(0)/x$ donne une intégrale nulle parce que le domaine d'intégration $\epsilon_m \leqslant |x| \leqslant \epsilon_n$ est symétrique alors que la fonction $1/x$ est impaire.

$$|a_m - a_n| \leqslant \Big| \int_{\epsilon_m<|x|<\epsilon_n} K \Big| = 2|\epsilon_n - \epsilon_m|K \tag{30.215}$$

Tout cela nous dit que T est bien définie. Nous devons encore étudier sa continuité.

Soit χ une fonction dans $C_c^\infty(\mathbb{R})$ valant 1 sur $[-1,1]$, paire et à valeurs dans $[0,1]$. Pour tout $\epsilon > 0$ nous avons $\int_{|x|>\epsilon} \frac{\chi(x)}{x} dx = 0$.

Nous avons aussi $\varphi = \chi\varphi + (1-\chi)\varphi$, et donc

$$\int_{|x|>\epsilon} \frac{\varphi(x)}{x} dx = \int_{|\epsilon|>0} \chi(x) \frac{\varphi(x)-\varphi(0)}{x} dx + \int_{|\epsilon|>0} \big(1-\chi(x)\big)\frac{\varphi(x)}{x} dx \tag{30.216a}$$

$$= \int_{|\epsilon|>0} \chi(x) \int_0^1 \underbrace{\varphi'(\theta x)}_{\leqslant \|\varphi'\|_\infty} d\theta + \int_{|x|\geqslant 1} \big(1-\chi(x)\big)\frac{\varphi(x)}{x} dx \tag{30.216b}$$

$$\leqslant \|\varphi'\|_\infty \int_{|x|\geqslant\epsilon} \chi(x) dx + \|\varphi\|_{L^1} \tag{30.216c}$$

$$= C\|\varphi'\|_\infty + \|\varphi\|_1. \tag{30.216d}$$

Cela est valable pour toute fonction $\varphi \in \mathscr{S}(\mathbb{R})$. Mais nous savons que si $\varphi_n \overset{\mathscr{S}(\mathbb{R})}{\to} 0$, alors $\|\varphi_n\|_\infty \to 0$, $\|\varphi_n'\|_\infty \to 0$ et $\|\varphi_n\|_1 \to 0$; donc si $\varphi_n \overset{\mathscr{S}(\mathbb{R})}{\to} 0$, alors

$$T(\varphi_n) = \lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \int_{|x|>\epsilon} \frac{\varphi(x)}{x} \leqslant C\|\varphi_n'\|_\infty + \|\varphi_n\|_1 \to 0. \tag{30.217}$$

△

30.8.1 Topologie

La topologie que nous mettons sur l'espace $\mathscr{S}'(\mathbb{R}^d)$ est le même type que celle que nous mettons sur $\mathscr{D}'(\mathbb{R}^d)$.

DEFooNLIKooYfafLw

Proposition-Définition 30.61 (Topologie sur $\mathscr{D}(\mathbb{R}^d)$)**.**
Si $\varphi \in \mathscr{S}(\mathbb{R}^d)$, alors l'application

$$\begin{aligned} p_\varphi : \mathscr{S}'(\mathbb{R}^d) &\to \mathbb{R} \\ T &\mapsto |T(\varphi)| \end{aligned} \tag{30.218}$$

est une seminorme.

Nous considérons sur $\mathscr{S}(\mathbb{R}^d)$ la topologie donnée par ces seminormes.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

La définition 30.31 et la proposition 30.32 restent.

PropQAuJstI

Proposition 30.62.
Nous avons $T_n \overset{\mathscr{S}'(\mathbb{R}^d)}{\longrightarrow} T$ si et seulement si pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$ nous avons $T_n(\varphi) \to T(\varphi)$.

Démonstration. Il s'agit encore de la proposition 7.344. Nous avons

$$T_n \overset{\mathscr{S}'(\mathbb{R}^d)}{\longrightarrow} T \Leftrightarrow \forall \varphi \in \mathscr{S}(\mathbb{R}^d) p_\varphi(T_n - T) \to 0 \tag{30.219a}$$

$$\Leftrightarrow \forall \varphi \in \mathscr{S}(\mathbb{R}^d) |(T_n - T)(\varphi)| \to 0 \tag{30.219b}$$

$$\Leftrightarrow \forall \varphi \in \mathscr{S}(\mathbb{R}^d) |T_n(\varphi) - T(\varphi)| \to 0 \tag{30.219c}$$

$$\Leftrightarrow \forall \varphi \in \mathscr{S}(\mathbb{R}^d) T_n(\varphi) \to T(\varphi) \tag{30.219d}$$

□

30.8.2 Distributions associées à des fonctions

Si $f \in L^1_{loc}(\mathbb{R}^d)$ alors nous lui associons une distribution $T_f \in \mathscr{D}'(\mathbb{R}^d)$ définie par la formule

$$T_f(\varphi) = \int_{\mathbb{R}^d} f(x)\varphi(x) dx. \tag{30.220}$$

Proposition 30.63.
L'application ainsi définie

$$\begin{aligned} L^1_{loc}(\mathbb{R}^d) &\to \mathscr{D}'(\mathbb{R}^d) \\ f &\mapsto T_f \end{aligned} \tag{30.221}$$

est injective.

Démonstration. Si $T_f = 0$ alors pour tout $\varphi \in \mathscr{D}$ nous avons $\int_{\mathbb{R}^d} f\varphi = 0$. En vertu de la proposition 27.190 cela implique $f = 0$ presque partout. □

30.8.3 Composition avec une fonction

PropBQUOcyw

Proposition 30.64 ([718], page 113 et 32)**.**
Soit $T \in \mathscr{S}'(\Omega)$ et $f \in C^k(A \times \Omega)$ où A est ouvert dans $\mathbb{R}^d$. Nous posons

$$\begin{aligned} F\colon A &\to \mathbb{R} \\ \lambda &\mapsto T\big(f(\lambda, .)\big). \end{aligned} \tag{30.222}$$

Alors $F \in C^k(A)$.

30.8.4 Transformée de Fourier d'une distribution tempérée

Définition 30.65.
*La **transformée de Fourier** de la distribution tempérée $T \in \mathscr{S}'(\mathbb{R}^d)$ est la distribution $\hat{T}$ définie par*

$$\hat{T}(\varphi) = T(\hat{\varphi}) \tag{30.223}$$

pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

Lemme 30.66.
Si $f \in \mathscr{S}(\mathbb{R}^d)$, nous avons $\hat{T}_f = T_{\hat{f}}$.

Démonstration. En utilisant les définitions,

$$\hat{T}_f(\varphi) = T_f(\hat{\varphi}) = \int_{\mathbb{R}^d} f(x)\hat{\varphi}(x)dx = \int_{\mathbb{R}^d} f(x) \left[\int_{\mathbb{R}^d} \varphi(y) e^{-iyx} dy \right] dx \tag{30.224}$$

où nous avons noté xy le produit scalaire $x \cdot y$. Nous permutons les intégrales en utilisant le théorème de Fubini 14.304 avec la fonction

$$(x, y) \mapsto f(x)\varphi(y)e^{-ixy} \tag{30.225}$$

qui est parfaitement dans $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Nous écrivons alors

$$\hat{T}_f(\varphi) = \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} f(x)\varphi(y)e^{-iyx} dx \right] dy = \int_{\mathbb{R}^d} \varphi(y)\hat{f}(y)dy = T_{\hat{f}}(\varphi). \tag{30.226}$$

□

30.8.5 Convolution d'une distribution par une fonction

Nous savons que si $\psi \in \mathscr{S}(\mathbb{R}^d)$ et si $x \in \mathbb{R}^d$ alors la fonction $y \mapsto \psi(x - y)$ est encore une fonction dans $\mathscr{S}(\mathbb{R}^d)$. Donc si $T \in \mathscr{S}'(\mathbb{R}^d)$ nous pouvons considérer la fonction $T * \psi = \psi * T$ définie par

$$(T * \psi)(x) = T\big(y \mapsto \psi(x - y)\big). \tag{30.227}$$ EQooOUXKooGHDSzL

Notons que $T * \psi$ est bien une fonction et non une distribution.

Le but de la définition est d'avoir

$$T_f * \psi = f * \psi. \tag{30.228}$$

En effet

$$(T_f * \psi)(x) = T_f\big(y \mapsto \psi(x-y)\big) = \int_{\mathbb{R}^d} f(y)\psi(x-y)dy = (f*\psi)(x). \tag{30.229}$$

Exemple 30.67.
La distribution de Dirac est le neutre pour le produit de convolution. En effet

$$(\delta * \psi)(x) = \delta\big(y \mapsto \psi(x-y)\big) = \psi(x), \tag{30.230}$$

c'est-à-dire $\delta * \psi = \psi$. △

PropZMKYMKI

Proposition 30.68 ([704]).
Si $T \in \mathscr{S}'(\mathbb{R}^d)$ et $\psi \in \mathscr{S}(\mathbb{R}^d)$, alors la distribution associée à la fonction $T * \psi$ est tempérée.

Démonstration. En agissant sur $\varphi \in \mathscr{D}(\mathbb{R}^d)$ nous avons

$$\begin{aligned}
T_{T*\psi}(\varphi) &= \int_{\mathbb{R}^d} T\big(y \mapsto (t_x\psi)(y)\big)\varphi(x)dx && \text{(30.231a)}\\
&= \int_{\mathbb{R}^d} T\big(y \mapsto \varphi(x)\psi(x-y)\big)dx && \text{(30.231b)}\\
&= T\left(y \mapsto \int_{\mathbb{R}^d} \varphi(x)\psi(x-y)dx\right) && \text{(30.231c)}\\
&= T\big(y \mapsto (\varphi * \check{\psi})(y)\big) && \text{(30.231d)}\\
&= T(\varphi * \check{\psi}). && \text{(30.231e)}
\end{aligned}$$

SubEqSVDsVTS

Attention :

¡¡ Avertissement/question à la lectrice !! 30.69
Le passage à la ligne (30.231c) n'est pas justifié.

□

30.8.6 Approximation de la distribution de Dirac

LEMooHEEOooFtKgfz

Lemme 30.70 ([1]).
Soient des fonctions $j_n : \mathbb{R} \to \mathbb{R}^+$ de classe C^∞ telles que

(1) Pour chaque n, la fonction $x \mapsto j_n\big(|x|\big)$ est strictement décroissante et converge ponctuellement vers zéro.
(2) Pour chaque x, la suite $n \mapsto j_n(x)$ est décroissante et converge vers 0.
ITEMooFQYXooEkUAIb
(3) Pour tout $M > 0$, la suite j_n converge vers zéro uniformément sur $B(0,M)^c$.
ITEMooFYCRooFeRRjE
(4) Pour tout δ et ϵ, il existe un $N \in \mathbb{N}$ tel que $|\int_{B(0,\delta)} j_n(x)dx - 1| \leqslant \epsilon$.
(5) Pour tout n, nous avons $\int_{\mathbb{R}} j_n = 1$.

Alors si $u \in \mathscr{S}(\mathbb{R})$ nous avons

$$\lim_{n\to\infty} \int_{\mathbb{R}} u(x)j_n(x)dx = u(0). \tag{30.232}$$

Démonstration. Nous posons

$$\begin{aligned}
I_n &= \int_{\mathbb{R}} j_n u && \text{(30.233a)}\\
I_{\delta,n} &= \int_{B(0,\delta)} j_n u && \text{(30.233b)}\\
Z_{\delta,n} &= \int_{B(0,\delta)} u(0) j_n && \text{(30.233c)}
\end{aligned}$$

Nous allons progressivement montrer qu'en prenant δ assez petit et n assez grand, les quantités $|I_n - I_{\delta,n}|$, $|I_{\delta,n} - Z_{\delta,n}|$ et $|Z_{\delta,n} - u(0)|$ peuvent être simultanément majorées par ϵ.

Soient $\delta > 0$ et $\epsilon > 0$; vu que $u \in \mathscr{S}(\mathbb{R})$, il existe M tel que $\int_{|x|>M} |u| < \epsilon$. Soit $N_1 \in \mathbb{N}$ tel que pour tout $n > N_1$ nous avons $|j_n(x)| < 1$ dès que $|x| > M$ (hypothèse (3)). Alors

$$\int_{|x|>M} |j_n(x)u(x)| < \epsilon. \tag{30.234}$$

De plus en posant $s = \max\{|u(x)| \text{ tel que } \delta \leqslant |x| \leqslant M\}$ (qui existe parce que u est continue et prise sur un compact) nous pouvons considérer N_2 tel que $j_n(x) < \epsilon/s$ pour tout $|x| > \delta$.

Avec $n > \max\{N_1, N_2\}$ nous avons

$$\begin{aligned} |\int_{B(0,\delta)} j_n u - \int_{\mathbb{R}} j_n u| &= |\int_{B(0,\delta)^c} j_n u| && \text{(30.235a)}\\ &\leqslant \int_{\delta\leqslant|x|\leqslant M} |j_n u| + \int_{|x|\geqslant M} |j_n u| && \text{(30.235b)}\\ &\leqslant \epsilon(1 + |M - \delta|). && \text{(30.235c)} \end{aligned}$$

En redéfinissant le ϵ nous avons donc montré que pour tout ϵ et δ, il existe un $N \in \mathbb{N}$ tel que

$$|I_{\delta,n} - I_n| \leqslant \epsilon \tag{30.236}$$

dès que $n \geqslant N$.

La fonction u est uniformément continue sur tout $\overline{B(0,\delta)}$, et nous pouvons donc choisir δ tel que $|u(0) - u(x)| \leqslant \epsilon$ pour tout $x \in \overline{B(0,\delta)}$. Pour ce δ, nous avons déjà trouvé un N tel que $|I_{\delta,n} - I_n| \leqslant \epsilon$ dès que $n > N$. Nous avons :

$$\begin{aligned} |I_{\delta,n} - Z_{\delta,n}| &\leqslant \int_{B(0,\delta)} |u(x) - u(0)| j_n(x) dx && \text{(30.237a)}\\ &\leqslant \epsilon \int_{B(0,\delta)} j_n && \text{(30.237b)}\\ &\leqslant \epsilon. && \text{(30.237c)} \end{aligned}$$

Nous avons donc prouvé que pour tout $\epsilon > 0$, il existe un δ et un N tels que

$$\begin{cases} |I_{\delta,n} - I_n| \leqslant \epsilon & \text{(30.238a)}\\ |I_{\delta,n} - Z_{\delta,n}| \leqslant \epsilon & \text{(30.238b)} \end{cases}$$

dès que $n \geqslant N$.

Enfin nous avons

$$|Z_{\delta,n} - u(0)| = u(0) \left(\int_{B(0,\delta)} j_n - 1 \right), \tag{30.239}$$

et par l'hypothèse (4) nous pouvons choisir n assez grand pour que la parenthèse soit plus petite que ϵ.

Pour ϵ donné, nous avons donc trouvé un δ et un N tels que

$$|I_n - u(0)| \leqslant |I_n - I_{\delta,n}| + |I_{\delta,n} - Z_{\delta,n}| + |Z_{\delta,n} - u(0)| \leqslant 3\epsilon. \tag{30.240}$$

En passant à la limite nous avons bien $I_n \to u(0)$ dans $\mathbb{R}$. □

Il va sans dire que nous connaissons de telles fonctions. Nous en donnons une maintenant.

Exemple 30.71 ([720]).
Nous introduisons la fonction f_ϵ ($\epsilon > 0$) donnée par

$$f_\epsilon(x) = e^{-\epsilon|x|}. \tag{30.241}$$

Nous calculons la transformée de Fourier de f_ϵ en divisant le domaine d'intégration :

$$\hat{f}_\epsilon(k) = \int_{\mathbb{R}} e^{-ikx}e^{-\epsilon|x|}dx = \int_{-\infty}^{0} e^{\epsilon x}e^{-ikx}dx + \int_{0}^{\infty} e^{-\epsilon x}\ e^{-ikx}dx \tag{30.242}$$

En décomposant les parties imaginaires et réelles, et avec un peu de changement de variables, nous pouvons utiliser les intégrales (18.99) et (18.100) pour obtenir

$$\hat{f}_\epsilon(k) = \frac{2\epsilon}{k^2+\epsilon^2}. \tag{30.243}$$

Sachant que $\arctan(x)$ est une primitive de $\frac{1}{x^2+1}$ et avec encore un peu de changement de variables, nous avons [30]

$$\int_{\mathbb{R}} \hat{f}_\epsilon(k)dk = \int_{-\infty}^{\infty} \frac{2\epsilon}{k^2+\epsilon^2} = 2[\arctan(x/\epsilon)]_{-\infty}^{\infty} = 2\pi. \tag{30.244}$$

Cela montre que si nous introduisons la fonction δ_ϵ donnée par

$$\delta_\epsilon(k) = \frac{1}{\pi}\frac{\epsilon}{\epsilon^2+k^2}, \tag{30.245}$$

alors nous avons une fonction qui tout en même temps ressemble à $\hat{f}_\epsilon$ et vérifie

$$\int_{\mathbb{R}} \delta_\epsilon(k)dk = 1 \tag{30.246}$$

pour tout ϵ.

Jusqu'ici nous avons montré que

$$\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon|x|}dx = 2\pi\delta_\epsilon(k). \tag{30.247}$$ EQooQOOQooBnfINi

Pour chaque $\epsilon > 0$ nous avons $\delta_\epsilon \in L^1(\mathbb{R})$. △

Proposition 30.72 ([1]).
Soit $g \in \mathscr{S}(\mathbb{R})$. Alors nous avons

$$\int_{\mathbb{R}}\int_{\mathbb{R}} g(x)e^{-ixy}dx\,dy = 2\pi g(0). \tag{30.248}$$ EQooBOJUooQGvMrk

Démonstration. Soit $u \in \mathscr{S}(\mathbb{R})$; nous multiplions l'équation (30.247) par $u(k)$ et nous intégrons par rapport à k :

$$\int_{\mathbb{R}} u(k)\left[\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon|x|}dx\right]dk = 2\pi\int_{\mathbb{R}} u(k)\delta_\epsilon(k)dk. \tag{30.249}$$ EQooTTQQooKxhxzl

Il s'agit de passer à la limite dans l'équation (30.249). Les intégrales à gauche peuvent être effectuées séparément parce qu'elles respectent le théorème de Fubini. En effet soit la fonction

$$f(k,x) = u(k)e^{-ikx}e^{-\epsilon|x|} \tag{30.250}$$

qui est dans $L^1(\mathbb{R}\times\mathbb{R})$ en vertu du critère du corolaire 14.303 et du fait que à la fois $k \mapsto |u(k)|$ et $x \mapsto e^{-\epsilon|x|}$ sont dans $L^1(\mathbb{R})$.

Nous pouvons donc grouper et dégrouper les intégrales et en particulier les inverser. Si nous effectuons d'abord l'intégrale sur k nous trouvons

$$\int_{\mathbb{R}} u(k)\left[\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon|x|}dx\right]dk = \int_{\mathbb{R}} e^{-\epsilon|x|}\int_{\mathbb{R}} u(k)e^{-ikx}dk\,dx = \int_{\mathbb{R}} e^{-\epsilon|x|}\hat{u}(x)dx. \tag{30.251}$$

La fonction $x \mapsto |e^{-\epsilon|x|}\hat{u}(x)|$ est majorée (uniformément en ϵ) par $x \mapsto \hat{u}(x)$ qui est intégrable parce que la transformée de Fourier d'une fonction de $\mathscr{S}$ est dans $\mathscr{S}$ par la proposition 29.20. Le

30. Et en écrivant correctement l'intégrale sur $\mathbb{R}$ comme une limite, etc.

théorème de la convergence dominée de Lebesgue 14.207 nous permet de permuter la limite $\epsilon \to 0$ avec l'intégrale et obtenir

$$\lim_{\epsilon\to 0}\int_{\mathbb{R}} u(k)\int_{\mathbb{R}} e^{-ikx}e^{-\epsilon|x|}dx\,dk = \int_{\mathbb{R}}\hat{u}(x)dx = \int_{\mathbb{R}}\int_{\mathbb{R}} u(k)e^{-ikx}dk\,dx. \tag{30.252}$$

Notons qu'en passant à la limite nous avons perdu le droit de permuter les intégrales.

Nous devons encore prouver que

$$\lim_{\epsilon\to 0}\int_{\mathbb{R}} u(k)\delta_\epsilon(k)dk = u(0). \tag{30.253}$$

Cela n'est rien d'autre que le lemme 30.70 appliqué à la suite de fonctions $j_n = \delta_{1/n}$. □

30.73.
Notons que les intégrales dans (30.248) ne peuvent pas être permutées parce que $\int_{\mathbb{R}} e^{-ixy}dy$ n'existe pas. Il faut avouer que, malgré tous les conseils du type « attention : permuter des intégrales doit être fait avec prudence », ce n'est pas tous les jours que nous trouvons des intégrales qui ne peuvent pas être permutées, autrement que dans des exemples fait exprès.

30.8.7 Peigne de Dirac

Proposition 30.74.
La formule

$$\Delta_a = \sum_{k\in\mathbb{Z}} \delta_{ka} \tag{30.254}$$

EqMEVmKyg

définit un élément de $\mathscr{D}'(\mathbb{R})$.

La forme linéaire Δ_a est le **peigne de Dirac** de pas a.

Démonstration. Nous utilisons le critère de continuité séquentielle en zéro du théorème 30.30. Soit une suite $\varphi_n \to 0$ dans $\mathscr{D}(\mathbb{R})$. Par le théorème 30.27 il existe un compact K de $\mathbb{R}$ pour lequel $\varphi_n \in \mathscr{D}(K)$ pour tout n et $\varphi_n \to 0$ dans $\mathscr{D}(K)$. La somme 30.254 est donc finie et nous pouvons la permuter avec une limite :

$$\lim_{n\to\infty}\Delta_a(\varphi_n) = \sum_{k\in\mathbb{Z}}\lim_{n\to\infty}\varphi_n(ka). \tag{30.255}$$

La limite $\varphi_n \to 0$ dans $\mathscr{D}(K)$ signifie que nous avons convergence uniforme de la fonction et de toutes ses dérivées vers 0. En particulier $\|\varphi_n\|_\infty \to 0$; disons que la somme (qui est finie) fasse s termes :

$$\sum_{k\in\mathbb{Z}}\varphi_n(ka) \leqslant s\|\varphi_n\|_\infty. \tag{30.256}$$

Le terme de droite tend vers zéro lorsque n tend vers l'infini. □

Donc Δ_a est bien une distribution au sens de la définition 30.29.

Lemme 30.75 ([678])**.**
Le peigne de Dirac vérifie la relation

$$\Delta_a = \frac{1}{a}\Delta_1 \circ D_a \tag{30.257}$$

où D_a est l'application $D_a\colon \mathscr{D}(\mathbb{R}) \to \mathscr{D}(\mathbb{R})$,

$$(D_a f)(x) = af(ax). \tag{30.258}$$

Démonstration. Pour $\varphi \in \mathscr{D}(\mathbb{R})$ nous avons

$$\Delta_a(\varphi) = \sum_{k\in\mathbb{Z}} \varphi(ka) = \frac{1}{a}\sum_{k\in\mathbb{Z}} (D_a\varphi)(k) = \frac{1}{a}\Delta_1(D_a\varphi). \tag{30.259}$$

□

Proposition 30.76.
Le peigne de Dirac est une distribution tempérée.

Notez qu'il y a plus de fonctions dans $\mathscr{S}(\mathbb{R})$ que dans $\mathscr{D}(\mathbb{R})$; il est donc plus difficile de rentrer dans $\mathscr{S}'(\mathbb{R})$ que dans $\mathscr{D}'(\mathbb{R})$: il est plus compliqué d'avoir existence de $T(\varphi)$ pour tout $\varphi \in \mathscr{S}(\mathbb{R})$ que pour tout $\varphi \in \mathscr{D}(\mathbb{R})$.

Démonstration. Soit $\varphi \in \mathscr{S}(\mathbb{R})$. Nous avons

$$|\Delta_a(\varphi)| = |\sum_k \varphi(ak)| = \left|\sum_k \frac{(1+a^2k^2)\varphi(ak)}{1+a^2k^2}\right| \leqslant \sup_{x\in\mathbb{R}} \left|(1+x^2)\varphi(x)\right| \sum_k \frac{1}{1+a^2k^2}. \tag{30.260}$$

La somme $\sum_k \frac{1}{1+a^2k^2}$ est une somme convergente, et le supremum est borné par la proposition 27.236 en prenant $Q(x) = 1 + x^2$. En effet sur $\overline{B(0,r)}$ la fonction $x \mapsto (1+x^2)\varphi(x)$ est bornée par ce que c'est une fonction continue sur un compact, et à l'extérieur de $B(0,r)$ cette fonction est alors bornée par 1. □

Si aucune ambigüité n'est à craindre, nous noterons f la distribution T_f.

Exemple 30.77.
La transformée de Fourier de la distribution de Dirac est la fonction constante : $\hat{\delta} = 1$. En effet si nous agissons sur une fonction test,

$$\hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int_{\mathbb{R}^d} \varphi(x)dx. \tag{30.261}$$

△

30.9 L'espace $C^\infty(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d))$

Dans cette section nous notons I un ouvert de $\mathbb{R}$ et Ω un ouvert de $\mathbb{R}^d$; si ψ est une fonction sur $I \times \Omega$ nous allons noter $\psi_t \colon \Omega \to \mathbb{R}$ la fonction $\psi_t(x) = \psi(t,x)$. C'est une notation plus légère que $\psi(t,.)$.

30.9.1 Propriétés générales

La définition de l'espace $C^\infty(I, \mathscr{S}'(\Omega))$ est encore la définition 7.363 et les propriétés énoncées dans la proposition 30.47 sont encore bonnes ici.

D'abord parlons un peu de continuité en recopiant la proposition 7.362 dans notre contexte.

PropBXFmvPs

Proposition 30.78.
Soient I un intervalle ouvert de $\mathbb{R}$ et $u \colon I \to \mathscr{S}'(\mathbb{R}^d)$ une fonction continue. Alors

ItemFTvVUEW

(1) Pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$, l'application $t \mapsto u_t(\varphi)$ est continue.

(2) Pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$, nous avons la limite

$$\lim_{t\to t_0} u_t(\varphi) = u_{t_0}(\varphi). \tag{30.262}$$

(3) Nous avons la limite dans $\mathscr{S}'(\mathbb{R}^d)$

$$\lim_{t\to t_0} u_t = u_{t_0}. \tag{30.263}$$

Démonstration. Ceci est un cas particulier de la proposition 7.362. □

LEMooFTRNooNqFplH

Lemme 30.79.
Nous avons $C^\infty(I, \mathscr{S}'(\Omega)) \subset C^\infty(I, \mathscr{D}'(\Omega))$.

Démonstration. Soit $(T_t) \in C^\infty(I, \mathscr{S}'(\Omega))$. Pour chaque t nous avons

$$T_t \in \mathscr{S}'(\Omega) \subset \mathscr{D}'(\Omega). \tag{30.264}$$

Ensuite il suffit de dire que pour tout $\varphi \in \mathscr{D}(\Omega)$ la fonction

$$t \mapsto T_t(\varphi) \tag{30.265}$$

est de classe C^∞ parce que c'est le cas pour toute fonction dans $\mathscr{S}(\Omega)$. La proposition 30.47 (en changeant $\mathscr{D}$ en $\mathscr{S}$) conclut que $(T_t) \in C^\infty(I, \mathscr{D}'(\Omega))$. □

PropIIAcyDq

Proposition 30.80.
L'espace $\mathscr{S}(\Omega)$ est complet et métrisable.

Démonstration. En ce qui concerne le métrisable nous reprenons la formule de l'écart (7.451). Dans notre cas pour l'écrire explicitement il faudrait une énumération de $\mathbb{N}^2$ à partir de 1 (et non de zéro). Cette formule donne bien une distance parce que si $d(\varphi_1 - \varphi_2) = 0$ alors en particulier $p_{00}(\varphi_1 - \varphi_2) = \|\varphi_1 - \varphi_2\|_\infty = 0$ et donc $\varphi_1 = \varphi_2$.

Nous montrons maintenant que $\mathscr{S}(\Omega)$ est complet en y considérant une suite de Cauchy (φ_n). Soit $\epsilon > 0$ et $\alpha, \beta \in \mathbb{N}$ ainsi que k, l assez grands pour que $\varphi_k - \varphi_l \in B_{\alpha\beta}(0, \epsilon)$. En particulier pour $\alpha = \beta = 0$ nous avons $\|\varphi_k - \varphi_l\|_\infty \leqslant \epsilon$, ce qui signifie que nous avons une suite vérifiant le critère de Cauchy uniforme 12.356. Elle converge donc uniformément vers une certaine fonction φ que la proposition 12.357 nous assure être continue. Il existe donc $\varphi \in C(\Omega)$ telle que

$$\varphi_k \xrightarrow{unif} \varphi. \tag{30.266}$$

Nous devons montrer que $\varphi \in \mathscr{S}(\Omega)$. Le fait que φ soit de classe C^∞ s'obtient en utilisant les seminormes $p_{0,\alpha}(\varphi) = \|\partial^\alpha \varphi\|_\infty$ de la même façon que dans la preuve que $\mathscr{D}(\Omega)$ était complet (proposition 30.28). Nous obtenons en particulier que

$$\partial^\alpha \varphi_k \xrightarrow{unif} \partial^\alpha \varphi \tag{30.267}$$

EqSZyYkqk

pour tout multiindice α. Montrons encore que φ est à décroissance rapide : nous devons montrer que pour tout α et β nous avons

$$p_{\alpha\beta}(\varphi) = \sup_{x \in \Omega} \left| x^\beta (\partial^\alpha \varphi)(x) \right| < \infty. \tag{30.268}$$

Étant donné que (φ_n) est de Cauchy dans $\mathscr{S}(\Omega)$ nous avons (pour ϵ fixé et k, l assez grands) :

$$\left| x^\beta (\partial^\alpha \varphi_k - \partial^\alpha \varphi_l)(x) \right| \leqslant \epsilon \tag{30.269}$$

pour tout $x \in \Omega$. En considérant l fixé et en prenant la limite $k \to \infty$ et en utilisant la convergence uniforme (30.267) nous trouvons que

$$\left| x^\beta (\partial^\alpha \varphi - \partial^\alpha \varphi_l)(x) \right| \leqslant \epsilon \tag{30.270}$$

Du coup nous pouvons faire la majoration

$$\sup_{x \in \Omega} \left| x^\beta (\partial^\alpha \varphi)(x) \right| \leqslant \sup_x \left| x^\beta (\partial^\alpha \varphi - \partial^\alpha \varphi_l)(x) \right| + \sup_x \left| (\partial^\alpha \varphi_l)(x) \right| \leqslant \epsilon + p_{\alpha\beta}(\varphi_l) < \infty \tag{30.271}$$

du fait que $p_{\alpha\beta}(\varphi_l) < \infty$ parce que $\varphi_l \in \mathscr{S}(\Omega)$.

Donc $\varphi \in \mathscr{S}(\Omega)$ et ce dernier est alors complet. □

Proposition 30.81.
Soit $(T_t) \in C^0\big(I, \mathscr{S}'(\Omega)\big)$ et $\psi \in \mathscr{S}(I \times \Omega)$. Alors la fonction

$$t \mapsto T_t(\psi_t) \tag{30.272}$$

EqULcaYjm

est continue sur I.

Démonstration. Soient $t_0 \in I$ et une suite convergente vers t_0 : $t_j \to t_0$ dans $\mathbb{R}$. Vu que (T_t) est continue en t, elle est en particulier séquentiellement continue et nous avons

$$T_{t_j} \xrightarrow{\mathscr{S}'(\Omega)} T_{t_0}. \tag{30.273}$$

Montrons que nous avons aussi $\psi_{t_j} \xrightarrow{\mathscr{S}(\Omega)} \psi_{t_0}$. Pour cela nous utilisons les seminormes [31] $p_{\alpha\beta}$ définies en (27.990) :

$$\begin{aligned}
p_{\alpha\beta}(\psi_{t_j} - \psi_{t_0}) &= \sum_{x\in\Omega} \Big| x^\beta \big(\partial^\alpha \psi(t_j, x) - \partial^\alpha \psi(t_0, x)\big)\Big| && \text{(30.274a)}\\
&\leqslant \sup_{x\in\Omega} \Big| x^\beta |t_0 - t_j| \sup_{t\in[t_0,t_j]} \big|\partial_t \partial^\alpha \psi(t, x)\big|\Big| && \text{(30.274b)}\\
&\leqslant |t_0 - t_j| \sup_{x\in\Omega} \sup_{t\in I} \big| x^\beta \partial_t \partial^\alpha \psi(t, x)\big| && \text{(30.274c)}\\
&\leqslant |t_0 - t_j| P_{(\alpha t),\beta}(\psi). && \text{(30.274d)}
\end{aligned}$$

Pour la première majoration nous avons utilisé le théorème des accroissements finie 12.314. Pour la dernière ligne nous avons noté $P_{\alpha\beta}$ les seminormes de $\mathscr{S}(I \times \Omega)$ et (αt) est le multiiindice α à qui on a ajouté la variable t à la fin. Étant donné que $P_{(\alpha t)\beta}(\psi) < \infty$ nous avons bien

$$p_{\alpha\beta}(\psi_{t_j} - \psi_{t_0}) \to 0 \tag{30.275}$$

et donc $\psi_{t_j} \xrightarrow{\mathscr{S}(\Omega)} \psi_{t_0}$.

Étant donné que $\mathscr{S}(\Omega)$ est métrisable et complet, le corolaire 27.6 nous dit que

$$T_{t_j}(\psi_{t_j}) \xrightarrow{\mathbb{C}} T_{t_0}(\psi_{t_0}), \tag{30.276}$$

ce qui est bien le critère de continuité séquentielle de la fonction (30.272). □

RemZYVkHRT

Remarque 30.82.
La proposition 30.49 nous dit, a fortiori, que si $(T_t) \in C^\infty\big(I, \mathscr{S}'(\Omega)\big)$ alors la formule

$$\tilde{T}(\psi) = \int_I T_t(\psi_t) \tag{30.277}$$

donne un élément $\tilde{T} \in \mathscr{D}'(I \times \Omega)$. Au cas où aucune confusion n'est à craindre, nous pourrons noter également T l'élément de $\mathscr{D}'(I \times \Omega)$ déduit de $T \in C^\infty\big(I, \mathscr{S}'(\Omega)\big)$.

Notons que ce T ne sera pas toujours une distribution tempérée comme le montre l'exemple suivant.

Exemple 30.83.
En posant $T_t(\varphi) = e^{t^2}\varphi(0)$ avec $I = \mathbb{R}$, l'intégrale

$$T(\psi) = \int_{\mathbb{R}} T_t(\psi_t) = \int_{\mathbb{R}} e^{t^2} \psi(t, 0) dt \tag{30.278}$$

ne converge pas pour tout $\psi \in \mathscr{S}(\mathbb{R} \times \Omega)$. En effet par rapport à t, la fonction $\psi(t, 0)$ décroît rapidement mais pas spécialement assez rapidement pour compenser e^{t^2}. △

31. Pas parce que nous en avons envie, mais bien parce qu'elles font partie de la définition de la convergence et de tous ces trucs.

30.9.2 Dérivation

PropGKoPbko

Proposition 30.84 ([718]).
Soit $T \in C^k\big(I, \mathscr{S}'(\Omega)\big)$ et $0 \leqslant l \leqslant k$. Pour tout $t_0 \in I$ l'application

$$\begin{aligned} T_{t_0}^{(l)} \colon \mathscr{S}(\Omega) &\to \mathbb{C} \\ \varphi &\mapsto \left(\frac{d^l}{dt} T_t(\varphi)\right)(t_0) \end{aligned} \tag{30.279}$$

EqZMDeZco

est bien définie, est une distribution et de plus

$$t \mapsto T_t^{(l)} \in C^{k-l}\big(I, \mathscr{S}'(\Omega)\big). \tag{30.280}$$

Attention que la formule (30.279) est bonne si $\varphi \in \mathscr{S}(\Omega)$. Si par contre $\psi \in \mathscr{S}(I \times \Omega)$ et qu'on veut regarder $u_t^{(1)}(\psi_t)$ alors il faut regarder la proposition 30.85 et utiliser la formule (30.285) dans laquelle se trouve $u_t^{(1)}(\psi_t)$.

Démonstration. Pour $k = 0$ nous avons $T_t^{(0)} = T_t$ et c'est bon. Pour le cas $k = 1$ et $l = 0$ c'est encore $T_t^{(0)} = T_t$ qui fonctionne.

Le premier cas non trivial à traiter est $k = 1$ et $l = 1$. Nous considérons $t_0 \in I$; par définition de la dérivée, pour tout $\varphi \in \mathscr{S}(\Omega)$, nous avons (pour peu que les limites existent) :

$$T_{t_0}^{(1)}(\varphi) = \frac{d}{dt}\Big[T_t(\varphi)\Big]_{t=t_0} = \lim_{j\to\infty} \frac{T_{t_0+\epsilon_j}(\varphi) - T_{t_0}(\varphi)}{\epsilon_j} = \lim_{j\to\infty} U_j(\varphi) \tag{30.281}$$

EqCTuAfXe

où

$$U_j = \frac{1}{\epsilon_j}\big(T_{t_0+\epsilon_j} - T_{t_0}\big). \tag{30.282}$$

et (ϵ_j) est une suite de réels tendant vers zéro.

Vu que $(T_t) \in C^k(I, \mathscr{S}'(\Omega))$, l'application $t \mapsto T_t(\varphi)$ est de classe C^k et en particulier l'expression (30.281) a une limite lorsque $j \to \infty$. Donc $T_{t_0}^{(1)}(\varphi)$ est bien définie. Le point (1) du corolaire 27.6 nous dit que $\lim_{j\to\infty} U_j(\varphi) = T_{t_0}^{(1)}(\varphi)$ et $T_{t_0}^{(1)}$ est une distribution (linéaire et continue).

Nous devons encore voir que $t \mapsto T_t^{(1)}$ est une application $C^0\big(I, \mathscr{S}'(\Omega)\big)$. Cela est une conséquence du fait que (T_t) soit de classe C^1, ce qui se traduit par le fait que l'application

$$t \mapsto \frac{d}{dt}\Big(T_t(\varphi)\Big) \tag{30.283}$$

est continue (définition de la dérivée et point (1) de la proposition 30.78 appliquée à la dérivée).

Les cas $k \geqslant 1$ se traitent par récurrence. □

PropUDkgksG

Proposition 30.85 ([718]).
Soit $(T_t) \in C^1\big(I, \mathscr{S}'(\Omega)\big)$ et $\psi \in \mathscr{S}(I \times \Omega)$. Alors la fonction

$$t \mapsto T_t\big(\psi(t, .)\big) \tag{30.284}$$

est de classe C^1 sur I et

$$\frac{d}{dt}\Big(T_t\big(\psi(t,.)\big)\Big) = T_t^{(1)}\big(\psi(t,.)\big) + T_t\left(\frac{\partial\psi}{\partial t}(t,.)\right) \tag{30.285}$$

EqSCNYYbE

Démonstration. Soient $t_0 \in I$ et $\epsilon_j \to 0$ une suite réelle. Le membre de gauche de (30.285), écrit en t_0, donne

$$\spadesuit = \lim_{j\to\infty} \frac{T_{t_0+\epsilon_j}\big(\psi(t_0+\epsilon_j, .)\big) - T_{t_0}\big(\psi(t_0, .)\big)}{\epsilon_j} \tag{30.286}$$

BJPHgwn

Afin d'alléger les notations nous allons écrire $\psi_t = \psi(t,.)$. Dans le numérateur de (30.286) nous ajoutons et soustrayons la quantité $T_{t_0+\epsilon_j}(\psi_{t_0})$ et nous découpons la limite en deux morceaux :

$$\spadesuit = \lim_{j\to\infty} \frac{T_{t_0+\epsilon_j}(\psi_{t_0+\epsilon_j} - \psi_{t_0})}{\epsilon_j} + \lim_{j\to\infty} \frac{(T_{t_0+\epsilon_j} - T_{t_0})(\psi_{t_0})}{\epsilon_j} \tag{30.287}$$

Le second terme vaut

$$\frac{d}{dt}\Big(T_t(\psi_{t_0})\Big)_{t=t_0} = T_{t_0}^{(1)}(\psi_{t_0}) \tag{30.288}$$

par la proposition 30.84. Occupons nous de l'autre morceau de $\spadesuit$. Nous posons $U_j = T_{t_0+\epsilon_j}$ et

$$\varphi_j = \frac{1}{\epsilon_j}(\psi_{t_0+\epsilon_j} - \psi_{t_0}). \tag{30.289}$$

Nous voulons utiliser le corolaire 27.6(4) pour obtenir

$$\lim_{j\to\infty} U_j(\varphi_j) = T_{t_0}\Big(\frac{\partial\psi}{\partial t}(t_0,.)\Big). \tag{30.290}$$

D'une part (T_t) est de classe C^∞ en t et nous avons donc la convergence $U_j \xrightarrow{\mathscr{S}'(\Omega)} T_{t_0}$. Reste à prouver que

$$\varphi_j \xrightarrow{\mathscr{S}(\Omega)} \frac{\partial\psi}{\partial t}(t_0,.). \tag{30.291}$$

Cela en remarquant bien que la variable de dérivation n'est pas celle par rapport à laquelle nous voulons la convergence Schwartz [32]. Soient α et β des naturels et calculons un peu :

$$p_{\alpha\beta}\big(\varphi_j - \frac{\partial\psi}{\partial t}(t_0,.)\big) = \sup_{x\in\Omega}\Big|x^\beta\partial^\alpha\Big(\frac{1}{\epsilon_j}\big(\psi(t_0+\epsilon_j,x) - \psi(t_0,x)\big) - \frac{\partial\psi}{\partial t}(t_0,x)\Big)\Big| \tag{30.292}$$

EqEBUYDRA

Il est à présent l'heure d'utiliser un développement de Taylor avec le reste de la proposition 12.472 :

$$\psi(t_0+\epsilon_j,x) = \psi(t_0,x) + \epsilon_j\frac{\partial\psi}{\partial t}(t_0,x) + \frac{\epsilon_j^2}{2}\frac{\partial^2\psi}{\partial t^2}(\bar{t},x) \tag{30.293}$$

pour un certain $\bar{t}\in[t_0,t_0+\epsilon_j]$. En mettant ça dans le calcul (30.292) nous restons avec

$$p_{\alpha\beta}\big(\varphi_j - \frac{\partial\psi}{\partial t}(t_0,.)\big) = \sup_{x\in\Omega}\Big|x^\beta\partial^\alpha\Big(\epsilon_j\frac{\partial^2\psi}{\partial t^2}(\bar{t},x)\Big)\Big| \leqslant \epsilon_j P_{\alpha,2;\beta,0}(\psi) \tag{30.294}$$

où $P_{\alpha,k;\beta,l}$ sont les seminormes de $\mathscr{S}(I\times\Omega)$ avec la notation plus ou moins évidente de prendre α dérivations sur x, k sur t puis de multiplier par $x^\beta t^l$. Au final nous avons bien

$$\lim_{j\to\infty} p_{\alpha\beta}\big(\varphi_j - \frac{\partial\psi}{\partial t}(t_0,.)\big) = 0 \tag{30.295}$$

et donc la convergence $\varphi_j \xrightarrow{\mathscr{S}(\Omega)} \frac{\partial\psi}{\partial t}(t_0,.)$. □

LemWRoRPIX

Lemme 30.86.
Soit $(T_t)\in C^1\big(I,\mathscr{S}'(\Omega)\big)$ alors si $\mathcal{F}$ dénote la transformée de Fourier nous avons

$$\mathcal{F}\big(T_t^{(1)}\big) = (\mathcal{F}T)_t^{(1)} \tag{30.296}$$

où $(\mathcal{F}T)$ est la famille de distributions $(\mathcal{F}T)_t = \mathcal{F}T_t$.

Démonstration. Pour la preuve il suffit de tester l'égalité sur une fonction $\varphi\in\mathscr{S}(\Omega)$:

$$(\mathcal{F}T_t^{(1)})(\varphi) = T_t^{(1)}(\mathcal{F}\varphi) = \frac{d}{dt}\Big(T_t(\mathcal{F}\varphi)\Big) = \frac{d}{dt}\Big((\mathcal{F}T_t)(\varphi)\Big) = (\mathcal{F}T)_t^{(1)}(\varphi). \tag{30.297}$$

□

32. Je ne sais pas si je me suis bien fait comprendre là.

Chapitre 31

Espaces de Sobolev, équations elliptiques

CHAPooVTIIooGOEvXT

31.1 Espaces de Sobolev

Rappel : la définition de la dérivée faible est 30.2.

31.1.1 Sur un intervalle de $\mathbb{R}$

Sauf mention du contraire dans cette section I est un intervalle borné ouvert $I =]a, b[$ de $\mathbb{R}$.

Définition 31.1.
*Soit $I =]a, b[$ un ouvert borné de $\mathbb{R}$. L'**espace de Sobolev** $H^1(I)$ est l'ensemble*

$$H^1(I) = \Big\{ u \in L^2(I) \text{ tel que } \exists g \in L^2(I) \text{ tel que } \forall \varphi \in C_c^\infty(I), \int_I u\varphi' = -\int_I g\varphi \Big\}. \tag{31.1}$$

L'unique élément g de $L^2(I)$ vérifiant $\int_I u\varphi' = -\int_I g\varphi$ est noté u' et est nommé **dérivée** ; nous verrons dans les prochaines pages pourquoi.

L'espace H^1 accepte le produit scalaire suivant :

$$\langle u, v \rangle = \int_I uv + \int_I u'v', \tag{31.2}$$

et nous notons $\|.\|_{H^1}$ la norme correspondante qui n'est autre que

$$\|u\|_{H^1}^2 = \langle u, u \rangle = \|u\|_{L^2}^2 + \|u'\|_{L^2}^2. \tag{31.3}$$

Nous introduisons l'espace $L^1_{loc}(I)$ des fonctions étant L^1 sur tout compact de I.

CorEVJYihj

Corolaire 31.2.
Si $u \in H^1(I)$ et si $u' = 0$ alors il existe une constante C telle que $u = C$ presque partout.

Démonstration. L'hypothèse $u' = 0$ signifie que pour toute fonction $\varphi \in C_c^\infty(I)$,

$$\int_I u\varphi' = \int_I u'\varphi = 0. \tag{31.4}$$

La proposition 27.190 nous dit alors qu'il existe une constante C telle que $u = C$ presque partout. □

LemMPkbZxX

Lemme 31.3.
Tout élément de $H^1(I)$ admet un unique représentant continu.

Nous verrons dans le corolaire 31.5 que ce représentant pourra être prolongé par continuité sur $\bar{I}$.

Démonstration. Soient $y_0 \in I$ et $u \in H^1(I)$. Nous considérons la fonction

$$\bar{u}(x) = \int_{y_0}^{x} u'(t)dt. \tag{31.5}$$

Notons que par définition, $u' \in L^2$ donc l'intégrale ne pose pas de problèmes. Montrons que $\bar{u}$ est continue sur $\bar{I}$. Pour cela nous considérons $x \in \bar{I}$ et h tel que $x + h \in \bar{I}$. Alors

$$\big|\bar{u}(x+h) - \bar{u}(x)\big| = \Big|\int_{x}^{x+h} u'\Big| \leqslant \int_{x}^{x+h} |u'|. \tag{31.6}$$

Mais la fonction $|u'|$ est dans $L^1_{loc}(I)$ par le lemme 27.36 ; elle est en particulier intégrable sur un ouvert contenant x et par conséquent la dernière intégrale tend vers zéro lorsque h tend vers 0.

Nous prouvons à présent que $\bar{u}$ est dans $H^1(I)$ et que sa dérivée est égale à u' ; pour cela nous allons montrer que pour tout $\varphi \in C^\infty_c(I)$,

$$\int_I \bar{u}\varphi' = -\int_I u'\varphi. \tag{31.7}$$

Nous avons

$$\int_I \bar{u}\varphi' = \int_I \left(\int_{y_0}^{x} u'(t)dt\right)\varphi'(x)dx = \int_a^{y_0} \left(\int_{y_0}^{x} u'(t)dt\right)\varphi'(x)dx + \int_{y_0}^{b} \left(\int_{y_0}^{x} u'(t)dt\right)\varphi'(x)dx. \tag{31.8}$$

Pour faire plus court, nous notons $f(t,x) = u'(t)\varphi'(x)$. La première intégrale vaut

$$\begin{aligned}
\int_a^{y_0} \left(\int_{y_0}^{x} u'(t)\varphi'(x)\right) &= \int_a^{y_0} \left(\int_{y_0}^{a} f(t,x)\mathbb{1}_{t<x}(t,x)dt\right)dx && \text{(31.9a)}\\
&= \int_{y_0}^{a}\int_{a}^{y_0} f(t,x)\mathbb{1}_{t<x}dxdt && \text{(31.9b)}\\
&= \int_{y_0}^{a}\int_{a}^{t} f(t,x)dxdt && \text{(31.9c)}\\
&= -\int_a^{y_0}\int_a^{t} u'(t)\varphi'(x)dx\,dt && \text{(31.9d)}
\end{aligned}$$

La permutation d'intégrales pour obtenir (31.9b) est due au théorème de Fubini 14.304(3). Par le même petit jeu, la seconde intégrale vaut

$$\int_{y_0}^{b}\int_{t}^{b} u'(t)\varphi'(x)dx\,dt. \tag{31.10}$$

En refaisant la somme,

$$\begin{aligned}
\int_I \bar{u}\varphi' &= -\int_a^{y_0} u'(t)\left(\int_a^t \varphi'(x)dx\right)dt + \int_{y_0}^{b} u'(t)\left(\int_t^b \varphi'(x)dx\right)dt && \text{(31.11a)}\\
&= -\int_a^{y_0} u'(t)\big(\varphi(t) - \varphi(a)\big)dt + \int_{y_0}^{b} u'(t)\big(\varphi(b) - \varphi(t)\big) && \text{(31.11b)}\\
&= -\int_a^b u'\varphi && \text{(31.11c)}\\
&= -\int_I u'\varphi. && \text{(31.11d)}
\end{aligned}$$

Notons que $\varphi(a) = \varphi(b) = 0$ parce que φ est à support compact dans $]a,b[$. Nous avons donc prouvé que $\bar{u}$ est dans $H^1(I)$ et que $\bar{u}' = u'$. Par le corolaire 31.2, nous avons une constante C telle que $\bar{u} = u + C$ presque partout, c'est-à-dire $u = \bar{u} + C$ dans $H^1(I)$.

En résumé, $\tilde{u} = \bar{u} + C$ est un représentant continu de u dans $L^2(I)$.

L'unicité du représentant continu est simplement le fait que deux fonctions continues égales presque partout sont égales (proposition 20.154).

□

PropGWOIoDg

Proposition 31.4.
Si $u \in H^1(I)$, alors

$$u(x) - u(y) = \int_y^x u' \tag{31.12}$$

pour tout $x, y \in I$.

Démonstration. Pour fixer les idées, nous supposons $x < y$. Nous considérons une suite $\varphi_n \in C_c^\infty(I)$ convergeant uniformément sur I vers $\mathbb{1}_{[x,y]}$. Nous exigeons de plus que

— φ_n' est positive sur $[a, x + \frac{1}{n}]$
— φ_n' est négative sur $[y - \frac{1}{n}, b]$
— $\varphi_n = 1$ sur $[x + \frac{1}{n}, y - \frac{1}{n}]$.
— $\varphi_n = 0$ sur $[a, x - 1/n]$ et sur $[y + 1/n, b]$.

Pour chaque n, nous découpons l'intégrale comme

$$-\int_I u'\varphi_n = \int_I u\varphi_n' = \int_a^{x-1/n} u\varphi_n' + \int_{x-1/n}^{x+1/n} u\varphi_n' + \int_{x+1/n}^{y-1/n} u\varphi_n' + \int_{y-1/n}^{y+1/n} u\varphi_n' + \int_{y+1/n}^{b} u\varphi_n'. \tag{31.13}$$

Par construction de φ_n, de ces 5 morceaux, il n'en reste que deux de non nulles :

$$\int_I u\varphi_n' = \underbrace{\int_{x-1/n}^{x+1/n} u(t)\varphi_n'(t)dt}_{A} + \underbrace{\int_{y-1/n}^{y+1/n} u(t)\varphi_n'(t)dt}_{B} \tag{31.14}$$

Soit $\epsilon > 0$ et n suffisamment grand pour avoir $u(t) \in B\big(u(x), \epsilon\big)$ pour tout $t \in B(x, \frac{1}{n})$ et (en même temps) $u(t) \in B\big(u(y), \epsilon\big)$ pour tout $t \in B(y, \frac{1}{n})$. C'est la continuité de u qui permet de trouver un tel n. Pour cette valeur de n, en tenant compte des hypothèses sur la positivité de φ_n' nous avons

$$\int_{x-1/n}^{x+1/n} \big(u(x) - \epsilon\big)\varphi_n'(t)dt \leqslant \int_{x-1/n}^{x+1/n} u(t)\varphi_n'(t)dt \leqslant \int_{x-1/n}^{x+1/n} \big(u(x) + \epsilon\big)\varphi_n'(t)dt, \tag{31.15}$$

mais par hypothèse sur φ_n nous trouvons

$$\int_{x-1/n}^{x+1/n} \varphi_n'(t)dt = \varphi_n(x + \frac{1}{n}) - \varphi(x + \frac{1}{n}) = 1. \tag{31.16}$$

donc

$$u(x) - \epsilon \leqslant \int_{x-1/n}^{x+1/n} u(t)\varphi_n'(t)dt \leqslant u(x) + \epsilon. \tag{31.17}$$

EqIYrpFdb

Pour encadrer la seconde, il faut être plus prudent avec les signes parce que φ_n' y est négative. En posant $\psi_n = -\varphi_n$ nous avons

$$-B = \int_{y-1/n}^{y+1/n} u(t)\psi_n(t)dt, \tag{31.18}$$

et donc

$$u(y) - \epsilon \leqslant -B \leqslant u(y) + \epsilon \tag{31.19}$$

ou encore

$$-\epsilon - u(y) \leqslant B \leqslant \epsilon - u(y). \tag{31.20}$$

En additionnant avec (31.17) nous voyons que pour tout $\epsilon > 0$ il existe un $N(\epsilon)$ tel que nous ayons

$$u(x) - u(y) - 2\epsilon \leqslant \int_I u'\varphi_n \leqslant u(x) - u(y) + 2\epsilon \tag{31.21}$$

EqEBwWUxm

pour tout $n \geqslant N$. Nous voulons évidemment prendre la limite $\epsilon \to 0$, c'est-à-dire $n \to \infty$. Étant donné que $\varphi_n(t) < 1$ pour tout t et pour tout n, la fonction $t \mapsto u'(t)\varphi_n(t)$ est dominée par u', qui est dans $L^1(I)$ par le lemme 27.36. Le théorème de la convergence dominée nous permet donc d'affirmer que

$$\lim_{n\to\infty} \int_I u'\varphi_n = \int_I u' \mathbb{1}_{[x,y]} = \int_x^y u', \tag{31.22}$$

et donc les inégalités (31.21) donnent le résultat, grâce au signe dans (31.13). □

CorCEPJGAu

Corolaire 31.5.
Si $[u] \in H^1(I)$, le représentant continu $u \in C^0(I)$ peut être prolongé par continuité en $u \in C^0(\bar{I})$.

Démonstration. Soit (x_n) une suite strictement croissante dans $]a, b[$ convergeant vers b. Nous voulons montrer que la suite $\big(u(x_n)\big)$ est de Cauchy dans $\mathbb{R}$, ce qui nous permettra de définir

$$u(b) = \lim_{n\to\infty} u(x_n). \tag{31.23}$$

qui sera évidemment continue. Cette construction ne dépendra pas du choix de la suite (x_n) parce que deux fonctions continues sur $\bar{I}$ et égales sur I sont égales sur $\bar{I}$.

En notant u' la dérivée de u dans H^1, nous avons par construction du représentant continu : $u(x) = \int_{y_0}^x u'(t)dt$. Et donc

$$\big|u(x_n) - u(x_{n+p})\big| = \left|\int_{y_0}^{x_n} u' - \int_{y_0}^{x_{n+p}} u'\right| = \left|\int_{x_n}^{x_{n+p}} u'\right|. \tag{31.24}$$

Vu que la suite (x_n) est de Cauchy et que u' est intégrable (même sur $\bar{I}$), la limite $n \to \infty$ de cela est zéro, quelle que soit la valeur de p. Donc $\big(u(x_n)\big)$ est de Cauchy dans $\mathbb{R}$ et est donc convergente. □

ThoESIyxfU

Proposition 31.6 ([104, 1])**.**
Quelques propriétés de l'espace de Sobolev $H^1(I)$ où $I =]a, b[$ est un ouvert borné de $\mathbb{R}$.
ITEMooITQUooKWwMwu

(1) $H^1(I)$ est un espace de Hilbert.

(2) $H^1(I)$ s'injecte de façon compacte dans $C^0(\bar{I})$.

(3) $H^1(I)$ s'injecte de façon continue dans $L^2(I)$.

Démonstration. Nous prouvons point par point.

(1) Le seul critère à vérifier est la complétude. Pour cela nous considérons une suite de Cauchy (u_n) dans $H^1(I)$. Si $\epsilon > 0$, alors il existe $N > 0$ tel que pour tout $p \geqslant 0$ nous ayons $\|u_{n+p} - u_n\|_{H^1}^2 \leqslant \epsilon$, c'est-à-dire

$$\|u_{n+p} - u_n\|_{L^2}^2 + \|u'_{n+p} - u'_n\|_{L^2}^2 \leqslant \epsilon. \tag{31.25}$$

En particulier les suites (u_n) et (u'_n) sont de Cauchy dans L^2 qui est complet par le théorème de Fischer-Riesz 27.45. Nous notons donc

$$u_n \xrightarrow{L^2} u \tag{31.26a}$$

$$u'_n \xrightarrow{L^2} v. \tag{31.26b}$$

Nous allons maintenant montrer quelques limites.

(1a) $\underline{u_n\varphi \xrightarrow{L^2} u\varphi}$ Si M est une constante qui majore φ alors $\|u_n\varphi - u\varphi\|_2 \leqslant M\|u_n - u\|_2 \to 0$.

(1b) $\underline{u'_n\varphi \xrightarrow{L^2} v\varphi}$ C'est la même chose avec $\|u'_n\varphi - v\varphi\|_2 \leqslant M\|u'_n - v\|_2 \to 0$.

(1c) $\underline{u \in H^1(I) \text{ avec } u' = v}$ Attendu le corolaire 27.37 qui permet de permuter intégrale et limite dans $L^2(I)$ et les limites que nous venons de prouver,

$$\int_I u\varphi' = \lim_{n\to\infty} \int_I u_n\varphi' = -\lim_{n\to\infty} \int_I u_n'\varphi = -\int_I v\varphi. \tag{31.27}$$

Cela signifie que v est la dérivée faible de u : $u' = v$.

(1d) $\underline{u_n \xrightarrow{H^1} u}$
Nous pouvons alors prouver que $u_n \to u$ dans $H^1(I)$:

$$\|u_n - u\|_{H^1(I)}^2 = \|u_n - u\|_{L^2}^2 + \|u_n' - u'\|_{L^2}^2. \tag{31.28}$$

Mais nous savons déjà que $u_n \to u$ dans L^2 (d'ailleurs c'est la définition de u) et que $u' = v$ alors que par définition de v, nous avons $u_n' \to v$ dans L^2.

Tout cela donne que $u_n \to u$ dans $H^1(I)$ et donc que $H^1(I)$ est un espace complet.

(2) L'application que nous allons prouver être compacte entre $H^1(I)$ et $C^0(\bar{I})$ est

$$\begin{aligned} \psi\colon H^1(I) &\to C^0(\bar{I}) \\ [u] &\mapsto \tilde{u} \end{aligned} \tag{31.29}$$

où $[u]$ désigne une classe de fonction dans $H^1(I)$ et $\tilde{u}$ est son représentant continu prolongé par continuité à $\bar{I}$ [1], qui existe par le lemme 31.3 et le corolaire 31.5. Cette application est une injection par l'unicité du représentant continu. Nous allons prouver que c'est une application compacte en utilisant le critère (2) de la proposition 27.3. Pour cela nous allons commencer par utiliser le théorème d'Ascoli sur l'ensemble $\tilde{\mathcal{B}}$ des représentants continus des éléments de $\mathcal{B}$, prolongés par continuité sur $\bar{I}$; c'est-à-dire $\tilde{B} \subset C^0(\bar{I})$.
Soit $u \in \tilde{\mathcal{B}}$; par la proposition 31.4, nous avons

$$\begin{aligned} \big|u(x) - u(y)\big| &= \Big|\int_y^x u'(t)dt\Big| && (31.30a)\\ &= \left|\int_I \mathbb{1}_{[x,y]}(t)u'(t)dt\right| && (31.30b)\\ &\leqslant \|\mathbb{1}_{[x,y]}\|_{L^2}\|u'\|_{L^2} && (31.30c)\\ &\leqslant \sqrt{|x-y|}\|u'\|_{H^1} && (31.30d)\\ &\leqslant \sqrt{|x-y|}. && (31.30e) \end{aligned}$$

où nous insistons sur le fait que la continuité n'impliquant pas la dérivabilité, le u' ici est la dérivé au sens de H^1, et non la dérivée usuelle. Quoi qu'il en soit, l'ensemble $\tilde{\mathcal{B}}$ est équicontinu [2]. Nous montrons à présent qu'il est également borné pour la norme uniforme. Soit $u \in \tilde{\mathcal{B}}$; vu la construction du représentant continu au lemme 31.3, nous avons

$$\begin{aligned} \big|u(x)\big| &= \left|\frac{1}{b-a}\int_a^b u(x)dy\right| && (31.31a)\\ &= \left|\frac{1}{b-a}\int_a^b \left(\int_y^x u'(t)dt - u(y)\right)dy\right| && (31.31b)\\ &= \left|\frac{1}{b-a}\int_a^b\int_y^x u'(t)dtdy - \frac{1}{b-a}\int_a^b u(y)dy\right| && (31.31c)\\ &\leqslant \frac{1}{b-a}\int_a^b\int_a^b |u'(t)|dt\,dy + \frac{1}{b-a}\int_a^b |u(y)|dy. && (31.31d) \end{aligned}$$

EqCFwSOxb

1. Encore que par souci d'économie d'encre nous n'allons pas écrire toujours les tildes et noter u le représentant continu prolongé à $\bar{I}$ par le corolaire 31.5.
2. Définition 7.317.

À ce niveau, il faut remarquer que dans la première intégrale, le passage de la valeur absolue à l'intérieur de l'intégrale en même temps que l'élargissement des bornes n'a rien d'innocent. Si $x < y$, les bornes ne sont pas « dans le bon ordre » et nous ne pouvons pas faire la majoration usuelle en entrant simplement la valeur absolue. Ici nous tenons compte de cela en élargissant les bornes, et en les mettant dans le bon ordre. Le passage exact est le suivant : si $x, y \in]a, b[$, nous avons

$$\left|\int_y^x f(t)dt\right| \leqslant \left|\int_y^x |f(t)|dt\right| \leqslant \left|\int_a^b |f(t)|dt\right| = \int_a^b |f(t)|dt. \tag{31.32}$$

Notons en particulier que dans le cas du passage vers l'équation (31.31d), le nombre x est fixé alors que y est une variable d'intégration. Donc l'ordre des deux est certainement de temps en temps le « mauvais ».

Quoi qu'il en soit, la première intégrale se réduit à une multiplication par $b-a$ et le calcul continue :

$$\begin{aligned}
|u(x)| &\leqslant \int_I |u'(t)|dt + \frac{1}{b-a}\int_I |u| && (31.33a)\\
&\leqslant \sqrt{b-a}\|u'\|_{L^2} + \frac{1}{\sqrt{b-a}}\|u\|_{L^2} && (31.33b)\\
&\leqslant \left(\sqrt{b-a} + \frac{1}{\sqrt{b-a}}\right)\left(\|u'\|_{L^2} + \|u\|_{L^2}\right) && (31.33c)\\
&\leqslant \left(\sqrt{b-a} + \frac{1}{\sqrt{b-a}}\right)\|u\|_{H^1} && (31.33d)\\
&= \sqrt{b-a} + \frac{1}{\sqrt{b-a}}. && (31.33e)
\end{aligned}$$

Donc $\tilde{\mathcal{B}}$ est borné pour la norme L^∞. Et c'est même borné par un nombre facilement calculable connaissant I. En particulier l'ensemble

$$\{u(x) \text{ tel que } u \in H^1\} \tag{31.34}$$

est, pour tout x, contenu dans la boule de rayon $\sqrt{b-a} + \frac{1}{\sqrt{b-a}}$ et donc est relativement compact dans $\mathbb{R}$. Par conséquent le théorème d'Ascoli 27.4 nous dit que l'ensemble $\tilde{\mathcal{B}}$ est relativement compact dans $C^0(I)$.

Par conséquent nous avons montré que l'image par ψ de la boule unité fermée $\mathcal{B}$ de $H^1(I)$ est relativement compacte dans $C^0(\bar{I})$, ce qui signifie que ψ est une application compacte.

(3) Les éléments de $H^1(I)$ sont des éléments de $L^2(I)$; donc l'identité est une injection. Nous devons seulement étudier la continuité. Si (u_n) est une suite dans H^1 convergeant dans H^1 vers u, alors

$$\|u_n - u\|_{L^2} \leqslant \|u_n - u\|_{L^2} + \|u_n' - u'\|_{L^2} = \|u_n - u\|_{H^1} \to 0. \tag{31.35}$$

Donc la suite des images (par l'identité) converge dans L^2. L'identité est donc continue.

□

31.1.2 Sur un ouvert de $\mathbb{R}^n$

Soit Ω, un ouvert de $\mathbb{R}^n$ et $v \in L^2(\Omega)$ (voir 27.91). Les fonctions considérées sont à valeurs réelles.

31.1.2.1 Définition

Définition 31.7 (Espace de Sobolev $H^1(\Omega)$)**.**
*Soit Ω une partie de $\mathbb{R}^n$. L'espace de **Sobolev** $H^1(\Omega)$ est :*

$$H^1(\Omega) = \{v \in L^2(\Omega) \text{ tel que } \forall i = 1, \dots, n, \partial_i v \in L^2(\Omega)\}. \tag{31.36}$$

Nous munissons cet espace d'un produit scalaire

$$(u,v)_{H^1} = \langle u,v\rangle_{L^2} + \langle \nabla u, \nabla v\rangle_{L^2}, \tag{31.37}$$

EQooQRMKooLaMpcp

où $\nabla u = \sum_i \partial_i u \in L^2$.

L'existence des intégrales dans le produit scalaire est assurée par le fait que u, v, ∇u et ∇v sont dans $L^2(\Omega)$. La définition du produit scalaire dans L^2 est la définition 27.337 (mais sans la conjugaison complexe).

Pour la même raison, $(u,u)_{H^1} = 0$ demande que chacun des deux termes est séparément nul, et nous avons $u = 0$ dans L^2, et donc aussi dans H^1.

Théorème 31.8 ([721]).
L'espace $H^1(\Omega)$ est un espace de Hilbert[3].

Démonstration. Nous devons nous assurer que l'espace H^1 est complet. Pour cela nous considérons une suite de Cauchy (u_n) dans H^1. Soit $\epsilon > 0$; il existe $N > 0$ tel que si $n,m > N$ alors $\|u_n - u_m\|_{H^1} < \epsilon$. Or en déballant la définition de la norme et du produit scalaire,

$$\begin{aligned}\|u_m - u_n\|_{H^1}^2 &= (u_m - u_n, u_m - u_n)_{H^1} && (31.38a)\\ &= \langle u_m - u_n, u_m - u_n\rangle + \langle \nabla(u_m - u_n), \nabla(u_m - u_n)\rangle && (31.38b)\\ &= \|u_m - u_n\|_{L^2}^2 + \|\nabla(u_m - u_n)\|_{L^2}^2. && (31.38c)\end{aligned}$$

En particulier les suites (u_n) et (∇u_n) sont de Cauchy dans L^2. Vu que L^2, lui, est complet (théorème 27.44), il existe $u \in L^2$ et $v_i \in L^2$ tels que

$$u_n \xrightarrow{L^2} u \tag{31.39a}$$
$$\partial_i u_n \xrightarrow{L^2} v_i. \tag{31.39b}$$

Nous savons que l'injection $i\colon L^2 \to \mathscr{D}'$ est continue par la proposition 30.35. Nous avons donc aussi les limites

$$T_{u_n} \xrightarrow{\mathscr{D}'} T_u \tag{31.40a}$$
$$T_{\partial_i u_n} \xrightarrow{\mathscr{D}'} T_{v_i}. \tag{31.40b}$$

SUBEQooMWLJooWakTkx

La dérivée étant une opération continue sur $\mathscr{D}'$ nous avons de plus

$$\partial_i(T_{u_n}) \xrightarrow{\mathscr{D}'} \partial_i(T_u) \tag{31.41}$$

En utilisant le lemme 30.38 nous avons alors

$$T_{\partial_i u_n} = \partial_i(T_{u_n}) \xrightarrow{\mathscr{D}} \partial_i(T_u) = T_{\partial_i u}. \tag{31.42}$$

En comparant avec (31.40b) et par l'unicité de la limite, nous avons $T_{v_i} = T_{\partial_i u}$. Cela implique $v_i = \partial_i u$.

Vu que $v_i \in L^2$ nous avons aussi $\partial_i u \in L^2$. Par conséquent $u \in H^1(\Omega)$ parce que ses dérivées sont dans L^2.

Nous devons maintenant prouver que $u_n \xrightarrow{H^1} u$. Nous avons

$$\|u_n - u\|_{H^1}^2 = \|u_n - u\|_{L^2}^2 + \|\nabla u_n - \nabla u\|_{L^2}^2 \tag{31.43}$$

Le premier terme tend vers zéro parce que $u_n \xrightarrow{L^2} u$ et le second parce que $\partial_i u_n \xrightarrow{L^2} \partial_i u$. □

3. Définition 25.3.

31.1.3 Espace de Sobolev fractionnaire

Définition 31.9.
*Pour $m \in \mathbb{N}$ et un ouvert Ω de $\mathbb{R}^d$ nous définissons l'**espace de Sobolev***

$$H^m(\Omega) = \{u \in L^2(\Omega) \text{ tel que } \partial^\alpha u \in L^2(\Omega) \, \forall |\alpha| \leqslant m\}. \tag{31.44}$$

Nous définissons également un produit scalaire sur H^m par

$$(u, v)_{H^m} = \sum_{|\alpha| \leqslant m} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2}. \tag{31.45}$$

En particulier la topologie est celle de la norme dérivée du produit scalaire :

$$\|u\|^2_{H^m(\Omega)} = \sum_{|\alpha| \leqslant m} \|\partial^\alpha u\|^2_{L^2(\Omega)}. \tag{31.46}$$

Le lemme suivant montre que la proposition 29.19 fonctionne encore avec L^2 au lieu de $\mathscr{S}$.

Lemme 31.10 (Lemme de transfert[722], thème 76)**.**
Soit $f \in H^m(\mathbb{R}^d)$. Alors pour tout multiindice α avec $|\alpha| \leqslant m$ nous avons

$$\mathcal{F}(\partial^\alpha f) = \big[\xi \mapsto i^{|\alpha|} \xi^\alpha \hat{f}(\xi)\big]. \tag{31.47}$$

Lemme 31.11.
Il existe des constantes c_1 et c_2 telles que pour tout $x \in \mathbb{R}^d$,

$$c_1(1 + \|x\|^2)^m \leqslant \sum_{|\alpha| \leqslant m} (x^\alpha)^2 \leqslant c_2(1 + \|x\|^2)^m. \tag{31.48}$$

Lemme 31.12.
Soit $u \in L^2(\mathbb{R}^d)$. Nous avons $u \in H^m(\mathbb{R}^d)$ si et seulement si l'application

$$\xi \mapsto \big(1 + |\xi|^2\big)^{k/2} \hat{u} \tag{31.49}$$

est dans $L^2(\mathbb{R}^d)$ pour tout $k \leqslant m$. Ici $|\xi|$ est la norme euclidienne de ξ dans $\mathbb{R}^d$.

Démonstration. Vu le lemme 31.10, il suffit de montrer que

$$\big(1 + |\xi|^2\big)^{k/2} \hat{u} \tag{31.50}$$

est dans L^2 pour tout $k \leqslant m$ si et seulement si

$$\xi^\alpha \hat{u} \tag{31.51}$$

l'est pour tout α avec $|\alpha| \leqslant m$.

L'expression (31.50) est une somme d'expressions du type (31.51). Donc l'implication dans un sens est montrée. Pour l'autre sens, nous savons que

$$\xi^\alpha = \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n}, \tag{31.52}$$

et donc

$$|\xi^\alpha| \leqslant |\xi_1|^{\alpha_1} \dots |\xi_n|^{\alpha_n}. \tag{31.53}$$

Or $|\xi|^{|\alpha|} = |\xi|^{\sum_i \alpha_i} = |\xi|^{\alpha_1} \dots |\xi|^{\alpha_n}$ et $|\xi| \geqslant |\xi_i|$ pour tout i, donc

$$|\xi^\alpha| \leqslant |\xi|^{|\alpha|}. \tag{31.54}$$

D'autre part pour tout $x \in \mathbb{R}^+$ et tout k positif nous avons

$$(1 + x^2)^{k/2} \geqslant x^k \tag{31.55}$$

qui est facile à vérifier en prenant le carré des deux membres.

En remettant tout ensemble,

$$|\xi^\alpha \hat{u}| \leqslant |\xi^\alpha||\hat{u}| \leqslant |\xi|^{|\alpha|}|\hat{u}| \leqslant \big(1 + |\xi|^2\big)^{|\alpha|/2}|\hat{u}|. \tag{31.56}$$

Donc si le membre de droite est de carré intégrable, celui de gauche l'est également. □

DEFooWEAQooAIWBwx

Définition 31.13 (Espace de Sobolev H^s[723]).
Pour $s > 0$ nous définissons l'espace de Sobolev $H^s(\mathbb{R}^d)$ par

$$H^s(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) \text{ tel que } \big(1 + \|\xi\|^2\big)^{s/2}\hat{u} \in L^2(\mathbb{R}^d)\}. \tag{31.57}$$

Nous y mettons le produit scalaire

$$(u, v)_{H^s} = \int_{\mathbb{R}^d} \hat{u}(\xi)\overline{\hat{v}(\xi)}(1 + \|\xi\|^2)^s d\xi. \tag{31.58}$$

31.14.
Vu que $\mathscr{D}(\mathbb{R}^d)$ est dense dans $L^2(\mathbb{R}^d)$ (théorème 27.51), on pourrait croire à la densité a fortiori dans $H^s(\mathbb{R}^d)$. Mais attention : $\mathscr{D}(\mathbb{R}^d)$ est dense dans L^2 pour la norme L^2. Nous n'avons encore rien dit pour la norme $H^s(\mathbb{R}^d)$.

PROPooMKAFooKDNTbO

Proposition 31.15 ([724]).
La partie $\mathscr{S}(\mathbb{R}^d)$ est dense dans $H^s(\mathbb{R}^d)$.

Démonstration. Soit $u \in H^s(\mathbb{R}^d)$. Par définition l'application

$$\xi \mapsto (1 + \|\xi\|^2)^{s/2}\hat{u} \tag{31.59}$$

est dans $L^2(\mathbb{R}^d)$. Elle peut donc être approximée au sens L^2 par des fonctions dans $\mathscr{D}(\mathbb{R}^d)$ (théorème 27.51(5)), c'est-à-dire qu'il existe des fonctions $\phi_n \in \mathscr{D}(\mathbb{R}^d)$ telles que

$$\phi_n \xrightarrow{L^2(\mathbb{R}^d)} (1 + \|\xi\|^2)^{s/2}\hat{u}. \tag{31.60}$$

Nous posons

$$\psi_n = \frac{\phi_n}{(1 + \xi^2)^{s/2}} \tag{31.61}$$

Cela est encore une fonction de $\mathscr{D}(\mathbb{R}^d)$, et donc de $\mathscr{S}(\mathbb{R}^d)$. Vu que la transformée de Fourier est une bijection de $\mathscr{D}(\mathbb{R}^d)$ (proposition 29.26), nous pouvons considérer une suite $\varphi_n \in \mathscr{D}(\mathbb{R}^d)$ telle que $\hat{\varphi}_n = \psi_n$, et nous allons montrer que $\varphi_n \xrightarrow{H^s(\mathbb{R}^d)} u$.

Nous avons :

$$\begin{aligned}
\|\varphi_n - u\|_{H^s}^2 &= \int_{\mathbb{R}^d} |\hat{\varphi}_n - \hat{u}|^2(1 + \xi^2)^s d\xi && \text{(31.62a)}\\
&= \int_{\mathbb{R}^d} \Big|\frac{\phi_n(\xi)}{(1 + \xi^2)^{s/2}} - \hat{u}(\xi)\Big|^2(1 + \xi^2)^s d\xi && \text{(31.62b)}\\
&= \int_{\mathbb{R}^d} |\phi_n(\xi) - \hat{u}(\xi)(1 + \xi^2)^{s/2}|^2 d\xi && \text{(31.62c)}\\
&= \|\phi_n - (1 + \xi^2)^{s/2}\hat{u}\|_{L^2}^2. && \text{(31.62d)}
\end{aligned}$$

Par définition de la suite ϕ_n nous avons donc bien

$$\|\varphi_n - u\|_{H^s}^2 = \|\phi_n - (1 + \xi^2)^{s/2}\hat{u}\|_{L^2}^2 \to 0. \tag{31.63}$$

Notons que même si ϕ_n est dans $\mathscr{D}(\mathbb{R}^d)$, nous n'avons pas prouvé la convergence $\phi_n \xrightarrow{H^s} u$, mais bien $\varphi_n \xrightarrow{H^s} u$. Or les fonctions φ_n sont dans $\mathscr{S}(\mathbb{R}^d)$, et rien n'assure qu'elles soient à support compact. Nous avons donc bien prouvé la densité de $\mathscr{S}$ et non celle de $\mathscr{D}$. □

Remarque 31.16.
Pour qui a tout compris, cela peut sembler une évidence, mais nous précisions que nous parlons de densité de $\mathscr{S}(\mathbb{R}^d)$ dans $H^s(\mathbb{R}^d)$, à aucun moment la topologie de $\mathscr{S}(\mathbb{R}^d)$ n'entre en compte.

Un peu moins évident : ce que nous avons réellement montré est la densité de $\iota\big(\mathscr{D}(\mathbb{R}^d)\big)$ dans $H^s(\mathbb{R}^d)$ où ι est l'application « prise de classe ». Nous n'avons pas insisté là-dessus, mais il faut dire que dans la preuve de la proposition 31.15, u est un représentant d'un élément choisi dans $H^s(\mathbb{R}^d)$.

Nous avons ensuite prouvé la convergence $\|\varphi_n - u\|_{H^s(\mathbb{R}^d)} \to 0$ qui est une convergence d'une suite dans $\mathbb{R}$, et dans laquelle l'opération $\|.\|_{H^s}$ est définie sur un espace de fonctions et n'est pas une norme (c'est pour que cela devienne une norme que l'on prend les classes).

Nous en avons déduit la convergence $\varphi_n \xrightarrow{H^s(\mathbb{R}^d)} u$ où maintenant φ_n et u sont des classes dans $H^s(\mathbb{R}^d)$.

PROPooLIQJooKpWtnV

Proposition 31.17.
La partie $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\big(H^s(\mathbb{R}^d), \|.\|_{H^s(\mathbb{R}^d)}\big)$.

Démonstration. Nous savons déjà que $\mathscr{S}(\mathbb{R}^d)$ est dense dans $H^s(\mathbb{R}^d)$ par la proposition 31.15. Nous devons seulement prouver que $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\big(\mathscr{S}(\mathbb{R}^d), \|.\|_{H^s(\mathbb{R}^d)}\big)$. Pour cela nous utilisons la densité de $\mathscr{D}(\mathbb{R}^d)$ dans $\mathscr{S}(\mathbb{R}^d)$ de la proposition 27.243. Soit donc $f \in \mathscr{S}(\mathbb{R}^d)$ et une suite f_k dans $\mathscr{D}(\mathbb{R}^d)$ telle que

$$f_k \xrightarrow{\mathscr{S}(\mathbb{R}^d)} f. \tag{31.64}$$

Vu que la transformée de Fourier est continue sur $\mathscr{S}(\mathbb{R}^d)$ (proposition 29.20) nous avons aussi

$$\hat{f}_k \xrightarrow{\mathscr{S}(\mathbb{R}^d)} \hat{f}, \tag{31.65}$$

et en particulier pour tout polynôme P nous avons la convergence uniforme

$$P\hat{f}_k \xrightarrow{unif} P\hat{f}. \tag{31.66}$$

EQooDTMUooQWphpR

D'autre part la fonction $\xi \mapsto |\hat{f}_k(\xi) - \hat{f}(\xi)|^2(1+\xi^2)^s$ est Schwartz et en tout point décroissante en k. Soientt $\epsilon > 0$ et $r > 0$ choisis de telle sorte à avoir

$$\int_{\|\xi\|>r} |\hat{f}_k(\xi) - \hat{f}(\xi)|(1+\xi^2)^s d\xi < \epsilon. \tag{31.67}$$

pour tout k. La convergence uniforme (31.66) permet de considérer k_0 tel que pour tout $k > k_0$,

$$|\hat{f}_k - \hat{f}|(1+\xi^2)^s < \frac{\epsilon}{\mathrm{Vol}\big(B(0,r)\big)} \tag{31.68}$$

dans $B(0,r)$. Avec tout cela, dès que $k > k_0$ nous avons

$$\|f_k - f\|_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}} |\hat{f}_k - \hat{f}|(1+\xi^2)^s d\xi = \int_{B(0,r)} \ldots + \int_{\|\xi\|>r} \ldots \leqslant 2\epsilon. \tag{31.69}$$

Donc nous avons bien $\|f_k - f\|_{H^s(\mathbb{R}^d)} \to 0$ et convergence de f_k vers f dans $H^s(\mathbb{R}^d)$. □

31.2 Trace

Définition 31.18 ([723])**.**
Nous définissons la ***trace*** *d'une fonction par*

$$\begin{aligned} \gamma_0 \colon \mathscr{D}(\mathbb{R}^d) &\to \mathscr{D}(\mathbb{R}^{d-1}) \\ (\gamma_0 v)(x_1, \ldots, x_{d-1}) &= v(x_1, \ldots, x_{d-1}, 0). \end{aligned} \tag{31.70}$$

THOooXEJZooBKtXBW

Théorème 31.19 ([725, 723])**.**
Si $s > \frac{1}{2}$, alors γ_0 accepte une unique extension en opérateur linéaire borné

$$\gamma_0 \colon H^s(\mathbb{R}^d) \to H^{s-\frac{1}{2}}(\mathbb{R}^{d-1}). \tag{31.71}$$

Démonstration. Nous subdivisons la preuve en plusieurs pas.

(1) **Une inégalité pour $\varphi \in \mathscr{D}(\mathbb{R}^d)$** Nous commençons par considérer $\varphi \in \mathscr{D}(\mathbb{R}^d)$ (fonction C^∞ à support compact). Nous allons alors prouver que

$$\|\gamma_0\varphi\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} \leqslant K\|\varphi\|_{H^s(\mathbb{R}^d)} \tag{31.72}$$

pour une certaine constante K (qui ne dépend en particulier pas de φ).
Nous avons

$$\|\gamma_0\varphi\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} = (\gamma_0\varphi, \gamma_0\varphi)_{H^{s-\frac{1}{2}}} \tag{31.73a}$$
$$= \int_{\mathbb{R}^{d-1}} |\widehat{\gamma_0\varphi(\xi)}|^2(1+\|\xi\|^2)^{s-\frac{1}{2}}d\xi \tag{31.73b}$$
$$= \int_{\mathbb{R}^{d-1}} \Big| \int_{\mathbb{R}^{d-1}} (\gamma_0\varphi)(x)e^{-i\xi x}dx\Big|^2(1+\|\xi\|^2)^{s-\frac{1}{2}}d\xi \tag{31.73c}$$
$$\tag{31.73d}$$

SUBEQooKLLYooNRiPwn

Nous appliquons la trace en appliquant la formule du corolaire 29.27,

$$(\gamma_0\varphi)(x) = \varphi(x,0) = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-iky}\varphi(x,y)dy\,dk \tag{31.74}$$

En remplaçant dans (31.73c) nous avons

$$\|\gamma_0\varphi\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} = \frac{1}{2\pi}\int_{\mathbb{R}^{d-1}} \Big| \int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-iky}e^{-i\xi x}\varphi(x,y)dy\,dk\,dx\Big|^2(1+\|\xi\|^2)^{s-\frac{1}{2}}d\xi. \tag{31.75}$$

Nous voudrions permuter les intégrales en k et en x. Pour cela nous étudions la fonction $u\colon \mathbb{R}\times\mathbb{R}^{d-1}\to\mathbb{C}$ donnée par

$$u(k,x) = e^{-ikx}\int_{\mathbb{R}} e^{-i\xi x}\varphi(x,y)dy \tag{31.76}$$

Effectuer l'intégrale par rapport à y revient à calculer la transformée de Fourier partielle dont nous parlons dans la proposition 29.21 [4]. Elle est donc une fonction Schwartz de k et de x (conjointement et non seulement séparément) et est donc dans $L^1(\mathbb{R}\times\mathbb{R}^{d-1})$. Les intégrales sur k et sur x peuvent donc être réunies et permutées par le théorème de Fubini 14.304 (n'oubliez tout de même pas de vous convaincre que la condition (2) est remplie).
Nous avons donc

$$\|\gamma_0\varphi\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} = \frac{1}{2\pi}\int_{\mathbb{R}^{d-1}} \Big| \int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}} e^{-iky}e^{-i\xi x}\varphi(x,y)dy\,dx\,dk\Big|^2(1+\|\xi\|^2)^{s-\frac{1}{2}}d\xi. \tag{31.77}$$

Étant donné que φ est à support compact, les intégrales sur x et sur y peuvent se réunir en utilisant encore le théorème de Fubini ; ces intégrales donnent :

$$\int_{\mathbb{R}^{d-1}\times\mathbb{R}} e^{-iky}e^{-i\xi x}\varphi(x,y)dx\otimes dy = \int_{\mathbb{R}^{d-1}\times\mathbb{R}} e^{-i(\xi,k)\,\cdot\,(x,y)}\varphi(x,y)dx\otimes dy = \hat{\varphi}(\xi,k). \tag{31.78}$$

Nous restons avec

$$\|\gamma_0\varphi\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} = \frac{1}{2\pi}\int_{\mathbb{R}^{d-1}} \Big| \int_{\mathbb{R}} \hat{\varphi}(\xi,k)dk\Big|^2(1+\|\xi\|^2)^{s-\frac{1}{2}}d\xi. \tag{31.79}$$

Nous allons maintenant traiter la partie du milieu :

$$\clubsuit = \Big| \int_{\mathbb{R}} \hat{\varphi}(\xi,k)dk\Big| = \Big| \int_{\mathbb{R}} \hat{\varphi}(\xi,k)(1+\xi^2+k^2)^{s/2}\frac{1}{(1+\xi^2+k^2)^{s/2}}dk\Big| = |\langle f_1, f_2\rangle_{L^2(\mathbb{R}^d)}| \tag{31.80}$$

4. Dont une relecture de la preuve ne serait vraiment pas de trop, ainsi que la preuve de 27.237.

Ici ξ est vu comme une constante et les fonctions f_1 et f_2 sont

$$f_1 : k \mapsto \hat{\varphi}(\xi,k)(1+\xi^2+k^2)^{s/2} \tag{31.81a}$$

$$f_2 : k \mapsto \frac{1}{(1+\xi^2+k^2)^{s/2}} \tag{31.81b}$$

Nous pouvons utiliser l'inégalité de Cauchy-Schwarz 10.1 :

$$\clubsuit \leqslant \left(\int_{\mathbb{R}} |\hat{\varphi}(\xi,k)|^2 (1+\xi^2+k^2)^s dk \right)^{1/2} \left(\int_{\mathbb{R}} \frac{1}{(1+\xi^2+k^2)^s} dk \right)^{1/2} \tag{31.82}$$

Nous notons $g(\xi)$ ce qui se trouve dans la seconde parenthèse (après intégration sur k). Avec cela nous continuons :

$$\| \gamma_0 \varphi \|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} \leqslant \frac{1}{2\pi} \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |g(\xi)| |\hat{\varphi}(\xi,k)|^2 (1+\xi^2+k^2)^s (1+\|\xi\|^2)^{s-\frac{1}{2}} dk\, d\xi. \tag{31.83}$$

Vu que $\hat{\varphi}$ est Schwartz, la fonction qui est à l'intérieur des deux intégrales est dans $L^1(\mathbb{R}^{d-1} \times \mathbb{R})$ et nous pouvons réunir les deux intégrales :

$$\| \gamma_0 \varphi \|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} \leqslant \frac{1}{2\pi} \int_{\mathbb{R}\times\mathbb{R}^{d-1}} |g(\xi)| |\hat{\varphi}(\xi,k)|^2 (1+\xi^2+k^2)^s (1+\|\xi\|^2)^{s-\frac{1}{2}} dk \otimes d\xi. \tag{31.84}$$

À ce point nous démontrons qu'en réalité la combinaison $g(\xi)(1+\xi^2)^{s-\frac{1}{2}}$ ne dépend pas de ξ. En effet

$$g(\xi)(1+\xi^2)^{s-\frac{1}{2}} = (1+\xi^2)^{s-\frac{1}{2}} \int_{\mathbb{R}} \frac{1}{(1+\xi^2+k^2)} dk \tag{31.85a}$$

$$= \frac{1}{(1+\xi^2)^{1/2}} \int_{\mathbb{R}} \left(\frac{1+\xi^2}{1+\xi^2+k^2} \right)^s dk \tag{31.85b}$$

$$= \frac{1}{(1+\xi^2)^{1/2}} \int \left(\frac{1}{1+\frac{k^2}{1+\xi^2}} \right)^s dk. \tag{31.85c}$$

Nous effectuons le changement de variables $t = \frac{k}{\sqrt{1+\xi^2}}$, $dk = (1+\xi^2)^{1/2} dt$, et le tout vaut

$$\int_{\mathbb{R}} \left(\frac{1}{1+t^2} \right)^s dt, \tag{31.86}$$

qui est effectivement indépendant de ξ. Nous nommons cela K (auquel nous ajoutons le $\frac{1}{2\pi}$) :

$$\| \gamma_0 \varphi \|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} \leqslant K \int_{\mathbb{R}\times\mathbb{R}^{d-1}} |\hat{\varphi}(\xi,k)|^2 (1+\xi^2+k^2)^s dk \otimes d\xi = K \| \varphi \|^2_{H^s(\mathbb{R}^d)}. \tag{31.87}$$

Nous avons donc prouvé pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$ (avec redéfinition du K) :

$$\| \gamma_0 \varphi \|_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})} \leqslant K \| \varphi \|_{H^s(\mathbb{R}^d)}. \tag{31.88}$$

EQooNWCIooALMivH

(2) **À propos de classes** Il serait tentant de conclure en disant que $\mathscr{D}(\mathbb{R}^d)$ est dense dans $H^s(\mathbb{R}^d)$. Hélas, techniquement, l'ensemble $\mathscr{D}(\mathbb{R}^d)$ n'est même pas un sous-ensemble de $H^s(\mathbb{R}^d)$ parce que ce dernier est un ensemble de *classes* de fonctions. Ce petit détail a ici son importance parce que γ_0 n'est pas une application qui descend aux classes. En effet, $\mathbb{R}^{d-1}$ étant de mesure nulle dans $\mathbb{R}^d$, deux fonctions de la même classe peuvent différer en *tous* les points de $\mathbb{R}^{d-1}$ en même temps.

Notons ι l'application qui consiste à prendre la classe, c'est-à-dire $\iota(f) = [f]$. Ce qui est dense dans $H^s(\mathbb{R}^d)$, c'est $\iota(\mathscr{D}(\mathbb{R}^d))$. Or chaque classe contient au maximum une seule fonction continue (qui sera même de classe C^∞ à support compact pour les éléments de $\iota(\mathscr{D})$).

L'application γ_0 considérée est l'application composée entre le γ_0 classique et le choix du représentant continu dans la classe. La formule (31.88) que nous venons de prouver est valide pour l'application γ_0 vue comme

$$\gamma_0 : \iota\big(\mathscr{D}(\mathbb{R}^d)\big) \to \mathscr{D}(\mathbb{R}^{d-1}). \tag{31.89}$$

(3) **Densité et conclusion** Ce que la majoration (31.88) prouve est la continuité de l'application

$$\gamma_0 \colon \big(\mathscr{D}(\mathbb{R}^d), \|.\|_{H^s(\mathbb{R}^d)}\big) \to \big(H^{s-\frac{1}{2}}(\mathbb{R}^{d-1}), \|.\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})}\big). \tag{31.90}$$

Mais la proposition 31.17 nous donne la densité de la partie $\mathscr{D}(\mathbb{R}^d)$ dans $H^s(\mathbb{R}^d)$. La proposition 17.127 nous donne alors une extension

$$\gamma_0 \colon \big(H^s(\mathbb{R}^d), \|.\|_{H^s(\mathbb{R}^d)}\big) \to \big(H^{s-\frac{1}{2}}(\mathbb{R}^{d-1}), \|.\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{d-1})}\big). \tag{31.91}$$

□

Remarque 31.20.
L'extension n'est pas évidente parce que les éléments de $H^s(\mathbb{R}^d)$ sont en général des classes de fonctions dont les valeurs sur le bord ne sont pas du tout fixées du fait que le bord soit de mesure nulle.

31.3 Théorème de plongement

L'objet des théorèmes de plongement de Sobolev est de montrer que si $s > \frac{d}{2} + k$ alors les éléments de $H^s(\mathbb{R}^d)$ possèdent des représentants de classe C^k. Avant de démontrer le théorème, pour alléger, nous allons donner deux lemmes.

LEMooZIBIooANHyPy

Lemme 31.21.
Soit (u_j) une suite dans $\mathscr{S}(\mathbb{R}^d)$ telle que

$$u_j \xrightarrow{H^s(\mathbb{R}^d)} u \tag{31.92}$$

avec $s > 0$. Alors nous avons aussi la convergence

$$u_j \xrightarrow{L^2(\mathbb{R}^d)} u. \tag{31.93}$$

Démonstration. Vu que $s > 0$ nous avons $(1 + k^2)^s > 1$ (ici nous écrivons k^2 pour $\|k\|^2$). Par conséquent

$$(u, v)_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \hat{u}\bar{\hat{v}}(1 + k^2)^s dk \geqslant \int_{\mathbb{R}^d} \hat{u}\bar{\hat{v}} dk = \langle \hat{u}, \hat{v} \rangle_{L^2(\mathbb{R}^d)}. \tag{31.94}$$

Nous avons alors

$$\begin{aligned}
\|u_j - u\|_{L^2} &= \frac{1}{(2\pi)^d} \|\hat{u}_j - \hat{u}\|_{L^2} && (31.95a)\\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{u}_j - \hat{u}|^2 && (31.95b)\\
&\leqslant \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{u}_j - \hat{u}|^2 (1 + k^2)^s dk && (31.95c)\\
&= \frac{1}{(2\pi)^d} \|u_j - u\|_{H^s(\mathbb{R}^d)}. && (31.95d)
\end{aligned}$$

□

LEMooGDTXooJRudME

Lemme 31.22.
Soient des fonctions $u_j \in \mathscr{S}(\mathbb{R}^d)$ telles que

$$u_j \xrightarrow{\left(C_0^0(\mathbb{R}^d), \|.\|_\infty\right)} v. \tag{31.96}$$

Alors nous avons la convergence

$$\int_{\mathbb{R}^d} u_j \varphi \to \int_{\mathbb{R}^d} v\varphi \tag{31.97}$$

pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

Démonstration. La suite (u_j) est équibornée. En effet il existe une queue de suite pour laquelle $\|u_j - v\|_\infty < \epsilon$; cette queue de suite est alors équibornée par $\|v\|_\infty + \epsilon$. Le début de la suite est un nombre fini de fonctions, toutes bornées. Le maximum des bornes donne alors une borne.

Soit donc $M > 0$ tel que $|u_j(x)| < M$ pour tout $x \in \mathbb{R}^d$ et pour tout $j \in \mathbb{N}$. Nous avons alors $|u_j\varphi| < M|\varphi|$ pour tout j et les fonctions $|u_j\varphi|$ sont majorées par la fonction $M|\varphi|$ qui est intégrable. Nous pouvons donc utiliser le théorème de la convergence dominée de Lebesgue 14.207 nous donne

$$\lim_{j\to\infty} \int_{\mathbb{R}^d} u_j\varphi = \int_{\mathbb{R}^d} v\varphi. \tag{31.98}$$

□

Nous pouvons écrire la conclusion du lemme 31.22 sous la forme

$$\langle u_j, \varphi\rangle_{L^2(\mathbb{R}^d)} \to \langle v, \varphi\rangle_{L^2(\mathbb{R}^d)} \tag{31.99}$$

pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$ (et non pour tout $\varphi \in L^2(\mathbb{R}^d)$.

THOooOHIPooXSEkVI

Théorème 31.23 (Théorème de Sobolev avec $k = 0$[726])**.**
Soit $s > \frac{d}{2}$ et $u \in H^s(\mathbb{R}^d)$. Alors u possède un représentant dans $C_0^0(\mathbb{R}^d)$ (les fonctions continues et qui s'annulent à l'infini). Nous écrivons cela $H^s(\mathbb{R}^d) \subset C_0^0(\mathbb{R}^d)$.

Démonstration. Nous commençons par supposer que $u \in H^s(\mathbb{R}^d) \cap \mathscr{S}(\mathbb{R}^d)$, et dans ce cas nous notons u le représentant dans $\mathscr{S}(\mathbb{R}^d)$. Nous allons prouver l'inégalité

$$\|u\|_\infty \leqslant c\|u\|_{H^s(\mathbb{R}^d)}. \tag{31.100}$$

La formule d'inversion de Fourier 29.26 appliquée à u_j donne

$$u(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ikx}\hat{u}(k)dk, \tag{31.101}$$

EQooDKYDooAxThqf

nous avons alors

$$\begin{aligned}
(2\pi)^d|u(x)| &\leqslant \int_{\mathbb{R}^d} |\hat{u}(k)|dk && \text{(31.102a)}\\
&= \int_{\mathbb{R}^d} (1+k^2)^{s/2}|\hat{u}(k)|(1+k^2)^{-s/2}dk && \text{(31.102b)}\\
&= \int_{\mathbb{R}^d} \underbrace{\left(|\hat{u}(k)|^2(1+k^2)^s\right)^{1/2}}_{f} \underbrace{\left((1+k^2)^{-s}\right)^{1/2}}_{g} dk && \text{(31.102c)}\\
&= \langle f, g\rangle_{L^2(\mathbb{R}^d)}. && \text{(31.102d)}
\end{aligned}$$

Ici il convient nous arrêter un instant pour nous convaincre que f et g sont réellement des éléments de L^2. En ce qui concerne f, c'est facile : $\hat{u}$ est une fonction Schwartz. En ce qui concerne g il faut l'intégrabilité de $|g|^2$, c'est-à-dire de $k \mapsto (1+k^2)^{-s}$. Cela a lieu si et seulement si $2s > d$ et donc a lieu dans les hypothèses du théorème. Nous utilisons le théorème de Cauchy-Schwarz [5] pour continuer :

$$\begin{aligned}
(2\pi)^d|u(x)| &\leqslant \|f\|_{L^2}\|g\|_{L^2} && \text{(31.103a)}\\
&= c\left(\int_{\mathbb{R}^d} |\hat{u}(k)|^2(1+k^2)^s dk\right)^{1/2} && \text{(31.103b)}\\
&= c\|u\|_{H^s(\mathbb{R}^d)}. && \text{(31.103c)}
\end{aligned}$$

Donc en introduisant le facteur $(2\pi)^d$ dans la constante c nous avons

$$\|u\|_\infty \leqslant c\|u\|_{H^s(\mathbb{R}^d)}. \tag{31.104}$$

5. Formule 10.2.

Cela est tout ce que nous voulions faire avec $u \in \mathscr{S}(\mathbb{R}^d)$.

Nous considérons maintenant $u \in H^s(\mathbb{R}^d)$. Vue la densité des fonctions Schwartz dans H^s (proposition 31.15) nous considérons une suite (u_j) dans $\mathscr{S}(\mathbb{R}^d)$ telle que

$$u_j \xrightarrow{H^s(\mathbb{R}^d)} u \tag{31.105}$$

Ici u est une classe, mais nous identifions u_j avec sa classe (parce qu'il ne faut pas exagérer non plus). La suite (u_j) est de Cauchy dans H^s, donc si $\epsilon > 0$ est donné, il existe N tel que si $n, m > N$, $\|u_m - u_n\| \leqslant \epsilon$. Nous avons alors aussi

$$\|u_m - u_n\|_\infty \leqslant c\epsilon, \tag{31.106}$$

ce qui signifie que (u_j) est également une suite de Cauchy dans $\big(C^0(\mathbb{R}^d), \|.\|_\infty\big)$ qui est un espace complet par la proposition 12.359.

Il existe donc une fonction $v \in C_0^0(\mathbb{R}^d)$ telle que

$$u_j \xrightarrow{\left(C_0^0(\mathbb{R}^d), \|.\|_\infty\right)} v. \tag{31.107}$$

EQooBNGCooBUAQYN

La question est de savoir si nous pouvons déduire que v est un représentant de u.

Par le lemme 31.21 nous avons également la convergence

$$u_j \xrightarrow{L^2(\mathbb{R}^d)} u. \tag{31.108}$$

Pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$ nous avons alors

$$\langle u_j, \varphi \rangle_{L^2} \to \langle u, \varphi \rangle_{L^2}. \tag{31.109}$$

Mais en même temps, la convergence (31.107) couplée au lemme 31.22 donne également

$$\langle u_j, \varphi \rangle_{L^2} \to \langle v, \varphi \rangle_{L^2}. \tag{31.110}$$

Par unicité de la limite (dans $\mathbb{R}$) nous avons

$$\langle v, \varphi \rangle_{L^2} = \langle u, \varphi \rangle_{L^2} \tag{31.111}$$

pour tout $\varphi \in \mathscr{S}(\mathbb{R}^d)$. La proposition 30.1 appliquée à $u - v$ montre alors que $u - v = 0$ presque partout, c'est-à-dire que v est bien un représentant de u.

Le représentant v de u est non seulement continu (comme limite uniforme de fonctions continues), mais également bornée, comme limite uniforme de fonctions Schwartz. □

Proposition 31.24 ([726])**.**
Si $u \in H^s(\mathbb{R}^d)$ ($s \in \mathbb{R}$) alors

(1) $\partial^\alpha u \in H^{s-|\alpha|}(\mathbb{R}^d)$,

(2) l'application

$$\partial^\alpha : H^s(\mathbb{R}^d) \to H^{s-|\alpha|}(\mathbb{R}^d) \tag{31.112}$$

est continue.

Note : ici ∂ est l'opération de dérivée faible.

Démonstration. Nous allons seulement prouver que $\partial_j : H^s(\mathbb{R}^d) \to H^{s-1}(\mathbb{R}^d)$ est bien définie[6] et continue. Par composition, la thèse suivra.

Soit $u \in H^s(\mathbb{R}^d)$ par le lemme 31.10 nous avons

$$\widehat{\partial_j u} = i\xi_j \hat{u}. \tag{31.113}$$

6. Au sens où l'espace d'arrivée est bien celui-là.

D'autre part, la fonction

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \frac{x_i}{1+\|x\|^2} \end{aligned} \tag{31.114}$$

est bornée (et même indépendamment de i) par une constante K. Donc nous avons pour tout [7] s :

$$k_i(1+\|k\|^2)^{-s} < K(1+\|k\|^2)^{-s+1}. \tag{31.115}$$

Avec cela nous pouvons calculer un peu : si $u \in H^s(\mathbb{R}^d)$, nous avons

$$\begin{aligned} \|\partial_j u\|_{H^{s-1}(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |\widehat{\partial_j u}|(1+k^2)^{s-1}dk && (31.116a)\\ &= \int_{\mathbb{R}^d} k_j|\hat{u}|(1+k^2)^{s-1}dk && (31.116b)\\ &\leqslant \int_{\mathbb{R}^d} K|\hat{u}|(1+k^2)^{s}dk && (31.116c)\\ &= K\|u\|_{H^s(\mathbb{R}^d)}. && (31.116d) \end{aligned}$$

Nous avons donc que $\|\partial_j u\|_{H^{s-1}(\mathbb{R}^d)}$ est fini lorsque $u \in H^s(\mathbb{R}^d)$.

La majoration $\|\partial_j u\| \leqslant K\|u\|$ donne la majoration suivante pour la norme de l'opérateur ∂_j :

$$\|\partial_j\| = \sup_{\|u\|_{H^s}=1} \|\partial_j u\|_{H^{s-1}} \leqslant K. \tag{31.117}$$

Le fait d'être borné implique d'être continu par la proposition 10.70. □

Théorème 31.25 (Théorème de plongement de Sobolev [726])**.**
Soient $k \in \mathbb{N}$ et $s > \frac{d}{2} + k$. Alors

$$H^s(\mathbb{R}^d) \subset C_0^k(\mathbb{R}^d). \tag{31.118}$$

EQooIJZQooZiYSnJ

Remarques :

— L'espace $C_0^k(\mathbb{R}^d)$ est l'ensemble des fonctions de classe C^k qui s'annulent à l'infini.

— L'inclusion (31.118) signifie que tout élément dans H^s possède un représentant dans $C_0^k(\mathbb{R}^d)$.

Démonstration. Pour $k = 0$, c'est le théorème 31.23. Si $|\alpha| < k$ nous savons que $\partial^\alpha u \in H^{s-k} \subset C_0^0(\mathbb{R}^d)$. Cela signifie que les dérivées faibles sont continues, mais pas qu'il existe un représentant qui est réellement k fois continument dérivable.

Soit $u \in H^s(\mathbb{R}^d)$ et une suite (u_j) dans $\mathscr{S}(\mathbb{R}^d)$ telle que

$$u_j \xrightarrow{H^s(\mathbb{R}^d)} u. \tag{31.119}$$

Vu que l'espace topologique $\big(C_0^k(\mathbb{R}^d), \|.\|_\infty\big)$ est complet il existe $v \in C_0^k$ tel que

$$u_j \xrightarrow{C_0^k} v. \tag{31.120}$$

Il reste à montrer que v est un représentant de u. Cela se fait comme plus haut en montrant que $u_j \xrightarrow{L^2} u$. □

7. Question : dans [726], il faut dépendre cette constante de s. Je ne comprends pas pourquoi.

Chapitre 32

Équations différentielles ordinaires

Une équation différentielle ordinaire est la recherche de toutes les fonctions définies sur une partie de $\mathbb{R}$ satisfaisant à une certaine égalité, faisant intervenir les dérivées de la fonction recherchée.

Définition 32.1.
*Si I est un intervalle de $\mathbb{R}$, une fonction sera **dérivable sur** I si elle est dérivable au sens usuel sur l'intérieur de I, et si elle est dérivable à droite (resp. à gauche) sur l'éventuel bord gauche (resp. droit) de I.*

Définition 32.2.
*Une **équation différentielle ordinaire d'ordre** n **sur** I est la recherche d'une fonction $y : I \to \mathbb{R}$ dérivable n fois, satisfaisant à une équation du type*

$$F(t, y(t), y'(t), \dots, y^{n\prime}(t)) = 0 \quad \text{pour tout } t \in I \tag{32.1}$$

eqequadiff

où I est un intervalle de $\mathbb{R}$ et $F : (I \times D) \subset (\mathbb{R} \times \mathbb{R}^{n+1}) \to \mathbb{R}$ est une fonction donnée.

Remarque 32.3.
L'équation différentielle (32.1) sera raccourcie sous la forme

$$F(t, y, y', \dots, y^{n\prime}) = 0 \tag{32.2}$$

où la dépendance en t est sous-entendue.

Exemple 32.4.
Soit $f : I \to \mathbb{R}$ une fonction continue fixée. L'équation différentielle

$$y' = f(t) \tag{32.3}$$

se ramène à la recherche des primitives de f sur l'intervalle I. △

Le lemme suivant sert de temps en temps.

LemuBVozy

Lemme 32.5 (Lemme de Grönwall)**.**
Soient ϕ et ψ deux fonctions telles que pour tout $t \in [t_0, t_1]$, $\phi(t) \geqslant 0$, $\psi(t) \geqslant 0$ et

$$\phi(t) \leqslant K + L \int_{t_0}^{t} \psi(s)\phi(s)ds \tag{32.4}$$

où K et L sont des constantes positives. Alors

$$\phi(t) \leqslant K \exp\big(L \int_{t_0}^{t} \psi\big). \tag{32.5}$$

LEMooUGZGooCczAmKa

Lemme 32.6 (Lemme de Grönwall[727]).
Si $u, a, b \in C^0\big([0,T], \mathbb{R}^+\big)$ sont telles que

$$u(t) \leqslant b(t) + \int_0^t a(s)u(s)ds \tag{32.6}$$

pour tout $t \in [0,T]$ alors pour tout $t \in [0,T]$ nous avons aussi

$$u(t) \leqslant b(t) + \int_0^t b(s)a(s)e^{\int_s^t a(u)du}ds. \tag{32.7}$$

32.1 Équation homogène, solution particulière

Voici un petit morceau d'algèbre linéaire. Soient des espaces vectoriels V et W ainsi qu'une application linéaire $D\colon V \to W$. Nous voulons résoudre $D(u) = v$, c'est-à-dire déterminer l'ensemble

$$D^{-1}(v) = \{u \in V \text{ tel que } Du = v\}. \tag{32.8}$$

LEMooEWUPooXNJMcc

Lemme 32.7.
Soient des espaces vectoriels V et W ainsi qu'une application linéaire $D\colon V \to W$. Si $u_P \in V$ satisfait à $Du_P = v$ alors

$$D^{-1}(v) = \ker(D) + u_P. \tag{32.9}$$

Démonstration. Si $u \in \ker(D) + u_p$ alors $u = k + u_p$ avec $Dk = 0$, ce qui donne tout de suite $Du = Dk + Du_P = v$. Donc $u \in D^{-1}(v)$.

Dans l'autre sens, si $u \in D^{-1}(v)$ alors nous pouvons écrire $u = (u - u_P) + u_P$. Vu que $u - u_P \in \ker(D)$ nous avons bien $u \in \ker(D) + u_P$. □

Ce petit lemme explique pourquoi la résolution d'équation différentielles passe par le principe « générale de l'homogène plus particulière de la non-homogène ». Cela marche autant pour les équations différentielles ordinaires que pour celles aux dérivées partielles.

Exemple 32.8.
Considérons l'équation différentielle ordinaire

$$y' - y = 4. \tag{32.10}$$

L'opérateur dont nous parlons est par exemple

$$\begin{aligned} D\colon C^\infty(\mathbb{R}) &\to C^\infty(\mathbb{R}) \\ y &\mapsto y' - y. \end{aligned} \tag{32.11}$$

Nous devons résoudre $Dy = 4$ où « 4 » est l'élément fonction constante égale à 4 dans $C^\infty(\mathbb{R})$. L'ensemble $\ker(D)$ sont les éléments $y \in C^\infty(\mathbb{R})$ tels que $y' = y$:

$$\ker(D) = \{t \mapsto Ke^t \text{ tel que } K \in \mathbb{R}\}. \tag{32.12}$$

Nous devons trouver un élément quelconque y_P de $D^{-1}(4)$. Facile : $y_P(t) = -4$.

Au final,

$$D^{-1}(4) = \{t \mapsto Ke^t - 4 \text{ tel que } K \in \mathbb{R}\}. \tag{32.13}$$

△

Dans cet exemple nous avons pris $V = W = C^\infty(\mathbb{R})$. Mais souvent nous sommes amenés à considérer des espaces plus subtils, parce qu'il existe simplement pas de solutions dans C^∞, ou alors parce que beaucoup de solutions n'y sont pas.

32.2 Que faire avec $f(z)dz = g(t)dt$?

SecFairedzdt

Dans de nombreux exercices d'équations différentielles, nous tombons sur $u' = f(t)$, et nous faisons formellement

$$\frac{du}{dt} = f(t) \quad \Rightarrow \quad du = f(t)dt, \tag{32.14}$$

et ensuite, il y a la formule un peu magique

$$u - u_0 = \int_{t_0}^{t} f(t)dt. \tag{32.15}$$

Voyons ce qu'il en est. Tout d'abord, il faut comprendre ce que signifie la formule

$$f(z)dz = g(t)dt. \tag{32.16}$$

EqDiffAstufzdz

Il s'agit d'une égalité entre deux formes différentielles sur $\mathbb{R}$ où z est une fonction de t. Étant donné que z est une fonction de t, il faut voir dz comme la différentielle de cette fonction. La différentielle d'une fonction à une variable est donné par la dérivée :

$$dz_t = z'(t)dt \tag{32.17}$$

Écrire l'équation (32.16) pour chaque t revient donc à écrire

$$f\big(z(t)\big)z'(t)dt = g(t)dt \tag{32.18}$$

Cela est une égalité entre deux formes différentielles. Nous avons donc égalité entre les intégrales des formes sur un chemin. Prenons un chemin tout simple de t_0 vers t :

$$\int_{t_0}^{t} f\big(z(t)\big)z'(t)dt = \int_{t_0}^{t} g(t)dt. \tag{32.19}$$

Dans le premier membre, nous faisons un changement de variable $\xi = z(t)$, $d\xi = z'(t)dt$, et nous obtenons

$$\int_{z_0}^{z(t)} f(\xi)d\xi = \int_{t_0}^{t} g(t)dt. \tag{32.20}$$

EqIntDiffAstuztz

où nous avons remplacé la constante $z(t_0)$ par z_0 dans la borne d'intégration. Si F est une primitive de f et G une primitive de g, nous avons

$$F(z) - F(z_0) = G(t) - G(t_0). \tag{32.21}$$

Si aucun problème de Cauchy n'est donné, les constantes $F(z_0)$ et $G(t_0)$ sont mises en une seule et nous écrivons la solution

$$F\big(z(t)\big) = G(t) + C, \tag{32.22}$$

qui est une équation implicite pour $z(t)$.

Nous trouvons assez souvent le cas simple

$$f(z)dz = dt. \tag{32.23}$$

EqAstfzdzdt

En remplaçant $g(t) = 1$ dans (32.20), nous trouvons la fameuse

$$t - t_0 = \int_{z_0}^{z} f(z)dz, \tag{32.24}$$

Eqttzint

dans laquelle il y a un abus de notation terrible entre le z de la borne (que les étudiants oublient souvent) et la variable d'intégration z !!

Le passage de (32.23) à (32.24) sera très souvent utilisé dans le cours de mécanique par exemple.

32.3 Équations linéaires du premier ordre

Une **équation différentielle linéaire** est une équation de la forme

$$y' + u(t)y = v(t). \tag{32.25}$$
EXooVVLGooPWaHUI

Exemple 32.9.
Tant qu'il n'y a pas de second membre, c'est facile. Prenons l'exemple suivant :

$$y' + 2ty = 0. \tag{32.26}$$

Nous mettons tous les t d'un côté et tous les y et y' de l'autre :

$$\frac{y'}{y} = -2t, \tag{32.27}$$

et puis on intègre sans oublier la constante d'intégration :

$$\ln(y) = -t^2 + C, \tag{32.28}$$

et donc $y(t) = Ke^{-t^2}$. △

Exemple 32.10.
Lorsqu'il y a un second membre, il y a une astuce. Prenons par exemple

$$y' + 2ty = 4t. \tag{32.29}$$
EqDiffExLin

L'astuce est de commencer par résoudre l'équation sans le second membre (l'équation homogène associée). Nous notons y_H la solution. Ici, la réponse est

$$y_H(t) = Ke^{-t^2}. \tag{32.30}$$

Ensuite le truc est d'essayer de trouver la solution de l'équation (32.29) sous la forme

$$y(t) = K(t)e^{-t^2}. \tag{32.31}$$
EqEssaiLin

L'idée est de prendre la même que la solution de l'équation homogène (sans second membre), mais en disant que K est une fonction. Afin de trouver la fonction K qui donne la solution, il suffit de remettre l'essai (32.31) dans l'équation (32.29) :

$$\underbrace{K'e^{-t^2} - 2tKe^{-t^2}}_{y'(t)} + \underbrace{2tKe^{-t^2}}_{2ty(t)} = 4t \tag{32.32}$$

Les deux termes avec K se simplifient et il reste

$$K'(t) = 4te^{t^2}, \tag{32.33}$$

ce qui signifie $K(t) = 2e^{t^2+C}$. Nous avons donc déterminé la fonction qui fait fonctionner l'essai, et la solution à l'équation est

$$y(t) = \big(2e^{t^2} + C\big)e^{-t^2} = 2 + Ce^{-t^2}. \tag{32.34}$$

△

La technique pour résoudre cette équation est de commencer par résoudre l'équation homogène associée. Si $U(t)$ est une primitive de $u(t)$, nous avons

$$\begin{aligned} y_H'(t) + u(t)y_H(t) &= 0 \\ \frac{y_H'}{y_H} &= -u(t) \\ \ln(y_H) &= -U(t) + C \\ y_H(t) &= e^{-U(t)+C} = Ke^{-U(t)} \end{aligned} \tag{32.35}$$

où $K = e^C$.

Cela fournit la solution générale de l'équation homogène. Il existe un truc génial qui permet d'en tirer la solution générale du système non homogène. Lorsque nous avons trouvé $y_H(t) = Ke^{-U(t)}$, le symbole K désigne une constante. La méthode de **variation des constantes** consiste à essayer la solution

$$y(t) = K(t)e^{-U(t)}, \tag{32.36}$$ EqEssayVarSctr

c'est-à-dire à dire que la constante est en réalité une fonction. Afin de trouver quelle fonction $K(t)$ fait en sorte que l'essai (32.36) soit une solution, nous la remplaçons dans l'équation de départ $y' + uy = v$. Maintenant,

$$y'(t) = K'(t)e^{-U(t)} - K(t)u(t)e^{-U(t)}. \tag{32.37}$$

En remettant dans l'équation,

$$y' + uy = K'e^{-U} - Kue^{-U} + uKe^{-U} = K'e^{-U} = v. \tag{32.38}$$

Notez que les termes en K se sont miraculeusement simplifiés. Cela est directement dû au fait que e^{-U} est solution de l'équation homogène. Nous restons avec l'équation

$$K' = \frac{v}{e^{-U}} \tag{32.39}$$

pour $K(t)$. La solution générale du problème non homogène est donc finalement donnée par

$$y(t) = \big(W(t) + C\big)e^{-U(t)} \tag{32.40}$$

si $W(t)$ est une primitive de $v(t)e^{U(t)}$.

Tout ceci est un peu heuristique. La proposition suivante dit dans quels cas ça fonctionne. PROPooZCXQooPQpkdQ

Proposition 32.11.
Soient u et v continues sur I et U, une primitive de u sur I et W une primitive de ve^{-U} sur I. Une fonction $y\colon I \to \mathbb{R}$ est solution de $y' + u(t)y = v(t)$ si et seulement si il existe une constante $C \in \mathbb{R}$ telle que

$$y(t) = \big(W(t) + C\big)e^{-U(t)} \tag{32.41}$$

pour tout $t \in I$.

32.3.1 Pourquoi la variation des constantes fonctionne toujours ?

Prenons une équation non homogène

$$z'(t) = f(t)z(t) + g(t), \tag{32.42}$$ EqAstNNHomo

et supposons avoir une solution de l'homogène associée sous la forme $z_H(t) = Ch(t)$. Le coup de la variation des constates consiste à essayer une solution pour l'équation non homogène sous la forme [1]

$$z(t) = K(t)h(t). \tag{32.43}$$

Nous injectons cette solution dans l'équation de départ en utilisant le fait que $z'(t) = K'(t)h(t) + K(t)h'(t)$:

$$K'(t)h(t) + K(t)h'(t) = f(t)K(t)h(t) + g(t). \tag{32.44}$$

Le terme $K(t)h'(t)$ se récrit en utilisant la propriété de définition de h, c'est-à-dire que $h'(t) = f(t)h(t)$. Nous voyons que les termes ne contenant pas de K' se simplifient ; il reste

$$K'h = g. \tag{32.45}$$

1. Je ne sais plus qui a eu l'idée de changer le nom de la constante de C vers K au moment de la transformer en fonction, mais c'est une bonne idée.

Cette équation a comme solution

$$K = \int \frac{f}{h} + C. \tag{32.46}$$

J'insiste sur la constante d'intégration ! En réalité, celles et ceux qui auront compris l'équation (32.24) sauront que K est donné par

$$K(t) = \int_{\xi_0}^{t} \frac{f(\xi)}{g(\xi)} d\xi \tag{32.47}$$

où ξ_0 joue le rôle de la constante d'intégration.

Quoi qu'il en soit, la solution générale de l'équation non homogène est

$$z(t) = K(t)h(t) = \left(\int \frac{g}{h} + C \right) h. \tag{32.48}$$ EqSolVarCosntCool

Cette solution comprend deux termes : Ch qui est solution de l'homogène, et $\left(\int \frac{g}{h}\right) h$ qui est une particulière de l'équation non homogène.

Quelques conclusions :

(1) Si vous avez encore du K (et pas que du K') dans votre équation qui donne K, c'est que vous n'êtes pas dans le cadre d'une équation de type (32.42). Le plus souvent, c'est que vous avez fait une faute de calcul quelque part.

(2) La méthode des variations des constantes n'est pas en contradiction avec le principe de « SGEH+SPENH ». En effet, la SGEP et la SPENH sont toutes deux dans la solution (32.48).

(3) La variation des constantes peut être vue comme une façon cool de trouver une solution particulière de l'équation non homogène.

(4) La simplification ne se fait que après avoir remplacé Kh' par Kfh, c'est-à-dire après avoir utilisé le fait que z_H est solution de l'homogène. Sinon, la simplification n'est pas du tout évidente à priori. Il se peut même que, visuellement, les termes Kh' et Kfh ne se ressemblent pas du tout. Un exemple de cela arrivera par exemple dans l'exemple 32.15, pour arriver à l'équation (32.69).

32.4 Équations à variables séparées

Secvarsep

Une **équation à variables séparées** est une équation de la forme

$$y' = u(t)f(y) \tag{32.49}$$ EqDiffSeparee

où $u\colon I \to \mathbb{R}$ et $f\colon J \to \mathbb{R}$ sont deux fonctions continues données. Les propositions 32.12 et 32.14 résolvent ce cas, mais avant de voir cela, nous allons donner quelques indications « pratiques ».

32.4.1 La méthode rapide

On peut évidemment mettre tous les y et y' d'un côté :

$$\frac{y'}{f(y)} = u(x). \tag{32.50}$$

Une fois que cela est fait, on écrit $y' = \frac{dy}{dx}$, et on envoie le dx du côté des x :

$$\frac{dy}{f(y)} = u(x)dx. \tag{32.51}$$

Maintenant il suffit de prendre l'intégrale des deux côtés : comme la position des dx et dy l'indiquent, il faut intégrer par rapport à y d'un côté et par rapport à dx de l'autre côté.

L'intégrale à gauche est facile : c'est $\ln(y)$. À droite, par contre, ça dépend tout à fait de u.

32.4.2 La méthode plus propre

$$y'(t) = u(t)f\big(y(t)\big). \tag{32.52}$$

Nous considérons U, une primitive de u sur I et G, une primitive de $1/f$ sur J. Si $I' \subseteq I$ et $y\colon I' \to J$, alors y est solution de (32.49) si et seulement si il existe une constante C telle que

$$G\big(y(t)\big) = U(t) + C. \tag{32.53}$$ EqSolSepThe

La recherche des solutions de l'équation différentielle se ramène donc à la recherche de primitives et de solutions d'une équation algébrique (il faut isoler $y(t)$ dans (32.53)). Réciproquement toute solution régulière de cette dernière relation est solution de l'équation différentielle.

Remarque : lorsque nous cherchons U et G, nous ne cherchons que *une* primitive. Il ne faut pas considérer des constantes d'intégration à ce niveau.

32.4.3 Les théorèmes

ProJLykrK

Proposition 32.12.
Soient des intervalles I et J de $\mathbb{R}$. Nous considérons des fonctions continues $u\colon I \to \mathbb{R}$ et $f\colon J \to \mathbb{R}$, et nous supposons que f ne s'annule pas sur J. Soit U, une primitive de u sur I, et G, une primitive de $1/f$ sur J.

Si I' est un intervalle, une application $y\colon I' \to \mathbb{R}$ est une solution de

$$y'(t) = u(t)f\big(y(t)\big) \tag{32.54}$$ EQooJVCQooGoVPVA

si et seulement si il existe $C \in \mathbb{R}$ telle que

$$G\big(y(t)\big) = U(t) + C. \tag{32.55}$$ EqSoluceEqDiffSep

Cette proposition dit que toutes les solutions qui ne s'annulent jamais sur un intervalle ont la forme $G\big(y(t)\big) = U(t) + C$ et peuvent donc être trouvées en calculant des primitives.

La formule (32.55) peut être obtenue de la façon heuristique suivante, en écrivant $y' = dy/dt$, et en passant le dt à droite. Nous trouvons successivement

$$\begin{aligned} y' &= u(t)f(y) \\ dy &= u(t)f(y)dt \\ \frac{dy}{f(y)} &= u(t)dt \\ \int \frac{dy}{f(y)} &= \int u(t)dt \\ G(y) &= U(t) + C. \end{aligned} \tag{32.56}$$

PROPooIGWTooULXrKI

Proposition 32.13.
Soient $0 < r < R$ ainsi que $\alpha \in \mathbb{R}$. Une fonction $y\colon]r, R[\to \mathbb{R}$ est une solution de

$$y'(x) = \frac{\alpha}{x}y(x), \tag{32.57}$$ EQooBQCTooTtShNw

si et seulement si il existe $K > 0$ tel que

$$y(x) = Kx^{\alpha}. \tag{32.58}$$

Démonstration. En deux parties.

(1) $\Rightarrow$ Vu que le logarithme est une primitive de la fonction inverse (proposition 15.85), nous pouvons utiliser la proposition 32.12 avec les fonctions $U(x) = \alpha \ln(x)$ et $G(t) = \ln(t)$. Si y est une solution de (32.57), alors il existe $C \in \mathbb{R}$ tel que

$$\ln\big(y(x)\big) = \alpha \ln(x) + C \tag{32.59}$$

pour tout $x \in]r, R[$. En prenant l'exponentielle des deux côtés, et en utilisant le fait que $\alpha \ln(x) = \ln(x^\alpha)$,

$$y(x) = e^C e^{\ln(x^\alpha)} \tag{32.60}$$

En posant $K = e^C$, nous avons un $K > 0$ tel que $y(x) = Kx^\alpha$.

(2) $\Leftarrow$ Il suffit de dériver en utilisant la proposition 14.294. Pas de soucis de domaine parce que nous ne considérons que $x \in]r, R[$.

□

PropOkmXmC

Proposition 32.14.
Soient u continue sur I et f continue sur J, et $f(\eta) \neq 0$ sur J. Soient $t_0 \in I$ et $y_0 \in J$. Alors il existe $I' \subset I$ avec $t_0 \in I'$ et $f \in C^1(I' \to J)$ tels que

(1) y est solution de (32.49) sur I' et vérifie $y(t_0) = y_0$,

(2) si z est une solution de (32.49) sur $I'' \subset I'$ avec $t_0 \in I''$ et $z(t_0) = y_0$, alors $I'' = I'$ et $z(t) = y(t)$ pour tout $t \in I''$.

ExYCPtxgZ

Exemple 32.15.
Résoudre l'équation différentielle

$$y - \cos(t)y' = \cos(t)\big(1 - \sin(t)\big)y^2. \tag{32.61}$$

La fonction $y = 0$ est solution. En posant $z = 1/y$, nous trouvons l'équation

$$z + \cos(t)z' = \cos(t)\big(1 - \sin(t)\big) \tag{32.62}$$

EDEqII107EqpourZ

à laquelle z doit satisfaire. L'équation homogène est

$$z'_H = -\frac{z_H}{\cos(t)}. \tag{32.63}$$

Ceci est une équation à variables séparées que nous résolvons en suivant les méthodes données plus haut : nous posons

$$\begin{aligned} u(t) &= \frac{1}{\cos(t)}, \\ f(z) &= -z, \\ U(t) &= \ln\left[\tan\left(\frac{\pi}{4} + \frac{t}{2}\right)\right] \quad \text{(voir formulaire)}, \\ G(z) &= \ln\left(\frac{1}{z}\right). \end{aligned} \tag{32.64}$$

EqEDufUGII107

La solution z_H est donnée par l'équation

$$\ln\left(\frac{1}{z}\right) = \ln\left[K \tan\left(\frac{\pi}{4} + \frac{t}{2}\right)\right], \tag{32.65}$$

c'est-à-dire

$$z_H(t) = \frac{K}{\tan\left(\frac{\pi}{4} + \frac{t}{2}\right)}. \tag{32.66}$$

Nous appliquons maintenant la méthode de variation des constantes sur cette solution afin de trouver la solution générale de l'équation (32.62). En utilisant la règle de Leibniz, $z' = K'z_H + Kz'_H$, nous trouvons

$$\frac{K}{\tan\left(\frac{\pi}{4} + \frac{t}{2}\right)} + \cos(t)\left(\frac{K'}{\tan\left(\frac{\pi}{4} + \frac{t}{2}\right)} - \frac{K}{2\sin^2\left(\frac{\pi}{4} + \frac{t}{2}\right)}\right) = \cos(t)\big(1 - \sin(t)\big). \tag{32.67}$$

Malgré leurs apparences, les deux termes en K se simplifient. En effet, en vertu de l'équation $z'_H = \frac{-z_H}{\cos(t)}$, nous avons

$$\frac{-K}{2\sin^2\left(\frac{\pi}{4}+\frac{t}{2}\right)} = \frac{-K}{\cos(t)\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)}. \tag{32.68}$$

Le travail de voir quel est le lien entre $\sin^2\left(\frac{\pi}{4}+\frac{t}{2}\right)$, $\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)$ et $\cos(t)$ est en réalité fait dans votre formulaire au moment où vous l'avez utilisé pour intégrer u pour obtenir le $U(t)$ de (32.64).

Après cette simplification durement méritée, nous trouvons l'équation suivante pour $K(t)$:

$$\frac{K'}{\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)} = 1-\sin(t). \tag{32.69}$$ EDEqFracII107exoVVprb

Résoudre cela revient à trouver la primitive de

$$\big(1-\sin(t)\big)\tan\left(\frac{\pi}{4}+\frac{t}{2}\right), \tag{32.70}$$

ce qui est relativement compliqué. La réponse est

$$\begin{aligned} K(t) = \ln\left(\sin\left(\frac{2\,x+\pi}{4}\right)+1\right) + \ln\left(\sin\left(\frac{2\,x+\pi}{4}\right)-1\right) \\ + 2\,\ln\sec\left(\frac{2\,x+\pi}{4}\right) + 2\,\sin^2\left(\frac{2\,x+\pi}{4}\right) \end{aligned} \tag{32.71}$$

Nous pouvons un peu simplifier en utilisant le fait que $\ln(a+b)+\ln(a-b)=\ln(a^2-b^2)$:

$$K(t) = \ln\left(-\cos^2\left(\frac{2\,x+\pi}{4}\right)\right) + 2\,\ln\sec\left(\frac{2\,x+\pi}{4}\right) + 2\,\sin^2\left(\frac{2\,x+\pi}{4}\right). \tag{32.72}$$

Il me semble toutefois qu'il faudrait prendre des valeurs absolues pour les logarithmes.

△

32.5 Équations linéaires d'ordre supérieur

32.5.1 Équations et systèmes linéaires à coefficients constants

SUBSECooMXLVooALNtge

Nous regardons l'équation

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = v(t) \tag{32.73}$$ EqLinConstantRappels

où les coefficients a_k sont maintenant des constantes. Il faut commencer par résoudre le polynôme caractéristique

$$r^n + a_1 r^{n-1} + \cdots + a_n = 0. \tag{32.74}$$

Si $\lambda_1,\ldots,\lambda_k$ sont les solutions avec multiplicité $\mu_1,\ldots,\mu_k$, alors le **système fondamental** de solutions linéairement indépendantes est l'ensemble suivant de solutions à l'équation homogène :

$$\begin{aligned} &e^{\lambda_1 t}, te^{\lambda_1 t},\ldots,t^{\mu_1-1}e^{\lambda_1 t} \\ &\vdots \\ &e^{\lambda_k t}, te^{\lambda_k t},\ldots,t^{\mu_k-1}e^{\lambda_k t}. \end{aligned} \tag{32.75}$$

Nous notons y_i ces solutions. La solution générale de l'équation homogène est donc donnée par

$$y_H = \sum_i c_i y_i. \tag{32.76}$$

Afin de trouver la solution générale de l'équation non homogène, nous appliquons la méthode de variation des constantes, en imposant les $n-1$ conditions

$$\sum_{i=1}^{n} c_i'(t)y_i^{(l)}(t) = 0 \tag{32.77}$$

avec $l = 0, \ldots, n-2$. Ces conditions plus l'équation de départ (32.73) forment un système de n équations différentielles pour les n fonctions inconnues $c_i(t)$.

Cette condition peut paraitre mystérieuse. Il est cependant encore possible de travailler sans poser la condition (32.77) en suivant la recette, en calculant des déterminants de Wronskien. Des exemples sont donnés dans les exercices sur le second ordre.

32.5.2 Si les coefficients ne sont pas constants ?

Une équation différentielle linéaire d'ordre n sur I est une équation de la forme

$$y^{(n)} + u_1(t)y^{(n-1)} + \cdots + u_{n-1}(t)y' + u_n(t)y = v(t) \tag{32.78}$$

où v et u_k sont des fonctions continues fixées de I vers $\mathbb{R}$.

Pour résoudre cette équation, il faut commencer par résoudre l'équation homogène correspondante (c'est-à-dire celle que l'on obtient en posant $v(t) = 0$). Ensuite, nous trouvons la solution de l'équation (32.78) en appliquant la méthode de la **variation des constantes**.

Donnons un exemple du pourquoi la méthode de variations des constantes est efficace. Soit l'équation

$$u' + f(t)u = g(t), \tag{32.79}$$

et disons que u_H est une solution de l'équation homogène. La méthode de variations des constantes consiste à poser $u(t) = K(t)u_H(t)$, et donc $u'(t) = K'u_H + Ku_H'$. En remettant dans l'équation de départ,

$$K'u_H + Ku_H' + fKu_H = g. \tag{32.80}$$

La somme $Ku_H' + fKu_H$ est nulle, par définition de u_H. Par conséquent, il ne reste que

$$K' = \frac{g(t)}{u_H}. \tag{32.81}$$

Lorsqu'on utilise la méthode de variation des constantes, nous trouvons toujours une simplification « miraculeuse ».

Dans l'immédiat, nous ne considérons que le cas où les u_i sont des constantes. Le cas où les u_i deviennent des fonctions de t sera vu plus tard.

32.6 Système d'équations linéaires

32.6.1 La magie de l'exponentielle...

Prenons l'équation différentielle très simple

$$y' = ay. \tag{32.82}$$

La solution est $y(t) = Ae^{at}$. Et si on a la donnée de Cauchy $y(t_0) = y_0$, alors

$$y(t) = Ae^{at}e^{-at_0}e^{at_0} = e^{a(t-t_0)}y(t_0). \tag{32.83}$$

Donc on a le facteur multiplicatif $e^{a(t-t_0)}$ qui sert à faire passer de $y(0)$ à $y(t)$. C'est un peu un opérateur d'évolution. Ce qui fait la magie de l'exponentielle, c'est son développement en série

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \tag{32.84}$$

qui est tel que chaque terme est la dérivée du terme suivant.

Maintenant, si on a un système

$$\bar{y}' = A\bar{y}, \tag{32.85}$$

il n'est pas du tout étonnant d'avoir comme solution $\bar{y}(t) = e^{At}$ où l'exponentielle de la matrice est définie exactement par la série (32.84). C'est un peu longuet, mais dans le cours, c'est effectivement ce qui est prouvé. La matrice résolvante $R(t, t_0)\colon \bar{y}_0 \to \bar{y}(t; t_0, y_0)$ est donnée par

$$R(t, t_0) = e^{(t-t_0)A}, \tag{32.86}$$

exactement comme dans l'équation (32.83).

32.6.2 ...mais la difficulté

Maintenant, il est suffisant de calculer des exponentielles de matrices pour résoudre des systèmes. Hélas, il est en général très difficile de calculer des exponentielles. Tu peux essayer de prouver les deux suivantes :

$$\begin{aligned} A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} &\rightsquigarrow e^A = \begin{pmatrix} \cos(a) & \sin(a) \\ -\sin(a) & \cos(a) \end{pmatrix} \\ S = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} &\rightsquigarrow e^S = \begin{pmatrix} \cosh(a) & \sinh(a) \\ \sinh(a) & \cosh(a) \end{pmatrix}. \end{aligned} \tag{32.87}$$

La première, tu vas la revoir si tu fais de la géométrie différentielle ou de la mécanique quantique : l'algèbre de Lie du groupe des matrices orthogonales de déterminant 1 est l'algèbre des matrices antisymétriques.

La seconde se retrouve en relativité parce que e^S est la matrice qui préserve $x^2 - y^2$, tout comme e^A préserve $x^2 + y^2$. Quelques mots sur l'uutilisation des fonctions hyperboliques en relativité dans 42.4.8.1.

32.6.3 La recette

Afin d'éviter de devoir calculer explicitement des exponentielles de matrices, nous faisons appel à toutes sortes de trucs, dont la forme de Jordan. Le résultat final est la méthode suivante. Soit le système homogène

$$\bar{y}' = A\bar{y}. \tag{32.88}$$

(1) D'abord, nous calculons les valeurs propres de A.

(2) Ensuite les vecteurs propres.

ItemRapSystDc

(3) Une bonne valeur propre, c'est une valeur propre dont l'espace propre a une dimension égale à sa multiplicité. C'est-à-dire que si λ est de multiplicité m, alors on a, dans les bons cas, m vecteur propres linéairement indépendants.

Dans ce cas, si $v_1, \dots, v_m$ sont les vecteurs, alors on a les solutions linéairement indépendantes suivantes :

$$\begin{pmatrix} \vdots \\ v_1 \\ \vdots \end{pmatrix} e^{\lambda t}, \dots, \begin{pmatrix} \vdots \\ v_m \\ \vdots \end{pmatrix} e^{\lambda t}. \tag{32.89}$$

Pour chaque bonne valeur propre, ça nous fait un tel paquet de solutions linéairement indépendantes.

(4) Si λ n'est pas une bonne valeur propre, alors les choses se compliquent. Mettons que λ ait k vecteurs propres en moins que sa multiplicité. Dans ce cas, il faut chercher des solutions sous la forme

$$\begin{pmatrix} a_1^{(k)} t^k + \dots + a_1^{(0)} \\ \vdots \\ a_n^{(k)} t^k + \dots + a_n^{(0)} \end{pmatrix} e^{\lambda t}. \tag{32.90}$$

EqEqRapAsTestPolk

C'est-à-dire qu'on prend comme coefficient de $e^{\lambda t}$, un vecteur de polynômes de degré k. Il faut mettre cela dans l'équation de départ pour voir quelles sont les contraintes sur les constantes $a_i^{(j)}$ introduites.

ItemRapSystDe

(5) Nous avons un cas particulier du cas précédent. Si λ est une valeur propre de multiplicité m qui n'a que un seul vecteur propre v, alors il faut chercher des polynômes de degré $m-1$, et on peut directement fixer le coefficient de t^{m-1}, ce sera l'unique vecteur propre :

$$\left[\begin{pmatrix}\vdots\\ v\\ \vdots\end{pmatrix}+\begin{pmatrix}a_1^{(m-2)}\\ \vdots\\ a_n^{(m-2)}\end{pmatrix}t^{m-2}+\dots\right]e^{\lambda t}. \tag{32.91}$$

Cela économise quelques calculs par rapport à poser brutalement (32.90).

32.6.4 Système d'équations linéaires avec matrice constante

Nous considérons l'équation différentielle

$$y'(t)=Ay(t) \tag{32.92}$$

EqOOsXZJ

pour la fonction $y\colon \mathbb{R}\to\mathbb{R}^n$ et A est une matrice ne dépendant pas de t. Nous supposons que A est diagonalisable pour les vecteurs propres v_i et les valeurs propres λ_i correspondantes.

La matrice

$$R(t)=\begin{bmatrix}e^{\lambda_1 t}v_1&\dots e^{\lambda_n t}v_n\end{bmatrix} \tag{32.93}$$

est la **matrice résolvante** du système. Alors la solution du système (32.92) pour la condition initiale $y(0)=y_0$ est

$$y(t)=R(t)y_0. \tag{32.94}$$

En effet

$$AR(t)=\left[A\begin{pmatrix}\uparrow\\ e^{\lambda_1 t}v_1\\ \downarrow\end{pmatrix}\dots A\begin{pmatrix}\uparrow\\ e^{\lambda_n t}v_n\\ \downarrow\end{pmatrix}\right]=R'(t). \tag{32.95}$$

Par conséquent $y'(t)=R'(t)y_0=AR(t)y_0=Ay(t)$.

32.6.5 Système d'équations linéaires avec matrice non constante

ThoNYEXqxO

Théorème 32.16 ([728]).
Soient I un intervalle de $\mathbb{R}$ et $M\colon\mathbb{R}\to\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ une fonction. Si les composantes M_{ij} sont des fonctions continues sur I alors :

(1) pour tout $t_0\in I$ et pour tout $y_0\in R^n$ le système

$$y'(t)=M(t)y(t) \tag{32.96}$$

EqKYDrMgv

admet une unique solution maximale définie sur I telle que $y(t_0)=y_0$;

(2) l'ensemble des solutions de l'équation (32.96) sur I est un espace vectoriel de dimension n.

32.7 Réduction de l'ordre

SecWGdleRM

Afin de diminuer l'ordre d'une équation dans laquelle le paramètre n'apparaît pas, il y a deux changements de variables très utiles. Le premier, le plus simple, est simplement de poser $z(t)=y'(t)$, ce qui donne

$$z'(t)=y''(t). \tag{32.97}$$

Le second, *qui n'est pas le même*, est $z\big(y(t)\big)=y'(t)$, qui entraîne

$$y''(t)=z'\big(y(t)\big)y'(t)=z'\big(y(t)\big)z\big(y(t)\big). \tag{32.98}$$

Dans ce second cas, il faut également changer de variable, et utiliser $y(t)$ comme variable au lieu de t.

Si ça ne marche pas, il faut suivre la procédure ci-après.

Nous supposons avoir une équation différentielle d'ordre p dans laquelle $y^{(p)}$ est isolée des autres dérivées :

$$y^{(p)}(t) = f\big(t, y(t), y'(t), \dots, y^{(p-1)}(t)\big) \tag{32.99}$$

où f est une fonction $f\colon \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}$, et la fonction cherchée est $y\colon \mathbb{R} \to \mathbb{R}$.

La méthode proposée ici consiste à transformer cette équation d'ordre p en un système d'équations d'ordre 1. Pour cela nous posons

$$\begin{aligned} F\colon \mathbb{R} \times \mathbb{R}^p &\to \mathbb{R}^p \\ (t, x) &\mapsto \begin{pmatrix} x_2 \\ \vdots \\ x_p \\ f(t, x_1, \dots, x_p) \end{pmatrix}. \end{aligned} \tag{32.100}$$

Nous considérons alors l'équation différentielle

$$Y'(t) = F\big(t, Y(t)\big). \tag{32.101}$$

pour $Y\colon \mathbb{R} \to \mathbb{R}^p$. La fonction $y = Y_1$ résout l'équation (32.99) si et seulement si la fonction Y résout l'équation (32.101).

De plus si l'équation (32.99) est donnée avec les conditions initiales $y^{(k)} = a_k$ $(k = 0, \dots, p-1)$ alors l'équation (32.101) vient avec les conditions initiales

$$Y(t_0) = \begin{pmatrix} a_0 \\ \vdots \\ a_{p-1} \end{pmatrix}, \tag{32.102}$$

c'est-à-dire $Y(t_0) = A_0$ avec $A_0 \in \mathbb{R}^p$.

Le théorème de Cauchy-Lipschitz 17.39 nous donne existence et unicité locale de la solution au système 32.101. Lorsque le système est linéaire, c'est-à-dire sous la forme $Y'(t) = M(t)Y(t)$, alors il y a mieux : le théorème 32.16.

Exemple 32.17.
Nous considérons l'équation différentielle

$$\begin{cases} -u''(t) - u(t) = 1 & \text{(32.103a)} \\ u(0) = a_0 \in \mathbb{R} & \text{(32.103b)} \\ u'(0) = a_1 \in \mathbb{R}, & \text{(32.103c)} \end{cases}$$

et nous voulons montrer que ce système accepte une unique solution. Vu que l'équation différentielle se présente sous la forme $u'' = f(t, u)$ avec $f(t, u) = -1 - u$ nous posons

$$\begin{aligned} F\colon \mathbb{R} \times \mathbb{R}^2 &\to \mathbb{R}^2 \\ (t, x) &\mapsto \begin{pmatrix} x_2 \\ -1 - x_1 \end{pmatrix}, \end{aligned} \tag{32.104}$$

et nous considérons l'équation différentielle

$$Y' = F(t, Y) \tag{32.105}$$

pour la fonction $Y\colon \mathbb{R} \to \mathbb{R}^2$. La fonction F est Lipschitz (et même globalement) par rapport à Y. En effet,

$$\|F(x) - F(y)\|^2 = (x_2 - y_2)^2 + (x_1 - y_1)^2 = \|x - y\|^2. \tag{32.106}$$

Le théorème de Cauchy-Lipschitz 17.39 s'applique et en posant $Y_0 = (a_0, a_1)$, il existe une unique solution à l'équation $Y' = F(t, Y)$ vérifiant $Y(0) = Y_0$. Nous notons $t \mapsto Y(t)$ cette solution.

En quoi cela nous aide ? Nous posons $u(t) = Y_1(t)$. Alors

$$u'(t) = Y_1'(t) = F_1'(t, Y) = Y_2(t). \tag{32.107}$$

En dérivant encore,

$$u''(t) = Y_2'(t) = F_2'(t, Y) = -1 - Y_1(t) = -1 - u(t), \tag{32.108}$$

ce qu'il fallait. La fonction $t \mapsto Y_1(t)$ est solution de notre équation de départ. Quid des conditions initiales ? Vu que $u = Y_1$ et $u' = Y_2$ nous avons [2]

$$u(0) = Y_1(0) = (Y_0)_1 = a_0 \tag{32.109}$$

et

$$u'(0) = (Y_0)_2 = a_1. \tag{32.110}$$

Toutes les prescriptions sont respectées.

Si vous voulez vraiment résoudre cette équation, il faudra plus de travail. D'abord résoudre l'équation homogène associée, c'est-à-dire l'équation caractéristique $r^2 + 1 = 0$, ce qui va donner

$$u_H(t) = Ae^{it} + Be^{-it}, \tag{32.111}$$

et ensuite faire le coup de la variation des constantes pour déterminer la solution générale du problème non homogène. △

EXooJNOMooYqUwTZ

Exemple 32.18.
Nous reprenons l'équation différentielle

EQooGUYBooQUSti

$$-u''(t) - u(t) = 1. \tag{32.112}$$

Nous avons déjà vu dans l'exemple 32.17 que cette équation avait une solution unique pour toute condition initiale. Cette fois nous voulons étudier les solutions lorsque nous imposons les conditions aux limites $u(0) = u(\pi) = 0$. Nous allons voir qu'il n'y a pas de telles solutions.

Pour ce faire, soit une solution u. D'abord u'' existant, la fonction u est de classe au moins C^1. Mais $u'' = -1 - u$, donc u'' est également C^1, ce qui donne la régularité C^3 pour u. En continuant ainsi nous trouvons que u est de classe C^∞.

Le truc est de considérer la fonction $v(t) = \sin(t)$ qui vérifie l'équation différentielle

$$\begin{cases} -v'' - v = 0 & \text{(32.113a)} \\ v(0) = 0 & \text{(32.113b)} \\ v'(0) = 1. & \text{(32.113c)} \end{cases}$$

Nous calculons le produit scalaire sur $L^2\big(]0, \pi[\big)$ de (32.112) avec v :

EQooQWGEooFbsVzq

$$\langle u'', v\rangle + \langle u, v\rangle = -\langle v, 1\rangle. \tag{32.114}$$

Le calcul de $\langle v, 1\rangle$ est simplement l'intégrale de $\sin(t)$ pour t allant de 0 à π, c'est-à-dire $\langle v, 1\rangle = 2$. Vu que u et v sont toutes deux des fonctions qui s'annulent en 0 et en π nous pouvons faire des intégrations par partie les yeux fermés et exprimer $\langle u'', v\rangle$ sans dérivées sur u :

$$\langle u'', v\rangle = -\langle u', v'\rangle = \langle u, v''\rangle = -\langle u, v\rangle \tag{32.115}$$

où la dernière égalité n'est autre que le fait que $v = \sin$, donc $v'' = -v$. Le membre de gauche de (32.114) vaut donc zéro alors que celui de droite vaut -2.

Nous concluons que le problème aux limites posé n'admet pas de solutions. △

2. Dans les équations suivantes, il faut faire attention aux notations : Y_0 est le vecteur Y_0, tandis que Y_1 est la première composante de la fonction Y.

32.8 Autour de Cauchy-Lipschitz

SECooNKICooDnOFTD

Dans cette section nous étudions les équations différentielles du type

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & (32.116a)\\ y(t_0) = y_0. & (32.116b)\end{cases}$$

32.8.1 Fuite des compacts et explosion en temps fini

Théorème 32.19 (Fuite des compacts[729, 507])**.**
Nous considérons l'équation différentielle

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & (32.117a)\\ y(t_0) = y_0, & (32.117b)\end{cases}$$

où $f\colon I \times \Omega \to \mathbb{R}^n$ est continue et Ω ouvert dans $\mathbb{R}^n$. Soit la solution maximale $y_M\colon J_M =]t_{min}, t_{max}[\to \Omega$. Si $t_{max} < \sup(I)$ alors $y_M(t)$ sort de tout compact de Ω lorsque $t \to t_{max}$.

Démonstration. Soit K un compact de Ω et nous considérons une suite (t_m) dans $]t_{min}, t_{max}[$ telle que $t_m \to t_{max}$. Si nous supposons que $y_M(t)$ ne sort pas de K alors nous avons $y_M(t_m) \in K$, c'est-à-dire une suite dans un compact. Quitte à passer à une sous-suite, nous supposons qu'elle est convergente. Soit $x_1 \in K$ la limite $\lim_{m\to\infty} y_M(t_m) = x_1$.

Vu que $t_{max} \in I$, la condition initiale $y(t_{max}) = x_1$ est valide et le théorème de Cauchy-Lipschitz 17.39 nous donne une unique solution maximale y_P définie sur un ouvert J_P autour de t_{max}.

Nous allons maintenant construire une solution au problème initial qui contredit la maximalité de y_M. Attention : il n'est pas évident à priori que $y_P(t) = y_M(t)$ sur l'intersection des domaines. Si c'était évident, la proposition serait démontrée.

Soit $\tilde{J} = J_M \cup J_P \cap]t_{min}, +\infty[$ et la fonction

$$\tilde{y}(t) = \begin{cases} y_M(t) & \text{si } t < t_{max}\\ y_P(t) & \text{si } t \geqslant t_{max}.\end{cases} \tag{32.118}$$

La fonction $\tilde{y}$ est continue par construction parce que

$$\lim_{t\to t_{max}} y_M(t) = x_1 = y_P(t_{max}). \tag{32.119}$$

Nous vérifions à présent que $\tilde{y}$ est une solution : $\tilde{y}'(t_{max}) = f\big(t_{max}, y(t_{max})\big)$:

$$\begin{aligned}\lim_{\epsilon\to 0} \frac{\tilde{y}(t_{max}-\epsilon) - \tilde{y}(t_{max})}{\epsilon} &= \lim_{\epsilon\to 0} \frac{y_M(t_{max}-\epsilon) - y_P(t_{max})}{\epsilon} && (32.120a)\\ &= \lim_{\epsilon\to 0} \frac{y_M(t_{max}-\epsilon) - y_P(t_{max}-\epsilon) + y_P(t_{max}-\epsilon) - y_P(t_{max})}{\epsilon} && (32.120b)\\ &= \lim_{\epsilon\to 0} \frac{y_P(t_{max}-\epsilon) - y_P(t_{max})}{\epsilon} && (32.120c)\\ &= y'_P(t_{max}). && (32.120d)\end{aligned}$$

Donc $\tilde{y}$ est solution pour la condition initiale $\tilde{y}(t_{max}) = x_1$ et coïncide avec y_P en t_{max} et avec y_M avant t_{max}. Donc en réalité y_P, y_M et $\tilde{y}$ sont identiques et cela contredit la maximalité de y_M. □

CorGDJQooNEIvpp

Corolaire 32.20 (Explosion en temps fini)**.**
Soit (y_m, J) la solution maximale du problème de Cauchy (17.146) :

$$\begin{cases} y' = f(t, y) & (32.121a)\\ y(t_0) = y_0, & (32.121b)\end{cases}$$

avec $f\colon U = I \times \Omega \to \mathbb{R}^n$ où I est ouvert dans $\mathbb{R}$ et Ω ouvert dans $\mathbb{R}^n$. Nous supposons que f est continue sur U et localement Lipschitz par rapport à y.

Si la solution maximale est définie sur $J =]t_{min}, t_{max}[$ alors nous avons l'alternative suivante :
ItemOLYYooJVkRfj

(1) Soit $t_{max} = \sup(I)$,
ITEMooUKFAooXwRNSB

(2) soit $t_{max} < \sup(I)$ et $\lim_{t\to t_{max}} \|y(t)\| = \infty$.

Le résultat tient aussi mutatis mutandis *pour $t_{\min}$.*

Remarque 32.21.
Attention : ceci n'est pas une simple paraphrase de la fuite des compacts. L'information supplémentaire que ce corolaire donne est que la solution sort de tout compact *pour ne plus y retourner.*

Démonstration. L'hypothèse $t_{max} < \sup(I)$ signifie que la solution finit d'exister avant que les hypothèses sur f cessent d'être vraies. C'est-à-dire que la solution maximale est moindre que ce que nous aurions pu espérer.

Soit un compact K. Supposons que que pour tout $t_0 < t_{max}$ il existe $t \in]t_0, t_{max}[$ tel que $y_M(t) \in K$. Alors cela crée une suite t_k dans J telle que $y_M(t_k)$ est dans K. Comme dans le théorème de la fuite des compacts nous concluons l'impossibilité de la chose.

Donc pour tout compact K de Ω, il existe $T < t_{max}$ tel que $y_M(t) \in \Omega\backslash K$ pour tout $t \in [T, t_{max}[$. En prenant des boules fermées de plus en plus grandes en guise de compacts nous concluons que

$$\lim_{t\to t_{max}} \|y_M(t)\| = \infty. \tag{32.122}$$

□
NORMooZROGooZfsdnZ

32.22.
Notons que si $t_{max} < \infty$, si nous sommes dans l'alternative 32.20(2) et si la solution maximale y est de classe C^1 (ce qui est le cas lorsqu'on utilise Cauchy-Lipschitz 17.39) alors la dérivée de y est également non bornée dans un voisinage de t_{max}.

Mais si f est globalement bornée, alors dans l'équation $y' = f(t, y)$, la dérivée y' sera globalement bornée. Dans ce cas, la solution ne peut pas exploser en temps fini et existe donc globalement.

¡¡ Avertissement/question au lecteur !! 32.23
Êtes-vous d'accord avec 32.22 ?

Exemple 32.24.
Soit l'équation différentielle

$$\begin{cases} y' = y(y-1)\sin(yt) & \text{(32.123a)} \\ y(0) = \dfrac{1}{2}. & \text{(32.123b)} \end{cases}$$

La fonction $f(t, y) = y(y-1)\sin(yt)$ ayant une dérivée bornée partout, elle est localement Lipschitz et le théorème de Cauchy-Lipschitz 17.39 s'applique. Pour toute condition initiale, une solution maximale unique existe.

Si nous oublions la condition initiale, il est facile de trouver des solutions constantes : $y' = 0$ avec $y(t) = k$ donne l'équation

$$0 = k(k-1)\sin(kt). \tag{32.124}$$

Les solutions $y_1(t) = 0$ et $y_2(t) = 1$ sont des solutions existant pour tout t.

Le graphe de la solution correspondante à la condition initiale $y(0) = \frac{1}{2}$ ne pouvant pas croiser les graphes de y_1 et y_2, elle est obligée d'exister pour tout t parce qu'elle ne peut pas exploser en temps fini. △

32.8.2 Écart entre deux conditions initiales

PROPooOPRRooQgYFDk

Proposition 32.25 ([727, 730]).
Soit une fonction $f\colon I\times\Omega\to\mathbb{R}^n$ continue et globalement Lipschitz en sa seconde variable (Ω est un ouvert de $\mathbb{R}^n$). Soient deux solutions $y_1\colon I_1\to\mathbb{R}^n$ et $y_2\colon I_2\to\mathbb{R}^n$ aux problèmes

$$\begin{cases} y_i'(t)=f\big(t,y_i(t)\big) & (32.125a)\\ y_i(t_0)=a_i. & (32.125b)\end{cases}$$

Alors pour tout $t\in I_1\cap I_2$ nous pouvons estimer l'écart entre y_1 et y_2 par la formule

$$\|y_1(t)-y_2(t)\|\leqslant e^{L|t-t_0|}\|a_1-a_2\|, \tag{32.126}$$

où L est la constante de Lipschitz de f.

Démonstration. Nous avons d'abord les majorations suivantes, qui semblent juste jouer avec les notations, mais qui utilisent le fait (contenu dans le théorème de Cauchy-Lipschitz) que y_i soit de classe C^1 :

$$\begin{aligned}\|y_1(t)-y_2(t)\|&=\|\int_{t_0}^t\big(y_1'(s)-y_2'(s)\big)\|ds & (32.127a)\\ &\leqslant\int_{t_0}^t\|f\big(s,y_1(s)\big)-f\big(s,y_2(s)\big)\|ds & (32.127b)\\ &\leqslant L\int_{t_0}^t\|y_1(s)-y_2(s)\|ds. & (32.127c)\end{aligned}$$

C'est à ce moment que nous utilisons le lemme de Grönwall. Vu que

$$\|y_1(t)-y_2(t)\|\leqslant L\int_{t_0}^t\|y_1(s)-y_2(s)\|ds, \tag{32.128}$$

nous sommes dans les hypothèses de Grönwall 32.6 en posant

$$\begin{aligned}u(t)&=\|y_1(t)-y_2(t)\| & (32.129a)\\ b(t)&=\|y_1(0)-y_2(0)\| & (32.129b)\\ a(t)&=L. & (32.129c)\end{aligned}$$

Nous avons la majoration

$$\|y_1(t)-y_2(t)\|\leqslant\|y_1(0)-y_2(0)\|+L\int_0^t\|y_1(0)-y_2(0)\|e^{L(t-s)}ds. \tag{32.130}$$

Le calcul de l'intégrale intérieure donne

$$\int_0^t e^{L(t-s)}ds=-\frac{1}{L}(e^{-Lt}-1). \tag{32.131}$$

Avec ça, nous avons

$$\|y_1(t)-y_2(t)\|\leqslant e^{Lt}\|y_1(0)-y_2(0)\|. \tag{32.132}$$

□

32.26.
Notons que la proposition 32.25 est plutôt une mauvaise nouvelle parce que les solutions restent seulement linéairement proches l'une de l'autre lorsqu'on rapproche les conditions initiales, mais elle divergent exponentiellement vite avec le temps. Donc deux trajectoires arbitrairement proches au départ finissent assez vite par être bien séparées.

Cette proposition est cependant cruciale parce qu'elle explique que pour des petits t, les solutions ne s'écartent pas beaucoup, c'est-à-dire que pour t fixé, l'application qui à une donnée initiale fait correspondre la solution en t est continue. C'est le premier pas pour parler de régularité du flot.

32.8.3 Flot d'un champ de vecteurs

Nous reprenons l'équation différentielle du théorème de Cauchy-Lipschitz 17.39. En ce qui concerne les notations, I est un intervalle ouvert de $\mathbb{R}$ contenant 0 et l'application $f\colon I\times\mathbb{R}^n\to\mathbb{R}^n$ est continue et localement Lipschitz en sa seconde variable. Pour $a\in\mathbb{R}^n$, nous notons (J_a,y_a) la solution maximale (donc $y_a\colon J_a\to\mathbb{R}^n$) du problème [EQooADJMooRAZWfm]

$$\begin{cases} y_a'(t)=f\big(t,y_a(t)\big) & (32.133a)\\ y_a(0)=a. & (32.133b)\end{cases}$$

Nous noterons aussi de temps en temps $\varphi(t,a)=y_a(t)$.

Nous savons que $t\mapsto y_a(t)$ est de classe C^1, et cela est directement dans le théorème de Cauchy-Lipschitz. Une question d'une toute autre difficulté est la régularité de $a\mapsto y_a(t)$ pour t fixé, et encore pire : celle de $(t,a)\mapsto y_a(t)$.

Il se fait que l'application $(t,a)\mapsto y_a(t)$ a la même régularité que celle de f, mais cela va être un peu long à prouver. En ce qui concerne la régularité C^1, ce sera le théorème 32.33 dont la démonstration, comme vous pouvez le voir sera copieuse et demandera des propositions intermédiaires pas simples.

Définition 32.27.
Si t est fixé, l'application

$$\begin{aligned}\varphi_t\colon\mathbb{R}^n&\to\mathbb{R}^n\\ x&\mapsto\varphi(t,x)=y_x(t),\end{aligned}\tag{32.134}$$

*est le **flot** du problème de Cauchy (32.133).*

L'application $t\mapsto\varphi_t$ est ce qui est appelé le groupe à un paramètre de flot, pour des raisons qui arriverons plus tard[3]*.*

Le but est d'étudier les propriétés du flot : est-il continu, un difféomorphisme, existe, pour quels t ? Où se cache le champ de vecteurs du titre dans l'équation différentielle ?

Nous posons

$$\mathcal{D}=\bigcup_{x\in\Omega}\big(J_x\times\{x\}\big).\tag{32.135}$$

Comme tout produit d'espaces métrique, l'ensemble $\mathcal{D}$ est muni d'une métrique via la définition 7.225.

[PROPooUDQWooNFrNOQ]

Proposition 32.28 ([727])**.**
Soit un intervalle I ouvert de $\mathbb{R}$ contenant 0 et Ω un ouvert connexe de $\mathbb{R}^n$. Soit une application $f\colon I\times\mathbb{R}^n\to\mathbb{R}^n$ continue et localement Lipschitz en sa seconde variable. Pour $a\in\mathbb{R}^n$, nous notons (J_a,y_a) la solution maximale (donc $y_a\colon J_a\to\mathbb{R}^n$) du problème

$$\begin{cases} y_a'(t)=f\big(t,y_a(t)\big) & (32.136a)\\ y_a(0)=a. & (32.136b)\end{cases}$$

Nous posons

$$\mathcal{D}=\bigcup_{x\in\Omega}\big(J_x\times\{x\}\big).\tag{32.137}$$

Nous définissons la fonction φ par $\varphi(t,x)=y_x(t)$ là où ça existe.

L'ensemble $\mathcal{D}$ est ouvert. L'application $\varphi\colon\mathcal{D}\to\Omega$ est localement Lipschitz.

Démonstration. Soit $(s,a)\in\mathcal{D}$ et (J_a,y_a) la solution maximale passant par a en $t=0$. Par définition de $\mathcal{D}$ nous avons $s\in J_a$. Nous considérons J, un compact inclus dans J_a et contenant 0 et s en son intérieur. Nous posons

$$K=J\times y_a(J).\tag{32.138}$$

3. ou pas...

Vu que y_a est continue, cela est un compact. Chaque point de K possède un voisinage ouvert sur lequel f est Lipschitz [4] ; nous considérons un sous recouvrement fini et le maximum des constantes de Lipschitz. Cela nous crée un voisinage V de K dans $I \times \Omega$ dans lequel f est Lipschitz.

Vu que V est ouvert et K est compact avec $K \subset V$, nous pouvons trouver un ouvert V' et un compact K' tels que

$$K \subset V' \subset K' \subset V. \tag{32.139}$$

Sur ce V', la fonction f est de plus bornée parce que continue sur le compact K'. Nous renommons V' en V. Sur V nous avons :

— $\|f\|_{\infty,V} \leqslant M$,

— f est Lipschitz en sa seconde variable, de constante de Lipschitz L.

En tant qu'espace produit, nous avons une distance sur $I \times \Omega$ donnée en 7.225 :

$$d\big((t,y),(t',y')\big) = \max\big\{|t-t'|, \|y-y'\|\big\}. \tag{32.140}$$

Nous posons

$$V_\epsilon(K) = \{z \in I \times \Omega \text{ tel que } d(z,K) < \epsilon\} \tag{32.141a}$$

$$W_\epsilon = \{(t,y) \in J \times \Omega \text{ tel que } \|y - y_a(t)\| < \epsilon\}. \tag{32.141b}$$

(1) $\overline{W_\epsilon} \subset \overline{V_\epsilon(K)}$

Soit $(t,y) \in W_\epsilon$. Nous avons :

$$d\big((t,y),K\big) = \inf_{(t',y')\in K} d\big((t,y),(t',y')\big) \tag{32.142a}$$

$$= \inf_{(t',y')\in K} \max\{|t-t'|, \|y-y'\|\}. \tag{32.142b}$$

SUBEQooUCZXooRpAitk

Mais demander $(t,y) \in W_\epsilon$ signifie que $t \in J$ et $\|y - y_a(t)\| \leqslant \epsilon$. Dans K nous avons l'élément $\big(t, y_a(t)\big)$ qui vérifie

$$d\big((t,y),(t,y_a(t))\big) = \|y - y_a(t)\| \leqslant \epsilon. \tag{32.143}$$

Donc l'infimum de (32.142b) est majoré par ϵ. Nous avons prouvé que $W_\epsilon \subset V_\epsilon(K)$ et donc même inclusion pour les fermetures.

(2) **Il existe $\epsilon > 0$ tel que $\overline{V_\epsilon(K)} \subset V$**

Supposons que $\overline{V_\epsilon(K)}$ ne soit inclus dans V pour aucun ϵ. Alors nous considérons

$$z_n \in \overline{V_{1/n}(K)} \backslash V. \tag{32.144}$$

Nous avons par définition $d(z_n,K) \leqslant \frac{1}{n}$. Vu que K est compact, il comprend (au moins) un élément réalisant la distance : soit $z'_n \in K$ tel que

$$d(z_n, z'_n) = d(z_n, K). \tag{32.145}$$

Nous avons $d(z_n,z'_n) \to 0$, de telle sorte que les valeurs d'adhérence de (z_n) et (z'_n) sont les mêmes. Et comme (z'_n) est une suite dans un compact, elle a des valeurs d'adhérence. Soit z_∞ l'une d'elles. Vu que c'est une valeur d'adhérence d'une suite contenue dans le compact K, elle est également dans K : $z_\infty \in K$. Mais en même temps, z_n est hors de l'ouvert V, et donc dans le fermé V^c. Les valeurs d'adhérences restent dans le fermé, c'est-à-dire $z_\infty \notin V$. Vu que $K \subset V$, il y a contradiction.

Donc il existe $\epsilon > 0$ tel que $\overline{V_\epsilon(K)} \subset V$.

(3) **Il existe ϵ tel que $\overline{W_\epsilon} \subset V$**

Il suffit de prendre le ϵ dont nous venons de parler pour avoir

$$\overline{W_\epsilon} \subset \overline{V_\epsilon(K)} \subset V. \tag{32.146}$$

4. Cela est à peu près la définition d'être localement Lipschitz : 12.319, voir aussi 12.320.

Soit le ϵ en question, et $T > 0$ tel que $J \subset [-T, T]$. Nous posons $r = \epsilon e^{-LT}$. Soit $b \in \overline{B(a,r)}$ et

$$X = \{\tau \in J_+ \text{ tel que }]0,\tau] \subset J_b \text{ et } \big(t, y_b(t)\big) \in \overline{W_\epsilon}\, \forall\, t \in [0,\tau]\}. \tag{32.147}$$

Nous allons prouver que $X = J_+$ en prouvant qu'il est ouvert, fermé et non vide dans $J_+ = J \cap]0,\infty[$. Nous parlons bien de la topologie de J_+, celle induite[5] de $\mathbb{R}$. Vu que $0 \in J$, l'ensemble J_+ est ouvert à gauche, mais comme il est compact, il ne va certainement pas jusqu'à $+\infty$, de telle sorte qu'il est fermé à droite. Les ouverts de J_+ sont les ensembles de la forme $\mathcal{O} \cap J_+$ où $\mathcal{O}$ est ouvert de $\mathbb{R}$. Il y en a de la forme $]0, m]$.

(1) **X est fermé**
C'est parce que $\overline{W_\epsilon}$ et J_b sont fermés.

(2) **X est ouvert**
Soit $\tau \in X$. Si $\tau = \sup J_+$ alors $X = J_+$ est un ouvert de J_+. Supposons donc que $0 < \tau < \sup J_+$. Dans ce cas nous avons

$$\big(\tau, y_b(\tau)\big) \in \overline{W_\epsilon} \subset V, \tag{32.148}$$

et nous pouvons résoudre localement le problème de Cauchy

$$\begin{cases} y'(t) = f\big(t, y(t)\big) & \text{(32.149a)} \\ y(\tau) = \varphi(\tau, b) = y_b(\tau). & \text{(32.149b)} \end{cases}$$

Ce y existe jusqu'à $\tau + \eta$ (pour au moins un petit η), et par l'unicité de la solution, $y = y_b$ sur $[\tau, \tau + \eta[$. Ceci pour dire que le flot $\varphi(.,b)$ existe au moins jusqu'à $\tau + \eta$.
Grâce à la proposition 32.25 nous pouvons évaluer

$$\|\varphi(\tau, b) - \varphi(\tau, a)\| = \|y_a(\tau) - y_b(\tau)\| \leqslant e^{L\tau}\|b - a\|. \tag{32.150}$$

Comme nous avions choisi $r = \epsilon e^{-LT}$ et $b \in \overline{B(a,r)}$ nous avons aussi $\|b - a\| \leqslant \epsilon e^{-LT}$ et donc

$$\|\varphi(\tau, b) - \varphi(\tau, a)\| \leqslant \epsilon e^{L(\tau - T)} < \epsilon \tag{32.151}$$

parce que nous avions $\tau < \sup J_+ \leqslant T$, ce qui garantit que $e^{L(\tau-T)} < 1$.
Est-ce que ceci nous garantit que $\tau + \eta \in X$? Il faudrait $\big(\tau + \eta, y_b(\tau + \eta)\big) \in \overline{W_\epsilon}$, c'est-à-dire $\|y_b(\tau + \eta) - y_a(\tau + \eta)\| \leqslant \epsilon$. L'ensemble J_+ étant fermé dans l'ouvert J_a, ce dernier déborde certainement. Prenons donc η assez petit pour que y_a existe jusqu'en $\tau + \eta$.
Vu que y_a et y_b sont continues, et qu'en τ elles sont distantes de moins de ϵ, en $\tau + \eta$, elles restent distantes de moins de ϵ (quitte à prendre encore η plus petit).
Ceci nous permet de conclure que X est ouvert.

(3) **X est non vide**
La solution y_b au problème

$$\begin{cases} y_b'(t) = f\big(t, y_b(t)\big) & \text{(32.152a)} \\ y_b(0) = b & \text{(32.152b)} \end{cases}$$

existe au moins localement et vérifie $\|y_b(0) - y_a(0)\| = \|b - a\| \leqslant \epsilon e^{-LT} < \epsilon$. Par continuité nous avons

$$\|y_b(t) - y_a(t)\| < \epsilon \tag{32.153}$$

pour tout t dans un voisinage de 0. Donc X est non vide.

(4) **Conclusion pour X**
La partie X est ouverte, fermée et non vide dans J_+ qui est connexe. Donc $X = J_+$ par la proposition 7.74(4).

5. Définition 7.26.

La conclusion $X = J_+$ nous enseigne que pour tout $t \in J_+$ nous avons $]0,t] \subset J_b$ et $\big(t, y_b(t)\big) \in \overline{W_\epsilon}$. Nous pouvons faire la même chose pour J_- et au final nous avons que pour tout $\tau \in J$ nous avons d'abord $\tau \in J_b$, ce qui prouve $J \subset J_b$. De plus pour tout $t \in J$ nous avons aussi

$$\big(t, y_b(t)\big) \in \overline{W_\epsilon} \subset V. \tag{32.154}$$

Nous en concluons que

$$J \times \overline{B(a,r)} \subset V. \tag{32.155}$$

Nous savons de plus que pour tout $b \in \overline{B(a,r)}$, $J \subset J_b$. Cela signifie que

$$J \times \overline{B(a,r)} \subset \mathcal{D}. \tag{32.156}$$

Mais $J \times \overline{B(a,r)}$ est un voisinage de (s,a) qui était au début de la preuve un point générique choisi dans $\mathcal{D}$. Donc $\mathcal{D}$ est ouvert parce qu'il contient un voisinage de chacun de ses points.

Il nous reste à voir que $\varphi\colon \mathcal{D} \to \Omega$ est localement Lipschitz. Soit donc le point générique (s,a) dans $\mathcal{D}$ et l'ensemble V qui avait été construit plus haut. Nous allons montrer que φ est Lipschitz sur $J \times \overline{B(a,r)} \subset V$. D'abord sur V, l'application f est Lipschitz, donc

$$\|\varphi(t,b_1) - \varphi(t,b_2)\| \leqslant e^{Lt}\|b_1 - b_2\| \leqslant e^{LT}\|b_1 - b_2\| \tag{32.157}$$

pour tout $t \in J$ et $b_1, b_2 \in \overline{B(a,r)}$.

Ensuite, f est bornée, majorée par M sur V, donc

$$\begin{aligned}
\|\varphi(t_1,b) - \varphi(t_2,b)\| &= |\int_{[t_1,t_2]} y'(s)ds| && \text{(32.158a)}\\
&= |\int_{[t_1,t_2]} f\big(s, y_b(s)\big)ds| && \text{(32.158b)}\\
&\leqslant \int_{[t_1,t_2]} |f\big(s, y_b(s)\big)|ds && \text{(32.158c)}\\
&\leqslant M|t_1 - t_2|. && \text{(32.158d)}
\end{aligned}$$

Et enfin nous prouvons que φ est localement Lipschitz. En posant $k = \max\{e^{LT}, M\}$ nous avons

$$\begin{aligned}
\|\varphi(t_1,b_1) - \varphi(t_2,b_2)\| &\leqslant \|\varphi(t_1,b_1) - \varphi(t_1,b_2)\| + \|\varphi(t_1,b_2) - \varphi(t_2,b_2)\| && \text{(32.159a)}\\
&\leqslant e^{LT}\|b_1 - b_2\| + M|t_1 - t_2| && \text{(32.159b)}\\
&\leqslant k\big(\|b_1 - b_2\| + |t_1 - t_2|\big) && \text{(32.159c)}\\
&\leqslant 2k \max\{\|b_1 - b_2\|, |t_1 - t_2|\} && \text{(32.159d)}\\
&= 2kd\big((b_1,t_1),(b_2,t_2)\big). && \text{(32.159e)}
\end{aligned}$$

Le flot φ est donc Lipschitz de constante $2k$. □

Exemple 32.29.

¡¡ Avertissement/question au lecteur !! 32.30
Cet exemple doit être lu attentivement. Il me semble prouver que le flot n'est pas dérivable en la condition initiale sans que f le soit. Le document [731] semble dire le contraire. Je ne suis pas assez sûr de mon coup pour contredire.

Il n'y a pas de raisons de penser que $a \mapsto y_a(t)$ soit mieux que continue sans hypothèses supplémentaires sur f. Pour illustrer cela nous considérons l'équation différentielle

$$\begin{cases}
\dfrac{\partial X}{\partial s} = f\big(X(s), s\big) & \text{(32.160a)}\\
X(t) = x & \text{(32.160b)}
\end{cases}$$

où t et x sont des paramètres fixés. Nous allons étudier la dérivabilité de X en x lorsque

$$f(x,t) = |x|. \tag{32.161}$$

Cela est un exemple typique de fonction autant Lipschitz que l'on veut sans être dérivable. L'équation différentielle est

$$\frac{\partial X}{\partial s}(s) = |X(s)|. \tag{32.162}$$

Si $x > 0$ alors $X(s) > 0$ dans un voisinage de $s = t$ et nous avons $X(s) = Ke^s$. La constante K se fixe par la condition initiale $X(t) = x$:

$$X(s) = xe^{s-t}. \tag{32.163}$$

Et cette solution tient en réalité pour tout s parce que $X(s)$ est alors toujours positif.

Si au contraire $x < 0$ nous avons la solution

$$X(s) = xe^{t-s}. \tag{32.164}$$

Au final,

$$X(s;x,t) = \begin{cases} xe^{t-s} & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ xe^{s-t} & \text{si } x > 0 \end{cases} \tag{32.165}$$

L'application $(s,x,t) \mapsto X(s;x,t)$ est continue. En ce qui concerne la dérivée partielle $\partial_x X$ en $x = 0$ nous avons :

$$\frac{\partial X}{\partial x}(s,0,t) = \lim_{\epsilon\to 0} \frac{X(s,\epsilon,t) - X(s,0,t)}{\epsilon} = \lim_{\epsilon\to 0} \frac{X(s,\epsilon,t)}{\epsilon}. \tag{32.166}$$

EQooTQXQooAtRxNT

La limite à droite donne :

$$\lim_{\epsilon\to 0^+} \frac{X(s,\epsilon,t)}{\epsilon} = \frac{\epsilon e^{s-t}}{\epsilon} = e^{s-t}. \tag{32.167}$$

La limite à gauche donne :

$$\lim_{\epsilon\to 0^-} \frac{X(s,\epsilon,t)}{\epsilon} = e^{t-s}. \tag{32.168}$$

Les deux limites n'étant pas égales, la limite (32.166) n'existe pas[6] et l'application $(s,x,t) \mapsto X(s,x,t)$ n'est pas dérivable par rapport à x. △

PROBooYUYQooSBsrNh

¡¡ Avertissement/question au lecteur !! 32.31

Je ne suis pas tout à fait certain de la définition du V dans le lemme 32.32. Il faudrait écrire la preuve complète pour s'assurer.

LEMooOJSNooXTJoEf

Lemme 32.32 ([732]).
Soit un application $A\colon \overline{B(t_0,\tau)} \times \overline{B(a,R)} \to \mathcal{L}(\mathbb{R}^n)$ continue par rapport à sa première variable ($t_0 \in \mathbb{R}$ et $a \in \mathbb{R}^n$). Nous considérons l'équation

$$\begin{cases} \dfrac{\partial \psi}{\partial t}(t,b) = A(t,b)\psi(t,b) & \text{(32.169a)} \\ \psi(t_0,b) = \psi_0. & \text{(32.169b)} \end{cases}$$

Alors il existe un voisinage V de a dans lequel nous avons l'estimation

$$\|\psi(t,v) - \psi(t,w)\| \leqslant \|\psi_0\|\tau \max_{s\in\overline{B(t_0,\tau)}} \|A(s,v) - A(s,w)\| \times \\ \times \exp\left(\tau \max_{s\in\overline{B(t_0,\tau)}} \max\{\|A(s,v)\|, \|A(s,w)\|\}\right) \tag{32.170}$$

pour tout $t \in \overline{B(t_0,\tau)}$ et $v, w \in V$.

6. Si vous comptez donner ça à manger au jury d'un concours, soyez prudente et n'écrivez pas l'équation (32.166) au tableau. Réfléchissez comment rédiger cela correctement.

THOooSTHXooXqLBoT

Théorème 32.33 (Régularité C^1 du flot [732])**.**
Soit un intervalle ouvert I de $\mathbb{R}$ et un ouvert connexe Ω de $\mathbb{R}^n$. Soit une fonction $f \in C^1\big(I\times\Omega, \mathbb{R}^n\big)$, $a \in \Omega$ et $t_0 \in I$.

Il existe un voisinage $W \times V = \overline{B(t_0,\tau)} \times \overline{B(a,r)}$ de (t_0, a) dans $I \times \Omega$ et une unique application $\varphi\colon W \times V \to \Omega$ telle que

$$\begin{cases} \dfrac{\partial\varphi}{\partial t}(t,x) = f\big(t,\varphi(t,x)\big) & \text{(32.171a)}\\ \varphi(t_0,x) = x & \text{(32.171b)} \end{cases}$$

pour tout $x \in V$.

L'application $(t, x) \mapsto \varphi(t, x)$ est de classe C^1.

¡¡ Avertissement/question au lecteur !! 32.34
La preuve qui suit doit être lue avec beaucoup d'attention, en particulier sur les incohérences possibles de notations, et sur les oublis possibles de précautions oratoires type « quitte à encore réduire les voisinages V et W ».

Démonstration. En termes de notations, pour $x \in \Omega$ fixé nous écrivons $y_x(t)$ pour $\varphi(t, x)$ et pour $t \in I$ fixé nous notons $\varphi_t(x)$ pour $\varphi(t, x)$.

De plus lorsque nous écrirons des choses comme $g\colon \mathbb{R} \to \mathbb{R}$, nous n'entendrons pas que g est effectivement définie sur tout $\mathbb{R}$. La notation « $g\colon \mathbb{R} \to \mathbb{R}$ » indiquera seulement que la variable de g est réelle, et que nous comptons préciser le domaine plus tard. Cette remarque s'applique seulement à cette démonstration et non à l'ensemble du livre.

Nous considérons $R > 0$ tel que $\overline{B(a, 2R)} \subset \Omega$ et ensuite nous posons $V = \overline{B(a, R)}$. La fonction y_x, solution pour la condition initiale $y_x(t_0) = x$ est définie sur $W = [t_0 - \tau, t_0 + \tau]$ et prend ses valeurs dans $\overline{B(x, R)}$. Ceci est parce que y_x est continue, alors en prenant τ assez petit, la valeur de $y_x(t)$ ne va pas s'éloigner de x lorsque t ne s'éloigne pas de t_0.

Nous savons déjà de la proposition 32.28 que φ est C^1 en t et localement Lipschitz en sa seconde variable, avec une constante Lipschitz uniforme sur $W \times V$. Elle est donc continue en tant que fonction

$$\varphi\colon V \times W \to \mathbb{R}^n. \tag{32.172}$$

(1) **La différentielle partielle Df** Pout t fixé nous notons $Df_{(t,x)}$ la différentielle de f par rapport à x. C'est-à-dire que

$$\begin{aligned} Df_{(t,x)}\colon \mathbb{R}^n &\to \mathbb{R}^n\\ u &\mapsto \frac{d}{ds}\Big[f(t, x + su)\Big]_{s=0}. \end{aligned} \tag{32.173}$$

C'est un élément de $\mathcal{L}(\mathbb{R}^n)$, l'ensemble des applications linéaires de $\mathbb{R}^n$ vers $\mathbb{R}^n$. Nous allons montrer que

$$(t, x) \mapsto Df_{(t,x)} \tag{32.174}$$

est continue en tant qu'application $\mathbb{R} \times \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n)$. Pour cela nous introduisons l'application d'inclusion $i\colon \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n$, $i(u) = (0, u)$. Elle donne

$$Df_{(t,x)}(u) = \frac{d}{ds}\Big[f\big((t, x) + s(0, u)\big)\Big]_{s=0} = df_{(t,x)} \circ i(u). \tag{32.175}$$

Autrement dit

$$Df_{(t,x)} = df_{(t,x)} \circ i. \tag{32.176}$$

Or l'application $(t, x) \mapsto df_{(t,x)}$ est continue par hypothèse (f est de classe C^1) et l'application

$$\begin{aligned} \mathcal{L}(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n) &\to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)\\ A &\mapsto A \circ i \end{aligned} \tag{32.177}$$

EQooZTAPooCduWcl

est également continue. Donc $(t,x) \mapsto Df_{(t,x)}$ est continue [7].

(2) **L'équation aux variations**

Soit $x \in \Omega$. Nous introduisons l'opérateur

$$\begin{aligned} S_x \colon \mathbb{R} \times \mathcal{L}(\mathbb{R}^n) &\to \mathcal{L}(\mathbb{R}^n) \\ S_x(t,\psi) &= Df_{(t,y_x(t))} \circ \psi. \end{aligned} \tag{32.178}$$

Par ce que nous avons raconté, cela est une fonction continue en sa première variable et Lipschitz en sa seconde variable. Nous identifions $\mathcal{L}(\mathbb{R}^n)$ à $\mathbb{R}^{n^2}$.

Toujours pour chaque x considéré nous posons l'équation différentielle ordinaire EQooQONGooBrxuSA

$$\begin{cases} \dfrac{\partial \psi}{\partial t}(t,x) = S_x\big(t,\psi(t,x)\big) & \text{(32.179a)} \\ \psi(t_0,x) = \mathrm{Id}\,. & \text{(32.179b)} \end{cases}$$

qui est une équation différentielle ordinaire pour $\psi \colon \mathbb{R} \times \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n)$ rentrant dans le cadre de Cauchy-Lipschitz.

Quel est le domaine de définition de ψ pour sa première variable ? C'est un ouvert autour de t_0. Nous réduisons W de telle sorte que la solution ψ soit définie sur W. Idem pour la variable x qui est dans un voisinage de a.

L'équation (32.179) s'appelle l'**équation aux variations**. Nous allons montrer dans la douleur que ψ est continue et est la différentielle de φ_t, c'est-à-dire que

$$(d\varphi_t)_b = \psi(t,b). \tag{32.180}$$

(3) **ψ est continue en (t,x) (début)**

Il s'agit de majorer les deux termes de

$$\|\psi(t_1,a_1) - \psi(t_2,a_2)\| \leqslant \|\psi(t_1,a_1) - \psi(t_2,a_1)\| + \|\psi(t_2,a_1) - \psi(t_2,a_2)\|. \tag{32.181}$$ EQooVUNUooExeQba

(3a) **Premier terme**

Nous avons

$$\begin{aligned} \|\psi(t_1,b) - \psi(t_2,b)\| &= \| \int_{[t_1,t_2]} \frac{\partial \psi}{\partial t}(s,b) ds\| & \text{(32.182a)} \\ &\leqslant \int_{[t_1,t_2]} \|Df_{(s,y_b(s))} \circ \psi(s,b)\| ds & \text{(32.182b)} \\ &\leqslant \int_{[t_1,t_2]} \|Df_{(s,t_b(s))}\| \|\psi(s,b)\| ds & \text{(32.182c)} \\ &\leqslant |t_1 - t_2| \max_{s\in[t_1,t_2]} \|Df_{(s,y_b(s))}\| \max_{s\in[t_1,t_2]} \|\psi(s,b)\|. & \text{(32.182d)} \end{aligned}$$ SUBEQooLYMAooRMaMhn

Nous allons majorer le second maximum. Prenons $t \in [0,\tau]$; et posons $A(u,b) = Df_{(u,y_b(u))}$ pour alléger les notations. Par l'équation de définition de ψ nous avons

$$\psi(t,b) = \psi(0,b) + \int_{[0,t]} A(u,b)\psi(u,b) du, \tag{32.183}$$

et donc

$$\|\psi(t,b)\| \leqslant \|\psi_0\| + \int_{[0,t]} \|A(u,b)\| \|\psi(u,b)\| du. \tag{32.184}$$

7. Si quelqu'un peut prouver ça de façon moins verbeuse, je suis preneur. Il me semble que quel que soit la façon dont on s'y prend, sous le capot, on passe par la continuité de l'application (32.177).

En y appliquant le lemme de Grönwall dans sa version 32.5 nous trouvons

$$\|\psi(s,b)\| \leqslant \|\psi_0\| \exp\left(\int_{[0,s]} \|A(u,b)\| du\right) \tag{32.185a}$$

$$\leqslant \|\psi_0\| \exp\left(s \max_{u\in[0,s]} \|A(u,b)\|\right). \tag{32.185b}$$

En retournant à (32.182d) nous avons $\psi_0 = \mathrm{Id}$ et donc $\|\psi_0\| = 1$ et

$$\max_{s\in[t_1,t_2]} \|\psi(s,b)\| \leqslant \max_{s\in[t_1,t_2]} \exp\left(s \max_{u\in[0,t]} \|Df_{(u,y_b(u))}\|\right) \tag{32.186}$$

Là dedans nous pouvons remplacer t par $\max\{|t_1|,|t_2|\}$. Posons enfin, pour alléger les expressions

$$\alpha(t_1,t_2,b) = \max_{s\in[t_1,t_2]} \|Df_{(s,y_b(s))}\|. \tag{32.187}$$

La majoration que nous retenons est :

$$\|\psi(t_1,b) - \psi(t_2,b)\| \leqslant |t_1 - t_2|\alpha(t_1,t_2,b) \exp\big(\max\{|t_1|,|t_2|\}\alpha(0,t,b)\big). \tag{32.188}$$

Cela tend vers zéro lorsque $t_1 \to t_2$.

(3b) **Deuxième terme**

En ce qui concerne le second terme,

$$\|\psi(t,b_1) - \psi(t,b_2)\| \tag{32.189}$$

nous utilisons le lemme 32.32 qui donne, pour $t \in [t_0 - \tau, t_0 + \tau]$,

$$\begin{aligned}\|\psi(t,b_1) - \psi(t,b_2)\| \leqslant \tau \max_{s\in\overline{B}(0,\tau)} &\|Df_{s,y_{b_1}(s)} - Df_{s,y_{b_2}(s)}\| \times \\ &\times \exp\left(\tau \max\{\|Df_{s,y_{b_1}(s)}\|, \|Df_{s,y_{b_2}(s)}\|\}\right).\end{aligned} \tag{32.190}$$

Dans notre cas, $t_0 = 0$, donc $t \in [-\tau,\tau]$. Vu la continuité de Df, nous avons

$$\max_{s\in\overline{B}(0,\tau)} \|Df_{s,y_{b_1}(s)} - Df_{s,y_{b_2}(s)}\| \to 0 \tag{32.191}$$

lorsque $b_1 \to b_2$.

(3c) **ψ est continue en (t,x) (fin)**

Les deux bons calculs faits, nous avons, en repartant de (32.181),

$$\lim_{(t_1,b_1)\to(t_2,b_2)} \|\psi(t_1,b_1) - \psi(t_2,b_2)\| = 0, \tag{32.192}$$

ce qui signifie que ψ est une fonction continue de ses deux variables en même temps.

(4) **Différentiabilité de φ (début)**

Nous montrons maintenant que $D\varphi(t,x)$ existe. Pour rappel, D est la différentielle par rapport à la seconde variable. Nous sommes à étudier l'existence de $D\varphi_{(t,b)} = d(\varphi_t)_b$. Nous posons

$$\theta(t,h) = \varphi(t,b+h) - \varphi(t,b) = y_{b+h}(t) - y_b(t) \tag{32.193}$$

où b est le point où nous étudions la différentiabilité. Il est dans un voisinage du point a fixé depuis le début et autour duquel il existe un voisinage qui donne un sens à tout ce que nous avons fait jusqu'à présent. La dépendance de θ en b est implicite. Vu que φ est Lipschitz en sa seconde variable, nous avons la majoration

$$\|\theta(t,h)\| \leqslant C\|h\| \tag{32.194}$$

EQooKYELooZlfeed

dès que $t \in V$ et $b, b+h \in W$.

De plus, parce que t_0 est le temps de la condition initiale nous avons

$$\theta(t_0, h) = y_{b+h}(t_0) - y_b(t_0) = a + h - a = h. \tag{32.195}$$

Et aussi, par définition de ψ :

$$\psi(t, b) = \psi_0 + \int_{t_0}^{t} \frac{\partial \psi}{\partial t}(s, b) ds = \psi_0 + \int_{t_0}^{t} Df_{(s,y_b(s))} \circ \psi(s, b) \tag{32.196}$$

En appliquant à h et en se souvenant que $\psi_0 = \mathrm{Id}$,

$$\psi(t, b)h = h + \int_{t_0}^{t} \Big(Df_{s,y_b(s)} \circ \psi(s, b) \Big) h \, ds. \tag{32.197}$$

Puis on peut faire un calcul assez classique en se souvenant que $\theta(t_0, h) = h$:

$$\theta(t, h) = \theta(t_0, h) + \int_{t_0}^{t} \big[\frac{\partial \varphi}{\partial t}(s, b+h) - \frac{\partial \varphi}{\partial t}(s, b) \big] ds \tag{32.198a}$$

$$= h + \int_{t_0}^{t} \big[f\big(s, y_{b+h}(s)\big) - f\big(s, y_b(s)\big) \big] ds. \tag{32.198b}$$

On fait la différence entre les deux :

$$\theta(t, h) - \psi(t, b)h = - \int_{t_0}^{t} \big[Df_{s,y_b(s)} \circ \psi(s, b)h - f\big(s, y_{b+h}(s)\big) + f\big(s, y_b(s)\big) \big] ds. \tag{32.199}$$

Nous y ajoutons et soustrayons $Df_{s,y_b(s)}\theta(s, h)$ et nous retenons la majoration suivante :

$$\begin{aligned} \|\theta(t, h) - \psi(t, b)h\| \leqslant & \int_{t_0}^{t} \|Df_{(s,y_b(s))}\psi(s, b) - Df_{(s,y_b(s))}\theta(s, h)\| ds \\ & + \int_{t_0}^{t} \|f(s, y_{b+h}(s)) - f(s, y_b(s)) + Df_{(s,y_b(s))}\theta(s, h)\| ds. \end{aligned} \tag{32.200}$$

EQooODHPooDYyBoH

Nous allons encore majorer ces deux termes séparément. Soit $\epsilon > 0$.

(4a) **Premier terme**

Ce qui est dans la norme à majorer est

$$Df_{(s,y_b(s))}\big(\psi(s, b)h - \theta(s, h)\big). \tag{32.201}$$

Vu que Df est continue et que y_b est continue [8], l'application $s \mapsto Df_{s,y_b(s)}$ est continue et donc de norme majorée sur le compact $[t_0, t]$. Nous rapellons la notation

$$\alpha(t_0, t, b) = \max_{s \in [t_0, t]} \|Df_{s,y_b(s)}\|, \tag{32.202}$$

et nous majorons encore et toujours. D'abord

$$\int_{t_0}^{t} \|Df_{(s,y_b(s))}\big(\psi(s, b)h - \theta(s, h)\big)\| ds \leqslant \alpha(t_0, t, b) \int_{t_0}^{t} \|\psi(s, b) - \theta(s, h)\| ds. \tag{32.203}$$

(4b) **Deuxième terme**

Pour traiter le deuxième terme, nous allons provisoirement noter $x = y_b(s)$ et $y = y_{b+h}(s)$; entre autres, $y - x = \theta(s, h)$. Ce qui est écrit dans le second terme de (32.200) est

$$f(s, y) - f(s, x) + Df_{(s,x)}\theta(s, h) = f(s, y) - f(s, x) + Df_{(s,x)}(y - x) \tag{32.204}$$

8. Il faut encore réduire les voisinages V et W pour que ceci ait un sens.

Comme D ne s'applique pas à la variable s, nous pouvons alléger la notation et déduire de la différentiabilité de f qu'il existe un $\eta > 0$ tel que $x, y \in W$ avec $\|y - w\| \leqslant \eta$ implique

$$\|f(y) - f(x) - Df_x(y - x)\| \leqslant \epsilon\|y - x\|. \tag{32.205}$$

Prenons $\|h\| \leqslant \eta/C$ (le C de (32.194)) ; en déballant les notations,

$$\|f\big(s, y_{b+h}(s)\big) - f\big(s, y_b(s)\big) - Df_{(s,y_b(s))}\theta(s, h)\| \leqslant \epsilon\|\theta(s, h)\| \leqslant \epsilon C\|h\|. \tag{32.206}$$

(4c) **Les deux termes ensemble**

En remettant les deux dans (32.200) nous trouvons la majoration

$$\|\theta(t, h) - \psi(t, b)h\| \leqslant |t - t_0|\epsilon C\|h\| + \alpha(t_0, t, v)\int_{t_0}^{t} \|\psi(s, b)h - \theta(s, h)\|ds \tag{32.207}$$

qui est encore de la graine à Grönwall avec

$$\begin{cases} u(t) = \|\theta(t, h) - \psi(t, b)\| & \text{(32.208a)} \\ b(t) = |t - t_0|\epsilon C\|h\| & \text{(32.208b)} \\ a(s) = \alpha(t_0, t, v), & \text{(32.208c)} \end{cases}$$

la troisième étant une fonction constante. Cela donne, pour $t \in [t_0 - \tau, t_0 + \tau]$,

$$\|\theta(t, h) - \psi(t, b)h\| \leqslant |t - t_0|\epsilon C\|h\| + \int_{t_0}^{t} (s - t_0)\epsilon C\|h\|\alpha(t_0, t, b) \exp\left(\int_{s}^{t} \alpha(t_0, t, b)du\right) ds. \tag{32.209}$$

En valeur absolue, la différence $s - t_0$ est majorée par τ, l'intégrale dans l'exponentielle vaut $(t - s)\alpha(t_0, t, b)$, et restons avec

$$\|\theta(t, h) - \psi(t, b)h\| \leqslant \tau\epsilon C\|h\| + \tau\int_{t_0}^{t} \epsilon C\|h\|\alpha(t_0, t, b)e^{\alpha(t_0,t,b)(t-s)}ds. \tag{32.210}$$

En supposant $t > t_0$ nous pouvons calculer l'intégrale. Si vous m'avez suivi jusqu'ici, vous devriez avoir de tels maux de tête que je vous donne la réponse :

$$\int_{t_0}^{t} a(t_0, t, b)e^{(t-s)\alpha(t_0,t,b)}ds = e^{(t_0-t)\alpha(t_0,t,b)} - 1. \tag{32.211}$$

En remettant dans l'expression,

$$\|\theta(t, h) - \psi(t, b)h\| \leqslant \tau\epsilon C\|h\| + \epsilon C\|h\|\alpha(t_0, t, b)\tau\big(e^{(t-t_0)} - 1\big) = \tau\epsilon C\|h\|e^{(t-t_0)\alpha(t,t_0,b)}. \tag{32.212}$$

Nous pouvons majorer $t - t_0$ par τ et $\alpha(t, t_0, b)$ par $\alpha(t_0 - \tau, t_0 + \tau, b)$ pour avoir la majoration

$$\|\theta(t, h) - \psi(t, b)h\| \leqslant \tau\epsilon C\|h\|e^{\tau\alpha(t_0-\tau,t_0+\tau,b)}. \tag{32.213}$$

(4d) **Différentiabilité de $\varphi(t, b)$ (fin)**

Nous écrivons la définition 10.243 de la différentiabilité : nous voulons vérifier que

$$\lim_{h\to 0} \frac{\varphi(t, b + h) - \varphi(t, b) - \psi(t, b)h}{\|h\|} = 0. \tag{32.214}$$

Nous remplaçons $\varphi(t, b+h)-\varphi(t, b)$ par $\theta(t, h)$ et prenons la norme avec les majorations données :

$$\lim_{h\to 0} \frac{\|\varphi(t, b + h) - \varphi(t, b) - \psi(t, b)h\|}{\|h\|} \leqslant \lim_{h\to 0} \tau\epsilon Ce^{\tau e(t_0-\tau,t_0+\tau,b)}. \tag{32.215}$$

Cela étant valable pour tout ϵ, nous en déduisons la nullité de la limite.

Nous avons démontré que φ était différentiable par rapport à sa deuxième variable et que

$$D\varphi_{(t,b)} = \psi(t, b). \tag{32.216}$$

EQooPJFQooHuQIuw

(5) **Conclusion : φ est de classe C^1**
Nous avons déjà prouvé que $(t,b)\mapsto\psi(t,b)$ est continue. Donc de (32.216) nous déduisons que les dérivées partielles $(t,b)\mapsto\frac{\partial\varphi}{\partial x_i}(t,b)$ sont continues. Mais comme φ est Lipschitz en t, la dérivée partielle $(t,b)\mapsto\frac{\partial\varphi}{\partial t}(t,b)$ est également continue.
La continuité de toutes les dérivées partielles de φ nous donne la classe C^1 pour φ par le théorème 12.294.

□
PROPooINLNooDVWaMn

Proposition 32.35 (Régularité C^p du flot[733])**.**
Soit un intervalle ouvert I de $\mathbb{R}$ et un ouvert connexe Ω de $\mathbb{R}^n$. Soit une fonction $f\in C^p\big(I\times\Omega,\mathbb{R}\big)$, ainsi que $a\in\Omega$ et $t_0\in I$.

Il existe un voisinage $W\times V=\overline{B(t_0,\tau)}\times\overline{B(a,r)}$ de (t_0,a) dans $I\times\Omega$ et une unique application $\varphi\colon W\times V\to\Omega$ telle que

$$\begin{cases}\dfrac{\partial\varphi}{\partial t}(t,x)=f\big(t,\varphi(t,x)\big) & \text{(32.217a)}\\ \varphi(t_0,x)=x & \text{(32.217b)}\end{cases}$$

pour tout $x\in V$.

L'application $(t,x)\mapsto\varphi(t,x)$ est de classe C^p.

Démonstration. Nous savons déjà par le théorème 32.33 que $(t,x)\mapsto\varphi(t,x)$ est de classe C^1. Nous supposons que f est de classe C^p avec $p\geqslant 2$.

Vu que φ et f sont de classe C^1, nous avons aussi que l'application $(t,x)\mapsto f\big(t,\varphi(t,x)\big)$ est de classe C^1. L'équation donne alors immédiatement le fait que

$$(t,x)\mapsto\frac{\partial\varphi}{\partial t}(t,x) \tag{32.218}$$

est de classe C^1.

En ce qui concerne la régularité par rapport aux autres variables, il faudra travailler un peu plus.

(1) **Une équation différentielle pour le flot** Nous allons commencer par un habile jeu d'écriture : la formule

$$\varphi(t,x)=x+\int_{t_0}^t f\big(s,\varphi(s,x)\big)ds \tag{32.219}$$

devient

$$\varphi_t(x)=x+\int_{t_0}^t f\big(s,\varphi_s(x)\big)ds. \tag{32.220}$$
EQooQGVOooYMEggM

Dans le même ordre d'idée nous notons $f_s(x)=f(s,x)$, et ce qui se trouve dans l'intégrale (32.220) n'est autre que la fonction

$$g_s(x)=(f_s\circ\varphi_s)(x). \tag{32.221}$$

Tout cela pour différentier l'égalité (32.220) par la proposition 17.25 :

$$\begin{aligned}(d\varphi_t)_x&=\mathrm{Id}+\int_{t_0}^t(dg_s)_x ds & \text{(32.222a)}\\ &=\mathrm{Id}+\int_{t_0}^t(df_s)_{\varphi_s(x)}\circ(d\varphi_s)_s ds. & \text{(32.222b)}\end{aligned}$$

Nous dérivons ensuite cela par rapport à t :

$$\frac{\partial}{\partial t}\Big((d\varphi_t)_x\Big)=(df_t)_{\varphi_t(x)}\circ(d\varphi_t)_x. \tag{32.223}$$
EQooBETGooXKWRxX

Cela est une égalité dans $\mathcal{L}(\mathbb{R}^n)$.

Nous introduisons la fonction

$$\begin{aligned} F\colon I \times \mathcal{L}(\mathbb{R}^n) \times \Omega &\to \mathcal{L}(\mathbb{R}^n) \\ (t, A, x) &\mapsto (df_t)_{\varphi_t(x)} \circ A. \end{aligned} \tag{32.224}$$

En fait, à la place de I et Ω il faut prendre des petits voisinages dans lesquels les choses ont un sens. Ce que dit l'équation 32.223 est que l'application

$$\begin{aligned} A\colon I \times \Omega &\to \mathcal{L}(\mathbb{R}^n) \\ (t, x) &\mapsto (d\varphi_t)_x \end{aligned} \tag{32.225}$$

vérifie l'équation différentielle EQooTOJMooLXLfVv

$$\begin{cases} \dfrac{\partial A}{\partial t}(t, x) = F\big(t, A(t, x), x\big) & \text{(32.226a)} \\ A(t_0, x) = \mathrm{Id}\,. & \text{(32.226b)} \end{cases}$$

(2) **Une autre équation différentielle**

Nous n'oublions pas l'équation différentielle pour la dérivée par rapport à t :

$$\frac{\partial \varphi}{\partial t}(t, x) = f\big(t, \varphi(t, x)\big). \tag{32.227}$$ EQooYOJPooKEgicc

(3) **Réécriture pour la différentielle**

Nous allons récrire l'équation (32.226) de façon à ce que le paramètre x soit inclus dans la condition initiale. De cette manière, la solution pourra profiter de la régularité C^1 du flot déjà prouvée dans le théorème 32.33.

Soit

$$\begin{aligned} g\colon I \times \big(\mathcal{L}(\mathbb{R}^n) \times \Omega\big) &\to \mathcal{L}(\mathbb{R}^n) \times \Omega \\ \big(t, (A, x)\big) &\mapsto \Big(F(t, a, x), 0\Big). \end{aligned} \tag{32.228}$$

Nous posons $E = \mathcal{L}(\mathbb{R}^n) \times \Omega$; c'est cet espace qui va jouer le rôle de Ω. Nous considérons à présent l'équation différentielle suivante pour $z_x\colon I \to E$: EQooSMYOooDaxgQx

$$\begin{cases} \begin{pmatrix} z_1'(t) \\ z_2'(t) \end{pmatrix} = z'(t) = g(t, z(t)) = \begin{pmatrix} F\big(t, z_1(t), z_2(t)\big) \\ 0 \end{pmatrix} & \text{(32.229a)} \\ z(t_0) = (\mathrm{Id}, x). & \text{(32.229b)} \end{cases}$$

Il devrait y avoir un indice (Id, x) à z parce que c'est sa condition initiale. La fonction g est de classe C^p, donc cette équation admet une unique solution dont le flot est de classe C^1. Autrement dit, si S est dans un voisinage de Id, l'application

$$(t, x) \mapsto z(t) \tag{32.230}$$

est de classe C^1. Nous allons montrer qu'en posant $A(t, x) = z_1(t)$, nous avons une solution de (32.226) (l'unicité de la solution impose que cette solution est effectivement la différentielle de $d\varphi_t$). D'abord, la seconde ligne de l'équation différentielle est $z_2'(t) = 0$, c'est-à-dire $z_2(t) = x$ pour tout t.

Sachant cela, la première équation devient

$$\begin{cases} z_1'(t) = F\big(t, z_1(t), x\big) & \text{(32.231a)} \\ z_1(t_0) = \mathrm{Id}, & \text{(32.231b)} \end{cases}$$

qui est l'équation différentielle pour A. Rappel : il y a partout une dépendance de z en sa condition initiale x que nous n'avons pas écrite pour des raisons de légèreté notationnelle. Il n'en reste pas moins que le flot de l'équation différentielle pour z est C^1, c'est-à-dire que $(t, x) \mapsto z_1(t)$ est de classe C^1.

Par conséquent, $(t, x) \mapsto A(t, x)$ est également C^1.

(4) **Régularité C^2 du flot**

Le fait que A soit C^1 n'implique pas que le flot le soit parce que le fait le flot suive la même équation différentielle que A ne n'implique pas qu'il soit égal. Il y a un raisonnement à faire.

Le fait est que si A est une solution de (32.226), alors $z(t) = \big(A(t,x), x\big)$ est solution de (32.229). C'est l'unicité de cette dernière qui permet de déduire l'unicité de la solution pour A.

Nous avons donc que l'unique solution A du système (32.226) est égale à $A(t,x) = (d\varphi_t)_x$ et est de classe C^1 par rapport à (t,x).

Donc $(t,x) \mapsto \varphi(t,x)$ est de classe C^2.

(5) **Régularité C^p**

Nous avons vu que le flot de $y' = f(t,y)$ est de classe C^2 dès que f est de classe C^2. Supposons que f soit de classe C^p et montrons que si le flot est de classe C^k $(k < p)$ alors il est de classe C^{k+1}.

Vu que le flot d'une équation différentielle de classe C^p est de classe C^k, en particulier celui de (32.229) est de classe C^k. Donc aussi la solution pour $A(t,x) = (d\varphi_t)_x$ est de classe C^k. Et vu que $(t,x) \mapsto (d\varphi_t)_x$ est de classe C^k, l'application φ est de classe C^{k+1}.

□

NORMooWEWVooXbGmfE

32.36.

Le théorème d'inversion locale 17.47 nous permet de dire que, pour t fixé, le flot $x \mapsto \varphi_t(x)$ est un C^p-difféomorphisme local.

PROPooPYHWooIZhQST

Proposition 32.37 (Cauchy-Lipschitz avec paramètre, régularité C^p[734, 735, 736])**.**
Soit un intervalle ouvert I de $\mathbb{R}$, un connexe ouvert Ω de $\mathbb{R}^n$ et un intervalle ouvert Λ de $\mathbb{R}^d$. Soit une fonction $f \in C^p(I \times \Omega \times \Lambda, \mathbb{R}^n)$ localement Lipschitz en Ω. Soient $t_0 \in I$, $y_0 \in \Omega$ et $\lambda_0 \in \Lambda$. Il existe un voisinage compact de (t_0, y_0, λ_0) sur lequel le problème

$$\begin{cases} y'_\lambda(t) = f\big(t, y_\lambda(t), \lambda\big) & \text{(32.232a)} \\ y_\lambda(t_0) = y_0 & \text{(32.232b)} \end{cases}$$

possède une unique solution. De plus $(t,\lambda) \mapsto y_\lambda(t)$ est de classe C^p par rapport à ses deux variables.

Démonstration. Nous récrivons immédiatement le problème pour la fonction $y\colon I \times \Lambda \to \mathbb{R}^n$ donné par $y(t,\lambda) = y_\lambda(t)$: SUBEQooXMYMooKfpqQW

$$\begin{cases} \dfrac{\partial y}{\partial t}(t,\lambda) = f\big(t, y(t,\lambda), \lambda\big) & \text{(32.233a)} \\ y(t_0) = y_0. & \text{(32.233b)} \end{cases}$$

Nous allons montrer que ce problème est en réalité équivalent à un problème sans paramètre. Nous posons $E = \Omega \times \Lambda$ et

$$\begin{aligned} g\colon I \times E &\to E \\ (t,x) &\mapsto \big(f(t,x_1,x_2), 0\big) \end{aligned} \tag{32.234}$$

où x_1 est la composante Ω de x et x_2 est la composante Λ de x. Pour une valeur $\mu \in \Lambda$ donnée nous considérons le problème au condition initiales SUBEQSooIBTNooHzYImh

$$\begin{cases} x'(t) = g\big(t, x(t)\big) & \text{(32.235a)} \\ x(t_0) = (y_0, \mu). & \text{(32.235b)} \end{cases}$$

Le théorème de Cauchy-Lipschitz que nous prenons sous la forme 32.35 nous indique que ce problème admet une unique solution maximale et que le flot $\big(t, (y_0,\mu)\big) \mapsto x_{(y_0,\mu)}(t)$ est de classe C^p.

Nous passons maintenant à la résolution du problème (32.233)

(1) **Existence d'une solution C^1**

Nous montrons à présent que la fonction y donnée par

$$y(t,\mu) = x_{(t_0,\mu)}(t)_1 \tag{32.236}$$

est solution de (32.233). Vu que x a deux composantes, nous pouvons un peu déballer l'équation. Afin d'éviter les notations laborieuses nous allons noter x pour $x_{(t_0,\mu)}$ et donc $x_1(t)$ pour $x_{(t_0,\mu)}(t)$. Nous avons l'équation différentielle

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} f\big(t, x_1(t), x_2(t)\big) \\ 0 \end{pmatrix} \tag{32.237}$$

avec la condition initiale

$$\begin{pmatrix} x_1(t_0) \\ x_2(t_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ \mu \end{pmatrix}. \tag{32.238}$$

La seconde ligne de l'équation donne immédiatement $x_2(t) = \mu$ pour tout t. En injectant dans la première ligne :

$$x_1'(t) = f\big(t, x_1(t), \mu\big). \tag{32.239}$$

Or vue la définition de y, le nombre $x_1'(t)$ n'est autre que $\frac{\partial y}{\partial t}(t,\mu)$. La fonction y que nous avons définie vérifie donc

$$\frac{\partial y}{\partial t}(t,\mu) = f\big(t, y(t,\mu), \mu\big) \tag{32.240}$$

et la condition initiale $y(t_0) = x_1(t_0) = y_0$. Elle est donc bien solution du problème initial. De plus l'application $(t,\mu) \mapsto y(t,\mu) = x_{(t_0,\mu)}(t)_1$ est de classe C^p.

(2) **Unicité**

Pour l'unicité, soit on invoque la proposition 17.42 qui donne l'unicité dans les fonctions continues et a fortiori dans les fonctions C^1. Soit on fait le jeu inverse : on prouve qu'à chaque solution de (32.233) correspond une solution de (32.235), et l'unicité de la solution x donne l'unicité du côté de y.

□

¡¡ Avertissement/question au lecteur !! 32.38

Je ne suis pas du tout sûr de la proposition 32.39. Écrivez-moi si vous avez un commentaire, et surtout si vous avez une preuve. La proposition 32.35 entre certainement dans la preuve.

PROPooKPCQooQNMLQL

Proposition 32.39.

Soient un intervalle I de $\mathbb{R}$, un ouvert connexe V de $\mathbb{R}^n$, une application $f\colon \mathbb{R}^2 \to \mathbb{R}$ de classe C^p, ainsi qu'une application $y\colon I \times V \times I \to \mathbb{R}$ telle que

$$\begin{cases} \dfrac{\partial y}{s}(s,x,t) = f\big(y(s,x,t), s\big) & \text{(32.241a)} \\ y(t,x,t) = x. & \text{(32.241b)} \end{cases}$$

Alors y est de classe C^p.

Démonstration. Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

¡¡ Avertissement/question au lecteur !! 32.40

L'énoncé du lemme 32.41 est douteux. À mon avis il doit être supprimé et remplacé par la proposition 32.39, pourvu que cette dernière soit vraie.

LEMooQWDNooOjNXhl

Lemme 32.41.

Soient t et x fixés, ainsi que f de classe C^p. Nous considérons la solution $y_{x,t}$ de EQooKHSKooEFCsMQ

$$\begin{cases} \dfrac{\partial y}{\partial s}(s) = f\big(y(s), s\big) & \text{(32.242a)} \\ y_{x,t}(t) = x & \text{(32.242b)} \end{cases}$$

Alors l'application $t \mapsto y_x(s)$ est de classe C^p.

Démonstration. Soit t fixé, et l'équation différentielle

$$\begin{cases} \frac{\partial z}{\partial s}(s) = -f\big(z(s), t-s\big) & (32.243a) \\ z(0) = x. & (32.243b) \end{cases}$$

Par le théorème 32.35, la solution z est de classe C^p en (s,x). En posant $y(s) = z(t-s)$ il est vite vérifié que y est solution de (32.242). C'est alors bien de classe C^p en t. □

32.8.4 Stabilité de Lyapunov

DefKMCGooOeFKlA

Définition 32.42.
*Dans le cas de l'équation différentielle $y'(t) = f\big(y(t),t\big)$ pour $y\colon \mathbb{R} \to \mathbb{R}^n$, un point $a \in \mathbb{R}^n$ est un **point d'équilibre** lorsque la fonction constante $y(t) = a$ est une solution.*

*Le point d'équilibre $a \in \mathbb{R}^n$ est **stable** si pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $\|y(0) - a\| < \delta$ implique $\|y(t) - a\| < \epsilon$ pour tout t.*

ThoBSEJooIcdHYp

Théorème 32.43 (Théorème de stabilité de Lyapunov[1, 737, 380, 729])**.**
Soit l'équation différentielle EqZZLBooKBkZkG

$$\begin{cases} y'(t) = f(y) & (32.244a) \\ y(0) = y_0 & (32.244b) \end{cases}$$

avec une fonction $f\colon \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 vérifiant $f(0) = 0$ et $y_0 \in \mathbb{R}^n$. Nous supposons que l'application linéaire df_0 n'a que des valeurs propres dont la partie réelle est strictement négative.

Alors ItemGZEAooAhxuDQi

(1) Il existe $k > 0$ tel que si $\|y_0\| < k$ alors la solution maximale est définie sur $\mathbb{R}$, ItemGZEAooAhxuDQii

(2) pour le même nombre $k > 0$, si $\|y_0\| < k$ alors $y(t) \xrightarrow{t\to\infty} 0$ exponentiellement vite,

(3) la solution $y = 0$ est un point d'équilibre attractif.

Démonstration. Placer ici une phrase intelligente[9].

(1) **Prolégomène**
Le théorème de Cauchy-Lipschitz 17.39 nous enseigne que l'équation différentielle considérée possède une unique solution maximale (entre autres parce qu'une fonction de classe C^1 est localement Lipschitz) et nous nommons J l'intervalle sur lequel elle est définie.

(2) **Système linéarisé**
Nous posons $A = df_0$. La fonction $y_L(t) = e^{tA}y_0$ est solution du système linéarisé

$$\begin{cases} y'(t) = Ay(t) & (32.245a) \\ y(0) = y_0. & (32.245b) \end{cases}$$

Pour évaluer la norme de y_L nous utilisons le lemme 15.154 : il existe un polynôme P tel que

$$\|y_L(t)\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{\operatorname{Re}\lambda_i t}\|y_0\|. \tag{32.246}$$

Mais par hypothèse, $\operatorname{Re}(\lambda_i) < 0$ et si nous posons $\lambda = \max\{\operatorname{Re}(\lambda_i)\}$ nous avons $\lambda < 0$ et

$$\|y_L(t)\| \leqslant P\big(|t|\big)e^{\lambda t}\|y_0\|. \tag{32.247}$$

Donc quel que soit y_0 nous avons $\lim_{t\to\infty}\|y_L(t)\| = 0$ c'est-à-dire $\lim_{t\to\infty} y_L(t) = 0$.

9. Parce que sinon l'environnement `description` qui suit donne un mauvais effet.

(3) **Une forme linéaire**

Nous définissons la forme bilinéaire suivante sur $\mathbb{R}^n$:

$$b(x,y)=\int_0^\infty \langle e^{tA}x,e^{tA}y\rangle dt. \tag{32.248}$$

D'abord cela est bien défini pour tout $x,y\in\mathbb{R}^n$ parce que

$$\big|\langle e^{tA}x,e^{tA}y\rangle\big|\leqslant \|e^{tA}x\|\|e^{tA}y\|\leqslant P_1\big(|t|\big)P_2\big(|t|\big)e^{2\lambda t}\|x\|\|y\|, \tag{32.249}$$

qui est intégrable entre 0 et ∞ à cause de la décroissance exponentielle [10]. Montrons que b est définie positive. Soit donc $x\neq 0$ et calculons

$$b(x,x)=\int_0^\infty \|e^{tA}x\|^2dt. \tag{32.250}$$

Ce qui est dans l'intégrale est forcément (pas strictement) positif pour tout t. Mais si $x\neq 0$ alors $\|x\|^2$ est strictement positif et sur un voisinage de $t=0$ nous avons aussi $\|e^{tA}x\|^2$ qui est strictement positif. Ergo $b(x,x)>0$ dès que $x\neq 0$, ce qui signifie que b est strictement définie positive (lemme 8.234).

Nous notons $q\colon V\to\mathbb{R}$ la forme quadratique associée à b et aussi la norme qui va avec : $\|x\|_q=\sqrt{q(x)}$. En ce qui concerne le gradient $\nabla q\colon V\to V$, nous avons le petit calcul suivant[737] qui se base sur une des nombreuses formules du lemme 12.255 [11] :

$$\begin{aligned}\nabla q(x)\cdot y&=\frac{d}{dt}\Big[q(x+ty)\Big]_{t=0} &(32.251a)\\ &=\frac{d}{dt}\Big[q(x)+t^2q(y)+2tb(x,y)\Big]_{t=0} &(32.251b)\\ &=2b(x,y). &(32.251c)\end{aligned}$$

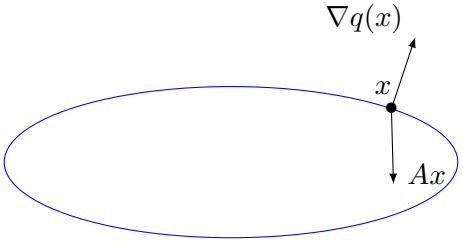
Nous avons aussi

$$\begin{aligned}\nabla q(x)\cdot Ax&=2b(x,Ax) &(32.252a)\\ &=2\int_0^\infty\langle e^{tA}x,e^{tA}Ax\rangle &(32.252b)\\ &=\int_0^\infty\frac{\partial}{\partial t}\Big(\langle e^{tA}x,e^{tA}x\rangle\Big)(t)dt &(32.252c)\\ &=\lim_{T\to\infty}\Big[\langle e^{tA}x,e^{tA}x\rangle\Big]_{t=0}^{t=T}. &(32.252d)\end{aligned}$$

Mais vu que $\|e^{tA}x\|\to 0$, pour $t\to\infty$ il ne reste que terme $t=0$ de la différence, c'est-à-dire

$$\nabla q(x)\cdot Ax=2b(x,Ax)=-\|x\|^2. \tag{32.253}$$ EqUCOGooEFxZSQ

Étant donné que $\nabla q(x)$ est le vecteur dirigé vers l'extérieur de l'ellipsoïde de la courbe de niveau de q au point x, le vecteur Ax est dirigé vers l'intérieur.



10. Proposition 15.113.
11. Le fait que q soit différentiable est simplement le fait que b soit bilinéaire.

(4) **Majoration de $q\big(y(t)\big)'$** Nous posons

$$\begin{aligned} r\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto f(x) - Ax. \end{aligned} \tag{32.254}$$

Soit y la solution maximale au problème (32.244) que nous pouvons aussi écrire sous la forme

$$y'(t) = r\big(y(t)\big) + Ay(t). \tag{32.255}$$

Calculons un peu ...

subeqsZCOLooOzTBLr

$$\begin{aligned} q\big(y(t)\big)' &= b\big(y(t), y(t)\big)' && \text{(32.256a)}\\ &= 2b(y, y') && \text{(32.256b)}\\ &= 2b\big(y, Ay\big) + 2b\big(y, r(y)\big) && \text{(32.256c)}\\ &= -\|y\|^2 + 2b\big(y, r(y)\big) && \text{(32.253) avec } x = y(t) \quad \text{(32.256d)}\\ &\leqslant -\|y\|^2 + 2\|y(t)\|_q \|r\big(y(t)\big)\|_q && \text{Cauchy-Schwarz : } |b(a,b)| \leqslant \|a\|_q\|b\|_q. \quad \text{(32.256e)} \end{aligned}$$

Chacun des deux termes peut encore être majoré. En ce qui concerne le premier, par équivalence des normes [12], il existe une constante C telle que $\|y\| \geqslant C\|y\|_q$. En renommant immédiatement C^2 en C, $\|y\|^2 \geqslant C\|y\|_q^2 = Cq(y)$.
Pour le second, nous allons utiliser la différentiabilité de r et le théorème des accroissements finis. Vu que $df_0 = A$ nous avons $dr_0 = df_0 - A = 0$ et de plus r est de classe C^1 parce que f l'est. Toutes les normes étant équivalentes [13] sur $\mathbb{R}^n$ nous pouvons exprimer la continuité de dr pour la norme $\|.\|_q$: si $\epsilon > 0$ est fixé alors il existe $\alpha > 0$ tel que $\|x\| < \alpha$ implique $\|dr_x\|_q < \epsilon$. Nous pouvons écrire les accroissements finis [14] pour la fonction r :

EqIDTHooCsMSVs

$$\|r(x) - r(0)\|_q \leqslant \sup_{a\in[0,x]} \|df_a\|\|x\|_q. \tag{32.257}$$

La chose facile à remarquer est que $r(0) = f(0) = 0$. En ce qui concerne les choses difficiles, vu que dr est continue (parce que r est C^1) il existe un $\delta > 0$ tel que $\|dr_a\|_q < \epsilon$ dès que $a \in B_q(0, \delta)$. Si nous prenons $\|x\|_q < \delta$ alors cette majoration est valable pour tous les éléments sur lequel est pris le supremum dans la formule (32.257). Donc

$$\|r(x)\|_q \leqslant \epsilon\|x\|_q \tag{32.258}$$

tant que $\|x\|_q \leqslant \delta$. Par conséquent, tant que $\|y(t)\|_q \leqslant \delta$ nous avons $\|r\big(y(t)\big)\| \leqslant \epsilon\|y(t)\|_q$. Nous continuons le calcul (32.256) :

$$\begin{aligned} q\big(y(t)\big)' &\leqslant Cq(y) + 2\epsilon\|y(t)\|_q^2 && \text{(32.259a)}\\ &= -(C - 2\epsilon)q(y). && \text{(32.259b)} \end{aligned}$$

Si ϵ est petit on a $C - 2\epsilon > 0$ et on pose $\beta = C - 2\epsilon$ pour écrire

EqEYJIooHxSBic

$$q\big(y(t)\big)' \leqslant -\beta q\big(y(t)\big) \tag{32.260}$$

tant que $\|y(t)\|_q < \delta$.

(5) **Deux nombres** Nous posons

subeqsFNPJooERJkxO

$$\begin{aligned} t_1 &= \min\{t > 0 \text{ tel que } q\big(y(t)\big) = \delta\} && \text{(32.261a)}\\ t_2 &= \max\{t < 0 \text{ tel que } q\big(y(t)\big) = \delta\}. && \text{(32.261b)} \end{aligned}$$

Le nombre t_1 est bien défini et est bien un minimum. J'en veux pour preuve que si $q\big(y(t_s)\big) = \delta$, on peut prendre le minimum seulement sur les $t \in [0, t_s]$; or par continuité $q\big(y(t)\big) = \delta$ définit un fermé. Bref t_1 est un infimum sur un compact (fermé borné) et donc bien un minimum atteint.

12. Définition 10.46 et théorème 10.49.
13. Théorème 10.49.
14. Théorème 10.292.

(6) **Si $q(y_0) < \delta$ alors $q\big(y(t)\big) < \delta$**

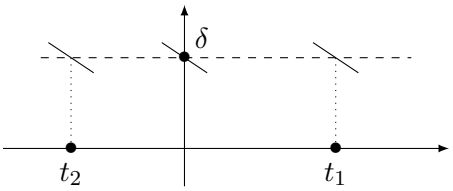
L'inégalité (32.260) est valable pour $t = 0$, $t = t_1$ et $t = t_2$; nous l'écrivons pour t_1 :

$$q\big(y(t)\big)'_{t=t_1} \leqslant -\beta q\big(y(t_1)\big) \leqslant -\beta\delta < 0 \tag{32.262}$$

Nous avons donc $q\big(y(t_1)\big) = \delta$ et $q\big(y(t)\big)'_{t=t_1} < 0$. Par conséquent pour tout t proche de t_1 avec $0 < t < t_1$ il y a $q\big(y(t)\big) > \delta$.

Pour la même raison, prise en $t = 0$ nous avons pour tout t proche de 0 avec $t > 0$ que $q\big(y(t)\big) < \delta$. Par continuité de $t \mapsto q\big(y(t)\big)$ cette fonction doit passer par la valeur δ dans $]t_2, 0[$ et $]0, t_1[$, ce qui contredit la maximalité de t_2 et la minimalité de t_1.

Ci-dessous, une partie de ce à quoi ressemble le graphe de $t \mapsto q\big(y(t)\big)$:



Deux conclusions :

— Vu que $q\big(y(t)\big)$ est borné pour tout $t \in \mathbb{R}$, nous sommes dans le cas (1) de l'alternative du théorème d'explosion en temps fini 32.20. Donc la solution $y(t)$ existe sur tout $\mathbb{R}$ pourvu que $\|y_0\|$ soit assez petit. Plus précisément par équivalence des normes, il existe un nombre $D > 0$ tel que $\|x\| \geqslant D\|x\|_q$ pour tout x. Si $\|y_0\| \leqslant D\delta$ alors

$$D\|y_0\|_q \leqslant \|y_0\| \leqslant D\delta, \tag{32.263}$$

qui donne immédiatement $\|y_0\|_q \leqslant \delta$, ce qui faut pour faire fonctionner l'existence de $y(t)$ pour tout t.

— Nous pouvons maintenant d'utiliser l'inégalité (32.260) pour tout $t \in \mathbb{R}$ sous la seule hypothèse que $q(y_0) < \delta$ au lieu de $q\big(y(t)\big) < \delta$.

La partie (1) de ce théorème est prouvée ; nous passons au reste à la partie (2). Pour cela nous supposons que $q(y_0) < \delta$.

(7) **À propos de $e^{\beta t}q(y)$**

En sous-entendant la dépendance en t dans y nous avons

$$\Big(e^{\beta t}q(y)\Big)' = \beta e^{\beta t}q(y) + e^{\beta t}q(y)' = e^{\beta t}\big(\beta q(y) + q(y)'\big), \tag{32.264}$$

mais nous avons déjà prouvé que $q(y)' \leqslant -\beta q(y)$ (équation (32.260)), donc

$$\Big(e^{\beta t}q(y)\Big)' \leqslant 0 \tag{32.265}$$

EqEJMEooFKuxTv

(8) **Décroissance exponentielle** Si $t \geqslant 0$, l'inégalité (32.265) donne

$$e^{\beta t}q\big(y(t)\big) \leqslant q(y_0), \tag{32.266}$$

c'est-à-dire

$$q\big(y(t)\big) \leqslant e^{-\beta t}q(y_0) \tag{32.267}$$

lorsque $t \geqslant 0$. Par équivalence des normes, nous avons des nombres D_1 et D_2 tels que

$$D_1\|x\|_q \leqslant \|x\| \leqslant D_2\|x\|_q \tag{32.268}$$

pour tout $x \in \mathbb{R}^n$. Nous avons donc pour tout $t \geqslant 0$ que

$$\|y(t)\| \leqslant D_2\|y(t)\|_q \leqslant D_2\|y_0\|_q e^{-\beta t}. \tag{32.269}$$

Pour rappel, $\beta > 0$, ce qui prouve la partie (2) du théorème.

(9) **Point d'équilibre**
Le point $y = 0$ est point d'équilibre (définition 32.42) parce que $f(0) = 0$, donc $y(t) = 0$ fonctionne. Dans ce cas, $y_0 = 0$.

(10) **Stabilité**
La stabilité est le fait que $\|y(t)\|_q \leqslant \delta$ dès que $\|y_0\|_q \leqslant \delta$.

□

32.8.5 Système proies-prédateurs de Lotka-Volterra

Le système de **Lotka-Volterra** est l'équation différentielle suivante :

$$\begin{cases} x' = ax - bxy & (32.270a) \\ y' = -cy + dxy & (32.270b) \end{cases}$$

où a, b, c, d sont des constantes positives, et avec la condition $x(t_0) > 0$, $y(t_0) > 0$.
En ce qui concerne l'interprétation des équations[738],

(1) $x(t)$ est le nombres de proies,
(2) $y(t)$ est le nombres de prédateurs,
(3) Les proies ont une reproduction rapide qui mène à une croissance exponentielle en absence de prédation (d'où le terme ax).
(4) Au contraire, les prédateurs meurent (ou migrent) rapidement lorsqu'ils n'ont pas de proies et nous supposons une décroissance exponentielle du nombre de prédateurs en l'absence de proies. D'où le terme $-cy$ avec le signe négatif.
(5) Les termes $-bxy$ et dxy sont les termes d'interaction entre les proies et les prédateurs. Ils sont proportionnels à la fréquence de leurs rencontres, lesquelles sont avantageuses pour les prédateurs et problématiques pour les proies.

ThoJHCLooHjeCvT

Théorème 32.44 (Lotka-Volterra[380]).
Soient des constantes positives a, b, c, d et le système d'équations différentielles

$$\begin{cases} x' = ax - bxy & (32.271a) \\ y' = -cy + dxy & (32.271b) \\ x(t_0) > 0, y(t_0) > 0. & (32.271c) \end{cases}$$

Alors

(1) Les solutions sont positives sur leur domaines.
(2) Les solutions existent sur $\mathbb{R}$.
(3) Les solutions sont périodiques.

Démonstration. Nous divisons la preuve.

(1) **Comment le théorème de Cauchy-Lipschitz s'applique** Le théorème de Cauchy-Lipschitz [15] ne peut pas s'appliquer tel quel parce qu'il demande une condition initiale pour avoir unicité.
En ce qui concerne les notations, ce qui est noté « y » dans le théorème est ici le couple x, y et la fonction f est alors

$$f\left(t, \begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} ax - bxy \\ -cy + dxy \end{pmatrix}. \quad (32.272)$$

C'est une fonction continue localement Lipschitz partout par le lemme 12.323 et la proposition 12.322.
Nous savons cependant que les solutions sont de classe C^1 et que moyennant la donnée d'une condition initiale, la solution est unique.

15. Théorème 17.39.

(2) **Les solutions restent positives** Supposons $x(s) = 0$ pour un certain $s > t_0$. Alors la solution

$$\begin{cases} x(t) = 0 & (32.273a) \\ y(t) = \exp(-ct) & (32.273b) \end{cases}$$

est une solution pour $[t_0, s + \epsilon]$. Par unicité de la solution avec condition initiale $x(s) = 0$, nous avons aussi $x(t_0) = 0$ pour toutes les solutions, ce qui contredit notre condition.

De la même façon, avoir $y(s) = 0$ donne une solution avec $y(t) = 0$ pour tout t et donc une contradiction.

(3) **Solutions sur $\mathbb{R}$**

Nous montrons maintenant que les solutions sont définies sur $\mathbb{R}$.

Nous avons $x' < ax$, donc pour tout t où la solution est définie,

$$0 < x(t) < x(t_0)e^{a(t-t_0)}, \tag{32.274}$$

c'est-à-dire que la solution ne peut pas exploser en temps fini [16] : elle est bornée par le haut et le bas. Elle doit donc exister pour tout $t \in \mathbb{R}$. Par ailleurs, $y' < dxy$ donc

$$0 < y(t) < y(t_0)e^{d\int_{t_0}^{t} x(s)ds} \tag{32.275}$$

qui est également contraire à l'explosion en temps fini.

(4) **4 zones : monotonie**

Nous divisons $\mathbb{R}^2$ en quatre zones d'après les signes de $a - by$ et $c - dx$. Nous montrons que dans chacune de ces zones, les solutions sont monotones. Prenons par exemple la partie

$$\{(x, y) \in \mathbb{R}^2 \text{ tel que } a - by > 0\} \times \{c - dx < 0\}. \tag{32.276}$$

Vu l'équation $x' = x(a - by)$, tant que $\big(x(t), y(t)\big)$ est dans cette zone, la fonction x' a le signe de x et est donc positive. Donc x est croissante dans cette zone.

De la même façon, $y' = -y(c - dx)$, et y' a un signe constant dans la zone.

(5) **4 zones : on bouge**

Nous prouvons à présent qu'une solution ne reste pas dans une zone.

(5a) Supposons que $\big(x(t_0), y(t_0)\big)$ soit dans la zone

$$\begin{aligned} \{a - by > 0\} \quad & \times \quad \{c - dx > 0\} & (32.277a) \\ x' > 0 \quad & \qquad\quad y' < 0 & (32.277b) \end{aligned}$$

et que la solution reste dans cette zone (pour les $t > t_0$). Nous avons en particulier $x' > 0$, donc x est croissante tout en ayant la borne supérieure $x < c/d$. Par conséquent x a une limite que nous appelons $x_1 \in [0, \frac{c}{d}]$.

De la même façon, ; y est décroissante et bornée vers le bas par zéro. Donc y a une limite que nous notons $y_1 \in [0, y(t_0)]$.

Vu que x est bornée et de classe C^1 nous avons forcément $\lim_{t\to\infty} x'(t) = 0$. Mais vu que $x' = ax - bxy$ nous devons avoir

$$ax_1 - bx_1y_1 = 0. \tag{32.278}$$

Mais ni $x_1 > 0$ donc $a - by_1 = 0$, ce qui donne $y_1 = \frac{a}{b}$ et aussi $x_1 = \frac{c}{d}$. Bref, y est décroissante et tend vers a/b ; donc $y(t_0) > a/b$, ce qui contredit que $y(t_0)$ soit dans la zone considérée.

Étant donné que $x' > 0$ et $y' < 0$, la solution sort de la zone pour entrer dans la zone ...

16. Voir le corolaire 32.20.

(5b) Supposons que $\big(x(t_0), y(t_0)\big)$ soit dans la zone

$$\{a - by > 0\} \quad \times \quad \{c - dx < 0\} \tag{32.279a}$$
$$x' < 0 \qquad\qquad y' > 0 \tag{32.279b}$$

et que la solution reste dans cette zone (pour les $t > t_0$). Les fonctions x et y sont convergentes. Par conséquent $\ln(y)$ converge aussi et vu que x est croissante,

$$\frac{y'}{y} = -c + dx \geqslant -x + dx(t_0) > 0 \tag{32.280}$$

Cela signifie que $\ln(y)'$ est toujours positive et bornée par le bas. Cela est impossible si y est borné.

Donc on sort de la zone pour entrer dans ...

(5c) Supposons que $\big(x(t_0), y(t_0)\big)$ soit dans la zone

$$\{a - by < 0\} \quad \times \quad \{c - dx < 0\} \tag{32.281a}$$
$$x' < 0 \qquad\qquad y' > 0 \tag{32.281b}$$

et que la solution reste dans cette zone (pour les $t > t_0$).

Le même type de raisonnement fait passer à la zone...

(5d) Supposons que $\big(x(t_0), y(t_0)\big)$ soit dans la zone

$$\{a - by < 0\} \quad \times \quad \{c - dx > 0\} \tag{32.282a}$$
$$x' < 0 \qquad\qquad y' < 0 \tag{32.282b}$$

et que la solution reste dans cette zone (pour les $t > t_0$). Encore une fois, cela nous fait sortir de la zone et retourne vers la première zone.

À ce moment nous voyons déjà que la relation entre proies et prédateurs, c'est un peu le mythe de Sisyphe...

(6) **Une intégrale première**

Posons la fonction

$$H(x, y) = by + dx - a\ln(y) - c\ln(x). \tag{32.283}$$

Une simple dérivation montre que $t \mapsto H\big(x(t), y(t)\big)$ est constante. Nous considérons la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ s &\mapsto H\big(\frac{c}{d}, s\big) \end{aligned} \tag{32.284}$$

dont la dérivée n'est autre que $f'(s) = b - \frac{a}{s}$. La fonction f est donc décroissante sur l'intervalle $[\frac{a}{b}, \infty[$ et donc injective. Sur les changements de zones, il existe un t_0 tel que

$$x(t_0) = \frac{d}{c} \tag{32.285a}$$
$$y(t_0) > 0. \tag{32.285b}$$

Pour cette valeur t_0 nous avons alors $H\big(x(t_0), y(t_0)\big) = f\big(y(t_0)\big)$. En posant $s_0 = y(t_0) > 0$ nous avons

$$H(x_0, y_0) = f(s_0) \tag{32.286}$$

et f étant injective, ce s_0 est la seule valeur de s à vérifier $H(x_0, y_0) = f(s)$.

(7) **Conclusion**

La fonction x passant d'une zone à l'autre, il existe un $t_1 > t_0$ tel que $x(t_1) = a/b$. Nous avons évidemment

$$H\big(x(t_1), y(t_1)\big) = H(x_0, y_0) \tag{32.287}$$

parce que H est constante le long du mouvement. Cela se traduit par

$$H\big(\frac{a}{b}, y(t_1)\big) = f(s_0), \tag{32.288}$$

et donc $y(t_1) = f(s_0) = y(t_0)$. Avec tout cela nous avons

$$\begin{cases} y(t_1) = y(t_0) & \text{(32.289a)} \\ x(t_1) = x(t_2) = \dfrac{a}{b}. & \text{(32.289b)} \end{cases}$$

Cela est donc un point par lequel la solution repasse. Par unicité de la solution, elle est donc périodique.

□

32.9 Équation du second ordre

32.9.1 Wronskien

Nous considérons ici une équation différentielle de la forme

$$y''(t) + q(t)y(t) = 0 \tag{32.290}$$

Dans ce point nous allons considérer la fonction q sans hypothèse de périodicité. L'équation de Hill (sous-section 32.9.4) sera la même équation, mais en supposant que q est périodique.

Nous commençons par argumenter que si q est continue, alors l'ensemble des solutions de l'équation (32.290) est un espace vectoriel de dimension deux. Pour cela il suffit d'appliquer la méthode de réduction de l'ordre (section 32.7) puis le théorème de dimension pour les systèmes linéaires (théorème 32.16). En effet si la fonction y_1 est solution de (32.290) si et seulement si le vecteur $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ est solution du système linéaire

$$Y'(t) = \begin{pmatrix} 0 & 1 \\ -q(t) & 0 \end{pmatrix} Y(t). \tag{32.291}$$

Soient deux solutions y_1 et y_2 de l'équation différentielle. Le **Wronskien** de ces deux solutions est le déterminant

$$W(t) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}. \tag{32.292}$$

Si nous considérons l'équation différentielle

$$y'' + py' + qy = 0, \tag{32.293}$$

le Wronskien peut être déterminé sans savoir explicitement y_1 et y_2 parce que $W = y_1y_2' - y_1'y_2$, et en dérivant,

$$\begin{aligned} W' &= y_1y_2'' + y_1'y_2' - y_1''y_2 - y_1'y_2' & \text{(32.294a)} \\ &= y_1(-py_2' - qy_2) - (-py_1' - qy_1)y_2 & \text{(32.294b)} \\ &= -p\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}, & \text{(32.294c)} \end{aligned}$$

c'est-à-dire

$$W' = -pW. \tag{32.295}$$

Il suffit donc de savoir une condition initiale pour obtenir une équation différentielle pour W.

32.9.2 Avec second membre

Une équation différentielle du second ordre avec un second membre se présente sous la forme

$$ay''(t) + by'(t) + cy(t) = v(t) \tag{32.296}$$

où $v(t)$ est une fonction donnée. Le truc est de commencer par résoudre l'équation différentielle sans second membre, c'est-à-dire trouver la fonction $y_H(t)$ telle que

$$ay''_H(t) + by'_H(t) + cy_H(t) = 0. \tag{32.297}$$

Cela se fait en utilisant la méthode du polynôme caractéristique.

Ensuite, il faut trouver une solution particulière $y_P(t)$ de l'équation avec le second membre. Une seule. Pour y parvenir, il faut du doigté et un peu de technique. Il faut faire des essais en fonction de ce à quoi ressemble le $v(t)$:

(1) Si $v(t)$ est un polynôme, alors il faut essayer un polynôme,
(2) Si $v(t) = \cos(\omega t)$ ou bien $v(t) = \sin(\omega t)$, alors essayer $y_P(t) = A\cos(\omega t) + B\sin(\omega t)$,
(3) Si $v(t) = e^{\omega t}$, alors essayer $y_P(t) = Ae^{\omega t}$.

32.9.3 Équation $y'' + q(t)y = 0$

subsecSyTwyM

Nous allons donner quelques propriétés des solutions de l'équation

$$y'' + qy = 0 \tag{32.298}$$

en fonction de telle ou telle hypothèse sur q.

Proposition 32.45.
Si $q\colon \mathbb{R}^+ \to \mathbb{R}$ est continue et si

$$\int_0^\infty |q(t)|dt \tag{32.299}$$

converge, alors

(1) toute solution bornée de $y'' + qy = 0$ vérifie $\lim_{t\to\infty} y'(t) = 0$,
(2) l'équation $y'' + qy = 0$ admet des solutions non bornées.

Démonstration. Soit y une solution bornée, et intégrons l'équation différentielle entre 0 et ∞ :

$$\int_0^\infty y''(t)dt = -\int_0^\infty q(t)y(t)dt. \tag{32.300}$$

La fonction y étant bornée, l'hypothèse sur q permet de dire que l'intégrale de droite existe. Par ailleurs,

$$\int_0^\infty y'' = \lim_{a\to\infty}\int_0^a y'' = \lim_{a\to\infty} y'(a) - y'(0). \tag{32.301}$$

Cela justifie que la limite $\lim_{t\to\infty} y'(t)$ existe. Posons $\alpha = \lim_{t\to\infty} y'(t)$ et supposons par l'absurde que $\alpha \neq 0$. Soit $\epsilon > 0$ et λ assez grand pour que

$$\|y' - \alpha\|_{[\lambda,\infty[} < \epsilon. \tag{32.302}$$

Soit aussi $x > \lambda$. Nous avons, pour $t \in [\lambda, x]$, que $|y'(t) - \alpha| < \epsilon$. Autrement dit, $\alpha \in]y'(t) - \epsilon, y'(t) + \epsilon[$ et donc que $y'(t) > \alpha - \epsilon$. Nous avons alors

$$y(x) = y(\lambda) + \int_\lambda^x y'(t)dt \tag{32.303a}$$
$$\geqslant y(\lambda) + \int_\lambda^x (\alpha - \epsilon)dt \tag{32.303b}$$
$$= y(\lambda) + (\alpha - \epsilon)(x - \lambda). \tag{32.303c}$$

En prenant la limite des deux côtés on voit que $y(x) \to \infty$ dès que $\alpha \neq 0$, ce qui est contraire aux hypothèses. Donc $\alpha = 0$.

Pour la seconde partie de la proposition, nous devons prouver que l'équation $y'' + qy = 0$ possède des solutions non bornées. Si l'équation a seulement des solutions bornées et si $\{u, v\}$ est une base de solutions, alors nous avons $u', v' \to 0$. Si nous reprenons l'équation (32.295) avec $p = 0$ nous savons que dans notre cas le Wronskien satisfait à $W' = 0$, c'est-à-dire qu'il est constant. Mais vu que u et v sont bornées et que les dérivées tendent vers zéro, nous avons $W(t) \to 0$ et donc $W(t) = 0$.

Or l'annulation identique du Wronskien contredit que $\{u, v\}$ serait une base de solutions. Donc il existe des solutions non bornées. □

PropMYskGa

Proposition 32.46.
Soit l'équation différentielle $y'' + qy = 0$. Si q est C^1, strictement positive et croissante, alors toutes les solutions sont bornées.

Démonstration. Soit y une solution et multiplions l'équation par $2y'$ (qui est non nulle par hypothèse) :

$$2y'y'' + 2qy'y = 0. \tag{32.304}$$

Nous allons intégrer cela en nous souvenant que $2y'y''$ est la dérivée de $(y')^2$. Pour tout $t > 0$ nous avons

$$0 = y'(t)^2 - y'(0)^2 + 2 \underbrace{\int_0^t q(t)y'(t)y(t)dt}_{\text{par partie}} \tag{32.305a}$$

$$= y'(t)^2 - y'(0)^2 + 2 \left([qy^2]_0^t - \int_0^t q'y^2 \right) \tag{32.305b}$$

$$\tag{32.305c}$$

Le terme qui nous intéresse est celui qui contient $y(t)$:

$$2q(t)y(t)^2 = -y'(t)^2 + y'(0)^2 + 2q(0)y(0)^2 + 2\int_0^t q'y^2 \tag{32.306}$$

Nous pouvons majorer $-y'(t)^2$ par zéro et remplacer toutes les constantes par K :

$$q(t)y(t)^2 \leqslant \int_0^t q'y^2 + K = \int_0^t \frac{q'}{q} qy^2. \tag{32.307}$$

C'est le moment d'utiliser le lemme de Grönwall 32.5 avec $\phi = qy^2$ et $\psi = q'/q$. Les hypothèses de croissance et de positivité ont été posées exprès. Bref, on a

$$qy^2 \leqslant K \exp\left(\int_0^t \frac{q'(s)}{q(s)} ds \right) \tag{32.308a}$$

$$= K \exp\left(\ln \frac{q(t)}{q(0)} \right) \tag{32.308b}$$

$$= K \frac{q(t)}{q(0)}. \tag{32.308c}$$

Notons que $q(0)$ est strictement positif. Nous déduisons que

$$y^2 \leqslant \frac{K}{q(0)} \tag{32.309}$$

et donc y est bornée. □

32.9.4 Équation de Hill

SubSecDWwVVPa

L'équation **de Hill** est une équation différentielle de la forme

$$y'' + qy = 0 \tag{32.310}$$

EqPQMyzEZ

où

(1) $q \in C^1(\mathbb{R}, \mathbb{R})$,

(2) q est paire et π-périodique

Nous nous intéressons aux solutions complexes de cette équation différentielle.

Nous nommons $W \subset C^2(\mathbb{R}, \mathbb{C})$ l'espace des solutions complexes de l'équation (32.310). Nous savons par ce qui a été dit en 32.9.3 que cet espace est de dimension deux. De plus avec les hypothèses faites ici sur q, nous savons que les solutions sont de classe C^3 parce que si y est une solution, alors l'équation $y'' = qy$ nous indique que y est C^1 parce que y'' existe (y' est dérivable et donc continue). Mais si y est de classe C^1, alors le membre de droite qy est C^1 et donc y'' est C^1, ce qui prouve que y est de classe C^3. La récurrence ne va pas plus loin parce que q est seulement de classe C^1.

Nous considérons l'application de translation

$$\begin{aligned} T\colon C^2(\mathbb{R}, \mathbb{C}) &\to C^2(\mathbb{R}, \mathbb{C}) \\ (Ty)(x) &= y(x + \pi). \end{aligned} \tag{32.311}$$

En utilisant la règle de dérivation de fonctions composées, $(Ty)' = Ty'$ et $(Ty)'' = Ty''$, de telle sorte que si u est solution de l'équation (32.310), alors Tu est également solution. Donc W est un espace stable par T.

Le théorème 32.16 nous permet de choisir une base de W en imposant des conditions. Nous choisissons une base $\{y_1, y_2\}$ telles que

$$\begin{aligned} y_1(0) &= 1 \quad y_2(0) = 0 \\ y_1'(0) &= 0 \quad y_2'(0) = 1. \end{aligned} \tag{32.312}$$

Le théorème 32.16 nous assure que deux telles solutions existent et qu'elles forment une base de W parce que W est de dimension 2.

IVLzNaU

Lemme 32.47 ([728]).
Avec ce choix de base $\{y_1, y_2\}$ la matrice de T est donnée par

$$T = \begin{pmatrix} y_1(\pi) & y_2(\pi) \\ y_1'(\pi) & y_2'(\pi) \end{pmatrix}. \tag{32.313}$$

De plus la fonction y_1 est paire et la fonction y_2 est impaire.

Démonstration. Cherchons la matrice de T dans cette base en associant $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ à y_1 et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ à y_2. Si $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors

$$Ty_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} = ay_1 + cy_2. \tag{32.314}$$

EqSZhBPCx

En évaluant cela en $t = 0$,

$$(Ty_1)(0) = ay_1(0) + cy_2(0) = a, \tag{32.315}$$

donc $a = (Ty_1)(0) = y_1(\pi)$. En dérivant (32.314), en tenant compte du fait que $(Ty_1)' = Ty_1'$ et en évaluant en $t = 0$, nous trouvons de même $c = y_1'(\pi)$. Puis le même cinéma avec y_2 donne

$$T = \begin{pmatrix} y_1(\pi) & y_2(\pi) \\ y_1'(\pi) & y_2'(\pi) \end{pmatrix}. \tag{32.316}$$

Passons maintenant à la parité de y_1 et y_2. Nous posons $\psi(t) = y_1(-t)$. Alors $\psi'(t) = -y_1'(-t)$ et $\psi''(t) = y_1''(-t)$, tant et si bien que

$$\psi''(t) + q(t)\psi(t) = y_1''(-t) + q(t)y_1(-t) = 0. \tag{32.317}$$

donc ψ est une solution de l'équation. Mais

$$\begin{cases} \psi(0) = y_1(0) & \text{(32.318a)} \\ \psi'(0) = -y_1'(0) = 0, & \text{(32.318b)} \end{cases}$$

donc ψ a les mêmes conditions initiales que y_1. Par conséquent $\psi = y_1$ (par le l'unicité donnée dans le théorème de Cauchy-Lipschitz 17.39) et y_1 est paire. Nous procédons de même en partant de $\varphi(t) = -y_2(-t)$ pour trouver que $\varphi = y_2$ et que donc que y_2 est impaire.

□

Remémorons nous toutefois, pour calmer tout enthousiasme excessif, que T dépend de deux solutions et donc de la fonction q donnée dans l'équation.

PropGJCZcjR

Proposition 32.48 ([104])**.**
Nous considérons l'équation $y'' + qy = 0$ et sa base de solutions $\{y_1, y_2\}$ en suivant les notations données plus haut.

(1) Si $|\operatorname{Tr}(T)| < 2$, alors toutes les solutions de l'équation sont bornées.
(2) Si $|\operatorname{Tr}(T)| = 2$ alors nous avons une solution non bornée.
(3) Si $|\operatorname{Tr}(T)| > 2$ alors toutes les solutions de l'équation sont non bornées.
(4) Le cas $|\operatorname{Tr}(T)| = 2$ se présente si et seulement si $y_1'(\pi)y_2(\pi) = 0$.

Démonstration. Remarquons que le déterminant de la matrice T est égal au Wronskien des solutions y_1 et y_2 calculé en $t = \pi$. Calculons sa valeur :

$$W(y_1, y_2) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} = y_1y_2' - y_1'y_2. \tag{32.319}$$

En dérivant et en remplaçant y_i'' par $-qy_i$, nous trouvons tout de suite $W(y_1, y_2)' = 0$. Donc le Wronskien est constant et il est facile de le calculer en $t = 0$:

$$W(y_1, y_2)(0) = 1 - 0 = 1. \tag{32.320}$$

Donc pour tout t nous avons $W(y_1, y_2)(t) = 1$. En particulier

$$\det(T) = W(y_1, y_2)(\pi) = 1, \tag{32.321}$$

et notons au passage que T est inversible.

Nous écrivons le polynôme caractéristique de T sous la forme $\chi_T = X^2 - \operatorname{Tr}(T)X + \det(T)$, c'est-à-dire

$$\chi_T = X^2 - \operatorname{Tr}(T)X + 1, \tag{32.322}$$

dont le discriminant est $\Delta = \operatorname{Tr}(A)^2 - 4$.

Nous passons à présent aux différents points de la proposition.

(1) Si $|\operatorname{Tr}(T)| < 2$, alors $\Delta < 0$ et χ_T a deux racines complexes conjuguées que nous notons ρ et $\bar{\rho}$. De plus le produit des racines étant le terme indépendant, $\rho\bar{\rho} = 1$; en particulier $|\rho| = |\bar{\rho}| = 1$. Notons $\{u, v\}$ une base de vecteurs propres : $Tu = \rho u$ et $Tv = \bar{\rho}v$. Il est vite vu que la fonction $|u|$ est π-périodique :

$$|u|(t + \pi) = |u(t + \pi)| = |(Tu)(t)| = |(\rho u)(t)| = |\rho||u|(t) = |u|(t). \tag{32.323}$$

La fonction $|u|$ est continue [17] et périodique ergo bornée. La fonction $|v|$ est bornée pour la même raison et par linéarité, toutes les fonctions de W sont bornées.

17. La fonction u elle-même n'est cependant pas garantie d'être périodique.

(2) Si $\mathrm{Tr}(T) = \pm 2$, alors $\Delta = 0$ et χ_T a une racine réelle double[18] qui doit être ± 1. Soit u un vecteur propre de T pour la valeur propre ± 1. Nous avons

$$|u|(t+\pi) = |Tu(t)| = |\pm u(t)|, \tag{32.324}$$

ce qui prouve encore que $|u|$ est périodique et donc bornée.
Notons que nous n'avons pas d'informations sur le fait qu'une autre solution soit ou non bornée.

(3) Si $|\mathrm{Tr}(T)| > 2$, alors χ_T a deux racines réelles distinctes r et r' avec $rr' = 1$ (toujours les relations coefficients-racines). En raison de quoi $r' = r^{-1}$ et quitte à échanger r et r' nous supposons $|r| > 1$. L'opérateur est maintenant diagonalisable et nous considérons $\{u, v\}$ une base de vecteurs propres pour les valeurs propres r et r'. Une solution non nulle de l'équation s'écrit donc sous la forme

$$y = \alpha u + \beta v \tag{32.325}$$

avec $(\alpha, \beta) \neq (0, 0)$.

— Si $\alpha = 0$, alors $\beta \neq 0$ et nous choisissons une valeur t telle que $v(t) \neq 0$. Dans ce cas,

$$y(t + n\pi) = \beta v(t + n\pi) = \beta(T^n v)(t) = \beta(r')^n v(t), \tag{32.326}$$

et en faisant $n \to -\infty$ nous obtenons $\pm\infty$ suivant le signe de β.

— Si $\alpha \neq 0$, alors nous fixons[19] t tel que $u(t) \neq 0$. Alors

$$y(t + n\pi) = \alpha r^n u(t) + \beta(r')^n(t). \tag{32.327}$$

En faisant $n \to \infty$, nous avons $(r')^n \to 0$ tandis que le premier terme tend vers $\pm\infty$ suivant le signe de α.

(4) D'abord le théorème de Cayley-Hamilton 8.120 nous indique que $\chi_T(T) = 0$, c'est-à-dire que

$$T^2 - \mathrm{Tr}(T)T + 1 = 0. \tag{32.328}$$

Nous avons déjà mentionné le fait que T était inversible. Multiplions donc (32.328) par T^{-1} :

$$T + T^{-1} = \mathrm{Tr}(T)\mathbb{1}_2. \tag{32.329}$$

Vu que T^{-1} est l'endomorphisme $T^{-1}u(t) = u(t - \pi)$, sa matrice est donnée par

$$T^{-1} = \begin{pmatrix} y_1(-\pi) & y_2(-\pi) \\ y_1'(-\pi) & y_2'(-\pi) \end{pmatrix} = \begin{pmatrix} y_1(\pi) & -y_2(\pi) \\ -y_1'(\pi) & y_2'(\pi) \end{pmatrix} \tag{32.330}$$

où nous avons utilisé le fait que y_1 était paire et y_2 impaire (lemme 32.47). Si nous notons $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors $T^{-1} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$ et

$$T + T^{-1} = \begin{pmatrix} 2a & 0 \\ 0 & 2b \end{pmatrix}. \tag{32.331}$$

L'équation (32.329) donne alors, vu que $\mathrm{Tr}(T) = a + d$,

$$\begin{pmatrix} 2a & 0 \\ 0 & 2b \end{pmatrix} = \begin{pmatrix} a+d & 0 \\ 0 & a+d \end{pmatrix}, \tag{32.332}$$

ce qui donne immédiatement $a = d$. La matrice de T a donc comme forme $T = \begin{pmatrix} a & b \\ c & a \end{pmatrix}$ et $\mathrm{Tr}(T) = 2a$.
Donc $\mathrm{Tr}(T) = \pm 2$ si et seulement si $a = \pm 1$ et vu que $1 = \det(T) = a^2 - bc$, nous avons $a = \pm 1$ si et seulement si $bc = 0$, ce qui signifie exactement $y_1'(\pi)y_2(\pi) = 0$.

□

18. Ce qui n'implique pas le fait d'avoir deux vecteurs propres pour cette valeur propre, mais tout de même au moins un, voir l'exemple 8.118.

19. Mais pas trop hein ; nous aurons encore besoin d'assigner à t d'autres valeurs dans d'autres théorèmes.

32.10 Différents types d'équations différentielles

32.10.1 Équation homogène

SubSecEqDiffHomo

Une équation différentielle **homogène** est une équation de la forme

$$y' = f(t, y) \tag{32.333}$$

où $f(\lambda t, \lambda y) = f(t, y)$ pour tout $\lambda \neq 0$.

Elle se présente sous la forme

$$y' = \frac{\text{degré } n \text{ en } t, y}{\text{degré } n \text{ en } t, y}, \tag{32.334}$$

avec pas de y' à droite : juste du y et du t.

Lemme 32.49.
L'équation $y' = f(t, y)$ est homogène si et seulement si $f(t, y)$ est une fonction de y/t seulement.

Pour résoudre l'équation homogène, on pose

$$z(t) = \frac{y(t)}{t}, \tag{32.335}$$

EqDiffHomoPoser

donc $tz = y$, et

$$y'(t) = tv'(t) + v(t), \tag{32.336}$$

à remettre dans l'équation de départ.

32.10.2 Équation de Bernoulli

SubSecBernh

C'est une équation du type

$$y' = a(t)y + b(t)y^\alpha \tag{32.337}$$

EqBerNDiffalp

où $\alpha \neq 0$ ou 1. Pour la résoudre, on divise l'équation par y^α, et on pose $u = y^{1-\alpha}$, et on tombe sur une équation linéaire

$$u' = (1 - \alpha)\big(a(t)u + b(t)\big). \tag{32.338}$$

32.10.3 Équation de Riccati

SubSecRicatti

Définition 32.50 ([739]).
C'est une équation de la forme

$$y' = a(t)y^2 + b(t)y + c(t). \tag{32.339}$$

EqDiffGFeneRicatti

En général, on ne peut pas la résoudre, mais si on en connaît *à priori* des solutions particulières, alors on peut s'en sortir.

(1) Si on sait que $y_1(t)$ est une solution, alors on pose

$$y(t) = y_1(t) + \frac{1}{u(t)}, \tag{32.340}$$

et on obtient une équation linéaire

$$u' = -\big(2y_1(t)a(t) + b(t)\big)u - a(t). \tag{32.341}$$

(2) Si y_1 et y_2 sont solutions, alors nous avons y sous forme implicite

$$\frac{y - y_1}{y - y_2} = Ke^{\int a(t)\left(y_1(t) - y_2(t)\right)dt}. \tag{32.342}$$

Pour résoudre une équation de Ricatti, il faut donc d'abord deviner une ou deux solutions.

32.10.4 Équation différentielle exacte

SubSecEqDiffExacte

32.10.4.1 Résolution lorsque tout va bien

Avant de vous lancer dans les équations différentielles exacte, vous devez lire la section sur les formes différentielles 12.21. Une équation différentielle exacte est de la forme $P(t,y)+Q(t,y)y'=0$ que nous allons écrire sous la forme

$$P(t,y)dt+Q(t,y)dy=0. \tag{32.343}$$

EqExacteDiff

Nous savons que si $\partial_y P=\partial_t Q$, alors il existe une fonction $f(t,y)$ telle que $Pdt+Qdy=df$. Pour trouver une telle fonction, nous pouvons simplement intégrer la forme $Pdt+Qdy$. En effet, si $\gamma\colon[0,1]\to\mathbb{R}^2$ est un chemin tel que $\gamma(0)=(0,0)$ et $\gamma(1)=(t,y)$, alors en définissant

$$f(t,y)=\int_\gamma [Pdt+Qdy]=\int_0^1 \big[(P\circ\gamma)(u)+(Q\circ\gamma)(u)\big]\big(\gamma'(u)\big)du, \tag{32.344}$$

nous avons $df=Pdt+Qdy$. N'importe quel chemin fait l'affaire. Calculons avec $\gamma(u)=(tu,yu)$. La dérivée de ce chemin est donnée par

$$\gamma'(u)=t\begin{pmatrix}1\\0\end{pmatrix}+y\begin{pmatrix}0\\1\end{pmatrix}. \tag{32.345}$$

Étant donné que $dt\begin{pmatrix}a\\b\end{pmatrix}=a$ et $dy\begin{pmatrix}a\\b\end{pmatrix}=b$, nous avons

$$\begin{aligned}f(t,y)&=\int_0^1 [Pdt+Qdy]\big(\gamma(u)\big)\left(t\begin{pmatrix}1\\0\end{pmatrix}+y\begin{pmatrix}0\\1\end{pmatrix}\right)du\\&=\int_0^1 P\big(\gamma(t)\big)tdu+\int_0^1 Q\big(\gamma(t)\big)ydu\\&=\int_0^1 \big[tP(tu,uy)+yQ(tu,yu)\big]du.\end{aligned} \tag{32.346}$$

Nous retrouvons exactement la formule (20.318). Si ça t'étonne, c'est que tu n'as pas compris ;) Dans le cas où nous avons la fonction f qui vérifie $P=\partial_t f$ et $Q=\partial_y f$, l'équation (32.343) devient

$$\frac{\partial f}{\partial t}+\frac{\partial f}{\partial y}\frac{dy}{dt}=0, \tag{32.347}$$

c'est-à-dire

$$\frac{d}{dt}\Big[f\big(t,y(t)\big)\Big]=0, \tag{32.348}$$

dont la solution

$$f\big(t,y(t)\big)=C \tag{32.349}$$

donne la solution $y(t)$ sous forme implicite.

32.10.4.2 Facteur intégrant (quand tout ne va pas bien)

Si la forme $Pdt+Qdy$ n'est pas exacte, il n'existe pas de fonction f qui résolve l'affaire. Nous pouvons toutefois essayer de trouver un **facteur intégrant**. Nous cherchons une fonction M telle que

$$(MP)dt+(MQ)dy \tag{32.350}$$

soit exacte. Nous cherchons donc $M(t,y)$ telle que $\partial_y(MP)=\partial_t(MQ)$. En utilisant la règle de Leibniz, nous trouvons l'équation suivante pour M :

$$M(\partial_y P-\partial_t Q)=Q(\partial_t M)-P(\partial_y M). \tag{32.351}$$

EqDuFacteurIntegrant

Cette équation est en générale extrêmement difficile à résoudre, mais dans certains cas particuliers, il est possible d'en trouver une solution à tâtons.

32.11 Distributions pour les équations différentielles

SecTNgeNms

Nous commençons par définir l'espace $C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ en disant que $t \mapsto u_t$ est dans cet espace si

(1) pour tout $t \in \mathbb{R}$ nous avons $u_t \in \mathscr{S}'(\mathbb{R}^d)$,
(2) l'application $t \mapsto u_t$ est de classe C^∞.

Pour définir ce que nous entendons par une fonction de classe C^k à valeurs dans $\mathscr{S}'(\mathbb{R}^d)$ nous nous souvenons de la proposition 30.62.

32.11.1 Équation de Schrödinger

ThoLDmNnBR

Théorème 32.51 (Équation de Schrödinger[104]).
Soit $g \in \mathscr{S}'(\mathbb{R}^d)$ et le problème

$$\begin{cases} \partial_t \tilde{u} - i\Delta\tilde{u} = 0 & \text{(32.352a)} \\ u_0 = g & \text{(32.352b)} \end{cases}$$

EqIKhGuiq

où $\tilde{u} \in C^\infty\big(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d)\big)$ est lié à u par la remarque 30.82. Alors

ItemVFracYji

(1) Il existe une unique solution dans $C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$.

ItemVFracYjiii

(2) Cette solution u vérifie de plus $\tilde{u} \in \mathscr{S}'(\mathbb{R} \times \mathbb{R}^d)$.

Démonstration. Nous allons donner explicitement une fonction $u \in C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ et nous allons vérifier l'équation (32.352a) en testant sur une fonction $\psi \in \mathscr{S}'(\mathbb{R}\times\mathbb{R}^d)$. Cela prouvera le point (2) ainsi que la partie existence de (1). Dans ce qui suit toutes les transformées de Fourier seront par rapport à la variable $x \in \mathbb{R}^d$ ou par rapport à ξ. Jamais par rapport à $t \in \mathbb{R}$.

(1) **Existence** Pour $t \in \mathbb{R}$ nous posons[20]

$$u_t = \mathcal{F}^{-1}(f_t \hat{g}) \tag{32.353}$$

où $f_t \in \mathscr{S}(\mathbb{R}^d)$ est la fonction $f_t(x) = e^{-it\|x\|^2}$. Pour toute fonction $\varphi \in \mathscr{S}(\mathbb{R}^d)$ nous avons

$$u_t(\varphi) = (f\hat{g})\big(\mathcal{F}^{-1}(\varphi)\big) = \hat{g}\big(f\mathcal{F}^{-1}(\varphi)\big) = g\Big(\mathcal{F}\big(f\mathcal{F}^{-1}(\varphi)\big)\Big). \tag{32.354}$$

Le fait que $\mathcal{F}^{-1}(\varphi)$ soit une fonction Schwartz fait partie de la proposition 29.26. Pour chaque t nous avons bien $u_t \in \mathscr{S}'(\Omega)$.
De plus la fonction $h(t,x) = e^{-it\|x\|^2}(\mathcal{F}^{-1}\varphi)(x)$ est dans $C^\infty(\mathbb{R}\times\mathbb{R}^d)$, et par conséquent l'application

$$t \mapsto \hat{g}\big(h(t,.)\big) \tag{32.355}$$

est également C^∞ par la proposition 30.64. Ceci pour dire que $u \in C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$. Il faut encore vérifier que cette fonction est bien une solution de notre problème. Nous testons cette équation sur $\psi \in \mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$. Pour alléger les notations nous posons $\psi_t \colon x \mapsto \psi(t,x)$ et par conséquent aussi $(\partial_t\psi_t)(x) = (\partial_t\psi)(t,x)$. Nous avons :

$$\begin{aligned} \heartsuit &= (\partial_t\tilde{u} - i\Delta\tilde{u})(\psi) && \text{(32.356a)} \\ &= -\tilde{u}(\partial_t\psi) - i\tilde{u}(\Delta\psi) && \text{(32.356b)} \\ &= -\int_{\mathbb{R}} u_t\big((\partial_t\psi_t) + i(\Delta\psi_t)\big)dt && \text{(32.356c)} \end{aligned}$$

Ici nous nous souvenons du lemme 29.29 qui nous dit que nous pouvons permuter $\mathcal{F}^{-1}$ et ∂_t. Et pour l'autre terme il faut utiliser le lemme 29.19 avec $|\alpha| = 2$ et une somme pour obtenir que

$$\widehat{\Delta\varphi}(x) = -\|x\|^2\hat{\varphi}(x), \tag{32.357}$$

20. En utilisant la définition (30.86) du produit d'une distribution par une fonction.

qui dans notre cas s'écrit sous la forme

$$\mathcal{F}^{-1}\Big((\Delta\psi_t)\Big)(x) = -\|x\|^2\mathcal{F}^{-1}\psi(t,x). \tag{32.358}$$

En remettant bout à bout,

$$\begin{aligned}\heartsuit &= -\int_{\mathbb{R}} (f_t\hat{g})\Big((\partial_t - i\|.\|^2)\mathcal{F}^{-1}\psi_t\Big)dt && (32.359a)\\ &= -\int_{\mathbb{R}} \hat{g}\Big(x\mapsto e^{-it\|x\|^2}(\partial_t - i\|x\|^2)(\mathcal{F}^{-1}\psi)(t,x)\Big)dt && (32.359b)\end{aligned}$$

Pour alléger les notations nous notons $\check{\psi}_t(x) = (\mathcal{F}^{-1}\psi)(t,x)$. Nous avons

$$\partial_t\left(e^{-it\|x\|^2}\check{\psi}_t(x)\right) = -i\|x\|^2e^{-it\|x\|^2}\check{\psi}_t(x) + e^{-it\|x\|^2}(\partial_t\check{\psi}_t)(x) = e^{-it\|x\|^2}\big(\partial_t - i\|x\|^2\big)\check{\psi}_t(x); \tag{32.360}$$

cela nous permet d'un peu factoriser une dérivée dans $\heartsuit$:

$$\begin{aligned}\heartsuit &= -\int_{\mathbb{R}} \hat{g}\left(\partial_t\Big(e^{-it\|.\|^2}\check{\psi}_t(.)\Big)\right)dt && (32.361a)\\ &= -\int_{\mathbb{R}} \partial_t\hat{g}\left(e^{-it\|.\|^2}\check{\psi}_t(.)\right)dt && (32.361b)\\ &= -\lim_{N\to\infty}\Big[\hat{g}\Big(e^{-it\|.\|^2}\check{\psi}_t(.)\Big)\Big]_{t=-N}^{t=N}. && (32.361c)\end{aligned}$$

Histoire de bien comprendre les notations, il ne s'agit pas de calculer $\hat{g}\big(e^{-it\|.\|^2}\check{\psi}_t\big)$ pour un t général et de remplacer ensuite t par N et $-N$. En effet la valeur de $\hat{g}\big(e^{-it\|.\|^2}\check{\psi}_t\big)$ pour un t donné est celle qu'on obtient en calculant $\hat{g}(\dots)$ après avoir remplacé t par ce que l'on veut. Par conséquent, en posant $\varphi(t,\xi) = e^{-i\|\xi\|^2}\check{\psi}_t(\xi)$ nous avons :

$$\begin{aligned}\heartsuit &= \lim_{N\to\infty}\left[g\left(x\mapsto\int_{\mathbb{R}} e^{-ix\xi}\varphi(t,\xi)d\xi\right)\right]_{t=-N}^{t=N} && (32.362a)\\ &= \lim_{N\to\infty} g\left(x\mapsto\int_{\mathbb{R}} e^{-ix\xi}\varphi(N,\xi)d\xi\right) - \lim_{N\to\infty} g\left(x\mapsto\int_{\mathbb{R}} e^{-ix\xi}\varphi(-N,\xi)d\xi\right) && (32.362b)\end{aligned}$$

La limite commute avec g parce que cette dernière est une distribution (continue). De plus la limite commute avec l'intégrale parce que ce qui est dedans est Schwartz. La fonction φ étant Schwartz, la limite est nulle. Donc

$$\heartsuit = 0. \tag{32.363}$$

Cela signifie que la fonction u proposée est bien une solution de l'équation de Schrödinger dans $C^\infty(\mathbb{R},\mathscr{S}'(\mathbb{R}^d))$.

(2) **Unicité**

Nous considérons deux solutions $u_1, u_2 \in C^\infty\big(\mathbb{R},\mathscr{S}'(\mathbb{R}^d)\big)$ et la fonction $u = u_1 - u_2$ doit satisfaire au problème

$$\begin{cases}(\partial_t\tilde{u} - i\Delta\tilde{u})(\psi) = 0 & (32.364a)\\ u_0 = 0. & (32.364b)\end{cases}$$

Nous allons montrer que seule la fonction $u_t = 0$ peut satisfaire à cela pour tout $\psi\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$. Nous allons même montrer qu'en imposant ces équations seulement sur la partie de $\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$ qui est à support compact par rapport à $\mathbb{R}$, la seule solution est $u_t = 0$. Soit donc $\psi\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$ à support compact vis-à-vis de sa variable t. Alors

$$0 = -\tilde{u}(\partial_t\psi + i\Delta\psi) = -\int_{\mathbb{R}} u_t\Big((\partial_t\psi_t) + i(\Delta\psi_t)\Big)dt \tag{32.365}$$

où encore une fois $\partial_t\psi_t$ est la fonction $x\mapsto(\partial_t\psi)(t,x)$. Maintenant nous utilisons la proposition 30.85 pour dire que

$$\frac{d}{dt}\Big(u_t(\psi_t)\Big)=u_t^{(1)}(\psi_t)+u_t\left(\frac{\partial\psi}{\partial t}(t,.)\right) \tag{32.366}$$

pour écrire

$$0=-\int_{\mathbb{R}}\frac{d}{dt}\big(u_t(\psi_t)\big)-u_t^{(1)}(\psi_t)+u_t\big(i(\Delta\psi)(t,.)\big)dt \tag{32.367}$$

Le premier terme est facile :

$$\int_{\mathbb{R}}\frac{d}{dt}\Big(u_t(\psi_t)\Big)dt=\lim_{N\to\infty}\Big[u_t(\psi_t)\Big]_{t=-N}^{t=N}=0 \tag{32.368}$$

parce que ψ est à support compact par rapport à t. Nous restons donc avec

$$\int_{\mathbb{R}}u_t^{(1)}(\psi_t)-iu_t\big((\Delta\psi)(t,.)\big)dt=0 \tag{32.369}$$

Nous traitons le terme en $u_t^{(1)}$ en utilisant le fait évident $T(\varphi)=(\mathcal{F}T)(\mathcal{F}^{-1}\varphi)$ et en remarquant le lemme 30.86 :

$$u_t^{(1)}(\psi_t)=(\mathcal{F}u_t^{(1)})(\mathcal{F}^{-1}\psi_t)=(\mathcal{F}u)_t^{(1)}(\mathcal{F}^{-1}\psi_t). \tag{32.370}$$

Pour l'autre terme on fait un peu la même chose en nous souvenant ce que fait la transformée de Fourier en traversant le laplacien :

$$u_t(\Delta\psi_t)=(\mathcal{F}u_t)(\mathcal{F}^{-1}\Delta\psi_t)=(\mathcal{F}u_t)\big(x\mapsto-\|x\|^2(\mathcal{F}^{-1}\psi_t)(x)\big). \tag{32.371}$$

En recollant encore :

$$\int_{\mathbb{R}}(\mathcal{F}u)_t^{(1)}(\mathcal{F}^{-1}\psi_t)+i(\mathcal{F}u_t)\big(\|.\|^2\mathcal{F}^{-1}\psi_t\big)dt=0. \tag{32.372}$$ EqHQGaGpt

Cette équation est valable tant que $\psi\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$ avec support compact en t. Nous allons nous en créer une super cool. D'abord nous choisissons $\varphi\in\mathscr{S}(\mathbb{R}^d)$ et $\chi\in\mathscr{D}(\mathbb{R})$ et nous considérons [21]

$$\psi(t,x)=\mathcal{F}\Big(\xi\mapsto e^{it\|\xi\|^2}\varphi(\xi)\chi(t)\Big)(x). \tag{32.373}$$ EqEVtIcnz

Notons que la transformée de Fourier conserve le fait qu'une fonction soit Schwartz [22], mais pas le fait d'avoir support compact. Cependant nous ne prenons que la transformée de Fourier par rapport à x. Le résultat est donc une fonction ψ qui est Schwartz par rapport à x et support compact par rapport à t. Nous pouvons donc écrire (32.372) en utilisant la fonction (32.373) :

$$0=\int_{\mathbb{R}}(\mathcal{F}u)_t^{(1)}\Big(x\mapsto e^{it\|x\|^2}\varphi(x)\chi(t)\Big)+i(\mathcal{F}u_t)\Big(x\mapsto\|x\|^2e^{it\|x\|^2}\varphi(x)\chi(t)\Big)dt. \tag{32.374}$$ EqHPUvZFz

Là dedans, $\chi(t)$ peut sortir à la fois de la transformée de Fourier et de l'application des distributions ; il doit seulement rester dans l'intégrale. Dans le second terme nous allons utiliser l'égalité (due entre autre à la proposition 30.85) : EqCRGfbLU

$$\begin{aligned}
\frac{d}{dt}\big(\hat{u}_t(e^{it\|.\|^2}\varphi)\big)&=\frac{d}{dt}\Big(u_t\big(\mathcal{F}e^{it\|.\|^2}\varphi\big)\Big) && (32.375a)\\
&=u_t^{(1)}\big(\mathcal{F}e^{it\|.\|^2}\varphi\big)+u_t\left(\frac{\partial}{\partial t}\mathcal{F}e^{it\|.\|^2}\varphi\right) && (32.375b)\\
&=(\mathcal{F}u_t^{(1)})\big(x\mapsto e^{it\|x\|^2}\varphi(x)\big)+(\mathcal{F}u_t)\big(x\mapsto i\|x\|^2e^{it\|x\|^2}\varphi(x)\big) && (32.375c)\\
&=(\mathcal{F}u)_t^{(1)}\big(x\mapsto e^{it\|x\|^2}\varphi(x)\big)+(\mathcal{F}u_t)\big(x\mapsto i\|x\|^2e^{it\|x\|^2}\varphi(x)\big)\ \text{lem. 30.86.} && (32.375d)
\end{aligned}$$

21. Le candidat qui parvient à effectivement présenter ça comme développement, il est fort.
22. Proposition 29.20.

Et là, magie c'est exactement ce qui est dans (32.374). Donc

$$\int_{\mathbb{R}} \frac{d}{dt}\hat{u}_t\big(x \mapsto e^{it\|x\|^2}\varphi(x)\big)\chi(t)dt = 0 \tag{32.376}$$

pour toute fonctions à support compact χ. Donc la proposition 30.1 nous dit que

$$\partial_t \hat{u}_t\big(x \mapsto e^{it\|x\|^2}\varphi(x)\big) = 0. \tag{32.377}$$

C'est zéro partout et non seulement presque partout parce qu'en plus nous avons la continuité. Par conséquent pour tout $t \in \mathbb{R}$ nous avons

$$\hat{u}_t\big(x \mapsto e^{it\|x\|^2}\varphi(x)\big) = \hat{u}_0\big(x \mapsto \varphi(x)\big) = 0. \tag{32.378}$$

Et cela est vrai pour toute fonction $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Nous considérons donc $t_0 \in \mathbb{R}$ et une fonction $\theta \in \mathscr{S}(\mathbb{R}^d)$ pour construire

$$\varphi(x) = e^{-it_0\|x\|^2}\theta(x). \tag{32.379}$$

Nous avons alors $\hat{u}_{t_0}\big(x \mapsto \theta(x)\big) = 0$, ce qui signifie que $\hat{u}_{t_0} = 0$. Du coup pour tout $\theta \in \mathscr{S}(\mathbb{R}^d)$ nous avons $u_{t_0}(\mathcal{F}\theta) = 0$, mais comme la transformée de Fourier est une bijection de $\mathscr{S}(\mathbb{R}^d)$ (proposition 29.26) nous avons en fait $u_{t_0}(\theta) = 0$ pour tout $\theta \in \mathscr{S}(\mathbb{R}^d)$, c'est-à-dire $u_{t_0} = 0$ pour tout $t_0 \in \mathbb{R}$ et au final $u = 0$.

□

32.12 Équations différentielles du premier ordre

Définition 32.52 (Équation différentielle du premier ordre)**.**
*Une **équation différentielle du premier ordre** est une équation qui, sur un intervalle donné, I, décrit la relation entre une variable réelle, notée x ou t dans I, une fonction $y : I \to \mathbb{R}$, et la dérivée première de y qui on note y'.*

Souvent on écrit «$y'(x)$ = une formule contenant x et $y(x)$», c'est à dire

$$y'(x) = f(x, y(x)), \quad \text{pour } x \in I, \tag{32.380}$$

ed_generale

où f est une fonction de deux variables réelles.

Remarque 32.53.
La théorie des fonctions de deux variables ne sera pas abordée dans ce cours, nous allons nous contenter de prendre f dans (32.380) comme une simple notation.

On peut presque toujours omettre d'écrire la dépendance de y en x et écrire simplement (32.380) sous la forme $y' = f(x, y)$.

Définition 32.54 (Solution particulière d'une équation différentielle du premier ordre)**.**
*Une **solution particulière** de l'équation (32.380) sur l'intervalle I est une fonction $z : I \to \mathbb{R}$ telle que :*

(1) z est dérivable sur I ;

(2) $z'(x) = f(x, z(x))$, pour tout $x \in I$.

Définition 32.55 (Solution générale d'une équation différentielle du premier ordre)**.**
*Résoudre une équation différentielle veut dire trouver l'ensemble qui contient toutes ses solutions particulières. Cet ensemble s'appelle **solution générale** de l'équation.*

Exemple 32.56. (1) Résoudre une équation du type $y'(x) = f(x)$ revient à trouver l'ensemble des primitives de la fonction f, qui est donc la solution générale de cette équation. Il y a donc une infinité de solutions particulières, déterminées par une constante additive.
Si $f(x) = \sin(x)$ alors la solution générale sera $\mathcal{Y} = \{-\cos(x) + C \,:\, C \in \mathbb{R}\}$.

(2) L'équation

$$y' = y, \qquad x \in \mathbb{R}, \tag{32.381}$$ equation_exponentielle

a peut-être été abordée dans votre cours de terminale lors de la définition de la fonction exponentielle. Sa solution générale est $\mathcal{Y} = \{Ce^x \,:\, C \in \mathbb{R}\}$. Ici aussi il y a une infinité de solutions particulières.

△

Remarque 32.57.
La solution générale d'une équation différentielle du premier ordre est une famille à un paramètre de fonctions.

Définition 32.58 (Équation différentielle du second ordre)**.**
Une ***équation différentielle du second ordre*** *est une équation qui, sur un intervalle donné, I, décrit la relation entre une variable réelle, notée x ou t dans I, une fonction $y : I \to \mathbb{R}$, et les dérivées première et seconde de y qui on note y' et y'' respectivement.*

On utilise la forme générale

$$y'' = f(x, y, y'), \quad \text{pour } x \in I. \tag{32.382}$$ ed_generale_second_ordre

où f est une fonction de trois variables réelles.

On peut définir de manière analogue les équations différentielles d'ordre supérieur. Les définitions de solution particulière et de solution générale se généralisent aux équations différentielles d'ordre supérieur à un.

Définition 32.59 (Trajectoire)**.**
La trajectoire tracée par une solution particulière y de l'équation (32.380) est le graphe de y en tant que fonction de x.

Exemple 32.60.
Nous allons regarder de plus près l'équation (32.381), $y' = y$, pour tout $x \in \mathbb{R}$. Nous savons que les solutions sont données par

$$\mathcal{Y} = \{Ce^x \,:\, C \in \mathbb{R}\}. \tag{32.383}$$

Soient y_1 et y_2 deux solutions distinctes de cette équation. Les seules solutions qui s'annulent en un point sont les solutions identiquement nulles. Nous supposons n'être pas dans ce cas y_1 et y_2 ne s'annulent pas.

Supposons qu'il existe un point $\bar{x}$ tel que $y_1(\bar{x}) = y_2(\bar{x})$ alors forcément $y_1(\bar{x})/y_2(\bar{x}) = 1$. Le rapport $y_1(\bar{x})/y_2(\bar{x})$ vaut C_1/C_2 et par conséquent $C_1 = C_2$. Ce résultat contredit l'hypothèse que les deux solutions soient distinctes. On a donc montré que *deux trajectoires distinctes de cette équation ne se croisent jamais.*

La figure 32.1 représente quelques trajectoires de l'équation. Si on les avait tracées toutes elles recouvriraient tout le plan. Cela veut dire que *par tout point (x, y) passe une et une seule trajectoire de l'équation (32.381).*

△

Définition 32.61 (Condition initiale)**.**
Une ***condition initiale*** *pour l'équation (32.380) sur l'intervalle I est un point $(\bar{x}, \bar{y}) \in I \times \mathbb{R}$.*

On dit que la solution particulière z de (32.380) satisfait la condition initiale $(\bar{x}, \bar{y}) \in I \times \mathbb{R}$ si $z(\bar{x}) = \bar{y}$.

Définition 32.62 (Problème de Cauchy)**.**
L'association d'une équation différentielle et d'une condition initiale est appelée ***problème de Cauchy***

$$\begin{cases} y' = f(x, y), & x \in I, \\ y(\bar{x}) = \bar{y}. \end{cases} \tag{32.384}$$ pbme_cauchy

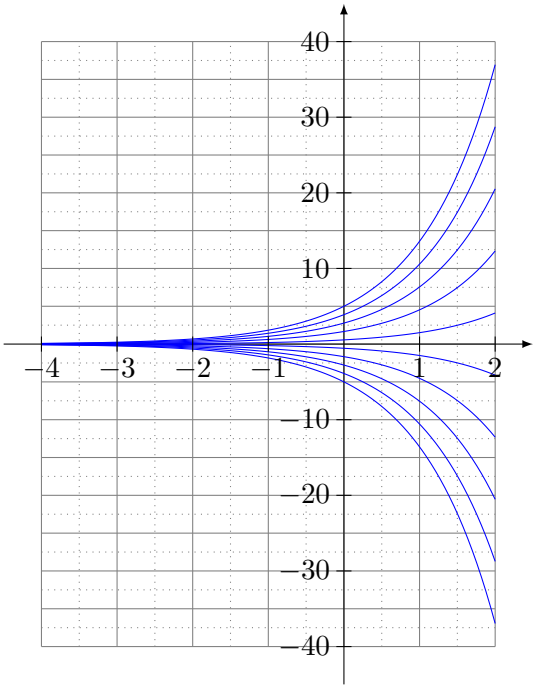


FIGURE 32.1: Quelques trajectoires de l'équation $y' = y$. LabelFigSBTooEasQsT

Remarque 32.63.
Sous des conditions assez générales qui seront toujours vérifiées dans ce cours, tout problème de Cauchy admet une et une seule solution.

Pour passer de la solution générale d'une équation différentielle de premier ordre à une solution particulière il faut choisir une valeur du paramètre. Comme il y a un seul paramètre une seule condition (la trajectoire de la solution doit passer par un point fixe du plan) peut suffire. Pour une équation différentielle de second ordre comme (32.382), nous aurons besoin de plus de conditions. Sans rentrer dans les détails, nous allons constater ce fait dans l'exemple suivant.

Exemple 32.64.
La solution générale de l'équation

$$y'' = -y, \tag{32.385}$$ eq_expcompl

est $\mathcal{Y} = \{C_1 \cos(x) + C_2 \sin(x) \ : \ C_1, C_2 \in \mathbb{R}\}$. Remarquez que l'équation est du second ordre et que sa solution générale est une famille d'équations à deux paramètres réels. Ce sera toujours les cas pour les équations abordées dans la section 32.15. Pour déterminer une solution particulière de (32.385) il faut fixer les valeurs des deux paramètres et donc, en général, il sera nécessaire de donner deux conditions. △

Remarque 32.65.
Une condition comme $y(0) = 4$ nous dit que la constante $C_1 = 4$ mais elle ne nous permet pas de trouver C_2. Il y a donc une infinité de solutions de (32.385) qui satisfont à la condition $y(0) = 4$.

On peut fixer les deux conditions de deux manières différentes.

(1) Problème de Cauchy : on fixe une terne de valeurs réels $\bar{x}, \bar{y}, \bar{y}'$ et on cherche la solution telle que $y(\bar{x}) = \bar{y}$, $y'(\bar{x}) = \bar{y}'$.

Exemple 32.66.
Les conditions $y(0) = 4$, $y'(0) = 15$ permettent de trouver la solution $z(x) = 4\cos(x) + 15\sin(x)$. △

(2) Problème aux bords : on fixe deux points dans le plan x-y, $A = (\bar{x}, \bar{y})$ et $B = (\tilde{x}, \tilde{y})$, et on cherche la solution dont la trajectoire passe par A et B, c'est à dire, on impose $y(\bar{x}) = \bar{y}$,

$y(\tilde{x}) = \tilde{y}$.

Exemple 32.67.
Les conditions $y(0) = 4$, $y(\pi/2) = 15$ permettent de trouver la solution $z(x) = 4\cos(x) + 15\sin(x)$. △

32.13 Premier ordre, variables séparables

Pour certaines équations différentielles la recherche d'une solution particulière se réduit à une recherche de primitive moyennant un changement de variables.

Définition 32.68 (Équation différentielle du premier ordre à variables separables)**.**
Une ***équation différentielle du premier ordre à variables séparables*** *est une équation qui, pour tous les x dans un intervalle donné, I, peut se mettre sous la forme*

$$f(y)y' = g(x), \tag{32.386}$$ eq_var_sep

où f et g sont deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Nous pouvons intégrer les deux côtés de l'égalité par rapport à x et obtenir

$$\int f(y(x))y'(x)\,dx = G(x) + C,$$

où G est une primitive de g et C une constante réelle. Il est facile à ce point d'effectuer un changement de variable dans le membre de gauche de l'équation en posant (sans surprise) $y = y(x)$ et donc $y'(x)\,dx = dy$.

$$\int f(y(x))y'(x)\,dx = \int f(y)\,dy = F(y(x)) + C,$$

où F est une primitive de f et C une constante réelle. En somme nous avons

$$F(y(x)) = G(x) + C,$$

et, si F admet une fonction réciproque, alors

$$y(x) = F^{-1}(G(x) + C). \tag{32.387}$$

Remarque 32.69.
L'expression de F^{-1} peut être difficile à calculer. Il sera alors préférable de garder y dans la forme implicite.

Exemple 32.70.
L'équation

$$3y^2y' = x, \qquad \text{pour tout } x \in \mathbb{R}, \tag{32.388}$$ ex_un_var_sep

est une équation à variables séparables. Pour reprendre les notations du début du chapitre, ici $f(y) = 3y^2$ et $g(x) = x$. En intégrant de deux côtés on trouve

$$y^3 = \frac{x^2}{2} + C.$$

La fonction $F(y) = y^3$ est une bijection de $\mathbb{R}$ dans $\mathbb{R}$, donc nous pouvons écrire la solution générale de l'équation (32.388) dans la forme

$$\mathcal{Y} = \left\{ \left(\frac{x^2}{2} + C \right)^{1/3} \text{ tel que } C \in \mathbb{R} \right\}.$$

△

Exemple 32.71.
En intégrant de deux côtés l'équation à variables séparables

$$2yy' = x, \text{pour tout } x \in \mathbb{R}, \tag{32.389}$$

on trouve

$$y^2 = \frac{x^2}{2} + C.$$

La fonction $F(y) = y^2$ est *n'est pas inversible* sur tout $\mathbb{R}$, et on sait que $\sqrt{y^2} = |y|$. Au moment de rendre y explicite on doit choisir entre

$$y = \left(\frac{x^2}{2} + C\right)^{1/2} \qquad \text{ou} \qquad y = -\left(\frac{x^2}{2} + C\right)^{1/2}.$$

Ce choix se fait suivant la condition initiale, si elle est donnée. Si il n'y a pas de condition initiale nous pouvons écrire que la solution générale est l'ensemble

$$\mathcal{Y} = \left\{ y \,:\, \mathbb{R} \to \mathbb{R} \text{ tels que } y^2 = \frac{x^2}{2} + C \text{ et } C \in \mathbb{R} \right\}.$$

△

Exemple 32.72.
On considère le problème de Cauchy

$$\begin{cases} e^y y' = \frac{1}{x+3}, & x \in]-\infty, -3[, \\ y(-4) = 0. \end{cases} \tag{32.390}$$

En intégrant des deux côtés nous trouvons

$$e^y = \ln(|x+3|) + C.$$

Nous pouvons alors imposer la condition initiale et obtenir $e^0 = \ln(|-4+3|) + C$, c'est à dire $C = 1 - \ln(1) = 1$.

Remarque 32.73.
L'énoncé du problème de Cauchy dit que x peut varier dans $]-\infty, -3[$, mais nous voyons maintenant que la solution n'est pas définie sur toute la demi-droite, parce que e^y est toujours positif et $\ln(|x+3|) + 1$ est positif seulement pour $x < -(1/e+3) \approx -3,3679$.

Donc la solution du problème de Cauchy est $y(x) = \ln(|x+3|) + 1$ pour tout $x \in]-\infty, -(1/e+3)[$. △

exemple_eq_hom

Exemple 32.74.
Attention, cet exemple est le plus important de la section !
On considère l'équation à variables séparables

$$y' = \sin(x)y, \qquad x \in \mathbb{R}. \tag{32.391}$$

EqYNXooFzYeZS

Dans ce cas, pour pouvoir écrire l'équation dans la forme (32.386) il faut pouvoir multiplier les deux côtés par $1/y$. Il faut donc éliminer tout de suite le cas où $y = 0$.

Si $y = 0$ alors $y' = 0$ et on a une solution constante (on dit souvent : une solution stationnaire) de l'équation. Par ailleurs les trajectoires des solutions ne peuvent pas se croiser ; donc si y_G est une solution non nulle de l'équation (32.391) alors $y_G(x) \neq 0$ pour tout x[23]. Il n'y a donc aucun danger à diviser par y dans la recherche d'une solution non identiquement nulle.

23. Ça vaut la peine de prendre un peu de temps pour bien comprendre cela.

Supposons maintenant que $y \neq 0$ et écrivons $y'/y = \sin(x)$. En intégrant des deux côtés on trouve

$$\ln(|y|) = -\cos(x) + C,$$

d'où

$$|y| = e^{-\cos(x)+C} = e^C e^{-\cos(x)}.$$

Si on avait imposé une condition initiale alors on pourrait déterminer une solution particulière de l'équation en choisissant une valeur de la constante C. Nous pouvons observer cependant que la fonction exponentielle est bijective de $\mathbb{R}$ dans $\mathbb{R}^{+,\star}$ et par conséquent il n'y a pas de perte de généralité en disant que la solution générale de l'équation est

$$\mathcal{Y} = \left\{ y \,:\, |y| = Ke^{-\cos(x)}, \text{ pour } K \in \mathbb{R}^{+,\star} \right\} \cup \{y \equiv 0\}.$$

Il n'empêche qu'il serait plus élégant d'écrire la solution générale de l'équation sous une forme plus explicite, sans valeur absolue. Nous pouvons le faire en nous nous rappelant que

$$|x| = \begin{cases} x & \text{si } x \geqslant 0, \\ -x & \text{si } x < 0, \end{cases}$$

Il suffit alors d'autoriser K dans $\mathbb{R}^\star$ pour éliminer la valeur absolue.

Pour écrire la solution générale de façon encore plus compacte nous observons que si $K = 0$ alors $y \equiv 0$, c'est à dire, on retrouve la solution constante nulle.

Finalement, la solution générale de cette équation sera toujours écrite sous la forme suivante

$$\mathcal{Y} = \left\{ y = Ke^{-\cos(x)}, \text{ pour } K \in \mathbb{R} \right\}. \tag{32.392}$$

△

32.14 Équations différentielles linéaires du premier ordre

Définition 32.75 (Équation différentielle linéaire du premier ordre)**.**
Soit $I \subset \mathbb{R}$ un intervalle .

*Une **équation différentielle linéaire du premier ordre** est une équation différentielle de la forme*

$$a(x)y' + b(x)y = c(x), \quad \text{pour } x \in I, \tag{32.393}$$

eq_lin_ordre_un

où a, b, c sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et $a \neq 0$ pour tout $x \in I$.

On dit que a, b, c sont les coefficients de l'équation (32.393).

remarque_lineaire

Remarque 32.76.
Une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite *linéaire* si pour tout x_1, x_2 dans $\mathbb{R}$ et pour tout couple de constantes λ et μ on a

$$f(\lambda x_1 + \mu x_2) = \lambda f(x_1) + \mu f(x_2). \tag{32.394}$$

eq_linearite

Ces équations différentielles sont dites linéaires parce que la partie de l'équation qui contient y (le membre de gauche) satisfait la propriété (32.394) par rapport à y. En effet par les propriétés de la dérivée nous avons que

$$a(x)(\lambda y_1 + \mu y_2)' + b(x)(\lambda y_1 + \mu y_2) = \lambda(a(x)y_1' + b(x)y_1) + \mu(a(x)y_2' + b(x)y_2).$$

Définition 32.77.
*L'équation (32.393) est dite **homogène** quand c est la fonction nulle. Si (32.393) n'est pas homogène on dit que l'équation*

$$a(x)y' + b(x)y = 0, \tag{32.395}$$

eq_lin_ordre_un_hom

*est son **équation homogène associée**.*

Toute équation linéaire du premier ordre homogène est une équation du premier ordre à variables séparables, comme nous en avons vu l'exemple 32.74. Nous n'allons pas répéter les détails du procédé pour trouver sa solution générale, qui aura la forme suivante

À retenir 32.78

$$\mathcal{Y}_h = \left\{ K e^{-\int \frac{b(x)}{a(x)}\,dx} \;:\; K \in \mathbb{R} \right\}. \tag{32.396}$$

solgeneqlinordre1

Proposition 32.79. *(1) Soit y_p une solution particulière de l'équation (32.393) et y_h une solution particulière de l'équation homogène associée (32.395). Alors la fonction somme $z = y_p + y_h$ est encore une solution particulière de l'équation (32.393).*

(2) Soient y_1 et y_2 deux solutions particulières de (32.393). Alors la fonction différence $w = y_1 - y_2$ est une solution particulière de (32.395).

Démonstration. (1)

$$a(x)\,(y_p + y_h)' + b(x)\,(y_p + y_h) - c(x) = \big(a(x)y_p' + b(x)y_p - c(x)\big) + \big(a(x)y_h' + b(x)y_h\big) = 0. \tag{32.397}$$

(2)

$$a(x)\,(y_1 - y_2)' + b(x)\,(y_1 - y_2) = \big(a(x)y_1' + b(x)y_1 - c(x)\big) - \big(a(x)y_2' + b(x)y_2 - c(x)\big) = 0. \tag{32.398}$$

□

Cette proposition permet de démontrer le théorème suivant, qui est le plus important de cette section.

Théorème 32.80.
Soit y_p une solution particulière de l'équation (32.393) et $\mathcal{Y}_h$ la solution générale de l'équation (32.395), alors la solution générale de l'équation (32.393) est l'ensemble

$$\mathcal{Y} = \mathcal{Y}_h + y_p = \{ z = y_h + y_p \;:\; y = h \in \mathcal{Y}_h \}. \tag{32.399}$$

À retenir 32.81
La résolution d'une équation différentielle linéaire du premier ordre comporte trois étapes :

(1) résolution de l'équation homogène associée ;

(2) recherche d'une solution particulière de l'équation non homogène ;

(3) somme de la solution générale de l'équation homogène et de la solution particulière trouvée au point précédent.

La partie qui nous manque encore est de savoir comment trouver une solution particulière de l'équation non homogène (32.393). Si la fonction c dans (32.393) est une constante ou un polynôme simple, ou une exponentielle alors on peut essayer de deviner. Cette méthode cependant n'est pas la plus sure pour des débutants.

Exemple 32.82.
On considère l'équation

$$y' - 5y = 10, \qquad x \in \mathbb{R}. \tag{32.400}$$

Comme tous les coefficients de l'équation sont constants on peut essayer de trouver une solution constante.

Toutes les fonctions constantes ont une dérivée nulle, par conséquent, si une solution constante existe elle doit satisfaire $-5y = 10$, ce qui veut dire que la solution constante est $y(x) \equiv -2$. △

Exemple 32.83.
On considère l'équation

$$xy' + y = x + 1, \qquad x \in \mathbb{R}^{+,\star}. \tag{32.401}$$

Comme le membre de droite de l'équation est un polynôme de degré un on cherche une solution de la forme $y(x) = Ax + B$ avec A et B dans $\mathbb{R}$.

Par substitution on obtient $Ax + (Ax + B) = x + 1$, c'est-à-dire que une solution particulière de l'équation est $y(x) = x/2 + 1$. △

Exemple 32.84.
L'équation

$$xy' - y = x + 1, \qquad x \in \mathbb{R}^{+,\star}. \tag{32.402}$$

ressemble beaucoup à celle de l'exemple précédent, cependant il n'existe pas un polynôme de degré un qui en soit solution.

Dans un cas comme celui-ci, il faut rapidement abandonner la divination et replier sur la méthode, plus technique mais plus sure, dite *variation de la constante*. △

32.14.1 Méthode de variation de la constante

— Soit $\mathcal{Y}_h$ la solution générale de l'équation homogène associé à (32.393). Il s'agit d'une famille à un paramètre de fonctions. La première étape de cette méthode consiste à construire un candidat solution particulière y_p en remplaçant le paramètre dans $\mathcal{Y}_h$ par une fonction $C : \mathbb{R} \to \mathbb{R}$ à déterminer.

Exemple 32.85.
L'équation homogène associée à $y' - y = \cos(x)$ est $y' - y = 0$, dont la solution générale est $\mathcal{Y}_h = \{Ce^x \ : \ C \in \mathbb{R}\}$. Le candidat solution sera alors $y_p = C(x)e^x$, avec C fonction à déterminer. △

— La deuxième étape de cette méthode consiste à injecter y_p dans l'équation. Cela permet de trouver une équation différentielle à variables séparables pour C, en principe plus facile à résoudre que l'équation de départ.

Exemple 32.86.
On continue avec l'exemple précédent. On a $y_p' = C'(x)e^x + C(x)e^x$, d'où

$$(C'(x)e^x + C(x)e^x) - C(x)e^x = \cos(x),$$

c'est-à-dire

$$C'(x) = \cos(x)e^{-x}.$$

△

— La troisième étape de la méthode consiste à trouver une solution particulière de l'équation différentielle pour C et, par conséquent déterminer une y_p.

Exemple 32.87.
La solution générale de

$$C'(x) = \cos(x)e^{-x}.$$

est $\mathcal{C} = \left\{ e^{-1\,\frac{(\sin(x)-\cos(x))}{2}} + K \ : \ K \in \mathbb{R} \right\}$. Il nous suffit une solution particulière, nous pouvons donc choisir $K = 0$ et alors la solution particulière de (32.393) sera $y_p(x) = \frac{\sin(x)-\cos(x)}{2}$. △

Remarque 32.88.
Le plus souvent en intégrant l'équation pour C on en trouvera la solution générale. Dans ce cas on peut remplacer C par cette solution générale et obtenir d'un seul coup la solution générale de l'équation (32.393) , c'est-à-dire sans faire la somme entre la solution générale de l'homogène associée et la solution particulière.

Exemple 32.89.
Dans l'exemple qu'on vient de voir la solution générale de (32.393) est

$$\mathcal{Y} = \mathcal{Y}_h + y_p = \left\{ Ce^x + \frac{(\sin(x) - \cos(x))}{2} \ : \ C \in \mathbb{R} \right\}. \tag{32.403}$$

On obtient le même résultat est écrivant $\mathcal{Y} = \left\{ e^{-x} \left(e^{-1} \frac{(\sin(x)-\cos(x))}{2} + K \right) \ : \ K \in \mathbb{R} \right\}$. Notez qu'on a changé le nom du paramètre de C à K seulement pour souligner qu'on obtient de même résultat par deux chemins différents, sinon les deux expressions sont équivalentes ! △

32.15 Équations différentielles linéaires du second ordre

Secordredeux

Définition 32.90 (Équation différentielle linéaire du second ordre)**.**
Une ***équation différentielle linéaire du second ordre*** *est une équation différentielle de la forme*

$$a(x)y'' + b(x)y' + c(x)y = d(x), \quad \text{pour } x \in I, \tag{32.404}$$

eq_lin_ordre_deux

où a, b, c et d sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et $a \neq 0$ pour tout $x \in I$.

On dit que a, b, c et d sont les coefficients de l'équation (32.404).

Dans ce cours nous allons étudier exclusivement le cas où a, b et c sont des fonctions constantes.

Définition 32.91 (Équation différentielle linéaire du second ordre homogène)**.**
Une ***équation différentielle linéaire du second ordre homogène*** *est une équation différentielle de la forme (32.404), telle que le coefficient d est nul.*

À toute équation de la forme (32.404) on peut associer une équation homogène exactement comme on a fait dans la section précédente pour les équations linéaires du premier ordre.

32.15.1 Équations différentielles linéaires du second ordre homogènes à coefficients constants

Remarque 32.92.
L'application qui à la fonction y fait correspondre $a(x)y'' + b(x)y' + c(x)y$ est linéaire, au sens de la remarque 32.76.

Cela nous dit en particulier, que si y_1 et y_2 sont deux solutions de l'équation homogène alors toute leur combinaison de la forme $z = \lambda y_1 + \mu y_2$, avec λ et μ dans $\mathbb{R}$, est encore une solution.

> Jusqu'ici nous avons toujours travaillé avec des fonctions définies sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. Dans cette section nous nous autorisons à passer par des fonctions définies sur $\mathbb{R}$ et à valeurs dans $\mathbb{C}$, mais cela sera uniquement une étape dans nos calculs. Au final toutes les solutions que nous allons considérer sont des fonctions à valeurs dans $\mathbb{R}$.

La solution générale **à valeurs dans les complexes** d'une équation de ce type a la forme

$$\mathcal{Y}_h^{\mathbb{C}} = \{ C_1 e^{r_1 x} + C_2 e^{r_2 x} \ : \ C_1, C_2 \in \mathbb{C}, \ x \in I \}, \tag{32.405}$$

sol_gen_ordre_deux_hom

où r_1 et r_2 sont aussi des nombres complexes. Remarquez que la solution générale est une famille à deux paramètres. Il faut aussi observer que en tout cas l'intervalle I dans lequel varie x est un intervalle dans $\mathbb{R}$, parce que I est une des données du problème.

À partir de cette information nous pouvons, pour toute équation donnée, chercher la solution générale **complexe** par substitution. Il suffit de remplacer y dans l'équation par e^{rx} et chercher les valeurs de r qui nous conviennent.

Si notre équation de départ est

$$ay'' + by' + cy = 0, \quad \text{pour } x \in I, \tag{32.406}$$ eq_lin_ordre_deux_hom

alors la substitution nous donne

$$e^{rx}\left(ar^2 + br + c\right) = 0.$$

Il est connu que la fonction exponentielle ne prend pas la valeur 0, par conséquent ce qui s'annule est le polynôme de degré deux $ar^2 + br + c$. Il est donc très facile de trouver les valeurs de r qu'on pourra utiliser comme r_1 et r_2 dans la solution générale **complexe**.

Si $b^2 - 4ac > 0$ **:** le polynôme admet deux solutions réelles et distinctes, r_1 et r_2 ;

Si $b^2 - 4ac < 0$ **:** le polynôme admet deux solutions complexes conjuguées, $r_1 = \alpha + i\beta$ et $r_2 = \alpha - i\beta$;

Si $b^2 - 4ac = 0$ **:** le polynôme admet une solution réelle double $r = r_1 = r_2$.

Il faut maintenant écrire la solution générale **réelle** de l'équation, qui est celle que nous intéresse vraiment. La façon de l'obtenir est différente dans les trois cas.

Si $b^2 - 4ac > 0$ **:** la solution générale réelle a la même forme que la solution complexe, (32.405), il suffit de prendre les paramètres C_1 et C_2 dans $\mathbb{R}$ plutôt que dans $\mathbb{C}$.

$$\mathcal{Y}_h = \{C_1 e^{r_1 x} + C_2 e^{r_2 x} \,:\, C_1, C_2 \in \mathbb{R},\, x \in I\}, \tag{32.407}$$ sol_gen_reelle_ordre_deux_hom

Si $b^2 - 4ac < 0$ **:** le polynôme admet deux solutions complexes conjuguées, $r_1 = \alpha + i\beta$ et $r_2 = \alpha - i\beta$; Il faut alors utiliser les formules suivantes

$$\begin{aligned} e^{\alpha+i\beta} &= e^{\alpha}(\cos(\beta) + i\sin(\beta)) \\ e^{\alpha-i\beta} &= e^{\alpha}(\cos(\beta) - i\sin(\beta)). \end{aligned} \tag{32.408}$$

La somme $e^{r_1 x} + e^{r_2 x}$, où x est dans $I \in \mathbb{R}$, vaut

$$e^{(\alpha+i\beta)x} + e^{(\alpha-i\beta)x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) + e^{\alpha x}(\cos(\beta x) - i\sin(\beta x)) = 2e^{\alpha x}\cos(\beta x)$$

et la différence $e^{r_1 x} - e^{r_2 x}$ vaut

$$e^{(\alpha+i\beta)x} - e^{(\alpha-i\beta)x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) - e^{\alpha x}(\cos(\beta x) - i\sin(\beta x)) = 2e^{\alpha x}\sin(\beta x).$$

Par ces deux calculs élémentaires nous avons trouvé deux fonctions à valeurs dans $\mathbb{R}$ qui n'ont pas de zéros en commun. Elles sont les génératrices de la famille des solutions réelles de l'équation différentielle (la solution générale)

$$\mathcal{Y}_h = \{e^{\alpha x}\left(C_1\cos(\beta x) + C_2\sin(\beta x)\right) \,:\, C_1, C_2 \in \mathbb{R},\, x \in I\}, \tag{32.409}$$ sol_gen_reelle_ordre_deux_hom_complconj

Si $b^2 - 4ac = 0$ **:** le polynôme admet une solution réelle double $r = r_1 = r_2$. Dans ce cas la solution générale de l'équation est la famille

$$\mathcal{Y}_h = \{(C_1 + C_2 x)e^{rx} \,:\, C_1, C_2 \in \mathbb{R},\, x \in I\}. \tag{32.410}$$ sol_gen_reelle_ordre_deux_hom_doublerac

Pour justifier cette formule nous observons d'abord que toute fonction $x \mapsto Ce^{rx}$, pour $C \in \mathbb{R}$ est une solution de l'équation différentielle (par construction). Ensuite nous utilisons la méthode de variation de la constante. On trouve rapidement que si une fonction de la forme $x \mapsto C(x)e^{rx}$ est une solution alors $C(x)$ est un polynôme de degré au plus 1, c'est-à-dire $C(x) = C_1 + C_2 x$ avec C_1 et C_2 dans $\mathbb{R}$.

32.15.2 Linéaires du second ordre à coefficients constants, non homogènes

Nous ne présentons pas une méthode générale pour la résolution de ces équations. Comme dans le cas des équations différentielles linéaires du premier ordre non homogènes, la solution générale de (32.404) est donnée par la somme d'une solution particulière et de la solution générale de l'équation homogène associée. La recherche d'une solution particulière est facilitée par le fait que les coefficients de (32.404) sont supposés constants, c'est-à-dire que a, b et c sont des fonctions constantes. Il faut essayer de deviner la forme d'une solution particulière à partir de la forme du second membre de l'équation, la fonction d. Si d est un polynôme il faut essayer avec un polynôme du même degré, si d est une exponentielle, par exemple $d(x) = e^{5x}$, on pourra essayer avec un multiple de la même fonction exponentielle, dans l'exemple $f(x) = ke^{5x}$, avec k à determiner. Si d est une combinaison linéaire de sinus et cosinus, comme par exemple $12\cos(x) + 2\sin(x)$, on peut essayer avec $k_1\cos(x) + k_2\sin(x)$.

Exemple 32.93.
On considère l'équation différentielle

$$y'' + 12y' + 36y = -192e^{2x}, \quad x \in \mathbb{R}. \tag{32.411}$$

exemple_non_hom

Son équation homogène associée est

$$y'' + 12y' + 36y = 0, \tag{32.412}$$

exemple_hom_ass

dont le polynôme caractéristique est $r^2 + 12r + 36$. Ce polynôme admet une racine double, qui est -6, par conséquent la solution générale de (32.412) est

$$\mathcal{Y}_h = \left\{(C_1 + C_2x)e^{-6x} \,:\, C_1,\, C_2 \in \mathbb{R},\, x \in \mathbb{R}\right\}.$$

Le membre de droite de (32.411) est une fonction exponentielle, nous allons donc chercher une solution particulière de (32.411) de la forme $f(x) = ke^{2x}$. Par substitution nous trouvons

$$ke^{2x}(4 + 12 \times 2 + 36) = -192e^{2x},$$

ce qui veut dire que k doit être -3.

La solution générale de l'équation (32.411) est donc

$$\mathcal{Y} = \left\{(C_1 + C_2x)e^{-6x} - 3e^{2x} \,:\, C_1,\, C_2 \in \mathbb{R},\, x \in \mathbb{R}\right\}.$$

△

Exemple 32.94.
Nous allons résoudre l'équation

$$y'' + 12y' + 36y = 12\cos(x) + 2\sin(x), \quad x \in \mathbb{R}. \tag{32.413}$$

Cette équation a comme homogène associée l'équation (32.412), comme dans l'exemple précédent. Il nous suffit donc de trouver une solution particulière de (32.411).

Nous pouvons essayer avec $f(x) = k_1\cos(x) + k_2\sin(x)$. Par substitution on trouve

$$\begin{aligned} -\left(k_1\cos(x) + k_2\sin(x)\right) + 12\left(-k_1\sin(x) + k_2\cos(x)\right) + 36\left(k_1\cos(x) + k_2\sin(x)\right) \\ = 12\cos(x) + 2\sin(x) \end{aligned}$$

Cette équation doit être satisfaite pour toute valeur de x, en particulier pour $x = 0$ et $x = \pi/2$. Cela revient à considère séparément les coefficients des fonctions sinus et cosinus. Il faut alors que k_1 et k_2 soient solutions du système

$$\begin{cases} -k_1 + 12k_2 + 36k_1 &= 12, \\ -k_2 - 12k_1 + 36k_2 &= 2. \end{cases}$$

On trouve $k_1 = 396/1369$ et $k_2 = 214/1369$, et la solution générale de notre équation est

$$\mathcal{Y} = \left\{ (C_1 + C_2 x)e^{-6x} + \frac{396}{1369}\cos(x) + \frac{214}{1369}\sin(x) \ : \ C_1,\, C_2 \in \mathbb{R},\, x \in \mathbb{R} \right\}.$$

△

Exemple 32.95.
Nous allons résoudre l'équation

$$y'' + 12y' + 36y = 10x^2 + 3, \quad x \in \mathbb{R}. \tag{32.414}$$

Cette équation a comme homogène associée l'équation (32.412), comme dans l'exemple précédent. Il nous suffit donc de trouver une solution particulière de (32.411).

Nous pouvons essayer avec $f(x) = k_1x^2 + k_2x + k_3$. Par substitution on trouve

$$(2k_1) + 12\left(2k_1x + k_2\right) + 36\left(k_1x^2 + k_2x + k_3\right) = 10x^2 + 3.$$

Pour trouver les bonnes valeurs des coefficients nous devons résoudre le système

$$\begin{cases} 36k_1 & = 10, \\ 24k_1 + 36k_2 & = 0, \\ 2k_1 + 12k_2 + 36k_3 & = 3, \end{cases}$$

ce qui donne $k_1 = 5/18$, $k_2 = -5/27$ et $k_3 = 7/54$. La solution générale de notre équation est

$$\mathcal{Y} = \left\{ (C_1 + C_2 x)e^{-6x} + \frac{5}{18}x^2 - \frac{5}{27}x + \frac{7}{54} \ : \ C_1,\, C_2 \in \mathbb{R},\, x \in \mathbb{R} \right\}.$$

△

32.16 Fonction de Green

Soit l'équation différentielle

$$\begin{cases} y''(x) = g(x) & \text{(32.415a)} \\ y(0) = y(1) = 0 & \text{(32.415b)} \end{cases}$$

pour $x \in]0,1[$ et où g est continue sur $]0,1[$.

Nous définissons la fonction de Green

$$G(x,t) = \begin{cases} t(x-1) & \text{si } 0 \leqslant t \leqslant x \leqslant 1 \\ x(t-1) & \text{si } 0 \leqslant x \leqslant t \leqslant 1, \end{cases} \tag{32.416}$$

et nous allons montrer que

$$y(x) = \int_0^1 G(x,t)g(t)dt \tag{32.417}$$

est l'unique solution.

(1) **Unicité**
Si y_1 et y_2 sont des solutions, alors $y_1'' = y_2''$ et donc $y_1(x) = y_2(x) + ax + b$. Les conditions aux bords donnent alors $0 = y_1(0) = y_2(0) + b = b$. D'où $b = 0$. En imposant $y_1(1) = 0$ nous trouvons alors immédiatement $a = 0$, ce qui donne $y_1 = y_2$.

(2) **Existence**

Il est vite vérifié qu'avec (32.417) nous avons $y(0) = y(1) = 0$ parce que $G(0,t) = G(1,t) = 0$ pour tout t. Nous fixons une valeur pour $x \in]0,1[$ et nous découpons l'intégrale :

$$y(x) = \int_0^x G(x,t)g(t)dt + \int_x^1 G(x,t)g(t)dt. \tag{32.418}$$

Pour calculer $y'(x)$, il faut dériver à la fois à travers l'intégrale et dans la borne. Si vous connaissez une formule pour faire cela, c'est bien pour vous. Nous allons faire ça à la main et poser

$$I(x,y) = \int_0^y t(x-1)g(t)dt. \tag{32.419}$$

La dérivation de I par rapport à x se fait en utilisant le théorème 17.24 :

$$\frac{\partial I}{\partial x}(x,y) = \int_0^y tg(t)dt. \tag{32.420}$$

Pour la dérivation par rapport à y, il s'agit du théorème fondamental de l'analyse, plus précisément le lien primitive et intégrale de la proposition 14.286 :

$$\frac{\partial I}{\partial y}(x,y) = y(x-1)g(y). \tag{32.421}$$

Maintenant nous considérons la fonction $\varphi_I(x) = I(x,x)$. Elle satisfait à

$$\varphi_I'(x) = \frac{\partial I}{\partial x}(x,x) + \frac{\partial I}{\partial y}(x,x) = \int_0^x tg(t) + x(x-1)g(x). \tag{32.422}$$

Le même jeu avec $J(x,y) = \int_y^1 x(t-1)g(t)dt$ donne

$$\varphi_J'(x) = \int_0^x (t-1)g(t)dt - x(x-1)g(x). \tag{32.423}$$

En remettant les bouts ensemble,

$$y'(x) = \int_0^x tg(t)dt + \int_1^x (1-t)g(t)dt. \tag{32.424}$$

Le calcul de la dérivée seconde donne alors

$$y''(x) = xg(x) + (1-x)g(x) = g(x). \tag{32.425}$$

Nous pouvons aussi, sur cette équation, estimer la variation de la solution en termes d'une variation de g. Soit donc une fonction continue δ_g sur $[0,1]$ et $\tilde{g} = g + \delta_g$. Nous considérons l'équation différentielle

$$\begin{cases} \tilde{y}''(x) = \tilde{g}(x) & \text{(32.426a)} \\ \tilde{y}(0) = \tilde{y}(1) = 0. & \text{(32.426b)} \end{cases}$$

Par ce que nous venons de faire, l'unique solution est

$$\tilde{y}(x) = \int_0^1 G(x,t)\tilde{g}(t)dt = \int_0^1 G(x,t)g(t)dt + \int_0^1 G(x,t)\delta_g(t)dt = y(x) + \delta_y(x) \tag{32.427}$$

où δ_y est une fonction continue ainsi définie :

$$\delta_y(x) = \int_0^1 G(x,t)\delta_g(t)dt. \tag{32.428}$$

Supposons que $\|\delta_g\|_\infty = \epsilon$. Alors des majorations donnent

$$|\delta_y(x)| \leqslant \epsilon \int_0^1 |G(x,t)|dt = \epsilon(1-x)\int_0^x tdt + \epsilon x \int_x^1 (1-t)dt = \frac{\epsilon}{2}x(1-x). \tag{32.429}$$

Mais la fonction $x \mapsto x(1-x)$ a son maximum en $x = \frac{1}{2}$, donc nous pouvons donner une majoration indépendante de x :

$$\|\delta_y\|_\infty \leqslant \frac{1}{8}\|\delta_g\|_\infty. \tag{32.430}$$

Notons que la majoration (32.430) en norme uniforme a l'air plus impressionnante, mais la majoration (32.429) donnant une majoration séparée pour chaque x est en réalité plus précise.

Chapitre 33

Équations aux dérivées partielles

33.1 Symbole principal, équation des caractéristiques

Soit l'équation différentielle semi-linéaire d'ordre k

$$\sum_{|\alpha|=k} a_\alpha(x)(\partial^\alpha u)(x) + F\Big(x, u(x), (Du)(x), \dots, (D^{k-1}u)(x)\Big) = 0 \tag{33.1}$$

EQooEZXSooOxZTFj

pour la fonction $u\colon \mathbb{R}^d \to \mathbb{R}$.

DEFooWPGHooNIMAuS

Définition 33.1.
*Le **symbole principal** de l'équation (33.1) est l'application*

$$\begin{aligned} \sigma\colon \mathbb{R}^d \times \mathbb{R}^d &\to \mathbb{R} \\ (x,\xi) &\mapsto \sum_{|\alpha|=k} a_\alpha(x)\xi^\alpha \end{aligned} \tag{33.2}$$

où si $\alpha = (\alpha_1, \dots, \alpha_d)$ et $\xi = (\xi_1, \dots, \xi_d)$ alors $\xi^\alpha = \xi_1^{\alpha_1} \dots \xi_d^{\alpha_d}$.

DEFooYYNOooZlZMxu

Définition 33.2.
*Une **caractéristique** de l'équation (33.1) est une surface S de $\mathbb{R}^d$ donnée par l'équation $\phi(x) = 0$ où ϕ satisfait à*

$$\begin{cases} \sigma\big(x, \nabla\phi(x)\big) = 0 & \text{(33.3a)} \\ \nabla\phi(x) \neq 0 & \text{(33.3b)} \end{cases}$$

pour tout $x \in S$. Ici, σ est le symbole principal (définition 33.1) de l'équation différentielle.

33.2 Méthode des caractéristiques pour l'ordre 1

SECooHKSLooOCYNDz

Nous[740, 741] voulons étudier l'équation d'ordre 1

$$a(x,y)\frac{\partial u}{\partial x}(x,y) + b(x,y)\frac{\partial u}{\partial y}(x,y) + c(x,y)u(x,y) = f(x,y) \tag{33.4}$$

EQooGJKZooCCEpRj

Le champ de vecteurs associé à cette équation est

$$v = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{33.5}$$

et l'équation peut être écrite sous la forme

$$(v \cdot \nabla u) + cu = f. \tag{33.6}$$

DEFooNYVSooZRjNcH

Définition 33.3.
*Le **flot** de ce champ de vecteurs sont les courbes paramétriques* $\gamma(t) = \big(x(t), y(t)\big)$ *vérifiant* $\gamma'(t) = v\big(\gamma(t)\big)$.

Les équations du flot pour l'équation (33.4) sont [EQooURBTooVdFzZR]

$$\begin{cases} x'(t) = a\big(x(t), y(t)\big) & (33.7a)\\ y'(t) = b\big(x(t), y(t)\big). & (33.7b)\end{cases}$$

Ce sont des équations différentielles ordinaires. Un système de deux équations couplées du premier ordre.

Quel est l'intérêt du flot ? Nous allons voir que sur la ligne $t \mapsto \gamma(t)$, la fonction u est constante. Or des solutions γ au système (33.7), il y en aura plusieurs : une pour chaque valeur des constantes d'intégration. Pour peu que ces lignes recouvrent tout le plan, nous pourrons résoudre l'équation de départ ligne par ligne.

Nous posons

$$\tilde{u}(t) = u\big(x(t), y(t)\big) \quad (33.8a)$$
$$\tilde{c}(t) = c\big(x(t), y(t)\big) \quad (33.8b)$$
$$\tilde{f}(t) = f\big(x(t), y(t)\big). \quad (33.8c)$$

La fonction $\tilde{u}$ est une fonction $\mathbb{R} \to \mathbb{R}$ normale qui se dérive normalement, en suivant la règle de dérivation des fonctions composées :

$$\begin{aligned}\tilde{u}'(t) &= \frac{\partial u}{\partial x}\big(x(t), y(t)\big)x'(t) + \frac{\partial u}{\partial y}\big(x(t), y(t)\big)y'(t) && (33.9a)\\ &= a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y} && (33.9b)\\ &= f\big(x(t), y(t)\big) - c\big(x(t), y(t)\big)u\big(x(t), y(t)\big) && (33.9c)\\ &= \tilde{f}(t) - \tilde{c}(t)\tilde{u}(t). && (33.9d)\end{aligned}$$

Nous avons pour $\tilde{u}$ l'équation différentielle ordinaire

$$\tilde{u}' + \tilde{c}\tilde{u} = \tilde{f} \quad (33.10)$$

qui est résolue par la proposition 32.11.

33.2.1 Un exemple complet un peu minimal

Nous considérons l'équation différentielle[740]

$$\frac{\partial u}{\partial x} - \frac{\partial u}{\partial y} - (x - y)u = 0. \quad (33.11)$$

Et nous allons la résoudre.

Les équations du flot, sont simples parce que les coefficients sont des constantes : $x'(t) = 1$, $y'(t) = -1$. Donc

$$x(t) = t + C_1 \quad (33.12a)$$
$$y(t) = -t + C_2. \quad (33.12b)$$

À priori nous avons une caractéristique pour chaque choix de (C_1, C_2) et nous espérons que le tout recouvre le plan $\mathbb{R}^2$. En fait seule une des deux constantes doit être laissée libre, l'autre consiste seulement en décaler le paramètre t. Nous posons donc $C_1 = 0$ et nous considérons les courbes caractéristiques

$$\gamma_C(t) = \begin{pmatrix} t \\ -t + C \end{pmatrix}. \quad (33.13)$$

Ces courbes recouvrent bien tout le plan. Pour savoir les valeurs de u sur la courbe γ_C, nous devons résoudre l'équation différentielle ordinaire

$$\tilde{u}'_C + \tilde{c}\tilde{u}_C = \tilde{f}, \tag{33.14}$$

en sachant que $\tilde{c}(t) = c\big(x(t), y(t)\big) = -\big(x(t) - y(t)\big) = 2t - C$. Cela se fait en suivant la méthode décrite dans l'exemple 32.9 et résumée dans la proposition 32.11.

En termes de notations, $\tilde{u}_C(t) = u\big(\gamma_C(t)\big)$. Récrivons l'équation :

$$\tilde{u}'(t) - (2t - C)\tilde{u}(t) = 0. \tag{33.15}$$

La méthode pour la résoudre est de mettre les $\tilde{u}$ d'un côté et les t de l'autre :

$$\frac{\tilde{u}'}{u} = 2t - C. \tag{33.16}$$

En intégrant par rapport à t des deux côtés,

$$\ln(\tilde{u}) = t^2 - Ct + K_C, \tag{33.17}$$

c'est-à-dire (avec redéfinition de K_C)

$$\tilde{u}(t) = K_C e^{t^2 - Ct} \tag{33.18}$$

ou encore

$$u\big(\gamma_C(t)\big) = K_C e^{t^2 - Ct} \tag{33.19}$$

EQooSSTJooEQfRnP

où C est le paramètre que nous déterminons en sachant sur quelle caractéristique se trouve le point (x, y) où nous voulons calculer $u(x, y)$ et K est une constante (strictement positive parce que si vous avez suivi le mouvement, c'est une exponentielle) qui doit être déterminée par les conditions initiales. Dès que K est fixé pour un des points de la courbe γ_C, alors il est fixé pour tous les points.

Ce que nous avons obtenu est qu'il existe un K_C tel que pour tout t nous avons

$$u\big(\gamma_C(t)\big) = K_C e^{t^2 - Ct}. \tag{33.20}$$

Soit donc un point $(x_0, y_0) \in \mathbb{R}^2$. Nous devons d'abord déterminer où ce point se trouve par rapport aux caractéristiques, c'est-à-dire quelle est la valeur de C pour laquelle (x_0, y_0) est sur la courbe γ_C, et ensuite déterminer pour quelle valeur de t nous aurons $\gamma_C(t) = (x_0, y_0)$. À résoudre :

$$\gamma_C(t_0) = \begin{pmatrix} t_0 \\ -t_0 + C \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \tag{33.21}$$

Donc $t_0 = x_0$ et $C = x_0 + y_0$. En reprenant (33.19) nous avons

$$u\big(\gamma_C(t_0)\big) = K e^{x_0^2 - (x_0 + y_0)x_0} = K e^{-x_0 y_0}. \tag{33.22}$$

Pour peu que des conditions soient données sur chaque caractéristique, nous pouvons déterminer K. Attention : ce K est une constante d'intégration de l'équation différentielle ordinaire pour $\tilde{u}$. Donc elle n'est valable que sur chaque caractéristique séparément. Cela n'est donc pas du tout une constante sur $\mathbb{R}^2$.

Nous pouvons maintenant écrire la solution générale de l'équation de départ. L'équation cartésienne de la courbe γ_C est

$$x + y = C. \tag{33.23}$$

Donc K est une fonction de $x + y$, pas de x et y séparément. Cela est important à comprendre. À priori nous avons

$$u(x, y) = K(x, y)e^{-xy} \tag{33.24}$$

où $K(x, y)$ est constante sur la courbe γ_C contenant (x, y). Nous avons

— si $x_1 + y_1 = x_2 + y_2$,

— alors il existe C tel que (x_1, y_1) et (x_2, y_2) sont sur γ_C,

— alors $K(x_1, y_1) = K(x_2, y_2)$.

Donc il existe une fonction $\mathbb{R} \to \mathbb{R}$ telle que $K(x, y) = f(x + y)$.

Au final, la solution générale de l'équation est

$$u(x, y) = f(x + y)e^{-xy} \tag{33.25}$$

où f est une fonction à déterminer par les conditions initiales qui peuvent être données. Typiquement nous espérons que les conditions imposent une et une seule valeur de u sur chacune des courbes γ_C.

33.2.2 Un théorème d'existence et d'unicité

La méthode des caractéristiques donne essentiellement une preuve de l'unicité des solutions aux équations de transport, et une méthode pour construire cette solution. En effet, la procédure suivante permet de construire $u(x_0, y_0)$.

— Trouver la caractéristique passant par le point (x_0, y_0)

— Calculer en quel point elle passe par une condition initiale donnée.

— Attribuer à $u(x_0, y_0)$ la valeur trouvée sur la caractéristique là où elle passe par une condition initiale.

Rien ne permet à priori de savoir que cette procédure construit effectivement une solution. En particulier, comment calculer $\partial_x u$? Le quotient différentiel serait

$$\frac{\partial u}{\partial x}(x, y) = \lim_{\epsilon \to 0} \frac{u(x + \epsilon, y) - u(x, y)}{\epsilon}, \tag{33.26}$$

mais la caractéristique donnant la valeur de $u(x + \epsilon, y)$ est différente pour chaque ϵ. Rien à priori ne permet d'affirmer que le calcul soit simple, ni qu'il arrive à une solution du problème donné.

D'où la nécessité d'avoir un résultat un peu rigoureux donnant des conditions sous lesquelles les choses vont bien.

PROPooVQLBooQyFfEH

Proposition 33.4 (Équation de transport à coefficients variables[741, 731]).
Soit une fonction $c\colon \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 en ses deux variables et uniformément Lipschitzienne en sa première variable[1] *et $g \in C^1(\mathbb{R})$. Alors l'équation aux dérivées partielles de premier ordre* SUBEQSooKWBFooKkp

$$\begin{cases} \dfrac{\partial u}{\partial x}(x, t) + c(x, t)\dfrac{\partial u}{\partial t}(x, y) = 0 & \text{(33.27a)} \\ u(x, 0) = h(y) & \text{(33.27b)} \end{cases}$$

admet une unique solution de classe C^1.

Cette solution est construite de la façon suivante[2]*. D'abord nous considérons la solution X au problème* SUBEQSooTOCEooHRhXiv

$$\begin{cases} \dfrac{\partial X}{\partial s}(s; x, t) = c\big(X(s; x, t), s\big) & \text{(33.28a)} \\ X(t; x, t) = x, & \text{(33.28b)} \end{cases}$$

et ensuite le problème (33.27) a pour unique solution

$$u(x, t) = h\big(X(0; x, t)\big). \tag{33.29}$$ EQooLFYNooJawBbE

1. Dans [731], on ne demande que continue puis uniformément Lipschitz. Moi je crois que ce n'est pas assez pour assurer la dérivabilité de X par rapport à x, et encore moins pour permuter les dérivées dans $\partial^2_{tx} Y$.

2. Le fait que la construction ait un sens fait partie des choses à prouver

Démonstration. Nous commençons par étudier l'existence et l'unicité de la fonction X définie par le problème 33.28. La fonction c ici est dans les hypothèses de la fonction f du théorème de Cauchy-Lipschitz global 17.40. D'où l'existence et l'unicité de la fonction $s \mapsto X(s;x,t)$ sur $\mathbb{R}$ pour chaque (x,y) donné[3].

La proposition 32.39 nous dit que X est de classe C^2 en (s,x,t). Donc nous pourrons dériver et permuter les dérivées autant que nous voudrons (sans exagérer : ordre 2 au maximum).

(1) **Unicité** Nous montrons que u doit être constante sur le chemin

$$\gamma_{(x,t)}(s) = \begin{pmatrix} X(s;x,t) \\ s \end{pmatrix}. \tag{33.30}$$

En effet, en posant

$$\varphi(s) = u\big(\gamma(s)\big) = u\big(X(s;x,t),s\big), \tag{33.31}$$

et en dérivant nous obtenons

$$\varphi'(s) = \frac{\partial u}{\partial t}\big(X(s;x,t),s\big)\frac{\partial X}{\partial s}(s;x,t) + \frac{\partial u}{\partial t}\big(X(s;x,t)\big) = 0. \tag{33.32}$$

Par conséquent la valeur commune de tous les $u\big(\gamma_{(x,t)}(s)\big)$ doit être celle en $\gamma_{(x,t)}(0) = h(X(0;x,t))$

Cela prouve l'unicité parce que la valeur de u est fixée en tout point. Nous devons encore vérifier que la fonction u ainsi construite est bien une solution du problème. C'est l'objet de la partie « existence » de la preuve.

(2) **Existence**

Même l'existence est divisée en plusieurs étapes.

(2a) **Mise en place**

Nous prouvons que la fonction u donnée par (33.29) est une solution du problème. Nous avons :

$$\frac{\partial u}{\partial t}(x,t) = h'\big(X(0;x,t)\big)\frac{\partial X}{\partial t}(0;x,t) \tag{33.33}$$

et

$$\frac{\partial u}{\partial x}(x,t) = h'\big(X(0;x,t)\big)\frac{\partial X}{\partial x}(0;x,t), \tag{33.34}$$

de sorte qu'en posant

$$g(s;x,t) = \frac{\partial X}{\partial t}(s;x,t) + c(x,t)\frac{\partial X}{\partial x}(s;x,t) \tag{33.35}$$

nous avons

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = h'\big(X(0;x,t)\big)g(0;x,t). \tag{33.36}$$

(2b) **Une équation différentielle pour g**

Nous allons prouver que

$$\frac{\partial g}{\partial s}(s;x,t) = \alpha_{(x,t)}(s)g(s;x,t) \tag{33.37}$$

avec[4]

$$\alpha_{(x,t)}(s) = \frac{\partial c}{\partial x}\big(X(s;x,t),s\big). \tag{33.38}$$

D'abord nous avons

$$\frac{\partial g}{\partial s}(s;x,t) = \frac{\partial^2 X}{\partial s\partial t}(s;x,t) + c(x,t)\frac{\partial^2 X}{\partial s\partial x}(s;x,t). \tag{33.39}$$

3. Avec l'énoncé tel que donné dans [741], il faut utiliser la technique de 32.22 pour l'existence globale, parce que la fonction b là-dedans n'est pas dans les mêmes hypothèses.

4. La ligne suivante est une de celles qui me font penser qu'il manque des hypothèses dans [731]. Il faut bien pouvoir dériver c.

Nous permutons les dérivées et nous tenons compte de (33.28) :

$$\frac{\partial g}{\partial s}(s;x,t) = \frac{\partial}{\partial t}\Big(c\big(X(s;x,t),s\big)\Big) + c(x,t)\frac{\partial}{\partial x}\Big(c\big(X(s;x,t),s\big)\Big). \tag{33.40}$$

Nous dérivons maintenant plus en profondeur. D'une part

$$\frac{\partial}{\partial t}\Big(c\big(X(s;x,t),s\big)\Big) = \frac{\partial c}{\partial x}\Big(X(s;x,t),s\Big)\frac{\partial X}{\partial t}(s;x,t) \tag{33.41}$$

et d'autre part,

$$\frac{\partial}{\partial x}\Big(c\big(X(s;x,t),s\big)\Big) = \frac{\partial c}{\partial x}\Big(X(s;x,t),s\Big)\frac{\partial X}{\partial x}(s;x,t), \tag{33.42}$$

de telle sorte que

$$\begin{aligned}\frac{\partial g}{\partial s}(s;x,t) &= \frac{\partial c}{\partial x}\big(X(s;x,t),s\big)\Big[\frac{\partial X}{\partial t}(s;x,t) + c(x,t)\frac{\partial X}{\partial x}(s;x,t)\Big] &\text{(33.43a)}\\ &= \alpha_{(x,t)}(s)g(s;x,t). &\text{(33.43b)}\end{aligned}$$

(2c) **Une condition initiale pour g**

Nous montrons maintenant que $g(t;x,t) = 0$. La condition initiale pour X est $X(t;x,t) = x$ pour tout $t, x \in \mathbb{R}$. Nous dérivons cette dernière par rapport à x et à t :

$$\frac{\partial X}{\partial x}(t;x,t) = 1 \tag{33.44a}$$

EQooQKVLooMpfEDC

$$\frac{\partial X}{\partial s}(t;x,t) + \frac{\partial X}{\partial t}(t;x,t) = 0. \tag{33.44b}$$

EQooNAWYooTzYkuQ

Mais $\partial_s X(t;x,t) = c\big(X(t;x,t),t\big)$, donc la relation (33.44b) donne

$$\frac{\partial X}{\partial t}(t;x,t) = -c\big(X(t;x,t),t\big) = -c(x,t) \tag{33.45}$$

EQooVOMVooUhxkJx

où nous avons tenu compte du fait que $X(t;x,t) = x$.
Voyons à présent ce que (33.44a) et (33.45) donnent pour $g(t;x,t)$:

$$g(t;x,t) = \frac{\partial X}{\partial t}(t;x,t) + c(x,t)\frac{\partial X}{\partial x}(t;x,t) = -c(x,t) + c(x,t) = 0. \tag{33.46}$$

(2d) **Conclusion pour g**

La fonction g vérifie l'équation différentielle

$$\begin{cases}\dfrac{\partial g}{\partial s}(s) = \alpha(s)g(s) &\text{(33.47a)}\\ g(t) = 0. &\text{(33.47b)}\end{cases}$$

Bien entendu, $g(s) = 0$ est une solution. Mais la solution à ce système est unique par Cauchy-Lipschitz 17.39. Ici nous utilisons le fait que

$$(s,y) \mapsto \alpha(s)y \tag{33.48}$$

est continue. C'est-à-dire entre autres que

$$s \mapsto \frac{\partial c}{\partial x}\big(X(s;x,t),s\big) \tag{33.49}$$

doit être continue. C'est le cas parce que c est de classe C^1 en sa première variable.

□

Les hypothèses de la proposition 33.4 sont loin d'être optimales. Voici un exemple dans lequel c n'est même pas dérivable par rapport à t et qui se passe très bien quand même.

Exemple 33.5.
Soit l'équation différentielle

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) + |t-1|\frac{\partial u}{\partial x}(x,t) = 0 & (33.50a) \\ u(x,0) = h(x) & (33.50b) \end{cases}$$

où h est une fonction bien régulière ; mettons C^p. En suivant la méthode de la proposition nous devrions poser l'équation différentielle

$$\begin{cases} \frac{\partial X}{\partial s}(s;x,t) = |s-1| & (33.51a) \\ X(t;x,t) = x. & (33.51b) \end{cases}$$

Cela est la caractéristique passant par (x,t). Cependant il sera plus simple de chercher les caractéristiques en demandant qu'elles passent par $(x_0,0)$. Nous allons donc plutôt résoudre pour X_{x_0} l'équation différentielle

$$\begin{cases} \frac{\partial X}{\partial s}(s) = |s-1| & (33.52a) \\ X_{x_0}(0) = x_0 & (33.52b) \end{cases}$$

et les courbes caractéristiques seront les chemins

$$\gamma_{x_0}(s) = \begin{pmatrix} X_{x_0}(s) \\ s \end{pmatrix}. \tag{33.53}$$

La résolution donne d'abord

$$X_{x_0}(s) = \begin{cases} -\frac{s^2}{2} + s + K_1 & \text{si } s < 1 \\ \frac{s^2}{2} - s + K_2 & \text{si } s > 1. \end{cases} \tag{33.54}$$

Vu que la condition initiale est donnée pour $s = 0$, nous fixons K_1 pour la condition initiale et K_2 pour la continuité :

$$X_{x_0}(s) = \begin{cases} -\frac{s^2}{2} + s + x_0 & \text{si } s < 1 \\ \frac{1}{2} + x_0 & \text{si } s = 1 \\ \frac{s^2}{2} - s + 1 + x_0 & \text{si } s > 1. \end{cases} \tag{33.55}$$

Soit $(x,t) \in \mathbb{R}^2$. Quelle caractéristique passe par là ? Nous allons déterminer la fonction $x_0(x,t)$ qui donne le x_0 tel que la caractéristique γ_{x_0} passe par (x,t). Nous devons résoudre

$$\begin{pmatrix} X_{x_0}(s) \\ s \end{pmatrix} = \begin{pmatrix} x \\ t \end{pmatrix}. \tag{33.56}$$

Directement : $s = t$. Et ensuite $X_{x_0}(t) = x$. Nous avons

$$x_0(x,t) = \begin{cases} x + \frac{t^2}{2} - t & \text{si } t < 1 \\ x - \frac{1}{2} & \text{si } t = 1 \\ x - \frac{t^2}{2} + t - 1 & \text{si } t > 1. \end{cases} \tag{33.57}$$

Le truc presque étonnant est que x_0 est de classe C^1. En effet le calcul de

$$\frac{\partial x_0}{\partial t}(1) = \lim_{\epsilon \to 0} \frac{x_0(x,1+\epsilon) - x_0(x,1)}{\epsilon} \tag{33.58}$$

se fait en séparant les limites $\epsilon \to 0^+$ et $\epsilon \to 0^-$. Le résultat est que $\partial_t x_0(1) = 0$. Nous avons donc

$$\frac{\partial x_0}{\partial t}(t) = \begin{cases} t-1 & \text{si } t < 1 \\ 0 & \text{si } t = 1 \\ -t+1 & \text{si } t > 1. \end{cases} \tag{33.59}$$

Cela étant continu, la fonction x_0 est de classe C^1 en t, et la dérivée en x étant toujours 1, elle est de classe C^1.

En ce qui concerne la solution de l'équation de départ,

$$u(x,t) = h\big(X_{x_0(x,t)}(0)\big) = h\big(x_0(x,t)\big). \tag{33.60}$$

Pourvu que h soit assez régulière, la fonction u est facilement de classe C^1. △

33.3 Méthode des caractéristiques pour l'ordre 2

33.3.1 Principe général

Soit l'opérateur différentiel agissant sur $C^2(\mathbb{R}^2)$:

$$D = a(x,y)\frac{\partial^2}{\partial x^2} + b(x,y)\frac{\partial^2}{\partial x \partial y} + c(x,y)\frac{\partial^2}{\partial y^2}. \tag{33.61}$$

Nous voulons résoudre des équations du type $Du = 0$ pour $u\colon \mathbb{R}^2 \to \mathbb{R}$.

Pour commencer[740], et c'est le point crucial, nous voyons D comme un polynôme en ∂_x et ∂_y et nous le factorisons : si

$$aX^2 + bXY + cY^2 = (\alpha X + \beta Y)(\gamma X + \delta Y) \tag{33.62}$$

alors nous avons

$$D = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)\left(\gamma\frac{\partial}{\partial x} + \delta\frac{\partial}{\partial y}\right) + \text{ termes d'ordre inférieurs.} \tag{33.63}$$

Les « termes d'ordre inférieurs » sont ceux de la forme $\alpha(x,y)\frac{\partial\delta}{\partial x}\frac{\partial}{\partial y}$.

L'astuce est de poser

$$v = (\gamma\partial_x + \delta\partial_y)u, \tag{33.64}$$

et de résoudre le système [SUBESQooGCMNooEDWQHd]

$$\begin{cases} (\alpha\partial_x + \beta\partial_y)v = 0 & \text{(33.65a)} \\ (\gamma\partial_x + \delta\partial_y)u = v. & \text{(33.65b)} \end{cases}$$

Cela sont deux équations différentielles du premier ordre pour lesquelles nous avons déjà des techniques décrites en la section 33.2.

Afin que les fonctions α, β, γ et δ soient réelles, il faut que $b^2 - 4ac \geqslant 0$. Sachant que $a = \alpha\gamma$, $b = \alpha\delta + \beta\gamma$ et $c = \beta\delta$ cette condition sur a, b et c donne

$$(\alpha\delta + \beta\gamma)^2 - 4\alpha\gamma\beta\delta \geqslant 0. \tag{33.66}$$

Cela revient à

$$(\alpha\delta - \beta\gamma)^2 \geqslant 0. \tag{33.67}$$

Nous supposons à présent que l'inégalité soit stricte (cas hyperbolique). Nous avons en particulier que

$$\alpha\delta - \beta\gamma \neq 0. \tag{33.68}$$

Cette condition implique que les équations

$$\frac{dx}{dt} = \alpha(x,y) \quad \frac{dy}{dt} = \beta(x,y) \tag{33.69}$$

sont indépendantes des équations

$$\frac{dx}{dt} = \gamma(x,y) \quad \frac{dy}{dt} = \delta(x,y) \tag{33.70}$$

Ce sont les équations caractéristiques des équations (33.65).

33.3.2 Exemple : l'équation d'onde

SUBSECooYBBKooUOIlCS

Nous considérons l'équation aux dérivées partielles

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \tag{33.71}$$

où c est une constante réelle. Nous en cherchons des solutions de classe C^2.

L'opérateur différentiel est donné par le polynôme $P(T, X) = T^2 - c^2 X^2$ qui se factorise en

$$P = (T + cX)(T - cX), \tag{33.72}$$

c'est-à-dire que nous pouvons récrire l'équation des ondes sous la forme

$$(\partial_t + c\partial_x)(\partial_t - c\partial_x)u = 0. \tag{33.73}$$

Nous posons donc $v = (\partial_t - c\partial_x)u$ et nous avons le système[730]

$$\begin{cases} (\partial_t + c\partial_x)v = 0 & \text{SUBEQooLQAPooVoJccp} \quad (33.74a) \\ (\partial_t - c\partial_x)u = v. & \text{SUBEQooPWXMooDjThlJ} \quad (33.74b) \end{cases}$$

La méthode des caractéristiques est efficace pour résoudre la première, et pour trouver la solution générale de l'homogène associée à la seconde.

Nous nous lançons dans la résolution de (33.74a). Le flot est $v = \begin{pmatrix} 1 \\ c \end{pmatrix}$, et nous cherchons ses courbes intégrales sous la forme $\varphi(t) = \big(t, x(t)\big)$. Immédiatement, $x'(t) = c$, ce qui donne

$$\gamma_C(t) = \begin{pmatrix} t \\ ct + C \end{pmatrix}. \tag{33.75}$$

Cela donne une caractéristique pour chaque valeur de C. En posant $\tilde{v}_C(t) = v(t, ct + C)$ nous avons

$$\tilde{v}_C'(t) = \frac{\partial v}{\partial t}\big(\gamma_C(t)\big) + c\frac{\partial v}{\partial t}\big(\gamma_C(t)\big) = 0. \tag{33.76}$$

Donc $\tilde{v}_C$ est une fonction constante. Donc v est constant sur la courbe γ_C dont l'équation cartésienne est $x - ct = C$. Cela implique que

$$v(t, x) = f(x - ct) \tag{33.77}$$

où f est une fonction de classe C^1. En effet si (t_1, x_1) et (t_2, x_2) vérifient $x_1 - ct_1 = x_2 - ct_2$ alors $v(t_1, x_2) = v(t_2, x_2)$. Le fait que f soit C^1 est une demande que v soit au final dans C^2.

Nous devons maintenant résoudre l'équation (33.74b)

$$(\partial_t - c\partial_x)u = v. \tag{33.78}$$

Nous allons agir conformément à la stratégie expliquée par le lemme 32.7. Nous devons résoudre $Du = v$ avec

$$\begin{aligned} D\colon C^2(\mathbb{R}) &\to D(C^2(\mathbb{R})) \\ u &\mapsto (\partial_t - c\partial_x)u. \end{aligned} \tag{33.79}$$

Par la même méthode des caractéristiques que celle déjà menée plus haut nous trouvons $\ker(D)$ comme solution générale de $(\partial_t - c\partial_x)u_G = 0$. C'est-à-dire

$$u_G = g(x + ct) \tag{33.80}$$

où g est une fonction quelconque de classe C^2.

Il nous faut maintenant une solution particulière de

$$(\partial_t - c\partial_x)u_P(t, x) = f(x - ct). \tag{33.81}$$

Si F est une primitive de f alors

$$u_P(t,x) = -\frac{1}{2c}F(x-ct) \tag{33.82}$$

fonctionne. Vu que f est quelconque dans $C^1(\mathbb{R})$, la fonction F est un élément quelconque de $C^2(\mathbb{R})$. Au final, la solution générale de l'équation des ondes est

$$u(t,x) = g_1(x+ct) + g_2(x-ct) \tag{33.83}$$

où g_1 et g_2 sont des éléments de $C^2(\mathbb{R})$.

33.4 Classification des équations du second ordre

Soit une équation générale d'ordre 2 sur $\mathbb{R}^2$:

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + \beta u = f. \tag{33.84}$$

En ce qui concerne son symbole principal nous avons

$$\sigma(x,y,\xi_1,\xi_2) = a(x,y)\xi_1^2 + b(x,y)\xi_1\xi_2 + c(x,y)\xi_2^2, \tag{33.85}$$

ce qui donne l'équation des caractéristiques

$$a\left(\frac{\partial\phi}{\partial x}\right)^2 + b\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y} + c\left(\frac{\partial\phi}{\partial y}\right)^2 = 0. \tag{33.86}$$

EQooSAFNooEimPhQ

Si nous nous posons sur un point (x_0,y_0) tel que $\nabla\phi(x_0,y_0) = 0$ et $\partial_x\phi(x_0,y_0) \neq 0$ alors, via le théorème de la fonction implicite [5], la condition $\phi(x,y) = 0$ définit une fonction $y \mapsto x(y)$ vérifiant

$$\phi\big(x(y),y\big) = 0 \tag{33.87}$$

pour tout y dans un voisinage de y_0.

Nous pouvons obtenir une équation différentielle ordinaire pour x de la façon suivante. D'abord nous posons $\varphi(y) = \phi\big(x(y),y\big)$ et ensuite nous calculons la dérivée de φ (qui est nulle par construction) :

$$0 = \varphi'(y) = \frac{\partial\phi}{\partial x}x' + \frac{\partial\phi}{\partial y}. \tag{33.88}$$

Nous pouvons donc remplacer $\partial_y\phi$ par $x'\partial_x\phi$ dans l'équation des caractéristiques (33.86) :

$$a\left(\frac{\partial\phi}{\partial x}\right)^2 + bx'\left(\frac{\partial\phi}{\partial x}\right)^2 + c\left(\frac{\partial\phi}{\partial x}\right)^2 (x')^2 = 0. \tag{33.89}$$

Vu que nous avons supposé $(\partial_x\phi) \neq 0$ sur un voisinage de (x_0,y_0) nous pouvons simplifier par $(\partial_x\phi)^2$ et avoir l'équation différentielle ordinaire

$$a\big(x(y),y\big) + b\big(x(y),y\big)x'(y) + c\big(x(y),y\big)x'(y)^2 = 0. \tag{33.90}$$

Notons que, conformément à ce que raconte le théorème des fonctions implicites, nous avons pris un voisinage de y_0 suffisamment petit pour que $x(y)$ reste dans un voisinage de x_0. Ce voisinage étant, nous pouvons le restreindre pour nous assurer du signe de a, b et c. Cela est évidemment très théorique parce que le théorème de la fonction implicite parle de l'existence de voisinages, mais pas de façon de les construire.

Nous donnons la classification suivante.

5. Théorème 17.48.

Définition 33.6 (elliptique, hyperbolique, parabolique)**.**
Types d'équations différentielles.

*(1) Si $b^2 - 4ac < 0$ alors l'équation est **elliptique**.*

*(2) Si $b^2 - 4ac > 0$ alors l'équation est **hyperbolique**.*

*(3) Si $b^2 - 4ac = 0$ alors l'équation est **parabolique**.*

Exemple 33.7.
Un exemple d'équation parabolique est l'équation de la chaleur

$$\frac{\partial u}{\partial t} - \alpha \frac{\partial^2 u}{\partial x^2} = 0 \tag{33.91}$$

où $\alpha > 0$ est une constante. Cette équation est avec $b = c = 0$, donc elle est parabolique. △

Une équation aux dérivées partielles peut changer de nature selon le point.

Exemple 33.8.
Soit l'équation

$$\frac{\partial^2 u}{\partial x^2} - (x^2 - y^2)\frac{\partial^2 u}{\partial y^2} = 0. \tag{33.92}$$

Son $b^2 - 4ac$ vaut $4(x^2 - y^2)$. Elle peut donc être hyperbolique, parabolique ou elliptique selon le point où l'on se trouve. △

33.4.1 Problème au limite

Dans les définitions qui suivent nous considérons un ouvert $\Omega \subset \mathbb{R}^d$ assez régulier et possédant en particulier un vecteur normal extérieur $n(x)$ pour tout point $x \in \partial\Omega$.

Définition 33.9.
*Un problème aux limites **de Dirichlet** est d'imposer la condition*

$$u(x) = g(x) \tag{33.93}$$

pour tout $x \in \Gamma \subset \partial\Omega$. C'est-à-dire imposer la valeur de u sur une partie du bord du domaine.

Définition 33.10.
*Un problème aux limites de **Von Neumann** est d'imposer*

$$\frac{\partial u}{\partial n} \cdot x = g(x) \tag{33.94}$$

pour tout $x \in \Gamma \subset \partial\Omega$. C'est-à-dire imposer les valeurs de la dérivée normale de u sur une partie du bord.

Exemple 33.11.
Lorsqu'on veut imposer un flux de chaleur aux bords d'un domaine pour l'équation de la chaleur, il s'agit de poser des conditions de type Von Neumann. △

Définition 33.12.
*Soit un domaine Ω de $\mathbb{R}^d$ et un opérateur différentiel L sur une partie de $\mathrm{Fun}(\Omega)$. Soit une fonction g sur $\partial\Omega$. Un problème **aux limites stationnaires** est un problème du type : trouver u définie sur Ω telle que*

$$\begin{cases} L(u) = f & \text{(33.95a)} \\ u|_{\partial\Omega} = g. & \text{(33.95b)} \end{cases}$$

Définition 33.13.
Un problème ***aux limites d'évolution*** *est du type : trouver $u \in \mathrm{Fun}\big(]0,\infty[\times \Omega\big)$ tel que*

$$\begin{cases} \frac{\partial^m u}{\partial t^m} + L(u) = f & \text{sur }]0,\infty[\times \Omega \\ u(t,.) = g(t,.) & \text{sur }]0,\infty[\times \partial\Omega \\ u(0,.) = u_0 & \text{sur } \Omega \end{cases} \tag{33.96}$$

où u_0 est une fonction sur Ω.

L'opérateur L ne doit pas opérer sur la partie « t » de u.

DEFooSNIRooBFYSFh

Définition 33.14 (Problème bien posé au sens de Hadamard)**.**
Un problème aux limites est ***bien posé au sens de Hadamard*** *si*

(1) Il admet une unique solution.

(2) La solution dépend de façon continue en les données du problème.

La continuité est au sens des normes sur les fonctions sur $\partial\Omega$ et sur Ω en ce qui concerne les fonctions « données » du problème et des normes pour les fonctions sur Ω ou $\bar{\Omega}$ en ce qui concerne la solution.

Exemple 33.15 (Un problème de Dirichlet bien posé)**.**
Trouver la fonction u définie sur $\Omega = [0,1]^2$ telle que

$$\begin{cases} -\Delta u = 0 & \text{sur } \Omega \\ u = g & \text{sur } \partial\Omega. \end{cases} \tag{33.97}$$

△

33.5 Principe du maximum

LEMooSPVUooDQOeom

Lemme 33.16 ([586])**.**
Soit un ouvert borné $\Omega \subset \mathbb{R}^n$. Soit une matrice symétrique strictement définie positive A telle que $A_{ij} \in C^0(\bar{\Omega})$ pour laquelle il existe $\lambda > 0$ minorant toutes les valeurs propres de toutes les matrices $A(x)$ pour $x \in \Omega$ [6]*.*

Nous posons $L' = -\sum_{ij} A_{ij}\partial_{ij}$. Si $u \in C^2(\Omega)$ atteint un minimum local en $x_0 \in \Omega$, alors

$$(L'u)(x_0) \leqslant 0. \tag{33.98}$$

Démonstration. Nous allons bien entendu diagonaliser A. Si T est une matrice nous avons, en posant $u(x) = v(Tx)$:

$$\frac{\partial u}{\partial x_i}(x) = \sum_k \frac{\partial v}{\partial x_k}(Tx)T_{ki} \tag{33.99}$$

et

$$\frac{\partial^2 u}{\partial x_j \partial x_i}(x) = \sum_{kl} T_{ki}T_{lj}\frac{\partial^2 v}{\partial x_l \partial x_k}(Tx). \tag{33.100}$$

Si T est en particulier une matrice orthogonale diagonalisant A (théorème 8.227) nous avons $T^{-1} = T^t$ et

$$\sum_{ij} T_{ki}A_{ij}T_{jl}^{-1} = D_{kl} = \lambda_k\delta_{kl} \tag{33.101}$$

où les λ_k sont les valeurs propres de A. Notons que partout ici, tout est fonction de x sur Ω : tant A que T que les λ_k. Avec tous ces résultats nous calculons vite que

$$(Lu)(x_0) = -\sum_k \lambda_k(\partial_k^2 v)(Tx_0). \tag{33.102}$$

6. Cela est plus que dire que toutes les $A(x)$ son symétriques strictement définie positive.

Si $x_0 \in \Omega$ est un minimum local de u, alors l'application v a un minimum local en $T^{-1}x_0$. Et donc

$$\frac{\partial^2 v}{\partial x_k^2}(Tx_0) \geqslant 0. \tag{33.103}$$

Du coup,

$$(Lu)(x_0) \leqslant 0. \tag{33.104}$$

□

LEMooDVUYooHEkrsL

Lemme 33.17 ([586]).
Soit un ouvert borné $\Omega \subset \mathbb{R}^n$. Soit une matrice symétrique strictement définie positive A telle que $A_{ij} \in C^0(\bar{\Omega})$ pour laquelle il existe $\lambda > 0$ minorant toutes les valeurs propres de toutes les matrices $A(x)$ pour $x \in \Omega$.

Nous posons

$$L = -\sum_{ij} A_{ij}\partial_{ij}^2 + \sum_i b_i\partial_i + c \tag{33.105}$$

où $b_i, c \in C^0(\bar{\Omega})$.

Soit $u \in C^0(\bar{\Omega}) \cap C^2(\Omega)$ telle que $Lu \geqslant 0$ sur Ω. Alors

ITEMooZNUVooNGPXVc

(1) Si $c = 0$ alors

$$\min_{\bar{\Omega}}(u) = \min_{\partial\Omega}(u). \tag{33.106}$$

EQooOJRWooIMemRN

ITEMooQOQIooJqTzdA

(2) Si $c \geqslant 0$ alors

$$\min_{\bar{\Omega}}(u) \geqslant \min_{\partial\Omega}(-u_-) \tag{33.107}$$

EQooGSXVooGUnRnt

où u_- est défini par

$$u_-(x) = \begin{cases} 0 & si\ u(x) \geqslant 0 \\ -u(x) & si\ u(x) \leqslant 0. \end{cases} \tag{33.108}$$

Démonstration. D'abord, vu que Ω est borné, la fermeture $\bar{\Omega}$ est compacte et u y atteint son minimum. De plus $\partial\Omega$ est également compact (borné et le complémentaire est ouvert parce que Ω est ouvert). Donc u y atteint également son minimum. Cela pour dire que les minimums écrits dans (33.106) et (33.107) ont un sens.

Pour (1).

(1) $\underline{Lu \geqslant \eta > 0}$

Nous supposons qu'il existe $\eta > 0$ tel que $(Lu)(x) \geqslant \eta$ pour tout $x \in \Omega$. Soit x_0 le point minimum sur $\bar{\Omega}$. Si $x_0 \in \Omega$ alors il est intérieur et $\partial_i u(x_0) = 0$ par la proposition 17.69. Dans ce cas nous avons

$$(Lu)(x_0) = (L'u)(x_0) > 0, \tag{33.109}$$

ce qui contredit le lemme 33.16. Nous en déduisons que le minimum de u sur $\bar{\Omega}$ n'est pas atteint dans Ω, mais sur $\partial\Omega$. Pour la définition de la frontière, voir 7.125. Ici nous avons $\bar{\Omega}\backslash\Omega = \partial\Omega$.

Cela prouve (33.106) dans ce cas.

(2) $\underline{Lu \geqslant 0 \textbf{ sur } \Omega}$

Nous prenons maintenant le cas général. Nous posons

$$u_{\gamma,\epsilon}(x) = u(x) - \epsilon e^{\gamma x_1}. \tag{33.110}$$

Nous avons [7]

$$L(e^{\gamma x_1}) = e^{\gamma x_1}\big(-A_{11}(x)\gamma^2 + b_1(x)\gamma\big). \tag{33.111}$$

7. Nous abusons un peu de l'écriture parce que ce que nous calculons vraiment est $L\big(x \mapsto e^{\gamma x_1}\big)$.

Soit γ suffisamment grand pour que

$$\lambda\gamma^2 - \|b_1\|_{\bar{\Omega}}\gamma > 0. \tag{33.112}$$

Ici λ minore toutes les valeurs propres des $A(x)$ et nous notons que γ ne dépend pas de ϵ. En utilisant l'inégalité du lemme 8.234(3), pour tout $\xi \in \mathbb{R}^n$ nous avons

$$\sum_{ij} A_{ij}\xi_i\xi_j \geqslant \lambda|\xi|^2, \tag{33.113}$$

ce qui donne avec $\xi = e_1$: $A_{11} \geqslant \lambda$, et même pour être plus précis : $A_{11}(x) \geqslant \lambda$ pour tout x. Ces inégalités donnent

$$-A_{11}(x)\gamma^2 + b_1(x)\gamma \leqslant -\lambda\gamma^2 + \|b_1\|_{\bar{\Omega}}\gamma < 0. \tag{33.114}$$

Nous avons donc

$$L(e^{\gamma x_1}) = e^{\gamma x_1}\big(-A_{11}(x)\gamma^2 + b_1(x)\gamma\big) \leqslant -\lambda\gamma^2 + \|b_1\|_{\bar{\Omega}}\gamma < 0. \tag{33.115}$$

EQooXCGQooLGMnyL

Nous posons

$$\eta = \epsilon\big(\lambda\gamma^2 - \|b_1\|_{\bar{\Omega}}\gamma\big)\min_{\bar{\Omega}}(e^{\gamma x_1}), \tag{33.116}$$

où le minimum a un sens parce que Ω est borné. Nous avons $\eta > 0$.
C'est le moment de calculer ce que u_ϵ peut pour nous :

$$\begin{aligned}
L(u_\epsilon) &= Lu - \epsilon L(e^{\gamma x_1}) && (33.117a)\\
&\geqslant -\epsilon L(e^{\gamma x_1}) && (33.117b)\\
&= -\epsilon e^{\gamma x_1}\big(-A_{11}(x)\gamma^2 + b_1(x)\gamma\big) && (33.117c)\\
&\geqslant -\epsilon e^{\gamma x_1}\big(\lambda\gamma^2 - \|b_1\|_{\bar{\Omega}}\gamma\big) && (33.117d)\\
&\geqslant \eta > 0. && (33.117e)
\end{aligned}$$

SUBEQooIZTWooVjlTZ

SUBEQooHBALooSXxUeL

Justification :

— (33.117b) parce que $Lu \geqslant 0$.

— (33.117d) par (33.115).

La fonction u_ϵ est donc dans le cas précédent et nous avons

$$\min_{x\in\bar{\Omega}}\big(u(x) - \epsilon e^{\gamma x_1}\big) = \min_{x\in\partial\Omega}\big(u(x) - \epsilon e^{\gamma x_1}\big). \tag{33.118}$$

Mentionnons le fait que le choix fait de γ ne dépend pas de ϵ. Nous pouvons donc encore faire varier ϵ sans toucher à γ et en maintenant toutes les inégalités prouvées jusqu'ici.
Vu que l'expression $e^{\gamma x_1}$ est majorable sur $\bar{\Omega}$ nous avons convergence uniforme

$$u_\epsilon \xrightarrow{\|.\|_{\bar{\Omega}}} u \tag{33.119}$$

pour $\epsilon \to 0$.
Supposons qu'aucun point de $\partial\Omega$ ne réalise le minimum de u. Alors il existe $x_0 \in \Omega$ tel que $u(x_0) > u(x)$ pour tout $x \in \partial\Omega$. Mais comme $\partial\Omega$ est compact, il existe $\eta > 0$ tel que

$$u(x_0) < u(x) + \eta. \tag{33.120}$$

Soit ϵ tel que $\|u - u_\epsilon\|_{\bar{\Omega}} < \eta/2$. Nous avons

$$u(x_0) < u(x) - \eta \tag{33.121}$$

et donc aussi

$$u_\epsilon(x_0) < u(x_0) + \frac{\eta}{2} < u(x) - \frac{\eta}{2} < u_\epsilon(x), \tag{33.122}$$

ce qui signifierait que u_ϵ prend son minimum dans Ω. Or nous savons que ce n'est pas le cas. Donc il existe un point de $\partial\Omega$ qui réalise le minimum de u.

Pour (2).

Supposons pour commencer que $u \geqslant 0$ sur Ω. Alors par continuité $u \geqslant 0$ sur $\bar{\Omega}$. Alors $u_- = 0$ et

$$\min_{\bar{\Omega}} u \geqslant 0, \tag{33.123}$$

ce qui fait que l'inégalité (33.107) est évidente.

Nous supposons donc que l'ensemble

$$\Omega_- = \{x \in \Omega \text{ tel que } u(x) < 0\} \tag{33.124}$$

est non vide. Notons que c'est également un ouvert.

Soit $\bar{L}u = Lu - cu$. Vu que $c \geqslant 0$ et que $Lu \geqslant 0$ nous avons $\bar{L}u \geqslant 0$ sur Ω_-. Par le point (1) nous avons

$$\min_{x\in\bar{\Omega}} u(x) = \min_{x\in\partial\Omega} u(x). \tag{33.125}$$

Mais le minimum de u est certainement atteint dans Ω_-, donc

$$\min_{x\in\bar{\Omega}_-} u(x) = \min_{x\in\bar{\Omega}} u(x). \tag{33.126}$$

Cela nous donne une première bonne égalité :

$$\min_{x\in\partial\Omega_-} u(x) = \min_{x\in\bar{\Omega}} u(x). \tag{33.127}$$

Nous pouvons la prolonger :

$$\begin{aligned}
\min_{x\in\bar{\Omega}} u(x) &= \min_{x\in\partial\Omega_-} u(x) && \text{(33.128a)}\\
&= \min_{x\in\partial\Omega_-} \big(-u_-(x)\big) && \text{(33.128b)}\\
&= \min_{x\in\partial\Omega_-\cap\partial\Omega} \big(-u_-(x)\big) && \text{(33.128c)}\\
&= \min_{x\in\partial\Omega} \big(-u_-(x)\big) && \text{(33.128d)}
\end{aligned}$$

SUBEQooUYFXooKKrHiF

SUBEQooKIHRooPjADmK

Justifications :

— Pour (33.128c) nous avons la décomposition

$$\partial\Omega_- = \big(\partial\Omega_- \cap \Omega\big) \cup \big(\partial\Omega_- \cap \partial\Omega\big). \tag{33.129}$$

Or sur $\partial\Omega_- \cap \Omega$ nous avons $u(x) = 0$ et donc pas le minimum.

— Pour (33.128d). Sur $\partial\Omega$, le minimum est atteint dans la partie $\partial\Omega_-$ parce que le reste ne contient que des valeurs positives de u.

Cela prouve ce que nous voulions. □

Théorème 33.18 (Principe du maximum fort)**.**
Soit un ouvert borné Ω de $\mathbb{R}^n$. Soit une matrice A dont

— *A_{ij} est dans $C^0(\bar{\Omega})$*

— *$A(x)$ est symétrique strictement définie positive pour tout x.*

— *Il existe $\lambda > 0$ minorant toutes les valeurs propres des toutes les matrices $A(x)$ sur $\bar{\Omega}$.*

Soit $b_i, c \in C^0(\bar{\Omega})$ avec $c(x) \geqslant 0$ sur $\bar{\Omega}$.

Soit $u \in C^0(\bar{\Omega}) \cap C^2(\Omega)$ telle que

$$\begin{cases} -\sum_{ij} A_{ij}\partial_{ij}u(x) + \sum_i b_i\partial_i u(x) \geqslant 0 & \text{(33.130a)}\\ u(x) \geqslant 0 \ \forall x \in \partial\Omega. & \text{(33.130b)} \end{cases}$$

Alors $u \geqslant 0$ sur $\bar{\Omega}$.

Démonstration. Nous appliquons le lemme 33.17(2) :

$$\min_{\bar{\Omega}} u \geqslant \min_{\partial\Omega}(-u_-). \tag{33.131}$$

Vu que $u(x) \geqslant 0$ sur $\partial\Omega$, nous avons $u_- = 0$ sur $\partial\Omega$ et donc $\min_{\bar{\Omega}} u \geqslant 0$. □

33.6 Quelques exemples

33.6.1 Un changement de variables

Soit l'équation différentielle

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \tag{33.132}$$

sur $\Omega =]0, \infty[\times \mathbb{R}$. Nous imposons la condition aux bords

$$u(0,x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x > 0. \end{cases} \tag{33.133}$$

Nous cherchons les solutions sous la forme

$$u(t,x) = f\left(\frac{x}{\sqrt{t}}\right). \tag{33.134}$$

Nous avons

$$\frac{\partial u}{\partial t} = -\frac{1}{2} x t^{-3/2} f'\left(\frac{x}{\sqrt{t}}\right) \tag{33.135}$$

ainsi que

$$\frac{\partial u}{\partial x} = \frac{1}{\sqrt{t}} f'\left(\frac{x}{\sqrt{t}}\right) \tag{33.136}$$

et

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{t} f''\left(\frac{x}{\sqrt{t}}\right). \tag{33.137}$$

En remettant le tout dans l'équation de départ et en simplifiant par $1/t$ (qui est permis parce que $t > 0$) :

$$-\frac{1}{2} f'\left(\frac{x}{\sqrt{t}}\right) \frac{x}{\sqrt{t}} - f''\left(\frac{x}{\sqrt{t}}\right) = 0. \tag{33.138}$$

Nous résolvons l'équation différentielle

$$\frac{z}{2} g'(z) + g''(z) = 0 \tag{33.139}$$

pour la fonction g de la variable réelle z. Cela fait, la réponse sera $f = g \circ z$ où z serait la fonction

$$z(x,t) = x/\sqrt{t}, \tag{33.140}$$

Pour résoudre (33.139) nous commençons par résoudre pour la dérivée $h = g'$, c'est-à-dire l'équation différentielle

$$\frac{z}{2} h(z) + h'(z) = 0 \tag{33.141}$$

qui donne

$$\frac{h'(z)}{h(z)} = -z/2. \tag{33.142}$$

Une intégration fournit $\ln\big(h(z)\big) = -z^2/4 + K$ et donc

$$h(z) = K e^{-z^2/4} \tag{33.143}$$

et

$$g(z) = C + K \int_0^z e^{-s^2/4} ds. \tag{33.144}$$

Nous ne pouvons pas aller plus loin parce que nous ne sommes pas capables de calculer la primitive demandée. Nous laissons donc g sous cette forme et nous posons

$$u(x,t) = g(x/\sqrt{t}), \tag{33.145}$$

en pleine confiance du fait que cela soit une solution de l'équation aux dérivées partielles (33.132).

Nous devons fixer K et C de telle façon à respecter les conditions aux bords (33.133). D'abord écrivons aussi explicitement que possible la fonction u :

$$u(t,x) = K + C\int_0^{x\sqrt{t}} e^{-s^2/4}ds. \tag{33.146}$$

Il faut calculer $u(0,x)$ en termes de C et K. Pour cela nous calculons, pour un x fixé :

$$\lim_{t\to 0^+} u(t,x) = K + C\lim_{t\to 0^+}\int_0^{x/\sqrt{t}} e^{-s^2/4}ds. \tag{33.147}$$

En utilisant un petit changement de variables sur l'intégrale gaussienne de l'exemple 14.307, et en remarquant que la fonction est symétrique,

$$\lim_{t\to 0^+} u(t,x) = \begin{cases} K + C\sqrt{\pi} & \text{si } x > 0 \\ K - C\sqrt{\pi} & \text{si } x < 0. \end{cases} \tag{33.148}$$

À résoudre :

$$\begin{cases} K + C\sqrt{\pi} = 1 & \text{(33.149a)} \\ K - C\sqrt{\pi} = 0. & \text{(33.149b)} \end{cases}$$

Solution : $C = 1/2\sqrt{\pi}$ et $K = 1/2$. Au final,

$$u(t,x) = \frac{1}{2} + \frac{1}{2\sqrt{\pi}}\int_0^{x/\sqrt{t}} e^{-s^2/4}. \tag{33.150}$$

Notons que cela donne une valeur pour $u(t,0)$:

$$u(t,0) = \frac{1}{2}. \tag{33.151}$$

Chapitre 34

Numérique

D'autres lectures agréables dans [742].

34.1 Introduction

À quels types de problèmes peut-on s'attendre lorsqu'on se lance dans du calcul numérique, et en particulier dans la résolution numérique d'équations (algébrique, différentielles ou aux dérivées partielles, etc) ?

Quelques réflexions en vrac sur ce sujet.

(1) Les erreurs de représentation de nombres : troncature et propagation de décalales (*drift*),

(2) Erreur de compensation (*cancellation*),

(3) Conditionnement, stabilité : les réponses peuvent fortement dépendre des paramètres,

(4) Si on utilise une méthode itérative, comment savoir à quel moment on s'arrête ? Calculer la différence $|x_k - x_{k-1}|$ mène t-il à une erreur de cancellation ?

(5) Lors d'une implémentation, les matrices des systèmes à résoudre sont souvent très grandes et/ou très creuses. Cela pose la question de la manière de les enregistrer.

(6) Pour la parallélisation, il faut faire attention au fait que parfois créer un nouveau processus demande plus de ressources que le mini-calcul qu'on voulait faire. Donc il ne faut pas toujours paralléliser tout ce qui est théoriquement parallélisable.

(7) Le fait que certaines méthodes sont non-déterministes (Monté-Carlo) mène à des problèmes pour les tests unitaires des implémentations.

34.2 Représentations numériques

Dans cette section, les séquences de chiffres écrites entre crochet sont à comprendre comme des séquences de chiffres qui représentent une quantité suivant un codage donné.

34.2.1 Entier relatif en complément à deux (binaire)

Si nous avons m bits pour coder un entier relatif, une idée serait de prendre le premier bit pour le signe (0 pour positif et 1 pour négatif) et les autres pour la valeur absolue. Deux inconvénients :

(1) Il y a deux codages pour le zéro, donc gaspillage.

(2) L'algorithme pour faire la somme passe mal. Par exemple pour faire $1 + (-1)$, le 1 est codé comme [001] et le -1 par [101] et la somme se ferait naïvement comme

$$\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array}$$

Donc le résultat est [110] qui s'interprète comme -2. Complètement faux.

Une solution est d'utiliser le **complément à deux**, qui est la façon usuelle de représenter des entiers signés.

Les entiers positifs se codent normalement, en laissant à zéro le premier bit (donc si nous disposons de m bits, nous codons sur $m-1$ bits).

Les entiers négatifs se codent en trois étapes.

- coder la valeur absolue
- inverser tous les bits (d'où le nom de « complément à deux »)
- soustraire 1.

Exemple 34.1.
Pour coder -1 nous faisons

- Nous codons 1 : [001]
- Nous inversons tous les bits : [110]
- Nous faisons -1 : [101].

△

Avec ce système, la somme passe bien : calculer $1+(-1)$ donne

$$\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array}$$

La réponse est donc [110] qu'il faut interpréter via le complément à deux.

$$110 \xrightarrow{+1} 111 \xrightarrow{\text{complément}} 000. \tag{34.1}$$

Et ce dernier [000] s'interprète comme zéro.

Définition 34.2 (Entier signé en complément à deux[743])**.**
La suite de bits $[a_{m-1}\dots a_0]$ *s'interprète via la formule*

$$-a_{m-1}2^{m-1}+\sum_{i=0}^{m-2}a_i2^i. \tag{34.2}$$

EQooXFHKooHRXDmZ

Le premier bit donne effectivement le signe du nombre, mais l'interprétation d'un nombre n'est pas aussi simple que ce que l'on pourrait croire de prime abord.

Exemple 34.3 (Entier signé en 8 bits)**.**
Que pouvons nous faire avec 8 bits ? Le plus grand nombre est codé par [01111111] qui vaut $\sum_{k=0}^{6}2^k=2^7-1=127$. (avez-vous utilisé la somme (1.687) ?)

Le plus petit nombre codable en 8 bits n'est pas [11111111] mais bien [10000000] (cela est plus clair en regardant la formule (34.2) qu'en tentant de suivre la construction du complément à deux) qui signifie $-2^7=-128$.

Nous pouvons donc coder tous les nombres de -128 à 127. △

Plus généralement un système qui codes des entiers signés en N bits utilisant le complément à deux peut coder de $-(2^{N-1})$ à $2^{N-1}-1$.

34.4 (Le dépassement)**.**
Que se passe-t-il lorsque nous commettons un dépassement ? Calculons sur 3 bits la somme [011] + [001] qui revient à ajouter 1 au nombre le plus grand :

$$\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 0 & 1 \\ \hline 1 & 0 & 0 \end{array}$$

qui signifie $-2^2 = -4$. Lors d'un dépassement, nous retombons automatiquement sur le plus petit.

Ce phénomène est bien connu des personnes qui programment sans faire attention dans certains languages de programmation qui ne font pas attention à votre place.

Définition 34.5 (Représentation en virgule fixe)**.**
Soit x un réel. On définit sa ***représentation en virgule fixe*** *par*

$$x = \{[x_n x_{n-1} ... x_0, x_{-1} ... x_{-m}], b, s\} \tag{34.3}$$

avec $b \in \mathbb{N}, b \geqslant 2$, $s \in \{0, 1\}$ et $x_j \in \mathbb{N}, x_j < b$ suivant la formule

$$x = (-1)^s \sum_{j=-m}^{n} x_j . b^j. \tag{34.4}$$

34.2.2 Représentation en virgule flottante

DEFooLYONooBNskty

Définition 34.6 (Représentation en virgule flottante[744])**.**
La ***représentation en virgule flottante normalisée*** *en base b d'un nombre est la donnée de*

(1) Un bit s pour le signe

(2) Un entier non signé *q de e chiffres pour l'exposant*

(3) Une suite de chiffres $[a_1 \dots a_m]$ pour la mantisse.

Ces données s'interprètent via la formule

$$\mathrm{fl}(s, q, [a_1, \dots, a_m]) = (-1)^s \sum_{j=1}^{m} b^j a_j \times b^{q-d} \tag{34.5}$$

EQooAGWJooBuBbBu

où $d = b^{e-1}$ est le ***décalage****.*

Une idée à retenir est que l'exposant est un entier non signé parce qu'il est plus simple d'introduire un décalage dans la formule (34.5) que de compliquer l'écriture de l'exposant.

34.2.3 Simple précision, IEEE-754

En écriture binaire, la représentation en virgule flottante est un peu différente parce qu'il y a une idée supplémentaire ; la simple précision que nous allons voir maintenant n'est donc pas un cas particulier de 34.6 avec $b = 2$.

Nous commençons par une description informelle de la précision simple avant de donner la définition. La représentation en **précision simple** d'un nombre se fait sur 32 bits répartis comme suit :

(1) 1 bit pour le signe,

(2) 8 bits pour l'exposant interprété comme nombre entier non signé

(3) 23 bits pour la mantisse

Soit le triple

$$\big(s, q, [a_1, \dots, a_{23}]\big) \tag{34.6}$$

Dans le cas générique, l'idée est de donner 24 bits pour la mantisse, mais en gardant en tête le fait que de toutes façons, le premier bit doit être 1, sinon il suffirait de décaler, c'est-à-dire changer l'exposant. Par conséquent la mantisse ne reçoit que 23 bits ; il y a un « 1 » sous-entendu en première position. Donc la mantisse $[a_1, \dots, a_{23}]$ est à lire comme le nombre

$$1, a_1 \dots a_{23} = 1 + \sum_{j=1}^{23} a_j 2^{-j}. \tag{34.7}$$

Exemple 34.7.
La mantisse $[011100 \dots 0]$ signifie $1,0111 = 1 + 2^{-2} + 2^{-3} + 2^{-4} = 1 + \frac{1}{4} + \frac{1}{8} + \frac{1}{16}$. △

Cela pour justifier la formule

$$\mathrm{sp}\big(s,q,[a_1,\dots,a_{23}]\big)=(-1)^s\big(1+\sum_{j=1}^{23}a_j2^{-j}\big)2^{q-127}. \tag{34.8}$$

Notons :

(1) Le « 1+ » dans la parenthèse correspond au 1 implicite en première position de la mantisse.

(2) Il y a un décalage de 127 dans l'exposant, parce que q est un entier non signé.

Notons que cette règle du 1 implicite dans la mantisse empêche d'écrire le nombre 0, et ne permet pas d'écrire des nombres franchement petits parce que le 1 implicite est en *première* position dans la mantisse.

D'où l'idée de donner une règle particulière lorsque l'exposant vaut 0. Lorsque l'exposant est $q=0$, alors nous ne considérons pas de 1 implicite dans la mantisse, et le décalage de l'exposant est -126 au lieu de -127. D'où la formule

$$\mathrm{sp}(s,q=0,[a_1\dots a_{23}])=(-1)^s2^{-126}\sum_{j=1}^{23}a_j2^{-j}. \tag{34.9}$$

En particulier, si $q=0$ et $a=[0\dots0]$, nous avons le nombre zéro exact (il y a deux possibilités pour le code).

Enfin, nous avons des cas particuliers lorsque l'exposant est maximum, c'est-à-dire lorsque $q=[1111\,1111]=2^8-1=255$. Dans ce cas, le nombre codé est soit $+\infty$ soit NaN. Nous posons

$$\mathrm{sp}(s,q=255,a=0)=+\infty \tag{34.10}$$

et

$$\mathrm{sp}(s,q=255,a\neq0)=NaN. \tag{34.11}$$

Il y a en réalité plusieurs valeurs différentes de NaN, mais nous n'entrons pas dans ces détails[745].

DEFooEIOZooYLDVjs

Définition 34.8 (Représentation en simple précision (binaire))**.**
La représentation en ***précision simple*** *d'un nombre se fait sur* 32 *bits répartis comme suit :*

(1) 1 *bit pour le signe,*

(2) 8 *bits pour l'exposant interprété comme nombre entier non signé*

(3) 23 *bits pour la mantisse*

Un nombre est représenté par un triple

$$\big(s,q,[a_1,\dots,a_{23}]\big) \tag{34.12}$$

Selon que l'exposant $q-d$ soit égal à 0, $2^8-1=255$ ou autre chose, les règles d'interprétation sont différentes. Il y a donc trois cas.

Exposant q générique[746] *Si $q\neq0$ et $q\neq255$ alors le nombre est* ***normalisé****. La règle de lecture est alors*

$$\mathrm{sp}\big(s,q,[a_1,\dots,a_{23}]\big)=(-1)^s\big(1+\sum_{j=1}^{23}a_j2^{-j}\big)2^{q-127}. \tag{34.13}$$

EQooEFEKooIrUaKj

Exposant q égal à 0 *Le nombre est dit* ***dénormalisé*** *et la règle de lecture est*

$$\mathrm{sp}\big(s,q,[a_1,\dots,a_{23}]\big)=(-1)^s2^{-126}\sum_{j=1}^{23}a_j2^{-j}. \tag{34.14}$$

EQooRTBFooIplydi

Exposant q égal à 255 *La règle de lecture est alors au cas pas cas ou à peu près.*

(1) $\mathrm{sp}(s,q=255,a=0)=+\infty$.

(2) $\mathrm{sp}(s, q = 255, a \neq 0) = \mathrm{NaN}$.

Vous pouvez jouer avec la simple précision dans [747].

Exemple 34.9 (Plus petit normalisé)**.**
Pour faire un nombre normalisé, il faut au minimum $q = 1$. En prenant $a_j = 0$ nous obtenons le plus petit nombre normalisé possible en simple précision. La formule (34.13) donne

$$\mathrm{sp}(1, q = 1, a = 0) = 2^{1-127} = 2^{-126} \simeq 1.17549435082229 \times 10^{-38}. \tag{34.15}$$

△

Exemple 34.10 (Plus grand normalisé)**.**
L'exposant q ne peut pas être maximum, sous peine de tomber dans les règles spéciales de $+\infty$ ou NaN. Donc $q = [1111\,1110] = 2^8 - 2 = 254$. En ce qui concerne la mantisse, il faut la prendre maximale, c'est-à-dire $a_j = 1$ pour tout j. Nous avons alors le nombre

$$\mathrm{sp}(1, q = 254, a = [1 \dots 1]) = \big(1 + \sum_{j=1}^{23} 2^{-j}\big)2^{254-127} = (1 - \frac{1}{2^{24}})2^{128} \tag{34.16a}$$

$$= 3.4028234663852885981170418348451692544 \times 10^{38} \tag{34.16b}$$

EQmoFRPYooRnxjFP

où nous avons utilisé la somme (1.687) (et Sage pour le dernier calcul). △

Notons ceci avec Sage :

```
1
2  SageMath Version 7.0, Release Date: 2016-01-19
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: A=(1- (1/2**24)  )*2**(128)
7 sage: type(A)
8 <type 'sage.rings.rational.Rational'>
```

tex/sage/sageSnip003.sage

La précision du nombre donné en (34.16b) aurait été embarrassante si le type avait été un nombre en simple précision. Précision technique : en Python, le type `int` n'a pas de limite supérieure à part la mémoire.

Exemple 34.11 (Plus petit non nul dénormalisé)**.**
Pour être dénormalisé il faut $q = 0$ (ce qui est toutefois assez logique si nous voulons un petit nombre), et pour ne pas être nul, il faut une mantisse non nulle. Donc $a = [0 \dots 01]$. La formule (34.14) donne alors

$$\mathrm{sp}(s = 0, q = 0, a = [0 \dots 01]) = 2^{-126}2^{-23} = 2^{-149} \simeq 1.40129846432482 \times 10^{-45}. \tag{34.17}$$

△

EXEMooRHENooGwumoA

Exemple 34.12 (Plus grand dénormalisé)**.**
Pour être dénormalisé il faut toujours $q = 0$, mais cette fois nous prenons la plus grande mantisse possible :

$$\mathrm{sp}(s = 0, q = 0, a = [1 \dots 1]) = 2^{-126} \sum_{j=1}^{23} 2^{-j} = 2^{-126}(1 - 2^{-23}) = 1.17549421069244 \times 10^{-38} \tag{34.18}$$

△

Notons ceci avec Sage :

```
sage: B=2**(-126)*(1-2**(-23))
sage: A=2**(-126)
sage: n(A-B)
1.40129846432482e-45
```

tex/sage/sageSnip004.sage

Vu que $2^{-23} \simeq 1.2 \times 10^{-7}$, approximer la parenthèse par 1 donne une faute sur la septième décimale, ce qui est visible en simple précision.

34.3 Problèmes pour écrire des nombres

Définition 34.13.
*L'**erreur relative** commise en remplaçant un nombre réel x par une valeur approchée $\hat{x}$ est définie par*

$$\epsilon_x := \left| \frac{x - \hat{x}}{x} \right|. \tag{34.19}$$

L'erreur relative n'est pas influencée par l'ordre de grandeur de x. En effet, l'ordre de grandeur de $\hat{x}$ est certainement la même que celle de x, dans la majorité des cas sans problèmes. Du coup si $x' = 200x$ alors $\hat{x'} \simeq 200\hat{x}$ et le 200 se simplifie.

Le nombre de chiffres significatifs correct dans l'approximation est donné par $-\log_{10}(\epsilon_x)$. La partie entière de ce nombre est le nombre de chiffres tout à fait exacts et la partie décimale donne une idée sur le fait que le chiffre suivant est plus ou moins bien.

Remarque 34.14.
Si nous voulons donner $x \in \mathbb{R}$ à un ordinateur, nous sommes soumis à deux erreurs :

(1) D'abord, vu que nous ne pouvons pas taper sur le clavier toutes les décimales de x, nous faisons une **erreur de troncature**.

(2) L'ordinateur devant convertir cela en base deux, il commet une seconde erreur, dite **erreur d'assignation**.

34.3.1 Troncature : la base

Supposons que nous voulions écrire le nombre (écrit ici en base 10)

$$0.4567894251 \tag{34.20}$$

de façon plus facile à lire, on peut demander de ne laisser que t chiffres significatifs. Disons $t = 3$.

Technique de troncature On garde 3 chiffres significatifs : 0.456. Facile.

Technique d'arrondi Vu que le premier qu'on supprime est un 7, le dernier qu'on garde est majoré de 1 : on écrit 0.457.

Que faire si le premier chiffre rejeté est un 5 ? En première approximation, nous pouvons prendre la règle suivante : si le premier chiffre rejeté est un 5, il faut augmenter de 1 de dernier chiffre gardé parce qu'il y a presque certainement encore un chiffre non nul derrière.

Remarque 34.15.
Les ordinateurs travaillent tous en mode d'arrondi.

Exemple 34.16.
Si on doit entrer le nombre 0.38358546 dans un ordinateur qui ne garde que 3 chiffres significatifs, il faut taper 0.384 au clavier (erreur classique dans les exercices). △

34.3.2 Troncature : le drift

Soit une machine ne pouvant retenir que 3 chiffres significatifs et effectuant les arrondis vers le haut lorsque le chiffre à éliminer est un 5. Nous notons $\oplus$ et $\ominus$ les opérations d'addition et soustraction avec arrondis[748]. Les égalités comprenant plus de trois chiffres significatifs sont des égalités au sens de la machine. Nous écrirons donc sans états d'âme :

$$1 \oplus 0.555 = 1.555 = 1.56. \tag{34.21}$$

Considérons la suite numérique

$$\begin{cases} x_0 = 1.00 & \text{(34.22a)} \\ x_n = (x_{n-1} \ominus y) \oplus y & \text{(34.22b)} \end{cases}$$

avec $y = -0.555$.

Nous avons

$$x_1 = (1 \oplus 0.555) \ominus 0.555 = 1.56 \ominus 0.555 = 1.005 = 1.01 \tag{34.23}$$

et ensuite

$$x_2 = (1.01 \oplus 0.555) \ominus 0.555 = 1.565 \ominus 0.555 = 1.57 \ominus 0.555 = 1.015 = 1.02. \tag{34.24}$$

Et ainsi de suite. La suite est donc croissante alors que la définition nous donnerait envie d'avoir $x_n = x_0$ pour tout n.

Remarque 34.17.
En réalité, cette suite se stabilise à $x_n = 10$ pour tout n à partir de $n = 845$. En effet,

$$(10 \oplus 0.555) \ominus 0.555 = 10.555 \ominus 0.555 = 10.6 \ominus 0.555 = 10.045 = 10. \tag{34.25}$$

Le fait est qu'à ce moment, l'erreur de troncature est assez loin dans les décimales pour que le premier chiffre négligé soit un "0" au lieu d'un "5".

Notons toutefois que cette stabilité n'est pas là pour nous rassurer parce qu'elle n'en est pas moins complètement fausse.

La règle de troncature adoptée dans Sage est d'arrondir au nombre pair le plus proche lorsque le premier nombre à négliger est un 5. Donc 12.5 s'arrondit à 12 plutôt que 13.

Exemple 34.18.
Soient les expressions (algébriquement égales) :

(1) $A = x(x + 1)$

(2) $B = x^2 + x$

Nous savons que

$$x = \text{fl}(x) = 10^{-30} \tag{34.26}$$

et

$$1 = \text{fl}(1) \tag{34.27}$$

parce que pour 1 et 10^{-30}, il n'y a pas d'erreurs d'assignation.

En précision simple, $10^{-30} + 1 = 1$ parce qu'en précision simple, il n'y a que 7 ou 8 chiffres significatifs [1].

Nous avons $A = 10^{-30}$, mais x^2 donne un `underflow` parce que 10^{-60} ne peut pas être représenté en précision simple. En pratique, beaucoup de logiciels en font 0. Dans ce cas, en réalité B donne effectivement 10^{-30} après avoir fait $x^2 + x = 0 + x = 10^{-30}$. △

1. Erreur de « relation normale ».

34.3.3 Quelques bonnes règles

(1) Si on a plusieurs nombres à additionner ou soustraire, il vaut mieux commencer par sommer ou soustraire ceux dont on sait qu'ils ont le même ordre de grandeur. Il n'y a donc pas tout à fait « associativité » des erreurs.

(2) Les opérations délicates sont l'addition et la soustraction. La multiplication et la division sont sans dangers, à part l'erreur de dépassement du maximum. Dans une multiplication, on perd au pire quelques chiffres significatifs, mais certainement les derniers, pas les premiers.

34.3.4 Erreur de "cancellation"

Lorsque deux nombres sont de même ordre de grandeur, avec plusieurs nombres significatifs identiques, la cancellation est le fait que, suite à la soustraction, tous les chiffres significatifs ou presque se sont simplifiés et qu'il ne reste plus que des chiffres non significatifs.

Exemple 34.19 ([749])**.**
Sur une machine ne gardant que 4 chiffres significatifs, faire

$$0.5678 \times 10^6 - 0.5677 \times 10^6 = 0.0001 \times 10^6 = 0.1000 \times 10^3. \tag{34.28}$$

Le fait est que les trois derniers zéros ne sont pas significatifs, mais maintenant la machine nous fait croire qu'ils le sont.

Une autre façon de voir ce problème est d'imaginer qu'il faille calculer la différence

$$0.5678\,289798 \times 10^6 - 0.5677\,3136907 \tag{34.29}$$

sur cette machine. Certes la machine nous autorise à avoir 4 chiffres significatifs, donc au moment d'entrer les nombres nous perdons un beau paquet de chiffres. Mais au moment de faire la différence, nous perdons (presque) tout le reste. Donc là où nous pouvions espérer avoir 4 chiffres significatifs de la différence, nous n'en avons que 1. Les trois derniers zéros de la réponse (0.1000×10^3) sont faux. △

REMooRQIJooNLdAZE

Remarque 34.20.
L'erreur de cancellation provoque des chiffres significatifs faux, mais ne provoque pas de faute dans l'*ordre de grandeur* des réponses[2]. Donc si nous voulons nous assurer que a et b sont égaux « à erreur numérique près », le test

$$|a - b| < \epsilon \tag{34.30}$$

est valide, malgré l'erreur de cancellation qui ne manquera pas de se produire dans le calcul de la différence.

Exemple 34.21.
Soit à résoudre l'équation $ax^2 + bx + c = 0$ avec $a, b, c \neq 0$ et $b^2 - 4ac > 0$. Solution :

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{34.31}$$

Supposons que $|4ac| \ll b^2$ avec tout de même pas tellement petit qu'on se perd dans la précision. Bref, on suppose que seules quelques dernières décimales de $b^2 - 4ac$ sont différentes de zéro.

On a :

$$\sqrt{b^2 - 4ac} = \sqrt{\tilde{b}^2} = |\tilde{b}| \tag{34.32a}$$

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{34.32b}$$

$$x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{34.32c}$$

2. Est-ce bien vrai, cela ?

Si $b > 0$, nous avons une erreur de cancellation dans x_2 parce qu'on fait la différence entre deux nombres presque égaux. Donc x_2 mal calculé. Par contre x_1 est bien calculé.

Si par contre $b < 0$, c'est le contraire.

Avec $a = 10^{-3}$, $b = 0.8$, $c = -1.2 \times 10^{-5}$. À la main nous obtenons : $x_1 = -800$, $x_2 = 1.5 \times 10^{-5}$, et un ordinateur se tromperait ...

```
1
2  SageMath Version 7.0, Release Date: 2016-01-19
3  Type "notebook()" for the browser-based notebook interface.
4  Type "help()" for help.
5
6 sage: f(x)=10**(-3)*x**2+0.8*x-1.2*10**(-5)
7 sage: solve(f(x)==0,x)
8 [x == -1/50*sqrt(400000030) - 400, x == 1/50*sqrt(400000030) - ↩
     400]
9 sage: numerical_approx(-1/50*sqrt(400000030))
10 -400.000015000000
11 sage: numerical_approx( 1/50*sqrt(400000030) - 400  )
12 0.0000149999996779115
```

tex/sage/sageSnip001.sage

Donc Sage ne tombe pas dans le piège. △

Comment résoudre ce problème ? Ou, autre façon de poser la question : comment Sage a fait pour résoudre le problème ?

Utilisons les relations coefficients-racines :

$$x_1 + x_2 = -b/a \tag{34.33a}$$
$$x_1 x_2 = c/a \tag{34.33b}$$

La première lie les deux racines par des opérations de addition et soustractions, et donc n'est pas intéressantes. La seconde est bien. Si nous connaissons x_1, nous calculons

$$x_2 = \frac{c}{ax_1}. \tag{34.34}$$

Quitte à redéfinir x_1 et x_2, la solution bien calculée est :

$$x_1 = \frac{-b - \operatorname{sgn}(b)\sqrt{b^2 - 4ac}}{2a}. \tag{34.35}$$

Exemple 34.22.
Nous considérons :

$$f(x) = cos(x + \delta) - \cos(x). \tag{34.36}$$

Cela a une erreur de cancellation lorsque $|\delta| \ll |x|$. On élimine l'erreur de cancellation par

$$f(x) = -2\sin(\delta/2)\sin\left(x + \frac{\delta}{2}\right). \tag{34.37}$$

¡¡ Avertissement/question à la lectrice !! 34.23
Pourquoi la condition pour avoir l'erreur est $\delta \ll x$ et non simplement $\delta \ll 1$?

△

Exemple 34.24.
Pour

$$f(x) = \sqrt{x + \delta} - \sqrt{x}. \tag{34.38}$$

On fait la coup du binôme conjugué :

$$f(x) = \frac{\delta}{\sqrt{x+\delta}+\sqrt{x}}. \tag{34.39}$$

Plus d'erreur de cancellation, vu qu'au dénominateur nous avons une somme de deux positifs. △

Les erreurs de cancellation ne se résolvent pas en augmentant la précision des nombres donnés.

Exemple 34.25 (Dans la vie réelle)**.**
La préparation de l'exemple 17.45 nous a porté à calculer la différence entre $\exp(x)$ et $f_{30}(x)$ où f_{30} est censée être une bonne approximation de l'exponentielle. Des erreurs de cancellation sont donc à craindre.

Et en effet, le code suivant produit un résultat non déterministe :

```
f=1/152444172305856930250752000000*x^28 + ↩
   1/10888869450418352160768000000*x^27 + ↩
   1/155112100433309859840000000*x^25 + ↩
   1/3102242008666197196800000*x^24 + 1/25852016738884976640000*x↩
   ^23 + 1/5109094217170944000*x^21 + 1/1216451004088320000*x^20↩
    + 1/121645100408832000*x^19 + 1/3556874280960000*x^17 + ↩
   1/10461394944000*x^16 + 1/1307674368000*x^15 + 1/6227020800*x↩
   ^13 + 1/239500800*x^12 + 1/39916800*x^11 + 1/362880*x^9 + ↩
   1/20160*x^8 + 1/5040*x^7 + 1/120*x^5 + 1/12*x^4 + 1/6*x^3 + x ↩
   - cos(x) + 2 -exp(x)
a=numerical_approx(10)
print(f(a))
```

tex/sage/sageSnip016.sage

Voir la question ici :
https://ask.sagemath.org/question/37946/undeterministic-numerical-approximation/
△

34.3.5 Calcul d'une dérivée

Pour calculer la dérivée de f en a, il est loisible d'utiliser la formule

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}. \tag{34.40}$$

Le numérateur est alors sujet à une erreur d'absorption dans le calcul de $a+h$ et ensuite une erreur de cancellation dans le calcul de la différence.

En utilisant la formule

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a-h)}{2h} \tag{34.41}$$

nous pouvons espérer avoir une erreur de cancellation plus petite.

34.3.6 Erreur d'absorption

L'addition d'un nombre avec un nombre très différent peut faire perdre de l'information sur le plus petit. Par exemple avec 4 chiffres significatifs,

$$0.5678 \oplus 0.0001237 = 0.5679 \tag{34.42}$$

où nous avons perdu presque toute l'information du petit nombre.

Une situation particulièrement ennuyeuse est celle où justement c'est le petit nombre qui nous intéresse parce que le grand est censé se simplifier :

$$(0.0001327 \oplus 0.5678) \ominus 0.5678 = 0.5679 \ominus 0.5678 = 0.0001 \tag{34.43}$$

qui ne possède qu'un seul chiffre significatif correct alors que voyant le calcul, la réponse aurait pu être trouvée.

Moralité : si certains manipulations algébrique peuvent faire apparaitre des simplifications avant de passer le calcul à la machine, il est bon de les effectuer.

34.4 Conditionnement et stabilité

DEFooYIFAooSJbMkC

Définition 34.26.
Soit F une fonction à valeurs réelles définie sur $X \times D$ où X et D sont des espaces vectoriels réels normés. Le problème de la recherche des solutions de

$$F(x, d) = 0 \tag{34.44}$$

*est dit **stable** autour de $d_0 \in D$ si*

(1) la solution $x = x(d)$ existe et est unique pour tout d ; ItemProbStableB

(2) Pour tout $\eta > 0$, et pour tout d_0, il existe un nombre $K > 0$ tel que $\|d - d_0\| < \eta$ entraine $\|x(d) - x(d_0)\| \leqslant K \|d - d_0\|$.

La seconde condition est le fait que x soit Lipschitz[3] *sur un voisinage de d_0.*

ExooXJONooTAYZVc

Exemple 34.27 (Stabilité de la différence)**.**
Prenons le problème qui consiste à calculer la différence entre deux nombres : $x = a - b$. Cela se traduit par

$$\begin{aligned} F\colon \mathbb{R} \times \mathbb{R}^2 &\to \mathbb{R} \\ \big(x, (a, b)\big) &\mapsto x - a + b. \end{aligned} \tag{34.45}$$

Nous avons :

$$\begin{aligned} \big|x(a, b) - x(a', b')\big| &= |a - b - a' + b'| && \text{(34.46a)} \\ &\leqslant |a - a'| + |b - b'| && \text{(34.46b)} \\ &= \|(a, b) - (a', b')\|_1 && \text{(34.46c)} \end{aligned}$$

où nous avons utilisé la norme $\|.\|_1$ sur $\mathbb{R}^2$. Par la proposition 10.47 sur les équivalences de normes, le nombre $K = \sqrt{2}$ fonctionne pour toute valeurs de η.

La problème de la différence est donc un problème stable. △

Exemple 34.28 (Stabilité de la multiplication)**.**
Si a est fixé, le problème de calculer ab (b est la donnée) est stable. En effet ce problème est donné par la fonction $F(x, b) = x - ab$, dont la solution est $x(b) = ab$. Nous avons donc

$$\big|x(b) - x(b')\big| = |ab - ab'| = |a||b - b'|. \tag{34.47}$$

La constante de Lipschitz de ce problème est donc $|a|$. △

Définition 34.29.
Le nombre

$$K_{abs}(d_0, \eta) := \sup_{d \text{ tel que } |d_0 - d| < \eta} \frac{\|x(d) - x(d_0)\|_X}{\|d - d_0\|_D} \tag{34.48}$$ EqDefAABSOLU

*est appelé le **conditionnement absolu** du problème autour de d_0.*

3. Définition 12.316.

Soit $F(x,d)=0$ un problème stable de conditionnement absolu $K_{abs}(d,\eta)$. Le conditionnement relatif est défini par

$$K_{rel}(d,\eta):=K_{abs}(d,\eta)\frac{\|d\|_D}{\|x(d)\|_X}. \tag{34.49}$$

DEFEQooSXDBooYbvGrC

Le problème est dit ***bien conditionné*** *près de d si $K_{rel}(d,\eta)$ est petit.*

Exemple 34.30 (Mauvais conditionnement de la différence)**.**
Reprenons le problème de la différence, mais en fixant a. Nous avons donc $x(b)=a-b$ et le conditionnement absolu est

$$\sup\frac{|x(b)-x(b_0)|}{|b-b_0|}=1 \tag{34.50}$$

Le conditionnement relatif est :

$$K_{rel}(b_0,\eta)=\frac{|b|}{|a-b|}. \tag{34.51}$$

Et donc le problème est mal conditionné autour de a.

Autrement dit, si a' est un nombre proche de a, calculer la différence $a-a'$ est un problème mal conditionné. △

Exemple 34.31 (Bon conditionnement de la multiplication)**.**
Pour le problème $F(x,b)=x-ab$ nous avons

$$K_{abs}=\sup_{b'}\frac{|ab-ab'|}{|b-b'|}=|a|. \tag{34.52}$$

Et aussi

$$K_{rel}=a\frac{|b|}{|ab|}=1. \tag{34.53}$$

Le conditionnement relatif du problème de la multiplication est donc toujours 1. Il est donc toujours un problème bien conditionné. △

Ne pas confondre :

Le conditionnement provient du problème lui-même.

La stabilité provient de l'algorithme de résolution.

Exemple 34.32 (Un problème mal conditionné)**.**
Le système

$$\begin{cases} 2.1x+3.5y=8 & \text{(34.54a)}\\ 4.19x+7.0y=15 & \text{(34.54b)}\end{cases}$$

Solution : $x=100$, $y=-57.714285\ldots$ (périodique)

Perturbons : nous remplaçons 4.19 par 4.192. L'erreur relative est : 4.77×10^{-4}.

Solution : $\bar{x}=125$, $\bar{y}=-72.714285\ldots$, avec donc erreur relative de 0.26. Autrement dit : l'erreur relative sur la solution est grande même avec une petite erreur relative sur la donnée.

C'est un problème mal conditionné.

Le fait est que c'est une intersection de deux droites presque parallèles. Donc effectivement une petite perturbation d'une des deux droites donne une grande perturbation du point d'intersection.

Le fait est qu'un ordinateur effectue *toujours* une perturbation, au moins de l'ordre 10^{-16} pour ne fut-ce que représenter les nombres. C'est-à-dire une perturbation sur les six nombres définissant le système. Il n'y a donc pas d'espoir d'obtenir un algorithme donnant une bonne réponse. △

Un résultat pratique pour étudier le conditionnement d'un problème est le suivant.

CorConditionnementNormeNabla

Corolaire 34.33.
Soit $x=x(d)$ un problème stable. Supposons $\mathbb{D}$ de dimension finie, supposons que U est ouvert dans $\mathbb{D}$. Supposons encore $x\colon U\to\mathbb{R}$ différentiable en d_0. Alors quand η est petit, on a

$$K^{\eta}_{abs}(d_0)\sim\|\nabla x(d_0)\|. \tag{34.55}$$

LemITCxqyS

Lemme 34.34.
Tout problème de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^1(\mathbb{R})$ est stable.

Démonstration. Il faut démontrer qu'une fonction C^1 sur $\mathbb{R}$ vérifie automatiquement la condition (2) de la définition de la stabilité. Pour cela, remarquons qu'une fonction C^1 possède une dérivée continue, et donc bornée sur tout compact [4]

Prenons $\eta > 0$ et $d_0 \in \mathbb{R}$ et puis un d tel que $|d - d_0| < \eta$. Par le théorème des bornes atteintes, la fonction x' est bornée sur l'intervalle $[d_0 - \eta, d_0 + \eta]$. Appelons K un majorant de x' sur cet intervalle. La fonction

$$x(d) = x(d_0) + K|d - d_0| \tag{34.56}$$

majore $x(d)$, et donc on a

$$\big|x(d) - x(d_0)\big| \leqslant K|d - d_0|. \tag{34.57}$$

Attention : vérifier si ce raisonnement est correct avec $d_0 > d$, et adapter au besoin. □

ExRZrOeoi

Exemple 34.35.
Un exemple de problème stable de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^0(\mathbb{R})\backslash C^1(\mathbb{R})$.

La fonction

$$x(d) = \begin{cases} 0 & \text{si } d \geqslant 0 \\ x & \text{si } d < 0 \end{cases} \tag{34.58}$$

est continue, mais pas C^1 (non dérivable en $x = 0$). La dérivée est partout bornée par 1, et donc le problème est stable.

Un autre exemple très classique serait de prendre $x(d) = |d|$. Dans ce cas, on peut prendre n'importe que η et $K = 1$. Le calcul est que

$$|x(d) - x(d_0)| < K|d - d_0| \tag{34.59a}$$
$$\big||d| - |d_0|\big| < |d - d_0|. \tag{34.59b}$$

Cette dernière inéquation est correcte, comme on peut le voir en mettant au carré les deux membres. △

PIluknK

Exemple 34.36.
Un exemple de problème instable de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^0(\mathbb{R})$.

Un exemple assez classique de fonction dont la dérivée n'est pas bornée sans pour autant que la fonction aie un comportement immoral [5] est $x \mapsto \sqrt{x}$. Afin d'avoir une fonction définie sur $\mathbb{R}$ tout entier, nous regardons la fonction

$$x(d) = \sqrt{|d|}. \tag{34.60}$$

Si nous considérons maintenant $d_0 = 0$ et n'importe quel η, nous avons

$$\frac{|x(d) - x(d_0)|}{|d - d_0|} = \frac{\sqrt{d}}{d} = \frac{1}{\sqrt{d}}. \tag{34.61}$$

Il n'est pas possible de trouver un K qui majore ce rapport. Le problème est donc mal conditionné.

Attention : dans ce calcul nous avons supposé $d > 0$. Pensez à adapter au cas $d < 0$. △

Exemple 34.37 (Problème bien conditionné avec algorithme instable)**.**
Soit à calculer

$$I_n = \frac{1}{e}\int_0^1 x^n e^x dx \tag{34.62}$$

4. Un compact est un ensemble fermé et borné, typiquement un intervalle du type $[a, b]$.
5. Penser à $x \mapsto x \sin(1/x)$.

avec $n \geqslant 0$. Par partie, nous obtenons :

$$I_n = 1 - nI_{n-1}. \tag{34.63}$$

D'autre part, $I_0 = \frac{e-1}{e}$, $I_1 = \frac{1}{e}$. Puis par récurrence, c'est tout en main.

Du côté de l'ordinateur, nous lui donnons forcément une approximation de I_1, parce que nous lui donnons une approximation de e. Soit l'erreur ϵ_1 sur I_1.

Sans démonstration :

Lemme 34.38.
Nous avons $\lim_{n\to\infty} I_n = 0$.

Mais numériquement, il n'est pas possible de rester longtemps sous ϵ_1 parce que nous n'espérons pas avoir une erreur plus petite que ça. Donc à partir du moment où $I_n < \epsilon_1$, les valeurs sont toutes complètement fausses. Cela est le mieux que l'on puisse espérer. Mais la réalité est pire.

En réalité, en lançant le calcul sur un ordinateur, les valeurs sont même croissantes avec n à partir d'un certain moment.

On peut étudier l'erreur et montrer que l'erreur est donnée par :

$$\epsilon_n = (-1)^{n-1} n! \epsilon_1. \tag{34.64}$$

Mais comme la factorielle est tellement forte que c'est sans espoir d'aller loin en essayant très fort de donner une petite erreur sur ϵ_1.

△

Il existe heureusement un algorithme stable pour cette intégrale. La formule est :

$$I_{n-1} = \frac{1}{n}(1 - I_n). \tag{34.65}$$

Si nous savons un I_N avec N grand, cette formule donne les I_i avec $i = N, N-1, \dots, 2$. Posons donc $I_N = a \in \mathbb{R}$ n'importe comment. Donc ϵ_N est grand. Mais il se trouve que l'erreur sur ϵ_1 est donnée par

$$\epsilon_1 = \frac{(-1)^{N-1}}{N!}\epsilon_N. \tag{34.66}$$

Donc même en prenant vraiment n'importe quoi pour I_N, nous obtenons de bonnes approximations pour I_i avec les petits i. Même avec $I_{20} = 1000$ (qui est complètement faux), nous trouvons énormément de chiffres significatifs corrects pour I_1.

34.4.1 Comment choisir et penser le K ?

La formule (34.48) contient une formule qui ressemble étrangement à la dérivée. La stabilité d'un problème est très liée à la dérivée de F. La stabilité et la dérivée ne sont pas les mêmes choses, mais il n'est pas mauvais de penser au K de la stabilité comme la dérivée. Ou plus précisément : le supremum de la dérivée.

Un fil conducteur du lemme 34.34 et des exemples 34.35, 34.36 est que l'on a un K qui fonctionne lorsque la dérivée est bornée sur l'intervalle $]d_0 - \eta, d_0 + \eta[$. Dans le cas où ce supremum existe, le prendre en guise de K fonctionne souvent.

Il faut cependant parfois faire acte d'imagination. La fonction $x \mapsto |x|$ n'est pas dérivable en 0. Il n'empêche que $K = 1$ fait fonctionner la définition de la stabilité. Remarquez que $K = 1$ est le supremum de la dérivée là où elle existe.

À partir du moment où c'est clair que le K est le supremum de la dérivée, on comprend pourquoi c'est le gradient qui arrive dans le corolaire 34.33. En effet, le gradient indique la direction de plus grande pente. C'est donc bien dans cette direction qu'il faut chercher la « plus grande dérivée ».

Proposition 34.39.
Pour le problème stable $x = x(d)$ avec $x \in C^1(\mathbb{R}^n, \mathbb{R})$, on a

$$K_{abs}(d) \sim \|dx_d\| \tag{34.67}$$

où dx_d désigne la différentielle de x en d et la norme est la norme opérateur.

34.5 Un peu de points fixes

SECooWUVTooMhmvaW

34.5.1 Choix de la fonction à point fixe

Pour l'équation $f(x) = 0$, il existe une infinité de fonctions g pour lesquelles l'équation est équivalente à $x = g(x)$.

Exemple : $f(x) = x^2 - 2 - \ln(x)$, nous pouvons faire

(1) $x = x^2 - 2 - \ln(x) + x$

(2) Poser $x^2 = 2 + \ln(x)$ et donc

$$x = -\sqrt{2 + \ln(x)} \tag{34.68a}$$

$$x = \sqrt{2 + \ln(x)}. \tag{34.68b}$$

(3) Ou encore

$$x = \frac{2 + \ln(x)}{x} \tag{34.69}$$

où nous savons déjà que $x \neq 0$ parce que $x = 0$ n'est pas dans le domaine de f.

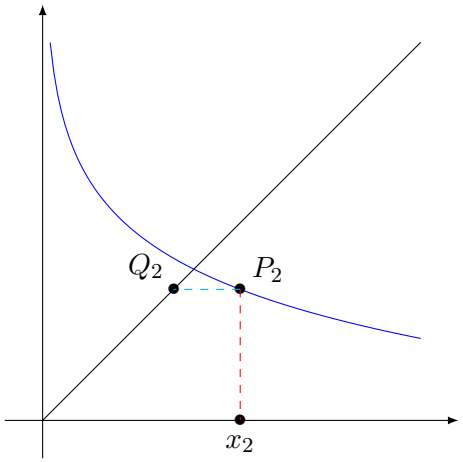
(4) Ou par l'exponentielle :

$$x = e^{x^2 - 2}. \tag{34.70}$$

Dans tous ces cas nous pouvons construire une suite (x_n) en posant un nombre arbitraire pour x_0 et ensuite la récurrence

$$x_{n+1} = g(x_n). \tag{34.71}$$

Graphiquement, la solution de l'équation est l'intersection entre les courbes $y = x$ et $y = g(x)$. Un petit dessin pour montrer la convergence :



Attention : cette méthode ne converge pas toujours. Parfois elle converge de façon monotone, et parfois pas. Le choix de la fonction g qui fait $x = g(x)$ peut énormément changer la vitesse de convergence.

Théorème 34.40 (Condition suffisante pour existence d'un point fixe)**.**
Une fonction continue $f : [a, b] \to [a, b]$ admet au moins un point fixe dans $[a, b]$.

Théorème 34.41 (Condition suffisante pour l'unicité)**.**
Soit f continue sur $[a, b]$ avec $g(x) \in [a, b]$ pour tout $x \in [a, b]$. Supposons qu'il existe $0 < k < 1$ tel que pour tout $x \in [a, b]$ nous ayons $|g'(x)| \leqslant k$ alors

(1) La fonction g possède un unique point fixe dans $[a, b]$.

(2) Pour tout $x_0 \in [a, b]$, tous les termes de la suite $x_{n+1} = g(x_n)$ sont dans $[a, b]$.

(3) Ladite suite (x_n) converge vers le point fixe.

Théorème 34.42.
Soit f continue sur $[a,b]$ avec $g(x)\in[a,b]$ pour tout $x\in[a,b]$. Supposons

(1) qu'il existe $0<k<1$ tel que pour tout $x\in[a,b]$ nous ayons $|g'(x)|\leqslant k$ et

(2) g est p fois dérivable sur $[a,b]$.

(3) $g'(\alpha)=g''(\alpha)=\ldots=g^{(p-1)}(\alpha)$ et $g^{(p)}(\alpha)\neq 0$ où α est l'unique point fixe.

Alors la suite (x_n) converge avec un ordre p.

Exemple 34.43.
Nous reprenons

$$f(x)=x^2-2-\ln(x). \tag{34.72}$$

Et nous voulons résoudre $f(x)=0$. Graphiquement c'est l'intersection entre $y=x^2-2$ et $y=\ln(x)$. Il est vite tracé de savoir qu'il y a deux solutions : $\alpha_1\in[0,1]$ et $\alpha_2\in[\sqrt{2},2]$.

Déjà un petit problème : l'intervalle $[0,1]$ ne va pas parce que f n'y est pas continue. Un petit raffinement d'analyse nous fournit $\alpha_1\in[e^{-2},1]$.

Nous avons au moins les fonctions de points fixes suivantes :

$$g_1(x)=\sqrt{2+\ln(x)} \tag{34.73a}$$

$$g_2(x)=e^{x^2-2}. \tag{34.73b}$$

Pour la première, il y avait un $\pm$ qui a été négligé parce que nous savons que les deux solutions cherchées sont positives. Travaillons avec la première. D'abord

$$g_1'(x)=\frac{1}{2x\sqrt{2-\ln(x)}}. \tag{34.74}$$

Nous avons $\lim_{x\to e^{-2}}g_1'(x)=+\infty$. Il ne sera donc pas possible de trouver $0<k<1$ tel que $|g'(x)|\leqslant k$. Tentons quand même la méthode :

$$x_0=0.5 \tag{34.75}$$

Il se fait que cela est plus proche de α_1 que de α_2. Mais en réalité la suite converge vers α_2.

Passons à la seconde méthode.

$$g_2'(x)=2xe^{x^2-2}. \tag{34.76}$$

Sur l'intervalle $[e^{-2},1]$, g_2' est croissante et prend toutes ses valeurs dans $[e^{-2},1]$. Nous pouvons prouver que

$$|g_2'(x)|\leqslant 2e^{-1}<1. \tag{34.77}$$

Donc poser $k=2e^{-1}$ fait fonctionner la proposition. Donc quel que soit le x_0 pris dans cet intervalle, nous aurons une suite convergente vers un point fixe à l'intérieur de l'intervalle. C'est-à-dire convergente vers α_1.

Cela est un exemple de problème pour lequel changer de fonction g change réellement la vie. △

34.5.2 Convergence quadratique

DEFooSUTRooAcXXjj

Définition 34.44.
*Une suite (x_n) a une convergence **quadratique** vers α si elle converge vers α et si il existe un C tel que pour tout n nous ayons*

EQooMWBIooLGashp

$$\|x_{n+1}-\alpha\|\leqslant C\|x_n-\alpha\|^2. \tag{34.78}$$

Il est bien entendu possible de parler de convergence quadratique si la relation (34.78) a lieu seulement à partir d'un certain indice.

Le lemme suivant donne l'importance du choix de point de départ lorsqu'on utilise une méthode itérative dont la convergence est quadratique.

LEMooLQMAooICcmrn

Lemme 34.45.
Soit une suite $x_n \to \alpha$ de convergence quadratique. Si $\|x_0 - \alpha\| \leqslant r$ alors

$$\|x_n - \alpha\| \leqslant \frac{1}{C}(Cr)^{2^n} \tag{34.79}$$

EQooVYRIooTxetPn

Démonstration. Nous pourrions directement prouver la formule (34.79) par récurrence, mais nous allons la reconstruire un peu. Nous cherchons

$$\|x_n - \alpha\| \leqslant C^{k(n)} r^{2^n}. \tag{34.80}$$

EQooZCZVooHqpkqs

Nous avons les inégalités

$$\begin{aligned}\|x_{n+1} - \alpha\| &\leqslant C\|x_n - \alpha\|^2 && \text{(34.81a)}\\ &\leqslant CC^{2k(n)} r^{2^{n+1}} && \text{(34.81b)}\\ &= C^{2k(n)+1} r^{2^{n+1}} && \text{(34.81c)}\end{aligned}$$

d'où nous voyons que la fonction k doit vérifier

$$\begin{cases} k(0) = 0 & \text{(34.82a)}\\ k(n+1) = 2k(n) + 1 & \text{(34.82b)}\end{cases}$$

La première équation est l'hypothèse $\|x_0 - \alpha\| \leqslant r$ comparée à la formule (34.80). Il est vite vérifié que $k(n) = 2^n - 1$. D'où le résultat. □

Si le point de départ est choisi de façon à avoir $Cr < 1$ alors nous avons là un très bon majorant parce qu'il s'agit d'un majorant convergeant très rapidement vers zéro. Si au contraire $Cr > 1$ alors ce majorant ne sert à rien.

34.46.
Le fait d'avoir une convergence quadratique signifie que le nombre décimales correctes double (environ) à chaque itération, dans n'importe quelle base. En effet supposons que x_n ait k décimales correctes ; cela signifie que $|x_n - \alpha| \sim 10^{-k}$. Donc

$$|x_{n+1} - \alpha| \lessapprox M10^{-2k}. \tag{34.83}$$

Cela est le double de décimales correctes de $|x_n - \alpha|$, moins l'ordre de grandeur de M.

Pour la méthode de bisection, le nombre de décimales augmente de 1 à chaque itération, mais seulement en base 2. En base 10, de façon générique [6] il faut entre 3 et 4 itérations pour avoir une décimale de plus.

NTooVXLXooXlAGEq

34.47 (Condition d'arrêt[750])**.**
D'autre part, lorsqu'une méthode a une convergence quadratique, nous avons un test d'arrêt. Pour ce voir, nous avons la limite

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} \leqslant \lim_{n\to\infty} \frac{C|x_n - \alpha|^2}{|x_n - \alpha|} = 0. \tag{34.84}$$

Cette limite est alors également valable sans les valeurs absolues et si nous soustrayons $x_n - \alpha$ au numérateur, la limite devient -1 :

$$-1 = \lim_{n\to\infty} \frac{x_{n+1} - x_n}{x_n - \alpha}. \tag{34.85}$$

Ou encore

$$\lim_{n\to\infty} \frac{x_n - x_{n+1}}{x_n - \alpha} = 1. \tag{34.86}$$

Cela a pour conséquence que si n est grand,

6. C'est-à-dire sauf coup de malchance ou coup de chance.

(1) x_{n+1} a le même ordre de grandeur que $x_n - \alpha$.

(2) $x_n - x_{n+1}$ et $x_n - \alpha$ ont le même signe.

Donc si nous voulons une approximation de α avec une erreur ϵ, il suffit d'arrêter le calcul lorsque $|x_{n+1} - x_n| \leqslant \epsilon$. Et ce faisant nous savons de plus si l'approximation est par excès ou par défaut.

34.5.3 Convergence

PROPooRPHKooLnPCVJ

Proposition 34.48 (Convergence d'une méthode de point fixe[750])**.**
Soit $g\colon \mathbb{R} \to \mathbb{R}$ de classe C^1 et α un point fixe attractif[7] *de g. Soit k tel que $|g'(\alpha)| < k < 1$ et δ tel que $\|g'\|_{B(\alpha,\delta)} < k$.*

Alors

ITEMooOQKMooTRSvUo

(1) La fonction g est k-contractante[8] *sur $B(\alpha, \delta)$.*

ITEMooFTAQooPBsBcR

(2) Nous avons $g\big(B(\alpha,\delta)\big) \subset B(\alpha,\delta)$.

ITEMooFSOAooKlcxih

(3) Pour tout $x_0 \in B(\alpha,\delta)$ la suite $x_{n+1} = g(x_n)$ converge vers α et

$$|x_n - \alpha| \leqslant |x_0 - \alpha|k^n. \tag{34.87}$$

Si de plus $g'(\alpha) = 0$ et g est de classe C^2 alors nous avons convergence quadratique (définition 34.44).

Démonstration. Vu que α est un point fixe attractif de g nous pouvons considérer un k tel que $|g'(\alpha)| < k < 1$. Et comme g est de classe C^1, la fonction g' est continue et donc bornée sur toute boule du type $\overline{B(\alpha,\delta)}$. Soit δ le plus grand nombre tel que $\|g'\|_{\overline{B(\alpha,\delta)}} \leqslant k$. Nous notons $I = \overline{B(\alpha,\delta)}$ pour cette valeur de δ.

Pour tout $x \in I$ nous avons, en utilisant le théorème des accroissements finis 12.192(2) : SUBEQooYXLHooSCnnRA

$$\begin{aligned}
|g(x) - \alpha| &= |g(x) - g(\alpha)| && \text{(34.88a)}\\
&\leqslant \sup_{t\in I} |g'(t)||x - \alpha| && \text{(34.88b)}\\
&\leqslant k|x - \alpha| && \text{(34.88c)}\\
&< \delta && \text{(34.88d)}
\end{aligned}$$

parce que $k < 1$ et $|x - \alpha| \leqslant \delta$. Par conséquent $g(x) \in B(\alpha,\delta)$. Cela prouve le point (2). Pour le point (1), soient $x, y \in B(\alpha,\delta)$ et

$$|g(x) - g(y)| \leqslant \sup_{a\in I} |g'(a)||x - y| \leqslant k|x - y|. \tag{34.89}$$

Pour le point (3) nous avons $|g(x_n) - \alpha| \leqslant k|x_n - \alpha|$, c'est-à-dire

$$|x_{n+1} - \alpha| \leqslant k|x_n - \alpha|. \tag{34.90}$$

Le résultat annoncé s'obtient par récurrence sur n.

En ce qui concerne la convergence quadratique, c'est du Taylor (proposition 12.471). Développons $g(x_n)$ autour de $g(\alpha)$:

$$g(x_n) = g(\alpha) + g'(\alpha)(x_n - \alpha) + \frac{1}{2}(x_n - \alpha)^2\epsilon(x_n - \alpha) \tag{34.91}$$

avec $\lim_{t\to 0} \epsilon(t) = 0$. En posant $C = \frac{1}{2}\sup_{t<\delta} |\epsilon(t)|$ nous avons $|g(x_n) - g(\alpha)| \leqslant C|x_n - \alpha|^2$, c'est-à-dire

$$|x_{n+1} - \alpha| \leqslant C|x_n - \alpha|^2. \tag{34.92}$$

□

7. Définition 17.29.
8. Définition 17.33

Ce corolaire est une paraphrase de la proposition 34.48. Il en retient seulement les points intéressants en pratique.

CORooHKZCooEXRzcW

Corolaire 34.49.
Soit α une solution de l'équation $x = g(x)$, avec g continue sur un voisinage de α et dérivable dans l'intérieur. Nous supposons que

$$|g'(\alpha)| < 1. \tag{34.93}$$

Alors il existe un rayon δ tel que si $x_0 \in B(\alpha, \delta)$, la suite (x_n) converge vers α.

Certes cette proposition demande moins d'hypothèses, mais en réalité, il ne donne pas de vrais moyens de choisir un point de départ x_0. Avec les deux théorèmes précédents, nous pouvions prendre x_0 n'importe où dans $[a, b]$. Le fait est que pour choisir x_0 nous pouvons tracer et donner à la main un x_0 proche de ce qui semble être α. Si ça ne converge pas, il faut donner un x_0 plus proche. La proposition nous assure que si nous jouons bien à choisir x_0 très proche, la suite finira par converger.

Notons que le corolaire 34.49 a encore l'inconvénient de demander de calculer $g'(\alpha)$ alors que α est inconnu. La résolution de l'inéquation $|g'(x)| < 1$ nous donne un certain nombre d'intervalles dans $\mathbb{R}$.

Soient I_n les intervalles solutions de l'inéquation. Si $\alpha \in I_n$ alors la méthode converge. Sinon, c'est pas garanti. En tout cas nous ne devons pas savoir réellement α pour appliquer le théorème. Il suffit de savoir que α est dans un des I_n.

34.6 Méthode de Newton

SECooIKXNooACLljs

L'objectif de la méthode de Newton est d'évaluer une racine α de l'équation $f(x) = 0$ lorsque nous avons déjà une approximation x_0 de la racine α.

C'est la méthode de Newton qui est à l'origine de la suite de la proposition 1.413 donnant une suite dans $\mathbb{Q}$ qui converge vers $\sqrt{A}$.

DEFooXSOQooAnWqKM

Définition 34.50.
*Le nombre α est une **racine simple** de l'équation $f(x) = 0$ si $f(\alpha) = 0$ et $f'(\alpha) \neq 0$. Le nombre α est une **racine multiple** d'ordre r de $f(x) = 0$ si*

$$f(\alpha) = f'(\alpha) = \ldots = f^{(r-1)}(\alpha) = 0 \tag{34.94}$$

et $f^{(r)}(\alpha) \neq 0$.

Exemple 34.51.
La fonction $x \mapsto x^3$ admet une racine d'ordre 3 en $x = 0$. △

34.6.1 « Justification » par la formule de Taylor

Soit une fonction f continue et dérivable sur $[a, b]$. Soit α une racine de f et x_n une de ses approximations. Nous notons l'erreur θ et nous avons $\alpha = x_n + \theta$. Du coup nous avons $f(x_n + \theta) = f(\alpha) = 0$.

Écrivons la série de Taylor du théorème 12.461 autour de x_n : il existe une fonction $\epsilon\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \epsilon(t) = 0$ telle que

$$f(\alpha) = f(x_n + \theta) = f(x_n) + \theta f'(x_n) + \theta^2 \epsilon(\theta). \tag{34.95}$$

EQooOPUBooYaznay

Nous isolons le θ du terme d'ordre 1 en nous souvenant que le membre de gauche est nul :

$$\theta = -\frac{f(x_n) - \theta^2 \epsilon(\theta)}{f'(x_n)} \tag{34.96}$$

Vu que $\alpha = x_n + \theta$, nous pouvons écrire

$$\alpha = x_n - \frac{f(x_n) + \theta^2\epsilon(\theta)}{f'(x_n)}. \tag{34.97}$$

Il est donc raisonnable de poser

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{34.98}$$

en espérant que cela soit une meilleure approximation de α que x_n.

En tout cas l'erreur sur x_{n+1} est

$$\alpha - x_{n+1} = x_n + \theta - x_n + \frac{f(x_n) + \theta^2\epsilon(\theta)}{f'(x_n)} = \theta + \frac{f(x_n) + \theta^2\epsilon(\theta)}{f'(x_n)}, \tag{34.99}$$

qui ne doit pas être fondamentalement plus grand que θ dès que θ est petit, surtout que si x_n est une approximation de α, nous pouvons espérer que $f(x_n)$ soit également petit. Là où les choses peuvent déraper en grand, c'est si $f'(x_n)$ est petit.

Cette méthode de Newton ne converge pas toujours. Le pire est lorsque par malheur il y a une bosse pas loin de la racine. Alors il y a un risque de tomber sur $f'(x_{n+1}) = 0$ ou en tout cas très proche de zéro. Dans ce cas le point x_{n+2} est envoyé très loin.

34.6.2 « Justification » par points fixes

SUBSECooIBLNooTujslO

Nous savons que pour résoudre $f(x) = 0$ par une méthode de point fixe, il y a de nombreux choix possibles de fonctions g telles que $g(x) = x$ donne la même solution que $f(x) = 0$. Soit α une solution de $f(x) = 0$ et cherchons une fonction g de la forme

$$g(x) = x - kf(x). \tag{34.100}$$

EQooYVFIooJXnJXa

Nous savons par la proposition 34.48 que la fonction g donne une convergence quadratique lorsque $g'(\alpha) = 0$. Pour la forme (34.100) nous avons $g'(\alpha) = 1 - kf'(\alpha)$, ce qui nous donne l'idée de poser $k = \frac{1}{f'(\alpha)}$.

Le fait est que $f'(\alpha)$ n'est pas connu, mais nous pouvons l'approximer par $f'(x)$ lorsque x est proche de α. D'où l'idée de considérer la fonction

$$g(x) = x - \frac{f(x)}{f'(x)}, \tag{34.101}$$

et donc la suite $x_{n+1} = g(x_n)$ c'est-à-dire

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{34.102}$$

Dès que x_n est proche de α, sous l'hypothèse (raisonnable par continuité) que $f'(x_n)$ soit proche de $f'(\alpha)$, la méthode devrait donner une convergence quadratique.

Remarque 34.52.
Cette justification par points fixes n'est pas vraiment différente de celle par Taylor parce que Taylor est utilisé dans la preuve de la proposition 34.48.

Définition 34.53 (Méthode de Newton)**.**
*La **méthode de Newton** pour la fonction f est la suite définie par récurrence*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{34.103}$$

Cette définition ne précise pas la valeur de x_0, ni de condition d'arrêt.

34.6.3 Convergence de la méthode de Newton

THOooDOVSooWsAFkx

Théorème 34.54 (Convergence quadratique de la méthode de Newton[750])**.**
Soit f une fonction continue vérifiant $f(\alpha)=0$ et $f'(\alpha)\neq 0$. Nous considérons la fonction

$$g(x)=x-\frac{f(x)}{f'(x)} \tag{34.104}$$

que nous supposons être de classe C^2.

Si C est une majoration de $\|g''\|$ sur un intervalle contenant α, alors en posant $\delta=1/C$ nous avons

ITEMooVXSKooWCVWQc

(1) *La boule $B(\alpha,\delta)$ est préservée par g : $g\big(B(\alpha,\delta)\big)\subset B(\alpha,\delta)$.*

ITEMooZPSXooGgbfhG

(2) *Pour tout $x_0\in B(\alpha,\delta)$ nous avons convergence quadratique vers α de la suite définie par $x_{n+1}=g(x_n)$.*

ITEMooZCXZooCjeWPl

(3) *Nous avons l'estimation*

$$|x_n-\alpha|\leqslant\frac{1}{C}\big(C|x_0-\alpha|\big)^{2^n} \tag{34.105}$$

EQooFAIPooDpoNWK

où C est la constante de la définition de convergence quadratique.

Démonstration. Nous commençons par calculer la dérivée de g :

$$g'(x)=\frac{f(x)f''(x)}{f'(x)^2}, \tag{34.106}$$

d'où nous déduisons que $g'(\alpha)=0$. Ensuite nous utilisons abondamment la formule des accroissements finis (théorème 12.314) en commençant par

$$|g(t)-g(\alpha)|\leqslant\|g'\|_{[t,\alpha]}|t-\alpha| \tag{34.107}$$

EQooZITHooEbGVKG

où par $\|f\|_A$ nous entendons la norme uniforme de f sur A, c'est-à-dire $\|f\|_A=\sup_{x\in A}\|f(x)\|$. Note : nous écrivons $[t,\alpha]$, mais ça pourrait être $[\alpha,t]$.

Si $x\in[t,\alpha]$ alors

$$\begin{aligned}|g'(x)|&=|g'(x)-g'(\alpha)| &(34.108a)\\ &\leqslant\|g''\|_{[x,\alpha]}|x-\alpha| &(34.108b)\\ &\leqslant\|g''\|_{[x,\alpha]}|t-\alpha| &(34.108c)\\ &\leqslant\|g''\|_{[t,\alpha]}|t-\alpha|. &(34.108d)\end{aligned}$$

En particulier, $\|g'\|_{[t,\alpha]}\leqslant\|g''\|_{[t,\alpha]}|t-\alpha|$, et nous pouvons continuer les majorations (34.107) :

$$|g(t)-g(\alpha)|\leqslant\|g''\|_{[t,\alpha]}|t-\alpha|^2. \tag{34.109}$$

La fonction g étant de classe C^2, la dérivée seconde g'' est bornée (nous supposons déjà travailler sur un compact contenant α). Soit C une borne. Nous sommes en mesure de prouver le point (1) avec $\delta=1/C$. En effet si $t\in B(\alpha,1/C)$ alors

$$|g(t)-\alpha|=|g(t)-g(\alpha)|\leqslant C|t-\alpha|^2\leqslant C\frac{1}{C^2}=\frac{1}{C}, \tag{34.110}$$

ce qui prouve que $g(t)\in B(\alpha,1/C)$.

Le point (2) se prouve de la même manière : si $x_n\in B(\alpha,1/C)$ alors

$$|x_{n+1}-\alpha|=|g(x_n)-g(\alpha)|\leqslant C|x_n-\alpha|^2, \tag{34.111}$$

ce qui est bien la convergence quadratique.

La majoration du point (3) s'obtient par récurrence sur n. Pour $n=0$, la relation (34.105) devient $|x_0-\alpha|\leqslant|x_0-\alpha|$ qui est vraie. Ensuite par la convergence quadratique et la récurrence,

$$|x_{n+1}-\alpha|\leqslant C|x_n-\alpha|^2\leqslant C\Big[\frac{1}{C}(C|x_0-\alpha|)^{2^n}\Big]^2=\frac{1}{C}\big[C|x_0-\alpha|\big]^{2^{n+1}}. \tag{34.112}$$

□

34.55.
Dans les cas pratiques, nous commençons souvent par résoudre l'équation $f(x) = 0$ par dichotomie. Au moment où nous sommes assez proche de la solution nous commençons Newton.

La raison est que la dichotomie fonctionne toujours : nous allons toujours nous approcher de la solution. Si par contre le point de départ est mal choisi, la méthode de Newton peut envoyer n'importe où, y compris très loin de la solution.

La proposition suivante nous indique que dans le cas d'une fonction convexe, le choix de point de départ de la méthode de Newton n'est pas tellement crucial parce qu'ils sont tous bons. De plus la convergence se faisant de façon décroissante (si on part de la droite), nous savons que le résultat sera une approximation par excès de α.

PROPooVTSAooAtSLeI

Proposition 34.56 (Newton dans le cas convexe)**.**
Soit f de classe C^2 et une racine α telle que $f'(\alpha) > 0$. Soit $b > \alpha$ tel que f soit convexe sur $[\alpha, b]$.

Alors pour tout $x_0 \in [\alpha, b]$ la suite de la méthode de Newton est

(1) décroissante

(2) reste dans $[\alpha, b]$

(3) converge vers α.

Démonstration. Nous savons par la proposition 17.83(2) que la fonction f' est croissante, et par hypothèse $f'(\alpha) > 0$, donc sur $[\alpha, b]$ nous avons $f' > 0$. Par conséquent, nous avons aussi $f > 0$ sur $[\alpha, b]$.

Le graphe de f est au dessus de la tangente de f en $x = x_n$ (proposition 17.89). Si nous nommons t_x la fonction qui donne la tangente en x nous avons $t_{x_n}(\alpha) < 0$ parce que $f(\alpha) = 0$. Par conséquent

$$t_{x_n}(x) = 0 \tag{34.113}$$

pour $\alpha < x < x_n$. Cela prouve que $x_{n+1} \in [\alpha, b]$, et que (x_n) est une suite décroissante

Étant donné que (x_n) est une suite décroissante dans le compact $[\alpha, b]$, elle est convergente. Notons β sa limite. Nous avons la relation de récurrence

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{34.114}$$

En passant à la limite $n \to \infty$ nous avons l'équation

$$\beta = \beta - \frac{f(\beta)}{f'(\beta)}. \tag{34.115}$$

Vu que $f(x) > 0$ sur $]\alpha, b]$ nous avons automatiquement $\beta = \alpha$. □

34.6.4 Racine carrée par la méthode de Newton

Nous avons vu dans la proposition 1.413 une suite dont la limite était $\sqrt{A}$.

EXooDLSVooMHPpcl

Exemple 34.57.
Utilisation de la méthode de Newton pour calculer $\sqrt{A}$. La première idée serait d'appliquer le théorème 34.54 à la fonction $f(x) = x - \sqrt{A}$. Je vous laisse voir pourquoi ça ne marche pas.

Nous faisons $f(x) = x^2 - A$. Le calcul donne :

$$g(x) = x - \frac{x^2 - A}{2x} = \frac{1}{2}\big(x + \frac{A}{x}\big). \tag{34.116}$$

Cela explique la formule (1.587). △

34.6.5 Formalisation de l'algorithme

La méthode de Newton consiste a exprimer la solution x de $f(x) = 0$ avec $f \in C^1(\mathbb{R})$ comme limite d'une suite $\{x_n\}_{n\in\mathbb{N}}$ définie par récurrence par la formule

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{34.117}$$

où x_0 est arbitraire.

Si on veut exprimer cela en termes d'algorithmes, nous disons que l'algorithme de Newton est donné par la suite de problèmes

$$F_n(x_{n+1}, x_n, f) = x_{n+1} - x_n + \frac{f(x_n)}{f'(x_n)}. \tag{34.118}$$ EqFPourNewtonUn

La donnée du problème est la fonction f, et rien que elle.

Plus précisément, une fois que la fonction f est donnée, il existe une infinité de problèmes : pour chaque $a \in \mathbb{R}$ nous avons le problème

$$G_a(x_n, f) = x - a + \frac{f(a)}{f'(a)}. \tag{34.119}$$

La méthode de Newton consiste à sélectionner une partie de ces problèmes de la façon suivante :

$$\begin{cases} F_0 = G_{x_0} & \text{(34.120a)} \\ F_n = G_{x_n}. & \text{(34.120b)} \end{cases}$$

Le problème F_0 fournit un nombre x_1 qui nous permet de sélectionner le problème G_{x_1} qui va fournir le nombre x_2, etc.

Au moment de calculer le conditionnement de F_n, nous ne devons pas voir x_{n-1} comme fonction de x_0 et de la donnée f. Il ne faut donc pas dériver à travers les x_n.

Proposition 34.58.
Si une racine est multiple, alors l'ordre de convergence de la méthode de Newton est 1.

Voici un algorithme possible :

```
1  def Newton(f,x0,toll,maxit):
2      fp=f.derivative()
3      n=0
4      x=x0
5      diff=toll+1
6      while abs(diff)>toll and n<maxit:
7          n=n+1
8          diff=-f(x)/fp(x)
9          x=x+diff
10     return x,n
```

tex/frido/codeSnip_2.py

Commentaires :

(1) Notons que dans un langage vraiment numérique comme Matlab, il faut passer f' en argument.

(2) Dans le `while` il faudrait mettre $x_{n+1} - x_n$ (en valeur absolue), mais cette différence est aussi utilisée pour calculer x_{n+1} donc on la calcule une seule fois.

(3) Il faudrait faire une vérification sur $f(x_n) \neq 0$. Il n'y a pas tellement de choix que de changer le point initial.

34.6.6 Caractéristiques

L'algorithme de Newton a les caractéristiques suivantes :

(1) Pour résoudre le problème numéro n, il faut avoir résolu le problème numéro $n-1$.
(2) Aucune des solutions x_n aux problèmes intermédiaires n'est une solution au problème de départ (à moins d'un coup de chance).
(3) Étant donné que la donnée du problème F_n est la fonction f de départ, nous avons $d_m = d_n = d$ pour tout m et n.

THOooMACHooLofCVu

Théorème 34.59.
Soit f continue sur un voisinage de α, racine simple. Alors il existe un voisinage de α de rayon σ tel que pour tout x_0 dans ce voisinage, la méthode converge vers α avec ordre de convergence $p = 2$.

Donc dès qu'on a continuité autour de la solution recherchée, il suffit de prendre x_0 assez proche pour que tout se passe bien. Cela se fait par localisation des racines, par exemples en traçant la fonction avec un bon niveau de zoom. Le fait est qu'on cherche disons 3 décimales à la main (travail sur ordinateur et graphique) et Newton donne les 20 décimales suivantes à la vitesse de la lumière.

34.6.7 Exemple de la racine carrée

Nous allons nous lancer dans un exemple : le cas de la racine carrée. Soit à calculer une approximation numérique de $\sqrt{2}$. Il s'agit d'une racine de la fonction $f(x) = x^2 + 2$. La fonction de la méthode de Newton associée est :

$$g(x) = x - \frac{f(x)}{f'(x)} = \frac{x^2 - 2}{2x}. \tag{34.121}$$

Cherchons un intervalle autour de $\sqrt{2}$ sur lequel nous avons convergence de la méthode de Newton. Cela s'obtient grace à la proposition 34.48 qui nous informe qu'il suffit de trouver un intervalle autour de $\sqrt{2}$ sur lequel $|g'(x)| \leqslant 1$.

Nous avons

$$g'(x) = \frac{x^2 - 2}{2x^2}, \tag{34.122}$$

et nous cherchons à résoudre $|g'(x)| \leqslant 1$. D'abord $g'(x) = 1$ n'a aucune solutions alors que $g'(\sqrt{2}) = 0$. Donc nous avons $g'(x) \leqslant 1$ pour tout $x \in \mathbb{R}^+$. Par contre l'équation $g'(x) = -1$ a des solutions : $x = \pm\sqrt{2/3}$.

Nous avons donc convergence de la méthode de Newton pour x_0 dans un intervalle de la forme

$$[\sqrt{2/3}, \sqrt{2} + \ldots] \tag{34.123}$$

où les les trois points représentent l'expression qu'il faut pour que ce soit symétrique autour de $\sqrt{2}$. La valeur précise n'a pas tellement d'importance parce, vu que nous sommes en train de chercher $\sqrt{2}$, il est peu probable que nous ayons déjà en main une bonne approximation de nombres du type $\sqrt{2/3}$.

Proposition 34.60.
La méthode de Newton pour la fonction $f(x) = x^2 - 2$ converge vers $\sqrt{2}$ pour toute valeur de départ dans $]0, +\infty[$.

Démonstration. La fonction $f(x) = x^2 - 2$ est convexe et $f'(\sqrt{2}) = 2\sqrt{2} > 0$. Donc la méthode converge vers $\sqrt{2}$ pour tout $x_0 \geqslant \sqrt{2}$ par la proposition 34.56.

Si par contre $x_0 \in]0, \sqrt{2}[$ nous avons

$$x_1 = \frac{x_0^2 + 2}{2x_0}. \tag{34.124}$$

En posant $h(x) = (x^2+2)/2x$ et en résolvant $h'(x) = 0$ nous trouvons $x = \sqrt{2}$. Et là, $h(\sqrt{2}) = \sqrt{2}$. Donc $h(x)$ est toujours plus grand que $\sqrt{2}$ pour tout $x \in]0, \sqrt{2}[$.

En d'autres termes, si $x_0 \in]0, \sqrt{2}[$ alors $x_1 \geqslant \sqrt{2}$ et nous retombons dans le premier cas. □

34.6.8 Si multiplicité

Supposons que α soit de multiplicité r (définition 34.50).

Cela se remarque en voyant que la méthode de Newton demande plutôt 20 itérations que 5. Le problème que cela pose est que chaque itération, les évaluations provoquent des erreurs. Donc moins d'itérations, c'est mieux.

Nous pouvons modifier la formule avec

$$x_{n+1} = x_n - r\frac{f(x_n)}{f'(x_n)}. \tag{34.125}$$

Il est possible de prouver que cette suite est à nouveau à convergence quadratique.

Ou alors on pose $F(x) = f^{(r-1)}(x)$ et α est une racine simple pour F. Donc faire Newton pour F est à nouveau quadratique, tout en donnant la même solution parce que $F(\alpha) = 0$ et $F'(\alpha) \neq 0$.

La seconde façon est bien parce que le théorème de localisation fonctionne 34.59

Et si r n'est pas connu ?

Il est toujours possible de faire $r = 2$ puis $r = 3$ et caetera jusqu'au moment où l'on remarque que le nombre d'itérations baisse un grand coup.

Mais ça demande beaucoup de calculs. Le mieux est de changer de méthode.

34.6.9 Et la dérivée ?

Un des problèmes de la méthode de Newton est que l'on doit pouvoir calculer la dérivée. Typiquement, il faut savoir f de façon analytique. Si cela n'est pas possible, nous pouvons changer de méthode et utiliser la méthode des sécantes décrite en 34.9.

34.6.10 Méthode de Newton : le cas général

LemXdObnV

Lemme 34.61.
Soient A et B deux matrices inversibles telles que la matrice $(A + \epsilon B)$ soit inversible pour tout ϵ assez petit. Alors il existe une matrice $X(\epsilon)$ telle que

$$(A + \epsilon B)^{-1} = (A^{-1} + \epsilon X) \tag{34.126}$$

et telle que $\lim_{\epsilon \to 0} X(\epsilon) = -A^{-1}BA^{-1}$.

Démonstration. Le candidat matrice X est relativement simple à trouver en écrivant

$$(A + \epsilon B)(A^{-1} + \epsilon X) = \mathbb{1} + \epsilon AX + \epsilon BA^{-1} + \epsilon^2 BX. \tag{34.127}$$

En imposant que cela soit $\mathbb{1}$, nous trouvons

$$X(\epsilon) = -(A + \epsilon B)^{-1}BA^{-1}. \tag{34.128}$$

La matrice $X(\epsilon)$ étant un inverse à droite de $(A + \epsilon B)$, son déterminant est non nul et X est inversible. Par conséquent elle est également inversible au sens usuel. Le calcul de la limite est direct :

$$\lim_{\epsilon \to 0} -(A + \epsilon B)^{-1}BA^{-1} = A^{-1}BA^{-1} \tag{34.129}$$

parce que l'inverse est une fonction continue sur $\mathrm{GL}(n, \mathbb{R})$. □

Remarque 34.62.
Un calcul naïf nous permet de trouver le même résultat de façon plus heuristique. En effet un développement usuel (dans $\mathbb{R}$) est

$$\frac{1}{a+\epsilon b}=\frac{1}{a}-\frac{\epsilon b}{a^2}+\dots \tag{34.130}$$

Si nous récrivons cela avec des matrices, nous écrivons (attention : passage heuristique !) :

$$(A+\epsilon B)^{-1}=A^{-1}-\epsilon A^{-1}BA^{-1}+\dots \tag{34.131}$$

Notons le choix de généraliser b/a^2 par $a^{-1}ba^{-1}$. Dans les réels les deux écritures sont équivalentes, mais pas dans les matrices.

Étudions si $A^{-1}-\epsilon A^{-1}BA^{-1}$ est bien un inverse à ϵ^2 près de $(A+\epsilon B)$:

$$(A+\epsilon B)(A^{-1}+\epsilon A^{-1}BA^{-1})=\mathbb{1}-\epsilon BA^{-1}+\epsilon BA^{-1}-\epsilon^2BA^{-1}BA^{-1}=\mathbb{1}-\epsilon^2BA^{-1}BA^{-1}. \tag{34.132}$$

Par conséquent, à des termes en ϵ^2 près la matrice $A^{-1}-\epsilon A^{-1}BA^{-1}$ est bien un inverse de $A+\epsilon B$.
ThoHGpGwXk

Théorème 34.63 (Méthode de Newton[751])**.**
Soit $f\colon \mathbb{R}^n\to\mathbb{R}^n$ une application de classe C^2 et un point $a\in\mathbb{R}^n$ tel que $f(a)=0$. Nous supposons que df_a est inversible.

Alors il existe un voisinage V de a tel que pour tout $x_0\in V$ la suite définie par récurrence

$$x_{n+1}=x_n-(df_a)^{-1}\big(f(x_n)\big) \tag{34.133}$$

converge vers a. De plus la vitesse est quadratique au sens où il existe $C>1$ tel que

$$\|x_n-a\|\leqslant C^{-1-2^n}. \tag{34.134}$$
EqtkiDXt

Démonstration. Étant donné que df_a est inversible et que df est continue, l'application df_x est continue[9] pour tout x dans un voisinage de a. Nous prenons $r>0$ tel que df_x est inversible pour tout $x\in B(a,r)$.

Nous considérons la fonction

$$\begin{aligned}F\colon B(a,r)&\to\mathbb{R}^n\\ x&\mapsto x-(df_x)^{-1}\big(f(x)\big).\end{aligned} \tag{34.135}$$

Cela est une application C^1. La clef est de montrer que l'application de F à un point $a+h$ rapproche de a pourvu que h soit assez petit. Nous avons la formule suivante :

$$F(a+h)-F(a)=h-\big(df_{a+h}\big)^{-1}\big(f(a+h)\big). \tag{34.136}$$
EqyDLQeE

Nous allons maintenant utiliser un développement de Taylor par rapport à h en suivant la formule (12.1333). Nous avons

$$f(a+h)=f(a)+df_a(h)+\|h\|^2\xi(h) \tag{34.137}$$

où $\xi\colon\mathbb{R}^n\to\mathbb{R}^n$ est une fonction qui tend vers une constante lorsque $h\to 0$. Nous avons aussi

$$df_{a+h}=df_a+\|h\|\tau(h) \tag{34.138}$$

où $\tau\colon\mathbb{R}^n\to\mathbb{M}(n,\mathbb{R})$ est une application qui tend vers une constante lorsque $h\to 0$. En ce qui concerne l'inverse nous utilisons le lemme[10] 34.61 :

$$\big(df_a+\|h\|\tau(h)\big)^{-1}=(df_a)^{-1}+\|h\|A(h) \tag{34.139}$$

9. Nous pouvons voir df comme l'application qui à x fait correspondre la matrice $df_x\in\mathbb{M}(n,\mathbb{R})$. Cette application étant continue et la non inversibilité d'une matrice étant donnée par l'annulation du déterminant, les matrices inversibles forment un ouvert dans l'ensemble des matrices.

10. Pour l'inversibilité de $\|h\|\tau(h)$, notons que df_a est inversible et que par hypothèse la somme $df_a+\|h\|\tau(h)$ est inversible.

où A est une autre matrice fonction de h qui tend vers une constante lorsque h tend vers zéro. En substituant le tout dans (34.136) nous trouvons

$$F(a+h)-F(a)=\|h\|^2(df_a)^{-1}\xi(h)+\|h\|\big(A(h)\circ df_a\big)(h)+\|h\|^3A(h)\xi(h). \tag{34.140}$$

En ce qui concerne la norme nous utilisons le fait que si T est un opérateur, $\|Tx\|\leqslant\|T\|\|x\|$. Nous trouvons

$$\begin{aligned}\|F(a+h)-F(a)\|&\leqslant\|h\|^2\|(df_a)^{-1}\|\|\xi(h)\|+\|h\|^2\|A(h)\circ df_a\|+\|h\|^3\|A(h)\|\|\xi(h)\| &\text{(34.141a)}\\ &=\|h\|^2\alpha(h) &\text{(34.141b)}\end{aligned}$$

pour une certaine fonction $\alpha\colon\mathbb{R}^n\to\mathbb{R}$ qui tend vers une constante lorsque $h\to0$.

En posant $C=\lim_{h\to0}\alpha(h)$ nous avons la majoration

$$\|F(x)-a\|\leqslant C\|x-a\|^2. \tag{34.142}$$ EqSYiuYF

Nous pouvons également supposer que $C>1$. Afin de prouver la vitesse de convergence (34.134), nous allons encore redéfinir r en demandant $r<1/C^2$. De cette manière nous avons

$$\|x_0-a\|\leqslant\frac{1}{C^2} \tag{34.143}$$

et la récurrence sur n est :

$$\|x_{n+1}-a\|=\|F(x_n)-a\|\leqslant C\|x_n-a\|^2\leqslant C\big(C^{-1-2^n}\big)^2=C^{-1-2^{n+1}}. \tag{34.144}$$

Note : ce dernier calcul est le lemme 34.45 appliqué à $r=(1/C^2)$. □

Remarque 34.64.
La valeur de la constante C a été fixée par l'équation (34.142). Certes nous pouvons toujours choisir C plus grand afin d'augmenter la vitesse de convergence, mais le point de départ x_0 devant être dans une boule de taille $1/C^2$ autour de a, demander C plus grand revient à demander un point de départ plus précis.

34.7 Estimation de l'ordre de convergence

Définition 34.65 ([752]).
*Nous disons que la suite (x_n) de limite x est **convergente d'ordre** q pour $q>1$ si il existe $\mu>0$ tel que*

$$\lim_{k\to\infty}\frac{|x_{k+1}-x|}{|x_k-x|^q}=\mu. \tag{34.145}$$

En particulier :

- *la convergence d'ordre 2 est dite quadratique,*
- *la convergence d'ordre 3 est dite cubique,*
- *la convergence d'ordre 4 est dite quartique.*

Comment estimer numériquement l'ordre p de convergence de la méthode ? Soit une suite (x_n) convergente vers α. Considérons les 4 termes x_{n-3}, x_{n-2}, x_{n-1}, x_n. Alors nous pouvons écrire l'approximation

$$\frac{|x_n-x_{n-1}|}{|x_{n-1}-x_{n-2}|}\simeq\left(\frac{|x_{n-1}-x_{n-2}|}{|x_{n-1}-x_{n-3}|}\right)^p. \tag{34.146}$$

Cette approximation ne serait pas trop mauvaise tant que n est assez grand pour que la convergence soit bien engagée. Passons au logarithme :

$$\ln\frac{|x_n-x_{n-1}|}{|x_{n-1}-x_{n-2}|}\simeq p\ln\left(\frac{|x_{n-1}-x_{n-2}|}{|x_{n-1}-x_{n-3}|}\right). \tag{34.147}$$

et donc

$$p \simeq \frac{\ln\left(\frac{|x_n-x_{n-1}|}{|x_{n-1}-x_{n-2}|}\right)}{\ln\left(\frac{|x_{n-1}-x_{n-2}|}{|x_{n-1}-x_{n-3}|}\right)}. \tag{34.148}$$

Avec cette approximation, en réalité nous calculons une suite (p_i) qui sont les approximations de p à partir des termes i à $i+3$ de la suite (x_n). Il s'agit d'une suite d'estimations de p.

(1) Dans le cas de la bisection, nous obtenons toujours $p_i = 1$.

(2) Dans le cas de la méthode de Newton (34.6) nous avons $p = 2$. Mais les premières valeurs de p_i peuvent être aussi bien 0 que 7. Après quelques itérations pourtant les p_i se regroupent autour de 2.

En tout cas, le plus important est de savoir si $p > 1$ ou non. Rappel : nous voulons la superlinéarité parce que nous voulons utiliser le test d'arrêt de la différence entre deux termes, voir 34.47.

34.8 Autres méthodes

34.8.1 Méthode de Schröder

La formule est

$$x_{n+1} = x_n - \frac{f(x_n)f'(x_n)}{f'(x_n)^2 - f(x_n)f''(x_n)} \tag{34.149}$$

Cette méthode est d'ordre 2 pour toute racine et toute valeur de multiplicité. Le problème de cette méthode est qu'elle demande 3 évaluations de f. Son efficacité :

$$E = \sqrt[3]{2} \simeq 1.25 \tag{34.150}$$

Cela est donc moins efficace que Newton.

34.8.2 Halley

Il a $p = 3$ lorsque α est racine simple. Mais encore $p = 1$ pour les racines multiples. Plus efficace que Newton pour les racines simples, mais même problème pour les racines multiples.

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)} \tag{34.151}$$

34.9 Méthode des sécantes variables

SECooIUEUooVcHAoc

Si nous n'avons pas de formule analytique pour f, mais seulement la possibilité de calculer $f(x)$ pour tout x. Newton ne fonctionne pas, mais la bisection fonctionne.

Nous pouvons approximer

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \tag{34.152}$$

En substituant dans la formule de Newton, nous obtenons

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}. \tag{34.153}$$

Il s'agit de prendre la droite qui passe par $(x_{n-1}, f(x_{n-1}))$ et par $(x_n, f(x_n))$ et de prendre l'intersection de cette droite avec l'axe $y = 0$. Cela donne le x_{n+1}.

Pour cette méthode, il ne faut pas seulement x_0 mais également x_1.

L'ordre de convergence est le nombre d'or

$$p = \frac{1-\sqrt{5}}{2} \simeq 1.618. \tag{34.154}$$

EQooQEFCooUsGViR

Cela est donc superlinéaire.

Le nombre d'évaluations est $s = 1$ (il y a deux apparitions de f dans la formule, mais l'une des deux est récupérée dans l'itération suivante). Donc l'efficacité est

$$E = p. \tag{34.155}$$

Donc bien efficace.

Proposition 34.66.
Si α est racine simple, il existe un voisinage de α tel que pour tout choix de x_0, x_1 dans ce voisinage, la méthode converge.

Psychologiquement, on est tenté de prendre x_0 et x_1 de part et d'autre de α (pensant à la bisection), mais en réalité ce n'est pas obligatoire du tout et n'a aucune influence. Il faut seulement les prendre très proches de α.

Remarque 34.67.
La méthode de la sécante est souvent écrite sous la forme

$$x_{n+1} = \frac{x_{n-1}f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}. \tag{34.156}$$

EQooYVKLooKTEjwv

C'est évidemment algébriquement équivalent.

Les formules (34.154) et (34.156) ont toutes deux des erreurs de cancellation. Laquelle est la plus grave ?

Dans la première, si la fraction est mal calculée, elle ne fait que modifier x_n. C'est-à-dire qu'on peut espérer qu'à la prochaine itération, ça aille mieux. En tout cas, dans ce cas si la fraction est mal calculée, ça ne détruit pas tout.

Dans la seconde, c'est la valeur elle-même qui risque d'être mal calculée. Et si la fraction est mal calculée, alors on casse complètement l'éventuelle bonne approximation que nous avions déjà.

34.9.1 Aitken

La méthode du Δ^2 de Aitken est une méthode d'accélération de la convergence.

Soit (x_n) une suite qui converge. Nous voudrions une nouvelle suite (y_n) telle que

$$\lim_{n\to\infty} \frac{y_n - \alpha}{x_n - \alpha} = 0. \tag{34.157}$$

C'est la définition d'une convergence accélérée.

La façon de faire est :

$$y_n = \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n} = x_n - \frac{(x_{n+1} - x_n)}{x_{n+2} - 2x_{n+1} + x_n}. \tag{34.158}$$

La première expression a deux cancellations (la seconde une seule) et de plus la première est y_n elle-même alors que la seconde est une correction.

Donc la seconde expression est numériquement meilleure.

L'opérateur Δ appliqué à une suite est :

$$(\Delta x)_n = x_{n+1} - x_n \tag{34.159}$$

Donc

$$(\Delta^2 x)_n = (\Delta x)_{n+1} - (\Delta x)_n = x_{n+2} - x_{n+1} - x_{n+1} + x_n = x_{n+2} - 2x_{n+1} + x_n. \tag{34.160}$$

L'accélération a alors la formule

$$y_n = \frac{(\Delta x)_n^2}{(\Delta^2 x)_n}. \tag{34.161}$$

Le problème est que ça accélère tellement que l'on arrive vite à des erreurs de cancellations, et donc à une précision en pics oscillants.

34.10 Équations algébrique

C'est une équation du type $P(x) = 0$ où P est un polynôme. Soit un polynôme de degré n. Nous en savons des choses.

(1) L'équation a exactement n solutions dans $\mathbb{C}$ en comptant les multiplicités.

(2) Les racines complexes arrivent par paire complexes conjuguée. Elles sont donc toujours en nombre pair.

Si donc nous avons $n = 3$, nous ne pouvons pas avoir 2 racine réelles. Il y en a donc 1 ou 3 réelles. Pas zéro ni deux.

Quelques méthodes : Müller, matrice compagnon, Laguerre.

34.10.1 Résoudre un système linéaire

Pour résoudre un système linéaire d'équations, nous échelonnons la matrice du système. Soit à résoudre le système $Ax = b$ où

$$A = \begin{pmatrix} 2 & 4 & -6 \\ 1 & 5 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \text{et} \quad b = \begin{pmatrix} -4 \\ 10 \\ 5 \end{pmatrix}. \tag{34.162}$$

En termes de problèmes, on écrit $F\big(x, (A, b)\big) = Ax - b$. La donnée de ce problème est le couple (A, b).

En ce qui concerne l'algorithme, on pose comme premier problème

$$F_1\big(x_1, (A_1, b_1)\big) = A_1x_1 - b_1 = 0 \tag{34.163}$$

avec $A_1 = A$ et $b_1 = b$.

Ensuite, on commence à échelonner et le second problème est

$$F_2\big(x_2, (A_2, b_2)\big) = A_2x_2 - b_2 = 0 \tag{34.164}$$

avec

$$A = \begin{pmatrix} 2 & 4 & -6 \\ 0 & 3 & 6 \\ 0 & 1 & 5 \end{pmatrix}, \quad \text{et} \quad b = \begin{pmatrix} -4 \\ 12 \\ 13 \end{pmatrix}. \tag{34.165}$$

Le troisième problème sera

$$F_3\big(x_3, (A_3, b_3)\big) = A_3x_3 - b_3 = 0 \tag{34.166}$$

avec

$$A = \begin{pmatrix} 2 & 4 & -6 \\ 0 & 3 & 6 \\ 0 & 0 & 3 \end{pmatrix}, \quad \text{et} \quad b = \begin{pmatrix} -4 \\ 12 \\ 3 \end{pmatrix}. \tag{34.167}$$

Ce problème est facile à résoudre « à la main ». Nous nous arrêtons donc ici avec l'algorithme, et nous trouvons le x_3 qui résous le problème F_3.

34.10.2 Caractéristiques

L'algorithme de résolution de systèmes linéaires d'équations a les propriétés suivantes, à mettre en contraste avec celles de Newton :

(1) Pour résoudre le problème numéro n, il n'a pas fallu résoudre le problème numéro $n - 1$.

(2) Toutes les solutions x_n des problèmes intermédiaires sont solutions du problème de départ. Nous avons $F_n(x, d_n) = 0$ pour tout n (ici, $d_n = (A_n, b_n)$).

(3) D'un problème à l'autre, les données changent énormément : la matrice échelonnée peut être très différente de la matrice de départ.

34.10.3 Définitions

Nous allons maintenant formaliser en donnant quelques définitions pour nommer les propriétés que nous avons vues. D'abord, un algorithme est une suite de problèmes. Un **algorithme** pour résoudre un problème $F(x,d)=0$ est une suite de problèmes $\{F_n(x_n,d_n)=0\}_{n\in\mathbb{N}}$.

Définition 34.68.
Un tel algorithme est dit ***fortement consistant*** *si pour toutes données admissibles d_n, on a*

$$F_n(x,d_n)=0 \quad \forall\, n, \tag{34.168}$$

où x est la solution de $F(x,d)=0$.

L'algorithme des matrices est fortement consistant, mais pas l'algorithme de Newton.

Définition 34.69.
Un algorithme est ***consistant*** *si $\lim_{n\to\infty} F_n(x,d_n)=0$.*

Dans le cas de l'algorithme de Newton, c'est plutôt une telle consistance qu'on attend.

L'algorithme est dit **stable** si pour tout n le problème correspondant est stable. Dans ce cas, on note K^{num} le **conditionnement relatif asymptotique** défini par

$$K^{\text{num}}=\limsup_n K_n \tag{34.169}$$

où K_n est le conditionnement relatif du problème $F_n(x_n,d_n)=0$.

DefAlgoConverge

Définition 34.70.
Un algorithme est dit ***convergent*** *(en d) si pour tout $\epsilon>0$, il existe $N=N(\epsilon)$ et $\delta=\delta(N,\epsilon)$ tels que pour $n\geqslant 0$ et $|d-d_n|<\delta$, on ait $|x(d)-x_n(d_n)|<\epsilon$.*

RemConvAlgoNewton

Remarque 34.71.
Dans le cas de l'algorithme de Newton, nous avons vu que la donnée d_n du problème F_n était en fait la même que la donnée initiale d, donc nous avons $d_n=d$, et par conséquent nous avons toujours $|d-d_n|<\delta$. Dans ce cas, la définition de la convergence revient à demander que la suite numérique des x_n converge vers la solution x.

Remarque 34.72.
Dans le cas des matrices par contre, les données sont très différentes les unes des autres, nous avons donc en général que $|d-d_n|>\delta$. Mais en revanche nous savons que tous les problèmes intermédiaires F_n acceptent une solution unique[11] $x_n(d_n)=x(d)$. Par conséquent, $|x_n(d_n)-x(d)|$ est toujours plus petit que ϵ. L'algorithme des matrices est donc toujours un algorithme convergent.

34.11 Équations non linéaires

Certaines équations non linéaires sont résolubles explicitement, par exemples les polynômes de degré jusqu'à 4 ou des choses comme

$$\sin^2(x)+3\sin(x)+5=0. \tag{34.170}$$

Mais ces exemples sont très rares.

Nous allons étudier des équations du type $f(x)=0$, dans $\mathbb{R}$.

(1) Un problème écrit sous la forme $x=g(x)$ peut utiliser des théorèmes de points fixes.

(2) Un problème sous la forme $f(x)=0$ peut utiliser des méthodes de bisection, Newton ou autres.

11. Nous n'envisageons que le cas où le déterminant est non nul.

Il y a évidemment beaucoup de façons de transformer un problème pour passer d'une forme à l'autre.

Exemple 34.73.
Soit $f(x) = x^2 - a = 0$ avec $a > 0$. Nous pouvons l'écrire

$$x^2 + x - a = x \tag{34.171}$$

qui donne une forme $g(x) = x$ pour $g(x) = x^2 + x - a$.

Ou encore $x = \frac{a}{x}$ et donc $g(x) = a/x$ (si par ailleurs on sait que $x \neq 0$). Notons que $x \neq 0$ n'est pas une hypothèse très forte parce qu'on la vérifie directement sur a. △

Exemple 34.74.
Soit l'équation à résoudre

$$f(x) = x^2 - 2 - \ln(x) = 0 \tag{34.172}$$

Les solutions de cette équation peuvent être vues comme les intersections avec l'axe X du graphe $y = x^2 - 2 - \ln(x)$. Tracer peut donc aider. Par ailleurs, il faut noter que

$$\lim_{x \to \pm\infty} f(x) = \infty, \tag{34.173}$$

donc les solutions sont certainement contenues dans un compact de $\mathbb{R}$.

À part tracer nous pouvons écrire

$$x^2 - 2 = \ln(x). \tag{34.174}$$

Et là, ce sont deux fonctions dont nous pouvons tracer le graphe pour trouver graphiquement les points d'intersection. Une étude de fonction montre vite qu'il y a exactement deux solutions, qu'elles sont strictement positives. Pour trouver des bornes, il faut calculer par exemple pour $x = 2$ les valeurs de $\ln(x)$ et $x^2 - 2$ pour voir si le graphe de $x^2 - 2$ est déjà plus haut. △

La majorité des méthodes numériques de résolution d'équations du type $f(x) = 0$ ou $x = g(x)$ seront sous la forme de suites. Avec questions à la clef :

(1) Quel point de départ choisir ?
(2) Convergence ?
(3) Est-ce que la limite est bien une solution ?
(4) Vu que la limite est unique, comment faire si l'équation a plusieurs solutions ? (souvent c'est le choix du point initial qui va jouer sur ce point)

34.75.
Si la fonction est très plate, il est possible d'avoir

$$|f(\tilde{\alpha})| \leqslant \epsilon \tag{34.175}$$

sans que $\tilde{\alpha}$ ne soit une bonne approximation.

Lorsqu'on fait tourner une méthode itérative résolvant $f(x) = 0$, il n'est pas suffisant de s'arrêter lorsque

$$f(x_n) \leqslant \epsilon_1. \tag{34.176}$$

Il faut aussi s'assurer que, si $\bar{x}$ est la solution exacte, $|x_n - \bar{x}| \leqslant \epsilon_2$. Ici ϵ_1 et ϵ_2 sont deux « précisions » que nous nous fixons au départ.

Évidemment, vérifier la condition $|x_n - \bar{x}| \leqslant \epsilon_2$, il faudrait savoir $\bar{x}$. Et savoir $\bar{x}$ c'est justement le problème. Nous sommes donc amenés à faire des estimations de $|x_n - \bar{x}|$.

34.76.
Lorsque nous effectuons une méthode itérative, il faut donc contrôler deux grandeurs :

$$|\bar{x} - x_n| \leqslant \epsilon_1 \tag{34.177a}$$

$$|x_{n+1} - x_n| \leqslant \epsilon_2. \tag{34.177b}$$

Proposition 34.77.
Soit p l'ordre de convergence de la suite (x_n) vers $\bar{x}$. Si $p > 1$ et $|x_{n+1} - x_n| \leqslant \epsilon_2$ alors $|\bar{x} - x_n| \leqslant \epsilon_2$.

34.11.1 Méthode de bisection

Il y a ce théorème des valeurs intermédiaires.

Théorème 34.78.
Soit f continue sur $[a,b]$ telle que $f(a)f(b)<0$. Alors il existe au moins une solution à l'équation $f(x)=0$ sur l'intervalle $]a,b[$.

Pour démarrer une bisection, il est toujours bon de prendre l'intervalle $[a,b]$ de façon à ne contenir qu'une seule solution.

Soit donc un premier intervalle $[a_0,b_0]$ tel que $f(a_0)f(b_0)<0$ et ne contenant qu'une seule solution. À chaque itération nous considérons la moitié de l'intervalle précédent, mais la moitié contenant la solution.

Le test d'arrêt de la méthode de bisection se base uniquement sur la taille de l'intervalle qui reste. En effet si nous avons

$$|b_n-a_n|\leqslant\epsilon \tag{34.178}$$

nous avons certainement

$$|\bar{x}-x_n|\leqslant\frac{\epsilon}{2} \tag{34.179}$$

où x_n est le point du milieu de $[a_n,b_n]$.

34.79.
La fonction f n'intervient dans la méthode que via son signe, pas via ses valeurs exactes.

34.80.
Notons que le théorème des valeurs intermédiaires n'est pas très puissant pour choisir l'intervalle de départ ; penser à la fonction

$$f(x)=x^2-5 \tag{34.180}$$

sur l'intervalle $[-10,10]$. Il y a bien deux solutions dans l'intervalle, mais elles sont invisibles du théorème des valeurs intermédiaires. La fonction $x\mapsto x^2$ a sa solution en $x=0$, mais elle aussi n'est pas visible.

34.81.
Certes la méthode de bisection assure la convergence vers une solution, mais elle n'assure pas la convergence monotone. Il peut arriver que $|\bar{x}-x_n|<|\bar{x}-x_{n+1}|$. C'est le cas lorsque la solution est très proche du milieu de l'intervalle choisi. Le x_0 est alors proche de $\bar{x}$ alors que x_1 sera à une distance de $\bar{x}$ d'environ un quart de l'intervalle de départ.

Supposons déjà avoir trouvé un intervalle $[a,b]$ dans lequel se trouve une unique solution à $f(x)=0$. Voici un algorithme possible.

```
1  from __future__ import division
2
3  def bissection(f, a, b, toll, mmax):
4      """
5      f : une fonction
6      a,b : les limites de l'intervalle
7      toll : la tolérance. C'est l'amplitude de l'intervalle à ↩
           partir duquel nous nous arrêtons.
8      nmax : le nombre maximum d'itérations.
9
10     Nous supposons que \( b>a\).
11
12     Retourne un tuple (x,n) où 'x' est la solution approchée et 'n↩
           ' est le nombre d'itérations effectuées
```

```
    """
    n = -1
    amp = toll + 1 # Pour s'assurer que l'on entre dans le cycle
    while amp > toll and n < nmax :
        n = n + 1
        amp=abs(b - a)
        x = a + amp / 2
        if f(a) * f(x) < 0:
            b = x
        elif f(a)*f(x) > 0:
            a = x
        else :       # Problème ZERO
            amp = 0
    return (x, n)
```

tex/frido/codeSnip_1.py

Plusieurs remarques :

(1) Le fait de retourner le nombre d'itérations effectuées permet à l'utilisateur de savoir la précision et si le nombre maximum d'itérations est dépassé. Si ce `n` retourné est égal à `nmax`, l'utilisateur sait que le `x` retourné n'est pas fiable.

(2) La ligne `from __future__ import division` fait en sorte que l'opération `/` est bien la division usuelle. Sinon, le défaut en python 2 est que `/` soit la division *entière*, c'est-à-dire que $1/2 = 0$ en python 2. En python 3, le symbole `/` désigne bien la division usuelle, mais Sage utilise Python 2 [12].

(3) Même si l'intervalle $[a, b]$ contient plus d'une solution, la méthode fonctionne et donne une solution. Il est simplement éventuellement très compliqué de savoir laquelle.

(4) Nous faisons `amp=toll+1` parce que nous voulons absolument lancer le cycle au moins une fois. Sinon, le `x` à retourner ne serait pas défini au moment de sortir du cycle (si le cycle n'est pas exécuté).

(5) Calculer le point milieu d'un intervalle $[a, b]$ est par la formule $(a + b)/2$ sauf que cette opération est numériquement dangereuse parce qu'à cause de l'arithmétique en précision finie, il est possible que cela tombe *exactement* sur a ou b. D'où le fait de calculer le point milieu par

$$x = a + \frac{amp}{2}. \tag{34.181}$$

(6) Dans le cas `Problème ZERO` nous déduisons $f(x) = 0$. Attention que c'est pas que $f(x) = 0$ mais simplement que en mettant x dans f, la *machine* retourne son zéro.
Il peut cependant avoir une fonction telle que $f(1) = 10^{-50}$ et $f(2) = 0$. L'algorithme de bisection risque de s'arrêter si $x_n = 1$. Parce que la machine risque de calculer $f(x_n) = 0$.
Quoi qu'il en soit, nous y mettons `amp=0` pour être sûr de sortir de la boucle dès la prochaine vérification.

(7) Il y a moyen de sauver les valeurs de $f(a)$ et $f(x)$ pour ne pas les recalculer, et en particulier au moment de faire `b=x` nous pouvons poser `fa=fx`.

Si τ est la précision de la solution voulue, nous pouvons fixer à priori le nombre d'itérations à faire grâce à la formule

$$n \geqslant \left\lceil \log_2 \left(\frac{b-a}{\tau} \right) \right\rceil. \tag{34.182}$$

Il y a un "$\geqslant$" et non une égalité parce qu'en arithmétique numérique, le nombre obtenu à droite pourrait ne pas être le bon à 1 près.

Ici pour $\nu \in \mathbb{R}$ le nombre $\lceil \nu \rceil$ est le plus petit entier à être plus grand ou égal à ν.

12. Cela était vrai au moment d'écrire l'exemple. Aujourd'hui, Sage utilise la version 3. Il n'y a aucune raison d'importer `future`.

34.82.
Notons l'importance de la continuité de f. Par exemple que ferait la bisection sur la fonction $f(x) = 1/x$ pour l'intervalle $[-3, 1]$?

Il y a changement de signe sans avoir de racine.

Vu que 2^{10} est déjà 1024. Donc si on veut de la précision de l'ordre de 1/1000, dix itérations suffisent. Si donc nous avons besoin de 200 itérations pour atteindre la précision voulue, c'est l'occasion de trouver un intervalle plus petit. Par exemple en traçant la fonction, en faisant un zoom et en trouvant des valeurs de a et b qui sont déjà proches.

34.83.
Dans le monde réel, il arrive souvent d'utiliser une méthode de bisection pour se donner un point de départ pour une autre méthode.

34.12 Efficacité

Définition 34.84.
*L'**efficacité** est le nombre*

$$E = \sqrt[s]{p} \tag{34.183}$$

où p est l'ordre de convergence de la méthode et s est le nombre de fois qu'il faut calculer une valeur de la fonction à chaque itération (nous ne comptons pas l'initialisation).

Que le nombre de d'évaluations de f intervienne est logique parce que chaque évaluation provoque une erreur possible.

Exemple 34.85 (Bisection)**.**
Pour la méthode de bisection, nous avons $s = 1$ parce que chercher x_{n+1}, il faut seulement calculer $f(x_n)$. △

Exemple 34.86 (Newton)**.**
Pour l'algorithme de Newton nous avons $p = 2$ et il y a deux évaluations à chaque itération (une fois f et une fois f'), donc $s = 2$ et $E = \sqrt{2}$. △

34.13 Exemples sous forme d'exercices

Exemple 34.87.
Nous supposons une machine acceptant 5 chiffres significatifs. Elle retient les nombres sous la forme $\pm 0.x \cdots \times 10^{\cdots}$.

$x = 1.403$, $y = 0.4112 \times 10^{-3}$, $z = -0.4111 \times 10^{-3}$.

Soient $a = (x \oplus y) \oplus z$ et $b = (y \oplus z) \oplus x$.

(1) D'abord la calculer à la main.

(2) Quel est le calcul préférable ?

(3) Donner l'erreur relative avec 3 chiffres significatifs.

Dans le cas du calcul à la main, il faut en faire un seul parce que, algébriquement, $a = b$.

Nous avons $x = 1.403$, $y = 0.0004112$ et $z = -0.0004111$. Et la somme donne :

$$a = b = 1.4030001 = 0.14030001 \times 10^1. \tag{34.184}$$

Faisons d'abord la normalisation de x, c'est-à-dire $\mathrm{fl}(x)$.

$$\mathrm{fl}(x) = 0.1403 \times 10^1. \tag{34.185}$$

et y est déjà normalisé :

$$\mathrm{fl}(y) = 0.4112 \times 10^{-3}. \tag{34.186}$$

Il n'y a pas d'erreurs d'assignation pour ces deux nombres.

Pour faire la somme, il faudra déjà un peu casser les nombres pour les écrire de façon à pouvoir les sommer. En effet, il faut écrire les deux nombres avec le même exposant de 10 (le plus grand), pour pouvoir les mettre en colonne :

$$0.1403 \times 10^1 \to 0.1403 \times 10^1 \tag{34.187a}$$

$$0.4112 \times 10^{-3} \to 0.00004 \times 10^1. \tag{34.187b}$$

La somme donne 0.14034×10^1. Et ça, c'est à nouveau arrondi. Le premier chiffre supprimé est un 4, donc

$$x \oplus y = 0.1403 \times 10^1. \tag{34.188}$$

Et là on remarque que nous avons la même chose que x. C'est un classique du calcul numérique.

Nous avons aussi

$$\text{fl}(z) = 0.4111 \times 10^{-3}. \tag{34.189}$$

Et pour faire la somme de cela avec $x \oplus y$ nous devons le remettre sous la forme d'un 10^1 :

$$\text{fl}(z) \to -0.00004 \times 10^1 \tag{34.190}$$

(erreur de conversion), et en sommant on trouve

$$(x \oplus y) \oplus z = 0.140216 \times 10^1, \tag{34.191}$$

qui est encore arrondi. Le premier chiffre supprimé est un 6, donc

$$\text{fl}(a) = 0.1403 \times 10^1, \tag{34.192}$$

Nous avons donc $x \oplus y = x$. Le nom de l'erreur qui consiste à avoir $x \oplus y = x$ est **relation anormale**.

Calculons b.

Les nombres y et z ont même ordre de grandeur, donc pas d'erreur au moment de les mettre sous forme sommable.

$$\text{fl}(y) + \text{fl}(z) = 0.00010 \times 10^{-3}. \tag{34.193}$$

Cela est renormalisé et arrondi : $\text{fl}(x) \oplus \text{fl}(y) = 0.1000 \times 10^{-6}$.

Notons que nous avons ici commis potentiellement une erreur de cancellation parce que entre y et z, il y a 3 chiffres sur 4 qui sont identiques. Seul le chiffre 1 est significatif en réalité.

Il faut maintenant ajouter x à cela. D'abord

$$\text{fl}(x) = 0.1403 \times 10^1. \tag{34.194}$$

Pour cette somme, il faudra remettre notre 0.1000×10^{-6} avec une puissance 10^1. Et là, nous obtenons zéro parce que vraiment ce nombre est trop petit pour être écrit avec 10^1. Résultat des courses :

$$\text{fl}(b) = 0.1403 \times 10^1. \tag{34.195}$$

Dans le premier calcul nous avons deux "relations anormales" et dans le second nous en avons une plus une cancellation.

Nous préférons avoir deux relations anormales, parce que l'erreur de cancellation est plus grave : elle consiste à une perte de chiffre significatifs. Le fait est que faisant la différence à l'ordinateur nous avons obtenu 0.1 qui est certes exact, mais qui est un coup de bol : la différence aurait aussi bien pu être 0.19 avec d'autres nombres, machinement égaux.

Note : avec les données ici, il n'y a en fait pas d'erreur de cancellation. Mais il y a une erreur potentielle de cancellation, potentiellement grave.

En ce qui concerne l'erreur relative. Dans la formule

$$\epsilon_r = \frac{|a - a^*|}{|a|}, \tag{34.196}$$

la différence ne peut pas être calculée à la calculatrice justement parce qu'elle est très potentiellement sujette à erreur de cancellation.

$$\epsilon_r = \frac{|0.1030001 \times 10^1 - 0.1403 \times 10^1|}{0.14030001 \times 10^1} = \frac{0.1 \times 10^{-6}}{0.14030001 \times 10^1} \simeq 0.712758 \times 10^{-7}. \tag{34.197}$$

En passant à 3 chiffres significatifs, 0.713×10^{-7} (le premier chiffre supprimé est un 7).

△

Exemple 34.88.
Soient $x = 0.1 \times 10^{21}$ et $y = 0.5 \times 10^{20}$ et les expressions

(1) $z_1 = \frac{x-y}{y} + \frac{x+y}{x}$

(2) $z_2 = \frac{x^2+y^2}{xy}$.

Ces deux expressions sont algébriquement équivalentes.

(1) Calculer les valeurs.

(2) On suppose une machine en précision simple. Laquelle des deux expressions est préférable ?

Pour z_1, en arithmétique exacte :

$$z_1 = \frac{0.5 \times 10^{20}}{0.5 \times 10^{20}} + \frac{1.5 \times 10^{20}}{1 \times 10^{20}} = 0.25 \times 10^1. \tag{34.198}$$

Le calcul exact de z_2 donne la même chose.

(1) **Calcul de z_1**
Les deux valeurs sont mémorisables et la différence $x - y$ se fait sans erreurs de cancellation. Idem pour la somme $x + y$. Idem pour les divisions.

(2) **Calcul de z_2**
Pour faire x^2, c'est pas possible parce que c'est de l'ordre de 10^{40} alors que nous sommes en précision simple. Idem pour le produit xy.

Morale : z_2 donne un `overflow` alors que z_1 fonctionne de façon exacte.

Remarque 34.89.
En réalité le z_1 n'est pas tout à fait calculable de façon exacte sur la machine parce qu'elle doit d'abord convertir en binaire, ce qui n'est pas toujours possible. Mais sur notre machine qui fonctionne en base 10, il n'y a pas de problèmes.

△

¡¡ Avertissement/question au lecteur !! 34.90
Dans l'exemple suivant, je ne sais plus du tout ce qu'est la fonction g. Par conséquent, je ne sais pas quel est le signe à indiquer dans $g(x_2)$ du dernier tableau.

Si vous avez tout suivi jusqu'ici, écrivez-moi pour éclaircir ces points.

Exemple 34.91.
Soit la fonction

$$f(x) = 2x^2 - 4x + 2 - e^{-x}. \tag{34.199}$$

(1) Identifier la plus grande des solutions réelles de $f(x) = 0$.

(2) Effectuer une bisection pour la savoir.

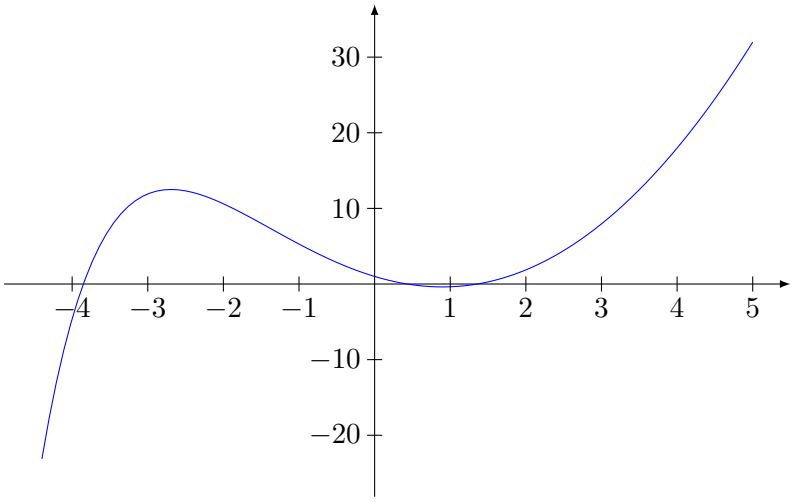
(3) Sachant que

$$\alpha \simeq 1.358500220734946, \tag{34.200}$$

quelle est l'erreur relative ?

Note : si c'est pour chercher à la main des approximations pour démarrer, il est évidemment préférable de dessiner $f_1(x) = 2x^2 - 4x + 2$ et $f_2(x) = -e^{-x}$ séparément.

Quoi qu'il en soit, voici un graphique :



Nous voyons trois racines : $\alpha_1 \in [0, 0.5,]$, $\alpha_2 \in [1, 1.5]$ et $\alpha_0 \in [-4, -3.5]$.

La plus grande solution est α_2. Nous pouvons déjà remplir le tableau des précisions :

n	a_n	b_n	x_n	$g(x_n)$	$\lvert b_n - a_n \rvert$
...	...	...	...	...	0.5
...	...	...	...	...	0.25
...	...	...	...	...	0.125

Et nous calculons les valeurs de f aux points d'extrémité de l'intervalle. Note que seul le signe nous importe :

$$f(1) \simeq -0.368 \tag{34.201a}$$

$$f(1.5) \simeq 0.278. \tag{34.201b}$$

Voici donc le tableau avec le signe de f indiqué :

n	a_n	b_n	x_n	$g(x_n)$	$\lvert b_n - a_n \rvert$
0	1(+)	1.5(−)	...	-0.162×10^0	0.5
1	...	...	...	...	0.25
2	...	...	...	...	0.125

Puis :

n	a_n	b_n	x_n	$g(x_n)$	$\lvert b_n - a_n \rvert$
0	1(+)	1.5(−)	1.25(−)	-0.162×10^0	0.5
1	...	...	..	...	0.25
2	...	...	...	...	0.125

Et enfin :

n	a_n	b_n	x_n	$g(x_n)$	$\lvert b_n - a_n \rvert$
0	1(+)	1.5(−)	1.25(−)	-0.162×10^0	0.5
1	1.25(−)	1.5(+)	1.375(+)	$+0.284 \times 10^{-1}$	0.25
2	1.25(−)	1.375(+)	1.3125(−)	0.738×10^{-1}	0.125

Note que les $f(x_n)$ restent toujours du même ordre de grandeur. Si un moment on voit un 1.6×10^7, c'est qu'une erreur a été commise.

En ce qui concerne le calcul de l'erreur relative, la première chose à faire est de vérifier que le α proposé est dans l'intervalle qui nous reste. Sinon c'est qu'une erreur a été commise.

De plus notre approximation est $x_n = 1.3125$, dont déjà deux chiffres sont corrects. En deux itérations de bisection en partant de 0.5, nous ne pouvons pas nous attendre à mieux.

△

34.14 Approximations de fonctions

(1) D'habitude on n'approxime pas une fonction sur tout son domaine, mais seulement sur une partie.

(2) Il y a le problème du choix de la classe des fonctions qui vont approximer. Nous allons travailler avec des polynômes.

(3) Il nous faut un critère disant si une approximation est bonne ou non.

34.14.1 Critère d'interpolation

À partir de $n+1$ abscisses points distinctes x_i, nous calculons $y_i = f(x_i)$. Il y a ce théorème qui dit qu'il existe un unique polynôme de degré (au plus) $n+1$ passant par les points (x_i, y_i).

Proposition-Définition 34.92 (Base de Lagrange)**.**
Étant donnés $n+1$ valeurs distinctes x_i, l'espace des polynômes de degré n admet la base

$$L_i^{(n)}(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j} \tag{34.202}$$

pour $i = 0, \dots, n$.

Soit par exemple les valeurs de f données dans

x_i	$y_i = f(x_i)$
$x_0 = 5$	1
$x_1 = -7$	-23
$x_2 = 6$	-54
$x_3 = 0$	-954

Les polynômes de Lagrange pour ces données dépendent seulement des x_i, pas des y_i. En particulier,

$$L_0^{(3)}(x) = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)}, \tag{34.203}$$

etc.

La réponse est que

$$L_0^{(3)}(x) = \frac{x^3 + 13x^2 + 42x}{660} \tag{34.204a}$$

$$L_1^{(3)}(x) = \frac{-x^3 - x^2 + 30x}{84} \tag{34.204b}$$

$$etc. \tag{34.204c}$$

Ce qu'il y a de bien avec cette base est que en posant $a_i = f(x_i)$ alors le polynôme

$$\sum_{i=0}^{n} a_i L_i^{(n)}(x) \tag{34.205}$$

passe par les points $(x_i, f(x_i))$. Du coup il suffit d'écrire

$$P_3(x) = L_0^{(3)}(x) - 23L_1^{(n)}(x) - 54L_2^{(n)}(x) - 954L_3^{(n)}(x) = 4x^3 + 35x^2 - 84x - 954. \tag{34.206}$$

Un inconvénient de cette base est qu'elle est complètement dépendante des points choisis. Si on ajoute un point ou qu'on en prend un à peine différent, tous les coefficients changent. Mais en pratique, ajouter des points est quelque chose qui arrive souvent parce que souvent, après avoir vu le résultat d'un polynôme d'interpolation, on veut ajouter un point pour avoir un meilleur résultat.

34.93.
Une habitude : le premier et le dernier nœud se choisissent aux extrémités de l'intervalle sur lequel nous voulons une approximation.

Le but d'une approximation est d'avoir des approximations de $f(x^*)$ pour des valeurs de x^* qui ne soit pas une des abscisses données (parce que sur ces points, le polynôme et la fonction sont égaux). Nous considérons donc

$$f(x^*) \simeq P_n(x^*). \tag{34.207}$$

Si x^* est dans l'intervalle $I = [x_{min}, x_{max}]$ alors nous disons que nous calculons f par **interpolation**. Si au contraire x^* est en dehors de cet intervalle nous parlons d'**extrapolation**.

Si x^* est pris à l'extérieur de I, alors l'erreur risque d'être très grande, surtout parce que les polynômes tendent tous vers $\pm\infty$ lorsque $x \to \pm\infty$.

Autant l'interpolation via polynômes est le plus souvent valable, il faut garder à l'esprit que les extrapolations sont souvent mauvaises si x^* est trop loin des extrémités de I.

34.14.2 Base de Newton

Après la base canonique et la base de Lagrange, nous voyons la base de Newton. Soient encore $n+1$ points donnés du graphe de f.

Définition 34.94.
*La **base de Newton** pour les abscisses x_i est l'ensemble des polynômes suivants :*

$$1, (x - x_0), (x - x_0)(x - x_1), \dots, (x - x_0)(x - x_1)\dots(x - x_{n-1}) \tag{34.208}$$

Notons que ces polynômes n'utilisent pas le dernier point des x_i. Le polynôme passant par les points est

$$P_n(x) = x_0 + \sum_{i=1}^{n} c_i \prod_{j=0}^{i-1}(x - x_j). \tag{34.209}$$

Le calcul des c_i n'est pas absolument évident. Mais si nous ajoutons un point d'interpolation, les polynômes déjà calculés sont encore bons ; en particulier

$$P_{n+1}(x) = P_n(x) + c_{n+1} \prod_{j=0}^{n}(x - x_j). \tag{34.210}$$

Et cela est bien, parce que ça donne une façon de les calculer par récurrence.

Il y a plusieurs façons de calculer les c_i.

Les différences divisées sont des façons d'approximer les dérivées.

Définition 34.95.
*Soient $n+1$ nœuds x_i pour la fonction f. La **différence divisée** sont :*

Ordre 0

$$f[x_i] = f(x_i) \tag{34.211}$$

Ordre 1

$$f[x_i, x_j] = \frac{f[x_i] - f[x_j]}{x_i - x_j}. \tag{34.212}$$

Ordre 2

$$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{x_i - x_k}. \tag{34.213}$$

Ordre n

$$f[x_0,\dots,x_n]=\frac{f[x_0,\dots,x_{n-1}]-f[x_1,\dots,x_n]}{x_0-x_n}. \tag{34.214}$$

Les ordres font référence à l'ordre de dérivation qui est approximé.

Nous avons alors

$$c_i=f[x_0,\dots,x_i]. \tag{34.215}$$

Cela donne effectivement une méthode de récurrence pour trouver les coefficients c_i.

Remarque 34.96.
Pour calculer c_0, il faut seulement calculer $f[x_0]=f(x_0)$. Mais pour calculer c_1 il faut $f[x_0]$ et $f[x_1]$. Et pour c_2 il faut $f[x_0,x_1,x_2]$ qui demande $f[x_0,x_1]$ et $f[x_1,x_2]$, qui demande etc.

Il faut donc calculer en réalité tous les $f[x_i]$ pour terminer le calcul. Par contre, pour ajouter un point, il ne faut pas tout recalculer, et même pas tout conserver en mémoire. Il faut seulement garder en mémoire la dernière diagonale.

Exemple 34.97.
Soit les nœuds

x	$f(x)$
3	1
1	-3
5	2
6	4

Trouver le polynôme d'interpolation via la base de Newton.

x_i	$f[x_i]$	$f[x_i,x_j]$	$f[x_i,x_j,x_k]$
3	1	$\frac{1+3}{2}=2$	$\frac{f[x_0,x_1]-f[x_1,x_2]}{x_0-x_2}=\frac{2-\frac{5}{4}}{-2}=-\frac{3}{8}$
1	-3	$\frac{-3-2}{-4}=\frac{5}{4}$	$\frac{\frac{5}{4}-2}{-5}=\frac{3}{20}$
5	2	$\frac{2-4}{-1}=2$	$\frac{f[x_0,x_1,x_2]-f[x_1,x_2,x_3]}{x_0-x_3}=\frac{4}{40}$
6	4	...	...

Le polynôme d'interpolation sera

$$P_3(x)=c_0+c_1(x-x_0)+c_2(x-x_0)(x-x_1)+c_3(x-x_0)(x-x_1)(x-x_2). \tag{34.216}$$

△

Un exercice typique serait de donner tout pour 3 points puis de demander le polynôme qui aurait un quatrième point.

34.14.3 Méthode des minimums quadratiques

Soient $m+1$ points connus sur le graphe de la fonction f que nous devons approximer. Au lieu d'exiger que notre approximation ne passe par tous les points, nous allons chercher une approximation qui minimise la somme des carrés des erreurs sur ces points.

Soit $\mathcal{F}$ une classe de fonctions dans laquelle nous allons chercher l'approximation. Nous cherchons $g\in\mathcal{F}$ qui minimise

$$E(g)=\sum_{i=0}^{m}\big(f(x_i)-g(x_i)\big)^2\omega_i \tag{34.217}$$

EQooJGQVooKwZZVJ

où $\omega_i>0$ est une pondération. Souvent on prend $\omega_i=1$, mais pas toujours. La fonction E sur $\mathcal{F}$ est la **fonction d'erreur**.

34.98.
À part dans les exercices à la main, le nombre de points est grand, du type du milliard. Il est bien

entendu pas envisageable de faire passer un polynôme *exactement* par un milliard de points, parce que cela demanderait un polynôme de degré un milliard.

Plus généralement, d'un point de vue scientifique, avoir n paramètres libres pour n données expérimentales, ça ne passe pas Popper.

Afin de faire de la science qui passe Popper nous nous restreignons à une classe de fonction $\mathcal{F}$ dont la dimension n'est pas grande : $\dim(\mathcal{F}) \ll m$. Et nous notons $\dim(\mathcal{F}) = n + 1$.

Exemple 34.99.
La qualité d'une expérience peut être influencée par des paramètres extérieurs comme l'humidité, le vent, etc. Donc il est normal d'avoir des mesures moins précises que d'autres. On les pèse moins. △

Soit $g \in \mathcal{F}$, et une base $\{g_i\}_{i=0,\dots,n}$ de $\mathcal{F}$. Nous écrivons

$$g = a_0 g_0 + a_1 g_1 + \cdots + a_n g_n \tag{34.218}$$

La fonction donnée E donnée en (34.217), est, à partir du moment où $\mathcal{F}$ et une base sont choisis, une fonction des paramètres a_i que nous nommons $F(a_0, \dots, a_n)$. Il faut minimiser F, c'est-à-dire poser

$$\frac{\partial F}{\partial a_j} = 0 \tag{34.219}$$

pour $j = 0, \dots, n$. Cela sont $n + 1$ équations pour $n + 1$ inconnues. Notons que ces équations sont linéaires parce que chacun des termes est du type

$$\left(f(x_i) - \sum_{j=0}^{m} a_j g_j(x_i) \right)^2, \tag{34.220}$$

et lors de la dérivation par rapport à a_j, nous obtenons du degré 1.

34.14.4 Notre espace de Hilbert

Nous allons maintenant formaliser un peu tout cela. Dans [753] il est expliqué que si ω est une fonction strictement positive, alors l'espace $L^2_\omega\big([a, b]\big)$ dérivé de la norme

$$\|f\|^2_{L^2_\omega} = \int_a^b |f(x)|^2 \omega(x) dx \tag{34.221}$$

est un espace de Hilbert. Nous allons tenter le coup avec $\omega = \sum_{i=0}^{m} \omega_i \delta_{x_i}$ où δ_a est la distribution de Dirac [13] centrée en a.

Sur l'ensemble des fonctions $\mathbb{R} \to \mathbb{R}$ nous considérons la relation d'équivalence $f \sim g$ si $f(x_i) = g(x_i)$ pour tout $i = 0, \dots, m$. Nous notons L^2_ω cet ensemble.

Proposition 34.100.
La formule

$$\langle f, g \rangle = \sum_{i=0}^{m} f(x_i) g(x_i) \omega_i \tag{34.222}$$

EQooVUKMooIinkXQ

définit un produit scalaire sur L^2_ω. Ce dernier devient un espace de Hilbert.

Démonstration. Pour être un produit scalaire (définition 8.168), la forme considérée doit être symétrique et strictement définie positive. La symétrie de la formule (34.222) ne fait pas de doute. Le fait que ce soit semi-défini positif non plus. Pour le strict,

$$\langle f, f \rangle = \sum_{i=0}^{m} |f(x_i)|^2 \omega_i. \tag{34.223}$$

13. Définition 30.59.

Étant donné que $\omega_i > 0$ pour tout i, l'annulation de $\langle f, f\rangle$ implique l'annulation de $f(x_i)$ pour tout i. Cela signifie que f est dans la classe de 0 et donc est nul dans L^2_ω.

En ce qui concerne la complétude, la proposition 7.293 répond à notre place, étant donné que L^2_ω est de dimension finie. Une base est donnée par exemple par $e_i(x) = \delta_{x,x_i}$. Ici le δ est celui de Kronecker, et non celui de Dirac. □

Lemme 34.101.
Si la classe de fonctions $\mathcal{F}$ est un sous-espace vectoriel de L^2_ω et si $f \in L^2_\omega$ il existe un unique élément g de $\mathcal{F}$ minimisant la distance à f.

Démonstration. Le théorème de projection (au choix 12.142 ou 25.7) nous assure l'existence et l'unicité d'un élément de $\mathcal{F}$ minimisant la distance à $f \in L^2_\omega$. □

34.102.
Ce lemme est gentil, mais ne nous donne pas de méthodes pour trouver ce minimum. Nous allons donc écrire explicitement un système d'équations permettant de le trouver. Si $\{g_\alpha\}$ est une base (finie) de $\mathcal{F}$ alors nous cherchons le minimisant sous la forme $g = \sum_\alpha a_\alpha g_\alpha$.

Nous devons minimiser

$$E(g) = \sum_{i=0}^{m} \big(f(x_i) - g(x_i)\big)^2 \omega_i = \sum_{i=0}^{m} \big(f(x_i) - \sum_\alpha a_\alpha g_\alpha(x_i)\big)^2 \omega_i. \tag{34.224}$$

Vu que cela est maintenant plutôt une fonction des coefficients a_α que de la fonction g nous la notons $F(a_0, \dots, a_n)$. Il s'agit d'étudier le système d'équations

$$\frac{\partial F}{\partial a_\alpha} = 0. \tag{34.225}$$

Un tout petit peu de calcul mène au système

$$\sum_i \sum_\beta a_\beta \omega_i g_\alpha(x_i) g_\beta(x_i) = \sum_i \omega_i g_\alpha(x_i) f(x_i). \tag{34.226}$$

À droite nous reconnaissons $\langle f, g_\alpha\rangle$. et à gauche, $\sum_\beta a_\beta \langle g_\alpha, g_\beta\rangle$. Donc le système s'écrit

$$\sum_\beta a_\beta \langle g_\alpha, g_\beta\rangle = \langle f, g_\alpha\rangle. \tag{34.227}$$

Il y a une équation pour chaque valeur de α.

La matrice $A \in \mathbb{M}(n+1, \mathbb{R})$ donnée par $\langle g_\alpha, g_\beta\rangle$ étant strictement définie positive (c'est un produit scalaire), le système a une unique solution. Et comme cette matrice est de plus symétrique, elle est diagonalisable par le théorème spectral 8.227. Toutes ses valeurs propres sont strictement positives.

Notons pour la curiosité que si l'on considère la matrice $B \in \mathbb{M}(m \times n)$ donnée par

$$B_{ij} = \sqrt{\omega_i} g_j(x_l), \tag{34.228}$$

alors nous avons $A = B^t B$.

34.14.5 Droite de régression

La droite de régression est le cas particulier $n = 1$, c'est-à-dire un système 2×2. Nous cherchons $P = a_0 + a_1 x$. Et la base choisie est $g_0(x) = 1$, $g_1(x) = x$. Nous avons

$$\langle g_0, g_0\rangle = \sum_i \omega_i g_0(x_i) g_0(x_i) = \sum_i \omega_i \tag{34.229a}$$

$$\langle g_0, g_1\rangle = \sum_i \omega_i x_i \tag{34.229b}$$

$$\langle g_1, g_1\rangle = \sum_i \omega_i x_i^2. \tag{34.229c}$$

Donc pour approximer une fonction f il faut résoudre le système

$$\begin{pmatrix} \sum_i \omega_i & \sum_i \omega_i x_i \\ \sum_i \omega_i x_i & \sum_i \omega_i x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \langle f, g_0 \rangle \\ \langle f, g_1 \rangle \end{pmatrix}. \tag{34.230}$$

Pour calculer les produits $\langle f, g_\alpha \rangle$ il suffit de savoir f sur les points x_i. Et encore heureux, parce que toute la méthode est basée sur le fait que nous ne connaissons pas f ailleurs. C'est pour cela que nous avons défini L^2_ω comme un ensemble quotient.

Exemple 34.103.
Faisons la droite de régression pour les données avec tous les poids $\omega_i = 1$.

x_i	$f(x_i)$
-5	18
-3	7
1	0
3	7
4	16
6	50
8	67

Nous avons

$$\langle f, g_0 \rangle = \sum_i f(x_i) = 165 \tag{34.231a}$$

$$\langle f, g_1 \rangle = \sum_i f(x_i) x_i = 810. \tag{34.231b}$$

et donc le système

$$\begin{pmatrix} 7 & 14 \\ 14 & 160 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 165 \\ 810 \end{pmatrix} \tag{34.232}$$

dont résolution donne la droite de régression. △

Proposition 34.104.
Si tous les poids sont identiques, alors la droite de régression passe par le barycentre des points donnés :

$$\begin{cases} x_M = \dfrac{1}{m+1} \displaystyle\sum_{i=0}^{m} x_i & (34.233a) \\ y_M = \dfrac{1}{m+1} \displaystyle\sum_{i=0}^{m} y_i. & (34.233b) \end{cases}$$

Cela donne une vérification possible de la réponse trouvée.

Définition 34.105.
*L'**erreur quadratique** est la fonction $F(a_0, \dots, a_n)$ dont il est question plus haut. Et si une solution est connue, son erreur quadratique est la valeur de F pour cette solution.*

34.15 Conditionnement d'une matrice

SECooQGLRooZQzzsA

Soit le système d'équations linéaires $Au = b$ avec la matrice inversible A ainsi que le système perturbé $(A + \Delta A)u' = (b + \Delta b)$. Nous notons $\Delta u = u' - u$ et nous voudrions pouvoir dire des choses de l'erreur relative $\frac{\|\Delta u\|}{\|u\|}$.

Exemple 34.106 ([389]).
Soit la matrice

$$A = \begin{pmatrix} 10 & 7 \\ 7 & 5 \end{pmatrix} \tag{34.234}$$

et $b = \begin{pmatrix} 32 \\ 23 \end{pmatrix}$. La solution de $Au = b$ est $u = \begin{pmatrix} -1 \\ 6 \end{pmatrix}$. Si nous conservons la même matrice mais nous considérons $b = \begin{pmatrix} 32.1 \\ 22.9 \end{pmatrix}$. La solution devient $u' = \begin{pmatrix} 0.2 \\ 4.3 \end{pmatrix}$

En norme $\|.\|_\infty$ nous avons [14]

$$\frac{\|\Delta b\|}{\|b\|} = \frac{0.1}{32} = 0.003125 \tag{34.235}$$

et

$$\frac{\|\Delta u\|}{\|u\|} = \frac{1.7}{6} = 0.28. \tag{34.236}$$

Cela montre environ amplification d'un facteur 100 entre l'erreur sur b et l'erreur sur la solution. △

DEFooBKQWooJuoCGX

Définition 34.107.
*Le **conditionnement** de la matrice inversible $A \in \mathrm{GL}(n, \mathbb{C})$ est le nombre positif*

$$\mathrm{Cond}(A) = \|A\|\|A^{-1}\|. \tag{34.237}$$

Cette dénomination sera justifiée par le corolaire 34.113 parce qu'il est évident que le conditionnement d'une matrice est lié au conditionnement du problème de résolution d'un système linéaire.

Remarque 34.108.
Le conditionnement dépend de la norme choisie, mais cette dependence est contrôlée par la proposition 10.47 qui nous indique que si le conditionnement d'une matrice est grand dans une norme, il sera grand dans une autre norme.

D'autre part, lorsque nous écrirons $\|A\|$ nous supposerons toujours que $\|.\|$ est une norme d'algèbre [15] et donc que nous avons toujours

$$\|AB\| \leqslant \|A\|\|B\|. \tag{34.238}$$

De plus nous supposerons toujours avoir une norme subordonnée à une norme sur l'espace $\mathbb{C}^n$, de telle sorte à avoir

$$\|Au\| \leqslant \|A\|\|u\| \tag{34.239}$$

pour tout $u \in \mathbb{C}^n$. Voir aussi le lemme 10.66.

Proposition 34.109 ([389]).
Si A est une matrice inversible et si $\alpha \in \mathbb{C}$ nous avons :

(1) $\mathrm{Cond}(A) \geqslant 1$
(2) $\mathrm{Cond}(A) = \mathrm{Cond}(A^{-1})$
(3) $\mathrm{Cond}(\alpha A) = \mathrm{Cond}(A)$.

Si $Q \in \mathrm{O}(n)$ alors

(1) Nous avons $\mathrm{Cond}_2(Q) = 1$ où Cond_2 est le conditionnement pour la norme $\|.\|_2$.
(2) Nous avons aussi

$$\mathrm{Cond}_2(A) = \mathrm{Cond}_2(AQ) = \mathrm{Cond}_2(QA). \tag{34.240}$$

14. La proposition 10.47(3) montre que si nous voulions des estimations en norme $\|.\|_2$, il y aurait au maximum un facteur $\sqrt{2}$ par-ci par là.
15. Définition 10.63.

Démonstration. Nous savons que $\text{Cond}(\mathbb{1}) = 1$ et donc

$$1 = \|\mathbb{1}\| \leqslant \|A\|\|A^{-1}\| \tag{34.241}$$

parce que la norme utilisée est une norme matricielle.

Les deux autres formules sont évidentes à partit du fait que la définition du conditionnement de A est symétrique entre A et A^{-1}.

En ce qui concerne les formules relatives à la matrice orthogonale Q nous savons par la proposition 8.46(3) qu'une matrice orthogonale est une bijection de l'ensemble $\{x \in \mathbb{R}^n \text{ tel que } \|x\| = 1\}$. Par conséquent

$$\|AQ\| = \sup_{x \text{ tel que } \|x\|=1} \|AQx\| = \sup_{Q^{-1}x \text{ tel que } \|x\|=1} \|AQQ^{-1}x\| = \|A\|. \tag{34.242}$$

Donc $\|AQ\| = \|A\|$. Les assertions s'ensuivent immédiatement en remarquant que Q^{-1} est également orthogonale. □

Soit une matrice inversible $A \in \text{GL}(n, \mathbb{C})$. La matrice A^*A est hermitienne[16] et le théorème 10.20 nous assure que ses valeurs propres sont réelles. Par la remarque 10.21, ses valeurs propres sont même positives.

PROPooNUAUooIbVgcN

Proposition 34.110 ([389]).
Soit une matrice inversible $A \in \text{GL}(n, \mathbb{C})$, et $\mu_1 \leqslant \ldots \leqslant \mu_n$ les valeurs propres de A^*A. Alors nous avons la formule

$$\text{Cond}_2(A) = \sqrt{\frac{\mu_n}{\mu_1}}. \tag{34.243}$$

Démonstration. Par le théorème 12.120, la norme de A est liée au au rayon spectral de A^*A par

$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\mu_n}. \tag{34.244}$$

Vu que le spectre de AA^* est le même que celui de A^*A (lemme 12.103) nous avons aussi

$$\|A^{-1}\|_2 = \sqrt{\rho\big((A^{-1})^*A^{-1}\big)} = \sqrt{\rho\big((A^*A)^{-1}\big)} = \frac{1}{\sqrt{\mu_1}} \tag{34.245}$$

parce que la plus grande valeur propre de $(A^*A)^{-1}$ est l'inverse de la plus petite de A^*A.

Ces deux calculs étant,

$$\text{Cond}_2(A) = \|A\|_2\|A^{-1}\|_2 = \sqrt{\frac{\mu_n}{\mu_1}}. \tag{34.246}$$

□

¡¡ Avertissement/question au lecteur !! 34.111
À mon avis ce qui est dans la proposition 17.114 est le conditionnement de la matrice ou sa racine carrée ou un truc du genre. Il faut voir le lien entre les valeurs propres de A et celles de AA^*.

34.15.1 Perturbation du vecteur

PROPooGIXFooAhJkIs

Proposition 34.112 (Système linéaire : perturbation du vecteur[389]).
Soit une matrice inversible A et les systèmes d'équations linéaires EQooYQIGooPXqWoX

$$Au = b \tag{34.247a}$$
$$Au' = b'. \tag{34.247b}$$

En notant $\Delta u = u' - u$ et $\Delta b = b' - b$ nous avons

$$\frac{\|\Delta u\|}{\|u\|} \leqslant \text{Cond}(A)\frac{\|\Delta b\|}{\|b\|}. \tag{34.248}$$ EQooESXRooMYuyRa

16. Définition 8.179.

Démonstration. En soustrayant les équations (34.247) nous avons $\Delta b = A\Delta u$, et donc $\Delta u = A^{-1}\Delta b$. D'une part nous avons alors

$$\|\Delta u\| \leqslant \|A^{-1}\|\|\Delta b\|. \tag{34.249}$$

Et d'autre part, $\|b\| \leqslant \|A\|\|u\|$, ce qui donne

$$\frac{\|b\|}{\|A\|} \leqslant \|u\|. \tag{34.250}$$

En mettant les deux ensemble,

$$\frac{\|\Delta u\|}{\|u\|} \leqslant \frac{\|A^{-1}\|\|\Delta b\|}{\|b\|}\|A\| = \mathrm{Cond}(A)\frac{\|\Delta b\|}{\|b\|}. \tag{34.251}$$

□

Le corolaire suivant justifie le nom « conditionnement » au conditionnement d'une matrice.

CORooXKPWooJVHVvh

Corolaire 34.113.
Soit $A \in \mathrm{GL}(n,\mathbb{C})$ fixée et le problème de résoudre $Au = b$, c'est-à-dire la fonction

$$F(u,b) = Au - b. \tag{34.252}$$

(1) Ce problème est stable pour toute valeur de b.

(2) Nous avons une majoration pour le conditionnement relatif[17] *:*

$$K_{rel}(\eta, b_0) \leqslant \mathrm{Cond}(A). \tag{34.253}$$

EQooZHQJooTMKYfr

Démonstration. (1) **Stabilité** Vu que A est inversible, il existe une solution unique à tout système de la forme $Au = b'$. De plus $u(b) = A^{-1}b$, donc

$$\|u(b) - u(b_0)\| = \|A^{-1}(b - b_0)\| \leqslant \|A^{-1}\|\|b - b_0\|, \tag{34.254}$$

de telle sorte que la condition 34.26(2) fonctionne avec $K = \|A^{-1}\|$.

(2) **Conditionnement** En partant de la définition 34.49, et en utilisant la majoration de la proposition 34.112 sous la forme

$$\|u(b) - u(b_0)\| \leqslant \mathrm{Cond}(A)\|u(b_0)\|\frac{\|\Delta b\|}{\|b_0\|}, \tag{34.255}$$

nous obtenons :

$$\begin{aligned}
K_{rel}(b_0,\eta) &= K_{abs}(b_0,\eta)\frac{\|b_0\|}{\|u(b_0)\|} && \text{(34.256a)}\\
&= \sup_{\|b-b_0\|\leqslant\eta}\frac{\|u(b)-u(b_0)\|}{\|b-b_0\|}\frac{\|b_0\|}{\|u(b_0)\|} && \text{(34.256b)}\\
&\leqslant \sup_b \mathrm{Cond}(A)\frac{\|u(b_0)\|}{\|b_0\|}\|\Delta b\|\frac{1}{\|b-b_0\|}\frac{\|b_0\|}{\|u(b_0)\|} && \text{(34.256c)}\\
&= \mathrm{Cond}(A). && \text{(34.256d)}
\end{aligned}$$

□

REMooAIKIooJEBEqi

Remarque 34.114.
La notion de conditionnement relatif dépend aussi de la norme choisie. Dans la formule (34.253) il faut prendre le conditionnement $\mathrm{Cond}(A)$ pour la norme dans laquelle le K_{rel} est écrit. Encore une fois, toutes les normes étant équivalentes, cette majoration est à constante près bonne pour toutes les normes. Si la dimension est très grande, cette constante peut par contre être grande.

17. Si vous doutez de la norme à prendre, lisez la remarque 34.114

34.15.2 Perturbation de la matrice

Proposition 34.115 (Système linéaire : perturbation de la matrice[389]).
Soient les systèmes linéaires

$$Au = b \tag{34.257a}$$
$$A'u' = b \tag{34.257b}$$

avec A et A' inversibles. Nous notons $\Delta A = A' - A$. Alors ITEMooJMTKooSEBavB

(1)

$$\frac{\|\Delta u\|}{\|u'\|} \leqslant \text{Cond}(A)\frac{\|\Delta A\|}{\|A\|} \tag{34.258}$$

(2)

$$\frac{\|\Delta u\|}{\|u\|} \leqslant \text{Cond}(A)\frac{\|\Delta A\|}{\|A\|}\big(1 + \alpha(\|\Delta A\|)\big) \tag{34.259}$$

où $\lim_{x\to 0}\alpha(x) = 0$.

Démonstration. D'abord nous avons

$$0 = A'u' - Au \tag{34.260a}$$
$$= (A' - A)u' - Au' - Au \tag{34.260b}$$
$$= \Delta A u' + A\Delta u. \tag{34.260c}$$

Par conséquent, $\Delta u = -A^{-1}(\Delta A)u'$ et

$$\|\Delta u\| \leqslant \|A^{-1}\|\|\Delta A\|\|u'\|. \tag{34.261}$$ EQooYYITooSSczEj

Donc

$$\frac{\|\Delta u\|}{\|u'\|} \leqslant \|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|} = \text{Cond}(A)\frac{\|\Delta A\|}{\|A\|}. \tag{34.262}$$

Cela est (1).

Pour l'autre inégalité, nous avons $A' = A + \Delta A$ et donc

$$\|A'^{-1}\| = \|(A + \Delta A)^{-1}\| \tag{34.263}$$

Nous repartons alors de (34.261) en changeant le rôle de A et A' (et donc aussi de u et u'). Ce changement étant, $\|\Delta u\|$ et $\|\Delta A\|$ ne changent pas. Nous avons :

$$\frac{\|\Delta u\|}{\|u\|} \leqslant \|A'^{-1}\|\|\Delta A\| \tag{34.264a}$$
$$= \|(A + \Delta A)^{-1}\|\|\Delta A\|\frac{\text{Cond}(A)}{\|A\|\|A^{-1}\|} \tag{34.264b}$$
$$= \frac{\|(A + \Delta A)^{-1}\|}{\|A^{-1}\|}\frac{\|\Delta A\|}{\|A\|}\text{Cond}(A). \tag{34.264c}$$

Il reste à voir que

$$\lim_{\|\Delta A\|\to 0}\frac{\|(A + \Delta A)^{-1}\|}{\|A^{-1}\|} = 1, \tag{34.265}$$

ou autrement dit que

$$\lim_{A\to A'}\frac{\|A'^{-1}\|}{\|A^{-1}\|} = 1 \tag{34.266}$$ EQooJURGooFxYiAs

où la limite est celle dans $\text{GL}(n, \mathbb{C})$. Par définition de la topologie, la norme est continue (quelle qu'elle soit par l'équivalence de norme 10.49). Par le théorème 10.300, l'application $A \mapsto A^{-1}$ est également continue et commute donc avec la limite. Nous avons donc

$$\lim_{A'\to A}\|A'^{-1}\| = \|(\lim_{A'\to A}A')^{-1}\| = \|A^{-1}\|. \tag{34.267}$$

Donc la limite du quotient (34.266) est bien 1. □

34.16 Système linéaires (généralités)

Soit un système d'équations linéaires $Ax = b$ avec $A \in \mathbb{M}(n, \mathbb{R})$. Le problème est évidemment de savoir si il existe une unique solution x et de la déterminer. Nous supposons l'existence et l'unicité. C'est-à-dire que les conditions équivalentes [18] sont vérifiées :

(1) A est inversible, c'est-à-dire qu'il existe une matrice notée A^{-1} telle que $AA^{-1} = A^{-1}A = \mathbb{1}$.
(2) $\det(A) \neq 0$.

Note : si nous avons un système pas carré du type $Bx = v$ avec $B \in \mathbb{M}(n \times m)$ alors nous pouvons nous ramener à un système carré en écrivant

$$B^t Bx = B^t v. \tag{34.268}$$

Mais attention : bien que $B^t B$ soit symétrique et semi-définie positive, certaines valeurs propres peuvent être nulles.

34.116.
Deux choses générales en calcul numérique :

(1) On ne calcule pas l'inverse d'une matrice.
(2) On ne calcule même pas son déterminant.

Par conséquent nous ne faisons pas $x = A^{-1}v$.

Il faut garder en tête le fait que dans la pratique, la matrice A possède des millions de lignes et colonnes, si pas pire. Pour une matrice de taille de l'ordre du million, il y a 1000 milliards d'entrées. Si on compte 32 bits par nombre (précision simple, définition 34.8), c'est-à-dire 4 octets, il faut 4000 giga-octets pour enregistrer la matrice. Même pour la mémoire actuellement disponible, ce n'est pas rien. Surtout que souvent, la précision simple n'est pas utilisée, mais la précision double, ce qui donne 8000 giga pour enregistrer la matrice.

Heureusement, dans la majorité des cas pratiques, les matrices géantes qui apparaissent sont pleines de zéros.

Définition 34.117.
Une matrice est ***creuse*** *si elle possède beaucoup de zéros. Une matrice non creuse est dite* ***dense****.*

Notons que lorsqu'on parle de matrice comprenant beaucoup de « zéros », nous pensons à des éléments très petits, et non de vrai zéros.

Les matrices creuses ne sont pas mémorisées entièrement, mais plutôt comme un dictionnaire (i, j, v) qui donne la valeur v de A_{ij}.

Définition 34.118.
Une matrice est de « grande dimension » si elle ne peut pas être mise en mémoire sur un ordinateur donné. Sur certains ordinateurs, ça commence à 5000 *inconnues. Mais sur des plus forts, on peut aller jusqu'au million ou le milliard.*

Si la matrice est de petite dimension, il est possible d'utiliser des méthodes dites « directes ». Sinon, il faudra utiliser des méthodes itératives.

34.16.1 Les méthodes directes

Une méthode directe consiste à successivement transformer un système $A^{(0)}x = b^{(0)}$ en de nouveaux systèmes $A^{(i)}x = b^{(i)}$ dont la solution est identique jusqu'à obtenir un système $A^{(n-1)}x = b^{(n-1)}$ qui est à résolution immédiate.

L'avantage d'une méthode directe est qu'elle fournit une réponse exacte, pour autant que les calculs intermédiaires soient bien faits (ce qui n'est pas le cas sur un ordinateur).

Une méthode directe fonctionne en général avec un nombre de pas fixés par la taille du système. Par exemple pour un système $n \times n$, la méthode de Gauss demande exactement n pas, et il n'y a

18. L'équivalence est la proposition 8.16(2).

pas moyen de faire mieux. Or chaque pas demande de recalculer tous les éléments de la matrice. Encore une fois, si la matrice a une taille de l'ordre du milliard, cela fait 10^{18} éléments à recalculer un milliard de fois (sans compter les éléments du vecteur b). Infaisable.

Souvent une méthode directe passe par une factorisation $A = BC$ avec $B, C \in \mathbb{M}(n \times n)$.

Quelques types de matrices dont la résolution est immédiate :

— Matrice diagonale.

— Matrice orthogonale parce que si A est orthogonale alors $Ax = v$ se résout par $x = A^t v$ qui n'est pas particulièrement lourd à faire numériquement.

— Matrice triangulaire.

Remarque 34.119.
Pour une matrice diagonale, le déterminant et l'inverse sont faciles. Mais également pour la triangulaire. Pour une matrice triangulaire, le déterminant est le produit des éléments diagonaux, et il se fait qu'il y a une algorithme facile pour calculer l'inverse.

Donc en fait les matrices à résolution immédiates sont des matrices pour lesquelles l'inverse et le déterminant sont facile à calculer.

34.16.2 Méthodes itératives

Si la matrice est trop grande, il n'est pas possible de faire des manipulations de matrices à chaque itération.

En général, les méthodes itératives ne convergent pas toujours. Mais lorsqu'une méthode converge, c'est une propriété de la matrice, et donc la convergence aura lieu pour tout vecteur de départ x_0. Cela est très différent du cas des équations non linéaires type Newton pour lesquelles la convergence peut fortement dépendre du point de départ.

34.17 Système linéaires (méthodes directes)

Les matrices que nous sommes autorisés à inverser sont les matrices

— orthogonales : l'inverse est la transposée

— diagonales : l'inverse est diagonale avec les inverses sur la diagonale

— triangulaires : nous en parlons maintenant.

34.17.1 Inversion de matrice triangulaire

Si T est une matrice triangulaire (mettons supérieure pour fixer les idées), il est possible d'en calculer l'inverse sans trop d'efforts. Notons B la matrice inverse que nous allons construire ligne par ligne. Vu que $BT = \mathbb{1}$ nous avons

$$\delta_{1j} = \sum_{k=1}^{n} B_{1k} T_{kj} = \sum_{k=1}^{j} B_{1k} T_{kj} \tag{34.269}$$

parce que $T_{kj} = 0$ pour $k > j$. Donc nous pouvons calculer les éléments B_{1j} un par un parce que chacun ne dépend que des précédents. Le même procédé fonctionne pour les autres lignes :

$$\delta_{ij} = \sum_{k=1}^{j} B_{ij} T_{kj}. \tag{34.270}$$

Et tu notes que le calcul peut être parallélisé : le calcul de la ligne numéro j ne dépend pas du résultat des autres lignes.

34.17.2 Transformation gaussienne

DEFooOJRKooDcrbeU

Définition 34.120 (Transformation gaussienne[754]).
*Soit $x \in \mathbb{R}^n$ avec $x_k \neq 0$. La k^e**transformation gaussienne** pour x est la matrice*

$$M_k(x) = \mathbb{1} - T_k(x) \tag{34.271}$$

où $T_k(x)$ est la matrice unité à qui on a ajouté le vecteur

$$\tau_k(x) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_{k+1}/x_k \\ \vdots \\ x_n/x_k \end{pmatrix} \tag{34.272}$$

à la k^ecolonne.

Autrement dit, la matrice $M_k(x)$ est la matrice

$$M_k(x) = \begin{pmatrix} 1 & & & & & \\ 0 & \ddots & & & & \\ \vdots & 0 & 1 & & & \\ \vdots & \vdots & -x_{k+1}/x_k & 1 & & \\ \vdots & \vdots & \vdots & 0 & \ddots & \\ 0 & 0 & -x_n/x_k & 0 & & 1 \end{pmatrix} \tag{34.273}$$

EQooMWXLooBDtsKS

En coordonnées nous avons

$$\big(M_k(x)\big)_{ij} = \delta_{ij} - \tau_k(x)_i\delta_{kj}. \tag{34.274}$$

34.121.
Les matrices de transformation gaussienne sont des matrices triangulaires de diagonale unitaire (c'est-à-dire avec des 1 sur la diagonale).

Lemme 34.122.
Si $x \in \mathbb{R}^n$ alors nous avons

$$M_k(x)x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{34.275}$$

Démonstration. Nous avons

$$\big(M_k(x)x\big)_i = \sum_l M_k(x)_{il}x_l = \sum_l \big(\delta_{il} - \tau_k(x)_i\delta_{kl}\big)x_l = x_i - \tau_k(x)_i x_k. \tag{34.276}$$

Si $i \leqslant k$ nous avons $\tau_k(x)_i = 0$ et donc $\big(M_k(x)x\big)_i = x_i$. Si par contre $i \geqslant k+1$ alors $\tau_k(x)_i = \frac{x_i}{x_k}$ et alors $\big(M_k(x)x\big)_i = 0$. □

LEMooPFWWooUmMsVH

Lemme 34.123.
Si $y \in \mathbb{R}^n$ vérifie $y_i = 0$ pour $i > k$ alors $M_{k+1}(x)y = y$.

Démonstration. C'est une simple vérification :

$$\big(M_{k+1}(x)y\big)_i = \sum_l \big(\delta_{il} - \tau_{k+1}(x)_i\delta_{k+1,l}\big)y_l = y_i - \tau_{k+1}(x)_i y_{k+1}. \tag{34.277}$$

Mais comme $y_{k+1} = 0$ il nous reste automatiquement y_i. □

Le sens de ce lemme est si un vecteur est déjà « gaussiannisé » au niveau k, alors en lui appliquant une transformation gaussienne de niveau plus élevé que k, il ne change pas. Ce fait est important parce qu'il assure que lorsque l'on avance dans le processus de Gauss, chaque étape ne détruit pas les précédentes.

Le lemme suivant nous indique que l'inverse d'une matrice de transformation gaussienne est facile à calculer [19].

LEMooFHZDooZiKdbr

Lemme 34.124.
L'inverse de la transformation gaussienne

$$M_k(x)_{ij} = \delta_{ij} - \tau_k(x)_i\delta_{kj}. \tag{34.278}$$

est la matrice donnée par

$$M_k(x)^{-1}_{ij} = \delta_{ij} + \tau_k(x)_i\delta_{kj}. \tag{34.279}$$

Autrement dit, il suffit de changer le signe de la partie non diagonale.

Démonstration. Il s'agit d'une simple vérification, utilisant le produit matriciel explicite, et en remarquant que $\tau_k(x)_k = 0$ pour tout k. □

34.17.3 Méthode de Gauss pour résoudre des systèmes d'équations linéaires

Pour résoudre un système d'équations linéaires, on procède comme suit :

(1) Écrire le système sous forme matricielle.

$$\text{p.ex. } \begin{cases} 2x + 3y &= 5 \\ x + 2y &= 4 \end{cases} \Leftrightarrow \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right)$$

(2) Se ramener à une matrice avec un maximum de 0 dans la partie de gauche en utilisant les transformations admissibles :

(2a) Remplacer une ligne par elle-même + un multiple d'une autre ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right) \overset{L_1 - 2.L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right)$$

(2b) Remplacer une ligne par un multiple d'elle-même ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right) \overset{-L_1 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 2 & 4 \end{array}\right)$$

(2c) Permuter des lignes.

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 0 & -2 \end{array}\right) \overset{L_1 \mapsto L_2' \text{ et } L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right)$$

(3) Retransformer la matrice obtenue en système d'équations.

$$\text{p.ex. } \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right) \Leftrightarrow \begin{cases} x &= -2 \\ y &= 3 \end{cases}$$

19. Elle rentre d'ailleurs dans la catégorie des matrices triangulaires dont nous avons déjà discuté l'inverse.

Remarques :

— Si on obtient une ligne de zéros, on peut l'enlever :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \end{array}\right)$$

— Si on obtient une ligne de zéros suivie d'un nombre non-nul, le système d'équations n'a pas de solution :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{array}\right) \Leftrightarrow \begin{cases} \cdots \\ \cdots \\ 0x + 0y + 0z = 7 \end{cases} \Rightarrow \textbf{Impossible}$$

— Si on moins d'équations que d'inconnues, alors il y a une infinité de solutions qui dépendent d'un ou plusieurs paramètres :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 1 & 0 & -2 & 2 \\ 0 & 1 & 3 & 0 \end{array}\right) \Leftrightarrow \begin{cases} x - 2z = 2 \\ y + 3z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 2 + 2\lambda \\ y = -3\lambda \\ z = \lambda \end{cases}$$

34.17.4 Méthode de Gauss sans pivot (décomposition LU)

La méthode de Gauss est souvent aussi appelée méthode « LU » qui va décomposer $A = LU$ où L est triangulaire inférieure et U est triangulaire supérieure. La décomposition est même plus précise que cela : on demande que L ait seulement des 1 sur la diagonale.

Si A est une matrice nous notons

$$\Delta_k(A) = (A_{ij})_{1\leqslant i,j\leqslant k} \tag{34.280}$$

la matrice tronquée dont nous ne gardons que le carré $k \times k$ en haut à gauche.

LEMooXEJFooGiYoyb

Lemme 34.125.
Soit S une matrice triangulaire inférieure. Soient également A et B telles que $B = SA$. Alors

$$\Delta_k(B) = \Delta_k(S)\Delta_k(A). \tag{34.281}$$

Démonstration. En effet nous avons

$$\Delta_k(B)_{ij} = \sum_{l=1}^{n} S_{il}A_{lj}. \tag{34.282}$$

EQooHBZZooHtjjsE

Dans la somme sur l il ne reste que les termes $l \leqslant i$. Mais dans le calcul des éléments de matrice $\Delta_k(B)_{ij}$, nous avons évidemment $i, j \leqslant k$. Donc $l \leqslant i \leqslant k$. Les seuls éléments de matrice de A qui sont utilisés dans la somme (34.282) sont les éléments A_{lj} avec $l, j \leqslant k$.

Nous pouvons donc limiter la somme à $l = k$ au lieu de $l = n$ et écrire $\Delta_k(A)_{lj}$ au lieu de A_{lj}.

Même chose en ce qui concerne S. À partir du moment où l est limité à k, les éléments S_{il} et $\Delta_k(S)_{il}$ sont les mêmes. □

THOooUXKJooYaPhiu

Théorème 34.126 (Décomposition LU[755, 1])**.**
Soit une matrice A inversible telle que $\det(\Delta_k(A)) \neq 0$ pour tout k. Alors il existe un unique couple de matrices (L, U) telles que

— *U soit triangulaire supérieure*

— *L soit triangulaire inférieure, de diagonale unité*

— *$A = LU$.*

De plus pour tout $k \leqslant n$ nous avons

$$\Delta_k(A) = \Delta_k(L)\Delta_k(U). \tag{34.283}$$

Démonstration. Nous allons prouver par récurrence le fait suivant : pour tout $1 \leqslant k \leqslant n-1$ il existe des matrices E_i $(i = 1, \dots, k)$ telles que en posant

$$A_k = E_k \dots E_1 A, \tag{34.284}$$

— E_j est une transformation gaussienne[20] pour la j^{e} colonne,

— pour tout $j \leqslant k$, $A_{ij} = 0$ dès que $i > j$. Autrement dit la matrice A_k est triangulaire supérieure jusqu'à (non comprise) la $(k+1)^{\text{e}}$ colonne (laquelle est quelconque). Exemple pour fixer les idées : pour une matrice $A \in \mathbb{M}(4 \times 4)$, la matrice A_2 doit avoir la forme

$$A_2 = E_2E_1A = \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & \circledast & * \\ 0 & 0 & * & * \end{pmatrix} \tag{34.285}$$

où les éléments notés $*$ sont à priori non nuls,

— l'élément de matrice $(A_k)_{k+1,k+1}$ est non nul (celui entouré dans l'exemple).

La chose un peu triste dans cette démonstration est que l'initialisation va être très ressemblante au pas de récurrence.

(1) **Initialisation : $k = 1$**

Vu que $\Delta_1(A)$ est inversible, l'élément A_{11} est non nul. Il existe donc une transformation gaussienne E_1 telle que la première colonne de la matrice $A_1 = E_1A$ soit nul sauf la première composante. En particulier $(A_1)_{21} = 0$.

Par le lemme 34.125, nous avons $\Delta_2(A_1) = \Delta_2(E_1)\Delta_2(A)$, donc[21]

$$\det\big(\Delta_2(A_1)\big) = \det\big(\Delta_2(E_1)\big)\det\big(\Delta_2(A)\big). \tag{34.286}$$

Étant donnée la forme (34.273), toutes les matrices du type $\Delta_k(E_i)$ ont un déterminant unité, et par hypothèse $\Delta_2(A)$ est inversible, donc de déterminant non nul. Par conséquent $\det\big(\Delta_2(A_1)\big) \neq 0$. Mais comme ce déterminant est le produit des éléments diagonaux (c'est une matrice triangulaire), ces derniers ne sont pas nuls. Finalement, $(A_1)_{22} \neq 0$.

(2) **Le pas de récurrence**

Nous supposons avoir $A_k = E_k \dots E_1 A$ avec $(A_k)_{k+1,k+1} \neq 0$. Alors il existe une transformation gaussienne E_{k+1} de la $(k+1)^{\text{e}}$ colonne telle que $A_{k+1} = E_{k+1}A_k$ soit une matrice dont la $(k+1)^{\text{e}}$ colonne n'ait que des zéros en dessous de la $(k+1)^{\text{e}}$ position. Vu le lemme 34.123, cette transformation n'affecte pas les colonnes précédentes.

La matrice A_{k+1} est donc triangulaire supérieure jusqu'à la $(k+1)^{\text{e}}$ colonne.

Vu que le produit $E_{k+1} \dots E_1$ est une matrice triangulaire inférieure, le lemme 34.125 fonctionne encore et nous avons

$$\Delta_{k+1}(A_{k+1}) = \Delta_{k+1}(E_{k+1} \dots E_1)\Delta_{k+1}(A). \tag{34.287}$$

En ce qui concerne les déterminants, par hypothèse, nous avons $\det\big(\Delta_{k+1}(A)\big) \neq 0$ ainsi que $\det\big(\Delta_{k+1}(E_{k+1} \dots E_1)\big) = 1$. Donc

$$\det\big(\Delta_{k+1}(A_{k+1})\big) \neq 0. \tag{34.288}$$

Cette matrice étant triangulaire de déterminant non nul, ses éléments diagonaux sont tous non nuls ; en particulier $(A_{k+1})_{k+2,k+2} \neq 0$.

20. Transformation gaussienne, définition 34.120.

21. Le déterminant est multiplicatif, proposition 8.16(1).

En poussant la récurrence jusqu'au bout, la matrice

$$A_{n-1} = E_{n-1} \dots E_n A \tag{34.289}$$

est triangulaire supérieure.

Nous posons alors $L = (E_{n-1} \dots E_n)^{-1}$ et $U = A_{n-1}$. Cela prouve l'existence parce que

$$A = (E_{n-1} \dots E_1)^{-1} A_{n_1}. \tag{34.290}$$

Encore une fois, le lemme 34.125 nous donne

$$\Delta_k(A) = \Delta_k\Big((E_{n_1} \dots E_1)^{-1}\Big)\Delta_k(A_{n-1}), \tag{34.291}$$

ou encore $\Delta_k(A) = \Delta_k(L)\Delta_k(U)$.

En ce qui concerne l'unicité, si $A = L_1U_1 = L_2U_2$ alors $L_2^{-1}L_1 = U_2U_1^{-1}$. Vu qu'à gauche nous avons une matrice triangulaire inférieure et que à droite nous avons une triangulaire inférieure, nous savons que les deux membres représentent une matrice diagonale. Mais à gauche, la diagonale est unitaire. Donc les deux membres représentent la matrice unité. □

34.127.
En pratique, pour résoudre $Ax = b$, il faut seulement appliquer les transformations gaussiennes à la matrice élargie $(A|b)$ pour finir sur un système du type

$$Ux = b' \tag{34.292}$$

qui est immédiatement soluble. Autrement dit, en effectuant les annulations de colonnes, la matrice U est « gratuite ».

Il n'est pas indispensable de calculer la matrice L qui, elle, demande à chaque étape de se souvenir de la matrice E_i utilisée. Si il faut résoudre plusieurs systèmes $Ax_i = b_i$, nous pouvons encore travailler avec la matrice encore plus élargie $(A|b_1 \dots b_m)$.

Si par contre nous ne connaissons pas à l'avance l'ensemble des vecteurs b avec lesquels il faudra résoudre le système, il est bon de calculer la décomposition $A = LU$ in extenso, c'est-à-dire de garder une trace des matrices L et U séparément. Dans ce cas, résoudre $Ax = b$ revient à résoudre $Ly = b$, et ensuite $Ux = y$. Ce sont deux systèmes de résolution directe parce que les matrices sont triangulaires.

34.128.
Le fait que

$$\Delta_k(A) = \Delta_k(L)\Delta_k(U) \tag{34.293}$$

nous dit que si après avoir calculé L et U nous remarquons que le système est un peu plus petit ou un peu plus grand que prévu, tout le travail n'est pas perdu. En particulier si le système est plus petit que prévu, l'adaptation de L et U est immédiate.

Notons que U et L sont inversibles, et que $\det(L) = 1$. Donc $\det(U) = \det(A)$.

Exemple 34.129.
Pour travailler la méthode de Gauss pour le système $Ax = b$, nous introduisons la matrice un peu augmentée $(A|b)$. Nous faisons un exemple. Soit à résoudre

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 3 & 10 \\ -2 & 1 & 73 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 11 \\ 28 \\ 3 \end{pmatrix}. \tag{34.294}$$

Nous introduisons la matrice augmentée

$$(A|b)^{(0)} = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 4 & 3 & 10 & 28 \\ -2 & 1 & 7 & 3 \end{pmatrix}. \tag{34.295}$$

Le premier pas consiste à annuler tous les éléments sous la diagonale de la première colonne. Autrement dit, nous prenons le 2 comme pivot. Nous introduisons les multiplicateurs $l_{ij} = \frac{A_{ij}}{A_{i1}}$. La nouvelle matrice est :

$$(A|b)^{(1)} = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 0 & 1 & 4 & 6 \\ 0 & 2 & 10 & 14 \end{pmatrix} \tag{34.296}$$

où nous avons utilisé les multiplicateurs $l_{21} = 2$, $l_{31} = -1$.

Et la matrice suivante est :

$$(A|b)^{(2)} = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 2 & 2 \end{pmatrix} \tag{34.297}$$

où nous avons utilisé le multiplicateur $l_{32} = 2$.

Cela est un système de résolution immédiate :

$$\begin{cases} 2x + y + 3z = 11 & \text{(34.298a)} \\ y + 4z = 6 & \text{(34.298b)} \\ 2z = 2. & \text{(34.298c)} \end{cases}$$

La troisième donne $z = 1$. Ensuite $y + 4 = 6$, donc $y = 2$. Et la première donne : $2x + 2 + 3 = 11$, c'est-à-dire $2x = 6$, enfin : $x = 3$.

Solution : $(x, y, z) = (3, 2, 1)$.

Nous notons surtout que dans $(A|b)^{(2)}$ nous avons une matrice triangulaire supérieure. Où est la matrice triangulaire inférieure ? En réalité la matrice L est la matrice des multiplicateurs :

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}. \tag{34.299}$$

△

Le problème de cette méthode est que faisant ainsi nous risquons d'avoir un zéro sur un des pivots. Par exemple tomber sur

$$(A|b) = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 0 & 0 & 4 & 6 \\ 0 & 2 & 10 & 14 \end{pmatrix}. \tag{34.300}$$

Le zéro sur la deuxième ligne nous ennuie si nous voulons tout faire dans l'ordre. Mais notons qu'en échangeant les deux dernières lignes, tout va bien : le système donné par

$$(A|b) = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 0 & 2 & 10 & 14 \\ 0 & 0 & 4 & 6 \end{pmatrix} \tag{34.301}$$

fonctionne très bien. Et même tellement bien qu'il est de résolution immédiate, dans ce cas.

Un autre problème est que si un des pivots est 10^{-14}, le multiplicateur sera de l'ordre 10^{14}, qui est mal représenté en mémoire. Il est donc bon de prendre les pivots le plus grand possible. Si le pivot est le plus grand nombre en valeur absolue d'une colonne, alors les nombres x_{k+i}/x_k qui entrent dans la matrice de transformation gaussienne sont des nombres dans $[-1, 1]$ qui sont bien représentés en mémoire.

Tout cela nous incite à développer une méthode de Gauss qui permet de tenir une trace des permutations.

34.17.5 Matrice de permutation élémentaire

Définition 34.130.
*Une **matrice de permutation élémentaire** est une matrice obtenue en permutant deux lignes de la matrice identité. Nous notons P_{ij} la matrice obtenue en inversant les lignes i et j de la matrice identité.*

Les autres opérations sur les lignes sont répertoriées en 4.88.

Exemple 34.131.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = P_{12}. \tag{34.302}$$

△

Lemme 34.132.
La matrice $P_{ij}A$ est la matrice A avec ses lignes i et j inversées.

Démonstration. Il suffit d'écrire

$$(P_{ij}A)_{kl} = \sum_m (P_{ij})_{km} A_{ml} \tag{34.303}$$

et de faire trois cas selon que $k = i$, $k = j$ ou k différent de i et j. Si $k = i$ alors $(P_{ij})_{im} = \delta_{mj}$ et si k est différent de i et j alors $(P_{ij})_{mk} = \delta_{km}$ (troisième cas similaire au premier). □

Et la matrice AP_{12} est la A avec ses deux premières *colonnes* échangées.
Avec ces notations, notre matrice $(A|b)^{0'}$ est

$$P_{12}(A|b)^{(0)}. \tag{34.304}$$

Puis la matrice $(A|b)^{(1')}$ est

$$P_{23}(A|b)^{(1)}. \tag{34.305}$$

Et la matrice P qui arrive dans $PA = LU$ est la matrice $P = P_{23}P_{21}$, qui est une matrice de permutations non élémentaire. Elle vaut :

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{34.306}$$

LEMooYIYIooYhnaOt

Lemme 34.133 ([756])**.**
Si $i, j > k$ alors les matrices de permutation élémentaires ont la relation de « commutation » suivante avec les transformations gaussiennes :

$$M_k(x)P_{ij} = P_{ij}M_k\big(P_{ij}(x)\big). \tag{34.307}$$

Démonstration. Il suffit de calculer les éléments de matrice :

$$\big(P_{ij}M_k(x)\big)_{st} = (P_{ij})_{st} - \sum_m (P_{ij})_{sm}\tau_k(x)_m\delta_{kt}, \tag{34.308}$$

mais

$$\sum_m (P_{ij})_{sm}\tau_k(x)_m = \big(P_{ij}\tau_k(x)\big)_s = \tau_k\big(P_{ij}(x)\big)_s \tag{34.309}$$

parce que $i, j > k$ implique que dans $P_{ij}\tau_k(x)$ nous inversons deux élément non nuls de $\tau_k(x)$, tout en laissant le k^{e}élément. Le dénominateur ne change pas et il s'agit réellement d'une inversion de ligne. Donc

$$\big(P_{ij}M_k(x)\big)_{st} = (P_{ij})_{st} - \tau_k\big(P_{ij}x\big)_s\delta_{kt}. \tag{34.310}$$

EQooIBVJooTQWCGT

De l'autre côté,

$$\big(M_k(y)P_{ij}\big)_{st} = (P_{ij})_{st} - \tau_k(y)_s(P_{ij})_{kt}. \tag{34.311}$$

Mais comme $i, j > k$ la k^{e}ligne de P_{ij} est la même que celle de la matrice unité, donc $(P_{ij})_{kt} = \delta_{kt}$.

$$\big(M_k(y)P_{ij}\big)_{st} = (P_{ij})_{st} - \tau_k(y)_s\delta_{kt}. \tag{34.312}$$

Cela correspond bien à (34.310). □

34.18 Méthode de Gauss avec pivot partiel (décomposition PLU)

34.18.1 L'idée

À chaque pas, nous faisons une permutation de ligne. Nous permutons à chaque pas la première ligne avec celle qui a le pivot le plus grand (en valeur absolue). Donc :

$$(A|b)^{(0)} = \begin{pmatrix} 2 & 1 & 3 & 11 \\ 4 & 3 & 10 & 28 \\ -2 & 1 & 7 & 3 \end{pmatrix} \tag{34.313}$$

Nous commençons par déplacer des lignes :

$$(A|b)^{(0')} = \begin{pmatrix} 4 & 3 & 10 & 28 \\ 2 & 1 & 3 & 11 \\ -2 & 1 & 7 & 3 \end{pmatrix}. \tag{34.314}$$

Les multiplicateurs sont $l_{21} = 1/2$ et $l_{31} = -1/2$. Le fait est que les multiplicateurs ont toujours le plus grand dénominateur possible et nous avons alors toujours $0 \leqslant |l_{ij}| \leqslant 1$, qui sont des nombres relativement petits, et bien représentés en mémoire.

Nous avons la nouvelle matrice

$$(A|b)^{(1)} = \begin{pmatrix} 4 & 3 & 10 & 28 \\ 0 & -1/2 & -2 & -3 \\ 0 & 5/2 & 12 & 17 \end{pmatrix}. \tag{34.315}$$

Le pivot serait $-1/2$. Nous cherchons un pivot plus grand en dessous de ce $-1/2$ (et pas au dessus, sinon on casserait les zéros déjà trouvés). Nous trouvons le $5/2$ qui est plus grand. Nous permutons donc les deux dernières lignes :

$$(A|b)^{(1')} = \begin{pmatrix} 4 & 3 & 10 & 28 \\ 0 & 5/2 & 12 & 17 \\ 0 & -1/2 & -2 & -3 \end{pmatrix} \tag{34.316}$$

où le pivot est maintenant $l_{32} = -1/5$. La matrice suivante :

$$(A|b)^{(1)} = \begin{pmatrix} 4 & 3 & 10 & 28 \\ 0 & 5/2 & 12 & 17 \\ 0 & 0 & 2/5 & 2/5 \end{pmatrix} \tag{34.317}$$

Dans ce cas, la matrice L n'est pas aussi simple à construire parce que nous avons permuté des choses. Dans ce cas, la matrice L est encore de la forme

$$L = \begin{pmatrix} 1 & 0 & 0 \\ . & 1 & 0 \\ . & . & 1 \end{pmatrix}. \tag{34.318}$$

Mais vu que nous avons permuté les lignes 2 et 3 au deuxième pas, nous devons permuter l_{21} et l_{31} avant de remplir la matrice L avec les multiplicateurs :

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 1/2 & -1/5 & 1 \end{pmatrix}. \tag{34.319}$$

Notons que ces L et U ne sont pas les mêmes que le LU obtenu sans pivot. Où est l'unicité ? Elle est que en fait maintenant nous n'avons pas $A = LU$, mais

$$PA = LU \tag{34.320}$$

où P est une matrice de permutation.

34.18.2 Le théorème

PROPooGCPAooDrlrGu

Proposition 34.134 (Méthode de Gauss avec pivot partiel[756]).
Soit une matrice inversible $A \in \mathbb{M}(n, \mathbb{C})$. Il existe

— *une matrice de permutations P*

— *une matrice triangulaire inférieure de diagonale unitaire L,*

— *une matrice triangulaire supérieure inversible U*

telles que

$$PA = LU. \tag{34.321}$$

Notons que cette proposition ne demande que l'hypothèse d'inversibilité pour A. Il n'y a pas d'hypothèses sur tous les mineurs comme c'était le cas avec Gauss sans pivot.

Démonstration. Nous prouvons par récurrence qu'il existe des matrices Q_k, E_1, …, E_k et A_k telles que

EQooOBXWooXxwXSe

$$Q_k A = E_1 \dots E_k A_k \tag{34.322}$$

avec

(1) Q_k est une matrice de permutation

(2) E_i est une transformation gaussienne sur la i^{e} ligne

(3) A_k est triangulaire supérieure jusqu'à la k^{e} colonne.

Sachant que $\det(Q_k) = \pm 1$, et que $\det(E_i) = 1$, le passage au déterminant dans (34.322) nous donne $\det(A_k) \neq 0$ et si nous notons $\Omega_k(A)$ la matrice tronquée de A, ne gardant que les entrées plus grandes que k, nous avons

$$\det(A_k) = \prod_{i=1}^{k} (A_k)_{i}i \det\big(\Omega_{k+1}(A_k)\big). \tag{34.323}$$

Donc : $(A_k)_{ii} \neq 0$ pour $i \leqslant k$ et $\det\big(\Omega_{k+1}(A_k)\big) \neq 0$.

Pour fixer les idées, voici une image de $k = 2$:

$$\begin{array}{c} \overbrace{}^{\Delta_k(A_2)} \\ \left(\begin{array}{ccccc} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{array}\right) \\ \underbrace{}_{\Omega_{k+1}(A_2)} \end{array} \tag{34.324}$$

Étant donné que $\det\big(\Omega_{k+1}(A_k)\big) \neq 0$, parmi les nombres $(A_k)_{i,k+1}$ $(i \geqslant k+1)$, au moins un est non nul et nous posons r_{k+1} tel que $|(A_k)_{r_{k+1},k+1}|$ soit maximum parmi ces éléments.

Le nombre r_{k+1} est enregistré parce qu'il servira à écrire la matrice P plus tard. Les matrices E_i ne sont pas enregistrées, parce que nous verrons qu'elles vont encore changer. Seule la dernière sera enregistrée.

La composante $(k+1, k+1)$ de la matrice

$$P_{r_{k+1},k+1} A_k \tag{34.325}$$

est non nulle et peut donc servir de pivot. Soit M_{k+1} la transformation gaussienne pour la $(k+1)^{\text{e}}$colonne de la matrice $P_{r_{k+1},k+1}A_k$. La matrice

$$A_{k+1} = M_{k+1}P_{r_{k+1},k+1}A_k \tag{34.326}$$

est alors une matrice triangulaire supérieure jusqu'à la $(k+1)^{\text{e}}$colonne. En posant $E_{k+1} = M_{k+1}^{-1}$ nous avons

$$P_{r_{k+1},k+1}E_{k+1}A_{k+1} = A_k, \tag{34.327}$$

et nous nous sentons en droit de récrire l'équation de départ (34.322) :

$$Q_kA = E_1 \dots E_kA_k = E_1 \dots E_kP_{r_{k+1},k+1}E_{k+1}A_{k+1}. \tag{34.328}$$

Le lemme 34.133 nous permet de ramener la matrice $P_{r_{k+1},k+1}$ en première position, quitte à modifier un peu (pas beaucoup) chacune des matrices E_i $(i = 1, \dots, k)$. C'est pour cela que nous n'enregistrons pas les matrices E_i. Nous avons donc

$$P_{r_{k+1},k+1}Q_kA = E_1' \dots E_k'E_{k+1}A_{k+1} \tag{34.329}$$

où

— Le produit $P_{r_{k+1},k+1}Q_k$ est encore une matrice de permutation, et mieux : elle vaut

$$\prod_{i=1}^{k+1} P_{r_i,i}. \tag{34.330}$$

Cela montre qu'il est suffisant d'enregistrer les nombres r_i pour reconstituer cette partie.

— La matrice E_i' est une transformation gaussienne pour la i^{e} colonne.

— La matrice A_{k+1} est triangulaire supérieure jusqu'à la $k+1^{\text{e}}$ ligne.

La récurrence est maintenant finie et nous pouvons écrire avec $k = n$:

$$Q_nA = E_1 \dots E_nA_n \tag{34.331}$$

où le produit $E_1 \dots E_n$ est triangulaire inférieure et A_n est triangulaire supérieur.

Maintenant nous enregistrons la matrice $U = A_n$, le produit $L = \prod_{i=1}^{n} E_n$ et les nombres r_i qui permettent de retrouver P. □

Note : dans l'équation (34.331) nous avons bien entendu massivement renommé les E_i' en E_i. En réalité la matrice E_1 vient avec n primes sur la tête.

Dans les exemples 34.138, 34.139 et 34.140, nous allons résoudre le système

$$\begin{pmatrix} 10^{-9} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{34.332}$$

d'abord de façon exacte, et ensuite en supposant une machine ne tenant que 8 chiffres significatifs en utilisant la méthode de Gauss avec ou sans pivot.

Commençons par voir comment se passe en pratique la décomposition $PA = LU$ de Gauss avec pivot partiel.

Exemple 34.135.
Décomposons la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 0 \\ 3 & 8 & 0 \end{pmatrix}. \tag{34.333}$$

Sur la première colonne, le plus grand nombre est 3. Nous commençons par permuter la première et la troisième ligne en utilisant la matrice de permutation $P_1 = P_{3,1}$ et nous enregistrons $r_1 = 3$. Nous avons alors la matrice

$$A_0' = \begin{pmatrix} 3 & 8 & 0 \\ 2 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}. \tag{34.334}$$

Pour trouver la matrice A_1 nous suivons l'équation (34.326). Bien que le résultat net soit des combinaisons de lignes : $L_2 \to L_2 - 2L_1/3$ et $L_3 \to L_3 - L_1/3$ (que nous pourrions savoir dès à présent), il est important de passer par la matrice gaussienne pour obtenir la matrice L_1.

La matrice de transformation gaussienne pour la première colonne de (34.334) est :

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \tag{34.335}$$

et $L_1 = M_1^{-1}$. Le lemme 34.124 nous dit comment calculer facilement cet inverse :

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{pmatrix} \tag{34.336}$$

En suivant l'équation (34.326) nous posons $A_1 = M_1 A'_0$:

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 8 & 0 \\ 2 & 5 & 0 \\ 1 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 8 & 0 \\ 0 & -1/3 & 0 \\ 0 & -2/3 & 3 \end{pmatrix} \tag{34.337}$$

EQooUBRPooRJbCYo

et nous avons

$$Q_1 A = L_1 A_1 \tag{34.338}$$

où L_1, A_1 et $r_1 = 3$ sont enregistrés. La matrice Q_1 peut être retrouvée en sachant r_1 parce que P est la matrice de permutation $P_{r_1,1}$.

Nous travaillons maintenant sur la deuxième colonne de A_1. Le plus grand élément en valeur absolue (sur ou sous la diagonale) est $-2/3$. Nous posons $r_2 = 3$ et

$$A'_1 = \begin{pmatrix} 3 & 8 & 0 \\ 0 & -2/3 & 3 \\ 0 & -1/3 & 0 \end{pmatrix} \tag{34.339}$$

et la matrice gaussienne pour la deuxième colonne est

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{pmatrix} \tag{34.340}$$

Le $-1/2$ provient du calcul $-\big((-1/3)/(-2/3)\big)$. L'inverse de cette matrice est facile :

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix} \tag{34.341}$$

et la matrice suivante à enregistrer est

$$A_2 = M_2 P_{3,2} A_1 = M_2 A'_1 = \begin{pmatrix} 3 & 8 & 0 \\ 0 & -2/3 & 3 \\ 0 & 0 & -3/2 \end{pmatrix}. \tag{34.342}$$

Notons toutefois que pour calculer cette matrice, seul le dernier élément demande un calcul. La première colonne ne change pas (par construction), la seconde gagne un zéro en dernière ligne (la matrice M_2 sert à ça) et sur la dernière colonne, seule la dernière ligne est sujette à changement.

Avec la matrice A_2, la trigonalisation supérieure est faite. La décomposition n'est cependant pas terminée. Nous devons encore trouver la partie triangulaire inférieure. Nous en sommes à

$$Q_1 A = L_1 A_1 = L_1 P_{3,2} L_2 A \tag{34.343}$$

où Q_1 est la première matrice de permutation.

Utilisant le lemme 34.133, il est facile de permuter L_1 avec $P_{3,2}$:

$$L_1P_{3,2} = P_{3,2}\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & 0 & 1 \end{pmatrix}}_{L_1'} \tag{34.344}$$

Nous avons donc

$$P_{3,2}P_{3,1}A = L_1'L_2A \tag{34.345}$$

Deux multiplications matricielles plus tard nous terminons :

$$PA = LU \tag{34.346}$$

avec

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & 1/2 & 1 \end{pmatrix}, \quad U = A_2 = \begin{pmatrix} 3 & 8 & 0 \\ 0 & -2/3 & 3 \\ 0 & 0 & -3/2 \end{pmatrix}. \tag{34.347}$$

△

Notons que Sage utilise la méthode de Gauss avec pivots :

```
sage: A=matrix([ [1,2,3],[2,5,0],[3,8,0] ])
sage: A
[1 2 3]
[2 5 0]
[3 8 0]
sage: A.LU()
(
[0 1 0]  [  1   0   0]  [   3    8    0]
[0 0 1]  [1/3   1   0]  [   0 -2/3    3]
[1 0 0], [2/3 1/2   1], [   0    0 -3/2]
)
```

tex/sage/sageSnip006.sage

Mais attention : Sage crée une décomposition $A = PLU$ et non $PA = LU$. D'où le fait que la matrice de permutations de Sage est l'inverse de celle donnée ici.

34.18.3 D'un point de vue algorithmique

¡¡ Avertissement/question à la lectrice !! 34.136

Je ne suis pas certain de l'optimalité de ce que je raconte ici. Je décris simplement ce que j'ai fait pour écrire mon programme finitediff.

Si vous êtes expert en calcul numérique, n'hésitez pas à donner votre avis.

Nous décrivons à présent la décomposition $A = PLU$ (du théorème 34.134, avec le P à droite). En suivant l'exemple 34.135 nous voyons assez bien comment créer les matrices U et P au fur et à mesure. La construction de L est peut-être moins évidente.

Écrivons un exemple très explicite pour

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 3 & 10 \\ -2 & 1 & 7 \end{pmatrix}. \tag{34.348}$$

Nous commençons par permuter des lignes pour avoir un grand pivot :

$$P_{12}A = \begin{pmatrix} 4 & 3 & 10 \\ 2 & 1 & 3 \\ -2 & 1 & 7 \end{pmatrix}. \tag{34.349}$$

Et nous effectuons l'élimination avec la matrice

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 1/2 & 0 & 1 \end{pmatrix}. \tag{34.350}$$

Cela donne le premier résultat :

$$M_1P_{12}A = \begin{pmatrix} 4 & 3 & 10 \\ 0 & -1/2 & -2 \\ 0 & 5/2 & 12 \end{pmatrix} \tag{34.351}$$

EQooKTBLooHeQkgk

Nous continuons avec P_{23} pour avoir un nouveau grand pivot :

$$P_{23}M_1P_{12}A = \begin{pmatrix} 4 & 3 & 10 \\ 0 & 5/2 & 12 \\ 0 & -1/2 & -2 \end{pmatrix}. \tag{34.352}$$

Nous utilisons la matrice

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/5 & 1 \end{pmatrix} \tag{34.353}$$

et au final :

$$M_2P_{23}M_1P_{12}A = \begin{pmatrix} 4 & 3 & 10 \\ 0 & 5/2 & 12 \\ 0 & 0 & 2/5 \end{pmatrix} = U. \tag{34.354}$$

L'égalité obtenue est

$$M_2P_{23}M_1P_{12}A = U. \tag{34.355}$$

Pour avoir la décomposition PLU il faut écrire

$$A = P_{12}M_1^{-1}P_{23}M_2^{-1}U, \tag{34.356}$$

EQooWKUYooUBQYtc

et permuter P_{23} avec M_1^{-1}, ce qui est facile par le lemme 34.133.

Remarque 34.137.
Nous ne devons permuter la matrice M_k avec une matrice de permutations qu'à partir de la deuxième étape. En effet l'équation (34.351) revient à

$$A = M_1^{-1}P_{12}m_U \tag{34.357}$$

qui est dans le bon ordre. Ce n'est qu'à partir de la seconde étape que des matrices de permutations apparaissent à droite des matrices gaussiennes.

Cependant dans un cas 4×4, cette méthode deviendrait fastidieuse parce que nous aurions encore des étapes à faire. En repartant de (34.356), mais avec m_U (la matrice pas encore tout à fait triangularisée) au lieu de U, nous aurons, pour un certain $k > 3$:

$$M_3P_{3k}M_2P_{23}M_1P_{12}A = U, \tag{34.358}$$

ce qui fait :

$$A = P_{12}M_1^{-1}P_{23}M_2^{-1}P_{3k}M_3^{-1}U. \tag{34.359}$$

Tous les P_{ij} peuvent être mis à gauche parce que leurs indices sont toujours strictement supérieurs à ceux des M_l placés devant eux. Mais c'est fastidieux.

Nous allons donc permuter à chaque étape pour ne retenir que l'important. Si à une certaine étape nous avons

$$A = P_{1,r_1} \dots P_{k,r_k} M_1^{-1} \dots M_k^{-1} m_U \tag{34.360}$$

avec

$$m_U = P_{k+1,r_{k+1}} M_{k+1}^{-1} U \tag{34.361}$$

alors nous allons directement permuter $P_{k+1,r_{k+1}}$ avec tous les M_i^{-1}. Si nous notons P_k la permutation (pas élémentaire) à l'étape k et L_k la matrice triangulaire inférieure à de l'étape k,

$$A = P_k L_k m_U = P_k L_k P_{k+1,r_{k+1}} M_{k+1}^{-1} m'_U. \tag{34.362}$$

Nous enregistrons alors $P_{k+1} = P_k P_{k+1,r_{k+1}}$ et pour L_{k+1} nous partons de L_k et nous faisons deux opérations suivantes :

— nous permutons, sur ses colonnes non triviales, les indices $k+1$ et r_{k+1},

— nous multiplions par M_{k+1}^{-1}, ce qui revient à simplement lui ajouter une colonne non triviale.

Notons que $r_{k+1} \geqslant k+1$, de telle sorte que sur les colonnes non triviales (qui sont jusqu'au numéro k), la permutation des lignes $k+1$ et r_{k+1} ne change pas l'aspect de la matrice : elle reste multi-gaussienne de dernière colonne k.

34.18.4 Exemples

Nous nous lançons dans la résolution du système

$$\begin{pmatrix} 10^{-9} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{34.363}$$

ExooNTECooXvTcoh

Exemple 34.138.
Nous commençons de façon exacte, par la méthode de Gauss sans pivot. La première transformation gaussienne est

$$E_1 = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix} \tag{34.364}$$

et nous calculons

$$E_1 A = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix} \begin{pmatrix} 10^{-9} & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 10^{-9} & 1 \\ 0 & 1 - 10^{-9} \end{pmatrix}. \tag{34.365}$$

Vu que cette dernière est triangulaire supérieure, nous avons fini la méthode de Gauss et $U = E_1 A$. En ce qui concerne la matrice L, elle est donnée par $L = E_1^{-1}$, c'est-à-dire

$$L = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 10^9 & 1 \end{pmatrix}. \tag{34.366}$$

Au final nous avons la décomposition $A = LU$ exacte suivante :

$$L = \begin{pmatrix} 1 & 0 \\ 10^9 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 10^{-9} & 1 \\ 0 & 1 - 10^9 \end{pmatrix}. \tag{34.367}$$

Résoudre le système $Ax = b$ revient à résoudre $LUx = b$ et donc résoudre successivement les systèmes

$$\begin{cases} Ly = b & \text{(34.368a)} \\ Ux = y. & \text{(34.368b)} \end{cases}$$

D'abord le système

$$\begin{pmatrix} 1 & 0 \\ 10^9 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{34.369}$$

donne $y_1 = 1$ et $y_2 = 2 - 10^9$.

Ensuite nous résolvons

$$\begin{pmatrix} 10^{-9} & 1 \\ 0 & 1 - 10^9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 - 10^9 \end{pmatrix}. \tag{34.370}$$

Cela donne

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{10^9}{1-10^9} \\ \frac{2-10^9}{1-10^9} \end{pmatrix} \simeq \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{34.371}$$

C'est également le résultat que trouve Sage :

```
sage: var('y')
y
sage: solve(    [10**(-9)*x+y==1,x+y==2],[x,y]    )
[[x == (1000000000/999999999), y == (999999998/999999999)]]
```

tex/sage/sageSnip007.sage

△

EXooNVRNooJgQmQc

Exemple 34.139 ([756])**.**
Nous recommençons tout le calcul avec une précision limitée à 8 chiffres significatifs, sans pivot.

Nous avons à nouveau la transformation gaussienne

$$E_1 = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix}, \tag{34.372}$$

mais pour calculer U nous effectuons le produit matriciel

$$U = E_1 A = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix} \begin{pmatrix} 10^{-9} & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 10^{-9} & 1 \\ 0 & * \end{pmatrix}. \tag{34.373}$$

Nous détaillons à présent le calcul de l'élément noté $*$. Le calcul de $10^9 \ominus 1$ donne

$$999999999 = 9.99999999 \times 10^8, \tag{34.374}$$

mais la précision étant limitée à 8 chiffres, un arrondi arrive. Étant donné que le premier chiffres supprimé est un 9 nous retombons sur 10^9, et donc notre machine à précision limitée donnera

$$U = \begin{pmatrix} 10^{-9} & 1 \\ 0 & -10^9 \end{pmatrix}. \tag{34.375}$$

Ensuite le calcul de $L = E_1^{-1}$ ne cause pas de problèmes :

$$L = \begin{pmatrix} 1 & 0 \\ -10^9 & 1 \end{pmatrix}. \tag{34.376}$$

Maintenant il s'agit de résoudre les systèmes $Ly = b$ et $Ux = y$. Du système

$$\begin{pmatrix} 1 & 0 \\ 10^9 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{34.377}$$

nous tirons tout de suite $y_1 = 1$ et ensuite $10^9 + y_2 = 2$, c'est-à-dire $y_2 = 2 - 10^9$, qui en précision limitée donne encore $y_2 = -10^9$. À résoudre maintenant :

$$\begin{pmatrix} 10^{-9} & 1 \\ 0 & -10^9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -10^9 \end{pmatrix}. \tag{34.378}$$

Cela donne immédiatement $x_2 = 1$ et ensuite

$$10^{-9}x_1 + 1 = 1, \tag{34.379}$$

donc $x_1 = 0$. La solution trouvée est

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{34.380}$$

EQooBGWEooVGSVoe

qui est complètement faux au niveau de la première variable. △

EXooNCRSooTfmPFr

Exemple 34.140 ([756]).
Nous résolvons encore le même système en précision limitée, mais en utilisant cette fois la méthode de Gauss avec pivot partiel.

Le plus grand élément de la première colonne est 1 ; nous utilisons donc la permutation $P_{1,2}$:

$$P_{1,2}A = \begin{pmatrix} 1 & 1 \\ 10^{-9} & 1 \end{pmatrix}. \tag{34.381}$$

La matrice de transformation gaussienne pour la première colonne de cette matrice est

$$M_1 = \begin{pmatrix} 1 & 0 \\ -10^{-9} & 1 \end{pmatrix} \tag{34.382}$$

et nous posons

$$A_1 = M_1P_{1,2}A = \begin{pmatrix} 1 & 0 \\ -10^{-9} & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 10^{-9} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -10^{-9}+1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = U \tag{34.383}$$

où un arrondi a eu lieu pour $-10^{-9} + 1 = 1$. En inversant M_1 nous avons

$$L_1 = M_1^{-1} = \begin{pmatrix} 1 & 0 \\ 10^{-9} & 1 \end{pmatrix}. \tag{34.384}$$

La décomposition est

$$A = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{P}\underbrace{\begin{pmatrix} 1 & 0 \\ 10^{-9} & 1 \end{pmatrix}}_{L}\underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{U} \tag{34.385}$$

Le moment de résoudre est venu. Vu que $PLUx = b$ nous devons résoudre les systèmes

$$\begin{cases} Pz = b & \text{(34.386a)} \\ Ly = z & \text{(34.386b)} \\ Ux = y. & \text{(34.386c)} \end{cases}$$

Pour z c'est facile :

$$z = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{34.387}$$

Pour y il y a un arrondi :

$$\begin{pmatrix} 1 & 0 \\ 10^{-9} & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{34.388}$$

Tout de suite : $y_1 = 2$ et ensuite $2 \times 10^{-9} + y_2 = 1$, ce qui donne $y_2 = 1 \ominus 2 \times 10^{-9} = 1$. Donc

$$y = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{34.389}$$

Et enfin pour x c'est le système

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{34.390}$$

Nous avons $x_2 = 1$ et ensuite $x_1 + 1 = 2$ c'est-à-dire $x_1 = 1$. Au final la solution trouvée est

$$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{34.391}$$

Cette solution est considérablement meilleure que (34.380). △

34.141.
L'utilisation du pivot non seulement assure le fait que la trigonalisation va bien se passer (on évite les zéros en pivot), mais aussi et surtout, en choisissant de prendre le plus grand pivot possible, nous obtenons une meilleur stabilité numérique.

34.19 Résolution de systèmes linéaires (suite)

34.19.1 Déterminant

Pour calculer un déterminant lorsque nous avons la décomposition $A = LU$ nous pouvons faire

$$\det(A) = \det(LU) = \det(L)\det(U) = \det(U) \tag{34.392}$$

parce que L est triangulaire avec des 1 sur la diagonale.

Si par contre nous avons fait des pivots, nous avons $PA = LU$. Il nous faut le déterminant de P, qui n'est autre que ± 1. Nous avons

$$\det(P) = (-1)^s \tag{34.393}$$

où s est le nombre de permutations effectives effectuées. Nous précisons « effectives » parce qu'il ne faut pas compter le pas où nous n'avons pas permuté (les cas où le bon pivot était présent du premier coup). Nous avons alors

$$\det(A) = (-1)^s \det(U). \tag{34.394}$$

34.19.2 Plusieurs termes indépendants

Mettons un système $Ax = b$ qu'il faut résoudre pour plusieurs b différents. C'est typiquement le cas où l'on voudrait calculer l'inverse de A. Mais on va directement se calmer. Soient donc à résoudre $Ax_1 = b_1$, ..., $Ax_n = b_n$.

Les opérations (avec ou sans pivot) que nous faisons ne dépendent que de la matrice A, mais aucune décision concernant les pivots ou la matrice des multiplicateurs ne dépend de b. Autre façon de dire : si le système $(A|b_1)$ devient $(U|y_1)$, le système $(A|b_i)$ devient $(U|y_i)$ avec le même U.

Nous ne sommes donc pas obligés de faire tout le travail autant de fois qu'il n'y a de systèmes à résoudre. Donc si on a plusieurs systèmes à résoudre avec la même matrice, on fait mieux de retenir une fois pour toute la décomposition LU (avec ou sans pivots), avant de vraiment résoudre.

Ou alors on peut aussi faire que, au lieu de faire $(A|b_i)$ plein de fois, faire une seule fois

$$(A|b_1 \dots b_n). \tag{34.395}$$

Et on fait tout le travail sur tous les vecteurs d'un en même temps.

Soit e_i la base canonique. Si nous notons x_n les solutions des problèmes $Ax_i = e_i$, tous les problèmes $Ax_i = e_i$ s'écrivent d'un seul coup

$$AX = Y \tag{34.396}$$

où X est la matrice des x_i en colonnes, et Y est celle des e_i en colonnes. Oh, mais $Y = \mathbb{1}$ évidemment. Donc

$$AX = \mathbb{1}. \tag{34.397}$$

Si nous supposons que A est inversible, alors ce X est l'inverse.

Donc pour calculer l'inverse d'une matrice de dimension non trop grande, il suffit d'utiliser la méthode de Gauss sur les vecteurs de la base canonique. Cette idée est la base du calcul de l'inverse par matrice companion. En effet, si nous partons du problème

$$(A|\mathbb{1}) \tag{34.398}$$

et nous appliquons la méthode de Gauss avec pivot, nous arrivons à

$$(U|L^{-1}P). \tag{34.399}$$

Attention : le produit $L^{-1}P$ est une permutation des *colonnes* de L^{-1}. Vu que L est triangulaire inférieure avec des 1 sur la diagonale, L^{-1} est triangulaire inférieure avec des 1 sur la diagonale. Donc si la matrice n'est pas trop grande, on peut assez facilement remettre les colonnes de $L^{-1}P$ dans l'ordre pour recomposer une matrice triangulaire inférieure avec 1 sur la diagonale.

Une autre façon de calculer l'inverse, si $A = LU$ est connue, il suffit de faire

$$A^{-1} = U^{-1}L^{-1}. \tag{34.400}$$

Et il existe un algorithme facile pour l'inverse d'une matrice triangulaire.

34.19.3 Cholesky

La méthode de Gauss s'applique sans hypothèses sur la matrice A, à part qu'elle doit être de petite dimension, comme pour toute méthode directe. Souvent nous savons des choses sur la matrice. Ici nous allons supposer que A est symétrique et définie positive.

Comment numériquement vérifier ces hypothèses ? En ce qui concerne la symétrie, il suffit de faire le test complet :

$$A^t = A. \tag{34.401}$$

La vérification de cela coûte au maximum n^2 comparaisons (et en fait la moitié de ça moins la diagonale).

Le fait que A soit définie positive est facile à vérifier pour utiliser Cholesky parce que il suffit de le faire, et si il n'y a pas de nombres complexes qui arrivent, c'est que la matrice était définie positive.

Un lemme très simple à mettre en oeuvre numériquement nous permet de traiter certains cas.

Lemme 34.142.
Une matrice symétrique possédant un élément négatif sur la diagonale n'est pas définie positive.

Démonstration. Un simple calcul ou effort d'imagination montre que $\langle Me_k, e_k\rangle = M_{kk}$. Donc si M doit être définie positive, M_{kk} doit être positive par le lemme 8.234. □

Ce lemme est un moyen déjà de faire quelques vérifications. Et si les éléments diagonaux de A sont tous négatifs, on peut prendre $-A$.

LEMooVEIYooZbShQb

Lemme 34.143.
Si A est une matrice symétrique strictement définie positive, alors pour tout k, la matrice tronquée $\Delta_k(A)$ l'est également.

Démonstration. Le fait que $\Delta_k(A)$ soit symétrique est évident. Le fait qu'elle soit définie positive l'est moins. Soit $y \in \mathbb{R}^k$ et le vecteur $\tau y \in \mathbb{R}^n$, qui est « complété » avec des zéros.

Nous avons $\langle \Delta_k(A)y, y\rangle_k = \langle A\tau y, \tau y\rangle_n$. En effet

$$\langle \Delta_k(A)y, y\rangle = \sum_{i=1}^{k}\sum_{l=1}^{k} A_{il}y_l y_i. \tag{34.402}$$

Et à droite :

$$\langle A\tau y, \tau y\rangle = \sum_{i=1}^{n}(A\tau y)_i(\tau y)_i = \sum_{i=1}^{k}(A\tau y)_i y_i = \sum_{i=1}^{k}\sum_{l=1}^{n} A_{il}(\tau y)_l y_i = \sum_{i=1}^{k}\sum_{l=1}^{k} A_{il}y_l y_i \tag{34.403}$$

où nous avons utilisé le fait que $(\tau y)_i = 0$ dès que $i > k$ et que $(\tau y)_i = y_i$ sinon.

En conséquence de quoi $\langle \Delta_k(A)y, u\rangle > 0$ pour tout $y \in \mathbb{R}^k$ et la matrice $\Delta_k(A)$ est strictement définie positive. □

LEMooLBQLooIYvacH

Lemme 34.144.
Si T est une matrice triangulaire, alors $(T_{ii})^{-1} = (T^{-1})_{ii}$.

Démonstration. Il suffit de se rendre compte que le coefficient ii de l'égalité $\mathbb{1} = TT^{-1}$ donne

$$1 = \sum_l T_{il}(T^{-1})_{li}. \tag{34.404}$$

Dans la somme il ne reste que le terme $l = i$. □

Nous allons chercher une décomposition de type LU sous la forme $A = LL^t$, c'est-à-dire $U = L^t$. Attention : maintenant nous n'avons plus des 1 sur la diagonale. Ce n'est donc pas exactement la décomposition LU dont nous parlions plus haut. C'est pour cela que nous n'allons pas la noter LL^t mais BB^t.

Théorème 34.145 (Cholesky[755])**.**
Soit une matrice réelle symétrique strictement définie positive. Il existe une unique matrice réelle B telle que

- *B est triangulaire inférieure,*
- *la diagonale de B est positive,*
- *$A = BB^t$.*

Démonstration. Par la décomposition LU du théorème 34.126 nous avons des matrices L et U telles que $A = LU$. Soit D la matrice diagonale donnée par

$$D_{ii} = \sqrt{U_{ii}}. \tag{34.405}$$

Cette définition fonctionne parce que $U_{ii} > 0$. En effet nous savons que $\Delta_k(A) = \Delta_k(L)\Delta_k(U)$, et en passant au déterminant,

$$\det\big(\Delta_k(A)\big) = \det\big(\Delta_k(U)\big). \tag{34.406}$$

Vu que $\Delta_k(A)$ est strictement définie positive par le lemme 34.143, son determinant est strictement positif[22] et nous avons

$$\det\big(\Delta_k(U)\big) > 0. \tag{34.407}$$

En appliquant cela à $k = 1$ nous avons $U_{11} > 0$ puis de proche en proche, $U_i i > 0$ pour tout i.

Nous posons :

$$B = LD \qquad \text{qui est triangulaire inférieure} \tag{34.408a}$$

$$C = D^{-1}U \qquad \text{qui est triangulaire supérieure.} \tag{34.408b}$$

Nous avons bien entendu $A = BC$ et nous allons prouver que $C = B^t$. Vu que $A = A^t$ nous pouvons identifier BC et C^tB^t :

$$BC = C^tB^t. \tag{34.409}$$

En mettant les matrices triangulaires supérieures à gauche et inférieures à droite :

$$C(B^t)^{-1} = B^{-1}C^t, \tag{34.410}$$

qui sont donc deux matrices diagonales. Nous montrons que cette diagonale est en réalité l'identité.

D'abord

$$B_{ii} = \sum_{l=1}^{n} L_{il}D_{li} = L_{ii}\sqrt{U_{ii}} = \sqrt{U_{ii}} \tag{34.411}$$

parce que $L_{ii} = 1$. Notons en passant que la diagonale de B est positive. Ensuite

$$C_{ii} = \sum_{l=1}^{n}(D^{-1})_{il}U_{li} = (D^{-1})_{ii}U_{ii} = \frac{1}{\sqrt{U_i i}}U_{ii} = \sqrt{U_{ii}}. \tag{34.412}$$

22. Le théorème 8.227 donne une diagonalisation par des matrices de déterminant 1. Vu que les valeurs propres forment sur la diagonale, et qu'elles sont toutes positives, el déterminant est positif.

Donc B et C ont des diagonales égales. Calculons alors la diagonale de $B^{-1}C^t$:

$$\left(B^{-1}C^t\right)_{ii} = \sum_l (B^{-1})_{il}(C^t)_{li} = (B^{-1})_{ii}C_{ii} \tag{34.413}$$

parce que encore une fois, de la somme il ne reste que le terme $l = i$.

Mais B est une matrice triangulaire qui tombe sous le coup du lemme 34.144. Donc $(B^{-1})_{ii} = (B_{ii})^{-1} = (C_{ii})^{-1}$. Nous avons alors

$$(B^{-1}C^t)_{ii} = 1. \tag{34.414}$$

Cela conclut l'existence de la décomposition de Cholesky.

En ce qui concerne l'unicité, soient $A = BB^t = CC^t$. Nous regroupons les supérieures et les inférieures :

$$B^t(C^t)^{-1} = B^{-1}C. \tag{34.415}$$

Ces deux matrices sont donc diagonales et nous posons $D = B^{-1}C$, c'est-à-dire $C = BD$. Nous remplaçons donc C par BD dans (34.415) :

$$A = BB^t = BD(BD)^t = BDD^tB^t. \tag{34.416}$$

Donc $DD^t = \mathbb{1}$, ce qui signifie que les éléments diagonaux de D sont ± 1. Nous montrons qu'ils sont positifs : à partir de $C = BD$ nous déballons

$$C_{ii} = \sum_l B_{il}D_{li}, \tag{34.417}$$

et donc

$$B_{ii}D_{ii} = C_{ii}. \tag{34.418}$$

En sachant que les conditions de la décomposition de Cholesky demandent les éléments diagonaux positifs nous en déduisons que D_{ii} est positif et donc égal à 1. Finalement $D = \mathbb{1}$ et $B = C$. □

Prenons la matrice

$$A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & -7 \\ -2 & -7 & 9 \end{pmatrix} \tag{34.419}$$

Elle est symétrique et définie positive. Nous posons

$$\begin{cases} l_{11} = \sqrt{a_{11}} & \text{(34.420a)} \\ l_{i1} = a_{i1}/l_{11} & \text{(34.420b)} \end{cases}$$

pour $i = 2, \ldots, n$. Et aussi

$$\begin{cases} l_{jj} = (a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2)^{1/2} & \text{(34.421a)} \\ l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk})/a_{jj} & \text{(34.421b)} \end{cases}$$

pour $i = j + 1, \ldots, n$.

Les formules (34.420) nous disent comment remplir la première colonne. Cela donne la matrice

$$L = \begin{pmatrix} \sqrt{4} = 2 & . & . \\ 2/2 = 1 & . & . \\ -2/2 = -1 & . & . \end{pmatrix} \tag{34.422}$$

Les formules (34.421) donnent les autres colonnes en fonction des précédentes.

Dans Sage :

```
sage: A=matrix(  [  [4,2,-2],[2,10,-7],[-2,-7,9]  ]   )
sage: A
[ 4  2 -2]
[ 2 10 -7]
[-2 -7  9]
sage: A.cholesky()
[ 2  0  0]
[ 1  3  0]
[-1 -2  2]
```

tex/sage/sageSnip005.sage

34.20 Système linéaire (méthodes itératives)

Nous trouvons des méthodes itératives lorsque les matrices sont grandes, ce qui arrive lorsque l'on discrétise une équation différentielle.

Nous allons chercher des méthodes de la forme $x_{n+1} = Bx_n + q$; ce sont des méthodes stationnaires. La convergence d'une méthode est toujours liée à la matrice B et en général, la convergence ne dépend pas du choix du vecteur initial. Nous faisons donc souvent $x_0 = 0$ et donc $x_1 = q$. Voilà donc une itération de faite gratuitement.

Nous notons e_k le **vecteur d'erreur** qui est défini par $e_k = x - x_k$. Et le **vecteur résidu** $r_k = q - Bx_k$. Attention : ici k n'est pas un indice mais un numéro de vecteur.

Notons que si x est solution, alors $q - Bx = 0$, ce qui motive le vecteur résidu.

Les conditions d'arrêt d'un algorithme seraient

$$\begin{cases} \|e_k\|_\infty \ll \epsilon_1 & \text{(34.423a)} \\ \|r_k\|_\infty < \epsilon_2 & \text{(34.423b)} \end{cases}$$

où ϵ_1 et ϵ_2 sont des précisions décidées à l'avance par l'utilisateur.

Proposition 34.146.
Si A est une matrice inversible, alors

$$\lim_{k\to\infty} e_k = \lim_{k\to\infty} r_k = 0. \tag{34.424}$$

Vu que $r_k = Ae_k$, si la matrice A est mal conditionnée, il peut arriver que r_k reste grand alors que e_k est déjà petit.

Remarque 34.147.
Dans les méthode stationnaires, nous avons $x_{n+1} = Bx_n + q$ avec B et q fixés au départ de l'algorithme. Il existe des méthodes non stationnaires pour lesquelles l'itération prend la forme $x_{n+1} = B_n x_n + q_n$ avec B_n et q_n qui changement avec les étapes.

PROPooAQSWooSTXDCO

Proposition 34.148.
Pour la méthode $x_{n+1} = Bx_n + q$ nous avons équivalence de

(1) La méthode converge pour tout x_0

(2) B est une matrice convergente[23]

(3) $\rho(B) < 1$ (rayon spectral).

De plus si $\|B\| < 1$ alors la méthode converge (quelle que soit la norme algébrique).

La norme d'une matrice (en tout cas, certaines normes) est quelque chose de facile à calculer à l'ordinateur. Typiquement $\|.\|_\infty$ est un simple maximum. Cependant si après avoir calculé $\|B\|_i$ pour des dizaines de normes i différentes, nous avons toujours $\|B\|_i \geqslant 1$, alors nous ne pouvons rien conclure.

23. C'est-à-dire $\lim_{k\to\infty} B^k = 0$.

34.20.1 La méthode générale

Nous décomposons la matrice A sous la forme $A = M - N$ avec M inversible. Le système $Ax = b$ devient

$$Mx - Nx = b \tag{34.425}$$

puis $Mx = Nx + b$ et finalement

$$x = M^{-1}Nx + M^{-1}b, \tag{34.426}$$

et voilà une méthode stationnaire avec $B = M^{-1}N$ et $q = M^{-1}b$.

Mais ici nous voyons que M doit être non seulement inversible, mais en plus doit être facilement calculable. En sachant que nous travaillons avec des grandes matrices, il n'est pas question d'inverser M avec une méthode de Gauss.

En bref, il faut choisir M triangulaire parce que c'est en gros la seule que nous pouvons inverser facilement [24].

Remarque 34.149.
La matrice B ne doit pas spécialement être inversible. Si elle ne l'est pas, ce n'est pas un problème.

34.20.2 Jacobi

Nous décomposons

$$A = D - E - F \tag{34.427}$$

EQooOCJYooCqsfQM

où D est la diagonale de A, $-F$ est la partie triangulaire supérieure (sans la diagonale) et $-E$ la triangulaire inférieure (sans la diagonale). Donc D, E et F sont simplement des extractions de parties de la matrice A (et quelques changements de signes).

La méthode de Jacobi prend $M = D$ et $N = (E + F)$. L'inverse de M est facile à calculer parce que M est diagonale. Nous notons B_J la matrice B de la méthode de Jacobi.

Remarque 34.150.
Il se peut que la matrice A ait des zéros sur la diagonale, même si elle est inversible. Et cela est un problème parce qu'alors la matrice D ici construite n'est pas inversible. Dans ce cas, avant de nous lancer dans la méthode de Jacobi, il faut permuter deux lignes de A et donc de b.

Attention cependant que l'on pourrait vouloir effectuer ces permutations en mettant sur la diagonale des nombres les plus grands possibles (parce qu'ensuite, ce qui rentre dans les calculs, c'est D^{-1} qui aura alors des petits nombres). Mais il faut toutefois faire en sorte que le rayon spectral de la matrice B résultante reste plus petit que 1.

Chaque changement dans A induit des changements dans B et donc sur la convergence de la méthode.

34.20.3 Gauss-Seidel

Nous partons de la même décomposition $A = D - E - F$ que dans (34.427). La méthode de Gauss-Seidel prend $M = (D - E)$ et $N = F$.

34.20.4 Autres

Voir la méthode des gradients, et des gradients conjugués.

34.21 Indices connectés, matrice irréductible

Définition 34.151.
*Soit $A \in \mathbb{M}(n, \mathbb{R})$ et $i, j \in \{1, \dots, n\}$. Nous disons que les indices i et j sont **directement connectés** si $A_{ij} \neq 0$ ou $A_{ji} \neq 0$.*

24. Les matrices orthogonales sont aussi facilement inversibles, mais ne se prêtent pas bien à une décomposition de type somme.

DEFooADAAooAAMscc

Définition 34.152.
*Soit $A \in \mathbb{M}(n,\mathbb{R})$ et $i,j \in \{1,\dots,n\}$. Nous disons que les indices i et j sont **connectés** si il existe un ensemble d'indices $i_0 = i, i_1, \dots, i_{r-1}, i_r = j$ tels que $A_{i_k,i_{k+1}} \neq 0$ pour tout $0 \leqslant k \leqslant r$.*

Par exemple pour que les indices 1 et 4 soient connectés, on peut avoir les éléments A_{13}, A_{32} A_{24} non nuls.

DEFooXIREooQtlzkO

Définition 34.153 ([757]).
*Une matrice carrée A est **réductible** si il existe une permutation σ telle que*

$$\sigma^t A \sigma = \begin{pmatrix} K & L \\ 0 & M \end{pmatrix} \tag{34.428}$$

EQooGGZKooUyXSJk

où K et M sont carrées.

Notons que par définition de la matrice d'une application linéaire,

$$B_{ij} = \langle e_i, Be_j \rangle = \langle e_i, \sigma^t A \sigma e_j \rangle = \langle \sigma e_i, A \sigma e_j \rangle = A_{\sigma(i),\sigma(j)}. \tag{34.429}$$

PROPooZTYDooZAxQxF

Proposition 34.154 ([757]).
Soit une matrice carrée A. Les faits suivants sont équivalents :

ITEMooYULAooVqgOnt

(1) A est réductible.

ITEMooNLVXooYSQKwO

(2) Il existe une partition non triviale I, J de $\{1,\dots,n\}$ telle que $I \cup J = \{1,\dots,n\}$, $I \cap J = \varnothing$ et pour tout $i \in I$, et pour tout $j \in J$, $A_{ij} = 0$.

ITEMooVNOHooRUNpwG

(3) La matrice A admet des indices non connectés (définition 34.152).

Démonstration. Dans plusieurs sens...

(1) **(1) implique (2)**
Nous notons j^* la taille de la matrice K dans (34.428). Nous avons $B_{ij} = 0$ si

$$\begin{cases} j^* + 1 \leqslant i \leqslant n & \text{(34.430a)} \\ 1 \leqslant j \leqslant j^*. & \text{(34.430b)} \end{cases}$$

Donc en posant $I = \sigma\{j^*+1,\dots,n\}$ et $J = \sigma\{1,\dots,j^*\}$ nous avons une partition non triviale de $\{1,\dots,n\}$ telle que si $i \in I$ et $j \in J$ alors $i = \sigma(i_0)$, $j = \sigma(j_0)$ et

$$A_{ij} = A_{\sigma(i_0),\sigma(j_0)} = B_{i_0,j_0} = 0. \tag{34.431}$$

(2) **(2) implique (1)** Soit une partition I, J comme indiquée dans l'hypothèse. Soit j^* le nombre d'éléments dans J. Soit σ une permutation de $\{1,\dots,n\}$ telle que $\sigma\{j^*+1,\dots,n\} = I$ et $\sigma\{1,\dots,j^*\} = J$. Nous posons ensuite $B = \sigma^t A \sigma$. Par construction si $i \in I$ et $j \in J$ alors $A_{ij} = 0$.
Mais si

$$\begin{cases} J^* + 1 \leqslant i \leqslant n & \text{(34.432a)} \\ 1 \leqslant j \leqslant j^*. & \text{(34.432b)} \end{cases}$$

alors $B_{ij} = A_{\sigma(i)\sigma(j)} = 0$. Donc B a la bonne forme.

(3) **(3) implique (2)**
Soient i et j deux indices non connectés : il n'existe pas de chaines partant de i et arrivant à j. Nous notons I l'ensemble des indices connectés à i, et J les autres. Pat hypothèses ces ensembles sont non vides.
Si $k \in I$ et $l \in J$ alors $A_{kl} = 0$ parce que sinon on aurait une chaine de i à k puis de k à l et donc de i à l, ce qui signifierait que l est connecté à i.

(4) **(2) implique (3)** Soit une partition I, J comme dans l'hypothèse. Si $j \in J$ est connecté à $i \in I$ alors il existe une chaine

$$i = i_0, i_1, \dots, i_r = j. \tag{34.433}$$

Si i_s est le premier dans J alors $i_{s-1} \in I$ et $A_{i_{s-1},i_s} = 0$, ce qui empêche la chaine de connecter j à i.

□

34.22 Localisation des valeurs propres

Sur l'ensemble $\mathbb{M}(n, \mathbb{R})$ des matrices $n \times n$ à coefficients réels nous introduisons l'ordre partiel [25] donné par $A \geqslant B$ lorsque $A_{ij} \geqslant B_{ij}$ pour tout i et j. Nous définissons de façon similaire les relations $A \leqslant B$, $A < B$ et $A > B$.

Si $x \in \mathbb{R}^n$ nous notons $|x| = \big(|x_1|, \dots, |x_n|\big)$ et $x \leqslant y$ lorsque $x_i \leqslant y_i$ pour tout i.

PROPooGVRVooZEvKcn

Proposition 34.155.
Soit $A \in \mathbb{M}(n, \mathbb{R})$ et $x, y \in \mathbb{R}^n$.

ITEMooXQOPooPVLjFh

(1) Si $A \geqslant 0$ et si $x \leqslant y$ alors $Ax \leqslant Ay$.

ITEMooQLCJooKIbws

(2) Si $A \geqslant 0$ alors $Ax \leqslant |Ax| \leqslant A|x|$.

Démonstration. Pour la première inégalité, pour tout i et k nous avons $A_{ik}x_k \leqslant A_{ik}y_k$ et donc

$$(Ax)_i = \sum_k A_{ik}x_k \leqslant \sum_k A_{ik}y_k = (Ay)_k. \tag{34.434}$$

Pour la seconde, d'abord l'inégalité $Ax \leqslant |Ax|$ est évidente. Ensuite vu que $A_{ik} \geqslant 0$ nous avons

$$|Ax|_i = |\sum_k A_{ik}x_k| \leqslant \sum_k A_{ik}|x_k| = \big(A|x|\big)_i. \tag{34.435}$$

□

Soit une matrice $A \in \mathbb{M}(n, \mathbb{R})$. Nous notons

$$r_i = \sum_{j \neq i} |A_{ij}|. \tag{34.436}$$

Notons la somme sur la ligne i, pas sur la colonne : la somme est horizontale.

Définition 34.156.
Les ensembles

$$D_i = \{z \in \mathbb{C} \text{ tel que } |z - A_{ii}| \leqslant r_i\} \tag{34.437}$$

*sont les **disques de Gershgorin**. Nous allons également noter $B_i = \text{Int}(D_i)$ les boules ouvertes correspondantes.*

THOooUJNFooHpvCCF

Théorème 34.157 (Gershgorin)**.**
Soit $A \in \mathbb{M}(n, \mathbb{R})$. Si $\lambda \in \mathbb{C}$ est valeur propre de A alors $\lambda \in D_i$ pour un certain i.

Démonstration. Soit une valeur propre λ et un de ses vecteurs propres $u \in \mathbb{R}^n$: $Au = \lambda u$ avec $u \neq 0$. Soit i un indice réalisant le maximum $|u_i| = \max\{|u_k|\}_k$. Nous écrivons la i^{e}ligne de $Au = \lambda u$:

$$\sum_k A_{ik}u_k = \lambda u_i, \tag{34.438}$$

25. Définition 1.12.

c'est-à-dire $A_{ii}u_i + \sum_{k\neq i} A_{ik}u_k = \lambda u_i$, ou encore

$$A_{ii} + \sum_{k\neq i} A_{ik}\frac{u_k}{u_i} = \lambda, \tag{34.439}$$

qui donne

$$|A_{ii} - \lambda| \leqslant \sum_{k\neq i} |A_{ik}|\frac{|u_k|}{|u_i|} \leqslant \sum_{k\neq i} |A_{ik}| \tag{34.440}$$

parce que $|u_i| \geqslant |u_k|$. Notons que sur la ligne précédente, $|.|$ est le module dans $\mathbb{C}$, pas la valeur absolue dans $\mathbb{R}$. □

THOooTXAPooQqsBCj

Théorème 34.158 (Gershgorin 2[757])**.**
Soit une matrice irréductible $A \in \mathbb{M}(n,\mathbb{R})$ et une valeur propre λ de A. Si elle est sur la frontière de l'union des disques de Gershgorin, alors elle est sur le bord de tous les disques.

Démonstration. Soit une valeur propre λ de A telle que $\lambda \in \partial\big(\bigcup_i D_i\big)$. Alors λ n'est dans l'intérieur d'aucune boule et nous avons $|\lambda - A_{ii}| \geqslant r_i$ pour tout i.

Soit un vecteur propre u de A tel que $\|u\|_\infty = 1$. Nous posons $I = \{1 \leqslant i \leqslant n \text{ tel que } |u_i| = 1\}$ et $J = \{1 \leqslant j \leqslant n \text{ tel que } |u_j| < 1\}$. Par hypothèse I n'est pas vide, et de plus $I \cap J = \varnothing$ et $I \cup J = \{1,\dots,n\}$ parce qu'aucune composante de u n'a un module[26] plus grand que 1.

La i^{e}composante de la relation $Au = \lambda u$ peut s'écrire

$$(A_{ii} - \lambda)u_i + \sum_{k\neq i} A_{ik}u_k = 0. \tag{34.441}$$

Forts de cela nous écrivons les inégalités suivantes :

$$r_i \leqslant |\lambda - A_{ii}| = \big|(\lambda - A_{ii})u_i\big| = |\sum_{k\neq i} A_{ik}u_k| \leqslant \sum_{k\neq i} |A_{ik}||u_k| \leqslant \sum_{k\neq i} |A_{ik}| = r_i. \tag{34.442}$$

Donc les inégalités sont des égalités :

$$r_i = |\lambda - A_{ii}| = \big|(\lambda - A_{ii})u_i\big| = |\sum_{k\neq i} A_{ik}u_k| = \sum_{k\neq i} |A_{ik}||u_k| = \sum_{k\neq i} |A_{ik}|. \tag{34.443}$$

EQooBIBJooFlscrx

En particulier l'égalité $\sum_{k\neq i} |A_{ik}||u_k| = \sum_{k\neq i} |A_{ik}|$ donne

$$\sum_{k\neq i} |A_{ik}|\big(|u_k| - 1\big) = 0. \tag{34.444}$$

Donc pour tout $k \in J$ nous avons $A_{ik} = 0$. Vu que A est irréductible, cela donnerait une partition impossible $\{1,\dots,n\} = I \cup J$. Nous en déduisons que J est vide et donc que $|u_j| = 1$ pour tout j. En repartant de (34.443) nous avons alors

$$r_i = \big|(\lambda - A_{ii})u_i\big| = |\lambda - A_{ii}||u_i| = |\lambda - A_{ii}|. \tag{34.445}$$

Cela prouve que $\lambda \in \partial D_i$ pour tout i. □

EXooUKQIooQqteHx

Exemple 34.159.
Soit la matrice

$$B = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -1/2 \\ -1 & 0 & 3 \end{pmatrix}. \tag{34.446}$$

D'abord nous rappelons que si vous voulez entrer cette matrice dans Sage (ou plus généralement dans Python2[27]), vous devez faire attention au 1/2 qui, tel quel, est évalué à 0. Nous vous rappelons donc que tous vos codes Sage doivent commencer par ceci :

26. Les composantes de u sont à priori dans $\mathbb{C}$, et non spécialement dans $\mathbb{R}$, même si A est une matrice réelle.
27. Que vous n'avez aucune raison d'utiliser autre que Sage.

```sage
# -*- coding: utf8 -*-
from __future__ import unicode_literals
from __future__ import division
```

tex/sage/sageSnip015.sage

Les éléments non nuls hors diagonale sont B_{13}, B_{31} et B_{23}. Elle n'est donc pas irréductible ; nous avons par exemple la partition $I = \{1, 3\}$, $J = \{2\}$ pour le critère de la proposition 34.154(2).

Les disques de Gershgorin sont

$$D_1 = \{z \in \mathbb{C} \text{ tel que } |z - 2| \leqslant 1\} \tag{34.447a}$$
$$D_2 = \{z \in \mathbb{C} \text{ tel que } |z - 1| \leqslant 1/2\} \tag{34.447b}$$
$$D_3 = \{z \in \mathbb{C} \text{ tel que } |z - 3| \leqslant 1\} \tag{34.447c}$$

Les valeurs propres de la matrice sont sur des bords de disques de Gershgorin, sans être sur tous les bords, comme ça aurait été le cas par le théorème 34.158 si la matrice avait été irréductible. Elles sont sur la figure 34.1 ; notez en particulier les valeurs propres λ_2 et λ_3 qui sont sur le bord de deux disques mais pas sur le bord des trois disques en même temps.

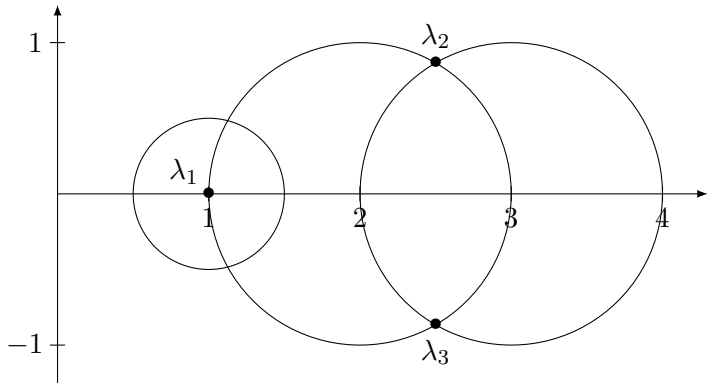


FIGURE 34.1: Les disques de Gershgorin et les valeurs propres pour l'exemple 34.159. LabelHgDNRRoo7WRHgOCw

△

EXooDQYDooPxqHjZ

Exemple 34.160.
Soit la matrice

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 2 \\ 3 & 0 & 2 \end{pmatrix}. \tag{34.448}$$

Nous avons

$$D_1 = \{z \in \mathbb{C} \text{ tel que } |z| \leqslant 1\}, \tag{34.449a}$$
$$D_2 = \{z \in \mathbb{C} \text{ tel que } |z - 1| \leqslant 2\}, \tag{34.449b}$$
$$D_3 = \{z \in \mathbb{C} \text{ tel que } |z - 2| \leqslant 3\}, \tag{34.449c}$$

Le polynôme caractéristique est

$$\chi(\lambda) = -\lambda^3 + 3\lambda^2 - 2\lambda - 6. \tag{34.450}$$

Une fois remarqué que $\lambda_1 = -1$ est une racine, les autres sont faciles à trouver (division euclidienne de $\chi(\lambda)$ par $\lambda + 1$) : $\lambda_2 = 2 + i\sqrt{2}$ et $\lambda_3 = 2 - i\sqrt{2}$.

La matrice A est irréductible. En effet les éléments non diagonaux non nuls sont A_{12}, A_{23} et A_{31}. Ils peuvent former une chaine reliant tous les indices entre eux.

Les contraintes sur la localisation des valeurs propres est donc qu'elles doivent être dans ou sur les disques de Gershgorin, mais que celles qui sont sur le bord d'un disque doivent être sur le

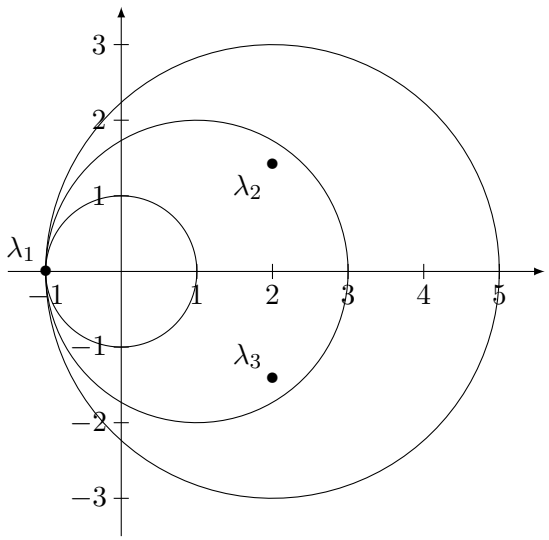


FIGURE 34.2: Les disques de Gershgorin et les valeurs propres pour l'exemple 34.160. LaDeFigDNHRooqGtffLkd

bord de tous les disques en même temps. C'est cela que nous observons sur la figure 34.2. Notez en particulier la position de la valeur propre λ_1.

△

34.22.1 Matrices à diagonale dominante

DEFooLSUTooHuXabV

Définition 34.161 ([758, 757])**.**
*La matrice $A \in \mathbb{M}(n, \mathbb{K})$ est **à diagonale dominante** si pour tout i,*

$$\sum_{j \neq i} |A_{ij}| \leqslant |A_{ii}| \tag{34.451}$$

où $|.|$ est la module dans $\mathbb{C}$ ou la valeur absolue dans $\mathbb{R}$.

*Elle est **à diagonale fortement dominante** si elle est à diagonale dominante et si il existe un i tel que*

$$\sum_{j \neq i} |A_{ij}| < |A_{ii}|. \tag{34.452}$$

*Elle est **à diagonale strictement dominante** si*

$$\sum_{j \neq i} |A_{ij}| < |A_{ii}| \tag{34.453}$$

EQooQLNLooJCLram

pour tout i (entier entre 1 et n).

Nous avons les inclusions suivantes :

$$\text{strictement dominante} \subset \text{fortement dominante} \subset \text{dominante}. \tag{34.454}$$

LEMooMQAEooUCkQxU

Lemme 34.162.
Si A est dans un des deux cas suivant :

- *diagonale strictement dominante,*
- *diagonale dominante et irréductible*[28]

alors $A_{ii} \neq 0$ pour tout i.

28. Définition 34.153.

Démonstration. Si A est à diagonale strictement dominante, alors l'inégalité stricte (34.453) n'est pas possible.

Si A est à diagonale dominante, alors si $A_{ii} = 0$, toute la ligne est nulle. Dans ce cas, la matrice ne peut pas être irréductible. □

Proposition 34.163 ([758]).
Une matrice à diagonale strictement dominante est inversible.

Démonstration. Soit une matrice A à diagonale strictement dominante. Soit x tel que $Ax = 0$. Le but est de montrer que $x = 0$. Soit un indice i_0 réalisant la norme maximum :

$$|x_{i_0}| = \|x\|_\infty. \tag{34.455}$$

Nous écrivons la composante i_0 de l'égalité $Ax = 0$:

$$\sum_k A_{i_0k}x_k = 0, \tag{34.456}$$

et nous séparons le terme $k = i_0$ des autres :

$$\sum_{k\neq i_0} A_{i_0k}x_k + A_{i_0i_0}x_{i_0} = 0. \tag{34.457}$$

Nous prenons le module et majorons les sommes :

$$|A_{i_0i_0}||x_{i_0}| \leqslant \sum_{k\neq i_0} |A_{i_0k}||x_k| \leqslant \sum_{k\neq i_0} |A_{i_0k}||x_{i_0}|. \tag{34.458}$$

Si $|x_{i_0}|$ est non nul nous pouvons simplifier :

$$|A_{i_0i_0}| \leqslant \sum_{k\neq i_0} |A_{i_0k}|. \tag{34.459}$$

Hélas, l'hypothèse de diagonale strictement dominante implique l'inégalité stricte dans le sens inverse. Impossible. Nous en déduisons que $|x_{i_0}| = 0$. Donc $\|x\|_\infty = 0$, ce qui signifie que $x = 0$.

Le fait que le noyau de A se réduise à $\{0\}$ implique l'inversibilité de A. □

PROPooTQWUooSLoniQ

Proposition 34.164.
Soit $A \in \mathbb{M}(n, \mathbb{C})$ une matrice qui est dans un des deux cas suivants :

- *à diagonale strictement dominante*
- *à diagonale dominante et irréductible*

Si $A = D - M$ où D est la diagonale de A (et M est « le reste ») alors D est inversible et

$$\rho(D^{-1}M) < 1 \tag{34.460}$$

où ρ est le rayon spectral (thème 40).

Démonstration. Le lemme 34.162 nous dit que les éléments diagonaux de A sont non nuls. Cela donne déjà le fait que la matrice D est inversible et que la produit $D^{-1}M$ ait un sens. Nous posons $T = D^{-1}M$. Nous avons alors

$$T_{ii} = \sum_k (D^{-1})_{ik}M_{ki}. \tag{34.461}$$

Si $k = i$ alors $M_{ki} = 0$ et si $k \neq i$ alors $D_{ik} = 0$. Donc $T_{ii} = 0$ pour tout i.

En ce qui concerne les autres éléments de T,

$$T_{ij} = \sum_k (D^{-1})_{ik}M_{kj} = \sum_k \frac{1}{A_{ik}}\delta_{ik}M_{kj} = -\frac{A_{ij}}{A_{ii}}. \tag{34.462}$$

Notes :

— Les hypothèses sur A jouent pour dire que $A_{ii} \neq 0$.

— Le signe moins est dû au fait que $M_{ij} = -A_{ij}$ lorsque $i \neq j$.

En faisant la somme des modules :

$$\sum_{j\neq i} |T_{ij}| = \sum_{j\neq i} \frac{|A_{ij}|}{|A_{ii}|} = \frac{1}{|A_{ii}|} \sum_{j\neq i} |A_{ij}| \leqslant 1. \tag{34.463}$$

La dernière inégalité est le fait que A soit à diagonale dominante.

(1) **Si A est à diagonale strictement dominante** Alors nous avons l'inégalité stricte

$$\sum_{j\neq i} |T_{ij}| < 1. \tag{34.464}$$

Et le théorème de Gershgorin 34.157 dit que le spectre de T est contenu dans l'union des disques

$$D_i = \{z \in \mathbb{C} \text{ tel que } |z - T_{ii}| \leqslant r_i\} \tag{34.465}$$

où

$$r_i = \sum_{j\neq i} |T_{ij}|. \tag{34.466}$$

Mais nous avons prouvé que pour tout i, $T_{ii} = 0$ et $\sum_{j\neq i} |T_{ij}| < 1$. Donc toutes ces boules sont contenues dans $B(0,1)$. Cela prouve que $\rho(T) < 1$.

(2) **Diagonale dominante, irréductible**
La matrice T est alors également irréductible parce que les éléments non nuls de A et de T sont les mêmes : $T_{ij} = -A_{ij}/A_{ii}$. Nous utilisons alors le second théorème de Gershgorin 34.158. Si λ est une valeur propre de T, alors soit

$$\lambda \in \bigcup_i B\big(0, r_i\big) \tag{34.467}$$

soit

$$\lambda \in \bigcap_i \partial B(0, r_i). \tag{34.468}$$

Vu que $r_i \leqslant 1$ pour tout i, dans le premier cas λ est dans l'union des boules *ouvertes* de rayon 1. Le nombre λ est donc une la boule ouverte de rayon 1. Bref, $|\lambda| < 1$.
Dans le second cas, l'intersection de deux cercles de même centre sont soit vide soit tout le cercle (auquel cas les rayons sont égaux). Dans le second cas, ledit rayon est certainement strictement plus petit que 1 parce que

$$r_i = \sum_{j\neq i} |T_{ij}| = \sum_{j\neq i} \frac{|A_{ij}|}{|A_{ii}|} < 1. \tag{34.469}$$

□

34.22.2 M-matrice

DEFooZAWWooEAujPy

Définition 34.165.
*Une matrice $A \in \mathbb{M}(n, \mathbb{R})$ est une **M-matrice** si*

(1) $A_{ii} > 0$ pour tout i,

(2) $A_{ij} \leqslant 0$ si $i \neq j$

(3) A est inversible et $A^{-1} \geqslant 0$.

PROPooWVHXooCfsvGq

Proposition 34.166.
Soit $A \in \mathbb{M}(n, \mathbb{R})$ telle que $A_{ii} > 0$ pour tout i et $A_{ij} \leqslant 0$ pour tout $i \neq j$. Nous posons $A = D - M$ où D est la diagonale de A.

La matrice A est une M-matrice si et seulement si $\rho(D^{-1}M) < 1$.

Démonstration. En deux morceaux.

(1) **Si $\rho(D^{-1}M)<1$** Nous posons encore $T=D^{-1}M$. Par le théorème 15.150, la matrice $\mathbb{1}-T$ est inversible et

$$(\mathbb{1}-T)^{-1}=\sum_{k=0}^{\infty}T^k. \tag{34.470}$$

D'autre part, via des calculs déjà faits, et les hypothèses sur les signes des éléments de A,

$$T_{ij}=-\frac{A_{ij}}{A_{ii}}\geqslant 0. \tag{34.471}$$

Donc tous les éléments de T sont positifs (ou nuls). Par conséquent $T^k\geqslant 0$ pour tout k et $(\mathbb{1}-T)^{-1}$ est positive.

Mais $A=D-M=D(\mathbb{1}-D^{-1}M)=D(\mathbb{1}-T)$. Vu que D et $\mathbb{1}-T$ sont inversibles, nous savons que A est inversible et

$$A^{-1}=(\mathbb{1}-T)^{-1}D^{-1}, \tag{34.472}$$

qui est un produit de matrices positives. Donc $A^{-1}\geqslant 0$.

Au final, A est une M-matrice.

(2) **Si A est une M-matrice**

Soit une valeur propre λ de $T=D^{-1}M$ et un vecteur propre u : $Tu=\lambda u$. Vu que $T\geqslant 0$ nous avons d'une part $|\lambda u|=|\lambda||u|$ et d'autre part $|\lambda u|=|Tu|\leqslant T|u|$, ce qui donne

$$|\lambda||u|\leqslant T|u|. \tag{34.473}$$

Dans cette inégalité nous substituons T par $\mathbb{1}-(\mathbb{1}-T)$ pour avoir

$$|\lambda||u|\leqslant |u|-(\mathbb{1}-T)|u| \tag{34.474}$$

ou encore

$$(\mathbb{1}-T)|u|\leqslant \big(1-|\lambda|\big)|u|. \tag{34.475}$$

EQooGFEQooBpiDJR

Mais $(\mathbb{1}-T)^{-1}=A^{-1}D\geqslant 0$ parce que A et D sont positives. Donc en appliquant $(\mathbb{1}-T)^{-1}$ à l'inégalité (34.475), elle est conservée (proposition 34.155(2)) :

$$|u|\leqslant (\mathbb{1}-T)^{-1}\big(1-|\lambda|\big)|u|. \tag{34.476}$$

Si $|\lambda|\geqslant 1$ alors toutes les composantes de $\big(1-|\lambda|\big)|u|$ sont négatives et l'inégalité n'est possible qu'avec $|u|=0$. Dans ce cas, λ n'est pas une valeur propre (le vecteur propre soit être non nul).

Nous en déduisons que $|\lambda|<1$ et donc que $\rho(T)=\rho(D^{-1}M)<1$.

□

Le théorème suivant résume ce que nous avons vu en donnant une condition suffisante facile à vérifier pour être une M-matrice.

THOooLZGSooSevggj

Théorème 34.167.
Soit $A\in\mathbb{M}(n,\mathbb{R})$ telle que

(1) $A_{ii}>0$

(2) $A_{ij}\leqslant 0$ *pour* $i\neq j$

(3) vérifiant une des deux conditions suivantes :

- *à diagonale strictement dominante*
- *à diagonale dominante et irréductible.*

Alors A est une M-matrice.

Démonstration. Au vu de la proposition 34.166, il suffira de montrer que $\rho(D^{-1}M) < 1$ où D et M sont la décomposition $A = D - M$ habituelle. C'est le cas grâce à la proposition 34.164. □

PROPooZDMQooIZAbKK

Proposition 34.168.
Soit $A \in \mathbb{M}(n, \mathbb{R})$, une M-matrice irréductible. Alors $A^{-1} > 0$.

Démonstration. Nous posons $T = D^{-1}M$. En comparant la définition 34.165 de M-matrice et la caractérisation de la proposition 34.166, nous avons $\rho(D^{-1}M) < 1$. Par conséquent

$$(\mathbb{1} - T)^{-1} = \sum_{k=0}^{\infty} T^k \tag{34.477}$$

par la proposition 34.166. D'autre part, $A^{-1} = (\mathbb{1} - T)^{-1}D$ où les éléments D sont strictement positifs. Donc nous devons encore prouver que $(\mathbb{1} - T)^{-1} > 0$. Nous savons que $T \geqslant 0$, et vu que

$$\big(\sum_k T^k\big)_{ij} = \sum_k (T^k)_{ij} \tag{34.478}$$

il nous suffit de prouver que pour chaque (ij), un des $(T^k)_{ij}$ est strictement positif. Soient donc deux indices i et j. Vu que A est irréductible, ils sont connectés par une suite d'indice $i = i_0, i_1, \dots, , i_r = j$ tels que

$$T_{i_k,i_{k+1}} = -\frac{A_{i_k,i_{k+1}}}{A_{i_k,i_k}} > 0. \tag{34.479}$$

Or les indices i_k sont choisis de telle sorte que les numérateurs soient non nuls et donc strictement négatifs. Nous avons, en général :

$$(T^k)_{ij} = \sum_{l_1,\dots,l_{r-1}} T_{i,l_1}T_{l_1,l_2}\dots T_{l_{r-1},j}. \tag{34.480}$$

Chacun des termes est positif ou nul, mais pour $k = r$, il y a entre autres le terme

$$T_{i,i_1}T_{i_1,i_2}\cdots T_{i_r,j} \neq 0. \tag{34.481}$$

Donc $(T^r)_{ij} > 0$ et $\sum_{k=0}^{\infty}(T^k)_{ij} > 0$. Et par conséquent

$$A^{-1} = (\mathbb{1} - T)^{-1}D > 0. \tag{34.482}$$

□

THOooWIFGooBQpddF

Théorème 34.169.
Soit une M-matrice $A \in \mathbb{M}(n, \mathbb{R})$ et $g \in \mathbb{R}^n$ tel que $(Ag)_i \geqslant 1$ pour tout i. Alors $\|A^{-1}\|_\infty \leqslant \|g\|_\infty$.

Démonstration. Nous posons $u = (1, \dots, 1)$ et considérons $x \in \mathbb{R}^n$. Vu que A est une M-matrice, nous avons $A^{-1} \geqslant 0$, donc

$$|A^{-1}x| \leqslant A^{-1}|x| \leqslant \|x\|_\infty A^{-1}u \leqslant \|x\|_\infty g. \tag{34.483}$$

Justifications :

— La première inégalité est la proposition 34.155(2).

— La seconde provient de

$$\big(B|x|\big)_i = \sum_k B_{ik}|x_k| \leqslant \sum_k B_{ik}\|x\|_\infty = \|x\|_\infty \sum_k B_{ik}u_k = \|x\|_\infty Bu. \tag{34.484}$$

— Étant donné que $A^{-1} \geqslant 0$ nous conservons l'inégalité et $Ag \geqslant u$ implique $g \geqslant A^{-1}u$ (c'est la proposition 34.155(1)).

En ce qui concerne la norme de A^{-1} nous avons donc

$$\|A^{-1}\|_\infty = \sup_{\|x\|_\infty=1} \|A^{-1}x\|_\infty \leqslant \sup_{\|x\|_\infty=1} \|x\|_\infty \|g\|_\infty = \|g\|_\infty. \tag{34.485}$$

□

PROPooQBWQooBbeZLO

Proposition 34.170.
Une matrice de $\mathbb{M}(n,\mathbb{R})$ *qui*

(1) est symétrique,

(2) Vérifie une des deux conditions suivantes

— *est irréductible à diagonale fortement dominante*

— *est à diagonale strictement dominante,*

(3) vérifie $A_{ii} > 0$ *pour tout* i

est strictement définie positive.

Démonstration. D'après le théorème de Gershgorin 34.157, chaque valeur propre de A est dans un des disques fermés

$$D_i = \{z \in \mathbb{C} \text{ tel que } |z - A_{ii}| \leqslant r_i\}. \tag{34.486}$$

Par hypothèse, les centres de ces disques sont réels et strictement positifs. Mais le fait que A soit à diagonale dominante donne que le rayon de ces cercles sont plus petits que A_{ii}. Donc D_i n'intersecte pas $]-\infty, 0[$. Mais le fait que A soit symétrique implique que les valeurs propres soient réelles (théorème 8.227(1)). Cela montre que les valeurs propres de A sont toutes dans $[0, \infty[$.

Si la matrice A est à diagonale strictement dominante, alors les inégalités sont strictes et le théorème est prouvé.

Sinon nous somme dans le cas irréductible à diagonale fortement dominante et nous avons le théorème de Gershgorin numéro 2 34.158. Soit une valeur propre λ. Soit elle est dans un des disques ouvert (qui est inclus dans $]0, \infty[$), soit elle est dans l'intersection des bords des disques. Mais au moins un des disques n'intersecte pas 0 (parce que la diagonale est fortement dominante). Dans ce cas non plus λ ne peut pas être nul.

Nous en déduisons que dans tous les cas, les valeurs propres sont toutes réelles strictement positives. □

Chapitre 35

Méthode des différences finies

¡¡ Avertissement/question à la lectrice !! 35.1
Dans toute la partie sur la méthode des différences finies, il y a un flottement entre Ω et $\bar{\Omega}$.

D'une part je ne vois pas bien pourquoi on ne peut pas se contenter de travailler avec une fonction $u\colon \Omega \to \mathbb{R}$ avec Ω ouvert dans $\mathbb{R}$ et u même pas définie sur le bord.

Mais d'autre part, de nombreuses sources demandent de la régularité sur un fermé, à commencer par wikipédia :

https://fr.wikipedia.org/wiki/Méthode_des_différences_finies

Si vous avez une idée sur la question, écrivez-moi, ou répondez directement sur la page de discussion de la page Wikipédia, sur laquelle j'ai laissé une question.

35.1 Problèmes de dimension un

Soit $u\colon \mathbb{R} \to \mathbb{R}$ une fonction et $h > 0$. Nous définissons les opérations suivantes (qui sont supposées approximer la dérivée $u'(x)$ lorsqu'elle existe).

Définition 35.2.
*La **différence progressive** est*

$$(D_h^+ u)(x) = \frac{u(x+h) - u(x)}{h}, \tag{35.1}$$

*la **différence régressive** est*

$$(D_h^- u)(x) = \frac{u(x) - u(x-h)}{h}, \tag{35.2}$$

*la **différence centrée** est*

$$(D_h^0 u)(x) = \frac{u(x+h) - u(x-h)}{2h}. \tag{35.3}$$

Nous ne noterons pas toujours la dépendance en h, c'est-à-dire que nous noterons D^+u au lieu de $D_h^+ u$ lorsque cela ne pose pas de problème.

Notons que u'' peut être approximé par $D_h^+ D_h^+ u$, $D_h^0 D_h^+$, $D_h^+ D_h^-$, et encore de nombreuses autres possibilités.

Voici un lemme qui dit que tout cela n'est pas si mal, pourvu que u soit assez régulière.

LEMooZECZooVKxOZZ

Lemme 35.3.
Soit un ouvert connexe Ω de $\mathbb{R}$, soit $x \in \Omega$ et $h > 0$ tel que $\overline{B(x,h)} \subset \Omega$.

(1) Si $u \in C^2(\Omega)$ alors

$$|u'(x) - D_h^+ u(x)| \leqslant \frac{h}{2} \|u''\|_\Omega \tag{35.4}$$

et

$$|u'(x) - D_h^- u(x)| \leqslant \frac{h}{2} \|u''\|_\Omega. \tag{35.5}$$

ITEMooSAWJooJUTWAb

(2) Si $u \in C^3(\Omega)$ alors

$$|u'(x) - D_h^0(x)| \leqslant \frac{h^2}{2}\|u^{(3)}\|_\Omega \tag{35.6}$$

ITEMooRWUHooZJLKuL

(3) Si $u \in C^4(\Omega)$ alors

$$|u''(x) - D_h^- D_h^+ u(x)| \leqslant \frac{h^2}{12}\|u^{(4)}\|_\Omega. \tag{35.7}$$

Démonstration. Nous prouvons le point (3). D'abord nous regardons de quoi nous avons besoin :

$$D^-D^+u(x) = \frac{(D^+u)(x) - (D^+u)(x-h)}{h} = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \tag{35.8}$$

EQooBLIIooWHXbqD

Nous allons y mettre les approximations de $u(x+h)$ et $u(x-h)$ par Taylor, proposition 12.472(2) :

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(x+\theta_1 h) \tag{35.9}$$

avec $\theta_1 \in [0,1]$. De même,

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{6}u'''(x) + \frac{h^4}{24}u^{(4)}(x-\theta_2 h) \tag{35.10}$$

avec $\theta_2 \in [0,1]$.

Donc

$$u(x+h) + u(x-h) - 2u(x) = h^2u''(x) + \frac{h^4}{4!}\Big(u^{(4)}(x+\theta_1 h) + u^{(4)}(x-\theta_2 h)\Big), \tag{35.11}$$

ce qui donne

$$(D^-D^+u)(x) = u''(x) + \frac{h^2}{4!}\Big(u^{(4)}(x+\theta_1 h) + u^{(4)}(x-\theta_2 h)\Big). \tag{35.12}$$

Chacun des deux termes dans la parenthèse peut être majoré par $\|u^{(4)}\|_\Omega$ parce que $x + \theta_1 h$ ne prend ses valeurs que dans $[x, x+h] \subset \overline{B(x,h)} \subset \Omega$. Quoi qu'il en soit nous ne pouvons pas dire mieux que

$$|u''(x) - D^-D^+u(x)| \leqslant \frac{h^2}{12}\|u^{(4)}\|_{\overline{\Omega}}. \tag{35.13}$$

□

Remarque 35.4 ([1]).
Si nous avons l'égalité

$$|u'(x) - D_h^+u(x)| \leqslant \delta \tag{35.14}$$

EQooHSPFooTJIoFx

pour tout x, il faut faire attention en écrivant

$$\|u' - D_h^+u\|_\infty \leqslant \delta \tag{35.15}$$

parce que l'inégalité (35.14) n'est valable que pour les x tels que $[x-h, x+h] \subset \Omega$, de telle sorte que l'inégalité n'est pas spécialement correcte sur $\overline{\Omega}$. Il faut donc d'abord se mettre d'accord sur ce que signifie $\|.\|_\infty$. Est-ce une norme supremum sur Ω ou sur $\bar{\Omega}$?

35.1.1 Un schéma à cinq points

35.1.1.1 Poser le système

Soit $\Omega =]0,1[$ et l'équation différentielle

$$\begin{cases} -u''(x) + c(x)u(x) = f \quad \text{sur } \Omega \\ u(0) = \alpha \\ u(1) = \beta \end{cases} \tag{35.16}$$

EQooXJBWooRhCsLv

où c est une fonction positive et $\alpha, \beta \in \mathbb{R}$. Nous considérons $h > 0$ assez petit pour que le reste ait un sens. Si nous cherchons des solutions dans $C^4(\Omega)$, le lemme 35.3 nous dit que

$$|u''(x) - D^-D^+u(x)| = \eta(h^2) \tag{35.17}$$

où η est une fonction telle que $\lim_{t\to 0} \eta(t) = 0$. Nous pouvons récrire l'équation différentielle sous la forme

$$-D^-D^+u(x) + c(x)u(x) = f(x) + \eta(h^2). \tag{35.18}$$

Si nous négligeons le terme $\eta(h^2)$ qui est supposé être petit nous pouvons tenter de résoudre pour la fonction u_h

$$-D^-D^+u_h(x) + c(x)u_h(x) = f(x). \tag{35.19}$$

Notons ici l'importance de la notion de problème bien posé parce qu'en remplaçant le paramètre (fonctionnel) f par $f + \eta(h^2)$, nous modifions les solutions. Dans la mesure où le problème est bien posé, cette petite modification ne modifiera pas trop la solution et nous pouvons espérer que $\|u - u_h\|$ soit petit pour une norme ou une autre.

Utilisant l'expression (35.8) pour D^-D^+ nous avons l'équation suivante pour u_h :

$$\frac{1}{h^2}\Big(2u_h(x) - u_h(x+h) - u_h(x-h)\Big) + c(x)u_h(x) = f(x). \tag{35.20}$$ EQooECLXooFxZEeA

Avons-nous gagné quelque chose ? Pas encore. L'idée de la discrétisation est de ne considérer u_h qu'en certains points, et de prendre ces points à intervalles réguliers de taille h. Soient donc N un nombre entier et $h = 1/(N+1)$. Nous posons

$$x_k = kh \tag{35.21}$$

pour $k = 0, \dots, N+1$. Avec cela nous avons

$$\overline{\Omega} = \bigcup_{k=0}^{N} [x_k, x_{k+1}] \tag{35.22a}$$

$$x_0 = 0 \tag{35.22b}$$

$$x_{N+1} = 1. \tag{35.22c}$$

Nous posons surtout

$$\Omega_h = \{x_i\}_{i=1,\dots,N} \tag{35.23}$$

et

$$\overline{\Omega_h} = \{x_i\}_{i=0,\dots,N+1}. \tag{35.24}$$

Enfin, nous ne considérons plus u_h que comme une fonction $u_h \colon \overline{\Omega_h} \to \mathbb{R}$. C'est-à-dire que u_h est un vecteur à $N+2$ composantes.

L'équation (35.20) devient

$$\frac{1}{h^2}\big(2u_h(x_i) - u_h(x_{i+1}) - u_h(x_{i-1}) + c(x_i)u_h(x_i)\big) = f(x_i) \tag{35.25}$$ EQooZMVMooTqIpkF

pour $i = 1, \dots, N$. Sur les bords, cette équation n'est pas possible parce que x_{i-1} ou x_{i+1} n'existerait pas. Au contraire, sur les bords nous avons les conditions aux bords

$$u_h(x_0) = \alpha \tag{35.26}$$

et

$$u_h(x_{N+1}) = \beta. \tag{35.27}$$

Nous posons $c_i = c(x_i)$ et $u_i = u_h(x_i)$. Les nombres u_0 et u_{N+1} sont donnés par les conditions aux bords, et les inconnues du problème sont donc les nombres u_i ($i = 1, \dots, N$). Pour les déterminer, nous devons résoudre un système d'équations linéaire.

L'écriture du système linéaire à résoudre consiste essentiellement à écrire (35.25) en séparant les cas $i = 1$ et $i = N$ parce que nous connaissons déjà les valeurs de u_0 et u_{N+1}. Le système que nous avons est :

$$\begin{cases} \left(\frac{2}{h^2} + c_1\right) u_1 - \frac{1}{h^2} u_2 = f_1 + \frac{\alpha}{h^2} & i = 1 \quad (35.28a) \\ \left(\frac{2}{h^2} + c_N\right) u_N - \frac{1}{h^2} u_{N-1} = f_N + \frac{\beta}{h^2} & i = N \quad (35.28b) \\ \left(\frac{2}{h^2} + c_i\right) u_i - \frac{1}{h^2} u_{i+1} - \frac{1}{h^2} u_{i-1} = f_i & \text{autres.} \quad (35.28c) \end{cases}$$

Cela se met sous la forme matricielle

$$L_h U_h = F_h \tag{35.29}$$

pour

$$F_h = \left(f_1 + \frac{\alpha}{h^2}, f_2, \ldots, f_{N-1}, f_N + \frac{\beta}{h^2}\right) \tag{35.30}$$

EQooMNTJooYPYoAj

et les éléments non nuls de L_h sont :

$$(L_h)_{i,i-1} = -\frac{1}{h^2} \quad \text{pour } i = 2, \ldots, N \tag{35.31a}$$

$$(L_h)_{i,i+1} = -\frac{1}{h^2} \quad \text{pour } i = 1, \ldots, N-1 \tag{35.31b}$$

$$(L_h)_{i,i} = \frac{2}{h^2} + c_i \quad \text{pour } i = 1, \ldots, N. \tag{35.31c}$$

Cette matrice est pleine de zéros, à part les trois diagonales centrales, et il existe des méthodes efficaces pour résoudre le système d'équations correspondant.

35.1.1.2 Propriétés du système

La matrice est la suivante :

$$L_h = \begin{pmatrix} \frac{2}{h^2} + c_1 & -1/h^2 & 0 & 0 & \cdots & 0 \\ -1/h^2 & \frac{2}{h^2} + c_2 & -1/h^2 & 0 & \cdots & 0 \\ 0 & -1/h^2 & \frac{2}{h^2} + c_3 & -1/h^2 & \ddots & \vdots \\ 0 & 0 & -1/h^2 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & -1/h^2 \\ 0 & 0 & \cdots & 0 & -1/h^2 & \frac{2}{h^2} + c_N \end{pmatrix} \tag{35.32}$$

où nous avions déjà posé l'hypothèse $c_i \geqslant 0$ pour tout i.

LEMooGGHQooNnVsuu

Lemme 35.5.
La matrice L_h est irréductible[1] *à diagonale fortement dominante*[2].

Démonstration. Nous décomposons la preuve en plusieurs parties, en notant L pour L_h, afin d'alléger les notations.

(1) **La première ligne** Sur la première ligne, seuls deux éléments sont non nuls et nous avons $|L_{11}| = \frac{2}{h^2} + c_1$ parce que c est une fonction positive et $|L_{12}| = \frac{1}{h^2}$. Nous avons donc

$$|L_{11}| - |L_{12}| = \frac{2}{h^2} + c_1 - \frac{1}{h^2} = \frac{1}{h^2} + c_1 > 0. \tag{35.33}$$

L'inégalité stricte est importante.

1. Caractérisation 34.154.
2. Définition 34.161.

(2) **La dernière ligne** Elle est semblable à la première.

(3) **Les autres lignes** Sur les autres lignes nous avons trois éléments non nuls et

$$\sum_{j\neq i}|L_{ij}|=\frac{2}{h^2}\leqslant\frac{2}{h^2}+c_i=L_{ii}. \tag{35.34}$$

(4) **Diagonale fortement dominante** Nous avons prouvé jusqu'à présent que L_h était une matrice à diagonale fortement dominante.

(5) **Irréductible** Nous allons utiliser la caractérisation de la proposition 34.154(3). Pour cela, nous considérons la chaine d'éléments non nuls

$$L_{12},L_{23},\dots,L_{N-1,N}=-\frac{1}{h^2}. \tag{35.35}$$

Soient deux indices i et j avec $i<j$. Cette suite d'indice (ou une sous-suite) rend i et j connectés.

Si par contre $i>j$, il faut considérer la suite inversée grâce au fait que L_h est symétrique :

$$L_{N,N-1},L_{N-1,N-2},\dots,L_{32},L_{21}=-\frac{1}{h^2}. \tag{35.36}$$

□

PROPooOQJVooJMTkVM

Proposition 35.6.
Soit le problème

$$\begin{cases}-u''(x)+c(x)u(x)=f\quad \text{sur } \Omega\\ u(0)=\alpha\\ u(1)=\beta\end{cases} \tag{35.37}$$

EQooEUHQooVHRelr

où c est une fonction positive et $\alpha,\beta\in\mathbb{R}$. Nous considérons $h>0$ assez petit pour que le reste ait un sens. Et nous approximons u'' par D^-D^+u.

La matrice L_h des différences finies associée à ce problème est

(1) une M-matrice,

(2) strictement définie positive,

(3) d'inverse $L_h^{-1}>0$.

Démonstration. Le théorème 34.167 dit que L_h est une M-matrice. La proposition 34.170 nous donne aussi que L_h est strictement définie positive.

Le lemme 35.5 dit que L_h est irréductible, ce qui permet à la proposition 34.168 de conclure que $L_h^{-1}>0$. □

Cela étant rappelé, nous pouvons continuer.

LEMooDXPRooOhwqSZ

Lemme 35.7.
Soit $\Omega=]0,1[$, soit $N\in\mathbb{N}$ et $h=1/(N+1)$. La solution $w_h\colon\Omega_h\to\mathbb{R}$ du problème discrétisé

SUBEQooFJKIooLvzM

$$\begin{cases}-(D^-D^+w_h)(x_k)=1 & \text{(35.38a)}\\ w_h(0)=0 & \text{(35.38b)}\\ w_h(1)=0 & \text{(35.38c)}\end{cases}$$

pour tout $x_k=kh$ $(k=1,\dots,N)$ donne les valeurs exactes des $w(x_k)$ lorsque w est la solution de

SUBEQooCRFWo

$$\begin{cases}-w''(x)=1 & \text{(35.39a)}\\ w(0)=0 & \text{(35.39b)}\\ w(1)=0. & \text{(35.39c)}\end{cases}$$

Démonstration. Un enseignement de la proposition 35.6 est que le système (35.38) peut être écrit sous la forme d'un système linéaire $L_h^0 w_h = F_h$ où L_h^0 est inversible. Il y a donc unicité de la solution.

D'autre part, la solution du système (35.39) est $w(x) = \frac{1}{2}(x - x^2)$, qui est de classe C^∞. Le lemme 35.3(3) dit que $D^- D^+ w = w''$. Donc les valeurs $w(x_k)$ résolvent aussi le système (35.38). □

Lemme 35.8 (Quelques estimations)**.**
La matrice L_h du problème sus-mentionné en (35.37) vérifie[3] *:*

(1) $\|L_h\|_\infty \leqslant \frac{4}{h^2} + \|c\|_\infty$

(2) $\|L_h^{-1}\|_\infty \leqslant \frac{1}{8}$.

Démonstration. Nous nous souvenons de la formule (10.169) :

$$\|A\|_\infty \leqslant \max_{i=1,\dots,n} \sum_{j=1}^{n} |A_{ij}|. \tag{35.40}$$

La première ligne a pour somme : $\frac{3}{h^2} + c_1$, la dernière a pour somme $\frac{3}{h^2} + c_n$ et les autres ont pour somme $\frac{4}{h^2} + c_i$. Elles sont donc toutes majorées par $\frac{4}{h^2} + \|c\|_\infty$.

Pour l'estimation de $\|L_h^{-1}\|_\infty$ nous allons nous appuyer sur le théorème 34.169.

Commençons par considérer le problème [SUBEQSooRENKooZaRjvL]

$$\begin{cases} -w'' = 1 & \text{(35.41a)} \\ w(0) = w(1) = 0. & \text{(35.41b)} \end{cases}$$

La première équation dit que w est un polynôme de degré 2. En écrivant $w(x) = ax^2 + bx + c$ et en imposant toutes les contraintes, nous trouvons l'unique solution

$$w(x) = -\frac{1}{2}(x^2 - x). \tag{35.42}$$

Le lemme 35.7 nous dit que la fonction w prise aux points $x_k = kh$ donne les valeurs de w_h.

La matrice L_h^0 est une M-matrice et le vecteur w_h vérifie $L_h^0 w_h = \mathbb{1}$. Donc le théorème 34.169 s'applique et

$$\|(L_h^0)^{-1}\| \leqslant \|w_h\|_\infty = \frac{1}{8}. \tag{35.43}$$

L'obtention de 1/8 n'est rien d'autre que la recherche du maximum (en valeur absolue) de la parabole $x \mapsto (x - x^2)/2$ pour $x \in [0, 1]$. Le maximum est atteint pour $x = 1/2$; calcul de dérivée et tout ça ...

Nous retournons maintenant à notre matrice originale L_h. Nous avons

$$L_h - L_h^0 = \operatorname{diag}(c_1, \dots, c_n) \geqslant 0, \tag{35.44}$$

et aussi

$$L_h^{-1} - (L_h^0)^{-1} = \underbrace{L_h^{-1}}_{\geqslant 0} \underbrace{(L_h^0 - L_h)}_{\leqslant 0} \underbrace{(L_h^0)^{-1}}_{\geqslant 0} \tag{35.45}$$

parce que L_h est une M-matrice. Donc tous les coefficients de $L_h^{-1} - (L_h^0)^{-1}$ sont négatifs. Cela implique

$$L_h^{-1} \leqslant (L_h^0)^{-1}. \tag{35.46}$$

Mais nous savons que les coefficients de L_h^{-1} sont positifs, donc le maximum de ses coefficients en valeur absolue est plus petit que ceux de $(L_h^0)^{-1}$, c'est-à-dire

$$\|L_h^{-1}\|_\infty \leqslant \|(L_h^0)^{-1}\|_\infty \leqslant \frac{1}{8}. \tag{35.47}$$

□

3. Dans le CTES d'analyse numérique de Marseille, l'estimation donnée est $\|L_h^{-1}\|_\infty \leqslant \frac{1}{4}$.

35.1.2 Exemple

Soit $\Omega =]0,1[$ et une fonction $u : \overline{\Omega} \to \mathbb{R}$ de classe C^4 vérifiant

$$\begin{cases} -u''(x) + u(x) = \sin(x) & (35.48a) \\ u(0) = 0 & (35.48b) \\ u(1) = 0. & (35.48c) \end{cases}$$

Nous allons écrire la méthode des différences finies pour $h = 1/4$. Nous posons donc les points

$$\begin{cases} x_0 = 0 & (35.49a) \\ x_1 = 1/4 & (35.49b) \\ x_2 = 1/2 & (35.49c) \\ x_3 = 3/4 & (35.49d) \\ x_4 = 1. & (35.49e) \end{cases}$$

Puisque nous avons supposé u de classe C^4, le lemme 35.3(3) nous donne [4]

$$u''(x) = (D^-D^+u)(x) + \alpha(h) \tag{35.50}$$

avec $\lim_{h\to 0} \alpha(h)/h = 0$. L'équation discrétisée serait alors SYSTooNEQHooOWJSbT

$$\begin{cases} -(D^-D^+u)(x) + u(x) = \sin(x) & (35.51a) \\ u(0) = u(1) = 0. & (35.51b) \end{cases}$$

où nous n'avons pas précisé l'indice h au bas des opérateurs D^+ et D^-. Les équations (35.51) ne doivent être posées que pour x_1, x_2 et x_3 parce que les valeurs en x_0 et x_4 sont déjà connues.

(1) **Pour x_1**

$$\frac{u_2 - 2u_1 + u_0}{h^2} + u_1 = \sin(x_1) \tag{35.52}$$

(2) **Pour x_2**

$$\frac{u_3 - 2u_2 + u_1}{h^2} + u_2 = \sin(x_2) \tag{35.53}$$

(3) **Pour x_3**

$$\frac{u_4 - 2u_3 + u_2}{h^2} + u_3 = \sin(x_3). \tag{35.54}$$

Nous tenons compte du fait que $u_0 = u_4 = 0$ et que $h = 1/4$ pour écrire le système

$$\begin{pmatrix} -31 & 16 & 0 \\ 16 & -31 & 16 \\ 0 & 16 & -31 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \tag{35.55}$$

où les s_k sont des nombres parfaitement connus : par exemple $s_1 = \sin(x_1) = \sin(1/4) \simeq 0.247403959254523$.

35.2 Problèmes de dimension deux

Nous allons considérer le système

$$\begin{cases} -\Delta u = f & \text{sur } \Omega \\ u = g & \text{sur } \partial\Omega \end{cases} \tag{35.56}$$

SYSooTANLooRgnIMp

où $\Omega =]0,a[\times]0,b[$.

Remarque 35.9.
Pourquoi un signe moins devant le laplacien ? Pour avoir la proposition 35.6 qui dira que la matrice correspondant aux différences finies appliquées à ce système est une M-matrice. Sinon, c'est la matrice $-L_h$ qui en serait une.

4. Nous ferions n'importe quoi pour ne pas écrire $u''(x) = (D^-D^+u)(x) + o(h^2)$. Notez que vous faites ce que vous voulez : écrivez avec la notation « petit o » si cela vous chante.

35.2.1 Discrétisation en croix

Nous allons maintenant déduire une discrétisation du laplacien en discrétisant les opérations ∂_x^2 et ∂_y^2. Nous discrétisons Ω en mailles carrées de côté h : $x_k = kh$ et $y_k = kh$. L'opération de dérivée partielle ∂_x est discrétisée par

$$(D_x^+ u)(x,y) = \frac{u(x+h,y) - u(x,y)}{h} \tag{35.57}$$

ou

$$(D_x^- u)(x,y) = \frac{u(x,y) - u(x-h,y)}{h} \tag{35.58}$$

ou

$$(D_x^0 u)(x,y) = \frac{u(x+h,y) - u(x-h,y)}{2h} \tag{35.59}$$

où le h est sous-entendu dans les opérateurs D^0, D^+ et D^-.

La dérivée partielle seconde $\partial_x^2 u$ peut être approximée par toutes les combinaisons imaginables, par exemple

$$(D_x^- D_x^+ u)(x,y) = \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2}. \tag{35.60}$$

Pour évaluer la différence entre $(\partial_x^2 u)(x,y)$ et $(D^-D^+u)(x,y)$, il est possible d'utiliser le théorème de Taylor en deux dimensions, mais nous pouvons également recycler ce qui a été fait. Nous posons $u_y(x) = u(x,y)$ et alors $(\partial_x^2 u)(x,y) = u_y''(x)$ et le lemme 35.3(3) donne, si u_y est de classe C^4,

$$|u_y''(x) - D^-D^+ u_y(x)| \leqslant \frac{1}{12} h^2 \|u_y^{(4)}\|_\infty. \tag{35.61}$$

Là, les opérateurs D^+ et D^- sont ceux à une dimension. Mais nous avons $(D^-D^+u_y)(x) = (D^-D^+u)(x,y)$ (à droite ce sont les opérateurs à deux dimensions), donc

$$\big|(\partial_x^2 u)(x,y) - (D^-D^+u)(x,y)\big| \leqslant \frac{1}{12} h^2 \|\partial_x^4 u\|_\infty \tag{35.62}$$

et nous pouvons écrire

$$(\partial_x^2 u)(x,y) = (D^-D^+u)(x,y) + h^2 R(x,y,h) \tag{35.63}$$

où R est une fonction qui dépend de x, y et h, mais aussi de u. Le point important est que R soit majoré par une quantité indépendante de h, de telle sorte que nous ayons quelques garanties que négliger ce terme soit une bonne approximation lorsque $h \to 0$.

Au niveau de la discrétisation, nous considérons x_i avec $i = 0, \dots, N_x$ et y_j avec $j = 0, \dots, N_y$. La discrétisation de $-(\Delta u)(x,y) = f(x,y)$ donne, pour $i = 1, \dots, N_x - 1$ et $j = 1, \dots, N_y - 1$,

$$\frac{1}{h^2}(-u_{i+1,j} + 4u_{ij} - u_{i-1,j} - u_{i,j+1} + u_{i,j-1}) = f_{ij}. \tag{35.64}$$

Les équations avec i ou j valant 0 ou N_x, N_y sont les valeurs aux bords.

35.10.
Nous notons pour référence ultérieure la discrétisation suivante du laplacien :

$$(\Delta_h u)(x_i, y_j) = \frac{1}{h^2}\big(-4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j-1} + u_{i,j+1}\big). \tag{35.65}$$

Elle vérifie

$$\Delta_h f = \Delta f + h^2 \alpha(h). \tag{35.66}$$

Cette discrétisation est dite « en croix » parce que les points exploités forment une croix.

35.2.2 Discrétisation en carré

L'opérateur laplacien est défini par $\Delta = \partial_x^2 + \partial_y^2$, mais il existe de nombreuses autres façons de l'écrire.

Lemme 35.11.
Le laplacien est invariant par changement de coordonnées orthogonales. Plus précisément, si A est une matrice orthogonale, en posant $u_i = \sum_k A_{ik} e_k$ nous avons

$$\sum_i \frac{\partial^2}{\partial u_i^2} f = \Delta f. \tag{35.67}$$

Démonstration. Nous avons :

$$\frac{\partial f}{\partial u_i} = \sum_k A_{ik} \frac{\partial f}{\partial x_k}, \tag{35.68}$$

et donc

$$\sum_i \frac{\partial^2 f}{\partial u_i^2} = \sum_i \frac{\partial}{\partial u_i} \left(\sum_k A_{ik} \frac{\partial f}{\partial x_k} \right) = \sum_{ijk} (A^t)_{ji} A_{ik} \partial^2 f = \sum_{jk} (A^t A)_{jk} \partial_{jk}^2 f. \tag{35.69}$$

En particulier si A est une matrice orthogonale, $(A^t A)_{jk} = \delta_{jk}$ et le résultat est prouvé. □

Les plus convaincus diront que $\Delta = \nabla \cdot \nabla$ et que le produit scalaire est invariant sous changement de coordonnées orthogonales.

Nous avons déjà déduit la discrétisation (35.65) du laplacien :

$$(\Delta_h u)(x_i, y_j) = \frac{1}{h^2} \big(-4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j-1} + u_{i,j+1} \big). \tag{35.70}$$

Nous allons maintenant en déduire une par l'idée de décomposer le laplacien dans la base $u = e_1 + e_2$, $v = e_1 - e_2$. Pour cela nous introduisons les opérations (le nombre h dont dépendent ces opérateurs est sous-entendu)

$$(D_u^+ f)(x, y) = \frac{f(x + h, y + h) - f(x, y)}{h} \tag{35.71a}$$

$$(D_v^+ f)(x, y) = \frac{f(x + h, y - h) - f(x, y)}{h} \tag{35.71b}$$

$$(D_u^- f)(x, y) = \frac{f(x, y) - f(x - h, y - h)}{h} \tag{35.71c}$$

$$(D_v^- f)(x, y) = \frac{f(x, y) - f(x - h, y + h)}{h}. \tag{35.71d}$$

Puisque

$$\partial_u^2 + \partial_v^2 = 2\Delta, \tag{35.72}$$

nous discrétisons le laplacien par

$$\Delta_h' = \frac{1}{2} \big(D_u^- D_u^+ + D_v^- D_v^+ \big). \tag{35.73}$$

Un peu de calcul donne : [EQooLHBDooSBFkho]

$$(\Delta_h' f)(x, y) = \frac{1}{2h^2} \Big(-4f(x, y) + f(x + h, y + h) + f(x - h, y + h) \tag{35.74a}$$

$$+ f(x + h, y - h) + f(x - h, y - h) \Big). \tag{35.74b}$$

35.2.3 Résolution de la discrétisation en croix

Les équations (35.64) forment un système d'équations linéaires à résoudre. Certaines peuvent être simplifiées parce qu'elles « touchent » le bord. Nous verrons cela un peu plus tard.

Nous allons d'abord numéroter correctement les équations de façon à ne pas avoir deux mais un seul indice. Notre fonction de numérotation sera

$$\varphi(i,j) = (j-1)(N_x-1) + i \tag{35.75}$$

avec $i = 1, \dots, N_x - 1$ et $j = 1, \dots, N_y - 1$. Cela correspond à numéroter les points de l'intérieur du quadrillage ligne par ligne de bas en haut, et de gauche à droite. Avec cela les équations (35.64) vont être numérotées par un seul indice I allant de $\varphi(1,1) = 1$ à $\varphi(N_x - 1, N_y - 1) = (N_x - 1)(N_y - 1)$.

Si $I = \varphi(i,j)$ alors nous avons vite

$$\varphi(i+1,j) = I + 1 \tag{35.76a}$$
$$\varphi(i,j+1) = I + N_x - 1 \tag{35.76b}$$
$$\varphi(i-1,j) = I - 1 \tag{35.76c}$$
$$\varphi(i,j-1) = I - N_x + 1. \tag{35.76d}$$

Nous posons $U_I = u_{\varphi^{-1}(I)}$, et l'équation (35.64) devient

$$\frac{1}{h^2}(-U_{I+1} + 4U_I - U_{I-1} - U_{I+N_x-1} - U_{I-N_x+1}) = f_I. \tag{35.77}$$

Pour écrire la matrice représentant ce système, nous devons simplifier les équations qui doivent l'être. Par exemple avec $I = 1$, le terme $U_{I-1} = U_0$ vaut $u_{0,1} = f_{0,1}$. Ce n'est donc pas réellement une inconnue de notre problème.

Nous voulons mettre les équations sous la forme du système

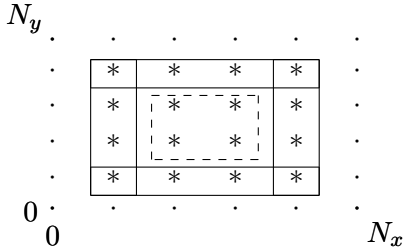
$$L_h U = F. \tag{35.78}$$

Sur la ligne numéro I de L_h, les éléments non nuls sont : SUBEQQooSRQNooYrCNhj

$$L_{I,I} = 4 \tag{35.79a}$$
$$L_{I,I+1} = -1 \tag{35.79b}$$
$$L_{I,I-1} = -1 \tag{35.79c}$$
$$L_{I,I+N_x-1} = -1 \tag{35.79d}$$
$$L_{I,I-N_x+1} = -1 \tag{35.79e}$$

pour peu qu'ils existent. Par exemple pour $I = 1$, il n'y a pas d'éléments $L_{I,I-1}$. Les indices I et J de $L_{I,J}$ vont de 1 à $\varphi(N_x - 1, N_y - 1) = (N_y - 1)(N_x - 1)$.

Voici un dessin de notre situation :



À chaque élément du quadrillage correspond une équation.

— Aux points simples sur le bord, correspondent des équations triviales parce que la fonction u y est directement donnée par les conditions aux bords.

— Aux points étoilés entourés en traits continus correspondent des équations « incomplètes » parce que certains termes de l'équation (35.64) sont donnés par les conditions aux bords. Elle correspondent aussi aux lignes incomplètes de la matrice L_h où certains éléments donnés en (35.79) n'existent pas.

Le membre de droite de ces équations est par contre enrichi de ce qui à gauche est « donné ».

— Aux points étoilés du centre entourés en traits discontinus correspondent des équations complètes.

Notons que $f_{0,0}$ ne joue aucun rôle dans notre histoire parce que dans les équations (35.64), chaque point (i, j) du maillage n'est lié qu'aux quatre points situés « à côté ».

PROPooWGTRooVjWhYY

Proposition 35.12.
La matrice L_h est

(1) irréductible et à diagonale fortement dominante[5],

ITEMooOOHPooDsvUPP

(2) une M-matrice,

(3) inversible avec $L_h > 0$,

(4) symétrique,

(5) strictement définie positive.

Démonstration. On divise la preuve.

(1) **Irréductible** Une matrice $n \times n$ dont les deux premières diagonales sont entièrement composées d'éléments non nuls est toujours irréductible. En effet, la première lie l'élément $(1, 2)$ à l'élément $(n - 1, n)$ et donc permet de dire que tous les $i < j$ sont connectés.

La seconde diagonale lie l'élément $(n, n - 1)$ à l'élément $(2, 1)$.

(2) **Diagonale fortement dominante** En ce qui concerne la dominance de la diagonale, il faut sommer sur les lignes. Or chaque ligne contient (en valeur absolue) un 4 sur la diagonale et au plus quatre éléments qui valent 1. D'où

$$|L_{II}| \geqslant \sum_{J \neq I} |L_{IJ}|. \tag{35.80}$$

La première ligne n'est jamais complète : elle contient un 4 sur l'élément $(1, 1)$ et au maximum, deux 1, plus à droite. Donc la matrice L_h est à diagonale fortement dominante.

(3) **M-matrice**
D'après ce que nous venons de voir (proposition 35.12), le théorème 34.167 s'applique et L_h est une M-matrice[6].

(4) **Inverse strictement positif** La proposition 34.168 nous assure qu'une M-matrice irréductible est d'inverse strictement positif. Donc $L_h^{-1} > 0$.

(5) **Symétrique** La ligne numéro I est

$$\big(\dots, \underbrace{-1}_{I-N_x+1}, \dots, -1, 4, -1, \dots, \underbrace{-1}_{I+N_x-1}, \dots\big) \tag{35.81}$$

Prenons par exemple l'élément $(I, I - N_x + 1)$ qui vaut -1. Son symétrique est l'élément $(I - N_x + 1, I)$ qui se trouve sur la ligne $I - N_x + 1$. Sur cette dernière ligne nous avons un -1 sur la colonne $I - N_x + 1 + N_x - 1 = I$. Donc l'élément $(I - N_x + 1, I)$ vaut bien -1 et la matrice est symétrique.

5. Définition 34.161. Le cas 1×1 est discutablement à diagonale fortement dominante, il faut avouer.

6. Notons que c'est ici que nous sommes content d'avoir posé $-\Delta u = f$ dans le système (35.56), avec un signe négatif devant le laplacien. Sinon tous les signes auraient changé, et la matrice $-L_h$ aurait été une M-matrice, au lieu de L_h.

(6) **Strictement définie positive** Vu que la matrice L_h est symétrique, irréductible à diagonale fortement dominante (proposition 35.12), et comme ses éléments diagonaux sont strictement positifs (ils valent 4), la proposition 34.170 nous dit que L_h est strictement définie positive.

□

35.3 Consistance, convergence

35.3.1 Définitions, mise en place

Soit un ouvert $\Omega \subset \mathbb{R}^n$ et un opérateur différentiel L sur Ω. Nous considérons le problème qui consiste à trouver une fonction u sur Ω telle que

$$Lu = f \tag{35.82}$$

pour une fonction f donnée.

¡¡ Avertissement/question au lecteur !! 35.13
La définition suivante est une invention personnelle, n'est pas précise et mérite des commentaires de la part du lecteur.

Définition 35.14 ([1]).
*Un **schéma numérique** de pas h pour $Lu = f$ est la donnée de*

(1) un nombre $h > 0$ supposé petit,

(2) une quantité N de points x_i dans Ω formant l'ensemble discret Ω_h,

(3) une matrice L_h de taille $N \times N$,

(4) une solution $u_h : \Omega_h \to \mathbb{R}$ de l'équation $(L_h u_h)(x_i) = f_i$ où nous avons posé $f_i = f(x_i)$.

35.15.
Évidemment pour qu'un schéma mérite le nom de schéma de pas h pour l'équation $Lu = f$, il faut que le nombre h soit lié au choix des points x_i, et que la matrice L_h soit liée à l'opérateur L. La définition n'impose pas formellement de tels liens, parce qu'il y a de nombreuses façons d'approximer une équation différentielle en un système linéaire, sans compter que même l'équation $(L_h u_h)_i = f_i$ peut se résoudre de beaucoup de façons, exactes ou approchées.

Cela pour dire que le lien entre la solution exacte u et la solution approchée n'a rien d'évident, et va dépendre des choix faits lors de la discrétisation et lors de la résolution du système linéaire. Nous allons supposer dans un premier temps que l'équation $L_h u_h = f$ est résolue exactement (nous avons un peu parlé de ces problèmes dans les sections 34.15 et suivantes).

Définition 35.16.
*L'erreur **de consistance** d'un schéma numérique est la fonction $\tau_h : \Omega_h \to \mathbb{R}$ définie par*

$$\tau_h(x_i) = (L_h u)_i - (Lu)(x_i). \tag{35.83}$$

EQooPFBBooJumDZQ

Il y a un jeu de notation pas tout à fait évident dans la définition (35.83). En effet, L_h est une matrice, et ne s'applique donc à priori pas immédiatement à une fonction. Ce que signifie la notation $(L_h u)_i$ est que l'on applique la matrice L_h au vecteur $j \mapsto u(x_j)$ et que l'on prend la composante i du résultat.

Définition 35.17.
*Nous disons que le schéma est **consistant** avec l'opérateur différentiel L lorsque*

$$\lim_{h \to 0^+} \|\tau_h\| = 0 \tag{35.84}$$

EQooMPQYooCZsaAT

où la norme $\|.\|$ est souvent la norme uniforme, c'est-à-dire $\|\tau_h\| = \max_i \tau_h(x_i)$.

Notons que le lien entre h et le choix des x_i fait partie de la définition du schéma. Sur un segment de longueur ℓ, lorsque h n'est pas un diviseur de ℓ, le schéma devrait expliquer ce que l'on fait pour que la limite (35.84) ait un sens.

Définition 35.18.
*Le schéma (Ω_h, L_h) est **consistant à l'ordre** p avec l'opérateur différentiel L pour la norme $\|.\|$ si il existe une constante C indépendante de h telle que*

$$\|\tau_h\| \leqslant Ch^p. \tag{35.85}$$

Définition 35.19.
*L'**erreur de discrétisation** entre la solution u du problème $Lu = f$ et la solution approchée u_h sur Ω_h est la fonction*

$$\begin{aligned} e_h \colon \Omega_h &\to \mathbb{R} \\ x_i &\mapsto u(x_i) - u_i. \end{aligned} \tag{35.86}$$

où $u_i = u_h(x_i)$ est la solution approchée.

*Le schéma discret $(L_h u_h)(x_i) = f_i$ est **convergent** si $\lim_{h\to 0} \|e_h\| = 0$. Si de plus, il existe une constante C, et $p > 0$, tels que*

$$\|e_h\| \leqslant Ch^p, \tag{35.87}$$

*alors nous disons que le schéma est convergent à l'**ordre** p.*

Si l'erreur de consistance est petite, le *problème* est bien approximé par le système linéaire. Cela n'implique cependant pas que la solution trouvée soit bien approximée.

Exemple 35.20 (Deux opérateurs différentiels proches dont les solutions sont loin)**.**
Soit la partie $\Omega =]0, \infty[$, et les problèmes

$$\begin{cases} L_1 u = u' = 0 & \text{(35.88a)} \\ u(0) = 1 & \text{(35.88b)} \end{cases}$$

et

$$\begin{cases} L_2 v = v' - \epsilon v = 0 & \text{(35.89a)} \\ v(0) = 1. & \text{(35.89b)} \end{cases}$$

Les solutions exactes sont $u(x) = 1$ et $v(x) = e^{\epsilon x}$.

En ce qui concerne les opérateurs, quelle que soit la norme utilisée, nous avons

$$\begin{aligned} \|L_1 - L_2\| &= \sup_{\|f\|=1} \|L_1(f) - L_2(f)\| && \text{(35.90a)} \\ &= \sup_{\|f\|=1} \|\epsilon f\| && \text{(35.90b)} \\ &= \epsilon. && \text{(35.90c)} \end{aligned}$$

Donc lorsque ϵ est petit, l'opérateur L_2 approxime bien l'opérateur L_1. Pour toutes les normes. Mais

$$\big|u(x) - v(x)\big| = |1 - e^{\epsilon x}|, \tag{35.91}$$

donc quel que soit ϵ nous avons $\|u - v\|_\infty = \infty$. Et d'ailleurs, quelle que soit la norme raisonnable que nous mettons sur l'espace des fonctions, avoir $\|u - v\| = \infty$ semble inévitable.

Donc deux opérateurs différentiels proches peuvent avoir des solutions lointaines. △

35.3.2 Exemple

Soit l'opérateur différentiel L donné par

$$Lu = -u'' + cu \tag{35.92}$$

où c est une fonction. Nous considérons sur $\Omega =]0,1[$ l'équation différentielle

$$Lu = 0. \tag{35.93}$$

En ce qui concerne la discrétisation, nous définissons le maillage $\Omega_h = \{x_i = ih\}$ avec $i = 0, \dots, N+1$. La solution approchée discrètement sera le vecteur v qui peut être vu comme fonction $v\colon \Omega_h \to \mathbb{R}$. Les nombres v_0 et v_{N+1} sont à priori donnés par les conditions aux bords. Pour les autres v_i nous avons les équations

$$(L_h v)_i = -\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + c(x_i)v_i. \tag{35.94}$$

Cela est la définition de l'opérateur L_h, et le vecteur v solution de $L_h v = 0$, est la solution du problème au sens de la méthode des différences finies (pour peu qu'il existe, soit unique et tout ça).

Pour calculer l'erreur de consistance, nous considérons une fonction u et nous posons $u_i = u(x_i)$. Le vecteur (u_i) ainsi construit est approximé par v (on espère). Nous avons :

$$\tau_h(x_i) = -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} - c(x_i)u_i - (Lu)(x_i). \tag{35.95}$$

Pour étudier cela nous développons $u_{i+1} = u(x_i + h)$ et $u_{i-1} = u(x_i - h)$ à l'ordre 4 : il existe $\alpha_i \in [x_i, x_i + h]$ et $\beta_i \in [x_i - h, x_i]$ tels que

$$u_{i+1} = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{3!}u^{(3)}(x_i) + \frac{h^4}{4!}u^{(4)}(\alpha_i) \tag{35.96}$$

et

$$u_{i-1} = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{3!}u^{(3)}(x_i) + \frac{h^4}{4!}u^{(4)}(\beta_i). \tag{35.97}$$

Après simplification de plusieurs termes,

$$\tau_h(x_i) = -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} - c_i u_i + u''(x_i) + c_i u_i = \frac{h^2}{4!}\big(u^{(4)}(\alpha_i) + u^{(4)}(\beta_i)\big). \tag{35.98}$$

Parler de la consistance du schéma demande d'étudier $\lim_{h\to 0^+} \|\tau_h\|$, et pour cela, il faut préciser la norme avec laquelle nous voulons travailler. L'ordre de consistance va dépendre de la norme utilisée.

Pour la norme $\|.\|_\infty$, les nombres $u^{(4)}(\alpha_i)$ et $u^{(4)}(\beta_i)$ se majorent par $\|u^{(4)}\|_\infty$ et nous avons

$$\|\tau_h\|_\infty \leqslant \frac{h^2}{12}\|u^{(4)}\|_\infty. \tag{35.99}$$

Nous avons consistance d'ordre 2.

Remarque 35.21.
La valeur de $\|\tau_h\|_\infty$ dépend de la fonction u sur laquelle nous la calculons. Cependant nous avons convergence $\|\tau_h\|_\infty \to 0$ pour toute fonction (de classe disons, C^4).

La constante C pour laquelle nous avons $\|\tau_h\| \leqslant Ch^2$ et donc qui nous vaut de pouvoir dire que la consistance est d'ordre 2 ne dépend pas de h, ni des valeurs ponctuelles de u ou de ses dérivées, mais dépend des normes de u et de ses dérivées (en l'occurrence seulement de la norme de $u^{(4)}$.)

Étudions la consistance pour la norme L_1 :

$$\|\tau_h\|_1 = \sum_i |\tau_h(x_i)| \leqslant \frac{h^2}{12}\sum_i \|u^{(4)}\|_\infty \tag{35.100}$$

où nous avons majoré chacun des $u^{(4)}(\alpha_i)$ par $\|u^{(4)}\|_\infty$. Combien de termes dans la somme ? Nous avons $h = 1/(N-1)$ et donc $N = (1+h)/h$, ce qui donne

$$\|\tau_h\|_1 \leqslant N\frac{h^2}{12}\|u^{(4)}\|_\infty = (1+h)Ch. \tag{35.101}$$

La constante $1+h$ se majore par n'importe quelle constante strictement plus grande que 1. Nous pouvons donc la rentrer dans C et écrire

$$\|\tau_h\|_1 \leqslant Ch \tag{35.102}$$

et donc avoir la consistance à l'ordre 1.

35.3.3 Consistance, stabilité et convergence

LEMooUOUMooPCoAtA

Lemme 35.22.
Soit un opérateur différentiel L, soit u la solution de $Lu = f$ et un schéma numérique (L_h, Ω_h) pour cette équation. Nous notons u_h la solution de $L_h u_h = f$. Alors nous avons

$$L_h e_h = \tau_h \tag{35.103}$$

Et si de plus L_h est inversible,

$$\|e_h\| \leqslant \|L_h^{-1}\|\|\tau_h\|. \tag{35.104}$$

Démonstration. Par définition u_h est solution de $L_h u_h = f$ en tant que fonction sur Ω_h. Nous avons donc

$$L_h e_h = L_h u_h - L_h u \tag{35.105}$$

où u doit être compris comme la restriction de u à Ω. En appliquant au point x_i,

$$(L_h e_h)(x_i) = \underbrace{(L_h u_h)(x_i)}_{=f_i} - (L_h u)(x_i), \tag{35.106}$$

mais $f_i = (Lu)(x_i)$ parce que u est solution de $Lu = f$. Donc

$$(L_h e_h)(x_i) = (Lu)(x_i) - (L_h u)(x_i) = \tau_h(x_i). \tag{35.107}$$

Si la matrice L_h est inversible nous avons $e_h = L_h^{-1}\tau_h$ et donc

$$\|e_h\| \leqslant \|L_h^{-1}\|\|\tau_h\| \tag{35.108}$$

par le lemme 10.66. □

Bien entendu, en tant qu'opérateur linéaire sur un espace de dimension finie, l'opérateur L_h^{-1} est borné pour chaque h. Mais si il n'y a pas une borne uniforme en h, alors le lemme 35.22 dit qu'il n'y a pas d'espoir de majorer $\|e_h\|$ de façon à passer à la limite $\lim_{h\to 0}\|L_h^{-1}\|$.

Définition 35.23.
*Un schéma numérique est **stable** si il existe une constante $C > 0$ indépendante de h telle que $\|L_h^{-1}\| \leqslant C$.*

THOooEPQQooUQMcgF

Théorème 35.24.
En deux parties.

(1) Si un schéma discret est consistant et stable, alors il est convergent.

(2) Si de plus il est consistant à l'ordre p, alors il est convergent à l'ordre p.

Démonstration. Nous savons du lemme 35.22 (qui s'applique parce que l'inversibilité de L_h est dans la définition de la stabilité) que $\|e_h\| \leqslant \|L_h^{-1}\|\|\tau_h\|$ et que $\|L_h^{-1}\| \leqslant C$. En passant à la limite [7],

$$\lim_{h\to 0} \|e_h\| \leqslant C \lim_{h\to 0} \|\tau_h\| = 0. \tag{35.109}$$

La dernière limite est le fait que le schéma soit consistant. Le schéma est donc convergent.

Si de plus il est consistant à l'ordre p, alors

$$\|e_h\| \leqslant C\|\tau_h\| \leqslant C' h^p \tag{35.110}$$

Il est donc également convergent à l'ordre p. □

35.3.4 Exemple : schéma à cinq points, laplacien en croix

Nous avions développé le schéma dont l'opérateur sur Ω_h est (voir (35.64))

$$(L_h u_h)(x_i, y_j) = \frac{1}{h^2}\big(- u_{i+1,j} - u_{i,j+1} + 4u_{i,j} - u_{i-1,j} - u_{i,j-1}\big). \tag{35.111}$$

Proposition 35.25.
Le schéma est :

(1) consistant à l'ordre 2,

(2) stable pour la norme uniforme et

$$\|L_h^{-1}\|_\infty \leqslant \frac{1}{8}, \tag{35.112}$$

(3) convergent à l'ordre 2 *pour la norme* $\|.\|_\infty$.

Démonstration. Cet opérateur avait été construit de telle sorte à avoir (voir (35.63))

$$(\Delta u)(x_i, y_j) = (L_h u)(x_i, y_i) + h^2 R(x, y, h) \tag{35.113}$$

où R peut être majoré indépendamment de h. En tant que fonctions sur Ω_h nous avons

$$\tau_h = \Delta u - L_h u = h^2 R(x, y, h), \tag{35.114}$$

et donc $\|\tau_h\|_\infty \leqslant Ch^2$, parce que le lemme 35.3(3) donne aussi

$$\|R\|_\infty \leqslant C \max\{\|\frac{\partial^4 u}{\partial x^4}\|_\infty, \|\frac{\partial^4 u}{\partial y^4}\|_\infty\}. \tag{35.115}$$

En ce qui concerne la stabilité nous allons utiliser le théorème 34.169. Nous considérons la fonction

$$g(x, y) = -\frac{1}{4}(x^2 + y^2), \tag{35.116}$$

qui vérifie $-\Delta g = 1$ sur le carré $[0, 1]^2$. Nous considérons le vecteur g_h d'indices $(i, j) \mapsto g_{ij} = g(x_i, y_j)$ sur lequel nous calculons L_h :

$$(L_h g_h)_{ij} = \frac{1}{h^2}(-g_{i+1,j} - g_{i,j+1} + 4g_{ij} - g_{i-1,j} - g_{i,j-1}); \tag{35.117}$$

en remplaçant les g par leurs valeurs en termes de x_i, x_{i-1}, x_{i+1}, y_j, y_{j-1} et y_{j+1}, et en tenant compte du fait que $x_k = kh$ et $y_l = lh$, nous avons :

$$(L_h g_h)_{ij} = \frac{1}{4}\big((i+1)^2 - j^2 + i^2 + (j+1)^2 - 4i^2 - 4j^2 + (i-1)^2 + j^2 + i^2 + (j-1)^2\big) = 1. \tag{35.118}$$

Donc $L_h g_h = 1$.

7. Toutes les limites $h \to 0$ sont en réalité des limites $h \to 0^+$, mais nous allégeons cette notation.

Comme L_h est une M-matrice (proposition 35.12(2)), le théorème 34.169 nous dit alors que L_h^{-1} vérifie

$$\|L_h^{-1}\|_\infty \leqslant \|g_h\|_\infty. \tag{35.119}$$

Mais

$$\|g_h\|_\infty \leqslant \|g\|_\infty = g\big(\frac{1}{2},\frac{1}{2}\big) = \frac{1}{8}. \tag{35.120}$$

Notons que cela est bien une inégalité et non une égalité parce que rien n'assure que le point $(1/2, 1/2)$ soit sur le maillage ; donc rien n'assure que la valeur $g(1/2, 1/2)$ ne soit parmi les valeurs du vecteur discrétisé g_h.

Notre schéma numérique est stable et consistant à l'ordre 2 pour la norme $\|.\|_\infty$. Le théorème 35.24 dit alors que le schéma est convergent à l'ordre 2 pour la même norme. □

35.4 Autres laplaciens

Nous avons vu le laplacien en croix (35.65) EQooOBWUooDEWKYv

$$(\Delta_h f)(x,y) = \frac{1}{h^2}\big(-4f(x,y) + f(x+h,y) + f(x-h,y) \tag{35.121a}$$

$$+ f(x,y+h) + f(x,y-h)\big) \tag{35.121b}$$

qui vérifie

$$\Delta_h f = \Delta f + Kh^2, \tag{35.122}$$ EQooQITHooZYJlVa

ainsi que le laplacien en carré (35.74) EQooUOYVooRpAMOC

$$(\Delta'_h f)(x,y) = \frac{1}{2h^2}\Big(-4f(x,y) + f(x+h,y+h) + f(x-h,y+h) \tag{35.123a}$$

$$+ f(x+h,y-h) + f(x-h,y-h)\Big) \tag{35.123b}$$

qui vérifie également $\Delta'_h f = \Delta f + Kh^2$.

À priori toutes combinaisons de la forme

$$a\Delta_h + b\Delta'_h \tag{35.124}$$

avec $a + b = 1$ sont valable comme tentative de discrétiser le laplacien. Ce sont des schémas à 9 points. Évidemment la matrice L_h correspondante va être moins creuse, mais nous pouvons espérer ajuster a et b de telle sorte à obtenir une consistance d'un ordre supérieur à 2.

Nous allons développer les $\Delta_h f$ et $\Delta'_h f$ à l'ordre 4 (reste à l'ordre 6). Quelques remarques avant de commencer.

(1) Allez relire la proposition 12.350 et les notations qui vont avec pour comprendre les différentielles.

(2) Écrivez les formules du type (12.971) pour $d^2 f$ et $d^4 f$.

(3) Allez relire le développement de Taylor du théorème 12.465.

(4) À l'ordre zéro, il n'y a rien, parce que le terme $-4f(x,y)$ compense les quatre termes d'ordre zéro des autres termes.

(5) Aux ordres impairs, il n'y a rien. En effet, prenons un nombre impair l et la formule

$$(d^l f)_x(h,\dots,h) = \sum_{i_1,\dots,i_l} h_{i_1}\dots h_{i_l}\frac{\partial^l f}{\partial x_{i_1}\dots\partial x_{i_l}}(x). \tag{35.125}$$

Nous avons

$$(d^l f)_x(h,\dots,h) + (d^l f)_x(-h,\dots,-h) = 0. \tag{35.126}$$

Or dans les expressions (35.123) et (35.121), les termes arrivent par paires opposées.

Commençons par calculer $h^2(\Delta_h f)(x,y)$.

Ordre 4 . Le premier terme est :

$$(d^2 f)_{(x,y)}\big((h,0),(h,0)\big) = h^2 (d^2 f)_{(x,y)}\big((1,0),(1,0)\big). \tag{35.127}$$

La formule (12.971) à peine adaptée permet de calculer ça explicitement.

Il y a encore les termes du même type avec $(0,1)$, $(-1,0)$ et $(0,-1)$.

Ordre 4 Cette fois, ce sont 4 termes du type

$$h^4 (d^4 f)_{(x,y)}\big((1,0),(1,0),(1,0),(1,0)\big) \tag{35.128}$$

à calculer.

Cela fait beaucoup de termes à calculer. Je vous laisse vous persuader que le programme suivant en Sage nous donne les coefficients.

```
#! /usr/bin/env python3
# -*- coding: utf8 -*-


x,y=var("x,y")

C=[]
C.append(x)
C.append(y)

def coef4(h):
    S=0
    for i in [0,1]:
        for j in [0,1]:
            for k in [0,1]:
                for l in [0,1]:
                    S=S+h[i]*h[j]*h[k]*h[l]*C[i]*C[j]*C[k]*C[l]
    return S

def coef2(h):
    S=0
    for i in [0,1]:
        for j in [0,1]:
            S=S+h[i]*h[j]*C[i]*C[j]
    return S

cross=[]
square=[]

cross.append(  [1,0] )
cross.append(  [-1,0] )
cross.append(  [0,1] )
cross.append(  [0,-1] )

square=[]
square.append(  [-1,-1] )
square.append(  [1,-1] )
square.append(  [-1,1] )
```

```
square.append(  [1,1] )


print("Cross scheme :")

K=sum(  coef2(v) for v in cross   )
L=sum(  coef4(v) for v in cross   )
print(K)
print(L)

print("square scheme :")

K=sum(  coef2(v) for v in square   )
L=sum(  coef4(v) for v in square   )
print(K)
print(L)
```

tex/sage/coefs.sage

Le résultat est, en utilisant la formule

$$\partial_x^4 f + 2\partial_{xxyy}^4 f + \partial_y^4 f = \Delta\Delta f, \tag{35.129}$$

nous avons

$$(\Delta_h f)(x,y) = \frac{1}{2}(2\partial_x^2 f + 2\partial_y^2 f)(x,y) + \frac{1}{4!}2h^2(\partial_x^4 f + \partial_y^4 f)(x,y) + Kh^4 \tag{35.130a}$$

$$= (\Delta f)(x,y) + \frac{1}{12}h^2(\partial_x^4 f + \partial_y^4 f)(x,y) + Kh^4 \tag{35.130b}$$

$$= \Delta f + \frac{h^2}{12}\Delta\Delta f - \frac{h^2}{12}2\partial_{xxyy}^4 f + Kh^4 \tag{35.130c}$$

$$= \Delta f + \frac{h^2}{12}\Delta\Delta f - \frac{h^2}{6}\partial_{xxyy}^4 f + Kh^4 \tag{35.130d}$$

où K est une constante qui peut être majorée en termes des dérivées quatrièmes de f. En particulier la plus grande des normes supremum de ces dérivées.

Le même genre de calculs donnent

$$(\Delta'_h f)(x,y) = \frac{1}{2}\Big[\frac{1}{2}4\Delta f + \frac{h^2}{4!}(4\partial_x^4 f + 24\partial_x^2\partial_y^2 f + 4\partial_y^2 f)\Big] + Kh^4. \tag{35.131}$$

Ça donne :

$$(\Delta'_h f) = \Delta f + \frac{h^2}{12}\Delta\Delta f + \frac{h^2}{3}\partial_{xxyy}^4 f + Kh^4 \tag{35.132}$$

avec redéfinition du K ; nous ne le préciserons plus à chaque fois.

Nous avons donc le résultat proposé dans [759] :

$$a\Delta_h f + b\Delta'_h f = (a+b)\Delta f + (a+b)\frac{h^2}{12}\Delta^2 f + h^2\frac{1}{6}(a-2b)\partial_{xxyy}^4 f + Kh^4. \tag{35.133}$$

L'idée est d'appliquer ça à une fonction u qui vérifie l'équation différentielle $-\Delta u = f$ (attention au clash de notation pour f). Le mieux est de supprimer le terme en $\partial_{xxyy} f$ en demandant $a - 2b = 0$. Nous avons donc à résoudre le système

$$\begin{cases} a + b = 1 & \text{(35.134a)} \\ a - 2b = 0. & \text{(35.134b)} \end{cases}$$

Qui propose une décomposition PLU pour résoudre ce système linéaire ? Quelle que soit la manière, la solution est

$$a = \frac{2}{3}, \quad b = \frac{1}{3}. \tag{35.135}$$

Nous allons donc étudier la discrétisation à neuf points

$$L_h = \frac{2}{3}\Delta_h + \frac{1}{3}\Delta'_h. \tag{35.136}$$

En faisant quelques additions nous trouvons que l'opération

$$(L_h u)(x_i, y_j) = \frac{1}{6h^2}\Big(-20u_{ij} + 4\big(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}\big) + u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1,j-1}\Big) \tag{35.137}$$

vérifie

$$L_h f = \Delta f + \frac{h^2}{12}\Delta^2 f + Kh^4. \tag{35.138}$$

35.4.1 Travail avec le laplacien à 9 points

Nous allons écrire un schéma numérique pour l'équation différentielle $-\Delta u = f$ utilisant la discrétisation à 9 points du laplacien. Nous recopions ses propriétés fondamentales (35.136), (35.137), (35.138) :

$$L_h = \frac{2}{3}\Delta_h + \frac{1}{3}\Delta'_h, \tag{35.139}$$

et

$$(L_h u)(x_i, y_j) = \frac{1}{6h^2}\Big(-20u_{ij} + 4\big(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}\big) + u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1,j-1}\Big), \tag{35.140}$$

et

$$L_h f = \Delta f + \frac{h^2}{12}\Delta^2 f + Kh^4. \tag{35.141}$$

Nous appliquons (35.141) à u et nous isolons Δu :

$$\Delta u = L_h u - \frac{h^2}{12}\Delta^2 u + Kh^4 = \frac{1}{6h^2}T_h u + \frac{h^2}{12}\Delta f + Kh^4 \tag{35.142}$$

où nous avons utilisé $\Delta^2 u = -\Delta f$ et avons noté

$$T_h u = -20u_{ij} + 4(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) + u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1,j-1}. \tag{35.143}$$

Nous imposons maintenant $\Delta u = -f$ en écrivant

$$\frac{1}{6h^2}T_h u = -f - \frac{h^2}{12}\Delta f + \alpha(h)h^4. \tag{35.144}$$

Une idée est de remplacer Δf par son approximation en croix (35.122) :

$$T_h u = -6h^2 f - \frac{h^4}{2}(\Delta_h f + Kh^2) + \alpha(h)h^6 \tag{35.145}$$

Avec quelques calculs nous trouvons le schéma numérique suivant :

$$20u_{ij} - 4(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} - u_{i,j-1}) - u_{i+1,j+1} - u_{i-1,j+1} - u_{i+1,j-1} - u_{i-1,j-1} \tag{35.146a}$$

$$= \frac{h^2}{2}(8f_{ij} + f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1}) + \alpha(h)h^6. \tag{35.146b}$$

En oubliant le terme en $\alpha(h)h^6$, nous obtenons un système d'équations linéaires.

¡¡ Avertissement/question à la lectrice !! 35.26
Il me semble que ce schéma donne une convergence d'ordre 6. *C'est correct ?*

Chapitre 36

Variables aléatoires et théorie des probabilités

36.1 Espace de probabilité

Définition 36.1.
*Une **mesure de probabilité** sur un espace mesurable*[1] $(\Omega, \mathcal{A})$ *est une mesure positive telle que* $P(\Omega) = 1$. *Dans ce cas, le triple* $(\Omega, \mathcal{A}, P)$ *est un **espace de probabilité**.*

Un point $\omega \in \Omega$ est une **observation**, une partie mesurable $A \in \mathcal{A}$ est un **événement**. L'ensemble $A \cup B$ représente l'événement A ou B tandis que l'ensemble $A \cap B$ représente l'événement A et B.

Si les A_n sont des événements, nous avons défini en 9.43 limite supérieure et la limite inférieure de la suite A_n par

$$\limsup_{n\to\infty} A_n = \bigcap_{n\geqslant 1} \bigcup_{k\geqslant n} A_k \tag{36.1}$$

et

$$\liminf_{n\to\infty} A_n = \bigcup_{n\geqslant 1} \bigcap_{k\geqslant n} A_k \tag{36.2}$$

Si $\omega \in \liminf A_n$, alors ω réalise tous les A_n sauf un nombre fini.

Nous avons

$$\limsup A_n = \{\omega \in \Omega \text{ tel que } \omega \in A_n \text{pour une infinité de } n\}. \tag{36.3}$$

Théorème 36.2 (Borel-Cantelli)**.**
Si

$$\sum_{n=1}^{\infty} P(A_n) < \infty \tag{36.4}$$

alors $P(\limsup A_n) = 0$.

Démonstration. La condition $\sum_{n\geqslant 1} P(A_n) < \infty$ signifie que la fonction

$$\varphi = \sum_{n\geqslant 1} \mathbb{1}_{A_n} \tag{36.5}$$

est P-intégrable. Par conséquent, elle est finie presque partout (au sens de P), c'est-à-dire

$$P(\varphi = \infty) = 0. \tag{36.6}$$

Les points ω sur lesquels $\varphi(\omega) = \infty$ sont ceux tels que

$$\sum_{n\geqslant 1} \mathbb{1}_{A_n}(\omega) = \infty, \tag{36.7}$$

1. Espace mesurable : 14.1, mesure positive : 14.18.

c'est-à-dire les ω qui appartiennent à une infinité d'ensembles A_n, ou encore les $\omega \in \limsup A_n$. Nous avons donc montré que

$$\{\omega \text{ tel que } \varphi(\omega) = \infty\} = \{\omega \in \Omega \text{ tel que } \omega \in A_n \text{pour une infinité de } n\} = \limsup A_n. \tag{36.8}$$

Or l'hypothèse signifie que la probabilité du membre de gauche est nulle. □

CORooUWLZooFLYmcY

Corolaire 36.3.
Si $\sum_{n=1}^{\infty} P(\complement A_n) < \infty$, alors presque surement tous les B_n sont réalisés à l'exception d'un nombre fini.

36.2 Variables aléatoires

DEFooSZRXooQUSZYU

Définition 36.4.
*Une **variable aléatoire** est une application mesurable*

$$X : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d)). \tag{36.9}$$

Nous convenons que $\mathbb{R}^1 = \bar{\mathbb{R}}$, c'est-à-dire que dans le cas où la variable aléatoire X est réelle, nous acceptons les valeurs $\pm\infty$.

DEFooRNKZooRczFwB

Définition 36.5 (densité[760])**.**
*Une variable aléatoire réelle $X : \Omega \to \mathbb{R}^d$ est **absolument continue** si il existe une fonction positive et intégrable $f : \mathbb{R} \to \mathbb{R}$ telle que pour tout borélien $A \subset \mathbb{R}^d$,*

$$P(X \in A) = \int_{\mathbb{R}^d} \mathbb{1}_A(x) f(x) dx. \tag{36.10}$$

*Nous disons alors que f est la **densité** de X.*

Cela ne devrait pas être sans rappeler la proposition 14.316 et la définition 14.213.

PROBooSWDQooHlSZPD

¡¡ Avertissement/question à la lectrice !! 36.6
Je ne suis pas 100% certain de la proposition 36.7. Écrivez-moi si vous avez un doute ou une preuve.

PROPooDXCVooBxaToe

Proposition 36.7 ([1])**.**
Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire à densité, alors sa densité est donnée par

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto P(X < x). \end{aligned} \tag{36.11}$$

36.2.1 Loi d'une variable aléatoire

DEFooBKJVooRJdMeA

Proposition-Définition 36.8 (Loi d'une variable aléatoire[761])**.**
Soient un espace probabilisé $(\Omega, \mathcal{F}, P)$ ainsi qu'une variable aléatoire $X : \Omega \to \mathbb{R}^n$. En définissant

$$\begin{aligned} P_X : \mathcal{B}or(\mathbb{R}^n) &\to \mathbb{R}^+ \\ B &\mapsto (P \circ X^{-1})(B), \end{aligned} \tag{36.12}$$

le triple $(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), P_X)$ est un espace probabilisé.

*La mesure P_X est la **loi** de la variable aléatoire X*

36.2.2 Indépendance

DEFooVYCUooKWvReO

Définition 36.9 ([762]).
*Nous disons que les événements $A_1, \dots, A_n$ sont **indépendants** si pour tout choix $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ nous avons*

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}). \tag{36.13}$$

*Les sous tribus $\mathcal{A}_1, \dots, \mathcal{A}_n$ sont **indépendantes** si pour tout choix $A_i \in \mathcal{A}_i$, les événements A_i sont indépendants.*

Exemple 36.10.
Soit $\Omega = [0,1] \times [0,1]$ muni de la mesure de Lebesgue. Soient $A = [0,a] \times [0,1]$ et $B = [0,1] \times [0,b]$. Nous avons $P(A) = a$ et $P(B) = b$ ainsi que $P(A \cap B) = ab$. △

LemIndepEvenCompl

Lemme 36.11.
Les événements $(A_i)_{i=0,\dots,n}$ sont indépendants si et seulement si les événements que nous obtenons en remplaçant certains des A_i par $\complement A_i$ le sont.

Démonstration. Sans perte de généralité, nous pouvons nous contenter de prouver que les événements $\complement A_0, A_1, \dots, A_n$ sont indépendants sous l'hypothèse que les événements $A_0, A_1, \dots, A_n$ sont indépendants. Soit I un sous-ensemble de $\{1, \dots, n\}$. Nous avons

$$\begin{aligned}
P\big(\complement A_0 \bigcap_{i\in I} A_i\big) &= P\big(\bigcap_{i\in I} A_i \backslash \bigcap_{i\in I} A_i \cap A_0\big) && \text{(36.14a)}\\
&= P\big(\bigcap_{i\in I} A_i\big) - P\big(\bigcap_{i\in I} A_i \cap A_0\big) && \text{(36.14b)}\\
&= P\big(\bigcap_{i\in I} A_i\big)\big(1 - P(\complement A_0)\big) && \text{(36.14c)}\\
&= P\big(\bigcap_{i\in I} A_i\big) P(\complement A_0). && \text{(36.14d)}
\end{aligned}$$

□

LemTribIndepProdProb

Lemme 36.12.
Les tribus $\mathcal{A}_1, \dots, \mathcal{A}_n$ sont indépendantes si et seulement si

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \dots P(A_n) \tag{36.15}$$

pour tout $A_i \in \mathcal{A}_i$.

Démonstration. L'implication dans le sens direct découle immédiatement des définitions.

Nous supposons avoir un choix $(A_i)_{i=1,\dots,n}$ avec $A_i \in \mathcal{A}_i$ et nous devons montrer que ces événements sont indépendants, c'est-à-dire que si $J \subset \{1, \dots, n\}$ alors les événements $(A_j)_{j\in J}$ sont indépendants. Sans perte de généralité, nous pouvons supposer que si $i \notin J$, $A_i = \Omega$. Alors nous avons

$$P\big(\bigcap_{j\in J} A_j\big) = P\big(\bigcap_{i=1}^{n} A_i\big) = \prod_{i=1}^{n} P(A_i) = \prod_{j\in J} P(A_j) \tag{36.16}$$

parce que $P(A_i) = P(\Omega) = 1$ lorsque i n'est pas dans J. □

DefNJUkotc

Définition 36.13.
*Nous disons que les variables aléatoires $X_k \colon \Omega \to \mathbb{R}^d$ sont **indépendantes** lorsque les tribus engendrées[2] $\mathcal{A}_{X_1}, \dots, \mathcal{A}_{X_n}$ le sont.*

2. Définition 14.6.

Remarque 36.14.
Il n'a de sens de dire que X_1 et X_2 sont indépendants que si X_1 et X_2 sont des applications dont l'espace de départ est identique.

Si nous voulons modéliser le jet de deux pièces indépendantes, le mauvais choix est de faire $\Omega = \{0,1\}$, y mettre la mesure d'équiprobabilité, et de considérer les deux variables aléatoires

$$X_i(\omega) = \begin{cases} f & \text{si } \omega = 0 \\ p & \text{si } \omega = 1. \end{cases} \tag{36.17}$$

Ces deux variables sont évidemment pas indépendantes. Il faut poser $\Omega = \{0,1\} \times \{0,1\}$, y mettre la mesure d'équiprobabilité et poser

$$X_1(x,y) = \begin{cases} f & \text{si } x = 0 \\ p & \text{si } x = 1 \end{cases}, \tag{36.18}$$

$$X_2(x,y) = \begin{cases} f & \text{si } y = 0 \\ p & \text{si } y = 1 \end{cases}, \tag{36.19}$$

Ces variables aléatoires sont indépendantes. Par exemple

$$X_1^{-1}\{p\} = \{(1,0),(1,1)\} \tag{36.20a}$$
$$X_2^{-1}\{p\} = \{(0,1),(1,1)\} \tag{36.20b}$$

et on a bien

$$P\big(X_1^{-1}\{p\} \cap X_2^{-1}\{p\}\big) = P\{(1,1)\} = \frac{1}{4} \tag{36.21}$$

ainsi que

$$P\{X_i^{-1}(p)\} = \frac{1}{2} \tag{36.22}$$

pour $i = 1$ et $i = 2$.

PropMLbfRTk

Proposition 36.15.
Soient $(X_k \colon \Omega \to \mathbb{R}^{d_k})$ des variables aléatoires indépendantes.

(1) Si $B_k \in \mathcal{B}or(\mathbb{R}^{d_k})$. Alors

$$P(X_k \in B_k \forall k \leqslant n) = P(X_1 \in B_1) \dots P(X_n \in B_n). \tag{36.23}$$

ItemHRjuTTii

(2) Les événements $\{X_i \in B_i\}$ sont indépendants.

ItemHRjuTTiii

(3) Les tribus engendrées par des X_i et d'autres sont indépendantes. Plus précisément, si I et J sont deux ensembles disjoints de $\mathbb{N}$ alors les tribus

$$\sigma(\{X_i, i \in I\}) \tag{36.24}$$

et

$$\sigma(\{X_i, i \in J\}) \tag{36.25}$$

sont indépendantes.

Démonstration. Lorsque nous écrivons $X_i \in B_i$, nous parlons de l'événement

$$(X_i \in B_i) = \{\omega \in \Omega \text{ tel que } X_i(\omega) \in B_i\} = X_i^{-1}(B_i) \in \mathcal{A}_{X_i}. \tag{36.26}$$

Vu que par hypothèse les tribus $(\mathcal{A}_i)$ sont indépendantes, le lemme 36.12 nous montre que

$$P\big(\bigcap_{i=1}^{n} X_i \in B_i\big) = \prod_i P(X_i \in B_i). \tag{36.27}$$

Il reste à voir que l'ensemble $X_i^{-1}(B_i)$ fait partie de la tribu $\mathcal{A}$ de départ. Cela est la définition du fait que l'application X_i soit une variable aléatoire : elle doit être mesurable en tant qu'application

$$X_i\colon (\Omega,\mathcal{A})\to(\mathbb{R}^d,\mathcal{B}\text{or}(\mathbb{R}^d)). \tag{36.28}$$

Les affirmations (2) et (3) ne sont que des façons alternatives d'exprimer la même chose. □

Proposition 36.16.
Les événements $(A_i)_{i=1,\dots,n}$ sont indépendants[3] *si et seulement si les variables aléatoires associées $\mathbb{1}_{A_1},\dots,\mathbb{1}_{A_n}$ le sont.*

Démonstration. La tribu engendrée par la variable aléatoire $\mathbb{1}_{A_k}$ est

$$\mathcal{A}_{\mathbb{1}_{A_k}}=\{\varnothing,A_k,\complement A_k,\Omega\}. \tag{36.29}$$

EqtribAAimtu

En effet si $1\in B$, alors $A_i\subset\mathbb{1}_{A_i}^{-1}(B)$, et si $0\in B$, alors $\complement A_i\subset\mathbb{1}_{A_i}^{-1}(B)$. Les éléments 0 et 1 sont tous deux soit dans B, soit hors de B. Cela donne les 4 possibilités énumérées dans (36.29).

Supposons que les événements (A_i) sont indépendants. Nous devons vérifier que les tribus le soient, c'est-à-dire que les événements A_i et $\complement A_j$ sont indépendants. Cela est une conséquence du lemme 36.11. □

PROPooSYNQooAlpGdI

Proposition 36.17.
Soient des variables aléatoires indépendantes $X_k\colon\Omega\to\mathbb{R}^{d_k}$ et des fonctions boréliennes $f_k\colon\mathbb{R}^{d_k}\to\mathbb{R}^{p_k}$. Alors les variables aléatoires

$$f_k\circ X_k \tag{36.30}$$

sont indépendantes.

Démonstration. Le théorème 14.119 assure que les applications

$$f_k\circ X_k\colon\Omega\to\mathbb{R}^{p_k}: \tag{36.31}$$

sont $\mathcal{A}_{X_k}$-mesurables. En particulier pour tout borélien $B\subset\mathbb{R}^{p_k}$, nous avons $X_k^{-1}\circ f_k^{-1}(B)\in\mathcal{A}_{X_k}$. Nous avons donc

$$\sigma(f_k\circ X_k)\subset\sigma(X_k), \tag{36.32}$$

et par conséquent les tribus $\sigma(f_k\circ X_k)$ sont indépendantes étant donné que les tribus $\sigma(X_k)$ le sont. □

LEMooWBAZooLVVRjO

Lemme 36.18.
Soient des variables aléatoires $X,Y\colon\Omega\to E$ que nous supposons être indépendantes et identiquement distribuées. Si $f\colon E\to F$ est mesurable,alors les variables aléatoires $f\circ X$ et $f\circ Y$ sont indépendantes et identiquement distribuées.

LemHOjqqw

Lemme 36.19 (Lemme de regroupement[763])**.**
Soit $(\Omega,\mathcal{A},P)$ un espace de probabilité et $(\mathcal{A}_i)_{i\in I}$ une famille de tribus indépendantes dans $\mathcal{A}$. Si $(M_j)_{j\in J}$ est une partition de I, alors les tribus

$$\mathcal{B}_j=\sigma\Big(\bigcup_{i\in M_j}\mathcal{A}_i\Big) \tag{36.33}$$

sont indépendantes.

Si les variables aléatoires $\{X_1,X_2,X_3,X_4,X_5\}$ sont indépendantes, et si f et g sont des fonctions mesurables, alors les variables aléatoires $f(X_2,X_3,X_5)$ et $g(X_1,X_4)$ sont indépendantes.

3. Événements indépendants, définition 36.9.

36.2.3 Lois conjointes et indépendance

Définition 36.20 (Loi conjointe et loi marginale)**.**
*Si nous considérons n variables aléatoires réelles $X_1, \dots, X_n : \Omega \to \mathbb{R}$, la loi du n-uplet $X = (X_1, \dots, X_n)$ est une variable aléatoire $X : \Omega \to \mathbb{R}^n$ appelée la **loi conjointe** des lois X_i. Dans ce cas, les variables aléatoires X_i elles-mêmes sont dites lois **marginales** de X.*

PROPooCXSWooByucuN

Proposition 36.21 ([1])**.**
Soient un espace de probabilité Ω ainsi que deux variables aléatoires indépendantes

$$X_1 : \Omega \to \mathbb{R}^k \tag{36.34a}$$
$$X_2 : \Omega \to \mathbb{R}^l \tag{36.34b}$$

Nous supposons que X_1 et X_2 ont des densités $f_1 : \mathbb{R}^k \to \mathbb{R}$ et $f_2 : \mathbb{R}^l \to \mathbb{R}$.

(1) En notant λ la mesure de Lebesgue sur $\mathbb{R}^n$, avons égalité des mesures[4]

$$P_X = f \cdot \lambda = P_{X_1} \otimes P_{X_2} \tag{36.35}$$

(2) La variable aléatoire

$$\begin{aligned} X : \Omega &\to \mathbb{R}^{k+l} \\ \omega &\mapsto \big(X_1(\omega), X_2(\omega)\big) \end{aligned} \tag{36.36}$$

a pour densité la fonction

$$\begin{aligned} f : \mathbb{R}^k \times \mathbb{R}^l &\to \mathbb{R} \\ (x, y) &\mapsto f_1(x) f_2(y). \end{aligned} \tag{36.37}$$

Démonstration. Nous commençons par considérer un borélien de la forme $B_1 \times B_2$ avec $B_1 \in \mathcal{B}\text{or}(\mathbb{R}^k)$ et $B_2 \in \mathcal{B}\text{or}(\mathbb{R}^l)$. Par définition de P_X et par indépendance (proposition 36.15) nous avons d'une part

$$P_X(B_1 \times B_2) = P(X \in B_1 \times B_2) \tag{36.38a}$$
$$= P(X_1 \in B_1 \cap X_2 \in B_2) \tag{36.38b}$$
$$= P(X_1 \in B_1) P(X_2 \in B_2). \tag{36.38c}$$

Et d'autre part nous pouvons utiliser le théorème de Fubini 14.301 parce que f_1 et f_2 sont positives :

$$(f \cdot \lambda)(B_1 \times B_2) = \int_{B_1 \times B_2} f(x, y)\, dx\, dy \tag{36.39a}$$
$$= \int_{B_1 \times B_2} f_1(x) f_2(y)\, dx\, dy \tag{36.39b}$$
$$= \int_{B_1} f_1(x) \int_{B_2} f_2(x) \tag{36.39c}$$
$$= P(X_1 \in B_1) P(X_2 \in B_2). \tag{36.39d}$$

Les mesures P_X et $f \cdot \lambda$ sont donc égales sur ces boréliens. La partie unicité du théorème 14.246(1) dit qu'alors elles sont égales sur tous les boréliens et qu'en plus, elles sont égales à $P_{X_1} \otimes P_{X_2}$. □

PROPooANJQooPjBPSy

Proposition 36.22 ([1])**.**
Soient des variables aléatoires indépendantes $X_i : \Omega \to \mathbb{R}$ de densités $f_i : \mathbb{R} \to \mathbb{R}$. Nous considérons la variable aléatoire

$$\begin{aligned} X : \Omega &\to \mathbb{R}^n \\ \omega &\mapsto \big(X_1(\omega), \dots, X_n(\omega)\big). \end{aligned} \tag{36.40}$$

Nous avons :

4. Voir les définitions 36.8 et 14.212.

(1) *La variable aléatoire X admet la densité*

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \prod_{i=1}^{n} f_i(x_i). \end{aligned} \tag{36.41}$$

ITEMooNIMQooGqFKNn

(2) *Nous avons, pour tout borélien $B \subset \mathbb{R}^n$:*

$$P(X \in B) = \int_B \prod_{i=1}^{n} f_i(x_i)dx \tag{36.42}$$

Démonstration. Il s'agit de faire une récurrence sur la proposition 36.21. □

PropPXXXPXPXPX

Proposition 36.23.
Les variables aléatoires $\{X_i\}$ sont indépendantes[5] *si et seulement si*

$$P_{(X_1,\dots,X_n)} = P_{X_1} \otimes \dots \otimes P_{X_n} \tag{36.43}$$

où les mesures P_{X_i} sont les lois des variables aléatoires introduites en 36.8 et le produit des mesures est définit dans 14.246.

DefFonrepConj

Définition 36.24 (densité conjointe)**.**
Soient $\{X_i\}_{1\leqslant i\leqslant n}$ des variables aléatoires réelles (pas spécialement indépendantes). Une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ qui satisfait

(1) *$f(x_1,\dots,x_n) \geqslant 0$ pour tout $(x_1,\dots,x_n) \in \mathbb{R}^n$,*
(2) *$\int_{\mathbb{R}^n} f = 1$,*

ItemDefFonrepConjiii

(3) *Si les parties A_i sont mesurables dans $\mathbb{R}$, nous avons*

$$P(\bigcap_{i=1}^{n} X_i \in A_i) = \int_{\prod_i A_i} f(x_1,\dots,x_n)dx_1\dots dx_n \tag{36.44}$$

est nommée ***densité conjointe*** *de $X_1,\dots,X_n$*

PROBooPVIXooNLpmIn

¡¡ Avertissement/question à la lectrice !! 36.25
Il manque un théorème d'unicité pour la densité conjointe. Écrivez-moi si vous en connaissez un.

36.26.
Il y a une différence entre la densité conjointe de $X_1,\dots,X_n$ et la densité de la variable aléatoire $\omega \mapsto \big(X_1(\omega,\dots,X_n(\omega))\big)$. La densité de cette dernière est une fonction f qui doit satisfaire

$$P(X \in B) = \int_B f(x)dx \tag{36.45}$$

pour tout borélien B. La densité conjointe ne vérifie cela que pour les boréliens très particuliers qui se présentent comme produits cartésiens de boréliens de $\mathbb{R}$.

PropDensiteConjIndep

Proposition 36.27.
Si les variables aléatoires $X_1,\dots X_n$ sont indépendantes et ont des densités $f_{X_1},\dots,f_{X_n}$, alors la variable aléatoire conjointe $X = (X_1,\dots,X_n)$ a pour densité conjointe la fonction

$$f_X(x_1,\dots,x_n) = f_{X_1}(x_1)\dots f_{X_n}(x_n). \tag{36.46}$$

5. Variables aléatoires indépendantes, définition 36.13.

Démonstration. En partant de la définition de l'indépendance et de la fonction de densité conjointe, ainsi qu'en utilisant le théorème de Fubini,

$$\begin{aligned}\int_{A_1\times\ldots\times A_n} f_X(x_1,\ldots,x_n)dx_1\ldots dx_n &= P(X_1\in A_1,\ldots,X_n\in A_n)\\ &= P(X_1\in A_1)\ldots P(X_n\in A_n)\\ &= \left(\int_{A_1} f_{X_1}(x_1)dx_1\right)\ldots\left(\int_{A_n} f_{X_n}(x_n)dx_n\right)\\ &= \int_{A_1\times\ldots\times A_n} f_{X_1}(x_1)\ldots f_{X_n}(x_n)dx_1\ldots dx_n.\end{aligned}\tag{36.47}$$

La fonction $(x_1,\ldots,x_n)\mapsto f_{X_1}(x_1)\ldots f_{X_n}(x_n)$ vérifie donc la condition (3) de la définition 36.24. La vérification des autres conditions est immédiate. □

LEMooLNIZooDWUvRN

Lemme 36.28 ([764])**.**
Si les variables aléatoires $X_1,\ldots,X_n$ sont indépendantes[6] *et identiquement distribuées de densité*[7] *f. Nous considérons la variable aléatoire*

$$\begin{aligned}X\colon\Omega &\to \mathbb{R}^n\\ \omega &\mapsto \big(X_1(\omega),\ldots,X_n(\omega)\big).\end{aligned}\tag{36.48}$$

Alors la densité de variable aléatoire X est

$$\begin{aligned}g\colon\mathbb{R}^n &\to \mathbb{R}\\ x &\mapsto n!\prod_{i=1}^n f(x_i)\mathbb{1}_{\{x_1<\ldots<x_n\}}.\end{aligned}\tag{36.49}$$

Démonstration. Nous devons vérifier les différents points de la définition 36.24.

(1) $g\geqslant 0$ Celle-là, elle est facile.

(2) **Les parties $\mathbb{R}^n_\sigma$** Soit $\sigma\in S_n$ (une permutation de $\{1,\ldots,n\}$). Nous considérons les parties

$$\mathbb{R}^n_\sigma=\{x\in\mathbb{R}^n \text{ tel que } x_{\sigma(1)}<\ldots<x_{\sigma(n)}\},\tag{36.50}$$

et la bijection

$$\begin{aligned}\varphi\colon\mathbb{R}^n_{\mathrm{Id}} &\to \mathbb{R}^n_\sigma\\ x &\mapsto \big(x_{\sigma^{-1}(1)},\ldots,x_{\sigma^{-1}(n)}\big).\end{aligned}\tag{36.51}$$

EQooLQDGooQJNLsd

Nous avons

$$\mathbb{R}^n=\bigcup_{\sigma\in S_n}\mathbb{R}^n_\sigma\cup Z\tag{36.52}$$

où Z est de mesure (le Lebesgue) nulle. Cela a pour conséquence que

$$\int_{\mathbb{R}^n} s(x)dx=\sum_{\sigma\in S_n}\int_{\mathbb{R}^n_\sigma} s(x)dx\tag{36.53}$$

EQooIEWWooZgORTM

pour toute fonction s intégrable.

(3) **Une formule intermédiaire** Nous prouvons, pour tout $\sigma\in S_n$, la formule $\int_{\mathbb{R}^n_\sigma}\prod_{i=1}^n f(x_i)=\int_{\mathbb{R}^n_{\mathrm{Id}}}\prod_{i=1}^n f(x_i)$. Pour cela nous fixons $\sigma\in S_n$.
Vu que le produit est commutatif nous avons

$$\prod_{i=1}^n f\big(x_{\sigma^{-1}(i)}\big)=\prod_{i=1}^n f(x_i)\tag{36.54}$$

EQooJEDAooQMmHrb

6. Variables aléatoires indépendantes, définition 36.13.
7. Définition 36.5.

(4) **Intégrale 1** Nous prouvons que $\int_{\mathbb{R}^n} g = 1$. Pour cela nous utilisons le théorème de Fubini 14.304(3) :

$$\int_{\mathbb{R}^n} \prod_{i=1}^{n} f(x_i)dx = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \prod_{i=1}^{n} f(x_i)dx_n \right) dx_1 \dots dx_{n-1} \tag{36.55a}$$

$$= \int_{\mathbb{R}^{n-1}} \Big(\prod_{i=1}^{n-1} f(x_i) \underbrace{\int_{\mathbb{R}} f(x_n)dx_n}_{=1} \Big) dx_1 \dots dx_{n-1} \tag{36.55b}$$

Par récurrence sur n, nous déduisons que $\int_{\mathbb{R}^n} g = 1$.

(5) $P(X \in A)$
Soit un borélien $A \subset \mathbb{R}^n$. Vu qu'à partie de mesure nulle près, A est l'union disjointe des A_σ, nous avons, en utilisant la proposition 36.22(2),

$$P(X \in A) = \int_A \prod_i f(x_i)dx = \sum_{\sigma \in S_n} \int_{A_\sigma} \prod_i f(x_i)dx. \tag{36.56}$$

Pour un σ donné, nous considérons l'application φ de (36.51). Nous avons $A_\sigma = \varphi(A_{\mathrm{Id}})$ et donc

$$P(X \in A) = \sum_\sigma \int_{\varphi_\sigma(A_{\mathrm{Id}})} \prod_i f(x_i)dx. \tag{36.57}$$

Là-dessus, nous utilisons φ_σ comme changement de variables dans l'intégrale, théorème 14.281(2) :

$$P(X \in A) = \sum_\sigma \int_{\varphi_\alpha(A_{\mathrm{Id}})} \prod_i f(x_i)dx \tag{36.58a}$$

$$= \sum_\sigma \int_{A_{\mathrm{Id}}} f\big(\varphi_\alpha(x)_i\big)dx \tag{36.58b}$$

$$= \sum_\sigma \int_{A_{\mathrm{Id}}} \prod_i f(x_i)dx \tag{36.58c}$$

$$= n! \int_{A_{\mathrm{Id}}} \prod_i f(x_i)dx \qquad \text{cf. justif.} \tag{36.58d}$$

SUBEQooEJBQooFyLbct

$$= n! \int_A \prod_i f(x_i) \mathbb{1}_{\mathbb{R}^n_{\mathrm{Id}}}(x)dx. \tag{36.58e}$$

Justification pour (36.58d). Les termes de la somme sur σ sont tous les mêmes, et il y en a $n!$.

□

36.2.4 Densité jointe

PROPooGXEOooWMAuIp

Proposition 36.29 ([765]).
Soit un ouvert A de $\mathbb{R}$. Si $f : A \to \mathbb{R}$ est dérivable, alors f' est mesurable.

Si $X, Y : \Omega \to \mathbb{R}$ sont des variables aléatoires, nous définissons l'application

$$\begin{aligned} F_{X,Y} : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto P\big(X \leqslant x, Y \leqslant y\big). \end{aligned} \tag{36.59}$$

DEFooUOKQooMIjDkr

Définition 36.30 ([766]).
*Lorsque cela a un sens, nous définissons la **densité jointe** de X et Y par*[8]

$$f_{X,Y}(x, y) = (\partial^2_{1,2} F_{X,Y})(x, y). \tag{36.60}$$

8. Vous préférez peut-être la notatoin $\partial_{xy} F_{X,Y}$.

¡¡ Avertissement/question à la lectrice !! 36.31
Pour la proposition 36.32, vous pouvez essayer de répondre aux questions suivantes :

(1) ahttps://math.stackexchange.com/questions/4961906/random-variable-density-from-basis-o
(2) https://math.stackexchange.com/questions/4961074/is-the-join-density-the-density-of-th noredirect=1&lq=1

Bref. Soyez très prudent. N'ayez confiance en personne.

PROPooPDIWooHzjqcZ

Proposition 36.32 ([1]).
Soient deux variables aléatoires $X, Y : \Omega \to \mathbb{R}$. Nous supposons

(1) *que la densité jointe*[9] *$f_{X,Y}$ soit bien définie.*
(2) *Pour toute borélien Z de mesure de Lebesgue nulle dans $\mathbb{R}$, nous avons*

$$P(X \in Z) = P(Y \in Z) = 0. \tag{36.61}$$

Alors pour tout borélien $B \subset \mathbb{R}^2$, nous avons

$$P\big((X,Y) \in B\big) = \int_B f_{X,Y}. \tag{36.62}$$

Démonstration. Vu que $F_{X,Y}$ est croissante en ses deux variables, la fonction $f_{X,Y}$ est positive. Et en tant que dérivée d'une fonction mesurable, elle est également mesurable (proposition 36.29). Nous pouvons donc appliquer le théorème de Fubini-Tonelli 14.301 :

$$\begin{aligned}\int_B f_{X,Y} &= \int_{\mathbb{R}\times\mathbb{R}} f_{X,Y}(x,y)\mathbb{1}_B(x,y)dxdy &\text{(36.63a)}\\ &= \int_{\mathbb{R}}\left[\int_{\mathbb{R}} f_{X,Y}(x,y)\mathbb{1}_B(x,y)dy\right]dx. &\text{(36.63b)}\end{aligned}$$

(1) **Sur un rectangle** Nous supposons que soit un rectangle ouvert : $B =]a,b[\times]c,d[$. Les intégrales deviennent :

$$\int_B f_{X,Y} = \int_a^b\left[\int_c^d f_{X,Y}(x,y)dy\right]dx. \tag{36.64}$$

Nous nous fixons à présent x et nous nous concentrons sur l'intégrale intérieure : $\int_c^d f_{X,Y}(x,y)dy$. Nous posons

$$\begin{aligned}G_x : \mathbb{R} &\to \mathbb{R}\\ y &\mapsto (\partial_x F_{X,Y})(x,y),\end{aligned} \tag{36.65}$$

de telle sorte que $F_{X,Y}(x,y) = G'_x(y)$ et que

$$\begin{aligned}\int_c^d f_{X,Y}(x,y)dy &= \int_c^d G'_x(y)dy &\text{(36.66a)}\\ &= G_x(d) - G_x(c) &\text{(36.66b)}\\ &= (\partial_x F_{X,Y})(x,d) - (\partial_x F_{X,Y})(x,c). &\text{(36.66c)}\end{aligned}$$

Nous passons à l'intégrale sur x. Pour le premier terme, nous posons

$$\begin{aligned}H : \mathbb{R} &\to \mathbb{R}\\ x &\mapsto F(x,d),\end{aligned} \tag{36.67}$$

de telle sorte à avoir $(\partial_x F_{X,Y})(x,d) = H'(x)$ et

$$\int_a^b (\partial_x F_{X,Y})(x,d) = H(b) - H(a) = F_{X,Y}(b,d) - F_{X,Y}(a,d). \tag{36.68}$$

9. Définition 36.30.

En faisant de même pour le second terme, nous trouvons

$$\int_B f_{X,Y} = F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c) \tag{36.69a}$$
$$= P(X \leqslant b, Y \leqslant d) - P(X \leqslant a, Y \leqslant d) - P(X \leqslant b, Y \leqslant c) + P(X \leqslant a, Y \leqslant c). \tag{36.69b}$$

SUBEQooLUZJooTpqbGO

Nous divisons $A = \{(x,y) \in \mathbb{R}^2 \text{ tel que } x \leqslant b, y \leqslant d\}$ en 4 parties :

$$A_1 = \{x \leqslant a, c \leqslant y \leqslant d\} \tag{36.70a}$$
$$A_2 = \{a \leqslant x \leqslant b, c \leqslant y \leqslant d\} = B \tag{36.70b}$$
$$A_3 = \{a \leqslant x \leqslant b, y \leqslant c\} \tag{36.70c}$$
$$A_4 = \{x \leqslant a, y \leqslant c\}. \tag{36.70d}$$

Faites quelques dessins pour voir quelle partie ressemble à quoi. Il n'est pas trop compliqué de voir que

$$\{X \leqslant b, Y \leqslant d\} = \{(X,Y) \in A_4\} \cup \{(X,Y) \in A_1\}. \tag{36.71}$$

En faisant de même pour les différents termes de (36.69b), nous trouvons

$$\int_B f_{X,Y} = P\big((X,Y) \in A_2\big) = P\big((X,Y) \in B\big). \tag{36.72}$$

(2) **Théorème d'extension** Nous voulons utiliser le théorème d'extension des mesures 14.34 en considérant $\mathcal{E}$, l'ensemble des rectangles ouverts de $\mathbb{R}^2$ de coordonnées rationnelles [10]. Nous posons

$$\mu_1(B) = P\big((X,Y) \in B\big) = \big(P \circ (X,Y)^{-1}\big)(B) \tag{36.73}$$

EQooPFXUooQumsUl

et

$$\mu_2(B) = \int_B f_{X,Y} \tag{36.74}$$

pour tout borélien B de $\mathbb{R}^2$. Nous allons prouver que μ_1 et μ_2 sont des mesures [11].

(3) **μ_1 est une mesure** La formule (36.73) a un sens parce que si B est mesurable dans Ω, alors

$$(X,Y)^{-1}(B) \in \mathcal{B}\text{or}(\mathbb{R}) \otimes \mathcal{B}\text{or}(\mathbb{R}) \qquad \text{prop. 14.129} \tag{36.75a}$$
$$= \mathcal{B}\text{or}(\mathbb{R}^2) \qquad \text{prop. 14.131.} \tag{36.75b}$$

À partir de là, μ_1 est une mesure par la proposition 14.47. En effet P est une mesure sur Ω et $(X,Y)\colon \Omega \to \mathbb{R}^2$ est mesurable.

(4) **μ_2 est une mesure** Vu que $F_{X,Y}$ est croissante en ses deux arguments, il n'y a pas de problèmes pour que $f_{X,Y}$ soit positive. Donc tant que B est borélien, $\int_B F_{X,Y}$ existe et est dans $[0,\infty]$.

Nous avons évidemment $\mu_2(\varnothing) = 0$. Ensuite, si les A_i sont des boréliens disjoints de $\mathbb{R}^2$, alors

$$\mu_2\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \int_{\bigcup_i A_i} f_{X,Y} \tag{36.76a}$$
$$= \sum_i \int_{A_i} f \tag{36.76b}$$
$$= \sum_i \mu_2(A_i). \tag{36.76c}$$

10. Pourquoi nous ne considérons que les rectangles de coordonnées rationnelles ? Nous allons comprendre cela plus tard.

11. Définition 14.11.

(5) $\sigma(\mathcal{E})$ **contient** $\mathcal{B}\text{or}(\mathbb{R}^2)$ Pour utiliser le théorème 14.34 nous devons encore prouver que la tribu engendrée par $\mathcal{E}$ contient les boréliens de $\mathbb{R}^2$. Vu que $\mathcal{E}$ est dénombrable, la proposition 14.35 dit que $\mathcal{B}\text{or}(\mathbb{R}) = \sigma(\mathcal{E})$.

(6) **Autres hypothèses** Vérifions vite fait les autres hypothèses du théorème 14.34. Si A et B sont des rectangles, $A \cap B$ est encore un rectangle. En posant $E_n =]-n, n[\times]-n, n[$, nous avons bien $\mu_1(E_n) = \mu_2(E_n) < \infty$. Et en plus $\lim E_n = \mathbb{R}^2$.

(7) **Conclusion** Nous avons $\mu_1 = \mu_2$ sur $\mathcal{B}\text{or}(\mathbb{R}^2)$, c'est-à-dire que pour tout $B \in Borelien(\mathbb{R}^2)$, nous avons

$$P\big((X,Y) \in B\big) = \int_B f_{X,Y}. \tag{36.77}$$

□

36.2.5 Espérance

DEFooQKFBooCBZtRG

Définition 36.33.
*Si $X \in L^1(\Omega, \mathcal{A}, P)$ est à valeurs dans $\mathbb{R}^n$, nous définissons l'**espérance** de X par*

$$E(X) = \int_\Omega X(\omega) dP(\omega). \tag{36.78}$$

EqdCBLst

Pour l'espérance de $f \circ X$, il y a le lemme de transfert 36.96.

PROPooESPWooAfLmeW

Proposition 36.34.
Soit une variable aléatoire X sur Ω. Nous avons :

ITEMooWSJVooAlZSBZ

(1) $E(X)$ existe et est finie si et seulement si $E(|X|) < \infty$.

ITEMooLDFWooXjyZyz

(2) Si $E(|X|) < \infty$, alors $|E(X)| \leqslant E(|X|) < \infty$.

ITEMooFJCTooEBeFaz

(3) Si $0 < r \leqslant s$ alors si $E(|X|^s) < \infty$ alors $E(|X|^r) < \infty$.

Démonstration. Point par point.

(1) **Pour (1)** C'est le lemme 14.188.

(2) **Pour (2)** C'est la proposition 14.189.

(3) **Pour (3)** C'est la proposition 27.35.

□

LEMooEHTYooWmMAgf

Lemme 36.35 ([1]).
Soit un espace de probabilité Ω ainsi qu'une variable aléatoire $X : \Omega \to \mathbb{R}$ prenant seulement les valeurs $\{y_k\}_{k\in\mathbb{N}}$. Alors

$$E(X) = \sum_{k=0}^{\infty} y_k P(X = y_k). \tag{36.79}$$

Démonstration. Vu que l'application $X : \Omega \to \mathbb{R}$ ne prend que les valeurs y_k, nous avons $\Omega =$

$\bigcup_{k=0}^{\infty} X^{-1}(y_k)$, avec une union disjointe. Nous avons

$$E(X) = \int_{\bigcup_i X^{-1}(y_i)} X(\omega) dP(\omega) \tag{36.80a}$$
$$= \sum_{i=0}^{\infty} \int_{X^{-1}(y_k)} \underbrace{X(\omega)}_{=y_k} dP(\omega) \tag{36.80b}$$
SUBEQooVLTMooQfnAJ
$$= \sum_{k=0}^{\infty} y_k \int_{X^{-1}(y_k)} dP(\omega) \tag{36.80c}$$
$$= \sum_{k=0}^{\infty} y_k P\big(X^{-1}(y_k)\big) \tag{36.80d}$$
SUBEQooXRHQooJVuFtY
$$= \sum_{k=0}^{\infty} y_k P(X = y_k) \tag{36.80e}$$
SUBEQooHDIBooNxnBwg

Justifications :

— Pour (36.80b), nous avons utilisé la σ-additivité de l'intégrale de la proposition 14.210.
— Pour (36.80d), l'intégrale de la fonction 1 donne la mesure de la partie, c'est le lemme 14.175.
— Pour (36.80e), la notation $P(X = y)$ signifie $P\big(\{\omega \in \Omega \text{ tel que } X(\omega) = y\}\big)$; c'est la mesure de $X^{-1}(y)$.

□
PROPooHJZWooGCpKLf

Proposition 36.36 ([767]).
Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire positive, alors

$$E(X) = \int_0^{\infty} P(X \geqslant t) dt. \tag{36.81}$$

Démonstration. Pour chaque $t \in \mathbb{R}$ nous définissons

$$\Omega_t = \{\omega \in \Omega \text{ tel que } X(\omega) \geqslant t\}. \tag{36.82}$$

(1) $\underline{P(X \geqslant t) = E(\mathbb{1}_{\Omega_t})}$ En utilisant le lemme 14.175, nous avons

$$E(\mathbb{1}_{\Omega_t}) = \int_{\Omega} \mathbb{1}_{\Omega_t} d\omega = P(\Omega_t) = P(X \geqslant t). \tag{36.83}$$

(2) **Une expresion pour $X(\omega)$** Soit $\omega \in \Omega$. Nous allons montrer que $X(\omega) = \int_0^{\infty} \mathbb{1}_{\Omega_t}(\omega) dt$. Pour cela il faut remarquer que :

$$\mathbb{1}_{\Omega_t}(\omega) = \begin{cases} 1 & \text{si } X(\omega) \geqslant t \\ 0 & \text{sinon.} \end{cases} \tag{36.84}$$

Avec cela nous avons

$$\int_{\mathbb{R}^+} \mathbb{1}_{\Omega_t}(\omega) dt = \int_0^{X(\omega)} 1 dt = X(\omega). \tag{36.85}$$

(3) **Le calcul** En utilisant les points précédents,

$$E(X) = \int_{\Omega} X(\omega) d\omega \tag{36.86a}$$
$$= \int_{\Omega} \int_0^{\infty} \mathbb{1}_{\Omega_t}(\omega) dt \tag{36.86b}$$
$$= \int_0^{\infty} \int_{\Omega} \mathbb{1}_{\Omega_t}(\omega) dt \qquad \text{Fubini-Tonelli (14.937)} \tag{36.86c}$$
$$= \int_0^{\infty} E(\mathbb{1}_{\Omega_t}) dt \tag{36.86d}$$
$$= \int_0^{\infty} P(X \geqslant t) dt. \tag{36.86e}$$

□

La proposition suivante provient du fait que la mesure d'une loi conjointe est le produit des mesures lorsque les variables aléatoires sont indépendantes (proposition 36.23).

PROPooDKQDooREiSSf

Proposition 36.37 ([432])**.**
Si les variables aléatoires réelles $X_1,\ldots,X_n$ sont intégrables et indépendantes, alors leur produit est intégrable et l'espérance[12] *du produit est égal au produit des espérances :*

$$E(X_1\cdots X_n)=E(X_1)\ldots E(X_n). \tag{36.87}$$

LEMooIVIWooUUVStW

Lemme 36.38.
Soit une variable aléatoire à densité $X\colon\Omega\to\mathbb{R}^n$. Soit une fonction $f\colon\mathbb{R}^n\to\mathbb{R}$ telle que pour tout choix $a_1<b_1<a_2<\ldots<a_n<b_n$ nous ayons

$$P(X\in I)=\int_I f \tag{36.88}$$

où $I=\prod_{i=1}^n]a_i,b_i]$.

Alors f est une densité pour X.

PROBooAHJSooWUuJhb

¡¡ Avertissement/question au lecteur !! 36.39
Je n'ai pas vérifié que ce lemme est correct. Soyez prudent. À mon avis il doit être bon parce que les pavés donnés engendrent la tribu des boréliens d'une façon ou d'une autre. C'est à peu près ce qui est dit dans [768].

Écrivez-moi si vous avez une preuve ou un contre-exemple.

36.2.6 Moment

DEFooTECPooMzGYgG

Définition 36.40.
*Soit $1\leqslant p<\infty$. Le **moment** d'ordre p d'une variable aléatoire $X\in$ sur $(\Omega,\mathcal{A},P)$ est l'espérance*

$$m_p(X)=E(X^p) \tag{36.89}$$

pour autant que X^p soit intégrable[13].

PROPooIQGWooZeQJgb

Proposition 36.41.
Soit une variable aléatoire X. Si elle admet un moment d'ordre $p\in\mathbb{N}$, elle admet un moment d'ordre k pour tout $k\leqslant p$.

Démonstration. L'hypothèse est que $\int_\Omega|X|^p<\infty$, c'est-à-dire que $E(|X|^p)<\infty$. La proposition 36.34(3) conclut que $E(|X|^s)<\infty$ pour tout $s\leqslant p$. □

36.2.7 Somme et produit de variables aléatoires indépendantes

subsecscnvommevariablsindep

À propos de la somme de variables aléatoires, nous avons la proposition suivante.

PROPooBNUEooOUvpdp

Proposition 36.42.
La densité de la somme de deux variables aléatoires réelles indépendantes est le produit de convolution des densités :

$$f_{X+Y}=f_X*f_Y \tag{36.90}$$

dès que X et Y sont indépendantes.

12. Espérance d'une variable aléatoire, définition 36.33.
13. C'est-à-dire que $\int_\Omega|X|^p<\infty$. Voir 14.187.

Démonstration. Soient X et Y, deux variables aléatoires réelles indépendantes. Nous voudrions étudier la loi de la variable aléatoire $S = X + Y$. Nous commençons par calculer la fonction de répartition en utilisant le résultat de la proposition 36.27 :

$$F_{X+Y}(z) = P(X + Y \leqslant z) = \int_{x+y\leqslant z} f_{X,Y}(x, y)dx\,dy \tag{36.91a}$$
$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{z-x} dy f_X(x)f_Y(y) \tag{36.91b}$$
$$= \int_{\mathbb{R}} \left(\int_{-\infty}^{z-x} f_Y(y)dy \right) f_X(x)dx \tag{36.91c}$$
$$= \int_{\mathbb{R}} F_Y(z - x)f_X(x)dx. \tag{36.91d}$$

Pour calculer la fonction de densité de S, nous dérivons la fonction de répartition :

$$f_{X+Y}(z) = \frac{dF_{X+Y}}{dz}(z) \tag{36.92a}$$
$$= \int_{\mathbb{R}} f_Y(z - x)f_X(x)dx, \tag{36.92b}$$

ce qui nous amène à dire que la densité de la somme est le produit de convolution [14] des densités :

EQooFKHHooWvUBMV
$$f_{X+Y}(x) = \int_{\mathbb{R}} f_Y(x - t)f_X(t)dt, \tag{36.93}$$

ou encore $f_{X+Y} = f_X * f_Y$. □

Notez que nous avons passé sous le silence la difficulté d'inverser la dérivée et l'intégrale. Un exemple sera donné au point 36.5.8.

LemEXYEXEYprodindep
Lemme 36.43.
Soient X et Y, deux variables aléatoires indépendantes. Alors [15]

$$E(XY) = E(X)E(Y). \tag{36.94}$$

Démonstration. Par indépendance et par proposition 36.27, la fonction de densité conjointe de X et Y vaut $f_{X,Y} = f_X f_Y$. Par conséquent l'utilisation de Fubini sous la forme (14.964) entraine

$$E(XY) = \int_{\mathbb{R}\times\mathbb{R}} xy f_{X,Y}(x, y)dxdy = E(X)E(Y). \tag{36.95}$$

□

Nous dirons tout un tas de chose sur l'indépendance et la variance en 36.5.13, mais pour l'instant nous allons mentionner et démontrer déjà ceci :

LemVarXpYsmindep
Lemme 36.44.
Soient X et Y deux variables aléatoires indépendantes et identiquement distribuées. Alors

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y). \tag{36.96}$$

Démonstration. Par définition, $\mathrm{Var}(X + Y) = E\big([X + Y - E(X) - E(Y)]^2\big)$. En développant le carré et en utilisant le lemme 36.43,

$$\mathrm{Var}(X + Y) = E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 = \mathrm{Var}(X) + \mathrm{Var}(Y). \tag{36.97}$$

□

14. Définition 27.65.
15. Espérance d'une variable aléatoire, définition 36.33.

ExWLzkuWd

Exemple 36.45.
Deux variables aléatoires non indépendantes dont la covariance est nulle. Nous considérons la variable aléatoire

$$Z\colon \Omega \to \{(1,0),(-1,0),(0,1),(0,-1)\} \tag{36.98}$$

de loi uniforme. C'est-à-dire que $P\big(Z=z\big)=\frac{1}{4}$ pour tout z. Ensuite nous considérons les variables aléatoires $X=\texttt{proj}_1\circ Z$ et $Y=\texttt{proj}_2\circ Z$. Toute personne étant capable de compter jusqu'à 4 voit que

$$P(X=1)=P(X=-1)=\frac{1}{4} \tag{36.99a}$$

$$P(X=0)=\frac{1}{2}, \tag{36.99b}$$

et les mêmes probabilités pour Y. De même $E(X)=E(Y)=0$. Par conséquent

$$\operatorname{Cov}(X,Y)=E(XY)=0 \tag{36.100}$$

parce que pour tout $\omega\in\Omega$ nous avons soit $X(\omega)=0$ soit $Y(\omega)=0$. Ces variables aléatoires X et Y ne sont donc pas corrélées.

Mais elles ne sont pas indépendantes pour autant, comme nous allons le voir pas plus tard qu'immédiatement. Nous avons

$$P(X=0|Y=0)=\frac{P(X=0,Y=0)}{P(Y=0)}=0 \tag{36.101}$$

parce que X et Y ne peuvent pas être simultanément nulles, tandis que

$$P(X=0)P(Y=0)=\frac{1}{4}. \tag{36.102}$$

△

Définition 36.46.
*Si $E(X)=0$ nous disons que la variable aléatoire est **centrée**. La variable aléatoire $X-E(X)$ est la variable aléatoire centrée associée à X.*

PropZBnsCgh

Proposition 36.47 ([769, 770]).
Soient deux variable aléatoires $X,Y\in L^1(\Omega,\mathcal{A},P)$. Alors

$$E(X+Y)=E(X)+E(Y). \tag{36.103}$$

Démonstration. Il suffit d'écrire

$$E(X+Y)=\int_\Omega (X+Y)(\omega)dP(\omega), \tag{36.104}$$

et de faire jouer la linéarité de l'intégrale de la proposition 14.193. □

Une application de l'inégalité de Hölder (proposition 27.33) est la suivante. Si X et Y sont des variables aléatoires intégrables alors

$$E(XY)\leqslant E(X^2)^{1/2}E(Y^2)^{1/2}. \tag{36.105}$$

En effet

$$E(XY)\leqslant \|XY\|_{L^1(\Omega)}\leqslant \|X\|_{L^2(\Omega)}\|Y\|_{L^2(\Omega)}. \tag{36.106}$$

EqEXYleqXdYdNormHolder

36.2.8 Variance

Si $X \in L^2(\Omega, \mathcal{A}, P)$ alors nous définissons la **variance** de X par

$$\mathrm{Var}(X) = E\big([X - E(X)]^2\big). \tag{36.107}$$

PrropVarAlterfrom

Proposition 36.48.
La variance de la variable aléatoire X peut être exprimée par la formule

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 \tag{36.108}$$

EqtWqMGB

où $X^2 = X \cdot X$ et $E(X)^2 = E(X) \cdot E(X)$ sont des produits scalaires dans $\mathbb{R}^d$.

Démonstration. De façon explicite, nous avons

$$E\big([X - E(X)]^2\big) = \int_\Omega \big(X(\omega) - E(X)\big) \cdot \big(X(\omega) - E(X)\big) dP(\omega) \tag{36.109}$$

où $E(X) \in \mathbb{R}^d$ est une constante. En développant le produit scalaire nous avons

$$\begin{aligned} E\big([X - E(X)]^2\big) &= E\big(X^2 - 2X \cdot E(X) + E(X)^2\big) && \text{(36.110a)} \\ &= E(X^2) - 2E(X)^2 + E(X)^2 && \text{(36.110b)} \\ &= E(X^2) - E(X)^2. && \text{(36.110c)} \end{aligned}$$

□

Définition 36.49 (écart-type)**.**
*Nous définissons l'**écart-type** de X par*

$$\sigma_X = \sqrt{\mathrm{Var}(X)}, \tag{36.111}$$

c'est à dire

$$\sigma_X = \|X - E(X)\|_{L^2}. \tag{36.112}$$

Définition 36.50 (moyenne quadratique)**.**
*On définit encore la **moyenne quadratique** de X par*

$$\|X\|_{L^2} = \big[E(X^2)\big]^{1/2}. \tag{36.113}$$

Définition 36.51 (moyenne empirique)**.**
*Si les $X_1, \dots, X_n$ sont des variables aléatoires on considère la **moyenne empirique***

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}. \tag{36.114}$$

Définition 36.52 (variance empirique)**.**
Si $(X_1, \dots, X_n)$ sont des variables aléatoires, nous notons $\bar{X}_n$ leur moyenne empirique, et la variable aléatoire

$$\bar{V}_n = \frac{1}{n} \sum_i (X_i - \bar{X}_n)^2 \tag{36.115}$$

*est la **variance empirique** de l'échantillon (X_i).*

LemEXYEXEYindep

Lemme 36.53.
Si X est une variable aléatoire,

(1) $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$ *pour tout $a \in \mathbb{R}$;*

(2) Si de plus Y est une variable aléatoire indépendante de X, alors $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Démonstration. Nous avons

$$\begin{aligned}\mathrm{Var}(X+Y) &= E(X^2+Y^2+2XY)-\big(E(X)+E(Y)\big)^2 &(36.116a)\\ &= E(X^2)+E(Y^2)+2E(XY)-E(X)^2-E(Y)^2+2E(X)E(Y). &(36.116b)\end{aligned}$$

Étant donné que X et Y sont indépendantes nous avons $E(XY)=E(X)E(Y)$ par le lemme 36.43. □

36.2.9 Covariance

Soient X et Y, deux variables aléatoires réelles. Leur **covariance** est définie par

$$\mathrm{Cov}(X,Y)=E\Big[\big(X-E(X)\big)\big(Y-E(Y)\big)\Big] \tag{36.117}$$

EqHUWttN

L'idée est que la covariance devient grande si X et Y s'écartent de leurs moyennes dans le même sens. Il existe une formule alternative :

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y) \tag{36.118}$$

En ce qui concerne les dimensions plus hautes, si $X\colon\Omega\to\mathbb{R}^d$ est un vecteur aléatoire de carré intégrable, nous définissons

$$\mathrm{Cov}(X)=E\Big[\big(X-E(X)\big)\otimes\big(X-E(X)\big)\Big] \tag{36.119}$$

EqZlyLWx

où par $a\otimes b$ nous entendons la matrice $(a\otimes b)_{ij}=a_ib_j$. Cela peut aussi être noté a^tb si l'on fait bien attention à qui est un vecteur colonne et qui est un vecteur ligne.

PropoVarXpYCov

Proposition 36.54.
Si X et Y sont deux variables aléatoires non spécialement indépendantes, nous avons

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y). \tag{36.120}$$

Démonstration. Il s'agit d'un calcul en partant de

$$\begin{aligned}\mathrm{Var}(X+Y)&=E\big((X+Y)^2\big)-E(X+Y)^2\\ &=E(X^2)+E(Y^2)+2E(XY)\\ &\quad+\big(E(X)+E(Y)\big)^2-2E(X)^2-2E(X)E(Y)\\ &\quad-2E(Y)E(X)-2E(Y)^2.\end{aligned} \tag{36.121}$$

À partir d'ici il s'agit de recombiner tous les termes pour former la formule annoncée. □

Plus généralement nous avons la formule

$$\mathrm{Var}(\sum_i X_i)=\sum_i\mathrm{Var}(X_i)+2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j). \tag{36.122}$$

36.2.10 Probabilité conditionnelle : événements

DEFooGJVHooVbhVYv

Proposition-Définition 36.55.
Soit $(\Omega,\mathcal{A},P)$ un espace de probabilité et $B\in\mathcal{A}$ avec $P(B)>0$. Alors avec la formule

$$P_B(A)=\frac{P(A\cap B)}{P(B)}, \tag{36.123}$$

EqProbCond

*l'espace $(\Omega,\mathcal{A},P_B)$ est un espace probabilisé. Nous notons $P(A|B)$ le nombre $P_B(A)$ et nous le nommons **probabilité conditionnelle** de A sachant B.*

Démonstration. On vérifie que $(\Omega, \mathcal{A}, P_B)$ est un espace de probabilité parce que $P_B(\Omega) = 1$ et

$$P_B(\bigcup_i A_i) = \sum_i P_B(A_i) \tag{36.124}$$

si les A_i sont deux à deux disjoints. □

Une conséquence immédiate de (36.123) est que si A et B sont des événements indépendants alors

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A). \tag{36.125}$$

La probabilité conditionnelle à B est quelque chose qui ne tient compte que de ce qui se passe dans B. Si K est un événement tel que $A \cap B = K \cap B$, alors

$$P(A|B) = P(K|B). \tag{36.126}$$

EqQVHCWom

EXooIAYTooFjFTrT

Exemple 36.56 ([771])**.**
Soient des événements A, B et X dont A et B sont indépendants. Intuitivement, nous devrions avoir [16]

$$P(A|B \cap X) = P(A|X). \tag{36.127}$$

Cela est faux. Prenez par exemple $\Omega = \{1, 2, 3, 4\}$ avec la mesure d'équiprobabilité. Ensuite $A = \{1, 2\}$, $B = \{1, 3\}$ et $X = \{1, 4\}$. Alors $P(A) = 1/2$, $P(B) = 1/2$ de telle sorte que

$$P(A)P(B) = \frac{1}{4} = P(A \cap B) \tag{36.128}$$

parce que $A \cap B = \{1\}$. Les événements A et B sont donc bien indépendants. Nous avons

$$P(A|B \cap X) = \frac{P(A \cap B \cap X)}{P(B \cap X)} = \frac{P(\{1\})}{P(\{1\})} = 1, \tag{36.129}$$

mais

$$P(A|X) = \frac{P(A \cap X)}{P(X)} = \frac{1/4}{1/2} = \frac{1}{2}. \tag{36.130}$$

△

ThoBayesEtAutres

Théorème 36.57.
Soient $(B_n)_{n \geqslant 1}$ une partition finie de Ω telle que $P(B_i) > 0$. Soit $A \in \mathcal{A}$ tel que $P(A) > 0$.

(1) Si A, B et C sont des événements, alors

$$P(A \cap B|C) = P(A|B \cap C)P(B|C). \tag{36.131}$$

(2) Si $P(B) > 0$, alors $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.

*(3) On a la **formule des probabilités totales** :*

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i) = \sum_i P(A \cap B_i). \tag{36.132}$$

*(4) On a la **formule de Bayes** :*

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_i P(A|B_i)P(B_i)}. \tag{36.133}$$

Démonstration. (1) En développant le membre de droite,

$$\begin{aligned} P(A \cap B|C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \frac{P(B \cap C)}{P(C)} \\ &= P(A \cap B|C). \end{aligned} \tag{36.134}$$

16. Probabilité conditionnelle, définition 36.55.

(2) C'est la définition de $P(A|B)$ et $P(B|A)$.

(3) Vu que les B_i forment une partition, nous avons

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i). \tag{36.135}$$

(4) En utilisant les deux premiers points, nous trouvons

$$\begin{aligned} P(A|B_k)P(B_k) &= P(A \cap B_k) \\ &= P(B_k|A)P(A) \\ &= P(B_k|A) \sum_i P(A|B_i)P(B_i). \end{aligned} \tag{36.136}$$

□

Lemme 36.58 ([1])**.**
Soient une variable aléatoire $X : \Omega \to \mathbb{R}$ ainsi qu'un réel $y \neq 0$ et qu'une partie $A \subset \Omega$ de mesure non nulle[17]*. Alors*

$$P(X\mathbb{1}_A = y) = P(X = y|A)P(A). \tag{36.137}$$

Démonstration. Juste pour être clair avec les notations,

— La notation $X = y$ désigne l'ensemble $\{\omega \in \Omega \text{ tel que } X(\omega) = y\} = X^{-1}(y)$

— De même $X\mathbb{1}_A = y$ désigne l'ensemble $\{\omega \in \Omega \text{ tel que } X(\omega)\mathbb{1}_A(\omega) = y\}$. Comme ici $y \neq 0$, il s'agit des éléments de A tels que $X(\omega) = y$, ou encore $X^{-1}(y) \cap A$.

Nous avons :

$$P(X = y|A) = \frac{P(X = y \cap A)}{P(A)} \tag{36.138a}$$

SUBEQooTJDIooHIiTD

$$= \frac{P(A\mathbb{1}_A = y)}{P(A)}. \tag{36.138b}$$

SUBEQooDAATooBeClLa

Justifications :

— Pour (36.138a), c'est la définition 36.55.

— Pour (36.138b), vu que $y \neq 0$ nous avons

$$\{\omega \in \Omega \text{ tel que } X(\omega) = y\} \cap A = \{\omega \in \Omega \text{ tel que } (X\mathbb{1}_A)(\omega) = y\}. \tag{36.139}$$

□

LEMooRDXRooQLMRGF

Lemme 36.59 ([1])**.**
Soit un espace de probabilité $(\Omega, \mathcal{F}, P)$. Nous considérons un mesurable $B \in \mathcal{F}$ ainsi que des mesurables $(A_i)_{i\in\mathbb{N}}$ deux à deux disjoints. Alors

$$P\Big(\bigcup_{i=0}^{\infty} A_i|B\Big) = \sum_{i=0}^{\infty} P(A_i|B). \tag{36.140}$$

Démonstration. Nous commençons par la définition 36.55 de la probabilité conditionnelle :

$$P\Big(\bigcup_i A_i|B\Big) = \frac{P\big(\bigcup_i A_i \cap B\big)}{P(B)} = \frac{P\big(\bigcup_i (A_i \cap B)\big)}{P(B)}. \tag{36.141}$$

EQooRUXMooNzxnxg

Ensuite, les A_i étant disjoints, les $A_i \cap B$ le sont aussi. Vu que P est une probabilité (et donc une mesure), elle vérifie la condition de la définition 14.18(2), qui donne ici

$$P\Big(\bigcup_{i=0}^{\infty}(A_i \cap B)\Big) = \sum_{i=0}^{\infty} P(A_i \cap B). \tag{36.142}$$

17. Ceci est exactement synonyme de « un événement de probabilité non nulle ».

En remettant dans (36.141),

$$P\big(\bigcup_i A_i|B\big) = \sum_{i=0}^{\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=0}^{\infty} P(A_i|B). \tag{36.143}$$

□

LEMooVYDTooGELPRY

Lemme 36.60 ([1]).
Soient des ensembles E, F, G. Soit une application $f\colon E \times F \to G$. Soient des éléments $a \in G$, $x \in E$. Nous considérons des variables aléatoires $X\colon \Omega \to E$ et $Y\colon \Omega \to F$.

Soit $B \subset \Omega$ tel que $X(\omega) = x$ pour tout $\omega \in B$. Alors nous avons

$$P\big(f(X,Y) = a|B\big) = P\big(f(x,Y) = a|B\big). \tag{36.144}$$

Démonstration. Posons $A = \{\omega \in \Omega \text{ tel que } f\big(X(\omega), Y(\omega)\big) = a\}$. Ce que nous cherchons à calculer est $P(A|B)$ qui est défini en 36.55. En posant $A' = \{\omega \in \Omega \text{ tel que } f\big(x, Y(\omega)\big) = a\}$, nous avons

$$\begin{aligned} A \cap B &= \{\omega \in \Omega \text{ tel que } f\big(X(\omega), Y(\omega)\big) = a, \omega \in B\} && \text{(36.145a)}\\ &= \{\omega \in \Omega \text{ tel que } f\big(x, Y(\omega)\big) = a, \omega \in B\} && \text{(36.145b)}\\ &= A' \cap B. && \text{(36.145c)} \end{aligned}$$

Donc

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A' \cap B)}{P(B)} = P(A'|B). \tag{36.146}$$

□

LEMooZTBSooLemswG

Lemme 36.61 ([1]).
Soient des ensembles E_i $(i = 1, \dots, n)$, F et E. Soient des variables aléatoires $X_i\colon \Omega \to E_i$ et $Y\colon \Omega \to E$, et une application $f\colon E_1 \times \dots \times E_n \to F$.

Nous supposons que Y est indépendante des X_i. Pour $y \in E$ et $z \in F$ nous avons

$$P\big(Y = y|f(X_1, \dots, X_n) = z\big) = P(Y = y). \tag{36.147}$$

LEMooWAOSooBsGucQ

Lemme 36.62 ([1]).
Soient des variables aléatoires $X_i\colon \Omega \to E_i$ $(i = 1, \dots, n)$ et $Y_i\colon \Omega \to F_i$. Nous supposons qu'elles sont toutes indépendantes. Nous considérons des applications $f\colon E_1 \times \dots \times E_n \to S$ et $g\colon F_1 \times F_m \to T$. Alors les variables aléatoires

$$\omega \mapsto f\big(X_1(\omega), \dots, X_n(\omega)\big) \tag{36.148}$$

et

$$\omega \mapsto g\big(Y_1(\omega), \dots, Y_m(\omega)\big) \tag{36.149}$$

sont indépendantes.

36.2.11 Espérance conditionnelle

ThoMWfDPQ

Théorème-Définition 36.63 (Définition de l'espérance conditionnelle[623]).
Soit un espace de probabilité $(\Omega, \mathcal{A}, P)$ et une variable aléatoire intégrable $X\colon \Omega \to \mathbb{R}$. Pour chaque sous tribu $\mathcal{F}$ de $\mathcal{A}$, il existe une (presque partout) unique variable aléatoire $Y\colon \Omega \to \mathbb{R}$ telle que

(1) Y est $\mathcal{F}$-mesurable

(2) Y est P-intégrable

(3) pour tout $B \in \mathcal{F}$,

$$\int_B X dP = \int_B Y dP. \tag{36.150}$$

EqBwBkgE

Cette variable aléatoire sera notée $E(X|\mathcal{F})$ pour des raisons qui apparaîtront plus tard.

Démonstration. Remarquons que prendre $Y = X$ ne fonctionne pas parce qu'en général si $\mathcal{O}$ est mesurable dans $\mathbb{R}$, alors $X^{-1}(\mathcal{O})$ est dans la tribu $\mathcal{A}$, mais n'est pas automatiquement dans la tribu $\mathcal{F}$. Il faudra donc un peu plus travailler.

(1) **Unicité** Si Y_1 et Y_2 vérifient tous les deux les conditions, l'ensemble $\{Y_1 < Y_2\}$ est un élément de $\mathcal{F}$ et nous avons

$$\int_{\{Y_1<Y_2\}} X = \int_{Y_1<Y_2} Y_1 = \int_{Y_1<Y_2} Y_2. \tag{36.151}$$

En particulier nous avons $\int_{\{Y_1<Y_2\}}(Y_1 - Y_2) = 0$ et donc

$$(Y_1 - Y_2)\mathbb{1}_{Y_1<Y_2} = 0 \tag{36.152}$$

presque partout. Le corolaire 14.200 montre alors que $Y_1 - Y_2 \geqslant 0$ presque partout. De la même manière, l'ensemble $\{Y_2 < Y_1\}$ est dans $\mathcal{F}$ et nous trouvons que $Y_2 - Y_1 \geqslant 0$ presque partout. Par conséquent $Y_1 = Y_2$ presque partout.

(2) **Existence dans le cas de carré intégrable**
Nous supposons à présent que $X \in L^2(\Omega, \mathcal{A}, P)$ et nous considérons K, le sous-ensemble de $L^2(\Omega, \mathcal{A}, P)$ des fonctions $\mathcal{F}$-mesurables. Le théorème des projections 25.7 nous indique que

$$L^2(\Omega, \mathcal{A}, P) = K \oplus K^\perp \tag{36.153}$$

par la décomposition $X = \mathtt{proj}_K X + (X - \mathtt{proj}_K X)$. La variable aléatoire $Y = \mathtt{proj}_K X$ a les propriétés d'être $\mathcal{F}$-mesurable et $\langle Y - X, Z\rangle = 0$ pour tout $Z \in K$. Soit $A \in \mathcal{F}$, si nous considérons $Z = \mathbb{1}_A$, la dernière condition signifie que

$$\int_\Omega X\mathbb{1}_A = \int_\Omega Y\mathbb{1}_A, \tag{36.154}$$

ou encore

$$\int_A Y = \int_A X. \tag{36.155}$$

La variable aléatoire $Y = \mathtt{proj}_K(X)$ répond donc à la question lorsque $X \in L^2(\Omega, \mathcal{A}, P)$.

(3) **Existence en général**
Nous considérons maintenant que $X \in L^1(\Omega, \mathcal{A}, P)$. Quitte à décomposer X en deux fonctions positives X_+ et X_- telles que $X = X_+ + X_-$, nous pouvons supposer que X est positive. Par hypothèse $X \in L^1(\Omega, \mathcal{A}, P)$; pour chaque $n \in \mathbb{N}$ nous posons

$$X_n(\omega) = \min\{X(\omega), n\}. \tag{36.156}$$

Étant donné que la mesure P est une mesure de probabilité, les constantes sont intégrables et $X_n \in L^2(\Omega, \mathcal{A}, P)$. De plus la suite (X_n) est croissante et

$$\lim_{n\to\infty} X_n(\omega) = X(\omega). \tag{36.157}$$

Si nous notons encore K l'ensemble des variables aléatoires dans $L^2(\Omega, \mathcal{A}, P)$ qui sont $\mathcal{F}$-mesurables, pour chaque n nous avons donc la variable aléatoire

$$Y_n = \mathtt{proj}_K X_n = E(X_n|\mathcal{F}) \tag{36.158}$$

qui est $\mathcal{F}$-mesurable et telle que

$$\int_A X_n = \int_A Y_n \tag{36.159}$$

pour tout $A \in \mathcal{F}$. Nous voudrions prouver que la variable aléatoire $Y = \lim_n Y_n$ existe et est la solution au problème, c'est-à-dire est $E(X|\mathcal{F})$.

Commençons par prouver que $Y_n \geqslant 0$ presque partout. Pour cela nous remarquons que l'ensemble $\{Y_n < 0\}$ est mesurable et

$$0 \geqslant \int_{Y_n<0} Y_n = \int_{Y_n<0} X_n \geqslant 0. \tag{36.160}$$

La première inégalité est évidente et la dernière est due au fait que X_n est positive. Par conséquent

$$\int_{Y_n<0} Y_n = 0 \tag{36.161}$$

et le lemme 14.200 conclut que $P(Y_n < 0) = 0$.

Soit $Z : \Omega \to \mathbb{R}$ une variable aléatoire positive dans $L^2(\Omega, \mathcal{A}, P)$. Montrons que $\texttt{proj}_K Z$ est encore positive. Pour cela nous considérons l'ensemble $A = \{\texttt{proj}_K Z < 0\}$ et les inégalités

$$0 \leqslant \int_A Z = \int_A \texttt{proj}_K Z \leqslant 0, \tag{36.162}$$

ce qui montre que $\int_A \texttt{proj}_K Z = 0$ et par conséquent que $P\{\texttt{proj}_K(Z) < 0\} = 0$. Cela nous montre que la projection depuis L^2 conserve la positivité.

Étant donné que $X_{n+1} - X_n \geqslant 0$ nous avons aussi

$$Y_{n+1} - Y_n \geqslant 0 \tag{36.163}$$

La suite de fonctions

$$n \mapsto Y_n = E(X_n|\mathcal{F}) \tag{36.164}$$

est croissante et vérifie le théorème de la convergence monotone :

$$\int_A X = \lim_{n\to\infty} \int_A X_n = \lim_{n\to\infty} \int_A E(X_n|\mathcal{F}) = \int_A \lim_{n\to\infty} E(X_n|\mathcal{F}) = \int_A Y. \tag{36.165}$$

Par conséquent $E(X|\mathcal{F})$ existe et

$$Y = \lim_{n\to\infty} E(X_n|\mathcal{F}) = E(X|\mathcal{F}). \tag{36.166}$$

□

NORMooHPHOooUuJWHR

36.64.

Vu la définition 36.63 nous pourrions croire que la variable aléatoire $E(X|\mathcal{F}) = X$ fait l'affaire. Il n'en est rien parce que la variable aléatoire X n'est pas spécialement $\mathcal{F}$-mesurable alors qu'il est requis que $E(X|\mathcal{F})$ le soit. Avec la tribu $\mathcal{F} = \{\emptyset, \Omega\}$, nous n'avons en général pas que $X^{-1}(B) \in \mathcal{F}$ pour tout borélien B.

Par contre si $\sigma(X)$ est la tribu engendrée par la variable aléatoire X, alors $E\big(X|\sigma(X)\big) = X$.

DefooKIHPooMhvirn

Définition 36.65.

*Soit Z une variable aléatoire. L'**espérance conditionnelle** « X sachant Z » est la variable aléatoire*

$$E(X|Z) = E(X|\sigma(Z)) \tag{36.167}$$

où $\sigma(Z)$ est la tribu engendrée par Z. Le membre de droite est une variable aléatoire définie en 36.63.

DEFooEYVCooCeyOXW

Définition 36.66.

Soient $A \in \mathcal{A}$ un événement et $\mathcal{F}$ une sous-tribu de $\mathcal{A}$. Nous définissons $P(A|\mathcal{F})$ par

$$P(A|\mathcal{F}) = E(\mathbb{1}_A|\mathcal{F}). \tag{36.168}$$

Notons que cela est une variable aléatoire et non un réel. Le membre de droite est l'espérance conditionnelle de la variable aléatoire $\mathbb{1}_A$ par rapport à $\mathcal{F}$ définie en 36.63.

Et l'espérance conditionnelle d'un événement par rapport à une variable aléatoire est :

$$E(A|X) = E\big(A|\sigma(X)\big). \tag{36.169}$$

Proposition 36.67.
Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$ ainsi qu'une variable aléatoire X à valeurs dans $\mathbb{R}^d$, et un événement A. Alors

$$E\big(P(A|X)\big) = P(A). \tag{36.170}$$

Démonstration. Tout le point de la preuve est de remarquer que $E(\mathbb{1}_A) = E(\mathbb{1}_A|X)$.

(1) **La formule $E(\mathbb{1}_A) = E(\mathbb{1}_A|X)$**
La notation $E(\mathbb{1}_A|X)$ est un raccourci pour écrire la variable aléatoire $E\big(\mathbb{1}_A|\sigma(X)\big)$. Cette dernière est l'application $\Omega \to \mathbb{R}^d$ telle que

$$\int_B E\big(\mathbb{1}_A|\sigma(X)\big) = \int_B \mathbb{1}_A \tag{36.171}$$

pour tout borélien B de $\mathbb{R}^d$ tout en étant $\sigma(X)$-mesurable. Comme expliqué en 36.64, il est tentant de dire $E\big(\mathbb{1}_A|\sigma(X)\big) = \mathbb{1}_A$, mais ce n'est pas le cas parce qu'il n'y a aucune raisons que $\mathbb{1}_A$ soit une application $\sigma(X)$-mesurable. Au niveau des espérances, par contre, l'égalité tient :

$$E\big(E(\mathbb{1}_A|X)\big) = \int_\Omega E(\mathbb{1}_A|X) = \int_\Omega \mathbb{1}_A = E(\mathbb{1}_A) \tag{36.172}$$

où nous avons utilisé le fait que Ω lui-même soit $\sigma(X)$-mesurable.

(2) **La preuve**
Nous avons alors

$$P(A) = E(\mathbb{1}_A) = E\big(E(\mathbb{1}_A|X)\big), \tag{36.173}$$

alors que $E(\mathbb{1}_A|X) = P(A|X)$. En mettant l'un dans l'autre :

$$P(A) = E\big(P(A|X)\big). \tag{36.174}$$

□

PropRGcscXj

Proposition 36.68 (Transitivité de l'espérance conditionnelle)**.**
Si $\mathcal{B}_2 \subseteq \mathcal{B}_1 \subset \mathcal{A}$ alors

$$E\Big(E(X|\mathcal{B}_1)|\mathcal{B}_2\Big) = E(X|\mathcal{B}_2). \tag{36.175}$$

Démonstration. Si $B \in \mathcal{B}_2$, nous avons

$$\int_B E\big(E(X|\mathcal{B}_1)|\mathcal{B}_2\big)dP = \int_B E(X|\mathcal{B}_1)dP = \int_B X dP. \tag{36.176}$$

La première égalité est la définition de l'espérance conditionnelle par rapport à $\mathcal{B}_2$. La seconde égalité est celle de l'espérance conditionnelle par rapport à $\mathcal{B}_1$ et le fait que $B \in \mathcal{B}_2 \subset \mathcal{B}_1$. Ce que nous avons prouvé est que

$$E\big(E(X|\mathcal{B}_1)|\mathcal{B}_2\big) \tag{36.177}$$

est une variable aléatoire $\mathcal{B}_2$-mesurable vérifiant la condition

$$\int_B E\big(E(X|\mathcal{B}_1)|\mathcal{B}_2\big) = \int_B E(X|\mathcal{B}_2) \tag{36.178}$$

pour tout $B \in \mathcal{B}_2$. C'est donc $E(X|\mathcal{B}_2)$ par la partie unicité du théorème 36.63. □

Proposition 36.69.
Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité, soit $\mathcal{A}$ une sous tribu de $\mathcal{F}$ et X, une variable aléatoire $\mathcal{F}$-mesurable et intégrable. Alors la variable aléatoire $E(X|\mathcal{A})$ du théorème 36.63 est l'unique (presque partout) variable aléatoire à être $\mathcal{A}$-mesurable telle que nous ayons

$$E\big(E(X|\mathcal{A})Y\big) = E(XY). \tag{36.179}$$

pour toute variable aléatoire Y $\mathcal{A}$-mesurable.

Démonstration. Supposons pour commencer que Y soit une fonction simple positive, alors $Y = \sum_{i=1}^{n} a_i \mathbb{1}_{E_i}$ et nous avons

$$\int_\Omega E(X|Y) = \sum_i a_i \int_{E_i} E(X|\mathcal{A}) \tag{36.180a}$$

$$= \sum_i a_i \int_{E_i} X \tag{36.180b}$$

$$= \int_\Omega XY. \tag{36.180c}$$

Maintenant si Y est mesurable et bornée, elle est limite croissante de fonctions étagées bornées (proposition 14.118) et le résultat tient par la convergence monotone, théorème 14.178.

Si Y n'est pas positive, nous séparons $Y = Y_+ - Y_-$.

Pour l'unicité, soit Z et Z' deux variables aléatoires telles que pour toute variable aléatoire Y,

$$\int_\Omega ZY = \int_\Omega XY = \int_\Omega Z'Y. \tag{36.181}$$

Si nous prenons $Y = \mathbb{1}_{\{Z \neq Z'\}}$, nous avons

$$0 = \int_\Omega (Z - Z')\mathbb{1}_{Z\neq Z'} = \int_{Z\neq Z'} Z - Z', \tag{36.182}$$

d'où le fait que $P(Z \neq Z') = 0$. □

Si X est une variable aléatoire dont la tribu engendrée est indépendante de la tribu $\mathcal{F}$, nous voudrions que la connaissance de $\mathcal{F}$ n'influence pas la connaissance de X, c'est-à-dire que

$$E(X|\mathcal{F}) = E(X). \tag{36.183}$$

Ce que nous avons est même mieux. Nous avons le lemme suivant qui dit que conditionner une variable aléatoire par rapport à une tribu indépendante ne change pas notre connaissance de la variable aléatoire.

LemxUZFPV

Lemme 36.70 ([432])**.**
Les tribus $\mathcal{F}_1$ et $\mathcal{F}_2$ sont indépendantes si et seulement si

$$E(U|\mathcal{F}_1) = E(U|\mathcal{F}_2) = E(U) \tag{36.184}$$

pour toute variable aléatoire U étant $\mathcal{F}_1$-mesurable.

Ici, par $E(U)$ nous entendons la variable aléatoire constante prenant la valeur numérique $E(U)$ en tout point de Ω.

Démonstration. Si $\mathcal{F}_1$ et $\mathcal{F}_2$ sont indépendantes, alors pour tout $B \in \mathcal{F}_2$ nous avons EqGGqgxl

$$\int_B U dP = E(U\mathbb{1}_B) \tag{36.185a}$$

$$= E(U)E(\mathbb{1}_B) \tag{36.185b}$$ subeqBZWLNS

$$= E(U)\int_\Omega \mathbb{1}_B dP \tag{36.185c}$$

$$= \int_B E(U)dP. \tag{36.185d}$$

Justifications.

— L'intégrale $\int_B UdP$ a un sens même si $B \in \mathcal{F}_2$ alors que U est $\mathcal{F}_1$-mesurable. Le supremum (14.468) définissant l'intégrale est tout de même bien défini, en particulier, l'ensemble sur lequel on prend le supremum est non vide.

— Pour (36.185b), la variable aléatoire U est $\mathcal{F}_1$-mesurable (donc la tribu engendrée par U est dans $\mathcal{F}_1$) alors que $\mathbb{1}_B$ est $\mathcal{F}_2$-mesurable. Les tribus engendrées étant indépendantes, les variables aléatoires le sont et nous pouvons décomposer l'espérance.

Ce que montre le calcul (36.185) est que $E(U)$ est une variable aléatoire $\mathcal{F}_2$-mesurable (parce que constante) dont l'intégrale sur chaque élément de $\mathcal{F}_2$ vaut l'intégrale de U. Par la partie unicité du théorème 36.63, nous déduisons que $E(U) = E(U|\mathcal{F}_2)$.

Pour l'autre sens, écrivez-moi si vous avez une démonstration. □

CorakyvMp

Corolaire 36.71.
Si X est une variable aléatoire et si $\mathcal{F}$ est une tribu, alors

$$E\big(E(X|\mathcal{F})\big) = E(X). \tag{36.186}$$

Démonstration. Il suffit d'appliquer la définition (36.150) à $B = \Omega$:

$$E\big(E(X|\mathcal{F})\big) = \int_\Omega E(X|\mathcal{F})(\omega)dP(\omega) = \int_\Omega X(\omega)dP(\omega) = E(X). \tag{36.187}$$

□

Exemple 36.72.
Soient X_1, X_2 deux variables aléatoires à valeurs dans $\{0,1\}$ avec probabilité $1/2$ et indépendantes. Nous considérons $S = X_1 + X_2$. La situation est modélisée par l'espace

$$\Omega = \{(0,0),(0,1),(1,0),(1,1)\} \tag{36.188}$$

et les variables aléatoires

$$X_i(\omega_1,\omega_2) = \omega_i \tag{36.189a}$$
$$S(\omega_1,\omega_2) = \omega_1 + \omega_2. \tag{36.189b}$$

Pour vérifier que de cette manière nous avons bien que X_1 est indépendante de X_2, nous commençons par voir les tribus associées. Un ouvert de $\mathbb{R}$ soit contient 0 et 1, soit contient un seul des deux soit n'en contient aucun des deux. En appliquant X_1^{-1} à chacune de ces quatre situations nous voyons que la tribu $\sigma(X_1)$ est

$$\mathcal{F}_1 = \sigma(X_1) = \big\{\{(0,0),(0,1)\},\{(1,0),(1,1)\},\Omega,\varnothing\big\}. \tag{36.190}$$

De la même façon nous avons

$$\mathcal{F}_2 = \sigma(X_1) = \big\{\{(0,0),(1,0)\},\{(0,1),(1,1)\},\Omega,\varnothing\big\}. \tag{36.191}$$

Nous posons

$$A_0 = \{(0,0),(0,1)\} \tag{36.192a}$$
$$A_1 = \{(1,0),(1,1)\} \tag{36.192b}$$
$$B_0 = \{(0,0),(1,0)\} \tag{36.192c}$$
$$B_1 = \{(0,1),(1,1)\}. \tag{36.192d}$$

Étant donné que $A_i \cap B_j = (i,j)$, nous avons toujours que $P(A_i \cap B_j) = \frac{1}{4} = P(A_i)P(B_j)$. L'indépendance est donc assurée.

Calculons l'espérance conditionnelle $E(S|\mathcal{F}_1)$. Une fonction $\mathcal{F}_1$-mesurable doit être constante sur A_0 et A_1, donc l'espérance conditionnelle est une fonction constante sur A_0 et A_1 dont l'intégrale sur ces ensembles est égale à l'intégrale de S. Nous avons en particulier

$$\int_{A_0} E(S|\mathcal{F}_1) = \int_{A_0} S, \tag{36.193}$$

c'est-à-dire

$$E(S|\mathcal{F}_1)(0,0) + E(S|\mathcal{F}_1)(0,1) = S(0,0) + S(0,1) = 1. \tag{36.194}$$

Nous en concluons que $E(S|\mathcal{F}_1)(0,0) = E(S|\mathcal{F}_1)(0,1) = \frac{1}{2}$. Cela correspond à l'intuition que si on est au point $(0,1)$ ou au point $(0,0)$ en ne sachant que X_1, nous ne savons que le premier zéro, et donc l'espérance de la somme est $\frac{1}{2}$.

Un calcul très similaire montre que

$$E(S|\mathcal{F}_1)(1,0) = E(S|\mathcal{F}_1)(1,1) = \frac{3}{2}. \tag{36.195}$$

Cela correspond au fait qu'en ces points, nous ne savons que le fait que le premier tirage a donné 1, et donc que l'espérance est $\frac{3}{2}$.

Complétons ce tour d'horizon en mentionnant que la tribu engendrée par X_1 et X_2 est la tribu des parties de Ω, de telle façon que l'espérance conditionnelle de S sachant X_1 et X_2 est égale à S. △

PropRNBtfql

Proposition 36.73 ([432])**.**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux variables aléatoires sur Ω réelles. Soit $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. Nous supposons que $X \in L^1(\Omega, \mathcal{A}, P)$, que $Z \in L^\infty(\Omega, \mathcal{B}, P)$ et que $XZ \in L^1(\Omega, P)$. Alors

$$E(ZX|\mathcal{B}) = ZE(X|\mathcal{B}) \tag{36.196}$$

presque surement.

Démonstration. Nous commençons par prouver que

$$\int_\Omega ZE(X|\mathcal{B}) = \int_\Omega ZX. \tag{36.197}$$

EqNDQWIea

Si $Z = \mathbb{1}_B$ pour un ensemble $B \in \mathcal{B}$, alors cette égalité est vraie par définition de l'espérance conditionnelle [18]. Donc cette égalité est correcte tant que Z est une variable aléatoire $\mathcal{B}$-mesurable et étagée. Nous considérons alors, grâce au lemme 14.116, une suite Z_n de variables aléatoires étagées et $\mathcal{B}$-mesurables avec $|Z_n| < Z$. Pour chaque n nous avons donc

$$\int_\Omega Z_n X = \int_\Omega Z_n E(X|\mathcal{B}). \tag{36.198}$$

EqNVpOSaH

Notre idée est de passer à la limite. Vu que Z et Z_n sont bornées (et donc intégrables sur Ω), pour chaque n nous avons $|Z_n X| \leqslant M|X|$ où M majore Z et donc tous les Z_n de façon uniforme vis-à-vis de n. Tout cela pour dire que le théorème de la convergence dominée fonctionne et que

$$\lim_{n\to\infty} \int_\Omega Z_n X = \int_\Omega ZX. \tag{36.199}$$

D'autre part vu que $X \in L^1$ et que $\Omega \in \mathcal{B}$ nous avons l'égalité $\int_\Omega E(X|\mathcal{B}) = \int_\Omega X$, ce qui prouve que $|E(X|\mathcal{B})|$ est intégrable. Cela nous permet d'utiliser encore la convergence dominée avec l'inégalité $|Z_n E(X|\mathcal{B})| \leqslant M|E(X|\mathcal{B})|$ pour écrire

$$\lim_{n\to\infty} \int_\Omega Z_n E(X|\mathcal{B}) = \int_\Omega ZE(X|\mathcal{B}). \tag{36.200}$$

18. Théorème 36.63.

En passant à la limite des deux côtés de (36.198) nous avons donc

$$\int_\Omega ZE(X|\mathcal{B}) = \int_\Omega ZX. \tag{36.201}$$

L'égalité (36.197) est prouvée.

Nous passons maintenant à la preuve de l'égalité demandée : $E(ZX|\mathcal{B}) = ZE(X|\mathcal{B})$. Pour cela il faut montrer que pour tout $B \in \mathcal{B}$ nous avons

$$\int_B ZE(X|\mathcal{B}) = \int_B ZX. \tag{36.202}$$

Cela n'est rien d'autre que l'égalité (36.197) que nous venons de prouver avec $Z\mathbb{1}_B$ au lieu de Z. □

Proposition 36.74.
Soit une variable aléatoire réelle $X \in L^1(\Omega, \mathcal{A}, P)$. Pour toute variable aléatoire $Y\colon \Omega \to \mathbb{R}^d$, il existe une fonction borélienne $\mathcal{A}_Y$-mesurable $h\colon \mathbb{R}^d \to \mathbb{R}$ telle que

$$E(X|Y) = h \circ Y. \tag{36.203}$$

Démonstration. Nous utilisons le résultat de Doob (théorème 14.119). Par définition $E(X|Y)$ est une variable aléatoire réelle $\mathcal{A}_Y$-mesurable, et il existe une fonction borélienne $h\colon \mathbb{R}^d \to \mathbb{R}$ telle que $E(X|Y) = h \circ Y$. □

Cette fonction $h\colon \mathbb{R}^d \to \mathbb{R}$ nous permet de définir

$$E(X|Z = z) = h(z). \tag{36.204}$$

Cela est l'espérance conditionnelle d'une variable aléatoire par rapport à une valeur donnée d'une autre variable aléatoire.

DEFooOMLCooJgrbpx

Définition 36.75 (Espérence conditionnelle à un événement[772])**.**
Si X est une variable aléatoire et si A est un événement, nous définissons

$$E(X|A) = \frac{E(X\mathbb{1}_A)}{P(A)}. \tag{36.205}$$

Pour rappel, la définition de l'espérance d'une variable aléatoire est 36.33.

LEMooRTVBooCEeIxL

Lemme 36.76 ([1])**.**
Soit un espace de probabilité $(\Omega, \mathcal{F}, P)$. Nous considérons une variable aléatoire X à valeurs réelles, prenant ses valeurs dans la partie dénombrable

$$\{0\} \cup \{y_k\}_{k\in\mathbb{N}} \tag{36.206}$$

avec $y_k \neq 0$. Si A est mesurable dans Ω, alors

$$E(X|A) = \sum_{k=0}^{\infty} y_k P(X = y_k|A), \tag{36.207}$$

c'est-à-dire que nous prenons la somme sur les valeurs non nulles atteintes par X.

Démonstration. Nous partons de la définition 36.75 de l'espérance conditionnelle : $E(X|A) = E(X\mathbb{1}_A)/P(A)$. La variable aléatoire $X\mathbb{1}_A$ peut prendre les valeurs 0 et y_k. Nous pouvons utiliser le lemme 36.35 pour écrire

$$E(X\mathbb{1}_A) = 0 \times P(X\mathbb{1}_A = 0) + \sum_{k=0}^{\infty} y_k P(X\mathbb{1}_A = y_k) = \sum_{k=0}^{\infty} y_k P(X\mathbb{1}_A = y_k). \tag{36.208}$$

EQooZSRJooZNtajK

Afin de traiter le dernier terme, nous prouvons que $\{X\mathbb{1}_A = y_k\} = X^{-1}(y_k) \cap A$. En effet si $\omega \in \{X\mathbb{1}_A = y_k\}$, c'est que

$$X(\omega)\mathbb{1}_A(\omega) = y_k. \tag{36.209}$$

Mais dans notre cas, $y_k \neq 0$, donc $\mathbb{1}_A(\omega) = 1$, ce qui signifie $\omega \in A$. Donc $\omega \in A$ et $X(\omega) = y_k$, ce qui signifie $\omega \in X^{-1}(y_k) \cap A$. Nous avons donc

$$P(X\mathbb{1}_A = y_k) = P\big(X^{-1}(y_k) \cap A\big). \tag{36.210}$$

En utilisant (à l'envers) la définition de la probabilité conditionnelle 36.55,

$$P(X\mathbb{1}_A = y_k) = P\big(X^{-1}(y_k)|A\big)P(A). \tag{36.211}$$

En remettant ça dans (36.208) et dans la définition de $E(X|A)$,

$$E(X|A) = \frac{\sum_{k=0}^{\infty} y_k P(X\mathbb{1}_A = y_k)}{P(A)} = \frac{1}{P(A)} \sum_{k=0}^{\infty} P(X = y_k|A)P(A) = \sum_{k=0}^{\infty} y_k P(X = y_k|A) \tag{36.212}$$

□

36.2.12 Probabilité conditionnelle : tribu

Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$.

LEMooXXTYooZCXiYr

Lemme 36.77.
Soit $(B_i)_{i\in\mathbb{N}}$ une partition de Ω en éléments de $\mathcal{A}$ deux à deux disjoints tels que $P(B_i) \neq 0$. Soit $\mathcal{F}$ la tribu engendrée par les B_i. Une variable aléatoire réelle est $\mathcal{F}$-mesurable si et seulement si elle est constante sur chaque B_i.

Proposition 36.78.
Soit $(B_i)_{i\in\mathbb{N}}$ une partition de Ω en éléments de $\mathcal{A}$ deux à deux disjoints tels que $P(B_i) \neq 0$. Soit $\mathcal{F}$ la tribu engendrée par les B_i. Soit une variable aléatoire X. Alors nous avons :

$$E(X|\mathcal{F}) = \sum_{i\in\mathbb{N}} \left(\frac{1}{P(B_i)} \int_{B_i} X dP \right) \mathbb{1}_{B_i}. \tag{36.213}$$

FqCibwoG

Démonstration. Si X est une variable aléatoire, alors la variable aléatoire $E(X|\mathcal{F})$ définie en 36.63 est une variable aléatoire $\mathcal{F}$-mesurable et elle est donc constante sur les ensembles B_i par le lemme 36.77 :

$$E(X|\mathcal{F}) = \sum_{i\in\mathbb{N}} a_i \mathbb{1}_{B_i}. \tag{36.214}$$

Étant donné que, par construction, B_i est $\mathcal{F}$-mesurable, nous avons

$$\int_{B_i} X dP = \int_{B_i} E(X|\mathcal{F}) = \sum_j a_j \int_{B_i} \mathbb{1}_{B_j} = \sum_j a_j \delta_{ij} P(B_j) = a_i P(B_i). \tag{36.215}$$

Par conséquent

$$a_i = \frac{1}{P(B_i)} \int_{B_i} X dP \tag{36.216}$$

et

$$E(X|\mathcal{F}) = \sum_{i\in\mathbb{N}} \left(\frac{1}{P(B_i)} \int_{B_i} X dP \right) \mathbb{1}_{B_i}, \tag{36.217}$$

ce qu'il fallait. □

Notons que si $B \in \mathcal{A}$ alors la tribu engendrée par B est aussi celle engendrée par la partition $\{B, \complement B\}$ de Ω. Cette circonstance nous permet d'aller plus loin.

Proposition 36.79.
Soit un espace probabilisé $(\Omega, \mathcal{A}, P)$ et un événement $B \in \mathcal{A}$ avec sa tribu engendrée $\mathcal{F} = \sigma(B)$. Alors

$$E(\mathbb{1}_A|\mathcal{F}) = P(A|B)\mathbb{1}_B + P(A|\complement B)\mathbb{1}_{\complement B}. \tag{36.218}$$

Démonstration. Nous allons particulariser la formule (36.213). Si $B \in \mathcal{A}$ nous considérons la partition $\{B, \complement B\}$ de Ω et la tribu engendrée

$$\mathcal{F} = \{\varnothing, B, \complement B, \Omega\}. \tag{36.219}$$

La formule (36.213) devient

$$E(X|\mathcal{F}) = \left(\frac{1}{P(B)}\int_B X dP\right)\mathbb{1}_B + \left(\frac{1}{P(\complement B)}\int_{\complement B} X dP\right)\mathbb{1}_{\complement B}. \tag{36.220}$$

Si nous considérons $A \in \mathcal{A}$, nous écrivons cette égalité avec $X = \mathbb{1}_A$ pour obtenir

$$E(\mathbb{1}_A|\mathcal{F}) = \frac{P(A \cap B)}{P(B)}\mathbb{1}_B + \frac{P(A \cap \complement B)}{P(\complement B)}\mathbb{1}_{\complement B} = P(A|B)\mathbb{1}_B + P(A|\complement B)\mathbb{1}_{\complement B} \tag{36.221}$$

parce que nous avons reconnu la probabilité conditionnelle $P(A|B)$ de la définition 36.55. □

Remarque 36.80.
Le nombre $P\big(A|\sigma(B)\big) = P\big(\mathbb{1}_A|\sigma(B)\big)$ n'est pas la probabilité conditionnelle de A sachant B.

Il nous reste à définir la probabilité conditionnelle d'un événement relativement à une variable aléatoire.

DEFooFRLFooNvXuPK

Définition 36.81.
Si la variable aléatoire X est à valeurs discrètes, nous disons que $P(A|X)$ est la variable aléatoire de valeur

$$P(A|X)(\omega) = P(A|X = X(\omega)). \tag{36.222}$$

Dans le cas d'une variable aléatoire à valeurs continues, cette définition ne fonctionne pas parce que la condition $X = X(\omega)$ est souvent de probabilité nulle, tandis que c'est toujours une mauvaise idée de conditionner par rapport à un événement de probabilité nulle. C'est la base du paradoxe de Borel. La bonne définition du conditionnement de l'événement A par rapport à la variable aléatoire X est

DEFooIUJMooBAVtMW

Définition 36.82.
Si A est un événement et X une variable aléatoire à valeurs continues dans $\mathbb{R}$, nous définissons

$$P(A|X) = P(A|\sigma(X)) = E\big(\mathbb{1}_A|\sigma(X)\big). \tag{36.223}$$

La première égalité est une notation. La seconde est la définition.

Cette définition s'appuie également sur la définition 36.63.

Proposition 36.83.
Si X est une variable aléatoire et si A est un événement, alors

$$E\big(P(A|X)\big) = P(A). \tag{36.224}$$

Démonstration. Nous commençons par le cas discret, c'est-à-dire $X : \Omega \to \mathbb{N}$. Nous notons $p_k = P(X = k)$. En décomposant l'intégrale sur Ω par rapport à l'union disjointe

$$\Omega = \bigcup_{k\in\mathbb{N}} A_k = \bigcup_{k\in\mathbb{N}} \{\omega \in \Omega \text{ tel que } X(\omega) = k\}, \tag{36.225}$$

nous obtenons

$$E\big(P(A|X)\big) = \int_\Omega P(A|X)(\omega)dP(\omega) \tag{36.226a}$$
$$= \sum_{k=0}^{\infty} \int_{A_k} P(A|X = X(\omega))dP(\omega) \tag{36.226b}$$
$$= \sum_k \int_{A_k} \frac{P(A \cap X = k)}{P(X = k)} dP(\omega) \qquad \text{dans } A_k,\ X(\omega) = k \tag{36.226c}$$
$$= \sum_k \frac{1}{p_k} P(A \cap X = k) \underbrace{\int_{A_k} 1 dP(\omega)}_{=P(A_k)=p_k} \tag{36.226d}$$
$$= \sum_k P(A \cap X = k) \tag{36.226e}$$
$$= P(A). \tag{36.226f}$$

Nous devons maintenant prouver la propriété dans le cas où X prend des valeurs continues. Pour cela il suffit d'appliquer le corolaire 36.71 :

$$E\big(E(\mathbb{1}_A|\sigma(A))\big) = E(\mathbb{1}_A) = P(A). \tag{36.227}$$

□

36.2.13 Variables de Rademacher indépendantes

SUBSECooWOOGooVxflVZ

Une variable aléatoire de Rademacher est une variable aléatoire qui prend les valeurs 1 et -1 avec probabilité $\frac{1}{2}$. Nous pouvons en décrire une explicitement de la façon suivante. L'espace probabilité est à deux éléments : $\Omega = \{a, b\}$ avec la mesure $P(\{a\}) = P(\{b\}) = \frac{1}{2}$. La variable aléatoire est alors l'application $X : \Omega \to \mathbb{R}$ donnée par $X(a) = 1$ et $X(b) = -1$.

Soient X et Y deux variables aléatoires de Rademacher indépendantes. Cela donne $\Omega = \{a, b\}^2$ et

$$\begin{array}{llll} X(a,a) = 1 & X(a,b) = 1 & X(b,a) = -1 & X(b,b) = -1 \\ Y(a,a) = 1 & Y(a,b) = -1 & Y(b,a) = 1 & Y(b,b) = -1 \end{array} \tag{36.228}$$

Remarque 36.84.
Si une variable aléatoire d'un certain type est donnée par une application $X : \Omega \to \mathbb{R}$, pour construire des variables aléatoires indépendantes identiquement distribuées, il faut considérer les variables aléatoires sur (au moins) le produit $\Omega \times \Omega$ muni de la mesure produit.

(1) **Tribu du produit XY**
Quelle est la tribu de la variable aléatoire produit XY ? Le produit XY peut prendre les valeurs 1 et -1. Nous avons

$$(XY)^{-1}(1) = \{(a, a), (b, b)\} \tag{36.229a}$$
$$(XY)^{-1}(-1) = \{(a, b), (b, a)\} \tag{36.229b}$$

La tribu est donc

$$\sigma(XY) = \{\Omega, \varnothing, A, B\} \tag{36.230}$$

avec $A = \{(a, a), (b, b)\}$ et $B = \{(a, b), (b, a)\}$.

(2) **Calcul de $E(X|XY)$**
La définition de l'espérance à calculer est le théorème 36.63. Pour chaque élément B de $\sigma(XY)$ nous avons besoin de $\int_B X = \int_B E(X|XY)$. Nous notons $V = E(X|XY)$ pour alléger la notation. Nous avons

$$4 \int_A V = V(a, a) + V(b, b) \tag{36.231}$$

et

$$4\int_A X = X(a,a) + X(b,b) = 0. \tag{36.232}$$

Pourquoi le facteur 4 ? Parce que sur Ω^2 nous avons la mesure produit de celle que dont nous avions parlé sur $\{a,b\}$. C'est la mesure d'équiprobabilité et donc chaque singleton a mesure 1/4. Pour plus de détails, il y a le théorème 14.246.

Nous en déduisons $V(a,a) + V(b,b) = 0$. Mais pour tout $t \in \mathbb{R}$ nous avons $V^{-1}(t) \in \sigma(XY)$ parce que la contrainte est que V soit XY-mesurable. En particulier

$$V^{-1}\big(V(a,a)\big) \tag{36.233}$$

est un mesurable qui contient (a,a). C'est donc soit Ω, soit $\{(a,a),(b,b)\}$. Dans les deux cas nous avons $V(a,a) = V(b,b)$ et nous en déduisons $V(a,a) = V(b,b) = 0$.

En faisant de même avec $\int_B V = V(a,b) + V(b,a)$ nous déduisons $V(a,b) = V(b,a) = 0$ et au final nous avons

$$E(X|XY) = 0. \tag{36.234}$$

Cette égalité signifie $E(X|XY)(\omega) = 0$ pour tout $\omega \in \Omega^2$.

(3) **Calcul de $E(X|X+Y)$** Il ne faudrait pas croire que, seulement parce que X a une espérance nulle, nous trouverons une espérance nulle quel que soit le conditionnement. Juste pour le plaisir, nous calculons $E(X|X+Y)$.

La variable aléatoire $X+Y$ peut prendre trois valeurs : -2, 0 et 2. La tribu engendrée par $X+Y$ doit en particulier contenir $A = \{(a,a)\}$, $B = \{(b,b)\}$ et $C = \{(a,b),(b,a)\}$.

Nous notons $V = E\big(X|\sigma(X+Y)\big)$. Vu que

$$\int_A V = \int_A X, \tag{36.235}$$

nous avons $V(a,a) = X(a,a) = 1$. Même chose pour B qui donne $V(b,b) = X(b,b) = -1$. En ce qui concerne l'intégrale sur C nous avons

$$V(a,b) + V(b,a) = X(a,b) + X(b,a) = 0. \tag{36.236}$$ EQooGWDYooTNxYAg

Par ailleurs l'ensemble $V^{-1}\big(V(a,b)\big)$ est un ensemble mesurable qui doit au moins contenir (a,b). Vu la tribu que nous avons, cela doit également contenir (b,a), de telle sorte que $V(a,b) = V(b,a)$. La relation (36.236) nous permet alors de conclure que $V(a,b) = V(b,a) = 0$.

Quoi qu'il en soit, l'espérance conditionnelle $E(X|X+Y)$ n'est pas nulle.

(4) **Calcul de $E(XY|\sigma(XY))$** . Celle-là, elle est facile par 36.64 : c'est XY.

Nous aurions pu croire que si X et Y sont indépendantes, alors

$$E(XY|\mathcal{A}) = E(X|\mathcal{A})E(Y|\mathcal{A}). \tag{36.237}$$ EQooHSVCooAcgRbw

L'exemple que nous venons de faire montre qu'il n'en est rien.

Exemple 36.85 ([773])**.**
Un autre exemple, peut-être plus simple, pour contredire l'équation (36.237). Soient X et Y des expériences indépendantes de pile ou face non truquées. Les résultats sont représentés par 0 et 1. Nous notons $\mathcal{A}$ la tribu engendrée par l'événement « les résultats des deux lancers sont différents » ; c'est-à-dire la tribu engendrée par l'événement $A = (1,0),(1,0)$. La variable aléatoire X et la tribu $\mathcal{A}$ sont indépendants (définition 36.9), donc, donc $E(X|\mathcal{A}) = E(X) = 1/2$. Pareil pour Y. En revanche, le produit XY est nul sur A donc $E(XY|\mathcal{A})$ aussi. Ça ne peut donc être égal à la constante $1/4 = (1/2)^2$. △

36.2.14 Un petit paradoxe

subSecGXVYooTDdZaB

Attention : ce qui est écrit ici est ma réflexion personnelle sur le sujet. Merci de me dire si je me trompe.

Soit une famille dont vous savez seulement qu'il y a exactement deux enfants. Trois situations :

ITEMooNUPAooWCXwBE

(1) Vous frappez, une fille ouvre la porte et dit « Bonjour, je suis l'aînée ». Quelle est la probabilité que l'autre enfant soit une fille ?

ITEMooBGIWooLnVCpm

(2) Vous frappez, une fille ouvre la porte et dit « Bonjour ». Quelle est la probabilité que l'autre enfant soit une fille ?

ITEMooCBNSooMaIwwB

(3) Vous demandez aux parents si il y a au moins une fille, ils répondent « oui ». Quelle est la probabilité que les deux enfants soient des filles ?

Dans les trois cas l'intuition dit que la probabilité est 1/2. Il semble que de plus la (2) et la (3) soient les mêmes parce que l'on sait qu'il y a une fille et on se demande quelle est la probabilité qu'il y ait deux filles.

Nous allons voir ça de plus près.

36.2.14.1 « Bonjour, je suis l'aînée »

Résolution Si nous notons X_0 et X_1 les variables aléatoires donnant le sexe des deux enfants, ce sont des variables aléatoires indépendantes et identiquement distribuées, avec $P(X_i = f) = \frac{1}{2}$. La formule (36.123) de la probabilité conditionnelle ainsi que l'indépendance donnent :

$$P(X_1 = f | X_2 = f) = \frac{P(X_1 = f, X_2 = f)}{P(X_2 = f)}. \tag{36.238}$$

Le numérateur vaut $\frac{1}{4}$ et le dénominateur vaut $\frac{1}{2}$; le résultat vaut $\frac{1}{2}$. Fin de l'histoire.

Simulation Voici un petit programme qui simule la situation. Il retourne clairement 1/2.

```
#! /usr/bin/python3
# -*- coding: utf8 -*-

"""
Vous frappez à la porte d'une famille qui a deux enfants. Une fille ouvre la porte et vous dit "Je suis l'aînée".
Quelle est la probabilité que l'autre soit une fille ?
"""


import random

def famille():
    """
    return a pair of 'f' and 'g'.
    """
    a=[random.choice( ['g','f'] )]
    a.append(random.choice( ['g','f'] ))
    return a

def toctoc():
    """
    - Create a family with two children.
    - Pick the second one, the elder.
```

```
    - if it is a 'g', return None.
    - if it is a 'f', return the other one.
    """
    F=famille()
    if F[1] != 'f':
        return None
    else :
        if F[0]=='f':
            return 1
        else :
            return 0

N_girl_opens=0
N_girl_other=0
for k in range(1,10000):
    res=toctoc()
    if res is not None:
        N_girl_opens = N_girl_opens+1
        N_girl_other = N_girl_other + res

proba=N_girl_other/N_girl_opens
print(proba)        # ~0.5, intuitively correct.
```

tex/sage/simul_famille_aine.py

36.2.14.2 « Bonjour »

Nous frappons à la porte, une fille ouvre en disant « bonjour », sans préciser si elle est la première ou la seconde. Quelle est la probabilité que l'autre soit une fille ? Naïvement on croirait que la probabilité est également $\frac{1}{2}$.

Un raisonnement moins naïf montre le contraire.

Et nous allons voir qu'un raisonnement encore moins naïf montre que la probabilité est bien $\frac{1}{2}$.

Premier raisonnement (incorrect) Voici le raisonnement qui est, à mon avis, faux. Vu que l'enfant qui ouvre la porte est une fille, la famille a une des compositions suivantes : fg, ff ou gf. Le cas où une fille ouvre la porte *et* que l'autre est également une fille est seulement le cas ff dont la probabilité est $\frac{1}{3}$.

Pour justifier cela nous considérons le couple de variables aléatoires (X_1, X_2) et le conditionnement $A = \{X_1 = f\} \cup \{X_2 = f\}$: évidemment $P(A) = \frac{3}{4}$. Nous calculons facilement la loi du couple (X_1, X_2) conditionné à A :

$$P(X_1 = f, X_2 = f|A) = \frac{P(\{X_1 = f, X_2 = f\} \cap A)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}. \tag{36.239}$$

Donc sachant A, la probabilité que la famille soit constituée de deux filles est $\frac{1}{3}$.

Comment faire mieux ? Ce calcul semble être correct, mais il ne l'est pas. Ce raisonnement fait l'hypothèse implicite que l'espace probabilisé décrivant la situation contient deux variables aléatoires X_1 et X_2 représentant les deux enfants. Or nous avons bien trois événements aléatoires dans l'histoires : le sexe des deux enfants et le *choix* de l'enfant qui ouvre la porte.

Certes, nous pouvons penser que cette troisième variable aléatoire ne change rien. Oui oui, on peut le penser. Mais ici, on ne doit pas penser, on doit *démontrer*.

Nous allons donc rédiger un calcul complet, en introduisant toutes les variables aléatoires, et en décrivant correctement l'espace probabilisé Ω et la mesure de probabilité P.

Peut-être que ça ne changera rien. Ou peut-être pas. Mais au moins nous serons surs d'avoir résolu le problème correctement.

La vraie réponse Nous considérons les variables aléatoires $X_0, X_1 : \Omega_E \to \{f, g\}$ avec probabilité $\frac{1}{2}$. De plus nous considérons une nouvelle variable aléatoire qui donne le numéro de l'enfant qui ouvre la porte :

$$\sigma : \Omega_C \to \{1, 2\}. \tag{36.240}$$

Notre espace de probabilité est donc l'ensemble $\Omega = \{f, g\} \times \{f, g\} \times 0, 1$ sur lequel nous considérons la mesure d'équiprobabilité [19].

Nous introduisons les variables aléatoires [20]

$$\begin{aligned} X_1 : \Omega &\to \{f, g\} \\ (s_1, s_2, n) &\mapsto s_1 \end{aligned} \tag{36.241}$$

et

$$\begin{aligned} X_2 : \Omega &\to \{f, g\} \\ (s_1, s_2, n) &\mapsto s_2 \end{aligned} \tag{36.242}$$

et

$$\begin{aligned} \sigma : \Omega &\to \{1, 2\} \\ (s_1, s_2, n) &\mapsto n \end{aligned} \tag{36.243}$$

Nous devons calculer

$$P\big(X_{1-\sigma} = f | X_\sigma = f\big) = \frac{P\big(X_{1-\sigma} = f, X_\sigma = f\big)}{P(X_\sigma = f)}. \tag{36.244}$$

Pour être explicite jusqu'au bout, nous énumérons tous les éléments de Ω :

(1) $g, g, 0$ (2) $g, g, 1$ (3) $g, f, 0$ (4) $g, f, 1$ (5) $f, g, 0$ (6) $f, g, 1$ (7) $f, f, 0$ (8) $f, f, 1$.

Et tant qu'à être explicite, l'événement vulgairement noté $\{X_\sigma = f\}$ est la partie

$$\{X_\sigma = f\} = \{\omega \in \Omega \text{ tel que } X_{\sigma(\omega)}(\omega) = f\} \tag{36.245a}$$
$$= \{(s_1, s_2, n) \text{ tel que } X_n(s_1, s_2, n) = f\} \tag{36.245b}$$
$$= \{(s_1, s_2, n) \text{ tel que } s_n = f\}. \tag{36.245c}$$

Méditez la dernière égalité ; elle n'est pas totalement indispensable au raisonnement, mais elle est cool.

Nous avons

$$\{X_\sigma = f\} = \{(g, f, 1), (f, g, 0), (f, f, 0), (f, f, 1)\}. \tag{36.246}$$

et

$$\{X_{1-\sigma} = f\} \cap \{X_\sigma = f\} = \{(f, f, 0), (f, f, 1)\}. \tag{36.247}$$

Donc

$$P\big(X_{1-\sigma} = f, X_\sigma = f\big) = \frac{2}{8} = \frac{1}{4} \tag{36.248}$$

et

$$P(X_\sigma = f) = \frac{4}{8} = \frac{1}{2}. \tag{36.249}$$

Au final,

$$P\big(X_{1-\sigma} = f | X_\sigma = f\big) = \frac{1/4}{1/2} = \frac{2}{4} = \frac{1}{2}. \tag{36.250}$$

19. C'est une hypothèse forte faisant appel d'un part ce que l'on sait de la reproduction humaine, et d'autre part ce que l'on sait de la sociologie de deux enfants qui entendent une sonnette.

20. Sur Ω, sur $\{f, g\}$ et sur $\{0, 1\}$ nous mettons la tribu des parties. Vérifiez que X_1, X_2 et σ sont mesurables.

Simulation Vous avez encore un doute ? Faites tourner la simulation suivante :

```
#! /usr/bin/python3
# -*- coding: utf8 -*-

"""
Vous frappez à la porte d'une famille qui a deux enfants. Une fille ouvre la porte.
Quelle est la probabilité que l'autre soit une fille ?
"""


import random

def famille():
    """
    return a pair of 'f' and 'g'
    """
    a=[random.choice( ['g','f'] )]
    a.append(random.choice( ['g','f'] ))
    return a

def toctoc():
    """
    - Create a family with two children.
    - Choose one (the one who opens the door)
    - if it is a 'g', return None.
    - if it is a 'f', return the other one.
    """
    F=famille()
    s=random.choice([0,1])
    if F[s] != 'f':
        return None
    else :
        t=(s+1)%2
        if F[t]=='f':
            return 1
        else :
            return 0

N_girl_opens=0
N_girl_other=0
for k in range(1,10000):
    res=toctoc()
    if res is not None:
        N_girl_opens = N_girl_opens+1
        N_girl_other = N_girl_other + res

proba=N_girl_other/N_girl_opens
print(proba)        # ~0.5, intuitively correct.
```

tex/sage/simul_famille_simple.py

Le faisant tourner, la réponse est sans appel : la fréquence observée est beaucoup plus proche

de 0.5 que de 0.33 ou 0.66.

36.2.14.3 Le parent qui répond aux questions

Nous avons une famille de deux enfants dont nous savons qu'au moins un des deux est une fille. Quelle est la probabilité que la famille contienne deux filles ? Cela est à priori la même question que celle où une fille ouvre la porte sans dire si elle est l'aînée ou non.

Simulation Commençons par la simulation :

```
#! /usr/bin/python3
# -*- coding: utf8 -*-

"""
Vous savez qu'une famille a deux enfants.
Vous demandez à un parent si il y a une fille.
Réponse : Oui.
Question : quelle est la probabilité que ce soient deux filles ?
"""


import random

def famille():
    """
    return a pair of 'f' and 'g'.
    """
    a=[random.choice( ['g','f'] )]
    a.append(random.choice( ['g','f'] ))
    return a

def toctoc():
    """
    - Create a family with two children.
    - If 'gg', there are no girls -> return None
    - If 'gf', there is a girl but the other is a boy -> 0
    - If 'fg', there is a girl but the other is a boy -> 0
    - If 'ff', there is a girl and the other is a girl -> 1
    """
    F=famille()
    if F==['g','g'] :
        return None
    if F==['g','f']:
        return 0
    if F==['f','g']:
        return 0
    if F==['f','f']:
        return 1

N_at_least_one_girl=0
N_two_girls=0
for k in range(1,10000):
    res=toctoc()
```

```
    if res is not None:
        N_at_least_one_girl=N_at_least_one_girl+1
        N_two_girls=N_two_girls+res

proba=N_two_girls/N_at_least_one_girl
print(proba)        # ~0.333. Beware !
```

tex/sage/simul_famille_une_fille.py

Et là, bing, la réponse est clairement plutôt 0.33 que 0.5.

Résolution Nous avons les variables aléatoires X_1 et X_2 qui valent 0 ou 1 suivant que l'enfant soit une fille ou un garçon ; ce sont des variables aléatoires indépendantes et identiquement distribuées. Nous définissons la variable aléatoire somme

$$S = X_1 + X_2 \tag{36.251}$$

qui compte le nombre de filles. La question est de calculer

$$P(S = 2|S \geqslant 1) = \frac{P(S = 2 \cap S \geqslant 1)}{P(S \geqslant 1)} = \frac{P(S = 2)}{P(S \geqslant 1)}. \tag{36.252}$$

L'événement $S = 2$ est réduit au singleton $\{ff\}$ et sa probabilité est $\frac{1}{4}$. Au contraire l'événement $S \geqslant 1$ est l'ensemble $\{fg, gf, ff\}$ et sa probabilité est $\frac{3}{4}$. Nous avons donc

$$P(S = 2|S \geqslant 1) = \frac{1/4}{3/4} = \frac{1}{3}. \tag{36.253}$$

Et là, la réponse est 1/3 et non 1/2 comme d'aucuns auraient pu le croire.

Précision Notons que l'événement $S \geqslant 1$ n'est pas le même que l'événement $X_\sigma = f$. En effet

$$S \geqslant 1 = \{(ff, 1), (ff, 2), (fg, 1), (fg, 2), (gf, 1), (gf, 2)\} \tag{36.254}$$

tandis que

$$\{X_\sigma = f\} = \{(ff, 1), (ff, 2), (fg, 1), (gf, 2)\}. \tag{36.255}$$

36.2.14.4 Conclusion

L'internet regorge de sites discutant du paradoxe des deux enfants [21].

Beaucoup insistent sur le fait que non seulement certaines informations apparemment anodines sont importantes, mais en plus *la façon* dont on obtient l'information est importante. Dans la situation « une fille ouvre », nous obtenons l'information « il y a au moins une fille » en en voyant une ; dans la situation « la parent dit qu'il y a au moins une fille », nous obtenons l'information « il y a au moins une fille » de façon plus « pure ».

Personnellement je ne souscris pas vraiment à cette façon de penser. Le fait est que la formule

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{36.256}$$

n'est pas seulement une formule dans laquelle il faut remplacer A par « la question » et B par « ce qu'on sait ». Il faut également remplacer P par « la bonne » mesure de probabilité.

Il est important de construire le bon espace de probabilité, avec la bonne mesure. Et pour cela, il faut bien s'assurer d'introduire une variable aléatoire pour chaque événement aléatoire se produisant dans l'histoire.

21. Par exemple [774].

36.2.14.5 Une analogie

Imaginez un parc de jeux réservée aux familles de deux enfants dont au moins une fille. Le videur à l'entrée a (au moins) trois stratégies possibles pour ne laisser entrer que les familles contenant au moins une fille.

(1) Regarder l'ensemble de la famille, et ne laisser entrer que les familles ayant au moins une fille.

(2) Ne regarder que l'aîné et ne laisser entrer que les familles dont l'aîné est une fille. ITEMooRKUAooGtfavc

(3) Prendre un des deux enfants au hasard et ne laisser rentrer la famille que si cet enfant est une fille.

Je vous laisse déterminer à quel scénario correspondent ces trois stratégies. Notez que la stratégie (3) contient un énorme biais de sélection : la moitié des familles fille-garçon et garçon-fille sont supprimées, alors que toutes les familles fille-fille entrent.

36.2.14.6 La belle au bois dormant

Vous avez aimé les questions avec les filles qui ouvrent la porte ? Vous trouvez que Monty Hall est trop facile ?

Cadeau de numperphile : « Sleeping Beauty Paradox » https://www.youtube.com/watch?v=cW27QJYNXtU

36.2.15 Inégalité de Jensen

PropABtKbBo

Proposition 36.86 (Inégalité de Jensen)**.**
Soit g une fonction convexe[22] *sur $\mathbb{R}$ et une variable aléatoire $Y \in L^1(\Omega, \mathcal{A}, P)$ telle que $g \circ Y$ soit également L^1. Alors*

$$g\big(E(Y|\mathcal{F})\big) \leqslant E\big((g \circ Y)|\mathcal{F}\big). \tag{36.257}$$

Démonstration. La convexité de g et la proposition 17.91 nous donnent deux suites (a_n) et (b_n) dans $\mathbb{R}$ telles que pour tout $x \in \mathbb{R}$,

$$g(x) = \sup_{n\in\mathbb{N}}(a_n x + b_n). \tag{36.258}$$

Nous avons alors

$$a_n E(Y|\mathcal{F}) + b_n \overset{p.s.}{=} E\big(a_n Y + b_n|\mathcal{F}\big) \leqslant E(g \circ Y|\mathcal{F}). \tag{36.259}$$ EqVAyCziG

L'inégalité est due au fait que $g \circ Y$ est le supremum sur les n de $a_n Y + b_n$. Pour chaque n, l'inégalité (36.259) est fausse sur un ensemble de mesure nulle $R_n \subset \Omega$. L'union

$$R = \bigcup_{n\in\mathbb{N}} R_n \tag{36.260}$$

est encore de mesure nulle. Sur $\Omega \backslash R$, nous avons

$$a_n E(Y|\mathcal{F}) + b_n \leqslant E(g \circ Y|\mathcal{F}). \tag{36.261}$$

Vu que cela est vrai presque partout et pour tout n nous passons au supremum et nous avons encore presque partout l'inégalité

$$\sup_{n\in\mathbb{N}}\big(a_n E(Y|\mathcal{F}) + b_n\big) \leqslant E(g \circ Y|\mathcal{F}). \tag{36.262}$$

□

Si nous ne nous intéressons pas à $E(Y|\mathcal{F})$ mais seulement à $E(Y)$, alors une démonstration plus simple est donnée sur Wikipédia[775].

22. Définition 17.76.

36.2.16 Fonction de répartition

DefooYAZVooNdxDCx

Définition 36.87.
*Si X est une variable aléatoire réelle, nous définissons sa **fonction de répartition** par*

$$\begin{aligned} F_X &: \mathbb{R} \to [0,1] \\ F_X(x) &= P(X \leqslant x). \end{aligned} \tag{36.263}$$

Remarque 36.88.
La fonction de répartition est discontinue en a si $P(X = a) > 0$. En particulier nous ne pouvons pas dire

$$P(X \geqslant a) = 1 - F_X(a). \tag{36.264}$$

LEMooMNXJooBdGXWg

Lemme 36.89.
La fonction de répartition de la loi exponentielle est

$$F(x) = 1 - e^{-\lambda x}, \tag{36.265}$$

36.2.17 Fonction caractéristique

DefooEIVXooNtHLQQ

Définition 36.90.
*La **fonction caractéristique** de la variable aléatoire $X : \Omega \to \mathbb{R}$ est la fonction réelle définie par*

$$\Phi_X(t) = E(e^{itX}). \tag{36.266}$$

Une autre façon d'écrire la définition est

$$\Phi_X(t) = \int_{\mathbb{R}} e^{itX} dP_X(x), \tag{36.267}$$

ou encore, si X a une densité f_X,

$$\Phi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x) dx \tag{36.268}$$

EqFnCaractfncadens

Nous reconnaissons la transformée de Fourier :

$$\Phi_X(t) = \hat{f}_X(-t/2\pi). \tag{36.269}$$

PropDerFnCaract

Proposition 36.91.
Soit X une variable aléatoire qui accepte un moment[23] *d'ordre $r \geqslant 1$. Alors la fonction caractéristique Φ_X est r fois continument dérivable et*

$$\Phi_X^{(r)}(t) = E\big((iX)^r e^{itX}\big). \tag{36.270}$$

EQooVOZTooFDkWjd

Démonstration. Nous étudions la fonction

$$\Phi(t) = \int_\Omega e^{itX(\omega)} dP(\omega). \tag{36.271}$$

Nous prouvons par récurrence que Φ est r fois dérivable et que

$$\Phi^{(k)}(t) = E\big((iX)^k e^{itX}\big) \tag{36.272}$$

pour tout $\leqslant r$.

23. Définition 36.40.

Lorsque $r = 0$ c'est immédiat. Supposons que la récurrence soit correcte pour $k < r$, et prouvons pour $k + 1$. Par hypothèse de récurrence,

$$\Phi^{(k)}(t) = E\big((iX)^k e^{itX}\big) = \int_\Omega \big(iX(\omega)\big)^k e^{itX(\omega)} d\omega. \tag{36.273}$$

Nous posons

$$\begin{aligned} f\colon \mathbb{R} \times \Omega &\to \mathbb{R} \\ (t, \omega) &\mapsto \big(iX(\omega)\big)^k e^{itX(\omega)}, \end{aligned} \tag{36.274}$$

et nous vérifions que ce f satisfait aux hypothèses du théorème 17.19.

(1) $\underline{f_t \in L(\Omega)}$ Pour chaque t fixé, nous posons $f_t(\omega) = f(t, \omega)$. Cette application est dans $L^1(\Omega)$ parce que

$$|f_t(\omega)| = |X(\omega)|^k. \tag{36.275}$$

La proposition 36.41 nous enseigne que X a un moment d'ordre k. Donc l'intégrale de $|f_t|$ existe et est finie.

(2) **Majoration** En posant $G(\omega) = |X(\omega)|^{k+1}$ nous avons $G \in L^1(\Omega)$ parce que $k + 1 \leqslant r$ et que X a des moments finis à tous les ordres jusqu'à r. Et G vérifie aussi

$$|(\partial_t f)(t, \omega)| = \big|\big(iX(\omega)\big)^{k+1} e^{itX(\omega)}\big| = |X(\omega)|^{k+1}. \tag{36.276}$$

Toutes les hypothèses étant vérifiées, nous avons

$$\Phi^{(k+1)}(t) = \frac{d}{dt}\big[f(t, \omega)\big]_{t=0} = \int_\Omega (\partial_t f)(t, \omega) d\omega = E\big((iX)^{k+1} e^{itX}\big), \tag{36.277}$$

ce qu'il fallait. □

Exemple 36.92.
Sachant la fonction caractéristique de X, nous pouvons calculer les moments. Par exemple en posant $r = 2$ et $t = 0$ dans la formule (36.270) nous avons

$$E(X^2) = -\Phi''_X(0). \tag{36.278}$$

△
ThonMxtTy

Théorème 36.93.
Si $\Phi_X = \Phi_Y$, alors $P_X = P_Y$.

Notons que cela n'implique pas que $X = Y$. En effet X et Y peuvent même être définis sur des espaces probabilisés différents.

Dans le cas d'une variable aléatoire vectorielle, nous définissons $\Phi_X\colon \mathbb{R}^d \to \mathbb{R}$ par

$$\Phi_X(v) = E\big(e^{i\langle v, X\rangle}\big) \tag{36.279}$$
EqydyDxg

36.2.18 Fonction génératrice des moments, transformée de Laplace

Soit X une variable aléatoire. Sa **transformée de Laplace** ou **fonction génératrice des moments** est la fonction

$$M_X(t) = E(e^{tX}) \tag{36.280}$$

pour chaque t tel que cette espérance existe.

Théorème 36.94 ([776]).
Soit X une variable aléatoire réelle et

$$I_X = \{t \in \mathbb{R} \text{ tel que } E(e^{tX}) \text{ existe}\}. \tag{36.281}$$

La fonction

$$\begin{aligned} M_X\colon I_X &\to \mathbb{R} \\ t &\mapsto E(e^{tX}) \end{aligned} \tag{36.282}$$

*est la **transformée de Laplace** de X.*

(1) I_X *est un intervalle contenant* 0.
(2) Si I_X *n'est pas réduit à* $\{0\}$ *alors* M_X *se développe en série entière*

$$M_X(t) = \sum_{n=0}^{\infty} \frac{E(X^n)}{n!} t^n. \tag{36.283}$$

(3) Si X *et* Y *sont des variables aléatoires indépendantes, alors* $I_{X+Y} = I_X \cap I_Y$ *et*

$$M_{X+Y} = M_X M_Y \tag{36.284}$$

sur I_{X+Y}.

Démonstration. Le fait que 0 soit dans I_X est évident : $E(1) = 1$. Pour montrer que I_X est un intervalle nous prenons $z \in I_X$ et $0 < s < z$ ou $z < s < 0$, puis nous montrons que $s \in I_X$. Il faut remarquer que dans tous les cas,

$$e^{sX} \leqslant 1 + e^{zX}. \tag{36.285}$$

En effet soit sX et zX sont tous deux à gauche de zéro et alors ils sont tous deux plus petit que 1 ; soit ils sont tous deux à droite de 0 et alors $e^{zX} > e^{sX}$ par croissance de l'exponentielle. Nous avons donc dans tous les cas que

$$E(e^{sX}) = \int_{\mathbb{R}} f_X(x) e^{sX} dx \leqslant \int_{\mathbb{R}} f_X(x)(1 + e^{zx}) = 1 + E(e^{zX}). \tag{36.286}$$

Soit maintenant $a > 0$ tel que $[-a, a] \subset I_X$. Étant donné que $e^{a|X|} < e^{aX} e^{-aX}$, l'espérance $E(e^{a|X|})$ existe toujours pour $|t|$. Nous avons

$$\left| M_X(t) - \sum_{n=0}^{N} \frac{E(X^n)}{n!} t^n \right| = \left| E\Big(e^{tX} - \sum_{n=0}^{N} \frac{X^n}{n!} t^n \Big) \right| \tag{36.287a}$$

$$= \left| E\Big(\sum_{n=N+1}^{\infty} \frac{X^n}{n!} t^n \Big) \right| \tag{36.287b}$$

$$\leqslant E\left(\sum_{n=N+1}^{\infty} \frac{|tX|^n}{n!} \right). \tag{36.287c}$$

Maintenant le but est de prendre la limite $N \to \infty$ en inversant la limite et l'espérance par le théorème de la convergence dominée (14.207). L'intégrale à traiter est

$$\lim_{N\to\infty} \int_{\Omega} \sum_{n=N+1}^{\infty} \frac{|tX(\omega)|^n}{n!} dP(\omega). \tag{36.288}$$

L'intégrande est uniformément borné (en N) par $e^{tX(\omega)}$, qui est intégrable par hypothèse (choix de t). Du coup

$$\lim_{N\to\infty} \int_{\Omega} \sum_{n=N+1}^{\infty} \frac{|tX(\omega)|^n}{n!} dP(\omega) = E\left(\lim_{N\to\infty} \sum_{n=N+1}^{\infty} \frac{|tX|^n}{n!} \right) = 0. \tag{36.289}$$

□

36.2.19 Théorème de transfert

LEMooHVFJooBDZYnT

Lemme 36.95.
Soient un espace de probabilité $(\Omega, \mathcal{A}, P)$ *ainsi qu'une variable aléatoire* $X : \Omega \to \mathbb{R}^n$. *Nous considérons un autre espace probabilisé* $(S, \mathcal{A}_S, \mu)$ *ainsi que l'application*

$$\begin{aligned} Y : \Omega \times S &\to \mathbb{R}^n \\ (\omega, s) &\mapsto X(\omega). \end{aligned} \tag{36.290}$$

Alors :

ITEMooDSIKooSqdyJb

(1) L'application Y *est une variable aléatoire sur*[24] $\big(\Omega \times S, \mathcal{A} \otimes \mathcal{A}_S, P \otimes \mu\big)$.

24. Produit de tribus, définition 14.127 ; produit de mesure, définition 14.246.

ITEMooQMMHooPcGKoD

(2) La loi de Y est la même[25] *que celle de X.*

Démonstration. Tout repose sur le fait que Y soit la même chose que X, mais « diagonalisé » sur S. Commençons par prouver que si $B \subset \mathbb{R}^n$, alors $Y^{-1}(B) = X^{-1}(B) \times S$.

(1) **Dans un sens** Soit $(\omega, s) \in X^{-1}(B) \times S$. Alors $Y(\omega, s) = X(\omega) \in B$. Nous avons donc bien $X^{-1}(B) \times S \subset Y^{-1}(B)$.

(2) **Dans l'autre sens** Soit $(\omega, s) \in Y^{-1}(B)$. Nous avons $X(\omega) = Y(\omega, s) \in B$, donc $\omega \in X^{-1}(B)$ et donc $(\omega, s) \in X^{-1}(B) \times S$.

Passons maintenant à la preuve proprement dite.

(1) **Pour (1)** Soit un borélien $B \subset \mathbb{R}^n$. Nous avons $Y^{-1}(B) = X^{-1}(B) \times S$. La partie $X^{-1}(B)$ est un borélien de Ω, et S est un borélien de S. Donc $X^{-1}(B) \times S$ est un borélien de $\Omega \times S$. L'application Y est donc mesurable. Elle est une variable aléatoire.

(2) **Pour (2)** Soit un borélien B de $\mathbb{R}^n$. Nous avons

$$\begin{aligned} P_Y(B) &= (P \otimes \mu)\big(Y^{-1}(B)\big) && (36.291a)\\ &= (P \otimes \mu)\big(X^{-1}(B) \times S\big) && (36.291b)\\ &= P\big(X^{-1}(B)\big)\mu(S) && (36.291c)\\ &= P\big(X^{-1}(B)\big) && (36.291d)\\ &= P_X(B). && (36.291e) \end{aligned}$$

Nous avons utilisé le fait que $\mu(S) = 1$.

□

La proposition suivante permet de calculer en pratique les intégrales qui définissent par exemple l'espérance mathématique d'une variable aléatoire.

PropintdPintdPXeR

Proposition 36.96 (Théorème de transfert[778])**.**
Si $X : \Omega \to \mathbb{R}^d$ est une variable aléatoire, alors dès que $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ est telle qu'une des deux intégrales existe,

$$E(f \circ X) = \int_\Omega f\big(X(\omega)\big) dP(\omega) = \int_{\mathbb{R}^d} f(x) dP_X(x) \tag{36.292}$$

où P_X est la loi de X[26]*.*

En particulier, ça marche si f est borélienne.

En utilisant cette proposition nous trouvons une formule pratique pour l'espérance d'une variable aléatoire réelle :

$$E(X) = \int_\Omega X(\omega) dP(\omega) = \int_{\mathbb{R}} x dP_X(x), \tag{36.293}$$

en vertu de la proposition 36.96 appliquée à la fonction $f(x) = x$.

PROPooNNTNooOmCLvy

Proposition 36.97.
Une variable aléatoire réelle X est intégrable si et seulement si $P(x = \pm\infty) = 0$ et

$$\int_{\mathbb{R}} |x| dP_X(x) < \infty. \tag{36.294}$$

Le lien entre la densité f_X de la variable aléatoire X et sa loi est

$$P_X(A) = \int_A f_X(x) dx \tag{36.295}$$

25. La même ; pas seulement la même à abus de notation près. Ce lemme est une partie du bout manquant dont il est question dans [777].
26. Définition 36.8.

pour tout ensemble mesurable $A \subset \mathbb{R}$. Le lien entre la mesure de Lebesgue et celle de la loi de X est alors donné par

$$dP_X(x) = f_X(x)dx. \tag{36.296}$$

En particulier l'espérance de X peut être calculée à partir de sa densité via la formule

$$E(X) = \int_{\mathbb{R}} x dP_X(x) = \int_{\mathbb{R}} x f_X(x)dx. \tag{36.297}$$

EqEspDensform

LEMooYVJAooXUkRTh

Lemme 36.98.
Soit une variable aléatoire $X\colon \Omega \to \mathbb{R}^n$ de densité g_X.

ITEMooBUWNooJkiRPp

(1) Nous avons

$$E(X) = \int_{\mathbb{R}^n} x g_X(x)dx. \tag{36.298}$$

ITEMooKZHAooKQQmno

(2) Si la fonction $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ est borélienne, alors

$$E(f \circ X) = \int_{\mathbb{R}^n} f(x) g_X(x)dx. \tag{36.299}$$

¡¡ Avertissement/question au lecteur !! 36.99
Je n'ai pas vérifié le lemme 36.98. Écrivez-moi si vous en connaissez une démonstration, ou si vous croyez qu'il est faux.

À mon avis le point (1) est une variation autour de (36.297) et (2) utilise ça et le théorème de transfert 36.96.

LEMooAVTFooIQRvop

Lemme 36.100.
Soient des variables aléatoires $X_i\colon \Omega \to \mathbb{R}^n$ $(i = 1, \dots, n)$. Nous supposons qu'il existe une application $g\colon \mathbb{R}^n \to \mathbb{R}$ telle que

$$E\big(f(X_1, \dots, X_n)\big) = \int_{\mathbb{R}^n} f(x)g(x)dx \tag{36.300}$$

pour tout fonction borélienne $f\colon \mathbb{R}^n \to \mathbb{R}$.

Alors g est une densité pour la variable aléatoire $(X_1, \dots, X_n)\colon \Omega \to \mathbb{R}$.

PROPooNYNUooTCJgpl

Proposition 36.101 ([779])**.**
Soient des variables aléatoires $X_i\colon \Omega \to \mathbb{R}$ $(i = 1, \dots, n)$. Nous posons $Z = (X_1, \dots, X_n)$ et nous supposons que Z est une variable aléatoire à densité notée f_Z.

Alors les X_i sont à densité et les densités de X_i sont

$$f_i(t) = \int_{\mathbb{R}^{n-1}} f_Z(x_1, \dots, t, x_{i+1}, \dots, x_n)dx_1, \dots, dx_{i-1}dx_{i+1} \dots dx_n. \tag{36.301}$$

PROPooVUTLooDPdhxK

Proposition 36.102.
Soient des variables aléatoires à densité $X_i\colon \Omega \to \mathbb{R}$ $(i = 1, \dots, n)$; nous notons f_i leurs densités. Nous posons $Z = (X_1, \dots, X_n)$.

(1) La variable aléatoire Z est à densité.

ITEMooUQEFooAAFXJY

(2) Les variables aléatoires X_i sont indépendantes si et seulement si $f_Z = \prod_{i=1}^n f_i$.

¡¡ Avertissement/question au lecteur !! 36.103
Je n'ai pas vérifié si cette proposition est vraie. Écrivez-moi si vous connaissez une démonstration, ou si vous savez que c'est faux.

36.2.20 Changement de variables

THOooDSTYooXMWfNw

Théorème 36.104.
Soit $\mathcal{O}$, un ouvert de $\mathbb{R}^n$ et $\mathcal{O}'$ un ouvert de $\mathbb{R}^m$ ainsi qu'un difféomorphisme C^1 $\varphi\colon \mathcal{O} \to \mathcal{O}'$. Soit $X\colon \Omega \to \mathbb{R}^n$ une variable aléatoire prenant presque surement ses valeurs dans $\mathcal{O}$. Si nous supposons que X a la densité f_X, alors la variable aléatoire $Y = \varphi(X)$ accepte la densité $f_Y\colon \mathcal{O}' \to \mathbb{R}$ donnée par

$$f_Y(v) = f_X\big(\varphi^{-1}(v)\big)|J_{\varphi^{-1}}(v)|. \tag{36.302}$$

Démonstration. Nous devons vérifier la relation

$$P(Y \in B) = \int_B f_Y(v)dv \tag{36.303}$$

pour tout borélien $B \subset \mathcal{O}'$. Nous avons

$$\begin{aligned}
P(Y \in B) &= \int_{\mathbb{R}^m} \mathbb{1}_B(v)dP_Y(v) && \text{(36.304a)}\\
&= E(\mathbb{1}_B \circ Y) && \text{(36.304b)}\\
&= E\big((\mathbb{1}_B \circ \varphi) \circ X\big) && \text{(36.304c)}\\
&= \int_{\mathbb{R}^n} (\mathbb{1}_B \circ \varphi)(u)dP_X(u) && \text{(36.304d)}\\
&= \int_{\mathbb{R}^n} (\mathbb{1}_B \circ \varphi)(u)f_X(u)du. && \text{(36.304e)}
\end{aligned}$$

À ce niveau, nous utilisons la formule de changement de variables du théorème 14.281. Nous trouvons alors

$$P(Y \in B) = \int_{\mathbb{R}^m} \mathbb{1}_B\big(\varphi^{-1}(v)\big)f_X\big(\varphi^{-1}(v)\big)|J_{\varphi^{-1}}(v)|dv. \tag{36.305}$$

□

36.3 Convergence

DEFooZKLFooZkKuMC

Définition 36.105 ([780, 672])**.**
Soit un espace de probabilité $(\Omega, \mathcal{A}, P)$. Soient des variables aléatoires[27] *$X_i\colon \Omega \to \mathbb{R} \cup \{\pm\infty\}$.*
Nous avons plein de modes de convergence.

Convergence presque sure *Nous disons que X_i converge **presque surement** vers la variable aléatoire X et nous notons*

$$X_n \xrightarrow{p.s.} X \tag{36.306}$$

si

$$P\big(\{\omega \in \Omega \text{ tel que } X_n(\omega) \to X(\omega)\}\big) = 1 \tag{36.307}$$

où la convergence $X_n(\omega) \to X(\omega)$ est la convergence usuelle dans $\mathbb{R} \cup \{\pm\infty\}$.

Convergence en probabilité *Nous disons que les variables aléatoires réelles X_n convergent **en probabilité** vers la variable aléatoire X si pour tout $\eta > 0$, on a*

$$P\big(|X_n - X| \geqslant \eta\big) \to 0, \tag{36.308}$$

et on note

$$X_n \xrightarrow{P} X. \tag{36.309}$$

27. Définition 12.24 pour la topologie et donc les boréliens sur la droite réelle achevée.

Convergence L^p *Si X_n et X sont de classe L^p, alors nous pouvons également parler de la convergence L^p :*

$$X_n \xrightarrow{L^p} X. \tag{36.310}$$

En notant les intégrales comme des espérances, cette convergence signifie

$$E\big(|X_n - X|^r\big) \to 0. \tag{36.311}$$

Convergence en loi *Nous disons que X_n converge vers X **en loi** vers la variable aléatoire X et nous notons*

$$X_n \xrightarrow{\mathscr{L}} X \tag{36.312}$$

EQooWSMPooHCyPwk

si pour toute fonction continue et bornée g nous avons

$$E\big(g(X_n)\big) \to E\big(g(X)\big). \tag{36.313}$$

La convergence en loi se distingue des autres modes de convergence en cela que les variables aléatoires X_n et X peuvent être définies sur des espaces probabilisés $(\Omega_n, \mathcal{A}_n, P_n)$ et $(\Omega, \mathcal{A}, P)$ tous différents.

THOooMXFMooVHdSFE

Théorème 36.106 ([672]).
Soient des variables aléatoires X_n et X sur l'espace probabilité $(\Omega, \mathcal{A}, P)$.

(1) Si $X_n \xrightarrow{p.s.} X$ alors $X_n \xrightarrow{P} X$.

ITEMooRPLIooHcFLCR

(2) Si $X_n \xrightarrow{P} X$ alors il existe une sous-suite (Y_n) de X_n telle que $Y_n \xrightarrow{p.s.} X$.

(3) Si les variables aléatoires X_n sont de classe L^r et si $X_n \xrightarrow{L^r} X$, alors $X_n \xrightarrow{P} X$.

ITEMooAYRCooCYLpWX

(4) Si $X_n \xrightarrow{P} X$ et si la famille $\{|X|^r\}$ est uniformément intégrable[28]*, alors $X_n \xrightarrow{L^r} X$.*

(5) Si $X_n \xrightarrow{P} X$ et si $\limsup_n E\big(|X_n|^r\big) \leqslant E(|X|^r)$ alors $X_n \xrightarrow{L^r} X$.

(6) Si $X_n \xrightarrow{L^r} X$, alors $X_n \xrightarrow{L^p} X$ pour tout $0 < p \leqslant r$.

ITEMooPRGRooUqNKjB

(7) Si $X_n \xrightarrow{P} X$ alors $X_n \xrightarrow{\mathscr{L}} X$.

(8) Si $X_n \xrightarrow{P} X$ alors toute sous-suite (Y_n) contient une sous-sous-suite (Y'_n) telle que $Y'_n \xrightarrow{p.s.} X$

Notez les implications

$$\textit{presque sure} \;\Rightarrow\; \textit{en probabilité} \;\Rightarrow\; \textit{en loi}. \tag{36.314}$$

PROPooXRVOooEYQzGr

Proposition 36.107 ([1]).
Si X_n et X est sont des variables aléatoires telles que $X_n \xrightarrow{\mathscr{L}} X$, alors

$$|X_n| \xrightarrow{\mathscr{L}} |X|. \tag{36.315}$$

LEMooADHZooGCxVvn

Lemme 36.108.
Nous avons $X_n \xrightarrow{p.s.} X$ si et seulement si il existe un événement $A \in \mathcal{A}$ tel que $P(A) = 1$ et tel que $X_n(\omega) \to X(\omega)$ pour tout $\omega \in A$.

LEMooNTNIooKbLwgX

Lemme 36.109.
Si X est une variable aléatoire à valeurs dans $\mathbb{R} \cup \{\pm\infty\}$, alors

(1) Pour tout $n \in \mathbb{N}$, l'application

$$\begin{aligned} \min(X, n) \colon \Omega &\to \mathbb{R} \cup \{\pm\infty\} \\ \omega &\mapsto \min\big(X(\omega), n\big) \end{aligned} \tag{36.316}$$

est une variable aléatoire.

28. Voir lemme 27.57.

(2) *Nous avons la convergence presque sure*[29]

$$\min(X,n) \xrightarrow{p.s.} X \tag{36.317}$$

Démonstration. Le fait que $\min(X,n)$ soit mesurable est le lemme 14.99. D'ailleurs pour la suite, nous la notons $X_n = \min(X,n)$. Sinon il faudrait écrire des choses comme $\min(X,n)(\omega)$ pour $X_n(\omega)$, et ça deviendrait compliqué.

Nous allons montrer que pour tout $\omega \in \Omega$, nous avons $X_n(\omega) \to X(\omega)$. Ainsi le lemme 36.108 conclura.

Soit $\omega \in \Omega$. Si $X(\omega) = \infty$, alors nous avons $X_n(\omega) = n$ pour tout n et alors $X_n(\omega) \xrightarrow{\mathbb{R}} \infty = X(\omega)$.

Si par contre $X(\omega) < \infty$. Alors en choisissant $N > X(\omega)$ dans $\mathbb{N}$ (lemme 1.438), nous avons $X_n(\omega) = X(\omega)$ pour tout $n \geqslant N$. Et donc aussi $X_n(\omega) \to X(\omega)$. □

36.110.
Dans certains textes, la variable aléatoire $\min(X,n)$ est notée $X \wedge n$. Donc le lemme 36.109 s'écrit

$$X \wedge n \xrightarrow{p.s.} X \tag{36.318}$$

PROPooDWTWooPygbmK

Proposition 36.111 ([1, 672]).
Si $X_n \xrightarrow{p.s.} X$ et si X est intégrable, alors

$$E(X) = E\big(\liminf_n(X_n)\big). \tag{36.319}$$

Démonstration. Soit

$$A = \{\omega \in \Omega \text{ tel que } X_n(\omega) \to X(\omega)\}. \tag{36.320}$$

Vu que $X_n \xrightarrow{p.s} X$, nous avons $P(A) = 1$ et donc, en posant $Z = \Omega \backslash A$, nous avons $P(Z) = 0$. Nous avons

$$\begin{aligned} E\big(\liminf(X_n)\big) &= \int_\Omega \liminf_n(X_n)d\omega && (36.321a)\\ &= \int_A \liminf_n(X_n)d\omega + \int_Z \liminf_n(X_n)d\omega && (36.321b)\end{aligned}$$

La seconde intégrale est nulle par le corolaire 14.168 ; vous noterez que rien ne garanti que $\liminf_n X_n(\omega) < \infty$ lorsque $\omega \in Z$ parce que rien ne garantit que $\liminf_n X_n(\omega) = X'\omega$.

Pour $\omega \in A$ nous avons $X_n(\omega) \to X(\omega)$, et à fortiori $\liminf_n X_n(\omega) = X(\omega)$. Donc

$$\begin{aligned} E\big(\liminf_n(X_n)\big) &= \int_A \liminf_n\big(X_n(\omega)\big)d\omega && (36.322a)\\ &= \int_A X(\omega)d\omega && (36.322b)\\ &= \int_A X(\omega)d\omega + \int_Z X(\omega)d\omega && (36.322c)\\ &= \int_\Omega X(\omega)d\omega && (36.322d)\\ &= E(X). && (36.322e)\end{aligned}$$

□

THOooEPYBooHlSuLI

Théorème 36.112 (Théorème de Vitali[672]).
Soient des variables aléatoires $X_n \in L^r(\Omega,\mathcal{A},P)$ avec $0 < r < \infty$. Nous supposons que[30] $X_n \xrightarrow{P} X$. *Nous avons équivalence entre*

29. Définition 36.105.
30. Convergence en probabilité, définition 36.105.

SPITEMooGKRVooQoDfKm

(1) *L'ensemble* $\{|X_n|^r\}$ *est uniformément intégrable*[31].

SPITEMooUYTKooBClYcR

(2) *Nous avons* $X \in L^r(\Omega)$ *et* $X_n \xrightarrow{L^r(\Omega)} X$.

SPITEMooQYLOooDiBvne

(3) $E\big(|X_n|^r\big) \to E\big(|X|^r\big)$.

SPITEMooRVFAooXUNQlf

(4) $\limsup_n E\big(|X_n|^r\big) \leqslant E\big(|X|^r\big)$.

Démonstration. En plusieurs parties.

(1) **(1) ⇒ (2)**

Par le théorème 36.106(2), il existe une sous-suite (Y_n) de (X_n) telle que $Y_n \xrightarrow{p.s.} X$. Nous avons donc également $|Y_n|^r \xrightarrow{p.s.} |X|^r$. Nous utilisons 36.111 :

$$E\big(|X|^r\big) = E\big(\liminf_n |Y_n|^r\big) \quad \text{prop. 36.111} \tag{36.323a}$$

$$\leqslant \liminf_n \Big(E\big(|Y_n|^r\big)\Big) \quad \text{Fatou 14.182.} \tag{36.323b}$$

Nous nous souvenons que la partie $\{|X_n|^r\}$ est uniformément intégrable ; à fortiori la partie $\{|Y_n|^r\}$ l'est également. Le théorème 27.61(1) dit qu'il existe $M > 0$ tel que $E\big(|Y_n|^r\big) < M$. Nous avons donc

$$E\big(|X|^r\big) \leqslant \liminf_n \Big(E\big(|Y_n|^r\big)\Big) < M. \tag{36.324}$$

Cela prouve que $|X|^r \in L^1(\Omega)$, c'est-à-dire que $X \in L^r(\Omega)$. Voilà une bonne chose de faite. Nous devons encore prouver que $X_n \xrightarrow{L^r(\Omega)} X$.

Soit $\omega \in \Omega$. La c-r inégalité (proposition 27.139) donne

EQooZYDGooKhtrqz

$$|X_n(\omega) - X(\omega)|^r \leqslant c_r\big(|X_n(\omega)|^r + |X(\omega)|^r\big). \tag{36.325}$$

Vu que $\{|X_n(\omega)|\}_{n\in\mathbb{N}}$ est uniformément intégrable, les différents points de la proposition 27.59 et l'inégalité 36.325 montrent que $\{|X_n - X|^r\}_{n\in\mathbb{N}}$ est uniformément intégrable.

Nous utilisons la proposition 27.61. Soit $\epsilon > 0$. Il existe $\delta_\epsilon > 0$ tel que pour tout n et pour toute partie $A \in \mathcal{A}$ telle que $P(A) \leqslant \delta_\epsilon$,

$$E\big(|X_n - X|^r \mathbb{1}_A\big) < \epsilon. \tag{36.326}$$

Vu que $X_n \xrightarrow{P} X$, si n est assez grand[32], nous avons $P\big(|X_n - X| \geqslant \epsilon\big) < \delta_\epsilon$. Pour de tels n, nous avons

$$E\Big(|X_n - X|^r \mathbb{1}_{\{|X_n - X| \geqslant \epsilon\}}\Big) < \epsilon. \tag{36.327}$$

Par la σ-additivité de l'intégrale (et le recouvrement en $\{|X_n - X| = \epsilon\}$), nous avons

$$E\big(|X_n - X|^r\big) \leqslant E\Big(|X_n - X|^r \mathbb{1}_{\{|X_n - X| \geqslant \epsilon\}}\Big) + E\Big(|X_n - X|^r \mathbb{1}_{\{|X_n - X| \leqslant \epsilon\}}\Big) \tag{36.328a}$$

SUBEQooHHJSooMueQQp

$$< \epsilon + E\Big(|X_n - X|^r \mathbb{1}_{\{|X_n - X| \leqslant \epsilon\}}\Big) \tag{36.328b}$$

Sur la partie $\{|X_n - X| \leqslant \epsilon\}$, nous pouvons majorer $|X_n(\omega) - X(\omega)| \leqslant \epsilon$. Vu que $P(\Omega) = 1$ nous avons

$$E\Big(|X_n - X|^r \mathbb{1}_{\{|X_n - X| \leqslant \epsilon\}}\Big) = \int_\Omega |X_n(\omega) - X(\omega)|^r \mathbb{1}_{\{|X_n - X| \leqslant \epsilon\}} dP(\omega) \tag{36.329a}$$

$$\leqslant \int_\Omega \epsilon^r dP(\omega) = \epsilon^r \tag{36.329b}$$

31. Définition 27.56.
32. Dans [672], il choisit ϵ petit et non n grand. Je ne comprends pas pourquoi.

Cela nous permet de majorer la seconde partie de (36.328b). Au final si n est grand,

$$E\big(|X_n - X|^r\big) \leqslant \epsilon + \epsilon^r. \tag{36.330}$$

Autrement dit, $E\big(|X_n - X|^r\big) \to 0$ ou encore

$$X_n \xrightarrow{L^r} X. \tag{36.331}$$

(2) **(2) ⇒ (3)**
C'est la proposition 27.64.

(3) **(3) ⇒ (4)**
Si la limite existe, alors, à fortiori, la limite supérieure existe et est égale.

(4) **(4) ⇒ (1)**
Nous supposons que $E\big(|X_n|^r\big) \leqslant E\big(|X|^r\big)$. Pour chaque $\lambda > 0$, nous considérons la fonction

$$\begin{aligned} f_\lambda\colon [0,\infty[&\to \mathbb{R} \\ x &\mapsto \begin{cases} x^r & \text{si } x^r \leqslant \lambda \\ \lambda^r - \lambda^r(x-\lambda) & \text{si } \lambda < x^r < \lambda + 1 \\ 0 & \text{si } x^r \geqslant \lambda + 1. \end{cases} \end{aligned} \tag{36.332}$$

(4a) **f_λ est continue** Si $x = \lambda$, la première ligne donne $f_\lambda(\lambda) = \lambda^r$ (parce que $\lambda > 0$), qui coïncide avec la valeur de la seconde ligne en $x = \lambda$. Si $x = \lambda + 1$, les seconde et troisième lignes donnent toutes deux 0.

(4b) **f_λ est bornée** Chacun des trois bouts de formule est borné.

(4c) **Une majoration avec f_λ** Soit $g\colon \Omega \to \mathbb{R}$. Nous montrons que

$$f_\lambda \circ |g| \leqslant |g|^r \mathbb{1}_{\{|g|^r \leqslant \lambda+1\}}. \tag{36.333}$$ EQooVLGYooYEiVoS

Il y a trois possibilités en fonction de ce que vaut $|g(\omega)|^r$.

(4c.iii) **Si $|g(\omega)| \leqslant \lambda$** Alors $(f_\lambda \circ |g|)(\omega) = |g(\omega)|^r$ alors que $\mathbb{1}_{\{|g|^r \leqslant \lambda+1\}}(\omega) = 1$.

(4c.iii) **Si $\lambda < |g(\omega)|^r < \lambda + 1$** Nous avons

$$\begin{aligned} (f_\lambda \circ g)(\omega) &= \lambda^r - \lambda^r\big(g(\omega) - \lambda\big) && && \text{(36.334a)} \\ &< \lambda^r - \lambda^r(\lambda - \lambda) && \text{pcq. } g(\omega) < \lambda && \text{(36.334b)} \\ &= \lambda^r && && \text{(36.334c)} \\ &\leqslant |g(\omega)|^r. && && \text{(36.334d)} \end{aligned}$$

(4c.iii) **Si $|g(\omega)|^r \geqslant \lambda + 1$** Nous avons $(f_\lambda \circ g)(\omega) = 0$.

(4d) **Une autre inégalité avec f_λ** Nous prouvons à présent que

$$|X|^r - (f_\lambda \circ X) \leqslant |X|^r \mathbb{1}_{\{|X|^r \geqslant \lambda\}}. \tag{36.335}$$ EQooNWRQooBcZZTo

(4d.iv) **Si $|X(\omega)| \leqslant \lambda$** Dans ce cas nous avons 0 à gauche de (36.335), et pas besoin de regarder à droite.

(4d.iv) **Si $\lambda < |X(\omega)|^r < \lambda + 1$** À gauche nous avons

$$\begin{aligned} \lambda^r - \lambda^r\big(|X(\omega)| - \lambda\big) &\leqslant \lambda^r - \lambda^r(\lambda - \lambda) && \text{(36.336a)} \\ &= \lambda^r && \text{(36.336b)} \\ &\leqslant |X(\omega)|^r. && \text{(36.336c)} \end{aligned}$$

Et à droite, nous avons la même chose.

(4d.iv) **Si $|X(\omega)| \geqslant \lambda + 1$** Alors nous avons 0 à gauche.

(4e) **Une limite avec l'espérance** Par hypothèse, $X_n \xrightarrow{P} X$. Par le théorème 36.106(7), cela implique la convergence en loi $X_n \xrightarrow{\mathscr{L}} X$. Et donc $|X_n| \xrightarrow{\mathscr{L}} |X|$ et aussi [33]

$$E(f_\lambda \circ |X_n|) = E(f_\lambda \circ |X|) \tag{36.337}$$

EQooJGSVooBXkEnC

parce que f_λ est continue et bornée.

Avec tout ça, nous pouvons calculer un petit peu :

$$\begin{aligned}
\limsup_n E\Big(|X_n|^r \mathbb{1}_{\{|X_n|^r>\lambda+1\}}\Big) &= \limsup_n \Big(E(|X_n|^r) - E\big(|X_n|^r \mathbb{1}_{\{|X_n|^r\leqslant\lambda+1\}}\big)\Big) && \text{(36.338a)}\\
&\leqslant E(|X|^r) - \limsup_n \Big(E\big(|X_n|^r \mathbb{1}_{\{|X_n|^r\leqslant\lambda+1\}}\big)\Big) && \text{(36.338b)}\\
&\leqslant E(|X|^r) - \liminf_n \Big(E\big(|X_n|^r \mathbb{1}_{\{|X_n|^r\leqslant\lambda+1\}}\big)\Big) && \text{(36.338c)}\\
&\leqslant E(|X|^r) - \liminf_n E(f_\lambda \circ |X_n|) && \text{de (36.333)}\\
& && \text{(36.338d)}\\
&\leqslant E(|X|^r) - E(f_\lambda \circ |X|) && \text{de (36.337)}\\
& && \text{(36.338e)}\\
&= E\big(|X|^r - (f_\lambda \circ |X|)\big) && \text{(36.338f)}\\
&\leqslant E\big(|X|^r \mathbb{1}_{\{|X|^r\geqslant\lambda\}}\big) && \text{(36.338g)}
\end{aligned}$$

Nous avons donc démontré que pour tout $\lambda > 0$,

$$\limsup_n E\Big(|X_n|^r \mathbb{1}_{\{|X_n|^r>\lambda+1\}}\Big) \leqslant E\big(|X|^r \mathbb{1}_{\{|X|^r\geqslant\lambda\}}\big). \tag{36.339}$$

Pour tester l'uniforme intégrabilité, il nous suffit maintenant de prendre la limite $\lambda \to \infty$. La proposition 27.60 dit que $E\big(|X|^r \mathbb{1}_{\{|X|^r\geqslant\lambda\}}\big) \xrightarrow{\lambda\to\infty} 0$. Nous avons donc bien l'uniforme intégrabilité.

□

36.3.1 Convergence et fonction caractéristique

PrpopCaractCvL

Proposition 36.113 (Helly-Bray[672])**.**
Deux autres caractérisations de la convergence en loi[34]*.*

(1) Nous avons $X_n \xrightarrow{\mathscr{L}} X$ si et seulement si

$$\Phi_{X_n}(v) \to \Phi_X(v) \tag{36.340}$$

pour tout $v \in \mathbb{R}^d$. Ici Φ_X est la fonction caractéristique de X.

(2) Dans la définition de la convergence en loi nous pouvons indifféremment utiliser les fonctions continues et bornées, les fonctions continues à support compact ou les fonctions bornées uniformément continues.

Dans le cas particulier $d = 1$ nous avons quelques critères supplémentaires.

PropoFnrepCvL

Proposition 36.114.
Supposons que les variables aléatoires X_n soient réelles, et notons F_n la fonction de répartition de X_n. Si $F_n(x) \to F(x)$ pour tout x dans l'ensemble des points de continuité de F, alors $X_n \xrightarrow{\mathscr{L}} X$.

PROPooNRQJooXoCcbZ

Proposition 36.115.
Si les X_n sont des variables aléatoires réelles positives, et si X est une variable aléatoire positive,

33. Proposition 36.107.
34. Définition 36.312

alors $X_n \xrightarrow{\mathscr{L}} X$ si les transformées de Laplace des fonctions de répartition convergent ponctuellement, c'est-à-dire si

$$E\big(e^{-\alpha X_n}\big) \to E\big(e^{-\alpha X}\big) \tag{36.341}$$

pour tout $\alpha \geqslant 0$.

PROPooZMHKooLwCwqs

Proposition 36.116.
Si les X_n et X sont des variables aléatoires réelles discrètes à valeurs dans $\{x_0, x_1, \ldots\}$ alors $X_n \xrightarrow{\mathscr{L}} X$ si et seulement si

$$P(X_n = x_k) \to P(X = x_k) \tag{36.342}$$

pour tout $k \in \mathbb{N}$.

PropXncvXFXcvFxt

Proposition 36.117 ([762])**.**
Soient X_n et X des variables aléatoires réelles. Nous avons

$$X_n \xrightarrow{\mathscr{L}} X \tag{36.343}$$

si et seulement si pour tout t où F_X est continue,

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t). \tag{36.344}$$

PropCvLfcvPsicst

Proposition 36.118 ([762])**.**
Soit X_n une suite de variables aléatoires $\Omega \to \mathbb{R}^d$ et $a \in \mathbb{R}^d$. Si $X_n \xrightarrow{\mathscr{L}} a$, alors

$$X_n \xrightarrow{P} a. \tag{36.345}$$

Démonstration. Quitte à passer aux composantes, nous pouvons supposer que $d = 1$. Soit $\eta > 0$; nous savons que l'inégalité $|x| > a$ a pour solution $x > a$ ou $x < -a$. Dans notre cas,

$$\begin{aligned}
P\big(|X_n - a| > \eta\big) &= P(X_n - a > \eta) + P(X_n - a < -\eta) && \text{(36.346a)}\\
&= P(X_n > \eta + a) + P(X_n < a - \eta) && \text{(36.346b)}\\
&= 1 - P(X_n \leqslant \eta + a) + P(X_n \leqslant a - \eta) - P(X_n = a - \eta) && \text{(36.346c)}\\
&\leqslant 1 - F_{X_n}(\eta + a) + F_{X_n}(a - \eta) && \text{(36.346d)}
\end{aligned}$$

où la majoration est l'oubli du terme $P(X_n = a-\eta)$, lequel est positif ou nul et F_{X_n} est la fonction de répartition de X_n, définition 36.87. Nous allons utiliser la proposition 36.117. La fonction de répartition de la variable aléatoire constante $X = a$ est donnée par

$$F_a(t) = P(a \leqslant t) = \mathbb{1}_{[0,\infty[}(t - a). \tag{36.347}$$

Par conséquent, la convergence en loi $X_n \xrightarrow{\mathscr{L}} a$ nous montre que

$$F_{X_n}(t) \to \mathbb{1}_{[0,\infty[}(t - a) \tag{36.348}$$

pour tout $t \neq a$ parce que $t = 0$ est un point de discontinuité de $\mathbb{1}_{[0,\infty[}$. Nous avons par conséquent

$$P\big(|X_n - a| > \eta\big) = 1 - \mathbb{1}_{[0,\infty[}(\eta) + \mathbb{1}_{[0,\infty[}(-\eta) = 1 - 1 + 0 = 0 \tag{36.349}$$

parce que $\eta > 0$. □

Le lemme de Slutsky sera utilisé en combinaison avec la proposition 36.120 pour calculer des intervalles de confiance, voir par exemple ce qui se passe autour de l'équation (37.126).

LemgXDlhs

Lemme 36.119 (Slutsky[781]).
Soient X_n et Y_n des suites de variables aléatoires réelles telles que

$$\begin{aligned} X_n &\xrightarrow{\mathscr{L}} X \\ Y_n &\xrightarrow{P} a \in \mathbb{R}. \end{aligned} \tag{36.350}$$

Alors $(X_n, Y_n) \xrightarrow{\mathscr{L}} (X, a)$.

Démonstration. Étant donné que $Y_n \xrightarrow{\mathscr{L}} a$, nous avons $Y_n \xrightarrow{P} a$ par la proposition 36.118. Soit une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}^2$; nous devons prouver que

$$E\big(f(X_n, Y_n)\big) \to E\big(f(X, a)\big). \tag{36.351}$$

Soit $\epsilon > 0$. Nous avons

EqEparXnYnfXa
$$E\big(\|f(X_n, Y_n) - f(X, a)\|\big) \leqslant E\big(\|f(X_n, Y_n) - f(X_n, a)\|\big) + E\big(\|f(X_n, a) - f(X, a)\|\big). \tag{36.352}$$

La fonction $g(t) = f(t, a)$ étant continue et bornée, la convergence en loi $X_n \xrightarrow{\mathscr{L}} X$ donne

$$E\big(\|f(X_n, a) - f(X, a)\|\big) \to 0. \tag{36.353}$$

Étudions à présent le premier terme du membre de droite de (36.352). Pour tout $\eta > 0$ et toute variables aléatoires Z et Z' nous avons

$$E(Z) = E(Z\mathbb{1}_{|Z'|<\eta}) + E(Z\mathbb{1}_{|Z'|\geqslant\eta}). \tag{36.354}$$

Nous décomposons donc le premier terme de (36.352) en

EqEXnADecomsecvolta
$$\begin{aligned} E\big(\|f(X_n, Y_n) - f(X_n, a)\|\big) = {} & E\big(\|f(X_n, Y_n) - f(X_n, a)\|\mathbb{1}_{|Y_n-a|<\eta}\big) \\ & + E\big(\|f(X_n, Y_n) - f(X_n, a)\|\mathbb{1}_{|Y_n-a|\geqslant\eta}\big). \end{aligned} \tag{36.355}$$

Choisissons maintenant une valeur de η telle que

$$|(x, y) - (x', y')| < \eta \Rightarrow |f(x, y) - f(x', y')| \leqslant \epsilon. \tag{36.356}$$

Un tel η existe par l'uniforme continuité de f. Dans le premier terme, $|Y_n - a| < \eta$, par conséquent

$$\|(X_n, Y_n) - (X_n, a)\| = |Y_n - a| < \eta \tag{36.357}$$

et donc

$$\|f(X_n, Y_n) - f(X_n, a)\| \leqslant \epsilon. \tag{36.358}$$

Le premier terme devient donc

$$E\big(\|f(X_n, Y_n) - f(X_n, a)\|\mathbb{1}_{|Y_n-a|<\eta}\big) \leqslant \epsilon E(\mathbb{1}_{|Y_n-a|<\eta}) \leqslant \epsilon \tag{36.359}$$

parce que $E(\mathbb{1}_A) = P(A) \leqslant 1$. Pour le second terme de (36.355) nous effectuons la majoration

$$\|f(X_n, Y_n) - f(X_n, a)\| \leqslant 2\|f\|_\infty \tag{36.360}$$

tandis que la convergence $Y_n \xrightarrow{P} a$ entraine

$$P\big(|Y_n - a| \geqslant \eta\big) \to 0. \tag{36.361}$$

□

PropcvLsousfonc

Proposition 36.120 ([782]).
Soient X_i des variables aléatoires telles que

$$X_i \xrightarrow{\mathscr{L}} X \tag{36.362}$$

et h, une fonction mesurable sur l'espace d'arrivée de X_i. Soit C l'ensemble des points de continuité de h au sens

$$C = \{\omega \in \Omega \text{ tel que } h \text{ est continue en } X_i(\omega)\}. \tag{36.363}$$

Alors si $P(X \in C) = 1$, nous avons

$$h(X_i) \xrightarrow{\mathscr{L}} h(X). \tag{36.364}$$

Une conséquence de cette proposition couplée au lemme de Slutsky est le résultat suivant, qui est donné sous le nom de théorème de Slutsky sur wikipédia.

CorINgTPH

Corolaire 36.121.
En reprenant les notations du lemme de Slutsky, si

$$X_n \xrightarrow{\mathscr{L}} X \tag{36.365a}$$

$$Y_n \xrightarrow{P} a, \tag{36.365b}$$

alors

$$X_n + Y_n \xrightarrow{\mathscr{L}} X + a \tag{36.366a}$$

$$X_n Y_n \xrightarrow{\mathscr{L}} aX \tag{36.366b}$$

$$Y_n^{-1} X_n \xrightarrow{\mathscr{L}} a^{-1} X \tag{36.366c}$$

(36.366d)

pourvu que a soit inversible.

LEMooXWMXooYEwWBn

Lemme 36.122 (Borel-Cantelli).
Soit (A_n) une suite d'événements (avec $A_n \in \mathcal{A}$ pour tout n).

(1) Si $\sum_{n=0}^{\infty} P(A_n)$ converge, alors

$$P(A_n \text{ i.s.}) = 0. \tag{36.367}$$

(2) Si la somme $\sum_n P(A_n)$ diverge, et si de plus les A_i sont indépendants, alors

$$P(A_n \text{ i.s.}) = 1. \tag{36.368}$$

La notation $P(A_n \text{ i.s.})$ signifie « infiniment souvent », c'est-à-dire

$$P(A_n \text{ i.s.}) = P\Big(\bigcap_{N\in\mathbb{N}} \bigcup_{k\geqslant N} A_k\Big) = P(\limsup A_n) \tag{36.369}$$

Une façon de paraphraser le lemme de Borel-Cantelli est que nous avons l'alternative

EqparaphrCantelli

$$P(\limsup A_n) = \begin{cases} 0 & \text{si } \sum_{n\geqslant 0} P(A_n) < \infty \\ 1 & \text{sinon.} \end{cases} \tag{36.370}$$

Proposition 36.123.
Soit X_n, une suite de variables aléatoires et X une variable aléatoire. Si

$$\sum_n P\big(\|X_n - X\| > \epsilon\big) < \infty \tag{36.371}$$

pour tout ϵ, alors $X_n \xrightarrow{p.s.} X$.

Démonstration. Fixons ϵ et considérons les événements $A_n = \|X_n - X\| > \epsilon$. L'hypothèse dit que

$$\sum_n P(A_{n,\epsilon}) < \infty \tag{36.372}$$

et le lemme de Borel-Cantelli implique que

$$P(\limsup \|X_n - X\| > \epsilon) = 0. \tag{36.373}$$

Un élément ω est dans $\limsup A_n$ si il est contenu dans tous les A_n, par conséquent, pour chaque ϵ nous avons l'inclusion

$$\{\omega \in \Omega \text{ tel que } X_n(\omega) \to X(\omega)\} \subset \complement \limsup A_n. \tag{36.374}$$

Nous pouvons aller plus loin et écrire

$$\{\omega \in \Omega \text{ tel que } X_n(\omega) \to X(\omega)\} = \complement\{\omega \in \Omega \text{ tel que } \|X_n - X\| > \epsilon, \forall \epsilon > 0\}. \tag{36.375}$$ EqProbomOmXnmXforalleps

Or la probabilité de l'ensemble

$$\{\omega \in \Omega \text{ tel que } \|X_n - X\| > \epsilon\} \tag{36.376}$$

est 0 pour chaque ϵ, et par conséquent la probabilité du membre de droite de (36.375) est 1. □

Exemple 36.124.
Considérons une suite de 0 et de 1 dans laquelle le 1 arrive avec une probabilité p et le 0 avec une probabilité $1 - p$. Une telle suite est modélisée par une suite de variables aléatoires de Bernoulli $(X_n)_{n\in\mathbb{N}}$ indépendantes de paramètre p.

Question : une telle suite contient elle une infinité de 1 ? Considérons les événements indépendants $A_n = \{X_n = 1\}$. Nous avons

$$\sum_n P(A_n) = \sum_n P(X_n = 1) = \sum_n p = \infty. \tag{36.377}$$

Par Borel-Cantelli et son expression (36.370), nous avons alors

$$P(\limsup A_n) = 1. \tag{36.378}$$

Donc une infinité d'événements A_n se produisent, et nous avons bien une infinité de 1 dans la suite.

Remarque : dans ce raisonnement nous pouvons considérer une probabilité non constante p_n tant que la série $\sum_n p_n$ diverge. △

EXooSRQUooWEpXRi

Exemple 36.125.
À propos de maximum. La fonction $h\colon \mathbb{R}^d \to \mathbb{R}$ donnée par $h(x_1, \dots, x_d) = \max_i\{x_i\}$ est une fonction continue. Nous voudrions prouver que si on a une famille (finie en $i = 1, \dots, l$) de suites (en n) de variables aléatoires $X_n^{(i)} \xrightarrow{p.s.} a$ convergeant toutes vers la même limite a, alors

$$M_n = \max_i\{X_n^{(i)}\} \xrightarrow{p.s.} a. \tag{36.379}$$

D'abord si nous avons l suites numériques $(x_n^{(i)})$, alors la suite

$$M_n = \max_i x_n^{(i)} \tag{36.380}$$

converge vers la même limite. En effet si $\epsilon > 0$ est donné, il suffit de prendre N_i l'entier tel que $|x_n^{(i)} - a| \leqslant \epsilon$ pour tout $n > N_i$. Et ensuite on prend $N > \max\{N_i\}$.

Si maintenant au lieu de suites numériques, nous avons des variables aléatoires, le résultat reste valable. Nous cherchons à prouver que

$$P\Big(\{\omega \in \Omega \text{ tel que } \max\{X_n^{(i)}(\omega)\} \to a\}\Big) = 1. \tag{36.381}$$ EqMZcnvTh

Par ce que nous venons de dire sur les suites numériques, un élément ω n'est pas dans cet ensemble seulement si il y a un i pour lequel $X_n^{(i)}$ ne converge pas vers a. Or cela, pour chaque i est un événement de probabilité zéro.

Les ω qui ne fonctionneront pas dans l'équation (36.381) sont ceux de la réunion d'un ensemble fini d'ensembles de probabilité nulle. C'est donc de probabilité nulle. △

36.4 Loi des grands nombres, théorème central limite

36.4.1 Loi des grands nombres

Lemme 36.126 (Inégalité de Markov[783])**.**
Soit une variable aléatoire $X \in L^p$ et $\epsilon > 0$. Nous avons

$$P(|X| \geqslant \epsilon) \leqslant \frac{1}{\epsilon^r} E(|X|^r). \tag{36.382}$$

Démonstration. Nous avons

$$E(|X|^r) \geqslant \int_{|X|\geqslant\epsilon} X(\omega) dP(\omega) \geqslant \epsilon^r \int_{|X|\geqslant\epsilon} dP = \epsilon^r P(|X| \geqslant \epsilon). \tag{36.383}$$

□
CorEWhIsBB

Corolaire 36.127 ([783])**.**
Soit ϕ, une fonction croissante et positive ou nulle sur l'intervalle I. Soit aussi une variable aléatoire $Y : \Omega \to \mathbb{R}$ telle que $P(Y \in I) = 1$. Alors pour tout $b \in I$ tel que $\phi(b) > 0$ nous avons

$$P(Y \geqslant b) \leqslant \frac{E\big[\phi(Y)\big]}{\phi(b)}. \tag{36.384}$$

ThoefQyKZ

Théorème 36.128 (Loi forte des grands nombres)**.**
Soit (X_n) une suite de variables aléatoires réelles

(1) indépendantes et identiquement distribuées,

(2) intégrables (c'est-à-dire dans L^1),

alors

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} E(X_1). \tag{36.385}$$

Note : étant donné que les variables aléatoires sont identiquement distribuées, nous avons évidemment $E(X_1) = E(X_2) = \dots$

¡¡ Avertissement/question à la lectrice !! 36.129
Est-ce que les variables aléatoires doivent vraiment être réelles ?

CORooAYTPooMdAdoN

Corolaire 36.130.
Si les variables aléatoires réelles X_n sont

(1) indépendantes et identiquement distribuées,

(2) dans L^2

alors

$$\bar{X}_n \xrightarrow{P} E(X_1). \tag{36.386}$$

Démonstration. Nous voulons prouver que pour tout $\eta > 0$,

$$P\big(|\bar{X}_n - E(X_1)| > \eta\big) \to 0. \tag{36.387}$$

Remarquons d'abord que les variables aléatoires X_n étant identiquement distribuées, $E(\bar{X}_n) = E(X_1)$ parce que $E(X_i) = E(X_1)$ pour tout i. L'inégalité de Markov avec $r = 2$ nous donne

$$P\big(|\bar{X}_n - E(\bar{X}_n)| > \eta\big) \leqslant \frac{1}{\eta^2} E\big(|\bar{X}_n - E(\bar{X}_n)|^2\big) \tag{36.388a}$$

$$= \frac{1}{\eta^2} \operatorname{Var}(\bar{X}_n) \tag{36.388b}$$

$$= \frac{1}{n\eta^2} \operatorname{Var}(X_1) \tag{36.388c}$$

où nous avons utilisé la la proposition 36.53 : $\text{Var}(\bar{X}_n) = \text{Var}(X_1)/n$. Au final nous avons prouvé que

$$P\big(|\bar{X}_n - E(\bar{X}_n)| > \eta\big) \leqslant \frac{1}{n\eta^2}\text{Var}(X_1), \tag{36.389}$$

qui tend vers zéro lorsque $n \to \infty$. □

Proposition 36.131.
Soient X_n des variables aléatoires indépendantes et identiquement distribuées avec $X_n \geqslant 0$. Nous acceptons $E(X_1) = \infty$, c'est-à-dire que nous relaxons la condition $X_n \in L^1$ par rapport à la loi des grands nombres.

Alors

$$\bar{X}_n \xrightarrow{p.s.} E(X_1) \in [0, \infty]. \tag{36.390}$$

Démonstration. Si $E(X_1) < \infty$, nous sommes dans le cas de la loi des grands nombres. Pour chaque $N \in \mathbb{N}$ nous considérons la suite de variables aléatoires

$$X_n^{(N)} = \min(X_n, N). \tag{36.391}$$

Nous avons évidement $\bar{X}_n^{(N)} \leqslant \bar{X}_n$. Les variables aléatoires $X_n^{(N)}$ étant bornées par N, elles vérifient la loi des grands nombres pour chaque N séparément. Par conséquent nous avons pour chaque N la limite

$$\bar{X}_n^{(N)} \to E(X_1^{(N)}) \tag{36.392}$$

EqbarXNbtoXnubus

Nous supposons que $E(X_1) = \infty$, par conséquent $\lim_{N\to\infty} E(X_1^{(N)}) = \infty$. Soit $\eta > 0$ et choisissons N de telle manière à avoir

$$E(X_1^{(N)}) > \eta + 1. \tag{36.393}$$

La limite (36.392) nous permet de trouver n_0 tel que pour tout $n > n_0$ nous ayons $\bar{X}_n^{(N)} > \eta$. Au final,

$$\eta < \bar{X}_n^{(N)} \leqslant \bar{X}_n, \tag{36.394}$$

ce qui montre que $\bar{X}_n \to \infty$. □

Exemple 36.132.
La loi des grands nombres justifie la pratique courante d'approximer une grandeur physique par la moyenne empirique d'un grand nombre de mesures. △

Exemple 36.133.
Citons ici le dernier paragraphe de *Le mystère de Marie Roget* par Edgar Allan Poe, traduit par Charles Baudelaire [35].

> Rien, par exemple, n'est plus difficile que de convaincre le lecteur non spécialiste que, si un joueur de dés a amené les six deux fois coup sur coup, ce fait est une raison suffisante de parier gros que le troisième coup ne ramènera pas les six. Une opinion de ce genre est généralement rejetée tout d'abord par l'intelligence. On ne comprend pas comment les deux coups déjà joués, et qui sont maintenant complètement enfouis dans le Passé, peuvent avoir de l'influence sur le coup qui n'existe que dans le Futur. La chance pour amener les six semble être précisément ce qu'elle était à n'importe quel moment, c'est-à-dire soumise seulement à l'influence de tous les coups divers que peuvent amener les dés. Et c'est là une réflexion qui semble si parfaitement évidente, que tout effort pour la controverser est plus souvent accueilli par un sourire moqueur que par une condescendance attentive.

Dans le cours de la nouvelle, Edgar Poe cite et utilise la théorie des probabilités avec une justesse inaccoutumée dans la littérature. Mais dans ce paragraphe final, Poe montre de façon la plus formelle qu'il n'a *rien* compris à la loi des grands nombres. △

35. Disponible sur https://fr.wikisource.org/wiki/Le_Mystère_de_Marie_Roget

36.4.2 Théorème central limite

Lemexpznznsurnton

Lemme 36.134.
Soit $z_n \to z$ une suite convergente dans $\mathbb{C}$. Alors

$$\left(1 + \frac{z_n}{n}\right)^n \to e^z. \tag{36.395}$$

ThoOWodAi

Théorème 36.135.
Si les variables aléatoires X_n sont

(1) indépendantes et identiquement distribuées de loi parente X,

(2) $X_1 \in L^2(\Omega, \mathcal{A})$,

alors nous notons $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, $m = E(X_1)$ et $\sigma^2 = \mathrm{Var}(X_1)$ et nous avons

$$\frac{\bar{X}_n - E(X)}{\sqrt{\mathrm{Var}(\bar{X}_n)}} = \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \xrightarrow{\mathscr{L}} \mathcal{N}(0,1). \tag{36.396}$$

Démonstration. Nous allons écrire la démonstration dans le cas de variables aléatoires réelles. La proposition 36.113 dit que la suite X_n converge en loi vers X si et seulement si les fonctions caractéristiques convergent ponctuellement. Nous devons donc prouver, pour chaque [36] $t \in \mathbb{R}$, que

EqPhitonbidNtznu

$$\Phi_{\frac{S_n - nm}{\sigma\sqrt{n}}}(t) \to \Phi_{\mathcal{N}(0,1)}(t). \tag{36.397}$$

Supposons dans un premier temps que $E(X_i) = 0$ et $\sigma(X_i) = 1$. Dans ce cas nous considérons la fonction

$$\begin{aligned}
\Phi_{\frac{S_n}{\sqrt{n}}} &= E\big(e^{i\frac{t}{\sqrt{n}}\sum_{k=1}^n X_k}\big) && (36.398a)\\
&= \prod_{k=1}^n E\big(e^{i\frac{t}{\sqrt{n}}X_1}\big) && (36.398b)\\
&= \prod_{k=1}^n \Phi_{X_1}\left(\frac{t}{\sqrt{n}}\right) && (36.398c)\\
&= \Phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n. && (36.398d)
\end{aligned}$$

Cette quantité est à priori complexe ; nous ne pouvons donc pas immédiatement passer au logarithme. Nous pouvons par contre utiliser un développement en puissances de t en nous servant de la proposition 36.91 et de l'hypothèse comme quoi $X_1 \in L^2$:

$$\Phi_{X_1}(t) = \Phi_{X_1}(0) + \Phi'_{X_1}(0)t + \Phi''_{X_1}(0)\frac{t^2}{2} + \alpha(t)t^2 \tag{36.399}$$

où α est une fonction qui a la propriété $\lim_{x\to 0}\alpha(x) = 0$.

En utilisant les hypothèses et la formule de dérivation de la fonction caractéristique,

$$\begin{aligned}
\Phi_{X_1}(0) &= 1 && (36.400a)\\
\Phi'_{X_1}(0) &= E\big((iX)\big) = 0 && (36.400b)\\
\Phi''_{X_1}(0) &= E(-X^2) = -\mathrm{Var}(X_1) = -1. && (36.400c)
\end{aligned}$$

Nous avons donc

$$\Phi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = \underbrace{1 - \frac{1}{2}\frac{t^2}{n}}_{\in\mathbb{R}} + \underbrace{\frac{t^2}{n}\alpha\left(\frac{t}{\sqrt{n}}\right)}_{\in\mathbb{C}}, \tag{36.401}$$

36. Chuck Norris peut *vraiment* le faire pour *chaque* $t \in \mathbb{R}$

de telle sorte que, en considérant une valeur fixée de t,

$$\Phi_{\frac{S_n}{\sqrt{n}}}(t) = \left(1 - \frac{\frac{t^2}{2} + \beta_n}{n}\right)^n \tag{36.402}$$

où $\beta_n = t^2\alpha(t/\sqrt{n})$. Nous avons bien entendu $\lim_{n\to\infty} \beta_n = 0$.

Nous pouvons appliquer le lemme 36.134 pour obtenir la limite

$$\lim_{n\to\infty} \Phi_{\frac{S_n}{\sqrt{n}}}(t) = e^{-t^2/2}. \tag{36.403}$$

La convergence (36.397) est par conséquent prouvée dans le cas où $E(X_i) = 0$ et $\text{Var}(X_i) = 1$.

Considérons maintenant des variables aléatoires avec $E(X_i) = m$ et $\text{Var}(X_i) = \sigma^2$. Elles peuvent être écrites sous la forme

$$X_i = \sigma X_i' + m \tag{36.404}$$

où X_i' est d'espérance nulle et de variance un. Nous avons alors

$$S_n = \sigma \sum_{i=1}^{n} X_i' + nm, \tag{36.405}$$

et

$$\frac{S_n - nm}{\sigma\sqrt{n}} = \frac{S_n'}{\sqrt{n}} \tag{36.406}$$

où $S_n' = \sum_i X_i'$. L'étude de la variable aléatoire

$$\frac{S_n - nm}{\sigma\sqrt{n}} \tag{36.407}$$

revient donc à celle de $S_n'/\sqrt{n}$ qui vient d'être effectuée. □

36.136.
À propos du théorème central limite 36.135. Si pour une certaine variable aléatoire X on a $E(X) = m$, alors nous n'avons pas forcément $P(X = m + a) = P(X = m - a)$. Est-ce que le théorème central limite permet cependant d'affirmer que dans un certaine mesure nous avons

$$P(\bar{X}_n = m + a) = P(\bar{X}_n = m - a) \tag{36.408}$$

lorsque n est grand ?

Tel quelle, l'équation (36.408) est en général fausse pour chaque n parce qu'il existe des distributions non symétriques. Mais bien entendu les deux membres tendent vers zéro pour $n \to \infty$. Mais cela n'est pas lié à la symétrie de la distribution gaussienne. C'est seulement le fait que la gaussienne n'a pas de masses ponctuelles.

Par contre, il y a effectivement une assurance de symétrie pour $\bar{X}_n$ lorsque $n \to \infty$. Le fait est que $X_n \xrightarrow{\mathscr{L}} X$ où X est une gaussienne. La fonction de répartition de X est continue partout et la proposition 36.117 nous dit que pour tout $x \in \mathbb{R}$,

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(t). \tag{36.409}$$

Vu que le nombre $P\big(\bar{X}_n \in B(m + a, \delta)\big)$ peut être exprimé avec des sommes et différences F_{X_n}, nous avons

$$\lim_{n\to\infty} P\big(\bar{X}_n \in B(m + a, \delta)\big) = P\big(X \in B(m + a, \delta)\big). \tag{36.410}$$

Par symétrie de la gaussienne le membre de droite est égal à $P\big(X \in B(m - a, \delta)\big)$ et nous avons bien

$$\lim_{n\to\infty} P\big(\bar{X}_n \in B(m + a, \delta)\big) = \lim_{n\to\infty} P\big(\bar{X}_n \in B(m - a, \delta)\big). \tag{36.411}$$

RemRHFDooGbaPYu

Remarque 36.137.
Le théorème central limite s'applique quelle que soit la distribution des variables aléatoires X_i (dans les limites de hypothèses) ; en particulier il ne dit rien sur la moyenne des X_i. Il dit seulement que l'écart de la moyenne « mesurée » à la moyenne « théorique » est une variable aléatoire gaussienne si on a mesuré assez de fois.

Autrement dit, si la durée d'attente à la poste est de 5 minutes, et si j'y vais 2000 fois, alors la probabilité que ma moyenne d'attente soit de 4 minutes est la même que la probabilité qu'elle soit de 6 minutes [37].

¡¡ Avertissement/question au lecteur !! 36.138
La remarque 36.137 est une interprétation personnelle. J'aimerais avoir l'avis de quelqu'un de plus compétent.

Remarque 36.139.
Nous pouvons obtenir la limite (36.403) d'une façon alternative. Nous considérons la détermination du logarithme complexe sur $\mathbb{C}\backslash\mathbb{R}^-$; cela est une fonction analytique (théorème 26.34) vérifiant l'équation

$$e^{\ln(z)} = z \tag{36.412}$$

pour tout $z \in \mathbb{C}\backslash\mathbb{R}^-$ et le développement

$$\ln(1+z) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{z^n}{n}. \tag{36.413}$$

En particulier, $\ln(1+z) = z + z\alpha(z)$ où $\lim_{z\to 0} \alpha(z) = 0$. Nous reprenons à l'équation (36.402) en fixant t. Nous avons

$$\Phi_{\frac{S_n}{\sqrt{n}}}(t) = \exp\left[\ln\left(\Phi_{\frac{S_n}{\sqrt{n}}}(t)\right)\right] \tag{36.414a}$$

$$= \exp\left[n \ln\left(1 + \frac{-\frac{t^2}{2} - \beta_n}{n}\right)\right] \tag{36.414b}$$

$$= \exp\left[-\frac{t^2}{2} - \beta_n + \left(-\frac{t^2}{2} - \beta_n\right)\alpha\left(\frac{-\frac{t^2}{2} - \beta_n}{n}\right)\right]. \tag{36.414c}$$

À la limite $n \to 0$ nous tombons sur $e^{-t^2/2}$.

Remarque 36.140.
Étant donné que la variable aléatoire

$$\frac{S_n - nm}{\sigma\sqrt{n}} \tag{36.415}$$

converge en loi vers $\mathcal{N}(0,1)$, nous avons la convergence des fonctions de répartition partout où la fonction de répartition de la normale est continue [38] (donc sur tout $\mathbb{R}$). En particulier,

$$\left| P\left(\frac{S_n - nm}{\sigma\sqrt{n}} \leqslant x\right) - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \right| \to 0. \tag{36.416}$$

Nous avons la borne de **Berry-Esséen** qui donne une estimation de la vitesse de convergence : si $X \in L^3$, alors il existe une constante C, indépendante de x, des X_i et de n telle que

$$\left| P\left(\frac{S_n - nm}{\sigma\sqrt{n}} \leqslant x\right) - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \right| \leqslant \frac{C\mu_3}{\sigma^3\sqrt{n}} \tag{36.417}$$

où $\mu_3 = E\big(|X_1 - m|^3\big)$ est le moment d'ordre 3 de X. La chose à retenir est que la convergence est à la vitesse de $1/\sqrt{n}$.

37. Et en l'occurrence, cette probabilité est nulle parce qu'on est en train de parler de variable aléatoire continue, mais vous voyez l'idée.
38. Proposition 36.117.

En dimension $d > 1$, nous avons encore un théorème central limite.

THOooADSIooFrWawC

Théorème 36.141.
Si $d > 1$, et si nous avons des variables aléatoires X_n à valeurs dans $\mathbb{R}^d$ avec

(1) les X_n sont indépendantes et identiquement distribuées

(2) les X_n sont dans L^2.

Alors nous notons $X_1 = (X_1^{(1)}, \dots, X_1^{(d)})$. Nous avons

$$\frac{S_n - nm}{\sqrt{n}} \xrightarrow{\mathscr{L}} \mathcal{N}(0, \Sigma) \tag{36.418}$$

où Σ est ma matrice de covariance du vecteur aléatoire X_1 :

$$\Sigma = \big(\mathrm{Cov}(X_1^{(1)}, \dots, X_1^{(d)}) \big)_{i,j=1,\dots,d}. \tag{36.419}$$

36.4.3 Marche aléatoire

Nous considérons un mobile qui se déplace sur l'axe $\mathbb{Z}$. À chaque pas de temps, nous supposons qu'il va faire un pas à gauche avec une probabilité p et un pas à droite avec une probabilité $(1-p)$. Nous nous demandons quel est le mouvement du mobile sur le long terme.

La position S_n du mobile à l'instant n est donnée par

$$S_n = \sum_{i=1}^{n} X_i \tag{36.420}$$

où X_i est le pas effectué à l'instant i. Ce sont des variables de Bernoulli indépendantes avec

$$X_i \overset{\mathscr{L}}{=} p\delta_{-1} + (1-p)\delta_1 \tag{36.421}$$

c'est-à-dire

$$P(X_i = -1) = p \tag{36.422a}$$
$$P(X_i = 1) = 1 - p. \tag{36.422b}$$

Ces variables vérifient les hypothèses de la loi des grands nombres :

(1) elles sont indépendantes et identiquement distribuées,

(2) elles sont intégrables.

Pour le second point, le calcul est

$$\int_\Omega |X_i| dP = \int_{\mathbb{R}} |x| dP_X = \int_{\mathbb{R}} |x|(p\delta_{-1} + (1-p)\delta_1) = |1-p| + |p| = 1. \tag{36.423}$$

Nous avons par conséquent

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} E(X_1) = (1-2p) \tag{36.424}$$

et

$$\frac{S_n}{n} \to (1-2p). \tag{36.425}$$

Si $p \neq 1/2$ nous pouvons conclure que

(1) si $p > 1/2$, alors $S_n \xrightarrow{p.s.} -\infty$

(2) si $p < 1/2$, alors $S_n \xrightarrow{p.s.} \infty$.

De plus nous connaissons la vitesse de divergence : elle est linéaire. Le mobile suit essentiellement l'équation $p(n) = (1-2p)n$.

Remarque 36.142.
Cela ne traite pas le cas $p = 1/2$. Dans ce cas, nous pouvons simplement dire que $S_n = o(n)$.

36.5 Les lois usuelles

36.5.1 Loi de Bernoulli

Une expérience de Bernoulli consiste à tirer au hasard un 0 ou un 1 avec une probabilité p de tomber sur 1 et $1-p$ de tomber sur zéro. Il s'agit donc d'une expérience qui réussit ou qui rate.

Le cas typique est une urne avec des boules indiscernables blanches ou noires. La probabilité p est la proportion de blanches dans l'urne (avec remise entre les tirages). Dans ce cas, nous avons l'espace de probabilité $(\Omega, \mathcal{A}, P)$ où Ω représente l'ensemble des boules, $\mathcal{A}$ est l'ensemble des parties de Ω et P est l'équiprobabilité sur Ω. Une variable aléatoire est une application

$$\begin{aligned} X\colon \Omega &\to \{0,1\} \\ \omega &\mapsto \text{couleur de la boule } \omega. \end{aligned} \tag{36.426}$$

Nous notons $\mathscr{B}(1,p)$ la loi de Bernoulli. Elle a une expression très simple :

$$\mathscr{B}(0,1)\big(\{1\}\big) = p \tag{36.427a}$$
$$\mathscr{B}(0,1)\big(\{0\}\big) = 1-p \tag{36.427b}$$

Une variable aléatoire réelle est de **Bernoulli** de paramètre p $(0 < p < 1)$ si

$$X\colon \Omega \to \mathbb{R} \tag{36.428}$$

avec $P(x=1) = p$ et $P(X=0) = 1-p$. En tant que mesure sur $\mathbb{R}$, nous avons

$$P_X = p\delta_1 + (1-p)\delta_0. \tag{36.429}$$

Une fonction h qui réalise le supremum de la formule (14.468) est par exemple une fonction en escalier qui vaut en x le plus petit entier plus grand ou égal à x. L'espérance d'une loi de Bernoulli est alors

$$E(x) = p. \tag{36.430}$$

Étant donné que la variable aléatoire X prend seulement les valeurs 0 et 1, nous avons pour tout ensemble mesurable B

$$P_{X^2}(B) = P(X^2 \in B) = P(X \in B), \tag{36.431}$$

et par conséquent $P_{X^2} = P_X$ et $E(X^2) = E(X)$. Nous trouvons donc la variance

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p). \tag{36.432}$$ EqVarBern

36.5.2 Loi binomiale

Une expérience binomiale consiste à répéter n expériences de Bernoulli de paramètre p et de compter le nombre de réussites. Une telle expérience peut être réalisée selon la procédure suivante.

Soit une urne contenant N boules dont une proportion p de 1 et $1-p$ de 0. Une expérience binomiale de paramètres n et p consistera à prendre n boules *avec remise* et à compter le nombre de 1 obtenus.

En termes d'espaces probabilisé, nous avons Ω qui est l'ensemble des tuples de taille n à valeurs dans $\{0,1\}$, la tribu $\mathcal{A}$ est l'ensemble des parties de Ω, et la probabilité P est l'équiprobabilité :

$$P(\omega) = \frac{1}{N^n} \tag{36.433}$$

si il y a N boules dans l'urne. Nous construisons alors la variable aléatoire

$$X(\omega) = \sum_{i=1}^{n} \omega_i \tag{36.434}$$

où ω est une suite de taille n de 0 et de 1.

Calculons $P(X = k)$. Il s'agit de considérer tous les sous-ensembles de taille n de Ω contenant exactement k fois 1. Il y a $\binom{n}{k}$ manière de décider lesquelles des n boules seront blanches. Ensuite, chaque boule blanche peut être choisie parmi les m boules disponibles, et chaque boule noire peut être choisie parmi les $(N - m)$ disponibles. Nous avons donc

$$P(X = k) = \binom{n}{k} \frac{m^k (N - m)^{n-k}}{N^k}. \tag{36.435}$$

En effet la mesure de probabilité sur Ω est la mesure de comptage renormalisée par le cardinal de Ω qui vaut N^m. Étant donné que $p = m/N$, nous transformons facilement (36.435) en

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}. \tag{36.436}$$

Une variable aléatoire de loi binomiale étant une somme de variables aléatoires de Bernoulli indépendantes, l'espérance [39] et la variance [40] s'obtiennent en sommant les espérances et variances termes à terme :

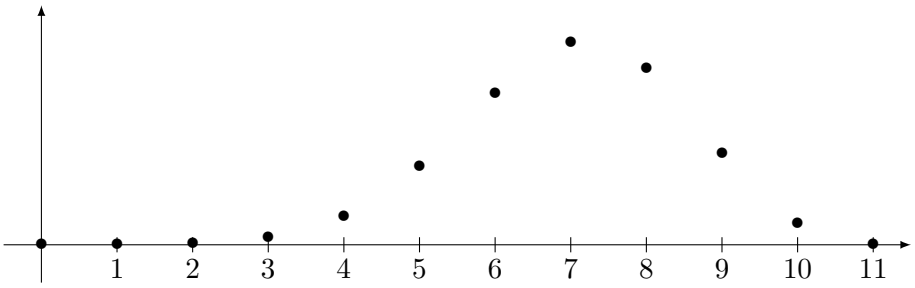
$$E(X) = np \tag{36.437}$$

et

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = np(1 - p) \tag{36.438}$$

en vertu de (36.432).

La loi binomiale lorsque $p = 0.7$ et $n = 10$.



36.5.3 Loi multinomiale

La loi multinomiale $\mathscr{M}(n; k; p_1, \dots, p_k)$ consiste à effectuer n épreuves d'une démarche aléatoire qui peut avoir k issues différentes avec probabilités $p_1, \dots, p_k$. Les variables aléatoires multinomiales sont N_i avec les contraintes

$$\sum_{i=1}^{k} N_i = n \tag{36.439a}$$

$$\sum_{i=1}^{k} p_i = 1. \tag{36.439b}$$

La fonction de probabilité multinomiale est

$$P(N_1 = n_1, \dots, N_k = n_k) = \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k}. \tag{36.440}$$

Chacune des N_i est une binomiale de probabilité p_i.

39. Valable même sans indépendance, proposition 36.47
40. Lemme 36.44.

36.5.4 Loi géométrique

Soit (X_n) une suite indépendante et identiquement distribuée de lois de Bernoulli de paramètre p. Alors la variable aléatoire

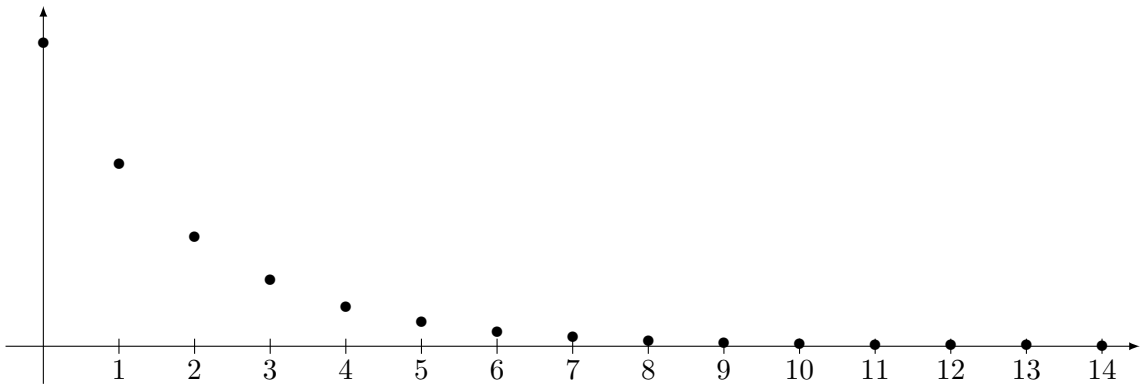
$$Z = \inf\{n \geqslant 1 \text{ tel que } X_n = 1\} \tag{36.441}$$

est une loi géométrique de paramètre p.

La loi géométrique compte donc le nombre d'expériences de Bernoulli à effectuer avant que le premier succès soit au rendez-vous. Nous avons

$$P(Z = k) = P(X_k = 1)P(X_1, \dots, X_{k-1} = 0) = p(1-p)^{k-1} \tag{36.442}$$

La loi géométrique de paramètre $p = 0.2$.



Note : si p est trop grand, on ne voit vite plus rien parce que la probabilité d'attendre longtemps est vite très faible.

36.5.5 Loi de Poisson

DEFooUWKHooUFKlcr

Définition 36.143 (Loi de Poisson)**.**
Une variable aléatoire Z suit une ***loi de Poisson*** *de paramètre λ, notée $\mathscr{P}(\lambda)$ si*

$$P(Z = k) = e^{-\lambda}\frac{\lambda^k}{k!} \tag{36.443}$$

EQooLGLUooLXohWe

pour tout $k \in \mathbb{N}$.

La loi de Poisson est une loi de probabilité discrète qui décrit le comportement du nombre d'évènements se produisant dans un laps de temps fixé, si ces évènements se produisent avec une fréquence moyenne connue et indépendamment du temps écoulé depuis l'évènement précédent.

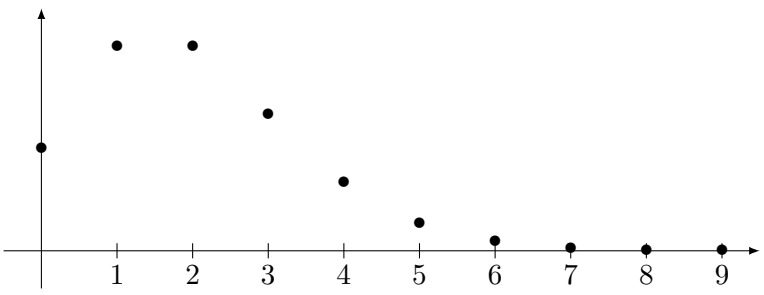
Si un événement se produit en moyenne p fois par seconde, la probabilité d'observer l'événement k fois durant n secondes est donnée par $P(Z = k)$ où Z est une loi de Poisson de paramètre $\lambda = pn$.

ThojDZjuj

Théorème 36.144 (wikipedia)**.**
L'espérance et la variance d'une variable aléatoire de Poisson sont λ :

$$E(X) = \lambda \tag{36.444a}$$
$$\text{Var}(X) = \lambda. \tag{36.444b}$$

La loi de Poisson de paramètre $\lambda = 2$.



36.5.6 Loi exponentielle

La loi exponentielle représente le temps qu'il faut attendre pour qu'une particule se désintègre si elle a en permanence une probabilité[41] λdt de se désintégrer entre t et $t+dt$. L'espérance est donc $1/\lambda$.

Il se passe donc en moyenne λ événements par seconde. La proposition 36.155 nous montrera que le nombre d'événements se produisant en une seconde suit une loi de Poisson de paramètre λ.

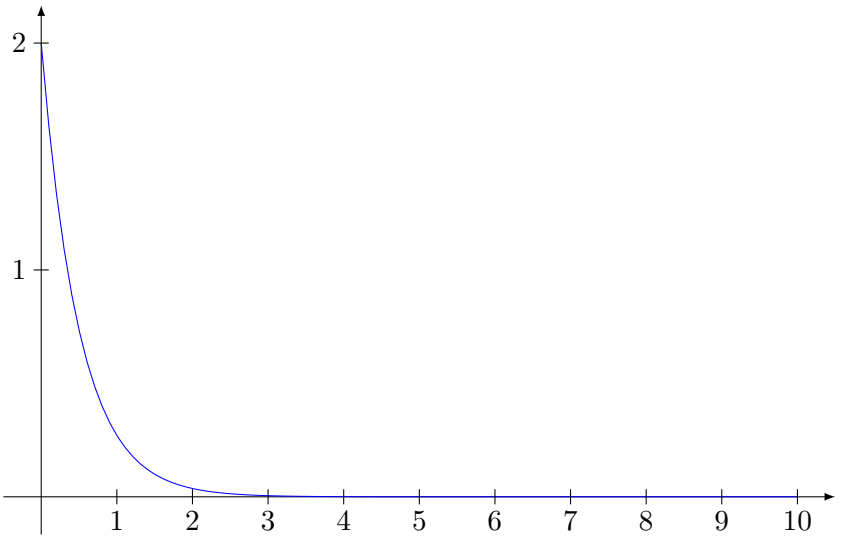
DEFooTSFNooULWNHY

Définition 36.145 (Loi exponentielle)**.**
*La **loi exponentielle** de paramètre $\lambda > 0$, notée $\mathscr{E}(\lambda)$ est la mesure de probabilité sur $\big(\mathbb{R}, \mathcal{B}or(\mathbb{R})\big)$ ayant pour densité*[42] *l'application*

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto \begin{cases} \lambda e^{-\lambda x} & si\ x \geqslant 0 \\ 0 & sinon. \end{cases} \end{aligned} \tag{36.445}$$

Voir la définition 36.8 pour ce qu'est la loi d'une variable aléatoire.

Densité de la loi exponentielle de paramètre $\lambda = 2$.



PROPooKLJIooUFfGwa

Proposition 36.146 ([1])**.**
Soit une variable aléatoire $X\colon \Omega \to \mathbb{R}$. Si pour tout $t > 0$ nous avons

$$P(X > t) = e^{-\lambda t}, \tag{36.446}$$

alors X suit une loi exponentielle.

PROPooXRMWooYjPMWN

Proposition 36.147 ([1])**.**
Soient des variables aléatoires exponentielles X et Y sur Ω. Si pour tout t et u nous avons

$$P(X > t, Y > u) = P(X > t)P(Y > u), \tag{36.447}$$

alors X et Y sont indépendantes.

PROPooBZREooUTvonZ

41. Plutôt une densité qu'une probabilité. Vu que λ est fini et que dt est infinitésimal, λdt est infinitésimal, et si on veut c'est un bon ordre de grandeur pour la probabilité qu'un événement arrive entre t et $t+dt$.
42. Densité, définition 36.5.

Proposition 36.148 (Somme de variables aléatoires exponentielles indépendantes[784]).
Si les variables aléatoires $X_1, \dots, X_n : \Omega \to \mathbb{R}$ sont indépendantes et suivent une même loi exponentielle de paramètre λ, alors la densité de la somme $\sum_{k=1}^{n} X_k$ est

EQooSCJHooUWxrgs
$$\begin{aligned} h_n : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \begin{cases} 0 & si\ t < 0 \\ \frac{(\lambda t)^{n-1}\lambda e^{-\lambda t}}{(n-1)!} & si\ t \geqslant 0. \end{cases} \end{aligned} \tag{36.448}$$

Démonstration. Nous y allons par récurrence sur n en commençant avec $n = 1$. Pour $n = 1$, il suffit de remarquer que la formule (36.448) se réduit à la densité d'une exponentielle :

$$\begin{aligned} h_1 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0. \end{cases} \end{aligned} \tag{36.449}$$

La proposition 36.42 dit que la densité de la somme sera le produit de convolution, et le lemme 27.67 nous assure que c'est commutatif :

$$h_{n+1}(t) = (h_n * h_1)(t) = (h_1 * h_n)(t). \tag{36.450}$$

Pour $t \geqslant 0$, nous avons le calcul suivant :

$$\begin{aligned} h_{n+1}(t) &= \int_{\mathbb{R}} h_1(t-x)h_n(x)dx && (36.451a) \\ &= \int_0^t \lambda e^{-\lambda(t-x)} \frac{(\lambda x)^{n-1}}{(n-1)!} \lambda e^{-\lambda x} && (36.451b) \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda t} \int_0^t x^{n-1} dx && (36.451c) \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda t} \frac{t^n}{n}, && (36.451d) \end{aligned}$$

qui est bien la valeur attendue dans (36.448) pour $t \geqslant 0$.

Si $t < 0$, alors nous avons

$$h_{n+1}(t) = \int_{\mathbb{R}} h_1(t-x)h_n(x)dx. \tag{36.452}$$

Nous avons $h_1(t-x) \neq 0$ seulement sur $t \geqslant x$. Et nous avons $h_n(x)$ seulement si $x \geqslant 0$. Bref, l'intégrale se fait sur un ensemble vide, et $h_{n+1}(t) = 0$, comme il fallait. □

PropTxGcWn
Proposition 36.149.
Soit $X \sim \mathscr{E}(\lambda)$. La fonction caractéristique est donnée par

EQooMDJFooKhZkJo
$$E(e^{itX}) = \frac{\lambda}{\lambda - it}. \tag{36.453}$$

L'espérance et la variance sont données par

$$\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \text{Var}(X) &= \frac{1}{\lambda^2}. \end{aligned} \tag{36.454}$$

Démonstration. Pour la fonction caractéristique,

$$E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx}\lambda e^{-\lambda x}\mathbb{1}_{[0,\infty[}(x)dx \tag{36.455a}$$
$$= \lambda\int_0^{\infty} e^{x(-\lambda+it)}dx \tag{36.455b}$$
$$= \lim_{A\to\infty}\left[\frac{e^{x(-\lambda+it)}}{it-\lambda}\right]_{x=0}^{x=A} \tag{36.455c}$$
$$= \lambda\lim_{A\to\infty}\frac{e^{A(it-\lambda)}}{it-\lambda} - \frac{\lambda}{it-\lambda}. \tag{36.455d}$$

Le premier terme est nul parce que si on prend la norme,

$$\left|\frac{e^{A(-\lambda+it)}}{-\lambda+it}\right| = \frac{e^{-\lambda A}}{|it-\lambda|} \to 0. \tag{36.456}$$

Cela prouve la formule (36.453).

En ce qui concerne l'espérance nous faisons le calcul suivant :

$$E(X) = \int_{\mathbb{R}} xf_X(x)dx = \lambda\int_{\mathbb{R}^+} xe^{-\lambda x}dx = \frac{1}{\lambda}. \tag{36.457}$$

Pour la variance, nous utilisons la formule (36.108). Nous avons

$$E(X^2) = \int_{\mathbb{R}} x^2 f_X(x)x = \int_0^{\infty} x^2\lambda e^{-\lambda x}dx = \frac{2}{\lambda^2}. \tag{36.458}$$

Donc $\text{Var}(X) = E(X^2) - E(X)^2 = \frac{1}{\lambda^2}$. □

La loi exponentielle est une loi **sans mémoire** en ce sens que

$$P(X > x+y|X > y) = P(X > x). \tag{36.459}$$

En effet nous utilisons la règle de la probabilité conditionnelle

$$P(A|B) = \frac{P(A\cap B)}{P(B)}. \tag{36.460}$$

Ici,

$$P(X > x+y|X > y) = \frac{P(X > x+y)}{P(X > y)} = e^{-\lambda x}. \tag{36.461}$$

La proposition suivante montre que la loi exponentielle est à peu près la seule à être sans mémoire. D'où son importance dans l'étude des machines dont les pièces ne subissent pas d'usure.

PropREXaIBg

Proposition 36.150.
Soit X, une variable aléatoire admettant une densité continue f par rapport à la mesure de Lebesgue. Si elle est sans mémoire, alors elle est exponentielle.

Démonstration. Nous posons $\varphi(x) = P(X \geqslant x)$. Cela est la fonction de répartition de X (à part que cette dernière est $1-\varphi(x)$ mais c'est pas grave), et est donnée par

$$\varphi(x) = 1 - \int_0^x f(t)dt. \tag{36.462}$$

Cette dernière intégrale vérifie les hypothèses du théorème 36.150, de telle sorte que φ soit une fonction dérivable et $\varphi'(x) = f(x)$.

D'autre part en utilisant la définition de la probabilité conditionnelle la propriété de ne pas avoir de mémoire donne

$$\varphi(x+y)=\varphi(x)\varphi(y) \tag{36.463}$$

et de plus $\varphi(0)=0$. Calculons la dérivée de φ :

$$\varphi'(x)=\lim_{\epsilon\to 0}\frac{\varphi(x+\epsilon)-\varphi(x)}{\epsilon}=\varphi(x)\lim_{\epsilon\to 0}\frac{\varphi(\epsilon)-1}{\epsilon}=\varphi(x)\varphi'(0). \tag{36.464}$$

Donc φ vérifie l'équation différentielle de l'exponentielle.

□

Exemple 36.151.
Une machine a une durée de vie représentée par une variable aléatoire suivant une loi de Poisson de paramètre λ. Soit T_y la variable aléatoire qui représente le temps de vie restant sachant que la machine a déjà vécu un temps y. Nous voulons trouver la fonction de répartition de T_y. Nous avons

$$P(T_y>x)=P(X>x+y|X>y)=P(X>x)=e^{-\lambda x}. \tag{36.465}$$

Dans ce cas, la loi de T_y ne dépend pas de y. Cela signifie que la machine ne vieillit pas et surtout que le modèle n'est pas réaliste. △

Proposition 36.152.
Si $X\sim\mathscr{E}(\lambda)$ et $Y\sim\mathscr{E}(\mu)$ sont indépendantes, alors

(1) $P(X<Y)=\frac{\lambda}{\lambda+\mu}$
(2) $P(X>Y)=\frac{\mu}{\lambda+\mu}$
(3) $P(X=Y)=0$
(4) $\min(X,Y)\sim\mathscr{E}(\lambda+\mu)$.

De plus les variables aléatoires exponentielles ont une propriété d'absence de mémoire :

$$P(X>t+s|X>s)=P(X>t)=e^{-\lambda t}. \tag{36.466}$$

Démonstration. Étant donné que X et Y sont indépendantes, la densité conjointe est le produit des densités (36.27). Nous avons donc

$$P(X>Y)=\int_D \lambda e^{-\lambda x}\mu e^{-\mu y}dxdy \tag{36.467}$$

où D est le domaine $D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } x,y>0, x>y\}$. Nous avons donc

$$P(X>Y)=\lambda\mu\int_0^\infty dx\int_0^x dy e^{-\lambda x}e^{-\mu y}=\frac{\mu}{\lambda+\mu}. \tag{36.468}$$

```
sage: var('a,b')
(a, b)
sage: f(x,y)=exp(-a*x)*exp(-b*y)
sage: assume(a>0)
sage: assume(b>0)
sage: a*b*f.integrate(y,0,x).integrate(x,0,oo)
(x, y) |--> a*b/(a^2 + a*b)
```

Pour trouver la loi de $\min(X,Y)$, nous écrivons

$$\begin{aligned}
P\big(\min(X,Y)>t\big)&=P(X>t,Y>t) && && (36.469a)\\
&=P(X>t)P(Y>t) && \text{par indépendance} && (36.469b)\\
&=\big(1-F_X(t)\big)\big(1-F_Y(t)\big) && && (36.469c)\\
&=e^{-(\lambda+\mu)t}\mathbb{1}_{[0,\infty[}(t) && && (36.469d)\\
&=1-F_Z(t) && && (36.469e)
\end{aligned}$$

où $Z\sim\mathscr{E}(\lambda+\mu)$. □

36.5.7 Approximation de la binomiale par une Poisson

Proposition 36.153.
Soit (X_n) une suite de variables aléatoires avec $X_n \sim \mathscr{B}(n, p_n)$ telle que np_n converge vers une constante $\lambda > 0$. Alors $X_n \xrightarrow{\mathscr{L}} \mathscr{P}(\lambda)$.

Démonstration. Commençons par écrire la loi binomiale sous une forme plus adaptée au passage à la limite :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{k!} p^k (1-p)^{n-k}. \tag{36.470}$$

Le produit au numérateur contient k termes dans lesquels nous mettons n en évidence. Nous trouvons

$$P(X = k) = \frac{(np)^k \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{k-1}{n}\right)}{k!} p^k (1-p)^{n-k}. \tag{36.471}$$

Lorsque nous passons à la limite, tous les facteurs du type $1-l/n$ tendent vers 1 ainsi que $(1-p_n)^{n-k}$. Les facteurs dont la limite n'est pas 1 sont donc

$$P(X_n = k) \simeq \frac{(np_n)^k}{k!}(1-p_n)^k. \tag{36.472}$$

Nous avons

$$\lim_{n\to\infty} (1-p_n)^n = \lim_{n\to\infty} \left(1 - \frac{np_n}{n}\right)^n = e^{-\lambda}. \tag{36.473}$$

La thèse est alors obtenue en remettant les morceaux ensemble. □

Exemple 36.154.
Considérons un serveur informatique qui reçoit des requêtes. Toutes les 10^{-3} s il reçoit une requête avec une probabilité $p = 0.05$. La variable aléatoire qui consiste à donner le nombre de requêtes effectivement effectuées en une seconde suit une loi binomiale $\mathscr{B}(1000, p)$.

Déterminons la probabilité que le serveur reçoive 20 requêtes en une seconde. Nous approximons $\mathscr{B}(1000, 0.05)$ par $\mathscr{P}(50)$, et la réponse est

$$e^{-50}\frac{50^{20}}{20!} \simeq 7 \cdot 10^{-7}. \tag{36.474}$$

△

36.5.8 Loi de Poisson et loi exponentielle

subsecPoissonetexpo

Soient $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes de loi exponentielle de paramètre λ. En utilisant le produit de convolution, nous pouvons trouver la fonction de densité de la somme (voir point 36.2.7). Commençons avec deux variables aléatoires X et Y. Les densités sont

$$f_X(x) = \mathbb{1}_{[x\geqslant 0]} \lambda e^{-\lambda x} \tag{36.475a}$$
$$f_Y(y) = \mathbb{1}_{[y\geqslant 0]} \lambda e^{-\lambda y}, \tag{36.475b}$$

et la densité conjointe est alors

$$f_{X+Y}(x) = \int_{\mathbb{R}} \mathbb{1}_{[x-t\geqslant 0]} \lambda e^{-\lambda(x-t)} \mathbb{1}_{[t\geqslant 0]} \lambda e^{-\lambda t} dt \tag{36.476a}$$
$$= \lambda^2 e^{-\lambda x} \int_0^x 1\, dt \tag{36.476b}$$
$$= x\lambda^2 e^{-\lambda x}. \tag{36.476c}$$

Par récurrence si $S = X_1 + \dots + X_n$ nous trouvons

$$f_S(x) = x^{n-1}\lambda^n e^{-\lambda x}. \tag{36.477}$$

PropGMntiy

Proposition 36.155 ([785]).
Soit $(T_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi $\mathscr{E}(\lambda)$. Nous considérons la variable aléatoire $S_n = \sum_{i=1}^n X_i$ et pour chaque $t \in \mathbb{R}^+$ nous considérons

$$N_t = \max\{n \geqslant 1 \text{ tel que } \sum_{i=1}^{n} T_i \leqslant t\}. \tag{36.478}$$

Alors $N_t \sim \mathscr{P}(\lambda t)$.

Démonstration. Ce que nous devons calculer est

$$P(N_t = k) = P(S_n \leqslant t \leqslant S_{n+1}). \tag{36.479}$$

Nous introduisons la variable aléatoire $V_{n=1} = (X_1, \dots, X_{n+1})$ ainsi que l'ensemble

$$A_{n+1} = \{x \in \mathbb{R}^{n+1} \text{ tel que } x_1 + \cdots + x_n \leqslant t \leqslant x_1 + \cdots + x_n + x_{n+1}\}. \tag{36.480}$$

Le problème est donc de calculer

$$P(S_n \leqslant t \leqslant S_{n+1}) = P(V_{n+1} \in A_{n+1}^+) = \int_{A_{n+1}^+} f_{n+1}(x)dx \tag{36.481}$$

où A_{n+1}^+ est la partie de A_{n+1} dans laquelle $x_i \geqslant 0$ pour tout i et f_{n+1} est la fonction de densité conjointe des variables aléatoires X_i. Nous effectuons le changement de variables

$$s_k = \sum_{i\leqslant k} x_i \tag{36.482a}$$

$$x_k = s_k - s_{k-1} \tag{36.482b}$$

dont le déterminant vaut 1. D'autre part par indépendance des variables aléatoires X_i, la fonction de partition jointe f_{n+1} s'exprime sous la forme

$$\begin{aligned} f_{n+1}(x_1, \dots, x_{n+1}) &= f_{X_1}(x_1) \dots f_{X_{n+1}}(x_{n+1}) && \text{(36.483a)}\\ &= \lambda^{n+1} e^{-\lambda(x_1+\cdots+x_{n+1})} && \text{(36.483b)}\\ &= \lambda^{n+1} e^{-\lambda s_{n+1}}. && \text{(36.483c)} \end{aligned}$$

En ce qui concerne les bornes de l'intégrale dans les variables s_i, nous voulons que tous les x_i soient positifs, par conséquent $s_1 \geqslant 0$ et ensuite l'équation $x_k = s_k - s_{k-1}$ demande $s_k \geqslant s_{k-1}$. Les bornes sont donc données par l'ensemble

$$0 \leqslant s_1 \leqslant s_2 \leqslant \dots \leqslant s_n \leqslant t \leqslant s_{n+1}, \tag{36.484}$$

c'est-à-dire $B_n \times]t, \infty[$ où

$$B_n = \{(s_1, \dots, s_n) \text{ tel que } 0 \leqslant s_1 \leqslant s_2 \leqslant \dots \leqslant s_n \leqslant t\}. \tag{36.485}$$

Le théorème de Fubini nous permet de décomposer l'intégrale :

$$\begin{aligned} P(S_n \leqslant t \leqslant S_{n+1}) &= \int_{B_n\times]t,\infty[} \lambda^{n+1} e^{-\lambda s_{n+1}} ds_1 \dots ds_{n+1} && \text{(36.486a)}\\ &= \lambda^{n+1} \left(\int_{B_n} ds_1 \dots ds_n \right) \underbrace{\left(\int_t^\infty e^{-\lambda s_{n+1}} ds_{n+1} \right)}_{=\lambda^{-1}e^{-\lambda t}} && \text{(36.486b)}\\ &= \lambda^n e^{-\lambda t} \operatorname{Vol}(B_n) && \text{(36.486c)} \end{aligned}$$

où $\mathrm{Vol}(B_n)$ est le volume de B_n qui reste à calculer. L'ensemble $C^n = [0,t]^n$ se décompose en cellules disjointes (à ensemble de mesure nulle près) de la forme

$$C_\sigma = \{0 \leqslant s_{\sigma(1)} \leqslant s_{\sigma(2)} \leqslant \dots \leqslant s_{\sigma(n)} \leqslant t\} \tag{36.487}$$

pour chaque permutation $\sigma \in S_n$. Il y a exactement $n!$ telles cellules dans C^n. Par conséquent

$$t^n = \mathrm{Vol}(C^n) = n!\,\mathrm{Vol}(C_\sigma) = n!\,\mathrm{Vol}(B_n) \tag{36.488}$$

et $\mathrm{Vol}(B_n) = \frac{t^n}{n!}$. Finalement nous avons

$$P(N_t = n) = P(S_n \leqslant t \leqslant S_{n+1}) = \frac{(\lambda t)^n}{n!} e^{-(\lambda t)}. \tag{36.489}$$

□

36.5.9 Loi normale

La loi normale de paramètres m et $\sigma > 0$, notée $\mathcal{N}(m,\sigma^2)$ est la loi donnée par la densité

$$\gamma_{m,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]. \tag{36.490}$$

Proposition 36.156.
Si la variable aléatoire réelle X suit une loi normale $\mathcal{N}(m,\sigma^2)$, alors nous avons $E(X) = m$ et $\mathrm{Var}(X) = \sigma^2$.

Démonstration. L'espérance d'une variable aléatoire se calcule à partir de la formule (36.297) :

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} x \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] dx \tag{36.491a}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \sigma \int_{\mathbb{R}} (\sigma u + m) e^{-u^2/2} \tag{36.491b}$$

où nous avons effectué le changement de variable $u = (x-m)/\sigma$. Nous utilisons ensuite l'intégrale remarquable

$$\int_{\mathbb{R}} e^{-u^2/2} du = \sqrt{2\pi}. \tag{36.492}$$

En ce qui concerne la variance, nous avons le même genre de calculs. □

La **loi normale réduite** est la densité

$$\gamma(x) = \gamma_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{36.493}$$

La variable aléatoire X suit la loi $\mathcal{N}(m,\sigma^2)$ si et seulement si la variable aléatoire $Z = \frac{X-m}{\sigma}$ suit la loi normale réduite $\mathcal{N}(0,1)$.

PropFnCaractNorm

Proposition 36.157.
La fonction caractéristique de la distribution normale $\mathcal{N}(m,\sigma^2)$ est

$$\Phi_{\mathcal{N}(m,\sigma^2)}(t) = \exp\left(itm - \frac{\sigma^2 t^2}{2}\right). \tag{36.494}$$

Démonstration. En suivant la formule (36.268), l'intégrale à calculer est

$$\Phi_{\mathcal{N}(m,\sigma^2)}(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx. \tag{36.495}$$

Nous reconnaissons une transformée de Fourier. Afin de la calculer sans encombres, nous passons par les fonctions intermédiaires suivantes :

$$\begin{aligned} g(x) &= e^{-\frac{1}{2}x^2} \\ h(x) &= g\left(\frac{x}{\sigma}\right) \\ k(x) &= h(x-m). \end{aligned} \tag{36.496}$$

La fonction caractéristique que nous cherchons est $\frac{1}{\sigma\sqrt{2\pi}}\hat{k}(t)$. Les formules liées à la transformée de Fourier nous donnent

$$\hat{k}(t) = \hat{h}(t)e^{itm} \tag{36.497a}$$
$$\hat{h}(t) = \sigma\hat{g}(\sigma t) \tag{36.497b}$$
$$\hat{g}(t) = \int_{\mathbb{R}} e^{-itx}e^{-\frac{1}{2}x^2}dx = \sqrt{2\pi}e^{-t^2/2}. \tag{36.497c}$$

Attention : l'intégrale à calculer est une transformée de Fourier *inverse*, d'où la formule (36.497a) qui a un signe de différence avec la formule usuelle. En recombinant toutes ces expressions nous trouvons

$$\Phi_{\mathcal{N}(m,\sigma^2)} = e^{-\sigma^2t^2/2}e^{itm}, \tag{36.498}$$

ce qu'il nous fallait. □

Lemme 36.158.
Si $Z \sim \mathcal{N}(0,1)$, alors l'espérance de $X = e^{\beta Z}$ est

$$E(e^{\beta Z}) = e^{\beta^2/2}. \tag{36.499}$$

Démonstration. L'espérance se calcule en remarquant que $x^2 - 2\beta x = (x-\beta)^2 - \beta^2$:

$$E(e^{\beta Z}) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{\beta x}e^{-x^2/2}dx \tag{36.500a}$$
$$= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{1}{2}(x-\beta)^2}e^{\beta^2/2} \tag{36.500b}$$
$$= \frac{e^{\beta^2/2}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-y^2/2}dy \tag{36.500c}$$
$$= e^{\beta^2/2}. \tag{36.500d}$$

□

36.5.10 Vecteurs gaussiens

Source : [432, 786].

Définition 36.159.
*Un vecteur aléatoire $X\colon \Omega \to \mathbb{R}^d$ est un **vecteur gaussien** si toutes les combinaisons linéaires de ses composantes sont des variables aléatoires normales. En d'autres termes, X est un vecteur gaussien si pour tout vecteur u, la variable aléatoire $u \cdot X$ est gaussienne.*

Le vecteur moyenne d'un vecteur gaussien est $E(X) = \big(E(X_1), \dots, E(X_n)\big)$ et sa matrice de **variance-covariance** est la matrice

$$K_X = \mathrm{Var}(X) = E\Big[\big(X - E(X)\big) \otimes \big(X - E(X)\big)\Big] \tag{36.501}$$

où l'opération $\otimes$ est celle introduite autour de l'équation (10.671). Cela n'est rien d'autre que la matrice de covariance de la variable aléatoire $X\colon \Omega \to \mathbb{R}^d$.

Lemme 36.160.
Si les variables aléatoires réelles $X_1, \dots, X_n$ sont des variables aléatoires réelles gaussiennes indépendantes, alors le vecteur $X = (X_1, \dots, X_n)$ est un vecteur gaussien.

Démonstration. Nous devons montrer que si X et Y sont des variables aléatoires gaussiennes indépendantes, alors $X+Y$ est encore gaussienne. L'indépendance nous assure les égalités suivantes pour la fonction caractéristique :

$$\Phi_{X+Y}(t) = E\big(e^{it(X+Y)}\big) = E\big(e^{itX}\big)E\big(e^{itY}\big). \tag{36.502}$$

Dans le cas où X et Y sont gaussiens nous trouvons

$$\Phi_{X+Y}(t) = \exp\left(im_1 t - \frac{\sigma_1^2 t^2}{2}\right)\exp\left(im_2 t - \frac{\sigma_2^2 t^2}{2}\right) = \exp\left(i(m_1+m_2)t - \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}\right). \tag{36.503}$$

Étant donné que la loi d'une variable aléatoire est entièrement déterminée par sa fonction caractéristique (théorème 36.93), nous déduisons que $X+Y$ est une normale de moyenne m_1+m_2 et de variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$. □

Propfmzuol

Proposition 36.161.
La fonction caractéristique d'un vecteur gaussien est donnée par

$$\Phi_X(u) = \exp\left(iu \cdot E(X) - \frac{1}{2}u \cdot K_X u\right) \tag{36.504}$$

où K_X est la matrice de covariance de X.

Démonstration. Nous considérons la variable aléatoire réelle gaussienne $u \cdot X$. Son espérance $m = E(u \cdot X) = u \cdot E(X)$. Nous commençons par établir la formule suivante :

$$u^t K_X u = u \cdot K_X u = E\Big[\big([X - E(X)] \cdot u\big)^2\Big]. \tag{36.505}$$

Utilisant la linéarité de l'espérance,

$$\sum_{kl} E(A_{kl})u_k u_l = \sum_{kl} E(A_{kl}u_k u_l), \tag{36.506}$$

nous trouvons

$$\begin{aligned}
K_X(u,u) &= E\Big([X - E(X)] \otimes [X - E(X)]\Big)(u,u) && \text{(36.507a)}\\
&= E\Big(\big([X - E(X)] \otimes [X - E(X)]\big)(u,u)\Big) && \text{(36.507b)}\\
&= E\Big(\big([X - E(X)] \cdot u\big)\big([X - E(X)] \cdot u\big)\Big) && \text{(36.507c)}\\
&= E\left(\big([X - E(X)] \cdot u\big)^2\right). && \text{(36.507d)}
\end{aligned}$$

Par la linéarité du produit scalaire et de l'espérance,

$$[X - E(X)] \cdot u = u \cdot X - E(u \cdot X), \tag{36.508}$$

ce qui nous ramène à la variable aléatoire $u \cdot X$. Nous avons alors

$$K_X(u,u) = E\left[\big([X - E(X)] \cdot u\big)^2\right] = \operatorname{Var}(u \cdot X). \tag{36.509a}$$

Nous avons donc obtenu une forme pour la variance de la variable aléatoire $u \cdot X$. Étant donné que $u \cdot X$ est gaussienne de moyenne $m = u \cdot E(X)$ et de variance $\sigma^2 = K_X(u,u)$, nous avons

$$\Phi_{u \cdot X}(t) = \exp\big(itm - \frac{1}{2}\sigma^2 t^2\big). \tag{36.510}$$

EqJJftQX

Par ailleurs nous avons $\Phi_X(u) = \Phi_{u \cdot X}(1)$ parce que

$$\Phi_X(u) = E\big(e^{iu \cdot X}\big) = \Phi_{u \cdot X}(1). \tag{36.511}$$

En utilisant la forme (36.510) pour $\Phi_{u \cdot X}$ nous trouvons

$$\Phi_X(u) = \Phi_{u \cdot X}(1) = \exp\big(im - \frac{1}{2}\sigma^2\big) = \exp\Big(iE(u \cdot X) - \frac{1}{2}u \cdot K_X u\Big). \tag{36.512}$$

□
ThoPRkxPdQ

Théorème 36.162.
Soit $X = (X_1, \dots, X_d)$, un vecteur gaussien. Les composantes sont indépendantes si et seulement si elles sont non corrélées.

Démonstration. Nous savons que les variables aléatoires indépendantes sont non corrélées. Nous devons donc surtout prouver le contraire. Le fait que les variables aléatoires X_i soient non corrélées signifie que la matrice de covariance est

$$K_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_d^2 \end{pmatrix} \tag{36.513}$$

où $\sigma_k^2 = \text{Var}(X_k)$. Notons $m_k = E(X_k)$. Si $u \in \mathbb{R}^d$, nous avons en vertu de la proposition 36.161 que

$$\begin{aligned} \Phi_X(u) &= \exp\Big(i(u_1m_1 + \cdots + u_dm_d) - \frac{1}{2}(\sigma_1^2u_1^2 + \cdots + \sigma_d^2u_d^2)\Big) && \text{(36.514a)} \\ &= \exp\big(iu_1m_1 - \frac{1}{2}\sigma_1^2u_1^2\big) \dots \exp\big(iu_dm_d - \frac{1}{2}\sigma_d^2u_d^2\big) && \text{(36.514b)} \\ &= \Phi_{X_1}(u_1) \dots \Phi_{X_d}(u_d). && \text{(36.514c)} \end{aligned}$$

Les variables aléatoires X_i sont donc indépendantes parce que la fonction caractéristique se factorise. □

Exemple 36.163.
Nous donnons à présent un exemple de deux variables aléatoires gaussiennes qui ne forment pas un vecteur gaussien. Pour ce faire nous devons chercher des variables aléatoires non indépendantes. Soit $Y \sim \mathcal{N}(0,1)$ et ϵ une variable aléatoire (indépendante de Y) donnée par $P(\epsilon = -1) = \frac{1}{2}$, $P(\epsilon = 1) = \frac{1}{2}$. Nous considérons le vecteur $(Y, \epsilon Y)$.

D'abord montrons que ϵY est une variable aléatoire gaussienne. Soit A un borélien de $\mathbb{R}$. Nous avons

$$P(\epsilon Y \in A) = P(\epsilon = 1, Y \in A) + P(\epsilon = -1, -Y \in A). \tag{36.515}$$

Par indépendance et par symétrie de Y [43] nous trouvons

$$\begin{aligned} P(\epsilon Y \in A) &= P(\epsilon = 1)P(Y \in A) + P(\epsilon = -1)P(-Y \in A) && \text{(36.516a)} \\ &= \frac{1}{2}P(Y \in A) + \frac{1}{2}P(-Y \in A) && \text{(36.516b)} \\ &= P(Y \in A). && \text{(36.516c)} \end{aligned}$$

Nous avons donc $Y \sim \epsilon Y$, et donc ϵY est gaussienne.

En ce qui concerne la covariance, nous savons que $E(Y) = E(\epsilon) = 0$, donc

$$\text{Cov}(Y, \epsilon Y) = E(Y \cdot \epsilon Y) = E(\epsilon Y^2) = E(\epsilon)E(Y^2) = 0. \tag{36.517}$$

Note : $E(Y^2) = 1$.

Les variables aléatoires Y et ϵY ne sont pas indépendantes. En effet si elles l'étaient, Y serait aussi indépendante de $(\epsilon Y)^2 = Y^2$, alors que Y et Y^2 ne sont pas indépendantes. Donc $X = (Y, \epsilon Y)$ n'est pas un vecteur gaussien. △

43. C'est-à-dire que $P(Y \in A) = P(Y \in -A) = P(-Y \in A)$.

THOooLRZNooFXWwqy

Théorème 36.164 ([432]).
Soit $m \in \mathbb{R}^d$ et $K \in \mathbb{M}(d,\mathbb{R})$ une matrice symétrique et positive. Alors il existe un vecteur gaussien de moyenne m et de matrice de covariance K.

Démonstration. Nous effectuons la preuve avec $m = 0$ [44]. Nous choisissons l'espace probabilisé $(\Omega,\mathcal{A}) = (\mathbb{R}^r, \mathcal{B}\text{or}(\mathbb{R}^r))$ où r est le rang de K muni de la probabilité de densité

$$\gamma(u) = \frac{1}{(2\pi)^{r/2}} \exp\big(-\frac{1}{2}\|u\|^2\big). \tag{36.518}$$

EqozQZiQ

Nous considérons la variable aléatoire

$$\begin{aligned} Y_0 \colon \Omega &\to \mathbb{R}^r \\ u &\mapsto u. \end{aligned} \tag{36.519}$$

C'est une variable aléatoire gaussienne de loi $P_{Y_0} = \gamma\lambda_r$ où λ_r est la mesure de Lebesgue sur $\mathbb{R}^r$. Sa densité (36.518) s'écrit comme le produit de r gaussiennes indépendantes ; sa matrice de covariance est donc $\mathbb{1}_{r\times r}$.

Étant donné que K est symétrique et positive, il existe une matrice $d \times r$ telle que $K = AA^t$. Pour voir cela, remarquons qu'il existe une matrice $d\times d$ qui fait le travail. En effet K se diagonalise par une orthogonale (théorème 8.227) :

$$K = ADA^t = A\sqrt{D}\sqrt{D}A^t \tag{36.520}$$

où D est une matrice diagonale contenant $d-r$ zéros et $\sqrt{D}$ est la matrice que l'on imagine. Donc la matrice $L = A\sqrt{D}$ est une matrice telle que $LL^t = K$. Maintenant, étant donné que les $d-r$ dernières lignes de D sont vides, les $d-r$ dernières lignes de L n'ont pas d'importance et peuvent être choisies nulles, voire même ne pas exister. La matrice $A \in \mathbb{M}(d\times r,\mathbb{R})$ qui réalise $AA^t = K$ est la « troncature » de L à ses r premières lignes.

Nous considérons la variable aléatoire $Y = AY_0 \colon \Omega \to \mathbb{R}^d$. Étant donné que AY_0 est une transformation linéaire d'un vecteur gaussien, c'est un vecteur gaussien. Nous avons encore $m(Y) = 0$ et

$$\text{Var}(Y) = E(Y\otimes Y) = E\big((AY_0)\otimes(AY_0)\big) = AE(Y_0\otimes Y_0)A^t = AA^t = K \tag{36.521}$$

parce que $E(Y_0\otimes Y_0) = \mathbb{1}$. Nous avons utilisé les formules du produit tensoriel introduit en (10.671), et en particulier la formule (10.674). □

Lorsqu'une matrice symétrique et positive K est donnée, nous avons créé un vecteur gaussien de covariance K en créant $X = AY$ où Y est le vecteur gaussien « le plus simple » et où A est donnée par $AA^t = K$. La proposition suivante montre l'inverse : un vecteur gaussien X peut se réduire au vecteur gaussien « le plus simple » en utilisant la transformation $Y = A^{-1}X$ où A est encore donnée par $AA^t = K$. Ce résultat nous permettra de voir les vecteurs gaussiens généraux comme des « changements de coordonnées » par rapport au vecteur gaussien simple $Y = (Y_1,\dots,Y_d)$ avec $Y_i \sim \mathcal{N}(0,1)$.

PropGacmRi

Proposition 36.165.
Soit $Y = (Y_1,\dots,Y_n)$ le vecteur gaussien formé des variables aléatoires indépendantes $Y_i \sim \mathcal{N}(0,1)$. Soit K une matrice symétrique et positive et la matrice A telle que $AA^t = K$. Alors le vecteur $X = AY$ est gaussien de covariance K.

44. À mon avis le cas $m \neq 0$ se déduit alors facilement. Écrivez-moi pour me donner la preuve quand vous l'avez.

Démonstration. Nous montrons que la covariance de $X = AY$ est donnée par K. Nous avons

$$\begin{aligned} K_X &= E\big([X - E(X)] \otimes [X - E(X)]\big) && (36.522a)\\ &= E\big(A(Y - E(Y)) \otimes A(Y - E(Y))\big) && (36.522b)\\ &= E\Big(A\big([Y - E(Y)] \otimes [Y - E(Y)]\big)A^t\Big) && (36.522c)\\ &= E(AA^t) && (36.522d)\\ &= K. && (36.522e) \end{aligned}$$

Nous avons utilisé le lemme 10.236 ainsi que le fait que

$$[Y - E(Y)] \otimes [Y - E(Y)] = K_Y = \mathbb{1}. \tag{36.523}$$

Notons que $E(Y) = 0$, mais cela ne joue pas ici. □

Théorème 36.166.
Un vecteur gaussien $X\colon \Omega \to \mathbb{R}^d$ possède une densité par rapport à la mesure de Lebesgue si et seulement si sa matrice de covariance est inversible. Dans ce cas nous avons la densité

$$f_X(x) = \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{|\det(K_X)|}} \exp\left(-\frac{1}{2}[X - E(X)] \cdot K_X^{-1}[X - E(X)]\right). \tag{36.524}$$

Démonstration. Nous supposons que $E(X) = 0$. En utilisant la proposition 36.165, nous posons $X = AY$ où Y est un vecteur gaussien $Y \sim \mathcal{N}(0, \mathbb{1})$ et $AA^t = K_X$. Si K n'est pas inversible, alors A n'est pas inversible non plus. Notons $r < d$ le rang de A. Étant donné que $X = AY$, la variable aléatoire X prend presque surement ses valeurs dans l'image de A, c'est-à-dire dans un sous-espace de dimension $r < d$ de $\mathbb{R}^d$. Ce sous-espace est de mesure nulle pour la mesure de Lebesgue, mais de mesure 1 pour la mesure P_X. La mesure P_X ne peut donc pas avoir de densité par rapport à celle de Lebesgue.

Supposons maintenant que K_X soit inversible. La matrice A l'est aussi. Nous anticipons l'utilisation du théorème de changement de variable 14.281. Ici le changement de variable sera la transformation linéaire A dont le jacobien vaut

$$|\det(A^{-1})| = \frac{1}{\det A} = \frac{1}{\sqrt{|\det K_X|}}. \tag{36.525}$$

Soit γ la densité de Y. Nous posons

$$f_X(x) = \frac{1}{\sqrt{|\det K_X|}} \gamma(A^{-1}X) \tag{36.526}$$

où γ est le produit des densités de d gaussiennes usuelles $\mathcal{N}(0, 1)$. Nous allons d'abord montrer que cette formule est bien la fonction (36.524) et ensuite que X admet f_X comme densité. Nous avons

$$\gamma(A^{-1}) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2}\|A^{-1}x\|^2\right), \tag{36.527}$$

et

$$\begin{aligned} \|A^{-1}x\|^2 &= \langle A^{-1}x, A^{-1}x\rangle && (36.528a)\\ &= \langle (A^{-1})^t A^{-1}x, x\rangle && (36.528b)\\ &= \langle (AA^t)^{-1}x, x\rangle && (36.528c)\\ &= x \cdot K_X^{-1}x, && (36.528d) \end{aligned}$$

ce qui nous donne bien la formule (36.524). Nous vérifions maintenant que f_X est bien une densité pour X. Soit B un borélien de $\mathbb{R}^d$. Nous avons d'abord

$$P(X \in B) = P(AY \in B) = P(Y \in A^{-1}B). \tag{36.529}$$

Ici nous avons utilisé le fait que A était bijectif. Nous avons ensuite

$$P(X \in B) = \int_{A^{-1}B} \gamma(t)dy = \int_B \gamma\big(A^{-1}x\big)|J_{A^{-1}}(x)|dx = \int_B f_X(x)dx. \tag{36.530}$$

C'est ici que nous avons utilisé le théorème de changement de variable 14.281. □

36.5.11 Variable aléatoire de Rademacher

Une variable aléatoire **de Rademacher** est une variable aléatoire de loi

$$\epsilon \sim \delta_0 - \delta_1, \tag{36.531}$$

sur l'ensemble $\Omega = \{0, 1\}$. C'est la variable aléatoire qui prend valeur 1 ou -1 avec probabilité $\frac{1}{2}$.

Parmi les propriétés évidentes de cette variable aléatoire nous avons $E(\epsilon) = 0$ et $E(\epsilon^2) = 1$.

PropCZRNRsf

Proposition 36.167 (Inégalité de Khintchine[104])**.**
Soient $r_1, \dots, r_n$ des variables aléatoires indépendantes et identiquement distribuées de Rademacher et une combinaison linéaire

$$X = \sum_{i=1}^{n} a_i r_i \tag{36.532}$$

avec $a_i \in \mathbb{R}$. Alors

EqYBZlMgs

$$\|X\|_2 \leqslant \sqrt{e} E\big(|X|\big). \tag{36.533}$$

Démonstration. Pour rappel la définition est que

$$\|X\|_2 = \sqrt{E\big(|X|^2\big)}. \tag{36.534}$$

Vu que c'est de la quantité $E\big(|X|\big)$ que nous voulons parler, nous notons l'inégalité

$$E\big(|X|\big) = \int_\Omega \big|X(\omega)\big| dP(\omega) \geqslant \Big| \int_\Omega X(\omega) dP(\omega) \Big| = \big|E(X)\big|. \tag{36.535}$$

Nous supposons que $\sum_{j=1}^{n} a_j^2 = 1$; sinon nous multiplions les a_j par ce qu'il faut pour l'avoir et ce facteur sortira des deux côtés de (36.533).

Nous allons passer par la variable aléatoire intermédiaire

$$Y = \prod_{j=1}^{n} (1 + i a_j r_j) \tag{36.536}$$

et pour presque tout $\omega \in \Omega$ nous avons

$$\begin{aligned}
|Y(\omega)| &= \prod_{j=1}^{n} |1 + i a_j r_j(\omega)| && \text{(36.537a)}\\
&= \prod_{j=1}^{n} \sqrt{1 + a_j^2 \underbrace{r_j(\omega)^2}_{=1}} && \text{(36.537b)}\\
&= \prod_{j=1}^{n} \sqrt{1 + a_j^2} && \text{(36.537c)}\\
&\leqslant \prod_{j=1}^{n} \sqrt{e^{a_j^2}} && \text{(36.537d)}\\
&= \sqrt{\prod_{j=1}^{n} e^{a_j^2}} && \text{(36.537e)}\\
&= \sqrt{e^{\sum a_j^2}} && \text{(36.537f)}\\
&= \sqrt{e}. && \text{(36.537g)}
\end{aligned}$$

Donc $\|Y\|_\infty \leqslant \sqrt{e}$ où $\|.\|_\infty$ est la norme supremum sur Ω. Cela nous permet de donner une première inégalité à propos de $E(|X|)$. D'abord

$$\big|E(XY)\big| = \Big| \int_\Omega X(\omega) Y(\omega) dP(\omega) \Big| \leqslant \int_\Omega \big|X(\omega)\big| \|Y\|_\infty dP(\omega) = \|Y\|_\infty E\big(|X|\big), \tag{36.538}$$

ensuite en remplaçant $\|Y\|_\infty$ par la majoration que nous venons de donner de $|Y(\omega)|$,

$$\sqrt{e}E\big(|X|\big) \geqslant \big|E(XY)\big|. \tag{36.539}$$

Il nous reste à prouver que $\big|E(XY)\big| \geqslant \|X\|_2$.

Pour ce faire nous commençons par noter que pour chaque j nous avons $E(r_j) = 0$ et $E(ia_j r_j^2) = ia_j$; en utilisant l'indépendance des r_j et le lemme 36.43 nous avons alors

$$\begin{aligned}
E(r_jY) &= E\Big(r_j(1 + ia_jr_j)\prod_{k\neq j}(1 + ia_kr_k)\Big) && (36.540a)\\
&= E\big(r_j(1 + ia_jr_j)\big)\prod_{k\neq j}E(1 + ia_kr_k) && (36.540b)\\
&= ia_j\prod_{k\neq j}\big(1 + ia_kE(r_k)\big) && (36.540c)\\
&= ia_j. && (36.540d)
\end{aligned}$$

Par conséquent, en utilisant la proposition 36.47 dans le cas non indépendant,

$$E(XY) = \sum_{j=1}^{n} a_jE(r_jY) = \sum_j a_jia_j = i\sum_j a_j^2 = i. \tag{36.541}$$

Nous pouvons compléter l'équation (36.539) en

$$\sqrt{e}E\big(|X|\big) \geqslant \big|E(XY)\big| = 1, \tag{36.542}$$

et nous nous empressons de montrer que $\|X\|_2 = 1$. En effet $\|X\|_2 = \sqrt{E(|X|^2)}$ alors que

$$\begin{aligned}
\|X\|_2^2 &= E\big((\sum_i a_ir_i)^2\big) && (36.543a)\\
&= E\Big(\sum_k a_k^2r_k^2 + 2\sum_{i\neq j}a_ia_jr_ir_j\Big) && (36.543b)\\
&= \sum_k a_k^2 + 2\sum_{i\neq j}a_ia_jE(r_ir_j) && (36.543c)\\
&= 1 && (36.543d)
\end{aligned}$$

□

36.5.12 Loi de Student

Définition 36.168.
La loi χ^2 à d degrés de liberté est la loi de la variable aléatoire $Y_1^2 + \cdots + Y_n^2$ si les (Y_i) sont des variables aléatoires normales indépendantes centrées et réduites.

*La loi de **Student** à d degrés de liberté est la loi de la variable aléatoire*

$$\frac{X}{\sqrt{K/d}} \tag{36.544}$$

où $X \sim \mathcal{N}(0,1)$ et $K \sim \chi^2(d)$ sont des variables aléatoires indépendantes. Cette loi est notée $\mathcal{T}(d)$

Nous avons une illustration de la densité de la loi $\chi^2(10)$ à la figure 36.1.

L'importance de cette loi sera dans le théorème de Cochrane 37.15.

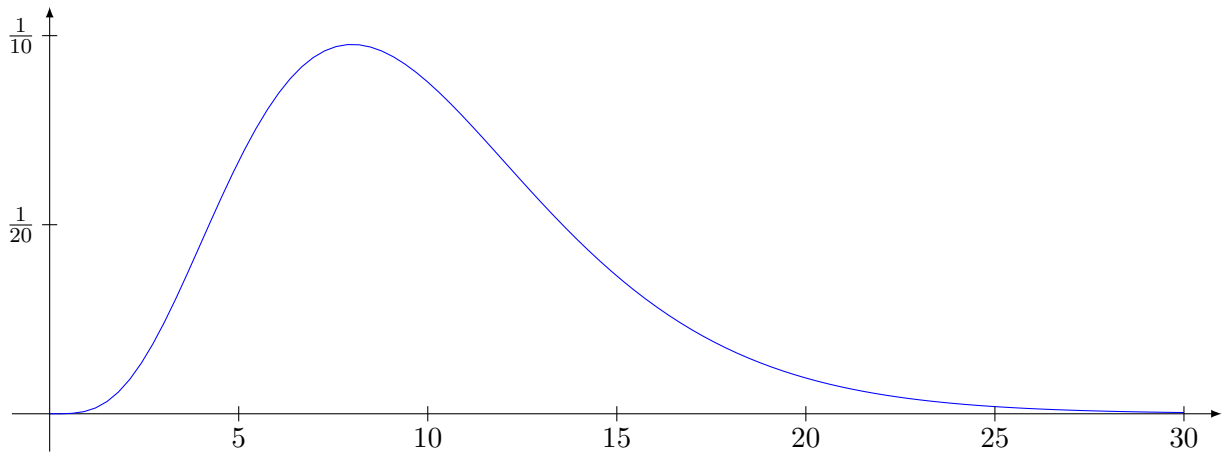


FIGURE 36.1: La densité de $\chi^2(10)$. LabelFigChiSquared

36.5.13 Indépendance, covariance et variance de somme

subsecTTHohur

Si X et Y sont des variables aléatoires réelles, nous avons défini la covariance par (36.117) :

$$\text{Cov}(X,Y) = E\Big[\big(X - E(X)\big)\big(Y - E(Y)\big)\Big] \tag{36.545}$$

Une clef est le lemme 36.43 qui dit que lorsque X et Y sont indépendantes, alors $E(XY) = E(X)E(Y)$. Nous avons les liens suivants.

(1) Si X et Y sont indépendantes, alors $\text{Cov}(X,Y) = 0$.

(2) La réciproque n'est pas vraie par l'exemple 36.45.

(3) Si Z est un vecteur gaussien, les composantes Z_i sont indépendantes si et seulement si la matrice de covariance est diagonale, c'est-à-dire que les Z_i sont deux à deux non corrélées ; c'est le théorème 36.162.

(4) En ce qui concerne la variance d'une somme,

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X,Y). \tag{36.546}$$

Donc lorsque X et Y ne sont pas corrélées, la variance est sympa avec la somme. En particulier lorsqu'elles sont indépendantes, mais pas seulement.

36.6 Estimation des grands écarts

Si S_n est une somme de variables aléatoires de Bernoulli X_i indépendantes de probabilité p, l'espérance de S_n/n est p, et nous voudrions savoir quelle est la probabilité d'avoir un taux de succès un peu plus important :

$$P\big(\frac{S_n}{n} \geqslant p + \epsilon\big). \tag{36.547}$$

La loi des grands nombres nous permet de dire que ça ne va pas être très grand. En effet si nous posons

$$Y_i = X_i - p - \epsilon \tag{36.548}$$

et

$$Z_n = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{S_n}{n} - p - \epsilon, \tag{36.549}$$

la loi des grands nombres (théorème 36.128) nous indique que

$$Z_n \xrightarrow{p.s.} E(Y_1) = -\epsilon. \tag{36.550}$$

Le théorème 36.106 sur le lien entre les types de convergence nous donne immédiatement la convergence en loi. C'est-à-dire que pour tout $\eta > 0$,

$$P\Big(|Z_n + \epsilon| \geqslant \eta\Big) \to 0. \tag{36.551}$$

En prenant $\eta = \frac{\epsilon}{2}$,

$$P(Z_n = 0) \leqslant P\big(|Z_n + \epsilon| \geqslant \frac{\epsilon}{2}\big) \to 0. \tag{36.552}$$

Tout cela montre que

$$\lim_{n\to\infty} P\big(\frac{S_n}{n} \geqslant p + \epsilon\big) = 0. \tag{36.553}$$

Autrement dit, la probabilité de tomber à une distance fixée de la moyenne tend vers zéro lorsque le nombre d'essais augmente. Rien d'étonnant.

Le théorème suivant nous indique la vitesse de convergence. Elle est exponentielle et le coefficient est donné en fonction de p et de ϵ.

ThoYYaBXkU

Théorème 36.169 ([104]).
Soient des variables aléatoires $X_i \sim \mathscr{B}(p)$ et

$$S_n = \sum_{i=1}^{n} X_i \sim \mathscr{B}(n, p). \tag{36.554}$$

Pour $\epsilon \in]0, 1 - p[$, nous définissons

$$h_+(\epsilon) = (p + \epsilon) \ln\left(\frac{p + \epsilon}{p}\right) + (1 - p - \epsilon) \ln\left(\frac{1 - p - \epsilon}{1 - p}\right). \tag{36.555}$$

Alors

ItemUWyKcpMi

(1) $h_+(\epsilon) > 0$.

ItemUWyKcpMii

(2) Pour tout $n \geqslant 1$, nous avons

$$P\left(\frac{S_n}{n} \geqslant p + \epsilon\right) \leqslant e^{-nh_+(\epsilon)}. \tag{36.556}$$

EqXFotLp

ItemUWyKcpMiii

(3) L'estimation (36.556) est optimale au sens que

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(P\big(\frac{S_n}{n} \geqslant p + \epsilon\big)\right) = -h_+(\epsilon). \tag{36.557}$$

Démonstration. Pour tout $t \geqslant 0$ nous avons :

$$\begin{aligned}
P\left(\frac{S_n}{n} \geqslant p + \epsilon\right) &= P\big(S_n \geqslant np + n\epsilon\big) && \text{(36.558a)}\\
&\leqslant E\Big(e^{t(S_n - np - n\epsilon)}\Big) && \text{subeqDUVpBeF (36.558b)}\\
&= e^{-nt(p+\epsilon)} E\left(e^{tS_n}\right) && \text{(36.558c)}\\
&= e^{-nt(p+\epsilon)} \sum_{k=0}^{\infty} e^{tk} P(S_n = k) && \text{subeqXEHKoqk (36.558d)}\\
&= e^{-nt(p+\epsilon)} \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1 - p)^{n-k} && \text{(36.558e)}\\
&= e^{-nt(p+\epsilon)} \big((1 - p) + pe^t\big)^n && \text{(36.558f)}\\
&= \exp\Big(- n\big(t(p + \epsilon) - \ln(1 - p + pe^t)\big)\Big). && \text{(36.558g)}
\end{aligned}$$

Justifications :

— L'inégalité (36.558b) est l'inégalité de Markov (corolaire 36.127) avec $\phi(x) = e^{tx}$.
— La ligne (36.558d) est une utilisation du théorème de transfert 36.96.

Nous posons maintenant

$$h = \sup_{t>0} \big(t(p+\epsilon) - \ln(1-p+pe^t)\big). \tag{36.559}$$

Pour cette valeur de h nous avons

$$P\left(\frac{S_n}{n} \geqslant p+\epsilon\right) \leqslant e^{-nh}. \tag{36.560}$$

Nous considérons la fonction

$$\begin{aligned} g\colon \mathbb{R}^+ &\to \mathbb{R} \\ t &\mapsto t(p+\epsilon) - \ln(1-p+pe^t). \end{aligned} \tag{36.561}$$

Elle vérifie $g(0) = 0$ et

$$g'(t) = (p+\epsilon) - \frac{pe^t}{1-p+pe^t} \tag{36.562a}$$

$$g'(0) = (p+\epsilon) - \frac{p}{1-p+p} = \epsilon > 0. \tag{36.562b}$$

Par conséquent sur un voisinage de $t = 0$ la fonction g est strictement croissante et nous concluons que g prend (au moins) quelques valeurs strictement positives. Du coup nous avons

$$h = \|g\|_\infty > 0. \tag{36.563}$$

Nous cherchons maintenant pour quelle valeur de t est réalisé le maximum de g. D'abord résoudre $g'(t) = 0$ donne

$$t_0 = \ln\left(\frac{(p+\epsilon)(1-p)}{p(1-p-\epsilon)}\right) \tag{36.564}$$

Vu que $g'(t) \to p+\epsilon-1 < 0$, nous avons $\lim_{t\to\infty} g(t) = -\infty$ et donc le t_0 trouvé est bien un maximum et non un minimum.

Il est maintenant loisible de calculer une valeur pour h : il suffit de calculer $g(t_0)$. La calcul n'est pas très compliqué et donne

$$h = g(t_0) = (p+\epsilon)\ln\left(\frac{p+\epsilon}{p}\right) + (p+\epsilon-1)\ln\left(\frac{1-p}{1-p-\epsilon}\right), \tag{36.565}$$

ce qui est bien $h = h_+(\epsilon)$. Cela démontre les points (1) et (2).

Nous montrons à présent l'aspect optimal de l'estimation. Nous savons déjà que

$$\frac{1}{n}\ln\left(P\big(\frac{S_n}{n} \geqslant p+\epsilon\big)\right) \leqslant h_+(\epsilon). \tag{36.566}$$

Nous posons $k_n = \lceil n(p+\epsilon)\rceil$. Vu que $p+\epsilon < 1$ et que $S_n \leqslant n$, nous avons

$$P\big(\frac{S_n}{n} \geqslant p+\epsilon\big) = P\big(S_n \geqslant n(p+\epsilon)\big) \geqslant P(S_n = k_n) = \binom{n}{k_n} p^{k_n}(1-p)^{n-k_n}. \tag{36.567}$$

C'est maintenant que nous utilisons la formule de Stirling (lemme 20.207) pour chacune des factorielles intervenant dans le coefficient binomial. Nous trouvons :

$$P\left(\frac{S_n}{n} \geqslant p+\epsilon\right) \geqslant P(S_n = k_n) = \clubsuit = \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}\alpha(n) p^{k_n}(1-p)^{n-k_n}}{\left(\frac{k_n}{e}\right)^{k_n}\sqrt{2\pi k_n}\alpha(k_n)\left(\frac{n-k_n}{e}\right)^{n-k_n}\sqrt{2\pi(n-k_n)}}. \tag{36.568}$$

Nous savons que $k_n = n(p+\epsilon) + \sigma(n)$ avec σ borné par 1. Par conséquent $n - k_n \to \infty$ et nous pouvons regrouper les coefficients en α en

$$\beta(n) = \frac{\alpha(n)}{\alpha(k_n)\alpha(n-k_n)} \to 1. \tag{36.569}$$

Nous remarquons aussi que les e se simplifient. Nous récrivons ♣ sous la forme

$$\clubsuit = \underbrace{\frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{k_n(n-k_n)}}\beta(n)}_{=A(n)} \frac{n^n p^{k_n}(1-p)^{n-k_n}}{k_n^{k_n}(n-k_n)^{n-k_n}} \tag{36.570a}$$

$$= A(n)n^n \left(\frac{p}{k_n}\right)^{k_n} \left(\frac{1-p}{n-k_n}\right)^{n-k_n} \tag{36.570b}$$

$$= A(n) \left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1-p)}{n-k_n}\right)^{n-k_n}. \tag{36.570c}$$

Nous passons au logarithme et nous étudions $\frac{1}{n}\ln\big(P(S_n = k_n)\big)$. Nous avons les termes suivants à étudier :

$$\begin{aligned}\frac{1}{n}\ln\big(P(S_n = k_n)\big) = &-\frac{1}{2n}\ln(2\pi) + \frac{1}{2n}\ln\left(\frac{n}{k_n(n-k_n)}\right)\\ &+ k_n\ln\left(\frac{np}{k_n}\right) + (n-k_n)\ln\left(\frac{n(1-p)}{n-k_n}\right) + \frac{1}{n}\ln\big(\alpha(n)\big).\end{aligned} \tag{36.571}$$

Nous étudions terme à terme la limite de cela lorsque $n \to \infty$.

(1) Le terme $\frac{1}{2n}\ln(2\pi)$ ne pose pas de problèmes. Il tend vers zéro.

(2) Si nous remplaçons k_n par $n(p+\epsilon)+\sigma(n)$ nous voyons que ce qui est dans le logarithme est majoré par $\frac{1}{P(n)}$ pour un certain polynôme P. Ce terme est dans le cas $\frac{\ln(P(n))}{n}$ qui tend vers zéro lorsque $n \to \infty$.

(3) Pour ce terme nous remplaçons k_n par $n(p+\epsilon)+k_n-n(p+\epsilon)$. Nous devons alors étudier la limite de

$$(p+\epsilon)\ln\left(\frac{np}{k_n}\right) + \frac{k_n - n(p+\epsilon)}{n}\ln\left(\frac{np}{k_n}\right). \tag{36.572}$$

EqZTmKdSn

Ce qui est dans les logarithmes est encadré de la façon suivante :

$$\frac{n(p+\epsilon)}{np} \leqslant \frac{k_n}{np} \leqslant \frac{n(p+\epsilon)+1}{np}. \tag{36.573}$$

Donc la limite de k_n/np est $(p+\epsilon)/p$. Les logarithmes restent bornés. Pour le second terme de (36.572), le numérateur du coefficient est borné par 1. Donc le second terme tend vers zéro et le tout tend vers

$$(p+\epsilon)\ln\left(\frac{p}{p+\epsilon}\right). \tag{36.574}$$

(4) Nous devons enfin étudier le dernier terme. La combinaison $\frac{n-k_n}{n}$ s'étudie de la façon suivante :

$$\frac{n-k_n}{n} = \frac{n-n(p+\epsilon)+n(p+\epsilon)-k_n}{n} = \frac{n(1-p-\epsilon)+n(p+\epsilon)-k_n}{n} \to 1-p-\epsilon \tag{36.575}$$

parce que $n(p+\epsilon)-k_n$ est borné par 1. Sachant cela, notre terme a pour limite

$$\frac{n-k_n}{n}\ln\left(\frac{n(1-p)}{n-k_n}\right) \to (1-p-\epsilon)\ln\left(\frac{1-p}{1-p-\epsilon}\right). \tag{36.576}$$

En remettant tous les morceaux bouts à bout,

$$\frac{1}{n}P(S_n = k_n) \to (1+\epsilon)\ln\left(\frac{p}{p+\epsilon}\right) + (1-p-\epsilon)\ln\left(\frac{1-p}{1-p-\epsilon}\right) = -h_+(\epsilon). \tag{36.577}$$

Étant donné que nous avions déjà prouvé que $P\big(\frac{S_n}{n} \geqslant p+\epsilon\big) \geqslant P(S_n = k_n)$, nous avons

$$\liminf_{n\to\infty} \frac{1}{n} \ln\Big(P(\frac{S_n}{n} \geqslant p+\epsilon)\Big) \geqslant \lim_{n\to\infty} \frac{1}{n} P(S_n = k_n) = -h_+(\epsilon). \tag{36.578}$$

En combinant avec (36.566) nous trouvons que

$$\lim_{n\to\infty} \frac{1}{n} \ln\Big(P(\frac{S_n}{n} \geqslant p+\epsilon)\Big) = h_+(\epsilon). \tag{36.579}$$

Pour cette dernière déduction nous utilisons le fait que si (a_n) est une suite telle que $a_n \leqslant l$ et $\liminf_{n\to\infty} a_n = l$, alors (a_n) admet une limite qui vaut l. □

36.7 Simulations de réalisations de variables aléatoires

Le générateur de base que possède un système informatique est un générateur de nombres pseudo-aléatoires de nombres entiers entre 0 et $m-1$ généré par une suite du type

$$x_{n+1} = (ax_n + b) \mod m. \tag{36.580}$$

36.7.1 Générateur uniforme

36.7.1.1 Première méthode

Une première façon de générer une variable aléatoire de loi uniforme sur $]0,1[$ est de diviser par $m-1$ la suite de x_n. En effet nous avons la proposition suivante.

Proposition 36.170.
Si (Y_n) est une suite de variables aléatoires indépendantes et identiquement distribuées de loi uniforme sur $\{0,\dots,n-1\}$. Alors

$$\frac{Y_n}{n} \xrightarrow{\mathscr{L}} \mathscr{U}[0,1]. \tag{36.581}$$

Démonstration. Nous prouvons la convergence en loi en passant par la fonction de répartition et la proposition 36.114. La fonction de répartition de la densité $\mathscr{U}[0,1]$ est

$$F_X(x) = x\mathbb{1}|_{[0,1]}. \tag{36.582}$$

La fonction de répartition de la variable aléatoire (discrète) $X_n = \frac{Y_n}{n}$ est

$$P(\frac{Y_n}{n} \leqslant x) = P(Y_n \leqslant nx) = \frac{\lfloor nx \rfloor}{n} \tag{36.583}$$

où $\lfloor a \rfloor$ est le plus grand entier inférieur à a. Nous avons évidemment

$$\lim_{n\to\infty} \frac{\lfloor nx \rfloor}{n} = x, \tag{36.584}$$

ce qui montre la convergence des fonctions de répartitions et donc la convergence en loi qui nous intéresse. □

36.7.1.2 Seconde méthode

Soit (Y_n) une suite de variables aléatoires indépendantes et identiquement distribuées selon la loi $Y_n \sim \mathscr{U}\{0,\dots,m-1\}$. Alors la série de variables aléatoires

$$Z = \sum_{k=0}^{\infty} \frac{Y_k}{m^{k+1}} \tag{36.585}$$

est une série qui converge presque surement parce que Y_k est borné par m. Avec probabilité zéro nous avons $Z = \sum_k 1/m^k$ qui converge. Nous avons

$$Z \sim \mathscr{U}[0,1]. \tag{36.586}$$

L'argument pour montrer cette loi est qu'en base m, la variable aléatoire Z a un développement décimal $Z = 0.Y_1Y_2Y_3Y_4\cdots$.

36.7.2 Simulation par inversion

Nous cherchons maintenant à simuler une loi X de fonction de répartition F.

Définition 36.171.
Soit $f\colon \mathbb{R} \to [0,1]$ une fonction croissante, continue à droite et telle que

$$\lim_{x\to-\infty} f(x) = 0 \tag{36.587a}$$
$$\lim_{x\to\infty} f(x) = 1. \tag{36.587b}$$

*L'**inverse généralisé** de f, notée f^{-1} est la fonction définie par*

$$f^{-1}(t) = \inf\{x \text{ tel que } f(x) \geqslant t\}. \tag{36.588}$$

Remarque 36.172.
L'inverse généralisé d'une fonction bijective est la vraie fonction réciproque usuelle.

PropInvgenecntddr

Proposition 36.173.
Soit f une fonction admettant un inverse généralisé f^{-1}. Alors nous avons $f^{-1}(t) \leqslant a$ si et seulement si $t \leqslant f(a)$.

La continuité à droite joue pour démontrer cette proposition.

Proposition 36.174.
Si F est la fonction de répartition de la variable aléatoire X et si U est une variable aléatoire de loi uniforme $\mathscr{U}[0,1]$, alors $F^{-1}(U)$ a la même loi que X.

Démonstration. Nous montrons que les fonctions de répartition de X et de $F^{-1}(U)$ sont identiques. En utilisant la proposition 36.173, nous avons

$$\begin{aligned} P\big(F^{-1}(U) \leqslant y\big) &= P\big(U \leqslant F(y)\big) && \text{(36.589a)}\\ &= F(y) && \text{(36.589b)}\\ &= P(X \leqslant y). && \text{(36.589c)} \end{aligned}$$

Donc $F^{-1}(U)$ est la fonction de répartition de X. □

La difficulté de la méthode par inversion est qu'il faut être capable de calculer l'inverse de la fonction de répartition de la loi à simuler.

36.7.2.1 Loi exponentielle

La loi exponentielle est une loi qui peut être simulée par inversion. Nous savons du lemme 36.89 que sa fonction de répartition [45] vaut

$$F(x) = 1 - e^{-\lambda x}, \tag{36.590}$$

et l'inverse vaut

$$F^{-1}(y) = -\frac{1}{\lambda}\ln(1-y). \tag{36.591}$$

Par conséquent, une bonne formule pour simuler une loi exponentielle est

$$-\frac{1}{\lambda}\ln(1-U). \tag{36.592}$$

Notez que U étant uniforme, nous pouvons tout autant prendre $-\ln(U)/\lambda$.

45. Définition 36.87.

36.7.3 Algorithme de Box-Muller

Il s'agit de simuler une loi gaussienne. La proposition est la suivante.

Proposition 36.175.
Si U et V sont des variables aléatoires indépendantes de même loi uniforme sur $[0,1]$, alors le couple

$$(X,Y) = \big(\sqrt{-2\ln(U)}\cos(2\pi V), \sqrt{-2\ln(U)}\sin(2\pi V)\big) \tag{36.593}$$

vérifie

(1) X est indépendante de Y

(2) X et Y sont de loi $\mathscr{N}(0,1)$.

Démonstration. Nous allons montrer la proposition en utilisant les fonctions tests. Soit donc $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction bornée et mesurable. Soient Z et W, deux variables aléatoires indépendantes de loi $\mathscr{N}(0,1)$. Nous allons montrer que

$$F\big(\varphi(X,Y)\big) = E\Big[\varphi\big(\sqrt{-2\ln(U)}\cos(2\pi V), \sqrt{-2\ln(U)}\sin(2\pi V)\big)\Big]. \tag{36.594}$$

Par indépendance de U et V, la densité du couple est le produit des densités, donc en passant aux coordonnées polaires,

$$\begin{aligned}
\diamondsuit &= E\big[\varphi(Z,W)\big] && \text{(36.595a)}\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)e^{-x^2/2}e^{-y^2/2}dxdy && \text{(36.595b)}\\
&= \frac{1}{2\pi}\int_0^{2\pi}d\theta\int_0^{\infty}\varphi(r\cos\theta, r\sin\theta)e^{-r^2/2}rdr. && \text{(36.595c)}
\end{aligned}$$

Nous posons $u = e^{-r^2/2}$ et $v = \frac{\theta}{2\pi}$. En particulier $r = \sqrt{-2\ln(u)}$ et

$$\begin{aligned}
\diamondsuit &= \int_0^1\int_0^1 \varphi\big(\sqrt{-2\ln(u)}\cos(2\pi v), \sqrt{-2\ln(u)}\sin(2\pi v)\big)dudv && \text{(36.596a)}\\
&= E\Big(\varphi\big(\sqrt{-2\ln(U)}\cos(2\pi V), \sqrt{-2\ln(U)}\sin(2\pi V)\big)\Big) && \text{(36.596b)}
\end{aligned}$$

parce que mesure $dudv$ est la densité de la loi uniforme. □

En pratique, la formule

$$(x,y) \mapsto \big(\sqrt{-2\ln x}\cos(2\pi y), \sqrt{-2\ln x}\sin(2\pi y)\big) \tag{36.597}$$

est une façon d'obtenir deux gaussiennes à partir de deux variables uniformes.

36.7.4 Méthode du rejet

La méthode du rejet permet de simuler des lois à densité. Soit f la densité de la loi à simuler. Nous faisons les hypothèses suivantes.

(1) Il existe une densité g d'une variable aléatoire facile à simuler.

(2) Il existe un $k \geqslant 0$ tel que $f(x) \leqslant kg(x)$.

Remarque 36.176.
Le k de la seconde hypothèse est nécessairement plus grand que 1. En effet,

$$1 = \int f \leqslant k\int g = k \tag{36.598}$$

parce que f et g sont des densités et ont donc une intégrale égale à 1.

Proposition 36.177.
Soient (X_n) et (U_n) des suites de variables aléatoires indépendantes au sens où non seulement les X_i et U_k sont indépendants entre eux, mais de plus X_i est indépendant de U_j pour tout i et j. Nous supposons que les X_i sont indépendantes et identiquement distribuées, de densité g et que les U_i sont indépendantes et identiquement distribuées de loi uniforme.

Nous introduisons la variable aléatoire à valeurs dans $\mathbb{N}$

$$p(\omega) = \inf\{n \geqslant 0 \text{ tel que } \alpha\big(X_n(\omega)\big) \geqslant U_n(\omega)\} \tag{36.599}$$

où α est la fonction définie par

$$\alpha(x) = \begin{cases} \frac{f(x)}{kg(x)} & si\ g(x) \neq 0 \\ 0 & si\ g(x) = 0. \end{cases} \tag{36.600}$$

Alors la variable aléatoire Y définie par

$$Y(\omega) = X_{p(\omega)}(\omega) \tag{36.601}$$

admet f pour densité.

Démonstration. D'abord étant donné que $f(x) \leqslant kg(x)$ nous avons $\alpha(x) \in [0,1]$. Nous pouvons à priori avoir $p(\omega) = \infty$, ce qui rendrait caduque la définition de $Y(\omega)$. Montrons donc pour commencer que $P(p = \infty) = 0$. En utilisant l'indépendance nous avons EqSubEqPnNinftyalphaUu

$$\begin{aligned} P\big(\alpha(X_n) < U_n, \forall n\big) &= \lim_{N\to\infty} \prod_{i=1}^{N} P\big(\alpha(X_i) < U_i\big) && \text{(36.602a)} \\ &= \lim_{N\to\infty} P\big(\alpha(X_1) < U_1\big)^N. && \text{(36.602b)} \end{aligned}$$

Pour conclure nous devons prouver que $P\big(\alpha(X_1) < U_1\big) < 1$. Pour cela nous calculons

$$\begin{aligned} P\big(\alpha(X_1) < U_1\big) &= \int_{\mathbb{R}} dx \int_0^1 du \mathbb{1}_{\alpha(x)<u} g(x) && \text{(36.603a)} \\ &= \int_{\mathbb{R}} g(x)\big(1 - \alpha(x)\big)dx && \text{(36.603b)} \\ &= \int_{\mathbb{R}} \left(g(x) - \frac{f(x)}{k}\right) dx && \text{(36.603c)} \\ &= 1 - \frac{1}{k} && \text{(36.603d)} \\ &< 1. && \text{(36.603e)} \end{aligned}$$

L'équation (36.602) nous permet donc de conclure que $P\big(\alpha(X_n) < U_n, \forall n\big) = 0$. Par conséquent la variable aléatoire $Y(\omega) = X_{p(\omega)}(\omega)$ a un sens.

Nous devons maintenant prouver que Y a bien f pour densité. Pour cela nous considérons un ensemble mesurable $A \in \mathcal{B}\text{or}(\mathbb{R})$ et nous montrons que $P(Y \in A) = \int_A f(x)dx$. Nous avons

$$P(Y \in A) = P(X_p \in A) = \sum_{i=1}^{\infty} P(X_j \in A, p = j). \tag{36.604}$$ EqdscPAYXjApj

Par ailleurs nous avons

$$\begin{aligned} P(X_j \in A, p = j) &= P\big(X_j \in A, \alpha(X_j) \geqslant U_j, \alpha(X_m) < U_m\, \forall m \leqslant j-1\big) \\ &= P\big(X_j \in A, \alpha(X_j) \geqslant U_j\big) P\big(\alpha(X_1) < U_1\big)^{j-1} \\ &= P\big(X_j \in A, \alpha(X_j) \geqslant U_j\big) \left(1 - \frac{1}{k}\right)^{j-1}. \end{aligned} \tag{36.605}$$ EqalPXjAvpjalphaXj

Étant donné que $g(x)dx$ est la densité de X_j et que du est la densité de U, nous avons

$$P\big(X_j \in A, \alpha(X_j) \geqslant U_j\big) = \int_{\mathbb{R}} \int_0^1 \mathbb{1}_{x\in A}\mathbb{1}_{\alpha(x)\geqslant u} g(x)dudx \tag{36.606a}$$

$$= \int_{\mathbb{R}} g(x)\mathbb{1}_{x\in A} \underbrace{\int_0^1 \mathbb{1}_{\alpha(x)\geqslant u} du}_{\alpha(x)}\, dx. \tag{36.606b}$$

$$= \int_{\mathbb{R}} \mathbb{1}_{x\in A}\frac{f(x)}{k}dx \tag{36.606c}$$

$$= \frac{1}{k}P(X \in A). \tag{36.606d}$$

En remplaçant dans l'équation (36.605) nous trouvons

$$P(X_j \in A, p = j) = \frac{1}{k}P(X \in A)\left(1 - \frac{1}{k}\right)^{j-1}. \tag{36.607}$$

Et enfin, l'équation (36.604) donne

$$P(Y \in A) = \frac{1}{k}P(X \in A)\sum_{i=1}^{\infty}\left(1 - \frac{1}{k}\right)^{j-1} = P(X \in A). \tag{36.608}$$

□

36.7.5 Simuler une loi géométrique à l'ordinateur

Si (X_n) est une suite de variables aléatoires indépendantes et identiquement distribuées avec $X_n \sim \mathscr{B}(p)$, alors

$$Z = \min\{k \geqslant 1 \text{ tel que } X_k = 1\} \sim \mathscr{G}(p). \tag{36.609}$$

Nous avons alors $P(Z = k) = (1-p)^k p$.

Si nous avons un générateur de lois de Bernoulli de paramètre p, alors nous simulons jusqu'à obtenir 1 et nous comptons combien de simulations ont été nécessaires.

36.7.6 Simuler une loi exponentielle à l'ordinateur

Nous pouvons utiliser la méthode de l'inversion. Étant donné que la fonction de répartition de la loi exponentielle est $F(x) = 1 - e^{-\lambda x}$, nous avons $F^{-1}(y) = \frac{1}{\lambda}\ln(1-y)$. Par conséquent à partir d'un générateur uniforme U, nous pouvons calculer

$$F^{-1}(U) = \frac{1}{\lambda}\ln(U) \tag{36.610}$$

qui suivra une loi exponentielle d'espérance $1/\lambda$.

36.7.7 Simuler une loi de Poisson à l'ordinateur

Nous savons du point 36.5.8 que si les T_i sont des variables aléatoires indépendantes et identiquement distribuées de loi $\mathscr{E}(\lambda)$, alors nous avons

$$\max\{n \geqslant 1 \text{ tel que } \sum_i T_i \leqslant 1\} \sim \mathscr{P}(\lambda). \tag{36.611}$$

La façon usuelle pour créer une loi exponentielle est d'avoir un générateur de loi uniforme U_i et d'écrire que

$$-\frac{1}{\lambda}\ln(U_i) \sim \mathscr{E}(\lambda). \tag{36.612}$$

Nous devons donc faire la somme de telles variables aléatoires et voir à partir de quel moment la somme dépasse 1. Le calcul est le suivant :

$$-\sum_{i=1}^{n}\frac{1}{\lambda}\ln(U_i)\leqslant 1 \tag{36.613}$$

implique

$$\prod_{i=1}^{n}U_i\leqslant e^{-\lambda}. \tag{36.614}$$

En pratique, la variable aléatoire qui se comporte comme une loi de Poisson de paramètre λ est

$$N=\max\{n\geqslant 1 \text{ tel que } \prod_{i=1}^{n}U_i\geqslant e^{-\lambda}\}. \tag{36.615}$$

Nous générons donc des nombres aléatoires entre 0 et 1 et nous effectuons le produit jusqu'à ce qu'il passe en dessous de $e^{-\lambda}$. À ce moment, nous retournons le nombre de nombres qu'il a fallu générer.

36.8 Sage

Nous allons montrer maintenant quelques trucs importants dans l'utilisation de Sage pour réaliser des petits graphiques.

Remarque 36.178.
Dans ce qui suit, nous allons parler de « Sage », mais en réalité nous allons surtout parler du module `scipy` qui fait partie des modules hyper-usuels de Python. Les remerciements vont donc au moins autant du côté de l'équipe de `scipy` que vers celle de Sage.

36.8.1 Loi exponentielle

Il faut savoir que la définition d'une loi continue retourne automatiquement la loi centrée réduite. Pour avoir une loi exponentielle de moyenne donnée, il faut donc préciser de façon plus maligne que ce que l'on croit.

```
from scipy import stats

X=stats.expon(scale=5)
print(X.mean())      # retourne 5

P=plot( X.pdf,x,0,10 )
show(P)              # Affiche le graphique
```

tex/frido/code_sage1.py

36.8.2 Inverser des lois

Pour trouver des intervalles de confiance, il faut souvent calculer des inverses de loi. Bien entendu Sage le fait. Ce que sage connaît, c'est l'inverse de la fonction de survie. Autrement dit si `X` est une variable aléatoire, `X.sf` est la fonction $x\mapsto 1-P(X<x)$ et `X.isf` en est l'inverse. Pour résoudre $P(X<\xi)=\alpha$, il faut résoudre $F(\xi)=\alpha$, c'est-à-dire

$$1-F(\xi)=1-\alpha, \tag{36.616}$$

ce qui se fait de la façon suivante : le programme suivant donne pour une loi normale centrée réduite la valeur de ξ pour laquelle $P(N<\xi)=0.05$:

```
from scipy import stats

N=stats.norm
print N.mean()        # 0
print N.var()         # 1

xi = N.isf(0.95)
print xi              # -1.64485

N.cdf(xi)             # Vérification : 0.05

# Graphiques de la fonction de densité et la cumulative.
P=plot(N.cdf,x,-10,10)
Q=plot(N.pdf,x,-10,10,color="red")
show(P+Q)
```

tex/frido/code_sage2.py

36.9 Monte-Carlo

Nous voudrions calculer une valeur approchée de l'intégrale

$$I = \int_a^b f(x)dx. \tag{36.617}$$

Les méthodes classiques consistent à discrétiser l'intervalle $[a, b]$ et en calculant une somme de la forme $\sum_i w_i f(x_i)$.

L'idée de Monté Carlo est de remplacer le découpage déterministe x_i par des variables aléatoires X_i en trois étapes.

(1) Pour cela nous commençons par écrire l'intégrale comme une espérance : $I = E(X)$ où X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ à déterminer. Une contrainte est évidemment d'avoir $X \in L^1(\Omega, \mathcal{F}, P)$.

(2) Nous générons une suite de variables aléatoires indépendantes et identiquement distribuées (X_n) de même loi que X et la loi (forte) des grands nombres implique que

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{p.s.} E(X) = I. \tag{36.618}$$

(3) Le dernier point sera de donner un intervalle de confiance.

ExempleIintfdxEXu

Exemple 36.179.
Nous voudrions déterminer de façon approchée l'intégrale $I = \int_0^1 f(x)dx$. Si $U \sim \mathscr{U}[0, 1]$, alors

$$I = E\big(f(U)\big) \tag{36.619}$$

et il suffit de faire

$$I \simeq \frac{1}{n}\sum_{k=1}^{n} f(U_i). \tag{36.620}$$

où mes U_i sont indépendantes et identiquement distribuées de loi $\mathscr{U}[0, 1]$. △

Exemple 36.180.
Supposons que la fonction à intégrer se présente sous la forme $f(x) = h(x)g(x)$ avec $g \geqslant 0$ et telle que l'intégrale $\int_{\mathbb{R}} g$ existe. Notons

$$c = \int_{\mathbb{R}} g \tag{36.621}$$

et

$$I = \int_{\mathbb{R}} h(x)c\frac{g(x)}{c}dx. \tag{36.622}$$

Nous avons alors $I = E\big(ch(Y)\big)$ où Y admet la densité $g(x)c$. △

Passons au cas de plusieurs variables et considérons l'intégrale

$$I = \int_{[0,1]^d} f(x_1, \dots, x_d)dx_1 \dots dx_d. \tag{36.623}$$

Nous écrivons

$$I = E\big(f(U_1, \dots, U_d)\big) \tag{36.624}$$

où les U_i sont de loi uniforme sur $[0, 1]$. En pratique, nous générons une suite de variables aléatoires (Z_k) de lois uniformes que nous regroupons par paquets :

$$V_k = (Z_{dk}, Z_{dk+1}, \dots, Z_{d(k+1)-1}). \tag{36.625}$$

Ces variables aléatoires V_k sont indépendantes et identiquement distribuées de loi $\mathscr{U}[0, 1]^d$. Ensuite la loi des grands nombres nous indique que

$$I \sim \frac{1}{n}\sum_{k=1}^{n} f(V_k). \tag{36.626}$$

36.9.1 Intervalle de confiance

36.9.1.1 Principe

Nous supposons que nous travaillons sur une approximation de Monte-Carlo telle que la variable aléatoire choisie soit dans L^2. La loi des grands nombres nous dit que $\bar{X}_n \sim I$ tandis que le théorème central limite nous enseigne que

$$\frac{\bar{X}_n - E(X)}{\sigma/\sqrt{n}} \xrightarrow{\mathscr{L}} \mathscr{N}(0, 1). \tag{36.627}$$

Par conséquent

$$P\left(\frac{\bar{X}_n - E(X)}{\sigma/\sqrt{n}} \in [-u, u]\right) \simeq P(-u \leqslant Z \leqslant u) \tag{36.628}$$

où $Z \sim \mathscr{N}(0, 1)$. En remplaçant $E(X)$ par I et en effectuant les manipulations usuelles, nous trouvons que $P(I \in J_\alpha) = 1 - \alpha$ si

$$J_\alpha = \big[\bar{X}_n - u_\alpha\frac{\sigma}{\sqrt{n}}, \bar{X}_n + u_\alpha\frac{\sigma}{\sqrt{n}}\big] \tag{36.629}$$

où σ^2 est la variance de X. Si σ n'est pas connue, alors nous le remplaçons par un estimateur

$$S'_n = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2 \tag{36.630}$$

et nous considérons l'intervalle

$$J'_\alpha = \big[\bar{X}_n - u_\alpha\frac{S'_n}{\sqrt{n}}, \bar{X}_n + u_\alpha\frac{S'_n}{\sqrt{n}}\big]. \tag{36.631}$$

Il y a deux façons de faire diminuer la longueur de l'intervalle de confiance : augmenter n ou diminuer σ. Pour le second point, le choix de X dans $I = E(X)$ est essentiel.

Exemple 36.181.
Soit à calculer

$$I = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\beta z} e^{-z^2/2} dz \tag{36.632}$$

avec $\beta > 0$. Nous introduisons la variable aléatoire $X = e^{\beta Z}$ avec $Z \sim \mathcal{N}(0, 1)$. Nous avons alors

$$I = E(X). \tag{36.633}$$

Par ailleurs l'intégrale demandée vaut $e^{\beta^2/2}$. En appliquant les formules vues plus haut nous trouvons

$$J_\alpha = \left[\bar{X}_n - u_\alpha \frac{\sigma}{\sqrt{n}}, \bar{X}_n + u_\alpha \frac{\sigma}{\sqrt{n}}\right] \tag{36.634}$$

où

$$\sigma^2 = \text{Var}(X) = E(e^{2\beta Z}) - E(e^{\beta z})^2 = e^{2\beta^2} - e^{\beta^2}. \tag{36.635}$$

Nous avons utilisé la formule (36.500). Si nous choisissons $\beta = 2$, nous trouvons $\sigma^2 \simeq 2926$. Donc si nous voulons une longueur de J_α plus petite que 10^{-2} tout en demandant $\alpha = 0.05$ (ce qui implique $u_\alpha = 1.96$), nous devons avoir

$$1.96 \frac{2973}{\sqrt{n}} < 10^{-2}, \tag{36.636}$$

c'est-à-dire environ $n = 10^{11}$, ce qui soit dit en passant est très largement au delà des capacités de la commande de scilab[787] [46]. △

Nous allons maintenant voir quelques méthodes pour réduire la variance.

36.9.1.2 Échantillonnage préférentiel

Nous devons calculer $I = \int_{\mathbb{R}} f(x)dx$. Pour cela nous introduisons une densité g et nous écrivons

$$I = \int_{\mathbb{R}} \frac{f(x)}{g(x)} g(x) dx = E\left(\frac{f(Y)}{g(Y)}\right) \tag{36.637}$$

où Y est de densité g. Il faut essayer de trouver g de telle sorte que

$$\text{Var}\left(\frac{f(Y)}{g(Y)}\right) \tag{36.638}$$

soit la plus petite possible.

36.9.1.3 Méthode de la variable de contrôle

Soit $I = E(X)$. Nous introduisons une variable aléatoire Z et nous écrivons

$$I = E(X - Z) + E(Z). \tag{36.639}$$

Il faut alors choisir Z de telle sorte que $E(Z)$ soit calculable et que $X - Z$ ait une variance plus faible. En particulier, Z ne peut pas être indépendante de X.

46. Je n'ai pas vérifié si c'est encore le cas avec les nouvelles versions de scilab. Écrivez-moi si vous le savez.

36.9.1.4 Variables antithétiques

Soit $I = \int_0^1 f(x)dx$. La première idée (exemple 36.179) est d'écrire

$$I = E\big(f(U)\big) \tag{36.640}$$

où $U \sim \mathscr{U}[0,1]$, mais nous n'avons pas de garanties sur la variance de $f(U)$. Nous pouvons écrire

$$I = E\big(f(U)\big) = E\Big[\frac{1}{2}\big(f(U) - f(1-U)\big)\Big]. \tag{36.641}$$

Ici $1-U$ est encore une variable aléatoire uniforme sur $[0,1]$, mais il se fait que la variable aléatoire

$$Z = \frac{1}{2}\big(f(U) + f(1-U)\big) \tag{36.642}$$

a une variance inférieure à $\text{Var}\big(f(U)\big)$. En effet, $f(U)$ et $f(1-U)$ ne sont pas indépendantes, par conséquent le résultat du lemme 36.44 n'est pas valide, par contre la proposition 36.54 reste vraie et nous avons

$$\text{Var}(Z) = \frac{1}{4}\text{Var}\big(f(U)\big) + \frac{1}{4}\text{Var}\big(f(1-U)\big) + \frac{1}{2}\text{Cov}\big(f(U), f(1-U)\big). \tag{36.643}$$

Nous avons $\text{Var}\big(f(1-U)\big) = \text{Var}\big(f(U)\big)$. En ce qui concerne le terme avec la covariance, nous lui appliquons l'équation (36.106) :

$$\begin{aligned}
\text{Cov}\big(f(U), f(1-U)\big) &= E\Big((f(U) - I)(f(1-U) - I)\Big) && \text{(36.644a)}\\
&\leqslant E\big((f(U) - I)^2\big)^{1/2} E\big((f(1-U) - I)^2\big)^{1/2} && \text{(36.644b)}\\
&= \text{Var}\big(f(U)\big) && \text{(36.644c)}
\end{aligned}$$

où nous avons utilisé le fait que $E\big(f(U)\big) = E\big(f(1-U)\big) = I$. Au final nous avons bien obtenu

$$\text{Var}(Z) \leqslant \text{Var}\big(f(U)\big). \tag{36.645}$$

36.10 Résultats qui se démontrent avec des variables aléatoires

36.10.1 Nombres normaux

Tout nombre $x \in [0,1[$ admet un unique[47] développement en base $b \geqslant 2$:

$$x = \sum_{n=1}^{\infty} \frac{\epsilon_n(x)}{b^n} \tag{36.646}$$

avec $\epsilon(x) \in \mathbb{D}_b$. Nous excluons $x = 1$ parce que son développement en puissances négatives de b est zéro.

Nous notons $\mathcal{A} = \{0, \dots, b-1\}$. Soit $k \geqslant 1$ et $r \in \mathcal{A}^k$; nous posons

$$N_x(r,n) = \text{Card}\big\{i \in \{1, \dots, n-k+1\} \text{ tel que } \epsilon_1(x) = r_1, \dots, \epsilon_{i+k-1} = r_k\big\}. \tag{36.647}$$

C'est le nombre d'occurrences du motif r (de longueur k) dans les n premières décimales de x.

Définition 36.182.
*Un nombre $x \in [0,1[$ est **normal** en base b si pour tout $r \in \{0, \dots, b-1\}^k$ nous avons*

$$\frac{N_x(r,n)}{n} \to \frac{1}{b^k}. \tag{36.648}$$

Un nombre est normal si il est normal en toute base.

47. Voir le théorème 10.146 pour plus de détails.

PropEEOXLae

Proposition 36.183 ([788, 104]).
Au sens de la mesure de Lebesgue, presque tous les nombres de $[0,1[$ *sont normaux.*

Démonstration. Pour $x \in [0,1[$, nous notons $\epsilon_n(x)$ son développement en base b. Cela nous donne des variables aléatoires $\epsilon_i : [0,1[\to \mathcal{A}$ dont la loi de probabilité est donnée par

$$P(\epsilon_1 = d) = P\big([\frac{d}{b}, \frac{d+1}{b}[\big) = \frac{1}{b} \tag{36.649}$$

parce que l'intervalle $[\frac{d}{b}, \frac{d+1}{d}[$ est l'ensemble des nombres de $[0,1[$ dont la première décimale est d. Pour la loi des ϵ_i, il faut un peu plus découper, mais ça donne le même résultat : $P(\epsilon_i = d) = 1/b$. Ces variables aléatoires sont indépendantes et identiquement distribuées. Nous considérons aussi la variable aléatoire

$$\begin{aligned} N(r,n) : [0,1[&\to \mathbb{N} \\ x &\mapsto N_x(r,n) \end{aligned} \tag{36.650}$$

Pour un $r \in \mathcal{A}$ fixé, nous définissons encore la variable aléatoire

$$\begin{aligned} X_j : \mathcal{A} &\to \{0,1\} \\ x &\mapsto \begin{cases} 1 & \text{si } \epsilon_j(x) = b \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{36.651}$$

Les variables aléatoires X_j sont des variables aléatoires de Bernoulli indépendantes et identiquement distribuées de paramètre $E(X_j) = P(X_j = 1) = P(\epsilon_1 = b) = \frac{1}{b}$. Nous pouvons utiliser dessus la loi forte des grands nombres (théorème 36.128). Pour dire que

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} E(X_1) = \frac{1}{b}. \tag{36.652}$$

EqNALwzsh

Mais en réalité nous avons aussi $\sum_{j=1}^{n} X_j = N(r,n)$ parce que en appliquant à $x \in [0,1[$:

$$\sum_{j=1}^{n} \begin{cases} 1 & \text{si } \epsilon_j(x) = r \\ 0 & \text{sinon} \end{cases} = \text{Card}\big\{i \in \{1,\dots,n\} \text{ tel que } \epsilon_i(x) = r\big\} = N_x(r,n), \tag{36.653}$$

de sorte que l'équation (36.652) nous dit exactement que pour tout $r \in \mathcal{A}$,

$$\lim_{n\to\infty} \frac{N_x(r,n)}{n} = \frac{1}{b} \tag{36.654}$$

pour presque tout $x \in [0,1[$.

Il reste à prouver la même chose pour tout $r \in \mathcal{A}^k$. Voyons avec $k = 2$ et $r = (u,v) \in \mathcal{A}^2$. Nous posons

$$Y_j = \mathbb{1}_{\{\epsilon_j = u, \epsilon_{j+1} = v\}}, \tag{36.655}$$

et $N(r,n) = \sum_{j=1}^{n-1} Y_j$. Les Y_i sont encore des binomiales de paramètre $\frac{1}{b^2}$, mais elles ne sont pas indépendantes. En effet pour avoir $Y_1(x) = Y_2(x) = 1$, il faut que les trois premières décimales de x soit en même temps de la forme $uv.$ et $.uv$, donc

$$P(Y_1, Y_2 = 1) = \delta_{u,v}/b^3 \tag{36.656}$$

alors que $P(Y_1 = 1)P(Y_2 = 2) = 1/b^4$. Nous pouvons contourner ce problème en remarquant que les ϵ_i, eux, sont indépendants. Donc le lemme de regroupement 36.19 nous dit que la famille $\{Y_{2n}\}$ est une famille de variables aléatoires indépendantes (et idem pour la famille Y_{2n-1}). En effet, les variables aléatoires Y_{2n} correspondent à la partition 23, 45, 67, etc.

Nous appliquons la loi des grands nombres sur les deux familles indépendamment :

$$\frac{1}{n}\sum_{j=1}^{n} Y_{2j-1} \xrightarrow{p.s.} \frac{1}{b^2} \tag{36.657}$$

et

$$\frac{1}{n}\sum_{j=1}^{n} Y_{2j} \xrightarrow{p.s.} \frac{1}{b^2} \tag{36.658}$$

Pour rappel, le but pour l'instant est d'établir la limite $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} Y_j = \frac{1}{b^2}$. Nous allons l'établir séparément pour les termes pairs et impairs de la suite. Pour les pairs :

$$\frac{1}{2n}\sum_{j=1}^{2n} Y_j = \frac{1}{2}\left(\frac{1}{n}\sum_{j=1}^{n} Y_{2j-1}\right) + \frac{1}{2}\left(\frac{1}{n}\sum_{j=1}^{n} Y_{2j}\right) \xrightarrow{p.s.} \frac{1}{b^2} \tag{36.659}$$

Pour les impairs [48],

$$\frac{1}{2n-1}\sum_{j=1}^{2n-1} Y_j = \frac{n}{2n-1}\left(\frac{1}{n}\sum_{j=1}^{n} Y_{2j-1}\right) + \frac{n}{2n-1}\left(\frac{1}{n}\sum_{j=1}^{n} Y_{2j}\right) \to \frac{1}{b^2} \tag{36.660}$$

parce que les deux parenthèses convergent vers $\frac{1}{b^2}$ alors que les coefficients devant convergent vers $\frac{1}{2}$.

Au final nous avons bien

$$\frac{N(r,n)}{n} = \frac{1}{n}\sum_{j=1}^{n-1} Y_j = \frac{n-1}{n}\left(\frac{1}{n-1}\sum_{j=1}^{n-1} Y_j\right) \to \frac{1}{b^2} \tag{36.661}$$

tant que $r \in \mathcal{A}^2$.

Pour prouver la même chose avec $r \in \mathcal{A}^k$, il suffit de faire le même raisonnement en divisant en plus de paquets : $\{Y_{kj+m}\}_{m=1,\dots,k-1}$ sont indépendants et nous utilisons k fois la loi des grands nombres.

Donc pour toute base b nous savons que les nombres non-normaux en base b forment un ensemble de mesure nulle dans $[0,1[$. Il reste à voir que leur union reste de mesure nulle. Cela est vrai parce que nous avons une union dénombrable et qu'une union dénombrable d'ensembles de mesure nulle est de mesure nulle par le lemme 14.28. □

RemUXAkcuH

Remarque 36.184.
Un nombre x est normal en base b si et seulement si la suite $u_k = xb^k$ est équirépartie modulo 1 sur $[0,1]$ (c'est-à-dire quel la suite des parties fractionnelles des u_k est équirépartie). Pour le nombre $0.2357873\dots$, nous parlons de la suite $0.2357873\dots$; $0.357873\dots$; $0.57897\dots$ etc. C'est la suite des queues de suites de la suite de ses décimales [49].

Deux vidéos de numberphile à propos de nombres normaux et calculables :

— https://www.youtube.com/watch?v=5TkIe60y2GI,

— https://www.youtube.com/watch?v=LsBQhfkw5ag

36.10.2 Théorème de Bernstein

ThoDJIvrty

Théorème 36.185 (Théorème de Bernstein[104])**.**
Soit $f \in C^0\big([0,1],\mathbb{C}\big)$ et son module de continuité

$$\begin{aligned}\omega\colon [0,1] &\to \mathbb{R}\\ h &\mapsto \sup\{|f(u)-f(v)| \text{ tel que } |u-v|<h\}.\end{aligned} \tag{36.662}$$

48. Ici dans [104], la seconde somme va jusqu'à $n-1$ et je ne comprends pas pourquoi.
49. C'est pas trop bien dit, mais on se comprend, non ?

Pour $n \geqslant 0$ nous définissons le n^e ***polynôme de Bernstein*** *de f par*

$$B_n(f)(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right). \tag{36.663}$$

Alors il existe C tel que pour tout $n \geqslant 1$:

(1)

$$\|f - B_n(f)\|_\infty \leqslant C\omega\left(\frac{1}{\sqrt{n}}\right). \tag{36.664}$$

EqZQgnVqM

(2)

$$B_n(f) \xrightarrow{unif} f \tag{36.665}$$

sur $[0,1]$.

(3) L'inégalité (36.664) est optimale : il existe une fonction $g \in C^0\big([0,1],\mathbb{C}\big)$ et $\delta > 0$ tels que pour tout $N \geqslant 1$, $\|g - B_n(g)\|_\infty \geqslant \frac{\delta}{\sqrt{n}}$. Cette fonction peut être choisie Lipschitzienne. Une telle fonction est donnée par exemple par $g(x) = |x - \frac{1}{2}|$.

(4) Les polynômes forment une partie dense dans $\Big(C^0\big([0,1]\big), \|.\|_\infty\Big)$.

Démonstration. Soit $x \in [0,1]$ et une suite de variables aléatoires de Bernoulli indépendantes[50] et identiquement distribuées $(X_i)_{i\geqslant 1}$ de paramètre x. Nous notons $S_n = \sum_{k=1}^n X_k$.

(1) Pour cette histoire de convergence, il faut majorer la quantité $\big|f(x) - B_n(f)(x)\big|$. Pour cela il y a trois astuces. La première est de se souvenir que $E\big(f(x)\big) = f(x)$, et la seconde est que le théorème de transfert 36.96 appliqué à $x \mapsto f(x/n)$ donne[51]

$$E\left(f\big(\frac{S_n}{n}\big)\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) P(S_n = k) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}, \tag{36.666}$$

c'est-à-dire que

$$B_n(f)(x) = E\left(f\big(\frac{S_n}{n}\big)\right). \tag{36.667}$$

Et enfin la troisième astuce est d'utiliser le lemme 10.311 pour avoir

$$\omega\left(|x - \frac{S_n}{n}|\right) = \omega\left(\frac{1}{\sqrt{n}}|\sqrt{n} - \frac{S_n}{\sqrt{n}}|\right) \leqslant \left(\sqrt{n}|x - \frac{S_n}{n}| + 1\right)\omega\big(\frac{1}{\sqrt{n}}\big). \tag{36.668}$$

À partir de là nous pouvons un peu calculer :

$$\begin{aligned}
\big|f(x) - B_n(f)(x)\big| &= \Big|E\left(f(x) - f\big(\frac{S_n}{n}\big)\right)\Big| && \text{(36.669a)}\\
&\leqslant E\left(|f(x) - f\big(\frac{S_n}{n}\big)|\right) && \text{(36.669b)}\\
&\leqslant E\left(\omega\Big(|x - \frac{S_n}{n}|\Big)\right) && \text{(36.669c)}\\
&\leqslant \omega\left(\frac{1}{\sqrt{n}}\right) E\left(|\sqrt{n}x - \frac{S_n}{n}| + 1\right). && \text{(36.669d)}
\end{aligned}$$

Le dernier facteur peut être récrit sous la forme

$$E\left(\sqrt{n}|x - \frac{S_n}{n}| + 1\right) = \sqrt{n}E\left(|x - \frac{S_n}{n}|\right) + 1, \tag{36.670}$$

50. Définition 36.13.
51. Nous avons aussi utilisé la formule de l'espérance pour les variables aléatoires discrètes.

et c'est là que nous pouvons utiliser l'inégalité de Hölder 27.33 :

$$E\big(|X|\big) = \|X\|_1 \leqslant \|X\|_2 \tag{36.671}$$

où $\|X\|_2$ désigne

$$\|X\|_2 = \sqrt{E\big(|X|^2\big)}. \tag{36.672}$$

Nous pouvons donc écrire

$$\big|f(x) - B_n(f)(x)\big| \leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{n}\|x - \frac{S_n}{n}\|_2 + 1\right). \tag{36.673}$$

Nous étudions maintenant de plus près la quantité $\|x - \frac{S_n}{n}\|_2$. D'abord

$$E\left(\big|x - \frac{S_n}{n}\big|^2\right) = x^2 - 2\frac{x}{n}E(S_n) + \frac{1}{n^2}E(S_n^2). \tag{36.674}$$

Ensuite nous savons l'espérance de S_n (qui vaut $E(S_n) = nx$) par (36.437) et le lemme 36.43 nous permet de calculer $E(S_n^2)$ par indépendance des X_i qui composent S_n. Nous avons alors

$$E\left(\big|x - \frac{S_n}{n}\big|^2\right) = x^2 - 2x^2 + \frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n} E(X_i)E(X_j) + \frac{1}{n^2}\sum_{i=1}^{n} E(X_i^2) \tag{36.675a}$$

$$= -x^2 + \frac{n^2 - n}{n^2}x^2 + \frac{nx}{n^2} \tag{36.675b}$$

$$= \frac{x(1-x)}{n}. \tag{36.675c}$$

Quelques justifications :

— $E(X_i) = E(X_i^2) = x$ parce que X_i est une variable aléatoire de Bernoulli de paramètre x.

— La première somme contient tous les couples (i, j) sauf les diagonaux ; il y en a donc $n^2 - n$.

En recombinant le tout, [subEqsRSuRoCJ]

$$\big|f(x) - B_n(f)(x)\big| \leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{n}\sqrt{\frac{x(1-x)}{n}} + 1\right) \tag{36.676a}$$

$$= \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{x(1-x)} + 1\right) \tag{36.676b}$$

$$\leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right). \tag{36.676c}$$

La dernière majoration est une rapide étude de la fonction $x(1-x)$.

Étant donné que les majorations (36.676) sont valables pour tout x, en passant au supremum nous avons

$$\|f - B_n(f)\|_\infty \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right) \to 0. \tag{36.677}$$

Ceci prouve les deux premiers points du théorème.

(2) Fait.

(3) Nous considérons la fonction

$$g(x) = |x - \frac{1}{2}| \tag{36.678}$$

et nous vérifions qu'elle vérifie toutes les conditions. D'abord si $u, v \in [0, 1]$ alors

$$\big|g(u) - g(v)\big| \leqslant |u - v| \tag{36.679}$$

et donc $\omega(h) \leqslant h$, ce qui signifie que g est 1-Lipschitz. Le principe de cette partie est de montrer que $\|g - B_n(g)\|_\infty$ est plus grand que d'autres trucs (et non plus petit que d'autres trucs comme d'habitude). Nous commençons par

$$\|g - B_n(g)\|_\infty \geqslant g(\frac{1}{2}) - B_n(g)(\frac{1}{2}). \tag{36.680}$$

Très vite nous nous rendons compte que $g(1/2) = 0$. Ensuite nous nous souvenons que

$$B_n(g)(\frac{1}{2}) = E\left(g(\frac{S_n}{n})\right) = E\left(|\frac{S_n}{n} - \frac{1}{2}|\right) = \frac{1}{2n}E\big(|2S_n - n|\big). \tag{36.681}$$

si nous posons $\epsilon_i = 2X_i - 1$, alors les ϵ_i sont des variables aléatoires de Rademacher indépendantes et identiquement distribuées qui satisfont à $2S_n - n = \sum_{i=1}^n \epsilon_i$. Nous utilisons la proposition 36.167 :

$$\|g - B_n(g)\|_\infty \geqslant \frac{1}{2n}E\big(|\sum_i \epsilon_i|\big) \geqslant \frac{1}{2n\sqrt{e}}\|\sum_{i=1}^n \epsilon_j\|_2. \tag{36.682}$$

Calculons ce qui est dans la norme :

$$\|\sum_{j=1}^n \epsilon_j\|_2^2 = E\left(\big(\sum_{j=1}^n \epsilon_j\big)^2\right) = \sum_{1\leqslant i\neq j\leqslant n} E(\epsilon_i)E(\epsilon_j) + \sum_{i=1}^n E(\epsilon_i^2) = 0 + n = n. \tag{36.683}$$

Nous finissons alors notre travail de majoration :

$$\|g - B_n(g)\|_\infty \geqslant \frac{1}{2n\sqrt{e}}\|\sum_{i=1}^n \epsilon_j\|_2 \geqslant \frac{1}{2\sqrt{n}\sqrt{e}} \geqslant \frac{1}{2\sqrt{e}}\omega\left(\frac{1}{\sqrt{n}}\right). \tag{36.684}$$

(4) Nous avons trouvé une suite de polynômes qui converge uniformément vers un élément arbitraire de $C^0\big([0,1]\big)$. Cela prouve la densité.

□

CORooCWLMooWwCOAP

Corolaire 36.186.
Dans $\mathbb{R}$, si $I = [a,b]$ alors les polynômes forment une partie dense dans $\big(C^0(I), \|.\|_\infty\big)$.

Démonstration. Nous supposons que $b > a$. Le cas $a = b$ est assez facile parce que l'espace des fonctions sur $\{a\}$ est de dimension 1.

Nous considérons une bijection affine $\varphi\colon [0,1] \to [a,b]$ telle que $\varphi(0) = a$ et $\varphi(1) = b$. Soit $f \in C^0(I)$.

Si $g = f \circ \varphi$, alors le théorème de Bernstein 36.185 nous donne une suite de polynômes g_k sur $[0,1]$ tels que

$$g_k \xrightarrow{unif} g. \tag{36.685}$$

Nous considérons $f_k = g_k \circ \varphi^{-1}$ qui est encore un polynôme parce que φ^{-1} est affine. Étant donné que φ^{-1} est une bijection, si h est une fonction sur $[0,1]$, nous avons

$$\sup_{x\in[a,b]} \|(h \circ \varphi^{-1})(x)\| = \sup_{y\in[0,1]} \|h(y)\|. \tag{36.686}$$

Cela nous permt le calcul suivant :

$$\begin{aligned}
\|f_k - f\|_\infty &= \|g_k \circ \varphi^{-1} - g \circ \varphi^{-1}\| && \text{(36.687a)}\\
&= \|(g_k - g) \circ \varphi^{-1}\| && \text{(36.687b)}\\
&= \sup_{x\in[a,b]} \|(g_k - g)\big(\varphi^{-1}(x)\big)\| && \text{(36.687c)}\\
&= \sup_{y\in[0,1]} \|(g_k - g)(y)\| && \text{(36.687d)}\\
&= \|g_k - g\|_\infty. && \text{(36.687e)}
\end{aligned}$$

Nous avons donc

$$\lim_{k\to\infty} \|f_k - f\|_\infty = 0, \tag{36.688}$$

ce qui prouve la densité. □

Chapitre 37

Statistiques

37.1 Notations et hypothèses

Nous notons X le caractère à étudier, et Ω l'ensemble des individus. Le caractère à étudier est vu comme une fonction sur Ω :

$$X : \Omega \to \mathbb{R}, \mathbb{N}, \{0,1\}, \dots \tag{37.1}$$

Les **statistiques descriptives** sont les techniques pour présenter et résumer les données : diagrammes, graphiques, indicateurs numériques : moyenne, écart-type, médiane, ...

Nous faisons les hypothèses suivantes :

(1) Chaque observation x_i est la réalisation de la variable aléatoire X qui sera de loi inconnue μ.

(2) Le n-uple $(x_1, \dots, x_n)$ est la réalisation de $(X_1, \dots, X_n)$ qui est l'échantillon de taille n.

(3) Les variables aléatoires X_i sont indépendantes et identiquement distribuées, de loi commune μ. La loi μ est la **loi parente** de l'échantillon.

Exemple 37.1.
Un échantillon de taille 1 consisterait à tirer au sort une personne dans une population et mesurer sa taille. △

Exemple 37.2.
Une échantillon de taille n consisterait à tirer au sort n personnes dans une population et de mesurer leurs tailles. △

L'**inférence statistique** est l'art de dégager des informations sur la population à partir d'informations partielles : intervalles de confiance, estimateurs, test d'hypothèses, ...

En théorie des probabilités, nous connaissons la loi de la variable aléatoire X et nous en déduisons des informations sur la réalisations de X : valeur la plus probable, moyenne, intervalle dans lequel $X(\omega)$ a le plus de chance d'appartenir. En statistique, au contraire, la loi est inconnue et nous cherchons des informations sur la loi à partir d'un échantillon de données numériques observées.

37.2 Modèle statistique

Un **modèle statistique** est un triplet

$$\mathcal{S} = \Big[(\Omega, \mathcal{F}, P), (X_\theta)_{\theta\in\Theta}, (\mu_\theta)_{\theta\in\Theta}\Big] \tag{37.2}$$

où $(\Omega, \mathcal{F}, P)$ est un espace probabilisé, (X_θ) est une famille de variables aléatoires définies sur Ω et telles que pour tout $\theta \in \Theta$, la variable aléatoire X_θ suive la loi μ_θ. Les μ_θ sont des mesures sur les boréliens de $\mathbb{R}$ et pour tout $B \in \mathcal{B}\text{or}(\mathbb{R})$ nous avons

$$P(X_\theta \in B) = \mu_\theta(B). \tag{37.3}$$

Remarque 37.3.
D'une certaine manière, l'introduction de μ_θ dans la définition est redondante parce que ces mesures sont déjà contenues dans la donnée des variables aléatoires X_θ.

Exemple 37.4 (Modèle statistique gaussien)**.**
Si nous savons que les variables aléatoires X_i suivent une loi gaussienne, alors nous considérons $\Theta = \mathbb{R}\times\mathbb{R}^+$ et $\theta = (m, \sigma^2)$. Dans ce cas, $\mu_\theta = \mathcal{N}(m, \sigma^2)$ et le but de la statistique est de déterminer la valeur de θ qui correspond à une population en partant de l'observation d'un échantillon. △

Définition 37.5.
Si $\Theta \subset \mathbb{R}^k$, nous disons que le modèle statistique est un modèle ***paramétrique****.*

Le modèle gaussien est un modèle paramétrique : dès que m et σ^2 sont déterminés, la loi du phénomène X est connue.

Définition 37.6.
Pour chaque $\theta \in \Theta$, nous disons qu'un ***échantillon*** *de taille n associé à un modèle statistique $\big[(\Omega, \mathcal{F}, P), (X_\theta), (\mu_\theta)\big]$ est un vecteur $\big(X_{\theta,1}, \dots, X_{\theta,n}\big)$ de taille n de variables aléatoires indépendantes et identiquement distribuées de la même loi que la variable aléatoire X_θ. La loi μ_θ est la* ***loi parente*** *de l'échantillon.*

Définition 37.7.
Un ***modèle d'échantillonnage*** *sur le modèle statistique $\mathcal{S}$ est une famille $(X_{\theta,1}, \dots, X_{\theta,n})_{\theta\in\Theta}$ d'échantillons de taille $n \geqslant 1$.*

Nous noterons souvent $(X_1, \dots, X_n)$ à la place de $(X_{\theta,1}, \dots, X_{\theta,n})$ un échantillon, mais il faut se souvenir que les X_i suivent toujours la même loi donnée par θ. La loi du vecteur $(X_1, \dots, X_n)$ est $\mu_\theta \otimes \dots \otimes \mu_\theta$ et est définie sur l'espace $(\Omega^n, \mathcal{F} \otimes \dots \otimes \mathcal{F}, P^{\otimes n})$.

Remarque 37.8.
Le travail du statisticien est de proposer un modèle statistique $\mathcal{S}$ à priori. Si nous étudions la taille d'une population, nous allons choisir un modèle gaussien. Plus le modèle est précis, plus l'espace Θ est petit mais plus il y a de risques que la vérité soit hors de l'ensemble considéré.

Exemple 37.9.
Soit X une variable aléatoire de carré intégrable que l'on sait simuler. Afin d'évaluer la moyenne μ de X, nous pouvons considérer la moyenne empirique des simulations : $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ où les variables aléatoires X_i sont indépendantes, identiquement distribuées et de même loi que X.

La loi des grands nombres nous enseigne que $\bar{X}_n \to \mu$. De plus,

$$\lim_{n\to\infty} P\left(\mu \in \left[\bar{X}_n - \frac{a\sigma}{\sqrt{n}}, \bar{X}_n + \frac{a\sigma}{\sqrt{n}}\right]\right) = \int_{-a}^{a} e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}. \tag{37.4}$$

En effet, la condition sur μ est équivalente à

$$-a \leqslant \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant a, \tag{37.5}$$

tandis que le théorème central limite nous enseigne que la variable aléatoire $\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}$ se comporte comme $\mathcal{N}(0, 1)$ lorsque n est grand. Dans ce cas, nous avons que

$$P\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \in [-a, a]\right) = \frac{1}{\sqrt{2\pi}}\int_{-a}^{a} e^{-x^2/2}dx. \tag{37.6}$$

Notons que dans ce calcul nous avons utilisé le fait que $\mu = E(X_1)$.

Montrons que la suite

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2 \tag{37.7}$$

converge presque surement vers σ^2. Le théorème central limite implique que

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p.s.} E(X_1^2) \tag{37.8}$$

et que

$$\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 \xrightarrow{p.s.} E(X_1)^2. \tag{37.9}$$

La différence converge donc presque surement vers σ^2 en vertu de la proposition 36.48.

Nous avons également $E(\sigma_n^2) = \sigma^2$. En effet, sachant que $E(X_i) = E(X_1) = \mu$ et que $E(X_i^2) = E(X_1^2) = \mu^2 + \sigma^2$,

$$E(\sigma_n^2) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) - \frac{1}{n^2}\left(\sum_{i=1}^{n} E(X_i^2)\right) + \sum_{i\neq j} E(X_iX_j)) \tag{37.10a}$$

$$= \sigma^2 + \mu^2 - \frac{1}{n^2}\big(n(\sigma^2+\mu^2) + (n^2-n)\mu^2\big) \tag{37.10b}$$

$$= \sigma^2 - \frac{1}{n}\sigma^2, \tag{37.10c}$$

dont la limite $n \to \infty$ donne bien σ^2.

Nous voudrions à présent montrer que

$$\frac{\bar{X}_n - \mu}{\sigma_n/\sqrt{n}} \xrightarrow{\mathscr{L}} \mathscr{N}(0,1). \tag{37.11}$$

Vu que le théorème central limite donne une convergence en loi, nous pouvons utiliser le lemme de Slutsky pour montrer que

$$\left(\frac{\bar{X}_n - \mu}{1/\sqrt{n}}, \sigma_n^2\right) \xrightarrow{\mathscr{L}} (\sigma Z, \sigma^2) \tag{37.12}$$

où $Z \sim \mathscr{N}(0,1)$. En vertu de la proposition 36.120 appliquée à la fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} \frac{x}{\sqrt{y}} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases} \tag{37.13}$$

nous avons la convergence en loi

$$f\left(\frac{\bar{X}_n - \mu}{1/\sqrt{n}}, \sigma_n^2\right) \xrightarrow{\mathscr{L}} f(\sigma Z, \sigma^2), \tag{37.14}$$

c'est-à-dire

$$\frac{\bar{X}_n - \mu}{\sigma_n/\sqrt{n}} \xrightarrow{\mathscr{L}} Z. \tag{37.15}$$

Afin d'être complet, précisons que

$$P\big((\sigma Z, \sigma) \in \mathbb{R} \times \{0\}\big) = 0. \tag{37.16}$$

△

PropLimxBNpxbxbsqrt

Proposition 37.10.
Soient des variables aléatoires indépendantes et identiquement distribuées X_i de loi parente $\mathscr{B}(n,p)$. Alors si $\bar{X}_n$ désigne la moyenne empirique,

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1). \tag{37.17}$$

Démonstration. Cela est une application de la loi des grands nombres, tu théorème central limite, du lemme de Slutsky et de la proposition 36.120.

D'abord, la loi des grands nombres nous indique que $\bar{X}_n \to p$ parce que p est l'espérance de Bernoulli. Ensuite nous avons

$$\frac{\bar{X}_n - E(X)}{\sqrt{\text{Var}(X)}/\sqrt{n}} = \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \tag{37.18}$$

parce que la variance d'une loi de Bernoulli est $p(1-p)$. Le théorème central limite nous indique par conséquent que

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1). \tag{37.19}$$

Le lemme de Slutsky implique alors la convergence du couple :

$$\left(\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}}, \bar{X}_n\right) \xrightarrow{\mathscr{L}} (Z, p). \tag{37.20}$$

Nous appliquons maintenant la proposition 36.120 avec la fonction

$$f(x,y) = \frac{\sqrt{p(1-p)}x}{\sqrt{y(1-y)}} \tag{37.21}$$

qui est une fonction dont l'ensemble des points de discontinuité est $C = \{0\}$. Étant donné que $P(\bar{X}_n = 0) = 0$, la proposition s'applique et nous avons

$$f\left(\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}}, \bar{X}_n\right) \xrightarrow{\mathscr{L}} f(Z, p), \tag{37.22}$$

c'est-à-dire

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1). \tag{37.23}$$

□

37.3 Modèles d'échantillonnages

Soit X, une variable aléatoire sur $(\Omega, \mathcal{F}, P)$. Un **échantillon** de taille n pour X est une suite de n variables aléatoires $(X_1, \dots, X_n)$ définies sur $(\Omega, \mathcal{F}, P)$ indépendantes et de même loi que X. Nous disons que la loi de X est la **loi parente** de la suite X_i.

Définition 37.11.
Soit

$$\mathcal{S} = \Big[(\Omega, \mathcal{F}, P), (X_\theta), (\mu_\theta)\Big]_{\theta\in\Theta}, \tag{37.24}$$

un modèle statistique. Un ***modèle d'échantillonnage*** *de taille n associé au modèle statistique $\mathcal{S}$ est la donnée d'une famille de n-échantillons $(X_{\theta,1}, \dots, X_{\theta,n})$ telle que pour tout $\theta \in \Theta$, l'échantillon $(X_{\theta,i})$ soit de variable parente X_θ.*

La **moyenne empirique** du n-échantillon (X_i) est la variable aléatoire

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{37.25}$$

La proposition suivante signifie que la moyenne empirique est une « bonne » façon d'approcher la variable aléatoire.

Proposition 37.12.
Soit X, une variable aléatoire dans $L^2(\Omega)$ (c'est-à-dire $E(X^2) < \infty$) d'espérance m et de variance σ^2. Alors

(1) $E(\bar{X}_n) = m$ et $\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$.

(2) Nous avons les convergences

$$\bar{X}_n \xrightarrow{p.s.} m \tag{37.26a}$$

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1). \tag{37.26b}$$

(3) Si X est de loi $\mathscr{N}(m, \sigma^2)$, alors $\bar{X}_n \sim \mathscr{N}(m, \frac{\sigma^2}{n})$, c'est-à-dire

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \sim \mathscr{N}(0,1). \tag{37.27}$$

Remarque 37.13.
L'intérêt de cette proposition en statistique descriptive expérimentale est le suivant. La taille moyenne des français est un nombre m qui existe, mais qui est largement hors de portée de l'expérience (mesurer $65 \cdot 10^6$ personnes risque de prendre un sacré temps). Si on mesure seulement n personnes dont les tailles sont $(x_i)_{i=1,\dots,n}$ (ici x_i est un *nombre expérimental*, pas une variable aléatoire), alors on peut calculer la moyenne $\bar{x}_n$ de ces n personnes-là. La proposition indique que si n est assez grand, alors $\bar{x}_n$ donne une bonne idée de m.

Ne pas confondre X_n qui est une variable aléatoire, c'est-à-dire une application mesurable, qui nous sert à démontrer des théorèmes en mathématique, avec x_n qui est un nombre mesuré sur le terrain, qui a une existence *physique* bien définie, mais aucun statut mathématique.

Si on croit que toute cette histoire de variables aléatoires, de tribu et de mesures décrit effectivement la réalité, alors on peut croire que le comportement de la suite $\bar{X}_n$ décrit bien le comportement de la suite $\bar{x}_n$ (cette dernière n'étant même pas une suite parce qu'on n'a jamais qu'un nombre fini de mesures expérimentales).

La **variance empirique** d'un échantillon est la variable aléatoire

$$V_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2. \tag{37.28}$$

La **variance empirique corrigée** est la variable aléatoire

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2. \tag{37.29}$$

Eqdefvarempicorri

Lemme 37.14.
La variance corrigée et la variance empirique ont comme espérances :

$$E(V_n^2) = \frac{n-1}{n}\sigma^2 \tag{37.30a}$$

$$E(S_n^2) = \sigma^2. \tag{37.30b}$$

Démonstration. Nous commençons par calculer l'espérance de la variance non corrigée. La première

étape est de la récrire sous la forme

$$\begin{aligned} V_n^2 &= \frac{1}{n}\sum_k (X_k^2 - 2X_k\bar{X}_n + \bar{X}_n^2) \\ &= \frac{1}{n}\sum_k X_k^2 - \frac{2}{n}\bar{X}_n \underbrace{\sum_k X_k}_{=n\bar{X}_n} + \frac{1}{n}\underbrace{\sum_k \bar{X}_n^2}_{=n\bar{X}_n^2} \\ &= \frac{1}{n}\sum_k X_k^2 - 2\bar{X}_n^2 + \bar{X}_n^2 \\ &= \frac{1}{n}\sum_k X_k^2 - \bar{X}_n^2. \end{aligned} \tag{37.31}$$

Nous calculons séparément l'espérance de ces deux termes. Si X est la loi parente des X_i, en utilisant l'indépendance des X_i nous trouvons

$$\begin{aligned} E\left(\frac{1}{n}\sum_k X_k^2\right) &= \frac{1}{n}\sum_{k=1}^{n} E(X_k^2) \\ &= E(X^2) \\ &= E(X)^2 - \operatorname{Var}(X). \end{aligned} \tag{37.32}$$

Nous devons à présent calculer l'espérance de $\bar{X}_n^2$:

$$E(\bar{X}_n^2) = E(\bar{X}_n)^2 + \operatorname{Var}(\bar{X}_n). \tag{37.33}$$

En utilisant le lemme 36.44,

$$\operatorname{Var}(\bar{X}_n) = \operatorname{Var}\left(\frac{1}{n}\sum_k X_k\right) \tag{37.34a}$$

$$= \frac{1}{n^2}\sum_k \operatorname{Var}(X_k) \tag{37.34b}$$

$$= \frac{1}{n}\operatorname{Var}(X). \tag{37.34c}$$

Par conséquent

$$E(V_n^2) = \operatorname{Var}(X)\left(1 - \frac{1}{n}\right) = \frac{n-1}{n}\operatorname{Var}(X). \tag{37.35}$$

En ce qui concerne la variance corrigée,

$$S_n^2 = \frac{1}{n-1}\sum_k (X_k - \bar{X}_n)^2 = \frac{n}{n-1}V_n^2, \tag{37.36}$$

par conséquent $E(S_n^2) = \frac{n}{n-1}E(V_n^2) = \operatorname{Var}(X)$. □

ThoCochraneChiStudent

Théorème 37.15 (Théorème de Cochran[432])**.**
Soient (X_i) des variables aléatoires gaussiennes indépendantes de loi $\mathcal{N}(m, \sigma^2)$ avec $\sigma > 0$. Alors

(1) $\bar{X}_n \sim \mathcal{N}(m, \frac{\sigma^2}{n})$,

ItemThoCochraneChiStudentii

(2) $\left(\frac{n-1}{\sigma^2}\right) S_n^2 = \left(\frac{n}{\sigma^2}\right) \bar{V}_n^2 \sim \chi^2(n-1)$,

(3) les variables aléatoires $\bar{X}_n$ et $\bar{V}_n$ sont indépendantes et

$$\frac{\bar{X}_n - m}{S_n/\sqrt{n}} = \frac{\bar{X}_n - m}{\bar{V}_n/\sqrt{(n-1)}} \sim \mathscr{T}(n-1). \tag{37.37}$$

Nous pouvons aussi écrire le dernier résultat en termes de la variance corrigée S_n, l'estimateur sans biais de la variance, parce que

$$\bar{V}_n\sqrt{\frac{1}{n-1}} = \frac{1}{\sqrt{n}}S_n \tag{37.38}$$

en vertu de la définition (37.29).

Proposition 37.16.
Soit X une variable aléatoire de variance $\mathrm{Var}(X) = \sigma$. Si $E(X^4) < \infty$, alors

(1) $S_n^2 \xrightarrow{p.s.} \sigma^2$.

(2)

$$\frac{S_n^2 - \sigma^2}{\sqrt{\frac{\mu^4-\sigma^4}{n}}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1) \tag{37.39}$$

où $\mu^4 = E(X^4)$ est le moment d'ordre 4 de X.

Théorème 37.17.
Si $(X_1, \dots, X_n)$ est un n-échantillon de loi parente $\mathscr{N}(m, \sigma^2)$, alors

(1) Les variables aléatoires

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \quad et \quad (n-1)\frac{S_n^2}{\sigma^2} \tag{37.40}$$

sont indépendantes.

(2) La loi de $(n-1)\frac{S_n^2}{\sigma^2}$ est $\chi^2(n-1)$.

Si un échantillon vérifie ces deux propriétés, alors les X_i sont de loi $\mathscr{N}(m, \sigma^2)$.

L'inégalité de Markov donne une borne supérieure à la probabilité qu'une variable aléatoire positive soit plus grande ou égale à une constante.

ThoInegMarkov

Théorème 37.18 (Inégalité de Markov)**.**
Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$ et $\varphi\colon \mathbb{R}^d \to [0, \infty[$. Alors

$$P\big(\varphi(X) \geqslant a\big) \leqslant \frac{E\big(\varphi(X)\big)}{a} \tag{37.41}$$

pour tout $a > 0$.

Démonstration. Calculons le second membre :

$$\begin{aligned}
\frac{E\big(\varphi(X)\big)}{a} &= \int_\Omega \frac{\varphi(X)}{a} dP \\
&= \int_{\varphi(X)\geqslant a} \underbrace{\frac{\varphi(X)}{a}}_{\geqslant 1} dP + \int_{\varphi(X)<a} \frac{\varphi(X)}{a} dP \\
&\geqslant \int_{\varphi(X)\leqslant a} dP \\
&= P\big(\varphi(X) \leqslant a\big).
\end{aligned} \tag{37.42}$$

D'où l'inégalité voulue. □

37.4 Estimation ponctuelle

Nous considérons un modèle statistique

$$\mathcal{S} = \big[(\Omega, \mathcal{F}, P), (X_\theta), (\mu_\theta)\big]_{\theta\in\Theta} \tag{37.43}$$

et pour tout θ nous notons $X = (X_{\theta,1}, \dots, X_{\theta,n})$ un échantillon de loi parente μ_θ. Tant que nous travaillerons avec un θ fixé, nous écrirons $X = (X_1, \dots, X_n)$ sans expliciter la paramètre θ. Nous noterons

$$E_\theta\big(\varphi(X_1, \dots, X_n)\big) = \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n) d\mu_\theta^{\otimes n}(x_1, \dots, x_n). \tag{37.44}$$

Dans cette notation nous plaçons le θ sur l'espérance, tandis qu'en réalité le θ devrait être sur chaque X_1. Tant qu'aucune confusion n'est possible nous ferons toujours cet abus d'écriture.

Le but de la théorie de l'estimation est de déduire la valeur de θ (et donc la loi μ_θ) à partir d'un échantillon de loi parente θ.

Nous posons les hypothèses suivantes.

(1) Le modèle statistique $\mathcal{S}$ est paramétré, c'est-à-dire que $\Theta \subset \mathbb{R}^d$ avec le plus souvent $d = 1, 2$. Typiquement les paramètres seront la moyenne et la variance.

(2) Le modèle statistique est **identifiable**, c'est-à-dire que pour tout couple $(\theta_1, \theta_2) \in \Theta^2$, si $\theta_1 \neq \theta_2$, alors $\mu_{\theta_1} \neq \mu_{\theta_2}$.

(3) Le modèle $\mathcal{S}$ est **dominé** par la mesure de Lebesgue si les lois μ_θ sont continues et par la mesure de comptage si les lois μ_θ sont discrètes.

Exemple 37.19.
Quelques familles identifiables :

— La famille des lois exponentielles $\big(\mathscr{E}(\lambda)\big)_{\lambda>0}$ est identifiable.

— Les lois gaussiennes $\big(\mathscr{N}(m, \sigma^2)\big)_{m\in\mathbb{R},\sigma^2>0}$ sont également identifiables.

En réalité il est assez compliqué de trouver un exemple de modèle non identifiable à moins de la faire exprès. Par exemple en paramétrant les lois exponentielles de la façon suivante : $\big(\mathscr{E}(\sin(\lambda))\big)_{\lambda\in\mathbb{R}}$. Cette famille n'est pas identifiable. △

Le corolaire 14.237 ainsi que l'hypothèse de modèle dominé impliquent que les lois ont des densités. Si la loi μ_θ est discrète, nous notons

$$p(x, \theta) = \mu_\theta(\{x\}) \tag{37.45}$$

la densité de μ_θ par rapport à la mesure de comptage. Si La loi μ_θ est continue, nous notons

$$p(x, \theta) = f_\theta(x) \tag{37.46}$$

la densité par rapport à la mesure de Lebesgue.

Si μ_θ est une loi discrète et si $(X_1, \dots, X_n)$ est un échantillon de taille n, alors pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$ nous avons

$$p_n(x_1, \dots, x_n; \theta) = \mu_\theta^{\otimes n}\big(\{x_1, \dots, x_n\}\big) = \mu_\theta(\{x_1\}) \dots \mu_\theta(\{x_n\}) = p(x_1, \theta) \dots p(x_n, \theta). \tag{37.47}$$

La première et la dernière égalité sont des notations ; la seconde est une conséquence de l'indépendance des X_i contenues dans l'échantillon. Pour une loi continue, nous adoptons la même notation. Le vecteur aléatoire $(X_1, \dots, X_n)$ admet la densité

$$(x_1, \dots, x_n) \mapsto p_n(x_1, \dots, x_n; \theta) = f_\theta(x_1) \dots f_\theta(x_n) = p(x_1, \theta) \dots p(x_n, \theta). \tag{37.48}$$

Exemple 37.20.
Soit $(X_1, \dots, X_n)$ un n-échantillon de loi $\mathscr{B}(1, \theta)$ avec $\theta \in]0, 1[$. C'est une loi discrète portée par l'ensemble $\{0, 1\}$. Nous avons

$$p(x, \theta) = \begin{cases} 0 & \text{si } 0 \neq x \neq 1 \\ 1 - \theta & \text{si } x = 0 \\ \theta & \text{si } x = 1. \end{cases} \tag{37.49}$$

De façon plus condensée nous pouvons écrire

$$p(x,\theta)=\theta^x(1-\theta)^{1-x}\mathbb{1}_{\{0,1\}}(x). \tag{37.50}$$

Pour tout $(x_1,\dots,x_n)\in\mathbb{R}^n$, la densité du n-échantillon est donnée par

$$p_n(x_1,\dots,x_n;\theta)=\theta^{x_1+\cdots+x_n}(1-\theta)^{n-\sum_i x_i}\mathbb{1}_{\{0,1\}^n}(x_1,\dots,x_n). \tag{37.51}$$

△

37.5 Statistiques et estimateurs

Définition 37.21.
*Une **statistique** sur un modèle d'échantillonnage est une variable aléatoire fonction de l'échantillon $(X_1,\dots,X_n)$ ne dépendant pas de θ*[1]*. C'est-à-dire une application borélienne $T\colon\mathbb{R}^n\to\mathbb{R}$ ne dépendant pas de θ. La statistique associée à cette application est $S=T(X_1,\dots,X_n)$.*

Les fonctions $T(X_1,\dots,X_n)$ données par $\sum_i X_i$, $e^{\sum X_i}$ sont des statistiques. La constante $\frac{1}{2}$ est également une statistique (mais elle est moins intéressante).

Un **estimateur** est une statistique qui prend ses valeurs dans Θ. Nous la noterons

$$\hat{\theta}_n=\theta(X_1,\dots,X_n). \tag{37.52}$$

La fonction $\hat{\theta}_n$ est borélienne à valeurs dans Θ.

Exemple 37.22.
Soit un n-échantillon de loi $\mathscr{B}(1,\theta)_{\theta\in[0,1]}$. Les fonctions $\hat{\theta}_n=\frac{1}{n}\sum_{i=1}^n X_i$ et $\hat{\varphi}_n=\frac{1}{2}$ sont des estimateurs. Cependant nous devinons que la première va être plus intéressante que la seconde. △

Pour la suite, nous travaillerons avec des estimateurs de carré intégrable, c'est-à-dire que

$$E_\theta\big(|\hat{\theta}_n(X_1,\dots,X_n)|^2\big)<\infty \tag{37.53}$$

pour tout $\theta\in\Theta$.

37.5.1 Qualité des estimateurs

Définition 37.23.
*Un estimateur est **convergeant** ou **consistant** si pour tout $\theta\in\Theta$, la suite de variables aléatoires $\hat{\theta}_n(X_1,\dots,X_n)$ converge en probabilité vers θ.*

En d'autres termes, l'estimateur $\hat{\theta}_n$ est convergent si pour tout $\theta\in\Theta$ et pour tout $\eta>0$,

$$\lim_{n\to\infty}P\big(|\hat{\theta}_n(X_1,\dots,X_n)-\theta|>\eta\big)=0. \tag{37.54}$$

La probabilité dans le membre de gauche est donnée par

$$\mu_\theta^{\otimes n}\Big(\big\{(x_1,\dots,x_n)\in\mathbb{R}^n\text{ tel que }|\hat{\theta}_n(x_1,\dots,x_n)-\theta|>\eta\big\}\Big). \tag{37.55}$$

Soit $\hat{\theta}_n$ un estimateur pour θ. Nous cherchons à minimiser l'erreur commise en remplaçant θ par $\hat{\theta}_n$. Nous introduisons donc le **risque quadratique** de l'estimateur $\hat{\theta}_n$ par

$$R(\hat{\theta}_n,\theta)=E_\theta\big((\hat{\theta}_n-\theta)^2\big). \tag{37.56}$$

Nous disons qu'un estimateur $\hat{\theta}_{n,1}$ est préférable à $\hat{\theta}_{n,2}$ si pour tout $\theta\in\Theta$ nous avons

$$R(\hat{\theta}_{n,1},\theta)<R(\hat{\theta}_{n,2},\theta). \tag{37.57}$$

1. Parce que d'habitude c'est ce qu'on cherche à estimer.

Lemme 37.24.
Une formule alternative pour le risque quadratique :

$$R(\hat{\theta}_n, \theta) = \mathrm{Var}(\hat{\theta}_n) + \big(E_\theta(\hat{\theta}_n) - \theta\big)^2 \tag{37.58}$$

Démonstration. Nous avons

$$E_\theta\Big((\hat{\theta}_n - \theta)^2\Big) = \mathrm{Var}(\hat{\theta}_n - \theta) + E_\theta(\hat{\theta}_n - \theta)^2. \tag{37.59}$$

D'une part $\mathrm{Var}(\hat{\theta}_n - \theta) = \mathrm{Var}(\hat{\theta}_n)$ et d'autre part $E_\theta(\hat{\theta}_n - \theta)^2 = \big[E(\hat{\theta}_n) - \theta\big]^2$. Par conséquent

$$R(\hat{\theta}_n, \theta) = \mathrm{Var}(\hat{\theta}_n) + \Big(E(\hat{\theta}_n) - \theta\Big)^2. \tag{37.60}$$

EqRisquestbaisiVar

□

Le **biais** de l'estimateur $\hat{\theta}_n$ est la quantité

$$E_\theta(\hat{\theta}_n) - \theta. \tag{37.61}$$

À ce niveau, nous rappelons que nous écrivons E_θ l'espérance calculée en supposant la valeur θ pour le paramètre des différentes variables aléatoires entrant dans le calcul. Voir la discussion autour de la définition (37.44).

ExytNlTq

Exemple 37.25.
Dans le cadre de la proposition 36.155, nous voulons savoir si $\frac{N_t}{t}$ est un estimateur sans biais de λ. Pour ce faire nous calculons

$$E_\lambda\left(\frac{N_t}{t}\right) = \frac{1}{t}E_\lambda(N_t) = \lambda \tag{37.62}$$

parce que $E(N_t) = \lambda t$. Ici nous avons calculé $E(N_t)$ en prenant λ pour valeur du paramètre du processus de Poisson, alors qu'en principe c'est justement le paramètre que nous voulons estimer. △

Exemple 37.26.
La moyenne empirique $\bar{X}_n$ est un estimateur sans biais de la moyenne. L'estimateur

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_i)^2 \tag{37.63}$$

est un estimateur sans biais de la variance. △

Un estimateur sans biais n'est pas toujours de meilleure qualité qu'un estimateur avec biais. En effet ce que nous voulons est de se donner un (petit) intervalle I autour de la bonne valeur de θ et de maximiser $P(\hat{\theta}_n \in I)$. Sur la figure 37.1, c'est l'estimateur biaisé rouge qui tombe plus souvent sur le bon intervalle que l'estimateur non biaisé bleu.

Nous allons maintenant étudier quelques manières de construire des estimateurs convergeant. Il vont évidemment s'appuyer sur la loi des grands nombres.

37.5.2 Méthode des moments

Sans surprises, un bon estimateur pour la moyenne est

$$\hat{\theta}_n(X_1, \dots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{37.64}$$

Plus généralement, nous supposons qu'il existe une fonction borélienne [2] $M\colon \mathbb{R} \to \mathbb{R}$ telle que

$$E_\theta\big(|M(X)|\big) < \infty \tag{37.65}$$

2. Définition 14.55.

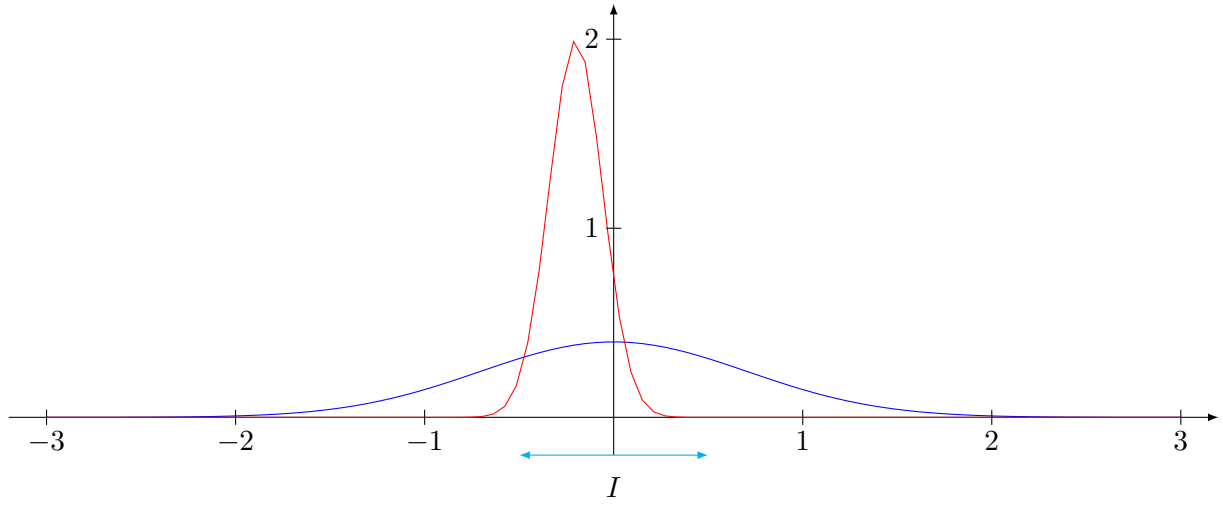


FIGURE 37.1: Un estimateur sans biais et un avec biais. LabelFigBiaisOuPas

où X est la loi parente de l'échantillon. Supposons également que la fonction

$$h(\theta) = E_\theta\big(M(X)\big) \tag{37.66}$$

soit inversible et continue sur Θ. Dans ce cas, pour estimer le paramètre θ, nous considérons l'estimateur

$$\hat{\theta}_n = h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} M(X_i)\right). \tag{37.67}$$

Cela est un estimateur convergent. En effet, la loi des grands nombres dit que

$$\frac{1}{n}\sum_i M(X_i) \xrightarrow{p.s.} E_\theta\big(M(X)\big). \tag{37.68}$$

En composant avec la fonction h, nous avons

$$\hat{\theta}_n \xrightarrow{p.s.} h^{-1}\Big(E_\theta\big(M(X)\big)\Big) = \theta. \tag{37.69}$$

Dans cette construction, $M(X)$ est le moment de X que l'on souhaite déterminer.

Exemple 37.27.
Soit $(X_1, \dots, X_n)$ un échantillon de loi $\mathscr{E}(\lambda)$. Construire $\hat{\lambda}_n$. Pour une loi exponentielle,

$$E(X) = \frac{1}{\lambda}. \tag{37.70}$$

Nous devons donc déterminer le moment d'ordre 1 de X (c'est-à-dire sa moyenne). Nous considérons donc la fonction $M(x) = x$; par conséquent

$$h(\lambda) = E(X) = \frac{1}{\lambda} \tag{37.71}$$

et $h^{-1}(\theta) = 1/\theta$. L'estimateur que nous considérons pour λ est finalement

$$\hat{\theta}_n = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i}. \tag{37.72}$$

△

37.5.3 Méthode de substitution

Supposons que nous connaissions un estimateur convergent $\hat{\theta}_n \to \theta$. Si $g\colon \mathbb{R}^d \to \mathbb{R}$ est une fonction continue, alors

$$g(\hat{\theta}_n) \to g(\theta). \tag{37.73}$$

37.5.4 Méthode du maximum de vraisemblance

ExVrasMaxLp

Exemple 37.28.
Nous désirons contrôler la qualité d'une chaine de production ; pour cela nous prélevons un échantillon de 10 pièces, et nous en trouvons 3 défectueuses. Que dire de la proportion de pièces défectueuses ?

Évidemment, le plus probable est que la proportion de pièces défectueuses soit de 1/3. Analysons en détail comment nous arrivons à ce résultat. Nous considérons que le fait de tirer 10 pièces revient à une expérience binomiale de paramètres 10 et de probabilité p inconnue. Dans ce cas, la probabilité d'observer exactement 3 pièces défectueuses est de

$$L(p) = P(X = 3) = \binom{10}{3} p^3 (1-p)^7. \tag{37.74}$$

Le maximum de $L(p)$ est $p = 3/10$.

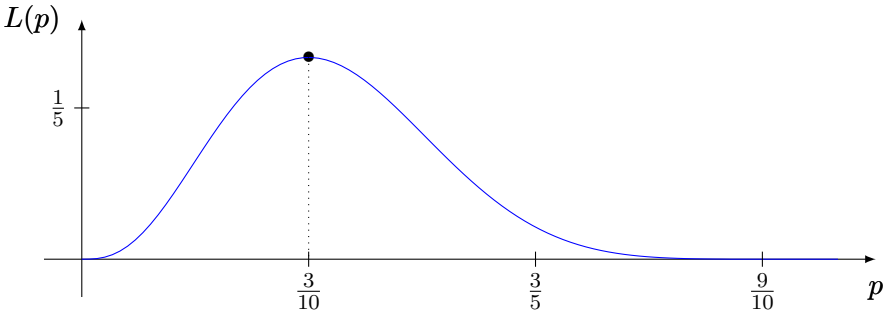


FIGURE 37.2: La fonction de vraisemblance de l'exemple 37.28. LabelFigMaxVraissLp

△

Soit $(x_1, \dots, x_n)$, une réalisation de l'échantillon $(X_1, \dots, X_n)$. L'application

$$\theta \mapsto p_n(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} p(x_i, \theta) \tag{37.75}$$

est la **vraisemblance** de l'échantillon. Nous définissons $\hat{\theta}_n$ par

$$p_n(x_1, \dots, x_n; \hat{\theta}_n) = \sup_{\theta \in \Theta} p_n(x_1, \dots, x_n; \theta). \tag{37.76}$$

Remarque 37.29.
Nous passons sous le silence le fait que la fonction sup soit une fonction mesurable, et que par conséquent $\hat{\theta}_n$ soit bien une variable aléatoire.

La variable aléatoire $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ est l'**estimateur de maximum de vraisemblance** de θ.

37.5.5 Exemples sous forme d'exercices

Exemple 37.30.
Soit $(X_1, \dots, X_n)$ un échantillon de loi $\mathscr{B}(1, \theta)$ avec $\theta \in [0, 1]$. Trouver l'estimateur de maximum de vraisemblance de θ.

En utilisant la densité de la loi multinomiale,

$$p_n(x_1, \dots, x_n; \theta) = \prod_i p(x_i, \theta) = \prod_i \theta^{x_i} (1-\theta)^{1-x_i} \mathbb{1}_{\{0,1\}^n}(x_1, \dots, x_n). \tag{37.77}$$

En passant au logarithme, si $x_i \in \{0, 1\}$,

$$L_n(\theta) = \sum_i x_i \ln(\theta) + (1 - x_i) \ln(1 - \theta) \tag{37.78}$$

En passant à la dérivée, nous trouvons que l'extrémum est donné par

$$\theta = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{37.79}$$

△

Exemple 37.31.
Soit $(X_1, \dots, X_n)$ un échantillon de loi $\mathscr{E}(\lambda)$ avec $\theta > 0$. Trouver l'estimateur de maximum de vraisemblance de θ.

Nous avons

$$p_n(x_1, \dots, x_n) = \prod_{i=1}^{n} \theta e^{-\theta x_i}. \tag{37.80}$$

En passant au logarithme la fonction à maximiser est

$$L(\theta) = n\ln(\theta) - \theta\sum_i x_i. \tag{37.81}$$

La maximisation donne

$$\theta = \frac{n}{\sum_i x_i}, \tag{37.82}$$

c'est-à-dire

$$\hat{\theta}_n(X_1, \dots, X_n) = \frac{n}{\sum_i X_i} = \frac{1}{\bar{X}_n}. \tag{37.83}$$

Ce résultat n'est pas étonnant vu que le paramètre λ de la loi exponentielle est l'inverse de la moyenne.

△

Exemple 37.32.
Soit $(X_1, \dots, X_n)$, un échantillon de loi parente uniforme sur $[0, \theta]$ avec $\theta > 0$.

(1) Montrer que la fonction de vraisemblance est donnée par

$$L(x_1, \dots, x_n; \theta) = \frac{1}{\theta^n}\mathbb{1}_{[0,\infty[}\big(\min(x_1, \dots, x_n)\big)\mathbb{1}_{[\max(x_1,\dots,x_n),\infty[}(\theta). \tag{37.84}$$

(2) Déterminer l'estimateur du maximum de vraisemblance de θ.

Lorsque nous parlons de paramètres qui peuvent prendre un spectre continu de valeurs, il est inutile de calculer la *probabilité* parce qu'elle est nulle. Le système de maximum de vraisemblance fonctionne avec les densités. Dans notre cas, la fonction de vraisemblance est le produit des densités :

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta) \tag{37.85a}$$

$$= \prod_i \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x_i) \tag{37.85b}$$

$$= \frac{1}{\theta^n}\mathbb{1}_{[0,\theta]^n}(x_1, \dots, x_n). \tag{37.85c}$$

De cette expression nous voyons que $\min\{x_1, \dots, x_n\}$ doit être positif en même temps que le maximum doit être plus petit que θ. Cette seconde condition peut s'écrire $\mathbb{1}_{[\max\{x_1,\dots,x_n\},\infty[}(\theta)$. Au final nous avons

$$L(x_1, \dots, x_n; \theta) = \frac{1}{\theta^n}\mathbb{1}_{[0,\infty[}\big(\min(x_1, \dots, x_n)\big)\mathbb{1}_{[\max\{x_1,\dots,x_n\},\infty[}(\theta). \tag{37.86}$$

Il n'est évidemment pas possible de dériver explicitement cette expression. Par contre pour que cette fonction soit non nulle, il faut obligatoirement $\theta \geqslant \max\{x_1, \dots, x_n\}$. Par conséquent elle prend son maximum pour $\theta = \max\{x_1, \dots, x_n\}$.

La conclusion est que l'estimateur de maximum de vraisemblance de θ est $\hat{\theta}_n = \max_i\{X_i\}$. △

Exemple 37.33.
Déterminer l'estimateur de maximum de vraisemblance de la moyenne pour un modèle statistique dans laquelle la famille de probabilités est

$$(\mu_\theta) = \{\mathcal{N}(\theta, \sigma^2) \text{ tel que } \theta \in \mathbb{R}\}. \tag{37.87}$$

Nous supposons que σ est connu.

La densité de la variable aléatoire conjointe $(X_1, \dots, X_n)$ au point $(x_1, \dots, x_n)$ est le produit des densités, donc

$$\prod_{i=1}^{n} \gamma_{m,\sigma}(x_1, \dots, x_n) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp -\frac{1}{2} \left(\sum_i \left(\frac{x_i - m}{\sigma} \right)^2 \right). \tag{37.88}$$

Étant donné que le but est de maximiser, nous pouvons oublier le facteur et passer au logarithme :

$$L_0(\bar{x}) = -\frac{1}{2} \sum_{i=0}^{n} \left(\frac{x_i - m}{\sigma} \right)^2. \tag{37.89}$$

Nous pouvons également supprimer le $\frac{1}{2}$ et le $1/\sigma^2$. La fonction à maximiser devient

$$L(x_1, \dots, x_n) = -\sum_i (x_i - m)^2, \tag{37.90}$$

dont la dérivée vaut $2nm - 2\sum_i x_i$. Par conséquent nous avons un maximum avec

$$m = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{37.91}$$

△

37.5.6 Estimation d'une fonction de répartition

ThoXAEMbTI

Théorème 37.34 (Glivenko-Cantelli[789])**.**
Soient X_i des variables aléatoires indépendantes et identiquement distribuées suivant une loi dont la fonction de distribution est F (inconnue). Nous définissons l'estimateur

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \leqslant x}. \tag{37.92}$$

EqMQNRVcT

Alors

$$P\big(\lim_{n\to\infty} \|F_n - F\|_\infty = 0 \big) = 1. \tag{37.93}$$

Autrement dit, pour presque tout $\omega \in \Omega$,

$$\lim_{n\to\infty} \|F_n(\omega, .) - F\|_\infty = 0. \tag{37.94}$$

C'est-à-dire qu'il y a presque certainement convergence en probabilité.

Notons que de façon générale lorsqu'on parle d'estimateurs, partout où il y a un « n » dans une variable aléatoire, il y a une dépendance sous-entendue en ω.

PropHSHFbEq

Proposition 37.35.
Pour presque tout x, l'estimateur $F_n(x)$ est sans biais par rapport à $F(x)$:

$$E\big(F_n(x)\big) = F(x). \tag{37.95}$$

Démonstration. C'est juste un calcul :

$$E\big(F_n(x)\big) = \frac{1}{n} \sum_{i=1}^{n} E\big(\mathbb{1}_{\{X_i \leqslant x\}}\big) = \frac{1}{n} \sum_{i=1}^{n} F(x) = F(x). \tag{37.96}$$

□

37.5.7 Exemples sous forme d'exercices

Exemple 37.36.
Dans cet exercice nous construisons un estimateur biaisé qui présente un risque quadratique inférieur à un estimateur non biaisé.

Soit un modèle statistique dont la famille de lois est $\{\mathcal{N}(m,\sigma^2)\}_{\theta\in]0,\infty[}$ où m est un paramètre réel connu. En ce qui concerne la variance nous considérons

$$\hat{\sigma}_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2. \tag{37.97}$$

(1) Montrer que $\hat{\sigma}_n^2$ est un estimateur sans biais de θ.

(2) Montrer que le risque quadratique de $\hat{\sigma}_n^2$ est

$$R(\hat{\sigma}_n^2,\theta)=\frac{2\theta^2}{n}. \tag{37.98}$$

(3) Considérer la famille d'estimateurs $c\hat{\sigma}_n^2$ avec $c>0$. Déterminer la valeur de c qui minimise le risque quadratique de l'estimateur.

(1) Nous calculons

$$\begin{aligned}E(\hat{\sigma}_n^2)&=\frac{1}{n}\sum_i E\big[(X_i-m)^2\big]\\&=\frac{1}{n}\sum_i \mathrm{Var}(X_i-m)-\frac{1}{n}\sum_i E(X_i-m^2)\\&=\frac{1}{n}\sum_i \mathrm{Var}(X_i)\\&=\sigma^2.\end{aligned} \tag{37.99}$$

(2) Par la formule (37.60) nous savons que le risque d'un estimateur sans biais est donné par sa variance :

$$R(\hat{\sigma}_n,\theta)=\mathrm{Var}(|\hat{\sigma}_n^2-\theta|)=\mathrm{Var}(\hat{\sigma}_n^2). \tag{37.100}$$

Étant donné que les variables aléatoires X_i sont indépendantes et identiquement distribuées, le lemme 36.44 nous enseigne que la variance de la somme est la somme des variances. Nous avons donc à calculer

$$\begin{aligned}\mathrm{Var}(|\hat{\sigma}_n^2-\theta|)&=\mathrm{Var}(\hat{\sigma}_n^2)\\&=\frac{1}{n^2}\sum_i \mathrm{Var}\big[(X_i-m)^2\big]\\&=\frac{1}{n^2}\sum_i E\big[(X_i-m)^4\big]-E\big[(X_i-m)^2\big]^2.\end{aligned} \tag{37.101}$$

Ces espérances ne sont pas très compliquées à calculer en utilisant la fonction caractéristique donnée par la proposition 36.157 :

$$\Phi_{X-m}(t)=E\big(e^{it(X-m)}\big)=e^{-itm}E(e^{itX})=e^{-itm}\Phi_X(t)=\exp\left(-\frac{\sigma^2t^2}{2}\right). \tag{37.102}$$

Nous avons $\Phi^{(4)}(0)=3\sigma^4$ et $\Phi''(0)=-\sigma^2$. Par conséquent

$$\mathrm{Var}(\hat{\sigma}_n^2)=\frac{1}{n^2}\sum_i 2\theta^2=\frac{2\theta^2}{n}. \tag{37.103}$$

Notez ici que $\theta=\sigma^2$.

(3) En tenant compte du fait que $\text{Var}(\hat{\sigma}_n^2) = \frac{2\theta^2}{n}$ et $E(\hat{\sigma}_n^2) = \theta$, nous avons

$$E\big((c\hat{\sigma}_n^2 - \theta)^2\big) = \text{Var}(c\hat{\sigma}_n^2 - \theta) + E(c\hat{\sigma}_n^2 - \theta)^2 \tag{37.104a}$$
$$= 2\frac{c^2\theta^2}{n} + (c-1)^2\theta^2. \tag{37.104b}$$

La dérivée (par rapport à c) de cela s'annule pour $c_0 = \frac{n}{n+2}$. Notons que nous n'avons pas tout à fait démontré que cela est bien un minimum. Calculons cependant le risque quadratique de notre estimateur pour cette valeur de c. Pour cela nous reportons $c = c_0$ dans l'expression (37.104b) :

$$E\big(\frac{n}{n+2}\hat{\sigma}_n^2\big) = \frac{2\theta^2}{n+2}. \tag{37.105}$$

Cela est effectivement plus petit que $R(\hat{\sigma}_n^2, \theta)$.

Nous avons ainsi construit un estimateur biaisé qui a un risque quadratique plus petit que l'estimateur non biaisé. △

Exemple 37.37.
Nous considérons la famille de probabilités $\mu_\theta = \mathcal{N}(\theta)$ où $\theta = (m, \sigma^2) \in \mathbb{R} \times]0, \infty[$. Déterminer l'estimateur de maximum de vraisemblance du paramètre $\theta = (m, \sigma^2)$.

Pour chaque observation x_i nous avons une densité gaussienne. Le produit donne

$$p\big(x_1, \dots, x_n; (m, \sigma^2)\big) = \frac{1}{\sigma^2(2\pi)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_i (x_i - m)^2\right]. \tag{37.106}$$

En passant au logarithme et en supprimant des facteurs inutiles à la maximisation,

$$L(m, \sigma) = -n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_i (x_i - m)^2. \tag{37.107}$$

L'annulation de la dérivée par rapport à m donne immédiatement $m = \frac{1}{n}\sum_i x_i$. L'annulation de la dérivée par rapport à σ donne

$$-n\sigma^2 + \sum_i (x_i - m)^2 = 0 \tag{37.108}$$

et donc

$$\sigma^2 = \frac{1}{n}\sum_i (x_i - \bar{x}_n). \tag{37.109}$$

L'estimateur de maximum de vraisemblance du couple $\theta = (m, \sigma^2)$ est donc

$$\hat{\theta}_n = \left(\frac{1}{n}\sum_i X_i, \frac{1}{n}\sum_i (X_i - \bar{X}_n)^2\right). \tag{37.110}$$

△

37.5.8 Espérance et variance d'un estimateur

Soit $T_n = T(X_1, \dots, X_n)$ un estimateur du paramètre θ dans un modèle d'échantillonnage. Les moyennes et variances de l'estimateur sont les variables aléatoires

$$m_{\theta,n} = E_\theta(T_n) = \int_{\mathbb{R}^n} T_n(x_1, \dots, x_n) d\mu_\theta^{\otimes n}(x_1, \dots, x_n), \tag{37.111a}$$
$$\text{Var}_\theta(T_n) = \int_{\mathbb{R}^n} [T_n(x_1, \dots, x_n) - m_{\theta,n}]^2 d\mu_\theta^{\otimes n}(x_1, \dots, x_n), \tag{37.111b}$$

Lemme 37.38.
Si l'estimateur T_n satisfait

$$\lim_{n\to\infty} E_\theta(T_n) = \theta \tag{37.112a}$$
$$\lim_{n\to\infty} \operatorname{Var}_\theta(T_n) = 0, \tag{37.112b}$$

alors il est convergent.

Démonstration. Nous utilisons l'inégalité de Markov (théorème 37.18) et nous introduisons l'espérance de l'estimateur :

$$P\big(|T_n(X_1,\dots,X_n)-\theta|>\epsilon\big) \leqslant \frac{1}{\epsilon}E\big(T_n(X_1,\dots,X_n)-\theta\big) \tag{37.113a}$$
$$\leqslant \frac{1}{\epsilon}E\big(|T_n-m_{n,\theta}|\big)+\frac{1}{\epsilon}E\big(|m_{n,\theta}-\theta|\big) \tag{37.113b}$$
$$\tag{37.113c}$$

Le second terme est l'espérance d'une constante. Nous majorons le premier terme en utilisant le fait que $\|.\|_1 \leqslant \|.\|_2$ (voir la remarque 27.34 après l'inégalité de Hölder) :

$$P\big(|T_n(X_1,\dots,X_n)-\theta|>\epsilon\big) \leqslant \frac{1}{\epsilon}E\big(|T_n-m_{n,\theta}|^2\big)^{1/2}+\frac{1}{\epsilon}|E_\theta(T_n)-\theta| \tag{37.114a}$$
$$= \frac{1}{\epsilon}\operatorname{Var}(T_n)^{1/2}+\frac{1}{\epsilon}|E_\theta(T_n)-\theta|. \tag{37.114b}$$

Les deux termes tendent séparément vers zéro par hypothèse. Nous avons par conséquent la convergence en probabilité $T_n \to \theta$. □

37.6 Estimation par intervalle de confiance

Nous voudrions estimer la proportion d'individus dans une population ayant un certain caractère déterminé par une variable booléenne : chaque individu a ou non le caractère étudié. L'échantillon sera donc une suite de 0 et de 1.

Pour tout $i \in \{1,\dots,n\}$ nous notons

$$x_i = \begin{cases} 1 & \text{si le } i\text{ème individu a le caractère} \\ 0 & \text{sinon.} \end{cases} \tag{37.115}$$

et $\bar{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$. Notre modèle statistique sera

$$\mathcal{S} = \Big[(\Omega,\mathcal{F},P),(X_\theta),\mathscr{B}(1,\theta)\Big] \tag{37.116}$$

où Ω est l'ensemble des individus étudiés, P est la manière de choisir les individus lors du sondage (essentiellement c'est une loi uniforme) et X_θ est la variable aléatoire

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \text{ a le caractère} \\ 0 & \text{sinon} \end{cases} \tag{37.117}$$

Cela est une variable aléatoire de distribution $\mathscr{B}(1,p)$ où p est inconnu. Ici, $\Theta = [0,1]$ est l'ensemble des p possibles.

Remarque 37.39.
Nous supposons que Ω est la population entière et que la variable aléatoire est l'opinion de la personne ω. En cela, nous considérons que le tirage de l'échantillon est sans remise. Le fait que nous modélisions par une variable aléatoire de Bernoulli signifie que nous considérons l'approximation dans laquelle la population globale est grande.

Nous supposons que nous ayons un échantillon $(X_1, \dots, X_n)$ dont nous avons observé une réalisation $(x_1, \dots, x_n)$ de fréquence $\bar{x}_n$. Nous voudrions déterminer un intervalle dans lequel $\bar{X}_n$ a de fortes chances de se trouver. Plus précisément nous considérons un petit α et nous cherchons ϵ tel que

$$P\big(p \in [\bar{X}_n - \epsilon, \bar{X}_n + \epsilon]\big) = 1 - \alpha. \tag{37.118}$$

Typiquement, $\alpha = 5\%$. Le nombre α est le **niveau de confiance** que nous nous fixons à priori.

Si nous trouvons un intervalle I tel que $P(p \in I) = 1-\alpha$, nous disons que l'intervalle est **exact**, si nous avons $P(p \in I) \geqslant 1 - \alpha$, nous disons que l'intervalle est **par excès**.

Il y a deux points de départ pour trouver l'intervalle. Le plus simple est d'utiliser le théorème central limite et considérer

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \xrightarrow{\mathscr{L}} \mathscr{N}(0,1). \tag{37.119}$$

La seconde est d'utiliser la loi exacte : $n\bar{X}_n = \sum_i X_i \sim \mathscr{B}(n,p)$.

Bien entendu la seconde donne lieu à des calculs plus compliqués.

Remarque 37.40.
Dans certaines vraies vies (par exemple en médecine), la taille des échantillons est très réduite. Dans ce cas le théorème central limite n'a aucun sens et les calculs exact s'imposent.

De plus dans de nombreux cas de la vraie vie, nous avons un ordinateur à disposition pour calculer avec la loi exacte. L'utilisation du théorème central limite dans le but de produire un intervalle de confiance semble donc de plus en plus être une survivance du passé.

Dans la suite, nous allons supposer que n est suffisamment grand pour justifier l'approximation normale du théorème central limite 36.135. Si Z est une variable aléatoire normale centrée réduite, notre premier essai est de faire

$$\begin{aligned}
1 - \alpha &= P\big(p \in [\bar{X}_n - \epsilon, \bar{X}_n + \epsilon]\big) && \text{(37.120a)}\\
&= P\left(\frac{-\epsilon\sqrt{n}}{\sqrt{p(1-p)}} \leqslant \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \leqslant \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) && \text{(37.120b)}\\
&\simeq P\left(-\frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} \leqslant Z \leqslant \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) && \text{(37.120c)}\\
&= 2P\left(Z \leqslant \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1. && \text{(37.120d)}
\end{aligned}$$

La dernière ligne utilise la symétrie de la distribution $\mathscr{N}(0,1)$. Le nombre ϵ que nous cherchons vérifie donc

$$P\left(Z \leqslant \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) = 1 - \frac{\alpha}{2}. \tag{37.121}$$

De nos jours, les ordinateurs donnent la loi de répartition inverse des normales. Cela nous fournit un nombre t_α tel que

$$\frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} = t_\alpha \tag{37.122}$$

où t_α est le nombre tel que $P(Z \leqslant t_\alpha) = 1 - \alpha/2$.

Conclusions de ce premier essai :

(1) Le problème est que nous ne pouvons pas déduire ϵ de l'équation (37.122) parce que p est inconnu.

(2) Cela ruine notre premier essai et nous demande de trouver mieux.

(3) L'astuce est évidemment de remplacer p par $\bar{X}_n$, mais il faut le justifier.

Méthode Slutsky Le point de départ du premier essai infructueux était la convergence

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \xrightarrow{\mathscr{L}} N \sim \mathscr{N}(0,1) \tag{37.123}$$

donnée par le théorème central limite 36.135. Ce que nous voudrions en réalité est la convergence

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathscr{L}} N \tag{37.124}$$

La loi des grands nombres nous donne

$$\bar{X}_n(1-\bar{X}_n) \xrightarrow{p.s.} p(1-p). \tag{37.125}$$

Par conséquent le lemme de Slutsky implique la convergence en loi du couple :

$$\left(\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}}, \sqrt{\bar{X}_n(1-\bar{X}_n)}\right) \xrightarrow{\mathscr{L}} \left(N, \sqrt{p(1-p)}\right). \tag{37.126}$$

À ce point des opérations nous pouvons utiliser la proposition 36.120 au couple avec la fonction

$$f(x,y) = \frac{x}{y}\sqrt{p(1-p)} \tag{37.127}$$

dont la probabilité d'être continue est 1 ($y = 0$ est de mesure nulle dans $\mathbb{R}^2$). La conclusion du théorème est que

$$\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathscr{L}} \mathscr{N}(0,1). \tag{37.128}$$

C'est à partir de là que nous pouvons construire notre intervalle de confiance :

$$1-\alpha = P(\bar{X}_n - \epsilon \leqslant p \leqslant \bar{X}_n + \epsilon) \tag{37.129a}$$

$$= P\left(\frac{\sqrt{n}\epsilon}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \geqslant \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \geqslant \frac{-\sqrt{n}\epsilon}{\sqrt{\bar{X}_n(1-\bar{X}_n)}}\right) \tag{37.129b}$$

$$\simeq P\left(\frac{\sqrt{n}\epsilon}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \geqslant N \geqslant \frac{-\sqrt{n}\epsilon}{\sqrt{\bar{X}_n(1-\bar{X}_n)}}\right). \tag{37.129c}$$

Nous cherchons maintenant dans les tables le ξ qui fait

$$P(-\xi \leqslant N \leqslant \xi) = 1-\alpha \tag{37.130}$$

puis nous cherchons ϵ de telle sorte à avoir

$$\frac{\sqrt{n}\epsilon}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} = \xi. \tag{37.131}$$

Dans cette équation tout est connu à part le ϵ qui se découvre.

Méthode piétonne Nous remarquons que l'événement

$$-\frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} \leqslant \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \leqslant \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} \tag{37.132}$$

est le même que l'événement

$$\left|\sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}}\right| \leqslant t_\alpha. \tag{37.133}$$

Vu que t_α est positif, cela est encore le même événement que

$$n\frac{(\bar{X}_n - p)^2}{p(1-p)} \leqslant t_\alpha^2 \tag{37.134}$$

ou encore

$$p^2(n + t_\alpha^2) - p(2n\bar{X}_n + t_\alpha^2) + n\bar{X}_n^2 \leqslant 0. \tag{37.135}$$

Les racines du polynôme du membre de gauche sont

$$p_\pm = \frac{2n\bar{X}_n + t_\alpha^2 \pm \sqrt{(2n\bar{X}_n + t_\alpha)^2 - 4(n + t_\alpha^2)n\bar{X}_n^2}}{2(n + t_\alpha^2)}. \tag{37.136}$$

Le but étant d'effectuer une limite $n \to \infty$, nous factorisons d'abord n. Après simplification

$$p_\pm = \frac{\bar{X}_n + \frac{t_\alpha^2}{2n} \pm \sqrt{\frac{t_\alpha^2}{4n^2} + \frac{t_\alpha \bar{X}_n(1-\bar{X}_n)}{n}}}{1 + \frac{t_\alpha^2}{2n}}. \tag{37.137}$$

Étant donné que nous considérons que n est grand, nous allons négliger les termes en $\frac{1}{n}$ en faisant attention à ce que le terme en $\frac{1}{n}$ sous la racine est en réalité $1/\sqrt{n}$ et ne doit pas être négligé. Nous trouvons, à cette approximation, que

$$p \in \Big[\bar{X}_n - t_\alpha\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + t_\alpha\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\Big] \tag{37.138}$$

avec une probabilité $1 - \alpha$.

37.6.1 Région de confiance

Soit un n-échantillon $(X_1, \dots, X_n)$ de loi parente μ_θ. Nous supposons $\Theta \subset \mathbb{R}$ avec Θ ouvert. Soit $\alpha \in [0, 1]$ un intervalle de confiance et une application mesurable

$$\begin{aligned} \Lambda\colon \mathbb{R}^n &\to \mathcal{B}\text{or}(\Theta) \\ (x_1, \dots, x_n) &\mapsto \Lambda(x_1, \dots, x_n). \end{aligned} \tag{37.139}$$

On appelle **région de confiance exacte** au niveau de confiance $1 - \alpha$ une région aléatoire $\Lambda(x_1, \dots, x_n)$ telle que

$$P\big(\theta \in \Lambda(x_1, \dots, x_n)\big) = 1 - \alpha. \tag{37.140}$$

Si $d = 1$, la région $\Lambda(x_1, \dots, x_n)$ est un intervalle.

37.6.2 Fonction pivotale

Soit $\hat{\theta}_n$, un estimateur de θ. Une fonction v sur $\Theta \times \Theta$ est **pivotale** pour θ si la loi de la variable aléatoire $v(\hat{\theta}_n, \theta)$ ne dépend pas de θ. Elle est **asymptotiquement pivotale** si

$$v(\hat{\theta}_n, \theta) \xrightarrow{\mathscr{L}} \xi \tag{37.141}$$

où ξ est une variable aléatoire indépendante de θ.

En pratique, nous essayons de faire apparaitre une variable aléatoire de loi connue qui ne dépend pas du paramètre que l'on recherche. Si la variance est connue et si l'échantillon est grand, le théorème central limite nous permet d'introduire une loi normale centrée réduite.

Exemple 37.41.
Soit $X_1, \dots, X_n$ des variables aléatoires de loi parente $\mathcal{N}(m, \sigma^2)$. Une fonction asymptotiquement pivotale pour m est

$$v(z_1, z_2) = \frac{z_1 - z_2}{\alpha/\sqrt{n}} \tag{37.142}$$

parce que la variable aléatoire

$$v(\bar{X}_n, m) = \frac{\bar{X}_n - m}{\alpha/\sqrt{n}} \tag{37.143}$$

tend vers $\mathcal{N}(0,1)$ qui ne dépend pas de m. △

Exemple 37.42.
Si (X_n) est une suite de variables aléatoires indépendantes et identiquement distribuées de moyenne m et d'écart type σ que nous supposons inconnus. Le fonction suivante est asymptotiquement pivotale pour m :

$$v(\bar{X}_n, m) = \frac{\bar{X}_n - m}{\sigma/\sqrt{n}}. \tag{37.144}$$

△

Soit $(X_1, \dots, X_n)$ un échantillon de loi $\mathcal{N}(m, \sigma^2)$ avec σ^2 connu. Nous cherchons un intervalle de confiance $1-\alpha$ pour m. Pour cela nous allons utiliser une fonction asymptotiquement pivotale, à savoir

$$\frac{\bar{X}_n - m}{\sigma^2/\sqrt{n}} \sim \mathcal{N}(0,1). \tag{37.145}$$

Nous devrions chercher des valeurs z_+ et z_- telles que

$$P\left(z_- \leqslant \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \leqslant z_+\right) = 1 - \alpha. \tag{37.146}$$

Pour des raisons de symétries (de la courbe gaussienne), nous allons chercher un intervalle symétrique : $z_- = -z_+$. Le nombre à chercher est donc le z_α tel que

$$P\big(|Z| \leqslant z_\alpha\big) = 1 - \alpha. \tag{37.147}$$

Si nous demandons $\alpha = 5\%$, la réponse est $z_\alpha = 1.96$, c'est-à-dire que

$$P\left(-1.96 \leqslant \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \leqslant 1.96\right) = 0.95. \tag{37.148}$$

Nous avons donc

$$P\left(m \in \big[\bar{X}_n - \frac{1.96\sigma}{\sqrt{n}}, \bar{X}_n + \frac{1.96\sigma}{\sqrt{n}}\big]\right) = 0.95. \tag{37.149}$$

Supposons maintenant que nous avons observé 100 valeurs numériques avec $\bar{x}_n = 12$ et $\sigma = 1$. La réalisation de l'intervalle de confiance pour m au niveau de confiance 0.95 est :

$$\big[12 - 0.196, 12 + 0.196\big]. \tag{37.150}$$

Cet intervalle est à interpréter de la façon suivante : si nous recommençons un grand nombre de fois le sondage, la moyenne tombera 95% des fois dans l'intervalle ainsi calculé. Mais il faut bien comprendre que la probabilité

$$P\left(m \in \big[12 - 0.196, 12 + 0.196\big]\right) \tag{37.151}$$

vaut zéro ou un.

Exemple 37.43.
Soit un échantillon $(X_1, \dots, X_n)$ de loi parente $\mathcal{N}(m, \sigma^2)$ où m et σ^2 sont inconnus. Déterminer un intervalle de confiance exact symétrique au niveau de confiance $1-\alpha$.

Déterminer un intervalle de confiance pour σ^2.

Nous savons que la moyenne empirique est un estimateur de la moyenne :

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{37.152}$$

Nous cherchons un intervalle du type $I = [\bar{X}_n - \epsilon, \bar{X}_n + \epsilon]$ pour lequel $P(m \in I) = 1 - \alpha$. Nous savons que la variable aléatoire

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \tag{37.153}$$

suit une loi $\mathscr{N}(0,1)$, mais la variance est inconnue. La subtilité à savoir est que la variable aléatoire

$$Z = \frac{\bar{X}_n - m}{S_n/\sqrt{\sigma}} \tag{37.154}$$

où $S_n^2 = \sum_i (X_i - \bar{X}_n)^2/(n-1)$ suit une loi de Student à n degrés de liberté $\mathscr{T}(n-1)$ en vertu du théorème de Cochran 37.15. Comme il est usuel de le faire, nous inversons l'intervalle :

$$1 - \alpha = P\left(-\epsilon \leqslant \bar{X}_n - m \leqslant \epsilon\right) \tag{37.155a}$$

$$= P\left(-\frac{\epsilon\sqrt{n}}{S_n} \leqslant Z \leqslant \frac{\epsilon\sqrt{n}}{S_n}\right). \tag{37.155b}$$

Les valeurs se trouvent dans des tables ; par exemple pour $n = 10$ et $\alpha = 5\%$ nous trouvons

$$\frac{\epsilon\sqrt{n}}{S_n} = 2.262. \tag{37.156}$$

Plus généralement nous notons $t_{n-1,1-\frac{\alpha}{2}}$ le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathscr{T}(n-1)$, c'est-à-dire le nombre tel que

$$P(Z \leqslant t_{n-1,1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \tag{37.157}$$

si $Z \sim \mathscr{T}(n-1)$. L'intervalle de confiance est alors donné par

$$I = \left[\bar{X}_n - \frac{t_{n-1,1-\frac{\alpha}{2}} S_n}{\sqrt{n}}, \bar{X}_n + \frac{t_{n-1,1-\frac{\alpha}{2}} S_n}{\sqrt{n}}\right]. \tag{37.158}$$

Cela est un intervalle exact pour m au niveau de confiance $1 - \alpha$.

Nous pouvons aussi trouver un intervalle asymptotique en utilisant le théorème central limite :

$$\frac{\bar{X}_n - m}{S_n/\sqrt{n}} \xrightarrow{\mathscr{L}} T \tag{37.159}$$

avec $T \sim \mathscr{N}(0,1)$.

Rappel : dire que I_n est un **intervalle de confiance asymptotique** signifie que

$$\lim_{n\to\infty} P(m \in I_n) = 1 - \alpha. \tag{37.160}$$

En ce qui concerne la variance σ^2, l'intervalle de confiance se construit en utilisant la partie (2) du théorème de Cochran 37.15. Nous introduisons la variable aléatoire pivot

$$Z = (n-1)\frac{S_n^2}{\sigma^2} \tag{37.161}$$

qui suit une loi $\chi^2(n-1)$. Cette loi n'étant pas symétrique (voir figure 36.1), nous n'allons pas chercher un intervalle de confiance symétrique. Nous cherchons c_1 et c_2 tels que

$$\begin{cases} P\big(\sigma^2 \in [c_1, c_2]\big) = 1 - \alpha & \text{(37.162a)} \\ P\big(\sigma^2 \in [0, c_1]\big) = \dfrac{\alpha}{2} & \text{(37.162b)} \\ P\big(\sigma^2 \in [c_2, \infty[\big) = \dfrac{\alpha}{2} & \text{(37.162c)} \end{cases}$$

La situation est représentée à la figure 37.3.

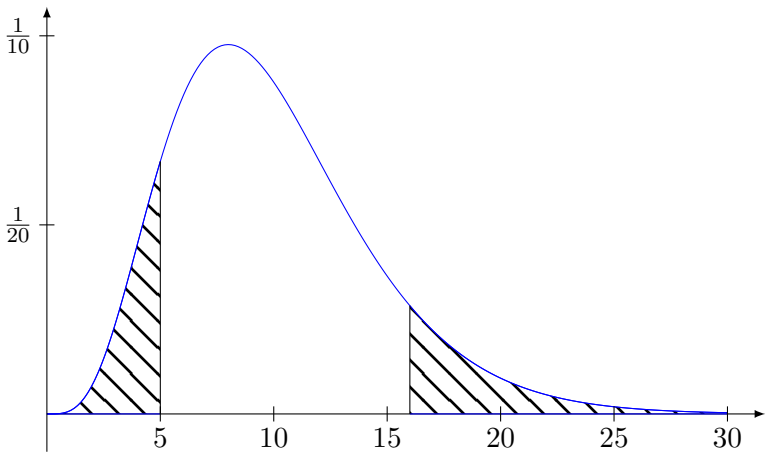


FIGURE 37.3: L'intervalle de confiance pour une χ^2. LabelFigChiSquaresQuantile

La construction des nombres c_1 et c_2 passe par la relation

$$P(c_1 \leqslant \sigma^2 \leqslant c_2) = P\left(\frac{(n-1)S_n^2}{c_2} \leqslant Z \leqslant \frac{(n-1)S_n^2}{c_1}\right). \tag{37.163}$$

Nous notons $t_{n-1,\frac{\alpha}{2}}$ et $t_{n-1,1-\frac{\alpha}{2}}$ les quantiles donnés sur la figure 37.3, c'est-à-dire

$$P(Z \leqslant t_{n-1,\frac{\alpha}{2}}) = \frac{\alpha}{2} P(Z \geqslant t_{n-1,1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}. \tag{37.164a}$$

Ce que nous obtenons est

$$\frac{(n-1)S_n^2}{c_2} = t_{n-1,\frac{\alpha}{2}} \tag{37.165a}$$

$$\frac{(n-1)S_n^2}{c_1} = t_{n-1,1-\frac{\alpha}{2}}, \tag{37.165b}$$

et par conséquent l'intervalle de confiance pour σ^2 est

$$I = \left[\frac{(n-1)S_n^2}{t_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S_n^2}{t_{n-1,1-\frac{\alpha}{2}}}\right] \tag{37.166}$$

avec $P(\sigma^2 \in I) = 95\%$.

Remarque 37.44.
Il n'est pas clair à priori que la longueur de l'intervalle I décroisse avec n parce qu'il y a n dans les t au numérateur.

△

37.6.3 Sondage de proportion

Une utilisation classique des statistiques est d'interpréter une proportion donnée par un sondage. Nous considérons une élection avec deux candidats A et B. Nous avons interrogés $n = 2500$ personnes et nous avons obtenu 51% pour le candidat A et 49% pour le candidat B. Que peut on dire ?

La modélisation de cette situation est que nous avons des variables aléatoires $X_i \sim \mathscr{B}(p_A)$ et que nous en avons observés n avec une moyenne

$$\bar{x}_n = 0.51. \tag{37.167}$$

La loi de $\bar{X}_n$ est une binomiale. Sa densité n'est pas symétrique, mais si n est grand, elle le devient. Nous cherchons un intervalle

$$I = [\bar{X}_n - \epsilon, \bar{X}_n + \epsilon] \tag{37.168}$$

tel que $P(p_A \in I) = 1 - \alpha$. Pour cela nous considérons le fait que $n = 2500$ est grand et nous utilisons la limite de la proposition 37.10 :

$$Z_n = \sqrt{n}\frac{\bar{X}_n - p_A}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathscr{L}} Z \sim \mathscr{N}(0,1). \tag{37.169}$$

La variable aléatoire Z_n est asymptotiquement pivotale et normale centrée réduite. Nous cherchons donc un intervalle symétrique pour $\bar{X}_n - p_A$:

$$1 - \alpha = P(-\epsilon \leqslant \bar{X}_n - p_A \leqslant \epsilon), \tag{37.170}$$

c'est-à-dire, si n est grand,

$$1 - \alpha = P\left(-\epsilon\frac{\sqrt{n}}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \leqslant Z_n \leqslant \epsilon\frac{\sqrt{n}}{\sqrt{\bar{X}_n(1-\bar{X}_n)}}\right) \tag{37.171}$$

où Z_n est une normale centrée réduite. Nous trouvons ainsi, via les tables que

$$\frac{\epsilon\sqrt{n}}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} = 1.96 \tag{37.172}$$

si nous voulons un intervalle à 5%. Par conséquent nous avons $\epsilon = 1.96\frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}}$ et l'intervalle de confiance est

$$I_C = \left[\bar{X}_n - \frac{1.96\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}}, \bar{X}_n + \frac{2.96\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}}\right]. \tag{37.173}$$

La propriété de cet intervalle est que

$$\lim_{n\to\infty} P(p_A \in I_c) = 1 - \alpha. \tag{37.174}$$

Remarque 37.45.
À quel moment avons nous fait une hypothèse sur la taille de la population globale ? En modélisant les sondés par des variables de Bernoulli et leur somme par une binomiale, nous supposons que le sondage est *avec remise*, sinon, elles ne seraient pas indépendantes. En supposant les sondés indépendants, nous avons donc fait comme si la population totale était infinie.

37.7 Estimer une densité lorsqu'on ne sait rien

Nous supposons avoir une série d'observations issues d'un processus complexe dont nous n'avons aucune idée de la loi parente, et nous voudrions nous faire une idée de la densité de cette loi inconnue.

Nous observons une suite de réalisations que nous modélisons comme étant des variables aléatoires $(X_1, \dots, X_n)$ de loi parente (inconnue) μ. Notre but est de trouver un estimateur $\hat{\mu}$ de μ. Par simplicité nous allons nous restreindre aux lois admettant une densité par rapport à la mesure de Lebesgue. C'est-à-dire que nous allons estimer μ par une suite de lois $\hat{\mu}_n$ qui sont toutes des lois acceptant une densité par rapport à la mesure de Lebesgue.

37.7.1 Distance entre des mesures

Si ν_1 et ν_2 sont deux mesures de densité sur $\mathbb{R}$, la **distance** entre ν_1 et ν_2 est définie par

$$d(\nu_1,\nu_2)=\sup_{A\in\mathcal{B}\text{or}(\mathbb{R})}\big|\nu_1(A)-\nu_2(A)\big| \tag{37.175}$$

où $\mathcal{B}\text{or}(\mathbb{R})$ désigne l'ensemble des boréliens de $\mathbb{R}$.

Théorème 37.46 (de Scheffé[790])**.**
Si f_1 et f_2 sont les densités de ν_1 et ν_2 par rapport à la mesure de Lebesgue, alors

$$d(\nu_1,\nu_2)=\int_{\mathbb{R}}(f_1-f_2)_+=\frac{1}{2}\int_{\mathbb{R}}|f_1-f_2|=\frac{1}{2}\|f_1-f_2\|_1 \tag{37.176}$$

où f_+ est la partie positive de f (pour la décomposition $f(x)=f_+(x)-f_-(x)$).

Démonstration. La dernière égalité est simplement une notation usuelle ; nous devons seulement prouver les deux premières. Pour la première nous commençons par prouver que le borélien réalisant le supremum est

$$B=\{x\in\mathbb{R}\text{ tel que }f_1(x)\geqslant f_2(x)\}. \tag{37.177}$$

En effet si A est un borélien nous avons

$$\nu_1(A)-\nu_2(A)=\int_A f_1-f_2\leqslant\int_{A\cap B}f_1-f_2\leqslant\int_B f_1-f_2=\int_B(f_1-f_2)_+=\int_{\mathbb{R}}(f_1-f_2)_+ \tag{37.178}$$

Justifications :

— f_1-f_2 négative sur $A\cap\complement B$.

— Vu que $f_1-f_2\geqslant 0$ sur B, l'intégrale augmente si on augmente le domaine.

— Sur B nous avons $f_1-f_2=(f_1-f_2)_+$.

Donc pour tout borélien A nous avons

$$d(\nu_1,\nu_2)\leqslant\int_{\mathbb{R}}(f_1-f_2)_+. \tag{37.179}$$

Mais pour $A=B$ nous avons égalité :

$$\nu_1(B)-\nu_2(B)=\int_B f_1-f_2=\int_B(f_1-f_2)_+=\int_{\mathbb{R}}(f_1-f_2). \tag{37.180}$$

Pour la seconde égalité nous savons que f_1 et f_2 s'intègrent toutes deux à 1 (parce que ce sont des densités de probabilité), donc

$$\int_{\mathbb{R}}f_1-f_2=0. \tag{37.181}$$

En particulier nous avons

$$\int_{\mathbb{R}}(f_1-f_2)_+=\int_{\mathbb{R}}(f_1-f_2)_-, \tag{37.182}$$

ce qui donne bien

$$\int_{\mathbb{R}}(f_1-f_2)_+=\frac{1}{2}\int_{\mathbb{R}}|f_1-f_2|. \tag{37.183}$$

□

37.7.2 Estimateur par fenêtres glissantes

Nous considérons les estimations suivantes de la fonction de répartition :

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leqslant x\}}, \tag{37.184}$$

et un nombre h_n qui sera la taille de la fenêtre glissante. Nous avons en tête de faire $\lim_{n\to\infty} h_n = 0$. Nous considérons ceci comme estimateur de la densité inconnue f des variables aléatoires X_i :

$$\hat{f}_n(x) = \frac{F_n(x+h_n) - F_n(x-h_n)}{2h_n}. \tag{37.185}$$

L'idée sous-jacente est de prendre la dérivée de la fonction de répartition comme densité.

Lemme 37.47 ([790]).
Pour tout $h_n > 0$, l'estimateur $\hat{f}_n$ est une densité de probabilité.

Démonstration. D'abord $\hat{f}_n$ est bien à valeurs positives ou nulle. Ensuite devons parler de son intégrale. Pour le numérateur nous avons

$$F_n(x+h_n) - F_n(x-h_n) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \in B(x,h_n)}. \tag{37.186}$$

En réalité cette égalité est valable seulement presque partout parce qu'elle n'est pas valable en $x = x \pm h_n$, mais cela ne va pas nous ennuyer dans la mesure où nous avons dans l'idée d'intégrer cela sur $\mathbb{R}$. Avant de nous lancer dans l'intégrale nous remarquons que $X_i \in B(x, h_n)$ est la même chose que $x \in B(X_i, h_n)$, c'est-à-dire que

$$\{X_i \in B(x,h_n)\} = \{x \in B(X_i, h_n)\}. \tag{37.187}$$

Donc

$$\int_{\mathbb{R}} \hat{f}_n = \frac{1}{2h_n}\frac{1}{n}\sum_{i=1}^{n} \underbrace{\int_{\mathbb{R}} \mathbb{1}_{B(X_i,h_n)}}_{2h_n} = \frac{1}{2nh_n}\sum_{i=1}^{n} 2h_n = 1. \tag{37.188}$$

□

LemTZopXDd

Lemme 37.48 ([790]).
L'estimateur $\hat{f}_n$ est déjà pas mal parce que

$$\lim_{h_n\to 0} E\big(\hat{f}_n(x)\big) = f(x) \tag{37.189}$$

pour presque tout $x \in \mathbb{R}$.

Démonstration. Nous commençons par nous rappeler le fait que $F_n(x)$ est un estimateur sans biais de $F(x)$ (proposition 37.35). Donc

$$E\big(\hat{f}_n(x)\big) = \frac{F(x+h_n) - F(x-h_n)}{2h_n}. \tag{37.190}$$

EqJEjrfFd

Nous devons prendre la limite de cela lorsque $h_n \to 0$, c'est-à-dire considérer la dérivée de F. Attention : rien ne dit que F soit dérivable, si ce n'est la proposition 14.328 qui indique qu'elle est dérivable presque partout avec f comme dérivée.

La limite $h_n \to 0$ dans (37.190) nous donne donc bien presque partout

$$\lim_{h_n\to 0} E\big(\hat{f}_n(x)\big) = f(x). \tag{37.191}$$

□

Proposition 37.49 ([790]).
Si la suite (h_n) est telle que $h_n \to 0$ et $nh_n \to \infty$, alors pour presque tout $x \in \mathbb{R}$ nous avons les convergences

$$\hat{f}_n(x) \xrightarrow{L^2(P)} f(x) \tag{37.192}$$

et

$$\hat{f}_n(x) \xrightarrow{P} f(x). \tag{37.193}$$

Démonstration. Il faut d'abord comprendre ce que signifie la convergence $L^2(P)$ pour presque tout x. Pour cela il faut comprendre que $\hat{f}_n(x)$ est en soi une variable aléatoire et est en réalité une fonction $\omega \mapsto \hat{f}_n(x, \omega)$. Ce que nous allons montrer est que pour presque tout x (maintenant fixé), cette variable aléatoire converge vers une constante (par rapport à ω) et que cette constante est $f(x)$.

La convergence $X_n \xrightarrow{L^2(P)} X$ signifie $E\big(|X_n - X|^2\big) \to 0$, c'est-à-dire

$$\int_\Omega \big|X_n(\omega) - X(\omega)\big|^2 dP(\omega) \to 0. \tag{37.194}$$

En faisant une décomposition biais-variance nous devons donc étudier

$$E\Big[\big(\hat{f}_n(x) - f(x)\big)^2\Big] = E\big[\hat{f}_n(x) - f(x)\big]^2 + \operatorname{Var}\big(\hat{f}_n(x) - f(x)\big) \tag{37.195}$$

Ici $f(x)$ doit être vue comme la variable aléatoire constante sur Ω. Par le lemme 37.48 et la proposition 36.47 le terme de biais converge vers zéro lorsque $n \to \infty$.

Pour traiter le terme de biais, nous savons déjà que

$$2nh_n\hat{f}_n(x) = \sum_{i=1}^{n} \mathbb{1}_{\{X_i \in B(x,h_n)\}}, \tag{37.196}$$

où le membre de droite (et donc aussi celui de gauche) est une variable aléatoire binomiale comptant le nombre de succès de l'expérience $X_i \in B(x, h_n)$ en n essais. Nous notons $p_{x,n} = P\big(X_i \in B(x, h_n)\big)$. Si μ est la loi parente des X_i, alors

$$p_{n,x} = P\big(X_i \in B(x, h_n)\big) = \mu\big(B(x, h_n)\big) = F(x + h_n) - F(x - h_n) \tag{37.197}$$

où F est la fonction de répartition (parente) des X_i.

Alors la variance de ladite binomiale est donnée par (36.438), c'est-à-dire $np_{x,n}(1 - p_{x,n})$. Nous avons alors

$$\operatorname{Var}\big(2nh_n\hat{f}_n(x)\big) = np_{x,n}(1 - p_{x,n}) \tag{37.198}$$

et

$$\operatorname{Var}\big(\hat{f}_n(x)\big) = \frac{1}{4n^2h_n^2} np_{x,n}(1 - p_{x,n}). \tag{37.199}$$

Nous pouvons faire la majoration $t(1 - t) \leqslant t$ qui est valable pour tout t et écrire

$$\operatorname{Var}(\hat{f}_n(x)) \leqslant \frac{1}{4nh_n} \frac{p_{x,n}}{h_n}. \tag{37.200}$$

Le premier facteur tend vers zéro parce que nous avons supposé que $nh_n \to \infty$. Pour le second facteur, il faut remarquer que l'expression (37.197) nous donne presque partout

$$\lim_{h_n \to 0} \frac{p_{n,x}}{h_n} = 2f(x), \tag{37.201}$$

qui est constant et certainement borné.

Nous avons maintenant prouvé que pour presque tout x nous avons $\hat{f}_n(x) \xrightarrow{L^2(P)} f(x)$. Montrons que cela implique la convergence en loi, c'est-à-dire que pour tout $\eta > 0$, nous avons la limite

$$P\big(|\hat{f}_n(x) - f(x)| > \eta\big) \to 0. \tag{37.202}$$

Si cela n'était pas vrai, nous aurions un nombre $\eta_0 > 0$ et $\epsilon > 0$ tel que pour tout n à partir d'une certaine taille,

$$P\Big(|\hat{f}_n(x) - f(x)|^2 > \eta_0\Big) > \epsilon, \tag{37.203}$$

et en particulier en notant A l'événement $|\hat{f}_n(x) - f(x)|^2 > \eta_0^2$, $P(A) > \epsilon$. Alors

$$\int_\Omega \big|\hat{f}_n(x,\omega) - f(x,\omega)\big|^2 dP(\omega) \geqslant \int_A \big|\hat{f}_n(x,\omega) - f(x,\omega)\big|^2 dP(\omega) \geqslant \int_A \eta_0^2 = \eta_0^2 P(A). \tag{37.204}$$

Cela signifie que

$$\|\hat{f}_n(x) - f(x)\|_{L^2(P)} \geqslant \eta_0 P(A), \tag{37.205}$$

ce qui contredit la première convergence démontrée. □

Note : l'hypothèse $nh_n \to \infty$ revient à dire que nous voulons que chaque boîte contienne de plus en plus d'observations. Si nous avions $nh_n \to 0$, alors avec n qui augmente, la majorité des boîtes deviendraient vides, ce qui reviendrait à une perte d'information.

37.8 Test d'hypothèses, prise de décision

37.8.1 Exemple : qualité des pièces d'usine

Une usine fabrique des composantes électroniques garantis un an. Le constructeur ne veut pas accepter que plus de 5% des pièces tombent en panne avant un an.

Nous supposons que la durée de vie T d'une pièce soit une variable aléatoire suivant une loi exponentielle de paramètre λ (qui est l'inverse de la moyenne : $\theta = 1/\lambda$). Le fabriquant veut donc s'assurer que

$$0.95 \leqslant P(T \geqslant 1), \tag{37.206}$$

ou encore

$$P(T \geqslant 1) = \int_0^\infty \frac{1}{\theta} e^{-x/\theta} dx = e^{-1/\theta} \geqslant 0.95, \tag{37.207}$$

donc le fabriquant doit s'assurer que

$$\theta \geqslant \frac{1}{\ln\left(\frac{1}{0.95}\right)}. \tag{37.208}$$

Nous posons donc $\theta_0 = 19.5$ et nous prenons comme modèle de décision que si $\theta < \theta_0$, alors la chaine de production doit être revue, et si $\theta > \theta_0$, alors l'usine peut continuer son travail.

Ce dont nous disposons n'est pas de θ, mais d'une estimation de θ à partir d'un échantillon. Cela étant il faudra aussi pouvoir estimer la probabilité de faire un mauvais choix.

37.8.2 Exemple : la résistance d'un fil

subsecExempLFilResituzz

Un artisan a besoin d'un fil qui a une résistance à une traction de 100 g en moyenne. Si la résistance est trop faible, le fil casse ; si elle est trop grande, c'est trop rigide et ça ne convient pas.

Remarque 37.50.
Dans l'exemple précédent, avoir $\theta > \theta_0$ ne dérange pas. Si la durée de vie moyenne est de 2 ans, le directeur de l'usine ne sera pas malheureux. Ici par contre l'artisan cherche une valeur précise et a donc une borne vers le haut et vers le bas.

L'artisan reçoit un lot de fils et souhaite savoir si il est conforme. Pour cela, il prend 4 fils au hasard et mesure une moyenne de 112 g. Est-ce que cela est cohérent avec une moyenne de 100 g ?

Nous faisons l'hypothèse que la résistance des fils suit une loi normale $\mathcal{N}(m, \sigma^2)$ avec m inconnu. Pour la simplicité nous supposons que σ est connu et vaut 10.

Nous devons prendre une décision entre deux hypothèses. L'hypothèse H_0 sera de dire que le lot a une résistance de 100 g et l'hypothèse alternative sera que le lot a une résistance différente.

Les 4 observations sont quatre variables aléatoires $(X_i)_{i=1,2,3,4}$, et le nombre 112 est une réalisation de la variable aléatoire

$$\bar{X}_4 = \frac{1}{4}(X_1 + X_2 + X_3 + X_4). \tag{37.209}$$

Nous supposons que H_0 est vraie, et nous calculons quel est l'intervalle autour de $m = 100$ qui a 95% de chances de contenir la moyenne observée. Si 112 est dedans, nous acceptons H_0 et si 112 est hors de cet intervalle, nous refusons H_0.

Compte tenue de l'hypothèse H_0, nous avons

$$\frac{\bar{X}_4 - 100}{\frac{10}{\sqrt{4}}} = \frac{\bar{X}_4 - 100}{5} \sim \mathcal{N}(0,1). \tag{37.210}$$

Nous commençons à connaitre par cœur l'intervalle de confiance à 95% d'une loi normale centrée réduite ; le quantile est à 1.96, c'est-à-dire

$$P\left(-1.96 \leqslant \frac{\bar{X}_4 - 100}{5} \leqslant 1.96\right) = 0.95, \tag{37.211}$$

ou encore

$$P\big(\bar{X}_4 \in [90.2, 109.8]\big). \tag{37.212}$$

Il y a donc moins de 5% de chances que la moyenne de ces quatre fils tombe en dehors de l'intervalle $[90.2, 109.8]$. L'artisan doit donc rejeter l'hypothèse et considérer que le lot est mauvais.

La région

$$]-\infty, 90.2] \cup [109.8, \infty[\tag{37.213}$$

est la **région de rejet**, ou **région critique**.

Ici, le nombre 5% représente le risque de refuser H_0 alors qu'elle était vraie. Notons que nous ne pouvons pas calculer le risque d'accepter H_0 alors qu'elle est fausse. En effet, si H_0 est fausse, nous ne savons pas quelles sont les valeurs de $\bar{X}_4$ acceptables parce qu'il y a une infinité de possibilités pour m qui soient alternatifs à $m = 100$.

Évidemment si la vraie moyenne est $100+10^{-7}$, l'hypothèse H_0 sera acceptée, mais nous n'avons aucun moyen de savoir si elle est vraie ou non.

37.8.3 Vocabulaire et théorie

Nous avons un modèle d'échantillonnage paramétrique $(X_{\theta,1}, \dots, X_{\theta,n})$ de taille n et de paramètre inconnu θ, de loi parente μ_θ appartenant à une famille paramétrique de lois $(\mu_\theta)_{\theta\in\Theta}$.

Soient H_0 et H_1 deux ensembles disjoints tels que $\Theta = H_0 \cup H_1$. L'ensemble H_0 sera nommé **hypothèse nulle** et l'ensemble H_1 sera l'**hypothèse alternative**.

Pour l'exemple des fils, nous avions $H_0 = \{100\}$ et $H_1 = \mathbb{R}\backslash\{100\}$. Si une hypothèse est réduite à un singleton, nous parlons d'hypothèse **simple** et sinon c'est une hypothèse **composite** ou **multiple**. Faire un tests consiste à déterminer une région critique.

Définition 37.51.
*Un **test** est une application mesurable δ qui à $(x_1, \dots, x_n) \in \mathbb{R}^n$ associe*

$$\delta(x_1, \dots, x_n) \in \{0, 1\}. \tag{37.214}$$

Si $\delta(x_1, \dots, x_n) = 0$ on accepte l'hypothèse H_0 pour l'échantillon $x_1, \dots, x_n$, et si $\delta(x_1, \dots, x_n) = 1$, alors on rejette H_0 et on choisit H_1. L'ensemble

$$W = \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } \delta(x_1, \dots, x_n) = 1\} \tag{37.215}$$

*est la **région de rejet** ou la **région critique**.*

L'ensemble $W = \delta^{-1}(1)$ est un borélien de $\mathbb{R}^n$ parce que δ est mesurable. L'événement auquel nous sommes intéressés est l'événement

$$R = \{(X_{\theta,1}, \dots, X_{\theta,n}) \in W\}. \tag{37.216}$$

Exemple 37.52.
Pour l'exemple de 37.8.2 nous avions

$$\delta(x_1, \dots, x_4) = \mathbb{1}_{\complement[90.2,109.8]}(\bar{x}_4). \tag{37.217}$$

△

37.8.4 Risque de première et seconde espèce

Le modèle de décision que nous avons introduit comprend deux façons de se tromper. Soit nous rejetons H_0 alors qu'elle est vraie (c'est le **risque de première espèce**), soit nous acceptions H_0 alors qu'elle est fausse (risque de **seconde espèce**). Nous pouvons formaliser ces concepts de la façon suivante.

Nous considérons un test de région critique W. Le risque de première espèce, noté α est la fonction

$$\begin{aligned} \alpha\colon H_0 &\to [0,1] \\ \theta &\mapsto P\big((X_{\theta,1}, \dots, X_{\theta,n}) \in W\big). \end{aligned} \tag{37.218}$$

Il s'agit de la probabilité de rejeter H_0 alors qu'elle es vraie. Le risque de seconde espèce est la fonction

$$\begin{aligned} \beta\colon H_1 &\to [0,1] \\ \theta &\mapsto P\big((X_{\theta,1}, \dots, X_{\theta,n}) \notin W\big). \end{aligned} \tag{37.219}$$

C'est la probabilité d'accepter H_0 alors qu'elle est fausse.

DefPuisszYkrQa

Définition 37.53.
*Soit δ un test de région critique W. La **puissance** du test est la fonction*

$$\begin{aligned} \eta\colon H_1 &\to [0,1] \\ \theta &\mapsto P\big((X_{\theta,1}, \dots, X_{\theta,n}) \in W\big). \end{aligned} \tag{37.220}$$

*La **courbe d'efficacité** du test est la fonction*

$$\begin{aligned} h\colon \Theta &\to [0,1] \\ \theta &\mapsto P\big((X_{\theta,1}, \dots, X_{\theta,n}) \notin W\big). \end{aligned} \tag{37.221}$$

La puissance d'un test est la probabilité de rejeter H_0 lorsque H_1 est vraie. Plus la puissance est grande, mieux c'est. La courbe d'efficacité du test est la probabilité d'accepter H_0 pour une certaine valeur de θ.

Soit un test δ. Une statistique $T_\theta = T_n(X_{\theta,1}, \dots, X_{\theta,n})$ est une **variable de décision** pour δ si $\complement W$ peut s'écrire d'une des façons suivantes

$$\complement W = \begin{cases} \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } T_n(x_1, \dots, x_n) < c\} & \text{test unilatéral à droite} \\ \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } T_n(x_1, \dots, x_n) > c\} & \text{test unilatéral à gauche} \\ \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } c_1 \leqslant T_n(x_1, \dots, x_n) < c_2\} & \text{test bilatéral.} \end{cases} \tag{37.222}$$

Le plus souvent la variable de décision sera la moyenne : $T_n(x_1, \dots, x_n) = \frac{1}{n}\sum_{i=1}^n x_i$. Les valeurs c, c_1, c_2 sont des **valeurs critiques**.

En ce qui concerne les notations, ici T_n représente la valeur mesurée sur un n-échantillon (d'où l'indice n) alors que T_θ est la valeur *théorique* de T lorsque θ est la vraie valeur du paramètre qu'on veut estimer.

Pour un test unilatéral à droite, nous fixons la valeur critique c de telle manière à avoir

$$P(T_\theta > c) \leqslant \alpha. \tag{37.223}$$

Pour un test unilatéral à gauche, nous fixons c de telle manière à avoir

$$P(T_\theta < c) \leqslant \alpha \tag{37.224}$$

et pour un test bilatéral nous fixons c_1 et c_2 de telle façon à avoir

$$P(T_\theta > c_2) = P(T_\theta < c_1) \leqslant \frac{\alpha}{2}. \tag{37.225}$$

37.8.5 Modèle paramétrique de loi gaussienne

Soit un modèle statistique paramétrique de lois gaussiennes $\mathscr{N}(m,1)$ de moyenne m inconnue avec $m \in \mathbb{R}^+$. Nous avons $\Theta = [0,\infty[$.

Nous observons un échantillon de taille $n = 36$. Avec un risque de première espèce de 5% nous voulons estimer l'hypothèse $H_0 = \{0\}$ contre l'hypothèse $H_1 =]0,\infty[$. De notre échantillon nous construisons la variable aléatoire

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{37.226}$$

dans laquelle les X_i sont les éléments de l'échantillon, elles sont indépendantes et identiquement distribuées de loi $\mathscr{N}(m,1)$ avec m inconnu.

Si $\bar{X}_n$ est proche de zéro nous acceptons H_0, sinon nous la rejetons. La région de rejet s'écrit donc sous la forme

$$W = \{(x_1,\dots,x_n) \in \mathbb{R}^n \text{ tel que } \bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i > u\} \tag{37.227}$$

dans lequel il faut fixer le u pour satisfaire au risque de première espèce de 5%. La contrainte est d'avoir

$$P(\bar{X}_n > u) = \alpha \tag{37.228}$$

si H_0 est vérifiée. Cela revient à dire que dans 5% des cas où H_0 est correcte, nous la rejetterons. Si H_0 est vraie alors $\bar{X}_n$ est une moyenne de gaussiennes de moyennes m et nous avons

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} \sim \mathscr{N}(0,1) \tag{37.229}$$

avec $m = 0$ et $\sigma = 1$. L'équation (37.228) devient donc

$$\alpha = P\left(\frac{\bar{X}_n}{1/\sqrt{n}} > \frac{u}{1/\sqrt{n}}\right) = P(T > \sqrt{n}u) \tag{37.230}$$

où $T \sim \mathscr{N}(0,1)$. Avec $n = 36$ et $\alpha = 5\%$ nous trouvons

$$u = \frac{1.645}{6} \simeq 0.274 \tag{37.231}$$

La règle de décision est donc la suivante : si $\bar{x}_n > 0.274$ alors nous rejetons H_0, et sinon nous l'acceptons.

Calculons la puissance de ce test (définition 37.53). C'est la fonction donnée par

$$\begin{aligned}\eta\colon H_1 &\to \mathbb{R}\\ m &\mapsto P\big((X_{1,m},\dots,X_{n,m}) \in W\big) = P\left(\frac{1}{n}\sum X_i > u\right).\end{aligned} \tag{37.232}$$

Dans ce calcul, les X_i sont d'une loi normale $\mathcal{N}(m,1)$, et non $\mathcal{N}(0,1)$. En retranchant m et en divisant par $1/\sqrt{n}$ nous trouvons

$$\eta(m) = P\left(\frac{\frac{1}{n}\sum X_i - m}{1/\sqrt{n}} > \frac{u-m}{1/\sqrt{n}}\right) \tag{37.233a}$$
$$= P(T > \sqrt{n}(u-m)) \tag{37.233b}$$
$$= P(T > 16.45 - 6m) \tag{37.233c}$$
$$= 1 - \Phi(1.645 - 6m) \tag{37.233d}$$

où Φ est la fonction de répartition de $\mathcal{N}(0,1)$. La fonction η a les propriétés suivantes :

$$\lim_{m\to-\infty} \eta(m) = 0 \tag{37.234a}$$
$$\lim_{m\to\infty} \eta(m) = 1 \tag{37.234b}$$
$$\eta(0) = \frac{5}{100}. \tag{37.234c}$$

Remarque 37.54.
Si nous regardons $m = 0.001$, le risque de seconde espèce est quasiment de 90%. En effet le risque de seconde espèce est d'accepter H_0 alors qu'il est faux. Lorsque $m = 0.001$, l'hypothèse H_0 est fausse, mais la probabilité qu'on l'accepte est grande. D'ailleurs les conséquences de l'accepter à tort ne sont peut-être pas si grandes que cela.

37.9 Tests paramétriques

La proposition suivante montre le lien entre région de confiance et les tests.

Proposition 37.55.
Soit $\Lambda(X_1,\dots,X_n)$ une région de confiance par excès de niveau de confiance $1-\alpha$. Alors il existe un test pur de niveau α pour tester $H_0 = \{\theta_0\}$ de région de rejet

$$W_n = \{x = (x_1,\dots,x_n) \in \mathbb{R}^n \text{ tel que } \theta_0 \notin \Lambda(x_1,\dots,x_n)\}. \tag{37.235}$$

Démonstration. L'hypothèse sur Λ signifie qu'avec les observations $(X_1,\dots,X_n)$, il y a une forte probabilité (plus grande que $1-\alpha$) que θ soit dans $\Lambda(X_1,\dots,X_n)$. Avec ou sans H_0 nous avons donc

$$P(\theta \in \Lambda) \geqslant 1-\alpha. \tag{37.236}$$

Supposons maintenant l'hypothèse H_0, alors

$$P\big((X_1,\dots,X_n) \in W_n\big) = P\big(\theta_0 \notin \Lambda(X_1,\dots,X_n)\big) \leqslant \alpha. \tag{37.237}$$

□

Remarque 37.56.
Soit W_n la région de confiance d'un test de niveau α pour tester $H_0 = \{\theta_0\}$. Alors

$$\Lambda = \{x \in \mathbb{R}^n \text{ tel que } x \notin W_n\} \tag{37.238}$$

est une région de confiance $1-\alpha$ pour θ.

Exemple 37.57.
Soit $(X_1,\dots,X_n)$ un échantillon de loi parente $\mathcal{N}(\theta,1)$ avec $\theta \in \{\theta_0,\theta_1\}$. Nous supposons $\theta_0 < \theta_1$. Nous voulons tester $H_0 = \{\theta_0\}$ contre $H_1 = \{\theta_1\}$. Nous proposons le test suivant. La variable de décision sera $\bar{X}_n$ et la région de rejet sera

$$W = \{(x_1,\dots,x_n) \in \mathbb{R}^n \text{ tel que } \frac{1}{n}\sum_i x_i > \frac{\theta_0+\theta_1}{2}\}. \tag{37.239}$$

(1) Donner le risque de première espèce de ce test.

(2) Soit $0 < \alpha < 1$. Pour quelle valeur de n le test a-t-il un risque de première espèce égal à α ?

(3) Donner la puissance du test.

Les réponses peuvent être exprimées en termes de la fonction de répartition F de la loi normale centrée réduite.

Le risque de première espèce est donné par

$$\alpha = P\left(\frac{1}{n}\sum_i X_i > \frac{\theta_0 + \theta_1}{2}\right) \tag{37.240}$$

où les X_i sont des variables aléatoires indépendantes et identiquement distribuées de loi parente $\mathcal{N}(\theta_0, 1)$. Cela est la probabilité d'être dans la région de rejet alors que l'hypothèse H_0 est vraie. La formule (37.240) se transforme en

$$\alpha = P\left(\frac{\frac{1}{n}\sum X_i - \theta_0}{1/\sqrt{n}} > \frac{\frac{\theta_0+\theta_1}{2} - \theta_0}{1/\sqrt{n}}\right) \tag{37.241a}$$

$$= P(T > \sqrt{n}\frac{\theta_1 - \theta_0}{2}). \tag{37.241b}$$

En termes de la fonction de répartition nous avons alors

$$\alpha = 1 - F\big(\sqrt{n}\frac{\theta_1 - \theta_0}{2}\big) \tag{37.242}$$

Il s'agit maintenant de trouver le nombre n qui réalise cette égalité. Pour cela nous utilisons l'inverse F^{-1} de la fonction de répartition de la normale :

$$n = \left(\frac{2}{\theta_1 - \theta_0}F^{-1}(1-\alpha)\right)^2. \tag{37.243}$$

Le risque de seconde espèce est la possibilité d'accepter H_0 lorsque H_1 est vraie, c'est-à-dire

$$\beta = P\left(\frac{1}{n}\sum_i X_i < \frac{\theta_0 + \theta_1}{2}\right) \tag{37.244}$$

sous l'hypothèse H_1. Dans le calcul de (37.244) nous prenons donc $X_i \sim \mathcal{N}(\theta_1, 1)$. Le résultat est que

$$\beta = F\big(\sqrt{n}\frac{\theta_0 - \theta_1}{2}\big). \tag{37.245}$$

Remarque 37.58.
L'expression (37.243) diminue lorsque θ_0 et θ_1 s'éloignent, ce qui est normal : plus les nombres à discerner sont éloignés, moins l'échantillon à prendre pour réaliser le travail doit être grand.

Notons aussi que $\theta_0 - \theta_1 < 0$, par conséquent augmenter n diminue la valeur de

$$\beta = F\big(\sqrt{n}\frac{\theta_0 - \theta_1}{2}\big) \tag{37.246}$$

△

37.10 Tests d'adéquation

Soit $X_1, \dots, X_n$ un échantillon de loi parente X finie prenant ses valeurs dans $\{a_1, \dots, a_k\}$. La loi de X est donnée par les nombres

$$p_i = P(X = a_i) \tag{37.247}$$

pour $i = 1, \dots, k$. Nous introduisons l'**effectif empirique**, la variable aléatoire N_i qui compte le nombre de fois que a_i est observé dans l'échantillon. La **fréquence empirique** est la variable aléatoire

$$F_i = \frac{N_i}{n}. \tag{37.248}$$

Nous savons que la loi de N_i est $\mathscr{B}(n, p_i)$, et la loi des grands nombres dit que

$$F_i \xrightarrow{p.s.} p_i \tag{37.249}$$

pour chaque i. Le théorème central limite nous indique de plus que

$$\frac{N_i - np_i}{\sqrt{np_i(1-p_i)}} \xrightarrow{\mathscr{L}} T \sim \mathscr{N}(0,1). \tag{37.250}$$

Nous considérons un cas où les p_i sont inconnus. Ils peuvent être approchés par $N_i \simeq np_i$. Le théorème de Pearson nous indique comment.

Théorème 37.59 (Théorème de Pearson)**.**
Nous avons

$$\sum_{i=1}^{k} \frac{(N_i - np_i)^2}{np_i} \xrightarrow{\mathscr{L}} K \sim \chi^2(k-1) \tag{37.251}$$

où la distribution $\chi^2(l)$ est la somme des carrés de l gaussiennes centrées réduites indépendantes.

Démonstration. Nous commençons par le cas $k = 2$. Dans ce cas nous avons $N_2 = n - N_1$ et $p_1 + p_2 = 1$. La somme que nous regardons est

$$\frac{(N_1 - np_1)^2}{np_1} + \frac{(N_2 - np_2)^2}{np_2} = \frac{(N_1 - np_1)^2}{np_1} + \frac{(N_1 - np_1)^2}{n(1-p_1)} \tag{37.252a}$$

$$= \frac{(N_1 - np_1)^2}{np_1(1-p_1)}. \tag{37.252b}$$

subeqHETRdC

Étant donné que N_1 est une variable aléatoire binomiale nous avons

$$\frac{N_1 - np_1}{\sqrt{np_1(1-p_1)}} \xrightarrow{\mathscr{L}} T \sim \mathscr{N}(0,1). \tag{37.253}$$

Par conséquent la limite de (37.252b) est

$$\left(\frac{N_1 - np_1}{\sqrt{np_1(1-p_1)}} \right)^2 \xrightarrow{\mathscr{L}} T^2 \simeq \chi^2(1). \tag{37.254}$$

Cela conclut le cas $k = 2$.

Passons à présent au cas général. Le k-uple $(N_1, \dots, N_k)$ est une variable aléatoire multinomiale de loi

$$\mathscr{M}(n; k; p_1, \dots, p_k). \tag{37.255}$$

Nous introduisons les variables aléatoires U_i données par $U_i \colon \Omega \to \mathbb{R}^k$ avec

$$P\big(U_i = (0, \dots, 1, \dots, 0)\big) = p_i; \tag{37.256}$$

c'est le vecteur aléatoire qui prend ses valeurs dans les vecteurs de la base canonique de $\mathbb{R}^k$ et qui prend la valeur e_i avec probabilité p_i. Par construction nous avons

$$(N_1, \dots, N_k) = \sum_{i=1}^{n} U_i. \tag{37.257}$$

Nous allons étudier la fonction caractéristique de $(N_1, \dots, N_k)$ définie par l'équation (36.279) :

$$\begin{aligned}\Phi_{(N_1,\dots,N_k)}\colon \mathbb{R}^k &\to \mathbb{C}\\ e_j &\mapsto E(e^{ie_j\cdot N}) = E(e^{iN_j}).\end{aligned} \tag{37.258}$$

Plus généralement,

$$\Phi_{(N_1,\dots,N_k)}(t_1,\dots,t_k) = E\big(e^{i\langle t,N\rangle_{\mathbb{R}^k}}\big). \tag{37.259}$$

Nous avons

$$e^{i\langle t,N\rangle} = e^{i\sum_j\langle t,U_j\rangle} = \prod_j e^{i\langle t,U_j\rangle} \tag{37.260}$$

et vu que les U_i sont indépendantes et identiquement distribuées nous pouvons écrire U_1 à la place de U_j de façon à avoir EqOhTHia

$$\Phi_{(N_1,\dots,N_k)}(t_1,\dots,t_k) = \prod_j E\big(e^{i\langle t,U_1\rangle}\big) \tag{37.261a}$$

$$= \prod_j \sum_l p_l e^{i\langle t,e_l\rangle} \tag{37.261b}$$

$$= \prod_j \sum_l p_l e^{it_l} \tag{37.261c}$$

$$= \left(\sum_{l=1}^k p_l e^{it_l}\right)^n. \tag{37.261d}$$

Nous allons montrer que

$$\lim_{n\to\infty}\Phi_{(N_1,\dots,N_n)}(t_1,\dots,t_k) = e^{-\frac{1}{2}\sum_{j=1}^n t_j^2} - \left(\sum_{j=1}^k t_j\sqrt{p_j}\right)^2. \tag{37.262}$$

Pour ce faire, nous allons effectuer un développement limité. D'abord nous introduisons les variables aléatoires

$$\alpha_j = \frac{N_j - np_j}{\sqrt{np_j}} \tag{37.263}$$ EqnqdROCD

et nous calculons

$$\Phi_{(\alpha_1,\dots,\alpha_k)}(t_1,\dots,t_k) = E\left[\exp\left(i\langle t,(\frac{N_1-np_1}{\sqrt{np_1}},\dots,\frac{N_k-np_k}{\sqrt{np_k}})\rangle\right)\right] \tag{37.264}$$

Étant donné que n et p_j sont des variables déterministes, nous pouvons les sortir de l'espérance. Nous avons alors

$$\Phi_{(\alpha_1,\dots,\alpha_k)}(t_1,\dots,t_k) = \exp\left(-i\sum_{j=1}^k t_j\sqrt{np_j}\right)\Phi_{(N_1,\dots,N_k)}\left(\frac{t_1}{\sqrt{np_1}},\dots,\frac{t_k}{\sqrt{np_k}}\right) \tag{37.265}$$ EqlzeCLS

parce que

$$E\left(e^{t_j\frac{N_j-np_j}{\sqrt{np_j}}}\right) = e^{\frac{-t_jnp_j}{\sqrt{np_j}}}E\left(e^{t_jN_j/\sqrt{np_j}}\right) \tag{37.266}$$

En remplaçant (37.261) dans (37.265) nous trouvons

$$\Phi_{(\alpha_1,\dots,\alpha_k)}(t_1,\dots,t_k) = \exp\left(-i\sum_{j=1}^k t_j\sqrt{np_j}\right)\underbrace{\left(\sum_{j=1}^k p_j e^{i\frac{t_j}{\sqrt{np_j}}}\right)^n}_{A} \tag{37.267}$$ EqrUYsaD

Nous analysons maintenant le terme A. Nous écrivons l'égalité $A = A + 1 - 1$ en tenant compte de $\sum_j p_j = 1$ sous la forme

$$A = \left(1 + \sum_{j=1}^{k} p_j \big(\exp(it_j/\sqrt{np_j}) - 1 \big) \right)^n, \tag{37.268}$$

Nous avons alors

$$\ln(A) = n \ln \left[1 + \sum_{j=1}^{\infty} p_j \big(e^{it_j/\sqrt{np_j}} - 1 \big) \right] \tag{37.269}$$

Nous développons l'exponentielle en

$$e^{it_j/\sqrt{np_j}} - 1 = \frac{it_j}{\sqrt{np_j}} - \frac{t_j^2}{2np_j} + \frac{1}{n}\epsilon(1/n) \tag{37.270}$$

et ensuite le logarithme selon la formule

$$\ln(1 + x) = x - \frac{x^2}{2} + x^2\alpha(x^2). \tag{37.271}$$

Nous avons

$$\ln(A) = n \ln \left[1 + \sum_j p_j \Big(\frac{it_j}{\sqrt{np_j}} - \frac{t_j^2}{2np_j} + \frac{1}{n}\epsilon(\frac{1}{n}) \Big) \right] \tag{37.272a}$$

$$= n \ln \left[1 + \sum_j p_j \Big(it_j \sqrt{\frac{p_j}{n}} - \frac{t_j^2}{2n} + \frac{1}{n}\epsilon(\frac{1}{n}) \Big) \right] \tag{37.272b}$$

$$= n \underbrace{\sum_{j=1}^{k} \left(it_j \sqrt{\frac{p_j}{n}} - \frac{t_j^2}{2n} + \frac{1}{n}\epsilon(1/n) \right)}_{K} \tag{37.272c}$$

$$- n\frac{1}{2} \left[\sum_j \left(it_j \sqrt{\frac{p_j}{n}} - \frac{t_j^2}{2n} + \frac{1}{n}\epsilon(1/n) \right)^2 \right] \tag{37.272d}$$

$$+ nK^2\alpha(K) \tag{37.272e}$$

Nous introduisons dans ϵ tous les termes en $1/n^2$ et nous trouvons

$$\ln(A) = \sum_j \left(it_j \sqrt{p_j n} - \frac{t_j^2}{2} \right) - \frac{1}{2} \left(\sum_j it_j \sqrt{p_j} \right)^2 + \epsilon(1/n) + K^2\alpha(K). \tag{37.273}$$

En remplaçant dans (37.267) et en passant à la limite pour $n \to \infty$,

$$\Phi_{(\alpha_1,\dots,\alpha_k)}(t_1, \dots, t_k) = \exp \left(-\frac{1}{2} \sum_j t_j^2 + \frac{1}{2} \big(\sum_j t_j \sqrt{p_j} \big)^2 \right). \tag{37.274}$$

EaqTXgCW

Nous reconnaissons des lois gaussiennes dans le premier terme de l'exponentielle. Nous allons maintenant nous atteler à identifier le second terme.

Soit C une matrice orthogonale dont la dernière ligne est $(\sqrt{p_1}, \dots, \sqrt{p_k})$. Nous considérons les vecteurs

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}, \quad t = \begin{pmatrix} t_1 \\ \vdots \\ t_k \end{pmatrix}. \tag{37.275}$$

et ensuite nous notons $U = (u_1, \dots, u_k)$ le vecteur Ct et $(\beta 1, \dots, \beta_k)$ le vecteur $C\alpha$. Étant donné que C est orthogonale, nous avons $\sum_{i=1}^k \alpha_i^2 = \sum_{i=1}^k \beta_i^2$ et $\sum_i t_i^2 = \sum_i u_i^2$. Nous avons alors

$$\Phi_{(\beta_1,\dots,\beta_k)} = E\big(e^{i\langle u,\beta\rangle}\big) = E\big(e^{i\langle t,\alpha\rangle}\big) = \Phi_{(\alpha_1,\dots,\alpha_k)}(t_1, \dots, t_k). \tag{37.276}$$

Nous pouvons récrire l'argument de l'exponentielle (37.274) de la façon suivante :

$$\sum_{i=1}^k t_j^2 = \sum_j u_j^2 \tag{37.277a}$$

$$\sum_{j=1}^k t_j \sqrt{p_j} = (Ct)_k, \tag{37.277b}$$

Nous avons alors

$$\lim_{n\to\infty} \Phi_{(\beta_1,\dots,\beta_k)}(t_1, \dots, t_k) = \lim_{n\to\infty} \Phi_{(\alpha_1,\dots,\alpha_k)}(u_1, \dots, u_k) \tag{37.278a}$$

$$= \exp\left(\frac{1}{2}\sum_{j=1}^{k-1} u_j^2\right) \tag{37.278b}$$

$$= \Phi_{(Z_1,\dots,Z_{k-1},0)}(u_1, \dots, u_k) \tag{37.278c}$$

où les Z_i sont des variables aléatoires indépendantes et identiquement distribuées de distribution normale centrée réduite. Nous avons donc montré que

$$(\beta_1, \dots, \beta_k) \xrightarrow{\mathscr{L}} (Z_1, \dots, Z_{k-1}, 0). \tag{37.279}$$

Étant donné que l'application $x \mapsto \|x\|^2$ est continue, nous avons aussi

$$\|(\beta_1, \dots, \beta_k)\|^2 \xrightarrow{\mathscr{L}} \|(Z_1, \dots, Z_{k-1}, 0)\|^2, \tag{37.280}$$

et par conséquent

$$\|(\alpha_1, \dots, \alpha_k)\|^2 \xrightarrow{\mathscr{L}} \sum_{j=1}^{k-1} Z_j^2 \sim \chi^2(k-1). \tag{37.281}$$

D'après la définition (37.263) nous avions

$$\|(\alpha_1, \dots, \alpha_k)\|^2 = \sum_{j=1}^k \frac{(N_j - p_j)^2}{np_j}. \tag{37.282}$$

□

Chapitre 38

Chaînes de Markov à temps discret

Mets tes deux pieds en canard, c'est la chaine de Markov qui se prépare.

38.1 Généralités

Les chaines de Markov interviennent pour la description des systèmes dont l'évolution future ne dépend que de l'état présent.

DEFooGDPFooWsvfRv

Définition 38.1.
Soit E un ensemble au plus dénombrable[1] *et $(\Omega, \mathcal{F}, P)$ un espace de probabilité. Une **chaine de Markov** à valeurs dans E est une famille $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires telles que pour tout $x_0, \dots, x_{n+1} \in E$,*

$$P(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = x_{n+1} | X_n = x_n). \tag{38.1}$$

Pour une chaine de Markov, il n'est pas important de savoir l'historique pour prédire le futur : X_{n+1} est seulement déterminé par X_n.

Remarque 38.2.
Il existe une théorie des chaines de Markov à temps continu ou avec E non dénombrable, mais ce n'est pas au programme.

38.3.
Vu que l'ensemble E des états est au plus dénombrable, nous rappelons très humblement à la lectrice la proposition 10.126 qui nous permet de changer des sommes sur E en des sommes sur $\mathbb{N}$ sans nous soucier de l'ordre sur E. Si f est une fonction sur E, nous nous écrirons

$$\sum_{x\in E} f(x) = \sum_{k=0}^{\infty} f(x_k) \tag{38.2}$$

sans citer 10.126 à chaque fois.

DEFooVVWUooKIBQDv

Définition 38.4.
Si (X_n) est une chaine de Markov[2]*, nous notons*

$$p_n(x, y) = P(X_{n+1} = y | X_n = x) \tag{38.3}$$

*la **probabilité de transition** de la chaine à l'instant n. Si cette probabilité ne dépend pas de n, nous disons que la chaine de Markov est **homogène**, et nous notons $p(x, y)$ au lieu de $p_n(x, y)$.*

1. Une chaine de Markov sur un ensemble indénombrable demanderait plus de technique à cause du lemme 10.122 qui fait que toutes les sommes sur des ensembles indénombrables sont infinies.
2. Définition 38.1.

DEFooKQROooYvJvvl

Définition 38.5 (Matrice de transition)**.**
Nous notons $Q^{(n)}$ la matrice de transition qui est éventuellement infinie :

$$Q_{xy}^{(n)} = p_n(x, y). \tag{38.4}$$

Si l'ensemble des états E est infini, ce n'est pas une matrice à proprement parler.

Le lemme suivant est intuitivement rien d'autre que le fait que la somme des probabilités doit être 1.

LEMooQNIWooQBMlge

Lemme 38.6.
Soit un processus de Markov (X_n) sur l'ensemble E. Pour chaque $x \in E$ pour pour tout $n \in \mathbb{N}$ nous avons

$$\sum_{y\in E} p_n(x, y) = 1. \tag{38.5}$$

Démonstration. Nous avons le calcul

$$\begin{aligned}
\sum_{y\in E} p_n(x, y) &= \sum_{y\in E} P(X_{n+1} = y | X_n = x) && \text{(38.6a)}\\
&= P(\Omega | X_n = x) && \text{(38.6b)}\\
&= \frac{P(\Omega \cap X_n = x)}{P(X_n = x)} && \text{(38.6c)}\\
&= 1. && \text{(38.6d)}
\end{aligned}$$

SUBEQooIRGEooHDOxOd

SUBEQooBRMTooQJiKfR

Justifications :

— Pour (38.6b), c'est le lemme 36.59 en observant que $\Omega = \bigcup_{y\in E}\{X_{n+1} = y\}$.

— Pour (38.6d), c'est $\Omega \cap A = A$.

□

Remarque 38.7.
Attention à ce que ce lemme 38.6 ne fonctionne que sur les colonnes de p_n. En effet, la somme $\sum_{x\in E} p(x, y)$ ne vaut pas spécialement 1. Si les états x_1 et x_2 arrivent tous les deux en y de façon certaine, alors nous avons $\sum_x p(x, y) \geqslant 2$. Il n'y a donc pas de limites aux sommes des lignes.

LEMooZIEPooXHGnvy

Lemme-Définition 38.8 (Produit de matrices de transition)**.**
Soient deux processus de Markov (X_n) et (Y_n) sur le même ensemble E. Si nous notons p et q leurs matrices de transition[3] *alors*

ITEMooWNWXooCKOYpE

(1) La somme $\sum_{x\in E} p(a, x)q(x, b)$ converge pour tout $a, b \in E$.

ITEMooEZIEooFEbwhj

(2) Nous notons pq la « matrice de transition »

$$(pq)(a, b) = \sum_{x\in E} p(a, x)q(x, b). \tag{38.7}$$

ITEMooKEFXooMLREkO

(3) La matrice pq vérifie

$$\sum_{y\in E} (pq)(x, y) = 1 \tag{38.8}$$

pour tout $x \in E$.

Démonstration. Nous considérons la mesure de comptage sur E (définition 14.263). Cela nous permet d'écrire la somme comme une intégrale.

3. Qui n'est pas spécialement une matrice, voir la définition 38.5.

(1) **Pour (1)** Nous devons donc démontrer la convergence de

$$\sum_{x\in E} p(a,x)q(x,b) = \int_E p(a,x)q(x,b)dm(x) \tag{38.9}$$

EQooCNNCooGJRWvi

où $dm(x)$ n'est pas du tout la mesure de Lebesgue (qui n'aurait aucun sens), mais bien la mesure de comptage en x.

Nous pouvons majorer $q(x,b)$ par 1 (tous les nombres sont strictement plus grands que zéro) : pour chaque x nous avons

$$p(a,x)q(x,b) \leqslant p(a,x), \tag{38.10}$$

alors que $\int_E p(a,x)dm(x) = 1$ par le lemme 38.6.

Vu que la fonction $x \mapsto p(a,x)q(x,b)$ est dominée par la fonction $x \mapsto p(a,x)$ et que cette dernière est intégrable, le lemme 14.192 conclut à l'intégrabilité de la première. Bref, l'intégrale (38.9) existe et est finie.

(2) **Pour (2)** Il n'y a rien à prouver, c'est seulement une définition.

(3) **Pour (3)** Nous devons calculer la valeur de

$$\sum_{b\in E}\Big(\sum_{x\in E} p(a,x)q(x,b)\Big) = \int_E \Big(\int_E p(a,x)q(x,b)dm(x)\Big)dm(b). \tag{38.11}$$

Pour cela nous allons utiliser le théorème de Fubini comme expliqué en 14.305, et nous partons de la somme dans le sens inverse. D'abord nous prouvons que $(x,b) \mapsto p(a,x)q(x,b)$ est dans $L^1(E\times E)$ en étudiant les intégrales en chaine :

$$\int_E \Big(\int_E |p(a,x)q(x,b)|dm(b)\Big)dm(x) = \int_E p(a,x)\Big(\underbrace{\int_E q(x,b)dm(b)}_{=1}\Big)dm(x) \tag{38.12a}$$

$$= \int_E p(a,x)dm(x) \tag{38.12b}$$

$$= 1. \tag{38.12c}$$

Donc la fonction est $L^1(E\times E)$ et nous pouvons fusionner et permuter les intégrales à volonté. Nous avons alors

$$1 = \int_E \Big(\int_E |p(a,x)q(x,b)|dm(b)\Big)dm(x) \tag{38.13a}$$

$$= \int_{E\times E} p(a,x)q(x,b)dm(x,b) \tag{38.13b}$$

$$= \int_E \Big(\int_E p(a,x)q(x,b)dm(x)\Big)dm(b). \tag{38.13c}$$

□

38.1.1 Matrice stochastique

DefGJEBooZvuIAV

Définition 38.9.
Une matrice dont tous les éléments sont positifs ou nuls et dont la somme de chaque ligne vaut 1 *est une* ***matrice stochastique****.*

Notons que l'ensemble des matrices stochastiques est un fermé dans l'ensemble des matrices.

Exemple 38.10.
Nous considérons une fourmi qui se déplace dans un appartement à trois pièces A, B, C. Supposons

qu'à chaque minute, elle a une probabilité 1/3 de rester dans la pièce et une probabilité 2/3 de se déplacer. Le plan de l'appartement est

$$A \longrightarrow B \longrightarrow C \tag{38.14}$$

De la pièce A il est donc uniquement possible d'aller vers la pièce B ; de la B il est possible d'aller en A et en C et de la C il est uniquement possible d'aller en B.

La matrice de transition de cette chaine de Markov est

$$Q = \begin{pmatrix} 1/3 & 2/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 2/3 & 1/3 \end{pmatrix} \tag{38.15}$$

△

38.2 Chaînes de Markov sur un ensemble fini

Définition 38.11.
Une chaine de Markov est ***finie*** *si l'ensemble E dans lequel elle prend ses valeurs est fini.*

PROBooNEMXooLbPpIN

¡¡ Avertissement/question à la lectrice !! 38.12
Je ne suis pas totalement certain du lemme suivant (38.13). Soyez prudente, et écrivez-moi si vous avez une démonstration ou un contre-exemple.

Le lemme 38.42 donne une partie de la preuve, je crois.

LEMooYKFZooUEOWbi

Lemme 38.13.
Une matrice est stochastique[4] *si et seulement si elle est la matrice de transition d'une chaine de Markov sur un ensemble fini.*

Définition 38.14.
Soit une matrice stochastique p. Une ***loi stationnaire*** *pour p est un vecteur μ tel que*

(1) $\mu = \mu P$,
(2) $\mu_j \geqslant 0$ *pour tout* j,
(3) $\sum_i \mu_i = 1$.

Lorsque nous parlons d'une chaine de Markov sur un ensemble fini E, la matrice stochastique est indexée par les éléments de E, et, par exemple, la condition $\mu = \mu P$ s'écrit

$$\mu_j = \sum_{i \in E} \mu_i p_{ij} \tag{38.16}$$

pour tout $j \in E$.

Notez que dans ce cas si $A \subset E$, nous pouvons définir $\mu(A) = \sum_{i \in A} \mu_i$. Cette application μ est alors une mesure positive[5] sur E.

DEFooGKRQooFNuwRm

Définition 38.15 ([791])**.**
Une chaine de Markov est ***irréductible*** *si pour tout $x, y \in E$, il existe n tel que $p^n(x, y) > 0$ où*

$$p^n(x, y) = P(X_n = y | X_0 = x). \tag{38.17}$$

Le nombre n peut dépendre de x et y.

Une chaine de Markov finie est ***régulière*** *si il existe un $n \in \mathbb{N}$ tel que P^n a uniquement des éléments strictement positifs.*

4. Matrice stochastique, définition 38.9.
5. Définition 14.18.

PropUMPpOHW

Proposition 38.16 ([792]).
Soit une chaine de Markov irréductible[6] *finie. Alors il existe une unique loi stationnaire π et de plus nous avons $\pi_i > 0$ pour tout état i de E.*

Je crois que pour prouver ça, il faut utiliser le théorème de Perron-Frobenius 9.110.

Théorème 38.17 ([791]).
Soit P la matrice de transition d'une chaine de Markov régulière[7] *sur un ensemble E de cardinal N. Alors il existe des nombres $\pi_1, \dots, \pi_N$ tels que*

(1) $\pi_i > 0$ *pour tout* $i = 1, \dots, N$

(2) $\pi_1 + \cdots + \pi_N = 1$

(3)

$$\lim_{n\to\infty} P^n = \Pi = \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_N \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_N \end{pmatrix} \tag{38.18}$$

De plus le vecteur $\pi = (\pi_1, \dots, \pi_N)$ est l'unique solution de

$$\pi P = \pi. \tag{38.19}$$

Démonstration. Si la chaine n'a qu'un seul état c'est évident parce que la probabilité de transition est toujours 1 ; fin de l'histoire.

(1) **Hypothèse**
Sinon nous supposons que P n'a que des éléments positifs, quitte à considérer P^n au lieu de P. Nous notons d le plus petit élément de P ; il vérifie $d \leqslant \frac{1}{2}$ parce que la somme des élément d'une ligne de la matrice P doit être égale à 1.

(2) **Les suites min et max**
Soit x un vecteur quelconque (de composantes positives). Nous notons $m_0 = \min\{x_i\}$ et $M_0 = \max\{x_i\}$. Étant donné que les éléments du vecteur Px sont des moyennes pondérées des éléments de x, si nous posons

$$m_k = \min\{(P^k x)_i\}_{i=1,\dots,N} \tag{38.20a}$$

$$M_k = \max\{(P^k x)_i\}_{i=1,\dots,N}, \tag{38.20b}$$

la suite (m_k) est croissante et la suite (M_k) est décroissante.

(3) **Stricte croissance et décroissance**
Si $M_{k+1} = M_k$, alors toutes les composantes de $P^k x$ sont égales à M_k et le théorème est prouvé. Cela est encore une propriété de la moyenne. Même remarque pour la suite (m_k).
Nous pouvons donc supposer que la suite (m_k) est strictement croissante et que la suite (M_k) est strictement décroissante. Elles sont toutes les deux bornées dans $[m_0, M_0]$. Le lemme 9.35 nous donne la convergence.

(4) **Égalité des limites**
Vu que les éléments de $P^k x$ ne sont pas tous les mêmes et s'étalent de m_k à M_k, pour majorer M_{k+1} nous mettons le plus petit coefficient possible (c'est-à-dire d) devant m_k et nous supposons que toutes les autres composantes sont M_k ; nous avons alors

$$M_{k+1} \leqslant dm_k + (1-d)M_k \tag{38.21}$$

parce que tous les autres coefficients de la ligne contenant le d (dans P^k) sont plus petits ou égaux à $1-d$. De la même façon nous avons la minoration

$$m_{k+1} \geqslant dM_k + (1-d)m_k. \tag{38.22}$$

6. Définition 38.15.
7. Définition 38.15.

En faisant la différence, et en nous souvenant que $0 < 1 - 2d < 1$,

$$M_{k+1} - m_{k+1} \leqslant (1 - 2d)(M_k - m_k), \tag{38.23}$$

ce qui signifie que

$$M_{k+1} - m_{k+1} \leqslant (1 - 2d)^k (M_0 - m_0), \tag{38.24}$$

et donc que les deux limites sont égales.

(5) **Conclusion pour la limite**

Pour tout vecteur x, la suite $P^k x$ tend vers un vecteur dont toutes les composantes sont égales. En particulier pour le vecteur e_i de la base canonique,

$$P^k e_i \to \begin{pmatrix} \pi_i \\ \vdots \\ \pi_i \end{pmatrix}. \tag{38.25}$$

Mais $P^k e_i$ est la i^{e} colonne de la matrice P^k. Cela prouve la convergence annoncée $P^k \to \Pi$.

Réglons rapidement le cas des deux autres allégations du théorème. D'abord les matrices P^k sont toutes des matrices stochastiques ; et l'ensemble des matrices stochastiques est fermé, donc la convergence se fait à l'intérieur de l'ensemble des matrices stochastiques. Cela prouve que $\pi_1 + \ldots + \pi_N = 1$.

Ensuite la suite (m_k) étant strictement croissante et m_0 étant égal à 0 dans le cas de e_i nous avons toujours $\pi_i > 0$ (strictement). □

38.2.1 Graphe de transition

Le **graphe de transition** d'une chaine de Markov est le graphe dont les sommets sont les éléments de l'espace des états de la chaine et dont les sommets sont reliés par des arêtes pondérées par la probabilité de transition correspondante.

Lemme 38.18.
Une chaine de Markov homogène est irréductible si et seulement si son graphe de transition est connexe.

Démonstration. Pour chaque couple $(x, y) \in E^2$ nous avons

$$\begin{aligned} p^n(x, y) &= \sum_{z_i \in E} P(X_n = y, X_{n-1} = z_{n-1}, \ldots, X_1 = z_1, X_0 = x) \\ &= \sum_{z_i} p(z_{n-1}, y) p(z_{n-2}, z_{n-1}) \ldots p(z_1, z_2) p(x, z_1). \end{aligned} \tag{38.26}$$

La positivité d'un des termes de la somme signifie que le graphe est connexe tandis que la positivité de $p^n(x, y)$ signifie que la chaine est irréductible. □

38.2.2 Nombre de visites

La fonction

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k = x\}} \tag{38.27}$$

est la **fréquence empirique** de la chaine de Markov.

Soit x un état récurrent, c'est-à-dire que $P\big(T(x) < \infty | X_0 = x\big) = 1$. Nous classons les visites de la façon suivante : SubEqsDefTirectempsretou

$$T_1(x) = T(x) = \inf\{k \geqslant 1 \text{ tel que } X_k = x\} \tag{38.28a}$$
$$T_2(x) = \inf\{k \geqslant 1 \text{ tel que } X_{T_1(x)+k} = x\} \tag{38.28b}$$
$$\vdots \tag{38.28c}$$
$$T_n(x) = \inf\{k \geqslant 1 \text{ tel que } X_{T_{n-1}(x)+k} = x\} \tag{38.28d}$$

La variable aléatoire T_i représente le temps entre la visite numéro $i-1$ et la visite numéro i (si $X_0 \neq x$, sinon il faut décaler). Nous définissons l'instant la visite numéro n :

$$S_n = \sum_{k=1}^{n} T_k(x). \tag{38.29}$$

Lemme 38.19.
Les variables aléatoires T_i sont indépendantes.

Démonstration. Nous choisissons n des T_i et nous calculons la probabilité

$$\spadesuit = P(T_{i_1} = k_1, T_{i_2} = k_2, \dots, T_{i_n} = k_n) \tag{38.30}$$

où nous supposons $i_1 > i_2 > \dots > i_n$. Nous décomposons cette probabilité en sommant sur toutes les histoires de la chaine de Markov compatibles avec les nombres k_i donnés :

$$\spadesuit = \sum_{\substack{\{z_j\}\\ \text{compatibles}}} P(X_j = z_j, j = 1, \dots, N). \tag{38.31}$$

Notons qu'ici, le numéro du dernier terme de la somme n'est pas certain parce que tous les T_i ne sont pas fixés. Nous l'avons noté N, mais en réalité il est différent d'un terme à l'autre de la somme. Il est certain que $z_N = x$ et $z_{N-k_1} = x$ et si $N - k_1 < j < N$, alors $z_j \neq x$. Cela est simplement le fait que nous demandions aux z_i de respecter les conditions données par les k_i. Nous avons

$$\spadesuit = \sum_{\{z_j\}} P(X_N = x, X_j = z_j, N - k_1 < j < N | X_j = z_j, j \leqslant N - k_1) P(X_j = z_j, j < N - k_1) \tag{38.32a}$$

$$= \sum_{\{z_j\}} P(X_N = x, X_j = z_j, N - k_1 < j < N | X_{N-k_1} = x) P(X_j = z_j, j < N - k_1) \tag{38.32b}$$

$$\tag{38.32c}$$

Le premier facteur est $P(T_{i_1} = k_1)$ tandis que le second facteur est précisément $P(T_{i_j} = k_j, j > 1)$. Nous avons donc montré que

$$P(T_{i_1} = k_1, T_{i_2} = k_2, \dots, T_{i_n} = k_n) = P(T_{i_1} = k_1) P(T_{i_j} = k_j, j > 1), \tag{38.33}$$

et donc les T_i sont indépendants. □

DEFooDWLAooINqgep

Définition 38.20 (Temps de première atteinte)**.**
Soit (X_k) une chaine de Markov dont l'espace des états est noté E. Pour chaque $x \in E$ nous notons

$$T(x) = \inf\{k \geqslant 1 \text{ tel que } X_k = x\}. \tag{38.34}$$

*Ce nombre est le **temps de première atteinte** de l'état x.*
*Si $X_0 = x$, alors $T(x)$ est le **temps de retour** en x. Si $p \in \mathbb{N}$ nous notons*

$$T_p(x) = \inf\{k \geqslant 1 \text{ tel que } X_{k+p} = x\} \tag{38.35}$$

C'est le temps mis pour atteindre x à partir de l'instant p.

LembyftKs

Lemme 38.21.
Soit (X_k) une chaine de Markov dont l'espace des états est noté E. Pour chaque $x \in E$ nous considérons les temps de premier atteinte[8] *T et T_p. Alors nous avons*

$$P(T_p(x) = k | X_p = y) = P(T(x) = k | X_0 = y). \tag{38.36}$$

8. Définition 38.20.

Remarque 38.22.
La loi des grands nombres est encore vraie sans l'hypothèse de variables aléatoires dans L^1 pourvu qu'elles soient positives. Alors dans la conclusion de la loi nous devons accepter la possibilité que l'espérance soit infinie.

PropMrkIrreLoishLCkpjkptXk

Proposition 38.23.
Soit (X_n) est une chaine de Markov irréductible. Nous notons T le temps de première atteinte[9] *Alors*

(1) Si x est un état récurrent, alors $T(X) < \infty$ presque surement.

(2) Nous avons une égalité entre les lois

$$\mathscr{L}\big(X_{k+T(x)}|T(x) < \infty\big) = \mathscr{L}(X_k|X_0 = x). \tag{38.37}$$

PropMrkIrreLoishLCkpjkptXkItemii

DefWknULk

Définition 38.24.
*Un état x est **récurrent** si $P(T(x) = \infty|X_0 = x) = 0$, c'est-à-dire si la probabilité de ne jamais retourner en x lorsqu'on y est passé est nulle. L'état x est **transient** ou **transitoire** dans le cas contraire.*

*Si x est un état récurrent, et si $E\big(T(x)|X_0 = x\big) < \infty$, nous disons que x est **récurrent positif**. Si $E\big(T(x)|X_0 = x\big) = \infty$ alors nous disons que est **récurrent nul**.*

38.2.3 Récurrent et transient

DEFooPVHHooBkIDpW

Définition 38.25.
Nous introduisons une variable aléatoire qui compte le nombre de fois que la chaine de Markov passe par l'état x :

$$N_x = \sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=x\}}. \tag{38.38}$$

EqDefNxmtuXkn

C'est une variable aléatoire à valeurs dans $\mathbb{N} \cup \{\infty\}$.

PropEquivEPrecuequiv

Proposition 38.26.
Les deux propriétés suivantes sont équivalentes à dire que x est récurrent[10] *:*

(1) $P(N_x < \infty|X_0 = x) = 0$

(2) $E(N_x|X_0 = x) = \infty$.

Les deux propriétés suivantes sont équivalentes à dire que x est transient :

ItemiMnGpD

(1) $P(N_x < \infty|X_0 = x) = 1$

(2) $E(N_x|X_0 = x) < \infty$.

Démonstration. En tant que événements, nous avons l'égalité

$$N_x < \infty = \bigcup_{n\in\mathbb{N}} \{\underbrace{X_n = x, X_{n+k} \neq x \forall k \geqslant 1}_{F_n}\}. \tag{38.39}$$

Nous avons donc

$$P(N_x < \infty|X_0 = x) = \sum_{n=0}^{\infty} P(F_n|X_0 = x), \tag{38.40}$$

Eqreprencalculstd

et

$$\begin{aligned}
P(F_n|X_0 = x) &= P(X_{n+k} \neq x, \forall k \geqslant 1, X_n = x|X_0 = x) && \text{(38.41a)}\\
&= P(X_{n+k} \neq x, k \geqslant 1|X_n = x, X_0 = x)P(X_n = x|X_0 = x) && \text{(38.41b)}\\
&= P(X_{n+k} \neq x, k \geqslant 1|X_n = x)P(X_n = x|X_0 = x) && \text{(38.41c)}\\
&= P(X_k \neq x, k \geqslant 1|X_0 = x)P(X_n = x|X_0 = x) && \text{(38.41d)}\\
&= P(T(x) = \infty|X_0 = x)P(X_n = x|X_0 = x) && \text{(38.41e)}
\end{aligned}$$

subEqPFnXkneqxii subEqPFnXkneqxi subEqPFnXkneqxiii

9. Définition 38.20.
10. La variable aléatoire N_x est définie par 38.25.

Justifications :

(1) Pour (38.41c), nous utilisons le fait que la chaine soit « sans mémoire ».

(2) Pour (38.41d), nous utilisons le fait que la chaine soit homogène.

(3) Pour (38.41e), l'événement $X_k \neq x$ pour tout $k \geqslant 1$ est exactement l'événement $T(x) = \infty$.

En nous servant de la proposition 14.308 (théorème de Fubini et mesure de comptage), nous permutons l'espérance et la somme dans l'expression EqPEEEntstq

$$\sum_{n=0}^{\infty} P(X_n = x|X_0 = x) = \sum_{n=0}^{\infty} E(\mathbb{1}_{\{X_n=x\}}|X_0 = x) \tag{38.42a}$$

$$= E\Big(\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n=x\}}|X_0 = x\Big) \tag{38.42b}$$

$$= E(N_x|X_0 = x). \tag{38.42c}$$

Voyons ce passage plus en détail. D'abord, en général nous avons

$$E(Y|X = x_0) = \int_{\{X=x_0\}} Y(\omega)dP(\omega) = \int_{\Omega} \mathbb{1}_{\{X=x_0\}}(\omega)Y(\omega)dP(\omega). \tag{38.43}$$

Dans notre cas,

$$E\big(\mathbb{1}_{\{X_n=x\}}|X_0 = x\big) = \int_{\Omega} \mathbb{1}_{X_0=x}(\omega)\mathbb{1}_{\{X_n=x\}}(\omega)dP(\omega). \tag{38.44}$$

La fonction qui correspond à la proposiiton 14.308 est

$$f(n, \omega) = f_n(\omega) = \delta_{X_0(\omega),x}\delta_{X_n(\omega),x}, \tag{38.45}$$

qui est bien une fonction positive et mesurable.

Nous reprenons à présent le calcul (38.40) en remplaçant les éléments par leurs valeurs que nous avons calculées :

$$P(N_x < \infty|X_0 = x) = P\big(T(x) = \infty|X_0 = x\big)E(N_x|X_0 = x). \tag{38.46}$$ EqPnxXzTxarn

Si x est récurrent, nous avons $P\big(T(x) = \infty|X_0 = x\big) = 0$, mais la relation (38.46) ne permet pas de conclure que le membre de gauche est nul parce qu'il reste la possibilité que $E(N_x|X_0 = x) = \infty$. Nous devons donc faire un pas en arrière et écrire cette espérance comme la limite des sommes partielles :

$$P(N_x < \infty|X_0 = x) = \lim_{N\to\infty} \sum_{n=0}^{N} P\big(T(x) = \infty|X_0 = x\big)P(X_n = x|X_0 = x) = 0 \tag{38.47}$$

parce que tous les termes de la suite des sommes partielles sont nuls. Nous avons donc bien que $P(N_x < \infty|X_0 = x) = 0$. Il s'ensuit immédiatement que $E(N_x|X_0 = x) = 1$.

Nous devons maintenant démontrer l'implication inverse. Supposons que $P(N_x < \infty|X_0 = x) = 0$. Dans ce cas nous avons immédiatement $P(N_x = \infty|X_0 = x) = 1$ et $E(N_x|X_0 = x) = \infty$. L'équation (38.46) nous indique alors que

$$P\big(T(x) = \infty|X_0 = x\big) = 0, \tag{38.48}$$

c'est-à-dire que x est récurrent. □

38.2.4 Chaînes irréductibles

Proptoustanstousrecirrsi

Proposition 38.27.
Soit (X_n) une chaine de Markov irréductible.

(1) Un état x est récurrent si et seulement si tous les états sont récurrents.

(2) Un état x est transient si et seulement si tous les états sont transients.

Démonstration. Soient x et y des états de la chaine de Markov. Nous devons tester la valeur de $P(X_n = y|X_0 = y)$. Afin d'exploiter l'hypothèse d'irréductibilité, nous considérons $r, s \in \mathbb{N}$ tels que

$$p^r(x, y) > 0 \tag{38.49a}$$
$$p^s(y, x) > 0 \tag{38.49b}$$

et nous minorons en passant par quelques intermédiaires :

$$\begin{aligned} P(X_{n+r+s} = y|X_0 = y) &\geqslant P(X_{n+r+s} = y, X_{n+s} = x, X_s = x|X_0 = y) && (38.50a)\\ &= P(X_{n+r+s} = y|X_{n+s} = x, X_s = x, X_0 = y) && (38.50b)\\ &\quad P(X_{n+s} = x|X_s = x, X_0 = y)P(X_s = x|X_0 = y). \end{aligned}$$

Les deux premiers facteurs se calculent en utilisant la propriété de Markov et l'homogénéité de la chaine. Pour le premier,

$$P(X_{n+s} = x|X_s = x, X_0 = y) = P(X_{n+s} = x|X_s = x) = P(X_n = x|X_0 = x). \tag{38.51}$$

Nous avons donc

$$\sum_{n\in\mathbb{N}} P(X_{n+r+s} = y|X_0 = y) \geqslant p^r(x, y)p^s(y, x) \sum_{n\in\mathbb{N}} P(X_n = x|X_0 = x). \tag{38.52}$$ EqRmuXtG

En réutilisant Fubini, avec la mesure de comptage, nous permutons somme et probabilité [11] comme ceci :

$$\sum_{n\in\mathbb{N}} P(X_{n+r+s} = y|X_0 = y) \geqslant KE(N_x|X_0 = x) \tag{38.53}$$

où K est une constante strictement positive, par hypothèse d'irréductibilité de la chaine de Markov.

Si x est un état récurrent, alors le membre de gauche est infini par la proposition 38.26 et donc

$$\sum_{n\in\mathbb{N}} P(X_{n+r+s} = y|X_0 = y) = \infty. \tag{38.54}$$

Aux $r + s$ premiers termes près (qui ne changent pas la somme), nous avons

$$\sum_{n\in\mathbb{N}} P(X_n = y|X_0 = y) = \infty, \tag{38.55}$$

ce qui signifie que y est récurrent. □

Nous rappelons [12] que $T(x)$ est le temps de première atteinte de l'état x. Nous notons

$$\pi(x) = \frac{1}{E\big(T(x)|X_0 = \infty\big)}. \tag{38.56}$$ EqKyuLYk

Étant donné que $T(x)$ est un entier positif ou nul nous avons $E\big(T(x)|X_0 = x\big) \in [1, \infty]$ et donc $\pi(x) \in [0, 1]$.

Si x est un état transient, alors $T(x) = \infty$ lorsque $X_0 = x$ et donc $E\big(T(x)|X_0 = x\big) = 0$ et $\pi(x) = 0$. Si x est récurrent par contre, $P\big(T(x) < \infty|X_0 = x\big) = 1$ et il n'y a pas de garanties sur la valeur de $E\big(T(x)|X_0 = x\big)$.

CorLhpRsk

Corolaire 38.28.
Un état récurrent est récurrent positif si et seulement si $\pi(x) > 0$. Un état récurrent est récurrent nul si et seulement si $\pi(x) = 0$.

11. Un autre exemple de telle utilisation dans (38.42).
12. Définition 38.20.

Démonstration. C'est la formule (38.56). □

PropjOjDux

Proposition 38.29.
Si (X_n) est une chaine de Markov irréductible et si $x \in E$ alors

$$\pi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k=x\}} \tag{38.57}$$

EqPiEndemproprokeu

presque surement.

Démonstration. Étant donné que la chaine est irréductible, les états sont soit tous transient soit tous récurrents par la proposition 38.27. Nous commençons par considérer que x est transient.

En comparant la définition (38.38) de N_x et le membre de droite de (38.57), nous avons pour chaque n l'inégalité

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k=x\}} \leqslant \frac{1}{n} E(N_x). \tag{38.58}$$

EqkkueusnfracunENx

Dans le cas d'un élément transient, nous avons $\pi(x) = 0$, donc il serait bon de montrer que $E(N_x) < \infty$, de sorte que prendre la limite $n \to \infty$ dans (38.58) donne zéro.

Nous décomposons le calcul en deux morceaux :

$$E(N_x) = E\big(N_x|T(x) = \infty\big)P\big(T(x) = \infty\big) + E\big(N_x|T(x) < \infty\big)P\big(T(x) < \infty\big). \tag{38.59}$$

Le fait que le premier terme soit fini découle immédiatement du fait que $T(x) = \infty$ implique $X_k \neq x$ pour tout $k \geqslant 1$. Dans ce cas l'espérance de N_x est évidemment finie.

Pour le second terme nous avons

$$E\big(N_x|T(x) < \infty\big) = E\big(\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=x\}}|T(x) < \infty\big) \tag{38.60a}$$

$$= \sum_{k=1}^{\infty} E\big(\mathbb{1}_{\{X_k=x\}}|T(x) < \infty\big). \tag{38.60b}$$

Pour inverser la somme et l'espérance, nous avons utilisé le théorème de Fubini-Tonelli qui est encore valable pour des fonctions qui prennent la valeur ∞. Le fait d'inverser ne signifie pas que ni la somme ni l'intégrale soit finie. D'ailleurs c'est exactement ce que nous sommes en train de déterminer.

Étant donné que nous voulons seulement savoir si cette somme est finie ou non, nous pouvons nous restreindre à la somme depuis $k = 1$ et oublier le premier terme. D'autre part nous avons

$$\sum_{k=1}^{\infty} \mathbb{1}_{\{X_k=x\}} = \sum_{j=0}^{\infty} \mathbb{1}_{\{X_{j+T(x)}=x\}} \tag{38.61}$$

parce que les $T(x)$ premiers termes sont par définition nuls. Nous regardons donc

$$\sum_{j=0}^{\infty} E\big(\mathbb{1}_{X_{j+T(x)}=x}|T(x) < \infty\big) = \sum_{j} P\big(X_{j+T(x)} = x|T(x) < \infty\big) \tag{38.62a}$$

$$= \sum_{j} P(X_j = x|X_0 = x) \tag{38.62b}$$

subeqsumkPXjXzsezii

$$= \sum_{j} E\big(\mathbb{1}_{\{X_j=x\}}|X_0 = x\big) \tag{38.62c}$$

$$= E\big(\sum_{j} \mathbb{1}_{X_j=x}|X_0 = x\big) \tag{38.62d}$$

$$= E(N_x|X_0 = x) \tag{38.62e}$$

$$< \infty \qquad \text{parce que } x \text{ est transient.} \tag{38.62f}$$

L'équation (38.62b) provient de la proposition 38.23 et plus précisément de l'égalité entre les lois (38.37). Nous avons terminé la preuve dans le cas où x est transient.

Nous passons maintenant au cas où x est récurrent, c'est-à-dire $P(T(x) < \infty | X_0 = x) = 1$. Les variables aléatoires T_i définies en (38.28) pour $i \geqslant 2$ sont indépendantes et identiquement distribuées et

$$\mathscr{L}\big(T_k(x)\big) \sim \mathscr{L}\big(T(X)|X_0 = x\big). \tag{38.63}$$

La loi des grands nombres nous indique que EqlgnMarkdemked

$$\frac{S_n}{n} = \frac{T_1(x)}{n} + \frac{1}{n}\sum_{k=2}^{n} T_k(x) \xrightarrow{p.s.} E\big(T_2(x)\big) \tag{38.64a}$$

$$= E\big(T(x)|X_0 = x\big). \tag{38.64b}$$

Nous posons pour $m \in \mathbb{N}$

$$n(m) = \sum_{j=1}^{m} \mathbb{1}_{\{X_j = x\}} \tag{38.65}$$ EqDennmsumjmtu

qui est le nombre de visites de x avant l'instant m. Nous avons évidemment $n(m) \leqslant m$. Mais S_n est l'instant de la $n$ième visite, par conséquent $S_{n(m)}$ est l'instant de la dernière visite avant le moment m. Pour tout m nous avons les inégalités

$$S_{n(m)} \leqslant m < S_{n(m)+1}. \tag{38.66}$$

Nous divisons par $n(m)$ et nous effectuons la limite $m \to \infty$:

$$\frac{S_n(m)}{n(m)} \leqslant \frac{m}{n(m)} \leqslant \frac{S_{n(m)} + 1}{n(m)} \tag{38.67}$$ EqdrambSum

En ce qui concerne la limite de $n(m)$, nous utilisons la définition (38.65) :

$$\lim_{m\to\infty} n(m) = \lim_{m\to\infty} \sum_{n=1}^{m} \mathbb{1}_{\{X_j = x\}} = \sum_{j=1}^{\infty} \mathbb{1}_{\{X_j = x\}} \xrightarrow{p.s.} \infty \tag{38.68}$$

par la proposition 38.26. Plus précisément, la limite vaut N_x qui vaut presque surement ∞ dans le cas où x est récurrent. Par ailleurs la loi des grands nombres (38.64) nous enseigne en particulier que

$$\frac{S_{n(m)}}{n(m)} \xrightarrow{p.s.} E\big(T(x)|X_0 = x\big). \tag{38.69}$$

Le terme de droite dans (38.67) se traite de façon usuelle :

$$\frac{S_{n(m)+1}}{n(m)} = \frac{S_{n(m)+1}}{n(m)+1}\frac{n(m)+1}{n(m)}. \tag{38.70}$$

Le dernier facteur tend vers 1 et le tout a pour limite $E\big(T(x)|X_0 = x\big)$. Par conséquent nous avons

$$\frac{m}{n(m)} \xrightarrow{p.s.} E\big(T(x)|X_0 = x\big) \tag{38.71}$$

et

$$\frac{n(m)}{n} = \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{\{X_j = x\}} \to \frac{1}{E\big(T(x)|X_0 = x\big)} = \pi(x). \tag{38.72}$$

□

La proposition suivante nous permet de parler de chaine de Markov **récurrence positive**. PropUyLCzp

Proposition 38.30.
Soit (x_n) une chaine de Markov irréductible.

(1) Un état x est transient si et seulement si tous les états sont transients.

(2) Un état est récurrent positif si et seulement si tous les états sont récurrents positifs.

Démonstration. Nous rappelons (proposition 38.29) que si la chaine est irréductible

$$\pi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{[X_k=x]} \tag{38.73}$$

Notons aussi que

$$\sum_{k=1}^{N} \mathbb{1}_{X_k=x} = \begin{cases} 0 & \text{si } N < T(x) \\ \sum_{k=0}^{N-T(x)} \mathbb{1}_{X_{k+T(x)}=x} & \text{si } N > T(x) \end{cases} \tag{38.74}$$

où dans la seconde ligne nous avons effectué le changement de variable de sommation $k' = k+T(x)$. Dans la limite (38.73) nous sommes toujours dans le cas où N est assez grand. Nous pouvons donc écrire

$$\pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-T(x)} \mathbb{1}_{X_{k+T(x)}=x}. \tag{38.75}$$

Nous pouvons aussi écrire

$$\frac{1}{N-T(x)} \sum_{k=0}^{N-T(x)} \mathbb{1}_{X_{k+T(x)}=x} = \frac{N}{N-T(x)} \frac{1}{N} \sum_{k=0}^{N-T(x)} \mathbb{1}_{X_{k+T(x)}=x}. \tag{38.76}$$

Dans cette dernière égalité le membre de droite tend vers $\pi(x)$ et nous avons

$$\lim_{N\to\infty} \frac{1}{N-T(x)} \sum_{k=0}^{N-T(x)} \mathbb{1}_{X_{k+T(x)}=x} = \pi(x) \tag{38.77}$$

ou encore

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N} \mathbb{1}_{X_{k+T(x)}=x} = \pi(x) \tag{38.78}$$

Étant donné que $\pi(x)$ est une constante nous avons évidemment $E(\pi(x)) = \pi(x)$. Nous pouvons cependant considérer les variables aléatoires

$$Z_n = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{X_{k+T(x)}=x} \tag{38.79}$$

et remarquer que $Z_n \xrightarrow{p.s.} \pi(x)$ avec $0 \leqslant Z_n \leqslant 1$. Le théorème de la convergence dominée (14.207) nous permet d'inverser la limite et l'espérance et écrire

$$\pi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} E\big(\mathbb{1}_{X_{k+T(x)}=x}\big) \tag{38.80a}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} P\big(X_{k+T(x)} = x\big). \tag{38.80b}$$

Par le lemme 38.21 nous avons

$$P(X_{k+T(x)} = x) = P(X_k = k | X_0 = x) \tag{38.81}$$

et $\pi(x)$ prend la forme

$$\pi(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} P(X_k = x | X_0 = x). \tag{38.82}$$

Soit maintenant un état x positif récurrent et y, un autre état. Par définition 38.24 et par corolaire 38.28 nous avons $\pi(x) > 0$. Nous devons prouver que $\pi(y) > 0$.

Étant donné que la chaine est irréductible il existe r et s tels que

$$\begin{cases} p^r(x,y) = P(X_r = y | X_0 = x) > 0 & \text{(38.83a)} \\ p^s(x,y) = P(X_s = x | X_0 = y) > 0 & \text{(38.83b)} \end{cases}$$

Nous reprenons l'équation (38.52) multipliée par $1/N$:

$$\frac{1}{N}\sum_{n=1}^{N} P(X_{r+s+n=y|X_0=y}) \geqslant \underbrace{p^r(x,y)p^s(y,x)}_{\geqslant 0} \underbrace{\frac{1}{N}\sum_{n=1}^{N} P(X_n = x | X_0 = x)}_{\to\pi(x)} \tag{38.84}$$

et nous prenons la limite lorsque $N \to \infty$. À $r+s$ termes près, nous trouvons à gauche l'expression (38.82) de $\pi(y)$. Par conséquent

$$\pi(y) \geqslant \lim_{N\to\infty} \frac{1}{n}\sum_{n=1}^{N} P(X_{r+s+n} = y | X_0 = y) \geqslant \alpha\pi(x) \tag{38.85}$$

où α est une constante positive. Le nombre $\pi(x)$ étant strictement positif par hypothèse nous avons montré que $\pi(y) > 0$, c'est-à-dire que y est récurrent positif. □

38.2.5 Périodique et irréductible

Nous voudrions savoir sous quelles conditions la variable aléatoire X_n converge en loi vers quelque chose lorsque $n \to \infty$. Une telle loi limite doit dépendre de la loi initiale [13] comme le montre l'exemple de la chaine de Markov

$$1 \circlearrowleft A \xleftarrow{1/2} C \xrightarrow{1/2} B \circlearrowright 1 \tag{38.86}$$

Si $X_0 = C$, alors la loi limite est

$$\frac{1}{2}(\delta_A + \delta_B). \tag{38.87}$$

Si par contre $X_0 = B$, la loi limite est δ_B. Notons que la chaine de Markov proposée ici est irréductible.

Notons qu'il n'y a pas toujours de lois limite comme le montre l'exemple

$$A \underset{1}{\overset{1}{\rightleftarrows}} B \tag{38.88}$$

avec $X_0 = A$. La loi en est

$$X_k = \begin{cases} \delta_A & \text{si } k \text{ est pair} \\ \delta_B & \text{si } k \text{ est impair.} \end{cases} \tag{38.89}$$

Lemme 38.31.
Si nous avons une loi limite

$$P(X_n = x) \to l(x), \tag{38.90}$$

EqnbQMAY

et que la chaine est irréductible, alors nous avons $l = \pi$.

Démonstration. D'après la proposition 38.29 nous avons

$$\frac{1}{n}\sum_{k=1}^{n} P(X_k = x) \to \pi(x). \tag{38.91}$$

13. Lorsque la loi limite ne dépend pas de la loi initiale, nous disons que la chaine de Markov est ergodique, nous y reviendrons.

Par le lemme 10.142 sur la moyenne de Cesàro et l'hypothèse (38.90), nous avons aussi

$$\frac{1}{n}\sum_{k=1}^{n} P(X_k = x) \to l(x). \tag{38.92}$$

Du coup $\pi(x) = l(x)$. □

Lemme 38.32 ([792]).
Si π est une loi stationnaire et si x est un étant transient, alors $\pi(x) = 0$.

Ce lemme (qui peut être prouvé rigoureusement) est principalement dû au fait que la chaine de Markov ne visite un état transitoire qu'un nombre fini de fois par la proposition 38.26(1).

DefCxvOaT

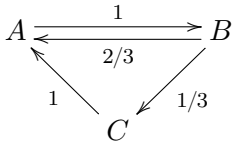
Définition 38.33.
*Un état $x \in E$ est **apériodique** si*

$$\text{pgcd}\{n \geqslant 1 \text{ tel que } p^n(x,x) > 0\} = 1. \tag{38.93}$$

Mettons que tous les n tels que $p^n(x,x) > 0$ ont 2 comme diviseur. L'état n'est alors pas apériodique, mais on voit que si $X_0 = x$, alors les états impairs ne peuvent pas être sur x. Cela est une forme de périodicité.

Si un état est apériodique, il existe r et s premiers entre eux tels que $p^r(x,x)$ et $p^s(x,x)$ sont non nuls. En particulier pour tout $n \in r\mathbb{N} + s\mathbb{N}$, $P(X_n = x) \neq 0$. Par conséquent la proposition 1.262 nous indique qu'à partir d'un certain moment tous les X_k pourraient être x.

L'état C de la chaine de Markov suivante est apériodique :



(38.94)

En effet $p^3(C,C) \neq 0$ par le chemin $C \to A \to B \to C$ tandis que $p^5(C,C) \neq 0$ également par le chemin $C \to A \to B \to A \to B \to C$. Or $\text{pgcd}\{3,5\} = 1$.

PropSaOysS

Proposition 38.34 ([792]).
Soit (X_n), une chaine de Markov irréductible. Un état x est apériodique si et seulement si il existe N tel que

$$p^k(x,x) = P(X_k = x | X_0 = x) > 0 \tag{38.95}$$

pour tout $k \geqslant N$.

DEFooBNBVooBuWQqi

Proposition-Définition 38.35.
Si une chaine de Markov est irréductible, alors un état est apériodique si et seulement si tous les états sont apériodiques.

*Dans ce cas nous disons que la chaine est **apériodique**.*

Démonstration. Soit x un état apériodique de la chaine de Markov $(X_n)_{n\in\mathbb{N}}$. En vertu de la proposition 38.34 il existe N_x tel que $p^k(x,x) \neq 0$ pour tout $k \geqslant N_x$. Soit $y \in E$. Étant donné que la chaine est irréductible, il existe r et s tels que $p^r(x,y) > 0$ et $p^s(y,x) > 0$. Nous avons

$$p^{k+r+s}(y,y) = P(X_{k+r+s} = y | X_0 = y) \geqslant p^s(x,y)P(X_k = x | X_0 = x)p^r(y,x). \tag{38.96}$$

Si k est assez grand, cette quantité est strictement positive. Donc il suffit de prendre $N_y = N_x + r + s$ pour savoir que y est également apériodique. □

Exemple 38.36.
Quelle est la différence entre une chaine irréductible et une chaine apériodique ? Une chaine est irréductible lorsque aucune sous-chaine ne peut piéger le système. Pour toute paire d'états $x, y \in E$,

il existe un n tel qu'il soit possible d'aller de x à y en n pas. Une chaine est apériodique lorsqu'après un temps suffisamment long, *tous* les états soient possibles en même temps.

Un exemple de chaine irréductible non apériodique :

$$A \underset{1}{\overset{1}{\rightleftarrows}} B \tag{38.97}$$

Cette chaine est irréductible parce que le graphe est connexe, par contre il n'est pas apériodique parce que si $X_0 = A$ il n'est pas possible d'être dans l'état A après un nombre impair de pas.

Plus formellement, $p^n(A, A) = 1$ dès que n est pair ; le PGCD de la définition 38.33 n'est donc certainement pas 1. △

ThoQSuLZoz

Théorème 38.37 ([792]).
Si (X_n) est une chaine de Markov finie, irréductible[14] *et apériodique*[15] *de loi stationnaire π, alors*

(1) La suite de matrices stochastiques p^k converge vers la matrice

$$p^k \to \Pi = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}. \tag{38.98}$$

(2) Nous avons convergence des variables aléatoires au sens où

$$P(X_k = \mu p^k) \to \pi. \tag{38.99}$$

38.3 Marche aléatoire sur $\mathbb{Z}$

Soit (Y_n) une suite de variables aléatoires indépendantes et identiquement distribuées valant -1 avec une probabilité p et 1 avec une probabilité $(1-p)$. La loi est

$$Y_n \sim p\delta_{-1} + (1-p)\delta_1. \tag{38.100}$$

Nous considérons la variable aléatoire

$$X_n = X_0 + \sum_{i=1}^{n} Y_i \tag{38.101}$$

où X_0 est une variable aléatoire indépendante des Y_i à valeurs dans $\mathbb{Z}$. Nous vérifions à présent que X_n est une chaine de Markov avec comme espace d'états $E = \mathbb{Z}$. Nous devons montrer que

$$P\big(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0\big) = P\big(X_{n+1} = x_{n+1} | X_n = x_n\big). \tag{38.102}$$

EqAVoirMarkovMAZ

Pour ce faire nous allons exprimer tout cela en termes des Y_i au lieu des X_i. D'abord étant donné que nous avons égalité des événements

$$\{X_{n+1} = x_{n+1}\} \cap \{X_n = x_n, \dots, X_0 = x_0\} = \{Y_{n+1} = x_{n+1} - x_n\} \cap \{X_n = x_n, \dots, X_0 = x_0\}, \tag{38.103}$$

nous pouvons, en vertu du principe (36.126), remplacer $X_{n+1} = x_{n+1}$ par $Y_{n+1} = x_{n+1} - x_n$ dans le membre de gauche de (38.102). Nous avons donc déjà

$$P\big(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0\big) = P\big(\underbrace{Y_{n+1} = x_{n+1} - x_n}_{A} \,|\, \underbrace{X_n = x_n, \dots, X_0 = x_0}_{B}\big). \tag{38.104}$$

14. Irréductible, définition 38.15.
15. Définition 38.35.

L'événement B est égal à l'événement

$$\{X_0 = x_0, Y_1 = x_1 - x_0, Y_2 = x_2 - x_1, \dots, Y_n = x_n - x_{n-1}\}, \tag{38.105}$$

qui n'est autre que l'ensemble

$$X_0^{-1}(x_0) \cap Y_1^{-1}(x_1 - x_0) \cap \dots \cap Y_n^{-1}(x_n - x_{n-1}) \tag{38.106}$$

qui est dans la tribu engendrée par les variables aléatoires $X_0, (Y_i)_{i=1,\dots,n}$. Le point délicat du raisonnement est de montrer que les événements A et B donnés par

$$A = \{Y_{n+1} = x_{n+1} - x_n\} \tag{38.107a}$$

$$B = \{X_0 = x_0\} \cap \bigcap_{i=1}^{n} \{Y_i = x_i - x_{i-1}\} \tag{38.107b}$$

sont indépendants. Nous ne pouvons pas montrer directement que $P(A \cap B) = P(A)P(B)$ parce que cela est la formule que nous voulons utiliser pour montrer que la chaine est de Markov. Nous passons donc par les tribus : [subesqqsABtribsYYXzY]

$$A \in \sigma(Y_{n+1}) \tag{38.108a}$$

$$B \in \sigma(X_0, Y_1, \dots, Y_n). \tag{38.108b}$$

Nous utilisons maintenant l'hypothèse d'indépendance des variables aléatoires X_0 et Y_i pour conclure que les deux tribus des équations (38.108) sont indépendantes. Les événements A et B sont par conséquent indépendants.

L'événement A est indépendant de l'événement $\{X_n = x_n\}$. Nous avons donc successivement

$$\begin{aligned}
&P\big(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0\big) \\
&\qquad = P\big(Y_{n+1} = x_{n+1} - x_n | X_n = x_n, \dots, X_0 = x_0\big) && && \text{(38.109a)}\\
&\qquad = P\big(Y_{n+1} = x_{n+1} - x_n | Y_i = x_i - x_{i-1}, X_0 = x_0\big) && && \text{(38.109b)}\\
&\qquad = P(Y_{n+1} = x_{n+1} - x_n) && \text{cf. justif.} && \text{(38.109c)}\\
&\qquad = P(Y_{n+1} = x_{n+1} - x_n | X_n = x_n) && \text{cf. justif.} && \text{(38.109d)}\\
&\qquad = P(Y_{n+1} = x_{n+1} - X_n | X_n = x_n) && \text{cf. justif.} && \text{(38.109e)}\\
&\qquad = P(X_{n+1} = x_{n+1} | X_n = x_n). && && \text{(38.109f)}
\end{aligned}$$

Justifications :

— (38.109c) parce que les tribus $\sigma(Y_{n+1})$ et $\sigma(Y_i, X_0)$ sont indépendantes.

— (38.109d) Nous avons

$$\{X_n = x_n\} \in \sigma(X_0, Y_1, \dots, Y_n) \tag{38.110}$$

tandis que

$$\{Y_{n+1} = x_{n+1} - x_n\} \in \sigma(Y_{n+1}); \tag{38.111}$$

ce sont donc deux événements issus de tribus indépendantes. Donc conditionner ou non l'événement $Y_{n+1} = x_{n+1} - x_n$ à l'événement $X_n = x_n$ ne change rien.

— (38.109e) est encore l'utilisation du fait que $P(A|B) = P(K|B)$ dès que $A \cap B = K \cap B$.

La chaine est par conséquent de Markov.

La matrice de transition de cette chaine de Markov est une matrice infinie « dans tous les sens » :

$$p(x, y) = \begin{cases} p & \text{si } y = x - 1 \\ (1 - p) & \text{si } y = x + 1 \\ 0 & \text{sinon.} \end{cases} \tag{38.112}$$

Remarque 38.38.
La plupart du temps lorsqu'il faut démontrer qu'une chaine est de Markov, il faut suivre la procédure que nous venons de suivre pour la marche aléatoire sur $\mathbb{Z}$.

— Écrire tout en fonction des incréments.

— Dire que les incréments conditionnés sont indépendants des incréments qui conditionnent (via les tribus engendrées).

— Écrire que la probabilité cherchée est égale à l'événement conditionné dans lequel on a juste remplacé l'incrément par sa valeur.

— Conditionner à nouveau par rapport au dernier incrément qui est indépendant.

— Changer la valeur du dernier incrément par la variable aléatoire.

Dans ce raisonnement nous utilisons deux fois le fait que $P(A|B) = P(K|B)$ si $A \cap B = K \cap B$.

38.3.1 Chaînes de Markov homogènes

PROPooYIDWooAKTVvS

Proposition 38.39.
Voici quelques propriétés des chaines de Markov homogènes[16].

ITEMooSDDUooVRnpjv

(1) La probabilité d'une trajectoire donnée est

$$P(X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = p(x_{n-1}, x_n) \dots p(x_0, x_1) P(X_0 = x_0) \tag{38.113}$$

où les $p(x, y)$ sont les probabilités de transitions introduits dans la définition 38.4.

ITEMooJZNRooXFQTQc

(2) La probabilité de transition « en n coups » est donnée par la puissance n^e de la matrice de transition :

$$P(X_n = x_n | X_0 = x_0) = Q^n_{x_0, x_n}. \tag{38.114}$$

ITEMooJUEMooWXEkBO

(3) Si l'espace des états E est fini, l'espérance d'une fonction bornée[17] *$f : E \to \mathbb{R}$ de l'état est donnée par*

$$E\big(f(X_{n+1}) | X_n = x_n, \dots, X_0 = x_0\big) = E\big(f(X_{n+1}) | X_n = x_n\big) = \sum_{y \in E} f(y) p(x_n, y) \tag{38.115}$$

pour tout $x_0, \dots, x_n \in E$.

Pour être précis, ce que nous notons « $f(X_{n+1})$ » est la composée $f \circ X_{n+1}$.

Démonstration. En plusieurs étapes.

(1) **Pour (1)** Nous écrivons la formule $P(A \cap B) = P(A|B)P(B)$ pour les événements $A = \{X_n = x_n\}$ et $B = \bigcap_{i=0}^{n-1} \{X_i = x_i\}$:

$$P(X_n = x_n, \dots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) P(X_{n-1} = x_{n-1}, \dots, X_0 = x_0). \tag{38.116}$$

Par la propriété de Markov, le premier facteur est

$$P(X_n = x_n | X_{n-1} = x_{n-1}) = p(x_{n-1}, x_n). \tag{38.117}$$

Le reste est une récurrence sur n.

16. Définition 38.4.
17. L'hypothèse de borne sur f n'est pas très chère parce que E est fini. Il suffit que f ne soit infinie en aucun point.

(2) **Pour (2)** Montrons avec $n = 2$. En utilisant les divers points du théorème 36.57, nous avons

$$P(X_2 = x_2|X_0 = x_0) = \sum_{y\in E} P(X_2 = x_2, X_1 = y|X_0 = x_0) \tag{38.118a}$$
$$= \sum_{y\in E} P(X_2 = x_2|X_1 = y, X_0 = x_0)P(X_1 = y|X_0 = x_0) \tag{38.118b}$$
$$= \sum_{y\in E} P(X_2 = x_2|X_1 = y)P(X_1 = y|X_0 = x_0) \tag{38.118c}$$
$$= \sum_{y\in E} p(y, x_2)p(x_0, y) \tag{38.118d}$$
$$= Q^2_{x_2,x_0}. \tag{38.118e}$$

subEqnExdyyxz

Nous avons utilisé l'homogénéité de la chaine de Markov au moment d'écrire l'expression (38.118d). En principe nous aurions dû écrire $p_2(y, x_2)p_1(x_0, y)$.

(3) **Pour (3)** Vu que E est fini, $f(E)$ est fini et nous notons $\{a_k\}_{k=1,\dots,N}$ l'ensemble des valeurs non nulles (dans $\mathbb{R}$) atteintes par f. Nous utilisons le lemme 36.76 pour la variable aléatoire [18] $f \circ X_{n+1} \colon \Omega \to \{0, a_k\}_{k=1,\dots,N}$:

$$E\big(f(X_{n+1})|X_n = x_n, \dots, X_0 = x_0\big) = \sum_{k=1}^{N} a_k P\big(f \circ X_{n+1} = a_k|X_n = x_n, \dots, X_0 = x_0\big). \tag{38.119}$$

EQooLRZJooJGceeV

L'événement $f \circ X_{n+1} = a_k$ signifie $X_{n+1} \in f^{-1}(a_k)$ ou encore

$$\{\omega \in \Omega \text{ tel que } X_{n+1}(\omega) \in f^{-1}(a_k)\}. \tag{38.120}$$

Vu que E est fini, l'ensemble $f^{-1}(a_k)$ est fini et nous écrivons

$$f^{-1}(a_k) = \{y_{k,1}, \dots, y_{k,N_k}\}. \tag{38.121}$$

Nous avons

$$\{\omega \in \Omega \text{ tel que } X_{n+1}(\omega) \in f^{-1}(a_k)\} = \bigcup_{i=1}^{N_k} \{\omega \in \Omega \text{ tel que } X_{n+1}(\omega) = y_{k,i}\}. \tag{38.122}$$

Nous pouvons utiliser le lemme 36.59 pour décomposer

$$P(f \circ X_{n+1} = a_k|\cdots) = \sum_{i=1}^{N_k} P(X_{n+1} = y_{k,i}|\cdots) \tag{38.123}$$

et continuer (38.119) pas

$$E\big(f(X_{n+1})|X_n = x_n, \dots, X_0 = x_0\big) = \sum_{k=1}^{N} a_k \sum_{i=1}^{N_k} P(X_{n+1} = y_{k,i}|X_n = x_n, \dots, X_0 = x_0) \tag{38.124a}$$
$$= \sum_{k=1}^{N} a_k \sum_{i=1}^{N_k} P(X_{n+1} = y_{k,i}|X_n = x_n) \tag{38.124b}$$
$$= \sum_{k=1}^{N} a_k P\big(f \circ X_{n+1} = a_k|X_n = x_n\big) \tag{38.124c}$$
$$= E\big(f(X_{n+1})|X_n = x_n\big). \tag{38.124d}$$

SUBEQooOWSUooTlUXAs

Pour (38.124b), nous avons utilisé la propriété de Markov.

□

18. La composée de fonctions mesurables est mesurable, proposition 14.45. De plus f est mesurable parce que E étant dénombrable, nous y mettons la tribu de toutes les parties.

38.3.2 Chaîne de Markov définie par récurrence

38.3.2.1 Le cas général

PropqiMdHh

Proposition 38.40.
Soit X_0 une variable aléatoire à valeurs dans E, un ensemble au plus dénombrable. Soit (Y_n) une suite de variables aléatoires réelles indépendantes et identiquement distribuées indépendantes de X_0.

Soit (X_n) la suite de variables aléatoires à valeurs dans E définie par récurrence selon la formule

$$X_{n+1} = G(X_n, Y_{n+1}) \tag{38.125}$$

où $G\colon E \times \mathbb{R} \to E$ est une fonction mesurable. Alors (X_n) est une chaine de Markov.

Démonstration. Soient $x_0, \dots, x_{n+1}$ des éléments de E. Nous devons calculer la valeur de

$$P(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0). \tag{38.126}$$

Commençons par préciser les espaces sur lesquels nos variable aléatoires sont définies. Nous avons

$$X_0\colon \Omega_0 \to E \tag{38.127}$$

et

$$Y_i\colon \Omega \to \mathbb{R}. \tag{38.128}$$

La variable aléatoire X_1 est donnée par

$$\begin{aligned} X_1\colon \Omega_0 \times \Omega &\to E \\ (\omega_0, \omega_1) &\mapsto G\big(X_0(\omega_0), Y_1(\omega_1)\big). \end{aligned} \tag{38.129}$$

La variable aléatoire X_2 est

$$\begin{aligned} X_2\colon \Omega_0 \times \Omega^2 &\to E \\ (\omega_0, \omega_1, \omega_2) &\mapsto G\big(X_1(\omega_0, \omega_1), Y_2(\omega_2)\big) \\ &= G\Big(G\big(X_0(\omega_0), \omega_1\big), Y_2(\omega_2)\Big) \end{aligned} \tag{38.130}$$

et ainsi de suite.

Considérons maintenant l'événement

$$\{X_1 = x_1, X_0 = x_0\} \subset \Omega_0 \times \Omega. \tag{38.131}$$

Il est donné explicitement par

$$\{X_1 = x_1, X_0 = x_0\} = \{(\omega_0, \omega_1) \text{ tel que } G\big(X_0(\omega_0), Y_1(\omega_1)\big) = x_1, X_0(\omega_0) = x_0\} \tag{38.132a}$$
$$= \{(\omega_0, \omega_1) \text{ tel que } G\big(x_0, Y_1(\omega_1) = x_1\big), X_0(\omega_0) = x_0\} \tag{38.132b}$$
$$= \{\omega_0 \in \Omega_0 \text{ tel que } X_0(\omega_0) = x_0\} \times \{\omega_1 \in \Omega \text{ tel que } G\big(x_0, Y_1(\omega_1)\big) = x_1\}. \tag{38.132c}$$

Le premier terme du produit cartésien est dans $\sigma(X_0)$, tandis que le second est dans $\sigma(Y_1)$. Étant donné la définition des tribus produit (définition 14.127) nous avons

$$\{X_1 = x_1, X_0 = x_0\} \in \sigma(X_0, Y_1). \tag{38.133}$$

Ce raisonnement se généralise immédiatement et nous trouvons que

$$\{X_n = x_n, \dots, X_0 = x_0\} \in \sigma(X_0, Y_1, \dots, Y_n). \tag{38.134}$$

Nous en sommes donc à calculer

$$\begin{aligned}\diamondsuit &= P(X_{n+1}=x_{n+1}|X_n=X_n,\dots,X_0=x_0) && (38.135a)\\ &= P\big(\underbrace{G(x_n,Y_{n+1})=x_{n+1}}_{\in\sigma(Y_{n+1})}\,|\,\underbrace{X_n=x_n,\dots,X_0=x_0}_{\in\sigma(X_0,Y_1,\dots,Y_n)}\big). && (38.135b)\end{aligned}$$

Les tribus $\sigma(Y_{n+1})$ et $\sigma(X_0,Y_1,\dots,Y_n)$ étant indépendantes nous avons

$$\begin{aligned}\diamondsuit &= P\big(G(x_n,Y_{n+1})=x_{n+1}\big) && (38.136a)\\ &= P\big(G(x_n,Y_{n+1})=x_{n+1}|X_n=x_n\big) && (38.136b)\\ &= P\big(G(X_n,Y_{n+1})=x_{n+1}|X_n=x_n\big) && (38.136c)\\ &= P(X_{n+1}=x_{n+1}|X_n=x_n). && (38.136d)\end{aligned}$$

Pour (38.136b) nous avons utilisé le fait que $\sigma(Y_{n+1})$ est indépendante de $\sigma(X_n)$. Nous avons prouvé que la chaine était de Markov. □

Les probabilités de transition de la chaine de Markov définie dans la proposition 38.40 sont

$$P(X_1=y|X_0=x)=P\big(G(X_0,Y_1)=y|X_0=x_0\big)=P\big(G(x_0,Y_1)=y\big). \tag{38.137}$$

Proposition 38.41 ([793]).
Soient :

1. *Un ensemble F, un ensemble au plus dénombrable E,*
2. *Des variables aléatoires indépendantes et identiquement distribuées $Z_n\colon\Omega\to F$,*
3. *Une variable aléatoire $X_0\colon\Omega\to E$, indépendante des Z_n*
4. *Une application mesurable $f\colon E\times F\to E$.*

Alors la définition par récurrence

$$\begin{aligned}X_{n+1}\colon\Omega&\to E\\ \omega&\mapsto f\big(X_n(\omega),Z_{n+1}(\omega)\big)\end{aligned} \tag{38.138}$$

est une chaine de Markov homogène dont la matrice de transition est donnée par

$$P_{ij}=P\big(f(i,Z_1)=j\big). \tag{38.139}$$

Démonstration. Nous commençons par construire des applications $g_n\colon E\times F^n\to E$ telles que $X_n(\omega)=g_n\big(X_0(\omega),Z_1(\omega),\dots,Z_n(\omega)\big)$. D'abord nous posons $g_1=f$. Ensuite, pour la récurrence nous posons

$$\begin{aligned}g_{n+1}\colon E\times F^{n+1}&\to E\\ (i,x_1,\dots,x_{n+1})&\mapsto f\big(g_n(i,x_1,\dots,x_n),x_{n+1}\big).\end{aligned} \tag{38.140}$$

Nous vérifions que c'est bon :

$$\begin{aligned}g_{n+1}(i,x_1,\dots,x_{n+1})&=f\big(g_n(i,x_1,\dots,x_n),x_{n+1}\big) && (38.141a)\\ &=f\big(X_n(\omega),Z_{n+1}(\omega)\big) && (38.141b)\\ &=X_{n+1}(\omega). && (38.141c)\end{aligned}$$

Et maintenant il faut vérifier que les X_n forment une chaine de Markov homogène. C'est un bon calcul, et nous commençons par une étape simple :

$$\begin{aligned}\clubsuit=P(X_{n+1}=j|X_n=x_n,\dots,X_0=x_0)&=P\big(f(X_n,Z_{n+1})=j|X_n=x_n,\dots,X_0=x_0\big) && (38.142a)\\ &=P\big(f(x_n,Z_{n+1})=j|X_n=x_n,\dots,X_0=x_0\big) && (38.142b)\end{aligned}$$

Pour la seconde ligne, nous avons utilisé le lemme 36.60. Pour la suite, nous nous souvenons que $X_k = g_k(X_0, Z_1, \dots, Z_k)$. Nous pouvons exprimer le conditionnement de la façon suivante :

$$\begin{aligned} B &= \{\omega \in \Omega \text{ tel que } X_n(\omega) = x_n, \dots, X_0(\omega) = x_0\} && (38.143a)\\ &= \bigcap_{k=1}^{n} \{X_k = x_k\} \cap \{X_0 = x_0\} && (38.143b)\\ &= \bigcap_{k=1}^{n} \{g_k(X_0, Z_1, \dots, Z_k)\} \cap \{X_0 = x_0\} && (38.143c)\\ & && (38.143d) \end{aligned}$$

Maintenant, en considérant l'application

$$\begin{aligned} g\colon E \times F^n &\to E^{n+1}\\ \big(i, (z_1, \dots, z_n)\big) &\mapsto \begin{pmatrix} i \\ g_1(i, z_1) \\ \vdots \\ g_n(i, z_1, \dots, z_n) \end{pmatrix}, \end{aligned} \tag{38.144}$$

nous avons

$$B = \{\omega \in \Omega \text{ tel que } g\big(X_0(\omega), Z_1(\omega), \dots, Z_n(\omega)\big) = (x_0, x_1, \dots, x_n)\}. \tag{38.145}$$

Nous avons

$$\clubsuit = P\big(f(x_n, Z_{n+1}) = j | g(X_0, Z_1, \dots, Z_n) = (x_0, \dots, x_n)\big). \tag{38.146}$$

Le lemme 36.62 dit que les variables aléatoires $f(X_n, Z_{n+1})$ et $g(X_0, Z_1, \dots, Z_n)$ sont indépendantes. Donc les événements $f(x_n, Z_{n+1}) = j$ et $g(X_0, Z_1, \dots, Z_n) = (x_0, \dots, x_n)$ sont indépendants par la proposition 36.15(2).

Nous avons donc montré que

$$\clubsuit = P(X_{n+1} = j | X_n = x_n, \dots, X_0 = x_0) = P\big(f(x_n, Z_{n+1}) = j\big). \tag{38.147}$$

Des calculs similaires[19] montrent que

$$P(X_{n+1} = j | X_n = x_n) = P\big(f(x_n, Z_{n+1}) = j\big). \tag{38.148}$$

Nous avons donc bien une chaine de Markov.

Nous montrons à présent que cette chaine est homogène en montrant que les probabilités de transitions $P(X_{n+1} = j | X_n = i)$ ne dépendent pas de n. Posons

$$A_n = \{\omega \in \Omega \text{ tel que } f\big(i, Z_{n+1}(\omega)\big) = j\}. \tag{38.149}$$

Nous voudrions prouver que $P(A_n) = P(A_1)$ pour tout n. Les variables aléatoires Z_n et Z_1 sont indépendantes et identiquement distribuées ; utilisant le lemme 36.18 avec

$$\begin{aligned} g\colon F &\to E\\ x &\mapsto f(i, x), \end{aligned} \tag{38.150}$$

nous voyons que les variables aléatoires $f(i, Z_n)$ et $f(i, Z_1)$ sont indépendantes et identiquement distribuées. Donc

$$P\big(f(i, Z_n) = j\big) = P\big(f(i, Z_1) = j\big), \tag{38.151}$$

et la chaine de Markov est homogène. □

19. Je n'ai pas vérifié ; écrivez-moi si il y a un problème.

LEMooQBOIooMNbJlV

Lemme 38.42.
Si P est une matrice stochastique[20] *$N \times N$ ainsi qu'un ensemble E de cardinal N, alors il existe une chaine de Markov dont la matrice de transition est P.*

Démonstration. Nous supposons que $E = \{1, \dots, N\}$. Si E est un ensemble quelconque, il faut considérer une bijection[21] $\varphi\colon E \to \{1, \dots, N\}$ et écrire $\varphi(i)$ partout au lieu de i.

Nous considérons des variables aléatoires $Z_n\colon \Omega \to [0,1]$ indépendantes et uniformes sur $[0,1]$. Nous considérons aussi une variable aléatoire $X_0\colon \Omega \to E$, indépendante des Z_n.

Ensuite nous construisons les variables aléatoires $X_n\colon \Omega \to E$ par récurrence en posant

$$X_{n+1}(\omega) = j \quad \text{si} \sum_{k=1}^{j-1} P_{X_n(\omega),k} \leqslant Z_{n+1}(\omega) < \sum_{k=1}^{j} P_{X_n(\omega),k}. \tag{38.152}$$

Notons que pour $j = 0$, la première somme est « vide » et vaut donc zéro. La proposition 38.41 appliquée à la fonction

$$\begin{aligned} f\colon E \times [0,1] &\to E \\ (i,t) &\mapsto \begin{cases} j & \text{si } \sum_{k=1}^{j-1} P_{ik} \leqslant t < \sum_{k=1}^{j} P_{ik} \\ N & \text{si } t = 1 \end{cases} \end{aligned} \tag{38.153}$$

implique que les X_k forment une chaine de Markov homogène dont la matrice de transition est donnée par

$$P_{ij} = P\big(f(i, Z_1) = j\big). \tag{38.154}$$

Nous calculons facilement cette probabilité :

$$\{f(i,Z_1) = j\} = \{\omega \in \Omega \text{ tel que } f\big(i, Z_1(\omega)\big) = j\} \tag{38.155a}$$

$$= \{\omega \in \Omega \text{ tel que } \sum_{k=0}^{j-1} P_{ik} \leqslant Z_1(\omega) < \sum_{k=0}^{j} P_{ik}\} \tag{38.155b}$$

$$= \{\omega \in \Omega \text{ tel que } Z_1(\omega) \in [\sum_{k=1}^{j-1} P_{ik}, \sum_{k=1}^{j} P_{ik}[\}. \tag{38.155c}$$

Vu que Z_1 est uniforme sur $[0,1]$, la probabilité de cet ensemble est la mesure de Lebesgue de l'intervalle $[\sum_{k=1}^{j-1} P_{ik}, \sum_{k=1}^{j} P_{ik}[$, c'est-à-dire P_{ij}. □

38.3.2.2 Exemple : la file de réparation de machines à laver

38.43.
David Madore parle de file d'attente[794]. En bref, il est inexplicable qu'une file d'attente soit durablement longue sans s'allonger à l'infini.

Nous considérons un magasin de réparation d'électroménager. Durant le jour n, un nombre aléatoire Z_n de machines en panne arrivent au magasin. Une machine est réparée chaque jour (aucune si le magasin est vide). Nous supposons que les Z_n soient indépendantes et identiquement distribuées, et nous posons X_n, le nombre de machines en magasin le jour n.

La loi d'avancement de X_n est

$$X_{n+1} = \begin{cases} X_n + Z_n - 1 & \text{si } X_n \neq 0 \\ Z_n & \text{si } X_n = 0. \end{cases} \tag{38.156}$$

Cela est une chaine de Markov en vertu de la proposition 38.40. Ici la fonction est

$$G(x,y) = x + y - \mathbb{1}_{x \neq 0}. \tag{38.157}$$

20. Définition 38.9.
21. Existence par la définition 1.126 du cardinal d'un ensemble fini.

Les probabilités de transitions sont

$$p(x,y)=\begin{cases}0 & \text{si } y\leqslant x-2\\ P(Z=0) & \text{si } y=x-1\\ P(Z=k) & \text{si } y=x+k-1\end{cases} \tag{38.158}$$

pour $x\neq 0$.

38.3.2.3 Cadeau numberphile

Vous aimez les chaines de Markov ? « The Beautiful Math of Snakes and Ladders » https://www.youtube.com/watch?v=nlm07asSU0c

Et si vous aimez les échelles, autre cadeau de numberphile : « Travelling to the Stars by Ladder » https://www.youtube.com/watch?v=hzcdcrsyA9Y Ce ne sont pas des math, mais ça détend.

38.4 Classification des états

Sauf mention expresse du contraire, nous considérons toujours une chaine de Markov homogène.

DEFooFIQUooScQDMI

Définition 38.44.
*Un état $x\in E$ est **absorbant** pour la chaine (X_n) si $p(x,x)=1$.*

Il n'est pas spécialement impossible d'arriver sur un état absorbant, mais il est impossible d'en sortir.

Proposition 38.45.
En notant T le temps de première atteinte[22], a loi de la variable aléatoire $[T_p(x)|X_p=x]$ est la même que celle de la variable aléatoire $[T(x)|X_0=x]$.

Démonstration. Nous devons montrer que

$$P(T_p(x)=k|X_p=x)=P(T(x)=k|X_0=x). \tag{38.159}$$

Cela est intuitivement évident du fait qu'une chaine de Markov soit un processus sans mémoire. Afin de prouver, nous allons sommer sur tous les états intermédiaires possibles :

$$P(T_p(x)=k|X_0=x)=P(X_{p+k}=x,X_{p+k-1}\neq x,\dots,X_{p+1}\neq x|X_p=x) \tag{38.160a}$$
$$=\sum_{z_i\neq x}P(X_{p+k}=x,X_{p+k-1}=z_{k-1},\dots,X_{p+1}=z_1|X_p=x) \tag{38.160b}$$
$$=\sum_{z_i}P(X_{p+k}=x,X_{p+k-i}=z_i|X_{p+1}=z_1,X_p=x)\underbrace{P(X_{p+1}=z_1|X_p=x)}_{=p(x,z_1)} \tag{38.160c}$$
$$=\sum_{z_i}P(X_{p+k}=x,X_{p+k-i}=z_i|X_{p+2}=z_2,X_{p+1}=z_1,X_p=x) \tag{38.160d}$$
$$\underbrace{P(X_{p+2}=z_2|X_{p+1}=z_1,X_p=x)}_{P(X_{p+2}=z_2|X_{p+1}=z_1)=p(z_1,z_2)}p(x,z_1) \tag{38.160e}$$
$$=\dots \tag{38.160f}$$
$$=\sum_{z_i}p(x,z_1)p(z_1,z_2)\dots p(z_{k-2},z_{k-1})p(z_{k-1},x). \tag{38.160g}$$

À ce point ci, nous avons éliminé toute référence à p grâce à l'homogénéité de la chaine. Nous

22. Définition 38.20.

pouvons refaire le calcul à l'envers pour reconstituer l'expression de départ sans le p :

$$\sum_{z_i} p(x,z_1)p(z_1,z_2)\dots p(z_{k-1},z_{k-1})p(z_{k-1},x) \tag{38.161a}$$
$$= P(x_k = x, X_{k-1} \neq x, \dots, X_1 \neq x | X_0 = x) \tag{38.161b}$$
$$= P(T(x) = k), \tag{38.161c}$$

ce qu'il fallait obtenir. □

Exemple 38.46.
Soit (X_n) une chaine de Markov de matrice de transition

$$P = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.1 & 0.8 \\ 0.5 & 0.2 & 0.3 \end{pmatrix}. \tag{38.162}$$

Calculer $P(X_3 = 1|X_0 = 1)$ et $P(X_7 = 0|X_4 = 0)$.

Déterminer, si il en existe, une loi stationnaire vers laquelle converge la chaine.

Nous avons

$$P^3 = \begin{pmatrix} 0.344 & 0.251 & 0.405 \\ 0.283 & 0.307 & 0.41 \\ 0.287 & 0.248 & 0.465 \end{pmatrix}. \tag{38.163}$$

La probabilité d'aller de l'état 1 à l'état 1 en trois étapes est donc 0.307. La chaine étant de Markov, sans mémoire, les probabilités entre les temps 4 et 7 sont les mêmes qu'entre 0 et 3. Nous avons alors

$$P(X_7 = 0|X_4 = 0) = 0.344. \tag{38.164}$$

La chaine est irréductible et n'a pas d'états absorbants. △

Si une chaine de Markov contient r état absorbants, nous convenons de numéroter les états de telle sorte à mettre en premier les $N - r$ états non absorbants. La matrice de transition s'écrit alors sous forme **canonique** :

$$P = \begin{pmatrix} Q & R \\ 0 & \mathbb{1} \end{pmatrix} \tag{38.165}$$

EQooGJVFooNLCgue

où Q est une matrice $N - r \times N - r$.

PROPooUNBDooVhXZxi

Proposition 38.47 ([791]).
Soit une chaine de Markov contenant des états absorbants. Nous considérons la matrice de transition sous forme canonique[23]*. Alors*

(1) $\lim_{n\to\infty} Q^n = 0$.

(2) La matrice $\mathbb{1} - Q$ *est inversible et*

$$(\mathbb{1} - Q)^{-1} = \sum_{k=0}^{\infty} Q^k. \tag{38.166}$$

PROPooCAORooMEaxtq

Proposition 38.48 ([791, 1]).
Si (X_n) *est une chaine de Markov irréductible*[24] *sur un ensemble fini, alors pour toute partie* $A \subset E$ *nous avons*

$$P(\tau_A < \infty) = \lim_{n\to\infty} P(\tau_A \leqslant n) = 1 \tag{38.167}$$

où $\tau_A = \min\{k \text{ tel que } X_k \in A\}$.

23. C'est à dire sous la forme (38.165).
24. Chaine de Markov irréductible, définition 38.15.

Démonstration. Soit P la matrice de transition de la chaine (X_n). Nous considérons une nouvelle matrice stochastique

$$P'_{ij} = \begin{cases} \delta_{ij} & \text{si } i \in A \\ P_{ij} & \text{sinon.} \end{cases} \tag{38.168}$$

Nous considérons (Y_k) la chaine de Markov construite par le lemme 38.42 à partir de la matrice P, ainsi que (Y'_k) celle construite à partir de P'.

Les chaines X et Y ont la même matrice de transition, mais l'intérêt de Y est d'être définie sur le même espace que Y'. Nous allons procéder de la façon suivante :

(1) Prouver le résultat sur Y'

(2) Le déduire pour Y parce que τ_A ne change pas entre Y' et Y.

(3) Le déduire pour X parce que la matrice de transition de X et Y sont les mêmes.

Pour la suite, voir [791] en utilisant la proposition 38.47. □

38.5 Mesure invariante

Définition 38.49.
*Une mesure de probabilité μ sur l'espace des états E d'une chaine de Markov est **invariante** si pour tout $x \in E$*

$$\mu(x) = \sum_{y \in E} p(y, x)\mu(y). \tag{38.169}$$

Remarque 38.50.
Une mesure invariante est une mesure de probabilité et nous noterons par abus $\mu(x)$ pour $\mu(\{x\})$. Si $A \subset E$ nous avons

$$\mu(A) = \sum_{x \in A} \mu(x). \tag{38.170}$$

RemwcRRFZ

Remarque 38.51.
Une loi invariante associée à une chaine de Markov est une loi associée à la matrice de transition de la chaine, mais pas à la loi de X_0. Par conséquent nous pouvons tester si μ est une mesure invariante pour une certaine chaine de Markov (X_k) en considérant la chaine (Y_k) avec $Y_k = X_k$ pour $k > 0$ et Y_0 arbitraire.

L'adjectif *invariant* provient du lemme suivant.

LemUVMwbM

Lemme 38.52.
Soit (X_n) une chaine de Markov homogène[25] *telle que $X_0 \sim \mu$ où μ est une mesure invariante sur l'espace des états. Alors $X_k \sim \mu$ pour tout k.*

Démonstration. Par hypothèse, $P(X_0 = x) = \mu(x)$. Ensuite nous avons

$$\begin{aligned} P(X_1 = y) &= \sum_{x \in E} P(X_1 = y | X_0 = x)P(X_0 = x) && \text{(38.171a)} \\ &= \sum_x p(x, y)\mu(x) && \text{(38.171b)} \\ &= \mu(y). && \text{(38.171c)} \end{aligned}$$

Par conséquent X_1 suit également la loi μ. Par récurrence tous les états suivent cette même loi. □

Définition 38.53.
*Si les états d'une chaine de Markov ont comme loi une mesure invariante, alors nous disons que la chaine est **stationnaire**.*

25. Définition 38.4.

RemcOEylF

Remarque 38.54.
Pour une chaine de Markov stationnaire de loi invariante μ nous avons

$$\mu(x) = \sum_y p(y,x)\mu(y) \tag{38.172}$$

et si l'ensemble E est fini cette équation signifie

$$\mu = Q\mu \tag{38.173}$$

où Q est la matrice de transition de la chaine de Markov.

Théorème 38.55 (Théorème ergodique)**.**
Une chaine de Markov irréductible est positive récurrente si et seulement si elle accepte une mesure invariante. Cette mesure est invariante est alors unique et vérifie $\mu = Q\mu$ où Q est la matrice de transition.

Démonstration. Nous allons seulement prouver le théorème ergodique dans le cas où E est fini. Soit (X_n) une chaine de Markov récurrente positive ; nous avons $\pi(x) > 0$ pour tout $x \in E$. Nous allons montrer que π est une mesure invariante.

Nous commençons par montrer que

$$\sum_{x\in E} \pi(x) = 1. \tag{38.174}$$

Pour cela nous reprenons la propriété de chaine irréductible pour écrire

$$\pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{X_k=x} \tag{38.175}$$

Étant donné que E est fini nous pouvons sommer sur $x \in E$ et permuter la somme avec la limite :

$$\sum_{x\in E} \pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \underbrace{\sum_{x\in E} \mathbb{1}_{X_k=x}}_{=1}. \tag{38.176}$$

Nous nous retrouvons donc avec $\lim_{N\to\infty} \frac{1}{N} N = 1$. La fonction π définit donc bien une mesure de probabilité sur E.

Nous montrons à présent que cette mesure est invariante, c'est-à-dire que

$$\pi(x) = \sum_{y\in E} p(y,x)\pi(y). \tag{38.177}$$

Pour cela nous utilisons encore le théorème de la convergence dominée pour permuter la limite et l'intégrale dans

$$\pi(x) = E(\pi(x)) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \underbrace{E\big(\mathbb{1}_{X_k=x}\big)}_{P(X_k=x)} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P(X_{k+1} = x). \tag{38.178}$$

EqcKxNcL

La dernière égalité découle du fait qu'en divisant par N et en faisant tendre N vers l'infini, le fait d'enlever un terme à la somme ne change pas la valeur de la limite. Nous pouvons substituer dans (38.178) la valeur

$$P(X_{k+1} = x) = \sum_{y\in E} p(y,x)P(X_k = y). \tag{38.179}$$

Nous avons alors

$$\pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \sum_{y\in E} p(y,x)P(X_k = y) \tag{38.180a}$$

$$= \sum_{y\in E} p(y,x) \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P(X_k = y) \tag{38.180b}$$

$$= \sum_{y\in E} p(y,x)\pi(y), \tag{38.180c}$$

ce qui signifie que π est une mesure invariante. Notons que nous avons encore utilisé le fait que E soit fini pour permuter avec la limite.

Il nous reste à montrer l'unicité de la mesure invariante sur la chaine de Markov. Soit μ une mesure invariante pour la chaine de Markov (X_k). Comme indiqué dans la remarque 38.51 nous pouvons supposer que X_0 suit la loi μ. Par le lemme 38.52 nous avons $P(X_k = x) = \mu(x)$ pour tout k. Par conséquent

$$\pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P(X_k = x) = \mu(x). \tag{38.181}$$

□

Théorème 38.56 (loi des grands nombres pour les chaines de Markov)**.**
Soit (X_n) une chaine de Markov irréductible acceptant une mesure invariante. Soit $f : E \to \mathbb{R}$ une fonction dans $L^1(E,\mu)$. Alors nous avons

$$\frac{1}{N} \sum_{k=1}^{N} f(X_k) \xrightarrow{p.s.} \sum_{x\in E} f(x)\mu(x). \tag{38.182}$$

En ce qui concerne les notations, l'hypothèse $f \in L^1(E,\mu)$ signifie

$$\sum_{x\in E} |f(x)|\mu(x) = \int_E |f(x)|d\mu(x) < \infty. \tag{38.183}$$

Démonstration. Nous prouvons le théorème dans le cas où E est fini. Si nous écrivons

$$f(X_k) = \sum_{y\in E} f(y)\mathbb{1}_{X_k=y}, \tag{38.184}$$

alors

$$\frac{1}{N} \sum_{k=1}^{N} f(X_k) = \sum_{y\in E} f(y) \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{X_k=y}. \tag{38.185}$$

Étant donné que E est fini nous pouvons permuter les sommes et prendre la limite $N \to \infty$:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k} f(X_k) = \sum_{y\in E} f(y) \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{X_k=y} = \sum_{y\in E} f(y)\pi(y). \tag{38.186}$$

□

38.6 Convergence vers l'équilibre

Théorème 38.57 (Convergence en loi des chaine de Markov)**.**
Si (X_n) est

(1) irréductible,

(2) récurrente positive,

(3) apériodique,

alors X_n converge en loi vers l'unique probabilité invariante π vérifiant

$$\pi(x) = \sum_{u\in E} p(y,x)\pi(y) = \frac{1}{E\big(T(x)|X_0 = x\big)}. \tag{38.187}$$

Cette convergence est indépendante de la loi de X_0 et on a

$$P(X_n = x|X_0 = y) \to_{n\to\infty} \pi(x). \tag{38.188}$$

38.7 Processus de Galton-Watson

SecBPmrPdtGalton

Nous considérons une maladie et notons Z_n le nombre de malades à l'instant n. Nous posons $Z_0 = 1$ et

$$Z_{n+1} = \begin{cases} 0 & \text{si } Z_n = 0 \\ \sum_{i=1}^{Z_n} \xi_i^{(n)} & \text{sinon} \end{cases} \tag{38.189}$$

EqByILKj

où $\xi_i^{(n)}$ est le nombre de personnes contaminées par le malade i à l'instant n. Nous supposons que ces variables aléatoires sont indépendantes et identiquement distribuées et admettent un moment d'ordre 1.

L'équation de propagation 38.189 signifie que nous supposons qu'une personne malade à l'instant n n'est plus malade à l'instant $n+1$. Par ailleurs les hypothèses d'indépendance signifient qu'à chaque instant, le nombre de personnes contaminées par le malade i est indépendant du nombre de personnes contaminées par le malade j. De plus la façon dont la contamination se passe à l'instant n est indépendante de la façon dont la contamination se passe à l'instant m. Ces hypothèses sont raisonnables tant que le nombre de personnes non contaminées est grand. À partir du moment où presque tout le monde est malade, l'approximation de Galton-Watson ne fonctionne plus.

Nous notons ξ la loi parente des $\xi_i^{(n)}$. Ensuite nous considérons

$$G(s) = E(s^\xi) \tag{38.190a}$$
$$m = E(\xi) \tag{38.190b}$$
$$G_n(s) = E(s^{Z_n}). \tag{38.190c}$$

Par le théorème de transfert (proposition 36.96) avec $f(t) = s^t$, ce que nous avons est

$$G_n(s) = E\big(f(Z_n)\big) = \int_{\mathbb{R}} s^x dP_{Z_n}(x) = \sum_{k=0}^{\infty} s^k P(Z_n = k) \tag{38.191}$$

EqNRtXdC

où l'intégrale s'est transformée en somme parce que la loi de Z_n est discrète : dP_{Z_n} est une somme de masses de Dirac. En particulier nous avons

$$G_n(s) = \sum_{k=0}^{\infty} s^k P(Z_n = k) \tag{38.192a}$$
$$G(0) = P(Z_n = 0) \tag{38.192b}$$

et

$$\eta = \lim_{n\to\infty} P(Z_n = 0) = \lim_{n\to\infty} G_n(0). \tag{38.193}$$

D'où l'intérêt d'étudier G_n.

LemezrOiI

Lemme 38.58.
Pour tout $n \in \mathbb{N}^*$ et pour tout $s \in [0,1]$, nous avons

$$G_n(s) = \underbrace{G \circ G \circ \ldots \circ G(s)}_{n \text{ fois}}. \tag{38.194}$$

Démonstration. Pour $n = 1$, nous avons $Z_1 = \xi_1^{(1)}$ et donc

$$G_1(s) = E(s^{\xi}) = G(s), \tag{38.195}$$

comme il se doit.

Si $n \neq 1$ nous écrivons [subEqsxhILKg]

$$G_n(s) = E(s^{Z_n}) \tag{38.196a}$$
$$= E\left(s^{\sum_{i=1}^{Z_{n-1}} \xi_i^{(n-1)}}\right) \tag{38.196b}$$
$$= E\left(\sum_{k=0}^{\infty} \mathbb{1}_{\{Z_{n-1}=k\}} s^{\sum_{i=1}^{k} \xi_i^{(n-1)}}\right). \tag{38.196c}$$

À ce niveau, nous voulons permuter la somme et l'espérance. Étant donné que le lemme est facile à vérifier pour $s = 1$, nous supposons $s < 1$. Du coup

$$s^{\sum_{i=1}^{k} \xi_i^{(n-1)}} < 1 \tag{38.197}$$

et ce qui se trouve dans l'espérance est majoré par

$$\sum_{k=0}^{\infty} \mathbb{1}_{Z_{n-1}=k} = 1. \tag{38.198}$$

La fonction constante 1 est intégrable sur Ω (ici nous utilisons à fond le fait que l'espace Ω soit un espace de probabilité) et nous pouvons utiliser le théorème de convergence dominée de Lebesgue 15.2 pour permuter la somme et l'intégrale. Nous continuons donc le calcul (38.196) :

$$G_n(s) = \sum_{k=0}^{\infty} E\left(\mathbb{1}_{\{Z_{n-1}=k\}} s^{\sum_{i=1}^{k} \xi_i^{(n-1)}}\right). \tag{38.199}$$

La tribu engendrée par la variable aléatoire $\mathbb{1}_{\{Z_{n-1}=k\}}$ est une fonction des variables aléatoires $\xi_i^{(m)}$ avec $m \leqslant n-2$ tandis que la variable aléatoire $s^{\sum_{i=1}^{k} \xi_i^{(n-1)}}$ est une fonction des variables aléatoires $\xi_i^{(n-1)}$. Par conséquent le lemme de regroupement 36.19 nous dit que ces variables aléatoires sont indépendantes, donc

$$G_n(s) = \sum_{k=0}^{\infty} \underbrace{E\big(\mathbb{1}_{\{Z_{n-1}=k\}}\big)}_{=P(Z_{n-1}=k)} E\big(s^{\sum_{i=1}^{k} \xi_i^{(n-1)}}\big). \tag{38.200}$$

Nous avons utilisé le fait que l'espérance d'une fonction indicatrice est la probabilité de l'événement.

En ce qui concerne la puissance de s, les événements ξ_i^{n-1} sont indépendants et suivent tous la même loi ξ, donc

$$s^{\sum_{i=1}^{k} \xi_i^{(n-1)}} = \prod_{i=1}^{k} s^{\xi_i^{(n-1)}} \tag{38.201}$$

et

$$E\big(\prod_{i=1}^{k} s^{\xi}\big) = E(s^{\xi})^k = G(s)^k. \tag{38.202}$$

En mettant tout bout à bout,

$$G_n(s) = \sum_{k=1}^{\infty} P(Z_{n-1} = k) G(s)^j = G_{n-1}\big(G(s)\big). \tag{38.203}$$

□

ThoJZnAOA

Théorème 38.59.
La probabilité d'extinction η est donnée par

$$\eta = P\left(\bigcup_{n\geqslant 1}(Z_n = 0)\right) = \lim_{n\to\infty} P(Z_n = 0). \tag{38.204}$$

Ce nombre est la plus petite solution positive de l'équation $G(s) = s$.

De plus la classification des cas est comme suit.

(1) Si $P(\xi = 0) = 0$ alors $\eta = 0$.

(2) Si $P(\xi = 0) \neq 0$ alors

(2a) si $m \leqslant 1$ alors $\eta = 1$,

(2b) si $m > 1$ alors $\eta \in]0, 1[$.

Le cas $m < 1$ est dit **sous-critique**, le cas $m = 1$ est dit **critique**. Le cas $m > 1$ est dit **sur-critique**.

Démonstration. Commençons par prouver que G est une fonction continue. En utilisant le théorème de transfert comme pour l'équation (38.191) nous trouvons que

$$G(s) = E(s^\xi) = \sum_{k=0}^{\infty} p_k s^k \tag{38.205}$$

EqQWTBfn

où nous avons noté $p_k = P(\xi = k)$. Si $r < 1$, alors la suite $p_k r^k$ est bornée, donc le critère d'Abel (15.18) nous indique que la série (38.205) converge absolument et la théorie générale des séries entières conclut que la fonction G est en particulier dérivable terme à terme pour tout $s \in]-1, 1[$.

(1) **La probabilité d'extinction est un point fixe de G**

En utilisant la continuité de G en 0 nous passons à la limite dans $G_{n+1}(0) = G\big(G_n(0)\big)$ et nous obtenons

$$\eta = G(\eta), \tag{38.206}$$

ce qui signifie que la probabilité d'extinction est un point fixe de G.

(2) **η est le plus petit point fixe de G**

Nous démontrons maintenant que η est plus précisément le plus petit point fixe de G sur $[0, 1]$. Nous allons effectuer cette partie en décomposant selon les valeurs de p_0 et de p_1.

Au vu de l'écriture (38.205), si $p_1 = 1$ alors $G(s) = s$ pour tout $s \in [0, 1]$. Mais dans ce cas nous savons par ailleurs que l'extinction est impossible. Zéro est bien la plus petite solution de $G(s) = s$.

Supposons maintenant que $p_1 < 1$ et $p_0 + p_1 = 1$. Alors $G(s) = p_0 + p_1 s$ et $s = 1$ est l'unique solution. Mais vu que nous savons que η est solution, c'est que $\eta = 1$ et l'extinction est certaine.

Nous passons au cas général : $p_0 + p_1 < 1$. D'abord nous remarquons que $s = 1$ est solution parce que

$$G(1) = p_0 + p_1 + \cdots = 1. \tag{38.207}$$

Remarquons aussi que dans ce cas $s = 0$ n'est plus solution.

La fonction G est strictement convexe sur $[0, 1]$ (parce que $G'' > 0$). Cela se voit en effectuant deux dérivations terme à terme (le rayon de convergence de la dérivée est le même que celui de la fonction). Cette stricte convexité entraine que l'équation $G(s) = s$ a au maximum une autre solution que $s = 1$. Nous nommons s_0 la plus petite solution dans $[0, 1]$. Étant donné que G est croissante on a

$$G(0) \leqslant G(s_0) = s_0. \tag{38.208}$$

En appliquant G à cette équation nous obtenons $G\big(G(s_0)\big) \leqslant G(s_0) = s_0$ et en appliquant n fois,

$$G_n(0) \leqslant s_0. \tag{38.209}$$

En passant à la limite, $\eta \leqslant s_0$ mais η étant solution, nous avons $\eta = s_0$. Nous avons donc prouvé que la probabilité d'extinction η est la plus petite solution de $G(s) = s$.

(3) **Classification des cas**

Nous devons encore discuter les cas. Si $P(\xi = 0) = 0$, alors $p_0 = 0$ et $G(0) = 0$, ce qui signifie que $s_0 = \eta = 0$ et l'extinction est impossible.

Nous passons au cas $p_0 \neq 0$. Si $p_0 + p_1 = 1$, alors $m = p_1 < 1$ et nous avions déjà vu que dans le cas $p_0 + p_1 = 1$, la probabilité d'extinction est $\eta = 1$.

Il nous reste à traiter le cas $p_0 + p_1 < 1$. Encore une fois, la courbe G est strictement convexe sur $[0, 1]$ et elle est en particulier plus grande que sa tangente en $s = 1$, c'est-à-dire

$$G(s) > G'(1)(s - 1) + G(1). \tag{38.210}$$

Nous savons que $G(1) = 1$. En ce qui concerne $G'(1)$, nous dérivons encore terme à termes :

$$G'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}, \tag{38.211}$$

donc

$$G'(1) = \sum_{k=1}^{\infty} k p_k = E(\xi) = m. \tag{38.212}$$

Ce que nous avons donc est

$$G(s) > 1 + m(s - 1). \tag{38.213}$$

Nous nous particularisons au cas sous-critique ($m \leqslant 1$). En nous rappelant que $s - 1 < 0$,

$$G(s) > 1 + (s - 1) = s, \tag{38.214}$$

donc $s = 1$ est la plus petite solution et effectivement nous avons déjà vu que $\eta = 1$ dans ce cas.

Si $m > 1$, alors on a

$$G(s) > 1 + m(s - 1). \tag{38.215}$$

Mais dire $m > 1$ revient à dire $G'(1) > 1$ et donc dans un voisinage de $s = 1$ on a

$$\frac{G(s) - G(1)}{s - 1} > 1, \tag{38.216}$$

ce qui implique que

$$G(s) < s - 1 + G(1) = s. \tag{38.217}$$

Nous avons donc $G(s) < s$ dans un voisinage de 1. Mais $G(0) - 0 = p_0 > 0$, donc la fonction $f(s) = G(s) - s$ est positive en 0 et négative proche de $s = 1$. Le théorème de la valeur intermédiaire nous indique alors qu'il existe un $s \in]0, 1[$ tel que $f(s) = 0$, c'est-à-dire tel que $G(s) = s$.

□

Chapitre 39

Martingales

39.1 Convergence de martingales

Définition 39.1.
*Si $\mathcal{A}$ est une tribu, une **filtration** de $\mathcal{A}$ est une suite croissante de sous-tribus $\mathcal{B}_i \subseteq \mathcal{B}_{i+1} \subseteq \mathcal{A}$.*

*Nous disons qu'une suite de variables aléatoires (X_n) est **adaptée** à une filtration $(\mathcal{F}_n)$ si X_i est $\mathcal{F}_i$-mesurable pour tout i.*

DEFooUNRAooOEmCsh

Définition 39.2 ([795])**.**
*Si nous avons les variables aléatoires $X_k : \Omega \to \mathbb{R}$, la **filtration naturelle** est la suite de sous-tribus*

$$\mathcal{F}_n = \sigma(X_1, \dots, X_n), \tag{39.1}$$

c'est-à-dire la tribu engendrée par les n premières variables aléatoires.

Ces définitions impliquent immédiatement que si (X_n) est adapté à $(\mathcal{F}_n)$ alors X_n est $\mathcal{F}_k$-mesurable pour $k \geqslant n$.

DEFooESAGooUffeDa

Définition 39.3.
*Une **martingale** adaptée à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires M_n telle que*

(1) $M_n \in L^1(\Omega, \mathcal{A}, P)$,

(2) M_n est $\mathcal{B}_n$-mesurable,

(3) $E(M_{n+1}|\mathcal{B}_n) = M_n$.

*Le processus M_n est une **sur-martingale** si $E(M_{n+1}|\mathcal{B}_n) \leqslant M_n$ pour tout n, et c'est une **sous-martingale** si $E(M_{n+1}|\mathcal{B}_n) \geqslant M_n$.*

Exemple 39.4.
Si $M \in L^1(\Omega, \mathcal{A}, P)$ et si $(\mathcal{B}_n)_{n\in\mathbb{N}}$ est une filtration, nous pouvons considérer la martingale $M_n = E(M|\mathcal{B}_n)$. △

ExtFFKTr

Exemple 39.5.
Soit $(X_i)_{i\geqslant 1}$ une suite de variables aléatoires indépendantes et centrées. On pose

$$S_n = X_1 + \cdots + X_n \tag{39.2}$$

et la filtration $\mathcal{B}_n = \sigma(X_1, \dots, X_n)$. Pour montrer que cela est une martingale, nous commençons par remarquer que

$$E(X_{n+1}|\mathcal{B}_n) = E(X_{n+1}) = 0 \tag{39.3}$$

par indépendance des tribus $\mathcal{B}_n$ et $\sigma(X_{n+1})$. Ici c'est le lemme 36.70 qui joue.

Ensuite nous argumentons que $E(X_1 + \cdots + X_n|\mathcal{B}_n) = X_1 + \cdots + X_n$. En effet d'une part $X_1 + \cdots + X_n$ est $\mathcal{B}_n$-mesurable et évidemment la condition intégrale de l'espérance conditionnelle est satisfaite.

Plus généralement si X est une variable aléatoire et si $\sigma(X) \subset \mathcal{B}$ alors $E(X|\mathcal{B}) = X$. △

LemqanhgJ

Lemme 39.6.
Soit (M_n) une martingale adaptée à la filtration $(\mathcal{F}_n)$ et $n \geqslant k$. Alors

$$E(M_n|\mathcal{F}_k) = M_k \tag{39.4a}$$
$$E(M_k|\mathcal{F}_n) = M_k. \tag{39.4b}$$

Démonstration. La seconde relation revient seulement à dire que M_k est $\mathcal{F}_n$-mesurable, ce qui est évident parce que $\mathcal{F}_k \subset \mathcal{F}_n$.

Nous prouvons la première par récurrence (à l'envers) sur k. D'abord si $k = n$, l'égalité $E(M_n|\mathcal{F}_n) = M_n$. Nous supposons maintenant que $E(M_n|\mathcal{F}_k) = M_k$, et nous prouvons que $E(M_n|\mathcal{F}_{k-1}) = M_{k-1}$. Si $B_{k-1} \in \mathcal{F}_{k-1}$, nous avons

$$\int_{B_{k-1}} M_{k-1} = \int_{B_{k-1}} M_k = \int_{B_{k-1}} M_n. \tag{39.5}$$

La première égalité est la définition d'une martingale, et la seconde est l'hypothèse de récurrence. □

ThobysyWI

Théorème 39.7 ([796, 797]).
Soit $(M_n)_{n\geqslant 0}$ une martingale bornée dans $L^2(\Omega)$, c'est-à-dire telle que

$$\alpha = \sup_{n\geqslant 0} E(M_n^2) < \infty. \tag{39.6}$$

Alors la suite M_n converge dans $L^2(\Omega)$.

Démonstration. Nous écrivons M_n en somme télescopique

$$M_n = M_0 + \sum_{k=1}^{n} \Delta_k \tag{39.7}$$

où $\Delta_k = M_k - M_{k-1}$. Nous commençons par montrer que les incréments sont orthogonaux au sens où $E(\Delta_n\Delta_k) = 0$. Pour $n > k$, la variable aléatoire $E\big(\Delta_n\Delta_k|\mathcal{F}_{n-1}\big)$ est la variable aléatoire $\mathcal{F}_{n-1}$-mesurable telle que

$$\int_{B_{n-1}} E\big(\Delta_n\Delta_k|\mathcal{F}_{n-1}\big) = \int_{B_{n-1}} \Delta_n\Delta_k \tag{39.8}$$

pour tout $B_{n-1} \in \mathcal{F}_{n-1}$. En particulier avec $B_{n-1} = \Omega$ nous trouvons

$$E\Big(E\big(\Delta_n\Delta_k|\mathcal{F}_{n-1}\big)\Big) = E(\Delta_n\Delta_k) \tag{39.9}$$

par la définition de l'espérance (36.78). Par conséquent, en utilisant le lemme 39.6 nous avons [1]

$$E(\Delta_n\Delta_k) = E\Big(E(\Delta_n\Delta_k|\mathcal{F}_{n-1})\Big) = E\Big(\Delta_k E(\Delta_n|\mathcal{F}_{n-1})\Big) = 0 \tag{39.10}$$

parce que $E(\Delta_n|\mathcal{F}_{n-1}) = E(M_n|\mathcal{F}_{n-1}) - E(M_{n-1}|\mathcal{F}_{n-1}) = 0$.

Utilisant l'orthogonalité des incréments, nous avons

$$E(M_n^2) = E(M_0^2) + \sum_{k=1}^{n} E(\Delta_k^2). \tag{39.11}$$

En prenant le supremum (par rapport à n des deux côtés),

$$E(M_0^2) + \sum_{k=1}^{\infty} E(\Delta_k^2) = \alpha < \infty. \tag{39.12}$$

1. À ce niveau je crois qu'il y a une faute dans [797] qui conditionne par rapport à $\mathcal{F}_n$.

Cela prouve que la suite $\sum_{k=1}^{n}\Delta_k$ converge dans $L^2(\Omega)$. Nous en déduisons immédiatement que (M_n) est de Cauchy dans $L^2(\Omega)$ parce que si $k,l>n$, nous avons (en utilisant encore l'orthogonalité des incréments)

$$E\big(|M_k-M_l|^2\big)=\sum_{i=k+1}^{l}E(\Delta_i^2)\leqslant\sum_{i=k+1}^{\infty}E(\Delta_i^2),\tag{39.13}$$

qui tend vers zéro lorsque $n\to\infty$. □

Le théorème suivant complète la conclusion du théorème 39.7.

ThofcttYW

Théorème 39.8 ([797]).
Soit $(M_n)_{n\in\mathbb{N}}$ une martingale bornée dans L^2. Alors (M_n) converge dans $L^2(\Omega)$ et presque surement vers une même variable aléatoire M_∞ qui vérifie

$$M_n=E(M_\infty|\mathcal{F}_n).\tag{39.14}$$

EqmDMfZf

Notons en particulier que la variable aléatoire M_∞ est presque surement finie parce qu'en vertu de (39.14) nous avons

$$\int_\Omega M_\infty=\int_\Omega M_n<\infty.\tag{39.15}$$

Exemple 39.9.
Soient des variables aléatoires indépendantes $V_k\sim\mathscr{E}(2^k\lambda)$ et la variable aléatoire somme

$$S_n=\sum_{k=1}^{n}V_k.\tag{39.16}$$

Nous allons montrer que $S_n\xrightarrow{p.s.}X$ où X est une variable aléatoire presque surement finie. Nous posons

$$M_n=S_n-\sum_{k=1}^{n}\frac{1}{2^k\lambda}\tag{39.17}$$

Cela est une martingale adaptée à la filtration $\mathcal{F}_n=\sigma(V_1,\dots,V_n)$ en vertu de l'exemple 39.5. Nous montrons à présent qu'elle est bornée dans $L^2(\Omega)$ au sens où $\sum_{n\geqslant1}E(M_n^2)<\infty$. Nous avons

$$E(M_n^2)=E\left(\big[S_n-\sum_k\frac{1}{2^k\lambda}\big]^2\right)=E\left(\big[\sum_k(V_k-\frac{1}{2^k\lambda})\big]^2\right).\tag{39.18}$$

La variable aléatoire $V_k-1/2^k\lambda$ est une variable aléatoire centrée de variance $1/(2^k\lambda)^2$ (voir proposition 36.149). Étant donné que M_n est centrée, $\mathrm{Var}(M_n)=E(M_n^2)$ et nous avons

$$E(M_n^2)=\sum_{k=1}^{n}\mathrm{Var}\left(V_k-\frac{1}{2^k\lambda}\right)=\sum_{k=1}^{n}\frac{1}{(2^k\lambda)^2},\tag{39.19}$$

cette dernière somme étant bornée par $l=\sum_{k=1}^{\infty}\frac{1}{(2^k\lambda)^2}$, nous avons

$$E(M_n^2)\leqslant l\tag{39.20}$$

avec l indépendant de n. C'est pour cela que $(M_n)_{n\in\mathbb{N}}$ est une martingale bornée dans $L^2(\Omega)$. Par le théorème 39.8 nous avons $M_n\to M_\infty$ et en faisant $n\to\infty$ dans

$$S_n=M_n+\sum_{k=1}^{n}\frac{1}{2^k\lambda},\tag{39.21}$$

nous trouvons

$$S_n\to M_\infty+\sum_{k=1}^{\infty}\frac{1}{2^k\lambda}=M_\infty+\frac{1}{\lambda}\tag{39.22}$$

qui est presque surement finie. △

PROPooSWTPooOvzYPQ

Proposition 39.10 ([795]).
Si $(M_n)_{n\in\mathbb{N}}$ est une sous-martingale, alors

$$E(M_0) \leqslant E(M_n) \tag{39.23}$$

pour tout n.

39.2 Temps d'arrêt et martingale terminée

DEFooJBYJooKYVAJG

Définition 39.11 (Temps d'arrêt).
*Soit $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$ un espace de probabilité filtré. Une application $T\colon \Omega \to \bar{\mathbb{N}} \cup \{\infty\}$ est un **temps d'arrêt** adapté à la filtration $(\mathcal{F}_n)$ si pour tout $n \in \mathbb{N}$ nous avons $\{T \leqslant n\} \in \mathcal{F}_n$.*

*Le temps d'arrêt T est **borné** si il existe $k \in \mathbb{N}$ tel que $T(\omega) \leqslant k$ pour presque tout $\omega \in \Omega$.*

39.2.1 Nombre de montées

DEFooIUOXooPfASWz

Définition 39.12 ([795, 1]).
Soient $a, b \in \mathbb{R}$ avec $a < b$. Soit une suite de variables aléatoires $Z_k\colon \Omega \to \mathbb{R}$. Nous définissons les applications T_k et S_k par récurrence en posant :

EQooVQSIooHGAWau

$$T_0 = 0 \tag{39.24a}$$
$$S_{j+1} = \min\{k \geqslant T_j \text{ tel que } Z_k \leqslant a\} \tag{39.24b}$$
$$T_{j+1} = \min\{k > S_{j+1} \text{ tel que } Z_k \geqslant b\}. \tag{39.24c}$$

Dans tout ceci, nous convenons que $\min(\varnothing) = +\infty$ et $\max(\varnothing) = 0$.

*Le **nombre de montées** de a) b est alors la variable aléatoire*

$$U_n(a, b) = \max\{j \leqslant n \text{ tel que } T_j \leqslant n\}. \tag{39.25}$$

PROPooTYOAooEsoEPH

Proposition 39.13 ([1]).
Soit une suite de variable aléatoires $(Z_k)_{k\in\mathbb{N}}$. Les applications T_k et S_k définies par (39.24) sont des variables aléatoires. Elles sont également des temps d'arrêts[2] pour la filtration naturelle[3] des $(Z_k)_{k\in\mathbb{N}}$.

De plus, tant que $S_j, T_j < \infty$ nous avons

$$0 = T_0 \leqslant S_1 < T_1 \tag{39.26}$$

EQooTFZKooPeJtgg

et, pour tout $j \geqslant 1$,

$$T_j < S_{j+1} < T_{j+1}, \tag{39.27}$$

EQooCOUAooSgWEmf

et

$$Z_{S_j} \leqslant a < b \leqslant Z_{T_j}. \tag{39.28}$$

EQooUNTOooXbDmWr

Démonstration. Pour montrer que ce sont des variables aléatoires, tout est dans le théorème 14.96 et les suivants[4].

(1) T_0 **est un temps d'arrêt**
Nous avons $\{T_0 \leqslant n\} = \{0 \leqslant n\} = \Omega$. Or Ω est dans toute tribu.

2. Définition 39.11.
3. Définition 39.2.
4. Si vous avez un problème, écrivez-moi.

(2) **Si T_j est un temps d'arrêt, alors S_{j+1} en est un**

Nous supposons que T_j est un temps d'arrêt pour la filtration naturelle des $(Z_k)_{k\in\mathbb{N}}$, c'est-à-dire que pour tout n,

$$\{\omega\in\Omega \text{ tel que } T_k(\omega)\leqslant n\}\in\mathcal{F}_n=\sigma(Z_1,\dots,Z_n). \tag{39.29}$$

Nous devons montrer que $\{\omega\in\Omega \text{ tel que } S_j(\omega)\leqslant n\}\in\mathcal{F}_n$. D'abord nous écrivons

$$\{\omega\in\Omega \text{ tel que } S_j(\omega)\leqslant n\}=\bigcup_{k=1}^{n}\{\omega\in\Omega \text{ tel que } S_j(\omega)=k\}. \tag{39.30}$$

EQooHBWRooGhAIan

Ensuite nous comprenons que $S_j(\omega)=k$ signifie que $Z_k(\omega)\leqslant a$ et $k\leqslant T_j(\omega)$, et que k est minimum pour ces deux propriétés. Nous avons donc

$$\begin{aligned}\{\omega\in\Omega \text{ tel que } S_j(\omega)=k\}=\{\omega\in\Omega \text{ tel que } k\leqslant T_j(\omega)\}\cap\{\omega\in\Omega \text{ tel que } Z_k(\omega)\leqslant a\}\\ \setminus\bigcup_{l=1}^{k}\{\omega\in\Omega \text{ tel que } T_j(\omega)=l, Z_l(\omega)>a\}.\end{aligned} \tag{39.31}$$

Nous rappelons que là dedans, $k\leqslant n$ et que nous voulons que le tout soit dans $\mathcal{F}_n=\sigma(Z_1,\dots,Z_n)$. Comme les tribus sont stables par complémentaires, unions dénombrables et intersections dénombrables [5], nous avons le droit de travailler chaque partie séparément. Chaque élément de l'intersection sur l est de la forme $\{T_j=l\}\cap\{Z_l>a\}$. Et nous avons

$$\{T_k=l\}=\{T_l\leqslant l\}\setminus\{T_k\leqslant l-1\}. \tag{39.32}$$

Nous prouvons bout par bout que pour tout $m\leqslant k\leqslant n$, les parties suivantes sont dans $\mathcal{F}_n$:

— $\{T_j\leqslant m\}\in\mathcal{F}_n$ parce que T_j est un temps d'arrêt.

— $\{T_j\geqslant k\}\in\mathcal{F}_n$ parce que c'est le complémentaire de $\{T_j<k\}=\{T_k\leqslant k-1\}$ et que $k-1\leqslant n$.

— $\{Z_k\leqslant a\}\in\mathcal{F}_n$ parce que $\mathcal{F}_n$ est engendrée, entre autres par Z_k. Ici nous utilisons le fait que $k\leqslant n$.

— $\{Z_l>a\}\in\mathcal{F}_n$ parce que $\mathcal{F}_n$ contient entre autres $\sigma(Z_l)$.

(3) **Si S_j est un temps d'arrêt alors T_{j+1} en est un**

Même raisonnement que l'autre [6].

Et maintenant, les inégalités (39.26), (39.27) et (39.28). Notons que dès que $T_j=\infty$, tous les autres sont ∞ et les inégalités n'ont plus vraiment de sens. Nous supposons donc toujours que $T_j,S_j<\infty$.

(1) $T_0\leqslant S_1$ Nous avons $S_1=\min\{k\geqslant T_0 \text{ tel que } Z_k\leqslant a\}$. En particulier $S_1\geqslant T_0$.

(2) $S_1<T_1$ Par définition.

(3) $T_j<S_{j+1}$ **si** $j\geqslant 1$ Par définition $S_{j+1}=\min\{k\geqslant T_j \text{ tel que } Z_k\leqslant a\}$. En particulier $S_{j+1}\geqslant T_j$. Voyons pourquoi pas $S_{j+1}=T_j$. Vue la définition de S_{j+1}, avoir $S_{j+1}=T_j$ demanderait $Z_{T_j}\leqslant a<b$. Mais $T_j=\min\{k>S_j \text{ tel que } Z_k\leqslant b\}$. Donc $Z_{T_j}\geqslant b$. Contradiction.

(4) $S_{j+1}<T_{j+1}$ Par définition.

(5) $Z_{S_j}\leqslant a$ Si $l=S_j(\omega)$, alors en particulier $l\in\{k\geqslant T_j(\omega) \text{ tel que } Z_k(\omega)\leqslant a\}$. En particulier $Z_l(\omega)\leqslant a$.

(6) $b\leqslant Z_{T_j}$ Si $l=T_j(\omega)$, alors $l\in\{k>S_j(\omega) \text{ tel que } Z_k(\omega)\geqslant b\}$. En particulier $Z_l(\omega)\geqslant b$.

□

5. Définition 14.1 et lemme 14.3.
6. Attention. J'ai pas vérifié. Je te conseille de vérifier, et de m'écrire si il y a une difficulté.

PROPooSGXJooQkbyMH

Proposition 39.14 ([795]).
Nous fixons $a < b$ dans $\mathbb{R}$. Le nombre de montées est croissant et en posant

$$U_n(\omega) = \lim_{n\to\infty} U_n(\omega), \tag{39.33}$$

l'application U_n est une variable aléatoire.

Démonstration. La définition est que $U_n(\omega) = \max\{j \leqslant n \text{ tel que } T_j(\omega) \leqslant n\}$. Nous avons

$$\{j \leqslant n \text{ tel que } T_j(\omega) \leqslant n\} \subset \{j \leqslant n+1 \text{ tel que } T_j(\omega) \leqslant n+1\}. \tag{39.34}$$

Donc le maximum de l'ensemble de droite est plus grand.

Le fait que la limite soit une variable aléatoire (c'est-à-dire mesurable) est la proposition 14.111. □

THOooZKNNooVdOxKu

Théorème 39.15 (Inégalité des montées de Doobs[795]).
Soit une sous-martingale $(Z_n)_{n\in\mathbb{N}}$. Soient $a < b$ dans $\mathbb{R}$. Nous avons

$$E\big(U_n(a,b)\big) \leqslant \frac{1}{b-a} E\Big((Z_n - a)^+\Big) \leqslant \frac{1}{b-a}\Big(E\big(|Z_n|\big) + |a|\Big). \tag{39.35}$$

PROPooMFRFooHMIjsy

Proposition 39.16 ([795]).
Soit une sous-martingale[7] *$(Z_n)_{n\in\mathbb{N}}$ telle que $\sup_n E(Z_n^+) < \infty$. Alors il existe $\Omega' \subset \Omega$ tel que $P(\Omega') = 1$ et une variable aléatoire $Z \in L^1(\Omega')$ telle que*

$$Z_n(\omega) \to Z \tag{39.36}$$

pour tout $\omega \in \Omega'$.

Démonstration. En plusieurs parties.

(1) $U(a,b) \in L^1(\Omega)$

Soient $a < b \in \mathbb{R}$. Nous considérons également le nombre de montées[8] entre a et b. C'est une suite croissante et la limite est une variable aléatoire (proposition 39.14). Le théorème de Beppo-Levi 14.178 nous dit alors que

$$\lim_{n\to\infty} E\big(U_n(a,b)\big) = E\big(U(a,b)\big) \in [0,\infty]. \tag{39.37}$$

Nous passons à la limite dans l'inégalité montées 39.15 :

$$E\Big(U_n(a,b)\Big) \leqslant \lim_{n\to\infty} \frac{1}{b-a}\Big(E\big(|Z_n|\big) + |a|\Big). \tag{39.38}$$

Dans le membre de droite nous majorons

$$E\big(|Z_n|\big) \leqslant E(Z_n^+) \leqslant \sup_n E(Z_n^+). \tag{39.39}$$

Nous avons donc

$$E\Big(U_n(a,b)\Big) \leqslant \frac{1}{b-a} \sup_{n\in\mathbb{N}} \Big(E(Z_n^+) + |a|\Big) < \infty. \tag{39.40}$$

La dernière inégalité étant l'hypothèse. Nous avons donc démontré que

$$E\big(U(a,b)\big) < \infty, \tag{39.41}$$

c'est-à-dire que $U(a,b) \in L^1(\Omega)$, et donc $U(a,b) < \infty$ presque surement. Donc

$$P\Big(\{\omega \in \Omega \text{ tel que } U(a,b)(\omega) = \infty\}\Big) = 0. \tag{39.42}$$

EQooJDDJooNYpzru

7. Définition 39.3.
8. Définition 39.12.

(2) **La partie $A_{a,b}$**
Pour chaque $a < b$ dans $\mathbb{Q}$, nous introduisons cette partie de Ω :

$$A_{a,b} = \{\omega \in \Omega \text{ tel que } \begin{cases} \liminf_n Z_n(\omega) \leqslant a \\ \limsup_n Z_n(\omega) \geqslant b \end{cases}\}. \tag{39.43}$$

(3) $T_j, S_j < \infty$ **sur** $A_{a,b}$
Soit $\omega \in A_{a,b}$. Nous prouvons que $T_j(\omega) < \infty$ et $S_j(\omega) < \infty$ par récurrence. Pour tout $N > 0$, il existe $k, > N$ tel que $Z_k(\omega) \leqslant a$ et $Z_l(\omega) \geqslant b$.
Si $T_j(\omega) < \infty$, alors $S_{j+1} < \infty$. En effet il existe $k > T_j(\omega)$ tel que $Z_k(\omega) \leqslant a$. Et donc $S_{j+1}(\omega) < \infty$.
De même si $S_j(\omega) < \infty$, alors il existe $k > S_j(\omega)$ tel que $Z_k(\omega) \geqslant b$. Donc $T_{j+1}(\omega) < \infty$.

(4) $U(a,b) = \infty$ **sur** $A_{a,b}$
Nous montrons à présent que $U(\omega) = \infty$ pour tout $\omega \in A_{a,b}$. Soient $\omega \in A_{a,b}$, et $m > 0$. Nous posons

$$N = \max\{T_m(\omega), m\}, \tag{39.44}$$

et nous considérons $n > N$. Nous allons prouver qu'alors $U_n(a,b) \geqslant m$. Pour rappel $U_n(\omega) = \max\{j \leqslant n \text{ tel que } T_j(\omega) \leqslant n\}$.. Vu que $T_m(\omega) < N < n$ et $m \leqslant N < n$, nous avons $m \in \{j \leqslant n \text{ tel que } T_j(\omega) \leqslant n\}$, et donc $U_n(\omega) \geqslant m$. Nous avons donc prouvé que

$$\lim_{n\to\infty} U_n(a,b)(\omega) = \infty \tag{39.45}$$

dès que $\omega \in A_{a,b}$. En vertu de (39.42), nous avons $P(A_{a,b}) = 0$ pour tout $a < b \in \mathbb{R}$.

(5) **L'union des** $A_{a,b}$
Nous posons

$$A = \{\omega \in \Omega \text{ tel que } \liminf_n Z_n(\omega) < \limsup_n Z_n(\omega)\}, \tag{39.46}$$

et nous prouvons à présent que $A = \bigcup_{\substack{a,b\in\mathbb{Q}\\ a<b}} A_{a,b}$.
Soit $\omega \in A$. Il existe $a, b \in \mathbb{Q}$ tels que $\omega \in A_{a,b}$. Nous avons donc

$$\liminf_n Z_n(\omega) \leqslant a < b \leqslant \limsup_n Z_n(\omega), \tag{39.47}$$

et donc $\omega \in A$. Dans l'autre sens, nous supposons que $\omega \in \Omega$ satisfasse $\liminf_n Z_n(\omega) < \limsup_n Z_n(\omega)$. En utilisant le lemme 1.441(1) (deux fois), nous considérons $a < b$ dans $\mathbb{Q}$ tels que

$$\liminf_n Z_n(\omega) < a < b < \limsup_n Z_n(\omega). \tag{39.48}$$

Cela prouve que $\omega \in A_{a,b} \subset A$.

(6) **Convergence presque sure**
En tant qu'union dénombrable de parties de mesures nulles, nous avons $P(A) = 0$. Le Ω' dont on parle dans l'énoncé sera A^c. Vu que $P(\Omega\backslash\Omega') = 0$, le mesure P restreinte aux parties de Ω' est une mesure de probabilité sur Ω'. Pour $\omega \in \Omega'$, nous avons

$$\liminf_n Z_n(\omega) = \limsup_n Z_n(\omega), \tag{39.49}$$

et donc la suite $(Z_n(\omega))$ converge dans $\bar{\mathbb{R}}$. Nous considérons la variable aléatoire

$$\begin{aligned} Z\colon \Omega' &\to \bar{\mathbb{R}} \\ \omega &\mapsto \lim_{n\to\infty} Z_n(\omega). \end{aligned} \tag{39.50}$$

Vu que $Z_n(\omega) \to Z(\omega)$, la proposition 9.48 dit que $\liminf_n |Z_n(\omega)| = |Z(\omega)|$. Nous pouvons calculer un peu d'espérance : SUBEQSooPEZEooTDliNe

$$\begin{aligned}
E\big(|Z|\big) &= \int_{\Omega'} |Z(\omega)| dP(\Omega) && (39.51a)\\
&= \int_{\Omega'} \liminf_n |Z_n(\omega)| dP(\omega) && (39.51b)\\
&\leqslant \int_{\Omega'} |Z_n(\omega)| dP(\omega) && \text{lem. Fatou 14.182} \quad (39.51c)\\
&= \liminf_n E\big(|Z_n|\big). && (39.51d)
\end{aligned}$$

(7) **Un peu de majoration**
Nous avons $|Z_n| = Z_n^+ + Z_n^- = 2Z_n^+ - Z_n$, et nous continuons les inégalités (39.51) :

$$\begin{aligned}
E(|Z|) &\leqslant \liminf_n \big(E(|Z_n|)\big) && (39.52a)\\
&= 2E(Z_n^+) - E(Z_n) && (39.52b)\\
&\leqslant 2E(Z_n^+) - E(Z_0) && \text{prop. 39.10} \quad (39.52c)\\
&\leqslant 2\sup_n E(Z_n^+) - E(Z_0) && (39.52d)\\
&< \infty. && (39.52e)
\end{aligned}$$

Cela termine de prouver que $Z \in L^1(\Omega', P)$.

□

39.2.2 Théorèmes d'arrêt

LemXYeCLXW

Lemme 39.17.
Si T est un temps d'arrêt presque surement fini, alors[9]

ItemIPPkxmAii

(1) $T \wedge n \xrightarrow{p.s.} T$,

(2) $\lim_{n\to\infty} E(T \wedge n) = E(T)$.

Note : il est d'usage assez classique de noter $a \wedge b$ le minimum de a et b.

Démonstration. Vu que T est presque surement finie, il suffit de prouver que

EqRVoKxsN
$$(T \wedge n)(\omega) \to T(\omega) \quad (39.53)$$

pour tout ω tel que $T(\omega) = k$ pour tout $k \in \mathbb{N}$. Soient donc $\omega \in \Omega$ tel que $T(\omega) = k$ et $n > k$. Nous avons

$$(T \wedge n)(\omega) = T(\omega) \wedge n = k = T(\omega). \quad (39.54)$$

En ce qui concerne la seconde assertion, la suite de variables aléatoires $X_n = T \wedge n$ est croissante et positive, donc le théorème de la convergence monotone 14.178 montre que

$$\lim_{n\to\infty} E(T \wedge n) = E(T). \quad (39.55)$$

□

Remarque 39.18.
Notons la différence subtile entre $S_T(\omega)$ et $(S_T)(\omega)$. La première est la variable aléatoire

$$\omega' \mapsto S_{T(\omega')}(\omega) \quad (39.56)$$

et la seconde est le nombre $S_{T(\omega)}(\omega)$.

9. Dans [104], dans le problème de la ruine du joueur, la seconde assertion est avec une limite sup et non avec une limite normale.

ThoQMsRbkp

Théorème 39.19 (Théorème d'arrêt borné[797])**.**
Soit (X_n) une sur-martingale et $S \leqslant T$, deux temps d'arrêts bornés. Alors

(1) les variables aléatoires X_S et X_T sont intégrables,

(2) $E(X_T|\sigma(S)) \leqslant X_S$ presque surement.

Si par contre (X_n) est une martingale alors X_S et X_T sont bornées, et

$$E(X_T|\sigma(S)) = X_S. \tag{39.57}$$

RemKCdpnid

Remarque 39.20.
Un cas particulier intéressant de ce théorème 39.19 est le cas $S = 0$ qui est un temps d'arrêt vérifiant $\mathcal{F}_0 = \{\Omega, \varnothing\}$. Si X est n'importe quelle variable aléatoire, la tribu engendrée $\sigma(X)$ est toujours indépendante de la tribu $\{\Omega, \varnothing\}$, donc le résultat $E(X_T|\mathcal{F}_S) = X_S$ donne

$$E(X_T) = X_0. \tag{39.58}$$

ThoZTrdjtZ

Théorème 39.21 (Premier théorème d'arrêt de Doob[798, 799, 795])**.**
Soient (X_n) une martingale et T un temps d'arrêt[10] *; tous deux pour la filtration $(\mathcal{F}_n)$. Nous supposons qu'une des trois propriétés suivantes soit vérifiée :*

(1) T est presque surement bornée : il existe $c \in \mathbb{R}$ tel que $T(\omega) < c$ pour presque tout ω. ItemQVWZuBkiii

(2) Il existe une constante c telle que $|X_{T\wedge n}| \leqslant c$ presque surement.

(3) Il existe $k > 0$ tel que $|X_n(\omega)| < k$ pour tout $n \in \mathbb{N}$ et $\omega \in \Omega$. Et $T < \infty$ presque surement.

(4) $E(T) < \infty$ et il existe une constante c telle que

$$E\big(|X_{n+1} - X_n|\,|\mathcal{F}_n\big) \leqslant c \tag{39.59}$$

sur l'événement $\{T \geqslant n\}$.

Alors X_T est une variable aléatoire presque surement bien définie nous avons

$$E(X_T) = E(X_0). \tag{39.60}$$

Si (X_n) est une sur-martingale, alors la conclusion est $E(X_T) \leqslant E(X_0)$ et si (X_n) est une sous-martingale, la conclusion est $E(X_T) \geqslant E(X_0)$.

¡¡ Avertissement/question au lecteur !! 39.22
D'après la page de discussion de l'article sur Wikipédia, il semblerait que la seconde condition soit mal énoncée. Je n'ai pas vérifié.

Remarque 39.23.
Sous l'hypothèse (2), il est possible d'avoir $T = \infty$ sur un ensemble de mesure non nulle. Sur cet ensemble, la variable aléatoire X_T doit être définie de façon plus fine.

Définition 39.24.
*Nous disons que la martingale $(M_n)_{n\geqslant 1}$ est **terminée** si il existe $M \in L^1(\Omega, \mathcal{A}, P)$ telle que $E(M|\mathcal{A}_n) = M$ pour tout $n > 1$.*

ThoEFbpVXb

Théorème 39.25.
Si (M_n) est une martingale, nous avons équivalence entre

(1) (M_n) converge dans L^1 ;

(2) (M_n) est terminée ;

(3) l'ensemble $\{M_n\}_{n\geqslant 1}$ est équi-intégrable[11]*.*

10. Définition 39.11.
11. Définition 27.56 et lemme 27.57.

La démonstration est sans doute par petits bouts dans [795].

Attention : en vertu de la proposition 27.18 et surtout de l'exemple 27.19, la convergence L^1 n'implique pas la convergence presque partout.

ThoHBvnTRk

Théorème 39.26 (Théorème de Doob[432])**.**
À propos de convergence de martingales.

(1) *Toute martingale terminée converge presque surement et pour la norme L^1.*
(2) *Toute martingale bornée dans L^2 converge presque surement et pour la norme L^2.*

PropAYJpGsc

Proposition 39.27 ([800])**.**
Soient (M_n) une martingale et T un temps d'arrêt (pour la même filtration $(\mathcal{B}_n)$). Alors le processus $V_n = M_{n\wedge T}$ est une martingale.

Démonstration. Nous décomposons V_n de la façon suivante :

$$V_n = M_{n\wedge T} = M_n \mathbb{1}_{T\geqslant n} + M_T \mathbb{1}_{T<n} = M_n \mathbb{1}_{T\geqslant n} + \sum_{k<n} M_k \mathbb{1}_{T=k}. \tag{39.61}$$

EqYJiUZrv

Nous avons, grâce au lemme 14.3,

$$\{T \geqslant n\} = \complement\{T < n\} = \complement\{T \leqslant n-1\} \in \mathcal{B}_{n-1} \tag{39.62}$$

et, si $k \leqslant n$,

$$\{T = k\} = \underbrace{\{T \leqslant k\}}_{\in\mathcal{B}_k} \setminus \underbrace{\{T \leqslant k-1\}}_{\in\mathcal{B}_{k-1}} \in \mathcal{B}_k \subset \mathcal{B}_n. \tag{39.63}$$

La forme (39.61) donne donc manifestement la $\mathcal{B}_n$-mesurabilité de V_n.

En ce qui concerne l'espérance nous devons calculer

$$E(V_{n+1}|\mathcal{B}_n) = E(M_{n+1}\mathbb{1}_{T\geqslant n+1}|\mathcal{B}_n) + \sum_{k<n+1} E(M_k \mathbb{1}_{T=k}|\mathcal{B}_n) \tag{39.64}$$

où nous avons utilisé la proposition 36.47. Étant donné que $\mathbb{1}_{T\geqslant n+1}$ et $\mathbb{1}_{T=k}$ sont des variables aléatoires $\mathcal{B}_n$-mesurables nous pouvons utiliser la proposition 36.73 pour les sortir :

$$E(V_{n+1}|\mathcal{B}_n) = \mathbb{1}_{T\geqslant n+1} M_n + \sum_{k\leqslant n} \mathbb{1}_{T=k} M_k = M_{T\wedge n} = V_n. \tag{39.65}$$

Pour cette ligne, nous avons aussi utilisé les égalités suivantes :

— $E(M_{n+1}|\mathcal{B}_n) = M_n$ parce que (M_n) est une martingale
— $E(M_k|\mathcal{B}_n) = M_k$ parce que M_k est $\mathcal{B}_n$-mesurable.

□

Définition 39.28 (Processus arrêté[801])**.**
Si (X_n) est un processus adapté à la filtration $(\mathcal{F}_n)$ et si T est un temps d'arrêt $\mathcal{F}_n$-mesurable alors le ***processus arrêté*** *à l'instant T est le processus*

$$X_n^T(\omega) = X_{\min\{n,T(\omega)\}}(\omega). \tag{39.66}$$

Il peut également être écrit $X_n^T = X_{n\wedge T}$.

Dans le cas continu, nous avons cette définition. Soient :

(1) *Un espace de probabilité $(\Omega, \mathcal{F}, P)$,*
(2) *Un espace mesurable $(X, \mathcal{A})$,*
(3) *Un processus stochastique $X\colon [0,\infty[\times \Omega \to X$,*
(4) *Un temps d'arrêt τ de X pour la filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ de $\mathcal{F}$.*

*Le **processus arrêté** X^τ est le processus*

$$X_t^\tau(\omega) = X_{\min\{t,\tau(\omega)\}}(\omega). \tag{39.67}$$

LEMooWTLJooUyMYoV

Lemme 39.29.
Si (X_n) est une martingale alors son processus arrêté est encore une martingale.

Démonstration. C'est la proposition 39.27. □

39.3 Décomposition de martingales

Définition 39.30 (Processus croissant prévisible[797]).
*Un processus X_n adapté à la filtration $\mathcal{F}_n$ est un processus **croissant prévisible** si*

(1) $A_0 = 0$
(2) $A_n \leqslant A_{n+1}$ *; c'est cette condition qui correspond à « croissant »,*
(3) A_{n+1} *est $\mathcal{F}_n$-mesurable ; c'est cette condition qui correspond à « prévisible ».*

Proposition 39.31 (Décomposition de Doob pour une sous-martingale[797]).
Toute sous-martingale (X_n) s'écrit de façon unique sous la forme

$$X_n = M_n + A_n \tag{39.68}$$

EqCCsAwbZ

où (M_n) est une martingale et (A_n) est un processus croissant prévisible.

Démonstration. Nous considérons le processus

EqQFlGRzo

$$\begin{cases} A_0 = 0 & \text{(39.69a)} \\ A_{n+1} = A_n + E(X_{n+1} - X_n|\mathcal{F}_n). & \text{(39.69b)} \end{cases}$$

Nous vérifions que cela est un processus croissant prévisible. D'abord

$$E(X_{n+1} - X_n|\mathcal{F}_n) = E(X_{n+1}|\mathcal{F}_n) - E(X_n|\mathcal{F}_n). \tag{39.70}$$

Le second terme est égal à X_n parce que cette variable aléatoire est $\mathcal{F}_n$-mesurable tandis que (X_n) étant une sous-martingale nous avons $E(X_{n+1}|\mathcal{F}_n) \geqslant X_n$. Nous avons donc bien $A_{n+1} \geqslant A_n$ et le processus (A_n) est croissant.

En ce qui concerne la prévisibilité nous devons prouver que A_{n+1} est $\mathcal{F}_n$-mesurable. D'une part A_n est $\mathcal{F}_n$-mesurable et d'autre part par définition de l'espérance conditionnelle, la variable aléatoire $E(X_{n+1} - X_n|\mathcal{F}_n)$ est également $\mathcal{F}_n$-mesurable.

Nous posons alors $M_n = X_n - A_n$ et nous devons prouver que cela est une martingale. Nous avons

$$E(M_{n+1} - M_n|\mathcal{F}_n) = E(X_{n+1} - X_n|\mathcal{F}_n) - E(A_{n+1} - A_n|\mathcal{F}_n). \tag{39.71}$$

Le second terme vaut

$$E(A_{n+1} - A_n|\mathcal{F}_n) = E\Big(E(X_{n+1} - X_n|\mathcal{F}_n)|\mathcal{F}_n\Big) = E(X_{n+1} - X_n|\mathcal{F}_n) \tag{39.72}$$

par la proposition 36.68. Le processus (M_n) est donc une martingale. La preuve de l'existence d'une décomposition (39.68) est achevée.

Nous passons maintenant à l'unicité en posant $X_n = M_n + A_n = M'_n + A'_n$. Nous avons $A_0 = A'_0 = 0$ et $A'_n = X_n - M'_n$, donc

$$A'_{n+1} - A'_n = X_{n+1} - X_n + M'_{n+1} - M'_n = X_{n+1} - X_n - (M'_{n+1} - M'_n). \tag{39.73}$$

Nous appliquons $E(.|\mathcal{F}_n)$ des deux côtés de cette égalité :

$$\underbrace{E(A'_{n+1} - A'_n|\mathcal{F}_n)}_{=A'_{n+1}-A'_n} = E(X_{n+1} - X_n|\mathcal{F}_n) - \underbrace{E(M'_{n+1} - M'_n|\mathcal{F}_n)}_{=0}. \tag{39.74}$$

Nous avons utilisé le que que (M_n) étant une martingale, $E(M_{n+1}-M_n|\mathcal{F}_n)=0$, et idem avec (M'_n). Donc

$$A'_{n+1}-A'_n=E(X_{n+1}-X_n|\mathcal{F}_n)=E(M_{n+1}-M_n|\mathcal{F}_n)+E(A_{n+1}-A_n|\mathcal{F}_n)=A_{n+1}-A_n. \quad (39.75)$$

Nous avons donc montré que $A_{n+1}-A_n=A'_{n+1}-A'_n$ et donc que $A_n=A'_n$ pour tout n. Nous en déduisons immédiatement que $M_n=M'_n$ pour tout n et l'unicité de la décomposition. □

LemPVgeKfc

Lemme 39.32.
Si (X_n) est une martingale de carré intégrable adaptée à la filtration $(\mathcal{F}_n)$ alors

(1) Le processus (X_n^2) est une sous-martingale.

(2) Si $X_n^2=M_n+A_n$ est la décomposition de Doob, alors

$$A_n=\sum_{i=1}^{n}\Big(E(X_i^2|\mathcal{A}_{i-1})-X_{i-1}^2\Big)=\sum_{i=1}^{n}E\Big((X_i-X_{i-1})^2|\mathcal{A}_{i-1}\Big). \quad (39.76)$$

EqSTGxVWF

Démonstration. Pour la première assertion, nous utilisons l'inégalité de Jensen 36.86 :

$$E(X_n^2|\mathcal{F}_{n-1})\geqslant\big(E(X_n|\mathcal{F}_{n-1})\big)^2=X_n^2 \quad (39.77)$$

parce que $E(X_n|\mathcal{F}_{n-1})=X_n$ du fait que (X_n) soit une martingale.

En ce qui concerne la seconde assertion, nous nous souvenons que le processus prévisible de la décomposition de Doob d'une sous-martingale est donné par la récurrence (39.69) que nous recopions ici :

$$\begin{cases} A_0=0 & (39.78a)\\ A_{n+1}=A_n+E(X_{n+1}^2-X_n^2|\mathcal{F}_n) & (39.78b)\end{cases}$$

Vu que X_n^2 est $\mathcal{F}_n$-mesurable, il peut sortir de l'espérance :

$$A_{n+1}=A_n+E(X_{n+1}^2|\mathcal{F}_n)-X_n^2 \quad (39.79)$$

et donc

$$A_n=\sum_{i=1}^{n}\Big(E(X_i^2|\mathcal{F}_{i-1})-X_{i-1}^2\Big). \quad (39.80)$$

Pour obtenir la dernière partie de (39.76) nous travaillons un peu :

$$\begin{aligned}E\big((X_i-X_{i-1})^2|\mathcal{F}_{i-1}\big)&=E\big(X_i^2+X_{i-1}^2-2X_iX_{i-1}|\mathcal{F}_{i-1}\big) && (39.81a)\\ &=E(X_i^2|\mathcal{F}_{i-1})+X_{i-1}^2-2E(X_iX_{i-1}|\mathcal{F}_{i-1}) && (39.81b)\\ &=E(X_i^2|\mathcal{F}_{i-1})+X_{i-1}^2-2X_{i-1}E(X_i|\mathcal{F}_{i-1}) && (39.81c)\\ &=E(X_i^2|\mathcal{F}_{i-1})+X_{i-1}^2-2X_{i-1}X_i && (39.81d)\end{aligned}$$

EqFBkXiJH

où nous avons utilisé la proposition 36.73 pour obtenir (39.81c). □

39.4 Problème de la ruine du joueur

SecMSOjfgM

Nous considérons un joueur compulsif qui joue à un jeu très simple [12] : il joue à pile ou face contre la banque avec une pièce truquée. Si pile sort, la banque donne 1 au joueur et si c'est face, c'est le joueur qui donne 1 à la banque. Nous nommons a la fortune initiale du joueur, b celle de la banque et p la probabilité d'obtenir pile.

Nous supposons que le jeu se poursuit jusqu'à la ruine du joueur ou de la banque. La modélisation est comme suit : nous considérons (Y_n) une suite de variables aléatoires indépendantes et identiquement distribuées de loi

$$Y_n\sim p\delta_1+(1-p)\delta_{-1}. \quad (39.82)$$

12. Le gros des choses dites à propos de la ruine du joueur provient de [104].

C'est le résultat financier pour le joueur du n^{e} jet. La fortune du joueur au bout de n jets est la variable aléatoire

$$S_n = a + \sum_{j=1}^{n} Y_j. \tag{39.83}$$

Nous notons $Y_0 = a$.

Nous considérons la filtration

$$\mathcal{A}_n = \sigma\big(S_i \text{ tel que } 0 \leqslant i \leqslant n\big) = \sigma\big(Y_i \text{ tel que } 0 \leqslant i \leqslant n\big), \tag{39.84}$$

et le temps d'arrêt du jeu :

$$T = \inf\{n \geqslant 1 \text{ tel que } S_n \in \{0, a+b\}\}; \tag{39.85}$$

c'est le temps qu'il faut pour que tout l'argent appartienne soit au joueur soit à la banque.

Nous voulons étudier les paramètres suivants :

(1) $\rho = P(S_T = a+b)$, c'est-à-dire la probabilité que ce soit le joueur qui gagne contre la banque.
(2) $P(T < \infty)$, c'est-à-dire la probabilité que le jeu se finisse.
(3) $E(T)$, la durée moyenne du jeu.

LemEOAmVyZ

Lemme 39.33.
Le processus S_n du problème de la ruine du joueur, en notant $q = 1 - p$, vérifie

$$E(S_n|\mathcal{A}_{n-1}) = S_{n-1} + p - q. \tag{39.86}$$

EqZQtwVMXyiB

De plus le processus S_n est

(1) une martingale si $p = q = \frac{1}{2}$,
(2) une sous-martingale si $p > q$.

Démonstration. Pour $n \geqslant 1$ nous avons

$$E(S_n|\mathcal{A}_{n-1}) = a + \sum_{j=1}^{n} E(Y_j|\mathcal{A}_{n-1}) = a + \sum_{j=1}^{n-1} E(Y_j|\mathcal{A}_{n-1}) + E(Y_n|\mathcal{A}_{n-1}). \tag{39.87}$$

Si $j \leqslant n-1$ alors $Y_j \in m(\mathcal{A}_{n-1})$. Mais nous savons que si X est $\mathcal{F}$-mesurable, alors $E(X|\mathcal{F}) = X$ (c'est la définition de l'espérance conditionnelle), donc $\sum_{j=1}^{n-1} E(Y_j|\mathcal{A}_{n-1}) = \sum_{j=1}^{n-1} Y_j$.

En ce qui concerne le terme $j = n$ nous utilisons le fait que $\sigma(Y_n)$ soit une tribu indépendante de $\mathcal{A}_{n-1}$; nous avons donc au final pour tout j que $E(Y_j|\mathcal{A}_{n-1})E(Y_j) = p - q$. Nous avons donc

$$E(S_n|\mathcal{A}_{n-1}) = S_{n-1} + p - q. \tag{39.88}$$

Si $p = q = \frac{1}{2}$ alors c'est une martingale, et si $p > q$ c'est une sous-martingale. □

LemXDlNxtE

Lemme 39.34.
La variable aléatoire T est un temps d'arrêt.

Démonstration. Par définition $T = \inf\{n \geqslant 1 \text{ tel que } S_n \in \{0, a+b\}\}$. Vu que les variables aléatoires S_i avec $i \leqslant n$ sont $\mathcal{A}_n$-mesurables, les ensembles $\big\{S_k \notin \{0, a+b\}\big\}$ avec $k \leqslant n$ sont $\mathcal{A}_n$-mesurables. Donc les ensembles

$$\{T = n\} = \bigcap_{k \leqslant n} \big\{S_k \notin \{0, a+b\}\big\} \cap \big\{S_n \in \{0, a+b\}\big\} \tag{39.89}$$

sont $\mathcal{A}_n$-mesurables. Nous en concluons que l'ensemble $\{T \leqslant n\}$ est également mesurable. □

39.4.1 Le cas où la pièce est truquée

Nous supposons être dans le cas $p > q$.

39.4.1.1 Introduction d'une martingale

Considérons le processus

$$\begin{cases} A_0 = 0 & (39.90a)\\ A_n = A_{n-1} + E(S_n - S_{n-1}|\mathcal{A}_{n-1}). & (39.90b)\end{cases}$$

Vu que $E(S_n|\mathcal{A}_{n-1}) = S_{n-1} + p - q$ (lemme 39.33) et que $E(S_{n-1}|\mathcal{A}_{n-1}) = S_{n-1}$ (parce que S_{n-1} est dans la tribu de $\mathcal{A}_{n-1}$), nous avons $A_n = A_{n-1} + (p-q)$ et donc

$$A_n = n(p-q). \tag{39.91}$$

Ce processus (A_n) est croissant et prévisible. Nous introduisons le processus

$$M_n = S_n - A_n \tag{39.92}$$

et nous montrons que c'est une martingale [13]. Nous conditionnons la définition (39.92) par rapport à $\mathcal{A}_{n-1}$:

$$\begin{aligned} E(M_n|\mathcal{A}_{n-1}) &= E(S_n|\mathcal{A}_{n-1}) - \underbrace{E(A_n|\mathcal{A}_{n-1})}_{=A_n} & (39.93a)\\ &= A_n - A_{n-1} + E(S_{n-1}|\mathcal{A}_{n-1}) - A_n & (39.93b)\\ &= E(S_{n-1}|\mathcal{A}_{n-1}) - A_{n-1}. & (39.93c)\end{aligned}$$

Mais S_{n-1} est $\mathcal{A}_{n-1}$-mesurable, donc $E(S_{n-1}|\mathcal{A}_{n-1}) = S_{n-1}$ et

$$E(M_n|\mathcal{A}_{n-1}) = S_{n-1} - A_{n-1} = M_{n-1}, \tag{39.94}$$

ce qui signifie que (M_n) est une martingale.

39.4.1.2 Finitude du temps d'arrêt

Nous montrons maintenant, en étudiant $M_{T\wedge n}$ que T est intégrable et nous prouvons que $P(T = \infty) = 0$.

Proposition 39.35.
La variable aléatoire T est intégrable, et $P(T = \infty) = 0$, c'est-à-dire que le jeu se termine presque certainement après un temps fini.

Démonstration. Nous rappelons que le lemme 39.34 nous indique que T est un temps d'arrêt. Le temps d'arrêt $T \wedge n$ est borné (par n évidemment) et nous pouvons donc lui appliquer le théorème d'arrêt 39.21 pour dire que

$$E(M_{T\wedge n}) = E(M_0). \tag{39.95}$$

Le membre de droite est simple parce que $M_0 = S_0 - A_0 = S_0 = a$ parce que c'est l'argent de départ du joueur. Pour l'autre :

$$E(M_{T\wedge n}) = E(S_{T\wedge n}) - E(A_{T\wedge n}). \tag{39.96}$$

D'une part, $E(A_{T\wedge n}) = E\big((T\wedge n)(p-q)\big)$ et d'autre part, $E(S_{T\wedge n}) \leqslant a + b$ parce que S_T vaut zéro ou $a + b$ (avec des probabilités encore inconnues [14]). En combinant avec ce qui était dit juste au dessus et remarquant que $(p-q)E(T\wedge n) \geqslant 0$ nous pouvons écrire

$$0 \leqslant (p-q)E(T\wedge n) \leqslant b. \tag{39.97}$$

13. Ceci est un peu le contraire de la décomposition de Doob.
14. Mais on y travaille.

La suite de variables aléatoires $T \wedge n$ est donc croissante, positive et intégrable [15] et donc nous avons du travail pour le théorème de la convergence monotone 14.178. La variable aléatoire T est alors mesurable et [16]

$$\lim_{n\to\infty} E(T \wedge n) = E(T). \tag{39.98}$$

Notons que nous n'avons pas encore prouvé que $E(T) < \infty$, mais en passant à la limite dans (39.97) nous écrivons

$$0 \leqslant (p-q)E(T) \leqslant b. \tag{39.99}$$

Maintenant nous avons prouvé que T est intégrable et même L^1. Par conséquent

$$P(T = \infty) = 0. \tag{39.100}$$

Le jeu se termine donc presque certainement après un temps fini. □

39.4.1.3 Temps moyen de jeu

Le lemme 39.17 nous indique que $S_{T\wedge n} \xrightarrow{p.s.} S_T$.

Nous avons les bornes $0 \leqslant S_{T\wedge n} \leqslant a+b$ et comme $a+b$ est intégrable, $S_{T\wedge n}$ l'est aussi et nous pouvons parler de $E(S_{T\wedge n})$. Repartons de (39.96) :

$$a = E(M_0) = E(M_{T\wedge n}) = E(S_{T\wedge n}) - E(A_{T\wedge n}) = E(S_{T\wedge n}) - (p-q)E(T\wedge n). \tag{39.101}$$

La variable aléatoire $S_{T\wedge n}$ est majorée par $a+b$ indépendamment de n ; donc le théorème de la convergence dominée 14.207 donne $\lim_{n\to\infty} E(S_{T\wedge n}) = E(S_T)$. En ce qui concerne le second terme, la convergence dominée ne fonctionne pas parce que $T \wedge n$ n'est pas à priori majoré par quelque chose d'indépendant de n. Heureusement, le théorème de la convergence monotone donne $\lim_{n\to\infty} E(T\wedge n) = E(T)$. Au final en passant à la limite dans (39.101) nous avons

$$a = E(S_T) - (p-q)E(T). \tag{39.102}$$

Étant donné que $T > 0$ et $p - q > 0$ nous pouvons récrire cela sous la forme

$$0 \leqslant (p-q)E(T) = E(S_T) - a. \tag{39.103}$$

Par définition de T nous avons aussi

$$E(S_T) = (a+b)P(S_T = a+b) + 0 \cdot P(S_T = 0) = \rho(a+b). \tag{39.104}$$

Nous déduisons

$$E(T) = \frac{(a+b)\rho - a}{p-q}. \tag{39.105}$$

Ne crions pas victoire trop vite : nous n'avons pas encore d'expression de $\rho = P(S_T = a+b)$. Le temps moyen de jeu n'est donc pas encore tout à fait connu.

39.4.1.4 Probabilité de victoire du joueur

Nous avons besoin d'exprimer ρ en termes de a, b et p. Pour cela nous introduisons la variable aléatoire [17]

$$U_n = \left(\frac{q}{p}\right)^{S_n}. \tag{39.106}$$

Nous commençons par prouver que c'est une martingale en calculant

$$E(U_n|\mathcal{A}_{n-1}) = E\left(\left(\frac{q}{p}\right)^{S_{n-1}}\left(\frac{q}{p}\right)^{Y_n}\Big|\mathcal{A}_{n-1}\right) \tag{39.107}$$

15. Je rappelle que les constantes sont des fonctions intégrables sur Ω. Oui, je sais, quand on est habitué à faire de l'analyse sur $\mathbb{R}^n$ c'est un truc qu'on perd toujours un peu de vue.

16. Dans [104], l'équation (39.98) vient avec une lim sup et non une limite normale. Je ne comprends pas pourquoi.

17. Nous dirons un mot sur ce choix dans le « petit complément » 39.4.3.

Nous utilisons la proposition 36.73. Dans notre cas, S_{n-1} et Y_n sont des variables aléatoires $\mathcal{A}_n$-mesurables ; la variable aléatoire Y_n est même $\mathcal{A}_{n-1}$-mesurable et sort donc du conditionnement ; nous avons donc

$$E(U_n|\mathcal{A}_{n-1}) = \left(\frac{q}{p}\right)^{S_{n-1}} E\left(\left(\frac{q}{p}\right)^{Y_n}\right) \tag{39.108}$$

Nous allons utiliser le théorème de transfert 36.96 :

$$E(s^{Y_n}) = \int_\Omega s^{Y_n(\omega)} dP(\omega) = \int_{Y_n=1} s dP(\omega) + \int_{Y_n=-1} \frac{1}{s} dP(\omega). \tag{39.109}$$

Mais nous savons que $P(Y_n = 1) = p$ et $P(Y_n = -1) = 1 - p = q$, donc

$$E(s^{Y_n}) = ps + \frac{1-p}{s}. \tag{39.110}$$

En posant $s = q/p$ et en tenant compte de $1 - p = q$, nous trouvons

$$E\left(\left(\frac{q}{p}\right)^{Y_n}\right) = p + q = 1. \tag{39.111}$$

Donc

$$E(U_n|\mathcal{A}_{n-1}) = \left(\frac{q}{p}\right)^{S_n - 1} = U_{n-1}, \tag{39.112}$$

ce qui prouve que (U_n) est une martingale.

Par définition nous avons toujours $S_n \geqslant 0$ tant que $n \leqslant T$ [18], donc $U_{T\wedge n} \in [0, 1]$. Il est donc évident que si $a \geqslant 1$ nous avons

$$\int_{|U_{T\wedge n}|>a} |U_{T\wedge n}| dP = 0 \tag{39.113}$$

parce que le domaine d'intégration est vide. Donc les variables aléatoires $V_n = U_{T\wedge n}$ sont équi-intégrables [19] et le théorème 39.25 montre que la martingale (V_n) est terminée ; par ricochet [20] le théorème de Doob 39.26 montre qu'il existe une variable aléatoire X telle que $V_n \xrightarrow{p.s.} X$. Nous allons prouver que $X = U_T$ presque partout. Nous savions déjà (voir l'équation (39.53) et ses alentours) que

$$S_{n\wedge T} \xrightarrow{p.s.} S_T. \tag{39.114}$$

Nous avons alors (au sens du presque surement) :

$$\lim_{n\to\infty} V_n = \lim_{n\to\infty} U_{T\wedge n} = \lim_{n\to\infty} \left(\frac{q}{p}\right)^{S_{T\wedge n}} = \left(\frac{q}{p}\right)^{S_T} = U_T. \tag{39.115}$$

Donc par unicité de la limite presque partout nous avons $X = U_T$ presque partout. Par le théorème de transfert 36.96 nous évaluons

$$E(U_T) = \left(\frac{q}{p}\right)^0 P(S_T = 0) + \left(\frac{q}{p}\right)^{a+b} P(S_T = a + b) = (1 - \rho) + \left(\frac{q}{p}\right)^{a+b} \rho. \tag{39.116}$$

La remarque 39.20 nous permet de dire que

$$E(U_{T\wedge n}) = U_0. \tag{39.117}$$

Mais par définition

$$U_0 = \left(\frac{q}{p}\right)^{S_0} = \left(\frac{q}{p}\right)^a, \tag{39.118}$$

18. Pour $n > T$ le jeu est terminé, donc on ne se pose pas la question.
19. Définition 27.56.
20. Nous rappellons que la convergence L^1 n'implique pas la convergence presque partout.

donc nous avons

$$E(U_{T\wedge n}) = \left(\frac{q}{p}\right)^a. \tag{39.119}$$

Nous voudrions passer à la limite $n \to \infty$ dans cette équation. Pour permuter la limite et l'espérance, il faut utiliser le théorème de la convergence dominée 14.207. Vu que nous avons choisi $q > p$, nous avons $q/p > 1$ et donc $U_{T\wedge n} \leqslant (q/p)^{a+b}$, ce qui montre que la fonction $\omega \mapsto (U_{T\wedge n})(\omega)$ est majorée par une constante (qui est une fonction intégrable). Nous pouvons donc permuter la limite et l'espérance :

$$\lim_{n\to\infty} E(U_{T\wedge n}) = E\Big(\lim_{n\to\infty} U_{T\wedge n}\Big). \tag{39.120}$$

Mais nous avion déjà montré que $U_{T\wedge n} \xrightarrow{p.s.} U_T$. Donc

$$E(U_T) = \left(\frac{q}{p}\right)^a. \tag{39.121}$$

En égalisant avec l'expression (39.116) de $E(U_T)$ nous trouvons

$$\rho = \frac{\left(\frac{q}{p}\right)^a - 1}{\left(\frac{q}{p}\right)^{a+b} - 1} \tag{39.122}$$

et ensuite nous trouvons $E(T)$ en remettant ce ρ dans l'expression (39.105) donnée plus haut.

39.4.2 Le cas où la pièce est non truquée

Maintenant $p = q = 1/2$.

39.4.2.1 Probabilité de gagner

Le lemme 39.33 nous indique alors que (S_n) est une martingale et le lemme 39.32 nous permet de dire que (S_n^2) est alors une sous-martingale. Le processus croissant prévisible de (S_n^2) est donné par (39.69) qui en adaptant les notations est

$$\begin{cases} B_0 = 0 & \text{(39.123a)} \\ B_n = B_{n-1} + E\Big((S_n - S_{n-1})^2 | \mathcal{A}_{n-1}\Big). & \text{(39.123b)} \end{cases}$$

subEqJQbqJEq

Nous avons toujours $S_n - S_{n-1} = \pm 1$ parce que soit le joueur gagne soit le joueur perd, mais de toutes façons sa fortune varie de 1 à chaque étape du jeu. Donc (39.123b) nous donne $B_n = B_{n-1}+1$ et

$$B_n = n. \tag{39.124}$$

Cela nous dit que la variable aléatoire

$$S_n^2 - B_n = S_n^2 - n \tag{39.125}$$

est une martingale (une sur-martingale moins son processus prévisible croissant). Nous lui appliquons le théorème d'arrêt 39.19 avec les temps d'arrêt 0 et $T \wedge n$:

$$E(S_{T\wedge n}^2 - T \wedge n | \mathcal{F}_0) = S_0^2 - 0 \tag{39.126}$$

EqINeyNUV

où $\mathcal{F}_0$ est la tribu engendrée par la variable aléatoire 0, c'est-à-dire $\{\Omega, \varnothing\}$. Cette tribu est indépendante de toute autre tribu et nous pouvons donc supprimer le conditionnement dans (39.126). Nous avons aussi $S_0 = a$ par définition. Avec tout ça nous avons la majoration

$$E(T \wedge n) = E(S_{T\wedge n}^2) - a^2 \leqslant (a+b)^2 - a^2 \tag{39.127}$$

EqQXeFPpq

parce que S_k est toujours positif et entre 0 et $a + b$. En utilisant le lemme 39.17 et en passant à la limite,

$$E(T) \leqslant (a+b)^2 - a^2. \tag{39.128}$$

En particulier, $T \in L^1(\Omega)$ et $P(T < \infty) = 1$.

En suivant exactement les mêmes étapes que dans le lemme 39.17(2) nous avons aussi

$$\lim_{n\to\infty} S_{T\wedge n} = S_T \tag{39.129}$$

presque partout. De plus nous savons que

$$0 \leqslant S_{T\wedge n}^2 \leqslant (a+b)^2, \tag{39.130}$$

et nous pouvons donc utiliser le théorème de la convergence dominée 14.207 pour dire que

$$\lim_{n\to\infty} E(S_{T\wedge n}^2) = E(S_T^2). \tag{39.131}$$

Nous montrons à présent que $S_{T\wedge n} \xrightarrow{L^2} S_T$. Pour cela nous devons évaluer la limite

$$\lim_{n\to\infty} \int_\Omega |S_{T\wedge n} - S_T|^2. \tag{39.132}$$

La fonction $|S_{T\wedge n} - S_T|^2$ est majorée par $(a+b)^2$ et nous pouvons à nouveau appliquer la convergence dominée :

$$\lim_{n\to\infty} E(|S_{T\wedge n} - S_T|^2) = \lim_{n\to\infty} \int_\Omega |S_{T(\omega)\wedge n}(\omega) - S_{T(\omega)}(\omega)|^2 dP(\omega) = \int_\Omega \lim_{n\to\infty} |S_{T\wedge n} - S_T|^2 = 0. \tag{39.133}$$

La même chose en n'écrivant pas les carrés montre que l'on a aussi $S_{T\wedge n} \xrightarrow{L^1} S_T$.

Il n'y a pas que $n \mapsto S_n^2 - n$ qui est une martingale. Il y a aussi (S_n) lui-même (lemme 39.33). Nous pouvons lui appliquer le théorème d'arrêt 39.19 pour les temps d'arrêts $T \wedge n$ et 0 :

$$E(S_{T\wedge n}) = E(S_0) = a. \tag{39.134}$$ EqGWrUwWE

En passant à la limite, $E(S_T) = a$. L'espérance $E(S_T)$ peut par ailleurs être calculée comme

$$E(S_T) = 0 \cdot P(S_T = 0) + (a+b)P(S_T = a+b). \tag{39.135}$$ EqDRXoLXt

En égalisant les valeurs (39.134) et (39.135) de $E(S_T)$ nous trouvons

$$\rho = \frac{a}{a+b}. \tag{39.136}$$ EqIHhbsCR

Cette formule est assez logique : la probabilité que le joueur gagne est égale à la proportion d'argent en jeu qu'il a amené.

39.4.2.2 Temps moyen de jeu

Nous calculons maintenant l'espérance $E(T)$ du temps de jeu (sans compter les pauses ni les jours de fermeture du casino [21]).

Nous recopions la première égalité de (39.127) sous la forme

$$a^2 = E(S_{T\wedge n}^2 - T \wedge n) \tag{39.137}$$

et nous passons à la limite [22] en sachant que $E(S_T^2) = \rho(a+b)^2$:

$$a^2 = \rho(a+b)^2 - E(T). \tag{39.138}$$

En reprenant la valeur (39.136) de ρ,

$$E(T) = ab. \tag{39.139}$$

Et là, on voit que si le joueur amène 1000 euros contre une banque qui en a un million, et si ils jouent toutes les secondes, on en a pour 32 ans de jeu en moyenne.

Voilà. C'est fini pour la ruine du joueur.

21. Le joueur est un *vrai* joueur compulsif.
22. Comme il est dit dans La Grande Illusion, à quoi sert un n ? À passer à la limite.

39.4.3 Un petit complément

SUBSECooACWWooBuSxGv

Nous avons introduit lors de l'équation (39.106) la variable aléatoire $U_n = (p/q)^{S_n}$. Sans aller jusqu'à motiver complètement ce choix, nous nous proposons maintenant de voir que parmi les variables aléatoires $U_n = s^{S_n}$, le choix $s = p/q$ est le seul qui donne une martingale.

Soit donc $U_n = s^{S_n}$ et exprimons le fait que ce soit une martingale. Nous avons

$$E(U_n|\mathcal{A}_{n-1}) = E(s^{S_{n-1}}s^{Y_n}|\mathcal{A}_{n-1}) \tag{39.140a}$$
$$= s^{S_{n-1}}E(s^{Y_n}|\mathcal{A}_{n-1}) \tag{39.140b}$$
subeqYQzYOxS
$$= s^{S_{n-1}}E(s^{Y_n}). \tag{39.140c}$$

Le passage à (39.140b) se justifie en disant que $s^{S_{n-1}}$ est une variable aléatoire bornée et $\mathcal{A}_{n-1}$-mesurable, et en invoquant proposition 36.73. La variable aléatoire Y_n vaut 1 avec probabilité p et -1 avec probabilité q ; donc l'espérance est vite vue :

$$E(s^{Y_n}) = ps + q\frac{1}{s} \tag{39.141}$$

et nous avons

$$E(U_n|\mathcal{A}_{n-1}) = \left(ps + q\frac{1}{s}\right)s^{S_{n-1}} = (ps + \frac{q}{s})U_{n-1}. \tag{39.142}$$

Pour que (U_n) soit une martingale il faut (et il suffit) que

$$ps + \frac{q}{s} = 1. \tag{39.143}$$
EqFMRHybk

Les solutions de cette équation sont $s \in \{1, \frac{p}{q}\}$. C'est évidemment $s = p/q$ qui donne une martingale non triviale. Attention pour être complet, il faut se demander ce qu'il se passe si $s = 0$ séparément parce que manifestement l'équation (39.143) ne traite pas ce cas. Encore une fois, en repartant du début, $s = 0$ ne se révèle pas être une martingale très excitante.

Bref, nous devons poser

$$U_n = \left(\frac{p}{q}\right)^{S_n} \tag{39.144}$$

pour avoir une martingale.

¡¡ Avertissement/question au lecteur !! 39.36
Les variables p et q sont inversées par rapport à (39.106). Écrivez-moi si vous voyez pourquoi.

Chapitre 40

Processus de Poisson

40.1 Processus de Poisson

SecHxbtzQ

On trouve différentes définitions de processus de Poisson dans la littérature. Nous allons définir des processus de a-Poisson, de c-Poisson et de n-Poisson et de o-Poisson. Ensuite nous démontrerons les équivalences.

DEFooEVRIooZQJRSZ

Définition 40.1 (processus de comptage)**.**
*Un **processus de comptage** est une famille de variables aléatoires $\{N_t : \Omega \to \mathbb{R}\}_{t\in\mathbb{R}^+}$ telles que*

(1) $N_0(\omega) = 0$ pour tout $\omega \in \Omega$.

(2) Pour tout $t \in \mathbb{R}^+$ et pour tout $\omega \in \Omega$, $N_t(\omega) \in \mathbb{N}$.

(3) Pour chaque $\omega \in \Omega$, la fonction $t \mapsto N_t(\omega)$ est croissante.

ITEMooXOEPooGewUBL

(4) si $N_t = n > 0$, alors il existe $0 < u < t$ tel que $N_u = n - 1$.

40.1.1 Processus a-Poisson

Définition 40.2.
*Une famille de variables aléatoires $(N_t)_{t\geqslant 0}$ forment un **processus de a-Poisson** d'intensité λ si*

(1) Pour tout choix $0 < t_0 < t_1 < \dots t_n$, les variables aléatoires $N_{t_{i+1}} - N_{t_i}$ sont indépendantes.

(2) Si $0 < s < t$ et $h > 0$ alors

$$N_{t+h} - N_{s+h} \stackrel{\mathcal{L}}{=} N_t - N_s, \tag{40.1}$$

c'est-à-dire que les accroissements décalés suivent les mêmes lois.

(3) Pour tout t nous avons N_t suit une loi de Poisson[1] : $N_t \sim \mathscr{P}(\lambda t)$.

LEMooSOUNooJDMQpM

Lemme 40.3 ([1])**.**
Si $\{N_t\}$ est un processus de a-Poisson, alors

(1) $N_0 = 0$

(2) $N_t - N_s \stackrel{\mathcal{L}}{=} N_{t-s} - N_0 = N_{t-s}$.

40.1.2 Processus c-Poisson

DEFooWDXDooRGCtXL

Définition 40.4 (processus de Poisson[768])**.**
*Un **processus de c-Poisson** d'intensité λ est un processus de comptage $(N_t)_{t\in\mathbb{R}^+}$ tel que*

(1) Pour tout choix de $0 \leqslant t_1 < \dots < t_m$, les variables aléatoires $N_{t_i} - N_{t_{i-1}}$ sont indépendantes[2].

(2) Pour tout $(s,t) \in \mathbb{R}^2$, la variable aléatoire $N_{s+t} - N_s$ suit une loi de Poisson[3] de paramètre λt.

1. Définition 36.143.
2. Ceci demande en particulier que l'ensemble Ω soit assez grand pour contenir une infinité non dénombrable de variables aléatoires indépendantes.
3. Définition 36.143.

40.1.3 Processus n-Poisson

DEFooWXHEooEHQUJU

Définition 40.5.
*Une famille de variables aléatoires $(N_t)_{t\geqslant 0}$ est un **processus de n-Poisson** d'intensité λ si il existe une suite de variables aléatoires indépendantes et identiquement distribuées $(T_k)_{k\in\mathbb{N}}$ de loi exponentielle*[4] *$\mathscr{E}(\lambda)$ telles que*

$$N_t = \sup\{n \geqslant 0 \text{ tel que } \sum_{k=1}^{n} T_k \leqslant t\}. \tag{40.2}$$

Plus explicitement, un processus de Poisson sur espace probabilisé $(\Omega, \mathcal{A}, P)$ est une famille de variables aléatoires $N_t\colon \Omega \to \mathbb{N}$ telles que qu'il existe des variables aléatoires $T_k\colon \Omega \to \mathbb{R}$ vérifiant

$$N_t(\omega) = \sup\{n \geqslant 0 \text{ tel que } \sum_{k=1}^{n} T_k(\omega) \leqslant t\} \tag{40.3}$$

pour tout $\omega \in \Omega$.

Dans le cas d'un processus de n-Poisson, nous posons

$$S_n = \sum_{k=1}^{n} T_k. \tag{40.4}$$

LEMooZTPGooYsFfNX

Lemme 40.6.
Soit un processus N_t de n-Poisson. Alors

(1) $N_t = \sum_{n=1}^{\infty} \mathbb{1}_{\{S_n\leqslant t\}}$.
(2) $N_t \sim \mathscr{P}(\lambda t)$.
(3) $\{s < S_n \leqslant t\} = \{N_t \geqslant n > N_s\}$.
(4) $\{N_t = n\} = \{S_n \leqslant t \leqslant S_{n+1}\}$

Note pour moi : le second point est probablement 36.155.

40.1.4 Processus o-Poisson

DEFooVJDDooHdFJJS

Définition 40.7 ([764])**.**
Un processus de comptage[5] *$N_t\colon \Omega \to \mathbb{N}$ est un processus de o-Poisson si*

(1) Pour tout $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$, les variables aléatoires $N_{t_i} - N_{t_{i+1}}$ sont indépenantes.
(2) Pour tout t, h, la loi de $N_{t+h} - N_t$ ne dépend que de h.
(3) Il existe des fonctions $\alpha_i\colon \mathbb{R} \to \mathbb{R}$ telles que $\lim_{h\to 0} \alpha_i(h)/h = 0$ et
  (3a) $P\big(N_{t+h} - N_t \geqslant 1\big) = \lambda h + \alpha_1(h)$
  (3b) $P\big(N_{t+h} - N_t \geqslant 2\big) = \alpha_2(h)$.
(4) Pour chaque $\omega \in \Omega$, l'application $t \mapsto N_t(\omega)$
  (4a) est continue à droite.
  (4b) a une limite à gauche.

PROPooGXCPooHsJKMH

Proposition 40.8 ([764])**.**
Soit un processus (N_t) de o-Poisson[6] *de paramètre λ. Pour tout $t > 0$, nous avons*

$$P(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}. \tag{40.5}$$

4. Définition 36.145.
5. Définition 40.1.
6. Définition 40.7.

Démonstration. La stratégie est de poser

$$\begin{aligned} p_n\colon \mathbb{R}^+ &\to \mathbb{R} \\ t &\mapsto P(N_t = n), \end{aligned} \tag{40.6}$$

et de chercher à écrire une équation différentielle pour p_n. Nous y allons par récurrence sur n en commençant par $n = 0$. Prouvons l'égalité d'ensembles

$$\{N_{t+h} = 0\} = \{N_t - N_0 = 0\} \cap \{N_{t+h} - N_t = 0\}. \tag{40.7}$$ EQooRIVUooPLTapP

(1) **Inclusion dans un sens** Si $N_{t+h}(\omega) = 0$, alors $N_s(\omega) = 0$ pour tout $s \leqslant t + h$. En particulier $N_t(\omega) = N_0(\omega) = 0$.

(2) **Inclusion dans l'autre sens** Nous savons que $N_0(\omega) = 0$. Donc si $\omega \in \{N_t - N_0 = 0\} \cap \{N_{t+h}(\omega) - N_t(\omega)\}$, nous avons $N_t(\omega) = 0$ et $(N_{t+h} - N_t)(\omega) = 0$. Cela donne $N_{t+h}(\omega) = 0$.

Les variables aléatoires $N_t - N_0$ et $N_{t+h} - N_t$ sont indépendantes [7] : leurs tribus sont indépendantes. Vu que $\{N_t - N_0 = 0\}$ est dans la tribu engendrée par $N_t - N_0$, les événements $\{N_t - N_0 = 0\}$ et $\{N_{t+h} - N_t = 0\}$ sont indépendants [8]. Cela pour dire que, en prenant la probabilité des deux côtés de (40.7), EQooHXOHooYSgikS

$$P(N_{t+h} = 0) = P\big(N_t - N_0 = 0 \cap N_{t+h} - N_t = 0\big) \tag{40.8a}$$
$$= P(N_t - N_0 = 0)P(N_{t+h} - N_t = 0) \tag{40.8b}$$

Pour faire apparaître une dérivée, nous calculons $P(N_{t+h} = 0) - P(N_t = 0)$. En utilisant (40.8) et en tenant compte du fait que $N_0 = 0$, nous avons SUBEQSooBYPTooUkROAX

$$P(N_{t+h} = 0) - P(N_t = 0) = P(N_t = 0)P(N_{t+h} - N_t = 0) - P(N_t = 0) \tag{40.9a}$$
$$= P(N_t = 0)\big(P(N_{t+h} - N_t = 0) - 1\big). \tag{40.9b}$$

De la définition de o-Poisson, nous avons

$$P(N_{t+h} - N_t = 0) = 1 - P(N_{t+h} - N_t \geqslant 1) = 1 - \lambda h + \alpha(t) \tag{40.10}$$

avec $\lim_{t\to 0} \alpha(t)/t = 0$. En substituant cela dans (40.9), nous trouvons

$$P(N_{t+h} = 0) - P(N_t = 0) = -P(N_t = 0)\big(\lambda t + \alpha(h)\big). \tag{40.11}$$

En ce qui concerne la fonction p_0 nous avons

$$\frac{p_0(t+h) - p_0(t)}{h} = -\frac{p_0(t)}{h}(\lambda h + \alpha(h)) = -\lambda p_0(t) + \frac{\alpha(h)}{h}, \tag{40.12}$$

et donc l'équation différentielle

$$\begin{cases} p_0'(t) = -\lambda p_0(t) & \text{(40.13a)} \\ p_0(0) = 1. & \text{(40.13b)} \end{cases}$$

La solution est unique et connue [9] :

$$P(N_t = 0) = p_0(t) = e^{-\lambda t}. \tag{40.14}$$ EQooICAVooCAMKgM

Voila qui prouve le cas $n = 0$.

Nous travaillons sur p_n avec une récurrence. Nous commençons par l'union disjointe

$$\{\omega \in \Omega \text{ tel que } N_{t+h}(\omega) = n\} = \bigcup_{k=0}^{n} \{\omega \in \Omega \text{ tel que } N_t(\omega) = n - k, N_{t+h}(\omega) - N_t(\omega) = k\}. \tag{40.15}$$

7. Définition 36.13.
8. Définition 36.9.
9. Proposition 15.80.

en tenant compte de $N_0(\omega) = 0$ et en prenant la probabilité des deux côtés,

$$P(N_{t+h} = n) = \sum_k P\big(N_t - N_0 = n - k \cap N_{t+h} - N_t = k\big) \tag{40.16a}$$
$$= \sum_{k=0}^{n} P(N_t - N_0 = n - k)P(N_{t+h} - N_t = k) \tag{40.16b}$$
$$= \sum_{k=0}^{n} P(N_t = n - k)P(N_{t+h} - N_t = k). \tag{40.16c}$$

Nous séparons les termes $k = 0$, $k = 1$ et les autres :

$$\begin{aligned} P(N_{t+h} = n) = {} & P(N_t = n)P(N_{t+h} - N_t = 0) \\ & + P(N_t = n - 1)P(N_{t+h} - N_t = 1) \\ & + \sum_{k=2}^{n} P(N_t = n - k)P(N_{t+h} - N_t = k). \end{aligned} \tag{40.17}$$

En utilisant la définition de o-Poisson, et (40.14),

$$P(N_{t+h} - N_t = 0) = P(N_h - N_0 = 0) = P(N_h = 0) = e^{-\lambda h} \tag{40.18}$$

et

$$P(N_{t+h} - N_t = 1) = P(N_{t+h} - N_t \geqslant 1) - P(N_{t+h} - N_t \geqslant 2) \tag{40.19a}$$
$$= \lambda h + \alpha_1(h) - \alpha_2(h) = \lambda h + \alpha(h) \tag{40.19b}$$

De plus, si $k \geqslant 2$ nous avons

$$\lim_{h\to 0} \left| \frac{P(N_{t+h} - N_t) = k}{h} \right| = \lim_{h\to 0} \frac{P(N_{t+h} - N_t = k)}{|h|} \tag{40.20a}$$
$$\leqslant \lim_{h\to 0} \frac{P(N_{t+h} - N_t \geqslant 2)}{|h|} \tag{40.20b}$$
$$\leqslant \lim_{h\to 0} \frac{\alpha(h)}{|h|} \tag{40.20c}$$
$$= 0. \tag{40.20d}$$

Donc nous avons

$$P(N_{t+h} - N_t = k) = \beta(h) \tag{40.21}$$

avec $\lim_{h\to 0} \beta(h) = 0$. En remplaçant (40.18), (40.19) et (40.21) dans (40.17), nous avons

$$p_n(t + h) = p_n(t)e^{-\lambda h} + p_{n-1}(t)(\lambda h + \alpha(h)) + \sum_{k=2}^{n} p_{n-k}(t)\beta(h). \tag{40.22}$$

En soustrayant $p_n(t)$, et en divisant par h nous avons :

$$\begin{aligned} \frac{p_n(t + h) - p_n(t)}{h} = {} & p_n(t)\frac{e^{-\lambda h} - 1}{h} \\ & + \lambda p_{n-1}(t) + p_{n-1}(t)\frac{\alpha(h)}{h} \\ & + \sum_{k=2}^{n} p_{n-k}(t)\frac{\beta(h)}{h}. \end{aligned} \tag{40.23}$$

C'est le moment de prendre la limite $h \to 0$. En utilisant la règle de l'Hospital,

$$\lim_{h\to 0} \frac{e^{-\lambda h} - 1}{h} = \lim_{h\to 0}(-\lambda e^{-\lambda h}) = -\lambda. \tag{40.24}$$

et nous trouvons

$$\begin{cases} p_n'(t) = \lambda\big(p_{n-1}(t) - p_n(t)\big) & (40.25a)\\ p_n(0) = 0 & (40.25b)\\ p_0(t) = e^{-\lambda t}. & (40.25c) \end{cases}$$

Nous devons prouver que ce système possède une unique solution donnée par

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}. \tag{40.26}$$

Nous avons déjà prouvé pour $n = 0$. Supposons que $p_k(t) = e^{-\lambda t}(\lambda t)^k/k!$ est l'unique solution pour tout k jusqu'à n et prouvons que nous avons existence et unicité pour $n + 1$. En ce qui concerne p_{n+1} nous devons prouver l'existence et l'unicité d'une solution à

$$\begin{cases} y'(t) = \lambda\big(p_n(t) - y(t)\big) & (40.27a)\\ y(0) = 0. & (40.27b) \end{cases}$$

En remplaçant $p_n(t)$ par l'hypothèse de récurrence, nous avons l'équation

$$y'(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^n}{n!} - \lambda y(t). \tag{40.28}$$

Si une telle fonction existe, elle est forcément dérivable. Mais sa dérivée est dérivable parce qu'elle est égale à la somme entre une partie C^∞ est λy qui est dérivable. Bref, si une telle fonction existe, elle est de classe C^∞.

En posant

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R}\\ (u, v) &\mapsto \lambda e^{-\lambda u}\frac{(\lambda u)^n}{n!} - \lambda v, \end{aligned} \tag{40.29}$$

nous avons l'équation différentielle $y'(t) = f\big(t, y(t)\big)$. Le théorème de Cauchy-Lipschitz 17.39 [10] dit qu'il existe une unique solution. Il n'est pas compliqué de vérifier que la fonction

$$y(t) = e^{-\lambda t}\frac{(\lambda t)^{n+1}}{(n+1)!} \tag{40.30}$$

vérifie toutes les conditions. □

PROPooGFBEooKgPfaB

Proposition 40.9 ([764]).
Soit un processus o-Poisson (N_t). Nous posons

$$T_n(\omega) = \inf\{t \in \mathbb{R} \text{ tel que } N_t(\omega) \geqslant n\}. \tag{40.31}$$

Alors nous avons l'égalité d'ensembles

$$\{N_t < n\} = \{T_n > t\}. \tag{40.32}$$

Démonstration. En deux parties.

(1) $\{N_t < n\} \subset \{T_n > t\}$ Soient $t > 0$ et $\omega \in \Omega$ tels que $N_t(\omega) < n$. Vu que l'application $u \mapsto N_u(\omega)$ est croissante, si $N_u(\omega) \geqslant n$, alors $u > t$. Donc $T_n(\omega)$ est l'infimum d'un ensemble contenant uniquement des nombres plus grands que t. Nous en déduisons que

$$T_n(\omega) = \inf\{u \in \mathbb{R} \text{ tel que } N_u(\omega) \geqslant n\} \geqslant t. \tag{40.33}$$

EQooIHSZooGkhSKz

Nous devons encore prouver qu'il est impossible d'avoir $T_n(\omega) = t$. Supposons que $T_n(\omega) = t$. Alors pour tout $\epsilon > 0$ nous avons $N_{t+\epsilon}(\omega) \geqslant n$. Par la continuité à droite, nous avons

$$N_t(\omega) = \lim_{\epsilon\to 0} N_{t+\epsilon}(\omega) \geqslant n. \tag{40.34}$$

Vu que nous avons supposé que $N_t(\omega) < n$, nous avons une contradiction et nous concluons que $T_n(\omega) \neq t$. Ça, avec (40.33), dit que $T_n(\omega) > t$.

10. Qui n'est pas cité dans [764]. Est-ce que j'en fais trop ?

(2) $\{T_n > t\} \subset \{N_t < n\}$ Soit $\omega \in \Omega$ tel que $T_n(\omega) > t$:

$$\inf\{u \geqslant 0 \text{ tel que } N_u(\omega) \geqslant n\} > t. \tag{40.35}$$

Si $N_t(\omega) \geqslant n$, alors $t \in \{u > 0 \text{ tel que } N_u(\omega) \geqslant n\}$ et donc $T_n(\omega) \leqslant t$, ce qui est contraire à l'hypothèse.

□

LEMooYKOIooAUuivH

Lemme 40.10 ([1]).
Soit une application croissante et continue à droite $f : [0, \infty[\to \mathbb{N}$. Nous supposons que si $f(x_1) = n > 0$, alors il existe $x_2 < x_1$ tel que $f(x_2) = n - 1$.

Nous posons

$$x_n = \inf\{x \in \mathbb{R} \text{ tel que } f(x) \geqslant n\}, \tag{40.36}$$

et nous supposons que x_n existe au sens où il existe x tel que $f(x) \geqslant n$. Alors

ITEMooHSCEooKVtXYn

(1) $f(x_n) = n$

(2) nous avons $x_n - x_{n-1} > \delta$ si et seulement si $f(x_{n-1} + \delta) - f(x_{n-1}) = 0$.

Démonstration. En plusieurs parties.

(1) $f(x_n) = n$ Vu que f est croissante, qu'il existe x tel que $f(x) \geqslant n$ et que f passe par les valeurs intermédiaires (entières), il existe t tel que $f(t) = n$. Donc nous avons

$$x_n = \inf\{u \in \mathbb{R} \text{ tel que } f(u) \geqslant n\} = \inf\{u \in \mathbb{R} \text{ tel que } f(u) = n\}. \tag{40.37}$$

Nous posons $A = \{u \in \mathbb{R} \text{ tel que } f(u) = n\}$, et nous savons que A n'est pas vide. Si A contient un seul élément, alors cet élément est l'infimum et nous avons $A = \{x_n\}$, de telle sorte que $f(x_n) = n$.

Nous supposons que A contienne au moins deux éléments. Il existe donc $\delta > 0$ tel que $x_n + \delta \in A$. Du coup $f(x_n + \delta) = n$ et par croissance, $f(x_n + \epsilon) = n$ pour tout $\epsilon \in]0, \delta]$. En prenant la limite $\epsilon \to 0^+$ et en utilisant la continuité à droite, $f(x_n) = n$.

(2) $\Rightarrow$ Nous supposons que $x_n - x_{n-1} > \delta$ et que $f(x_{n-1} + \delta) \neq f(x_{n-1})$. Par (1), nous savons que $f(x_{n-1} = n - 1)$, et par croissance, nous savons alors que $f(x_{n-1} + \delta) \geqslant n$. Donc

$$x_n = \inf\{x \text{ tel que } f(x) \geqslant n\} \leqslant x_{n-1} + \delta. \tag{40.38}$$

Cela signifie que $x_n - x_{n-1} \leqslant \delta$, ce qui est contraire à l'hypothèse. Contradiction.

(3) $\Leftarrow$ Nous y allons encore par l'absurde en supposant $f(x_{n-1}+\delta) = f(x_{n-1})$ et $x_n - x_{n-1} \leqslant \delta$. Vu que f est croissante et que $x_{n-1} + \delta \geqslant x_n$, nous avons

$$f(x_{n-1} + \delta) \geqslant f(x_n) = n > n - 1 = f(x_{n-1}). \tag{40.39}$$

Nous avons prouvé que $f(x_{n-1} + \delta) > f(x_{n-1})$, qui est contraire à l'hypothèse.

□

PROPooVLEEooLvCknn

Proposition 40.11 ([764]).
Soit un processus o-Poisson (N_t). Nous posons

$$T_n(\omega) = \inf\{t \in \mathbb{R} \text{ tel que } N_t(\omega) \geqslant n\}. \tag{40.40}$$

Alors

$$N_{T_n(\omega)}(\omega) = n \tag{40.41}$$

pour tout $\omega \in \Omega$.

Démonstration. En fixant $\omega \in \Omega$, et en posant $f(t) = N_t(\omega)$, la fonction f vérifie les hypothèses du lemme 40.10. Nous concluons avec 40.10(1). □

PROPooIQPSooHtbFBt

Proposition 40.12 ([764]).
Soit un processus o-Poisson (N_t). Nous posons

$$T_n(\omega) = \inf\{t \in \mathbb{R} \text{ tel que } N_t(\omega) \geqslant n\}. \tag{40.42}$$

Alors

(1) Pour chaque n, la variable aléatoire $T_n - T_{n-1}$ suit une loi exponentielle de paramètre λ.

(2) Les variables aléatoires $T_n - T_{n-1}$ sont indépendantes.

Démonstration. Nous considérons la proposition 40.9 avec $n = 1$: $P(N_t = 1) = P(T_1 > t)$. Vu que $N_t(\omega) \in \mathbb{N}$ pour tout ω, nous avons $\{N_t < 1\} = \{N_t = 0\}$ et donc, en utilisant la proposition 40.8,

$$P(T_1 > t) = P(N_t = 0) = e^{-\lambda t}. \tag{40.43}$$

En fixant $\omega \in \Omega$, et en posant $f(t) = N_t(\omega)$, la fonction f vérifie les hypothèses du lemme 40.10. Pour $t > 0$ fixé, nous avons $T_n(\omega) - T_{n-1}(\omega) > t$ si et seulement si $N_{T_{n-1}(\omega)+t}(\omega) - N_{T_{n-1}(\omega)}(\omega) = 0$. Autrement dit,

EQooVHJOooTVvnzw

$$\{\omega \in \Omega \text{ tel que } T_n(\omega) - T_{n-1}(\omega) > t\} = \{\omega \in \Omega \text{ tel que } N_{T_{n-1}(\omega)+t}(\omega) - N_{T_{n-1}(\omega)}(\omega) = 0\}. \tag{40.44}$$

Et donc, en ce qui concerne les probabilités,

$$P(T_n - T_{n-1} > t) = P(N_{T_{n-1}+t} - N_{T_{n-1}} = 0). \tag{40.45}$$

En utilisant les accroissements indépendants,

$$\begin{aligned} P(T_n - T_{n-1} > t) &= P(N_{T_{n-1}+t} - N_{T_{n-1}} = 0) && \text{(40.46a)}\\ &= P(N_t - N_0 = 0) && \text{(40.46b)}\\ &= P(N_t = 0) && \text{(40.46c)}\\ &= e^{-\lambda t}. && \text{(40.46d)} \end{aligned}$$

Cela prouve que $T_n - T_{n-1}$ suit une loi exponentielle par la proposition 36.146.

Nous prouvons à présent que les $T_n - T_{n-1}$ sont indépendants. Nous repartons de (40.44) que nous recopions de façon plus compacte :

$$\{T_n - T_{n-1} > t\} = \{N_{T_{n-1}+t} - N_{T_{n-1}} = 0\}. \tag{40.47}$$

Nous avons donc

$$\{T_n - T_{n-1} > t\} \cap \{T_k - T_{k-1} > u\} = \{N_{T_{n-1}+t} - N_{n-1} = 0\} \cap \{N_{T_{k-1}+u} - N_{T_{k-1}} = 0\}. \tag{40.48}$$

À droite, les événements sont indépendants, donc

$$\begin{aligned} P(T_n - T_{n-1} > t, T_k - T_{k-1} > u) &= P(N_{T_{n_1}+t} - N_{T_{n-1}} = 0)P(N_{T_{k-1}+u} - N_{T_{k-1}} = 0) && \text{(40.49a)}\\ &= P(T_n - T_{n-1} > t)P(T_k - T_{k-1} > u). && \text{(40.49b)} \end{aligned}$$

La proposition 36.147 conclut que $T_n - T_{n-1}$ et $T_k - T_{k-1}$ sont indépendants. □

PROBooNZCUooMteOpL

¡¡ Avertissement/question à la lectrice !! 40.13
Je ne suis pas sûr du tout de l'énoncé de la proposition 40.14. D'abord la source [764] ne donne pas le facteur e^{t-t_n} (mais je ne suis pas convaincu de la preuve donnée). Ensuite, j'aurais été nettement plus convaincu avec un changement de signe dans l'exponentielle : $e^{-(t-t_n)}$.

Essayez de répondre à cette question :

https://math.stackexchange.com/questions/4957480/density-of-the-vector-of-jump-times-in-a-poisson-process

PROPooEUVXooERgwIe

Proposition 40.14 ([764, 1]).
La densité conditionnelle de $(T_1, \dots, T_n)$ sachant $\{N_t = n\}$ est donnée par

$$f_n(t_1, \dots, t_n) = \mathbb{1}_{\{0 \leqslant t_1 \leqslant \dots \leqslant t_n\}} \frac{n!}{t^n} e^{t - t_n}. \tag{40.50}$$

Démonstration. Nous commençons par prouver cette égalité d'ensembles :

$$\bigcap_{i=1}^{n} \{T_i = t_i\} \cap \{N_t = n\} = \bigcap_{i=1}^{n} \{T_i - T_{i-1} = t_i - t_{i-1}\} \cap \{T_{n-1} - T_n = t - t_n\}. \tag{40.51}$$

EQooOKVVooTZSpGq

(1) **Dans un sens** Nous considérons $\omega \in \Omega$ tel que

$$\begin{cases} T_i(\omega) = t_i & i = 1, \dots, n \\ N_t(\omega) = n. \end{cases} \begin{matrix} \text{SUBEQooIZXNooNxrpnC} \\ \text{SUBEQooYSCVooMGOMqZ} \end{matrix} \begin{matrix} (40.52a) \\ (40.52b) \end{matrix}$$

et nous voulons prouver que

$$\begin{cases} T_i(\omega) - T_{i-1}(\omega) = t_i - t_{i-1} & i = 1, \dots, n \\ T_{n-1}(\omega) - T_n(\omega) > t - t_n. \end{cases} \begin{matrix} \text{SUBEQooDBKLooSixBrh} \\ \text{SUBEQooEHYAooCSFDXS} \end{matrix} \begin{matrix} (40.53a) \\ (40.53b) \end{matrix}$$

Les égalités (40.53a) sont immédiates à partir des égalités (40.52a). En ce qui concerne (40.53b), nous avons $T_{n+1}(\omega) - T_n(\omega) = T_{n+1}(\omega) - t_n$. Nous devons donc prouver que $T_{n+1}(\omega) > t$.
En utilisant la proposition 40.11 et (40.52b), nous avons $N_{T_{n+1}(\omega)}(\omega) = n + 1$ et $N_t(\omega) = n$. Par croissance de $u \mapsto N_u(\omega)$, nous avons $T_{n+1}(\omega) > t$.

(2) **Dans l'autre sens** Nous supposons maintenant que

$$\begin{cases} T_i(\omega) - T_{i-1}(\omega) = t_i - t_{i-1} & i = 1, \dots, n \\ T_{n-1}(\omega) - T_n(\omega) > t - t_n, \end{cases} \begin{matrix} \text{SUBEQooXKXPooWHrXlb} \\ \\ \end{matrix} \begin{matrix} (40.54a) \\ (40.54b) \end{matrix}$$

et nous prouvons que

$$\begin{cases} T_i(\omega) = t_i & i = 1, \dots, n \\ N_t(\omega) = n. \end{cases} \begin{matrix} (40.55a) \\ (40.55b) \end{matrix}$$

Nous savons que $T_0(\omega) = 0$. Donc (40.54a) avec $i = 1$ donne $T_1(\omega) - 0 = t_1 - t_0$, c'est-à-dire $T_1(\omega) = t_1$. Nous continuons de proche en proche pour obtenir $T_i(\omega) = t_i$. Le dernier est

$$T_n(\omega) - T_{n-1}(\omega) = t_n - t_{n-1}, \tag{40.56}$$

et donc encore $T_n(\omega) = t_n$.
Nous avons aussi

$$T_{n+1}(\omega) - T_n(\omega) > t - t_n, \tag{40.57}$$

et donc $T_{n+1}(\omega) - t_n > t - t_n$, ce qui signifie $T_{n+1}(\omega) > t$, et donc $N_t(\omega) < n + 1$. Mais comme $t \geqslant t_n$ et $N_{t_n}(\omega) = n$, nous avons bien $N_t(\omega) \geqslant n$ et donc $N_t(\omega) = n$.

Cela prouve l'égalité (40.51).

En notant $X_i = T_i - T_{i-1}$, nous avons

$$T_k = \sum_{l=1}^{k} X_l. \tag{40.58}$$

Donc la loi de T_k est celle d'une somme de variables aléatoires exponentielles indépendantes.

Oh zut ! cette démonstration n'est pas terminée ! Si tu sais comment la terminer, écris-moi. □

PROPooGMBBooCIkVCB

Proposition 40.15 ([768]).
Soit un processus de o-Poisson (N_t). Pour tout $t \geqslant 0$ nous avons

$$\lim_{h\to 0^+} P\big(N_{t+h} - N_t \geqslant 1\big) = 0. \tag{40.59}$$

LEMooXTVNooSBnnAd

Lemme 40.16 ([1]).
Soit un processus de o-Poisson $\{N_t : \Omega \to \mathbb{N}\}_{t\in\mathbb{R}^+}$ d'intensité λ. Pour chaque $n \in \mathbb{N}$ nous posons

$$\begin{aligned} T_n : \Omega &\to \mathbb{R} \\ \omega &\mapsto \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant n\}. \end{aligned} \tag{40.60}$$

Soit $h > 0$.

ITEMooXVSQooSQmIUv

(1) Il existe une partie de mesure nulle[11] *$Z \in \Omega$ telle que $\{T_n \leqslant h\} \subset \{N_h \geqslant n\} \cup Z$.*

ITEMooIPKCooIfjUzS

(2) Nous avons $\{N_h \geqslant n\} \subset \{T_n \leqslant h\}$.

(3) Au final,

$$P(T_n \leqslant h) = P(N_h \geqslant n). \tag{40.61}$$

Démonstration. En plusieurs parties

(1) $\{T_n \leqslant h\} \subset \{N_n \geqslant n\} \cup Z$ Soit $\omega \in \{T_n \leqslant h\}$. Cela signifie que

$$\inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant n\} \leqslant h. \tag{40.62}$$

Si $T_n(\omega) < h$, alors $N_h(\omega) \geqslant n$, et le point est démontré.
Nous prouvons maintenant que $Z = \{T_n = h\} \cap \{N_h \leqslant n\}$ est de mesure nulle. Soient $\omega \in Z$ ainsi que $\epsilon > 0$. Vu que $T_n(\omega) = h$, nous avons $N_{h+\epsilon}(\omega) \geqslant n$. Donc, pour $\omega \in Z$ nous avons

$$N_{h+\epsilon}(\omega) - N_h(\omega) \geqslant 1. \tag{40.63}$$

En posant $Z_\epsilon = \{N_{h+\epsilon} - N_h \geqslant 1\}$, nous avons donc $Z \subset Z_\epsilon$ pour tout ϵ. Nous avons alors, pour tout ϵ, que $P(Z) \leqslant P(Z_\epsilon)$. Mais la proposition 40.15 nous indique que $\lim_{\epsilon\to 0} P(Z_\epsilon) = 0$. Donc $P(Z) = 0$.

(2) $\{N_h \geqslant n\} \subset \{T_n \leqslant h\}$ Soit $\omega \in \Omega$ satisfaisant $N_h(\omega) \geqslant n$. C'est direct parce que si $N_h(\omega) \geqslant n$, alors

$$T_n(\omega) = \inf\{y \geqslant 0 \text{ tel que } N_t(\omega) \geqslant n\} \leqslant h. \tag{40.64}$$

(3) **Les probabilités** Le point (1) donne

$$P(T_n \leqslant h) \leqslant P(N_h \geqslant n) + P(Z) = P(N_h \geqslant n), \tag{40.65}$$

le point (2) donne $P(N_h \leqslant n) \leqslant P(T_n \leqslant h)$. Au final, nous avons l'égalité des probabilités.

□

LEMooJHMTooQWvOaI

Lemme 40.17 ([803]).
Soient $\lambda > 0$ et $t > 0$. Nous posons

$$I_n = \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} dx. \tag{40.66}$$

Alors

$$I_{n+1} = I_n - e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \tag{40.67}$$

11. Cette subtilité est manquante dans [802, 764, 803]. Plus généralement, je ne l'ai vue nulle part. Si vous savez démontrer que les événements $T_n \leqslant h$ et $N_h \geqslant n$ sont égaux, dites-le moi.

Démonstration. Nous intégrons I_{n+1} par partie en posant $u(x) = x^n$ et $v'(x) = \lambda e^{-\lambda x}$; $u'(x) = nx^{n-1}$, $v(x) = -e^{-\lambda x}$. Cela donne

$$I_{n+1} = \frac{\lambda^n}{n!} \int_0^t v'(x)u(x)dx \tag{40.68a}$$

$$= \frac{\lambda^n}{n!} \left[-x^n e^{-\lambda x} \right]_0^t + \frac{\lambda^n}{n!} \int_0^t nx^{n-1} e^{-\lambda x} dx \tag{40.68b}$$

$$= -\frac{(\lambda t)^n}{n!} e^{-\lambda t} + I_n \tag{40.68c}$$

□

Lemme 40.18 ([1])**.**
Soit un processus de o-Poisson $\{N_t : \Omega \to \mathbb{N}\}_{t\in\mathbb{R}^+}$ d'intensité λ. Pour chaque $n \in \mathbb{N}$ nous posons

$$\begin{aligned} T_n : \Omega &\to \mathbb{R} \\ \omega &\mapsto \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant n\}. \end{aligned} \tag{40.69}$$

Soit $t \geqslant 0$. Pour tout n nous avons

$$P(T_n \leqslant t) = \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} dx. \tag{40.70}$$ EQooRHQLooVxEfiV

Démonstration. Nous y allons par récurrence sur n. D'abord pour $n = 0$ nous savons que $T_0 = 0$, donc

$$P(N_n = 0) = P(T_0 \leqslant t) - P(T_1 \leqslant t). \tag{40.71}$$ EQooNCMMooHnvIpQ

Nous savons que $N_{s+t} - N_s$ suit une loi de Poisson de paramètre λt. En particulier avec $s = 0$, nous voyons que N_t suit une loi de Poisson de paramètre λt. Donc $P(N_t = 0) = e^{-\lambda t}$ (définition 36.143).

L'équation 40.71 donne donc $e^{-\lambda t} = 1 - P(T_1 \leqslant t)$. Par ailleurs, pour $n = 1$, l'intégrale (40.70) est facile à calculer et donne bien $1 - e^{-\lambda t}$.

Pour la récurrence, en reprenant la notation du lemme 40.17, nous supposons que $P(T_n \leqslant t) = I_n$, et nous prouvons que $P(T_{n+1} \leqslant t) = I_{n+1}$. Pour cela nous repartons de

$$P(N_t = n) = P(T_n \leqslant t) - P(T_{n+1} \leqslant t), \tag{40.72}$$

et nous substituons $P(N_t = n) = e^{-\lambda t}(\lambda t)^n/n!$ ainsi que $P(T_n \leqslant t) = I_n$. Il reste

$$P(T_{n+1} \leqslant t) = I_n - e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \tag{40.73}$$

Et le lemme 40.17 nous indique que le tout est égal à I_{n+1}. □

Voici un théorème donnant des propriétés d'un processus de o-Poisson. Une réciproque partielle est donnée dans la proposition 36.155.

THOooYRIMooSREVEO

Théorème 40.19 ([804, 768, 1])**.**
Soit un processus de o-Poisson $\{N_t : \Omega \to \mathbb{N}\}_{t\in\mathbb{R}^+}$ d'intensité λ. Pour chaque $n \in \mathbb{N}$ nous posons

$$\begin{aligned} T_n : \Omega &\to \mathbb{R} \\ \omega &\mapsto \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant n\}. \end{aligned} \tag{40.74}$$

et $S_k = T_k - T_{k-1}$.

Alors

ITEMooSLWOooRjYhBA

(1) T_n est une variable aléatoire.

ITEMooHGXTooOdonXa

(2) *Pour tout n, la variable aléatoire $(T_1, \dots, T_n)$ a pour densité*[12] *la fonction*

$$\begin{aligned} g\colon \mathbb{R}^n &\to \mathbb{R} \\ (x_1, \dots, x_n) &\mapsto \mathbb{1}_{0\leqslant x_1\leqslant \dots \leqslant x_n} \lambda^n e^{-\lambda x_n}. \end{aligned} \tag{40.75}$$

ITEMooGBIXooSjXpsx

(3) *Les variables aléatoires S_i sont à densité.*

ITEMooUTTWooOUhkZh

(4) *La variable aléatoire $(S_1, \dots, S_n)$ est à densité.*

ITEMooUZGFooAbCvvG

(5) *(S_n) est une suite de variables aléatoires indépendantes et identiquement distribuées de loi exponentielle*[13] *$\mathscr{E}(\lambda)$.*

ITEMooDXZFooXkklhG

(6) *La variable aléatoire T_n suit la loi de densité*

$$\begin{aligned} f_n\colon \mathbb{R} &\to \mathbb{R} \\ s &\mapsto \begin{cases} \frac{\lambda}{(n-1)!} e^{-\lambda s} (\lambda s)^{n-1} & si\ s \geqslant 0 \\ 0 & sinon. \end{cases} \end{aligned} \tag{40.76}$$

Démonstration. En plusieurs parties.

Pour (1)

Pour (2)

Soient des nombres a_i et b_i tels que

$$0 < a_1 < b_1 < a_2 < \dots < a_n < b_n. \tag{40.77}$$

Pour éviter les petits effets de bord, nous supposons que $n > 2$. Nous considérons les événements

$$A_n = \bigcap_{i=1}^{n} \{T_i \in]a_i, b_i[\} \tag{40.78}$$

et

$$B_n = \bigcap_{i=1}^{n-1} \{N_{b_i} - N_{a_i} = 1\} \cap \bigcap_{i=1}^{n-1} \{N_{a_{i+1}} - N_{b_i} = 0\} \cap \{N_{b_n} - N_{a_n} \geqslant 1\} \cap \{N_{a_1} = 0\}, \tag{40.79}$$

et nous allons prouver qu'il existe des parties Y et Z de Ω telles que[14]

$$\begin{cases} A_n \subset B_n \cup Y & (40.80a) \\ B_n \subset A_n \cup Z & (40.80b) \\ P(Y) = P(Z) = 0. & (40.80c) \end{cases}$$

Courage.

(1) **Quelques valeurs au bord dans A_n** Soit $\omega \in A_n$.

(1a) **Si $i < n$** Nous montrons que si $i < n$, alors SYSooCWKCooLtrVhn

$$\begin{cases} N_{a_i}(\omega) = i - 1 & (40.81a) \\ N_{b_i}(\omega) \in \{i-1, i\}. & (40.81b) \end{cases}$$

12. Définition 36.5.
13. Définition 36.145.
14. Je suis impressionné à quel point cette subtilité est absente de [768]. Soit il y a un truc macroscopique qui m'échappe, soit il y a un gros trou dans les livres sur lesquels est basé [768], soit il y a quelque chose que je ne comprends pas.

Pour cela, rien de tel qu'une récurrence. D'abord

$$T_1(\omega) = \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant 1\} \in]a_1, b_1]. \tag{40.82}$$

Donc $T_1(\omega) > a_1$ et nous déduisons que $N_{a_1}(\omega) = 0$. De la même manière [15],

$$T_2(\omega) = \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant 2\} \in]a_2, b_2]. \tag{40.83}$$

En particulier $T_2(\omega) > a_2 > b_1$ et nous avons $N_{b_1}(\omega) < 2$. Donc $N_{b_1} \in \{0, 1\}$.

Soit $i < n-1$, et supposons que (40.81) soit valide. Nous avons $T_i(\omega) \leqslant b_i$, et donc $N_{b_i+h}(\omega) \geqslant i$ pour tout $h > 0$. En particulier $N_{a_{i+1}}(\omega) \geqslant i$. Mais $T_{i+1}(\omega) > a_{i+1}$. Donc $N_{a_{i+1}}(\omega) < i+1$. Nous en déduisons que $N_{a_{i+1}}(\omega) = i$.

Comme $i + 1 < n$ nous pouvons encore écrire

$$T_{i+2}(\omega) > a_{i+2} > b_{i+1}, \tag{40.84}$$

donc $N_{b_{i+1}}(\omega) < i+2$. Vu que $N_{b_{i+1}}(\omega) \geqslant N_{a_{i+1}}(\omega) = i$, nous avons $N_{b_{i+1}}(\omega) \in \{i, i+1\}$.

(1b) **Pour $i = n$** De la même manière, nous avons $N_{a_n} < n$, et $T_{n-1}(\omega) < a_n$. Nous en déduisons $N_{a_n}(\omega) \geqslant n-1$ et donc $N_{a_n} = n-1$.

Nous avons aussi $N_{b_n}(\omega) \geqslant n-1$ et, vu que $T_n(\omega) \in]a_n, b_n]$, nous avons $N_{b_n+h}(\omega) \geqslant n$ pour tout $h > 0$.

(2) **Une partie de mesure nulle** Pour $i = 1, \dots, n$, nous posons $Z_i = \{N_{b_i} = i-1\}$. Ensuite nous définissons $Z = \bigcup_{i=1}^n Z_i$. Nous démontrons que $P(Z \cap A_n) = 0$. Si $\omega \in Z \cap A_n$, alors $N_{b_i}(\omega) = i-1$ et $N_{b_i+h}(\omega) \geqslant i$ pour tout $h > 0$. Donc

$$Z_i \cap A_n \subset \{N_{b_i+h} - N_{b_i} \geqslant 1\}. \tag{40.85}$$

Donc

$$P(Z_i \cap A_n) \leqslant P\big(\{N_{b_i+h} - N_{b_i} \geqslant 1\}\big). \tag{40.86}$$

En prenant la limite $h \to 0^+$, la proposition 40.15 fait que

$$P(Z_i \cap A_n) = 0. \tag{40.87}$$

Nous en déduisons que $P(Z \cap A_n) = 0$.

(3) $A_n \subset B_n \cup Z$ Soit $\omega \in A_n$. Nous devons montrer que si $\omega \notin Z$, alors $\omega \in B_n$, c'est à dire

(3a) $N_{b_i}(\omega) - N_{a_i}(\omega) = 1$ pour $i = 1, \dots, n-1$,

(3b) $N_{a_{i+1}}(\omega) - N_{b_i}(\omega) = 0$ pour $i = 1, \dots, n-1$,

(3c) $N_{b_n}(\omega) - N_{a_n}(\omega) \geqslant 1$

(3d) $N_{a_1}(\omega) = 0$

Si $i \leqslant n-1$, alors $N_{a_i}(\omega) = i-1$ et $N_{b_i}(\omega) \in \{i-1, i\}$. Mais si ω n'est pas dans Z, alors $N_{b_i}(\omega) = i$. Dans ce cas, nous avons bien $\omega \in \{N_{b_i} - N_{a_i} = 1\}$.

De même si ω est dans A_n mais pas dans Z, il vérifie $N_{a_{i+1}}(\omega) = i = N_{b_i}(\omega) = i$.

Pour $i = n$, nous avons $N_{a_n}(\omega) = n-1$ et $N_{b_n}(\omega) \geqslant n-1$. Et encore une fois, si ω n'est pas dans Z, alors $N_{a_n}(\omega) - N_{b_n}(\omega) \geqslant 1$.

Enfin pour $\omega \in A_n$, nous avons $T_1(\omega) = \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant 1\} > a_1$. Donc $N_{a_1}(\omega) = 0$.

Nous avons fini de prouver que $A_n \subset B_n \cup Z$.

(4) $P(A_n) \leqslant P(B_n)$ En utilisant le lemme 14.23(2), et (4), nous avons

$$P(A_n) \leqslant P(B_n \cup Z) \leqslant P(B_n) + P(Z) = P(B_n). \tag{40.88}$$

EQooGLVAooGGSZYn

15. Nous utilisons le fait que $n > 2$ pour être sûr ; mais je crois que ce n'est pas nécessaire.

(5) **Quelques valeurs au bord pour B_n** Soit $\omega \in B_n$. Nous avons $N_{a_1}(\omega) = 0$. Donc

$$T_1(\omega) = \inf\{t \geqslant 0 \text{ tel que } N_t(\omega) \geqslant 1\} \geqslant a_1. \tag{40.89}$$

De même nous prouvons les valeurs suivantes. Pour $i < n$ nous avons :

$$\begin{cases} N_{a_i}(\omega) = i - 1 & \text{(40.90a)} \\ N_{b_i}(\omega) = i & \text{(40.90b)} \end{cases}$$

et pour $i = n$ nous avons

$$\begin{cases} N_{a_n}(\omega) = n - 1 & \text{(40.91a)} \\ N_{b_n}(\omega) \geqslant n. & \text{(40.91b)} \end{cases}$$

Nous avons donc l'appartenance

$$T_i(\omega) \in [a_i, b_i] \tag{40.92}$$ EQooVYVJooHEshGE

pour tout $i \leqslant n$.
Cela n'est pas suffisant pour dire que $\omega \in A_n$ parce que la possibilité $T_i(\omega) = a_i$ n'est pas exclue.

(6) **Une partie de mesure nulle** Nous posons $Y_i = \{\omega \in \Omega \text{ tel que } T_i(\omega) = a_i\}$, et nous prouvons que $P(Y_i \cap B_n) = 0$. Pour $\omega \in Y_i \cap B_n$ nous avons $N_{a_i}(\omega) = i - 1$ et $N_{a_i+h}(\omega) = i$ pour tout $h > 0$. Donc pour tout $h > 0$ nous avons

$$Y_i \cap B_n \subset \{N_{a_i+h} - N_{a_i} \geqslant 1\}, \tag{40.93}$$

et donc

$$P(Y_i \cap B_n) \leqslant P\big(N_{a_i+h} - N_{a_i} \geqslant 1\big). \tag{40.94}$$

En utilisant à nouveau le lemme 40.15, nous trouvons $P(Y_i \cap B_n) = 0$.
Nous posons $Y = \bigcup_{i=1}^{n} Y_i$.

(7) $P(B_n) \leqslant P(A_n)$ Si $\omega \in B_n$, nous avons (40.92) : $T_i(\omega) \in [a_i, b_i]$. Par définition de Y, si $\omega \in B_n \backslash Y$ nous avons

$$T_i(\omega) \in]a_i, b_i]. \tag{40.95}$$

Autrement dit, $B_n \backslash Y \subset A_n$ et donc $B_n \subset A_n \cup Y$. Nous en déduisons que

$$P(B_n) \leqslant P(A_n \cup Y) \leqslant P(A_n) + P(Y) = P(A_n). \tag{40.96}$$ EQooWQZPooPBjPsG

(8) $P(A_n) = P(B_n)$ En mettant ensemble (40.88) et (40.96) nous avons

$$P(A_n) = P(B_n). \tag{40.97}$$

Calculons la probabilité de B_n. Vu que $N_0 = 0$, l'événement $N_{a_i} = 0$ peut être écrit $N_{a_i} - N_0 = 0$. Donc B_n peut être écrit comme intersection d'incréments indépendants. Nous pouvons donc écrire $P(A_n)$ comme un produit :

$$P(A_n) = \prod_{i=1}^{n-1} P(N_{b_i} - N_{a_i} = 1) \times \prod_{i=1}^{n-1} P(N_{a_{i+1}} - N_{b_i} = 0) \times P(N_{b_n} - N_{a_n} \geqslant 1) \times P(N_{a_1} = 0). \tag{40.98}$$

En posant $l_i = b_i - a_i$ et $\delta_i = a_{i+1} - b_i$ et en utilisant la formule (36.443) pour

$$P(N_{s+t} - N_s = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \tag{40.99}$$

nous trouvons

$$P(A_n) = \prod_{i=1}^{n-1} (\lambda l_i) e^{-\lambda l_i} \times \prod_{i=1}^{n-1} e^{-\lambda \delta_i} \times e^{-\lambda a_1} (1 - e^{-\lambda l_n}). \tag{40.100}$$

Note : pour le dernier nous avons posé $X = N_{b_n} - N_{a_n}$ qui suit une loi exponentielle de paramètre λl_n et nous avons dit $P(X \geqslant 1) = 1 - P(X = 0) = 1 - e^{\lambda l_n}$. Remarquez que $l_i + \delta_i = a_{i+1} - a_i$; en regroupant les exponentielles, nous avons une somme téléscopique ; au final

$$P(A_n) = \lambda^{n-1} e^{-\lambda a_n}(1 - e^{-\lambda l_n}) \prod_{i=1}^{n-1}(l_i). \tag{40.101}$$

Pour le plaisir d'ajouter des notations, nous posons $I_n =]a_1, b_1] \times \ldots \times]a_n, b_n]$. Cela permet d'écrire

$$A_n = \{(T_1, \ldots, T_n) \in I_n\} \tag{40.102}$$

et nous incite à calculer $\int_{I_n} g$. Pour alléger le calcul suivant, notez que

$$\int_a^b e^{-\lambda t} dt = \left[-\frac{1}{\lambda} e^{-\lambda t}\right]_a^b = \frac{1}{\lambda} e^{-\lambda a}\big(1 - e^{-\lambda(b-a)}\big). \tag{40.103}$$

EQooFXPWooRDBkUE

Ensuite nous calculons pour de vrai :

$$\begin{aligned}
\int_{I_n} g(x_1, \ldots, x_n) &= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} \mathbb{1}_{x_1 \leqslant \ldots \leqslant x_n} \lambda^n e^{-\lambda x_n} \, dx_1 \ldots dx_n && (40.104a)\\
&= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} \lambda^n e^{-\lambda x_n} \, dx_1 \ldots dx_n && \text{cf. justif.} \quad (40.104b)\\
&= \Big(\prod_{i=1}^{n-1} l_i\Big) \int_{a_n}^{b_n} \lambda^n e^{-\lambda x_n} dx_n && (40.104c)\\
&= \Big(\prod_{i=1}^{n-1} l_i\Big) \lambda^n \frac{1}{\lambda} e^{-\lambda a_n}\big(1 - e^{-\lambda(b_a - a_n)}\big) && \text{eq. (40.103)} \quad (40.104d)\\
&= \Big(\prod_{i=1}^{n-1} l_i\Big) \lambda^{n-1} e^{-\lambda a_n}\big(1 - e^{-\lambda l_n}\big) && (40.104e)\\
&= P(A_n) && (40.104f)\\
&= P\big((T_1, \ldots, T_n) \in I_n\big). && (40.104g)
\end{aligned}$$

SUBEQooPMLZooJhEjds

Justifications :

— Pour (40.104b). Vu que nous avons choisi $a_1 < b_1 < \ldots < a_n < b_n$, nous avons toujours $\mathbb{1}_{0 \leqslant x_1 \leqslant \ldots \leqslant x_n} = 1$. La fonction à intégrer ne dépend pas de $x_1, \ldots, x_{n-1}$. Les intégrales sur $x_1, \ldots, x_{n-1}$ sont immédiates, et nous avons

Nous savons que les T_i sont à valeurs dans $\mathbb{R}^+$. Donc sur les pavés N contenant des coordonnées négatives, nous avons $P\big((T_1, \ldots, T_n) \in N\big) = 0$. Heureusement nous avons aussi $\int_N g(x)dx = 0$.

En prenant tous les pavés I_n ainsi que tous ceux avec des coordonnées négatives, nous engendrons les boréliens. Le lemme 36.38 conclut que la fonction g proposée est une densité pour $(T_1, \ldots, T_n)$.

Pour (3)

Pour (4)

Pour (5)

(1) **Densité pour $(S_1, \ldots, S_n)$**

Le point (4) nous indique que $(S_1, \ldots, S_n)$ est à densité. Nous notons $h : \mathbb{R}^n \to \mathbb{R}$ sa densité. Soit une application borélienne $f : \mathbb{R}^n \to \mathbb{R}$. En posant

$$\begin{aligned}
m : \mathbb{R}^n &\to \mathbb{R}^n\\
(t_1, \ldots, t_n) &\mapsto (t_1, t_2 - t_1, \ldots, t_n - t_{n-1}),
\end{aligned} \tag{40.105}$$

nous avons

$$\begin{aligned}
E\big(f(S_1,\dots,S_n)\big) &= E\big(f(T_1,T_2-T_1,\dots,T_n-T_{n-1})\big) && \text{(40.106a)}\\
&= E\big((f\circ m)(T_1,\dots,T_n)\big) && \text{(40.106b)}\\
&= \int_{\mathbb{R}^n} (f\circ m)(x)g(x)dx && \text{(40.106c)}\\
&= \int_{\mathbb{R}^n} (f\circ m\circ\phi)(x)(g\circ\phi)(x)dx && \text{(40.106d)}\\
&= \int_{\mathbb{R}^n} f(x)(g\circ\phi)(x)dx && \text{(40.106e)}\\
& && \text{(40.106f)}
\end{aligned}$$

Justifications.

— Pour (40.106c), c'est le lemme 36.98(2).

— Pour (40.106d), nous effectuons le changement de variables [16] pour la bijection linéaire

$$\begin{aligned}
\phi\colon \mathbb{R}^n &\to \mathbb{R}^n\\
(t_1,\dots,t_n) &\mapsto \begin{pmatrix} t_1\\ t_1+t_2\\ \vdots\\ t_1+\dots+t_n \end{pmatrix}.
\end{aligned} \tag{40.107}$$

Notons que ϕ est de déterminant 1 : c'est une matrice triangulaire avec des 1 sur la diagonale.

— Pour (40.106e). Parce que $(m\circ\phi)(x)=x$ dans $\mathbb{R}^n$.

Bref, pour toute application borélienne $f\colon\mathbb{R}^n\to\mathbb{R}$, nous avons

$$E\big(f(S_1,\dots,S_n)\big) = \int_{\mathbb{R}^n} f(x)(g\circ\phi)(x)dx. \tag{40.108}$$

Le lemme 36.100 nous indique que $g\circ\phi$ est une densité pour $(S_1,\dots,S_n)$. Plus explicitement, la densité de $(S_1,\dots,S_n)$ est

$$\begin{aligned}
h\colon \mathbb{R}^n &\to \mathbb{R}\\
x &\mapsto \mathbb{1}_{\mathbb{R}^n_+}(x)\lambda^n e^{-\lambda(x_1+\dots+x_n)}.
\end{aligned} \tag{40.109}$$

(2) **Densités marginales** La proposition 36.101 nous assure que les variables aléatoires S_i sont à densité et que pour trouver les densités, il suffit d'intégrer. La densité de S_i est

$$h_i(t) = \int_{\mathbb{R}^{n-1}} \mathbb{1}_{\mathbb{R}^n_+}(x_1,\dots,t,\dots,x_n)e^{-\lambda x_1-\dots-\lambda t-\dots-\lambda x_n} \tag{40.110}$$

Les différentes intégrales sont indépendantes et valent

$$\begin{aligned}
\int_{\mathbb{R}} \mathbb{1}_{[0,\infty[}(u)e^{-\lambda u}du &= \left[-\frac{1}{\lambda}e^{-\lambda u}\right]_0^\infty && \text{(40.111a)}\\
&= -\frac{1}{\lambda}(0-1) && \text{(40.111b)}\\
&= \frac{1}{\lambda}. && \text{(40.111c)}
\end{aligned}$$

Au final, nous avons

$$h_i(u) = \mathbb{1}_{[0,\infty[}(u)\lambda e^{-\lambda u}. \tag{40.112}$$

16. Formule (14.856).

(3) **Conclusions** Nous voyons que (40.112) est la densité d'une loi exponentielle (définition 36.145). La proposition 36.102(2) nous dit que les S_i sont indépendantes.

Pour (6)

C'est la proposition 36.148.

□

Proposition 40.20 (Loi des grands nombres)**.**
Si (N_t) est un processus de o-Poisson d'intensité λ, alors

$$\lim_{t\to\infty} N_t = +\infty \tag{40.113}$$

presque surement. De plus

$$\lim_{t\to\infty} \frac{N_t}{t} = \lambda \tag{40.114}$$

presque surement.

.

Démonstration. Par définition nous savons que

$$N_t = \sup\{n \geqslant 0 \text{ tel que } S_n \leqslant t\}. \tag{40.115}$$

Évidemment la fonction $t \mapsto N_t$ est croissante, donc la limite

$$\lim_{t\to\infty} N_t(\omega) \tag{40.116}$$

existe dans $[0,\infty]$. Nous pouvons nous restreindre à $t \in \mathbb{N}$ et considérer $L(\omega) = \lim_{n\to\infty} N_n(\omega)$. Par somme télescopique avec $N_0 = 0$,

$$\frac{N_n}{n} = \frac{\sum_{k=1}^{n}(N_k - N_{k-1})}{n}. \tag{40.117}$$

Étant donné que le processus est de Poisson, les variables aléatoires $(N_k - N_{k-1})_{k=1,\dots,n}$ sont indépendantes et suivent toutes la loi de $N_1 - N_0$, c'est-à-dire la loi de N_1. Encore par le fait que N_t soit de Poisson nous savons que $N_1 \sim \mathscr{P}(\lambda)$. La loi des grands nombres (36.128) appliquée aux variables aléatoires $N_k - N_{k-1}$ nous dit que

$$\frac{N_n}{n} \xrightarrow{p.s.} E(N_1) = \lambda > 0. \tag{40.118}$$

Du coup $N_n \to \infty$ et $L(\omega) = \infty$.

Nous démontrons maintenant la loi des grands nombres pour les processus de Poisson. Étant donné que pour les entiers $N_n/n \to \lambda$, pour les réels, si la limite existe, ça ne peut pas être autre chose. Si nous notons $\bar{t}$ la partie entière de $t \in \mathbb{R}^+$,

$$\frac{N_t}{t} = \frac{N_t - N_{\bar{t}}}{t} + \frac{N_{\bar{t}}}{t}. \tag{40.119}$$

Le second terme est relativement simple à traiter :

$$\frac{N_{\bar{t}}}{t} = \underbrace{\frac{N_{\bar{t}}}{\bar{t}}}_{\to\lambda} \cdot \underbrace{\frac{\bar{t}}{t}}_{\to 1}. \tag{40.120}$$

où nous avons utilisé le premier point, $\bar{t}$ étant entier. Pour le premier terme nous savons que $t \mapsto N_t$ est croissante et donc que

$$\frac{N_t - N_{\bar{t}}}{t} \leqslant \frac{N_{\bar{t}+1} - N_{\bar{t}}}{t} = \frac{N_{\bar{t}+1} - N_{\bar{t}}}{\bar{t}+1}\frac{\bar{t}+1}{t}. \tag{40.121}$$

Le second facteur tend vers 1 lorsque $t \to \infty$. Le premier s'écrit

$$\frac{N_n - N_{n-1}}{n} \tag{40.122}$$

et tend vers zéro en tant que terme général de la série (40.117) qui converge. □

40.1.5 Équivalence de définitions

40.1.6 Estimateurs

Proposition 40.21.
La variable aléatoire N_t/t est un estimateur sans biais de λ. De plus il converge vers λ en moyenne quadratique.

Démonstration. Vu que $N_t/t \to \lambda$ presque surement, la variable aléatoire N_t/t est un estimateur de λ. Le fait qu'il soit sans biais a été fait dans l'exemple 37.25.

D'autre part nous avons (voir théorème 36.144)

$$\operatorname{Var}\left(\frac{N_t}{t}\right) = \frac{1}{t^2}\operatorname{Var}(N_t) = \frac{\lambda}{t}. \tag{40.123}$$

En appliquant la formule $\operatorname{Var}(X) = E(X^2) - E(X)^2$ à $X = N_t/t$ nous trouvons

$$E\left(\frac{N_t^2}{t^2}\right) = \frac{\lambda}{t} + \lambda^2. \tag{40.124}$$

Cela montre que $\frac{N_t}{t} \xrightarrow{L^2} \lambda$. □

Pour le théorème central limite d'un processus de Poisson, nous visons un résultat du style de

$$\frac{\frac{1}{n}\sum_i X_i - mn}{\sigma\sqrt{n}} \xrightarrow{\mathscr{L}} \mathscr{N}(0,1). \tag{40.125}$$

Nous écrivons le théorème central limite pour le nombre de sauts que le processus de Poisson a connu en un temps t. Le rôle de la moyenne empirique est joué par N_t. Nous considérons avoir fait *une seule expérience* qui a duré un temps t. Donc le rôle de n est joué par 1 (et non t comme on pourrait le croire). Pour le reste, le nombre de succès en un temps t d'une variable aléatoire exponentielle de paramètre λ est une variable aléatoire de Poisson de paramètre λt, en vertu de ce qui est raconté au point 36.5.8. C'est cela qui motive l'énoncé suivant.

ThoCSuLLo

Théorème 40.22 (Théorème central limite pour les processus de Poisson)**.**
Si $(N_t)_{t>0}$ est un processus de Poisson de paramètre λ, alors nous avons

$$\frac{N_t - \lambda t}{\sqrt{\lambda t}} \xrightarrow{\mathscr{L}} \mathscr{N}(0,1). \tag{40.126}$$

Remarque 40.23.
Avant de nous lancer dans la démonstration, remarquons que si nous nous limitons à $t \in \mathbb{N}$, alors nous avons

$$\frac{N_n - \lambda n}{\sqrt{\lambda n}} = \frac{\sum_{k=1}^{n}(N_k - N_{k-1}) - \lambda n}{\sqrt{\lambda n}} \tag{40.127}$$

or par définition nous avons les égalités de lois

$$N_k - N_{k-1} \sim N_1 \sim \mathscr{P}(\lambda), \tag{40.128}$$

donc

$$\frac{S_n - \lambda n}{\sqrt{\lambda n}} = \frac{\frac{1}{n}S_n - \lambda}{\frac{\sqrt{\lambda n}}{n}} = \frac{\frac{1}{n}S_n - \lambda}{\sqrt{\lambda}/\sqrt{n}}, \tag{40.129}$$

ce qui est exactement le théorème central limite pour une suite de lois de Poisson [17].

17. Au fait près que nous devrions encore montrer que S_n est de carré intégrable.

Démonstration. Nous écrivons $\bar{t}$ la partie entière de t et nous décomposons :

$$\frac{N_t-\lambda t}{\sqrt{\lambda t}}=\underbrace{\frac{N_t-N_{\bar{t}}}{\sqrt{\lambda t}}}_{A}+\underbrace{\frac{N_{\bar{t}}-\lambda\bar{t}}{\sqrt{\lambda t}}}_{B}+\underbrace{\frac{\lambda\bar{t}-\lambda t}{\sqrt{\lambda t}}}_{C}. \tag{40.130}$$

En ce qui concerne le terme B, nous avons

$$B=\sqrt{\frac{\bar{t}}{t}}\frac{N_{\bar{t}}-\lambda\bar{t}}{\sqrt{\lambda\bar{t}}}\to\mathcal{N}(0,1). \tag{40.131}$$

Notons que nous utilisons le fait que si $a_n\to 1$ (en tant que suite de nombres) et si $X_n\to\mathcal{N}(0,1)$ (limite en loi), alors $a_nX_n\to\mathcal{N}(0,1)$ en loi.

Le terme C est également facile parce que $\lambda\bar{t}-\lambda t$ est majoré en norme par λ. Du coup

$$-\frac{\lambda}{\sqrt{\lambda t}}\leqslant C\leqslant\frac{\lambda}{\sqrt{\lambda t}}. \tag{40.132}$$

Donc $\lim_{t\to\infty}C=0$.

Reste à travailler sur A. Vu que $t\mapsto N_t$ est croissante, la différence $N_t-N_{\bar{t}}$ est positive. Soit $\eta>0$, nous avons

$$P(|A|>\eta)=P(N_t-N_{\bar{t}}>\sqrt{\lambda t}\eta)\leqslant P\big(N_{\bar{t}+1}-N_{\bar{t}}\geqslant\sqrt{\lambda t}\eta\big)=P(N_1\geqslant\eta\sqrt{\lambda t}) \tag{40.133}$$

parce que nous savons que $N_{\bar{t}+1}-N_{\bar{t}}\sim N_1\sim\mathscr{P}(\lambda)$. En vertu des propriétés de la loi de Poisson,

$$\lim_{t\to\infty}P(N_1\geqslant\eta\sqrt{\lambda t})=0. \tag{40.134}$$

En effet si Z est une variable aléatoire de Poisson de paramètre λ nous avons

$$P(Z>l)=\sum_{k=l}^{\infty}P(Z=k)=e^{-\lambda}\sum_{k=l}^{\infty}\frac{\lambda^k}{k!}. \tag{40.135}$$

Nous reconnaissons la queue de série de e^{λ}, qui tend donc vers zéro lorsque $l\to\infty$. Nous avons donc prouvé que

$$\lim_{t\to\infty}P\big(|A|>\eta\big)=0, \tag{40.136}$$

c'est-à-dire la convergence en probabilité de A vers zéro.

Nous avons montré que

$$B+C\xrightarrow{\mathscr{L}}U\sim\mathcal{N}(0,1) \tag{40.137a}$$

$$A\xrightarrow{P}0. \tag{40.137b}$$

Le lemme de Slutsky (36.119) nous avons une convergence du couple

$$(A,B+C)\xrightarrow{\mathscr{L}}(0,U). \tag{40.138}$$

Utilisant le corolaire 36.121, nous trouvons la convergence en loi

$$A+(B+C)\xrightarrow{\mathscr{L}}0+U, \tag{40.139}$$

ce qu'il fallait. □

PROPooWAVPooHDVsER

Proposition 40.24.
Si N_t est un processus de Poisson[18]*, alors les variables aléatoires $X_n=N_n$ avec $n\in\mathbb{N}$ forment une chaine de Markov*[19]*.*

J'espère qu'un jour une preuve arrivera ici : https://math.stackexchange.com/q/4563414

18. Définition 40.5.
19. Définition 38.1.

40.2 Quelques trucs sur la simulation

Le théorème ergodique dit que

$$\pi(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{X_k=x}. \tag{40.140}$$

C'est avec cela qu'on calcule $\pi(x)$ à partir d'une simulation de chaine de Markov.

40.2.1 Le théorème central limite pour Markov

THOooBCKQooCkoUEV

Théorème 40.25 (Version allégée)**.**
Si (X_n) est irréductible et positive récurrente, alors pour toute fonction f,

$$\frac{1}{\sqrt{N}} \left[\sum_{k=1}^{N} -N \int f d\pi \right] \xrightarrow{\mathscr{L}} \mathscr{N}(0, \sigma^2) \tag{40.141}$$

où σ^2 dépend de la fonction f et de la chaine de Markov.
Ici, $\int f d\pi = \sum_{x\in E} f(x)\pi(x)$.

Nous allons simuler la variable aléatoire

$$Z = \frac{1}{\sqrt{N}} \left[\sum f(X_k) - N \sum_{x\in E} f(x)\pi(x) \right] \tag{40.142}$$

et puis on va mettre sa réalisation dans un histogramme. Dans le cas où on prend $f(i) = \mathbb{1}_{i=i_0}$, il y a de la simplification dans l'intégrale qui devient

$$Z = \frac{1}{\sqrt{N}} \left[\sum_{i=1}^{N} \mathbb{1}_{X_k=i_0} - N\pi(i_0) \right]. \tag{40.143}$$

40.2.2 Feuille 5

On pose

$$D_n = \sqrt{n} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)|. \tag{40.144}$$

On en génère un millier de fois D_n, on note $D_n^{(k)}$ ces réalisations, et on regarde ce que vaut

$$\frac{1}{1000} \sum_{k=1}^{1000} \mathbb{1}_{D_n^{(k)}\geqslant c}. \tag{40.145}$$

Cela nous donne une approximation de

$$P\big(\sqrt{n} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \geqslant c\big). \tag{40.146}$$

Note que chacun des $D_n^{(k)}$ demande de créer un nouveau vecteur Y_i de lois qu'on veut regarder. Par exemple de loi exponentielle.

40.2.3 Feuille 6

Pour créer une fonction qui renvoie i avec probabilité p_i pour $i = 1, 2, 3$, on peut faire

$$U \sim \mathscr{U}[0, 1] \tag{40.147}$$

et puis on a

$$P(U < p_0) = p_0 \tag{40.148a}$$
$$P(p_0 < U < p_0 + p_1) = p_1 \tag{40.148b}$$
$$P(p_0 + p_1 < U < p_2) = p_2. \tag{40.148c}$$

Une façon de faire une loi uniforme $[0, 1]$ est de faire `rand`

40.2.4 Feuille 7

L'échantillon est $(Y_1, \dots, Y_n)$ et nous écrivons le vecteur

$$Y = X\beta + \epsilon \tag{40.149}$$

où $Y \sim \mathcal{N}(X\beta, \sigma^2 \mathrm{Id})$ et $\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$. Nous utilisons le principe de maximum de vraisemblance. Soit $(y_1, \dots, y_n)$ un échantillon et

$$P_\theta(y_1, \dots, y_n) = \prod_i \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y_i - X_i^t\beta}{\sigma}\right)^2\right]. \tag{40.150}$$

L'astuce est de faire que $y_i - X_i^t\beta$ est la $i$ième composante du vecteur $Y - X\beta$ et donc la somme qui est dans l'exponentielle devient la norme de $Y - X\beta$:

$$f_\theta(y_1, \dots, y_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2}\|Y - X\beta\|^2\right]. \tag{40.151}$$

On passe au logarithme et on dérive par rapport à σ^2. Attention : la variable est σ^2, donc la dérivée de σ^2 est 1 et non 2σ. Bref, on trouve

$$\sigma^2 = \frac{1}{2n}\|U + X\beta\|. \tag{40.152}$$

40.2.5 Simuler des lois conditionnelles

Nous voulons générer des couples (X, Y) tels que Y prend les valeurs 0 ou 1 et tels que

$$\begin{cases} P(X|Y=0) \sim \mathscr{E}(\lambda_0) & \text{(40.153a)} \\ P(X|Y=1) \sim \mathscr{E}(\lambda_1). & \text{(40.153b)} \end{cases}$$

Le plus simple est de générer une liste

$$(X_1, 0) \qquad (X_4, 1) \tag{40.154a}$$
$$(X_2, 0) \qquad (X_5, 1) \tag{40.154b}$$
$$(X_3, 0) \qquad (X_6, 1) \tag{40.154c}$$

avec $X_1, X_2, X_3 \sim \mathscr{E}(\lambda_0)$ et $X_4, X_5, X_6 \sim \mathscr{E}(\lambda_1)$.

Avec cette méthode cependant la liste est triée et en plus on a autant de 1 que de 0. On peut faire un peu plus technologique pour corriger cela. Pour créer un couple, on commence par $Y \sim \mathscr{B}(p)$ et puis suivant que $Y = 0$ ou $y = 1$, on génère $X \sim \mathscr{E}(\lambda_0)$ ou $X \sim \mathscr{E}(\lambda_1)$.

Chapitre 41

Langages

41.1 Alphabets et mots

Définition 41.1.
*Un **alphabet** est un ensemble fini de symboles appelés **lettres**.*

On utilise aussi parfois le terme **vocabulaire** pour désigner un alphabet.

Définition 41.2.
*Un **mot** sur l'alphabet Σ est une suite finie et ordonnée, éventuellement vide, de lettres de Σ. Le **mot vide** est toujours noté ε.*

Définition 41.3.
*La **longueur d'un mot** w, noté $|w|$, est le nombre de lettres constituant le mot w. Le mot vide a une longueur de 0.*

Soit w un mot de longueur k, on peut désormais noter $w = w_1 \cdots w_k$, où chacun des $w_i, 1 \leqslant i \leqslant k$ représente une lettre de w. Par convention, si $k = 0$, alors le mot w est le mot vide.

Définition 41.4.
*Soient w un mot sur l'alphabet Σ et $a \in \Sigma$ une lettre, le **nombre d'occurrences** de la lettre a dans le mot w, noté $|w|_a$, est le nombre de fois où apparaît la lettre a dans le mot w, c'est-à-dire le cardinal de l'ensemble $\{i \mid w_i = a, 1 \leqslant i \leqslant |w|\}$.*

Définition 41.5.
*Soit Σ un alphabet, l'**ensemble des mots non-vides** sur l'alphabet Σ, noté Σ^+, est l'ensemble :*

$$\Sigma^+ = \{w = w_1 \dots w_n, n > 0\} \tag{41.1}$$

Définition 41.6.
*Soit Σ un alphabet, l'**ensemble des mots** sur l'alphabet Σ, noté Σ^*, est l'ensemble :*

$$\Sigma^* = \{w = w_1 \dots w_n, n \geqslant 0\} \tag{41.2}$$

Des deux définitions précédentes, on tire l'égalité suivante :

$$\Sigma^* = \Sigma^+ \cup \{\varepsilon\} \tag{41.3}$$

Définition 41.7.
*Soient Σ un alphabet et $x, y \in \Sigma^*$ deux mots sur l'alphabet Σ de longueur respective n et m, le **produit** w de x et y, noté $x \cdot y$ est défini par $w = x_1 \dots x_n y_1 \dots y_m$.*

Le produit est également appelé **concaténation**.

Proposition 41.8 (Longueur du produit de deux mots)**.**
La longueur du produit de deux mots x et y est la somme des longueurs des mots x et y.

$$|x \cdot y| = |x| + |y| \tag{41.4}$$

Proposition 41.9 (Monoïde $(\Sigma^*, \cdot, \varepsilon)$)**.**
L'ensemble Σ^* muni de l'opération produit d'élément neutre ε est un monoïde[1]*.*

Démonstration. Soient $x, y, z \in \Sigma^*$, avec les définitions précédentes, on peut vérifier facilement que :

— le produit est une loi interne : $x \cdot y \in \Sigma^*$;

— le produit est associatif : $x \cdot (y \cdot z) = (x \cdot y) \cdot z$;

— ε est l'élément neutre du produit : $x \cdot \varepsilon = \varepsilon \cdot x = x$.

□

Le produit n'est pas commutatif.

Définition 41.10.
*Soient Σ un alphabet et $w \in \Sigma^*$, la **puissance** n^ed'un mot w, notée w^n, est définie par :*

$$w^n = \begin{cases} \varepsilon & si\ n = 0 \\ w \cdot w^{n-1} & si\ n > 0 \end{cases}$$

41.2 Langages

Définition 41.11.
*Un **langage** sur un alphabet Σ est un sous-ensemble de Σ^*. C'est un ensemble de mots sur l'alphabet Σ.*

Un langage étant défini comme un ensemble, on peut appliquer toutes les notions de la théorie des ensembles aux langages.

Définition 41.12.
*Le **langage vide**, noté $\varnothing$ est le langage qui ne contient aucun mot.*

Définition 41.13.
*Le **langage unité** est le langage qui contient uniquement le mot vide : $\{\varepsilon\}$.*

Définition 41.14.
*Soient Σ un alphabet et $L_1, L_2 \subseteq \Sigma^*$ deux langages sur l'alphabet Σ, on définit le **produit** L de L_1 et L_2, noté $L_1.L_2$ par :*

$$L = L_1.L_2 = \{u_1 \cdot u_2, u_1 \in L_1, u_2 \in L_2\} \tag{41.5}$$

Le produit de langages est également appelé **concaténation**. Il ne faut pas confondre le produit de langage avec le produit cartésien de deux ensembles. Le langage unité est l'élément neutre du produit de langages.

Proposition 41.15 (Distributivité du produit de langage par rapport à l'union)**.**
Le produit de langage est distributif par rapport à l'union. Soient Σ un alphabet et $L_1, L_2, L_3 \subseteq \Sigma^*$, alors :

$$L_1.(L_2 \cup L_3) = (L_1.L_2) \cup (L_1.L_3) \text{ et } (L_1 \cup L_2).L_3 = (L_1.L_3) \cup (L_2.L_3) \tag{41.6}$$

Démonstration. Soit $w \in L_1.(L_2 \cup L_3)$, montrons que $w \in (L_1.L_2) \cup (L_1.L_3)$. $\exists w_1 \in L_1, w' \in L_2 \cup L_3, w = w_1 \cdot w'$. Donc $w' \in L_2$ ou $w' \in L_3$. Si $w' \in L_2$, alors $w = w_1 \cdot w' \in L_1.L_2$. Si $w' \in L_3$, alors $w = w_1 \cdot w' \in L_1.L_3$. Donc, $w \in (L_1.L_2) \cup (L_1.L_3)$. Donc $L_1.(L_2 \cup L_3) \subseteq (L_1.L_2) \cup (L_1.L_3)$.

1. Définition 6.23.

Soit $w \in (L_1.L_2) \cup (L_1.L_3)$, montrons que $w \in L_1.(L_2 \cup L_3)$. $w \in L_1.L_2$ ou $w \in L_1.L_3$. Si $w \in L_1.L_2, i$ avec $i \in \{2,3\}$ alors $\exists w_1 \in L_1, w_i \in L_i, w = w_1 \cdot w_i$. Donc $w \in L_1.(L_2 \cup L_3)$. Donc, $(L_1.L_2) \cup (L_1.L_3) \subseteq L_1.(L_2 \cup L_3)$

Donc $(L_1.L_2) \cup (L_1.L_3) = L_1.(L_2 \cup L_3)$

L'autre partie de la proposition se montre de manière analogue. □

Définition 41.16.
*Soient Σ un alphabet et $L \subseteq \Sigma^*$, la **puissance** n^edu langage L, notée L^n est définie par :*

$$L^n = \begin{cases} \{\varepsilon\} & si\ n = 0 \\ L.L^{n-1} & si\ n > 0 \end{cases}$$

Définition 41.17 ([805])**.**
*L'**étoile de Kleene** est un opérateur unaire noté $*$. L'**itéré** d'un langage L, noté L^*, est l'application de l'étoile de Kleene à un langage L et est défini par :*

$$L^* = \bigcup_{i \geqslant 0} L^i \tag{41.7}$$

En particulier, on remarque que le mot vide fait toujours partie de l'itéré d'un langage, y compris quand ce même langage ne contient pas le mot vide.

Définition 41.18.
*L'**itéré strict** d'un langage L, noté L^+, est défini par :*

$$L^+ = \bigcup_{i > 0} L^i \tag{41.8}$$

Proposition 41.19 (Relations entre itéré et itéré strict)**.**
Soit L un langage, alors on a :

$$L^* = L^+ \cup \{\varepsilon\} \tag{41.9}$$

$$L^+ = L.L^+ = L^+.L \tag{41.10}$$

Chapitre 42

Utilisation dans les autres sciences

Dans ce chapitre nous donnons des applications de divers théorèmes dans les autres sciences que la mathématique.

42.1 Démystification du MRUA

SecMRUAsecondeGGdQoT

42.1.1 Preuve de la formule

Nous sommes maintenant en mesure de donner une démonstration complète de la formule du MRUA :

$$x(t) = \frac{at^2}{2} + v_0 t + x_0. \tag{42.1}$$

EqMRUAINT

Au niveau de la physique, nous considérons un mobile qui se déplace avec une accélération constante a. Nous notons par v_0 sa vitesse initiale et par x_0 sa position initiale.

Nous savons que, pour tout mouvement, si $x(t)$ est la position en fonction du temps, et si $v(t)$ et $a(t)$ représentent la vitesse et l'accélération en fonction du temps, alors

$$v(t) = x'(t) \quad \text{et} \quad a(t) = v'(t) = x''(t). \tag{42.2}$$

Afin de trouver $x(t)$ en connaissant $a(t)$, il « suffit » donc de prendre deux fois la primitive. Essayons ça dans le cas facile du MRUA où $a(t) = a$ est constante.

La vitesse $v(t)$ doit être une primitive de la constante a. Il est facile de voir que $v(t) = at$ est une primitive de a. Par le corolaire 12.199(bis),

$$v(t) = at + C_1 \tag{42.3}$$

EqvtatC

pour une certaine constante C_1. Afin de fixer C_1, il faut faire appel à la physique : d'après la formule (42.3), la vitesse initiale est $v(0) = C_1$. Donc il faut identifier C_1 à la vitesse initiale : $C_1 = v_0$. Nous avons donc déjà obtenu que

$$v(t) = at + v_0. \tag{42.4}$$

Afin de trouver $x(t)$, il faut trouver une primitive de $v(t)$. Il n'est pas très difficile de voir que $at^2/2 + v_0 t$ fonctionne, donc il existe une constante C_2 telle que

$$x(t) = \frac{at^2}{2} + v_0 t + C_2. \tag{42.5}$$

Encore une fois, regardons la condition initiale : la formule donne comme position initiale $x(0) = C_2$, et donc nous devons identifier C_2 avec la position initiale x_0. En définitive, nous avons bien

$$x(t) = \frac{at^2}{2} + v_0 t + x_0. \tag{42.6}$$

Cette formule est donc maintenant *démontrée* à partir de la seule définition de la vitesse comme dérivée de la position et de l'accélération comme dérivée de la vitesse.

Remarquons cependant que la preuve complète fut *très* longue. En effet, nous avons utilisé les règles de dérivation de la proposition 12.172, pour la démonstration desquels, les résultats de la proposition 12.6 ont étés utiles. Mais nous avons surtout utilisé le corolaire 12.199(bis) qui repose sur le théorème de Rolle 12.189, qui lui-même demande le théorème de Borel-Lebesgue 9.19 dans lequel la notion d'ensemble compact a été cruciale.

42.1.2 Interprétation graphique

La distance parcourue $x(t)$ en un temps t est la primitive de la vitesse. Nous savons, par ailleurs, que la primitive est liée à la surface[1] sous la courbe. Pour reprendre les mêmes notations, nous notons $S_v(t)$ la surface contenue en dessous de la fonction v entre 0 et x. Nous ne serions donc pas étonné que

$$S_v(t) = \frac{at^2}{2} + v_0 t + x_0 \tag{42.7}$$ EqEncoreMRUASvt

soit la surface en dessous de la fonction $v(t) = at + v_0$. Nous voyons que la surface totale sous la fonction $v(t) = at + v_0$ est exactement

$$S_v(t) = \frac{at^2}{2} + v_0 t. \tag{42.8}$$

Cela est un bon début, mais hélas nous ne retrouvons pas le terme « $+x_0$ » de la formule (42.7). Cela n'est pas tout à fait étonnant parce que nous savons que la surface sous une fonction était *une* primitive de la fonction, mais nous n'avons pas dit *laquelle*. D'après le fameux corolaire 12.199(bis), la primitive n'est définie qu'à une constante près. Ici, c'est la constante x_0 qu'on a perdue en chemin.

Nous parlerons plus en détail du lien entre les surfaces et les primitives dans la section dédiée à l'intégration.

42.2 Relativité en mécanique newtonienne

42.2.1 Relativité du mouvement

Prenons quelqu'un qui court le cent mètres en onze secondes. Par rapport à un spectateur dans les gradins, il se sera déplacé de cent mètres. Mais si je cours à côté de lui de telle façon à avoir parcouru 80 mètres le temps qu'il en fasse cent, alors par rapport à moi l'athlète ne se sera déplacé que de 20 mètres. Par contre, par rapport à mon chronomètre, il aura également mis onze secondes : ce n'est pas parce que je cours que mon chronomètre s'affole !

Entre moi et les spectateurs, on a donc une loi de transformation

$$x' = x - vt \qquad\qquad t' = t. \tag{42.9}$$ EqTransGal

C'est-à-dire que la distance x' qu'aura parcouru l'athlète par rapport à moi vaut la distance x parcourue par le spectateur moins la distance que j'ai courue moi-même, c'est-à-dire moins vt.

42.2.2 Bob et Alice

Formalisons le concept de changement de repères. Pour cela, prenons deux amoureux, Bob et Alice[2]. Mettons que Bob reste assis sur un banc pendant qu'Alice court en ligne droite à une vitesse v. Tout deux déclenchent leur chronomètre quand Alice passe devant Bob. À tout moment, Bob et Alice ont leur repères de temps et d'espace. Par exemple si après un temps t, Alice voit une peau de banane à 1 mètre devant elle, elle va dire « Il y a une peau de banane à un mètre. », tandis que Bob va dire « Il y a une peau de banane à $(1 + vt)$ mètres ».

Plus généralement, si il se passe quelque chose à la position x au temps t pour Bob, ce quelque chose se passera au temps $t' = t$ à l'endroit $x' = x - vt$ pour Alice parce qu'en un temps t, elle aura déjà avancé d'une distance vt.

Ça c'est ce dont tout le monde était persuadé depuis Galilée jusqu'au début du vingtième siècle.

1. L'intégrale d'une fonction entre deux valeurs de t donne l'aire sous la courbe.
2. C'est plus poétique que dire « soient A et B deux observateurs ».

42.3 Invariance de la vitesse de la lumière

42.3.1 Champ de gravitation et électrique

Nous savons que la force de gravitation s'écrit :

$$F_{grav} = G\frac{mm'}{r^2},$$

tandis que la force électrique entre deux charges q et q' est donnée par

$$F_{elec} = k\frac{qq'}{r^2}. \tag{42.10}$$ EqRappelFelec

Nous avons aussi fait remarquer que dans le cas de la gravitation, la force a l'air d'être instantanée, et que cela posait quelques problèmes conceptuels. La force électrique a apparemment le même problème. Une différence entre les deux est qu'une charge électrique c'est tout petit et qu'on peut expérimenter à souhait, tandis que pour avoir une masse dont on peut mesurer le champ de gravitation correctement, il faut quelque chose grand comme la Terre [3].

42.3.1.1 Finitude de la vitesse de propagation de la force électrique

Si un micro est placé juste à côté de ton oreille, et qu'il commence à faire biiiiip, tu l'entends directement. Quand il s'arrête, tu ne l'entends plus. Si le micro est placé à 600 m de toi, tu ne commenceras à l'entendre que deux secondes après le commencement du son, et tu continueras à l'entendre deux secondes après qu'il a fini.

Eh bien, pour la force électrique, on a pu mesurer que c'est la même chose (sauf que ça va beaucoup plus vite). Si on place une charge quelque part, on ne ressent la force (42.10) qu'après qu'elle a eut le temps d'arriver. Si on déplace la charge électrique, on continue à ressentir la même force pendant un certain temps : il faut que la modification du champ électrique ait le temps d'arriver. Exactement comme quand on fait des remous quelque part dans un étang : il faut du temps que les remous arrivent plus loin.

On a pu faire des dizaines d'expériences de ce type avec l'électricité, le magnétisme et la lumière ; et les résultats sont clairs : il faut du temps pour que ça se déplace. Tout cela provoque des ondes électromagnétiques qui se déplacent à une vitesse finie. On peut produire de telles ondes avec n'importe quel courant électrique alternatif.

42.3.1.2 Pourquoi pas la gravitation ?

La gravitation telle que donnée par Newton pose le même problème de vitesse de propagation que l'électricité. Est-ce qu'en réalité la gravitation se propage également à une vitesse finie ?

Avec la gravitation c'est beaucoup plus compliqué parce qu'elle est beaucoup plus faible, et donc c'est beaucoup plus difficile à détecter. D'après la théorie d'Einstein de la gravitation, la gravitation devrait également produire des ondes gravitationnelles. Seulement, si un simple courant électrique suffit pour mesurer une onde électromagnétique, afin de mesurer une onde gravitationnelle, il faudrait un déplacement de masse de l'ampleur d'une étoile qui explose. Or ça, on ne sait pas produire dans un laboratoire.

Des ondes gravitationnelles ont été observées pour la première fois en 2016[806].

42.3.2 Support du champ : pas d'éther

Nous avons dit qu'une onde électromagnétique se propage comme une onde sonore (quoique beaucoup plus vite). Une question se pose alors. En effet, une onde sonore est matérialisée par de l'air qui vibre. Qu'est-ce qui vibre pour une onde électromagnétique ?

3. Une autre différence fondamentale est qu'il existe des charges électriques négatives, mais pas de masses négatives ; de ce fait on ne peut pas construire d'isolant gravitationnel, contrairement aux isolants électriques qui existent. Cela augmente encore la difficulté de faire des expériences avec la gravitation.

Étant donné que les ondes électromagnétiques se propagent dans le vide (c'est pour ça que la radio fonctionne dans l'espace), la question est problématique. Les physiciens ont donc supposé que tout l'univers était rempli d'un fluide invisible appelé **l'éther**. L'électromagnétisme consiste en une vibration de l'éther, exactement comme l'acoustique consiste en une vibration de l'air.

En fait, vérifier cette hypothèse n'est pas très compliqué. En effet il n'y a aucune raison que l'éther suive la Terre dans son mouvement. Or la Terre se déplace à environ 30 km/s autour du Soleil. Donc les ondes électromagnétiques doivent se propager plus vite dans le sens du mouvement de la Terre que dans le sens perpendiculaire. Tout comme le son se propage plus vite dans le sens du vent.

La célèbre expérience d'interférométrie de Michelson-Morley[807] a mesuré cet effet ...et ce fut la consternation : il n'y a aucun effet ! Or, la lumière se déplace à 300.000 km/s [4] ; une variation de 30 km/s devrait être détectable !

Mais rien ! On a recommencé les expériences dans tous les sens, à tous les mois de l'année, à tous les endroits de la Terre. On n'a pas observé un poil de variation de la vitesse de la lumière. Et ça, ça pose un gros problème à la physique.

42.3.3 Le problème

Si je joue au football dans un train qui avance à 100 km/h et que je lance une balle à 20 km/h, quelqu'un au sol mesura la vitesse de la balle soit à 120 km/h soit à 80 km/h d'après que l'on ait shooté vers l'avant ou l'arrière du train. Cela paraît logique. Mais ce qu'on vient de voir c'est que ça ne marche pas avec la lumière.

Si un train avance à 100.000 km/s et qu'on y allume une lampe de poche, la lumière avancera à 300.000 km/s par rapport au train et 400.000 km/s par rapport au sol. Non ! Justement pas ! La lumière avancera quand même à 300.000 km/s par rapport au sol.

Là encore, on a fait des dizaines d'expériences partout, sur Terre, dans des avions, dans l'espace avec des atomes, des lampes de poche et des horloges atomiques, et dans tous les sens, le sens de déplacement de la Terre, le sens inverse, le sens perpendiculaire, vers le haut, vers le bas : rien ! Personne n'a jamais observé un rayon de lumière se déplacer à une autre vitesse que 300.000 km/s.

Le problème est que le principe d'addition des vitesses est faux pour la lumière. Puisque l'expérience nous force, nous devons faire avec.

LoiVitLum

Loi numéro 1.
La réalité est que la vitesse de la lumière est la même dans tous les référentiels. On note c cette vitesse. C'est une constante fondamentale de la Nature.

Étant donné que c'est une loi expérimentale, nous n'y pouvons rien. C'est la nature qui est comme ça. En particulier tu ne peux pas en vouloir à ton prof de physique d'avoir inventé une théorie compliquée. Ce n'est pas lui qui l'a inventée et ce n'est pas de sa faute.

42.4 Conséquences

C'est maintenant que les choses vraiment graves commencent (cela soit dit sans vouloir te faire peur). Afin d'un peu simplifier les choses, nous n'allons étudier que les mouvements en une dimension, c'est-à-dire sur une droite.

42.4.1 Ligne d'univers

Un événement a une coordonnée (t, x). Si je pose un objet juste à mes pieds (disons en $x = 0$), ses coordonnées seront à tout moment $(t, 0)$. Il est bon de voir cette coordonnée comme l'équation paramétrique d'une droite horizontale dans le plan des coordonnées t et x. Plus généralement quand un mobile effectue un mouvement $x(t)$, on appelle la **ligne d'univers** du mobile la ligne (pas forcément droite) $(t, x(t))$. Dans le premier exemple, on avait $x(t) = 0$ pour tout t.

4. 299.792.458 m/s très exactement

Le cas d'un mobile se déplaçant à vitesse constante v donne comme ligne d'univers la droite[5] $(t, x_0 + vt)$, et un objet qui se déplace selon un MRUA a comme ligne d'univers

$$\big(t, x_0 + v_0 t + \frac{at^2}{2}\big).$$

42.4.2 Transformations de Lorentz

Reprenons les amours scientifiques de Bob et Alice, mais cette-fois, analysons celles-ci en tenant compte du fait que la vitesse de la lumière soit invariante. Maintenant, si Bob voit se passer quelque chose au temps t à l'endroit x, on va dire qu'Alice voit cette chose au temps t' à la position x', et on va chercher (t', x') en fonction de (t, x).

Posé en termes mathématiques, le problème s'énonce ainsi : trouver les fonctions f et g telles que les formules SubEqLorGen

$$t' = f(t, x) \tag{42.11a}$$
$$x' = g(t, x) \tag{42.11b}$$

donnent les coordonnées vues par Alice pour un événement vu par Bob à l'instant t au point x. Une première étape importante est franchie par la proposition suivante[6].

Proposition 42.1.
Les fonctions f et g contenues dans les transformations (42.11) sont nécessairement linéaires (affines), c'est-à-dire qu'elles doivent s'écrire sous la forme

$$t' = \alpha t + \beta x + p$$
$$x' = \gamma t + \delta x + q$$

pour certaines fonctions α, β, γ, δ, p et q de la vitesse d'Alice relativement à Bob.

Démonstration. Pendant qu'Alice court et que Bob la regarde, Ève tente de lancer une pierre sur Alice (Ève est jalouse). Bob et Alice regardent deux événements. Le premier est la pierre qui quitte la main de Ève, et le second est la pierre qui percute le sol. Pour Bob, le jet s'est passée au temps t_0 au point x_0, et la pierre touche le sol un petit peu plus tard, au temps $t_0 + \Delta t$ et un peu plus loin, au point $x_0 + \Delta x$. Bob écrit donc ceci sur sa feuille de papier :

$$E_1 = (t_0, x_0)$$
$$E_2 = (t_0 + \Delta t, x_0 + \Delta x),$$

tandis qu'Alice, en observant les mêmes deux événements, aura noté

$$E'_1 = \big(f(t_0, x_0), g(t_0, x_0)\big)$$
$$E'_2 = \big(f(t_0 + \Delta t, x_0 + \Delta x), g(t_0 + \Delta t, x_0 + \Delta x)\big).$$

Bob et Alice se demandent combien de temps la pierre est restée en l'air et quelle distance elle a parcourue. Par le principe général d'homogénéité, les deux seules quantités pertinentes (qui ont un sens physique) pour Bob sont $(t_0 + \Delta t) - t_0$ et $(x_0 + \Delta x) - x_0$, c'est-à-dire Δt et Δx. En effet, si Bob avait choisi de s'asseoir autre part et si Alice avait commencé à courir un peu plus tard, ça n'aurait rien changé à la longueur du jet de Ève.

D'une façon ou d'une autre, il doit exister une façon de déduire les mesures de Alice en connaissant celles de Bob ; je ne connais pas avec quelles formules, mais ces formules ne peuvent contenir que Δt, Δx et v parce que ce sont les seules quantités qui définissent tous les événements.

Cela dit, Alice va caractériser le mouvement de la pierre avec la différence des coordonnées entre le jet et la chute sur le sol mesurées par elle-même. En d'autres termes, pour Alice ce qui compte c'est la différence entre E'_1 et E'_2, soit

$$\big(f(t_0 + \Delta t, x_0 + \Delta x), g(t_0 + \Delta t, x_0 + \Delta x)\big) - \big(f(t_0, x_0), g(t_0, x_0)\big). \tag{42.12}$$ EqAlfxdelmoins

5. bon exercice de révision de ton cours de math de vérifier que c'est une droite.
6. dont je te suggère fortement de ne pas lire la preuve si tu ne veux pas que ton cerveau éclate.

Mais nous venons de signaler que ce qu'Alice mesurait devait pouvoir être exprimé en termes de Δt et Δx. Nous concluons que la différence (42.12) ne dépend en fait pas de x et t mais seulement de Δx et Δt.

Prenons maintenant une notation plus compacte et notons $X = (t, x)$, $\Delta X = (\Delta t, \Delta x)$ puis $F = (f, g)$. Avec ça, l'expression (42.12) se note $F(X + \Delta X) - F(X)$. Comme mentionné, cette expression ne dépend que de Δx. En particulier, elle ne dépend pas de X.

Maintenant tu vas comprendre pourquoi on apprend les dérivées dans ton cours de math. Comme $F(X+\Delta X)-F(X)$ ne dépend pas de X, le rapport $\big(F(X+\Delta X)-F(X)\big)/\Delta X$ non plus. La limite de ce rapport quand ΔX tend vers zéro non plus :

$$\lim_{\Delta X \to 0} \frac{F(X + \Delta X) - F(X)}{\Delta X} \tag{42.13}$$

ne dépend pas de X. Tu reconnais là la dérivée de F au point X. En d'autres termes, $F'(X)$ est constante, elle ne dépend pas de X. Disons donc que $F'(X) = a$. Tu connais beaucoup de fonctions dont la dérivée est constante ? Non ? En effet, il n'y en a pas beaucoup. Les fonctions qui vérifient $F'(X) = a$ sont toutes de la forme

$$F(X) = aX + b.$$

À ce niveau, il convient de re-déballer les notations compactes : si $a = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ et $b = (p, q)$ on trouve [EqLoUn]

$$f(t, x) = \alpha t + \beta x + p \tag{42.14a}$$
$$g(t, x) = \gamma t + \delta x + q, \tag{42.14b}$$

comme annoncé.

□

Nous savons que lorsque $(t, x) = (0, 0)$, alors $(t', x') = (0, 0)$. En effet, Bob et Alice ont lancé leurs chronos en même temps au moment où ils étaient au même endroit. En mettant $(t, x) = (0, 0)$ dans les équations (42.14), on trouve $(t', x') = (p, q)$, et donc $p = q = 0$. Ça fait une chose de réglée ; on se retrouve avec [EqLoDeux]

$$\begin{cases} t' = \alpha t + \beta x & \text{(42.15a)} \\ x' = \gamma t + \delta x. & \text{(42.15b)} \end{cases}$$

Quelles sont les contraintes à vérifier pour que ces transformations décrivent correctement la physique que l'on cherche à écrire ?

(1) Il faut que les transformations décrivent correctement que Alice avance à une vitesse v par rapport à Bob,

(2) dans le même ordre d'idée, il faut que l'on trouve que Bob avance à la vitesse $-v$ par rapport à Alice,

(3) il faut que si Alice et Bob observent un rayon lumineux, ce rayon aille à la vitesse c par rapport à Alice et à la *même* vitesse c par rapport à Bob,

(4) enfin, il faut avoir le principe de relativité, c'est-à-dire que comme les équations (42.15) disent ce que Alice voit en fonction de ce que Bob voit, on demande que les équations qui disent ce que Bob voit en fonction de ce que Alice voit soient les mêmes. En d'autres termes, il faut que les transformations et les transformations inverses soient les mêmes au changement près du signe de v.

Étudions une à une ce que chacune de ses contraintes impose. Rappelons que (t, x) et (t', x') sont les coordonnées que Bob et Alice mettent sur le même événement. Par exemple sur l'événement qui consiste à ce que Ève, par jalousie envers Bob, jette une peau de banane sous les pieds d'Alice. Cet événement a lieu à un certain moment, à un certain endroit. C'est ce moment et cet endroit qui sont notés (t, x) et (t', x').

(1) Les coordonnées (t,x) et (t',x') peuvent décrire n'importe quoi. Regardons les coordonnées de Alice qui cours. Pour Alice, cela correspond à $(t',x')=(t',0)$ parce que si x' désigne la position de Alice par rapport à Alice, alors x' est toujours nul. Pour Bob par contre, Alice ne reste pas en place, mais se déplace à une vitesse v. C'est-à-dire que si (t,x) sont les coordonnées de Alice pour Bob, alors $x/t=v$. Écrivons les équations (42.15) en tenant compte de tout ça : avec $x'=0$, la seconde équation donne

$$0=\gamma t+\delta x, \tag{42.16}$$

d'où on déduit que $x/t=-\gamma/\delta$. En imposant que cela soit v, on trouve $\gamma=-v\delta$, et on ré-écrit les transformations en tenant compte de ça :

$$\begin{cases} t'=\alpha t+\beta x & \text{(42.17a)}\\ x'=-v\delta t+\delta x. & \text{(42.17b)}\end{cases}$$

Nous voilà débarrassé d'un paramètre.

(2) Maintenant, on regarde ce qu'il se passe quand (t,x) et (t',x') décrivent les positions de Bob. On a $(t,x)=(t,0)$ parce que selon Bob, Bob est au repos. Les équations deviennent :

$$t'=\alpha t \qquad\qquad x'=-v\delta t. \tag{42.18}$$

La vitesse de Bob par rapport à Alice est $-v$, donc on exige que $x'/t'=-v$, c'est-à-dire que

$$\frac{-v\delta t}{\alpha t}=-v,$$

ce qui implique que $\delta=\alpha$. On avance encore un peu. Écrivons à nouveau les lois de transformation en en tenant compte :

$$\begin{cases} t'=\alpha t+\beta x & \text{(42.19a)}\\ x'=-v\alpha t+\alpha x. & \text{(42.19b)}\end{cases}$$

(3) Si maintenant Bob et Alice regardent un même rayon de lumière (comme c'est romanesque !), alors (t,x) et (t',x') expriment les coordonnées d'un rayon lumineux. Le fait que Bob regarde un rayon lumineux fait que $x=ct$, et donc que les coordonnées du rayon lumineux, observé par Alice sont :

$$t'=\alpha t+\beta ct \qquad\qquad x'=-\alpha vt+\alpha ct. \tag{42.20}$$

L'invariance de la vitesse de la lumière exige que Alice mesure une vitesse c pour le rayon de lumière, c'est-à-dire $x'=ct'$. On exige donc que

$$-\alpha vt+\alpha ct=c\alpha t+\beta c^2t,$$

ce qui implique que

$$\beta=-\frac{\alpha v}{c^2}.$$

Une fois de plus, l'avant-dernière, on ré-écrit les lois de transformations en tenant compte de ce fait ; mais cette fois, on fait l'effort d'écrire aussi les transformations inverses :

$$t'=\alpha t-\frac{\alpha v}{c^2}x \qquad\qquad t=\frac{1}{\Delta}\big(\alpha t'+\frac{\alpha v}{c^2}x'\big) \tag{42.21}$$

EqLorAvd

$$x'=-\alpha vt+\alpha x \qquad\qquad x=\frac{1}{\Delta}(\alpha vt'+\alpha x') \tag{42.22}$$

où $\Delta=\alpha^2-\frac{\alpha^2v^2}{c^2}$ que tu noteras au passage être toujours positif, et nul uniquement quand $v=c$.

(4) Maintenant il reste à imposer le principe de relativité. Les transformations (42.21) montrent comment Alice voit le monde (c'est-à-dire (t',x')) en fonction de la façon dont Bob voit le monde (c'est-à-dire (t,x)). On se demande donc quelle seraient, pour Bob, les coordonnées (t,x) d'un point vu en (t',x') par Alice. Cela signifie que l'on impose que les deux systèmes (42.21) soient en réalité les mêmes, à un changement de signe près.
Attention : il est à priori faux de dire qu'en changeant le signe de v dans $\alpha v/c^2$, j'obtiens $-\alpha v/c^2$ parce que α est une fonction de v. En réalité, il faut noter $\alpha(v)v/c^2$ et donc le changement de signe de v donne $-\alpha(-v)v/c^2$. Ceci étant clair, on peut un petit peu calculer.
Commençons par égaliser le coefficient de x dans t' à celui de x' dans t, en changeant le signe de v :

$$\frac{\alpha(-v)v}{c^2}=\frac{\alpha(v)v}{c^2},$$

et donc $\alpha(v)=\alpha(-v)$. Ça c'est une bonne nouvelle. Égalisons maintenant le coefficient de t dans t' à celui de t' dans t en changeant le signe de v :

$$\alpha(-v)=\frac{\alpha(v)}{\Delta(v)}=\frac{\alpha(v)}{\alpha(v)^2\big(1-\frac{v^2}{c^2}\big)}.$$

Comme $\alpha(-v)=\alpha(v)$, on en déduit que

$$\alpha(v)=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{42.23}$$ EqalphaLo

Maintenant qu'on a tout, on peut écrire les transformations de Lorentz. On met donc l'expression (42.23) dans les lois de transformations (42.21) :

$$\begin{aligned} t'&=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\Big(t-\frac{v}{c^2}x\Big) & t&=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\Big(t'+\frac{v}{c^2}x'\Big)\\ x'&=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x-vt) & x&=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(vt'+x'). \end{aligned} \tag{42.24}$$ EqTrLorentz

Tu remarqueras que $\Delta=1$; si tu ne sais pas ce qu'est le déterminant d'une application linéaire, ça n'a pas d'importance. Mais si tu sais ce qu'est le déterminant d'une application linéaire, alors ce $\Delta=1$ est crucial !

Afin d'avoir des équations un peu plus courtes, à partir de maintenant nous allons noter

$$\gamma(v)=\sqrt{1-\frac{v^2}{c^2}}.$$

42.4.3 Conditions d'existence

Comme tu vois une racine carrée et un dénominateur dans ces formules, tu dois te demander quelles sont les conditions d'existence. Étant donné que $v<c$, on a $v^2/c^2\leqslant 1$ et en particulier, $v^2/c^2=1$ si et seulement si $v=c$.

Ce qui se trouve dans la racine carrée ne pose donc jamais de problèmes parce que ce n'est jamais négatif.

Le dénominateur est par contre plus problématique : quand $v=c$ il n'y a plus rien qui fonctionne. Quelle est la physique de ce problème ? Pour le comprendre, il faut se souvenir ce que représente v. Nous avons dit que v est la vitesse à laquelle Alice court. Ce que la condition d'existence nous enseigne, c'est que personne ne peut courir à la vitesse de la lumière. C'est une vitesse que l'on ne peut pas atteindre.

Dit en termes plus savants, on ne peut pas choisir un repère qui se déplace à la vitesse de la lumière.

La question qui se pose alors est « ah bon, on ne peut pas atteindre la vitesse de la lumière ! Et la lumière, comment elle fait ? ». Bonne question, merci de l'avoir posée. Hélas la réponse sort du cadre de ce cours.

loivitlumlimite

Loi numéro 2.
Aucun objet ne peut atteindre la vitesse de la lumière.

Loi numéro 3.
Tu ne dois pas te demander pourquoi la lumière elle-même se déplace à la vitesse de la lumière malgré la loi numéro 2.

42.4.4 La notion d'intervalle

Un **événement** est quelque chose qui se passe à un endroit à un certain moment. C'est donc caractérisé par le moment et le lieu. Comme on travaille à une dimension, c'est un couple de réels (t, x).

Regardons un rayon de lumière. Un événement est le fait d'allumer une lampe de poche, et un autre est le fait que la lumière arrive sur l'objet qu'on éclaire. Appelons-les (t_1, x_1) et (t_2, x_2). Comme d'habitude, on note $\Delta t = t_2 - t_1$ et $\Delta x = x_2 - x_1$. Comme le rayon de lumière va à la vitesse c, on a $c = \Delta x/\Delta t$, ou encore

$$c^2\Delta t^2 - \Delta x^2 = 0.$$

Pour cette raison, on va dire que l'**intervalle** entre deux événements (t_1, x_1) et (t_2, x_2) vaut en général

$$s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2. \tag{42.25}$$

Par invariance de la vitesse de la lumière, si un intervalle est nul pour un observateur, il sera nul pour tous les observateurs.

42.4.4.1 En mécanique newtonienne

Afin de voir un peu mieux l'enjeu de l'invariance de l'intervalle, regardons un exemple chiffré. Si par exemple je me déplace de 10 m en 5 s, mon intervalle mesuré par une personne extérieure est

$$c^2\Delta t^2 - \Delta x^2 = (300.000.000)^2 \cdot (5)^2 - (10)^2 = 2,25 \cdot 10^{18}\,\text{m}.$$

Si je fais le calcul pour moi, j'ai que $\Delta x' = 0$ parce que je ne me déplace pas, et $\Delta t' = 5$ parce que je me suis déplacé en 5 secondes. Le truc est que à côté de $(300.000.000)^2$, l'intervalle spatial Δx ne pèse pas grand chose. Ça ne change presque rien qu'il soit de 5 mètres ou de zéro. Ça ne change pas grand chose, mais ça change quand même ! Entre moi qui calcule ou une personne extérieure, l'intervalle change de 100 sur un nombre de la grandeur de 200.000.000.000.000.000.0000 !

Reprenons plus clairement le raisonnement. D'après la mécanique classique, l'intervalle mesuré par deux personnes est différent, mais très peu différent. Inutile de dire que du temps de Newton, on n'avait pas les moyens techniques de mesurer si cet intervalle est effectivement différent ou bien si il est en réalité égal. C'est un peu comme si on te mettait un spot dans les yeux et qu'on te demandait si c'est un spot de 1000 W ou de 1001 W. Bonne chance pour le dire !

42.4.4.2 En mécanique relativiste

Maintenant qu'on a des moyens techniques nettement plus poussés que Newton, on a pu mesurer que l'intervalle est égal. L'intervalle est un invariant. Cela n'est pas un nouveau principe physique parce qu'il découle des transformations de Lorentz.

42.4.5 Le cône de lumière d'un point

Il est intéressant de dessiner dans le plan (t, x) l'ensemble des points atteints par le rayon lumineux. Le point (t, x) est atteint si $c^2t^2 - x^2 = 0$, ou encore si $x = \pm ct$. Cela forme deux droites dans le plan tracé par les coordonnées t et x. Ces deux droites forment ce qu'on appelle le **cône de lumière** du point $(0, 0)$.

42.4.6 Contraction des longueurs

Bob prend un morceau de bois qu'il mesure de longueur l et le dépose devant lui. À l'instant t (de Bob), les deux extrémités sont aux coordonnées $e_1 = (t, 0)$ et $e_2 = (t, l)$.

Afin de savoir quelle est la longueur de ce même morceau de bois pour Alice, il faut qu'elle mesure les deux extrémités en même temps (pour elle), et qu'elle fasse la différence. Comme Bob et Alice déclenchent leurs chronomètres en même temps, le plus simple est de faire la mesure à cet instant.

Pour Bob, c'est clair : les coordonnées des deux extrémités sont $e_1 = (0, 0)$ et $e_2 = (0, l)$. La longueur du bois est l. Pour savoir quelle est la longueur mesurée par Alice, on utilise les transformations de Lorentz qui donnent les coordonnées e'_1 et e'_2 relatives à Alice. On trouve $e'_1 = (0, 0)$ et

$$e'_2 = \left(\frac{-vl/c^2}{\gamma(v)}, \frac{l}{\gamma(v)} \right). \quad \text{Eqepdeuxfaut} \tag{42.26}$$

En d'autres termes, on a $x_1 = 0$ et $x_2 = l/\gamma(v)$, ce qui fait que la longueur observée par Alice est $l' = x_2 - x_1 = l/\gamma(v)$.

Eh bien ce résultat est faux. Si tu vois pourquoi sans lire la suite, tu es très fort.

Pour mesurer la longueur d'un objet, il faut mesurer la position des deux bouts *en même temps* puis faire la différence entre les deux. Effectivement, e_1 et e_2 sont en même temps pour Bob, et donc Bob peut mesurer la longueur de son bout de bois en faisant la différence $x_2 - x_1$. Mais comme le montre les coordonnées (42.26), les événements e'_1 et e'_2 ne se passent pas en même temps pour Alice ! Eh oui : $t'_1 = 0$ et $t'_2 = -vl/c^2\gamma(v)$; c'est pas la même chose.

Il faut donc trouver un événement qui pour Alice correspond à l'extrémité du bout de bois au temps $t' = 0$. Comme l'événement général qui correspond au bout du bois pour Bob est (t, l), l'événement général est pour Alice

$$t' = \frac{t - \frac{v}{c^2}l}{\gamma(v)} \qquad x' = \frac{l - vt}{\gamma(v)}. \tag{42.27}$$

Afin d'avoir $t' = 0$, il faut $t = vl/c^2$. En mettant cette valeur de t dans x', on trouve

$$x' = \frac{l - v\left(\frac{vl}{c^2}\right))}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{l\left(1 - \frac{v^2}{c^2}\right))}{\sqrt{1 - \frac{v^2}{c^2}}} = l\gamma(v).$$

Et là, c'est la bonne formule. Si un objet a une longueur l dans le référentiel où il est au repos, il aura une longueur

$$l' = l\sqrt{1 - \frac{v^2}{c^2}} \tag{42.28}$$

dans un référentiel qui se déplace à la vitesse v par rapport à l'objet.

42.4.7 Dilatation des intervalles de temps

Encore un petit effort et nous passons à une application concrète que tu connais des bizarreries de la relativité. Mais en attendant, regarde bien ta montre, tu ne va pas en croire tes yeux !

Reprenons Bob et Alice. On se rappelle que Bob et Alice avaient déclenché leurs chronomètres en même temps quand Alice était passée devant Bob. Un peu plus tard, Alice regarde sa montre qui indique un temps t. Et elle se demande si Bob a aussi à ce moment une montre qui indique un temps t.

Ce serait dingue que non hein ! ? ! En effet, si je synchronise ma montre avec quelqu'un et que je pars faire un tour, ma montre ne sera pas tout d'un coup désynchronisée. Oui, mais Alice, elle cours presque à la vitesse de la lumière ...et à ces vitesses-là, on a déjà vu des choses incroyables. Calculons donc pour en avoir le cœur net.

Le fait qu'Alice regarde sa montre est un événement qui se passe pour Alice aux coordonnées $(t', 0)$ (le zéro c'est parce que par rapport à elle-même, Alice est toujours au repos). À quelles coordonnées (t, x) pour Bob correspond cet événement ?

L'équation de t en fonction de t' et x' dans les transformations de Lorentz (42.24) prise avec $x' = 0$ donnent

$$t = \frac{t'}{\gamma(v)}.$$

Et si, juste pour le plaisir, on faisait l'inverse ? Bob regarde sa montre, il voir un temps t et sa coordonnée spatiale est $x = 0$. À quel temps d'Alice cela correspond ? Mettons $x = 0$ dans la transformation de Lorentz de t' en fonction de t et x. Ce qu'on obtient c'est

$$t' = \frac{t}{\gamma(v)}.$$

N'est-ce pas génial ? C'est la même ! Évidemment, ça ne pouvait pas être autre chose : le principe de relativité demande qu'on ne puisse pas faire la différence entre Alice qui cours vers la droite avec Bob assis et Alice assise avec Bob qui cours vers la gauche. C'est exactement pour ça que dans une gare, quand le train d'à côté démarre, il t'arrive de croire que c'est ton train qui démarre : tu ne peux pas faire la différence, c'est un principe physique.

42.4.8 Invariance de l'intervalle

Dans deux secondes, je vais te montrer comment une utilisation intelligente des exponentielles permet de trouver un résultat très fort en relativité. Quoi ? Les exponentielles, les mêmes qu'au cours de math ? Eh oui : la même exponentielle que celle qu'on t'a introduit avec des populations de bactéries qui se multiplient, cette même exponentielle qui la la miraculeuse propriété d'être égale à sa propre dérivée.

Mais n'anticipons pas.

Nous avons déjà signalé que si la quantité $\Delta s^2 = c^2\Delta t^2 - \Delta x^2$ était nulle pour un observateur, alors elle était nulle pour tous les observateurs. Supposons deux événements A et B observés par Alice et Bob. Bob les note aux coordonnées (t_a, x_a), et (t_b, x_b) tandis qu'Alice les note en (t'_a, x'_a) et (t'_b, x'_b).

L'intervalle entre les deux événements mesuré par Bob sera

$$s^2 = c^2(t_b - t_a)^2 - (x_b - x_a)^2,$$

tandis que ce même intervalle mesuré par Alice sera

$$s'^2 = c^2(t'_b - t'_a)^2 - (x'_b - x'_a)^2.$$

On peut bien entendu remplacer dans la première équation les t_a, t_b, x_a et x_b par leurs valeurs en termes de t'_a, t'_b, x'_a et x'_b données par les transformations de Lorentz. Tu paries que les trois quart des termes dans le calcul se simplifient et qu'il restera exactement s'^2 ? Je te dis que oui, et je te conseille de me croire sur parole, sinon tu vas devoir lire le calcul suivant :

$$\begin{aligned} s^2 = c^2(t_b - t_a)^2 - (x_b - x_a)^2 = c^2 &\left(\frac{1}{\gamma(v)}(t'_b + \frac{v}{c^2}x'_b) - \frac{1}{\gamma(v)}(t'_a + \frac{v}{c^2}x'_a)\right)^2 \\ &- \left(\frac{1}{\gamma(v)}(vt'_b + x'_b) - \frac{1}{\gamma(v)}(vt'_a + x'_a)\right)^2. \end{aligned}$$

Jusqu'ici, on n'a fait que remplacer les choses par leurs valeurs données par les transformations de Lorentz. Maintenant on regroupe à l'intérieur de chaque parenthèse les termes de façon à faire apparaitre $\Delta x'$ et $\Delta t'$:

$$\begin{aligned} s^2 &= \frac{c^2}{\gamma(v)^2}\left((t'_b - t'_a) + \frac{v}{c^2}(x'_b - x'_a)\right)^2 \\ &\quad - \frac{1}{\gamma(v)^2}\big((x'_b - x'_a) + v(t'_b - t'_a)\big)^2 \\ &= \frac{c^2}{\gamma(v)^2}\left((\Delta t')^2 + 2\frac{v}{c^2}\Delta t'\Delta x' + \frac{v^2}{c^4}(\Delta x')^2\right) \\ &\quad - \frac{1}{\gamma(v)^2}\Big((\Delta x')^2 + 2v\Delta x'\Delta t' + v^2(\Delta t')^2\Big). \end{aligned}$$

Là, on a utilisé le produit remarquable $(a+b)^2 = a^2+2ab+b^2$, et on a systématiquement renommé tous les intervalles avec la notation Δ pour être plus compact. Maintenant, on va regrouper tous les termes contentant $(\Delta t')^2$ ensemble, tous ceux qui contiennent $\Delta t'\Delta x'$ ensemble et ceux qui contiennent $(\Delta x')^2$ ensemble. Autre manière de le dire, on met les Δ en évidence comme on peut. On trouve ceci :

$$(\Delta t')^2\left(\frac{c^2}{\gamma(v)^2}-\frac{v^2}{\gamma(v)^2}\right)+\Delta t'\Delta x'\left(\frac{2vc^2}{\gamma(v)^2c^2}-\frac{2v}{\gamma(v)^2}\right)+(\Delta x')^2\left(\frac{c^2v^2}{c^4\gamma(v)^2}-\frac{1}{\gamma(v)^2}\right).$$

À partir de là, je te laisse vérifier (en utilisant le fait que $\gamma(v)^2 = 1-v^2/c^2$) que les coefficients se simplifient beaucoup et valent finalement respectivement c^2, 0 et -1 comme il se doit. Avec tout ça, nous avons montré le résultat très important suivant :

> L'intervalle entre deux événements est invariant sous les changements de repères d'inertie, c'est-à-dire que la valeur mesurée par n'importe qui qui se déplace en MRU sera toujours la même.

Pourquoi cela est tellement important ? À cause de Pythagore et d'une petite démonstration à coups d'exponentielles [7].

42.4.8.1 Rappel de trigonométrie hyperbolique

SUBSUBSECooZVHLooYwuhAj

Les fonctions de trigonométrie hyperboliques sont :

$$\cosh(x) = \frac{e^x+e^{-x}}{2} \qquad \sinh = \frac{e^x-e^{-x}}{2}. \tag{42.29}$$

Elles ont pas mal de propriétés en commun avec les sinus cosinus et normaux. D'abord, leurs dérivées sont faciles à calculer :

$$\cosh'(x) = \sinh(x)$$
$$\sinh'(x) = \cosh(x)$$

où tu noteras qu'il n'y a pas de signe moins qui apparaît, contrairement au cas de la trigonométrie normale. Une autre propriété qui ressemble fort à une propriété de la trigonométrie est :

Proposition 42.2.
Pour tout $x \in \mathbb{R}$,

$$\cosh^2(x) - \sinh^2(x) = 1 \tag{42.30}$$

avec un signe moins comme différence avec la trigonométrie.

Démonstration. La preuve revient simplement à calculer en utilisant le produit remarquable de $(a+b)^2$. D'abord, on a :

$$\cosh^2(x) = \frac{1}{4}\big(e^x+e^{-x}\big)^2 = \frac{1}{4}\big(e^{2x}+2e^xe^{-x}+e^{-2x}\big) = \frac{1}{4}(e^{2x}+2+e^{-2x})$$

où l'on a utilisé le fait que $(e^x)^2 = e^{2x}$ et que $e^xe^{-x} = 1$. Il te reste à faire la même chose pour $\sinh^2(x)$, la réponse est :

$$\sinh^2(x) = \frac{1}{4}\big(e^{2x}-2+e^{-2x}\big).$$

En faisant la différence entre les deux, il reste 1. □

Une propriété qui est par contre très différente entre la trigonométrie plane et la trigonométrie hyperbolique, c'est la périodicité. Les fonctions usuelles cos et sin sont périodiques. Pas les fonctions hyperboliques.

7. oui oui tout ton cours de math va finir par y passer.

Proposition 42.3.
La fonction $\sinh\colon \mathbb{R} \to \mathbb{R}$ *est bijective.*

Démonstration. Il faut démontrer que sinus hyperbolique est injective et surjective. Calculons d'abord les limites. Comme tu sais que $\lim_{x\to\infty} e^x = \infty$ et $\lim_{x\to-\infty} e^x = 0$, tu vois facilement que

$$\lim_{x\to-\infty} \sinh(x) = -\infty \qquad \lim_{x\to\infty} \sinh(x) = \infty. \tag{42.31}$$

Par ailleurs, la fonction sinus hyperbolique est continue et respecte donc le théorème de la valeur intermédiaire[8]. Soit $y \in \mathbb{R}$. Voyons si il existe un $x \in \mathbb{R}$ tel que $\sinh(x) = y$. Les deux limites indiquent qu'il existe $x_1 \in \mathbb{R}$ tel que $\sinh(x_1) < y$ et $x_2 \in \mathbb{R}$ tel que $\sinh(x_2) > y$. Le théorème de la valeur intermédiaire conclut qu'il existe un x entre x_1 et x_2 tel que $\sinh(x) = y$. Cela prouve la surjectivité.

Pour l'injectivité, on va utiliser le théorème de Rolle 12.189 et une petite preuve par l'absurde. Supposons que $\sinh(x_1) = \sinh(x_2)$ avec $x_1 \neq x_2$. Dans ce cas, le théorème de Rolle nous dit qu'il existe un x entre x_1 et x_2 tel que $\sinh'(x) = 0$. La dérivée de sinus hyperbolique étant cosinus hyperbolique, il faut se demander il existe un x tel que $\cosh(x) = 0$. Étant donné que $e^x > 0$ pour tout x, en fait le cosinus hyperbolique ne s'annule jamais. □

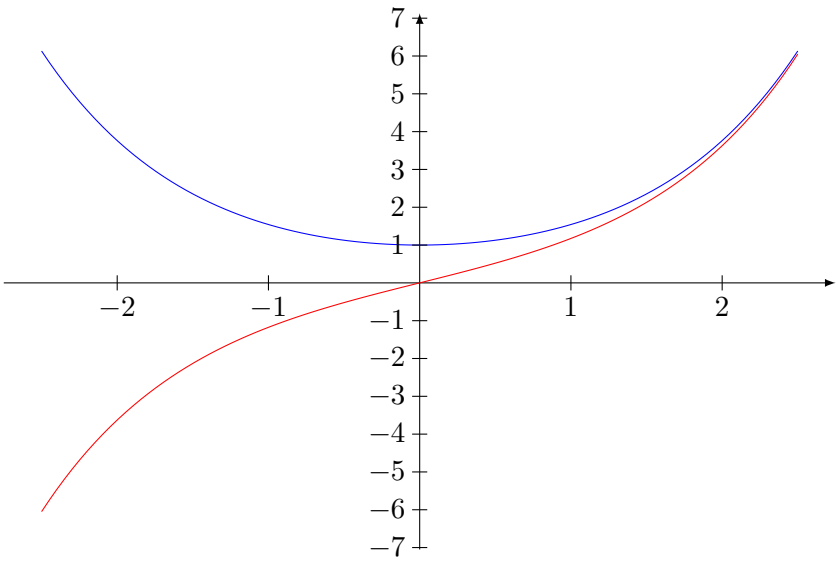


FIGURE 42.1: En rouge, la fonction $x \mapsto \sinh(x)$ et en bleu, la fonction $x \mapsto \cosh(x)$. LabelFig5YNKa8ZBuEWsWw

Un très bon exercice serait de faire un étude complète des fonctions cosinus et sinus hyperbolique. Leur graphes sont donnés à la figure 42.1

Un corolaire de la surjectivité de sinh sur $\mathbb{R}$ est que si je prends n'importe quel deux nombres dont la différence des carrés vaut 1, alors ces carrés sont représentables avec des fonctions hyperboliques :

$$\forall x, y \in \mathbb{R} \text{ tels que } x^2 - y^2 = 1,\ \exists \xi \in \mathbb{R} \text{ tel que } x^2 = \cosh(\xi) \text{ et } y^2 = \sinh(\xi).$$

La **tangente hyperbolique** est définie par

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}. \tag{42.32}$$

Un bon exercice est de prouver les deux relations suivantes :

$$\sinh(x) = \frac{\tanh(x)}{\sqrt{1-\tanh^2(x)}} \qquad \cosh(x) = \frac{1}{\sqrt{1-\tanh^2(x)}}. \tag{42.33}$$ EqRelSinhthcosh

8. Théorème 9.93.

42.4.8.2 Les transformations de Lorentz (bis)

Nous avons prouvé qu'en relativité, l'intervalle est un invariant. Pour cela, nous avons utilisé les transformations de Lorentz démontrées à partir de l'hypothèse d'invariance de la vitesse de la lumière. Eh bien, oublions un instant que la vitesse de la lumière soit invariante, et posons à la place comme hypothèse que l'intervalle soit invariant. C'est-à-dire que si Bob mesure un événement aux coordonnées (t,x) et Alice en (t',x'), alors $c^2t^2-x^2=c^2(t')^2-(x')^2$.

Théorème 42.4.
Les transformations de Lorentz sont les seules qui laissent l'intervalle invariant.

Démonstration. Toute la partie comme quoi les transformations doivent êtres linéaires reste parce que cette partie ne demandait pas l'invariance de la vitesse de la lumière.

Nous cherchons donc les transformations entre Alice et Bob sous la forme

$$\begin{aligned}t'&=\alpha t+\beta x\\ x'&=\gamma t+\delta x\end{aligned}$$

telles que $c^2(t')^2-(x')^2=c^2t^2-x^2$. Lorsque Alice passe devant Bob, ils déclenchent tous deux leurs chronomètre et leurs axes. C'est-à-dire que si à ce moment un événement se trouve à droite pour Alice, il est aussi à droite pour Bob. On doit donc avoir, quand $t=t'=0$, que $x>0$ implique $x'>0$. Cela donne la contrainte que $\delta>0$. D'autre part, comme leurs chronomètres vont dans le même sens (ils choisissent tout les deux de *compter* le temps et non *décompter*), on a $\alpha>0$.

En développant l'expression de $(s')^2$ en termes de t et x, on trouve la condition d'invariance de l'intervalle sous la forme :

$$c^2(\alpha^2t^2+2\alpha\beta tx+\beta^2x^2)-(\gamma^2t^2+2\gamma\delta tx+\delta^2x^2)=c^2t^2-x^2, \tag{42.34}$$ EqCondInvINter

qui doit être valable pour tout t et pour tout x. En $t=0$ on trouve la condition

$$\delta^2-c^2\beta^2=1. \tag{42.35}$$

Cela implique qu'il existe un $\xi\in\mathbb{R}$ tel que $\delta^2=\cosh^2(\xi)$ et $c^2\beta^2=\sinh(\xi)$. La première équation donne $\delta=\cosh(\xi)$ (il faut rejeter $\delta=-\cosh(\xi)$ parce qu'on a demandé que $\delta>0$). Pour la seconde, on trouve $c\beta=\sinh(\xi)$ où l'on peut oublier la possibilité $c\beta=-\sinh(\xi)$ parce que cela revient juste à renommer $\xi\to-\xi$ (la fonction sinus hyperbolique est impaire). Bref, il existe un ξ tel que

$$\begin{aligned}\delta&=\cosh(\xi)\\ \beta&=\frac{\sinh(\xi)}{c}\end{aligned} \tag{42.36}$$ EqCondxi

En mettant maintenant $x=0$ dans la condition (42.34), on trouve la condition

$$\alpha^2-\frac{\gamma^2}{c^2}=1.$$

Pour les mêmes raisons qu'avant, il existe un $\eta\in\mathbb{R}$ tel que

$$\begin{aligned}\alpha&=\cosh(\eta)\\ \gamma&=c\sinh(\eta).\end{aligned} \tag{42.37}$$ EqCondeta

Rien qu'en regardant deux cas très particuliers, on a déjà bien avancé, non ? Remettons maintenant les valeurs (42.36) et (42.37) dans la condition (42.34). En utilisant l'identité $\cosh^2(x)-\sinh^2(x)=1$, et en séparant les termes en t^2, x^2 et tx pour satisfaire la condition, il faut

$$\cosh(\eta)\sinh(\xi)=\sinh(\eta)\cosh(\xi) \tag{42.38}$$ EqConsetaxi

parce que les termes en t^2 et x^2 donnent exactement $c^2t^2 - x^2$ et qu'il faut que le terme en tx s'annule. Mettons la condition (42.38) au carré, et substituons $\cosh^2(\eta) = 1 + \sinh^2(\eta)$ et $\cosh^2(\xi) = 1 + \sinh^2(\xi)$, il reste

$$\sinh^2 \xi = \sinh^2 \eta,$$

ce qui signifie $\sinh \xi = \pm \sinh \eta$, ou encore $\xi = \pm\eta$. On voit que $\xi = -\eta$ ne fonctionne pas dans (42.38), donc on reste avec $\xi = \eta$ et les transformations prennent la forme

$$\begin{aligned} t' &= \cosh(\xi)t + \frac{\sinh(\xi)}{c}x \\ x' &= c\sinh(\xi)t + \cosh(\xi)x. \end{aligned} \tag{42.39}$$

EqLorxi

Ce que nous avons prouvé, c'est qu'il existe un $\xi \in \mathbb{R}$ tel que les transformations entre Alice et Bob aient cette forme. Il faut trouver ce que vaut ξ en fonction de la vitesse v à laquelle Alice court.

Pour ce faire, étudions le mouvement d'Alice. Bob la voit aux coordonnées (t, vt), ce qui correspond à

$$x' = c\sinh(\xi)t + \cosh(\xi)vt$$

pour Alice. Mais ces coordonnées sont celles de Alice elle-même, donc $x' = 0$, ce qui donne[9] $vt = -c\sinh(\xi)t/\cosh(\xi)$, ou encore

$$\tanh(\xi) = -\frac{v}{c} \tag{42.40}$$

En utilisant les relations (42.33), on trouve

$$\cosh(\xi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \sinh(\xi) = \frac{-v/c}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{42.41}$$

En remettant ces valeurs dans les transformations (42.39), on trouve

$$t' = \frac{t - \frac{v}{c^2}x}{\gamma(v)} \tag{42.42}$$

$$x' = \frac{x - vt}{\gamma(v)}, \tag{42.43}$$

exactement les transformations de Lorentz !

□

Ce résultat est important pour une raison assez simple : maintenant, la théorie de la relativité est indépendante de toute considérations sur la lumière. En effet, ce que nous venons de prouver, c'est que si il existe une vitesse c telle que $c^2t^2 - x^2 = c^2(t')^2 - (x')^2$, alors (t, x) et (t', x') sont liés par les transformations de Lorentz.

42.4.9 Vitesse limite

Afin de nous passer de l'hypothèse d'invariance de la vitesse de la lumière, nous avons prouvé que l'hypothèse d'invariance de l'intervalle était suffisante. Mais il faut avouer que cette hypothèse n'est pas très intuitive. Nous allons montrer maintenant que l'existence d'une vitesse limite est une troisième hypothèse qui peut être utilisée comme alternative aux deux premières.

42.5 Applications

Une première application sympa est le logiciel[10] *lightspeed*. Si tu es sous Ubuntu-Linux, installe juste le paquet nommé lightspeed, et régales-toi ! Tu verras c'est marrant. Si tu utilise des fenêtres, je laisse faire l'adage « Windows c'est facile ».

9. Conditions d'existence : $\cosh(\xi) \neq 0$; heureusement, nous avons vu que le cosinus hyperbolique ne s'annule jamais.

10. jeu de mot sur « application » ! ah ah !

42.5.1 Le GPS

Pour qu'un système GPS puisse te localiser, en gros, il t'envoie un signal, tu lui réponds et il mesure le temps qu'il a fallu à la lumière pour faire l'aller-retour. Déjà, tu remarques que cela n'est possible que grâce au fait que la vitesse de la lumière soit finie. Sinon, le GPS ne fonctionnerait pas. Mais il y a mieux.

Comme pour te localiser il faut plusieurs satellites en plus de ton appareil, il faut que les horloges internes de tout ce petit monde soient bien synchronisées, sinon pour mesurer des intervalles de temps et calculer des distances, c'est mal parti. Eh mais un satellite, ça bouge assez vite (surtout que les mesures doivent être très précises), et en plus ça ne fait même pas un MRU, vu que ça tourne en rond. Comme tu vois tout le travail qu'il a fallu faire pour trouver les transformations de Lorentz d'un MRU, tu t'imagines le travail pour un mouvement circulaire ! Eh bien ce travail a été fait, et le résultat est que si on en tient pas compte, les contractions temporelles liées à la relativité sont suffisamment grandes pour complètement dérégler le GPS.

42.5.2 Les ondes électromagnétiques

Tu te souviens qu'au début du chapitre, nous avons dit que le problème qui a amené la relativité était la propagation du champ électrique. Maintenant que nous avons déjà vu une partie des conséquences du problème, il est temps de se rendre compte que les champs électriques et magnétiques sont les objets les plus soumis aux bizarreries relativistes du monde : elles se propagent à la vitesse de la lumière. Regarde un coup autour de toi ; tout ce qui est champ électromagnétique a besoin de la relativité pour être bien compris : GSM, lumière, four à micro-onde, radio, wifi, fibre optique, ...

Si un jour un ingénieur te dit qu'il n'y a pas besoin de connaitre la relativité pour inventer la radio (c'est vrai : la radio a été inventée avant la relativité), ni pour construire une fibre optique, dis lui en pensant à moi qu'il utilise tout le temps les équations de Maxwell [11], et que ces équations sont relativistes.

Bref, soit convaincu que tu vis dans un monde relativiste et que les transformations de Lorentz te suivent à chacun de tes pas.

42.6 Mécanique relativiste

Cela est bien beau, mais la dilatation du temps, et les contractions de longueurs doivent bien avoir des répercussions sur la cinématique et la dynamique des objets. Est-ce que le théorème de l'énergie cinétique est encore valable ? est-ce que les lois de Newton tiennent encore la route ?

42.6.1 Des problèmes, toujours des problèmes

Attardons-nous un peu pour faire quelques commentaires sur cette citation du chevalier pégase dans les chevaliers du zodiaque :

> Ses coups vont à la vitesse de la lumière et pourtant je les vois distinctement arriver.

Est-ce possible ? Nous avons vu qu'il y avait des dénominateurs qui s'annulent quand des objets se déplacent plus vite que la lumière ; or pour voir venir un rayon de lumière qui vient vers soi, il faudrait que le rayon émette de la lumière devant elle. Ça semble un peu mal parti pour respecter les lois de la relativité, non ?

Cela pose en tout cas une question qu'il faudra résoudre. On *entend* venir une ambulance parce qu'elle émet du son qui avance plus vite qu'elle. Pas de problèmes avec ça. Mais quid de la *voir* venir ?

On peut voir venir un tram parce qu'il émet de la lumière ; cette lumière allant plus vite que le tram, elle arrive à nos yeux avant le tram lui-même. Cela est très bien. Mettons que le tram avance à 50 km/h ; pour le conducteur, la lumière de son phare avant avance devant lui à la vitesse

11. C'est sous ce nom là qu'on nomme l'ensemble des équations de l'électromagnétisme comme la loi de l'induction.

c. Par conséquent pour un observateur au sol, cette même lumière devrait avancer à la vitesse $c+50$. Encore une fois, on a un problème d'invariance de la vitesse de la lumière ; mais comme c'est de la lumière, on est habitué à ce que des trucs bizarres arrivent. On ne sera pas étonné que $c+50=c$ d'une manière ou d'une autre[12]. Pire. Si un vaisseau spatial avance à la vitesse 200000 km/s et qu'il envoie en reconnaissance un vaisseau devant lui à la vitesse de 150000 km/s, le vaisseau de reconnaissance ira à la vitesse 150000 km/s par rapport au vaisseau principal. Et par rapport au sol, il ira à la vitesse $150000+200000=350000$ km/s, ce qui est impossible. Il faudra trouver quelque chose pour que ça se passe bien.

Un autre problème maintenant.

Prenons une masse m que l'on soumet à une force constante F. Par la loi de Newton, $a=F/m$ est constante et la vitesse après un temps t vaut $v=Ft/m$. Pas de bol, ça devient plus grand que la vitesse de la lumière à partir du temps $t=cm/F$. Ça est un problème hein ? Il faut trouver un truc pour qu'avec une force constante, l'accélération diminue.

42.6.2 Loi d'addition des vitesses

Si Bob observe un objet se déplacer à la vitesse V, alors Alice devrait l'observer bouger à la vitesse $V-v$. Tout comme si une vache voit passer un train à 90 km/h, alors le vélo qui avance à 25 km/h le voit passer à 65 km/h.

Maintenant, tu es habitué à ce que rien ne se passe comme d'habitude, donc tu te doutes bien qu'en réalité la bonne formule ne va pas être $V-v$.

Bob observe l'objet aux coordonnées (t,Vt), ce qui fait pour Alice :

$$\left(\frac{t-\frac{v}{c^2}Vt}{\gamma(v)},\frac{Vt-vt}{\gamma(v)}\right).$$

En divisant le x' d'Alice par le t' d'Alice, on trouve la vitesse mesurée par Alice :

$$V'=\frac{(V-v)t}{\gamma(v)}\frac{\gamma(v)}{t\left(1-\frac{vV}{c^2}\right)}=\frac{V-v}{1-\frac{vV}{c^2}}.$$

La loi de transformation des vitesses relativiste est donc

$$V'=\frac{V-v}{1-\frac{vV}{c^2}}. \tag{42.44}$$ EqAddRelVit

Qu'en est-il de notre $c+50=c$? Disons que Bob lance un bisou à Alice pendant qu'elle arrive vers lui. Le bisou arrive à la vitesse de la lumière (càd $V=c$) tandis que Alice s'approche de Bob à la vitesse 50 m/s (càd $v=50$). Donc la vitesse à laquelle Alice devrait voir arriver le bisou est bien $c+50$. En utilisant la formule d'addition relativiste des vitesses (42.44), nous trouvons

$$V'=\frac{c-50}{1-\frac{50c}{c^2}}=\frac{c-50}{1-\frac{50}{c}}=\frac{c(c-50)}{c-50}=c.$$

Donc effectivement en relativité quand on additionne des vitesses il faut penser à la règle du « $c+50=c$ ».

42.6.3 L'action d'une force

L'équation fondamentale de la mécanique classique est

$$F=ma.$$

Or tu n'es pas sans savoir que l'accélération est la dérivée seconde de la position par rapport au temps. Nous noterions donc $F=mx''(t)$. Le problème est évidemment que si F est constante, on trouve $v=Ft/m$ qui dépasse toujours la vitesse c quand t est assez grand. Il faudra donc modifier

12. et je ne te cache pas que c'est ce qui va arriver.

la loi $F = ma$. Pour cela, posons-nous des questions sur la dérivée $x'(t)$. On dérive par rapport au temps ; oui mais nous avons vu que le temps n'est pas le même pour tout le monde. Introduisons donc la notation

$$v = \frac{dx}{dt} \tag{42.45}$$

qui ne signifie rien d'autre que nous dérivons x par rapport à t et non par rapport au temps t' de quelqu'un d'autre. Dans le cadre de la relativité, ce que signifie l'équation (42.45) est que v est la dérivée de x par rapport à t. Dans le cas où x et t sont les coordonnées de la position d'Alice mesurées par Bob, cela signifie qu'on dérive la position *mesurée par Bob* par rapport au temps *mesuré par Bob*.

Ce que dit la relativité est que cette quantité v ne peut pas varier proportionnellement à la force sous peine de dépasser la vitesse de la lumière. La subtilité est de modifier la loi de Newton en disant que la quantité qui varie sous l'action d'une force n'est plus $dx/dt = v$, mais

$$\frac{dx}{dt'} = \frac{v}{\sqrt{1-\frac{v^2}{c^2}}},$$

c'est-à-dire la dérivée de la position *mesurée par Bob* par rapport au temps *mesuré par Alice* ! La loi de Newton $v = Ft/m$ devient donc

$$\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{Ft}{m}. \tag{42.46}$$

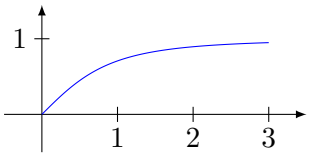
Est-ce que cela résoud le problème ? Pour le savoir, regardons la vitesse acquise par le mobile de masse m soumit à la force F pendant un temps t Il faut résoudre l'équation (42.46) par rapport à v et voir si cela reste bien toujours inférieur à c. On commence par mettre la racine à droite et à élever toute l'équation au carré :

$$\begin{aligned} v^2 &= \frac{F^2t^2}{m^2}\left(1-\frac{v^2}{c^2}\right) \\ v^2\left(1+\frac{F^2t^2}{c^2m^2}\right) &= \frac{F^2t^2}{m^2} \\ v &= \frac{\sqrt{F^2t^2/m^2}}{\sqrt{1+\frac{F^2t^2}{c^2m^2}}}, \end{aligned}$$

et donc finalement

$$v(t) = \frac{Ft}{m\sqrt{1+\frac{F^2t^2}{c^2m^2}}}. \tag{42.47}$$

Tu dois remarquer que si F et t ne sont pas trop grands, l'expression F^2t^2/c^2m^2 est minuscule parce que c est énorme. Si on fait l'approximation $F^2t^2/c^2m^2 = 0$ dans cette expression, on retrouve $v = Ft/m$. Cela montre qu'à moins de faire des expérience avec de très grandes forces pendant énormément de temps, on ne peut pas voir la différence entre la mécanique de Newton et la mécanique relativiste.



Sur le graphe suivant, la vitesse en fonction du temps lorsqu'une particule de masse $m = 1$ est soumise à une force constante. Pour les besoins du graphique, nous avons mis à 1 la vitesse c. Tu vois que quand la vitesse n'est pas très grande, le graphique est presque celui d'une droite ; et à partir d'un certain moment, la courbe s'infléchit pour tendre vers 1 sans l'atteindre.

Remarque que si on maintient une accélération constante égale à celle de la gravité terrestre pendant deux heures, on arrive déjà sur la Lune, à une vitesse de 75 km/s, c'est-à-dire encore rien par rapport à la vitesse de la lumière ! Cela pour te dire que la formule (42.47) a l'air d'être très différente de la formule classique $v = Ft/m$, mais en réalité tant qu'on n'atteint pas des forces énormes, elle ressemble très fort.

Vérifions maintenant que la formule (42.47) n'est pas en contradiction avec l'impossibilité de dépasser la vitesse de la lumière. Pour cela, regardons ce qu'il se passe si on applique une force constante F sur un objet de masse m pendant un temps très long. C'est-à-dire : calculons la limite

$$\lim_{t\to\infty} v(t).$$

Tu vois tout de suite qu'on est sur un cas $\frac{\infty}{\infty}$, ce qui t'oblige à utiliser la règle de l'Hospital. On peut cependant un peu réfléchir et deviner la réponse sans passer par des math trop compliquées.

En effet, quand t est vraiment énorme, l'expression $\frac{F^2t^2}{m^2c^2}$ devient très grande, et le 1 qui se trouve à côté ne vaut plus grand chose, on peut le négliger.

$$\begin{aligned}\lim_{t\to\infty} \frac{Ft}{m\sqrt{1+\frac{F^2t^2}{c^2m^2}}} &= \lim_{t\to\infty} \frac{Ft}{m\sqrt{\frac{F^2t^2}{m^2c^2}}} \\ &= \lim_{v\to\infty} \frac{Ft}{m\frac{Ft}{cm}} \\ &= c.\end{aligned} \tag{42.48}$$

Tout est bien : on arrive au maximum à la vitesse de la lumière, mais il faut un temps infini pour y parvenir. Conclusion : il n'est pas possible d'accélérer un objet jusqu'à atteindre la vitesse de la lumière.

42.6.4 Équivalence entre la masse et l'énergie

Le moment est venu de montrer ce que signifie la fameuse formule $E = mc^2$.

42.7 Principe de correspondance

Nous ne sommes pas parvenu à démontrer la formule (42.46) de la mécanique relativiste qui montre comme un objet accélère sous l'effet d'une force constante. Nous avons juste montré qu'il fallait modifier la loi $v = Ft/m$ et nous avons prit la première modification qui nous soit tombée sous la main, à savoir qu'il faut dériver la position par rapport au temps de l'objet qu'on observe plutôt que par rapport au temps de l'observateur.

En fait, il est possible de prouver rigoureusement [13] la formule

$$\frac{Ft}{m} = \frac{\alpha v}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Mais il n'y a pas moyen de trouver la valeur de la constante α. Tout ce qu'il y a moyen de trouver avec l'hypothèse de l'invariance de la vitesse de la lumière est l'existence d'une constante telle que cette formule soit vraie.

Afin de fixer la constante α, il faut faire intervenir un principe physique supplémentaire, le **principe de correspondance**

Loi numéro 4.
Lorsque la vitesse d'une particule est faible, les équations doivent être en première approximation les mêmes que celles de la mécanique classique.

13. Mais il n'existe pas de démonstrations simples à ma connaissance.

Que signifie *en première approximation* ? Tu sais qu'une fonction $x \mapsto f(x)$ peut être approximée (pour des petits x) par la formule

$$f(x) \simeq f(0) + xf'(0).$$

Nous voudrions donc que Ft/m soit en première approximation égal à v. Nous devons étudier la fonction

$$f(v) = \frac{\alpha v}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Voir ce que vaut cette fonction en première approximation lorsque v est petit est un exercice de dérivation. En utilisant la règle de dérivation des fractions, on trouve que

$$f'(v) = \frac{\alpha}{\sqrt{1 - \frac{v^2}{c^2}}} + \frac{\alpha v^2}{c^2 \left(1 - \frac{v^2}{c^2}\right)^{3/2}},$$

et donc que $f'(0) = \alpha$. Bien entendu, $f(0) = 0$. En première approximation, nous trouvons donc

$$f(v) \simeq \alpha v \tag{42.49}$$

qui doit être égal à la quantité non relativiste v. Nous en déduisons qu'il faut fixer $\alpha = 1$, et on tombe sur la formule relativiste proposée plus haut

$$\frac{Ft}{m} = \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}}$$

L'utilisation cruciale du principe de correspondance a une répercussion énorme sur notre vision de la physique. En effet, la relativité d'Einstein ne parvient pas à *remplacer* la mécanique de Newton. On a besoin d'invoquer la mécanique de Newton pour fixer la théorie. On peut écrire l'axiome suivant :

$$\lim_{v \to 0} \text{Einstein} = \text{Newton}. \tag{42.50}$$

Cela n'est pas une propriété de la théorie d'Einstein, mais un de ses axiomes !

La relativité ne fait donc pas table rase des principes physiques de la mécanique newtonienne : elle les complète et les contient.

Chapitre 43

Exemples avec Sage

Ce chapitre est un foure-tout de choses que l'on peut faire avec Sage.

43.1 Graphiques

Pour afficher le graphe d'une fonction, vous pouvez faire

```
+--------------------------------------------------------------------+
| SageMath version 8.1, Release Date: 2017-12-07                     |
| Type "notebook()" for the browser-based notebook interface.        |
| Type "help()" for help.                                            |
+--------------------------------------------------------------------+
sage: plot(cos(x),0,5)
Launched png viewer for Graphics object consisting of 1 graphics primitive
sage: f(x)=sin(x)
sage: f.plot(-pi,pi)
Launched png viewer for Graphics object consisting of 1 graphics primitive
```

Un programme externe se lance automatiquement pour afficher le graphique que vous avez demandé.

Il se peut qu'aucun programme ne se lance et vous ayez, au lieu de `Launched png viewer for Graphics object ...` uniquement `Created graphics object ...`. Disons pour faire court que Sage a produit un `png` et qu'il ne sait pas quel programme externe utiliser pour l'afficher.

La solution est à l'adresse http://doc.sagemath.org/html/en/reference/misc/sage/misc/viewer.html

43.1.1 Autres

Dans le but d'automatiser certaines tâches, j'ai écrit ce module, nomé `outilsINGE.sage`, dans le cadre d'un cours de première année donné à des ingénieurs. Certaines des fonctions définies ici sont utilisée dans les exemples qui suivent.

```
1  # -*- coding: utf8 -*-
2  from sage.all import *
3
4  """
5  This module provides _pragmatic_ tools for solving exercise for
6  a first year in general mathematics.
7  """
8
9  # TODO : trouver une bonne traduction pour "point de selle."
10
```

```
def automatedVar(symbol,n):
  """ If symbol = "x" and n=4, return the string 'x1,x2,x3,x4'
     """
  s = ",".join([ symbol+str(i) for i in range(1,n+1)])
  return s

class SolveLinearSystem(object):
  """
  Solve Ax=v and print it in a nice way

  Example :

  A=matrix([ [1,-2,3,-2,0],[3,-7,-2,4,0],[4,3,5,2,0] ])
  v=vector((0,0,0,0,0))
  print SolveLinearSystem(A,v)
  """
  def __init__(self,A,v):
    self.matrix = A
    self.vector = v
    self.nvars = A.ncols()
    s = automatedVar("x",self.nvars)
    self.xx=var(s)
  def equations(self):
    """Return the equations corresponding to the
                self.matrix and self.vector as a list of equations
                   """
    X=matrix( [self.xx[i] for i in range(0,self.nvars) ]).
       transpose()
    eqs=[]
    for i in range(0,self.matrix.nrows()):
      exp = (self.matrix*X)[i][0]==self.vector[i]
      eqs.append(exp)
    return eqs
  def solutions(self):
    return solve(self.equations(),self.xx)
  def latex(self):
    """Return the LaTeX's code of the system."""
    a=[]
    a.append(r"""

        \item
          $
          \left\{
          \begin{array}{ll}
    """)
    for eq in self.equations():
      a.append("          "+str(eq).replace("x","x_").replace("*"
         ,"").replace("==","=")+"\\\\ \n")
    a.append(r"""
          \end{array}
          \right.
          $
          """)
```

```
60      return "".join(a)
61    def __str__(self):
62      a = []
63      a.append("The given matrix corresponds to the system")
64      for eq in self.equations():
65        a.append(str(eq))
66      a.append("And the solutions are")
67      a.append(str(self.solutions()))
68      return "\n".join(a)
69
70  def QuadraticMap(A,v):
71    """
72    Return the result of the quadratic form associated
73          with A applied on the vector v, that is the number
74    A_ij v^iv^j
75    using the summation convention.
76    """
77    n = A.nrows()
78    if not A.is_symmetric():
79      print "Warning : Given matrix is not symmetric"
80    if not A.is_square():
81      raise TypeError,"Error : The matrix A is not square"
82    if not v.degree()==n :
83      raise TypeError,"The size do not agree"
84    return sum([ A[i,j]*v[i]*v[j] for i in range(n) for j in range(↩
       n)  ]).simplify_full()
85
86  class SymmetricMatrix(object):
87    """
88    Provide informations about the matrix A assuming it is symmetric↩
         .
89    """
90    def __init__(self,A):
91      if not A.is_square():
92        print "Error : A symmetric matrix must be square"
93        raise TypeError
94      self.matrix = A
95      self.degree = A.nrows()
96      self.matrix.set_immutable()
97    def primary_principal_submatrix(self,n):
98      """
99      Return the primary principal submatrix of order n, that is the↩
          matrix obtained
100                 by removing the n last lines and columns from self↩
                       .
101     """
102     taille=self.degree-n
103     v=[]
104     for i in range(0,taille):
105       v.append(self.matrix[i][0:taille])
106     return matrix(v)
107   def principal_minors(self):
108     """
```

```
109     Return the list of principal minors. The principal minor of order k is
110                 the determinant of the primary principal matrix of order k.
111     """
112     a=[]
113     for i in range(self.degree):
114       a.append(self.primary_principal_submatrix(i).determinant())
115     return a
116   def genre_list(self):
117     """
118     Return the genius of the matrix as a list of booleans in the order
119     positive defined, negative defined;
120                 semidefinite positive, semidefinite negative, indefinite.
121
122     """
123     defpos = True
124     defneg = True
125     semidefpos = True
126     semidefneg = True
127     indefinie=True
128     mineurs = self.principal_minors()
129     for i in range(len(mineurs)):
130       m = mineurs[i]
131       if m == 0:
132         defneg=False
133         defpos=False
134       if m < 0:
135         defpos=False
136         semidefpos=False
137         if i%2==0:
138           defneg=False
139       if m > 0:
140         semidefneg=False
141         if i%2==1:
142           defneg=False
143     if 0 not in mineurs:
144       semidefneg=False
145       semidefpos=False
146     if (defpos==True)or(defneg==True)or(semidefpos==True)or(semidefneg==True):indefinie=False
147     return [defpos,defneg,semidefpos,semidefneg,indefinie]
148   def __str__(self):
149     return str(self.matrix)
150
151 class QuadraticForm(SymmetricMatrix):
152   """
153   From a symmetric matrix A, provide informations concerning the associated quadratic form.
154   """
155   def __init__(self,A):
```

```
156     SymmetricMatrix.__init__(self,A)
157     if not A.is_symmetric():
158       print "Warning : matrix is not symmetric"
159   def evaluate(self,v):
160     """
161     Return the value of the quadratic form on the vector v.
162     """
163     return QuadraticMap(self.matrix,v)
164   def diagonalizing_martrix(self):
165     """
166     Return the matrix B such that B^tAB is diagonal.
167     """
168     # The transposition is because, in the matrix B, the ↩
          eigenvectors have
169     # to be read as column while Sage's matrix constructor takes ↩
          rows.
170     return matrix(self.orthonormal_basis()).transpose()
171   def new_variables(self):
172     """
173     Give the change of variables needed to put the quadratic form ↩
          under its normal form
174     X=BY
175     where X are the "old" variables
176     """
177     variables = var(automatedVar("y",self.degree))
178     Y = vector(variables)
179     return self.diagonalizing_martrix()*Y
180   def eigenmatrix_left(self):
181     return self.matrix.eigenmatrix_left()
182   def eigenvectors(self):
183     """
184     Return a list of eigenvectors of the matrix.
185
186     As the matrix is symmetric, that list has to be a basis.
187     """
188     D,P = self.eigenmatrix_left()
189     return [P[i] for i in range(P.nrows())]
190   def eigenvalues(self):
191     """
192     Return a list of eigenvalues of the matrix in the same order ↩
          as the list of eigenvectors given in
193       self.eigenvectors()
194     """
195     D,P = self.eigenmatrix_left()
196     return [ D[i,i] for i in range(D.nrows()) ]
197   def orthonormal_basis(self):
198     """
199     Return a basis of eigenvectors normalised to 1 as a list.
200     """
201     M,mu = matrix(self.eigenvectors()).gram_schmidt()
202     return [ v/v.norm() for v in M ]
203   def verification(self):
204     """
```

```
205       return the value of the quadratic form on the vector ↩
              new_variables()
206       """
207       return self.evaluate(self.new_variables())
208
209     def __str__(self):
210       a = []
211       a.append("Hi guy; I'm the quadratic form associated with the ↩
              matix")
212       a.append(str(self.matrix))
213       a.append("My eigenvalues and eigenvectors are : ")
214       veps = self.eigenvectors()
215       vaps = self.eigenvalues()
216       for i in range(len(veps)):
217         a.append("%s -> %s"%(str(vaps[i]),str(veps[i])))
218       a.append("I've the following orthonormal basis of eigenvectors↩
              :")
219       for v in self.orthonormal_basis():
220         a.append(str(v))
221       a.append("A matrix B such that B^tAB is diagonal is ")
222       a.append(str(self.diagonalizing_martrix()))
223       a.append("I'm quite pretty in the following variables ...")
224       for i in range(self.degree):
225         a.append("x%s = %s"%(str(i+1),str(self.new_variables()[i]))↩
              )
226       a.append("Look at me when I wear my cool variables")
227       a.append(str(self.verification()))
228       return "\n".join(a)
229
230 class Extrema(object):
231     """
232     From a function f, provides the informations for the study of ↩
          the extrema :
233     partial derivative
234     critical points
235     Hessian matrix at the critical points
236     Genius of the Hessian and conclusion as local min/max
237
238     Dear student : remember that this class does not furnish any ↩
          informations
239         concerning *global* extrema. The latter have to be found
240       among the critical points OR on the border of the domain.
241     """
242     def __init__(self,f):
243       var('x,y')
244       self.fun = f
245       self.gx=self.fun.diff(x).full_simplify()
246       self.gy=self.fun.diff(y).full_simplify()
247       self.gxx=self.gx.diff(x).simplify_full()
248       self.gxy=self.gx.diff(y).full_simplify()
249       self.gyy=self.gy.diff(y).full_simplify()
250       self.cp = solve( [self.gx(x,y)==0,self.gy(x,y)==0],[x,y] )
251     def critical_points(self):
```

```
252     """
253     Return the critical points as a list of tuples (x,y)
254     """
255     a = []
256     for pt in self.cp :
257       try :
258         px = SR(pt[0].rhs())
259         py = SR(pt[1].rhs())
260         a.append((px,py))
261       except TypeError :
262         a.append("  I'm not able to solve these equations.")
263     return a
264   def hessienne(self,a,b):
265     return matrix(SR,2,2,[self.gxx(a,b),self.gxy(a,b),self.gxy(a,↩
          b),self.gyy(a,b)])
266   def __str__(self):
267     a = []
268     a.append ("The function :")
269     a.append(str(self.fun))
270     a.append ("Derivative x and y :")
271     a.append(str(self.gx))
272     a.append(str(self.gy))
273     a.append ("Hessian matrix :")
274     a.append(str(self.hessienne(x,y)))
275     a.append ("Critical points :")
276     for pt in self.critical_points() :
277       a.append(str(pt))
278     for pt in self.critical_points():
279       try :
280         px = pt[0]
281         py = pt[1]
282         a.append("At (%s,%s), the Hessian is"%(str(px),str(py)))
283         try :
284           Hess = SymmetricMatrix(self.hessienne(px,py))
285           for l in Hess.matrix:
286             a.append("  "+str(l))
287           a.append("  Primary principal minors are %s"%str(Hess.↩
                principal_minors()))
288           l = Hess.genre_list()
289           if l[0]==True:
290             a.append("  Hessian positive defined")
291             a.append("  local minimum")
292           if l[1]==True:
293             a.append("  Hessian negative defined")
294             a.append("  local maximum")
295           if l[2]==True:
296             a.append("  Hessian positive semidéfinite")
297             a.append("  I don't conclude")
298           if l[3]==True:
299             a.append("  Hessian negative semidefinite")
300             a.append("  I don't conclude")
301           if l[4]==True:
302             a.append("  Undefinite Hessian")
```

```
303             a.append("  «selle» point")
304         except RuntimeError,data :
305           a.append("  "+str(data))
306       except TypeError :
307         a.append("  I'm not able to solve these equations.")
308     return "\n".join(a)
```

tex/sage/outilsINGE.sage

ExBCRXooDVUdcf

Exemple 43.1.
Calculer la limite

$$\lim_{x\to\infty} \frac{\sin(x)\cos(x)}{x} \tag{43.1}$$

```
var('x')
f(x)=sin(x)*cos(x)/x
limit(f(x),x=oo)
```

La première ligne déclare que la lettre `x` désignera une variable. Pour la troisième ligne, notez que l'infini est écrit par deux petits « o ». △

ExCWDRooKxnjGL

Exemple 43.2.
Quelques limites et graphes avec Sage.

(1) $\lim_{x\to 0} \frac{\sin(\alpha x)}{\sin(\beta x)}$.
Pour effectuer cet exercice avec Sage, il faut taper les lignes suivantes :

```
sage: var('x,a,b')
(x, a, b)
sage: f(x)=sin(a*x)/sin(b*x)
sage: limit( f(x),x=0  )
a/b
```

Notez qu'il faut déclarer les variables `x`, `a` et `b`.

(2) $\lim_{x\to\pm\infty} \frac{\sqrt{x^2+1}-x}{x-2}$

```
sage: f(x)=(sqrt(x**2+1))/(x-2)
sage: limit(f(x),x=oo)
1
sage: limit(f(x),x=-oo)
-1
```

Notez la commande pour la racine carré : `sqrt`. Étant donné que cette fonction diverge en $x = 2$, si nous voulons la tracer, il faut procéder en deux fois :

```
sage: plot(f,(-100,1.9))
Launched png viewer for Graphics object consisting of 1 graphics primitive
sage: plot(f,(2.1,100))
Launched png viewer for Graphics object consisting of 1 graphics primitive
```

La première ligne trace de -100 à 1.9 et la seconde de 2.1 à 100. Ces graphiques vous permettent déjà de voir les limites. Attention : ils ne sont pas des *preuves* ! Mais ils sont de sérieux indices qui peuvent vous inspirer dans vos calculs.

△

exJMGTooZcZYNy

Exemple 43.3.
Calculer les dérivées partielles $\partial_x f$, $\partial_y f$, $\partial^2_x f$, $\partial^2_{xy} f$, $\partial^2_{yx} f$ et $\partial^2_y f$ des fonctions suivantes.

(1) $2x^3 + 3x^2y - 2y^2$
(2) $\ln(xy^2)$
(3) $\tan(x/y)$
(4) $\frac{xy^2}{x+y}$

Le script Sage suivant (`exoDV002.sage`) résout l'exercice :

```
# -*- coding: utf8 -*-

def LesCalculs(f):
  print "Pour la fonction %s"%str(f)
  print "d_x",f.diff(x).simplify_full()
  print "d_y",f.diff(y).simplify_full()
  print "d^2_x",f.diff(x).diff(x).simplify_full()
  print "d_xd_y",f.diff(x).diff(y).simplify_full()
  print "d_yd_x",f.diff(y).diff(x).simplify_full()
  print "d^2_y",f.diff(y).diff(y).simplify_full()
  print ""

def exercise_DV002():
  var('x,y')
  fa(x,y)=2*x**3+3*x**2*y-2*y**2
  fb(x,y)=ln(x*y**2)
  fc(x,y)=tan(x/y)
  fd(x,y)=x*y**2/(x+y)
  LesCalculs(fa)
  LesCalculs(fb)
  LesCalculs(fc)
  LesCalculs(fd)
```

tex/sage/exoDV002.sage

La sortie est :

```
Pour la fonction (x, y) |--> 2*x^3 + 3*x^2*y - 2*y^2
d_x (x, y) |--> 6*x^2 + 6*x*y
d_y (x, y) |--> 3*x^2 - 4*y
d^2_x (x, y) |--> 12*x + 6*y
d_xd_y (x, y) |--> 6*x
d_yd_x (x, y) |--> 6*x
d^2_y (x, y) |--> -4

Pour la fonction (x, y) |--> log(x*y^2)
d_x (x, y) |--> 1/x
d_y (x, y) |--> 2/y
d^2_x (x, y) |--> -1/x^2
d_xd_y (x, y) |--> 0
d_yd_x (x, y) |--> 0
d^2_y (x, y) |--> -2/y^2

Pour la fonction (x, y) |--> tan(x/y)
```

```
d_x (x, y) |--> 1/(y*cos(x/y)^2)
d_y (x, y) |--> -x/(y^2*cos(x/y)^2)
d^2_x (x, y) |--> 2*sin(x/y)/(y^2*cos(x/y)^3)
d_xd_y (x, y) |--> -(2*x*sin(x/y) + y*cos(x/y))/(y^3*cos(x/y)^3)
d_yd_x (x, y) |--> -(2*x*sin(x/y) + y*cos(x/y))/(y^3*cos(x/y)^3)
d^2_y (x, y) |--> 2*(x^2*sin(x/y) + x*y*cos(x/y))/(y^4*cos(x/y)^3)

Pour la fonction (x, y) |--> x*y^2/(x + y)
d_x (x, y) |--> y^3/(x^2 + 2*x*y + y^2)
d_y (x, y) |--> (2*x^2*y + x*y^2)/(x^2 + 2*x*y + y^2)
d^2_x (x, y) |--> -2*y^3/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d_xd_y (x, y) |--> (3*x*y^2 + y^3)/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d_yd_x (x, y) |--> (3*x*y^2 + y^3)/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d^2_y (x, y) |--> 2*x^3/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
```

△

exKGDIooVefujD

Exemple 43.4.
Résoudre les systèmes suivants.

(1) $\begin{cases} x_1 - 2x_2 + 3x_3 - 2x_4 = 0 \\ 3x_1 - 7x_2 - 2x_3 + 4x_4 = 0 \\ 4x_1 + 3x_2 + 5x_3 + 2x_4 = 0 \end{cases}$

(2) $\begin{cases} 2x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 2x_2 - x_3 + 3x_4 = 4 \\ 3x_1 + 3x_2 + 3x_3 - 3x_4 = 9 \end{cases}$

(3) $\begin{cases} x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ 3x_1 - x_2 - 4x_3 = 0 \end{cases}$

(4) $\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ x_1 + 4x_2 + 7x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 0 \end{cases}$

(5) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 - x_3 + x_4 = -4 \\ x_1 - x_2 + x_3 + x_4 = 2 \end{cases}$

(6) $\begin{cases} x_1 + 3x_2 + 3x_3 = 1 \\ x_1 + 3x_2 + 4x_3 = 0 \\ x_1 + 4x_2 + 3x_3 = 3 \end{cases}$

(7) $\begin{cases} x_1 - 3x_2 + 2x_3 = -6 \\ -3x_1 + 3x_2 - x_3 = 17 \\ 2x_1 - x_2 = 3 \end{cases}$

(8) $\begin{cases} x_1 - 2x_2 + 3x_3 - 2x_4 = 0 \\ 3x_1 - 7x_2 - 2x_3 + 4x_4 = 0 \\ 4x_1 + 3x_2 + 5x_3 + 2x_4 = 0 \end{cases}$

(9) $\begin{cases} 2x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 2x_2 - x_3 + 3x_4 = 4 \\ 3x_1 + 3x_2 + 3x_3 - 3x_4 = 9 \end{cases}$

(10) $\begin{cases} x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ 3x_1 - x_2 - 4x_3 = 0 \end{cases}$

(11) $\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ x_1 + 4x_2 + 7x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 0 \end{cases}$

(12) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 - x_3 + x_4 = -4 \\ x_1 - x_2 + x_3 + x_4 = 2 \end{cases}$

(13) $\begin{cases} x_1 + 3x_2 + 3x_3 = 1 \\ x_1 + 3x_2 + 4x_3 = 0 \\ x_1 + 4x_2 + 3x_3 = 3 \end{cases}$

(14) $\begin{cases} x_1 - 3x_2 + 2x_3 = -6 \\ -3x_1 + 3x_2 - x_3 = 17 \\ 2x_1 - x_2 = 3 \end{cases}$

Nous résolvons les systèmes en utilisant Sage avec le script suivant.

```
1 # -*- coding: utf8 -*-
2 """
3 Ce script Sage résout un certain nombre
4 de systèmes d'équations linéaires du cours INGE1121
5 """
6
```

```
import outilsINGE

def exercise_1_1_bcdefhi():
  # Exercice 1.1.b (INGE1121)
  A=matrix([ [1,-2,3,-2,0],[3,-7,-2,4,0],[4,3,5,2,0] ])
  v=vector((0,0,0,0,0))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.c (INGE1121)
  A=matrix([ [2,1,-2,3],[3,2,-1,3],[3,3,3,-3] ])
  v=vector((0,4,9))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.d (INGE1121)
  A=matrix([ [1,2,-3],[2,5,2],[3,-1,-4] ])
  v=vector((0,0,0))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.e (INGE1121)
  A=matrix([ [1,2,-1],[2,5,2],[1,4,7],[1,3,3] ])
  v=vector((0,0,0,0))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.f (INGE1121)
  A=matrix([ [1,1,1,1],[1,1,1,-1],[1,1,-1,1],[1,-1,1,1] ])
  v=vector((0,4,-4,2))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.h (INGE1121)
  A=matrix([ [1,3,3],[1,3,4],[1,4,3] ])
  v=vector((1,0,3))
  print outilsINGE.SolveLinearSystem(A,v)
  # Exercice 1.1.i (INGE1121)
  A=matrix([ [1,-3,2],[-3,3,-1],[2,-1,0] ])
  v=vector((-6,17,3))
  print outilsINGE.SolveLinearSystem(A,v)
```

tex/sage/exo11.sage

Le résultat est le suivant :

```
The given matrix corresponds to the system
x1 - 2*x2 + 3*x3 - 2*x4 == 0
3*x1 - 7*x2 - 2*x3 + 4*x4 == 0
4*x1 + 3*x2 + 5*x3 + 2*x4 == 0
And the solutions are
[
[x1 == -23/16*r19, x2 == -5/16*r19, x3 == 15/16*r19, x4 == r19, x5 == r18]
]
The given matrix corresponds to the system
2*x1 + x2 - 2*x3 + 3*x4 == 0
3*x1 + 2*x2 - x3 + 3*x4 == 4
3*x1 + 3*x2 + 3*x3 - 3*x4 == 9
And the solutions are
[
[x1 == 3*r20 - 7, x2 == -4*r20 + 11, x3 == r20, x4 == 1]
]
The given matrix corresponds to the system
x1 + 2*x2 - 3*x3 == 0
2*x1 + 5*x2 + 2*x3 == 0
```

```
3*x1 - x2 - 4*x3 == 0
And the solutions are
[
[x1 == 0, x2 == 0, x3 == 0]
]
The given matrix corresponds to the system
x1 + 2*x2 - x3 == 0
2*x1 + 5*x2 + 2*x3 == 0
x1 + 4*x2 + 7*x3 == 0
x1 + 3*x2 + 3*x3 == 0
And the solutions are
[
[x1 == 9*r21, x2 == -4*r21, x3 == r21]
]
The given matrix corresponds to the system
x1 + x2 + x3 + x4 == 0
x1 + x2 + x3 - x4 == 4
x1 + x2 - x3 + x4 == -4
x1 - x2 + x3 + x4 == 2
And the solutions are
[
[x1 == 1, x2 == -1, x3 == 2, x4 == -2]
]
The given matrix corresponds to the system
x1 + 3*x2 + 3*x3 == 1
x1 + 3*x2 + 4*x3 == 0
x1 + 4*x2 + 3*x3 == 3
And the solutions are
[
[x1 == -2, x2 == 2, x3 == -1]
]
The given matrix corresponds to the system
x1 - 3*x2 + 2*x3 == -6
-3*x1 + 3*x2 - x3 == 17
2*x1 - x2 == 3
And the solutions are
[
[x1 == 37, x2 == 71, x3 == 85]
]
```

△

ExBGCEooPIQgGW

Exemple 43.5.
Pour chacun des systèmes suivants $A \cdot X = B$,

(1) Résoudre le système par échelonnement,

(2) Calculer A^{-1},

(3) Vérifier votre réponse en calculant $A^{-1}B$. Qu'êtes-vous censé obtenir ?

Les énoncés sont

(1)

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 3 & 2 & 2 \\ 5 & 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 10 \\ 1 \\ 4 \end{pmatrix} \tag{43.2}$$

Nous utilisons Sage pour fournir la réponse. Le code suivant résout le système et donne l'inverse de la matrice :

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_1_3():
  A=matrix([[2,1,-2],[3,2,2],[5,4,3]])
  v=vector((10,1,4))
  print outilsINGE.SolveLinearSystem(A,v)
  print "Matrice inverse :"
  print A.inverse()
```

tex/sage/exo13.sage

La sortie est ici :

```
The given matrix corresponds to the system
2*x1 + x2 - 2*x3 == 10
3*x1 + 2*x2 + 2*x3 == 1
5*x1 + 4*x2 + 3*x3 == 4
And the solutions are
[
[x1 == 1, x2 == 2, x3 == -3]
]
Matrice inverse :
[  2/7  11/7  -6/7]
[ -1/7 -16/7  10/7]
[ -2/7   3/7  -1/7]
```

△

exBNGVooIvKfTT

Exemple 43.6.
Sachant que $(-1,0,1,0)$ est un vecteur propre de la matrice

$$A=\begin{pmatrix}2&1&-1&1\\1&0&1&1\\-1&1&2&1\\1&1&1&0\end{pmatrix} \tag{43.3}$$

(1) Diagonaliser A au moyen d'une matrice orthogonale
(2) Écrire la forme quadratique X^tAX sous forme d'une somme pondérée de carrés.

Calculons Av afin de savoir la valeur propre associée au vecteur donné :

$$\begin{pmatrix}2&1&-1&1\\1&0&1&1\\-1&1&2&1\\1&1&1&0\end{pmatrix}\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}=\begin{pmatrix}-3\\0\\3\\0\end{pmatrix}. \tag{43.4}$$

La valeur propre est donc 3. Nous savons donc que $(\lambda-3)$ pourra être factorisé dans le polynôme caractéristique.

Pour le reste de l'exercice c'est standard et c'est résolu de la façon suivante :

```
# -*- coding: utf8 -*-
```

```
2
3 import outilsINGE
4
5 def exercise_6_5():
6   A=matrix(QQ,4,4,[2,1,-1,1,1,0,1,1,-1,1,2,1,1,1,1,0])
7   x=outilsINGE.QuadraticForm(A)
8   print x
```

tex/sage/exo65.sage

qui retourne

```
Hi guy; I'm the quadratic form associated with the matix
[ 2  1 -1  1]
[ 1  0  1  1]
[-1  1  2  1]
[ 1  1  1  0]
My eigenvalues and eigenvectors are :
3 -> (1, 0, -1, 0)
3 -> (0, 1, 2, 1)
-1 -> (1, 0, 1, -2)
-1 -> (0, 1, 0, -1)
I've the following orthonormal basis of eigenvectors :
(1/2*sqrt(2), 0, -1/2*sqrt(2), 0)
(1/2, 1/2, 1/2, 1/2)
(1/6*sqrt(6), 0, 1/6*sqrt(6), -1/3*sqrt(6))
(-1/2*sqrt(1/3), 3/2*sqrt(1/3), -1/2*sqrt(1/3), -1/2*sqrt(1/3))
A matrix B such that B^tAB is diagonal is
[   1/2*sqrt(2)            1/2    1/6*sqrt(6) -1/2*sqrt(1/3)]
[             0            1/2             0  3/2*sqrt(1/3)]
[  -1/2*sqrt(2)            1/2    1/6*sqrt(6) -1/2*sqrt(1/3)]
[             0            1/2   -1/3*sqrt(6) -1/2*sqrt(1/3)]
I'm quite pretty in the following variables ...
x1 = -1/2*sqrt(1/3)*y4 + 1/2*sqrt(2)*y1 + 1/6*sqrt(6)*y3 + 1/2*y2
x2 = 3/2*sqrt(1/3)*y4 + 1/2*y2
x3 = -1/2*sqrt(1/3)*y4 - 1/2*sqrt(2)*y1 + 1/6*sqrt(6)*y3 + 1/2*y2
x4 = -1/2*sqrt(1/3)*y4 - 1/3*sqrt(6)*y3 + 1/2*y2
Look at me when I wear my cool variables
3*y1^2 + 3*y2^2 - y3^2 - y4^2
```

△

exZHGRooTQpVpq

Exemple 43.7.

Rechercher les extrémums des fonctions suivantes

(1) $f(x,y) = 2 - \sqrt{x^2+y^2}$

(2) $f(x,y) = x^3 + 3xy^2 - 15x - 12y$

(3) $f(x,y) = \frac{x^3}{3} + \frac{4y^3}{3} - x^2 - 3x - 4y - 3$

Les corrigés sont créés par le script Sage `exo101.sage`

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_1_A():
```

```
   var('x,y')
   f(x,y)=2-sqrt(x**2+y**2)
   print outilsINGE.Extrema(f)
def exercise_10_1_B():
   var('x,y')
   f(x,y)=x**3+3*x*y**2-15*x-12*y
   print outilsINGE.Extrema(f)
def exercise_10_1_C():
   var('x,y')
   f(x,y)=x**3/3+4*y**3/3-x**2-3*x-4*y-3
   print outilsINGE.Extrema(f)
```

Des réponses :

(1)
```
The function :
(x, y) |--> -sqrt(x^2 + y^2) + 2
Derivative x and y :
(x, y) |--> -x/sqrt(x^2 + y^2)
(x, y) |--> -y/sqrt(x^2 + y^2)
Hessian matrix :
[-sqrt(x^2 + y^2)*y^2/(x^4 + 2*x^2*y^2 + y^4)                          x*y/(x^2 + y^2)^(3
[                        x*y/(x^2 + y^2)^(3/2) -sqrt(x^2 + y^2)*x^2/(x^4 + 2*x^2*y^2 + y
Critical points :
(0, 0)
At (0,0), the Hessian is
   power::eval(): division by zero
```

Ici nous voyons que Sage a du mal à calculer la matrice hessienne en $(0,0)$. En effet, nous tombons sur une division par zéro. Pour résoudre l'exercice, il faut se rendre compte que la fonction $(x,y) \mapsto \sqrt{x^2+y^2}$ est toujours positive et est nulle seulement au point $(0,0)$. Donc f est toujours plus petite ou égale à deux tandis que $f(0,0) = 2$. Le point est donc un maximum global.

(2)
```
The function :
(x, y) |--> x^3 + 3*x*y^2 - 15*x - 12*y
Derivative x and y :
(x, y) |--> 3*x^2 + 3*y^2 - 15
(x, y) |--> 6*x*y - 12
Hessian matrix :
[6*x 6*y]
[6*y 6*x]
Critical points :
(2, 1)
(1, 2)
(-1, -2)
(-2, -1)
At (2,1), the Hessian is
   (12, 6)
   (6, 12)
   Primary principal minors are [108, 12]
   Hessian positive defined
   local minimum
At (1,2), the Hessian is
   (6, 12)
   (12, 6)
   Primary principal minors are [-108, 6]
```

```
   Undefinite Hessian
   «selle» point
At (-1,-2), the Hessian is
   (-6, -12)
   (-12, -6)
   Primary principal minors are [-108, -6]
   Undefinite Hessian
   «selle» point
At (-2,-1), the Hessian is
   (-12, -6)
   (-6, -12)
   Primary principal minors are [108, -12]
   Hessian negative defined
   local maximum
```

(3)
```
The function :
(x, y) |--> 1/3*x^3 - x^2 + 4/3*y^3 - 3*x - 4*y - 3
Derivative x and y :
(x, y) |--> x^2 - 2*x - 3
(x, y) |--> 4*y^2 - 4
Hessian matrix :
[2*x - 2       0]
[      0     8*y]
Critical points :
(3, 1)
(-1, 1)
(3, -1)
(-1, -1)
At (3,1), the Hessian is
   (4, 0)
   (0, 8)
   Primary principal minors are [32, 4]
   Hessian positive defined
   local minimum
At (-1,1), the Hessian is
   (-4, 0)
   (0, 8)
   Primary principal minors are [-32, -4]
   Undefinite Hessian
   «selle» point
At (3,-1), the Hessian is
   (4, 0)
   (0, -8)
   Primary principal minors are [-32, 4]
   Undefinite Hessian
   «selle» point
At (-1,-1), the Hessian is
   (-4, 0)
   (0, -8)
   Primary principal minors are [32, -4]
   Hessian negative defined
   local maximum
```

△

exHWIHooOAvaDQ

Exemple 43.8.
Déterminer les valeurs extrêmes et les points de selle des fonctions suivantes.

(1) $f(x,y) = x^2 + 4x + y^2 - 2y$.
(2) $f(x,y) = e^{x^2+xy}$.
(3) $f(x,y) = e^x \sin(y)$.

Certains corrigés de cet exercice ont étés réalisés par Sage. Le script utilisé est `exo103.sage`

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_3_A():
  var('x,y')
  f(x,y)=x**2+4*x+y**2-2*y
  print outilsINGE.Extrema(f)

def exercise_10_3_H():
  var('x,y')
  f(x,y)=exp(x**2+x*y)
  print outilsINGE.Extrema(f)

def exercise_10_3_Q():
  var('x,y')
  f(x,y)=exp(x)*sin(y)
  print outilsINGE.Extrema(f)
```

tex/sage/exo103.sage

Des réponses :

(1)
```
The function :
(x, y) |--> x^2 + y^2 + 4*x - 2*y
Derivative x and y :
(x, y) |--> 2*x + 4
(x, y) |--> 2*y - 2
Hessian matrix :
[2 0]
[0 2]
Critical points :
(-2, 1)
At (-2,1), the Hessian is
   (2, 0)
   (0, 2)
   Primary principal minors are [4, 2]
   Hessian positive defined
   local minimum
```

(2)
```
The function :
(x, y) |--> e^(x^2 + x*y)
Derivative x and y :
(x, y) |--> (2*x*e^(x^2) + y*e^(x^2))*e^(x*y)
(x, y) |--> x*e^(x^2 + x*y)
Hessian matrix :
[(4*x*y*e^(x^2) + y^2*e^(x^2) + 2*(2*x^2 + 1)*e^(x^2))*e^(x*y)                    (x*y*e
```

```
[                  (x*y*e^(x^2) + (2*x^2 + 1)*e^(x^2))*e^(x*y)
Critical points :
(0, 0)
At (0,0), the Hessian is
   (2, 1)
   (1, 0)
   Primary principal minors are [-1, 2]
   Undefinite Hessian
   «selle» point
```

(3)
```
The function :
(x, y) |--> e^x*sin(y)
Derivative x and y :
(x, y) |--> e^x*sin(y)
(x, y) |--> e^x*cos(y)
Hessian matrix :
[ e^x*sin(y)  e^x*cos(y)]
[ e^x*cos(y) -e^x*sin(y)]
Critical points :
   I'm not able to solve these equations.
   I'm not able to solve these equations.
At (   ,I), the Hessian is
   I'm not able to solve these equations.
At (   ,I), the Hessian is
   I'm not able to solve these equations.
```

Ici, Sage n'est pas capable de résoudre les équations qui annulent le jacobien. Les équations à résoudre sont pourtant faciles :

$$\begin{cases} e^x \cos(y) = 0 & \text{(43.5a)} \\ e^x \sin(y) = 0 & \text{(43.5b)} \end{cases}$$

Étant donné que l'exponentielle ne s'annule jamais, il faudrait avoir en même temps $\cos(y) = 0$ et $\sin(y) = 0$, ce qui est impossible. La fonction n'a donc aucun extrémums local.

△

exEEHPooKDxLTJ

Exemple 43.9.
Considérons la fonction

$$f(x, y) = xy^2 e^{-(x^2+y^2)/4}. \tag{43.6}$$

(1) Montrer qu'il y a une infinité de points critiques.

(2) Déterminer leur nature.

Voici la fonction Sage qui fournit les informations :

```sage
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_4():
  var('x,y')
  f(x,y)=x*y**2*exp(-(x**2+y**2)/4)
  print outilsINGE.Extrema(f)
```

tex/sage/exo104.sage

La sortie est

```
The function :
(x, y) |--> x*y^2*e^(-1/4*x^2 - 1/4*y^2)
Derivative x and y :
(x, y) |--> -1/2*(x^2 - 2)*y^2*e^(-1/4*x^2 - 1/4*y^2)
(x, y) |--> -1/2*(x*y^3 - 4*x*y)*e^(-1/4*x^2 - 1/4*y^2)
Hessian matrix :
[                1/4*(x^3 - 6*x)*y^2*e^(-1/4*x^2 - 1/4*y^2) 1/4*((x^2 - 2)*y^3 - 4*(x^2 - 2
[1/4*((x^2 - 2)*y^3 - 4*(x^2 - 2)*y)*e^(-1/4*x^2 - 1/4*y^2)        1/4*(x*y^4 - 10*x*y^2 + 
Critical points :
(r17, 0)
(-sqrt(2), -2)
(sqrt(2), -2)
(-sqrt(2), 2)
(sqrt(2), 2)
At (r17,0), the Hessian is
   (0, 0)
   (0, 2*r17*e^(-1/4*r17^2))
   Primary principal minors are [0, 0]
   Hessian positive semidéfinite
   I don't conclude
   Hessian negative semidefinite
   I don't conclude
At (-sqrt(2),-2), the Hessian is
   (4*sqrt(2)*e^(-3/2), 0)
   (0, 4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), 4*sqrt(2)*e^(-3/2)]
   Hessian positive defined
   local minimum
At (sqrt(2),-2), the Hessian is
   (-4*sqrt(2)*e^(-3/2), 0)
   (0, -4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), -4*sqrt(2)*e^(-3/2)]
   Hessian negative defined
   local maximum
At (-sqrt(2),2), the Hessian is
   (4*sqrt(2)*e^(-3/2), 0)
   (0, 4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), 4*sqrt(2)*e^(-3/2)]
   Hessian positive defined
   local minimum
At (sqrt(2),2), the Hessian is
   (-4*sqrt(2)*e^(-3/2), 0)
   (0, -4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), -4*sqrt(2)*e^(-3/2)]
   Hessian negative defined
   local maximum
```

Notez la présence de `r1` comme paramètres dans les solutions. Tous les points avec $y = 0$ sont des points critiques. Cependant, Sage[1] ne parvient pas à conclure la nature de ces points $(x, 0)$.

Notons que le nombre $f(x, y)$ a toujours le signe de x parce que y^2 et l'exponentielle sont positives. Toujours ? En tout cas lorsque $x \neq 0$. Prenons un point $(a, 0)$ avec $a > 0$. Dans un

1. ou, plus précisément, le programme que j'ai écrit avec Sage.

voisinage de ce point, nous avons $f(x,y) > 0$ parce que si $a > 0$, alors $x > 0$ dans un voisinage de a. Le point $(a,0)$ est un minimum local parce que $0 = f(a,0) \leqslant f(x,y)$ pour tout (x,y) dans un voisinage de $(a,0)$.

De la même façon, les points $(a,0)$ avec $a < 0$ sont des maximums locaux parce que dans un voisinage, la fonction est négative.

Le point $(0,0)$ n'est ni maximum ni minimum local. C'est un point de selle.

△

exRNZKooUIOfPU

Exemple 43.10.
Dériver les fonctions suivantes.

(1) $\sin\big(\ln(x)\big)$
(2) $\dfrac{\sin x}{x}$;
(3) e^{x^2}
(4) $\cos(x)^{\sin(x)}$

Le programme suivant par Sage résout l'exercice :

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

var('x')
f=sin(ln(x))
print f.diff(x)
f=sin(x)/x
print f.diff(x)
f=exp(x**2)
print f.diff(x)
f=cos(x)**(sin(x))
print f.diff(x)
```

tex/sage/corrDerive_0002.sage

Le résultat est :

```
cos(log(x))/x
cos(x)/x - sin(x)/x^2
2*x*e^(x^2)
(log(cos(x))*cos(x) - sin(x)^2/cos(x))*cos(x)^sin(x)
```

△

exLFYFooNCXCJz

Exemple 43.11.
Donner une approximation de ln(1.0001).

```
----------------------------------------------------------------------
| Sage Version 4.5.3, Release Date: 2010-09-04                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: numerical_approx(ln(1.0001))
0.0000999950003332973
```

△

Chapitre 44

Épilogue : la constante de Weiner

Nous voici à la fin du Frido. Nous avons étudié beaucoup de math, et beaucoup reste à voir. En guise de conclusion, je voudrais vous parler de la constante de Weiner, introduite dans [808]. Il s'agit d'une constante, qui comme π ou e intervient dans à peu près tous les domaines de la mathématique.

Comme toujours, il existe énormément de définitions équivalentes différentes ; nous choisissons celle-ci, motivée par le théorème de Weinersmith 27.46.

DEFooXVXSooVJDTPy

Définition 44.1.
*La **constante de Weiner** W_c est l'unique réel p tel que l'espace $L^p(\mathbb{R}^7)$ soit un espace de Hilbert.*

Cette constante intervient de façon centrale dans de nombreux résultats dans tous les domaines ; nous en citons quelques-uns.

44.0.1 Résultats prouvés dans le Frido

(1) La moyenne de tout couple de réels peut être calculée en divisant leur somme par la constante de Weiner [1].

(2) La constante de Weiner donne l'indice du groupe alterné dans le groupe symétrique pour tous les ordres non triviaux, théorème 5.45.

(3) La constante de Weiner donne une borne inférieure optimale pour l'ensemble des nombres premiers.

(4) La constante de Weiner est à la fois l'unique point fixe non trivial de la fonction factorielle et le nombre de points fixes de la même fonction.

(5) Tout automorphisme d'anneau a un polynôme minimal dont le degré est donné par la constante de Weiner.

(6) Le théorème de Jordan 27.215 affirme que le nombre de composantes connexes du complémentaire d'une courbe de Jordan dans $\mathbb{R}^2$ est donné par la constante de Weiner.

(7) L'égalité $ab = 0$ dans un anneau n'implique pas spécialement $a = 0$ ou $b = 0$ lorsque la caractéristique de l'anneau est égale à la constante de Weiner, et seulement dans ce cas.

(8) Pour l'anecdote, la constante de Weiner donne le rapport τ/π ; elle est aussi la partie entière de e.

(9) C'est le degré maximal d'un polynôme irréductible sur $\mathbb{R}$, et le degré de la clôture algébrique de $\mathbb{R}$, voir les propositions 12.82 et 6.138.

1. C'est historiquement la première propriété énoncée de la constante de Weiner ; elle suggère également une notion de constante de Weiner généralisée pour moyenner un nombre arbitraire de nombres. La construction des nombres de Weiner généralisés est en projet dans la section 1.3.

44.0.2 Autres résultats

Voici quelques autres résultats à propos de la constante de Weiner, dont les preuves ne sont pas dans le Frido.

(1) La constante de Weiner est le plus petit n tel que tout monoïde libre sur un nombre fini ou même dénombrable de générateurs soit un sous-monoïde du monoïde libre à n générateurs. Les conséquences pour l'informatique sont laissées en exercice[809].

(2) C'est le nombre de valeurs possibles pour la probabilité d'un événement dépendant asymptotiquement d'une suite de variables aléatoires indépendantes[176].

(3) Comme chacun sait, l'équation diophantienne $a^n + b^n = c^n$ n'a de solutions non-triviales que pour n inférieur à la constante de Weiner[176].

(4) Un lien entre π et la constante de Weiner : le minimum de π_p est atteint lorsque p vaut la constante de Weiner[810].

(5) Pour chaque $0 < p < \infty$ nous notons π_p le rapport entre le diamètre du cercle (pour la norme p dans $\mathbb{R}^2$) et sa longueur. La valeur de p qui minimise π_p est la constante de Weiner[810].

(6) La constante de Wiener est l'unique nombre premier à être hautement composé[811].

44.0.3 Conjectures

Il reste encore de nombreuses conjectures mettant en valeur la constante de Weiner :

(1) La fameuse droite critique de la conjecture de Riemann est donnée par l'inverse de la constante de Weiner.

(2) Soit $\mathcal{P}$ l'ensemble de nombres premiers. Est-ce que l'ensemble

$$\{(p, q) \in \mathcal{P} \times \mathcal{P} \text{ tel que } 0 < |p - q| = W_c\} \tag{44.1}$$

est fini ?

D'aucuns pourraient objecter que tout cela n'est que fantaisie et trivialité. Il n'en est rien. La preuve que la constante de Weiner est centrale en mathématique est précisément qu'elle avait déjà un nom et un symbole réservé bien avant le début de l'histoire des mathématiques.

Le fait est que toutes les mathématiques que vous connaissez se basent sur les nombres entiers ; cela n'est pas du tout une trivialité.

Chapitre 45

Développements possibles

Nous donnons ici quelques idées de développements associés aux leçons mises à jour en 2021. Parfois, il est bon d'ajouter quelques lemmes au développement proposé, si il est trop court. Si l'un ou l'autre ne vous semble pas adapté à l'énoncé de la leçon, faites le moi savoir.

45.1 Algèbre et géométrie

Exemples d'équations en arithmétique.

PGCD et PPCM, algorithmes de calcul. Applications.

Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications.

Endomorphismes diagonalisables en dimension finie.

Formes linéaires et dualité en dimension finie. Exemples et applications.

Distances et isométries d'un espace affine euclidien.

Formes quadratiques réelles. Coniques. Exemples et applications.

Exemples d'utilisation de techniques d'algèbre en géométrie.

— Partitions d'un entier en parts fixées, théorème 26.99.

— Théorème de Sophie Germain, théorème 6.19.

Groupes abéliens et non abéliens finis. Exemples et applications.

— Structure des groupes d'ordre pq, théorème 5.25.

— Les groupes abéliens finis, théorème 5.23.

Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications.

— Le groupe alterné A_n est simple, théorème 5.51.

— Structure des groupes d'ordre pq, théorème 5.25.

Méthodes combinatoires, problèmes de dénombrement.

— Coloriage de roulette (18.10.15.1) et composition de colliers (18.10.15.2).

— Nombres de Bell, théorème 15.170.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

Barycentres dans un espace affine réel de dimension finie, convexité. Applications.

— Théorème de Carathéodory 11.41.

— Points extrémaux de la boule unité dans $\mathcal{L}(E)$, théorème 13.37.

— Enveloppe convexe du groupe orthogonal 13.38.

Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.

— Le lemme au lemme de Morse, lemme 17.118, voir le lemme de Morse lui-même 20.198.

— Connexité des formes quadratiques de signature donnée, proposition 17.117.

— Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 13.40 ou proposition 13.41.

— Ellipsoïde de John-Loewner, proposition 17.122.

Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.

— Algorithme des facteurs invariants 4.114.

— Méthode du gradient à pas optimal 17.114.

Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).

— Décomposition polaire 13.32.

— Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 13.40 ou proposition 13.41.

— Théorème 8.227 sur la diagonalisation de matrices symétriques.

Matrices symétriques réelles, matrices hermitiennes.

— Racine carrée d'une matrice hermitienne positive, proposition 13.27.

— Le lemme au lemme de Morse, lemme 17.118.

— Connexité des formes quadratiques de signature donnée, proposition 17.117.

— Théorème 8.227 sur la diagonalisation de matrices symétriques.

Endomorphismes trigonalisables. Endomorphismes nilpotents.

— Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 8.313.

— Décomposition de Dunford, théorème 8.265.

— Théorème de Lie-Kolchin 12.106.

Exponentielle de matrices. Applications.

— Décomposition de Dunford, théorème 8.265.

— Théorème de Von Neumann 17.61.

Endomorphismes diagonalisables en dimension finie.

— Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.

— Racine carrée d'une matrice hermitienne positive, proposition 13.27, parce qu'un utilise le résultat de diagonalisation simultanée.

— Équation de Hill $y'' + qy = 0$, proposition 32.48.

— Décomposition de Dunford, théorème 8.265.

— Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 8.301.

Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.

— Équation de Hill $y'' + qy = 0$, proposition 32.48.

— Décomposition de Dunford, théorème 8.265.

— Théorème de Lie-Kolchin 12.106.

Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.

— Racine carrée d'une matrice hermitienne positive, proposition 13.27.

— Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.

— Décomposition de Dunford, théorème 8.265.

— Algorithme des facteurs invariants 4.114.

Déterminant. Exemples et applications.

— Forme alternées de degré maximum, proposition 8.9, parce que c'est ce théorème qui donne l'unicité du déterminant du fait que l'espace est de dimension un.

— Théorème de Rothstein-Trager 20.97 parce que le résultant en est un.

— Ellipsoïde de John-Loewner, proposition 17.122.

Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.

— Forme alternées de degré maximum, proposition 8.9, parce que c'est ce théorème qui donne l'unicité du déterminant du fait que l'espace est de dimension un.

— Théorème de la dimension 4.14.

— Extrema liés, théorème 17.73.

— Théorème 8.227 sur la diagonalisation de matrices symétriques.

— Stabilité du rang par extension des scalaires, proposition 8.284.

Exemples d'actions de groupes sur les espaces de matrices.

— Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96.

— Lemme de Morse, lemme 20.198.

— Sous-groupes compacts de $\mathrm{GL}(n,\mathbb{R})$, lemme 13.40 ou proposition 13.41.

— Algorithme des facteurs invariants 4.114.

Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.

— Irréductibilité des polynômes cyclotomiques, proposition 19.23.

— Polynômes irréductibles sur $\mathbb{F}_q$.

Extensions de corps. Exemples et applications.

— Polynômes séparables, proposition 6.163.
— Lien entre les racines (multiples) de P et P', proposition 6.163.
— Théorème de l'élément primitif 6.172.
— À propos d'extensions de $\mathbb{Q}$, le lemme 6.183.
— Polynômes irréductibles sur $\mathbb{F}_q$.
— Polygones réguliers constructibles, théorème de Gauss-Wantzel, 19.87.

Corps finis. Applications.

— Théorème de Chevalley-Warning 19.43.
— Loi de réciprocité quadratique 19.39.
— $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.32.
— Polynômes irréductibles sur $\mathbb{F}_q$.

Nombres premiers. Applications.

— Structure des groupes d'ordre pq, théorème 5.25.
— Divergence de la somme des inverses des nombres premiers, théorème 15.120.
— RSA, section 19.2, plus l'exponentielle rapide, plus la recherche de couples de Bézout.
— Forme faible du théorème de Dirichlet (avec ses deux lemmes) 19.28.
— $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.32, peut-être redondant avec les groupes d'ordre pq.
— Irréductibilité des polynômes cyclotomiques, proposition 19.23.
— Théorème des deux carrés, théorème 6.25.

Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.

— RSA, section 19.2, plus l'exponentielle rapide, plus la recherche de couples de Bézout.
— Forme faible du théorème de Dirichlet (avec ses deux lemmes) 19.28.
— $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.32.
— Groupes d'ordre pq, théorème 5.25.
— Irréductibilité des polynômes cyclotomiques, proposition 19.23.

Exemples de parties génératrices d'un groupe. Applications.

— RSA, section 19.2. Assez indirect : la système RSA se base sur la formule $\varphi(pq) = (p-1)(q-1)$, laquelle se base sur l'isomorphisme $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \simeq \mathbb{Z}/pq\mathbb{Z}$ et leurs générateurs.
— Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96, parce que c'est avec lui qu'on montre les générateurs du groupe modulaire dans le corolaire 23.97.
— Générateurs du groupe diédral, proposition 18.175
— Table des caractères du groupe diédral, section 18.17.
— Le groupe alterné est simple, théorème 5.51.
— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Représentations et caractères d'un groupe fini sur un $\mathbb{C}$-espace vectoriel. Exemples

— Table des caractères du groupe diédral, section 18.17.
— Table des caractères du groupe symétrique S_4, section 16.5.

Groupe linéaire d'un espace vectoriel de dimension finie E , sous-groupes de GL(E). Applications.

- Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.
- Décomposition de Bruhat, théorème 13.39.
- Le lemme au lemme de Morse, lemme 17.118.
- Décomposition polaire 13.32.
- Enveloppe convexe du groupe orthogonal 13.38.
- Sous-groupes compacts de $\mathrm{GL}(n,\mathbb{R})$, lemme 13.40 ou proposition 13.41.
- Théorème de Von Neumann 17.61.

Groupe des permutations d'un ensemble fini. Applications.

- RSA, section 19.2, plus l'exponentielle rapide, plus la recherche de couples de Bézout.
- Coloriage de roulette (18.10.15.1) et composition de colliers (18.10.15.2).
- Forme alternées de degré maximum, proposition 8.9.
- Décomposition de Bruhat, théorème 13.39.
- Polynômes semi-symétriques, proposition 8.20.
- Table des caractères du groupe diédral, section 18.17.
- Table des caractères du groupe symétrique S_4, section 16.5.
- Isométries du cube, section 5.7.
- Le groupe alterné est simple, théorème 5.51.

Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.

- Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.
- Forme faible du théorème de Dirichlet (avec ses deux lemmes) 19.28 (parce qu'on parle de polynômes cyclotomiques qui sont basés sur les racines de l'unité).
- Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96, parce qu'on y utilise un peu les propriétés des nombres du type $|z| = 1$.
- Générateurs du groupe diédral, proposition 18.175.
- Irréductibilité des polynômes cyclotomiques, proposition 19.23.
- Théorème de Wedderburn 19.29.

Groupe opérant sur un ensemble. Exemples et applications.

- Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.
- Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96.
- Polynômes semi-symétriques, proposition 8.20.
- Lemme de Morse, lemme 20.198.
- Générateurs du groupe diédral, proposition 18.175.
- Sous-groupes compacts de $\mathrm{GL}(n,\mathbb{R})$, lemme 13.40 ou proposition 13.41.
- Théorème de Wedderburn 19.29.
- Isométries du cube, section 5.7.
- Algorithme des facteurs invariants 4.114.

45.2 Analyse

Extremums : existence, caractérisation, recherche. Exemples et applications.

Approximation d'une fonction par des fonctions régulières. Exemples et applications.

Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.

Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.

Équations différentielles ordinaires. Exemple de résolution et d'études de solutions en dimension 1 et 2.

Exemples d'équations aux dérivées partielles linéaires

Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.

Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.

Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

Analyse numérique matricielle : résolution approchée de systèmes linéaires, recherche de vecteurs propres, exemples.

Fonctions et espaces de fonctions Lebesgue-intégrables.

Problèmes d'interversion de limites et d'intégrales.

Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.

Fonctions d'une variable complexe. Exemples et applications

Loi d'une variable aléatoire : caractérisations, exemples, applications.

Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications.

Variables aléatoires discrètes. Exemples et applications.

Exemples d'études et d'applications de fonctions usuelles et spéciales.

Illustration de la notion d'indépendance en probabilité.

Exemples d'utilisation de courbes en dimension 2 ou supérieure.

Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.

- Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.
- Les théorèmes sur les fonctions définies par des intégrales, section 17.4.
- Lemme de Morse, lemme 20.198.
- Prolongement méromorphe de la fonction Γ d'Euler.

Utilisation de la notion de convexité en analyse.

— Ellipsoïde de John-Loewner, proposition 17.122.

— Peut-être la méthode de Newton, théorème 34.63, mais je ne sais pas très bien pourquoi.

Transformation de Fourier. Applications.

— Formule sommatoire de Poisson, proposition 29.15.

— Équation de Schrödinger, théorème 32.51.

Séries de Fourier. Exemples et applications.

— Formule sommatoire de Poisson, proposition 29.15.

— Inégalité isopérimétrique, théorème 28.23.

— Fonction continue et périodique dont la série de Fourier ne converge pas, proposition 28.21.

Suites et séries de fonctions. Exemples et contre-exemples.

— Formule sommatoire de Poisson, proposition 29.15.

— Théorème taubérien de Hardy-Littlewood 17.6.

— Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.

— La proposition 27.18 qui donne des indications sur la notion de classes dans L^p.

— Théorème de Montel 27.246.

— Prolongement méromorphe de la fonction Γ d'Euler.

Fonctions monotones. Fonctions convexes. Exemples et applications.

— La proposition 32.46 donne un résultat sur $y'' + qy = 0$ à partir d'une hypothèse de croissance.

— L'inégalité de Jensen, proposition 36.86.

— Méthode de Newton, théorème 34.63, si on parvient à expliquer quelle est le lien entre la méthode de Newton et la convexité.

— Ellipsoïde de John-Loewner, proposition 17.122.

— Extrema liés, théorème 17.73.

— Théorème d'inversion locale, théorème 17.47.

— Lemme de Morse, lemme 20.198.

Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.

— Espace de Sobolev $H^1(I)$, théorème 31.6.

— Inégalité isopérimétrique, théorème 28.23.

— Dual de $L^p\big([0,1]\big)$ pour $1 < p < 2$, proposition 27.184.

Espaces de fonctions : exemples et applications.

— Théorème de Fischer-Riesz 27.45.

— Espace de Sobolev $H^1(I)$, théorème 31.6.

— Théorème de Cauchy-Lipschitz 17.39.

— Dual de $L^p\big([0,1]\big)$ pour $1 < p < 2$, proposition 27.184.

Utilisation de la notion de compacité.

— Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.

— Suite telle que $\lim_{k\to\infty} d(u_{k+1}, u_k) = 0$, théorème 7.304.

— Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 13.40 ou proposition 13.41.

— Théorème de Montel 27.246.

— Ellipsoïde de John-Loewner, proposition 17.122.

Connexité. Exemples et applications.

— Théorème de Runge 26.70.

— Suite telle que $\lim_{k\to\infty} d(u_{k+1}, u_k) = 0$, théorème 7.304.

— Théorème de Brouwer en dimension 2 via l'homotopie 26.66.

— Théorème de Lie-Kolchin 12.106.

Espaces complets. Exemples et applications.

— La proposition 27.18 qui donne des indications sur la notion de classes dans L^p.

— Prolongement de fonction définie sur une partie dense, théorème 17.128

— Complétion d'un espace métrique, théorème 17.135.

— Théorème de Fischer-Riesz 27.45.

— Théorème de Cauchy-Lipschitz global 17.40.

Prolongement de fonctions. Exemples et applications.

— Prolongement de fonction définie sur une partie dense, théorème 17.128

— Lemme de Borel 15.168.

— Prolongement méromorphe de la fonction Γ d'Euler.

— Théorème de Tietze 27.178.

Espaces vectoriels normés, applications linéaires continues. Exemples.

— Théorème de Fischer-Riesz 27.45.

— Théorème de Banach-Steinhaus 10.152.

— Dual de $L^p\big([0,1]\big)$ pour $1 < p < 2$, proposition 27.184.

45.3 Anciennes leçons

Transformation de Fourier. Applications.

— Théorème de Cauchy-Lipschitz global 17.40.

Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.

— Équation de Hill $y'' + qy = 0$, proposition 32.48.

— Théorème de stabilité de Lyapunov 32.43.

— Le système proie-prédateur de Lotka-Volterra 32.44

— Théorème de Cauchy-Lipschitz global 17.40, si on parvient à réexprimer le théorème dans le cas linéaire.

Exemples de parties denses et applications.

— Prolongement de fonction définie sur une partie dense, théorème 17.128
— Complétion d'un espace métrique, théorème 17.135.
— Points extrémaux de la boule unité dans $\mathcal{L}(E)$, théorème 13.37.
— Critère de Weyl, proposition 28.10.
— Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 36.185.
— Enveloppe convexe du groupe orthogonal 13.38.

Applications des formules de Taylor.

— Méthode de Newton, théorème 34.63
— Lemme de Morse, lemme 20.198.

Continuité et dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications.

— Divergence de la somme des inverses des nombres premiers, théorème 15.120.
— Formule sommatoire de Poisson, proposition 29.15.
— Théorème taubérien de Hardy-Littlewood 17.6.
— Nombres de Bell, théorème 15.170.
— Partitions d'un entier en parts fixes, proposition 26.99.
— Théorème d'Abel angulaire 20.88.

Espaces L^p, $1 \leqslant p \leqslant \infty$

— La proposition 27.18 qui donne des indications sur la notion de classes dans L^p.
— Théorème de Fischer-Riesz 27.45.
— Espace de Sobolev $H^1(I)$, théorème 31.6.
— Dual de $L^p\big([0,1]\big)$ pour $1 < p < 2$, proposition 27.184.

Convergence des séries entières, propriétés de la somme. Exemples et applications.

— Processus de Galton-Watson, théorème 38.59.
— Formule sommatoire de Poisson, proposition 29.15.
— Nombres de Bell, théorème 15.170.
— Théorème d'Abel angulaire 20.88.

Applications des nombres complexes à la géométrie.

— Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96.
— Générateurs du groupe diédral, proposition 18.175
— Le groupe circulaire, proposition 23.95.

Sous-groupes discrets de $\mathbb{R}^2$. Réseaux. Exemples

— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Isométries d'un espace affine euclidien de dimension finie. Applications en dimensions 2 et 3.

— Isométries du cube, section 5.7.

Exemples d'équations diophantiennes.

— Dans 3.2.6, nous résolvons $ax + by = c$ en utilisant Bézout (théorème 1.236).
— L'exemple 3.92 résout l'équation $x^2 + 2 = y^3$ en parlant de l'extension $\mathbb{Z}[i\sqrt{2}]$ et de stathme.
— Les propositions 3.96 et 3.98 parlent de triplets pythagoriciens.
— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

Exemples de sous-groupes distingués et de groupes quotients. Applications.

— Suites de décomposition et théorème de Jordan-Hölder 2.21.
— Groupes d'ordre pq, théorème 5.25.
— Le groupe alterné est simple, théorème 5.51.
— Théorème de Lie-Kolchin 12.106.

Sous-groupes finis de $\mathrm{O}(2,\mathbb{R})$ et $\mathrm{O}(3,\mathbb{R})$. Applications

— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Groupes finis. Exemples et applications.

— RSA, section 19.2, plus l'exponentielle rapide, plus la recherche de couples de Bézout.
— Théorème de Wedderburn 19.29.
— Théorème de Sylow 5.11. Tout le théorème, c'est un peu long. On peut se contenter de la partie qui dit que G contient un p-Sylow.
— Coloriage de roulette (18.10.15.1) et composition de colliers (18.10.15.2).
— Suites de décomposition et théorème de Jordan-Hölder 2.21.
— Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n,\mathbb{C})$, théorème 8.313.
— $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.32.
— Groupes d'ordre pq, théorème 5.25.
— Générateurs du groupe diédral, proposition 18.175.
— Le groupe alterné est simple, théorème 5.51.
— Isométries du cube, section 5.7.

Angles : Définitions et utilisation en géométrie

— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Isométries d'un espace affine euclidien de dimension finie. Formes réduites. Applications

— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Applications affines

— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Utilisation des groupes en géométrie.

— Coloriage de roulette (18.10.15.1) et composition de colliers (18.10.15.2).
— Forme alternées de degré maximum, proposition 8.9, parce que c'est ce théorème qui donne l'unicité du déterminant du fait que l'espace est de dimension un.
— Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96.
— Générateurs du groupe diédral, proposition 18.175
— Isométries du cube, section 5.7.
— Les groupes de pavage de $\mathbb{R}^2$, théorème 18.219.

Opérations élémentaires sur les lignes et les colonnes d'une matrice. Exemples et applications.

— Décomposition de Bruhat, théorème 13.39.

— Algorithme des facteurs invariants 4.114.

Exemples de décompositions remarquables dans le groupe linéaire. Applications

— Décomposition polaire 13.32.

— Décomposition de Dunford, théorème 8.265.

Résultant. Applications.

— Théorème de Rothstein-Trager 20.97.

— Théorème de Kronecker 8.28.

Matrices équivalentes. Matrices semblables. Applications.

— Racine carrée d'une matrice hermitienne positive, proposition 13.27.

— Sous-groupes compacts de $\mathrm{GL}(n,\mathbb{R})$, lemme 13.40 ou proposition 13.41.

Exemples d'utilisation de la notion de dimension d'un espace vectoriel.

— Forme alternées de degré maximum, proposition 8.9, parce que c'est ce théorème qui donne l'unicité du déterminant du fait que l'espace est de dimension un.

— Théorème de la dimension 4.14, bien que ce soit plutôt dans la définition de la dimension que dans l'utilisation.

— Théorème de Carathéodory 11.41.

Corps des fractions rationnelles à une indéterminée sur un corps commutatif. Applications.

— Théorème de Rothstein-Trager 20.97.

— Partitions d'un entier en parts fixes, proposition 26.99.

Anneau de séries formelles. Applications.

— Nombres de Bell, théorème 15.170.

— Partitions d'un entier en parts fixes, proposition 26.99.

Formes linéaires et hyperplans en dimension finie. Exemples et applications.

— Extrema liés, théorème 17.73.

— Enveloppe convexe du groupe orthogonal 13.38.

— Une forme canonique pour les transvections et dilatations, théorème 13.8.

Algèbre des polynômes d'un endomorphisme en dimension finie. Applications.

— Racine carrée d'une matrice hermitienne positive, proposition 13.27.

Formes quadratiques réelles. Exemples et applications.

— Le lemme au lemme de Morse, lemme 17.118.

— Connexité des formes quadratiques de signature donnée, proposition 17.117.

— Sous-groupes compacts de $\mathrm{GL}(n,\mathbb{R})$, lemme 13.40 ou proposition 13.41.

— Ellipsoïde de John-Loewner, proposition 17.122.

Polynômes d'endomorphisme en dimension finie. Applications à la réduction d'un endomorphisme en dimension finie.

— Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 8.301.

Applications des nombres complexes à la géométrie. Homographies.

— Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.96. Parce que l'action est avec des homographies.

Anneaux principaux. Applications

— Polynôme minimal d'endomorphisme semi-simple, théorème 8.110.

— Théorème de Bézout, corolaire 3.73.

— Théorème des deux carrés, théorème 6.25.

— Algorithme des facteurs invariants 4.114.

Représentations de groupes finis de petit cardinal.

— Table des caractères du groupe diédral, section 18.17.

— Table des caractères du groupe symétrique S_4, section 16.5.

Algèbre des polynômes à n indéterminées ($n \geqslant 2$). Polynômes symétriques. Applications.

— À propos d'extensions de $\mathbb{Q}$, le lemme 6.183.

— Polynômes semi-symétriques, proposition 8.20.

— Théorème de Chevalley-Warning 19.43.

— Théorème de Kronecker 8.28.

Racines d'un polynômes. Fonctions symétriques élémentaires. Localisation des racines dans les cas réel et complexe.

— À propos d'extensions de $\mathbb{Q}$, le lemme 6.183.

135 - Isométries d'un espace affine euclidien de dimension finie. Forme réduite. Applications en dimensions 2 et 3.

— Points extrémaux de la boule unité dans $\mathcal{L}(E)$, théorème 13.37.

— Générateurs du groupe diédral, proposition 18.175

— Isométries du cube, section 5.7.

248 - Approximation des fonctions numériques par des fonctions polynomiales. Exemples.

— Théorème taubérien de Hardy-Littlewood 17.6.

— Théorème de Runge 26.70.

— Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 36.185.

250 - Loi des grands nombres. Théorème central limite. Applications.

— Presque tous les nombres sont normaux, proposition 36.183.

— Estimation des grands écarts, théorème 36.169.

251 - Indépendance d'événements et de variables aléatoires. Exemples.

— Presque tous les nombres sont normaux, proposition 36.183.

— Estimation des grands écarts, théorème 36.169.

— Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 36.185.

— Problème de la ruine du joueur, section 39.4.

252 - Loi binomiale. Loi de Poisson. Applications.

— Estimation des grands écarts, théorème 36.169.

— Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 36.185.

— Problème de la ruine du joueur, section 39.4.

219 - Problèmes d'extrémums.

— Extrema liés, théorème 17.73.

— Lemme de Morse, lemme 20.198.

— Ellipsoïde de John-Loewner, proposition 17.122.

Équations différentielles $X' = f(t, X)$. Exemples d'étude des solutions en dimension 1 et 2.

— Théorème de Cauchy-Lipschitz 17.39.

— Théorème de stabilité de Lyapunov 32.43.

— Équation de Hill $y'' + qy = 0$, proposition 32.48.

— Le système proie-prédateur de Lotka-Volterra 32.44

223 - Convergence des suites numériques. Exemples et applications.

— Calcul d'intégrale par suite équirépartie 28.10.

— Théorème taubérien de Hardy-Littlewood 17.6.

— Méthode de Newton, théorème 34.63

— Théorème d'Abel angulaire 20.88.

224 - Comportement asymptotique de suites numériques. Rapidité de convergence. Exemples.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

— Divergence de la somme des inverses des nombres premiers, théorème 15.120.

— Formule sommatoire de Poisson, proposition 29.15, grâce à l'exemple 29.16.

— Méthode de Newton, théorème 34.63

— Estimation des grands écarts, théorème 36.169.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

226 - Comportement d'une suite réelle ou vectorielle définie par une itération $u_{n+1} = f(u_n)$. Exemples.

— Processus de Galton-Watson, section 38.7.

— Méthode de Newton, théorème 34.63

— Méthode du gradient à pas optimal 17.114.

245 - Fonctions holomorphes et méromorphes sur un ouvert de $\mathbb{C}$. Exemples et applications.

— Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.

— Théorème de Montel 27.246.

— Prolongement méromorphe de la fonction Γ d'Euler.

238 - Méthodes de calcul approché d'intégrales et de solutions d'équations différentielles.

— Calcul d'intégrale par suite équirépartie 28.10.

222 - Exemples d'équations différentielles. Solutions exactes ou approchées.

— Équation $y'' + qy = 0$, 32.9.3.

225 - Étude locale de surfaces. Exemples.

— Lemme de Morse, lemme 20.198.

Exemples de problèmes d'interversion de limites.

— Théorème taubérien de Hardy-Littlewood 17.6 parce que l'énoncé revient à montrer la limite $\lim_{x\to 1^-} \sum_{n\in\mathbb{N}} a_n x^n = \sum_{n\in\mathbb{N}} a_n$.

— Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.

— La proposition 27.18 qui donne des indications sur la notion de classes dans L^p. Ça utilise la convergence monotone pour pour permuter une somme et une intégrale.

— Les théorèmes sur les fonctions définies par des intégrales, section 17.4.

— Nombres de Bell, théorème 15.170.

254 - Espaces de Schwartz et distributions tempérées.

— Formule sommatoire de Poisson, proposition 29.15.

— Équation de Schrödinger, théorème 32.51.

Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications.

— Extrema liés, théorème 17.73.

— Théorème d'inversion locale, théorème 17.47.

— Lemme de Morse, lemme 20.198.

— Théorème de Von Neumann 17.61.

Continuité et dérivabilité des fonctions réelles d'une variable réelle. Exemples et contre-exemples.

— Les théorèmes sur les fonctions définies par des intégrales, section 17.4.

— Lemme de Borel 15.168.

232 - Méthodes d'approximation des solutions d'une équation $F(X) = 0$. Exemples.

— Méthode de Newton, théorème 34.63

— Méthode du gradient à pas optimal 17.114.

Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables réelles.

— Calcul d'intégrale par suite équirépartie 28.10.

— Théorème de Rothstein-Trager 20.97.

256 - Transformation de Fourier dans $\mathscr{S}(\mathbb{R}^d)$ et $\mathscr{S}'(\mathbb{R}^d)$.

— Formule sommatoire de Poisson, proposition 29.15.

— Équation de Schrödinger, théorème 32.51.

Espaces de Schwartz $\mathscr{S}(\mathbb{R}^d)$ et distributions tempérées. Transformation de Fourier dans $\mathscr{S}(\mathbb{R}^d)$ et $\mathscr{S}'(\mathbb{R}^d)$

— Formule sommatoire de Poisson, proposition 29.15.

— Équation de Schrödinger, théorème 32.51.

Espaces de Schwartz. Distributions. Dérivation au sens des distributions.

— L'équation $(x-x_0)^\alpha u = 0$ pour $u \in \mathscr{D}'(\mathbb{R})$, théorème 30.51.

— Espace de Sobolev $H^1(I)$, théorème 31.6.

— Équation de Schrödinger, théorème 32.51.

Fonctions développables en série entière, fonctions analytiques. Exemples.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples, théorème 26.99.

Théorèmes de point fixe. Exemples et applications.

— Processus de Galton-Watson, théorème 38.59.

— Théorème d'inversion locale, théorème 17.47.

— Théorème de Brouwer en dimension 2 via l'homotopie 26.66.

— Théorème de Picard 17.34 et l'inséparable théorème de Cauchy-Lipschitz 17.39

— Inégalité isopérimétrique, théorème 28.23.

— Théorème des quatre sommets, théorème 21.106.

— Formule sommatoire de Poisson, proposition 29.15.

Sous-variétés de $\mathbb{R}^n$. Exemples.

— Extrema liés, théorème 17.73.

— Théorème de Von Neumann 17.61.

— Lemme de Morse, lemme 20.198.

Suites et séries de fonctions intégrables. Exemples et applications.

— La proposition 27.18 qui donne des indications sur la notion de classes dans L^p.

— Le théorème de Weierstrass sur la limite uniforme de fonctions holomorphes, théorème 26.106.

— Les théorèmes sur les fonctions définies par des intégrales, section 17.4.

— Théorème de Fischer-Riesz 27.45.

— Prolongement méromorphe de la fonction Γ d'Euler.

— Problème de la ruine du joueur, section 39.4.

Utilisation en probabilités du produit de convolution et de la transformation de Fourier ou de Laplace.

— Processus de Galton-Watson, lemme 38.58 et théorème 38.59.

— Fonction caractéristique 36.91.

— Théorème central limite 36.135.

Suites de variables de Bernoulli indépendantes.

— Processus de Galton-Watson, section 38.7.

— Estimation des grands écarts, théorème 36.169.

— Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 36.185.

— Problème de la ruine du joueur, section 39.4.

Part II

Giulietta

Chapter 46

Some old results

46.1 Some old questions

46.2 Still some questions and problems

(1) 83.1, page 4154 Où vient le ∇ dans la définition de WKB ?

(2) 84.4, page 4185 à voir que quand on a une courte suite exacte d'espaces qui s'injèctent les uns dans les autres, on a une longue suite exacte dans les espaces de cohomologie.

(3) 84.10, page 4189 En réalité, c'est sans doute l'avant-dernière ligne de la preuve du théorème 3.1 dans Semisimple Symplectic Symmetric spaces. Ou bien (i.2) de la proposition 4.1

(4) 79.30, page 3971 Encore le problème de θ comme automorphisme interne.

(5) 83.3, page 4155 Une référence pour l'équivalence entre Moyal et Weyl

(6) 83.20, page 4170 Un énoncé exact du produit sur $\mathrm{SL}(2,\mathbb{R})$.

(7) 83.19, page 4168 Prouver que $\mathcal{T}_\theta$ entrelace réellement ρ_ν et dL.

(8) A friend said me that the condition (85.32) means that the form is closed in the sense of Chevalley and thus its left invariant prolongation is closed in the sense of de Rahm.

(9) L'espace symétrique M est vu comme G/H où G est la composante connexe à l'identité de $Aff(M)$, c'est-à-dire le plus grand groupe connexe de transformations affines (c'est le théorème 66.27), où $Aff(M)$ =groupe des transformations affines de M.

A-t-on forcement : composante connexe à l'identite de $Aff(M)$ = larger connex group of $Aff(M)$? (genre, pourrait-on avoir un ss-groupe connexe qui ne passe pas par l'identité, plus grand que tous ceux qui passent par l'identité ?)

On sait que toutes les symétries s_x sont des transformations affines, mais le contraire n'est pas vrai ; en particulier, $Aff(M)$ peut contenir des elements qui ne sont pas des symétries. Ceci pour savoir la definition du théorème 66.27, était la même que celle du théorème 2.4 du papier de Bieliavsky, où il exprime $M = G/H$, avec cette fois G = transvection group de M, qui est INCLUS dans $Aff(M)$.

Dans ton théorème 66.27, G be the LARGER connex group of affine transformation, tandis que chez Bieliavsky (i), le transvection group G is the SMALLEST subgroup of Aff(M) which is transitive etc.

46.3 Some statements waiting for proofs

We provide here some results that are at undergrad level. Some of them are contained in the part "Pour l'agrégation" in French.

lem:Schur

Lemma 46.1 (Schur's lemma)**.**
If $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is irreducible, then the only endomorphism of V which commutes with all $\phi(\mathfrak{g})$ are multiples of identity.

Now we consider $W = \{u \in V \text{ tel que } \exists n \in \mathbb{N} : A^n u = 0\}$ and $v \in V$. There exists a $v' \in V$ such that $A^p(v) = A^{p+1}(v')$. Writing $v = A(v') + (v - A(v'))$, we find

$$V \subset A(V) + W \tag{46.1}$$

eq:VAW

because $A^p(v) - A^{p+1}(v') = 0$, $v - A(v') \in W$.

If we apply A on this, we find $A(V) \subset A^2(V) + A(W)$. Reinserting it into the right hand side of (46.1), we find $V \subset A^2(V + W)$ and repeating p times this process, we find $V = A^p(V) + W$ and the sum is direct because none of the elements of $A^p(V)$ is annihilated by A:

$$V = A^p(V) \oplus W. \tag{46.2}$$

eq:ApoplusW

ThoCayleyHamilton

Theorem 46.2 (Cayley-Hamilton)**.**
A square matrix on $\mathbb{R}$ or $\mathbb{C}$ satisfies its own characteristic equation. If $A \in \mathbb{M}_n(\mathbb{K})$, we consider the polynomial $p(\lambda) = \det(A - \lambda\mathbb{1})$. Thus $p(A) = 0$.

For a proof see Wikipedia, theorem 8.120 and theorem 13.25.

PropMtrDiagablaUnit

Proposition 46.3 ([812])**.**
If M is a complex $n \times n$ matrix, then there exists an unitary matrix U such that U^*MU is upper triangular.

Definition 46.4.
*A **positive defined** matrix is a matrix B such that*

$$\sum_{ij} B_{ij}\overline{x_i}x_j \tag{46.3}$$

is real and positive for every complex vector x.

Proposition 46.5.
A positive defined matrix is Hermitian.

Proof. We define the Hermitian matrices $M = (B + B^*)/2$ and $N = (B - B^*)/2i$, so $B = M + iN$ and

$$\bar{x}Bx = \bar{x}Mx + i\bar{x}Nx. \tag{46.4}$$

The matrices M and B being Hermitian, the numbers $\bar{x}Mx$ and $\bar{x}Nx$ are real. If $\bar{x}Bx$ has to be real, we need $\bar{x}Nx = 0$ for every x. This shows that $N = 0$, so that $B = N$. □

Theorem 46.6.
Let G be a Lie group and H a subgroup (with no special other structures) of G. If H is a closed subset of G then there exists an unique analytic structure on H such that H is a topological Lie subgroup of G.

Helgason2.3

This comes from [813], chapter 2, theorem 2.3.

Lemma 46.7.
Let G be a connected Lie group with Lie algebra $\mathcal{G}$ and let φ be an analytic homomorphism of G into a Lie group X with Lie algebra $\mathcal{X}$. Then

(1) The kernel $\varphi^{-1}(e)$ is a topological Lie subgroup of G. Its Lie algebra is the kernel of $d\varphi_e$.

(2) The image $\varphi(G)$ is a Lie subgroup of $\mathcal{X}$ with Lie algebra $d\varphi(\mathcal{G}) \subset \mathcal{X}$.

Helgason5.1

This comes from [813], chapter 2, lemma 5.1.

Lemma 46.8.
Let G and H be two Lie group, whose Lie algebra are $\mathcal{G}$ and $\mathcal{H}$. If $\theta : G \to H$ is a surjective map, then we have $\mathcal{H} \simeq \mathcal{G}/Ker\, d\theta_e$.

1203r1

1503t1

Theorem 46.9.
Let us consider $Ad\colon SU(2) \to GL(3)$, $Ad(U)X = UXU^{-1}$. *We have the following properties:*

(1) Ad is a linear homomorphism,

(2) it takes his values in $\mathrm{SO}(3)$*; then we can write* $Ad\colon SU(2) \to \mathrm{SO}(3)$,

(3) it is surjective,

(4) $Ker\, Ad = \mathbb{Z}_2$,

(5) all these properties show that

$$\mathrm{SO}(3) = \frac{SU(2)}{\mathbb{Z}_2}.$$

Definition 46.10.
If (a_k) *is a sequence in* $\mathbb{R}$, *its* ***upper limit*** *is the real number*

$$\limsup_{n\to\infty} a_n = \lim_{l\to\infty} \sup\{a_k : k \geqslant l\}.$$

Lemma 46.11.
If ω *is a* k*-form (not specially a symplectic one), and* ∇ *a torsion free connection, one has*

$$(d\omega)(X_0,\dots,X_k) = \sum_{i=0}^{k} (-1)^i (\nabla_{X_i}\omega)(X_0,\dots,\hat{X}_i,\dots X_k). \tag{46.5}$$

eq:d_omega_nabla

Remarque 46.12.
The link between d and ∇ comes from the fact that in the left hand side of (46.5) appears some commutators $[X_i, X_j]$, but since the connection is torsion-free,

$$[X_i, X_j] = \nabla_{X_i} X_j - \nabla_{X_j} X_i$$

The main consequence of this lemma is that $\nabla\omega = 0$ implies $d\omega = 0$.

prop:fdefint

Proposition 46.13.
Consider a function $f\colon X \times E \to \overline{\mathbb{R}}$ *and* $z_0 \in E$ *such that*

— *for all* $z \in E$, *the function* $x \to (x, z)$ *is integrable,*

— *for (almost) all* $x \in X$, *the function* $z \to f(x, z)$ *is continuous at* z_0,

— *there exists a function* $g \geqslant 0$ *such that for all* $z \in E$, $|f(x, z)| \leqslant g(x)$ *almost everywhere in* X.

Then the function $h\colon E \to \mathbb{R}$ *defined by* $h(z) = \int_X f(x, z)$ *is continuous at* z_0.

This proposition is (up to "almost") the theorem 17.15.

Chapter 47

Categories

chap_category

47.1 Functor

47.1.1 Functor and equivalence

For reference, see [814, 815].

If $\mathscr{C}$ and $\mathscr{D}$ are categories, a **(covariant) functor** is a map $F\colon \mathscr{C} \to \mathscr{D}$ such that

(1) if $f\colon A \to B$ is an arrow in $\mathscr{C}$, then $Ff\colon FA \to FB$ is an arrow in $\mathscr{D}$ (A and B are objects in $\mathscr{C}$),

(2) $F(\mathrm{Id}_A) = \mathrm{Id}_{FA}$,

(3) if $g\colon B \to C$, then $F(g \circ f) = Fg \circ Ff$.

A **contravariant functor** is $F\colon \mathscr{C} \to \mathscr{D}$ with $F(X) \in \mathscr{D}$ and $F(f) \in \hom\big(F(Y), F(X)\big)$ for every $X \in \mathscr{C}$ and $f \in \hom(X, Y)$ such that

PgPropFnctConvtra

(1) $F(\mathrm{Id}_X) = \mathrm{Id}_{F(X)}$,

(2) $F(g \circ f) = F(f) \circ F(g)$.

Notice that in the contravariant case, $F(f) \in \hom\big(F(Y), F(X)\big)$, and not $\hom\big(F(X), F(Y)\big)$

A functor is an **isomorphism** if it has an inverse.

When A and B are objects in the category $\mathscr{C}$, we denote by $\hom_{\mathscr{C}}(A, B)$ the set[1] of arrows from A to B. Such a set is called an **homset**. We say that a functor is **faithful** when its restriction to each homset is injective. It is **full** when it is surjective on each homset, i.e. for every $f \in \hom(FA, FB)$, there exists a $g \in \hom(A, B)$ such that $f = Fg$. The functor F between $\mathscr{C}$ and $\mathscr{D}$ is an **equivalence** when

(1) it is full,

(2) it is faithful

(3) for every object B in $\mathscr{D}$, there exists an object A in $\mathscr{C}$ such that $F(A)$ is isomorphic to B.

The notion of equivalence of categories applies on categories whose objects have a notion of isomorphism, such like category of groups, manifold or vector space.

47.2 Direct limit

SecDirectLimit

47.2.1 Direct limit of sets

Let (A_i) be a family of set labeled by an ordered set I, and functions $f_{ij}\colon A_i \to A_j$ for $i \leqslant j$ such that

— f_{ii} is the identity on A_i,

— $f_{ik} = f_{jk} \circ f_{ij}$ for every $i \leqslant j \leqslant k$.

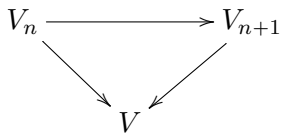
1. In general, it is actually not a set.

We say that (I, A_i, f_{ij}) is an **direct system**. The **direct limit** denoted by $A = \lim_{\rightarrow} A_i$ is defined as the quotient of the disjoint union $\bigsqcup_{i\in I} A_i$ by the equivalence relation $x_i \sim x_j$ if and only if there exists a $k \geqslant i, j$ such that $f_{ik}(x_i) = f_{jk}(x_j)$ ($x_i \in A_i$ and $x_j \in A_j$).

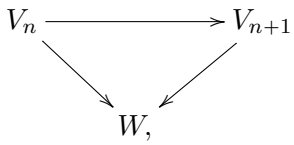
If one sees the maps f_{ij} as inclusions, equivalent elements are the ones which finish to be equal.

47.2.2 Direct limit of vector spaces

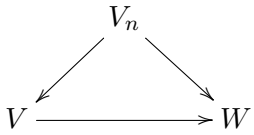
Consider a diagram $V_1 \to V_2 \to \cdots$ of vector spaces and linear maps. A *direct limit* of that system is a vector space V endowed with a collection of maps $V_n \to V$ such that the diagram



commutes. Moreover, we impose the following universal property: for every vector space W and compatible maps



there exists an unique map $V \to W$ which is compatible with everything, i.e. such that



commutes.

One can show that the direct limit exists and is unique in the category of vector spaces. If the diagram is $V_1 \subset V_2 \subset \cdots$, each of them being a vector subspace of V, then the direct limit is given by $\bigcup_n V_n$.

PgExDirectLimVS

As an example, let consider a sequence of matrix algebras M_1, M_2,... For simplicity, we suppose that these are full matrix algebras: $M_k \simeq \mathbb{M}_{n_k}(\mathbb{C})$. Put $A_n = M_1 \otimes_{\mathbb{C}} \ldots \otimes_{\mathbb{C}} M_n$; each of these A_n is a matrix algebra and we define $A_n \to A_{n+1}$,

$$T_1 \otimes \ldots \otimes T_n \mapsto T_1 \otimes \ldots \otimes T_n \otimes \mathbb{1}.$$

For instance if $M_1 = M_2 = \mathbb{M}_2(\mathbb{C})$, we have $M_1 \otimes M_2 \simeq \mathbb{M}_4(\mathbb{C})$ and $A_1 \to A_2$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a\mathbb{1}_2 & b\mathbb{1}_2 \\ c\mathbb{1}_2 & d\mathbb{1}_2 \end{pmatrix}.$$

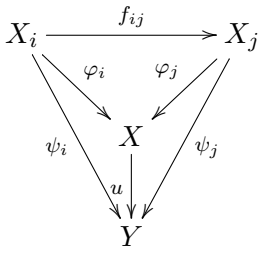
Notice that the rank is not preserved, for example

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}.$$

47.2.3 Direct limit in categories

Let now (X_i, f_{ij}) be a direct system in a category $\mathscr{C}$ (the definition is the same as in the case of sets). The direct limit of this system is an object X of $\mathscr{C}$ with arrows $\varphi_j \colon X_j \to X$ satisfying $\varphi_i = \varphi_j \circ f_{ij}$ whenever $i \leqslant j$ and which fulfils the following universal property: for every object

$Y \in \mathscr{C}$ with arrows $\psi_j \colon X_j \to X$ such that $\psi_i = \psi_j \circ f_{ij}$ (when $i \leqslant j$), there exists an arrow $u \colon Y \to X$ which makes the following diagram commutative when $i \leqslant j$



(47.1)

All the arrows here are arrows of the category $\mathscr{C}$. One can prove that the direct limit is unique in every category, but the existence is not always guarantee. In the case of the category of sets however it exists as we saw an explicit construction as quotient space.

47.3 Categories with tensor product

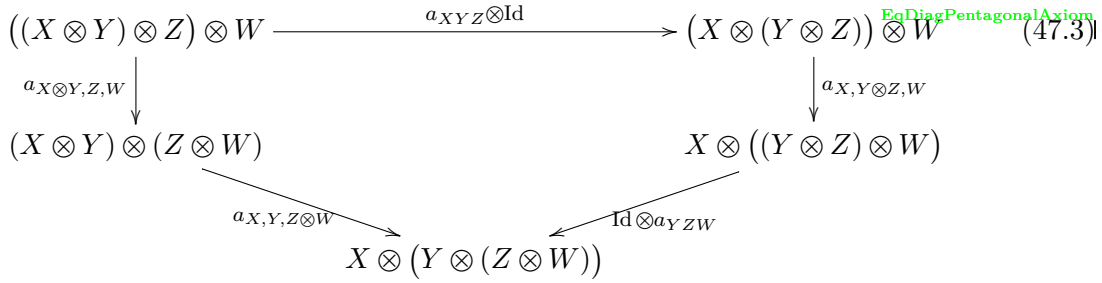
Definition 47.1 ([816]).
A category $\mathscr{C}$ with a functor $\otimes \colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ and a collection of morphisms

$$a_{XYZ} \colon (X \otimes Y) \otimes Z \to X \otimes (Y \otimes Z) \tag{47.2}$$

*for $X, Y, Z \in \operatorname{Ob} \mathscr{C}$ is a **monoidal** if*

(1) a_{XYZ} is an isomorphism for every X, Y and Z;

(2) for every $X, Y, Z, W \in \operatorname{Ob} \mathscr{C}$ the diagram



(47.3) EqDiagPentagonalAxiom

commutes;

*(3) there exists an object I named **identity object** such that $I \otimes I = I$ and the functors $X \mapsto X \otimes I$ and $X \mapsto I \otimes X$ are autoequivalences of $\mathscr{C}$.*

The commutativity of diagram (47.3) is sometimes called the **pentagonal axiom**. The transformation a is the **associativity constraint**.

Exemple 47.2.
The category of finite dimensional vector spaces over $\mathbb{K}$ with the usual tensor product is monoidal if we set

$$\begin{aligned} a_{XYZ} \colon (X \otimes Y) \otimes Z &\to X \otimes (Y \otimes Z) \\ (x \otimes y) \otimes z &\mapsto x \otimes (y \otimes z), \end{aligned} \tag{47.4}$$

i.e. the identity map. The unit object is $\mathbb{K}$ itself seen as one dimensional $\mathbb{K}$-vector space. △

A monoidal category is **strict** if $a_{XYZ} = \mathrm{Id}$ for every objects X, Y and Z and if there exists an object $\mathbb{1}$ such that

(1) $X \otimes \mathbb{1} = \mathbb{1} \otimes X = X$;

(2) the diagram

$$\begin{array}{ccc} X\otimes Y & \xrightarrow{\mathrm{Id}} & (X\otimes\mathbb{1})\otimes Y \\ {\scriptstyle\mathrm{Id}}\downarrow & & \downarrow{\scriptstyle a_{X,\mathbb{1},Y}} \\ X\otimes Y & \xrightarrow[\mathrm{Id}]{} & X\otimes(\mathbb{1}\otimes Y) \end{array} \tag{47.5}$$

commutes.

The category of finite dimensional $\mathbb{K}$-vector space becomes a strict monoidal category when we set $\mathbb{1} = I = \mathbb{K}$.

An object X in a strict monoidal category is **rigid** if there exists an object Y and a pair of morphisms

$$\begin{aligned} i_X &: \mathbb{1} \to X\otimes Y \\ e_X &: Y\otimes X \to \mathbb{1} \end{aligned} \tag{47.6}$$

such that the diagrams

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{Id}} & \mathbb{1}\otimes X \\ {\scriptstyle\mathrm{Id}}\downarrow & & \downarrow{\scriptstyle i_X\otimes\mathrm{Id}_X} \\ X & \xleftarrow[\mathrm{Id}_X\otimes e_X]{} & X\otimes Y\otimes X \end{array} \tag{47.7}$$

and

$$\begin{array}{ccc} Y & \xrightarrow{\mathrm{Id}} & Y\otimes\mathbb{1} \\ {\scriptstyle\mathrm{Id}}\downarrow & & \downarrow{\scriptstyle \mathrm{Id}_Y\otimes i_X} \\ Y & \xleftarrow[e_X\otimes\mathrm{Id}_Y]{} & Y\otimes X\otimes Y \end{array} \tag{47.8}$$

commute. In this case we say that Y is **dual** of X and we write $Y = X^*$.

Exemple 47.3.
In the case of $\mathbb{K}$-vector spaces, the object X^* is the usual dual of X. The map e_X is the evaluation

$$\begin{aligned} e_X &: X^*\otimes X \to \mathbb{K} \\ &\alpha\otimes v \mapsto \alpha(v) \end{aligned} \tag{47.9}$$

for $\alpha\in X^*$ and $v\in X$. The map i_X is

$$\begin{aligned} i_X &: \mathbb{K} \to X\otimes X^* \\ &z \mapsto z\sum_i e_i\otimes e_i^* \end{aligned} \tag{47.10}$$

where $\{e_i\}$ is a basis of X and $\{e_i^*\}$ is the dual basis (that is $e_i^*(e_j) = \delta_{ij}$). Notice that i_X is independent of the choice of the basis. Indeed let us apply $\sum_i e_i\otimes e_i^*$ to the element $\alpha\otimes v\in X^*\otimes X$. We have

$$\sum_i e_i\otimes e_i^*(\alpha\otimes v) = \sum_i e_i(\alpha)e_i^*(v) = \sum_i \alpha_i v_i \tag{47.11}$$

where α_i and v_i are the coordinates of α and v with respect to the basis $\{e_i^*\}$ and $\{e_i^*\}$. The number $\sum_i \alpha_i v_i$ is nothing else than the trace of the map $X\to X$, $w\mapsto \alpha(w)v$ and thus is independent on the choice of the basis.

These definitions satisfy the axioms of a rigid monoidal category since

$$\begin{array}{ccc} v & \longrightarrow & 1\otimes v \\ \downarrow & & \downarrow \\ \sum_i e_i\otimes e_i^*(v) & \longleftarrow & \sum_i e_i\otimes e_i^*\otimes v \end{array} \tag{47.12}$$

commutes because, identifying X with $X\otimes\mathbb{K}$, we have $\sum_i e_i^*(v)e_i = v$. The other diagram is the same and we found that it commutes because $\sum_i \alpha(e_i)e_i^* = \alpha$. △

A strict monoidal category is rigid if all the objects are rigid and if for every object X we have an object Y such that X is isomorphic to Y^*.

Exemple 47.4.
The category of finite dimensional $\mathbb{K}$-vector spaces is rigid since $X^{**} = X$. △

Let C be a strict monoidal category. A **braiding** on C is a collection $c = \{c_{X,Y}\}$ of functorial isomorphisms between the functors $(X, Y) \mapsto X \otimes Y$ and $(X, Y) \mapsto Y \otimes X$ such that

(1) $x_{X,\mathbb{1}} = c_{\mathbb{1},X} = \mathrm{Id}_X$;
(2) for every objects X, Y and Z the diagrams

$$\begin{array}{ccc} & (X\otimes Y)\otimes Z & \\ {\scriptstyle c_{X\otimes Y,Z}}\swarrow & & \searrow{\scriptstyle \mathrm{Id}} \\ Z\otimes(X\otimes Y) & & X\otimes(Y\otimes Z) \\ {\scriptstyle \mathrm{Id}}\downarrow & & \downarrow{\scriptstyle \mathrm{Id}_X\otimes c_{Y,Z}} \\ (Z\otimes X)\otimes Y & & (X\otimes Z)\otimes Y \\ {\scriptstyle \mathrm{Id}}\searrow & & \swarrow{\scriptstyle c_{x,Z}\otimes \mathrm{Id}_Y} \\ & (Z\otimes X)\otimes Y & \end{array} \tag{47.13}$$

and

$$\begin{array}{ccc} & (X\otimes Y)\otimes Z & \\ {\scriptstyle c_{X,Y\otimes Z}}\swarrow & & \searrow{\scriptstyle \mathrm{Id}} \\ (Y\otimes Z)\otimes X & & (X\otimes Y)\otimes Z \\ {\scriptstyle \mathrm{Id}}\uparrow & & \downarrow{\scriptstyle c_{X,Y}\otimes \mathrm{Id}_Z} \\ Y\otimes(Z\otimes X) & & (Y\otimes X)\otimes Z \\ {\scriptstyle \mathrm{Id}_Y\otimes c_{X,Z}}\searrow & & \swarrow{\scriptstyle \mathrm{Id}} \\ & Y\otimes(X\otimes Z) & \end{array} \tag{47.14}$$

commute.

The commutativity of these two diagrams is sometimes called the **hexagonal identities**.

Exemple 47.5.
The category of finite dimensional $\mathbb{K}$-vector spaces is a braided category with

$$\begin{aligned} c_{XY} \colon X \otimes Y &\to Y \otimes X \\ x \otimes y &\mapsto y \otimes x. \end{aligned} \tag{47.15}$$

△

Let C be a braided strict monoidal category. A **balancing** on C is a collection $b = \{b_X\}$ of automorphisms of the identity functor such that the diagram

$$\begin{array}{ccc} X\otimes Y & \xrightarrow{b_{X\otimes Y}} & X\otimes Y \\ {\scriptstyle b_X\otimes b_y}\downarrow & & \downarrow{\scriptstyle c_{XY}} \\ X\otimes Y & \xleftrightarrow[b_Y\otimes b_X]{} & Y\otimes X \end{array} \tag{47.16}$$

commutes.

In a balanced category one can define the **trace** of an arrow $f \colon X \to X$. The trace is the arrow

$$\mathrm{Tr}_X(f) \colon \mathbb{1} \to \mathbb{1} \tag{47.17}$$

given by

$$\mathbb{1} \xrightarrow{i_X} X \otimes X^* \xrightarrow{f\otimes \mathrm{Id}} X \otimes X^* \xrightarrow{b_X \otimes \mathrm{Id}_{X^*}} X \otimes X^* \xrightarrow{e_{X^*}} \mathbb{1}. \tag{47.18}$$

Exemple 47.6.
In the case of the finite dimensional $\mathbb{K}$-vector spaces, we have the balancing $b_X = \mathrm{Id}_X$. We check that the trace of a linear map $f\colon X \to X$ is the well known map:

$$1 \longrightarrow \sum_i e_i \otimes e_i^* \longrightarrow \sum_i f(e_i) \otimes e_i^* \longrightarrow \sum_i e_i^*\big(f(e_i)\big) = \sum_i f(e_i)_i. \tag{47.19}$$

△

Definition 47.7.
A rigid balanced braided monoidal category is a ***ribbon category****.*

Chapter 48

Topology

chap_topology

48.1 Main definitions

As for notations, when we have a topological space X and a point $x \in X$, we denote by $\mathcal{V}(x)$ the set of neighbourhoods of x. If F is a subset of X, we denote by $\overline{F}$ the closure of F and by $\operatorname{Int} F$, its interior.

Definition 48.1.
*Let $\mathcal{O}$ be a topology on a set X. A subset $\mathcal{B} \subset \mathcal{O}$ (the elements of $\mathcal{B}$ are subsets of X) is a **basis of topology** $\mathcal{O}$ of X when any element of $\mathcal{O}$ is an union of elements of $\mathcal{B}$.*

Definition 48.2.
*If X is a topological space, a subset $A \subset X$ is a **topological subspace** when we consider on A the **induced topology** from X. This is the topology on A where then open sets are the subsets of A which can be written as $A \cap \mathcal{O}$ for an open set $\mathcal{O} \subset X$ in the sense of the topology on X.*

Remarque 48.3.
In some literature, a topological subspace of X is just a subset A endowed with a topology for which the inclusion $\iota \colon A \to X$ is continuous. This is a less restrictive condition.

From this definition, we see that the topology on a topological subspace is at least the induced one. Indeed if $\mathcal{O}$ is an open subset of X, the set $i^{-1}(\mathcal{O})$ must be open in A, but it is clear that $i^{-1}(\mathcal{O}) = A \cap \mathcal{O}$.

48.1.1 Separability axioms

The following definitions are the **separability axioms**. A topological space X is said to be T_0 if for every pair of points in X, at least one of them has an open neighbourhood which does not contain the other.

The topological space X is T_1 if every point is closed or, equivalently, when for every two points, each has a neighbourhood which does not contain the other.

The topological space X is T_2 if every two points belong to disjoint neighbourhoods.

48.1.2 Axioms of countability

A topological space X is

(1) **first countable** if every point has a countable neighbourhood basis,
(2) **second countable** if the topology has a countable basis,
(3) **separable** if there exists a countable dense subspace.

Lemma 48.4.
If $\varphi \colon X \to Y$ is a linear map between two normed vector spaces such that

$$\|\varphi(A)\| \leqslant \|A\|,$$

then φ is continuous. lem:lin_vec_cont

Proof. Let $\mathcal{U}$ be open in Y, namely, consider that it is a ball centered at $y \in Y$ with ray δ; if $x \in \varphi^{-1}(\mathcal{U})$, we have to prove that A posses a neighbourhood contained in $\varphi^{-1}(\mathcal{U})$. Let's see the condition on $a \in X$ for $x+a$ to belongs to $\varphi^{-1}(\mathcal{U})$. We have

$$\|\varphi(x+a) - y\| \leqslant \|\varphi(x) - y\| + \|a\|,$$

but $\|\varphi(x)-y\| \leqslant \delta$. Taking $\|a\|$ suitably small, it is always possible to $x+a$ to belong to $\varphi^{-1}(\mathcal{U})$. □

Definition 48.5.
*If V and W are topological vector spaces, the map $f\colon V \to W$ is **uniformly continuous** when for every neighbourhood B of zero in W, there exists a neighbourhood A of zero in V such that*

$$v_1 - v_2 \in A \Rightarrow f(v_1) - f(v_2) \in B.$$

def:unif_cont

Definition 48.6.
*A topological space is **separable** when it admits a countable basis.*

Lemma 48.7.
Let Γ be a family of subsets of a set X such that $X = \bigcup_{A\in\Gamma} A$.

There exists a topology on X of basis Γ if and only if the intersection of any finite *subfamily of Γ can be written as the union of elements of Γ.* lem:topo_base

Lemma 48.8.
A metric space (E,d) is separable if and only if it contains a dense sequence. lem:sep_metric

Proof. Necessary condition. If the countable basis is (B_i), we take a x_i in each B_i. For any $x \in E$ and open set A containing x, there is a $B_i \subset A$ and then a $x_i \in A$. Thus the sequence (x_i) is dense in E.
Sufficient condition. Let (x_i) be the dense sequence. We consider $B_{i,k}$, the open ball centred in x_i of radius $1/k$. This is a countable system of open set in E. Let A be an open set in E and S, the union of all the $B_{i,k}$ which are contained in A. It is clear that $S \subset A$. Let $x \in A$ and D, an open ball of radius δ chosen in order to have $D \subset A$. We choose $k \in \mathbb{N}$ such that $1/k < \delta$ and x_i such that $d(x_i, x) < 1/k$. Then $x \in S$; namely x belongs to a ball of center x_i and of radius $1/k$. □

Definition 48.9.
*A subset of a topological space is **connected** if it cannot be written as an union of disjoint open sets.*

If A is a part of the topological space X, the **boundary** is the subset of X defined by

$$\mathrm{Fr}(A) = \overline{A} \cap \overline{\complement A}$$

Lemma 48.10.
Every path connected part of a topological space is connected.

Theorem 48.11.
Let E be a topological space and A, B two connected parts of E. If A intersects B and $\complement B$, then A intersects $\mathrm{Fr}(B)$. tho:doine

Proposition 48.12.
Let X be a topological space and A, B two subsets of X such that $A \cup B$ is closed and there exist A', B', two disjoint sets containing respectively the closure of A and B. Then A and B are separately closed. prop:sep_ferme

Proof. We consider a converging sequence in A. If x is the limit, $x \in A \cup B$ and $x \in \overline{A} \subset A'$. But it is clear from the assumptions that $A' \cap (A \cup B) = A$. Then $x \in A$, which proves that A is closed. □

Proposition 48.13.
Let X be a topological space, $K \subset X$ a compact and $\phi\colon K \to X$ a continuous map. Then $\phi(K)$ is compact in X.

DEFooERRQooNOpCXD

Definition 48.14.
*A family $\{X_i\}_{i\in I}$ of subsets of a topological space E is said to be **locally finite** if each $x \in E$ has a neighbourhood V such that $V \cap X_i \neq \varnothing$ for only finitely many values of i.*

LEMooGRPUooVhmqDH

Lemma 48.15 ([1])**.**
Let $\{X_i\}_{i\in I}$ be an open locally finite[1] covering of a topological space X. Then $\{\bar{X}_i\}_{i\in I}$ is a locally finite covering.

Proof. Let $x \in X$. We have an open neighbourhood V such that $I' = \{i \in I \text{ tel que } V \cap X_i \neq \varnothing\}$ is finite. We show that $\{i \in I \text{ tel que } V \cap \bar{X}_i \neq \varnothing\}$ is I'.

Let $i \in I$ and $a \in V \cap \bar{X}_i$. Since $a \in V$ and since V is open, a has an open neighbourhood $A \subset V$. Since $a \in \bar{X}_i$, we have $A \cap X_i \neq \varnothing$. Thus $V \cap X_i \neq \varnothing$ and $i \in I'$. □

DefParacompact

Definition 48.16.
*A topological space E is **paracompact** if every open cover admits a locally finite subcovering.*

Definition 48.17.
*A topological space E is **locally compact** if $\forall x \in E$, there exists a compact neighbourhood of x in E.*

For example, a (non-empty) open subset of $\mathbb{R}^n$ is *never* compact[2] but *always* locally compact.

Definition 48.18.
*A subset A of a metric space is **relatively compact** when $\overline{A}$ is compact.*

Definition 48.19.
*An **Hausdorff space** is a topological space in which any two distinct elements have disjoint neighbourhoods.*

Lemma 48.20.
If a set is Hausdorff for a topology τ_1 and compact in a weaker topology τ_2, then $\tau_1 = \tau_2$. lem:Hausweak

Proposition 48.21.
If K is a compact topological space and E a separated[3] one and if $f\colon K \to E$ is a continuous bijection, then f is an homeomorphism. lem:wiki

Definition 48.22.
*A space is **normal** when for all disjoint closed subset E and F of the topological space X, there exist neighbourhoods $\mathcal{U}$ and $\mathcal{V}$ of E and F which are disjoint too. A topological result says that a compact Hausdorff space is normal.*

lem:Urysohn

Lemma 48.23.
If X is a normal space, F a closed subset of X and $\mathcal{U}$ an open set containing F, then there exists a continuous function $f\colon X \to [0,1]$ such that $f(F) = 0$ and $f(X\backslash\mathcal{U}) = 1$.

Definition 48.24.
*A metric space is **complete** when any Cauchy sequence is converging in the space.*

1. Definition 48.14.
2. The compacts subsets of $\mathbb{R}^n$ are the closed and bounded subsets.
3. Comment on dit un espace séparé?

Pay attention not to be confused with a characterization of a *closed* subset: the limit of any converging sequence lies in the space.

def:separe

Definition 48.25.
Let E and F be two spaces in duality[4]*; if $A \subset E$ and $B \subset F$, we say that A **separates** B if $\forall x \neq y \in B$, $\exists a \in A$ such that $\hat{a}(x) \neq \hat{a}(y)$ where the hat denotes the duality.*

pg_def_Cz

For an Hausdorff X space, we denote by $C(X)$ the space of all continuous functions on X. For a locally compact space X, the set $C_0(X)$ contains continuous functions which decrease to zero at infinity in the sense that for any $\varepsilon > 0$, there exists a compact set K such that $|f(X)| < \varepsilon$ for all x outside K.

Definition 48.26.
*If X is an Hausdorff compact space, one can define the **supremum norm** on $C(X)$ by*

$$\|f\|_\infty := \sup_{x \in X} |f(x)|. \tag{48.1}$$

def:sup_norm

Definition 48.27.
*Let E be a vector space on $\mathbb{R}$ or $\mathbb{C}$. A **seminorm** on E is a map $p\colon E \to \mathbb{R}$ such that $\forall x, y \in E$, $\forall \lambda \in \mathbb{C}$,*

(1) $p(x) \geqslant 0$

(2) $p(\lambda x) = |\lambda| p(x)$

(3) $p(x + y) \leqslant p(x) + p(y)$.

A seminorm is a norm without the condition $p(x) = 0$ if and only if $x = 0$.

Definition 48.28.
*An **ecart**[5] on a set E is a map $d\colon E \times E \to \mathbb{R}$ such that $\forall x, y, z \in E$,*

(1) $d(x, x) = 0$

(2) $d(x, y) \geqslant 0$

(3) $d(x, y) = d(y, x)$

(4) $d(x, z) \leqslant d(x, y) + d(y, z)$.

It is a distance minus the condition $d(x, y) = 0$ if and only if $x = y$. It is clear that $d(x, y) = p(x - y)$ is an ecart on E when p is a seminorm. Moreover, this d is invariant under the translations and fulfils $d(\lambda x, \lambda y) = |\lambda| d(x, y)$.

Proposition 48.29.
Let p be a seminorm on a topological vector space. Then p is continuous if and only if there exists a neighbourhood of o in which p is bounded.

prop:semi_norm_cont

From a family of ecarts, we can define a topology. Let $(d_\alpha)_{\alpha \in I}$ be a family of ecarts on the set E. For an element $a \in E$, a *finite* subset $(\alpha_j)_{j=1\ldots m}$ of I and some numbers $(r_j)_{j=1\ldots m} > 0$, we write

$$B(a; (\alpha_j), (r_j)) = \{x \in E * d_{\alpha_j}(a, x) < r_j, \forall j = 1, \ldots, m\}.$$

topo_semi_norm

A subset $U \subset E$ is said **open** by respect of the family $(\alpha_i)_{i \in I}$ when for any $x \in U$, there exists a sub-family (α_j) and some numbers $(r_j) > 0$ such that $B(x; (\alpha_j), (r_j)) \subset U$.

Definition 48.30.
*When a topology can be defined from a family of ecarts, it is said an **uniformisable** topology.*

4. I'm not sure what it means.
5. C'est pas joli la traduction du mot "Écart".

If we have a family $(p_\alpha)_{\alpha\in I}$ of seminorm, we can build the ecarts $d_\alpha(x,y) = p_\alpha(x-y)$. One can show that this topological structure is compatible with the vector space structure.

Definition 48.31.
*A metric space is **separable** if there exists a subset at most countable which is everywhere dense.*

Proposition 48.32.
If E is a metric locally compact space, then the following are equivalent:

(1) There exists an increasing sequence (A_n) of open relatively compacts subset of E such that $\overline{A_n} \subset A_{n+1}$ and $E = \bigcup_n A_n$,

(2) E is a countable union of compact subsets,

(3) E is separable.

If A is a topological space with an equivalence relation $\sim$ and the canonical projection $\varphi\colon A \to A/\sim$, then $V \subset A/\sim$ is open in then **quotient topology** if and only if $\varphi^{-1}(V) \subset A$ is open.

48.2 Topology and convergence

When we have a topological space X, we can define a **convergence notion** for the sequences.

Definition 48.33.
We say that the sequence (x_n) in X converges to $x \in X$ if $\forall V \in \mathcal{V}(x)$, $\exists n_0$ such that $n \geqslant n_0$ implies $x_n \in V$. def:convergence

Definition 48.34.
*If X and Y are two topological spaces, a map $f\colon X \to Y$ is said **continuous** if $f^{-1}(\mathcal{U})$ is open in Y for any open set $\mathcal{U}$ in X.*

If f is continuous and $\mathcal{U}$ open in X, it is not true that $f(\mathcal{U})$ is open in Y. Simple counter-example are given by the constants functions from $\mathbb{R}$ to $\mathbb{R}$. So it is false that continuity preserves the openness. But it preserves the convergence.

Proposition 48.35.
Let a_k be a converging ($a_k \to a$) sequence in X and $g\colon X \to Y$, a continuous map. Then $g(a) = \lim_{k\to\infty} g(a_k)$. prop:continu_cv

Proof. If $\mathcal{U}$ is an open in Y which contains $g(a)$, $g^{-1}(\mathcal{U})$ is an open subset of X which contains a. Then from a certain N, $a_k \in g^{-1}(\mathcal{U})$, so that $g(a_k) \in \mathcal{U}$; it proves that the sequence $g(a_k)$ in Y converges to $g(a)$ in the sense of the topology of Y. □

Let us define a notion which depend only on the convergence notion.

Definition 48.36.
*An **usinage** around $x \in X$ is a subset $\mathcal{U} \subset X$ which contains $x \in \mathcal{U}$ and such that for any convergent sequence $(a_n) \to x$, there exists n_0 such that $n > n_0$ implies $a_n \in \mathcal{U}$.*

Lemma 48.37.
Let X be a metric topological space. The topology is completely determined by his convergence notion.

An other way to state this proposition: there is only one metric topology which gives rise to a given convergence notion. prop:usinage

Proof. We will show that, provided that the topology comes from a metric, the notion of neighbourhood is the same as the notion of usinage.

On the one hand, from topological definition 48.33 of convergence, it is clear that any neighbourhood is an usinage.

On the other hand, any usinage must contains a metric ball centred at x and then is a neighbourhood. Indeed, if $\mathcal{U}$ doesn't contains any ball, then

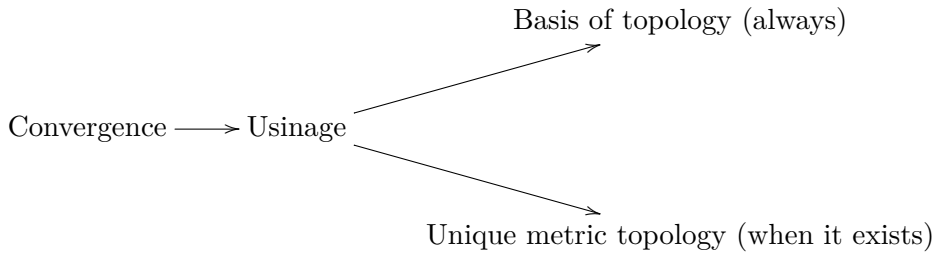
$$\forall \epsilon > 0, \exists a(\epsilon) \text{ such that } d(a(\epsilon), x) < \epsilon \text{ and } a(\epsilon) \notin \mathcal{U}.$$

In such a case, any sequence $(a(\epsilon_1), \dots)$ with $\epsilon_n \to 0$ contradict the fact that $\mathcal{U}$ is an usinage. □

Proposition 48.38.
For a given notion of usinage (i.e. for a given notion of convergence) on any set, there exists a topology (maybe no metric!) for which the usinage form a basis.

Proof. This is a direct application of lemma 48.7 and the fact that the intersection of two usinages is an usinage. □

The fact that a convergence notion gives rise to a topology is summarized by the following scheme:



Definition 48.39.
*Let $\mathcal{U}$ be an open subset of $\mathbb{R}^n$. A increasing sequence (K_m) of compacts subset of $\mathcal{U}$ is **fundamental** if $K_m \subset \operatorname{Int} K_{m+1}$ for all m and*

$$\bigcup_m K_m = \mathcal{U}$$

48.3 Locally convex

DEFooCGJBooSvDpyC

Definition 48.40 ([817])**.**
*A topological vector space is called **locally convex** if the origin has a neighborhood basis consisting of convex sets.*

Proposition 48.41.
A topological vector space is locally convex if and only if its topology can be defined from a family of seminorms.

LEMooDVZWooWKRQWC

Lemma 48.42 ([1])**.**
Let X be a locally convex vector space. Let K be compact in X and let U be an open set containing 0. There exists $\lambda \in \mathbb{R}$ such that $K \subset \lambda U$.

LEMooAHUMooBwxzPj

Lemma 48.43.
Let U be convex, $u_1, u_2 \in U$ and $a, b \in \mathbb{R}^+$. Then there exists $u \in U$ such that

$$au_1 + bu_2 = (a + b)u. \tag{48.2}$$

Proof. We have

$$au_1 + bu_2 = (a + b)\left(\frac{a}{a + b}u_1 + \frac{b}{a + b}u_2\right). \tag{48.3}$$

Since $a, b \in \mathbb{R}^+$, the coefficients $a/(a + b)$ and $b/(a + b)$ belong to $[0, 1]$ and they sum to 1. From convexity of U, the element $\frac{a}{a+b}u_1 + \frac{b}{a+b}u_2$ belongs to U. □

prop:topo_E

Proposition 48.44.
On $\mathscr{E}(\mathcal{U})$, there exists one and only one structure of locally convex metrisable Hausdorff space whose convergence notion is:

$(f_k) \to 0$ in $\mathscr{E}(\mathcal{U})$ when for any compact $K \subset \mathcal{U}$ and any multi-indices ν, the sequence of restriction $(D^\nu f_k|_K)$ converges uniformly to 0.

Here the D^ν denotes the multi-derivation with respect to the coordinates contained in the multi-index ν.

Proof. The unicity part is given by lemma 48.37. Now, we will build such a topology. First remark that we can find a fundamental sequence K_m in $\mathcal{U}$. Next, for any couple of integer $s \geqslant 0$, $m > 0$ and for any function $f \in \mathscr{E}^{(r)}(\mathcal{U})$ with $r \geqslant s$, we define

$$p_{s,m}(f) = \sup_{\substack{x \in K_m \\ |\nu| \leqslant s}} |(D^\nu f)(x)|.$$

One can see that these $p_{s,m}$ are seminorms, moreover $p_{s,m} \leqslant p_{r,m}$. Remark that if $p_{0,m}(f) = 0$ for any m, $f = 0$ on any K_m, and then $f = 0$ in $\mathcal{U}$. Thus $p_{0,m} = 0$ for all m implies $p_{s,m}(f) = 0$.

Now, we consider only the $p_{s,m}$ with $0 \leqslant s \leqslant m$ and the induced topology on $\mathscr{E}(\mathcal{U})$.

First, we prove that this topology is Hausdorff. Let us consider f, $g \in \mathscr{E}(\mathcal{U})$, and suppose that any ball centred at f contains g. Then we must obtain $f = g$. For any sequences (s_i), (r_i) with $0 \leqslant s_i \leqslant r_i$ and for any sequence $(\epsilon_j) > 0$,

$$g \in B(f; (p_{s_i,r_i}), (\epsilon_j)) = \{x \in \mathscr{E}(\mathcal{U}) |\, d_{p_{s_i,r_i}}(f, x) < \epsilon_j,\ \forall j\}.$$

In particular, for a sequence $\epsilon_j \to 0$ and $s_i = 0$, the condition becomes $p_{0,r_i}(f - g) < \epsilon_j$ for any f. Then $p_{0,r_i}(f - g) = 0$ and $f = g$.

By the way, remark that this topology is almost the "supremum norm" topology, then one can guess that the convergence notion will be something as the uniform convergence. There are just some subtleties as "for any compact" or "for any derivatives". But these are just what we need in the statement. Let us be more precise.

Definition (5) of a compact subset makes that any compact subset of $\mathcal{U}$ is a subset of one of the K_m. This makes our topology independent of the choice of the fundamental sequence K_m.

Let us consider (f_k), a sequence in $\mathscr{E}(\mathcal{U})$. The condition to converge to 0 in the sense of our new topology is that if A is an open subset of $\mathscr{E}(\mathcal{U})$ which contains 0, there exists a k_0 such that $k > k_0$ implies $f_k \in A$.

Then for any real sequences $(\epsilon_j) > 0$ and p_{s_i,r_i} with $0 \leqslant s_i \leqslant r_i$, $f_k \in B(0; (p_{s_i,r_i}), (\epsilon_j))$. Thus, $\exists k_0$ such that $k > k_0$ implies that $\forall s_j, r_j$,

$$\sup_{\substack{x \in K_{r_j} \\ |\nu| \leqslant s_j}} |D^\nu f_k(x)| < \epsilon_j.$$

Taking a sequence $\epsilon_j \to 0$, we get the thesis. □

Definition 48.45.
*A subset $H \subset \mathscr{E}(\mathcal{U})$ is **bounded** is each of the seminorm $p_{s,m}$ is bounded in H.* def:bounded

One can see that this notion only depend on the topology of $\mathscr{E}(\mathcal{U})$. Let us once again give some properties without proof.

Proposition 48.46.
There exists a sequence of functions in $\mathscr{E}(\mathcal{U})$ with compact support which is dense in each $\mathscr{E}^{(r)}(\mathcal{U})$ and in $\mathscr{E}(\mathcal{U})$. With others words, the spaces $\mathscr{E}^{(r)}(\mathcal{U})$ and $\mathscr{E}(\mathcal{U})$ are Fréchet separable spaces. prop:E_Frechet

Proposition 48.47.
Any bounded subset of $\mathscr{E}(\mathcal{U})$ is relatively compact in $\mathscr{E}(\mathcal{U})$.

Proposition 48.48.
For any multi-index ν, the linear map $D^\nu \colon \mathscr{E}(\mathcal{U}) \to \mathscr{E}(\mathcal{U})$, $f \to D^\nu f$ is continuous.

48.3.1 Topological group

DEFooCHZVooHnvSgW

Definition 48.49.
*A **topological group** is a group G equipped with a topological structure such that the maps $(x, y) \in G^2 \to xy \in G$ and $x \in G \to x^{-1} \in G$ are continuous.*

rem:ouvert

Remarque 48.50.
From the existence of an unique inverse for any element of G, the multiplication and the inversion are also open maps.

DEFooUEBTooDqipcL

Definition 48.51 (Uniform continuity)**.**
*Let G be a topological group and (E, d) a metric space. A map $f\colon G \to E$ is **uniformly continuous** if for every $\epsilon > 0$, there exists a neighbourhood V of e in E such that*

$$y^{-1}x \in V \Rightarrow d\big(f(x), f(y)\big) < \epsilon. \tag{48.4}$$

PROPooSHBAooVRdAFM

Proposition 48.52.
If K is compact in the topological group G, and if f is continuous on K, then it is uniformly continuous[6] *on K.*

48.4 Metrisable groups

sec:metrisable_groups

Let G be any group. The function $f\colon G \times G \to \mathbb{R}$ is said **left invariant** if for any x, y, $z \in G$, $f(xy, xz) = f(y, z)$. If G is abelian, the notion of left invariant and of right invariant are the same; in this case, we say it to be **translation invariant**.

For example, when a distance is left invariant, the left translations are isometries. For example, in a normed space vector, the distance $d(x, y) = \|x - y\|$ is always translation invariant: $\|(x + a) - (y + a)\| = \|x - y\|$.

Proposition 48.53.
Let G be a topological group[7]*.*

- *G is metrisable if and only if there exists a countable fundamental system of neighbourhoods of e (the neutral) whose intersection is only e.*
- *In this case, the topology of G ca be defined by a left invariant distance or a right invariant distance.*

On a metrisable topological group, one has thus two types of distance. But in general, one can not find a distance which defines the topology being left and right invariant.

Proposition 48.54.
Let G be a metrisable group. A sequence (a_n) in G is a Cauchy sequence for a left invariant distance if and only if
$\forall\, V \in \mathcal{V}(e)$, $\exists n_0 > 0$ *such that* $x_n^{-1}x_m \in V$ $\forall n, m \geqslant n_0$. ⌈ prop:Cauchy_metrisable

The point is that if a sequence is a Cauchy sequence for a left invariant distance, then it is a Cauchy sequence for *any* left invariant distance. This yields the important conclusion: in metrisable groups, the notion of Cauchy sequence is a topological notion which only depends on the neighbourhood notion.

A right invariant Cauchy sequence is the same but with $x_n x_m^{-1}$ instead of $x_n^{-1}x_m$. By definition of a topological group, the map $x \to x^{-1}$ is continuous. Thus, if all the left Cauchy sequences are convergent, then all the right Cauchy sequences are convergent. In this case, the group G is said **complete**.

If the group is abelian, we have only one notion of Cauchy sequence, and it is independent of the metric.

6. Definition 48.51.
7. Definition 48.49.

48.5 One point compactification

DEFooAKSWJcomWjkcMfv

Definition 48.55.
*Let X be a non compact topological space. We consider the "extended" space $\tilde{X} := X \cup \{\infty\}$ on which we put the following topology. Open sets of X are open in $\tilde{X}$ and when $\mathcal{O} \subset \tilde{X}$ contains ∞, it is open if and only if the complementary in is compact in X. It is easy to see that $\tilde{X}$ is compact. It is the **one point compactification** of X.*

Let us now consider $\tilde{X}$, any compact space and $X = \tilde{X}\backslash\{\infty\}$ where ∞ is a point in $\tilde{X}$. If we put on X the induced topology from $\tilde{X}$, we get a locally compact space whose compactification is $\tilde{X}$.

Exemple 48.56.
When $X = \mathbb{R}$ we add *one* infinity to get the compactification $\tilde{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ which is isomorphic to the circle S^1.

This is not the same space as $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ which is not compact. △

In this context we make the difference between ∞ and $+\infty$. When we write $+\infty$ we are speaking of the infinity in $\bar{\mathbb{R}}$ while writing ∞ with no sign we are speaking about the infinity in the one point compactification. The difference is important since for the one point compactification of $\mathbb{R}$ we have the limit

$$\lim_{x\to 0} \frac{1}{x} = \infty \tag{48.5}$$

while for the topology on $\bar{\mathbb{R}}$ the limite does not exist.

48.5.1 Example: the Riemann sphere

SEBSECooLJSEooNlyFYv

We already defined the Riemann sphere as a projective space in the definition 23.58.

DEFooMAHAooGUYyqU

Definition 48.57.
*The **Riemann sphere** is the one point compactification of $\mathbb{C}$. We denote it by $\hat{\mathbb{C}}$.*

See the definition 48.55 for the topology.

PROPooDTPKooYcSzYq

Proposition 48.58.
Let $\gamma\colon \mathbb{R} \to \hat{\mathbb{C}}$ be a path in $\hat{\mathbb{C}}$. We have $\lim_{t\to\infty} \gamma(t) = \infty$ if and only if $\|\gamma(t)\| \to +\infty$.

Proof. To be clear: the first limit is a limit in $\hat{\mathbb{C}}$ for its one point compactification topology while the second limit is an usual limit in $\mathbb{R}$.

(1) $\Rightarrow$ Let $M > 0$. Since $\gamma(t) \to \infty$, there exists a T such that $t > T$ implies

$$\gamma(t) \in \overline{B(O, M)}^c \tag{48.6}$$

because this is an open set around ∞. For that $t > T$ we thus have $\|\gamma(t)\| > M$. This proves that $\|\gamma(t)\| \to +\infty$.

(2) $\Leftarrow$ Let $\mathcal{O}$ be an open neighbourhood of ∞ in $\hat{\mathbb{C}}$. This $\mathcal{O}^c$ is compact in $\mathbb{C}$; $\mathcal{O}^c$ is then bounded [8] and there exists M such that $\mathcal{O}^c \subset B(0, M)$.
There also exists T such that $t > T$ implies $\|\gamma(t)\| > M$. Thus for $t > T$ we also have $\gamma(t) \in B(0, M)^c \subset \mathcal{O}$.

□

As far as arithmetic is concerned, the proposition 48.58 makes us define $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$.

8. The theorem 9.24 applies because $\mathbb{C}$ has the topology of $\mathbb{R}^2$.

Remarque 48.59.

```
 SageMath Version 7.0, Release Date: 2016-01-19
 Type "notebook()" for the browser-based notebook interface.
 Type "help()" for help.

sage: f(x)=1/x
sage: limit(f,x=0)
x |--> Infinity
sage: g(x)=x**2
sage: limit(g,x=oo)
x |--> +Infinity
```

tex/sage/sageSnip002.sage

Sage makes the difference between $+\infty$, $-\infty$ and ∞.

In the first case, $1/x$ is seen in $\hat{\mathbb{C}}$ and the limit is ∞, while in the second case, the limit is understood in $\mathbb{R}$ and sage provides $+\infty$.

DEFooPXYYooOMZYOT

Definition 48.60 (Pole of a complex function [818])**.**
Suppose U is an open subset of the complex plane $\mathbb{C}$, p is an element of U and $f\colon U\backslash\{p\} \to \mathbb{C}$ is an holomorphic function over its domain. If there exists a holomorphic function $g\colon U \to \mathbb{C}$, such that $g(p)$ is nonzero, and a positive integer n, such that for all $z \in U\backslash\{p\}$,

$$f(z) = \frac{g(z)}{(z-p)^n} \tag{48.7}$$

*holds, then p is called a **pole** of f. The smallest such n is called the **order**. A pole of order 1 is called a **simple pole**.*

*A function $f\colon \hat{\mathbb{C}} \to \mathbb{C}$ which is holomorphic on a neighbourhood of ∞ has a **pole** at $z = \infty$ if the function $z \mapsto f(1/z)$ has a pole at $z = 0$.*

48.6 Topological approximation

48.6.1 Introduction

In this section, we follow [819]. Let consider a particle moving on a circle S^1, and suppose that we have three detectors, each of them cover a surface, let

$$\mathcal{U}_1 =]-\frac{\pi}{3}, \frac{2\pi}{3}[\qquad \mathcal{U}_2 =]\frac{\pi}{3}, \frac{4\pi}{3}[, \qquad \mathcal{U}_3 =]\pi, 2\pi[.$$

If the detectors $\mathcal{U}_1$ and $\mathcal{U}_2$ are react, we know that the particle is in $\mathcal{U}_1 \cap \mathcal{U}_2$ (which is an open set), but no more; while if the detector $\mathcal{U}_1$ reacts alone, we know that the particle *is not* in $\mathcal{U}_2$ or $\mathcal{U}_3$, so that the particle belongs to

$$\mathcal{U}_1\backslash(\mathcal{U}_2 \cap \mathcal{U}_3),$$

which is closed. Our experimental setting allows us to distinguish six parts of S^1 that we name in the following way

$$\begin{aligned} \mathcal{U}_1 \cap \mathcal{U}_3 &\mapsto \alpha & \mathcal{U}_1\backslash(\mathcal{U}_2 \cup \mathcal{U}_3) &\mapsto a \\ \mathcal{U}_1 \cap \mathcal{U}_2 &\mapsto \beta & \mathcal{U}_2\backslash(\mathcal{U}_1 \cup \mathcal{U}_3) &\mapsto b \\ \mathcal{U}_2 \cap \mathcal{U}_3 &\mapsto \gamma & \mathcal{U}_3\backslash(\mathcal{U}_1 \cup \mathcal{U}_2) &\mapsto c. \end{aligned}$$

Notice that each point in S^1 belongs to one and only one of these sets. One says that a point of S^1 is open when it belongs to α, β or γ and that it is closed when it belongs to a, b, c. Now let us consider the six point space

$$P = \{\alpha, \beta, \gamma, a, b, c\},$$

on which we put the topology induced from S^1. A basis of that topology is given by

$$\{\alpha\}, \{\beta\}, \{\gamma\}, \{\alpha, a, \beta\}, \{\alpha, c, \gamma\}, \{\beta, b, \gamma\}.$$

48.6.2 Generalization

Let $(X, \mathcal{U}_\lambda)$, a topological space. We define the equivalence relation $x \sim y$ if and only if $x \in \mathcal{U}_\lambda \Leftrightarrow y \in \mathcal{U}_\lambda$ for every λ. In other words, x is equivalent to y when they cannot be distinguished by the topology. Now we consider the quotient space

$$P_{\mathcal{U}}(X) = X/\sim$$

with its quotient topology: $A \subset P_{\mathcal{U}}(X)$ is open if and only if $\pi^{-1}(A)$ is open in X where $\pi\colon X \to P_{\mathcal{U}}(X)$ is the canonical projection. It is the finest topology in which π is continuous.

When X is compact, the covering $\mathcal{U}_\lambda$ can be finite, so that $P_{\mathcal{U}}(X)$ is finite. If X is only locally finite, the space $P_{\mathcal{U}}(X)$ will be countable and it is said to be **unitary**. Notice that $P_{\mathcal{U}}(X)$ is not Hausdorff in general, for example in $P_6(S^1)$, the point a cannot be separated from α. Neither it is T_1 because points α, β and γ are open (in genera, it is even possible to get points which are open neither closed). The space $P_{\mathcal{U}}(X)$ is however always T_0.

48.6.3 Order and topology

When P is a finite topological space, the collections τ of open sets in P is closed under arbitrary unions and intersections. For each $x \in P$ we can consider the set

$$\Lambda(x) = \bigcap \{\mathcal{O} \in \tau \text{ tel que } x \in \mathcal{O}\} \tag{48.8}$$

which is the smallest neighbourhood of x. We define the relation $x \preceq y$ if and only if $\Lambda(x) \subseteq \Lambda(y)$, or equivalently if and only if every open neighbourhood of y contains x. Yet another way to express it is $x \preceq y$ if and only if $y \in \overline{\{x\}}$ where the bar denotes the closure. The relation $\preceq$ is reflexive and transitive.

The space P is T_0 by assumptions, so that for every couple of points $x, y \in P$, there always is a neighbourhood of x which does not contain y (or *vice versa*). That makes the relation antisymmetric:

$$x \preceq y \text{ and } y \preceq \Rightarrow x = y.$$

So the relation $\preceq$ is reflexive, antisymmetric and transitive. Such a relation is a **partial order**; a partially ordered set is often called **poset** for short.

The construction can be carried in the inverse sense: if $(P, \preceq)$ is a poset, we define $\Lambda(x) = \{y \in P \text{ tel que } y \preceq x\}$ and we define an open set in P as an union of such subsets. The resulting topological space is T_0.

A map $f\colon P \to Q$ between two posets is continuous if and only if it preserves order:

$$x \preceq_P y \Rightarrow f(x) \preceq_Q f(y). \tag{48.9}$$

When $x \preceq y$ and $x \neq y$, we write $x \prec y$.

With every poset, there is a **Hasse diagram** associated. It is obtained by drawing one point for each element of the poset on different levels and lines between them respecting the following two rules:

(1) if $x \prec y$, then the point associated with x has to be one level below the one of y,

(2) if $x \prec y$ and if there does not exist z such that $x \prec z \prec y$ then the point of x in one level bollow the one of y and the two points are connected by a line.

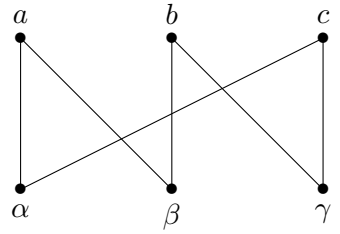
Let us draw the Hasse diagram for $P_6(S^1) = \{a, b, c, \alpha, \beta, \gamma\}$. We have

$$\begin{array}{lll} \Lambda(\alpha) = \{\alpha\} & \Lambda(\beta) = \{\beta\} & \Lambda(\gamma) = \{\gamma\} \\ \Lambda(a) = \{\alpha, a, \beta\} & \Lambda(b) = \{\beta, b, \gamma\} & \Lambda(c) = \{\alpha, c, \gamma\} \end{array}$$

and

$$\alpha \prec a, \qquad \alpha \prec c \qquad \beta \prec a \qquad \beta \prec b \qquad \gamma \prec b \qquad \gamma \prec c,$$

so that the Hasse diagram of $P_6(S^1)$ is



The smallest open set containing a point x is the set of points below x which are connected to x by a sequence of lines.

48.7 Partition of unity

DEFooKFXLooFRLaBG

Definition 48.61 ([820]).
*Let X be a topological space. A **partition of unity** of X is a family of continuous functions $\{\psi_j\}_{j\in J}$ such that*

(1) $\psi_j : X \to [0,1]$;

(2) for every $x \in X$ there exists a neighbourhood of x in X in which only a finite number of the ψ_j's is non zero;

(3) for every $x \in X$ we have

$$\sum_{j\in J} \psi_j(x) = 1. \tag{48.10}$$

Definition 48.62 ([820]).
*Let X be a topological space and $\{U_i\}_{i\in I}$ be a locally finite cover of X. A partition of unity is **subordinate** to that cover if it is indexed by I ($J = I$ in the definition 48.61) and such that $\operatorname{supp}(\psi_i) \subset U_i$ for every $i \in I$.*

THOooPCHDooITWKpC

Theorem 48.63 ([820]).
Let Ω be an open set in $\mathbb{R}^d$ and $\{U_i\}_{i\in I}$ be an open cover of Ω. There exists

(1) a C^∞ partition of unity $\{\psi_j\}_{j\in J}$ such that $\operatorname{supp}(\psi_j)$ is compact in one of the U_i;

ITEMooFGMJooQPLqGY

(2) a C^∞ partition of unity $\{\alpha_i\}_{i\in I}$ subordinated to the cover, such that for every compact K only a finite number of these ψ_i's is non zero.

Remarque 48.64.
This theorem does not furnish a smooth compactly supported partition of unity subordinated to the given cover. Either you choose the partition to be compactly supported, either you choose them subordinated to the cover.

Definition 48.65 ([820]).
*Let X be a topological space and $\{U_i\}_{i\in I}$ be a locally finite cover of X. A partition of unity is **subordinate** to that cover if it is indexed by I ($J = I$ in the definition 48.61) and such that $\operatorname{supp}(\psi_i) \subset U_i$ for every $i \in I$.*

CORooMSWPooCxvuhm

Corollary 48.66.
If Ω is bounded and $\{U_i\}_{i\in I}$ is an open cover of Ω, there exists a partition of unity subordinated to $\{U_i\}_{i\in I}$ such that each ψ_i belongs to $\mathscr{D}(U_i)$.

Proof. If Ω is bounded in $\mathbb{R}^d$ we can consider $U_i' = U_i \cap \Omega$ and use the point (2) of theorem 48.63 for the cover $\{U_i'\}_{i\in I}$. So we have a partition of unity subordinated to that cover with supports in the U_i''s. Since the support is closed and the U_i's are bounded, the supports are compact. The functions of this partition of unity are also subordinated to the original U_i's. □

DEFooTUYIooCRkgDm

Definition 48.67 ([821])**.**
*Let X be a topological space. A part $\mathcal{A} \subset C^0(X)$ is **normal** is for every disjoint parts $A, B \subset X$, there exists $f : X \to [0,1]$ such that*

(1) $f(X) \in [0,1]$,

(2) $f(A) = \{0\}$

(3) $f(B) = \{1\}$.

LEMooDZGCooIGFXnA

Lemma 48.68.
If X is a locally compact topological space, the part $C_c^0(X)$ is normal[9]

DEFooEDFIooHKnGdE

Definition 48.69.
*Let X be a topological space and $U = \{U_1, \dots, U_n\}$ be an open covering of X. A **partition of unity subordinated** to U is a set of functions $\{\phi_i\}_{i=1,\dots,n}$ such that*

(1) $\phi_i(X) \subset [0,1]$

(2) $\text{supp}(\phi_i) \subset U_i$

(3) $\sum_{i=1}^n \phi(x) = 1$ for every $x \in X$.

If $\mathcal{A} \subset C^0(X)$ we say that $\{\phi_i\}$ is a $\mathcal{A}$-partition of unity of $\phi_i \in \mathcal{A}$ for every i.

THOooUGQCooFVySMP

Theorem 48.70 ([821])**.**
Let X be a topological space. Let $\mathcal{A} \subset C^0(X)$ be

(1) normal

(2) closed under finite sums

(3) if $f, g \in \mathcal{A}$ and $g(x) \neq 0$ for every $x \in X$, then $f/g \in \mathcal{A}$.

For every finite open covering $U_1, \dots, U_n$ of X, there exists a $\mathcal{A}$-partition of unity[10] *subordinated to $\{U_i\}$.*

48.8 Directed sets and net

Definition 48.71.
*A **directed set** is a pre-ordered set (i.e. a set with a reflexive and transitive binary relation) such that every pair of elements has an upper bound.*

As a consequence, when $a_1, \dots a_n$ are elements of the directed set A, then there exists a $a \in A$ such that $a \geqslant a_i$.

Definition 48.72.
*A **net** is a map $A \to X$ from a directed set to a topological space. We denote by x_α the element of X which corresponds to $\alpha \in A$.*

As example, if S is any set, the set A of finite subsets of S with the inclusion is an example of net.

There is a notion of **convergence** of net. We say that the net $\alpha \mapsto x_\alpha$ converges to x and we write $x_\alpha \to x$ if and only if for every open set $\mathcal{U} \subseteq X$ containing x, there exists a $\alpha \in A$ such that $\alpha' \geqslant \alpha$ implies $x_{\alpha'} \in \mathcal{U}$.

9. Definition 48.67.
10. Definition 48.69.

So a topology implies a convergence notion for nets, as well as for sequences. However, there exists different topologies which have the same notion of convergence of sequences, but two topologies having the same notion of convergence of nets are the same.

48.9 Homotopy group

Let X be a topological space with a base point b, and S^n be the n-sphere. The nth **group of homotopy** on the point b of X is

$$\pi_n(X,b) = \{\text{homotopy classes of maps } f\colon S^n \to X \text{ such that } f(a) = b\}. \tag{48.11}$$

The classes are taken up to homotopy, i.e. continuous deformations. In an equivalent way, $\pi_n(X,b)$ can be seen as the set of classes of maps $p\colon [0,1]^n \to X$ which sent the whole border of the cube to b.

48.10 Covering spaces

DEFooQBDWooVVrkkh

Definition 48.73.
*Let M be an Hausdorff connected topological space. Let X be a topological space. A map $\pi\colon X \to M$ is a **covering** if*

(1) it is continuous

(2) it is surjective

(3) for every $p \in M$, there exists a neighbourhood V of p such that $\pi^{-1}(V)$ is a union of disjoint open sets $\{A_i\}$ in X such that the restriction $\pi\colon A_i \to V$ is homeomorphic[11].

Proposition 48.74 (lifting property)**.**
Let $\rho\colon C \to X$ be a covering and $\gamma\colon [0,1] \to X$, a continuous map. Let $c \in \rho^{-1}\big(\gamma(0)\big)$. Then there exists one unique path σ in C such that $\sigma \circ \rho = \gamma$ and $\sigma(0) = 0$.

Proof. No proof. □

If x and y in X are connected by a path, the lifted path provides a bijection between the fibres $\rho^{-1}(x)$ and $\rho^{-1}(y)$.

PROPooPWIFooFAZhVe

Proposition 48.75.
If M is a simply connected manifold, then every covering is an homeomorphism.

48.10.1 Universal covering

Definition 48.76.
*One says that a covering $q\colon D \to X$ is **universal** if D is simply connected.*

The following proposition states that an universal covering is a covering that covers all other coverings.

Proposition 48.77.
Let $q\colon D \to X$ be an universal covering, and $\rho\colon C \to X$ be a covering of X with C being connected. Then there exists a covering map $f\colon D \to C$ such that $\rho \circ f = q$.

The following proposition states that the universal covering is essentially unique.

Proposition 48.78.
Let $q_i\colon D_i \to X$ (with $i = 1, 2$) be two universal coverings of the topological space X. Then there exists an homeomorphism $f\colon D_1 \to D_2$ such that $q_2 \circ f = q_1$.

11. Definition 7.41.

48.10.2 Monodromy action

sssMonodromyact

Let $\rho: C \to X$ be a covering with C being connected and locally arc connected. First, that shows that X has these two properties too. Now, let $x \in X$ and $c \in \rho^{-1}(x)$ and a path $\gamma: [0,1] \to X$ with $\gamma(0) = \gamma(1) = x$. By the lifting property, that path lifts to an unique path in C starting at c, while it is not guarantee that the lifted path will *end* at x. One only knows that the lifted path will end in $\rho^{-1}(x)$.

It turns out that the end point of the lifted path only depends on the class of γ in $\pi(X,x)$. Thus we define an action if $\pi(X,x)$ on the fibre over x. This is the **monodromy action**. Notice that by taking that action pointwise on X, the group $\pi(X)$ acts on C.

48.11 Haar measure

THOooWXQFooQrWcLY

Theorem 48.79 (Kakutani fixed point[822, 823])**.**
Let X be an Hausdorff locally convex[12] topological vector space. Let K be compact in X and $\mathcal{S}$ be a group of linear endomorphisms of X such that

(1) $\mathcal{S}$ is equicontinuous[13] on K,

(2) $\mathcal{S}(K) \subset K$.

There exists a point $p \in K$ such that $T(p) = p$ for every $T \in \mathcal{S}$.

Proof. We'll use the Zorn lemma. Let

$$\mathcal{A} = \big\{L \subset K \text{ tel que } L \text{ is non empty, compact, convex and} \mathcal{S}(L) \subset L\big\}. \tag{48.12}$$

The set $\mathcal{A}$ is ordered with $L_1 \leqslant L_2$ if and only if $L_2 \subset L_1$ (this is the opposite of what we usually do).

(1) **$(\mathcal{A}, \leqslant)$ is inductive** Let $\mathcal{F} \subset \mathcal{A}$ be a totally ordered subset. We prove that

$$M = \bigcap_{A \in \mathcal{F}} A \tag{48.13}$$

is a upper bound of $\mathcal{F}$.

(1a) **Finite intersection property** We prove that the family $\mathcal{F}$ has the finite intersection property[14]. Let $\mathcal{I}$ be finite in $\mathcal{F}$. The set $(\mathcal{I}, \leqslant)$ is totally ordered and the lemma 1.15 says that $\mathcal{I}$ has a maximum. Here the order is the inverse inclusion, so that the maximum is in fact the smaller set, which is contained in all the others. In other words, the maximum, which is an element of $\mathcal{I}$ is the intersection $\bigcap_{A \in \mathcal{I}} A$.

(1b) **M is non empty** The family $\mathcal{F}$ is a family of closed[15] parts of X which has the finite intersection property. Thus the intersection of $\mathcal{F}$ is non empty by theorem 7.109.

(1c) **M is compact** Since X is Hausdorff, every intersection of compacts parts is compact by the proposition 7.99.

(1d) **M is convex** Proposition 7.155 says that every intersection of convex is convex.

(1e) **$\mathcal{S}(M) \subset \mathcal{S}(M)$** Let $x \in M$. For every $A \in \mathcal{F}$, we have $x \in A$. Since $\mathcal{S}(A) \subset A$ we also have $\mathcal{S}(x) \in A$ and $\mathcal{S}(x) \in \bigcap_{A \in \mathcal{F}} A$.

(2) **Zorn, definition of K_1** Since $\mathcal{A}$ is inductive we use the Zorn lemma 1.23 and we let K_1 be a maximal element of $\mathcal{A}$.

(3) **If K_1 has only one point** Let us suppose that K_1 contains only one element p. We prove that $T(p) = p$ for every $T \in \mathcal{S}$. Since $K_1 \in \mathcal{A}$ we have $T(K_1) \subset K_1$ and since $K_1 = \{p\}$ we have $T(p) \in \{p\}$ which means $T(p) = p$.

12. Locally convex, definition 48.40.
13. Definition 7.317.
14. Definition 7.108.
15. Compacts are closed, see 7.96.

We suppose that K_1 has at least 2 points and we will reach a contradiction. As far as the notations are concerned, we write $K_1 - K_1 = \{k_1 - k_2 \text{ tel que } k_1, k_2 \in K_1\}$.

(1) **$K_1 - K_1$ has a non zero element** The part K_1 contains at least two elements. Let k_1 and k_2 be two distinct elements in K_1. Then $k_1 - k_2 \neq 0$ in $K_1 - K_1$.

(2) **Definition of V** From lemma 7.64, we consider in X a neighbourhood V of 0 such that $\bar{V}$ is not contained in $K_1 - K_1$.

(3) **Definition of V_1** Since X is locally convex, 0 has a basis of convex neighbourhoods. Let $V_1 \subset V$ be a convex neighbourhood of 0.

(4) **Definition of U_1** The set $\mathcal{S}$ is equicontinuous on K and then equicontinuous on K_1. There exists a neighbourhood U_1 of 0 in X such that for every $k_1, k_2 \in K_1$ such that $k_2 - k_1 \in U_1$, we have $\mathcal{S}(k_1 - k_2) \subset V_1$. In other words,

$$\mathcal{S}\big(U_1 \cap (K_1 - K_1)\big) \subset V_1. \tag{48.14}$$

EQooPELQooJgcrzr

(5) **Open map** Since $\mathcal{S}$ is a group, each of the linear maps $T \in \mathcal{S}$ is invertible and its inverse is part of $\mathcal{S}$. In other words, the elements of $\mathcal{S}$ are continuous linear maps with continuous inverse.

In particular, if $\mathcal{O}$ is open in X, then $T(\mathcal{O})$ is open in X.

(6) **Définition of U_2** We define

$$U_2 = \mathcal{C}\text{onv}\Big(\mathcal{S}\big(U_1 \cap (K_1 - K_1)\big)\Big). \tag{48.15}$$

Since the elements $T \in \mathcal{S}$ are linear and inversible, we have $T(A \cap B) = T(A) \cap T(B)$ and $T\big(\mathcal{C}\text{onv}(A)\big) = \mathcal{C}\text{onv}\big(T(A)\big)$ for every parts $A, B \subset X$. Thus we have

$$\mathcal{S}(U_2) = U_2. \tag{48.16}$$

(7) **$K_1 - K_1 \subsetneq U_2$** The set U_2 is the convex hull of $\mathcal{S}\big(U_1 \cap (K_1 - K_1)\big)$ which is a part of the convex set V_1 by (48.14). Thus $U_2 \subset V_1$. We have the inclusions chain

$$U_2 \subset V_1 \subset V. \tag{48.17}$$

Since $K_1 - K_2$ is not included in V, it is not included in U_2.

(8) **Definition of U** Since the operators T are continuous and since $T(U_2) \subset U_2$, we also have $\mathcal{S}(\bar{U}_2) \subset \bar{U}_2$. The part U_2 is an open convex containing 0 and $K_1 - K_2$ is compact. Thus lemma 48.42 implies the existence of $\lambda > 0$ such that $K_1 - K_2 \subset \lambda U_2$. Let

$$\delta = \inf\{\lambda > 0 \text{ tel que } K_1 - K_1 \subset \lambda U_2\}. \tag{48.18}$$

Notice that $\delta < \infty$ because of lemma 48.42 and $\delta > 1$ because $K_1 - K_1 \not\subset U_2$. We define

$$U = \delta U_2. \tag{48.19}$$

(9) **Some sums** Let $0 < \epsilon < 1$. For every $\lambda > \delta$ we have $K_1 - K_1 \subset \lambda U_2$. In particular with $\lambda = (1 + \epsilon)\delta$. Thus

$$K_1 - K_1 \subset (1 + \epsilon)\delta U_2 = (1 + \epsilon)U. \tag{48.20}$$

In the same way, for every $\lambda < \delta$ we have $K_1 - K_1 \not\subset \lambda U_2$. Thus

$$K_1 - K_2 \not\subset (1 - \epsilon)\bar{U}. \tag{48.21}$$

EQooNYNTooQScVJL

For $k \in K_1$ we consider $\mathcal{O}_k = \frac{1}{2}U + k$. Since $0 \in U$ the set $\{\mathcal{O}_k\}_{k \in K_1}$ is an open covering of K_1. The set K_1 being compact we can extract a finite subcovering: there exists $k_1, \dots, k_n \in K_1$ such that

$$K_1 \subset \bigcup_{i=1}^{n} (\frac{1}{2}U + k_i). \tag{48.22}$$

Let

$$p = \frac{k_1 + \dots + k_n}{n}. \tag{48.23}$$

Let $k \in K_1$. Two statements:

(9a) For each l we have $k_l - k \in (1+\epsilon)U$.

(9b) There exists i such that $k_i - k \in \frac{1}{2}U$.

Thus there exist $u_1, \dots, u_n \in U$ such that [EQSooZFRQooJoFhXF]

$$\begin{cases} k_i - k = \frac{1}{2}u_i & (48.24a) \\ k_l - k = (1+\epsilon)u_l. & (48.24b) \end{cases}$$

Inserting the values of k_i from (48.24) in the definition of p,

$$\begin{aligned} p &= \frac{1}{n}\sum_{j=1}^{n} k_j && (48.25a) \\ &= \frac{1}{n}\Big(\frac{1}{2}u_i + k + \sum_{j\neq i}\big[(1+\epsilon)u_l + k\big]\Big) && (48.25b) \\ &= \frac{1}{n}\Big(\frac{1}{2}u_i + \sum_{j\neq i}(1+\epsilon)u_l\Big) && (48.25c) \end{aligned}$$

The set U is convexe because $U = \delta U_2$ where U_2 is convex. Thus if $u_1, u_2 \in U$ we have $\frac{u_1+u_2}{2} \in U$ and $u_1 + u_2 \in 2U$. In the same way, there exists $u' \in U$ such that

$$\sum_{j\neq i} u_l = (n-1)u'. \tag{48.26}$$

So we write

$$p = \frac{1}{n}\big(\frac{1}{2}u_i + (n-1)(1+\epsilon)u'\big) + k. \tag{48.27}$$

Using lemma 48.43, there exists $u \in U$ such that

$$p = \frac{1}{n}\big(\frac{1}{2} + (n-1)(1+\epsilon)\big)u + k. \tag{48.28}$$

We write that with $\epsilon = \frac{1}{4(n-1)}$. The coefficient becomes

$$\frac{1}{2} + (n-1)(1+\epsilon) = \frac{4n-1}{4} \tag{48.29}$$

and we have proven that, for each $k \in K_1$, there exists $u \in U$ such that

$$p = \frac{4n-1}{4n}u + k. \tag{48.30}$$

(10) **Definition of K_2** Let, for each $k \in K_1$,

$$A_k = \frac{4n-1}{4n}\bar{U} + k \tag{48.31}$$

and

$$K_2 = K_1 \cap \bigcap_{k\in K_1} A_k. \tag{48.32}$$

We give some properties of K_2.

(10a) **K_2 is non empty** We know that $p \in K_2$.

(10b) **$K_2 \subsetneq K_1$** From what we said around equation (48.21), there exists $k_1, k_2 \in K_1$ such that

$$k_1 - k_2 \notin \left(1 - \frac{4n-1}{4n}\right)\bar{U}, \tag{48.33}$$

which means that $k_1 \notin A_{k_2}$. In particular $k_1 \notin K_2$. Thus the inclusion $K_2 \subset K_1$ is strict.

(10c) K_2 **is compact and convex** The sets K_1 and A_k are closed. The set K_2 is closed as intersection of closed sets. Since K_2 is closed in the compact K_1, it is compact (lemma 7.96(1)).

The set K_2 is convex as intersection of convexes.

(10d) $\mathcal{S}(K_2) \subset K_2$ Let $T \in \mathcal{S}$ and $u \in \bar{U}$. We have $u = \delta u_2$ with $u_2 \in \bar{U}_2$. Then

$$T(u) = \delta T(u_2) \in \delta T(\bar{U}_2) \subset \delta \bar{U}_2 = \bar{U} \tag{48.34}$$

because $T(U_2) \subset U_2$ and T is continuous. We proved that $T(\bar{U}) \subset \bar{U}$.

Let $a \in K_2$. We have $a \in K_1$ and, since $a \in A_k$ for each $k \in K_1$, we have a map $u\colon K_1 \to \bar{U}$ such that

$$\lambda u(k) + k = a \tag{48.35}$$

EQooFDOZooNTThUj

where $\lambda = \frac{4n-1}{4n}$. Consider $l \in K_1$ and let us show that $T(a) \in A_l$. Since $T(K_1) = K_1$, there exists $l' \in K_1$ such that $T(l') = l$. Writing (48.35) with l' we have

$$T(a) = \lambda T\big(u(l')\big) + T(l') = \lambda T\big(u(l')\big) + l. \tag{48.36}$$

Since $u(l') \in \bar{U}$ and $T(\bar{U}) \subset \bar{U}$, we have

$$T(a) \in \lambda \bar{U} + l = A_l. \tag{48.37}$$

We have proved that $T(a) \in A_l$ for each $l \in K_1$.

(11) **Conclusion** The part K_2 belong to $\mathcal{A}$ and satisfies the strict inclusion $K_2 \subsetneq K_1$. This means that $K_2 > K_1$, which contradicts the maximality of K_1.

We conclude that K_1 has only one point, so that $\mathcal{S}$ has a unique fixed point as shown before.

□

We recall the regular left action. If G is a group, u is a map $u\colon G \to E$ and $s \in G$, we write

$$(L_s u)(x) = u(s^{-1}x). \tag{48.38}$$

DEFooSBRZooUbzMnN

Definition 48.80 (Haar measure)**.**
Let G be a compact topological group[16]*. A linear form $m\colon C^0(G,\mathbb{C}) \to \mathbb{C}$ is a* ***Haar functional***[17] *if it satisfies*

(1) $m(f) \geqslant 0$ *for every* $f \geqslant 0$.

(2) $m(\mathbb{1}) = 1$ *where* $\mathbb{1}$ *is the constant function* $\mathbb{1}(x) = 1$.

(3) $m(L_s f) = m(f)$ *for every* $f \in C^0(G,\mathbb{C})$ *and* $s \in G$.

DEFooXIKEooWOxHlr

Definition 48.81.
Let G be a topological group. A ***Haar measure*** *on G is a positive measure*[18] *μ such that*

(1) $\mu(gA) = \mu(A)$ *for every* $g \in G$ *and measurable part* $A \subset G$,

ITEMooNETPooYTwhBx

(2) $\mu(K) < \infty$ *for every compact part* $K \subset G$.

The link between measure and functional is given by the Riesz-Markov representation theorem 48.88.

LEMooRWOFooDOSUYo

Lemma 48.82 ([824])**.**
A Haar functional[19] *on a compact topological group G is continuous on $\big(C^0(G,\mathbb{C}), \|.\|_\infty\big)$.*

16. Definition 48.49.
17. Several place say Haar *measure*, but we prefer to keep the word "measure" for the good old concept of measure, and give a Haar measure in definition 48.81.
18. Definition 14.18.
19. Definition 48.80.

Proof. Let $f \in C^0(G, \mathbb{C})$, and let m be a Haar functional on G.

Support for the moment that f is real valued. For every $x \in G$ we have

$$-\|f\| \leqslant f(x) \leqslant \|f\|. \tag{48.39}$$

If we consider $\|f\|$ as the constant function on G we can write the inequalities $-\|f\| \leqslant f \leqslant \|f\|$ in $C^0(G, \mathbb{C})$. Since m is positive, we can apply m to these inequalities :

$$m(-\|f\|) \leqslant m(f) \leqslant m(\|f\|). \tag{48.40}$$

By linearity of m we have $-\|f\|m(\mathbb{1}) \leqslant m(f) \leqslant \|f\|m(\mathbb{1})$, and then

$$-\|f\| \leqslant m(f) \leqslant \|f\|. \tag{48.41}$$

Now we suppose that f is complex-valued and that $f(x) = u(x) + iv(x)$ where u and v are real-valued. We have: SUBEQooKMQEooOoydWP

$$\begin{aligned}
|m(f)| &= |m(u) + im(v)| && \text{(48.42a)}\\
&\leqslant |m(u)| + |m(v)| && \text{(48.42b)}\\
&\leqslant \|u\| + \|v\| && \text{(48.42c)}\\
&\leqslant 2\|f\|. && \text{(48.42d)}
\end{aligned}$$

We used the result for real valued functions and lemma 12.355. The inequality (48.42) shows that m is continuous at 0. Since m is linear, it is continuous. □ THOooHZNRooLWmJMB

Theorem 48.83 (Haar functional[822, 824])**.**
Let G be a compact topological group.

(1) There exists a unique Haar functional[20] *on G.*

(2) The Haar functional satisfies $m(R_s f) = m(f)$.

Proof. Here is a summary of the proof.

(1) We prove the existence by building a Haar functional m.

(2) We show that m satisfies $m(R_s f) = m(f)$.

(3) We show the unicity by showing that every Haar functional is equal to m.

For $f \in C^0(G, \mathbb{C})$ we consider C_f, the set of convex combinations of the functions of the form $L_s(f)$. In other words, $g \in C_f$ is there exists $\{(a_i, s_i) \in \mathbb{R}^+ \times G\}_{i=1,\dots,N}$ such that $a_i > 0$, $\sum_i a_i = 1$ and $g(x) = \sum_i a_i f(s_i x)$.

Since f is continuous in the compact space G, we define

$$\|f\| = \max_{x \in G} |f(x)|. \tag{48.43}$$

If $g \in C_f$ we have $\|g\| \leqslant \|f\|$ because

$$\begin{aligned}
\|g\| &= \max_{x \in G} |\sum_i a_i f(s_i x)| && \text{(48.44a)}\\
&\leqslant \sum_i \max_{x \in G} |a_i f(s_i x)| && \text{(48.44b)}\\
&\leqslant \sum_i a_i \max_{x \in G} |f(s_i x)| && \text{(48.44c)}\\
&= \sum_i a_i \|f\| && \text{(48.44d)}\\
&= \|f\|. && \text{(48.44e)}
\end{aligned}$$

20. Definition 48.80.

For $x \in G$ we also define

$$C_f(x) = \{g(x) \text{ tel que } g \in C_f\}. \tag{48.45}$$

This is a bounded set in $\mathbb{C}$ because each element is bounded by $\|f\|$. The closure $\overline{C_f(x)}$ is then compact in $\mathbb{C}$. The function f being continuous, it is uniformly continuous on $\overline{C_f(x)}$ (proposition 48.52.). Let $\epsilon > 0$ and an open neighbourhood V of e as in the definition of uniform continuity. If $x, y \in V$ we have

$$|(L_s f)(x) - (L_s f)(y)| = |f(s^{-1}x) - f(s^{-1}y)| < \epsilon \tag{48.46}$$

because $(s^{-1}y)^{-1}(s^{-1}x) = y^{-1}x$.

This shows that the set $\{L_s(f)\}_{s\in G}$ is equicontinuous[21]. Now we prove that C_f is equicontinuous. Let $g \in C_f$ and $x, y \in V$. We have

$$\begin{aligned}
|g(x) - g(y)| &= \Big|\sum_i a_i\big[f(s_i x) - f(x_i y)\big]\Big| && \text{(48.47a)}\\
&\leqslant \sum_i a_i |s_{s_i^{-1}}(x) - L_{s_i^{-1}}(y)| && \text{(48.47b)}\\
&\leqslant \sum_i a_i \epsilon && \text{(48.47c)}\\
&= \epsilon, && \text{(48.47d)}
\end{aligned}$$

so C_f is equicontinuous. The Ascoli theorem 27.4 implies that C_f is relatively compact in $C^0(G, \mathbb{C})$, which means that $K_f = \overline{C_f}$ is compact.

The action

$$\begin{aligned} L\colon G \times C^0(G,\mathbb{C}) &\to C^0(G,\mathbb{C})\\ (g,u) &\mapsto L_g(u) \end{aligned} \tag{48.48}$$

is isometric: $\|L_s(u)\| = \|u\|$ and leaves C_f invariant.

(1) $\underline{L_s(C_f) \subset C_f}$ We write

$$\begin{aligned} \psi\colon \mathbb{R}^N \times G^N &\to C_f\\ (a_i),(s_i) &\mapsto \sum_i a_i L_{s_i}(f). \end{aligned} \tag{48.49}$$

If $g \in C_f$ we have

$$g(x) = \psi\big((a_i),(s_i)\big)(x) = \sum_i a_i f(s_i x), \tag{48.50}$$

so that

$$\begin{aligned}
(L_s g)(x) &= \sum_i a_i f(s_i s^{-1} x) && \text{(48.51a)}\\
&= \psi\big((a_i),(s_i s^{-1})\big)(x). && \text{(48.51b)}
\end{aligned}$$

This shows that $L_s(g) \in C_f$.

(2) $\underline{C_f \subset L_s(C_f)}$ The same kind of computation shows that

$$\psi\big((a_i),(s_i)\big) = L_s\Big(\psi\big((a_i),(s_i s)\big)\Big). \tag{48.52}$$

(3) **Kakutani** Let us summarise some facts.

— $L_s(K_f) = K_f$ by lemma 7.169.

— K_j is compact and convex in $C^0(G, \mathbb{C})$.

— G is a group of linear transformations acting on K_f : $G(K_f) \subset K_f$.

— G is isomectric on K_f, and then equicontinuous on K_f by lemma 7.318.

21. Definition 7.317.

The Kakutani theorem 48.79 is then valid and there exists a fixed point $h \in K_f$ for the action of G. In other words, $L_s(h) = h$ for every $s \in G$. The function h is constant because for every $s \in G$ we have

$$h(s) = (L_{s^{-1}}h)(e) = h(e). \tag{48.53}$$

Thus K_f contains one constant function. We prove now that it contains only one constant.

(4) **K_f contains only one constant** We have just seen that K_f contains a constant function. We are going to show that K_f contains only one constant. Let $c \in \mathbb{C}$ be such a constant, and let $\epsilon > 0$. Since $c \in K_f = \bar{C}_j$, there exists an element $g \in C_f$ such that $\|c - g\| < \epsilon$. More precisely, there exists $a_i > 0$ and $\{s_1, \dots, s_N\} \in G$ such that SUBEQSooDGOPooQLxVjw

$$\begin{cases} g(x) = \sum_{i=1}^{N} a_j f(s_i x) & \text{(48.54a)} \\ \sum_{i=1}^{N} a_i = 1 & \text{(48.54b)} \\ |c - g(x)| < \epsilon. & \text{(48.54c)} \end{cases}$$

The third condition holds for every $x \in G$.

We can make the whole work with the right action instead of the left action of G on $C^0(G, \mathbb{C})$. That leads us to build C'_f, K'_f and a constant function $c' \in K'_f$, elements $b_j \geqslant 0$ and $t_j \in G$ such that SUBEQSooBQIZooApBecK

$$\begin{cases} |c' - \sum_{j=1}^{N} b_j f(x t_j)| < \epsilon & \text{(48.55a)} \\ \sum_{j=1}^{N} b_j = 1. & \text{(48.55b)} \end{cases}$$

You may wonder why the conditions (48.55) hold for the same N as the conditions (48.54). In fact, we took N as the maximum of the two and we set $a_i = 0$ or $b_i = 0$ for the indices i between the «correct» N and N.

Let us fix i, multiply the first condition (48.55) by a_i and set $x = s_i$:

$$|c' a_i - \sum_{j=1}^{N} a_i b_j f(s_i t_j)| < \epsilon a_i. \tag{48.56}$$

Making the sum over i we get [22]

$$|\sum_i c' a_i - \sum_{ij} a_i b_j f(s_i t_j)| < \sum_i \epsilon a_i, \tag{48.57}$$

and then

$$|c' - \sum_{ij} a_i b_j f(s_i t_j)| < \epsilon. \tag{48.58}$$

We make the same from (48.54) : multiply by b_j, $x = t_j$ and sum over j :

$$|c - \sum_{ij} a_i b_j f(s_i t_j)| < \epsilon. \tag{48.59}$$

And we deduce that $c = c'$ because

$$|c - c'| \leqslant |c - \sum_{ij} a_i b_j f(s_i t_j)| + |\sum_{ij} a_i b_j f(s_i t_j) - c'| < 2\epsilon. \tag{48.60}$$

We have shown that if c is a constant in K_f and if c' is a constant in K'_f, then $c = c'$. We deduce that the set of constants in K_f and K'_f are the same and contain only one element.

22. We can enter the sum in the module because if $|a_i| < b_i$, then $|\sum_i a_i| \leqslant \sum ß|a_i| < \sum_i b_i$.

(5) **Definition of m** We denote by $m(f)$ the unique $c \in K_f$ such that $\|c - C_f\| = 0$. This defines a map

$$m\colon C^0(G,\mathbb{C}) \to \mathbb{C}. \tag{48.61}$$

We prove now a lot of properties.

(6) $m(1) = 1$ We consider the constant function 1. A function in C_1 has the form

$$g(x) = \sum_i a_i 1(x_i x) = \sum_i a_i = 1. \tag{48.62}$$

Thus $C_1 = K_1 = \{1\}$, so that $m(1) = 1$.

(7) $m(f) \geqslant 0$ **if** $f \geqslant 0$ An element of C_f has the form $g(x) = \sum_i a_i f(s_i x)$ where $a_i \geqslant 0$. Thus all functions of C_f are positive. A function in K_f cannot take a strictly negative value. In particular, the constant one is positive.

(8) $m(\lambda f) = \lambda m(f)$ An element of $C_{\lambda f}$ has the form $g(x) = \sum_i a_i(\lambda f)(s_i x) = \lambda \sum_i a_i f(s_i x)$, so that $C_{\lambda f} = \lambda C_f$ and $K_{\lambda f} = \lambda K_f$. If c is the constant in K_f, then λc is the constant in $K_{\lambda f}$.

(9) $m(L_s f) = m(f)$ An element in $C_{L_s f}$ has the form

$$g(x) = \sum_i a_i (L_s f)(s_i x) = \sum_i f(s^{-1} s_i x). \tag{48.63}$$

This shows that $C_{L_s f} \subset C_f$. The same shows that $C_f \subset C_{L_s f}$.

(10) $m(R_s f) = m(f)$ Same as $m(L_s f) = m(f)$.

(11) m **is linear** We have already shown that $m(\lambda f) = \lambda m(f)$. So let $f, g \in C^0(G,\mathbb{C})$ and let us prove that $m(f+g) = m(f) + m(g)$. Let $\epsilon > 0$; we consider some reals a_i and some elements $s_i \in G$ such that

$$|m(f) - \sum_i a_i f(s_i x)| < \epsilon, \tag{48.64}$$

EQooDDZXooWIPAhX

and we define

$$h(x) = \sum_i a_i g(s_i x). \tag{48.65}$$

We have $h \in C_g$, so that $C_h \subset C_g$ and $K_h \subset K_g$. Since there is only one constant in K_g and only one constant in K_h, we deduce that $m(h) = m(g)$. There exist a function $l \in C_h$ such that $\|h - l\|_\infty < \epsilon$; in other words we can choose $b_j \in \mathbb{R}^+$ and $t_j \in G$ such that

$$|m(h) - \sum_j b_j h(t_j x)| < \epsilon \tag{48.66}$$

for every $x \in G$. We inject therein the definition of h and the fact that $m(h) = m(g)$:

$$|m(g) - \sum_j b_j \sum_i a_i g(s_i t_j x)| < \epsilon. \tag{48.67}$$

EQooZZOSooQwJtip

We multiply (48.64) by b_j, we evaluate it at $t_j x$ and we make the sum over j :

$$|\underbrace{\sum_j b_j m(f)}_{=m(f)} - \sum_{ij} a_i b_j f(s_i t_j x)| < \epsilon. \tag{48.68}$$

Combining with (48.67) we have the majorations :

$$\begin{aligned} |m(f) + m(g) - \sum_{ij} a_i b_j (f+g)(s_i t_j x)| &\leqslant |m(f) - \sum_{ij} a_i b_j f(s_i t_j x)| && \text{(48.69a)} \\ &\quad + |m(g) - \sum_{ij} a_i b_j g(s_i t_j x)| \\ &< 2\epsilon. && \text{(48.69b)} \end{aligned}$$

This proves that the constant $m(f) + m(g)$ belong to K_{f+g}. Since K_{f+g} contains only one constant, which is $m(f+g)$, we deduce that $m(f) + m(g) = m(f+g)$.

Up to now we build a Haar functional $m\colon C^0(G,\mathbb{C})\to\mathbb{C}$ which satisfies $m(R_sf)=m(f)$. We prove the unicity.

Let $m_2\colon C^0(G,\mathbb{C})\to\mathbb{C}$ be an other Haar functional on G. We will prove that $m_2(f)=m(f)$ for every f.

(1) m_2 **is constant on** C_f Let $g\in C_f$. We have $g(x)=\sum_i a_iL_{s_i(f)}(x)$, so that

$$m_2(g)=\sum_i a_im_2\big(L_{s_i}(f)\big)=\sum_i a_im_2(f)=m_2(f). \tag{48.70}$$

We used the left invariance of m_2 and the fact that $\sum_i a_i=1$.

(2) m_2 **is constant on** K_f We know from lemma 48.82 that m_2 is continuous on $C^0(G,\mathbb{C})$. Let $g,h\in K_f$. There exists sequences (g_i) ang (h_i) in C_f such that $g_i\xrightarrow{C^0(G,\mathbb{C})}g$ and $h_i\xrightarrow{C^0(G,\mathbb{C})}h$. For every i we have $m_2(g_i)=m_2(h_i)$ (because m_2 is constant on C_f). By continuity of m_2 we also have $m_2(g)=m_2(h)$.

(3) $m_2=m$ Let $f\in C^0(G,\mathbb{C})$. We consider the constant function

$$\begin{aligned}g\colon G&\to\mathbb{C}\\ x&\mapsto m(f).\end{aligned} \tag{48.71}$$

This function belongs to K_f; this is even, by definition, the only constant function in K_f. Thus $m_2(f)=m_2(g)$ because m_2 is constant on K_f. Now we have

$$m_2(f)=m_2(g)=m_2\big(m(f)\mathbb{1}\big)=m(f)m_2(\mathbb{1})=m(f). \tag{48.72}$$

Thus every Haar functional on G is equal to m and we deduce that there exists only one Haar functional on G. □

48.11.1 Modular function

LEMooLBEZooVuQnan

Lemma 48.84.
If μ and ν are Haar measures on $(G,\mathcal{A})$, then there exists $\lambda\in\mathbb{R}^+$ such that

$$\mu(A)=\lambda\nu(A) \tag{48.73}$$

for every $A\in\mathcal{A}$.

THOooKVLYooBmeMHM

Theorem-Definition 48.85 ([825])**.**
Let G be a compact Hausdorff topological group. There exists a unique group morphism $\Delta\colon G\to\mathbb{R}^+$ such that for every Haar measured space $(G,\mathcal{A},\mu)$ we have

$$\mu(At)=\Delta(t)^{-1}\mu(A) \tag{48.74}$$

for every $t\in G$ and every $A\in\mathcal{A}$.

This morphism is the ***modular function*** *of G.*

Proof. Let $(G,\mathcal{A},\mu)$ be a Haar measure on G. We prove the existence of a unique morphism $\Delta\colon G\to\mathbb{R}^+$ such that $\mu(At)=\Delta(t)^{-1}\mu(A)$. Then we will prove that Δ_0 has the same property for the other Haar measures $(G,\mathcal{A}',\mu')$.

Let $t\in G$. We define

$$\begin{aligned}\nu\colon\mathcal{A}&\to\mathbb{R}^+\\ A&\mapsto\mu(At).\end{aligned} \tag{48.75}$$

It is easy to check that ν is measure[23] on $(G,\mathcal{A})$, among others because the sets A_it are disjoint when the sets A_i are. Moreover ν is left invariant :

$$\nu(gA)=\mu(gAt)=\mu(At)=\nu(A). \tag{48.76}$$

23. Check definition 14.18.

Moreover if K is compact in G, the part Kt is also compact, so that ν is a Haar measure[24] on G.

Lemma 48.84 says that there exists $\lambda \in \mathbb{R}^+$ such that $\nu(A) = \lambda\mu(A)$. We denote by $\Delta(t)^{-1}$ this number. Thus for every $A \in \mathcal{A}$ we have

$$\mu(At) = \Delta(t)^{-1}\mu(A). \tag{48.77}$$

We have to prove that Δ is a morphism. Let $s, t \in G$. We have

$$\mu(Ast) = \Delta(st)^{-1}\mu(A) \tag{48.78}$$

and

$$\mu(Ast) = \Delta(t)^{-1}\mu(As) = \Delta(t)^{-1}\Delta(s)^{-1}\mu(A). \tag{48.79}$$

Thus $\Delta(st)^{-1} = \Delta(s)^{-1}\Delta(t)^{-1}$ and Δ is a morphism.

Now we see that Δ has the property $\mu'(At) = \Delta(t)^{-1}\mu'(A)$ for every Haar measure λ'. From lemma 48.84, there exists λ such that $\mu' = \lambda\mu$. Then

$$\mu'(At) = \lambda\mu(At) = \lambda\Delta(t)^{-1}\mu(A) = \Delta(t)^{-1}\mu(A). \tag{48.80}$$

□

PROPooYUQRooJfoApW

Proposition 48.86 ([825]).
The modular function[25] $\Delta\colon G \to \mathbb{R}^+$ *is continuous and satisfy*

$$\Delta(t)\int_G (st)d\mu(s) = \int_G f(s)d\mu(s). \tag{48.81}$$

LEMooRNXNooOPbzfN

Lemma 48.87 ([825]).
Let G be a compact Hausdorff topological group. Then its modular function is identically equal to 1.

Proof. Suppose that $t \in G$ satisfy $\Delta(t) \neq 1$. We can suppose that $\Delta(t) > 1$; if not, $\Delta(t^{-1}) > 1$. Since Δ is a morphism, we have

$$\Delta(t^n) = \Delta(t)^n \to \infty, \tag{48.82}$$

so that Δ is not bounded.

But wait. The modula function is continuous (proposition 48.86) on the compact G. It has to be bounded. We conclude that $\Delta(t) = 1$ for every $t \in G$. □

48.11.2 Riesz-Markov representation theorem

THOooTWZWooHqGDAx

Theorem 48.88 (Riesz-Markov representation theorem[826, 827, 828, 1]).
Let X be a Hausdorff locally compact space. Let m be a positive linear form on $C_c^0(X, \mathbb{R})$. There exists a unique measured[26] *space $(X, \mathcal{A}, \mu)$ containing the Borel sets such that*

ITEMooAKLQooWLetDk

(1) The integral build on μ satisfies

$$m(f) = \int_X f d\mu \tag{48.83}$$

for every $f \in C_c^0(X, \mathbb{R})$.

ITEMooFLYHooTNUENu

(2) $\mu(K) < \infty$ for every compact set $K \subset X$.

ITEMooKDTLooJuUTaW

(3) If $E \in \mathcal{A}$ and $\mu(E) < \infty$, then

$$\mu(E) = \inf\{\mu(V) \text{ tel que } E \subset V \text{ and } V \text{ is open in } X\}, \tag{48.84}$$

24. Definition 48.81.
25. Definition 48.85.
26. I'm not completely sure about the unicity of the σ-algebra. At least if $(\mathcal{A}_1, \mu_1)$ and $(\mathcal{A}_2, \mu_2)$ satisfy the requirements, we have $\mu_1 = \mu_2$ on $\mathcal{A}_1 \cap \mathcal{A}_2$. Please, contact me if you have an opinion.

ITEMooVKDEooXGJYgg

(4) *If $E \in \mathcal{A}$ and $\mu(E) < \infty$, then*

$$\mu(E) = \sup\{\mu(K) \text{ tel que } K \subset E \text{ and } K \text{ is compact in } X\}, \tag{48.85}$$

ITEMooLOQQooFcHyaM

(5) *The measured space $(X, \mathcal{A}, \mu)$ is complete*[27]

Proof. We initiate with the unicity.

Unicity

Let $\mathcal{A}$ be a σ-algebra containing the Borel set and let μ_1, μ_2 be measures on $(X, \mathcal{A})$ satisfying the conditions.

(1) **Unicity on compacts parts**
Let K be compact in X. Since it is closed, it is a Borel set and $K \in \mathcal{A}$. Let $\epsilon > 0$. By (2) and (3) there exists an open set V such that $\mu_1(V) < \mu_1(K) + \epsilon$. From Urysohn's lemma 15.165, there exists a map $f \in C_c^0(X, \mathbb{R})$ such that $f(K) = \{1\}$ and $f(x) = 0$ when $x \notin V$. We have

SUBEQooWVQVooLbgIDZ

$$\begin{aligned}
\mu_2(K) &= \int_X \mathbb{1}_K d\mu_2 && \text{(48.86a)}\\
&\leqslant \int_X f d\mu_2 && \text{(48.86b)}\\
&= m(f) && \text{(48.86c)}\\
&= \int_X f d\mu_1 && \text{(48.86d)}\\
&\leqslant \int_X \mathbb{1}_V d\mu_1 && \text{(48.86e)}\\
&= \mu_1(V) && \text{(48.86f)}\\
&< \mu_1(K) + \epsilon. && \text{(48.86g)}
\end{aligned}$$

SUBEQooXKADooTpXNlS

Justifications.

— For (48.86a). Link between integral and measure, lemma 14.175.

— For (48.86c). Property (1).

Thus for every $\epsilon > 0$ we have $\mu_2(K) \leqslant \mu_1(K) + \epsilon$. Permuting the roles of μ_1 and μ_2 we also get $\mu_1(K) \leqslant \mu_2(K) + \epsilon$.
We conclude that $\mu_1(K) = \mu_2(K)$.

(2) **Unicity** Let $E \in \mathcal{A}$. We have

$$\{\mu_1(K) \text{ tel que } K \text{ is compact in } V\} = \{\mu_2(K) \text{ tel que } K \text{ is compact in } V\}, \tag{48.87}$$

thus $\mu_1(V) = \mu_2(V)$.

Existence

Existence is the big part. Let us introduce some notations. If $f \in C_c^0(X, \mathbb{R})$ and $E \subset X$ we write

(1) $f \prec E$ if $\text{supp}(f) \subset E$ and $f(X) \subset [0, 1]$.

(2) $E \prec f$ if $f(E) = \{1\}$ and $f(X) \subset [0, 1]$.

27. Definition 14.68.

For open sets $V \subset X$ we define

$$\mu^o(V) = \sup\{m(f) \text{ tel que } f \prec V\}. \tag{48.88}$$

Then, for every part E we define

$$\mu^+(E) = \inf\{\mu^o(V) \text{ tel que } \ E \subset V \text{ and } V \text{ open}\}, \tag{48.89}$$

and

$$\mu^-(E) = \sup\{\mu^+(K) \text{ tel que } \ K \subset E \text{ and } K \text{ compact}\}, \tag{48.90}$$

EQooSIKUooJeeIqJ

and

$$\mathcal{A}_F = \{E \subset X \text{ tel que } \mu^-(E) = \mu^+(E) < \infty\}, \tag{48.91}$$

and

$$\mathcal{A} = \{E \subset X \text{ tel que } E \cap K \in \mathcal{A}_F \, \forall \text{ compact } K \text{ in } X\}, \tag{48.92}$$

and finally, for $E \in \mathcal{A}$ we define

$$\mu(E) = \mu^+(E). \tag{48.93}$$

We have to prove that $(X, \mathcal{A}, \mu)$ is a complete measured space which satisfies all the conditions. We subdivise the proof into an incredible number of steps.

(1) **μ^o is increasing** ITEMooOMDSooQiYmaP Let $V_1 \subset V_2$ be open sets. If $f \prec V_2$, then $f \prec V_1$, so that

$$\{m(f) \text{ tel que } f \prec V_1\} \subset \{m(f) \text{ tel que } f \prec V_2\} \tag{48.94}$$

and we deduce that $\mu^o(V_1) \leqslant \mu^o(V_2)$.

(2) **μ^+ is increasing** ITEMooGODTooQNQzfz Let $E \subset F$. Then $\{V \text{ tel que } F \subset V\} \subset \{V \text{ tel que } E \subset V\}$. The infimum of the largest set is smaller, then $\mu^+(E) \leqslant \mu^+(F)$.

(3) **μ^- is increasing**
Same argument as for μ^+, but with the inclusions inverted.

(4) **If V is open, $\mu^o(V) = \mu^+(V)$** ITEMooBRBTooAJBxGO
Let V be open and let us show $\mu^o(V) = \mu^+(V)$. Per definition

$$\mu^+(V) = \inf\{\mu^o(W) \text{ tel que } \ U \subset W \text{ and } W \text{ is open}\}. \tag{48.95}$$

If W is open with $V \subset W$ we have $\mu^o(V) \leqslant \mu^o(W)$. Thus $\mu^o(V)$ is a lower bound of the set whose infimum is $\mu^+(V)$. Thus $\mu^o(V) \leqslant \mu^+(V)$.
Since $V \subset V$, the number $\mu^o(V)$ is part of the set $\{\mu^o(W) \text{ tel que } \ V \subset W \text{ and } W \text{ is open}\}$. Thus $\mu^+(V) \leqslant \mu^o(V)$.

(5) **For every part, $\mu^-(E) \leqslant \mu^+(E)$** ITEMooIHLQooCjqhPY
Let K be compact and V open such that $K \subset E \subset V$. Using points (2) and (4),

$$\mu^+(K) \leqslant \mu^+(E) \leqslant \mu^+(V) = \mu^o(V). \tag{48.96}$$

In particular, for each open set V containing E we have

$$\mu^-(E) = \sup\{\mu^+(K) \text{ tel que } K \subset E\} \leqslant \mu^o(V). \tag{48.97}$$

Taking the infimum on V we get

$$\mu^-(E) \leqslant \inf\{\mu^o(V) \text{ tel que } E \subset V\} = \mu^+(E). \tag{48.98}$$

(6) **If K is compact, $\mu^+(K) < \infty$** ITEMooQVOGooQvTTzw
From lemma 7.87, we consider an open set V such that $K \subset V$ and $\bar{V}$ is compact. Urysohn's lemma provides a function g such that $\bar{V} \prec g$. If $f \in C_c^0(X, \mathbb{R})$ satisfies $f \prec V$, we have $f \leqslant g$. Indeed if $x \in \bar{V}$, $g(x) = 1 \geqslant f(x)$ and if $x \notin \bar{V}$, then $f(x) = 0 \leqslant g(x)$.
We have

$$\mu^+(V) = \mu^o(V) = \sup\{m(f) \text{ tel que } f \prec V\}. \tag{48.99}$$

And since $K \subset V$ we have

$$\mu^+(K) \leqslant \mu^+(V) = \sup\{m(f) \text{ tel que } f \prec V\} \leqslant m(g) < \infty. \tag{48.100}$$

(7) $\underline{\mu^+(K) = \inf\{m(f) : K \prec f\}}$ Let $\epsilon > 0$. The definition is $\mu^+(K) = \inf\{\mu^o(V)$ tel que $K \subset V\}$. There exists an open set V such that $\mu^+(K) > \mu^o(V) - \epsilon$ and $K \subset V$. From Urysohn's lemma we consider a function $f \in C_c^0(X)$ such that $K \prec f \prec V$. In particular $\mu^o(V) \geqslant m(f)$. Putting the two inequalities together,

$$m(f) \leqslant \mu^o(V) < \mu^+(K) + \epsilon. \tag{48.101}$$

We proved that for every $\epsilon > 0$, there exists f such that $K \prec f$ and $m(f) < \mu^+(K) + \epsilon$. This means that

$$\mu^+(K) \geqslant \inf\{m(f) \text{ tel que } K \prec f\}. \tag{48.102}$$

Now we prove the reverse inequality.
We want to prove that

$$\inf\{\mu^o(V) : K \subset V\} \leqslant \inf\{m(f) : K \prec f\}. \tag{48.103}$$

Let $\epsilon > 0$ and $f \in C_c^0(X)$ such that $K \prec f$. We will build an open set V such that

$$\begin{cases} \mu^o(V) < m(f) + \epsilon & \text{(48.104a)} \\ K \subset V. & \text{(48.104b)} \end{cases}$$

Let $0 < \alpha < 1$ and $V_\alpha = \{x \text{ tel que } f(x) > \alpha\}$. This is an open set satisfying $K \subset V_\alpha$. We have

$$\mu^+(V_\alpha) = \mu^o(V_\alpha) = \sup\{m(g) : g \prec V_\alpha\}. \tag{48.105}$$

Let $g \prec V_\alpha$. For every $x \in V_\alpha$ we have

$$\alpha g(x) \leqslant \alpha < f(x), \tag{48.106}$$

and, if $x \notin V_\alpha$ we have $g(x) = 0$. Thus $\alpha g \leqslant f$ and

$$m(\alpha g) \leqslant m(f). \tag{48.107}$$

That inequality is valid for every $g \prec V_\alpha$, so that

$$\alpha \sup\{m(g) : g \prec V_\alpha\} \leqslant m(f). \tag{48.108}$$

In other words,

$$\alpha \mu^o(V_\alpha) \leqslant m(f) \tag{48.109}$$

for every $0 < \alpha < 1$. For $\alpha = 1 - \epsilon$ we have

$$\mu^o(V_\alpha) \leqslant m(f) + \epsilon \mu^o(V_\alpha). \tag{48.110}$$

Each V_α is contained in the compact $\operatorname{supp}(f)$, if $s = \mu^+\big(\operatorname{supp}(f)\big)$, we have $s < \infty$ and $\mu^o(V_\alpha) = \mu^+(V_\alpha) < s$ for every α. Writing ϵ' instead of ϵ, we have proved that for every ϵ', we have

$$\mu^o(V_{1-\epsilon'}) \leqslant m(f) + s\epsilon'. \tag{48.111}$$

Now let $\epsilon > 0$. Taking ϵ' as small as $\epsilon' s < \epsilon$ we have

$$\mu^o(V_{1-\epsilon'}) \leqslant m(f) + \epsilon' s < m(f) + \epsilon. \tag{48.112}$$

So for every $\epsilon > 0$, there exists an open V_α such that

$$\begin{cases} \mu^o(V_\alpha) < m(f) + \epsilon & \text{(48.113a)} \\ K \subset V_\alpha. & \text{(48.113b)} \end{cases}$$

This shows that

$$\inf\{\mu^o(V) : V \text{ is open and } K \subset V\} < m(f) + \epsilon, \tag{48.114}$$

and than that

$$\mu^+(K) < m(f) + \epsilon. \tag{48.115}$$

Since this is true for every ϵ, we deduce

$$\mu^+(K) \leqslant m(f), \tag{48.116}$$

and then

$$\mu^+(K) \leqslant \inf\{m(f) : K \prec f\}. \tag{48.117}$$

(8) $K \in \mathcal{A}_F$ ITEMooSMHRooIxgdeO We already have $\mu^+(K) < \infty$ (point (6)). The equality $\mu^-(K) = \mu^+(K)$ is from the definition (48.90) itself: the set K realises the supremum.

(9) $K \in \mathcal{A}$ ITEMooNAIAooDnPhPN If E is compact, the set $E \cap K$ is compact (proposition 7.98) and then belongs to $\mathcal{A}_F$. Thus if E is compact, $E \in \mathcal{A}$.

(10) $\mu^+(V) = \mu^o(V) = \mu^-(V)$ ITEMooEDOSooPwvyAO
Let V be open in X. The equality $\mu^+(V) = \mu^o(V)$ is the point (4).
Let $0 < \alpha < \mu^o(V)$. From the definition of $\mu^o(V)$, there exists $f \prec V$ such that $m(f) > \alpha$. We write $K = \text{supp}(f)$; since $f \prec V$ we have $K \subset V$.
Let W be an open set with $K \subset W$. We have $\text{supp}(f) = K \subset W$ and then $f \prec W$, so that $m(f) \leqslant \mu^o(W)$. Here is a summary of the inequalities so far:

$$\alpha \leqslant m(f) \leqslant \mu^o(W). \tag{48.118}$$

These inequalities hold for every open W such that $K \subset W$. Thus we have

$$\alpha \leqslant \inf\{\mu^o(W) : \ W \text{ is open and } K \subset W\} = \mu^+(K) \leqslant \mu^+(V) = \mu^o(V). \tag{48.119}$$

These inequalities hold for every $\alpha < \mu^o(V)$. So for every $\alpha < \mu^o(V)$, there exists a compact $K \subset V$ such that $\alpha \leqslant \mu^+(K) \leqslant \mu^o(V)$. We deduce that

$$\sup\{\mu^+(K) : \ K \text{ compact in } V\} \geqslant \mu^o(V). \tag{48.120}$$

For the reverse inequality, if $K \subset V$ we have $\mu^+(K) \leqslant \mu^+(V) = \mu^o(V)$.

(11) **If $\mu^o(X) < \infty$, then $V \in \mathcal{A}_F$**
Let V be open in X. From points (4) and (10) we have $\mu^o(V) = \mu^+(V) = \mu^-(V)$. Since μ^o is increasing (point (1)), we have

$$\mu^-(V) = \mu^+(V) = \mu^o(V) < \mu^o(X) < \infty, \tag{48.121}$$

so $V \in \mathcal{A}_F$.

(12) $\mu^+(V_1 \cup V_2) \leqslant \mu^+(V_1) + \mu^+(V_2)$ ITEMooLPYWooQNCYTi Let V_1 and V_2 be open sets in X. Since $V_1 \cup V_2$ is open, using point (4) we have

$$\mu^+(V_1 \cup V_2) = \mu^o(V_1 \cup V_2) = \sup\{m(f) : f \prec (V_1 \cup V_2)\}. \tag{48.122}$$

Let $\epsilon > 0$. There exists $g \in C_C^0(X)$ such that $g \prec V_1 \cup V_2$ and

$$\mu^o(V_1 \cup V_2) - \epsilon \leqslant m(g). \tag{48.123}$$ LEMooFYAAooDaOYUy

Let $K = \text{supp}(g)$; we have $K \subset V_1 \cup V_2$.
We consider $V_1 \cup V_2$ as topological vector space with the induced topology. Since $V_1 \cup V_2$ is open in X, the induced topology makes no difficulties: the open parts of $V_1 \cup V_2$ are the same as the open parts of X. The sets V_1 and V_2 themselve form an open covering of $V_1 \cup V_2$. We use a $C_c^0(V_1 \cup V_2)$-partition of unity[28], i.e. functions $\phi_i \colon V_1 \cup V_2 \to [0,1]$ such that $\phi_i \in C_c^0(V_1 \cup V_2)$ and $\text{supp}(\phi_i) \subset V_i$.

28. Theorem 48.70 and lemma 48.68.

We consider the extensions

$$\begin{aligned} h_i\colon X &\to [0,1] \\ x &\mapsto \begin{cases} \phi_i(x) & \text{is } x \in V_i \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{48.124}$$

The functions h_i belong to $C_c^0(X)$ (lemma 17.124). Here are their properties:

(12a) $h_i \in C_c^0(X)$

(12b) $h_i \prec V_i$

(12c) $h_1 + h_2 = 1$ on $V_1 \cup V_2$ and, in particular, on K.

(12d) $h_i g \prec V_i$ because $\text{supp}(h_i) \subset V_i$

(12e) $h_1 g + h_2 g = g$ on K.

We have

$$m(h_i g) \leqslant \sup\{m(f) \text{ tel que } f \prec V_i\} = \mu^o(V_i). \tag{48.125}$$

Using the linearity of m,

$$m(g) = m(h_1 g) + m(h_2 g) \leqslant \mu^o(V_1) + \mu^o(V_2) \tag{48.126}$$

Combining with (48.123) we find

$$\mu^o(V_1 \cup V_2) - \epsilon \leqslant m(g) \leqslant \mu^o(V_1) + \mu^o(V_2). \tag{48.127}$$

Since it is true for every $\epsilon > 0$ we deduce

$$\mu^o(V_1 \cup V_2) \leqslant \mu^o(V_1) + \mu^o(V_2). \tag{48.128}$$

(13) $\mu^+(\bigcup_i E_i) \leqslant \sum_i \mu^+(E_i)$ ITEMooQZTEooJhWFna

Let E_i be parts of X. Let $\epsilon > 0$. There exists an open set V_i such that $E_i \subset V_i$ and

$$\mu^o(V_i) < \mu^+(E_i) + \frac{\epsilon}{2^i}. \tag{48.129}$$

Let $E = \bigcup_i E_i$ and $V = \bigcup_i V_i$. We have $E \subset V$ while V is open as union of open parts. Thus $\mu^+(E) \leqslant \mu^+(V)$. Let $f \prec V$. We have $\mu^+(V) = \mu^o(V) \leqslant m(f)$. Using point (12), and the geometric series [29]

$$\mu^o(V) \leqslant \sum_i \mu^o(V_i) \leqslant \sum_i \left(\mu^+(E_i) + \frac{\epsilon}{2^i}\right) = \sum_i \mu^+(E_i) + \epsilon. \tag{48.130}$$

Since this is valid for every ϵ, we deduce

$$\mu^+(E) \leqslant \mu^+(V) \leqslant \sum_i \mu^+(E_i). \tag{48.131}$$

(14) $\mu^+(A \cup B) = \mu^+(A) + \mu^+(B)$ **when** A, B **are disjoint compacts** ITEMooCVWSooCmPoPL

Let A, B be disjoint compacts. We know from point (8) that $A, B \in \mathcal{A}$, so that $\mu(A) = \mu^+(A) = \mu^-(A)$, and the same for B. Point (13) says that $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$. We have to prove the reverse inequality.

Let $\epsilon > 0$. We consider an open set W such the $K \subset W$ and

$$\mu^+(W) = \mu(K) + \epsilon. \tag{48.132}$$

The set A is closed [30] and W is open. Thus $V = W \backslash A$ is open. Since A and B are disjoint, we also have $B \backslash V$.

29. Proposition 10.135(2).
30. Lemma 7.96.

The set V is open and contains the compact B. By lemma 7.88, there exists an open set V_2 such that

$$B \subset V_2 \subset \bar{V}_2 \subset V \tag{48.133}$$

with $\bar{V}_2$ being compact.

Let $V_1 = W \backslash \bar{V}_2$. Since $\bar{V}_2$ is closed and W is open, V_1 is open. Moreover we have $V_1 \cap V_2 = \varnothing$. Since $A \subset W$ and since $\bar{V}_2 \cap A = \varnothing$, $A \subset V_1$. Let us summarise the important properties of V_1 and V_2 :

— $A \subset V_1$

— $B \subset V_2$

— $V_1 \cap V_2 = \varnothing$.

Let finally

$$V_1' = V_1 \cap \big(W \backslash B\big) \tag{48.134a}$$
$$V_2' = V_2 \cap \big(W \backslash A\big). \tag{48.134b}$$

These open sets have the same properties as V_1 and V_2

Let $f_i \prec V_i'$ such that $\mu^+(V_i') < m(f_i) + \epsilon$. Since $V_1' \cap V_2' = \varnothing$ we also have

$$f_1 + f_2 \prec V_1 \cup V_2 \subset A \cup B \subset W. \tag{48.135}$$

Now we can make the small computation

$$\begin{aligned}
\mu^+(A) + \mu^+(B) &\leqslant \mu^+(V_1') + \mu^+(V_2') && (48.136a)\\
&\leqslant m(f_1) + m(f_2) + 2\epsilon && (48.136b)\\
&= m(f_1 + f_2) + 2\epsilon && (48.136c)\\
&\leqslant \mu^+(W) + 2\epsilon && (48.136d)\\
&\leqslant \mu^+(K) + 3\epsilon. && (48.136e)
\end{aligned}$$

These inequalities hold for every $\epsilon > 0$, then $\mu^+(A) + \mu^+(B) \leqslant \mu^+(K)$.

(15) $\mu^+(\bigcup_{n=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu^+(E_i)$ **is E_i are disjoint in $\mathcal{A}_F$** ITEMooJHIAooVoHLqt

Let $\{E_i\}_{i=1,\dots,\infty}$ be disjoint elements of $\mathcal{A}_F$ We write $E = \bigcup_{i=1}^{\infty}$. Let $\epsilon > 0$. For each i, there exist a compact $K_1 \subset E_i$ such that

$$\mu^+(E_i) - \frac{\epsilon}{2^i} \leqslant \mu^+(K_i). \tag{48.137}$$

Let $H_n = \bigcup_{i=1}^{n} K_i$. We have $H_n \subset E$ for every n and we make the following computation: SUBEQooHCNFooFmH.

$$\begin{aligned}
\sum_{i=1}^{n} \mu^+(E_i) - \epsilon &= \sum_{i=1}^{n} \Big(\mu^+(E_i) - \frac{\epsilon}{2^i}\Big) && \text{SUBEQooILZMooYwLCkC} \quad (48.138a)\\
&\leqslant \sum_{i=1}^{n} \mu^+(K_i) && (48.138b)\\
&= \mu^+\Big(\bigcup_{i=1}^{n} K_i\Big) && \text{SUBEQooSPIXooJqDeCj} \quad (48.138c)\\
&= \mu^+(H_n) && (48.138d)\\
&\leqslant \mu^+(E) && \text{SUNEQooFUWQooEzxDxQ} \quad (48.138e)
\end{aligned}$$

Justification:

— For (48.138a). Geometric series, proposition 10.135(3).

— For (48.138c). Disjoint compacts, point (14).

The inequality $\sum_{i=1}^{n}\mu^+(E_i)-\epsilon\leqslant\mu^+(E)$ holds for every $\epsilon>0$ and every n. Taking the limit $n\to\infty$ and $\epsilon\to 0$ we get

$$\sum_{i=1}^{\infty}\mu^+(E_i)\leqslant\mu^+(E). \tag{48.139}$$

The opposite inequality is the point (13).

(16) E_i **disjoint in** $\mathcal{A}_F$ **and** $\mu^+(E)<\infty$ **implies** $\bigcup_i E_i\in\mathcal{A}_F$ ITEMooKVQIooKqHXGO

Let E_i be disjoint elements of $\mathcal{A}_F$. We write $E=\bigcup_{i=1}^{\infty}E_i$. We suppose that $\mu^+(E)<\infty$, and we prove that $E=\bigcup_{i=1}^{\infty}\in\mathcal{A}_F$.

Let $\epsilon>0$. We already know by (15) that $\mu^+(E)=\sum_{i=1}^{\infty}\mu^+(E_i)$. It means that there exists $N>0$ such that

$$\mu^+(E)<\sum_{i=1}^{N}\mu^+(E_i)+\epsilon. \tag{48.140}$$

The computation (48.138) shows, among other things, that $\sum_{i=1}^{N}\mu^+(E_i)<\mu^+(H_N)+\epsilon$. Thus we have the computation

$$\begin{aligned}
\mu^+(E)&<\sum_{i=1}^{N}\mu^+(E_i)+\epsilon && \text{(48.141a)}\\
&<\big(\mu^+(H_N)+\epsilon\big)+\epsilon && \text{(48.141b)}\\
&=\mu^+(H_N)+2\epsilon && \text{(48.141c)}\\
&\leqslant\sup\{\mu^+(K)\text{ tel que }K\text{ is compact in }E\}+2\epsilon\quad H_N\text{ is compact in }E && \text{(48.141d)}\\
&=\mu^-(E)+2\epsilon. && \text{(48.141e)}
\end{aligned}$$

We proved that $\mu^+(E)\leqslant\mu^-(E)$. Since we always have $\mu^+(E)\geqslant\mu^-(E)$, we deduce that $\mu^+(E)=\mu^-(E)$ and then that $E\in\mathcal{A}_F$.

(17) $\mu^+(V\cup K)=\mu^+(V)+\mu^+(K)$ **if** $V\cap K=\varnothing$ ITEMooZSUIooPrPhtD

Let V be open and K be compact such that $V\cap K=\varnothing$. If L is compact in V, we have $L\cap K=\varnothing$ and then $\mu^+(K\cup L)=\mu^+(K+\mu^+(L)$. Using point (14), we have:

$$\begin{aligned}
\mu^+(V\cup K)&=\sup\{\mu^+(M)\text{ tel que }\ M\text{ is compact and }M\subset V\cup K\ \} && \text{(48.142a)}\\
&\geqslant\sup\{\mu^+(L\cup K)\text{ tel que }\ L\text{ is compact and }L\subset V\ \} && \text{(48.142b)}\\
&=\sup\{\mu^+(K)+\mu^+(L)\text{ tel que }\ L\text{ is compact and }L\subset V\ \} && \text{(48.142c)}\\
&=\mu^+(K)+\sup\{\mu^+(L)\text{ tel que }\ L\text{ is compact and }L\subset V\ \} && \text{(48.142d)}\\
&=\mu^+(K)+\mu^+(V). && \text{(48.142e)}
\end{aligned}$$

We proved that $\mu^+(V\cup K)\geqslant\mu^+(K)+\mu^+(V)$. The reverse inequality is from point (13).

(18) $\mu^+(A\cap V)+\mu^+(A\backslash V)\leqslant\mu^+(A)$, A **open** ITEMooMWJBooCHUrot

Let A,V be open. We consider $K\subset A\cap V$ and we define $W=A\backslash K$. The set W is open and we have the inclusions

$$A\backslash V\subset A\backslash K=W. \tag{48.143}$$

Thus we have the following computation: SUBEQSooOSJAooBgpKec

$$\begin{aligned}
\mu^+(K)+\mu^+(A\backslash V)&\leqslant\mu^+(K)+\mu^+(A\backslash K) && \text{(48.144a)}\\
&=\mu^+\big(K\cup(A\backslash K)\big) && \text{(48.144b)}\quad\text{SUBEQooUEFQooMbLiTC}\\
&=\mu^+(A). && \text{(48.144c)}
\end{aligned}$$

Justifications:

— For (48.144b): point (17).

Taking the supremum over compacts K in $A \cap V$ we have

$$\begin{aligned} \mu^o(A \cap V) + \mu^+(A\backslash V) &= \mu^-(A \cap V) + \mu^+(A\backslash V) && (48.145a) \\ &= \sup\{\mu^+(K) \text{ tel que } K \text{ is compact and } K \subset A \cap V\} && (48.145b) \\ &\leqslant \mu^+(A). && (48.145c) \end{aligned}$$

SUBEQooHHXTooAcuBNJ SUBEQooRLOHooVCPVeG

Justifications:

— For (48.145a): point (10)

— For (48.145c): inequality (48.144)

(19) $\mu^+(E \cap V) + \mu^+(E\backslash V) \leqslant \mu^+(E)$, E **arbitrary** ITEMooHQPDooPEsgYN

Let E be an arbitrary subset of X. Let W be an open set containing E: $E \subset W$. We have

$$\begin{aligned} \mu^+(E \cup V) + \mu^+(E\backslash V) &\leqslant \mu^+(W \cap V) + \mu^+(W\backslash V) && (48.146a) \\ &\leqslant \mu^+(W) && (48.146b) \\ &= \mu^+(W). && (48.146c) \end{aligned}$$

ITEMooLMDTooTDppPU

Justifications:

— For (48.146b), point (18).

Taking the supremum,

$$\mu^+(E \cap V) + \mu^+(E\backslash V) \leqslant \inf\{\mu^o(W) \text{ tel que } \ W \text{ is open, } E\backslash W \ \} = \mu^+(E). \tag{48.147}$$

(20) **If $E \in \mathcal{A}_F$, there exists $K \subset E \subset V$ such that $\mu^+(V\backslash K) < \epsilon$** ITEMooTWGLooHVvkUW

Let $E \in \mathcal{A}_F$ and $\epsilon > 0$. By definition, $\mu^-(E) = \mu^+(E) < \infty$. Thus:

(20a) There exists compact K such that $K \subset E$ and $\mu^-(E) - \epsilon < \mu^+(K)$.

(20b) There exists open V such that $E \subset V$ and $\mu^+(E) + \epsilon < \mu^o(V)$.

Thus we have $K \subset E \subset V$ with $\mu^o(V) - \epsilon \leqslant \mu^+(E) \leqslant \mu^+(K) + \epsilon$, and in particular

$$\mu^+(V) - \mu^+(K) \leqslant 2\epsilon. \tag{48.148}$$

EQooKILTooVSeQsv

Using point (19) we have

$$\mu^+(V \cap K) + \mu^+(V\backslash K) \leqslant \mu^+(V) \tag{48.149}$$

and then

$$\mu^+(K) + \mu^+(V\backslash K) \leqslant \mu^+(V), \tag{48.150}$$

and combining with (48.148),

$$\mu^+(V\backslash K) \leqslant \mu^+(V) - \mu^+(K) < 2\epsilon. \tag{48.151}$$

(21) **If $A, B \in \mathcal{A}_F$, then $A\backslash B \in \mathcal{A}_F$** ITEMooDSKEooPQmrcW

First $A\backslash B \subset A$, while, by hypothesis, $A \in \mathcal{A}_F$. Thus, using the point (2),

$$\mu^+(A\backslash B) \leqslant \mu^+(A) < \infty. \tag{48.152}$$

We still have to prove that $\mu^+(A\backslash B) = \mu^-(A\backslash B)$. Let $\epsilon > 0$ and V_A, V_B, K_A, K_B be as in point (20). We have

$$A\backslash B \subset (V_A\backslash K_B) \cup (K_A\backslash V_B) \cup (V_B\backslash K_B). \tag{48.153}$$

Using now point (13) on this decomposition, we have an estimation of $\mu^+(A\backslash B)$: SUBEQSooCULSooCFFloL

$$\begin{aligned} \mu^+(A\backslash B) &\leqslant \mu^+\Big((V_A\backslash K_B) \cup (K_A\backslash V_B) \cup (V_B\backslash K_B)\Big) && (48.154a) \\ &\leqslant \mu^+(V_A\backslash K_A) + \mu^+(K_A\backslash V_B) + \mu^+(V_B\backslash K_B) && (48.154b) \\ &\leqslant 2\epsilon + \mu^+(K_A\backslash V_B). && (48.154c) \end{aligned}$$

The part $K_A \backslash V_B$ is closed in the compact K_A; it is thus compact by lemma 7.96(1). So $K_A \subset V_B$ is a compact inside A. We can continue the inequalities (48.154) :

$$\begin{aligned}
\mu^+(A\backslash B) &\leqslant 2\epsilon + \mu^+(K_A\backslash V_B) && \text{(48.155a)}\\
&\leqslant 2\epsilon + \sup\{\mu^+(K) \text{ tel que } K \text{ is compact and } K \subset A\backslash B\} && \text{(48.155b)}\\
&= 2\epsilon + \mu^-(A\backslash B). && \text{(48.155c)}
\end{aligned}$$

Since this is true for every ϵ, we have

$$\mu^+(A\backslash B) \leqslant \mu^-(A\backslash B). \tag{48.156}$$

From point (5) we deduce $\mu^+(A\backslash B) = \mu^-(A\backslash B)$ and then $A\backslash B \in \mathcal{A}_F$.

(22) **If $A, B \in \mathcal{A}_F$, then $A \cup B \in \mathcal{A}_F$**
We have the decomposition $A \cup B = (A\backslash B) \cup B$ where $A\backslash B$ and B are disjoint elements of $\mathcal{A}_F$ (see point (21)). By (15) we have

$$\mu^+(A \cup B) = \underbrace{\mu^+(A\backslash B)}_{<\infty} + \underbrace{\mu^+(B)}_{<\infty} < \infty \tag{48.157}$$

Point (16) concludes that $A \cup B \in \mathcal{A}_F$.

(23) **If $A, B \in \mathcal{A}_F$, then $A \cap B \in \mathcal{A}_F$** ITEMooGJBSooQsLmAg
Simply write $A \cap B = A\backslash(A\backslash B)$. Using twice the point (21), we get the result.

(24) $\mathcal{A}_F \subset \mathcal{A}$ ITEMooPOJQooCwbpkE
Let $E \in \mathcal{A}_F$. We show that $E \in \mathcal{A}$. Let K be compact in X; by point (9), $K \in \mathcal{A}_F$. Now by point (23), $E \cap K \in \mathcal{A}_F$, which means that $E \in \mathcal{A}$.

(25) $\mathcal{A}_F = \{E \in \mathcal{A} \textbf{ tel que } \mu^+(E) < \infty\}$ ITEMooNXZTooOyKbfy
Let $E \in \mathcal{A}$ be such that $\mu^+(E) < \infty$. We have to prove that $\mu^+(E) = \mu^-(E)$. Let $\epsilon > 0$. There exist an open set V such that $\mu^o(V) < \mu^+(E) + \epsilon$. Since $\mu^+(E) < \infty$, we have $V \in \mathcal{A}_F$. From point (20) applied to V itself, there exists an open set V' and a compact set K such that $K \subset V \subset V'$ and $\mu^+(V'\backslash K) < \epsilon$. Since $V \subset V'$ we also have

$$\mu^+(V\backslash K) < \epsilon. \tag{48.158}$$

Since K is compact we have $K \in \mathcal{A}_F$ and then $E \cap K \in \mathcal{A}_F$ (points (8) and (23)). There exists a compact H in $E \cap K$ such that

$$\mu^-(E \cap K) < \mu^+(H) + \epsilon. \tag{48.159}$$

From the "decomposition" $E \subset (E \cap K) \cup (V\backslash K)$ we have

$$\begin{aligned}
\mu^+(E) &\leqslant \mu^+(E \cap K) + \mu^+(V\backslash K) && \text{(48.160a)}\\
&\leqslant \mu^+(H) + 2\epsilon && \text{(48.160b)}\\
&\leqslant \sup\{\mu^+(M) \text{ tel que } M \text{ compact in } E\} + 2\epsilon && \text{(48.160c)}\\
&= \mu^-(E) + 2\epsilon. && \text{(48.160d)}
\end{aligned}$$

Thus we have $\mu^+(E) \leqslant \mu^-(E)$ and thus $\mu^+(E) = \mu^-(E)$.

(26) $X \in \mathcal{A}$ ITEMooOQZLooMqPYTP
Let K be compact. We have $X \cap K = K \in \mathcal{A}_F$.

(27) **If $E \in \mathcal{A}$, then $X\backslash E \in \mathcal{A}$** ITEMooCJRFooJAEZIV
Let $E \in \mathcal{A}$ and K be compact in X. Since $E \in \mathcal{A}$, we have $K \cap E \in \mathcal{A}$. The trick lies in the equality

$$K \cap E^c = K\backslash(E \cap K). \tag{48.161}$$

We know that K and $E \cap K$ belong to $\mathcal{A}_F$, so the whole right hand side belong to $\mathcal{A}_F$.

(28) **If $A_k \in \mathcal{A}$, then $A = \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$** ITEMooUGTIooWqVFIo

Let $A_k \in \mathcal{A}$ and K be compact. We have to show that

$$\big(\bigcup_{k=1}^{\infty} A_k\big) \cap K \in \mathcal{A}_F. \tag{48.162}$$

We build the auxiliary sequence (B_k) in the following way [31]:

$$\begin{cases} B_1 = A_1 \cap K & \text{(48.163a)} \\ B_n = (A_n \cap K) \backslash \big(B_1 \cup \ldots \cup B_{n-1}\big). & \text{(48.163b)} \end{cases}$$

Using the fact that $A_k \in \mathcal{A}$ for every k (in particular $A_n \cap K \in \mathcal{A}_F$), we prove by recursion that $B_n \in \mathcal{A}_F$ for every n. Moreover, the B_n are disjoint. A key observation:

$$K \cap A_n = \bigcup_{k=1}^{n} B_n. \tag{48.164}$$

Here is a computation:

$$\begin{aligned} K \cap A &= K \cap \big(\bigcup_{k=1}^{\infty} A_k\big) && \text{(48.165a)} \\ &= \bigcup_{k=1}^{\infty} (K \cap A_k) && \text{(48.165b)} \\ &= \bigcup_{k=1}^{\infty} B_k \cup (B_1 \cup \ldots B_{k-1}) && \text{(48.165c)} \\ &= \bigcup_{k=1}^{\infty} B_k. && \text{(48.165d)} \end{aligned}$$

Since $K \cap A$ is the disjoint union of the B_k, point (15) gives

$$\begin{aligned} \mu^+(A \cap K) &= \sum_{k=1}^{\infty} \mu^+(B_n) && \text{(48.166a)} \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \mu^+(B_n) && \text{(48.166b)} \\ &= \lim_{n\to\infty} \mu^+\big(\bigcup_{k=1}^{n} B_n\big) && \text{(48.166c)} \\ &= \lim_{n\to\infty} \mu^+(K \cap A_n) && \text{(48.166d)} \\ &\leqslant \mu^+(K) && \text{(48.166e)} \\ &< \infty. && \text{(48.166f)} \end{aligned}$$

The sets B_k are disjoint and $\mu^+(\bigcup_{k=1}^{\infty} B_k) < \epsilon$. Point (16) shows that

$$K \cap A = \bigcup_{k=1}^{\infty} B_k \in \mathcal{A}_F. \tag{48.167}$$

Thus $A \in \mathcal{A}$.

(29) **$\mathcal{A}$ is a σ-algebra containing all the Borel sets**

The fact that $\mathcal{A}$ is a σ-algebra [32] is already checked by points (26), (27) and (28).

Let C be closed in X. If K is compact, $K \cap C$ is closed in the compact K and is then compact. Thus $K \cap C \in \mathcal{A}_F$. Thus $C \in \mathcal{A}$.

By complementarity (point (27)), the σ-algebra $\mathcal{A}$ contains all the open sets and then all the Borel sets.

31. See 1.49 to make it formal.
32. Definition 14.1.

(30) **μ is a positive measure on $\mathcal{A}$**

We have to check the conditions of the definition 14.18.

(30a) **$\mu(0) = 0$** We have

$$\begin{aligned}\mu(\varnothing) &= \mu^+(\varnothing) && (48.168a)\\ &= \inf\{\mu^o(V) \text{ tel que } V \text{ is open}\} && (48.168b)\\ &\leqslant \mu^o(\varnothing) && (48.168c)\\ &= m(0) && (48.168d)\\ &= 0. && (48.168e)\end{aligned}$$

SUBEQooFITMooZexFio SUBEQooFMDQooDWPdDk

For (48.168c), the set $\varnothing$ is open. For (48.168d), function $f = 0$ is the only one to satisfy $f \prec \varnothing$.

(30b) **$\mu(E) \geqslant 0$ for every $E \in \mathcal{A}$** The function μ^+ is increasing while $\mu^+(\varnothing) = 0$ as we have just seen.

(30c) **Countable additive** Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of measurable sets (i.e. elements of $\mathcal{A}$). We have to show that

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k). \tag{48.169}$$

EQooNYATooWKSkoG

If $A_k \in \mathcal{A}_F$ for all k, the equality (48.169) holds by point (15).

Now suppose that $A_n \in \mathcal{A}\backslash\mathcal{A}_F$ for some $n \in \mathbb{N}$. By point (25) we have $\mu(A_n) = \infty$, and then the right hand side of (48.169) is ∞.

Since $A_n \subset \bigcup_{k=1}^{\infty} A_k$, since μ^+ is increasing [33],

$$\infty = \mu(A_n) = \mu(A_n) \leqslant \mu\Big(\bigcup_{k=1}^{\infty} A_k\Big), \tag{48.170}$$

so that $\mu\big(\bigcup_{k=1}^{\infty} A_k\big) = \infty$.

From now on we can say that $(X, \mathcal{A}, \mu)$ is a measure space. This is quite a achievement, and you are allowed to relax yourself having a cup of tea.

Done ? We continue.

(31) **$(X, \mathcal{A}, \mu)$ is a complete measure space**

Let $E \in \mathcal{A}$ satisfying $\mu(E) = 0$. Let $A \subset E$. We have to show that $A \in \mathcal{A}$. Since μ^+ is increasing we have

$$\mu^+(A) \leqslant \mu^+(E) = 0. \tag{48.171}$$

Thus $\mu^+(A) = 0$. The same works with μ^-: since $E \in \mathcal{A}_F$ we have $\mu^-(E) = \mu^+(E)$ and

$$\mu^-(A) \leqslant \mu^-(E) = \mu^+(E) = 0. \tag{48.172}$$

Thus we have $\mu^+(A) = \mu^-(A) = 0 < \infty$ and $A \in \mathcal{A}_F \subset \mathcal{A}$. We used point (24).

(32) **Summary**

The points (2), (3), (4) and (5) are proved. We still have to show the point (1) about the relation between m and the integral.

(33) **We have $m(f) \leqslant \int_X f d\mu$ for every $f \in C_c^0(X)$**

Let $f \in C_c^0(X)$ and denote by K its support. First we show that $m(f) \leqslant \int_X f d\mu$. Since f is continuous on the compact K, it is bounded; let $a, b \in \mathbb{R}$ such that $f(X) \subset [a, b]$. Let $\epsilon > 0$ and consider numbers $\{y_i\}_{i=0,\dots,n}$ such that

$$y_0 < a < y_1 < \ldots < y_n = b \tag{48.173}$$

33. We have $\mu^+ = \mu$ for every sets we are interested in.

and $y_i - y_{i-1} < \epsilon$. We consider the sets

$$E_i = f^{-1}\big(]y_{i-1}, y_i]\big). \tag{48.174}$$

Since f is continuous, it is Borel measurable and the sets E_i are Borel sets. In particular $E_i \in \mathcal{A}$. The sets E_i are covering K without intersections, so $\sum_{i=1}^n \mu(E_i) = \mu(K)$.
Now consider open sets V_i' such that $\mu(V_i') < \mu(E_i) + \frac{\epsilon}{n}$ and $E_i \subset V_i'$. Then define

$$V_i = V_i' \backslash f^{-1}\big(]y_i + \epsilon, \infty[\big). \tag{48.175}$$

The set V_i is open and satisfy

$$\mu(V_i) \leqslant \mu(V_i') < \mu(E_i) + \frac{\epsilon}{n} \tag{48.176}$$

while $E_i \subset V_i$.
Since the sets E_i form a covering of $K = \operatorname{supp}(f)$, the open sets V_i are an open covering of the compact K. We consider a partition of unity (theorem 27.75). These are functions $h_i \in C^0(X)$ such that $h_i \prec V_i$ and

$$\sum_{i=1}^n h_i(x) = 1 \tag{48.177}$$

for every $x \in K$. We have

$$h_i f \leqslant h_i(y_i + \epsilon) \tag{48.178}$$

and we start with a long computation:

$$\begin{aligned}
m(f) &= \sum_{i=1}^n m(h_i f) && (48.179a)\\
&\leqslant \sum_{i=1}^n (y_i + \epsilon) m(h_i) && (48.179b)\\
&= \sum_{i=1}^n (|a| + y_i + \epsilon) m(h_i) - |a| \sum_{i=1}^n m(h_i) && (48.179c)
\end{aligned}$$

Since $h_i \prec V_i$ we have $m(h_i) \leqslant \mu(V_i) < \mu(E_i) + \frac{\epsilon}{n}$ and we continue the computation:

$$m(f) \leqslant \sum_{i=1}^n (|a| + y_i + \epsilon)\big(\mu(E_i) + \frac{\epsilon}{n}\big) - |a| m\big(\sum_i h_i\big) \tag{48.180}$$

Since $\sum_i h_i = 1$ on K we have $K \prec \sum_i h_i$ and $m(\sum_i h_i) \geqslant \mu(K) = \sum_{i=1}^n \mu(E_i)$, so that

$$\begin{aligned}
m(f) &\leqslant \sum_{i=1}^n (|a| + y_i + \epsilon)\big(\mu(E_i) + \frac{\epsilon}{n}\big) - |a| m\big(\sum_i h_i\big) && (48.181a)\\
&= \sum_i (|a| + y_i + \epsilon)\big(\mu(E_i) + \frac{\epsilon}{n}\big) - |a|\mu(K) && (48.181b)\\
&= \sum_i (|a| + y_i + \epsilon)\mu(E_i) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(K) && (48.181c)\\
&= \sum_i (|a| + y_i - \epsilon)\mu(E_i) + 2\sum_i \epsilon\mu(E_i) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(K) && (48.181d)\\
&= \sum_i (|a| + y_i - \epsilon)\mu(E_i) + 2\epsilon\mu(K) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(K) && (48.181e)\\
&= \sum_i |a|\mu(E_i) + \sum_i (y_i - \epsilon)\mu(E_i) + 2\epsilon\mu(K) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(K) && (48.181f)\\
&= |a|\mu(K) + \sum_i (y_i - \epsilon)\mu(E_i) + 2\epsilon\mu(K) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(K) && (48.181g)\\
&= \sum_i (y_i - \epsilon)\mu(E_i) + 2\epsilon\mu(K) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} && (48.181h)\\
& && (48.181i)
\end{aligned}$$

We find a majoration for the first and third terms. For the first term, recall that for every $x \in E_i$ we have

$$f(x) > y_{i+1} > y_i - \epsilon, \tag{48.182}$$

so that

$$\sum_{i=1}^{n}(y_i - \epsilon)\mu(E_i) \leqslant \int_X f d\mu. \tag{48.183}$$

And for the last,

$$\frac{\epsilon}{n}\sum_{i=1}^{n}(|a| + y_i + \epsilon) = \frac{\epsilon}{n}\sum_{i=1}^{n}(|a| + \epsilon) + \frac{\epsilon}{n}\sum_i y_i \tag{48.184a}$$

$$\leqslant \epsilon(|a| + \epsilon) + \frac{\epsilon}{n}nb \tag{48.184b}$$

$$= \epsilon(|a| + \epsilon + b) \tag{48.184c}$$

because each of the y_i satisfy $y_i \leqslant b$.
We continue the majoration of $m(f)$:

$$m(f) \leqslant \sum_i (y_i - \epsilon)\mu(E_i) + 2\epsilon\mu(K) + \sum_i (|a| + y_i + \epsilon)\frac{\epsilon}{n} \tag{48.185a}$$

$$\leqslant \int_X f d\mu + 2\epsilon\mu(K) + \epsilon(|a| + \epsilon + b) \tag{48.185b}$$

$$= \int_X f d\mu + \epsilon(2\mu(K) + |a| + \epsilon + n) \tag{48.185c}$$

Taking the limit $\epsilon \to 0$ we get

$$m(f) \leqslant \int_X f d\mu. \tag{48.186}$$

That inequality is correct for every $f \in C_c^0(X)$.
Now we prove the opposite inequality. We have

$$-m(f) = m(-f) \leqslant \int_X (-f)d\mu = -\int_X f d\mu. \tag{48.187}$$

Removing the negative signs and flipping the inequality,

$$\int_X df\mu \leqslant m(f). \tag{48.188}$$

We proved that $m(f) = \int_X f d\mu$.

□

ThoBZBooOTxqcI

Theorem 48.89.
Let G be a Hausdorff compact topological.

ITEMooVFCCooXevpMx

(1) G accepts a unique normalized[34] *Haar measure.*

ITEMooWDIHooHEkcDz

(2) The Haar measure is unimodular:

$$\mu(Ag) = \mu(gA) = \mu(A) \tag{48.189}$$

for every measurable subset $A \subset G$ and every $g \in G$.

ITEMooJQMBooCsSWws

(3) For every $s \in G$, the integral satisfies

$$\int_G f(xs)d\mu(x) = \int_G f(sx)d\mu(x) = \int_G f(x)d\mu(x). \tag{48.190}$$

EQooGHNFooZMEcob

34. Normalized in the sense that $\mu(G) = 1$.

Each time we will speak about a Haar measure on a compact Lie group, we are speaking about this one.

Proof. The theorem 48.83 says that G accepts a unique Haar functional $m\colon C^0(G,\mathbb{C})\to\mathbb{C}$. We also know from lemma 48.82 that this functional is continuous for the supremum norm on $C^0(G,\mathbb{C})$. We restrict m to the real functions. The Riesz-Markov representation theorem 48.88 provides the measured space $(G,\mathcal{A},\mu)$ such that

$$m(f)=\int_G f d\mu. \tag{48.191}$$

Since G is compact, we have $\mu(G)<\infty$ (this is the condition 48.81(2)). If $\mu(G)\neq 1$, we consider $\mu'=\mu/\mu(G)$. This measure is normalized and shows the existence part of item (1).

As far as the unicity is concerned, we recall the lemma 48.84. If μ and ν are normalized Haar measures, we have $\mu=\lambda\nu$ and then $1=\mu(G)=\lambda\nu(G)=\lambda$, so that $\lambda=1$.

Now we prove item (2). The equality $\mu(gA)=\mu(A)$ is a requirement for every Haar measure. The specific feature to prove here is $\mu(Ag)=\mu(A)$. Using the associated integral,

$$\mu(Ag)=\int_G \mathbb{1}_{Ag}(t)d\mu(t). \tag{48.192}$$

Since $\mathbb{1}_{Ag}(x)=\mathbb{1}_A(xg^{-1})$ for every $x\in G$ we can compute

$$\begin{aligned}
\mu(Ag)&=\int_G \mathbb{1}_{Ag}(t)d\mu(t) && && (48.193a)\\
&=\int_G \mathbb{1}_A(tg^{-1})d\mu(A) && && (48.193b)\\
&=\Delta(g^{-1})\int_G \mathbb{1}_A(t)d\mu(t) && \text{prop. 48.86} && (48.193c)\\
&=\Delta(g^{-1})\mu(A) && && (48.193d)\\
&=\mu(A) && \text{lem. 48.87} && (48.193e)
\end{aligned}$$

Now we prove item (3). First using the fact that $m(L_{g^{-1}}f)=m(f)$ (inside definition 48.80) we have

$$\begin{aligned}
\int_G f(sx)d\mu(x)&=\int_G (L_{s^{-1}}f)(x)d\mu(x) && (48.194a)\\
&=m(L_{s^{-1}}f) && (48.194b)\\
&=m(f) && (48.194c)\\
&=\int_G f(x)d\mu(x). && (48.194d)
\end{aligned}$$

For the other equality,

$$\begin{aligned}
\int_G f(xs)d\mu(w)&=\Delta(s^{-1})\int_G f(x)d\mu(x) && \text{prop. 48.86} && (48.195a)\\
&=\int_G f(x)d\mu(x) && \text{lem. 48.87.} && (48.195b)
\end{aligned}$$

□

THOooCXARooScQJdU

Theorem 48.90 ([829])**.**
Every representation of a compact Hausdorff topological group an be made unitary with a choice of the scalar product on the representation space.

Proof. Let (ρ, V) be a finite dimensional representation of the compact Hausdorff topological group G. We consider the Haar measure on G whose existence is guaranteed by theorem 48.89.

Let $\langle .,. \rangle$ be an hermitian product [35] on V. It is very unlikely that $\rho(s)$ is unitary on $\big(V, \langle .,. \rangle\big)$ for every $s \in G$. But one can define a new hermitian product $\cdot$ such that $\rho(s)$ is unitary on $(V, \cdot)$. Here it is:

$$v \cdot w = \int_G \langle \rho(g)v, \rho(g)w \rangle d\mu(g). \tag{48.196}$$

Using the linearity of $\rho(g)$, the linearity of the integral and the properties of the hermitian product $\langle .,. \rangle$, we see that $\cdot$ is an hermitian product on V.

Now we show that $\rho(s)$ is unitary. Let $v, w \in V$. For the sake of simplicity, define

$$\begin{aligned} g \colon G &\to \mathbb{C} \\ g &\mapsto \langle \rho(g)v, \rho(g)w \rangle. \end{aligned} \tag{48.197}$$

We have

$$\rho(s)v \cdot \rho(s)w = \int_G \langle \rho(g)\rho(s)v, \rho(g)\rho(s)w \rangle d\mu(g) \tag{48.198a}$$

$$= \int_G f(gs) d\mu(g) \tag{48.198b}$$

$$= \int_G f(g) d\mu(g) \qquad \text{eqs. (48.190)} \tag{48.198c}$$

$$= v \cdot w. \tag{48.198d}$$

□

35. Definition 8.176.

Chapter 49

Manifolds

Chapitre_FB

49.1 Topological manifold

We will often define the manifold structure by giving the charts. By doing so, we do not provide directly the topology, but we are referring to theorem 49.10. The following definition is seldom directly checked.

DEFooJOMAooZscKwn

Definition 49.1 ([830])**.**
*A **topological manifold** of dimension n is a topological space M such that:* ITEMooZZXHooPICcjz

(1) M is Hausdorff, ITEMooKOSMooSawqbB

(2) M is second-countable[1] ITEMooJOPDooPiOadZ

(3) For each $m \in M$, there exists (E, U, φ, V) where
  (3a) E is a normed vector space of dimension n
  (3b) U is open in E,
  (3c) V is open in M and contains m,
  (3d) the map $\varphi\colon U \to V$ is an homeomorphism[2]*.*

49.2.
Most of the sources only define manifolds with maps from $\mathbb{R}^n$. Here we allow the charts being from an arbitrary normed vector space. We will see that this does not change the existing manifolds (up to isomorphism). In some cases however the charts are easier to be defined from other vector spaces than $\mathbb{R}^n$. As an example, the charts for $\mathrm{GL}(n, \mathbb{R})$ will be taken from $\mathrm{End}(\mathbb{R}^n)$, not from $\mathbb{R}^{n^2}$.

In the same way, tensor product of vector bundle take their maps from $\mathbb{R}^{k_1} \otimes \mathbb{R}^{k_2}$ instead of $\mathbb{R}^{k_1 k_2}$.

49.3.
Since there is no set containing all normed vector spaces, we cannot consider the set of all charts of a given manifold.

DEFooKUJWooXHoCFI

Definition 49.4.
*Let M be a topological manifold. A **chart** on M is a couple (U, φ) where U is an open part in a normed vector space, and $\varphi\colon U \to M$ is homeomorphic between U and $\varphi(U)$.*

49.5.
Some notations.

(1) We will often write $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ a set of charts such that $M = \bigcup_{\alpha\in\Lambda} \varphi_\alpha(U_\alpha)$. Such a set is an **atlas** of M.

(2) If $\varphi_\alpha\colon U_\alpha \to M$ is a chart, we write $M_\alpha = \varphi_\alpha(U_\alpha)$.

1. Has a countable basis of topology.
2. Definition 7.41. This is an isomorphism of topological spaces. Bijective, continuous and continuous inverse.

49.1.1 Isomorphisms

DEFooZPUEooPcSxNs

Definition 49.6 ([1]).
Let M and N be manifolds. A topological space isomorphism[3] *$f\colon M \to N$ is an **isomorphism** if there exist atlases $\{\varphi_\alpha\}_{\alpha\in\Lambda}$ and $\{\phi_i\}_{i\in I}$ such that*

$$\phi_i^{-1} \circ f \circ \varphi_\alpha\colon \varphi_\alpha^{-1}(M_{\alpha i}) \to (\phi_i^{-1} \circ f)(M_{\alpha i}) \tag{49.1}$$

is an isomorphism for every $\alpha \in \Lambda$ and $i \in I$.

This definition requires

$$\phi_i^{-1} \circ f \circ \varphi_\alpha\colon \varphi_\alpha^{-1}(M_{\alpha i}) \to (\phi_i^{-1} \circ f)(M_{\alpha i}) \tag{49.2}$$

for an atlas. We see now that this is enough. In some sense, if f is an isomorphism with respect to an atlas, this is an isomorphism with respect to every chart.

PROPooNOZNooZFYOwK

Proposition 49.7 ([1]).
Let $f\colon M \to N$ be a topological manifold isomorphism. For every charts $\varphi\colon U \to M$ and $\phi\colon V \to N$, we consider $M' = \varphi(U) \cap \phi(V)$. Then the map

$$\phi^{-1} \circ f \circ \varphi\colon \varphi^{-1}(M') \to (\phi^{-1} \circ f)(M') \tag{49.3}$$

is a topological spaces isomorphism.

Now we prove that every manifold is isomorphic to a manifold with maps from $\mathbb{R}^n$.

PROPooNVVSooAVVLOM

Proposition 49.8 ([1]).
If $\big(M, \{\varphi_\alpha\}_{\alpha\in\Lambda}\big)$ is a n-dimensional topological manifold, there exists maps $\{\phi_\alpha\colon V_\alpha \to M\}_{\alpha\in\Lambda}$ such that V_α is open in $\mathbb{R}^n$ and such that $\big(M, \{\phi_\alpha\}_{\alpha\in\Lambda}\big)$ is isomorphic to $\big(M, \{\varphi_\alpha\}_{\alpha\in\Lambda}\big)$.

Proof. Let $\alpha \in \Lambda$. There is a normed vector space E_α (dimension n), an open part U_α an injective chart $\varphi_\alpha\colon U_\alpha \to M$. By proposition 10.54, we have an open part $V_\alpha \subset \mathbb{R}^n$ and a bijective isometry $g_\alpha\colon V_\alpha \to U_\alpha$. We consider the map

$$\begin{aligned} \phi_\alpha\colon V_\alpha &\to M \\ x &\mapsto (\varphi_\alpha \circ g_\alpha)(x). \end{aligned} \tag{49.4}$$

The map $\phi_\alpha\colon V_\alpha \to \phi_\alpha(V_\alpha)$ is homeomorphic as composition of homeomorphic maps. We still have to prove that we have an isomorphism of manifolds. More specifically we prove that $\mathrm{Id}\colon M \to M$ is an isomorphism of manifolds.

Let $\alpha, \beta \in \Lambda$ and consider $M_{\alpha\beta} = \varphi_\alpha(U_\alpha) \cap \phi_\alpha(V_\alpha)$. We have to prove that

$$\phi_\beta^{-1} \circ \mathrm{Id} \circ \varphi_\alpha\colon M_{\alpha\beta} \to (\phi_\beta^{-1} \circ \mathrm{Id})(M_{\alpha\beta}) \tag{49.5}$$

is homeomorphic. Since $\phi_\beta^{-1} \circ \varphi_\alpha = g_\alpha^{-1} \circ \varphi_\alpha^{-1} \circ \varphi_\alpha = g_\alpha^{-1}$, this is an homomorphism. □

There are at least three occasions in which we consider charts that are not from $\mathbb{R}^n$.

(1) The charts for $\mathrm{GL}(n, \mathbb{R})$ will be defined on $\mathrm{End}(\mathbb{R}^n)$.

(2) The charts of a tensor product of vector bundle will be defined from $(U_\alpha \times \mathbb{R}^{k_1} \otimes \mathbb{R}^{k_2})$ instead of $U_\alpha \times \mathbb{R}^{k_1 k_2}$.

(3) In some case, the charts for Lie subgroups will be defined from the Lie algebra.

3. Homeomorphism : continuous with continuous inverse.

PROPooPYLIooFpCPcu

Proposition 49.9 ([1]).
Let M be a topological manifold with charts $\varphi_\alpha : U_\alpha \to M$ where U_α is open in (V_α, N_α) is a normed vector space[4].

For each α we consider an other norm N'_α. Then the manifold with the maps

$$\phi_\alpha : (U_\alpha, N'_\alpha) \to M \tag{49.6}$$

is isomorphic to M.

49.1.2 Manifold from maps

THOooFIHIooLiSUxH

Theorem 49.10 ([487, 1]).
Let X be a set. We consider maps $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ such that

ITEMooDWSWooWdcDdI
(1) U_α is open in a normed vector space E_α.

ITEMooPEXDooNuJBKH
(2) $\bigcup_{\alpha\in\Lambda} \varphi_\alpha(U_\alpha) = X$,

ITEMooSRPQooNUPzlj
(3) For each $\alpha \in \Lambda$, the map φ_α is injective,

ITEMooWFAWooAqQfzZ
(4) For every $\alpha, \beta \in \Lambda$, the part

$$\varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \tag{49.7}$$

is open in E_α.

ITEMooZHLXooBpWSXr
(5) For every $\alpha, \beta \in \Lambda$, the map

$$\varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \to \varphi_\beta^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\nu)\big) \tag{49.8}$$

is continuous.

For each $\alpha \in \Lambda$ we define

$$\tau_\alpha = \{\varphi_\alpha(A) \text{ tel que } A \text{ is open in } U_\alpha\} \tag{49.9}$$

and

$$\tau_0 = \bigcup_{\alpha\in\Lambda} \tau_\alpha. \tag{49.10}$$

Then we define τ_X as the topology generated by τ_0.

Then:

ITEMooLNAPooIZVFVZ
(1) τ_0 is a basis of topology for τ_X.

ITEMooRSZOooFfNrCQ
(2) If $m \in \varphi_\alpha(U_\alpha)$, then τ_α is a basis of topology around m.

ITEMooXGDRooRWiVnF
(3) The maps $\varphi_\alpha : U_\alpha \to X$ are continuous.

ITEMooPPZIooViYPck
(4) If we set $M_\alpha = \varphi_\alpha(U_\alpha)$, the maps $\varphi_\alpha : U_\alpha \to M_\alpha$ are homeomorphic.

ITEMooCMUKooBvlXbs
(5) τ_X is the unique topology on X such that the maps $\varphi_\alpha : U_\alpha \to X$ are homeomorphic.

ITEMooAZHPooWzAjlS
(6) If this topology is Hausdorff and second countable, then X is a topological manifold, and $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ is an atlas.

Proof. Several steps.

(1) **Finite intersections**
Let $S_\alpha \in \tau_\alpha$ and $S_\beta \in \tau_\beta$ be such that $S_\alpha \cap S_\beta \neq \varnothing$. We prove that there exists $S \in \tau_0$ such that $S \subset S_\alpha \cap S_\beta$. There exists open parts $A_\alpha \subset U_\alpha$ and $A_\beta \subset U_\beta$ such that $S_\alpha = \varphi_\alpha(A_\alpha)$ and $S_\beta = \varphi_\beta(A_\beta)$.

4. N_α is the norm.

By hypothesis, the part $(\varphi_\alpha^{-1}\circ\varphi_\beta)(A_\beta)$ is open in U_α. The intersection

$$A = (\varphi_\alpha^{-1}\circ\varphi_\beta)(A_\beta)\cap A_\alpha \tag{49.11}$$

is open and non empty. Setting $S=\varphi_\alpha(A)$ we have $S\in\tau_\alpha\subset\tau_0$ and

$$\begin{aligned} S=\varphi_\alpha(A) &= \varphi_\alpha\big((\varphi_\alpha^{-1}\circ\varphi_\beta)(A_\beta)\cap A_\alpha\big) && (49.12a)\\ &= \varphi_\alpha\big((\varphi_\alpha^{-1}\circ\varphi_\beta)(A_\beta)\big)\cap\varphi_\alpha(A_\alpha) && (49.12b)\\ &= \varphi_\beta(A_\beta)\cap\varphi_\alpha(A_\alpha) && (49.12c)\\ &= S_\alpha\cap S_\beta. && (49.12d)\end{aligned}$$

If, for each $i\in\{1,\dots,n\}$ we have $S_{\alpha_i}\in\tau_{\alpha_i}$, using what we just did, we have a $S_2\in\tau_{\alpha_1}$ such that $S_2\subset S_{\alpha_1}\cap S_{\alpha_2}$. Then we have $S_3\in\tau_{\alpha_1}$ with $S_3\subset S_2\cap S_{\alpha_3}$, and so on.

(2) **Proof of (1) : τ_0 is a basis of topology** Let $\mathcal{O}\in\tau_X$. There exists a finite part $\Lambda_0\subset\Lambda$ such that

$$\bigcap_{\alpha\in\Lambda_0}\mathcal{O}_\alpha\subset\mathcal{O} \tag{49.13}$$

with $\mathcal{O}_\alpha\in\tau_\alpha$ and the intersection being non empty [5]. By the finite intersection part, there exists $S\in\tau_0$ such that

$$S\subset\bigcap_{\alpha\in\Lambda_0}\mathcal{O}_\alpha\subset\mathcal{O}. \tag{49.14}$$

Thus τ_0 is a basis of topology for τ_X.

(3) **Proof of (2)**
Let $m\in\varphi_\alpha(U_\alpha)$ and let $\mathcal{O}\in\tau_X$ containing m. Since τ_0 is a basis of topology of τ_X, there exists $\beta\in\Lambda$ and an open part $A_\beta\subset U_\beta$ such that $m\in\varphi_\beta(A_\beta)\subset\mathcal{O}$. Let $y=\varphi_\beta^{-1}(m)$. We have

$$\varphi_\alpha\big((\varphi_\alpha^{-1}\circ\varphi_\beta)(y)\big)=m, \tag{49.15}$$

and then $(\varphi_\alpha^{-1}\circ\varphi_\beta)(y)=x$. Setting $A_\alpha=(\varphi_\alpha^{-1}\circ\varphi_\beta)(A_\beta)$, the part A_α is open in U_α and we have

$$m\in\varphi_\alpha(A_\alpha)\subset\varphi_\beta(A_\beta)\subset\mathcal{O}. \tag{49.16}$$

(4) **Proof of (3)**
Let $\mathcal{O}$ be open in X. We prove that $\varphi_\alpha^{-1}(\mathcal{O})$ is open in U_α by proving that it contains a neighbourhood of each of its points. Let $x\in\varphi_\alpha^{-1}(\mathcal{O})$ and $m=\varphi_\alpha(x)$. Since τ_α is a basis of topology around m, there exists an open part $A_\alpha\subset U_\alpha$ such that $m\in\varphi_\alpha(A_\alpha)\subset\mathcal{O}$. This means that

$$x\in A_\alpha\subset\varphi_\alpha^{-1}(\mathcal{O}), \tag{49.17}$$

and $\varphi_\alpha^{-1}(\mathcal{O})$ contains a neighbourhood of x.

(5) **Proof of (4)** The maps φ_α are injective by hypothesis, and surjective on their image. We know from point (3) that φ_α is continuous. The map φ^{-1} is continuous too since $\varphi_\alpha(A_\alpha)$ is open when $A_\alpha\subset U_\alpha$ is open.

(6) **Proof of (5)** Let τ'_X be a topology on X such that the maps φ_α are homeomorphic. We prove that $\tau_X=\tau'_X$.

(6a) $\tau_X\subset\tau'_X$ If $A_\alpha\subset U_\alpha$ is open, then $\varphi_\alpha(A_\alpha)\in\tau'_X$ because φ_α^{-1} is continuous on (X,τ'_X). Thus $\tau_\alpha\subset\tau'_X$. We deduce that $\tau_0\subset\tau'_X$, and since τ_X is the topology generated by τ_0 it is contained in all the topologies containing τ_0. Thus $\tau_X\subset\tau'_X$.

(6b) $\tau'_X\subset\tau_X$
Let $\mathcal{O}\in\tau'_X$. For each $\alpha\in\lambda$, the part $\varphi_\alpha(\mathcal{O})$ is open in U_α. We write $A_\alpha=\varphi_\alpha(\mathcal{O})$. We have $\varphi_\alpha(A_\alpha)\in\tau_X$ and

$$\mathcal{O}=\bigcup_{\alpha\in\Lambda}\varphi_\alpha(A_\alpha)\in\tau_X. \tag{49.18}$$

5. The set $\mathcal{O}$ is a union of such intersections, see proposition 7.12.

(7) **Proof of (6)**
Definition 49.1 of a topological manifold.

□

49.1.3 Topology

PROPooDEVUooCATTZI

Proposition 49.11 ([1]).
Let M be a topological manifold. We denote by Λ the set of all the charts on M and

$$\tau_\Lambda = \{S \subset M \text{ tel que } \forall \alpha \in \Lambda, \varphi_\alpha^{-1}(S \cap M_\alpha) \in \tau_n\}. \tag{49.19}$$

Then τ_Λ is the topology of M.

Proof. We denote by τ_M the topology on M.

(1) $\tau_\Lambda \subset \tau_M$ Let $A \in \tau_\Lambda$. In order to prove that A is open, we prove that each $a \in A$ is contained in an open part contained in A (theorem 7.8).
Let $a \in A$. We consider a chart $\varphi_a : U_a \to M$ such that $a \in \varphi_a(U_a)$. By definition, the part $\varphi_a^{-1}(A \cap M_a)$ is open in E_α. Since φ_α is homeomorphic, the image of an open by φ_α is open. Namely,

$$\varphi_\alpha\Big(\varphi_\alpha^{-1}(A \cap M_a)\Big) = A \cap M_a \tag{49.20}$$

is open in M. Thus the par $A \cap M_a$ satisfy $a \in A \cap M_a \subset$. This proves that A is open.

(2) $\tau_M \subset \tau_\Lambda$ Let $A \in \tau_M$. Let $\alpha \in \Lambda$; we have to check that $\varphi_\alpha^{-1}(A \cap M_\alpha)$ is open. The part M_α is open (because it is the image of the open U_α by the homeomorphism φ_α). Thus the intersection $A \cap M_\alpha$ is open and finally the part $\varphi_\alpha^{-1}(A \cap M_\alpha)$ is open.

□

PROPooGAZZooWIPVuf

Proposition 49.12 ([1]).
Let M be a topological manifold. Let $m \in M$ and $\varphi : U \to M$ be a chart such that $m = \varphi(0)$. The set

$$\{\varphi\big(B(0,\epsilon)\big)\}_{\epsilon>0} \tag{49.21}$$

is a basis of neighbourhood of m in M.

49.1.4 Covering with compact closure

LEMooDHIHooQtAWUJ

Lemma 49.13 ([487, 830, 1]).
Let M be a topological manifold. There exists a countable collection of open parts $\{X_i\}_{i\in\mathbb{N}}$ such that

(1) $\bar{X}_i$ is compact,
(2) $\bar{X}_i \subset X_{i+1}$
(3) $M = \bigcup_{i\in\mathbb{N}} X_i$.

Proof. We suppose that M is connected and we prove the result in several steps.

(1) **Countable basis** Since M is second countable, we consider a countable basis of topology $\{\mathcal{O}_j\}_{j\in J}$. We pick among them the ones with compact closure and we denote them by $\{Y_i\}_{i\in\mathbb{N}}$.

(2) Y_i **covers** M We prove that $\bigcup_{i\in\mathbb{N}} Y_i = M$. Let $p \in M$ and a chart $\varphi : U \to M$ around p. We write $x = \varphi^{-1}(p)$. Let S be a bounded open neighbourhood of x in U. The part $\varphi(S)$ is open, and since $\{\mathcal{O}_j\}_{j\in J}$ is a basis of topology, there exists $j \in J$ such that $\mathcal{O}_j \subset \varphi(S)$. Using lemma 7.43,

$$\bar{\mathcal{O}}_j \subset \overline{\varphi(S)} = \varphi(\bar{S}). \tag{49.22}$$

Since $\bar{S}$ is closed and bounded, it is compact (Borel-Lebesgue 9.24). Nous the part $\bar{\mathcal{O}}_j$ is closed in a compact so that it is compact by lemma 7.96(1).
The set $\mathcal{O}_j$ has compact closure and is then one of the $\{Y_i\}_{i\in\mathbb{N}}$.

(3) **The function m** We consider R, the set of parts of M whose closure is compact. Then we define

$$\begin{aligned} m\colon R &\to \mathbb{N} \\ X &\mapsto \min\{n \in \mathbb{N} \text{ tel que } \bar{X} \subset \bigcup_{i=1}^{n} Y_i\}. \end{aligned} \tag{49.23}$$

Now we argue that the minimum exists. The set $\bar{X}$ is compact and covered by the $\{Y_i\}_{i\in\mathbb{N}}$. Thus there exists a finite subset $I_X \subset \mathbb{N}$ such that $\bar{X} \subset \bigcup_{i\in I_X} Y_i$. The finite integer $\max(I_X)$ belongs to the set $\{n \in \mathbb{N} \text{ tel que } \bar{X} \subset \bigcup_{i=0}^{n} Y_i\}$. Thus this set has a minimum.

(4) **Union and compacity** SPITEMooJGIXooJPVjP The part $\bigcup_{i=1}^{n} Y_i$ is compact as finite union of compacts; it is thus closed and we have

$$\overline{\bigcup_{i=1}^{n} Y_i} \subset \overline{\bigcup_{i=1}^{n} \bar{Y}_i} = \bigcup_{i=1}^{n} \bar{Y}_i. \tag{49.24}$$

The part $\overline{\bigcup_{i=1}^{n} Y_i}$ is compact as a closed subset of a compact set.

(5) **The function g** We define

$$\begin{aligned} g\colon R &\to R \\ X &\mapsto \bigcup_{i=0}^{m(X)} Y_i. \end{aligned} \tag{49.25}$$

The function g takes its values in R from point (4).

(6) **A recursion** We define the X_i by the following recursion:

$$\begin{cases} X_1 = Y_1 & \text{(49.26a)} \\ X_{i+1} = g(X_i) & \text{(49.26b)} \end{cases}$$

and the theorem 1.48. Now we have to prove that these X_i satisfy the properties.

(7) **X_i is open** The part X_i is the union of open sets Y_k.

(8) **$\bar{X}_i$ is compact** This is by construction because the image of g is in R.

(9) **$\bar{X}_i \subset X_{i+1}$** This is by construction of m.

(10) **$M = \bigcup_{i=1}^{n} Y_i$** At this point, I remember you that we are dealing with the hypothesis that M is connected. There are two possibilities: either there exists k such that $X_{k+1} = X_k$, or it does not.

(10a) **If there is k such that $X_{k+1} = X_k$** For such a k we have $\bar{X}_k \subset X_{k+1} = X_k \subset \bar{X}_k$, so that $\bar{X}_k = X_k$ and X_k is closed. Since X_k is also open, it is equal to M by proposition 7.74(4).

(10b) **If not** We have $m(X_{i+1}) > m(X_i)$. Since m is a $\mathbb{N}$-valued function, this means that $m(X_i) \to \infty$. In other words, every Y_i belong to some X_i. Since $\{Y_i\}_{i\in\mathbb{N}}$ is a covering of M, we also have that $\{X_i\}_{i\in\mathbb{N}}$ is a covering of M.

What if M is not connected ? There are at most countably many connected components. For each component, we make the whole construction and we get the covering $\{X_{i\in\mathbb{N}}^{(k)}\}$ for the k-th connected component. The we set

$$X_i = \bigcup_{k\leqslant i} X_i^{(k)}. \tag{49.27}$$

(1) **$\bar{X}_i$ is compact** We have

$$\bar{X}_i = \overline{\bigcup_{k\leqslant i} X_i^{(k)}} = \bigcup_{k\leqslant i} \overline{X_i^{(k)}}, \tag{49.28}$$

which is compact as finite union of compacts.

(2) $\underline{\bar{X}_i \subset X_{i+1}}$ We have

$$\bar{X}_i = \overline{\bigcup_{k\leqslant i} X_i^{(k)}} \tag{49.29a}$$
$$= \bigcup_{k\leqslant i} \overline{X_i^{(k)}} \tag{49.29b}$$
$$\subset \bigcup_{k\leqslant i} X_{i+1}^{(k)} \tag{49.29c}$$
$$\subset \bigcup_{k\leqslant i} X_{i+1}^{(k)} \cup X_{i+1}^{(i+1)} \tag{49.29d}$$
$$= X_{i+1}. \tag{49.29e}$$

(3) $\underline{M = \bigcup_{i\in\mathbb{N}} X_i}$ Let $p \in M$. Since $\{X_i^{(k)}\}_{i\in\mathbb{N}}$ is a covering of the k-th connected component, there exists $k, i \in \mathbb{N}$ such that $p \in X_i^{(k)}$. In this case, $p \in X_i$.

□

DEFooBDJJooAxszUA

Definition 49.14 ([487]).
*Let $\mathcal{W} = \{W_\alpha\}_{\alpha\in A}$ be a covering of a set E. A covering $\mathcal{Y} = \{Y_i\}_{i\in I}$ is a **refinement** if for every $i \in I$, there exits a $\alpha \in A$ such that $Y_i \subset W_\alpha$.*

PROPooKAXIooNhNkNB

Proposition 49.15 ([487]).
Let M be a topological manifold. We consider a basis $\mathcal{B} = \{B_s\}_{s\in S}$ of the topology of M and an open covering $\mathcal{W} = \{W_\alpha\}_{\alpha\in A}$.

There exists a countable part I of S such that $\{B_i\}_{i\in I}$ is

(1) a refinement[6] of $\mathcal{W}$,

(2) locally finite.

Proof. We consider a covering $\{X_i\}_{i\in\mathbb{N}}$ as in lemma 49.13.

(1) $\underline{\bar{X}_{i+1}\backslash X_i \text{ **is compact**}}$ For each i, the part $\bar{X}_{i+1}\backslash X_i$ is closed (lemma 7.7(1)) in the compact $\bar{X}_{i+1}$, and then compact (lemma 7.96).

(2) $\underline{\bar{X}_{i+1}\backslash X_i \subset X_{i+2}\backslash \bar{X}_{i-1}}$ Because $\bar{X}_{i+1} \subset X_{i+1}$ and $\bar{X}_{i-1} \subset X_i$.

(3) $\underline{X_{i+2}\backslash \bar{X}_{i-1} \text{ **is open**}}$ By lemma (2)(2).

(4) $\underline{\textbf{The open covering } \mathcal{O}}$ We show that the set

$$\mathcal{O} = \{X_4, X_{i+2}\backslash \bar{X}_{i-1}\}_{i=3,4,\dots} \tag{49.30}$$

is an open covering of M. Let $p \in M$. From the construction (49.13) of the X_i, the point p belongs to one of the X_i. If p belongs to X_1, X_2, X_3 or X_4, it is in X_4.
Suppose that $p \in X_i$ with $i > 4$. Let $m = \min\{i \text{ tel que } p \in X_i\}$. We have $p \in X_m\backslash \bar{X}_{m-1} \subset X_m\backslash \bar{X}_{m-3} \in \mathcal{O}$.

(5) $\underline{\textbf{The basis } \mathcal{B}'}$ We consider the set

$$\mathcal{B}' = \{B \in \mathcal{B} \text{ tel que } \begin{cases} \exists \alpha \in A & \text{tq } B \subset W_\alpha \\ \exists i \in \mathbb{N} & \text{tq } B \subset \mathcal{O}_i \end{cases}\} \tag{49.31}$$

where we denoted

$$\mathcal{O}_i = \begin{cases} X_4 & \text{if } i = 2 \\ X_{i+2}\backslash \bar{X}_{i-1} & \text{otherwise.} \end{cases} \tag{49.32}$$

We prove now that $\mathcal{B}'$ is a basis of the topology on M. Let $x \in M$ and D be an open set containing x. Since $\mathcal{W}$ and $\mathcal{O}$ are coverings of M,we consider $\alpha \in A$ and $i \in \mathbb{N}$ such that $x \in W_\alpha \cap \mathcal{O}_i$. Now $W_\alpha \cap \mathcal{O}_i \cap D$ is an open set containing x. Since $\mathcal{B}$ is a basis, there exists $B \in \mathcal{B}$ such that $B \subset W_\alpha \cap \mathcal{O}_i \cap D$. This set belongs to $\mathcal{B}'$.

6. Definition 49.14.

(6) **Next steps** For each part $\mathcal{O}_i$ we are going to build a finite subset $S_i \subset S$ such that $\mathcal{O}_i \subset \bigcup_{s\in S_i} B_s$. The whole will be a countable subset of S covering M as we wish.

(7) S_3 Since X_4 is open and $\mathcal{B}'$ is a basis of topology, the set

$$\{B \in \mathcal{B}' \text{ tel que } B \subset X_4\} \tag{49.33}$$

is an open covering of X_4, and then an open covering of the compact $\bar{X}_3$. We extract a finite covering; in other words we consider a finite subset $S_3 \subset S$ such that

(7a) $B_s \in \mathcal{B}'$ for every $s \in S_3$,

(7b) $B_s \subset X_4$ for every $s \in S_3$,

(7c) $\bar{X}_3 \subset \bigcup_{s\in S_3} B_s$.

(8) S_i **with** $i \geqslant 4$ The part $X_{i+2}\backslash\bar{X}_{i-1}$ is open, thus the set

$$\{B \in \mathcal{B}' \text{ tel que } B \subset X_{i+2}\backslash\bar{X}_{i-1}\} \tag{49.34}$$

is an open covering of $X_{i+2}\backslash\bar{X}_{i-1}$. Then it is also an open covering of the compact part $\bar{X}_{i+1}\backslash X_i \subset X_{i+2}\backslash X_{i+2}\backslash\bar{X}_{i_1}$. We extract a finite subcovering, that is a finite subset $S_i \subset S$ such that

(8a) $B_s \in \mathcal{B}'$ for every $s \in S_i$,

(8b) $B_s \subset X_{i+2}\backslash\bar{X}_{i-1}$ for every $s \in S_i$,

(8c) $\bar{X}_{i+1}\backslash X_i \subset \bigcup_{s\in S_i} B_s$.

(9) **The answer** We set $I = \bigcup_{i=3}^{\infty} S_i$. This is a countable subset of S such that $B_i \in \mathcal{B}'$ for every $i \in I$. Our work now is to prove that I satisfies the requested conditions.

(10) **Covering** We prove that $\mathcal{I} = \bigcup_{i\in I} B_i$ is a covering of M. In this purpose, we see that the elements of $\mathcal{O}$ are covered by $\mathcal{I}$.

(10a) $X_4 \subset \mathcal{I}$ We know that $\bar{X}_3 \subset \bigcup_{s\in S_3} B_s$. It remains to prove that $X_4\backslash\bar{X}_3 \subset \mathcal{I}$. We have

$$\bar{X}_4\backslash X_3 \subset X_5\backslash\bar{X}_2 \subset \bigcup_{s\in B_4} B_s \subset \mathcal{I}. \tag{49.35}$$

(10b) **The others** We have $X_{i+2}\backslash\bar{X}_{i-1} \subset \mathcal{I}$ by construction.

(11) **Locally finite** We prove that $\{B_i\}_{i\in I}$ is locally finite[7]. Let $x \in M$. The exists a $i \in I$ such that $x \in B_i$.

(11a) **If** $x \in B_s$ **with** $s \in S_3$ If $i \geqslant 5$, we have $B_s \cap X_{i+2}\backslash\bar{X}_{i_1} = \varnothing$. Thus an element of B_s is covered by (at most) the sets B_t with $t \in S_3 \cup S_4$. Since S_3 and S_4 are finite, there are only finitely many elements of $\mathcal{I}$ containing x.

(11b) **If** $x \in S_i$ **with** $i \geqslant 4$ We have $B_s \subset X_{i+2}\backslash\bar{X}_{i-1}$, and then

$$B_s \cap \left(X_{i+5}\backslash\bar{X}_{i+2}\right) = \varnothing. \tag{49.36}$$

Thus $B_s \cap B_t = \varnothing$ for every $t \geqslant i+3$. This shows that $x \in B_s$ is contained at most in the sets B_t with $t \in S_2 \cup \ldots \cup S_{i+3}$ which is finite.

(12) **Refinement** It remains to be shown that

$$\mathcal{B}' = \{B_i \text{ tel que } i \in I\}' \tag{49.37}$$

is a refinement of $\mathcal{W}$. We have

$$\mathcal{B}'' \subset \mathcal{B}' = \{B \in \mathcal{B} \text{ tel que } \begin{cases} \exists \alpha \in A & \text{tq } B \subset W_\alpha \\ \exists i \in \mathbb{N} & \text{tq } B \subset \mathcal{O}_i. \end{cases}\} \tag{49.38}$$

Thus every element of $\mathcal{B}''$ is contained in some W_α.

7. Definition 48.14.

□

PROPooYJKOooRwbOXF

Proposition 49.16 (Countable locally finite atlas[1])**.**
Let M be a topological manifold and let $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an atlas. There exists an atlas $\{(V_\beta, \phi_\beta)\}_{\beta\in\Sigma}$ such that

(1) Σ is countable.

(2) $\{\phi_\beta(V_\beta)\}$ is a locally finite covering.

(3) There exists a map $f\colon \Sigma \to \Lambda$ such that

(3a) $V_\beta \subset U_{f(\beta)}$

(3b) $\phi_\beta(x) = \varphi_{f(\beta)}(x)$.

Proof. Let $\{B_s\}_{s\in S}$ be a basis of topology of M. The set $\{\varphi_\alpha(U_\alpha)\}$ is an open covering of M. Proposition 49.15 says that there exists a countable subset $\Sigma \subset S$ such that $\{B_\beta\}_{\beta\in\Sigma}$ is a refinement of $\{\varphi_\alpha(U_\alpha)\}_{\alpha\in\Lambda}$. In other words, there exists a map $f\colon \Sigma \to \Lambda$ such that

$$B_\beta \subset \varphi_{f(\beta)}(U_{f(\beta)}). \tag{49.39}$$

We define

$$V_\beta = \varphi^{-1}_{f(\beta)}(B_\beta) \subset U_{f(\beta)}. \tag{49.40}$$

and

$$\begin{aligned} \phi_\beta\colon V_\beta &\to M \\ x &\mapsto \varphi_{f(\beta)}(x). \end{aligned} \tag{49.41}$$

□

49.2 Differentiable manifolds

49.2.1 Definition, charts

DEFooRUVIooSZDjlE

Definition 49.17.
Let M be a topological manifold[8]*. An **atlas** is a set $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ such that*

(1) For every $\alpha \in \Lambda$, the part U_α is open in some normed vector space E_α, and $\varphi_\alpha(U_\alpha)$ is open in M,

(2) The map φ_α is homeomorphic.

(3) For every $p \in M$, there exists $\alpha \in \Lambda$ such that $p \in \varphi_\alpha(U_\alpha)$.

DEFooVMWRooGQYJwl

Definition 49.18.
*Let $\mathcal{A}$ be a class of functions as C^∞, C^k, or analytic. A n-dimensional $\mathcal{A}$-**manifold** is a topological manifold M with an atlas*[9] *$\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ such that for every $\alpha, \beta \in \Lambda$, the map*

$$(\varphi_\beta^{-1} \circ \varphi_\alpha)\colon \varphi_\alpha^{-1}(\varphi_\alpha(\mathcal{U}_\alpha) \cap \varphi_\beta(\mathcal{U}_\beta)) \to \mathcal{U}_\beta \tag{49.42}$$

is in the class $\mathcal{A}$ as map from E_α to E_β.

In most of the cases, we will use C^∞ or analytic manifolds. The expression "smooth manifold" means C^∞ manifold.

The maps $(U_\alpha, \varphi_\alpha)$ are said "definition charts", but the definition 49.20 and the proposition 49.27 will show that they are not really special.

Exemple 49.19.
Any open set of $\mathbb{R}^n$ is a smooth manifold if we choose the identity map as maps. △

8. Definition 49.1.
9. Definition 49.17.

Most of surfaces $z = f(x, y)$ in $\mathbb{R}^3$ are manifolds, depending on certain regularity conditions on f.

DEFooQLPIooPGagtz

Definition 49.20 ([1])**.**
*Let M be a manifold with its maps $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in I}$. A C^k-**chart** of M is a couple (V, ψ) where V is open in $\mathbb{R}^n$ and $\psi\colon V \to M$ is such that for every $\alpha \in I$, the maps*

$$\psi^{-1} \circ \varphi_\alpha\colon \varphi_\alpha^{-1}\big(\psi(V)\big) \to V \tag{49.43}$$

and

$$\varphi_\alpha^{-1} \circ \psi\colon \psi^{-1}\big(\varphi_\alpha(U_\alpha)\big) \to U_\alpha \tag{49.44}$$

are of class C^k.

DEFooUFHTooTXUVpN

Definition 49.21 ([1])**.**
Let $\mathcal{A}$ be a class of functions (C^k, smooth, analytic). Let M, N be two manifolds. We consider the charts (U, φ) and (V, ϕ) of M and N. We say that a map $f\colon M \to N$ is in the class $\mathcal{A}$ with respect to these charts if the map

$$\phi^{-1} \circ f \circ \varphi\colon \varphi^{-1}\big(\phi(V)\big) \to V \tag{49.45}$$

is in the class $\mathcal{A}$.

If A is a set of charts of M and B is a set of charts of N, we say that f is in the class $\mathcal{A}$ if it is in the class $\mathcal{A}$ for every choice of charts in A and B.

PROPooCWPAooKDnwHR

Proposition 49.22.
If f and g are C^k functions on M, then the functions $f + g$ and fg are C^k.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

DEFooMLNQooEgEfdq

Definition 49.23 ([1])**.**
*Let M be a manifold. An **atlas** for M is a set of charts*[10] *$\{(U_j, \phi_j)\}_{j\in J}$ such that $\bigcup_{j\in J} \phi_j(U_j) = M$.*

49.24.
An atlas is a set of charts which is sufficient to reach each point of M. It is however not required that every open sets in M is the image of a chart in the atlas.

PROPooUFGQooACIjVL

Proposition-Definition 49.25.
Let A be an atlas[11] *of M and B an atlas of N. If a map $f\colon M \to N$ is in the class*[12] *$\mathcal{A}$ for these atlas, it is in the same class $\mathcal{A}$ for every charts of M and N.*

In this case, one say that the map f is in the class $\mathcal{A}$.

LEMooGAMVooIWUzmy

Lemma 49.26.
Let $(M, \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in I})$ be a manifold. If (V, ψ) is a chart, the set

$$\psi^{-1}\big(\varphi_\alpha(U_\alpha)\big) \tag{49.46}$$

is open in $\mathbb{R}^n$.

Proof. By definition 49.20, the map $\varphi_\alpha \circ \psi\colon V \to U_\alpha$ is C^k and, in particular, continuous[13]. Thus if $\mathcal{O}$ is open in U_α, then $(\varphi_\alpha^{-1} \circ \psi)^{-1}(\mathcal{O})$ is open in V. Since V is open in $\mathbb{R}^n$, an open set in V is open in $\mathbb{R}^n$. The set U_α is in particular open in U_α, thus the part

$$(\varphi_\alpha^{-1} \circ \psi)^{-1}(U_\alpha) = (\psi^{-1} \circ \varphi_\alpha)(U_\alpha) \tag{49.47}$$

is open in $\mathbb{R}^n$. □

10. Definition 49.20.
11. Definition 49.23.
12. Being in a class, definition 49.21.
13. Definition 7.36(2).

PROPooUDVFooEJeluM

Proposition 49.27 ([1])**.**
If (V_1, ψ_1) and (V_2, ψ_2) are C^k-charts, then the map

$$\psi_2^{-1} \circ \psi_1 \colon \psi_1^{-1}\big(\psi_2(V_2)\big) \to V_2 \tag{49.48}$$

is of class C^k.

Proof. Let $q \in \psi_1^{-1}\big(\psi_2(V_2)\big)$. We will prove that $\psi_2^{-1} \circ \psi_1$ is C^k on a neighbourhood of q. Let $p = \psi_1(q)$ and $(U_\alpha, \varphi_\alpha)$ be a definition chart around p. We consider

$$A = \varphi_\alpha(U_\alpha) \cap \psi_1(V_1) \cap \psi_2(V_2). \tag{49.49}$$

The point p belongs to A. We set $U' = \varphi_\alpha^{-1}(A)$, $V_1' = \psi_1^{-1}(A)$ and $V_2' = \psi_2^{-1}(A)$. We show that

$$\psi_2^{-1} \circ \psi_1 \colon V_1' \to V_2' \tag{49.50}$$

is C^k.

We have

$$\psi_2^{-1} \circ \psi_1 = \psi_2^{-1} \circ \varphi \circ \varphi^{-1} \circ \psi_1. \tag{49.51}$$

Since ψ_1 and ψ_2 are charts, the maps $\psi_2^{-1} \circ \varphi$ and $\varphi^{-1} \circ \psi_1$ are C^k, so that the compound function is C^k by theorem 10.277. □

LEMooOPPJooXezOHS

Lemma 49.28.
Let M be a C^k manifold and $p \in M$. There exists a chart (V, ψ) of M around p such that $0 \in V$ and $p = \psi(0)$.

Proof. By definition there exists a chart $(U_\alpha, \varphi_\alpha)$ around p and $u \in U_\alpha$ such that $p = \varphi_\alpha(u)$. Let $V = U - u$ and $\psi(x) = \varphi_\alpha(x + u)$.

Then V is an open set and $\psi(0) = \varphi_\alpha(u) = p$. □

49.2.2 Topology

PROBooGNZMooFknpBn

¡¡ Avertissement/question à la lectrice !! 49.29
The proposition-definition 49.30 should not exist because of the proposition 49.11.

DEFooHGNOooNqGmxE

Proposition-Definition 49.30 (Topology on a differntiable manifold)**.**
Let M be a manifold with the set of all the charts $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in \Lambda}$. We denote by τ_n the topology on $\mathbb{R}^n$. The set

$$\tau_M = \{S \subset M \text{ tel que } \forall \alpha \in \Lambda, \varphi_\alpha^{-1}\big(S \cap \varphi_\alpha(U_\alpha)\big) \in \tau_n\} \tag{49.52}$$

*is a **topology** on M.*

Proof. First we prove that the open system defines a topology. For this, remark that φ_α^{-1} is injective (if not, there should be some multivalued points). Then $\varphi^{-1}(A \cap B) = \varphi^{-1}(A) \cap \varphi^{-1}(B)$. If V_1 and V_2 are open in M, then

$$\varphi^{-1}(V_1 \cap V_2 \cap \varphi(\mathcal{U})) = \varphi^{-1}(V_1 \cap \varphi(\mathcal{U})) \cap \varphi^{-1}(V_2 \cap \varphi(\mathcal{U})) \tag{49.53}$$

which is open in $\mathbb{R}^n$. The same property works for the unions. □

THOooIAXUooDqMrav

Theorem 49.31.
Let M be a manifold. Its topology has the following properties.

(1) the charts maps are continuous,

(2) the sets $\varphi_\alpha(\mathcal{U}_\alpha)$ are open.

Proof. We proof the continuity of $\varphi: \mathcal{U} \to M$; for an open set V in M, we have to show that $\varphi^{-1}(V)$ is open in $\mathcal{U} \subset \mathbb{R}^n$. But the definition of the topology on M, is precisely the fact that $\varphi^{-1}(V \cap \varphi(\mathcal{U}))$ is open. □

LEMooGDMZooLCtnuA

Lemma 49.32.
The topology of $\mathbb{R}^n$ as manifold is the same as the usual one.

PROPooCXXKooJAOmOl

Proposition 49.33 ([1]).
If $\{\varphi_\alpha\}_{\alpha\in\Lambda}$ is the set of all charts of the smooth manifold M, then the parts $\{\varphi_\alpha(U_\alpha)\}$ is a basis of topology of M.

49.2.3 Regularity

DEFooMELXooEkEnwz

Definition 49.34.
*If M and M are two analytic manifolds, a map $\phi: M \to N$ is **regular** at $p \in M$ if it is analytic at p and $d\phi_p: T_pM \to T_{\phi(p)}N$ is injective.*

PgDefRegular

DEFooFNTHooEwsqXB

Definition 49.35 ([831]).
Let M and N be C^k-manifolds. We say that a map $h: M \to N$ is C^k if the two following conditions hold:

(1) The map h is continuous[14].

(2) for every $p \in M$, there exists charts (U, φ) around p and (V, ψ) around $h(p)$ such that

(2a) $h\big(\varphi(U)\big) \subset \psi(V)$

SUBITEMooXQFUooRxMVnw

(2b) the map $\psi^{-1} \circ h \circ \varphi: U \to V$ is C^k (definition 10.261).

PROPooGQVUooWmPmQV

Proposition 49.36 ([1]).
Let M and N be C^k manifolds. If for every $m \in M$ there exists a chart (U_m, φ_m) such that $m \in \varphi_m(U_m)$ and $f: U_m \to \varphi_m(U_m)$ is C^k (here $\varphi_m(U_m)$ is seen as a manifold with this chart), then f is C^k on M.

PROPooODGHooFYluSu

Proposition 49.37 (Composition[1]).
Let X, Y and Z be C^k manifolds. We consider C^k maps $f: X \to Y$ and $g: Y \to Z$. Then the composition $g \circ f: X \to Z$ is C^k.

For the sake of the following lemma, we say «mC^k» for «manifold»-C^k (definition 49.35) and «uC^k» for «usual» C^k (definition 10.261 for normed vector spaces). The lemma will say that the two notions are the same, so that we can say «C^k» without more precisions.

LEMooXGFXooBcmIbK

Lemma 49.38 ([1]).
We consider the manifolds $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$ with the identify as charts. A map $f: M \to N$ is mC^k if and only if it is uC^k.

Proof. For the sake of notations, we set the charts $U = \mathbb{R}^m$, $V = \mathbb{R}^n$ and $\psi_m: U \to M$, $\psi_n: V \to N$. The maps ψ_n and ψ_m are the identity.

(1) $\Rightarrow$ The hypothesis that $f: M \to N$ is mC^k says that the map $\psi_n^{-1} \circ f \circ \psi_m$ is uC^k. Thus the map

$$f = \psi_n \circ \psi_n^{-1} \circ f \circ \psi_m \circ \psi_m^{-1} \tag{49.54}$$

is uC^k too.

(2) $\Leftarrow$ Since f is uC^k and since ψ_n and ψ_m are the identity, the map $\psi_n^{-1} \circ f \circ \psi_m$ is uC^k, which means that f is mC^k.

□

14. With respect to the topology of the definition 49.30.

49.2.4 Partition of unity

DEFooGEAJooInejRf

Definition 49.39 (Partition of unity[487]).
*Let M be a smooth manifold. A **partition of unity** on M is a collection of smooth functions $\{\phi_i : M \to \mathbb{R}\}_{i\in I}$ such that*

ITEMooWBHEooSHUQvv

(1) The sets $\{\text{supp}(\phi_i)\}_{i\in I}$ is a locally finite covering of M.

ITEMooSPHGooQSDFTz

(2) $\phi_i(m) \geqslant 0$ for every $i \in I$ and $m \in m$.

ITEMooGWGPooOGdAGU

(3) $\sum_{i\in I} \phi_i(m) = 1$ for every $m \in M$.

The last sum is defined by 10.112, but in most cases we will be reduced to a finite sum.

DEFooPJKOooFnCkuk

Definition 49.40 ([487]).
*Let $\{\phi_i\}_{i\in I}$ be a partition of unity on the smooth manifold M. Let $\mathcal{W} = \{W_\alpha\}_{\alpha\in\Lambda}$ be an open covering of M. We say that the partition $\{\phi_i\}_{i\in I}$ is **subordinate** to the covering $\mathcal{W}$ if for every $i \in I$, there exists $\alpha \in \Lambda$ such that $\text{supp}(\phi_i) \subset W_\alpha$.*

49.41.
The definition of "subordinate" partition of unity in [487] and [830] are not the same. There are several conventions.

LEMooBBBNooVmpLok

Lemma 49.42 ([487]).
Let M be a smooth manifold. There exists a basis of topology $\{B_i\}_{i\in I}$ such that

(1) $\bar{B}_i$ is compact for every $i \in I$,

(2) For every $i \in I$, there exists a smooth function $\phi_i : M \to \mathbb{R}$ such that

(2a) $\phi_i(x) = 0$ if $x \notin B_i$

(2b) $\phi_i(x) > 0$ if $x \in B_i$

Proof. We consider the open cubes in $\mathbb{R}^n$:

$$C(r) = \{x \in \mathbb{R}^n \text{ tel que } x_i \in]-r, r[\}. \tag{49.55}$$

(1) **The set Λ_p**
We consider the set $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ of all the charts of M. For each $p \in M$ we define

$$\Lambda_p = \{\alpha \in \Lambda \text{ tel que } p = \varphi_\alpha(0), C(3) \subset U_\alpha\}. \tag{49.56}$$

Notice that any chart containing p can be translated and dilated in order to satisfy the conditions. For every p, the set Λ_p is non empty and, in fact, infinite.

(2) **The set $\mathcal{R}(p)$** We define $\mathcal{R}(p) = \{\varphi_\alpha\big(C(2)\big)\}_{\alpha\in\Lambda_p}$ and we prove that this is a basis of neighbourhood at p. Let A be a neighbourhood of $p \in M$. Since the charts make a basis of topology (proposition 49.33), exists a chart $\varphi_\alpha : U_\alpha \to M$ with $p \in \varphi_\alpha(U_\alpha) \subset A$. Then we consider the translated chart

$$\begin{aligned} \varphi_\beta : U_\alpha - x &\to M \\ y &\mapsto \varphi_\alpha(y + x) \end{aligned} \tag{49.57}$$

where $x = \varphi_\alpha^{-1}(p)$. Notice that

$$\varphi_\beta\big(B(0, r)\big) \subset \varphi_\beta(U_\beta) = \varphi_\alpha(U_\alpha) \subset A. \tag{49.58}$$

We choose a $r > 0$ such that $B(0, r) \subset U_\beta$, and for each $k \in \mathbb{N}$, we consider the chart

$$\begin{aligned} \varphi_k : B(0, kr) &\to M \\ x &\mapsto \varphi_\beta(x/k). \end{aligned} \tag{49.59}$$

This is well defined because, if $x \in B(0, kr)$, we have $x/k \in B(0, r) \subset U_\beta$. We choose k large enough to fulfill the two conditions

— $C(3) \subset B(0, kr)$

— $\varphi_k\big(C(2)\big) \subset A$.

The first condition ensures that φ_k belongs to $\{\varphi_\alpha\}_{\alpha\in\Lambda_p}$. If $x \in C(3)$, we have $\|x\| \leqslant 3\sqrt{n}$, so that taking $k > 3\sqrt{n}/r$ satisfy the first condition. For the second condition, taking $k > 2\sqrt{n}/r$ ensures that

$$\begin{aligned}
\varphi_k\big(C(2)\big) &\subset \varphi_k\big(B(0, 2\sqrt{n})\big) && \text{(49.60a)}\\
&= \varphi_\beta\big(B(0, 2\sqrt{n}/k)\big) && \text{(49.60b)}\\
&\subset \varphi_\beta\left(B\big(0, \frac{2\sqrt{n}}{2\sqrt{n}/r}\big)\right) && \text{(49.60c)}\\
&= \varphi_\beta\big(B(0, r)\big) && \text{(49.60d)}\\
&\subset A. && \text{(49.60e)}
\end{aligned}$$

(3) **Basis of topology** Taking $I = \bigcup_{m\in M} \Lambda_m$, the parts $B_i = \varphi_i\big(C(2)\big)$ with $i \in I$ form a basis of the topology of M.

(4) **Bump function** We consider the Urysohn's lemma [15] with the compact $C(2)$ and the open $\text{Int}\big(C(3)\big)$. There exists a smooth function $b\colon \mathbb{R}^n \to [0, 1]$ such that $b(x) = 1$ if $x \in C(2)$ and $C(x) = 0$ if $x \notin C(3)$.

For each $i \in I$ we define

$$\begin{aligned}
\phi_i\colon M &\to \mathbb{R}\\
x &\mapsto \begin{cases} b \circ \varphi_i^{-1}(x) & \text{if } x \in \varphi_i(U_i)\\ 0 & \text{otherwise.}\end{cases}
\end{aligned} \tag{49.61}$$

We have to check that ϕ_i is smooth in the sense of definition 49.35. The proposition 49.36 allows us to prove more or less pointwise while the lemma 49.38 says that b is smooth in the sense of manifolds. We will also use the proposition 49.37 to deal with the composition.

Let $m \in M$. There are two possibilities: $m \in \varphi_i(U_i)$ or not. If $m \in \varphi_i(U_i)$, then the part $\varphi_i(U_i)$ is an open neighbourhood of m on which ϕ_i is smooth because for every chart $\varphi_\beta\colon U_\beta \to M$ we have $\varphi_\beta \circ \phi_i = \varphi_\beta \circ b \circ \varphi_i^{-1}$ which is smooth.

Now we suppose that m does not belong to $\varphi_i(U_i)$. We have two (sub-)cases: there is a neighbourhood of m which does not intersect $\varphi_i(U_i)$ or every neighbourhood of m intersects $\varphi_i(U_i)$. In the first case we have a neighbourhood of m on which $\phi_i = 0$, and ϕ_i is smooth.

Now we suppose that every neighbourhood of m intersects $\varphi_i(U_i)$. We use the lemma 7.105: M is Hausdorff, $\varphi_i(\overline{C(3)}) \subset \varphi_i(U_i)$ where $\varphi_u\big(\overline{C(3)}\big)$ is compact and $\varphi_i(U_i)$ is open. Since m is outside $\varphi_i(U_i)$, there exists a neighbourhood A of m which do not intersect $\overline{C(3)}$. We have $\phi_i = 0$ on this neighbourhood.

□

THOooEJDHooTkRvsn

Theorem 49.43 (Existence of a partition of unity[487] [16])**.**
Let M be a smooth manifold. Let $\mathcal{W}$ be an open covering. There exists a partition of unity $\{\phi_i\}_{i\in I}$ such that

(1) I is countable,

(2) subordinate to $\mathcal{W}$ [17],

(3) $\text{supp}(\phi_i)$ is compact for every $i \in I$.

(4) $\{\text{supp}(\phi_i)\}_{i\in I}$ is a locally finite covering of M [18].

15. Proposition 27.74.
16. Most of the time, you want to use the proposition 49.44 which allows to choose a gentle atlas.
17. Definition 49.40.
18. This is not really a point because this is part of the definition of a partition of unity.

Proof. The lemma 49.42 provides a basis of topology $\{B_i\}_{i\in I}$ and smooth maps $\phi_i\colon M\to\mathbb{R}$ such that

(1) $\bar{B}_i$ is compact
(2) $\phi_i = 0$ outside B_i
(3) $\phi_i > 0$ on B_i.

The proposition 49.15 says that there exists a countable part $J\subset I$ such that $\{B_j\}_{j\in J}$ is a locally finite refinement of $\mathcal{W}$.

(1) **The function ϕ**
We define
$$\phi(x) = \sum_{j\in J}\phi_j(x). \tag{49.62}$$
Since $\{B_j\}_{j\in J}$ is locally finite, for each x, the sum is finite.

(2) **ϕ is smooth** Let $m\in M$. There exists a neighbourhood A of m and a finite subset $J'\subset J$ such that $A\cap B_j\neq\varnothing$ only when $j\in J'$. Thus the map $\phi\colon A\to\mathbb{R}$ is a finite sum and is smooth as finite sum of smooth functions.

(3) **$\phi(x)\neq 0$** We prove that ϕ is everywhere non vanishing. Let $m\in M$. There is a $j_0\in J'$) such that $m\in B_{j_0}$. Then $\phi_{j_0}(m)>0$ and
$$\phi(m) = \sum_{j\in J}\phi_j(m)\geqslant\phi_{j_0}(m)>0. \tag{49.63}$$

(4) **The functions ψ_j** Since $\phi(x)\neq 0$, we are allowed to define, or each $j\in J'$
$$\begin{aligned}\psi_j\colon M&\to\mathbb{R}\\ x&\mapsto\frac{\phi_j(x)}{\phi(x)}.\end{aligned} \tag{49.64}$$
We show that this is our partition of unity subordinate to $\mathcal{W}$. We have
$$\operatorname{supp}(\psi_j) = \operatorname{supp}(\phi_j) = \bar{B}_i. \tag{49.65}$$ EQooVRCOooKnBAmA

(4a) **Condition 49.39(1)** The covering $\{B_j\}_{j\in J'}$ is locally finite. Lemma 48.15 says that $\{\bar{B}_j\}_{j\in J'}$ is locally finite.

(4b) **Condition 49.39(2)**
We have $\phi(x)>0$ for every x and $\phi_j(x)\geqslant 0$ for every $x\in M$.

(4c) **Condition 49.39(3)**
Let $m\in M$. We have
$$\sum_{j\in J'}\psi_j(m) = \frac{1}{\phi(m)}\sum_{j\in J'}\phi_j(m) = 1. \tag{49.66}$$

(4d) **Subordinate to $\mathcal{W}$ (definition 49.40)**
Let $W\in\mathcal{W}$. We have chosen $\{B_j\}_{j\in J}$ to be a refinement of $\mathcal{W}$; thus $\{\bar{B}_j\}_{j\in J'}$ is also a refinement.

(4e) **$\operatorname{supp}(\psi_j)$ is compact** By equation 49.65 and the fact that we choose B_i in such a way that $\bar{B}_i$ is compact.

□ PROPooGEWIooTwCfPp

Proposition 49.44 ([1]).
Let M be a smooth manifold. There exists an atlas $\{(U_\alpha,\varphi_\alpha)_{\alpha\in\Lambda}\}$ and a partition of unity $\{\phi_i\}_{i\in I}$ such that

(1) The atlas $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ is countable and locally finite.

(2) The partition of unity $\{\phi\}_{i\in I}$ is countable and subordinate to $\mathcal{W} = \{\varphi_\alpha(U_\alpha)\}_{\alpha\in\Lambda}$.

(3) The supports $\operatorname{supp}(\phi_i)$ are compacts.

Proof. The proposition 49.16 gives an atlas $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ which is countable and for which $\mathcal{W} = \{\varphi_\alpha(U_\alpha)\}_{\alpha\in\Lambda}$ is locally finite. Now the theorem 49.43 provides a partition of unity $\{\phi_i\}_{i\in I}$ which is subordinate to $\mathcal{W}$ with compact supports. □

PROPooOVQHooMNRQDH

Proposition 49.45 ([1]).
Let $\{U_\alpha\}_{\alpha\in\Lambda}$ be an open covering of a smooth manifold. There exists a partition of unity $\{\rho_\alpha\}_{\alpha\in\Lambda}$ such that for every $\alpha \in \Lambda$,

(1) $\operatorname{supp}(\rho_\alpha) \subset U_\alpha$,

(2) $\operatorname{supp}(\rho_\alpha)$ is compact.

Proof. We consider a partition of unity $\{\phi_i\}_{i\in I}$ subordinate to the open covering $\{U_\alpha\}$ as in theorem 49.43. For each $i \in I$, there exists a $\alpha \in \Lambda$ such that $\operatorname{supp}(\phi_i) \subset U_\alpha$. We consider a map $f : I \to \Lambda$ such that [19]

$$\operatorname{supp}(\phi_i) \subset U_{f(i)}. \tag{49.67}$$

Now for each x we define

$$\rho_\alpha(x) = \sum_{i\in f^{-1}(\alpha)} \phi_i(x). \tag{49.68}$$

EQooBMYNooBnXQHm

Since the covering $\{\operatorname{supp}(\phi_i)\}_{i\in I}$ is locally finite, the sum in the right hand side of (49.68) is finite. We check that $\{\phi_\alpha\}_{\alpha\in\Lambda}$ is the requested partition of unity.

(1) **Sum to 1**
For each $i \in I$, there exists a unique $\alpha \in \Lambda$ such that $i \in f^{-1}(\alpha)$, thus

$$\sum_{\alpha\in\Lambda} \rho_\alpha(x) = \sum_{i\in I} \phi_i(x) = 1. \tag{49.69}$$

(2) **Support** For every $i \in f^{-1}(\alpha)$ we have $\operatorname{supp}(\phi_i) \subset U_\alpha$. Thus

$$\operatorname{supp}(\rho_\alpha) = \operatorname{supp}\Big(\sum_{i\in f^{-1}(\alpha)} \phi_i\Big) \subset U_\alpha. \tag{49.70}$$

□

49.2.5 Product of manifolds

DEFooYOLXooDPrnHa

Definition 49.46.
TODO: Definition of $M \times N$ if M and N are manifolds.

PROPooCHVLooVFScOl

Proposition 49.47 ([1]).
Let $\mathcal{A}$ be a class of functions: C^k, smooth or analytic. Let M and N be $\mathcal{A}$-manifolds.

(1) The permutation

$$\begin{aligned}\sigma : M \times N &\to N \times M\\ (p, q) &\mapsto (q, p)\end{aligned} \tag{49.71}$$

is in the class $\mathcal{A}$.

ITEMooRFFAooRSeBP1

(2) The projection

$$\begin{aligned}\pi_1 : M \times N &\to M\\ (p, q) &\mapsto p\end{aligned} \tag{49.72}$$

is in the class $\mathcal{A}$.

19. Such a map is not unique, and its existence maybe requires some form of axiom of choice.

(3) *Let $p \in M$. The inclusion map*

$$\begin{aligned} \iota\colon N &\to M \times N \\ q &\mapsto (p,q) \end{aligned} \tag{49.73}$$

is in the class $\mathcal{A}$.

49.3 Tangent vector

Definition 49.48.
An open set $\Omega \subset \mathbb{R}^n$ is a manifold with the identify chart (Ω, φ) where $\phi\colon \Omega \to \Omega$ is the identity.

LEMooMFQNooMTtpRp

Lemma 49.49.
Let $\gamma\colon \mathbb{R} \to M$ be a C^1 path. Let $f\colon M \to \mathbb{R}$ be a C^1-function. Then the map $f \circ \gamma\colon \mathbb{R} \to \mathbb{R}$ is C^1.

Proof. Let $a \in \mathbb{R}$ and $p = \gamma(a)$. We consider a chart (U, φ) of M around p. We decompose

$$f \circ \gamma = f \circ \varphi \circ \varphi^{-1} \circ \gamma. \tag{49.74}$$

Since φ is a chart, the maps $f \circ \varphi$ and $\varphi^{-1} \circ \gamma$ are C^1. □

49.3.1 Tangent vector

DEFooJJVIooDUBwDJ

Definition 49.50.
Let $\gamma\colon \mathbb{R} \to M$ be a C^1 path[20]*. We consider the operator*[21]

$$\begin{aligned} \nabla_\gamma\colon C^k(M,\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto (f \circ \gamma)'(0) \end{aligned} \tag{49.75}$$

We define

$$T_aM = \{\nabla_\gamma \text{ tel que } \gamma \in C^k(\mathbb{R}, M), \gamma(0) = a\}. \tag{49.76}$$

*This is the **tangent space** at a.*

If $\gamma\colon \mathbb{R} \to M$ is a path, we also use the notation

$$\nabla_\gamma = \frac{d}{dt}\Big[\gamma(t)\Big]_{t=0}. \tag{49.77}$$

EQooJQVRooLziKoH

If one sees $\mathbb{R}$ as a manifold, then the expression $\frac{d}{dt}\Big[2t+1\Big]_{t=0}$ can stand for the number 1 (usual derivative of $t \mapsto 2t+1$ at $t=0$) or for the operator

$$\frac{d}{dt}\Big[2t+1\Big]_{t=0}\phi = 2\phi'(1). \tag{49.78}$$

REMooJQFHooQuoZxt

Remarque 49.51.
The notation $\gamma'(0)$ for the tangent vector to the curve γ has to be taken with caution. In particular, $\gamma'(0)$ is not defined by the limit

$$\lim_{\epsilon \to 0} \frac{\gamma(\epsilon) - \gamma(0)}{\epsilon} \tag{49.79}$$

EQooVMGFooFUCNEY

because when M is a manifold, there is in general no notion of difference between the points of M, so that the difference $\gamma(\epsilon) - \gamma(0)$ has no meaning.

The only definition of $\gamma'(0)$ is as differential operator.

20. The concept of map of class C^k between two manifolds is the definition 49.35.
21. The map $f \circ \gamma$ is differentiable by lemma 49.49.

LEMooMHSQooQyTZCg

Lemma 49.52 ([1]).
Let $\gamma: I \to M$ be a C^1 path. Let $u \in \mathbb{R}$. We consider the path

$$\begin{aligned} \sigma: I &\to M \\ t &\mapsto \gamma(ut). \end{aligned} \tag{49.80}$$

Then we have $\sigma'(0) = u\gamma'(0)$.

Proof. Let f be a C^1 map on M. We consider the map

$$\begin{aligned} m: \mathbb{R} &\to \mathbb{R} \\ t &\mapsto ut \end{aligned} \tag{49.81}$$

and we use the usual chain derivation rule :

$$\nabla_\sigma(f) = (f \circ \sigma)'(0) \tag{49.82a}$$
$$= \Big((f \circ \gamma) \circ m\Big)'(0) \tag{49.82b}$$
$$= (f \circ \gamma)'\big(m(0)\big)m'(0) \tag{49.82c}$$
$$= \nabla_\gamma(f)u. \tag{49.82d}$$

□

49.3.2 Vector space structure on the tangent space

The fact that $\{\partial_i\}$ is a basis of T_pM will be the proposition 49.56.

PROPooEJBWooSbvypo

Definition 49.53 (T_pM is a vector space[1]).
Let M be a C^k manifold. If X and Y are elements of T_pM, and if $\lambda \in \mathbb{R}$, we introduce the operations[22]

$$\begin{aligned} \lambda X: C^\infty(M) &\to \mathbb{R} \\ f &\mapsto \lambda X(f) \end{aligned} \tag{49.83}$$

and

$$\begin{aligned} X + Y: C^\infty(M) &\to \mathbb{R} \\ f &\mapsto X(f) + Y(f). \end{aligned} \tag{49.84}$$

Then:

1. *We have $\lambda X \in T_pM$*
2. *We have $X + Y \in T_pM$* ITEMooDRUQooWrsaDd
3. *With these operations, $(T_pM, +, \cdot)$ is a vector space.* ITEMooJEYEooLuJiqz
4. *The dimension of T_pM is the dimension of M.*

Proof. Let $p \in M$ and a chart (U, φ) of M around p such that $p = \varphi(0)$ (lemma 49.28). We consider $X, Y \in T_pM$ and paths $\gamma: \mathbb{R} \to M$, $\sigma: \mathbb{R} \to M$ such that $X = \nabla_\gamma$ and $Y = \nabla_\sigma$.

(1) **Sum**
We aim to find a path $s: \mathbb{R} \to M$ such that $\nabla_s = X + Y$. We set

$$\begin{aligned} s: \mathbb{R} &\to M \\ t &\mapsto \varphi\Big(\varphi^{-1}\big(\gamma(t)\big) + \varphi^{-1}\big(\sigma(t)\big)\Big). \end{aligned} \tag{49.85}$$

EQooYNAOooKkKmxo

22. These are the usual definitions of sum and product by a scalar of maps $E \to \mathbb{R}$ where E is whatever. The point here is not these definitions, but the fact that they are defining elements that keep in T_pM.

This path satisfy $s(0)=\varphi\Big(\varphi^{-1}(p)+\varphi^{-1}(p)\Big)=\varphi(0)=p$. In order to prove that $\nabla_s=X+Y$ we consider $f\in C^k(M,\mathbb{R})$ and we compute $\nabla_s(f)$: SUBEQSooKOGNooGISCax

$$\frac{d}{dt}\Big[(f\circ s)(t)\Big]_{t=0}=\frac{d}{dt}\Big[(f\circ\varphi)\Big(\varphi^{-1}\big(\gamma(t)\big)+\varphi^{-1}\big(\sigma(t)\big)\Big)\Big]_{t=0} \tag{49.86a}$$

$$=\sum_k\partial_k(f\circ\varphi)(0)\frac{d}{dt}\Big[\varphi^{-1}\big(\gamma(t)\big)_k+\varphi^{-1}\big(\sigma(t)\big)_k\Big]_{t=0}. \tag{49.86b}$$

We used the theorem 12.301 with the maps $f\circ\varphi\colon\mathbb{R}^n\to\mathbb{R}$ and

$$\begin{aligned}g\colon\mathbb{R}&\to\mathbb{R}^n\\ t&\mapsto\varphi^{-1}\big(\gamma(t)\big)+\varphi^{-1}\big(\sigma(t)\big).\end{aligned} \tag{49.87}$$

We focus on one term:

$$\sum_k\partial_k(f\circ\varphi)(0)\frac{d}{dt}\Big[\varphi^{-1}\big(\gamma(t)\big)_k\Big]_{t=0}=\frac{d}{dt}\Big[(f\circ\varphi)\big(\varphi^{-1}(\gamma(t))\big)\Big]_{t=0} \tag{49.88a}$$

$$=\frac{d}{dt}\Big[f\big(\gamma(t)\big)\Big]_{t=0} \tag{49.88b}$$

$$=\nabla_\gamma(f). \tag{49.88c}$$

The two terms of (49.86) sum to $\nabla_\gamma(f)+\nabla_\sigma(f)$, so that $\nabla_s=\nabla_\gamma+\nabla_\sigma=X+Y$.

(2) **Product**

We ail to find a path $s\colon\mathbb{R}\to M$ such that $\nabla_s=\lambda X$. The answer is easy: $s(t)=\gamma(\lambda t)$. Indeed:

$$\frac{d}{dt}\Big[(f\circ s)(t)\Big]_{t=0}=\frac{d}{dt}\Big[(f\circ\gamma)(\lambda t)\Big]_{t=0}=\lambda\frac{d}{dt}\Big[(f\circ\gamma)(t)\Big]_{t=0}=\lambda X(f). \tag{49.89}$$

We used the lemma 12.176.

(3) **It is a vector space**

We have to check that T_pM is a real vector space[23]. So we have to check the conditions of definition 1.330. In particular we have to check that $(T_pM,+)$ is a group.

The neutral is the constant null operator $X(f)=0$ which is ∇_γ with γ begin the constant path $\gamma(t)=p$. Lemma 49.52 shows an inverse for the sum : $\nabla_\gamma+\nabla_\sigma=0$ when σ is defined by

$$\begin{aligned}\sigma\colon I&\to M\\ t&\mapsto\gamma(-t).\end{aligned} \tag{49.90}$$

□

49.3.3 Basis vectors

DEFooUWWWooXhHVXP

Definition 49.54.
Let M be a C^k manifold. Let $\varphi_\alpha\colon U_\alpha\to M$ be a chart around $p\in M$. We consider the path

$$\begin{aligned}\sigma\colon\mathbb{R}&\to U_\alpha\\ t&\mapsto\varphi_\alpha^{-1}(p)+te_i.\end{aligned} \tag{49.91}$$

We define the tangent vector $\partial_{\alpha,i}$ as

$$\partial_{\alpha,i}=(\varphi_\alpha\circ\sigma)'(0). \tag{49.92}$$ EQooHBUBooSuKyLm

In other words, if $f\colon M\to\mathbb{R}$ is a C^k function,

$$\partial_{\alpha,i}(f)=(f\circ\varphi_\alpha\circ\sigma)'(0)=\frac{d}{dt}\Big[\varphi_\alpha\Big(\varphi_\alpha^{-1}(p)+te_i\Big)\Big]_{t=0} \tag{49.93}$$

23. A module over $\mathbb{R}$, see definition 1.332.

PROPooCGKRooLjlULU

Proposition 49.55.
For every $X_k, Y_k \in \mathbb{R}$ we have

$$\left(\sum_k X_k \partial_k\right) + \left(\sum_k Y_k \partial_k\right) = \sum_k (X_k + Y_k)\partial_k. \tag{49.94}$$

Proof. This is merely the associativity of the sum of operators :

$$\Big(\sum_k X_k\partial_k + \sum_k Y_k\partial_k\Big)(f) = \sum_k X_k(\partial_k f) + \sum_k Y_k(\partial_k f) \tag{49.95a}$$

$$= \sum_k (X_k + Y_k)(\partial_k f). \tag{49.95b}$$

The first equality is the definition of the sum of operators. The second equality is the commutativity and associativity of the sum of reals. □

PROPooAAAXooKAMsfK

Proposition 49.56.
Let $X \in T_pM$.

(1) There exists a unique choice of $X_k \in \mathbb{R}$ such that

$$X(f) = \sum_{i=1}^{n} X_i(\partial_i f) \tag{49.96}$$

for every C^k function f.

ITEMooLWEHooRQwZQW

(2) If $k \neq l$, then $\partial_k \neq \partial_l$.

ITEMooJRPCooOGopSu

(3) The vectors

$$\partial_i = \frac{d}{dt}\left[\varphi_\alpha(s_0 + te_i)\right]_{t=0} \tag{49.97}$$

form a basis of the vector space[24] *T_pM.*

ITEMooPJHVooOHNSNP

(4) We have

$$\dim(T_pM) = \dim(M) \tag{49.98}$$

Proof. Two parts.

(1) **Existence**
Let $\varphi\colon U \to M$ be a chart of a neighbourhood of p with $\varphi^{-1}(p) = 0$. Let $f\colon M \to \mathbb{R}$ be a C^k function. We say $X = \nabla_\gamma$ with $\gamma\colon \mathbb{R} \to M$ and we define

$$\begin{aligned} \tilde{\gamma}\colon \mathbb{R} &\to U_\alpha \\ t &\mapsto (\varphi^{-1} \circ \gamma)(t). \end{aligned} \tag{49.99}$$

Let's compute:

$$X(f) = (f \circ \gamma)'(0) \tag{49.100a}$$

$$= \big((f \circ \varphi) \circ \tilde{\gamma}\big)'(0) \tag{49.100b}$$

$$= \sum_k \partial_k(f \circ \varphi)\big(\tilde{\gamma}(0)\big)\tilde{\gamma}'(0)_k \tag{49.100c}$$

SUBEQooJVQNooEuipsc

$$= \sum_k (\partial_k f)\tilde{\gamma}'(0)_k. \tag{49.100d}$$

SUBEQooBQCSooSQldjW

Note. In (49.100c), the symbol ∂_k stands for the usual partial derivative of the function $f \circ \varphi\colon U \to \mathbb{R}$ while in (49.100d), the symbol ∂_k stands for the tangent vector defined in (49.92).
We see that $X_k = \tilde{\gamma}'(0)_k$ makes the job. This shows the existence part.

24. Proposition 49.55 for the vector space structure on T_pM.

(2) **Unicity** Suppose $X = \sum_k X_k \partial_k$. With $X = \nabla_\gamma$ and $\gamma\colon \mathbb{R} \to M$. We consider

$$\begin{aligned} f_i\colon M &\to \mathbb{R} \\ m &\mapsto (\texttt{proj}_i \circ \varphi^{-1})(m). \end{aligned} \tag{49.101}$$

We have

$$X(f_i) = \sum_k X_k(\partial_k f_i) \tag{49.102a}$$
$$= \sum_k (f_i \circ \varphi \circ \sigma_i)'(0) \tag{49.102b}$$
$$= \sum_k \sum_l \partial_l(f_i \circ \varphi)\big(\sigma_i(0)\big)\big(\sigma_i'(0)\big)_l. \tag{49.102c}$$

On the one hand, from the expression $\sigma_i(t) = \varphi^{-1}(p) + te_i$ we get $\sigma_i'(0)_l = \delta_{il}$. On the other hand we have

$$f_i \circ \varphi = \texttt{proj}_i \circ \varphi^{-1} \circ \varphi = \texttt{proj}_i, \tag{49.103}$$

and $\partial_k(\texttt{proj}_i) = \delta_{ki}$. Using these two relations we have

$$X(f_i) = \sum_k X_k \delta_{ki} = X_i. \tag{49.104}$$

Now we compute the same, but using γ. We define $\tilde{\gamma} = \varphi^{-1} \circ \gamma]$. And we consider the functions $a_k\colon \mathbb{R} \to \mathbb{R}$ defined by

$$\tilde{\gamma}(t) = \sum_k a_k(t) e_k. \tag{49.105}$$

We have

$$X(f_i) = (f_i \circ \varphi \circ \tilde{\gamma})'(0) \tag{49.106a}$$
$$= (\texttt{proj}_i \circ \tilde{\gamma})'(0) \tag{49.106b}$$
$$= a_i'(0). \tag{49.106c}$$

Thus we have $X_i = a_i'(0)$, which proves the unicity.

(3) **For (2)** Let $k \neq l$. Apply ∂_k and ∂_l on the function f_k. We have $\partial_k(f_k) = 1$ and $\partial_l(f_k) = 0$.

(4) **For (3)** The fact that T_pM is a vector space is proposition 49.53(3). We know that every element in T_pM can be written as a finite linear combination of $\{\partial_k\}_{k=1,\dots,n}$ in a unique way. Proposition 4.9 says that $\{\partial_k\}_{k=1,\dots,n}$ is a basis of T_pM.
If you doubt about the cardinal, let $B = \{\partial_k\}_{k=1,\dots,n}$. The map

$$\begin{aligned} \psi\colon \{1,\dots,n\} &\to B \\ k &\mapsto \partial_k \end{aligned} \tag{49.107}$$

is obviously surjective. It is injective by point (2). Thus it is a bijection. The definition 1.126 says that $\mathrm{Card}(B) = n$.

(5) **For (4)** The vector space T_pM has a basis of cardinal n. Thus its dimension is n (definition 4.16).

□
PROPooJVSQooGvNqIx

Proposition 49.57 ([1]).
Let V be a n-dimensional vector space.

(1) V has a structure of n-dimensional manifold.

(2) For every $a \in V$, the map

$$\begin{aligned} i\colon V &\to T_aV \\ v &\mapsto \sum_i v_i \partial_i \end{aligned} \tag{49.108}$$

is a vector space isomorphism.

49.3.4 Tangent vector on a product manifold

LEMooTONEooFiysTA

Lemma 49.58.
Let M and N be C^k manifolds[25]*. Let $\gamma_M : I \to M$ and $\gamma_N : I \to N$ be C^k paths. We define*

$$\begin{aligned}\gamma : I &\to M \times N \\ t &\mapsto \big(\gamma_M(t), \gamma_N(t)\big).\end{aligned} \tag{49.109}$$

We have:

(1) A vector space isomorphism $T_{(a,b)(M\times N)} \simeq T_aM \times T_bN$.

(2) Under that isomorphism,

$$\frac{d}{dt}\Big[\gamma(t)\Big]_{t=0} = \big(\gamma'_M(0), \gamma'_N(0)\big). \tag{49.110}$$

49.4 Differential of a map

DEFooDRGUooDPFIJa

Proposition-Definition 49.59 (Differential of a map[1])**.**
Let M, N be C^1-manifolds. Let $a \in M$. Let $\phi \in C^1(M,N)$. We consider two paths γ and σ such that $\gamma(0) = \sigma(0) = a$ and

$$\nabla_\gamma = \nabla_\sigma. \tag{49.111}$$

Then

$$\nabla_{\phi\circ\gamma} = \nabla_{\phi\circ\sigma}. \tag{49.112}$$

We define

$$\begin{aligned}d\phi_a : T_aM &\to T_{\phi(a)}N \\ \nabla_\gamma &\mapsto \nabla_{\phi\circ\gamma}.\end{aligned} \tag{49.113}$$

EQooQNZPooMVaSQC

We have the formula

$$d\phi_a(\nabla_\gamma)f = \nabla_{\phi\circ\gamma}(f) = \frac{d}{dt}\Big[(f\circ\phi\circ\gamma)(t)\Big]_{t=0} \tag{49.114}$$

EQooEWMRooFsSVpb

where $a = \gamma(0)$.

Proof. Let $f \in C^1(N,\mathbb{R})$. The map $f\circ\phi : M \to \mathbb{R}$ belongs to $C^1(M,\mathbb{R})$ so that we can apply ∇_γ and ∇_σ on it. By hypothesis,

$$\nabla_\gamma(f\circ\phi) = \nabla_\sigma(f\circ\phi). \tag{49.115}$$

Using the definition of ∇,

$$\nabla_\gamma(f\circ\phi) = \frac{d}{dt}\Big[(f\circ\phi\circ\gamma)(t)\Big]_{t=0} = \nabla_{\phi\circ\gamma}(f). \tag{49.116}$$

and

$$\nabla_\sigma(f\circ\phi) = \frac{d}{dt}\Big[(f\circ\phi\circ\sigma)(t)\Big]_{t=0} = \nabla_{\phi\circ\sigma}(f). \tag{49.117}$$

□

PROPooALATooGgcVQV

Proposition 49.60 ([1])**.**
Let M, N be C^1-manifolds. Let I be an interval around 0 in $\mathbb{R}$ and a C^1 path $\gamma : I \to M$. We consider a C^1 map $\phi : M \to N$ and a function $f : N \to \mathbb{R}$.

We have the formula

$$\frac{d}{dt}\Big[(f\circ\phi)\big(\gamma(t)\big)\Big]_{t=t_0} = d\phi_{\gamma(t_0)}\big(\gamma'(t_0)\big)f, \tag{49.118}$$

EQooYHBZooUqInIC

or

$$d\phi_{\gamma(t_0)}\big(\gamma'(t_0)\big) = \frac{d}{dt}\Big[(\phi\circ\gamma)(t)\Big]_{t=t_0}. \tag{49.119}$$

EQooVLUIooAbGZEj

25. The product of manifolds is defined in 49.46.

Proof. It's a computation. We write $s(t) = \gamma(t_0 + t)$ and

SUBEQooMYOQooTByFTu
$$\frac{d}{dt}\Big[(f\circ\phi)\big(\gamma(t)\big)\Big]_{t=t_0} = \frac{d}{dt}\Big[(f\circ\phi)\big(s(t)\big)\Big]_{t=0} \tag{49.120a}$$
SUBEQooNINCooIwordP
$$= d\phi_{s(0)}\big(s'(0)\big)f \tag{49.120b}$$
$$= d\phi_{\gamma(t_0)}\big(\gamma'(t_0)\big)f. \tag{49.120c}$$

Justifications.

— For (49.120a). The left hand side of is a classical derivative of the map $f\circ\phi\circ\gamma\colon I\to\mathbb{R}$.

— For (49.120b). This is formula (49.114).

□

LEMooBOZBooNJMccB

Lemma 49.61 ([1])**.**
Let M, N be C^1-manifolds. We consider C^1-maps $f\colon M\to N$ and $\phi\colon N\to\mathbb{R}$. Let $\gamma\colon\mathbb{R}\to M$ be a C^1-path. We have

$$\nabla_\gamma(\phi\circ f) = \nabla_{f\circ\gamma}(\phi). \tag{49.121}$$

Proof. By definition,

$$\nabla_\gamma(\phi\circ f) = \frac{d}{dt}\Big[(\phi\circ f\circ\gamma)(t)\Big]_{t=0} = \frac{d}{dt}\Big[\phi\big((f\circ\gamma)(t)\big)\Big]_{t=0} = \nabla_{f\circ\gamma}(\phi). \tag{49.122}$$

□

Here we prove that df_a is linear. The proposition 49.77 will provide the reciprocal map : $(df_a)^{-1} = (df^{-1})_{f(a)}$.

PROPooNRLVooChhiIS

Proposition 49.62.
Let M and N be C^1-manifolds. Let $f\in C^1(M,N)$ and $a\in M$. The map $df_a\colon T_aM\to T_{f(a)}N$ is linear.

Proof. We consider two tangent vectors ∇_γ and ∇_σ to M at a. We know that there exists a path $s\colon\mathbb{R}\to M$ such that $\nabla_\gamma+\nabla_\sigma = \nabla_s$. This is proposition 49.53, see equation (49.85).

Let $\phi\in C^1(N,\mathbb{R})$ be a test function; we have

EQooIMZIooQlZQDR
$$df_a(\nabla_\gamma+\nabla_\sigma)\phi = df_a(\nabla_s)\phi \tag{49.123a}$$
EQooXMHHooGnbAge
$$= \nabla_s(\phi\circ f) \tag{49.123b}$$
$$= \nabla_\gamma(\phi\circ f) + \nabla_\sigma(\phi\circ f) \tag{49.123c}$$
EQooHMZAooArgfTN
$$= \nabla_{f\circ\gamma}(\phi) + \nabla_{f\circ\sigma}(\phi) \tag{49.123d}$$
$$= df_a(\nabla_\gamma)\phi + df_a(\nabla_\sigma)\phi. \tag{49.123e}$$

Justifications:

— For (49.123a): definition od the path s.

— For (49.123b): definition (49.113) of df_a.

— For (49.123d): lemma 49.61.

□

The following lemma will be generalized to vector fields in 49.86.

LEMooSCVHooYPiGse

Lemma 49.63.
Let $\varphi\colon U\to M$ be a chart around $a\in M$ with $\varphi(s) = a$ $(s\in U)$. We have $v\in T_aM$ if and only if there exists reals numbers $\{v_k\}_{k=1,\dots,n}$ such that

EQooNEDSooOhyrCZ
$$v(f) = \sum_{k=1}^{n} v_k\partial_k(f\circ\varphi)(s). \tag{49.124}$$

Proof. Two parts.

(1) $\Rightarrow$ Let $\gamma\colon \mathbb{R} \to M$ be a path for the vector v: $v = \nabla_\gamma$. By the regularity hypothesis [26], the maps $f \circ \varphi\colon U \to \mathbb{R}$ and $\varphi^{-1} \circ \gamma\colon \mathbb{R} \to U$ are differentiable usual maps.

Thus, using the equalities of lemma 12.255 and theorem 10.276, we can write

$$\begin{aligned}
\nabla_\gamma f &= \frac{d}{dt}\Big[(f \circ \gamma)(t)\Big]_{t=0} && \text{(49.125a)}\\
&= \frac{d}{dt}\Big[(f \circ \varphi) \circ (\varphi^{-1} \circ \gamma)(t)\Big]_{t=0} && \text{(49.125b)}\\
&= d(f \circ \varphi)_{(\varphi^{-1}\circ\gamma)(0)}\big((\varphi^{-1} \circ \gamma)'(0)\big) && \text{(49.125c)}\\
&= d(f \circ \varphi)_s\big((\varphi^{-1} \circ \gamma)'(0)\big) && \text{(49.125d)}\\
&= \sum_k \partial_k(f \circ \varphi)(s)(\varphi^{-1} \circ \gamma)'(0)_k && \text{(49.125e)}
\end{aligned}$$

Let $v_k = (\varphi^{-1} \circ \gamma)'(0)_k$ and we have the result.

(2) $\Leftarrow$

□

PROPooLHMSooMMXrSS

Proposition 49.64.
Let M be a n-dimensional manifold. Let $X \in T_pM$ be a tangent vector. There exists a unique $v \in \mathbb{R}^n$ such that

$$X = \frac{d}{dt}\left[\varphi_\alpha(s_0 + tv)\right]_{t=0} \tag{49.126}$$

where $s_0 = \varphi_\alpha^{-1}(p)$.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

LEMooZXEFooZgXbNP

Lemma 49.65.
Let $\varphi\colon U \to M$ be a chart around $a \in M$ with $\varphi(s) = a$ $(s \in U)$. Let $v \in T_aM$. We have

$$v(f) = \sum_{k=1}^{n} v_k \partial_k(f \circ \varphi)(s) \tag{49.127}$$

with $v = (\varphi^{-1} \circ \gamma)'(0)$.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

NORMooXAJGooDNyxjv

49.66.
If $v \in \mathbb{R}^n$ and if $a \in \mathbb{R}^n$, we can speak of $v \in T_a\mathbb{R}^n$ with the abuse of notation $v = \nabla_\gamma$ where

$$\begin{aligned}
\gamma\colon \mathbb{R} &\to \mathbb{R}^n\\
t &\mapsto a + tv.
\end{aligned} \tag{49.128}$$

You have to keep in mind that v is an element of $\mathbb{R}^n$ (a list of numbers) while ∇_γ is an element of $T_a\mathbb{R}^n$ (a differential operator). Writing «$v = \nabla_\gamma$» is an abuse of notation.

The object ∇_γ is what one could name «the vector v tied to the point a».

This remark is formalised by the following proposition which provides a canonical isomorphism between $\mathbb{R}^n$ and $T_a\mathbb{R}^n$.

PROPooRXIIooFmhqJd

Proposition 49.67.
Let M be a C^k manifold and $a \in M$. We consider a chart $\varphi\colon U \to M$ around a. For $v \in \mathbb{R}^n$ we define

$$\gamma_{a,v}(t) = \varphi\big(\varphi^{-1}(a) + tv\big). \tag{49.129}$$

26. The manifold M is C^k, the charts maps are C^k and all that.

(1) *The map* $\gamma_{a,v}: \mathbb{R} \to M$ *is* C^k.

(2) *The map*

$$\begin{aligned} \psi: \mathbb{R}^n &\to T_aM \\ v &\mapsto \nabla_{\gamma_{a,v}} \end{aligned} \tag{49.130}$$

is a bijection.

(3) *We have the equalities*

$$\psi(v)f = d\varphi_s(v)f = \nabla_{s,v}(f) = \frac{d}{dt}\Big[(f\circ\varphi)\big(\varphi^{-1}(a)+tv\big)\Big]_{t=0} = \sum_k v_k \partial_k (f\circ\varphi)\big(\varphi^{-1}(x)\big) \tag{49.131}$$

where $\nabla_{s,v}$ *is the operator defined in proposition 49.67.*

Proof. The map $\gamma_{a,v}$ is continuous as composed of continuous maps. The set $\mathbb{R}$ is a manifold with the identity as charts. Thus the condition 49.35(2b) to be checked reduces to

$$\tilde{\gamma}_{a,v} = \varphi^{-1} \circ \gamma_{a,v}. \tag{49.132}$$

We consider the C^∞ map

$$\begin{aligned} l: \mathbb{R} &\to \mathbb{R}^n \\ t &\mapsto \varphi^{-1}(a) + tv. \end{aligned} \tag{49.133}$$

We have $\tilde{\gamma}_{a,v} = \varphi^{-1} \circ \varphi \circ l$. Thus $\tilde{\gamma}$ is C^∞.

Remain to prove that ψ is a bijection.

(1) **Injective** Suppose that $\psi(v) = \psi(w)$. For every $f \in C^k(M)$ we have $\nabla_{a,v}(f) = \nabla_{a,w}(f)$, which means

$$\frac{d}{dt}\Big[f\Big(\varphi\big(\varphi^{-1}(a)+tv\big)\Big)\Big]_{t=0} = \frac{d}{dt}\Big[f\Big(\varphi\big(\varphi^{-1}(a)+tw\big)\Big)\Big]_{t=0}. \tag{49.134}$$

EQooMXJWooQKpzCG

We apply this to the function $f = \texttt{proj}_k \circ \varphi^{-1}$:

$$f\Big(\varphi\big(\varphi^{-1}(a)+tv\big)\Big) = \varphi^{-1}(a)_k + tv_k, \tag{49.135}$$

so that

$$\nabla_{a,v}(f) = \frac{d}{dt}\Big[\varphi^{-1}(a)_k + tv_k\Big]_{t=0} = v_k. \tag{49.136}$$

The equation (49.134) implies $v_k = w_k$ for every k.

(2) **Surjective** The map $f \circ \varphi: U \to \mathbb{R}$ is a usual C^k function. The formulas of the lemma 12.255 are valid; in particular the ones concerning the directional derivative. We have

$$\begin{aligned} \nabla_{a,v}(f) &= \frac{d}{dt}\Big[(f\circ\varphi)\big(\varphi^{-1}(a)+tv\big)\Big]_{t=0} && \text{(49.137a)} \\ &= \partial_v(f\circ\varphi)\big(\varphi^{-1}(a)\big) && \text{(49.137b)} \\ &= \sum_k v_k \partial_k (f\circ\varphi)\big(\varphi^{-1}(a)\big). && \text{(49.137c)} \end{aligned}$$

This is the general form (49.124) for the action of a tangent vector on f.

□

PROPooKMCGooDEuaWz

Proposition 49.68.
Let $\varphi: U \to M$ *be a chart satisfying* $\varphi(s) = a$.

ITEMooSFUBooNXgGuu

(1) *The set*[27] $\{d\varphi_s(e_i)\}_{i=1,\dots,n}$ *is a basis of* T_aM.

27. We use the abuse of notation of 49.66.

ITEMooPYPVooKkHrkQ

(2) For every $X \in T_aM$, there exists an unique $v \in \mathbb{R}^n$ such that

$$X = d\varphi_s(v). \tag{49.138}$$

Proof. For part (1), we have to prove that the vectors of the form $d\varphi_s(e_i)$ are linearly independent and spanning T_aM.

(1) **Spanning** Lemma 49.63 says that, if $v \in T_aM$, there exist numbers v_k such that

$$v(f) = \sum_{k=1}^{n} v_k \partial_k (f \circ \varphi)(s) \tag{49.139}$$

We consider the path

$$\begin{aligned} \gamma_k \colon \mathbb{R} &\to U \\ t &\mapsto s + te_k. \end{aligned} \tag{49.140}$$

Now we have

$$v(f) = \sum_{k=1}^{n} v_k \partial_k (f \circ \varphi)(s) \tag{49.141a}$$

$$= \sum_k v_k \frac{d}{dt}\Big[(f \circ \varphi)(s + te_k)\Big]_{t=0} \tag{49.141b}$$

$$= \sum_k v_k \frac{d}{dt}\Big[(f \circ \varphi \circ \gamma_k)(t)\Big]_{t=0} \tag{49.141c}$$

$$= \sum_k v_k d\varphi_{\gamma_k(0)}(\nabla_{\gamma_k}) f \tag{49.141d}$$

$$= \sum_k v_k d\varphi_s(\nabla_{\gamma_k}) f \tag{49.141e}$$

$$= \sum_k v_k d\varphi_s(e_k). \tag{49.141f}$$

The last equality is the abuse of notation explained in 49.66.

(2) **Independent** We suppose that

$$\sum_{k=1}^{n} v_k d\varphi_s(e_k) = 0 \tag{49.142}$$

for some numbers v_k. It means that for every C^k functions $f \colon M \to \mathbb{R}$ we have

$$\sum_k v_k \partial_k (f \circ \varphi)(s) = 0. \tag{49.143}$$

Since the function

$$\begin{aligned} \mathtt{proj}_i \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto x_i \end{aligned} \tag{49.144}$$

is C^∞, we can choose $f = \mathtt{proj}_k \circ \varphi^{-1}$. Then we have

$$0 = \sum_k v_k \partial_k (\mathtt{proj}_i \circ \varphi^{-1} \circ \varphi)(s) \tag{49.145a}$$

$$= \sum_k v_k \partial (s \mapsto s_i)(s) \tag{49.145b}$$

$$= \sum_k v_k \delta_{ki} \tag{49.145c}$$

$$= v_i. \tag{49.145d}$$

We conclude that $v_i = 0$ for every i and we are done.

Now we prove part (2). From part (1) there is a set of numbers v_i such that

$$X = \sum_i v_i \tag{49.146}$$

We pose $v = \sum_i v_i e_i$ and we use the linearity of $d\varphi_s$:

$$X = \sum_i v_i d\varphi_s(e_i) = \sum_i d\varphi_s(v_i e_i) = d\varphi_s\big(\sum_i v_i e_i\big) = d\varphi_s(v). \tag{49.147}$$

For the unicity, let $v, w \in \mathbb{R}^n$ such that $d\varphi_s(v) = d\varphi_s(w)$. Thus we have $d\varphi_s(v - w) = 0$ which implies $v = w$. □

Lemma 49.69 ([1]).
Let a and b be different points in the manifold M. Then

$$T_aM \cap T_bM = \{0\}. \tag{49.148}$$

Proof. The operator which maps every function to zero belongs to T_aM and T_bM. This is easy. The tricky part is the contrary. Let $v \in T_aM$ and $w \in T_bM$ both non zero. We consider charts $\varphi\colon U \to M$ and $\psi\colon V \to M$ such that

— $U \cap V = \varnothing$
— $\varphi(U) \cap \psi(V) = \varnothing$
— $a = \varphi(s)$ and $b = \psi(t)$ with $s \in U$ and $t \in V$.

Let $\phi \in C^k(M, \mathbb{R})$. By lemma 49.63 we have

$$v(\phi) = \sum_k v_k \partial_k(\phi \circ \varphi)(s) \tag{49.149}$$

EQooTQKZooIeQNaU

and

$$w(\phi) = \sum_k w_k \partial_k(\phi \circ \varphi)(t). \tag{49.150}$$

The trick now is to build a function ϕ for which $v(\phi) \neq w(\phi)$.

Let $r > 0$ such that $\overline{B(s,r)} \subset U$. We use the Urysohn lemma 27.74 to create a function $p\colon U \to \mathbb{R}$ such that

— $p = 1$ on a neighbourhood of s,
— $p = 0$ outside $\overline{B(s,r)}$,
— $p \in C^\infty(U)$.

We also consider a C^∞ function $q\colon U \to \mathbb{R}$ such that $(\partial_k q)(s) = \alpha_k$ for some numbers α_k to be fixed later.

Finally we build

$$\begin{aligned} \phi\colon M &\to \mathbb{R} \\ x &\mapsto \begin{cases} (pq)\big(\varphi^{-1}(x)\big) & \text{if } x \in \varphi(U) \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{49.151}$$

This function is C^k. Indeed there are two possibilities : $x \in \varphi(U)$ or $x \notin \varphi(U)$. In the first case ϕ is the composition of pq (which is C^∞) with φ^{-1} which is C^k. If $x \notin \varphi(U)$, then it is in particular outside $\varphi\big(\overline{B(s,r)}\big)$ which is closed.

The set of points outside of $\varphi\big(\overline{B(s,r)}\big)$ is open. Thus there is a neighbourhood of x which does not intersect $\varphi\big(\overline{B(r,s)}\big)$. The function ϕ is zero in this neighbourhood, so that ϕ is C^k.

We can compute the values of $v(\phi)$ and $w(\phi)$. The easiest if $w(\phi) = 0$ because $\phi = 0$ on a neighbourhood of b. For $v(\phi)$ we use the formula 49.149. We have

$$\partial_k(\phi \circ \varphi)(s) = \partial_k(q)(s) = \alpha_k \tag{49.152}$$

because $p = 1$ on a neighbourhood of $\varphi^{-1}(a)$. Thus we have

$$v(\phi) = \sum_k v_k \alpha_k. \tag{49.153}$$

We can choose the α_k in such a way that $v(\phi) \neq 0$ because $v \neq 0$. □

PROPooLJYEooMjevio

Proposition 49.70.
Let M be a submanifold of the manifold N. If $p \in M$, then there exists a coordinate system $\{x_1, \dots, x_n\}$ on a neighbourhood of p in N such that $x_1(p) = \dots = x_n(p) = 0$ and such that the set

$$U = \{q \in V \text{ tel que } x_j(q) = 0 \, \forall \, m+1 \leqslant j \leqslant n\}$$

gives a local chart of M containing p.

The sense of this proposition is that one can put p at the center of a coordinate system on N such that M is just a submanifold of N parametrised by the fact that its last $m - n$ components are zero.

lem:var_cont_diff

Lemma 49.71.
Let V, M be two manifolds and $\varphi\colon V \to M$, a differentiable map. We suppose that $\varphi(V)$ is contained in a submanifold S of M. If $\varphi\colon V \to S$ is continuous[28]*, then it is differentiable.*

Proof. Let $p \in V$. By proposition 49.70, we have a coordinate system $\{x_1, \dots, x_m\}$ valid on a neighbourhood N of $\varphi(p)$ in M such that the set

$$\{r \in N \text{ tel que } x_j(r) = 0 \, \forall s < j \leqslant m\}$$

with the restriction of $(x_1, \dots x_s) \in N_S$ form a local chart which contains $\varphi(p)$. From the continuity of φ, there exists a chart (W, ψ) around p such that $\varphi(W) \subset N_S$. The coordinates $x_j(\varphi(q))$ are differentiable functions of the coordinates of q in W. In particular, the coordinates $x_j(\varphi(q))$ for $1 \leqslant j \leqslant s$ are differentiable and $\varphi\colon V \to S$ is differentiable because its expression in a chart is differentiable. □

A consequence of this lemma: if V and S are submanifolds of M with $V \subset S$, and if S has the induced topology from M, then V is a submanifold of S. Indeed, we can consider the inclusion $\iota\colon V \to S$: it is differentiable from V to M and continuous from V to S then it is differentiable from V to S by the lemma. Thus $V = \iota^{-1}(S)$ is a submanifold of S (this is a classical result of differential geometry).

PROPooNAACooIUMcNw

Proposition 49.72 ([1])**.**
Let $f\colon M \to \mathbb{R}$ be a smooth manifold. Then

$$df_m = \sum_{k=1}^{n} (\partial_k f)(m) \partial_k^* \tag{49.154}$$

where ∂_k is the vector field.

49.4.1 Chain rule and inverse

LEMooEGITooXbAPDe

Lemma 49.73.
If M is a C^k manifold, and if $\mathrm{Id}\colon M \to M$ is the identity map, we have

$$d\,\mathrm{Id}_a = \mathrm{Id}_{T_aM}\,. \tag{49.155}$$

In other words, the differential of the identity map is the identity.

28. This hypothesis states the continuity for the topology of S, which is different from the continuity with respect to the topology of M.

LEMooGRRAooXxDMuw

Lemma 49.74 (Chain rule[832]).
Let M_i be C^k manifolds. If the maps $g\colon M_1 \to M_2$ and $f\colon M_2 \to M_3$ are C^k, thus the composition $f \circ g$ is C^k and for every $a \in M_1$ we have

$$d(f \circ g)_a = df_{g(a)} \circ dg_a. \tag{49.156}$$

PROBooIRONooNAYXjo

¡¡ Avertissement/question à la lectrice !! 49.75
I'm not totally sure of the statement for the local inverse 49.76.

THOooUTSRooTqBegO

Theorem 49.76 (Local inverse[1]).
Let M and N be C^p-manifolds. Let $f \in C^p(M,N)$ and $a \in M$ such that df_a is a bicontinuous bijection. Then there exists open neighbourhood V of a and an open neighbourhood W of $f(a)$ such that

(1) $f\colon V \to W$ is bijective and C^p.
(2) $f^{-1}\colon W \to v$ is C^p.

PROPooPEMLooPQcywG

Proposition 49.77 ([832]).
Let $f\colon M \to N$ be a diffeomorphism between the C^k manifolds M and N. For every $a \in M$, the map $df_a\colon T_aM \to T_{f(a)}N$ is a vector space isomorphism and the inverse is given by

$$(df_a)^{-1} = (df^{-1})_{f(a)}. \tag{49.157}$$

Proof. The linearity of df_a is the proposition 49.62. Since f is a diffeomorphism we have the equality $f^{-1} \circ f = \mathrm{Id}_M$. Using the chain rule of lemma 49.74 and the differential of the identity of lemma 49.73, we get

$$(df^{-1})_{f(a)} \circ df_a = \mathrm{Id}\,. \tag{49.158}$$

The same with $f \circ f^{-1} = \mathrm{Id}$ provides

$$df_a \circ (df^{-1})_{f(a)} = \mathrm{Id}\,. \tag{49.159}$$

This proves that df_a is invertible and that its inverse is $(df^{-1})_{f(a)}$.

□

49.4.2 Topology on a tangent space

PROPooHJOXooMGANfd

Proposition-Definition 49.78 ([833]).
Let M be a C^k manifold. Let $a \in M$ and $\varphi\colon U \to M$ be a chart around a. We define $s = \varphi^{-1}(a)$. For $v \in T_aM$ we define

$$\|v\|_{T_aM} = \|(d\varphi_s)^{-1}(v)\|_{\mathbb{R}^n}. \tag{49.160}$$

This is a norm on the vector space T_aM.

The topology on T_aM is the one induced by this norm[29].

Proof. Several points.

(1) $\|v\| \geqslant 0$ From the very definition, yes.
(2) $\|v\| = 0$ **si et seulement si** $v = 0$ If $\|v\| = 0$, then $\|(d\varphi_s)^{-1}(v)\|_{\mathbb{R}^n} = 0$. Since the norm on $\mathbb{R}^n$ is a norm, this implies $(d\varphi_s)^{-1}(v) = 0$. And since $d\varphi_s$ is a linear bijection, we conclude $v = 0$.
(3) $\|\lambda v\| = |\lambda|\|v\|$ Because $d\varphi_s^{-1}$ is linear.
(4) $\|v + w\| \leqslant \|v\| + \|v\|$ Because $d\varphi_s^{-1}$ is linear and the corresponding property on $\mathbb{R}^n$.

□

Now we are allowed to write $\|v\|$ when $v \in T_aM$. But we have to keep in mind that it depends on the choice of a local chart.

29. Keep in mind that, since T_aM is finite dimensional, all the norm are equivalent (theorem 10.49), so that this norm is not special.

49.4.3 Smooth invariance of the domain

DEFooLXVNooIpoIBp

Definition 49.79 (Regularity on non open domains[834]).
Let M be a differential manifold and $A \subset M$. We say that a map $f\colon A \to \mathbb{R}$ is of class C^k if there exists an open S such that $A \subset S \subset M$ and a map $f_0\colon S \to \mathbb{R}$ such that $f_0 = f$ on A and f_0 is of class C^k on S.

THOooGUELooSBQhRu

Theorem 49.80 (Smooth invariance of the domain[835]).
Let M and N be smooth manifolds of same dimension. Let $U \subset M$ be open and $S \subset N$. We consider a C^1-diffeomorphism[30] *$f\colon U \to S$. Then S is an open part of N.*

Proof. We denote by τ_M the topology[31] of M and by τ_N the topology of N

Let $s \in S$. We will prove that there exists an open $A \in \tau_N$ such that $s \in A \subset S$. The theorem 7.8 will conclude.

Let $p = f^{-1}(s)$. We know that f^{-1} is C^1 at s in the sense of extensions. Thus there exists opens parts $V \in \tau_N$ and $W \in \tau_M$ containing s and a diffeomorphism $g\colon V \to W$ such that $g|_{S\cap V} = f^{-1}$. Using the chain rule (lemma 49.74), for every $q \in V \cap S$ we have

$$dg_{f(q)} \circ df_q = \mathrm{Id}\,. \tag{49.161}$$

In particular this is true for $q = p$: we have $dg_{f(p)} \circ df_p = \mathrm{Id}$. We deduce that the map $df_p\colon T_pM \to T_{f(p)}N$ is injective. By proposition 49.56 we have $\dim(T_pM) = \dim(T_{f(p)}N)$ and df_p is a vector space isomorphism.

The local inverse theorem 49.76 says that f is a local C^1-diffeomorphism: there is $\mathcal{O} \in \tau_M$ containing p such that $f(\mathcal{O}) \in \tau_N$. More generally, if $\mathcal{O}' \in \tau_M$ is contained in $\mathcal{O}$, then $f(\mathcal{O}')$ is open in N.

Taking $\mathcal{O}' = U \cap \mathcal{O}$, the part $f(\mathcal{O}')$ belongs to τ_N, contains s and is contained in S. □

49.5 Vector field

LEMooXFNQooXwCMNB

Lemma 49.81.
Let M be a manifold and (U, φ) be a chart and γ be a path. We set $a = \varphi^{-1}\big(\gamma(0)\big)$, $v = (\varphi^{-1}\circ\gamma)'(0)$ and

$$\begin{aligned}\sigma\colon \mathbb{R} &\to M\\ t &\mapsto \varphi(a + tv).\end{aligned} \tag{49.162}$$

Then we have $\nabla_\gamma = \nabla_\sigma$.

Proof. Let $\phi \in C^k(M, \mathbb{R})$.

$$\begin{aligned}
\nabla_\sigma(\phi) &= \frac{d}{dt}\Big[\phi\big(\sigma(t)\big)\Big]_{t=0} && (49.163a)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi)\big(a + t(\varphi^{-1}\circ\gamma)'(0)\big)\Big]_{t=0} && (49.163b)\\
&= \sum_k \partial_k(\phi\circ\varphi)(a)(\varphi^{-1}\circ\gamma)'(0)_k && (49.163c)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi)\big((\varphi^{-1}\circ\gamma)(t)\big)\Big]_{t=0} && (49.163d)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi\circ\varphi^{-1}\circ\gamma)(t)\Big]_{t=0} && (49.163e)\\
&= \frac{d}{dt}\Big[(\phi\circ\gamma)(t)\Big]_{t=0} && (49.163f)\\
&= \nabla_\gamma(\phi). && (49.163g)
\end{aligned}$$

SUBEQooFDVFooFzACbX

30. Definition 10.265. This is a C^1 map which is bijective and whose inverse is C^1. Here the notions of continuity and differentiability is taken on S via the "extension" technique, definition 49.79.

31. The set of open parts.

For (49.163c): the maps $\phi \circ \varphi \colon U \to \mathbb{R}$ and $\varphi^{-1} \circ \gamma \colon \mathbb{R} \to U$ are C^k maps, so that we can apply the formula (12.811): □

LEMooGPCBooXMTddG

Lemma 49.82.
Let M be a manifold, (U, φ) be a chart, $a \in U$ and $v, w \in \mathbb{R}^n$. We define $\gamma_v(t) = \varphi(a + tv)$ and $\gamma_w(t) = \varphi(a + tw)$. If $v \neq w$ then

$$\nabla_{\gamma_v} \neq \nabla_{\gamma_w}. \tag{49.164}$$

Proof. We suppose $v \neq w$. If w is a multiple of v, an adaptation of the «product» part of proposition 49.53 shows that $\nabla_{\gamma_v} \neq \nabla_{\gamma_w}$.

Of v and w are nor aligned, we consider a basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ such that $e_1 = v$ and $e_2 = w$ (theorem 4.25). Now we consider the function

$$\begin{aligned} f \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto (x - a)_1. \end{aligned} \tag{49.165}$$

We will show that the result of ∇_{γ_v} and ∇_{γ_w} on the function $f \circ \varphi^{-1}$ are not equal. First we have

$$\begin{aligned}
\nabla_{\gamma_v}(f \circ \varphi^{-1}) &= \frac{d}{dt}\Big[(f \circ \varphi^{-1})\big(\gamma_v(t)\big)\Big]_{t=0} && \text{(49.166a)} \\
&= \frac{d}{dt}\Big[(f \circ \varphi^{-1} \circ \varphi)(a + tv)\Big]_{t=0} && \text{(49.166b)} \\
&= \frac{d}{dt}\Big[f(a + tv)\Big]_{t=0} && \text{(49.166c)} \\
&= \frac{d}{dt}\Big[(tv)_1\Big]_{t=0} && \text{(49.166d)} \\
&= \frac{d}{dt}\Big[t\Big]_{t=0} && \text{(49.166e)} \\
&= 1. && \text{(49.166f)}
\end{aligned}$$

In the same way we get

$$\nabla_{\gamma_w}(f \circ \varphi^{-1}) = \frac{d}{dt}\Big[f(a + tw)\Big]_{t=0} = \frac{d}{dt}\Big[(tw)_1\Big]_{t=0} = \frac{d}{dt}\Big[0\Big]_{t=0} = 0. \tag{49.167}$$

□

PROPooOSCIooFPRtRg

Proposition 49.83 ([1])**.**
Let $\varphi_\alpha \colon U_\alpha \to M$ be a chart on the smooth manifold M. We have

$$\frac{d}{dt}\left[\phi_\alpha(x + tv)\right]_{t=0} = \sum_i v_i \partial_{\beta,i}. \tag{49.168}$$

PROPooMEPPooRonxuh

Proposition 49.84.
Let M be a manifold, (U, φ) a chart. If $s \in U$ the map[32]

$$\begin{aligned} \psi \colon \mathbb{R}^n &\to T_{\varphi(s)}M \\ v &\mapsto \frac{d}{dt}\Big[\varphi(s + tv)\Big]_{t=0} \end{aligned} \tag{49.169}$$

EQooJMTJooZNzREv

is a vector space isomorphism.

Proof. We need to prove that ψ is surjective, injective and linear.

(1) **Surjective** Lemma 49.81.

(2) **Injective** Lemme 49.82.

32. We use the notation (49.77).

(3) **Linear** Let $v, w \in \mathbb{R}^n$. If set $\sigma(t) = s + t(v + w)$, we have

$$\begin{aligned}
\psi(v+w)\phi &= \frac{d}{dt}\Big[(\phi\circ\varphi)\big(s+t(v+w)\big)\Big]_{t=0} && (49.170a)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi)\circ\sigma(t)\Big]_{t=0} && (49.170b)\\
&= \sum_k \partial_k(\phi\circ\varphi)\big(\sigma(0)\big)\sigma'(0)_k && (49.170c)\\
&= \sum_k \partial_k(\phi\circ\varphi)\big(\sigma(0)\big)(v_k + w_k) && (49.170d)\\
&= \sum_k \partial_k(\phi\circ\varphi)\big(\sigma(0)\big)v_k + \sum_k \partial_k(\phi\circ\varphi)\big(\sigma(0)\big)w_k && (49.170e)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi)\big(\sigma(0)+tv\big)\Big]_{t=0} + \frac{d}{dt}\Big[(\phi\circ\varphi)\big(\sigma(0)+tw\big)\Big]_{t=0} && (49.170f)\\
&= \psi(v)\phi + \psi(w)\phi. && (49.170g)
\end{aligned}$$

In the same way, if $\lambda \in \mathbb{R}$ we set $\sigma(t) = s + t\lambda v$ and we have

$$\begin{aligned}
\psi(\lambda v)\phi &= \frac{d}{dt}\Big[(\phi\circ\varphi)(s+t\lambda v)\Big]_{t=0} && (49.171a)\\
&= \frac{d}{dt}\Big[(\phi\circ\varphi)\big(\sigma(t)\big)\Big]_{t=0} && (49.171b)\\
&= d(\phi\circ\varphi)_{\sigma(0)}\sigma'(0) && (49.171c)\\
&= \lambda d(\phi\circ\varphi)_{\sigma(0)}(v) && (49.171d)\\
&= \lambda\psi(v)\phi. && (49.171e)
\end{aligned}$$

□

PROPooXURIooYPytwa

Proposition 49.85.
Let M be a C^k manifold, and $\varphi\colon U \to M$ be a local chart. A map $X\colon M \to TM$ is a C^k vector field on $\varphi(U)$ if and only it can be written under the form

$$X_x = \sum_{i=1}^n X_i(x)\partial_i \tag{49.172}$$

for some C^k maps $X_i\colon M \to \mathbb{R}$.

LEMooZWFAooDlYaJm

Lemma 49.86 ([1])**.**
Let M be a C^k manifold, and $\varphi\colon U \to M$ be a local chart. A map $X\colon M \to TM$ is a C^k vector field on $\varphi(U)$ if and only it can be written under the form

$$X_x(f) = \sum_{k=1}^n v_k(x)\partial_k(f\circ\varphi)\big(\varphi^{-1}(x)\big) \tag{49.173}$$

for some C^k maps $v_k\colon M \to \mathbb{R}$.

LEMooIQZWooOSLNXB

Lemma 49.87.
If $x \in M$ and $v \in \mathbb{R}^n$. The isomorphism ψ of proposition 49.84 satisfies

$$\psi(v)f = \sum_k v_k\partial_k(f\circ\varphi)(s) \tag{49.174}$$

EQooBVOBooBTfYWC

where $s = \varphi^{-1}(x)$.

Proof. We use the linearity of proposition 49.84:

$$\psi(v)f = \sum_k v_k \psi(e_k)f \tag{49.175a}$$
$$= \sum_k v_k \frac{d}{dt}\Big[f\big(\varphi(s+te_k)\big)\Big]_{t=0} \tag{49.175b}$$
$$= \sum_k v_k \frac{d}{dt}\Big[(f\circ\varphi)(s+te_k)\Big]_{t=0} \tag{49.175c}$$
$$= \sum_k v_k \partial_k (f\circ\varphi)(s). \tag{49.175d}$$

□

49.5.1 Vector field

DEFooAATTooLhNqDb

Definition 49.88 (Vector field)**.**
*Let M be a C^∞ manifold. A **vector field** is a map $X\colon M \to TM$ such that $X(p) \in T_pM$ for every $p \in M$.*

We will often write X_p instead of $X(p)$.

PROPooGYWRooPIyocN

Proposition 49.89.
Let M be a C^{k+1} manifold, X be a C^k vector field and $f\colon M \to \mathbb{R}$ be a C^k function. Then the map

$$\begin{aligned} X(f)\colon M &\to \mathbb{R} \\ p &\mapsto X_p(f) \end{aligned} \tag{49.176}$$

is C^{k-1}.

LEMooLNIAooCmbLQp

Lemma 49.90.
Let X be a smooth vector field on M. Let $p \in M$ such that $X_p \neq 0$. There exists a local chart $\varphi\colon U \to M$ around p such that $X = \partial_1$ in that chart. More precisely, for every $q \in \varphi(U)$ and every smooth function f on M we have

$$X_q(f) = \partial_1(f\circ\varphi)\big(\varphi^{-1}(q)\big). \tag{49.177}$$

49.6 Tangent and cotangent bundle

If M is a n dimensional manifold, as set the tangent bundle is the *disjoint* union of tangent spaces

$$TM = \bigcup_{x\in M} T_xM. \tag{49.178}$$

LEMooLZVTooSvHmjn

Lemma 49.91 ([1])**.**
Let M be a manifold. We consider the definition charts $\{U_\alpha, \varphi_\alpha\}$ of M, and the map

$$\begin{aligned} \psi_\alpha\colon U_\alpha \times \mathbb{R}^n &\to TM \\ (x,v) &\mapsto \frac{d}{dt}\Big[\varphi_\alpha(x+tv)\Big]_{t=0}. \end{aligned} \tag{49.179}$$

We have

$$\psi_\alpha(U_\alpha \times \mathbb{R}^n) = \bigcup_{m\in\varphi_\alpha(U_\alpha)} T_mM. \tag{49.180}$$

Proof. The inclusion $\psi_\alpha(U_\alpha \times \mathbb{R}^n) \subset \bigcup_{m\in\varphi_\alpha(U_\alpha)} T_mM$ comes from the fact that the formula $\frac{d}{dt}\left[\varphi_\alpha(x+tv)\right]_{t=0}$ is a vector in $T_{\varphi_\alpha(x)}M$ by proposition 49.84.

For the reverse inclusion, let $m = \varphi_\alpha(x)$, and $X \in T_mM$. Proposition 49.84 shows that there exists $v \in \mathbb{R}^n$ such that $X = \frac{d}{dt}\left[\varphi_\alpha(x+tv)\right]_{t=0} = \psi_\alpha(x,v)$. □

THOooTSQXooLvJMQb

Theorem 49.92 ([1]).
Let M be a manifold. We consider the definition charts $\{U_\alpha, \varphi_\alpha\}$ of M. Then the set TM becomes a manifold[33] *with the maps $V_\alpha = U_\alpha \times \mathbb{R}^n$ and*

$$\begin{aligned} \psi_\alpha \colon U_\alpha \times \mathbb{R}^n &\to TM \\ (x, v) &\mapsto \frac{d}{dt}\Big[\varphi_\alpha(x + tv)\Big]_{t=0}. \end{aligned} \tag{49.181}$$

Proof. We have to check the definition 49.18. In order to prove that TM is a topological manifold, we use theorem 49.10.

(1) Every element of TM has the form $\frac{d}{dt}\left[\varphi_\alpha(x + tv)\right]_{t=0}$ by proposition 49.84.

(2) The map ψ_α is injective from the same proposition 49.84.

(3) We have to check that

$$\psi_\alpha^{-1}\Big(\psi_\alpha(U_\alpha \times \mathbb{R}^n) \cap \psi_\beta(U_\beta \times \mathbb{R}^n)\Big) \tag{49.182}$$

is open. Using lemma 49.91, and writing $M_\alpha = \varphi_\alpha(U_\alpha)$,

$$\psi_\alpha(U_\alpha \times \mathbb{R}^n) \cap \psi_\beta(U_\beta \times \mathbb{R}^n) = \bigcup_{m \in M_\alpha \cap M_\beta} T_m M \tag{49.183a}$$

$$= \psi_\alpha\Big(\varphi_\alpha^{-1}(M_\alpha \cap M_\beta) \times \mathbb{R}^n\Big). \tag{49.183b}$$

Thus

$$\psi_\alpha^{-1}\Big(\psi_\alpha(U_\alpha \times \mathbb{R}^n) \cap \psi_\beta(U_\beta \times \mathbb{R}^n)\Big) = \varphi_\alpha^{-1}(M_\alpha \cap M_\beta) \times \mathbb{R}^n \tag{49.184}$$

is open.

(4) Now we prove that $\psi_\beta^{-1} \circ \psi_\beta$ is continuous. Let $v = \sum_i v_i e_i \in \mathbb{R}^n$. We have[34]

$$\psi_\beta(x, v) = \frac{d}{dt}\left[\varphi_\beta(x + tv)\right]_{t=0} = \sum_i v_i \partial_{\beta,i} = \sum_i v_i (d\varphi_\beta)_x e_i. \tag{49.185}$$

Let use check that

$$\sum_i (d\varphi_\beta)_x e_i = \psi_\alpha\Big((\varphi_\alpha^{-1} \circ \varphi_\beta)(x), \sum_i v_i d(\varphi_\alpha^{-1} \circ d\varphi_\beta)_x e_i\Big). \tag{49.186}$$

This is a computation

$$\psi_\alpha\Big((\varphi_\alpha^{-1} \circ \varphi_\beta)(x), (d\varphi_\alpha^{-1} \circ d\varphi_\beta)_x e_i\Big) \tag{49.187a}$$

$$= \frac{d}{dt}\left[\varphi_\alpha\Big((\varphi_\alpha^{-1} \circ \varphi_\beta)(x) + t\sum_i v_i (d\varphi_\alpha^{-1} \circ d\varphi_\beta)_x e_i\Big)\right]_{t=0} \tag{49.187b}$$

$$= (d\varphi_\alpha)_{(\varphi_\alpha^{-1} \circ \varphi_\beta)(x)}\Big(\sum_i v_i (d\varphi_\alpha^{-1} \circ d\varphi_\beta)_x e_i\Big) \tag{49.187c}$$

$$= \sum_i v_i d(\varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_\beta)_x e_i \tag{49.187d}$$

$$= \sum_i v_i (d\varphi_\beta)_x e_i. \tag{49.187e}$$

This shows that

$$(\psi_\alpha^{-1} \circ \psi_\beta)(x, v) = \Big((\varphi_\alpha^{-1} \circ \varphi_\beta)(x), \sum_i v_i (d\varphi_\alpha^{-1} \circ d\varphi_\beta)_x e_i\Big). \tag{49.188}$$

EQooHAUMooRgHKlL

Since the maps $\varphi_\alpha^{-1} \circ \varphi_\beta$ are continuous (even smooth), we see that the maps $\psi_\alpha^{-1} \circ \psi_\beta$ are continuous.

33. For the vector bundle structure, see proposition 56.13.
34. Proposition 49.83.

At this point, we have proved that TM is a topological manifold. The definition 49.18 still requires $\psi_\alpha^{-1} \circ \psi_\beta$ to be smooth. The formula (49.188) shows that. □

PROPooMDDPooXeQwrH

Proposition 49.93.
The projection

$$\begin{aligned} \pi\colon TM &\to M \\ T_pM &\mapsto p \end{aligned} \tag{49.189}$$

is a submersion.

The structure is easy to guess. If $\varphi_\alpha\colon \mathcal{U}_\alpha \to M$ is a coordinate system on M (with $\mathcal{U}_\alpha \subset \mathbb{R}^n$), we define $\psi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{R}^n \to TM$ by

$$\psi(\underbrace{x_1, \dots x_n}_{\in \mathcal{U}_\alpha}, \underbrace{a_1, \dots a_n}_{\in \mathbb{R}^n}) = \sum_i a_i \left.\frac{\partial}{\partial x_i}\right|_{\varphi(x_1,\dots,x_n)}.$$

The map $\psi_\beta^{-1} \circ \psi_\beta$ is differentiable because

$$(\psi_\beta^{-1} \circ \psi_\beta)(x, a) = (y(x), \sum_i a_i \left.\frac{\partial y_j}{\partial x_i}\right|_{y(x)})$$

which is a composition of differentiable maps. The set TM endowed with this structure is called the **tangent bundle**.

49.6.1 Dual basis

DEFooZNHIooKIpND

Definition 49.94 (dual basis)**.**
The vectors[35] $\{\partial_{\alpha,i}\}_{i=1,\dots,n}$ *make a basis of* T_pM*. We denote by* $\{\partial^*_{\alpha,i}\}_{i=1,\dots,n}$ *the dual basis.*

LEMooSZTOooBIzMCc

Lemma 49.95 ([1])**.**
Let M *be a* C^k *manifold. We consider two charts* $\varphi_\alpha\colon U_\alpha \to M$ *and* $\varphi_\beta\colon U_\beta \to M$*. For* $x \in \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$*, we have*

EQooQNRRooKkCkLE

$$\partial^*_{\alpha,i}(\partial_{\beta,j})_x = (d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{\varphi_\beta^{-1}(x)})_{ij}, \tag{49.190}$$

and

EQooAGFKooWwWdAf

$$(\partial^*_{\alpha,1} \wedge \dots \wedge \partial^*_{\alpha,n})(\partial_{\beta,1}, \dots, \partial_{\beta,n}) = \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{\varphi_\beta^{-1}(x)}\right). \tag{49.191}$$

Proof. Let $x \in \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta) = M_{\alpha\beta}$, $a = \varphi_\alpha^{-1}(x)$ and $b = \varphi_\beta^{-1}(x)$. If t is small enough to keep $\varphi_\beta(b + te_j) \in M_{\alpha\beta}$ we have

SUBEQSooYFBPooPFeuLy

$$\begin{aligned} \partial_{\beta,j} &= \frac{d}{dt}\left[\varphi_\beta(b + te_j)\right]_{t=0} && \text{(49.192a)} \\ &= \frac{d}{dt}\left[\varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_\beta(b + te_j)\right]_{t=0} && \text{(49.192b)} \\ &= (d\varphi_\alpha)_a \frac{d}{dt}\left[\varphi_{\alpha\beta}(b + te_j)\right]_{t=0} && \text{(49.192c)} \\ &= (d\varphi_\alpha)_a\Big((d\varphi_{\alpha\beta})_b(e_j)\Big). && \text{(49.192d)} \end{aligned}$$

Dropping some index, we denote by $(d\varphi_{\alpha\beta})_{ij}$ the matrix elements of the linear map $(d\varphi_{\alpha\beta})_b\colon \mathbb{R}^n \to \mathbb{R}^n$, so that

$$(d\varphi_{\alpha\beta})e_j = \sum_k (d\varphi_{\alpha\beta})_{kj} e_k. \tag{49.193}$$

35. Definition 49.56.

Plugging in (49.192) we find

$$\partial_{\beta,j} = \sum_k (d\varphi_{\alpha\beta})_{kj} \underbrace{(d\varphi_\alpha)e_k}_{=\partial_{\alpha,k}} = \sum_k (d\varphi_{\alpha\beta})_{kj}\partial_{\alpha,k}. \tag{49.194}$$

We can finish the computation:

$$\partial^*_{\alpha,i}(\partial_{\beta,j}) = \sum_k (d\varphi_{\alpha\beta})_{kj} \underbrace{\partial^*_{\alpha,i}(\partial_{\alpha,k})}_{=\delta_{ik}} = (d\varphi_{\alpha\beta})_{ij}. \tag{49.195}$$

This proves equation (49.190). Now the formula (49.191) is given by proposition 10.214. □

49.6.2 Field of basis

PROPooMLSCooEGemdf

Proposition-Definition 49.96.
Let M be a C^k manifold, and $\varphi_\alpha : U_\alpha \to M$ be a chart. For $p \in M$ we define

$$(\partial_{\alpha,i})_p = \frac{d}{dt}\left[\varphi_\alpha(t_0 + te_i)\right]_{t=0} \tag{49.196}$$

where $t_0 = \varphi_\alpha^{-1}(p)$ and $\{e_i\}$ is the canonical basis of $\mathbb{R}^n$.

The vector field $p \mapsto (\partial_i)_{\alpha,p}$ is C^k.

Most the time we will simply we will write ∂_i, dropping the indices α and p.

PROPooGSTWooPvvwIF

Proposition-Definition 49.97.
Let M be a C^k manifold, and $\varphi_\alpha : U_\alpha \to M$ be a chart. For each $p \in M$ we consider the basis[36] $\{(\partial_i)_{\alpha,p}\}$ of T_pM, and the dual basis $\{(\alpha_i)_{\alpha,p}\}$ of T_p^*M.

$$\begin{aligned} \alpha_i : M &\to T^*M \\ p &\mapsto (\alpha_i)_{\alpha,p} \end{aligned} \tag{49.197}$$

are C^k.

LEMooVCSJooEuDZFz

Lemma 49.98.
Let M and N be smooth manifolds of dimension m and n with charts $\varphi : U \to M$ and $\psi : V \to N$ around $p \in M$ and $f(p) \in N$. We consider basis $\{e_i\}_{i=1,\dots,m}$ of $\mathbb{R}^m$ and $\{e'_\alpha\}_{\alpha=1,\dots,n}$ of $\mathbb{R}^n$.

The matrix of $df_p : T_pM \to T_{f(p)}N$ in the basis $\{d\varphi_{\varphi^{-1}(p)}(e_i)\}$ and $\{d\psi_{\psi^{-1}(f(p))}(e'_\alpha)\}$ is the same as the matrix of $d(\psi^{-1} \circ f \circ \varphi)_{\varphi^{-1}(p)}$ as map from $\mathbb{R}^m$ to $\mathbb{R}^n$.

Proof. Let subdivise.

(1) **Notations** As a preliminary remark, the fact that the proposed sets are basis is the proposition 49.68(1). For the notations, we write

$$\frac{\partial}{\partial x_i} = d\varphi_{\varphi^{-1}(p)}(e_i), \tag{49.198a}$$

$$\frac{\partial}{\partial y_\alpha} = d\psi_{\psi^{-1}\left(f(p)\right)}(e'_\alpha). \tag{49.198b}$$

(2) **Component** Let $v \in T_{f(p)}N$. We prove that

$$v_\alpha = \left((d\psi^{-1})_{f(p)}v\right)_\alpha \tag{49.199}$$

EQooISXNooJOzUmS

where in the left-hand side we are speaking of component with respect to the basis $\{\partial_{y_\alpha}\}$ while in the right-hand side, the ones with respect to the basis $\{e'_\alpha\}$.

36. See definition 49.96.

First we decompose v:

$$v = \sum_\alpha v_\alpha \frac{\partial}{\partial y_\alpha} = \sum_\alpha v_\alpha d\psi_{\psi^{-1}\left(f(p)\right)} e'_\alpha, \tag{49.200}$$

then we apply $d\psi^{-1}_{f(p)}$ to that equation:

$$d\psi^{-1}_{f(p)} v = \sum_\alpha v_\alpha e'_\alpha. \tag{49.201}$$

Taking the αþ component on both side we have our result (49.199).

(3) **Matrix** The matrix of a linear map is defined by the proposition 4.71. In our case,

$$(df_p)_{\alpha i} = \left(df_p\big(\frac{\partial}{\partial x_i}\big)\right)_\alpha = \left(df_p \circ d\varphi_{\varphi^{-1}(p)} e_i \right)_\alpha. \tag{49.202}$$

Using the formula (49.199),

$$\begin{aligned}
(df_p)_{\alpha i} &= \left(df_p \circ d\varphi_{\varphi^{-1}(p)} e_i \right)_\alpha && \text{(49.203a)}\\
&= \left((d\psi^{-1})_{f(p)} \circ df_p \circ d\varphi_{\varphi^{-1}(p)} e_i \right)_\alpha && \text{(49.203b)}\\
&= \left((d\psi^{-1})_{f(p)} \circ df_p \circ d\varphi_{\varphi^{-1}(p)} \right)_{\alpha i} && \text{(49.203c)}\\
& && \text{(49.203d)}
\end{aligned}$$

□

LEMooPSWEooVKLWMQ

Lemma 49.99.
If X is a c^k vector field on the C^k manifold M and if f is a c^k function, then the formula

$$(Xf)(x) = X_x(f) \tag{49.204}$$

defines a C^{k-1} function Xf on M.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooVZNUooVoyFnx

Proposition 49.100.
Let $\{e_i\}$ be a C^k field of basis on an open set around $x \in M$. Let X be a X^k vector field on M. For each i, the map

$$\begin{aligned} g\colon M &\to \mathbb{R} \\ x &\mapsto X(x)_i \end{aligned} \tag{49.205}$$

is C^k.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooLAQTooRHYmqy

Proposition 49.101.
Let M be a C^k manifold. Let $\{e_i\}$ be a C^k field of basis of TM. The projection map

$$\begin{aligned} \mathit{proj}_i\colon TM &\to \mathbb{R} \\ v &\mapsto v_i \end{aligned} \tag{49.206}$$

is C^k.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooYPLLooZaQEMx

Proposition 49.102.
Let $f\colon M \to N$ be a C^k map. The map

$$\begin{aligned} \det(ds)\colon M &\to \mathbb{R} \\ x &\mapsto \det(ds_x) \end{aligned} \tag{49.207}$$

is C^{k-1}

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooNTSLooCXQQXy

Proposition 49.103 ([1])**.**
Let $f\colon M \to N$ be a C^k map between C^k manifolds. If X is a C^k vector field on M, then the map

$$\begin{aligned} f\colon M &\to TN \\ x &\mapsto (df)_x X_x \end{aligned} \tag{49.208}$$

is C^{k-1}.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

LEMooPWMUooRalWxC

Lemma 49.104.
Let M be a C^k manifold. Let $\varphi\colon U \to M$ be a chart around $a \in M$. Let $\psi_a\colon \mathbb{R}^n \to T_aM$ be the map defined in proposition 49.84.

A map $X\colon M \to TM$ is a C^k vector field on $\varphi(U)$ if and only if there exist a C^k function $v\colon M \to \mathbb{R}^n$ such that

$$X = \psi \circ v \tag{49.209}$$

where it is understood that $(\psi \circ v)(x) = \psi_x\big(v(x)\big)$.

49.6.3 Commutator of vector fields

DEFooHOTOooRaPwyo

Proposition-Definition 49.105.
Let M be a C^k manifold with $k \geqslant 2$. Let X, $Y \in \mathfrak{X}(M)$.

ITEMooZKKUooQjYftU

(1) For every $x \in M$, the operator $[X,Y]_x$ defined by

$$[X,Y]_x f = X_x(Yf) - Y_x(Xf) \tag{49.210}$$

EQooDSKWooXdjPPP

is an element of T_xM.
Here, Yf and Xf are defined by virtue of lemma 49.99.

ITEMooPGPLooQrKxWY

(2) The map $x \mapsto [X,Y]_x$ is a vector field of class C^{k-1}.

*The so-defined vector field $[X,Y]$ is the **commutator** of X and Y.*

Proof. Point by point.

(1) **(1)** From lemma 49.86, we have C^k functions $v_k\colon M \to \mathbb{R}$ such that

$$X_x(f) = \sum_k v_k(x)\partial_k(f \circ \varphi)\big(\varphi^{-1}(x)\big), \tag{49.211}$$

and the same for Y :

$$Y_y(f) = \sum_l w_l(y)\partial_l(f \circ \varphi)\big(\varphi^{-1}(y)\big). \tag{49.212}$$

We need to compute $X_x(Yf)$, that is

$$X_x(Yf) = \sum_k v_k(x)\partial_k(Yf \circ \varphi)\big(\varphi^{-1}(x)\big). \tag{49.213}$$

EQooAHMYooWRttQr

The easy part of that is

$$(Yf\circ\varphi)(s)=(Yf)\big(\varphi(s)\big)=\sum_l(w_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s). \tag{49.214}$$

This is a product of two functions $w_l\circ\varphi\colon U\to\mathbb{R}$ and $f\circ\varphi\colon U\to\mathbb{R}$. For computing the partial derivative, we use the usual Leibniz rule :

$$\partial_k(Yf\circ\varphi)(s)=\sum_l\partial_k(w_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s)+\sum_l(w_l\circ\varphi)(s)\partial_k\partial_l(f\circ\varphi)(s). \tag{49.215}$$

We can put that in (49.213), with the definition $s=\varphi^{-1}(x)\in U\subset\mathbb{R}^n$:

$$\begin{aligned}X_x(Yf)&=\sum_k v_k(x)\sum_l\partial_k(w_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s) &\text{(49.216a)}\\&\quad+\sum_{kl}v_k(x)(w_l\circ\varphi)(s)\partial_k\partial_l(f\circ\varphi)(s) &\text{(49.216b)}\\&=\sum_{kl}v_k(x)\partial_k(w_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s) &\text{(49.216c)}\\&\quad+\sum_{kl}v_k(x)w_l(x)\partial_k\partial_l(f\circ\varphi)(s). &\text{(49.216d)}\end{aligned}$$

The expression of $Y_x(Xf)$ is the same, permuting v and w. The commutator has 4 terms :

$$\begin{aligned}[X,Y]_xf&=\sum_{kl}v_k(x)\partial_k(w_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s)\\&\quad+\sum_{kl}v_k(x)w_l(x)\partial_k\partial_l(f\circ\varphi)(s)\\&\quad-\sum_{kl}w_k(x)\partial_k(v_l\circ\varphi)(s)\partial_l(f\circ\varphi)(s)\\&\quad-\sum_{kl}w_k(x)v_l(x)\partial_k\partial_l(f\circ\varphi)(s).\end{aligned} \tag{49.217}$$

By virtue of theorem 12.343, the two terms with second derivatives cancel out because the maps are of class C^k with $k\geqslant 2$. Only two terms remain :

$$[X,Y]_xf=\sum_l\Big[\sum_k v_k(x)\partial_k(w_l\circ\varphi)(s)-w_k(x)\partial_k(v_l\circ\varphi)(s)\Big]. \tag{49.218}$$

We pose

$$\begin{aligned}u_l\colon M&\to\mathbb{R}\\x&\mapsto\sum_k\big[v_k(x)\partial_k(w_l\circ\varphi)\big(\varphi^{-1}(x)\big)-w_k(x)\partial_k(v_l\circ\varphi)\big(\varphi^{-1}(x)\big)\big].\end{aligned} \tag{49.219}$$

For each x we have

$$[X,Y]_x(f)=\sum_l u_l(x)\partial_l(f\circ\varphi)(s). \tag{49.220}$$

EQooVUOKooAyGeae

By lemma 49.63, this means that $[X,Y]_x\in T_xM$.

(2) **(2)**

Now the functions u_l are of of class C^{k-1} (theorem 12.330) which satisfies, so that the lemma 49.99 makes $[X,Y]$ a vector field of class C^{k-1}.

□

Lemma 49.106.
Let M be a smooth manifold[37]*. Let X, Y be smooth vector fields given by $X = \psi \circ v$ and $Y = \psi \circ w$. Then we have*

$$[\psi \circ v, \psi \circ w] = \psi \circ u \tag{49.221}$$

with

$$\begin{aligned} u_l \colon \varphi(U) &\to \mathbb{R} \\ x &\mapsto X_x(w_l) - Y_x(v_l). \end{aligned} \tag{49.222}$$

This equation can be shorthanded into

$$u = X(w) - Y(v). \tag{49.223}$$

Proof. We start from the expression 49.220 in which we substitute the values of (49.86) :

$$u_l(x) = \sum_k \underbrace{v_k(x)\partial_k(w_l \circ \varphi)(s)}_{=X_x(w_l)} - w_k(x)\partial_k(v_l \circ \varphi)(s) = X_x(w_l) - Y_x(v_l). \tag{49.224}$$

□

49.7 Submanifold

DEFooLQHWooMOTgzq

Definition 49.107.
*If M is a differentiable manifold and S, a subset of M, we say that S is a **submanifold** of dimension k if $\forall\, p \in S$, there exists a chart $\varphi \colon \mathcal{U} \to M$ around p such that*

$$\varphi^{-1}(\varphi(\mathcal{U}) \cap S) = \mathbb{R}^k \cap \mathcal{U} := \{(x_1, \dots, x_k, 0 \dots, 0) \in \mathcal{U}\}. \tag{49.225}$$ EQooLWQRooQJCQbA

Let $k \leqslant n$ in $\mathbb{N}$. We consider the maps

$$\begin{aligned} \mathtt{proj}_k \colon \mathbb{R}^n &\to \mathbb{R}^k \\ x &\mapsto (x_1, \dots, x_k). \end{aligned} \tag{49.226}$$

and

$$\begin{aligned} j \colon \mathbb{R}^k &\to \mathbb{R}^n \\ x &\mapsto (x_1, \dots, x_k, 0, \dots 0). \end{aligned} \tag{49.227}$$

Proposition 49.108 ([836])**.**
Let S be a smooth submanifold of the smooth manifold M. Let $p \in S$. We consider a map $\varphi \colon U \to M$ around p such that[38]

$$\varphi^{-1}\big(\varphi(U) \cap S\big) = U \cap j(\mathbb{R}^k). \tag{49.228}$$

Let $X = \varphi(U) \cap S$ and $Y = (\mathtt{proj}_k \circ \varphi^{-1})(X)$. Then we define

$$\begin{aligned} \psi \colon Y &\to X \\ x &\mapsto (\varphi \circ j)(x). \end{aligned} \tag{49.229}$$

We have:

(1) ψ is bijective,

(2) for every chart $\varphi_\alpha \colon U_\alpha \to M$ of M, the maps

$$\psi^{-1} \circ \varphi_\alpha \colon \varphi_\alpha^{-1}\big(\psi(Y)\big) \to Y \tag{49.230}$$ EQooBAGFooDnnctJ

and

$$\varphi_\alpha^{-1} \circ \psi \colon \psi^{-1}\big(\varphi_\alpha(U_\alpha)\big) \to U_\alpha \tag{49.231}$$ EQooKQIUooDCCczD

are smooth.

37. When one deal with commutators, the natural setting is (at least) C^∞ manifolds since the bracket diminishes by 1 the regularity of the vector fields.

38. we write the condition 49.225 in a more condensed way.

Proof. In several parts.

(1) ψ **is injective** Let $a, b \in Y \subset \mathbb{R}^k$ such that $\psi(a) = \psi(b)$, that is $(\varphi \circ j)(a) = (\varphi \circ j)(b)$. Since φ and j are injective, we have $a = b$.

(2) $j \circ \mathtt{proj}_k = \mathrm{Id}\,|_{j(\mathbb{R}^k)}$ If $a \in j(\mathbb{R}^k)$, we have $a = (a_1, \dots, a_k, 0, \dots, 0)$, so that

$$j \circ \mathtt{proj}_k(a) = j(a_1, \dots, a_k) = (a_1, \dots, a_k, 0, \dots, 0) = a. \tag{49.232}$$

(3) ψ **is surjective** Let $x \in X$. By definition of Y, we have $y = \mathtt{proj}_k\big(\varphi^{-1}(x)\big) \in Y$. We prove that $\psi(y) = x$. We have

$$\psi(y) = (\psi \circ \mathtt{proj}_k \circ \varphi^{-1})(x) = (\varphi \circ j \circ \mathtt{proj}_k \circ \varphi^{-1})(x). \tag{49.233}$$

EQooFFSTooCddIyX

Since $x \in X = \varphi(U) \cap S$ we have $\varphi^{-1}(x) \subset U \cap j(\mathbb{R}^k)$. We know that $j \circ \mathtt{proj}_k = \mathrm{Id}\,|_{j(\mathbb{R}^k)}$, so that

$$(j \circ \mathtt{proj}_k)\big(\varphi^{-1}(x)\big) = \varphi^{-1}(x). \tag{49.234}$$

We continue (49.233):

$$\psi(y) = (\varphi \circ j \circ \mathtt{proj}_k \circ \varphi^{-1})(x) = \varphi\big(\varphi^{-1}(x)\big) = x. \tag{49.235}$$

So ψ is surjective. And we proved that ψ is a bijection.

By the way, we have a formula for the inverse of ψ :

$$\psi^{-1} = \mathtt{proj}_k \circ \varphi^{-1} \colon X \to Y. \tag{49.236}$$

EQooQSZUooWsAqCz

(4) **First smoothness** We prove that (49.230) is smooth. First notice that

$$\varphi_\alpha\Big(\varphi_\alpha^{-1}\big(\psi(Y)\big)\Big) \subset \psi(Y) = X = \varphi(U) \cap S \subset \varphi(U). \tag{49.237}$$

Thus it makes sense to write

$$\psi^{-1} \circ \varphi_\alpha = \psi^{-1} \circ \varphi \circ \varphi^{-1} \circ \varphi_\alpha \tag{49.238}$$

EQooOTSUooXHGAZg

as maps defined on $\varphi_\alpha^{-1}\big(\psi(Y)\big)$. Since φ and φ_α are charts, the map $\varphi^{-1} \circ \varphi_\alpha$ is smooth. We know the inverse of ψ from equation (49.236). We have

$$\psi^{-1} \circ \varphi = \mathtt{proj}_k \circ \varphi^{-1} \circ \varphi = \mathtt{proj}_k, \tag{49.239}$$

which is smooth.

Equation (49.238) is now the composition of two smooth functions.

(5) **Second smoothness** We prove that (49.231) is smooth. We have

$$\varphi_\alpha^{-1} \circ \psi = \varphi_\alpha^{-1} \circ \varphi \circ j \tag{49.240}$$

while φ_α^{-1}, φ and j are smooth.

□

The following proposition shows that a submanifold is a manifold for itself.

PROPooRZIHooXIhnpq

Proposition 49.109 ([1])**.**
Let S be a smooth submanifold of the smooth manifold M. For each $p \in S$, we consider the set $\{\varphi_{p,i}\}_{i \in I}$ of the charts $\varphi_{p,i} \colon U_{p,i} \to M$ around p such that

$$\varphi_{p,i}^{-1}\big(\varphi(U_{p,i}) \cap S\big) = U_{p,i} \cap j(\mathbb{R}^k). \tag{49.241}$$

Let $X_{p,i} = \varphi_{p,i}(U_{p,i}) \cap S$ and $Y_{p,i} = (\mathtt{proj}_k \circ \varphi_{p,i}^{-1})(X_{p,i})$. Then we define

$$\begin{aligned} \psi_{p,i} \colon Y_{p,i} &\to X_{p,i} \\ x &\mapsto (\varphi_{p,i} \circ j)(x). \end{aligned} \tag{49.242}$$

The couple $\big(S, \{(U_{p,i}, \varphi_{p,i})\}_{i \in I}\big)$ is a manifold[39]*.*

When S is a submanifold, we will always consider this manifold structure on S.

39. Well. The index set I may depend on p, but the notations are already complicated enough.

DEFooZKUIooXWVGvh

Definition 49.110.
Let a map $f\colon M_1 \to M_2$.

*(1) It is an **immersion** at $p \in M_1$ if $df_p\colon T_pM_1 \to T_{f(p)}M_2$ is injective*[40].

*(2) It is a **submersion** if df_p is surjective.*

PROPooEWUCooTStAvb

Proposition 49.111 ([836]).
If S is a submanifold of M, the inclusion map $\iota\colon S \to M$ is an immersion[41].

PROPooZACHooCNgLSl

Proposition 49.112.
A manifold S is a submanifold of M if $S \subset M$ (as sets) and the identity $\iota\colon S \to M$ is regular[42].

prop:topo_sub_manif

Proposition 49.113.
The own topology of a submanifold is finer than the induced one from the manifold.

Proof. Let M be a manifold of dimension n and N a submanifold[43] of dimension $k < n$. We consider V, an open subset of N for the induced topology, so $V = N \cap \mathcal{O}$ for a certain open subset $\mathcal{O}$ of M. The aim is to show that V is an open subset in the topology of N.

Let us define $\mathcal{P} = \varphi^{-1}(\mathcal{O})$. The charts of N are the projection to $\mathbb{R}^k$ of the ones of M. We have to consider $W = \varphi^{-1}(V)$, since N is a submanifold, $\varphi^{-1}(\mathcal{O} \cap N) = \mathbb{R}^k \cap \mathcal{P}$. It is clear that $W = \mathbb{R}^k \cap \mathcal{P}$ is an open subset of $\mathbb{R}^k$ because it is the projection on the k first coordinates of an open subset of $\mathbb{R}^n$.

The subset V of N will be open in the sense of the own topology of N if ${\varphi'}^{-1}(V \cap \varphi'(\mathcal{U}'))$ is open in $\mathbb{R}^k$ where φ' is the restriction of φ to his k first coordinates: $\varphi'(a) = \varphi(a, 0)$ and $\mathcal{U}'$ is the projection of $\mathcal{U}$. □

prop:subvar_ouvert

Proposition 49.114.
A submanifold is open if and only if it has the same dimension as the main manifold.

Proof. Necessary condition. We consider some charts $\varphi_i\colon U_i \to M$ on some open subsets U_i of $\mathbb{R}^n$. If N is open in M, then this can be written as

$$N = \bigcup_i U_i.$$

If we choose the charts on M in such a manner that $\varphi_i\colon U_i \cap \mathbb{R}^k \to N$ are charts of N, we must have $\varphi_i(U_i \cap \mathbb{R}^k) = \varphi_i(U_i)$. Then it is clear that $k = n$ is necessary.
Sufficient condition. If N has same dimension as M, the charts $\varphi_i\colon U_i \to M$ are trivially restricted to N. □

The following result allow to extend a smooth function defined on a submanifold to an open set of the «larger» manifold.

PROPooOTZQooIfboXV

Proposition 49.115.
Let N be a submanifold of M and $f \in C^\infty(N)$. Let $p \in N$. There exists a neighbourhood W of p in M and a function $\tilde{f} \in C^\infty(W)$ such that

$$\tilde{f}(n) = f(n) \tag{49.243}$$

for every $n \in N$.

40. Differential of map, definition 49.59.
41. Definition 49.110
42. Definition 49.34.
43. In the whole proof, we should say "there exists a sub-neighbourhood such that..."

Proof. Since N is a submanifold of M, the definition 49.107 provides a chart $\varphi\colon U \to M$ around p such that

$$\varphi^{-1}\big(\varphi(U) \cap N\big) = \{(x_1, \dots, x_n, 0, \dots, 0)\}. \tag{49.244}$$

From the function $f\colon N \to \mathbb{R}$ we consider

$$\begin{aligned} f_1\colon \varphi^{-1}\big(\varphi(U) \cap N\big) &\to \mathbb{R} \\ f_1 &= f \circ \varphi \end{aligned} \tag{49.245}$$

This is the function f seen trough the chart. The function f_1 is only defined on the "N" part of the chart, but can be extended as

$$\begin{aligned} \tilde{f}_1\colon U &\to \mathbb{R} \\ (x_1, \dots, x_m) &\mapsto f_1(x_1, \dots, x_n, 0, \dots, 0), \end{aligned} \tag{49.246}$$

which is a good definition since $(x_1, \dots, x_n, 0, \dots, 0)$ is in $\varphi^{-1}\big(\varphi(U) \cap N\big)$.

Finally we write

$$\begin{aligned} \tilde{f}\colon \varphi(U) &\to \mathbb{R} \\ \tilde{f} &= \tilde{f}_1 \circ \varphi^{-1}. \end{aligned} \tag{49.247}$$

This is the extension we are searching for. Indeed it is defined on $\varphi(U)$ which is an open set in M which contains p and if $q \in N \cap \varphi(U)$ we have $q = \varphi(x_1, \dots, x_n, 0, \dots, 0)$ and then

$$\begin{aligned} \tilde{f}(q) &= (\tilde{f}_\circ\varphi^{-1})\varphi(x_1, \dots, x_n, 0, \dots, 0) && \text{(49.248a)} \\ &= \tilde{f}_1(x_1, \dots, x_n, 0, \dots, 0) && \text{(49.248b)} \\ &= f_1(x_1, \dots, x_n, 0, \dots, 0) && \text{(49.248c)} \\ &= (f \circ \varphi)(x_1, \dots, x_n, 0, \dots, 0) && \text{(49.248d)} \\ &= f(q). && \text{(49.248e)} \end{aligned}$$

Thus $\tilde{f} = f$ on $\varphi(U) \cap N$. □

49.8 Flow and integral curve

PROPooJACTooXBSxfE

Proposition-Definition 49.116 (Integral curve, flow of a vector field[1, 837])**.**
Let X be a C^k vector field on M and let $p \in M$. There exists a unique maximal interval $I \subset \mathbb{R}$ and C^k map[44] *$\Phi_p\colon I \to M$ such that*

(1) $\Phi_p(0) = p$,

(2) $\frac{d}{dt}\Big[\Phi_p(t)\Big]_{t=t_0} = X_{\Phi_p(t_0)}$.

The map Φ satisfy

$$\Phi_{t+u} = \Phi_t \circ \Phi_u. \tag{49.249}$$

EQooCHYXooCbECRN

*We say that Φ_p is the **integral curve** of X at p, and that Φ is the **flow** of X. We will often write $\Phi(p,t)$ instead of $\Phi_p(t)$. When we are interested in only one point p, we can write γ_X or simply γ.*

When we want to specify the vector field, we write $\Phi_g^X(t)$.

Proof. In several parts.

(1) **Local unicity** ITEMooZPSEooPivDiw

Let $p \in M$, we show that there exists an interval $I \subset \mathbb{R}$ such that a C^k integral curve $\gamma\colon \mathbb{R} \to M$ satisfying $\gamma(0) = p$ is unique. Here the existence is not yet discussed.

Let J be an interval around 0 in $\mathbb{R}$ and $\gamma\colon J \to M$, an integral curve of X such that $\gamma(0) = p$.

44. The map Φ_p is C^k. We are not speaking of the regularity with respect to p. For more regularity, see the proposition 49.117.

Let $\varphi: U \to M$ be a local chart around p. We know from lemma 49.86 that there exist functions $v_k \in C^k(M, \mathbb{R})$ such that

$$X_x(f) = \sum_{k=1}^{n} v_k(x)\partial_k(f \circ \varphi)\big(\varphi^{-1}(x)\big). \tag{49.250}$$

We consider the map $\tilde{\gamma}: I \to U$ defined by

$$\tilde{\gamma} = \varphi^{-1} \circ \gamma, \tag{49.251}$$

and, if f is a function on M we define $\tilde{f}: U \to \mathbb{R}$ by $\tilde{f} = f \circ \varphi$.
On the one hand,

$$\gamma'(a)f = \frac{d}{dt}\Big[(f \circ \varphi \circ \varphi^{-1} \circ \gamma)(t)\Big]_{t=a} = \frac{d}{dt}\Big[(\tilde{f} \circ \tilde{\gamma})\Big]_{t=a}. \tag{49.252}$$

On the other hand,

$$X_{\gamma(a)}(f) = \sum_k (\partial_k \tilde{f})\big(\tilde{\gamma}(a)\big)\frac{d}{dt}\Big[\tilde{\gamma}_k(t)\Big]_{t=a} = \sum_k (\partial_k \tilde{f})\big(\varphi^{-1}(p)\big)\tilde{\gamma}_k'(a) = \sum_k (\partial_k \tilde{f})\big(a\big)\tilde{\gamma}_k'(a) \tag{49.253}$$

where $\tilde{\gamma}_k'$ is an usual derivative.
So, in order to be an integral curve, we need the equality SUBEQSooXSIYooPlVcEI

$$\begin{cases} v_k\big(\varphi(a)\big) = \tilde{\gamma}_k'(a) & \text{(49.254a)} \\ \tilde{\gamma}(0) = \varphi^{-1}(p). & \text{(49.254b)} \end{cases}$$

for every $a \in I$. Each map $\tilde{\gamma}_k$ satisfy the equations SUBEQSooBLWCooATQGEn

$$\begin{cases} \tilde{\gamma}_k' = v_k \circ \varphi & \text{(49.255a)} \\ \tilde{\gamma}_k(0) = \varphi^{-1}(p)_k. & \text{(49.255b)} \end{cases}$$

The function $v_k \circ \varphi: U \to \mathbb{R}$ is a C^k function. We know from the Cauchy-Lipschitz theorem 17.39 that the system (49.255) has a unique solution on an interval I around 0 in $\mathbb{R}$.

(2) **Local existence**
Consider a solution of (49.255).

(3) **Maximum interval**

□

PROPooQVQAooJVdwOa

Proposition 49.117 ([1]).
Let X be a C^k vector field on M.
Let $\Phi: M \times \mathbb{R} \to M$ be a map satisfying

(1) $\Phi_p(0) = p$,

(2) $\frac{d}{dt}\Big[\Phi_p(t)\Big]_{t=t_0} = X_{\Phi(t_0)}$.

Then Φ is C^k.

DEFooMZSBooFkfKkS

Definition 49.118.
*A vector field X is **complete** if for every $p \in M$, there exists an integral curve $\gamma: \mathbb{R} \to M$ such that $\gamma(0) = p$.*

The point of a complete vector field is that the integral curves are defined on $\mathbb{R}$, not just a neighbourhood of 0.

The integral curve given by point (1) is only valid on a neighbourhood given by a local chart.

PROPooQNWOooBOxYtu

Proposition 49.119 ([1]).
Let X be a complete C^k vector field on M. For every $p \in M$, there is an unique integral curve $\gamma_p \colon \mathbb{R} \to M$ such that $\gamma_p(0) = p$.

Proof. The existence is not an issue because it is part of the definition of a complete vector field. The trick is the unicity. We already know from point (1) that we have the unicity on each chart. The difficulty is to glue them together.

Let $\gamma \colon \mathbb{R} \to M$ and $\sigma \colon \mathbb{R} \to M$ be integral curves satisfying $\gamma(0) = \sigma(0) = p$.

Let K be a compact interval containing 0 in $\mathbb{R}$. For each $a \in K$, the point (1) provides us an open interval I_a containing a and such that there exists a unique $\alpha \colon I_a \to M$ satisfying EQSooCEPBooJCuFJL

$$\begin{cases} \alpha \text{ is an integral curve of} X & (49.256a) \\ \alpha(a) = \gamma(a). & (49.256b) \end{cases}$$

Since γ satisfy these requirements, we know that every map satisfying the conditions (49.256) is equal to γ on I_a.

The intervals I_a make a cover of K which is compact; we extract a finite subcover $\{I_i\}_{i=1,\dots,n}$. Lemma 9.54 allows us to sort these intervals in such a way that $0 \in I_1$ and

$$I_m \cap \bigcup_{i=1}^{m-1} I_i \neq \varnothing \tag{49.257}$$

for every $m = 1, \dots, n$.

From unicity on I_1 we know that $\gamma = \sigma$ on I_1 because $\gamma(0) = \sigma(0)$ and $0 \in I_1$. We make a induction. We suppose that, for some m we have $\gamma = \sigma$ on $\bigcup_{i=1}^{m} I_i$. Let

$$t_0 \in I_{m+1} \cap \bigcup_{i=1}^{m} I_i. \tag{49.258}$$

From the induction hypothesis, we have $\gamma(t_0) = \sigma(t_0)$ because $t_0 \in \bigcup_{i=1}^{m} I_i$.

The curve σ is, on I_{m+1} an integral curve of X satisfying $\sigma(t_0) = \gamma(t_0)$ with $t_0 \in I_{m+1}$. Thus $\sigma = \gamma$ on I_{m+1}. We deduce that $\sigma = \gamma$ on $\bigcup_{i=1}^{n} I_i$ and in particular $\sigma = \gamma$ on K.

Since the whole is valid for every compact $K \subset \mathbb{R}$ we conclude that $\sigma = \gamma$ on $\mathbb{R}$. □

Let X be a vector field on M and $p \in M$. We denote by $\Phi(p,t) = \Phi_t(p) \in M$ the point $\gamma_p(t)$ where γ_p is the integral curve of proposition 49.119. In other terms, for every $p \in M$, there exist an interval I_p such that

$$\frac{d}{dt}\Big[\Phi(p,t)\Big]_{t=t_0} = X_{\Phi(p,t_0)}. \tag{49.259}$$

and for every $p \in M$, we have $\Phi(p,0) = p$.

PROPooDPXIooTvXOIP

Proposition 49.120 ([837]).
Let X, Y be smooth vectors fields on the manifold M. Let Φ^X and Φ^Y be their flows[45]*. The following are equivalent:*

(1) $[X,Y] = 0$,

(2) $(d\Phi_t^X)Y = Y$

(3) For every t, u in the domains, $\Phi_t^X \circ \Phi_u^Y = \Phi_u^Y \circ \Phi_t^X$.

45. Proposition 49.116.

49.8.1 Next stuff

The manifold could, of course have some additional structure which allows to write the differential quotient (49.79). This is the case when $M = \mathbb{R}^n$ or when M is a matrix group. In these cases, the question of the link between $\gamma'(0)$ and the "true" derivative of γ has to be studied.

In that case we have the same notational problem with "df". Let $f\colon M \to \mathbb{R}$ where M is a manifold like $\mathbb{R}^n$. The symbol df_a with $a \in M$ can be the differential of f as function $M \to \mathbb{R}$, so that df_a is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$. But df_a can also be the linear map $df_a\colon T_aM \to T_{f(a)}\mathbb{R}$ where the spaces T_aM and $T_{f(a)}\mathbb{R}$ are made of differential operators.

This is the point of the section 53.2.

Using the chain rule $d(g\circ f)(a) = dg(f(a))\circ df(a)$ for the differentiation in $\mathbb{R}^n$, one sees that this equivalence notion doesn't depend on the choice of φ. In other words, if φ and $\tilde{\varphi}$ are two charts for a neighbourhood of x, then $(\varphi^{-1}\circ\gamma)'(0) = (\varphi^{-1}\circ\sigma)'(0)$ if and only if $(\tilde{\varphi}^{-1}\circ\gamma)'(0) = (\tilde{\varphi}^{-1}\circ\sigma)'(0)$. The space of all tangent vectors at x is denoted by T_xM. There exists a bijection $[\gamma] \leftrightarrow (\varphi^{-1}\circ\gamma)'(0)$ between T_xM and $\mathbb{R}^n$, so T_xM is endowed with a vector space structure.

If $(\mathcal{U}, \varphi)$ is a chart around $X(0)$, we can express Xf using only well know objects by defining the function $\tilde{f} = f \circ \varphi$ and $\tilde{X} = \varphi^{-1} \circ X$

$$Xf = \frac{d}{dt}\Big[(\tilde{f}\circ\tilde{X})(t)\Big]_{t=0} = \frac{\partial\tilde{f}}{\partial x^\alpha}\Big|_{x=\tilde{X}(0)} \frac{d\tilde{X}^\alpha}{dt}\Big|_{t=0}.$$

In this sense, we write

$$X = \frac{d\tilde{X}^\alpha}{dt}\frac{\partial}{\partial x^\alpha} \tag{49.260}$$

and we say that $\{\partial_1, \dots, \partial_n\}$ is a basis of T_xM. As far as notations are concerned, from now a tangent vector is written as $X = X^\alpha\partial_\alpha$ where X^α is related to the path $X\colon \mathbb{R} \to M$ by $X^\alpha = d\tilde{X}^\alpha/dt$. We will no more mention the chart φ and write

$$Xf = \frac{d}{dt}\Big[f(X(t))\Big]_{t=0}.$$

Correctness of this short notation is because the equivalence relation is independent of the choice of chart. When we speak about a tangent vector to a given path $X(t)$ without specification, we think about $X'(0)$.

All this construction gives back the notion of tangent vector when $M \subset \mathbb{R}^m$. In order to see it, think to a surface in $\mathbb{R}^3$. A tangent vector is precisely given by a derivative of a path: if $c\colon \mathbb{R} \to \mathbb{R}^n$ is a path in the surface, a tangent vector to this curve is given by

$$\lim_{t\to 0}\frac{c(t_0) - c(t_0 + t)}{t}$$

which is a well know limit of a difference in $\mathbb{R}^n$. pg:vecto_vecto

Let us precise how does a tangent vector acts on maps others than $\mathbb{R}$-valued functions. If V is a vector space and $f\colon M \to V$, we define

$$Xf = (Xf^i)e_i$$

where $\{e_i\}$ is a basis of V and the functions $f^i\colon M \to \mathbb{R}$, the decomposition of f with respect to this basis. If we consider a map $\varphi\colon M \to N$ between two manifolds, the natural definition is $Xf := dfX$. More precisely, if we consider local coordinates x^α and a function $f\colon M \to \mathbb{R}$,

$$(d\varphi X)f = \frac{d}{dt}\Big[(f\circ\varphi\circ X)(t)\Big]_{t=0} = \frac{\partial f}{\partial x^\alpha}\frac{\partial\varphi^\alpha}{\partial x^\beta}\frac{dX^\beta}{dt}. \tag{49.261}$$ eq:dvp_phi

Now we are in a notational trouble: when we write $X = X^\alpha\partial_\alpha$, the "$X^\alpha$" is the derivative of the "X^α" which appears in the path $X(t) = (X^1(t), \dots, X^n(t))$ which gives X by $X = X'(0)$. So equation (49.261) gives

$$X(\varphi) := d\varphi X = X^\beta(\partial_\beta\varphi^\alpha)\partial_\alpha. \tag{49.262}$$

49.8.2 Differential of a map

Let $f\colon M_1 \to M_2$ be a differentiable map, $x \in M_1$ and $X \in T_xM_1$, i.e. $X\colon \mathbb{R} \to M_1$ with $X(0) = x$ and $X'(0) = X$. We can consider the path $Y = f \circ X$ in M_2. The tangent vector to this path is written df_xX.

Proposition 49.121.
If $f\colon M_1 \to M_2$ is a differentiable map between two differentiable manifolds, the map

$$\begin{aligned} df_x\colon T_xM_1 &\to T_{f(x)}M_2 \\ X'(0) &\mapsto (f \circ X)'(0) \end{aligned} \tag{49.263}$$

is linear.

Proof. We consider local coordinates $x\colon \mathbb{R}^n \to M_1$ and $y\colon \mathbb{R}^m \to M_2$. The maps $f\colon M_1 \to M_2$ and $y^{-1} \circ f \circ x\colon \mathbb{R}^n \to \mathbb{R}^m$ will sometimes be denoted by the same symbol f. We have $(x^{-1} \circ X)(t) = (x_1(t), \dots, x_n(t))$ and $(y^{-1} \circ Y)(t) = \big(y_1(x_1(t), \dots, x_n(t), \dots, y_m(x_1(t), \dots, x_n(t)\big)$, so that

$$Y'(0) = \left(\sum_{i=1}^{n} \frac{\partial y_1}{\partial x_i} x_i'(0), \dots, \sum_{i=1}^{n} \frac{\partial y_m}{\partial x_i} x_i'(0) \right) \in \mathbb{R}^m$$

which can be written in a more matricial way under the form

$$Y'(0) = \left(\frac{\partial y_i}{\partial x_j} x_j'(0) \right).$$

So in the parametrisations x and y, the map df_x is given by the matrix $\partial y^i/\partial x_j$ which is well defined from the only given of f. □

Let $x\colon \mathcal{U} \to M$ and $y\colon \mathcal{V} \to M$ be two charts systems around $p \in M$. Consider the path $c(t) = x(0, \dots, t, \dots 0)$ where the t is at the position k. Then, with respect to these coordinates,

$$c'(0)f = \frac{d}{dt}\Big[f(c(t)) \Big]_{t=0} = \frac{\partial f}{\partial x^i} \frac{dc^i}{dt} = \frac{\partial f}{\partial x^k},$$

so $c'(0) = \partial/\partial x^k$. Here, implicitly, we wrote $c^i = (x^i)^{-1} \circ c$ where $(x^i)^{-1}$ is the ith component of x^{-1} seen as element of $\mathbb{R}^n$. We can make the same computation with the system y. With these abuse of notation,

$$\frac{\partial}{\partial x^i} = \sum_j \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \tag{49.264}$$

as it can be seen by applying it on any function $f\colon M \to \mathbb{R}$. More precisely if $x\colon \mathcal{U} \to M$ and $y\colon \mathcal{U} \to M$ are two charts (let $\mathcal{U}$ be the intersection of the domains of x and y), let $f\colon M \to \mathbb{R}$ and $\overline{f} = f \circ x$, $\tilde{f} = f \circ y$. The action of the vector ∂_{x^i} of the function f is given by

$$\partial_{x^i} f = \frac{\partial \overline{f}}{\partial x^i}$$

where the right hand side is a real number that can be computed with usual analysis on $\mathbb{R}^n$. This real *defines* the left hand side. Now, $\overline{f} = \tilde{f} \circ y^{-1} \circ x$, so that

$$\frac{\partial \overline{f}}{\partial x^i} = \frac{\partial(\tilde{f} \circ y^{-1} \circ x)}{\partial x^i} = \frac{\partial \tilde{f}}{\partial y^j} \frac{\partial y^j}{\partial x^i}$$

where $\frac{\partial \tilde{f}}{\partial y^j}$ is precisely what we write now by $\partial_{y^j} f$ and $\frac{\partial y^j}{\partial x^i}$ must be understood as the derivative with respect to x^i of the function $(y^{-1} \circ x)\colon \mathbb{R}^n \to \mathbb{R}^n$.

Let $f\colon M \to N$ and $g\colon N \to \mathbb{R}$; the definitions gives

$$(df_xX)g = \frac{d}{dt}\Big[(g \circ f)(X(t)) \Big]_{t=0} = \frac{\partial g}{\partial y^i} \frac{\partial f^i}{\partial x^\alpha} \frac{dX^\alpha}{dt}.$$

This shows that $\frac{\partial f^i}{\partial x^\alpha} \frac{dX^\alpha}{dt}$ is $(df_xX)^i$. But dX^α/dt is what we should call X^α in the decomposition $X = X^\alpha \partial_\alpha$ then the matrix of df is given by $\frac{\partial f^i}{\partial x^\alpha}$. So we find back the old notion of differential.

49.8.3 Examples

49.8.3.1 Example: the sphere

The sphere S^n is the set

$$S^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \text{ tel que } \|x\| = 1\}$$

for which we consider the following open set in $\mathbb{R}^n$:

$$\mathcal{U} = \{(u_1, \dots, u_n) \in \mathbb{R}^n \text{ tel que } \|u\| < 1\}$$

and the charts $\varphi_i \colon \mathcal{U} \to S$, and $\tilde{\varphi}_i \colon \mathcal{U} \to S$

$$\varphi_i(u_1, \dots, u_n) = (u_1, \dots, u_{i-1}, \sqrt{1 - \|u\|^2}, u_i, \dots, u_n) \tag{49.265a}$$

$$\tilde{\varphi}_i(u_1, \dots, u_n) = (u_1, \dots, u_{i-1}, -\sqrt{1 - \|u\|^2}, u_i, \dots, u_n). \tag{49.265b}$$

These map are clearly injective. To see that $\varphi(\mathcal{U}) \cup \tilde{\varphi}(\mathcal{U}) = S$, consider $(x_1, \dots, x_{n+1}) \in S$. Then at least one of the x_i is non zero. Let us suppose $x_1 \neq 0$, thus $x_1^2 = 1 - (x_2^2 + \cdots + x_{n+1}^2)$ and

$$x_1 = \pm\sqrt{1 - (\dots)}. \tag{49.266}$$

eq:xnnm

If we put $u_i = x_{i+1}$, we have $x = \varphi(u)$ or $x = \tilde{\varphi}(u)$ following the sign in relation (49.266). The fact that $\varphi^{-1} \circ \tilde{\varphi}$ and $\tilde{\varphi}^{-1} \circ \varphi$ are differentiable is a "first year in analysis exercise".

49.8.3.2 Example: projective space

On $\mathbb{R}^{n+1}\backslash\{o\}$, we consider the equivalence relation $v \sim \lambda w$ for all non zero $\lambda \in \mathbb{R}$, and we put

$$\mathbb{R}P^n = \left(\mathbb{R}^{n+1}\backslash\{o\}\right) / \sim .$$

This is the set of all the one dimensional subspaces of $\mathbb{R}^{n+1}$. This is the real **projective space** of dimension n. We set $\mathcal{U} = \mathbb{R}^n$ and

$$\varphi_i(u_1, \dots, u_n) = \text{Span}\{(u_1, \dots, u_{i-1}, 1, u_i, \dots, u_n)\}.$$

One can see that this gives a manifold structure to $\mathbb{R}P^n$. Moreover, the map

$$A \colon S^n \to \mathbb{R}P^n \; v \quad \mapsto \text{Span}\, v \tag{49.267}$$

is differentiable.

Let us show how to identify $\mathbb{R} \cup \{\infty\}$ to $\mathbb{R}P^1$, the set of directions in the plane $\mathbb{R}^2$. Indeed consider any vertical line l (which does contain the origin). A non vertical vector subspace of $\mathbb{R}^2$ intersects l in one and only one point, while the vertical vector subspace is associated with the infinite point.

49.8.4 Some Leibniz formulas

Lemma 49.122 ([838]).
If M and N are two manifolds, we have a canonical isomorphism

$$T_{(p,q)}(M \times N) \simeq T_pM + T_qN.$$

lemLeibniz

Proof. A $Z \in T_{(p,q)}(M \times N)$ is the tangent vector to a curve $(x(t), y(y))$ in $M \times N$. We can consider $X \in T_pM$ given by $X = x'(0)$ and $Y \in T_qN$ given by $Y = y'(0)$. The isomorphism is the identification $(X, Y) \simeq Z$. Indeed, let us define $\overline{X} \in T_{(p,q)}(M \times N)$, the tangent vector to the curve $(x(t), q)$, and $\overline{Y} \in T_{(p,q)}(M \times N)$, the tangent vector to the curve $(p, y(t))$. Then $Z = \overline{X} + \overline{Y}$ because for any $f \colon M \times N \to \mathbb{R}$,

$$Zf = \frac{d}{dt} f(x(t), y(t))\Big|_{t=0} = \frac{d}{dt} f(x(t), y(0))\Big|_{t=0} + \frac{d}{dt} f(x(0), y(t))\Big|_{t=0} = \overline{X}f + \overline{Y}f. \tag{49.268}$$

□

Leibniz

Proposition 49.123 (Leibniz formula)**.**
Let us consider M, N, V, three manifold; a map $\varphi\colon M \times N \to V$ and a vector $Z \in T_{(p,q)}(M \times N)$ which corresponds (lemma 49.122) to $(X, Y) \in T_pM + T_qN$.

*If we define $\varphi_1\colon M \to V$ and $\varphi_2\colon N \to V$ by $\varphi_1(p') = \varphi(p', q)$ and $\varphi_2(q') = \varphi(p, q')$, we have the **Leibniz formula**:*

$$d\varphi(Z) = d\varphi_1(X) + d\varphi_2(Y). \tag{49.269}$$

Proof. Since $Z = \overline{X} + \overline{Y}$, we just have to remark that

$$d\varphi(\overline{X}) = \frac{d}{dt}\varphi(x(t), q)\Big|_{t=0} = d\varphi_1(X),$$

so $d\varphi(Z) = d\varphi(\overline{X} + \overline{Y}) = d\varphi_1(X) + d\varphi_2(Y)$. □

One of the most important application of the Leibniz rule is the corollary 56.51 on principal bundles.

49.8.5 Cotangent bundle

A form on a vector space V is a linear map $\alpha\colon V \to \mathbb{R}$. The set of all forms on V is denoted by V^* and is called the **dual space** of V. On each point of a manifold, one can consider the tangent bundle which is a vector space. Then one can consider, for each $x \in M$ the dual space $T_x^*M := (T_xM)^*$ which is called the **cotangent bundle**. A 1-**differential form** on M is a smooth map $\omega\colon M \to T^*M$ such that $\omega_x := \omega(x) \in T_x^*M$. So, for each $x \in M$, we have a 1-form $\omega_x\colon T_xM \to \mathbb{R}$.

Here, the smoothness is the fact that for any smooth vector field $X \in \mathfrak{X}(M)$, the map $x \to \omega_x(X_x)$ is smooth as function on M. One often considers vector-valued forms. This is exactly the same, but $\omega_x X_x$ belongs to a certain vector space instead of $\mathbb{R}$. The set of V-valued 1-forms on M is denoted by $\Omega(M, V)$ and simply $\Omega(M)$ if $V = \mathbb{R}$ The cotangent space T_p^*M of M at p is the dual space of T_pM, i.e. the vector space of all the (real valued) linear [46] 1-forms on T_pM. In the coordinate system $x\colon \mathcal{U} \to M$, we naturally use, on T_p^*M, the dual basis of the basis $\{\partial/\partial_{x^i}, \dots \partial/\partial_{x^i}\}$ of T_pM. This dual basis is denoted by $\{dx_1, \dots, dx_n\}$, the definition being as usual:

eq:dx_v

$$dx_i(\partial^j) = \delta_i^j. \tag{49.270}$$

The notation comes from the fact that equation (49.270) describes the action of the differential of the projection $x_i\colon \mathcal{U} \to \mathbb{R}$ on the vector ∂^j.

If $(\mathcal{U}_\alpha, \varphi_\alpha)$ is a chart of M, then the maps

$$\phi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{R}^n \to T^*M \ (x, a) \ \mapsto a^i dx_i|_x \tag{49.271}$$

give to T^*M a $2n$ dimensional manifold structure such that the canonical projection $\pi\colon T^*M \to M$ is an immersion.

When V is a finite-dimensional vector space, we denote by V^* its dual [47] and we often use the identifications $V \simeq V^* \simeq T_vV \simeq T_wV \simeq T_v^*V$ where v and w are any elements of V. Note however that there are no *canonical* isomorphism between these spaces, unless we consider some basis.

49.8.6 Immersion, embedding

DEFooZEWNooMVOzWI

Definition 49.124 ([839])**.**
*A smooth map $f\colon M \to N$ between the manifolds M and N is an **immersion** if its differential $df_p\colon T_pM \to T_{f(p)N}$ is injective for every $p \in M$.*

46. When we say *a form*, we will always mean *a linear form*.
47. The vector space of all the linear map $V \to \mathbb{R}$.

Definition 49.125 (Topological embedding[840]).
*Let X, Y be topological spaces. A map $f\colon X \to Y$ is a **topological embedding** if*

(1) f is continuous

(2) f is injective

(3) $f\colon X \to f(X)$ is an homeomorphism when $f(X)$ is equipped with the induced topology from Y.

DEFooQLGLooNyXaOV

Definition 49.126 (Embedding[840]).
*Let M and N be smooth manifolds. A smooth function $f\colon M \to N$ is an **embedding** if*

(1) f is an immersion,

(2) f is a topological embedding.

The following theorem says that immersions are local embeddings.

THOooXAOUooRKHMBm

Theorem 49.127 ([841]).
Let M be a n-dimensional manifold and S be a m-dimensional manifold. Let $f\colon M \to S$ be an immersion. For every $a \in M$, there exists an open set $\mathcal{O}$ containing a in M such that $f\colon \mathcal{O} \to f(\mathcal{O})$ is an embedding.

Proof. We consider maps $\varphi\colon U \to M$ and $\psi\colon V \to S$ around a and $f(a)$ respectively. Then we consider :

(1) $\tilde{f} = \psi^{-1} \circ f \circ \varphi$

(2) We rename U in such a way that $f \circ \varphi(U) \subset \psi(V)$ and we restrict φ to that subset of U.

(3) With that restriction of U, the map $\tilde{f}\colon U \to V$ is defined on U.

(4) $q = \varphi^{-1}(a)$.

By hypothesis, f is an immersion. It means that the map $df_a\colon T_aM \to T_{f(a)}S$ is injective. The differential $d\tilde{f}_q\colon \mathbb{R}^n \to \mathbb{R}^m$ of $\tilde{f}\colon U \to V$ is injective too. This has two consequences :

$$\det(d\tilde{f}_q) \neq 0 \tag{49.272}$$

and its image has dimension n. We suppose that

$$\text{Image}(d\tilde{f}_q) = \text{Span}\{e_1, \dots, e_n\}. \tag{49.273}$$

If not, we consider a linear bijection map $h\colon \mathbb{R}^m \to \mathbb{R}^m$ which maps $\text{Image}(d\tilde{f}_q)$ to $\text{Span}\{e_1, \dots, e_n\}$ and we rewrite the whole proof with $\tilde{f} \circ h$ instead of $\tilde{f}$.

Let $k = m - k$. We consider the map

$$\begin{aligned} \phi\colon U \times \mathbb{R}^k &\to \mathbb{R}^m \\ (x_1, \dots, x_n, t_1, \dots, t_k) &\mapsto \big(\tilde{f}(x)_1, \dots, \tilde{f}(x)_n, \tilde{f}(x)_{n+1} + t_1, \dots, \tilde{f}(x)_m + t_k\big). \end{aligned} \tag{49.274}$$

It satisfies $\phi(x, 0) = \tilde{f}(x)$ and the matrix of its differential is

$$d\phi_{(q,0)} = \begin{pmatrix} d\tilde{f}_q & 0 \\ M & \mathbb{1}_k \end{pmatrix} \tag{49.275}$$

where M a is some $k \times n$ matrix. Since it is block-diagonal, its determinant is given by

$$\det\big(d\phi_{(q,0)}\big) = \det(d\tilde{f}_q)\det(\mathbb{1}_k) = \det(d\tilde{f}_q) \neq 0. \tag{49.276}$$

Since $\det(d\phi_{(q,0)}) \neq 0$, we can apply the local inversion theorem 17.47 for ϕ at $(a, 0)$. We have a neighbourhood W of $(q, 0)$ in $U \times \mathbb{R}^k$ such that $\phi\colon W \to \phi(W)$ is a diffeomorphism.

Let $U' = \texttt{proj}_{\mathbb{R}^n}(W)$. We will prove that $\tilde{f}\colon U' \to \tilde{f}(U')$ is a diffeomorphism.

(1) **Injective** Let $x, y \in U'$ such that $\tilde{f}(x) = \tilde{f}(y)$. It means $\phi(x,0) = \phi(y,0)$. The points $(x,0)$ and $(y,0)$ belong to W while ϕ is injective on W, thus $(x,0) = (y,0)$ and $x = y$.

(2) **Surjective** This is not an issue since $\tilde{f}$ is obviously surjective on $\tilde{f}(U')$.

(3) **The inverse** The prove that the inverse of $\tilde{f}$ is $\texttt{proj}_{\mathbb{R}^n} \circ \phi^{-1}$. Let $w \in U' = \texttt{proj}_{\mathbb{R}^n}(W)$ and compute

$$(\texttt{proj}_{\mathbb{R}^n} \circ \phi^{-1} \circ \tilde{f})(w) = (\texttt{proj}_{\mathbb{R}^n} \circ \phi^{-1})\big(\omega(w,0)\big) = \texttt{proj}_{\mathbb{R}^n}(w,0) = w. \tag{49.277}$$

(4) **Diffeomorphism** Since $\texttt{proj}_{\mathbb{R}^n}$ and ϕ^{-1} are differentiable, the inverse of $\tilde{f}$ is differentiable and $\tilde{f}$ is a diffeomorphism.

The fact for $\tilde{f}$ to be a diffeomorphisms implies that f itself (restricted to $\varphi(U')$) is bijective, differentiable with the inverse being differentiable. □

PROPooYHOKooYASzRL

Proposition 49.128.
If $f\colon M \to N$ is an embedding, the image Image(f) *is a smooth* $\dim(M)$*-dimensional submanifold of N.*

49.9 Distribution and Frobenius theorem

LEMooQXRSooHDRequ

Lemma 49.129.
Let $a, b \in C^\infty(M)$. We have

$$[a\partial_i, b\partial_j] = a(\partial_i b)\partial_j - b(\partial_j a)\partial_i. \tag{49.278}$$

In particular, $[\partial_i, \partial_j] = 0$.

LEMooGHQZooMIWAfn

Lemma 49.130 ([1])**.**
Let M be a n-dimensional manifold. Let $\varphi\colon U \to M$ be a chart, and $X_1, \dots, X_r$ $(r < n)$ be vectors fields such that for each $p \in \varphi(U)$, the set $\{X_i(p)\}_{i=1,\dots,r}$ is linearly independent in T_pM.

There exist vector fields $X_{r+1}, \dots, X_n$ such that $\{X_1(p), \dots, X_n(p)\}$ is a basis of T_pM for every $p \in \varphi(U)$.

DEFooYOMHooZJvsSt

Definition 49.131.
Let M be a manifold. A k dimensional ***distribution*** *is the data, for each $p \in M$ of a k dimensional vector subspace $\mathcal{D}_p$ of T_pM.*

The distribution $\mathcal{D}$ is said to be ***smooth*** *if for each $p \in M$, there exists a neighbourhood S of p and smooth vector fields*[48] *$\{X_i\}_{i=1,\dots,k}$ in S such that*

$$\mathcal{D}_p = \text{Span}\{X_i(p)\}. \tag{49.279}$$

We denote by $\Gamma^\infty(\mathcal{D})$ the set of smooth vector fields with values in $\mathcal{D}$. These are the smooth vector fields X such that $X_p \in \mathcal{D}_p$ for every p. More precisely,

$$\Gamma^\infty(\mathcal{D}) = \{X \in \Gamma(TM) \text{ tel que } X_p \in \mathcal{D}_p \, \forall p \in M\}. \tag{49.280}$$

Definition 49.132.
Let $\mathcal{D}$ be a distribution on M. A smoothly immersed[49] *submanifold $N \subset M$ is an* ***integral manifold*** *of $\mathcal{D}$ if for every $p \in N$ we have*

$$di_p(T_pN) = \mathcal{D}_p \tag{49.281}$$

where $i\colon N \to M$ is the immersion.

When an integral submanifold exists, we say that the distribution is ***integrable****.*

48. Definition 49.88.
49. Definition 49.110.

Definition 49.133.
*A distribution $\mathcal{D}$ is **involutive** if we have*

$$[X, Y] \in \Gamma^\infty(\mathcal{D}) \tag{49.282}$$

for every $X, Y \in \Gamma^\infty(TM)$. In other words, for every $p \in M$ we have $[X, Y]_p \in \mathcal{D}_p$.

PROPooDYJNooAwnaFK

Proposition 49.134.
An smooth integrable distribution is involutive.

Proof. Let N be a smooth integral manifold for the distribution $\mathcal{D}$. Let $X, Y \in \Gamma^\infty(\mathcal{D})$. Since $T_pN = \mathcal{D}_p$, for each $p \in M$, we have $X, Y \in \Gamma^\infty(TN)$. So

$$[X, Y] \in \Gamma^\infty(TN) = \Gamma^\infty(\mathcal{D}) \tag{49.283}$$

and $\mathcal{D}$ is involutive. □

THOooDVBHooGRhuGl

Theorem 49.135 (Frobenius[832, 837])**.**
An involutive distribution on a smooth manifold is integrable.

Proof. Let $\mathcal{D}$ be an involutive r-dimensional distribution of the smooth n-dimensional manifold M. Let $p \in M$. We will build a submanifold N of M such that $T_qN = \mathcal{D}_q$ for every q in a neighbourhood of p. By definition of a smooth distribution, there exists a neighbourhood S of p and vector fields $\{Y_i\}_{i=1,\dots,r}$ such that

$$D_q = \operatorname{Span}\{Y_i(q)\}_{i=1,\dots,r} \tag{49.284}$$

for every $q \in S$. The vectors $Y_i(q)$ are linearly independent since there are r of them and they span the r-dimensional vector subspace $\mathcal{D}_q$ of T_qM.

We may reduce S to fit in a local chart $\varphi\colon U \to M$ around p. We denote by ∂_i the basis vector associated with the chart φ. For some $a_{ij}(q)$ we have

$$Y_i = \sum_{j=1}^{n} a_{ij}\partial_j. \tag{49.285}$$

We use the lemma 49.130 to consider $Y_{r+1}, \dots, Y_n$ such that, for every $q \in \varphi(U)$, $\operatorname{Span}\{Y_1(q), \dots, Y_n(q)\} = T_qM$. Each of these Y_i can be decomposed in the basis $\{\partial_i\}$:

$$Y_i(q) = \sum_{j=1}^{n} a_{ij}(q)\partial_j(q). \tag{49.286}$$

The matrix $a_{ij}(p)$ is invertible. By lemma 4.97, let $\sigma \in S_n$ be such that the matrix $A \in \mathbb{M}(r, \mathbb{R})$ given by

$$A_{ij} = a_{i,\sigma^{-1}(j)}(p) \tag{49.287}$$

is invertible. We also consider the associated linear bijection

$$\begin{aligned} \sigma\colon \mathbb{R}^n &\to \mathbb{R}^n \\ e_i &\mapsto e_{\sigma(i)}. \end{aligned} \tag{49.288}$$

Now we consider a new chart $U' = \sigma^{-1}(U)$ with

$$\begin{aligned} \varphi'\colon U' &\to M \\ x &\mapsto (\varphi \circ \sigma)(x). \end{aligned} \tag{49.289}$$

By continuity of the determinant, there exists an open neighbourhood S of p such that $\det\left(a_{i\sigma^{-1}(j)}(q)\right) \neq 0$ for every $q \in S$. We restrict U' in such a way for $\varphi'(U')$ to fit in S.

We denote by ∂'_i the corresponding vectors in TM. We have $\partial'_i(q) = \partial_{\sigma(i)}(q)$; indeed let $q = \varphi'(v')$ and compute:

$$\partial'_i(q)f = \frac{d}{dt}\Big[(f\circ\varphi')(v'+te_i)\Big]_{t=0} \tag{49.290a}$$
$$= \frac{d}{dt}\Big[(f\circ\varphi)\big(\sigma(v')+te_{\sigma(i)}\big)\Big]_{t=0} \tag{49.290b}$$
$$= \partial_{\sigma(i)}(q)f \tag{49.290c}$$

because $\varphi\big(\sigma(v')\big) = q$. We can express Y_i in terms of ∂'_j instead of ∂_j:

$$Y_i(q) = \sum_{j=1}^{n} a_{ij}(q)\partial_j(q) = \sum_j a_{ij}(q)\partial'_{\sigma(j)}(q) = \sum_j a_{i\sigma^{-1}(j)}(q)\partial'_j(q). \tag{49.291}$$

For the sake of simplifying the notations, we redefine U for being U', φ for φ' and so on. Thus we have

$$Y_i(q) = \sum_{j=1}^{n} a_{ij}\partial_j(q) \tag{49.292}$$

with $(a_{ij})_{i,j=1,\dots,r}$ being invertible.

Let $A\in \mathrm{GL}(r,\mathbb{R})$ be the matrix given by $A_{ij} = a_{ij}$. We consider $B = A^{-1}$ and, for $i = 1,\dots,r$, we define

$$X_i = \sum_{k=1}^{r} B_{ik}Y_k. \tag{49.293}$$

We can express them in terms of the ∂_i's:

$$X_i = \sum_{k=1}^{r} B_{ik}Y_k \tag{49.294a}$$
$$= \sum_{k=1}^{r} B_{ik}\sum_j a_{kj}\partial_j \tag{49.294b}$$
$$= \sum_{k=1}^{r} B_{ik}\Big(\sum_{j=1}^{r} a_{kj}\partial_j + \sum_{j=r+1}^{n} a_{kj}\partial_j\Big) \tag{49.294c}$$
$$= \sum_{k,j=1}^{r} B_{ik}a_{kj}\partial_j + \sum_{k=1}^{r}\sum_{j=r+1}^{n} B_{ik}a_{kj}\partial_j \tag{49.294d}$$
$$= \sum_{j=1}^{r}\delta_{ij}\partial_j + \sum_{j=r+1}^{n} c_{ij}\partial_j \tag{49.294e}$$
$$= \partial_j + \sum_{j=r+1}^{n} c_{ij}\partial_j \tag{49.294f}$$

where c_{ij} is matrix whose value has no importance here.

We know that $\mathcal{D}_q = \mathrm{Span}\{Y_1(q),\dots,Y_r(q)\}$ for every q. Since the Y_i's are linearly independent and since the matrix B is invertible, the vectors $\{X_i(q)\}_{i=1,\dots,r}$ are linearly independent and belong to $\mathcal{D}_q$. Thus we have

$$\mathcal{D}_q = \mathrm{Span}\{X_1(q),\dots,X_r(q)\}. \tag{49.295}$$

We compute the commutators of the X_i's using the lemma 49.129:

$$[X_i,X_j] = [\partial_i,\partial_j] + \sum_{k=r+1}^{n}[\partial_i,c_{ik}\partial_k] + \sum_{k=r+1}^{n}[c_{ik}\partial_k,\partial_j] + \sum_{k=r+1}^{n}\sum_{l=r+1}^{n}[c_{ik}\partial_k,c_{jl}\partial_l] \tag{49.296a}$$
$$= \sum_{k=r+1}^{n} d_k\partial_k \tag{49.296b}$$

for some functions d_k. Now the distribution $\mathcal{D}$ is involutive, so there exists functions f_k such that

$$[X_i, X_j] = \sum_{k=1}^{r} f_k X_k = \sum_{k=1}^{r} f_k\big(\partial_k + \sum_{l=r+1}^{n} c_{kl}\partial_l\big). \tag{49.297}$$

Thus we have

$$\sum_{k=r+1}^{n} d_k\partial_k = \sum_{k=1}^{r} f_k\big(\partial_k + \sum_{l=r+1}^{n} c_{kl}\partial_l\big). \tag{49.298}$$

On the left hand side there are no components in $\{\partial_k\}_{k=1,\dots,r}$; then on the right hand side we deduce $\sum_{k=1}^{r} f_k\partial_k = 0$ and then $f_k = 0$ for every $k = 1, \dots, r$. We proved that

$$[X_i, X_j] = 0. \tag{49.299}$$

We consider the flows Φ^i of each X_i (proposition 49.116), and the map

$$\begin{aligned} s\colon I &\to M \\ (t_1, \dots, t_r) &\mapsto (\Phi^1_{t_1} \circ \dots \circ \Phi^t_{t_2})(p) \end{aligned} \tag{49.300}$$

where I is an open neighbourhood of $(0, \dots, 0)$ in $\mathbb{R}^r$. The differential at $0 \in \mathbb{R}^r$ is a map $ds_0\colon T_0\mathbb{R}^r \to T_pM$ and satisfy

$$ds_0(\partial_i)\frac{d}{dt}\Big[\Phi^i_t(p)\Big]_{t=0} = X_i(p). \tag{49.301}$$

Since the vectors $\{X_i(p)\}$ are linearly independent, the map ds_0 is injective. This means that $s\colon I \to M$ is an immersion [50]. Theorem 49.127 says that, reducing the size of I, the map s becomes an embedding.

We define $N = \text{Image}(s)$. This is a r-dimensional submanifold of M by proposition 49.128. Since $s(0) = p$ and $ds_0(\partial_i) = X_i$ we have $T_pN = \mathcal{D}_p$. Ou task is to prove that $\mathcal{D}_q = T_qN$ for every $q \in N$.

Let $q \in N$. There exists $(t_1, \dots, t_r) \in \mathbb{R}^r$ such that

$$q = s(t_1, \dots, t_r) = (\Phi^1_{t_1} \circ \dots \circ \Phi^r_{t_r})(p). \tag{49.302}$$

We fix i. As a recall, the vector $\partial_i \in T_{(t_1,\dots,t_r)}\mathbb{R}^n$ is the derivative of the path $(t_1, \dots, t_r) + te_i$. We have

$$\begin{aligned} ds_{(t_1,\dots,t_r)}(\partial_i) &= \frac{d}{dt}\Big[s\big((t_1, \dots, t_r) + te_i\big)\Big]_{t=0} && \text{(49.303a)} \\ &= \frac{d}{dt}\Big[(\Phi^i_t \circ \Phi^1_{t_1} \circ \dots \circ \Phi^r_{t_r})(p)\Big]_{t=0} && \text{(49.303b)} \\ &= \frac{d}{dt}\Big[\Phi^i_t(q)\Big]_{t=0} && \text{(49.303c)} \\ &= X_i(q). && \text{(49.303d)} \end{aligned}$$

Justifications.

— For (49.303b). We used the property (49.249) of proposition 49.116 and the commutation of proposition 49.120.

— For (49.303c). By definition $q = (\Phi_{t_1} \circ \dots \circ \Phi^r_{t_r})(p)$.

So $\text{Span}\{X_1(q), \dots, X_r(q)\}$ is a r-dimensional subspace of T_qM which is included in the r-dimensional subspace T_qN. We conclude that

$$T_qN = \text{Span}\{X_1(q), \dots, X_r(q)\} = \mathcal{D}_qN. \tag{49.304}$$

□

50. Definition 49.124.

THOooVRDYooIusxwW

Theorem 49.136 (Frobenius[832, 837]).
About distributions.

ITEMooLDUZooJCwZek

(1) A distribution is integrable if and only if it is involutive.

ITEMooCQEAooRsLSOV

(2) If $\mathcal{D}$ is involutive, for every $p \in M$, there exists a unique maximal[51] *connected integral submanifold containing p.*

Proof. Point (1) is given by the proposition 49.134 and the theorem 49.135.
Proof of point (2) still to be done... □

49.9.1 Analytic

PROPooKAXOooZCwcwi

Proposition 49.137 ([842]).
Let M, N and S be analytic manifolds and $f\colon M \to N$ be analytic.

(1) We suppose that f is an immersion and that $g\colon S \to M$ is continuous. The map g is analytic if and only if $f \circ g$ is analytic.

(2) We suppose that f is a submersion and that $g\colon N \to S$ is continuous. The map g is analytic if and only if $g \circ f$ is analytic.

49.9.2 Tensor algebra

DEFooHPQXooETvEyn

Definition 49.138 (Tensor algebra).
*Let V be a vector space over $\mathbb{C}$. The **tensor algebra** of V is the vector space*

$$T(V) = \bigoplus_{n \geqslant 0} (\otimes^n V) = \mathbb{C} \oplus V \oplus (V \otimes V) \oplus \dots \tag{49.305}$$

49.10 Rank theorem

We proof a generalization of the rank theorem 4.47.

Definition 49.139.
*Let M and N be smooth manifolds of dimension m and n. Let a smooth map $f\colon M \to N$. The **rank** of f at $p \in M$ is the rank of the linear map $df_p\colon T_pM \to T_{f(p)N}$.*

Lemma 49.140.
Let a smooth map $f\colon M \to N$ and $p \in M$. Let $\varphi\colon U \to M$ and $\psi\colon V \to N$ be chats around p and $f(p)$. Then

$$\operatorname{rk}(df_p) = \operatorname{rk}\left(f_{\varphi^{-1}(p)}(\psi^{-1} \circ f \circ \varphi)\right) \tag{49.306}$$

where the rank on the right han side is the usual rank of a map $\mathbb{R}^m \to \mathbb{R}^n$.

Proof. By proposition 4.110 we can compute the rank of a linear map in whatever base. When a basis is chosen in $\mathbb{R}^m$ and $\mathbb{R}^n$ we know from lemma 49.98 that the matrix of df_p is the same as the one of $f_{\varphi^{-1}(p)}(\psi^{-1} \circ f \circ \varphi)$. Since these two linear maps have the same matrix, they have the same rank. □

THOooSWKVooTJQsXc

Theorem 49.141 (Rank theorem[843]).
Let M and N be smooth manifolds of dimension m and n. Let $f\colon M \to N$ be a smooth map. Let $p \in M$. We suppose that the rank of f is equal to k at every point x in a neighbourhood of p.
There exists charts $\varphi\colon U \to M$ around $p \in M$ and $\psi\colon V \to N$ around $f(p) \in M$ such that

(1) $\varphi(0) = p$,

(2) $\psi(0) = f(p)$

51. As far as understand, here «maximal» means that every integral manifold containing p is contained in that one. It is used in that sense during the proof the theorem 53.82.

(3) the function f is more or less trivialized in the sense that

$$(\psi^{-1} \circ f \circ \varphi)(x_1, \dots, x_m) = (x_1, \dots, x_k, 0, \dots, 0) \tag{49.307}$$

for every $(x_1, \dots, x_m) \in U$.

Proof. We prove in two parts. Fist we consider the case in which M and N are open sets of $\mathbb{R}^m$ and $\mathbb{R}^n$. Then we will generalize to any smooth manifolds.

(1) **The case of $\mathbb{R}^m$ and $\mathbb{R}^n$** Let W be open in $\mathbb{R}^m$, W' be open in $\mathbb{R}^n$. We consider a smooth map $f : W \to W'$ such that $f(0) = 0$ and $\operatorname{rk}(f) = k$ on W.
By hypothesis, the rank of df_0 is k, so that is one chooses good bases on $\mathbb{R}^m$ and $\mathbb{R}^n$ we can suppose that the matrix of df_0 has a upper-left square $k \times k$ with non-zero determinant. We write A that square matrix:

$$A_{ij} = \frac{\partial f_i}{\partial x_j}(0) \tag{49.308}$$

with $i, j \leqslant k$.

(1a) **On the $\mathbb{R}^m$ side**
We consider the map

$$\begin{aligned} &\varphi : W \subset \mathbb{R}^m \to \mathbb{R}^m \\ &(x_1, \dots, x_m) \mapsto \big(f_1(x_1, \dots, x_m), \dots, f_k(x_1, \dots, x_m), x_{k+1}, \dots, x_m\big). \end{aligned} \tag{49.309}$$

We have $\varphi(0) = 0$ because $f_i(0) = f(0)_i = 0$. The matrix of the differential is

$$d\varphi_0 = \begin{pmatrix} A & * \\ 0 & \mathbb{1}_{n-k} \end{pmatrix} \tag{49.310}$$

where A is $k \times k$ and $*$ is some $k \times (n-k)$ matrix. Thus we have $\det(d\varphi_0) = \det(A) \neq 0$. From the inverse function theorem 17.47, the map φ is a local diffeomorphism, more precisely there exists an open set $W_1 \subset W \subset \mathbb{R}^m$ such that the restriction

$$\varphi : W_1 \to W_1 \tag{49.311}$$

is a diffeomorphism. From now on we only consider φ as being that restriction.
The vector $(y_1, \dots, y_m)$ such that $\varphi^{-1}(x_1, \dots, x_m) = (y_1, \dots, y_m)$ has the property that

$$\varphi(y_1, \dots, y_m) = (x_1, \dots, x_m), \tag{49.312}$$

which means that [52]

$$f_i(y_1, \dots, y_m) = x_i \tag{49.313}$$

when $i = 1, \dots, k$ and

$$y_l = x_l \tag{49.314}$$

when $l = k + 1, \dots, m$.

(1b) **On the middle side**
Thus we have

$$\begin{aligned} (f \circ \varphi^{-1})(x_1, \dots, x_m) &= f(y_1, \dots, y_m) && \text{(49.315a)} \\ &= \big(x_1, \dots, x_k, f_{k+1}(y_1, \dots, y_m), \dots, f_n(y_1, \dots, y_m)\big) && \text{(49.315b)} \\ &= \big(x_1, \dots, x_k, \tilde{f}_{k+1}(x), \dots, \tilde{f}_n(x)\big) && \text{(49.315c)} \end{aligned}$$

EQooAQJGooLqbpXJ

where $\tilde{f}_i = f_i \circ \varphi^{-1} : W_1 \to \mathbb{R}$ are some smooth functions.

52. At this point, it is really important that f takes its values in $\mathbb{R}^n$, not in a general manifold: if (y) was in a manifold, the expression $f_i(y)$ would not make sense.

For every $x \in W_1$ we have

$$f(f \circ \varphi^{-1})_x = \begin{pmatrix} \mathbb{1}_{k\times k} & 0 \\ * & d\tilde{f}_x \end{pmatrix} \tag{49.316}$$

EQooEDJJooLyPslk

where $d\tilde{f}_x$ is the matrix whose elements are $\left(\frac{\partial \tilde{f}_i}{\partial x_s}\right)$ with $i = k+1, \dots, n$ and $s = k+1, \dots, m$. This is not a square matrix by the way. We have, by theorem 10.276,

$$d(f \circ \varphi^{-1})_x = df_{\varphi^{-1}(x)} \circ (d\varphi^{-1})_x \tag{49.317}$$

while $(d\varphi^{-1})_x$ is invertible. Thus

$$\operatorname{rk}\left(d(f \circ \varphi^{-1})_x\right) = \operatorname{rk}\left(df_{\varphi^{-1}(x)}\right) = k. \tag{49.318}$$

So the rank of $f \circ \varphi^{-1}$ is k all over W_1. But the image of $d(f \circ \varphi^{-1})_x$ is spanned by the columns of its differential given by (49.316). The k columns spanned by the identity matrix are obviously linearly independent; these are thus a basis of the image. Since the vectors in the "$d\tilde{f}_x$" part are linearly independent of these k vectors, they must be vanishing:

$$\frac{\partial \tilde{f}_i}{\partial x_s}(x) = 0 \tag{49.319}$$

for every $x \in W_1$, $i = k+1, \dots, m$ and $s = k+1, \dots, n$.

(1c) **On the $\mathbb{R}^n$ side**

We do not know if $n \geqslant m$ or $m \geqslant n$. If $n \geqslant n$, we choose V_1 such that the projection of V_1 on its m first components is included in W_1. If $n < m$ we choose V_1 such that the projection of W_1 on its n first components is included in V_1 [53].

With that choice of V_1 in mind we can remember the functions $\tilde{f}_i \colon W_1 \to \mathbb{R}$. If $y \in V_1$, we define $\tilde{f}(y)$ as $\tilde{f}(x)$ with $x \in W_1$ created from y either by adding zeroes or by projecting on $\mathbb{R}^m$. In both cases, the resulting y belongs to V_1.

So now we consider the map

$$\begin{aligned} T \colon V_1 &\to \mathbb{R}^n \\ (y_1, \dots, y_n) &\mapsto \big(y_1, \dots, y_k, y_{k+1} + \tilde{f}_{k+1}(y), \dots, y_n + \tilde{f}_n(y)\big). \end{aligned} \tag{49.320}$$

EQooKEZQooSQTBlo

If $y \in V_1$, the differential is the matrix given by

$$(dT_y)_{ij} = \frac{\partial T_i}{\partial y_j}(y) \tag{49.321}$$

where

— The upper-left $k \times k$ corner is $\mathbb{1}_{k\times k}$.

— The upper-right $k \times (n-k)$ corner (non square in general) is given by elements of the form

$$\frac{\partial y_i}{\partial y_j} \tag{49.322}$$

with $i \leqslant k$ and $j > k$. So this is vanishing.

— The lower-left (non square in general) corner is made of

$$\frac{\partial (y_i + \tilde{f}_i(y))}{\partial y_j} = \frac{\partial \tilde{f}_i(y)}{\partial y_j} \tag{49.323}$$

with $i > k$ and $j \leqslant k$. The elements in this pars are some numbers.

53. This precision about the choice of V_1 is not done in [843] and seems strange to me. Am I correct ? By the way, there could be a misprint in the definition of T in [843]: y must have n components, not m.

— The lower-right square $(n-k)\times(n-k)$ corner is made of

$$\frac{\partial(y_i+\tilde{f}_i(y))}{\partial y_j}=\delta_{ij}+\frac{\partial\tilde{f}_i}{\partial y_j} \tag{49.324}$$

with $i>k$ and $j\geqslant k$. For these elements we have $\frac{\partial\tilde{f}_i(y)}{\partial y_j}=0$ and then the identity matrix.

With all that,

$$dT_y=\begin{pmatrix}\mathbb{1}_{k\times k} & 0\\ * & \mathbb{1}_{n-k}\end{pmatrix}. \tag{49.325}$$

Moreover $T(0)=0$ because

$$\tilde{f}_i(0)=f_i\big(\varphi^{-1}(0)\big)=f_i(0)=0. \tag{49.326}$$

We deduce that there exist an open set $V\subset\mathbb{R}^n$ included in V_1 such that $T\colon V\to T(V)$ is a diffeomorphism. We restrict V in such a way that $T(V)\subset V_1$.

(1d) **The final map**

Finally we consider the map

$$T^{-1}\circ f\circ\varphi^{-1}\colon W_1\to V. \tag{49.327}$$

If $(x_1,\dots,x_m)\in W_1$ from (49.315a) we have

$$(f\circ\varphi^{-1})(x_1,\dots,x_m)=\big(x_1,\dots,x_k,\tilde{f}_{k+1}(x),\dots,\tilde{f}_n(x)\big). \tag{49.328}$$

Using the definition (49.320) we see that

$$T(x_1,\dots,x_k,0,\dots,0)=\big(x_1,\dots,x_k,\tilde{f}_{k+1}(x),\dots,\tilde{f}_n(x)\big) \tag{49.329}$$

which proves that

$$T^{-1}\big(x_1,\dots,,x_k,\tilde{f}_{k+1}(x),\dots,\tilde{f}_n(x)\big)=(x_1,\dots,x_k,\dots,0). \tag{49.330}$$

(2) **The general case**

Now we consider the manifolds M and N with the map $f\colon M\to N$. Let $p\in M$ and charts $\varphi_0\colon U_0\to M$, $\psi_0\colon V_0\to N$ where U_0 is a neighbourhood of 0 in $\mathbb{R}^m$ and V_0 a neighbourhood of 0 in $\mathbb{R}^n$. We suppose that $\varphi_0(0)=p$ and $\psi_0(0)=f(p)$.

Now we consider the function $\tilde{f}=\psi_0^{-1}\circ f\circ\varphi_0$ from U_0 to V_0 and we are left in the previous case.

□

If the differential is bijective, then the function is a local diffeomorphism.

THOooDWEXooMClWVi

Theorem 49.142 (local inversion[1])**.**
Let M,N be C^k manifolds. Consider a C^k map $f\colon M\to N$. If $df_p\colon T_pM\to T_{f(p)}N$ is a bijection, the f is a local C^k-diffeomorphism.

More precisely, there exists a neighbourhood V_1 of p in M and a neighbourhood V_2 of $f(p)$ in M such that $f\colon V_1\to V_2$ is a C^k-diffeomorphism.

Due to proposition 52.5, a local diffeomorphism can be often converted into a global diffeomorphism.

49.11 Exterior calculus

49.11.1 The exterior algebra

Definition 49.143 (Exterior product)**.**
*If V is a vector space, we denote by $\Lambda^k V^*$ the space of all the k-form on V. We define the **exterior product** $\wedge : \Lambda^k V^* \times \Lambda^l V^* \to \Lambda^{k+l} V^*$ by*

$$(\omega^k \wedge \eta^l)(v_1, \dots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} (-1)^\sigma \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, v_{\sigma(k+1)}) \tag{49.331}$$

If $\{e_1, \dots, e_n\}$ is a basis of V, the dual basis $\{\sigma^1, \dots, \sigma^n\}$ of V^* is defined by $\sigma^i(e_j) = \delta^i_j$.
If $I = \{1 \leqslant i_1 \leqslant \dots i_k \leqslant n\}$, we write $\sigma^I = \sigma^{i_1} \wedge \dots \sigma^{i_k}$ any k-form can be decomposed as

$$\omega = \sum_I \omega_I \sigma^I.$$

The exterior algebra is provided with the **interior product** denoted by ι. It is defined by[pg_DefProdExt]

$$\begin{aligned} \iota(v_0) &: \Lambda^k W \to \Lambda^{k-1} W \\ (\iota(v_0)\omega)(v_1, \dots, v_{k-1}) &= \omega(v_0, v_1, \dots, v_{k-1}). \end{aligned} \tag{49.332}$$

Lemma 49.144.
Let σ be an element of the symmetric group[54] of the set $\{a_1, \dots, a_n\}$ where the a_i are integers. Then

$$(dx_{a_1} \wedge \dots \wedge dx_{a_n})(e_{\sigma(a_1)}, \dots, e_{\sigma(a_n)}) = (-1)^\sigma. \tag{49.333}$$

Proof. We make it by induction over n. With $n = 1$ the only permutation is the identity; the claim reduces to $dx_{a_1}(e_{a_1}) = 1$. Let us try with $n = 2$. Up to renumbering we have

$$(dx_1 \wedge dx_2)(e_1, e_2) = 1 \tag{49.334}$$

and

$$(dx_1 \wedge dx_2)(e_2, e_1) = -1. \tag{49.335}$$

We pass to the induction. Let $\sigma \in S_n$. We have

$$(dx_{a_1} \wedge dx_{a_n})(e_{\sigma(a_1)}, \dots, e_{\sigma(a_n)}) = dx_{a_1} \wedge (dx_{a_n})(e_{\sigma(a_1)}, \dots, e_{\sigma(a_n)}) \tag{49.336a}$$

$$= \sum_{\phi \in S_n} (-1)^\phi \frac{1}{(n-1)!} dx_{a_1}\big(e_{\phi\sigma(a_1)}\big)(dx_{a_2} \wedge \dots \wedge dx_{a_n})(e_{\phi\sigma(a_2)}, \dots, e_{\phi\sigma(a_n)}) \tag{49.336b}$$

$$= \sum_{\phi \in S_n} (-1)^\phi \frac{1}{(n-1)!} \delta_{a_1, \phi\sigma(a_1)} (-1)^{\phi\sigma} \tag{49.336c}$$

$$= \sum_{\phi \in S_n} \delta_{a_1, \phi\sigma(a_1)} (-1)^\sigma \frac{1}{(n-1)!} \tag{49.336d}$$

where we used the fact that the sign of a permutation provides a morphism between S_n and $\{-1, 1\}$ (proposition 1.300(1)). In the sum over S_n, only the ϕ that make $\sigma(a_1) \to a_1$ remain; there are $|S_{n-1}| = (n-1)!$ such elements. Thus the whole evaluates to $(-1)^\sigma$. □

LEMooICRXooFKPCRd

Lemma 49.145 ([1])**.**
Let $\tau_i : \mathbb{R}^n \to \mathbb{R}^{n-1}$ defined by

$$\tau_i(v)_k = \begin{cases} v_k & \text{if } k < i \\ v_{k+1} & \text{if } k \geqslant i. \end{cases} \tag{49.337}$$

54. Definition 1.274.

Then we have

$$(dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n)(v_1, \ldots, \widehat{v_i}, \ldots, v_n) = \det\Big(\tau_i(v_1), \ldots, \widehat{\tau_i(v_i)}, \ldots, \tau_i(v_n)\Big) \tag{49.338}$$

where the hat denotes a non present term.

Proof. We extend τ_i to the dual : $\tau_i \colon (\mathbb{R}^n)^* \to (\mathbb{R}^{n-1})^*$ is defined by

$$\tau_i(dx_k) = \begin{cases} dy_k & \text{if } k < i \\ dy_{k-1} & \text{if } k > i \end{cases} \tag{49.339}$$

(not defined on dx_i). It is easy to check that, if $k \neq i$,

$$\tau_i(dx_k)\tau_i(v) = dx_k(v). \tag{49.340}$$

The value of $(dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n)(v_1, \ldots, \widehat{v_i}, \ldots, v_n)$ is a polynomial in the variables $dx_k(v_l)$ (with $k \neq l$). Since $dx_k(v_l) = \tau ß(dx_k)\big(\tau_i v_l\big)$, the same polynomial will give the value of

$$(\tau_i dx_1 \wedge \ldots \wedge \widehat{\tau_i dx_i} \wedge \ldots \wedge \tau_i dx_n)(\tau_i v_1, \ldots, \widehat{\tau_i v_i}, \ldots, \tau_i v_n). \tag{49.341}$$

Thus we have

$$\begin{aligned}
&(dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n)(v_1, \ldots, \widehat{v_i}, \ldots, v_n) && \text{(49.342a)}\\
&\quad = (\tau_i dx_1 \wedge \ldots \wedge \widehat{\tau_i dx_i} \wedge \ldots \wedge \tau_i dx_n)(\tau_i v_1, \ldots, \widehat{\tau_i v_i}, \ldots, \tau_i v_n) && \text{(49.342b)}\\
&\quad = (dy_1 \wedge \ldots \wedge dy_{n-1})(\tau_i v_1, \ldots, \widehat{\tau_i v_i}, \ldots, \tau_i v_n) && \text{(49.342c)}\\
&\quad = \det\big(\tau_i v_1, \ldots, \widehat{\tau_i v_i}, \ldots, \tau_i v_n\big) && \text{(49.342d)}
\end{aligned}$$

SUBEQooQGSKooSgfxJk

The last equality is because (49.342c) is is a $(n-1)$-form applied to $n-1$ vectors in $\mathbb{R}^{n-1}$ and so is the determinant. □

49.11.2 Differential of k-forms

The differential of a k-form is defined by the following theorem.

Theorem 49.146.
Let M be a differentiable manifold. Then for each $k \in \mathbb{N}$, there exists an unique map

$$d \colon \Omega^k(M) \to \Omega^{k+1}(M)$$

such that

(1) d is linear,
(2) for $k = 0$, we find back the $d \colon C^\infty(M) \to \Omega^1(M)$ previously defined,
(3) if f is a function and ω^k a k-form, then

$$d(f\omega^k) = df \wedge \omega^k + f d\omega^k, \tag{49.343}$$

(4) $d(\omega^k \wedge \eta^l) = d\omega^k \wedge \eta^l + (-1)^k \omega^k \wedge d\eta^l$,
(5) $d \circ d = 0$.

An explicit expression for $d\omega^k$ is actually given by

$$d\omega^k = \sum d\omega_I \wedge dx^I \tag{49.344}$$

if $\omega^k = \sum \omega_I dx^I$. An useful other way to write it is the following. If ω is a k-form and $X_1, \ldots, X_{p+1}$ some vector fields,

$$\begin{aligned}
(k+1)d\omega(X_1, \ldots, X_{p+1}) = &\sum_{i=1}^{p+1} (-1)^{i+1} X_i \omega(X_1, \ldots \hat{X}_i, \ldots, X_{p+1}) \\
&+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \hat{X}_j, \ldots, X_{p+1}).
\end{aligned} \tag{49.345}$$

eq:formule_domega

Let us show it with $p = 1$. Let $\omega = \omega_i dx^i$ and compute $d\omega(X,Y) = \partial_i\omega_j(dx^i \wedge dx^j)(X,Y)$. For this, we have to keep in mind that the ∂_i acts only on ω_j while, in equation (49.345), a term $X\omega(Y)$ means –pointwise– the action of X on the function $\omega(Y)\colon M \to \mathbb{R}$. So we have to use Leibniz formula:

$$(\partial_i\omega_j)X^iY^j = (X\omega_j)Y^j = X(\omega_jY^j) - \omega_jXY^j.$$

On the other hand, we know that $[X,Y]^i = XY^i - YX^i$, so

$$d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]). \tag{49.346}$$

49.11.2.1 Hodge dual operator

Let us take a manifold M endowed with a metric g. We can define a map $r\colon T_x^*M \to T_xM$ by, for $\alpha \in T_x^*M$,

$$\langle r(\alpha), v\rangle = \alpha(v).$$

for all $v \in T_xM$, where $\langle\ \cdot\ ,\ \cdot\ \rangle$ stands for the product given by the metric g. If we have $\alpha, \beta \in T_x^*M$, we can define

$$\langle \alpha, \beta\rangle = \langle r(\alpha), r(\beta)\rangle.$$

With this, we define an inner product on $\Lambda^p(T_x^*M)$:

$$\langle \alpha_1 \wedge \dots \alpha_p, \beta_1 \wedge \dots \beta_p\rangle = \det_{ij}\langle \alpha_i, \beta_j\rangle.$$

DEFooUOJQooSzKjNR

Definition 49.147.
*The **Hodge operator** is* $\star\colon \Lambda^p(T_x^*M) \to \Lambda^{n-p}(T_x^*M)$ *such that for any* $\phi \in \Lambda^p(T_x^*M)$,

$$\phi \wedge (\star\psi) = \langle \phi, \psi\rangle\Omega = \langle \phi, \psi\rangle\sqrt{|\det(g)|}dx^1 \wedge \dots \wedge dx^n. \tag{49.347}$$

EXooCIYIooFPMLMU

Exemple 49.148.
We consider $\mathbb{R}^n$ with the euclidian metric. If $\sigma = dx_j$, then we expect $\star\sigma$ to be $s dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n$ for a certain factor s to be fixed (something like $(-1)^j$).

For every 1-form ϕ we need $\phi \wedge (\star\sigma) = \langle \phi, \sigma\rangle dx_1 \wedge \dots \wedge dx_n$. A basis of $\bigwedge^1(TM)$ is $\{dx_k\}_{k=1,\dots,n}$, so we test on dx_k.

First we have

$$\langle dx_k, dx_j\rangle = \langle e_k, e_j\rangle = \delta_{kj}. \tag{49.348}$$

Then

$$s\, dx_k \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n = s\delta_{kj}(-1)^{j+1}dx_1 \wedge \dots \wedge dx_n. \tag{49.349}$$

Thus we need $s = (-1)^{j+1}$ and we have

$$\star dx_j = (-1)^{j+1}dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n. \tag{49.350}$$

△

49.11.3 Musical isomorphism

subsec_musique

In some literature, we find the symbols $v^\flat$ and $\alpha^\sharp$. What does it mean ? For $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^2(M)$, the **flat** operation $v^\flat \in \Omega^1(M)$ is simply defined by the inner product:

$$v^\flat = i(v)\omega \tag{49.351}$$

EQooBTWXooTooNxa

In the same way, we define the **sharp** operation by taking a 1-form α and defining $\alpha^\sharp$ by

$$i(\alpha^\sharp)\omega = \alpha. \tag{49.352}$$

An immediate property is, for all $v \in \mathfrak{X}(M)$, $v^{\flat\sharp} = v$, and for all $\alpha \in \Omega^1(M)$, $\alpha^{\sharp\flat} = \omega$.

49.11.4 Pull-back and push-forward

49.149.
Let $\varphi\colon M \to N$ be a smooth map, α a k-form on N, and Y a vector field on N. Consider the map $d\varphi\colon T_xM \to T_{\varphi(x)}M$. The aim is to extend it to a map from the tensor algebra [55] of T_xM to the one of $T_{\varphi(x)}M$.

The **pull-back** of φ on a k-form α is the map

$$\varphi^*\colon \Omega^k(N) \to \Omega^k(M)$$

defined by

$$(\varphi^*\alpha)_m(v_1,\dots,v_k) = \alpha_{\varphi(m)}(d\varphi_m v_1,\dots,d\varphi_m v_k) \tag{49.353}$$

for all $m \in M$ and $v_i \in \mathfrak{X}(M)$.

Note the particular case $k = 0$. In this case, we take –instead of α– a function $f\colon N \to \mathbb{R}$ and the definition (49.353) gives $\varphi^*f\colon M \to \mathbb{R}$ by

$$\varphi^*f = f \circ \varphi.$$

The **push-forward** of φ on a k-form is the map

$$\varphi_*\colon \Omega^k(M) \to \Omega^k(N)$$

defined by $\varphi_* = (\varphi^{-1})^*$. For $v \in T_nN$, we explicitly have:

$$(\varphi_*\alpha)_n(v) = \alpha_{\varphi^{-1}(n)}(d\varphi_n^{-1}v).$$

Let now $\varphi\colon M \to N$ be a diffeomorphism. The **pull-back** of φ on a vector field is the map

$$\varphi^*\colon \mathfrak{X}(N) \to \mathfrak{X}(M)$$

defined by

$$(\varphi^*Y)(m) = [(d\varphi^{-1})_m \circ Y \circ \varphi](m),$$

or

$$(\varphi^*Y)_{\varphi^{-1}(n)} = (d\varphi^{-1})_n Y_n,$$

for all $n \in N$ and $m \in M$. Notice that

$$(d\varphi^{-1})_n\colon T_nN \to T_{\varphi^{-1}(n)}M,$$

and that $\varphi^{-1}(n)$ is well defined because φ is an homeomorphism.

The **push-forward** is, as before, defined by $\varphi_* = (\varphi^{-1})^*$. In order to show how to manipulate these notations, let us prove the following equation:

$$f_{*\xi} = (df)_\xi.$$

For $\varphi\colon M \to N$ and Y in $\mathfrak{X}(N)$, we just defined $\varphi^*\colon \mathfrak{X}(N) \to \mathfrak{X}(M)$, by

$$(\varphi^*Y)_{\varphi^{-1}(n)} = (d\varphi^{-1})_n Y_n. \tag{49.354}$$

Take $f\colon M \to N$; we want to compute $f_* = (f^{-1})^*$ with $(f^{-1})^*\colon \mathfrak{X}(M) \to \mathfrak{X}(N)$. Replacing the "$-1$" on the right places, the definition (49.354) gives us

$$\Big[(f^{-1})^*X\Big]_{f(m)} = (df)_m X_m,$$

if $X \in \mathfrak{X}(M)$, and $m \in M$.

We can rewrite it without any indices: the coherence of the spaces automatically impose the indices: $(f^{-1})^*X = (df)X$. It can also be rewritten as $(f^{-1})^* = df$, and thus $f_* = df$. From there to $f_{*\xi} = (df)_\xi$, it is straightforward.

55. Definition 49.138

49.12 Lie derivative

Consider $X \in \mathfrak{X}(M)$ and $\alpha \in \Omega^p(M)$. Let $\varphi_t \colon M \to M$ be the flow of X. The **Lie derivative** of α is

$$\mathcal{L}_X \alpha = \lim_{t\to 0} \frac{1}{t}[(\varphi_t^* \alpha) - \alpha] = \frac{d}{dt}\varphi_t^* \alpha \Big|_{t=0}. \tag{49.355}$$ liesurforme

More explicitly, for $x \in M$ and $v \in T_x M$,

$$(\mathcal{L}_X \alpha)_x(v) = \lim_{t\to 0} \frac{1}{t}\left[(\varphi_t^* \alpha)_x(v) - \alpha_x(v)\right]$$

In the definition of the **Lie derivative** for a vector field, we need an extra minus sign:

$$(\mathcal{L}_X Y)_x = \frac{d}{dt}\varphi_{-t*} Y_{\varphi_t(x)} \Big|_{t=0}. \tag{49.356}$$ EqDefLieDerivativeVect

Why a minus sign ? Because $Y_{\varphi_t(x)} \in T_{\varphi_t(x)}M$, but $(d\varphi_{-t})_a \colon T_a M \to T_{\varphi_{-t}(a)}M$ so that, if we want, $\varphi_{-t*} Y_{\varphi_t(x)}$ to be a vector at x, we can't use φ_{t*}.

These two definitions can be embedded in only one. Let $X \in \mathfrak{X}(M)$ and φ_t its integral curve [56]. We know that φ_{t*} is an isomorphism $\varphi_{t*} \colon T_{\varphi^{-1}(x)}M \to T_x M$. It can be extended to an isomorphism of the tensor algebras at $\varphi^{-1}(x)$ and x. We note it $\tilde{\varphi}_t$. For all tensor field K on M, we define

$$(\mathcal{L}_X K)_x = \lim_{t\to 0}[K_x - (\tilde{\varphi}_t K)_x].$$

On a Riemannian manifold (M, g), a vector field X is a **Killing vector field** if $\mathcal{L}_X g = 0$.

Lemma 49.150.
Let $f \colon (-\epsilon, \epsilon) \times M \to \mathbb{R}$ be a differentiable map with $f(0, p) = 0$ for all $p \in U$. Then there exists $g \colon (-\epsilon, \epsilon) \times M \to \mathbb{R}$, a differentiable map such that $f(t, p) = tg(t, p)$ and

$$g(0, q) = \frac{\partial f(t, q)}{\partial t}\Big|_{t=0}.$$

Proof. Take

$$g(t, q) = \int_0^1 \frac{\partial f(ts, p)}{\partial (ts)} ds,$$

and use the change of variable $s \to ts$. □

Lemma 49.151.
If φ_t is the integral curve of X, for all function $f \colon M \to \mathbb{R}$, there exists a map g, $g_t(p) = g(t, p)$ such that $f \circ \varphi_t = f + tg_t$ and $g_0 = Xf$.

Proof. Consider $\overline{f}(t, p) = f(\varphi_t(p)) - f(p)$, and apply the lemma:

$$f \circ \varphi_t = tg_t(p) + f(p).$$

Thus we have

$$Xf = \lim_{t\to 0} \frac{1}{t}[f(\varphi_t(p)) - f(p)] = \lim_{t\to 0} g_t(p) = g_0(p).$$

□

One of the main properties of the Lie derivative is the following:

56. *i.e.* for all $x \in M$, $\varphi_0(x) = x$ and $\frac{d}{dt}\varphi_{u+t}(x)\big|_{t=0} = X_{\varphi_u(x)}$.

ThoLieDerrComm

Theorem 49.152.
Let X, $Y \in \mathfrak{X}(M)$ and φ_t be the integral curve of X. Then

$$[X,Y]_p = \lim_{t\to 0} \frac{1}{t}[Y - d\varphi_t Y](\varphi_t(p)),$$

or

$$\mathcal{L}_X Y = [X,Y]. \tag{49.357}$$

where the commutator is given by the definition 49.105.

Proof. Take $f\colon M \to \mathbb{R}$ and the function given by the lemma: $g_t\colon M \to \mathbb{R}$ such that $f\circ\varphi_t = f + tg_t$ and $g_0 = Xf$. Then put $p(t) = \varphi_t^{-1}(p)$. The rest of the proof is a computation:

$$(\varphi_{t*}Y)_p f = Y(f\circ\varphi_t)_{p(t)} = (Yf)_{p(t)} + t(Yg_t)_{p(t)},$$

so

$$\begin{aligned}\lim_{t\to 0}\frac{1}{t}[Y_p - (\varphi_{t*}Y)_p]f &= \lim_{t\to 0}\frac{1}{t}[(Yf)_p - (Yf)_{p(t)}] - \lim_{t\to 0}(Yg_t)_{p(t)}\\ &= X_p(Yf) - Y_p g_0\\ &= [X,Y]_p f.\end{aligned} \tag{49.358}$$

□

A second important property is

Theorem 49.153.
For any function $f\colon M \to V$,

$$\mathcal{L}_X f = Xf.$$

Proof. If $X(t)$ is the path which defines the vector X, it is obvious that at $t = 0$, $X(t)$ is an integral curve to X, so that we can take $X(t)$ instead of φ_t in (49.355). Therefore we have:

$$\mathcal{L}_X f = \frac{d}{dt}\varphi_t^* f\Big|_{t=0} = Xf \tag{49.359}$$

by definition of the action of a vector on a function. □

49.12.1 Notions of topology

Let X be a set. A **topology** on X is a choice of a set τ of parts of X such that

— $X \in \tau$, $\varnothing \in \tau$,
— $\bigcap_{\text{finite}} A_i \in \tau$ if $A_i \in \tau$,
— $\bigcup A_i \in \tau$ if $A_i \in \tau$.

Elements of τ are said **open** in X for the topology τ. Let us point out the difference between the second and the third point: the third point allows infinite unions. If we consider $A_i =]0, 1 + \frac{1}{i}[\subset \mathbb{R}$, we have

$$\bigcap_{i\in\mathbb{N}_0} A_i =]0,1]$$

which is not open in $\mathbb{R}$.

If $E \subset X$, the **induced** topology on E by X is the one for which a subset A of E is open in E if and only if there exists a subset $\mathcal{U}$ of X which is open in X and such that $A = E \cap \mathcal{U}$. In other words, $A \in \tau_E$ if and only if $A = E \cap \mathcal{U}$ for a certain $\mathcal{U} \in \tau_X$.

When X and Y are topological spaces, we say that $f\colon X \to Y$ is **continuous** if for all $A \in \tau_Y$, $f^{-1}(A) \in \tau_X$, that is, if the inverse image of any open set is open. Let us see a non trivial example.

We consider X as $\mathbb{R}$ endowed with the usual topology and Y as $\mathbb{R}$ with the topology $\tau_Y = \{\varnothing, \mathbb{R}\}$. We have $X = Y$ as sets, but $X \neq Y$ as topological spaces. The identity map $\mathrm{Id}\colon X \to Y$ is continuous because $\mathrm{Id}^{-1}(\varnothing) = \varnothing \in \tau_X$ and $\mathrm{Id}^{-1}(\mathbb{R}) = \mathbb{R} \in \tau_X$. But the identity $\tilde{\mathrm{Id}}\colon Y \to X$ is not continuous because $\tilde{\mathrm{Id}}^{-1}(]0,1[) =]0,1[$ which is not open in Y.

49.12.2 Differential, path and extensions

Let's take $X_0 \in T_{x_0}C$, a map $f\colon C \to C'$ and an extension $\hat{f}$ of f on a neighbourhood V around x_0. If one consider the path $\eta\colon \mathbb{R} \to V$, $\eta(t) = x_0 + tX_0$, we know that

$$\begin{aligned} d\hat{f}_{x_0}\big(\eta'(0)\big) &= \frac{d}{dt}\Big[\hat{f}\big(\eta(t)\big)\Big]_{t=0} \\ &= \frac{\partial \hat{f}}{\partial x^i}\frac{d\eta^i}{dt}. \end{aligned}$$

On the latter expression, we see that if $\eta(0) = \gamma(0) = x_0$ and $\gamma'(0) = \eta'(0)$, then $df_{x_0}\big(\gamma'(0)\big) = df_{x_0}\big(\eta'(0)\big)$. In particular, we can chose a path γ such that $\gamma(t) \in c$ for all t. For this path, we can put $f\big(\gamma(t)\big)$ in the derivative instead of $\hat{f}\big(\gamma(t)\big)$.

Hence we can consider the differential of a map $f\colon C \to C'$ as an object intrinsically defined on C without extensions. So we set $df_{x_0}\colon T_{x_0}C \to T_{f(x_0)}C$,

$$df_{x_0}(\eta'(0)) = \frac{d}{dt}\Big[(f \circ \eta)(t)\Big]_{t=0}.$$

49.12.3 Vector fields

A *tangent vector field* along an embedded curve C in $\mathbb{R}^2$ is a rule which assigns to every point $x \in C$ a vector X_x tangent to C at x. Moreover, one requires the dependence on x "to be smooth".

More precisely, let $T(C)$ denote the set of all vectors tangent to C:

$$T(C) = \bigcup_{x \in C} T_x(C) \tag{49.360}$$

eq_TC_biguu

where the union is a *disjoint* union. A **tangent vector field** on C is a map $X\colon C \to T(C)$, $x \mapsto X_x$ such that

(1) for all x, one has

$$X_x \in T_x(C), \tag{49.361}$$

eq_XxTxC

(2) the map $C \to \mathbb{R}^2 \times \mathbb{R}^2$,

$$x \mapsto (x, X_x) \tag{49.362}$$

eq_xmapstoXx

is smooth.

The latter is a non trivial point since C is not open in $\mathbb{R}^2$.

Exercise 1 Some students noticed that equation (49.362) was a wrong because of (49.361): firstly T_xC is vector space of dimension one (so X_x cannot belong to a $\mathbb{R}^2$) and secondly, the union (49.360) being a disjoint union, if x and y are different points of C, X_x and X_y belongs to two different spaces: T_xC and T_yC.

Try to understand the problem and to find a solution.

49.12.4 Cohomology

When C is a curve in $\mathbb{R}^2$, the **cohomology** of C is the quotient

$$H^1(C) = \frac{\Omega^1(C)}{dC^\infty(C)}$$

of differential forms on C by forms which are differential of some functions. When $\alpha \in \Omega^1(C)$,

$$[\omega] = \{\omega + df \,|\, f \in C^\infty(C)\}.$$

We are going to study two examples: $H^1(\mathbb{R})$ and $H^1(S^1)$.

49.12.4.1 Cohomology of $\mathbb{R}$

We begin by defining the form dt on $\mathbb{R}$ by the following action of a tangent vector $\dot\gamma(t) = v \in \mathbb{R}$: $dt(\dot\gamma(t)) = v$. Each form on $\mathbb{R}$ can be written as

$$\alpha = a\,dt$$

where a is a smooth function on $\mathbb{R}$. Why? Because for each $x \in \mathbb{R}$, $T^*_x\mathbb{R}$ has dimension one. We are going to prove the following:

Theorem 49.154.
The cohomology of $\mathbb{R}$ *is zero:*

$$H^1(\mathbb{R}) = 0.$$

In other words, each form α on $\mathbb{R}$, reads df for a suitable function $f \in C^\infty(\mathbb{R})$. For the proof, we explicitly give a function f such that $\alpha = df$ for any $\alpha \in \Omega^1(\mathbb{R})$.

Let $\alpha \in \Omega^1(\mathbb{R})$ and let us write it as $\alpha_x = a(x)\,dt \in T^*_x(\mathbb{R})$. We pose

$$f(x) = \int_0^x a(u)\,du$$

and we compute that $df = \alpha$. Let γ be a path in $\mathbb{R}$,

$$\begin{aligned} df_{\gamma(0)}\dot\gamma(0) &= \frac{d}{dt}\Big[(f\circ\gamma)(t)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\int_0^{\gamma(t)} a(u)\,du\Big]_{t=0} \\ &= a(\gamma(0))\gamma'(0). \end{aligned} \tag{49.363}$$

On the other hand,

$$\begin{aligned} \alpha_{\gamma(0)}\dot\gamma(0) &= a(x)\big(dt\dot\gamma(0)\big) \\ &= a(\gamma(0))\dot\gamma(0); \end{aligned} \tag{49.364}$$

this is the same!

49.12.4.2 Cohomology of S^1

In the whole part about S^1, γ will denote the path

$$\gamma(t) = e^{it},$$

and σ will denote a parametrization of S^1 used in integrals.

We begin to show that $H^1(S^1) \neq 0$: there exists some 1-forms on S^1 that cannot be written under the form df. For this, we begin to prove that

prop_intdfz

Proposition 49.155.
For all $f \in C^\infty(S^1)$,

$$\int_{S^1} df = 0.$$

Proof. If $\sigma\colon [0,2\pi] \to S^1$ is a parametrization of S^1, we have

$$\begin{aligned} \int_{S^1} df &= \int_{[0,2\pi]} df_{\sigma(t)}\dot\sigma(t)\,dt \\ &= \int_0^{2\pi} \frac{d}{du}\Big[(f\circ\sigma)(u)\Big]_{u=t}\,dt \\ &= (f\circ\sigma)(2\pi) - (f\circ\sigma)(0) \\ &= 0. \end{aligned} \tag{49.365}$$

□

We define the differential 1-form α by

$$\alpha(\dot{\gamma}(t)) = 1.$$

So

$$\int_{s^1} \alpha = \int_0^{2\pi} \alpha_{\sigma(t)}(\dot{\sigma}(t))\, dt = 2\pi \neq 0.$$

Then this α is not an exact form.

There exists an alternative way to define this form. We consider the function $\theta\colon S^1 \to \mathbb{R}$ given by

$$\theta(e^{i\eta}) = \eta. \tag{49.366}$$

eq_deftheta

This is the function which gives the angular coordinate of a point in S^1. As a map from S^1 to $\mathbb{R}$, one can compute its differential applied on a tangent vector; for example on $\dot{\gamma}(t)$:

$$d\theta\big(\dot{\gamma}(t)\big) = \frac{d}{dt}\Big[\theta(e^{it})\Big]_{t=0} = \frac{d}{dt}\Big[t\Big]_{t=0} = 1,$$

so this is the previously defined form α. By the way, it proves that this form is a smooth differential form.

We have proved that $d\theta$ is the differential of the function θ and that it is not an exact form (because its integral is not zero). It is a problem. In fact, equation (49.366) doesn't *globally* defines a function on S^1 because $\theta(e^{i\,\cdot\,0}) = 0$ but $\theta(e^{i\,\cdot\,2\pi}) = 2\pi$. In other words, $\theta(1)$ is not well defined. Then we can only define the function θ on a local chart. For example on the half circle, we define $\theta\colon S^1 \to]-\frac{\pi}{2}, \frac{\pi}{2}[$ by $\theta(e^{i\sigma}) = \sigma$. This is a well defined function.

The following theorem is the main part of the proof that $H^1(S^1) = \mathbb{R}$.

Theorem 49.156.
A form $\alpha \in \Omega^1(S^1)$ fulfil

$$\int_{S^1} \alpha = 0$$

if and only if there exists a function $f\colon S^1 \to \mathbb{R}$ such that $df = \alpha$.

Proposition 49.155 is the first half of the proof. Let's see the second one

Proof. Let $\alpha \in \Omega^1(S^1)$ such that $\int_{S^1} \alpha = 0$. It can be written under the form

$$\alpha_{e^{i\theta}} = a(e^{i\theta})\, d\theta$$

and we define $f\colon S^1 \to \mathbb{R}$ by

$$f(e^{i\theta}) = \int_0^{\theta} a(e^{i\sigma})\, d\sigma.$$

Since $\alpha_{e^{i\theta}}\big(\dot{\gamma}(\sigma)\big) = a(e^{i\theta})\, d\theta\big(\dot{\gamma}(\sigma)\big) = a(e^{i\theta})$, in fact f is

$$f(e^{i\theta}) = \int_0^{\theta} \alpha. \tag{49.367}$$

It is important to remark that this function has not the problem of definition of the previously θ because $\int_{S^1} \alpha = 0$.

Now we want to prove that $df = \alpha$. For this we consider a path $c(t) = e^{i\theta(t)}$ in S^1 and we compute $df_{e^{i\theta(0)}}c'(0)$. We have:

$$\begin{aligned} df_{e^{i\theta(0)}}c'(0) &= \frac{d}{dt}\Big[(f \circ c)(t)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\int_0^{\theta(t)} a(e^{i\sigma})\, d\sigma\Big]_{t=0} \\ &= a(e^{i\theta(0)})\theta'(0). \end{aligned} \tag{49.368}$$

On the other hand,

$$\begin{aligned}\alpha_{e^{i\theta(0)}c'(0)} &= a(e^{i\theta(0)})\,d\theta\big(c'(0)\big)\\ &= a(e^{i\theta(0)})\frac{d}{dt}\Big[\theta(e^{i\theta(t)})\Big]_{t=0}\\ &= a(e^{i\theta(0)})\frac{d}{dt}\Big[\theta(t)\Big]_{t=0}\\ &= a(e^{i\theta(0)})\theta'(0).\end{aligned} \tag{49.369}$$

This concludes the proof.

□

It is now easy to see that

$$I : [\alpha] \mapsto \int_{S^1} \alpha$$

defines a vector space isomorphism between $H^1(S^1)$ and $\mathbb{R}$. We check that I is well defined: if $[\alpha] = [\beta]$, we must have $I([\alpha]) = I([\beta])$. It is correct because $[\alpha] = [\beta]$ means that there exists a function f such that $\beta = \alpha + df$. Then

$$I([\beta]) = \int_{S^1} \beta = \int_{S^1} (\alpha + df) = \int_{S^1} \alpha = I([\alpha]).$$

49.12.5 Integral curve

A map $c\colon \mathbb{R} \to \mathbb{R}^n$ is an **integral curve** of the vector field $X \in \mathfrak{X}(\mathbb{R}^n)$ if for all t (in the domain of c),

$$c'(t) = X_{c(t)}. \tag{49.370}$$ eq_defintcurv

As a vector field, X is a map $X\colon \mathbb{R}^n \to \mathbb{R}^n$, so equation (49.370) reads

$$c'(t) = X\big(c(t)\big).$$

Now, "to be an integral curve" means "to satisfy a differential equation". You know some existence and unicity theorems; for example this one:

tho_exuniintcur

Theorem 49.157.
Let $u\colon I \to \mathbb{R}$ and $f\colon J \to \mathbb{R}$ where J is open in $\mathbb{R}$ and I has non empty interior. We consider the following differential equation:

$$y'(t) = u(t)f(y(t)). \tag{49.371}$$ eq:eqdiff

Suppose that u is continuous on I and f is continuous on J with $f(\eta) \neq 0$ for all $\eta \in J$. Let $t_0 \in I$ and $y_0 \in J$. Then there exists an interval $I' \subset I$ with non empty interior with $t_0 \in I'$ and a C^1 function $y\colon I' \to J$ such that

(1) y is solution of (49.371) on I' and fulfils the initial condition $y(t_0) = y_0$,

(2) if z is a solution of (49.371) in an interval $I'' \subset I$ with $t_0 \in I''$ and $z(t_0) = y_0$, then $I'' \subset I'$ and $z(t) = y(t)$ for $t \in I''$.

When one says that an embedded curve C is an integral curve of a vector field X, one means that there exists a system of charts $\varphi_\alpha\colon \mathcal{U}_\alpha \to C$ on this curve which solve condition (49.370) as map $\varphi_\alpha\colon \mathcal{U}_\alpha \to \mathbb{R}^n$. In particular, X must be tangent to C on each point.

49.12.6 Christoffel symbols

When $f : \mathcal{U} \to \mathbb{R}^3$ is a parametrization of a surface S in $\mathbb{R}^3$, there exists two interesting tangent vectors fields:

$$e_x|_{x_0,y_0} = \frac{d}{dt}\Big[f(x_0+t,y_0)\Big]_{t=0} \text{ and } e_y|_{(x_0,y_0)} = \frac{d}{dt}\Big[f(x_0,y_0+t)\Big]_{t=0}. \tag{49.372}$$

eq_basvecto

They form a basis of $T_{f(x_0,y_0)}S$. These vectors are often called ∂_i. You can ask yourself why this partial derivative notation.

The problem of finding the first fundamental form (see exercise 13) reduce to computes products $e_1 \cdot e_j$. These products form the matrix g:

$$g_{ij} := e_i \cdot e_j.$$

Now, there are no reasons for $\partial_i e_j$ to be a tangent vector. In fact, it is even difficult to *define* the object $\partial_i e_j$ because it should be something like

$$\partial_i e_j = \lim_{t\to 0} \frac{e_j(x+t) - e_j(x)}{t}.$$

But $e_j(x)$ and $e_j(x+t)$ are not vectors on the same points, therefore the difference on the right hand side is not well defined as difference in TS. This difference is however well defined in the ambient space $\mathbb{R}^3$. So in order to define the derivative of a tangent vector field, we have to consider the surface as a part of $\mathbb{R}^3$. In other words, we have to extend the coordinate system $(x,y) \to S$ into a coordinate system $(x,y,z) \to \mathbb{R}^3$ which reduce to original coordinates when $z = z_0$ and such that the vector e_z is normal.

As an example, polar coordinates on the (unit) sphere in $\mathbb{R}^2$ are given by

$$(\varphi,\theta) \to (\cos\varphi\sin\theta, \cos\varphi\cos\theta, \sin\varphi)$$

and an extension of the system is the full spherical coordinates on $\mathbb{R}^3$:

$$(\varphi,\theta,r) \to (r\cos\varphi\sin\theta, r\cos\varphi\cos\theta, r\sin\varphi);$$

it is easy to see that e_r is normal. Now, our new set of e_i's form a basis of the whole $\mathbb{R}^3$, then we define the **Christoffel symbols** Γ_{ij}^k of the new coordinates system

$$\partial_i e_j = \sum_k \Gamma_{ij}^k e_k. \tag{49.373}$$

eq_def_sch

In this formula, k is summed over *all* coordinates, including the normal one.

There exists a funny formula to explicitly compute these Christoffel symbols. We begin by taking the derivative ∂_k of $g_{ij} = e_i \cdot e_j$, where k, i and j are tangent coordinates. So we do not take care about the normal coordinate. Remark that in formula

$$\partial_k g_{ij} = \Gamma_{ki}^l e_l \cdot e_j + e_i \cdot \Gamma_{kj}^l e_l,$$

the summed index l take *a priori* all values including the normal coordinate, but the scalar product with e_j and e_i kill the normal component. So in the sequel we are left with only tangent coordinates. Replacing the products $e_l \cdot e_j$ and $e_i \cdot e_l$ by the corresponding elements of the matrix g,

$$\partial_k g_{ij} = \Gamma_{ki}^l g_{lk} + \Gamma_{kj}^l g_{il}.$$

Computing the combination $\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}$ and taking the symmetry property $\Gamma_{ij}^k = \Gamma_{ji}^k$, $g_{ij} = g_{ji}$ into account, we find the relation

$$2\Gamma_{ki}^l g_{lj} = \partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}.$$

If we write $g^{ab} := (g^{-1})_{ab}$, we can rewrite it as

$$\Gamma_{ij}^k = \frac{1}{2} g^{lk}\big(\partial_i g_{jl} + \partial_j g_{li} - \partial_l g_{ij}\big). \tag{49.374}$$

eq_Gamaform

This is the promised formula. Now we are able to write down elements Γ_{ij}^k by simply derive and invert the matrix g.

In formula (49.374), indices (including the summed one) only take values in tangent coordinates. The normal one is completely forgotten.

Information Here is the definition given in the oral course. Let $r : U \subset \mathbb{R}^3 \to \Sigma \subset \mathbb{R}^3$ be a chart on an embedded surface Σ. Set $r_i := \partial_{x_i} r$; $N := \frac{1}{\|r_1 \times r_2\|} r_1 \times r_2 \quad (i = 1, 2)$. Then the **Christoffel symbols** Γ_{ij}^k are defined by the relations:

$$\partial_{x_i}(r_j) \ := \ \sum_k \Gamma_{ij}^k r_k \ + \ \nu_{ij} N \ .$$

Do you understand why the latter formula has two terms while formula (49.373) has only one? Is there *really* one term less in (49.373)?

49.12.7 Commutator of vector field

Let M be a surface and $X, Y \in \mathfrak{X}(M)$. We want to define $[X, Y] \in \mathfrak{X}(M)$, so for each $x \in M$ and $f \in C^\infty(M)$ we have to define $[X, Y]_x f$. We know that X_x is an element in $T_x M$ which acts on f by

$$X_x(f) = \frac{d}{dt}\Big[f\big(X_x(t)\big)\Big]_{t=0} = df_x X_x \quad \in \mathbb{R},$$

so one can consider Xf as a function, $(Xf)(x) = X_x f$. We can apply the vector Y_x to this function and get a real:

$$\begin{aligned} Y_x(Xf) &= \frac{d}{dt}\Big[(Xf)(Y_x(t))\Big]_{t=0} \\ &= \frac{d}{dt}\Big[f\big(X_{Y_x(t)}(u)\big)\Big]_{t=0} s0. \end{aligned}$$

This allows us to define $[X, Y]_x f = X_x(Yf) - Y_x(Xf)$. If, in a local coordinate system, $X = X^i \partial_i$ and $Y = Y^j \partial_j$, we have

$$\begin{aligned} (XY)f &= X^i \partial_i (Y^j \partial_j f) \\ &= X^i (\partial_i Y^j) \partial_j f + X^i Y^j \partial_{ij}^2 f. \end{aligned}$$

The same computation with $X \leftrightarrow Y$ shows that the second derivative disappears in the commutator $[X, Y]$. So $[X, Y]$ is a combination of the e_i's and therefore is a vector fields.

The set of vector fields endowed with the bracket $[.,.]$ is an algebra. It is even a *Lie* algebra because the bracket

(1) is bilinear,

(2) is skew-symmetric: $[X, Y] = -[Y, X]$,

(3) fulfills the Jacobi identity: $[[X, Y], Z] + [[Y, Z], Z] + [[Z, X], Y] = 0$,

for all X, Y, $Z \in \mathfrak{X}(M)$.

49.13 Exercises

Exercise 2 exo002

Find a smooth parametrization of the set

$$E := \{(t, |t|)\}_{t \in]-1,1[}.$$

Is E an embedded curve in $\mathbb{R}^2$?

A **parametrization** is like a chart, but without the regularity condition. corr002

Correction of the exercise 2

The parametrization given in the question is not smooth because the derivative is discontinuous at $t = 0$. Remark that tangent vectors in this parametrization must be parallel to $(1, -1)$ in the left part and to $(1, 1)$ in the right part. The only way for a function (namely the vertical part of the derivative of the parametrization) to change sign without discontinuity is to vanish. One

thus needs a parametrization whose derivative vanishes at $(0,0)$. This already answers the second question: this is not an embedded curve.

Consider the parametrization $\varphi(t) = \big(t^2, x(t)\big)$ where

$$x(t) = \begin{cases} -t^2 & \text{when } t \in]-1,0] \\ t^2 & \text{when } t \in]0,1[. \end{cases}$$

This parametrization doesn't work because the second derivative has a discontinuity at $t=0$. Try to convince yourself that a correct parametrization is given by

$$\varphi(t) = \begin{cases} (1,-1)f(t) & \text{if } t \in]-1,0] \\ (1,1)f(t) & \text{if } t \in]0,1[\end{cases}$$

where $f\colon \mathbb{R} \to \mathbb{R}$ is a function which gives 1 at 1 and -1 and vanishes at zero with all its derivatives. Such a function can be built from Cauchy's regularization function $\rho(x) = e^{-1/x^2}$.

The annihilation of *all* derivatives of the parametrization at $(0,0)$ is required by the fact that on the left, a derivative of any order must be a multiple of $(1,-1)$ while on the right it must be a multiple of $(1,1)$.

Exercise 3 exo001

Is the following set E an embedded curve in $\mathbb{R}^2$?

$$E := \{(t, \sin(\frac{1}{t}))\}_{t\in]0,1[} \cup \{(0,s)\}_{s\in[-1,1]}.$$

corr001

Correction of the exercise 3

Let us denote $E_1 = \{(t, \sin(\frac{1}{t}))\}_{t\in]0,1[}$ (figure 49.1) and $E_2 = \{(0,s)\}_{s\in[-1,1]}$. If one wants E to be an embedded curve in $\mathbb{R}^2$, one has to find maps $\varphi_\alpha\colon \mathcal{U}_\alpha \to E$ such that the famous three conditions hold.

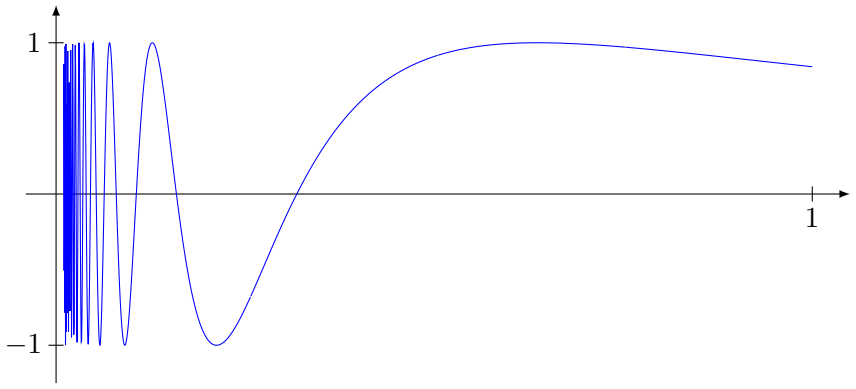


Figure 49.1: The graph of the function $x \mapsto \sin(1/x)$. LabelFigTZCISko

One of these maps must contain the point $a = (0,-1)$; let's say $\varphi(1) = a$. Consider the open set $B(a,r)$ of radius $r < \frac{1}{2}$ around a. If φ is continuous, there exists a δ such that $\epsilon \leqslant \delta$ implies $\varphi(1+\epsilon) \in B(a,r)$.

It is possible to find an open neighbourhood B' of a which does not contain the oscillation on which b lies. Let $b' = \varphi(1+\epsilon') \in B'$. Since φ is continuous, the path $c\colon [\epsilon', \epsilon] \to C$ given by $c(t) = \varphi(1+t)$ is continuous (as path in $\mathbb{R}^2$). A point of the image of c has the form $\varphi(1+s)$ with $s \leqslant \epsilon \leqslant \delta$. So $c\big([\epsilon', \epsilon]\big) \subset B(a,r)$.

But the fact that c reaches b' from b which is not on the same oscillation forces at least one point of c to be on the top of an oscillation and then to get out from $B(a,r)$. This contradicts the continuity.

Exercise 4 exo004 Consider the set

$$E = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \theta \in \mathbb{R} \right\}.$$

(1) It this set an embedded curve in $\mathbb{R}^4$?

(2) Write a general tangent vector at $\theta = 0$. Is it surprising ? [enum004ii]

Now, fix a $v \in \mathbb{R}^2$, and let $f_v \colon E \to \mathbb{R}^2$ be defined by

$$f_v(\theta) = \theta v$$

with a slight abuse of notation between θ and the element of E which is defined by θ. You have to guess the product in the right hand side.

(1) Give an explicit form of $(df_v)_{\mathbb{1}}$ where $\mathbb{1}$ is the 2×2 unit matrix.

(2) Can you give a geometric interpretation as "infinitesimal" rotation ? Does it help to answer point (2) ?

For general culture, remark that E is a group for the matrix product and that the anti-symmetric matrices form an algebra over $\mathbb{R}$ for the product $A \cdot B := AB - BA$ where the dot denotes the algebra product and the product in the right hand side is the usual matrix product.

The elements of E depend smoothly on one parameter, turning E into what is called a **Lie group** of dimension 1. The set of antisymmetric matrices (which form the tangent space to E at the identity), endowed with the dot product, forms the **Lie algebra** associated with the Lie group E. Note that in this case, the dot product always gives zero. This is common to all Lie algebras of dimension 1.

More generally, consider a curve $x(t)$ in $O(n)$ (the group of orthogonal matrices) such that $x(0) = \mathbb{1}$. For each t, one has $x(t)x(t)^T = \mathbb{1}$. Taking the derivative of this equation with respect to t at $t = 0$, Leibniz rule yields $\dot{x}(0)x(0)^T + x(0)\dot{x}(0)^T = 0$, or, taking $x(0) = \mathbb{1}$ into account,

$$\dot{x}(0) + \dot{x}(0)^T = 0.$$

This proves that $\dot{x}(0)$ is an antisymmetric matrix. So the Lie algebra of the Lie group of orthogonal matrices is the algebra of antisymmetric matrices. When $n \geqslant 3$, the algebra product $AB - BA$ is no longer identically zero. [corr004]

Correction of the exercise 4

The natural mistake is to say "θ takes values in $[0, 2\pi[$ which is not an open interval. Then I cannot parametrize E ". One can use more than one local chart ! For instance it is possible to take

$$\begin{aligned} \mathcal{U}_1 &=]0, 2\pi[\,, \\ \mathcal{U}_2 &=]\frac{3\pi}{2}, \frac{5\pi}{2}[\end{aligned}$$

and in both cases

$$\varphi_i(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

In order to write a tangent vector at $\theta = 0$, we consider a general path in E given by a path $\theta \colon \mathbb{R} \to \mathbb{R}$ such that $\theta(0) = 0$. We have

$$\begin{aligned} \frac{d}{dt}\Big[\begin{pmatrix} \cos\theta(t) & \sin\theta(t) \\ -\sin\theta(t) & \cos\theta(t) \end{pmatrix} \Big]_{t=0} &= \begin{pmatrix} -\theta'(0)\sin\theta(0) & \theta'(0)\cos\theta(0) \\ -\theta'(0)\cos\theta(0) & -\theta'(0)\sin\theta(0) \end{pmatrix} \\ &= \theta'(0) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{49.375}$$

which is a general anti-symmetric matrix. Formally, you have to interpret it as a vector in $\mathbb{R}^4$, but matrix realization is fruitful to interpret the sequel.

Now the map $(df_v)_{\mathbb{1}}$ has to be applied to a tangent vector at point $\mathbb{1}$, so one computes (we set $v = (v_x, v_y)$)

$$\begin{aligned}(df_v)_{\mathbb{1}} \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} &= \frac{d}{dt}\Big[f_v \begin{pmatrix} \cos at & \sin at \\ -\sin at & \cos at \end{pmatrix} \Big]_{t=0} \\ &= \frac{d}{dt}\Big[\begin{pmatrix} v_x \cos at + v_y \sin at \\ -v_x \sin at + v_y \cos at \end{pmatrix} \Big]_{t=0} \\ &= a \begin{pmatrix} v_y \\ -v_x \end{pmatrix}.\end{aligned} \tag{49.376}$$

If you apply a "very little" rotation on the vector $\begin{pmatrix} v_x \\ v_y \end{pmatrix}$, the displacement is given by the vector $\begin{pmatrix} v_y \\ -v_x \end{pmatrix}$.

Exercise 5 exo017 Consider the algebra $C^\infty(S^2)$ of smooth functions on the sphere S^2 and the Lie algebra $\mathfrak{X}(S^2)$ of smooth vector fields. Consider the vector field

$$Y_i = \sum_{j,k=1}^{3} \epsilon_{ijk} x_j \frac{\partial}{\partial x_k}$$

with $\sum_j (x_j)^2 = 1$. Check that Y_i is actually a tangent vector and prove that these vector fields generate the whole algebra. Then, prove that this set is not free because

$$\sum_j x_j Y_j = 0.$$

This proves that the tangent space of S^2 is non trivial. corr017

Correction of the exercise 5

Let us pick a point $x = (x_1, x_2, x3) \in S^2$ and check that the vector $Y_i(x)$ is tangent to S^2 at the point x. For that, we just have to prove that the product of $Y_i(x)$ and x is zero (when you are tangent to a sphere, you are perpendicular to the radius):

$$Y_i(x) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \sum_{j,k=1}^{3} \epsilon_{ijk} x_j x_k = 0$$

because contraction of symmetric indices (jk in $x_j x_k$) and skew-symmetric indices (jk in ϵ_{ijk}) is zero.

Exercise 6 exo003

Show that the set, $T(S^1)$, of all tangent vectors to $S^1 \subset \mathbb{C}$ is naturally diffeomorphic to a cylinder in $\mathbb{R}^3$. Describe, in terms of the cylinder, the differential of the map $\phi\colon S^1 \to S^1$,

$$\phi(z) = z^N.$$

where $N \in \mathbb{N}_0$ is fixed. corr003

Correction of the exercise 6

A point $z \in S^1$ can be written under the form $z = e^{i\theta_0}$. A tangent vector of S^1 at the point z can be expressed by means of the derivative of a path $X\colon \mathbb{R} \to S^1$ such that $X(0) = z$. Let us consider a general path $X(t) = e^{i\theta(t)}$ with $\theta_0 = \theta_0$. Its tangent vector is

$$X = \frac{d}{dt}\Big[e^{i\theta(t)} \Big]_{t=0} = i\theta'(0) e^{i\theta_0} = \theta'(0) e^{i\left(\frac{\pi}{2} + \theta_0\right)}. \tag{49.377}$$

eq:corrX

Let us point out the fact that the velocity $\theta'(0)$ of the path describes the norm of the tangent vector while $e^{i\theta_0}$ is the point of S^1 on which X is fixed. The differential of ϕ at z on the vector X reads

$$\begin{aligned} d\phi_z X &= \frac{d}{dt}\Big[\phi(X(t))\Big]_{t=0} \\ &= \frac{d}{dt}\Big[e^{iN\theta(t)}\Big]_{t=0} \\ &= N\theta'(0)e^{i(N-1)\theta_0}e^{i\left(\frac{\pi}{2}+\theta_0\right)} \\ &= Nz^{N-1}X. \end{aligned} \tag{49.378}$$ eq:dphiz

Now let us see TS^1 as subset of $\mathbb{R}^4$. Any $X \in TS^1$ can be written as (z, v) where $z \in S^1 \subset \mathbb{C} \simeq \mathbb{R}^2$ is the point at which X is "fixed" and $v \in \mathbb{C} \simeq \mathbb{R}^2$ is the vector itself. Relation (49.377) contains strong constraints on elements $(z, v) \in \mathbb{R}^4$: if $z = e^{i\theta_0} = (\cos\theta, \sin\theta)$, the element v must be of the form $he^{i\left(\frac{\pi}{2}+\theta_0\right)} = (-h\sin\theta, h\cos\theta)$. Now TS^1 is seen as the subset of $\mathbb{R}^4$ of vectors of the form

$$(z, X) = (\cos\theta, \sin\theta, -h\sin\theta, h\cos\theta).$$

At this point it is important to underline two points.

— The map h can be negative as well as positive.

— An element in $\mathbb{C}$ does not uniquely determines an element in TS^1 although tangent spaces of $e^{i\theta}$ and $e^{i(\pi+\theta)}$ look like the same (draw a picture if you are unsure). When one deals with a tangent vector, one has to keep trace of the point on which the vector is fixed. If one does not, one can believe that $T_{z_1}S^1 = T_{z_2}S^1$ and then that finally, $TS^1 = \mathbb{C}$, see figure 49.2.

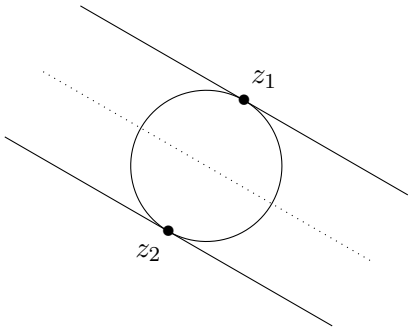


Figure 49.2: A possible mistake is to confuse $T_{z_1}S^1$ and $T_{z_2}S^1$. LabelFigALIzHFm

The way to see a cylinder (of radius 1) as a subset of $\mathbb{R}^3$ is to take the coordinates $(\cos\theta, \sin\theta, h)$. A diffeomorphism $f\colon TS^1 \to \text{Cyl}$ is given by

$$f(\cos\theta, \sin\theta, -h\sin\theta, h\cos\theta) = (\cos\theta, \sin\theta, h). \tag{49.379}$$

We have to check that it is diffeomorphic: it is smooth, bijective and the inverse is smooth too. This needs the "new" definition of smooth map. Indeed the map f is defined on a rather non trivial subset of $\mathbb{R}^4$: the set of (x, y, u, v) such that

$$x^2 + y^2 = 1 \quad \text{and} \quad \begin{cases} v = -ux/y & \text{if } y \neq 0 \\ v \in \mathbb{R} & \text{if } y = 0 \end{cases}$$

which is not open. We have to consider pointwise a neighbourhood and a prolongation. Let's take the part $A_0 \equiv x \neq 0, y \neq 0$ of the domain of f. On this part of the domain, f reads $(x, y, u, v) \mapsto (x, y, \frac{v}{x})$. We can find an open set $A \subset \mathbb{R}^4$ which contains A_0 and such that x and y don't vanish on A. The prolongation of f from A_0 to A by the same formula $(x, y, u, v) \mapsto (x, y, \frac{v}{x})$ is a smooth map.

When $x = 0$, the map f reads $(0, 1, 0, h) \mapsto (1, 0, h)$ which can be prolongated to the smooth map $(x, y, z, h) \mapsto (y, x, h)$. The same idea is true on $y = 0$ where f reads $(1, 0, 0, h) \mapsto (1, 0, h)$.

The inverse map f^{-1} is $(x, y, h) \mapsto (x, y, -hy, hx)$ with the constraint $x^2 + y^2 = 1$ on the domain.

The following diagram is the *definition* of the dotted line; this is what one means when one says "$d\phi$ seen on the cylinder":

$$\begin{array}{ccc} p \in \mathrm{Cyl} & \xrightarrow{\quad f^{-1} \quad} & f^{-1}(p) \in TS^1_z \\ {\scriptstyle d\phi}\,\vdots\downarrow & & \downarrow{\scriptstyle d\phi_z} \\ (f \circ d\phi \circ f^{-1})(p) \in \mathrm{Cyl} & \xleftarrow[f]{} & d\phi_z\big(f^{-1}(p)\big) \in T_{\phi(z)}S^1. \end{array}$$

In this, we suppose that $p \in \mathrm{Cyl}$ is the image by f of a tangent vector on S^1 at the point z.

In order to express the $d\phi_z$ of equation (49.378) in terms of the cylinder, we pick $X = (-\sin\theta, \cos\theta, r) \in \mathrm{Cyl}$, i.e. $X = he^{i(\frac{\pi}{2}+\theta)} \in T_{e^{i\theta}}S^1$ and we apply $d\phi_z$ on it:

$$\begin{aligned} d\phi_z\left(he^{i\left(\frac{\pi}{2}+\theta\right)}\right) &= Ne^{i(N-1)\theta}he^{i\left(\frac{\pi}{2}+\theta\right)} \\ &= Nhe^{i\left(\frac{\pi}{2}+N\theta\right)} \\ &= (-\sin N\theta, \cos N\theta, Nh). \end{aligned}$$

Exercise 7 exo005

Let C be an embedded curve in $\mathbb{R}^2$ globally traced out by a regular path $\gamma\colon]-1, 1[\to \mathbb{R}^2$. It means that γ is a global chart of C. Prove that the velocity field $\{\dot\gamma(t)\}_{t\in]-1,1[}$ naturally defines a tangent vector field $\dot\gamma$ on C. corr005

Correction of the exercise 7

For each t_0, it is clear that $\dot\gamma(t_0) \in T_{t_0}C$. The problem is to prove that the function $f\colon C \to \mathbb{R}^2 \times \mathbb{R}^2$, $f(x) = (x, X_x)$ is smooth. Let x be a point of C and $\mathcal{U}$ an open neighbourhood of x. We have to find a smooth extension $\tilde f\colon \mathcal{U} \to \mathbb{R}^2 \times \mathbb{R}^2$ of f. In particular, we have to find an extension $\hat X$ of X in a neighbourhood $\hat C$ of C in $\mathbb{R}^2$.

First method Since γ is a chart, it can be extended to a smooth bijection $\tilde\gamma\colon]-1, 1[\times]-1, 1[\to \mathcal{U}$ such that $\tilde\gamma(0, t) = \gamma(t)$. For each $y = \tilde\gamma(u_0, t_0) \in \mathcal{U}$, we define

$$\tilde X_y = \frac{d}{dt}\Big[\tilde\gamma(u, t)\Big]_{t=t_0}.$$

Now we define $\tilde f(y) = (y, \tilde X_y)$. The vector field $\tilde X$ being a directional derivative of a smooth function, it is itself a smooth function.

Here is a partial proof of the existence of the extension $\hat\gamma$.

Lemma 49.158.
Let φ be a chart of an embedded curve C in $\mathbb{R}^2$; is is seen as a map $\varphi\colon \mathcal{U} \to \hat C$ where $\hat C$ is an open neighbourhood of C in $\mathbb{R}^2$.

Then there exists an open set $W \subset \mathbb{R}^2$ and a smooth map $\varphi_2\colon W \to \hat C$ such that $\tilde\varphi = \varphi_2^{-1} \circ \varphi\colon \mathcal{U} \to W$ has the form

$$\tilde\varphi(x) = (x, 0).$$

We will not prove it here.

Now we can extend this $\tilde\varphi$ to the map $\hat{\tilde\varphi}\colon \mathcal{U} \times I \to W$ by $\hat{\tilde\varphi}(x, y) = (x, y)$ for a suitable neighbourhood I of 0 in $\mathbb{R}$. Then we look at

$$\hat\varphi = \varphi_2 \circ \hat{\tilde\varphi}\colon \mathcal{U} \times I \to \hat C$$

which is smooth.

How to choice I ? For each $x \in \mathcal{U}$, the point $(x,0)$ belongs to the open set $W \subset \mathbb{R}^2$. Therefore there exists a non empty neighbourhood I_x of 0 such that $(x, I_x) \subset W$. Taking if necessary a smallest chart $\mathcal{U}'$ instead of $\mathcal{U}$, one can take the non empty set $I = \bigcap_{x\in\mathcal{U}} I_x$.

Exercise 8 exo007

Prove that the data of a tangent vector field X on a curve C is equivalent to the data of a smooth map $\Xi\colon C \to \mathbb{R}^2$ with $\Xi(x) \in T_xC$, the identification being given by $\Xi(x) = X_x$. In other words, locally, a tangent vector field on a curve embbeded in $\mathbb{R}^2$ is the restriction to C of a vector field $\hat{\Xi}$ on some open neighbourhood U of C in $\mathbb{R}^2$. Draw a picture illustrating this fact. corr007

Correction of the exercise 8

Let us consider a smooth map $\Xi\colon C \to \mathbb{R}^2$ with $\Xi(x) \in T_x(C)$. It defines the tangent vector field $X_x = \Xi(x)$. It is smooth because Ξ is, and thus so is $x \mapsto (x, X_x)$. The inverse sense is the same.

Exercise 9 exo006

Prove that, identifying $T(S^1)$ with the cylinder $\mathrm{Cyl} := S^1 \times \mathbb{R} \subset \mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$, the data of a tangent vector field X on S^1 is equivalent to the data of a smooth map $\xi\colon S^1 \to \mathrm{Cyl}$ such that $\mathtt{proj}_1 \circ \xi = \mathrm{Id}_{S^1}$ where $\mathtt{proj}_1 : \mathbb{R}^3 \to \mathbb{R}^2 : (x,y,z) \mapsto (x,y)$. corr006

Correction of the exercise 9

A general form of a function $\xi\colon S^1 \to \mathrm{Cyl}$ is

$$\xi(x,y) = (x, y, h(x,y))$$

only defined on $x^2 + y^2 = 1$. Here, h is a smooth function. We naturally associate to ξ the field

$$X_{(x,y)} = h(x,y)e^{i\left(\frac{\pi}{2}+\theta\right)}$$

where θ is the angular polar coordinate of (x,y). We have to prove that this field is smooth. For this, we have to prove that

$$(x,y) \to (x, y, h(x,y)e^{i\left(\frac{\pi}{2}+\theta\right)}) \tag{49.380}$$

eq:xytohxy

is smooth at each point of $x^2 + y^2 = 1$. Since h is smooth, by definition it can be extended. So we can extend equation (49.380) keeping the same form around any $(x_0, y_0) \in S^1$. The result is smooth in the sense of functions from an open set of $\mathbb{R}^2$ to an open set in $\mathbb{R}^3$.

Let us now prove that a vector field defines a function. The general form of a vector field on S^1 is

$$X_{(x,y)} = h(x,y)e^{i\left(\frac{\pi}{2}+\theta\right)} \tag{49.381}$$

where h is smooth on S^1 and can thus be extended into a smooth function on a neighbourhood of each point of S^1. The map $\xi\colon S^1 \to \mathrm{Cyl}$ we were looking for is thus

$$\xi(x,y) = (x, y, h(x,y)).$$

A second way is possible The reader should remark that the construction given up to here is essentially the construction of exercise 6. We can direclty exploit the diffeomorphism $f\colon TS^1 \to \mathrm{Cyl}$. So in a first time, we pick a $\xi\colon S^1 \to \mathrm{Cyl}$ and we want to buils a vector field X^ξ, i.e. a map $X^\xi\colon S^1 \to TS^1$. The natural candidate is

$$X^\xi = f^{-1} \circ \xi.$$

On the one hand, smoothness comes from smoothness of f^{-1} and ξ and on the other hand, $X^\xi(x) \in T_xS^1$ because of projection property of ξ.

In a second time, we pick a vector field $X\colon S^1 \to TS^1$ and we want to build a map $\xi^X\colon S^1 \to \mathrm{Cyl}$. The candidate is

$$\xi^X = f \circ X.$$

You have to check that $\texttt{proj}_1 \circ \xi^X = \mathrm{Id}$; it comes from construction of f.

Now in order to say that the data of the function is "the same" that the data of the vector field, we have to prove that the two constructions are inverse each other. In other words, you are now able to do two things: if I give you a vector field, you can give me a function and if I give you a function, you can give me a vector field. Let me give you the function ξ. So you give me a vector field X^ξ. Now I give you X^ξ; will you give back the original function ξ ? If not, the whole construction has no interest.

$$\xi \to X^\xi \to \xi^{X^\xi} \stackrel{?}{=} \xi.$$

From definitions, $\xi^{X_\xi} = f \circ X^\xi = f \circ f^{-1} \circ \xi = \xi$. This is a good point. The other is

$$X \to \xi^X \to X^{\xi^X} \stackrel{?}{=} X;$$

from definitions, $X^{\xi^X} = f^{-1} \circ \xi^X = f^{-1} \circ f \circ X = X$.

Exercise 10 exo009

Let C be a curve in $\mathbb{R}^2$ and X, $Y \in \mathfrak{X}(C)$ with $X_x \neq 0$ for all $x \in C$. Then there exists a smooth curve $f : C \to \mathbb{R}$ such that $Y = fX$, i.e. $Y_x = f(x)X_x$ for each $x \in C$.

Hint Show the property in the case $C = \mathbb{R}$ and $C = S^1$. In the general case, you can use a diffeomorphism to transform C into $\mathbb{R}$ or S^1. corr009

Correction of the exercise 10

When $C = \mathbb{R}$, a vector field is just a smooth function $\mathbb{R} \to \mathbb{R}$ (be sure that you deeply understand this assertion !). If g_X is the function of the vector field X and f_Y the one of Y, the answer is given by

$$f(x) = \frac{g_Y(x)}{g_X(x)} \tag{49.382}$$

which is smooth as quotient of smooth functions.

We turn our attention to the case $C = S^1$. The formula, for each $x \in C$

$$Y_x = f(x)X_x$$

completely defines f because X doesn't vanishes. The problem is to prove that it is smooth (extensions and all that). The natural way to begin is to say that X and Y admit smooth extensions $\hat{X}$ and $\hat{Y}$ in terms of which, we should be able to build an extension $\hat{f}$. The problem is that you have no guarantee that $\hat{X}$ and $\hat{Y}$ are parallel outside C.

One can build normal vectors to the curve and translate X and Y along these lines. Then one gets smooth parallel extensions. Such a construction was given in the course during the proof of the lemma just below the inverse function theorem. From there, one defines the extension of f by

$$\hat{Y}_z = \hat{f}(z)\hat{X}_z$$

for all z in a neighbourhood of C. But how to prove that it is smooth ? Maybe it is possible. I don't know, but this formula seems to define a smooth function since f is a quantity which relates two smooth quantities each other.

A possible way to solve the exercise is to use results of exercise 9. Let $\xi_X : S^1 \to \mathrm{Cyl}$ be the map associted to X, and ξ_Y the one associated with Y. They can be written under the form

$$\xi_X(x) = (.,.,h_X(x)) \tag{49.383a}$$
$$\xi_Y(x) = (.,.,h_Y(x)) \tag{49.383b}$$

where the dots represent some non essential functions. The functions h_X and h_Y are smooth and we define

$$\hat{f}(z) = \frac{\hat{h}_Y(z)}{\hat{h}_X(z)}$$

which is smooth and a correct extension of f.

Now a general curve C is diffeomorphic of $\mathbb{R}$ or S^1. We denote by $\alpha\colon C \to E$ the diffeomorphism where E denotes $\mathbb{R}$ or S^1. The map α allows us to transform a complicated problem on C into a solved problem on E. We consider $\tilde{X} = d\alpha X$ and $\tilde{Y} = d\alpha Y$, two vector fields on E. Then there exists $\tilde{f}\colon E \to \mathbb{R}$ such that

eq_tildeffx

$$\tilde{Y}_x = \tilde{f}(x)\tilde{X}_x \tag{49.384}$$

for each $x \in E$.

Let us summarize the functions that we have at hand:

$$\alpha\colon C \to E$$
$$\tilde{f}\colon E \to \mathbb{R}$$

and we are searching for

$$f\colon C \to \mathbb{R}.$$

Obvioulsly the candidate is $f = \tilde{f} \circ \alpha$ which is smooth because it is a composition of smooth functions. In order to check that it is the right function, we have to prove that, for all $x \in C$, $Y_x = (\tilde{f} \circ \alpha)(x)X_x$.

The defining property of $\tilde{f}$ is

$$d\alpha_x Y_x = \tilde{f}(\alpha(x))d\alpha_x X_x.$$

Since α is diffeomorphic, $d\alpha_x$ is a linear injective map for each x; so we can "simplify" both sides by $d\alpha_x$.

Exercise 11 exo010

Prove the following.

lem_XYparall

Lemma 49.159.
Let X be a smooth non vanishing vector field on an open set $\mathcal{U}$ in $\mathbb{R}^2$, $f\colon \mathcal{U} \to \mathbb{R}$ a smooth non vanishing function and Y the smooth vector field on $\mathcal{U}$ defined by $Y_x = f(x)X_x$. If $c\colon I \to \mathbb{R}^2$ is an integral curve of X, then there exists a smooth map $\alpha\colon I \to \mathbb{R}$ such that $c \circ \alpha$ is an integral curve of Y.

In other words, if one changes the norm of a vector field, then one just has to change the parametrization of integral curves.

corr010

Correction of the exercise 11

By definition of c as an integral curve of X, we have

$$\dot{c}(t) = X_{c(t)}$$

for all t. The requirement for $c \circ \alpha$ to be an integral curve of Y reads

$$\begin{aligned}(c \circ \alpha)'(t) &= c'\big(\alpha(t)\big)\alpha'(t)\\ &\overset{!}{=} Y_{(c\circ\alpha)(t)}\\ &= f\big((c \circ \alpha)(t)\big)X_{(c\circ\alpha)(t)}.\end{aligned} \tag{49.385}$$

Since we supposed that X is non vanishing, $c'(t) \neq 0$ everywhere. So we find the following differential equation for α:

$$\alpha'(t) = (f \circ c)(\alpha(t)). \tag{49.386}$$

The existence of a solution for this equation is given by the following powerful theorem 49.157.

Remark that this theorem only deals with continuity (not smoothness). It should be an exercise to find an even more powerful theorem which ensures the smoothness of α.

Exercise 12 exo008

In the setting of exercise 8, prove that the curve C is an **integral curve** of $\hat{\Xi}$ i.e. for all initial data $x_0 \in C$ the solution $x = x(t),\ t \in]-\epsilon, \epsilon[$ to the Cauchy problem

$$\dot{x} = \hat{\Xi}\,; \quad x(0) = x_0$$

is a local chart on C around x_0. We suppose that X is non vanishing.

Hints Draw a circle and any smooth tangent vector field. Extend it in any smooth manner. Then pick a point on the circle. Is it possible to find an integral curve of your vector field which is not the circle ? Try and convince yourself that the integral curves of a vector field X are defined by the field of *directions* and are independent of the *norm* of X_x. corr008

Correction of the exercise 12

Let us begin by *almost* true proof. Why "almost" ? Because we will not care about *locality* problems. We will speak about C, $\mathbb{R}^2$,... while we should speak about "a neigborhood of $x_0 \in C$"... You should do it as an exercise.

Let $\varphi\colon \mathcal{U} \to C$ be a paramertization of C with tangent vector field V. We want to find a new parametrization $\phi\colon \mathcal{U}' \to C$ with tangent vector field $X_x = f(x)V_x$ where the smooth function f is given. In other words, we want a $a\colon \mathcal{U}' \to \mathcal{U}$ and $\phi = \varphi \circ a$ with

$$\phi'(t) = (\varphi \circ a)'(t) = \varphi'(a(t))a'(t).$$

The function a is the solution of the differential equation

$$a'(t) = f\big(a(t)\big).$$

So another way to state this result is to say that there always exists a reparametrization of a curve for which the tangent vector field is given.

Exercise 13 exo011

Prove that on an embedded surface $\Sigma \subset \mathbb{R}^3$, the following formula

$$\mathcal{I}_x(v, w) := v \cdot w$$

for each $x \in \Sigma$ and $v,\ w \in T_x(\Sigma) \subset \mathbb{R}^3$ defines a Riemannian metric, $\mathcal{I}$, on Σ called the **first fundamental form** of Σ. corr011

Correction of the exercise 13

The fact for $\mathcal{I}$ to be bilinear, symmetric, positive definite and nondegenerate comes from the corresponding properties of the inner product on $\mathbb{R}^3$. The eventual problem is smoothness. As usual, we have to find extensions. Let $X,\ Y \in \mathfrak{X}(\Sigma)$; the problem is to find a (local) smooth extension of the function $\mathcal{I}(X, Y)\colon \Sigma \to \mathbb{R},\ x \mapsto \mathcal{I}_x(X_x, Y_x) = X_x \cdot Y_x$. If $\hat{X}$ and $\hat{Y}$ are extensions of X and Y, what do you think about

$$\hat{\mathcal{I}}(X, Y)(y) = \hat{X}_y \cdot \hat{Y}_y \quad ?$$

Exercise 14 exo012

Give the expression of the first fundamental form of the sphere S^2:

(1) within spherical coordinates (θ, ϕ);

(2) within stereographical coordinates $z \in \mathbb{C}$.

corr012

Correction of the exercise 14

The chart $f\colon \mathbb{R}^2 \to S^2$ of S^2 that we chose is

$$f(\theta, \varphi) = (\cos\theta \sin\varphi, \cos\theta \cos\varphi, \sin\theta).$$

In these coordinates, a basis of $T_{(\theta_0,\varphi_0)}S^2$ is given by

$$\frac{d}{dt}\Big[f(\theta_0+t,\varphi_0)\Big]_{t=0} \quad \text{and} \quad \frac{d}{dt}\Big[f(\theta_0,\varphi_0+t)\Big]_{t=0}.$$

Computations give

$$\partial_\theta = -\begin{pmatrix}\sin\theta_0\sin\varphi_0\\ \sin\theta_0\cos\varphi_0\\ -\cos\theta_0\end{pmatrix}, \tag{49.387a}$$

$$\partial_\varphi = \begin{pmatrix}\cos\theta_0\cos\varphi_0\\ -\cos\theta_0\sin\varphi_0\\ 0\end{pmatrix}. \tag{49.387b}$$

One can check that $\partial_\theta \cdot f(\theta_0,\varphi_0) = \partial_\varphi \cdot f(\theta_0,\varphi_0) = 0$: tangent vectors are orthogonal to radius. It is easy to see that $\partial_\varphi \cdot \partial_\theta = 0$, $\partial_\theta \cdot \partial_\theta = 1$ and $\partial_\varphi \cdot \partial_\varphi = \cos^2\theta$.

Exercise 15 exo013 Compute the first fundamental form of the torus $T^2 \subset \mathbb{R}^3$. corr013

Correction of the exercise 15

Let us parametrize the torus by two angles θ (radius R) and φ (radius r). We suppose that θ descibes a circle in the plane xz (or a parallel plane when φ is non zero) while the plane of the rotation φ depends of θ.

A simple rotation (an isometry of $\mathbb{R}^3$) allows us to only consider points with $\theta = 0$. So we expect that the result will not depend of θ. Let us explicitly give the parametrization $f\colon \mathbb{R}^2 \to T^2$. When φ_0 is fixed, the curve $f(\theta,\varphi_0)$ is a circle of radius $R + r\cos\varphi_0$ in a plane parallel to xz with center at $(0, r\sin\varphi_0, 0)$. Thus we have

$$f(\theta,\varphi) = \Big((R+r\cos\varphi)\sin\theta, r\sin\varphi, (R+r\cos\varphi)\cos\theta\Big). \tag{49.388}$$

We find

$$\partial_\varphi|_{(\theta_0,\varphi_0)}\frac{d}{dt}\Big[f(\theta_0,\varphi+t)\Big]_{t=0} = -r\Big(\sin_0\varphi\sin\theta_0, -\cos\varphi_0, \sin\varphi_0\cos\theta_0\Big) \tag{49.389a}$$

$$\partial_\theta|(\theta_0,\varphi_0) = \frac{d}{dt}\Big[f(\theta_0+t,\varphi_0)\Big]_{t=0} = (R+r\cos\varphi_0)\Big(\cos\theta_0, 0, -\sin\theta_0\Big). \tag{49.389b}$$

Given under a matrix form, the result is

$$g = \begin{pmatrix} r^2 & 0\\ 0 & (R+r\cos\varphi)^2\end{pmatrix}.$$

Can you geometrically understand the dependance in r, R and φ ?

Note that the norm of an angular coordinate is the radius of the corresponding circle.

Exercise 16 exo014 Same question for a surface of revolution generated by the curve $[y = f(x)] \cap [z = 0]$. corr014

Correction of the exercise 16

We parametrize the surface by x and θ:

$$f(x,\theta) = (x, f(x)\cos\theta, f(x)\sin\theta).$$

In this parametrization, the basis vectors are given by

$$e_x = (1, f'(x)\cos\theta, f'(x)\sin\theta) \tag{49.390a}$$

$$e_\theta = (0, -f(x)\sin\theta, f(x)\cos(\theta)), \tag{49.390b}$$

and the matrix g is given by

$$g = \begin{pmatrix}1+f'(x)^2 & 0\\ 0 & f(x)^2\end{pmatrix}$$

Christoffel symbols are obtained by formula

$$\Gamma_{ij}^k = \frac{1}{2} g^{lk}(\partial_i g_{jl} + \partial_j g_{li} - \partial_l g_{ij})$$

where the g^{lk} are elements of the *inverse* matrix of g. The non zero symbols are

$$\Gamma_{\theta\theta}^x = -\frac{ff'}{1+f'^2}, \quad \Gamma_{\theta x}^\theta = \Gamma_{x\theta}^\theta = \frac{f'}{f}, \quad \Gamma_{xx}^x = \frac{f'f''}{1+f'^2}. \tag{49.391}$$

Covariant derivatives of basis vectors are

$$\nabla_x e_x = \Gamma_{xx}^x e_x \qquad \nabla_\theta e_x = \Gamma_{\theta x}^\theta e_\theta \tag{49.392}$$

$$\nabla_x e_\theta = \Gamma_{x\theta}^\theta e_\theta \qquad \nabla_\theta e_\theta = \Gamma_{\theta\theta}^x e_x. \tag{49.393}$$

As far as Riemann tensor and related curvature issues are concerned, we do not need to compute $\nabla_x \nabla_x e_i$ and $\nabla_\theta \nabla_\theta e_i$. Computations are as follows:

$$\begin{aligned} \nabla_x \nabla_\theta e_x &= \partial_x(\Gamma_{\theta x}^\theta) e_\theta + \Gamma_{x\theta}^\theta \Gamma_{x\theta}^\theta e_\theta \\ &= \frac{f''}{f} e_\theta. \end{aligned}$$

Other results —up to personal faults— are

$$\begin{aligned} \nabla_x \nabla_\theta e_\theta &= \Big(\frac{3ff'^2 f''}{(1+f'^2)^2} = \frac{f'^2 + ff''}{1+f'^2} \Big) e_x \\ \nabla_\theta \nabla_x e_x &= \Big(\frac{f'f''}{1+f'^2} \Big) \frac{f'}{f} e_\theta \\ \nabla_\theta \nabla_x e_\theta &= -\frac{f'}{f} \frac{ff'}{1+f'^2} e_x. \end{aligned} \tag{49.394}$$

Isometry criterion All this part is "without proof".

We consider the following tensor:

$$R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$$

and the **Ricci curvature** pointwise defined by

$$r(Y,Z) = \operatorname{Tr}\big[X \mapsto R(X,Y)Z\big].$$

It is a trace in the sense of the trace of a matrix of a linear operator in a vector space:

$$R(.,Y)Z \colon T_x\Sigma^f \to T_x\Sigma^f. \tag{49.395}$$

For a general $A \colon V \to V$, it is defined by

$$\operatorname{Tr} A = \sum_i \langle e_i, Ae_i \rangle$$

where $\{e_i\}$ is a basis of V. We define the **scalar curvature** by

$$\rho = \sum_{a=1}^{2} r(e_a, e_a).$$

This defines a real valued function $\rho^f \colon \Sigma^f \to \mathbb{R}$ for the surface Σ^f and a corresponding one $\rho^g \colon \Sigma^g \to \mathbb{R}$ for Σ^g. From symmetry of the problem, these functions do not depend on the angular part of the parametrization of Σ^f and Σ^g. So one can only consider the restrictions

$$\rho^f \colon]a,b[\to \mathbb{R}, \tag{49.396a}$$

$$\rho^g \colon]c,d[\to \mathbb{R}. \tag{49.396b}$$

The criterion for Σ^f and Σ^g to be isometric is the existence of a diffeomorphism $\varphi\colon]a,b[\to]c,d[$ such that

$$\rho^f = \rho^g \circ \varphi.$$

Explicit proof of existence of such a φ can be very hard.

Exercise 17 exo015

Compute the Christoffel symbols in the following cases: the plane $XoY \subset \mathbb{R}^3$, the sphere parametrized by spherical coordinates (θ, ϕ), a surface of revolution. corr015

Correction of the exercise 17

The polar coordinate on the sphere is given by

$$f(\theta,\varphi) = r\begin{pmatrix}\cos\theta\cos\varphi\\ \cos\theta\sin\varphi\\ \sin\theta\end{pmatrix}$$

where r is the (constant) radius. Simple derivations shows that

$$e_\theta = r\begin{pmatrix}-\sin\theta\cos\varphi\\ -\sin\theta\sin\varphi\\ \cos\theta\end{pmatrix}, \quad e_\varphi = r\begin{pmatrix}-\cos\theta\sin\varphi\\ \cos\theta\cos\varphi\\ 0\end{pmatrix}.$$

The metric matrix is given by

$$g = r^2\begin{pmatrix}1 & \\ & \cos^2\theta\end{pmatrix}.$$

The non zero Christoffel symbols are

$$\begin{aligned}\Gamma^\theta_{\varphi\varphi} &= \cos\theta\sin\theta,\\ \Gamma^\varphi_{\theta\varphi} &= -\tan\theta.\end{aligned} \tag{49.397}$$

Exercise 18 exo016

Since basis vectors given in equation (49.372) are just derivatives with respect to coordinates, it is obvious that they commute: $[e_i, e_j] = 0$. The fact that basis vectors commute seems to prove that any vector field should commute. Find the mistake. Can you find two vector fields on the sphere which doesn't commute ? corr016

Correction of the exercise 18

When coeficient of a vector *field* are non constant functions, the computation of XYf shows that X act not only on f but also on the coeficients of X. For example $X = e_\theta$ and $Y = \sin\theta\, e_\varphi$. At (θ_0, φ_0), the path which define these vectors are

$$Y_{(\theta_0,\varphi_0)}(t) = (\theta_0, \varphi_0 + t\sin\theta_0), \tag{49.398a}$$
$$X_{(\theta_0,\varphi_0)}(t) = (\theta_0 + t, \varphi_0). \tag{49.398b}$$

So $Xf = \frac{\partial f}{\partial\theta}$ and $Yf = \sin\theta\frac{\partial f}{\partial\varphi}$,

$$\begin{aligned}(XY)f &= \frac{\partial^2 f}{\partial\theta\,\partial\varphi}\sin\theta + \frac{\partial f}{\partial\varphi}\cos\theta\\ (YX)f &= \sin\theta\frac{\partial^2 f}{\partial\theta\,\partial\varphi}.\end{aligned} \tag{49.399}$$

Chapter 50

Representations

50.1 Definitions

The concept of representation of a group is already defined in 4.135.

Definition 50.1.
*[513] If V is a locally convex space, a **continuous representation** of a Lie group G on V is a left invariant action $\pi\colon G \times V \to V$ such that for any $x \in G$, the map $\pi(x)\colon V \to V$ is a linear endomorphism of V.*

Definition 50.2.
*If $\mathfrak{g}$ is a Lie algebra, a **representation** of $\mathfrak{g}$ in V is a bilinear map $\sigma\colon \mathfrak{g} \to \mathrm{End}(V)$ such that*

$$\sigma([X,Y])v = [\sigma(X), \sigma(Y)]v = \sigma(X)\sigma(Y)v - \sigma(Y)\sigma(X)v. \tag{50.1}$$

In other words, $\sigma\colon \mathfrak{g} \to \mathrm{End}\, V$ is an algebra homomorphism.

A vector space equipped with a representation of a Lie algebra $\mathfrak{g}$ is a **$\mathfrak{g}$-module**. A *complete* locally convex space equipped with a representation of a Lie group is a **G-module**.

If π is a representation of G in a (eventually complex) vector space V, an **invariant subspace** is a vector subspace $W \subset V$ such that $\pi(x)W \subset W$ for any $x \in G$. A continuous representation in a complete locally compact vector space V is **irreducible** if $\{0\}$ and V are the only two invariant closed subspaces of V.

In the case of finite dimensional vector space, any subspace is closed; in this class, we find back the usual notion of irreducibility.

Definition 50.3.
*An **unitary** representation of G is a continuous representation π of G in a complex (or real) Hilbert space H such that $\pi(x)$ is unitary for any $x \in G$. This is: π is unitary if and only if $\forall x \in G$, v, $w \in H$,*

$$\langle \pi(x)v, w\rangle = \langle v, \pi(x)^{-1}w\rangle. \tag{50.2}$$

*A continuous and finite dimensional representation is **unitarisable** if there exists an hermitian product for which the representation is unitary.*

50.1.1 Complex conjugate representation

Definition 50.4.
*Let $(V, +, \cdot\,)$ be a vector space on $\mathbb{C}$ (the dot is the multiplication by a scalar). The **complex conjugate** vector space is $(V, \oplus, \odot)$ with the definitions*

(1) $x \oplus y = x + y$ for every $x, y \in V$

(2) $\lambda \odot x = \bar{\lambda}\dot{x}$ for $\lambda \in \mathbb{C}$ and $x \in V$.

We usually denote by $\bar{V}$ the complex conjugate vector space, but as sets, $V = \bar{V}$.

LEMooEKTRooBApWlp

Lemma 50.5.
If the map $f: V \to W$ is linear, then the (same) map $f: \bar{V} \to \bar{W}$ is linear.

Proof. There is no requirements for the sum since the complex conjugation of the sum is the same. For the product,

$$f(\lambda \odot x) = f(\bar{\lambda}x) = \bar{\lambda}f(x) = \lambda \odot f(x). \tag{50.3}$$

□

Definition 50.6.
*If G is a group and (ρ, V) a representation of G, the **complex conjugate** is the representation $\bar{\rho}: G \to \mathrm{GL}(\bar{V})$ given by*

$$\bar{\rho}(g) = \rho(g) \tag{50.4}$$

by the lemma 50.5.

50.1.2 Faithful representation

DEFooAFSAooGDSDBb

Definition 50.7.
*The representation (V, T) is **faithful** if the map $T: G \to \mathrm{GL}(V)$ is injective. It means that $T(g) = \mathrm{Id}$ only when g is the neutral in G.*

When G is a subgroup of $\mathrm{GL}(V)$ we say that

$$\begin{aligned} \rho: G &\to \mathrm{GL}(V) \\ g &\mapsto g \end{aligned} \tag{50.5}$$

is the **definition representation**. This is in particular the case when G is a group of matrix; the definition representation of G is then the action of these matrices on $\mathbb{R}^n$ or $\mathbb{C}^n$. This is a representation by proposition 4.137.

Definition 50.8.
*Let G be a group. Two representations (T_1, V_1), (T_2, V_2) of G are **equivalent** is there exists a vector space isomorphism*[1] *$m: V_1 \to V_2$ such that*

$$T_2(g) \circ m = m \circ T_1(g) \tag{50.6}$$

for every $g \in G$.

Definition 50.9.
*If $T: G \to \mathrm{GL}(V)$ is a representation, a subspace $W \subset V$ is **invariant** under T if*

$$T(g)W \subset W \tag{50.7}$$

for every $g \in G$.

*A representation is **irreducible** if V and $\{0\}$ are the only invariant subspaces of V.*

50.1.3 Schur lemma

Lemma 50.10 (Schur[844])**.**
Let (T, V) and (S, W) be two irreducible representations of G. Let $\alpha: V \to W$ be a linear map such that

$$\alpha \circ T(g) = S(g) \circ \alpha \tag{50.8}$$

for every $g \in G$. If $\alpha \neq 0$, then α is a bijection.

Proof. Let L be the image of α: $L = \alpha(V)$.

1. Linear bijection

(1) L **is invariant under** S Let $y \in \alpha(V)$: there exist $x \in V$ such that $y = \alpha(x)$. Then we have
$$S(g)\big(\alpha(x)\big) = \big(\alpha \circ T(g)\big)x = \alpha\big(T(g)x\big) \in \alpha(V). \tag{50.9}$$

(2) **Two possibilities** Since S is irreducible, there are two possibilities: $\alpha(V) = \{0\}$ and $\alpha(V) = W$.

(3) **First:** $\alpha(V) = \{0\}$ In this case, $\alpha = 0$.

(4) **Second:** $\alpha(V) = W$ This means that α is surjective. We show that α is injective too. The space $\ker(\alpha)$ is invariant under T because, if $z \in \ker(\alpha)$, then
$$\alpha\big(T(g)z\big) = S(g)\big(\alpha(z)\big) = S(g)(0) = 0. \tag{50.10}$$
There are two possibilities: $\ker(\alpha) = \{0\}$ and $\ker(\alpha) = V$. The possibility $\ker(\alpha) = V$ says $\alpha = 0$ while we are in the case $\alpha(V) = W$. So we deduce $\ker(\alpha) = \{0\}$ which means that α is injective.

□

50.1.4 Irreducible representations

THOooXHVHooDQdgDI

Theorem 50.11 ([844]).
Let (T, V) be an irreducible finite dimensional representation of G. Let $\alpha \in \mathrm{End}(V)$ be such that
$$\alpha \circ T(g) = T(g) \circ \alpha \tag{50.11}$$
for every $g \in G$. There exists $\lambda \in \mathbb{C}$ such that $\alpha = \lambda\,\mathrm{Id}$.

Proof. The endomorphism α has at least one eigenvalue $\lambda \in \mathbb{C}$ (proposition 12.96). Let $V_\lambda = \{x \in V \text{ tel que } \alpha(x) = \lambda x\}$ be the associated eigenspace.

We prove that V_λ is an invariant subspace: if $x \in V_\lambda$, then $T(g)x$ has eigenvalue λ for α. Let's do that:
$$\alpha\big(T(g)x\big) = T(g)\big(\alpha(x)\big) = T(g)(\lambda x) = \lambda T(g)x. \tag{50.12}$$

Since T is irreducible, the invariant subspace V_λ is $\{0\}$ or V. The space V_λ has at least dimension one because λ is an eigenvalue. Thus $V_\lambda = V$ and $\alpha(x) = \lambda x$. □

THOooFFJGooCekFQc

Theorem 50.12 ([844]).
Every irreducible representation of an abelian group has dimension 1.

Proof. Let G be an abelian group and (T, V) an irreducible representation of G. If $g_1, g_2 \in G$ we have $g_1 g_2 = g_2 g_1$ and thus
$$T(g_1)T(g_2) = T(g_2)T(g_1). \tag{50.13}$$
The endomorphism $\alpha = T(g_1)$ commutes with every $T(g)$ and is then a multiple of Id by the theorem 50.11. So there exists a map $\lambda\colon G \to \mathbb{C}$ such that $T(g) = \lambda(g)\,\mathrm{Id}$.

Every subspace of V is invariant. In particular a subspace of dimension 1 is the whole space V. □

50.1.5 Direct sum of representations

DEFooGKALooIjJcpV

Definition 50.13.
Let G be a group and (ρ_k, V_k) be representations of G. We define the ***direct sum representation*** *$(\rho_1 \oplus \ldots \oplus \rho_n)$ on $V_1 \oplus \ldots \oplus V_n$ by the formula*
$$(\rho_1 \oplus \ldots \oplus \rho_n)(g)(x_1 + \ldots + x_n) = \rho_1(g)x_1 + \ldots + \rho_n(g)x_n. \tag{50.14}$$

Definition 50.14.
*A representations (ρ, V) of the group G is **completely reducible** if it is a direct sum*[2] *of irreducible representations.*

Lemma 50.15 ([844])**.**
A finite dimensional unitary representation is completely reducible.

Proof. Let (ρ, V) an unitary representation for the hermitian product of V. If ρ is irreducible, the theorem is proved. If not, we consider an invariant non trivial subspace $M \subset V$ and

$$M^\perp = \{x \in V \text{ tel que } \langle x, y\rangle = 0 \forall y \in M\}. \tag{50.15}$$

We know from proposition 8.185 that $V = M \oplus M^\perp$. We show that $M^\perp$ is invariant under the representation ρ; namely we take $x \in M^\perp$ and we show that $\rho(g)x \in M^\perp$. Let $y \in M$; we have

$$\begin{aligned}\langle \rho(g)x, y\rangle &= \langle \rho(g^{-1})\rho(g)x, \rho(g)^{-1}y\rangle && (50.16a)\\ &= \langle x, \rho(g^{-1})y\rangle && (50.16b)\\ &= 0. && (50.16c)\end{aligned}$$

Justifications:

— We can insert $\rho(g^{-1}) = \rho(g)^{-1}$ in the hermitian product because ρ is unitary.
— Since M is invariant, $\rho(g^{-1})y \in M$.
— Since $x \in M^\perp$, we have $\langle x, \rho(g^{-1})y\rangle = 0$.

Thus ρ is a direct sum of is restrictions to M and $M^\perp$. □

Proposition 50.16 ([844])**.**
A completely reducible representation finite dimensional is irreducible if and only if the only operators commuting with every operator in the representation are multiple of the identity.

Proof. In the direct sense, we suppose that the representation is irreducible. The theorem 50.11 concludes.

In the reverse sense we suppose that (ρ, V) is completely reducible and the only operator commuting with the representation is the identity. We have to prove that ρ is irreducible. For that we suppose that ρ is not irreducible and we will found an operator which is not a multiple of the identity while commuting with $\rho(g)$ for every $g \in G$.

Let V_1 and V_2 be invariant subspaces such that $V = V_1 \oplus V_2$. We consider $T = \lambda_1 \mathrm{Id}_{V_1} \oplus \lambda_2 \mathrm{Id}_{V_2}$ where $\lambda_1 \neq \lambda_2$ are two non vanishing numbers.

Let $g \in G$ and $x \in V$. We can write $x = x_1 + x_2$ with $x_1 \in V_1$ and $x_2 \in V_2$. We have:

$$\begin{aligned}\big(T \circ \rho(g)\big)(x) &= T\big(\rho(g)x_1 + \rho(g)x_2\big) && (50.17a)\\ &= \lambda_1\rho(g)x_1 + \lambda_2\rho(g)x_1 && (50.17b)\\ &= \rho(g)(\lambda_1 x_1) + \rho(g)(\lambda_2 x_2) && (50.17c)\\ &= \rho(g)(\lambda_1 x_1 + \lambda_2 x_2) && (50.17d)\\ &= \big(\rho(g) \circ T\big)(x_1 + x_2) && (50.17e)\\ &= \big(\rho(g) \circ T\big)x. && (50.17f)\end{aligned}$$

So T commutes with the representation while not being a multiple of the identity. □

50.1.6 Tensor product of representations

The notion of tensor product of vector spaces is given in 10.170. The tensor product of two maps is defined, when $f\colon E \to E$ and $g\colon F \to F$ are linear,

$$\begin{aligned}f \otimes f\colon E \otimes F &\to E \otimes F\\ x \otimes y &\mapsto f(x) \otimes f(y).\end{aligned} \tag{50.18}$$

2. Definition 50.13.

Definition 50.17.
Let ρ_1 and ρ_2 be representations of G on V_1 and V_2. We define the representation $\rho_1 \otimes \rho_2$ on $V_1 \otimes V_2$ by

$$(\rho_1 \otimes \rho_2)(g) = \rho_1(g) \otimes \rho_2(g). \tag{50.19}$$

Proposition 50.18 ([844])**.**
Let (ρ_1, V_1) be a representation of G_1 and (ρ_2, V_2) be representations of G_2. The formula

$$\rho(g_1, g_2) = \rho_1(g_1) \otimes \rho_2(g_2) \tag{50.20}$$

defines a representation of $G_1 \times G_2$ on $V_1 \otimes V_2$.

Chapter 51

Lie algebras

DEFooVBPKooGxlDBn

Definition 51.1.
*A **Lie algebra** is a vector space $\mathfrak{g}$ on $\mathbb{K}(=\mathbb{R},\mathbb{C})$ endowed with a bilinear operation $(x,y)\mapsto[x,y]$ from $\mathfrak{g}\times\mathfrak{g}$ with the properties*

(1) $[x,y]=-[y,x]$

(2) $\big[x,[y,z]\big]+\big[y,[z,x]\big]+\big[z,[x,y]\big]=0$.

*The second condition is the **Jacobi identity**.*

DEFooHINCooYoxPFj

Definition 51.2 (Representation of a Lie algebra[845])**.**
*If $\mathfrak{g}$ is a Lie algebra, a **representation** is a map a vector space V with a linear map $\rho\colon\mathfrak{g}\to\operatorname{End}(V)$ such that*

$$\rho\big([X,Y]\big)=\rho(X)\rho(Y)-\rho(Y)\rho(X) \tag{51.1}$$

for every $X,Y\in\mathfrak{g}$.

Notice that a representation of a Lie algebra takes its value in $\operatorname{End}(V)$ instead of $\operatorname{GL}(V)$. The reason is that $\operatorname{End}(V)$ is a Lie algebra while $\operatorname{GL}(V)$ is not.

DEFooDUEUooZLhKdv

Definition 51.3.
***derivation** of $\mathfrak{a}$:*

$$D[X,Y]=[DX,Y]+[X,DY] \tag{51.2}$$

for every X, $Y\in\mathfrak{a}$.

LemadhomomadXadYadXY

Lemma 51.4.
The adjoint map is an homomorphism $\operatorname{ad}\colon\mathfrak{g}\to\operatorname{GL}(\mathfrak{g})$. In other terms for every $X,Y\in\mathfrak{g}$ we have

$$\big[\operatorname{ad}(X),\operatorname{ad}(Y)\big]=\operatorname{ad}\big([X,Y]\big) \tag{51.3}$$

as operators on $\mathfrak{g}$. In particular the algebra acts on itself and $\mathfrak{g}$ carries a representation of each of its subalgebra.

Proof. Using the fact that $\operatorname{ad}(X)$ is a derivation and Jacobi, for $Z\in\mathfrak{g}$ we have

$$\begin{aligned}\big[\operatorname{ad}(X),\operatorname{ad}(Y)\big]Z&=\operatorname{ad}(X)\operatorname{ad}(Y)Z-\operatorname{ad}(Y)\operatorname{ad}(X)Z &(51.4a)\\ &=\big[[X,Y],Z\big]+\big[Y,[X,Z]\big]-\big[[Y,X],Z\big]-\big[X,[Y,Z]\big] &(51.4b)\\ &=\operatorname{ad}\big([X,Y]\big)Z. &(51.4c)\end{aligned}$$

□

PROPooSWQSooSEfTuX

Proposition 51.5 ([846])**.**
The set of all the smooth[1] vector fields on a manifold is a Lie algebra[2].

1. Smooth means C^∞.
2. Definition 51.1.

Proof. If X and Y are smooth vector fields, the commutator $[X,Y]$ is still smooth by proposition 49.105. From the definition (49.210), it is apparent that $[X,Y] = -[Y,X]$. The Jacobi identity remains to be proven.

Let X, Y, Z be smooth vector fields. If $f\colon M \to \mathbb{R}$ is a smooth function, we write XYf the function $x \mapsto X_x(Yf)$ and similarly $XYZf$ and other combinations. In particular we have

$$[X,Y]f = X(Yf) - Y(Xf) = XYf - YXf \tag{51.5}$$

which is shorthanded into $[X,Y] = XY - YX$. We also have

$$\big[X,[Y,Z]\big] = [X, YZ - ZY] = XYZ - XZY - YZX + ZYX. \tag{51.6}$$

Making the same with the two other terms we see that

$$\big[X,[Y,Z]\big] + \big[Y,[Z,X]\big] + \big[Z,[X,Y]\big] = 0. \tag{51.7}$$

□

51.1 Adjoint representation

Let G be a Lie group and $g \in G$ we consider the map $\mathbf{Ad}(g)\colon G \to G$ given by $\mathbf{Ad}(g)h = ghg^{-1}$. This is an analytic automorphism of G. We define:

$$\mathrm{Ad}(g) = d\,\mathbf{Ad}(g)_e.$$

Using equation $\varphi(\exp X) = \exp d\varphi_e(X)$ with $\varphi = I(g)$,

$$ge^X g^{-1} = \exp[\mathrm{Ad}(g)X] \tag{51.8}$$ eq:sigma_X_sigma

for every $g \in G$ and $X \in \mathfrak{g}$. The map $g \to \mathrm{Ad}(g)$ is a homomorphism from G to $\mathrm{GL}(\mathfrak{g})$. This homomorphism is called the **adjoint representation** of G.

Proposition 51.6.
The adjoint representation is analytic.

Proof. We have to prove that for any $X \in \mathfrak{g}$ and for any linear map $\omega\colon \mathfrak{g} \to \mathbb{R}$, the function $\omega(\mathrm{Ad}(g)X)$ is analytic at $g = e$. Indeed if we take as ω , the projection to the ith component and X as the jth basis vector ($\mathfrak{g}$ seen as a vector space), and if we see the product $\mathrm{Ad}(g)X$ as a product matrix times vector, $(\mathrm{Ad}(g)X)_i$ is just $\mathrm{Ad}(g)_{ij}$. Then our supposition is the analyticity of $g \to \mathrm{Ad}(g)_{ij}$ at $g = e$.[3]

Now we prove it. Consider $f \in C^\infty(G)$, analytic at $g = e$ and such that $Yf = \omega(Y)$ for any $Y \in \mathfrak{g}$. Using equation (51.8),

$$\omega(\mathrm{Ad}(g)X) = (\mathrm{Ad}(g)X)f = \frac{d}{dt}\Big[f(e^{t\,\mathrm{Ad}(g)X})\Big]_{t=0} = \frac{d}{dt}\Big[f(ge^{tX}g^{-1})\Big]_{t=0}, \tag{51.9}$$

which is well analytic at $g = e$. □

Proposition 51.7.
Let G be a connected Lie group and H, an analytic subgroup of G. Then H is a normal subgroup of G if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Proof. We consider X, $Y \in \mathfrak{g}$. Formula $\exp tX \exp tY \exp -tY = \exp(tY + t^2[X,Y] + o(t^3))$ and equation (51.8) give

$$\exp\Big(\mathrm{Ad}(e^{tX})tY\Big) = \exp\Big(tY + t^2[X,Y] + o(t^3)\Big).$$

3. L'analicité de Ad, elle vient par prolongement analytique depuis juste un point ?

Since it is true for any X, $Y \in \mathfrak{g}$, $\mathrm{Ad}(e^{tX})tY = tY + t^2[X,Y]$; thus

$$\mathrm{Ad}(e^{tX}) = \mathbb{1} + t[X,Y] + o(t^2). \tag{51.10}$$

Since we know that $d\,\mathrm{Ad}_e\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a homomorphism (Ad is seen as a map $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$), taking the derivative of the last equation with respect to t gives

$$d\,\mathrm{Ad}_e(X) = \mathrm{ad}\,X. \tag{51.11}$$

Then $\mathrm{Ad}(e^X) = e^{\mathrm{ad}\,X}$. Since is connected, an element of G can be written as $\exp X$ for a certain $X \in \mathfrak{g}$[4]. The purpose is to prove that $g \exp X g^{-1} = \exp(\mathrm{Ad}(g)X)$ remains in H for any $g \in G$ if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. In other words, we want $\mathrm{Ad}(g)X \in \mathfrak{h}$ if and only if $\mathfrak{h}$ is an ideal. We can write $g = e^Y$ for a certain $Y \in \mathfrak{g}$. Thus

$$\mathrm{Ad}(g)X = \mathrm{Ad}(e^Y)X = e^{\mathrm{ad}\,Y}X.$$

Using the expansion

$$e^{\mathrm{ad}\,Y} = \sum_k \frac{1}{k!}(\mathrm{ad}\,Y)^k, \tag{51.12}$$

we have the thesis. □

51.2 Representation of the complex algebra

When $\mathfrak{g}$ is a real Lie algebra, the corresponding complex Lie algebra is defined as

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}. \tag{51.13}$$

We have, as an example, in the lemma 55.63 the equality $\mathfrak{su}(2)_{\mathbb{C}} = \mathfrak{sl}(2,\mathbb{C})$.

LEMooIGAFooTSUsJR

Lemma 51.8 ([1])**.**
Let $\mathfrak{g}$ be a real Lie algebra and $\rho\colon \mathfrak{g} \to \mathrm{GL}(V)$ be an irreducible representation. There exists an irreducible representation $\rho'\colon \mathfrak{g}_{\mathbb{C}} \to \mathrm{GL}(V)$ such that $\rho = \rho'|_{\mathfrak{g}}$.

Proof. We prove that the map

$$\begin{aligned} \rho'\colon \mathfrak{g}_{\mathbb{C}} &\to \mathrm{GL}(V) \\ X + iY &\mapsto \rho(X) + i\rho(Y) \end{aligned} \tag{51.14}$$

is the representation we are searching for.

(1) **This is a representation** Using the linearity, $[X+iY, Z+iT] = [X,Z] - [Y,T] + i[X,T] + i[Y,Z]$. On the one hand we have

$$\rho'\big([X+iY, Z+iT]\big) = \rho\big([X,Z] - [Y,T]\big) + i\rho\big([X,T] + [Y,Z]\big); \tag{51.15}$$

while on the other hand,

$$\begin{aligned} \big[\rho'(X+iY), \rho'(Z+iT)\big] &= \big[\rho(X) + i\rho(Y), \rho(Z) + i\rho(T)\big] && \text{(51.16a)} \\ &= \big[\rho(X), \rho(Z)\big] + i\big[\rho(X), \rho(T)\big] && \text{(51.16b)} \\ &\quad + i\big[\rho(Y), \rho(Z)\big] - \big[\rho(Y), \rho(T)\big] && \text{(51.16c)} \\ &= \rho\big([X,Z]\big) - \rho\big([Y,T]\big) + i\rho\big([X,T] + [Y,Z]\big). && \text{(51.16d)} \end{aligned}$$

This proves that the map ρ' commutes with the Lie bracket, so that ρ' is a representation.

(2) **Irreducible** Let W be a subspace of V and suppose that W is invariant under ρ'. This means that, for every $X, Y \in \mathfrak{g}_{\mathbb{C}}$, we have $\rho'(X+iY)W \subset W$. In particular, with $Y = 0$ we have

$$\rho(X)W = \rho'(X)W \subset W. \tag{51.17}$$

This shows that W is invariant under ρ. Since ρ is irreducible, the subspace W must be $\{0\}$ or V. Thus ρ' is irreducible.

□

4. Because G is generated by any neighbourhood of e and there exists such a neighbourhood of e which is diffeomorphic to a subset of $\mathfrak{g}$ by exp.

51.3 Jordan decomposition

If V is a finite dimensional space, a subspace W in V is **invariant** under a subset $G \subset \mathrm{Hom}(V,V)$ if $sW \subset W$ for any $s \in G$. The space V is **irreducible** when V and $\{0\}$ are the only two invariant subspaces. The set G is **semisimple** if any invariant subspace has an invariant complement. In this case, the vector space split into $V = \sum_i V_i$ with V_i invariant and irreducible.

Theorem 51.9 (Jordan decomposition)**.**
Any element $A \in \mathrm{Hom}(V,V)$ is decomposable in one and only one way as $A = S + N$ with S semisimple and N nilpotent and $NS = SN$. Furthermore, S and N are polynomials in A. More precisely:

If V is a complex vector space and $A \in \mathrm{Hom}(V,V)$ with $\lambda_1, \dots, \lambda_r$ his eigenvalues, we pose

$$V_i = \{v \in V \text{ tel que } (A - \lambda_i \mathbb{1})^k v = 0 \text{ for large enough } k\}.$$

Then tho:jordan

(1) $V = \sum_{i=1}^r V_i$,

(2) each V_i is invariant under A,

(3) the semisimple part of A is given by

$$S(\sum_{i=1}^r v_i) = \sum_{i=1}^r \lambda_i v_i,$$

for $v_i \in V_i$,

(4) the characteristic polynomial of A is

$$\det(\lambda \mathbb{1} - A) = (\lambda - \lambda_1)^{d_1} \dots (\lambda - \lambda_r)^{d_r}$$

where $d_i = \dim V_i$ $(1 \leqslant i \leqslant r)$.

Now we give a great theorem without proof.

Theorem 51.10 (Jordan decomposition)**.**
Let V be a finite dimensional vector space and $x \in \mathrm{End}\, V$.

(1) There exists one and only one choice of $x_s, x_n \in \mathrm{End}(V)$ such that $x = x_s + x_n$, x_s is semisimple, n_n is nilpotent and $[x_s, x_n] = 0$.

(2) There exists polynomials p and q without independent term such that $x_s = p(x)$, $x_n = q(x)$; in particular if $y \in \mathrm{End}\, V$ commutes with x, then it commutes with x_s and x_n.

(3) If $A \subset B \subset V$ are subspaces of V and if $x(B) \subset A$, then $x_s(B) \subset A$ and $x_n(B) \subset A$.

prop:Jordan_decomp
lem:Jordan_ad

Lemma 51.11.
Let $x \in \mathrm{End}\, V$ with his Jordan decomposition $x = x_s + x_n$. Then the Jordan decomposition of $\mathrm{ad}\, x$ is

$$\mathrm{ad}\, x = \mathrm{ad}\, x_s + \mathrm{ad}\, x_n. \tag{51.18}$$

eq:ad_x_xs_xn

Proof. We already know that $\mathrm{ad}\, x_s$ is semisimple and $\mathrm{ad}\, x_n$ is nilpotent. They commute because $[\mathrm{ad}\, x_s, \mathrm{ad}\, x_n] = \mathrm{ad}[x_s, x_n] = 0$. Then the unicity part of Jordan theorem 51.10 makes (51.18) the Jordan decomposition of $\mathrm{ad}\, x$. □

lem:M_nil

Lemma 51.12.
Let $A \subset B$ be two subspace of $\mathfrak{gl}(V)$ with $\dim V < \infty$. We pose

$$M = \{x \in \mathfrak{gl}(V) \text{ tel que } [x, B] \subset A\},$$

and we suppose that $x \in M$ verify $\mathrm{Tr}(x \circ y) = 0$ for all $y \in M$. Then x is nilpotent.

Proof. We use the Jordan decomposition $x = x_s + x_n$ and a basis in which x_s takes the form $diag(a_1, \dots, a_m)$; let $\{v_1, \dots, v_m\}$ be this basis. We denotes by E the vector space on $\mathbb{Q}$ spanned by $\{a_1, \dots, a_m\}$. We want to prove that $x_s = 0$, i.e. $E = 0$. Since E has finite dimension, it is equivalent to prove that its dual is zero. In other words, we have to see that any linear map $f : E \to \mathbb{Q}$ is zero.

We consider $y \in \mathfrak{gl}(V)$, an element whose matrix is $diag(f(a_1), \dots, f(a_m))$ and (E_{ij}), the usual basis of $\mathfrak{gl}(V)$. We know that

$$(\operatorname{ad} x_s)E_{ij} = (a_i - a_j)E_{ij}, \tag{51.19a}$$
$$(\operatorname{ad} y)E_{ij} = (f(a_i) - f(a_j))E_{ij}. \tag{51.19b}$$

It is always possible to find a polynomial r on $\mathbb{R}$ without constant term such that $r(a_i - a_j) = f(a_i) - f(a_j)$. Note that this is well defined because of the linearity of f: if $a_i - a_j = a_k - a_l$, then $f(a_i) - f(a_j) = f(a_k) - f(a_l)$. Since $\operatorname{ad} x_s$ is diagonal, $r(\operatorname{ad} x_s)$ is the matrix with $r(\operatorname{ad} x_s)_{ii}$ on the diagonal and zero anywhere else. Then $r(\operatorname{ad} x_s) = \operatorname{ad} y$. By lemma 51.11, $\operatorname{ad} x_s$ is the semisimple part of $\operatorname{ad} x$, then $\operatorname{ad} y$ is a polynomial without constant term with respect to $\operatorname{ad} x$ (second point of theorem 51.10).

Since $(\operatorname{ad} y)B \subset A$, $y \in M$ and $\operatorname{Tr}(xy) = 0$. It is easy to convince ourself that the s_n part of x will not contribute to the trace because x_n is strictly upper triangular and y is diagonal. From the explicit forms of x_s and y,

$$\operatorname{Tr}(xy) = \sum_i a_i f(a_i) = 0.$$

This is a $\mathbb{Q}$-linear combination of element of E: we have to see it as a_i being a basis vector and $f(a_i)$ a coefficient, so that we can apply f on both sides to find $0 = \sum_i f(a_i)^2$. Then for all i, $f(a_i) = 0$, so that $f = 0$ because the a_i spans E. □

Definition 51.13 (semisimple endomorphism)**.**
pg:def_semisimple
*If V is a finite dimensional vector space, we say that an element $u \in \operatorname{End} V$ is **semisimple** if every u-invariant subspace of V has a complementary u-invariant.*

Proposition 51.14 ([847])**.**
If V a vector space over an algebraically closed field, an endomorphism is semisimple if and only if it is diagonalizable.

pg:E_ij
Let E_{kl} be the $(n+2) \times (n+2)$ matrix with a 1 at position (k, l) and 0 anywhere else: $(E_{kl})_{ij} = \delta_{ki}\delta_{lj}$. An easy computation show that

$$E_{kl}E_{ab} = \delta_{la}E_{kb}, \tag{51.20}$$
EqFormMulEmtr

and

$$[E_{kl}, E_{rs}] = \delta_{lr}E_{ks} - \delta_{sk}E_{rl}. \tag{51.21}$$
comm_de_E

51.4 Killing form

Definition 51.15.
*The **Killing form** of $\mathcal{G}$ is the symmetric bilinear form:*

$$B(X, Y) = Tr(\operatorname{ad} X \circ \operatorname{ad} Y). \tag{51.22}$$

Proposition 51.16.
*It is **invariant** in the sense of*

$$B\big((\operatorname{ad} S)X, Y\big) = -B\big(X, (\operatorname{ad} S)Y\big), \tag{51.23}$$
eq:Killing_invariant

$\forall X, Y, S \in \mathcal{G}$.

PropAutomInvarB

Proposition 51.17.
If $\varphi\colon \mathcal{G} \to \mathcal{G}$ is an automorphism of $\mathcal{G}$, then

$$B(\varphi(X), \varphi(Y)) = B(X, Y).$$

prop:auto_2

Proof. The fact that φ is an automorphism of $\mathcal{G}$ is written as $\varphi \circ \operatorname{ad} X = \operatorname{ad}(\varphi(X)) \circ \varphi$, or

$$\operatorname{ad}(\varphi(X)) = \varphi \circ \operatorname{ad} X \circ \varphi^{-1}.$$

Then

$$\begin{aligned} \operatorname{Tr}(\operatorname{ad}(\varphi(X)) \circ \operatorname{ad}(\varphi(Y))) &= \operatorname{Tr}(\varphi \circ \operatorname{ad} X \circ \varphi^{-1} \circ \varphi \operatorname{ad} Y \circ \varphi^{-1}) \\ &= \operatorname{Tr}(\operatorname{ad} X \circ \operatorname{ad} Y). \end{aligned} \tag{51.24}$$

□

Remarque 51.18.
The Killing 2-form is a map $B\colon \mathcal{G} \times \mathcal{G} \to \mathbb{R}$. When we say that it is preserved by a map $f\colon G \to G$, we mean that it is preserved by df: $B(df\ \cdot\ , df\ \cdot\ ) = B(\ \cdot\ ,\ \cdot\ )$.

Remarque 51.19.
The Killing form is *a priori* only defined on $\mathcal{G} = T_eG$. For A, $B \in T_gG$, one naturally defines

$$B_g(A, B) = B(dL_{g^{-1}} A, dL_{g^{-1}} B). \tag{51.25}$$

This assures the left invariance of B. Now we prove the right invariance.

An other important property of the Killing form is its bi-invariance.

Lemma 51.20.
Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{i}$ an ideal in $\mathfrak{g}$. Let $B\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ be the Killing form on $\mathfrak{g}$ and $B'\colon \mathfrak{i} \times \mathfrak{i} \to \mathbb{R}$, the one of $\mathfrak{i}$. Then $B' = B|_{\mathfrak{i}\times\mathfrak{i}}$, i.e. the Killing form on $\mathfrak{g}$ descent to the ideal $\mathfrak{i}$. lem:Killing_descent_ideal

Proof. If W is a subspace of a (finite dimensional) vector space V and $\phi\colon V \to W$ and endomorphism, then $\operatorname{Tr}\phi = \operatorname{Tr}(\phi|_W)$. Indeed, if $\{X_1, \dots, X_n\}$ is a basis of V such that $\{X_1, \dots, X_r\}$ is a basis of W, the matrix element ϕ_{kk} is zero for $k > r$. Then

$$\operatorname{Tr}\phi = \sum_{i=1}^{n} \phi_{ii} = \sum_{i=1}^{r} \phi_{ii} = \operatorname{Tr}(\phi|_W).$$

Now consider X, $Y \in \mathfrak{i}$; $(\operatorname{ad} X \circ \operatorname{ad} Y)$ is an endomorphism of $\mathfrak{g}$ which sends $\mathfrak{g}$ to $\mathfrak{i}$ (because $\mathfrak{i}$ is an ideal). Then

$$B'(X, Y) = \operatorname{Tr}\left((\operatorname{ad} X \circ \operatorname{ad} Y)|_{\mathfrak{i}}\right) = \operatorname{Tr}(\operatorname{ad} X \circ \operatorname{ad} Y) = B(X, Y).$$

□

51.5 Solvable and nilpotent algebras

If $\mathfrak{g}$ is a Lie algebra, the **derived Lie algebra** is

$$\mathscr{D}\mathfrak{g} = \operatorname{Span}\{[X, Y] \text{ tel que } X, Y \in \mathfrak{g}\}.$$

We naturally define $\mathscr{D}^0\mathfrak{g} = \mathfrak{g}$ and $\mathscr{D}^n\mathfrak{g} = \mathscr{D}(\mathscr{D}^{n-1}\mathfrak{g})$ this is the **derived series**. Each $\mathscr{D}^n\mathfrak{g}$ is an ideal in $\mathfrak{g}$. We also define the **central decreasing sequence** by $\mathfrak{a}^0 = \mathfrak{a}$, $\mathfrak{a}^{p+1} = [\mathfrak{a}, \mathfrak{a}^p]$.

Definition 51.21.
*The Lie algebra $\mathfrak{g}$ is **solvable** if there exists a $n \geqslant 0$ such that $\mathscr{D}^n\mathfrak{g} = \{0\}$. A Lie group is solvable when its Lie algebra is.*

*The Lie algebra $\mathfrak{g}$ is **nilpotent** if $\mathfrak{g}^n = 0$ for some n. We say that $\mathfrak{g}$ is* ad*-nilpotent if* $\mathrm{ad}(X)$ *is a nilpotent endomorphism of $\mathfrak{g}$ for each $X \in \mathfrak{g}$.*

Do not confuse *nilpotent* and *solvable* algebras. A nilpotent algebra is always solvable, while the algebra spanned by $\{A, B\}$ with the relation $[A, B] = B$ is solvable but not nilpotent.

If $\mathfrak{g} \neq \{0\}$ is a solvable Lie algebra and if n is the smallest natural such that $\mathscr{D}^n\mathfrak{g} = \{0\}$, then $\mathscr{D}^{n-1}\mathfrak{g}$ is a non zero abelian ideal in $\mathfrak{g}$. We conclude that a solvable Lie algebra is never semisimple (because the center of a semisimple Lie algebra is zero).

A Lie algebra is said to fulfil the **chain condition** if for every ideal $\mathfrak{h} \neq \{0\}$ in $\mathfrak{g}$, there exists an ideal $\mathfrak{h}_1$ in $\mathfrak{h}$ with codimension 1.

Lemma 51.22.
A Lie algebra is solvable if and only if it fulfils the chain condition.

Proof. Necessary condition. The Lie algebra $\mathfrak{g}$ is solvable (then $\mathscr{D}\mathfrak{g} \neq \mathfrak{g}$) and $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. We consider $\mathfrak{h}_1$, a subspace of codimension 1 in $\mathfrak{h}$ which contains $\mathscr{D}\mathfrak{h}$. It is clear that $\mathfrak{h}_1$ is an ideal in $\mathfrak{h}$ because $[H_1, H] \in \mathscr{D}\mathfrak{h} \subset \mathfrak{h}_1$.
Sufficient condition. We have a sequence

$$\{0\} = \mathfrak{g}_n \subset \mathfrak{g}_{n-1} \subset \ldots \subset \mathfrak{g}_0 = \mathfrak{g} \tag{51.26}$$

eq:solvable_chaine

where $\mathfrak{g}_r$ is an ideal of codimension 1 in $\mathfrak{g}_{r-1}$. Let A be the unique vector in $\mathfrak{g}_{r-1}$ which don't belong to $\mathfrak{g}_r$. When we write $[X, Y]$ with X, $Y \in \mathfrak{g}_{r-1}$, at least one of X or Y is not A (else, it is zero) then at least one of the two is in $\mathfrak{g}_r$. But $\mathfrak{g}_r$ is an ideal; then $[X, Y] \in \mathfrak{g}_r$. Thus $\mathscr{D}(\mathfrak{g}_{r-1}) \subset \mathfrak{g}_r$ and

$$\mathscr{D}^n\mathfrak{g} = \mathscr{D}^{n-1}\mathscr{D}\mathfrak{g} \subset \mathscr{D}^{n-1}\mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_n = 0.$$

□

tho:Lie_Vu

Theorem 51.23 (Lie theorem)**.**
Consider $\mathfrak{g}$, a real (resp. complex) solvable Lie algebra and a real (resp. complex) vector space $V \neq \{0\}$. If $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a homomorphism, then there exists a non zero vector in V which is eigenvector of all the elements of $\pi(\mathfrak{g})$.

¡¡ Avertissement/question au lecteur !! 51.24
It is strange to be stated for real and complex Lie algebras. Following [848], this is only true for complex Lie algebras while there exists other versions for reals ones.

Proof. Let us do it by induction on the dimension of $\mathfrak{g}$. We begin with $\dim\mathfrak{g} = 1$. In this case, π is just a map $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ such that $\pi(aX) = a\pi(X)$. We have to find an eigenvector for the homomorphism $\pi(X)\colon V \to V$. Such a vector exists from the Jordan decomposition 51.9. Indeed, if there are no eigenvectors, there are no spaces V_i and the decomposition $V = \sum V_i$ can't be true.

Now we consider a general solvable Lie algebra $\mathfrak{g}$ and we suppose that the theorem is true for any solvable Lie algebra with dimension less that $\dim\mathfrak{g}$. Since $\mathfrak{g}$ is solvable, there exists an ideal $\mathfrak{h}$ of codimension 1 in $\mathfrak{g}$; then there exists a $e_0 \neq 0 \in V$ which is eigenvector of all the $\pi(H)$ with $H \in \mathfrak{h}$. So we have $\lambda\colon \mathfrak{h} \to \mathbb{R}$ naturally defined by

$$\pi(H)e_0 = \lambda(H)e_0.$$

Now we consider $X \in \mathfrak{g}\backslash\mathfrak{h}$ and $e_{-1} = 0$, $e_p = \pi(X)^p e_0$ for $p = 1, 2, \ldots$ We will show that $\pi(H)e_p = \lambda(H)e_p \mod (e_0, \ldots, e_{p-1})$ for all $H \in \mathfrak{h}$ and $p \geqslant 0$. It is clear for $p = 0$. Let us suppose that it is

true for p. Then

$$\begin{aligned}\pi(H)e_{p+1} &= \pi(H)\pi(X)e_p \\ &= \pi([H,X])e_p + \pi(X)\pi(H)e_p \\ &= \lambda([H,X])e_p + \pi(X)\lambda(H)e_p \\ &\quad \mod (e_0,\dots,e_{p-1},\pi(X)e_0,\dots,\pi(X)e_{p-1}).\end{aligned} \tag{51.27}$$

But we can put $\pi([H,X])$ and $\pi(X)e_i$ into the modulus. Thus we have

$$\pi(H)e_{p+1} = \lambda(H)e_{p+1} \mod (e_0,\dots,e_p).$$

Now we consider the subspace of V given by $W = \text{Span}\{e_p\}_{p=1,\dots}$. The algebra $\pi(\mathfrak{h})$ leaves W invariant and our induction hypothesis works on $(\pi(\mathfrak{h}), W)$; then one can find in W a common eigenvector for all the $\pi(H)$. This vector is the one we were looking for. □

Corollary 51.25.
Let $\mathfrak{g}$ be a solvable Lie group and π a representation of $\mathfrak{g}$ on a finite dimensional vector space V. Then there exists a basis $\{e_1,\dots,e_n\}$ of V in which all the endomorphism $\pi(X)$, $X \in \mathfrak{g}$ are upper triangular matrices. cor:de_Lie_Vu

Proof. Consider $e_1 \neq 0 \in V$, a common eigenvector of all the $\pi(X)$, $X \in \mathfrak{g}$. We consider $E_1 = \text{Span}\{e_1\}$. The representation π induces a representation π_1 of $\mathfrak{g}$ on the space V/E_1. If $V/E_1 \neq \{0\}$, we have a $e_2 \in V$ such that $(e_2 + E_1) \in V/E_1$ is an eigenvector of all the $\pi_($X).

In this manner, we build a basis $\{e_1,\dots,e_n\}$ of V such that $\pi(X)e_i = 0 \mod (e_1,\dots,e_i)$ for all $X \in \mathfrak{g}$. In this basis, $\pi(X)$ has zeros under the diagonal. □

Theorem 51.26.
Let V be a real or complex vector space and $\mathfrak{g}$, a subalgebra of $\mathfrak{gl}(V)$ made up with nilpotent elements. Then

(1) $\mathfrak{g}$ is nilpotent;

(2) $\exists v \neq 0$ in V such that $\forall Z \in \mathfrak{g}$, $Zv = 0$;

(3) There exists a basis of V in which the elements of $\mathfrak{g}$ are matrices with only zeros under the diagonal.

tho:trois_nil

Proof. First item. We consider a $Z \in \mathfrak{g}$ and we have to see that $\text{ad}_{\mathfrak{g}} Z$ is a nilpotent endomorphism of $\mathfrak{g}$. Be careful on a point: an element X of $\mathfrak{g}$ is nilpotent as endomorphism of V while we want to prove that $\text{ad}\, X$ is nilpotent as endomorphism of $\mathfrak{g}$. We denote by L_Z and R_Z, the left and right multiplication; since we are in a matrix algebra, the bracket is given by the commutator: $\text{ad}\, Z = L_Z - R_Z$. We have

$$(\text{ad}\, Z)^p(X) = \sum_{i=0}^{p} (-1)^p \binom{p}{i} Z^{p-i} X Z^i \tag{51.28}$$

There exists a $k \in \mathbb{N}$ such that $Z^k = 0$. For this k, $(\text{ad}\, Z)^{2k+1}$ is a sum of terms of the form $Z^{p-i}XZ^i$: either $p-i$ either i is always bigger than k. But $\text{ad}_{\mathfrak{g}} Z$ is the restriction of $\text{ad}\, Z$ (which is defined on $\mathfrak{gl}(V)$) to $\mathfrak{g}$. Then $\mathfrak{g}$ is nilpotent.

Second item. Let $r = \dim \mathfrak{g}$. If $r = 1$, we have only one $Z \in \mathfrak{g}$ and $Z^k = 0$ for a certain (minimal) $k \in \mathbb{N}$. We take v such that $w = Z^{k-1}v \neq 0$ (this exists because k is the minimal natural with $Z^k = 0$). Then $Zw = 0$.

Now we suppose that the claim is valid for any algebra with dimension less than r. Let $\mathfrak{h}$ be a strict subalgebra of $\mathfrak{g}$ with maximal dimension. If $H \in \mathfrak{h}$, $\text{ad}_{\mathfrak{g}} H$ is a nilpotent endomorphism of $\mathfrak{g}$ which sends $\mathfrak{h}$ onto itself. Thus $\text{ad}_{\mathfrak{g}} H$ induces a nilpotent endomorphism H^* on the vector space $\mathfrak{g}/\mathfrak{h}$. We consider the set $\mathcal{A} = \{H^* \text{ tel que } H \in \mathfrak{h}\}$; this is a subalgebra of $\mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$ made up with nilpotent elements which has dimension strictly less than r.

The induction assumption gives us a non zero $u \in \mathfrak{g}/\mathfrak{h}$ which is sent to 0 by all $\mathcal{A}$, i.e. $(\mathrm{ad}_{\mathfrak{g}} H)u = 0$ in $\mathfrak{g}/\mathfrak{h}$. In other words, $u \in \mathfrak{g}\backslash\mathfrak{h}$ is such that $(\mathrm{ad}_{\mathfrak{g}} H)u \in \mathfrak{h}$.

The space $\mathfrak{h} + \mathbb{K}X$ (here, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$) of $\mathfrak{g}$ is a subalgebra of $\mathfrak{g}$. Indeed, with obvious notations,

$$[H + kX, H' + k'X] = [H, H'] + \mathrm{ad}\, H(k'X) - \mathrm{ad}\, H'(kX) + kk'[X, X]. \tag{51.29}$$

eq:H_k_X

The first term lies in $\mathfrak{h}$ because it is a subalgebra; the second and third therms belongs to $\mathfrak{h}$ by definition of X. The last term is zero. Since $\mathfrak{h}$ is maximal, $\mathfrak{h} + \mathbb{K}X = \mathfrak{g}$. Then (51.29) shows that $\mathfrak{h}$ is also an ideal. Now we consider

$$W = \{e \in V \text{ tel que } \forall H \in \mathfrak{h}, He = 0\}.$$

Since $\dim \mathfrak{h} < r$, $W \neq \{0\}$ from our induction assumption. Furthermore, for $e \in W$, $HXe = [H, X]e + XHe = 0$. Then $X \cdot W \subset W$. The restriction of X to W is nilpotent. Then there exists a $v \in W$ such that $Xv = 0$. For him $Hv = 0$ because $v \in W$ and $Xv = 0$ by definition of X. Then $Gv = 0$ for any $G \in \mathfrak{h} + \mathbb{K}X = \mathfrak{g}$.

Third item. Let e_1 be a non zero vector in V such that $Ze_1 = 0$ for any $Z \in \mathfrak{g}$ (the existence comes from the second item). We consider $E_1 = \mathrm{Span}\, e_1$. Any $Z \in \mathfrak{g}$ induces a nilpotent endomorphism Z^* on the vector space V/E_1. If $V/E_1 \neq \{0\}$, we take a $e_2 \in V\backslash E_1$ such that $e_2 + E_1 \in V/E_1$ fulfils $Z^*(e_2 + E_1) = 0$ for all $Z \in \mathfrak{g}$. By going on so, we have $Ze_1 = 0$, $Ze_i = 0 \mod (e_1, \dots, e_{i-1})$. In this basis, the matrix of Z has zeros on and under the diagonal. □

Corollary 51.27.
Let us consider V, a finite dimensional vector space on $\mathbb{K}$ and $\mathfrak{g}$, a subalgebra of $\mathfrak{gl}(V)$ made up with nilpotent elements. Then if $s \geqslant \dim V$ and $X_i \in \mathfrak{g}$, we have $X_1 X_2 \dots X_s = 0$. cor:nil_XXX

Proof. We write the X_i's in a basis where they have zeros on and under the diagonal. It is rather easy to see that each product push the non zero elements into the upper right corner. □

Corollary 51.28.
A nilpotent algebra is solvable.

Proof. The algebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g})$ made up with nilpotent endomorphisms of $\mathfrak{g}$. The product of s (see notations of previous corollary) such endomorphism is zero. In particular $\mathfrak{g}$ is solvable. □

We recall the definition of the central decreasing sequence: $\mathfrak{a}^0 = \mathfrak{a}$, $\mathfrak{a}^{p+1} = [\mathfrak{a}, \mathfrak{a}^p]$.

Corollary 51.29.
A Lie algebra $\mathfrak{a}$ is nilpotent if and only if $\mathfrak{a}^m = \{0\}$ for $m \geqslant \dim \mathfrak{a}$. cor:nil_Gn

Proof. The direct sense is easy: we use corollary 51.27 with $\mathfrak{g} = \mathrm{ad}(\mathfrak{a})$ ($\dim \mathfrak{g} = \dim \mathfrak{a}$). Since $\mathfrak{g}$ is nilpotent, for any $X_i \in \mathfrak{g}$ we have $X_1 \dots X_s$, so that $\mathfrak{a}^m = 0$. The inverse sense is trivial.

□

Corollary 51.30.
A nilpotent Lie algebra $\mathfrak{a} \neq \{0\}$ has a non zero center

Proof. If m is the smallest natural such that $\mathfrak{a}^m = 0$, $\mathfrak{a}^{m-1}$ is in the center. □

Lemma 51.31.
If $\mathfrak{i}$ and $\mathfrak{j}$ are ideals in $\mathfrak{g}$, then we have a canonical isomorphism $\psi\colon (\mathfrak{i} + \mathfrak{j})/\mathfrak{j} \to \mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$ given by

$$\psi([x]) = \bar{i}$$

if $x = i + j$ with $i \in \mathfrak{i}$ and $j \in \mathfrak{j}$. Here classes with respect to $\mathfrak{j}$ are denoted by $[.]$ and the one with respect to $(\mathfrak{i} \cap \mathfrak{j})$ by a bar. lem:pre_trois_resoluble

Proof. We first have to see that ψ is well defined. If $x' = i + j + j'$, $\psi([x]) = \bar{i}$ because $j + j' \in \mathfrak{j}$. If $x = i' + j'$ (an other decomposition for $x = i + j$), $\bar{i} = \bar{j}$, $j' - j = i - i' \in \mathfrak{j} \cap \mathfrak{i}$. Then $\bar{i} = \overline{i' + j' - j} = \bar{i'}$.

Now it is easy to see that ψ is a homomorphism. □

Proposition 51.32.
Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras.

(1) If $\mathfrak{g}$ is solvable then any subalgebra is solvable and if $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a Lie algebra homomorphism, then $\phi(\mathfrak{g})$ is solvable in $\mathfrak{g}'$.

(2) If $\mathfrak{i}$ is a solvable ideal in $\mathfrak{g}$ such that $\mathfrak{g}/\mathfrak{i}$ is solvable, then $\mathfrak{g}$ is solvable.

(3) If $\mathfrak{i}$ and $\mathfrak{j}$ are solvable ideals in $\mathfrak{g}$, then $\mathfrak{i} + \mathfrak{j}$ is also a solvable ideal in $\mathfrak{g}$.

prop:trois_resoluble

Proof. First item. If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $\mathscr{D}^k\mathfrak{h} \subset \mathscr{D}^k\mathfrak{g}$, so that $\mathfrak{h}$ is solvable. Now consider $\mathfrak{h} = \phi(\mathfrak{g}) \subset \mathfrak{g}'$. This is a subalgebra of $\mathfrak{g}'$ because $[h, h'] = [\phi(g), \phi(g')] = \phi([g, g']) \in \mathfrak{h}$. It is clear that $\mathscr{D}(\phi(\mathfrak{g})) \subset \phi(\mathscr{D}(\mathfrak{g}))$ and

$$\mathscr{D}^2(\phi(\mathfrak{g})) = \mathscr{D}\big(\mathscr{D}\phi(\mathfrak{g})\big) \subset \mathscr{D}(\phi\mathscr{D}(\mathfrak{g})) \subset \phi\mathscr{D}\mathscr{D}(\mathfrak{g}) = \phi(\mathscr{D}^2(\mathfrak{g})). \tag{51.30}$$

Repeating this argument, $\mathscr{D}^k(\mathfrak{h}) \subset \phi(\mathscr{D}^k\mathfrak{g})$. So $\mathfrak{h}$ is also solvable. Note that

$$\phi([g, g']) = [\phi(g), \phi(g')] \subset \mathscr{D}(\pi(\mathfrak{g})). \tag{51.31}$$

Then

$$\mathscr{D}^k\pi(\mathfrak{g}) = \pi(\mathscr{D}^k\mathfrak{g}). \tag{51.32}$$

Second item. Let n be the smallest integer such that $\mathscr{D}^n(\mathfrak{g}/\mathfrak{i}) = 0$; we look at the canonical homomorphism $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{i}$. This satisfies $\mathscr{D}^n(\pi(\mathfrak{g})) = \pi(\mathscr{D}^n\mathfrak{g}) = 0$. Then $\mathscr{D}^n(\mathfrak{g}) \subset \mathfrak{i}$. If $\mathscr{D}^m\mathfrak{i} = 0$, then $\mathscr{D}^{m+n}\mathfrak{g} = 0$.
Third item. The space $\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$ is the image of $\mathfrak{i}$ by a homomorphism, then it is solvable and $(\mathfrak{i} + \mathfrak{j})/\mathfrak{j}$ is also solvable. The second item makes $\mathfrak{i} + \mathfrak{j}$ solvable. □

Now we consider $\mathfrak{g}$, any Lie algebra and $\mathfrak{s}$ a maximum solvable ideal i.e. it is included in none other solvable ideal. Let us consider $\mathfrak{i}$, an other solvable ideal in $\mathfrak{g}$. Then $\mathfrak{i} + \mathfrak{s}$ is a solvable ideal; since $\mathfrak{s}$ is maximal, $\mathfrak{i} + \mathfrak{s} = \mathfrak{s}$. Thus there exists an unique maximal solvable ideal which we call the **radical** of $\mathfrak{g}$. It will be often denoted by $\operatorname{Rad}\mathfrak{g}$. If β is a symmetric bilinear form, his **radical** is the set

$$S = \{x \in \mathfrak{g} \text{ tel que } \beta(x, y) = 0\ \forall y \in \mathfrak{g}\}. \tag{51.33}$$

The form β is nondegenerate if and only if $S = \{0\}$.

Proposition 51.33.
Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras.

(1) If $\mathfrak{g}$ is nilpotent, then his subalgebras are nilpotent and if $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a Lie algebra homomorphism, then $\phi(\mathfrak{g})$ is nilpotent.

(2) If $\mathfrak{g}/\mathcal{Z}(\mathfrak{g})$ is nilpotent, then $\mathfrak{g}$ is nilpotent. For recall,

$$\mathcal{Z}(\mathfrak{g}) = \{z \in \mathfrak{g} \text{ tel que } [x, z] = 0\ \forall x \in \mathfrak{g}\}.$$

(3) If $\mathfrak{g}$ is nilpotent, then $\mathcal{Z}(\mathfrak{g}) \neq 0$.

prop:nil_homom_nil

Proof. The proof of the first item is the same as the one of 51.32. Now if $(\mathfrak{g}/\mathcal{Z}(\mathfrak{g}))^n = 0$, then $\mathfrak{g}^n/\mathcal{Z}(\mathfrak{g}) = 0$; thus $\mathfrak{g}^n \subset \mathcal{Z}(\mathfrak{g})$, so that $\mathfrak{g}^{n+1} = [\mathfrak{g}, \mathcal{Z}(\mathfrak{g})] = 0$. Finally, if n is the smallest natural such that $\mathfrak{g}^n = 0$, then $[\mathfrak{g}^{n-1}, \mathfrak{g}] = 0$ and $\mathfrak{g}^{n-1} \subset \mathcal{Z}(\mathfrak{g})$. □

The condition to be nilpotent can be reformulated by $\exists n \in \mathbb{N}$ such that $\forall X_i$, $Y \in \mathfrak{g}$,

$$(\operatorname{ad} X_1 \circ \ldots \circ \operatorname{ad} X_n)Y = 0,$$

in particular for any $X \in \mathfrak{g}$, there exists a $n \in \mathbb{N}$ such that $(\operatorname{ad} X)^n = 0$. An element for which such a n exists is **ad-nilpotent**. If $\mathfrak{g}$ is nilpotent, then all his elements are ad-nilpotent.

Some results without proof:

lem:pre_Engel

Lemma 51.34.
If $X \in \mathfrak{gl}(V)$ is a nilpotent endomorphism, then $\operatorname{ad} X$ *is nilpotent.*

Lemma 51.35.
If $x \in \mathfrak{gl}(V)$ is semisimple, then $\operatorname{ad}(x)$ *is also semisimple.*

Proof. We choose a basis $\{v_1, \cdots, v_n\}$ of V in which x is diagonal with eigenvalues $a_1, \ldots, a_n$. For $\mathfrak{gl}(V)$, we consider the basis $\{E_{ij}\}$ in which E_{ij} is the matrix with a 1 at position (i, j) and zero anywhere else. This satisfies $[E_{kl}, E_{rs}] = \delta_{lr}E_{ks} - \delta_{sk}E_{rl}$. We easily check that $E_{kl}(v_i) = \delta_{li}v_k$. Since we are in a matrix algebra, the adjoint action is the commutator: $(\operatorname{ad} x)E_{ij} = [x, E_{ij}]$; as we know that $x = a_k E_{kk}$,

$$(\operatorname{ad} x)E_{ij} = a_k[E_{kk}, E_{ij}] = (a_i - a_j)E_{ij} \tag{51.34}$$

which proves that $\operatorname{ad} x$ has a diagonal matrix in the basis $\{E_{ij}\}$ of $\mathfrak{gl}(V)$. Furthermore, we have an explicit expression for his matrix: the eigenvalues are $(a_i - a_j)$. □

Remarque 51.36.
The inverse implication is not true, as the unit matrix shows.

tho:Engel

Theorem 51.37 (Engel,[848])**.**
A Lie algebra is nilpotent if and only if all his elements are ad-nilpotent.

PropBDrongP

Proposition 51.38.
If an algebra $\mathfrak{g} \subset \mathfrak{gl}(V)$ is made up with nilpotent endomorphisms of V, then $\mathfrak{g}$ is nilpotent as Lie algebra.

PropKillingTraceDeuxn

Proposition 51.39 ([849, 848])**.**
On the Lie algebra $\mathfrak{gl}(\mathbb{R}^n)$, the following formula holds:

$$B(X, Y) = 2n \operatorname{Tr}(XY). \tag{51.35}$$

Proof. We consider a simple subalgebra $\mathfrak{g}$ of $\mathfrak{gl}(V)$ for a certain vector space V and a nondegenerate ad-invariant symmetric 2-form f. Then there exists a $S \in \mathrm{GL}(\mathfrak{g})$ such that

$$f(X, Y) = B(SX, Y) \tag{51.36a}$$
eq:S_un

$$B(SX, Y) = B(X, SY). \tag{51.36b}$$
eq:S_deux

If we consider a basis of $\mathfrak{g}$, we can write $f(X, Y)$ (and the Killing) in a matricial form [5] as

$$f(X, Y) = f_{ij}X^iY^j, \qquad B(X, Y) = B_{ij}X^iY^j.$$

Since B is nondegenerate, we can define the matrix (B^{ij}) by $B^{ij}B_{jk} = \delta^i_k$. It is easy to see that the searched endomorphism of $\mathfrak{g}$ is given by $S^k_l = f_{kj}B^{jl}$.

Using the invariance (51.23) of the Killing form and (51.36b), we find

$$B\big((\operatorname{ad} X \circ S)Y, Z\big) = -B\big((S \circ \operatorname{ad} X)Z, Y\big)$$

5. We systematically use the sum convention on the repeated subscript.

for any X, Y, $Z \in \mathfrak{g}$. Now using (51.36a),

$$f\big((S^{-1} \circ \operatorname{ad} X \circ S)Y, Z\big) = -f\big((\operatorname{ad} X)Z, Y\big) = f\big((\operatorname{ad} Z)X, Y\big) = f\big(Z, (\operatorname{ad} X)Y\big). \tag{51.37}$$

Since f is nondegenerate, we find $\operatorname{ad} X \circ S = S \circ \operatorname{ad} X$. It follows from Schurs'lemma that $S = \lambda I$. Note that $f(X, Y) = \lambda B(X, Y)$; this proves a certain unicity of the Killing form relatively to his invariance properties.

Now we consider $f(X, Y) = \operatorname{Tr}(XY)$. This is symmetric because of the cyclic invariance of the trace and this is *ad*-invariant because of the formula $\operatorname{Tr}([a, b]c) = \operatorname{Tr}(a[b, c])$ which holds for any matrices a, b, c.

The next step to show that f is nondegenerate; we define

$$\mathfrak{g}^{\perp} = \{X \in \mathfrak{g} \text{ tel que } f(X, Y) = 0 \, \forall Y \in \mathfrak{g}\}.$$

The simplicity of $\mathfrak{g}$ ($\mathfrak{g}$ has no proper ideals) makes $\mathfrak{g}$ equal to 0 or $\mathfrak{g}$. Indeed consider $Z \in \mathfrak{g}^{\perp}$. For any X, $Y \in \mathfrak{g}$, we have

$$0 = f(Z, [X, Y]) = f([Z, X], Y).$$

Then $[Z, X] \in \mathfrak{g}^{\perp}$ and $\mathfrak{g}^{\perp}$ is an ideal. We will see that the reality is $\mathfrak{g}^{\perp} = 0$ (cf. error 88.6). Let us suppose $\mathfrak{g}^{\perp} = \mathfrak{g}$ and consider the lemma 51.12 with $A = B = \mathfrak{g}$. We define

$$M = \{X \in \mathfrak{g} \text{ tel que } [X, \mathfrak{g}] \subset \mathfrak{g}\} = \mathfrak{g}.$$

If $X \in M$ satisfies $\operatorname{Tr}(XY) = 0$ for any $Y \in M$, then X is nilpotent. Here, $X \in M$ is not a true condition because $M = \mathfrak{g}$. Since $\mathfrak{g}^{\perp} = \mathfrak{g}$, the trace condition is also trivial. Then $\mathfrak{g}$ is made up with nilpotent endomorphisms of V. Then lemma 51.34 makes all the $X \in \mathfrak{g}$ ad-nilpotent, so that $\mathfrak{g}$ is nilpotent by proposition 51.38.

By the third item of proposition 51.33, $\mathcal{Z}(\mathfrak{g}) \neq 0$ which contradicts the simplicity of $\mathfrak{g}$. Then $\mathfrak{g}^{\perp} = 0$ and f is nondegenerate. Finally,

$$B(X, Y) = \lambda \operatorname{Tr}(X, Y) \tag{51.38}$$

for a certain real number λ. With a certain amount of work, one can determine the exact value of λ when $\mathfrak{g}$ is the Lie algebra of $n \times n$ matrices with vanishing trace. □

51.6 Flags and nilpotent Lie algebras

Here we give a "flag description" of some previous results. In particular the chain (51.26). If V is a vector space of dimension $n < \infty$, a **flag** in V is a chain of subspaces $0 = V_0 \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = \mathfrak{g}$ with $\dim V_k = k$. If $x \in \operatorname{End} V$ fulfils $x(V_i) \subset V_i$, then we say that x **stabilise** the flag.

Theorem 51.40.
If $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(V)$ in which the elements are nilpotent endomorphisms and if $V \neq 0$, then there exists a $v \in V$, $v \neq 0$ such that $\mathfrak{g}v = 0$.

Proof. This is the second item of theorem 51.26. □

Corollary 51.41.
Under the same assumptions, there exists a flag (V_i) stable under $\mathfrak{g}$ such that $\mathfrak{g}V_i \subset V_{i-1}$. In other words, there exists a basis of V in which the matrices of $\mathfrak{g}$ are nilpotent; this basis is the one given by the flag.

Proof. Let $v \neq 0$ such that $\mathfrak{g}v = 0$ which exists by the theorem and $V_1 = \operatorname{Span} v$. We consider $W = V/V_1$; the action of $\mathfrak{g}$ on W is also made up with nilpotent endomorphisms. Then we go on with V_1 and $W_1 = W/V_2$,... □

Lemma 51.42.
If $\mathfrak{g}$ is nilpotent and if $\mathfrak{i}$ is an non trivial ideal in $\mathfrak{g}$, then $\mathfrak{i} \cap \mathcal{Z}(\mathfrak{g}) \neq 0$.

Proof. Since $\mathfrak{i}$ is an ideal, $\mathfrak{g}$ acts on $\mathfrak{i}$ with the adjoint representation. The restriction of an element $\operatorname{ad} X$ for $X \in \mathfrak{g}$ to $\mathfrak{i}$ is in fact a nilpotent element in $\mathfrak{gl}(\mathfrak{i})$. Then we have a $I \in \mathfrak{i}$ such that $\mathfrak{g}I = 0$. Thus $I \in \mathfrak{i} \cap \mathcal{Z}(\mathfrak{g})$. □

Theorem 51.43.
Let $\mathfrak{g}$ be a solvable Lie subalgebra of $\mathfrak{gl}(V)$. If $V \neq 0$, then V posses a common eigenvector for all the endomorphisms of $\mathfrak{g}$. tho:sol_ss_dem

Proof. This is exactly the Lie theorem 51.23 □

Corollary 51.44 (Lie theorem)**.**
Let $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then $\mathfrak{g}$ stabilize a flag of V. tho:Lie_Vd

Proof. This corollary is the corollary given in 51.25.

We consider v_1 the vector given by theorem 51.43. Since it is eigenvector of all $\mathfrak{g}$, $\operatorname{Span} v_1$ is stabilised by $\mathfrak{g}$. Next we consider v_2 in the complementary which is also a common eigenvector,... □

Corollary 51.45.
If $\mathfrak{g}$ is a solvable Lie algebra, then there exists a chain of ideals in $\mathfrak{g}$

$$0 = \mathfrak{g}_0 \subset \mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_n = \mathfrak{g}$$

with $\dim \mathfrak{g}_k = k$.

Proof. If $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a finite-dimensional representation of $\mathfrak{g}$, then $\phi(\mathfrak{g})$ is solvable by proposition 51.33. Then $\phi(\mathfrak{g})$ stabilises a flag of V. Now we take as ϕ the adjoint representation of $\mathfrak{g}$. A stable flag is the chain of ideals; indeed if $\mathfrak{g}_i$ is a part of the flag, then $\forall H \in \mathfrak{g}$ $\operatorname{ad} H \mathfrak{g}_i \subset \mathfrak{g}_i$ because the flag is invariant. □

Corollary 51.46.
If $\mathfrak{g}$ is solvable then $X \in \mathscr{D}\mathfrak{g}$ implies that $\operatorname{ad}_{\mathfrak{g}} X$ is nilpotent. In particular $\mathscr{D}\mathfrak{g}$ is nilpotent.

Proof. We consider the ideals chain of previous corollary and an adapted basis: $\{X_1, \ldots, X_n\}$ is such that $\{X_1, \ldots, X_i\}$ spans $\mathfrak{g}_i$. In such a basis the matrices of $\operatorname{ad}(\mathfrak{g})$ are upper triangular and it is easy to see that in this case, the matrices of $[\operatorname{ad}\mathfrak{g}, \operatorname{ad}\mathfrak{g}]$ are *strictly* upper triangular: they have zeros on the diagonal. But $[\operatorname{ad}\mathfrak{g}, \operatorname{ad}\mathfrak{g}] = \operatorname{ad}_{\mathfrak{g}}[\mathfrak{g}, \mathfrak{g}]$. Then for $X \in \operatorname{ad}_{\mathfrak{g}} \mathscr{D}\mathfrak{g}$, $\operatorname{ad}_{\mathfrak{g}} X$ is nilpotent. *A fortiori*, $\operatorname{ad}_{\mathscr{D}\mathfrak{g}} X$ is nilpotent and by the Engels'theorem 51.37, $\mathscr{D}\mathfrak{g}$ is nilpotent. □

The following lemma is computationally useful because it says that if X is a nilpotent element of a Lie algebra, then $g \cdot X$ is also nilpotent with (at most) the same order.

Lemma 51.47.
The following formula

$$\operatorname{ad}(g \cdot X)^n Y = g \cdot \operatorname{ad}(X)^n (g^{-1} \cdot Y) \tag{51.39}$$

holds for all $g \in G$ and X,$Y \in \mathfrak{g}$, lem:nil_Ad

The proof is a simple induction on n.

51.7 Semisimple Lie algebras

A useful reference to go trough semisimple Lie algebras is [850]. Very few proofs, but the statements of all the useful results with explanations.

Definition 51.48.
*A Lie algebra is **semisimple** if it has no proper abelian invariant Lie subalgebra. A Lie algebra is **simple** if it is not abelian and has no proper Lie subalgebra.*

In that definition, we say that a Lie subalgebra $\mathfrak{h}$ is **invariant** if $\mathrm{ad}(\mathfrak{g})\mathfrak{h} \subset \mathfrak{h}$.

There are a lot of equivalent characterisations. Here are some that are going to be proved (or not) in the next few pages. A Lie algebra is semisimple if an only if one of the following conditions is respected.

(1) The Killing form is nondegenerate.
(2) The radical of $\mathfrak{g}$ is zero (theorem 51.57).
(3) There are no abelian proper invariant subalgebra.

¡¡ Avertissement/question au lecteur !! 51.49
I think that in the following I took the degenerateness of Killing as definition.

The Killing form is a convenient way to define a Riemannian metric on a semisimple[6] Lie group.

Proposition 51.50.
Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{a}$ an ideal in $\mathfrak{g}$, and $\mathfrak{a}^\perp = \{X \in \mathfrak{g} \text{ tel que } B(X,A) = 0 \forall A \in \mathfrak{a}\}$. Then

(1) $\mathfrak{a}^\perp$ is an ideal,
(2) $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$,
(3) $\mathfrak{a}$ is semisimple,

prop:a_aperp

Proof. First item. We have to show that for any $X \in \mathfrak{g}$ and $P \in \mathfrak{a}^\perp$, $[X,P] \in \mathfrak{a}^\perp$, or $\forall\, Y \in \mathfrak{a}$, $B(Y,[X,P]) = 0$. From invariance of B,

$$B(Y,[X,P]) = B(P,[Y,X]) = 0.$$

Second item. Since B is nondegenerate, $\dim \mathfrak{a} + \dim \mathfrak{a}^\perp = \dim \mathfrak{g}$. Let us consider $Z \in \mathfrak{g}$ and X, $Y \in \mathfrak{a} \cap \mathfrak{a}^\perp$. We have $B(Z,[X,Y]) = B([Z,X],Y) = 0$. Then $[X,Y] = 0$ because $B(Z,[X,Y]) = 0$ for any Z and B is nondegenerate. Thus $\mathfrak{a} \cap \mathfrak{a}^\perp$ is abelian. It is also an ideal because $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are.

Now we consider $\mathfrak{b}$, a complementary of $\mathfrak{a} \cap \mathfrak{a}^\perp$ in $\mathfrak{g}$, $Z \in \mathfrak{g}$ and $T \in \mathfrak{a} \cap \mathfrak{a}^\perp$. The endomorphism $E = \mathrm{ad}\, T \circ \mathrm{ad}\, Z$ sends $\mathfrak{a} \cap \mathfrak{a}^\perp$ to $\{0\}$. Indeed consider $A \in \mathfrak{a} \cap \mathfrak{a}^\perp$; $(\mathrm{ad}\, Z)A \in \mathfrak{a} \cap \mathfrak{a}^\perp$ because it is an ideal, and then $(\mathrm{ad}\, T \circ \mathrm{ad}\, Z)A = 0$ because it is abelian.

The endomorphism E also sends $\mathfrak{b}$ to $\mathfrak{a} \cap \mathfrak{a}^\perp$ (it may not be surjective); then $\mathrm{Tr}(\mathrm{ad}\, T \circ \mathrm{ad}\, Z) = 0$ and $\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$. Since B is nondegenerate, $\dim \mathfrak{a} + \dim \mathfrak{a}^\perp = \dim \mathfrak{g}$. Then $\mathfrak{a} \oplus \mathfrak{a}^\perp = \mathfrak{g}$ is well a direct sum.

Third item. From lemma 51.20, the Killing form of $\mathfrak{g}$ descent to the ideal $\mathfrak{a}$; then it is also nondegenerate and $\mathfrak{a}$ is also semisimple. □

Corollary 51.51.
A semisimple Lie algebra has center $\{0\}$.

cor:ss_no_centre

Proof. If $Z \in \ker \mathfrak{g}$, $\mathrm{ad}\, Z = 0$. So $B(Z,X) = 0$ for any $X \in \mathfrak{g}$. Since B is nondegenerate, it implies $Z = 0$. □

Corollary 51.52.
If $\mathfrak{g}$ is a semisimple Lie algebra, it can be written as a direct sum

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$$

where the $\mathfrak{g}_i$ are simples ideals in $\mathfrak{g}$. Moreover each simple ideal in $\mathfrak{g}$ is a direct sum of some of them.

cor:decomp_ideal

6. In this case, B is nondegenerate.

Proof. If $\mathfrak{g}$ is simple, the statement is trivial. If it is not, we consider $\mathfrak{a}$, an ideal in $\mathfrak{g}$. Proposition 51.50 makes $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. Since $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are semisimple, we can once again brake them in the same way. We do it until we are left with simple algebras.

For the second part, consider $\mathfrak{b}$ a simple ideal in $\mathfrak{g}$ which is not a sum of $\mathfrak{g}_i$. Then $[\mathfrak{g}_i, \mathfrak{b}] \subset \mathfrak{g}_i \cap \mathfrak{b} = \{0\}$. Then $\mathfrak{b}$ is in the center of $\mathfrak{g}$. This contradict corollary 51.51. □

Proposition 51.53.
If $\mathfrak{g}$ is semisimple then

$$\mathrm{ad}(\mathfrak{g}) = \partial(\mathfrak{g}),$$

i.e. any derivation is an inner automorphism: ⌈ prop:ss_derr_int

Proof. We saw at page 3208 that $\mathrm{ad}(\mathfrak{g}) \subset \partial(\mathfrak{g})$ holds without assumptions of (semi)simplicity. Now we consider D, a derivation: $\forall X \in \mathfrak{g}$,

$$\mathrm{ad}(DX) = [D, \mathrm{ad}\, X].$$

Then $\mathrm{ad}(\mathfrak{g})$ is an ideal in $\partial(\mathfrak{g})$ because the commutator of $\mathrm{ad}\, X$ with any element of $\partial(\mathfrak{g})$ still belongs to $\mathrm{ad}(\mathfrak{g})$. Let us denote by $\mathfrak{a}$ the orthogonal complement of $\mathrm{ad}(\mathfrak{g})$ in $\partial(\mathfrak{g})$ (for the Killing metric). The algebra $\mathrm{ad}(\mathfrak{g})$ is semisimple because of it isomorphic to $\mathfrak{g}$. Since the Killing form on $\mathrm{ad}(\mathfrak{g})$ is nondegenerate, $\mathfrak{a} \cap \mathrm{ad}(\mathfrak{g}) = \{0\}$. Finally $D \in \mathfrak{a}$ implies $[D, \mathrm{ad}\, X] \in \mathfrak{a} \cap \mathrm{ad}(\mathfrak{g}) = \{0\}$. Then $\mathrm{ad}(DX) = 0$ for any $X \in \mathfrak{g}$, so that $D = 0$. This shows that $\mathfrak{a} = \{0\}$, so that $\mathrm{ad}(\mathfrak{g}) = \partial(\mathfrak{g})$. □

51.7.1 Cartan criterion

Let us recall a result: $\mathscr{D}\mathfrak{g} = \mathfrak{g}^1$, $[\mathscr{D}\mathfrak{g}, \mathscr{D}\mathfrak{g}] \subset \mathfrak{g}^2$; then $\mathscr{D}^k\mathfrak{g} \subset \mathfrak{g}^k$. Thus if $\mathfrak{g}$ is nilpotent, it is solvable. On the other hand, by the Engel theorem 51.37, $\mathscr{D}\mathfrak{g}$ is nilpotent if and only if all the $\mathrm{ad}_{\mathscr{D}\mathfrak{g}} x$ are nilpotent for $x \in \mathscr{D}\mathfrak{g}$.

Theorem 51.54 (Cartan criterion)**.**
Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. We suppose that $\mathrm{Tr}(xy) = 0$ $\forall x \in \mathscr{D}\mathfrak{g}, y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.

Proof. It is sufficient to prove that $\mathscr{D}\mathfrak{g}$ is nilpotent indeed if we write $\mathscr{D}^k\mathfrak{g} \subset \mathfrak{g}^k$ with $\mathscr{D}\mathfrak{g}$ instead of $\mathfrak{g}$, $\mathscr{D}^{k+1}\mathfrak{g} \subset (\mathscr{D}\mathfrak{g})^k$. If $\mathscr{D}\mathfrak{g}$ is nilpotent, $(\mathscr{D}\mathfrak{g})^n = 0$ and $\mathscr{D}^{n+1}\mathfrak{g} = 0$ so that $\mathfrak{g}$ is solvable.

Let us consider $x \in \mathscr{D}\mathfrak{g}$. We have to prove that it is ad-nilpotent (see the Engel theorem 51.37). Let $A = \mathscr{D}\mathfrak{g}$, $B = \mathfrak{g}$ and $M = \{x \in \mathfrak{gl}(V)$ tel que $[x\mathfrak{g}] \subset \mathscr{D}\mathfrak{g}\}$. By definition of $\mathscr{D}\mathfrak{g}$, $\mathfrak{g} \subset M$. The lemma 51.12 will conclude that $x \in \mathscr{D}\mathfrak{g}$ is nilpotent if $\mathrm{Tr}(xy) = 0$ for any $y \in M$. Here we just have this equality for $y \in \mathfrak{g}$.

A typical generator of $\mathscr{D}\mathfrak{g}$ is $[x, y]$ with x, $y \in \mathfrak{g}$. Take a $z \in M$; by the formula $\mathrm{Tr}([x, y]z) = \mathrm{Tr}(x[y, z])$, the trace that we have to check is

$$\mathrm{Tr}([x, y]z) = \mathrm{Tr}(x[y, z]) = \mathrm{Tr}([y, z]x). \tag{51.40}$$

But with $z \in M$, $[y, z] \in \mathscr{D}\mathfrak{g}$, then $\mathrm{Tr}([x, y]z) = \mathrm{Tr}([y, z]x) = 0$. Thus we are in the situation of the lemma. □

cor:ad_g_sol

Corollary 51.55.
A Lie algebra $\mathfrak{g}$ for which $\mathrm{Tr}(\mathrm{ad}\, x \circ \mathrm{ad}\, y) = 0$ for all $x \in \mathscr{D}\mathfrak{g}$, $y \in \mathfrak{g}$ is solvable.

Proof. We consider $\mathfrak{h} = \mathrm{ad}\, \mathfrak{g}$; this is a subalgebra of $\mathfrak{gl}(V)$ such that $a \in \mathscr{D}\mathfrak{h}$ and $b \in \mathfrak{h}$ imply $\mathrm{Tr}(ab) = 0$. In order to see it, remark that $a \in \mathscr{D}\mathfrak{h}$ can be written as $a = [\mathrm{ad}\, x, \mathrm{ad}\, y] = \mathrm{ad}[x, y]$ for certain x, $y \in \mathfrak{g}$. Then $\mathrm{Tr}(ab) = \mathrm{Tr}(\mathrm{ad}[x, y]\, \mathrm{ad}\, z)$ with x, y, $z \in \mathfrak{g}$; this is zero from the hypothesis. Then $\mathfrak{h} = \mathrm{ad}\, \mathfrak{g}$ is solvable.

It is also known that $\ker(\mathrm{ad}) = \mathcal{Z}(\mathfrak{g})$ is also solvable. Now we consider $\mathfrak{m}$ a complementary of $\mathcal{Z}(\mathfrak{g})$ in $\mathfrak{g}$: $\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{m}$. The Lie algebra $\mathrm{ad}(\mathfrak{m})$ is solvable and the homomorphism ϕ: $\mathrm{ad}\, \mathfrak{m} \to \mathfrak{m}$ defined by $\phi(\mathrm{ad}\, x) = x$ is well defined. From the first item of the proposition 51.32, $\mathfrak{m}$ is solvable. With obvious notations, an element of $\mathscr{D}\mathfrak{m}$ can be written as $[m, m']$ (because $\mathcal{Z}(\mathfrak{g})$ don't contribute to $\mathscr{D}\mathfrak{g}$). Then $\mathscr{D}\mathfrak{g} = \mathscr{D}\mathfrak{m}$, so that $\mathfrak{g}$ is as much solvable than $\mathfrak{m}$. □

Lemma 51.56.
The radical of a Lie algebra is non zero if and only if it has at least non zero abelian ideal. lem:ss_ideal

Proof. The radical of $\mathfrak{g}$ is its unique maximal solvable ideal. An eventually non empty abelian ideal should be in the radical.

Let us now consider that the radical is non zero, and consider the derived series of $\operatorname{Rad}\mathfrak{g}$. Since $\operatorname{Rad}\mathfrak{g}$ is solvable, we can consider n, the minimal integer such that $\mathscr{D}^n \operatorname{Rad}\mathfrak{g} = 0$. Then $\mathscr{D}^{n-1}\operatorname{Rad}\mathfrak{g}$ is a non zero abelian ideal. □

ThoRadicalEquivSS

Theorem 51.57.
A Lie algebra is semisimple if and only if its radical is zero.

Proof. Direct sense. We suppose $\operatorname{Rad}\mathfrak{g} = 0$ and we consider S, the radical of the Killing form:

$$S = \{X \in \mathfrak{g} \text{ tel que } B(X,Y) = 0 \,\forall Y \in \mathfrak{g}\}.$$

By definition, for any $X \in S$ and $Y \in \mathfrak{g}$, $\operatorname{Tr}(\operatorname{ad}X \circ \operatorname{ad}Y) = 0$. The Cartan criterion makes $\operatorname{ad}S$ solvable and the corollary 51.55 makes S solvable.

Now, the *ad*-invariance of the Killing form turns S into an ideal, so that $S \subset \operatorname{Rad}(\mathfrak{g})$ because any solvable ideal is contained in $\operatorname{Rad}\mathfrak{g}$. From the assumptions, $\operatorname{Rad}S = 0$, then $S \subset \operatorname{Rad}\mathfrak{g} = 0$. This shows that the Killing form is nondegenerate.
Inverse sense. We suppose $S = 0$ and we will show that any abelian ideal of $\mathfrak{g}$ is in S. In this case, if A is a solvable ideal with $\mathscr{D}^n A = 0$, then $\mathscr{D}^{n-1}A$ is an abelian ideal, so that $\mathscr{D}^{n-1}A = 0$. By induction, $A = 0$.

Let I be an abelian ideal of $\mathfrak{g}$, $X \in I$ and $Y \in \mathfrak{g}$. Then $\operatorname{ad}X \circ \operatorname{ad}Y$ is nilpotent because for $Z \in \mathfrak{g}$,

$$(\operatorname{ad}X \operatorname{ad}Y \operatorname{ad}X \operatorname{ad}Y)Z = (\operatorname{ad}X \operatorname{ad}Y)\underbrace{([X,[Y,Z]])}_{=X_1\in I} = (\operatorname{ad}X)\underbrace{[Y,X_1]}_{=X_2\in I} = (\operatorname{ad}X)X_2 = 0. \quad (51.41)$$

Then $0 = \operatorname{Tr}(\operatorname{ad}X \operatorname{ad}Y) = B(X,Y)$ and $X \in S$, so that $I \subset S = 0$.

□

51.7.2 More about radical

If $\mathfrak{g}$ is a Lie algebra whose radical is $\mathfrak{r}$, we say that a subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ is a **Levi subalgebra** if $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}$.

Any Lie algebra posses a Levi subalgebra [7].

Lemma 51.58.
If $\mathfrak{a}$ is an ideal in a Lie algebra $\mathfrak{g}$, then

$$\operatorname{Rad}\mathfrak{a} = (\operatorname{Rad}\mathfrak{r}) \cap \mathfrak{a}.$$

lem:rad_ideal

Before to begin the proof, let us recall that lemma 51.31 gives us an isomorphism $\psi\colon (\mathfrak{a}+\mathfrak{b})/\mathfrak{a} \to \mathfrak{b}/(\mathfrak{a}\cap\mathfrak{b})$ when $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in $\mathfrak{g}$.

Proof of the lemma. If $\mathfrak{r}$ is the radical of $\mathfrak{g}$, then the radical of $\mathfrak{g}/\mathfrak{r}$ is zero, so that $\mathfrak{r}/\mathfrak{r}$ is semisimple. Let $\mathfrak{a}$ be an ideal in $\mathfrak{g}$, then $(\mathfrak{a}+\mathfrak{r})/\mathfrak{r}$ is an ideal in the semisimple Lie algebra $\mathfrak{g}\mathfrak{r}$, so that it is also semisimple. From the isomorphism, $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{r})$ is also semisimple and $\mathfrak{a}\cap\mathfrak{r}$ must contains the radical of $\mathfrak{a}$. Indeed if a solvable ideal of $\mathfrak{a}$ where not in $\mathfrak{a}\cap\mathfrak{r}$, then this should give rise to a non zero solvable ideal in $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{r})$ although the latter is semisimple. Then $\mathfrak{a}\cap\mathfrak{r} = \operatorname{Rad}\mathfrak{a}$. □

7. Reference needed.

Proposition 51.59.
If A is a compact group of automorphisms of the Lie algebra $\mathfrak{g}$, then there exists a Levi subalgebra of $\mathfrak{g}$ which is invariant under A.

Proof. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$; we will split our proof into two cases following $[\mathfrak{r},\mathfrak{r}]=0$ or not.
The radical is abelian. In this first case we consider an induction with respect to the dimension of $\mathfrak{g}$. We consider $\overline{\mathfrak{g}}=\mathfrak{g}/[\mathfrak{r},\mathfrak{r}]$ and $\overline{\mathfrak{r}}=\mathfrak{r}/[\mathfrak{r},\mathfrak{r}]$: these are algebras with one less dimension that $\mathfrak{g}$ and $\mathfrak{r}$. We denote by $\pi\colon\mathfrak{g}\to\overline{\mathfrak{g}}$ the natural projection.

We begin to prove that $\overline{\mathfrak{r}}$ is the radical of $\overline{\mathfrak{g}}$. It is clear from the Lie algebra structure on a quotient that $\overline{\mathfrak{r}}$ is an ideal because $\mathfrak{r}$ is. It is also clear that $\overline{\mathfrak{r}}$ is solvable. We just have to see that $\overline{\mathfrak{r}}$ is maximal in $\overline{\mathfrak{g}}$. For this, suppose that $\overline{\mathfrak{r}}\cup\overline{X}$ is a solvable ideal in $\overline{\mathfrak{g}}$. Then it is easy to see that $\mathfrak{r}\cup X$ is an ideal in $\mathfrak{g}$. Taking commutators in $\overline{\mathfrak{r}}\cup\overline{X}$, we always finish in $\overline{0}\in\overline{\mathfrak{g}}$, i.e. in $[\mathfrak{r},\mathfrak{r}]$. Taking again some commutators, we finish on $0\in\mathfrak{g}$ because $\mathfrak{r}$ is solvable. This contradict the maximality of $\mathfrak{r}$.

Since A is made up of automorphisms, it leaves $\mathfrak{r}$ invariant, so that it also acts on $\overline{\mathfrak{g}}$ as an automorphism group: $a\overline{X}=\overline{aX}$ for $a\in A$ and $X\in\mathfrak{g}$. From the induction assumption, we can find a Levi subalgebra $\overline{\mathfrak{s}}$ in $\overline{\mathfrak{g}}$: $\overline{\mathfrak{s}}\oplus\overline{\mathfrak{r}}=\overline{\mathfrak{g}}$. In this case, the radical of $\pi^{-1}(\overline{\mathfrak{s}})$ is $[\mathfrak{r},\mathfrak{r}]$. Indeed in the one hand, $\overline{\mathfrak{r}}\cap\overline{\mathfrak{s}}=0$, so that $\pi^{-1}(\overline{\mathfrak{r}}\cap\overline{\mathfrak{s}})=[\mathfrak{r},\mathfrak{r}]$. In the other hand $\pi^{-1}(\overline{\mathfrak{r}}\cap\overline{\mathfrak{s}})=\pi^{-1}(\overline{\mathfrak{r}})\cap\pi^{-1}(\overline{\mathfrak{s}})=\mathfrak{r}\cap\pi^{-1}(\overline{\mathfrak{s}})$. The lemma 51.58 conclude that $\operatorname{Rad}\pi^{-1}(\overline{\mathfrak{s}})=[\mathfrak{r},\mathfrak{r}]$.

Now A is a compact group of automorphism which leaves invariant $\pi^{-1}(\overline{\mathfrak{s}})$, so we have a Levi subalgebra $\mathfrak{s}$ of $\pi^{-1}(\overline{\mathfrak{s}})$ invariant under A. We will see that this is in fact a Levi subalgebra of the whole $\mathfrak{g}$, i.e. we have to prove that $\mathfrak{s}\oplus\mathfrak{r}=\mathfrak{g}$. From the definition of $\mathfrak{s}$,

$$\mathfrak{s}\oplus[\mathfrak{r},\mathfrak{r}]=\pi^{-1}(\overline{\mathfrak{s}}),$$

and by definition of $\overline{\mathfrak{s}}$,

$$\overline{\mathfrak{s}}\oplus\frac{\mathfrak{r}}{[\mathfrak{r},\mathfrak{r}]}=\overline{\mathfrak{g}}.$$

Then

$$\mathfrak{g}=\pi^{-1}(\overline{\mathfrak{s}})\oplus\mathfrak{r}+[\mathfrak{r},\mathfrak{r}]=\mathfrak{s}\oplus[\mathfrak{r},\mathfrak{r}]\oplus\mathfrak{r}+[\mathfrak{r},\mathfrak{r}]=\mathfrak{s}\oplus\mathfrak{r}. \tag{51.42}$$

We can now pass to the second case: $[\mathfrak{r},\mathfrak{r}]=0$.
The radical is not abelian. Let $\mathfrak{s}_0$ and $\mathfrak{s}$ be Levi subalgebras of $\mathfrak{g}$. For $X\in\mathfrak{s}_0$, we write

$$X=f(X)+X_{\mathfrak{s}}$$

with respect to the decomposition $\mathfrak{g}=\mathfrak{s}\oplus\mathfrak{r}$. This defines a linear map $f\colon\mathfrak{s}_0\to\mathfrak{r}$. For any X, $Y\in\mathfrak{s}_0$, $[X_{\mathfrak{s}},X_{\mathfrak{s}}]=[X,Y]-[X,f(y)]-[f(X),Y]$ because $\mathfrak{r}$ is abelian. Since [8], $[X_{\mathfrak{s}},X_{\mathfrak{s}}]=[X,Y]_{\mathfrak{s}}$,

$$f([X,Y])=[X,f(Y)]-[f(X),Y]. \tag{51.43}$$ eq:f_presque_isom

Now let us consider a map $f\colon\mathfrak{s}_0\to\mathfrak{r}$ which satisfy this equation. Then the map $X\to X-f(X)$ is an isomorphism between $\mathfrak{s}_0$ and his image which is a Levi subalgebra of $\mathfrak{g}$. Indeed

$$\begin{aligned}[X,Y]\to[X,Y]-f([X,Y])&\\ =[X,Y]-[X,f(Y)]-[f(X),Y]&\\ =[X-f(X),Y-f(Y)].&\end{aligned} \tag{51.44}$$

Now we consider V, the space of all the linear maps $\mathfrak{s}_0\to\mathfrak{r}$ which fulfil the condition (51.43). We have a bijection between V and the Levi subalgebras of $\mathfrak{g}$: for any Levi subalgebra we associate the map $f\in V$ given by $X=f(X)+X_{\mathfrak{s}}$.

So our proof can be reduced to find a fixed point of V under the action of A. In order to do that, we will see that A is a group of *affine* transformations on V. Consider a $\alpha\in A$ and f_0, f_0^α,

8. C'est pas clair pourquoi on a $\mathfrak{a}$.

f^α be the elements of V corresponding to $\mathfrak{s}_0$, $\mathfrak{s}$ and $\alpha(\mathfrak{s})$. We take a $X \in \mathfrak{s}_0$ and we denote by $\overline{\alpha}(X)$ the $\mathfrak{s}_0$-component of $\alpha(X)$ with respect to the decomposition $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}_0$:

$$\alpha(X) = \overline{\alpha}(X) + \beta(X).$$

This also defines $\beta\colon \mathfrak{g} \to \mathfrak{r}$ and $-\beta(X)$ is the $\mathfrak{r}$-component of $\overline{\alpha}(X)$ with respect to $\mathfrak{g} = \mathfrak{r} \oplus \alpha(\mathfrak{s}_0)$. Since f_0^α just correspond to this decomposition, $f_0^\alpha(\overline{\alpha}(X)) = -\beta(X)$, so that

$$\begin{aligned} \overline{\alpha}(X) &= f_0^\alpha(\overline{\alpha}(X)) + \alpha(X) \\ &= f_0^\alpha(\overline{\alpha}(X)) + \alpha(f(X)) - \alpha(f(X)) + \alpha(X). \end{aligned} \tag{51.45}$$

Since $X - f(X) \in \mathfrak{s}$, $\alpha(X) - \alpha(f(X)) \in \alpha(\mathfrak{s})$, then $f_0^\alpha(\overline{\alpha}(X)) + \alpha(X)$ is the $\mathfrak{r}$-component of $\overline{\alpha}(X)$ with respect to $\mathfrak{g} = \mathfrak{r} \oplus \alpha(\mathfrak{s})$. Then

$$f_0^\alpha(\overline{\alpha}(X)) + \alpha(f(X)) = f^\alpha(\overline{\alpha}(X)) = f^\alpha(\overline{\alpha}(X)).$$

Since X was taken arbitrary, $f^\alpha = f_0^\alpha + \alpha \circ f \circ \overline{\alpha}^{-1}$. Then the map $V \to V$, $f \to f^\alpha$ is an affine transformation with translation equals to f_0^α and linear part being $f \to \alpha \circ f \circ \overline{\alpha}$.

A general result shows that a compact group of affine transformations on a vector space has a fixed point.

□

An other result that will be used:

lem:Killing_ss_descent

Lemma 51.60.
If G is a semisimple Lie group and H a semisimple subgroup of G, the restrictions on H of the Killing form of G is nondegenerate.

51.7.3 Lorentz algebra

LemCommsopqAlg

Lemma 51.61 ([851])**.**
The matrices of $\mathfrak{so}(p,q)$ satisfy the definition relation

$$M^t\eta + \eta M = 0, \tag{51.46}$$

and if M^{ab} is the "rotation" in the place of directions a and b (i.e. a trigonometric or an hyperbolic rotation following that a and b are of the same type or not), then the action on $\mathbb{R}^{(p,q)}$ is given by $(x')^\mu = (M^{ab})^\mu_\nu x^\nu$ with

$$(M^{ab})^\mu_\nu = \eta^{a\mu}\delta^b_\nu - \eta^{b\mu}\delta^a_\nu. \tag{51.47}$$

The commutation relations are given by

$$[M^{ab}, M^{cd}] = -\eta^{ac}M^{bd} + \eta^{ad}M^{bc} + \eta^{bc}M^{ad} - \eta^{bd}M^{ac}. \tag{51.48}$$

Notice that $M^{ab} = -M^{ba}$.

There is an other physical reason (which is in fact the same, but differently presented) justifying the study of the Clifford algebra. The quantum field theory need representation of the Lorentz algebra [9]

$$[J^{\mu\nu}, J^{\rho\sigma}] = i(\eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\rho} + \eta^{\mu\sigma}J^{\nu\rho}).$$

Dirac had a trick to find such J matrices from a representation of the Clifford algebra. If we have $n \times n$ matrices γ_μ such that

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1}_{n\times n},$$

9. When one think to real infinitesimal rotation matrices, the presence of i seems not natural, but one redefines $J \to iJ$ for formalism reasons.

a n-dimensional representation of the Lorenz algebra is obtained by

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu].$$

By a simple redefinition $J = iM$, one obtains

$$[J, J] = i\eta J \tag{51.49}$$

EqJJietaJcomm

instead of $[M, M] = \eta M$, and the matrices J are Hermitian. Here η is the matrix

$$\eta = diag(\underbrace{+, \dots, +}_{p\text{times}}, \underbrace{-, \dots, -}_{q \text{ times}}). \tag{51.50}$$

As convention, we say that a direction corresponding to a *positive* entry in the metric is a *time* direction, while the spatial directions are negative. That corresponds to the convention of page 3941 to say that a *time-like* vector has positive norm.

51.8 Clifford algebra

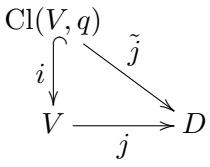
51.8.1 Definition and universal problem

Theorem 51.62.
Let E be an unital associative algebra and $j: V \to E$ a linear map such that

$$j(v) \cdot j(v) = q(v)1. \tag{51.51}$$

102r1

Then we have an unique extension of j to a homomorphism $\tilde{j}$: $\mathrm{Cl}(V, q) \to E$. Moreover, $\mathrm{Cl}(V, q)$ is the unique associative algebra which have this property for all such E.



tho_Cliffunif

This theorem can be seen as a definition of $\mathrm{Cl}(V, q)$.

Proof. The proof shall belongs two parts: the first one will show how to extend j and why it is unique, and the second one will prove the unicity of $\mathrm{Cl}(V, q)$.

We begin by define the extension of j. First note that any linear map $f: V \to E$ can be extended to an algebra homomorphism $\overline{f}: T(V) \to E$ in only one way. Indeed, the homomorphism condition require that $\overline{f}(v \otimes w) = f(v) \cdot f(w)$. The whole map $\overline{f}$ is then well defined by the data of f alone.

As far as the map j is concerned, we have the relation (51.51) which says that $\overline{j}(\mathcal{I}) = 0$. Indeed,

$$\overline{j}(v \otimes v - q(v) \cdot (1)) = \overline{j}(v) \cdot \overline{j}(v) - q(v)\overline{j}(1) = j(v) \cdot j(v) - q(v)1 = 0. \tag{51.52}$$

Thus $\overline{j}: T(V) \to E$ is a class map for $\mathcal{I}$, and we can descent $\overline{j}$ from $T(V)$ to $\mathrm{Cl}(V, q)$, We define $\tilde{j}: \mathrm{Cl}(V, q) \to E$ by

$$\tilde{j}[x] = \overline{j}(x) \tag{51.53}$$

where $[x]$ is the class of x. That's for the existence part.

The unicity is clear: $f_1 = f_2$ on V implies that $\overline{f_1} = \overline{f_2}$ on $T(V)$. Thus $\tilde{f}_1 = \tilde{f}_2$ on $\mathrm{Cl}(V, q)$.

We turn now our attention to the unicity of $\mathrm{Cl}(C, q)$. Let D be an unital associative algebra such that

(1) $V \subset D$,

(2) For any unital associative algebra E and for any $f: D \to E$ such that $f(v) \cdot f(v) = -q(v)1$, there exists only one homomorphic map $\tilde{f}: D \to E$ which extend f.

We should find a homomorphic map $\tilde{k}\colon D \to \mathrm{Cl}(V,q)$. Let i be the canonical injection $i\colon V \to D$. Clearly, we have a homomorphism $V \to i(V)$. Now, as a space E, we can take $\mathrm{Cl}(V,q)$; i can be seen as a linear map $i\colon V \to \mathrm{Cl}(V,q)$ such that $i(v)\cdot i(v) = q(v)1$. The assumptions say that i can be extended (in only one way) to a homomorphic map $\tilde{i}\colon D \to \mathrm{Cl}(V,q)$.

The Clifford algebra is thus unique up to a homomorphism.

□

What we proved is the following: if for any E and for any $j\colon V \to E$ such that $j(v)\cdot j(v) = q(v)1$, there exist an unique $\tilde{j}\colon D \to E$ which extend j, then $D = \mathrm{Cl}(V,q)$ up to a homomorphism. One ays that $\mathrm{Cl}(V,q)$ solve an **universal problem**.

51.8.2 Representations of the algebra $\mathfrak{su}(2) = \mathfrak{so}(3)$

subsecPJmtqrG

If one knows a load of theory, it is possible to determine the irreducible representations of $\mathfrak{so}(3)$ in a very short way; This will be done in example 51.174. We are now going to determine the irreducible representations of $\mathfrak{so}(3)$ in a quite explicit way.

51.8.2.1 Ladder operators

The algebra $\mathfrak{su}(2)$ is the real algebra generated by the matrices of the form $\begin{pmatrix} \alpha & \beta \\ -\beta^* & -\alpha \end{pmatrix}$ with $\alpha, \beta \in \mathbb{C}$. A convenient basis is given by

$$u_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad u_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad u_3 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{51.54}$$

EqGenssudeux

That algebra satisfies the commutation relations

$$[u_i, u_j] = \epsilon_{ijk} u_k. \tag{51.55}$$

The trick to build finite dimensional representations of that algebra is common (see [852] for example). The first step is to perform a change of basis $J_k = iu_k$ that brings the algebra under the form (see section 53.19 to understand why)

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{51.56}$$

EqAlgsuiepsijk

We are going to construct all the finite dimensional irreducible representations of the algebra (51.56). The key point of that new basis is that one can define the **ladder operators**

$$J_\pm = J_1 \pm iJ_2 \tag{51.57}$$

that have the property that

$$[J_3, J_\pm] = \pm J_\pm. \tag{51.58}$$

Notice that for every i, we have $(J_i)^* = J_i$, so that $(L^\pm)^* = L^\mp$. An other important property is that, defining $J^2 = J_1^2 + J_2^2 + J_3^2$, we have

$$[J_i, J^2] = 0, \tag{51.59}$$

which show that J^2 is a Casimir operator, and is thus by Schur's lemma a multiple of identity. Notice that we are using an abuse of notation between J_i as element of $\mathfrak{su}(2)$ and J_i as the operator that represent J_i. In the first case, products like J_iJ_j make no sense[10], but it makes sense as operator composition.

10. In fact, one has to understand these products as elements of the universal enveloping algebra. What we are building is a reprensentation of that algebra, which, obviously, restricts to a representation of the algebra. When we use the Schur's lemma, in fact we invoke it in $\mathcal{U}\big(\mathfrak{so}(3)\big)$

The subalgebra $\{J^2, J_3\}$ being abelian, we can simultaneously diagonalise J^2 and J_3. Let $|m,\sigma\rangle$ be an orthonormal basis of the eigenspace of J_3 associated with the eigenvalue m. The index σ is for a possible degenerateness to be studied later. We have

$$J_3|m,\sigma\rangle = m|m,\sigma\rangle.$$

Using the commutation relations between J_3 and the ladder operators, we have

$$J_3 J_\pm|m,\sigma\rangle = \big(\pm J_\pm + J_\pm J_3\big)|m,\sigma\rangle = (m\pm 1)J_\pm|m,\sigma\rangle. \tag{51.60}$$

EqJtroisJpmmplusun

Thus $J_\pm|m,\sigma\rangle$ is an eigenvector of J_3 with the eigenvalue $m\pm 1$, which means that $J_\pm|m,\sigma\rangle$ is a linear combination of the vectors $|m\pm 1,\sigma\rangle$ with different values of σ. This is the reason of the name of the *ladder* operators: they raise and lower the eigenvalue of J_3.

We can now prove that one has to drop the index σ because eigenvalues of J_3 cannot be degenerated. For, compute

$$J_+J_- = (J_1+iJ_2)(J_1-iJ_2) = J^2 - J_3^2 + i[J_2,J_1] = J^2 - J_3^2 + J_3, \tag{51.61}$$

JpJmJcarrerelation

so that

$$J_+J_-|m,\sigma\rangle = (\alpha - m^2 + m)|m,\sigma\rangle$$

where α is defined by $J^2 = \alpha\mathbb{1}$. That proves that the space generated by $|m,\sigma\rangle$ and the action of J_3, J_+ and J_- is invariant under the representation, while one cannot obtain $|m,\sigma'\rangle$ by action of $J_\pm$ on $|m,\sigma\rangle$. Since we are looking for *irreducible* representations, that space must actually be all the representation space. That rules out the possibility to have two different vectors $|m,\sigma_1\rangle$ and $|m,\sigma_2\rangle$.

The explicit matrix form of $J_\pm$ are:

$$J_+ = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad J_- = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{51.62}$$

Since we are searching for finite dimensional representations, there exists a maximal eigenvalue of J_3. Let us denote by j that maximal eigenvalue and by $|j\rangle$ the corresponding eigenvector. The relation (51.60) shows that if $J_+|j\rangle \neq 0$, then $J_+|j\rangle$ is an eigenvector for J_3 with eigenvalue $j+1$, which contradicts maximality. Then we have $J_+|j\rangle = 0$.

Since we know the action of J_3 and J_+ on $|j\rangle$, it is convenient to write J^2 in terms of these two operators. This is done in the same way as probing equation (51.61):

$$J^2 = J_3^2 + J_3 + J_-J_+, \tag{51.63}$$

so that

$$J^2|j\rangle = j(j+1)|j\rangle. \tag{51.64}$$

EqJcarreiiplusun

We know that $J^2 = \alpha\mathbb{1}$ and that α is a characteristic of the representation. What equation (51.64) tells us is that the maximal eigenvalue of J_3 is related to α by $j(j+1) = \alpha$.

We are now able to determine the proportionality constant of relation $J_\pm|m\rangle \propto |m\pm 1\rangle$. Since $(J_-)^* = J_+$, we have

$$\|J_-|m\rangle\|^2 = \langle m|J_+J_-|m\rangle = j(j+1) - m^2 + m. \tag{51.65}$$

EqnormeJmoinun

Then one has

$$J_-|m\rangle = \sqrt{j(j+1) - m(m-1)}\,|m-1\rangle, \tag{51.66a}$$

$$J_+|m\rangle = \sqrt{j(j+1) - m(m+1)}\,|m+1\rangle. \tag{51.66b}$$

EqJmoinsmanglemmointun

As expected, $J_-|-j\rangle = 0$ and $J_+|j\rangle = 0$. Notice that we avoid the possibility $J_-|m\rangle = -\sqrt{\cdots}|m-1\rangle$ by a simple redefinition $|m-1\rangle \to -|m-1\rangle$.

Equation (51.65) shows that the norm of $|m\rangle$ becomes negative for $m < -j$ and $m > j+1$. We conclude that the minimal eigenvalue of J_3 is $-j$. Since $|j\rangle$ has to be reached from $|-j\rangle$ by action of J_+, the difference $j-(-j)$ must be an integer. Thus $j \in \mathbb{N}/2$. The number j is the **spin** of the representation.

Let us give the explicit example with spin one half. When $j = \frac{1}{2}$, the vector space is generated by the vectors $|1/2\rangle$ and $|-1/2\rangle$, and the operators are given by

$$J_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad J_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{51.67}$$

from which we deduce

$$J_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Notice that we have $J_i = \frac{1}{2}\sigma_i$ with the **Pauli matrices**,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{51.68}$$

These matrices fulfil the relation

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k. \tag{51.69}$$

51.8.2.2 Weight vectors

subSubSecweightsotrois

The algebra $\mathfrak{so}(3)$ does not contain abelian subalgebra of dimension bigger than one, so a Cartan subalgebra is generated by J_3. The unique (up to dilatation) element of $\mathscr{H}^*$ is thus given by $\alpha(J_3) = 1$. The relation $[J_z, J_\pm]$ provides the root spaces:

$$\begin{aligned} \mathfrak{so}(3)_1 &= \{J_+\} \\ \mathfrak{so}(3)_{-1} &= \{J_-\}, \end{aligned} \tag{51.70}$$

thus $\mathfrak{n}^\pm$ is generated by $J_\pm$.

51.9 Cartan subalgebras in complex Lie algebras

SecCartaninComplex

About Cartan algebra, one can read [853, 854, 855, 848].

In this section $\mathfrak{g}$ will always denotes a complex finite dimensional Lie algebra.

PgDefCentralisateur

Definition 51.63.
*When $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, the **centralizer** of $\mathfrak{h}$ is the set*

$$\mathcal{Z}(\mathfrak{h}) = \{x \in \mathfrak{g} \text{ tel que } [x, \mathfrak{h}] \subset \mathfrak{h}\}. \tag{51.71}$$

More generally if $\mathfrak{g}$ is a Lie algebra and if $\mathfrak{a}$, $\mathfrak{b}$ are two subset of $\mathfrak{g}$, the centraliser of $\mathfrak{a}$ in $\mathfrak{b}$ is

$$\mathcal{Z}_{\mathfrak{b}}(\mathfrak{a}) = \{X \in \mathfrak{b} \text{ tel que } [X, \mathfrak{a}] = 0\}. \tag{51.72}$$

*If $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$, its **normalizer** is*

$$\mathfrak{n}_{\mathfrak{a}} = \{X \in \mathfrak{g} \text{ tel que } [X, \mathfrak{a}] \subset \mathfrak{a}\}. \tag{51.73}$$

One can check that $\mathfrak{a}$ is an ideal in $\mathfrak{n}_{\mathfrak{a}}$.

DEFooUNWJooVsBZgJ

Definition 51.64.
*A subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a **Cartan subalgebra** if it is nilpotent and if it is its own centralizer: $[x, \mathfrak{h}] \subset \mathfrak{h}$ implies $x \in \mathfrak{h}$.*

Our first task is to show that every Lie algebra has a Cartan algebra.

Lemma 51.65 (Primary decomposition theorem)**.**
Let V be a complex vector space and $A\colon V \to V$ be linear map. Then we have the direct sum decomposition

$$V = \bigoplus_{\lambda \in \mathbb{C}} V_\lambda(A) \tag{51.74}$$

EqPrimDecomTho

where $V_\lambda(A) = \{v \text{ tel que } (A - \lambda\mathbb{1})^n v = 0 \text{ for some } n \in \mathbb{N}\}$

This is the result that restricts ourself to *complex* Lie algebras when proving that Cartan subalgebras exist. Notice that the sum in (51.74) is reduced to the eigenvalues of A since $\mathfrak{g}_\lambda(A) = 0$ when λ is not an eigenvalue. Indeed if $\big(A - \lambda\mathbb{1}\big)^n Y = 0$ then $(A - \lambda\mathbb{1})^{n-1} Y$ is an eigenvector for A with eigenvalue λ.

For any $\lambda \in \mathbb{C}$ and $X \in \mathfrak{g}$ we consider the space

$$\mathfrak{g}_\lambda(X) = \{Y \in \mathfrak{g} \text{ tel que } \big(\operatorname{ad}(X) - \lambda\mathbb{1}\big)^n Y = 0 \text{ for some } n\}. \tag{51.75}$$

The primary decomposition theorem implies the decomposition

$$\mathfrak{g} = \bigoplus_{\lambda} \mathfrak{g}_\lambda(X) \tag{51.76}$$

EqDecomplGpRimDecombjjk

for each $X \in \mathfrak{g}$.

A small useful formula: if u is a derivation of the Lie bracket and if $[X, Y]$ is any bracket, then

$$(u - \lambda\mathbb{1})[X, Y] = \big[(u - \lambda\mathbb{1})X, Y\big] + [X, uY]. \tag{51.77}$$

EqWGujmeF

LemVZzSnUW

Lemma 51.66.
For each $X \in \mathfrak{g}$ and $\lambda, \mu \in \mathbb{C}$ we have

$$\big[\mathfrak{g}_\lambda(X), \mathfrak{g}_\mu(X)\big] \subset \mathfrak{g}_{\lambda+\mu}(X). \tag{51.78}$$

Proof. Let $X_\lambda \in \mathfrak{g}_\lambda(X)$ and $X_\mu \in \mathfrak{g}_\mu(X)$. Using the fact that $\operatorname{ad}(X)$ is a derivation we have

$$\operatorname{ad}(X)[X_\lambda, X_\mu] - (\lambda + \mu)[X_\lambda, X_\mu] = \Big[\big(\operatorname{ad}(X) - \mu\mathbb{1}\big)X_\lambda, X_\mu\Big] + \Big[X_\lambda, \big(\operatorname{ad}(X) + \mu\mathbb{1}\big)X_\mu\Big] \tag{51.79}$$

Let us show by induction the following equality for all n:

$$\big(\operatorname{ad}(X) - (\lambda + \mu)\mathbb{1}\big)^n [X_\lambda, X_\mu] = \sum_{i=0}^{\infty} \binom{n}{i} \Big[\big(\operatorname{ad}(X) - \lambda\mathbb{1}\big)^i X_\lambda, \big(\operatorname{ad}(X) - \mu\mathbb{1}\big)^{n-i} X_\mu\Big]. \tag{51.80}$$

EqPIzsRhb

In order to prove that, it is sufficient to apply $\big(\operatorname{ad}(X) - (\lambda + \mu)\mathbb{1}\big)$ to that equality and use the fact that $\operatorname{ad}(X)$ is a derivation of the Lie bracket. Then apply formula (51.77).

The expression (51.80) vanishes when n is large enough. □

We say that X is **regular** if $\dim \mathfrak{g}_0(X)$ is the smallest with respect to the others $\dim \mathfrak{g}_0(Y)$.

The following proposition shows that every complex Lie algebra has a Cartan Lie subalgebra.

Proposition 51.67.
If X is regular in $\mathfrak{g}$ then the subalgebra $\mathfrak{g}_0(X)$ is Cartan.

Proof. Since $X \in \mathfrak{g}_0(X)$ we have $\mathrm{ad}(X)\mathfrak{g}_\lambda(X) \subset \mathfrak{g}_\lambda(X)$. Thus we see $\mathrm{ad}(X)$ as a linear operator on $\mathfrak{g}_\lambda(X)$. The operator $\mathrm{ad}(X)|_{\mathfrak{g}_\lambda(X)}$ is nonsingular [11] when $\lambda \neq 0$. Indeed all the eigenvalues of $\mathrm{ad}(X)$ on $\mathfrak{g}_\lambda(X)$ are equal to λ because

$$\big(\mathrm{ad}(X) - \mu\mathbb{1}\big)Y = 0 \tag{51.81}$$

implies $Y \in \mathfrak{g}_\mu(X)$. If $Y \in \mathfrak{g}_\lambda(X)$ it only occurs when $\mu = \lambda$ since the sum (51.74) is direct.

For each eigenvalue λ we have a neighborhood $\mathcal{U}_\lambda$ of X in $\mathfrak{g}_0(X)$ such that for all $Y \in \mathcal{U}_\lambda$, $\mathrm{ad}(Y)$ is nonsingular on $\mathfrak{g}_\lambda(X)$. We consider $\mathcal{U} = \bigcap_\lambda \mathcal{U}_\lambda$ which is a non empty open set since the intersection is taken over the eigenvalues of $\mathrm{ad}(X)$ that are in finite numbers.

Let us prove that the restriction to $\mathfrak{g}_0(X)$ of the linear operator $\mathrm{ad}(Y)$ is nilpotent for each $Y \in \mathcal{U}$. First we have

$$\mathfrak{g}_0(Y) \subseteq \mathfrak{g}_0(X) \tag{51.82}$$

EqLgzsubsetlgX

because by construction $\mathrm{ad}(Y)$ cannot be nilpotent on the other spaces $\mathfrak{g}_\lambda(X)$. But by hypothesis the element X is regular, thus the inclusion (51.82) cannot be strict. Thus $\mathfrak{g}_0(X) \subset \mathfrak{g}_0(Y)$ which means that $\mathrm{ad}(Y)$ is nilpotent on $\mathfrak{g}_0(X)$.

Now the fact for $\mathrm{ad}(Y)$ to be nilpotent means the vanishing of a polynomial determined by the coefficients of the matrix of $\mathrm{ad}(Y)$. Since this polynomial vanishes on the open set $\mathcal{U}$, it vanishes identically, so that $\mathrm{ad}(Y)$ is nilpotent on $\mathfrak{g}_0(X)$. It results that $\mathfrak{g}_0(X)$ is a ad-nilpotent algebra and the Engel's theorem 51.37 concludes that $\mathfrak{g}_0(X)$ is nilpotent.

We still have to prove that $\mathfrak{g}_0(X)$ is its own centralizer. Since $\mathfrak{g}_0(X)$ is a subalgebra we have the inclusion

$$\mathfrak{g}_0(X) \subseteq \mathcal{Z}\big(\mathfrak{g}_0(X)\big). \tag{51.83}$$

Let $Z \in \mathcal{Z}\big(\mathfrak{g}_0(X)\big)$. For each $Y \in \mathfrak{g}_0(X)$ we have $[Z, Y] \in \mathfrak{g}_0(X)$. In particular with $Y = X$ we have $\mathrm{ad}(X)Z \in \mathfrak{g}_0(X)$. Thus

$$\mathrm{ad}(X)^n Z = \mathrm{ad}(X)^{n-1} \underbrace{\mathrm{ad}(X)Z}_{\in \mathfrak{g}_0(X)} \tag{51.84}$$

and there exists a n such that $\mathrm{ad}(X)^{n-1}\,\mathrm{ad}(X)Z = 0$. □

If $\mathfrak{g}$ is a Lie algebra, the group of **inner automorphism** is the subgroup of $\mathrm{Aut}(\mathfrak{g})$ generated by the elements of the form $e^{\mathrm{ad}(X)}$ with $X \in \mathfrak{g}$. This definition is motivated in the context of matrix groups by the fact that when $g = e^Y \in G$ and $X \in \mathfrak{g}$ we have

$$gXg^{-1} = e^{\mathrm{ad}(Y)}X. \tag{51.85}$$

Exemple 51.68.
If

$$g = \begin{pmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & a & b \\ -a & 0 & 0 \\ -b & 0 & 0 \end{pmatrix}, \tag{51.86}$$

then one checks that $g = e^Y$ with

$$Y = \begin{pmatrix} 0 & t & 0 \\ -t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{51.87}$$

and

$$gXg^{-1} = e^{\mathrm{ad}(Y)}X = \begin{pmatrix} 0 & a & b\cos(t) \\ -a & 0 & -b\sin(t) \\ -b\cos(t) & b\sin(t) & 0 \end{pmatrix}. \tag{51.88}$$

△

Theorem 51.69.
The group of inner automorphisms of $\mathfrak{g}$ acts transitively on the set of Cartan subalgebras.

11. it means that $\mathrm{ad}(Y)$ is invertible.

For a proof, see [856]. In particular they have all the same dimension and the definition of the **rank** as the dimension of its Cartan algebra make sense. In [856] we have a more abstract definition of the rank, see page III-2.

PropCartanLzXtjs

Proposition 51.70.
If $\mathfrak{h}$ is a Cartan subalgebra of the complex Lie algebra $\mathfrak{g}$, there exists a regular element $X \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}_0(X)$.

For a proof, see [856].

prop:Cartan_max_nil

Proposition 51.71.
A Cartan subalgebra is a maximal nilpotent subalgebra.

Proof. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ and $\mathfrak{n}$, a nilpotent algebra which contains $\mathfrak{h}$. Let $\{X_1, \dots, X_n\}$ be a basis of $\mathfrak{g}$ chosen in such a way that the p first vectors form a basis of $\mathfrak{h}$ while the r first, a basis of $\mathfrak{n}$ ($r > p$ of course). As notational convention, the subscript i, j are related to $\mathfrak{h}$ and u, t to $\mathfrak{n} \ominus \mathfrak{h}$.

Let us first suppose $\dim \mathfrak{n} = \dim \mathfrak{h} + 1$ and let X_u be the basis vector of $\mathfrak{n}$ which is not in $\mathfrak{h}$. Since $\mathfrak{h}$ is Cartan, we can find $X_i \in \mathfrak{h}$ such that $Y = [X_u, X_i] \notin \mathfrak{h}$. Then Y has a X_u-component and this contradict the fact that $\operatorname{ad} X_i$ is nilpotent.

The next case is $\mathfrak{n} = \mathfrak{h} \oplus X_u \oplus X_t$. In this case we can find a $X_i \in \mathfrak{h}$ such that $Y = [X_u, X_i] \notin \mathfrak{h}$. The fact to be nilpotent makes that Y has no X_u-component, so that it has a X_t-component. Now it is clear that for any $X_j \in \mathfrak{h}$, $[Y, X_j]$ still has no X_u-component (because $(\operatorname{ad} X_i \circ \operatorname{ad} X_j)$ has to be nilpotent), but has also no X_t-component. Then for any $X \in \mathfrak{h}$, $[Y, X] \in \mathfrak{h}$ with $Y \notin \mathfrak{h}$. There is a contradiction.

Now the step to the general case is easy: if $\dim \mathfrak{n} = \dim \mathfrak{h} + m$, we consider $X_1, \dots, X_m \in \mathfrak{h}$ and $A = (\operatorname{ad} X_1 \circ \operatorname{ad} X_m) X_u$. This is not in $\mathfrak{h}$ although $[A, X] \in \mathfrak{h}$ for any $X \in \mathfrak{h}$. □

Proposition 51.72.
If $\mathfrak{g}$ is a semisimple Lie algebra, a subalgebra $\mathfrak{h}$ is Cartan if and only if the two following conditions are satisfied:

(1) $\mathfrak{h}$ is a maximal abelian subalgebra

(2) the endomorphism $\operatorname{ad}(H)$ is diagonalizable for every $H \in \mathfrak{h}$.

51.10 Root spaces in semisimple complex Lie algebras

SecRootcomplexss

In this section we particularize ourself to complex semisimple Lie algebras. A very good reference about complex semisimple algebras including the reconstruction *via* the Cartan matrix and Chevalley-Weyl basis is [856].

51.10.1 Introduction and notations

Real and complex Lia algebras deserve quite different treatment with root space. We review here the main steps in both cases, emphasising the differences. We restrict ourself to semisimple Lie algebras. See [850].

51.10.1.1 Complex Lie algebras

If $\mathfrak{g}$ is a complex semisimple Lie algebra, we choose a Cartan subalgebra $\mathfrak{h}$ and the root spaces are given by

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \text{ tel que } [H, X] = \alpha(H) X \forall H \in \mathfrak{h}\}. \tag{51.89}$$

The dimension of $\mathfrak{h}$ is the rank of $\mathfrak{g}$. Then the root space decomposition reads

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \tag{51.90}$$

where Φ is the set of roots.

51.10.1.2 Real Lie algebras

If $\mathfrak{g}$ is a real semisimple Lie algebra we consider a Cartan involution and the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Then we choose a maximally abelian subalgebra $\mathfrak{a}$ in $\mathfrak{p}$ and we define

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \text{ tel que } [J, X] = \alpha(J)X \forall J \in \mathfrak{a}\}. \tag{51.91}$$

The rank of $\mathfrak{g}$ is the dimension of $\mathfrak{a}$. The root space decomposition then reads

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda \tag{51.92}$$

where Σ is the set of $\lambda \in \mathfrak{a}^*$ such that $\lambda \neq 0$ and $\mathfrak{g}_\lambda \neq 0$.

51.10.1.3 Notations

SubsecNotationRootsDel

We summarize the notations that will be used later. Let $\mathfrak{h}$ be a Cartan algebra in the complex semisimple Lie algebra $\mathfrak{g}$. An element $\alpha \in \mathfrak{h}^*$ is a root if the space

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \text{ tel que } \operatorname{ad}(H)X = \alpha(H)x, \forall H \in \mathfrak{h}\} \tag{51.93}$$

is non empty.

(1) Φ is the set of all the roots. We consider an ordering notion on Φ and $\Phi^+ = \Pi$ is the set of positive roots.
(2) An element in Φ^+ is simple if it cannot be written as the sum of two positive roots.
(3) Δ is the set of simple roots [12]. The simple roots are denoted by $\{\alpha_1, \dots, \alpha_l\}$.

51.10.2 Root spaces

We are considering a complex semisimple Lie algebra $\mathfrak{g}$ with a Cartan subalgebra $\mathfrak{h}$.

DefRootSpace

Definition 51.73.
For each $\alpha \in \mathfrak{h}^*$ we define

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \text{ tel que } \forall h \in \mathfrak{h}, \big(\operatorname{ad} h - \alpha(h)\big)^n x = 0 \text{ for some } n \in \mathbb{N}\}. \tag{51.94}$$

eq:lG_alpha_nss

*If $\mathfrak{g}_\alpha$ is not reduced to 0, we say that α is a **root** and $\mathfrak{g}_\alpha$ is a **root space**.*

Corollary 51.81 will provide an easier formula for the root spaces when the algebra $\mathfrak{g}$ is complex and semisimple.

TholGCartalphaplusbeta

Theorem 51.74.
Let $\mathfrak{g}$ be a complex Lie algebra with Cartan subalgebra $\mathfrak{h}$. If $\alpha, \beta \in \mathfrak{g}^*$ then

ItemTholGCartalphaplusbetai

(1) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$,
(2) $\mathfrak{g}_0 = \mathfrak{h}$.

prop:deux_racine

Proof. For $z \in \mathfrak{h}$ and x, $y \in \mathfrak{g}$ we have

$$\big(\operatorname{ad} z - (\alpha + \beta)(z)\big)[x, y] = [(\operatorname{ad} z - \alpha(z))x, y] + [x, (\operatorname{ad} z - \beta(z))y]. \tag{51.95}$$

Using the same induction as in the proof of lemma 51.66 we show that

$$\big(\operatorname{ad} z - (\alpha + \beta)(z)\big)^n [x, y] = \sum_k \binom{k}{n} \Big[(\operatorname{ad} z - \alpha(z))^k (x), (\operatorname{ad} z - \beta(z))^{n-k}(y)\Big]. \tag{51.96}$$

12. The symbol Δ has not a fixed signification in the literature. As example, in [857] the symbol Δ is the set of roots while in [858] it denotes the set of simple roots.

This formula shows that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. Indeed let $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_\beta$ and n be large enough,

$$(\operatorname{ad} z - (\alpha+\beta)(z))^n [x,y] = 0. \tag{51.97}$$

Now we turn our attention to the second part. Let us apply the Lie theorem 51.23 to the action of $\mathfrak{g}$ on the quotient $\mathfrak{g}_0/\mathfrak{h}$. There exists $[X_0] \in \mathfrak{g}_0/\mathfrak{h}$ such that $h[X_0] = \lambda(h)[X_0]$ where the bracket stand for the class. Since $\mathfrak{h}$ is nilpotent on $\mathfrak{g}_0$ we have $\lambda = 0$ identically. Looking outside the class, the existence of a non vanishing $[X_0] \in \mathfrak{g}/\mathfrak{h}$ such that $h[X_0] = 0$ means that there exists $X_0 \in \mathfrak{g}_0 \backslash \mathfrak{h}$ such that $[h, X_0] \in \mathfrak{h}$ for every $h \in \mathfrak{h}$. This contradicts the fact that $\mathfrak{h}$ is its own centralizer. □

Proposition 51.75.
The complex Lie algebra decomposes into the root spaces as

$$\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha. \tag{51.98}$$

Proof. Let $H \in \mathfrak{h}$. We consider the primary decomposition (51.76) with respect to the operator $\operatorname{ad}(H)$:

$$\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda(H). \tag{51.99}$$

If $H' \in \mathfrak{h}$ the operator $\operatorname{ad}(H')$ acts the space $\mathfrak{g}_\lambda(H)$ because $H' \in \mathfrak{g}_0(H)$ so that

$$[H', \mathfrak{g}_\lambda(H)] \subset \mathfrak{g}_\lambda(H). \tag{51.100}$$

Thus we can write the primary decomposition of $\mathfrak{g}_\lambda(H)$ with respect to the operator $\operatorname{ad}(H')$ knowing that

$$\big(\mathfrak{g}_\lambda(H)\big)_\mu(H') = \{X \in \mathfrak{g}_\lambda(H) \text{ tel que } \big(\operatorname{ad}(H') - \mu\big)^n X = 0\} = \mathfrak{g}_\lambda(H) \cap \mathfrak{g}_\mu(H'). \tag{51.101}$$

What we get is the decomposition

$$\mathfrak{g} = \bigoplus_\lambda \bigoplus_\mu \mathfrak{g}_\lambda(H) \cap \mathfrak{g}_\mu(H'). \tag{51.102}$$

We continue the decomposition with $H'', H''', \ldots$ until each $\operatorname{ad}(H)$ with $H \in \mathfrak{h}$ has only one eigenvalue on each of the summand of the decomposition

$$\mathfrak{g} = \bigoplus_{\lambda_1,\ldots,\lambda_l} \mathfrak{g}_{\lambda_1}(H_1) \cap \ldots \cap \mathfrak{g}_{\lambda_l}(H_l). \tag{51.103}$$

For each l-uple $(\lambda_1, \ldots, \lambda_l)$, the eigenvalue of H_i on $\mathfrak{g}_{\lambda_1} \cap \ldots \cap \mathfrak{g}_{\lambda_l}$ is λ_i. Thus we can see λ as a 1-form on $\mathfrak{h}$ and write

EqdirectsumGRoots
$$\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda \tag{51.104}$$

with

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \text{ tel que } \big(\operatorname{ad}(H) - \lambda(H)\big)^n X = 0\}. \tag{51.105}$$

□

cor:Bxy_zero
Corollary 51.76.
If $X_\alpha \in \mathfrak{g}_\alpha$ and $X_\beta \in \mathfrak{g}_\beta$ with $\alpha + \beta \neq 0$, then $B(X_\alpha, X_\beta) = 0$.

Proof. From the second point of proposition 51.74, we have $\operatorname{ad} X_\alpha \circ \operatorname{ad} X_\beta \colon \mathfrak{g}_\mu \to \mathfrak{g}_{\mu+\alpha+\beta}$. If $\alpha + \beta \neq 0$, the fact that the sum (51.104) is direct makes the trace of $\operatorname{ad} X_\alpha \circ \operatorname{ad} X_\beta$ zero. □

Since $\mathfrak{g}$ is semisimple, the restriction of the Killing form on $\mathfrak{h}$ is nondegenerate [13]. Thus we can introduce, for each linear function $\phi \colon \mathfrak{h} \to \mathbb{C}$, the unique element $t_\phi \in \mathfrak{h}$ such that

$$\phi(h) = B(t_\phi, h) \tag{51.106}$$

13. Because the Killing form is zero on each space $\mathfrak{g}_\alpha$ with $\alpha \neq 0$.

for every $h \in \mathfrak{h}$. This element is nothing else that the dual ϕ^* with respect to the Killing form. Indeed

$$t_\phi^*(h) = B(t_\phi, h) = \phi(h), \tag{51.107}$$

so that $t_\phi^* = \phi$. Incidentally, this proves that when ϕ runs over a basis of $\mathfrak{h}^*$, the vector t_ϕ runs over a basis of $\mathfrak{h}$. The space $\mathfrak{h}^*$ is endowed with an inner product defined by

EqDefInnprHestrar

$$(\alpha, \beta) = B(t_\alpha, t_\beta) = \beta(t_\alpha) = \alpha(t_\beta). \tag{51.108}$$

PropoLaplXYBXYtalpha

Lemma 51.77.
If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then

$$[X, Y] = B(X, Y)t_\alpha. \tag{51.109}$$

Proof. By theorem 51.74(1), $[X, Y] \in \mathfrak{g}_0 = \mathfrak{h}$. Now we consider $h \in \mathfrak{h}$ and the invariance formula (51.23). We find:

$$B\big(h, [X, Y]\big) = -B\big([X, h], Y\big) = \alpha(h)B(X, Y) = B(h, t_\alpha)B(X, Y) = B\big(h, B(X, Y)t_\alpha\big). \tag{51.110}$$

The lemma is proven since it is true for any $h \in \mathfrak{h}$ and B is nondegenerate on $\mathfrak{h}$. □

The elements t_α allow to introduce an inner product on $\mathfrak{h}^*$ and hence on the roots by defining

$$(\alpha, \beta) = B(t_\alpha, t_\beta). \tag{51.111}$$

Leminnerabequaaggb

Lemma 51.78.
If α and β are roots we have the formula

$$(\alpha, \beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)(\alpha, \gamma)(\beta, \gamma). \tag{51.112}$$

Proof. We consider for $\mathfrak{g}$ a basis in which all the elements are part of one of the root spaces and we look at the endomorphism $\mathrm{ad}(t_\alpha)$ of $\mathfrak{g}$. This is diagonal and has zeros on the entries corresponding to $\mathfrak{h}$. The other entries on the diagonal are of the form $\gamma(t_\alpha)$. Thus

$$B(t_\alpha, t_\beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)\gamma(t_\alpha)\gamma(t_\beta). \tag{51.113}$$

Thus we have $(\alpha, \beta) = B(t_\alpha, t_\beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)(\alpha, \gamma)(\beta, \gamma)$. □

PropScalrooTsQ

Proposition 51.79 ([857])**.**
Let α and β be roots. We have

(1) $(\alpha, \beta) \in \mathbb{Q}$,

(2) $(\alpha, \alpha) \geqslant 0$.

Proof. Let $\alpha, \beta \in \Phi$ and consider the space

$$V = \bigoplus_{m \in \mathbb{Z}} \mathfrak{g}_{\beta + m\alpha}. \tag{51.114}$$

If $X_\alpha \in \mathfrak{g}_\alpha$ and $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ with $[X_\alpha, X_{-\alpha}] = t_\alpha$ we have, for all $v \in V$,

$$[X_\alpha, v] \in V \tag{51.115a}$$
$$[X_{-\alpha}, v] \in V \tag{51.115b}$$
$$[t_\alpha, v] \in V. \tag{51.115c}$$

Thus we can consider the restrictions to V of the operators $\mathrm{ad}(X_\alpha)$, $\mathrm{ad}(X_{-\alpha})$ and $\mathrm{ad}(t_\alpha)$. Since ad is an homomorphism we have, as operator on V,

$$\mathrm{ad}(t_\alpha) = \big[\mathrm{ad}(X_\alpha), \mathrm{ad}(X_{-\alpha})\big], \tag{51.116}$$

and then $\mathrm{Tr}\left(\mathrm{ad}(t_\alpha)|_V\right) = 0$.

Let us compute that trace on the basis $\{v_k^{(i)}\}$ where $v_k^{(i)} \in \mathfrak{g}_{\beta+k\alpha}$. Since

$$\mathrm{ad}(t_\alpha)v_k^{(i)} = (\beta + k\alpha)(t_\alpha)v_k^{(i)} \tag{51.117}$$

we have

$$\begin{aligned} 0 &= \mathrm{Tr}\left(\mathrm{ad}(t_\alpha)|_V\right) && \text{(51.118a)}\\ &= \sum_{k\in\mathbb{Z}} \dim \mathfrak{g}_{\beta+k\alpha}(\beta + k\alpha)(t_\alpha) && \text{(51.118b)}\\ &= \sum_{k\in\mathbb{Z}} \dim_{\beta+k\alpha}\big((\alpha,\beta) + (\alpha,\alpha)\big) && \text{(51.118c)} \end{aligned}$$

and

$$\underbrace{\left(\sum_{k\in\mathbb{Z}} \dim \mathfrak{g}_{\beta+k\alpha}\right)}_{A\in\mathbb{N}}(\alpha,\beta) = -(\alpha,\alpha)\underbrace{\left(\sum_{k\in\mathbb{Z}} k\dim \mathfrak{g}_{\beta+k\alpha}\right)}_{B\in\mathbb{Z}}. \tag{51.119}$$

EqunderbAabmaaB

If $(\alpha,\alpha) = 0$ then we have $(\beta,\alpha) = 0$ for every $\beta \in \Phi$, hence $B(t_\alpha, t_\beta) = 0$ which contradicts non degeneracy of the Killing form. We conclude that $(\alpha,\alpha) \neq 0$. By the formula of lemma 51.78 we get

$$(\alpha,\alpha) = \sum_{\beta\in\Phi} \dim \mathfrak{g}_\beta(\alpha,\beta)^2. \tag{51.120}$$

Replacing in that formula the value of (α,β) taken from formula (51.119) we found

$$(\alpha,\alpha) = \sum_{\beta\in\Phi} \dim \mathfrak{g}_\beta \frac{B^2}{A^2}(\alpha,\alpha)^2 \tag{51.121}$$

and then $(\alpha,\alpha) \in \mathbb{Q}^+$. The fact that (α,β) is rational follows.

Notice that the sign of B is not guaranteed because it's not sure because we do not know whether there are more positive or negative terms in the sum of the right hand side of (51.119). □

Proposition 51.80.
Let α be a root of the complex semisimple Lie algebra $\mathfrak{g}$. Then

(1) $\dim \mathfrak{g}_\alpha = 1$,

(2) the only integer multiple of α to be roots are $\pm\alpha$.

Proof. Let $X_\alpha \in \mathfrak{g}_\alpha$ and consider the vector space

$$V = \mathbb{C}t_\alpha \oplus \mathbb{C}X_\alpha \oplus \bigoplus_{m<0} \mathfrak{g}_{m\alpha}. \tag{51.122}$$

Let $y \in \mathfrak{g}_{-\alpha}$ be chosen in such a way that $[X_\alpha, y] = t_\alpha$; by lemma 51.77 this is only a matter of normalization. The space V is invariant under $\mathrm{ad}(X_\alpha)$ and $\mathrm{ad}(y)$. Indeed

(1) $\mathrm{ad}(X_\alpha)t_\alpha = -\alpha(t_\alpha)X_\alpha \in \mathbb{C}X_\alpha$;

(2) $\mathrm{ad}(X_\alpha)X_\alpha = 0$;

(3) $\mathrm{ad}(X_\alpha)\mathfrak{g}_{m\alpha} \subset \mathfrak{g}_{(m+1)\alpha}$; if $m < -1$, $(m+1) < 0$, while if $m = -1$ we know that the commutator $[X_\alpha, \mathfrak{g}_{-\alpha}]$ is included in $\mathbb{C}t_\alpha \in V$;

(4) $\mathrm{ad}(y)t_\alpha \in \mathfrak{g}_{-\alpha}$

(5) $\mathrm{ad}(y)X_\alpha = -t_\alpha$ by definition;

(6) $\mathrm{ad}(y)\mathfrak{g}_{m\alpha} \subset \mathfrak{g}_{(m-1)\alpha}$.

Since $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{GL}(\mathfrak{g})$ is an homomorphism (lemma 51.4) we have

$$\big[\mathrm{ad}(X_\alpha), \mathrm{ad}(y)\big] = \mathrm{ad}(t_\alpha) \tag{51.123}$$

and then $\mathrm{Tr}\big(\mathrm{ad}(t_\alpha)\big) = 0$ because the trace of a commutator is zero[14]. Since V is an invariant subspace, the trace of $\mathrm{ad}(t_\alpha)$ restricted to V is also vanishing. Let us compute that trace on the basis $\{X_\alpha, t_\alpha, X^i_{m\alpha}\}_{m<0}$ where i takes as many values as the dimension of $\mathfrak{g}_{m\alpha}$.

We have $\mathrm{ad}(t_\alpha)X_{-\alpha} = -\alpha(t_\alpha)X_{-\alpha}$, $\mathrm{ad}(t_\alpha)t_\alpha = 0$ and $\mathrm{ad}(t_\alpha)X^i_{m\alpha} = m\alpha(t_\alpha)X_{m\alpha}$, thus the trace is

$$0 = \alpha(t_\alpha)\Big(-1 + \sum_{m=1}^{\infty} m \dim \mathfrak{g}_{m_\alpha}\Big). \tag{51.124}$$

Eqzequalmulsumnotinfty

Notice that the sum is in fact finite since the dimension of $\mathfrak{g}$ is finite. We know that $\alpha(t_\alpha) = B(t_\alpha, t_\alpha) \neq 0$, so that equation (51.124) is only possible with $\dim \mathfrak{g}_\alpha = 1$ and $\dim \mathfrak{g}_{m\alpha} = 0$ for $m \neq 1$. □

A very similar proof can be found in [857], page 827.

CorCoolWrlGbalpha

Corollary 51.81.
In the case of semisimple complex Lie algebra,

(1) the root spaces are given by

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \text{ tel que } \forall h \in \mathfrak{h}, [h, X] = \alpha(h)X\}; \tag{51.125}$$

EqExpWeightSemiSimple

(2) for every $x_\alpha \in \mathfrak{g}_\alpha$, and for every $h \in \mathfrak{h}$, we have

$$[h, x_\alpha] = \alpha(h)x_\alpha. \tag{51.126}$$

Proof. Let $X \in \mathfrak{g}_\alpha$, we have

$$\big(\mathrm{ad}(h) - \alpha(h)\big)^n X = 0, \tag{51.127}$$

so

$$\big(\mathrm{ad}(h) - \alpha(h)\big)\underbrace{\big(\mathrm{ad}(h) - \alpha(h)\big)^{n-1}X}_{v} = 0. \tag{51.128}$$

Eqadhalphahnnnuux

In particular the vector $v = \big(\mathrm{ad}(h) - \alpha(h)\big)^{n-1}X$ belongs to $\mathfrak{g}_\alpha$. Since the latter space has dimension one, the vector v is a multiple of X and consequently equation (51.128) shows that

$$\big(\mathrm{ad}(h) - \alpha(h)\big)v = \big(\mathrm{ad}(h) - \alpha(h)\big)X = 0. \tag{51.129}$$

The second point is only an other way to write the same. □

LemHzesialphaHz

Lemma 51.82.
If H is an element of $\mathfrak{h}$ with $\alpha(H) = 0$ for every root, then $H = 0$

Proof. Consider the decompositions (not unique) $H = \sum_{\alpha\in\Phi} a_\alpha t_\alpha$ and $H' = \sum_{\beta\in\phi} a'_\beta t_\beta$. Then

$$\begin{aligned} B(H, H') &= \sum_{\alpha,\beta} a_\alpha a'_\beta B(t_\alpha, t_\beta) && (51.130a)\\ &= \sum_{\alpha,\beta} a'_\beta \beta(a_\alpha, t_\alpha) && (51.130b)\\ &= \sum_{\beta} a'_\beta \beta(H) && (51.130c)\\ &= 0. && (51.130d) \end{aligned}$$

Such an element is thus Killing-orthogonal to the whole space $\mathfrak{h}$ but we already know the $\mathfrak{h}$ is orthogonal to each space $\mathfrak{g}_\alpha$ ($\alpha \neq 0$). By non degeneracy of the Killing form we must have $H = 0$. □

14. From the cyclic invariance of the trace.

Proposition 51.83.
The set of roots of a complex semisimple Lie algebra spans the dual space $\mathfrak{h}^*$.

Proof. Consider a basis $\{H_i\}$ of $\mathfrak{h}$ with $\{H_0, \dots, H_m\} = \mathrm{Span}(\Phi)$ and $\{H_{m+1}, \dots, H_r\}$ be outside of $\mathrm{Span}\,\Phi$. A root reads $\alpha = \sum_{k=0}^{m} a_k H_k^*$. Thus $\alpha(H_{m+1}) = 0$, which implies that $H_{m+1} = 0$ by lemma 51.82. □

Corollary 51.84.
A Cartan algebra $\mathfrak{h}$ of a complex semisimple Lie algebra is abelian.

Proof. Let $H', H'' \in \mathfrak{h}$ and consider $H = [H', H'']$, a root α and $X_\alpha \in \mathfrak{g}_\alpha$. On the one hand we have

$$\big[[H', H''], X_\alpha\big] = -\alpha(H')[X_\alpha, H'] + \alpha(H')[X_\alpha, H''] = 0 \tag{51.131}$$

and on the other hand we have $\big[[H', H''], X_\alpha\big] = [H, X_\alpha] = \alpha(H)X_\alpha$. We deduce that $\alpha(H) = 0$ for every root and then that $H = 0$ by lemma 51.82. □

We denote by Φ the set of roots. These are the elements $\lambda \in \mathfrak{h}^*$ such that $\mathfrak{g}_\lambda$ is non trivial. We suppose to have chosen a positivity notion on $\mathfrak{h}^*$, so that we can speak of Φ^+, the set of **positive roots**.

A positive root is **simple** is it cannot be written as the sum of two positive roots.

51.10.3 Generators

We are going to build the Chevalley basis of the complex semisimple Lie algebra $\mathfrak{g}$. That will essentially be a choice of a basis vector in each of the root spaces. We are following the notations summarized in point 51.10.1.3.

Now, for each root α, we pick $e_\alpha \in \mathfrak{g}_\alpha$. We will see that, up to renormalization, we can set the in nice commutation relations.

LemBalpahbetaef

Lemma 51.85.
If α and β are roots such that $\alpha + \beta \neq 0$, then

$$B(e_\alpha, e_\beta) = 0. \tag{51.132}$$

If $f_\alpha \in \mathfrak{g}_{-\alpha}$ we also have $B(e_\alpha, f_\alpha) \neq 0$.

Proof. By definition $B(e_\alpha, e_\beta) = \mathrm{Tr}\big(\mathrm{ad}(e_\alpha) \circ \mathrm{ad}(e_\beta)\big)$. If we apply $\mathrm{ad}(e_\alpha) \circ \mathrm{ad}(e_\beta)$ to an element of e_γ (including $\mathfrak{g}_0 = \mathfrak{h}$), we get an element of $\mathfrak{g}_{\alpha+\beta+\gamma}$. Thus the trace defining the Killing form is zero and $B(e_\alpha, e_\beta) = 0$ when $\alpha + \beta = 0$.

Since the Killing form is nondegenerate, we conclude that $B(e_\alpha, e_{-\alpha}) \neq 0$. □

CorrExistInverseRoot

Corollary 51.86.
Let $\mathfrak{g}$ be a semisimple complex Lie algebra and $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Let α be a root of $\mathfrak{g}$ and $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. There exist an unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$.

Proof. We have $[e_\alpha, f_\alpha] = B(e_\alpha, f_\alpha)t_\alpha$ and the lemma 51.85 shows that the Killing form is non zero. Multiplying by a suitable number provides the result. □

The element $H_\alpha \in \mathfrak{h}$ defined in this lemma is the **inverse root** of α.

Lemalphaakbetaknimport

Lemma 51.87.
Let $\{\beta_1, \dots, \beta_l\}$ be a choice of elements in $\mathfrak{h}^*$ such that the set $\{t_{\beta_1}, \dots, t_{\beta_l}\}$ is a basis of $\mathfrak{h}$. Thus the roots can be decomposed as

$$\alpha = \sum_{k=1}^{l} a_k \beta_k \tag{51.133}$$

with $a_k \in \mathbb{Q}$.

Proof. Let $\alpha = \sum_{k=1}^{l} a_k\beta_k$. We know that the vectors t_{β_i} form a basis of $\mathfrak{h}$, so we have the decomposition $t_\alpha = \sum_k a_k t_{\beta_k}$. Indeed

$$B\big(h, \sum_k a_k t_{\beta_k}\big) = \sum_k a_k B(h, t_{\beta_k}) = \sum_k a_k\beta_k(h) = \alpha(h). \tag{51.134}$$

For each $k = 1, 2, \dots, l$ we have

$$(\alpha_k, \alpha) = \sum_{j=1}^{l} a_k(\alpha_k, \alpha_j). \tag{51.135}$$

This is a system of linear equations for the l variables a_k. Since the coefficients (α_k, α) and (α_k, α_j) are rational by proposition 51.79, the solutions are rational too. □

Remarque 51.88.
The lemma 51.87 deals with a quite general basis of $\mathfrak{h}$. We will see in the proposition 51.103 that in the case of the basis of simple roots, the coefficients a_k are integers, either all positive or all negative.

51.10.4 Subalgebra $\mathfrak{sl}(2)_i$

SubSecCopiedeSLdansGi

For each nonzero root $\alpha \in \mathfrak{h}^*$, we choose $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ in such a way to have

$$B(e_\alpha, f_\alpha) = \frac{2}{B(t_\alpha, t_\alpha)}, \tag{51.136}$$

and then we pose

$$h_\alpha = \frac{2}{B(t_\alpha, t_\alpha)} t_\alpha. \tag{51.137}$$

Notice that these choices are possible because the Killing form is non degenerated on $\mathfrak{h}$.

PropWEzZYzC

Proposition 51.89 ([858]).
For each root, the set $\{e_\alpha, f_\alpha, h_\alpha\}$ generates an algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$, i.e. they satisfy

$$[h_\alpha, e_\alpha] = 2e_\alpha \tag{51.138a}$$
$$[h_\alpha, f_\alpha] = -2f_\alpha \tag{51.138b}$$
$$[e_\alpha, f_\alpha] = h_\alpha \tag{51.138c}$$
$$\tag{51.138d}$$

Proof. Since $\alpha(t_\alpha) = B(t_\alpha, t_\alpha)$ we have $\alpha(h_\alpha) = 2$. Now the computations are quite direct. The first is

$$[h_\alpha, e_\alpha] = \alpha(h_\alpha)e_\alpha = 2e_\alpha. \tag{51.139}$$

For the second,

$$[h_\alpha, f_\alpha] = -\alpha(h_\alpha)f_\alpha = -2f_\alpha. \tag{51.140}$$

For the third, we know that $[e_\alpha, f_\alpha] \in \mathfrak{h}$; thus $B\big(X, [e_\alpha, f_\alpha]\big) = 0$ whenever $X \in \mathfrak{g}_\lambda$ with $\lambda \neq 0$. Let $h \in \mathfrak{h}$. Using the invariance of the Killing form,

$$B\big(h, [e_\alpha, f_\alpha]\big) = B\big([h, e_\alpha], f_\alpha\big) = \alpha(h)B(e_\alpha, f_\alpha) = B(t_\alpha, t_\alpha)B(e_\alpha, f_\alpha) = B\big(B(e_\alpha, f_\alpha)t_\alpha, h\big). \tag{51.141}$$

Thus

$$[e_\alpha, f_\alpha] = B(e_\alpha, f_\alpha)f_\alpha = h_\alpha. \tag{51.142}$$

□

Remark that we used the non degeneracy of the Killing form in a crucial way. The copy of $\mathfrak{sl}(2, \mathbb{R})$ formed by $\{e_\alpha, f_\alpha, h_\alpha\}$ is denoted by $\mathfrak{sl}(2, \mathbb{R})_\alpha$.

Proposition 51.90.
In the universal enveloping algebra,

$$[h_j, f_i^{k+1}] = -(k+1)\alpha_i(h_j) f_i^{k+1} \tag{51.143}$$

Eqhifikplusun

as generalisation of the previous one.

Proof. We use an induction over k. Since $\mathrm{ad}(h_j)$ is a derivation in $\mathcal{U}(\mathfrak{g})$, the induction hypothesis and the definition relation $[h, f_i] = -\alpha_i(h) f_i$ with $h = h_i$, we have

$$\begin{aligned} \mathrm{ad}(h_j) f_i^{k+1} &= \big(\mathrm{ad}(h_j) f_i^k\big) f_i + f_i^k \,\mathrm{ad}(h_j) f_i. \\ &= -k\alpha + i(h_j) f_i^k f_i - \alpha_i(h_j) f_i^{k+1} \\ &= -(k+1)\alpha_i(h_j) f_i^{k+1}. \end{aligned} \tag{51.144}$$

□

Now the Lie algebra $\mathfrak{g}$ can be seen as a $\mathfrak{sl}(2,\mathbb{R})$-module. As an example, for each choice of $\beta \in \Phi$, the algebra $\mathfrak{sl}(2)_\alpha$ acts on the vector space

$$V = \bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{\beta+k\alpha}. \tag{51.145}$$

The vector space $\mathfrak{g}$ carries thus several representations of $\mathfrak{sl}(2)$; this fact will be used in a crucial way during the proof of proposition 51.101.

51.10.5 Chevalley basis

The Chevalley basis corresponds to an other choice of normalization of the element e_α, h_α. If we set

$$\begin{cases} H_\alpha = K_\alpha t_\alpha & \text{(51.146a)} \\ E_\alpha = N_\alpha e_\alpha & \text{(51.146b)} \end{cases}$$

with

$$\begin{aligned} K_\alpha &= \frac{2}{(\alpha,\alpha)} \\ N_\alpha &= \sqrt{\frac{2}{B(e_\alpha, e_{-\alpha})(\alpha,\alpha)}}, \end{aligned} \tag{51.147}$$

then we have the **Chevalley relations**:

EqsChevalleuRels

$$[E_\alpha, E_{-\alpha}] = H_\alpha \tag{51.148a}$$

$$[H_\alpha, E_\beta] = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} E_\beta \tag{51.148b}$$

$$[H_\alpha, H_\beta] = 0. \tag{51.148c}$$

The last relation is nothing else than the fact that the Cartan subalgebra $\mathfrak{h}$ is abelian. Notice that we don't give relations between E_α and E_β. Of course $[E_\alpha, E_\beta] \sim E_{\alpha+\beta}$ but the spaces $\mathfrak{g}_\alpha$ and $\mathfrak{g}_\beta$ being Killing orthogonal, the Killing does not provides a natural relative normalisation between E_α and E_β.

DefORftFjP

Definition 51.91.
If $\{\alpha_i\}_{i=1,\dots,l}$ is the set of simple roots, we consider the notation $X_i^+ = E_{\alpha_i}$, $X_i^- = E_{-\alpha_i}$, $H_i = H_{\alpha_i}$ and we introduce the ***Cartan matrix***

$$A_{ij} = \frac{2(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}. \tag{51.149}$$

Reduced to the simple roots the relations (51.148) become

$$\begin{aligned}[X_i^+, X_j^-] &= \delta_{ij} H_i \\ [H_i, X_j^\pm] &= \pm A_{ij} X_j^\pm \\ [H_i, H_j] &= 0.\end{aligned} \tag{51.150}$$ EqChevalleySimple

The first relation comes from the fact that $\alpha_i - \alpha_j$ is not a root when α_i and α_j are simple roots.

Remarque 51.92.
The idea behind the Chevalley relations is that the algebra $\mathfrak{g}$ is generated by the elements $X_i^\pm$, H_i and the commutation relations (51.150). Even if these elements do not form a basis (while the elements E_α, H_α do), one can define a function on $\mathfrak{g}$ by giving its values on $X_i^\pm$ and H_i providing one has a canonical way to extend it on commutators.

The definition 54.147 of standard cobracket works in this way.

Exemple 51.93.
The basis $\{H, E, F\}$ given by

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{51.151}$$

satisfy the relations subEqsSBhuAWx

$$[H, E] = 2E \tag{51.152a}$$
$$[H, F] = -2F \tag{51.152b}$$
$$[E, F] = H. \tag{51.152c}$$

which are nothing else than the relations (51.150).

The only positive root is $\alpha(H) = 2$. The Cartan matrix[15] reduces to one number:

$$A = A_{11} = \frac{2(\alpha, \alpha)}{(\alpha, \alpha)} = 2. \tag{51.153}$$

The Killing form, in the basis $\{H, E, F\}$ is given by

$$B = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix} \tag{51.154}$$

and the element t_α is then

$$t_\alpha = \frac{1}{4} H. \tag{51.155}$$

The inner product on $\mathfrak{h}^*$ is then

$$(\alpha, \alpha) = B(t_\alpha, t_\alpha) = \frac{1}{2}. \tag{51.156}$$ Eqinnerhstarsldc

△

Remarque 51.94.
Notice that these relations do not give the value of

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta} \tag{51.157}$$

when $\alpha + \beta$ is a root.

¡¡ Avertissement/question au lecteur !! 51.95
It has to be possible to compute $N_{\alpha,\beta}$, but I do not know how. The answer is given in equation (51.395) but I don't know where I got them. Maybe there are some hints in [857] (Il faut ajouter Cornwell à la bibliographie et enlever le problème 66.9).

15. Definition 51.91.

¡¡ Avertissement/question à la lectrice !! 51.96
It seems that A_{ij} is the larger integer k such that $\alpha_i + k\alpha_j$ is a root. This is the justification of the other Serre's relations that read

$$\mathrm{ad}^{1-A_{ij}}(X_i^{\pm})X_j^{\pm} = 0. \tag{51.158}$$

That relation has to be written with the Chevalley's ones.

One can choose the coefficients in a more scientific way[856]. Let α be a positive root, let H_α be the inverse root of α and $e_\alpha \in \mathfrak{g}_\alpha$. We have

$$[e_\alpha, e_\beta] = \begin{cases} N_{\alpha,\beta}e_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root} \\ 0 & \text{if } \alpha+\beta \text{ is not a root.} \end{cases} \tag{51.159}$$

We are going to find a multiple E_α of e_α in such a way to have in the same time

$$\begin{cases} [E_\alpha, E_{-\alpha}] = H_\alpha & \text{(51.160a)} \\ N_{\alpha,\beta} = -N_{-\alpha,-\beta}. & \text{(51.160b)} \end{cases}$$

Let σ be an involutive automorphism of $\mathfrak{g}$ such that $\sigma|_{\mathfrak{h}} : -\mathrm{Id}$. First we have $\sigma(\mathfrak{g}_\alpha) = \mathfrak{g}_{-\alpha}$ because

$$[h, \sigma(e_\alpha)] = \sigma[\sigma(h), e_\alpha] = -\sigma\alpha(h)e_\alpha = -\alpha(h)\sigma(e_\alpha) \tag{51.161}$$

for every $h \in \mathfrak{h}$ and $e_\alpha \in \mathfrak{g}_\alpha$. From corollary 51.86 there exist a number a_α such that

$$[e_\alpha, \sigma(e_\alpha)] = a_\alpha H_\alpha. \tag{51.162}$$

We pose

$$\begin{cases} E_\alpha = \dfrac{1}{\sqrt{-a_\alpha}}e_\alpha & \text{(51.163a)} \\ E_{-\alpha} = -\sigma(E_\alpha). & \text{(51.163b)} \end{cases}$$

With that choice we immediately have $[E_\alpha, E_{-\alpha}] = H_\alpha$. We also have $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$; in order to see it, consider

$$[\sigma E_\alpha, \sigma E_\beta] = \sigma[E_\alpha, E_\beta] = N_{\alpha,\beta}\sigma(E_{\alpha+\beta}) = -N_{\alpha,\beta}E_{-\alpha-\beta}. \tag{51.164}$$

But the same is also equal to

$$[-E_{-\alpha}, -E_{-\beta}] = [E_{-\alpha}, E_{-\beta}] = N_{-\alpha,-\beta}E_{-\alpha,-\beta}. \tag{51.165}$$

Proposition 51.97.
With these choices we have

$$N_{\alpha,\beta} = \pm(p+1) \tag{51.166}$$

where p is the largest integer j such that $\beta - j\alpha$ is a root.

¡¡ Avertissement/question au lecteur !! 51.98
I don't know a proof of that, but [856] gives a reference.

From proposition 51.77 we know that $t_\alpha \in \mathfrak{h}_\alpha$, so that H_α is a multiple of H_α. The proportionality factor is easy to fix since

$$\begin{aligned} \alpha(H_\alpha) &= 2 && \text{definition of } H_\alpha \\ \alpha(t_\alpha) &= (\alpha,\alpha) && \text{definition (51.108).} \end{aligned} \tag{51.167}$$

EqRealsHalpatalphaNorsmd EqRealsHalpatalphaNorsmd

Thus $H_\alpha = \frac{2}{(\alpha,\alpha)}t_\alpha$ and

$$[H_\alpha, E_\beta] = \beta(H_\alpha)E_\beta = \frac{2}{(\alpha,\alpha)}\beta(t_\alpha) = \frac{2(\alpha,\beta)}{(\alpha,\beta)} \tag{51.168}$$

again by the definition (51.108).

51.10.6 Coefficients in the Cartan matrix

In this section we search to give the form of the coefficients in the Cartan matrix. We will show that the values of (α, β) are quite restricted.

Remarque 51.99.
The notations are not standard. Here the symbol Δ denotes the set of *simple* roots while the set of all roots is denoted by Φ. In the book [857], the symbol Δ is the set of all roots. This makes quite a difference!

Definition 51.100.
*If α and β are roots of the complex semisimple Lie algebra $\mathfrak{g}$, then the α-**string** containing β is the set of roots of the form $\alpha + k\beta$ with $k \in \mathbb{Z}$.*

Among other things, the following proposition shows that a string has no gap.

Proppqasbabaa

Proposition 51.101.
Let $\alpha, \beta \in \Phi$. Then there exits integers p, q such that $\{\beta + k\alpha\}_{-p \leqslant k \leqslant q}$ is the α-string containing β. The numbers p and q satisfy

$$p - q = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \tag{51.169}$$

Eq2qbaaapmq

and the form

$$\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \tag{51.170}$$

is a nonzero root.

Proof. We consider the vector space

$$V = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha} \tag{51.171}$$

and the Lie algebra $\mathfrak{sl}(2)_\alpha = \langle e_\alpha, f_\alpha, h_\alpha \rangle$ defined in subsection 51.10.4. The latter acts on V. Simple computation using the fact that $\beta(h_\alpha) = 2(\alpha, \beta)/(\alpha, \alpha)$ shows that

$$[\frac{1}{2}h_\alpha, e_{\beta + k\alpha}] = \left(\frac{(\alpha, \beta)}{(\alpha, \alpha)} + k \right) e_{\beta + k\alpha}. \tag{51.172}$$

Thus the matrix of $\mathrm{ad}(\frac{1}{2}h_\alpha)$ is diagonal and has no multiplicity in its eigenvalues. We deduce that the representation if irreducible. From general theory of irreducible representations of $\mathfrak{sl}(2)$ we know that there exists an half-integer number j such that the diagonal entries of $\mathrm{ad}(\frac{1}{2}h_\alpha)$ take *all* the values from $-j$ to j by integer steps. Thus the α-string containing β has the form $\{\beta + k\alpha\}_{-p \leqslant k \leqslant q}$ where p and q satisfy

$$\begin{cases} \dfrac{(\alpha, \beta)}{(\alpha, \alpha)} - p = -j & \text{(51.173a)} \\ \dfrac{(\alpha, \beta)}{(\alpha, \alpha)} + q = j. & \text{(51.173b)} \end{cases}$$

Summing we get

$$p - q = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}. \tag{51.174}$$

If λ is an eigenvalue of $\mathrm{ad}(\frac{1}{2}h_\alpha)$, then $-\lambda$ is also an eigenvalue (this is still from the irreducible representation theory of $\mathfrak{sl}(2)$). The number $(\alpha, \beta)/(\alpha, \beta)$ is obviously an eigenvalue (with $k = 0$), thus the string contains a k such that

$$\frac{(\alpha, \beta)}{(\alpha, \alpha)} + k = -\frac{(\alpha, \beta)}{(\alpha, \alpha)}. \tag{51.175}$$

The solution is $k = -2(\alpha,\beta)/(\alpha,\alpha)$ and we deduce that

$$\beta - 2\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha \tag{51.176}$$

is a root of $\mathfrak{g}$. □

Proposition 51.102.
Let α, β be two roots. Then we have

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 0, \pm 1, \pm 2, \pm 3. \tag{51.177}$$

Proof. First, equation (51.169) shows that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is integer. If $\alpha = \pm\beta$, the result is 2. If $\alpha \neq \pm\beta$, the vectors t_α and t_β are linearly independent and the Schwarz inequality shows

$$(\alpha,\beta)^2 = |B(t_\alpha, t_\beta)| < B(t_\alpha, t_\alpha)B(t_\beta, t_\beta) = (\alpha,\alpha)(\beta,\beta). \tag{51.178}$$

Thus

$$\left|\frac{2(\alpha,\beta)}{(\alpha,\alpha)}\right|\left|\frac{2(\alpha,\beta)}{(\beta,\beta)}\right| < \frac{4|(\alpha,\beta)(\alpha,\beta)|}{(\alpha,\beta)^2} = 4. \tag{51.179}$$

Consequently the number $|2(\alpha,\beta)/(\alpha,\alpha)|$ being integer can only take the values 0, 1, 2 and 3. Notice that the inequality in (51.178) and (51.179) are strict since α_i is not collinear to α_j. □

51.10.7 Simple roots

As seen before, Φ admits an order inherited from $\mathfrak{h}^*_{\mathbb{R}}$. A root $\alpha > 0$ is **simple** if it cannot be written as a sum of two positive roots.

Theorem 51.103.
Let $\{\alpha_1, \ldots, \alpha_l\}$ be the set of simple roots. Then every root $\beta \in \Phi$ can be decomposed into

$$\beta = \sum_{i=1}^{l} n_i \alpha_i \tag{51.180}$$

where non vanishing the numbers $n_i \in \mathbb{Z}$ are either all positive or all negative.

Proof. Let β be positive. If it is not simple, the one can decompose it into two positive roots:

$$\beta = \gamma + \delta \tag{51.181}$$

with $\gamma, \delta > 0$. If γ and/or δ are not simple, they can be decomposed further. This process has to be finite, indeed if the process is not finite, the decomposition of at least one positive root has to contains itself (because there are finitely many of them) while it is impossible to have $\gamma = \gamma + \alpha$ with $\alpha > 0$. □

Two corollaries: a root is either positive or negative (this is part of the definition of positivity) and when a root is positive, its decomposition into simple roots has only positive coefficients.

51.10.8 Weyl group

References about Weyl group: [859]. See also [857], page 530.
If α is a root of $\mathfrak{g}$ we define the **symmetry** of α as

$$\begin{aligned} s_\alpha &: \mathfrak{h}^* \to \mathfrak{h}^* \\ &\beta \mapsto \beta - \beta(H_\alpha)\alpha \end{aligned} \tag{51.182}$$

where $H_\alpha \in \mathfrak{h}$ is the inverse root of α. Since $\alpha(H_\alpha) = 2$ we have $s_\alpha(\alpha) = -\alpha$. The group generated by the symmetries and the identity is the **Weyl group**.

From what is said around equation (51.167) and the definition $(\alpha,\beta)=\alpha(t_\beta)$, we have

$$s_\alpha(\beta)=\beta-\frac{2(\alpha,\beta)}{(\alpha,\alpha)}\alpha. \tag{51.183}$$

We know from proposition 51.101 that $s_\alpha(\beta)$ is a root while there are only finitely many roots; thus the Weyl group is finite since there are only a finite number of maps from a finite set to itself.

The symmetries associated to roots are involutive:

$$s_\alpha^2=\mathrm{Id}\,. \tag{51.184}$$

Indeed

$$\begin{aligned}s_\alpha^2(\beta)&=s_\alpha\big(\beta-\beta(H_\alpha)\alpha\big)\\&=\beta-\beta(H_\alpha)-\big(\beta-\beta(H_\alpha)\alpha\big)(H_\alpha)\alpha\\&=\beta\end{aligned} \tag{51.185}$$

if we take into account $\alpha(H_\alpha)=2$.

Relative to the symmetry s_{α_i} we have the symmetry s_i on $\mathfrak{h}$ defined by

$$s_i(h)=h-\alpha_i(h)H_i \tag{51.186}$$ EqSymsiReltosalphai

where $h\in\mathfrak{h}$ and H_i is the inverse root of α_i.

Remarque 51.104.
The simple roots α_i **are not** orthogonal.

Let Δ be a reduced abstract root system on a real finite dimensional vector space V. The group W generated by the $s_\alpha:\alpha\in\Delta$ is the **Weyl group**.

PropWeylIsomalphai

Proposition 51.105.
The elements s_{α_i} are isometries of $\mathfrak{h}^*$, i.e.

$$\big(s_{\alpha_i}(\alpha),s_{\alpha_i}(\beta)\big)=(\alpha,\beta). \tag{51.187}$$

Proof. For the sake of shortness, let us write

$$n_{i,\alpha}=\frac{2(\alpha_i,\alpha)}{(\alpha_i,\alpha_i)}. \tag{51.188}$$

We have $t_{s_{\alpha_i}(\alpha)}=t_\alpha-n_{i,\alpha}t_{\alpha_i}$. Thus

$$B\big(t_{s_{\alpha_i}(\alpha)},t_{s_{\alpha_i}(\beta)}\big)=B(t_\alpha-n_{i,\alpha}t_{\alpha_i},t_\beta-n_{i,\beta}t_{\alpha_i}) \tag{51.189}$$

distributing and taking into account the fact all the relations like $B(t_\alpha,t_{\alpha_i})=(\alpha,\alpha_i)$, the right hand side reduces to $B(t_\alpha,t_\beta)=(\alpha,\beta)$. □

When Φ is the root system, one can chose many different notions of positivity; each of them bring to different simple systems. It turns out that the action of the Weyl group on a simple system produces the simple system of an other choice of positivity on Φ.

LemalphajsPhipinjsasbab

Lemma 51.106.
If $s_{\alpha_i}\alpha=s_{\alpha_i}\beta$, then $\alpha=\beta$.

Proof. The hypothesis $s_{\alpha_j}(\alpha-\beta)=0$ provides

$$0=\alpha-\beta-\frac{2(\alpha-\beta,\alpha_j)}{(\alpha_j,\alpha_j)}\alpha_j \tag{51.190}$$

so that $\alpha=\beta+z\alpha_j$ for some $z\in\mathbb{C}$. Thus we have

$$s_{\alpha_j}(\alpha)=s_{\alpha_j}(\beta)+zs_{\alpha_j}(\alpha_j)=s_{\alpha_j}(\alpha)-z\alpha_j. \tag{51.191}$$

Thus $z=0$ and $\alpha=\beta$. □

PropsalphaisurjPhipmaj

Proposition 51.107.
Let α_i a simple root. The set $\Phi^+\backslash\{\alpha_i\}$ is stable under s_{α_i}, i.e.

$$s_{\alpha_i}\big(\Phi^+\backslash\{\alpha_i\}\big) = \Phi^+\backslash\{\alpha_i\}. \tag{51.192}$$

Proof. Let $\lambda \in \Phi^+$ be a positive root. By theorem 51.103 we have

Eqllamsumajajajnos

$$\lambda = \sum_j a_j\alpha_j \tag{51.193}$$

with $a_j \geqslant 0$. Since $\lambda \neq \alpha_i$ we have $a_j > 0$ for some $j \neq i$. Indeed the only multiple of α_i to be a root are 0 and $\pm\alpha_i$. Since $\lambda \in \Phi^+$ and $\lambda \neq \alpha_i$, none of these three solutions are taken into consideration.

Let's apply s_{α_i} on both sides of (51.193):

eqalialphaillamfracalphaj

$$\begin{aligned} s_{\alpha_i}(\lambda) &= s_{\alpha_i}\Big(\sum_j a_j\alpha_j\Big) \\ &= \sum_{j\neq i} a_j s_{\alpha_i}(\alpha_j) + a_i \underbrace{s_{\alpha_i}(\alpha_i)}_{-\alpha_i} \\ &= \sum_{j\neq i} a_j\alpha_j - \sum_{j\neq i} a_j \frac{2(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}\alpha_i - a_i\alpha_i \end{aligned} \tag{51.194}$$

Since a root is either positive or negative, the coefficients are either *all* positive or *all* negative. Since all the coefficients (apart for the one of α_i) are the same as the ones of λ, the root (51.194) is positive.

We still have to prove that $s_{\alpha_i}(\lambda) \neq \alpha_i$. Indeed if $s_{\alpha_i}(\lambda) = \alpha_i$ we have

$$\lambda = s_{\alpha_i}s_{\alpha_i}(\lambda) = s_{\alpha_i}(\alpha_i) = -\alpha_i, \tag{51.195}$$

which contradicts the positivity of λ.

Up to now we proved that $s_{\alpha_i}\big(\Phi^+\backslash\{\alpha_i\}\big) \subset \Phi^+\backslash\{\alpha_i\}$. If $\lambda \in \Phi^+\backslash\{\alpha_i\}$, then

$$\sigma = s_{\alpha_i}(\lambda) \in s_{\alpha_i}\big(\Phi^+\backslash\{\alpha_i\}\big) \subset \Phi^+\backslash\{\alpha_i\} \tag{51.196}$$

and $s_{\alpha_i}(\sigma) = \lambda$, so that λ is the image by s_{α_i} of $\sigma \in \Phi^+\backslash\{\alpha_i\}$. □

ThosajBijSurpPpsmaj

Theorem 51.108.
The map $s_{\alpha_j}\colon \Phi^+\backslash\{\alpha_j\} \to \Phi^+\backslash\{\alpha_j\}$ is bijective.

Proof. Surjectivity is proposition 51.107 while injectivity is lemma 51.106. □

Lemma 51.109 ([857], page 533)**.**
We consider the half sum of the positive roots:

$$\delta = \frac{1}{2}\sum_{\alpha\in\Phi^+} \alpha. \tag{51.197}$$

We have

(1) If α_j is a simple root, $s_{\alpha_j}\delta = \delta - \alpha_j$.
(2) If α_j is a simple root, $(\delta,\alpha_j) = \frac{1}{2}(\alpha_j,\alpha_j)$.

Proof. We compute $s_{\alpha_j}\delta$ dividing the sum into two parts:

$$s_{\alpha_j}\delta = \frac{1}{2}\sum_{\substack{\alpha\in\Phi^+\\ \alpha\neq\alpha_j}} s_{\alpha_j}(\alpha) + \frac{1}{2}s_{\alpha_j}(\alpha_j) \tag{51.198a}$$

$$= \frac{1}{2}\sum_{\substack{\alpha\in\Phi^+\\ \alpha\neq\alpha_j}} \alpha - \frac{1}{2}\alpha_j. \tag{51.198b}$$

The second inequality is from the fact that s_{α_j} is bijective on $\Phi^+\backslash\{\alpha_j\}$ by theorem 51.108. Adding a subtracting $\frac{\alpha_j}{2}$ we get

$$s_{\alpha_j}\delta = \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha - \frac{\alpha_j}{2} - \frac{\alpha_j}{2} = \delta - \alpha_j \tag{51.199}$$

Using the proposition 51.105, we have

$$(\delta,\alpha_j) = (s_{\alpha_j}\delta, s_{\alpha_j}\alpha_j) = (\delta-\alpha_j, -\alpha_j) = -(\delta,\alpha_j) + (\alpha_j,\alpha_j), \tag{51.200}$$

consequently, $2(\delta,\alpha_j) = (\alpha_j,\alpha_j)$ and the result follows. □

51.10.9 Abstract root system

The material about abstract root system mainly comes from [859].

DefAbsRootSystSerre

Definition 51.110.
*An **abstract root system** in a finite dimensional vector space V endowed with an is a subset Φ of V such that*

- *Φ is finite and $\operatorname{Span}\Phi = V$,*
- *For every $\alpha\in\Phi$, there is a symmetry s_α of vector α leaving Φ stable.*
- *For every $\alpha,\beta\in\Phi$, the vector $s_\alpha(\beta)-\beta$ is an integer multiple of α.*

The abstract system is **reduced** when $\alpha\in\Phi$ implies $2\alpha\notin\Phi$. It is **irreducible** is Φ doesn't admits non trivial decomposition as $\Phi = \Phi'\cup\Phi''$ with $(\alpha,\beta)=0$ for any $\alpha\in\Phi'$ and $\beta\in\Phi''$. We use the notation $\Phi := \Phi\cup\{0\}$.

The following is a consequence of all we did up to now.

Theorem 51.111.
The root system of a complex semisimple Lie algebra is a reduced abstract root system.

The **Weyl group** of Φ is the subgroup of $\mathrm{GL}(V)$ generated by the transformations s_α with $\alpha\in\Phi$.

51.10.9.1 Link with other definitions

The definition 51.110 is not the "usual" one (in [850], page 14 for example). We show now that we retrieve the usual features of an abstract.

Lemma 51.112.
An abstract root system admits a bilinear positive symmetric non degenerate form which is invariant under its Weyl group.

Proof. If $(.,.)_1$ is a bilinear positive non degenerate symmetric form on the vector space V, the form

$$(\alpha,\beta) = \sum_{w\in W}(w\alpha, w\beta)_1 \tag{51.201}$$

is invariant under the Weyl group. This construction is possible since the Weyl group is finite. □

Definition 51.113.
Let V be a vector space and $v\in V$ a non vanishing vector. A symmetry of vector v is an automorphism $s\colon V\to V$ such that

(1) $s(v) = -v$;

(2) the set $H = \{w\in V$ tel que $\alpha(w) = w\}$ is an hyperplane in V.

A symmetry of vector v induces the decomposition $V = H\oplus\mathbb{R}v$. The symmetries are of order 2: $s^2 = \mathrm{Id}$.

Lemma 51.114.
let v be a nonzero vector of V and A be a finite part of V such that $\mathrm{Span}(A) = V$. *Then there exists at most one symmetry of vector v leaving A invariant.*

Proof. Let s and s' be two such symmetries and consider $u = ss'$. We immediately have $u(A) = A$ and $u(v) = v$. Let us prove that u induce the identity on the quotient $V/\mathbb{R}v$. A general vector in V can be written (in a non unique way) under the form

$$h + h' + v \tag{51.202}$$

with $h \in H$ and $h' \in H'$. Let $h = h'_1 + \beta v$ be the decomposition of h in $H' \oplus \mathbb{R}v$ and $h' = h_1 + \gamma v$ be the decomposition of h' with respect to the direct sum $V = H \oplus \mathbb{R}v$. Then we have

$$\begin{aligned}
ss'(h + h' + \alpha v) &= ss'\big((h'_1 + \beta v) + h' + \alpha v\big) && \text{(51.203a)}\\
&= s\big((h'_1 - \beta v) + h' + \alpha v\big) && \text{(51.203b)}\\
&= s(h - 2\beta v + h_1 + \gamma v + \alpha v) && \text{(51.203c)}\\
&= h + 2\beta v + h_1 - \gamma v + \alpha v && \text{(51.203d)}\\
&= h + h' + (\alpha - 2\gamma + 2\beta)v. && \text{(51.203e)}
\end{aligned}$$

Thus at the level of the quotient, u leaves invariant $h + h'$.

It is not guaranteed that u is the identity, but the eigenvalues of u are 1. For each $x_i \in A$, there exists $n_i \in \mathbb{N}$ such that $u^{n_i}x_i = x$. If n is a common multiple of all the n_i (these are finitely many), we have $u^n(x) = x$ for every $x \in A$. Since A generates V, we have $u^n = \mathrm{Id}$ and then u is diagonalizable.

We already mentioned the fact that the eigenvalues of u are 1. Since u is diagonalizable, it is the identity and $s = s'$. □

The invariant form give to V a structure of euclidian vector space for which the elements of the Weyl group are orthogonal matrices. Thus the symmetries read

$$s_\alpha(x) = x - 2\frac{(x,\alpha)}{(\alpha,\alpha)}\alpha. \tag{51.204}$$ EqSymparrmusul

This is the only transformation which makes $s_\alpha(\alpha) = -\alpha$ in the same time as being implemented by an orthogonal matrix. The symmetry s_α is nothing else than the orthogonal symmetry with respect to the hyperplane orthogonal to α.

The expression (51.204) has the consequence that

$$s_\alpha(\beta) - \beta = -\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha. \tag{51.205}$$

By the definition of an abstract root system, the latter has to be an integer multiple of α, so

$$\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}. \tag{51.206}$$

Definition 51.115.
Two abstract root systems Φ on V and Φ' on V' are ***isomorphic*** *is there exists an isomorphism of vector space $\psi\colon V \to V'$ such that $\psi(\Phi) = \Phi'$ and*

$$2\frac{(\alpha,\beta)}{(\alpha,\alpha)} = 2\frac{\big(\psi(\alpha),\psi(\beta)\big)}{\big(\psi(\alpha),\psi(\alpha)\big)} \tag{51.207}$$

for every $\alpha, \beta \in \Phi$.

51.10.9.2 Basis of abstract root system

The part about basis of abstract root system comes from [856].

DefbasisabsRoot

Definition 51.116.
*Let Φ be an abstract root system. A part $S \subset \Phi$ is a **basis** of Φ if*

(1) S is a basis of V as vector space;

(2) every $\beta \in \Phi$ can be written under the form

$$\beta = \sum_{\alpha \in S} m_\alpha \alpha \tag{51.208}$$

where m_α are all integers of the same sign.

The set Δ of simple roots of the root system of a complex semisimple Lie algebra is a basis.

We are going to build a basis of an abstract root system. Let $h \in V^*$ be such that $\alpha(h) \neq 0$ for every $\alpha \in \Phi$ and define

$$\Phi_h^+ = \{\alpha \in \Phi \text{ tel que } \alpha(h) > 0\}. \tag{51.209}$$

We have $\Phi = \Phi_h^+ \cup -\Phi_h^+$. We say that an element $\alpha \in \Phi_h^+$ is **decomposable** if there exist $\beta, \gamma \in \Phi_h^+$ such that $\alpha = \beta + \gamma$. We write S_h the set of undecomposable elements in Φ_h^+.

LemShPhihpCBLSh

Lemma 51.117.
Any element in Φ_h^+ is a linear combination with positive coefficients of elements of S_h.

¡¡ Avertissement/question à la lectrice !! 51.118
It seems to me that Serre's book[856] has a misprint here. At page V-11 he writes:

> *Tout élément de R_t^+ est combinaison linéaire, à coefficients entiers $\geqslant 0$ des éléments de S.*

Shouldn't he have written S_t.

Proof. Let I be the set of $\alpha \in \Phi_h^+$ that cannot be written under such a decomposition. We choose $\alpha \in I$ such that $\alpha(h)$ is minimal. If α is undecomposable, then $\alpha \in S_h$ and the condition $\alpha \in I$ is contradicted. Thus α is decomposable. Let $\beta, \gamma \in \Phi_h^+$ be such that $\alpha = \beta + \gamma$. Since $\alpha(h)$ is minimal,

$$\begin{aligned} \beta(h) &\leqslant \alpha(h) \\ \gamma(h) &\leqslant \alpha(h). \end{aligned} \tag{51.210}$$

Thus we have $\beta(h) = \alpha(h) - \gamma(h) < 0$ which contradicts $\beta \in \Phi^+$. We conclude that I is empty. □

51.10.9.3 Properties

The main properties of an abstract root system are given in the following proposition.

PropPropAbstrRootviiiikl

Proposition 51.119.
If Φ is an abstract root system in a vector space V, one has the following properties:

enubi

(1) If $\alpha \in \Phi$ then $-\alpha \in \Phi$.

enubii

(2) If $\alpha \in \Phi$, the multiples of α which could also be in Φ are either $\pm\alpha$, or $\pm\alpha$ and $\pm 2\alpha$ or $\pm\alpha$ and $\pm\frac{1}{2}\alpha$.

enubiii

(3) If $\alpha\beta \in \Phi$ then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ can take the nonzero values ± 1, ± 2, ± 3 or ± 4. The case ± 4 can only arise if $\beta = \pm 2\alpha$.

enubiv

(4) If $\alpha, \beta \in \Phi$ are not proportional each other and if $|\alpha| \leqslant |\beta|$, then $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ equals 0 or ± 1.

enubv

(5) If $\alpha, \beta \in \Phi$ and $(\alpha, \beta) > 0$, then $\alpha - \beta \in \Phi$ and if $(\alpha, \beta) < 0$, the $\alpha + \beta \in \Phi$.

enubvi

(6) If $\alpha, \beta \in \Phi$ and neither $\alpha + \beta$ neither $\alpha - \beta$ belongs to Φ, then $(\alpha, \beta) = 0$.

enubvii

(7) *If $\alpha \in \Phi$ and $\beta \in \Phi$, the $n \in \mathbb{Z}$ such that $\beta + n\alpha \in \Phi$ fulfils $-p \leqslant n \leqslant q$ for certain $p, q \geqslant 0$. Moreover there are no gap,*

$$p - q = \frac{2(\alpha, \beta)}{(\alpha, \alpha)},$$

and there are at most four roots in the set $\{\beta + n\alpha\}_{-p \leqslant n \leqslant q}$.

enubviii

(8) *If Φ is reduced,*

enubviiia

(8a) *If $\alpha \in \Phi$, the only multiples of α to lies in Φ are $\pm\alpha$,*

enubviiib

(8b) *If $\alpha \in \Phi$ and $\beta \in \Phi$, then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ can be equal to 0, ± 1, ± 2 or ± 3.*

prop:Cartan_matr

The proof will not use the fact that Φ spans V.

Proof. (1) $s_\alpha \alpha = -\alpha$.

(2) If $\beta = c\alpha$ with $|c| < 1$, then

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 2c$$

must belongs to $\mathbb{Z}$, then $c = 0, \pm\frac{1}{2}$. If $|c| > 1$, we use exactly the same with $\alpha = \frac{1}{c}\beta$, so that $\frac{1}{c} = 0; \pm\frac{1}{2}$. Now if 2α is a root, it is clear that $\frac{1}{2}\alpha$ can't be.

If Φ is reduced, the fact that $\frac{1}{2}\alpha \in \Phi$ implies that $\alpha \notin \Phi$, so that $\pm\frac{1}{2}\alpha$ is excluded if $\alpha \in \Phi$, under the same assumption, 2α is also excluded. This proves (8a).

(3) The Schwartz inequality $|(\alpha, \beta)| \leqslant |\alpha||\beta|$ gives

$$\left| \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \frac{2(\beta, \alpha)}{(\beta, \beta)} \right| \leqslant 4.$$

The equality only holds for $\beta = c\alpha$. In this case, we just saw that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 2c$ with $c = 2$ at most. If the equality is strict, then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ and $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ are two integers whose product is $\leqslant 3$. The possibilities are 0, ± 1, ± 2, ± 3.

(4) If $|\alpha| \leqslant |\beta|$, then the following integer inequality holds:

$$\left| \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \right| \leqslant \left| \frac{2(\beta, \alpha)}{(\beta, \beta)} \right|.$$

Since the product of the two is $\leqslant 3$, the smallest is 0 or 1.

(5) If $\beta = c\alpha$, then $c = \pm\frac{1}{2}, \pm 2, \pm 1$. All the cases are easy. If $(\alpha, \beta) > 0$, then $c > 0$ and $\alpha - \beta = \alpha - \frac{1}{2}\alpha = \frac{1}{2}\alpha$ or $\alpha - \beta = \alpha - 2\alpha = -\alpha$.

Then we can suppose that α and β are not proportional each other. We consider $\alpha, \beta \in \Phi$ and $(\alpha, \beta) > 0$ (the other case is proved in much the same way). We just saw in (4) that $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ could be equals to 0 or ± 1, then the fact that $(\alpha, \beta) > 0$ imposes $\frac{2(\beta,\alpha)}{(\beta,\beta)} = 1$, so that $s_\beta(\alpha) = \alpha - \beta$.

If $|\beta| \leqslant |\alpha|$, we use

$$s_\alpha(\beta) = \beta - \frac{2(\beta, \alpha)}{(\beta, \beta)}\alpha = \beta - \alpha, \tag{51.211}$$

(6) is an immediate consequence of the previous point.

(7) Let $-p$ and q be the smallest and the largest values of n such that $\beta + n\alpha \in \Phi$. They exist because Φ is a finite set. Suppose that there is a gap between r and s ($r < s - 1$), i.e. $\beta + r\alpha \in \Phi$, $\beta + s\alpha \in \Phi$, but $\beta + (r + 1)\alpha, \beta + (s - 1)\alpha \notin \Phi$.

By the point (5), $(\beta + r\alpha, \alpha) \geqslant 0$ and $(\beta + s\alpha, \alpha) \leqslant 0$. Making the difference between these two inequalities,

$$(r - s)|\alpha|^2 \geqslant 0,$$

then $r \geqslant s$, which contradict the definition of r and s. So there is no gap. Now let us compute

$$\begin{aligned} s_\alpha(\beta + n\alpha) &= \beta + n\alpha - \frac{2(\alpha, \beta + n\alpha)}{(\alpha, \alpha)}\alpha \\ &= \beta + n\alpha - \left(\frac{2(\alpha, \beta)}{(\alpha, \alpha)} + 2n\right)\alpha \\ &= \beta - n\alpha - \frac{2(\alpha, \beta)}{(\alpha, \alpha)}\alpha \in \Phi. \end{aligned} \tag{51.212}$$

Then for any n in $-p \leqslant n \leqslant q$,

$$-q \leqslant n + \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \leqslant p.$$

With $n = q$, the second inequality gives $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \leqslant p - q$ while the first one with $n = -p$ gives $p - q \leqslant \frac{2(\alpha,\beta)}{(\alpha,\alpha)}$.

The last point is to check the length of the string of root. We can suppose $q = 0$ (i.e to look the string of $\beta - q\alpha$ instead of the one of α; of course this is the same), then the length is $p + 1$ and

$$p = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}.$$

If α and β are not proportional, the point (3) makes it equals at most to 3. If they are proportional, then the possibilities are $\alpha = \pm\beta, \pm\frac{1}{2}\beta, \pm 2\beta$. The string $\beta + n\alpha$ with $\alpha = \beta$ is at most $\{\beta, 2\beta\}$, if $\alpha = \frac{1}{2}\beta$, this is just $\{\beta\}$ and if $\alpha = 2\beta$, this is $\{\beta, -\beta\}$.

The proof is complete. □

LemShabShablesz

Lemma 51.120.
If $\alpha, \beta \in S_h$, then $(\alpha, \beta) \leqslant 0$.

Proof. If $(\alpha, \beta) \geqslant 0$, then proposition 51.119(5) shows that $\gamma = \alpha - \beta$ is a root. There are two possibilities: $\gamma \in \pm\Phi_h^+$. If $\gamma \in \Phi_h^+$, then $\alpha = \gamma + \beta$ is decomposable; contradiction. If $\gamma \in -\Phi_h^+$, then $\beta = \alpha - \gamma$ is decomposable; contradiction. □

LemIndepAhVstar

Lemma 51.121 (Lemme 4 page V-12)**.**
Let $h \in V^*$ and $A \subset V$ be a subset satisfying

(1) $\alpha(h) > 0$ for every $\alpha \in A$;

(2) $(\alpha, \beta) \leqslant 0$ for every $\alpha, \beta \in A$.

Then the elements in A are linearly independent.

Proof. Let us consider a vanishing linear combination of elements in A:

$$\sum_{\alpha \in A} m_\alpha \alpha = 0. \tag{51.213}$$

EqNullCombinsumAmAuAd

We can sort the terms following that m_α is positive or negative and cut the sum in two parts:

$$\sum_{\beta \in A_1} y_\beta \beta = \sum_{\gamma \in A_2} z_\gamma \gamma \tag{51.214}$$

with $y_\beta, z_\gamma \geqslant 0$ and where A_1 and A_2 are disjoint subsets of A. Let us consider $\lambda = \sum_{\beta \in A_1} y_\beta \beta$ and compute

$$(\lambda, \lambda) = \sum_{\substack{\beta \in A_1 \\ \gamma \in A_2}} y_\beta z_\gamma (\beta, \gamma). \tag{51.215}$$

EqllamllamnprofAunAdeuxsom

By hypothesis (β, γ) is lower than zero and by construction the product y_β, z_γ is positive. Thus the right hand side of equation (51.215) is negative. We conclude that $\lambda = 0$. Thus

$$0 = \lambda(h) = \sum_{\beta \in A_1} y_\beta \beta(h). \tag{51.216}$$

Since all the terms in the sum are larger than zero we have $y_\beta = 0$. In the same way we get $z_\gamma = 0$. The vanishing linear combination (51.213) is then trivial and the elements of A are linearly independent. □

PropSestShsi

Proposition 51.122.
The elements of S_h form a basis of Φ in the sense of definition 51.116. Conversely, if S is a basis of Φ and if $h \in V^*$ is such that $\alpha(h) > 0$ for every $\alpha \in S$,we have $S = S_h$.

Proof. The set S_h satisfies the conditions of lemma 51.121 since by definition $\alpha(h) > 0$ for every $\alpha \in S_h$ and by lemma 51.120 the inner products are all negative. Thus S_h is a free set. It is generating by lemma 51.117. Again by lemma 51.117, every element in Φ can be written as sum of elements of S_h with all coefficients of the same sign. Here we use the fact that v is positive if and only if $-v$ is negative and that every vector is either positive or negative.

For the second part, let S be a basis and $h \in V^*$ such that $\alpha(h) > 0$ for all $\alpha \in S$. Let

$$\Phi^+ = \{\sum_{\alpha\in S} m_\alpha \alpha \text{ with } m_\alpha \in \mathbb{N}\}. \tag{51.217}$$

We have $\Phi^+ \subset \Phi_h^+$ and $-\Phi^+ \subset -\Phi_h^+$. Since $\Phi = \Phi^+ \cup -\Phi^+$ we also have $\Phi^+ = \Phi_h^+$. Since elements of S are indecomposable in Φ^+, they are indecomposable in Φ_h^+ and we have $S \subset S_h$.

The sets S and S_h have the same number of elements because they both are basis of V, thus $S = S_h$. □

LemwShShpahwahp

Lemma 51.123.
If h and h' are elements of V^* related by $\alpha(h) = (w\alpha)h'$, then $w(S_h) = S_{h'}$ (if these space can be defined).

Proof. Let $\alpha \in S_h$. The element $w(\alpha)$ belongs to $\Phi_{h'}^+$ because

$$w(\alpha)h' = \alpha(h) > 0 \tag{51.218}$$

because $\alpha \in \Phi_h^+$. We still have to check that $w(\alpha)$ is undecomposable in $\Phi_{h'}^+$. If $w(\alpha) = \beta + \gamma$ with $\beta, \gamma \in \Phi_{h'}^+$, we have $\alpha = w^{-1}\beta + w^{-1}\gamma$. From the link between h and h' we have

$$(w^{-1}\beta)(h) = (ww^{-1}\beta)h' = \beta(h') > 0. \tag{51.219}$$

Thus $w^{-1}\beta \in \Phi_h^+$ which is a contradiction because we supposed that α is undecomposable. □

Lemswwsbwemucirc

Lemma 51.124.
If $\alpha, \beta \in \Phi$ and if $w \in W_S$, then $s_{w(\beta)} = w \circ s_\beta \circ w^{-1}$.

Proof. Using the fact that the symmetries are isometries of the inner product,

$$s_{w(\beta)}(\alpha) = \alpha - \frac{\big(w(\beta), \alpha\big)}{\big(w(\beta), w(\beta)\big)} w(\beta) = \alpha - \frac{(\beta), w^{-1}\alpha}{(\beta, \beta)} w\beta. \tag{51.220}$$

Applying that to $w(\alpha)$ instead of α and applying w^{-1}, we have

$$\begin{aligned} w^{-1}s_{w(\beta)}\big(w(\alpha)\big) &= w^{-1}\left(w\alpha - \frac{(\beta, w^{-1}w\alpha)}{(\beta,\beta)} w\beta\right) && \text{(51.221a)}\\ &= \alpha - \frac{(\beta,\alpha)}{(\beta,\beta)} w^{-1}w\beta && \text{(51.221b)}\\ &= s_\beta(\alpha). && \text{(51.221c)} \end{aligned}$$

□

ThoWeylGenere

Theorem 51.125 ([856]).
Let W be the Weyl group of the abstract root system Φ. Let S a basis of Φ and W_S the subgroup of W generated by s_α with $\alpha \in S$. Then

ItemThoWeylGenerei
(1) *for every $h \in V^*$, there exists $w \in W_S$ such that $(w\alpha)(h) \geqslant 0$ for every $\alpha \in S$.*
ItemThoWeylGenereii
(2) *If S' is a basis of Φ, the there exists a $w \in W_S$ such that $w(S') = S$.*
ItemThoWeylGenereiii
(3) *For every $\beta \in \Phi$ there exists $w \in W_S$ such that $w(\beta) \in S$.*
ItemThoWeylGenereiv
(4) *The group W is generated by the symmetries s_α with $\alpha \in S$.*

Proof. For item (1), consider $h \in V^*$ and $\delta = \frac{1}{2}\sum_{\gamma\in S}\gamma$. Let $w \in W_S$ be such that $w(\delta)h$ is the largest possible [16]. If $\alpha \in S$ we have

$$w(\delta)h \geqslant ws_\alpha(\delta)h = w(\delta)h - w(\alpha)h, \tag{51.222}$$

so that $w(\alpha)h \geqslant 0$ for every $\alpha \in S$. This proves our first assertion.

We pass to point (2). Let $h' \in V^*$ be such that $\alpha'(h') > 0$ for every $\alpha' \in S'$. By the first item there exists $w \in W_S$ such that

$$(w\alpha)(h') \geqslant 0 \tag{51.223}$$

for every $\alpha \in S$. In fact we even have $w\alpha h' > 0$ for every $\alpha \in S$. Indeed $w\alpha$ can be decomposed as $\sum_{\alpha'\in S'} m_{\alpha'}\alpha'$ where all the $m_{\alpha'}$ have the same sign. In this case

$$(w\alpha)h' = \sum_{\alpha'} m_{\alpha'}\alpha'(h') \neq 0 \tag{51.224}$$

Eqwapsurmaraphp

because each of $\alpha'(h')$ is strictly positive while all the terms of the sum have the same sign. This means, by the way, that $S' = S_{h'}$ following the proposition 51.122.

We define $h \in V^*$ by the relation

$$\alpha(h) = (w\alpha)(h'). \tag{51.225}$$

By what we said in equation (51.224) we have $\alpha(h) > 0$ for every $\alpha \in S$, so that we have $S = S_h$. Finally by lemma 51.123, $w(S_h) = S_{h'}$.

We prove now the point (3). For $\gamma \in \Phi$ we consider the hyperplane

$$L_\gamma = \{h \in V^* \text{ tel que } \gamma(h) = 0\}. \tag{51.226}$$

Consider a particular $\beta \in \Phi$ the hyperplanes L_γ with $\gamma \neq \pm\beta$ do not coincide with L_β and there is only finitely many of them, so there exists a $h_0 \in L_\beta$ such that h_0 do not belong to any L_γ for any $\gamma \neq \pm\beta$. In particular we have $\beta(h_0) = 0$ and $\gamma(h_0) \neq 0$ for every $\gamma \in \Phi$, $\gamma \neq \pm\beta$. If we choose ϵ small enough, there exists h near from h_0 such that

$$\begin{cases} \beta(h) = \epsilon > 0 & \text{(51.227a)} \\ |\gamma(h)| > \epsilon \qquad \text{if } \gamma \neq \pm\beta. & \text{(51.227b)} \end{cases}$$

Let S_h be the basis associated with this h. We have $\beta \in S_h$. Indeed first $\beta(h) = \epsilon > 0$ and if $\beta = \gamma + \rho$, we would have

$$\gamma(h) = \beta(h) - \rho(h) = \epsilon - \rho(h) < 0, \tag{51.228}$$

so that β is undecomposable in Φ_h^+. Now from point (2) there exists $w \in W_S$ such that $w(S_h) = S$. In particular $w(\beta) \in S$.

We turn our attention to the item (4). We are going to prove that $W = W_S$. Since W is generated by the symmetries s_β ($\beta \in \Phi$), it is sufficient to prove that W_S generates the symmetries s_β.

Let $\beta \in \Phi$ and consider the element $w \in W_S$ such that $\alpha = w(\beta) \in S$. From lemma 51.124 we have

$$s_\alpha = s_{w(\beta)} = w \circ s_\beta \circ w^{-1}, \tag{51.229}$$

so that

$$s_\beta = w^{-1} \circ s_\alpha \circ w \in W_S. \tag{51.230}$$

□

16. We can consider that w because W is finite.

What this theorem says in the case of complex semisimple Lie algebras is that if $\{\alpha_1, \dots, \alpha_l\}$ is the set of simple roots, the symmetries s_{α_i} generate the Weyl group. Now, since any root can be mapped on a simple one using the Weyl group, any root can be recovered from a simple one acting with the Weyl group that is generated by the simple ones.

Thus one can determine all the roots from the data of the simple ones by computing $s_{\alpha_i}\alpha_j$ and then acting again with the s_{α_i} on the results and again and again. This is the fundamental reason from which the root system can be recovered for the Cartan matrix.

When we have the Cartan matrix A of a semisimple complex Lie algebra, the first point is to find the norm of the roots by finding the diagonal matrix D. We have $(\alpha_i, \alpha_i) = D_{ii}$. For the other products we write

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{D_{ii}}, \tag{51.231}$$

thus

$$(\alpha_i, \alpha_j) = \frac{D_{ii}A_{ij}}{2}. \tag{51.232}$$

51.10.10 Abstract Cartan matrix

The following proposition summarize the properties of the of the Cartan matrix.

DeabstrCartanmatr

Definition 51.126.
*A matrix $(A_{ij})_{1\leqslant i,j\leqslant l}$ satisfying the following conditions is an **abstract Cartan matrix***

(1) $A_{ij} \in \mathbb{Z}$,

(2) $A_{ii} = 2$,

ItempoprCartaniii

(3) $A_{ij} \leqslant 0$ *if* $i \neq j$,

(4) $A_{ij} = 0$ *if and only if* $A_{ji} = 0$,

ItempoprCartanv

(5) there exists a diagonal matrix D with positive coefficients such that DAD^{-1} is symmetric and positive defined.

The classification of abstract Cartan matrix will be performed in subsection 51.10.11. The data of an abstract Cartan matrix defines an abstract root system. For a proof, see [860].

Proposition 51.127.
The Cartan matrix of a semisimple complex Lie algebra is an abstract Cartan matrix.

Proof. The first two points are already done. For the point (3), note that the sign of (α, β) is not sure when α is any root. However here we are speaking of simple roots. Let us consider the root

$$\lambda = \alpha_i - \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}\alpha_i \tag{51.233}$$

Since it is a root, proposition 51.103 says that the coefficients in the decomposition in simple roots have to be all integer and of the same sign. Thus the combination $(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)$ has to be negative.

The point (5) is also non trivial. Consider the diagonal matrix $D = \operatorname{diag}\big((\alpha_i, \alpha_i)\big)_{i=1,\dots,l}$. We have

$$(DAD^{-1})_{ij} = \sum_{kl} D_{ik}A_{kl}(D^{-1})_{lj} \tag{51.234a}$$

$$= \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)^{1/2}(\alpha_j, \alpha_j)^{1/2}}. \tag{51.234b}$$

This is a symmetric matrix. In order to proof that this is positive defined, we are going to provide a matrix B such that $DAD^{-1} = BB^t$. Let $\{\lambda_i\}$ be an orthonormal basis of $\mathfrak{h}^*$ and consider the matrix b given by the decomposition of the simple roots in this basis:

$$\alpha_i = \sum_j b_{ij}\lambda_j. \tag{51.235}$$

In particular we have $(\alpha_i, \alpha_j) = \sum_k b_{ik} b_{jk}$. Then we consider the matrix

$$B_{ij} = \frac{b_{ij}}{(\alpha_i, \alpha_i)^{1/2}} \tag{51.236}$$

which is non degenerate since the α_i are simple and thus a re linearly independent. Small computation shows that

$$(BB^t)_{ij} = \sum_k \frac{b_{ik}}{(\alpha_i, \alpha_i)^{1/2}} \frac{b_{jk}}{(\alpha_j, \alpha_j)^{1/2}} \tag{51.237a}$$

$$= \frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)^{1/2} (\alpha_j, \alpha_j)^{1/2}} \tag{51.237b}$$

$$= (DAD^{-1})_{ij}. \tag{51.237c}$$

But BB^t is positive defined, then DAD^{-1} is. □

51.10.11 Dynkin diagrams

SubsecDynkindiam

The sources for Dynkin diagrams is [858, 850].

We are going to associate to each abstract Cartan matrix, a diagram that will uniquely correspond to an abstract root system. In other words what we are going to do is to classify the matrix satisfying the conditions of definition 51.126.

If A is an abstract Cartan matrix we build the **Dynkin diagram** of A with the following rules.

(1) We put l vertices (one for each root)

(2) The vertex i and j are joined with $A_{ij}A_{ji}$ lines.

A step by step construction is available in [850].

In the following we are considering an abstract Cartan matrix A and its associated abstract root system $\{\alpha_i\}$.

LesmabsCartDynk

Lemma 51.128.
A abstract Cartan matrix with its abstract root system and its Dynkin diagram have the following properties.

ItemLesmabsCartDynki

(1) If one remove the ith line an column of an abstract Cartan matrix, one still has an abstract Cartan matrix. In other words, each subdiagram of a Dynkin diagrm is a Dynkin diagram.

ItemLesmabsCartDynkii

(2) Two vertices are linked by at most *three lines.*

ItemLesmabsCartDynkiii

(3) Each Dynkin diagram has more vertices than linked pairs.

ItemLesmabsCartDynkiv

(4) A Dynkin diagram has no loop.

ItemLesmabsCartDynkv

(5) A vertex in a Dynkin diagram has at most three lines attached (including multiplicities). Note: this is a generalization of point (2).

ItemLesmabsCartDynkvi

(6) Two root linked by a simple edge have equal ***weight****, that is* $(\alpha_i, \alpha_i) = (\alpha_j, \alpha_j)$.

ItemLesmabsCartDynkvii

(7) If the two roots α_i, α_i are connected by a simple edge, we can collapse them, removing the connecting edge and conserving all the other edges.

Proof. For point (2) we have

$$A_{ij}A_{ji} = 4 \frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} \frac{(\alpha_j, \alpha_i)}{(\alpha_j, \alpha_j)} < 4 \tag{51.238}$$

by Cauchy-Schwarz inequality. We insist on the fact that the inequality is strict since α_i and α_j are not collinear: they are simple roots.

For point (3) consider the form

$$\gamma = \sum_{i=1}^{l} \alpha_i (\alpha_i, \alpha_i)^{1/2}. \tag{51.239}$$

Since the simple roots are linearly independent, this sum is nonzero and we have $0 < (\gamma, \gamma)$. We have

$$\begin{aligned} 0 < (\gamma, \gamma) &= \sum_{ij} \frac{(\alpha_i, \alpha_j)}{\sqrt{(\alpha_i, \alpha_i)(\alpha_j, \alpha_j)}} \\ &= 2 \sum_{i<j} \frac{(\alpha_i, \alpha_j)}{\sqrt{(\alpha_i, \alpha_i)(\alpha_j, \alpha_j)}} + \text{number of nodes} \\ &= -\sum_{i<j} (A_{ij} A_{ji})^{1/2} + \text{number of nodes}. \end{aligned} \tag{51.240}$$

Since for each linked pair (i, j) we have a term $A_{ij}A_{ji} \geqslant 1$, we have

$$-\sum_{i<j} (A_{ij} A_{ji})^{1/2} \leqslant \text{number of pairs} \tag{51.241}$$

and the positivity of the sum shows that

$$\text{number of nodes} > \sum_{ij} A_{ij} A_{ji} \geqslant \text{number of pairs}. \tag{51.242}$$

For item (4), suppose that a loop is given by the roots $\alpha_1, \ldots, \alpha_n$. Since any sub-Dynkin diagram is a Dynkin diagram (from point (1)), we can consider only the loop. This is a diagram with n vertices and n pairs, which contradicts point (3).

We pass to item (5). Let α_0 be a root linked to n simple lines, m double lines and p triple lines. For notational convenience, we write $v_i = \alpha_i/(\alpha_i, \alpha_i)$, $\{v_i\}_{1 \leqslant i \leqslant n}$ is the set of "simply" linked roots to α_0, $\{v'_i\}_{1 \leqslant i \leqslant m}$ the set of "doubly" linked and $\{v''_i\}_{1 \leqslant i \leqslant p}$ the set of "triply" ones. Consider the vector

$$\gamma = v_0 + \sum_{i=1}^{n} f_i v_i + \sum_{i=1}^{m} g_i v'_i + \sum_{i=1}^{p} h_i v''_i \tag{51.243}$$

where f_i, g_i and h_i are constant to be determined. In order to compute the norm of γ, notice that since there are no loops, no lines join v_i, v'_i and v''_i together, so we have $(v_i, v'_j) = (v_i, v''_j) = (v'_i, v''_j) = 0$ and from the number of lines, $(v_0, v_i) = -1/2$, $(v_0, v'_i) = -1/\sqrt{2}$ and $(v_0, v''_i) = -\sqrt{3}/2$. Thus we have

$$(\gamma, \gamma) = 1 + \sum_{i=1}^{m} (f_i^2 - f_i) + \sum_{i=1}^{m} (g_i^2 - \sqrt{2} g_i) + \sum_{i=1}^{p} (h_i^2 - \sqrt{3} h_i). \tag{51.244}$$

The minimum is realised with $f_i = 1/2$, $g_i = \sqrt{2}/2$ and $h_i = \sqrt{3}/2$ and for these values we have

$$(\gamma, \gamma) = 1 - \frac{n + 2m + 3p}{4}. \tag{51.245}$$

Since the inner product has to be positive we must have $n + 2m + 3p < 4$, the is the number of lines issued from α_0 has to be lower or equal to 3.

In order to proof (6), remark that if α_i and α_j are connected by a simple edge, then $A_{ij}A_{ji} = 1$, which is only possible with $A_{ij} = A_{ji} = -1$. In particular we have $2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i) = 2(\alpha_j, \alpha_i)/(\alpha_j, \alpha_j)$, which proves that $(\alpha_i, \alpha_i) = (\alpha_j, \alpha_j)$.

Proof of item (7). Since the two roots have same weight, the item (6) says that up to permutation the Cartan matrix has a block 2×2 looking like

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{51.246}$$

The proposed move consist to replace that block with the 1×1 matrix (2). As an example,

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & -1 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}. \tag{51.247}$$

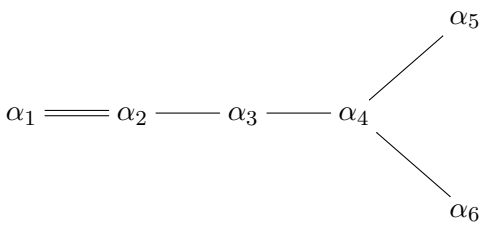
It is clear that the obtained matrix is still an abstract Cartan matrix. □

From these properties we can deduce much constrains on the Dynkin diagrams. First, the only diagram containing a triple edge is

$$\alpha_1 \equiv\equiv \alpha_2 \tag{51.248}$$

Let pass to the diagrams with only simple and double edges. If there is a double, there cannot be a triple point: the following is impossible

(51.249)



since collapsing the roots α_2, α_3 and α_4 should create a point with four edges. Thus a diagram with a double edge is only possible inside a straight chain. Let us study the diagram

$$\alpha_1 - \alpha_2 \equiv\equiv \alpha_3 - \alpha_4 - \alpha_5 \tag{51.250}$$

Once again we denote $v_i = \alpha_i/|\alpha_i|$ and we consider the (non vanishing) vector

$$\gamma = v_1 + bv_2 + cv_3 + dv_4 + ev_5 \tag{51.251}$$

whose norm is given by

$$(\gamma, \gamma) = 1 + b^2 + c^2 + d^2 + e^2 - b - \sqrt{2}bc = cd = de. \tag{51.252}$$

Equating all the partial derivative to zero provides the point

$$b = 2 \quad c = \frac{3}{\sqrt{2}} \quad d = \sqrt{2} \quad e = \frac{1}{\sqrt{2}}. \tag{51.253}$$

One check that with these values $(\gamma, \gamma) = 0$ which is impossible. The diagram (51.250) is thus impossible. By the collapsing principle, all the diagrams of the form

$$\alpha_1 - \alpha_2 \equiv\equiv \alpha_3 - \alpha_4 - \ldots - \alpha_l \tag{51.254}$$

are impossible. The only possible diagrams with double edge are thus

$$\alpha_1 - \alpha_2 \equiv\equiv \alpha_3 - \alpha_4 \tag{51.255a}$$

$$\alpha_1 \equiv\equiv \alpha_2 - \ldots - \alpha_l \tag{51.255b}$$

$$\alpha_1 - \alpha_2 - \ldots - \alpha_{l-1} \equiv\equiv \alpha_l. \tag{51.255c}$$

The diagrams (51.255b) and (51.255c) are the same. They however do not completely determine the abstract Cartan matrix because the diagram (51.255c) induces an asymmetry between α_1 and α_2. The so written Dynkin diagram cannot distinguish between the matrices

$$\begin{pmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \tag{51.256}$$

Thus we split the diagram (51.255c) into [suBeqdfynkabGP]

$$\alpha_1 \text{ —— } \alpha_2 \text{ —— } \ldots \text{ —— } \alpha_{l-1} \Longrightarrow \alpha_l. \tag{51.257a}$$

$$\alpha_1 \text{ —— } \alpha_2 \text{ —— } \ldots \text{ —— } \alpha_{l-1} \Longleftarrow \alpha_l. \tag{51.257b}$$

In which the arrow points to the biggest root. The first one means that $|\alpha_1| = \ldots = |\alpha_{l-1}| = 1$, $\alpha_l = 2$ while the second diagram means $|\alpha_1| = \ldots = |\alpha_{l-2}| = |\alpha_l| = 1$, $\alpha_{l-1} = 2$.

We'll have to come back on this point later in subsection 51.10.12. Notice that this is the only diagram on which that problem occurs.

We are left to study the diagrams with only single edge. The following diagram is the simplest possible one:

$$\alpha_1 \text{ —— } \alpha_2 \text{ —— } \ldots \text{ —— } \alpha_l. \tag{51.258}$$

We have to know under what conditions one can have a triple point. We already know that there can be only one triple point.

If a diagram has a triple point, then one of the branch is of length 1. Indeed if not we would have the following diagram:

$$\begin{array}{ccccccc} & & & & \alpha_2 & \text{——} & \alpha_5 \\ & & & \nearrow & & & \\ \alpha_7 & \text{——} & \alpha_4 & \text{——} & \alpha_1 & & \\ & & & \searrow & & & \\ & & & & \alpha_6 & \text{——} & \alpha_6 \end{array} \tag{51.259}$$

Looking at the vector $\gamma = 3v_1 + 2(v_2 + v_3 + v_4) + v_5 + v_6 + v_7$ provides $(\gamma, \gamma) = -3$ which is impossible. Thus the diagrams with a branch are straight chains with one unique triple point which has a branch of length one. The question is: where can happen that branch? The diagram

$$\begin{array}{ccccccccccccc} \alpha_1 & \text{——} & \alpha_2 & \text{——} & \alpha_3 & \text{——} & \alpha_4 & \text{——} & \alpha_5 & \text{——} & \alpha_6 & \text{——} & \alpha_7 \\ & & & & & & \vert & & & & & & \\ & & & & & & \alpha_8 & & & & & & \end{array} \tag{51.260}$$

cannot happen since the corresponding vector $v_1 + 2v_2 + 3v_3 + 4v_4 + 3v_5 + 2v_6 + v_7 + 2v_8$ has norm zero. Thus on a triple point, one branch has one branch of length 1 and at least one other to be of length 1 or 2. It turns out that all the diagrams of the form

$$\begin{array}{ccccccc} & & & & \alpha_{l-1} & & \\ & & & \nearrow & & & \\ \alpha_1 & \text{——} \ \ldots \ \text{——} & \alpha_{l-2} & & & & \\ & & & \searrow & & & \\ & & & & \alpha_l & & \end{array} \tag{51.261}$$

are possible. We are thus left with diagrams with a triple point with a branch of length 1 and a branch of length 2:

$$\begin{array}{ccccccccccc} \alpha_1 & \text{——} & \alpha_2 & \text{——} & \alpha_3 & \text{——} & \alpha_5 & \text{——} \ \ldots \ \text{——} & \alpha_l \\ & & & & \vert & & & & \\ & & & & \alpha_4 & & & & \end{array} \tag{51.262}$$

The diagram with a branch of length 5

$$\begin{array}{ccccccccccccc} \alpha_1 & \text{——} & \alpha_2 & \text{——} & \alpha_3 & \text{——} & \alpha_5 & \text{——} & \alpha_6 & \text{——} & \alpha_7 & \text{——} & \alpha_8 & \text{——} & \alpha_9 \\ & & & & \vert & & & & & & & & & & \\ & & & & \alpha_4 & & & & & & & & & & \end{array} \tag{51.263}$$

does not exist. We achieve the proof of that fact using for example this code for sage:

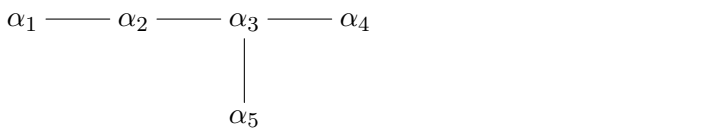
```
--------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                     |
| Type notebook() for the GUI, and license() for information.      |
--------------------------------------------------------------------
sage: a=[var('a'+str(i-1)) for i in range(1,11)]
sage: l=9
sage: a[1]=1
sage: squares = sum( [a[i]**2 for i in range(1,l+1)] )     # The sum goes to l

# The sum goes up to l-2
sage: lines = sum(  [a[i]*a[i+1] for i in range(1,l-1)  ]   )+a[3]*a[9]
sage: f=symbolic_expression(squares - lines)
sage: X = solve( [f.diff(a[i])==0 for i in range(2,l+1)],
                 [ a[i] for i in range(2,l+1)  ]  )
sage: print X[0]
[a2 == 2, a3 == 3, a4 == (5/2), a5 == 2,
a6 == (3/2), a7 == 1, a8 == (1/2),
a9 == (3/2)]
sage: f(*tuple(  [  X[0][i].rhs() for i in range(0,l-1)]) )
0
```

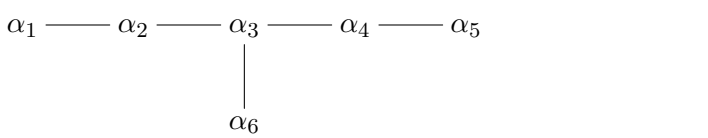
This proves that the vector $v_1 + 2v_2 + 3a_3 + \frac{5}{2}v_4 + 2v_5 + \frac{3}{2}v_6 + v_7 + \frac{1}{2}v_8 + \frac{3}{2}v_9$ has vanishing norm, which is impossible.

¡¡ Avertissement/question au lecteur !! 51.129
This code raises a deprecation warning that I'm not able to solve.

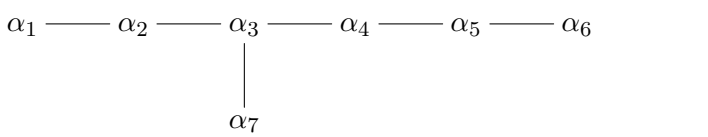
We are finally left with the diagrams with one triple point with one branch of length 1, one branch of length 2 and the third branch with length 1, 2, 3 or 4:



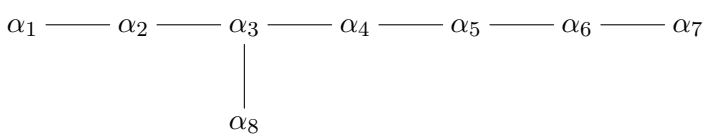
(51.264a)



(51.264b)



(51.264c)



(51.264d)

In order to list all the possible complex semisimple Lie algebra, we have to check each of the left Dynkin diagrams if they give rise to an abstract Cartan matrix.

```
# -*- coding: utf8 -*-
from sage.all import *


def Dynkin():
```

```
    """
    This proves that the Dynkin diagram with a triple point whose branches are
    of length 1,3,5 does not exist.

    The result is zero.
    """
    a=[var('a'+str(i-1)) for i in range(1,11)]
    l=9
    a[1]=1
    squares = sum( [a[i]**2 for i in range(1,l+1)] )      # The sum goes to l
    lines = sum(  [a[i]*a[i+1] for i in [1,2,3,4,5,6,7]  ]   )+a[3]*a[9]  # The sum goes up to l-1
    f=symbolic_expression(squares - lines)
    X = solve( [f.diff(a[i])==0 for i in range(2,l+1)],[ a[i] for i in range(2,l+1)  ]  )   # We put a1=1, so we don't solve with respect to a1
    print f(   *tuple(  [  X[0][i].rhs() for i in range(0,l-1)  ]  )   )


def gXgemu():
    a,b,t=var('a,b,t')
    X=matrix(3,3,[0,a,b,-a,0,0,-b,0,0])
    g=matrix(3,3,[cos(t),sin(t),0,-sin(t),cos(t),0,0,0,1])
    A=g*X*g.inverse()
    for i in range(0,3):
        for j in range(0,3):
            print i,j," -- ",A[i,j].simplify_full()
    Y=matrix(3,3,[0,t,0,-t,0,0,0,0,0])
    g=Y.exp()
    print g
    for i in range(0,3):
        for j in range(0,3):
            print i,j," -- ",g[i,j].real_part().simplify_full()
```

tex/research/calculs.py

51.10.12 Example of reconstruction by hand

subsecRecbyhanfd

We turn now our attention on the difference between the two diagrams (51.257). The Cartan matrix of the diagram $\alpha_1 \text{——} \alpha_2 \Longrightarrow \alpha_3$ is given by

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -1 & 2 \end{pmatrix}. \tag{51.265}$$

The diagonal matrix D of definition 51.126 is

$$D = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix} \tag{51.266}$$

and the length of the roots are $\|\alpha_1\| = \|\alpha_2\| = 1$ and $|\alpha_3| = 2$. Let us compute the angles between the roots. In order to compute (α_1, α_2) we look at A_{12}:

$$A_{12} = -1 = 2\frac{(\alpha_1, \alpha_2)}{(\alpha_1, \alpha_1)}, \tag{51.267}$$

and the same computation with A_{23} provides

$$(\alpha_1, \alpha_2) = -\frac{1}{2} \tag{51.268a}$$

$$(\alpha_2, \alpha_3) = -1 \tag{51.268b}$$

We compute all the roots using the theorem 51.125 which basically says that acting with the "simple" Weyl group W_S on the simple roots generates all the roots. On the first strike we have

$$\begin{aligned} &s_1(\alpha_2) = \alpha_2 + \alpha_1 \quad s_2(\alpha_1) = \alpha_1 + \alpha_2 \quad s_3(\alpha_1) = \alpha_1 \\ &s_1(\alpha_3) = \alpha\alpha_3 \qquad s_2(\alpha_3) = \alpha_3 + 2\alpha_2 \quad s_3(\alpha_2) = \alpha_2 + \alpha_3. \end{aligned} \tag{51.269}$$

We discovered the roots $\alpha_2 + \alpha_1$, $\alpha_3 + 2\alpha_2$ and $\alpha_2 + \alpha_3$. Acting again on these roots by s_{α_1}, s_{α_2} and s_{α_3} the only new results are

$$\begin{aligned} s_1(\alpha_3 + \alpha_2) &= \alpha_1 + \alpha_2 + \alpha_3 \\ s_1(\alpha_3 + 2\alpha_2) &= 2\alpha_1 + 2\alpha_2 + \alpha_3. \end{aligned} \tag{51.270}$$

Acting again we find only one new root:

$$s_{\alpha_2}(\alpha_1 + \alpha_2 + \alpha_3) = \alpha_1 + 2\alpha_2 + \alpha_3. \tag{51.271}$$

We check that acting once again with the three simple roots on this last one does not brings new roots. Thus we have 9 positive roots. Adding the negative ones, we are left with 18 root spaces of dimension one. The Cartan algebra has dimension 3, so the algebra we are looking at has dimension 21.

Now take a look at the similar Dynkin diagram and its Cartan matrix:

$$\alpha_1 \longrightarrow \alpha_2 \Longleftarrow \alpha_3 \qquad\qquad A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{pmatrix} \tag{51.272a}$$

The inner products are

$$\begin{aligned} |\alpha_1| = |\alpha_3| = 1, |\alpha_2| = 2 \\ (\alpha_1, \alpha_2) = -1/\sqrt{2}, (\alpha_2, \alpha_3) = -1 \end{aligned} \tag{51.273}$$

and the roots are

$$\alpha_1 \tag{51.274a}$$

$$\alpha_2 \tag{51.274b}$$

$$\alpha_3 \tag{51.274c}$$

$$\alpha_1 + \alpha_2 \tag{51.274d}$$

$$\alpha_2 + \alpha_3 \tag{51.274e}$$

$$\alpha_2 + 2\alpha_3 \tag{51.274f}$$

$$\alpha_1 + \alpha_2 + \alpha_3 \tag{51.274g}$$

$$\alpha_1 + \alpha_2 + 2\alpha_3 \tag{51.274h}$$

$$\alpha_1 + 2\alpha_2 + 2\alpha_3. \tag{51.274i}$$

We see that the inner products are already not the same. Notice that the roots are really different: it is not simply a renaming $\alpha_2 \leftrightarrow \alpha_3$.

Thus the two Dynkin diagrams (51.255c) are describing two different Lie algebras.

51.10.13 Reconstruction

The construction theorem is the following.

Theorem 51.130.
*Let R be an abstract root system in a complex vector space V^* and $\{\alpha_1, \dots, \alpha_n\}$ be a basis of R. We denote by $H_i \in V$ the **inverse root** of α_i (i.e. $\alpha(H_\alpha) = 2$). We define the Cartan matrix*

$$A_{ij} = \alpha_j(H_i). \tag{51.275}$$

Let $\mathfrak{g}$ be the Lie algebra defined by the $3n$ generators X_i, Y_i, H_i and the relations

$$[H_i, H_j] = 0 \tag{51.276a}$$
$$[X_i, Y_j] = \delta_{ij} H_i \tag{51.276b}$$
$$[H_i, X_j] = A_{ij} X_j \tag{51.276c}$$
$$[H_i, Y_j] = -A_{ij} Y_j \tag{51.276d}$$

and, for $i \neq j$,

$$\mathrm{ad}(X_i)^{-A_{ij}+1}(X_j) = 0 \tag{51.277a}$$
$$\mathrm{ad}(Y_i)^{-A_{ij}+1}(Y_j) = 0. \tag{51.277b}$$

EqSerrea

Then $\mathfrak{g}$ is a semisimple Lie algebra in which a Cartan subalgebra is generated by $H_1, \dots, H_n$ and its root system is R.

A complete proof can be found in [856] at page VI-19. We are going to give some ideas.

We consider $\mathfrak{g}$, the Lie algebra generated by the elements H_i, X_i and Y_i. We denote by $\mathfrak{h}$ the abelian Lie algebra generated by the elements H_i.

LemadXiNilpotent

Lemma 51.131.
The endomorphism $\mathrm{ad}(X_i)$ and $\mathrm{ad}(Y_i)$ are nilpotent.

Proof. Let V_i the subspace of $\mathfrak{g}$ of elements z such that $\mathrm{ad}(X_i)^k z = 0$ for some $k \in \mathbb{N}$. The space V_i is a Lie subalgebra of $\mathfrak{g}$ because

$$\mathrm{ad}(X_i)[z, z'] = -[z, \mathrm{ad}(X_i)z'] + [z', \mathrm{ad}(X_i)z]. \tag{51.278}$$

Acting with $\mathrm{ad}(X_i)^n$ we get terms of the form $[\mathrm{ad}(X_i)^k z, \mathrm{ad}(X_i)^l z']$ with $k + l = n$. If n is large enough, all the terms vanish.

From the relation (51.277a) we see that $X_j \in V_i$ for every j. Since $[X_i, H_j]$ is proportional to X_i we also have $H_j \in V_i$ and then $Y_j \in V_i$ because $[X_i, Y_j] = \delta_{ij} H_i \in V_i$. Thus the Lie algebra V_i contains all the Chevalley generators and then $V_i = \mathfrak{g}$. □

For $\lambda \in \mathfrak{h}^*$ we define

$$\mathfrak{g}_\lambda = \{z \in \mathfrak{g} \text{ tel que } \mathrm{ad}(h)z = \lambda(h)z \forall h \in \mathfrak{h}\}. \tag{51.279}$$

Then one prove that $\dim \mathfrak{g}_{\alpha_i} = 1$ and $\dim \mathfrak{g}_{m\alpha_i} = 0$ if $m \neq \pm 1, 0$. This corresponds to the fact that we have a reduced root system, which is always the case in complex semisimple Lie algebras [17]. We denote by Φ the subset of $\lambda \in \mathfrak{h}^*$ such that $\mathfrak{g}_\lambda \neq 0$.

It turns out that we have the direct sum decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha. \tag{51.280}$$

One of the key ingredients in this building is the following lemma.

17. However, at this point we have not proved yet that $\mathfrak{g}$ is semisimple and has that root system.

Lemma 51.132.
If λ and μ are related by an element of the Weyl group, then $\dim \mathfrak{g}_\lambda = \mathfrak{g}_\mu$.

Proof. Lemma 51.131 allows us to introduce the automorphism

$$\theta_i = e^{\operatorname{ad}(X_i)} e^{-\operatorname{ad}(Y_i)} e^{\operatorname{ad}(X_i)} \tag{51.281}$$

of $\mathfrak{g}$. We see that the restriction of θ_i to $\mathfrak{h}$ is the symmetry associated to α_i (see (51.186)). Indeed the first exponential reduces to

$$e^{\operatorname{ad}(X_i)} H_k = H_k - A_{ki} X_i \tag{51.282}$$

where $A_{ki} = \alpha_i(H_k)$. The second exponential gives

$$\begin{aligned} e^{\operatorname{ad}(-Y_i)}(H_k - A_{ki}X_i) &= H_k - A_{ki}X_i + (-A_{ki}Y_i - A_{ki}H_i) + \frac{1}{2}(2A_{ki}Y_i) \\ &= H_k - A_{ki}H_i - A_{ki}X_i. \end{aligned} \tag{51.283}$$

Notice the simplification of $A_{ki}Y_i$. The third exponential then provides the result (after some simplifications):

$$e^{\operatorname{ad}(X_i)}(H_k - A_{ki}H_i - A_{ki}X_i) = H_k - A_{ki}H_i = H_k - \alpha_i(H_k)H_i. \tag{51.284}$$

We proved that $\theta_i(H_k) = s_1(H_k)$. We deduce that $\theta_i e_\alpha \in \mathfrak{g}_{s_{\alpha_i}(\alpha)}$ whenever $e_\alpha \in \mathfrak{g}_\alpha$. Since θ_i is an automorphism of $\mathfrak{g}$ we have

$$[H_k, \theta_i e_\alpha] = \theta_i[\theta_i^{-1} H_k, e_\alpha]. \tag{51.285}$$

Since θ_i reduces to the involutive automorphism s_i on $\mathfrak{h}$ we have $\theta_i^{-1} H_k = \theta_i H_k = s_i(H_k)$. Then we have

$$[H_k, \theta_i e_\alpha] = \theta_i[s_i(H_k), e_\alpha] = \theta_i \alpha\big(s_i(H_k)\big) e_\alpha. \tag{51.286}$$

The eigenvalue of $\theta_i e_\alpha$ for $\operatorname{ad}(H_k)$ is thus $\alpha\big(s_i(H_k)\big)$. Using the definition and $A_{ki} = \alpha_i(H_k)$ we have

$$\begin{aligned} \alpha\big(s_i(H_k)\big) &= \alpha(H_k) - \alpha_i(H_k)\alpha(H_i) \\ &= \big(\alpha - \alpha(H_i)\alpha_i\big) H_k \\ &= s_{\alpha_i}(\alpha) H_k. \end{aligned} \tag{51.287}$$

At the end we got

$$[H_k, \theta_i e_\alpha] = s_{\alpha_i}(H_k) \theta_i e_\alpha \tag{51.288}$$

and then $\theta_i e_\alpha \in \mathfrak{g}_{s_{\alpha_i}(\alpha)}$. Thus the automorphism θ_i transforms $\mathfrak{g}_\lambda$ into $\mathfrak{g}_\mu$ when $\mu = s_i(\lambda)$ and

$$\dim \mathfrak{g}_\lambda = \dim \mathfrak{g}_{s_i(\lambda)}. \tag{51.289}$$

□

From here we prove that $\dim \mathfrak{g}_\alpha = 1$ for every root α[18].

Now if $\alpha + \beta = \gamma + \mu$, the elements $[E_\alpha, E_\beta]$ and $[E_\gamma, E_\mu]$ are proportional since they belong to the one-dimensional space $\mathfrak{g}_{\alpha+\beta}$.

RemChevDefmapCommXH

Remarque 51.133.
A linear map $\phi\colon \mathfrak{g} \to V$ from $\mathfrak{g}$ to a vector space V can be defined on the generators X_i, Y_i and H_i among with a formula giving $\phi([X, Y])$ in terms of $\phi(X)$ and $\phi(Y)$.

¡¡ Avertissement/question à la lectrice !! 51.134
This remark could be made more precise. I'm thinking to the proposition 54.57 giving the standard bialgebra structure on a Lie algebra.

The classification of complex semisimple Lie algebras is the following:

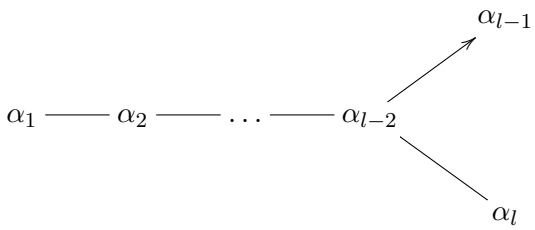
18. [856] page VI-23. Be careful: this is not the statement of page VI-2.

(1) A_l avec $l = 1, 2, \ldots$.
   — $A_l = \mathfrak{sl}(l+1, \mathbb{C})$
   — $\dim(A_l) = l(l+2)$
   — $\alpha_1 \text{——} \alpha_2 \text{——} \ldots \text{——} \alpha_l$

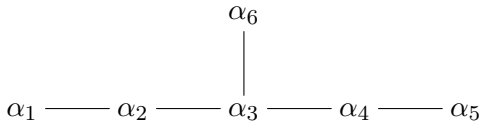
(2) B_l avec $l = 2, 3, \ldots$.
   — $B_l = \mathfrak{o}(2l+1, \mathbb{C})$
   — $\dim(B_l) = l(2l+1)$
   — $\alpha_1 \Longrightarrow \alpha_2 \text{——} \ldots \text{——} \alpha_{l-1} \text{——} \alpha_l$.

(3) C_l avec $l = 3, 4, \ldots$
   — $C_l = \mathfrak{sp}(l, \mathbb{C})$
   — $\dim(C_l) = l(2l+1)$
   — $\alpha_1 \longrightarrow \alpha_2 \text{——} \ldots \text{——} \alpha_{l-1} \Longrightarrow \alpha_l$.

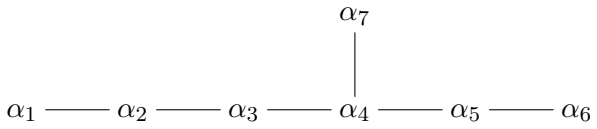
(4) D_l avec $l = 4, 5 \ldots$
   — $D_l = \mathfrak{o}(2l, \mathbb{C})$
   — $\dim(D_l) = l(2l-1)$
   —



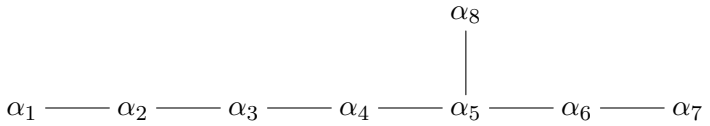
(5) E_6
   — E_6
   — $\dim(E_6) = 78$
   —



(6) E_7
   — E_7
   — $\dim(E_7) = 133$
   —



(7) E_8
   — E_8
   — $\dim(E_8) = 248$
   —



(8) F_4
   — F_4
   — $\dim(F_4) = 52$
   — $\alpha_1 \text{——} \alpha_2 \Longrightarrow \alpha_3 \text{——} \alpha_4$

(9) G_2
   — G_2
   — $\dim(G_2) = 14$
   — $\alpha_1 \lll \alpha_2$

51.10.14 Cartan-Weyl basis

Let us study the eigenvalue equation

$$\mathrm{ad}(A)X = \rho X. \tag{51.290}$$

EqvalradArho

The number of solutions with $\rho = 0$ depends on the choice of $A \in \mathfrak{g}$.

Lemma 51.135.
If A is chosen in such a way that $\mathrm{ad}(A)X = 0$ has a maximal number of solutions, then the number of solutions is equal to the rank of $\mathfrak{g}$ and the eigenvalue $\alpha = 0$ is the only degenerated one in equation (51.290).

We suppose A to be chosen in order to fulfill the lemma. Thus we have linearly independent vectors H_i $(i = 1, \dots l)$ such that

$$[A, H_i] = 0 \tag{51.291}$$

where l is the rank of $\mathfrak{g}$. Since $[A, A] = 0$, the vector A is a combination $A = \lambda^i H_i$. Since $\mathrm{ad}(A)$ is diagonalisable, one can find vectors E_α with

$$[A, E_\alpha] = \alpha E_\alpha, \tag{51.292}$$

and such that $\{H_i, E_\alpha\}$ is a basis of $\mathfrak{g}$. Using the fact that $\mathrm{ad}(A)$ is a derivation, we find

$$[A, [H_i, E_\alpha]] = \alpha [H_i, E_\alpha], \tag{51.293}$$

The eigenvalue $\alpha = 0$ being the only one to be degenerated, one concludes that $[H_i, E_\alpha]$ is a multiple of E_α:

$$[H_i, E_\alpha] = \alpha_i E_\alpha. \tag{51.294}$$

Replacing $A = \lambda^i H_i$, we have

$$\alpha E_\alpha = [\lambda^i H_i, E_\alpha] = \lambda^i \alpha_i E_\alpha, \tag{51.295}$$

thus $\alpha = \lambda^i \alpha_i$ (with a summation over $i = 1, \dots, l$).

Before to go further, notice that the space spanned by $\{H_i\}_{i=1,\dots,l}$ is a maximal abelian subalgebra of $\mathfrak{g}$, so that it is a Cartan subalgebra that we, naturally denote by $\mathfrak{h}^*$. Thus, what we are doing here is the usual root space construction. In order to stick the notations, let us associate the form $\sigma_\alpha \in \mathfrak{h}^*$ defined by $\sigma_\alpha(H_i) = \alpha_i$. In that case,

$$\sigma_\alpha(A) = \sigma_\alpha(\lambda^i H_i) = \lambda^i \alpha_i = \alpha \tag{51.296}$$

and we have

$$[A, E_\alpha] = \sigma_\alpha(A) E_\alpha. \tag{51.297}$$

On the other hand, we have $[H_i, E_\alpha] = \alpha_i E_\alpha = \sigma_\alpha(H_i) E_\alpha$, so that the eigenvalue α is identified to the root α, and we have $E_\alpha \in \mathfrak{g}_\alpha$.

Let us now express the vectors t_α in the basis of the H_i. The definition property is $B(t_\alpha, H_i) = \alpha(H_i) = \alpha_i$. If $t_\alpha = (t_\alpha)^i H_i$, we have

$$\alpha_i = B(t_\alpha, H_i) = B_{kl}(t_\alpha)^k \underbrace{(H_i)^l}_{=\delta_i^l} = B_{ki}(t_\alpha)^k. \tag{51.298}$$

If (B^{ij}) are the matrix elements of B^{-1}, we have

$$(l_\alpha)^l = \alpha_i B^{il} = \alpha^l \tag{51.299}$$

where α^l is defined by the second equality. Using proposition 51.77, we have

$$[E_\alpha, E_{-\alpha}] = B(E_\alpha, E_{-\alpha}) \alpha^l H_l. \tag{51.300}$$

Thus one can renormalise E_α in such a way to have

$$\begin{aligned}[H_i, H_j] &= 0,\\ [E_\alpha, E_{-\alpha}] &= \alpha^i H_i\\ [H_i, E_\alpha] &= \alpha_i E_\alpha = \alpha(H_i) E_\alpha\\ [E_\alpha, E_\beta] &= N_{\alpha\beta} E_{\alpha+\beta}\end{aligned} \tag{51.301}$$

where the constant $N_{\alpha\beta}$ are still undetermined. A basis $\{H_i, E_\alpha\}$ of $\mathfrak{g}$ which fulfill these requirements is a basis of **Cartan-Weyl**.

51.10.15 Cartan matrix

We follow [861]. We denote by Π the system of simple roots of $\mathfrak{g}$. All the positive roots have the form

$$\sum_{\alpha\in\Pi} k_\alpha \alpha \tag{51.302}$$

with $k_\alpha \in \mathbb{N}$.

Theorem 51.136.
Let α and β be simple roots Thus

(1) $\alpha - \beta$ is not a simple root

(2) we have

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = -p \tag{51.303}$$

EqabSuraaStrictNEf

where p is a strictly positive integer.

Partial proof. We are going to prove that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is an integer. Let α and γ be non vanishing roots such that $\alpha+\gamma$ is not a root, and define

$$E'_{\gamma-j\alpha} = \mathrm{ad}(E_{-\alpha})^k E_\gamma \in \mathfrak{g}_{\gamma-k\alpha}. \tag{51.304}$$

Since there are a finite number of roots, there exists a minimal positive integer g such that $\mathrm{ad}(E_{-\alpha})^{g+1} E_\gamma = 0$. We define the constants μ_k (which depend on γ and α) by

$$[E_\alpha, E'_{\gamma-k\alpha}] = \mu_k E'_{\gamma-(k-1)\alpha}. \tag{51.305}$$

Using the definition of $E'_{\gamma-k\alpha}$ and Jacobi, one founds

$$\mu_k E'_{\gamma-(k-1)\alpha} = \big[E'_\alpha, [E_{-\alpha}, E'_{\gamma-(k-1)\alpha}]\big] = \alpha^i [H_i, E'_{\gamma-(k-1)\alpha}] + \mu_{k-1} E'_{\gamma-(k-a)\alpha}, \tag{51.306}$$

so that $\mu_k = \alpha^i\gamma_i - (k-1)\alpha^i\alpha_i + \mu_{k-1}$, and we have the induction formula

$$\mu_k = (\alpha,\gamma) - (k-1)(\alpha,\alpha) + \mu_{k-1} \tag{51.307}$$

for $k \geqslant 2$. If we define $\mu_0 = 0$, that relation is even true for $k = 1$. The sum for $k = 1$ to $k = j$ is easy to compute and we get

$$\mu_j = j(\alpha,\gamma) - \frac{j(j-1)}{2}(\alpha,\alpha). \tag{51.308}$$

Since $\mu_{g+1} = 0$, we have

$$(\alpha,\gamma) = g(\alpha,\alpha)/2, \tag{51.309}$$

Eqalphagammaparqdeux

and thus

$$\mu_j = \frac{j(g-j+1)(\alpha,\alpha)}{2}. \tag{51.310}$$

Let β be any root and look at the string $\beta + j\alpha$. There exists a maximal $j \geqslant 0$ for which $\beta + j\alpha$ is a root while $\beta + (j+1)\alpha$ is not a root. Now we consider $\gamma = \beta + j\alpha$ with that maximal j. Putting $\gamma = \alpha + j\beta$ in (51.309), one finds

$$(\alpha, \beta) = \frac{(g - 2j)(\alpha, \alpha)}{2}, \tag{51.311}$$

and finally,

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = g - 2j, \tag{51.312}$$

which is obviously an integer.

□

From the inner product on $\mathfrak{h}^*$, we deduce a notion of **angle**:

$$\cos(\theta_{\alpha,\beta}) = \frac{(\alpha, \beta)}{\sqrt{(\alpha, \alpha)(\beta, \beta)}}. \tag{51.313}$$

The **length** of the root α is the number $\sqrt{(\alpha, \alpha)}$.

Lemma 51.137.
If α and β are roots, then

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}, \tag{51.314}$$

and

$$\beta - \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \tag{51.315}$$

is a root too.

If α and β are non vanishing, then the α-string which contains β contains at most 4 roots. Finally, the ratio

$$\frac{2(\alpha, \beta)}{(\alpha, \beta)} \tag{51.316}$$

takes only the values 0, ± 1, ± 2 or ± 3.

Let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a system of simple roots. The **Cartan matrix** is the $l \times l$ matrix with entries

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}. \tag{51.317}$$

EqDefMatriceCartan

Notice that, in the literacy, one find also the convention $A_{ij} = 2(\alpha_i, \alpha_j)/(\alpha_j, \alpha_j)$, as in [862], for example.

LemRatdjaijdjaji

Lemma 51.138.
There exist positive rational numbers d_i such that

$$d_i A_{ij} = d_j A_{ji} \tag{51.318}$$

EqdiAijdjAji

where A is the Cartan matrix.

Proof. The numbers are given by

$$d_i = \frac{(\alpha_i, \alpha_i)}{(\alpha_1, \alpha_1)}. \tag{51.319}$$

The relations (51.318) are easy to check using the definition (51.317). The fact that d_i is a strictly positive rational number comes from (51.303). □

¡¡ Avertissement/question au lecteur !! 51.139
I think that there is a property saying (something like) that A_{ij} is the larger integer k such that $\alpha_i + k\alpha_j$ is a root.

51.11 Other results

51.11.1 Abstract Cartan matrix

As before if we chose a basis $\{\varphi_1 \dots \varphi_l\}$ of V, we can consider a lexicographic ordering on V. A root is **simple** when it is positive and can't be written as as sum of two positive roots. As in a non abstract case, abstract simple root also have the following property:

Proposition 51.140.
If $\dim V = l$, one has only l simple roots $\alpha_1, \dots, \alpha_l$; they are linearly independent and if $\beta \in \Phi$ expands into $\beta = \sum c_j \alpha_j$, the c_j's all are integers and the non zero ones all have the same sign.

An ordering on V gives a notion of simple roots. The $l \times l$ matrix whose entries are

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}$$

is the **abstract Cartan matrix** of the abstract root system and the given ordering.

Theorem 51.141.
The main properties are

(1) $A_{ij} \in \mathbb{Z}$,

(2) $A_{ii} = 2$,

(3) if $i \neq j$, *then* $A_{ij} \leqslant 0$ *and* A_{ij} *can only take the values* $0, -1, -2$ *or* -3,

(4) if $i \neq j$, $A_{ij}A_{ji} < 4$ *(no sum),*

(5) $A_{ij} = 0$ *is and only if* $A_{ji} = 0$,

(6) $\det A$ *is integer and positive.*

Proof. The last point is the only non immediate one. The matrix A is the product of the diagonal matrix with entries $2/|\alpha_i|^2$ and the matrix whose entries are (α_i, α_j). The fact that the latter is positive definite is a general property of linear algebra. If $\{e_i\}$ is a basis of a vector space V, the matrix whose entry ij is given by (e_i, e_i) is positive definite. Indeed one can consider an orthonormal basis $\{f_i\}$ and a nondegenerate change of basis $e_i = B_{ik} f_k$. Then $(e_i, e_j) = (BB^t)_{ij}$. It is easy to see that for all $v \in V$, we have $(BB^t)_{ij} v^i v^j = \sum_k (v^i B_{ik})^2 > 0$.

The fact that the determinant is integer is simply the fact that this is a polynomial with integer variables. □

If we have an ordering on V we define Φ^+, the set of positive roots. From there, one can consider Π, the set of simple roots. Any element of Φ expands to a sum of elements of Π. Note that the knowledge of Π is sufficient to find Φ^+ back because $\alpha > 0$ implies $\alpha = \sum c_i \alpha_i$ with $c_i \geqslant 0$.

We can make this reasoning backward. Let us consider $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a basis of V such that any $\alpha \in \Phi$ expands as a sum of α_i with all coefficients of the same sign. Such a Π is a **simple system**. From such a Π, we can build a Φ^+ as the set of elements of the form $\alpha = \sum c_i \alpha_i$ with $c_i \geqslant 0$.

Proposition 51.142.
The so build Φ^+ is the set of positive roots for a certain ordering.

Proof. If we consider on V the lexicographic ordering with respect to the basis Π, a positive element $\alpha = \sum c_i \alpha_i$ has at least one positive coefficient among the c_i. If $\alpha \in \Phi$, we can say (by definition of Π) that in this case *all* the coefficients are positive, then the positive roots exactly form the set Φ^+. □

From now when we speak about a Φ^+, it will always be with respect to a simple system. The advantage is the fact that there are no more implicit ordering.

Lemma 51.143.
Let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a simple system and $\alpha \in \Phi^+$. Then

$$s_{\alpha_i} = \begin{cases} -\alpha_i & \text{if } \alpha = \alpha_i \\ > 0 & \text{if } \alpha \neq \alpha_i. \end{cases}$$

Proof. The first case is well know from a long time. For the second, compute

$$\begin{aligned} s_{\alpha_i}(\sum c_j\alpha_j) &= \sum_{j\neq i} c_j\alpha_j + c_i\alpha_i - 2c_i\alpha_i - \sum_{j\neq i} \frac{2c_j}{|\alpha_i|^2}(\alpha_j, \alpha_i)\alpha_i \\ &= \sum_{j\neq i} + \left(- \sum_{i\neq j} \frac{2c_j}{|\alpha_i|^2}(\alpha_j, \alpha_i) + c_i \right) \alpha_i. \end{aligned} \tag{51.320}$$

We see that between $\sum c_k\alpha_k$ and $s_{\alpha_i}(\sum c_k\alpha_k)$, there is just the coefficient of α_i which changes. Then if $\alpha \neq \alpha_i$, the positivity is conserved.

□

Proposition 51.144.
Let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a simple system. Then W is generate by the s_{α_i}'s. If $\alpha \in \Phi$, then there exists a $\alpha_i \in \Pi$ and $s \in W$ such that $s\alpha_j = \alpha$.

Proof. We denote by W' the group generate by the s_{α_i}'s; the purpose is to show that $W = W'$. We begin to show that if $\alpha > 0$, then $\alpha = s\alpha_j$ for certain $s \in W'$ and $\alpha_j \in \Pi$. For this, we write $\alpha = \sum c_j\alpha_j$ and we make an induction with respect to $\text{Level}(\alpha) = \sum c_j$. If $\text{Level}(\alpha) = 1$, then $\alpha - \alpha_j$ and $s = \text{Id}$ works. Now we suppose that it works for $\text{Level} < \text{Level}(\alpha)$. We have

$$0 < (\alpha, \alpha) = \sum c_i(\alpha, \alpha_i).$$

Since all the c_i are positive, it assures the existence of a i_0 such that $(\alpha, \alpha_{i_0}) > 0$. Then from the lemma, $\beta = s_{\alpha_{i_0}}(\alpha) > 0$ ($\alpha \neq \alpha_{i_0}$ because $\text{Level}(\alpha) > 1$). The root β can be expanded as

$$\beta = \sum_{j\neq i_0} c_j\alpha_j + \left(c_{i_0} - \sum_{j\neq i_0} \frac{c_j}{|\alpha_{i_0}|^2}(\alpha, \alpha_{i_0}) \right) \alpha_{i_0}. \tag{51.321}$$

Since $(\alpha, \alpha_{i_0}) > 0$, it implies $\text{Level}(\beta) < \text{Level}(\alpha)$ and thus $\beta = s'\alpha_j$ for a certain $s' \in W'$. So $\alpha = s_{\alpha_{i_0}}s'\alpha_j$ with $s_{\alpha_{i_0}}s' \in W'$. This conclude the induction. For $\alpha < 0$, the same result holds by writing $-\alpha = s\alpha_j$ and $\alpha = ss_{\alpha_j}\alpha_j$.

Now it remains to prove that $W' \subseteq W$. For a $\alpha \in \Phi$, we write $\alpha = s\alpha_j$ with $s \in W'$. Then

$$s_\alpha = ss_{\alpha_j}s^{-1} \in W'.$$

□

51.11.2 About group representations

Let π be a representation of a group G. The **character** of π is the function

$$\begin{aligned} \chi_\pi \colon G &\to \mathbb{C} \\ g &\mapsto \text{Tr}\big(\pi(g)\big). \end{aligned} \tag{51.322}$$

From the cyclic invariance of trace, it fulfils $\chi_\pi(gxg^{-1}) = \chi_\pi(x)$, so that the character is a central function.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. We denote by $Z_\pm$ the subgroup of G generated by $\mathfrak{n}^\pm$. The **Cartan subgroup** D of G is the maximal abelian subgroup of G which has $\mathfrak{h}$ as Lie algebra.

A **character** of an abelian group is a representation of dimension one.

Let T be a representation of G on a complex vector space V. One say that $\xi \in V$ is a **highest weight** if

— $T(z)\xi = \xi$ for every $z \in Z_+$,

— $T(g)\xi = \alpha(g)\xi$ for every $g \in D$.

The function $\alpha\colon D \to \mathbb{C}$ is the **highest weight** of the representation T.

Lemma 51.145.
The function α is a character of the group D.

Proof. The number $\alpha(gg')$ is defined by $T(gg')\xi = \alpha(gg')\xi$. Using the fact that T is a representation, one easily obtains $T(gg')\xi = \alpha(g)\alpha(g')\xi$. □

51.11.3 Modules and reducibility

As far as terminology is concerned, one can sometimes find the following definitions. A $\mathfrak{g}$-module is **simple** when the only submodules are $\mathfrak{g}$ and 0. It is **semisimple** when it is isomorphic to a direct sum of simple modules. The module is **indecomposable** if it is not isomorphic to the direct sum of two non trivial submodules.

An vector space endomorphism $a\colon V \to V$ is **semisimple** if V is semisimple as module for the associative algebra spanned by A. In this text, we will not use this terminology but the one in terms of reducibility. It is clear that $\mathfrak{g}$ is itself a $\mathfrak{g}$-module for the adjoint representation. From this point of view, a $\mathfrak{g}$-submodule is an ideal. Then a simple Lie algebra is an irreducible $\mathfrak{g}$-module and a semisimple Lie algebra is completely reducible by corollary 51.52. This explains the terminology correspondence

$$\begin{array}{ccc} \textit{simple} & \leftrightarrow & \textit{irreducible} \\ \textit{semisimple} & \leftrightarrow & \textit{completely reducible.} \end{array}$$

Theorem 51.146 (Weyl's theorem)**.**
A representation of a semisimple Lie algebra is completely reducible.

51.11.4 Weight and dual spaces

In general, when $T\colon V \to V$ is an endomorphism of the vector space V and $\lambda \in \mathbb{K}$ ($\mathbb{K}$ is the base field of V), we define

$$V_\lambda = \{v \in V \text{ tel que } (T - \lambda\mathbb{1})^n v = 0 \text{ for a } n \in \mathbb{N}\}). \tag{51.323}$$

If $V(\lambda) \neq 0$, we say that λ is a **weight** and $V(\lambda)$ is a weight space.

Let now particularize to the case where $\mathfrak{g}$ is a Lie algebra, and $\mathfrak{g}^*$ its dual space (the space of all the complex linear forms on $\mathfrak{g}$). Let ρ be a representation of $\mathfrak{g}$ on a complex vector space V (seen as a $\mathfrak{g}$-module), and $\gamma \in \mathfrak{g}^*$. For each $x \in \mathfrak{g}$, we have $\rho(x)\colon V \to V$ and $\gamma(x) \in \mathbb{C}$; then it makes sense to speak about the operator $\rho(x) - \gamma(x)\colon V \to V$ and to define

$$V_\gamma = \{v \in V \text{ tel que } \forall x \in \mathfrak{g}, \exists n \in \mathbb{N} \text{ tel que } \big(\rho(x) - \gamma(x)\big)^n v = 0\}. \tag{51.324}$$

If $V_\gamma \neq 0$, we say that γ is a **weight** for the representation ρ while V_γ is the corresponding **weight space**.

Notice that a root is a weight space for the adjoint representation, see definition 51.73. We denote by Φ the set of non empty root spaces.

Lemma 51.147.
Let End(V) *be the algebra of linear endomorphism of a vector space V. Let $x_1, \dots, x_k, y_1, \dots, y_k \in$* End$(V)$ *and*

$$e = \sum_i [x_i, y_i].$$

If e commutes with all x_i, then it is nilpotent. lem:EndV_e

A proof of this lemma can be found in [855]

Theorem 51.148.
Let $\mathfrak{g}$ be a Lie algebra of linear endomorphisms of a finite dimensional vector space V. We suppose that V is a completely reducible $\mathfrak{g}$-module and we denote the center of $\mathfrak{g}$ by $\mathcal{Z}$. Then

(1) $[\mathfrak{g},\mathfrak{g}] \cap \mathcal{Z} = 0$,

(2) $L/\mathcal{Z}$ *has a non zero abelian ideal,*

(3) any element of $\mathcal{Z}$ is a semisimple endomorphism.

¡¡ Avertissement/question au lecteur !! 51.149
The following proof seems me to be quite wrong.

Proof. Let A be the associative algebra spanned by $\mathfrak{g}$ and the identity on V. It is clear that the A-stable subspaces are exactly the $\mathfrak{g}$-stable ones. Then V is a completely reducible A-module and it has no non zero nilpotent left ideal. Indeed let B be a left ideal in A such that $BB = 0$. In this case, $B \cdot V$ is a A-submodule of V (because B is an ideal) and $V = B \cdot V \oplus W$ for a certain A-submodule W. Since $B \cdot V$ is a A-submodule,

$$B \cdot W \subset (B \cdot V) \cap W$$

(because W is stable under A) which implies $B \cdot W = 0$ and $B \cdot V = B \cdot (BV + W)0$. Consequently, $B = 0$.

Let T be the center of A; this is an ideal, so that T has no non zero nilpotent elements. To see it, consider a nilpotent element $z \in T$. Remark that $T = Az$ is a nilpotent ideal because $AzAz = Az^2A$. Now, we prove that z is a semisimple linear endomorphism of V. There exists n such that $z^n(V) = z^{n+1}(V)$. Let $q = \sum_{v\in V}$ The space $V_s = z^n(V)$ is not zero because z is not nilpotent. Let W be the set of elements of V which are annihilated by a certain power of z. Equation (46.2) makes z semisimple because V_s and W are z-stables.

By lemma 51.147, any element of $[A, A] \cap T$ is nilpotent; but we just saw that it has no non zero nilpotent elements then $[A, A] \cap T = 0$, so that

$$[\mathfrak{g},\mathfrak{g}] \cap \mathcal{Z} = 0.$$

This proves the first point.

Now we consider an ideal J such that $[J, J] \subset \mathcal{Z}$. Then $[J, J] = [J, J] \cap \mathcal{Z} = 0$. We looks at the abelian ideal $[\mathfrak{g}, J]$ of $\mathfrak{g}$. This is an ideal because $[[g, j], h] = -[[j, h], g] - [[h, g], j]$. By the lemma, the elements of $[\mathfrak{g}, J]$ are nilpotent and the associative algebra generated by $[G, J]$ is also nilpotent because $[\mathfrak{g}, J]$ is abelian.

The elements of B are polynomials with respect to elements of $[\mathfrak{g}, J]$, then $AB \subset BA + B$ because AB is made up with elements of the form $a(hj - jh)^n$ which itself is made up with elements ah^kj^l. By commutating j^l, we get

$$j^lah^k + \text{elements of } [\mathfrak{g}, J],$$

but J is an ideal and $j^l \in J$. By induction,

$$(AB)^k \subset B^kA + B^k. \tag{51.325}$$

Since B is nilpotent, AB is a nilpotent left ideal. Then $AB = 0$ which in turn implies $B = 0$. In particular $[\mathfrak{g}, J] = 0$, so that $J \subset \mathcal{Z}$. But any abelian ideal in $\mathfrak{g}/\mathcal{Z}$ is the canonical projection of an ideal J of $\mathfrak{g}$ such that $[J, J] \in \mathcal{Z}$. We conclude that $\mathfrak{g}/\mathcal{Z}$ has no non zero abelian ideal.

□

Now we are able to prove a third version of Lie's theorem:

tho:Lie_trois

Theorem 51.150 (Lie)**.**
If $\mathfrak{g}$ is a solvable ideal, then any completely reducible $\mathfrak{g}$-module is annihilated by $[\mathfrak{g},\mathfrak{g}]$.

Proof. Let V be such a $\mathfrak{g}$-module, ρ the representation of $\mathfrak{g}$ on V and $\mathcal{A} = \rho(\mathfrak{g}) \subset \mathrm{End}(V)$. By assumption, $\mathfrak{a}$ is a solvable subalgebra of $\mathrm{End}(V)$; let $\mathcal{Z}$ be the center of $\mathfrak{a}$. It is clear that $\mathfrak{a}/\mathcal{Z}$ is solvable, so that it has no non zero abelian ideal. But the fact that $\mathfrak{a}/\mathcal{Z}$ is solvable makes one of the $\mathscr{D}^k(\mathfrak{a}/\mathcal{Z})$ an abelian ideal. The conclusion is that $\mathfrak{a}/\mathcal{Z} = 0$, or $\mathfrak{a} = \mathcal{Z}$. Clearly this makes $[\mathfrak{a}, \mathfrak{a}] = 0$. □

51.11.5 Dynkin diagram

Proposition 51.151.
If α and β are simple roots, then the angle $\theta_{\alpha,\beta}$ can only take the values $90°$, $120°$, $135°$ or $150°$.

Proof. No proof. □

In order to draw the **Dynkin diagram** of a Lie algebra, one draws a circle for each simple root, and one joins the roots with 1, 2 or 3 lines, following that the value of the angle is $120°$, $135°$ or $150°$. If the roots are orthogonal (angle $90°$), they are not connected. If the length of a root is maximal, the circle is left empty. If not, it is filled.

One easily determines the number of lines between two roots by the following proposition.

PropProdNbLignes

Proposition 51.152.
If α and β are two simple roots with $(\alpha, \alpha) \leqslant (\beta, \beta)$, then

$$\frac{(\alpha,\alpha)}{(\beta,\beta)} = \begin{cases} 1 & \text{if } \theta_{\alpha,\beta} = 120° \\ 2 & \text{if } \theta_{\alpha,\beta} = 135° \\ 3 & \text{if } \theta_{\alpha,\beta} = 150°. \end{cases} \tag{51.326}$$

Proof. No proof. □

If M is a weight of a representation, its **Dynkin coefficients** are

$$M_i = \frac{2(M, \alpha_i)}{(\alpha_i, \alpha_i)}, \tag{51.327}$$

and we can compute the Dynkin coefficients from one weight to another by the simple formula

EqCofDynMmoisAlpha

$$(M - \alpha_j)_i = M_i - A_{ij}. \tag{51.328}$$

A weight is **dominant** if all its Dynkin coefficients are strictly positive.

prop:trois_poids

Proposition 51.153.
Let $\mathfrak{g}$ be a nilpotent complex Lie algebra and ρ, a representation of $\mathfrak{g}$ on a finite dimensional vector space V. Then

(1) $\forall \gamma \in \mathfrak{g}^*$, the space V_γ is a $\mathfrak{g}$-submodule of V,

(2) if γ is a weight, then there exists a nonzero vector $v \in V_\gamma$ such that $\forall x \in \mathfrak{g}$, $x \cdot v = \gamma(x)v$, prop:trois_poids:de ItemVSEZIIIviii

(3) $V = \bigoplus_\gamma V_\gamma$ where the sum is taken over the set of weight.

Proof. Since ρ is a representation,

$$\big(\rho(x) - \gamma(x)\big)\rho(y) = \rho(y)\big(\rho(x) - \gamma(x)\big) + \rho([x, y]).$$

Now let us suppose that $\big(\rho(x) - \gamma(x)\big)^m \rho(y)$ is a sum of endomorphism of the form

$$\rho((\mathrm{ad}\, x)^p y)\big(\rho(x) - \gamma(x)\big)^q$$

with $p + q = m$. We just saw that it was true for $m = 1$. Let us check for $m + 1$:

$$\begin{aligned} \rho(x)\rho((\mathrm{ad}\, x)^p y)\big(\rho(x) - \gamma(x)\big)^q &= \rho([x, (\mathrm{ad}\, x)^p y])\big(\rho(x) - \gamma(x)\big)^q \\ &\quad + \rho((\mathrm{ad}\, x)^p y)\rho(x)\big(\rho(x) - \gamma(x)\big)^q. \end{aligned} \tag{51.329}$$

Then, since $\mathfrak{g}$ is nilpotent, the space V_γ is a submodule of V because for large enough m and for all y, $\big(\rho(x)-\gamma(x)\big)^m\rho(y)v=0$ if $v\in V_\gamma$.

Any nilpotent algebra is solvable, then from Lie theorem 51.150, the restrictions of $\rho(x)$ (with $x\in\mathfrak{g}$) to irreducible submodules commute. By Schur's lemma 46.1, they are multiples of identity. But if all $\mathfrak{g}$ is the identity on an irreducible module, then the module has dimension one. In particular, *any irreducible submodule of V_γ has dimension one*[19].

Then, in the weight space V_γ, there is a v which fulfils $\rho(x)v=\lambda(x)v$ for all $x\in\mathfrak{g}$. It is rather clear that it will only works for $\lambda=\gamma$. Our conclusion is that there exists a $v\in V_\gamma$ such that $\rho(x)v=\gamma(x)v$.

Now we consider $\gamma_1,\dots,\gamma_k$, distinct weights. They are linear forms; then there exists a $x\in\mathfrak{g}$ such that $\gamma_1(x),\dots,\gamma_k(x)$ are distinct numbers. In fact, the set $\{h\in\mathfrak{h}$ tel que $\alpha_i(h)=\alpha_j(h)$ for a certain pair $(i,j)\}$ is a finite union of hyperplanes in $\mathfrak{h}$; then the complementary is non empty.

With this fact we can see that the sum $V_{\gamma_1}+\cdots+V_{\gamma_k}$ is direct. Indeed let $v\in V_{\gamma_i}\cap V_{\gamma_j}$; for the chosen $x\in\mathfrak{g}$ and for suitable m,

$$\big(\rho(x)-\gamma_i(x)\big)^m v=\big(\rho(x)-\gamma_j(x)\big)^m v=0 \tag{51.330}$$

which implies $\gamma_i(x)=\gamma_j(x)$ or $v=0$. In particular one has only a finitely many roots and we can suppose that our choice of γ_i is complete.

For $a\in\mathbb{C}$, we define V_a as the set of elements in V which are annihilated by some power of $\rho(x)-a$ with our famous x. By the first lines of the proof, V_a is a $\mathfrak{g}$-submodule of V.

For the same reasons as before[20], if $V_a\neq0$, there exists a $v\in V_a$ and a weight γ_i such that $\forall y\in\mathfrak{g}$,

$$\rho(y)v=\gamma_i(y)v.$$

But as v is annihilated by a power of $\big(\rho(x)-a\big)$, it is clear that $a=\gamma_i(x)$, and some theory of linear endomorphism[21] shows that V is the sum of the V_a's:

$$V=\sum_{i=1}^{k}V_{\gamma_i(x)}. \tag{51.331}$$

It remains to be proved that $V_{\gamma_i(x)}\subset V_{\gamma_i}$. Let $y\in\mathfrak{g}$ and

$$V_{i,a}=\{v\in V_{\gamma_i(x)}\text{ tel que }\exists n:(\rho(y)-a)^nv=0\}.$$

As usual[22] if $V_{i,a}\neq0$, there exists a $v\in V_{i,a}$ and a weight γ_j such that $\rho(z)v=\gamma_j(z)v$ for any $z\in\mathfrak{g}$. Then $a=\gamma_j(y)=\gamma_i(y)$. But $V_{\gamma_i(x)}$ being the sum of the $V_{i,a}$'s, we have $V_{\gamma_i(x)}=V_{i,\gamma_i(y)}$ for any $y\in\mathfrak{g}$. This makes $V_{\gamma_i(x)\subset V_{\gamma_i}}$.

□

From proposition 51.153(3), an element $y\in\mathfrak{g}$ can be decomposed as

eq:decomp_racine

$$y=\sum_{\beta\in\Phi}y_\beta \tag{51.332}$$

with $y_\beta\in\mathfrak{g}_\beta$.

tho:Killing_Cartan

Theorem 51.154.
Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$ and B the Killing form of $\mathfrak{g}$. Then for all x, $y\in\mathfrak{h}$,

$$B(x,y)=\sum_{\gamma\in\Phi}d_\gamma\gamma(x)\gamma(y) \tag{51.333}$$

where $g_\gamma=\dim\mathfrak{g}_\gamma$.

19. Encore que soit pas bien clair pourquoi un tel module existerait... donc l'affiramation suivante ne me semble pas trop justifiée
20. Celles que je n'ai pas bien comprises
21. théorie que je ne connais pas trop
22. et comme d'hab, l'argument que je ne saisit pas

Proof. We are seeing $\mathfrak{g}$ as a $\mathfrak{h}$-module for the adjoint representation. In particular, proposition 51.153 makes $\mathfrak{g}$ a direct sum of the $\mathfrak{h}$-submodules $\mathfrak{g}_\gamma$. Then

$$B(x,y) = \mathrm{Tr}(\mathrm{ad}\, x^2) = \sum_{\gamma\in\Phi} \mathrm{Tr}(\mathrm{ad}\, x|_\gamma^2) \tag{51.334}$$

where $\mathrm{ad}\, x|_\gamma$ means the restriction of $\mathrm{ad}\, x$ to $\mathfrak{g}_\gamma$. It is clear that $\mathrm{ad}\, x|_\gamma - \gamma(x)$ is nilpotent, then $\mathrm{ad}\, x|_\gamma^2 - \gamma(x)^2$ is also nilpotent because

$$\mathrm{ad}\, x|_\gamma^2 - \gamma(x)^2 = (\mathrm{ad}\, x|_\gamma + \gamma(x))(\mathrm{ad}\, x|_\gamma - \gamma(x))$$

and the fact that these two terms commute. The trace of a nilpotent endomorphism is zero, then $\mathrm{Tr}(\mathrm{ad}\, x|_\gamma^2 - \gamma(x)^2) = 0$ or for all $x \in \mathfrak{g}$,

$$B(x,x) = \sum_{\gamma\in\Phi} d_\gamma \gamma(x)^2. \tag{51.335}$$

eq:Bxx

on the other hand, we know that a quadratic form determines only one bilinear form. Here the form (51.335) gives

$$B(x,y) = \sum_{\gamma\in\Phi} d_\gamma \gamma(x)\gamma(y).$$

□

tho:six_Cartan

Theorem 51.155.
If α, β are roots of a semisimple Lie algebra $\mathfrak{g}$ with respect to a Cartan subalgebra $\mathfrak{h}$, then

(1) if $x_\alpha \neq 0 \in \mathfrak{g}_\alpha$ fulfils $[h, x_\alpha] = \alpha(h)x_\alpha$ for all $h \in \mathfrak{h}$, then $\forall y \in \mathfrak{g}_{-\alpha}$

$$[x_\alpha, y] = B(x_\alpha, y)h_\alpha,$$

ite:six_deux

(2) $\alpha(h_\alpha)$ is rational and positive. Moreover

$$\alpha(h_\alpha) \sum_{\gamma\in\Phi} (\gamma_\alpha - \gamma^\alpha)^2 = 4,$$

(3) $2\beta(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$, ite:six_trois

(4) the forms $0, \alpha, -\alpha$ are the only integer multiples of α which are roots, ite:six_quatre

(5) $\dim \mathfrak{g}_\alpha = 1$, ite:six_cinq

(6) any k which makes $\beta + k\alpha$ a root lie between $-\beta_\alpha$ and β^α. In other words, $\beta + k\alpha \in \Phi$ is only true with $-\beta_\alpha \leqslant k \leqslant \beta^\alpha$. ite:six_six

Proof. The fact that $y \in \mathfrak{g}_{-\alpha}$ and that $x \in \mathfrak{g}_\alpha$ make $[x,y] \in \mathfrak{g}_0 = \mathfrak{h}$. Now we consider $h \in \mathfrak{h}$ and the invariance formula (51.23). We find:

$$B(h,[x_\alpha,y]) = -B([x_\alpha,h],y) = \alpha(h)B(x_\alpha,y) = B(h,h_\alpha)B(x_\alpha,y) = B(h,B(x_\alpha,y)h_\alpha). \tag{51.336}$$

Since it is true for any $h \in \mathfrak{h}$ and B is nondegenerate on $\mathfrak{h}$ we find the first point. In order to prove (2), we consider

$$U = \bigoplus_{-\beta_\alpha \leqslant m \leqslant \beta^\alpha} \mathfrak{g}_{\beta+m\alpha}.$$

By definition of α_β and α^β, each term of the sum is a root space. If $z \in \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$, then U is stable under $\mathrm{ad}\, z$ because the terms in $\mathrm{ad}\, zU$ are of the form $[z, x_{\beta+m\alpha}] \in \mathfrak{g}_{\beta+m\alpha\pm\alpha}$. Note however that this $\mathrm{ad}\, zU$ is not *equal* to U.

Let $x_\alpha \neq 0 \in \mathfrak{g}_\alpha$. There exists a $y \in \mathfrak{g}_{-\alpha}$ such that $[x_\alpha, y] = B(x_\alpha, y)h_\alpha$ (here we use semi-simplicity). By fitting the norm of y, we can choose it in order to get $[x_\alpha, y] = h_\alpha$, so that

$$\mathrm{ad}\, h_\alpha = [\mathrm{ad}\, x_\alpha, \mathrm{ad}\, y].$$

Now we look at the restriction of $\operatorname{ad} h_\alpha$ to U:

$$\operatorname{Tr}(\operatorname{ad} h_\alpha) = \operatorname{Tr}(\operatorname{ad} x_\alpha \circ \operatorname{ad} y) - \operatorname{Tr}(\operatorname{ad} y \circ \operatorname{ad} x_\alpha) = 0. \tag{51.337}$$

Since $h_\alpha \in \mathfrak{h} = \mathfrak{g}_0$, we have $\operatorname{ad} h_\alpha \colon U \to U$, so that the annihilation of the trace of $\operatorname{ad} h_\alpha$ can be particularised to

$$\operatorname{Tr}(\operatorname{ad} h_\alpha|_U) = 0.$$

On the other hand, by definition $\operatorname{ad} h_\alpha - (\beta + m\alpha)(h_\alpha)$ is nilpotent on $\mathfrak{g}_{\beta+m\alpha}$. Then it has a vanishing trace:

$$\sum_m \operatorname{Tr}(\operatorname{ad} h_\alpha - (\beta + m\alpha)h_\alpha) = 0.$$

But we had yet seen that the term with $\operatorname{ad} h_\alpha$ is zero; then

$$\sum_{-\beta_\alpha \leqslant m \leqslant \beta^\alpha} (\beta + m\alpha)h_\alpha \dim \mathfrak{g}_{\beta+m\alpha} = 0. \tag{51.338}$$

eq:trace_U

If we suppose that $\alpha(h_\alpha) = 0$ this gives $\beta(h_\alpha) = 0$. Since this conclusion is true for any root β, we find $B(h, h_\alpha) = 0$ for any $h \in \mathfrak{h}$. In other words, $\alpha(h) = 0$ for any $h \in \mathfrak{h}$. This contradicts the assumption, so that we conclude $\alpha(h_\alpha) \neq 0$.

Let $V = \mathfrak{h} + (x_\alpha) + \sum_{m<0} \mathfrak{g}_{m\alpha}$ where (x_α) is the one dimensional space spanned by x_α. On the one hand, from the definition of x_α, $\operatorname{ad} x_\alpha \mathfrak{h} \subset (x_\alpha)$ and $\operatorname{ad} x_\alpha \mathfrak{g}_{m\alpha} \subset \mathfrak{g}_{(m+1)\alpha}$. On the other hand, $y \in \mathfrak{g}_{-\alpha}$ is defined by the relation $[x_\alpha, y] = h_\alpha$, then $\operatorname{ad} y\mathfrak{h} \subset \mathfrak{g}_{-\alpha} \subset \sum_{m<0} \mathfrak{g}_{m\alpha}$, $\operatorname{ad} y(x_\alpha) \subset \mathfrak{g}_0 = \mathfrak{h}$ and $\operatorname{ad} y \sum_{m<0} \mathfrak{g}_{m\alpha} = \sum_{m<0} \mathfrak{g}_{(m-1)\alpha}$. All this make V invariant under $\operatorname{ad} x_\alpha$ and $\operatorname{ad} y$.

Since $\operatorname{ad} h_\alpha = [\operatorname{ad} x_\alpha, \operatorname{ad} y]$, the trace of $\operatorname{ad} h_\alpha$ is zero so that the invariance of V gives

$$\operatorname{Tr}(\operatorname{ad} h_\alpha|_V) = 0.$$

By the definition of x_α particularised to $h \to h_\alpha$, we have $\operatorname{Tr}(\operatorname{ad} h_\alpha|_{(x_\alpha)}) = \alpha(h_\alpha)$. By the definition of $\mathfrak{g}_0$, for any $x \in \mathfrak{h}$ and $v \in \mathfrak{g}_0$, $\operatorname{ad} x$ is nilpotent on v. Taking h_α as x, we see that $(\operatorname{ad} h_\alpha)h$ don't contain "h-component". Then $\operatorname{Tr}(\operatorname{ad} h_\alpha|_\mathfrak{h}) = 0$. Finally the operator $(\operatorname{ad} h_\alpha - m\alpha(h_\alpha))$ is nilpotent on $\mathfrak{g}_{m\alpha}$, so that $\operatorname{Tr}(\operatorname{ad} h_\alpha|_{\mathfrak{g}_{m\alpha}}) = \operatorname{Tr}(m\alpha(\alpha)|_{\mathfrak{g}_{m\alpha}}) = m\alpha(h_\alpha) \dim \mathfrak{g}_{m\alpha}$. All this gives

$$\alpha(h_\alpha)\left(1 + \sum_{m<0} m \dim \mathfrak{g}_{m\alpha}\right) = 0. \tag{51.339}$$

As we saw that $\alpha(h_\alpha) \neq 0$, we conclude that $\dim \mathfrak{g}_{m\alpha} = 0$ for $m < -1$ and $\dim \mathfrak{g}_{-\alpha} = 1$. This proves (5).

This also prove (4) in the particular case of *integer* multiples. It is rather simple to get relations such that $0_\alpha = 1$, $0^\alpha = 1$, $\alpha_\alpha = 2$, $(-\alpha)_\alpha = 0$, and it is easy to check (3) in the cases $\beta = -\alpha, 0, \alpha$. Now we turn our attention to the case in which β is not an integer multiple of α. By (4) applied to $\alpha \to \beta + m\alpha$, we have $\dim \mathfrak{g}_{\beta+m\alpha} = 1$ whenever $-\beta_\alpha \leqslant m \leqslant \beta^\alpha$.

From equation (51.338), $\sum_{-\beta_\alpha \leqslant m \leqslant \beta^\alpha} (\beta(h_\alpha) + m\alpha(h_\alpha)) = 0$, then

$$(\beta_\alpha + \beta^\alpha + 1)\beta(h_\alpha) = (\sum_m m)\alpha(h_\alpha) = \left(\frac{\beta_\alpha(\beta_\alpha + 1)}{2} - \frac{(\beta^\alpha - 1)\beta^\alpha}{2}\right)\alpha(h_\alpha). \tag{51.340}$$

This gives (3). Now we consider the formula of theorem 51.154 in the case $x = y = h_\alpha$ and we use the fact that $B(h, h_\alpha) = \alpha(h)$ in the case $h = h_\alpha$:

$$B(h_\alpha, h_\alpha) = \alpha(h_\alpha) = \sum_{\gamma \in \Phi} \dim \mathfrak{g}_\gamma \gamma(h_\alpha)^2 = \sum_{\gamma \in \Phi} \gamma(h_\alpha)^2. \tag{51.341}$$

Since $\beta(h_\alpha) = \frac{1}{2}(\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$, we find (2). In order to prove (4), we consider $\beta = c\alpha$ for a $c \in \mathbb{C}$. By (3), $2c\alpha(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$, so that c is an half integer: $c = p/2$ with $p \in \mathbb{Z}$. If c is non zero, we can interchange α and β and see that $\alpha = c^{-1}\beta$ implies $c^{-1} = q/2$ with $q \in \mathbb{Z}$. It is clear the $pq = 4$. But we had already discussed the case of integer multiples of α, so that we can suppose

that p is odd. The only odd p such that $pq = 3$ with $q \in \mathbb{Z}$ are $p = 1, -1$, which are two excluded cases: they are $\alpha = \pm 2\beta$ which lies in the case of integer multiples.

It remains to prove (6). By definition of β^α, the form $\beta + (\beta^\alpha + 1)\alpha$ is not a root. But it remains possible that $\beta + (\beta^\alpha + 2)\alpha$ is. We suppose that $k_1, \dots, k_p$ are the p positive integers such that $\beta + k_i\alpha \in \Phi$. We pose

$$W = \bigoplus_{i=1}^{p} \mathfrak{g}_{\beta + k_i\alpha}.$$

As usual we see that W is stable under $\operatorname{ad} x_\alpha$ and $\operatorname{ad} y$ (because $k_i \geqslant \beta^\alpha + 2$). The trace of $\operatorname{ad} g_\alpha$ on W is zero, thus

$$0 = \sum_{i=1}^{p} (\beta + k_i\alpha)(h_\alpha). \tag{51.342}$$

By (3), we find

$$p(\beta_\alpha - \beta^\alpha)\alpha(h_\alpha) = 2(k_1 + \cdots + k_p) > p(\beta^\alpha + 1).$$

This is not possible because it would gives $-\beta^\alpha - \beta_\alpha > 2$. □

51.11.5.1 Strings of roots

Let α, β be two roots with respect to $\mathfrak{h}$ and suppose $\beta \neq 0$. We denote by α^β the largest integer m such that $\alpha + m\beta$ is a root and by α_β the one such that $\alpha - m\beta$ is a root. Let $x \in \mathfrak{g}_\alpha$; since the Killing form is nondegenerate, there exists a $y \in \mathfrak{g}$ such that $B(x, y) \neq 0$. Using the root space decomposition (51.332) for y and corollary 51.76, $B(x, y) = B(x, y_{-\alpha})$. Then [pg:root_ss]

$$\forall x \in \mathfrak{g}_\alpha, \exists y \in \mathfrak{g}_{-\alpha} \text{ such that } B(x, y) \neq 0.$$

In particular if α is a root, $-\alpha$ is also a root and the restriction of B to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate because $\mathfrak{h} = \mathfrak{g}_0$. So

$$\forall \mu \in \mathfrak{h}^*, \exists! h_\mu \in \mathfrak{h} \text{ such that } \forall h \in \mathfrak{g}, B(h, h_\mu) = \mu(h).$$

This is a general result about nondegenerate (here we use the semi-simplicity assumption) bilinear forms on a vector space. If $B(x, y) = B_{ij}x^i y^j$ and $a(x) = a_i x^i$, then a vector v such that $B(x, v) = a(x)$ exists, is unique and is given by coordinates $v^k = B^{ki}a_i$ where the matrix (B^{ij}) is the inverse of (B_{ij}).

We will sometimes use the following notation if α and β are roots:

$$(\alpha, \beta) = B(h_\alpha, h_\beta), \qquad |\alpha|^2 = (\alpha, \alpha).$$

By proposition 51.155, the roots come by pairs $(\alpha, -\alpha)$. For each of them, we choose $x_\alpha \in \mathfrak{g}_\alpha$. Our choice of $x_{-\alpha}$ is made as following. From discussion at page 3117 we can find a $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $B(x_{-\alpha}, x_\alpha) = 1$. Note that this choice is unambigous: if we had chosen first $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$, this construction would have given the same x_α than our starting point. Note also that $h_{-\alpha} = -h_\alpha$. These x_α fulfil $[x_\alpha, x_{-\alpha}] = h_\alpha$.

¡¡ Avertissement/question au lecteur !! 51.156
Here the notation Δ does not follow our convention of subsection 51.10.1.3.

Let Δ be the set of non zero roots. We define an antisymmetric map $c\colon \Delta \times \Delta \to \mathbb{C}$ as following. If $\alpha, \beta \in S$ are such that $\alpha + \beta \notin \Delta$, we pose $c(\alpha, \beta) = 0$. If $\alpha + \beta \in \Delta$,

$$[x_\alpha, x_\beta] = c(\alpha, \beta)x_{\alpha+\beta}. \tag{51.343}$$

It is easy to see that $c(\alpha, \beta) = -c(\beta, \alpha)$.

Proposition 51.157.
If $\alpha, \beta, \alpha + \beta \in \Delta$, then

(1)

$$c(-\alpha,\alpha+\beta)=c(\alpha+\beta,-\beta)=c(-\beta,-\alpha),$$

enuai

enuaii

(2) If $\alpha,\beta,\gamma,\delta\in\Delta$ and $\alpha+\beta+\gamma+\delta=0$ wile δ is neither $-\alpha$, nor $-\beta$ nor $-\gamma$, then

$$c(\alpha,\beta)c(\gamma,\delta)+c(\beta,\gamma)c(\alpha,\delta)+c(\gamma,\alpha)c(\beta,\delta)=0, \tag{51.344}$$

eq:enuaii

(3) if $\beta\neq\alpha\neq-\beta$, then

$$c(\alpha,\beta)+c(-\alpha,-\beta)=c(\alpha,-\beta)c(-\alpha,\beta)-B(h_\alpha,h_\beta),$$

enuaiii

enuaiv

(4) if $\alpha+\beta\neq0$ then

$$2c(\alpha,\beta)c(-\alpha,-\beta)=\beta^\alpha(1+\beta_\alpha)\alpha(h_\alpha). \tag{51.345}$$

eq:enuaiv

prop:enua

Proof. From our choice of x_α, we find that $B(x_\beta,x_{-\beta})=B(x_{-\alpha},x_\alpha)=B(x_{\alpha+\beta},x_{\alpha-\beta})=1$, but

$$\begin{aligned}B\big(c(-\alpha,\alpha+\beta)x_\beta,x_{-\beta}\big)&=B\big(x_{-\alpha},c(\alpha+\beta,-\beta)x_\alpha\big)\\&=B\big(x_{\alpha+\beta},c(-\beta,-\alpha)x_{-\alpha-\beta}\big).\end{aligned} \tag{51.346}$$

This proves (1). In order to prove (2), suppose that

$$c(\alpha,\beta)c(\gamma,\delta)=B\Big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\Big) \tag{51.347}$$

eq:amontrer

Then the Jacobi identity gives the result:

$$\begin{aligned}0&=B\Big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\Big)+B\Big(\big[[x_\beta,x_\gamma],x_\alpha\big],x_\delta\Big)+B\Big(\big[[x_\gamma,x_\alpha],x_\beta\big],x_\delta\Big)\\&=c(\alpha,\beta)c(\gamma,\delta)+c(\beta,\gamma)c(\alpha,\delta)+c(\gamma,\alpha)c(\beta,\delta),\end{aligned} \tag{51.348}$$

Here, we used the hypothesis $-\gamma\neq\delta\neq-\beta$ by supposing that (51.347) still hold after permutation of α,β,γ. Now we show the (51.347) is true. The assumptions imply $\alpha+\beta=-(\gamma+\delta)\neq0$, then

$$\begin{aligned}B\big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\big)&=B\big([x_\alpha,x_\beta],[x_\gamma,x_\delta]\big)\\&=c(\alpha,\beta)c(\gamma,\delta)B(x_{\alpha+\beta},x_{\gamma+\delta})\\&=c(\alpha,\beta)c(\gamma,\delta).\end{aligned} \tag{51.349}$$

Now we turn our attention to (3). If α and β fulfil the condition $\beta\neq\alpha\neq-\beta$, we can apply (2) on the quadruple $(\alpha,\beta,-\alpha,-\beta)$ to get $c(\alpha,\beta)c(-\alpha,-\beta)=-B\big([x_\alpha,x_\beta],[x_{-\alpha},x_{-\beta}]\big)$. If we replace β by $-\beta$ and if we make the difference between the two expressions,

$$\begin{aligned}c(\alpha,\beta)c(-\alpha,-\beta)&=-B\big([x_\alpha,x_\beta],[x_{-\alpha},x_{-\beta}]\big)+B\big([x_\alpha,x_{-\beta}],[x_{-\alpha},x_\beta]\big)\\&=B\big([x_\alpha,[x_{-\beta},x_{-\alpha}]],x_\beta\big)-B\big([x_{-\alpha},[x_\alpha,x_{-\beta}]],x_\beta\big)\\&=-B\big([x_{-\alpha},x_\alpha],[x_{-\beta},x_\beta]\big)\\&=-B(h_\alpha,h_\beta).\end{aligned} \tag{51.350}$$

In order to prove (4), we consider $\alpha+\beta\neq0$ and we pose

$$d(\alpha,\beta)=c(\alpha,\beta)c(-\alpha,-\beta)-\frac{1}{2}\beta^\alpha(1+\beta_\alpha)\alpha(h_\alpha).$$

Our aim is to prove that it is zero. We will do it by induction on β^α. First $\beta^\alpha=0$ means that $\beta+\alpha=0$, so that $c(\alpha,\beta)=0$ and $d(\alpha,\beta)=0$. Now we suppose that $\beta^\alpha>0$ and that (4) is yet checked for lower cases. Note that $\beta+\alpha\in\Delta$ and $(\beta+\alpha)+\alpha\neq0$ because -2α is not a root. Then $\beta=2\alpha$ is not possible. From the fact that $(\beta+\alpha)^\alpha=\beta^\alpha-1$, we conclude $d(\alpha,\beta+\alpha)=0$. Then

$$c(\alpha,\alpha+\beta)c(-\alpha,-\alpha-\beta)=c(\alpha,-\alpha-\beta)c(-\alpha,\alpha+\beta)-B(h_\alpha,h_{\alpha+\beta}).$$

On the other hand, (1) and the antisymmetry of c give

$$c(-\alpha,\alpha+\beta)=c(-\beta,-\alpha)=-c(-\alpha,-\beta) \tag{51.351a}$$

and

$$c(\alpha,-\alpha-\beta)=c(\beta,\alpha)=-c(\alpha,\beta) \tag{51.351b}$$

With all this

$$\begin{aligned} d(\alpha,\beta+\alpha)&=c(\alpha,\alpha+\beta)c(-\alpha,-\alpha-\beta)-\frac{1}{2}(\alpha+\beta)^\alpha(1+(\alpha+\beta)_\alpha)\alpha(h_\alpha)\\ &=c(\alpha,\beta)c(-\alpha,-\beta)-k(\alpha,\beta)\end{aligned} \tag{51.352}$$

where $k(\alpha,\beta)=B(h_\alpha,h_{\alpha+\beta})+\frac{1}{2}(\alpha+\beta)^\alpha(1+(\alpha+\beta)_\alpha)\alpha(h_\alpha)$. But $h_{\alpha+\beta}$ is definied in order to have $B(h,h_{\alpha+\beta})=(\alpha+\beta)(h)$ for any $h\in\mathfrak{h}$. Then using $2\beta(h_\alpha)=(\beta_\alpha-\beta^\alpha)\alpha(h_\alpha)$, we find $k(\alpha,\beta)=\frac{1}{2}\alpha(h_\alpha)\beta^\alpha(1+\beta_\alpha)$. □

prop:lHeR

Proposition 51.158.
Let

$$\mathfrak{h}_\mathbb{R}=\sum_{\alpha\in\Delta}\mathbb{R}h_\alpha. \tag{51.353}$$

Then:

(1) any root is real on $\mathfrak{h}_\mathbb{R}$,
(2) the Killing form is real and strictly positive definite on $\mathfrak{h}_\mathbb{R}$,
(3) $\mathfrak{h}=\mathfrak{h}_\mathbb{R}\oplus i\mathfrak{h}_\mathbb{R}$.

The last item shows that $\mathfrak{h}_\mathbb{R}$ is a real form of $\mathfrak{h}$. Remark also that $\mathfrak{h}_\mathbb{R}$ can also be written as

$$\mathfrak{h}_\mathbb{R}=\{h\in\mathfrak{h}\text{ tel que }\alpha(h)\in\mathbb{R}\,\forall\alpha\in\Phi\}.$$

Proof. Let $\beta\in\Delta$; we looks at $\beta(h_\alpha)$. From (2) of theorem 51.155, we know that $\alpha(h_\alpha)$ is real and positive, and (3) makes $\beta(h_\alpha)$ real. From the formula $B(h_\alpha,h_\beta)=\sum_{\gamma\in\Delta}\gamma(h_\alpha)\gamma(h_\beta)$, the Killing form is real and positive definite on $\mathfrak{h}_\mathbb{R}\times\mathfrak{h}_\mathbb{R}$. If $B(h,h)=0$ for a certain $h\in\mathfrak{h}_\mathbb{R}$, we find $\alpha(h)=0$ for all $\alpha\in\Delta$. Then any $x=x^\alpha x_\alpha\in\mathfrak{g}$ commutes with h because

$$[h,x]=\sum_{\alpha\in\Phi}a^\alpha(\operatorname{ad}h)x_\alpha=\sum_\alpha a^\alpha\alpha(h)=0.$$

So h is in the center of $\mathfrak{g}$ and so $h=0$ be cause $\mathfrak{g}$ is semisimple. Thus the Killing form is strictly positive definite on $\mathfrak{h}_\mathbb{R}\times\mathfrak{h}_\mathbb{R}$.

Now we are going to show that $\mathfrak{h}=\mathfrak{h}_\mathbb{R}\oplus i\mathfrak{h}_\mathbb{R}$. If $h\in\mathfrak{h}_\mathbb{R}\cap i\mathfrak{h}_\mathbb{R}$, it can be written as $h=ih'$ with $h,h'\in\mathfrak{h}_\mathbb{R}$. Then

$$0<B(h,h)=B(ih',ih')=-B(h',h')<0,$$

so that $h=0$ because B is nondegenerate. This shows that $\mathfrak{h}_\mathbb{R}\cap i\mathfrak{h}_\mathbb{R}=0$. It is clear that $\sum_{\alpha\in\Delta}\mathbb{C}h_\alpha\subset\mathfrak{h}$; thus it remains to be proved that $\mathfrak{h}\subset\sum_{\alpha\in\Delta}\mathbb{C}h_\alpha$. It it is not, we can build a linear function $\lambda\colon\mathfrak{h}\to\mathbb{C}$ which is not identically zero but which is zero on the subspace $\sum_{\alpha\in\Delta}\mathbb{C}h_\alpha$. Then there exists (only one) $h_\lambda\in\mathfrak{h}$ such that $B(h,h_\lambda)=\lambda(h)$ for every $h\in\mathfrak{h}$. In particular, $\alpha(h_\lambda)=0$ for every $\alpha\in\Delta$ because $\alpha(h_\lambda)=B(h_\alpha,h_\lambda)=\lambda(h_\alpha)$. This implies that $h_\lambda=0$, so that $\lambda\equiv0$.

□

One interest in the third point of this proposition is that we are now able to see Δ as a subset of $\mathfrak{h}_\mathbb{R}^*$ because the definition of $\alpha\in\Delta$ on $\mathfrak{h}_\mathbb{R}$ only is sufficient to define α on the whole $\mathfrak{h}$.

If $\{e_i\}$ is a basis of a vector space V, we say that $x=x^ie_i>y=y^ie_i$ if $x-y=a^ie_i$ and the first non zero a^i is positive. This is the **lexicographic order** on V. It is clear that it doesn't works on a complex vector space (because in this case we should first define $a^i>0$), but we can anyway get an order on Δ by seeing it as a subset of $\mathfrak{h}_\mathbb{R}$.

Lemma 51.159.
If $\alpha - \beta$ are simple roots with $\alpha \neq \beta$, then $\beta - \alpha$ is not a root and $B(h_\alpha, h_\beta) \leqslant 0$.

Proof. Define $\gamma = \beta - \alpha \in \Delta$ (and not Φ because $\alpha \neq \beta$). If $\gamma > 0$, the fact that $\beta = \gamma + \alpha$ contradict the simplicity of β while if $\gamma < 0$, in the same way $\alpha = \beta - \alpha$ contradict the simplicity of α.

Since $\beta - \alpha$ is not a root, $\beta_\alpha = 0$ and $\beta^\alpha \geqslant 0$ thus formula $2\beta(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$ gives

$$2B(h_\alpha, h_\beta) = \underbrace{(\beta_\alpha - \beta^\alpha)}_{\leqslant 0} B(h_\alpha, h_\alpha). \tag{51.354}$$

Now proposition 51.158 gives the result. □

Lemma 51.160.
The simple roots are linearly independent.

Proof. In the definition of a simple root, we need an order notion on Δ which is then seen as a subset of $\mathfrak{h}_\mathbb{R}$. But the roots are real thereon. Then the right notion of "independence" for the simple root is the independence with respect to *real* combinations.

If one has a combination $c^i\alpha_i = 0$ (sum over i) with at least one non zero among the c^i's by putting the negative c^i's at right, one can write

$$a^i\alpha_i = b^j\alpha_j$$

with $a^i, b^j \geqslant 0$. Let us consider $\gamma = a^i\alpha_i$ and h_γ. For every $h \in \mathfrak{h}$, we have

$$B(h, h\gamma) = \gamma(h) = a^i\alpha_i(h_\gamma).$$

but $h_\gamma = a^j h_\alpha$, then

$$B(h_\gamma, h_\gamma) = a^i a^j \alpha_i(h_{\alpha_j}) = a^i a^j B(h_{\alpha_i}, h_{\alpha_i}). \tag{51.355}$$

Since the α_i are all simple roots, the right hand side is negative, but proposition 51.158 makes the left hand side positive. Thus $\gamma = 0$. □

Theorem 51.161.
If $\{\alpha_1, \dots, \alpha_r\}$ is the set of all the simple roots, then $\dim \mathfrak{h}_\mathbb{R} = r$ and every $\beta \in \Phi$ can be decomposed as

$$\beta = \sum_{i=1}^{r} n_i\alpha_i$$

where the n_i are integers either all positive either all negative.

Proof. Let β be a non simple positive root. Then it can be decomposed as $\beta = \gamma + \delta$ with $\gamma, \delta > 0$.We can also separately decompose γ and δ and continue so until we are left with simple roots. We have to see why the process stops. Since there are only a finite number of positive root, if the process does not stop, then the decomposition of (at least) one of the positive roots γ contains γ itself. So we have a situation $\gamma = \gamma + \alpha$ for a certain positive α. This contradict the notion of order.

In particular $\mathrm{Span}_\mathbb{N}\{\alpha_i\} = \{\text{positive roots}\}$. Thus it is clear that

$$\mathrm{Span}_\mathbb{R}\{\alpha_i\} = \Phi.$$

□

tho:Phi_base

Theorem 51.162.
The Cartan algebra of a complex semisimple Lie algebra is abelian and the dual is spanned by the roots: $\mathrm{Span}\,\Phi = \mathfrak{h}^*$.

Proof. Let α be a non zero root; from the point (2) of proposition 51.153, there exists a $v \in \mathfrak{g}_\alpha$ such that for any $x \in \mathfrak{h}$, $[x, v] = \alpha(x)v$. Since $\dim \mathfrak{g}_\alpha = 1$ it is in fact true for any $v \in \mathfrak{g}_\alpha$. In particular $\forall v \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{h}$, we have $[h, x] = \alpha(h)x$.

Let $\mathfrak{n} \subset \mathfrak{h}$ be the set of elements which are annihilated by all the roots:

$$\mathfrak{n} = \{H \in \mathfrak{h} \text{ tel que } \alpha(H) = 0 \, \forall \alpha \in \Phi\}. \tag{51.356}$$

First remark that

$$[\mathfrak{g}_\alpha, \mathfrak{n}] = 0 \tag{51.357}$$

eq:GNz

because for $x \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{n} \subset \mathfrak{h}$, we have $[h, x] = \alpha(h)x = 0$. An other property of $\mathfrak{n}$ is

$$[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{n}. \tag{51.358}$$

eq:HHN

Indeed consider a root α and $x \in \mathfrak{g}_\alpha$. We have

$$\begin{aligned} -\alpha([h, h'])x &= [x, [h, h']] = [h, [h', x]] + [h', [x, h]] = \alpha(h)[h', x] + \alpha(h')[x, h] \\ &= \alpha(h)\alpha(h') - \alpha(h')\alpha(h) = 0. \end{aligned} \tag{51.359}$$

If $x \in \mathfrak{g}$ is decomposed as $x = \sum_{\alpha \in \Phi} x_\alpha$ and if $n \in \mathfrak{n}$, then

$$[x, n] = \sum_\alpha [x_\alpha, n] = \sum_\alpha \alpha(n) x_\alpha = 0.$$

In particular, $\mathfrak{n}$ is an ideal[23]. Moreover, the fact that $\mathfrak{n} \subset \mathfrak{h}$ makes $\mathfrak{n}$ a *nilpotent* ideal in the semisimple Lie algebra $\mathfrak{g}$. Then $\mathfrak{n} = 0$. Equation (51.357) makes $\mathfrak{h}$ abelian while equation (51.358) says that no element of $\mathfrak{h}$ is anihilated by all the roots. This implies that $\operatorname{Span} \Phi = \mathfrak{h}^*$. To see it more precisely, if Φ don't span a certain (dual) basis element e_i^* of $\mathfrak{h}^*$, then a basis of $\operatorname{Span} \Phi$ is at most $\{e_j\}_{j \neq i}$. Then it is clear that $\alpha(e_i) = 0$ for any root α. □

The following important result is the fact that a complex semisimple Lie algebra is determined by its root system.

Theorem 51.163.
Let $\mathfrak{g}$ and $\mathfrak{g}'$ be two semisimple complex Lie algebras; $\mathfrak{h}$ and $\mathfrak{h}'$, Cartan subalgebras. We suppose that we have a bijection $\Phi \to \Phi'$, $\alpha \to \alpha'$ which preserve the root system:

- *$\alpha' + \beta' = 0$ if and only if $\alpha + \beta = 0$,*
- *$\alpha' + \beta'$ is not a root if and only if $\alpha + \beta$ is also not a root,*
- *$(\alpha + \beta)' = \alpha' + \beta'$ whenever $\alpha + \beta$ is a root.*

Then we have a Lie algebra isomorphism $\eta \colon \mathfrak{g} \to \mathfrak{g}'$ such that $\eta(\mathfrak{h}) = \mathfrak{h}'$ and $\alpha' \circ \eta|_{\mathfrak{h}} = \alpha$.

Proof. From the assumptions, $\beta^\alpha = (\beta')^{\alpha'}$ and $\beta_\alpha = (\beta')_{\alpha'}$ and the point (2) of theorem 51.155 makes $\alpha'(h_{\alpha'}) = \alpha(h_\alpha)$. The fourth point of the same theorem then gives

$$\beta'(h\alpha') = \beta(h_\alpha). \tag{51.360}$$

eq:beta_h_beta

Now we choose a maximally linearly independent set $(\alpha_1, \dots, \alpha_R)$ of roots of $\mathfrak{g}$. Because of theorem 51.162, this is a basis of $\mathfrak{h}^*$. For notational convenience, we put $h_r = h_{\alpha_r}$ and naturally, $h'_r = h_{\alpha'_r}$. It is easy to see that the set of h_r is a basis of $\mathfrak{h}$. Indeed if $a^r h_r = 0$ (with sum over r), then $B(h, a^r, h_r) = a^r \alpha_r(h) = 0$ which implies that $a^r \alpha_r|_{\mathfrak{h}} = 0$ but it is impossible because the α_r are free in $\mathfrak{h}^*$.

$$\begin{aligned} &\{\alpha_1, \dots, \alpha_r\} \text{ is a basis of } \mathfrak{h}^*, \\ &\{h_1, \dots, h_r\} \text{ is a basis of } \mathfrak{h}. \end{aligned}$$

23. Ça me semble quand même fort de prouver que c'est le centralisateur pour dire que c'est un idéal. D'autant plus que je pourais directement dire que $\mathfrak{n}$ est centralisateur dans un semisimple et donc nulle.

Then the matrix $(A_{ij}) = \alpha_i(h_j)$ has non zero determinant. Since $\alpha'_i(h'_j) = \alpha_i(h_j)$, the set $\{\alpha'_1, \dots, \alpha'_r\}$ is free and $\{h'_1, \dots, h'_r\}$ is a basis of $\mathfrak{h}'$.

$$\{\alpha'_1, \dots, \alpha'_r\} \text{ is a basis of } \mathfrak{h}'^*,$$
$$\{h'_1, \dots, h'_r\} \text{ is a basis of } \mathfrak{h}'.$$

Then can define an isomorphism $\eta_\mathfrak{h} : \mathfrak{h} \to \mathfrak{h}'$ by $\eta_\mathfrak{h}(h_i) = h'_i$. If $x \in \mathfrak{h}$ is decomposed as $x = a^r h_r$, from equation (51.360) we have $(\alpha'_i \circ \eta_\mathfrak{h})(a^r h_r) = a^r \alpha'(h'_r) = \alpha_i(h_r)$. Then

$$\alpha'_i \circ \eta_\mathfrak{h} = \alpha_i.$$

Let $\alpha \in \Phi$; we can write $\alpha = c_i \alpha_i$ and $\alpha' = c'_i \alpha'_i$ (with a sum over i). We have

$$c_i \alpha_i(h_k) = \alpha(h_j) = \alpha'(h_j) = c'_i \alpha_i(h_j). \tag{51.361}$$

As the determinant of $(\alpha_i(h_j))$ is non zero, this implies $c_i = c'_i$, so that

$$\alpha' \circ \eta_\mathfrak{h} = \alpha \tag{51.362}$$

because $\alpha' \circ \eta_\mathfrak{h} = c'_i(\alpha'_i \circ \eta_\mathfrak{h}) = c_i \alpha_i = \alpha$. Now we "just" have to extend $\eta_\mathfrak{h}$ into a Lie algebra isomorphism $\eta : \mathfrak{g} \to \mathfrak{g}'$. As before for each $\alpha \in \Delta$ we choose $x_\alpha \in \mathfrak{g}_\alpha$ such that $B(x_\alpha, x_{-\alpha}) = -1$ and $[x_{-\alpha}, x_\alpha] = h_\alpha$. We naturally do the same for $x_{\alpha'} \in \mathfrak{g}'_{\alpha'}$. We also consider the function c as before: $[x_\alpha, x_\beta] = c(\alpha, \beta) x_{\alpha+\beta}$. Since $\mathfrak{h} = \mathfrak{g}_0$, these x_α form a basis of $\mathfrak{g} \ominus \mathfrak{h}$ and η can be defined by the date of $\eta(x_\alpha)$. We set $\eta(x_\alpha) = a_\alpha x_{\alpha'}$ (without sum).

The condition $\eta\big([x_\alpha, x_\beta] = [\eta(x_\alpha), \eta(x_\beta)]\big)$ gives

$$c(\alpha, \beta) a_{\alpha+\beta} = c(\alpha', \beta') a_\alpha a_\beta \qquad \text{if } \alpha + \beta \neq 0 \tag{51.363a}$$

and

$$a_\alpha a_{-\alpha} = 1 \qquad \forall \alpha \in \Phi. \tag{51.363b}$$

These two conditions are necessary and also sufficient. Indeed there are three cases of $[x, y]$ to check: x, $y \in \mathfrak{h}$, one of these two is out of $\mathfrak{h}$ or x, y are booth out of $\mathfrak{h}$. In the third case, using (51.363a),

$$\eta([x_\alpha, x_\beta]) = c(\alpha, \beta) a_{\alpha+\beta} x_{\alpha'+\beta'} = x(\alpha', \beta') a_\alpha a_\beta x_{\alpha'+\beta'} = a_\alpha a_\beta [x_{\alpha'}, a_{\beta'}] = [\eta(x_\alpha), \eta(x_\beta)]. \tag{51.364}$$

If x, $y \in \mathfrak{h}$, then from theorem 51.162, $\eta([x, y]) = 0 = [\eta(x), \eta(y)]$. Using the fact that $[h, x_\alpha] = \alpha(h) x_\alpha$, we find the third case:

$$\eta\big([h_\beta, x_\alpha]\big) = \eta\big(\alpha(h_\alpha) x_\alpha\big) = \eta\big(\alpha'(h_{\beta'}) x_\alpha\big) = a_\alpha [h_{\beta'}, x_{\alpha'}] = [\eta(h_\beta), \eta(x_\alpha)]. \tag{51.365}$$

Now we are going to find some $a_\alpha \in \mathbb{C}$ such that

— $a_\alpha a_{-\alpha} = 1$ for any α,
— $c(\alpha, \beta) a_{\alpha+\beta} = c(\alpha', \beta') a_\alpha a_\beta$ if $\alpha + \beta \neq 0$.

We consider the lexicagraphic order on Φ: this is the order on Φ seen as a subset of $\mathfrak{h}_\mathbb{R}$ on which we put the lexicagraphic order. For a root $\alpha > 0$, we will fix the coefficient a_α by an induction with respect to the order and put $a_{-\alpha} = a_\alpha^{-1}$. Let us consider $\rho > 0$ and suppose that a_α is already defined for $-\rho < \alpha < \rho$ in such a manner that $a_\alpha a_{-\alpha} = 1$ and $c(\alpha, \beta) a_{\alpha+\beta} = c(\alpha', \beta') a_\alpha a_\beta$ for every α, β such that α, β and $\alpha + \beta$ are stricly between $-\rho$ and ρ. We have to define a_ρ in such a way that if $a_{-\rho} = a_\rho^{-1}$, the second condition holds for every α, β such that α, β and $\alpha + \beta$ are no zero roots between $-\rho$ and ρ.

If such a pair (α, β) doesn't exist, there are no problem to put $a_\rho = a_{-\rho} = 1$. Let us suppose that such a pair exists: $\alpha + \beta = \rho$. Then $\beta^\alpha \neq 0$ and the point (3) of proposition 51.157 shows that $c(\alpha, \beta) \neq 0$; in the same way, $(\beta')^{\alpha'} = \beta^\alpha \neq 0$ implies $c(\alpha', \beta') \neq 0$. We define

$$a_\rho = c(\alpha, \beta)^{-1} c(\alpha', \beta') a_\alpha a_\beta, \tag{51.366a}$$
$$a_{-\rho} = a_\rho^{-1}. \tag{51.366b}$$

Since the value of the right hand side of (51.345) doesn't change under $\alpha \to \alpha'$ and $\beta \to \beta'$, it gives $c(\alpha,\beta)c(-\alpha,-\beta) = c(\alpha',\beta')c(-\alpha',-\beta')$ and thus

$$\begin{aligned} c(-\alpha,-\beta)a_{-\rho} &= c(-\alpha,-\beta)c(\alpha,\beta)c(\alpha',\beta')^{-1}a_{-\alpha}a_{-\beta} \\ &= c(\alpha',\beta')c(-\alpha',-\beta')c(\alpha',\beta')^{-1}a_{\alpha}a_{-\beta} \\ &= c(-\alpha',-\beta')a_{-\alpha}a_{-\beta}. \end{aligned} \tag{51.367}$$

Thus the definition (51.366) fulfils the requirements for the pair (α,β). It should be shown whether that works as well with another pair (γ,δ) such that $-\rho \leqslant \gamma,\delta \leqslant \rho$ and $\gamma+\delta=\rho$. If this second pair is really different that (α,β), then δ is neither α nor β; it is allso clear that δ is not $-\gamma$. Then formula (51.344) works with the quadruple $(-\alpha,-\beta,\gamma,\delta)$:

$$c(-\alpha,-\beta)c(\gamma,\delta) + c(-\beta,\gamma)c(-\alpha,\delta) + c(\gamma,-\alpha)c(-\beta,\delta) = 0. \tag{51.368}$$ eq:c_un

If $\alpha < 0$, the assumption $\alpha+\beta=\rho$ makes $\beta > \rho$, which is in contradiction with $-\rho \leqslant \beta \leqslant \rho$. Then $\alpha,\beta,\gamma,\delta > 0$ and moreover, the difference of any two of them is strictly between $-\rho$ and ρ. Since $\delta-\alpha = -(\gamma-\beta)$, if $\gamma-\beta$ is a root, $\delta-\alpha$ is also a root and the induction hypothesis gives eq:c_deux_un

$$c(\gamma,-\beta)a_{\gamma-\beta} = c(\gamma',-\beta')a_{\gamma}a_{-\beta}, \tag{51.369a}$$ eq:c_deux_un_a

$$c(-\alpha,\delta)a_{-\alpha+\delta} = c(-\alpha',\delta')a_{-\alpha}a_{\delta}. \tag{51.369b}$$ eq:c_deux_un_b

If we take for the convention $a_\mu = 1$ whenever μ is not a root, these relations still hold if $\gamma-\beta$ is not a root. In the same way, eq:c_deux_deux

$$c(\gamma,-\alpha)a_{\gamma-\alpha} = c(\gamma',-\alpha')a_{-\alpha}a_{\gamma}, \tag{51.370a}$$ eq:c_deux_deux_a

$$c(-\beta,\delta)a_{-\beta+\delta} = c(-\beta',\delta')a_{-\beta}a_{\delta}. \tag{51.370b}$$ eq:c_deux_deux_b

As $\delta-\alpha = -(\gamma-\beta)$, we have $a_{\delta-\alpha}a_{\gamma-\beta}=1$ and in the same way, $a_{\gamma-\alpha}a_{\delta-\beta}=1$. Taking it into account and multiplicating (51.369a) by (51.369b) and (51.370a) by (51.370a), we find: eq:c_deux_trois

$$c(-\beta,\gamma)c(-\alpha,\delta) = c(-\beta',\gamma')c(-\alpha',\delta')a_{-\alpha}a_{-\beta}a_{\gamma}a_{\delta} \tag{51.371a}$$ eq:c_deux_trois_a

$$c(\gamma,-\alpha)c(-\beta,\delta) = c(\gamma',-\alpha')c(-\beta',\delta')a_{-\alpha}a_{-\beta}a_{\gamma}a_{\delta}. \tag{51.371b}$$

We can use it to rewrite equation (51.368). After multiplication by $a_\alpha a_\beta a_{-\gamma}a_{-\delta}$,

$$c(-\alpha,-\beta)c(\gamma,\delta)a_\alpha a_\beta a_{-\gamma}a_{-\delta} + c(-\beta',\gamma')c(-\alpha',\delta') + c(\gamma',-\alpha')c(-\beta',\delta') = 0. \tag{51.372}$$

But equation (51.368) is also true for $(\alpha',\beta',\gamma',\delta')$ instead of $(\alpha,\beta,\gamma,\delta)$, so that the last two terms can be replaced by only one term to give

$$c(-\alpha,-\beta)c(\gamma,\delta)a_\alpha a_\beta a_{-\gamma}a_{-\delta} - c(-\alpha',-\beta')c(\gamma',\delta') = 0.$$

Since the pair (α,β) fulfils $c(-\alpha,-\beta)a_{-\alpha-\beta} = c(-\alpha',-\beta')a_{-\alpha}a_{-\beta}$, using $\alpha+\beta=\gamma+\delta$, we find

$$c(\gamma,\delta)a_{\gamma+\delta} = c(\gamma',\delta')a_\gamma a_\delta.$$

□

Corollary 51.164.
The elements $x_\alpha \in \mathfrak{g}_\alpha$ can be chosen in order to satisfy

- $B(x_\alpha, x_{-\alpha}) = 1$,
- $[x_\alpha, x_{-\alpha}] = h_\alpha$,
- $c(\alpha,\beta) = c(-\alpha,-\beta)$.

These vectors $x_\alpha \in \mathfrak{g}_\alpha$ are called **root vectors**.

Proof. We consider the isomorphism $\alpha \to \alpha$ from Φ to Φ; by the theorem this induces an isomorphism $\eta\colon \mathfrak{g} \to \mathfrak{g}$ given by some constants c_α:

$$\eta(x_\alpha) = c_{-\alpha} x_{-\alpha}$$

without sum on α, because of course $\eta(x_\alpha) \in \mathfrak{g}_{-\alpha}$. We choose $ax_\alpha \in \mathbb{C}$ in such a way that

$$a_\alpha^2 = -c_{-\alpha} \tag{51.373a}$$
$$a_\alpha a_{-\alpha} = 1, \tag{51.373b}$$

and then we put $y_\alpha = a_\alpha x_\alpha$. It is immediate that $B(y_\alpha, y_{-\alpha}) = 1$, thus the redefinition $x_\alpha \to y_\alpha$ doesn't change the obtained relations. Acting on y_α, the isomorphism η gives

$$\eta(y_\alpha) = a_\alpha c_{-\alpha} x_{-\alpha} = -a_{-\alpha} x_{-\alpha} = -y_{-\alpha}. \tag{51.374}$$

If $\alpha, \beta, \alpha+\beta \in \Delta$, we naturally define $c'(\alpha, \beta)$ by

$$[y_\alpha, y_\beta] = c'(\alpha, \beta) y_{\alpha,\beta}.$$

Using the fact that η is a Lie algebra automorphism of $\mathfrak{g}$ we have:

$$-c'(\alpha,\beta) y_{-(\alpha+\beta)} = \eta\big(c'(\alpha,\beta) y_{\alpha+\beta}\big) = [-y_{-\alpha}, -y_{-\beta}] = c'(-\alpha,-\beta) y_{-(\alpha+\beta)}. \tag{51.375}$$

□

From now we always our x_α in this way.

Remarque 51.165.
It is also possible to choice the x_α in such a way that

— $B(x_\alpha, x_{-\alpha}) = -1$,
— $c(\alpha, \beta) = c(-\alpha, -\beta)$.

This is the choice of the reference [855].

Here is a characterization for Cartan subalgebras of semisimple Lie algebras. This is sometimes taken as the *definition* of a Cartan subalgebra in books devoted to semisimple Lie algebras (for example in [813]).

PropCartanMaxAnel

Proposition 51.166.
A subalgebra $\mathfrak{h}$ of a semisimple Lie algebra $\mathfrak{g}$ is a Cartan subalgebra if and only if

— *$\mathfrak{h}$ is maximally abelian in $\mathfrak{g}$,*
— *the endomorphism $\operatorname{ad} h$ is semisimple*[24] *for every $h \in \mathfrak{h}$.*

Proof. Necessary condition. We know from theorem 51.162 that $\mathfrak{h}$ is abelian and from proposition 51.71 that it is maximally nilpotent. Then it is maximally abelian. On the other hand, let $h \in \mathfrak{h}$; the endomorphism $\operatorname{ad} h$ is diagonalisable with respect to the decomposition $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Delta} \mathfrak{h}_\alpha$.
Sufficient condition. Firstly it is clear that a maximal abelian subalgebra is nilpotent and the $\operatorname{ad} h_i$ are simultaneously diagonalisable for the different $h_i \in \mathfrak{h}$. Let $\{x_1, \dots, x_n\}$ be a basis of $\mathfrak{g}$ which diagonalise all the $\operatorname{ad} h_i$. In this basis, if $(\operatorname{ad} h)_{ii} = 0$ for any $h \in \mathfrak{h}$, then $x_i \in \mathfrak{h}$: if it was not, $\mathfrak{h} \cup \{x_i\}$ would be abelian.

Let $x \in \mathfrak{g}$ such that $(\operatorname{ad} h)x \in \mathfrak{h}$ for every $h \in \mathfrak{h}$. Suppose that x has a x_i-component with $x_i \notin \mathfrak{h}$. There is a $h \in \mathfrak{h}$ with $(\operatorname{ad} h)_{ii} \neq 0$. Then $(\operatorname{ad} h)x$ has a x_i-component and can't lies in $\mathfrak{h}$.

□

This characterization of Cartan subalgebras is used to prove the existence of Cartan subalgebra for any complex semisimple Lie algebra.

24. If the base field is $\mathbb{C}$, this means "diagonalizable".

51.11.6 Weyl: other results

Proposition 51.167.
Two immediate properties of the Weyl group are

(1) W is a finite group of orthogonal transformations of V,

(2) if r is an orthogonal transformation of V, the $s_{r\alpha} = rs_\alpha r^{-1}$.

Proof. First item. By definition of an abstract root system, W leaves Δ invariant; since V is spanned by V, it implies that W also leaves V invariant. From an easy computation, $(s_\alpha\varphi, s_\alpha\phi) = (\varphi, \phi)$. Since Δ is a finite set, there are only a finite number of common permutations of elements of Δ *a fortiori* W is finite.
Second item. It is easy to see that $s_{r\alpha}(r\varphi) = rs_\alpha\varphi$, then $s_{r\alpha} = r \circ s_\alpha \circ r^{-1}$. □

We introduce the **root reflexion** $s_\alpha \colon \mathfrak{h}^*_{\mathbb{R}} \to \mathfrak{h}^*_{\mathbb{R}}$ for $\alpha \in \Phi$ and $\varphi \in \mathfrak{h}^*_{\mathbb{R}}$ by

$$s_\alpha(\varphi) = \varphi - \frac{2(\varphi, \alpha)}{|\alpha|^2}\alpha. \tag{51.376}$$

Proposition 51.168.
If $\alpha \in \Phi$, then s_α leaves Φ invariant.

Proof. If α or φ is zero, then it is clear that $s_\alpha(\varphi)$ belongs to Φ. Thus we can suppose that $\alpha \in \Delta$ and proof that s_α leaves Δ invariant. For, we use the theorem 51.155 to find

$$s_\alpha\beta = \beta - \frac{2(\beta, \alpha)}{|\alpha|^2}\alpha = \beta - (\beta_\alpha - \beta^\alpha)\alpha. \tag{51.377}$$

If $\beta_\alpha - \beta^\alpha > 0$, we are in a case $\beta - n\alpha$ with $\beta_\alpha - \beta^\alpha < \beta_\alpha$, so that $s_\alpha\beta$ is a root. The case $\beta^\alpha > \beta_\alpha$ is treated in the same way. It just remains to check that if $\alpha, \beta \in \Delta$, then $s_\alpha\beta \neq 0$. The problem is to show that the equation (with a given α in Δ)

$$\beta = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}\alpha \tag{51.378}$$

eq:beta_frZ_alpha

has no solution in Δ (the indeterminate is β). The only nonzero multiples of β which are roots are $\pm\beta$, then if we set $\beta = r\alpha$, equation (51.378) gives $r = \pm\frac{1}{2}$, which is impossible.

□

Proposition 51.169.
The Weyl group permutes simply transitively the simple systems.

51.11.7 Longest element

Let $w \in W$. The **length** of w is the smallest k such that w can be written as a composition of k reflexions s_{α_i}. That is the smallest k such that

$$w = s_{\alpha_{i_1}} s_{\alpha_{i_2}} \dots s_{\alpha_{i_k}}. \tag{51.379}$$

Lemma 51.170.
If w and w' are elements of the Weyl group,

(1) $l(w) = l(w^{-1})$,

(2) $l(w) = 0$ if and only if $w = \mathrm{Id}$,

(3) $l(ww') \leqslant l(w) + l(w')$,

(4) $l(ww') \geqslant l(w) - l(w')$,

(5) $l(w) - 1 \leqslant l(ws_{\alpha_i}) \leqslant l(w) + 1$.

Let $n(w)$ be the number of positive simple roots that are send to a negative root:

$$n(w) = Card\, \Pi \cap w^{-1}(-\Pi). \tag{51.380}$$

Proposition 51.171.
Let Δ be a system of simple roots and Π the associated positive system. The following conditions on an element w of the Weyl group are equivalent:

(1) $w\Pi = \Pi$;

(2) $w\Delta = \Delta$;

(3) $l(w) = 0$;

(4) $n(w) = 0$;

(5) $w = \mathrm{Id}$.

For a proof see page 15 in [863].

Theorem 51.172.
If w is an element of the Weyl group,

$$l(w) = n(w). \tag{51.381}$$

Proof. No proof. □

51.11.8 Weyl group and representations

This subsection comes from [857].

Thoirrepllamifffmor

Theorem 51.173.
There exists an irreducible representation of highest weight Λ if and only if

$$\Lambda_\alpha = \frac{2(\Lambda, \alpha)}{(\alpha, \alpha)} \in \mathbb{N} \tag{51.382}$$

for every simple root α. Moreover, if ξ is a highest weight vector and if α is a simple root, then

$$E^k_{-\alpha}\xi \begin{cases} \neq 0 & \text{if } k \leqslant \Lambda_\alpha \\ = 0 & \text{if } k > \Lambda_\alpha. \end{cases} \tag{51.383}$$

Proof. No proof. □

ExHESKimc

Exemple 51.174.
We already studied the irreducible representations if $\mathfrak{so}(3)$ in the subsection 51.8.2. The theorem 51.173 allows to determine them in a much more synthetic way.

In the case of $\mathfrak{so}(3)$, the Cartan subalgebra is one dimensional, and one has only one simple root: $\alpha = J^*_{12}$. If $\Lambda = aJ^*_{12}$, one has $(\Lambda, \alpha) = a$, and theorem 51.173 says that Λ is highest weight of an irreducible representation if and only if $a \in \mathbb{N}/2$. △

ThoLOngestlowestrepres

Theorem 51.175.
If Λ is the highest weight of a representation and if w_0 is the longest element of the Weyl group, then $w_0\Lambda$ is the lowest weight.

¡¡ Avertissement/question à la lectrice !! 51.176
It is still not clear for me how does the proof works. Questions to be answered:

(1) existence, unicity

(2) w_0 is the longest element of the Weyl group

(3) if Λ is the highest weight, then $w_0\Lambda$ is the lowest.

51.11.9 Chevalley basis (deprecated)

See [864].

Let Φ be the finite set of roots of $\mathfrak{g}$. Then chose a positivity notion on $\mathfrak{h}^*$ and consider Φ^+, the positive subset of Φ. We also take Δ, a basis of the roots. An element of Φ^+ is a **simple root** if it cannot be written under the form of a sum of two elements of Φ^+. Every positive root is a sum of simple roots.

Let

$$\{\alpha_1, \dots, \alpha_l\} \tag{51.384}$$

be a basis of $\mathfrak{h}^*$ made of simple roots and

$$\{h_1, \dots, h_l\}, \tag{51.385}$$

the dual basis. One can choose the α_i in such a way that $\{h_1, \dots, h_l\}$ is orthogonal with respect to the Killing form[25]. One consequence of that is that

$$B(h_i, h) = \alpha_i(h) \tag{51.386}$$

EqBhihalphaih

for every $h \in \mathfrak{h}$. Indeed, h can be written, in the basis, as $h = h^j h_j$ where $h^j = B(h_j, h)$. Thus one has

$$B(h_i, h) = h^i = h^j \delta_{ij} = \alpha_i(h^j h_j) = \alpha_i(h). \tag{51.387}$$

We consider $\{\alpha_1, \dots, \alpha_m\}$, the positive roots (the roots α_1,...,α_l are some of them). One knows that $\mathfrak{g}_{\alpha_i}$ is one dimensional, so one take $e_i \in \mathfrak{g}_{\alpha_i}$ and $f_i \in \mathfrak{g}_{-\alpha_i}$ as basis of their respective spaces. If we denote by $\mathfrak{n}^+ = \text{Span}\{e_1, \dots, e_m\}$ and $\mathfrak{n}^- = \text{Span}\{f_1, \dots, f_m\}$, we have the decomposition

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+. \tag{51.388}$$

It $\{\alpha_i\}$ are the simple roots, we consider the following new basis for $\mathfrak{h}$:

$$H_{\alpha_i} = \frac{2\alpha_i^*}{(\alpha_i, \alpha_i)} \tag{51.389}$$

where α_i^* is the dual of α_i with respect to the inner product on $\mathfrak{h}^*$, this means

$$\alpha_j(\alpha_i^*) = (\alpha_i, \alpha_j). \tag{51.390}$$

Since $\mathfrak{h}$ is abelian (proposition 51.166), we have

$$[H_{\alpha_i}, H_{\alpha_j}] = 0. \tag{51.391}$$

Each root is a combination of the simple roots. If $\beta = \sum_{i=1}^{l} k_i \alpha_i$, we generalise the definition of H_{α_i} to

$$H_\beta = \frac{2\beta^*}{(\beta, \beta)} = \sum_i k_i \frac{(\alpha_i, \alpha_i)}{(\beta, \beta)} H_{\alpha_i}. \tag{51.392}$$

The element H_β is the **co-weight** associated with the weight β.

Using the inner product $(.,.)$, we have the decomposition $\beta = \sum_i (\beta, \alpha_i)\alpha_i$ of the roots. An immediate consequence is that

$$\beta(\alpha_i^*) = (\alpha_i, \beta). \tag{51.393}$$

If β is any root, we denote by β_i the result of β on H_{α_i}:

$$\beta_i = \beta(H_{\alpha_i}) = \frac{2(\alpha_i, \beta)}{(\alpha_i, \alpha_i)}. \tag{51.394}$$

EqbetaialphaiH

25. Why?

Theorem 51.177 (Chevalley basis)**.**
For each root β, one can found an eigenvector E_β of $\mathrm{ad}(H_\beta)$ *such that*

$$\begin{aligned}
[H_\beta, H_\gamma] &= 0 \\
[E_\beta, E_{-\beta}] &= H_\beta \\
[E_\beta, E_\gamma] &= \begin{cases} \pm(p+1)E_{\beta+\gamma} & \text{if } \beta+\gamma \\ 0 & \text{otherwise} \end{cases} \\
[H_\beta, E_\gamma] &= 2\frac{(\beta,\gamma)}{(\beta,\beta)}E_\gamma
\end{aligned} \tag{51.395}$$

EqChevalleyBasis

where p is the biggest integer j such that $\gamma + j\beta$ is a root. Moreover, if α_i and α_j are simple roots, the latter becomes

$$[H_{\alpha_i}, E_{\pm\alpha_j}] = \pm A_{ij} E_{\pm\alpha_j} \tag{51.396}$$

where A is the Cartan matrix.

An important point to notice is that, for each positive root α, the algebra generated by $\{H_\alpha, E_\alpha, E_{-\alpha}\}$ is $\mathfrak{sl}(2)$. This is the reason why the representation theory of $\mathfrak{g}$ reduces to the representation theory of $\mathfrak{sl}(2)$.

51.12 Iwasawa decomposition of Lie groups

In this section, we show the main steps of the Iwasawa decomposition for a semisimple Lie group. For proofs, the reader will see [865] VI.4 and [813] III,§ 3,4 and VI,§ 3. In the whole section, G denotes a semisimple group, and $\mathfrak{g}$ its real (finite dimensional) Lie algebra. The two main examples that are widely used are $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{SO}(2,n)$.

51.12.1 Cartan decomposition

If $\mathfrak{g}$ is a finite dimensional Lie algebra and X, $Y \in \mathfrak{g}$, the composition of the adjoint $\mathrm{ad}\,X \circ \mathrm{ad}\,Y : \mathfrak{g} \to \mathfrak{g}$ makes sense.

Definition 51.178.
An involutive automorphism θ on a real *semi simple Lie algebra $\mathfrak{g}$ for which the form B_θ,*

$$B_\theta(X,Y) := -B(X,\theta Y) \tag{51.397}$$

(B is the Killing form on $\mathfrak{g}$) is positive definite is a ***Cartan involution****.*

Proposition 51.179.
There exists a Cartan involution for every real semisimple Lie algebra.

¡¡ Avertissement/question à la lectrice !! 51.180
The theorem 4.1 in [813] is maybe a proof of this proposition.

See [813], theorem 4.1. Since $\theta^2 = id$, the eigenvalues of a Cartan involution are ± 1, and we can define the **Cartan decomposition** $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p} \tag{51.398}$$

into ± 1-eigenspaces of θ in such a way that $\theta = (-\mathrm{Id})|_{\mathfrak{p}} \oplus \mathrm{Id}\,|_{\mathfrak{k}}$. These eigenspaces are subject to the following commutation relations:

$$[\mathfrak{k},\mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k},\mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p},\mathfrak{p}] \subseteq \mathfrak{k}. \tag{51.399}$$

leq:comm_KP

The dimension of a maximal abelian subalgebra of $\mathfrak{p}$ is the **rank** of $\mathfrak{g}$. One can prove that it does not depend on the choices (Cartan involution and maximal abelian subalgebra). We denote by $\mathfrak{a}$ such a maximal abelian subalgebras of $\mathfrak{p}$.

Lemma 51.181.
If $\mathfrak{g}_0$ is a real semisimple Lie algebra and θ a Cartan involution, then for all $X \in \mathfrak{g}_0$,

$$(\operatorname{ad} X)^* = -\operatorname{ad}(\theta X), \tag{51.400}$$

where the star on an operator on $\mathfrak{g}$ is defined by

$$B_\theta(X, AY) = B_\theta(A^*X, Y). \tag{51.401}$$

Lemma 51.182.
The set of operators $\operatorname{ad}(\mathfrak{a})$ is an abelian algebra whose elements are self-adjoint.

Proof. We have to prove that $(\operatorname{ad} H)^* = (\operatorname{ad} H)$ and $[\operatorname{ad} H, \operatorname{ad} I] = 0$ for every H, $I \in \mathfrak{a}$. First, note that $H \in \mathfrak{a} \subset \mathfrak{p}$, thus $\theta H = -H$, and $(\operatorname{ad} H)^* = -\operatorname{ad}(\theta H) = \operatorname{ad} H$.

For the second, $\operatorname{ad} H \circ \operatorname{ad} I = \operatorname{ad}(H \circ I)$ so that $[\operatorname{ad} H, \operatorname{ad} I] = \operatorname{ad}[H, I] = 0$ because $\mathfrak{a}$ is abelian. □

51.12.2 Root space decomposition

From the lemma, the operators $\operatorname{ad}(H)$ with $H \in \mathfrak{a}$ are simultaneously diagonalisable. That means that there exists a basis $\{X_i\}$ of $\mathfrak{g}$ and linear maps $\lambda_i \colon \mathfrak{a} \to \mathbb{R}$ such that

$$\operatorname{ad}(H)X_i = \lambda_i(H)X_i.$$

For each $\lambda \in \mathfrak{a}^*$, we define

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} | (\operatorname{ad} H)X = \lambda(H)X, \forall H \in \mathfrak{a}\}. \tag{51.402}$$

Elements $0 \neq \lambda \in \mathfrak{a}^*$ such that $\mathfrak{g}_\lambda \neq 0$ are called **restricted roots** of $\mathfrak{g}$. The set of restricted roots is denoted by Σ.

prop:somme_de_G

Proposition 51.183.
The restricted root together with $\mathfrak{a}$ itself span the whole space:

$$\mathfrak{g} = \mathfrak{g}_0 \oplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda, \tag{51.403}$$

eq:somme_de_G

*This decomposition is called the **restricted root space decomposition**.*

Proof. We first prove that the sum is direct. If the sum is not so, we can find a $H^* \in \mathfrak{g}_0$ and $X_i \in \mathfrak{g}_{\lambda_i}$ ($\lambda_i \in \Sigma$) such that

$$H^* + \sum_i X_i = 0 \tag{51.404}$$

eq:2906r2

Let us consider

$$N = \{H \in \mathfrak{g}_0 | \text{ the } \lambda_i(H) \text{ are all differents and not zero}\}$$

A H which is not in N fulfils some relations as $\lambda_i(H) = \lambda_j(H)$ which are linear equations, so the complement of N is an union of hyperplanes and thus N is not empty. This allows us to consider a $H \in N$.

We have choice the X_i in $\mathfrak{g}_{\lambda_i}$, *i.e.*

$$(\operatorname{ad} A)X_i = \lambda_i(A)X_i \tag{51.405}$$

eq:2906r1

for all $A \in \mathfrak{a}$. In other words, X_i diagonalise $\operatorname{ad} A$ with eigenvalues $\lambda_i(A)$. Now, let us consider $\operatorname{ad} H$ for a $H \in N$. Since all the $\lambda_i(H)$ are different and not zero, the equation (51.405) implies that all the X_i (and H^*) are in separate eigenspaces of $\operatorname{ad} H$. Thus they are linearly independent, hence the equation (51.404) is not possible. The sum (51.403) is thus a direct sum. For the rest of the proof, see [813] theorem 4.2. □

Other properties of the root spaces are listed in the following proposition.

Proposition 51.184.
The spaces $\mathfrak{g}_{\lambda_i}$ satisfy also:

(1) $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$,

(2) $\theta\mathfrak{g}_\lambda = \mathfrak{g}_{-\lambda}$; *in particular, if λ belongs to Σ, then $-\lambda$ belongs to Σ too,*

(3) $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathcal{Z}_{\mathfrak{k}}(\mathfrak{a})$ *orthogonally.*

¡¡ Avertissement/question au lecteur !! 51.185
Il faut définir quelque part ce qu'est cet espace $\mathcal{Z}_{\mathfrak{k}}(\mathfrak{a})$

51.12.3 Positivity, convex cone and partial ordering

SubsecPosiCconePartOrder

Definition 51.186 ([865]).
*Let V be a vector space. A **positivity notion** is the data of a subset V^+ of V such that*

(1) for every nonzero $v \in V$, we have $v \in V^+$ xor $-v \in V^+$,

(2) for every v, $w \in V^+$ and every $\lambda \in \mathbb{R}^+$, the elements $v + w$ and λv are positive.

If $v \in V^+$, we say that v is **positive** and we note $v > 0$.

DefConvexCone

Definition 51.187.
*A **convex cone** in a vector space A is a subset C such that*

(1) $x \in C$ and $t \in \mathbb{R}^+$ imply $tx \in C$, enuli

(2) $x, y \in C$ implies $x + y \in C$, enulii

(3) $C \cap (-C) = \{0\}$. enuliii

To a convex cone C is attached a notion of positivity by defining $x \geqslant 0$ if and only if $x \in C$. The converse is also true: if we have a notion of positivity on V, we define the corresponding convex cone by PgConeAndPositive

$$V^+ = \{x \in V \text{ tel que } x \geqslant 0\}. \tag{51.406}$$

A **linear partial ordering relation** is a partial ordering $\leqslant$ such that

— $A \leqslant B$ implies $A + C \leqslant B + C$ for all C,

— $\lambda A \leqslant \lambda B$ for all $\lambda \in \mathbb{R}^+$.

From a positivity notion gives rise to a linear partial ordering on V by defining $x \geqslant y$ if and only if $y - x \geqslant 0$.

51.12.4 Iwasawa decomposition

Let us consider a notion of positivity on $\mathfrak{a}^*$ and denote by Σ^+ the set of positive roots. We define

$$\mathfrak{n} := \oplus_{\lambda \in \Sigma^+} \mathfrak{g}_\lambda. \tag{51.407}$$

The **Iwasawa decomposition** is given by the following theorem ([865], theorem 5.12):

Theorem 51.188.
Let G be a linear connected semisimple group and $A = \exp \mathfrak{a}$, $N = \exp \mathfrak{n}$ where $\mathfrak{a}$ and $\mathfrak{n}$ are the previously defined algebras. Then A, N and AN are simply connected subgroups of G and the multiplication map

$$\begin{aligned} \phi\colon A \times N \times K &\to G \\ (a, n, k) &\mapsto ank \end{aligned} \tag{51.408}$$

is a global analytic diffeomorphism. In particular, the Lie algebra $\mathfrak{g}$ decomposes as vector space direct sum

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{n} \oplus \mathfrak{k}. \tag{51.409}$$

*The group AN is a solvable subgroup of G which is called the **Iwasawa group**, or Iwasawa component of G.* ThoIwasawaVrai

Remarque 51.189.
It can be proved that this theorem is independent of the choices: the Cartan involution, the maximal abelian subalgebra $\mathfrak{a}$ and the notion of positivity.

Notice that A, N and K are unique up to isomorphism. Their matricial representation of course depend on choices.

This theorem from [813], chapter VI, Theorem 3.4. will be useful.

Theorem 51.190.
The Lie algebra $\mathfrak{a} \oplus \mathfrak{k}$ is solvable.

This theorem implies that the group AN is solvable.[26] Before to go into concrete situations, let us prove an useful property of the $\mathfrak{k}$ part of $\mathfrak{g}$:

Theorem 51.191.

$$\mathrm{Stab}(\mathfrak{k}) = K$$

for the adjoint action of G on $\mathfrak{k}$. tho:Stab_K

The proof of it is given by two lemmas. [866]

Lemma 51.192.
For any $k \in K$,

$$\mathrm{Ad}(k)\mathfrak{k} = \mathfrak{k},$$

lem:stab_1

and

Lemma 51.193.
If for any $L \in \mathfrak{k}$, $\mathrm{Ad}(x)L$ belongs to $\mathfrak{k}$, then $x \in K$. lem:stab_2

Proof of lemma 51.192. Let us take a $L \in \mathfrak{k}$ and define $M \in K$ $k = e^M$. We have $\mathrm{Ad}(k)L = e^{\mathrm{ad}\, M}L$. But in general, we have the relations (51.399) which give $e^{\mathrm{ad}\, M}L \in \mathfrak{k}$. Then $\mathrm{Ad}(k)\mathfrak{k} \subset \mathfrak{k}$.

In order to show that $\mathfrak{k} \subset \mathrm{Ad}(k)\mathfrak{k}$, let us consider a $L \in \mathfrak{k}$. We have to find a $N \in \mathfrak{k}$ such that $\mathrm{Ad}(k)N = L$. It is clear that $N = \mathrm{Ad}(k^{-1})L$ fulfils the conditions. □

Proof of lemma 51.193. Let us consider $X \in \mathfrak{g}$ such that $x = e^X$. We have $e^{\mathrm{ad}\, X}L \in \mathfrak{k}$ for all $L \in \mathfrak{k}$. This implies that all the terms of the expansion of $e^{\mathrm{ad}\, X}L$ are in $\mathfrak{k}$. In particular, $[X, L] \in \mathfrak{k}$ for all $L \in \mathfrak{k}$. Let us consider the Cartan decomposition of X: $X = X_k + X_p$. We need X such that

$$[X_k, L] + [X_p, L] \in \mathfrak{k}$$

for any $L \in \mathfrak{k}$. But inclusions (51.399) make $[X_p, L] \in \mathfrak{p}$. Then the X_p part of X must vanish (because $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is a direct sum). □

51.13 Other results about Cartan algebras

Lemma 51.194.
A Cartan subalgebra[27] *of a semisimple complex Lie algebra is maximally abelian.*

Proof. If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, proposition 51.70 provides $H_0 \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}_0(H_0)$; in particular $H_0 \in \mathfrak{h}$. We are going to prove that if $H_1, H_2 \in \mathfrak{h}$, then for every $Y \in \mathfrak{g}$ we have $B\big([H_1, H_2], Y\big) = 0$, so that the non degeneracy of the Killing form will conclude that $[H_1, H_2] = 0$.

Let $X \in \mathfrak{g}(H_0, \lambda)$, $H \in \mathfrak{h}$. The map $\mathrm{ad}\, X \circ \mathrm{ad}\, H$ sends $\mathfrak{g}(H_0, \mu)$ to $\mathfrak{g}(H_0, \lambda + \mu)$. If we choose a basis of $\mathfrak{g}$ made up with basis of the spaces $\mathfrak{g}(H_0, \lambda_i)$ (by the primary decomposition

26. J'esère que ce que je raconte ici n'est pas trop débile pcq j'ai pas été fouiller à fond.
27. Definition 51.64.

theorem) it is clear that $B(H,X)=\mathrm{Tr}(\mathrm{ad}\,H\circ\mathrm{ad}\,X)=0$. In particular with $H=[H_1,H_2]$ we get $B\big([H_1,H_2],X\big)=0$.

On the other hand, $\mathfrak{h}$ is solvable because it is nilpotent. Since the adjoint action provides a representation of $\mathfrak{h}$ on $\mathfrak{h}$, corollary 51.25 says that we have basis of $\mathfrak{h}$ in which all the matrices of are upper triangular. Now if A, B and C are upper triangular matrices, ABC and BAC have same elements on the diagonal;in particular they traces are the equal: $\mathrm{Tr}(ABC)=\mathrm{Tr}(BAC)$. Let us consider H_1, H_2, $H\in\mathfrak{h}$ By Jacobi, $\mathrm{ad}[H_1,H_2]=[\mathrm{ad}\,H_1,\mathrm{ad}\,H_2]$, then

$$\begin{aligned}\mathrm{Tr}(\mathrm{ad}[H_1,H_2]\,\mathrm{ad}\,H)&=\mathrm{Tr}(\mathrm{ad}\,H_1\,\mathrm{ad}\,H_2\,\mathrm{ad}\,H)-\mathrm{Tr}(\mathrm{ad}\,H_2\,\mathrm{ad}\,H_1\,\mathrm{ad}\,H)\\&=\mathrm{Tr}(\mathrm{ad}\,H_2\,\mathrm{ad}\,H_1\,\mathrm{ad}\,H)-\mathrm{Tr}(\mathrm{ad}\,H_1\,\mathrm{ad}\,H_2\,\mathrm{ad}\,H)\\&=0.\end{aligned}\tag{51.410}$$

Up to now we had seen that $B([H_1,H_2],H)=0$ and $B(H,X)=0$ if $X\in\oplus_{\lambda\neq0}\mathfrak{g}(H_0,\lambda)$. In the latter, we can consider $[H_1,H_2]$ as H. Then

$$B([H_1,H_2],Y)=0$$

for all $Y\in\mathfrak{g}$. Then $[H_1,H_2]=0$ because the Killing form is nondegenerate ($\mathfrak{g}$ is semisimple). This proves that $\mathfrak{h}$ is abelian.

Now it remains to see that $\mathfrak{h}$ is contained in no larger abelian subalgebra of $\mathfrak{g}$. For this, we naturally consider a larger abelian subalgebra $\mathfrak{h}'$ of $\mathfrak{g}$. For any $H'\in\mathfrak{h}'$ and $H\in\mathfrak{h}$, we have $[H,H']=0$. In particular $[H',H_0]=0$; the property

$$\mathfrak{h}=\mathfrak{g}(H_0,0)=\{X\in\mathfrak{g}\text{ tel que }(\mathrm{ad}\,H_0)^kX=0\text{ for a cerain }k\in\mathbb{N}\}.$$

makes $H'\in\mathfrak{h}$. □

prop:G_x_central

Proposition 51.195.
Let $\mathfrak{g}$ be a Lie algebra, $x\in\mathfrak{g}$ and

$$\mathfrak{g}^x=\{y\in\mathfrak{g}\text{ tel que }\exists n\in\mathbb{N}:(\mathrm{ad}\,x)^ny=0\}.\tag{51.411}$$

Then $\mathfrak{g}^x$ is a subalgebra of $\mathfrak{g}$ which is its own centralizer in $\mathfrak{g}$.

Proof. Since $\mathrm{ad}(x)$ is a derivation [28] of $\mathfrak{g}$

$$(\mathrm{ad}\,x)^n([u,v])=\sum_{k=0}^{n}\binom{n}{k}[(\mathrm{ad}\,x)^ku,(\mathrm{ad}\,x)^{n-k}v];$$

then $[\mathfrak{g}^x,\mathfrak{g}^x]\subset\mathfrak{g}^x$. This proves that $\mathfrak{g}^x$ is a subalgebra of $\mathfrak{g}$. Let $y\in\mathfrak{g}$ be such that $[y,\mathfrak{g}^x]\subset\mathfrak{g}^x$. Clearly $[x,y]\in\mathfrak{g}^x$ (because $x\in\mathfrak{g}^x$) then $(\mathrm{ad}\,x)^ny=(\mathrm{ad}\,x)^{n-1}[x,y]$, so that $y\in\mathfrak{g}^x$. □

Proposition 51.196.
Let $\mathfrak{g}$ be a Lie algebra and $x\in\mathfrak{g}$. Then there exists a subspace $\mathfrak{g}_x$ of $\mathfrak{g}$ such that $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$ and $[\mathfrak{g}^x,\mathfrak{g}_x]\subset\mathfrak{g}_x$. prop:G_x_G_x

Proof. We claim that the space is given by

$$\mathfrak{g}_x=(\mathrm{ad}\,x)^p\mathfrak{g}\tag{51.412}$$

where p is taken large enough to have $(\mathrm{ad}\,x)^p\mathfrak{g}=(\mathrm{ad}\,x)^{p+1}\mathfrak{g}$. The lemma and the discussion below show the correctness of the definition of $\mathfrak{g}_x$ and that $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$. It remains to be proved that $[\mathfrak{g}^x,\mathfrak{g}_x]\subset\mathfrak{g}_x$. For we will prove (by induction with respect to m) for any m that $(\mathrm{ad}\,x)^my=0$ implies $(\mathrm{ad}\,y)\mathfrak{g}_x\subset\mathfrak{g}_x$.

For $m=1$, the induction assumption becomes $[x,y]=0$ and Jacobi gives $\mathrm{ad}\,x\circ\mathrm{ad}\,y=\mathrm{ad}\,y\circ\mathrm{ad}\,x$, then $(\mathrm{ad}\,y)\mathfrak{g}_x=(\mathrm{ad}\,x)^p(\mathrm{ad}\,y)\mathfrak{g}\subset\mathfrak{g}_x$. Now we suppose that $(\mathrm{ad}\,x)^{m-1}z=0$ implies

28. Definition 51.3.

$(\operatorname{ad} z)\mathfrak{g}_x \subset \mathfrak{g}_x$ and we consider $y \in \mathfrak{g}$ such that $(\operatorname{ad} x)^m y = 0$ and $u \in \mathfrak{g}_x$. We are going to show that $(\operatorname{ad} y)u \in \mathfrak{g}_x$. Let f be the characteristic polynomial of $\operatorname{ad} x$:

$$f(t) = \det\big(\operatorname{ad} x - t\mathbb{1}\big)$$

where $\operatorname{ad} x$ and $\mathbb{1}$ are taken on $\mathfrak{g}_x$. Since $(\operatorname{ad} x)u = 0$, $f(0) \neq 0$ and by the Cayley-Hamilton theorem, $f(\operatorname{ad} x)u = 0$. Then

$$(f(\operatorname{ad} x)\operatorname{ad} y)u = (f(\operatorname{ad} x)\operatorname{ad} y - (\operatorname{ad} y)f(\operatorname{ad} x))u, \tag{51.413}$$

and, on the other hand, $\forall q \in \mathbb{N}$,

$$(\operatorname{ad} x)^q \operatorname{ad} y - \operatorname{ad} y(\operatorname{ad} x)^q = \sum_{r=0}^{q-1} (\operatorname{ad} x)^r (\operatorname{ad}[x,y])(\operatorname{ad} x)^{q-r-1}.$$

It follows that $f(\operatorname{ad} x)\operatorname{ad} y - (\operatorname{ad} y)f(\operatorname{ad} x)$ is a linear combination of terms of the form

$$(\operatorname{ad} x)^a (\operatorname{ad}[x,y])(\operatorname{ad} x)^b$$

and the induction hypothesis shows that $f(\operatorname{ad} x)(\operatorname{ad} y)u \in \mathfrak{g}_x$.

Now we consider a n such that $(\operatorname{ad} x)^n \mathfrak{g}^x = 0$; the fact that $f(0) \neq 0$ implies the existence of polynomials $g(t)$ and $h(t)$ such that $g(t)t^n + h(t)f(t) = 1$. If we decompose $(\operatorname{ad} y)u = v + w$ with respect to $\mathfrak{g} = \mathfrak{g}_x \oplus \mathfrak{g}^x$ we find

$$\begin{aligned}(\operatorname{ad} y)u &= [g(\operatorname{ad} x)(\operatorname{ad} x)^n + h(\operatorname{ad} x)f(\operatorname{ad} x)](\operatorname{ad} y)u \\ &= f(\operatorname{ad} x)(\operatorname{ad} x)^n v + h(\operatorname{ad} x)f(\operatorname{ad} x)(\operatorname{ad} y)u \in \mathfrak{g}_x.\end{aligned} \tag{51.414}$$

□

Proposition 51.197.
Let $\mathfrak{g}$ be a Lie algebra and $x \in \mathfrak{g}$ such that $\mathfrak{g}^x$ is as small as possible. Then $\mathfrak{g}^x$ is a Cartan subalgebra.

Proof. From proposition 51.195, it is sufficient to prove that $\mathfrak{g}^x$ is nilpotent. Let $y \in \mathfrak{g}^x$ and $f_y(t)$ be the characteristic polynomial of $\operatorname{ad} y$. Since it is a subalgebra, $\mathfrak{g}^x$ is stable under $\operatorname{ad} y$ and proposition 51.196 makes $\mathfrak{g}_x$ also stable under $\operatorname{ad} y$. Then $\operatorname{ad} y$ can be written under a bloc-diagonal form with respect to the decomposition $\mathfrak{g} = \mathfrak{g}_x \oplus \mathfrak{g}^x$, so that the characteristic polynomial can be factorised as

$$f_y(t) = g_y(t)h_y(t) \tag{51.415}$$

where g_y and h_y are the characteristic polynomials of the restrictions of $\operatorname{ad} y$ to $\mathfrak{g}^x$ and $\mathfrak{g}_x$. Let $(y_1, \dots, y_m)$ be a basis of $\mathfrak{g}^x$ and t^n, the greatest power of t which divide all the $g_y(t)$ with $y \in \mathfrak{g}^x$. The coefficient of t^n in $g_{c^i y_i}(t)$ is a polynomial with respect to the c^i because of the expression

$$g_{c^i y_i}(t) = \det\Big(\operatorname{ad}(c^i y_i) - t\mathbb{1}\Big).$$

Let u be this polynomial: $g_{c^i y_i}(t) = \dots + u(c^1, \dots, c^m)t^n$. By definition of n, this is not an identically zero polynomial and there are no terms with t^{n-1}. For the same reasons, we have a polynomial v such that

$$h_{c^i y_i}(0) = v(c^1, \dots, c^m). \tag{51.416}$$

We know that none of the non-zero elements in $\mathfrak{g}_x$ are annihilated by $\operatorname{ad} x$ (because of the definition of $\mathfrak{g}^x$). Then $h_x(0) \neq 0$ and v is not identically zero. With all this we can find some $c^i \in \mathbb{C}$ such that $u(c^1, \dots, c^m)v(c^1, \dots, c^m) \neq 0$. If $y = c^i y_i$, the coefficient of t^n in $f_y(t)$ is $u(c)v(c) \neq 0$, so that $f_y(t)$ is not divisible by t^{n+1}.

But in the other hand $\mathfrak{g}^x$ has minimal dimension, then $\dim \mathfrak{g}^y \geqslant m = \dim \mathfrak{g}^x$. Moreover $t^{\dim \mathfrak{g}^y}$ divide $f_y(t)$ because there is a certain power of $\operatorname{ad} y$ which has zero as eigenvalue with

multiplicity $\dim \mathfrak{g}^y$ [29]. Since $f_y(t)$ can not be divided by t^{n+1} this shows that $n+1 > \dim \mathfrak{g}^y$ and $n \geqslant \dim \mathfrak{g}^y \geqslant m$.

Now we consider y, any element of $\mathfrak{g}^x$. From the fact that t^n divide all the $g_y(t)$ and that $n \geqslant m$, we see that t^m divide $g_y(t)$. But the degree of $g_y(t)$ is $\dim \mathfrak{g}^x = m$. Finally, $g_y(t) = m$ and $\operatorname{ad} y$ is nilpotent on $\mathfrak{g}^x$ for any $y \in \mathfrak{g}^x$.

The Engel's theorem 51.37 makes $\mathfrak{g}^x$ nilpotent. □

The following holds for complex or real Lie algebras and comes from [849] see also [848]. We denote by $\mathbb{K}$ the base field of $\mathfrak{g}$, i.e. $\mathbb{R}$ or $\mathbb{C}$. For $X \in \mathfrak{g}$ and $\lambda \in \mathbb{K}$ we define

$$\mathfrak{g}(X,\lambda) = \{Y \in \mathfrak{g} \text{ tel que } (\operatorname{ad} X - \lambda \mathbb{1})^n Y = 0 \text{ for a certain } n \in \mathbb{N}\}. \tag{51.417}$$

A first useful result is given in

Lemma 51.198.
If $Z \in \mathfrak{g}$, then

$$[\mathfrak{g}(Z,\lambda), \mathfrak{g}(Z,\mu)] \subset \mathfrak{g}(Z, \lambda+\mu),$$

in particular $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. lem:lambda_mu_plus

Proof. We consider $X_\lambda \in \mathfrak{g}(Z,\lambda)$ and $X_\mu \in \mathfrak{g}(Z,\mu)$. We have

$$\begin{aligned}\big(\operatorname{ad} Z - (\lambda+\mu)I\big)[X_\lambda, X_\mu] = {} & [(\operatorname{ad} Z - \lambda I)X_\lambda, X_\mu] \\ & + [X_\lambda, (\operatorname{ad} Z - \mu I)X_\mu].\end{aligned} \tag{51.418}$$

By induction,

$$\big(\operatorname{ad} Z - (\lambda+\mu)\mathbb{1}\big)^n [X_\lambda, X_\mu] = \sum_{i=0}^{\infty} \binom{n}{i} [(\operatorname{ad} Z - \lambda I)^i X_\lambda, (\operatorname{ad} Z - \mu I)^{n-i} X\mu]. \tag{51.419}$$

It will become zero for large enough n. □

An element $X \in \mathfrak{g}$ is **regular** if $\dim \mathfrak{g}(X,0)$ is minimum [30]. This minimum is the **rank** of $\mathfrak{g}$.

Proposition 51.199.
If $X \in \mathfrak{g}$ is a regular element then the algebra

$$\mathfrak{h} = \mathfrak{g}(X,0) = \{Y \in \mathfrak{g} \textit{ tel que } (\operatorname{ad} X)^n Y = 0 \textit{ for some } n \in \mathbb{N}\} \tag{51.420}$$

is nilpotent.

Proof. We have to show that for any $H \in \mathfrak{h}$, the endomorphism $\operatorname{ad} H$ of $\mathfrak{h}$ is nilpotent. Consider the characteristic polynomial of $\operatorname{ad} X$

$$p(t) = \det(\operatorname{ad} X - t\mathbb{1}) = t^r q(t)$$

where t^r is the maximal factorization of t; in other words, $q(t)$ is not divisible by t and $r = \dim \mathfrak{h}$. In particular

$$\mathfrak{h} = \{Y \in \mathfrak{g} \text{ tel que } (\operatorname{ad} X)^r Y = 0\}. \tag{51.421}$$

Let

$$\mathfrak{k} = \{Y \in \mathfrak{g} \text{ tel que } q(\operatorname{ad} X)Y = 0\} \tag{51.422}$$

From the Cayley-Hamilton theorem (46.2), $p(\operatorname{ad} X) = 0$, then $(\operatorname{ad} X)^r q(\operatorname{ad} X) = 0$ and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}$. Moreover $\mathfrak{h}$ and $\mathfrak{k}$ are $\operatorname{ad} X$-invariants: $(\operatorname{ad} X)\mathfrak{h} \subseteq \mathfrak{h}$ and $(\operatorname{ad} X)\mathfrak{k} \subseteq \mathfrak{k}$.

29. This is not a good reason.
30. Anglais ?

Every weight of $\operatorname{ad} X$ are in $\mathbb{C}$. As we know that $\mathfrak{h}$ is Cartan in $\mathfrak{g}$ if and only if $\mathfrak{h}^{\mathbb{C}}$ is Cartan in $\mathfrak{g}^{\mathbb{C}}$, we can suppose that $\mathfrak{g}$ is a complex algebra by considering $\mathfrak{g}^{\mathbb{C}}$ if $\mathfrak{g}$ is real. So all the weight are in the base field and we can define

$$\mathfrak{k} = \sum_{\lambda \in \Delta} \mathfrak{g}(X, \lambda).$$

where Δ is the set of all the non zero weight of $\operatorname{ad} X$. A property [31] of the weight space is that

$$\mathfrak{g} = \mathfrak{g}(X, \lambda_1) \oplus \ldots \oplus \mathfrak{g}(X, \lambda_m)$$

if the λ_i's are the weight of $\operatorname{ad} X$. Now we prove that $\sum_{\lambda \neq 0} \mathfrak{g}(X, \lambda) = \mathfrak{k}$. First consider a $Y \in \mathfrak{g}(X, \lambda)$ which can be decomposed as $Y = H + K$ with $H \in \mathfrak{h}$ and $K \in \mathfrak{k}$. Then $(\operatorname{ad} X - \lambda \mathbb{1})^n (H + K) = (\operatorname{ad} X - \lambda \mathbb{1})^n H + (\operatorname{ad} X - \lambda \mathbb{1})^n K$ where the first term is not zero (because $H \in \mathfrak{h}$) and lies in $\mathfrak{h}$ while the second term lies in $\mathfrak{k}$. Then the sum can be zero only if $H = 0$.

□

Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $H \in \mathfrak{g}$ and $0 = \lambda_0, \lambda_1, \ldots, \lambda_r$, the eigenvalues of $\operatorname{ad} H$. For any $\lambda \in \mathbb{C}$, one can consider

$$\mathfrak{g}(H, \lambda) = \{X \in \mathfrak{g} \text{ tel que } (\operatorname{ad} H - \lambda I)^k X = 0\}. \tag{51.423}$$

From the Jordan decomposition, $g(H, \lambda) = 0$ except if λ is one of the λ_i, and

$$\mathfrak{g} = \bigoplus_{i=0}^{r} g(H, \lambda_i). \tag{51.424}$$ eq:g_sum_g_H

An element $H \in \mathfrak{g}$ is **regular** if

$$\dim g(H, 0) = \min_{X \in \mathfrak{g}} \dim g(X, 0).$$

Let H_0 be a regular element and $\mathfrak{h} = \mathfrak{g}(H_0, 0)$.

Lemma 51.200.
The algebra $\mathfrak{h} = \mathfrak{g}(H_0, 0)$ is nilpotent

Proof. Let $0 = \lambda_0, \lambda_1, \ldots, \lambda_r$ be the eigenvalues of $\operatorname{ad} H_0$ and

$$\mathfrak{g}' = \sum_{i=1}^{r} \mathfrak{g}(H_0, \lambda_i)$$

which is a subspace of $\mathfrak{g}$. From the lemma,

$$[\mathfrak{g}(H_0, 0), \mathfrak{g}(H_0, \lambda_i)] \subset \mathfrak{g}(H_0, \lambda_i) \subset \mathfrak{g}'.$$

For each $H \in \mathfrak{h}$, we denote H', the restriction of $\operatorname{ad} H$ to $\mathfrak{g}'$ and $d(H) = \det H'$. The function $H \to d(H)$ is a polynomial on $\mathfrak{h}$ in the sense of the coordinates on $\mathfrak{h}$ as vector space. If H_0' has a zero eigenvalue we would have $\operatorname{ad}(H_0)X = 0$ for some $X \in \mathfrak{g}'$. In this case $[H_0, X] = 0$, but $X \in \mathfrak{g}(H_0, \lambda_i)$, then for a certain k, $(\operatorname{ad} H_0 - \lambda_i)^k X = 0$, so that $\lambda_i X = 0$. Since $\mathfrak{g}$ is defined by excluding λ_0, $X = 0$. Thus H_0' has only non zero eigenvalues and $d(H_0) \neq 0$.

We know that a polynomial which is zero on an open set is identically zero; then on any open set of $\mathfrak{h}$, d has a non zero value somewhere. In particular,

$$S = \{H \in \mathfrak{h} \text{ tel que } d(H) \neq 0\}$$

is dense in $\mathfrak{h}$. We consider a $H \in S$. The endomorphism H' has only non zero eigenvalues, so that $\mathfrak{g}(H, 0) \subset \mathfrak{h}$ from lemma 51.198; but H_0 is regular, then $\mathfrak{g}(H, 0) \subset \mathfrak{h}$. Thus the restriction of $\operatorname{ad} H$ to $\mathfrak{h}$ is nilpotent because it is nilpotent on $\mathfrak{g}(H, 0)$ [32].

If $l = \dim \mathfrak{h}$, then $(\operatorname{ad}_{\mathfrak{h}} H)^l = 0$ because $\operatorname{ad}_{\mathfrak{h}} H$ is nilpotent. By continuity, this equation is true for any $H \in \mathfrak{h}$ from the density of S in $\mathfrak{h}$. Then $\mathfrak{h}$ is nilpotent.

□

31. Que je dois encore faire, cf Sagle
32. Ce paragraphe n'est pas vraiment clair...

Here is an alternative proof (that I do not really understand) for theorem 51.74.

Theorem 51.201.
Let $\mathfrak{g}$ be a complex Lie algebra with Cartan subalgebra $\mathfrak{h}$. Then $\mathfrak{g}_0 = \mathfrak{h}$.

Proof. Since $\mathfrak{h}$ is Cartan, it is nilpotent. So $\mathfrak{h} \subset \mathfrak{g}_0$. If $v \in \mathfrak{g}_0$, there exists a n such that for any $z \in \mathfrak{h}$, $(\operatorname{ad} z)^n v = 0$. The fact that $\mathfrak{h}$ is nilpotent makes $(\operatorname{ad} z_n) \circ \ldots \circ (\operatorname{ad} z_1)v = 0$ for any $z \in \mathfrak{g}_0$ and for all $z_1, \ldots, z_n \in \mathfrak{h}$. If we write $(\operatorname{ad} z_1)v$ with $v \in \mathfrak{g}_0 \backslash \mathfrak{h}$, we can always choose z_1 in order the result to *not* be $\mathfrak{h}$. Next we can choose $z_2 \in \mathfrak{h}$ such that $(\operatorname{ad} z_2) \circ (\operatorname{ad} z_1)v$ is also not in $\mathfrak{h}$ and so on... Since $\mathfrak{g}_0$ is nilpotent, we always finish on zero. If n is the maximum of adjoint that we can take before to fall into zero; we have

$$[h, (\operatorname{ad} z_{n-1}) \circ (\operatorname{ad} z_1)v] = 0$$

for all $h \in \mathfrak{h}$ and with a good choice of z_i, it contradicts the fact that $\mathfrak{h}$ is Cartan. □

51.14 Representations

References for Lie algebras and their modules are [858, 848, 867, 861, 868, 862, 864].

Since $\mathfrak{h}$ is abelian, the operators H_{α_j} $(j = 1, \ldots, l)$ are simultaneously diagonalisable. In that basis of the representation space W, the basis vectors are denoted by $|u_\Lambda\rangle$ and have the property

$$H_{\alpha_i}|u_\Lambda\rangle = \Lambda(H_{\alpha_i})|u_\Lambda\rangle, \tag{51.425}$$

and, as notation, we note $\Lambda_i = \Lambda(H_{\alpha_i})$. The root Λ is a **weight** of the vector $|u_\Lambda\rangle$. The vector $E_\beta|u_\Lambda\rangle$ is of weight $\beta + \Lambda$, indeed,

$$H_{\alpha_i}E_\beta|u_\Lambda\rangle = \big([H_{\alpha_i}, E_\beta] + E_\beta H_{\alpha_i}\big)|u_\Lambda\rangle = \left(\frac{2(\alpha_i, \beta)}{(\alpha_i, \alpha_i)} + \Lambda_i\right) E_\beta|u_\Lambda\rangle. \tag{51.426}$$

Thus the eigenvalue of $E_\beta|u_\Lambda\rangle$ for H_{α_i} is, according to the relation, (51.394), $\beta(H_{\alpha_i}) + \Lambda(H_{\alpha_i})$.

We suppose that the roots α_i are given in increasing order:

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_l, \tag{51.427}$$

and one says that a weight is **positive** if its first non vanishing component is positive. Then one choose a basis of W

$$|u_{\Lambda^{(1)}}\rangle, \ldots, |u_{\Lambda^{(N)}}\rangle \tag{51.428}$$

of weight vectors. One say that this basis is **canonical** if

$$\Lambda^{(1)} \geqslant \ldots \geqslant \Lambda^{(N)}. \tag{51.429}$$

Theorem 51.202.
A vector if weight Λ which is a combination of vectors of weight $\Lambda^{(k)}$ all different of Λ vanishes.

Proof. No proof. □

A consequence of that theorem is that, if W is a representation of dimension N of $\mathfrak{g}$, there are at most N different weights. When several vectors have the same weight, the number of linearly independent such vectors is the **multiplicity** of the weight. A weight who has only one weight vector is **simple**.

Proposition 51.203.
The weights Λ and $\Lambda - 2\alpha(\Lambda, \alpha)/(\alpha, \alpha)$ have the same multiplicity for every root α.

Theorem 51.204.
Two representation are equivalent when they have the same highest weight.

PropMakalMmoinsinten

Proposition 51.205.
For any weight M and root α,

$$\frac{2(M,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}, \tag{51.430}$$

and

$$M - \frac{2(M,\alpha)}{(\alpha,\alpha)}\alpha \tag{51.431}$$

is a weight.

Notice, in particular, that for every weight M, the root $-M$ is also a weight.

51.14.1 List of the weights of a representation

We consider a representation of highest weight Λ. For each weight M, we define

$$\delta(M) = 2\sum_{\alpha_i \in \Pi} M_{\alpha_i} \tag{51.432}$$

where, as usual, $M_\alpha = 2(M,\alpha)/(\alpha,\alpha)$. For any root α, we define

$$\gamma(\alpha) = \frac{1}{2}\big(\delta(\Lambda) - \delta(\alpha)\big). \tag{51.433}$$

Proposition 51.205 shows in particular that $\gamma(\alpha)$ is an integer.

PropgammasubMLLweight

Proposition 51.206.
When M is a weight, $\gamma(M)$ is the number of simple roots that have to be subtracted from the highest weight Λ in order to get M.

Proof. No proof. □

Let us consider the sets

$$\Delta_\phi^k = \{M \text{ tel que } \gamma(M) = k\}. \tag{51.434}$$

That set is the **layer** of order k. Of course, there exists a $T(\phi)$ such that

$$\Delta_\phi = \Delta_\phi^0 \cup \Delta_\phi^1 \cup \ldots \cup \Delta_\phi^{T(\phi)}. \tag{51.435}$$

That $T(\phi)$ is the **height** of the representation ϕ. If Λ is the highest weight and Λ' is the lowest weight, then we have $\gamma(\Lambda) = 0$ and $\gamma(\Lambda') = T(\phi)$.

A corollary of proposition 51.206 is that, if $M \in \Delta_\phi^r$ and if α is a simple root, then $M+\alpha \in \Delta_\phi^{r-1}$, and $M - \alpha \in \Delta_\phi^{r+1}$.

Let us denote by $S_k(\phi)$ the multiplicity of the layer of order k; we have

$$S_0 + S_1 + \cdots + S_T = N, \tag{51.436}$$

where N is the dimension of the representation ϕ. The number

$$III(\phi) = \max S_k(\phi) \tag{51.437}$$

is the **width** of the representation.

Lemma 51.207.
If Λ is the highest weight and Λ' is the lowest weight, then $\delta(\Lambda) + \delta(\Lambda') = 0$.

Proof. No proof. □

From that lemma and the definition of $\gamma(M)$, we deduce that $\delta(\Lambda) - \delta(\Lambda') = 2\gamma(\Lambda') = T(\phi)$, so that $\delta(\Lambda) = T(\phi)$ and

$$\delta(M) = T(\phi) - 2\gamma(M). \tag{51.438}$$

In particular, $\delta(M)$ has a fixed parity for a given representation ϕ. It is the **parity** (even or odd) of the representation.

ThoLLralphatablefo

Theorem 51.208.
If Λ is the highest weight of the irreducible representation ϕ, then

$$T(\phi) = \sum_{\alpha_i \in \Pi} r_{\alpha_i} \Lambda_\alpha \tag{51.439}$$

where the coefficients r_{α_i} only depend on the algebra, and in particular not on the representation.

Proof. No proof. □

The coefficients r_{α_i} are known for all the simple Lie algebra, see for example page 105 of [861].

51.14.1.1 Finding all the weights of a representation

The following can be found in [862, 861].

Theorem 51.209.
If Δ_ϕ is the weight system of the irreducible representation ϕ, then

$$S_k = S_{T-k} \tag{51.440}$$

and

$$S_r \geqslant S_{r-1} \geqslant \ldots \geqslant S_2 \geqslant S_1 \tag{51.441}$$

where $r = \frac{T}{2} + 1$.

The theorem says that when $T(\phi)$ is even (let us say $T(\phi) = 2r$), then $III(\phi) = S_r(\phi)$ and when $T(\phi)$ is odd (let us say $T(\phi) = 2r + 1$), then

$$III(\phi) = S_r(\phi) = S_{r+1}(\phi). \tag{51.442}$$

Let α be a root. The **α-series** trough the weight M is the sequence of weights

$$M - r\alpha, \ldots, M + q\alpha \tag{51.443}$$

such that $M - (r+1)\alpha$ and $M + (q+1)\alpha$ do not belong to Δ_ϕ.

PropweightCondprstring

Proposition 51.210.
Let M be a weight of the representation ϕ and α, any root of $\mathfrak{g}$. If the α-series trough M begins at $M - r\alpha$ and ends at $M + q\alpha$, then

$$\frac{2(M,\alpha)}{(\alpha,\alpha)} = r - q, \tag{51.444}$$

or, more compactly, $M_\alpha + q = r$.

Notice that, in that proposition, q and r are well defined functions of M and α.

We are now able to determine all the weights of the representation ϕ. Let us suppose that we already know all the layers $\Delta_\phi^0, \ldots, \Delta_\phi^{r-1}$. We are going to determine the weights in the layer Δ_ϕ^r.

An element of Δ_ϕ^r has the form $M - \alpha$ with $M \in \Delta_\phi^{r-1}$ and α, a root. Thus, in order to determine Δ_ϕ^r, we have to test if $M - \alpha$ is a weight for each choice of $M \in \Delta_\phi^{r-1}$ and $\alpha \in \Pi$. Using proposition 51.210, if[33]

$$M_\alpha + q \geqslant 1, \tag{51.445}$$

33. At page 104 of [861], that condition is (I think) wrongly written $M_\alpha + q \geqslant 0$; that mistake is repeated in the example of page 106.

then $M-\alpha\in\Delta_\phi$. The number $M_\alpha-q(M,\alpha)$ is the **lucky number** of the root $M-\alpha$. The root is a weight if its lucky number is bigger or equal to 1. Notice that $q(M,\alpha)$ depends on the representation we are looking at.

Since $M+k\alpha\in\Delta_\phi^{r-k}$, the value of q is known when one knows the "lower" layers. We are thus able to determine, by induction, all the layers from Δ_ϕ^0 which only contains the highest weight. For this one, by definition, we always have $q=0$.

The Dynkin coefficients of one weights can be more easily computed using the following formula, which is a direct consequence of definition of the Cartan matrix:

$$(M-\alpha_j)_i=M_i-A_{ji}. \tag{51.446}$$ EqCoefDynkMalpha

As example, let us determine the weights of the representation of $\mathfrak{su}(3)$. The algebra $\mathfrak{su}(3)$ has two simple roots α and β whose inner products are $(\alpha,\alpha)=(\beta,\beta)=1$ and $(\alpha,\beta)=-1/2$. The highest weight of $\phi=$ is $\Lambda=(\alpha+2\beta)/3$.

We first test if $\Lambda-\alpha$ is a weight. Easy computations show that $\Lambda_\alpha=0$ wile $q=0$; thus $\Lambda-\alpha$ is not a weight. The same kind of computations show that $\Lambda_\beta=1$, so that $\Lambda_\beta=q(\Lambda,\beta)=1$. That shows that $\Delta_\phi^1=\{\Lambda-\alpha\}$.

Let now $M=\Lambda-\beta=(\alpha-\beta)/3$. Since $M+\alpha\notin\Delta_\phi$, we have $q(M,\alpha)=0$. On the other hand, $M_\alpha=1$, so that $M-\alpha\in\Delta_\phi^2$. The last one to have to be tested is $M-\beta$. Since $M+\beta=\Lambda$, we have $q(M,\beta)=1$, but $M_\beta=-1$. Thus $M_\beta+q(M,\beta)=0$ and $M-\beta$ is not a weight.

We can obviously continue in that way up to find $\Delta_\phi^r=0$, but there is an escape to be more rapid. Indeed, using theorem 51.208 with coefficients r_α that can be found in tables (for example in [861]), we find

$$T(\phi)=2\Lambda_\alpha+3\Lambda_\beta=2, \tag{51.447}$$

thus we immediately know that Δ_ϕ^3 does not exist.

On the other hand, one knows the width $III(\phi)=\max S_k(\phi)$ because (since $T(\phi)=2r$, with $r=1$), we have $III(\phi)=S_1(\phi)$. Thus, once $\Delta^1(\phi)$ is determined, we know that the next ones will never have more elements.

In the example, when we know that $M-\alpha$ is a weight, we do not have to test $M-\beta$. LeTravail

51.14.2 Tensor product of representations

51.14.2.1 Tensor and weight

Let ϕ and ϕ' be representations of $\mathfrak{g}$ on the vector spaces R and R' of dimensions n and m. If $A\in\mathbb{M}_n(R)$ and $B\in\mathbb{M}_m(R')$, the **tensor product**, also know as the **Kronecker product** of A and B is the matrix $A\otimes B\in\mathbb{M}_{mn}(R\otimes R')$ whose elements are given by

$$C_{ik,jl}=A_{ij}B_{kl}. \tag{51.448}$$

The principal properties of that product are

$$(A_1A_2)\otimes(B_1B_2)=(A_1\otimes B_1)(A_2\otimes B_2) \tag{51.449a}$$
$$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1} \tag{51.449b}$$
$$\mathbb{1}_R\otimes\mathbb{1}_{R'}=\mathbb{1}_{R\otimes R'} \tag{51.449c}$$

If φ_1 and φ_2 are two representations of a group G, the **tensor product** is defined by

$$(\varphi_1\otimes\varphi_2)(g)=\varphi_1(g)\otimes\varphi_2(g). \tag{51.450}$$

If ϕ and ϕ' are two representations of a Lie algebra $\mathfrak{g}$, the **tensor product** representation is defined by

$$(\phi\otimes\phi')(X)(v\otimes v')=\big(\phi(X)v\big)\otimes v'+v\otimes\big(\phi'(X)v'\big). \tag{51.451}$$

If $\{\phi_k\}$ are the irreducible representations, a natural question that arise is to determine the coefficients Γ which decompose $\phi \otimes \phi'$ into irreducible representations:

$$\phi \otimes \phi' = \sum_k \Gamma_k(\phi, \phi')\phi_k \tag{51.452}$$

Let W and W' be the representation spaces and consider the following decompositions in weight spaces:

$$W = \bigoplus_{\Lambda \in \Delta_1} W_\Lambda, \qquad W' = \bigoplus_{\Lambda \in \Delta_2} W'_\Lambda. \tag{51.453}$$

By definition,

$$(W \otimes W')_\alpha = \{v \otimes v' \text{ tel que } (\phi \otimes \phi')(h)(v \otimes v') = \alpha(h)(v \otimes v')\}. \tag{51.454}$$

If $\big(\phi(h)v\big) \otimes v' + v \otimes \big(\phi'(h)v'\big)$ is a multiple of $v \otimes v'$, one requires that

$$\phi(h)v = \alpha_1(h)v, \tag{51.455a}$$
$$\phi'(h)v = \alpha_2(h)v' \tag{51.455b}$$

for the weights α_1 and α_2 of ϕ and ϕ'. Thus we have

$$(W \otimes W')_{\alpha_1+\alpha_2} = W_{\alpha_1} \otimes W_{\alpha_2}. \tag{51.456}$$

We have in particular that the simple root system $\Delta_{\phi\otimes\phi'}$ of the representation $\phi \otimes \phi'$ is given by

EqDeldelDElphitens

$$\Delta_{\phi\otimes\phi'} = \Delta_\phi + \Delta_{\phi'}. \tag{51.457}$$

What we proved is[34]

Propphihwrepplullllam

Proposition 51.211.
If ϕ is a representation of highest weight Λ and ϕ' is a representation of highest weight Λ', then $\phi \otimes \phi'$ is a representation of height weight $\Lambda + \Lambda'$.

If, moreover, ϕ and ϕ' are irreducible, then $\phi \otimes \phi'$ is irreducible.

An irreducible representation that cannot be written under the form of a tensor product of irreducible representations is a **basic representation**.

Lemma 51.212.
A representation is basic if and only if its highest weight Λ is such that the Λ_{α_i} are all zero but one which is 1.

The basic representations of $\mathfrak{so}(10)$ are given by the Dynkin diagrams of figure 51.1. All the irreducible representations are obtained by tensor products of the basic ones. An **elementary** is a basic representation which has his "1" on a terminal point of the Dynkin diagram.

51.14.2.2 Decomposition of tensor products of representations

Proposition 51.211 allows us to decompose a tensor product of representations into irreducible representations. Let us do it on a simple example in $\mathfrak{su}(3)$. We consider the representations $\phi =$ and $\phi' =$. The first representation has weights

$$\Delta_\phi = \left\{ \frac{\alpha + 2\beta}{3}, \frac{\alpha - \beta}{3}, \frac{-(2\alpha + \beta)}{3} \right\}, \tag{51.458}$$

and the second one has

$$\Delta_{\phi'} = \left\{ \frac{\alpha + 2\beta}{3}, \frac{\alpha - \beta}{3}, \frac{-(2\alpha + \beta)}{3} \right\}. \tag{51.459}$$

34. The second part is not proved.

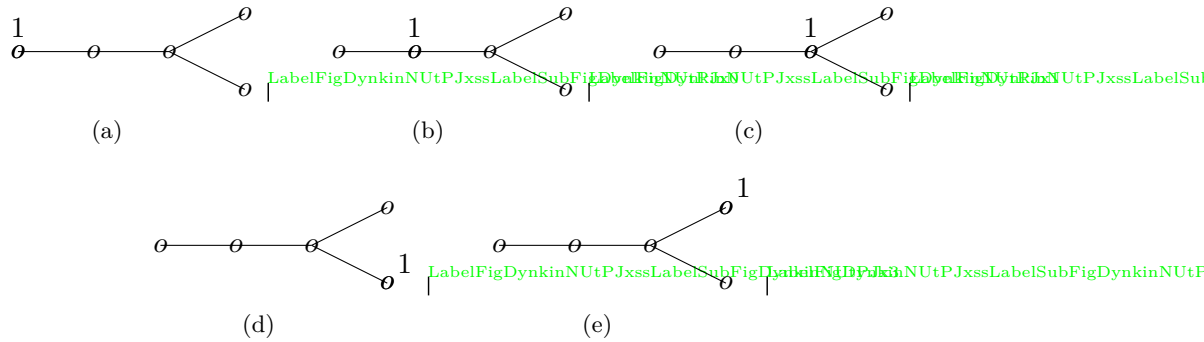


Figure 51.1: Basic representations of $\mathfrak{so}(10)$

According to equation (51.457), we have 9 weights in the representation $\phi \otimes \phi'$ (all the sums of one element of Δ_ϕ with a one of $\Delta_{\phi'}$). The highest one is

$$\frac{2\alpha + 4\beta}{3},$$

which is the double of the highest weight in , so $\phi \otimes \phi'$ contains the representation . Now, we remove from the list of weights of $\phi \otimes \phi'$ the list of weight of ; the result is

$$\frac{2\alpha + \beta}{3}, \frac{-(\alpha - \beta)}{3}, \frac{-(\alpha + 2\beta)}{3}, \tag{51.460}$$

which are the weights of . The conclusion is that [35]

(51.461)

That procedure of decomposition is quite long because it requires to compute the complete set of weights for some intermediate representations.

51.14.2.3 Symmetrization and anti symmetrization

Let ϕ be a irreducible representation. We want to compute the symmetric and antisymmetric parts of the representation $\phi^{\otimes k} = \underbrace{\phi \otimes \ldots \otimes \phi}_{k \text{ times}}$. These symmetric and antisymmetric parts are denoted by $\phi_s^{\otimes k}$ and $\phi_a^{\otimes k}$ respectively.

Proposition 51.213.
If $\{\xi_1, \ldots, \xi_N\}$ is a canonical basis of ϕ and if we denote by Λ_i the weight of the vector ξ_i, the followings hold:

(1) the weight system of $\phi_a^{\otimes k}$ is

$$\Lambda_{i_1} + \Lambda_{i_2} + \cdots + \Lambda_{i_k} \tag{51.462}$$

with $i_k > \ldots > i_2 > i_1$, and the highest weight is

$$\Lambda_1 + \cdots + \Lambda_k. \tag{51.463}$$

The dimension of the representation $\phi_a^{\otimes k}$ is

$$N\left(\phi_a^{\otimes k}\right) = \binom{n}{k}. \tag{51.464}$$

35. I guess the following line is a typo and should be 1–o times o–1.

(2) The weight system of the representation $\phi_s^{\otimes k}$ is

$$\Lambda_{i_1} + \Lambda_{i_2} + \cdots + \Lambda_{i_k} \tag{51.465}$$

with $i_k \geqslant \ldots \geqslant i_2 \geqslant i_1$, and the highest weight is

$$k\Lambda_1 \tag{51.466}$$

The dimension of the representation $\phi_s^{\otimes k}$ is

$$N\big(\phi_s^{\otimes k}\big) = \binom{n+k}{k}. \tag{51.467}$$

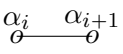
Proof. No proof. □

The representations $\phi_a^{\otimes k}$ and $\phi_s^{\otimes k}$ might be decomposable and we denote by $\phi_{s>}^{\otimes k}$ and $\phi_{a>}^{\otimes k}$ their highest weight parts.

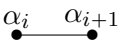
Let α be a terminal point in a Dynkin diagram. The **branch** of α is the sequence of points of the Dynkin diagram $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_k$ defined by the following properties.

— The point α_i is connected with (and only with) the points α_{i-1} and α_{i+1},

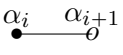
— the connexion between α_i and α_{i+1} is of one of the following forms



(51.468a)



(51.468b)



(51.468c)

— the sequence $\alpha_1, \ldots, \alpha_k$ is maximal in the sense that no α_{k+1} can be added without violating one of the two first rules.

Proposition 51.214.
Let α be a terminal point in a Dynkin diagram and $\alpha_1, \ldots, \alpha_k$ be the corresponding branch. Then we have

$$\phi_{\alpha_r} \simeq \phi_{\alpha\, a>}^{\otimes r} \tag{51.469}$$

for every $r = 1, 2, \ldots, k$.

51.15 Semi-direct product

51.15.1 From Lie algebra point of view

subsec:semi_Lie

Here, the matter comes from [865, 869]. When $\mathfrak{a}$ and $\mathfrak{b}$ are Lie algebras, one can consider $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ as vector space, and define a Lie algebra structure on $\mathfrak{g}$ by

$$[(a, b), (a', b')] = ([a, a'], [b, b']).$$

This is the **direct sum** of $\mathfrak{a}$ and $\mathfrak{b}$.

An endomorphism $\mathcal{D}$ of the Lie algebra $\mathfrak{a}$ is a **derivation** when

$$\mathcal{D}[X, Y] = [\mathcal{D}X, Y] + [X, \mathcal{D}Y].$$

The set of the derivations of $\mathfrak{a}$ is written $\mathtt{Der}\,\mathfrak{a}$.

prop:Lie_derr

Proposition 51.215.
Let $\mathfrak{a}$ be a Lie algebra

(1) $\mathtt{Der}\,\mathfrak{a}$ is a Lie algebra for the usual commutator,

(2) $\mathrm{ad}\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{a} \subseteq \mathrm{End}\,\mathfrak{a}$ *is a Lie algebra homomorphism.*

Proof. For the first statement, we just have to compute to see that if $\mathcal{D}, \mathcal{E} \in \mathtt{Der}\,\mathfrak{a}$,

$$[\mathcal{D}, \mathcal{E}][X, Y] = (\mathcal{D}\mathcal{E} - \mathcal{E}\mathcal{D})[X, Y] = [[\mathcal{D}, \mathcal{E}]X, Y] + [X, [\mathcal{D}, \mathcal{E}]Y].$$

The second comes from the fact that $\mathrm{ad}\,X \in \mathtt{Der}\,\mathfrak{a}$ for any $X \in \mathfrak{a}$ and $\mathrm{ad}[X, Y] = \mathrm{ad}\,X\,\mathrm{ad}\,Y - \mathrm{ad}\,Y\,\mathrm{ad}\,X$. □

Let us now consider the vector space direct sum $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. Let us suppose moreover that $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$ and that $\mathfrak{b}$ is an idea in $\mathfrak{g}$. So we have that

$$\mathrm{ad}\,|_{\mathfrak{b}} \in \mathtt{Der}\,\mathfrak{b}.$$

By proposition 51.215, we have a homomorphism $\pi\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{b}$, $\pi(A) = \mathrm{ad}\,A|_{\mathfrak{b}}$. So if $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$, $[A, B] = \pi(A)B$. The conclusion is that the Lie algebra structure of $\mathfrak{g}$ is given by $\mathfrak{a}$, $\mathfrak{b}$ and π. In this case, we write $\mathfrak{g} = \mathfrak{a} \oplus_\pi \mathfrak{b}$, and we say that $\mathfrak{g}$ is the semidirect product of $\mathfrak{a}$ and $\mathfrak{b}$. The following theorem gives the general definition of semidirect product.

Theorem 51.216.
Let $\mathfrak{a}$ and $\mathfrak{b}$ be two Lie algebras, and $\pi\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{b}$, a Lie algebra homomorphism. There exists an unique Lie algebra structure on the vector space $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ such that

— *the commutators on $\mathfrak{a}$ and $\mathfrak{b}$ are the old ones,*

— *$[A, B] = \pi(A)B$ for any $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$.*

In this case, in the so defined Lie algebra $\mathfrak{g}$, $\mathfrak{a}$ is a subalgebra and $\mathfrak{b}$ is an ideal.

The vector space $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ endowed with this Lie algebra structure is the **semidirect product** of $\mathfrak{a}$ and $\mathfrak{b}$, it is denoted by

$$\mathfrak{g} = \mathfrak{a} \oplus_\pi \mathfrak{b}$$

One also often speak about **split extension** of $\mathfrak{a}$ by $\mathfrak{b}$, with the splitting map π.

Proof. The unicity part is clear: the Lie algebra structure is completely defined by the two conditions and the condition of antisymmetry. The matter is just to see that this structure is a Lie algebra structure: we have to check Jacobi. If in $[[X, Y], Z]$, X, Y, Z are all three in $\mathfrak{a}$ or $\mathfrak{b}$, it is trivial. The two other cases are:

— X, $Y \in \mathfrak{a}$ and $Z \in \mathfrak{b}$. In this case, we use $\pi([X, Y]) = \pi(X)\pi(Y) - \pi(Y)\pi(x)$ (because π is a Lie algebra homomorphism) to find

$$[[X, Y], Z] = \pi([X, Y])Z = -[[Y, Z], X] - [[Z, X], Y].$$

— The second case is X, $Y \in \mathfrak{b}$ and $Z \in \mathfrak{a}$. Here, we use the fact that $\pi(Z)$ is a derivation of $\mathfrak{b}$. The computation is also direct.

It is clear that $\mathfrak{b}$ is an ideal because for any $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$, $[B, A] = -[A, B] = -\pi(A)B \in \mathfrak{b}$. □

The theory of split extension is often used in the following sense. We have a Lie algebra $\mathfrak{g}$ which decomposes (as vector space) into a direct sum $\mathfrak{a} \oplus \mathfrak{b}$. If in $\mathfrak{g}$ the map $a \mapsto \mathrm{ad}(a)$ is an action of $\mathfrak{a}$ on $\mathfrak{b}$, we say that $\mathfrak{g}$ is a split extension

$$\mathfrak{g} = \mathfrak{a} \oplus_{\mathrm{ad}} \mathfrak{b}.$$

This way to use split extensions is used for example in the proof of proposition 84.13.

51.15.1.1 Example: extensions of the Heisenberg algebra

Let $\mathscr{H}(V,\Omega) = V \oplus \mathbb{R}E$ be the Heisenberg algebra. A derivation is a map $D\colon \mathscr{H} \to \mathscr{H}$ such that

$$D[X,Y] = [DX, YT] + [X, DY]. \tag{51.470}$$

Let us look at the derivations under the form

$$D = \begin{pmatrix} X & v \\ \xi & a \end{pmatrix} \tag{51.471}$$

where $a \in \mathbb{R}$, $X \in \mathrm{End}(V)$, $v \in V$ and $\xi \in V^*$. The left hand side of the condition (51.470) reads

$$D[w + zE, w' + z'E] = D\big(\Omega(w,w')E\big) = \Omega(w,w')(v + aE). \tag{51.472}$$

Now, using $Dw = Xw + \xi(w)E$ and $D(zE) = v + aE$, the right hand side is

$$\big(\Omega(Xw,v') + \Omega(zv,v') + \Omega(w,Xw') + \Omega(w,z'v)\big)E. \tag{51.473}$$

Equating (51.472) and (51.473) we find $v = 0$ and

$$\Omega(Xw,w') + \Omega(w,Xw') = a\Omega(w,w'). \tag{51.474}$$

If we write it as matrices, we find

$$X^t\Omega + \Omega X = a\Omega. \tag{51.475}$$

The derivations with $a = 0$ form the algebra

$$\mathtt{Der}(\mathscr{H})_0 = \mathfrak{sp}(\Omega, V) \times V^*. \tag{51.476}$$

If $a \neq 0$, we find the symplectic conform group

$$\mathrm{Conf}(V,\Omega) = \{A\colon V \to V \text{ tel que } \Omega(Av,Aw) = \lambda\Omega(v,w) \text{ with } \lambda \in \mathbb{R}_0^+\}. \tag{51.477}$$

Taking the derivative of the group condition, we find

$$\frac{d}{dt}\Big[\Omega\big(A(t)v, A(t)w\big)\Big]_{t=0} = \frac{d}{dt}\Big[\lambda(t)\Omega(v,w)\Big]_{t=0}, \tag{51.478}$$

which produces the condition (51.474) with $X = \dot{A}$ and $a = \dot{\lambda}$.

(1) If $X = \mathrm{Id}$ and $\xi = 0$, then we must have $a = 2$ and we have the derivation

$$H = \mathrm{Id}\,|_V \oplus 2\,\mathrm{Id}\,|_{\mathbb{R}E}. \tag{51.479}$$

(2) If $\xi = 0$, $a = 0$ and X if exchange the Lagrangian in the decomposition $V = W \oplus \bar{W}$.

51.15.2 Group algebra

Let $\mathcal{A}$ and $\mathcal{B}$ be abelian algebras and $\rho\colon \mathcal{A} \to \mathtt{Der}\,\mathcal{B}$ be a homomorphism. We want to put a group structure on the set $\mathcal{A} \times \mathcal{B}$ in such a way that the Lie algebra of $\mathcal{A} \times \mathcal{B}$ has Lie bracket given by

$$\big[(A+B),(A'+B')\big] = [A,B'] + [B,A'] = \rho(A)B' - \rho(A')B. \tag{51.480}$$

We claim that the group law should be

$$(a,b)(a',b') = (a + a', e^{\rho(a)}b' + b) \tag{51.481}$$

whose inverse is

$$(a,b)^{-1} = (-a, -e^{-\rho(a)b}) \tag{51.482}$$

Indeed, the general form of the commutator is

$$[X,Y] = \frac{d}{dt}\frac{d}{ds}\Big[\,\mathbf{Ad}(X(t))Y(s)\Big]_{\substack{s=0\\t=0}}$$

with respect to the group law. A path in $\mathcal{A} \times \mathcal{B}$ with tangent vector (a,b) is (at,bt). Then

$$\begin{aligned}\big[(a,b),(a',b')\big] &= \frac{d}{dt}\frac{d}{ds}\Big[(at,bt)(a's,b's)(at,bt)^{-1}\Big]_{\substack{s=0\\t=0}} \\ &= \big(0, -\rho(a)b + \rho(a)b'\big).\end{aligned} \tag{51.483}$$

Chapter 52

Lie groups

Do you know what is violet and commutative? Answer in the footnote [1].

52.1 Lie groups

The notion of topological group is defined in 48.49.

DEFooGDWTooTvINuw

Definition 52.1 ([1, 870, 871])**.**
*Let $\mathcal{A}$ be a class of functions: C^k, C^∞, analytic, etc. A $\mathcal{A}$-**Lie group** is a group G which is in the same times a $\mathcal{A}$-manifold such that the group operations*

$$\begin{aligned} i\colon G &\to G \\ g &\mapsto g^{-1} \end{aligned} \tag{52.1}$$

and

$$\begin{aligned} m\colon G\times G &\to G \\ (g,h) &\mapsto gh \end{aligned} \tag{52.2}$$

are in the class $\mathcal{A}$.

*In particular, a Lie group is **analytic** if the manifold is analytic and the group operations are analytic.*

For a Lie subgroup, see definition 52.24.

52.2 (Smooth or analytic ?)**.**
At this point, when we say "Lie group", we intend a smooth Lie group. That has the advantage of not being obliged to prove analyticity of the maps. But at some point we want to use power expansions of the exponential map, which are typical question requiring analyticity.

Thus we will turn to analytic Lie groups.

The theorem 53.66 will help us saying that every single smooth Lie group we already studied can be turned into an analytic Lie group by choosing a good atlas. So we will not lose any results in the "conversion".

PROPooAXYRooWVhXRa

Proposition 52.3 (Leibniz rule[1])**.**
Let G be a Lie group. We consider two smooth paths $\gamma\colon I\to G$ and $\sigma\colon I\to G$ where I is an interval containing 0. We suppose that $\gamma(0)=\sigma(0)=e$. We have

$$\frac{d}{dt}\Big[\gamma(t)\sigma(t)\Big]_{t=0} = \gamma'(0)+\sigma'(0). \tag{52.3}$$

PROPooSBVCooOZnszF

Proposition 52.4 (2-torus)**.**
Consider the equivalence relation $\sim$ on $\mathbb{R}^2$ $(x,y)\sim(x',y')$ if there exists integers $k,l\in\mathbb{Z}$ such that $x'=x+k$ and $y'=y+k$. We have:

1. An abelian grape!

(1) $\mathbb{R}^2/\sim$ *is a manifold.*

(2) $\mathbb{R}^2/\sim$ *is a Lie group.*

52.1.1 Connected component of Lie groups

PropUssGpGenere

Proposition 52.5.
If G is a connected Lie group and $\mathcal{U}$, a neighbourhood of the identity e, then G is generated by $\mathcal{U}$ in the sense that $\forall g \in G$, there exists a finite number of $g_i \in \mathcal{U}$ such that

$$g = g_1 \dots g_n. \tag{52.4}$$

Notice that the number n is function of g in general.

Proof. Eventually passing to a subset, we can suppose that $\mathcal{U}$ is open. In this case, $\mathcal{U}^{-1}$ is open because it is the image of $\mathcal{U}$ under the homeomorphism $g \mapsto g^{-1}$. Now we consider $V = \mathcal{U} \cap \mathcal{U}^{-1}$. The main property of this set is that $V = V^{-1}$. Let

$$[V] = \{g_1 \dots g_n \text{ tel que } g_i \in V\};$$

we will prove that $[V] = G$ by proving that it is closed and open in G (the fact that G is connected then concludes).

We begin by openness of $[V]$. Let $g_0 = g_1 \cdots g_n \in [V]$. We know that g_0V is open because the multiplication by g_0 is an homeomorphism. It is clear that $g_0V \subset [V]$ and that $g_0 = g_0e \in g_0V$. Hence $g_0 \in g_0V \subseteq [V]$. It proves that $[V]$ is open because g_0V is a neighbourhood of g_0 in $[V]$.

We now turn our attention to the closeness of $[V]$. Let $h \in \overline{[V]}$. The set hV is an open set which contains h and $hV \cap [V] \neq \varnothing$ because an open which contains an element of the closure of a set intersects the set (it is almost the definition of the closure). Let $g_0 \in hV \cap [V]$. There exists a $h_1 \in V$ such that $g_0 = hh_1$. For this h_1, we have $hh_1 = g_0 = g_1 \cdots g_n$, and therefore

$$h = g_1 \cdots g_n h_1^{-1} \in [V].$$

This proves that $h \in [V]$ because $h_1^{-1} \in V$ from the fact that $V = V^{-1}$. □

Remark that this proof emphasises the topological aspect of a Lie group: the differential structure was only used to prove thinks like that A^{-1} is open when A is open.

Proposition 52.6.
Let G be a Lie group and G_0, the identity component of G. We have the following:

(1) *G_0 is an open invariant subgroup of G,*

(2) *G_0 is a Lie group,*

(3) *the connected components of G are lateral classes of G_0. More specifically, if x belongs to the connected component G_1, then $G_1 = xG_0 = G_0x$.*

Proof. We know that when M_1 is open in the manifold M, one can put on M_1 a differential structure of manifold of same dimension as M with the induced topology. Since G_0 is open, it is a smooth manifold. In order for G_0 to be a Lie group, we have to prove that it is stable under the inversion and that $gh \in G_0$ whenever g, $h \in G_0$.

First, G_0^{-1} is connected because it is homeomorphic to G_0 in G. The element e belongs to the intersection of G_0 and G_0^{-1}, so $G_0 \cup G_0^{-1}$ is connected as non-disjoint union of connected sets. Hence $G_0 \cup G_0^{-1} = G_0$ and we conclude that $G_0^{-1} \subseteq G_0$. The set G_0G_0 is connected because it is the image of $G_0 \times G_0$ under the multiplication map, but $e \in G_0G_0$, so $G_0G_0 \subseteq G_0$ and G_0 is thus closed for the multiplication. Hence G_0 is a Lie group.

For all $x \in G$, we have $e = xex^{-1} \in xG_0x^{-1}$, but xG_0x^{-1} is connected. Hence $xG_0x^{-1} \subseteq G_0$, which proves that G_0 is an invariant subset of G.

Lateral classes xG_0 are connected because the left multiplication is an homeomorphism. They are moreover *maximal* connected subsets because, if $xG_0 \subset H$ (proper inclusion) with a connected

H, then $G_0 \subset x^{-1}H$ (still proper inclusion). But the definition of G_0 is that this proper inclusion is impossible. Therefore, the sets of the form xG_0 are maximally connected sets. It is clear that $\bigcup_{g\in G} gG_0 = G$.

Notice that the last point works with G_0x too. □

52.1.2 What is $g^{-1}dg$?

SubSecgmudg

The expression $g^{-1}dg$ is often written in the physical literature. In our framework, the way to gives a sense to this expression is to consider it pointwise acting on a tangent vector. More precisely, the framework is the data of a manifold M, a Lie group G and a map $g\colon M \to G$. Pointwise, we have to apply $g(x)^{-1}dg_x$ to a tangent vector $v \in T_xM$.

Note that $dg_x\colon T_xM \to T_{g(0)}G \neq T_eG$, so $dg_x \notin \mathcal{G}$. But the product $g(x)^{-1}dg_xv$ is defined by

$$g(x)^{-1}dg_xv = \frac{d}{dt}\Big[g(x)^{-1}g(v(t))\Big]_{t=0} \in \mathcal{G}. \tag{52.5}$$

52.1.3 Product of Lie groupe

PROPooKENCooMKJJjV

Proposition-Definition 52.7.
Let G, H be Lie groups. Then $G \times H$ is a Lie group for the product manifold structure[2] *and the product group structure.*

When we speak about the Lie group $G \times H$, we are intending that structure.

PROPooKITOooTcsIiu

Proposition 52.8.
If G and H are smooth Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, then the Lie algebra of $G \times H$ is $\mathfrak{g} \times \mathfrak{h}$.

52.1.4 Left and right translations

LEMooPIUFooHjyXln

Lemma 52.9.
Let G be a Lie group. Let $g, h \in G$. We consider the ***left translation***

$$\begin{aligned} L_h\colon G &\to G \\ g &\mapsto hg. \end{aligned} \tag{52.6}$$

(1) The map L_h is a diffeomorphism.
(2) The differential $(dL_h)_g$ is a vector space isomorphism.

52.2 Two words about Lie algebra

52.2.1 Adjoint map

The ideas of this short note comes from [872]. A more traumatic definition of the adjoint group can be found in [813], chapter II, §5.

Definition 52.10.
Let G be a Lie group, and $\mathcal{G}$, its Lie algebra. We define the ***adjoint map*** *at the point $x \in G$ by*

$$\begin{aligned} \boldsymbol{Ad}_x\colon G &\to G \\ \boldsymbol{Ad}_x\, y &= xyx^{-1} \end{aligned} \tag{52.7}$$

Then we define

$$Ad_x := (d\,\mathbf{Ad}_x)_e\colon \mathcal{G} \to \mathcal{G};$$

the chain rule applied on $\mathbf{Ad}_{xy} = \mathbf{Ad}_x \circ \mathbf{Ad}_y$ leads to $Ad_{xy} = Ad_x \circ Ad_y$, and thus we can see Ad as a group homomorphism $Ad\colon G \to GL(\mathcal{G})$, $Ad(x) = Ad_x$.

2. Product of manifolds, definition 49.46.

Definition 52.11.
*This homomorphism is the **adjoint representation** of the group G in the vector space $\mathcal{G}$.*

Finally, we define

$$ad := d(Ad)_1 \colon \mathcal{G} \to L(\mathcal{G}, \mathcal{G})$$

where we identify $T_1 GL(\mathcal{G})$ with $L(\mathcal{G}, \mathcal{G})$.

LEMooEALFooJOeOgk

Lemma 52.12.
If $f \colon G \to G$ is an automorphism of G (i.e.: $f(xy) = f(x)f(y)$), then df_e is an automorphism of $\mathcal{G}$: $df[X, Y] = [dfX, dfY]$

Proof. First, remark that $f(\mathbf{Ad}_x\, y) = \mathbf{Ad}_{f(x)}\, f(y)$. Now, $\mathrm{Ad}_x\, X = (d\,\mathbf{Ad}_x)_e X$, so that one can compute:

$$\begin{aligned} df(\mathrm{Ad}_x\, X) &= \frac{d}{dt}\Big[f(\mathbf{Ad}_x\, X(t)) \Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathbf{Ad}_{f(x)}\, f(X(t)) \Big]_{t=0} \\ &= (d\,\mathbf{Ad}_{f(x)})_{f(e)} df X \\ &= \mathrm{Ad}_{f(x)}\, df X. \end{aligned} \tag{52.8}$$

On the other hand, we need to understand how does the ad work.

$$\mathrm{ad}\, XY = \frac{d}{dt}\Big[\mathrm{Ad}_{X(t)} \Big]_{t=0} Y = \frac{d}{dt}\Big[\mathrm{Ad}_{X(t)}\, Y \Big]_{t=0}$$

because $\mathrm{Ad}_{X(t)} \colon \mathcal{G} \to \mathcal{G}$ is linear, so that Y can enter the derivation (for this, we identify $\mathcal{G}$ and $T_X \mathcal{G}$). Since $\mathrm{Ad}_{X(t)}\, Y$ is a path in $\mathcal{G}$ the *true space* is

$$(\mathrm{ad}\, X)Y = \frac{d}{dt}\Big[\mathrm{Ad}_{X(t)}\, Y \Big]_{t=0} \in T_{[X,Y]}\mathcal{G} \simeq \mathcal{G}.$$

For the same reason of linearity, df can get in the derivative in the expression $df \frac{d}{dt}\Big[\mathrm{Ad}_{X(t)}\, Y \Big]_{t=0}$. Thus

$$\begin{aligned} (\mathrm{ad}\, X)Y &= \frac{d}{dt}\Big[df\big(\mathrm{Ad}_{X(t)}\, Y \big) \Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathrm{Ad}_{f(X(t))}\, dfY \Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathrm{Ad}_{f(X(t))} \Big]_{t=0} dfY \\ &= \mathrm{ad}(dfX) dfY \\ &= [dfX, dfY] \end{aligned} \tag{52.9}$$

because $f(X(t))$ is a path which gives dfX. □

Corollary 52.13.
An automorphism of a semisimple Lie group is an isometry for the Killing metric. Stated in other words,

$$\mathrm{Aut}(G) \subset \mathrm{Iso}\, G. \tag{52.10}$$ eq:Aut_Iso

Proof. By lemma 52.12, if f is an automorphism of G, df is an automorphism of $\mathcal{G}$. Now, by proposition 51.17, f is an isometry of G. □

Theorem 52.14.
The Killing form is bi-invariant on G. tho:bi_invariance

Proof. Because of the left invariance,

$$B(dR_g X, dR_g Y) = B(dL_{g^{-1}} dR_g X, dL_{g^{-1}} dR_g Y) = B(\mathrm{Ad}_{g^{-1}} X, \mathrm{Ad}_{g^{-1}} Y).$$

But $\mathrm{Ad}_{g^{-1}} = d(\mathbf{Ad}_{g^{-1}})$ and $\mathbf{Ad}_{g^{-1}}$ is an automorphism of G. Thus by lemma 52.12 and proposition 51.17,

$$B\big(\mathrm{Ad}(g^{-1})X, \mathrm{Ad}(g^{-1})Y\big) = B(X,Y). \tag{52.11}$$

eq_KillAdinvariant

□

Lemma 52.15.
In the case of Lie algebra, the bracket is given by the derivative of the adjoint action:

$$\frac{d}{dt}\Big[\,\mathrm{Ad}(e^{tX})Y\Big]_{t=0} = [X,Y] \tag{52.12}$$

Proof. Let us make $[\tilde{X},\tilde{Y}]_e$ act on a function f. Using the definition (49.356) and the property of theorem 49.152, we have

$$\begin{aligned}
[\tilde{X},\tilde{Y}]_e f &= \frac{d}{dt}\Big[(d\varphi^X_{-t})\tilde{Y}\Big]_{t=0} f \\
&= \frac{d}{dt}\Big[(d\varphi^X_{-t})_{\varphi^X_t(e)}\big(\tilde{Y}_{\varphi^X_t(e)}\big)\Big]_{t=0} f \\
&= \frac{d}{dt}\Big[\tilde{Y}_{e^{tX}} \cdot (f\circ\varphi^X_{-t})\Big]_{t=0}
\end{aligned} \tag{52.13}$$

Now, we use the fact that, by definition, $\varphi^X_t(x) = xe^{tX}$, so that $\varphi^Y_s(e^{tX}) = e^{tX}e^{sY}$ and we get

$$\begin{aligned}
[\tilde{X},\tilde{Y}]_e f &= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f\big(\varphi^X_{-t}(e^{tX}e^{sY})\big)\Big]_{s=0}\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(e^{tX}e^{sY}e^{-tX})\Big]_{s=0}\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f\big(e^{s\,\mathrm{Ad}(e^{tX})Y}\big)\Big]_{s=0}\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\big(\mathrm{Ad}(e^{tX})Y\big)_e \cdot f\Big]_{t=0}
\end{aligned} \tag{52.14}$$

□

52.3 Universal enveloping algebra

subsec:env_alg

Let $\mathcal{A}$ be a Lie algebra. One knows that the composition law $(X,Y)\to[X,Y]$ is often non associative. In order to build an associative Lie algebra which "looks like" $\mathcal{A}$, one considers $T(\mathcal{A})$, the tensor algebra of $\mathcal{A}$ (as vector space) and $\mathcal{J}$ the two-sided ideal in $T(\mathcal{A})$ generated by elements of the form

$$X\otimes Y - Y\otimes X - [X,Y]$$

for $X, Y\in\mathcal{A}$. The **universal enveloping algebra** of $\mathcal{A}$ is the quotient

$$U(\mathcal{A}) = T(\mathcal{A})/\mathcal{J}. \tag{52.15}$$

For $X\in\mathcal{A}$, we denote by X^* the image of X by canonical projection $\pi\colon T(\mathcal{A})\to U(\mathcal{A})$ and by 1 the unit in $U(\mathcal{A})$. One has $1\neq 0$ if and only if $\mathcal{A}\neq\{0\}$.

Proposition 52.16 ([813])**.**
Le V be a vector space on K. Then there is a natural bijection between the representations of $\mathcal{A}$ on V and the ones of $U(\mathcal{A})$ on V. If ρ is a representation of $\mathcal{A}$ on V, the corresponding ρ^* of $U(\mathcal{A})$ is given by

$$\rho(X) = \rho^*(X^*)$$

($X\in\mathcal{A}$).

Let $\{X_1, \dots, X_n\}$ be a basis of $\mathcal{A}$. For a n-uple of complex numbers (t_i), one defines

$$X^*(t) = \sum_{i=1}^{n} t_i X_i^*. \tag{52.16}$$

On the other hand, we consider a n-uple of positive integers $M = (m_1 + \dots m_n)$, and the notation

$$\begin{aligned} |M| &= m_1 + \cdots + m_n \\ t^M &= t_1^{m_1} \cdots t_n^{m_n}. \end{aligned} \tag{52.17}$$

When $|M| > 0$, we denote by $X^*(M) \in U(\mathcal{A})$ the coefficient of t^M in the expansion of $(|M|!)^{-1}(X^*(t))^{|M|}$. If $|M| = 0$, the definition is $X^*(0) = 1$. Once again a proposition without proof.

Proposition 52.17.
The smallest vector subspace of $U(\mathcal{A})$ which contains all the elements of the form $X^*(M)$ is $U(\mathcal{A})$ itself:

$$U(\mathcal{A}) = \mathrm{Span}\{X^*(M) : M \in \mathbb{N}^n\}.$$

cor:/24

Corollary 52.18.
Let $\mathcal{A}$ be a Banach algebra of dimension n, $\mathcal{B}$ a Banach subalgebra of dimension $n - r$ and a basis $\{X_1, \dots, X_n\}$ of $\mathcal{A}$ such that the $n - r$ last basis vectors are in $\mathcal{B}$. We denotes by B the vector subspace of $U(\mathcal{A})$ spanned by the elements of the form $X^*(M)$ with $m = (0, \dots, 0, m_{r+1}, \dots, m_n)$. Then B is a subalgebra of $U(\mathcal{A})$.

Definition 52.19.
*Two Lie groups G and G' are **isomorphic** when there exists a differentiable group isomorphism between G and G'.*

*They are **locally isomorphic** when there exists neighbourhoods $\mathcal{U}$ and $\mathcal{U}'$ of e and e' and a differentiable diffeomorphism $f : \mathcal{U} \to \mathcal{U}'$ such that*

$\forall x, y, xy \in \mathcal{U},\ f(xy) = f(x)f(y)$,

and

$\forall x', y', x'y' \in \mathcal{U}',\ f^{-1}(x'y') = f^{-1}(x')f^{-1}(y')$.

The following universal property of the *universal* enveloping algebra explains the denomination:

Proposition 52.20.
Let $\sigma : \mathcal{G} \to \mathcal{U}(\mathcal{G})$ the canonical inclusion and A an unital complex associative algebra. A linear map $\varphi : \mathcal{G} \to A$ such that

$$\varphi[X, Y] = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) \tag{52.18}$$

can be extended in only one way to an algebra homomorphism $\varphi_0 : \mathcal{U}(\mathcal{G}) \to A$ such that $\varphi_0 \circ \sigma = \varphi$ and $\varphi(1) = 1$ prop:extunifmap

For a proof, see [859].

52.3.1 Adjoint map in $\mathcal{U}(\mathcal{G})$

ssadjunif

We know that $\mathrm{Ad}(g) : \mathcal{G} \to \mathcal{G}$ fulfils

$$\mathrm{Ad}(g)[X, y] = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y],$$

and we can define $\mathrm{Ad}(g) : \mathcal{G} \to \mathcal{U}(\mathcal{G})$ by $\mathrm{Ad}(g)X = X$ where in the right hand side, X denotes the class of X for the quotients of the tensor algebra which defines the universal enveloping algebra.

When $[A, B]$ is seen in $\mathcal{U}(\mathcal{G})$, we have $[A, B] = A \otimes B - B \otimes A$. Then $\mathrm{Ad}(g) : \mathcal{G} \to \mathcal{U}(\mathcal{G})$ fulfils proposition 52.20 and is extended in an unique way to $\mathrm{Ad}(g) : \mathcal{U}(\mathcal{G}) \to \mathcal{U}(\mathcal{G})$ with $\mathrm{Ad}(g)1 = 1$.

Lemma 52.21.
If $D \in \mathcal{U}(\mathcal{G})$, the following properties are equivalent:

— $D \in \mathcal{Z}(\mathcal{G})$
— $D \otimes X = X \otimes D$ *for all* $X \in \mathcal{G}$
— $e^{\mathrm{ad}\, X} D = D$ *for all* $X \in \mathcal{G}$
— $\mathrm{Ad}(g)D = D$ *for all* $g \in G$.

lem:equivDAd

52.3.2 Invariant fields

If $X \in \mathfrak{g}$, we have the associated left invariant vector field on G given by $\tilde{X}_x = dL_x X$. That field is left invariant as operator on the functions because

$$\tilde{X}_x(u) = \tilde{X}_e(L_x^* u) \tag{52.19}$$

as the following computation shows

$$\tilde{X}_e(L^* u) = \frac{d}{dt}\Big[(L_x^* u)\big(e^{tX}\big)\Big]_{t=0} = \frac{d}{dt}\Big[u\big(xe^{tX}\big)\Big]_{t=0} = \frac{d}{dt}\Big[u\big(\tilde{X}_x(t)\big)\Big]_{t=0} = \tilde{X}_x(u) \tag{52.20}$$

because the path defining $\tilde{X}_x$ is xe^{tX}.

We can perform the same construction in order to build left invariant fields based on $\mathcal{U}(\mathfrak{g})$. If X and Y are elements of $\mathfrak{g}$, the differential operator on $C^\infty(G)$ associated to $XY \in \mathcal{U}(\mathfrak{g})$ is given by

$$(XY)(f) = \frac{d}{dt}\frac{d}{ds}\Big[f\big(X(s)Y(t)\big)\Big]_{\substack{s=0\\t=0}} \tag{52.21}$$

The path defining the field $\widetilde{XY}$ is

$$\widetilde{XY}_x = xX(s)Y(t). \tag{52.22}$$

Thus we have

$$\widetilde{(XY)}_e(L^* u) = \widetilde{(XY)}_x u \tag{52.23}$$

EqInvarUgField
LemAdesthioo

Lemma 52.22.
If $X, Y \in \mathfrak{g}$ we have

$$[\mathrm{ad}(X), \mathrm{ad}(Y)] = \mathrm{ad}([X, Y]). \tag{52.24}$$

Proof. Let $f \in \mathfrak{g}$ and compute the action of $[\mathrm{ad}(X), \mathrm{ad}(Y)]$:

$$\begin{aligned}
[\mathrm{ad}(X), \mathrm{ad}(Y)]f &= \mathrm{ad}(X)[Yf, fY] - \mathrm{ad}(Y)(Xf - fX) && \text{(52.25a)}\\
&= (XY - YX)f + f(YX - XY) && \text{(52.25b)}\\
&= \mathrm{ad}([X, Y])f. && \text{(52.25c)}
\end{aligned}$$

□

52.3.3 Representation of Lie groups

Proposition 52.23.
Let G be a Lie group and $\mathcal{G}$ its Lie algebra. A representation $\varphi\colon G \to \mathrm{End}(V)$ of the group induces a representation $\phi\colon \mathcal{U}(\mathcal{G}) \to \mathrm{End}(V)$ of the universal enveloping algebra with the definitions

$$\begin{aligned}
\phi(X) &= d\varphi_e(X), && \text{(52.26a)}\\
\phi(XY) &= \phi(X) \circ \phi(Y) && \text{(52.26b)}
\end{aligned}$$

where e is the unit in G and X, Y are any elements of $\mathcal{G}$.

Proof. We have

$$\phi(X) = \frac{d}{dt}\Big[\varphi(e^{tX})v\Big]_{t=0} = d\varphi_e(X)v. \tag{52.27}$$

Notice that, by linearity of the action of $\varphi(e^{tX})$ on v, one can leave v outside the derivation. Now, neglecting the second order terms in t in the derivative, and using the Leibniz formula [3], we have

$$\begin{aligned}
\phi([X,Y])v &= \frac{d}{dt}\Big[\varphi(e^{tXY}e^{-tXY})\Big]_{t=0}v \\
&= \frac{d}{dt}\Big[\varphi(e^{tXY})\varphi(\mathbb{1})\Big]_{t=0}v + \frac{d}{dt}\Big[\varphi(\mathbb{1})\varphi(e^{-tXY})\Big]_{t=0}v \\
&= \phi(XY)v - \phi(YX)v \\
&= \big(\phi(X)\phi(Y) - \phi(Y)\phi(X)\big)v \\
&= [\phi(X),\phi(Y)]v,
\end{aligned} \tag{52.28}$$

which is the claim. □

52.4 Lie subgroup

DEFooGCHDooHUMSju

Definition 52.24 ([871]).
*We say that H is a **Lie subgroup** of the Lie group G if*

(1) H is a subgroup of G,

(2) H is a Lie group by itself,

(3) the inclusion map from $\iota\colon H \to G$ is an injective immersion [4] and a group morphism.

Remarque 52.25.
A Lie subgroup is not always a submanifold. See the proposition 52.26.

PROPooJGYRooSKPVSX

Proposition 52.26 (Lie subgroup which is not a submanifold[871]).
Let G be the 2-torus [5]. Let $\alpha \in \mathbb{R}$ be irrational and consider $H = \{(x,\alpha x) \in \mathbb{R}^2\}/\sim$. Then :

(1) As sets $H \neq G$.

(2) The set H is a subgroup of G.

(3) H is not a submanifold of G.

(4) The closure of H in G is G.

(5) H is a Lie subgroup of G.

PROPooFXZJooCOFXZX

Proposition 52.27.
A closed Lie subgroup is a submanifold [6].

Proof. Let H be a Lie subgroup of the Lie group G; we have to prove that H is a submanifold of G. As far as the notation are concerned we write $\dim(H) = m$ and $\dim(G) = n$. Let $h \in H$. We know from the definition 52.24 that the inclusion $\iota\colon H \to G$ is an injective immersion, and in particular the map

$$d\iota_h\colon T_hH \to T_{\iota(h)}G \tag{52.29}$$

is injective. In particular the rank of $d\iota_h$ is equal to $\dim(H)$ and the rank theorem 49.141 applies. There exists charts $\varphi\colon U \to H$ and $\psi\colon V \to G$ such that

$$\begin{aligned}
&\psi^{-1}\circ\iota\circ\varphi\colon U \to V \\
&\quad (x_1,\dots,x_m) \mapsto (x_1,\dots,x_m,0,\dots,0).
\end{aligned} \tag{52.30}$$

3. Proposition 52.3.
4. Definition 49.110.
5. Definition 52.4.
6. Definition 49.107.

What we need is a map $\alpha\colon W \to G$ such that

$$\alpha^{-1}\big(\alpha(W) \cap H\big) = \{(x_1, \dots, x_m, 0, \dots, 0)\} \cap W. \tag{52.31}$$

We prove that $\alpha = \psi$ with $W = V$ works. Indeed suppose that $x \in V$ satisfies $\psi(x) \in H$. In that case we have $\psi(x) = \varphi(x')$ for some $x' \in U$ and

$$\psi^{-1}(\psi(x)) = \psi^{-1}\big(\varphi(x')\big) = (\psi^{-1} \circ \iota \circ \varphi)(x') = (x'_1, \dots, x'_m, 0, \dots, 0). \tag{52.32}$$

□

CORooMCWWooXkpkNO

Corollary 52.28.
If H_1 and H_2 are two Lie subgroups of the Lie group G such that $H_1 = H_2$ as topological groups, then $H_1 = H_2$ as Lie groups.

Proposition 52.29.
Let G_1 and G_2 be two Lie groups with same Lie algebra such that $\pi_0(G_1) = \pi_0(G_2)$ and $\pi_1(G_1) = \pi_1(G_2)$, then G_1 and G_2 are isomorphic.

Proof. The assumptions of equality of Lie algebras and of the π_0 make that the universal covering $\tilde{G}_1$ and $\tilde{G}_2$ of G_1 and G_2 are the same. But we know that $G_i = \tilde{G}_i/\pi_1(G_i)$. Now equality $\pi_1(G_1) = \pi_1(G_2)$ concludes that $G_1 = G_2$. □

52.5 Semi-direct product of Lie groups

Definition 52.30.
*A subgroup H is **normal** in the group G if for any $g \in G$ and $a \in H$, $gag^{-1} \in H$.*

If G is a group, N a normal subgroup and L a subgroup, we have $LN = NL$ where, by notation, if A and B are subsets of G, $AB = \{xy | x \in A, y \in B\}$.

If N and L are groups, an **extension** of N by G is a short exact sequence

$$e \longrightarrow N \xrightarrow{i} G \xrightarrow{\pi} L \xrightarrow{L} e \tag{52.33}$$

which means that

(1) i is injective because only e_N is sent to e_G,

(2) π is surjective because the whole L is sent to e.

One often say that G is an extension of N by L. In the most common case, i is the inclusion, $L = G/N$ and π is the natural projection.

We say that the extension is **split** when there exists a *split homomorphism* $\rho\colon L \to G$ such that $\rho \circ \pi = \mathrm{Id}_G$.

Definition 52.31.
*We say that G is the **semidirect product** of N and L when any $g \in G$ can be written in one and only one way as $g = nl$ with $n \in N$ and $l \in L$.*

Definition 52.32.
*A **Lie group homomorphism** between G and G' is a map $u\colon G \to G'$ which is a group homomorphism and a morphism between G and G' as differentiable manifolds.*

Lemma 52.33.
Any continuous (group) homomorphism between two Lie groups is a Lie *group homomorphism.*

We consider G, a connected Lie group; N, a closed normal subgroup; and L, a connected immersed Lie group. Moreover, we suppose that G is semidirect product of N and L.

Proposition 52.34.
The restriction to L of the canonical projection $\pi\colon G \to G/N$ is continuous for the induced topology from G to L.

Proof. The definition of an open set $\mathcal{U}$ in G/N is that $\pi^{-1}(\mathcal{U})$ is open in G. Then it is clear that π is continuous. The matter is to check it for $\pi|_L$. Let $\mathcal{U}$ be a subset of $\pi(L)$. It is unclear that $\pi^{-1}(\mathcal{U}) \subset L$, but it is true that $\pi|_L^{-1}(\mathcal{U}) \subset L$.

As far as the induced topology on L is concerned, $A \subset L$ is open when $A = \mathcal{O} \cap L$ for a certain open set $\mathcal{O}$ in G.

Let $\mathcal{U}$ be an open subset of $\pi|_L(L)$; this is $\pi^{-1}(\mathcal{U})$ is open in G. We have to compare $\pi^{-1}(\mathcal{U})$ and $\pi|_L^{-1}(\mathcal{U})$. Since

$$\pi|_L^{-1}(\mathcal{U}) = \{x \in L | \pi(x) \in \mathcal{U}\},$$

we have $\pi|_L^{-1}(\mathcal{U}) = \pi^{-1}(\mathcal{U}) \cap L$. But $\pi^{-1}(\mathcal{U})$ is open in G, then $\pi^{-1}(\mathcal{U}) \cap L$ is open in L. □

Proposition 52.35.
The group G is the semidirect product of N and L if and only if $G = NL$ and $N \cap L = \{e\}$.

Proof. If G is semidirect product of N and L, $G = NL$ is clear. In this case, if $e \neq z \in N \cap L$, $z = ez = ze$, thus $z \in G$ can be written in two ways as xy with $x \in N$ and $y \in L$.

For the converse, let us consider $n'l' = nl$. Then $x^{-1}x' = yy'^{-1} \in N \cap L = \{e\}$. Thus $x' = x$ and $y' = y$. □

Now, we consider N, a normal subgroup of G. If $\pi\colon G \to G/N$ is the canonical homomorphism, the restriction $\pi|_L\colon L \to G/N$ is an isomorphism. Indeed, on the one hand, this is surjective because $G = NL$ yields $[g] = [nl] = [l] = \pi|_L(l)$. On the other hand, $\pi|_L(l) = \pi|_L(l')$ implies that $l = nl'$ for a certain $n \in N$. Then $ll'^{-1} = n \in N \cap L = \{e\}$. So $n = e$ and $l = l'$.

Remarque 52.36.
If N is any normal subgroup of G, there doesn't exist in general any subgroup L of G such that G should be the semidirect product of N and L.

If G is the semidirect product of N and L, for any $y \in L$, $\sigma_y\colon x \to yxy^{-1}$ is an automorphism of N. The point is that $\sigma_y(a) \in N$ for all $a \in N$ because N is a normal subgroup.

It is also clear that $\forall\, u, v \in L$, $\sigma_{uv} = \sigma_u \circ \sigma_v$. Then $\sigma\colon L \to \operatorname{Aut} N$ [7] is a homomorphism. Moreover, the data of σ, N and L determines the law in G (provided the fact that the product NL is seen as formal) because any element of G can be written as nl; thus a product GG is $(nl)(n'l') = (n\sigma_y(n'))(ll')$

Proposition 52.37.
Let N and L be two Lie groups and $\sigma\colon L \to \operatorname{Aut} N$ a homomorphism. With the law

$$(x, y)(x', y') = (x\sigma_y(x'), yy'),$$

the set $S = N \times L$ is a group.

Proof. □

The set $N \times L$ endowed with this inner product is denoted

$$N \times_\sigma L.$$

Proposition 52.38.
If G is the semidirect product of N and L, then G is isomorphic to $N \times_\sigma L$.

7. $\operatorname{Aut} N$ is the set of all the automorphism of N.

Proof. The isomorphism is $T\colon N\times_\sigma L\to G$, $T(x,y)=xy$. On the one hand, it is bijective because an element of G can be written as nl with $n\in N$ and $l\in L$ in only one way. On the other hand, it is easy to check that $T((x,y)(x',u'))=T(x,y)T(x',y')$. □

One can now give the final definition. Let us consider two connected Lie groups N, L and a Lie group homomorphism $\sigma\colon L\to \operatorname{Aut} N$. The map $N\times L\to N$, $(x,y)\to\sigma_y(x)$ is C^∞. So, the group structure on $N\times L$ given by

$$(x,y)(x',y')=(x\sigma_y(x'),yy') \tag{52.34}$$

eq:prod_semi_direct

is compatible with the C^∞ structure of $N\times L$ (seen as a Lie group). The manifold $N\times L$ endowed with the group structure (52.34) is the **semidirect product** on N and L; this is denoted by

$$N\times_\sigma L.$$

52.5.1 Introduction by exact short sequence

52.5.1.1 General setting

Let G_0, G_1 and G_2 be tree connected Lie groups. A **short exact sequence** between them is two group homomorphisms

$$\begin{aligned}\iota&\colon G_0\to G_1\\ \pi&\colon G_1\to G_2\end{aligned} \tag{52.35}$$

such that $\operatorname{Image}(\iota)=\ker(\pi)$. In that case, one says that G_1 is an **extension** of G_2 by G_0.

Since the group $\iota(G_0)$ is the kernel of an homomorphism, it is normal and we write $\iota(G_0)\lhd G_1$. Moreover, $\iota(G_0)=\pi^{-1}(e_2)$ and is then closed in G_1. As group, we have

$$G_2=G_1/\iota(G_0). \tag{52.36}$$

The extension is **split** if there exists a Lie group homomorphism $j\colon G_2\to G_1$ such that

$$\pi\circ j=\operatorname{Id}|_{G_2}. \tag{52.37}$$

This condition imposes j to be injective. In that case we have an action of G_2 on G_0 defined by

$$\begin{aligned}R&\colon G_2\to\operatorname{Aut}(G_0)\\ R_{g_2}(g_0)&=\iota^{-1}\Big(\mathbf{Ad}\big(j(g_2)\big)\iota(g_0)\Big).\end{aligned} \tag{52.38}$$

Notice that $\mathbf{Ad}\big(j(g_2)\big)\iota(g_0)$ belongs to $\iota(G_0)$ because the latter is normal.

As manifold we consider

$$G=G_0\times G_2 \tag{52.39}$$

and we define the multiplication law

$$\begin{aligned}\cdot&:G\times G\to G\\ (g_0,g_2)\cdot(g_0',g_2')&=\big(g_0R_{g_2}(g_0'),g_2g_2'\big).\end{aligned} \tag{52.40}$$

For associativity we have

$$(g_0,g_2)\cdot\big((g_0',g_2')\cdot(g_0'',g_2'')\big)=\Big(g_0R_{g_2}\big(g_0'R_{g_2'}(g_0'')\big),g_2g_2'g_2''\Big) \tag{52.41}$$

while

$$\big((g_0,g_2)\cdot(g_0',g_2')\big)\cdot(g_0'',g_2'')=\big(g_0R_{g_2}(g_0')R_{g_2g_2'}(g_0''),(g_2'g_2'')\big). \tag{52.42}$$

Thus the product is associative if and only if

$$g_0R_{g_2}\big(g_0'R_{g_2'}(g_0'')\big)=\big(g_0R_{g_2}(g_0')\big)R_{g_2g_2'}(g_0''). \tag{52.43}$$

That equality is in fact true because R is a morphism from G_2 to $\mathrm{Aut}(G_0)$, so that $R_{g_2}R_{g_2'} = R_{g_2g_2'}$.

The neutral in G is (e_0, e_2).

Since $R_{g_2}(g_0)$ is smooth with respect to both variables, the product is smooth. In that way, G becomes a Lie group named the **semi direct product** of G_2 by G_0 and is denoted by

$$G_0 \rtimes_R G_2. \tag{52.44}$$

All the construction is still valid when R is an homomorphism which does not comes from a split extension.

We define the product $G_0 \times G_2 \to G$ by

$$g_0 \cdot g_2 = (g_0, e_2) \cdot (e_0, g_2) \tag{52.45}$$

The diagram

$$\begin{array}{ccccc} & & G_1 & & \\ & \overset{\iota}{\nearrow} & \uparrow & \overset{\pi}{\searrow} & \\ G_0 & & \varphi & & G_2 \\ & \underset{\mathrm{Id}\times\{e\}}{\searrow} & & \underset{\mathtt{proj}_2}{\nearrow} & \\ & & G & & \end{array} \tag{52.46}$$

suggests us to define the map

$$\begin{aligned} \varphi\colon G_0 \times G_2 &\to G_1 \\ (g_0, g_2) &\mapsto \iota(g_0)j(g_2) \end{aligned} \tag{52.47}$$

This is a Lie group homomorphism because on the one hand

$$\varphi(g_0, g_2) \cdot \varphi(g_0', g_2') = \iota(g_0)j(g_2) \cdot \iota(g_0')j(g_2'), \tag{52.48}$$

while on the other hand

$$\begin{aligned} \varphi\big((g_0, g_2) \cdot (g_0', g_2')\big) &= \varphi\big(g_0R_{g_2}(g_0'), g_2g_2'\big) \\ &= \varphi\Big(g_0\iota^{-1}\big(\mathbf{Ad}(j(g_2))\iota(g_0')\big), g_2g_2'\Big) \\ &= \iota\Big(g_0\iota^{-1}\Big(\mathbf{Ad}(j(g_2))\iota(g_0')\Big)\Big)j(g_2g_2') \\ &= \iota(g_0)j(g_2)\iota(g_0')j(g_2') \end{aligned} \tag{52.49}$$

because ι and j are homomorphisms.

The Leibniz rule on $\iota(g_0)j(g_2)$ provides the differential

$$(d\varphi)_e = (d\iota)_{e_0} \oplus (dj)_{e_2}. \tag{52.50}$$

This is injective because j is injective. The kernel of φ is the set

$$\ker(\varphi) = \{(g_0, g_2) \text{ tel que } \iota(g_0) = j(g_2)^{-1}\}. \tag{52.51}$$

Since $\iota(G_0)$ and $j(G_2)$ have no intersections [8] (a part the identity), we have that the kernel reduces to the identity:

$$\ker(\varphi) = \{e\}. \tag{52.52}$$

The differentials provide the diagram

$$\mathcal{G}_0 \xrightarrow{(d\iota)_{e_0}} \mathcal{G}_1 \underset{(dj)_{e_2}}{\overset{(d\pi)_{e_1}}{\rightleftarrows}} \mathcal{G}_2. \tag{52.53}$$

We have $(d\pi)_{e_1} \circ (dj)_{e_2} = \mathrm{Id}\,|_{\mathcal{G}_2}$ and the map

$$(d\varphi)_e\colon \mathcal{G}_1 \to \mathcal{G}_0 \oplus \mathcal{G}_2 \tag{52.54}$$

is an algebra homomorphism (as differential of group homomorphism). It is also an isomorphism by dimension counting. The inverse theorem then shows that φ is a local diffeomorphism: $\varphi(G)$ contains a neighborhood of the identity and then is surjective by proposition 52.5.

We conclude that φ is a Lie group isomorphism.

8. They are transverse because $j \circ \pi = \mathrm{Id}\,|_{G_2}$.

Chapter 53

Lie group and Lie algebra

Here are the results which relate Lie groups and Lie algebras.

53.1 Lie algebra of a Lie group

DEFooSSDYooOwjHso

Lemma-Definition 53.1.
Let G be a smooth Lie group and $X \in T_eG$. The vector field X^L defined by

$$X_g^L = dL_X v \tag{53.1}$$

*is smooth. This is the **left invariant** vector field associated with the vector $X \in T_eG$.*

Definition 53.2.
*Let G be a group. The **left translation** by g on G is the map*

$$\begin{aligned} L_g \colon G &\to G \\ h &\mapsto gh. \end{aligned} \tag{53.2}$$

*The **right translation** by g on G is the map*

$$\begin{aligned} R_g \colon G &\to G \\ h &\mapsto hg. \end{aligned} \tag{53.3}$$

DEFooYHKXooVoJalX

Definition 53.3 ([873])**.**
*If G is a Lie group, a vector field $X \in \Gamma^\infty(TG)$ is **left invariant** if*

$$(dL_g)X = X, \tag{53.4}$$

which means that for every $g, h \in G$,

$$(dL_h)_g X_g = X_{hg}. \tag{53.5}$$

*In the same way, the vector field Y is **right invariant** if*

$$(dR_g)Y = Y. \tag{53.6}$$

PROPooLEIAooTnnYRw

Proposition-Definition 53.4.
Let G be a Lie group and $X \in T_eG$.

(1) The vector field X^L defined by

$$X_g^L = (dL_g)_e X. \tag{53.7}$$

DEFooYPUIooAzcdiP

is left-invariant.

(2) *The vector field X^L defined by*

$$X^R_g = (dR_g)_e X. \tag{53.8}$$

is right-invariant.

They are called the left and right-invariant vector fields ***associated with*** X.

DEFooKDCPooZOJsMD

Proposition-Definition 53.5.
Let G be a smooth Lie group. If $X, Y \in T_eG$ we define the bracket[1]

$$[X, Y] = [X^L, Y^L]_e. \tag{53.9}$$

The set T_eG with this bracket is a Lie algebra. This is the ***Lie algebra of the Lie group*** G. *It will usually be denoted by $\mathfrak{g}$.*

The topology on $\mathfrak{g} = T_eG$ is the usual one of the tangent spaces, definition 49.78.

Proof. We know from proposition 49.53 that T_eG is a vector space. We have to define a Lie bracket on it. For that we use the left-invariant vector field. Let $X \in T_eG$ and $g \in M$ we define

$$X^L_g = dL_g X \tag{53.10}$$

where $L_g \colon G \to G$ is the left translation: $L_g(h) = gh$. If $X, Y \in T_eG$ we define

$$[X, Y] = [X^L, Y^L]_e \tag{53.11}$$

where the bracket on the right hand side is the commutator of vector field defined in 49.105. It defines a Lie algebra structure by the proposition 51.5. □

In order to make the notations clear, let us write the formula explicitly. If $X, Y \in T_eG$ are given by $X = \alpha'(0)$ and $Y = \beta'(0)$ we have SUBEQSooHKWMooQbeStl

$$\begin{aligned}
(XY)f &= X(Y(f)) && \text{(53.12a)}\\
&= \frac{d}{dt}\Big[(Yf)\big(\alpha(t)\big)\Big]_{t=0} && \text{(53.12b)}\\
&= \frac{d}{dt}\Big[Y_{\alpha(t)}(f)\Big]_{t=0} && \text{(53.12c)}\\
&= \frac{d}{dt}\Big[Y^L_{\alpha(t)}(f)\Big]_{t=0} && \text{(53.12d)}\\
&= \frac{d}{dt}\frac{d}{ds}\Big[f\big(\alpha(t)\beta(u)\big)\Big]_{\substack{s=0\\t=0}}. && \text{(53.12e)}
\end{aligned}$$

Now a great theorem without proof:

tho:loc_isom

Theorem 53.6.
Two Lie groups are locally isomorphic if and only if their Lie algebras are isomorphic.

ThoSubGpSubAlg

Theorem 53.7.
If G is a Lie group, then

ThoSubGpSubAlgi

(1) *if $\mathfrak{h}$ is the Lie algebra of a Lie subgroup H of G, then it is a subalgebra of $\mathfrak{g}$,*

(2) *Any subalgebra of $\mathfrak{g}$ is the Lie algebra of one and only one connected Lie subgroup of G.*

¡¡ Avertissement/question à la lectrice !! 53.8
À mon avis, il faut dire "connexe et simplement connexe", et non juste "connexe".

1. On the right, this is the bracket of vector fields defined in 49.105.

Proof. First item. Let $i\colon H \to G$ be the identity map; it is a homomorphism from H to G, thus di_e is a homomorphism from $\mathfrak{h}$ to $\mathfrak{g}$. Conclusion: $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.

Characterization for $\mathfrak{h}$. Before to go on with the second point, we derive an important characterization of $\mathfrak{h}$:

$$\mathfrak{h} = \{X \in \mathfrak{g} : \text{the map } t \to \exp tX \text{ is a path in } H\}. \tag{53.13}$$

eq:path_alg

For that, consider $\exp_H\colon \mathfrak{h} \to H$ and $\exp_G\colon \mathfrak{g} \to G$; from unicity of the exponential, for any $X \in \mathfrak{h}$, $\exp_H X = \exp_G X$, so that one can simply write "exp" instead of "$\exp_h$" or "$\exp_G$".

Now, if $X \in \mathfrak{h}$, the map $t \to \exp tX$ is a curve in H. But it is not immediately clear that such a curve in H is automatically build from a vector in $\mathfrak{h}$ rather than in $\mathfrak{g}$. More precisely, consider a $X \in \mathfrak{g}$ such that $t \to \exp tX$ is a path (continuous curve) in H. By lemma 49.71, the map $t \to \exp tX$ is differentiable and thus by derivation, $X \in \mathfrak{h}$. The characterisation (53.13) is proved.

Thus $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.

Second item. For the second part, we consider $\mathfrak{h}$ any subalgebra of $\mathfrak{g}$ and H, the smallest subgroup of G which contains $\exp\mathfrak{h}$. We also consider a basis $\{X_1, \dots, X_n\}$ of $\mathfrak{g}$ such that $\{X_{r+1}, \dots, X_n\}$ is a basis of $\mathfrak{h}$.

By corollary 52.18, the set of linear combinations of elements of the form $X(M)$ with $M = (0, \dots, 0, m_{r+1}, \dots, m_r)$ form a subalgebra of $U(\mathfrak{g})$. If $X = x_1X_1 + \cdots + x_nX_n$, we define $|X| = (x_1^2 + \cdots + x_n^2)^{1/2}$ $(x_i \in \mathbb{R})$.

Let us consider a $\delta > 0$ such that exp is a diffeomorphism (normal neighbourhood) from $B_\delta = \{X \in \mathfrak{g} : |X| < \delta\}$ to a neighbourhood N_e of $e \in G$ and such that $\forall x, y, xy \in N_e$,

$$(xy)_k = \sum_{M,N} C_{MN}^{[k]} x^M y^N \tag{53.14}$$

eq:coord_xy

holds[2]. We note $V = \exp(\mathfrak{h} \cap B_\delta) \subset N_e$. The map

$$\exp(x_{r+1}X_{r+1} + \cdots + x_nX_n) \to (x_{r+1}, \dots, x_n)$$

is a coordinate system on V for which V is a connected manifold. But $\mathfrak{h} \cap B_\delta$ is a submanifold of B_δ, then V is a submanifold of N_e and consequently of G.

Let x, $y \in V$ such that $xy \in N_e$ (this exist: $x = y = e$); the canonical coordinates of xy are given by (53.14). Since $x_k = y_k = 0$ for $1 \leqslant k \leqslant r$, $(xy)_k = 0$ for the same k because for $(xy)_k$ to be non zero, one need $m_1 = \ldots = m_r = n_1 = \ldots = n_r = 0$ – otherwise, x^M or y^N is zero. Now we looks at $C_{MN}^{[k]}$ for such a k (say $k = 1$ to fix ideas): $[k] = (\delta_{11}, \dots, \delta_{1k}) = (1, 0, \dots, 0)$ and by definition of the C's,

$$X(M)X(N) = \sum_P C_{MN}^P X(P).$$

But we had seen that the set of the $X(A)$ with $A = (0, \dots, 0, a_{r+1}, \dots, a_n)$ form a subalgebra of $U(\mathfrak{g})$. Then, only terms with $P = (0, \dots, 0, p_{r+1}, \dots, p_n)$ are present in the sum; in particular, $C_{MN}^{[k]} = 0$ for $k = 1, \dots, r$. Thus $VV \cap N_e \subset V$.

The next step is to consider $\mathcal{V}$, the set of all the subset of H whose contains a neighbourhood of e in V. We can check that this fulfils the six axioms of a topological group:

(1) The intersection of two elements of $\mathcal{V}$ is in $\mathcal{V}$;

(2) the intersection of all the elements of $\mathcal{V}$ is $\{e\}$;

(3) any subset of H which contains a set of $\mathcal{V}$ is in $\mathcal{V}$;

(4) If $\mathcal{U} \in \mathcal{V}$, there exists a $\mathcal{U}_1 \in \mathcal{V}$ such that $\mathcal{U}_1\mathcal{U}_1 \subset \mathcal{U}$ because $VV \cap N_e \subset V$;

(5) if $\mathcal{U} \in \mathcal{V}$, then $\mathcal{U}^{-1} \in \mathcal{V}$ because the inverse map is differentiable and transforms a neighbourhood of e into a neighbourhood of e;

(6) if $\mathcal{U} \in \mathcal{V}$ and $h \in H$, then $h\mathcal{U}h^{-1} \in \mathcal{V}$.

2. The validity of this second condition is assured during the proof of theorem 53.6 which is not given here.

To see this last item, we denote by log the inverse map of $\exp\colon B_\delta \to N_e$. By definition of V, it sends V on $\mathfrak{h} \cap B_\delta$. If $X \in \mathfrak{g}$, there exists one and only one $X' \in \mathfrak{g}$ such that $he^{tX}h^{-1} = e^{tX'}$ for any $t \in \mathbb{R}$. Indeed we know that $he^X h^{-1} = e^{\mathrm{Ad}_h X}$, then X' must satisfy $e^{tX'} = e^{\mathrm{Ad}_h tX}$. If it is true for any t, then, by derivation, $X' = \mathrm{Ad}_h X$.

The map $X \to X'$ is an automorphism of $\mathfrak{g}$ which sent $\mathfrak{h}$ on itself. So one can find a δ_1 with $0 < \delta_1 < \delta$ such that

$$h \exp(B_{\delta_1} \cap \mathfrak{h})h^{-1} \subset V.$$

Indeed, $he^{\mathfrak{h}}h^{-1} \subset \mathfrak{h}$, so that taking $\delta_1 < \delta$, we get the strict inclusion. We can choose δ_1 even smaller to satisfy $he^{B_{\delta_1}}h^{-1} \subset N_e$. Since the map $X \to \log(he^X h^{-1})$ from $B_{\delta_1 \cap \mathfrak{h}}$ to $B_\delta \cap \mathfrak{h}$ is regular, the image of $B_{\delta_1} \cap \mathfrak{h}$ is a neighbourhood of 0 in $\mathfrak{h}$. Thus $he^{B_{\delta_1} \cap \mathfrak{h}}h^{-1}$ is a neighbourhood of e in V. Finally, $h\mathcal{U}h^{-1} \in \mathcal{V}$ and the last axiom of a topological group is checked.

This is important because there exists a topology on H such that H becomes a topological group and $\mathcal{V}$ is a family of neighbourhood of e in H. In particular, V is a neighbourhood of e in H.

For any $z \in G$, we define the map $\phi_z\colon zN_e \to B_\delta$ by

$$\phi_z(ze^{x_1X_1+\cdots+x_nX_n}) = (x_1, \ldots, x_n), \tag{53.15}$$

and we denote by φ_z the restriction of ϕ_z to zV. If $z \in H$, then φ_z sends the neighbourhood zV of z in H to the open set $B_\delta \cap \mathfrak{h}$ in $\mathbb{R}^{n-r}$. Indeed, an element of zV is a ze^Z with $Z \in \mathfrak{h} \cap B_\delta$ which is sent by φ_z to an element of $\mathfrak{h} \cap B_\delta$. (we just have to identify $x_1X_1 + \cdots + x_nX_n$ with $(x_1, \ldots, x_n)$).

Moreover, if $z_1, z_2 \in H$, the map $\varphi_{z_1} \circ \varphi_{z_2}^{-1}$ is the restriction to an open subset of $\mathfrak{h}$ of $\phi_{z_1} \circ \phi_{z_2}$. Then $\varphi_{z_1} \circ \varphi_{z_2}^{-1}$ is differentiable. Conclusion: $(H, \varphi_z : z \in H)$ is a differentiable manifold.

Recall that the definition of $\mathfrak{h}$ was to be a subalgebra of $\mathfrak{g}$; therefore $V = e^{\mathfrak{h} \cap B_\delta}$ is a submanifold of G. But the left translations are diffeomorphism of H and H is the smallest subgroup of G containing $e^{\mathfrak{h}}$. Thus H is a manifold on which the multiplication is diffeomorphic and consequently, H is a Lie subgroup of G.

Rest to prove that the Lie algebra of H is $\mathfrak{h}$ and the unicity part of the theorem.

We know that $\dim H = \dim \mathfrak{h}$ and moreover for $i > r$, the map $t \to \exp tX_i$ is a curve in H. Now, the fact that $\mathfrak{h}$ is the set of $X \in \mathfrak{g}$ such that $t \to \exp tX$ is a path in H show that $X_i \in \mathfrak{h}$. Then the Lie algebra of H is $\mathfrak{h}$ and H is a connected group because it is generated by $\exp \mathfrak{h}$ which is a connected neighbourhood of e in H.

We turn our attention to the unicity part. Let H_1 be a connected Lie subgroup of G such that $T_eH_1 = \mathfrak{h}$. Since $\exp_{\mathfrak{h}} X = \exp_{\mathfrak{h}_1} X$, $H = H_1$ as set. But exp is a differentiable diffeomorphism from a neighbourhood of 0 in $\mathfrak{h}$ to a neighbourhood of e in H and H_1, so as Lie groups, H and H_1 are the same.

Let us consider an element $X \in \mathfrak{g}$ such that $\exp tX \in H$ for every $t \in \mathbb{R}$, and the map $\varphi\colon \mathbb{R} \to G$, $\varphi(t) = \exp tX$. This is continuous, then there exists a connected neighbourhood $\mathcal{U}$ of 0 in $\mathbb{R}$ such that $\varphi(\mathcal{U}) \subset V$. Then $\varphi(\mathcal{U}) \subset H \cap V$ and the connectedness of $\varphi(\mathcal{U})$ makes $\varphi(\mathcal{U}) \subset \exp \mathcal{U}_h$. But $\exp \mathcal{U}_h$ is an arbitrary small neighbourhood of e in H; the conclusion is that φ is a continuous map from $\mathbb{R}$ into H. Indeed, we had chosen X such that $\exp tX \in H$.

Moreover, we know that

$$e^{(t_0+\epsilon)X} = e^{t_0X}e^{\epsilon X},$$

but $\exp \epsilon X$ can be as close to e as we want (this proves the continuity at t_0). Then φ is a path in H.

In definitive, we had shown that $\exp tX \in H$ implies that $t \to \exp tX$ is a path. Now equation (53.13) gives the result. □

Corollary 53.9.
Let G be a Lie group and H_1, H_2, two subgroups both having a finite number of connected components (each for his own topology). If $H_1 = H_2$ as sets, then $H_1 = H_2$ as Lie groups.

Proof. The proposition shows that H_1 and H_2 have same Lie algebra. But any Lie subalgebra of $\mathfrak{g}$ is the Lie algebra of exactly one connected subgroup of G (theorem 53.7). Then as Lie groups, $H_{10} = H_{20}$. Since H_1 and H_2 are topological groups, the equality of they topology on one connected component gives the equality everywhere (because translations are differentiable). □

pg:ex_topo_Lie

Consider the group $T = S^1 \times S^1$ and the continuous map $\gamma\colon \mathbb{R} \to T$ given by

$$\gamma(t) = (e^{it}, e^{i\alpha t})$$

with a certain irrational α in such a manner that γ is injective and $\Gamma = \gamma(\mathbb{R})$ is dense in T.

The subset Γ is not closed because his complementary in T is not open: any neighbourhood of element $p \in T$ which don't lie in Γ contains some elements of Γ. We will show that the inclusion map $\iota\colon \Gamma \to T$ is continuous. An open subset of T is somethings like

$$\mathcal{O} = (e^{iU}, e^{iV})$$

where U, V are open subsets of $\mathbb{R}$. It is clear that

$$\iota^{-1}(\mathcal{O}) = \{\gamma(t) \text{ tel que } t \in U + 2k\pi, \alpha t \in V + 2m\pi\},$$

but the set of elements t of $\mathbb{R}$ which satisfies it is clearly open. Then Γ has at least the induced topology from T (as shown in proposition 49.113). In fact, the own topology of Γ is *more* than the induced: the open subsets of Γ whose are just some small segments clearly doesn't appear in the induced topology. Thus the present case is an example (and not a counter-example) of theorem 53.113.

This example show the importance of the condition for a topological subspace to have *exactly* the induced topology. If not, any Lie subgroup were a topological Lie subgroup because a submanifold has at least the induced topology. We will go further with this example after the proof.

53.2 Matrix Lie group and its algebra

SECooTSAJooNtjgMD

In this section we deal with Lie groups made from matrices, that is subgroups of $\mathrm{GL}(n, \mathbb{C})$ (typically $\mathrm{SO}(n)$ or $\mathrm{SU}(n)$) and their Lie algebra. We will denote the identity either by e or by $\mathbb{1}$.

NORMooHZGKooJEiamo

53.10.
It is time to reread the remark 49.51. In this section, when γ is a path in the matrix group G, we denote by $\gamma'(0)$ the "usual" derivative of γ: that is the component-wise derivative; not the differential operator.

We denote by D_γ the differential operator

$$\begin{aligned} D_\gamma\colon C^\infty(G) &\to \mathbb{R} \\ f &\mapsto \frac{d}{dt}\Big[f\big(\gamma(t)\big)\Big]_{t=0}. \end{aligned} \tag{53.16}$$

We aim to study the link between D_γ and $\gamma'(0)$.

From the Lie group of matrix G we can build (at least) two Lie algebras[3]:

— The usual Lie algebra of the group: T_eG with the definition 53.5. As set, this is

$$T_eG = \{D_\gamma \text{ such that } \gamma(0) = e\} \tag{53.17}$$

with the implicit that γ is a smooth path in G.

3. Definition 51.1.

— The set of "usual" derivatives of the paths in G:

$$G' = \{\gamma'(0) \text{ tel que } \gamma(0) = e\}. \tag{53.18}$$

This is a set of matrices on which we can use the bracket $[X, Y] = XY - YX$ (matrix product). We will see the following facts.

— The set G' is a Lie algebra in proposition 53.13,

— The Lie algebras G' and T_eG are isomorphic as Lie algebras in theorem 53.15 for the case $G = \mathrm{GL}(n, \mathbb{C})$

— When H is a Lie subgroup of $\mathrm{GL}(n, \mathbb{C})$, the Lie algebras H' and T_eH are isomorphic as Lie algebras in proposition 53.16 for the Lie subgroups of $\mathrm{GL}(n, \mathbb{C})$.

Lemma 53.11 ([1])**.**
Let G be a matrix Lie group, et $g \in G$ and $X \in G'$. Then $gXg^{-1} \in G'$.

Proof. Let $x\colon \mathbb{R} \to G$ be a smooth path such that $X = x'(0)$. Then we the derivative of the path given by the matrix product

$$t \mapsto gx(t)g^{-1} \tag{53.19}$$

is gXg^{-1}. □

LEMooHQUYooSoiKbI

Lemma 53.12 ([1])**.**
Let G be a matrix Lie group. Then G' is a vector space on $\mathbb{R}$.

Proof. Let $X, Y \in G'$ be the derivatives of the paths x and y. If we set $\varphi_1(t) = x(t)y(t)$ we have

$$\varphi_1'(0) = x'(0)y(0) + x(0)y'(0). \tag{53.20}$$

Since $x(0) = y(0) = e$ we have $\varphi'(0) = X + Y$, so that $X + Y \in G'$.

For the product by a scalar, let the path $\varphi_2(t) = x(\lambda t)$. The component-wise derivative

$$\varphi_2'(0) = \lambda x'(0) = \lambda X, \tag{53.21}$$

so that $\lambda X \in G'$. □

PROPooUKITooLnEKZW

Proposition 53.13.
Let G be a matrix Lie group. The vector space G' is a Lie algebra for the matrix commutator.

Proof. We already know that G' is a real vector space by lemma 53.12. The fact that $(X, Y) \mapsto XY - YX$ satisfies the axioms of a Lie algebra is easy to check. The only point is to show that if $X, Y \in G'$, then $[X, Y] = XY - YX \in G'$.

Let

$$\varphi(t) = x(t)Yx(-t). \tag{53.22}$$

EQooJDTLooGWsDig

This is for sure a path in the full matrix vector space, and this is derivable because x is derivable while the matrix product is linear. So the derivative $\varphi'(0)$ is still a matrix. The question is: why $\varphi'(0) \in G'$?

By lemma 53.12, for each t we have

$$\frac{\varphi(t) - \varphi(0)}{t} \in G'. \tag{53.23}$$

Now, G' is a vector subspace of $\mathbb{M}(n, \mathbb{C})$ which is finite dimensional; is is thus closed and the limit belongs to G'.

Is is now a simple computation to show that $\varphi'(0) = [X, Y]$. □

53.14.
The following theorem is a Giulietta's masterpiece in the following sense:

— It is fundamental because the Lie algebra isomorphism between $T_eGL(n,\mathbb{R})$ and the matrices is used everywhere one says «The Lie algebra of SO(3) is the set of skew-symmetric traceless matrices».

— Either I'm idiot, either I never seen that theorem even stated (let alone being proved) [4].

— I think that the fundamental misunderstanding [5] is that in the context of Lie groups, people *define* $[X,Y]$ as being $\mathrm{ad}(X)Y$ while ad is defined as the "second differential" of $\mathbf{Ad}(g)h = ghg^{-1}$. In that case, obviously we get $[X,Y] = XY - YX$ with the matrix product. This way fails to make the link with the commutator of vector fields as defined by 49.105.

— So you must read this proof with much care and write me if you see any mistake or unclear point.

So here it is with the notations explained in 53.10.

THOooWQGMooHyjRtx

Theorem 53.15.
Let $G = \mathrm{GL}(n,\mathbb{C})$ be the group of invertible matrices. The map

$$\begin{aligned}\phi\colon G' &\to T_eG\\ \gamma'(0) &\mapsto D_\gamma\end{aligned} \tag{53.24}$$

is

(1) well defined,

(2) bijective,

(3) linear,

(4) a Lie algebra isomorphism.

Proof. Several points to be proved.

(1) ϕ **is well defined** Let α and β be paths in G such that $\alpha'(0) = \beta'(0)$ and let $f\colon G \to \mathbb{R}$ be a smooth function. We have to prove that $D_\alpha(f) = D_\beta(f)$.

We consider a chart $\varphi\colon \mathcal{U} \to \mathcal{O}$ where $\mathcal{U}$ is a neighbourhood of 0 in $\mathbb{R}^m$ and $\mathcal{O}$ is a neighbourhood of e in $\mathrm{GL}(n,\mathbb{C})$. We suppose that $\varphi(0) = e$. We set $\tilde{f} = f \circ \varphi$, $\tilde{\alpha} = \varphi^{-1} \circ \alpha$ and $\tilde{\beta} = \varphi^{-1} \circ \beta$. We have

$$\begin{aligned}D_\alpha(f) &= \frac{d}{dt}\Big[f\big(\alpha(t)\big)\Big]_{t=0} && \text{(53.25a)}\\ &= \frac{d}{dt}\Big[\tilde{f}\big(\tilde{\alpha}(t)\big)\Big]_{t=0} && \text{(53.25b)}\\ &= \sum_{i=1}^{m}\frac{\partial\tilde{f}}{\partial x_i}\big(\tilde{\alpha}(0)\big)\tilde{\alpha}_i(0). && \text{(53.25c)}\end{aligned}$$

Since $\tilde{\alpha}(0) = \tilde{\beta}(0)$ we still have to prove that $\tilde{\alpha}_i'(0) = \tilde{\beta}_i'(0)$. As you remember, $\tilde{\alpha}$ is a map from $\mathbb{R}$ to $\mathbb{R}^m$, so that the following derivative is quite usual:

$$\begin{aligned}\tilde{\alpha}'(0) &= \frac{d}{dt}\Big[(\varphi^{-1}\circ\alpha)(t)\Big]_{t=0} && \text{(53.26a)}\\ &= d\varphi^{-1}_{\alpha(0)}\big(\alpha'(0)\big) && \text{(53.26b)}\\ &= d\varphi^{-1}_{\beta(0)}\big(\beta'(0)\big). && \text{(53.26c)}\end{aligned}$$

Thus the map ϕ is well defined.

(2) ϕ **is linear** This is from the linearity of the derivation.

4. There is in fact a third possibility: this theorem is a classic one but cannot be found *on internet*.
5. Once again, either I'm idiot either everybody is wrong but me... well ...

(3) **ϕ is injective** If $\phi(\alpha') = \phi(\beta')$, then $D_\alpha(f) = D_\beta(f)$ for every function f. In that case,

$$\sum_{i=1}^{m} \frac{\partial \tilde{f}}{\partial x_i}(e)\tilde{\alpha}_i'(0) = \sum_{i=1}^{m} \frac{\partial \tilde{f}}{\partial x_i}(e)\tilde{\beta}_i'(0). \tag{53.27}$$

That equation must be satisfied for every function. Taking the projection on the components, we get $\tilde{\alpha}_i'(0) = \tilde{b}_i'(0)$, which means $\alpha'(0) = \beta'(0)$ because φ^{-1} is bijective.

(4) **ϕ is surjective** Every element of T_eG is of the form D_α for some path α, so ϕ is surjective.

(5) **ϕ is a Lie algebra isomorphism** Let $X, Y \in G'$ being the derivative of the paths α and β. We have to prove that

$$[\phi(X), \phi(Y)] = \phi[X, Y]. \tag{53.28}$$

If t is small enough, the paths

$$\alpha(t) = \mathbb{1} + tX \tag{53.29a}$$
$$\beta(t) = \mathbb{1} + tY \tag{53.29b}$$
$$\tag{53.29c}$$

are good ones because $\det(\mathbb{1}) \neq 0$, so that the determinant of $\mathbb{1} + tX$ remains different from zero when t is small, whatever X is. So α and β are paths in $\mathrm{GL}(n, \mathbb{C})$. Using the general definition in differential geometry, SUBEQSooCYRDooFOdLrn

$$[\phi(X), \phi(Y)]f = [\phi(X)^L, \phi(Y)^L]_e f \tag{53.30a}$$
$$= \phi(X)^L_e\big(\phi(Y)^L(f)\big) - \phi(Y)^L_e\big(\phi(X)^L(f)\big). \tag{53.30b}$$ SUBEQooOPUAooZYsZlX

We focus on the first term: SUBEQooTUNFooFkDmuP

$$\phi(X)^L\big(\phi(Y)^L(f)\big) = \frac{d}{dt}\Big[\phi(Y)^L_{\phi(X)^L_e(t)}(f)\Big]_{t=0} \tag{53.31a}$$
$$= \frac{d}{dt}\frac{d}{ds}\Big[f\big((\mathbb{1} + tX)(\mathbb{1} + sY)\big)\Big]_{\substack{s=0\\t=0}} \tag{53.31b}$$
$$= \frac{d}{dt}\frac{d}{ds}\Big[f(\mathbb{1} + tX + sY + tsXY)\Big]_{\substack{s=0\\t=0}} \tag{53.31c}$$
$$= \frac{d}{dt}\Big[df_{\mathbb{1}+tX}\big((\mathbb{1} + tX)Y\big)\Big]_{t=0} \tag{53.31d}$$ SUBEQooLHPBooTnXiZd
$$= \frac{d}{dt}\Big[df_{\mathbb{1}+tX}(Y)\Big]_{t=0} + \frac{d}{dt}\Big[df_{\mathbb{1}+tX}(tXY)\Big]_{t=0} \tag{53.31e}$$ SUBEQooMXJJooBFTLsM

where we have used the linearity of $df_{\mathbb{1}+tX}$ and where XY stands for the matrix product. In the expression (53.31d), the symbol df stands for the differential of f as function from $\mathbb{M}(n, \mathbb{C})$ (as vector space), not for the differential of f on G as manifold. This is why we are allowed to put an expression as the matrix Y as argument of $df_{\mathbb{1}+tX}$ while Y is not an element of $T_{\mathbb{1}+tX}G$.

The expression (53.31e) is still made of two terms. The second one is

$$\frac{d}{dt}\Big[df_{\mathbb{1}+tX}(tXY)\Big]_{t=0} = \frac{d}{dt}\Big[tdf_{\mathbb{1}+tX}(XY)\Big]_{t=0} = df_{\mathbb{1}}(XY) \tag{53.32}$$

where we used the Leibniz rule [6].

The first term in (53.31e) is computed as

$$\frac{d}{dt}\Big[df_{\mathbb{1}+tX}(Y)\Big]_{t=0} = \frac{d}{dt}\frac{d}{ds}\Big[f(\mathbb{1} + tX + sY)\Big]_{\substack{s=0\\t=0}}. \tag{53.33}$$

6. In general, notice that $\frac{d}{dt}\Big[tf(t)\Big]_{t=0} = f(0)$

We set

$$\begin{aligned}\gamma\colon \mathbb{R}^2 &\to G\\ (t,s) &\mapsto \mathbb{1}+tX+sY,\end{aligned} \tag{53.34}$$

so that

$$\frac{d}{dt}\Big[df_{\mathbb{1}+tX}(Y)\Big]_{t=0} = \frac{d}{dt}\frac{d}{ds}\Big[f(\mathbb{1}+tX+sY)\Big]_{\substack{s=0\\t=0}} \tag{53.35a}$$

$$= \frac{d}{dt}\frac{d}{ds}\Big[(\tilde{f}\circ\varphi^{-1}\circ\gamma)(t,s)\Big]_{\substack{s=0\\t=0}} \tag{53.35b}$$

$$= \frac{d}{dt}\frac{d}{ds}\Big[g(t,s)\Big]_{\substack{s=0\\t=0}} \tag{53.35c}$$

where the function $g=\tilde{f}\circ\varphi^{-1}\circ\gamma$ is a smooth function from $\mathbb{R}^2$ to $\mathbb{R}$.
The expression (53.31) is now

$$\phi(X)^L\big(\phi(Y)^L(f)\big) = \frac{d}{dt}\frac{d}{ds}\Big[g(t,s)\Big]_{\substack{s=0\\t=0}} + df_{\mathbb{1}}(XY). \tag{53.36}$$

The commutator we have to compute, with the same computations is

$$[\phi(X),\phi(Y)]f = \frac{d}{dt}\frac{d}{ds}\Big[g(t,s)\Big]_{\substack{s=0\\t=0}} + df_{\mathbb{1}}(XY) - \frac{d}{dt}\frac{d}{ds}\Big[g(s,t)\Big]_{\substack{s=0\\t=0}} - df_{\mathbb{1}}(YX). \tag{53.37}$$

The function g being C^∞, the derivative commute and the corresponding termes annihilate each other and we are left with

$$[\phi(X),\phi(Y)]f = df_{\mathbb{1}}(XY) - df_{\mathbb{1}}(YX) = df_{\mathbb{1}}(XY-YX) \tag{53.38}$$

where we used the linearity of the differential.
In the other sense,

$$\phi[X,Y]f = \frac{d}{dt}\Big[f(\mathbb{1}+tXY-tYX)\Big]_{t=0} = df_{\mathbb{1}}\big([X,Y]\big) \tag{53.39}$$

where, once again, df stands for the "usual" differential.

□

Ok. This is proved for $G=\mathrm{GL}(n,\mathbb{C})$, the full matrix group. What about subgroups ? Here is the result.

PROPooSQHLooGQAykc

Proposition 53.16 ([1])**.**
Let H be a closed Lie subgroup of $\mathrm{GL}(n,\mathbb{C})$. With the same notations as above, the map

$$\begin{aligned}\phi\colon H' &\to T_eH\\ \gamma'(0) &\mapsto D_\gamma\end{aligned} \tag{53.40}$$

is a Lie algebra isomorphism.

Proof. We have to prove that

$$\phi[X,Y]f = [\phi(X),\phi(Y)]f \tag{53.41}$$

EQooRLBBooYgHhtH

for every $X,Y\in H'$ and $f\in C^\infty(H)$. For that, we will see the left and right hand sides of (53.41) in $G=\mathrm{GL}(n,\mathbb{C})$, and use the already proved result, theorem 53.15.

If $X,Y\in H'$ we know from proposition 53.13 that $[X,Y]\in H'$. Thus there exists a path $\gamma\colon\mathbb{R}\to H$ such that $[X,Y]=\gamma'(0)$. We consider the extension [7] $\tilde{f}\colon W\to\mathbb{R}$ of f such that $\tilde{f}=f$

7. The proposition 49.115 can be used since H is a submanifold of G by 52.27.

on H and W is an open set around e in $\mathrm{GL}(n,\mathbb{C})$. For the sake of making things complicated we also define $\tilde{\gamma} = \iota \circ \gamma$ where $\iota\colon H \to \mathrm{GL}(n,\mathbb{C})$ is the inclusion. With all that we have

$$\phi[X,Y]f = \frac{d}{dt}\Big[f\big(\gamma(t)\big)\Big]_{t=0} = \frac{d}{dt}\Big[\tilde{f}\big(\tilde{\gamma}(t)\big)\Big]_{t=0} = \clubsuit. \tag{53.42}$$

At this point, notice that $[X,Y] \in \mathrm{GL}(n,\mathbb{C})'$ and $[X,Y] = \tilde{\gamma}'(0)$, so that if we consider the map $\tilde{\phi}\colon \mathrm{GL}(n,\mathbb{C}) \to T_e\mathrm{GL}(n,\mathbb{C})$ we also have

$$\clubsuit = \frac{d}{dt}\Big[\tilde{f}\big(\tilde{\gamma}(t)\big)\Big]_{t=0} = \tilde{\phi}[X,Y]\tilde{f} = \big[\tilde{\phi}(X),\tilde{\phi}(Y)\big]\tilde{f} \tag{53.43}$$

where we used the result 53.15 on $\mathrm{GL}(n,\mathbb{C})$.

We still have to prove that $\tilde{\phi}(X)\tilde{\phi}(Y)\tilde{f} = \phi(X)\phi(Y)f$. Using, among others the formula 53.12 adapted to $\tilde{\phi}(X)$ instead of X:

$$\begin{aligned}
\tilde{\phi}(X)\tilde{\phi}(Y)\tilde{f} &= \frac{d}{dt}\Big[\big(\tilde{\phi}(Y)^L\tilde{f}\big)\big(\alpha(t)\big)\Big]_{t=0} && (53.44a)\\
&= \frac{d}{dt}\Big[\tilde{\phi}(Y)^L_{\alpha(t)}\tilde{f}\Big]_{t=0} && (53.44b)\\
&= \frac{d}{dt}\frac{d}{ds}\Big[\tilde{f}\big(\alpha(t)\beta(u)\big)\Big]_{\substack{s=0\\t=0}}. && (53.44c)
\end{aligned}$$

At this point, notice that $\alpha(t)$ and $\beta(u)$ are elements in H which is a group, so $\tilde{f}\big(\alpha(t)\beta(u)\big) = f\big(\alpha(t)\beta(u)\big)$. Thus

$$\begin{aligned}
\tilde{\phi}(X)\tilde{\phi}(Y)\tilde{f} &= \frac{d}{dt}\frac{d}{ds}\Big[\tilde{f}\big(\alpha(t)\beta(u)\big)\Big]_{\substack{s=0\\t=0}} && (53.45a)\\
&= \frac{d}{dt}\frac{d}{ds}\Big[f\big(\alpha(t)\beta(u)\big)\Big]_{\substack{s=0\\t=0}} && (53.45b)\\
&= \phi(X)\phi(y)f. && (53.45c)
\end{aligned}$$

□

Lemma 53.17 ([1])**.**
Let G be a Lie group of matrices and $X \in T_eG$ such that

$$df_e(X) = 0 \tag{53.46}$$

for every smooth function $f\colon G \to \mathbb{R}$. Then $X = 0$.

Proof. We consider the functions $\mathtt{proj}_{ij}\colon G \to \mathbb{R}$ defined by $\mathtt{proj}_{ij}(A) = A_{ij}$. If $g\colon \mathbb{R} \to G$ is a path, for every t we have $\mathtt{proj}_{ij}\, g(t) = g(t)_{ij}$ and then

$$\frac{d}{dt}\Big[\mathtt{proj}_{ij}\, g(t)\Big]_{t=0} = g'(0)_{ij}. \tag{53.47}$$

Then we build

$$\begin{aligned}
f\colon G &\to \mathbb{R}\\
A &\mapsto \mathtt{proj}_{11}(A)\,\mathtt{proj}_{ij}(A).
\end{aligned} \tag{53.48}$$

If $g\colon \mathbb{R} \to G$ is a path such that $g(0) = e$ and $g'(0) = X$, then we have

$$\begin{aligned}
\frac{d}{dt}\Big[f\big(g(t)\big)\Big]_{t=0} &= \frac{d}{dt}\Big[\mathtt{proj}_{11}\big(g(t)\big)\,\mathtt{proj}_{ij}\big(g(t)\big)\Big]_{t=0} && (53.49a)\\
&= \mathtt{proj}_{11}\, g(0)\frac{d}{dt}\Big[\mathtt{proj}_{ij}\, g(t)\Big]_{t=0} + \frac{d}{dt}\Big[\mathtt{proj}_{11}\, g(t)\Big]_{t=0}\mathtt{proj}_{ij}\, g(0) && (53.49b)\\
&= X_{ij} + \delta_{ij}X_{11} && (53.49c)\\
&= X_{ij} + \delta_{ij}X_{11}. && (53.49d)
\end{aligned}$$

We know that this is zero for every choice of ij:

$$X_{ij} + \delta_{ij} X_{11} = 0 \tag{53.50}$$

In particular with $i = j = 1$ we have $2X_{11} = 0$, so that $X_{11} = 0$. Then we are left with $X_{ij} = 0$ for every ij. □

53.3 Fundamental vector field

sec:fond_vec

Definition 53.18.
If $\mathcal{G}$ is the Lie algebra[8] *of a Lie group G acting on a manifold M (the action of g on x being denoted by $x \cdot g$), the **fundamental vector field** associated with $A \in \mathcal{G}$ is given by*

$$A_x^* = \frac{d}{dt}\Big[x \cdot e^{-tA}\Big]_{t=0}. \tag{53.51}$$

EqDefChmpFond

If the action of G is transitive, the fundamental vectors at point $x \in M$ form a basis of T_xM. More precisely, we have the

Lemma 53.19.
For any $v \in T_xM$, there exists a $A \in \mathcal{G}$ such that $v = A_x^*$, in other terms

$$\text{Span}\{A_x^* \text{ tel que } A \in \mathcal{G}\} = T_xM.$$

LemFundSpansTan

Proof. The vector v is given by a path $v(t)$ in M. Since the action is transitive, one can write $v(t) = x \cdot c(t)$ for a certain path c in G which fulfills $c(0) = e$. We have to show that v depends only on $c'(0) \in \mathcal{G}$. We consider

$$\begin{aligned} R\colon G \times M &\to M \\ R(g, x) &= x \cdot g, \end{aligned} \tag{53.52}$$

eq_def_RGM

so

$$v = \frac{d}{dt}\Big[R(c(t), x)\Big]_{t=0} = dR_{(e,x)}\big[(d_t c(t), x) + (c(0), x)\big]. \tag{53.53}$$

eqv_Rc

□

lem:As_Bs_A_B

Lemma 53.20.
If A, $B \in \mathcal{G}$ are such that $A^* = B^*$, and if the action is effective, then $A = B$.

Proof. We consider once again the map (53.52) and we look at

$$v = \frac{d}{dt}\Big[R(c(t), x)\Big]_{t=0} = (dR)_{(e,x)} \frac{d}{dt}\Big[(c(t), x)\Big]_{t=0},$$

keeping in mind that $c(t) = e^{-tA}$. In order to treat this expression, we define

$$R_1\colon G \to M, \quad R_1(h) = R(h, x), \tag{53.54a}$$
$$R_2\colon M \to M, \quad R_2(y) = R(g, y). \tag{53.54b}$$

So

$$v = dR_1(X) + dR_2(0) = dR_1 c'(0)$$

and the assumption $A_x^* = B_x^*$ becomes $dR_1 A = dR_1 B$. This makes, for small enough t,

$$R_1(e^{tA}e^{-tB}) = x \cdot e^{tA}e^{-tB} = x; \tag{53.55}$$

if the action is effective, it imposes $A = B$. □

8. Lie algebra of a Lie group, definition 53.5.

Lemma 53.21.
If we consider the action of a matrix group, R_g acts on the fundamental field by

$$dR_g(A^*_\xi) = \big(\operatorname{Ad}(g^{-1})A\big)^*_{\xi \cdot g}.$$

lem:dRgAstar

Proof. Just notice that $e^{-t\operatorname{Ad}(g^{-1})A} = \mathbf{Ad}_{g^{-1}}(e^{-tA}) = g^{-1}e^{-tA}g$, thus

$$\big(\operatorname{Ad}(g^{-1})A\big)^*_{\xi \cdot g} = \frac{d}{dt}\Big[\xi \cdot ge^{-t\operatorname{Ad}(g^{-1})A}\Big]_{t=0} = dR_g(A^*_\xi). \tag{53.56}$$

□

53.4 Invariant vector fields

LEMooWTNRooCjlYMJ

Lemma 53.22 ([870]).
Let G be a Lie group and $X_1, \dots, X_k$ be linearly independent vectors in its Lie algebra $\mathfrak{g}$. For every $g \in G$, the vectors[9] *$X_i^L(g)$ are linearly independent in T_gG.*

Proof. Let λ_i be reals such that $\sum_{i=1}^k \lambda_i X_i^L(g) = 0$; in other terms,

$$0 = (dL_g)_e\big(\sum_i \lambda_i X_i\big). \tag{53.57}$$

We apply $(dL_{g^{-1}})_e$ on both side: $0 = \sum_i \lambda_i X_i$. By hypothesis, $\lambda_i = 0$ for every i. This proves that the vectors $X_i^L(g)$ are linearly independent. □

Theorem 53.23 ([873]).
The map $\varphi\colon X \mapsto X^L$ where $X^L_g = (dL_g)_e X$ is a bijection from T_eG to the set of left-invariant vector fields.

Proof. Two parts.

(1) **Surjective** Let X be a left-invariant vector field. We have $X = (X_e)^L$ because

$$(X_e)^L_g = (dL_g)X_e = X_g. \tag{53.58}$$

The first equality is the definition of the left-invariant associated vector field (equation (53.7) applied to X_e) and the second equality is the fact that X is left-invariant. Thus X is the left-invariant vector field associated with X_e.

(2) **Injective** Let $X, Y \in T_eG$ be such that $X^L = Y^L$. In particular $X^L_e = Y^L_e$, which means $X = Y$.

□

Proposition 53.24 ([873, 1]).
Let G be a Lie group. The map

$$\begin{aligned}\varphi\colon G \times \mathfrak{g} &\to TG\\ (g, X) &\mapsto X^L_g\end{aligned} \tag{53.59}$$

is a bijection.

Moreover for each $g \in G$, the map

$$\begin{aligned}\varphi_g\colon \mathfrak{g} &\to T_gG\\ X &\mapsto X^L_g\end{aligned} \tag{53.60}$$

is a vector space isomorphism.

9. Left invariant vector field, definition 53.3.

Proof. Several points.

(1) **φ is surjective** Let $X \in TG$; there is some $g \in G$ such that $X \in T_gG$. Since $X = (dL_g)_e(dL_{g^{-1}})_g X$ we have

$$X = (dL_{g^{-1}}X)^L_g = \varphi(g, dL_{g^{-1}}X). \tag{53.61}$$

(2) **φ is injective** If $\varphi(g, X) = \varphi(h, Y)$, we have $X^L_g = Y^L_h$, so that $g = h$. The equality $X^L_g = Y^L_g$ means $(dL_g)_e X = (dL_g)_e Y$. Applying $(dL_{g^{-1}})_g$ on both sides we get $X = Y$.

(3) **φ_g is bijective** These are the same verifications.

(4) **φ_g is linear** The map φ_g is nothing else than $(dL_g)_e$, so it is linear.

□

PROPooWWXKooWEBpMf

Proposition 53.25 ([1]).
Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. The map

$$\begin{aligned} L\colon \mathfrak{g} &\to \Gamma(TG) \\ X &\mapsto X^L \end{aligned} \tag{53.62}$$

is linear.

53.5 Exponential map

53.5.1 Integral curve

The integral curve of a vector field on a manifold is given by definition 49.116. Here we are dealing with special manifolds: Lie groupe. From definition 53.3 we know that one can create vector fields (invariant) on G from an element of T_eG.

We want to prove that the vector space of left invariant vector fields is isomorphic to the tangent vector space T_eG to G at identity. If $X \in T_eG$, we introduce the left invariant vector field $X^L = dLX$, more explicitly:

$$X^L_g = \frac{d}{dt}\Big[gX(t)\Big]_{t=0}. \tag{53.63}$$

Then we consider $\alpha_X\colon I \to G$ the integral curve of maximal length to X^L trough X_e. Here, I is the interval on which α_X is defined. This is the solution of SUBEQSooBJUVooEeOVVA

$$\begin{cases} \dfrac{d}{dt}\Big[\alpha_X(t_0 + t)\Big]_{t=0} = X_{\alpha_X(t_0)} & \text{(53.64a)} \\ \alpha_X(0) = e. & \text{(53.64b)} \end{cases}$$

PROPooWEYCooCvyHNr

Proposition 53.26.
Let $X \in T_eG$. The integral curve of X^L is defined on $\mathbb{R}$ and for every $s, t \in \mathbb{R}$,

$$\alpha_X(s + t) = \alpha_X(s)\alpha_X(t). \tag{53.65}$$

Proof. Let α be any integral curve for X^L and $y \in G$. If we put $\alpha_1(t) = y\alpha(t)$, we have

$$\frac{d}{dt}\Big[\alpha_1(t)\Big]_{t=0} = X^L_y, \tag{53.66}$$

so that α_1 is an integral curve for X^L trough the point y.

Let now I be the maximal domain of α_X, and $t_1 \in I$. If we set $x_1 = \alpha_X(t_1)$, the path

$$\alpha_1(t) = x_1\alpha_X(t) \tag{53.67}$$

is an integral curve of X^L trough x_1 and has the same maximal definition domain I. On the other hand, the maximal integral curve starting at e being α_X, the maximal integral curve starting at $\alpha_X(t_1)$ is

$$\alpha_2 : t \mapsto \alpha_X(t + t_1). \tag{53.68}$$

Its domain is $I - t_1$, but since it starts at x_1, it has to be the same as α_1, then $I \subset I - t_1$ which proves that $I = \mathbb{R}$.

For each s and t in $\mathbb{R}$, the maximal integral curve starting at $\alpha_X(s)$ can be written as

$$c(t) = \alpha_X(s)\alpha_X(t) \tag{53.69}$$

as well as

$$d(t) = \alpha_X(s + t), \tag{53.70}$$

so again by unicity, $\alpha_X(s + t) = \alpha_X(s)\alpha_X(t)$. □

LEMooNZCJooSltfLg

Lemma 53.27 ([1])**.**
Let G be a Lie group and $X \in \mathfrak{h}$. For every $u \in \mathbb{R}$ we have

$$\Phi_e^{(uX)^L}(t) = \Phi_e^{X^L}(tu). \tag{53.71}$$

EQooMNQLooHNDpJS

Proof. First observation: proposition 53.26 ensures that the flow exists for every u, so that the left and right hand sides of (53.71) makes sense for every $u \in \mathbb{R}$.

Let $u \in \mathbb{R}$. We define

$$\begin{aligned} \gamma : \mathbb{R} &\to G \\ t &\mapsto \Phi_e^{X^L}(tu) \end{aligned} \tag{53.72}$$

and we prove that this is the integral curve of $(uX)^L$ by checking the two conditions of proposition 49.116. The first condition is easy:

$$\gamma(0) = \Phi_e^{X^L}(0) = e. \tag{53.73}$$

The second condition to be checked is

$$\gamma'(t_0) = (uX)^L_{\gamma(t_0)}. \tag{53.74}$$

We have

$$\begin{aligned} \gamma'(t_0) &= u(\Phi_e^{X^L})'(t_0 u) && \text{lem. 49.52} && (53.75a)\\ &= u(X^L)_{\Phi_e^{X^L}(t_0 u)} && && (53.75b)\\ &= u(X^L)_{\gamma(t_0)} && && (53.75c)\\ &= (uX)^L_{\gamma(t_0)} && \text{prop. 53.25} && (53.75d) \end{aligned}$$

□

PROPooUXFQooIwimav

Proposition 53.28.
The flow[10] *of a smooth left-invariant vector field X is given by*

$$\Phi(t, g) = g\Phi(t, e). \tag{53.76}$$

Proof. In few steps.

(1) **Left invariance**
The left invariance means that $(dL_g)_h X_h = X_{gh}$ for every $h \in G$. We write that condition with $h = \Phi_e(t_0)$:

$$(dL_g)_{\Phi_e(t_0)} X_{\Phi_e(t_0)}(f) = X_{g\Phi_e(t_0)}(f). \tag{53.77}$$

EQooPGWKooWNKslJ

for every smooth function $f : M \to \mathbb{R}$.

10. Definition 49.116.

(2) **The path** In order to prove that $\Phi_g(t) = \Phi_e(t)$ we consider the path

$$\begin{aligned}\gamma\colon I &\to M\\ t &\mapsto g\Phi_e(t)\end{aligned} \tag{53.78}$$

and we prove that it satisfy the properties to be the integral curve of X at g, that is the two conditions of definition 49.116.

(3) **First condition** This is the easy one: $\gamma(0) = g\Phi_e(0) = ge = g$.

(4) **Second condition** We apply $\gamma'(0)$ to a function f:

$$\begin{aligned}
\gamma'(0)f &= \frac{d}{dt}\Big[(f\circ\gamma)(t)\Big]_{t=t_0} && && (53.79a)\\
&= \frac{d}{dt}\Big[f\big(g\Phi_e(t)\big)\Big]_{t=t_0} && && (53.79b)\\
&= \frac{d}{dt}\Big[(f\circ L_g)\big(\Phi_e(t)\big)\Big]_{t=t_0} && && (53.79c)\\
&= (dL_g)_{\Phi_e(t_0)}\big(\Phi_e'(t_0)\big)f && \text{proposition 49.60} && (53.79d)\\
&= (dL_g)_{\Phi_e(t_0)}X_{\Phi_e(t_0)}(f) && && (53.79e)\\
&= X_{g\Phi_e(t_0)}(f) && \text{by (53.77)} && (53.79f)\\
&= X_{\gamma(t_0)}(f). && && (53.79g)
\end{aligned}$$

SUBEQooOMDQooFZaCLp

Justifications.

— For (53.79e), by definition of an integral curve, $\Phi_e'(t_0) = X_{\Phi_e(t_0)}$

□

Proposition 53.29.
The flows of X^L and X^R are defined on $\mathbb{R}$.

53.5.2 Integral curve and exponential

DEFooOLLZooMHRgsz

Definition 53.30 (Exponential map)**.**
*If G is a Lie group with algebra $\mathfrak{g}$, we define the **exponential** is the map*

$$\begin{aligned}\exp\colon \mathfrak{g} &\to G\\ X &\mapsto \Phi_e^{X^L}(1)\end{aligned} \tag{53.80}$$

EqdefExpoLieTgFGp

where Φ is the flow defined in 49.116 and X^L is the left-invariant vector field defined in 53.4.

53.31.
This definition works on Lie groups thanks to the group structure that allows to build a natural vector field X^L from the data of a single vector X. On general manifolds, one has not a notion of exponential. However, if one has a Riemannian manifold, one consider the geodesic.

In the case of groups for which the Killing form defines a scalar product, the notion of exponential associated with the Riemannian structure propagated from the Killing form coincides with the definition (53.80).

PROPooMAGKooInwNom

Proposition 53.32.
The map $\exp\colon \mathfrak{g} \to G$ is continuous.

53.5.3 Flow and exponential

LEMooEQFHooRjUAin

Lemma 53.33.
We have

$$\exp(0) = e. \tag{53.81}$$

LEMooGKDKooFTsDSr

Lemma 53.34 ([1]).
Let G be a Lie group, $\mathfrak{h}$ its Lie algebra. Let $t \in \mathbb{R}$. We have

$$\exp(tX) = \Phi_e^{X^L}(t). \tag{53.82}$$

In other words, $t \mapsto \exp(tX)$ is an integral curve of X.

Proof. Computation:

$$\exp(tX) = \Phi_e^{(tX^L)}(1) \qquad \text{definition} \tag{53.83a}$$
$$= \Phi_e^{X^L}(t) \qquad \text{lem. 53.27.} \tag{53.83b}$$

□

PROPooMIMZooAwxvkB

Proposition 53.35 ([870]).
Let G be a smooth Lie group. The exponential map $\exp\colon \mathfrak{g} \to G$ is smooth.

Proof. We consider the product manifold [11] $M = G \times \mathfrak{g}$ and the vector field $Y_{(g,X)} = (X_g^L, 0)$, and the map

$$\begin{aligned} \Phi\colon \mathbb{R} \times G \times \mathfrak{g} &\to G \times \mathfrak{g} \\ (t, g, X) &\mapsto (ge^{tX}, X). \end{aligned} \tag{53.84}$$

We prove the Φ is the flow [12] of the vector field Y. Since $\exp(0) = e$ (lemma 53.33), we have $\Phi(0, g, X) = (g, X)$. For the second condition, we use the lemma 49.58:

$$\frac{d}{dt}\Big[\Phi_{(g,X)}(t)\Big]_{t=0} = \frac{d}{dt}\Big[\big(g\exp(tX), X\big)\Big]_{t=0} \tag{53.85a}$$
$$= \Big(\frac{d}{dt}\Big[g\exp(tX)\Big]_{t=0}, \frac{d}{dt}\Big[X\Big]_{t=0}\Big) \tag{53.85b}$$
$$= (X_g^L, 0) \tag{53.85c}$$
$$= Y_{(g,X)}. \tag{53.85d}$$

Since the vector field Y is smooth, the proposition 49.117 says that Φ is smooth.

Let

$$\begin{aligned} \alpha\colon \mathfrak{g} &\to G \times \mathfrak{g} \\ X &\mapsto \Phi(1, e, X) = \big(\exp(X), X\big). \end{aligned} \tag{53.86}$$

As restriction of Φ, the map α is smooth.

The projection map

$$\begin{aligned} \pi_1\colon G \times \mathfrak{g} &\to G \\ (g, X) &\mapsto g \end{aligned} \tag{53.87}$$

is smooth from the proposition 49.47(2). As composition, the map

$$\exp = \pi_1 \circ \alpha \tag{53.88}$$

is smooth. □

LEMooKSTKooNCnxzB

Lemma 53.36.
If $\varphi\colon G \to H$ is a Lie group morphism, then

$$d\varphi_e(X)^L = d\varphi(X^L). \tag{53.89}$$

More explicitly, for every $g \in G$ we have

$$d\varphi_e(X)^L_{\varphi(g)} = d\varphi_g X_g^L. \tag{53.90}$$

11. See the proposition 49.57 for the manifold structure on the vector space $\mathfrak{g}$ and the definition 49.46 for the product of manifolds.
12. Definition 49.116.

Proof. Let $h \in H$. We compute the vector field $d\varphi_e(X)^L$ at h:

$$\begin{aligned}
d\varphi_e(X)^L_h &= (dL_h)_e\big(d\varphi_e(X)\big) && (53.91a)\\
&= \frac{d}{dt}\Big[(L_h \circ \varphi)\big(\exp(tX)\big)\Big]_{t=0} && (53.91b)\\
&= \frac{d}{dt}\Big[\varphi\big(\varphi^{-1}(h)e^{tX}\big)\Big]_{t=0} && (53.91c)\\
&= d\varphi_{\varphi^{-1}(h)}\frac{d}{dt}\Big[L_{\varphi^{-1}(h)}e^{tX}\Big]_{t=0} && (53.91d)\\
&= d\varphi_{\varphi^{-1}(h)}\big(dL_{\varphi^{-1}(h)}\big)_e X && (53.91e)\\
&= d\varphi_{\varphi^{-1}(h)}X^L_{\varphi^{-1}(h)}. && (53.91f)
\end{aligned}$$

SUBEQooEWJUooMNJdHj

Justifications.

— For (53.91c). Since φ is a group morphism, $h\varphi(g) = \varphi\big(\varphi^{-1}(h)g\big)$.

□

LEMooHQYMooJjdKkG

Lemma 53.37 ([870])**.**
Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. We have

$$(d\exp)_0 \circ j = \mathrm{Id} \tag{53.92}$$

where j is given by[13]

$$\begin{aligned} j\colon \mathfrak{g} &\to T_0\mathfrak{g}\\ X &\mapsto \frac{d}{dt}\Big[tX\Big]_{t=0}. \end{aligned} \tag{53.93}$$

Proof. Let $\gamma(t) = tX$, so that $j(X) = \gamma'(0)$. We have

$$\begin{aligned}
(d\exp)_0 j(X) &= (d\exp)_0\gamma'(0) && (53.94a)\\
&= \frac{d}{dt}\Big[\exp\big(\gamma(t)\big)\Big]_{t=0} && (53.94b)\\
&= \frac{d}{dt}\Big[\exp(tX)\Big]_{t=0} && (53.94c)\\
&= \frac{d}{dt}\Big[\Phi_e^{X^L}(t)\Big]_{t=0} && \text{lem. 53.34} \quad (53.94d)\\
&= X^L_e && (53.94e)\\
&= X. && (53.94f)
\end{aligned}$$

□

lemsur5d

Lemma 53.38 ([872])**.**
Let G, H be two Lie groups with algebras[14] *$\mathcal{G}$ and $\mathcal{H}$. Let $\phi\colon G \to H$ be a group morphism which is differentiable at $e \in G$. Then for every $X \in \mathfrak{g}$, the following formula holds:*

$$\phi(\exp X) = \exp(d\phi_e X). \tag{53.95}$$

PROPooKYNDooVAEzFw

Proposition 53.39 ([870, 872])**.**
Let $\varphi\colon G \to H$ be a Lie groupe morphism. We have

$$\exp \circ d\varphi_0 = \varphi \circ \exp. \tag{53.96}$$

13. See also the proposition 49.57.
14. Lie algebra of a Lie group, definition 53.5.

Proof. Let $x \in \mathfrak{g}$ and define

$$\begin{aligned} \phi_X : \mathbb{R} &\to H \\ t &\mapsto (\varphi \circ \exp)(tX). \end{aligned} \tag{53.97}$$

We prove that this is an integral curve of $d\varphi(X^L)$. Using lemma 53.34 and the fact that φ is a group morphism ($\varphi(e) = e$), we have the first condition:

$$\phi_X(0) = (\varphi \circ \exp)(0) = \varphi(e) = e. \tag{53.98}$$

The second condition to be checked is, for every t_0 in a neighborhood of $0 \in \mathbb{R}$,

$$\frac{d}{dt}\Big[\phi_X(t)\Big]_{t=t_0} = d\varphi(X^L)_{\phi_X(t_0)}. \tag{53.99}$$

To prove that, we write $h = \phi_X(t_0)$. We have

$$\begin{aligned} \frac{d}{dt}\Big[\phi_X(t)\Big]_{t=t_0} &= \frac{d}{dt}\Big[\varphi\big(\exp(tX)\big)\Big]_{t=t_0} && (53.100a)\\ &= d\varphi_{\exp(t_0X)}\frac{d}{dt}\Big[\exp(tX)\Big]_{t=t_0} && (53.100b)\\ &= d\varphi_{\exp(t_0X)}\frac{d}{dt}\Big[\Phi_e^{X^L}(t)\Big]_{t=t_0} && \text{lem. 53.34} \quad (53.100c)\\ &= d\varphi_{\exp(t_0X)}X^L_{\exp(t_0X)} && (53.100d)\\ &= d\varphi_{\varphi^{-1}\left(\phi_X(t_0)\right)}X^L_{\varphi^{-1}\left(\phi_X(t_0)\right)} && (53.100e)\\ &= d\varphi_e(X)^L_{\phi_X(t_0)} && \text{lem. 53.36} \quad (53.100f) \end{aligned}$$

Justifications.

— For (53.100d). From the definition of the flow,

$$\frac{d}{dt}\Big[\Phi_e^{X^L}(t)\Big]_{t=t_0} = X^L_{\Phi_e^{X^L}(t_0)} = X^L_{\exp(t_0X)}. \tag{53.101}$$

— By definition of ϕ_X, $\exp(t_0X) = \varphi^{-1}\big(\phi_X(t_0)\big)$.

We proved that ϕ_X is an integral curve of $d\varphi(X)$. But lemma 53.34 says that the integral curve of $d\varphi(X)$ is $t \mapsto \exp\big(td\varphi(X)\big)$. We deduce that these two curves are equal, and that, for every t we have

$$\varphi\big(\exp(tX)\big) = \exp\big(d\varphi_e(tX)\big). \tag{53.102}$$

□

The following proposition is a generalization of 15.70.

PROPooNRVJooEDCpOI

Proposition 53.40.
If $X \in \mathfrak{g}$ and $s, t \in \mathbb{R}$ we have

$$e^{sX}e^{tX} = e^{(s+t)X}. \tag{53.103}$$

LEMooRPHVooAtZJnz

Lemma 53.41.
Let G be a Lie group with algebra $\mathfrak{g}$. If $n \in \mathbb{N}$ we have

$$\exp(X)^n = \exp(nX). \tag{53.104}$$

Proof. Apply n times the proposition 53.40. □

LEMooLMTZooCvunSl

Lemma 53.42.
Let G be a Lie group and $X \in G$. We have

$$X^R_g = \frac{d}{dt}\Big[\exp(tX)g\Big]_{t=0} \tag{53.105}$$

EQooNBENooPXLENs

and

$$X^L_g = \frac{d}{dt}\Big[g\exp(tX)\Big]_{t=0} \tag{53.106}$$

53.5.4 Invariant vector and derivation

Proof. We prove the one about X^R; the other one is similar. We have

$$\frac{d}{dt}\Big[\exp(tX)g\Big]_{t=0} = \frac{d}{dt}\Big[R_g\big(\exp(tX)\big)\Big]_{t=0} \tag{53.107a}$$
$$= (dR_g)_e \frac{d}{dt}\Big[\exp(tX)\Big]_{t=0} \qquad \text{by (49.119)} \tag{53.107b}$$
$$= (dR_g)_e X. \tag{53.107c}$$

□

NORMooSATDooIhwXXr

53.43.
We will often write the relation (53.105) under the form

$$X_g^R(t) = e^{tX}g. \tag{53.108}$$

This is a way to implies that $t \mapsto e^{tX}g$ is a path for the vector X_g^R. It is a common abuse of notation to write the vector and a path representing the vector with the same symbol.

You may want to know how the exponential can be used to write some formulas linking left-invariant vector field and derivation of functions. Here you are.

53.44.
Let $X \in \mathfrak{g}$, $g \in G$ and $u \in \mathbb{R}$. Let $f\colon G \to \mathbb{R}$ be a smooth function. Using the abuse of notation described in 53.43 and the proposition 53.40,

$$(X^L f)(ge^{uX}) = \frac{d}{dt}\Big[f\big(X^L_{ge^{uX}}(t)\big)\Big]_{t=0} \tag{53.109a}$$
$$= \frac{d}{dt}\Big[f\big(ge^{uX}e^{tX}\big)\Big]_{t=0} \tag{53.109b}$$
$$= \frac{d}{dt}\Big[f\big(ge^{(t+u)X)}\big)\Big]_{t=0} \tag{53.109c}$$
$$= \frac{d}{dt}\Big[f(ge^{tX})\Big]_{t=u}. \tag{53.109d}$$

The formula

$$(X^L f)(ge^{uX}) = \frac{d}{dt}\Big[f(ge^{tX})\Big]_{t=u} \tag{53.110}$$

means that X^L derives f in the direction of the path e^{tX} at right.

53.45.
By the way, we recall that, if f is a function and X a vector field, (Xf) is a new function, given by

$$(Xf)(a) = X_a(f). \tag{53.111}$$

In that sense we can write combinations like XYf or $(X^2 + X)f$ where X and Y are vector fields.

PROPooKSIDooVIFkiM

Proposition 53.46 ([874])**.**
Let G be a Lie group with Lie algebra $\mathfrak{g}$. We consider $X, Y \in \mathfrak{g}$ and a smooth function $f\colon G \to \mathbb{R}$. We have[15]

$$\big((X^R)^n(Y^L)^m f\big)(e^{sX}e^{tY}) = \frac{d^n}{du^n}\frac{d^m}{dv^m}\Big(f(e^{uX}e^{vY})\Big)_{\substack{u=s\\v=t}}. \tag{53.112}$$

Proof. We have to do a proof by induction on (n, m). We start with $(n, m) = (0, 0)$ and we prove the steps $(n, m) \to (n + 1, m)$ and $(n, m) \to (n, m + 1)$.

(1) $\underline{(0,0)}$ With $(n, m) = (0, 0)$ we are okay.

15. My source [874] seems to write $(X^R)^n(Y^R)^m$ instead of $(X^R)^n(Y^L)^m$. Let me know where I'm wrong.

(2) $\underline{(n+1,m)}$ We have

$$\Big((X^R)^{n+1}(Y^L)^m f\Big)(e^{sX}e^{tY}) = \big((X^R)(X^R)^n(Y^L)^m f\big)(e^{sX}e^{tY}). \tag{53.113}$$

We will apply the induction hypothesis on the function $(X^R)^n(Y^L)^m f$, but in a first time we just apply the vector field X^R to the function $(X^R)^n(Y^L)^m$ and we evaluate at $e^{sX}e^{tY}$. Here is a couple of computations:

$$\begin{aligned}
\Big((X^R)(X^R)^n(Y^L)^m f\Big)(e^{sX}e^{tY}) &= \frac{d}{du}\Big[\Big((X^R)^n(Y^L)^m f\Big)\big(X^R_{e^{sX}e^{tY}}(u)\big)\Big]_{u=0} && \text{(53.114a)}\\
&= \frac{d}{du}\Big[\Big((X^R)^n(Y^L)^m f\Big)(e^{uX}e^{sX}e^{tY})\Big]_{u=0} && \text{(53.114b)}\\
&= \frac{d}{du}\Big[\Big((X^R)^n(Y^L)^m f\Big)(e^{uX}e^{tY})\Big]_{u=s}. && \text{(53.114c)}
\end{aligned}$$

At this point we use the induction hypothesis:

$$\begin{aligned}
\frac{d}{du}\Big[\Big((X^R)^n(Y^L)^m f\Big)(e^{uX}e^{tY})\Big]_{u=s} &= \frac{d}{du}\left(\frac{d^n}{dw^n}\frac{d^m}{dv^m}\big(f(e^{wX}e^{vY})\big)_{\substack{w=u\\v=t}}\right)_{u=s} && \text{(53.115a)}\\
&= \frac{d^{n+1}}{dw^{n+1}}\frac{d^m}{dv^m}\big(f(e^{wX}e^{vY})\big)_{\substack{w=s\\v=t}}. && \text{(53.115b)}
\end{aligned}$$

(3) $\underline{(n,m+1)}$ Same kind of computations.

□

53.6 Exponential as bijection on a neighborhood

PROPooYFZZooLUOuOj

Proposition 53.47 ([875]).
Let G be a smooth Lie group. There exists a neighbourhood U of 0 in $\mathfrak{g}$ and a neighbourhood V of e in G such that

$$\exp\colon U \to V \tag{53.116}$$

is a C^∞-diffeomorphism[16].

Proof. Proposition 53.35 shows that the exponential is smooth. On the other hand, lemma 53.37 says that the differential is invertible. The local inversion theorem 49.142 makes exp a local C^∞-diffeomorphism. □

THOooFMFLooCnLJPr

Theorem 53.48.
Let G be an analytic Lie group. There exists a neighbourhood U of 0 in $\mathfrak{g}$ and a neighbourhood V of e in G such that

$$\exp\colon U \to V \tag{53.117}$$

is an analytic diffeomorphism[17].

53.6.1 Analytic Lie group, Taylor formula

In this subsection we study the analytic functions over an analytic Lie group.

LEMooPILVooHQbtAH

Lemma 53.49 ([874]).
Let G be an analytic Lie group. We consider an analytic function $f\colon G \to \mathbb{R}$, an element $X \in \mathfrak{g}$, a basis $\{X_i\}$ of $\mathfrak{g}$ and $g \in G$. There exists an absolutely converging power series P such that

$$f(ge^{x_1X_1+\dots+x_nX_n}) = P(x_1,\dots,x_n). \tag{53.118}$$

16. exp is C^∞, invertible and he inverse is C^∞ as well.
17. exp is analytic, invertible and its inverse is analytic too.

Proof. First we make the proof for $g = e$.

We consider a basis $\{e_i\}$ of $\mathfrak{g}$. Let U be a neighbourhood of 0 in $\mathfrak{g}$ and V a neighbourhood of e in G such that $\exp\colon U \to V$ is an analytic diffeomorphism [18].

We consider U', the open set in $\mathbb{R}^n$ which correspond to U via the basis $\{e_i\}$. The map

$$\begin{aligned} \varphi\colon U' &\to V \\ (x_1, \dots, x_n) &\mapsto \exp(x_1 e_1 + \dots + x_n e_n) \end{aligned} \tag{53.119}$$

is analytic chart of V.

The fact that f is analytic means that the composition of f with the charts are analytic. In our case, the map $\tilde{f} = f \circ \varphi$ is analytic from $U' \subset \mathbb{R}^n$ to $\mathbb{R}$. Thus there exists an absolutely converging power series P such that

$$\tilde{f}(x_1, \dots, x_n) = P(x_1, \dots, x_n). \tag{53.120}$$

We conclude:

$$f\big(\exp(x_1 e_1 + \dots + x_n e_n)\big) = f\big(\varphi(x_1, \dots, x_n)\big) = P(x_1, \dots, x_n). \tag{53.121}$$

If g is not e, we consider the neighbourhood gV and the map

$$\begin{aligned} \varphi\colon U &\to gV \\ (x_1, \dots, x_n) &\mapsto g\exp(x_1 e_1 + \dots + x_n e_n) \end{aligned} \tag{53.122}$$

is a chart, so that

$$f(g e^{x_1 e_1 + \dots + x_n e_n}) = \tilde{f}(x_1, \dots, x_n) \tag{53.123}$$

which is a power series. □

PROPooIYWQooZJtKiu

Proposition 53.50 (Taylor formula[874])**.**
Let G be an analytic Lie group. We suppose that $f\colon G \to \mathbb{R}$ is an analytic functions. For $g \in G$ and $X \in \mathfrak{g}$ we have

$$f(g e^X) = \sum_{k=0}^{\infty} \frac{1}{n!}\big((X^R)^n f\big)(g). \tag{53.124}$$

Proof. We know from proposition 53.49 that $f(ge^X) = P(x_1, \dots, x_n)$ for some power series P. We consider a neighbourhood U of 0 in $\mathfrak{g}$ and V of g in G such that

$$\begin{aligned} \varphi\colon U &\to V \\ X &\mapsto g e^X \end{aligned} \tag{53.125}$$

is an analytic diffeomorphism (i.e. an analytic chart for G around g). Let $X \in U$ and δ such that $tX \in U$ for all $t \in]-\delta, \delta[$. Notice that $\delta > 1$. Now, X being fixed, the value of $P(tx_1, \dots, tx_n)$ is an absolutely convergent power series of t. We have

$$f(g e^{tX}) = P(tx_1, \dots, tx_n) = \sum_{k=0}^{\infty} \frac{a_m}{m!} t^m \tag{53.126}$$

for some constants $a_m \in \mathbb{R}$.

But considering the function

$$\begin{aligned} r\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(g e^{tX}), \end{aligned} \tag{53.127}$$

there is an unicity of its power series expansion; thus a_m is the m-th derivative of r at $t = 0$.

18. By theorem 53.48.

But we also know from proposition 53.46 that

$$\big((X^L)^m f\big)(ge^{tX}) = \frac{d^m}{du^m}\big(f(ge^{uX})\big)_{u=t}; \tag{53.128}$$

taking that at $t = 0$ we have

$$a_m = \big((X^L)^m f\big)(g) \tag{53.129}$$

and the Taylor formula

$$f(ge^{tX}) = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{d^k}{du^k}\big(f(ge^{uX})\big)_{u=0} t^m. \tag{53.130}$$

Finally taking $t = 1$ (recall that $\delta > 1$, so it is valid):

$$f(ge^{X}) = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{1}{k!}\big((X^L)^k f\big)(g). \tag{53.131}$$

□

DEFooEUTBooVqlEGM

Definition 53.51 (normal neighbourhood)**.**
*Let M be a differentiable manifold. If V is a neighbourhood of zero in T_pM on which the exponential $\exp_p\colon T_pM \to M$ is a diffeomorphism, then $\exp_p V$ is **normal neighbourhood** of p.*

LEMooWKFIooRHsrFX

Lemma 53.52.
Let G be an analytic Lie group with algebra $\mathfrak{g}$. We consider a basis $\{e_i\}_{i=1,\dots,n}$ of $\mathfrak{g}$ and the functions

$$\begin{aligned} f_i\colon U &\to \mathbb{R} \\ \exp(x_1e_1 + \dots + x_ne_n) &\mapsto x_i \end{aligned} \tag{53.132}$$

defined on a normal neighbourhood[19] U of e.

If $X, Y \in \mathfrak{g}$ satisfy

$$Xf_i = Yf_i \tag{53.133}$$

for every i, then $X = Y$.

Proof. If $X = \sum_k X_k e_k$ we have

$$X(f_i) = \frac{d}{dt}\Big[f_i(e^{tX})\Big]_{t=0} = \frac{d}{dt}\Big[f_i\big(e^{t\sum_k X_k e_k}\big)\Big]_{t=0} = \frac{d}{dt}\Big[tX_i\Big]_{t=0} = X_i. \tag{53.134}$$

□

LEMooDHXPooGOmNYT

Lemma 53.53 ([1])**.**
Let G be a Lie group. Let smooth paths $\gamma\colon \mathbb{R} \to G$ and $\sigma\colon \mathbb{R} \to G$ satisfying $\gamma(0) = \sigma(0) = e$. If $\nabla_\gamma = \nabla_\sigma$, then there exists a smooth map $\alpha\colon \mathbb{R} \to G$ such that

$$\begin{cases} \sigma(t) = \alpha(t)\gamma(t) & \text{(53.135a)} \\ \alpha(0) = e & \text{(53.135b)} \\ \alpha'(0) = 0 & \text{(53.135c)} \end{cases}$$

LEMooGMMNooVlDkNm

Lemma 53.54 ([1])**.**
Let G be an analytic Lie group with Lie algebra $\mathfrak{g}$. Let $X, Y \in \mathfrak{g}$.

(1) We have the formula

$$[X, Y] = \frac{d}{dt}\Big[\exp\big(\sqrt{t}X\big)\exp\big(\sqrt{t}Y\big)\exp\big(-\sqrt{t}X\big)\exp\big(-\sqrt{t}Y\big)\Big]_{t=0}, \tag{53.136}$$

EQooDCUJooYDDZHD

19. Definition 53.51.

(2) There exists a smooth map $\alpha\colon \mathbb{R} \to G$ *such that* $\alpha(0) = e$, $\alpha'(0) = 0$ *and*

$$\exp\big(t[X,Y]\big) = \alpha(t)\exp\big(\sqrt{t}X\big)\exp\big(\sqrt{t}Y\big)\exp(-\sqrt{t}X)\exp(-\sqrt{t}Y) \tag{53.137}$$

LEMooMJBRooMOuJpa

Lemma 53.55 ([874]).
Let G *be an analytic Lie group with Lie algebra* $\mathfrak{g}$. *For* $X, Y \in \mathfrak{g}$ *we have:*

ITEMooHVOIooKDrUSw

(1) $\exp(tX)\exp(tY) = \exp\big(t(X+Y) + \frac{t^2}{2}[X,Y] + t^2\alpha(t)\big)$,

ITEMooWIQIooHphJcP

(2) $\exp\big(t(X+Y)\big) = \exp(tX)\exp(tY)\exp(t\alpha(t))$

ITEMooVMDCooExpIrp

(3) $\exp(-tX)\exp(-tY)\exp(tX)\exp(tY) = \exp\big(t^2[X,Y] + t^3\alpha(t)\big)$.

In both formulas, α *is a function* $\alpha\colon \mathbb{R} \to \mathfrak{g}$ *satisfying* $\lim_{t\to 0}\alpha(t) = 0$.

Proof. Several steps.

(1) **A good function**
Let $\{e_i\}_{i=1,\dots,n}$ be a basis of $\mathfrak{g}$. We consider a neighbourhood U of 0 in $\mathfrak{g}$ and V of e in G such that $\exp\colon U \to V$ is an analytic diffeomorphism. On that U we consider the function

$$\begin{aligned} f\colon U &\to \mathbb{R} \\ \exp(x_1e_1 + \dots + x_ne_n) &\mapsto x_i \end{aligned} \tag{53.138}$$

for some fixed i. This function is analytic and satisfies $f(e) = 0$.

(2) **Some Taylor expansions** Using proposition 53.46 we have

$$\big((X^R)^n(X^L)^m f\big)(e^{sX}e^{tY}) = \frac{d^n}{du^n}\frac{d^m}{dv^m}\big(f(e^{uX}e^{vY})\big)_{\substack{u=s\\v=t}}. \tag{53.139}$$

Considering the function $q(s,t) = f(e^{sX}e^{tY})$, we have the Taylor expansion

$$f(e^{sX}e^{tY}) = q(s,t) = \sum_{m,n\geqslant 0}\frac{s^n}{n!}\frac{t^m}{m!}\big((X^R)^n(Y^L)^m f\big)(e) = \sum_{m,n\geqslant 0}\frac{s^n}{n!}\frac{t^m}{m!}\big(X^nY^m f\big)(e). \tag{53.140}$$

EQooNBOIooRxlZmP

Here the second equality is due to the fact that $(X^Lf)(e) = (X^Rf)(e) = X(f)$.

(3) **The function Z**
On the other hand, when t is small enough, the element $e^{tX}e^{tY}$ belongs to a normal neighbourhood of e, so that there exists an element $Z(t) \in \mathfrak{g}$ satisfying

$$e^{tX}e^{tY} = e^{Z(t)}. \tag{53.141}$$

The element $Z(t)$ is given by

$$Z(t) = \exp^{-1}\big(e^{tX}e^{tY}\big). \tag{53.142}$$

Since the exponential is an analytic diffeomorphism [20] (the inverse is analytic), Z is an analytic function around $t = 0$. Thus there exists a function $\alpha\colon \mathbb{R} \to \mathfrak{g}$ such that

$$Z(t) = tZ_1 + t^2Z_2 + t^2\alpha(t) \tag{53.143}$$

EQooRPGCooXtZzFy

and $\lim_{t\to 0}\alpha(t) = 0$. Notice that $Z(0) = 0$, which explain the absence of constant term in (53.143).

(4) **A formula for $f\big(e^{Z(t)}\big)$**
We pose $Z_1 = \sum_k a_{1k}e_k$, $Z_2 = \sum_k a_{2k}e_k$ and $\alpha(t) = \sum_k \sigma_k(t)e_k$, so that

$$Z(t) = \sum_k \big(ta_{1k} + t^2a_{2k} + t^2\alpha_k(t)\big)e_k. \tag{53.144}$$

Applying f we have

$$f\big(e^{Z(t)}\big) = ta_{1i} + t^2a_{2i} + t^2\alpha_i(t) = f\big(e^{tZ_1+t^2Z_2}\big) + t^2\alpha_i(t). \tag{53.145}$$

20. Theorem 53.48.

(5) **Some more Taylor expansions**

We use the Taylor expansion of proposition 53.50 with $g = e$ and $X = Z(t)$:

$$f(e^{Z(t)}) = \sum_k \frac{1}{k!}\big([tZ_1^L + t^2 Z_2^L]^k f\big)(e) + t^2\alpha_i(t). \tag{53.146}$$

Once again we can drop the L exponent since $(X^L f)(e) = X(f)$. We collect out of (53.146) the terms with t and t^2:

$$f(e^{tX}e^{tY}) = f(e^{Z(t)}) = tZ_1(f) + t^2 Z_2 + \frac{t^2}{2}Z_1^2 + t^2\beta(t) \tag{53.147}$$

with $\lim_{t\to 0}\beta(t) = 0$.

(6) **Comparison**

The formulas (53.140) with $s = t$ and (53.147) are Taylor expansions of the same quantity. They are equal; we copy them here:

$$\sum_{m,n}\frac{t^{m+n}}{m!n!}\big((X^R)^n(Y^L)^m f\big)(e) = tZ_1(f) + t^2 Z_2 + \frac{t^2}{2}Z_1^2 + t^2\beta(t) \tag{53.148}$$

On the left hand side, the terms with t and t^2 are obtained when (n, m) is among the possibilities $(0, 1)$, $(1, 0)$, $(2, 0)$, $(0, 2)$, and $(1, 1)$. Collecting we have on the left

$$(X + Y)f + XYf + \frac{1}{2}X^2 f + \frac{1}{2}Y^2 f \tag{53.149}$$

where we used the fact that $\big((X^R)^2 f\big)(e) = X(Xf) = X^2 f$.
Using lemma 53.52 we have $Z_1 f = (X + Y)f$, so that $Z_1 = X + Y$ and then

$$\frac{1}{2}[X, Y] = Z_2. \tag{53.150}$$

At this point we proved that

$$e^{tX}e^{tY} = e^{t(X+Y)+\frac{t^2}{2}[X,Y]+t^2\alpha(t)}. \tag{53.151}$$

This is (1).

For point (2), we are searching for a function β such that

$$e^{tX}e^{tY}e^{t\beta(t)} = e^{t(X+Y)}. \tag{53.152}$$

We replace in the left-hand side the value of $e^{tX}e^{tY}$ given by the point (1) (this is the reason why we write β instead of α) and we isolate $e^{t\beta(t)}$:

$$e^{t\beta(t)} = e^{t(X+Y)}e^{-t(X+Y)-t^2[X,Y]/2-t^2\alpha(t)}. \tag{53.153}$$

So now our aim is to show that the right-hand side of (53.153) can be written as only one exponential with an argument of the form $t\beta(t)$ satisfying $\beta(t) \to 0$. For that, we use (1) once again with $X + Y$ instead of X and $-(X + Y) - t[X, Y]/2 - t\alpha(t)$ instead of Y. What we get is

$$e^{t\beta(t)} = \exp\big(t(-t[X, Y]/3 - t\alpha(t)) + \frac{t^2}{2}\big[X + Y, -(X + Y) - t[X, Y]/2 - t\alpha(t)\big]\big) \tag{53.154a}$$

$$= \exp\big(-\frac{t^2}{2}[X, Y] - t^2\alpha(t) - \frac{t^3}{4}\big[X + Y, [X, Y]\big] - \frac{t^3}{2}\alpha(t)\big). \tag{53.154b}$$

We are done with (2).

et

(3)

□

53.6.2 Exponential and subspace

LEMooFXKBooRnzZKQ

Lemma 53.56 ([876]).
Let G be an analytic Lie group and H be a closed subgroup of G. The set

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } e^{tX} \in H \forall t \in \mathbb{R}\} \tag{53.155}$$

is a vector subspace of $\mathfrak{g}$.

Proof. Let $X, Y \in \mathfrak{h}$, $\lambda \in \mathbb{R}$ and $t \in \mathbb{R}$. The element λX belongs to $\mathfrak{h}$. We have to prove that $X + Y \in \mathfrak{h}$.

For every $n \in \mathbb{N}$, we have $e^{tX/n}e^{tY/n} \in H$ because H is a subgroup. For every n we also have

$$\left(e^{tX/n}e^{tY/n}\right)^n \in H. \tag{53.156}$$

we have to following computation : SUBEQSooMDRVooQXBwiS

$$\left(e^{tX/n}e^{tY/n}\right)^n = \left[\exp\left(\frac{t(X+Y)}{n} + \frac{t^2}{n^2}\alpha(t/n)\right)\right]^n \tag{53.157a}$$ SUBEQooUYNCooJVIWMi

$$= \exp\left(t(X+Y) + \frac{t^2}{n}\alpha(t/n)\right)$$ (53.157b) SUBEQooUYKKooXtGaxL

Justifications.

— For (53.157a). This is lemma 53.55(2).

— For (53.157b). The n enters the exponential from lemma 53.41.

We have the limit

$$\lim_{n\to\infty} \exp\left(t(X+Y) + \frac{t^2}{n}\alpha(t/n)\right) = e^{t(X+Y)} \tag{53.158}$$ EQooTJWDooVJsDJt

because the exponential is continuous [21] and the properties of α. Since the limit exists on the right hand side of (53.157), the limit exists on the left hand side too.

The limit

$$\lim_{n\to\infty} \left(e^{tX/n}e^{tY/n}\right)^n \tag{53.159}$$

is a limit of elements in H. Since H is closed, this is an element of H. We deduce that (53.158) is an element of H. □

LEMooFDIIooCkSJpY

Lemma 53.57 ([876]).
Let G be an analytic Lie group and H be a closed subgroup of G. We consider the set

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } e^{tX} \in H \forall t \in \mathbb{R}\}. \tag{53.160}$$

Let (X_i) be a sequence in $\mathfrak{g}$ such that [22] ITEMooWQZAooTTnenM

(1) $X_i \to 0$

(2) $e^{X_i} \in H$ *for every* i. ITEMooFLLRooCUXnHb

(3) The limit $\lim_{i\to\infty} \frac{X_i}{\|X_i\|}$ *exists. We name it* X.

Then $X \in \mathfrak{h}$.

Proof. We consider $n_i = \text{int}(t/\|X_i\|)$ where int is the «integer part» defined in 1.476, and we write

$$n_i X_i = \text{int}\left(\frac{t}{\|X_i\|}\right) X_i = \|X_i\| \,\text{int}\left(\frac{t}{\|X_i\|}\right) \frac{X_i}{\|X_i\|}. \tag{53.161}$$

21. See 53.47.
22. The topology and the norm on $\mathfrak{h}$ are given in definition 49.78.

Using the hypothesis (1) and (3) and the limit of lemme 12.13 we have

$$\lim_{i\to\infty} n_i X_i = \lim_{i\to\infty} \|X_i\| \operatorname{int}\Big(\frac{t}{\|X_i\|}\Big)\frac{X_i}{\|X_i\|} = tX. \tag{53.162}$$

Using the continuity of the exponential (it commutes with the limit), we have

$$e^{tX} = \exp\Big(\lim_{i\to\infty} n_i X_i\Big) \tag{53.163a}$$
$$= \lim_{i\to\infty} e^{n_i X_i} \tag{53.163b}$$
$$= \lim_{i\to\infty} \exp(X_i)^{n_i}. \tag{53.163c}$$

Since $X_i \in \mathfrak{h}$, we have $\exp(X_i) \in H$, and $\exp(X_i)^{n_i} \in H$ because H is a group.

Thus the sequence $\exp(X_i)^{n_i}$ is a sequence contained in the closed space H and which has a limit $(\exp(tX))$. Thus the limit is contained in H by proposition 7.56. We conclude that $e^{tX} \in H$. □

LEMooNPLBooBrwNWY

Lemma 53.58 ([875])**.**
Let G be a Lie group, $\mathfrak{g}$ its Lie algebra and H, a closed subgroup of G. We consider

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } \exp(tX) \in H \forall t \in \mathbb{R}\}. \tag{53.164}$$

There exists a neighbourhood U of 0 in $\mathfrak{h}$ and V of e in H such that

$$\exp\colon U \to V \tag{53.165}$$

is a bijection.

Proof. Several steps.

(1) **The map ϕ** Let $\mathfrak{h}'$ be a subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}'$ and consider the map

$$\begin{aligned}\phi\colon \mathfrak{h} \oplus \mathfrak{h}' &\to G\\ X + Y &\mapsto \exp(X)\exp(Y).\end{aligned} \tag{53.166}$$

(2) **$\phi\colon \mathfrak{g} \to G$ is a local bijection** Let $X \in \mathfrak{h}$ and $Y \in \mathfrak{h}'$. Using the Leibniz rule 52.3 we have

$$d\phi_0(X+Y) = \frac{d}{dt}\Big[\phi(tX, tY)\Big]_{t=0} = \frac{d}{dt}\Big[\exp(tX)\exp(tY)\Big]_{t=0} = X + Y. \tag{53.167}$$

In other words, $d\phi_0 = \mathrm{Id}$. Since the exponential is continuous [23], the local inversion theorem 49.142 says that ϕ is a local bijection.

Let U be a neighbourhood of 0 in $\mathcal{G}$ and v a neighbourhood of e in G such that $\phi\colon U \to V$ is a bijection.

(3) **ϕ is a local bijection between $\mathfrak{h}$ and H**

We recall that the definition 53.32 says that the topology on $\mathfrak{g}$ is any norm topology. Suppose that ϕ is not a local bijection between $\mathfrak{h}$ and H. That means that if U is a neighbourhood of 0 in $\mathfrak{h}$ and V is a neighbourhood of e in H, then ϕ is not a bijection between U and V.

Let $U_i = B(0, r_i)$ in $\mathfrak{g}$ with $r_i \to 0$ be as small as $U_i \subset U$ for every i. Let $V_i = \phi(U_i)$. The part $B(0, r_i) \cap \mathfrak{h}$ is a neighbourhood of 0 in $\mathfrak{h}$ and $V_i \cap H$ is a neighbourhood of e in H. This ϕ is not a bijection between $B(0, r_i) \cap \mathfrak{h}$ and $V_i \cap H$.

Since $\phi(\mathfrak{h}) \subset H$, we have $\phi\big(B(0, r_i) \cap \mathfrak{h}\big) \subset V_i \cap \mathfrak{h}$. Since the map ϕ is injective and not bijective, it has to be not surjective. Thus there exists $h_i \in (V_i \cap H) \backslash \phi\big(B(0, r_i) \cap \mathfrak{h}\big)$. Now every element of V_i is the image of an element of $B(0, r_i)$, thus there exits $Z_i \in B(0, r_i)$ such that $\phi(Z_i) = h_i$.

The element Z_i can be decomposed into $Z_i = X_i + Y_i$ with $X_i \in \mathfrak{h}$ and $Y_i \in \mathfrak{h}'$. Since h_i is not in the image of $B(0, r_i) \cap \mathfrak{h}$ we deduce that $Y_i \neq 0$. Let us summarize:

23. Proposition 53.32.

— $X_i + Y_i \xrightarrow{\mathfrak{g}} 0$
— $\phi(X_i + Y_i) \in H$
— $Y_i \neq 0$
— $\exp(X_i) \in H$.

We have

$$\underbrace{\phi(X_i + Y_i)}_{\in\mathfrak{h}} = \underbrace{\exp(X_i)}_{\in\mathfrak{h}} \exp(Y_i), \tag{53.168}$$

and then $\exp(Y_i) \in \mathfrak{h}$.

Now we consider the limit Lemme 53.57 says that the limit $Y = \lim_{i\to\infty} Y_i/\|Y_i\|$ exists and belongs to $\mathfrak{h}$. But $\mathfrak{h}'$ is closed (because finite dimensional vector subspace), so that a limit in $\mathfrak{h}'$ remains in $\mathfrak{h}'$. We deduce that

$$Y \in \mathfrak{h} \cap \mathfrak{h}' = \{0\}. \tag{53.169}$$

This is impossible because $\|Y\| = 1$.

Then ϕ is a local bijection between $\mathfrak{h}$ and H: there exists a neighbourhood U of 0 in $\mathfrak{h}$ and a neighbourhood V of e in H such that $\phi\colon U \to V$ is a bijection.

(4) **Conclusion**

The map $\exp\colon U \to V$ is bijective because, on $\mathfrak{h}$ we have $\exp = \phi$.

□

LEMooOBIMooVvIDnb

Lemma 53.59 ([875])**.**
Let G be a Lie group and H, a Lie subgroup of G ($\mathfrak{g}$ and $\mathfrak{h}$ are the corresponding Lie algebras). If H is a topological subspace[24] *of G, then there exists an open neighbourhood U of 0 in $\mathcal{G}$ such that*

(1) $\exp$ *is a diffeomorphism between U and an open neighbourhood of e in G,* ITEMooVGGXooRIXOQc

(2) $\exp\colon U \cap \mathfrak{h} \to \exp(U) \cap H$ *is a smooth diffeomorphism*[25]*.*

PROPooVMLTooFpXoeh

Proposition 53.60 ([1])**.**
Let G be a Lie group. We denote by $\mathfrak{g}$ its Lie algebra. There exists a neighbourhood U of 0 in $\mathfrak{g}$ such that the maps

$$\begin{aligned} \varphi_g\colon U &\to G \\ X &\mapsto g\exp(X) \end{aligned} \tag{53.170}$$

form an atlas of G.

LEMooEBQUooKXkCda

Lemma 53.61.
Let $\mathfrak{g}$ admit a direct sum decomposition (as vector space) $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{n}$. Then there exists open and bounded neighbourhoods $\mathcal{U}_m$ and $\mathcal{U}_n$ of 0 in $\mathfrak{m}$ and $\mathfrak{n}$ such that the map

$$\begin{aligned} \phi\colon \mathcal{U}_m \times \mathcal{U}_n &\to G \\ (A, B) &\mapsto e^A e^B \end{aligned} \tag{53.171}$$

is a diffeomorphism between $\mathcal{U}_m \times \mathcal{U}_n$ and an open neighbourhood of e in G.

Corollary 53.62.
Let G be a Lie group, and K, H two differentiable subgroups of G. We suppose $K \subset H$. Then K is a differentiable subgroup of the Lie group H.

Proof. The Lie algebras of K and H are respectively denoted by $\mathfrak{k}$ and $\mathfrak{h}$. We denote by K^* the differentiable subgroup of H which has $\mathfrak{k}$ as Lie algebra. The differentiable subgroups K and K^* have same Lie algebra, and then coincide as Lie groups. □

24. This is a part of G with the induced topology, definition 7.28.
25. Bijection, smooth, inverse is smooth.

PROPooSOOTooMrtYmP

Proposition 53.63 ([870]).
Let G be a Lie group and H be a closed subgroup of G. Then

ITEMooQUKXooISAtJV

(1) The subgroup H with its induced topology from G is a topological manifold.

ITEMooYGZFooUshihY

(2) H is a smooth Lie group.

Proof. Several parts.

(1) **For (1)**
We have to define an atlas and check the condition of proposition 49.1. We consider on H the induced topology from G. Since G is Hausdorff and second countable, the topological space H is Hausdorff and second-countable (proposition 7.67). Now we define the charts. Let $\mathfrak{g}$ be the Lie algebra of G. We consider

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } \exp(tX) \in H \forall t \in \mathbb{R}\}. \tag{53.172}$$

From lemma 53.58, there exists a neighbourhood U of 0 in $\mathfrak{g}$

$$\exp\colon U \cap \mathfrak{h} \to \exp(U) \cap H \tag{53.173}$$

is a bijection. For each $h \in H$ we define

$$\begin{aligned} \varphi_h\colon U \cap \mathfrak{h} &\to H \\ X &\mapsto h\exp(X). \end{aligned} \tag{53.174}$$

It is useful to write $\varphi_h = L_h \circ \exp$ and $\varphi_h^{-1} = \exp^{-1} \circ L_h^{-1} = \exp^{-1} \circ L_{h^{-1}}$. Lemma 53.59(2) says that φ_h is homeomorphic.

(2) **For (2) : H is a smooth Lie group** We have to check the definition 52.1.

(2a) **The charts are smooth** The charts are smooth : if we set $H_h = \varphi_h(U \cap \mathfrak{h})$, then the map

$$\varphi^{-1} \circ \varphi_g\colon \varphi_g^{-1}(H_h \cap H_g) \to U \cap \mathfrak{h} \tag{53.175}$$

is given by $\varphi_h^{-1} \circ \varphi_g = \exp^{-1} \circ L_{h^{-1}g} \circ \exp$ and is smooth by lemma 53.59(2).

(2b) **Inclusion** We prove that the inclusion map $\tau\colon H \to G$ is smooth with smooth inverse.

(2c) **Inverse** We have to prove that the map

$$\begin{aligned} j\colon H &\to H \\ h &\mapsto h^{-1} \end{aligned} \tag{53.176}$$

is smooth. Since G is a Lie group, we know that the map

$$\begin{aligned} i\colon G &\to G \\ g &\mapsto g^{-1} \end{aligned} \tag{53.177}$$

is smooth. We express j as a composition of smooth maps : $j = \tau^{-1} \circ i \circ \tau$.

□

THOooDEJHooVKJYBL

Theorem 53.64 (Cartan[877, 870]).
Let G be a smooth Lie group. If H is a closed subgroup of G.

(1) There exists a unique structure of smooth manifold on H such that the inclusion map $\iota\colon H \to G$ is a smooth embedding[26].

(2) With that structure, H is a smooth Lie subgroup of G.

Proof. By proposition 53.47 ... □

26. Definition 49.126.

53.6.3 Campbell-Baker-Hausdorff formula

THOooYJPEooSpKHnC

Theorem 53.65.
Let G be a Lie group with Lie algebra $\mathfrak{g}$. For every $X, Y \in \mathfrak{g}$, there exists $Z \in \mathfrak{g}$ such that

$$e^X e^Y = e^Z. \tag{53.178}$$

53.6.4 Smooth and analytic Lie group

We recall that a manifold is a set with its charts. Usually we write "Let M be a manifold" where we should write "Let $(M, \{\varphi_\alpha\}_{\alpha\in I})$ be a manifold". Changing the set of charts can completely change the properties of a manifold. That subtlety is crucial in the following theorem which basically says "every smooth Lie group is analytic".

THOooSQVCooCyEPOS

Theorem 53.66 ([878]).
Let $(G, \{\varphi_\alpha\}_{\alpha\in A})$ be a smooth Lie group[27].

(1) There exists an atlas $\{\phi_i\}_{i\in I}$ of G for which the transition functions are analytic.

(2) The multiplication and the inverse of G are analytic on the manifold $\big(G, \{\phi_i\}_{i\in I}\big)$.

¡¡ Avertissement/question à la lectrice !! 53.67
We have to complete the proof; in particular we have to state and prove all the quoted results.

Proof. We initiate by considerate a vector space isomorphism $\sigma\colon \mathbb{R}^n \to \mathfrak{g}$ where $\mathfrak{g}$ is the Lie algebra of G. Let U' be a neighbourhood of 0 in $\mathfrak{g}$ from which the exponential is a smooth diffeomorphism. We write $U = \sigma^{-1}(U')$. For $g \in G$ we consider the chart

$$\begin{aligned} \phi_g\colon U &\to G \\ v &\mapsto g\exp\big(\sigma(v)\big). \end{aligned} \tag{53.179}$$

Here are some facts[28].

(1) These maps ϕ_g are smooth and then are charts of $(G, \{\varphi_\alpha\})$.

(2) The set $\{\phi_g\}_{g\in G}$ is an atlas of $(G, \{\varphi_\alpha\})$.

(3) The transitions functions $\phi_h^{-1} \circ \phi_g$ are analytic[29] where they are defined.

(4) The multiplication and the inverse are analytic for the manifold $(G, \{\phi_g\}_{g\in G})$.

Here is an hint for the last point. Let g, h be in a neighbourhood of e such that $g, h \in \phi_e(U)$. Let $g = \exp(X_0)$ and $h = \exp(X)$. We look at the map

$$\begin{aligned} \mu\colon \mathfrak{g} &\to G \\ X &\mapsto g\exp(X). \end{aligned} \tag{53.180}$$

The element $\phi_e\big(\mu(X)\big)$ satisfy

$$\exp(X_0)\exp(X) = \exp\Big(\phi_e^{-1}\big(\mu(X)\big)\Big). \tag{53.181}$$

In other terms, $\phi_e^{-1}(X)$ is given by the Campbell-Baker-Hausdorff formula, which is analytic[30], theorem 53.65. □

NORMooKCBMooGWQZJY

53.68.
From now on, we will always consider analytic Lie groups. The interesting point is that, if we show that some charts are smooth, theorem 53.66 allows us to say that the group is analytic by choosing the correct atlas.

27. Definition 52.1.
28. that need some justifications.
29. As functions from an open set of $\mathbb{R}^n$ to an open set of $\mathbb{R}^n$.
30. This affirmation should require some more work.

53.6.5 Topological Lie subgroup

Remarque 53.69.
A *topological* Lie subgroup is stronger that a common Lie subgroup because it needs to be a topological subgroup: it must carry *exactly* the induced topology. In our definition of a Lie group, this feature doesn't appears.

THOooXVXBooZDJzQo

Theorem 53.70.
Let G be a Lie group whose Lie algebra is $\mathfrak{g}$ and H, a closed subgroup (not specially a Lie *subgroup) of G. Then there exists one and only one analytic structure on H for which H is a topological Lie subgroup of G.*

Proof. The unicity part comes from the corollary 52.28.

We will work with

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } \forall t \in \mathbb{R},\, e^{tX} \in H\}. \tag{53.182}$$

eqJH_de_G

□

53.6.6 Action from Lie algebra to Lie group

A very important point[879] is that when G is acting on M, one can reconstruct the action of G only knowing the action of $\mathfrak{g}$. Let $X \in \mathfrak{g}$ and $x \in M$. We consider the path

$$\begin{aligned} \gamma\colon \mathbb{R} &\to M \\ t &\mapsto \exp(tX)(x). \end{aligned} \tag{53.183}$$

This map satisfies $\gamma(0) = x$. We also have, using proposition 53.26,

$$\gamma(t_0 + u) = \big(\exp(uX) \circ \exp(t_0 X)\big)(x) = \exp(uX)\gamma(t_0). \tag{53.184}$$

In that, we used the fact that G acts on M, so that we have transformed the product inside the group $\exp\big((u + t_0)X\big) = \exp(uX)\exp(t_0 x)$ into a composition of map. Then

$$\frac{d}{du}\Big[\gamma(t_0 + u)\Big]_{u=0} = X\big(\gamma(t_0)\big) \tag{53.185}$$

We conclude that γ satisfies the differential equation

$$\gamma'(t) = -X\big(\gamma(t)\big). \tag{53.186}$$

EQooFGSIooUplbmN

When $M = \mathbb{R}^n$, the Cauchy-Lipschitz theorem 17.39 provides unicity of the solution on a maximal domain providing the map $X\colon \mathbb{R}^n \to \mathbb{R}^n$ has nice properties.

NORMooMGAUooIoLtjW

53.71.
If we have to determine the transformations of $\mathbb{R}^n$ that satisfies some properties, the strategy is then the following:

- Suppose the searched group to be a connected Lie group.
- Write the condition with the groupe element $\exp(tX)$ and differentiate with respect to t. This point is what physicist call "consider an infinitesimal transformation and neglect the higher order terms".
- This provides an equation for X. Typically a differential equation for the map $X\colon \mathbb{R}^n \to \mathbb{R}^n$. Solve it.
- The group action is then retrieved solving the differential equation (53.186).

Using that technique we will determine the isometries of $\mathbb{R}^n$ in proposition 53.73 and determine the conformal group around definition 71.2.

Remarque 53.72.
When the Lie algebra is made of linear transformations, the last differential equation to solve is actually exponentiating a matrix.

53.6.7 Example: determining the smooth isometries of the flat vector space

We know from theorem 8.157 that the isometries of $\mathbb{R}^n$ are affine functions. We give now an alternative proof of that result.

PROPooDVIWooAFDNPy

Proposition 53.73.
The smooth[31] *isometries of* $(\mathbb{R}^n, \|.\|)$ *are affine maps.*

Proof. The condition for a diffeomorphism $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ to be an isometry is

$$\|\phi(x) - \phi(y)\|^2 = \|x - y\|^2. \tag{53.187}$$

EQooRKYWooFIKfYZ

We write $\phi_t(x) = e^{-tX}x$ and take the derivative of (53.187) with respect to t at $t = 0$ taking into account that $\phi_0(x) = x$:

$$\big(X(y) - X(x)\big) \cdot (x - y) = 0. \tag{53.188}$$

EQooXEKMooGOktOj

We used the fact that $\frac{d}{dt}\Big[\phi_t(x)\Big]_{t=0} = -X(x)$.

We write the condition (53.188) with tx and take the derivative with respect to t: $dX_0(x) \cdot y + X(y) \cdot x = 0$. The same with y gives

$$dX_0(x) \cdot y + dX_0(y) \cdot x = 0. \tag{53.189}$$

Taking $x = e_i$ and $y = e_j$ this equation reads

$$\frac{\partial X_j}{\partial x_i} + \frac{\partial X_i}{\partial x_j} = 0. \tag{53.190}$$

With $i = j$ we get $\frac{\partial X_i}{\partial x_i} = 0$. The we compute

$$\frac{\partial}{\partial x_j}\frac{\partial X_i}{\partial x_j} = -\frac{\partial}{\partial x_j}\left(\frac{\partial X_j}{\partial x_i}\right) = -\frac{\partial}{\partial x_i}\frac{\partial X_j}{\partial x_j} = 0. \tag{53.191}$$

We used the fact that X_j is of class C^2 in order to permute the derivatives (lemma 12.504). We proved that

$$\frac{\partial^2 X_i}{\partial x_j} = 0 \tag{53.192}$$

for all i, j. Thus X is linear. □

53.6.8 Other stuff

Lemma 53.74.
Let $\mathfrak{g}$ *be a Lie algebra and* A*, a linear operator on* $\mathfrak{g}$ *(see as a common vector space) such that* $\forall t \in \mathbb{R}$*, the map* e^{tA} *is an automorphism of* $\mathfrak{g}$*. Then* A *is a derivation of* $\mathfrak{g}$*.*

lem:autom_derr

Proof. Let us consider $X, Y \in \mathfrak{g}$; the assumption is

$$e^{tA}[X, Y] = [e^{tA}X, e^{tA}Y].$$

Since e^{tA} is a linear map, it has a "good behavior" with the derivations:

$$\frac{d}{dt}\Big(e^{tA}[X, Y]\Big)_{t=0} = \frac{d}{dt}\Big(e^{tA}\Big)_{t=0}[X, Y] = A[X, Y].$$

Using on the other hand the linearity of ad, we can see

$$(\mathrm{ad}(e^{tA}X))(e^{tA}Y)$$

31. In fact we only need C^2.

as a product "matrix times vector". Then

$$\begin{aligned}\frac{d}{dt}\Big([e^{tA}X,e^{tA}Y]\Big)_{t=0}&=\frac{d}{dt}\Big((\operatorname{ad}e^{tA}X)Y\Big)_{t=0}+\frac{d}{dt}\Big((\operatorname{ad}X)(e^{tA}Y)\Big)_{t=0}\\&=(\operatorname{ad}AX)Y+(\operatorname{ad}X)(AY).\end{aligned}\tag{53.193}$$

Finally, $A[X,Y]=[AX,Y]+[X,AY]$.

□

As notational convention, if G and H are Lie groups, their Lie algebra are denoted by $\mathfrak{g}$ and $\mathfrak{h}$.

LemAlgEtGroupesGenere

Lemma 53.75.
Let $\mathfrak{g}$ be a Lie algebra ans $\mathfrak{s}$ be a subset of $\mathfrak{g}$. The algebra of the group generated by $e^{\mathfrak{s}}$ is the algebra generated by $\mathfrak{s}$.

Invariant vector fields are also often used in order to transport a structure from the identity of a Lie group to the whole group by $A_g(X_g)=A_e(dL_{g^{-1}}X_g)$ where A_e is some structure and X_g, a vector at g.

Ad_e

Corollary 53.76.
An useful formula:

$$\operatorname{Ad}(e^X)=e^{\operatorname{ad}X}.$$

Corollary 53.77.
Another useful corollary of lemma 53.38 is the particular case $\phi=\boldsymbol{Ad}(e^X)$:

$$e^Xe^Ye^{-X}=e^{Ad(e^Y)X}.$$

cor:eXeYe-X

Proposition 53.78.
Let G be a connected Lie group.

(1) All the left invariant vector fields are complete. That means that the map $X\mapsto e^X$ is defined for every $X\in\mathcal{G}$.

(2) The map $\exp\colon\mathcal{G}\to G$ is a local diffeomorphism in a neighbourhood of 0 in $\mathcal{G}$.

Proof. (1) The flow is a one parameter subgroup. Thus if e^{tX} is defined for $t\in[0,a]$, by composition, e^{2a} is defined. So e^{tX} is defined for every value of t in $\mathbb{R}$.

(2) Let us consider the manifold $G\times\mathcal{G}$ and the vector field Ξ defined by

$$\Xi_{(g,X)}=\tilde{X}_g\oplus 0\in T_g(G)\oplus T_X\mathcal{G}\simeq T_{(g,X)}(G\times\mathcal{G}).\tag{53.194}$$

The flow of that vector field is given by

$$\Phi_t(g,X)=\big(g\exp(tX),X\big).\tag{53.195}$$

In particular, Ξ is a complete vector field, and we consider the global diffeomorphism

$$\begin{aligned}\Phi_1\colon G\times\mathcal{G}&\to G\times\mathcal{G}\\(g,X)&\mapsto\big(g\exp(X),X\big).\end{aligned}\tag{53.196}$$

On the point (e,X) we have $\Phi_1(e,X)=(\exp(X),X)$. Thus the exponential is the projection on the first component of $\Phi_1(e,X)$ and we can write

$$\exp(X)=\mathtt{proj}_1\circ\Phi_1(e,X).\tag{53.197}$$

It is a smooth function since both the projection and Φ_1 are smooth.
Now, the differential $(d\exp)_0$ is the identity on $\mathcal{G}$, so that the theorem of inverse function [32] makes exp a local diffeomorphism.

□

32. Theorem 17.47.

Lemma 53.79.
In a Lie group, e is an isolated fixed point for the inversion.

Proof. One can use an exponential map in a neighbourhood of e. In this neighbourhood, an element g can be written as $g = e^X$ for a certain $X \in \mathfrak{g}$. The equality $g = g^{-1}$ gives (because the exponential is a diffeomorphism) $X = -X$, so that $X = 0$ and $g = e$. □

53.6.9 Lie algebra of a Lie subgroup

¡¡ Avertissement/question au lecteur !! 53.80
The proposition 53.88 is the same as 53.81, but with more assumptions. Why ?

PROPooLUEJooAAyHVh

Proposition 53.81 ([836]).
Let G be a Lie group and H be a Lie subgroup of G. If $\mathfrak{g}$ is the Lie algebra of G and $\mathfrak{h}$ the one of H, then we have

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } \exp(tX) \in H \, \forall t \in \mathbb{R}\}. \tag{53.198}$$

53.7 Lifting a Lie subalgebra

THOooXALIooGiPVdD

Theorem 53.82 (Chevalley[870]).
We consider a class $\mathcal{A}$ (smooth, analytic) of functions. Let G be a $\mathcal{A}$-Lie group and $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$. There exists an unique connected $\mathcal{A}$-Lie subgroup H whose Lie algebra is $\mathfrak{h}$.

Proof. Let $X_1, \dots, X_k$ be a basis of $\mathfrak{h}$ in $\mathfrak{g}$. We consider the corresponding left invariant vector fields $X_i^L(g) = (dL_g)_e(X_i)$.

(1) **A distribution** By lemma 53.22, the vectors $X_i^L(g)$ are linearly independents in T_gG. Thus by setting

$$\mathcal{D}_g = \text{Span}\{X_i^L(g)\} \tag{53.199}$$

we define a k dimensional distribution [33] on G.

(2) **Involutive** Since $(dL_g)_e$ is linear we have

$$[X_i^L, X_j^L] = [X_i, X_j]^L, \tag{53.200}$$

and since $\mathfrak{h}$ is a Lie algebra, we have $[X_i, X_j] \in \mathfrak{h}$, so that $[X_i, X_j]^L \in \mathcal{D}$.

(3) **Frobenius** The Frobenius theorem 49.136 says that $\mathcal{D}$ is integrable : there exists an unique maximal integral connected manifold H trough e. There exists in particular an immersion $i\colon H \to G$.

(4) **H is a group** Note that, if $g \in G$, then $L_g(H)$ is still an integral manifold of $\mathcal{D}$. Let $h_1, h_2 \in H$. We have

$$h_1 = L_{h_1}(e) \in H \cap L_{h_1}(H) \tag{53.201}$$

because $h_1 \in H$ and $L_{h_1}(e) \in L_{h_1}(H)$. We have:

— The integral manifold H is maximal,

— $L_{h_1}(\mathcal{D})$ is still an integral manifold

— The intersection $L_{h_1}(\mathcal{D}) \cap H$ is not empty.

From these properties we deduce $L_{h_1}(H) \subset H$. Since $h_2 \in H$, we have in particular $L_{h_1}(h_2) \in H$, so that $h_1 h_2 \in H$.

For the inverse, let $h \in H$. We have $L_{h^{-1}}(h) = e \in H$. So $L_{h^{-1}}(H)$ is an integral manifold with an intersection with H. Thus $L_{h^{-1}}(H) \subset H$. In particular $h^{-1} = L_{h^{-1}}(e) \in H$.

We conclude that H is a subgroup of G.

33. Definition 49.131.

(5) G **is a Lie group** The manifold structure on H is the one of submanifold of G. The fact that the multiplication and the inverse are in the class $\mathcal{A}$ as maps on the manifold G implies that they are in the same class $\mathcal{A}$ for the manifold H.
Thus H is a $\mathcal{A}$-Lie group.

(6) **The inclusion is injective** This is always true for an inclusion map.

(7) H **is a Lie group** The manifold structure on H is given by the charts ψ_i of proposition 49.109 which are composition of charts of G with the inclusion. The product and the inverse of H are the restrictions of the product and the inverse on G; just check that their composition with the charts ψ_i are smooth.

(8) **The inclusion is an immersion** The inclusion of a submanifold is an immersion from proposition 49.111.

(9) H **is a Lie subgroup of** G The conditions of the definition 52.24 are satisfied.

The existence part is proven. Now we prove the unicity.

Let K be a Lie subgroup of G whose algebra is $\mathfrak{h}$.

(1) $T_kK = \mathcal{D}_k$ The map $(dL_g)_e : T_eG \to T_gG$ is an isomorphism of vector spaces. Thus we have

$$T_kK = (dL_k)_e(\mathfrak{h}) = \text{Span}\{(dL_k)_e(X_i)\} = \text{Span}\{X_i^L(k)\} = \mathcal{D}_k. \tag{53.202}$$

(2) **Maximality** Thus K is an integral manifold of $\mathcal{D}$ trough e. Thanks to the maximality of H, we have $K \subset H$. Since $T_eK = T_eH$, the inclusion $\iota : K \to H$ is a local isomorphism. There exists a neighbourhood V of e in G such that $K \cap V = H \cap V$.
Since a connected Lie group is generated by any neighbourhood of e, we have $K = H$.

□

53.8 Covering

LEMooSYVQooTjkgBL

Lemma 53.83 ([870]).
Let G and H be connected Lie groups. We consider map $\Phi : G \to H$ such that

(1) Φ is a smooth Lie group diffeomorphism[34].

(2) The map $d\Phi_e : \mathfrak{g} \to \mathfrak{h}$ is bijective.

Then Φ is a covering[35].

Proof. Several points.

(1) Φ **is surjective** Let $h \in H$. From proposition 53.47, there exists a neighbourhood V of e in H on which the exponential is surjective. From proposition 52.5, there exist $h_1, \dots, h_n$ in V such that $h = \prod_{i=1}^n h_i$.
Since $\exp : \mathfrak{h} \to V$ is surjective, there exist $Y_i \in \mathfrak{h}$ such that $h_i = \exp(Y_i)$. We know that $d\Phi : \mathfrak{g} \to \mathfrak{h}$ is surjective, so there exist $X_i \in \mathfrak{g}$ such that $Y_i = d\Phi_e(X_i)$. Thus we have

$$h_i = \exp\big(d\Phi(X_i)\big) = \Phi\big(\exp(X_i)\big) \tag{53.203}$$

from lemma 53.38.
As far as a product is concerned,

$$\begin{aligned} h_ih_j &= \Phi\big(\exp(X_i)\big)\Phi\big(\exp(X_j)\big) && \text{(53.204a)}\\ &= \Phi\big(\exp(X_i)\exp(X_j)\big) && \text{(53.204b)}\\ &= \Phi\big(\exp(Z)\big). && \text{(53.204c)} \end{aligned}$$

SUBEQooRTXGooZNryUY
SUBEQooFHKGooGHUITs

Justifications.

34. See definition 10.265. Everywhere "smooth" means C^∞.
35. Definition 48.73.

— For (53.204b). The map Φ is a group morphism.

— For (53.204c). The element Z is given by the Campbell-Baker-Hausdorff formula, theorem 53.65.

(2) **Continuous** The map Φ is continuous as part of the definition of smooth diffeomorphism.

(3) **The third condition** Since the differential $d\Phi_e\colon \mathfrak{g} \to \mathfrak{h}$ is bijective, the inversion theorem 49.142 says that there exist open neighbourhood U_0 of e and V_0 of $\Phi(e)$ such that the restriction $\Phi\colon U_0 \to V_0$ is bijective.

Since U_0 and U_0^{-1} are open sets, the intersection is still open and still contain e. We consider $U = U_0 \cap U_0^{-1}$, and $V = \Phi(U)$. Now the restriction $\Phi\colon U \to V$ is still a bijection.

Let $\Gamma = \Phi^{-1}(e)$. Since Φ is not injective, this Γ can be a set.

(3a) **Γ is a subgroup** Let $g_1, g_2 \in \Gamma$. Since Φ is a group morphism, we have $\Phi(g_1g_2) = \Phi(g_1)\Phi(g_2) = ee = e$, so that $g_1g_2 \in \Gamma$.

(3b) **As an union** Let $g \in \Phi^{-1}(V)$. There exists a $g_0 \in U$ such that $\Phi(g_0) = \Phi(g)$. Recall that Φ is not injective; g_0 and g are element of G that are mapped on the same point in V. Since Φ is a morphism,

$$\Phi(gg_0^{-1}) = \Phi(g)\Phi(g_0^{-1}) = \Phi(g)\Phi(g_0)^{-1} = e. \tag{53.205}$$

Thus $gg_0^{-1} \in \Gamma$. Since $U^{-1} = U$ we have proved that if $g \in \Phi^{-1}(V)$, there exists $a \in \Gamma$ and $s \in U$ such that $as = g$. In other words,

$$\Phi^{-1}(V) = \bigcup_{a\in\Gamma} L_a(U). \tag{53.206}$$

EQooEZBDooBIfRJx

(3c) **Disjoint** It remain to be proven that the union (53.206) is disjoint. Let $a, b \in \Gamma$ such that $L_a(U) \cap L_b(U) \neq \varnothing$. We'll prove the $a = b$. Let $x \in L_a(U) \cap L_(U)$; we have

$$L_{b^{-1}}(x) \in L_{b^{-1}a}(U) \cap U. \tag{53.207}$$

So, with $c = b^{-1}a$, we have $c \in \Gamma$ and

$$L_c(U) \cap U \neq \varnothing. \tag{53.208}$$

EQooVNIPooBtixRi

Thus there exist $p_1, p_2 \in U$ such that $cp_1 = p_2$ (these are two ways to write an element of (53.208)). In particular we have $\Phi(cp_1) = \Phi(p_2)$. Since $c \in \Gamma$ we also have

$$\Phi(p_2) = \Phi(cp_1) = \Phi(c)\Phi(p_1) = \Phi(p_1). \tag{53.209}$$

But $\Phi\colon U \to V$ is injective. We deduce that $p_1 = p_2$, so that $c = e$ and $b^{-1}a = e$. Thus $a = b$.

□

PROPooIORNooLeuXPW

Proposition 53.84 ([880]).
Let G, H be smooth Lie groups. If $f\colon G \to H$ is a smooth morphism[36]*, then it has constant rank.*

Proof. Let G and H be Lie groups, let $f\colon G \to H$ be a smooth morphism and $g_0 \in G$. We will prove that df_{g_0} has the same rank as df_e.

Since f is a morphism we have

$$(f \circ L_{g_0})(g) = f(g_0g) = f(g_0)f(g) = L_{f(g_0)}\big(f(g)\big) = \big(L_{f(g_0)} \circ f\big)(g). \tag{53.210}$$

Thus we have $f \circ L_{g_0} = L_{f(g_0)} \circ f$. We take the differential on both sides. On the left hand side:

$$d(f \circ L_{g_0})_e = df_{L_{g_0}(e)} \circ (dL_{g_0})_e, \tag{53.211}$$

36. The map is smooth with respect to the manifold structure while being a morphism of groups.

and on the right hand side:

$$d\big(L_{f(g_0)} \circ f\big) = (dL_{f(g_0)})_{f(g_0)} \circ df_e \tag{53.212}$$

We know from lemma 52.9 that for every $g, h \in G$, the differential $(dL_h)_g$ is a vector space isomorphism. The composition of a linear map with a vector space isomorphism does not change the rank. Thus

$$\mathrm{rk}\Big[(dL_{f(g_0)})_{f(g_0)} \circ df_e\Big] = \mathrm{rk}(df_e), \tag{53.213}$$

and

$$\mathrm{rk}\big[df_g \circ (dL_{g_0})_e\big] = \mathrm{rk}(df_g), \tag{53.214}$$

so that $\mathrm{rk}(df_g) = \mathrm{rk}(df_e)$. □

Proposition 53.85 ([870])**.**
Let $f\colon G \to H$ be a smooth morphism of Lie groups[37]*. Then*

(1) The set $\ker(f)$ is a closed subgroup of G.

(2) The group $\ker(f)$ is a Lie group.

(3) The Lie algebra of $\ker(f)$ is $\ker(df)$.

Proof. In several parts.

(1) $\ker(f)$ **is a subgroup** If $g_1, g_2 \in \ker(f)$, then

$$f(g_1g_2) = f(g_1)f(g_2) = ee = e, \tag{53.215}$$

so that $g_1g_2 \in \ker(f)$. Moreover $f(e) = e$, so that $e \in \ker(f)$. This proves that $\ker(f)$ is a subgroup of G.

(2) $\ker(f)$ **is closed** We are going to prove that $G\backslash\ker(f)$ is open. For that we consider $g \in \overline{G\backslash\ker(f)}$ and we prove that there exists a neighbourhood of g contained in $G\backslash\ker(f)$.
Let $g \in G\backslash\ker(f)$. We know that $f(g) \neq e$ in H, so that we can consider an open neighbourhood V of $f(g)$ in H such that $e \notin V$.
Since f is continuous, $f^{-1}(V)$ is an open set which contains g. We have $f^{-1}(V) \cap \ker(f) = \varnothing$ because if $x \in f^{-1}(V) \cap \ker(f)$, we would have $f(x) \in V$ and $f(x) = e$ which is impossible. We deduce that $f^{-1}(V) \cap \ker(f) = \varnothing$. Thus $G\backslash\ker(f)$ contains $f^{-1}(V)$ which is a neighbourhood of g. We conclude that $G\backslash\ker(f)$ is open, so that $\ker(f)$ is closed.

(3) $\ker(f)$ **is a Lie group** From the Cartan theorem 53.64, we know that $\ker(f)$ is a Lie subgroup of G.

(4) **Lie algebra** Proposition 53.81 says that the Lie algebra of $\ker(f)$ is

$$\mathfrak{k} = \{X \in \mathfrak{g} \text{ tel que } \exp(tX) \in \ker(f)\, \forall t \in \mathbb{R}\}. \tag{53.216}$$

We have to prove that $\mathfrak{k} = \ker(df)$. Let $X \in \mathfrak{g}$. We have:

$$\begin{aligned}
X \in \ker(df) &\Leftrightarrow df(X) = 0 && \text{(53.217a)}\\
&\Leftrightarrow df(tX) = 0\, \forall t \in \mathbb{R} && \text{(53.217b)}\\
&\Leftrightarrow \exp\big(df(tX)\big) = e\, \forall t \in \mathbb{R} && \text{(53.217c)}\\
&\Leftrightarrow f\big(\exp(tX)\big) = e\, \forall t \in \mathbb{R} && \text{(53.217d)}\\
&\Leftrightarrow tX \in \ker(f)\, \forall t \in \mathbb{R} && \text{(53.217e)}\\
&\Leftrightarrow X \in \mathfrak{k}. && \text{(53.217f)}
\end{aligned}$$

SUBEQooVVVKooVHdrfC
SUBEQooCUSGooKTkJn

Justifications.

— For (53.217c). The exponential is a bijection between a neighbourhood of 0 in $\mathfrak{g}$ and a neighbourhood of e in G (proposition 53.47). Thus if $X \neq 0$ we can choose $t_1, t_2 \in \mathbb{R}$ such that $\exp(t_1X) \neq \exp(t_2X)$.

37. This is a morphism of groups which is smooth.

— For (53.217d). Lemma 53.38.

□

Proposition 53.86 ([870]).
A continuous morphism between Lie groups is smooth.

Proof. Let $\phi\colon G \to H$ be a continuous morphism. We consider

$$\Gamma = \{(g,h) \in G \times H \text{ tel que } \phi(g)h^{-1} = e\}. \tag{53.218}$$

(1) **Γ is a subgroup of $G \times H$** Let (g_1,h_1) and (g_2,h_2) be in Γ. We have $(g_1,h_1)(g_2,h_2) = (g_1g_2, h_1h_2)$ while

$$\phi(g_1g_2)(h_1h_2)^{-1} = \phi(g_1)\underbrace{\phi(g_2)h_2^{-1}}_{=e}h_1^{-1} = e \tag{53.219}$$

(2) **Γ is closed** We consider

$$\begin{aligned} f\colon G \times H &\to H \\ (g,h) &\mapsto \phi(g)h^{-1}. \end{aligned} \tag{53.220}$$

This is a continuous map because ϕ is continuous (as well as the product and the inverse in H) and $\Gamma = \ker(f)$. Thus Γ is closed.

(3) **Γ is a Lie subgroup** The Cartan theorem 53.64 says that Γ being a closed subgroup, it is a smooth Lie subgroup.

(4) **The projection** We introduce the projection

$$\begin{aligned} p\colon \Gamma &\to G \\ (g,h) &\mapsto g \end{aligned} \tag{53.221}$$

and we given some properties.

(4a) **Smooth** By proposition 49.47, the projection is smooth.

(4b) **Bijective** Let $(g_1,h_1),(g_2,h_2) \in \Gamma$ such that $p(g_1,h_1) = p(g_2,h_2)$. Then $g_1 = g_2$. But for each i we have $\phi(g_i) = h_i$, so $h_1 = \phi(g_1) = \phi(g_2) = h_2$.

(4c) **Morphism** Because

$$p\big((g_1,h_1)(g_2,h_2)\big) = p(g_1g_2, h_1h_2) = g_1g_2 = p(g_1,h_1)p(g_2,h_2). \tag{53.222}$$

For the inverse, $(g,h)^{-1} = (g^{-1},h^{-1})$, so that

$$p\big((g,h)^{-1}\big) = p(g^{-1},h^{-1}) = g^{-1} = p(g,h)^{-1}. \tag{53.223}$$

So p is a group morphism.

(5) **Smooth diffeomorphism** The differential of p at (e,e) is the map[38] $dp_{(e,e)}\colon \mathfrak{g} \times \mathfrak{h} \to \mathfrak{g}$ given by

$$dp_{(e,e)}(X,Y) = \frac{d}{dt}\Big[p(e^{tX},e^{tX})\Big]_{t=0} = \frac{d}{dt}\Big[e^{tX}\Big]_{t=0} = X, \tag{53.224}$$

so that $dp_{(e,e)}$ is surjective on $\mathfrak{g}$. Its rank is the dimension of $\mathfrak{g}$.
Proposition 53.84 shows that p has constant rank. We know that this rank is maximal, so that the theorem 49.142 thus says that p is a local smooth diffeomorphism everywhere.
Since p is invertible, it is a global smooth diffeomorphism.

(6) **Conclusion** We have

$$\phi = \mathtt{proj}_2 \circ p^{-1}. \tag{53.225}$$

Since p^{-1} and $\mathtt{proj}_2$ are smooth, the map ϕ is smooth.

38. Proposition 52.8.

□

THOooZAEYooXCdxKI

Theorem 53.87 ([870]).
Let G and H be Lie groups. We suppose that G is connected and simply connected. Let $\rho\colon \mathfrak{g} \to \mathfrak{h}$ be a morphism of Lie algebra.

There exists an unique smooth morphism $f\colon G \to H$ such that $df_e = \rho$.

Proof. Let $\mathfrak{s}$ be the graph of ρ:

$$\mathfrak{s} = \{(X, Y) \in \mathfrak{g} \oplus \mathfrak{h} \text{ tel que } Y = \rho(X)\}. \tag{53.226}$$

Since ρ is linear, the set $\mathfrak{s}$ is a vector subspace of $\mathfrak{g} \oplus \mathfrak{h}$. Moreover $\mathfrak{s}$ is a Lie subalgebra. Indeed, if $Y_i = \rho(X_i)$ we have

$$[Y_1, Y_2] = \big[\rho(X_1), \rho(X_2)\big] = \rho\big([X_1, X_2]\big) \tag{53.227}$$

and then

$$\big[(X_1, Y_1), (X_2, Y_2)\big] = \big([X_1, X_2], [Y_1, Y_2]\big) = \Big([X_1, X_2], \rho\big([X_1, X_2]\big)\Big) \in \mathfrak{s}. \tag{53.228}$$

The Chevalley theorem 53.82 says that there exists a unique connected Lie subgroup S of $G \times H$ whose Lie algebra is $\mathfrak{s}$.

We consider the inclusion $\iota\colon S \to G \times H$ and the map

$$\begin{aligned} \varphi = \mathtt{proj}_1 \circ \iota\colon S &\to G \\ (g, h) &\mapsto g. \end{aligned} \tag{53.229}$$

This is a smooth morphism, so that its differential is a morphism of Lie algebra.

(1) $d\varphi_{(e,e)}$ **is bijective** We study the map

$$d\varphi_{(e,e)} = (d\,\mathtt{proj}_1)_{(e,e)} \circ d\iota_{(e,e)}\colon \mathfrak{s} \to \mathfrak{g}. \tag{53.230}$$

(1a) **Injection** Let (X_1, Y_1) and (X_2, Y_2) be elements of $\mathfrak{s}$ such that $d\varphi_{(e,e)}(X_1, Y_1) = d\varphi_{(e,e)}(X_2, Y_2)$. Since $d\iota_{(e,e)}$ is the identity, we have

$$(d\,\mathtt{proj}_1)_{(e,e)}(X_1, Y_2) = (d\,\mathtt{proj}_1)_{(e,e)}(X_2, Y_2). \tag{53.231}$$

We deduce $X_1 = X_2$. By definition of $\mathfrak{s}$ we also have $Y_i = \rho(X_i)$ and then

$$Y_1 = \rho(X_1) = \rho(X_2) = Y_2. \tag{53.232}$$

(1b) **Surjection** Let $X \in \mathfrak{g}$. We consider $Y = \rho(X)$, so that $(X, Y) \in \mathfrak{s}$ and

$$d\varphi_{(e,e)}(X, Y) = X. \tag{53.233}$$

(2) **Diffeomorphism** We know that $\varphi\colon S \to G$ is a smooth morphism and that $d\varphi_{(e,e)}$ is bijective. Theorem 49.142 concludes that φ is a local smooth diffeomorphism. So there exists a neighbourhood V_1 of (e, e) in S and a neighbourhood V_2 of e in G such that $V_2 = \varphi(V_1)$. By proposition 52.5, the map φ is surjective on the connected component of e in G. Since G is connected, φ is a global smooth diffeomorphism.

(3) **The map Φ** We define

$$\begin{aligned} \Phi\colon G &\to H \\ \Phi &= \mathtt{proj}_2 \circ \varphi^{-1}. \end{aligned} \tag{53.234}$$

As composition of smooth morphisms, the map Φ is a smooth morphism. We have to prove that $d\Phi_e = \rho$.

(4) **The map $d\varphi_e^{-1}$** Let $X \in \mathfrak{g}$. We have

$$d\varphi_{(e,e)}\big(X,\rho(X)\big) = (d\,\mathtt{proj}_1)_e \circ d\iota_{(e,e)}\big(X,\rho(X)\big) \tag{53.235a}$$
$$= (d\,\mathtt{proj}_1)_e\big(X,\rho(X)\big) \tag{53.235b}$$
$$= X. \tag{53.235c}$$

The proposition 49.77 concludes that

$$(d\varphi^{-1})_e(X) = \big(X,\rho(X)\big). \tag{53.236}$$

(5) **Conclusion** Now we conclude with the differential of Φ; if $X \in \mathfrak{g}$ we have

$$d\Phi_e(X) = (d\,\mathtt{proj}_2)_{\varphi^{-1}(e)} \circ (d\varphi^{-1})_e(X) \tag{53.237a}$$
$$= (d\,\mathtt{proj}_2)_{\varphi^{-1}(e)}\big(X,\rho(X)\big) \tag{53.237b}$$
$$= \rho(X) \tag{53.237c}$$

and Φ is the map we were searching for.

□

PROPooCRKMooIIbKUM

Proposition 53.88.
Let G be a Lie group and H, a Lie subgroup of G ($\mathfrak{g}$ and $\mathfrak{h}$ are the corresponding Lie algebras). We suppose that H has at most a countable number of connected components. Then

$$\mathfrak{h} = \{X \in \mathfrak{g} : \forall t \in \mathbb{R}, e^{tX} \in H\} \tag{53.238}$$

Proof. We will once again use the lemma 53.61 with $\mathfrak{n} = \mathfrak{h}$ and $\mathfrak{m}$, a complementary vector space of $\mathfrak{h}$ in $\mathfrak{g}$. We define

$$V = \exp\mathcal{U}_m \exp\mathcal{U}_h$$

where $\mathcal{U}_m$ and $\mathcal{U}_h$ are the sets given by the lemma. We consider on V the induced topology from G. If we define

$$\mathcal{A} = \{A \in \mathcal{U}_m : e^A \in H\},$$

we have

$$H \cap V = \bigcup_{A\in\mathcal{A}} e^A e^{\mathcal{U}_h}. \tag{53.239}$$

eq:union_A

First, the definition of V makes clear that the elements of the form $\exp A \exp\mathcal{U}_h$ are in V. They are also in H because $\exp A \in H$ (definition of $\mathcal{A}$) and $\exp\mathcal{U}_h$ still by definition. In order to see the inverse inclusion, let us consider a $h \in H \cap V$. We know that

$$(A,B) \to \exp A \exp B \tag{53.240}$$

eq:AB_to_exp

is a diffeomorphism between $\mathcal{U}_m \times \mathcal{U}_h$ and a neighbourhood of e in G which we called V. Thus any element of V (*a fortiori* in $V \cap H$) can be written as $\exp A \exp B$ with $A \in \mathcal{U}_m$ and $B \in \mathcal{U}_h$. Then $h = e^A e^B$ for some $A \in \mathcal{U}_m$, $B \in \mathcal{U}_h$. Since H is a group and $e^B \in H$, in order the product to belongs to H, e^A must lies in H: $A \in \mathcal{A}$.

rem:union_disj

Remarque 53.89.
Note that since (53.240) is diffeomorphic, the union in right hand side of (53.239) is disjoint. Each member of this union is a neighbourhood in H because it is a set $h\exp\mathcal{U}_h$ where $\exp\mathcal{U}_h$ is a neighbourhood of e in H.

Now we consider the map $\pi\colon V \to \mathcal{U}_m$,

$$\pi(e^X e^Y) = X$$

if $X \in \mathcal{U}_m$ and $Y \in \mathcal{U}_h$. This is a continuous map which sends $H \cap V$ into $\mathcal{A}$. The identity component of $H \cap V$ (in the sense of topology of V) is sent to a countable subset of $\mathcal{U}_m$. Indeed by remark 53.89, identity component of $H \cap V$ is only one of the terms in the union (53.239), namely $A = 0$. But we know that $\pi^{-1}(o) = \exp\mathcal{U}_h$, thus $\exp\mathcal{U}_h$ is the identity component of $H \cap V$ for the topology of V. □

Now we take back our example with $G = S^1 \times S^1$, $H = \gamma(\mathbb{R})$. In this case, the theorem doesn't works. Let us see why as deep as possible. We have $\mathfrak{g} = \mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$ and $\mathfrak{h} = \mathbb{R}$, a one-dimensional vector subspace of $\mathfrak{g}$. ($\mathfrak{h}$ is a "direction " in $\mathfrak{g}$) First, we build the neighbourhood V of 0 in $\mathfrak{g}$. It is standard to require that exp is diffeomorphic between V and an open around $(1,1) \in S^1 \times S^1$. It also must satisfy $e^{V \cap \mathfrak{h}} = e^V \cap H$. This second requirement is impossible.

Intuitively. We can see $V \subset \mathfrak{g}$ as a little disk tangent to the torus. The exponential map deposits it on the torus, as well that e^V covers a little area on G. Then $e^V \cap H$ is one of these amazing open subset of Γ which are dense in a certain domain of G.

On the other hand, $V \cap \mathfrak{h}$ is just a little vector in $\mathfrak{h}$; the exponential deposits it on a small line in G. This is not the same at all.

Let $W_0 \subset \mathfrak{g}$ be a neighbourhood of 0 which is in bijection with an open around e in G. We consider N_0, an open subset of H such that $N_0 \subset W_0$ and N_0 is in bijection with N_e, a neighbourhood of e in G. Until here, no problems. But now the proof says that there exists an open U_e in G such that $N_e = U_e \cap H$. This is false in our case. Indeed, $N_e = e^{N_0}$ is just a segment in G while any subset of G of the form $U_e \cap H$ is an "amazing" open.

So we see that deeply, the obstruction for a Lie subgroup to be a topological Lie subgroup resides in the fact that the topology of a submanifold is *more* than the induced topology, so that we can't automatically find the open U_e in G.

Note that two groups which have the same Lie algebra are not necessarily isomorphic. For example the sphere S^2 and $\mathbb{R}^2$ both have $\mathbb{R}^2$ as Lie algebra. But two groups with same Lie algebra are locally the same. More precisely, we have the following lemma.

Lemma 53.90.
If G is a Lie group and H, a topological subgroup of G with the same Lie algebra ($\mathfrak{h} = \mathfrak{g}$), then there exists a common neighbourhood A of e of G and G on which the products in G and H are the same.

Proof. The exponential is a diffeomorphism between $U \subset \mathfrak{g}$ and $V \subset G$ and between $U' \subset \mathfrak{h}$ and $W \subset H$ (obvious notations). We consider an open $\mathcal{O} \subset \mathfrak{h}$ such that $\mathcal{O} \subset U \subset U'$. The exponential is diffeomorphic from $\mathcal{O}$ to a certain open A in G and H. Since H is a subgroup of G, the product $e^X e^Y$ of elements in A is the same for H and G. (cf error 88.5) □

Under the same assumptions, we can say that H contains at least the whole G_0 because it is generated by any neighbourhood of the identity. Since H is a subgroup, the products keep in H.

For a semisimple Lie group, the Lie algebras $\partial(\mathfrak{g})$ and $\mathrm{ad}(\mathfrak{g})$ are the same. Then $\mathrm{Int}(\mathfrak{g})$ contains at least the identity component of $\mathrm{Aut}(\mathfrak{g})$. Since $\mathrm{Int}(\mathfrak{g})$ is connected, for a semisimple group, it is the identity component of $\mathrm{Aut}(\mathfrak{g})$.

PropCompactKilling

Proposition 53.91.
Let $\mathfrak{g}$ be a real Lie algebra.

1. *If $\mathfrak{g}$ is semisimple, then $\mathfrak{g}$ is compact if and only if the Killing form is strictly negative definite.*
2. *If it is compact then it is a direct sum*

$$\mathfrak{g} = \mathcal{Z} \oplus [\mathfrak{g}, \mathfrak{g}] \tag{53.241}$$

where $\mathcal{Z}$ is the center of $\mathfrak{g}$ and the ideal $[\mathfrak{g}, \mathfrak{g}]$ is compact and semisimple.

Proof. If the Killing form is nondegenerate. We consider $\mathfrak{g}$, a Lie algebra whose Killing form is strictly negative definite. Up to some dilatations (and a sign), this is the euclidian metric. Then $O(B)$, the group of linear transformations which leave B unchanged is compact in the topology of $\mathrm{GL}(\mathfrak{g})$: this is almost the rotations. From equation (52.10), $\mathrm{Aut}(\mathfrak{g}) \subset O(B)$. With this, $\mathrm{Aut}(\mathfrak{g})$ is closed in a compact, then it is compact. Then $\mathrm{Int}(\mathfrak{g})$ is closed in $\mathrm{Aut}(\mathfrak{g})$ –here is the assumption of semi-simplicity– and $\mathrm{Int}(\mathfrak{g})$ is compact.

If $\mathfrak{g}$ is compact. Since $\mathfrak{g}$ is compact, $\mathrm{Int}(\mathfrak{g})$ is compact in the topology of $\mathrm{Aut}(\mathfrak{g})$; then there exists an $\mathrm{Int}(\mathfrak{g})$-invariant quadratic form Q. In a suitable basis $\{X_1, \dots, X_n\}$ of $\mathfrak{g}$, we can write this form

as

$$Q(X) = \sum x_i^2$$

for $X = \sum x_i X_i$. In this basis the elements of $\mathrm{Int}(\mathfrak{g})$ are orthogonal matrices and the matrices of $\mathrm{ad}(\mathfrak{g})$ are skew-symmetric matrices (the Lie algebra of orthogonal matrices). Let us consider a $X \in \mathfrak{g}$ and denote by $a_{ij}(X)$ the matrix of $\mathrm{ad}(X)$. We have

$$B(X,X) = \mathrm{Tr}(\mathrm{ad}\, X \circ \mathrm{ad}\, X) = \sum_i \sum_j a_{ij}(X)a_{ji}(X) = -\sum_{ij} a_{ij}(X)^2 \leqslant 0. \tag{53.242}$$

Then the Killing form is negative definite [39]. On the other hand, $B(X,X) = 0$ implies $\mathrm{ad}(X) = 0$ and $X \in \mathcal{Z}(\mathfrak{g})$. Thus $\mathfrak{g}^\perp \subset \mathcal{Z}$. If $\mathfrak{g}$ is semisimple, this center is zero; this conclude the first item of the proposition.

Now $\mathcal{Z}$ is an ideal and corollary 51.52 decomposes $\mathfrak{g}$ as

$$\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{g}'. \tag{53.243}$$

Let us suppose that the restriction of B to $\mathfrak{g}' \times \mathfrak{g}'$ is actually the Killing form on $\mathfrak{g}'$ (we will prove it below). Then the Killing form on $\mathfrak{g}'$ is strictly negative definite; then $\mathfrak{g}'$ is compact.

Now we prove that the Killing form on $\mathfrak{g}$ descent to the Killing form on $\mathfrak{g}'$. Remark that $\mathcal{Z}$ is invariant under all the automorphism. Indeed consider $Z \in \mathcal{Z}$, i.e. $[X,Z] = 0$. If σ is an automorphism,

$$[X, \sigma Z] = \sigma[\sigma^{-1}X, Z] = 0.$$

Here the difference between $\mathrm{Int}(\mathfrak{g})$ and $\mathrm{Aut}(\mathfrak{g})$ is the fact that $\mathrm{Int}(\mathfrak{g})$ is compact; then we can construct a $\mathrm{Int}(\mathfrak{g})$-invariant quadratic form Q, but not a $\mathrm{Aut}(\mathfrak{g})$-invariant one. We consider an orthogonal complement (with respect to Q) $\mathfrak{g}'$ of $\mathcal{Z}$:

$$\mathfrak{g} = \mathfrak{g}' \oplus_\perp \mathcal{Z}. \tag{53.244}$$

The algebra $\mathfrak{g}'$ is also invariant because for any $Z \in \mathcal{Z}$,

$$Q(Z, \sigma X) = Q(\sigma^{-1}(Z), X) = 0.$$

It is also clear that $\mathcal{Z}$ is invariant under $\mathrm{ad}\,\mathfrak{g}$ because $(\mathrm{ad}\, X)Z = 0$. Finally $\mathfrak{g}'$ is invariant as well under $\mathrm{ad}(\mathfrak{g})$. Indeed $a \in \mathrm{ad}(\mathfrak{g})$ can be written as $a = a'(0)$ for a path $a(t) \in \mathrm{Int}(\mathfrak{g})$. We identify $\mathfrak{g}$ and his tangent space (as vector spaces),

$$aX = \frac{d}{dt}\Big[a(t)X\Big]_{t=0}.$$

If $X \in \mathfrak{g}'$, $a(t)X \in \mathfrak{g}'$ for any t because $\mathfrak{g}'$ is invariant under $\mathrm{Int}(\mathfrak{g})$ [40]. Thus $a(t)X$ is a path in $\mathfrak{g}'$ and his derivative is a vector in $\mathfrak{g}'$.

All this make $\mathfrak{g}'$ an ideal in $\mathfrak{g}$; then the Killing form descent by lemma 51.20. Now if $X \in \mathfrak{g}$, we have

$$B(X,X) = \mathrm{Tr}(\mathrm{ad}\, X \circ \mathrm{ad}\, X) = \sum_{ij} a_{ij}(X)a_{ji}(X) = -\sum_{ij} a_{ij}(X)^2; \tag{53.245}$$

then $B(X,X) \leqslant 0$ and the equality holds if and only if $\mathrm{ad}\, X = 0$ i.e. if and only if $X \in \mathcal{Z}$. Thus B is strictly negative definite on $\mathfrak{g}'$.

Up to now we have proved that $\mathfrak{g}'$ is semisimple (because B is nondegenerate) and compact (because B is strictly negative definite).

39. Here we use "negative definite" and "*strictly* negative definite"; in some literature, the terminology is slightly different and one says "*semi* negative definite" and "negative definite".

40. As physical interpretation, if something is invariant under a group of transformations, it is invariant under the infinitesimal transformations as well.

It remains to be proved that $\mathfrak{g}' = [\mathfrak{g},\mathfrak{g}] = \mathscr{D}(\mathfrak{g})$. From corollary 51.52, $\mathscr{D}\mathfrak{g}$ has a complementary $\mathfrak{a}$ which is also an ideal: $\mathfrak{g} = \mathscr{D}\mathfrak{g} + \mathfrak{a}$. Then $[\mathfrak{g},\mathfrak{a}] \subset \mathscr{D}\mathfrak{g}$ and $[\mathfrak{g},\mathfrak{a}] \subset \mathfrak{a} \cap \mathscr{D}\mathfrak{g} : \{0\}$. Then $\mathfrak{a} \subset \mathcal{Z}$, so that

$$\mathfrak{g} = \mathcal{Z} + \mathscr{D}\mathfrak{g} \qquad \text{(non direct sum)}. \tag{53.246}$$

cor:G_Z_R

Now we have to prove that the sum is actually direct. The ideal $\mathcal{Z}$ has a complementary ideal $\mathfrak{b}$: $\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{b}$ and

$$\mathscr{D}\mathfrak{g} = [\mathfrak{g},\mathfrak{g}] \subset \underbrace{[\mathfrak{g},\mathcal{Z}]}_{=0} + [\mathfrak{g},\mathfrak{b}] \subset \mathfrak{b}.$$

Then $\mathscr{D}\mathfrak{g} \subset \mathfrak{b}$ which implies that $\mathscr{D}\mathfrak{g} \cap \mathcal{Z} = \{0\}$ because the sum $\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{b}$ is direct. Then the sum (53.246) is direct.

□

Proposition 53.92.
A real Lie algebra $\mathfrak{g}$ is compact if and only if one can find a compact Lie group G which Lie algebra is isomorphic to $\mathfrak{g}$.

prop:alg_grp_compact

Proof. Direct sense. Since $\mathfrak{g}$ is compact, $\mathfrak{g} = \mathcal{Z} \oplus \mathscr{D}\mathfrak{g}$ with $\mathscr{D}\mathfrak{g} = \mathfrak{g}'$ compact and semisimple; in particular, the center of $\mathfrak{g}'$ is $\{0\}$. Since $\mathcal{Z}$ is compact and abelian, it is isomorphic to the torus $S^1 \times \ldots \times S^1$. Since $\mathfrak{g}'$ is compact, $\mathrm{Int}(\mathfrak{g}')$ is compact, but the Lie algebra if $\mathrm{Int}(\mathfrak{g}')$ is –by definition– $\mathrm{ad}(\mathfrak{g}')$. The center of a semisimple Lie algebra is zero; then $\operatorname{ad} X' = 0$ implies $X = 0$ (for $X \in \mathfrak{g}'$). Then ad is an isomorphism between $\mathfrak{g}'$ and $\operatorname{ad}\mathfrak{g}'$.

All this shows that –up to isomorphism– $\mathcal{Z}$ and $[\mathfrak{g},\mathfrak{g}]$ are Lie algebras of compact groups. We know from lemma 49.122 that the Lie algebra of $G \times H$ is $\mathfrak{g} \oplus \mathfrak{h}$. Thus, here, $\mathfrak{g}$ is the Lie algebra of the compact group $S^1 \times \ldots \times S^1 \times \mathrm{Int}(\mathfrak{g})$.

Reverse sense. We consider a compact group G and we have to see the its Lie algebra $\mathfrak{g}$ is compact. If G is connected, Ad_G is an analytic homomorphism from G to $\mathrm{Int}(\mathfrak{g})$. If G is not connected, the Lie algebra of G is T_eG_0 (G_0 is the identity component of G) where G_0 is connected and compact because closed in a compact. □

Proposition 53.93.
Let $\mathfrak{g}$ be a real Lie algebra and $\mathcal{Z}$, the center of $\mathfrak{g}$. We consider $\mathfrak{k}$, a compactly embedded in $\mathfrak{g}$. If $\mathfrak{k} \cap \mathcal{Z} = \{0\}$ then the Killing form of $\mathfrak{g}$ is strictly negative definite on $\mathfrak{k}$.

prop:K_Z_Killing

Proof. Let B be the Killing form on $\mathfrak{g}$ and K the analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. By assumption, K is a compact Lie subgroup of $\mathrm{GL}(\mathfrak{g})$. Then there exists a quadratic form on $\mathfrak{g}$ invariant under K, and a basis in which the endomorphisms $\mathrm{ad}_{\mathfrak{g}}(T)$ for $T \in \mathfrak{k}$ are skew-symmetric because the matrices of K are orthogonal. If the matrix of $\operatorname{ad} T$ is (a_{ij}), then

$$B(T,T) = \sum_{ij} a_{ij}(T)a_{ji}(T) = -\sum_{ij} a_{ij}^2(T) \leqslant 0, \tag{53.247}$$

and the equality hold only if $\operatorname{ad} T = 0$ i.e. if $T \in \mathcal{Z}$. From the assumptions, $\mathfrak{k} \cap \mathcal{Z} = \{0\}$; then $B(T,T) = 0$ if and only if $T = 0$. □

53.9 Representations from the Lie algebra to the Lie group

PROPooXCGMooKlJlwp

Proposition 53.94 ([844])**.**
Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. Let (ρ, V) be a smooth representation of G. We consider the map

$$\begin{aligned} s&\colon \mathfrak{g} \to \mathrm{End}(V) \\ s(X)v &= \frac{d}{dt}\Big[\rho(e^{tX})v\Big]_{t=0}. \end{aligned} \tag{53.248}$$

We have the equality

$$\rho(e^{tX}) = e^{ts(X)} \tag{53.249}$$

as operators on V.

Proof. Let $X \in \mathfrak{g}$. We define $M(t) = \rho(e^{tX})$ and $N(t) = e^{ts(X)}$. These are maps from $\mathbb{R}$ to $\mathrm{End}(V)$; the proposition 15.72 helps to derive them.

For N we immediately have

$$N'(t) = s(X)e^{ts(X)}. \tag{53.250}$$

For M, we have few more prudence. We fix $\epsilon > 0$ and we write (thanks to proposition 15.70)

$$M(t+\epsilon) = \rho(e^{(t+\epsilon)X}) = \rho(e^{tX}e^{\epsilon X}) = \rho(e^{tX})e^{\epsilon X}. \tag{53.251}$$

Now for the differential quotient,

$$\frac{M(t+\epsilon) - M(t)}{\epsilon} = \rho(e^{tX})\frac{\rho(e^{\epsilon X}) - \mathrm{Id}}{\epsilon}. \tag{53.252}$$

If we compute the limit $\epsilon \to 0$, the second factor goes, by definition to $s(X)$. So

$$M'(t) = \rho(e^{tX})s(X) = s(X)M(t). \tag{53.253}$$

So M and N satisfy the same differential equation

$$\begin{cases} y' = s(X)y(t) & \text{(53.254a)} \\ y(0) = \mathbb{1} & \text{(53.254b)} \end{cases}$$

for the function $y\colon \mathbb{R} \to \mathrm{End}(V)$. This is a work for Cauchy-Lipschitz, theorem 17.40. So we define $f(t,m) = s(X)m$ and we show that this is Lipschitz with respect to m :

$$\|f(t_0,m_0) - f(t,m)\| = \|s(X)(m_0 - m)\| \leqslant \|s(X)\|\|m_0 - m\|. \tag{53.255a}$$

So the function f is Lipschitz with respect to m with a Lipschitz constant bounded by $\|s(X)\|$.

The unicity part of Cauchy-Lipschitz shows that $M(t) = N(t)$ for every t for which the expressions make sense. □

THOooLVSNooOpzYgO

Theorem 53.95 ([844, 1])**.**
Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. If (ρ, V) is a finite dimensional representation of G, the map

$$\begin{aligned} s\colon \mathfrak{g} &\to \mathrm{End}(V) \\ s(X)v &= \frac{d}{dt}\Big[\rho(e^{tX})v\Big]_{t=0} \end{aligned} \tag{53.256}$$

is a representation[41] *of $\mathfrak{g}$ on V.*

Proof. We choose a basis of V and we consider the matrix representation associated with ρ, namely[42]

$$\tilde{\rho}\colon G \to \mathrm{GL}(n,\mathbb{C}). \tag{53.257}$$

We have

$$\rho\big(\exp(tX)\big)v = \sum_{kl} \tilde{\rho}\big(\exp(tX)\big)_{kl} v_l e_k. \tag{53.258}$$

41. Representation of a Lie algebra, definition 51.2.
42. If you want details, this is $\tilde{\rho} = \psi^{-1} \circ \rho$ where $\psi\colon \mathbb{M}(n,\mathbb{C}) \to \mathrm{GL}(V)$ is defined in 4.71.

and then

$$s(X)v = \frac{d}{dt}\Big[\sum_{kl}\tilde{\rho}\big(\exp(tX)\big)_{kl}v_k e_l\Big]_{t=0} \tag{53.259a}$$
$$= \sum_{kl}\left(\frac{d}{dt}\Big[\tilde{\rho}\big(\exp(tX)\big)\Big]_{t=0}\right)_{kl} v_k e_l \tag{53.259b}$$
$$= \sum_{kl}\left(\frac{d}{dt}\Big[\exp\big(td\tilde{\rho}_e(X)\big)\Big]_{t=0}\right)_{kl} v_k e_l \qquad \text{lem. 53.38} \tag{53.259c}$$
$$= \sum_{kl} d\tilde{\rho}_e(X)_{kl}v_k e_l \tag{53.259d}$$
$$= d\tilde{\rho}_e(X)v. \tag{53.259e}$$

We keep in mind that, expressed with matrices, we have

$$s(X)v = d\tilde{\rho}_e(X)v. \tag{53.260}$$

(1) **s is linear** This is equation (53.260).

(2) **Lie bracket** We have to prove that $s\big([X,Y]\big)v = [s(X), s(Y)]v$. On the left-hand side, the bracket is the commutator in $\mathrm{GL}(V)$. We start from the expression (53.259b) and we focus on what lies inside the derivative.

The trick is to use the formulas of lemma 53.54 to enter $d\tilde{\rho}_e$ inside the commutator via the lemma 53.38. Here is the computation:

$$\tilde{\rho}\big(\exp(t[X,Y])\big) = \tilde{\rho}\Big(\alpha(t)\exp(\sqrt{t}X)\exp(\sqrt{t}Y)\exp(-\sqrt{t}X)\exp(-\sqrt{t}Y)\Big) \tag{53.261a}$$
$$= \tilde{\rho}\big(\alpha(t)\big)\tilde{\rho}\big(\exp(\sqrt{t}X)\big)\tilde{\rho}\big(\exp(\sqrt{t}Y)\big) \tag{53.261b}$$
$$\tilde{\rho}\big(\exp(-\sqrt{t}X)\big)\tilde{\rho}\big(\exp(-\sqrt{t}Y)\big)$$
$$= \tilde{\rho}\big(\alpha(t)\big)\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big) \tag{53.261c}$$
$$\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)$$

Thus, using the Leibniz rule, ans the properties $\alpha(0) = e$, $\alpha'(0) = 0$ we have

$$\frac{d}{dt}\Big[\tilde{\rho}\Big(\exp\big(t[X,Y]\big)\Big)\Big]_{t=0} = \frac{d}{dt}\Big[\tilde{\rho}\big(\alpha(t)\big)\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)\dots\Big]_{t=0} \tag{53.262a}$$
$$= \frac{d}{dt}\Big[\tilde{\rho}\big(\alpha(t)\big)\Big]_{t=0} \tag{53.262b}$$
$$+ \tilde{\rho}\big(\alpha(0)\big)\frac{d}{dt}\Big[\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)\dots\Big]_{t=0}$$
$$= \frac{d}{dt}\Big[\exp\big(\sqrt{t}d\tilde{\rho}_e(X)\big)\exp\big(\sqrt{t}d\tilde{\rho}_e(Y)\big) \tag{53.262c}$$
$$\exp\big(-\sqrt{t}d\tilde{\rho}_e(X)\big)\exp\big(-\sqrt{t}d\tilde{\rho}_e(Y)\big)\Big]_{t=0}$$
$$= \big[d\tilde{\rho}_e(X), d\tilde{\rho}_e(Y)\big] \qquad \text{eq. (53.136)} \tag{53.262d}$$
$$= \big[s(X), s(Y)\big] \qquad \text{eq. (53.260).} \tag{53.262e}$$

□

53.10 Old stuff

I move here the theorems which are not really well proven and which are replaced by better ones.

53.11 Cosets

We consider G, a Lie group and H, a closed subgroup. Then from theorem 53.70, there exists an unique analytic structure on H for which H is a topological Lie subgroup of G. We naturally consider this structure on H. We also consider $\mathfrak{g}$ and $\mathfrak{h}$, the Lie algebras of G and H, and $\mathfrak{m}$ be a subspace of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$.

Now we will study the structure of the coset space G/H on which we put the topology such that π is continuous and open; this is the **natural topology**. pg:natur_topo As notations, we define $p_0 = \pi(e)$ and $\psi\colon \mathfrak{m} \to G$, the restriction to $\mathfrak{m}$ of the exponential.

Lemma 53.96.
The dimension of G/H is $\dim(G/H) = \dim G - \dim H$. lem:dim_G_H

Proof. We decompose the Lie algebra $\mathfrak{g}$ as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, and we will see that there exists a real vector space isomorphism $\psi\colon T_{[e]}(G/H) \to \mathfrak{m}$ given by

$$\psi(X) = \frac{d}{dt}\Big[e^{m(t)}\Big]_{t=0} \tag{53.263}$$

if $X(t) = [g(t)]$ with $g(t) = e^{m(t)}e^{h(t)}$ where $m(t) \in \mathfrak{m}$ and $h(t) \in \mathfrak{h}$ (the existence of such a decomposition in reasonably small neighbourhood of e is given by lemma 53.61). The fact that ψ is surjective is clear. The injectivity is also easy: $\psi(X) = 0$ implies that $\exp m(t)$ is a constant. Thus

$$X = \frac{d}{dt}\Big[[cst\, e^{h(t)}]\Big]_{t=0} = \frac{d}{dt}\Big[[cst]\Big]_{t=0} = 0.$$

□

lem:vois_U

Lemma 53.97.
There exists a neighbourhood U of 0 in $\mathfrak{m}$ such that

(1) ψ is homeomorphic on U,

(2) π sends homeomorphically $\psi(U)$ on a neighbourhood of p_0 in G/H.

Proof. By lemma 53.61, we consider bounded, open and connected neighbourhoods $\mathcal{U}_m$ and $\mathcal{U}_h$ of 0 in $\mathfrak{m}$ and $\mathfrak{h}$ such that

$$\phi\colon (A, B) \to e^A e^B$$

is a diffeomorphism from $\mathcal{U}_m \times \mathcal{U}_h$ to an open neighbourhood of e in G. Since H has the induced topology from G, we can find a neighbourhood V of e in G such that $V \cap H = \exp \mathcal{U}_h$.

Now we take U, a compact neighbourhood of 0 in $\mathcal{U}_m$ such that

$$e^{-U}e^{U} \subset V. \tag{53.264}$$

eq:UUV

So, ψ is an homeomorphism from U to $\psi(U)$. Indeed for $X \in U$, $\psi(X) = e^X = \phi(X, 0)$ and ϕ is diffeomorphic.

On the other hand, π is bijective on $\psi(U)$. In order to see that it is injective, let us consider X, $Y \in U$ such that $\pi(e^X) = \pi(e^Y)$. Then $\exp X$ and $\exp Y$ are in the same class with respect to H: $\exp X \in [\exp Y]$. Then $\exp(-X)\exp Y \in H$, and reversing the role[43] of X and Y, $\exp(-Y)\exp X \in H$. Since $X', X'' \in U$ and (53.264),

$$e^{-Y}e^{X} \in V \cap H.$$

Then there exists a Z in $\mathcal{U}_h$ such that $\exp X = \exp Y \exp Z$, but U is a subset of $\mathcal{U}_m$ (so that $(A, B) \to e^A e^B$ is diffeomorphic), then $X = Y$ and $Z = 0$.

Since π is bijective on $\psi(U)$, it is homeomorphic because the topology is build in order for π to be open and continuous.

On a third hand, $U \times \mathcal{U}_h$ is a neighbourhood of $(0, 0)$ in $\mathcal{U}_m \times \mathcal{U}_h$, so that $e^U e^{\mathcal{U}_h}$ is a neighbourhood of e in G. Since π is open, $\pi(\exp U \exp \mathcal{U}_h) = \pi(\psi(U))$ is a neighbourhood of p_0 in G/H. □

43. Anglais ?

Let N_0 be the interior of $\pi(\psi(U))$ and $\{X_1, \dots, X_r\}$ a basis of $\mathfrak{m}$. If $g \in G$, we looks at the map

$$\pi(g \cdot e^{x_1 X_1 + \dots + x_r X_r}) \to (x_1, \dots, x_r).$$

This is an homeomorphism from $g \cdot N_0$ to an open subset of $\mathbb{R}^r$ because π is homeomorphic from U. With this chart, G/H is an analytic manifold and moreover if $x \in G$, the map

$$\tau(x) \colon [y] \to [xy] \tag{53.265}$$

eq:tau_x_y

is an analytic diffeomorphism of G/H. Let us prove it. If we consider $[x] \in G/H$, we can write $x = gm$ for a certain $m \in \psi(U)$. Hence the chart around $[x]$ will be around $[gm] = [ge^{x_1 X_1 + \dots + x_r X_r}]$ (in other word, we can find an open set around $[x]$ on which can be parametrised so). We can forget the g because the action is a diffeomorphism. Then we looks at the chart $\varphi \colon G/H \to \mathbb{R}^r$, $\varphi[e^{x_1 X_1 + \dots + x_r X_r}] = (x_1, \dots, x_r)$. The map (53.265) makes

$$(y_1, \dots, y_r) \to (CBH_1(x_1, \dots, x_r, y_1, \dots y_r), \dots, CBH_r(x_1, \dots, x_r, y_1, \dots y_r)). \tag{53.266}$$

But CBH is a diffeomorphism.

thoHelgasonA2

Theorem 53.98 ([813])**.**
Let G be a Lie group, H a closed subgroup of G and G/H with the natural topology. Then G/H has an unique analytic structure with the property that G is a Lie transformation group of G/H.

Proof. We denote by $\mathring{U}$ the interior of the U given by the lemma 53.97, and $B = \psi(\mathring{U}) \subset G$. Since $\phi \colon (A, B) \to \exp A \exp B$ is a diffeomorphism, $\psi(\mathring{U}) = \phi(U, 0)$ is a submanifold of G. We consider the following diagram:

$$\begin{tikzcd} G \times B \arrow[r, "\Phi"] \arrow[d, "I \times \pi"'] & G \arrow[d, "\pi"] \\ G \times N_0 & G/H \end{tikzcd}$$

with, for $g \in G$ and $x \in B$,

$$I \times \pi \colon (g, x) \mapsto (g, [x])$$

and

$$\Phi \colon (g, x) \mapsto gx.$$

The classes $[x]$ are taken with respect to H. The map $\mu \colon G \times N_0 \to G/H$, $\mu(g, [x]) = [gx]$ can be written under the form

$$\mu = \pi \circ \Phi \circ (I \times \pi)^{-1}$$

which is analytic [44]. So G is a Lie transformation group on G/H.

□

lem:categ

Lemma 53.99 (Category theorem)**.**
If a locally compact space M can be written as a countable union

$$M = \bigcup_{n=1}^{\infty} M_n \tag{53.267}$$

eq:M_union

where each M_i is closed in M, then at least one of them contains an open subset of M.

Proof. We suppose that none of the M_i contains an open subset of M. Let U_1 be an open whose closure is compact, $a_1 \in U_1 \backslash M_1$ and a neighbourhood U_2 of a_1 such that $\overline{U}_2 \subset U_1$ and $\overline{U_2} \cap M_1 = \varnothing$. Let $a_2 \in U_2 \backslash M_2$ and a neighbourhood U_3 of a_2 such that $\overline{U_3} \subset U_2$ and $\overline{U_3} \cap M_2 = \varnothing \dots$ and so on. The existence of the a_i comes from the fact that U_j is open, so that it is contained in no one of the M_k.

The decreasing sequence $\overline{U}_1, \overline{U}_2, \dots$ is made up from non empty compacts sets. Then $\bigcap_{i=1}^{\infty} U_i \neq \varnothing$ and the elements of this intersection are in none of the M_i; this contradict (53.267). □

44. Notice that the inverse of $I \times \pi$ exists because π is homeomorphic on the spaces considered here.

tho:homeo_action

Theorem 53.100.
Let G be a locally compact group with a countable basis. Suppose that it is a transitive, locally compact and Hausdorff topological group of transformation on M. Consider $p \in M$ and $H = \{g \in G \text{ tel que } g \cdot p = p\}$. Then

(1) H is closed,

(2) the map $[g] \to g \cdot p$ is homeomorphic between G/H and M.

Proof. By definition of a group action, the map $\varphi\colon G \to M$, $\varphi(g) = g \cdot p$ is continuous. Then $H = \varphi^{-1}(p)$ is closed in G.

As usual, the topology considered on G/H is a topology which makes the canonical projection $\pi\colon G \to G/H$ continuous and open. Now we study the map $\psi\colon G/H \to M$, $\psi([g]) = g \cdot p$ which is well defined because H fixes p by definition. It is clearly injective, and it is surjective because the action is transitive.

Now remark that $\psi = \varphi \circ \pi^{-1}$. Since π is continuous and open, and φ is continuous, it just remains to be proved that φ is open in order for ψ to be continuous and open. In order to do it, consider V, an open subset of G, $g \in V$ and a compact neighbourhood U of e in G such that $U = U^{-1}$ and $gU^2 \subset V$. If U is small and u, $v \in U$ close to e, then guv can keep in V, so that such a U exists.

We can find a sequence (g_n) in G such that $G = \bigcup_n g_n U$; the transitivity of G on M implies that

$$M = \bigcup_n g_n U \cdot p.$$

Each term in this union is compact, and therefore closed in M. By lemma 53.99, one of the $g_n U \cdot p$ contains an open subset of M. Since the action "$g \cdot$ " is continuous, $U \cdot p$ also contains an open subset in M. The conclusion is that one can find a $u \cdot p$ in the interior of M, and p is then an interior point of $u^{-1}U \cdot p \subset U^2 \cdot p$. Then $g \cdot p$ in in the interior of $V \cdot p$ and φ is therefore open. □

Proposition 53.101.
Let G be a transitive transformation Lie group on a C^∞ manifold M. Consider $p_0 \in M$ and H, the stabilizer of p_0:

$$H = \{g \in G \text{ tel que } g \cdot p_0 = p_0\}.$$

Let

$$\begin{aligned} \alpha\colon G/H &\to M \\ [g] &\mapsto g \cdot p_0. \end{aligned} \tag{53.268}$$

We have:

(1) The stabilizer H is closed in G.

(2) If α is homeomorphic, then it is diffeomorphic (if G/H has the analytic structure of theorem 53.98).

(3) If α is homeomorphic and if M is connected, then G_0, the identity component of G, is transitive on M.

propHelgason4.3

This comes from [813], chapter 2, proposition 4.3. The interest of this theorem is the fact that one only has to check the continuity of α and α^{-1} in order to have a diffeomorphism $M \simeq G/H$.

Proof. The group H is closed in G. We consider the map $\varphi\colon G \to M$, $\varphi(g) = g \cdot p_0$. This is continuous; therefore $\varphi^{-1}(p_0)$ is closed. Remark that we are in the situation of theorem 53.100
First item. We will use lemma 53.97. Se denotes by $\mathfrak{h}$, the Lie algebra of H and we consider a $\mathfrak{m}$ such that $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$; the lemma 53.97 assure us that we have a neighbourhood U of 0 in $\mathfrak{m}$ on which ψ is homeomorphic and such that π sends homeomorphically $\psi(U)$ to a neighbourhood of p_0 in G/H. We define $\mathring{U}$, the interior of U, $B = \psi(\mathring{U})$ and N_0, the interior of $\pi(\psi(U))$.

The set B is a submanifold of G, diffeomorphic to N_0 by π because everything is continuous and then everything respect the interiors.

¡¡ Avertissement/question au lecteur !! 53.102
C'est n'importe quoi comme justification. C'est lié au problème 53.104. prob:diffeo_1

Consider $\iota\colon B \to G$, the identity and $\beta\colon G \to M$, $\beta(g) = g \cdot p_0$. The restriction α_{N_0} of α to N_0 is an homeomorphism (this is a part of the assumptions) from N_0 to an open subset of M: N_0 is open (this is an interior), then its image by an homeomorphism is open.

Now we can see that α_{N_0} is differentiable. The reason is that it can be written as $\alpha_{N_0} = \beta \circ \iota \circ \pi^{-1}\colon N_0 \to M$. The construction makes π a diffeomorphism and β a diffeomorphism when G is a Lie group of transformations (as it is the case here); ι is clear. Now we have to see that the whole α is also differentiable, and then we will have to prove the same for α^{-1}.

By definition, $\alpha([g]) = g \cdot p_0$ (the classes $[g]$ is taken with respect to H). Consider $[n] \in H$; for any $g \in G$, one can write $[g] = [gn^{-1}n]$. Then

$$\alpha([g]) = \alpha([gn^{-1}n]) = gn^{-1}n \cdot p_0 = gn^{-1}\alpha([n]) = gn^{-1} \cdot \alpha_{N_0}([n]), \tag{53.269}$$

but the last dot denotes a differentiable action, and α_{N_0} is differentiable. Thus α is differentiable.

In order for α to be a diffeomorphism, we still have to prove that α^{-1} is differentiable., we begin to show that the Jacobian of β at $g = e$ has rank $r_\beta = \dim M$. We looks at $d\beta_e\colon \mathfrak{g} \to T_{p_0}M$, and consider $X \in \ker(d\beta_e)$. For $f \in C^\infty(M)$, we compute

$$0 = (d\beta_e X)f = X(f \circ \beta) = \frac{d}{dt}\Big[f(e^{tX} \cdot p_0)\Big]_{t=0}. \tag{53.270}$$

Let $s \in \mathbb{R}$, and we write this equation for f^* instead of f, which f^* defined by $f^*(q) = f(e^{sX} \cdot q)$ for each $q \in M$:

$$0 = \frac{d}{dt}\Big[f^*(e^{tX} \cdot p_0)\Big]_{t=0} = \frac{d}{dt}\Big[f(e^{(s+t)X} \cdot p_0)\Big]_{t=0} = \frac{d}{dt}\Big[f(e^{tX} \cdot p_0)\Big]_{t=s}. \tag{53.271}$$

Thus $f(e^{sX} \cdot p_0)$ is a constant with respect to s. Since f is arbitrary, $e^{sX} \cdot p_0 = p_0$ for any s. So $X \in \mathfrak{h}$ because $\exp sX \in H$ for any s. Then $\ker d\beta_e \subset \mathfrak{h}$.

On the other hand, $\mathfrak{h} \subset \ker d\beta_e$ is clear, then

$$\ker d\beta_e = \mathfrak{h}$$

and $r_\beta = \dim \mathfrak{g} - \dim \mathfrak{h}$.

Since α is an homeomorphism, the dimension of the origin and the target space are the same: $\dim G/H = \dim M$. On the other hand, lemma 53.96 gives $\dim G/H = \dim \mathfrak{g} - \dim \mathfrak{h}$, so that $r_\beta = \dim M$.

Now we prove that α^{-1} is differentiable. Remark that $\beta(g) = g \cdot p_0$ and $\alpha([g]) = g \cdot p_0 = \beta(g)$ is a good definition for α because the class are taken with respect to the stabilizer of p_0. Since $r_\beta = \dim M$, the map β is locally a diffeomorphism from a neighbourhood of e to a neighbourhood of p_0.

If $p = g \cdot p_0$, $\alpha^{-1}(p) = [g]$ because $[k] \in \alpha^{-1}(o)$ if $\alpha([k]) = p$, i.e. $k \cdot p_0 = p$. But $k = gr$ for a certain $r \in G$. It is clear that $p = k \cdot p_0 = gr \cdot p_0$. In particular, $g \cdot (r \cdot p_0)$. We know that in general $g \cdot p = g \cdot q$ implies $p = q$; here it gives us $r \in H$, so that $k \in [g]$.

We consider a $n \in G$ such that $n \cdot$ and $n^{-1} \cdot$ are diffeomorphic. We can make the following manipulation:

$$\alpha^{-1}(p) = [g] = [gnn^{-1}] = \pi(gn)\alpha^{-1}(n^{-1} \cdot p_0). \tag{53.272}$$

Under this form, α^{-1} is diffeomorphic.

Second item. If α is an homeomorphism, then β is open. Let us denote by G_0 the identity component of G. There exists a subset $\{x_\gamma \text{ tel que } \gamma \in I\}$ of G such that

$$G = \bigcup_{\gamma \in I} G_0 x_\gamma.$$

This comes from the fact that the components are all some left translations of the identity component (this is true for any Lie group). Each orbit $G_0 x_\gamma \cdot p_0$ is open in M and two orbits are either disjoint either equals. Since M is connected, all these orbits must coincide; thus each orbit contains the whole M. In particular, the orbit $G_0 \cdot p_0 = M$: G_0 is transitive on M.

□

53.12 Connected components

lem:vp_G_X

Lemma 53.103.
Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. If $\varphi\colon G \to X$ is an analytic homomorphism (X is a Lie group with Lie algebra $\mathfrak{x}$), then

(1) The kernel $\varphi^{-1}(e)$ is a topological Lie subgroup of G; his algebra is the kernel of $d\varphi_e$.

(2) The image $\varphi(G)$ is a Lie subgroup of X whose Lie algebra is $d\varphi(\mathfrak{g}) \subset \mathfrak{x}$.

(3) The quotient group $G/\varphi^{-1}(e)$ with his canonical analytic structure is a Lie group. The map $g\varphi^{-1}(e) \mapsto \varphi(g)$ is an analytic isomorphism $G/\varphi^{-1}(e) \to \varphi(G)$. In particular the map $\varphi\colon G \to \varphi(G)$ is analytic.

Proof. First item. We know that a subgroup H closed in G admits an unique analytic structure such that H becomes a topological Lie subgroup of G. This is the case of $\varphi^{-1}(e)$. We know that $Z \in \mathfrak{g}$ belongs to the Lie algebra of $\varphi^{-1}(e)$ if and only if $\varphi(\exp tZ) = e$ for any $t \in \mathbb{R}$. But $\varphi(\exp tZ) = \exp(td\varphi(Z)) = e$ if and only if $d\varphi(Z) = 0$.
Second item. Consider X_1, the analytic subgroup of X whose Lie algebra is $d\varphi(\mathfrak{g})$. The group $\varphi(G)$ is generated by the elements of the form $\varphi(\exp Z)$ for $Z \in \mathfrak{g}$. The group X_1 is generated by the $\exp(d\varphi Z)$. Because of lemma 53.38, these two are the same. Then $\varphi(G) = X_1$ and their Lie algebras are the same.
Third item. We consider H, a closed normal subgroup of G; this is a topological subgroup and the quotient G/H has an unique analytic structure such that the map $G \times G/H \to G/H$, $(g, [x]) \to [gx]$ is analytic. We consider a decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ and we looks at the restriction $\psi\colon \mathfrak{m} \to G$ of the exponential. Then there exists a neighbourhood U of 0 in $\mathfrak{m}$ which is homomorphically send by ψ into an open neighbourhood of e in G and such that $\pi\colon G \to G/H$ sends homomorphically $\psi(U)$ to a neighbourhood of $p_0 \in G/H$ (cf. lemma 53.97).

We consider $\mathring{U}$, the interior of U and $B = \psi(\mathring{U})$. The following diagram is commutative:

$$\begin{array}{ccc} G \times G/H & \xrightarrow{\Phi} & G/H \\ & {\scriptstyle \pi \times I} \searrow \quad \nearrow {\scriptstyle \alpha} & \\ & G/H \times G/H & \end{array} \tag{53.273}$$

with $\Phi(g, [x]) = [g^{-1}x]$, $(\pi \times I)(g, [x]) = ([g], [x])$ and $\alpha([g], [x]) = [g^{-1}x]$. Indeed,

$$\alpha \circ (\pi \times I)(g, [x]) = \alpha([g], [x]) = [g^{-1}x].$$

In order to see that α is well defined, remark that if $[h] = [g]$ and $[y] = [x]$ $[g^{-1}x] = [h^{-1}y]$ because H is a normal subgroup of G.

Now, we consider $g_0, x_0 \in G$ and the restriction of $(\pi \times I)$ to $(g_0 B) \times (G/H)$. Since π is homeomorphic on $\psi(U)$ and $B = \psi(\mathring{U})$, on $g_0 B$, π is a diffeomorphism (because the multiplication is diffeomorphic as well)

prob:diffeo_2

¡¡ Avertissement/question au lecteur !! 53.104
Why is the π a diffeomorphism? I understand why it is qn homeomorphism, but no more.

This diffeomorphism maps to a neighbourhood N of $([g_0], [x_0])$ in $G/H \times G/H$. From the commutativity, we know that $\alpha = \Phi \circ (\pi \times I)^{-1}$, so that α is analytic. Consequently, G/H is a Lie group. On N, α is analytic, then $\alpha(N)$ is analytic.

All this is for a closed normal subgroup H of G. Now we consider $H = \varphi^{-1}(e)$ and $\mathfrak{h}$, the Lie algebra of H. From the first item, we know that the Lie algebra of H is the kernel of $d\varphi$: $\mathfrak{h} = d\varphi^{-1}(0)$ which is an ideal in $\mathfrak{g}$.

From the second point, the Lie algebra of G/H is $d\pi(\mathfrak{g})$ which is isomorphic to $\mathfrak{g}/\mathfrak{h}$; the bijection is $\gamma(d\pi(X)) = [X] \in \mathfrak{g}/\mathfrak{h}$. In order to prove the injectivity, let us consider $\gamma(A) = \gamma(B)$; $A = d\pi(X)$, $B = d\pi(Y)$. The condition is $[X] = [Y]$; thus it is clear that $d\pi(X) = d\pi(Y)$

Let us consider on the other hand the map $Z+\mathfrak{h} \to d\varphi(Z)$ for $Z \in \mathfrak{g}$[45]. In other words, the map is $[Z] \to d\varphi(Z)$. This is an isomorphism $\mathfrak{g}/\mathfrak{h} \to d\varphi(\mathfrak{g})$, which gives a local isomorphism between G/H and $\varphi(G)$. This local isomorphism is $[g] \to \varphi(g)$ for g in a certain neighbourhood of e in G.

Since $[g] \to \varphi(g)$ has a differential which is an isomorphism, this is analytic at e. Then it is analytic everywhere.

□

Corollary 53.105.
If G is a connected Lie group and if Z is the center of G, then

(1) Ad_G is an analytic homomorphism from G to $\mathrm{Int}(G)$, with kernel Z,

(2) the map $[g] \to \mathrm{Ad}_G(g)$ is an analytic isomorphism from G/Z to $\mathrm{Int}(\mathfrak{g})$ (the class $[g]$ is taken with respect to Z).

cor:Ad_homom

Proof. First item. A connected Lie group is generated by a neighbourhood of identity, and any element of a suitable such neighbourhood can be written as the exponential of an element in the Lie algebra. So $\mathrm{Int}(\mathfrak{g})$ is generated by elements of the form $\exp(\mathrm{ad}\, X) = \mathrm{Ad}(\exp X)$; this shows that $\mathrm{Int}(\mathfrak{g}) \subset \mathrm{Ad}(G)$. In order to find the kernel, we have to see $\mathrm{Ad}_G^{-1}(e)$ by the formula

$$e^{\mathrm{Ad}(g)X} = ge^X g^{-1}.$$

We have to find the $g \in G$ such that $\forall X \in \mathfrak{g}$, $\mathrm{Ad}_G(g)X = X$. We taking the exponential of the two sides and using (51.8),

$$ge^X g^{-1} = e^X. \tag{53.274}$$

Then g must commute with any $e^X \in G$: in other words, g is in the kernel of G.
Second item. This is contained in lemma 53.103. Indeed G is connected and we had just proved that $\mathrm{Ad}_G : G \to \mathrm{Int}(\mathfrak{g})$ with kernel Z; the third item of lemma 53.103 makes G/Z a Lie group and the map $[g] \to \mathrm{Ad}_G(g)$ an analytic isomorphism from G/Z to $\mathrm{Ad}_G(G) = \mathrm{Int}(\mathfrak{g})$. □

Lemma 53.106.
Let G_1 and G_2 be two locally isomorphic connected Lie groups with trivial center (i.e. $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ and $Z(G_i) = \{e\}$). In this case, we have $G_1 = G_2 = \mathrm{Int}(\mathfrak{g})$ where $\mathrm{Int}\,\mathfrak{g}$ stands for the group of internal automorphism of $\mathfrak{g}$.

Proof. We denote by G_0 the group $\mathrm{Int}\,\mathfrak{g}$. The adjoint actions $\mathrm{Ad}_i : G_i \to G_0$ are both surjective because of corollary 53.105. Let us give an alternative proof for injectivity. Let $Z_i = \ker(\mathrm{Ad}_i) = \{g \in G_i \text{ tel que } \mathrm{Ad}(g)X = X, \forall X \in \mathfrak{g}\}$. Since G_i is connected, it is generated by any neighbourhood of the identity in the sense of proposition 52.5; let V_0 be such a neighbourhood. Taking eventually a subset we can suppose that V_0 is a normal coordinate system. So we have

$$g \exp_{G_i}(X) g^{-1} = \exp_{g_i}(X)$$

for every $X \in V_0$. Using proposition 52.5 we deduce that $gxg^{-1} = x$ for every $x \in G_i$, thus $g \in Z(G_i)$. That proves that $\ker(\mathrm{Ad}_i) \subset Z(G_i)$. The assumption of triviality of $Z(G_i)$ concludes injectivity of Ad_i. □

CORooDBIGooTUplRL

Corollary 53.107.
Let $\mathfrak{g}$ be a real Lie algebra with center $\{0\}$. Then the center of $\mathrm{Int}(\mathfrak{g})$ is only composed of the identity.

Proof. We note $G' = \mathrm{Int}(\mathfrak{g})$ and Z his center; ad is the adjoint representation of $\mathfrak{g}$ and Ad', ad', the ones of G' and $\mathrm{ad}(\mathfrak{g})$ respectively. We consider the map $\theta : G'/Z \to \mathrm{Int}(\mathrm{ad}(\mathfrak{g}))$, $\theta([g]) = \mathrm{Ad}'(g)$. By the second item of the corollary 53.105, $[g] \to \mathrm{Ad}_{G'}(g)$ is an analytic homomorphism from G' to $\mathrm{Int}(\mathfrak{g}')$ where $\mathfrak{g}'$ is the Lie algebra of G'; this is $\mathrm{ad}(\mathfrak{g})$. So $\theta : G'/Z \to \mathrm{Int}(\mathfrak{g}')$ is isomorphic.

45. Note that $\mathfrak{g}$ and $\mathfrak{h}$ are not groups; by $[X]$, we mean $[X] = \{X + h \text{ tel que } h \in \mathfrak{h}\}$.

Now we consider the map $s\colon \mathfrak{g} \to \mathrm{ad}(\mathfrak{g})$, $s(X) = \mathrm{ad}(X)$; this is an isomorphism. We also consider $S\colon G' \to \mathrm{GL}(\mathrm{ad}(\mathfrak{g}))$, $S(g) = s \circ g \circ s^{-1}$. The Lie algebra of $S(G')$ is $\mathrm{ad}(\mathfrak{g}') = \mathrm{ad}\big(\mathrm{ad}(\mathfrak{g})\big)$. Then $S(G')$ is the subset of $\mathrm{GL}(\mathrm{ad}\,\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}\big(\mathrm{ad}\,\mathfrak{g}\big)$, i.e. exactly $\mathrm{Int}(\mathrm{ad}\,\mathfrak{g})$. So S is an isomorphism $S\colon G' \to \mathrm{Int}(\mathrm{ad}\,\mathfrak{g})$. From all this,

$$S(e^{\mathrm{ad}\,X}) = s \circ e^{\mathrm{ad}\,X} \circ s^{-1} = e^{\mathrm{ad}'(\mathrm{ad}\,X)} = \mathrm{Ad}'(e^{\mathrm{ad}\,X}). \tag{53.275}$$

With this equality, $S^{-1} \circ \theta\colon G'/Z \to G'$ is an isomorphism which sends $[g]$ on g for any $g \in Z$. Then Z can't contains anything else than the identity. □

53.108.
The corollary 53.107 does not hold i we relax the assumptions of the trivial center. We have a counter-example with $\mathfrak{g} = \mathbb{R}^3$ and the commutations relation

$$[X_1, X_2] = X_3, \quad [X_1, X_3] = [X_2, X_3] = 0.$$

The group $\mathrm{Int}(\mathfrak{g})$ is abelian; then his center is the whole group, although $\mathfrak{g}$ is not abelian.

53.13 Adjoint group, inner automorphisms

sec:adj_gp

Let $\mathfrak{a}$ be a *real* Lie algebra. We denote by $GL(\mathfrak{a})$ the group of all the nonsingular endomorphisms of $\mathfrak{a}$: the linear and nondegenerate operators on $\mathfrak{a}$ as vector space. An element $\sigma \in \mathrm{GL}(\mathfrak{a})$ does not specially fulfils somethings like $\sigma[X, Y] = [\sigma X, \sigma Y]$. The Lie algebra $\mathfrak{gl}(\mathfrak{a})$ is the vector space of the endomorphisms (without non degeneracy condition) endowed with the usual bracket $(\mathrm{ad}\,A)B = [A, B] = A \circ B - B \circ A$. The map $X \to \mathrm{ad}\,X$ is a homomorphism from $\mathfrak{a}$ to the subalgebra $\mathrm{ad}(\mathfrak{a})$ of $\mathfrak{gl}(\mathfrak{a})$.

The group $\mathrm{Int}(\mathfrak{a})$ is the analytic Lie subgroup of $\mathrm{GL}(\mathfrak{a})$ whose Lie algebra is $\mathrm{ad}(\mathfrak{a})$ by theorem 53.7. This is the **adjoint group** of $\mathfrak{a}$.

Proposition 53.109.
The group $\mathrm{Aut}(\mathfrak{a})$ *of all the automorphisms of* $\mathfrak{a}$ *is a closed subgroup of* $\mathrm{GL}(\mathfrak{a})$.

Proof. The property which distinguish the elements in $\mathrm{Aut}(\mathfrak{a})$ from the "commons" elements of $\mathrm{GL}(\mathfrak{a})$ is the preserving of structure: $\varphi[A, B] = [\varphi A, \varphi B]$. These are equalities, and we know that a subset of a manifold which is given by some equalities is closed. □

Now, theorem 53.70 provides us an unique analytic structure on $\mathrm{Aut}(\mathfrak{a})$ in which it is a topological Lie subgroup of $\mathrm{GL}(\mathfrak{a})$. From now we only consider this structure. We denote by $\partial(\mathfrak{a})$ the Lie algebra of $\mathrm{Aut}(\mathfrak{a})$: this is the set of the endomorphisms D of $\mathfrak{a}$ such that $\forall t \in \mathbb{R}$, $e^{tD} \in \mathrm{Aut}(\mathfrak{a})$. By differencing the equality

$$e^{tD}[X, Y] = [e^{tD}X, e^{tD}Y] \tag{53.276}$$

eq:exp_der

with respect to t, we see [46] that D is a derivation [47] of $\mathfrak{a}$

Conversely, consider D, any derivation of $\mathfrak{a}$; by induction,

$$D^k[X, Y] = \sum_{i+j=k} \frac{k!}{i!j!}[D^i X, D^j Y] \tag{53.277}$$

where by convention, D^0 is the identity in $\mathfrak{a}$. This relation shows that D fulfils condition (53.276), so that any derivation of $\mathfrak{a}$ lies in $\partial(\mathfrak{a})$. Then

$$\partial(\mathfrak{a}) = \{\text{derivations of } \mathfrak{a}\}.$$

46. As usual, if we consider a basis of $\mathfrak{a}$ as vector space, the expression in the right hand side of

$$[e^{tD}X, e^{tD}Y] = \mathrm{ad}(e^{tD}X)e^{tD}X$$

can be seen as a product matrix times vector, so that Leibniz works.

47. Definition 51.3.

The Jacobi identities show that

$$\mathrm{ad}(\mathfrak{a}) \subset \partial(\mathfrak{a}).$$

From this, we deduce [48]:

eq:int_sub_aut

$$\mathrm{Int}(\mathfrak{a}) \subset \mathrm{Aut}(\mathfrak{a}). \tag{53.278}$$

Indeed the group $\mathrm{Int}(\mathfrak{a})$ being connected, it is generated [49] by any neighbourhood of e; note that $\mathrm{Aut}(\mathfrak{a})$ has not specially this property. We take a neighbourhood of e in $\mathrm{Int}(\mathfrak{a})$ under the form $\exp V$ where V is a sufficiently small neighbourhood of 0 in $\mathrm{ad}(\mathfrak{a})$ to be a neighbourhood of 0 in $\partial(\mathfrak{a})$ on which exp is a diffeomorphism. In this case, $\exp V \subset \mathrm{Aut}(\mathfrak{a})$ and then $\mathrm{Int}(\mathfrak{a}) \subset \mathrm{Aut}(\mathfrak{a})$.

Elements of $\mathrm{ad}(\mathfrak{a})$ are the **inner derivations** while the ones of $\mathrm{Int}(\mathfrak{a})$ are the **inner automorphisms.**

Let $\mathcal{O}$ be an open subset of $\mathrm{Aut}(\mathfrak{a})$; for a certain open subset U of $\mathrm{GL}(\mathfrak{a})$, $\mathcal{O} = U \cap \mathrm{Aut}(\mathfrak{a})$. Then

$$\iota^{-1}(\mathcal{O}) = \mathcal{O} \cap \mathrm{Int}(\mathfrak{a}) = U \cap \mathrm{Aut}(\mathfrak{a}) \cap \mathrm{Int}(\mathfrak{a}) = U \cap \mathrm{Int}(\mathfrak{a}). \tag{53.279}$$

The subset $U \cap \mathrm{Int}(\mathfrak{a})$ is open in $\mathrm{Int}(\mathfrak{a})$ for the topology because $\mathrm{Int}(\mathfrak{a})$ is a Lie [50] subgroup of $\mathrm{GL}(\mathfrak{a})$ and thus has at least the induced topology. This proves that the inclusion map $\iota\colon \mathrm{Int}(\mathfrak{a}) \to \mathrm{Aut}(\mathfrak{a})$ is continuous.

The lemma 49.71 and the consequence below makes $\mathrm{Int}(\mathfrak{a})$ a Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$. Indeed $\mathrm{Int}(\mathfrak{a})$ and $\mathrm{Aut}(\mathfrak{a})$ are both submanifolds of $\mathrm{GL}(\mathfrak{a})$ which satisfy (53.278).

By definition, $\mathrm{Aut}(\mathfrak{a})$ has the induced topology from $\mathrm{GL}(\mathfrak{a})$. Then $\mathrm{Int}(\mathfrak{a})$ is a submanifold of $\mathrm{Aut}(\mathfrak{a})$. This is also a subgroup and a topological group : $\mathrm{Int}(\mathfrak{a})$ is not a topological subgroup of $\mathrm{Aut}(\mathfrak{a})$. Then $\mathrm{Int}(\mathfrak{a})$ is a Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$. Schematically, links between $\mathrm{Int}\,\mathfrak{g}$, $\mathrm{ad}\,\mathfrak{g}$, $\mathrm{Aut}\,\mathfrak{g}$ and $\partial\mathfrak{g}$ are

eq:schem_ad_int

$$\mathrm{Int}\,\mathfrak{g} \longleftarrow \mathrm{ad}\,\mathfrak{g} \tag{53.280a}$$

$$\mathrm{Aut}\,\mathfrak{g} \longrightarrow \partial\mathfrak{g}. \tag{53.280b}$$

Remark that the sense of the arrows is important. By definition $\partial\mathfrak{g}$ is the Lie algebra of $\mathrm{Aut}\,\mathfrak{g}$, then there exist some algebras $\mathfrak{g}$ and $\mathfrak{g}'$ with $\mathrm{Aut}\,\mathfrak{g} \neq \mathrm{Aut}\,\mathfrak{g}'$ but with $\partial\mathfrak{g} = \partial\mathfrak{g}'$, because the equality of two Lie algebras doesn't implies the equality of the groups. The case of $\mathrm{Int}\,\mathfrak{g}$ and $\mathrm{ad}\,\mathfrak{g}$ is very different: the group is defined from the algebra, so that $\mathrm{ad}\,\mathfrak{g} = \mathrm{ad}\,\mathfrak{g}'$ implies $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}'$ and $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}'$ if and only if $\mathrm{ad}\,\mathfrak{g} = \mathrm{ad}\,\mathfrak{g}'$.

A result about the group of inner automorphism which will be useful later:

lem:Int_g_gR

Lemma 53.110.
If $\mathfrak{g}$ is a complex semisimple Lie algebra, then $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}^{\mathbb{R}}$.

Proof. If $\{X_i\}$ is a basis of $\mathfrak{g}$, then $\{X_j, iX_j\}$ is a basis of $\mathfrak{g}^{\mathbb{R}}$. We define $\psi\colon \mathrm{ad}\,\mathfrak{g} \to \mathrm{ad}\,\mathfrak{g}^{\mathbb{R}}$ by

$$\psi(\mathrm{ad}(a^j X_j)) = \mathrm{ad}(a^j X_j).$$

It is clearly surjective. On the other hand, if $\mathrm{ad}(a^j X_j)\,\mathrm{ad}(b^k X_k)$ as elements of $\mathrm{ad}\,\mathfrak{g}^{\mathbb{R}}$, then they are equals as elements of $\mathrm{ad}\,\mathfrak{g}$. The discussion following equations (53.280) finishes the proof. □

Corollary 53.111.
Any two real compact form of a complex semisimple Lie algebra are conjugate by an inner automorphism.

Proof. We know that any real form of $\mathfrak{g}$ induces an involution (the conjugation) and that if the real form is compact, the involution is Cartan on $\mathfrak{g}^{\mathbb{R}}$. Let $\mathfrak{u}_0$ and $\mathfrak{u}_1$ be two compact real forms of $\mathfrak{g}$ and τ_0, τ_1 the associated involutions of $\mathfrak{g}$ (which are Cartan involutions of $\mathfrak{g}^{\mathbb{R}}$). For a suitable $\varphi \in \mathrm{Int}\,\mathfrak{g}^{\mathbb{R}}$,

$$\tau_0 = \varphi\tau_1\varphi^{-1}.$$

48. See error 88.3
49. See proposition 52.5
50. Is it true??

The fact that $\operatorname{Int}\mathfrak{g} = \operatorname{Int}\mathfrak{g}^{\mathbb{R}}$ (lemma 53.110) finishes the proof. □

Proposition 53.112.
The group $\operatorname{Int}(\mathfrak{a})$ *is a normal subgroup of* $\operatorname{Aut}(\mathfrak{a})$.

Proof. Let us consider a $s \in \operatorname{Aut}(\mathfrak{a})$. The map $\sigma_s\colon \operatorname{Aut}(\mathfrak{a}) \to \operatorname{Aut}(\mathfrak{a})$, $\sigma_s(g) = sgs^{-1}$ is an automorphism of $\operatorname{Aut}(\mathfrak{a})$. Indeed, consider g, $h \in \operatorname{Aut}(\mathfrak{a})$; direct computations show that $\sigma_s(gh) = \sigma_s(g)\sigma_s(h)$ and $[\sigma_s(g), \sigma_s(h)] = \sigma_s([g,h])$. From this, $(d\sigma_s)_e$ is an automorphism of $\partial(\mathfrak{a})$, the Lie algebra of $\operatorname{Aut}(\mathfrak{a})$. For any $D \in \partial(\mathfrak{a})$ we have

$$(d\sigma_s)_e D = \frac{d}{dt}\Big[sD(t)s^{-1}\Big]_{t=0} = sDs^{-1}. \tag{53.281}$$ eq:ad_s_2

Since s is an automorphism of $\mathfrak{a}$ and $\operatorname{ad}(\mathfrak{a})$, a subalgebra of $\mathfrak{gl}(\mathfrak{a})$,

$$s \operatorname{ad} X s^{-1} = \operatorname{ad}(sX) \tag{53.282}$$ eq:ad_s_1

for any $X \in \mathfrak{a}$, $s \in \operatorname{Aut}(\mathfrak{a})$. Since $\operatorname{ad}(\mathfrak{a}) \subset \partial(\mathfrak{a})$, we can write (53.281) with $D = \operatorname{ad} X$ and put it in (53.282):

$$(d\sigma)_e \operatorname{ad} X = s \operatorname{ad} X s^{-1} = \operatorname{ad}(s \cdot X).$$

We know from general theory of linear operators on vector spaces that if A, B are endomorphism of a vector space and if A^{-1} exists, then $Ae^B A^{-1} = e^{ABA^{-1}}$. We write it with $A = s$ and $B = \operatorname{ad} X$:

$$\sigma_s \cdot e^{\operatorname{ad} X} = se^{\operatorname{ad} X}s^{-1} = e^{s \operatorname{ad} X s^{-1}} = e^{\operatorname{ad}(s \cdot X)},$$

sot that

$$\sigma_s \cdot e^{\operatorname{ad} X} = e^{\operatorname{ad}(sX)}. \tag{53.283}$$ eq:sigma_aut_s

Ont the other hand, we know that $\operatorname{Int}(\mathfrak{a})$ is connected, so it is generated by elements of the form $e^{\operatorname{ad} X}$ for $X \in \mathfrak{a}$. Then $\operatorname{Int}(\mathfrak{a})$ is a normal subgroup of $\operatorname{Aut}(\mathfrak{a})$; the automorphism s of $\mathfrak{a}$ induces the isomorphism $g \to sgs^{-1}$ in $\operatorname{Int}(\mathfrak{a})$ because of equation (53.283). □

More generally, if s is an isomorphism from a Lie algebra $\mathfrak{a}$ to a Lie algebra $\mathfrak{b}$, then the map $g \to sgs^{-1}$ is an isomorphism between $\operatorname{Aut}(\mathfrak{a})$ and $\operatorname{Aut}(\mathfrak{b})$ which sends $\operatorname{Int}(\mathfrak{a})$ to $\operatorname{Int}(\mathfrak{b})$. Indeed, consider an isomorphism $s\colon \mathfrak{a} \to \mathfrak{b}$ and $g \in \operatorname{Aut}(\mathfrak{a})$. If $g \in \operatorname{Int}(\mathfrak{a})$, we have to see that $sgs^{-1} \in \operatorname{Int}(\mathfrak{b})$. By definition, $\operatorname{Int}(\mathfrak{a})$ is the analytic subgroup of $\operatorname{GL}(\mathfrak{a})$ which has $\operatorname{ad}(\mathfrak{a})$ as Lie algebra. We have $g = e^{\operatorname{ad} A}$, then $sgs^{-1} = e^{\operatorname{ad}(sA)}$ which lies well in $\operatorname{Int}(\mathfrak{b})$.

Theorem 53.113.
Let G be a Lie group and H, a Lie subgroup of G.

(1) If H is a topological Lie subgroup of G, then it is closed in G,

(2) If H has at most a countable number of connected components and is closed in G, then H is a topological subgroup of G.

tho:H_ferme

Proof. First point. It is sufficient to prove that if a sequence $h_n \in H$ converges (in G) to $g \in G$, then $g \in H$ (this is almost the definition of a closed subset). We consider V, a neighbourhood of 0 in $\mathfrak{g}$ such that

— exp is diffeomorphic between V and an open neighbourhood of e in G,

— $\exp(V \cap \mathfrak{h}) = (\exp V) \cap H$.

This exists by the lemma 53.59; we can suppose that V is bounded. Consider $\mathcal{U}$, an open neighbourhood of 0 in $\mathfrak{g}$ contained in V such that $\exp -\mathcal{U} \exp \mathcal{U} \subset \exp V$.

Since $h_n \to g$, there exists a $N \in \mathbb{N}$ such that $n \geqslant N$ implies $h_n \in g \exp \mathcal{U}$ (i.e h_n is the product of g by an element rather close to e; since the multiplication is differentiable, the notion of "not so far" is good to express the convergence notion). From now we only consider such elements in the sequence. So, $h_N^{-1}h_n \in (\exp V) \cap H$ $(n \geqslant N)$ because

$$h_N^{-1}h_n \in \exp -\mathcal{U} g^{-1} g \exp \mathcal{U} \subset \exp V.$$

(note that H is a group, then $h_i^{-1} \in H$) From the second point of the definition of V, there exists a $X_n \in V \cap \mathfrak{h}$ such that $h_N^{-1} h_n = \exp X_n$ for any $n \geqslant N$.

Since V in bounded, there exists a subsequence out of (X_i) (which is also called X_i) converging to a certain $Z \in \mathfrak{g}$. But $\mathfrak{h}$ is closed in $\mathfrak{g}$ because it is a vector subspace (we are in a finite dimensional case), then $Z \in \mathfrak{h}$ and thus the sequence (h_i) converges to $h_N \exp Z$; therefore $g \in H$.

Second point. The subgroup H is closed in G and has a countable number of connected component. Since H is closed, theorem 53.70 it has an analytics structure for which it is a topological Lie subgroup of G. We denotes by H' this Lie group.

The identity map $I\colon H \to H'$ is continuous [51] (see error 88.4). Thus any connected component of H is contained in a connected component of H', the it has only a countable number of connected components. By corollary 52.28, $H = H'$ as Lie group. □

With the notations and the structure of theorem 53.70, the subgroup H is discrete if and only if $\mathfrak{h} = \{0\}$. Indeed, recall the definition (53.182):

$$\mathfrak{h} = \{X \in \mathfrak{g} : \forall t \in \mathbb{R}, e^{tX} \in H\},$$

and the fact that there exists a neighbourhood of e in H on which the exponential map is a diffeomorphism.

Theorem 53.114.
Let G and H be two Lie groups and $\varphi\colon G \to H$ a continuous homomorphism. Then φ is analytic.

Proof. The Lie algebra of the product manifold $G \times H$ as $\mathfrak{g} \times \mathfrak{h}$ is given in 49.122. We define

$$K = \{(g, \varphi(g)) : g \in G\} \subset G \times H. \tag{53.284}$$

It is clear that K is closed in $G \times H$ because G is closed and φ is continuous. By theorem 53.70, there exists an unique differentiable structure on $G \times H$ such that K is a topological Lie subgroup of $G \times H$ (i.e.: Lie subgroup [52] + induced topology). The Lie algebra of K is

$$\mathfrak{k} = \{(X, Y) \in \mathfrak{g} \times \mathfrak{h} : \forall t \in \mathbb{R}, (e^{tX}, e^{tY}) \in K\}. \tag{53.285}$$

Let N_0 be an open neighbourhood of 0 in $\mathfrak{h}$ such that exp is diffeomorphic between N_0 and an open neighbourhood N_e of e in H. We define M_0 and M_e in the same way, for G instead of H. We can suppose $\varphi(M_e) \subset N_e$: if it is not, we consider a smaller M_e: the openness of N_e and the continuity of φ make it coherent.

The lemma 53.59 allow us to consider M_0 and N_0 small enough to say that

$$\exp\colon (M_0 \times N_0) \cap \mathfrak{k} \to (M_e \times N_e) \cap K$$

is diffeomorphic. Now, we are going to show that for any $X \in \mathfrak{g}$, there exists an unique $Y \in \mathfrak{h}$ such that $(X, Y) \in \mathfrak{k}$. The unicity is easy: consider $(X, Y_1), (X, Y_2) \in \mathfrak{k}$; then $(0, Y_1 - Y_2) \in \mathfrak{k}$ (because a Lie algebra is a vector space). Then the definition of $\mathfrak{k}$ makes for any $t \in \mathbb{R}$, $(e, \exp t(Y_1 - Y_2)) \in K$. Consequently, $\exp t(Y_1 - Y_2) = \varphi(e) = e$ and then $Y_1 - Y_2 = 0$.

In order to show the existence, let us consider a $r > 0$ such that $X_r = (1/r)X$ keeps in M_0. This exists because the sequence $X_r \to 0$ (then it comes M_0 from a certain r). From the definitions, exp is diffeomorphic between M_0 and M_e, then $\exp X_r \in M_e$ and $\varphi(\exp X_r) \in N_e$ because $\varphi(M_e) \subset N_e$.

From this, there exists an unique $Y_r \in N_0$ such that $\exp Y_r = \varphi(\exp X_r)$ and an unique $Z_r \in (M_0 \times N_0) \cap \mathfrak{k}$ satisfying $\exp Z_r = (\exp X_r, \exp Y_r)$. But exp is bijective from $M_0 \times N_0$, so that $Z_r = (X_r, Y_r)$ and we can choose $Y = rY_r$ as a $Y \in \mathfrak{h}$ such that $(X, Y) \in \mathfrak{k}$ (it is not really a choice: the unicity was previously shown). We denotes by $\psi\colon \mathfrak{g} \to \mathfrak{h}$ the map which gives the unique $Y \in \mathfrak{h}$ associated with $X \in \mathfrak{g}$ such that $(X, Y) \in \mathfrak{k}$. This is a homomorphism between $\mathfrak{g}$ and $\mathfrak{h}$.

51. pourquoi ?
52. Definition 52.24.

By definition, $(X, \psi(X)) \in \mathfrak{k}$, i.e. $(\exp tX, \exp t\psi(X)) \in K$ or

$$\varphi(\exp tX) = \exp t\psi(X). \tag{53.286}$$

Let us now consider a basis $\{X_1, \dots, X_n\}$ of $\mathfrak{g}$. Since φ is a homomorphism,

$$\varphi\big((\exp t_1X_1)(\exp t_2X_2)\dots(\exp t_nX_n)\big) = \big(\exp t_1\psi(X_1)\big)\dots\big(\exp t_n\psi(X_n)\big) \tag{53.287}$$ eq:coord_vp_exp

Now, we apply lemma 53.61 on the decomposition of $\mathfrak{g}$ into the n subspace spanned by the n vector basis (this is n applications of the lemma), the map

$$(\exp t_1X_1)\dots(\exp t_nX_n) \to (t_1, \dots, t_n)$$

is a coordinate system around e in G. In this case, the relation (53.287) shows that φ is differentiable at e. Then it is differentiable anywhere in G. □

53.14 Compact Lie algebra

pg:compact_Lie

We consider $\mathfrak{g}$, a real Lie algebra and $\mathfrak{h}$, a subalgebra of $\mathfrak{g}$. Let K^* be the analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ which corresponds to the subalgebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{h})$ of $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$.

DEFooROMGooTLicyL

Definition 53.115.
*We say that $\mathfrak{h}$ is **compactly embedded** in $\mathfrak{g}$ if K^* is compact. A Lie algebra is **compact** when it is compactly embedded in itself.*

The analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ which corresponds to $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$, by definition, is $\mathrm{Int}(\mathfrak{g})$. Then the compactness of $\mathfrak{g}$ is the one of $\mathrm{Int}(\mathfrak{g})$.

Remarque 53.116.
The compactness notion on a Lie group is defined from the topological structure of the Lie group seen as a manifold. It is all but trivial that the compactness on a Lie group is related to the compactness on its Lie algebra; the proposition 53.92 will however make the two notions related in the natural way.

Remarque 53.117.
The topology on K^* is not necessary the same as the induced one from $\mathrm{Int}(\mathfrak{g})$ and $\mathrm{Int}(\mathfrak{g})$ has also not necessary the induced topology from $\mathrm{GL}(\mathfrak{g})$. However the next proposition will show that the compactness notion is well the one induced from $\mathrm{GL}(\mathfrak{g})$.

Proposition 53.118.
We consider $\widetilde{K}$, the same set and group as K^*, but with the induced topology from $\mathrm{GL}(\mathfrak{g})$. Then $\widetilde{K}$ is compact if and only if K^* is compact.

Note however that K^* and $\widetilde{K}$ are not automatically the same as manifold.

Proof. K^* compact implies $\widetilde{K}$ compact. The identity map $\iota\colon K^* \to \mathrm{GL}(\mathfrak{g})$ is analytic, and then is continuous because $\mathrm{Int}(\mathfrak{g})$ is by definition an analytic subgroup of $\mathrm{GL}(\mathfrak{g})$ and K^* an analytic subgroup of $\mathrm{Int}(\mathfrak{g})$. If we have a covering of $\widetilde{K}$ with open set $\mathcal{O}_i \cap \widetilde{K}$ of $\widetilde{K}$ ($\mathcal{O}_i$ is open in $\mathrm{GL}(\mathfrak{g})$), the continuity of ι make the finite subcovering of K^* good for $\widetilde{K}$.
$\widetilde{K}$ compact implies K^* compact. If $\widetilde{K}$ is compact, then it is closed in $\mathrm{GL}(\mathfrak{g})$. As set, K^* is closed in $\mathrm{GL}(\mathfrak{g})$ and by definition it is connected. Then by the theorem 53.113, K^* is a topological subgroup of $\mathrm{GL}(\mathfrak{g})$. Consequently, K^* and $\widetilde{K}$ are homeomorphic and they have same topology. □

Lemma 53.119 ([1])**.**
If G is a compact group in $\mathrm{GL}(n, \mathbb{R})$, then there exists a G-invariant quadratic form on $\mathbb{R}^n$ [53].

53. Is it true ? I've not even found a precise statement of this claim.

53.15 Real Lie algebras

53.15.1 Real and complex vector spaces

If V is a real vector space, the **complexification** of V is the vector space

$$V^{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}.$$

If $\{v_i\}$ is a basis of V on $\mathbb{R}$, then $\{v_i \otimes 1\}$ is a basis of $V^{\mathbb{C}}$ on C. Then

$$\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}.$$

Let W be a complex vector space. If one restrains the scalars to $\mathbb{R}$, we find a real vector space denoted by $W^{\mathbb{R}}$. If $\{w_j\}$ is a basis of W, then $\{w_j, iw_j\}$ is a basis of $W^{\mathbb{R}}$ and

$$\dim_{\mathbb{C}} W = \frac{1}{2} \dim_{\mathbb{R}} W^{\mathbb{R}}.$$

Note that $(V^{\mathbb{C}})^{\mathbb{R}} = V \oplus iV$.

A real vector space V is a **real form** of a complex vector space W if $W^{\mathbb{R}} = V \oplus iV$. If V is a real form of W, the map $\varphi\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ given by the identity on V and the multiplication by -1 on iV is the **conjugation** of $V^{\mathbb{C}}$ with respect of the real form V.

53.15.2 Real and complex Lie algebras

For notational convenience, if not otherwise mentioned, $\mathfrak{g}$ will denote a complex Lie algebra and $\mathfrak{f}$ a real one. If $\mathfrak{f}$ is a real Lie algebra and $\mathfrak{f}^{\mathbb{C}} = \mathfrak{f} \otimes \mathbb{C}$, its complexification (as vector space), we endow $\mathfrak{f}^{\mathbb{C}}$ with a Lie algebra structure by defining

$$[(X \otimes a), (Y \otimes b)] = [X, Y] \otimes ab.$$

This is a bilinear extension of the Lie algebra bracket of $\mathfrak{f}$. It is rather easy to see that $[\mathfrak{f}, \mathfrak{f}]^{\mathbb{C}} = [\mathfrak{f}^{\mathbb{C}}, \mathfrak{f}^{\mathbb{C}}]$.

Now we turn our attention to the Killing form. Let $\mathfrak{f}$ be a real Lie algebra with a Killing form $B_{\mathfrak{f}}$. A basis of $\mathfrak{f}$ is also a basis of $\mathfrak{f}^{\mathbb{C}}$. Then the matrix $B_{ij} = \operatorname{Tr}(\operatorname{ad} X_i \circ \operatorname{ad} X_j)$ of the Killing form is the same for $\mathfrak{f}^{\mathbb{C}}$ than for $\mathfrak{f}$. In conclusion:

$$B_{\mathfrak{f}^{\mathbb{C}}}|_{\mathfrak{f} \times \mathfrak{f}} = B_{\mathfrak{f}}.$$

Let us study the inverse process: $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{g}^{\mathbb{R}}$ is the real Lie algebra obtained from $\mathfrak{g}$ by restriction of the scalars. If $\mathcal{B} = \{v_j\}$ is a basis of $\mathfrak{g}$, $\mathcal{B}' = \{v_j, iv_j\}$ is a one of $\mathfrak{g}^{\mathbb{R}}$. For a certain $X \in \mathfrak{g}$ we denote by (c_{kl}) the matrix of $\operatorname{ad}_{\mathfrak{g}} X$. Now we study the matrix of $\operatorname{ad}_{\mathfrak{g}^{\mathbb{R}}} X$ in the basis $\mathcal{B}'$ by computing

$$(\operatorname{ad}_{\mathfrak{g}} X)v_i = c_{ik}v_k = \big[\operatorname{Re}(c_{ik}) + i \operatorname{Im}(c_{ik})\big]v_k = a_{ik}v_k + b_{ik}(iv_k) \tag{53.288}$$

if $a = \operatorname{Re} c$ and $b = \operatorname{Im} c$. Then the columns of $\operatorname{ad}_{\mathfrak{g}^{\mathbb{R}}}$ which correspond to the $v_i \in \mathcal{B}'$'s are given by

$$\operatorname{ad}_{\mathfrak{g}^{\mathbb{R}}} X = \begin{pmatrix} a & \cdot \\ b & \cdot \end{pmatrix}$$

where the dots denote some entries to be find now:

$$(\operatorname{ad}_{\mathfrak{g}} X)(iv_i) = i\big(a_{ik}v_k + b_{ik}(iv_k)\big) = a_{ik}(iv_k) - b_{ik}v_k, \tag{53.289}$$

so that the complete matrix of $\operatorname{ad}_{\mathfrak{g}^{\mathbb{R}}} X$ in the basis $\mathcal{B}'$ is given by

$$\operatorname{ad}_{\mathfrak{g}^{\mathbb{R}}} X = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

So,

$$\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X \circ \mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X' = \begin{pmatrix} aa' - bb' & \cdot \\ \cdot & aa' - bb' \end{pmatrix}.$$

Then $B(X, X') = 2\,\mathrm{Tr}(aa' - bb')$ while

$$B(X, Y) = \mathrm{Tr}\left((a + ib)(a' + ib')\right) = \mathrm{Tr}(aa' - bb') + i\,\mathrm{Tr}(ab' + ba'). \tag{53.290}$$

Thus we have

$$B_{\mathfrak{g}^{\mathbb{R}}} = 2\,\mathrm{Re}\,B_{\mathfrak{g}}, \tag{53.291}$$

so that $\mathfrak{g}^{\mathbb{R}}$ is semisimple if and only if $\mathfrak{g}$ is semisimple.

53.15.3 Split real form

Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{h}$ a Cartan subalgebra, Φ the set roots, Δ the set of non zero roots and B, the Killing form. From property (51.345) and the fact that $c(-\alpha, -\beta) = c(\alpha, \beta)$, we find $c(\alpha, \beta)^2 = \frac{1}{2}\beta^\alpha(1 + \beta_\alpha)|\alpha|^2$, so that $c(\alpha, \beta)^2 \geqslant 0$ which gives $c(\alpha, \beta) \in \mathbb{R}$. We can define

$$\mathfrak{g}_{\mathbb{R}} = \mathfrak{h}_0 \bigoplus_{\alpha \in \Phi} \mathbb{R} x_\alpha.$$

Remark that $\mathfrak{g}_\alpha$ has dimension one with respect to $\mathbb{C}$, not $\mathbb{R}$; then $\mathbb{R}x_\alpha \neq \mathfrak{g}_\alpha$, but $\mathbb{C}x_\alpha = \mathfrak{g}_\alpha$ and $\mathfrak{g}_\alpha = \mathbb{R}x_\alpha \oplus i\mathbb{R}x_\alpha$. Since it is clear that $\bigoplus_{\alpha\in\Delta}(\mathbb{R}x_\alpha \oplus i\mathbb{R}x_\alpha) = \bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha$, the proposition 51.158 gives

$$\mathfrak{g} = \mathfrak{g}_{\mathbb{R}} \oplus i\mathfrak{g}_{\mathbb{R}}. \tag{53.292}$$

Any real form of $\mathfrak{g}$ which contains the $\mathfrak{h}_{\mathbb{R}}$ of a certain Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is said a **split real form**. The construction shows that any complex semisimple Lie algebra admits a split real form.

53.15.4 Compact real form

Definition 53.120.
*A **compact real form** of a complex Lie algebra is a real form which is compact as Lie algebra*[54].

Theorem 53.121.
Any complex semisimple Lie algebra contains a compact real form.

Proof. Let $\mathfrak{h}$ be a Cartan algebra of the complex semisimple Lie algebra $\mathfrak{g}$ and x_α, some root vectors. We consider the space

$$\mathfrak{u}_0 = \underbrace{\sum_{\alpha\in\Phi} \mathbb{R}ih_\alpha}_{A} + \underbrace{\sum_{\alpha\in\Phi} \mathbb{R}(x_\alpha - x_{-\alpha})}_{B} + \underbrace{\sum_{\alpha\in\Phi} \mathbb{R}i(x_\alpha + x_{-\alpha})}_{C}. \tag{53.293}$$

Since $\mathfrak{u}_0 \oplus i\mathfrak{u}_0$ contains all the $\mathbb{C}h_\alpha$, $\mathfrak{h} \subset \mathfrak{u}_0 \oplus i\mathfrak{u}_0$; it is also rather clear that $\mathfrak{u}_0$ is a real form of $\mathfrak{g}$ (as vector space), for example, $i\mathbb{R}(x_\alpha - x_{-\alpha}) + \mathbb{R}i(x_\alpha + x_{-\alpha}) = \mathbb{R}ix_\alpha$. Now we have to check that $\mathfrak{u}_0$ is a real form of $\mathfrak{g}$ as Lie algebra, i.e. that $\mathfrak{u}_0$ is closed for the Lie bracket. This is a lot of

54. Definition 53.115.

computations:

$$\begin{aligned}
[ih_\alpha, ih_\beta] &= 0,\\
[ih_\alpha, (x_\alpha - x_{-\alpha})] &= i(\alpha(h_\alpha)x_\alpha - (-\alpha)(h_\alpha)x_{-\alpha})\\
&= i\alpha(h_\alpha)(x_\alpha + x_{-\alpha}) \in C,\\
[ih_\alpha, i(x_\alpha + x_{-\alpha})] &= -\alpha(h_\alpha)(x_\alpha - x_{-\alpha}) \in B,\\
[(x_\alpha - x_{-\alpha}), (x_\beta - x_{-\beta})] &= c(\alpha,\beta)(x_{\alpha+\beta} - x_{-(\alpha+\beta)}) \in B\\
&\quad - c(\alpha,\beta)(x_{\alpha-\beta} - x_{\beta-\alpha}) \in B,\\
[(x_\alpha - x_{-\alpha}), i(x_\beta + x_{-\beta})] &= ic(\alpha,\beta)(x_{\alpha+\beta} + x_{-(\alpha+\beta)}) \in C\\
&\quad + ic(\alpha,-\beta)(x_{\alpha-\beta)} + x_{-\alpha+\beta)}) \in C\\
[ih_\alpha, (x_\beta - x_{-\beta})] &= i\beta(h_\alpha)(x_\beta - x_{-\beta}) \in C\\
[ih_\alpha, i(x_\beta + x_{-\beta})] &= -\beta(h_\alpha)(x_\beta - x_{-\beta}) \in B\\
[i(x_\alpha + x_{-\alpha}), i(x_\beta + x_{-\beta})] &= -c(\alpha,\beta)(x_{\alpha+\beta} - x_{-(\alpha+\beta)})\\
&\quad - c'(\alpha,-\beta)(x_{\alpha-\beta} - x_{-\alpha+\beta}).
\end{aligned}$$

From proposition 53.91, it just remains to prove that the Killing form of $\mathfrak{u}_0$ is strictly negative definite. We know that $B_\mathfrak{g}(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ if $\alpha, \beta \in \Phi$ and $\alpha + \beta \neq 0$; then $A \perp B$ and $A \perp C$. It is a lot of computation to compute the Killing form; we know that B is strictly positive definite on $\sum_{\alpha\in\Delta} \mathbb{R}h_\alpha$ (and then strictly negative definite on A) a part this, the non zero elements are (recall that if $\alpha \neq 0$, $B(x_\alpha, x_\alpha) = 0$ from corollary 51.76)

$$\begin{aligned}
B((x_\alpha - x_{-\alpha}), (x_\alpha - x_{-\alpha})) &= -2B(x_\alpha, x_{-\alpha}) = -2\\
B(i(x_\alpha + x_{-\alpha}), i(x_\alpha, x_{-\alpha})) &= -2.
\end{aligned}$$

What we have in the matrix of $B_\mathfrak{g}|_{\mathfrak{u}_0\times\mathfrak{u}_0}$ is a negative definite block (corresponding to A), -2 on the rest of the diagonal and zero anywhere else. Then it is well negative definite and $\mathfrak{u}_0$ is a compact real from of $\mathfrak{g}$. □

53.15.5 Involutions

Let $\mathfrak{g}$ be a (real or complex) Lie algebra. An automorphism $\sigma : \mathfrak{g} \to \mathfrak{g}$ which is not the identity such that σ^2 is the identity is a **involution**. An involution $\theta : \mathfrak{f} \to \mathfrak{f}$ of a *real* semisimple Lie algebra $\mathfrak{f}$ such that the quadratic form B_θ defined by

$$B_\theta(X, Y) := -B(X, \theta Y)$$

is positive definite is a **Cartan involution**.

Proposition 53.122.
Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{u}_0$ a compact real form and τ, the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{u}_0$. Then τ is a Cartan involution of $\mathfrak{g}^\mathbb{R}$. [prop:conj_invol]

Proof. From the assumptions, $\mathfrak{g} = \mathfrak{u}_0 \oplus i\mathfrak{u}_0$, $\tau_{\mathfrak{u}_0} = id$ and $\tau_{i\mathfrak{u}_0} = -id$; then it is clear that $\tau^2_{\mathfrak{g}^\mathbb{R}} = id|_{\mathfrak{g}^\mathbb{R}}$. If $Z \in \mathfrak{g}$, we can decompose into $Z = X + iY$ with $X, Y \in \mathfrak{u}_0$. For $Z \neq 0$, we have

$$B_\mathfrak{g}(Z, \tau Z) = B_\mathfrak{g}(X + iY, X - iY) = B_\mathfrak{g}(X, X) + B_\mathfrak{g}(Y, Y) = B_{\mathfrak{u}_0}(X, X) + B_{\mathfrak{u}_0}(Y, Y) < 0 \quad (53.294)$$

because B restricts itself to $\mathfrak{u}_0$ which is compact. Then

$$(B_{\mathfrak{g}^\mathbb{R}})_\tau(Z, Z') = B_{\mathfrak{g}^\mathbb{R}}(Z, \tau Z) = -2\,\mathrm{Re}\, B_\mathfrak{g}(Z, \tau Z') \quad (53.295)$$

is positive definite because $(B_\mathfrak{g})_\tau$ is negative definite. Thus τ is a Cartan involution of $\mathfrak{g}^\mathbb{R}$. □

Lemma 53.123.
If φ and ψ are involutions of a vector space V (we denote by V_{ψ^+} and V_{ψ^-} the subspaces of V for the eigenvalues 1 and -1 of ψ and similarly for φ), then

$$[\varphi,\psi]=0 \quad \textit{iff} \quad \begin{cases} V_{\varphi^+} = (V_{\varphi^+}\cap V_{\psi^+})\oplus(V_{\varphi^+}\cap V_{\psi^-}) \\ V_{\varphi^-} = (V_{\varphi^-}\cap V_{\psi^+})\oplus(V_{\varphi^-}\cap V_{\psi^-}), \end{cases}$$

i.e. if and only if the decomposition of V with respect to φ is "compatible" with the one with respect to ψ. lem:invol_compat

Proof. Direct sense. Let us first see that φ leaves the decomposition $V=V_{\psi^+}\oplus V_{\psi^-}$ invariant. If $x=x_{\psi^+}+x_{\psi^-}$,

$$\varphi(x_{\psi^+})=(\varphi\circ\psi)(x_{\psi^+})=(\psi\circ\varphi)(x_{\psi^+}).$$

Then $\varphi(x_{\psi^+})\in V_{\psi^+}$, and the matrix of φ is block-diagonal with respect to the decomposition given by ψ. Thus V_{ψ^+} and V_{ψ^-} split separately into two parts with respect to φ.
Inverse sense. If $x\in V$, we can write $x=x_{++}+x_{+-}+x_{-+}+x_{--}$ where the first index refers to ψ while the second one refers to ψ; for example, $x_{+-}\in V_{\psi^+}\cap V_{\varphi^-}$. The following computation is easy:

$$\begin{aligned}(\varphi\circ\psi)(x) &= \varphi(x_{++}+x_{+-}-x_{-+}-x_{--}) \\ &= x_{++}-x_{+-}-x_{-+}+x_{--} \\ &= \psi(x_{++}-x_{+-}-x_{-+}-x_{--}) \\ &= (\psi\circ\varphi)(x).\end{aligned} \tag{53.296}$$

□

cor:Cartan_conj_inner

Corollary 53.124.
Any two Cartan involutions of a real semisimple Lie algebra are conjugate by an inner automorphism.

Proof. Let σ and σ' be two Cartan involutions of $\mathfrak{f}$. We can find a $\varphi\in\inf\mathfrak{f}$ such that $[\varphi\sigma\varphi^{-1},\sigma']=0$. Thus it is sufficient to prove that any two Cartan involutions which commute are equals. So let us consider θ and θ', two Cartan involutions such that $[\theta,\theta']=0$. By lemma 53.123, we know that the decompositions into $+1$ and -1 eigenspaces with respect to θ and θ' are compatibles. If we consider $X\in\mathfrak{f}$ such that $\theta X=X$ and $\theta' X=-1$ (it is always possible if $\theta\neq\theta'$), we have

$$\begin{aligned}0<B_\theta(X,X)&=-B(X,\theta X)=-B(X,X)\\ 0<B_{\theta'}(X,X)&=-B(X,\theta' X)=B(X,X)\end{aligned}$$

which is impossible. □

Theorem 53.125.
Let $\mathfrak{f}$ be a real semisimple Lie algebra, θ a Cartan involution on $\mathfrak{f}$ and σ, another involution (not specially Cartan). Then there exists a $\varphi\in\operatorname{Int}\mathfrak{f}$ such that $[\varphi\theta\varphi^{-1},\sigma]=0$ tho:sigma_theta_un

Proof. If θ is a Cartan involution, then B_θ is a scalar product on $\mathfrak{f}$. Let $\omega=\sigma\theta$. By using $\sigma^2=\theta^2=1$, $\theta=\theta^{-1}$ and the invariance property 51.17 of the Killing form,

$$B(\omega X,\theta Y)=B(X,\omega^{-1}\theta Y)=B(X,\theta\sigma\theta Y)=B(X,\theta\omega Y). \tag{53.297}$$

Then $B_\theta(\omega X,Y)=B_\theta(X,\omega Y)$. This is a general property of scalar product that in this case, the matrix of ω is symmetric while the one of ω^2 is positive definite. If we consider the classical scalar product whose matrix is (δ_{ij}), the property is written as $A_{ij}v_jw_j=v_iA_{ij}w_j$ (with sum over i and j); this implies the symmetry of A. To see that A^2 is positive definite, we compute (using the symmetry):

$$A_{ij}A_{jk}v_iv_k=v_iA_{ij}v_kA_{kj}=\sum_j(v_iA_{ij})^2>0.$$

The next step is to see that there is an unique linear transformation $A\colon \mathfrak{f} \to \mathfrak{f}$ such that $\omega^2 = e^A$, and that for any $t \in \mathbb{R}$, the transformation e^{tA} is an automorphism of $\mathfrak{f}$.

We choose an orthonormal (with respect to the inner product B_θ) basis $\{X_1, \dots, X_n\}$ of $\mathfrak{f}$ in which ω is diagonal. In this basis, ω^2 is also diagonal and has positive real numbers on the diagonal; then the existence and unicity of A is clear. Now we take some notations:

$$\omega(X_i) = \lambda_i X_i \tag{53.298a}$$
$$\omega^2(X_i) = e^{a_i} X_i, \tag{53.298b}$$

(no sum at all) where the a_i are the diagonals elements of A. The structure constants are as usual defined by

$$[X_i, X_j] = c_{ij}^k X_k. \tag{53.299}$$

Since σ and θ are automorphisms, ω^2 is also one. Then

$$\omega^2[X_i, X_j] = c_{ij}^k \omega^2(X_k) = c_{ij}^k e^{a_k} X_k$$

can also be computed as

$$\omega^2[X_i, X_j] = [\omega^2 X_i, \omega^2 X_j] = e^{a_i} e^{a_j} c_{ij}^k X_k,$$

so that $c_{ij}^k e^{a_k} = c_{ij}^k e^{a_i} e^{a_j}$, and then $\forall t \in \mathbb{R}$,

$$c_{ij}^k e^{t a_k} = c_{ij}^k e^{t a_i} e^{t a_j},$$

which proves that e^{tA} is an automorphism of $\mathfrak{f}$. By lemma 53.74, A is thus a derivation of $\mathfrak{f}$. The semi-simplicity makes $\partial\mathfrak{f} = \operatorname{ad}\mathfrak{f}$, then $A \in \operatorname{ad}\mathfrak{f}$ and $e^{tA} \in \operatorname{Int}\mathfrak{f}$ because it clearly belongs to the identity component of $\operatorname{Aut}\mathfrak{f}$.

Now we can finish de proof by some computations. Remark that $\omega = e^{A/2}$ and $[e^{tA}, \omega] = 0$ because it can be seen as a common matrix commutator. Since $\omega^{-1} = \theta\sigma$, we have $\theta\omega^{-1}\theta = \sigma\theta$, or $\theta\omega^2\theta = \omega^2$ and

$$e^A \theta = \theta e^{-A}. \tag{53.300}$$ eq:eAth

From this, one can deduce that $e^{tA}\theta = \theta e^{-tA}$. Indeed, as matricial identity, equation (53.300) reads

$$(e^A\theta)_{ik} = (e^A)_{ij}\theta_{jk} = e^{a_i}\theta_{ik} = e^{-a_k}\theta_{ik}.$$

Then for any ik such that $\theta_{ik} \neq 0$, we find $e^{a_i} = e^{-a_k}$ and then also $e^{ta_i} = e^{-ta_k}$. Thus $(e^{tA}\theta)_{ik} = (e^{tA})_i j\theta_{jk} = e^{ta_i}\theta_{ik} = \theta_{ik}e^{-ta_k} = (\theta e^{-tA})_{ik}$. So we have

$$e^{tA}\theta = \theta e^{-tA}. \tag{53.301}$$

Now we consider $\varphi = e^{A/4} \in \operatorname{Int}\mathfrak{f}$ and $\theta_1 = \varphi\theta\varphi^{-1}$. We find $\theta_1\sigma = e^{A/2}\omega^{-1}$ and $\sigma\theta^{-1} = e^{-A/2}\omega$. Since $\omega^2 = A$, we have $e^{A/2} = e^{-A/2}\omega^2$ and thus $\theta_1\sigma = \sigma\theta_1$.

□

Corollary 53.126.
Every real Lie algebra has a Cartan involution.

Proof. Let $\mathfrak{f}$ be a real Lie algebra and $\mathfrak{g}$ be his complexification: $\mathfrak{g} = \mathfrak{f}^{\mathbb{C}}$. Let $\mathfrak{u}_0$ be a compact real form of $\mathfrak{g}$ and τ the induced involution (the conjugation) on $\mathfrak{g}$. By the proposition 53.122, we know that τ is a Cartan involution of $\mathfrak{g}^{\mathbb{R}}$. We also consider σ, the involution of $\mathfrak{g}$ with respect to the real form $\mathfrak{f}$. It is in particular an involution on the real Lie algebra $\mathfrak{f}$. Then one can find a $\varphi \in \operatorname{Int}\mathfrak{g}^{\mathbb{R}}$ such that $[\varphi\tau\varphi^{-1}, \sigma] = 0$ on $\mathfrak{g}^{\mathbb{R}}$. Let $\mathfrak{u}_1 = \varphi\mathfrak{u}_0$ and $X \in \mathfrak{u}_1$. We can write $X = \varphi Y$ for a certain $Y \in \mathfrak{u}_0$. Then

$$\varphi\tau\varphi^{-1}X = \varphi\tau Y = \varphi Y = X,$$

so that $\varphi\tau\varphi^{-1} = id|_{\mathfrak{u}_1}$. Note that $\mathfrak{u}_1$ is also a real compact form of $\mathfrak{g}$ because the Killing form is not affected by φ. Let τ_1 be the involution of $\mathfrak{g}$ induced by $\mathfrak{u}_1$. We have

$$\tau_1|_{\mathfrak{u}_1} = \varphi\tau\varphi_{\mathfrak{u}_1}^{-1} = \mathrm{Id}\,|_{\mathfrak{u}_1}.$$

Since φ is $\mathbb{C}$-linear, we have in fact $\tau_1 = \varphi\tau\varphi^{-1}$. Now we forget $\mathfrak{u}_0$ and we consider the compact real form $\mathfrak{u}_1$ with his involution τ_1 of $\mathfrak{g}$ which satisfy $[\tau_1, \sigma] = 0$ on $\mathfrak{g}^{\mathbb{R}}$ This relation holds also on $i\mathfrak{g}^{\mathbb{R}}$, then

$$[\tau_1, \sigma] = 0$$

on $\mathfrak{g} = \mathfrak{f}^{\mathbb{C}}$. Let $X \in \mathfrak{f}$, i.e. $\sigma X = X$; it automatically fulfils

$$\sigma\tau_1 X = \tau_1\sigma X = \tau_1 X,$$

so that τ_1 restrains to an involution on $\mathfrak{f}$ (because $\tau_1\mathfrak{f} \subset \mathfrak{f}$). Let $\theta = \tau_1|_{\mathfrak{f}}$. For X, $Y \in \mathfrak{f}$, we have

$$B_\theta(X, Y) = -B_{\mathfrak{f}}(X, \theta Y) = -B_{\mathfrak{f}}(X, \tau Y) = \frac{1}{2}(B_{\mathfrak{g}^{\mathbb{R}}})_{\tau_1}(X, Y), \tag{53.302}$$

which shows that θ is a Cartan involution. The half factor on the last line comes from the fact that $\mathfrak{g}^{\mathbb{R}} = (\mathfrak{f}^{\mathbb{C}})^{\mathbb{R}} = \mathfrak{f} \oplus i\mathfrak{f}$.

□

53.15.6 Cartan decomposition

Examples of Cartan and Iwasawa decomposition are given in sections 53.17, 66.4.3,66.5, 55.8 and 55.7. An example of how it works to prove isomorphism of Lie algebras is provided in subsection 66.5.7.

Let $\mathfrak{f}$ be a real semisimple Lie algebra. A vector space decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ is a **Cartan decomposition** if the Killing form is negative definite on $\mathfrak{k}$ and positive definite on $\mathfrak{p}$ and the following commutators hold:

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}. \tag{53.303}$$ eq:comm_Cartan

If $X \in \mathfrak{k}$ and $Y \in \mathfrak{p}$, we have $(\operatorname{ad} X \circ \operatorname{ad} Y)\mathfrak{k} \subseteq \mathfrak{p}$ and $(\operatorname{ad} X \circ \operatorname{ad} Y)\mathfrak{p} \subseteq \mathfrak{k}$, therefore $B_{\mathfrak{f}}(X, Y) = 0$.

Let $\theta\colon \mathfrak{f} \to \mathfrak{f}$ be a Cartan involution, $\mathfrak{k}$ its $+1$ eigenspace and $\mathfrak{p}$ his -1 one. It is easy to see that the relations (53.303) are satisfied for the decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$. For example, for $X, X' \in \mathfrak{k}$, using the fact that θ is an automorphism,

$$[X, X'] = [\theta X, \theta X'] = \theta[X, X'],$$

which proves that $[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}$. Since θ is a Cartan involution, B_θ is positive definite. For $X \in \mathfrak{k}$,

$$B(X, X) = B(X, \theta X) = -B_\theta(X, X)$$

proves that B is negative definite on $\mathfrak{k}$; in the same way we find that B is also positive definite on $\mathfrak{p}$. Then the Cartan involution gives rise to a Cartan decomposition. We are going to prove that any Cartan decomposition defines a Cartan involution.

Let us now do the converse. Let $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition of the real semisimple Lie algebra $\mathfrak{f}$. We define $\theta = \mathrm{Id}\,|_{\mathfrak{k}} \oplus (-\mathrm{Id})|_{\mathfrak{p}}$. If $X, X' \in \mathfrak{k}$, the definition of a Cartan algebra makes $[X, X'] \in \mathfrak{k}$ and so

$$\theta[X, X'] = [X, X'] = [\theta X, \theta X'],$$

and so on, we prove that θ is an automorphism of $\mathcal{F}$. It remains to prove that B_θ is positive definite. If $X \in \mathfrak{k}$,

$$B_\theta(X, X) = -B(X, \theta X) = -B(X, X).$$

Then B_θ is positive definite on $\mathfrak{k}$ because on this space, B is negative definite by definition of a Cartan involution. The same trick shows that B_θ is also positive definite on $\mathfrak{p}$. We had seen that $\mathfrak{p}$ and $\mathfrak{k}$ where B_θ-orthogonal spaces. Thus B_θ is positive definite and θ is a Cartan involution.

Let $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition. Then it is quite easy to see that $\mathfrak{k} \oplus i\mathfrak{p}$ is a compact real form of $\mathfrak{g} = (\mathfrak{f}^{\mathbb{C}})$.

Proposition 53.127.
Let $\mathfrak{L}$ and $\mathfrak{q}$ be the $+1$ and -1 eigenspaces of an involution σ. Then σ is a Cartan involution if and only if $\mathfrak{L} \oplus i\mathfrak{q}$ is a compact real form of $\mathfrak{f}^{\mathbb{C}}$.

Proof. First remark that $\mathfrak{L} \oplus i\mathfrak{q}$ is always a real form of $\mathfrak{f}^{\mathbb{C}}$. The direct sense is yet done. Then we suppose that $B_{\mathfrak{f}^{\mathbb{C}}}$ is negative definite on $\mathfrak{L} \oplus i\mathfrak{q}$ and we have to show that $\mathfrak{L} \oplus \mathfrak{q}$ is a Cartan decomposition of $\mathfrak{f}$. The condition about the brackets on $\mathfrak{L}$ and $\mathfrak{q}$ is clear from their definitions. If $X \in \mathfrak{L}$, $B(X, X) < 0$ because B is negative definite on $\mathfrak{L}$. If $Y \in \mathfrak{q}$, $B(Y, Y) = -B(iY, iY) > 0$ because B is negative definite on $i\mathfrak{q}$. □

53.16 Root spaces in the real case

Let $\mathfrak{f}$ be a real semisimple Lie algebra with a Cartan involution θ and the corresponding Cartan decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$. We consider B, a "Killing like" form, i.e. B is a symmetric nondegenerate invariant bilinear form on $\mathfrak{f}$ such that $B(X, Y) = B(\theta X, \theta Y)$ and $B_\theta := -B(X, \theta X)$ is positive definite. Then B is negative definite on the compact real form $\mathfrak{k} \oplus i\mathfrak{p}$. Indeed if $Y \in \mathfrak{p}$,

$$B(iY, iY) = -B(\theta Y, \theta Y) = B(Y, \theta Y) = -B_\theta(Y, Y) < 0. \tag{53.304}$$

The case with $X \in \mathfrak{k}$ is similar. It is easy to see that B_θ is in fact a scalar product on $\mathfrak{f}$, so that we can define the orthogonality and the adjoint from B_θ. If $A \colon \mathfrak{f} \to \mathfrak{f}$ is an operator on $\mathfrak{f}$, his adjoint is the operator A^* given by the formula

$$B_\theta(AX, Y) = B_\theta(X, A^*Y)$$

for all X, $Y \in \mathfrak{f}$.

Proposition 53.128.
With this definition, when $X \in \mathfrak{f}$, the adjoint operator of $\operatorname{ad} X$ is given by means of the Cartan involution:

$$(\operatorname{ad} X)^* = \operatorname{ad}(\theta X).$$

Proof. This is a simple computation

$$B_\theta\big((\operatorname{ad}\theta X)Y, Z\big) = -B\big(Y, [\theta X, \theta Y]\big) = -B_\theta(Y, [X, Z]) = -B_\theta\big((\operatorname{ad} X)^* Y, Z\big). \tag{53.305}$$

□

Let $\mathfrak{a}$ be a maximal abelian subalgebra of $\mathfrak{p}$ (the existence comes from the finiteness of the dimensions). If $H \in \mathfrak{a}$, the operator $\operatorname{ad} H$ is self-adjoint because

$$(\operatorname{ad} H)^* X = (-\operatorname{ad}\theta H)X = [H, X] = (\operatorname{ad} H)X, \tag{53.306}$$

where we used the fact that $H \in \mathfrak{p}$. For $\lambda \in \mathfrak{a}^*$, we define the space

$$\mathfrak{f}_\lambda = \{X \in \mathfrak{f} \text{ tel que } \forall H \in \mathfrak{a},\ (\operatorname{ad} H)X = \lambda(H)X\}. \tag{53.307}$$

If $\mathfrak{f}_\lambda \neq 0$ and $\lambda \neq 0$, we say that λ is a **restricted root** of $\mathfrak{f}$. We denote by Σ the set of restricted roots of $\mathfrak{f}$. We may sometimes write $\Sigma_{\mathfrak{f}}$ if the Lie algebra is ambiguous.

The main properties of the real root spaces are given in the following proposition.

PropPropRacincesReelles

Proposition 53.129.
The set Σ of the restricted roots of a real semisimple Lie algebra $\mathfrak{f}$ has the following properties: prop:enuc

(1) $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma} \mathfrak{f}_\lambda$, enuci

(2) $[\mathfrak{f}_\lambda, \mathfrak{f}_\mu] \subseteq \mathfrak{f}_{\lambda+\mu}$, enucii

(3) $\theta\mathfrak{f}_\lambda = \mathfrak{f}_{-\lambda}$, enuciii

enuciv

(4) $\lambda \in \Sigma$ *implies* $-\lambda \in \Sigma$,

enucv

(5) $\mathfrak{f}_0 = \mathfrak{a} \oplus \mathfrak{m}$ *where* $\mathfrak{m} = \mathcal{Z}_{\mathfrak{k}}(\mathfrak{a})$ *and* $\mathfrak{a} \perp \mathfrak{m}$.

Proof. Proof of (1). The operators $\operatorname{ad} H$ with $H \in \mathfrak{a}$ form an abelian algebra of self-adjoint operators, then they are simultaneously diagonalisable. Let $\{X_i\}$ be a basis which realize this diagonalisation, and $\mathfrak{f}_i = \operatorname{Span} X_i$, so that $\mathfrak{f} = \oplus_i \mathfrak{f}_i$. We have $(\operatorname{ad} H)\mathfrak{f}_i = \mathfrak{f}_i$ and then $(\operatorname{ad} H)X_i = \lambda_i(H)X_i$ for a certain $\lambda_i \in \mathfrak{a}^*$. This shows that $\mathfrak{f}_i \subseteq \mathfrak{f}_{\lambda_i}$.[55]

Proof of (2). Let $H \in \mathfrak{a}$, $X \in \mathfrak{f}_\lambda$ and $Y \in \mathfrak{f}_\mu$. We have

$$(\operatorname{ad} H)[X,Y] = [[H,X],Y] + [X,[H,Y]] = \big(\lambda(H) + \mu(H)\big)[X,Y]. \tag{53.308}$$

Proof of (3). Using the fact that $\theta H = -H$ because $H \in \mathfrak{p}$,

$$(\operatorname{ad} H)\theta X = \theta[\theta H, X] = -\theta\lambda(H)X = -\lambda(H)(\theta X). \tag{53.309}$$

Proof of (4). It is a consequence of (3) because if $\mathfrak{f}_\lambda \neq 0$, then $\theta\mathfrak{f}_\lambda \neq 0$.

Proof of (5). By (3), $\theta\mathfrak{f}_0 = \mathfrak{f}_0$, then $\mathfrak{f}_0 = (\mathfrak{k} \cap \mathfrak{f}_0) \oplus (\mathfrak{p} \cap \mathfrak{f}_0)$. If $X \in \mathfrak{f}_0$, then it commutes with all the elements of $\mathfrak{a}$ and by the maximality property of $\mathfrak{a}$, provided that $X \in \mathfrak{p}$, it also must belongs to $\mathfrak{a}$. This fact makes $\mathfrak{a} = \mathfrak{p} \cap \mathfrak{f}_0$. Now,

$$\mathfrak{m} = \mathcal{Z}_{\mathfrak{k}}(\mathfrak{a}) = \{X \in \mathfrak{k} \text{ tel que } [X,\mathfrak{a}] = 0\} = \mathfrak{k} \cap \mathfrak{f}_0.$$

All this gives $\mathfrak{f}_0 = \mathcal{Z}_{\mathfrak{k}}(\mathfrak{a}) \oplus \mathfrak{a}$. □

We choose a positivity notion on $\mathfrak{a}^*$, we consider Σ^+, the set of restricted positive roots and we define

$$\mathfrak{n} = \bigoplus_{\lambda \in \Sigma^+} \mathfrak{f}_\lambda.$$

From finiteness of the dimension, there are only a finitely many forms $\lambda \in \mathfrak{a}^*$ such that $\mathfrak{f}_\lambda \neq 0$. Then, taking, more and more commutators in $\mathfrak{n}$, the formula $[\mathfrak{f}_\lambda, \mathfrak{f}_\mu] \subseteq \mathfrak{f}_{\lambda+\mu}$ shows that the result finish to fall into a $\mathfrak{f}_\mu = 0$. On the other hand, since $\mathfrak{a} \subset \mathfrak{f}_0$, we have $[\mathfrak{a},\mathfrak{n}] = \mathfrak{n}$. If $a_1, a_2 \in \mathfrak{a}$ and $n_1, n_2 \in \mathfrak{n}$,

$$[a_1 + n_1, a_2 + n_2] = \underbrace{[a_1,a_2]}_{=0} + \underbrace{[a_1,n_2]}_{\in\mathfrak{n}} + \underbrace{[n_1,a_2]}_{\in\mathfrak{n}} + \underbrace{[n_1,n_2]}_{\in\mathfrak{n}}, \tag{53.310}$$

then $[\mathfrak{a} \oplus \mathfrak{n}, \mathfrak{a} \oplus \mathfrak{n}] = \mathfrak{n}$. This proves the three following important properties:

(1) $\mathfrak{n}$ is nilpotent.

(2) $\mathfrak{a}$ is abelian.

(3) $\mathfrak{a} \oplus \mathfrak{n}$ is a solvable Lie subalgebra of $\mathfrak{f}$.

53.16.1 Iwasawa decomposition

Theorem 53.130.
Let $\mathfrak{f}$ be a real semisimple Lie algebra and $\mathfrak{k}$, $\mathfrak{a}$, $\mathfrak{n}$ as before. Then we have the following direct sum:

$$\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}. \tag{53.311}$$

This is the **Iwasawa decomposition** for the real semisimple Lie algebra $\mathfrak{f}$.

55. pourquoi ça n'implique pas que $\dim \mathfrak{f}_{\lambda_i} = 1$? Réponse par Philippe: tu as oublié les valeurs propres nulles dans ta base ce qui doit entrainer quelques modifs dans ton texte(par ex. $adH f_i = f_i$ pas toujours)

Proof. We yet know the direct sum $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma} \mathfrak{f}_\lambda$. Roughly speaking, in $\mathfrak{n}$ we have only vectors of Σ^+, in $\theta\mathfrak{n}$, only of Σ^- and in $\mathfrak{a}$, only in "zero". Then the sum $\mathfrak{a}\oplus\mathfrak{n}\oplus\theta\mathfrak{n}$ is direct.

Now we prove that the sum $\mathfrak{k}+\mathfrak{a}+\mathfrak{n}$ is also direct. It is clear that $\mathfrak{a}\cap\mathfrak{n}=0$ because $\mathfrak{a}\subseteq\mathfrak{f}_0$. Let $X\in\mathfrak{k}\cap(\mathfrak{a}\oplus\mathfrak{n})$. Then $\theta X = X$. But $\theta X\in\mathfrak{a}\oplus\theta\mathfrak{n}$. Thus $X\in\mathfrak{a}\oplus\mathfrak{n}\cap\mathfrak{a}\oplus\mathfrak{n}$ which implies $X\in\mathfrak{a}$. All this makes $X\in\mathfrak{p}\oplus\mathfrak{k}$ and $X=0$.

Now we prove that $\mathfrak{k}\oplus\mathfrak{a}\oplus\mathfrak{n}=\mathfrak{f}$. An arbitrary $X\in\mathfrak{f}$ can be written as

$$X = H + X_0 + \sum_{\lambda\in\Sigma} X_\lambda$$

where $H\in\mathfrak{a}$, $X_0\in\mathfrak{m}$ and $X_\lambda\in\mathfrak{f}_\lambda$. Now there are just some manipulations...

$$\sum_{\lambda\in\Sigma} X_\lambda = \sum_{\lambda\in\Sigma^+}(X_{-\lambda}+X_\lambda) = \sum_{\lambda\in\Sigma^+}(X_{-\lambda}+\theta X_{-\lambda}) + \sum_{\lambda\in\Sigma^+}(X_\lambda+\theta X_{-\lambda}), \tag{53.312}$$

but $\theta(X_{-\lambda}+\theta X_{-\lambda}) = X_{-\lambda}+\theta X_{-\lambda}$, then $X_{-\lambda}+X_{-\lambda}\in\mathfrak{k}$. Moreover, $X_\lambda, \theta X_{-\lambda}\in\mathfrak{f}_\lambda$, then $X_\lambda - \theta X_{-\lambda}\in\mathfrak{f}_\lambda\subseteq\mathfrak{n}$. Then

$$X = X_0 + \sum_{\lambda\in\Sigma^+}(X_{-\lambda}+\theta X_{-\lambda}) + H + \sum_{\lambda\in\Sigma^+}(X_\lambda - \theta X_{-\lambda}) \tag{53.313}$$

where the two first term belong to $\mathfrak{k}$, $H\in\mathfrak{a}$ and the last term belongs to $\mathfrak{n}$. □

Lemma 53.131.
There exists a basis $\{X_i\}$ of $\mathfrak{f}$ in which

enudi

(1) The matrices of $\operatorname{ad}\mathfrak{k}$ are symmetric,

enudii

(2) The matrices of $\operatorname{ad}\mathfrak{a}$ are diagonal and real,

enudiii

(3) The matrices of $\operatorname{ad}\mathfrak{n}$ are upper triangular with zeros on the diagonal.

Proof. We have the orthogonal decomposition $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma}\mathfrak{f}_\lambda$ given by proposition 53.129. Let $\{X_i\}$ be an orthogonal basis of $\mathfrak{f}$ compatible with this decomposition and in such an order that $i<j$ implies $\lambda_i\geqslant\lambda_j$. From the orthogonality of the basis it follows that the matrix of B_θ is diagonal. Thus the adjoint is the transposition.

(1) If $X\in\mathfrak{k}$, $(\operatorname{ad}X)^t = (\operatorname{ad}X)^* = -\operatorname{ad}\theta X = -\operatorname{ad}X$.

(2) Each X_i is a restricted root; then $(\operatorname{ad}H)X_i = \lambda_i(H)X_i$, then the diagonal of $\operatorname{ad}H$ is made of $\lambda_i(H)$ whose are real.

(3) If $Y_i\in\mathfrak{f}_{\lambda_i}$ with $\lambda_i\in\Sigma^+$, $(\operatorname{ad}Y_i)X_j$ has only components in $\mathfrak{f}_{\lambda_i+\lambda_j}$ with $\lambda_i+\lambda_j>\lambda_j$ because $\lambda_i\in\Sigma^+$. □

Lemma 53.132.
Let $\mathfrak{h}$ be a subalgebra of the real semisimple Lie algebra $\mathfrak{f}$. Then $\mathfrak{h}$ is a Cartan subalgebra if and only if $\mathfrak{h}^{\mathbb{C}}$ is Cartan in $\mathfrak{f}^{\mathbb{C}}$.

Proof. Direct sense. If $\mathfrak{h}$ is nilpotent in $\mathfrak{f}$, it is cleat that $\mathfrak{h}^{\mathbb{C}}$ is nilpotent in $\mathfrak{f}^{\mathbb{C}}$. We have to prove that $[x,\mathfrak{h}^{\mathbb{C}}]\subseteq\mathfrak{h}^{\mathbb{C}}$ implies $x\in\mathfrak{h}^{\mathbb{C}}$. As set, $\mathfrak{f}^{\mathbb{C}} = \mathcal{F}\oplus i\mathfrak{f}$ (but not as vector space!), then we can write $x=a+ib$ with $a,b\in\mathfrak{f}$. The assumption makes that for any $h\in\mathfrak{h}$, there exists $h',h''\in\mathfrak{h}$ such that

$$[a+ib,h] = h+ih''.$$

This equation can be decomposed in $\mathfrak{f}$-part and $i\mathfrak{f}$-part: for any $h\in\mathfrak{h}$, there exists a $h'\in\mathfrak{h}$ such that $[a,h]=h'$, and for any $h\in\mathfrak{h}$, there exists a $h''\in\mathfrak{h}$ such that $[b,h]=h''$. Thus a, $b\in\mathfrak{h}$ because $\mathfrak{h}$ is Cartan in $\mathfrak{f}$.

Inverse sense. The assumption is that $[x,\mathfrak{h}^{\mathbb{C}}]\subset\mathfrak{h}^{\mathbb{C}}$ implies $x\in\mathfrak{h}^{\mathbb{C}}$. In particular consider a $x\in\mathfrak{h}$ such that $[x,\mathfrak{h}]\subset\mathfrak{h}$. Then $x\in\mathfrak{h}^{\mathbb{C}}$ because $[x,\mathfrak{h}^{\mathbb{C}}]\subset\mathfrak{h}^{\mathbb{C}}$. But $\mathfrak{h}^{\mathbb{C}}\cap\mathfrak{f}=\mathfrak{h}$. □

In the complex case, the Cartan subalgebras all have same dimensions because they are maximal abelian.

53.17 The group $SL(2,\mathbb{R})$ and its algebra

SecToolSL

The study of $\mathrm{SL}(2,\mathbb{R})$ and $\mathfrak{sl}(2,\mathbb{C})$ is required before to go further in the general study because of proposition 51.79 that will reduce the study of genera Lie algebras into combinations of $\mathfrak{sl}(2,\mathbb{C})$ algebras.

53.17.1 Iwasawa decomposition

Let $G = \mathrm{SL}(2,\mathbb{R})$ the group of 2×2 matrices with unit determinant. The Lie algebra $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{R})$ is the algebra of matrices with vanishing trace:

$$\mathfrak{g} = \{X \in \mathrm{End}(\mathbb{R}^2) \text{ tel que } \mathrm{Tr}(X) = 0\} = \left\{ \begin{pmatrix} x & y \\ z & -x \end{pmatrix} \text{ with } x,y,z \in \mathbb{R} \right\}. \tag{53.314}$$

The following elements will be intensively used:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{53.315}$$

EqsXPdnZlG

where $T = E - F$ has been introduced for later convenience. The commutators are

EqTableSLdR

$$[H,E] = 2E \qquad [T,H] = -2T \tag{53.316a}$$
$$[H,F] = -2F \qquad [T,E] = H \tag{53.316b}$$
$$[E,F] = H \qquad [T,F] = H. \tag{53.316c}$$

The exponentials can be easily computed and the result is

$$e^{tH} = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \qquad e^{tE} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad e^{tF} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}. \tag{53.317}$$

EqExpMatrsSLdeuxR

Notice that the sets $\{H,E,F\}$, $\{H,E,F\}$ and $\{H,E+F,T\}$ are basis. A Cartan involution is given by $\theta(X) = -X^t$, and the corresponding Cartan decomposition is

$$\mathfrak{k} = \mathrm{Span}\{T\}, \qquad \mathfrak{p} = \mathrm{Span}\{H, E+F\}. \tag{53.318}$$

Indeed, we are in a matrix algebra, then $\mathrm{Tr}(XY)$ is proportional to $\mathrm{Tr}(\mathrm{ad}\,X \circ \mathrm{ad}\,Y)$. In order to see that θ is a Cartan involution, we have to prove that $B|_{\mathfrak{k}\times\mathfrak{k}}$ is negative definite and $B|_{\mathfrak{p}\times\mathfrak{p}}$ positive. It is true because for $X \in \mathfrak{k}$,

$$\mathrm{Tr}(\mathrm{ad}\,X \circ \mathrm{ad}\,X) = \mathrm{Tr}(XX) = \mathrm{Tr}\begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix} < 0,$$

and for $Y \in \mathfrak{p}$,

$$\mathrm{Tr}(YY) = \mathrm{Tr}\begin{pmatrix} x & y \\ y & -x \end{pmatrix}\begin{pmatrix} x & y \\ y & -x \end{pmatrix} = \mathrm{Tr}\begin{pmatrix} x^2+y^2 & 0 \\ 0 & x^2+y^2 \end{pmatrix} > 0.$$

pg_iwasldr

Up to some choices, the Iwasawa decomposition of the group $\mathrm{SL}(2,\mathbb{R})$ is given by the exponentiation of $\mathfrak{a}$, $\mathfrak{n}$ and $\mathfrak{k}$

$$\mathfrak{a} = \mathrm{Span}\{H\} \quad \mathfrak{n} = \mathrm{Span}\{E\} \quad \mathfrak{k} = \mathrm{Span}\{T\}, \tag{53.319}$$

so that

$$A = \begin{pmatrix} e^a & 0 \\ 0 & e^{-a} \end{pmatrix} \quad N = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \quad K = \begin{pmatrix} \cos k & \sin k \\ -\sin k & \cos k \end{pmatrix}. \tag{53.320}$$

eq:expo_ANK

A common parametrization of AN by $\mathbb{R}^2$ is provided by

$$(a,l) = \begin{pmatrix} e^a & le^a \\ 0 & e^{-a} \end{pmatrix}. \tag{53.321}$$

EqParmalSL

One immediately has the following formula for the left action of AN on itself:

$$L_{(a,l)}(a',l') = \begin{pmatrix} e^{a+a'} & e^{a+a'}l' + e^{a-a'}l \\ 0 & e^{-a-a'} \end{pmatrix} = (a+a', l' + e^{-2a'}l).$$

In this setting, the inverse is given by $(a,l)^{-1} = (-a, -le^{2a})$.

Let's give some formulas in $\mathrm{SL}(2,\mathbb{R})$. Using corollary 53.76 and exponentiating commutation relations, eq_eaHsldr

$$\mathrm{Ad}(e^{aH})E = e^{2a}E, \tag{53.322a}$$
$$\mathrm{Ad}(e^{aH})F = e^{-2a}F, \tag{53.322b}$$
$$\mathrm{Ad}(e^{aH})T = e^{2a}E - e^{-2u}E + e^{-2u}T \tag{53.322c}$$
$$\mathrm{Ad}(e^{tE})H = H - 2tE, \tag{53.322d}$$ eq_AdetE
$$\mathrm{Ad}(e^{tE})T = -tH + (t^2+1)E - E + T \tag{53.322e}$$
$$\mathrm{Ad}(e^{tT})H = \cos(2t)H + \sin(2t)(2E - T) \tag{53.322f}$$
$$\mathrm{Ad}(e^{tT})E = \frac{1}{2}\Big(\sin(2t)H + \cos(2t)(2E-T) + T\Big) \tag{53.322g}$$

where z belongs to the center: $z = \pm\mathbb{1}$.

53.17.2 A companion: $A\bar{N}$

We can consider the Iwasawa decomposition which is θ-conjugated to the AN that we just saw. That decomposition is generated by

$$\bar{\mathfrak{n}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{53.323}$$

The exponentiation produces

$$\bar{N} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \tag{53.324}$$

and the Iwasawa group is given by

$$A\bar{N} = \begin{pmatrix} e^{a} & 0 \\ le^{-a} & e^{-a} \end{pmatrix}. \tag{53.325}$$ EqGeneANbarSLdeuxR

53.17.3 Killing form

In the basis $\{H, E, T\}$, the adjoint operators are given by

$$\mathrm{ad}\,H = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix}, \mathrm{ad}\,E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ -2 & 0 & 0 \end{pmatrix}, \text{ and } \mathrm{ad}\,T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 0 \end{pmatrix}.$$

so that the Killing form can be computed directly from definition $B(X,Y) = \mathrm{Tr}(\mathrm{ad}\,X, \mathrm{ad}\,Y)$. The result is

$$B(T,H) = 0 \qquad B(H,H) = 8 \tag{53.326a}$$
$$B(T,E) = -4 \qquad B(E,E) = 0 \tag{53.326b}$$
$$B(H,E) = 0 \qquad B(T,T) = -4. \tag{53.326c}$$

Expressed in the basis $\{H, E, F\}$, the matrix of the Killing form reads

$$B = \begin{pmatrix} 8 & & \\ & & 4 \\ & 4 & \end{pmatrix} \tag{53.327}$$

while, in the basis $\{H, E+F, T\}$, we find

$$B = \begin{pmatrix} 8 & & \\ & 8 & \\ & & -8 \end{pmatrix}. \tag{53.328}$$

The latter is the reason of the name of the vector T: the sign of its norm is different, so that T is candidate to be a time-like direction.

53.17.4 Abstract root space setting

Looking on the table (53.316) from an abstract point of view, we see that E and F are eigenvectors of $\mathrm{ad}(H)$ with eigenvalues 2 and -2. So $\mathfrak{a} = \mathfrak{g}_0 = \mathbb{R}H$; $\mathfrak{g}_2 = \mathbb{R}E$; and $\mathfrak{g}_{-2} = \mathbb{R}F$. Using a more abstract notation, the table of $\mathrm{SL}(2,\mathbb{R})$ becomes

$$[A_0, A_2] = 2A_2 \tag{53.329a}$$
$$[A_0, A_{-2}] = -2A_{-2} \tag{53.329b}$$
$$[A_2, A_{-2}] = A_0. \tag{53.329c}$$

53.17.5 Isomorphism

As pointed out in the chapter II, §6 of [859], the map (seen as a conjugation in $\mathrm{SL}(2,\mathbb{C})$)

$$\begin{aligned} \psi\colon \mathrm{SU}(1,1) &\to \mathrm{SL}(2,\mathbb{R}) \\ U &\mapsto AUA^{-1} \end{aligned} \tag{53.330}$$

with $A = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ is an isomorphism between $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{SU}(1,1)$.

53.18 The complex algebra $\mathfrak{sl}(2,\mathbb{C})$ and its representations

The book [881] contains the representations of $\mathfrak{sl}(2,\mathbb{C})$.

The algebra $\mathfrak{sl}(2,\mathbb{C})$ is the complex algebra of complex 2×2 matrices with vanishing trace. As generating matrices, one can take the elements u_i of (51.54) and complete them by

$$v_1 = \frac{1}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad v_2 = \frac{1}{2}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \qquad v_3 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

which satisfy the commutation relations

$$[v_i, v_j] = -\epsilon_{ikj}u_k \tag{53.331a}$$
$$[v_i, u_j] = \epsilon_{ikj}v_k. \tag{53.331b}$$

Remarque 53.133.
This is not the algebra $\mathfrak{sl}(2,\mathbb{C})$ used in physics. The latter is the *four*-dimensional *real* algebra of trace vanishing 2×2 complex matrices. There is one more generator and the representation theory is different. Moreover the physics works with the *group* instead of the *algebra*.

The change of basis

$$x_j = \frac{1}{2}(u_j + iv_j), \qquad y_j = \frac{1}{2}(u_j - iv_j) \tag{53.332}$$

provides the simplification

$$[x_i, x_j] = \epsilon_{ijk}x_k \qquad [y_i, y_j] = \epsilon_{ijk}y_k \qquad [x_i, y_j] = 0, \tag{53.333}$$

so that, as algebras, we have the isomorphism

$$\mathfrak{sl}(2,\mathbb{C}) = \mathfrak{su}(2) \oplus \mathfrak{su}(2). \tag{53.334}$$

Thus the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ is determined by the one of $\mathfrak{su}(2)$.

Consider the space $\mathcal{P}_m$ of homogeneous polynomials of degree m in two variables with complex coefficients[[882]]. The dimension of $\mathcal{P}_m$ is $m+1$ and we have the following representation of $\mathrm{SL}(2,\mathbb{C})$ thereon:

$$\big(\pi_m(g)f\big)(u,v) = f\Big(g\begin{pmatrix}u\\v\end{pmatrix}\Big) = f(au+bv, cu+dv) \tag{53.335}$$

if $g = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$. We are going to determine the corresponding representation ρ_m of the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ as algebra over complex numbers.

A basis of $\mathfrak{sl}(2,\mathbb{C})$ over $\mathbb{C}$ is given by the matrices $\{H,E,F\}$ given in equation (53.315) and are subject to the commutation relations (51.152).

Using the exponentiation (53.317), we find

$$\big(\pi_m(e^{tH})f\big)(u,v) = f(e^t u, e^{-t}v),$$

so that

$$\big(\rho_m(H)f\big)(u,v) = u\frac{\partial f}{\partial u} - v\frac{\partial f}{\partial v}. \tag{53.336}$$

In the same way, we find

$$\big(\rho_m(E)f\big)(u,v) = \frac{d}{dt}\Big[\big(\pi_m(e^{tE})f\big)(u,v)\Big]_{t=0} = v\frac{\partial f}{\partial u}, \tag{53.337}$$

and

$$\big(\rho_m(F)f\big)(u,v) = u\frac{\partial f}{\partial v}. \tag{53.338}$$

A natural basis of $\mathcal{P}_m$ is given by the monomials $f_j(u,v) = u^j v^{m-j}$ with $j = 0,\dots,m$. The representation ρ_m on this basis reads

EqRepresentsldeuxC

$$\begin{aligned}\rho_m(H)f_j &= (2j-m)f_j\\ \rho_m(E)f_j &= (m-j)f_{j+1}\\ \rho_m(F)f_j &= jf_{j-1}.\end{aligned} \tag{53.339}$$

ProprhomirredsldeuxC

Proposition 53.134.
The representation ρ_m is irreducible.

Proof. Let $W \neq \{0\}$ be an invariant subspace of $\mathcal{P}_m$. If $p \in W$, from invariance, $\rho_m(H)(p) \in W$. If p is a linear combination of $\{f_j\}_{j\in I}$ ($I \subseteq \{0,\dots m\}$), then $\rho_m(H)p$ is still a linear combination q of elements in the same set. Thus there exists a linear combination of p and $\rho_m(H)p$ which is a linear combination of $\{f_j\}_{j\in J}$ with $J \subset I$ (strict inclusion). Using the same trick with q and $\rho_m(H)q$, we still reduce the number of basis elements. Proceeding in the same way at most m times, we find that one of the f_j belongs to W. From there, acting with $\rho_m(E)$ and $\rho_m(F)$, one generates the whole $\mathcal{P}_m$. That proves that $W = \mathcal{P}_m$ and thus that ρ_m is irreducible. □

Theorem 53.135 ([882])**.**
Every $\mathbb{C}$-linear irreducible finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ is equivalent to one of the ρ_m.

These results will be proved also in the quantum case in theorems 73.54 and 73.55.

53.19 The group SO(3) and its Lie algebra

SubSecTheGroupSotrois

Proposition 53.136 ([883]).
An element of SO(3) *has exactly one eigenvector with eigenvalue* 1. *That vector is the* ***rotation axis****.*

The generator of rotation around the axis n (unit vector) is given by the matrix

$$\begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}. \tag{53.340}$$

That form results form the requirement that $Nr = n \times r$. If we denote by $R(n,\theta)$ the operator of rotation in $\mathbb{R}^3$ by an angle θ around the axis n, one shows that

$$R(b,\theta) = \mathbb{1} + \sin(\theta)N + \big(1 - \cos(\theta)\big)N^2. \tag{53.341}$$

53.19.1 Rotations of functions

Consider any function $f\colon \mathbb{R}^3 \to \mathbb{C}$; we define the **rotation operator** $U(n,\theta)$ by

$$\big(U(n\theta)f\big)(r) = f\big(R(n,\theta)^{-1}r\big). \tag{53.342}$$

EqRotFunSOtrois

These operators form a group, and we have in particular that

$$U(n,\theta_1)U(n,\theta_2) = U(n,\theta_1+\theta_2).$$

We are interested in *infinitesimal* rotations, that is rotations of angle $d\theta$ for which $(d\theta)^2 \ll d\theta$, or in other words, we are interested in a development of equation (53.342) restricted to linear terms in θ. What one obtains is

$$\big(U(n,d\theta)f\big)(r) = \big((1 - id\theta\, n\cdot l)f\big)(r) \tag{53.343}$$

where the operator l is defined by

$$l = -ir\times\nabla. \tag{53.344}$$

Its components $l_i = -i\epsilon_{ijk}r_j\partial_k$ satisfy commutation relations

$$[l_i, l_j] = i\epsilon_{ijk}l_k. \tag{53.345}$$

EqAlldlepsl

The operator $n\cdot l$ is refereed as the **generator of infinitesimal rotations**. One can derive an expression of $U(n,\theta)$ in terms of $n\cdot l$ by the following:

$$U(n,\theta+d\theta)f = U(n,\theta)U(n,d\theta)f = U(n,\theta)(1 - id\theta\, n\cdot l)f,$$

so that we have the differential equation

$$\frac{dU}{d\theta}(n,\theta) = -iU(n,\theta)n\cdot l \tag{53.346}$$

with the initial condition $U(n,0) = 1$. The solution is

$$U(n,\theta) = e^{-i\theta\, n\cdot l}. \tag{53.347}$$

53.20 Verma module

When $\mathfrak{g}$ is a semisimple Lie algebra, we have the usual decomposition[884]

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+, \tag{53.348}$$

where each of the three components are Lie algebras. In particular, the universal enveloping algebra $\mathcal{U}(\mathfrak{n}^-)$ makes sense. Let $\mu \in \mathfrak{h}^*$. We build a representation π_μ of $\mathfrak{g}$ on $V_\mu = \mathcal{U}(\mathfrak{n}^-)$ in the following way

— If $Y_\alpha \in \mathfrak{n}^-$, we define

$$\pi_\mu(Y_\alpha)1 = Y_\alpha \tag{53.349a}$$
$$\pi_\mu(Y_{\alpha_1} \dots Y_{\alpha_n}) = Y_\alpha Y_{\alpha_1} \dots Y_{\alpha_n}, \tag{53.349b}$$

— if $H \in \mathfrak{h}$, we define

$$\pi_\mu(H)1 = \mu(H) \tag{53.350a}$$
$$\pi_\mu(Y_{\alpha_1} \dots Y_{\alpha_k}) = \big(\mu(H) - \sum_{j=1}^{k} \alpha_j(H)\big) Y_{\alpha_1} \dots Y_{\alpha_k}, \tag{53.350b}$$

— and if $X_\alpha \in \mathfrak{n}^+$, we define

$$\pi_\mu(X_\alpha)1 = 0 \tag{53.351a}$$
$$\pi_\mu(X_\alpha)Y_{\alpha_1} \dots Y_{\alpha_k} = Y_{\alpha_1}\big(\pi_\mu(X_\alpha)Y_{\alpha_2} \dots Y_{\alpha_k}\big) \tag{53.351b}$$
$$- \delta_{\alpha,\alpha_1} \sum_{j=1}^{k} \alpha_j(H_\alpha) Y_{\alpha_1} \dots Y_{\alpha_k}. \tag{53.351c}$$

In the last one, we do an inductive definition.

Lemma 53.137.
The couple (π_μ, V_μ) is a representation of $\mathfrak{g}$ on V_μ.

Proof. No proof. □

That representation is one **Verma module** for $\mathfrak{g}$. If the algebra $\mathfrak{g}$ is an algebra over the field $\mathbb{K}$, the field $\mathbb{K}$ itself is part of $\mathcal{U}(\mathfrak{n})^-$, so that the scalars are vectors of the representation. In that context, the multiplicative unit $1 \in \mathbb{K}$ is denoted by v_0.

Theorem 53.138.
The representation (π_μ, V_μ) of the semisimple Lie algebra $\mathfrak{g}$ is a cyclic module of highest weight, with highest weight μ and where v_0 is a vector of weight μ.

Proof. No proof. □

The Verma module is, *a priori*, infinite dimensional and non irreducible, thus one has to perform quotients of the Verma module in order to build finite dimensional irreducible representations.

53.21 Cyclic modules and representations

An example over $\mathfrak{so}(3)$ is given in subsection 51.8.2.2. The case of $\mathfrak{so}(5)$ is treated in subsection 55.5.1. Let $\mathfrak{g}$ be a semisimple Lie algebra with a Cartan subalgebra $\mathfrak{h}$ and a basis Δ for its roots $\Phi = \Phi^+ \cup \Phi^-$. Let W be a finite dimensional $\mathfrak{g}$-module.

Lemma 53.139.
If $\mathfrak{g}$ is a nilpotent complex algebra and if γ is a weight, then there exists a v in V_γ such that $c \cdot v = \gamma(x)v$ for every $x \in \mathfrak{g}$.

This is the proposition 51.153. Notice that a Cartan algebra is nilpotent, thus one has at least one vector of W which is a common eigenvector of every elements of $\mathfrak{h}$, in other words, $\exists \mu \in \mathfrak{h}^*$ and $\exists w \in W$ such that

$$hw = \mu(h)w \tag{53.352}$$

for every $h \in \mathfrak{h}$, and $w \neq 0$. If w is such and if $x \in \mathfrak{g}_\alpha$, we have

$$(hx) \cdot w = [h, x] \cdot w + (xh) \cdot w = \alpha(h)x \cdot w + x\mu(h)w = (\alpha + \mu)(h)x \cdot w. \tag{53.353}$$

If we define

$$S = \{w \in W \text{ tel que } \exists \mu \in \mathfrak{h}^* \text{ tel que } hw = \mu(h)w\}, \tag{53.354}$$

this is not a vector space, but the vector space $\operatorname{Span} S$ generated by S is invariant under $\mathfrak{g}$ because S itself is invariant under all the $\mathfrak{g}_\alpha$ with $\alpha \in \mathfrak{g}^*$.

On the other hand, we suppose that $\mathfrak{g}$ and W are finite dimensional, so that their dual are isomorphic. Since a Cartan subalgebra is chosen, we have the decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha \tag{53.355}$$

where $\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \text{ tel que } [h, x] = \alpha(h)x \, \forall g \in \mathfrak{h}\}$. When $\alpha \in \mathfrak{h}^*$, the two following spaces are independent of the choice of the Cartan subalgebra $\mathfrak{h}$:

$$\begin{aligned} W_\alpha &= \{v \in W \text{ tel que } hv = \alpha(h)v \, \forall h \in \mathfrak{h}\} \\ \mathfrak{g}_\alpha &= \{x \in \mathfrak{g} \text{ tel que } [h, x] = \alpha(h)x \, \forall h \in \mathfrak{h}\}. \end{aligned} \tag{53.356}$$

If $v_\alpha \in W_\alpha$ and $x_\beta \in \mathfrak{g}_\beta$, we have

$$h(x_\beta v_\alpha) = \big([h, x_\beta] + x_\beta h\big)v_\alpha = \big(\beta(h) + \alpha(h)\big)x_\beta v_\alpha, \tag{53.357}$$

so $x_\beta v_\alpha \in W_{\alpha+\beta}$. Thus x_β is a map

$$x_\alpha \colon W_\alpha \to W_{\alpha+\beta}. \tag{53.358}$$

Since W is finite dimensional, there exists a maximal α such that $W_\alpha \neq 0$. We name it λ. For every $\beta \in \Phi^+$, we have $W_{\lambda+\beta} = \{0\}$. In particular, if $v_\lambda \in W_\lambda$,

$$x_\alpha x_\lambda = 0 \tag{53.359}$$

for every $\alpha \in \Phi^+$, and, of course,

$$hv_\lambda = \lambda(h)v_\lambda. \tag{53.360}$$

On the other hand, for every vector $v \in W$, and for v_λ in particular, the space $\mathcal{U}(\mathfrak{g})v$ is invariant, so

$$W = \mathcal{U}(\mathfrak{g})v_\lambda \tag{53.361}$$

by irreducibility. One say that W is the **cyclic module** generated by v_λ.

53.21.1 Choice of basis

ThoBaseUGxxmono

Theorem 53.140.
Let $\mathfrak{g}$ be a Lia algebra on a field of characteristic zero. If $\{x_i\}$ is an ordered basis of $\mathfrak{g}$, then

$$\{x_{i_1} \cdots x_{i_n} \text{ tel que } i_1 \leqslant \ldots \leqslant i_n\} \tag{53.362}$$

is a basis for the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ of $\mathfrak{g}$.

One can find a proof in [867].

53.21.2 Roots and highest weight vectors

PropoIrrrgenfffIamble

Proposition 53.141.
An irreducible cyclic module is generated by the elements of the form $f_1^{i_1} \cdots f_m^{i_m} v_\lambda$.

Proof. From theorem 53.140, the monomials of the form

$$(f_1^{i_1} \cdots f_m^{i_m}) \cdot (h_1^{j_1} \cdots h_l^{j_l}) \cdot (e_1^{k_1} \cdots e_m^{k_m}) \tag{53.363}$$

form a basis of $\mathcal{U}(\mathfrak{g})$. When one act with such an element on v_λ, the e_i kill it, while the h_i do not act (a part of changing the norm). Thus, in fact, the module W is generated by the only elements $f_1^{i_1} \cdots f_m^{i_m} v_\lambda$ □

In very short, one can write

EqWnmoinvlambldarootmodul

$$W = (\mathfrak{n}^-)^n v_\lambda. \tag{53.364}$$

Since $f_k v_\alpha \in \mathfrak{g}_{\alpha-\alpha_k}$, we have

Eqfmlaphamoinsmouns

$$f_1^{i_1} \cdot f_m^{i_m} v_\lambda \in \mathfrak{g}_{\lambda-(i_m\alpha_m-\dots i_1\alpha_1)}. \tag{53.365}$$

The set of roots is ordered by

$$\mu_1 \prec \mu_2 \quad \text{iff} \quad \mu_2 - \mu_1 = \sum_i k_i \alpha_i \tag{53.366}$$

with $\alpha_i > 0$ and with $k_i \in \mathbb{N}$. Equation (53.365) means that

$$\mu \prec \lambda \tag{53.367}$$

for every weight μ of W.

Definition 53.142.
Let $\mathfrak{g}$ *be a finite dimensional Lia algebra. A* ***cyclic module of highest weight*** *for* $\mathfrak{g}$ *is a module (not specially of finite dimension) in which there exists a vector* v_+ *such that* $x_+ v_+ = 0$ *for every* $x_+ \in \mathfrak{n}^+$ *and* $hv_+ = \lambda(h)v_+$ *for every* $h \in \mathfrak{h}$.

Proposition 53.143.
Every submodule of a cyclic highest weight module is a direct sum of weight spaces.

Proof. No proof. □

From the relation $x_+ v_+ = 0$, we know that all the weight spaces satisfy V_μ satisfy $\mu \prec \lambda$, and, since a module is the sum of all its submodules,

EqVsumValpha

$$V = \bigoplus V_\mu. \tag{53.368}$$

Notice that if v_+ is in a submodule, then that submodule is the whole V, thus the sum of two proper submodules is a proper submodule. We conclude that V has an unique maximal submodule, and has thus an unique irreducible quotient.

53.21.3 Dominant weight

SubSecDomiunSei

We know that every representation is defined by a highest weight. The following proposition shows that every root cannot be a highest weight of an irreducible representation.

Proposition 53.144 ([885])**.**
The highest weight of an irreducible representation of a simple complex Lie algebra is an integral dominant weight.

Proof. Let α_i be a simple root and consider the corresponding copy of $\mathfrak{sl}(2,\mathbb{C})$ generated by $\{e_i, f_i, h_i\}$ (see proposition 51.89). The following part of $L(\Lambda)$ is a $\mathfrak{sl}(2,\mathbb{C})_i$-module:

$$V(\alpha_i) = \bigoplus_{n\in\mathbb{Z}} V_{\Lambda+n\alpha_i} = V_\Lambda \oplus V_{\Lambda-\alpha_i} \oplus V_{\Lambda-2\alpha_i} \oplus \dots \oplus V_{\Lambda-r\alpha_i} \tag{53.369}$$

for some positive integer r. Notice that the sum over $n \in \mathbb{Z}$ does not contain terms with $n < 0$ because Λ being an highest weight, $V_{\Lambda+k\alpha_i} = \varnothing$ when $k > 0$. We know that in a $\mathfrak{sl}(2,\mathbb{C})$-module the eigenvalues of h run from $-m$ to m (see equations (53.339) for example). Thus here

$$\Lambda(h_i) = -(\Lambda - r\alpha_i)(h_i). \tag{53.370}$$

By construction $\alpha_i(h_i) = 2$, so $\Lambda(h_i) = r$ and the proof is finished. □

Proposition 53.145.
If Λ is the highest weight of the representation $L(\Lambda)$ of the complex simple Lie algebra $\mathfrak{g}$ and if w_0 is the longest elements of the Weyl group, then $w_0\Lambda$ is the lowest weight.

Proof. First remember that whenever λ is a weight of a representation and w is an element of the Weyl group, the root $w\lambda$ is a weight [56]; in particular $w_0\Lambda$ is a weight of $L(\Lambda)$. Let $v \in L(\Lambda)_{w_0\Lambda}$; we want to show that $X_i^- v = 0$.

If $X_i^- v \neq 0$, then $w_0\Lambda - \alpha_i$ is a weight and $w_0\big(w_0\Lambda - \alpha_i\big) = \Lambda - w_0\alpha_i$ is a weight too. Here we used the fact that $w_0^2 = \mathrm{Id}$. □

¡¡ Avertissement/question au lecteur !! 53.146
Still to be shown:

(1) $w\lambda$ is a weight

(2) $w_0^2 = \mathrm{Id}$

53.21.4 Verma modules

Let us consider

$$\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+, \tag{53.371}$$

and take $\alpha \in \mathfrak{h}^*$. Now, we define $\mathbb{C}_\alpha$ as the vector space $\mathbb{C}$ (one dimensional, generated by $z_+ \in \mathbb{C}$) equipped with the following action of $\mathfrak{b}$:

$$\big(h + \sum_{\mu<0} x_\mu\big) z_+ = \alpha(h) z_+. \tag{53.372}$$

The vector space $\mathbb{C}_\alpha$ becomes a left $\mathcal{U}(\mathfrak{b})$-module. On the other hand, $\mathcal{U}(\mathfrak{g})$ is a free right $\mathcal{U}(\mathfrak{b})$-module because $\mathcal{U}(\mathfrak{b}) \cup \mathcal{U}(\mathfrak{g}) \subseteq \mathcal{U}(\mathfrak{g})$. As $\mathcal{U}(\mathfrak{b})$-module, a basis of $\mathcal{U}(\mathfrak{g})$ is given by $\mathfrak{n}^-$, i.e. by $\{f_1^{i_1} \cdots f_m^{i_l}\}$. The **Verma module** is the cyclic module

$$\mathrm{Verm}(\alpha) = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{b})} \mathbb{C}_\alpha \tag{53.373}$$

which has a highest weight vector $v_\lambda = 1 \otimes z_+$. The tensor product over $\mathcal{U}(\mathfrak{b})$ beans that, when $X \in \mathcal{U}(\mathfrak{g})$, then

$$\big(h + \sum_\mu x_\mu\big) X \otimes_{\mathcal{U}(\mathfrak{b})} z z_+ = X \otimes \big(h + \sum_\mu x_\mu\big) z z_+ = X \otimes_{\mathcal{U}(\mathfrak{g})} z\alpha(h) z_+ = \alpha(h) X \otimes_{\mathcal{U}(\mathfrak{b})} z z_+. \tag{53.374}$$

The Verma module is generated by $1 \otimes z_+$ and the fact that

$$zX(1 \otimes z_+) = X \otimes z z_+. \tag{53.375}$$

Proposition 53.147.
Two irreducible cyclic modules with same highest weight are isomorphic.

56. To be proved.

Proof. Let V and W be two highest weight cyclic modules with highest weight λ and highest weight vectors v_λ and w_λ. In the module $V \oplus W$, the vector $v_\lambda \oplus w_\lambda$ is a highest weight vector of weight λ. Let us consider the module

$$Z = \mathcal{U}(\mathfrak{g})(v_\lambda \oplus w_\lambda). \tag{53.376}$$

That module is a highest weight cyclic module. The projections onto $V = Z/W$ and $W = Z/V$ are non vanishing surjective homomorphisms, so V and W are irreducible quotients of Z. But we saw bellow equation (53.368) that Z can only accept one irreducible quotient. Thus V and W are isomorphic. □

We denote by $\mathrm{Irr}_{\mathfrak{g}}(\lambda)$ the unique cyclic highest weight $\mathfrak{g}$-module with highest weight λ.

Chapter 54

A lot of algebra

54.1 Algebraic structures

54.1.1 Algebraic structures

A set with operations $(R, +, \cdot)$ is a **ring** if

— $(R, +)$ is an abelian group (we denote the neutral by 0),

— the multiplication $(R, \cdot)$ is unital and associative,

— the multiplication is distributive with respect to addition.

A ring is not always a vector space because there are no multiplication by scalar.

A **field** is a commutative ring whose nonzero elements form a group under multiplication.

An **algebra** is a vector space $\mathfrak{A}$ on $\mathbb{R}$ or $\mathbb{C}$ (or any field $\mathbb{K}$) with an operation $\times : \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ such that

— $\times$ is distributive at left and right with respect to addition (vector structure),

— for all $z, z' \in \mathfrak{A}$ and A, $B \in \mathfrak{A}$,

$$zA \times z'B = (zz')(A \times B).$$

Most of time, the element $A \times B$ is just denoted by AB. In the case of Lie algebra, $X \times Y$ is $[X, Y]$. We will often treat with unital associative algebras on $\mathbb{C}$; these algebras are in fact rings.

When $\mathfrak{A}$ is an algebra, the **opposite** $\mathfrak{A}^0$ is given by $\mathfrak{A}^0 = \mathfrak{A}$ as sets, but multiplication in $\mathfrak{A}^0$ is

$$a^0 b^0 := (ba)^0$$

where the left hand side product is the one in $\mathfrak{A}^0$ while the right hand side one is the product in $\mathfrak{A}$. A **trace** aver an algebra is a linear form which satisfies the cyclic property $\tau(xy) = \tau(yx)$ for every elements x and y in the algebra.

54.1.2 Morphisms and such

When $(E, \times)$ and $(F, \cdot)$ are two set with composition law and appropriate differential structure, a map $\varphi : E \to F$ is a

monomorphism if $\ker \varphi = 0$,

epimorphism if φ is surjective,

homomorphism if $\varphi(x \times y) = \varphi(x) \cdot \varphi(y)$,

homeomorphism if φ and φ^{-1} are continuous (φ^{-1} needs to exist!).

If $(E, \times)$ and $(F, \cdot)$ are two sets with a composition law, a map $\varphi : E \to F$ is a **homomorphism** if it satisfies

$$\varphi(x \times y) = \varphi(x) \cdot \varphi(y).$$

It is a **monomorphism** when $\ker \varphi = 0$ and an **epimorphism** if $\varphi(E) = F$.

If M and N are differentiable manifolds of respective dimension m and n, a map $f\colon M \to M$ is **differentiable** if the coordinate associated map $f_{\alpha i}\colon \mathcal{U}_\alpha \to \mathcal{U}_i$ is of class C^∞ for any chart $\mathcal{U}_\alpha$ of M and $\mathcal{U}_i$ of N.

The map $f\colon M \to N$ is a **diffeomorphism** if f is bijective and if g and g^{-1} are both differentiable. An **homeomorphism** is a continuous map which has a continuous inverse.

Remarque 54.1.
Take care to not confuse *homomorphism* and *homeomorphism*. The first is only a map which respect an algebraic structure (a composition law) while the second also satisfies topological conditions: it must be continuous and open.

54.1.3 Group ring

Let G be a group and R, a ring. The set $R[G]$, called the **group ring** of G is

$$R[G] = \{\sum_{g\in G} a_g g\} \tag{54.1}$$

where $a_g \in R$ is non vanishing only for a finite number of g. The set $R[G]$ becomes a group by combining the composition law of G and R in the following way:

$$\Big(\sum_g a_g g\Big)\Big(\sum_h b_h h\Big) = \sum_{g,h}(a_g b_h)gh. \tag{54.2}$$

One can show that this definition is the only one for which the property

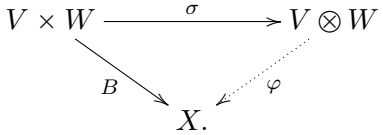
$$(1g)(1h) = (1gh)$$

holds when R has an unit 1. If R and G are commutative, then $R[G]$ is commutative.

54.2 Tensor products

54.2.1 Tensor product of vector spaces

When V and W are vector spaces, it is well known that the Cartesian product $V \times W$ is a vector space. A **tensor product** of V and W is a vector space $V \otimes W$ endowed with a bilinear map $\sigma\colon V \times W \to V \otimes W$ such that for every vector space X and bilinear map $B\colon V \times W \to X$, there exists an unique linear map $\varphi\colon V \otimes W \to X$ such that $\varphi \circ \sigma = B$, i.e. the following diagram commutes:



One can exhibit an explicit construction for the tensor product. Let $\{v_\alpha\}$ and $\{w_\beta\}$ be basis of V and W respectively and define $V \otimes W$ as the free vector space generated by the symbols $v_\alpha \otimes w_\beta$ endowed with the bilinear map $V \times W \to V \otimes W$,

$$\big(\sum_\alpha a_\alpha v_\alpha, \sum_\beta b_\beta w_\beta\big) \mapsto \sum_{\alpha\beta} a_\alpha b_\beta\, v_\alpha \otimes w_\beta.$$

One can form in that way tensor product of any *finite* number of vector space. For the infinite product case, we make a direct limit, namely if V_k are normed vector spaces, we choose unit vectors $v_k \in V_k$ and we define

$$\bigotimes_{k=1}^{\infty} V_k = \text{completion of } \left(\lim_{\rightarrow} \bigotimes_{k=1}^{N} V_k\right), \tag{54.3}$$

the maps being $w_1 \mapsto w_1 \otimes v_2$, and so on...

Lemma 54.2.
If V_i are finite dimensional vector spaces, there exists an isomorphism $\phi\colon (V_1 \otimes V2)^* \to V_1^* \otimes V_2^*$.

Proof. Indeed an element of $(V_1 \otimes V_2)^*$ reads $F = F^{i\alpha}(e_i \otimes e_\alpha)^*$ where $\{e_i\}$ is a basis of V_1 and $\{e_\alpha\}$ is a basis of V_2. Then we define

$$\phi(F) = F^{i\alpha} e_i^* \otimes e_\alpha^*.$$

□

Lemma 54.3.
When V_1 and V_2 are finite dimensional, an inner product on $V_1 \otimes V_2$ such that

$$\langle w_1 \otimes w_2, z_1 \otimes z_2 \rangle = \langle w_1, z_1 \rangle \langle w_2, z_2 \rangle$$

exists and is unique.

54.2.2 Some conversions

SUBSECooAASYooVHZEhz

Let A, B be vector spaces and a bilinear map

$$\psi\colon A \times B \to B. \tag{54.4}$$

This can also be seem as a map $\psi\colon A \otimes B \to B$. Here the tensor product is taken with respect to the base field $\mathbb{R}$. We consider a basis $\{e_k\}$ of A and $\{f_\alpha\}$ of B and we define $\psi_{k\alpha} \in B$ by $\psi_{k\alpha} = \psi(e_k, f_\alpha)$. Thus if $a = \sum_k a_k e_k$ and $b = \sum_\alpha b_\alpha f_\alpha$ we have

$$\psi(a, b) = \sum_{k\alpha} \psi_{k\alpha} a_k b_\alpha. \tag{54.5}$$

This induces a map

$$\varphi\colon B \to B \otimes A^* \tag{54.6}$$

defined by

$$\varphi(b) = \sum_{k\alpha} (\psi_{k\alpha} b_\alpha) \otimes e_k^*. \tag{54.7}$$

It can also be seen as a map $\varphi(b)\colon A \to B$ by

$$\varphi(b)a = \varphi(b)(A \otimes a) = \sum_{k\alpha} \big((\psi_{k\alpha} b_\alpha) \otimes e_k^*\big)(1 \otimes a) = \sum_{k\alpha} \psi_{k\alpha} b_\alpha a_k = \psi(a, b). \tag{54.8}$$

Here "1" is the B-valued 1-form over B that is defined by $1(b) = b$.

54.2.3 Contraction of tensors

Let (M, g) be a (pseudo)Riemannian manifold and $X \in \Gamma(TM)$, $\omega \in \Gamma(T^*M)$. We can consider the tensor product

$$X \otimes \omega \in \Gamma(TM) \otimes \Gamma(T^*M). \tag{54.9}$$

Definition 54.4.
*The **contraction** of $X \otimes \omega$ is the function*

$$C(X \otimes \omega) = \sum_{kj} (g^{-1})_{kj} \omega_j X_k \tag{54.10}$$

EQooDODKooOxCzZP

where we have decomposed $X = \sum_k X_k \partial_k$ and $\omega = \sum_l \omega_l dx_l$.

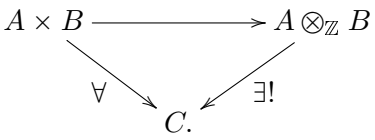
The fact to have used the inverse of the metric allows to use the Einstein's summation convention:

$$C(X \otimes \omega) = \omega^i X_i = g^{ik} \omega_k X_i \tag{54.11}$$

where $g^{ik} = (g^{-1})_{ik}$.

54.2.4 Tensor product of groups

Let A and B be two abelian groups, we define the tensor product of A and B as the abelian group $A \otimes_{\mathbb{Z}} B$ endowed with a map $A \times B \to A \otimes_{\mathbb{Z}} B$ which is an homomorphism onto A when $b \in B$ is fixed and an homomorphism onto B when $a \in A$ is fixed, and such that for every morphism $A \times B \to C$ to an abelian group C, there exists a morphism $A \otimes_{\mathbb{Z}} B \to C$ which makes the following diagram commute



Let R be a ring; A, a right R-module and B, a left R-module. We define

$$A \otimes_R B = \frac{A \otimes_{\mathbb{Z}} B}{\text{subgroup generated by elements of the form } ar \otimes_{\mathbb{Z}} b - a \otimes_{\mathbb{Z}} rb}. \tag{54.12}$$ EqdefAtensRB

54.3 Modules over algebras

Some more details can be found in [819].

Let $\mathfrak{A}$ be an algebra on $\mathbb{C}$. A vector space $\mathcal{E}$ is a **right module** on $\mathfrak{A}$ if it carry a right representation of $\mathfrak{A}$. In other terms, if we have a right multiplication

$$\begin{aligned} \mathcal{E} \times \mathfrak{A} &\to \mathcal{E} \\ (\eta, a) &\mapsto \eta a \end{aligned} \tag{54.13}$$

such that

$$\eta(ab) = (\eta a)b \tag{54.14a}$$
$$\eta(a+b) = \eta a + \eta b \tag{54.14b}$$
$$(\eta + \xi)a = \eta a + \xi a \tag{54.14c}$$

for all $\eta, \xi \in \mathcal{E}$ and a, $b \in \mathfrak{A}$. Let $\mathcal{E}$ and $\mathcal{F}$ be two right $\mathfrak{A}$-module. A **morphism** of $\mathcal{E}$ on $\mathcal{F}$ if a $\mathfrak{A}$-linear map $\rho\colon \mathcal{E} \to \mathcal{F}$ which fulfils $\rho(\eta a) = \rho(\eta)a$ for all $a \in \mathfrak{A}$ and $\eta \in \mathcal{E}$. When $\mathcal{E}$ is a right $\mathfrak{A}$-module, we have a left $\mathfrak{A}^0$-module $\mathcal{E}^0$ defined by $a^0 \eta a$.

Let $\mathfrak{A}$ and $\mathcal{B}$ be two algebras; the vector space $\mathcal{E}$ is a $\mathfrak{A}$-$\mathcal{B}$-**bimodule** if $\mathcal{E}$ is a left $\mathfrak{A}$-module, a right $\mathcal{B}$-module and if the two actions commute: $(a\xi)b = a(\xi b)$ for every $\xi \in \mathcal{E}$, $a \in \mathfrak{A}$ and $b \in cB$.

The algebra $\mathfrak{A}^e := \mathfrak{A} \otimes_{\mathbb{C}} \mathfrak{A}^0$ is the **enveloping algebra** of $\mathfrak{A}$. A $\mathfrak{A}$-bimodule can be seen as a right $\mathfrak{A}^e$-module with definition $\eta(a \otimes b^0) = b\eta a$.

Let Γ be a directed set; a family $(e_t)_{t\in\Gamma}$ is **generating** for the right module $\mathcal{E}$ if any element of $\mathcal{E}$ can be written under the form $\sum_{t\in\Gamma} e_t a_t$ with $a_t \in \mathfrak{A}$ with only a finite non zero terms in the sum. The family (e_t) is **free** if it is made of $\mathfrak{A}$-linearly independent elements. A family is a **basis** when it is both free and generating. A module is of **finite type** if it possesses a generating family of finite cardinality. Notice that the decomposition is in general not unique.

Let us now consider the set $\mathfrak{A}^N = \mathbb{C}^N \otimes_{\mathbb{C}} \mathfrak{A}$. An element of this space can be written under the form

$$\eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix}$$

with $\eta_i \in \mathfrak{A}$. A module structure on $\mathfrak{A}^N$ is given by the map $(\mathbb{C}^N \otimes_{\mathbb{C}} \mathfrak{A}) \times \mathfrak{A} \to \mathbb{C}^N \otimes_{\mathbb{C}} \mathfrak{A}$,

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} a = \begin{pmatrix} \eta_1 a \\ \vdots \\ \eta_N a \end{pmatrix}.$$

A basis of this module can be identified with the canonical basis of $\mathbb{C}^N$ because

$$\begin{pmatrix}\eta_1\\ \vdots\\ \eta_N\end{pmatrix} = \eta_1\begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix} + \cdots + \eta_N\begin{pmatrix}0\\ \vdots\\ 1\end{pmatrix}.$$

Hence the module $\mathfrak{A}^N$ is free and of finite type. The inclusion map (inverse of p in a certain sense) $\lambda\colon \mathcal{E} \to \mathbb{C}^N \otimes_{\mathbb{C}} \mathfrak{A}$ will also be sometimes used. A general element of $\mathfrak{A}^N$ reads $\sum_{j=1}^N e_j \otimes_{\mathbb{C}} a_j$ where $a_j \in \mathfrak{A}$ and $\{e_j\}$ is the canonical basis of $\mathbb{C}^N$. If we pose $f_i = e_i \otimes_{\mathbb{C}} 1$, every element of $\mathfrak{A}^N$ is $\sum_j j_j a_j$, so that $\mathfrak{A}^N$ is a finitely generated module.

PgdeflambdaMod

Since p is a module homomorphism,

$$p\big(\sum_j f_j a_j\big) = \sum_j p(f_j)a_j$$

with $f_j \in \mathfrak{A}^N$ and $a_j \in \mathfrak{A}$. We deduce that $\{p(f_j)\}$ is a basis of the finitely generated module $p\mathfrak{A}^N$.

If $\mathcal{E}$ is a $\mathfrak{A}$-bimodule, the **center** of $\mathcal{E}$ is the set

$$\mathcal{Z}(\mathcal{E}) = \{\xi \in \mathcal{E} \text{ tel que } a\xi = \xi a,\ \forall a \in \mathfrak{A},\ \xi \in \mathcal{E}\}. \tag{54.15}$$

Proposition 54.5.
If we define the maps

$$\begin{aligned} j\colon \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A} &\to \mathcal{E} \otimes_{\mathbb{C}} \mathfrak{A}\\ \xi \otimes_{\mathfrak{A}} \delta a &\mapsto \xi \otimes_{\mathbb{C}} a - \xi a \otimes_{\mathbb{C}} 1 \end{aligned} \tag{54.16}$$

and

$$\begin{aligned} m\colon \mathcal{E} \otimes_{\mathbb{C}} \mathfrak{A} &\to \mathcal{E}\\ \xi\otimes_{\mathbb{C}} &\mapsto \xi a, \end{aligned} \tag{54.17}$$

the sequence

$$0 \longrightarrow \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1 \xrightarrow{\ j\ } \mathcal{E} \otimes_{\mathbb{C}} \mathfrak{A} \xrightarrow{\ m\ } \mathcal{E} \longrightarrow 0$$

is exact

Proof. An element in the kernel of m has the form $\xi \otimes_{\mathbb{C}} ba - \xi b \otimes_{\mathbb{C}} a$ which is the image by j of $\xi \otimes_{\mathfrak{A}} (\delta b)a$. Indeed

$$\begin{aligned} j\big(\xi \otimes_{\mathfrak{A}} (\delta b)a\big) &= j\big(\xi \otimes_{\mathfrak{A}} \delta(ba) - \xi \otimes_{\mathfrak{A}} b\delta a\big)\\ &= \xi \otimes_{\mathbb{C}} ba - \xi ba \otimes_{\mathbb{C}} A - \xi b \otimes_{\mathbb{C}} a + \xi ba \otimes_{\mathbb{C}} 1\\ &= \xi_{\mathbb{C}} ba - \xi b \otimes_{\mathbb{C}} a. \end{aligned}$$

□

54.3.1 Endomorphism

Let us prove that an endomorphism $A \in \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ is identified with a matrix $A \in \mathbb{M}_{N\times N}(\mathfrak{A})$ such that $A = Ap = pA = pAp$. The action of A must satisfy

$$A\big(\sum_i p(f_i)a_i\big) = \sum_i A\big(p(f_i)\big)a_i,$$

so that A acts on the generating part $\{p(f_i)\}$ of $\mathcal{E}$. We pose

$$Ap(f_i) = \sum_k p(f_i)A_{ik} \tag{54.18}$$

for some $A_{ik} \in \mathfrak{A}$. So we have

$$Ap\big(\sum_i e_i \otimes_{\mathbb{C}} a_i\big) = \sum_i\sum_k p(f_i)A_{ik}a_i = \sum_i p(f_i)\sum_k A_{ik}a_i = p\Big(\sum_i e_i \otimes_{\mathbb{C}} \big(\sum_k A_{ik}a_i\big)\Big).$$

Using the notation $\xi = p\big(\sum_i e_i \otimes_{\mathbb{C}} \xi^i\big)$ we have

$$A\xi = p\Big(\sum_i e_i \otimes_{\mathbb{C}} (A\xi)_i\Big), \tag{54.19}$$

that we note $pA\xi$ by abuse of notation. For the same reason of notations, the equation $\xi = p\big(\sum_i e_i \otimes_{\mathbb{C}} \xi^i\big)$ is denoted by $\xi = p\xi$, thus we have

$$A = Ap = pA,$$

where A and p have different meaning.

54.3.2 Dimension of a module

subsec_DimofModule

There exists some free modules which admit several basis of different cardinality; for them, there are no way to define the notion of dimension. For a free module whose all basis have same cardinality, the latter is the **dimension** of the module.

If $\mathcal{E}$ is a module of finite type, there exists an integer N and a surjective map $\rho\colon \mathfrak{A}^N \to \mathcal{E}$. Indeed, the fact the $\mathcal{E}$ is of finite type gives a generating part $\{e_1, \dots, e_N\}$. Then the definition

$$\rho\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \sum_{i=1}^{N} e_i a_i$$

works. In general, it is not possible to extract a free part of this generating part. Let us see an example. Consider $C^\infty(S^2)$, the algebra of smooth functions on the sphere and the Lie algebra $\mathfrak{X}(S^2)$. The latter is a module of finite type on $C^\infty(S^2)$ and a generating part is given by the three following vector fields:

$$Y_i(x) = \sum_{j,k=1}^{3} \epsilon_{ijk} x_j \partial_k.$$

These vectors are tangent vectors because scalar product $Y_i(x) \cdot x$ vanishes:

$$Y_i \cdot x = \sum_l (Y_i)_l x_l = \sum_{l,j} \epsilon_{ijl} x_j x_l = 0.$$

In order to prove that this set is generating, let us suppose that Y_1 is proportional to Y_2 and Y_3. If we impose $bY_3 = Y_1$, we find, for each k:

$$\epsilon_{1jk} x_j \partial_k = b\epsilon_{3jk}\partial_k,$$

taking $k = 1$, we find $x_2 = 0$. The same with $Y_1 = aY_2$ leads to $x_3 = 0$. So the only candidate point where the set of Y_i can fail to be generating is $x_1 = \pm 1$, $x_2 = x_3 = 0$. At this point, $Y_2 = \pm\partial_3$ and $Y_3 = \pm\partial_2$ whose are linearly independent. Hence the set of Y_j is a generating part. It is not a free part (with respect to $C^\infty(S^2)$) because

$$\sum_{j=1}^{3} x_j Y_j = 0.$$

We can however not find two vector fields X_1 and X_2 which form a global basis of $\mathfrak{X}S^2$ because we would have $X_1 - X_2 \neq 0$ everywhere on S^2.

54.3.3 Tensor product of modules (first)

Let $\mathcal{E}_1, \dots, \mathcal{E}_n$, and $\mathcal{F}$, some modules on $\mathbb{C}$. We denote by $L^n(\mathcal{E}_1, \dots, \mathcal{E}_n; \mathcal{F})$ the module of n-multi-linear maps $f: \mathcal{E}_1 \times \dots \times \mathcal{E}_n \to \mathcal{F}$. Let $\mathfrak{M}$ the free module generated by n-uples $(x_1, \dots, x_n)$, $x_i \in \mathcal{E}_i$; we denote by $\mathfrak{N}$ the submodule generated by elements of the form

$$(x_1, \dots, x_i + x_i', \dots, x_n) + (x_1, \dots, x_i, \dots, x_n) - (x_1, \dots, x_i', \dots, x_n) \tag{54.20a}$$

$$(x_1, \dots, zx_i, \dots, x_n) - z(x_1, \dots, x_n) \tag{54.20b}$$

for all $x_i, x_i' \in \mathcal{E}_i$ and $z \in \mathbb{C}$.

We have a canonical injection $\mathcal{E}_1 \times \dots \times \mathcal{E}_n \to \mathfrak{M}$ and $\mathfrak{M} \to \mathfrak{M}/\mathfrak{N}$. We denote by $\varphi: \mathcal{E}_1 \times \dots \times \mathcal{E}_n \to \mathfrak{M}/\mathfrak{N}$ the composition of these two. This φ is the **tensor product** $\mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_n$, and we write

$$\xi_1 \otimes_{\mathbb{C}} \xi_2 = \varphi(\xi_1, \xi_2).$$

The tensor product has an universal property.

Proposition 54.6.
If $f: \mathcal{E}_1 \times \dots \times \mathcal{E}_n \to G$ is a n-multi-linear, there exists a map $f_*: \mathfrak{M}/\mathfrak{N} \to G$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{E}_1 \times \dots \times \mathcal{E}_n & \xrightarrow{\varphi} & \mathfrak{M}/\mathfrak{N} \\ & \searrow^{f} & \downarrow{\scriptstyle f_*} \\ & & G \end{array} \tag{54.21}$$

Proof. It is first clear that there exists a $f': \mathfrak{M} \to G$ such that $i \circ f' = f$. This map takes the value 0 on $\mathfrak{N}$ because $i^{-1}(\mathfrak{N}) = 0$. So f' can be factorised by $\mathfrak{N}$ to give f_*. □

54.3.4 Tensor product of module (second)

Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two $\mathfrak{A}$-bimodule. We consider $\mathfrak{M}$, the free $\mathfrak{A}$-bimodule generated by $\mathcal{E}_1 \times \mathcal{E}_2$, i.e. the set of (formal) linear combinations of elements of the form

$$a(\xi_1, \xi_2) \text{ and } (\xi_1, \xi_2)a$$

with $xi_i \in \mathcal{A}$ and $a \in \mathcal{A}$. We define the sub-module $\mathfrak{N}$ generated by elements of the form

$$i(\xi_1 + \eta_1, \xi_2 + \eta_2) - i(\xi_1, \xi_2) - i(\xi_1, \eta_2) - i(\eta_1, \xi_2) - i(\eta_1, \eta_2) \tag{54.22a}$$

$$ai(\xi_1, \xi_2) \tag{54.22b}$$

$$ai(\xi_1, \xi_2) - i(a\xi_1, \xi_2) \tag{54.22c}$$

$$ai(\xi_1, \xi_2) - i(\xi_1, a\xi_2) \tag{54.22d}$$

$$i(\xi_1, \xi_2)a - i(\xi_1, \xi_2 a) \tag{54.22e}$$

$$i(\xi_1, \xi_2)a - i(\xi_1 a, \xi_2). \tag{54.22f}$$

We define $\varphi: \mathcal{E}_1 \times \mathcal{E}_2 \to \mathfrak{M}/\mathfrak{N}$ as

$$\varphi = \pi \circ i$$

where i is the inclusion of $\mathcal{E}_1 \times \mathcal{E}_2$ in $\mathfrak{M}$. The definition $\xi_1 \otimes_{\mathfrak{A}} \xi_2 = \varphi(\xi_1, \xi_2)$ has an universal property.

Proposition 54.7.
Let G be a $\mathfrak{A}$-bimodule and a $\mathfrak{A}$-linear map $f: \mathcal{E}_1 \times \mathcal{E}_2 \to G$. Then there exists one and only one $\mathfrak{A}$-linear map $f_*: \mathfrak{M}/\mathfrak{N} \to G$ such that the diagram

$$\begin{array}{ccc} \mathcal{E}_1 \times \mathcal{E}_2 & \xrightarrow{\varphi} & \mathfrak{M}/\mathfrak{N} \\ & \searrow^{f} & \downarrow{\scriptstyle f_*} \\ & & G \end{array} \tag{54.23}$$

eq__diagprodtensunif

is commutative.

Proof. We first consider a map $f' : \mathfrak{M} \to G$ such that $f' \circ i = f$. From $\mathfrak{A}$-linearity of f', we conclude that

$$\begin{aligned} f'\big(ai(\xi_A, \xi_2) - i(a\xi_1, \xi_2)\big) &= a(f' \circ i)(\xi_1, \xi_2) - (f' \circ i)(a\xi_1, \xi_2) \\ &= af(\xi_2, \xi_2) - f(a\xi_2, \xi_2) \\ &= 0. \end{aligned}$$

So $f'(\mathfrak{N}) = 0$, and we can consider the quotient map $f_* = f'/\mathfrak{N}$ which gives commutativity of diagram (54.23). We still have to prove the unicity part. Let $g_* : \mathfrak{M}/\mathfrak{N} \to G$ be an other map which has the same commutativity property that f_*:

$$f = f_* \circ \varphi = g_* \circ \varphi$$

and $f_*(\mathfrak{N}) = g_*(\mathfrak{N}) = 0$. A general element in $\mathfrak{M}/\mathfrak{N}$ has the form (of linear combination of) $\varphi(\xi_1, \xi_2)$, so we following computation concludes the proof:

$$f_*\varphi(\xi_1, \xi_2) - g_*\varphi(\xi_1, \xi_2) = f(\xi_1, \xi_2) - f(\xi_2, \xi_2) = 0.$$

□

54.3.5 Tensor product of modules (third)

This definition of tensor mainly product comes from [886].

We suppose the algebra $\mathfrak{A}$ to be unital and associative, so it is a ring. Note that $\mathbb{C}$ is a ring too; so we will in the same time define $\otimes_{\mathbb{C}}$ and $\otimes_{\mathfrak{A}}$. Let $\mathcal{E}$ be a right $\mathfrak{A}$-module and $\mathcal{F}$ a left $\mathfrak{A}$-module on the ring R, in particular, recall that $\mathcal{E}$ and $\mathcal{F}$ are vector spaces on $\mathbb{C}$. For $\eta \in \mathcal{E}$, $\xi \in \mathcal{F}$, and $a \in R$, we have maps

$$\begin{aligned} \mathcal{E} \times R &\to \mathcal{E} \\ (\eta, a) &\mapsto \eta a, \end{aligned} \tag{54.24}$$

and

$$\begin{aligned} R \times \mathcal{F} &\to \mathcal{F} \\ (a, \xi) &\mapsto a\xi. \end{aligned} \tag{54.25}$$

A **balanced product** of $\mathcal{E}$ and $\mathcal{F}$ is an abelian group $(P, +)$ with a map $f : \mathcal{E} \times \mathcal{F} \to P$ such that

$$f(\eta + \eta', \xi) = f(\eta, \xi) + f(\eta', \xi) \tag{54.26a}$$
$$f(\eta, \xi + \xi') = f(\eta, \xi) + f(\eta, \xi') \tag{54.26b}$$
$$f(\eta a, \xi) = f(\eta, a\xi). \tag{54.26c}$$

This balanced product is denoted by (P, f). The function f must fulfil

$$f(0, \xi) = 0 = f(\eta, 0)$$

because $f(0, \xi) = f(0 + 0, \xi) = f(0, \xi) + f(0\xi)$. The same kind of reason leads to

$$f(-\eta, \xi) = -f(\eta, \xi)$$

where the minus sign in the right hand side is in the sense of the inverse in the additive group $(P, +)$.

When (P, f) and (P, g) are two balanced products, a **morphism** is a group homomorphism $\varphi : P \to Q$ such that

$$g(\eta, \xi) = \varphi\big(f(\eta, \xi)\big).$$

Definition 54.8.
*A **tensor product** of $\mathcal{E}$ and $\mathcal{F}$ is a balanced product $(A, \otimes_R)$ such that for each balanced product (P, f), there exists an unique morphism $(A, \otimes_R) \to (P, f)$ such that*

$$\otimes_R(\eta, \xi) = f(\eta, \xi).$$

The element $\otimes_R(\eta,\xi)$ is often denoted by $\eta\otimes_R\xi$.

Unicity of tensor product is given by the following proposition.

Proposition 54.9.
Let $(A,\otimes_1)$ and $(B,\otimes_2)$ be two tensor products of $\mathcal{E}$ and $\mathcal{F}$. There exists an unique isomorphism $\varphi\colon A\to B$ such that

$$\varphi\big(\otimes_1(\eta,\xi)\big)=\otimes_2(\eta,\xi).$$

Proof. Since $(A,\otimes_1)$ is a tensor product and $(B,\otimes_2)$ is a balanced product, there exists an unique morphism $\varphi\colon(A,\otimes_1)\to(B,\otimes_2)$ such that

$$\varphi(\eta\otimes_1\xi)=\eta\otimes_2\xi.$$

We have to prove that this morphism is an isomorphism. There also exists an unique morphism $\phi\colon(B,\otimes_2)\to(A,\otimes_1)$ such that

$$\phi(\eta\otimes_2\xi)=\eta\otimes_1\xi.$$

Then

$$(\varphi\circ\phi)(\eta\otimes_2\xi)=\varphi(\eta\otimes_1\xi)=\eta\otimes_2\xi$$

and φ is the inverse of ϕ. □

Proposition 54.10.
If $(A,\otimes)$ is a tensor product of $\mathcal{E}$ and $\mathcal{F}$, then the group A is generated by products $\eta\otimes\xi$.

Proof. Let us suppose the existence of $\omega\in A$ which cannot be written under the form $\sum_i\eta_i\otimes\xi_i$. If (P,f) is a balanced product, a morphism $\varphi\colon A\to P$ is not completely defined by the requirement $\varphi(\eta\otimes\xi)=f(\eta,\xi)$ because $\varphi(\omega)$ can take any value in A.

□

54.3.6 Explicit building of tensor product

We will now explicitly build a tensor product $\mathcal{E}\otimes_R\mathcal{F}$ that will be named *the* tensor product. First we consider the abelian free group M of basis $\mathcal{E}\times\mathcal{F}$, i.e. the set of formal sums

$$n_1(\eta_1,\xi_1)+\cdots+n_r(\eta_r,\xi_r)$$

with $n_i\in\mathbb{Z}$, $\eta_i\in\mathcal{E}$ and $\xi_i\in\mathcal{F}$. We put the usual addition and if $(\eta_i,\xi_i)\neq(\eta_j,\xi_j)$ for all $i\neq j$, this sum is zero only if $n_i=0$ for all i. In M, we consider the subgroup N generated by

$$(\eta+\eta',\xi)-(\eta,\xi)-(\eta',\xi) \tag{54.27a}$$
$$(\eta,\xi+\xi')-(\eta,\xi)-(\eta,\xi') \tag{54.27b}$$
$$(\eta a,\xi)-(\eta,a\xi) \tag{54.27c}$$

with $\eta,\eta'\in\mathcal{E}$, $\xi,\xi'\in\mathcal{F}$ and $a\in R$. Now we define

$$\mathcal{E}\otimes_R\mathcal{F}=M/N \tag{54.28}$$

and

$$\eta\otimes_R\xi=[(\eta,\xi)]. \tag{54.29}$$

Proposition 54.11.
The group $(\mathcal{E}\otimes_R\mathcal{F},\otimes_R)$ is a tensor product.

Proof. We begin by proving that $(\mathcal{E}\otimes_R\mathcal{F},\otimes_R)$ is a balanced product. From definition of the equivalence defining the class $[(\eta,\xi)]$, we have

$$\begin{aligned}\otimes_R(\eta+\eta',\xi)&=\otimes_R(\eta,\xi)-\otimes_R(\eta',\xi)\\&=[(\eta+\eta',\xi)]-[(\eta,\xi)]-[(\eta',\xi)]\\&=[0].\end{aligned}$$

and

$$\otimes_R(\eta a, \xi) = [(\eta a, \xi)] = [(\eta, a\xi)] = \otimes_R(\eta, a\xi).$$

Now let (P, f) be a balanced product. Since M is generated by elements of the form (η, ξ), the prescription

$$\begin{aligned} \varphi\colon M &\to M \\ (\eta, \xi) &\mapsto f(\eta, \xi) \end{aligned} \tag{54.30}$$

completely defines φ. Now we try to defines something on M/N from this φ. Let K be the kernel of φ. Elements of the form

$$\begin{aligned} (\eta + \eta', \xi) - (\eta, \xi) - (\eta', \xi) \\ (\eta, \xi + \xi') - (\eta, \xi) - (\eta, \xi') \\ (\eta, a, \xi) - (\eta, a\xi) \end{aligned}$$

belongs to K. Then $N \subset K$ and one has a well defined morphism

$$\begin{aligned} M/N &\to P \\ [(\eta, \xi)] &\mapsto f(\eta, \xi). \end{aligned} \tag{54.31}$$

Uniqueness comes from the fact that elements of the form $[(\eta, \xi)]$ are generating M/N.

□

54.3.7 Unitary group

SubsecUnitGroup

Let $\mathcal{E} = p\mathfrak{A}^N$ be an Hermitian finite projective module. The algebra of **endomorphism** is

$$\mathrm{End}_{\mathfrak{A}}(\mathcal{E}) = \{T\colon \mathcal{E} \to \mathcal{E} \text{ tel que } T(\eta a) = T(\eta)a\}.$$

We naturally define the involution $*\colon \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \to \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ defined by

$$\langle T^*\eta, \xi\rangle = \langle \eta, T\xi\rangle \tag{54.32}$$

If $A \in \mathbb{M}_{N\times N}(\mathfrak{A})$ is an endomorphism of $\mathcal{E}$, in particular for each $\xi \in \mathcal{E}$, the element $A\xi$ must belongs to $\mathcal{E}$, so that $pA\xi = A\xi$. That proves that $pA = A$. Since $\xi = p\xi$, we have moreover $pAp = Ap$. Finally

$$A = Ap = pA = pAp.$$

So we have an isomorphism $\mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \simeq p\mathbb{M}_{N\times N}(\mathfrak{A})p$.

Definition 54.12.
*An endomorphism u is **unitary** if $u^*u = uu^* = 1$. We denote by $U(\mathcal{E})$ the space of unitary endomorphism of $\mathcal{E}$.*

As an example, let us take M, a differentiable manifold and $C^\infty(M)$, the algebra of smooth functions over M. We want to study $U_N\big(C^\infty(M)\big) = U\big(C^\infty(M)^N\big)$. An element of that algebra is a $N \times N$ matrix whose entries are functions on M or, equivalently, a function $u\colon M \to \mathbb{M}_{N\times N}(\mathbb{C})$. The 1 in the unitary condition $u^*u = uu^* = 1$ has to be understood as the function on M with $\mathbb{1}$ as constant value. So we have the isomorphism

$$U_N\big(C^\infty(M)\big) \simeq C^\infty\big(M, U(N)\big)$$

where $U(N)$ is the usual unitary group.

54.4 Module over unital ring

SecModUnitalAnneau

¡¡ Avertissement/question au lecteur !! 54.13
This section is a mess. Several propositions have to be merged. We have to check that everything said in the algebra case hold in the ring case, and then change algebra to ring everywhere. The module over algebra part mainly comes from [819].

54.4.1 Projective module of finite type on unital algebra

Let $\mathfrak{A}$ be an unital algebra.

PropEquivProjModule

Proposition 54.14.
Let $\mathcal{E}$ be a right $\mathfrak{A}$-module where $\mathfrak{A}$ is an unital algebra. The three following are equivalent:

(1) If $\rho\colon \mathcal{F}_1 \to \mathcal{F}_2$ is surjective homomorphism of right $\mathfrak{A}$-module, then every homomorphism $\lambda\colon \mathcal{E} \to \mathcal{F}_2$ can be lifted into a homomorphism $\tilde{\lambda}\colon \mathcal{E} \to \mathcal{F}_1$ which satisfies $\rho \circ \tilde{\lambda} = \lambda$.

eq_diag_proj_mod

$$\begin{array}{ccc} & & \mathcal{F}_1 \\ & \overset{\tilde{\lambda}}{\nearrow} & \big\downarrow \rho \\ \mathcal{E} & \xrightarrow[\lambda]{} & \mathcal{F}_2 \end{array} \tag{54.33}$$

*That property is the **lifting property**.*

ItemTroisCarecterisationProjectif

(2) If $\rho\colon \mathcal{F}_1 \to \mathcal{E}$ is a surjective morphism of module, there exists a module morphism $s\colon \mathcal{E} \to \mathcal{F}_1$ such that $\rho \circ s = \mathrm{Id}\,|_{\mathcal{E}}$.

$$\mathcal{F}_1 \underset{s}{\overset{\rho}{\rightleftarrows}} \mathcal{E}$$

prop_def_proj_module_iii

(3) There exists a free module $\mathcal{F}$ which decomposes as a direct sum which one component is $\mathcal{E}$:

$$\mathcal{F} = \mathcal{E} \oplus \mathcal{E}'.$$

In this case, $\mathcal{E}'$ is also free.

prop_def_proj_module

Proof. For a proof, see [819] at page 59. □

A right module which fulfils these properties is said to be **projective**.

54.4.2 Projective module of finite type on unital ring

Let R be an unital ring, and $\mathcal{E}$, $\mathcal{F}$, ... be left modules over R. One says that $\mathcal{E}$ is a **projective module** if for every surjective module homomorphism $\lambda\colon \mathcal{F}_2 \to \mathcal{F}_1$, there exists a lifting map $\tilde{\lambda}\colon \mathcal{E} \to \mathcal{F}_2$ such that the following diagram commutes:

EqLiftPropProjModules

$$\begin{array}{ccc} & & \mathcal{F}_2 \\ & \overset{\tilde{\lambda}}{\nearrow} & \big\downarrow \lambda \\ \mathcal{E} & \longrightarrow & \mathcal{F}_1 \end{array} \tag{54.34}$$

where the double arrow denotes the surjectivity.

PropFGPRkP

Proposition 54.15.
Every finitely generated projective module over the unital ring R has the form

$$R^k P$$

for some $k \in \mathbb{N}$ and some idempotent matrix $P \in \mathbb{M}_k(R)$. The choices of k and P are not unique.

Proof. We suppose that $\mathcal{E}$ is projective and finitely generated by $\{\xi_1, \dots, \xi_k\}$. A module map $T\colon R^k \to R^k$ is given by a matrix r defined by

$$T(0, \dots, 1, \dots, 0) = (r_{1j}, \dots, r_{kj})$$

where, in the left hand side, the 1 is on the jth position, and T is the right matrix multiplication

$$T(a_1,\dots,a_k)=\begin{pmatrix}a_1 & \dots & a_k\end{pmatrix}\begin{pmatrix}r_{11} & r_{21} & \dots & r_{k1}\\ r_{12} & r_{22} & \dots & r_{k2}\\ \vdots & \vdots & & \vdots\\ r_{k1} & r_{k2} & \dots & r_{kk}\end{pmatrix}. \tag{54.35}$$

Item (2) of proposition 54.14 assures the existence of a direct sum decomposition $R^k=\mathcal{E}\oplus\mathcal{F}$ of submodules of R^k, then, associated with the projection onto $\mathcal{E}$, there is an idempotent matrix $P\in\mathbb{M}_k(R)$ such that

$$\mathcal{E}\simeq R^kP. \tag{54.36}$$

EqMRkPdec

□

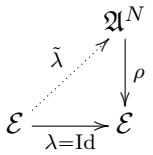
The following proposition shows that every finitely generated projective module over R come from that construction.

Proposition 54.16.
A module $\mathcal{E}$ is projective of finite type on $\mathfrak{A}$ if and only if there exists a matrix $p\in\mathbb{M}_{N\times N}(\mathfrak{A})$ such that

(1) p is idempotent: $p^2=p$,

(2) $\mathcal{E}=p\mathfrak{A}^N$.

Proof. Since $\mathcal{E}$ is projective as well as of finite type, we have a surjective map $\rho\colon\mathfrak{A}^N\to\mathcal{E}$. Let us draw the diagram (54.33) in the particular case with $\mathcal{F}_1=\mathfrak{A}^N$ and $\mathcal{F}_2=\mathcal{E}$:



(54.37)

Hence we have a map $\tilde{\lambda}\colon\mathcal{E}\to\mathfrak{A}^N$ such that $\rho\circ\tilde{\lambda}=\mathrm{Id}\,|_{\mathcal{E}}$. The map $p=\tilde{\lambda}\circ\rho\colon\mathfrak{A}^N\to\mathfrak{A}^N$ fulfils $p^2=\mathrm{Id}\,|_{\mathcal{E}}$ and is therefore idempotent in $\mathrm{End}(\mathfrak{A}^N)$. It allows us to decompose $\mathfrak{A}^N$ as a sum of submodules

$$\mathfrak{A}^N=p\mathfrak{A}^N+(1-p)\mathfrak{A}^N.$$

We know that ρ and $\tilde{\lambda}$ are two isomorphism between $\mathcal{E}$ and $p\mathfrak{A}^N$. We now prove that ρ is bijective on $p\mathfrak{A}^N$. Surjectivity is clear; we have to prove that $\rho\colon p\mathfrak{A}^N\to\mathcal{E}$ is injective. Let $\rho(pa)=\rho(pb)$ and apply $\tilde{\lambda}$ on both sides of this equality. Since $\tilde{\lambda}\rho=p$, we find $p^2a=p^2b$, en therefore $pa=pb$. Remark that this does not imply $a=b$.

For the inverse sense, we suppose a module $\mathcal{E}$ on $\mathfrak{A}$ and a map $p\in M_{N\times N}(\mathfrak{A})$ such that $p^2=p$ and $\mathcal{E}=p\mathfrak{A}^N$. The module $\mathcal{E}$ fulfils point (3) of proposition 54.14 with

$$\mathfrak{A}^N=p\mathfrak{A}^N\oplus(1-p)\mathfrak{A}^N.$$

The fact that $\mathcal{E}$ is of finite type is clear because $\mathfrak{A}^N$ is such.

□

An element of $\mathcal{E}$ can be seen as a p-invariant column vector with entry in $\mathfrak{A}$:

$$\xi=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_N\end{pmatrix}$$

with $\xi_i\in\mathfrak{A}$ and $p\xi=\xi$. Remark that it does *not* mean that $p\xi_i=\xi_i$ for each i: such a condition is in fact senseless.

A beautiful result that provides a concrete realisation of projective finite module is the following.

Theorem 54.17 (Serre-Swan)**.**
Let M be a compact finite dimensional manifold. A $C^\infty(M)$-module $\mathcal{E}$ is isomorphic to the space of sections $\Gamma(M,E)$ of a vector bundle $E \to M$ if and only if it is finite and projective.

In other words, to any projective finite module, we can associate a vector bundle whose sections give $\mathcal{E}$.

Proof. No proof □

LemRklQkR

Lemma 54.18.
If $R^k P \simeq \mathcal{E} \simeq R^l Q$, then there are matrices $U \in \mathbb{M}_{k\times l}(R)$ and $V \in \mathbb{M}_{l\times k}(R)$ such that $UV = P$ and $VU = Q$.

Proof. No proof. □

This lemma says that if we perform two times the constructions which lead to decomposition (54.36) with different choices, then the two results are in a certain sense equivalent.

Let $E \subseteq F$ be a R-submodule. The **closure** of E in F, denoted by $\mathrm{Cl}_F(E)$, or $\bar{E}$ when there are no confusion, is

$$\mathrm{Cl}_F(E) = \{f \in F \text{ tel que } \varphi(f) = 0, \forall \text{ module map } \varphi\colon F \to R \text{ such that } \varphi|_E = 0\}. \tag{54.38}$$

We are not interested in the topology on F that this definition provides.

The lemmas 54.19 and 54.20 will help to reduce numerous problems to singly generated modules.

LemEprojEEEfproj

Lemma 54.19.
If E is a finite projective module over R, then $E\oplus\ldots\oplus E$ is a finite projective module over $\mathbb{M}_n(R)$.

If F is a submodule of E, then

$$\mathrm{Cl}_{E\oplus\ldots\oplus E}(F\oplus\ldots\oplus F) = \mathrm{Cl}_E(F)\oplus\ldots\oplus\mathrm{Cl}_E(F).$$

Proof. No proof. □

LemFGenEEEsingleGen

Lemma 54.20.
If E is generated by $\{e_1,\ldots,e_n\}$, then $E\oplus\ldots\oplus E$ is generated over $\mathbb{M}_n(R)$ by the single element $(e_1,\ldots,e_n)$.

Proof. No proof. □

Now we suppose that we have a trace map $\varphi\colon R \to \mathbb{A}$ where $\mathbb{A}$ is a commutative ring (or any abelian group). If S is a square matrix of elements in R, we define

$$\varphi(S) = \sum_i \varphi(S_{ii}), \tag{54.39}$$

and one checks that $\varphi(ST) = \varphi(TS)$, so that φ is a trace over $\mathbb{M}_k(R)$ for every k. Now, we define

$$\varphi(\mathcal{E}) = \varphi(P) \tag{54.40}$$

EqPreDefDimModuleRA

where P is defined by the condition $\mathcal{E} = R^k P$. Lemma 54.18 assures that this condition does not depend on the choice.

DefHDhfMGJ

Definition 54.21.
*The **dimension function** associated with the trace φ is*

$$\begin{aligned}\dim_\varphi\colon \{\textit{isomorphism class of finitely generated projective modules}\} &\to \mathbb{A}\\ \mathcal{E} &\mapsto \varphi(\mathcal{E}).\end{aligned} \tag{54.41}$$

DefDimFunctModule

Lemma 54.22.
It fulfills the condition

$$\dim_\varphi(\mathcal{E}\oplus\mathcal{F}) = \dim_\varphi(\mathcal{E}) + \dim_\varphi(\mathcal{F}). \tag{54.42}$$

We will show in section 74.12 that the construction of the dimension function (54.41) extends to all modules over von Neumann algebras.

Lemma 54.23.
If H_1 and H_2 are modules over the ring R, they are isomorphic if and only if $H_1 \oplus \ldots \oplus H_1 \simeq H_2 \oplus \ldots \oplus H_2$ as module over $\mathbb{M}_n(R)$.

54.5 Bialgebras and co-properties

54.5.1 Bialgebras

Some generalities about bialgebras are taken from [887].

Let $\mathfrak{A}$ be an unital algebra and $r \in \mathfrak{A}\times\mathfrak{A}$. There exists a (non unique) decomposition $r = \sum_i \alpha_i\otimes\beta_i$ with α_i, $\beta_i \in \mathfrak{A}$. There exists three ways to extend that in order to embed r in $\mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A}$. We introduce the **Sweedler notation** in order to distinguish them:

$$\begin{aligned} r_{12} &= \sum_i \alpha_i\otimes\beta_i\otimes 1\\ r_{13} &= \sum_i \alpha_i\otimes 1\otimes\beta_i\\ r_{23} &= \sum_i 1\otimes\alpha_i\otimes\beta_i. \end{aligned}$$

One can define r_{ij} in the same way in $\mathfrak{A}^{\otimes n}$. We denote by $\tau\colon \mathfrak{A}\otimes\mathfrak{A}\to\mathfrak{A}\otimes\mathfrak{A}$,

$$\tau(a\otimes b) = b\otimes a.$$

If we denote by $\mu\colon \mathfrak{A}\otimes\mathfrak{A}\to\mathfrak{A}$ the algebra law in $\mathfrak{A}$, the unit 1 defines (and is defined by) the map

$$\begin{aligned} \eta\colon \mathbb{K} &\to \mathfrak{A}\\ \eta(z) &= z1. \end{aligned} \tag{54.43}$$

The fact that 1 is an unit is expressed by the equalities $1a = a1 = a$ for all $a\in\mathfrak{A}$, or equivalently by the commutativity of the diagrams

$$\begin{array}{ccc} \mathbb{K}\otimes\mathfrak{A} & \xrightarrow{\eta\otimes\mathrm{Id}} & \mathfrak{A}\otimes\mathfrak{A}\\ & \searrow^{\psi} & \downarrow{\scriptstyle\mu}\\ & & \mathfrak{A} \end{array} \qquad \begin{array}{ccc} \mathfrak{A}\otimes\mathbb{K} & \xrightarrow{\mathrm{Id}\otimes\eta} & \mathfrak{A}\otimes\mathfrak{A}\\ & \searrow^{\psi} & \downarrow{\scriptstyle\mu}\\ & & \mathfrak{A} \end{array} \tag{54.44}$$

EqDiagUnitHopf

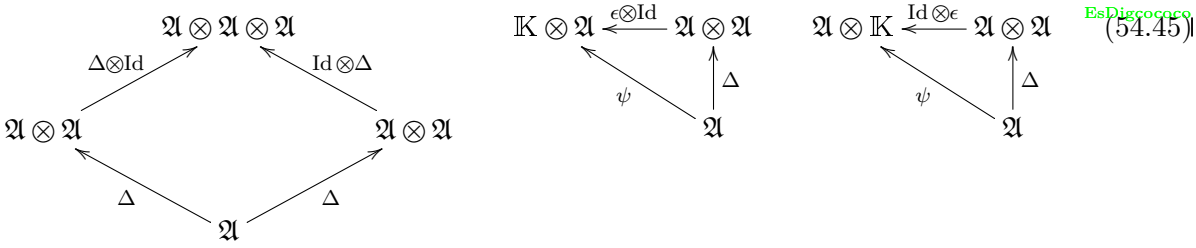
where we indifferently denoted by ψ the identification $\psi(z\otimes a) = za$ and $\psi(a) = 1\otimes a$.

So an algebra can be defined as a vector space endowed with operations μ and η such that the latter two diagrams commute. One define a **coalgebra** taking the dual of that construction. So a coalgebra is a vector space $\mathfrak{A}$ endowed with two linear maps

$$\begin{aligned} \epsilon&\colon \mathfrak{A}\to\mathbb{K},\\ \Delta&\colon \mathfrak{A}\to\mathfrak{A}\otimes_{\mathbb{K}}\mathfrak{A} \end{aligned}$$

such that the following diagrams commute

EsDizcococo



(54.45)

$$(\mathrm{Id}\otimes\Delta)\Delta=(\Delta\otimes\mathrm{Id})\Delta \qquad (\mathrm{Id}\otimes\epsilon)\Delta=\mathrm{Id} \qquad (\epsilon\otimes\mathrm{Id})\Delta=\mathrm{Id}\,. \tag{54.46}$$

The first diagram expresses the **coassociativity** of Δ while the two last ones are the **counit** definition for ϵ. When these three diagrams commute, we say that $(\mathfrak{A},\Delta,\epsilon)$ is a coalgebra.

Notice that, since the tensor product is defined by a quotient, we have $1\otimes za=z\otimes a$ for every $z\in\mathbb{C}$ and $a\in\mathfrak{A}$. We also have $(\epsilon\otimes\mathrm{Id})(a\otimes b)=\epsilon(a)\otimes b=1\otimes\epsilon(a)b$.

DefBialgebra

Definition 54.24.
*A **bialgebra** structure on $\mathfrak{A}$ is the data of $(\mu,\eta,\Delta,\epsilon)$ such as before (the six diagrams commute) and which fulfills the following compatibility conditions:*

EqSixBialgebraformdef

$$\Delta(hg)=\Delta(h)\Delta(g) \qquad \epsilon(hg)=\epsilon(h)\epsilon(g) \tag{54.47}$$

$$\Delta(1)=1\otimes 1 \qquad \epsilon(1)=1 \tag{54.48}$$

*where in the last equality, the 1 of the left hand side denotes the unit in $\mathfrak{A}$ while the 1 in the right hand side is the unit in $\mathbb{K}$. Such an operation Δ is said to be a **coproduct** and such an operation ϵ is a **counit**.*

It is interesting to write the formula $\Delta(ab)=\Delta(a)\Delta(b)$ in a more abstract way. The product

$$(a\otimes b)(c\otimes d)=ac\otimes bd \tag{54.49}$$

reads $(\mu\otimes\mu)(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id})(a\otimes b\otimes c\otimes d)$. Thus we consider

EqDefmdeuxAAotAAmtAA

$$\begin{aligned} m_2\colon (A\otimes A)\otimes(A\otimes A)&\to A\otimes A\\ m_2&=(\mu\otimes\mu)\circ(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id}) \end{aligned} \tag{54.50}$$

and we have the property

EqDelMumdex

$$\Delta\circ\mu=m_2(\Delta\otimes\Delta). \tag{54.51}$$

Among other interesting formulas, the identification $A\otimes\mathbb{K}=A$ allows us to write

EqInterAmongOtheaaensb

$$a=(\mathrm{Id}\otimes\epsilon)\Delta(a)=\sum_i(\mathrm{Id}\otimes\epsilon)(a_i\otimes b_i)=\sum_i a_i\epsilon(b_i). \tag{54.52}$$

LemUnicityCounit

Lemma 54.25.
A coalgebra has at most one counit.

Proof. Let ϵ_1 and ϵ_2 be counits of (A,Δ). We have

$$\epsilon_1=\epsilon_1\circ(\mathrm{Id}\otimes\epsilon_2)\Delta=(\epsilon_1\otimes\epsilon_2)\Delta=\epsilon_2(\epsilon_1\otimes\mathrm{Id})\Delta=\epsilon_2. \tag{54.53}$$

□

54.5.2 Other co-properties

subSecOtherCoPropoerties

We usually denote by σ the flip operation:

$$\begin{aligned}\sigma\colon \mathfrak{A}\otimes\mathfrak{A} &\to \mathfrak{A}\otimes\mathfrak{A}\\ a\otimes b &\mapsto b\otimes a.\end{aligned} \tag{54.54}$$

As a general rule, one defines a "coproperty" by writing a property as a commutative diagram and then reversing all the arrows. A map $\varphi\colon \mathfrak{A}\otimes\mathfrak{A}\to\mathfrak{A}$ is skew symmetric if the following diagram commutes:

$$\begin{array}{ccc}\mathfrak{A}\otimes\mathfrak{A} & \xrightarrow{\varphi} & \mathfrak{A}\\ {\scriptstyle\sigma}\downarrow & & \uparrow{\scriptstyle -1}\\ \mathfrak{A}\otimes\mathfrak{A} & \xrightarrow[\varphi]{} & \mathfrak{A}\end{array} \tag{54.55}$$

Then one say that a map $\phi\colon \mathfrak{A}\to\mathfrak{A}\otimes\mathfrak{A}$ is **co-skew symmetric** if the diagram

$$\begin{array}{ccc}\mathfrak{A}\otimes\mathfrak{A} & \xleftarrow{\phi} & \mathfrak{A}\\ {\scriptstyle\sigma}\uparrow & & \downarrow{\scriptstyle -1}\\ \mathfrak{A}\otimes\mathfrak{A} & \xleftarrow[\phi]{} & \mathfrak{A}\end{array} \tag{54.56}$$

commutes. In formula, it means that for every $a\in\mathfrak{A}$, we have $\phi(a)=\sigma\phi(-a)$ or

$$(\sigma+\mathrm{Id})\circ\phi=0. \tag{54.57}$$

EqDefCoSkewSym

A typical example of co-skew symmetric map is $\phi(a)=1\otimes a-a\otimes 1$.

Let ξ be the cyclic permutation operator

$$\begin{aligned}\xi\colon \mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A} &\to \mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A}\\ a\otimes b\otimes c &\mapsto b\otimes c\otimes a.\end{aligned} \tag{54.58}$$

A map $\phi\colon \mathfrak{A}\otimes\mathfrak{A}\to\mathfrak{A}$ satisfies the **Jacobi** relation

$$\varphi\big(a,\varphi(b,c)\big)=-\varphi\big(b,\varphi(c,a)\big)-\varphi\big(c,\varphi(a,b)\big) \tag{54.59}$$

if the diagram

$$\begin{array}{ccccc}\mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A} & \xrightarrow{\mathrm{Id}\otimes\varphi} & \mathfrak{A}\otimes\mathfrak{A} & \xrightarrow{\varphi} & \mathfrak{A}\\ {\scriptstyle -\xi-\xi^2}\downarrow & & & & \uparrow{\scriptstyle\varphi}\\ \mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A} & & \xrightarrow[\mathrm{Id}\otimes\varphi]{} & & \mathfrak{A}\otimes\mathfrak{A}\end{array} \tag{54.60}$$

commutes. Thus one says that a map $\phi\colon \mathfrak{A}\to\mathfrak{A}\otimes\mathfrak{A}$ satisfies the **co-Jacobi** relation if the diagram

$$\begin{array}{ccccc}\mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A} & \xleftarrow{\mathrm{Id}\otimes\phi} & \mathfrak{A}\otimes\mathfrak{A} & \xleftarrow{\phi} & \mathfrak{A}\\ {\scriptstyle -\xi-\xi^2}\uparrow & & & & \downarrow{\scriptstyle\phi}\\ \mathfrak{A}\otimes\mathfrak{A}\otimes\mathfrak{A} & & \xleftarrow[\mathrm{Id}\otimes\phi]{} & & \mathfrak{A}\otimes\mathfrak{A}\end{array} \tag{54.61}$$

commutes. In formula,

$$(\mathrm{Id}+\xi+\xi^2)\circ(\mathrm{Id}\otimes\phi)\circ\phi=0. \tag{54.62}$$

EqDefCoJacobi

54.6 Hopf algebras

54.6.1 Convolution product

We define the **convolution product** between two linear maps $f, g\colon \mathfrak{A} \to \mathfrak{A}$ by

$$f \star g = \mu(f \otimes g)\Delta \tag{54.63}$$

Lemma 54.26.
The map $\eta \circ \epsilon$ is a neutral for the convolution product.

Proof. First, notice that in general, $(\eta \circ \epsilon)(a) = \epsilon(a)1$. Let now consider $a \in \mathfrak{A}$ and a_1, $a_2 \in \mathfrak{A}$ such that $\Delta(a) = a_1 \otimes a_2$. We look at

$$\mu(f \otimes g)\Delta(a) = f(a_1)g(a_2) \tag{54.64}$$

which, using linearity of g, becomes $\mu(\eta \circ \epsilon \otimes g)\Delta(a) = \epsilon(a_1)1g(a_2) = \epsilon(a_1)g(a_2) = g\big(\epsilon(a_1)a_2\big)$. Now if we see $\epsilon(a_1)a_2 \in \mathbb{K} \otimes \mathfrak{A}$, using the commutativity of the second diagram in (54.45), we deduce $\Delta\big(\epsilon(a_1)a_2\big) = a_1 \otimes a_2 = \Delta(a)$, so that $\epsilon(a_1)a_2 = a$ by injectivity of Δ.

In the same way, we prove that $a_1\epsilon(a_2) = a$ and that, consequently,

$$\mu(f \otimes \eta \circ \epsilon)\Delta(a) = f\big(a_1\epsilon(a_2)\big) = f(a).$$

□

A bialgebra is a Hopf algebra if the identity map $\mathrm{Id}\colon \mathfrak{A} \to \mathfrak{A}$ is invertible for the convolution product. In that case the inverse is denoted by S and is called **antipode**. The defining property of S is

$$S \star \mathrm{Id} = \mathrm{Id} \star S = \eta \circ \epsilon.$$

That can equivalently be expressed by the commutativity of the diagram

$$\begin{array}{ccccc} A \otimes A & \xleftarrow{\Delta} & A & \xrightarrow{\Delta} & A \otimes A \\ {\scriptstyle \mathrm{Id}\otimes S}\downarrow & & {\scriptstyle \eta\circ\epsilon}\downarrow & & \downarrow{\scriptstyle S\otimes \mathrm{Id}} \\ A \otimes A & \xrightarrow[\mu]{} & A & \xleftarrow[\mu]{} & A \otimes A. \end{array} \tag{54.65}$$

DefHopfAlgebra

Definition 54.27.
*A tuple $(A, \mu, \eta, \Delta, \epsilon, S)$ is a **Hopf algebra** if the maps*

$$\begin{array}{lll} \mu\colon A \otimes A \to A & \Delta\colon A \to A \otimes A & S\colon A \to A \\ \eta\colon \mathbb{K} \to A & \epsilon\colon A \to A & \end{array} \tag{54.66}$$

satisfy

(1) (A, μ, η) is an associative algebra:

$$\mu(\mu \otimes \mathrm{Id}) = \mu(\mathrm{Id} \otimes \mu) \tag{54.67a}$$
$$\mu(\eta \otimes \mathrm{Id}) = \mu(\mathrm{Id} \otimes \eta), \tag{54.67b}$$

(2) (A, Δ, ϵ) is a coassociative coalgebra:

$$(\mathrm{Id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{Id})\Delta \tag{54.68a}$$
$$(\mathrm{Id} \otimes \epsilon)\Delta = (\epsilon \otimes \mathrm{Id})\Delta = \mathrm{Id}, \tag{54.68b}$$

EqformCounitDef

(3) the compatibility relations (54.47) which turn $(A, \mu, \eta, \Delta, \epsilon)$ into a bialgebra:

$$\Delta\mu = \mu(\Delta \otimes \Delta) \qquad \epsilon\mu = \mu(\epsilon \otimes \epsilon) \tag{54.69a}$$
$$\Delta(1) = 1 \otimes 1 \qquad \epsilon(1) = 1, \tag{54.69b}$$

(4) S is an antipode:

$$\mu(\mathrm{Id}\otimes S)\Delta = \mu(S\otimes \mathrm{Id})\Delta = \eta\epsilon. \tag{54.70}$$

EqDefPorpAntipode

In all these conditions, we identify $A = A\otimes\mathbb{K} = \mathbb{K}\otimes A$.

LemSuuSetaeta

Lemma 54.28.
We have

ItemLemSuuSetaetai

(1) $S(1) = 1$ and $S\circ\eta = \eta$;

ItemLemSuuSetaetaii

(2) $S\mu = \mu(S\otimes S)\sigma$ where μ is the multiplication on A and σ is the flip operator on $A\otimes A$.

Proof. On the one hand we have $1 = (\eta\circ\epsilon)(1)$ and on the other hand, by the property (54.70) of the antipode,

$$(\eta\circ\epsilon)(1) = \mu(S\otimes\mathrm{Id})\Delta(1) = S(1) \tag{54.71}$$

because $\Delta(1) = 1\otimes 1$. Thus $S(1) = 1$. Now if $z\in\mathbb{K}$, we have $\eta(z) = z\cdot 1$ and $S(z\cdot 1) = zS(1) = z\cdot 1$. This proves (1).

The statement (2) is a reformulation of the identity $S(ab) = S(b)S(a)$. □

Proposition 54.29.
The antipode satisfies

$$\Delta\circ S = \sigma(S\otimes S)\Delta = (S\otimes S)\sigma\Delta. \tag{54.72}$$

Proof. We will prove that ΔS and $\sigma(S\otimes S)\Delta$ are both inverses of Δ in a well chosen algebra. We consider $F = L(A, A\otimes A)$ with the product

$$f\star g = m_2(f\otimes g)\Delta \tag{54.73}$$

where

$$m_2 = (\mu\otimes\mu)(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id}). \tag{54.74}$$

The product m_2 is the one introduced in equation (54.50). Note that $f\star g\colon A\to A\otimes A$. The unit of that algebra is $\eta_2\epsilon$ where $\eta_2 = \eta\otimes\eta$. That function $A\to A\otimes A$ has to be understood as identifying $\mathbb{K} = \mathbb{K}\otimes\mathbb{K}$ and then

$$\eta_2\epsilon(a) = (\eta\otimes\eta)\epsilon(a)(1\otimes 1) = \epsilon(a)1\otimes 1. \tag{54.75}$$

In the latter formula the "1" is the unit in A. Let us show that this is the unit of the algebra F. As far as the notations are concerned, we write $\Delta(a) = \sum_i a_i\otimes b_i$ and $f(b_i) = \sum_k f_k(b_i)\otimes g_k(b_i)$, so

$$\begin{aligned}
(\eta_2\epsilon\star f)(a) &= m_2(\eta_2\epsilon\otimes f)\Delta(a)\\
&= \sum_i (\mu\otimes\mu)(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id})(\eta_2\epsilon\otimes f)(a_i\otimes b_i)\\
&= \sum_i (\mu\otimes\mu)(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id})\big(\epsilon(a_i)1\otimes 1\otimes f(b_i)\big)\\
&= \sum_{ik} \epsilon(a_i)(\mu\otimes\mu)(\mathrm{Id}\otimes\sigma\otimes\mathrm{Id})\big(1\otimes 1\otimes f_k(b_i)\otimes g_k(b_i)\big)\\
&= \sum_{ik} \epsilon(a_i)(\mu\otimes\mu)\big(1\otimes f_k(b_i)\otimes 1\otimes g_k(b_i)\big)\\
&= \sum_{ik} \epsilon(a_i)f_k(b_i)\otimes g_k(b_i)\\
&= \sum_i \epsilon(a_i)f(b_i)\\
&= \sum_i f\big(\epsilon(a_i)b_i\big)\\
&= f(a).
\end{aligned} \tag{54.76}$$

Thus $\eta_2\epsilon\star f = f$.

Now in order to check that $\Delta \circ S$ is an inverse of Δ, we have to check that

$$m_2(\Delta \otimes \Delta S)\Delta = \eta_2 \epsilon. \tag{54.77}$$

Notice that η_2 satisfies $\Delta\eta\epsilon = \eta_2\epsilon$. Indeed $\Delta(1) = 1 \otimes 1$ and $\eta\epsilon(a) = \epsilon(a)1$, so

$$\Delta\eta\epsilon(a) = \epsilon(a)1 \otimes \epsilon(a)1 = (\eta \otimes \eta)(a). \tag{54.78}$$

Using the formula (54.51) we have

$$\begin{aligned} m_2(\Delta \otimes \Delta S)\Delta &= m_2(\Delta \otimes \Delta)(\mathrm{Id} \otimes S)\Delta \\ &= \Delta\mu(\mathrm{Id} \otimes S)\Delta \\ &= \Delta\eta\epsilon \\ &= \eta_2\epsilon. \end{aligned} \tag{54.79}$$

Thus $\Delta \circ S$ is an inverse of Δ in the algebra F. Let us now check that $\sigma(S \otimes S)\Delta$ is also an inverse of Δ. We introduce the short hand notation $\Delta^3 = (\Delta \otimes \Delta)\Delta$. Let us show that this operator satisfies

$$(\Delta \otimes \Delta)\Delta = (\mathrm{Id} \otimes \Delta \otimes \mathrm{Id})(\mathrm{Id} \otimes \Delta)\Delta. \tag{54.80}$$ EqDDDrelDifete

Using the relation $(\mathrm{Id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{Id})\Delta$ we have

$$\begin{aligned} (\Delta \otimes \Delta)\Delta &= (\Delta \otimes \mathrm{Id} \otimes \mathrm{Id})(\mathrm{Id} \otimes \Delta)\Delta \\ &= (\Delta \otimes \mathrm{Id} \otimes \mathrm{Id})(\Delta \otimes \mathrm{Id})\Delta \\ &= \big((\Delta \otimes \mathrm{Id})\Delta \otimes \mathrm{Id}\big)\Delta \\ &= \big((\mathrm{Id} \otimes \Delta)\Delta \otimes \mathrm{Id}\big)\Delta \\ &= (\mathrm{Id} \otimes \Delta \otimes \mathrm{Id})(\Delta \otimes \mathrm{Id})\Delta \\ &= (\mathrm{Id} \otimes \Delta \otimes \mathrm{Id})(\mathrm{Id} \otimes \Delta)\Delta. \end{aligned} \tag{54.81}$$

We have

$$\begin{aligned} \diamondsuit &= m_2(\Delta \otimes \sigma(S \otimes S)\Delta)\Delta \\ &= m_2\Big[(\mathrm{Id} \otimes \mathrm{Id}) \otimes \big(\sigma(S \otimes S)\big)\Big](\Delta \otimes \Delta)\Delta \\ &= m_2\sigma_{34}(\mathrm{Id} \otimes \mathrm{Id} \otimes S \otimes S)\Delta^3 \\ &= (\mu \otimes \mu)(\mathrm{Id} \otimes \sigma \otimes \mathrm{Id})\sigma_{34}(\mathrm{Id} \otimes \mathrm{Id} \otimes S \otimes S)\Delta^3 \end{aligned} \tag{54.82}$$

where σ_{34} is the flip of the third and fourth component in $A \otimes A \otimes A \otimes A$. Now we have

$$(\mathrm{Id} \otimes \sigma \otimes \mathrm{Id})\sigma_{34}(a \otimes b \otimes c \otimes d) = (\mathrm{Id} \otimes \sigma \otimes \mathrm{Id})(a \otimes b \otimes d \otimes c) = a \otimes d \otimes b \otimes c, \tag{54.83}$$

thus $(\mathrm{Id}\,\sigma \otimes \mathrm{Id})\sigma_{34} = \sigma_{234}$ where σ_{234} is the cyclic permutation $a \otimes b \otimes c \otimes d \mapsto a \otimes d \otimes b \otimes c$. Using that and the relation (54.80) we continue the computation:

$$\diamondsuit = (\mu \otimes \mu)\sigma_{234}(\mathrm{Id} \otimes \mathrm{Id} \otimes S \otimes S)\Delta^3 \tag{54.84a}$$

$$= \Big[\mu(\mathrm{Id} \otimes S) \otimes \mu(\mathrm{Id} \otimes S)\Big]\sigma_{234}(\Delta \otimes \Delta)\Delta \tag{54.84b}$$

$$= \Big[\mu(\mathrm{Id} \otimes S) \otimes \mu(\mathrm{Id} \otimes S)\Big]\sigma_{234}(\mathrm{Id} \otimes \Delta \otimes \mathrm{Id})(\mathrm{Id} \otimes \Delta)\Delta \tag{54.84c}$$

$$= \Big[\mu(\mathrm{Id} \otimes S) \otimes \mu(\mathrm{Id} \otimes S)\Big](\mathrm{Id} \otimes \mathrm{Id} \otimes \Delta)(\mathrm{Id} \otimes \sigma)(\mathrm{Id} \otimes \Delta)\Delta. \tag{54.84d}$$

On the last line we did $\sigma_{234}(\mathrm{Id} \otimes \Delta \otimes \mathrm{Id}) = (\mathrm{Id} \otimes \mathrm{Id} \otimes \Delta)(\mathrm{Id} \otimes \sigma)$; one checks that equality by

applying both sides to $a\otimes b\otimes c$. In the following lines we use $(\mathrm{Id}\otimes\epsilon)\sigma=\sigma(\epsilon\otimes\mathrm{Id})$:

$$\begin{aligned}
\diamondsuit &= \Big[\mu(\mathrm{Id}\otimes S)\otimes \underbrace{\mu(\mathrm{Id}\otimes S)\Delta}_{\eta\circ\epsilon}\Big](\mathrm{Id}\otimes\sigma)(\mathrm{Id}\otimes\Delta)\Delta && (54.85a)\\
&= \Big[\mu(\mathrm{Id}\otimes S)\otimes\eta\Big](\mathrm{Id}\otimes\mathrm{Id}\otimes\epsilon)(\mathrm{Id}\otimes\sigma)(\mathrm{Id}\otimes\Delta)\Delta && (54.85b)\\
&= \Big[\mu(\mathrm{Id}\otimes S)\otimes\eta\Big](\mathrm{Id}\otimes\sigma\underbrace{(\epsilon\otimes\mathrm{Id})\Delta}_{1\otimes\mathrm{Id}})\Delta && (54.85c)\\
&= \Big[\mu(\mathrm{Id}\otimes S)\otimes\eta\Big](\mathrm{Id}\otimes\mathrm{Id}\otimes 1)\Delta && (54.85d)\\
&= \mu(\mathrm{Id}\otimes S)\Delta\otimes 1 && (54.85e)\\
&= \eta\epsilon\otimes 1. && (54.85f)
\end{aligned}$$

The last line is the map $a\mapsto\epsilon(a)1\otimes 1$ and then is the unit for $L(A,A\otimes A)$. □

54.6.2 Opposite algebra

From a Hopf algebra (A,Δ) we define the **opposite algebra** $(A,\Delta)^{op}=(A^{op},\Delta)$ and the coopposite coalgebra $(A,\Delta)^{cop}=(A,\sigma\circ\Delta)$. The opposite and coopposite algebras are not always Hopf algebras.

LemATSopHofounon

Lemma 54.30.
Let (A,Δ) be a Hopf algebra and a linear map $T\colon A\to A$. The following statements are equivalent:

ItemLemATSopHofounoni

(1) the bialgebra $(A,\Delta)^{op}$ is Hopf with T as antipode;

ItemLemATSopHofounonii

(2) $\mu\circ\sigma\circ(T\otimes\mathrm{Id})\circ\Delta=\mu\circ\sigma\circ(\mathrm{Id}\otimes T)\circ\Delta=\eta\circ\epsilon$;

ItemLemATSopHofounoniii

(3) $\sum_i b_iT(a_i)=\sum_i T(b_i)a_i=\eta\circ\epsilon(a)$ where $\Delta a=\sum_i a_i\otimes b_i$;

ItemLemATSopHofounoniv

(4) $\mu\circ(\mathrm{Id}\otimes T)\circ\sigma\circ\Delta=\mu\circ(T\otimes\mathrm{Id})\circ\sigma\circ\Delta=\eta\circ\epsilon$;

ItemLemATSopHofounonv

(5) the bialgebra $(A,\Delta)^{cop}$ is Hopf with T as antipode.

Proof. The equivalence (1) ⇔ (2) is the definition of an antipode. For the equivalence (1) ⇔ (3), following the definition of an antipode on $(A,\Delta)^{op}$, the two following lines are equal to $\eta\epsilon(a)$:

$$\mu\sigma(T\otimes\mathrm{Id})\Delta a=\sum_i\mu\sigma T(a_i)\otimes b_i=\sum_i b_iT(a_i) \tag{54.86}$$

while on the other hand

$$\mu\sigma(\mathrm{Id}\otimes T)\Delta a=\mu\sigma a_i\otimes T(b_i)=\sum_i T(b_i)a_i. \tag{54.87}$$

The equivalence (2) ⇔ (4) is the fact that $\sigma(\alpha\otimes\beta)=(\beta\otimes\alpha)\sigma$ whenever $\alpha,\beta\colon A\to A$.

The equivalence (4) ⇔ (5) is the definition of the antipode on $(A,\Delta)^{cop}$ since in the latter algebra the comultiplication is given by $\sigma\circ\Delta$ instead of Δ.

The proof is finished. □

PropAAopAcopHopfSSemu

Proposition 54.31.
Let (A,Δ) be a Hopf algebra. The following statements are equivalent:

ItemPropAAopAcopHopfSSemui

(1) S is bijective;

ItemPropAAopAcopHopfSSemuii

(2) $(A,\Delta)^{op}$ is a Hopf algebra with antipode S^{-1};

ItemPropAAopAcopHopfSSemuiii

(3) $(A,\Delta)^{cop}$ is a Hopf algebra with antipode S^{-1}.

In this case the antipode of $(A,\Delta)^{op}$ and $(A,\Delta)^{cop}$ is given by S^{-1}.

Proof. First we suppose that S is invertible and we prove that S^{-1} is an antipode for $(A,\Delta)^{op}$; in this case it will be an antipode for $(A,\Delta)^{cop}$ by the points (1) and (5) of lemma 54.30. We have to check that

$$\mu^{op}(\mathrm{Id}\otimes S^{-1})\Delta = \eta\circ\epsilon. \tag{54.88}$$

First we have

$$\mu^{op}(\mathrm{Id}\otimes S^{-1})\sum_i a_i\otimes b_i = S^{-1}(b_i)a_i. \tag{54.89}$$

Let us take the antipode on both sides and use the fact that S is an antipode for (A,Δ,μ):

$$S\mu^{op}(\mathrm{Id}\otimes S^{-1})\Delta(a) = \sum_i S(a_i)b_i = \mu(S\otimes\mathrm{Id})\Delta a = \eta\big(\epsilon(a)\big). \tag{54.90}$$

Thus

$$\mu^{op}(\mathrm{Id}\otimes S^{-1})\Delta = S^{-1}\circ\eta\circ\epsilon, \tag{54.91}$$

but $S^{-1}\circ\eta = \eta$ by lemma 54.28, so we have proved (1) $\Rightarrow$ (2),(3).

Let us now prove that (2) $\Rightarrow$ (1). For this part of the proof we follow[888]. Let $(A,\Delta)^{op}$ be a Hopf algebra with antipode T. We are going to prove that ST is an inverse of S for the convolution product, so that $ST = \mathrm{Id}$ and $T = S^{-1}$ and S is bijective. What we have to check is

$$\mu(ST\otimes S)\Delta = \eta\epsilon. \tag{54.92}$$

Using $\mu(S\otimes S)\sigma = S\mu$ we have

$$\begin{aligned}\mu(ST\otimes S)\Delta &= \mu(S\otimes S)(T\otimes\mathrm{Id})\Delta\\ &= S\mu\sigma(T\otimes\mathrm{Id})\Delta\\ &= S\mu(\mathrm{Id}\otimes T)\sigma\Delta\\ &= S\circ\eta\circ\epsilon\end{aligned} \tag{54.93}$$

because $\mu(\mathrm{Id}\otimes T)\sigma\Delta = \eta\circ\epsilon$ from the fact that T is an antipode for $(A,\Delta)^{cop}$. Now we conclude because $S\eta = \eta$. □

54.6.3 Dual of a Hopf algebra

Let $(\mathfrak{A},\mu,\Delta,\eta,\epsilon,S)$ be a Hopf algebra and $\mathfrak{A}^*$ the dual as vector space. We can create a Hopf algebra structure $(m^*,\mu^*,\Delta^*,\epsilon^*,S^*)$ on $\mathfrak{A}^*$ in the following way. First we extend the duality by $\langle f\otimes g, x\otimes y,\rangle = \langle f,x\rangle\langle g,y\rangle$ and then we define the following

subeqDualHopfmult

$$\langle m^*(f\otimes g),x\rangle = \langle f\otimes g,\Delta(x)\rangle \tag{54.94a}$$
$$\langle\Delta^*(f),x\otimes y\rangle = \langle f,xy\rangle \tag{54.94b}$$
$$\langle\eta^*(\alpha),x\rangle = \alpha\epsilon(x) \tag{54.94c}$$
$$\epsilon^*(f) = \langle f,1\rangle \tag{54.94d}$$
$$\langle S^*(f),x\rangle = \langle f,S(x)\rangle \tag{54.94e}$$

for every f, $g\in\mathfrak{A}^*$ and x, $y\in\mathfrak{A}$. One can show that these definitions fulfill all the condition of a Hopf algebra.

We say that a pair $(\mathfrak{A},\mathcal{B})$ is a **dual pair** of $*$-Hopf algebra is there exists a map $\langle .,.\rangle\colon \mathfrak{A}\otimes\mathcal{B}\to\mathbb{C}$ which fulfils the conditions

(1) $\langle x\otimes y,\Delta(f)\rangle = \langle ab,f\rangle$,
(2) $\epsilon(f) = \langle 1,f\rangle$,
(3) $\langle\Delta(x),f\otimes g\rangle = \langle x,fg\rangle$,
(4) $\langle x,S(f)\rangle = \langle S(x),f\rangle$,
(5) $\langle x,f^*\rangle = \overline{\langle S(x)^*,f\rangle}$

for every $x,y\in\mathfrak{A}$ and $f,g\in\mathcal{B}$

54.6.4 Involution, $*$-Hopf algebra

subsecHopfInvolution

We follow [887]. Let $\mathbb{K}$ be a commutative ring with unity.

Definition 54.32.
*An **involution** on a complex vector space V is a map $a \mapsto a^*$ such that*

(1) $(a+b)^* = a^* + b^*$ *for* $a, b \in V$;

(2) $(\lambda a)^* = \overline{\lambda} a^*$ *for every* $\lambda \in \mathbb{K}$ *and* $a \in A$;

(3) $(a^*)^* = a$.

An involution on a $\mathbb{K}$-algebra A is an involution $a \mapsto a^*$ on A (as vector space) which satisfies $(ab)^* = b^*a^*$ for every $a, b \in A$. An algebra equipped with an involution is a $*$-algebra.

¡¡ Avertissement/question au lecteur !! 54.33
In [889], one does not require $a^{**} = a$ while it required in [887]. In [881] they include $S\big(S(x)^*\big)^* = x$ in the definition of a Hopf$*$-algebra.

Definition 54.34.
*A **$*$-coalgebra** $(A, \Delta, *)$ is a coalgebra equipped with an involution such that*

$$\Delta(a^*) = \Delta(a)^* \tag{54.95}$$

where $(a \otimes b)^* = a^* \otimes b^*$.

Definition 54.35.
*A Hopf algebra (A, Δ) equipped with an involution $*$ such that $(A, *)$ is a $*$-algebra and $(A, \Delta, *)$ is a $*$-coalgebra is a **$*$-Hopf algebra**.*

LemcounitstarHopfalg

Lemma 54.36.
The counit of a counital $*$-coalgebra is $*$-linear, namely we have $\epsilon(a^*) = \overline{\epsilon(a)}$ where the bar stands for the involution in $\mathbb{K}$.

Proof. Let $(A, \Delta, *)$ be a $*$-coalgebra with counit ϵ. We consider the map

$$\begin{aligned} \epsilon^* \colon A &\to \mathbb{K} \\ a &\mapsto \overline{\epsilon(a^*)}. \end{aligned} \tag{54.96}$$

We prove that ϵ^* is a counit for A. We have to check the property (54.68b) for ϵ^*. If $\Delta a = \sum_i a_i \otimes b_i$,

$$\begin{aligned} (\mathrm{Id} \otimes \epsilon^*)\Delta(a) &= \sum_i a_i \otimes \overline{\epsilon(b_i^*)} \\ &= \sum_i a_i \overline{\epsilon(b_i^*)} && A \otimes \mathbb{K} = A \\ &= \sum_i \Big(\epsilon(b_i^*) a_i^*\Big)^* \\ &= (a^*)^* && \text{by formula (54.52)} \\ &= a. \end{aligned} \tag{54.97}$$

By unicity of the counit (lemma 54.25) we have $\epsilon^* = \epsilon$ and then $\overline{\epsilon(a^*)} = \epsilon(a)$. □

Theorem 54.37.
The antipode of a $*$-Hopf algebra is a bijection and $S \circ * \circ S \circ * = \mathrm{Id}$.

Proof. We consider the map

$$\begin{aligned} S^* \colon A &\to A \\ a &\mapsto S(a^*)^* \end{aligned} \tag{54.98}$$

and we show that S^* is an antipode for the opposite Hopf algebra $(A,\Delta)^{op}$. We have to check the definition property (54.70) in A^{op}. The multiplication in A^{op} being

$$\mu^{op}(a\otimes b)=ba \tag{54.99}$$

we have

$$\begin{aligned}\mu^{op}(S^*\otimes \mathrm{Id})\Delta(a)&=\sum_i b_iS^*(a_i)\\&=\sum_i b_iS(a^*)^*\\&=\sum_i\Big(S(a^*)b_i^*\Big)^*.\end{aligned} \tag{54.100}$$

Since S is an antipode for (A,Δ) the property (54.70) provides

$$\sum_i S(a_i^*)b_i^*=\eta\epsilon(a^*), \tag{54.101}$$

thus we have

$$\mu^{op}(S^*\otimes \mathrm{Id})\Delta(a)=\big(\eta\epsilon(a^*)\big)^*=\eta\epsilon(a). \tag{54.102}$$

So S^* is an antipode for $(A,\Delta)^{op}$ which becomes a Hopf algebra. Using the proposition 54.31 it proves that S is bijective and that S^* is the inverse of S. □

Definition 54.38.
Let A be a Hopf algebra. The action of A on itself defined by

$$\begin{aligned}\mathrm{ad}(a)\colon A&\to A\\ b&\mapsto\sum_i a_ibS(b_i)\end{aligned} \tag{54.103}$$

*where $a\in A$ and $\Delta(a)=\sum_i a_i\otimes b_i$ is the **adjoint action**.*

54.6.5 Example: universal enveloping algebra

SUBSECooTKZAooWVXXug

The set of function on a Lie group and the universal enveloping algebra of a Lie algebra are the two classical examples of Hopf algebras. The important example of Hopf algebra of continuous functions needs some more analysis and is the reported to definition 72.126.

The following puts a structure of Hopf algebra on the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ when G is a Lie group:

(1) $X\cdot Y=X\otimes Y$, the ordinary multiplication in $\mathcal{U}(\mathfrak{g})$,
(2) $\Delta X=X\otimes 1+1\otimes X$,
(3) $\eta(\lambda)=\lambda 1$,
(4) $\epsilon(1)=1$ and $\epsilon(X)=0$ otherwise
(5) $S(X)=-X$.

ItemCounitUg

It is proven in [890] that the two latter constructions are dual. See also the remark 62.55 for an example.

PropHopfSurDual

Proposition 54.39.
Let A be a finite dimensional Hopf algebra on a field of characteristic zero. The following structure defines a canonical Hopf algebra structure on the dual A^*

$$(f_1f_2)(a)=(f_1\otimes f_2)\Delta a \tag{54.104a}$$
$$(\Delta f)(a\otimes b)=f(ab) \tag{54.104b}$$
$$S(f)a=f\big(S(a)\big) \tag{54.104c}$$
$$\epsilon(f)=f(1) \tag{54.104d}$$

DefHopfSurAstar

where $(f_1\otimes f_2)(a\otimes b)=f_1(a)f_2(b)$

Proof. We check the property $\Delta\big(S(f)\big) = (S\otimes S)\Delta'(f)$ where $\Delta' = \sigma\Delta$ and $\sigma(a\otimes b) = b\otimes a$. On the one hand we have

$$\Delta\big(S(f)\big)(a\otimes b) = S(f)(ab) = f\big(S(ab)\big) = f\big(S(b)S(a)\big). \tag{54.105}$$

On the other hand, if

$$\Delta(f) = \sum_i f_i\otimes g_i \tag{54.106}$$

we have

$$\begin{aligned}\clubsuit = (S\otimes S)(\Delta' f)(a\otimes b) &= (S\otimes S)\sum_i \sigma(f_i\otimes g_i)(a\otimes b)\\ &= \sum_i S(g_i)\otimes S(f_i)(a\otimes b)\\ &= \sum_i g_i\big(S(a)\big) f_i\big(S(b)\big).\end{aligned} \tag{54.107}$$

Using the fact that $\mathbb{K}$ is commutative (it is a field), we continue

$$\begin{aligned}\clubsuit &= \sum_i f_i\big(S(b)\big) g_i\big(S(a)\big)\\ &= \sum_i (f_i\otimes g_i)\big(S(b)\otimes S(a)\big)\\ &= (\Delta f)\big(S(b)\otimes S(a)\big)\\ &= f\big(S(b)S(a)\big).\end{aligned} \tag{54.108}$$

□

DefInvolutionHopf

Definition 54.40.
*Let A be a $\mathbb{K}$-algebra with involution. A A-module V is **unitarizable** if there exists an Hermitian scalar product $V\otimes V\to\mathbb{K}$ such that*

$$\langle av_1, v_2\rangle = \langle v_1, a^* v_2\rangle \tag{54.109}$$

for every $a\in A$, $v_1, v_2\in V$.

PropAstarstaralg

Proposition 54.41.
Let A be a $*$-Hopf algebra. The dual A^* becomes a $*$-Hopf algebra with the involution $l\mapsto l^*$ defined by

$$l^*(a) = \overline{l\big(S(a)^*\big)} \tag{54.110}$$

with $l\in A^*$ and $a\in A$.

Proof. We check the condition $\Delta(l^*) = \Delta(l)^*$. We recall the definition (54.104b): $\Delta(l)(a\otimes b) = l(ab)$. On the one hand we have

$$\begin{aligned}\Delta(l^*)(a\otimes b) &= l^*(ab)\\ &= \overline{l\big(S(ab)^*\big)}\\ &= \overline{l\Big(\big(S(b)S(a)\big)^*\Big)}\\ &= \overline{l\big(S(a)^*S(b)^*\big)}\end{aligned} \tag{54.111}$$

EqCalculDellstraotb

On the other hand if we write $\Delta(l) = \sum_i f_i \otimes g_i$ we have

$$\begin{aligned}\Delta(l)^*(a\otimes b) &= \sum_i (f_i^*\otimes g_i^*)(a\otimes b)\\ &= \sum_i \overline{f_i\big(S(a)^*\big)g_i\big(S(b)^*\big)}\\ &= \sum_i \overline{(f_i\otimes g_i)\big(S(a)^*\otimes S(b)^*\big)}\\ &= \overline{\Delta(l)\big(S(a)^*\otimes S(b)^*\big)}\\ &= \overline{l\big(S(a)^*S(b)^*\big)},\end{aligned} \tag{54.112}$$

which is the same as the last line of (54.111). □

Let A be a Hopf algebra. A linear form $\alpha \in A^*$ is **left invariant** if

$$(\mathrm{Id}\otimes\alpha)\circ\Delta = \eta\circ\alpha. \tag{54.113}$$

The linear form α is **right invariant** if

$$(\alpha\otimes\mathrm{Id})\circ\Delta = \eta\circ\alpha. \tag{54.114}$$

54.6.6 Modules and representations

SubSecMOdulREepe

Definition 54.42.
*Let A be a Hopf algebra. A A-**module** is a module for the* algebra *structure.*

Let V_1 and V_2 be two A-modules. The **tensor product** is the $\mathbb{K}$-module $V_1\otimes V_2$ with the action of A given by

$$a\cdot(v_1\otimes v_2) = \Delta(a)(v_1\otimes v_2). \tag{54.115}$$

If A is a Hopf algebra on $\mathbb{K}$, we say that a A-module is **finite dimensional** if it is a free $\mathbb{K}$-module of finite type. The homomorphism $\pi\colon A\to \mathrm{End}\,V$ corresponding to the module structure is a **finite dimensional representation**.

The following definitions are from [888].

If V is a left A-module, the dual V^* becomes a right A-module by the definition

$$(\alpha\cdot a)(v) = \alpha(a\cdot v) \tag{54.116}$$

for every $a\in A$, $\alpha\in V^*$ and $v\in V$. If V is a right module, we turn it into a left module using the antipode:

$$a\cdot v = v\cdot S(a). \tag{54.117}$$

In such a way, V^* is a left A-module by $a\cdot\alpha = \alpha\cdot S(a)$, that is

$$(a\cdot\alpha)(v) = \alpha\big(S(a)\cdot v\big). \tag{54.118}$$

EqDefacctrleftUoGLstat

Yet another way to make V^* a left A-module is to use S^{-1} instead of S. This will also produce a left module structure since $S^{-1}(b)S^{-1}(a) = S^{-1}(ab)$. This representation of A will be denoted by $\mathcal{L}$:

$$\begin{aligned}&\mathcal{L}\colon A\to\mathrm{End}(V^*)\\ &\big(\mathcal{L}(a)f\big)(v) = f\big(S^{-1}(a)v\big).\end{aligned} \tag{54.119}$$

EqDefTroisleftmodAstar

Remarque 54.43.
This is not the regular left representation of definition 54.45. One difference is that the regular left representation is well defined by itself while the action $\mathcal{L}(a)\colon V^*\to V^*$ is only defined when a representation of A is given on V.

54.6.7 Matrix elements of modules

Let $\mathbb{K}$ be a commutative ring with unit and A, an unital $\mathbb{K}$-algebra. Let V be a A-module. A pair $(l,v) \in V^* \times V$ defines a linear functional on A by

$$a \mapsto l(av). \tag{54.120}$$

This functional is the **matrix element** of the representation $\pi\colon A \to \mathrm{End}_{\mathbb{K}}(V)$ corresponding to the pair (l,v). It is denoted by $c^{\pi}_{l,v}$ or $c^{V}_{l,v}$.

In this section we suppose that the A-modules are projective as $\mathbb{K}$-modules.

PropHopfDual

Proposition 54.44.
Let A be a Hopf algebra on $\mathbb{K}$. The matrix elements of the finite dimensional A-modules form a Hopf subalgebra of the dual $A^\star = \mathrm{Hom}_{\mathbb{K}}(A,\mathbb{K})$.

The Hopf algebra defined in proposition 54.44 is denoted by A^* and is the **dual Hopf algebra** of A.

DefUMXgVdT

Definition 54.45.
*Let A be a Hopf algebra. The **regular left representation** of A on A^* is given by*

$$\mathcal{R}(a)f = \langle \mathrm{Id} \otimes a, \Delta(f)\rangle. \tag{54.121}$$

*The **regular right** is*

$$\mathcal{L}(a)f = \langle S^{-1}(a) \otimes \mathrm{Id}, \Delta(f)\rangle \tag{54.122}$$

where Δ is the antipode of A^*.

More explicitly, for every $a, x \in U_q\mathfrak{g}$,

$$\big(\mathcal{R}(a)f\big)(x) = (\Delta f)(x \otimes a) \tag{54.123}$$

and

$$\big(\mathcal{L}(a)f\big)(x) = (\Delta f)\big(S^{-1}(a) \otimes x\big). \tag{54.124}$$

One common idea is to see A^* as a $U_q\mathfrak{g} \otimes U_q\mathfrak{g}$-modules by the representation $\mathcal{L} \otimes \mathcal{R}$. With that structure we have

$$\big((a \otimes b)f\big)(x) = \big(\mathcal{L}(a)\mathcal{R}(b)f\big)(x). \tag{54.125}$$

EqmRmLAAsurAstar

Using the definition of the coproduct given in proposition 54.39,

$$\begin{aligned}\big(\mathcal{L}(a)\mathcal{R}(b)f\big)(x) &= \Delta(\mathcal{R}(b)f)\big(S^{-1}(a) \otimes x\big)\\ &= \big(\mathcal{R}(b)f\big)\big(S^{-1}(a)x\big)\\ &= \Delta(f)\big(S^{-1}(a)x \otimes b\big)\\ &= f\big(S^{-1}(a)xb\big).\end{aligned} \tag{54.126}$$

EqmRmLabf

54.6.8 Quasitriangular Hopf algebra

Definition 54.46.
*A **quasitriangular** Hopf algebra (or a **braid**) is a Hopf algebra $(\mathfrak{A}, \mu, \eta, \Delta, \epsilon, S)$ together with an invertible element $R \in \mathfrak{A} \otimes \mathfrak{A}$ such that*

(1) $R\Delta(x)R^{-1} = \tau \circ \Delta(x)$,

ItemCondUnifRi

(2) (2a) $(\Delta \otimes \mathrm{Id})(R) = R_{13}R_{23}$,

(2b) $(\mathrm{Id} \otimes \Delta)(R) = R_{13}R_{12}$

ItemCondUnifRiib

*In that situation, the element R is said to be an **universal R-matrix**.*

Theorem 54.47 (Yang-Baxter equation)**.**
An universal R-matrix satisfies

$$R_{12}R_{13}R_{23} = R_{23}R_{13}R_{12}. \tag{54.127}$$

Proof. We compute the quantity $(\mathrm{Id}\otimes\tau\circ\Delta)R$ in two different ways. First we compute it using (2b), and then we first use (1) before (2b). In the first case we have (a sum is implied over repeated indices)

$$\begin{aligned}(\mathrm{Id}\otimes\tau\circ\Delta)R &= (\mathrm{Id}\,\tau)\circ(\mathrm{Id}\otimes\Delta)R\\ &= (\mathrm{Id}\otimes\tau)(R_{13}R_{12})\\ &= (\mathrm{Id}\otimes\tau)(a_i\otimes 1\otimes b_i)(a_k\otimes b_k\otimes 1)\\ &= (\mathrm{Id}\otimes\tau)(a_ia_k\otimes b_k\otimes b_i)\\ &= a_ia_k\otimes b_i\otimes b_k\\ &= (a_i\otimes b_i\otimes 1)(a_k\otimes 1\otimes b_k)\\ &= R_{12}R_{13}.\end{aligned}$$

Now if we denote $\Delta(b_i) = b_{i1}\otimes b_{i2}$ and $R^{-1} = \alpha_k\otimes\beta_k$, remark that we have $R^{-1}_{ij} = (R^{-1})_{ij}$, so that

$$\begin{aligned}(\mathrm{Id}\otimes\tau\circ\Delta)(a_i\otimes b_i) &= a_i\otimes\Big((a_j\otimes b_j)(b_{i1}\otimes b_{i2}(\alpha_k\otimes\beta_k))\Big)\\ &= \Big(a_i\otimes\big((a_i\otimes b_j)(b_{i1}\otimes b_{i2})\big)\Big)(1\otimes\alpha_k\otimes\beta_k)\end{aligned}$$

The first factor of the last line can be written as

$$a_i\otimes a_jb_{i1}\otimes b_jb_{i2} = (1\otimes a_j\otimes b_j)(a_i\otimes b_{i1}\otimes b_{i2}) = R_{23}\big((\mathrm{Id}\otimes\Delta)R\big) = R_{23}R_{13}R_{12},$$

while the second factor is $(R^{-1})_{23}$. Finally we find $(\mathrm{Id}\otimes\tau\circ\Delta)(a_i\otimes b_i) = R_{23}R_{13}R_{12}R_{23}^{-1}$. Equating with the first value obtained, we find the claim. □

54.7 Convolution semigroup

Definition 54.48.
*A **convolution semigroup** of linear functionals on B is a set of functionals φ_t such that*

(1) $\varphi_0 = \epsilon$,

(2) $\lim_{t\searrow 0}\varphi_t(b) = \epsilon(b)$

(3) $\varphi_s * \varphi_t = \varphi_{s+t}$

for every $b\in B$.

Definition 54.49.
*If j_1 and j_2 are linear functionals on a coalgebra C taking values in an algebra A, we define the **convolution** by*

$$j_1 * j_2 = m_A\circ(j_1\otimes j_2)\circ\Delta. \tag{54.128}$$

Lemma 54.50.
Let C be a coalgebra. We have

(1) If $\psi\colon C\to\mathbb{C}$ is a linear functional on C, then the series

$$\exp(\psi)a = \sum_{n=0}^{\infty}\frac{\psi^{*n}}{n!}(a) = \epsilon(a)+\psi(a)+\frac{1}{2}(\psi*\psi)(a)+\dots \tag{54.129}$$

converges for every $a\in C$. This defines the map $\exp(\psi)\colon C\to\mathbb{C}$.

(2) Let $(\varphi_t)_{t\geqslant 0}$ be a convolution semigroup on C. Then the limit

$$L(a) = \lim_{t\searrow 0}\frac{1}{t}\big(\varphi_t(a)-\epsilon(a)\big) \tag{54.130}$$

exists for every $a\in C$. Moreover we have

$$\varphi_t = \exp(tL) \tag{54.131}$$

for $t\geqslant 0$.

54.8 Modules bialgebras

54.8.1 Module (bi)algebras

subSecModulebialgebra

This subsection comes from [891] and we follow the notations of [892]. The notions of module algebra and module coalgebra will be used when we will study twists in section 81.1.

Let B be a bialgebra and $\mathcal{E}$, a left B-module. For each $b \in B$, we have the map

$$\begin{aligned} b_l &: \mathcal{E} \to \mathcal{E} \\ b_l(m) &= b \cdot m. \end{aligned} \tag{54.132}$$

In the same way, if $\mathcal{F}$ is a right B-module, we have the map

$$\begin{aligned} b_r &: \mathcal{F} \to \mathcal{F} \\ b_r(m) &= m \cdot b. \end{aligned} \tag{54.133}$$

In that case, $\mathcal{E} \otimes \mathcal{E}$ becomes a right $B \otimes B$-module by

$$\begin{aligned} (b \otimes b')_l &: \mathcal{E} \otimes \mathcal{E} \to \mathcal{E} \otimes \mathcal{E} \\ (b \otimes b')_l(m \otimes m') &= (b \cdot m) \otimes (b' \cdot m'). \end{aligned} \tag{54.134}$$

DefBModuleAlgebra

Definition 54.51.
*A left B-module $\mathbb{A}$ is a **algebra B-module** if for every $b \in B$ and $a, a' \in \mathbb{A}$,*

(1) $b \cdot 1_{\mathbb{A}} = \epsilon(b) \cdot 1_{\mathbb{A}}$. This equality has to be seen in $\mathbb{K} \subset B$

(2) $b \cdot (aa') = \sum (b_{(1)} \cdot a)(b_{(2)} \cdot a')$ where $\Delta_B(b) = \sum b_{(1)} \otimes b_{(2)}$.

The second condition is the commutativity of the diagram

EqDiagModAlg

$$\begin{array}{ccc} \mathbb{A} \otimes \mathbb{A} & \xrightarrow{\mu_{\mathbb{A}}} & \mathbb{A} \\ {\scriptstyle \left(\Delta_B(b)\right)_l}\downarrow & & \downarrow{\scriptstyle b_l} \\ \mathbb{A} \otimes \mathbb{A} & \xrightarrow{\mu_{\mathbb{A}}} & \mathbb{A} \end{array} \tag{54.135}$$

for every $b \in B$.

Definition 54.52.
We say that C is a coalgebra and B is a bialgebra, we say that C is a B-module coalgebra if it is a B-module satisfying

ItemDefCoalgModuleUn

(1) $\Delta(a \cdot b) = \sum (a_{(1)} \cdot b_{(1)}) \otimes (a_{(2)} \otimes b_{(2)})$,

(2) $\epsilon(a \cdot b) = \epsilon(a)\epsilon(b)$

for any $a \in C$ and $b \in B$.

From a notational point of view, notice that

$$\begin{aligned} \Delta_C(a) &= \sum_i a^i_{(1)} \otimes a^i_{(2)} \\ \Delta_B(b) &= \sum_j b^j_{(1)} \otimes b^j_{(2)} \end{aligned} \tag{54.136}$$

and the sum in the point (1) is a double sum. The formula is equivalent to commutativity of the diagram

EqDiagModCoAlg

$$\begin{array}{ccc} C \otimes C & \xleftarrow{\Delta} & C \\ {\scriptstyle \Delta(b)_r}\uparrow & & \uparrow{\scriptstyle b_r} \\ C \otimes C & \xleftarrow{\Delta} & C \end{array} \tag{54.137}$$

for every $b \in B$, which is the dual diagram of (54.135).

54.8.2 Tensor product of representations

Let (ψ_1, V_1) and (ψ_2, V_2) be two representations of the algebra $\mathfrak{A}$. We want to define the tensor product $(\psi_1, V_) \otimes (\psi_2, V_2)$. The two first ideas are

$$a(v_1 \otimes v_2) = \big(\psi_1(a)v_1\big) \otimes \big(\psi_2(a)v_2\big) \tag{54.138a}$$
$$a(v_1 \otimes v_2) = \psi_1(a)v_1 \otimes v_2 + v_1 \otimes \psi_2(a)v_2. \tag{54.138b}$$

The bad news is that the first possibility cannot be linear while the second one fails in genera to complete the homomorphism condition. Notice however that, on a Lie algebra, the second possibility works because the product is skew-symmetric.

The trick in order to construct a tensor product of two representations is to take a map $\Delta\colon \mathfrak{A} \to \mathfrak{A} \otimes \mathfrak{A}$ and to define the tensor representation by

$$\Psi = (\psi_1 \otimes \psi_2)\Delta, \tag{54.139}$$

in other words if $a \in \mathfrak{A}$ and $\Delta(a) = a_1 \otimes a_2$, we define

$$\Psi(a)(v_1 \otimes v_2) = (\psi_1 \otimes \psi_2)\Delta(a)(v_1 \otimes v_2) = \psi_1(a_1)v_1 \otimes \psi_2(a_2)v_2. \tag{54.140}$$

Proposition 54.53.
If we impose Ψ to be

— *linear,*

— *homomorphic,*

— *associative, i.e.* $\big((\psi_1, V_1) \otimes (\psi_2, V_2)\big) \otimes (\psi_3, V_3) = (\psi_1, V_1) \otimes \big((\psi_2, V_2) \otimes (\psi_3, V_3)\big)$,

then Δ must be a comultiplication.

Proof. No proof. □

54.9 Poisson structure

A **Poisson algebra** is a commutative algebra $(\mathfrak{A}, m, \eta)$ endowed with a map $\{.,.\}\colon \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ called the *Poisson bracket* such that

(1) $\big(\mathfrak{A}, \{.,.\}\big)$ is a Lie algebra,

(2) $\{ab, c\} = a\{b, c\} + \{a, c\}b$

for all a, b, $c \in \mathfrak{A}$. The second condition shows that $\{a, .\}$ is a derivation of $\mathfrak{A}$. Since the bracket is bilinear, it defined and is defined by a map $\gamma\colon \mathfrak{A} \otimes \mathfrak{A} \to \mathfrak{A}$ and the properties of the Poisson bracket are translated in terms of algebra structure in the following way.

First the antisymmetry is encoded by the requirement

$$\gamma \circ \tau = \gamma. \tag{54.141}$$

The Jacobi identities $\{a, \{b, c\}\} + \{b, \{c, a\}\} + \{c, \{a, b\}\} = 0$ becomes

$$\gamma\big(a \otimes \gamma(b \otimes c)\big) + \gamma\big(b \otimes \gamma(c \otimes a)\big) + \gamma\big(c \otimes \gamma(a \otimes b)\big) = 0,$$

or after rearrangements,

$$\gamma(1 \otimes \gamma)\big(1 \otimes 1 \otimes 1 + (1 \otimes \tau)(\tau \otimes 1) + (\tau \otimes 1)(1 \otimes \tau)\big) = 0 \tag{54.142}$$

and the last condition reads

$$\gamma(m \otimes 1) = m(1 \otimes \gamma)\big(1 \otimes 1 \otimes 1 + (1 \otimes \tau)\big). \tag{54.143}$$

54.9.1 Homomorphism

Let $\big(\mathfrak{A}, m_{\mathfrak{A}}, \{.,.\}_{\mathfrak{A}}\big)$ and $\big(\mathfrak{A}, m_{\mathcal{B}}, \{.,.\}_{\mathcal{B}}\big)$ two Poisson algebras. An **homomorphism** is a map $f\colon \mathfrak{A} \to \mathcal{B}$ such that

(1) $f(ab) = f(a)f(b)$,

(2) $f\big(\{a,b\}_{\mathfrak{A}}\big) = \{f(a), f(b)\}_{\mathcal{B}}$

for every a, $b \in \mathfrak{A}$. These conditions can be written in terms of $m_{\mathfrak{A}}$, $m_{\mathcal{B}}$, $\gamma_{\mathfrak{A}}$ and $\gamma_{\mathcal{B}}$:

— $f \circ \gamma_{\mathfrak{A}} = m_{\mathcal{B}} \circ (f \otimes f)$,

— $f \circ \gamma_{\mathfrak{A}} = \gamma_{\mathcal{B}} \circ (f \otimes f)$.

54.9.2 Tensor product

We can build the tensor product $\mathfrak{A} \otimes \mathcal{B}$ and define the product $m_{\mathfrak{A}\otimes\mathcal{B}}$ by

$$(a_1 \otimes b_1) \cdot (a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2$$

which can immediately be rewritten under the form

$$m_{\mathfrak{A}\otimes\mathcal{B}} = (m_{\mathfrak{A}} \otimes m_{\mathcal{B}})(1 \otimes \tau \otimes 1). \tag{54.144}$$

We define the Poisson bracket over the tensor product by

$$\{a_1 \otimes b_1, a_2 \otimes b_2\} = \{a_1, a_2\} \otimes b_1 b_2 + a_1 a_2 \otimes \{b_1, b_2\}$$

which is nothing else that

$$\gamma_{\mathfrak{A}\otimes\mathcal{B}} = (\gamma_{\mathfrak{A}} \otimes m_{\mathcal{B}} + m_{\mathfrak{A}} \otimes \gamma_{\mathcal{B}})(1 \otimes \tau \otimes 1). \tag{54.145}$$

Definition 54.54.
*A **co-Poisson bialgebra** is a co-commutative bialgebra $(\mathfrak{A}, m, \Delta, \eta, \epsilon)$ with a map $\delta\colon \mathfrak{A} \to \mathfrak{A} \otimes \mathfrak{A}$ such that*

(1) $\tau \circ \delta = -\delta$,

(2) $\big(1 \otimes 1 \otimes 1 + (1 \otimes \tau)(\tau \otimes 1) + (\tau \otimes 1)(1 \otimes \tau)\big)(1 \otimes \gamma)\delta = 0$,

(3) $(\Delta \otimes 1)\delta = (1 \otimes 1 \otimes 1 + \tau \otimes 1)(1 \otimes \delta)\Delta$,

(4) $(m \otimes m) \circ \delta_{\mathfrak{A}\otimes\mathfrak{A}} = \delta \circ m$

where $\delta_{\mathfrak{A}\otimes\mathfrak{A}} = (1 \otimes \tau \otimes 1)(\delta \otimes \Delta + \Delta \otimes \delta)$ is the co-Poisson structure associated with the tensor product space.

54.10 Lie bialgebra

Definition 54.55.
*A vector space $\mathfrak{g}$ equipped with a map $\phi\colon \mathfrak{g} \to \mathfrak{g} \otimes \mathfrak{g}$ is a **Lie coalgebra**[893] if the map ϕ satisfies the co-skew symmetry and co-Jacobi properties, that is if*

(1) $(\sigma + \mathrm{Id}) \circ \phi = 0$;

(2) $(\mathrm{Id} + \xi + \xi^2) \circ (\mathrm{Id} \otimes \phi) \circ \phi = 0$

where σ is the flip operator in $\mathfrak{g} \otimes \mathfrak{g}$, $\sigma(u \otimes v) = v \otimes u$ and ξ is the cyclic permutation operator on $\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}$, $\xi(u \otimes v \otimes w) = v \otimes w \otimes u$. See subsection 54.5.2 for a justification of these expressions.

Definition 54.56.
*A pair $(\mathfrak{g}, \phi)$ in which $\mathfrak{g}$ is a Lie algebra and ϕ is a Lie coalgebra structure on $\mathfrak{g}$ is a **Lie bialgebra** if the commutator and ϕ satisfy the following compatibility property:*

$$\phi\big([X,Y]\big) = X \cdot \phi(Y) - Y \cdot \phi(X) \tag{54.146}$$

where $X \cdot (Y \otimes Z) = [X,Y] \otimes Z + Y \otimes [X,Z]$.

PropStandardBialgStruct

Proposition 54.57.
Let $\mathfrak{g}$ be a complex simple Lie algebra with its Killing form B and a Cartan subalgebra $\mathfrak{h}$. Let $n = \dim \mathfrak{h}$ be the rank of $\mathfrak{g}$. Let $X_i^\pm$ and H_i $(i = 1, \dots, n)$ be the Chevalley generators and d_i be the numbers such that $d_i A_{ij} = d_j A_{ji}$ given by the lemma 51.138. Then the cobracket $\phi\colon \mathfrak{g} \to \mathfrak{g} \otimes \mathfrak{g}$

$$\begin{aligned} \phi(H_i) &= 0 \\ \phi(X_i^\pm) &= d_i X_i^\pm \wedge H_i \end{aligned} \tag{54.147}$$

EqDefCobrackStandard

is a Lie bialgebra structure on $\mathfrak{g}$.

This structure is the **standard bialgebra structure** on $\mathfrak{g}$. The definition (54.147) is given in the spirit of the remark 51.133.

Proof. There are three conditions to be satisfied.

(1) If one apply $(\sigma + \mathrm{Id}) \circ \phi$ to H_i, we get zero. If one apply this to $X_i^\pm$, one gets

$$\begin{aligned} (\sigma + \mathrm{Id}) \circ \phi(X_i^\pm) &= d_i(\sigma + \mathrm{Id})(X_i^\pm \otimes H_i - H_i \otimes X_i^\pm) \\ &= d_i(H_i \otimes X_i^\pm - X^\pm \otimes H_i + X_i^\pm \otimes H_i - H_i \otimes X_i^\pm) \\ &= 0. \end{aligned} \tag{54.148}$$

(2) The co-Jacobi identity is easy to check on H_i. For $X_i^\pm$ we have

$$\begin{aligned} (\mathrm{Id} + \xi + \xi^2) \circ (\mathrm{Id} \otimes \phi) \circ \phi(X_i^\pm) &= d_i(\mathrm{Id} + \xi + \xi^2) \circ (\mathrm{Id} \otimes \phi)(X_i^\pm \otimes H_i - H_i \otimes X_i^\pm) \\ &= -d_i^2(\mathrm{Id} + \xi + \xi^2)(H_i \otimes X_i^\pm \otimes H_i - H_i \otimes H_i \otimes X_i^\pm) \\ &= -d_i^2(H_i \otimes X_i^\pm \otimes H_i - H_i \otimes H_i \otimes X_i^\pm) \\ &\quad + d_i^2(X_i^\pm \otimes H_i \otimes H_i - H_i \otimes X_i^\pm \otimes H_i) \\ &\quad + d_i^2(H_i \otimes H_i \otimes X_i^\pm - X_i^\pm \otimes H_i \otimes H_i) \\ &= 0. \end{aligned} \tag{54.149}$$

(3) We have to check the value of $\phi\big([H_i, X_j^\pm]\big)$ and $\phi\big([X_i^+, X_j^-]\big)$. For the first one we know that $[H_i, X_j^\pm] = \alpha_j(H_i) X_j^\pm$, so that

$$\phi\big([H_i, X_j^\pm]\big) = \pm \alpha_j(H_i) \phi(X_j^\pm) = \pm d_i \alpha_j(H_i) X_j^\pm \wedge H_i. \tag{54.150}$$

EqphiHXcobrack

On the other hand

$$\begin{aligned} H_i \cdot \phi(X_j^\pm) - X_j^\pm \cdot \phi(H_i) &= d_j H_i \cdot (X_j^\pm \otimes H_j - H_j \otimes X_j^\pm) \\ &= d_j [H_i, X_j^\pm] \otimes H_j - d_j H_j \otimes [H_i, X_j^\pm] \\ &= d_j [H_i, X_j^\pm] \wedge H_j, \end{aligned} \tag{54.151}$$

which is the same as what we found in (54.150).
We check the property for $\phi\big([X_i^+, X_j^-]\big) = \phi(\delta_{ij} H_i) = 0$ in the same way (but there are more terms).

□

54.11 Poisson-Lie group

Definition 54.58.
*A Lie group is a **Poisson-Lie group** if the space of smooth functions is Poisson Hopf algebra.*

Notice that the Hopf structure on $C^\infty(G)$ is defined by the group structure of G, so that the fact to be a Lie-Poisson group only imposes conditions on the Poisson structure. The condition to be fulfilled is

$$\{\Delta(a),\Delta(b)\}_{\mathfrak{A}\otimes\mathfrak{A}} = \Delta\big(\{a,b\}_{\mathfrak{A}}\big). \tag{54.152}$$

EqCondcmpDelPoiss

We know on the other hand that every Poisson structure on $C^\infty(G)$ reads

$$\{f,h\}(g) = \sum_{ij} \eta^{ij}(g) X_i(g)(f) X_j(g)(h) \tag{54.153}$$

Poissgeneclafo

for a certain η. That Poisson bracket can be rewritten as

$$\{f,h\}(g) = \eta(g)(df_g \otimes dh_g) \tag{54.154}$$

where we defined

$$\begin{aligned} \eta &\colon G \to \mathcal{G}\otimes\mathcal{G} \\ \eta(g) &= \eta^{ij}(g)(X_i\otimes X_j). \end{aligned} \tag{54.155}$$

Let $C^n(G,\mathcal{G})$ be the set of maps $\lambda\colon G\times\cdots\times G\to\mathcal{G}\otimes\mathcal{G}$. The union of that spaces can be turned into a complex taking the coboundary

$$\begin{aligned} \delta_G &\colon C^n(G,\mathcal{G}) \to C^{n+1}(G,\mathcal{G}) \\ (\delta_G\lambda)(g_1,\cdots,g_{n+1}) &= g\cdot\lambda(g_2,\cdots,g_{n+1}) \\ &\quad + \sum_{i=1}^{n}(-1)^i\lambda(g_1,\cdots,g_ig_{i+1},\cdots,g_{n+1}) \\ &\quad + (-1)^{n+1}\lambda(g_1,\cdots,g_n). \end{aligned} \tag{54.156}$$

For that, we defined the action of G on $\mathcal{G}\otimes\mathcal{G}$ by

$$g\cdot(X\otimes Y) = \mathrm{Ad}(g)X\otimes\mathrm{Ad}(g)Y = \big(\mathrm{Ad}(g)\otimes\mathrm{Ad}(g)\big)(X\otimes Y),$$

and we can prove that $\delta_G^2=0$. The interest of that complex is that the compatibility condition (54.152) reads $\delta_G\eta=0$, or more specifically

$$(\delta_G\eta)(g_1,g_2) = g_1\cdot\eta(g_2) - \eta(g_1,g_2) + \eta(g_1) = 0.$$

The claim is that for the Poisson structure (54.153) to be compatible with the natural Hopf algebra structure, one needs

$$\delta_G\eta = 0. \tag{54.157}$$

That is only compatibility. Every such η does not define a Poisson-Lie group. Let us introduce the function $d\eta_e$, that we denote by ϕ_η:

$$\begin{aligned} \phi_\eta &\colon \mathcal{G} \to \mathcal{G}\otimes\mathcal{G} \\ X &\mapsto \frac{d}{dt}\Big[\eta(e^{tX})\Big]_{t=0}. \end{aligned} \tag{54.158}$$

Theorem 54.59.
In order to define a Poisson-Lie group G, the map ϕ_η must define a Lie bialgebra structure on $\mathfrak{g}$.

One way to find such a η is to take a $r\in\mathcal{G}\otimes\mathcal{G}$ and to define $\eta=\delta_G r$. By the definitions, we have $\delta_G r(g) = r - g\cdot r$ and thus

$$\phi_\eta(X) = \frac{d}{dt}\Big[\eta(e^{tX})\Big]_{t=0} = \frac{d}{dt}\Big[r - e^{tX}r\Big]_{t=0} = -X\cdot r.$$

A generic r reads $r=r^{ij}(X_i\otimes X_j)$ and then we have

$$\phi_\eta(X) = [r, 1\otimes X + X\otimes 1]. \tag{54.159}$$

54.12 Cyclic cohomology

SecCyclicHomol

Let $\mathfrak{A}$ be an algebra over $\mathbb{C}$. We consider the complex (C_λ^n, b) defined by

— C_λ^n is the set of $(n+1)$-linear functionals φ on $\mathfrak{A}$ such that

$$\varphi(a^1, \cdots, a^n, a^0) = (-1)^n \varphi(a^0, \cdots, a^n), \tag{54.160}$$

— the **Hochschild coboundary** b is defined by

$$\begin{aligned}(b\varphi)(a^0, \cdots, a^{n+1}) = &\sum_{j=0}^{n} (-1)^j \varphi(a^0, \cdots, a^j a^{j+1}, \cdots, a^{n+1}) \\ &+ (-1)^{n+1} \varphi(a^{n+1} a^0, \cdots, a^n).\end{aligned} \tag{54.161}$$

We denote by C_λ^0 the set of linear functionals.

We denote by $HC^*(\mathfrak{A})$ the cohomology of this complex.

DefCycleCoh

Definition 54.60.
*A n-dimensional **cycle** is a triple $(\Omega, d, \int)$ with*

— $\Omega = \sum_{j=0}^n \Omega^j$ *is a graded algebra over $\mathbb{C}$,*

— d *is graduate derivative of degree* 1,

— $\int : \Omega^n \to \mathbb{C}$ *is a **closed trace**, that is $\int d\omega = 0$.*

When $\mathfrak{A}$ is an algebra over $\mathbb{C}$, we say that a cycle over $\mathfrak{A}$ *is a cycle endowed with a homomorphism $\rho : \mathfrak{A} \to \Omega^0$.*

An important subset of C_λ^n is the set

$$Z_\lambda^n(\mathfrak{A}) = \{\varphi \text{ tel que } b\varphi = 0\}. \tag{54.162}$$

When $n = 0$, the condition $b\varphi = 0$ becomes $(b\varphi)(a^0, a^1) = \varphi(a^0 a^1) - \varphi(a^1 a^0) = 0$, in such a way that the elements of Z_λ^0 are exactly the traces on $\mathfrak{A}$.

If $(\Omega, d, \int)$ is a cycle and $\rho : \mathfrak{A} \to \Omega^0$ a homomorphism, then the **character** of ρ is the functional τ defined by

$$\tau(a^0, \cdots, a^n) = \int \rho(a^0) d\big(\rho(a^1)\big) \cdots d\big(\rho(a^n)\big). \tag{54.163}$$

54.12.1 Example: de Rham homology

Let M be a differentiable manifold and C, a closed de Rham q-current on M with $q \leqslant \dim M$. If ω is a form on M, we denote by $C(\omega)$ the evaluation of C on ω. We consider $\Omega^i = C^\infty\big(M, \bigwedge^i T^*M\big)$, the space of differential forms of degree i. It provides a graded differential algebra (Ω, d) on which we can put the trace $\int : \Omega^q \to \mathbb{C}$,

$$\int \omega = C(\omega).$$

54.12.2 Hochschild cohomology

Let $\mathfrak{A}$ be an algebra over $\mathbb{C}$ and $\mathfrak{A}_\mathbb{1} = \mathfrak{A} \oplus \mathbb{C}\mathbb{1}$ be the algebra obtained by adduction of an unity to $\mathfrak{A}$. We define

$$\Omega^1(\mathfrak{A}) = \mathfrak{A}_\mathbb{1} \times_\mathbb{C} \mathfrak{A}$$

and an $\mathfrak{A}$-bimodule structure by

$$x\Big((a + \lambda\mathbb{1}) \times b\Big)y := (xa + \lambda x) \times by - (xa + \lambda xb) \times y$$

with a, b, x, $y \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$. Then we consider

$$\begin{aligned} d\colon \mathfrak{A} &\to \Omega^1(\mathfrak{A}) \\ da &= 1 \otimes a \end{aligned} \tag{54.164}$$

which can be checked to be a derivation.

Proposition 54.61.
Let $\mathcal{E}$ be a $\mathfrak{A}$-bimodule and $\delta\colon \mathfrak{A} \to \mathcal{E}$, a derivation. Then there exists a bimodule morphism $\rho\colon \Omega^1(\mathfrak{A}) \to \mathcal{E}$ such that $\delta = \rho \circ d$.

Proof. No proof. □

This proposition says that $\big(\Omega^1(\mathfrak{A}), d\big)$ is an universal derivation in a $\mathfrak{A}$-bimodule. From $\Omega^1(\mathfrak{A})$, we define $\Omega^0(\mathfrak{A}) = \mathfrak{A}$.

$$\Omega^n(\mathfrak{A}) = \Omega^1(\mathfrak{A}) \otimes_{\mathfrak{A}} \cdots \otimes_{\mathfrak{A}} \Omega^1(\mathfrak{A}),$$

and the differential extends to an unique graded derivation of $\Omega^*(\mathfrak{A})$ such that $d^2 = 0$. Note the isomorphism

$$\begin{aligned} j\colon \mathfrak{A}_{\mathbb{1}} \otimes \mathfrak{A}^{\otimes n} &\to \Omega^n(\mathfrak{A}) \\ (a^0 + \lambda\mathbb{1}) \otimes a^1 \otimes \cdots \otimes a^n &\mapsto a^0 da^1 \cdots da^n + \lambda da^1 \cdots da^n \end{aligned} \tag{54.165}$$

for each $a^j \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$.

Lemma 54.62.
The cohomology of the complex $\big(\Omega^*(\mathfrak{A}), d\big)$ is zero in any dimension, including zero.

Proof. No proof. □

A product in $\Omega^*(\mathfrak{A})$ is defied as usual by juxtaposition and rearrangement using the fact that d is a derivation, see for example the proof of proposition 80.1:

$$\begin{aligned} (a^0 da^1 \cdots da^n)(a^{n+1} da^{n+2} \cdots da^n) = &\sum_{j=1}^{n} (-1)^{n-j} a^0 da^1 \cdots d(a^j a^{j+1}) \cdots da^n da^{n+1} \cdots da^m \\ &+ (-1)^n a^0 a^1 da^2 \cdots da^m. \end{aligned}$$

It is nothing else than define the product in such a way that $\Omega^*(\mathfrak{A})$ is a right $\mathfrak{A}$-module and d a derivation.

Proposition 54.63.
Let τ be a $(n+1)$-linear functional on $\mathfrak{A}$. The three following properties, where $a^0, \cdots a^n$ are some elements of $\mathfrak{A}$, are equivalent

(1) There exists a cycle $(\Omega, d, \int)$ of dimension n and a homomorphism $\rho\colon \mathfrak{A} \to \Omega^0$ such that

$$\tau(a^0, \cdots, a^n) = \int \rho(a^0) d\big(\rho(a^1)\big) d\big(\rho(a^n)\big)$$

for all $a^0, \cdots, a^n \in \mathfrak{A}$. In other words, φ is the character of a cycle.

(2) There exists a closed graded trace $\hat{\tau}$ of dimension n on $\Omega^*(\mathfrak{A})$ such that

$$\tau(a^0, \cdots, a^n) = \hat{\tau}(a^0 da^1, \cdots, da^n)$$

(3) The functional τ satisfies $b\tau = 0$ and $\tau \circ \gamma = \epsilon(\gamma)\tau$ where γ is any cyclic permutation of $\{0, 1, \cdots, n\}$ and $\epsilon(\gamma)$ is its parity, or more explicitly,

$$\tau(a^1, \cdots, a^n, a^0) = (-1)^n \tau(a^0, \cdots, a^n)$$

and

$$\sum_{i=0}^{n} (-1)^n \tau(a^0, \cdots a^i a^{i+1}, \cdots, a^{n+1}) + (-1)^{n+1} \tau(a^{n+1} a^0, \cdots, a^n) = 0.$$

The most compact way to express this condition is just $\tau \in Z^n_\lambda$.

54.12.3 Hochschild groups of cohomology

Definition 54.64.
*Let $\mathfrak{A}^e = \mathfrak{A} \otimes \mathfrak{A}^0$, the tensor product of $\mathfrak{A}$ with its opposite algebra. Let $\mathfrak{M}$ be a module over $\mathfrak{A}$. The **Hochschild cohomology** is defined as follows.*

— *Let $C^n(\mathfrak{A}, \mathfrak{M})$ is the set of n-linear maps from $\mathfrak{A}$ to $\mathfrak{M}$.*

— *A cochain is an element of $C^*(\mathfrak{A}, \mathfrak{M})$ and the differential is given by*

$$\begin{aligned}(bT)(a^1, \cdots, a^{n+1}) = a^1 T(a^2, \cdots, a^{n+1}) + \sum_{i=1}^{n} (-1)^i T(a^1, \cdots, a^i a^{i+1}, \cdots, a^{n+1}) \\ + (-1)^{n+1} T(a^1, \cdots, a^n) a^{n+1}.\end{aligned} \tag{54.166}$$

— *The Hochschild cohomology of $\mathfrak{A}$ with coefficients in $\mathfrak{M}$ is finally defined as the cohomology of the complex $\big(C^*(\mathfrak{A}, \mathfrak{M}), b\big)$*

Let us study a particular case. The space $\mathfrak{A}^*$ of functionals on $\mathfrak{A}$ becomes a $\mathfrak{A}$-bimodule when one defines $(a\varphi b)(c) = \varphi(abc)$ for each a, b, $c \in \mathfrak{A}$.

Let $T \in C^n(\mathfrak{A}, \mathfrak{A}^*)$; this is a n-linear functional from $\mathfrak{A}$ to $\mathfrak{A}^*$ which can be seen as a $(n+1)$-linear function τ_T on $\mathfrak{A}$ by

$$\tau_T(a^0, a^1, \cdots, a^n) = T(a^1, \cdots, a^n)(a^0).$$

If one defines b acting on τ by

$$b(\tau)(a^0, \cdots, a^{n+1}) = \sum_{i=0}^{n} (-1)^i \tau(a^0, \cdots, a^i a^{i+1}, \cdots, a^{n/1}) \tag{54.167}$$

$$+ (-1)^{n+1} \tau(a^{n+1} a^0, \cdots, a^n), \tag{54.168}$$

we have $\tau_{bT} = b\tau_T$.

Let $A \colon C^n(\mathfrak{A}, \mathfrak{A}^*) \to C^n(\mathfrak{A}, \mathfrak{A}^*)$, the linear map defined by

$$(A\varphi) = \sum_{\gamma \in \Gamma} \epsilon(\gamma)(\varphi \circ \gamma)$$

where Γ is the group of cyclic permutations of $\{0, 1, \cdots, n\}$, and $\epsilon(\gamma)$, the parity of γ. The image of A is $C^n_\lambda(\mathfrak{A})$.

Lemma 54.65.
If one defines $b' \colon C^n(\mathfrak{A}, \mathfrak{A}^*) \to C^{n+1}(\mathfrak{A}, \mathfrak{A}^*)$ by

$$(b'\varphi)(a^0, \cdots, a^{n+1}) = \sum_{j=0}^{n} (-1)^j \varphi(a^0, \cdots, a^j a^{j+1}, \cdots, a^n), \tag{54.169}$$

we have $b \circ A = A \circ b'$.

Proof. No proof. □

As corollary,

Corollary 54.66.
The complex $\big(C^n_\lambda(\mathfrak{A}), b\big)$ is a subcomplex of the Hochschild complex.

Proof. No proof. □

54.12.4 Homomorphisms

Let $\mathfrak{A}$ and $\mathcal{B}$ be two algebras. From a homomorphism $\rho\colon \mathfrak{A} \to \mathcal{B}$, one induces a homomorphism of complex $\rho^*\colon C^n_\lambda(\mathcal{B}) \to C^n_\lambda(\mathfrak{A})$ with the definition

$$(\rho^*\varphi)(a^0, \cdots, a^n) = \varphi\big(\rho(a^0), \cdots, \rho(a^n)\big).$$

In order for ρ^* to pass to quotient and define a homomorphism $\rho^*\colon HC^n(\mathcal{B}) \to HV^n(\mathfrak{A})$, we need that, for each φ, a certain η fulfills $\rho^* b\varphi = b'\eta$.

Proposition 54.67.
Let $u \in \mathfrak{A}^\times$, and $\theta\colon \mathfrak{A} \to \mathfrak{A}$ defined by $\theta(x) = uxu^{-1}$. Then the induced map $\theta^*\colon HC^*(\mathfrak{A}) \to HC^*(\mathfrak{A})$ is the identity.

Proof. No proof. □

The so defined map θ is called the **inner automorphism** associated with u.

Lemma 54.68.
Let $\mathfrak{A}$ be an unital algebra for which there exists a homomorphism ρ and an invertible element X in $M_2(\mathfrak{A})$ such that

$$X \begin{pmatrix} a & 0 \\ 0 & \rho(a) \end{pmatrix} X^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \rho(a) \end{pmatrix}$$

for all $a \in \mathfrak{A}$. Then $HC^n(\mathfrak{A}) = 0$ for all n.

Proof. No proof. □

We say that the cycle $(\Omega, d, \int)$ **vanishes** if the algebra Ω^0 fulfills the hypothesis of the latter lemma.

54.12.5 The cup product

In general, the space $\Omega^*(\mathfrak{A} \otimes \mathcal{B})$ is different to the space $\Omega^*(\mathfrak{A}) \otimes \Omega^*(\mathcal{B})$, but the first if a left $\mathfrak{A}$-module and a right $\mathcal{B}$-module and $\Omega^*(\mathfrak{A} \otimes \mathcal{B})$ is the universal algebra with this property. So we have a homomorphism $\pi\colon \Omega^*(\mathfrak{A} \otimes \mathcal{B}) \to \Omega^*(\mathfrak{A}) \otimes \Omega^*(\mathcal{B})$ such that $\pi \circ d_{\mathfrak{A}\otimes\mathcal{B}} = (d_{\mathfrak{A}} \otimes d_{\mathcal{B}}) \circ \pi$.

Let now φ be a $(n+1)$-linear functional on $\mathfrak{A}$. One define the linear functional $\hat{\varphi}$ on $\Omega^n(\mathfrak{A})$ by the formula

$$\hat{\varphi} \circ j\big((a^0 + \lambda\mathbb{1}) \otimes a^1 \otimes \cdots \otimes a^n\big) = \varphi(a^0, a^1, \cdots, a^n). \tag{54.170}$$

This definition can be of course be adapted to $\mathcal{B}$. Now we consider $\varphi \in C^n(\mathfrak{A}, \mathfrak{A}^*)$ and $\psi \in C^m(\mathcal{B}, \mathcal{B}^*)$ and the corresponding hat functions $\hat{\varphi}\colon \Omega^n(\mathfrak{A}) \to \mathfrak{A}^*$ and $\hat{\psi}\colon \Omega^m(\mathcal{B}) \to \mathcal{B}^*$. We can also consider the tensor product $\hat{\varphi} \otimes \hat{\psi}$. The homomorphism $\pi\colon \Omega^*(\mathfrak{A} \otimes \mathcal{B}) \to \Omega^*(\mathfrak{A}) \otimes \Omega^*(\mathcal{B})$ allows us to consider the composition map $(\hat{\varphi} \otimes \hat{\psi}) \circ \pi$ which has to be the hat function of a certain multilinear functional on $\mathfrak{A} \otimes \mathcal{B}$. The latter is denoted by $\varphi \sharp \psi$:

$$\widehat{\varphi \sharp \psi} = (\hat{\varphi} \otimes \hat{\psi}) \circ \pi. \tag{54.171}$$

The map $\varphi \sharp \psi$ and is called the **cup product** of φ and ψ.

An element of $\Omega^l(\mathfrak{A} \otimes \mathcal{B})$ is of the form $\omega_1 \otimes \omega_2 \otimes \cdots \otimes \omega_l$ with $\omega_i \in \Omega^1(\mathfrak{A} \otimes \mathcal{B})$. But an element of the latter is of the form

$$\big((a \otimes b) + \lambda\mathbb{1}\big) \otimes (a' \otimes b')$$

where a, $a' \in \mathfrak{A}$ and b, $b' \in \mathcal{B}$. Notice that taking $b = b' = 1$, one can identify the result to an element of $\Omega^1(\mathfrak{A})$. Hence an element of $\Omega^n(\mathfrak{A}) \otimes \Omega^m(\mathcal{B})$ can be seen as a very special element of $\Omega^{n+m}(\mathfrak{A} \otimes \mathcal{B})$. From this point of view, we can see π as a map

$$\pi\colon \Omega^{n+m}(\mathfrak{A} \otimes \mathcal{B}) \to \Omega^n(\mathfrak{A}) \otimes \Omega^m(\mathcal{B}).$$

Proposition 54.69.
The cup product enjoys the following main properties.

(1) The map $\varphi \otimes \psi \mapsto \varphi \sharp \psi$ is a homomorphism

$$HC^n(\mathfrak{A}) \otimes HC^m(\mathcal{B}) \to HC^{n+m}(\mathfrak{A} \otimes \mathcal{B}).$$

(2) The character of the tensor product of two cycles is the cup product of their character.

Chapter 55

Examples of groups and representations

ChapThoComsGroupes

55.1 Connectedness of some groups

The following is a general result about Lie groups:

LemConnSpecMo

Lemma 55.1.
If G is a Lie group and G_0 is its identity component, the connected components of G are lateral classes of G_0. More specifically, if $x \in G_1$, then $G_1 = xG_0 = G_0x$.

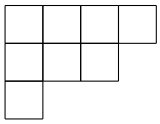
LemOHjzfsL

Lemma 55.2 ([894])**.**
Connectedness of some usual groups:

- *The groups $\mathrm{SU}(p,q)$, $\mathrm{SU}^*(2n)$, $\mathrm{SO}^*(2n)$, $p(n,R)$, and $\mathrm{SP}(p,q)$ are connected.*
- *The group $\mathrm{SO}(p,q)$ $(0 < p < p+q)$ have exactly two connected components.*

55.2 Young tableau and diagrams

We follow [895]. Let $n \in \mathbb{N}$. A **Young diagram** is a diagram of n boxes arranged in rows of decreasing number of boxes. Here is an example with $n = 8$:



A **Young tableau** is a Young diagram in each box of which we placed a number between 1 and n such that

- in each line, the numbers are increasing from left to right,
- in each line, the numbers are increasing from up to down.

It is **standard** if each of the integers $\{1, \dots, n\}$ appears one and only one time. Here is an example:

1	2	3	4	7	8
3	5	6	9		
10					

Let S_n be the symmetric group of all the permutations of n objects. Let us see how does a Young tableau define symmetry properties of tensors. If λ is a Young tableau with n boxes, its applies on the tensor $T_{\mu_1,\dots,\mu_n}$ by symmetrization over the set of indices in each line, and antisymmetrization over the sets of indices in each column. Now, the group S_n can act on the so-obtained tensor in order to produce $n!$ new tensors, which are not independent.

Lemma 55.3.
The space spanned by

$$\{g\tilde{T} \text{ tel que } g \in S_n\} \tag{55.1}$$

where $\tilde{T}$ is λT is an irreducible representation of S_n. Moreover, two representations coming from two different tableau with the same diagram are equivalent.

Proof. No proof. □

As an example, consider the following tableau

1	3
2	4

The corresponding symmetrization of the tensor T_{abcd} is as follows. First, we symmetrize over ac (i.e the row $1, 3$):

$$\frac{1}{2}(T_{abcd} + T_{cbad}), \tag{55.2}$$

and then we symmetrize over bd:

$$\frac{1}{4}(T_{abcd} + T_{cbad} + T_{adcb} + T_{cdab}). \tag{55.3}$$

Then we have to take the ansisymmetrization over ab (8 terms), and we finish by antisymmetrization over cd to get the 16 terms result.

55.2.1 Representations of $\mathbf{GL}(V)$

If $v \in \mathbb{R}^D$, then the action of all the non-singular $D \times D$ matrices on v creates a representation space for $GL(D)$. In the same way, the 2-indices tensors T_{ab} form a D^2-dimensional representation of $GL(V)$ by acting separately on each index:

$$(g \cdot T)_{ab} = g^{ac} g^{bd} T_{cd}. \tag{55.4}$$ EqYoungAgitIndices

That representation is, however, reducible because a symmetric tensor remains symmetric; indeed

$$(g \cdot T)_{ba} = g^{bc} g^{ad} T_{cd} = g^{bc} g^{ad} T_{dc} = g^{bd} g^{ac} T_{cd} = (g \cdot T)_{ab}. \tag{55.5}$$

The same is true for the antisymmetric tensors. Thus,

$$V \otimes V \simeq (V \odot V) \oplus (V \wedge V). \tag{55.6}$$

More generally, when $GL(D)$ acts on $V^{\otimes n}$, the irreducible representations are the different combinations of (anti)symmetrizations, namely: the irreducible representations of S_n.

Let λ be a Young tableau with n boxes. The **Schur module** $V_\lambda^{GL(D)}$ is the vector space of tensor $\tilde{T}$ of rank n such that

— $\tilde{T}$ is completely antisymmetric for all the labels of each column of λ,
— the antisymmetrization of all the labels of one column of λ with another label on the right vanishes.

55.2.2 Representations of $O(p, q)$

Theorem 55.4 ([895])**.**
If the sum of the length of the two first column of a Young diagram λ is strictly bigger than D, then the irreducible representation of $O(D)$ given by λ is identically vanishing.

Young diagrams such that the sum of the lengths of the first two columns does not exceed D are said to be **allowed**.

ThoOpqrepreTens

Theorem 55.5.
Every non vanishing finite dimensional irreducible representation of $O(p, q)$ is isomorphic to a completely traceless tensor representation whose symmetry properties are given by an allowed Young diagram.

55.3 Two words about $\mathfrak{su}(3)$

Using the Cartan matrix of $\mathfrak{su}(3)$ and formula (51.446), we will determine the Dynkin coefficients of the representation $\overset{1}{o}\!\!-\!\!\!-\!\!o$ without even explicitly compute the weights. For that, we follow the construction of [862]. The Cartan matrix is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{55.7}$$

The Dynkin coefficients of the highest weight is given by

$$\Lambda_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{55.8}$$

Since Λ is highest weight, we have $q(\Lambda, \alpha_i) = 1$, so that $\Lambda_1 + q(\Lambda, \alpha_1) = 1$ and $\Lambda_2 + q(\Lambda, \alpha_2) = 0$. Thus the only weight of the first layer is $M = \Lambda - \alpha_1$. Using formula (51.446), we find

$$(\Lambda - \alpha_1)_i = \Lambda_i - A_{1i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{55.9}$$

We also have, by construction, $p(M, \alpha_1) = 1$ and $p(M, \alpha_2) = 0$, so that $M_1 + p(M, \alpha_1) = -1 + 1 = 0$ and $M_2 + p(M, \alpha_2) = 1$. We conclude that $M - \alpha_2$ is a weight, and its Dynkin coefficients are given by

$$(M - \alpha_2)_i = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \tag{55.10}$$

55.4 Representations of $\mathfrak{so}(n)$

55.4.1 Ladder operators

We follow [896], and we restrict ourself to the case where n is even. We will use two abuse of language. The first one is to denote by the same symbol the matrices J_{ij} of $\mathfrak{so}(n)$ and the endomorphisms of V that represent them (and that we are searching for). The second one is to improperly speak about J_{ij}^2 as elements of $\mathfrak{so}(3)$ when they are in fact in the universal covering. If we denote by (ij) the rotation in the plane generated by the directions i and j, the algebra has the block structure (here for $\mathfrak{so}(8)$)

$$(12)\quad \begin{array}{|cc|}\hline (13) & (14) \\ (23) & (24) \\ \hline \end{array} \quad \begin{array}{|cc|}\hline (15) & (16) \\ (25) & (26) \\ \hline \end{array} \quad \begin{array}{|cc|}\hline (17) & (38) \\ (27) & (28) \\ \hline \end{array}$$
$$(34) \quad \begin{array}{|cc|}\hline (35) & (36) \\ (45) & (46) \\ \hline \end{array} \quad \begin{array}{|cc|}\hline (37) & (38) \\ (47) & (48) \\ \hline \end{array} \tag{55.11}$$
$$(56) \quad \begin{array}{|cc|}\hline (57) & (58) \\ (67) & (68) \\ \hline \end{array}$$
$$(78)$$

The elements $A_p = J_{2p-1,2p}$ are simultaneously diagonalisable (because they form an abelian subalgebra). We denote by $|m_1, m_2, \dots\rangle$ the eigenvector for the corresponding eigenvalue

$$A_p|\overline{m}\rangle = m_p|\overline{m}\rangle. \tag{55.12}$$

The remaining four operators in each block can be arranged in the following way (here in the block between A_1 and A_2): subEqLppmmettouteca

$$L_{12}^{++} = -J_{13} + iJ_{14} + iJ_{23} + J_{24} \tag{55.13a}$$
$$L_{12}^{+-} = J_{13} + iJ_{14} - iJ_{23} + J_{24} \tag{55.13b}$$
$$L_{12}^{-+} = J_{13} - iJ_{14} + iJ_{23} + J_{24} \tag{55.13c}$$
$$L_{12}^{--} = -J_{13} - iJ_{14} - iJ_{23} + J_{24}. \tag{55.13d}$$

Using the fact that the J_{ij} are Hermitian, we see that $(L_{pq}^{\pm\pm})^* = L_{pq}^{\mp\mp}$. These operators are build in such a way to have [SubEqCommsLppmmettoutca]

$$[A_p, L_{pq}^{\pm\cdot}] = \pm L_{pq}^{\pm\cdot} \tag{55.14a}$$
$$[A_q, L_{pq}^{\cdot\pm}] = \pm L_{pq}^{\cdot\pm} \tag{55.14b}$$

These relations make that

$$L_{pq}^{\pm\pm}|\overline{m}\rangle \propto |\dots, m_p \pm 1, \dots, m_q \pm 1, \dots\rangle. \tag{55.15}$$

To find the proportionality coefficient is one part of our work now. We denote by j_p the maximal eigenvalue of A_p. Using the matricial expression $J_{ij} = i(E_{ij} - E_{ji})$ and the multiplication formula (51.20), we find

$$J^2 = \sum_{ij} (J_{ij})^2 = (n-1)\mathbb{1} \tag{55.16}$$ [EqJsqmmunmtu]

which is thus a Casimir operator that is represented as a multiple of identity [1].

Since we are searching for finite dimensional representations, each of m_p have minimal and maximal value. To find these values is the second part of our work now. We denote by j_p the maximal eigenvalue of A_p and by

Using the commutation relations and the fact that products like $J_{13}J_{24}$ vanish, a simple computation yields

$$L_{12}^{++}L_{12}^{--} = \sum_{(i,j)\in b_{12}} J_{ij}^2 + 2(A_1 + A_2) \tag{55.17}$$ [EqLppLmmundeux]

where b_{12} denotes the block associated with A_1 and A_2, namely $\sum_{(i,j)\in b_{12}} = J_{13}^2 + J_{23}^2 + J_{14}^2 + J_{24}^2$. In the case of even n, we have $n/2$ elements A_p, each of them generating $(\frac{n}{2} - 1)$ blocks, so that formula (55.17) can be summed up on all the blocks to have

$$\sum_b L_b^{--}L_b^{++} = J^2 + (n-2)\sum_p A_p - \sum_p A_p^2. \tag{55.18}$$

On the other hand, notice that in the definition representation, for each block the matrix $\sum_{(i,j)\in b} J_{ij}^2$ is diagonal, so that its belongs to the center of $\mathfrak{so}(n)$. From, its representative is a multiple of identity and we can in fact simultaneously diagonalise the operators

$$\{A_p, J_b^2, J^2\}. \tag{55.19}$$

If we take J_a^2 in the equation (55.17), and apply that to $|\overline{j}\rangle$ we obtain

$$J_a^2|\overline{j}\rangle = \Big(J^2 - \sum_{b\setminus a} J_b^2 - \sum_p A_p\Big)|\overline{j}\rangle = \Big((n-2)\sum_p j_p - 2\sum_p j_p^2 - \sum_{b\setminus a} j_b^2\Big)|\overline{j}\rangle, \tag{55.20}$$

so that we have the following constrain over the j_b^2:

$$\sum_b j_j^2 = (n-2)\sum_p j_p - 2\sum_p j_p^2. \tag{55.21}$$ [EqCondCompajscsumjdeux]

Once these j_b^2 are fixed, the representation is fixed because

$$\|L_a^{++}|\overline{m}\rangle\|^2 = \langle\overline{m}|L_a^{--}L_a^{++}|\overline{m}\rangle = \langle\overline{m}|\big(J_a^2 + 2(A_{a_1} + A_{a_2})\big) = j_a^2 + 2(m_{a_1} + m_{a_2}), \tag{55.22}$$

Thus we have

$$L_a^{++}|\overline{m}\rangle = \sqrt{j_a^2 + 2(m_{a_1} + m_{a_2})}|\dots, m_{a_1} + 1, \dots, m_{a_2} + 1, \dots\rangle. \tag{55.23}$$ [EqLppketmsqrtjijldots]

We are now going to prove that one can choice the values of j_a^2.

1. Take care that the expression (55.16) is an expression for J^2 in its definition representation. This is not the expression of J^2 acting on $|\overline{m}\rangle$. Indeed we make a systematic abuse between elements J_{ij} in $\mathfrak{so}(n)$ and the operators which represent them on the vector space V generated by the ket $|\overline{m}\rangle$.

Proposition 55.6.
Different choices of j_a^2 give rise to equivalent representations.

¡¡ Avertissement/question au lecteur !! 55.7
Je ne sais pas où est la faute dans la preuve qui suit, mais cette proposition me semble fausse. Doit y avoir un moyen de fixer ces j_a^2.

Proof. Let π_1 and π_2 be two representations of $\mathfrak{so}(n)$ that differ only by the value of J_a (for every a). We have

$$\pi_1(L_a^{++})|\overline{m}\rangle = k_a^{++}|\overline{n}\rangle \tag{55.24a}$$
$$\pi_2(L_a^{++})|\overline{m}\rangle = l_a^{++}|\overline{n}\rangle. \tag{55.24b}$$

Notice that the numbers k and l are not arbitrary: they have to be of the form of the square root of equation (55.23) for some j_a^2 that fulfil compatibility condition (55.21). We are searching for an operator $A\colon V \to V$ such that

$$\pi_1(L_a^{++})A|\overline{m}\rangle = A\pi_2(L_a^{++})|\overline{m}\rangle \tag{55.25a}$$
$$\pi_1(J_p)A|\overline{m}\rangle = A\pi_2(J_p)|\overline{m}\rangle \tag{55.25b}$$

where $\pi_1(J_p) = \pi_2(J_p)$. The numbers $k_a^{\pm\pm}$ and $l_a^{\pm\pm}$ are given by the representations π_i and are non zero. We write A under the form

$$A|\overline{m}\rangle = x(\overline{m})|\overline{m}\rangle, \tag{55.26}$$

so that the second condition is automatically satisfied. We have

$$\pi_1(L_a^{++})A|\overline{m}\rangle = x(\overline{m})\pi_1(L_a^{++})|\overline{m}\rangle = x(\overline{m})k_a^{++}(\overline{m})|\overline{m}_a^{++}\rangle \tag{55.27}$$

where $\overline{m}_a^{++} = (\dots, m_{a_1} + 1, \dots, m_{a_2} + 1, \dots)$. The constrain is

$$A\pi_2(L_a^{++})\overline{m} = Al_a^{++}(\overline{m})|\overline{m}_a^{++}\rangle = l_a^{++}x(\overline{m}_a^{++})|\overline{m}_a^{++}\rangle. \tag{55.28}$$

One can choose one of the $x(\overline{m})$ and then construct the other with the law

$$x(\overline{m}_a^{\pm\pm}) = \frac{k_a^{\pm\pm}(\overline{m})}{l_a^{\pm\pm}(\overline{m})}x(\overline{m}). \tag{55.29}$$

□

One choice of j_a^2 that respect the condition (55.21) is

$$j_a^2 = \frac{n-2}{2}(j_{a_1} + j_{a_2}) - (j_{a_1}^2 + j_{a_2}^2). \tag{55.30}$$

55.4.2 Root spaces and Dynkin diagram for $\mathfrak{so}(2l)$

The Cartan algebra of $\mathfrak{so}(2l)$ is the abelian algebra generated by $\{A_1, \dots, A_l\}$ where $A_p = J_{2p-1,2p}$. The roots are of the form

$$\pm A_p^* \pm A_q^* \tag{55.31}$$

with $p \neq q$. Among of them, the positive roots are

$$\alpha_p = A_p^* \pm A_q^* \tag{55.32}$$

with $p < q$. The positive root $A_p + A_{p+k}$ is not simple, indeed

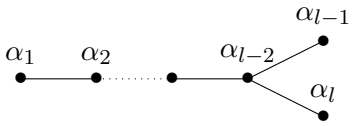
$$\begin{aligned} A_p^* + A_{p+k}^* = {} & (A_p^* - A_{p+1}^*) + (A_{p+1}^* - A_{p+2}^*) + \cdots + (A_{p+k-1}^* - A_{p+k}^*) \\ & + 2(A_{p+k}^* - A_{k-1}^*) + \cdots + 2(A_{l-2}^* - A_{l-1}^*) \\ & + (A_{l-1}^* - A_l^*) + (A_{l-1}^* + A_l^*). \end{aligned} \tag{55.33}$$

The simple roots of $\mathfrak{so}(2l)$ are thus $A^*_p - A^*_{p+1}$ and $A^*_{l-1} + A^*_l$ with $p = 1, \dots, l-1$.

All the length are maximal (and equal to $1/\sqrt{2}$) and all the angles between two "consecutive" simple roots are the same, but the last one:

$$(A^*_{l-1} - A^*_l, A^*_{l-1} + A^*_l) = 0, \tag{55.34}$$

thus the Dynkin diagram of $\mathfrak{so}(2l)$ is given by



55.5 Orthogonal algebra in the odd dimensional case

55.5.1 An odd dimensional example: $\mathfrak{so}(5)$

SubSecsocinq

55.5.1.1 Root spaces

Elements of $\mathfrak{so}(5)$ are arranged in the following picture:

$$\begin{array}{ccc} (12) & \boxed{\begin{array}{cc}(13) & (14)\\ (23) & (24)\end{array}} & \boxed{\begin{array}{c}(15)\\(25)\end{array}} \\ & (34) & \boxed{\begin{array}{c}(35)\\(45)\end{array}} \end{array} \tag{55.35}$$

Because of the fact that J_{12} and J_{34} commute and because of commutators (55.14), the subalgebra

$$\mathfrak{h} = \{J_{12}, J_{34}\} \tag{55.36}$$

is a Cartan subalgebra of $\mathfrak{so}(5)$.

Using the commutators (51.49), and the convention [2] that $X^{\alpha\beta} = 0$ when $\beta \leqslant a$, we find

$$[J_{12}, X^{\alpha\beta}J_{\alpha\beta}] = i(X^{2\beta}J_{1\beta} - X^{1\beta}J_{2\beta}).$$

Thus, if we are searching for vectors X such that $[J_{12}, X] = \lambda X$, we are lead to the equation

$$i(X^{2\beta}J_{1\beta} - X^{1\beta}J_{2\beta}) = \lambda X^{ab}J_{ab}, \tag{55.38}$$

and therms proportional to J_{ab} with a and b different to 1 and 2 have to vanish. There only two possible values for λ, apart of zero, are ± 1. What we find is that $[A_1, X] = X$ implies that $X = \sum_{b=3,4}(J_{2b} + iJ_{1b})$. One finds similar results for A_2. One also finds that $[A_1, X] = 0$ implies that $X^{2\beta} = X^{1\beta} = 0$.

Putting all that together, one finds that EqRootSOImpaircinq

$$\mathfrak{so}(5)_{(\pm 1,0)} = \{J_{25} \pm iJ_{15}\} \tag{55.39a}$$
$$\mathfrak{so}(5)_{(0,\pm 1)} = \{J_{45} \pm iJ_{35}\} \tag{55.39b}$$
$$\mathfrak{so}(5)_{(1,\pm 1)} = \{\mp J_{13} + iJ_{14} \pm iJ_{23} + J_{24}\} \tag{55.39c}$$
$$\mathfrak{so}(5)_{(-1,\pm 1)} = \{\pm J_{13} - iJ_{14} \pm iJ_{23} + J_{24}\}. \tag{55.39d}$$

2. One can take that convention because $J_{\alpha\beta} = -J_{\beta\alpha}$. One other possible convention would has been $X^{\alpha\beta} = -X^{\beta\alpha}$. In that case, the result is

$$[J_{12}, X^{\alpha\beta}J_{\alpha\beta}] = 2i(X^{\beta 1}J_{2\beta} - X^{\beta 2}J_{1\beta}). \tag{55.37}$$

The positivity of the root (a,b) is defined by the positivity of a, or the one of b when a is vanishing. The positive roots of $\mathfrak{so}(5)$ are $(1,0)$, $(1,\pm1)$ and $(0,1)$:

$$\begin{aligned}\alpha_1 &= (1,0), \quad \alpha_2 = (0,1),\\ \alpha_3 &= (1,1), \quad \alpha_4 = (1,-1),\end{aligned} \tag{55.40}$$

among which α_2 and α_4 are simple. Indeed, the root $(0,1)$ is obviously simple. The root $(1,1)$ is not simple, as sum of $(1,0)$ and $(0,1)$; the root $(1,0)$ is the sum of $(1,1)$ and $(1,-1)$ and is thus not simple. We are left with the simple roots $(1,-1)$ and $(0,1)$. So we have

$$\alpha_2 = A_2^*, \tag{55.41a}$$
$$\alpha_4 = A_1^* - A_2^* \tag{55.41b}$$

and the dual basis is

$$h_2 = A_1 + A_2. \tag{55.42a}$$
$$h_4 = A_1 \tag{55.42b}$$

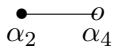
One shows easily compute the inner product (51.108):

$$\big((a,b),(c,d)\big) = \frac{1}{4}(ac+bd). \tag{55.43}$$ EqPridScaJacbdroot

The length of (a,b) is thus $\frac{1}{4}(a^2+b^2)$. The angle between α_1 and α_2 is given by

$$\cos(\theta_{\alpha_1,\alpha_2}) = \frac{(\alpha_1,\alpha_2)}{\sqrt{(\alpha_1,\alpha_2)(\alpha_4,\alpha_4)}} = \frac{\frac{1}{4}(-1)}{\sqrt{\frac{1}{4}\frac{1}{4}2}} = -\frac{1}{\sqrt{2}}. \tag{55.44}$$

Thus we have $\theta_{\alpha_1,\alpha_2} = 135^\circ$, and the Dynkin diagram of $\mathfrak{so}(5)$ is



55.5.1.2 Irreducible representations

We saw around equation (53.364) that an irreducible representation of $\mathfrak{so}(5)$ has to look like $f_1^{i_1}\cdots f_m^{i_m}v\lambda$. In our case, we consider λ, the highest weight and we denote by $|\lambda\rangle$ the cyclic vector which has the following properties

$$L^+_{1(5)}|\lambda\rangle = L^+_{2(5)}|\lambda\rangle = L^{1+}_{12}|\lambda\rangle = 0 \tag{55.45a}$$
$$A_p|\lambda\rangle = \lambda(A_p)|\lambda\rangle. \tag{55.45b}$$

Proposition 53.141 says that a general element of the carrying space W of an irreducible representation of $\mathfrak{so}(5)$ is given by

$$|i_1i_2i_3i_4\rangle = \big(L^-_{1(5)}\big)^{i_1}\big(L^-_{2(5)}\big)^{i_2}\big(L^{--}_{12}\big)^{i_3}\big(L^{-+}_{12}\big)^{i_4}|\lambda\rangle \tag{55.46}$$

In order to know to which root space it belongs, we apply A_p on that [3]. Using the relation (51.143) and the fact that $\alpha_{1(5)}(A_1) = 1$, we find

$$A_1\big(L^-_{1(5)}\big)^{i_1} = -i_1\big(L^-_{1(5)}\big)^{i_1} + \big(L^-_{1(5)}\big)^{i_1}A_1. \tag{55.47}$$

Since $\alpha_{2(5)}(A_1) = 0$, we also have $A_1\big(L^-_{2(5)}\big)^{i_2} = \big(L^-_{2(5)}\big)A_1$. One found the other ones in the same way. The result is

$$A_1|i_1i_2i_3i_4\rangle = \big(\lambda(A_1) - i_1 - i_3 - i_4\big)|i_1i_2i_3i_4\rangle, \tag{55.48}$$

which has to be seen as a straight generalisation of (51.60).

3. In fact, equation (53.365) answers the question. The computation that we give here is more pedestrian, but still is independent of any explicit matricial choices, and is easily generalisable to others $\mathfrak{so}(n)$ algebras.

55.5.1.3 List of irreducible representations

Let us determine the possible irreducible representations of $\mathfrak{so}(5)$ using theorem 51.173. The simple roots are $\alpha_2 = (0,1)$ and $\alpha_4 = (1,-1)$. A simple computation with equation (55.43) and $\Lambda = (a,b)$ gives

$$\begin{aligned} \Lambda_{(1,0)} &= 2b \\ \Lambda_{(1,-1)} &= a - b. \end{aligned} \tag{55.49}$$

The requirement these two to be nonnegative integer produces the list of the irreducible representations:

$$\begin{aligned} b &\in \mathbb{N}/2 \\ a &= b, b+1, \dots \end{aligned} \tag{55.50}$$

EqIrrepsso5

In particular, a and b are both integer or half-integer.

55.5.2 Cartan matrix and Dynkin diagram of $\mathfrak{so}(2l+1)$

The Cartan subalgebra of $\mathfrak{g} = \mathfrak{so}(2l+1)$ is generated by the elements $\mathfrak{h} = \{J_{12}, J_{34}, \dots, J_{2l-1,2l}\}$. We denote them by

$$B_p = J_{2p-1,2p} \tag{55.51}$$

with $p = 1, \dots, l$. As explained in equations (55.39), the roots are $\pm B_p^* \pm B_q^*$ and $\pm B_p^*$ with $p = 1, \dots l$. The root spaces of the roots $\pm B_p^*$ are the combinations $J_{1,2l+1} \pm i J_{2,2l+1}$, $J_{3,2l+1} \pm i J_{4,2l+1}, \dots$

The positive roots are

$$B_p^* \pm B_q^* \quad \text{and} \quad B_p^* \tag{55.52}$$

with $p < q$. In order to determine the simple roots, first remark that $B_p^* + B_q^*$ is the sum of B_p^* and B_q^*. The root B_p^* (with $p \neq l$) is not simple because

$$B_p^* = (B_p^* - B_q^*) + B_q^*. \tag{55.53}$$

If $k > 1$, we have

$$B_p^* - B_{p+k}^* = (B_p^* - B_{p+k-1}^*) + (B_{p+k-1}^* - B_{p+k}^*). \tag{55.54}$$

Thus the simple roots for $\mathfrak{so}(2l+1)$ are

$$\alpha_i = B_i^* - B_{i+1}^* \quad \text{and} \quad \alpha_l = B_l^*, \tag{55.55}$$

EqRacinesSimplesso1

$i = 1, \cdots, l-1$. The inversion of that is easy:

$$B_i^* = \alpha_i + \alpha_{i+1} + \cdots + \alpha_l, \tag{55.56}$$

EqsonBenFnDesAlpha

and in particular, $B_l^* = \alpha_l$. Now we have to determine the basis $\{t_{\alpha_i}\}$ defined by

$$\alpha_i(h) = B(t_{\alpha_i}, h) \tag{55.57}$$

for every $h \in \mathfrak{h}$. In the dual basis of $\{\alpha_i\}$ if $h - h_j \alpha_j^*$, we have $\alpha_i(h) = h_j \delta_{ij} = h_i$. We have

$$h_i = B(t_\alpha, h) = h_j B_{kl} (t_{\alpha_i})_k \underbrace{(\alpha_j^*)_l}_{=\delta_{jl}} = h_j B_{kj} (t_{\alpha_i})_k, \tag{55.58}$$

so we have

$$(t_{\alpha_i})_l = B^{il} \tag{55.59}$$

where B^{ij} are the component of the inverse of the Killing form matrix in the basis $\{\alpha_i^*\}$ of $\mathfrak{h}$. The inner product on $\mathfrak{h}^*$ is thus determined, in function of the Killing form, by

$$B(t_{\alpha_i}, t_{\alpha_j}) = B_{kl} (t_{\alpha_i})_k (t_{\alpha_j})_l = B_{kl} B^{ik} B^{jl} = B^{ji}, \tag{55.60}$$

so that we write

$$(\alpha_i, \alpha_j) = B^{ij}. \tag{55.61}$$

We know that, for a vector space V, if $\xi_i = B_{ij}e_j^*$ is a basis of the dual V^*, then the dual basis of V is given by $v_k = A_{kl}e_l$ with $A = (B^t)^{-1}$. In our case, $\alpha_i = B_i^* - B_{i+1}^*$, so that

$$B = \begin{pmatrix} 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & 1 & -1 & & \\ & & & \ddots & \ddots & \\ & & & & 0 & 1 \end{pmatrix} \tag{55.62}$$

Inverting that matrix, one finds that

$$\alpha_k = B_1^* + \cdots + B_k^*. \tag{55.63}$$

Now, we have to find the matrix of the Killing for in the basis $\{\alpha_i^*\}$. Since $B(B_k, B_l) = -2\delta_{kl}$, we have $B(\alpha_k^*, \alpha_l^*) = -2\min(k,l)$, and the inverse of that matrix is given by

$$-B^i j = \begin{pmatrix} 1 & -\frac{1}{2} & & & & \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & & \\ & -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & & \frac{1}{2} & \end{pmatrix}, \tag{55.64}$$

so that

$$(\alpha_k, \alpha_{k+1}) = \frac{1}{2}, \quad (\alpha_k, \alpha_k) = -1, \quad (\alpha_l, \alpha_l) = -\frac{1}{2} \tag{55.65}$$

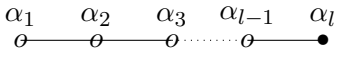
EqProdsoimpairrac

when $k \leqslant l-1$. The Cartan matrix and the Dynkin diagram can now be written down easily:

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2 \end{pmatrix} \tag{55.66}$$

EqCartanSOimpair

One sees form the products (55.65) that all the roots but α_l are of maximal length, so that the Dynkin diagram has $l-1$ withe circles and one black one. Using proposition 51.152, we find that the Dynkin diagram of $\mathfrak{so}(2l+1)$ is



(55.67)

55.5.3 Irreducible representations of $\mathfrak{so}(2, l-1)$

We use the same technique as the one from which we deduced the values (55.50), using theorem 51.173. The simple roots $\{\alpha_k\}$ are given in (55.55) and the inner product is (55.65). We are thus able to compute the numbers Λ_{α_k} for every $\Lambda = \sum_{i=1}^l a_i B_i^*$. Let $\Lambda = \sum_{i=1}^l a_i B_i^*$; taking (55.56) into account, we compute first Λ_{α_j}. With B_k^* $(k < l)$, we have

$$(B_k^*, \alpha_l) = (\alpha_{l-1} + \alpha_l, \alpha_l) = \frac{1}{2} - \frac{1}{2} = 0, \tag{55.68}$$

so we have $\sum_{i=1}^l a_i (B_i^*, \alpha_l) = -\frac{a_l}{2}$, and

$$\Lambda_{\alpha_l} = 2a_l. \tag{55.69}$$

From that, we deduce $a_l \in \mathbb{N}/2$. The computation of Λ_{α_1} is easy too because $(\alpha_k, \alpha_1) \neq 0$ only when $k = 1$ or $k = 2$. So $\Lambda_{\alpha_1} = -2(a_1 B_1^* + a_2 B_2^*, \alpha_1) = a_1 - a_2$. That result generalises to

$$\Lambda_{\alpha_k} = a_k - a_{k+1}. \tag{55.70}$$

Since a_l must belongs to $\mathbb{N}/2$, we deduce that the irreducible representations of $\mathfrak{so}(2l+1)$ are parametrized by

$$a_l \geqslant a_{l-1} \geqslant \ldots \geqslant a_1 \tag{55.71}$$

which are all integer or half integer.

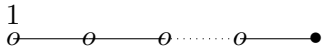
55.5.4 Elementary and basic representations of $\mathfrak{so}(2l+1)$

¡¡ Avertissement/question à la lectrice !! 55.8
Attention : ici j'utilise la formule (51.328), et je crois que je l'utilise avec la transposée de la matrice de Cartan; j'ai peut-être une faute de convention qui invalide la fin de la récurrence.

The Dynkin diagram of $so(2l+1)$ is

EqDynkSOimpair (55.72)

We are looking at the elementary representation

$$\tau_1 = \overset{1}{}$$


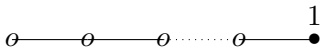
(55.73)

Other basic representations on the same branch are given by

$$\tau_k\left(\tau_1^{\otimes k}\right)_1. \tag{55.74}$$

EqReprBasTausimpair

Notice, however, that the last line in the diagram (55.72) is not of the kind which is allowed to form a branch. Thus the process (55.74) does not produce the representation



(55.75)

which is called the **spinor representation**. Let us now determine the list of weights for the representation τ_1. As showed on the diagram (55.74), the Dynkin coefficients of the highest weight are

$$\Lambda_i = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, \tag{55.76}$$

and we have $q(\Lambda, \alpha_i) = 0$ for every i. Thus $\Lambda_i + q(\Lambda, \alpha_i) \geqslant 1$ only for $i = 1$, and the only weight in the first layer is

$$M^{(1)} = \Lambda - \alpha_1. \tag{55.77}$$

Using formula (51.446) and the first line of the Cartan matrix (55.66), we find

$$M_i^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \tag{55.78}$$

and, by construction,

$$q(M^{(1)}, \alpha_i) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}. \tag{55.79}$$

In order to know what belongs to the second layer, we look at the lucky numbers of $M^{(1)} - \alpha_i$:

$$M_i^{(1)} + q(M^{(1)}, \alpha_i) = \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \tag{55.80}$$

and we deduce that $M^{(2)} = M^{(1)} - \alpha_2$ is the only weight in the second layer. Its Dynkin coefficients are given by

$$M_i^{(2)} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} - \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ \vdots \end{pmatrix}. \tag{55.81}$$

We can now proceed by induction. Let us suppose that

$$M_i^{(k)} = \begin{pmatrix} 0 \\ \vdots \\ -1 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \quad \text{and} \quad q(M^{(j)}, \alpha_i) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{55.82}$$

where in both, the first non zero element is on the kth line. The, we see that the only weight in the next layer is

$$M^{(k+1)} = M^{(k)} - \alpha_{k+1}, \tag{55.83}$$

and, using the corresponding line of the Cartan matrix, we find

$$M_i^{(k+1)} = \begin{pmatrix} 0 \\ \vdots \\ -1 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} - \begin{pmatrix} 0 \\ \vdots \\ -1 \\ 2 \\ -1 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -1 \\ 1 \\ 0 \\ \vdots \end{pmatrix}. \tag{55.84}$$

The induction finishes with

$$M^{(l)} = M^{(l-1)} - \alpha_l \tag{55.85}$$

which as the following Dynkin coefficients:

$$M_i^{(l)} = M_i^{(l-1)} - A_{li} = \begin{pmatrix} \vdots \\ 0 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} \vdots \\ 0 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}. \tag{55.86}$$

Notice that

$$M_i^{(l)} + q(M^{(l)}, \alpha_i) = \begin{pmatrix} \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{55.87}$$

so that

$$M^{(l+1)} = M^{(l)} - \alpha_{l-1} \tag{55.88}$$

is a weight, and its Dynkin coefficients are

$$\begin{pmatrix} \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} \vdots \\ -1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} \vdots \\ 1 \\ -1 \\ 0 \end{pmatrix} \tag{55.89}$$

with

$$q(M^{(l+1)}, \alpha_i) = \begin{pmatrix} \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{55.90}$$

At this point, a new induction can be done up to get

$$M^{(2l-1)} = M^{(2l-2)} - \alpha_1. \tag{55.91}$$

Thus we have

$$M_i^{(2l-1)} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{55.92}$$

and

$$q(M^{(2l-1)}, \alpha_i) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{55.93}$$

We immediately check that there are no $M^{(2l)}$. Thus the Dynkin coefficients of the weights of the representation are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad \begin{pmatrix} \vdots \\ 1 \\ -1 \\ \vdots \end{pmatrix}, \quad \begin{pmatrix} \vdots \\ -1 \\ 1 \\ \vdots \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}. \tag{55.94}$$

We have $l-1$ weights of each of the two middle types. Thus the representations is $2l$-dimensional.

55.6 Explicit choices in $\mathfrak{so}(2, l-1)$

In the case of AdS_4 the matrices are 5×5. We will write them down, but the general form are entirely similar. Our choice of Iwasawa decomposition is

$$\mathcal{N} = \{W_i, V_j, M, L\} \tag{55.95a}$$

$$\mathcal{A} = \{J_1, J_2\}. \tag{55.95b}$$

The basis of $\mathfrak{so}(2, n)$ in which we want to decompose all our elements is the root space one:

$$\mathcal{G} = \mathrm{Span}\{J_1, q_1, X, Y, V, W, M, N, F, L\} \tag{55.96}$$

note in particular that $\mathcal{G}_{(0,0)} = \mathrm{Span}\{J_1, q_1\}$ and $W, J_1 \in \mathcal{H}$.

$$\frac{1}{2}(W-Y) = \begin{pmatrix} & 0 & & & \\ 0 & 0 & 0 & 0 & 1 \\ & 0 & & & \\ & 0 & & & \\ & 1 & & & \end{pmatrix}, \qquad \frac{1}{2}(V+X) = \begin{pmatrix} & & & & 0 \\ & & & & 0 \\ & & & & 1 \\ & & & & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}, \tag{55.97}$$

$$\frac{1}{2}(W+Y) = \begin{pmatrix} & & & & 0 \\ & & & & 0 \\ & & & & 0 \\ & & & & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}, \qquad q_3 = \frac{1}{2}(V-X) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \\ 1 & & & & \end{pmatrix}. \tag{55.98}$$

55.6.1 Decompositions and commutators for $\mathcal{Q}$

First, the root space decomposition of the basis $\{q_i\}$ of $\mathcal{Q}$:

$$q_0 = \frac{1}{4}(M+N+L+F) \qquad q_2 = \frac{1}{4}(N+F-M-L) \tag{55.99a}$$

$$= \frac{1}{4}(X_{11}+X_{1,-1}+X_{-1,1}+X_{-1,-1}) \qquad = \frac{1}{4}(X_{-1,1}+X_{-1,-1}-X_{11}-X_{1,-1}) \tag{55.99b}$$

$$q_1 = q_1 = J_2 \qquad q_3 = \frac{1}{2}(V-X) \tag{55.99c}$$

$$= \frac{1}{2}(X_{01}-X_{0,-1}). \tag{55.99d}$$

The commutators:

$$[q_0,q_1] = \frac{1}{4}(L+F-M-N) \qquad [q_1,q_2] = \frac{1}{4}(L+N-F-M) \tag{55.100a}$$

$$= \frac{1}{4}(X_{1,-1}+X_{-1,-1}-X_{11}-X_{-1,1}) \qquad = \frac{1}{4}(X_{1,-1}+X_{-1,1}-X_{-1,1}-X_{11}) \tag{55.100b}$$

$$[q_0,q_2] = -J_1 \qquad [q_1,q_3] = \frac{1}{2}(V+X) \tag{55.100c}$$

$$[q_0,q_3] = \frac{1}{2}(Y-W) \qquad [q_2,q_3] = \frac{1}{2}(W+Y) \tag{55.100d}$$

55.6.2 Commutators between root spaces and $\mathcal{Q}$

$$[q_0,J_1] = \frac{1}{4}(N+F-M-L) \qquad [q_1,J_1] = 0 \qquad [q_2,J_1] = q_0 \tag{55.101a}$$

$$= \frac{1}{4}(X_{-1,1}+X_{-1,-1}-X_{11}-X_{1,-1}) \tag{55.101b}$$

$$= q_2 \tag{55.101c}$$

$$[q_0,q_1] = \frac{1}{4}(L+F-M-N) \qquad [q_2,q_1] = \frac{1}{4}(F+M-L-N) \tag{55.101d}$$

$$= \frac{1}{4}(X_{1,-1}+X_{-1,-1}-X_{11}-X_{-1,1}) \tag{55.101e}$$

$$[q_0,X] = \frac{1}{2}(W-Y) \qquad [q_1,X] = -X \qquad [q_2,X] = -\frac{1}{2}(W+Y) \tag{55.101f}$$

$$[q_0,Y] = \frac{1}{2}(X-V) \qquad [q_1,Y] = 0 \qquad [q_2,Y] = \frac{1}{2}(V-X) \tag{55.101g}$$

$$[q_0,V] = \frac{1}{2}(Y-W) \qquad [q_1,V] = V \qquad [q_2,V] = \frac{1}{2}(W+Y) \tag{55.101h}$$

$$[q_0,W] = \frac{1}{2}(V-X) \qquad [q_1,W] = 0 \qquad [q_2,W] = \frac{1}{2}(V+X) \tag{55.101i}$$

$$[q_0,M] = q_1+J_1 \qquad [q_1,M] = M \qquad [q_2,M] = q_1+J_1 \tag{55.101j}$$

$$[q_0,N] = q_1-J_1 \qquad [q_1,N] = N \qquad [q_2,N] = -q_1+J_1 \tag{55.101k}$$

$$[q_0,L] = -q_1+J_1 \qquad [q_1,L] = -L \qquad [q_2,L] = -q_1+J_1 \tag{55.101l}$$

$$[q_0,F] = -q_1-J_1 \qquad [q_1,F] = -F \qquad [q_2,F] = q_1+J_1 \tag{55.101m}$$

$$[q_3, J_1] = 0 \qquad [q_3, M] = W \qquad [q_3, V] = -q_1 \tag{55.102a}$$
$$[q_3, q_1] = -\frac{1}{2}(V + X) \qquad [q_3, N] = -Y \qquad [q_3, W] = \frac{1}{2}(M + L) \tag{55.102b}$$
$$[q_3, L] = W \qquad [q_3, X] = -q_1 \tag{55.102c}$$
$$[q_3, F] = -Y \qquad [q_3, Y] = -\frac{1}{2}(N + F) \tag{55.102d}$$

55.6.3 Commutators in the root spaces

$$[J_1, q_1] = 0 \tag{55.103a}$$
$$[J_1, X] = 0 \qquad [q_1, X] = -X \tag{55.103b}$$
$$[J_1, Y] = -Y \qquad [q_1, Y] = 0 \qquad [X, Y] = F \tag{55.103c}$$
$$[J_1, V] = 0 \qquad [q_1, V] = V \qquad [X, V] = 2q_1 \tag{55.103d}$$
$$[J_1, W] = W \qquad [q_1, W] = 0 \qquad [X, W] = -L \tag{55.103e}$$
$$[J_1, M] = M \qquad [q_1, M] = M \qquad [X, M] = -2W \tag{55.103f}$$
$$[J_1, N] = -N \qquad [q_1, N] = N \qquad [X, N] = 2Y \tag{55.103g}$$
$$[J_1, F] = -F \qquad [q_1, F] = -F \qquad [X, F] = 0 \tag{55.103h}$$
$$[J_1, L] = L \qquad [q_1, L] = -L \qquad [X, L] = 0 \tag{55.103i}$$

$$[V, W] = M \tag{55.104a}$$
$$[V, M] = 0 \qquad [W, M] = 0 \tag{55.104b}$$
$$[V, N] = 0 \qquad [W, N] = -2V \qquad [M, N] = 0 \tag{55.104c}$$
$$[V, F] = -2Y \qquad [W, F] = 2X \qquad [M, F] = -4q_1 - 4J_1 \tag{55.104d}$$
$$[V, L] = 2W \qquad [W, L] = 0 \qquad [M, L] = 0 \tag{55.104e}$$

$$[N, F] = 0 \tag{55.105a}$$
$$[N, L] = -4q_1 + 4J_1 \qquad [F, L] = 0 \tag{55.105b}$$

55.6.4 Killing form

The adopted definition is $B(x, y) = \operatorname{Tr}(\operatorname{ad} x \circ \operatorname{ad} y)$ with no one half or such coefficient.

$$\begin{aligned} B(J_1, q_1) &= 0 & B(V, X) &= -12 \\ B(J_1, J_1) &= 6 & B(N, L) &= -24 \\ B(W, Y) &= -12 & B(M, F) &= -24 \end{aligned} \tag{55.106}$$

Some easy computations show that for $g \in \mathrm{SO}(2)$,

$$dL_g q_0 = \begin{pmatrix} -\sin u & \cos u \\ -\cos u & \sin u \end{pmatrix}, \quad dL_g q_1 = \begin{pmatrix} 0 & 0 & \cos u \\ 0 & & -\sin u \\ 1 & & \end{pmatrix}$$

$$dL_g H_1 = \begin{pmatrix} 0 & 0 & \sin u \\ 0 & & \cos u \\ 0 & 1 & \end{pmatrix}$$

$$dR_g J_1 = \begin{pmatrix} 0 & & & & \\ 0 & & 0 & 0 & 1 \\ 0 & & & & \\ -\sin u & \cos u & & & \end{pmatrix}, \quad dR_g J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & & \\ \cos u & \sin u & \end{pmatrix}$$

So

$$dR_g J_1 = -\sin u\, dL_g q_2 + \cos u\, dL_g H_2 \tag{55.107a}$$
$$dR_g J_2 = \sin u\, dL_g H_1 + \cos u\, dL_g q_1. \tag{55.107b}$$

and

$$B_{[g]}(J_1^*, J_1^*) = 6\sin^2 u \tag{55.108a}$$
$$B_{[g]}(J_2^*, J_2^*) = 6\cos^2 u. \tag{55.108b}$$

55.7 Iwasawa decomposition for $\mathfrak{sl}(2,\mathbb{C})$

SecIwasldeuxC

Matrices of $\mathfrak{sl}(2,\mathbb{C})$ are acting on $\mathbb{C}^2$ as

$$\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \begin{pmatrix} a+bi \\ c+di \end{pmatrix} = \begin{pmatrix} (\alpha_1 a - \alpha_2 b + \beta_1 c - \beta_2 d) + i(\alpha_2 a + \alpha_1 b + \beta_2 c + \beta_1 d) \\ (\gamma_1 a - \gamma_2 b - \alpha_1 c + \alpha_2 d) + i(\gamma_2 a + \gamma_1 b - \alpha_2 c - \alpha_1 d) \end{pmatrix}$$

if $\alpha = \alpha_1 + i\alpha_2$. Our aim is to embed $\mathrm{SL}(2,\mathbb{C})$ in $\mathrm{SP}(2,\mathbb{R})$ (see sections 55.8 and 79.27), so that we want a four dimensional realization of $\mathfrak{sl}(2,\mathbb{C})$. It is easy to rewrite the previous action under the form of $\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}$ acting of the vertical four component vector (a,b,c,d). The result is that a general matrix of $\mathfrak{sl}(2,\mathbb{C})$ reads

$$\mathfrak{sl}(2,\mathbb{C}) \rightsquigarrow \begin{pmatrix} \boxed{\begin{matrix} \alpha_1 & -\alpha_2 \\ \alpha_2 & \alpha_1 \end{matrix}} & \begin{matrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{matrix} \\ \begin{matrix} \gamma_1 & -\gamma_2 \\ \gamma_2 & \gamma_1 \end{matrix} & \boxed{\begin{matrix} -\alpha_1 & \alpha_2 \\ -\alpha_2 & -\alpha_1 \end{matrix}} \end{pmatrix}. \tag{55.109}$$

EqConslMatr

The boxes are drawn for visual convenience. Using the Cartan involution $\theta(X) = -X^t$, we find the following Cartan decomposition:

$$\begin{aligned} \mathcal{K}_{\mathfrak{sl}(2,\mathbb{C})} &\rightsquigarrow \begin{pmatrix} \boxed{\begin{matrix} 0 & -\alpha_2 \\ \alpha_2 & 0 \end{matrix}} & \begin{matrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{matrix} \\ \begin{matrix} -\beta_1 & -\beta_2 \\ \beta_2 & -\beta_1 \end{matrix} & \boxed{\begin{matrix} 0 & \alpha_2 \\ -\alpha_2 & 0 \end{matrix}} \end{pmatrix}, \\ \mathcal{P}_{\mathfrak{sl}(2,\mathbb{C})} &\rightsquigarrow \begin{pmatrix} \boxed{\begin{matrix} \alpha_1 & 0 \\ 0 & \alpha_1 \end{matrix}} & \begin{matrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{matrix} \\ \begin{matrix} -\beta_1 & -\beta_2 \\ \beta_2 & -\beta_1 \end{matrix} & \boxed{\begin{matrix} 0 & \alpha_2 \\ -\alpha_2 & 0 \end{matrix}} \end{pmatrix}. \end{aligned} \tag{55.110}$$

We have $\dim \mathcal{P}_{\mathfrak{sl}(2,\mathbb{C})} = 3$ and $\dim \mathcal{P}_{\mathfrak{sl}(2,\mathbb{C})} = 3$. A maximal abelian subalgebra of $\mathcal{P}_{\mathfrak{sl}(2,\mathbb{C})}$ is the one dimensional algebra generated by

$$A_1 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.$$

The corresponding root spaces are

— $\mathfrak{sl}(2,\mathbb{C})_0$:

$$I_1 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & -1 & 0 \end{pmatrix}$$

— $\mathfrak{sl}(2,\mathbb{C})_2$:

$$D_1 = \begin{pmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}, \quad D_2 = \begin{pmatrix} & & 0 & -1 \\ & & 1 & 0 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}$$

— $\mathfrak{sl}(2,\mathbb{C})_{-2}$

$$C_1 = \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 1 & 0 & & \\ 0 & 1 & & \end{pmatrix}, \quad C_2 = \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 0 & -1 & & \\ 1 & 0 & & \end{pmatrix}.$$

It is natural to choose $\mathfrak{sl}(2,\mathbb{C})_2$ as positive root space system. In this case, $\mathcal{N}_{\mathfrak{sl}(2,\mathbb{C})} = \{D_1, D_2\}$, $\mathcal{A}_{\mathfrak{sl}(2,\mathbb{C})} = \{I_1\}$ and the table of $\mathcal{A} \oplus \mathcal{N}$ is

$$[I_1, D_1] = 2D_1 \qquad [D_1, D_2] = 0 \tag{55.111}$$
$$[I_1, D_2] = 2D_2 \tag{55.112}$$

The full table is

$$[I_1, D_1] = 2D_1 \quad [I_2, D_1] = 2D_2 \quad [D_1, D_2] = 0 \tag{55.113}$$
$$[I_1, D_2] = 2D_2 \quad [I_2, D_2] = -2D_1 \quad [D_1, C_1] = I_1 \tag{55.114}$$
$$[I_1, C_1] = -2C_1 \quad [I_2, C_1] = -2C_2 \quad [D_1, C_2] = I_2 \tag{55.115}$$
$$[I_1, C_2] = -2C_2 \quad [I_2, C_2] = 2C_1 \quad [D_2, C_1] = I_2 \tag{55.116}$$
$$[D_2, C_2] = -I_1. \tag{55.117}$$

55.8 Symplectic group

SecSympleGp

55.8.1 Iwasawa decomposition

A simple computation shows that 4×4 matrices subject to $A^t\Omega + \Omega A = 0$ are given by

$$\begin{pmatrix} A & B \\ C & -A^t \end{pmatrix}$$

where A is any 2×2 matrix while B and C are symmetric matrices. Looking at general form (55.109), we see that the operation to invert the two last column and then to invert the two last lines provides a homomorphism $\phi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{sp}(2,\mathbb{R})$. The aim is now to build an Iwasawa decomposition of $\mathfrak{sp}(2,\mathbb{R})$ which "contains" the one of $\mathfrak{sl}(2,\mathbb{C})$.

Using the Cartan involution $\theta(X) = -X^t$, we find the Cartan decomposition

$$\mathcal{K}_{\mathfrak{sp}(2,\mathbb{R})} \rightsquigarrow \begin{pmatrix} A & S \\ -S & A \end{pmatrix}, \qquad \mathcal{P}_{\mathfrak{sp}(2,\mathbb{R})} \rightsquigarrow \begin{pmatrix} S & S' \\ S' & -S \end{pmatrix} \tag{55.118}$$

where S and S' are any symmetric matrices while A is a skew-symmetric one. As far as the dimensions are concerned, we have $\dim \mathcal{K}_{\mathfrak{sp}(2,\mathbb{R})} = 4$ and $\dim \mathcal{P}_{\mathfrak{sp}(2,\mathbb{R})} = 6$. It turns out that $\phi(\mathcal{K}_{\mathfrak{sl}(2,\mathbb{C})}) \subset \mathcal{K}_{\mathfrak{sp}(2,\mathbb{R})}$ and $\phi(\mathcal{P}_{\mathfrak{sl}(2,\mathbb{C})}) \subset \mathcal{P}_{\mathfrak{sp}(2,\mathbb{R})}$. A maximal abelian subalgebra of $\mathcal{P}_{\mathfrak{sp}(2,\mathbb{R})}$ is spanned by the matrices A'_1 and A'_2 listed below and the corresponding root spaces are:

— $\mathfrak{sp}(2,\mathbb{R})_{(0,0)}$:

$$A_1' = \begin{pmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & -1 & 0 \\ & & 0 & -1 \end{pmatrix}, \quad A_2' = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & -1 \\ & & -1 & 0 \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(0,2)}$:

$$X' = \begin{pmatrix} 1 & -1 & & \\ 1 & -1 & & \\ & & -1 & -1 \\ & & 1 & -1 \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(0,-2)}$:

$$V' = \begin{pmatrix} 1 & 1 & & \\ -1 & -1 & & \\ & & -1 & 1 \\ & & -1 & 1 \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(2,0)}$:

$$W' = \begin{pmatrix} & & 1 & 0 \\ & & 0 & -1 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(2,2)}$:

$$L' = \begin{pmatrix} & & 1 & 1 \\ & & 1 & 1 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(2,-2)}$:

$$M' = \begin{pmatrix} & & 1 & -1 \\ & & -1 & 1 \\ 0 & 0 & & \\ 0 & 0 & & \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(-2,0)}$

$$Y' = \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 1 & 0 & & \\ 0 & -1 & & \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(-2,2)}$:

$$N' = \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 1 & -1 & & \\ -1 & 1 & & \end{pmatrix}$$

— $\mathfrak{sp}(2,\mathbb{R})_{(-2,-2)}$:

$$F' = \begin{pmatrix} & & 0 & 0 \\ & & 0 & 0 \\ 1 & 1 & & \\ 1 & 1 & & \end{pmatrix}$$

It is important to notice how do the root spaces of $\mathfrak{sl}(2,\mathbb{C})$ embed:

$$\phi(I_1) = A'_1 \qquad \phi(I_2) = \frac{V' - X'}{2} \tag{55.119}$$

$$\phi(D_1) = \frac{L' - M'}{2} \qquad \phi(D_2) = -W' \tag{55.120}$$

$$\phi(C_1) = \frac{F' - N'}{2} \qquad \phi(C_2) = Y'. \tag{55.121}$$

So $\mathcal{N}_{\mathfrak{sp}(2,\mathbb{R})}$ must at least contain the elements L', M' and W'. We complete the notion of positivity by V'. The Iwasawa algebra reads

$$\begin{aligned}\mathcal{A}_{\mathfrak{sp}(2,\mathbb{R})} &= \{B_1, B_2\}\\ \mathcal{N}_{\mathfrak{sp}(2,\mathbb{R})} &= \{L', M', W', V'\}\end{aligned}$$

with

$$\begin{aligned}[L', V'] &= -4W' & [W', V'] &= -2M'\\ [B'_1, L'] &= 2L' & [B'_2, M'] &= 2M'\\ [B'_1, W'] &= W' & [B'_2, W'] &= W'\\ [B'_1, V'] &= -V' & [B'_2, V'] &= V'\end{aligned}$$

where $B'_1 = \frac{1}{2}(A'_1 + A'_2)$ and $B_2 = \frac{1}{2}(A'_1 - A'_2)$. The generators of $\mathcal{K}_{\mathfrak{sp}(2,\mathbb{R})}$ are

$$K'_t = \begin{pmatrix} & & 1 & 0\\ & & 0 & 1\\ -1 & 0 & & \\ 0 & -1 & & \end{pmatrix} \qquad K'_1 = \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & 0 & 1\\ & & -1 & 0\end{pmatrix}$$

$$K'_2 = \begin{pmatrix} & & 0 & 1\\ & & 1 & 0\\ 0 & -1 & & \\ -1 & 0 & & \end{pmatrix} \qquad K'_3 = \begin{pmatrix} & & 1 & 0\\ & & 0 & -1\\ -1 & 0 & & \\ 0 & 1 & & \end{pmatrix}.$$

Notice that $[K'_t, K'_i] = 0$ for $i = 1, 2, 3$.

55.8.2 Isomorphism

SubSecIsosp

The following provides an isomorphism $\psi\colon\ \mathfrak{so}(2,3) \to \mathfrak{sp}(2,\mathbb{R})$:

$$\begin{aligned}\psi(H_i) &= B'_i & \psi(u) &= K'_t\\ \psi(W) &= W' & \psi(R_1) &= \frac{1}{2}K'_1\\ \psi(M) &= M' & \psi(R_2) &= \frac{1}{2}K'_2\\ \psi(L) &= L' & \psi(R_3) &= \frac{1}{2}K'_3\\ \psi(V) &= \frac{1}{2}V'\end{aligned}$$

where the R_i's are the generators of the $\mathfrak{so}(3)$ part of $\mathcal{K}_{\mathfrak{so}(2,3)}$ satisfying the relations $[R_i, R_j] = \epsilon_{ijk}R_k$. It is now easy to check that the image of the embedding $\phi\colon\ \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{sp}(2,\mathbb{R})$ is exactly $\mathfrak{so}(1,3)$, so that

$$\psi^{-1} \circ \phi\colon\ \mathfrak{sl}(2,\mathbb{C}) \to \mathcal{H} \tag{55.122}$$

is an isomorphism which realises $\mathcal{H}$ as subalgebra of $\mathfrak{sp}(2,\mathbb{R})$. This circumstance will be useful in defining a spin structure on AdS_4.

One can prove that the kernel of the adjoint representation of $\mathrm{SP}(2,\mathbb{R})$ on its Lie algebra is $\pm\mathbb{1}$, in other words, $\mathrm{Ad}(a) = \mathrm{Id}$ if and only if $a = \pm\mathbb{1}$. We define a bijective map $h\colon \mathrm{SO}(2,3) \to \mathrm{SP}(2,\mathbb{R})/\mathbb{Z}_2$ by the requirement that

$$\psi\big(\mathrm{Ad}(g)X\big) = \mathrm{Ad}\big(h(g)\big)\psi(X) \tag{55.123}$$

Eqdefhspsl

for every $X \in \mathfrak{so}(2,3)$. The following is true for all $\psi(X)$:

$$\begin{aligned}\mathrm{Ad}\big(h(gg')\big)\psi(X) &= \psi\Big(\mathrm{Ad}(g)\big(\mathrm{Ad}(g')X\big)\Big)\\ &= \mathrm{Ad}\big(h(g)\big)\psi\big(\mathrm{Ad}(g')X\big)\\ &= \mathrm{Ad}\big(h(g)h(g')\big)\psi(X),\end{aligned}$$

the map h is therefore a homomorphism. If an element $a \in \mathrm{SP}(2,\mathbb{R})$ reads $a = e^{X_A}e^{X_N}e^{X_K}$ in the Iwasawa decomposition, the property $\mathrm{Ad}(a)\psi(X) = \psi\big(\mathrm{Ad}(g)X\big)$ holds for the element $g = e^{\psi^{-1}X_A}e^{\psi^{-1}X_N}e^{\psi^{-1}X_K}$ of $\mathrm{SO}(2,3)$. This shows that h is surjective.

PgSolhpsiSP

55.8.3 Reductive structure on the symplectic group

SubSecRedspT

A lot of structure of $\mathfrak{so}(2,3)$, such as the reductive homogeneous space decomposition as $\mathcal{Q}\oplus\mathcal{H}$, can be immediately transported from $\mathfrak{so}(2,3)$ to $\mathfrak{sp}(2,\mathbb{R})$. Indeed, let $\mathcal{T} = \psi(\mathcal{Q})$ and $\mathcal{I} = \phi\big(\mathfrak{sl}(2,\mathbb{C})\big)$. We have the direct sum decomposition

$$\mathfrak{sp}(2,\mathbb{R}) = \mathcal{T}\oplus\mathcal{I}.$$

Let $X \in \mathcal{T}\cap\mathcal{I}$, then $\psi^{-1}X$ belongs to $\mathcal{Q}\cap\mathcal{H}$ which only contains 0. The fact that ψ is an isomorphism yields that $X = 0$. Since ψ preserves linear independence, a simple dimension counting shows that the sum actually spans the whole space.

Putting $g = h^{-1}(a)$ in the definition (55.123) of h, we find

$$\psi\left(\mathrm{Ad}\left(h^{-1}(a)\right)X\right) = \mathrm{Ad}(a)\psi(X).$$

Considering a path $a(t)$ with $a(0) = e$, we differentiate this expression with respect to t at $t = 0$ we find

$$\mathrm{ad}(dh^{-1}\dot{a})X = d\psi^{-1}\big(\mathrm{ad}(\dot{a})\psi(X)\big) = \mathrm{ad}(d\psi^{-1}\dot{a})(d\psi^{-1}\psi X),$$

but $d\psi = \psi$ because ψ is linear, hence $[dh^{-1}\dot{a}, X] = [\psi^{-1}\dot{a}, X]$ for all $X \in \mathfrak{so}(2,3)$ and $\dot{a} \in \mathfrak{sp}(2,\mathbb{R})$. We deduce that $(dh^{-1})_e = \psi^{-1}$. We define

$$\begin{aligned}\theta_{\mathfrak{sp}} &= \mathrm{Id}\,|_{\mathcal{K}_{\mathfrak{sp}}} \oplus (-\mathrm{Id})|_{\mathcal{P}_{\mathfrak{sp}}}\\ \sigma_{\mathfrak{sp}} &= \mathrm{Id}\,|_{\mathcal{T}} \oplus (-\mathrm{Id})|_{\mathcal{I}}.\end{aligned}$$

We can check that $\psi^{-1}\circ\theta_{\mathfrak{sp}}\circ\psi = \theta$ and $\psi^{-1}\circ\theta_{\mathfrak{sp}}\circ\psi = \theta$. Then it is clear that

$$[\sigma_{\mathfrak{sp}}, \theta_{\mathfrak{sp}}] = 0$$

using the corresponding vanishing commutator in $\mathfrak{so}(2,3)$. We denote $\mathcal{T}_a = dL_a\mathcal{T}$ and the fact that $dp = d\pi\circ dh^{-1} = d\pi\circ\psi^{-1}$ shows that $dp(\mathcal{T}_a)$ is a basis of $T_{p(a)}(G/H)$. So we consider the basis $t_i = \psi(q_i)$ of $\mathcal{T}$ and the corresponding left invariant vector fields $\tilde{t}_i(a) = dL_a t_i$.

55.9 Heisenberg group and algebra

Let V be a symplectic vector space with the symplectic form Ω. The Heisenberg algebra build on V is the vector space

$$\mathscr{H}(V,\Omega) = V\oplus\mathbb{R}E \tag{55.124}$$

endowed with the bracket defined by

(1) $[\mathscr{H}(V,\Omega),E]=0$,

(2) $[v,w]=\Omega(v,w)E$ for every $v,w\in V$.

The first conditions makes E central in $\mathscr{H}$.

The Heisenberg group is, as set, the same as the algebra: $H=V\oplus\mathbb{R}E$ with the product

$$g_1\cdot g_2=g_1+g_2+\frac{1}{2}[g_1,g_2] \tag{55.125}$$ EqProduitHeisenbergGp

where the bracket is the one in the Lie algebra. Direct computations show that this product is associative, the neutral is $(0,0)$ and that the inverse is given by

$$g^{-1}=-g. \tag{55.126}$$

We are now going to prove that the Lie algebra of that group actually is $\mathscr{H}(V,\Omega)$.

55.9.1 The exponential mapping

Let us build the exponential map between the Heisenberg algebra and its group. Let $(x,\tau)\in T_eH$ and consider $g(s)=e^{s(x,\tau)}=\big(v(s),h(s)\big)$. This map is subject to the following three relations:

(1) $g(s)g(t)=g(s+t)$,

(2) $g(0)=0$,

(3) $g'(01)=(x,t)$.

Taking the derivative of the first one with respect to s and taking into account $v'(0)=x$ and $h'(0)=\tau$, we find

$$\frac{d}{ds}\Big[g(s)+g(t)+\frac{1}{2}\big[g(s),g(t)\big]\Big]_{s=0}=\frac{d}{ds}\Big[\big(v(s+t),h(s+t)\big)\Big]_{s=0} \tag{55.127a}$$

$$\frac{d}{ds}\Big[v(s)+v(t),h(s)+h(t)+\frac{1}{2}\Omega\big(v(s),v(t)\big)\Big]_{s=0}=\big(v'(t),h'(t)\big) \tag{55.127b}$$

$$\Big(x,\tau+\frac{1}{2}\Omega\big(x,v(t)\big)\Big)=\big(v'(t),h'(t)\big) \tag{55.127c}$$

We deduce that $v'(t)=x$ and $h'(t)=\tau+\frac{1}{2}\Omega\big(x,v(t)\big)$, so that $v(t)=tx$ and $h(t)=t\tau$. The exponential mapping is thus given by the identity:

$$\exp(x,\tau)=(x,\tau). \tag{55.128}$$

In order to prove that the law (55.125) accepts the Heisenberg algebra as Lie algebra, we need to compute the adjoint action.

$$\begin{aligned}\mathrm{Ad}(e^{t(x,\tau)})(x',\tau')&=\frac{d}{ds}\Big[\,\mathbf{Ad}(e^{t(x,\tau)})e^{s(x',\tau')}\Big]_{s=0}\\&=\frac{d}{ds}\Big[(tx,t\tau)(sx',s\tau')(-tx,-t\tau)\Big]_{s=0}\\&=\frac{d}{ds}\Big[\big(tx+sx',t\tau+s\tau'+\frac{ts}{2}\Omega(x,x')\big)(-tx,t\tau)\Big]_{s=0}\\&=\big(x',\tau'+t\Omega(x,x')\big).\end{aligned} \tag{55.129}$$

Now, the Lie algebra bracket is given by

$$\begin{aligned}\big[(x,\tau),(x',\tau')\big]&=\frac{d}{dt}\Big[\,\mathrm{Ad}(e^{t(x,\tau)})(x',\tau')\Big]_{t=0}\\&=\big(0,\Omega(x,x')\big)\\&=\Omega(x,x')E,\end{aligned} \tag{55.130}$$

which is the bracket of $\mathscr{H}(V,\Omega)$.

55.10 Hermitian conjugate, unitary operators

Lemma 55.9.
Let $\mathscr{H}$ be an Hilbert space. Let A be a linear continuous operator on $\mathscr{H}$. The map

$$\begin{aligned} S_A\colon \mathscr{H}' &\to \mathscr{H}' \\ \alpha &\mapsto \alpha\circ A \end{aligned} \tag{55.131}$$

is well defined and continuous.

Proof. We have to prove two elements: firstly $\alpha\circ A$ must be continuous, so that S_A takes its values in $\mathscr{H}'$, and secondly we want S_A itself to be continuous.

The fact that $\alpha\circ A$ is continuous is simply the fact that α and A are continuous.

Now we compute the norm of S_A; first we have

$$\|S_A\| = \sup_{\alpha\in\mathscr{H}'} \frac{\|S_A(\alpha)\|}{\|\alpha\|}. \tag{55.132}$$

Then we compute

$$\|S_A(\alpha)\| = \sup_{\|u\|=1} \|S_A(\alpha)u\| = \sup_{\|u\|=1} \|\alpha(Au)\| \leqslant \|\alpha\|\|A\|\|u\| = \|\alpha\|\|A\|. \tag{55.133}$$

Then we have

$$\|S_A\| = \sup_{\alpha\in\mathscr{H}'} \frac{\|S_A(\alpha)\|}{\|\alpha\|} \leqslant \sup_{\alpha\in\mathscr{H}'} \frac{\|\alpha\|\|A\|}{\|\alpha\|} = \|A\| < \infty. \tag{55.134}$$

The proposition 10.70 shows that S_A is continuous because it is bounded. □

We will use the map

$$\begin{aligned} \Phi\colon \mathscr{H} &\to \mathscr{H}' \\ y &\mapsto \Phi_y \end{aligned} \tag{55.135}$$

where Φ_y is defined by $\Phi_y(x) = \langle x, y\rangle$. From the Riesz representation theorem 25.18 we know that Φ is a bijective isometry. For the sake of notational convenience we will write $\Phi(u)$ for Φ_u.

Notice the following formula:

$$\langle u, \Phi^{-1}(\alpha)\rangle = \alpha(u) \tag{55.136}$$

EQooHWQPooNeYokT

for every $u\in\mathscr{H}$ and $\alpha\in\mathscr{H}'$.

DEFooAAKCooZJCHPS

Proposition-Definition 55.10 ([1])**.**
Let $\mathscr{H}$ be an Hilbert space over $\mathbb{C}$. We consider a continuous $A\colon \mathscr{H}\to\mathscr{H}$. There exists an unique linear operator $B\colon \mathscr{H}\to\mathscr{H}$ such that

$$\langle Au, v\rangle = \langle u, Bv\rangle \tag{55.137}$$

for every $u, v\in\mathscr{H}$.

*The so-defined operator B is the **hermitian conjugate of** A and is denoted $A^\dagger$. In other words, $A^\dagger$ is defined by the equality*

$$\langle Au, v\rangle = \langle u, A^\dagger v\rangle \tag{55.138}$$

EQooPTUWooPCbNxA

for every $u, v\in\mathscr{H}$.

Proof. In two parts.

(1) **Unicity** We have to prove that, v being given, there exists an unique w such that

$$\langle Au, v\rangle = \langle u, w,\rangle \tag{55.139}$$

EQooVBJXooOtdmlQ

for every u. Be clear: the same w must works for every u. The condition (55.139) can be written as

$$\Phi(v)Au = \Phi(w)u \tag{55.140a}$$
$$S_A\big(\Phi(v)\big)u = \Phi(w)u \tag{55.140b}$$
$$S_A\big(\Phi(v)\big) = \Phi(w), \tag{55.140c}$$

and finally

$$w = \Phi^{-1}\Big(S_A\big(\Phi(v)\big)\Big). \tag{55.141}$$

EQooPRVGooUfIELg

This proves the unicity: Bv must be given by the expression (55.141).

(2) **Existence** We check that the formula

$$A^\dagger = \Phi^{-1} \circ S_A \circ \Phi. \tag{55.142}$$

EQooOIROooXUiCWL

satisfy the properties. Using the formula (55.136) we have:

$$\langle u, (\Phi^{-1} \circ S_A \circ \Phi)(v)\rangle = (S_A \circ \Phi)(v)u \tag{55.143a}$$
$$= \Phi(v)Au \tag{55.143b}$$
$$= \langle Au, v\rangle, \tag{55.143c}$$

so that we see that the operator given by (55.142) makes the work.

□

DEFooOKGXooFCzCHu

Definition 55.11 (Unitary, hermitian)**.**
*An operator $A\colon \mathscr{H} \to \mathscr{H}$ is **unitary** if it satisfies*

$$A^\dagger A = AA^\dagger = \mathrm{Id}\,. \tag{55.144}$$

*An operator $A\colon \mathscr{H} \to \mathscr{H}$ is **hermitian** if it satisfies*

$$A^\dagger = A. \tag{55.145}$$

This was already the definition 8.179.

Lemma 55.12.
An unitary operator is an isometry: it preserves the hermitian product on an Hilbert space.

Proof. Let $u, v \in \mathscr{H}$ and A be an unitary operator on the Hilbert space $\mathscr{H}$. We have

$$\langle Au, Au\rangle = \langle A^\dagger Au, v\rangle = \langle u, v, \rangle. \tag{55.146}$$

□

LEMooJYGRooPTMZwY

Lemma 55.13 ([1])**.**
The hermitian conjugation[4] *satisfies:*

(1) $\langle A^\dagger u, v\rangle = \langle u, Av\rangle$ *for every* $u, v \in \mathscr{H}$
(2) $(A^\dagger)^\dagger = A$.
(3) $(AB)^\dagger = B^\dagger A^\dagger$.

ITEMooULJLooAqnbHI

Proof. We have

$$\langle A^\dagger u, v\rangle = \overline{\langle v, A^\dagger u\rangle} = \overline{\langle Av, u\rangle} = \langle u, Av\rangle. \tag{55.147}$$

This proves the first point.

4. Definition 55.10.

For the second point, $(A^\dagger)^\dagger$ is the unique operator satisfying

$$\langle A^\dagger u, v\rangle = \langle u, (A^\dagger)^\dagger v\rangle \tag{55.148}$$

for every $u, v \in \mathscr{H}$. Using the first point,

$$\langle u, (A^\dagger)^\dagger v\rangle = \langle u, Av\rangle. \tag{55.149}$$

This shows that $\Phi\big((A^\dagger)^\dagger v\big) = \Phi(Av)$. Since Φ is bijective, $(A^\dagger)^\dagger v = Av$ for every $v \in \mathscr{H}$.

For (3), for every x, y we have

$$\langle (AB)x, y\rangle = \langle Bx, A^\dagger y\rangle = \langle x, B^\dagger A^\dagger y\rangle. \tag{55.150}$$

□

LEMooZILJooFfDEAF

Lemma 55.14 ([1])**.**
The map

$$\begin{aligned} f\colon \mathrm{GL}(n,\mathbb{C}) &\to \mathrm{GL}(n,\mathbb{C}) \\ U &\mapsto U^\dagger \end{aligned} \tag{55.151}$$

is continuous.

Proof. Let $U \in \mathrm{GL}(n,\mathbb{C})$ and a sequence $U_k \xrightarrow{\mathrm{GL}(n,\mathbb{C})} U$. We want to prove that $U_k^\dagger \xrightarrow{\mathrm{GL}(n,\mathbb{C})} U^\dagger$. Let $x, y \in \mathbb{C}^n$; from the definition 55.10 we have

$$\langle (U_k^\dagger - U^\dagger)(x), y\rangle = \langle x, (U_k - U)y\rangle. \tag{55.152}$$

Since $U_k \to U$, we have

$$\langle (U_k^\dagger - U^\dagger)(x), y\rangle \xrightarrow{\mathbb{R}} 0. \tag{55.153}$$

Since the scalar product is non degenerate, this implies $U_k^\dagger - U^\dagger \xrightarrow{\mathrm{GL}(n,\mathbb{C})} 0$. Notice that we used the fact that $(U_k - U)^\dagger = U_k^\dagger - U^\dagger$. □

55.11 The group GL(V, W) of invertible operators

The group $\mathrm{GL}(V,W)$ is defined in 10.272. This is the group of isomorphisms between V and W. Here, $\mathrm{GL}(n,\mathbb{R})$ is a shortcut for $\mathrm{GL}(\mathbb{R}^n,\mathbb{R}^n)$.

PROPooQJANooMaqLvV

Proposition 55.15 ([1])**.**
Let V be a finite dimensional vector space.

(1) The group $\mathrm{GL}(V)$ is a topological manifold with the unique map

$$\begin{aligned} \varphi\colon U &\to \mathrm{GL}(V) \\ a &\mapsto a \end{aligned} \tag{55.154}$$

where[5]

$$U = \{a \text{ invertible in } \mathrm{End}(V)\}. \tag{55.155}$$

(2) The topological manifold $\mathrm{GL}(V)$ is smooth.

(3) The smooth manifold $\mathrm{GL}(V)$ is analytic.

Proof. We check the conditions of theorem 49.10. Here, Λ contains only one element and the unique chart is the identity. All Simple.

(1) **Hypothesis (1)** The part U is open by proposition 10.76.

5. The part U is $\mathrm{GL}(V)$ itself, but I prefer to see it as a part of $\mathrm{End}(V)$. The difference is that $\mathrm{End}(V)$ is a vector space while $\mathrm{GL}(V)$ is not.

(2) **Hypothesis (2)** The identity is a bijection between U and $\mathrm{GL}(V)$ because $U = \mathrm{GL}(V)$.

(3) **Hypothesis (3)** The identity is a bijection between U and $\mathrm{GL}(V)$ because $U = \mathrm{GL}(V)$.

(4) **Hypothesis (4)** Since φ is the identity, we have

$$\varphi^{-1}\Big(\varphi\big(\mathrm{End}(V)\big) \cap \varphi\big(\mathrm{End}(V)\big)\Big) = \mathrm{End}(V), \tag{55.156}$$

is open in $\mathrm{End}(V)$.

(5) **Hypothesis (5)** The identity on a finite dimensional normed vector space is continuous.

Now $\mathrm{GL}(V)$ is a topological manifold.

Following the definition 49.18, it is a smooth manifold since the unique chart is the identity which is smooth. Since the identity is analytic, the manifold is analytic. □

PROPooPZABooXxQkFi

Proposition 55.16.
The group $\mathrm{GL}(n,\mathbb{R})$ *is an analytic Lie group.*

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

For the proof, it will be sufficient to prove that $\mathrm{GL}(n,\mathbb{R})$ is a smooth Lie group because of theorem 53.66.

PROPooWRVKooLfqLfV

Proposition 55.17.
The group $\mathrm{GL}(n,\mathbb{C})$ *is an analytic Lie group.*

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

In the remaining we will always consider analytic charts on $\mathrm{GL}(n,\mathbb{C})$.

PROPooRRNXooXJZYSB

Proposition 55.18 ([1])**.**
Let V and W be finite dimensional vector spaces. Let S be open in the manifold M. We have $A \in C^p\big(S, \mathrm{End}(V,W)\big)$ if an only if $A_{ij} \in C^p(S,\mathbb{R})$ for every i,j.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooLANVooKPiLuu

Proposition 55.19 ([1])**.**
Let S be open in the manifold M. If $A \in C^p\big(S, \mathrm{GL}(\mathbb{R},k)\big)$ and $B \in C^p\big(S, \mathrm{GL}(\mathbb{R},l)\big)$, then[6]

$$A \otimes B \in C^p\big(S, \mathrm{GL}(\mathbb{R}^k \otimes \mathbb{R}^l)\big). \tag{55.157}$$

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

55.12 The group $\mathrm{SU}(n)$ of special unitary operators

LEMooKFQOooBVtyoW

Lemma 55.20 ([1])**.**
The part $\mathrm{SU}(n)$ *is closed in* $\mathrm{GL}(n,\mathbb{C})$.

Proof. The definition of the hermitian adjoint is in the proposition 55.10. The map $U \to U^\dagger$ is continuous by lemma 55.14. Since $(U,U^\dagger) \mapsto UU^\dagger$ is continuous too, the set of operators satisfying $UU^\dagger \neq \mathbb{1}$ is open. By complementarity, $\mathrm{SU}(n)$ is closed. □

6. Definition 10.198 for the tensor product of linear maps.

55.12.1 Some settings

We particularize ourself to the finite dimensional Hilbert space $\mathscr{H} = \mathbb{C}^n$. The vector space $\mathbb{C}^n$ has a canonical basis; so we use it, and the linear maps $\mathbb{C}^n \to \mathbb{C}^n$ are identified with their matrices in that very basis. Using the conventions described in the proposition 4.71 we have

$$A(e_j) = \sum_i A(e_j)_i e_i = \sum_i A_{ij} e_i \tag{55.158}$$

where $\{e_i\}_{i=1,\dots,n}$ is the canonical basis of $\mathbb{C}^n$.

When $u \in \mathbb{C}^n$, we denote by $u_k \in \mathbb{C}$ the k^{e} component of u with respect to the canonical basis. The vector space $\mathbb{C}^n$ has an hermitian product [7] given by

$$\langle u, v \rangle = \sum_{k=1}^{n} u_k \bar{v}_k. \tag{55.159}$$

LEMooKEUZooUjQVmp

Lemma 55.21.
Let A be a linear continuous operator on the Hilbert space $\mathscr{H}$. We have

$$\det(A^\dagger) = \overline{\det(A)} \tag{55.160}$$

where the bar stands for the complex conjugate.

If A is unitary, then $\det(A) \in S^1$ where S^1 is the set of elements of norm 1 in $\mathbb{C}$.

Proof. We use for the determinant the formula given by lemma 10.5:

$$\det(A^\dagger) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \langle e_{\sigma(i)}, A^\dagger e_i \rangle = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \langle A e_{\sigma(i)}, e_i \rangle = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \overline{\langle e_i, A e_{\sigma(i)} \rangle}. \tag{55.161}$$

At this point we re-index the product and we make a change of variable $\sigma \to \sigma^{-1}$ for the sum (as in the proof for the transposed matrix, lemma 4.81).

If the operator is unitary,

$$1 = \det(AA^\dagger) = \det(A)\det(A^\dagger) = |\det(A)|^2. \tag{55.162}$$

Thus $|\det(A)| = 1$. □

Recall: the hermitian conjugate is definition 55.10.

Lemma 55.22.
As for the matrices,

$$(U^\dagger)_{ij} = \overline{U_{ji}}. \tag{55.163}$$

Proof. The matrix convention are summarized in -1.1. We write the definition (55.138) for the basis vectors:

$$\langle U e_i, e_j \rangle = \langle e_i, U^\dagger e_j \rangle. \tag{55.164}$$

This means $U_{ji} = \overline{\langle U^\dagger e_j, e_i \rangle} = \overline{U^\dagger_{ij}}$. □

DEFooVIQUooQbnYMu

Definition 55.23 ([897]).
The group of the unitary operators with determinant 1 is $\mathrm{SU}(n)$. More explicitly,

$$\mathrm{SU}(n) = \{U \in \mathrm{GL}(n, \mathbb{C}) \text{ tel que } U^\dagger U = \mathrm{Id}, \det(U) = 1\}. \tag{55.165}$$

PROPooYXPRooBgikdE

Proposition 55.24.
A unitary endomorphism of $\mathbb{C}^n$ is diagonalizable by a unitary operator.

7. Definition 8.176.

Proof. A unitary operator U is normal (definition 12.101), so that the spectral theorem 12.102 provides an unitary matrice V such that $V^\dagger UV$ is diagonal. □

PROPooZBJSooEIguXR

Proposition 55.25.
A special unitary matrix is the exponential of a skew-hermitian matrix with vanishing trace.

Proof. Let $U \in \mathrm{SU}(n)$. We prove that here exists an hermitian operator H with $U = e^{iH}$ and $\mathrm{Tr}(H) = 0$. Then iH is the requested skew-hermitian operator.

By mean of the proposition 55.24, we diagonalize it with the unitary operator S. We have

$$U = SDS^\dagger, \tag{55.166}$$

and

$$\det(U) = 1, \tag{55.167}$$

and

$$D = \begin{pmatrix} e^{i\varphi_1} & & \\ & \ddots & \\ & & e^{i\varphi_n} \end{pmatrix}. \tag{55.168}$$

Since $U \in \mathrm{SU}(n)$ we have

$$1 = \det(U) = \det(S)\det(D)\det(S^\dagger) = |\det(S)|^2 \det(D) = \det(D). \tag{55.169}$$

Thus $\det(D) = 1$ and we deduce

$$\sum_{i=1}^{n} \varphi_i = 2k\pi \tag{55.170}$$

for some $k \in \mathbb{Z}$. We consider the two following operators:

$$\Delta = \begin{pmatrix} \varphi_1 & & \\ & \ddots & \\ & & \varphi_n \end{pmatrix}, \tag{55.171}$$

and $H = S\Delta S^\dagger$. Since $\Delta^\dagger = \Delta$, we have $H^\dagger = H$, so that H is hermitian. We also have $U = e^{iH}$. Indeed, since $(S\Delta S^\dagger)^k = S\Delta^k S^\dagger$ we have

$$e^{iH} = \sum_{k=0}^{\infty} \frac{(iH)^k}{k!} = \sum_k \frac{i^k}{k!}(S\Delta S^\dagger)^k = \sum_k \frac{i^k}{k!} S\Delta^k S^\dagger = Se^{i\Delta}S^\dagger = SDS^\dagger = U. \tag{55.172}$$

We also have $\mathrm{Tr}(H) = 2k\pi$.

This means that H is almost the operator we are searching for. It is easy to modify H in order to get our answer. We set

$$\Delta' = \begin{pmatrix} \varphi_1 & & \\ & \ddots & \\ & & \varphi_n - 2k\pi \end{pmatrix}. \tag{55.173}$$

This matrix satisfies $e^{i\Delta'} = e^{i\Delta}$ and $\mathrm{Tr}(\Delta') = 0$. If we set $H' = S\Delta' S^\dagger$ we still have $\mathrm{Tr}(H') = \mathrm{Tr}(\Delta') = 0$ because of the cyclic invariance of the trace. And finally the operator H' satisfies

$$e^{iH'} = Se^{i\Delta'}S^\dagger = e^{iH} = U. \tag{55.174}$$

□

55.12.2 The center of SU(n)

PROPooLMGHooKrKpsa

Proposition 55.26.
The center[8] *of the group* SU(n), $\mathcal{Z}\big(\mathrm{SU}(n)\big)$, *is the subgroup of the element of the form*

$$e^{i\varphi}\,\mathrm{Id}\,. \tag{55.175}$$

In particular, the center of SU(2) *is* $\{\mathrm{Id}, -\mathrm{Id}\}$.

Proof. Let $Z \in \mathcal{Z}\big(\mathrm{SU}(n)\big)$. It can be diagonalized by an unitary matrix S:

$$SZS^\dagger = D. \tag{55.176}$$

The operator S belongs to U(n) while Z only commutes with the elements of SU(n). Thus we cannot immediately deduce $Z = D$. However, the operator $S_0 = S/\det(S)$ diagonalizes Z as well as S: $S_0 Z S_0^\dagger = D$. But since S_0 is in SU(n) we can deduce $Z = D$.

Long story short, the elements of $\mathcal{Z}\big(\mathrm{SU}(n)\big)$ are diagonal. Let $Ze_i = \lambda_i e_i$. We consider the operator A of SU(n) given by the formula

$$Ae_1 = -e_2 \tag{55.177a}$$
$$Ae_2 = e_1 \tag{55.177b}$$
$$Ae_k = e_k \qquad \text{otherwise.} \tag{55.177c}$$

This is the element of SU(n) which permutes e_1 with e_2 (with a sign for the sake of the determinant). Since $ZA = AZ$, we have

$$AZe_1 = \lambda_1 Ae_1 = -\lambda_1 e_1 \tag{55.178}$$

and

$$ZAe_1 = Z(-)e_2 = -\lambda_2 e_2. \tag{55.179}$$

So we have $\lambda_1 = \lambda_2$. The same being true not only for 1 and 2 but for the other ones, we know that the numbers λ_i are all equal.

So far, $Z = \lambda\,\mathrm{Id}$. The value of λ is not arbitrary: we must impose $Z \in \mathrm{SU}(n)$. The determinant condition provides

EQooZZPJooCeKPDD

$$1 = \det(Z) = \lambda^n \tag{55.180}$$

and the unitary condition imposes

$$1 = \langle e_i, e_i\rangle = \langle Ze_i, Ze_i\rangle = |\lambda|^2, \tag{55.181}$$

so that $|\lambda|^2 = 1$. This conditions imposes $\lambda = e^{i\varphi}$ for some $\varphi \in \mathbb{R}$ while the condition (55.180) furnishes

$$e^{in\varphi} = 1. \tag{55.182}$$

This shows the existence of $k \in \mathbb{Z}$ such that $in\varphi = 2ki\pi$ and finally

$$\varphi = \frac{2k\pi}{n}. \tag{55.183}$$

In the case $n = 2$ we have $\varphi = k\pi$ and $e^{i\varphi} = \pm 1$. □

55.12.3 Lie group

PROPooPXRJooLNnMFn

Proposition 55.27 (Lie group structure for SU(n)[1])**.**
There exists a unique analytic manifold structure on SU(n) *such that :*

ITEMooHZQRooFGCVjP

(1) The group SU(n) *is Lie subgroup*[9] *of* GL(n, ℂ).

8. The set of elements which commute with all the elements, see definition 1.170.
9. Definition 52.24.

(2) There exists an atlas in which $\mathrm{SU}(n)$ *is an analytic Lie group.*

(3) With such an atlas, the inclusion map $\iota\colon \mathrm{SU}(n) \to \mathrm{GL}(n,\mathbb{C})$ *is analytic.*

Proof. Since $\mathrm{SU}(n)$ is closed in $\mathrm{GL}(n,\mathbb{C})$ (lemma 55.20), the Cartan theorem 53.64 provides us a unique smooth manifold structure on $\mathrm{SU}(n)$ such that the inclusion map $\iota\colon \mathrm{SU}(n) \to \mathrm{GL}(n,\mathbb{C})$ is a smooth embedding. The Cartan theorem adds that $\mathrm{SU}(n)$ is then a Lie subgroup of $\mathrm{GL}(n,\mathbb{C})$

Theorem 53.66 allows us to choose an atlas for $\mathrm{SU}(n)$ in which it becomes analytic.

¡¡ Avertissement/question à la lectrice !! 55.28
Thus proof is not finished: lacks of the analytic statement. I may have to prove an analytic version of the Cartan theorem.

□

55.13 Representations of $U(1)$

Proposition 55.29.
Let V *be a complex vector space of dimension* 1*. Let* $\xi \neq 0 \in V$*. If* $\langle .,. \rangle$ *is an hermitian product on* V*, there exists* $m \in \mathbb{R}$ *such that*

$$\langle z_1\xi, z_2\xi \rangle = m z_1 \bar{z}_2. \tag{55.184}$$

Proof. Every element of V can be written under the form $z\xi$ for some $z \in \mathbb{C}$. The properties of an hermitian product say

$$\langle z_1\xi, z_2\xi \rangle = z_1\bar{z}_2\langle \xi, \xi \rangle \tag{55.185}$$

and $\langle \xi, \xi \rangle \in \mathbb{R}$. □

55.30.
What we write S^1 and $\mathrm{U}(1)$ are the same thing. When we write S^1 we have in mind the geometrical object (with a measure) made of the complex numbers of norm 1; when we write $\mathrm{U}(1)$ we have in mind the group structure. But it's the same.

However the generalizations S^2 and $\mathrm{U}(2)$ are not the same.

55.31 ([1])**.**
We are searching now for the irreducible representations of $U(1)$. More precisely we will determine the irreducible continuous representations of $U(1)$. Here the fact to be «continuous» means that $\rho\colon U(1) \to \mathrm{GL}(V)$ is continuous; in particular, V has to be a topological vector space.

This is not a restriction because the theorem 50.12 shows that every irreducible representation of $U(1)$ has dimension 1.

PROPooLWWEooUmqbRA

Proposition 55.32 (Irreducible representations of $U(1)$)**.**
Let $m \in \mathbb{Z}$*. We consider*

$$\begin{aligned} T_m\colon U(1) &\to \mathrm{GL}(\mathbb{C}) \\ T_m(g)z &= g^m z. \end{aligned} \tag{55.186}$$

EQooXPXKooJasMxY

(1) The formula (55.186) defines a representation of $U(1)$*.*

(2) The representation T_m *is irreducible.*

(3) The representation T_m *is continuous.*

(4) If $m \neq l$*, then the representations* T_m *and* T_l *are not equivalent.*

ITEMooUPVQooQddQOJ

(5) Every continuous[10] *irreducible representation of* $U(1)$ *is equivalent to one of them.*

Proof. Several points.

(1) **It is a representation** Since $U(1)$ is abelian, $(g_1g_2)^m = g_1^m g_2^m$.

10. For the norm of proposition 10.53 on V and the corresponding operator norm one on $\mathrm{GL}(V)$, definition 10.55.

(2) **Irreducible** The representation T_m is irreducible because the vector space is $\mathbb{C}$ which has dimension 1.

(3) **Continuous** Let $g_k \xrightarrow{U(1)} g$. We have

$$\|T_m(g_k) - T_m(g)\| = \sup_{|z|1=1} |g_k^m - g^m z| = |g_k^m - g^m| \to 0. \tag{55.187}$$

This shows that T_m is continuous.

(4) **Non equivalence** Let $\psi\colon \mathbb{C} \to \mathbb{C}$ be a linear map such that $T_m(g) \circ \psi = \psi \circ T_l(g)$ for every $g \in U(1)$. This implies

$$g^m \psi(z) = \psi(g^l z), \tag{55.188}$$

hence

$$g^m \psi(z) = g^l \psi(z). \tag{55.189}$$

Taking z such that $\psi(z) \neq 0$ we have $g^m = g^l$ for every $g \in U(1)$, or $g^{l-m} = 1$. Writing $g = e^{ix}$ we have

$$e^{i(l-m)x=1} \tag{55.190}$$

for every $x \in \mathbb{R}$. We conclude $l - m = 2k\pi$. Since $l - m \in \mathbb{Z}$ the only solution is $l - m = 0$.

At this point, it "remains" to prove the point (5). Let (ρ, V) be a continuous irreducible representation of $U(1)$.

(1) **The function λ** Since $U(1)$ is abelian, $\dim(V) = 1$ (theorem 50.12). So there exist a function $\lambda\colon U(1) \to \mathbb{C}$ such that

$$\rho(g) = \lambda(g)\,\mathrm{Id}\,. \tag{55.191}$$

(2) **λ is continuous**

The spaces V, $\mathrm{GL}(V)$ and $\mathbb{C}$ are metric, thanks to the restriction we imposed on the topology of V. Let $g_k \to g$ in $U(1)$. Since ρ is continuous we have $\rho(g_k) \xrightarrow{\mathrm{GL}(V)} \rho(g)$. From the definition of the operator norm, that implies, for each $v \in V$ that

$$\rho(g_k) \xrightarrow{V} \rho(g)v, \tag{55.192}$$

which means

$$\lambda(g_k)v \xrightarrow{V} \lambda(g)v. \tag{55.193}$$

EQooTPAXooAPSgxP

Using the definition of the topology on V,

$$\|\lambda(g_k)v - \lambda(g)v\| = \|\big(\lambda(g_k) - \lambda(g)\big)v\| = \|\lambda(g_k) - \lambda(g)\|\|v\|. \tag{55.194}$$

The convergence (55.193) means

$$|\lambda(g_k) - \lambda(g)|\|v\| \xrightarrow{\mathbb{R}} 0, \tag{55.195}$$

which implies the convergence $\lambda(g_k) \to \lambda(g)$, hence the continuity of λ.

(3) **λ takes values in S^1**

We show that $|\lambda(g)| = 1$ pour tout $g \in U(1)$. An element of $U(1)$ reads $g = e^{2\pi i x}$ with $x \in \mathbb{R}$ [11].

If $x \in \mathbb{Z}$ we have $g = 1$, so that $\lambda(g) = 1$. If $x = 1/n$ $(n \in \mathbb{Z})$ we have $g^n = 1$, but

$$\lambda(g^n) = \lambda(g)^n, \tag{55.196}$$

so that $|\lambda(g)|^n = 1$ which proves that $|\lambda(g)| = 1$.

From here we know that $|\lambda(e^{2\pi i q})| = 1$ for every $q \in \mathbb{Q}$.

Since λ is continuous, the function $x \mapsto |\lambda(e^{2\pi i x})|$ is continuous. A continuous function whose value is 1 over $\mathbb{Q}$ is constant.

11. Proposition 18.61(1).

(4) **Functional equation** For $x, y \in \mathbb{R}$ we have $\rho(e^{ix})\rho(e^{it}) = \rho(e^{ix}e^{iy}) = \rho(e^{i(x+y)})$. We introduce the notation

$$\alpha = \lambda \circ \varphi \tag{55.197}$$

where $\varphi\colon \mathbb{R} \to U(1)$ is $\varphi(x) = e^{ix}$. The function

$$\begin{aligned} \alpha\colon \mathbb{R} &\to S^1 \\ x &\mapsto \lambda(e^{ix}) \end{aligned} \tag{55.198}$$

satisfies

$$\begin{cases} \alpha(0) = 1 & \text{(55.199a)} \\ \alpha(x+y) = \alpha(x)\alpha(y). & \text{(55.199b)} \end{cases}$$

The proposition 12.437 shows that there exists $m \in \mathbb{R}$ such that

$$\alpha(x) = e^{imx}. \tag{55.200}$$

Since $\alpha(2\pi) = \alpha(0) = 1$ we have $e^{2\pi i m} = 1$ and we conclude that $m \in \mathbb{Z}$.

(5) **The value of λ**
We have defined $\alpha = \lambda \circ \varphi$. Since φ is not a bijection, we cannot write $\lambda = \alpha \circ \varphi^{-1}$. However for every $x \in \mathbb{R}$ we have

$$\lambda(e^{ix}) = \alpha(x) = e^{imx} = (e^{ix})^m. \tag{55.201}$$

So

$$\lambda(g) = g^m. \tag{55.202}$$

(6) **Equivalence**
We prove that ρ is equivalent to the representations T_m. Let $\{v\}$ be a basis of V; we consider the linear map

$$\begin{aligned} \psi\colon V &\to \mathbb{C} \\ sv &\mapsto s. \end{aligned} \tag{55.203}$$

We have

$$\big(T_m(g) \circ \psi\big)(sv) = T_m(g)s = g^m s \tag{55.204}$$

while

$$\big(\psi \circ \rho(g)\big)(sv) = \psi\big(g^m sv\big) = g^m s\psi(v) = g^m s. \tag{55.205}$$

So we have

$$T_m(g) \circ \psi = \psi \circ \rho(g) \tag{55.206}$$

which proves that T_m and ρ are equivalent.

□

55.14 The group $SU(2)$

The group SU(2) is already defined in 55.23.

PROPooZMPLooUFyAPW

Proposition 55.33 ([897, 898])**.**
The matrices of SU(2) *are*

$$\mathrm{SU}(2) = \{\begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \text{ tel que } \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1\}. \tag{55.207}$$

Proof. We initiate with a matrix $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathbb{M}(2, \mathbb{C})$. Then we impose the conditions. The unitary property gives:

$$UU^\dagger = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{pmatrix} = \begin{pmatrix} \alpha\overline{\alpha} + \beta\overline{\beta} & \alpha\bar{\gamma} + \beta\bar{\delta} \\ \gamma\bar{\alpha} + \delta\bar{\beta} & \gamma\overline{\gamma} + \delta\bar{\delta} \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{55.208}$$

Among with the determinant conditions, we have the system SUBEQSooGUDNooOoxdSO

$$\begin{cases} \alpha\delta - \gamma\beta = 1 & \text{(55.209a)} \\ |\alpha|^2 + |\beta|^1 = 1 & \text{(55.209b)} \\ |\gamma|^2 + |\delta|^2 = 1 & \text{(55.209c)} \\ \alpha\bar{\gamma} + \beta\bar{\delta} = 0. & \text{(55.209d)} \end{cases}$$ SUBEQooSPRRooWjAUNi

We multiply (55.209d) by γ, and we substitute $\gamma\bar{\gamma} = 1 - |\delta|^2$ and $\gamma\beta = \alpha\delta - 1$. What we get is

$$\alpha(1 - |\delta|^2) + (\alpha\delta - 1)\bar{\delta} = 0. \tag{55.210}$$

If you develop the products, you see some simplifications and you remain with $\delta = \bar{\alpha}$.

Now we substitute $\delta = \bar{\alpha}$ in (55.209d) again. We obtain

$$\alpha(\bar{\gamma} + \beta) = 0. \tag{55.211}$$ EQooNOVWooYSTXqJ

There are two possibilities: $\alpha = 0$ or $\alpha \neq 0$.

(1) **If $\alpha \neq 0$** In that case the equality (55.211) produces $\gamma = -\bar{\beta}$ and the result is proved.

(2) **If $\alpha = 0$** The system (55.209) reduces to

$$\begin{cases} \gamma\beta = -1 & \text{(55.212a)} \\ |\beta|^2 = 1 & \text{(55.212b)} \\ |\gamma|^2 = 1. & \text{(55.212c)} \end{cases}$$ SUBEQooVCDOooTGejzi SUBEQooWJJMooItqDsi

There exist $\theta \in \mathbb{R}$ such that $\beta = e^{i\theta}$. The equation (55.212a) then shows that $\gamma = -e^{-i\theta}$. The condition (55.212c) is automatically satisfied. At the end we have the matrix

$$U = \begin{pmatrix} 0 & \beta \\ -\bar{\beta} & 0 \end{pmatrix}. \tag{55.213}$$

□

55.14.1 SU(2) as compact group

PROPooGLPQooKOfrjl

Proposition 55.34.
The group SU(2) *is compact.*

Proof. The proposition 55.33 show that SU(2) is contained in a bounded subset of $\mathbb{R}^8$, and it is clear that SU(2) is closed in $\mathbb{R}^8$ because it is defined by equalities. □

55.14.2 Pauli matrices

We denote by V the set of the hermitian traceless matrices. These are the elements $u \in \mathbb{M}(2, \mathbb{C})$ such that

$$u^\dagger = u \tag{55.214a}$$
$$\text{Tr}(u) = 0. \tag{55.214b}$$

This is not the Lie algebra $\mathfrak{su}(2)$ because the elements of $\mathfrak{su}(2)$ are anti-hermitian [12]. This set V is a vector space over $\mathbb{R}$, but not over the field $\mathbb{C}$ because if $u \in V$, then $(iu)^\dagger = -iu^\dagger = -iu \neq iu$.

12. Proposition 55.57.

DEFooRNTDooTVkPtB

Definition 55.35.
We consider the following matrices:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{55.215}$$

*The **Pauli matrices** are the matrices σ_1, σ_2 and σ_3.*

The matrix $\sigma_0 = \mathrm{Id}$ is introduced for later use.

For the sake of notations, we write σ the vector of V^3 given by $\sigma = (\sigma_1, \sigma_2, \sigma_3)$. This allows us to write combinations like

EQooXNRGooZaRQoZ

$$a \cdot \sigma = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \tag{55.216}$$

when $a \in \mathbb{R}^3$. It must be noticed however that the notation «$a \cdot \sigma$» is not a scalar product. In particular, the formula (55.216) depends on the chosen basis $\{\sigma_i\}$ of V and $\{e_i\}$ on $\mathbb{R}^3$.

LEMooZNCQooLgoReX

Lemma 55.36.
The Pauli matrice form a basis[13] *of the real vector space V of hermitian traceless matrices.*

Proof. An element of V is a matrice of the form

$$u = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{55.217}$$

with $a, b, c, d \in \mathbb{C}$. The condition $u = u^\dagger$ imposes the relations $a = \bar{a}$, $d = \bar{d}$ and $c = \bar{b}$, so that

$$u = \begin{pmatrix} x & z \\ \bar{z} & y \end{pmatrix}. \tag{55.218}$$

The trace condition imposes $x = -y$. Finally a general element of V has the form

$$u = \begin{pmatrix} x & z \\ \bar{z} & -x \end{pmatrix} \tag{55.219}$$

with $x \in \mathbb{R}$ and $z \in \mathbb{C}$. We have:

$$u = -\operatorname{Im}(z)\sigma_1 + \operatorname{Re}(z)\sigma_2 + x\sigma_3. \tag{55.220}$$

This proves that $\{\sigma_i\}_{i=1,2,3}$ spans V.

We still have to prove that $\{\sigma_i\}_{i=1,2,3}$ is free. For that, consider $a, b, c \in \mathbb{R}$ such that $a\sigma_1 + b\sigma_2 + c\sigma_3 = 0$:

$$\begin{pmatrix} c & a - bi \\ a + bi & -c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{55.221}$$

We immediately deduce $c = 0$, $a + bi = 0$ and $a - bi = 0$. Thus $a = b = c = 0$. □

Notice that the hermitian matrices do not form a vector space over $\mathbb{C}$ because, if X is hermitian,

$$(\lambda X)^\dagger = \bar{\lambda} X^\dagger = \bar{\lambda} X \neq \lambda X. \tag{55.222}$$

Lemma 55.37.
The matrices $i\sigma_k$ are unitary.

Proof. A simple computation show that $\sigma_k^2 = \mathbb{1}$ and $\sigma_k^\dagger = \sigma_k$, so that

$$(i\sigma_k)(i\sigma_k)^\dagger = (i\sigma_k)(-i\sigma_k) = \sigma_k^2 = \mathbb{1}. \tag{55.223}$$

□

13. Definition 4.5.

55.14.2.1 Some relations

LEMooIBJMooTYnooZ

Lemma 55.38.
For every $i, j = 1, 2, 3$ we have the formula

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i \sum_m \epsilon_{ijm} \sigma_m. \tag{55.224}$$

In particular $\sigma_i^2 = \mathbb{1}$.

Proof. Explicit matricial computation. □

LEMooJRWXooMkzRnk

Lemma 55.39.
We have the commutator[14]

$$[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk} \sigma_k. \tag{55.225}$$

Proof. This is a computation using the lemma 55.38:

$$[\sigma_i, \sigma_j] = \delta_{ij} \mathbb{1} + \sum_k i \epsilon_{ijk} \sigma_k - \delta_{ji} \mathbb{1} - \sum_k i \epsilon_{jik} \sigma_k = 2i \sum_k \epsilon_{ijk} \sigma_k \tag{55.226}$$

where we used the fact that $\epsilon_{jik} = -\epsilon_{ijk}$. □

LEMooLNCSooPHsVut

Lemma 55.40.
If $a, b \in \mathbb{R}^3$ we have

$$(a \cdot \sigma)(b \cdot \sigma) = (a \cdot b)\mathbb{1} + i(a \times b) \cdot \sigma. \tag{55.227}$$

Proof. We use lemme 55.38:

$$(a \cdot b)(b \cdot \sigma) = \big(\sum_i a_i \sigma_i\big)\big(\sum_j b_j \sigma_j\big) \tag{55.228a}$$

$$= \sum_{ij} a_i b_j \sigma \sigma_j \tag{55.228b}$$

$$= \sum_{ij} a_i b_j (\delta_{ij} \mathbb{1} + \sum_m i \epsilon_{ijm} \sigma_m) \tag{55.228c}$$

Then we use the property $\sum_{kl} a_k b_l \epsilon_{klm} = (a \times b)_m$. □

55.14.2.2 Isomorphism with $\mathbb{R}^3$

The following lemme does not aims to provide a norm on V. The norm on V is already the operator norm:

$$\|X\| = \sup_{|z|=1} |Xz| \tag{55.229}$$

where $z \in \mathbb{C}$ and $X \in V$. This is not something new.

In the same perspective, the elements of SU(2) are normed to 1 because, if $U \in \mathrm{SU}(2)$,

$$\|U\|_{\mathrm{SU}(2)} = \sup_{|z|=1} |Uz| = \sup \langle Uz, Uz \rangle = \sup_{|z=1|} \langle U^\dagger U z, z \rangle = \sup_{|z|=1} \langle z, z \rangle = 1. \tag{55.230}$$

LEMooRFBTooIRDbEq

Lemma 55.41.
The map

$$\begin{aligned} \phi \colon \mathbb{R}^3 &\to V \\ a &\mapsto a \cdot \sigma \end{aligned} \tag{55.231}$$

is a vector space isomorphism and satisfies

$$\det\big(\phi(x)\big) = -\|x\|^2. \tag{55.232}$$

14. We are adult here; I believe you will not confuse the i of the index and the i of the imaginary numbers.

Proof. Some immediate facts:

— ϕ is linear,

— ϕ is bijective because $\{\sigma_i\}_{i=1,2,3}$ is a basis (lemme 55.36).

— Thus ϕ is a vector space isomorphism.

The formula $\det\big(\phi(x)\big) = -\|x\|^2$ is a computation:

$$\begin{aligned}
\det\big(\phi(x)\big) &= \det\begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} && (55.233a)\\
&= -x_3^2 - (x_1 - ix_2)(x_1 + ix_2) && (55.233b)\\
&= -(x_1^2 + x_2^2 + x_3^2) && (55.233c)\\
&= -\|x\|^2. && (55.233d)
\end{aligned}$$

□

Some more properties about the Pauli matrices and the map ϕ.

Proposition 55.42 ([899]).
Let $x, y \in \mathbb{R}^3$. We have

ITEMooDDRNooGZASBN

(1) $[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk}\sigma_k$.

ITEMooXORKooXFwQhR

(2) $\phi(x \times y) = \frac{1}{2i}[\phi(x), \phi(y)]$.

ITEMooREMBooLPVnxz

(3) $\operatorname{Tr}\big(\phi(x)\phi(y)\big) = 2x \cdot y$.

Proof. Several points.

(1) **Formula (1)** We use the product formula of lemma 55.38:

$$[\sigma_i, \sigma_j] = \sigma_i\sigma_j - \sigma_j\sigma_i = \delta_{ij}\,\mathrm{Id} + i\sum_k \epsilon_{ijk}\sigma_k - \delta_{ij}\,\mathrm{Id} - i\sum_k \epsilon_{jik}\sigma_k. \tag{55.234}$$

Using the fact that $\epsilon_{ijk} = -\epsilon_{jik}$ we get the result.

(2) **Formula (2)** We have

$$[\phi(x), \phi(y)] = [x \cdot \sigma, y \cdot \sigma] = \sum_{ij} x_i y_j [\sigma_i, \sigma_j]. \tag{55.235}$$

Substituting the first result and using the formula $\sum_{ij} x_i y_j \epsilon_{ijk} = (x \times y)_k$ [15] we get

$$[\phi(x), \phi(y)] = \sum_{ijk} x_i y_j 2i\epsilon_{ijk}\sigma_k = 2i\sum_k (x \times y)_k \sigma_k = 2i(x \times y) \cdot \sigma = 2i\phi(x \times y). \tag{55.236}$$

(3) **Formula (3)** Using formula of lemma 55.40,

$$\begin{aligned}
\operatorname{Tr}\big(\phi(x)\phi(y)\big) &= \operatorname{Tr}\big((x \cdot \sigma)(y \cdot \sigma)\big) && (55.237a)\\
&= \sum_i \big[(x \cdot \sigma)(y \cdot \sigma)\big]_{ii} && (55.237b)\\
&= \operatorname{Tr}\big((x \cdot y)\mathbb{1}_2 + i(x \times y) \cdot \sigma\big) && (55.237c)\\
&= (x \cdot y)\operatorname{Tr}(\mathbb{1}_2) && (55.237d)\\
&= 2(x \cdot y). && (55.237e)
\end{aligned}$$

SUBEQooRJKBooNrLhhV

We used the fact that the Pauli matrice have vanishing trace, so that the second term in (55.237c) is zero.

□

15. Definition 10.28.

55.14.2.3 Path connection

PROPooLEKXooSXPhRX

Proposition 55.43.
The Lie groups SO(3) *and* SU(2) *are path connected.*

Proof. Particular cases of theorem 13.4 and proposition 13.6. □

55.14.2.4 One representation

We still consider V, the real vector space of hermitian matrices with vanishing trace. Thanks to the lemma 55.41 we define the following norm on V:

$$\|X\| = \|\phi^{-1}(X)\|_{\mathbb{R}^3} = -\det(X). \tag{55.238}$$

For the sake of notational convenience in the proof of the next proposition, we introduce the maps L and R.

LEMooQVYXooQFNaGc

Lemma 55.44.
The maps

$$\begin{aligned} L\colon \mathrm{GL}(V) &\to \mathrm{End}\big(\mathrm{End}(V)\big) \\ L(A)X &= AX \end{aligned} \tag{55.239}$$

and

$$\begin{aligned} R\colon \mathrm{GL}(V) &\to \mathrm{End}\big(\mathrm{End}(V)\big) \\ R(A)X &= XA \end{aligned} \tag{55.240}$$

are continuous.

Proof. We prove our statement for L. Let $A_k \xrightarrow{\mathrm{End}(V)} A$. We want to prove that

$$\|L(A_k) - L(A)\|_{\mathrm{End}\left(\mathrm{End}(V)\right)} \to 0. \tag{55.241}$$

Using the definition of the operator norm [16], and the fact that it is an algebra norm (lemme 10.69),

$$\|L(A_k) - L(A)\| = \sup_{X\in\mathrm{End}(V)} \frac{\|L(A_k)X - L(A)X\|}{\|X\|} \tag{55.242a}$$

$$= \sup_{X\in\mathrm{End}(V)} \frac{\|(A_k - A)X\|_{\mathrm{End}(V)}}{\|X\|} \tag{55.242b}$$

$$\leqslant \sup_{X\in\mathrm{End}(V)} \|A_k - A\| \tag{55.242c}$$

$$\to 0. \tag{55.242d}$$

Thus L is continuous by proposition 7.249. □

PROPooRQUZooAoZzwx

Proposition 55.45.
We still consider V, the real vector space of hermitian matrices [17] *with vanishing trace. Let*

$$\begin{aligned} \rho\colon \mathrm{SU}(2) &\to \mathrm{End}(V) \\ \rho(U)X &= UXU^\dagger. \end{aligned} \tag{55.243}$$

Then

ITEMooLZBSooZUQGgJ

(1) The map ρ is well defined: $\rho(U)X \in V$ for every $U \in \mathrm{SU}(2)$ and $X \in V$.

(2) The map ρ is a representation of SU(2) *on V by isometries,*

16. Definition 10.55.
17. Definition 55.11.

ITEMooBZUQooNXNVfs

(3) *for each U the map $\rho(U)\colon V \to V$ is continuous,*

ITEMooGHZYooQuabWb

(4) *the map $\rho\colon \mathrm{SU}(2) \to \mathrm{End}\big(\mathrm{End}(V)\big)$ is continuous.*

Proof. Let us prove that $\rho(U)X \in V$. First, using the properties of lemma 55.13,

$$(UXU^\dagger)^\dagger = (U^\dagger)^\dagger X^\dagger U^\dagger = UXU^\dagger. \tag{55.244}$$

Then, with the cyclic invariance of the trace (lemma 4.65),

$$\mathrm{Tr}(UXU^\dagger) = \mathrm{Tr}(U^\dagger UX) = \mathrm{Tr}(X) = 0. \tag{55.245}$$

So $UXU^{-1} = UXU^\dagger \in V$.

The fact that $\rho(U)$ is linear is a small computation. It is a representation because

$$\rho(U_1)\rho(U_2)X = \rho(U_1)U_2XU_2^\dagger = U_1U_2XU_2^\dagger U_1^\dagger = \rho(U_1U_2)X. \tag{55.246}$$

For the isometry part, the determinant being multiplicative (proposition 8.249),

$$\|UXU^\dagger\| = -\det(UXU^\dagger) = -\det(U)\det(X)\det(U^\dagger) = -|\det(U)|\det(X). \tag{55.247}$$

Since $U \in \mathrm{SU}(2)$ we have $|\det(U)| = 1$ and then $\|\rho(U)X\| = \|X\|$.

The last point to check is the continuity of $\rho\colon \mathrm{SU}(2) \to \mathrm{End}(V)$. With the notations of lemma 55.44 we have $\rho(U) = L(U) \circ R(U^\dagger)$ while $L(U)$ and $R(U)$ are continuous [18]. This is point (3) of continuity.

The point (4) of the continuity statement is more subtle. It is done in proposition 12.113. □

55.14.3 Link with SO(3)

PROPooGEHAooPCReoU

Proposition 55.46 ([844])**.**
The map

$$\begin{aligned} f\colon \mathrm{SU}(2) &\to \mathrm{SO}(3) \\ U &\mapsto \phi^{-1} \circ \rho(U) \circ \phi \end{aligned} \tag{55.248}$$

EQooSOZTooTIkQNx

where

$$\begin{aligned} \phi\colon \mathbb{R}^3 &\to V \\ x &\mapsto \langle x, \sigma \rangle \end{aligned} \tag{55.249}$$

is

(1) *continuous,*

(2) *a group homomorphism,*

ITEMooZSSHooDUCqSQ

(3) *surjective,*

(4) $\ker(f) = \{\pm\mathbb{1}\}$.

Proof. Several points.

(1) **Continuous** We know from proposition 55.45 that $U \mapsto \rho(U)$ is continuous. The inequality

$$\|\phi^{-1} \circ \rho(U_k) \circ \phi - \phi^{-1} \circ \rho(U) \circ \phi\| \leqslant \|\phi^{-1}\|\|\phi(U_k) - \rho(U)\|\|\phi\| \tag{55.250}$$

shows that f is continuous too.

18. They are matrix multiplication.

(2) **Group homomorphism** Let $U_1, U_2 \in \mathrm{SU}(2)$ and $x \in \mathbb{R}^3$. We have

$$\begin{aligned} f(U_1U_2)x &= \phi^{-1}\big(U_1U_2\phi(x)U_2^{-1}U_1^{-1}\big) && (55.251a)\\ &= \phi^{-1}\Big(\rho(U_1)\big(U_2\phi(x)U_2^{-1}\big)\Big) && (55.251b)\\ &= \phi^{-1}\big(\rho(U_1)\circ\rho(U_2)\phi(x)\big) && (55.251c)\\ &= \big(\phi^{-1}\circ\rho(U_1)\circ\rho(U_2)\circ\phi\big)(x). && (55.251d) \end{aligned}$$

We can suppress the dependency on x and continue:

$$\begin{aligned} f(U_1U_2) &= \phi^{-1}\circ\rho(U_1)\circ\rho(U_2)\circ\phi && (55.252a)\\ &= \phi^{-1}\circ\rho(U_1)\circ\phi\circ\phi^{-1}\circ\rho(U_2)\circ\phi && (55.252b)\\ &= f(U_1)\circ f(U_2). && (55.252c) \end{aligned}$$

Since $\rho(\mathrm{Id}) = \mathrm{Id}$ we also have $f(\mathrm{Id}) = \mathrm{Id}$. Thus f is a group homomorphism.

(3) **Surjective** Elements of $\mathrm{SO}(3)$ are compositions of two reflexions by corollary 18.89. A generic element of $\mathrm{SO}(3)$ has the form ST where S and T are reflexions. They have the form

$$\begin{aligned} S\colon \mathbb{R}^3 &\to \mathbb{R}^3\\ x &\mapsto x - 2(x\cdot n_1)n_1 \end{aligned} \tag{55.253}$$

and

$$\begin{aligned} T\colon \mathbb{R}^3 &\to \mathbb{R}^3\\ x &\mapsto x - 2(x\cdot n_2)n_2 \end{aligned} \tag{55.254}$$

with $\|n_1\| = \|n_2\| = 1$.
Let $M = \phi(n_1) = n_1\cdot\sigma$ and $Q = \phi(n_2) = n_2\cdot\sigma$. We will prove that $MQ \in \mathrm{SU}(2)$ and $f(MQ) = S\circ T$.

(3a) $M^2 = \mathbb{1}$ We have

$$M^2 = (n_1\cdot n_1)\mathbb{1}_2 + i(n_1\times n_1)\cdot\sigma = \mathbb{1}_2. \tag{55.255}$$

(3b) $\det(M) = -1$ We know from lemma 55.41 that $\det(M) = \det\big(\phi(n_1)\big) = -\|n_1\|^2 = -1$.

(3c) $MQ \in \mathrm{SU}(2)$ First, $\det(MQ) = \det(M)\det(Q) = -1$. Second, since M and Q are hermitian, $(MQ)^\dagger = Q^\dagger M^\dagger = QM$ and then

$$(MQ)^\dagger(MQ) = QMMQ = \mathbb{1}_2 \tag{55.256}$$

because $M^2 = Q^2 = 1$.

(3d) $\phi(Sx) = -M\phi(x)M$ Let $x \in \mathbb{R}^3$. We have EQooSHKEooAhOxfH

$$\begin{aligned} \phi(Sx) &= \phi\big(x - 2(x\cdot n_1)n_1\big) && (55.257a)\\ &= \phi(x) - 2(x\cdot n_1)\phi(n_1) && (55.257b)\\ &= \phi(x) - 2(x\cdot n_1)M. && (55.257c) \end{aligned}$$

Using the formula $(a\cdot\sigma)(b\cdot\sigma) = (a\cdot b)\mathbb{1} + i(a\times b)\cdot\sigma$ we have

$$\phi(x)M = \phi(x)\phi(n_1) = (x\cdot n_1)\mathbb{1} + i(x\times n_1)\cdot\sigma \tag{55.258}$$

and

$$M\phi(x) = (n_1\cdot\sigma)(x\cdot\sigma) = (n_1\cdot x)\mathbb{1} + i(n_1\times x)\cdot\sigma, \tag{55.259}$$

si that

$$\phi(x)M + M\phi(x) = 2(x\cdot n_1)\mathbb{1}. \tag{55.260}$$

Multiplying that by M and using $M^2 = \mathbb{1}$ we deduce

$$\phi(x) = 2(x\cdot n_1)M - M\phi(x)M. \tag{55.261}$$

Now we substitute that into (55.257) in order to see that

$$\phi(Sx) = -M\phi(x)M. \tag{55.262}$$

(3e) **Conclusion (surjective)** We can now compute the action of $f(MQ)$ on $x \in \mathbb{R}^3$:

$$\begin{aligned} f(MQ)x &= \big(\phi^{-1} \circ \rho(MQ) \circ \phi\big)x && (55.263a)\\ &= \phi^{-1}\big(MQ\phi(x)QM\big) && (55.263b)\\ &= \phi^{-1}\big(M\phi(Tx)M\big) && (55.263c)\\ &= \phi^{-1}\big(\phi(STx)\big) && (55.263d)\\ &= STx. && (55.263e) \end{aligned}$$

So we have $f(MQ) = ST$ and f is surjective.

(4) **Kernel** A less technological proof will be given for fun in the lemma 55.102. Let $U \in \mathrm{SU}(2)$ be such that $f(U) = \mathbb{1}_3 \in \mathrm{SO}(3)$. For every $x \in \mathbb{R}^3$ we have $x = f(U)x$ while

$$f(U)x = \phi^{-1}\big(U\phi(x)U^{-1}\big). \tag{55.264}$$

We conclude that $U\phi(x)U^{-1} = \phi(x)$ for every $x \in \mathbb{R}^3$. Since f is surjective on the vector space V of hermitian matrices with vanishing trace, we have

$$UXU^\dagger = X \tag{55.265}$$

for every $X \in V$. In particular $UX = XU$. Since the matricial product is continuous, we can commute U and the infinite sum and get

$$U\sum_{k=0}^{\infty}\frac{(iX)^k}{k!} = \lim_{N\to\infty}\sum_{k=0}^{N}\frac{U(iX)^k}{k!} = \lim_{N\to\infty}\sum_{k=0}^{N}\frac{(iX)^kU}{k!} = e^{iX}U. \tag{55.266}$$

So we have $[U, e^{iX}] = 0$ for every $X \in V$. Since proposition 55.25 says that every element of $\mathrm{SU}(2)$ is the exponential of an element in V the element U is in the center of $\mathrm{SU}(2)$. The center of $\mathrm{SU}(2)$ is $\{\pm \mathrm{Id}\}$ by the proposition 55.26.

Until now we have $\ker(f) \subset \{\mathrm{Id}, -\mathrm{Id}\}$. It is a simple verification to check that $\{\mathrm{Id}, -\mathrm{Id}\}$ are in the kernel of f. We conclude that $\ker(f) = \{\pm\mathbb{1}_2\}$.

□

PROPooDKPTooBnLflt

Proposition 55.47.
We have the group isomorphism

$$\mathrm{SO}(3) = \frac{\mathrm{SU}(2)}{\mathbb{Z}_2}. \tag{55.267}$$

Proof. We use the first isomorphism theorem 2.6 with θ being the map $f\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ defined by the proposition 55.46. It says that

$$\frac{\mathrm{SU}(2)}{\ker(f)} = \mathrm{Image}(f). \tag{55.268}$$

The known properties of f are that $\ker(f) = \mathbb{Z}_2$ and $\mathrm{Image}(f) = \mathrm{SO}(3)$. This is the expected result. □

LEMooSYGUooVWxGYX

Lemma 55.48.
The images of the unitary matrices $i\sigma_k$ by f are

$$f(i\sigma_1) = \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}, f(i\sigma_2) = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, f(i\sigma_3) = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}. \tag{55.269}$$

Proof. We know that

$$\phi(x) = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}. \tag{55.270}$$

We have

$$(i\sigma_1)\phi(x)(i\sigma_1)^\dagger = \sigma_1\phi(x)\sigma_1 = \begin{pmatrix} -x_3 & x_1 + ix_2 \\ x_1 - ix_2 & x_3 \end{pmatrix} = \phi \begin{pmatrix} x_1 \\ -x_2 \\ -x_3 \end{pmatrix}. \tag{55.271}$$

This shows that

$$f(i\sigma_1) = \begin{pmatrix} x_1 \\ -x_2 \\ -x_3 \end{pmatrix}, \tag{55.272}$$

so that the matrix of $i\sigma_1$ is

$$f(i\sigma_1) = \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}. \tag{55.273}$$

The same kind of computations provide the result. □

Proposition 55.49.
Let $f\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ be the map of the proposition 55.46:

$$\begin{aligned} f\colon \mathrm{SU}(2) &\to \mathrm{SO}(3) \\ f(U) &= \phi^{-1} \circ \rho(U) \circ \phi. \end{aligned} \tag{55.274}$$

There exist no group homomorphism $g\colon \mathrm{SO}(3) \to \mathrm{SU}(2)$ such that $f \circ g = \mathrm{Id}$.

Proof. Let g be such an homomorphism and let's derive a contradiction. Since g is an homomorphism it satisfies $g(\mathbb{1}_3) = \mathbb{1}_2$. Let

$$T_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in \mathrm{SO}(3). \tag{55.275}$$

The map f is surjective, so there exist $U \in \mathrm{SU}(2)$ such that $f(U) = T_x$. From lemma 55.48 we have

$$f(i\sigma_1) = f(-i\sigma_1) = T_x. \tag{55.276}$$

Thus $g(T_x) = i\sigma_1$ or $g(T_x) = -i\sigma_1$. In both cases we have a contradiction. Indeed, since $T_x^2 = \mathbb{1}$ and $g(T_x^2) = g(T_x)^2$ we must have $g(T_x)^2 = \mathbb{1}$ while

$$(i\sigma_1)^2 = (-i\sigma_1)^2 = \mathbb{1}. \tag{55.277}$$

□

LEMooRCSSooTvAaJY

Lemma 55.50.
Let $\alpha_0 \in \mathrm{SO}(3)$ and $U \in \mathrm{SU}(2)$ such that $f(U) = \alpha_0$. There exists a neighborhood $\mathcal{O}$ of α_0 in $\mathrm{SU}(2)$ such that

$$f^{-1}(\mathcal{O}) = V_1 \cup V_2 \tag{55.278}$$

where V_1 is a neighborhood of U, V_2 is a neighborhood of $-U_1$ and $V_1 \cap V_2 = \varnothing$.

Proof. We know that $f(U) = f(-U) = \alpha_0$. Let W_1 be a neighborhood of U and W_2 be a neighborhood of $-U$ such that $W_1 \cap W_2 = \varnothing$.

The part $-W_2$ is a neighborhood of U. we consider V_1, a neighborhood of U contained in $W_1 \cap -W_2$. Then we set $V_2 = -V_1$. This is a neighborhood of $-U$ contained in W_2.

Thus we have $V_1 \cap V_2 = \varnothing$ and $f(V_1) = f(V_2)$ is a neighborhood of α_0. It remains to define $\mathcal{O} = f(V_1)$. □

PROPooHQENooUsQeiZ

Proposition 55.51.
Let $U \in \mathrm{SU}(2)$ be such that

$$UX = XU \tag{55.279}$$

for every $X \in V$. Then $U \in \{\mathbb{1}, -\mathbb{1}\}$.

Proof. The proof is a simple computation. Let $a, b, c, d \in \mathbb{C}$ such that $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We have

$$U\sigma_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ d & c \end{pmatrix} \tag{55.280}$$

while

$$\sigma_1 U = \begin{pmatrix} c & d \\ a & b \end{pmatrix}. \tag{55.281}$$

We deduce $b = c$ and $a = d$ and $U = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$. Taking that into account, the same work with σ_2 provides

$$U\sigma_2 = \begin{pmatrix} bi & -ia \\ ia & -ib \end{pmatrix} \tag{55.282}$$

and

$$\sigma_2 U = \begin{pmatrix} -bi & -ia \\ ia & bi \end{pmatrix}, \tag{55.283}$$

so that $b = 0$. Now $U = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ for some $a \in \mathbb{C}$.

The constrain $U\sigma_3 = \sigma_3 U$ does not provide new informations.

Since $U \in \mathrm{SU}(2)$ we have $\det(U) = 1$ which implies $a^2 = 1$, which in turn means $a = \pm 1$. □

LEMooBHVBooEPbWwZ

Lemma 55.52 ([1]).
Let $U_k \in \mathrm{SU}(2)$ such that

(1) $\rho_{U_k} \xrightarrow{\mathrm{End}(V)} \mathrm{Id}_V$

(2) there exists a neighborhood of $-\mathbb{1}$ in $\mathrm{SU}(2)$ which contains no element U_k.

Then we have $U_k \xrightarrow{\mathrm{SU}(2)} \mathbb{1}$.

Proof. Let A_k be a converging subsequence of U_k (we will see later that it exists) with $A_k \xrightarrow{\mathrm{SU}(2)} A$. For each $X \in V$ we have $A_k X A_k^{-1} \xrightarrow{\mathrm{SU}(2)} X$, so that

$$\begin{aligned} \|A_k X - X A_k\| &= \|A_k X A_k^{-1} A_k - X A_k\| && \text{(55.284a)} \\ &\leqslant \|A_k X A_k^{-1} - X\| \|A_k\| && \text{(55.284b)} \\ &\leqslant \|A_k X A_k^{-1} - X\| M \to 0 && \text{(55.284c)} \end{aligned}$$

where M is some constant majoration of $\|A_k\|$. Thus we have $A_k X - X A_k \to 0$ which means

$$XA = AX. \tag{55.285}$$

If it is true for every X, we conclude that $A = \pm\mathbb{1}$ (proposition 55.51). Since there is a neighborhood of $-\mathbb{1}$ in which there are no elements U_k, we cannot have $A_k \to -\mathbb{1}$, so we have $A = \mathbb{1}$.

Now U_k is a sequence in the compact $\mathrm{SU}(2)$ (proposition 55.34), so that every subsequence has a converging subsequence [19]. We are in the case of the lemma 7.66 and we conclude $U_k \xrightarrow{\mathrm{SU}(2)} \mathbb{1}$. □

19. Bolzano-Weierstrass 9.40.

LEMooMNWSooAjmBQK

Lemma 55.53 ([1]).
Let $U \in \mathrm{SU}(2)$. We consider the linear map[20]

$$\begin{aligned} \rho_U : V &\to V \\ v &\mapsto UvU^{-1}. \end{aligned} \tag{55.286}$$

Then $\|\rho_U\| \leqslant 1$.

Proof. By definition, the norme of $\rho_U : V \to V$ is

$$\|\rho_U\| = \sup_{\|v\|=1} \|\rho_U v\| = \sup_{\|v\|=1} \|UvU^{-1}\|. \tag{55.287}$$

In the last expression, the norms are in $\mathrm{End}(\mathbb{C}^2)$ because U and v are both 2×2 complex matrices. The operator norm is an algebra norm[21], so that

$$\|UvU^{-1}\| \leqslant \|U\|\|U^{-1}\|\|v\| = \|v\| \tag{55.288}$$

because the éléments of $\mathrm{SU}(2)$ are normed to 1. Thus

$$\|\rho_1\| \leqslant \sup_{\|v\|=1} \|v\| = 1. \tag{55.289}$$

□

LEMooHPQQooIGwljm

Lemma 55.54.
Let U_k be a sequence in $\mathrm{SU}(2)$ and $U \in \mathrm{SU}(2)$ such that

(1) $\rho_{U_k} \xrightarrow{\mathrm{End}(V)} \rho_U$

(2) there exists a neighborhood[22] *of $-A$ in $\mathrm{SU}(2)$ which contains no element U_k.*

Then we have $U_k \xrightarrow{\mathrm{SU}(2)} U$.

Proof. Several steps.

(1) $\rho_{U_k^{-1}U} \to \mathrm{Id}$ We start by proving that $\rho_{U_k^{-1}U} \to \mathrm{Id}$. For each $v \in V$ we have

$$\|\rho_{U_k^{-1}U} v - v\| = \|\rho_{U_k^{-1}}\big(\rho_U v - \rho_{U_k} v\big)\| \leqslant \|\rho_{U_k^{-1}}\|\rho_U v - \rho_{U_k} v\|\| \tag{55.290}$$

Thus we have

$$\begin{aligned} \|\rho_{U_k^{-1}U} - \mathrm{Id}\,\| &= \sup_{\|v\|=1} \|\rho_{U_k^{-1}U} v - v\| && \text{(55.291a)} \\ &\leqslant \|\rho_{U_k^{-1}}\| \sup_{\|v\|=1} \|\rho_U v - \rho_{U_k} v\| && \text{(55.291b)} \\ &= \|\rho_{U_k}\|\|\rho_u - \rho_{U_k}\| && \text{(55.291c)} \\ &\leqslant \|\rho_U - \rho_{U_k}\|. && \text{(55.291d)} \end{aligned}$$

We used lemme 55.53. In conclusion,

$$\|\rho_{U_k^{-1}U} - \mathrm{Id}\,\| \leqslant \|\rho_U - \rho_{U_k}\| \to 0. \tag{55.292}$$

20. See proposition 55.45.
21. Lemma 10.69.
22. Ultimately, the topological properties of SU(2) are given by its analytic structure defined in the proposition 55.27.

(2) **No neighborhood of $-\mathbb{1}$** We prove that there exists a neighborhood of $-\mathbb{1}$ which contains no elements of the sequence $U_k^{-1}U$. Suppose that each neighborhood of $-\mathbb{1}$ contains one of the $U_k^{-1}U$. At this point we have a subsequence (B_k) of (U_k) such that

$$B_k^{-1}U \to -\mathbb{1}. \tag{55.293}$$

Since the multiplication and the inverse are continuous operations [23] we also have

$$B_k^{-1} \to -U^{-1} \tag{55.294}$$

and

$$B_k \to -U. \tag{55.295}$$

This prove that for every neighborhood of $-U$ we have a B_k and then a U_k, which is a contradiction with the hypothesis.

(3) **Conclusion** The sequence $U_k^{-1}U$ satisfies the lemma 55.52 and we conclude $U_k^{-1}U \to \mathbb{1}$. Thus $U_k \to U$.

□

PROPooHCVZooMOSzTm

Proposition 55.55 ([844, 1])**.**
Let $\alpha_0 \in \mathrm{SO}(3)$ and $\mathcal{O}$ be a neighborhood of α_0 such that $f^{-1}(\mathcal{O}) = V_1 \cup V_2$ with $V_1 \cap V_2 = \varnothing$ [24]*.*
The map

$$\begin{aligned} \varphi\colon \mathcal{O} &\to \mathrm{SU}(2) \\ \alpha &\mapsto f^{-1}(\alpha) \cap V_1 \end{aligned} \tag{55.296}$$

is continuous.

Proof. Let $\sigma_k \xrightarrow{\mathrm{SO}(3)} \alpha$ with $\alpha_k \in \mathcal{O}$. We have to prove that $\varphi(\alpha_k) \xrightarrow{\mathrm{SU}(2)} \varphi(\alpha)$.

(1) **General setting** First we suppose that α_k converges to the identity. For each k we have

$$f\big(\varphi(\alpha_k)\big) = \alpha_k, \tag{55.297}$$

with the map (55.248). That means, for each k:

$$\phi^{-1} \circ \rho_{\varphi(\alpha_k)} \circ \phi = \alpha_k, \tag{55.298}$$

or

$$\rho_{\varphi(\alpha_k)} = \varphi \circ \alpha_k \circ \phi^{-1} \tag{55.299}$$

as operator on V.

(2) **Norm convergence** We have the following computation:

$$\begin{aligned} \|\rho_{\varphi(\alpha_k)} - \rho_{\varphi(\alpha)}\|_{\mathrm{End}(V)} &= \|\phi \circ \alpha_k \circ \phi^{-1} - \phi \circ \alpha \circ \phi^{-1}\| && \text{(55.300a)} \\ &= \|\phi \circ (\alpha_k - \alpha) \circ \phi^{-1}\| && \text{(55.300b)} \\ &\leqslant \|\phi\|\|\alpha_k - \alpha\|\|\phi^{-1}\| \to 0. && \text{(55.300c)} \end{aligned}$$

This shows that

$$\rho_{\varphi(\alpha_k)} \to \rho_{\varphi(\alpha)}. \tag{55.301}$$

EQooCEFUooTCoczj

(3) **Conclusion**
The sequence $U_k = \varphi(\alpha_k)$ and the element $U = \varphi(\alpha)$ satisfy the lemma 55.54, so that $\varphi(\alpha_k) \to \varphi(\alpha)$.

□

23. The group SU(2) is a Lie group, proposition 55.27(1).
24. Such choice is possible by the lemma 55.50.

Proposition 55.56.
The map $f\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a representation of $\mathrm{SU}(2)$ on $\mathbb{R}^3$, but is not faithful[25].

Proof. The function f is written as

$$f(U) = \phi^{-1} \circ \rho_U \circ \phi. \tag{55.302}$$

On the other hand, we have

$$\rho_{U_1U_2} v = U_1U_2v(U_1U_2)^{-1} = U_1U_2vU_2^{-1}U_1^{-1} = (\rho_{U_1} \circ \rho_{U_2})v. \tag{55.303}$$

Thus

$$f(U_1U_2) = \phi^{-1} \circ \rho_{U_1} \circ \rho_{U_2} \circ \phi = \underbrace{\phi^{-1}\rho_{U_1}\phi}_{f(U_1)} \underbrace{phi^{-1}\rho_{U_2}\phi}_{f(U_2)} = f(U_1) \circ f(U_2). \tag{55.304}$$

Thus f is a representation.

It is not faithful because $f(\mathbb{1}) = f(-\mathbb{1}) = \mathrm{Id}$. □

55.14.4 The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$

PROPooSERWooFtxBgV

Proposition 55.57.
The Lie algebra $\mathfrak{su}(n)$ of $\mathrm{SU}(n)$ is isomorphic to the algebra of traceless anti-hermitian matrices[26].

$$\mathfrak{su}(n) = \{X \in \mathfrak{gl}(2,\mathbb{C}) \text{ tel que } X^\dagger = -X, \mathrm{Tr}(X) = 0\}. \tag{55.305}$$

Proof. Let consider $G = \mathrm{SU}(n)$; the elements are complexes $n \times n$ matrices U such that $UU^\dagger = \mathbb{1}$ and $\det U = 1$. An element of the Lie algebra is given by a path $u\colon \mathbb{R} \to G$ in the group with $u(0) = \mathbb{1}$. Since $\mathrm{SU}(n)$ is a Lie subgroup of $\mathrm{GL}(n,\mathbb{C})$[27], by the proposition 53.16, it is sufficient to compute the usual derivative of such a path. Since for all t, $u(t)u(t)^\dagger = \mathbb{1}$,

$$0 = \frac{d}{dt}\Big[u(t)u(t)^\dagger\Big]_{t=0} \tag{55.306a}$$

$$= u(0)\frac{d}{dt}\Big[u(t)^\dagger\Big]_{t=0} + \frac{d}{dt}\Big[u(t)\Big]_{t=0}u(0)^\dagger \tag{55.306b}$$

$$= [d_t u(t)]^\dagger + [d_t u(t)]. \tag{55.306c}$$

So a general element of the Lie algebra $\mathfrak{su}(n)$ is an anti-hermitian matrix.

An element of $\mathrm{SU}(n)$ has also a determinant equal to 1. What condition does it implies on the elements of the Lie algebra?

Let X be an element of $\mathfrak{su}(2)$. For each t, the element e^{tX} is part of the Lie group and satisfy $\det(e^{tX}) = 1$. Using the formula[28]

$$\frac{d}{dt}\Big[\det(e^{tX})\Big]_{t=0} = \mathrm{Tr}(X) \tag{55.307}$$

we deduce $\mathrm{Tr}(X) = 0$.

An other way to prove the same result is to consider a path in $\mathrm{SU}(n)$ and derive; let's do it. If $g(t)$ is a path in $\mathrm{SU}(n)$ with $g(0) = \mathbb{1}$. For each t we have $\det\big(g(t)\big) = 1$.

Using the formula expression the determinant with the minors,

$$\det\begin{pmatrix} g_{11}(t) & g_{12}(t) & \dots \\ f_{21}(t) & g_{22}(t) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} = g_{11}(t)M_{11}(t) + g_{12}(t)M_{12}(t) + \ldots = 1 \tag{55.308}$$

25. Définition 50.7.
26. See the definitions 53.5 and 49.50.
27. Proposition 55.27(1).
28. Corollary 13.24.

where M_{ij} is the minor of g. If we derive the left hand side we get

$$g'_{11}(0)M_{11}(0) + g_{11}(0)M'_{11}(0) + g'_{12}(0)M_{12}(0) + g_{12}(0)M'_{12}(0) + \dots \tag{55.309}$$

where the numbers $g'_{ij}(0)$ are the matrix entries of the tangent matrix, that is the matrix elements of a general element in $\mathfrak{su}(n)$. Since $g(0) = \mathbb{1}$ we have $M_{11}(0) = 1$, $g_{11}(0) = 1$, $M_{12}(0) = 0$ and $g_{12}(0) = 0$. Thus we have

$$(\det g)'(0) = X_{11} + M'_{11}(0) \tag{55.310}$$

where $X = g'(0)$. By induction we found that the trace of X appears. Thus the elements of $\mathfrak{su}(n)$ have vanishing trace. □

55.58.
The space V spanned by the matrices σ_i is not $\mathfrak{su}(2)$.

PROPooDNNEooMOdrkq

Proposition 55.59.
The Lie algebra $\mathfrak{so}(n)$ is the vector space of antisymmetric matrices.

Proof. As said in the proposition 53.16, the Lie algebra $\mathfrak{so}(n)$ can be seen as $\mathrm{SO}(n)'$, the Lie algebra of the matrices obtained by componentwise derivate paths in $\mathrm{SO}(n)$.

(1) **Inclusion in one sense** So let be a path $g\colon \mathbb{R} \to \mathrm{SO}(n)$. For each t we have the equality

$$g(t)g(t)^t = \mathbb{1}. \tag{55.311}$$

We differentiate that equation with respect to t at $t = 0$ taking into account $g(0) = \mathbb{1}$:

$$g'(0) + g'(0)^t = 0. \tag{55.312}$$

This shows that the matrices of the Lie algebra $\mathfrak{so}(n)$ are skew-symmetric.

(2) **Inclusion in the other sense** Now we prove that every skew-symmetric matrix is of the form $\gamma'(0)$ for some path $\gamma\colon \mathbb{R} \to \mathrm{SO}(n)$. Let X be a skew-symmetric matrix. We consider the path

$$\gamma(t) = e^{tX} \tag{55.313}$$

defined in the proposition 10.301. We have to prove that $\gamma(t) \in \mathrm{SO}(3)$ for every t (at least in a neighborhood of $t = 0$) and that $\gamma(0) = \mathbb{1}$.

(2a) $\gamma(0) = \mathbb{1}$ This is the proposition 15.69(1).

(2b) $\gamma(t)$ **is orhogonal** By proposition 15.69(3) we know that

$$(e^{tX})^t = e^{tX^t}. \tag{55.314}$$

Since X is skew-symmetric we also have $[X, X^t] = 0$, because

$$(XX^t)_{ij} = \sum_k X_{ik}X^t_{kj} = \sum_k X_{ik}X_{jk} \tag{55.315}$$

while

$$(X^tX)_{ij} = \sum_k X^t_{ik}X_{kj} = \sum_k X_{ki}X_{kj} = \sum_k X_{ik}X_{jk}. \tag{55.316}$$

The last equality accounts the fact that X is skew-symmetric. Since X and X^t commute we can use the theorem 15.73:

$$e^{tX}(e^{tX})^t = e^{tX}e^{tX^t} = e^{t(X+X^t)} = e^0 = \mathbb{1}. \tag{55.317}$$

(2c) $\gamma(t)$ **is special** We prove that $\det\big(\gamma(t)\big) = 1$. The proposition 13.23 provides

$$\det(e^{tX}) = e^{\mathrm{Tr}(tX)} = e^0 = 1. \tag{55.318}$$

By the way, these equalities are equalities in $\mathbb{R}$, not equalities on $\mathrm{GL}(n, \mathbb{R})$.

(2d) **Pause** We finished to prove that $\gamma(t) \in \mathrm{SO}(n)$ for every $t \in \mathbb{R}$. We still have to prove that $\gamma'(0) = X$.

(2e) $\gamma'(0) = X$ This is from proposition 15.72 :

$$\frac{d}{dt}\Big[e^{tX}\Big]_{t=0} = X. \tag{55.319}$$

□

55.60.
Notice that antisymmetric matrices are automatically with vanishing trace.

PROPooHOOLooOrcquD

Proposition 55.61.
Two Lie algebras.

ITEMooFSTMooGSjovL

(1) The Lie algebra of $\mathrm{U}(n)$ *is the set* $\mathfrak{u}(n)$ *of anti-hermitian matrices.*

ITEMooYEFMooRmGmlF

(2) The Lie algebra of $\mathrm{SU}(n)$ *is the set* $\mathfrak{su}(n)$ *of anti-hermitian matrices with vanishing trace*[29].

ITEMooXXTRooQZzCfs

(3) A basis of $\mathfrak{su}(2)$ *is given by the matrices* $t_k = -i\sigma_k$ *where* σ_k *are the Pauli matrices*[30]*, that is*

$$t_1 = i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad t_2 = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad t_3 = i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{55.320}$$

(4) The commutation relations in $\mathfrak{su}(2)$ *are*

EQooFJIDooRtQGjA

$$[t_i, t_j] = 2\sum_k \epsilon_{ijk} t_k. \tag{55.321}$$

Proof. The point (1) is the same kind of proof that the one of proposition 55.59; the only difference is that one starts with $g(t)g(t)^\dagger = \mathbb{1}$. Then one use $e^{X^\dagger} = (e^X)^\dagger$.

The point (2), is already proved in the proposition 55.57.

For the point (3), an explicit computation shows that the matrices t_k belong to $\mathfrak{su}(2)$ and are linearly independent. Now there are two ways to proceed.

One way is to prove that $\mathfrak{su}(2)$ has dimension 3. For that, you can write down an explicit manifold structure on $\mathrm{SU}(2)$ and show that it is a manifold of dimension 3. Then the Lie algebra has the same dimension.

An other way is to make it by hand. We consider a matrix $\begin{pmatrix} a & c \\ d & b \end{pmatrix} \in \mathbb{M}(\mathbb{C}, 2)$. The fact to be traceless imposes $a = -b$. Then we have

$$\begin{pmatrix} a & c \\ d & -a \end{pmatrix}^\dagger = \begin{pmatrix} \bar{a} & \bar{d} \\ \bar{c} & -\bar{a} \end{pmatrix}. \tag{55.322}$$

The condition to be skew-hermitian is

$$\begin{pmatrix} a & c \\ d & -a \end{pmatrix} = -\begin{pmatrix} \bar{a} & \bar{d} \\ \bar{c} & -\bar{a} \end{pmatrix}. \tag{55.323}$$

That provide the constrains that a is purely imaginary and that, if $c = x + iy$, then $d = -x + iy$. Thus a generic matrix in $\mathfrak{su}(2)$ is given by

$$\begin{pmatrix} \lambda i & x + iy \\ -x + iy & -\lambda i \end{pmatrix} = \lambda \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = x(i\sigma_2) + y(i\sigma_1) + \lambda(i\sigma_3). \tag{55.324}$$

with $x, y, \lambda \in \mathbb{R}$. That shows that $\{i\sigma_k\}$ is a basis of $\mathfrak{su}(2)$. In the same time this is a proof a proof that $\mathfrak{su}(2)$ has dimension 3.

29. Just to be clear: as set this is the skew-hermitian matrices. As vector space, this is a real vector space. The fact to be skew-hermitian is not preserved by a multiplication by i.

30. Definition 55.35.

The lemma 55.39 provides the commutators for the Pauli matrices. We can adapt them for our basis of $\mathfrak{su}(2)$:

$$[t_i, t_j] = [i\sigma_i, i\sigma_j] = -[\sigma_i, \sigma_j] = -2i\sum_k \epsilon_{ijk}\sigma_k = 2\sum_k \epsilon_{ijk}t_k. \tag{55.325}$$

□

Proposition 55.62.
The Lie algebras $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic.

Proof. The propositions 55.59 and 55.61 provide a description of $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$. The easiest way to prove the isomorphism is to show an explicit isomorphism. A basis of $\mathfrak{so}(3)$ is

$$O_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad O_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad O_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{55.326}$$

These matrices satisfy

$$[O_i, O_j] = \sum_k \epsilon_{ijk}O_k. \tag{55.327}$$ EQooWJMUooOtFAkW

A basis of $\mathfrak{su}(2)$ is given by $t_k = i\sigma_k$. Our isomorphism is

$$\begin{aligned} \varphi\colon\ \mathfrak{so}(3) &\to \mathfrak{su}(2) \\ O_i &\mapsto -\frac{i\sigma_k}{2}. \end{aligned} \tag{55.328}$$

The fact that φ is a bijection derives from the fact that it maps a basis on a basis. We have to check that φ is a morphism, that is

$$\big[\varphi(O_i), \varphi(O_j)\big] = \varphi\big([O_i, O_j]\big). \tag{55.329}$$

This is done by virtue of the commutators (55.327) and of lemma 55.39. □

55.14.5 Irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$

We are not here to joke or to be funny. We are here to make quantum fields theory. So we need (among maaaaaany other things) the irreducible representations of the groups $\mathrm{SL}(2, \mathbb{C})$ and $\mathrm{SU}(2)$. Here is a good news: at the Lie algebra level, these two are more or less related by the lemma 51.8.

LEMooVEJZooUVNdmE

Lemma 55.63 ([897]).
As sets,

$$\mathfrak{su}(2)_{\mathbb{C}} = \mathfrak{su}(2) \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{sl}(2, \mathbb{C}) = \{\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \text{ tel que } \alpha, \beta, \gamma \in \mathbb{C}\}. \tag{55.330}$$

The first equality is a definition for the notation $\mathfrak{su}(2)_{\mathbb{C}}$.

Proof. A basis of $\mathfrak{su}(2)$ is $\{t_1, t_2, t_3\}$; so $\mathfrak{su}(2)_{\mathbb{C}} = \mathbb{C}t_1 \oplus \mathbb{C}t_2 \oplus \mathbb{C}t_3$, that is

$$\mathfrak{su}(2)_{\mathbb{C}} = \{\begin{pmatrix} 0 & ix \\ ix & 0 \end{pmatrix} + \begin{pmatrix} 0 & y \\ -y & 0 \end{pmatrix} + \begin{pmatrix} iz & 0 \\ 0 & -iz \end{pmatrix} \text{ tel que } x, y, z \in \mathbb{C}\} \tag{55.331a}$$

$$= \{\begin{pmatrix} iz & ix + y \\ ix - y & -iz \end{pmatrix}' x, y, z \in \mathbb{C}\} \tag{55.331b}$$

$$= \{\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \text{ tel que } \alpha, \beta, \gamma \in \mathbb{C}\}. \tag{55.331c}$$

In order to determine the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ of $\mathrm{SL}(2,\mathbb{C})$ we use the proposition 53.16 to allow ourself to work at the matrix level. Let g be a smooth path in $\mathrm{SL}(2,\mathbb{C})$ such that $g(0) = \mathbb{1}$. A generic element of $\mathfrak{sl}(2,\mathbb{C})$ has the form $g'(0)$. We have

$$g(t) = \begin{pmatrix} \alpha(t) & \beta(t) \\ \gamma(t) & \delta(t) \end{pmatrix} \tag{55.332}$$

with

$$\alpha(t)\delta(t) - \gamma(t)\beta(t) = 1 \tag{55.333}$$

EQooMNXMooVkbfDg

for every t. Moreover $g(0) = \mathbb{1}$ implies $\alpha(0) = 1$, $\beta(0) = 0$, $\gamma(0) = 0$ and $\delta(0) = 1$. Now we differentiate (55.333) with respect to t at $t = 0$:

$$\alpha'(0)\delta(0) + \alpha(0)\delta'(0) - \gamma'(0)\beta(0) - \gamma(0)\beta'(0) = 0 \tag{55.334}$$

which reduces to $\alpha'(0) + \delta'(0) = 0$. An element of $\mathfrak{sl}(2,\mathbb{C})$ is thus of the form $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ with $a, b, c \in \mathbb{C}$. □

55.64.
Lemma 55.63 speaks about the *sets* of $\mathfrak{su}(2)_{\mathbb{C}}$ and $\mathfrak{sl}(2,\mathbb{C})$. The Lie bracket, in both cases, is the matrix commutator. As a vector space, one can consider on $\mathfrak{sl}(2,\mathbb{C})$ a vector space structure on $\mathbb{C}$ or on $\mathbb{R}$. If you want $\{t_1, t_2, t_3\}$ to be a basis, you have to consider the complex linear combinations. If you really want real linear combinations, you need a larger basis.

We consider the following basis for $\mathfrak{sl}(2,\mathbb{C})$: EQSooORIBooAsgdDp

$$h_3 = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \tag{55.335a}$$

$$h_+ = \begin{pmatrix} 0 & \sqrt{2}/2 \\ 0 & 0 \end{pmatrix} \tag{55.335b}$$

$$h_- = \begin{pmatrix} 0 & 0 \\ \sqrt{2}/2 & 0 \end{pmatrix} \tag{55.335c}$$

They satisfy the commutation relations SUBEQSooXMMVooKtnRXW

$$[h_3, h_+] = h_+ \tag{55.336a}$$

$$[h_3, h_-] = -h_- \tag{55.336b}$$

$$[h_+, h_-] = h_3. \tag{55.336c}$$

LEMooDGUYooPUkDNr

Lemma 55.65 ([897])**.**
Let (ρ, V) be a representation of $\mathfrak{sl}(2,\mathbb{C})$. If V_λ is the eigenspace of the eigenvalue $\lambda \in \mathbb{C}$ for $\rho(h_3)$, then

$$\rho(h_+)V_\lambda \subset V_{\lambda+1} \tag{55.337a}$$

$$\rho(h_-)V_\lambda \subset V_{\lambda-1} \tag{55.337b}$$

Proof. Let $w \in V_\lambda$. We test the eigenvalue of $\rho(h_3)$ on $\rho(H_+)w$:

$$\phi(h_3)\rho(h_+)w = \big([\rho(h_3), \rho(h_+)] + \rho(h_+)\rho(h_3)\big) \tag{55.338a}$$

$$= \rho(h_+)w + \lambda\rho(h_+)w \tag{55.338b}$$

$$= (\lambda + 1)\rho(h_+)w, \tag{55.338c}$$

so that $\phi(h_+)w \in V_{\lambda+1}$.

The computation is the same for the other one. □

LEMooWXDYooUyijnm

Lemma 55.66.
Let (V,ρ) be a finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ over the complex vector space V. There exists $\lambda_0\in\mathbb{C}$ such that $V_{\lambda_0}\neq\{0\}$ and $\rho(h_+)V_{\lambda_0}=\{0\}$.

Proof. A vector $w\in V$ belong to V_λ if $\rho(h_3)w=\lambda w$, which meas that $\big(\rho(h_3)-\lambda\,\mathrm{Id}\big)w=0$. The equation $\det\big(\rho(h_3)-\lambda\,\mathrm{Id}\big)$ has (at least) one solution $\lambda\in\mathbb{C}$. So there exists $\lambda\in\mathbb{C}$ such that $V_\lambda\neq\{0\}$.

Let λ be such a number and a non vanishing vector $w\in V_\lambda$.

Now the sequence of elements $w_k=\rho(h_+)w$ satisfy $w_k\in V_{\lambda+k}$. Since V is finite dimensional only a finite number of the $V_{\lambda+k}$ are different to $\{0\}$. The space V_λ on the other hand contains $w\neq 0$. Let k_0 be the lowest natural such that $V_{\lambda+k_0}=\{0\}$. What we have is $V_{\lambda+k_0-1}\neq\{0\}$ and $V_{\lambda_0+k_0}=\{0\}$.

The proposition is done with $\lambda_0=\lambda+k_0$. □

PROPooZCAOooHHGxQk

Proposition 55.67 ([897]).
Let (ρ,V) be a finite dimensional complex representation of $\mathfrak{sl}(2,\mathbb{C})$. Let λ_0 be such that $V_{\lambda_0}\neq\{0\}$ and $\rho(h_+)V_{\lambda_0}=\{0\}$. Let $w_0\in V_{\lambda_0}$ and

$$w_k=\rho(h_-)^k w_0. \tag{55.339}$$

Then

ITEMooBPPFooKdGyqO
(1) $w_k\in V_{\lambda_0-k}$

ITEMooHNULooHoTgEa
(2) $\rho(h_+)w_k=\frac{1}{2}k(2\lambda_0+1-k)w_{k-1}$

ITEMooHDAPooClASpy
(3) There exists $n\in\mathbb{N}$ such that $w_n\neq 0$ and $w_{n+1}=0$.

ITEMooJBZFooGqallS
(4) $\lambda_0=n/2$.

Proof. Notice that the existence of λ_0 such that $V_{\lambda_0}\neq\{0\}$ and $\rho(h_+)V_{\lambda_0}\neq\{0\}$ is provided by lemma 55.66.

(1) **For (1)** By recursion, using lemma 55.65.

(2) **For (2)** We'll have a recursion. Just to be clear here are two facts that are not yet proved:
- The spaces V_{λ_0+k} are one-dimensional.
- $\rho(h_+)w_{k+1}$ is a multiple of w_k.

We will now prove by recursion that the first fact is true. The second one is, in general, false. We will see later that it is true when the representation is irreducible.

Ok. So let's begin our work. For $k=0$ we already have

$$\rho(h_+)w_0=0. \tag{55.340}$$

Let work out the case of $k=1$.

$$\begin{aligned}\rho(h_+)w_1&=\rho(h_+)\rho(h_-)w_0 && (55.341a)\\ &=\big(\underbrace{[\rho(h_+),\rho(h_-)]}_{=\rho(h_3)}+\rho(h_-)\rho(h_+)\big)w_0 && (55.341b)\\ &=\rho(h_3)w_0 && (55.341c)\\ &=\lambda_0 w_0. && (55.341d)\end{aligned}$$

For the recursion, suppose that $\rho(h_+)w_k=f(k)w_{k-1}$ for some function $f\colon\mathbb{N}\to\mathbb{C}$. Then we compute $\rho(h_+)w_{k+1}$:

$$\begin{aligned}\rho(h_+)w_{k+1}=\rho(h_+)\rho(h_-)w_k&=\big(\underbrace{[\rho(h_+),\rho(h_-)]}_{=\rho(h_3)}+\rho(h_-)\rho(h_+)\big)w_k && (55.342a)\\ &=\lambda_0 w_k+f(k)\rho(h_-)w_{k-1} && (55.342b)\\ &=\big(\lambda_0+f(k)\big)w_k. && (55.342c)\end{aligned}$$

This shows that $\rho(h_+)w_{k+1}$ is a multiple of w_k and that the proportionality factor $f(k)$ satisfy [SUBEQSooHGQNooRjMCap]

$$\begin{cases} f(k+1) = f(k) + \lambda_0 - k & (55.343a) \\ f(1) = \lambda_0. & (55.343b) \end{cases}$$

The function f is defined by recursion and you see that at each step k, we substrat k and add λ_0. The guess is

$$f(k) = k\lambda_0 - \frac{k(k-1)}{2}. \tag{55.344}$$

Check that this satisfy (55.343).

(3) **For (3)** The sequence of elements $w_k \in V_{\lambda_0 - k}$ has to finish on 0 because the space V is finite dimensional.

(4) **For (4)** Let $n \in \mathbb{N}$ such that $w_n \neq 0$ and $w_{n+1} = 0$. This means $f(k+1) = 0$. Solving

$$(n+1)\big(\lambda_0 + \frac{1-(n+1)}{2}\big) = 0 \tag{55.345}$$

we get $\lambda_0 = n/2$.

□

This is quite an achievement because we proved not only that $\rho(h_3)$ has a real eigenvalue, but that it has an eigenvalue in $\mathbb{N}/2$.

[PROPooDAIQooPZVjju]

Proposition 55.68.
Let λ_0 and w_0 be as before. We suppose that the representation is irreducible. Then

$$V = \mathrm{Span}\{\rho(h_-)^k w_0\}_{k=0,\dots,2\lambda_0} \tag{55.346}$$

The eigenspaces of $\rho(h_3)$ are one-dimensional.

Proof. Let $W = \mathrm{Span}\{\rho(h_-)^k w_0\}_{k=0,\dots,2\lambda_0}$. This space is invariant under ρ because of the definitions and the proposition 55.67:

$$\begin{aligned} \rho(h_3)w_k &= (\lambda_0 - k)w_k \\ \rho(h_+)w_k &= \frac{1}{2}k(2\lambda_0 + 1 - k)w_{k-1} \\ \rho(h_-)w_k &= w_{k+1} \end{aligned} \tag{55.347}$$

with the convention that w_{k+1} could be 0.

Since W is a non trivial invariant subspace, it has to be V. So $W = V$. □

55.69.
In the proposition 55.67 and 55.68, the vectors w_k are more or less enumerated in the reverse order: the larger k is, the lower is the eigenvalue. That leads to missleading formula like $\rho(h_-)v_k = v_{K+1}$. In the following theorem, we make it in the correct order and one has to think v_m as being $w_{\lambda_0 - m}$.

Notice that up to now, the results we have collected are «if a representation of $\mathfrak{sl}(2,\mathbb{C})$ exists». The next theorem 55.70 will show that a representation exists.

Here is the theorem which provides every irredicible finite-dimensional representations of the Lie algebra $\mathfrak{sl}(2,\mathbb{C}) = \mathfrak{su}(2)_{\mathbb{C}}$.

[THOooSRQYooXQDZpT]

Theorem 55.70.
Let $j \in \mathbb{N}/2$. Let V_j be a complex vector space of dimension $2j+1$; we label a basis of V_j in the following way: $\{v_m\}_{m=j,j-1,\dots,-j}$.

We define the map $\rho_j\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V_j)$ *by*

$$\rho_j(h_3)v_m = mv_m, \tag{55.348a}$$

$$\rho_j(h_+)v_m = \begin{cases} 0 & \text{if } m = j \\ \frac{1}{2}(j-m)(j+m+1)v_{m+1} & \text{otherwise}, \end{cases} \tag{55.348b}$$

$$\rho_j(h_-)v_m = \begin{cases} v_{m-1} & \text{if } m \neq -j \\ 0 & \text{if } m = -j. \end{cases} \tag{55.348c}$$

Two statements.

(1) The map ρ_j *is a representation of* $\mathfrak{sl}(2,\mathbb{C})$.

(2) Every finite dimensional complex irreducible representation is isomorphic to ρ_j *for some* $j \in \mathbb{N}/2$.

Proof. For the first item we have to check the algebra of $\mathfrak{su}(2)$ given by (55.336). There are three computations.

We begin to check $[\rho_j(h_3), \rho_j(h_+)] = \rho_j(h_+)$. The bracket in the left-hand side is the commutator of operators in $\mathrm{End}(V)$. We have:

$$\big(\rho_j(h_3)\rho_j(h_+) - \rho(h_+)\rho(h_3)\big)v_m = \rho_j(h_3)\frac{1}{2}(j-m)(j+m+1)v_{m+1} - m\rho_j(h_+)v_m \tag{55.349a}$$

$$= (m+1)\frac{1}{2}(j-m)(j+m+1)v_{m+1} \tag{55.349b}$$

$$- m\frac{1}{2}(j-m)(j+m+1)v_{m+1} \tag{55.349c}$$

$$= \frac{1}{2}(j-m)(j+m+1)v_{m+1} \tag{55.349d}$$

$$= \rho_j(h_+)v_m. \tag{55.349e}$$

The two other ones are checks with the same kind of computations.

For the second item, we consider an irreducible finite-dimensional representations (ρ, V) of $\mathfrak{sl}(2,\mathbb{C})$. Combining the propositions 55.67 and 55.68 we have:

— A vector $w_0 \in V$ such that $\rho(h_+)w_0 = 0$.

— Letting $n = 2\lambda_0$ we have $n \in \mathbb{N}$,

— we let $j = \lambda_0$ (pure notational purpose),

— $V = \mathrm{Span}\{w_k = \rho(h_-)^k w_0\}$,

— $w_k \in V_{\lambda_0 - k}$ where V_λ is the eigenspace of $\rho(h_3)$ for the eigenvalue λ,

— $w_k \neq 0$ if and only if $k = 0, \dots, n$, so $\dim(V) = 2n+1$

— $\rho(h_3)w_k = (\lambda_0 - k)w_k$

— $\rho(h_+)w_k = \frac{1}{2}k(2\lambda_0 + 1 - k)w_{k-1}$

We choose $j = \lambda_0 \in \mathbb{N}/2$. Do you believe that the map

$$\begin{aligned} \phi\colon V &\to V_j \\ w_k &\mapsto v_{j-k} \end{aligned} \tag{55.350}$$

provides an equivalence of representations between ρ_j and ρ ? No ? Ok. We check that for every $X \in \mathfrak{sl}(2,\mathbb{C})$ we have

$$\phi \circ \rho(X) = \rho_j(X) \circ \phi. \tag{55.351}$$

For $X = h_3$ we have

$$\phi \circ \rho(h_3)w_k = \phi\big((\lambda_0 - k)w_k\big) = (\lambda_0 - k)v_{j-k} = (j-k)v_{j-k} \tag{55.352}$$

while

$$\rho_j(h_3)\phi(w_k) = \rho_j(h_3)v_{j-k} = (j-k)v_{j-k}. \tag{55.353}$$

Ok for the first one. Next: $X = h_+$. We have

$$\phi \circ \rho(h_+)w_k = \phi\big(\frac{1}{2}k(2j+1-k)w_{k-1}\big) = \frac{1}{2}k(2j+1-k)v_{j-k+1} \tag{55.354}$$

while

$$\rho_j\phi(w_k) = \rho_j(h_+)v_{j-k} = \frac{1}{2}\big(j-(j-k)\big)\big(j+(j-k)+1\big)v_{j-k+1} = \frac{1}{2}k(2j-k+1)v_{j-k+1}. \tag{55.355}$$

Ok again. And last one: $X = h_-$; we have

$$\phi \circ \rho(h_-)w_k = \phi(w_{k+1}) = v_{j-k-1} \tag{55.356}$$

while

$$\rho_j(h_-)\phi(w_k) = \rho_j(h_-)v_{j-k} = v_{j-k-1}. \tag{55.357}$$

Done [31]. □

55.71.
The representations ρ_j of the theorem 55.70 are not yet hermitian for two reasons.

— The representations we expect to be hermitian are the ones of $\mathfrak{su}(2)$. The basis $\{h_2, h_+, h_-\}$ of $\mathfrak{su}(2)_{\mathbb{C}}$ defined by (55.335) is made of elements which do not belong to $\mathfrak{su}(2)$.

— We did not defined a scalar product on V_j; thus the notion of «hermitian» makes no sene.

55.14.6 Representations of $\mathfrak{su}(2)$

We know every representations of $\mathfrak{su}(2)_{\mathbb{C}}$ by the theorem 55.70. Let $\rho\colon \mathfrak{su}(2) \to \mathrm{End}(V)$ be an irreducible representation of $\mathfrak{su}(2)$. By lemma 51.8, there exists an irreducible representation $\rho'\colon \mathfrak{su}(2)_{\mathbb{C}} \to \mathrm{End}(V)$ such that $\rho = \rho'|_{\mathfrak{su}(2)}$.

Thus there exists a j such that $\rho(X) = \rho_j(X)$.

55.14.7 Haar measure on $\mathrm{SU}(2)$

The quaternion field $\mathbb{H}$ can be embed in $\mathbb{M}_2(\mathbb{C})$ as a genera element reads

$$q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{55.358}$$

with $\alpha, \beta \in \mathbb{C}$. Under that isomorphism, we have

$$|q|^2 = |\alpha|^2 + |\beta|^2 = \det q.$$

Thus we have the identification

$$\mathrm{SU}(2) = \{q \in \mathbb{H} \text{ tel que } |q| = 1\}. \tag{55.359}$$

We can act on $\mathbb{H}$ by $\mathrm{SU}(2) \times \mathrm{SU}(2)$ by

$$(u, v) \cdot q = uqv^{-1} \tag{55.360}$$

for every $(u, v) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$ and $q \in \mathbb{H}$. That action defines an homomorphism from $\mathrm{SU}(2) \times \mathrm{SU}(2)$ onto $O(4)$.

31. Now that we reached the end, I recognize that I did not belive neither until the last check.

Proposition 55.72.
The previously defined homomorphism

$$\phi\colon \mathrm{SU}(2)\times\mathrm{SU}(2)\to O(4).$$

is surjective over $\mathrm{SO}(4)$ *(which is the identity component of* $O(4)$*) and, moreover, the kernel is* $\big\{(e,e),(-e,-e)\big\}$.

Proof. The group $\mathrm{SU}(2)\times\mathrm{SU}(2)$ being connected, its image can only be included in $\mathrm{SO}(4)$. Let us first determine the kernel of ϕ. If $(u,v)\in\ker\phi$, we have $uqv^{-1}=q$ for every $q\in\mathbb{H}$. In particular, with $q=1$, we find $u=v$. Then the relation $uqu^{-1}=q$ means that u belongs to the center of $\mathbb{H}$, which is $\mathbb{R}$. We conclude that $u=\pm1$. That proves that $\ker\phi=\big\{(e,e),(-e,-e)\big\}$.

The differential $(d\phi)_{(e,e)}$ is an homomorphism

$$d\phi\colon \mathfrak{su}(2)\oplus\mathfrak{su}(2)\to\mathfrak{so}(4).$$

Let $(S,T)\in\mathfrak{su}(2)\oplus\mathfrak{su}(2)$, we have

$$d\phi(S,T)q=\frac{d}{dt}\Big[\phi(e^{t(S,T)})q\Big]_{t=0}=\frac{d}{dt}\Big[\phi(e^{tS},e^{tT})q\Big]_{t=0}=\frac{d}{dt}\Big[e^{tS}qe^{-tT}\Big]_{t=0}=Sq-qT,$$

on which one sees that $d\phi$ is injective. Moreover we have $\dim\big(\mathfrak{su}(2)\oplus\mathfrak{su}(2)\big)=6=\dim\mathfrak{so}(4)$. An injective map between vector space of same dimension being an isomorphism, the image of ϕ contains a neighborhood of identity in $\mathrm{SO}(4)$. From connectedness of $\mathrm{SO}(4)$, that neighborhood generates the whole group (see proposition 52.5), so that ϕ is in fact surjective. □

Since the map $\phi\colon \mathrm{SU}(2)\times\mathrm{SU}(2)\to\mathrm{SO}(4)$ is a surjective homomorphism with a discrete kernel, we have an isomorphism at the algebra level:

$$\mathfrak{so}(4)\simeq\mathfrak{su}(2)\oplus\mathfrak{su}(2).$$

55.14.8 Building some representations for $\mathrm{SU}(2)$

Since $\mathrm{SU}(2)$ acts on $\mathbb{C}^2$, we can build a representation of $\mathrm{SU}(2)$ on functions on $\mathbb{C}^2$. We define $T\colon SU(2)\to\mathrm{End}\big(C^\infty(\mathbb{C}^2)\big)$ by

$$(T(U)f)(\xi)=f(U^{-1}\xi),$$

if $f\in C^\infty\big(\mathbb{C}^2\big)$, $\xi\in\mathbb{C}^2$ and $U\in SU(2)$.

Let V_j be the space of the homogeneous polynomials of degree j on $\mathbb{C}^2$; a basis of this space is given by the ϕ_{pq}, $p+q=2j$ defined by

$$\phi_{pq}(\xi)=\xi_1^p\xi_2^q \tag{55.361}$$

($\xi=\xi_1+i\xi_2$). If j is fixed, we will often write ϕ_m instead of ϕ_{pq}. The signification is $p=j+m$, $q=j-m$, and m takes its values in $-j,\dots,j$. Note that $p-q=2m$. It is clear that if A is any invertible 2×2 matrix , and $f\in V_j$, then

$$\rho(A)f:=f(A^{-1}\,\cdot\,)$$

is still an element of V_j. This representation ρ is defined on the whole $C^\infty(\mathbb{C}^2)$. We will descent it to V_j later. Now, we fix j and a m between $-j$ and j.

Consider the diagonal matrix

$$U_{-\theta}=\begin{pmatrix}e^{-i\theta}&0\\0&e^{i\theta}\end{pmatrix}\in\mathrm{SU}(2).$$

One has

$$\big(\rho(U_{-\theta})\phi_{pq}\big)(\xi)=\phi_{pq}\begin{pmatrix}e^{i\theta}\xi_1\\e^{-i\theta}\xi_2\end{pmatrix}=e^{pi\theta}e^{-qi\theta}\xi_1^p\xi_2^q=e^{2mi\theta}\phi_{pq}(\xi). \tag{55.362}$$

First conclusion: the ϕ's are eigenvectors of $\rho(U_{-\theta})$ because

$$\rho(U_{-\theta})\phi_m = e^{2mi\theta}\phi_m.$$

Second, the trace of $\rho(U_{-\theta})$ is

$$\chi_j(\theta) = \sum_{m=-s}^{s} e^{2mi\theta}. \tag{55.363}$$

By the way, the χ_j are the characters of the representation ρ.

From considerations about the Haar [32] invariant measure on SU(2), one knows that the good notion product between functions is:

$$(f_1, f_2)_{\mathrm{SU}(2)} = \frac{2}{\pi}\int_0^{\pi} f_1(\theta)\overline{f_2(\theta)}\sin^2\theta\, d\theta, \tag{55.364}$$

so that $(\chi_j, \chi_j) = 1$. This and the fact that SU(2) is compact make the theorem of Peter-Weyl (cf. [900]) applicable, thus the restrictions of ρ to the V_j's are irreducible and moreover, these provide *all* the irreducible representations.

55.14.9 Special case: $j = \frac{1}{2}$

Consider a matrix $A \in \mathrm{SU}(2)$:

$$A = \begin{pmatrix} \overline{\alpha} & -\beta \\ \overline{\beta} & \alpha \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}. \tag{55.365}$$

A basis of $V_{\frac{1}{2}}$ is given by ϕ_{10} and ϕ_{01}. Let us see how $\rho(A)$ acts on. Since

$$A^{-1}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \alpha\xi_1 + \beta\xi_2 \\ -\overline{\beta}\xi_1 + \overline{\alpha}\xi_2 \end{pmatrix},$$

we find

$$\begin{aligned} (\rho(A)\phi_{10})(\xi) &= \alpha\xi_1 + \beta\xi_2 = (\alpha\phi_{10} + \beta\phi_{01})(\xi) \\ (\rho(A)\phi_{01})(\xi) &= -\overline{\beta}\xi_1 + \overline{\alpha}\xi_2 = (-\overline{\beta}\phi_{10} + \overline{\alpha}\phi_{01})(\xi). \end{aligned} \tag{55.366}$$

Thus in the basis $\{\phi_{10}, \phi_{01}\}$, the matrix of $\rho(A)$ is given by

$$\rho(A) = \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} = \overline{A}. \tag{55.367}$$

Up to here, we were looking at the representation ρ of SU(2) on the whole set of functions on $\mathbb{C}^2$, and more precisely, its restriction to V_j. We could define the representation $\rho_{\frac{1}{2}}$ as $\rho_{\frac{1}{2}} = \rho|_{V_{\frac{1}{2}}}$, but we will not do it. Our definition is

$$\rho_{\frac{1}{2}}(A) = \rho(\overline{A})|_{V_{\frac{1}{2}}}. \tag{55.368}$$

Note that

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} A \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \overline{A},$$

thus the representation $A \to \rho(\overline{A})|_{V_{\frac{1}{2}}}$ is equivalent to $A \to \rho(A)|_{V_{\frac{1}{2}}}$. This equivalence can also be seen because these two representations have the same characters [33].

The basis ϕ_{pq} is orthogonal; we will build an orthonormal one: $e_m(\xi)$ is the vector whose coordinates are

$$e_m^j(\xi) = \frac{\xi_1^{j+m}\xi_2^{j-m}}{\sqrt{(j+m)!(j-m)!}} \tag{55.369}$$

32. In project...
33. In project...

for $m = -j, -j+1, \dots, j$. The metric to take in order to define (e_m, e_n) is the unique one on V_j which is SU(2)-invariant.

The Newton's formula for the binomial yields:

$$\begin{aligned}\sum_{m=-s}^{s} e_m(\xi)\overline{e_m(\eta)} &= \sum \frac{\xi_1^{j+m}\xi_2^{j-m}\overline{\eta}_1^{j+m}\overline{\eta}_2^{j-m}}{(j+m)!(j-m)!} \\ &= \frac{1}{(2j)!}(\xi_1\overline{\eta}_1 + \xi_2\overline{\eta}_2)^2 \\ &= \frac{1}{(2j)!}\langle \xi, \eta\rangle^{2j}.\end{aligned} \tag{55.370}$$

But we know that $A \in \mathrm{SU}(2)$ preserves the scalar product: $\langle A\xi, A\eta\rangle = \langle \xi, \eta\rangle$. Therefore:

$$\sum (\rho_j(A)e_m)(\xi)\overline{(\rho_j(A)e_m)(\eta)} = \sum e_m(\xi)\overline{e_m(\eta)}. \tag{55.371}$$

eq:produit_e_m

Now, instead of considering the matrices $\rho_j(A)$ on V_j for the basis ϕ_m, we looks at the ones with respect to the basis e_m:

$$\rho_j(A)e_m = r(A)_m^k e_k; \tag{55.372}$$

in others words, we looks at the representation $A \to r(A)$. The equations (55.371) makes

$$\sum_{m=-j}^{j} \left(r(A)_m^l e_l(\xi)\overline{r(A)_m^k e_k(\eta)} - \delta_m^l e_l(\xi)\delta_m^k e_k(\eta)\right) = 0.$$

Since the functions

$$\begin{aligned} e_k \otimes \overline{e_l} \colon \mathbb{C}^2 \times \mathbb{C}^2 &\to \mathbb{C} \\ (\xi, \eta) &\mapsto e_k(\xi)\overline{e_l(\eta)} \end{aligned} \tag{55.373}$$

are linearly independent, one gets $\sum_m r(A)_l^k \overline{r(A)_m^l} = \delta^{kl}$, or

$$r(A)r(A)^* = \mathbb{1}, \tag{55.374}$$

the conclusion is that in this basis, the matrices $\rho_j(A)$ are unitary.

55.14.10 Clebsch-Gordan

From the knowledge of the characters of ρ_j, one can decompose the product $\rho_s \otimes \rho_r$ into irreducible representations. For example,

$$V_{\frac{1}{2}} \otimes V_{\frac{1}{2}} = V_0 \oplus V_1.$$

More generally,

$$V_s \otimes V_r = V_{|r-s|} \oplus V_{|r-s|+1} \oplus \dots \oplus V_{r+s}. \tag{55.375}$$

For this reason, the representation ρ_j is sometimes called the **spin** j representation of SU(2).

55.15 Representations of SO(3)

The group SO(3) is strongly linked with $SU(2)$ by the following property:

$$\mathrm{SO}(3) = \frac{SU(2)}{\mathbb{Z}_2}. \tag{55.376}$$

proved in proposition 55.47.

lem:SO_3

Lemma 55.73.
A representation ρ_j of $SU(2)$ is a representation of SO(3) *if and only if $\rho_j(X) = \mathrm{Id}$ for any X in the kernel of the homomorphism $SU(2) \to \mathrm{SO}(3)$, namely: $\rho_j(\pm\mathbb{1}) = \mathrm{Id}$.*

Proof. We consider $\rho_j : SU(2) \to \operatorname{End} V_j$. By proposition 55.47 we have $SO(3) = \mathrm{SU}(2)/\mathbb{Z}_2$ and there exists a group homomorphism[34] $\psi : SU(2) \to \mathrm{SO}(3)$ such that $\psi(\mathbb{1}) = \psi(-\mathbb{1}) = \mathbb{1}$, which is an important equation because it ensures us that the rest of the expressions are well defined with respect to the class representative.

If $\rho_j(-\mathbb{1}) = \mathbb{1}$, we define $d_j : \mathrm{SO}(3) \to \operatorname{End} V$ by $d_j([x]) = \rho_j(x)$ (check that this is well defined). With this,

$$d_j([x])d_j([y]) = \rho_j(x)\rho_j(y) = \rho_j(xy) = d_j([xy]).$$

Now let us suppose that $d_j([x]) = \rho_j(x)$ is a representation. Thus

$$\rho_j(x) = d_j([x]) = d_j([-x]) = \rho_j(-x) = \rho_j(-\mathbb{1})\rho_j(x),$$

so $\rho_j(-\mathbb{1}) = \mathrm{Id}_{V_j}$. □

Moreover, any representation of $\mathrm{SO}(3)$ comes from a representation $\tilde{\rho}$ of $SU(2)$ by setting $\tilde{\rho}(-\mathbb{1}) = \mathrm{Id}$ and $\tilde{\rho}(x) = \rho([x])$.

Now, we research the representations of $SU(2)$ for which the matrix $-\mathbb{1}$ is represented by the identity operator. These will be representations of $\mathrm{SO}(3)$. The spin j representations of $SU(2)$ is given by

$$\rho_j(X)\phi_{pq}(\xi) = \phi_{pq}(X^{-1}\xi). \tag{55.377}$$

With $X = -\mathbb{1}$, this gives: $\phi_{pq}(-\xi) = (-1)^{p+q}\phi_{pq}(\xi)$. If we want it to be equal to $\phi_{pq}(\xi)$, we need $p + q = 2j$ even. This is true if and only if $j \in \mathbb{N}$.

NORMooHWAYooPlSDOp

55.74.
The conclusion is that the irreducible representations of $\mathrm{SO}(3)$ are the integer spin irreducible representations of $\mathrm{SU}(2)$. Note that the non relativistic mechanics has $\mathrm{SO}(3)$ as group of space symmetry. Thus there are no hope to find any half integer spin in a non relativistic theory.

55.16 Lorentz group

Definition 55.75.
We consider the vector space $\mathbb{R}^4$ with its usual scalar product $\langle .,. \rangle$ which is positive defined. Using the matrix[35]

$$\eta = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \tag{55.378}$$

we introduce the ***Minkowskian product***

$$x \cdot y = \langle \eta x, y \rangle. \tag{55.379}$$

EQooQAXNooXbGUOV

This product is not positive defined and is often called «pseudo-scalar product».

LEMooICEYooNcjJjD

Lemma 55.76.
A map $\Lambda : \mathbb{R}^4 \to \mathbb{R}^4$ such that

$$\Lambda x \cdot \Lambda y = x \cdot y \tag{55.380}$$

for every $x, y \in \mathbb{R}^4$ is linear.

Proof. The bilinear form

$$\begin{aligned} b : \mathbb{R}^4 \times \mathbb{R}^4 &\to \mathbb{R} \\ x, y &\mapsto x \cdot y \end{aligned} \tag{55.381}$$

is non degenerated. An element of $\mathrm{O}(3,1)$ satisfy $b(\Lambda x, \Lambda y) = b(x, y)$. Thus the theorem 8.157 says that Λ must be linear. □

34. Defined and studied in proposition 55.46.
35. The plus/minus convention is, by far, not universal. As an example, [844] uses the opposite.

Lemma-Definition 55.77.
The set of maps $\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4$ *such that*

$$\Lambda x \cdot \Lambda y = x \cdot y \tag{55.382}$$

EQooLPXWooNgrAXs

for every $x, y \in \mathbb{R}^4$ *is a group.*

This group is named the ***Lorentz group*** *and is denoted by* $\mathrm{O}(3,1)$ *or* L*.*

Proof. From lemma 55.76 we know that the elements of $\mathrm{O}(3,1)$ are linear operators.

(1) The identify is part of $\mathrm{O}(3,1)$.

(2) The product is associative.

(3) The only tricky part is to prove that if $\Lambda \in \mathrm{O}(3,1)$, then Λ is invertible and $\Lambda^{-1} \in \mathrm{O}(3,1)$.

Let $z \neq 0 \in \mathbb{R}^4$ being such that $\Lambda z = 0$, then there exists $y \in \mathbb{R}^4$ such that $z \cdot y \neq 0$ while obviously $\Lambda z \cdot \Lambda y = 0$. Thus every element in $\mathrm{O}(3,1)$ is invertible.

Let $\Lambda \in \mathrm{O}(3,1)$, $x, y \in \mathbb{R}^4$. Using the fact that $\Lambda \in \mathrm{O}(3,1)$ we have

$$x \cdot y = \Lambda(\Lambda^{-1}x) \cdot \Lambda(\Lambda^{-1}y) = \Lambda^{-1}x \cdot \Lambda^{-1}y, \tag{55.383}$$

so that $\Lambda^{-1} \in \mathrm{O}(3,1)$. □

55.16.1 Adjoint map

We have a notational issue here. We already defined the adjoint map A^* by

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \tag{55.384}$$

for the usual scalar product. Since the main product we consider now in $\mathbb{R}^4$, is the Minkowskian one, we will define Λ^* by $\Lambda x \cdot y = x \cdot \Lambda^* y$ and leave the notation A^t for the adjoint with respect to the usual scalar product.

Proposition-Definition 55.78.
Let $A \in \mathcal{L}(\mathbb{R}^4)$*. There exists a unique operator* $B \in \mathcal{L}(\mathbb{R}^4)$ *such that*

$$Ax \cdot y = x \cdot By \tag{55.385}$$

for every $x, y \in \mathbb{R}^4$*.*

This operator is called ***adjoint****, is written* A^* *and is given by*

$$A^* = \eta A^t \eta. \tag{55.386}$$

EQooPFPGooXiGcXs

Proof. For the existence, we just have to check that $B = \eta A^t \eta$ works. Using the fact that $\eta^t = \eta$ and $\eta^2 = \mathbb{1}$,

$$\begin{aligned} x \cdot \eta A^t \eta y &= \langle \eta x, \eta A^t \eta y\rangle && (55.387a)\\ &= \langle x, A^t \eta y\rangle && (55.387b)\\ &= \langle Ax, \eta y\rangle && (55.387c)\\ &= \langle \eta Ax, y\rangle && (55.387d)\\ &= Ax \cdot y. && (55.387e) \end{aligned}$$

For the unicity, we suppose $Ax \cdot y = x \cdot By$ for every $x, y \in \mathbb{R}^4$. We have

$$x \cdot By = Ax \cdot y = x \cdot \eta A^t \eta y \tag{55.388}$$

Since the product is non degenerate, this implies $B = \eta A^t \eta$. □

LEMooVRWJooPsDRwU

Lemma 55.79.
The adjoint operator satisfy $(\Lambda^*)^* = \Lambda$*.*

Proof. We use the formula (55.386): $(\Lambda^*)^* = \eta(\Lambda^*)^t\eta = \eta(\eta\Lambda^t\eta)^t\eta = \eta\eta^t\Lambda\eta^t\eta = \Lambda$. □

LEMooDLWDooWCXlWq

Lemma 55.80 ([1])**.**
If $\Lambda \in \mathcal{L}(\mathbb{R}^4, \mathbb{R}^4)$, the following are equivalent:

(1) $\Lambda \in \mathrm{O}(3,1)$, ITEMooWHGKooFPfujT
(2) $\Lambda^* \in \mathrm{O}(3,1)$, ITEMooNISDooMajEMS
(3) $\Lambda^*\Lambda = \mathbb{1}$, ITEMooNLZGooUINRiP
(4) $\Lambda\Lambda^* = \mathbb{1}$, ITEMooFFRVooOwLmnz
(5) $\Lambda^t\eta\Lambda = \eta$, ITEMooOYTDooCWImBJ
(6) $\Lambda\eta\Lambda^t = \eta$. ITEMooAEEYooDiJuEi

Proof. We prove the equivalences.

(1) **(1) implies (3)** Since $\Lambda \in \mathrm{O}(3,1)$ we have $x \cdot y = \Lambda x \cdot \Lambda y = x \cdot \Lambda^*\Lambda y$ for every $x, y \in \mathbb{R}^4$. This implies $\Lambda^*\Lambda = \mathbb{1}$.

(2) **(3) implies (1)** Since $\Lambda^*\Lambda = \mathbb{1}$ and the fact that $(\Lambda^*)^*$ (lemma 55.79) we have

$$x \cdot y = \Lambda^*\Lambda x \cdot y = \Lambda x \cdot \Lambda y. \tag{55.389}$$

This shows that $\Lambda \in \mathrm{O}(3,1)$.

(3) **(3) if and only if (4)** Le corolaire 4.51 nous dit que $AB = \mathbb{1}$ if and only if $BA = \mathbb{1}$.

(4) **(2) if and only if (4)** Same proof as (1) if and only if (3).

At this point, we have the equivalence between (1), (2), (3) and (4).

(1) **(3) implies (5)** We plug the expression (55.386) of the adjoint in the equation $\Lambda^*\Lambda = \mathbb{1}$:

$$\mathbb{1} = \Lambda^*\Lambda = \eta\Lambda^t\eta\Lambda. \tag{55.390}$$

Multiplying by η on the left, we get the result.

(2) **(5) implies (3)** We write $\Lambda^t\eta\Lambda = \eta$ and we multiply by η.

(3) **(4) if and only if (6)** It is the same as (5) if and only if (3).

□

55.16.2 Structure

LEMooBMDUooHoYYae

Lemma 55.81.
Elements $\Lambda \in \mathrm{O}(3,1)$ satisfy $\det(\Lambda) = \pm 1$.

Proof. Taking the determinant on both sides of $\Lambda^*\eta\Lambda = \eta$ (lemma 55.80), we get $\det(\Lambda^*)\det(\Lambda) = 1$. Since the operator Λ is real, the proposition 8.39 says that $\det(\Lambda^*) = \det(\Lambda)^* = \det(\Lambda)$. Thus $\det(\Lambda)^2 = 1$. The result follows. □

LEMooHRNXooJOgfpy

Lemma 55.82.
Let $\Lambda \in \mathrm{O}(3,1)$. We have

$$\Lambda_{00}^2 - \sum_{k=0}^{3}(\Lambda_{k0})^2 = 1 \tag{55.391a}$$ SUBEQooFLECooUFywOy

$$\Lambda_{00}^2 - \sum_{k=0}^{3}(\Lambda_{0k})^2 = 1 \tag{55.391b}$$ SUBEQooBLTOooPUTztZ

In particular, $\Lambda_{00}^2 \geqslant 1$.

Proof. Just write the 00 component of the equation $\Lambda^t\eta\Lambda = \eta$ (lemma 55.80):

$$1 = \sum_{kl} \Lambda^t_{0k} \underbrace{\eta_{kl}}_{\eta_{kk}\delta_{kl}} \Lambda_{l0} \tag{55.392a}$$

$$= \sum_{k} \Lambda^t_{0k}\eta_{kk}\Lambda_{k0} \tag{55.392b}$$

$$= \Lambda^t_{00}\Lambda_{00} - \sum_{k=1}^{3} \Lambda^t_{0k}\Lambda_{k0} \tag{55.392c}$$

$$= \Lambda^2_{00} + \sum_{k=1}^{3} \Lambda^2_{k0}. \tag{55.392d}$$

This is (55.391a); the same computation from $\Lambda\eta\Lambda^t = \eta$ provides (55.391b). □

LEMooEKXWooLEMBIj

Lemma 55.83.
The set

$$L^\uparrow = \mathrm{O}(3,1)^\uparrow = \{\Lambda \in \mathrm{O}(3,1) \text{ tel que } \Lambda_{00} \geqslant 1\} \tag{55.393}$$

is a subgroup of $\mathrm{O}(3,1)$.

Elements of $\mathrm{O}(3,1)^\uparrow$ *are said* ***orthochronous****. We will not use the notation* $L^\uparrow$.

Proof. Let $A, B \in L^\uparrow$. We have

$$(AB)_{00} = \sum_{k=0}^{3} A_{0k}B_{k0} = A_{00}B_{00} + \sum_{k=1}^{3} A_{0k}B_{k0}. \tag{55.394}$$

EQooBUMZooEGCDEa

Using lemma 55.82, we write

$$A^2_{00} = 1 + \sum_{k=1}^{3} A^2_{k0} \tag{55.395}$$

EQooBFWTooZNIdIP

and

$$B^2_{00} = 1 + \sum_{k=1}^{3} B^2_{0k}. \tag{55.396}$$

Since $A_{00} \geqslant 1$, taking the square root of (55.395) does not require additional caution:

$$A_{00} > \sqrt{\sum_{k=1}^{3} A^2_{k0}}. \tag{55.397}$$

The same holds for B. Using these (strict) inequalities in (55.394) we have

$$(AB)_{00} > \sqrt{\sum_{k=1}^{3} A^2_{k0}}\sqrt{\sum_{k=1}^{3} B^2_{0k}} + \sum_{k=1}^{3} A_{0k}B_{k0}. \tag{55.398a}$$

If we set $a = \sum_{k=1}^{3} A_{k0}e_k$ and $b = \sum_{k=1}^{3} B_{0k}e_k$ (here $e_i \in \mathbb{R}^3$), we have

$$(AB)_{00} > \|a\|\|b\| + \langle a, b\rangle \geqslant \|a\|\|b\| - |\langle a, b\rangle| \geqslant 0 \tag{55.399}$$

because of the Cauchy-Schwarz identity, theorem 10.1. The strict inequality $(AB)_{00} > 0$ implies the inequality $(AB)_{00} \geqslant 1$ because $AB \in \mathrm{O}(3,1)$ (see lemma 55.82). □

LEMooLJMMooOXCyOl

Lemma 55.84.
About adjoint.

(1) If $\Lambda \in \mathrm{O}(3,1)$*, then* $\Lambda^* \in \mathrm{O}(3,1)$.

(2) If $\Lambda \in \mathrm{SO}(3,1)$*, then* $\Lambda^* \in \mathrm{SO}(3,1)$.

(3) If $\Lambda \in \mathrm{O}(3,1)^{\uparrow}$, *then* $\Lambda^* \in \mathrm{O}(3,1)^{\uparrow}$.

In particular, if $\Lambda \in \mathrm{SO}(3,1)^{\uparrow}$, *then* $\Lambda^* \in \mathrm{SO}(3,1)^{\uparrow}$.

Proof. Just compute Λ^*_{00} and $\det(\Lambda^*)$ with the formula (55.386). □

DEFooVQLPooWyINoc

Definition 55.85.

A ***boost*** *in the direction* x *is an element* $\Lambda \in \mathrm{SO}(3,1)^{\uparrow}$ *such that* $\Lambda(e_2) = e_2$ *and* $\Lambda(e_3) = e_2$. *In other words, this is a transformations which only involves the components* t *and* x.

The boost in the directions y *and* z *are defined in a similar way.*

A ***spatial rotation*** *is an element* $\Lambda \in \mathrm{O}(3,1)$ *such that* $\Lambda(e_0) = e_0$.

Lemma 55.86.

An operator $\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4$ *is a boost if and only if there exists* $\gamma \in \mathbb{R}$ *such that the matrix of* Λ *has the form*

$$\Lambda = \begin{pmatrix} \cosh(\gamma) & \sinh(\gamma) & 0 & 0 \\ \sinh(\gamma) & \cosh(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{55.400}$$

Proof. It is easy to see that the proposed matrix is a boost. The only difficult part is the direct sense. We suppose that Λ is a boost. The conditions $\Lambda e_2 = e_2$ and $\Lambda e_3 = e_3$ imply that the matrix of Λ has the form

$$\Lambda = \begin{pmatrix} . & . & 0 & 0 \\ . & . & 0 & 0 \\ . & . & 1 & 0 \\ . & . & 0 & 1 \end{pmatrix} \tag{55.401}$$

where the dots are to be determined. Since $\Lambda \in \mathrm{O}(3,1)$ we have

$$0 = e_0 \cdot e_2 = \Lambda e_0 \cdot \Lambda e_2 = \Lambda e_0 \cdot e_2 = -\Lambda_{20}. \tag{55.402}$$

The same shows that $\Lambda_{20} = \Lambda_{21} = \Lambda_{30} = \Lambda_{31} = 0$. The matrix of Λ is block diagonal:

$$\Lambda = \begin{pmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{55.403}$$

where a, b, c and d are still to be determined. Here are the constrains.

First the operator Λ preserve the product, so that $\Lambda e_0 \cdot \Lambda e_0 = 1$, $\Lambda e_1 \cdot \Lambda e_1 = -1$, and $\Lambda e_0 \cdot \Lambda e_1 = 0$. These conditions are translated into

SUBEQooWEJSooPWfmNS
$$a^2 - c^2 = 1 \tag{55.404a}$$

SUBEQooKLZEooCovszD
$$d^2 - b^2 = 1 \tag{55.404b}$$

SUBEQooUZQWooUxUCSe
$$ab - cd = 0. \tag{55.404c}$$

Second, the operator Λ has determinant equals to 1:

EQooJAEKooTCZaIG
$$ad - bc = 1, \tag{55.405}$$

and finally the element Λ is orthochronous: $\Lambda_{00} \geqslant 0$, so that

EQooQCMYooPhHeas
$$a \geqslant 0. \tag{55.406}$$

The proposition 15.128 about hyperbolic functions along with the conditions (55.404a) and 55.404b show that there exist $x \in \mathbb{R}$, $y \in \mathbb{R}$, $\sigma \in \{\pm 1\}$ and $\epsilon \in \{\pm 1\}$ such that

$$a = \sigma \cosh(x) \tag{55.407a}$$

$$b = \sinh(y) \tag{55.407b}$$

$$c = \sinh(x) \tag{55.407c}$$

$$d = \epsilon \cosh(y). \tag{55.407d}$$

(1) $\underline{\sigma = 1}$ The condition (55.406) show that $\sigma = 1$ because the hyperbolic cosine is always strictly positive.

(2) $\underline{\epsilon = 1}$ The determinant condition (55.405) provides

$$\epsilon \cosh(x)\cosh(y) - \sinh(x)\sinh(y) = 1. \tag{55.408}$$

If $\epsilon = -1$ we are left with

$$1 = -(\cosh(x)\cosh(y) + \sinh(x)\sinh(y)) \tag{55.409}$$

Using the formula of proposition 15.125(6) we get $1 = -\cosh(x+y)$ which is impossible because the hyperbolic cosine is always positive.

(3) $\underline{x = y}$ The orhogonality condition (55.404c) implies

$$0 = \cosh(x)\sinh(y) - \sinh(x)\cosh(y) = -\sinh(x-y). \tag{55.410}$$

Since the hyperbolic sine is bijective (proposition 15.127) we deduce $x - y = 0$ and then $x = y$.

□

We define the projection from $\mathbb{R}^4$ to $\mathbb{R}^3$ as

$$\texttt{proj}(x) = (x_1, x_2, x_3) \in \mathbb{R}^3 \tag{55.411}$$

and, if $b \in \mathbb{R}^3$ we write

$$\bar{b} = (0, b_1, b_2, b_3). \tag{55.412}$$

So if $b \in \mathbb{R}^3$ we have

$$\bar{b} \cdot x = -\langle b, \texttt{proj}(x)\rangle. \tag{55.413}$$

EQooOIBWooAyxfYz
PROPooYADMooQOTpWX

Proposition 55.87 (Standard decomposition[844])**.**
Every operator $\Lambda \in \mathrm{SO}(1,3)^\uparrow$ can be decomposed as

$$\Lambda = RLS \tag{55.414}$$

where R and S are spatial rotations and L is a boost in the x direction[36]*.*

Proof. We initiate with $a = \texttt{proj}(\Lambda e_0)$. If $a \neq 0$ we define $b_1 = \frac{a}{\|a\|} \in \mathbb{R}^3$ and we consider b_2 and b_3 in $\mathbb{R}^3$ such that $\{b_1, b_2, b_3\}$ is an orthonormal basis of $\mathbb{R}^3$ with the same orientation[37] as the canonical basis..

(1) **The first spatial rotation** Now we define the spatial rotation $R\colon \mathbb{R}^4 \to \mathbb{R}^4$ by

$$\begin{cases} Re_0 = e_0 & \text{(55.415a)} \\ Re_i = \bar{b}_i. & \text{(55.415b)} \end{cases}$$

Since the basis $\{\bar{b}_i\}$ is positive-oriented, the determinant of R is positive[38] and since $Re_0 = e_0$, we have $R_{00} = 1$, so that $R \in \mathrm{SO}(3,1)^\uparrow$.

(2) **One property** We prove that $\Lambda^* Re_i \cdot e_0 - 0$ for $i = 2, 3$. For that:

$$\begin{aligned} \Lambda^* Re_i \cdot e_0 &= \Lambda^* \bar{b}_i \cdot e_0 && \text{(55.416a)} \\ &= \bar{b}_i \cdot \Lambda e_0 && \text{(55.416b)} \\ &= -\langle b_i, \texttt{proj}(\Lambda e_0)\rangle && \text{(55.416c)} \\ &= -\langle b_i, \|a\| b_1\rangle && \text{(55.416d)} \\ &= 0. && \text{(55.416e)} \end{aligned}$$

We used the relation (55.413) and the fact that $\{b_i\}$ is an orthonormal basis of $\mathbb{R}^3$.

36. Definition 55.85.
37. Definition 8.29.
38. Proposition 8.34.

(3) **A new basis** We define the following vectors:

$$f_0 = e_0 \tag{55.417a}$$
$$f_2 = \Lambda^* R e_2 \tag{55.417b}$$
$$f_3 = \Lambda^* R e_3. \tag{55.417c}$$

We check that these vectors are orthonormal. First:

$$\begin{aligned} f_0 \cdot f_2 &= e_0 \cdot f_2 && (55.418a)\\ &= e_0 \cdot \Lambda^* R e_2 && (55.418b)\\ &= R e_2 \cdot \Lambda e_0 && (55.418c)\\ &= \bar{b}_2 \cdot \Lambda e_0 && (55.418d)\\ &= -\langle b_2, \mathtt{proj}(\Lambda e_0)\rangle && (55.418e)\\ &= 0. && (55.418f) \end{aligned}$$

We get $f_0 \cdot f_3 = 0$ in the same way. Second:

$$f_2 \cdot f_3 = \Lambda^* R e_2 \cdot \Lambda^* R e_3 = e_2 \cdot e_3 = 0. \tag{55.419}$$

Now we fix f_1 in such a way that $\{f_0, f_1, f_2, f_3\}$ is a pseudo-orthonormal basis of $(\mathbb{R}^4, \cdot)$. Up to redefinition $f_1 \to -f_1$ we also suppose that $\{f_1, f_2, f_3\}$ is a basis of $\mathbb{R}^3$ with the same orientation as the canonical basis.

(4) **The second spatial rotation** We define the spatial rotation $S\colon \mathbb{R}^4 \to \mathbb{R}^4$ by

$$S f_i = e_i \tag{55.420}$$

for $i = 0, 1, 2, 3$.

Due to our choice of orientation for f_1 we have $S \in \mathrm{SO}(3,1)^\uparrow$.

(5) **Boost** We show that $S\Lambda^* R$ is a boost in the x direction. We have

$$S\Lambda^* R e_2 = S f_2 = e_2 \tag{55.421}$$

and the same for e_3: $S\Lambda^* R e_3 = e_3$.

We made some choices such that S and R belong to $\mathrm{SO}(3,1)^\uparrow$. Moreover by hypothesis $\Lambda \in \mathrm{SO}(3,1)^\uparrow$ and the lemma 55.84 implies that $\Lambda^* \in \mathrm{SO}(3,1)^\uparrow$. The whole shows that $S\Lambda^* R$ is a boost.

Finally, $\Lambda^* = S^{-1} L R^{-1}$ and taking onto account the fact that the adjoint is the inverse (lemma 55.80),

$$\Lambda = R L^{-1} S. \tag{55.422}$$

EQooSJNLooWZztHU

The operator L^{-1} is a boost because L is a boost.

The decomposition (55.422) is the requested one. □

55.16.3 Isomorphism $\mathrm{SO}(3,1)^\uparrow = \mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$

.

We are now going to prove the isomorphism

$$\mathrm{SO}(3,1)^\uparrow = \frac{\mathrm{SL}(2,\mathbb{C})}{\mathbb{Z}_2}. \tag{55.423}$$

We know from lemma 55.36 that the Pauli matrices form a basis of the hermitian 2×2 matrices. Here we introduce a new Pauli matrix $\sigma_0 = \mathrm{Id}$:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{55.424}$$

We denote by $\mathcal{H}$ the spanned vector space $\mathcal{H} = \mathrm{Span}\{\sigma_i\}_{i=0,1,2,3}$.

55.88.
We recall that $x \cdot x = x_0^2 - \sum_{i=1}^3 x_i^2$; we could write it as $\|x\|^2$ but we cannot figure out what $\|x\|$ should mean.

LEMooXKHYooFTzHhg

Lemma 55.89.
We define

$$\begin{aligned} \phi\colon \mathbb{R}^4 &\to \mathcal{H} \\ \sum_{i=0}^{3} x_i e_i &\mapsto \sum_i x_i \sigma_i. \end{aligned} \tag{55.425}$$

Then

(1) *The map ϕ is a vector space isomorphism*
ITEMooISFKooItVpre
(2) *We have $\det\big(\phi(x)\big) = x \cdot x$ for every $x \in \mathbb{R}^4$.*

55.90.
We will often write $\|x\|^2$ for $x \cdot x$, but we know that this is an abuse since we don't really want to define $\|x\|$ itself.

In an euclidian space like $\mathbb{R}$ endowed with its scalar product, we always have $x \cdot x > 0$ so that one can define $\|x\| = \sqrt{x \cdot x}$.

In the Minkowskian space, the situation is quite different because the number $x \cdot x$ can be positive as well as negative. Thus we will often use the notation $\|x\|^2$, and if you realty want to, $\|x\| = \sqrt{x \cdot x}$ with the convention that $\sqrt{-r} = i\sqrt{r}$ for $r \in \mathbb{R}^+$. So, if you want, $\|e_1\| = i$.

Proof. We have

$$\phi(x) = \begin{pmatrix} x_0 + x_3 & x - 1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix} \tag{55.426}$$

and then

$$\begin{aligned} \det\big(\phi(x)\big) &= (x_0 + x_3)(x_0 - x_3) - (x_1 + ix_2)(x_1 - ix_2) && \text{(55.427a)} \\ &= x_0^2 - x_3^2 - x_1^2 - x_2^2. && \text{(55.427b)} \end{aligned}$$

□
LEMooHPSJooEVIaoE

Lemma 55.91.
Let $S \in \mathrm{SL}(2, \mathbb{C})$. We define the map

$$\begin{aligned} g_S\colon \mathcal{H} &\to \mathbb{M}(2, \mathbb{C}) \\ X &\mapsto SXS^\dagger. \end{aligned} \tag{55.428}$$

(1) *The map g_S is linear,*
(2) *the target space is included in $\mathcal{H}$: $g_S(\mathcal{H}) \subset \mathcal{H}$.*

Proof. For the linearity,

$$g_S(X + Y) = S(X + Y)S^\dagger = SXS^\dagger + SYS^\dagger = g_S(X) + g_S(Y) \tag{55.429}$$

and

$$g_S(\lambda X) = S\lambda XS^\dagger = \lambda SXS^\dagger = \lambda g_S(X). \tag{55.430}$$

Since the map g_S is linear, it is sufficient to check that $g_S(\sigma_i) \in \mathcal{H}$ for each $i = 0, 1, 2, 3$. Each of these matrices is hermitian, we have

$$g_S(\sigma_i)^\dagger = (S\sigma_i S^\dagger)^\dagger = S\sigma_i^\dagger S^\dagger = S\sigma_i S^\dagger = g_S(\sigma_i). \tag{55.431}$$

Thus we have $g_S(\sigma_i) = g_S(\sigma_i)$ and we deduce that $g_S(\sigma_i)$ is hermitian.

When $i = 0$ we have $g_S(\sigma_0) = S\sigma_0 S^\dagger = SS^\dagger = \mathbb{1} = \sigma_0$.

When $i \neq 0$, using the cyclic invariance of the trace[39], we also have

$$\mathrm{Tr}\left(g_S(\sigma_i)\right) = \mathrm{Tr}(S\sigma_i S^\dagger) = \mathrm{Tr}(S^\dagger S\sigma_i) = \mathrm{Tr}(\sigma_i) = 0. \tag{55.432}$$

Thus $g_S(\sigma_i)$ is an hermitian matrix with null trace, so that $g_S(\sigma_i) \in \mathcal{H}$. □

Now we can consider the map $g_S : \mathcal{H} \to \mathcal{H}$.

LEMooEDIMooNIURdn

Lemma 55.92.
Let $S \in \mathrm{SL}(2,\mathbb{C})$. We consider the map of the lemma 55.91

$$\begin{aligned} g_S : \mathcal{H} &\to \mathcal{H} \\ X &\mapsto SXS^\dagger \end{aligned} \tag{55.433}$$

Then $\det\left(g_S(X)\right) = \det(X)$.

Proof. We use the proposition 4.103 :

$$\det(SXS^\dagger) = \det(S)\det(X)\det(S^\dagger) = \det(X) \tag{55.434}$$

because the determinant of $S \in \mathrm{SL}(2,\mathbb{C})$ is 1. □

LEMooXDPPooPImUQX

Lemma 55.93.
For $S \in \mathrm{SL}(2,\mathbb{C})$ we define

$$\begin{aligned} \Lambda(S) : \mathbb{R}^4 &\to \mathbb{R}^4 \\ x &\mapsto \phi^{-1}\Big(g_S\big(\phi(x)\big)\Big). \end{aligned} \tag{55.435}$$

Then $\Lambda(S) \in \mathrm{O}(3,1)$.

Proof. The lemma 55.89(2) shows that $\det\big(\phi(y)\big) = \|y\|^2$ for every $y \in \mathbb{R}^4$. We use that equation with $y = \Lambda(S)x$:

$$\begin{aligned} \|\Lambda(S)x\|^2 &= \|\phi g_S \phi^{-1} x\|^2 && \text{(55.436a)} \\ &= \det(\phi\phi^{-1} g_S \phi x) && \text{(55.436b)} \\ &= \det(g_S \phi x) && \text{(55.436c)} \\ &= \det\big(\phi(x)\big) && \text{(55.436d)} \\ &= \|x\|^2 && \text{(55.436e)} \end{aligned}$$

We used the fact that $\det\big(g_S(X)\big) = \det(X)$ by lemma 55.92. □

LEMooJBYDooSFEUFr

Lemma 55.94.
The map $\Lambda : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{O}(3,1)$ is a group homomorphism.

Proof. Let $S, T \in \mathrm{SL}(2,\mathbb{C})$ and $X \in \mathcal{H}$. We have

$$g_{ST}(X) = STX(ST)^\dagger = STXT^\dagger S^\dagger = Sg_T(X)S^\dagger = (g_S g_T)(X). \tag{55.437}$$

Now we can prove that Λ is an homomorphism:

$$\Lambda(ST) = \phi g_{ST}\phi^{-1} = \phi g_S g_T \phi^{-1} = \phi g_S \phi^{-1}\phi g_T \phi^{-1} = \Lambda(S)\Lambda(T). \tag{55.438}$$

□

LEMooTLQKooGntuRH

Lemma 55.95.
The map $\Lambda : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{O}(3,1)$ is continuous.

39. Lemma 4.65.

Proof. In several steps.

(1) **If $S_i \xrightarrow{\mathrm{SL}(2,\mathbb{C})} \mathbb{1}$ then $g_{S_i} \xrightarrow{\mathrm{End}(\mathcal{H})} \mathrm{Id}$** For each $X \in \mathcal{H}$ we have $g_{S_i}X = S_iXS_i^\dagger$. Using the fact that the matrix product is continus (proposition 9.34), we deduce

$$\lim_{i\to\infty} g_{S_i}(X) = X. \tag{55.439}$$

The proposition 12.118 concludes that $g_{S_i} \xrightarrow{\mathrm{End}(\mathcal{H})} \mathrm{Id}$.

(2) **If $S_i \xrightarrow{\mathrm{SL}(2,\mathbb{C})} \mathbb{1}$ then $\Lambda(S_i) \xrightarrow{\mathrm{O}(3,1)} \mathrm{Id}$** For each i we have

$$\Lambda(S_i) = \phi^{-1} \circ g_{S_i} \circ \phi. \tag{55.440}$$

Since ϕ and ϕ^{-1} are continuous, they commute with the limit and we have

$$\lim_{i\to\infty} \Lambda(S_i) = \phi^{-1} \circ \lim_{i\to\infty} g_{S_i} \circ \phi = \phi^{-1} \circ \mathrm{Id} \circ \phi = \mathrm{Id}\,. \tag{55.441}$$

(3) **If $S_i \xrightarrow{\mathrm{SL}(2,\mathbb{C})} g$ then $\Lambda(S_i) \xrightarrow{\mathrm{O}(3,1)} \Lambda(g)$** We successiveny have :

$$S_i \xrightarrow{\mathrm{SL}(2,\mathbb{C})} g \tag{55.442a}$$
$$g^{-1}S_i \xrightarrow{\mathrm{SL}(2,\mathbb{C})} \mathbb{1} \tag{55.442b}$$
$$\Lambda(g^{-1}S_i) \xrightarrow{\mathrm{O}(3,1)} \mathrm{Id} \tag{55.442c}$$
$$\Lambda(g^{-1})\Lambda(S_i) \xrightarrow{\mathrm{O}(3,1)} \mathrm{Id} \tag{55.442d}$$
$$\Lambda(S_i) \xrightarrow{\mathrm{O}(3,1)} \Lambda(g^{-1})^{-1} = \Lambda(g). \tag{55.442e}$$

Justifications:

— For (55.442b), the group law is continuous; thus we can multiply by g^{-1} to both sides

— For (55.442c), we apply Λ and we use the previous step.

— For (55.442d), we use the fact that Λ is an homomorphism (lemma 55.94)

— For (55.442e), we multiply both sides by $\Lambda(g^{-1})^{-1}$ and Λ and we use once again the fact that Λ is an homomorphism : $\Lambda(h)^{-1} = \Lambda(h^{-1})$.

(4) **Λ is continuous** Since $\mathrm{SL}(2,\mathbb{C})$ and $\mathrm{O}(3,1)$ are metric spaces, the proposition 7.255 makes sequential continuity (the one we) equivalent to the continuity. Thus the convergence (55.442e) says that Λ is continuous.

□

Lemma 55.96.
We have $\Lambda\big(\mathrm{SL}(2,\mathbb{C})\big) \subset \mathrm{SO}(3,1)^\uparrow$.

Proof. The group $\mathrm{SL}(2,\mathbb{C})$ is connected from proposition 13.18, and Λ is continuous by lemme 55.95. Thus the image of $\mathrm{SL}(2,\mathbb{C})$ by Λ is connected by lemma 7.218.

Thus $\Lambda\big(\mathrm{SL}(2,\mathbb{C})\big)$ is contained in the part of $\mathrm{SO}(3,1)$ which is connected to $\mathbb{1}$, namelly in $\mathrm{SO}(3,1)^\uparrow$. □

Before to prove that $\Lambda\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}(3,1)^\uparrow$ is surjective, we need some computations.

Lemma 55.97.
We have

$$e^{t\sigma_1} = \cosh(t)\mathbb{1} + \sinh(t)\sigma_1 \tag{55.443}$$

where σ_1 is the Pauli matrix (55.35).

Proof. We have $\sigma_1^2 = \mathbb{1}$ and, by definition,

$$e^{t\sigma_1} = \sum_{k\in\mathbb{N}} \frac{t^k\sigma_1^k}{k!}. \tag{55.444}$$

We use the proposition 10.120 to split the sum into even and odd terms:

$$\begin{aligned} e^{t\sigma_1} &= \sum_{k \text{ even}} \frac{t^k\sigma_1^k}{k!} + \sum_{k \text{ odd}} \frac{t^k\sigma_1^k}{k!} && (55.445a)\\ &= \sum_{k \text{ even}} \frac{t^k}{k!}\sigma_1 + \sum_{k \text{ odd}} \frac{t^k}{k!}\mathbb{1} && (55.445b)\\ &= \sinh(t)\sigma_1 + \cosh(t)\mathbb{1}. && (55.445c) \end{aligned}$$

□

LEMooPKYXooWGZkkG

Lemma 55.98.
Some results with the matrices $\{\sigma_i\}_{i=0,\dots,3}$.

(1) $e^{\gamma\sigma_1/2}\sigma_0 e^{\gamma\sigma_1/2} = e^{\gamma\sigma_1}$

(2) $e^{\gamma\sigma_1/2}\sigma_1 e^{\gamma\sigma_1/2} = \cosh(\gamma)\sigma_1 + \sinh(\gamma)\mathbb{1}$

(3) $e^{\gamma\sigma_1/2}\sigma_2 e^{\gamma\sigma_1/2} = \sigma_2$

(4) $e^{\gamma\sigma_1/2}\sigma_3 e^{\gamma\sigma_1/2} = \sigma_3$

Proof. These are computations using the formulas of lemma (55.38) as well as the hyperbolic trigonometry relations of proposition 15.125. As an example,

$$\begin{aligned} &\big(\cosh(\gamma/2)\mathbb{1} + \sinh(\gamma/2)\sigma_1\big)\sigma_2\big(\cosh(\gamma/2)\mathbb{1} + \sinh(\gamma/2)\sigma_1\big) && (55.446a)\\ &= \big(\cosh(\gamma/2)\sigma_2 + i\sinh(\gamma/2)\sigma_3\big)\big(\cosh(\gamma/2)\mathbb{1} + \sinh(\gamma/2)\sigma_1\big) && (55.446b)\\ &= \cosh(\gamma/2)\sigma_2 - i\cosh(\gamma/2)\sinh(\gamma/2)\sigma_3 + i\sinh(\gamma/2)\cosh(\gamma/2)\sigma_3 + i\sinh^2(\gamma/2)i\sigma_2 && (55.446c)\\ &= \big(\cosh^2(\gamma/2) - \sinh^2(\gamma/2)\big)\sigma_2 && (55.446d)\\ &= \sigma_2. && (55.446e) \end{aligned}$$

□

The following lemma provides a clearly non trivial element in $\Lambda\big(\mathrm{SL}(2,\mathbb{C})\big)$.

LEMooGURFooRTBBmi

Lemma 55.99 ([844])**.**
Let $\gamma \in \mathbb{R}$. We write $S(\gamma) = e^{\gamma\sigma_1/2}$. We have

$$\Lambda\big(S(\gamma)\big) = \begin{pmatrix} \cosh(\gamma) & \sinh(\gamma) & 0 & 0 \\ \sinh(\gamma) & \cosh(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{55.447}$$

Proof. Let $x \in \mathbb{R}^4$. We have

$$\begin{aligned} \Lambda\big(S(\gamma)\big)x &= (\phi^{-1}g_{S(\gamma)}\phi)(x) && (55.448a)\\ &= (\phi^{-1}\circ g_{S(\gamma)})\big(\sum_{k=0}^{3} x_k\sigma_k\big) && (55.448b)\\ &= \sum_k x_k\phi^{-1}\big(S(\gamma)\sigma_k S(\gamma)^\dagger\big) && (55.448c)\\ &= \sum_k x_k\phi^{-1}\big(S(\gamma)\sigma_k S(\gamma)\big). && (55.448d) \end{aligned}$$

At this point, we use the relations of lemma 55.98. We have:

$$\begin{aligned}
\Lambda\big(S(\gamma)\big)x &= x_0\phi^{-1}\big(e^{\gamma\sigma_1}\big) + x_1\phi^{-1}\big(\cosh(\gamma)\sigma_1 + \sinh(\gamma)\mathbb{1}\big) + x_2\phi^{-1}(\sigma_2) + x_3\phi^{-1}(\sigma_3) && (55.449a)\\
&= x_0\cosh(\gamma)e_0 + x_0\sinh(\gamma)e_1 + x_1\cosh(\gamma)e_1 + x_1\sinh(\gamma)e_0 + x_2e_2 + x_3e_3 && (55.449b)\\
&= \big(x_0\cosh(\gamma) + x_1\sinh(\gamma)\big)e_0 && (55.449c)\\
&\quad + \big(x_0\sinh(\gamma) + x_1\cosh(\gamma)\big)e_1 && (55.449d)\\
&\quad + x_2e_2 && (55.449e)\\
&\quad + x_3e_3. && (55.449f)
\end{aligned}$$

The matrix of $\Lambda\big(S(\gamma)\big)$ is now easy to write. □

LEMooTZPAooWGjMgU

Lemma 55.100.
We have $\mathrm{SO}(3) \subset \Lambda\big(\mathrm{SL}(2,\mathbb{C})\big)$ *where* $\mathrm{SO}(3)$ *denotes the block-diagonal matrices of the form*

$$\begin{pmatrix} 1 & 0 \\ 0 & \bar{A} \end{pmatrix} \tag{55.450}$$

in which $\bar{A}$ *is a* 3×3 *matrix of* $\mathrm{SO}(3)$.

Proof. A matrix of $\mathrm{SO}(3)$ in our context has the form

$$\begin{pmatrix} 1 & \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \bar{A} \end{pmatrix} \tag{55.451}$$

where $\bar{A}$ is a 3×3 matrix of the authentic $\mathrm{SO}(3)$. Now we use the proposition 55.46(3) addind bars everywhere to distinguish the objects about $\mathbb{R}^3$ from the objects about $\mathbb{R}^4$. There exists an element $U \in \mathrm{SU}(2)$ such that

$$\big(\bar{\phi}^{-1} \circ \rho(U) \circ \bar{\phi}\big)(\bar{y}) = \bar{A}\bar{y} \tag{55.452}$$

for every $\bar{y} \in \mathbb{R}^3$. In other words,

$$\bar{\phi}^{-1}\big(U\bar{\phi}(\bar{y})U^\dagger\big) = \bar{A}\bar{y}. \tag{55.453}$$

We know that $\mathrm{SU}(2) \subset \mathrm{SL}(2,\mathbb{C})$ and we will prove that $\Lambda(U) = A$, that is $\Lambda(U)y = Ay$ for every $y \in \mathbb{R}^4$. Since $\Lambda(U)$ and A are linear, we can prove the cases $y = e_k$ with $k = 0$ and $k \neq 0$ separately.

(1) $\underline{y = e_0}$ We have

$$\begin{aligned}
\Lambda(U)e_0 &= \phi^{-1}\big(U\phi(e_0)U^\dagger\big) && (55.454a)\\
&= \phi^{-1}\big(U\sigma_0U^\dagger\big) && (55.454b)\\
&= \phi^{-1}(\sigma_0) && (55.454c)\\
&= e_0. && (55.454d)
\end{aligned}$$

because $\sigma_0 = \mathrm{Id}$.

(2) $\underline{y = e_k \text{ with } k \neq 0}$ We have

$$f(U)e_k = \phi^{-1}\big(U\phi(e_k)U^\dagger\big) = \phi^{-1}(U\sigma_kU^\dagger). \tag{55.455}$$

Since $k \neq 0$, we know from proposition 55.45(1) that $U\sigma_kU^\dagger \in V$. Thus it makes sense to consider $\bar{\phi}^{-1}\big(U\sigma_kU^\dagger\big)$, and in fact we have

$$\phi^{-1}\big(U\sigma_kU^\dagger\big) = \begin{pmatrix} 0 \\ \bar{\phi}(U\sigma_kU^\dagger) \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{A}\bar{e}_k \end{pmatrix} = Ae_k. \tag{55.456}$$

It is now proven that $A = \Lambda(U)$, so that $\mathrm{SO}(3) \subset \Lambda\big(\mathrm{SL}(2,\mathbb{C})\big)$. □

Notice that we even have $\mathrm{SO}(3) \subset \Lambda\big(\mathrm{SU}(2)\big)$.

PROPooFPSLooSzvSYF

Proposition 55.101.
The map $\Lambda\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}(1,3)^\uparrow$ *is surjective.*

Proof. Let $A \in \mathrm{SO}(1,3)^\uparrow$. The standard decomposition 55.87 provides two rotations R and S and a boost L such that $A = RLS$. The lemmas 55.100 and 55.99 show that R, S and L are part of $\Lambda\big(\mathrm{SL}(2,\mathbb{C})\big)$. Since Λ is an homomorphism (lemma 55.94), we have the result. □

LEMooQKCQooTIaesa

Lemma 55.102.
We have $\ker(\Lambda) = \mathbb{Z}_2$.

Proof. We can use the same proof as the one given in 55.46, but we present here a simpler one. Let $S \in \mathrm{SL}(2,\mathbb{C})$ be an element of $\ker(\Lambda)$, that is $\Lambda(S) = \mathrm{Id}$, or

$$\Lambda(S)y = \phi^{-1}\big(S\phi(y)S^\dagger\big) = y. \tag{55.457}$$

We begin our investigations with $y = e_0$. Since $\phi(e_0) = \sigma_0 = \mathrm{Id}$, it provides $\phi^{-1}(SS^\dagger) = e_0 = \sigma_0$, so that

$$SS^\dagger = \mathrm{Id}\,. \tag{55.458}$$

We conclude that $S \in \mathrm{SU}(2)$. From here we can continue as in proposition 55.46, but we provide a simpler computation (just for fun).

For every $i = 1, 2, 3$ we have $\phi^{-1}\big(S\phi(e_i)S^\dagger\big) = e_i$. We apply ϕ on both sides:

EQooFMQQooDHtAGl

$$S\sigma_i S^\dagger = \phi(e_i) = \sigma_i. \tag{55.459}$$

Since $S \in \mathrm{SU}(2)$ we have $S^\dagger = S^{-1}$. We multiply both sides of (55.459) by S:

$$S\sigma_i = \sigma_i S. \tag{55.460}$$

Let $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a, b, c, d \in \mathbb{C}$. We have

$$S\sigma_1 = \begin{pmatrix} b & a \\ d & c \end{pmatrix} \tag{55.461a}$$

$$\sigma_1 S = \begin{pmatrix} c & d \\ a & b \end{pmatrix}, \tag{55.461b}$$

so that $b = c$ and $a = d$ and we are left with $S = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$. The same with σ_2:

$$S\sigma_2 = i\begin{pmatrix} b & -a \\ a & -b \end{pmatrix} \tag{55.462a}$$

$$\sigma_2 S = i\begin{pmatrix} -b & -a \\ a & b \end{pmatrix}, \tag{55.462b}$$

so that $b = 0$ and we are left with $S = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$. Since $S \in \mathrm{SU}(2)$ we have $1 = \det(S) = a^2$.

Thus a is an element of $\mathbb{C}$ such that $a^2 = 1$. Theorem 3.130 says that a polynomial of degree 2 has at most 2 roots. Thus $a^2 = 1$ implies $a = \pm 1$. No other possibilities. Thus $S = \pm\mathbb{1}$. It is easy to check that $\mathbb{1}$ and $-\mathbb{1}$ are part of $\ker(\Lambda)$.

We conclude that $\ker(\Lambda) = \pm\mathbb{1} = \mathbb{Z}_2$. □

Theorem 55.103.
We have the group isomorphism

$$\mathrm{SO}(1,3)^{\uparrow} = \frac{\mathrm{SL}(2,\mathbb{C})}{\mathbb{Z}_2}. \tag{55.463}$$

Proof. The proof follows the same lines as the proof of proposition 55.47. We use the first isomorphism theorem 2.6 with θ being the map $\Lambda\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}(1,3)$ defined by the lemma 55.93. It says that

$$\frac{\mathrm{SL}(2,\mathbb{C})}{\ker(\Lambda)} = \mathrm{Image}(\Lambda). \tag{55.464}$$

Lemma 55.102 says that $\ker(\Lambda) = \mathbb{Z}_2$ and proposition 55.101 says that $\mathrm{Image}(\Lambda) = \mathrm{SO}(1,3)^{\uparrow}$. □

55.16.4 Lie group structure

LEMooISDKooUnafZR

Lemma 55.104.
Some manifolds.

1. *The part* $\mathrm{O}(3,1)$ *is a* 6*-dimensional submanifold of* $\mathrm{GL}(4)$.
2. *The part* $\mathrm{SO}(3,1)$ *is a* 6*-dimensional submanifold of* $\mathrm{O}(3,1)$.
3. *The part* $\mathrm{SO}(3,1)^{\uparrow}$ *is a* 6*-dimensional submanifold of* $\mathrm{SO}(3,1)$.

Recall that $\mathrm{GL}(n\mathbb{R})$ is an analytic Lie group by proposition 55.16.

55.16.5 Lie algebra of $\mathrm{SO}(3,1)$

LEMooEEDYooKuRxbW

Lemma 55.105.
If $\omega \in \mathbb{M}(4,\mathbb{R})$ *satisfy* $\eta\omega = -\omega^t\eta$*, then for every* $k \in \mathbb{N}$ *we have*

$$\eta\omega^k = (-1)^k(\omega^t)^k\eta. \tag{55.465}$$

Proof. By induction. For $k = 0$ or $k = 1$ it's okay. For the induction,

$$\eta\omega^{k+1} = \eta\omega^k\omega = (-1)^k(\omega^t)^k\eta\omega = (-1)^k(\omega^t)^k(-)\omega^t\eta = (-1)^{k+1}(\omega^t)^{k+1}\eta. \tag{55.466}$$

□

PROPooGNHJooKILwuI

Proposition 55.106 ([1, 844])**.**
The Lie algebra $\mathfrak{so}(3,1)$ *of* $\mathrm{SO}(3,1)$ *is the set of the* 4×4 *matrices* ω *satisfying*

$$\omega^t\eta + \eta\omega = 0. \tag{55.467}$$

Proof. We are going to determine the set $\mathrm{SO}(3,1)'$ of the componentwise derivative of paths of matrices in $\mathrm{SO}(3,1)$. See 53.10 and the proposition 53.16.

Let

$$\mathcal{A} = \{\omega \in \mathbb{M}(4,\mathbb{R}) \text{ tel que } \omega^t\eta + \eta\omega = 0\}. \tag{55.468}$$

(1) $\underline{SO(3,1)' \subset \mathcal{A}}$
Let Λ be a path in $\mathrm{SO}(3,1)$; for each $t \in \mathbb{R}$ we have

$$\Lambda(t)^t\eta\Lambda(t) = \eta. \tag{55.469}$$

Taking the componentwise derivative at $t = 0$ we have

$$\Lambda'(0)^t\eta\Lambda(0) + \Lambda(0)^t\eta\Lambda'(0) = 0. \tag{55.470}$$

Notice that the derivative and the transposition commute. Since $\Lambda(0) = \mathbb{1}$, the element $\omega = \Lambda'(0)$ satisfy

$$\omega\eta + \omega^t\eta = 0. \tag{55.471}$$

This shows that the elements ω of $\mathfrak{so}(3,1)$ satisfy $\omega^t\eta + \eta\omega = 0$.

(2) $\mathcal{A} \subset \mathrm{SO}(3,1)$ Let $\omega \in \mathcal{A}$. We consider the path $\Lambda(t) = e^{\omega t}$. Here the exponential is a «normal» exponential of matrix [40]. The proposition 15.72 says that $\Lambda'(0) = \omega$. We have to prove that $\Lambda(t) \in \mathrm{SO}(3,1)$ for every t in a neighborhood of 0. For that, our first task is to prove that $\eta e^{t\omega} = e^{-t\omega^t}\eta$:

$$\eta e^{t\omega} = \eta \sum_{k=0}^{\infty} \frac{t^k \omega^k}{k!} \tag{55.472a}$$

$$= \sum_{k} \frac{t^k}{k!} \eta \omega^k \tag{55.472b}$$

SUBEQooBFPHooSVCAoQ

$$= \sum_{k} \frac{t^k}{k!} (-1)^k (\omega^t)^k \eta \tag{55.472c}$$

SUBEQooGDTPooZPNJOg

$$= e^{-t\omega^t}\eta \tag{55.472d}$$

SUBEQooMZOAooJFqwgT

Justifications:

— For (55.472b). Proposition 10.106.

— For (55.472c). Lemma 55.105.

— For (55.472d). Proposition 10.106 again.

We are now able to prove that $\Lambda(t) \in \mathrm{SO}(3,1)$. Using propositions 15.69(3) and 15.71, we see that

$$\Lambda(t)^t \eta \Lambda(t) = e^{t\omega^t} e^{-\omega^t} \eta = \eta. \tag{55.473}$$

This proves that $\Lambda(t) \in \mathrm{O}(3,1)$.

For the determinant, $\det\big(\Lambda(0)\big) = 1$ and $t \mapsto \det\big(\Lambda(t)\big)$ is continuous. Lemma 55.81 says that $\det\big(\Lambda(t)\big)$ can take only the values ± 1. By continuity, $\det\big(\Lambda(t)\big) = 1$ for every t.

□

55.107.

The link between the matrices ω satisfying $\omega^t\eta + \eta\omega = 0$ and the authentic Lie algebra of $\mathrm{SO}(3,1)$ is that the element $\omega \in T_e \mathrm{SO}(3,1)$ associated with the matrix ω is given by

$$\omega(f) = \frac{d}{dt}\Big[f(e^{\omega t})\Big]_{t=0} \tag{55.474}$$

when f is a function on $\mathrm{SO}(3,1)$.

LEMooVAYBooHcPKHU

Lemma 55.108 ([1, 844, 901])**.**

Let the matrices $\{M_{\lambda\mu}\}_{\lambda,\mu=0,1,2,3}$ given by the following matrix elements

$$(M_{\lambda\mu})_{\alpha\beta} = \delta_{\mu\alpha}\eta_{\lambda\beta} - \delta_{\lambda\alpha}\eta_{\mu\beta}. \tag{55.475}$$

(1) They belong to $\mathfrak{so}(3,1)$.

(2) The commutator is given by

$$[M_{\lambda\mu}, M_{\rho\sigma}] = \eta_{\mu\sigma}M_{\lambda\rho} - \eta_{\mu\rho}M_{\lambda\sigma} + \eta_{\lambda\rho}M_{\mu\sigma} - \eta_{\lambda\sigma}M_{\mu\rho}. \tag{55.476}$$

Proof. By proposition 55.106, we have to prove that $M_{\lambda\mu}\eta$ is antisymmetric. Here you are:

$$(M_{\lambda\mu}\eta)_{ij} = \sum_{k=0}^{3} (M_{\lambda\mu})_{ik}\eta_{kj} \tag{55.477a}$$

$$= \sum_{k} (\delta_{i\mu}\eta_{\lambda k} - \delta_{\lambda i}\eta_{\mu k})\eta_{kj} \tag{55.477b}$$

$$= \delta_{i\mu}(\eta\eta)_{\lambda j} - \delta_{\lambda i}(\eta\eta)_{\mu k} \tag{55.477c}$$

$$= \delta_{i\mu}\delta_{\lambda j} - \delta_{\lambda i}\delta_{\mu_j} \tag{55.477d}$$

SUBEQooWIBYooLnzvMw

40. Definition 15.60.

We used the fact that $\eta\eta = \delta$. The expression (55.477d) is antisymmetric with respect to i, j.

The commutation relation is a computations:

$$[M_{\lambda\mu}, M_{\rho\sigma}]_{ij} = \sum_k (M_{\lambda\mu})_{ik}(M_{\rho\sigma})_{kj} - \sum_k (M_{\rho\sigma})_{ik}(M_{\lambda\mu})_{kj} \tag{55.478a}$$

$$= \delta_{i\mu}\eta_{\lambda\sigma}\eta_{\rho j} - \delta_{i\mu}\eta_{\lambda\rho}\eta_{\sigma j} - \delta_{\lambda i}\eta_{\mu\sigma}\eta_{\rho j} + \delta_{\lambda i}\eta_{\mu\rho}\eta_{\sigma j} \tag{55.478b}$$

$$- \delta_{i\sigma}\eta_{\rho\mu}\eta_{\lambda i} + \delta_{i\sigma}\eta_{\rho\lambda}\eta_{\mu j} + \delta_{\rho i}\eta_{\sigma\mu}\eta_{\lambda j} - \delta_{\rho i}\eta_{\sigma\lambda}\eta_{\mu j} \tag{55.478c}$$

$$= \eta_{\lambda\sigma}(M_{\rho\mu})_{ij} + \eta_{\lambda\rho}(M_{\mu\sigma})_{ij} + \eta_{\mu\rho}(M_{\sigma\lambda})_{ij} + \eta_{\sigma\mu}(M_{\lambda\rho})_{ij}. \tag{55.478d}$$

Keep in mind that $M_{\lambda\mu} = -M_{\mu\lambda}$ and you get the result. □

Lemma 55.109.
We set

$$\begin{aligned} M_1 &= M_{32} & N_1 &= M_{01} \\ M_2 &= M_{13} & N_2 &= M_{02} \\ M_3 &= M_{21} & N_3 &= M_{03} \end{aligned} \tag{55.479}$$

We have the commutation relations

$$[M_i, M_j] = \sum_k \epsilon_{ijk} M_k \tag{55.480a}$$

$$[M_i, N_j] = \sum_k \epsilon_{ijk} N_k \tag{55.480b}$$

$$[N_i, N_j] = -\sum_k \epsilon_{ijk} M_k \tag{55.480c}$$

$$\tag{55.480d}$$

Proof. Direct computation using the commutators of lemma 55.108. An alternative is to ask Sage[902] to make the job.

First define the matrices:

```
"""Some definitions for the Lie algebra so(3,1)."""


from sage.all import matrix


def fun_delta(i, j):
    """The function under the matrix delta."""
    if i == j:
        return 1
    return 0


def matrix_delta():
    """Return the matrix delta."""
    return matrix(4, fun_delta)


def fun_eta(i, j):
    """The function under the matrix eta."""
    if i == j == 0:
        return 1
    if i == j:
```

```
        return -1
    return 0


def matrix_eta():
    """Return the matrix eta."""
    return matrix(4, fun_eta)


def matrix_M(lam, mu):
    """The matrix M_{lam, mu}"""

    def fun_M(alpha, beta):
        """The function for the matrix alpha, beta"""
        delta = matrix_delta()
        eta = matrix_eta()
        return delta[mu, alpha] * eta[lam, beta] \
            - delta[lam, alpha] * eta[mu, beta]

    return matrix(4, fun_M)


def epsilon(i, j, k):
    """Return epsilon_{ijk}."""
    if i == j:
        return 0
    if i == k:
        return 0
    if j == k:
        return 0
    if (i, j, k) == (1, 2, 3):
        return 1
    if (i, j, k) == (1, 3, 2):
        return -1
    if (i, j, k) == (2, 1, 3):
        return -1
    if (i, j, k) == (2, 3, 1):
        return 1
    if (i, j, k) == (3, 2, 1):
        return -1
    if (i, j, k) == (3, 1, 2):
        return 1


def epsilon_matrices(i, j, matrices):
    """return sum_k epsilon_{ijk}martices[k]."""
    return sum([epsilon(i, j, k) * matrices[k] for k in [1, 2, ↩
       3]])


def com(A, B):
    """Return the commutator of A and B"""
    return A * B - B * A
```

tex/sage/sageSnip020.py

The make the checks:

```
"""
Some check for the Lie algebra so(1,3).

This file has to be attached to sage's terminal.

sage: import os
sage: os.chdir('/path/to/tex/sage')
sage: attach('sageSnip021.sage')

This should only print a list of 'True'.
"""


from sageSnip020 import matrix_M
from sageSnip020 import epsilon_matrices
from sageSnip020 import com


def do_work():
    """Do the work."""
    M = {1: matrix_M(3, 2),
         2: matrix_M(1, 3),
         3: matrix_M(2, 1)}
    N = {1: matrix_M(0, 1),
         2: matrix_M(0, 2),
         3: matrix_M(0, 3)}

    # Check [M_i, M_j]=epsilon_{ijk}M_k
    for kk in range(1, 4):
        for ll in range(1, 4):
            ans = epsilon_matrices(kk, ll, M)
            print(com(M[kk], M[ll]) == ans)

    # Check [M_i, N_j]=epsilon_{ijk}N_k
    for kk in range(1, 4):
        for ll in range(1, 4):
            ans = epsilon_matrices(kk, ll, N)
            print(com(M[kk], N[ll]) == ans)

    # Check [N_i, N_j]=-epsilon_{ijk}M_k
    for kk in range(1, 4):
        for ll in range(1, 4):
            ans = - epsilon_matrices(kk, ll, M)
            print(com(N[kk], N[ll]) == ans)


do_work()
```

tex/sage/sageSnip021.sage

□

Lemma 55.110.
We set

$$L_i = \frac{1}{2}(M_i + iN_i) \tag{55.481a}$$
$$S_i = \frac{1}{2}(M_i - iN_i) \tag{55.481b}$$

and

$$\mathcal{B}_1 = \mathrm{Span}_{\mathbb{C}}\{L_i\} \tag{55.482a}$$
$$\mathcal{B}_2 = \mathrm{Span}_{\mathbb{C}}\{S_i\} \tag{55.482b}$$

(1) The sets $\mathcal{B}_1$ et $\mathcal{B}_2$ are Lie algebras.

(2) The commutation relations are

$$[L_i, L_j] = \sum_k \epsilon_{ijk} L_k \tag{55.483a}$$
$$[S_i, S_j] = \sum_k \epsilon_{ijk} S_k \tag{55.483b}$$
$$[L_i, S_j] = 0. \tag{55.483c}$$

Proof. It is sufficient to check the commutation relations. One again, these are only computations. You can do it with Sage:

```
"""
Some check for the Lie algebra so(1,3).

This file has to be attached to sage's terminal.

sage: import os
sage: os.chdir('/path/to/tex/sage')
sage: attach('sageSnip022.sage')

This should only print a list of 'True'.
"""

from sage.all import I
from sageSnip020 import matrix_M
from sageSnip020 import epsilon_matrices
from sageSnip020 import com


def do_work():
    """Do the verifications."""
    M = {1: matrix_M(3, 2),
         2: matrix_M(1, 3),
         3: matrix_M(2, 1)}
    N = {1: matrix_M(0, 1),
         2: matrix_M(0, 2),
         3: matrix_M(0, 3)}
    L = {k: (M[k] + I * N[k])/2 for k in [1, 2, 3]}
    S = {k: (M[k] - I * N[k])/2 for k in [1, 2, 3]}
```

```
29
30      # Check [L_i, L_j] = epsilon_{ijk}L_k
31      for kk in range(1, 4):
32          for ll in range(1, 4):
33              ans = epsilon_matrices(kk, ll, L)
34              print(com(L[kk], L[ll]) == ans)
35
36      # Check [L_i, S_j] = 0
37      for kk in range(1, 4):
38          for ll in range(1, 4):
39              print(com(L[kk], S[ll]) == 0)
40
41      # Check [S_i, S_j] = epsilon_{ijk}S_k
42      for kk in range(1, 4):
43          for ll in range(1, 4):
44              ans = epsilon_matrices(kk, ll, S)
45              print(com(S[kk], S[ll]) == ans)
46
47  do_work()
```

tex/sage/sageSnip022.sage

□

55.17 Representations of $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{SU}(2)$

The representation π_m of $\mathrm{SL}(2,\mathbb{C})$ restricts to $\mathrm{SL}(2,\mathbb{R})$.

Lemma 55.111.
The representation π_m of $\mathrm{SL}(2,\mathbb{R})$ is irreducible.

Proof. If W is an invariant space under $\pi_m\big(\mathrm{SL}(2,\mathbb{R})\big)$, then is is invariant under the derived representation $\rho_m\big(\mathfrak{sl}(e,\mathbb{R})\big)$. The proof of proposition 53.134 still holds here, so that $W = \mathcal{P}_m$. □

Theorem 55.112.
Let π be an irreducible representation of $G = \mathrm{SL}(2,\mathbb{R})$ or $\mathrm{SU}(2)$ in a complex finite dimensional vector space V. Then π is equivalent to one of the π_m.

Proof. Let $\mathfrak{g}$ be the Lie algebra of G. One important property shared by $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{SU}(2)$ is that $G = \exp(\mathfrak{g})$. It is clear that the representation $d\pi$ on $\mathfrak{sl}(2,\mathbb{R})$ extends $\mathbb{C}$-linearly to a representation ρ of $\mathfrak{sl}(2,\mathbb{C})$. Looking on the basis (51.54), one sees that in fact the same is true for $\mathfrak{su}(2)$ which $\mathbb{C}$-linearly extends to $\mathfrak{sl}(2,\mathbb{C})$.

Let us prove that ρ is irreducible. Let $W \neq \{0\}$ be a subspace of V invariant under $\rho(\mathfrak{g})$. Then W is invariant under $e^{\rho(X)} = \pi(e^X)$ for every $X \in \mathfrak{g}$. Since $\exp(\mathfrak{g}) = G$, the space W is in fact invariant under $\pi(G)$, and is therefore equal to V.

Since ρ is irreducible, we have $\rho = \rho_m$ for a certain m. Thus there exists an intertwining operator $A\colon V \to \mathcal{P}_m$ such that

$$A\rho(X) = \rho_m(X)A$$

for every $X \in \mathfrak{g}$. By linearity, for every $N \in \mathbb{N}$, we have $A\rho\big(\sum_{k=1}^n X^k/k!\big) = \rho_m\big(\sum_{k=1}^n X^k/k!\big)A$, and at the limit, we have

$$Ae^{\rho(X)} = e^{\rho_m(X)A}. \tag{55.484}$$

From that we deduce that $A\pi(e^X) = \pi_m(e^X)A$ which means that

$$A\pi(g) = \pi_m(g)A.$$

That shows that A intertwines π and π_m, so that π is equivalent to π_m. □

55.18 Representations of $\mathfrak{so}(2, d-1)$

Here we deal with the representations of $\mathfrak{so}(2, d-1)$. For singleton theory as field theory, see the section 70.5 (you are welcome if you can fill that section which is for the moment almost empty).

55.18.1 Verma module

One can find text about these representations in [903, 896, 904], while we will mainly follow the developments of [905, 906, 907]. We are going to study representations of the algebra $\mathfrak{so}(2, d-1)$ which fulfills the commutation relations described in lemma 51.61:

EqCommsodeuxdmoinsun

$$[M_{ab}, M_{cd}] = -i\eta_{ac}M_{bd} + i\eta_{ad}M_{bc} + i\eta_{bc}M_{ad} - i\eta_{bd}M_{ac}. \tag{55.485}$$

Notice that $M_{ab} = -M_{ba}$. As convention, the indices a, b,... run over $\{0, 0', 1, 2, \dots d-1\}$ while r, s, ... run over $\{1, 2, \dots, d-1\}$. As on page 3941, we choose the convention

$$\eta = \begin{pmatrix} \mathbb{1}_{2\times 2} & \\ & -\mathbb{1}_{(d-1)\times(d-1)} \end{pmatrix}. \tag{55.486}$$

Notice that this convention numerically holds for the matrix η_{st} as well as for its inverse η^{st}.

The algebra separates into two parts: the compact and the non compact part. The maximal compact subalgebra is $\mathfrak{so}(2) \oplus \mathfrak{so}(d-1)$ which is generated by $E = M_{00'}$ and $J_{rs} = M_{rs}$. The non compact generators are $M_{0'r}$ and M_{0r} that we rearrange into ladder operators

$$L_r^{\pm} = M_{0r} \mp iM_{0'r}. \tag{55.487}$$

Using commutation relations (55.485), one computes the commutators in the new basis. For example

$$[E, L_r^{\pm}] = [M_{0'0}, M_{0r}] \mp i[E_{0'0}, M_{0'r}] = \pm M_{0r} - iM_{0'r} = \pm L_r^{\pm}.$$

The table of $\mathfrak{so}(2, d-1)$ in this basis is

SubEqsCommssodeuxd

$$[E, L_r^{\pm}] = \pm L_r^{\pm} \tag{55.488a}$$
$$[J_{rs}, L_t^{\pm}] = -i(\delta_{rt}L_s^{\pm} - \delta_{st}L_r^{\pm}) \tag{55.488b}$$
$$[L_r^-, L_s^+] = 2(iJ_{rs} + \delta_{rs}E) \tag{55.488c}$$
$$[J_{rs}, J_{tu}] = -i\delta_{ac}M_{bd} + i\delta_{ad}M_{bc} + i\delta_{bc}M_{ad} - i\delta_{bd}M_{ac}. \tag{55.488d}$$

The unitary properties are $(M_{rs})^\dagger = M_{rs}$, $E^\dagger = E$ and $(L_r^{\pm})^\dagger = L_r^{\mp}$. From these commutators, we deduce the following rules that will be always used

$$L_r^- L_s^+ = L_s^+ L_r^- + 2(iJ_{rs} + \delta_{rs}E) \tag{55.489a}$$
$$J_{rs}L_t^+ = L_t^+ J_{rs} - i(\delta_{rt}L_s^+ - \delta_{st}L_r^+) \tag{55.489b}$$
$$EL_r^+ = L_r^+ E + L_r^+. \tag{55.489c}$$

The Cartan algebra of $\mathfrak{so}(d)$ is given by the elements $A_p = M_{2p-1,2p}$ with $p = 1, \dots, r$ for $\mathfrak{so}(2r)$ and $\mathfrak{so}(2r+1)$.

The unitary irreducible representations of $\mathfrak{so}(2, n)$ have the form $\mathcal{D}(e_0, \bar{\jmath})$. It is given by a basis vector $|e_0, \bar{\jmath}\rangle$ on which E and J_{rs} act by their respective representations (of $\mathfrak{so}(n)$ and $\mathfrak{so}(2)$). The **energy** of the vector $|e, \overline{m}\rangle$ is its eigenvalue for the operator E, namely e:

$$E|e, \overline{m}\rangle = e|e, \overline{m}\rangle. \tag{55.490}$$

Using the commutators (55.488), we find $L_r^{\pm} = (E \pm 1)L_r^{\pm}$, so that

$$EL_r^{\pm}|e, \overline{m}\rangle = (e \pm 1)|e, \overline{m}\rangle. \tag{55.491}$$

We see that the ladder operator L_r^+ raises the value of the energy of one unit, while the operator L_r^- lower the energy of one unit. The vector $|e_0, \bar{\jmath}\rangle$ is the **vacuum vector**, it has the lowest energy in the sense that $L_r^-|e_0, \jmath\rangle = 0$. A **scalar representation** is a representation with $\bar{\jmath} = 0$. They are, logically, denoted by $\mathcal{D}(e_0)$ and its vacuum is $|e_0\rangle$ which satisfies

$$J_{rs}|e_0\rangle = 0 \qquad (E - e_0)|e_0\rangle = 0 \qquad L_r^-|e_0\rangle = 0. \tag{55.492}$$

Eqaldefketezerovac

Then one build the generalised Verma module

$$\mathcal{V}(e_0, 0) \equiv \{L_{r_1}^+ \dots L_{r_n}^+|e_0\rangle\}_{n=0}^{\infty}. \tag{55.493}$$

EqmVVermaldots

Notice that the Verma module is not automatically irreducible. We will soon build irreducible representations by taking quotient of the Verma module by its singular module.

In order to compute the norm of $L_s^+ L_r^+|e_0\rangle$, we compute $L_r^- L_s^- L_s^+ L_r^+|e_0\rangle = 4\big(E + E^2 + \delta_{rs}E^2 + (J_{rs})^2\big)|e_0\rangle$. In order to get that result, we moved all the L^- on the right using the commutation relation, and we taken into account the simplifications induced by the definition relations (55.492). Now, using the relation $J_{rs}|e_0\rangle = 0$, we have

$$4\big(E + E^2 + \delta_{rs}E^2\big)|e_0\rangle. \tag{55.494}$$

We also have

$$(E - e_0)L_s^+ L_s^+|e_0\rangle = 0. \tag{55.495}$$

Proposition 55.113.
The vectors $L_{r_1}^+ \dots L_{r_k}^+|e_0\rangle$ and $L_{t_1}^+ \dots L_{t_l}^+|e_0\rangle$ are orthogonal if $k \neq l$.

That proposition says that different layers are orthogonal [41]

Proof. We proceed by induction. We suppose that the result is proved for $k, l \geqslant n$, and we prove that

$$L_{t_1}^- \dots L_{t_{n+1}}^- L_{r_1}^+ \dots L_{r_n}^+|e_0\rangle = 0. \tag{55.496}$$

First, remark that, using the commutation relations and the fact that $J_{rs}|e_0\rangle = 0$ and $E|e_0\rangle = e_0|e_0\rangle$, the vectors

SubEqsJELLket

$$J_{st}L_{r_1}^+ \dots L_{r_k}^+|e_0\rangle \tag{55.497a}$$
$$EL_{r_1}^+ \dots L_{r_k}^+|e_0\rangle \tag{55.497b}$$

are combinations of vectors of the form $L_{a_1}^+ \dots L_{a_k}^+|e_0\rangle$. Now, we have

$$L_{t_1}^- \dots L_{t_{n+1}}^- L_{r_1}^+ \dots L_{r_n}^+|e_0\rangle = L_{t_1}^- \dots L_{t_n}^-\big(L_{r_1}^+ L_{t_{n+1}}^- + 2i(J_{t_{n+1},r_1} + \delta_{t_{n+1},r_1}E)\big)L_{r_2}^+ \dots L_{r_n}^+|e_0\rangle \tag{55.498}$$

which decomposes in three terms. The first one is

$$L_{t_1}^- \dots L_{t_n}^- L_{r_1}^+ L_{t_{n+1}}^- L_{r_2}^+ \dots L_{r_n}^+|e_0\rangle, \tag{55.499}$$

and according to equations (55.497), the two other terms reduce to zero. Continuing that way, the operator $L_{t_{n+1}}^-$ advance of one position at each step and finishes to kill himself on $|e_0\rangle$. □

55.18.2 Singular module

Let us compute the norm of the general vector $L_{r_1}^+ \dots L_{r_k}^+|s_0, s\rangle$. We have

$$\begin{aligned} L_{r_k}^- \dots L_{r_1}^- L_{r_1}^+ \dots L_{r_k}^+|e_0, s\rangle &= L_{r_k}^- \dots L_{r_2}^-(L_{r_1}^+ L_{r_2}^- + 2E)L_{r_2}^+ \dots L_{r_k}^+|e_0, s\rangle \\ &= L_{r_k}^- \dots L_{r_2}^- L_{r_1}^+ L_{r_2}^- L_{r_2}^+ \dots L_{r_k}^+|e_0, s\rangle \\ &\quad + 2(e_0 + k - 1)L_{r_k}^- \dots L_{r_2}^- L_{r_2}^+ \dots L_{r_k}^+|e_0, s\rangle \end{aligned} \tag{55.500}$$

41. À justifier en analysant qui est exactement ɧ et les racines simples, mais ça me semble ok.

Using again and again the commutation relations, we eliminate all the operators L_r^+ and we obtain a sum of terms of the form $(e_0 + k - l)$. It will, obviously be positive for large enough e_0. Thus, unitarity of the representation is enforced for large values of e_0, and there exists a lower bound $E_0(s)$ such that negative norm states appears when $e_0 < E_0(s)$. If $e_0 = E_0(s)$, then these vectors have a vanishing norm.

Let us consider a limit representation: $e_0 = E_0$; there are vectors of vanishing norm, but no vectors with negative norm. In that case, if $v \cdot v = 0$, then $v \cdot w = 0$ for every other vector w. Indeed, if $v \cdot w \neq 0$, we have

$$(v - w) \cdot (v - w) = w \cdot w - v \cdot w - w \cdot v, \tag{55.501}$$

which holds for every positive multiple λv and μw. Choosing a big λ and a small μ, the norm of $\lambda v - \mu w$ becomes negative. What we proved is

Lemma 55.114.
If the energy e_0 of a representation saturates the unitary condition, then a vector with vanishing norm is orthogonal to every other vectors. Moreover, the vectors with vanishing norm form an invariant subspace.

The second part is the fact that, if $A \in \mathfrak{g}$, and $\||\psi\rangle\| = 0$ then $\|A|\psi\rangle\| = 0$, because it is the scalar product of $|\psi\rangle$ with the vector $A^\dagger A|\psi\rangle$. The submodule made of vectors of zero norm is the **singular submodule**, and is denoted by $\mathcal{S}(e_0, s)$.

PropSinModRedSSIADesNuls

Proposition 55.115.
*A module is reducible if and only if it possesses a vector $|v\rangle$ (different from $|e_0, s\rangle$) such that $L_r^-|v\rangle = 0$ for every r. Such a vector is said to be **null**.*

Proof. Since the energy is bounded from bellow, applying several times the lowering operators L_r^- on any vector ends up on zero. Thus, any submodule contains a vector $|v\rangle$ such that $L^-|v\rangle = 0$ for every r. If that vector is not $|e_0, s\rangle$, then the submodule is a proper submodule.

If $|v\rangle \neq |e_0, s\rangle$, then it is of the form $L_{\bar{r}}^+|e_0, s\rangle$ and its norm is given by

$$\|L_{\bar{r}}^+|e_0, s\rangle\| = \langle e_0, s|L_{\bar{r}}^- L_{\bar{r}}^+|e_0, s\rangle = 0 \tag{55.502}$$

because $L_{\bar{r}}^- L_{\bar{r}}^+|e_0, s\rangle = 0$ by assumption. □

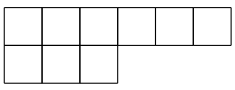
From the Verma module (55.493), we thus extract the irreducible representation taking the quotient by the singular module:

$$\mathcal{H}(e_0) = \mathcal{V}(e_0)/\mathcal{S}(e_0). \tag{55.503}$$

Most of time, we have only one extra vacuum, let $|e_0', s'\rangle$, and in this case, the whole singular module is generated by vectors of the form

$$L_{r_1}^+ \dots L_{r_k}^+|e_0', s'\rangle. \tag{55.504}$$

Let $|e_0, s\rangle$ be the vacuum with $s = (s_1, s_2, 0, \dots, 0)$, corresponding to the Young diagram



(55.505)

where the first line has s_1 boxes and the second one has s_2 boxes. Thanks to theorem 55.5, it can be realized with the tensor

$$v_{a_1, \dots a_{s_2} b_1 \dots s_1}(e_0) \tag{55.506}$$

which is separately symmetric in the indices a and b, in the same time as being antisymmetric in the couples a_i, b_i when $i \leqslant s_2$, for example,

$$v_{a_1 \dots a_{s_2}, b_1 \dots b_{s_1}} = -v_{b_1 a_1 \dots a_{s_2}, b_2 \dots b_{s_1}}. \tag{55.507}$$

In particular, if we symmetrise v on s_1+1 indices, we always found zero. Moreover, all the traces vanishes. If η is the metric of $O(D-1)$, we have for example

$$\eta^{b_1 b_2} v_{a_1 \dots a_{s_2}, b_1, \dots b_{s_2}} = 0. \tag{55.508}$$

The vectors of the first level are the ones of the form $L_r^+|e_0, s\rangle$. As far as notations are concerned, we have

$$L_r^+ v_{a_1 \dots a_{s_2}, b_1 \dots b_{s_2}} = (L_r^+ v)_{a_1 \dots a_{s_2}, b_1 \dots b_{s_2}}. \tag{55.509}$$

Remark that the operators $\{L_r^+\}_{r=1,\dots,D-1}$ carry a representation of $o(D-1)$, namely the vector representation. Thus, the states of the first level form the representation given by the tensor product of $(s_1, s_2, 0, \dots)$ and the vector representation. In order to see the irreducible components of that representation, we have to know what are the symmetry properties that we can give to the indices

$$r, a_1, \dots, a_{s_2}, b_1, \dots, b_{s_1}. \tag{55.510}$$

There are three possibilities: we can contract the r with one of the a_i (by symmetry, all of these contractions are equivalent), or with one of the b_i, or add one box in the Young diagram. The latter possibility splits into three cases: the diagram $(s_1, s_2, 0, \dots)$ can be transformed in $(s_1+1, s_2, 0, \dots)$, $(s_1, s_2+1, 0, \dots)$ or $(s_1, s_2, 1, 0, \dots)$. So we have 5 irreducible component in the $o(D-1)$ representation carried by the level one.

The question that naturally arises is to know if one of these have a singular vacuum. In other words, if Π_β are the projections to the irreducible components, do we have

$$L_t^- \Pi_\alpha \big(L_r^+ v_{\bar{a}, \bar{b}}(e_0) \big) = 0 \tag{55.511}$$

for a certain e_0?

Notice that the contraction with the last b_i's is not the same as the one with the firsts ones because of the symmetry properties with respect to the a_i's. The first representation with cell cut is given by

$$v^1_{\bar{a}, b_1 \dots b_{s_1-1}} = \eta^{rt} L_r^+ \big\{ v_{\bar{a}, b_1 \dots b_{s_1-1} t}(e_0) + \frac{s_2}{s_1 - s_2 + 1} v_{c a_1 \dots a_{s_2-1}, b_1 \dots b_{s_1-1} a_{s_2}(e_0)} \big\}, \tag{55.512}$$

while the second representation with cell cut is easier:

$$v^2_{a_1 \dots a_{s_2-1}, \bar{b}} = \eta^{rt} L_r^+ v_{t a_1 \dots a_{s_2-1}, \bar{b}}(e_0). \tag{55.513}$$

Now, the sport is to compute $L_q^- v^1_{\bar{a}, b_1 \dots b_{s_1-1}}$ and $L_q^- v^2_{a_1 \dots a_{s_2-1}, \bar{b}}$.

¡¡ Avertissement/question à la lectrice !! 55.116
Il y a du calcul non terminé, ici.

55.18.3 The quotient for the scalar singleton

The value of the energy which saturates the unitary condition is 1 when $s = \frac{1}{2}$ and $\frac{1}{2}$ when $s = 0$. That is the reason why we consider the two special representations

$$\text{Di} = \mathcal{D}(1, \frac{1}{2}) \quad \text{Rac} \quad = \mathcal{D}(\frac{1}{2}, 0). \tag{55.514}$$

We are now interested in the scalar case, the *Rac*.

We know that, when ϵ_0 is the value of the energy which saturates the unitary condition $e_0 \geqslant \frac{d-3}{2}$ (in the scalar case, then

(1) the vectors $L_s^+ L_s^+ |\epsilon_0\rangle$ are singular vectors,

(2) the vectors $L_{r_1}^+ \cdots L_{r_n}^+ L_s^+ L_s^+ |\epsilon_0\rangle$ is orthogonal to all other states, it is a null vector.

On the other hand, we know from proposition 55.115 that a module is reducible if and only if it has a vector $|v\rangle \neq |e_0, s\rangle$ such that $L_r^-|v\rangle = 0$ for every r. Thus one constructs irreducible representations by taking the quotient of the Verma module by the singular module.

What is the dimension of the scalar singleton? We have to count how many different vectors we have in the Verma module $\mathcal{V}(e_0, 0) \equiv \left\{L_{r_1}^+ \dots L_{r_n}^+ |e_0\rangle\right\}_{n=0}^{\infty}$, and which *are not* build over $L_s^+ L_s^+ |e_0\rangle$. In the case of SO(2, 3), we have the generators L_1^+, L_2^+ and L_3^+ (which are commuting), so the only vectors that are left after removing the singular modules are the seven following ones: $L_1^+|e_0\rangle$, $L_2^+|e_0\rangle$,$L_3^+|e_0\rangle$, $L_1^+ L_2^+|e_0\rangle$, $L_1^+ L_3^+|e_0\rangle$,$L_2^+ L_3^+|e_0\rangle$, and $L_1^+ L_2^+ L_3^+|e_0\rangle$. The scalar singleton representation is thus 7 dimensional.

Chapter 56

Fiber bundle

56.1 Vector bundle

DEFooFRPLooTxMUzg

Definition 56.1 ([1, 908]).
*Let M be a smooth manifold with charts $\{\varphi_\alpha : I_\alpha \to M\}_{\alpha\in\tau}$. A k-dimensional **vector bundle** on M is the given of*

ITEMooIVKKooNEfxVi

(1) A vector space V_x for each $x \in M$.
Formally, we suppose that the vector space V_x has the form $V_x = \{(x, v) \text{ tel que } v \in W_x\}$ for some vector space W_x. The reason is that we want the union $E = \bigcup_{x\in M} V_x$ to be disjoint.

(2) For each chart $\varphi_\alpha : U_\alpha \to M$, a map

$$\sigma_\alpha : U_\alpha \times \mathbb{R}^k \to \bigcup_{x\in\varphi_\alpha(U_\alpha)} V_x. \tag{56.1}$$

With the following conditions.

ITEMooVIYBooMXFaKb

(a) The maps σ_α give a manifold structure on E.

ITEMooXHOBooERubro

(b) The projection

$$\begin{aligned} \pi : E &\to M \\ (x, v) &\mapsto x \end{aligned} \tag{56.2}$$

is smooth for the differential structure of E given by the atlas $\{\sigma_\alpha\}_{\alpha\in\tau}$.

ITEMooXJIVooVJTorK

(c) For every $s \in U_\alpha$ and $v \in \mathbb{R}^k$ we have $\sigma_\alpha(s, v) \in V_{\varphi_\alpha(s)}$.

ITEMooOBCMooOCqoQo

(d) For each $s \in U_\alpha$, the map

$$\begin{aligned} \tau_{\alpha,s} : \mathbb{R}^k &\to V_{\varphi_\alpha(s)} \\ v &\mapsto \sigma_\alpha(s, v) \end{aligned} \tag{56.3}$$

is a vector space isomorphism (linear bijection).

ITEMooPNLXooLQWGqc

(e) The map $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Fun}(\mathbb{R}^k, \mathbb{R}^k)$ defined by

$$(\sigma_\alpha^{-1} \circ \sigma_\beta)(s, v) = \big(s, g_{\alpha\beta}(s)v\big) \tag{56.4}$$

belongs to $C^\infty\big(U_\alpha \cap U_\beta, \mathrm{GL}(\mathbb{R}, k)\big)$.

Most of the time, we will denote the whole vector bundle only by the couple (E, π).
*The maps $g_{\alpha\beta}$ are the **transition functions**.*

DEFooDSKKooEFgmPU

Definition 56.2.
*Let E_1, E_2 be two vector bundles over the same manifold M. A map $f : E_1 \to E_2$ is a **vector bundle isomorphism** if*

(1) f is a manifold smooth diffeomorphism.

(2) $\pi_1 \circ f = \pi_2$

(3) $f\big(\pi_1^{-1}(x)\big) = \pi_2^{-1}(x)$ *for every* $x \in M$.

(4) *The restriction* $f\colon \pi_1^{-1}(x) \to \pi_2^{-1}(x)$ *is a vector space isomorphism.*

PROPooJZLEooYFgjnY

Proposition 56.3.
Let V be a k-dimensional vector space. Let $f\colon \mathbb{R}^k \to V$ be an isomorphism and M, a manifold. We define

$$V_x = \{(x, v) \text{ tel que } v \in V\} \tag{56.5}$$

and

$$\begin{aligned} \sigma_\alpha\colon U_\alpha \times \mathbb{R}^k &\to \bigcup_{x \in \varphi_\alpha(U_\alpha)} \\ (s, v) &\mapsto \big(\varphi_\alpha(s), f(v)\big). \end{aligned} \tag{56.6}$$

We have:

ITEMooYHUVooXUSryK

(1) *This is a vector bundle structure.*

ITEMooCZLRooIJjQUn

(2) *If f and g are two isomorphisms between $\mathbb{R}^k$ and V, the corresponding vector bundle are isomorphic*[1].

Proof. For (1), this is a long verification[2].

We prove (2). We name E_1 the vector bundle created by f and E_2 the one created by g. We define

$$\begin{aligned} h\colon E_1 &\to E_2 \\ (x, v) &\mapsto \big(x, (f \circ g^{-1})(v)\big). \end{aligned} \tag{56.7}$$

It satisfies $\pi_1 \circ h = \pi_2$ because $\pi_i(x, v) = x$ for every v. The map h is a bijection because f and g are bijections. Since f and g are linear, the restriction map $h\colon V_x \to V_x$ is a vector space isomorphism.

We prove that h is a manifold isomorphism: h and h^{-1} have to be smooth.

(1) **h is smooth** In order to say that $h\colon E_1 \to E_2$ is smooth, we have to check that

$$\sigma_{f,\alpha}^{-1} \circ h \circ \sigma_{g,\beta}\colon U_\alpha \cap U_\beta \times \mathbb{R}^k \to U_\alpha \cap U_\beta \times \mathbb{R}^k \tag{56.8}$$

is smooth. We have

$$\begin{aligned} \sigma_{f,\alpha}^{-1} \circ h \circ \sigma_{g,\beta}(s, v) &= (\sigma_{f,\alpha} \circ h)\big(\varphi_\beta(s), g(v)\big) && \text{(56.9a)} \\ &= \sigma_{f,\alpha}^{-1}\Big(\varphi_\beta(s), (f \circ g^{-1})\big(g(v)\big)\Big) && \text{(56.9b)} \\ &= \sigma_{f,\alpha}^{-1}\big(\varphi_\beta(s), f(v)\big) && \text{(56.9c)} \\ &= \big((\varphi_\alpha^{-1} \circ \varphi_\beta)(s), v\big). && \text{(56.9d)} \end{aligned}$$

The last expression is smooth with respect to (s, v).

(2) **h^{-1} is smooth** The verification is the same for $\sigma_{g,\alpha}^{-1} \circ h^{-1} \circ \sigma_{g,\beta}$.

□

56.1.1 Transition functions

Proposition 56.4.
Let (E, π) be a vector bundle. The transition functions satisfy

$$g_{\alpha\beta} \circ g_{\alpha\gamma} \circ g_{\gamma\alpha} = \mathbb{1}. \tag{56.10}$$

eq:g_compat

1. Definition 56.2
2. Make it, write it and send me your proof, so that I can add it here.

56.1.2 Inverse construction

subsec:inv_g

Let us consider a manifold M, an open covering $\{\mathcal{U}_\alpha : \alpha \in I\}$ and some functions $g_{\alpha\beta}\colon \mathcal{U}_{\alpha\beta} \to GL(r,\mathbb{K})$ which fulfill relations (56.10). We will build a vector bundle $E \xrightarrow{p} M$ whose transition functions are the $g_{\alpha\beta}$'s. Let $\tilde{E}$ be the disjoint union

$$\tilde{E} = \bigsqcup_{\alpha\in I} \mathcal{U}_\alpha \times \mathbb{K}^r,$$

i.e. triples of the form $(x,v,\alpha) \in M \times \mathbb{K}^r \times I$ with the condition that $x \in \mathcal{U}_\alpha$. We define an equivalence relation on $\tilde{E}$ by $(x,v,\alpha) \sim (y,w,\beta)$ if and only if $x = y$ and $w = g_{\alpha\beta}(x)v$. Next, we define $E = \tilde{E}/\sim$ and $\omega\colon \tilde{E} \to E$, the canonical projection. The projection $p\colon E \to M$ is naturally defined by $p([x,v,\alpha]) = x$. The chart diffeomorphism is $\varphi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{K}^r \to p^{-1}(\mathcal{U}_\alpha)$,

$$\varphi_\alpha(x,v) = \omega(x,v,\alpha).$$

Now we have to prove that E endowed with the φ_α's is a vector bundle.

First we prove that φ_α is surjective. For this we remark that a general element in $p^{-1}(\mathcal{U}_\alpha)$ can be written under the form $\omega(x,v,\alpha)$ with $x \in \mathcal{U}_{\alpha\beta}$. But

$$\begin{aligned}\varphi_\alpha(x, g_{\alpha\beta}(x)w) &= \omega(x, g_{\alpha\beta}(x)w, \alpha)\\ &= \omega(x, g_{\alpha\beta}(x)g_{\alpha\beta}(x)w, \beta)\\ &= \omega(x, w\beta),\end{aligned} \tag{56.11}$$

then φ_α is surjective. Now we suppose $\varphi_\alpha(x,v) = \varphi_\alpha(y,w)$. Then $\omega(x,v,\alpha) = \omega(y,w,\alpha)$ and $x = y$, $w = g_{\alpha\alpha}v$ which immediately gives $v = w$. Then φ_α is injective.

Finally, we have

$$(\varphi\alpha \circ \varphi_\beta^{-1})(\omega(x,v,\alpha)) = \varphi_\alpha(x, g_{\alpha\beta}(x)v) = \omega(x, g_{\alpha\beta}(x)v, \alpha), \tag{56.12}$$

which proves that the maps g are the transition functions of the vector bundle E.

56.1.3 Equivalence of vector bundle

Let $E \xrightarrow{p} M$ and $F \xrightarrow{p'} M$ be two vector bundles on M. They are **equivalent** if there exists a smooth diffeomorphism $f\colon E \to F$ such that

— $p' \circ f = p$,

— $f|_{E_x}\colon E_x \to F_x$ is a vector space isomorphism.

Let E and F be two equivalent vector bundles, $\{\mathcal{U}_\alpha$ tel que $\alpha \in I\}$, an open covering which trivialize E and F in the same time and ϕ_α^E, ϕ_α^F the corresponding trivializations. A map $f\colon E \to F$ reads "in the trivialization" as $\phi_\alpha^F \circ f|_{p^{-1}(\mu_\alpha)} \circ \phi^{E-1}_\alpha\colon \mathcal{U}_\alpha \times \mathbb{K}^r \to \mathcal{U}_\alpha \times \mathbb{K}^r$ and defines a map $\lambda_\alpha\colon \mathcal{U}_\alpha \to GL(r,\mathbb{K})$ by

$$(\phi_\alpha^F \circ f|_{p^{-1}(\mu_\alpha)} \circ \phi^{E-1}_\alpha)(x,v) = (x, \lambda_\alpha(x)v). \tag{56.13}$$

If we denote by g^E the transition functions for E (and g^F for F),

$$\phi_\alpha^F \circ \phi_\beta^{F-1} = (\phi_\alpha^F \circ f \circ \phi_\alpha^{E-1}) \circ (\phi_\alpha^E \circ \phi_\beta^{E-1}) \circ (\phi_\beta^E \circ f^{-1} \circ \phi_\beta^{E-1}),$$

so that

$$g^F_{\alpha\beta}(x) = \lambda_\alpha(x) g^E_{\alpha\beta}(x) \lambda(x)^{-1}. \tag{56.14}$$

eq:g_l_g_l

Once again we have an inverse construction. We consider a vector bundle E on M with transition functions g^E and some maps $\lambda_\alpha\colon \mathcal{U}_\alpha \to GL(r,\mathbb{K})$; then we define $g^F_{\alpha\beta}(x)$ by equation (56.14).

From subsection 56.1.2, one can construct a vector bundle F on M whose transition functions are these g^F. With the trivializations ϕ^F of F, one can define $f\colon E \to F$ by

$$(\phi_\alpha^F \circ f \circ \phi^{E-1}_\alpha)(x,v) = (x, \lambda_\alpha(x)v).$$

When a basis space B is given, we denote by $\mathrm{Vect}(B)$ the set of isomorphism classes of vector bundles over B. In the complex case, we denote it by $\mathrm{Vect}_{\mathbb{C}}(B)$.

Proposition 56.5.
Any vector bundle over $\mathbb{R}^n$ is trivial.

Proof. Let $p\colon F \to M$ be a vector bundle on $M = \mathbb{R}^n$ and $\{\mathcal{U}_\alpha\}$ be covering of $\mathbb{R}^n$ by local trivializations. Now consider a partition of unity related to the covering $\mathcal{U}_\alpha$: a set of functions $f_\alpha\colon M \to \mathbb{R}$ such that

— $f_\alpha > 0$,
— $\forall x \in M$, one can find a neighbourhood of x in which only a *finite* number of f_α is non zero,
— $\forall x \in M$, $\sum_\alpha f_\alpha(x) = 1$.
— $f_\alpha = 0$ outside of $\mathcal{U}_\alpha$.

Using that partition of unity, we build the trivialization function $f\colon F \to \mathbb{R}^n \times V$ by $f(l) = (x, \sum_\alpha f_\alpha(x)\phi_{\alpha x}(l))$. □

The following two propositions have some importance in K-theory.

PropEoplusEprimetriv

Proposition 56.6.
Let $\pi\colon E \to B$ be a complex vector bundle over a basis compact, Hausdorff, connected basis B. Then there exists a vector bundle E' such that $E \oplus E'$ is trivial.

PropmapfEEsun

Proposition 56.7.
Let $f\colon A \to B$ be a map between the topological spaces A and B, and consider a vector bundle $\pi\colon E \to B$. Then there exists one and only one vector bundle $\pi'\colon E' \to A$ and a map $f'\colon E' \to E$ such that $f'|_{E'_x}\colon E'_x \to E_{f(x)}$ is an isomorphism. The vector bundle E' is unique up to isomorphism.

Proofs can be found in [909]. Let us denote by $f^*(E)$ the function given by proposition 56.7. It satisfies the following properties

$$\begin{aligned} (fg)^*(E) &= g^*\big(f^*(E)\big) \\ \mathrm{Id}^*(E) &= E \\ f^*(E_1 \oplus E_2) &= f^*(E_1) \oplus f^*(E_2) \\ f^*(E_1 \otimes E_2) &= f^*(E_1) \otimes f^*(E_2). \end{aligned} \tag{56.15}$$

EqPropfstarEVect

56.1.4 Sections of vector bundle

DEFooULPGooZeVheH

Definition 56.8.
*A **section** of the vector bundle $p\colon E \to M$ is a smooth map $s\colon M \to E$ such that $p \circ s = \mathrm{Id}|_M$. The set of all the sections is denoted by $\Gamma^\infty(M)$ or simply $\Gamma(E)$.*

PROPooMCJWooNAqmIQ

Proposition 56.9.
Let $E = \big(\{V_x\}_{x\in M}, \{\sigma_\alpha\}\big)$ be a vector bundle. We consider a set of sections $e_i\colon M \to E$ such that for each $x \in M$, the set $\{e_i(x)\}$ is a basis of V_x.

For every $v \in \Gamma^\infty(E)$, there exist smooth functions $v_i \in C^\infty(M)$ such that

$$v(x) = \sum_i v_i(x) e_i(x) \tag{56.16}$$

for every $x \in M$.

If $(\mathcal{U}_\alpha, \phi_\alpha)$ is a local trivialization, one can describe the section s by a function $s_\alpha\colon \mathcal{U}_\alpha \to V$ defined by $\phi_\alpha(s(x)) = (x, s_\alpha(x))$, or equivalently by

$$s(x) = \phi_\alpha^{-1}(x, s_\alpha(x)).$$

As usual when we define such a local quantity, we have to ask ourself how are related s_α and s_β on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. The best is $s_\alpha = s_\beta$, but most of the time it is not. Here, we compute

$$\phi_\beta \circ \phi_\alpha^{-1} \circ \phi_\alpha(s(s)) = (x, g_{\alpha\beta}(x) s_\alpha(x)),$$

which is obviously also equal to $(x, s_\beta(x))$. Then

$$s_\beta(x) = g_{\alpha\beta}(x)s_\alpha(x) \tag{56.17}$$

eq:tr_sec

without summation.

56.1.5 Tensor product

DEFooCSDZooJuzGuE

Proposition-Definition 56.10 ([1, 910]).
Let M be a smooth manifold with charts $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$. We consider two vector bundles[3] *$V_x^{(i)}$, $\sigma_\alpha^{(i)} : U_\alpha \times \mathbb{R}^{k_i} \to \bigcup_{x\in M} V_x^{(i)}$ with transition functions $g_{\alpha\beta}^{(i)} : U_\alpha \cap U_\beta \to \mathrm{GL}(\mathbb{R}, k)$.*

Let $r : \mathbb{R}^{k_1} \otimes \mathbb{R}^{k_2} \to \mathbb{R}^{k_1 k_2}$ be a vector space isomorphism[4].

We consider the following :

(1) $V_x = V_x^{(1)} \otimes V_x^{(2)}$,

(2) $T = \bigcup_{x\in M} V_x$

(3) For each $\alpha \in \Lambda$ we define

$$\begin{aligned} \sigma_\alpha : U_\alpha \times \mathbb{R}^{k_1 k_2} &\to T \\ \big(s, r(\sum_{ij} a_{ij} v_i \otimes w_j)\big) &\mapsto \sum_{ij} a_{ij} \sigma_\alpha^{(1)}(s, v_i) \otimes \sigma_\alpha^{(2)}(s, w_j) \end{aligned} \tag{56.18}$$

EQooXHURooBpOMIU

for every $s \in U_\alpha$, $v_i \in \mathbb{R}^{k_1}$ and $w_j \in \mathbb{R}^{k_2}$.

We have:

ITEMooDTGOooSCsElj

(1) There exists a unique topology τ_T on T such that the maps σ_α are homeomorphic.

ITEMooRALUooFvLQaV

(2) The topological space (T, τ_T) is a topological manifold[5]

(3) The set $\{\sigma_\alpha\}_{\alpha\in\Lambda}$ is an atlas of T.

(4) The topological manifold T with the atlas $\{\sigma_\alpha\}_{\alpha\in\Lambda}$ is a smooth vector bundle of dimension $k_1 k_2$ over M.

(5) The vector bundle does not depend on the choice of the isomorphism r in the sense that two vector bundles build from two different isomorphisms r_1 and r_2 are equivalent.

Proof. There are numerous points to be checked.

Some properties of σ_α.

Before to begin with all the verifications, we check few properties of σ_α.

We introduce the maps

$$\begin{aligned} \tau_{\alpha,s} : \mathbb{R}^{k_1 k_2} &\to V_{\varphi_\alpha(s)} \\ \xi &\mapsto \sigma_\alpha(s, \xi). \end{aligned} \tag{56.19}$$

EQooYICAooRDhYnz

And we prove that $\tau_{\alpha,s}$ is a linear bijection. If $\xi = r\big(\sum_{ij} a_{ij} v_i \otimes w_j\big)$,

$$\tau_{\alpha,s}(\xi) = \sum_{ij} a_{ij} \sigma_\alpha^{(1)}(s, v_i) \otimes \sigma_\alpha^{(2)}(s, w_j). \tag{56.20}$$

3. Definition 56.1.
4. For the dimension, see the proposition 10.179.
5. Definition 49.1.

(1) $\tau_{\alpha,s}$ **is linear** ITEMooCMQEooReGOdu For the linearity, for $\lambda \in \mathbb{R}$ we have

$$\tau_{\alpha,s}(\lambda\xi) = \tau_{\alpha,s}\Big(r\big(\sum_{ij} a_{ij}(\lambda v_i)\otimes w_j\big)\Big) \tag{56.21a}$$
$$= \sum_{ij} a_{ij}\sigma_\alpha^{(1)}(s,\lambda v_i)\otimes\sigma_\alpha^{(2)}(s,w_j) \tag{56.21b}$$
$$= \lambda\sum_{ij} \sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j) \tag{56.21c}$$
$$= \lambda\tau_{\alpha,s}(\xi) \tag{56.21d}$$

because $\sigma_\alpha^{(1)}(s,\lambda v_i) = \lambda\sigma_\alpha^{(1)}(s,v_i)$. The same kind of computations hold to prove that $\tau_{\alpha,s}(\xi_1+\xi_2) = \tau_{\sigma,s}(\xi_1) + \tau_{\alpha,s}(\xi_2)$.

(2) $\tau_{\alpha,s}$ **is injective** Suppose $\tau_{\alpha,s}(\xi) = 0$. Let $\{e_i\}$ be a basis of V and $\{e'_j\}$ be a basis of W. By proposition 10.179(1), the set $\{e_i\otimes e'_j\}$ is a basis of $V\otimes W$. We write $\xi = r\big(\sum_{ij} a_{ij}e_i\otimes e'_j\big)$, so that

$$\sum_{ij} a_{ij}\sigma_\alpha^{(1)}(s,e_i)\otimes\sigma_\alpha^{(2)}(s,e'_j) = 0. \tag{56.22}$$ EQooVGWPooTRDuJW

Since $v\mapsto\sigma_\alpha^{(i)}(s,v)$ is a linear bijection, the set $\{\sigma_\alpha^{(i)}(s,v)\}$ is a basis of $V^{(i)}_{\varphi_\alpha(s)}$. We conclude that the set

$$\{\sigma_\alpha^{(1)}(s,e_i)\otimes\sigma_\alpha^{(2)}(s,e'_j)\} \tag{56.23}$$

is a basis of $V_{\varphi_\alpha(s)} = V^{(1)}_{\varphi_\alpha(s)}\otimes V^{(2)}_{\varphi_\alpha(s)}$. Thus equation (56.22) implies that $a_{ij} = 0$ for every i,j. We conclude that $\xi = 0$ and that $\tau_{\alpha,s}$ is injective.

(3) $\tau_{\alpha,s}$ **is surjective** The set $\{\sigma_\alpha^{(1)}(s,e_i)\otimes\sigma_\alpha^{(2)}(s,e'_j)\}$ is a basis and is in the image of $\tau_{\alpha,s}$ which is linear. Thus $\tau_{\alpha,s}$ is surjective.

(4) σ_α **is injective** ITEMooXSPDooKMXlBN Let $\sigma_\alpha(s_1,t_1) = \sigma_\alpha(s_2,t_2)$. In particular we have

$$\phi\big(\sigma_\alpha(s_1,t_1)\big) = \tag{56.24}$$

Now we introduce the projection map

$$\begin{aligned}\pi\colon T &\to M\\ \xi &\mapsto x \quad \text{if } \xi\in V_x.\end{aligned} \tag{56.25}$$

Let $s\in U_\alpha$ and $t\in\mathbb{R}^{k_1k_2}$. For some v_i and w_j we have

$$\sigma_\alpha(s,t) = \sum_{ij} a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j)\in V^{(1)}_{\varphi_\alpha(s)}\otimes V^{(2)}_{\varphi_\alpha(s)}. \tag{56.26}$$

Thus we have

$$(\pi\circ\sigma_\alpha)(s,t) = \varphi_\alpha(s). \tag{56.27}$$ EQooRSRJooWNtWIA

Proof of (1): topology

We prove that the set T with the maps σ_α satisfy the theorem 49.10.

(1) **Theorem 49.10, hypothesis (1)** The domain of σ_α is $U_\alpha\times\mathbb{R}^{k_1k_2}$ which is open in $\mathbb{R}^{n+k_1k_2}$.

(2) **Theorem 49.10, hypothesis (2)** A generic element of T has the form

$$\xi = \sum_{ij} a_{ij}v_i\otimes w_j = \sum_{ij}(a_{ij}v_i)\otimes w_j \tag{56.28}$$

with $v_i \in V_x^{(1)}$ and $w_j \in V_x^{(2)}$ for some $x \in M$. There exists a α such that $x \in \varphi_\alpha(U_\alpha)$. Let $s \in U_\alpha$ such that $x = \varphi_\alpha(s)$. There exist vectors $v'_{ij} \in \mathbb{R}^{k_1}$ such that

$$a_{ij} v_i = \sigma_\alpha^1(s, v'_{ij}), \tag{56.29}$$

and $w'_j \in \mathbb{R}^{k_2}$ such that $w_j = \sigma_\alpha^2(s, w'_j)$. Now we have

$$\sigma_\alpha\big(s, r(\sum_{ij} v'_{ij} \otimes w'_j)\big) = \sum_{ij} \sigma_\alpha^{(1)}(s, v'_{ij}) \otimes \sigma_\alpha^{(2)}(s, w'_j) \tag{56.30a}$$

$$= \sum_{ij} a_{ij} v_i \otimes w_j \tag{56.30b}$$

$$= \xi. \tag{56.30c}$$

This proves that ξ belongs to $\sigma_\alpha(U_\alpha \times \mathbb{R}^{k_1 k_2})$.

(3) **Theorem 49.10, hypothesis (3)** The fact that σ_α is injective is already proved.

(4) **Theorem 49.10, hypothesis (4)** We have to prove that, for every $\alpha, \beta \in \Lambda$,

$$\sigma_\alpha^{-1}\Big(\sigma_\alpha(U_\alpha \times \mathbb{R}^{k_1 k_2}) \cap \sigma_\beta(U_\beta \times \mathbb{R}^{k_1 k_2})\Big) \tag{56.31}$$

is open in $U_\alpha \times \mathbb{R}^{k_1 k_2}$.

First we prove that

$$\sigma_\alpha(U_\alpha \times \mathbb{R}^{k_1 k_2}) \cap \sigma_\beta(U_\beta \times \mathbb{R}^{k_1 k_2}) = \bigcup_{x \in \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)} V_x^{(1)} \otimes V_x^{(2)}. \tag{56.32}$$

EQooNWFCooWcxQxa

(4a) **First inclusion** Let $\xi \in \sigma_\alpha(S_\alpha \times X_\alpha) \cap \sigma_\beta(S_\beta \times X_\beta)$. There exists $s \in S_\alpha$ and $v \in X_\alpha$ such that, using the condition 56.1(c),

$$\xi = \sigma_\alpha(s, v) \in V_{\varphi_\alpha(s)} = V_{\varphi_\alpha(s)}^{(1)} \otimes V_{\varphi_\alpha(s)}^{(2)}. \tag{56.33}$$

In the same way, there exists $t \in U_\beta$ and $w \in \mathbb{R}^{k_1 k_2}$ such that

$$\xi = \sigma_\beta(t, w) \in V_{\varphi_\beta(t)} = V_{\varphi_\beta(t)}^{(1)} \otimes V_{\varphi_\beta(t)}^{(2)}. \tag{56.34}$$

We deduce that $\varphi_\alpha(s) = \varphi_\beta(t)$, which is, by the way, $\pi(\xi)$. Thus, setting $x = \varphi_\alpha(s)$ we have

$$\xi \in V_x^{(1)} \otimes V_x^{(2)}. \tag{56.35}$$

(4b) **Second inclusion** Let $\xi \in V_x^{(1)} \otimes V_x^{(2)}$ with $x \in \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$. We can write $\xi = \sum_{ij} a_{ij} v_i \otimes w_j$ with $v_i \in V_x^{(1)}$ and $w_j \in V_x^{(2)}$. Since the map $\sigma_\alpha^{(1)} : U_\alpha \times \mathbb{R}^{k_1} \to \bigcup_{x \in M} V_x^{(1)}$ is a chart, there exist $v'_i \in \mathbb{R}^{k_1}$ such that

$$v_i = \sigma_\alpha^{(1)}\big(\varphi_\alpha^{-1}(x), v'_i\big). \tag{56.36}$$

In the same way, there exists $w'_{ij} \in \mathbb{R}^{k_2}$ such that

$$a_{ij} w_j = \sigma_\alpha^{(2)}\big(\varphi_\alpha^{-1}(x), w'_{ij}\big). \tag{56.37}$$

Few computations:

$$\xi = \sum_{ij} a_{ij} v_i \otimes w_j \tag{56.38a}$$

$$= \sum_{ij} v_i \otimes (a_{ij} w_j) \tag{56.38b}$$

$$= \sum_{ij} \sigma_\alpha^{(1)}\big(\varphi_\alpha^{-1}(x), v'_i\big) \otimes \sigma_\alpha^{(2)}\big(\varphi_\alpha^{-1}(x), w'_{ij}\big) \tag{56.38c}$$

$$= \sigma_\alpha\Big(\varphi_\alpha^{-1}(x), r\big(\sum_{ij} v'_i \otimes w'_{ij}\big)\Big) \tag{56.38d}$$

$$\in \sigma_\alpha(U_\alpha \times \mathbb{R}^{k_1 k_2}). \tag{56.38e}$$

In the same way we see that $\xi \in \sigma_\beta(U_\beta \times \mathbb{R}^{k_1 k_2})$. Equation (56.32) is proven.

Now let A be a part of U_α. We prove that

$$\sigma_\alpha^{-1}\Big(\bigcup_{x\in\varphi_\alpha(A)} V_x^{(1)}\otimes V_x^{(2)}\Big) = A\times\mathbb{R}^{k_1k_2}. \tag{56.39}$$

(4a) **First inclusion** Let $\xi\in V_x^{(1)}\otimes V_x^{(2)}$ and let $s\in U_\alpha$ be such that $x=\varphi_\alpha(s)$. We have $\xi=\sum_{ij}a_{ij}v_i\otimes w_j$ with $v_i=\sigma_\alpha^{(1)}(s,v_i')$ and $a_{ij}w_j=\sigma_\alpha^{(2)}(s,w_{ij}')$. Now we have

$$\begin{aligned}\xi &= \sum_{ij}\sigma_\alpha^{(1)}(s,v_i')\otimes\sigma_\alpha^{(2)}(s,w_{ij}') && (56.40a)\\ &= \sigma_\alpha\big(s,r(\sum_{ij}v_i'\otimes w_{ij}')\big) && (56.40b)\\ &\in \sigma_\alpha\big(\varphi_\alpha^{-1}(x),\mathbb{R}^{k_1k_2}\big). && (56.40c)\end{aligned}$$

(4b) **Second inclusion** Let $(s,\omega)\in A\times\mathbb{R}^{k_1k_2}$. We have to prove that $\sigma_\alpha(s,\omega)\in\bigcup_{x\in\varphi_\alpha(A)}V_x^{(1)}\otimes V_x^{(2)}$. For that, notice that $\omega\in\mathbb{R}^{k_1k_2}$ can be written as $\omega=r\big(\sum_{ij}a_{ij}v_i\otimes w_j\big)$ for some $v_i\in\mathbb{R}^{k_1}$ and $w_j\in\mathbb{R}^{k_2}$. Now compute

$$\sigma_\alpha(s,\omega)=\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2))}(s,w_j). \tag{56.41}$$

Since $\sigma_\alpha^{(1)}(s,v_i)\in V_{\varphi_\alpha(s)}$ and $\sigma_\alpha^{(2)}(s,w_j)\in V^{(2)}_{\varphi_\alpha(s)}$ we found that

$$\sigma_\alpha(s,\omega)\in V^{(1)}_{\varphi_\alpha(s)}\otimes V^{(2)}_{\varphi_\alpha(s)} \tag{56.42}$$

The equality (56.39) is proved.

Recall what we have in equations (56.32) and (56.39) :

$$\begin{aligned}&\sigma_\alpha(U_\alpha\times\mathbb{R}^{k_1k_2})\cap\sigma_\beta(U_\beta\times\mathbb{R}^{k_1k_2}) && (56.43a)\\ &\quad=\bigcup_{x\in\varphi_\alpha(U_\alpha)\cap\varphi_\beta(U_\beta)} V_x^{(1)}\otimes V_x^{(2)} && (56.43b)\\ &\quad=\Big(\bigcup_{x\in\varphi_\alpha(U_\alpha)} V_x^{(1)}\otimes V_x^{(2)}\Big)\cap\Big(\bigcup_{x\in\varphi_\beta(U_\beta)} V_x^{(1)}\otimes V_x^{(2)}\Big) && (56.43c)\end{aligned}$$

and

$$\sigma_\alpha^{-1}\Big(\bigcup_{x\in\varphi_\alpha(A)} V_x^{(1)}\otimes V_x^{(2)}\Big) = A\times\mathbb{R}^{k_1k_2}. \tag{56.44}$$

The map σ_α^{-1} is injective, so that we can use proposition 1.8. Now we make the final computation, writing $M_\alpha=\varphi_\alpha(U_\alpha)$:

$$\begin{aligned}&\sigma_\alpha^{-1}\big(\sigma_\alpha(U_\alpha\times\mathbb{R}^{k_1k_2})\cap\sigma_\beta(U_\beta\times\mathbb{R}^{k_1k_2})\big) && (56.45a)\\ &=\sigma_\alpha^{-1}\Big(\bigcup_{x\in\varphi_\alpha(U_\alpha)} V_x^{(1)}\otimes V_x^{(2)}\Big)\cap\sigma_\alpha^{-1}\Big(\bigcup_{x\in\varphi_\beta(U_\beta)} V_x^{(1)}\otimes V_x^{(2)}\Big) && \text{eq. (56.43c)} \quad (56.45b)\\ &=\big(U_\alpha\cap\mathbb{R}^{k_1k_2}\big)\cap\sigma_\alpha^{-1}\Big(\bigcup_{x\in\varphi_\alpha\left((\varphi_\alpha^{-1}\circ\varphi_\beta)(U_\beta)\right)} V_x^{(1)}\otimes V_x^{(2)}\Big) && \text{eq. (56.44)} \quad (56.45c)\\ &=\big(U_\alpha\times\mathbb{R}^{k_1k_2}\big)\cap\big((\varphi_\alpha^{-1}\circ\varphi_\beta)(U_\beta)\times\mathbb{R}^{k_1k_2}\big) && \text{eq. (56.44)} \quad (56.45d)\\ & && (56.45e)\end{aligned}$$

We know that $(\varphi_\alpha^{-1}\circ\varphi_\beta)(U_\beta)$ is open in U_α since φ_α and φ_β are charts of M.

(5) **Theorem 49.10, hypothesis (5)**
We have to prove that

$$\sigma_\beta^{-1}\circ\sigma_\alpha\colon \sigma_\alpha^{-1}\Big(\sigma_\alpha(U_\alpha\times\mathbb{R}^{k_1k_2})\cap\sigma_\beta(U_\beta\times\mathbb{R}^{k_1k_2})\Big)\to U_\beta\times\mathbb{R}^{k_1k_2} \tag{56.46}$$

is continuous. In fact, we will prove that this map is even smooth. It is the same price and we will need it later.

(5a) **First step**
We have :

$$(\sigma_\beta^{-1}\circ\sigma_\alpha)\Big(s,r\big(\sum_{ij}a_{ij}v_i\otimes w_j\big)\Big)=\sigma_\beta^{-1}\Big(\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j)\Big). \tag{56.47}$$

(5b) **Notation A_1 and A_2** Small point about the notations. When we have a map $A\colon U_\alpha\times\mathbb{R}^{k_1}\to U_\beta\times\mathbb{R}^{k_2}$, we write $A_1(s,v)$ the U_β component of $A(s,v)$ and $A_2(s,v)$ the component in $\mathbb{R}^{k_2}$.
With that notation we have

$$(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_1(s,v)=(\varphi_\beta^{-1}\circ\varphi_\alpha)(s). \tag{56.48}$$

Indeed, by definition 56.1(d), we have $\sigma_\alpha^{(1)}(s,v)\in V^{(1)}_{\varphi_\alpha(s)}$, and by the same point,

$$\sigma_\beta^{(1)-1}\big(V_{\varphi_\alpha(s)}\big)=\big((\varphi_\beta^{-1}\circ\varphi_\alpha)(s),\mathbb{R}^{k_1}\big). \tag{56.49}$$

The first component is $(\varphi_\beta^{-1}\circ\varphi_\alpha)(s)$ as expected.

(5c) **Next step** We continue (56.47) by proving

$$\begin{aligned}&\sigma_\beta^{-1}\Big(\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j)\Big)\\&=\Big((\varphi_\beta^{-1}\circ\varphi_\alpha)(s),r\big(\sum_{ij}a_{ij}(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_2(s,v_i)\otimes(\sigma_\beta^{(2)-1}\circ\sigma_\alpha^{(2)})_2(s,w_j)\big)\Big).\end{aligned} \tag{56.50}$$

Notice that, setting $x=(\varphi_\beta^{-1}\circ\varphi_\alpha)(s)$, the equation (56.48) says that $(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_1(s,v_i)=x$, so that

$$\big(x,(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_2(s,v_i)\big)=(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})(s,v_i). \tag{56.51}$$

Now we prove (56.50) by applying σ_β to the right hand side :

$$\sigma_\beta\Big(x,r\big(\sum_{ij}a_{ij}(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_2(s,v_i)\otimes(\sigma_\beta^{(2)-1}\circ\sigma_\alpha^{(2)})_2(s,w_j)\big)\Big) \tag{56.52a}$$

$$=\sum_{ij}a_{ij}\sigma_\beta^{(1)}\big(x,(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})_2(s,v_i)\big)\otimes\sigma_\beta^{(2)}\big(x,(\sigma_\beta^{(2)-1}\circ\sigma_\alpha^{(2)})_2(s,w_i)\big) \tag{56.52b}$$

$$=\sum_{ij}a_{ij}\sigma_\beta^{(1)}\big((\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(1)})(s,v_i)\big)\otimes\sigma_\beta^{(2)}\big((\sigma_\beta^{(2)-1}\circ\sigma_\alpha^{(2)})(s,w_i)\big) \quad \text{eq. (56.51)} \tag{56.52c}$$

$$=\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j). \tag{56.52d}$$

(5d) **Conclusion** We have

$$(\sigma_\beta^{-1}\circ\sigma_\alpha)\Big(s,r\big(\sum_{ij}a_{ij}v_i\otimes w_j\big)\Big) \tag{56.53a}$$

$$=\sigma_\beta^{-1}\Big(\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j)\Big) \tag{56.53b}$$

$$=\Big((\varphi_\beta^{-1}\circ\varphi_\alpha)(s),r\big[\sum_{ij}a_{ij}(\sigma_\beta^{(2)-1}\circ\sigma_\alpha^{(1)})_2(s,v_i)\otimes(\sigma_\beta^{(1)-1}\circ\sigma_\alpha^{(2)})_2(s,w_j)\big]\Big). \tag{56.53c}$$

Since the maps φ_α, φ_β are charts of M, the composition $\varphi_\beta^{-1}\circ\varphi_\alpha$ is smooth. The same way, the maps $\sigma_\alpha^{(i)}$ and $\sigma_\beta^{(i)}$ are charts of $\bigcup_{x\in M} V_x^{(i)}$, the composition $\sigma_\beta^{(i)-1}\circ\sigma_\alpha^{(}i)$ is smooth.

At this point, the theorem 49.10(5) gives us a unique topology on T such that the maps σ_α are homeomorphic. In order to turn T into a topological manifold, we have to show that this topology is Hausdorff and second countable.

(1) **Hausdorff** Let $\xi_1,\xi_2\in T$. There are two possibilities: either $\pi(\xi_1)=\pi(\xi_2)$ or not.
If $\pi(\xi_1)=\pi(\xi_2)$, then by writing $\pi(\xi_1)=\pi(\xi_2)=\varphi_\alpha(s)$, we have $\xi_1=\sigma_\alpha(s,t_1)$ and $\xi_2=\sigma_\alpha(s,t_2)$. Let V be an open neighbourhood of s. Since $\mathbb{R}^{k_1k_2}$ is Hausdorff, we can consider open parts A and B such that $t_1\in A$, $t_2\in B$ and $A\cap B=\varnothing$. Then the open parts $\sigma_\alpha(V,A)$ and $\sigma_\alpha(V,B)$ are open sets separating ξ_1 and ξ_2.
If $\pi(\xi_1)\neq\pi(\xi_2)$, we use the fact that M is Hausdorff. Let A be an open neighbourhood of $\pi(\xi_1)$ and B be an open neighbourhood of $\pi(\xi_2)$ with $A\cap B=\varnothing$. Then the open parts $\sigma_\alpha(A,\mathbb{R}^{k_1k_2})$ and $\sigma_\alpha(B,\mathbb{R}^{k_1k_2})$ are separating ξ_1 and ξ_2 because σ_α is injective, see point (4).

(2) **Second countable**

(2a) **The set $\mathcal{B}$** By proposition 49.16, the manifold M has a countable atlas $\{(W_s,\phi_s)\}_{s\in\Sigma}$ with a map $f:\Sigma\to\Lambda$ such that $W_s\subset U_{f(s)}$ and $\phi_s=\varphi_{f(s)}$. When $s\in\Sigma$, we write $\sigma_s=\sigma_{f(s)}$.
Let $\{B_i\}_{i\in\mathbb{N}}$ be a countable basis of topology of $\mathbb{R}^{k_1k_2}$. Each W_s being part of $\mathbb{R}^n$, we consider countable basis of topology $\{W_{s,i}\}_{i\in\mathbb{N}}$ of W_s. Then we consider

$$\mathcal{A}=\{\sigma_s(W_{s,i}\times B_j)\}_{\substack{s\in\Sigma\\ i,j\in\mathbb{N}}}. \tag{56.54}$$

The set $\mathcal{A}$ is countable as product of countable sets by the proposition 1.141(2). Finally we define $\mathcal{B}$ as the set of finite intersections of elements of $\mathcal{A}$. The set $\mathcal{B}$ is countable as union of finites parts of a countable set (lemma 1.139).

(2b) **Elements of τ_0**
We show that every element of τ_0 contains an element of $\mathcal{B}$. Let $\mathcal{O}\in\tau_0$. There exists an open $A\subset U_\alpha\times\mathbb{R}^{k_1k2}$. Then we can choose $s\in\Sigma$, $i,j\in\mathbb{N}$ such that

$$W_{s,i}\times B_j\subset A. \tag{56.55}$$

Notice that $f(s)=\alpha$. Then the element $\sigma_\alpha(W_{s,i}\times B_j)$ is contained in $\mathcal{O}$.

(2c) **Finite intersection**
Let $\mathcal{O}_1=\sigma_\alpha(A_1)$ and $\mathcal{O}_2=\sigma_\beta(A_2)$ be two elements of τ_0. Here A_1 is open in $U_\alpha\times\mathbb{R}^{k_1k_2}$ and A_2 is open in $U_\beta\times\mathbb{R}^{k_1k_2}$. There exist $s,t\in\Sigma$, $i,j,k,l\in\mathbb{N}$ such that

$$\begin{aligned}W_{s,i}\times B_k&\subset A_1\\ W_{t,j}\times B_l&\subset A_2.\end{aligned} \tag{56.56}$$

So the part

$$\sigma_\alpha(W_{s,i}\times B_k)\cap\sigma_\beta(W_{t,j}\times B_l)\subset\mathcal{O}_1\cap\mathcal{O}_2. \tag{56.57}$$

(2d) **Union**
If the sets $\{\mathcal{O}_i\}_{i\in I}$ are (each) intersections of elements of τ_0, then, in particular $\mathcal{O}_1\subset\bigcup_{i\in I}\mathcal{O}_i$ and we can find an element in $\mathcal{B}$ included in $\mathcal{O}_1$.

Since the spaces $V_x^{(i)}$ are distinct, the spaces $V_x=V_x^{(1)}\otimes V_x^{(2)}$ are distinct. We have to check the conditions of the definition 56.1 of a vector bundle. The set T is a topological manifold and the maps σ_α are smooth (see equation (56.46) and bellow). Thus this is a smooth manifold.

(1) **Condition (a)**
Already done, mainly by the theorem 49.10.

(2) **Condition (b)**

We have to prove that [6]

$$\begin{aligned}\pi\colon \bigcup_{x\in M} V_x^{(1)}\otimes V_x^{(2)} &\to M\\ \sum_{ij} a_{ij}v_i\otimes w_j &\mapsto x\end{aligned} \tag{56.58}$$

is smooth. In other words, for each α and β, the map

$$\varphi_\beta^{-1}\circ\pi\circ\sigma_\alpha\colon U_\alpha\to U_\beta \tag{56.59}$$

must be smooth. We have

$$(\varphi_\beta^{-1}\circ\pi\circ\sigma_\alpha)\big(s,r(\sum_{ij}a_{ij}v_i\otimes w_j)\big) \tag{56.60a}$$

$$=(\varphi_\beta^{-1}\circ\pi)\Big(\underbrace{\sum_{ij}a_{ij}\sigma_\alpha^{(1)}(s,v_i)\otimes\sigma_\alpha^{(2)}(s,w_j)}_{\in V_{\varphi_\alpha(s)}}\Big) \tag{56.60b}$$

$$=(\varphi_\beta^{-1}\circ\varphi_\alpha)(s). \tag{56.60c}$$

So we have

$$(\varphi_\beta^{-1}\circ\pi\circ\sigma_\alpha)(s,\xi)=(\varphi_\beta^{-1}\circ\varphi_\alpha)(s), \tag{56.61}$$

which leads us to $\varphi_\beta^{-1}\circ\pi\circ\sigma_\alpha=\varphi_\beta^{-1}\circ\varphi_\alpha\circ\texttt{proj}$ where $\texttt{proj}\colon\mathbb{R}^n\times\mathbb{R}^{k_1k_2}\to\mathbb{R}^n$ is the projection. Since `proj` is smooth and $\varphi_\beta^{-1}\circ\varphi_\alpha$ is smooth, we conclude that $\varphi_\beta^{-1}\circ\pi\circ\sigma_\alpha$ is smooth.

(3) **Condition (c)** Since $\sigma_\alpha^{(i)}$ is a chart for the vector bundle $\bigcup_{x\in M}V_x^{(i)}$ we have $\sigma_\alpha^i(s,v)\in V_{\varphi_\alpha(s)}$. Thus for right hand side of (56.18) is an element of $V^{(1)}_{\varphi_\alpha(s)}\otimes V^{(2)}_{\varphi_\alpha(s)}=V_{\varphi_\alpha(s)}$.

(4) **Condition (d)** We fix $s\in U_\alpha$. We have to prove that the map $\tau_{\alpha,s}$ is a linear bijection. This is already done after the definition ((1)), with the items (1) and the following ones.

(5) **Condition (e)** Two points to be proven.

(5a) $g_{\alpha\beta}(s)\in\mathbf{GL}(\mathbb{R},k_1k_2)$ The condition

$$(\sigma_\alpha^{-1}\circ\sigma_\beta)(s,v)=\big(s,g_{\alpha\beta}(s)v\big) \tag{56.62}$$

for every $v\in\mathbb{R}^{k_1k_2}$ says

$$g_{\alpha\beta}(s)v=(\tau_{\alpha,s}^{-1}\circ\tau_{\alpha,s})(v). \tag{56.63}$$

Thus $g_{\alpha\beta}(s)$ is linear and invertible. This means that for every s, the map $g_{\alpha\beta}(s)$ belongs to $\mathrm{GL}(\mathbb{R},k_1k_2)$.

(5b) $g_{\alpha\beta}$ **is smooth** We are going to prove that

$$g_{\alpha\beta}(s)=r\circ\big(g_{\alpha\beta}^{(1)}(s)\otimes g_{\alpha\beta}^{(2)}(s)\big)\circ r^{-1}. \tag{56.64}$$

EQooAUMXooCoqPTQ

Let $\{e_i\}_{i\in I}$ be a basis of $\mathbb{R}^k_1$ and $\{e'_j\}_{j\in J}$ be a basis of $\mathbb{R}^{k_2}$. We consider $v=r(e_i\otimes e'_j)$

6. Here we use the fact that the spaces $V_x^{(i)}$ have the form $\{(x,v)\text{ tel que }v\in W_x^{(i)}\}$, so that when we write $v_i\otimes w_j$, we could have written $(x,v_i)\otimes(x,w_j)$. Just keep in mind that the base point x is contained in the vector v_i in the following formula.

and we compute [EQSooUIMFooMBbBvp]

$$\begin{aligned}
\sigma_\alpha\Big(s, r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big)\circ r^{-1}v\Big) &= \sigma_\alpha\Big(s, r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big)(e_i\otimes e'_j)\Big) && (56.65a)\\
&= \sigma_\alpha\Big(s, r\big(g^{(1)}_{\alpha\beta}(s)e_i\otimes g^{(2)}_{\alpha\beta}(s)e'_j\big)\Big) && (56.65b)\\
&= \sigma^{(1)}_\alpha\big(s, g^{(1)}_{\alpha\beta}(s)e_i\big)\otimes\sigma^{(2)}_\alpha\big(s, g^{(2)}_{\alpha\beta}e'_j\big) && (56.65c)\\
&= \sigma^{(1)}_\beta(s, e_i)\otimes\sigma^{(2)}_\beta(s, e'_j) && \text{cf. justif.}\\
& && \text{SUBEQooSZDEooLQBCoQ}\ (56.65d)\\
&= \sigma_\beta\big(s, r(e_i\otimes e'_j)\big) && (56.65e)\\
&= \sigma_\beta\big(s, r(v)\big). && (56.65f)
\end{aligned}$$

Justifications.

— For (56.65d). We have $\big(s, g^{(1)}_{\alpha\beta}(s)e_i\big) = \big(\sigma^{(1)-1}_\alpha\circ\sigma_\beta\big)(s, e_i)$, and similar for e'_j.

Putting σ_α on the right hand side the equalities (56.65) show that

$$(\sigma^{-1}_\alpha\circ\sigma_\beta)\big(s, r(v)\big) = \Big(s, r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big)r^{-1}v\Big). \tag{56.66}$$

The left hand side can be expressed in terms og $g_{\alpha\beta}(s)$. We have

$$\big(g_{\alpha\beta}(s)\circ r\big)(e_i\otimes e'_j) = r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big)(e_i\otimes e'_j). \tag{56.67}$$

Since the operators $g_{\alpha\beta}(s)\circ r$ and $r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big)$ are linear and since the elements $e_i\otimes e'_j$ form a basis we have the operator equality

$$g_{\alpha\beta}(s)\circ r = r\circ\big(g^{(1)}_{\alpha\beta}(s)\otimes g^{(2)}_{\alpha\beta}(s)\big), \tag{56.68}$$

which is the announced equality (56.64).

(5c) **Conclusion** The maps $g^{(1)}_{\alpha\beta}$ and $g^{(2)}_{\alpha\beta}$ are smooths by hypothesis. By proposition 55.19, the tensor product is smooth. The maps r and r^{-1} are linear on finite dimensional vector spaces; they are smooth too.

□

PROPooEYSWooOeQNyX

Proposition 56.11 ([1])**.**
The tensor product[7] *of vector bundles is associative : if E_1, E_2 and E_3 are vector bundles, then*

$$(E_1\otimes E_2)\otimes E_3 = E_1\otimes(E_2\otimes E_3). \tag{56.69}$$

56.2 Tensor bundle

Definition 56.12.
*Let (k, l) be integers. The k, l-**tensor bundle** is*

$$\bigotimes^{k,l}(TM) = (TM)^{\otimes k}\otimes(T^*M)^{\otimes l}. \tag{56.70}$$

*A (k, l)-**tensor field** is a smooth section of $\bigotimes^{k,l}(TM)$.*

We know from theorem 49.92 that TM is a manifold. Here we prove that this is a vector bundle.

7. Definition 56.10.

PROPooIMOJooKzEDXA

Proposition 56.13.
The set TM becomes a vector bundle with the charts

$$\begin{aligned}\sigma_\alpha : U_\alpha \times \mathbb{R}^n &\to TM \\ (s, v) &\mapsto \frac{d}{dt}\left[\varphi_\alpha(s + tv)\right]_{t=0}.\end{aligned} \tag{56.71}$$

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

56.2.1 Dual basis

The proposition 49.56 says that the vectors

$$\partial_i = \frac{d}{dt}\left[\varphi_\alpha(s_0 + te_i)\right]_{t=0} \tag{56.72}$$

form a basis of T_pM where $p = \varphi_\alpha(s_0)$. We denote by $\{\partial_i^*\}$ the dual basis.

56.14.
When $I = (i_1, \ldots, i_k)$ is a multiindex, we write

$$\partial^*_{\wedge I} = \partial_{i_1} \wedge \ldots \wedge \partial_{i_k}. \tag{56.73}$$

56.15.
The object $\partial^*_{\wedge I}$ is build from the elements of the basis ∂_i^* which is the dual of the basis

$$\partial_i = \frac{d}{dt}\left[\varphi_\alpha(s + te_i)\right]_{t=0}. \tag{56.74}$$

Thus $\partial^*_{\wedge I}$ is a function of s.

LEMooKCBSooPjEwEl

Lemma 56.16 ([1])**.**
Let (U, φ) be a chart of a smooth manifold. We have the local expression

$$\partial_i^* = df_i \tag{56.75}$$

where f_i is the function

$$\begin{aligned}f_i : M &\to \mathbb{R} \\ q &\mapsto \varphi^{-1}(q)_i.\end{aligned} \tag{56.76}$$

Proof. Let $q = \varphi(x) \in M$. At q we have $\partial_j = \frac{d}{dt}\left[\varphi(x + te_j)\right]_{t=0}$ and, using the definition $f_i = \texttt{proj}_i \circ \varphi^{-1}$, we have

$$\begin{aligned}df_i(\partial_j) &= \frac{d}{dt}\left[f_i\big(\varphi(x + te_j)\big)\right]_{t=0} && (56.77a)\\ &= \frac{d}{dt}\left[(\texttt{proj}_i \circ \varphi^{-1} \circ \varphi)(x + te_j)\right]_{t=0} && (56.77b)\\ &= \frac{d}{dt}\left[\texttt{proj}_i(x + te_j)\right]_{t=0} && (56.77c)\\ &= \frac{d}{dt}\left[x_i + t\delta_{ij}\right]_{t=0} && (56.77d)\\ &= \delta_{ij}, && (56.77e)\end{aligned}$$

which proves that $df_i = \partial_i^*$. □

LEMooGEFSooVlPLOs

Lemma 56.17 ([1])**.**
Let f be a smooth function over the manifold M. We consider a chart (U, φ). The differential of f is

$$df = \sum_i (\partial_i f)\partial_i^*. \tag{56.78}$$

Proof. Let X be the vector $X = \sum_i \partial_i$. Since $\{\partial_i^*\}$ is the dual basis, we express the components of X under the form $X_i = \partial_i^*(X)$. Now, by definition,

$$df(X) = X(f) = \sum_i X_i \partial_i f = \sum_i (\partial_i f)\partial_i^*(X). \tag{56.79}$$

Since it is valid for every vector X, we have the equality. □

LEMooIAASooDnnhRA

Lemma 56.18.
The map

$$\partial_{\wedge I}^* \colon U_\alpha \to \bigwedge^k (T^*M) \tag{56.80}$$

is smooth and satisfy

$$\partial_{\wedge I}^*(s) \in \bigwedge^k \big(T^*_{\varphi_\alpha(s)} M\big). \tag{56.81}$$

NORMooDQEZooHaJPoa

56.19.
Let $\varphi_\alpha \colon U_\alpha \to M$ be a chart. We introduce two maps to make the link between the real world and the manifold world. One for the forms :

$$\begin{aligned} \tau_{\alpha,s} \colon \bigwedge^k (\mathbb{R}^n)^* &\to \bigwedge^k T^*_{\varphi_\alpha(s)} M \\ \sum_{I \in C_k} \omega_I e^*_{\wedge I} &\mapsto \sum_{I \in C_k} \omega_I \partial^*_{\wedge I}. \end{aligned} \tag{56.82}$$

An one for the vectors:

$$\begin{aligned} \tau_{\alpha,s} \colon \mathbb{R}^n &\to T_{\varphi_\alpha(s)} M \\ v &\mapsto \sum_i v_i \partial_i. \end{aligned} \tag{56.83}$$

These two maps have the same name «$\tau_{\alpha,s}$», but their arguments are different enough not to cause confusion.

DEFooZELVooFfosEn

Proposition-Definition 56.20 (Exterior bundle)**.**
Let M be a n-dimensional manifold. We write[8]

$$C_k = \{(i_1, \dots, i_k) \text{ tel que } 1 \leqslant i_1 < \dots < i_k \leqslant n\}. \tag{56.84}$$

The set $\bigwedge^k T^*M$ is a vector bundle with the charts

$$\begin{aligned} \sigma_\alpha \colon U_\alpha \times \bigwedge^k (\mathbb{R}^n)^* &\to \bigwedge^k T^*M \\ (s, \sum_{I \in C_k} \omega_I e^*_{\wedge I}) &\mapsto \Big(\varphi_\alpha(s), \sum_{I \in C_k} \omega_I \partial^*_{\wedge I}\Big). \end{aligned} \tag{56.85}$$

EQooOMWMooKKWhqf

In other words,

$$\sigma_\alpha(s, \omega') = \big(\varphi_\alpha(s), \tau_{\alpha,s}(\omega')\big). \tag{56.86}$$

Two notes.

(1) There is an hidden choice of isomorphism between $\mathbb{R}^m$ and $\bigwedge^k \mathbb{R}^n$ with $m = \binom{n}{k}$.

(2) The vector space $V_{\varphi_\alpha(s)}$ is the set of elements of the form $\big(\varphi_\alpha(s), \omega\big)$ with $\omega \in \bigwedge^k T^*_{\varphi_\alpha(s)} M$. The purpose of that is to make explicit the fact that all the vector spaces V_x are distinct. See the point (1) of the definition 56.1 of a vector bundle.

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

8. This provides a basis of $\bigwedge^k T^*m$, see proposition 10.216.

56.3 Differential form

DEFooBRSSooWOgoov

Definition 56.21.
A k***-differential-form*** *is a section of* $\bigwedge^k T^*M$. *We denote by* $\Omega^k(M)$ *the set of* k*-differential forms.*

When the context is clear, we could simply say "k-form" instead of k-differential form.

56.22.
Formally, a k-differential form ω is a smooth map

$$\omega\colon M \to M \times \bigwedge^k T^*M \tag{56.87}$$

where the M-component of ω is the identity.

We will, however, often make the abuse of notation and drop that M-component.

PROPooQOUXooZfFHNb

Proposition-Definition 56.23.
If $\omega \in \Omega^k(M)$ *and* $\eta \in \Omega^l(M)$ *we define* $\omega \wedge \eta$ *pointwise by*

$$(\omega \wedge \eta)_p = \omega_p \wedge \eta_p \tag{56.88}$$

where the exterior product is defined by 10.211.

We have $\omega \wedge \eta \in \Omega^{k+l}(M)$.

LEMooUVRQooVZDKir

Lemma 56.24 ([1])**.**
If f *is a function and* ω *is a* k*-differential form we have*

EQooYHCKooNLgJWG

$$f \wedge \omega = f\omega. \tag{56.89}$$

where, on the left hand side, f *is seen as a* 0*-differential form while, on the right hand side this is the pointwise product.*

Proof. We compute the left-hand side of (56.89) at $x \in M$. From definition 56.23,

$$(f \wedge \omega)_x = f(x) \wedge \omega_x. \tag{56.90}$$

Now we are reduced to check $(\lambda \otimes \omega) = \lambda\omega$ for $\lambda \in \mathbb{R}$ and $\omega \in \bigwedge^k(V^*)$ for some vector space V. Using the definitions 10.211 and 10.206,

$$\begin{aligned}
\lambda \wedge \omega &= \frac{(0+k)!}{0!k!} \operatorname{Alt}(\lambda \otimes \omega) && \text{(56.91a)}\\
&= \operatorname{Alt}(\lambda \otimes \omega) && \text{(56.91b)}\\
&= \frac{1}{k!} \sum_{\pi \in S_k} (-1)^\pi (\lambda \otimes \omega)^\pi. && \text{(56.91c)}
\end{aligned}$$

Each term in the sum is subject to the relation $(\lambda \otimes \omega)(v_1, \dots, v_k) = \lambda\omega(v_1, \dots, v_k)$. □

The following proposition works for a manifold without boundary. If you have a manifold with boundary, see proposition 57.45.

PROPooNOGKooQedJba

Proposition 56.25.
Let $\omega \in \Omega^k(M)$ *and* $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in \Lambda}$ *be an atlas of* M. *For each* $\alpha \in \Lambda$, *the exist functions* $\omega_{\alpha,I} \in C^\infty(M)$ *for each*

$$I \in C_k = \{(i_1, \dots, i_k) \text{ tel que } 1 \leqslant i_1 < \dots < i_k \leqslant n\} \tag{56.92}$$

such that

$$\omega_x = \sum_{I \in C_k} \omega_{\alpha,I}(x) \partial^*_{\alpha,\wedge I}. \tag{56.93}$$

for every $x \in \varphi_\alpha(U_\alpha)$. *Notes:*

— *When $I = (i_1, \dots, i_k)$, the notation $a_{\wedge I}$ means $a_{i_1} \wedge \dots \wedge a_{i_k}$.*

— *Even if it not written explicitly, ∂_i^* is a vector field, and then a function of x.*

Proof. No proof for the moment. You can help sending me a proof based on theorems already stated before. □

PROPooHTKMooVzYQdW

Proposition 56.26.
We consider a k-form $\omega \in \Omega^k(M)$ and smooth vector fields $X_1, \dots, X_k$. The map

$$\begin{aligned} \omega(X_1, \dots, X_k) \colon M &\to \mathbb{R} \\ x &\mapsto \omega_x\big((X_1)_x, \dots, (X_k)_x\big) \end{aligned} \tag{56.94}$$

is smooth.

Proof. We have to prove that $\omega(X_1, \dots, X_k) \circ \varphi_\alpha$ is smooth as function from U_α to $\mathbb{R}$. Using proposition 56.25, we have smooth functions $\omega_I \in C^{\infty(M,\mathbb{R})}$ such that

$$\omega_x = \sum_{i \in C_k} \omega_I(x) \partial^*_{\wedge I}. \tag{56.95}$$

From proposition 49.85, there exist smooth functions $X_{ij} \in C^{\infty(M)}$ such that

$$X_i(x) = \sum_j X_{ij}(x) \partial_j. \tag{56.96}$$

We have

$$\omega(X_1, \dots, X_k)\varphi_\alpha(s) = \sum_{I \in C_k} (\omega_I \circ \varphi_\alpha)(s) \partial^*_{\wedge I}\Big((X_1 \circ \varphi_\alpha)(s), \dots, (X_k \circ \varphi_\alpha)(s)\Big) \tag{56.97a}$$

$$= \sum_I (\omega_I \circ \varphi_\alpha)(s)(\partial^*_{i_1} \wedge \dots \partial^*_{i_k})\Big((X_1 \circ \varphi_\alpha)(s), \dots, (X_k \circ \varphi_\alpha)(s)\Big). \tag{56.97b}$$

SUBEQooKPRLooFzrsMJ

The function $\omega_I \circ \varphi_\alpha$ is smooth. The rest of (56.97b) are sum of products of elements of the form [9]

$$\partial^*_{i_m}(x)\Big((X_l \circ \varphi_\alpha)(s)\Big) = \partial^*_{i_m}(x) \sum_{j=1}^{n} (X_{lj \circ \varphi_\alpha})(s) \partial_j(x) \tag{56.98a}$$

$$= \sum_j (X_{lj} \circ \varphi_\alpha)(s) \delta_{j,i_m} \tag{56.98b}$$

$$= (X_{l,i_m} \circ \varphi_\alpha)(s). \tag{56.98c}$$

The last line is a smooth function of s since X_{ij} are smooth. The proposition 49.22 about sum and product of smooth functions makes $\omega(X_1, \dots, X_k)$ smooth. □

LEMooMEZCooGXpOOa

Lemma 56.27.
Let $v \in \mathbb{R}^n$. The maps of 56.19 satisfy

ITEMooZSFDooAuUnPd

(1)

$$\tau_s(e_i^*)\big(\tau_s(v)\big) = e_i^*(v) = v_i. \tag{56.99}$$

ITEMooSVEFooBKhHnc

(2) For every miltiindex $I = (i_1, \dots, i_k)$ and $X_j \in T_{\varphi_\alpha(s)}M$, we have

$$\tau_{\alpha,s}(e^*_{\wedge I})(X_1, \dots, X_k) = e^*_{\wedge I}\big(\tau^{-1}_{\alpha,s}(X_1), \dots, \tau^{-1}_{\alpha,s}(X_k)\big). \tag{56.100}$$

ITEMooQUZCooGcJrLe

(3) For every $\omega' \in \bigwedge^k (\mathbb{R}^n)^*$ and $X_i \in T_{\varphi_\alpha(s)M}$ we have

$$\tau_{\alpha,s}(\omega')(X_1, \dots, X_k) = \omega'\big(\tau^{-1}_{\alpha,s}(X_1), \dots, \tau^{-1}_{\alpha,s}(X_k)\big). \tag{56.101}$$

9. We write the dependence of ∂_i^* in $x = \varphi_\alpha(s)$.

Proof. Each point is a generalisation of the previous one.

(1) **For (1)** We have $e_i^* \in \bigwedge^k \mathbb{R}^n$ with $k=1$; we can apply τ_s to it. We have

$$\tau_s(e_i^*)\big(\tau_s(v)\big) = \partial_i^*\big(\sum_j v_j\partial_j\big) = \sum_j v_j\delta_{ij} = v_i = e_i^*(v). \tag{56.102}$$

(2) **For (2)** Let $X_j \in T_{\varphi_\alpha(s)}M$. Thanks to proposition 49.56, we write $X_j = \sum_l X_{jl}\partial_l$. Applying ∂_m^* we have

$$\partial_m^*(X_j) = \sum_l X_{jl}\delta_{ml} = X_{jm}. \tag{56.103}$$

EQooBGWWooXQoAdc

Now we can compute. On the one hand,

$$\begin{aligned}
\tau_{\alpha,s}(e^*_{\wedge I})(X_1,\dots,X_k) &= \partial^*_{\wedge I}(X_1,\dots,X_k) && (56.104a)\\
&= (\partial^*_{i_1}\wedge\dots\wedge\partial^*_{i_k})(X_1,\dots,X_k) && (56.104b)\\
&= k!\,\mathrm{Alt}(\partial^*_{i_1}\otimes\dots\otimes\partial^*_{i_k})(X_1,\dots,X_k) && \text{lem. 10.213.} \quad (56.104c)\\
&= \sum_{\pi\in S_k}(-1)^\pi(\partial^*_{i_1}\otimes\dots\partial^*_{i_k})(X_{\pi(1)},\dots,X_{\pi(k)}) && \text{def. 10.206} \quad (56.104d)\\
&= \sum_{\pi\in S_k}(-1)^\pi\partial^*_{i_1}(X_{\pi(1)})\dots\partial_{i_k}(X_{\pi(k)}) && (56.104e)\\
&= \sum_{\pi\in S_k}(-1)^\pi X_{i_1\pi(1)}\dots X_{i_k\pi(k)} && \text{by (56.103).} \quad (56.104f)
\end{aligned}$$

On the other hand we can make the same calculation [10] for $e^*_{\wedge I}\big(\tau^{-1}_{\alpha,s}(X_1),\dots,\tau^{-1}_{\alpha,s}(X_k)\big)$ using the fact that

$$e_i^*\big(\tau_{\alpha,s}(X_j)\big) = e_i^*\big(\sum_l X_{jl}e_l\big) = X_{ji}. \tag{56.105}$$

(3) **For (3)** Using the linearity of $\tau_{\alpha,s}$ and point (2).

□

LEMooOLSHooZNRoZs

Lemma 56.28.
Let $\omega \in \Omega^k(M)$. There exists a smooth $\omega' : U_\alpha \to \bigwedge^k(\mathbb{R}^n)^*$ such that

$$\omega_{\varphi_\alpha(s)} = \Big(\varphi_\alpha(s), \tau_{\alpha,s}\big(\omega'(s)\big)\Big). \tag{56.106}$$

Proof. Saying that the map $\omega : M \to M \times \bigwedge^k T^*M$ is smooth means that the map $f = \sigma_\beta^{-1}\circ\omega\circ\varphi_\alpha$ is smooth as map from U_α to U_β. In particular with $\beta=\alpha$ we are looking at $f(s) = (\sigma_\alpha^{-1}\circ\omega\circ\varphi_\alpha)(s)$:

$$\begin{aligned}
f(s) &= \sigma_\alpha^{-1}\big((\omega\circ\varphi_\alpha)(s)\big) && (56.107a)\\
&= \sigma_\alpha^{-1}\big(\varphi_\alpha(s),\xi(s)\big) && (56.107b)\\
&= \big(s,\omega'(s)\big) && (56.107c)
\end{aligned}$$

for some $\xi(s) \in \bigwedge^k(T^*_{\varphi_\alpha(s)})$ and $\omega'(s) \in \bigwedge^k(\mathbb{R}^n)^*$. Since f is smooth, the map

$$\omega' : U_\alpha \to \bigwedge^k(\mathbb{R}^n)^* \tag{56.108}$$

is smooth. From the definition of f,

$$(\omega\circ\varphi_\alpha)(s) = (\sigma_\alpha\circ f)(s) = \sigma_\alpha\big(s,\omega'(s)\big) = \Big(\varphi_\alpha(s),\tau_{\alpha,s}\big(\omega'(s)\big)\Big). \tag{56.109}$$

We proved that

$$(\omega\circ\varphi_\alpha)(s) = \Big(\varphi_\alpha(s),\tau_{\alpha,s}\big(\omega'(s)\big)\Big). \tag{56.110}$$

EQooMAINooBKsaex

□

10. I did not. Check and send me a message saying if it works or not.

PROPooFWEQooIqTFSa

Proposition 56.29.
If $\omega\colon U_\alpha \to \bigwedge^k(\mathbb{R}^n)^*$ is smooth, then the map

$$\begin{aligned} f\colon U_\alpha &\to \bigwedge^k \big(T^*_{\varphi_\alpha(s)}M\big) \\ s &\mapsto \tau_{\alpha,s}\big(\omega(s)\big) \end{aligned} \tag{56.111}$$

is smooth.

Proof. Proposition 10.219 says that there exist smooth functions $\omega_I\colon U_\alpha \to \mathbb{R}$ such that

$$\omega(s) = \sum_{I\in C_k} \omega_I(s) e^*_{\lambda I}. \tag{56.112}$$

Thus

$$\tau_{\alpha,s}\big(\omega(s)\big) = \sum_{I\in C_k} \omega_I(s)\partial^*_{\wedge I}. \tag{56.113}$$

The functions ω_I are smooth as well as $\partial^*_{\wedge I}$ (lemma 56.18). □

The following definition is in the same spirit of definition 10.227.

DEFooZMQNooRNhWXk

Definition 56.30.
Let $\omega' \in \bigwedge^k(\mathbb{R}^n)^*$, $p = \varphi_\alpha(s) \in M$ and $X \in T_pM$. We define $i_X(\omega)$ by

$$i_X(\omega)(X_1, \dots, X_{k-1}) = \omega(X, X_1, \dots, X_{k-1}). \tag{56.114}$$

Let X be a smooth vector field and ω a smooth k-differential form. We defin $i_X(\omega)$ for each $p \in M$ by

$$i_X(\omega)_p = i_{X_p}(\omega_p). \tag{56.115}$$

PROPooLCFAooLrbqeh

Proposition 56.31.
Let $\omega' \in \bigwedge^k(\mathbb{R}^n)^*$, $v \in \mathbb{R}^n$ and $p = \varphi_\alpha(s) \in M$. We have

$$\tau_{\alpha,s}\big(i_v(\omega')\big) = i_{\tau_{\alpha,s}(v)}\big(\tau_{\alpha,s}(\omega')\big). \tag{56.116}$$

Proof. Using lemma 56.27(3) twice (once in each direction) we have

$$\begin{aligned} i_v(\omega')(X_1, \dots, X_{k-1}) &= i_v(\omega')\big(\tau^{-1}_{\alpha,s}(X_1), \dots, \tau^{-1}_{\alpha,s}(X_{k_1})\big) && \text{(56.117a)} \\ &= \omega'\big(v, \tau^{-1}_{\alpha,s}(X_1), \dots, \tau^{-1}_{\alpha,s}(X_{k-1})\big) && \text{(56.117b)} \\ &= \omega'\big(\tau^{-1}_{\alpha,s}\tau_{\alpha,s}(v), \tau^{-1}_{\alpha,s}(X_1), \dots, \tau^{-1}_{\alpha,s}(X_{k-1})\big) && \text{(56.117c)} \\ &= (\tau_{\alpha,s}\omega')\big(\tau_{\alpha,s}(v), X_1, \dots, X_{k-1}\big) && \text{(56.117d)} \\ &= i_{\tau_{\alpha,s}(v)}\big(\tau_{\alpha,s}(\omega')\big)(X_1, \dots, X_{k-1}). && \text{(56.117e)} \end{aligned}$$

□

LEMooEKYDooKkgZJB

Lemma 56.32 ([1]).
Let $v\colon U_\alpha \to \mathbb{R}^n$ and $\omega\colon U_\alpha \to \bigwedge^k(\mathbb{R}^n)^*$ be smooth. Then the map

$$\begin{aligned} r\colon U_\alpha &\to \bigwedge^{k-1}(\mathbb{R}^n)^* \\ s &\mapsto i_{v(s)}\omega(s) \end{aligned} \tag{56.118}$$

is smooth.

Proof. Notice that $\bigwedge^{k-1}(\mathbb{R}^n)^*$ is a part of $\text{End}\left((\mathbb{R}^n)^{k-1},\mathbb{R}\right)$, thus we can consider

$$r\colon U_\alpha \to \text{End}\left((\mathbb{R}^n)^{k-1},\mathbb{R}\right). \tag{56.119}$$

If $\omega(s)=\sum_{I=(i_1,\dots,i_k)}(s)e^*_{\otimes I}$, proposition 10.229 gives

$$r(s)=\sum_{J=(j_1,\dots,j_{k-1})}\omega'_J(s)e^*_{\otimes J} \tag{56.120}$$

with

$$\omega'_J=\sum_i \omega_{iJ}(s)v(s)_i. \tag{56.121}$$

Thus we write

$$r(s)=\sum_J\sum_i \omega_{iJ}(s)v(s)_i e^*_{\otimes J}. \tag{56.122}$$

EQooQKQHooGXUzHY

The map $s\mapsto\omega_{iJ(s)}$ is smooth as well as $s\mapsto v(s)_i$. The whole expression (56.122) is smooth. □

PROPooTBFLooVCqycE

Proposition-Definition 56.33 ([1])**.**
Let X be a smooth vector field and ω a smooth k-differential form. Then the map

$$i_X(\omega)\colon M\to\bigwedge^k(T^*M) \tag{56.123}$$

is smooth and is then a k-differential form.

Proof. Let $p=\varphi_\alpha(s)\in M$. We have, for some $\omega'(s)\in\bigwedge^k(\mathbb{R}^n)^*$,

$$\begin{aligned}\big(i_X(\omega)\big)_p &= i_{X_p}(\omega_p) && \text{(56.124a)}\\ &= i_{X_p}\tau_{\alpha,s}\big(\omega'(s)\big) && \text{(56.124b)}\\ &= \tau_{\alpha,s}\Big(i_{\tau^{-1}_{\alpha,s}}(X_p)\omega'\Big) \qquad \text{prop. 56.31.} && \text{(56.124c)}\end{aligned}$$

Since $\tau_{\alpha,s}$ is smooth by proposition 56.29, it remains to prove that the map

$$\begin{aligned}r\colon U_\alpha &\to \bigwedge^{k-1}(\mathbb{R}^n)^*\\ s&\mapsto i_{\tau^{-1}_{\alpha,s}(X_{\varphi_\alpha(s)})}\omega'(s)\end{aligned} \tag{56.125}$$

is smooth. Since X is a smooth vector field, proposition 49.85 says that we can write

$$X_{\varphi_\alpha(s)}=\sum_{i=1}^n X_i\big(\varphi_\alpha(s)\big)\partial_i \tag{56.126}$$

for some smooth functions $X_i\colon M\to\mathbb{R}$. With these notations we have

$$\tau^{-1}_{\alpha,s}(X_{\varphi_\alpha(s)})=\sum_i (X_i\circ\varphi_{\alpha,s})(s)e_i. \tag{56.127}$$

The latter is smooth with respect to s. □

56.3.1 Interior product

In the algebraic setting (without regularity and manifolds), the interior product we introduced in definition 10.227, and we already had the Leibniz rule in proposition 10.230. Recall: the we denote by $\Omega^k(M)$ the space of k-differential forms on the smooth manifold M. These are the sections of $\bigwedge^k(T^*M)$, see definition 56.21.

PROPooIQIUooTDNJdB

Proposition 56.34 ([832]).
Let X be a smooth vector field. Let $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M)$ be k and l differential forms. Then we have

$$i_X(\omega \wedge \eta) = i_X(\omega) \wedge \eta + (-1)^k \omega \wedge i_X(\eta). \tag{56.128}$$

Proof. For each $p \in M$ we have

$$\Big(i_X(\omega \wedge \eta)\Big)_p = i_{X_p}(\omega_p \wedge \eta_p) \tag{56.129a}$$
$$= i_{X_p}(\omega_p) \wedge \eta_p + (-1)^k \omega_p \wedge (i_{X_p}\eta) \qquad \text{prop. 10.230} \tag{56.129b}$$
$$= \Big(i_X(\omega) \wedge \eta\Big)_p + (-1)^k \Big(\omega \wedge i_X\eta\Big)_p. \tag{56.129c}$$

□

56.3.2 Exterior derivative

PROPooTXFRooMtaVrU

Proposition 56.35 (exterior derivative[1, 911]).
There exists a unique set of maps $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$ $(k \in \mathbb{N})$ such that

ITEMooPQYCooHMjgIR
(1) d is $\mathbb{R}$-linear : $d(\lambda\omega) = \lambda d\omega$ and $d(\omega + \eta) = d\omega + d\eta$ for every $\lambda \in \mathbb{R}$ and $\omega, \eta \in \Omega^k(M)$.

ITEMooRKJRooWwHUiS
(2) If $f \in \Omega^0(M) = C^\infty(M)$, then df is the usual differential of f.

ITEMooBBFLooWpZJuT
(3) Leibniz rule : if $\omega \in \Omega^k(M)$, we have $d(\omega \wedge \eta) = (d\omega) \wedge \eta = (-1)^k \omega \wedge d\eta$

ITEMooJXMFooHFTFYi
(4) $d(d\omega) = 0$.

Proof. First unicity, then existence.

Unicity

We initiate by proving the unicity by showing that d has a fixed form in a chart. Let (U, φ) be a chart. From 56.25 a generic differential form reads $\omega_x = \sum_I \omega_I(x) \partial^*_{\wedge I}$.

(1) $\underline{d(\partial^*_{\wedge I}) = 0}$ Let $I = (i_1, \dots, i_k)$ and compute $d(\partial^*_{\wedge I})$. We denote $\omega = \partial^*_{i_1} \wedge \dots \wedge \partial^*_{i_{k-1}}$:

$$d(\partial^*_{\wedge I}) = d\big(\omega \wedge \partial^*_{i_k}\big) \tag{56.130a}$$
$$= (d\omega) \wedge \partial^*_{i_k} + (-1)^{k-1} \omega \wedge d(\partial^*_{i_k}) \qquad \text{by item (3).} \tag{56.130b}$$
$$= d(\partial^*_{i_1} \wedge \dots \partial^*_{i_{k-1}}) \wedge \partial^*_{i_k} \qquad \text{see bellow.} \tag{56.130c}$$

SUBEQooOWTHooIZcNWw

Justification for (56.130c). From lemma 56.16, $\partial^*_{i_k} = df_{i_k}$, so that item (4) gives $d(\partial^*_{i_{k+1}}) = d(df_i) = 0$.

By recursion we have $d(\partial^*_{\wedge I}) = 0$.

(2) **Computation of $d(f\partial^*_{\wedge I})$** From lemma 56.24 we use the wedge product $f\partial^*_{\wedge I} = f \wedge \partial^*_{\wedge I}$ and we can use the Leibniz formula of item (3) :

$$d(f\partial^*_{\wedge I}) = d(f \wedge \partial^*_{\wedge I}) = df \wedge \partial^*_{\wedge I} + f \wedge \underbrace{d(\partial^*_{\wedge I})}_{=0} = df \wedge \partial^*_{\wedge I}. \tag{56.131}$$

(3) **The general case** A general differential k-form ω is given by proposition 56.25 which is a real linear combination of elements of the form $f\partial^*_{\wedge I}$.

$$d\omega = d\Big(\sum_I a_I \partial^*_{\wedge I}\Big) = \sum_I da_I \wedge \partial^*_{\wedge I}. \tag{56.132}$$

EQooMGAVooCndFmv

(4) **Unicity** If a map d satisfying all the conditions exists, it has to be of the form

$$d\Big(\sum_I a_I \partial^*_{\wedge I}\Big) = \sum_I da_I \wedge \partial^*_{\wedge I} \tag{56.133}$$

on each coordinate patch. This proves the unicity.

Local existence

For each chart $(U_\alpha, \varphi_\alpha)$ we define the operators

$$\begin{aligned} d_\alpha \colon \Omega^k\big(\varphi_\alpha(U_\alpha)\big) &\to \Omega^{k+1}\big(\varphi_\alpha(U_\alpha)\big) \\ \sum_I a_I \partial^*_{\wedge I} &\mapsto \sum_I da_I \wedge \partial^*_{\wedge I}, \end{aligned} \tag{56.134}$$

and we prove that they satisfy the conditions.

(1) **Property (1)** We have

$$d(\lambda\omega) = \big(\sum_I (\lambda a_I)\partial^*_{\wedge I}\big) = \sum_I d(\lambda a_I) \wedge \partial^*_{\wedge I} = \lambda \sum_I da_I \wedge \partial^*_{\wedge I}. \tag{56.135}$$

Same kind of computations for $d(\omega + \eta) = d\omega + d\eta$.

(2) **Property (2)** If ω is a 0-form, the expression (56.132) has no sum over I and reduces to $\omega = f$.

(3) **Property (3)** Let $\omega = \sum_{I \in C_k} a_I \partial^*_{\wedge I}$ and $\eta = \sum_{J \in C_l} b_J \partial^*_{\wedge J}$. Since the wedge product is linear we can group the sums:

$$\omega \wedge \eta = \sum_I \sum_J a_I b_J \partial^*_{\wedge I} \wedge \partial^*_{\wedge J}. \tag{56.136}$$

Using the product rule for usual differential, we have

$$\begin{aligned} d_\alpha(\omega \wedge \eta) &= \sum_{I,J} d(a_I b_J) \wedge (\partial^*_{\wedge I} \wedge \partial^*_{\wedge J}) && \text{(56.137a)} \\ &= \sum_{I,J} b_J (da_I) \wedge (\partial^*_{\wedge I} \wedge \partial^*_{\wedge J}) + \sum_{I,J} (a_I db_J) \wedge (\partial^*_{\wedge I} \wedge \partial^*_{\wedge J}) && \text{(56.137b)} \\ &= \sum_{I,J} (da_I \wedge \partial^*_{\wedge I}) \wedge (b_J \partial^*_{\wedge J}) + \sum_{I,J} a_I (-1)^k \partial^*_{\wedge I} \wedge db_J \wedge \partial^*_{\wedge J} \quad \text{cf. bellow} && \text{(56.137c)} \\ &= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. && \text{(56.137d)} \end{aligned}$$

SUBEQooFXQAgoCziZnU

For (56.137c). In the first sum, the function b_J commutes with the operation $\wedge$. In the second sum, the differential form db_J commutes with the wedge products giving a minus sign. This explains the $(-1)^k$.

(4) $d(df) = 0$ Let $f \colon M \to \mathbb{R}$ be a smooth function. Using lemma 56.17, we compute $d_\alpha(df)$:

$$d_\alpha(df) = \sum_i \sum_j (\partial_j \partial_i f) \partial^*_j \wedge \partial^*_i. \tag{56.138}$$

Since $\partial_i \partial_i f$ is symmetric with respect to i, j and $\partial^*_j \wedge \partial^*_i$ is anti-symmetric with respect to i, j, the sum is zero.

(5) **Property (4)** Let $\omega = \sum_I a_I \partial^*_{\wedge I}$. First we have

$$d_\alpha \omega = \sum_I da_I \wedge \partial^*_{\wedge I}. \tag{56.139}$$

Then

$$d_\alpha(d_\alpha \omega) = \sum_I \underbrace{d(da_I)}_{=0} \wedge \partial^*_{\wedge I} - \sum_I da_I \wedge \underbrace{d(\partial^*_{\wedge I})}_{=0}. \tag{56.140}$$

Global existence

By the unicity part, we have $d_\alpha = d_\beta$ on the part $\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$. Thus the local formula

$$d\big(\sum_I a_I \partial^*_{\wedge I}\big) = \sum_I da_I \wedge \partial^*_{\wedge I} \tag{56.141}$$

globally defines d. □

PROPooUOQRooAdvYqx

Proposition 56.36 ([1, 912, 913, 914]).
Let ω be a k-differential form and $X_1, \dots, X_{k+1}$ be vector fields. We have

$$\begin{aligned}(d\omega)(X_1,\dots,X_{k+1}) = &\sum_i (-1)^{i+1} X_i\big(\omega(X_1,\dots,\hat{X}_i,\dots X_{k+1})\big)\\ &+\sum_{i<j}(-1)^{i+j}\omega\big([X_i,X_j],X_1,\dots\hat{X}_i,\dots,\hat{X}_j,\dots,X_{k+1}\big)\end{aligned} \tag{56.142}$$

where $d\omega$ is the exterior derivative defined by proposition 56.35.

Proof. We defined $s\colon \Omega^k(M) \to \Omega^{k+1}(M)$ by

$$\begin{aligned}(s\omega)(X_1,\dots,X_{k+1}) = &\sum_i (-1)^{i+1} X_i\big(\omega(X_1,\dots,\hat{X}_i,\dots X_{k+1})\big)\\ &+\sum_{i<j}(-1)^{i+j}\omega\big([X_i,X_j],X_1,\dots\hat{X}_i,\dots,\hat{X}_j,\dots,X_{k+1}\big)\end{aligned} \tag{56.143}$$

EQooZPNOooAXitUC

and we will show that s has the same local form as d:

$$s(f\partial^*_{\wedge I}) = df \wedge \partial^*_{\wedge I}. \tag{56.144}$$

Then, as we seen in proposition 56.35, it makes s satisfy the defining properties of the exterior derivative and the unicity part will conclude that $s = d$.

———— $s(\omega)$ is $C^\infty(M)$-linear ————

In this part, we prove that $s(\omega)(fX_1, X_2, \dots, X_{k+1}) = fs(\omega)(X_1, \dots, X_{k+1})$ for every smooth function $f\colon M \to \mathbb{R}$.

(1) **Splitting the sums**
We split the sums in (56.143). We separate the terms $i = 1$ in the sums in equation (56.143):

$$\begin{aligned}s(\omega)(fX_1,X_2,\dots,X_{k+1}) = &fX_1\big(\omega(X_2,\dots,X_{k+1})\big)\\ &+\sum_{i>1}(-1)^{i+1}X_i\big(\omega(fX_1,X_2,\dots,\hat{X}_i,\dots,x_{k+1})\big)\\ &+\sum_{j>1}(-1)^{1+j}\omega\big([fX_1,X_j],X_2,\dots,\hat{X}_j,\dots,X_{k+1}\big)\\ &+\sum_{1<i<j}(-1)^{i+j}\omega\big([X_i,X_j],fX_1,\dots,\hat{X}_j,\dots,X_{k+1}\big).\end{aligned} \tag{56.145}$$

Since ω is $C^\infty(M)$-linear, we can move f outside ω when it is a simple coefficient:

$$\begin{aligned}s(\omega)(fX_1,X_2,\dots,X_{k+1}) = &fX_1\big(\omega(X_2,\dots,X_{k+1})\big)\\ &+\sum_{i>1}(-1)^{i+1}X_i\big(f\omega(X_1,X_2,\dots,\hat{X}_i,\dots,x_{k+1})\big)\\ &+\sum_{j>1}(-1)^{1+j}\omega\big([fX_1,X_j],X_2,\dots,\hat{X}_j,\dots,X_{k+1}\big)\\ &+\sum_{1<i<j}(-1)^{i+j}f\omega\big([X_i,X_j],X_1,\dots,\hat{X}_j,\dots,X_{k+1}\big).\end{aligned} \tag{56.146}$$

(2) **Apply X_i** Using the product rule,

$$\begin{aligned}X_i\big(f\omega(X_1,\dots,\hat{X}_i,\dots,X_{k+1})\big) = &(X_if)\omega(X_1,\dots,\hat{X}_i,\dots,X_{k+1})\\ &+fX_i\big(\omega(X_1,\dots,\hat{X}_i,\dots,X_{k+1})\big)\end{aligned} \tag{56.147}$$

(3) **The commutator** If v is a vector we have

$$[fX_1, X_j](v) = (fX_1)(X_jv) - X_j\big(X_j(X_1v)\big) \tag{56.148a}$$
$$= (fX_1X_j)(v) - X_j(f)X_1(v) - fX_jX_1v \tag{56.148b}$$
$$= f[X_1, X_j](v) - X_j(f)X_1(v). \tag{56.148c}$$

So we have $[fX_1, X_k] = f[X_1, X_j] - X_j(f)X_1$. Beware: the expression $f(X_1v)$ means the function $x \mapsto f(x)(X_1)_xv$, not the function f applied to X_1v (which makes no sense).
Notice that $X_j(f)$ is a function and can get out of ω.

(4) **Last computation** Using the small computations we've just made,

$$s(\omega)(fX_1, \dots, X_{k+1}) = fX_1\big(\omega(X_2, \dots, X_{k+1})\big) \tag{56.149a}$$
$$+ \sum_{i>1}(-1)^{i+1}(X_if)\omega(X_1, \dots, \hat{X}_i, \dots, X_{k+1}) \tag{56.149b}$$
$$+ \sum_{i>1}(-1)^{i+1}fX_1\big(\omega(X_1, \dots, \hat{X}_i, \dots, X_{k+1})\big) \tag{56.149c}$$
$$+ \sum_{j>1}(-1)^{1+j}f\omega\big([X_1, X_j], X_2, \dots, \hat{X}_j, \dots, X_{k+1}\big) \tag{56.149d}$$
$$- \sum_{j>1}(-1)^{1+j}X_j(f)\omega(X_1, \dots, \hat{X}_j, \dots, X_{k+1}) \tag{56.149e}$$
$$+ \sum_{1<i<j}(-1)^{i+j}f\omega\big([X_i, X_j], X_1, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{k+1}\big). \tag{56.149f}$$

The terms (56.149c) and (56.149e) simplify. The other can be grouped to get

$$s(\omega)(fX_1, \dots, X_{k+1}) = \sum_i(-1)^{i+1}fX_i\big(\omega(X_1, \dots, \hat{X}_i, \dots, X_{k+1})\big) \tag{56.150a}$$
$$+ \sum_{i<j}(-1)^{i+j}f\omega\big([X_i, X_j], X_1, \dots, \hat{X}_j, \dots, \hat{X}_j, \dots, X_{k+1}\big)$$
$$= fs(\omega)(X_1, \dots, X_{k+1}). \tag{56.150b}$$

Do you believe that one can make the same work with X_l instead of X_1 ? Okay we are done for the first part: $s(\omega)$ is $C^\infty(M)$-linear.

Local expression

Several steps.

(1) **Setting**
Since $s(\omega)$ is $C^\infty(M)$-linear, it is sufficient to compute its local expression when acting on the coordinate vectors $(\partial_{i_1}, \dots, \partial_{i_{k+1}})$. So consider a multiindex $J = (j_1, \dots, j_{k+1})$, let [11]

$$\omega = \sum_{I \in C_k} a_I(\partial^*_{\wedge I}) \tag{56.151}$$

and let's compute $s(\omega)(\partial_J)$.

(2) **Commutators**
Since $[\partial_i, \partial_i] = 0$ for every i and j, the term containing the commutator in $s(\omega)$ vanishes when applied to ∂_J. We are left with

$$s(\omega)(\partial_J) = \sum_l(-1)^{l+1}\sum_I \partial_{j_l}\big(a_I\partial^*_{\wedge I}(\partial_{j_1}, \dots, \hat{\partial}_{j_l}, \dots, \partial_{j_{k+1}})\big). \tag{56.152}$$

Notice that the numbers $i_1 < i_2 < \dots < i_k$ because the sum is over $I \in C_k$, see proposition 56.25.

11. We use compact notations like $\partial_I = (\partial_{i_1}, \dots, \partial_{i_k})$ and $\partial_{\wedge I} = \partial_{i_1} \wedge \dots \wedge \partial_{i_k}$.

(3) **Intermediate result** If I is a multiindex, if η is a 1-form and if k is not part of I, let us prove that

$$(\eta \otimes \partial^*_{\wedge I})(\partial_k, \partial_I) = (\eta \wedge \partial^*_{\wedge I})(\partial_k, \partial_I). \tag{56.153}$$

The wedge product $\eta \wedge \partial^*_{\wedge I}$ contains terms with $\eta \otimes \wedge^*_{\wedge I}$ and with $\partial^*_{\wedge I} \wedge \eta$. The second term is zero because ∂_k is part of the arguments of $\partial^*_{\wedge I}$ while k is not part of I.

(4) **If J is sorted** Let us suppose that $j_1 < \ldots < j_{k+1}$. We have:

$$s(\omega)(\partial_J) = \sum_l (-1)^{l+1} \sum_I \partial_{j_l}\big(a_I \partial^*_{\wedge I}(\partial_{J_{\hat{l}}})\big) \tag{56.154}$$

where $J_{\hat{l}} = (j_1, \ldots, \hat{j}_l, \ldots, j_{k+1})$. The only terms in the sum over I which do not vanish are the ones with

$$\{i_1, \ldots, i_k\} = \{j_1, \ldots, j_{k+1}\} \backslash \{j_l\}. \tag{56.155}$$

Since the multiindices I and J are supposed to be sorted, the set equality (56.155) says that only one term is non zero: the term $(i_1, \ldots, i_k) = (j_1, \ldots, \hat{j}_l, \ldots, j_{k+1}) = J_{\hat{l}}$. We have:

$$s(\omega)(\partial_J) = \sum_l (-1)^{l+1} \partial_{j_l}\big(a_{J_{\hat{l}}} \partial^*_{\wedge J_{\hat{l}}}(\partial_{J_{\hat{l}}})\big). \tag{56.156}$$

In this expression, the factor $\partial^*_{\wedge J_{\hat{l}}}(\partial_{J_{\hat{l}}})$ is a constant. The derivative ∂_{i_l} does not act on it. We continue the computation:

$$\begin{aligned}
s(\omega)(\partial_J) &= \sum_l (-1)^{l+1} \partial_{j_l}\big(a_{J_{\hat{l}}} \partial^*_{\wedge J_{\hat{l}}}(\partial_{J_{\hat{l}}})\big) && \text{(56.157a)}\\
&= \sum_l (-1)^{l+1} \underbrace{\partial_{j_l}(a_{J_{\hat{l}}})}_{=(da_{J_{\hat{l}}})(\partial_{j_l})} \partial^*_{\wedge J_{\hat{l}}}(\partial_{J_{\hat{l}}}) && \text{(56.157b)}\\
&= \sum_l (-1)^{l+1} (da_{J_{\hat{l}}})(\partial_{j_l})(\partial^*_{j_1} \wedge \ldots \hat{\partial}^*_{j_l} \ldots \wedge \partial^*_{j_{k+1}})(\partial_{j_1}, \ldots, \hat{\partial}_{j_l}, \ldots, \partial_{j_{k+1}}) && \text{(56.157c)}\\
&= \sum_l (-1)^{l+1} \big[(da_{I_{\hat{l}}} \otimes \partial_{\wedge J_{\hat{l}}})\big](\partial_{j_l}, \partial_{J_{\hat{l}}}) && \text{(56.157d)}
\end{aligned}$$

We use the intermediate result (56.153) :

$$\begin{aligned}
s(\omega)(\partial_J) &= \sum_l (-1)^{l+1} (da_{J_{\hat{l}}} \wedge \partial^*_{\wedge J_{\hat{l}}})(\partial_{j_l}, \partial_{J_{\hat{l}}}) && \text{(56.158a)}\\
&= \sum_l (da_{J_{\hat{l}}} \wedge \partial^*_{\wedge J_{\hat{l}}})(\partial_{j_1}, \ldots, \partial_{j_{k+1}}). && \text{(56.158b)}
\end{aligned}$$

The second equality correspond to move the argument ∂_{j_l} to the l^{e} position: it requires $l-1$ transpositions, each of them adding an extra minus sign.

(5) **Conclusion (J sorted)** We are still with J sorted. We have to compare (56.158b) with the known expression of $d\omega$. We have

$$\begin{aligned}
(d\omega)(\partial_J) &= \sum_I (da_I \wedge \partial^*_{\wedge I})(\partial_J) && \text{(56.159a)}\\
&= \sum_I \sum_l (\partial_l a_I)(\partial^*_l \wedge \partial^*_{\wedge I})(\partial_J) && \text{lem. 56.17} \quad \text{(56.159b)}\\
&= \sum_l \sum_I (\partial_l a_I)(\partial^*_l \wedge \partial^*_{\wedge I})(\partial_J) && \text{(56.159c)}\\
&= \sum_l (\partial_l a_{J_{\hat{l}}}) \partial_l \wedge \partial^*_{\wedge J_{\hat{l}}}(\partial_J) && \text{see bellow} \quad \text{(56.159d)}\\
&= \sum_l da_{J_{\hat{l}}} \wedge \partial^*_{\wedge J_{\hat{l}}}(\partial_J). && \text{(56.159e)}
\end{aligned}$$

Explanation for (56.159d). The only non vanishing term in the sum over I is the one with $\{l, i_1, \dots, i_k\} = \{j_1, \dots, j_{k+1}\}$, which is $I = J_{\hat{l}}$.
Now we see that the expression (56.158b) and (56.159e) are the same, so that

$$(d\omega)(\partial_J) = (s\omega)(\partial_J). \tag{56.160}$$

(6) <u>**J is not sorted**</u> We suppose now that J is not sorted. Let π be a permutation such that $\pi(J)$ is sorted. Using the fact that $s(\omega)$ and $d\omega$ are anti-symmetric, we have:

$$\begin{aligned} s(\omega)(\partial_J) &= (-1)^\pi s(\omega)(\partial_{\pi(J)}) && (56.161a)\\ &= (-1)^\pi (d\omega)(\partial_{\pi(J)}) && (56.161b)\\ &= (d\omega)(\partial_J). && (56.161c) \end{aligned}$$

□
LEMooJCXRooMQxauV

Lemma 56.37 ([1])**.**
Let M be a smooth manifold. We consider a chart $\varphi\colon U \to M$ and the maps

$$\begin{aligned} x_i\colon M &\to \mathbb{R} \\ p &\mapsto (\boldsymbol{proj}_i \circ \varphi^{-1})(p). \end{aligned} \tag{56.162}$$

We have $\partial_i^* = dx_i$.

Proof. We prove that $dx_i(\partial_j) = \delta_{ij}$:

$$\begin{aligned} dx_i(\partial_j) &= \frac{d}{dt}\big[x_i\big(\varphi(s + te_j)\big)\big]_{t=0} && (56.163a)\\ &= \frac{d}{dt}\big[\mathtt{proj}_i(s + te_j)\big]_{t=0} && (56.163b)\\ &= \frac{d}{dt}\big[s_i + t\delta_{ij}\big]_{t=0} && (56.163c)\\ &= \delta_{ij}. && (56.163d) \end{aligned}$$

□
LEMooQWTXooOkSZmV

Lemma 56.38.
Let $f_1, \dots, f_k$ be smooth functions we have

$$d(df_1 \wedge \dots df_k) = 0. \tag{56.164}$$

Proof. Proof by induction over k. If $k = 1$, we have $d(df_1) = 0$ by definition 56.35(4). If not we set $\omega = df_2 \wedge \dots \wedge df_k$ and we use the Leibniz rule 56.35(3) :

$$d(df_1 \wedge \dots df_k) = (ddf_1) \wedge \omega - df_1 \wedge d\omega. \tag{56.165}$$

In the first term, $ddf_1 = 0$ and $d\omega = 0$ by the induction hypothesis. □
LEMooAONJooDgEJjA

Lemma 56.39 ([1])**.**
Let a and $f_1, \dots, f_k$ be smooth functions. If $\omega = df_1 \wedge \dots \wedge df_k$, then

$$d(a\omega) = da \wedge \omega. \tag{56.166}$$

Proof. Point (3) of definition 56.35 says $d(a\omega) = da \wedge \omega + a \wedge d\omega$. Lemma 56.38 implies that the second term is zero. □
PROPooINQYooXZcfou

Proposition 56.40 ([1])**.**
Let M and N be smooth manifolds. Let $f\colon M \to N$ and $g_i\colon N \to \mathbb{R}$ be smooth functions. Let X_i be smooth vector fields on M. We have

$$(dg_1 \wedge \dots \wedge dg_k)(df_x X_1, \dots, df_x X_k) = d(g_1 \circ f)_x \wedge \dots \wedge d(g_k \circ f)_x(X_1, \dots, X_k). \tag{56.167}$$

Proof. First take a look at the formulas of corollary 10.215. Then we prove by induction over k. We set $\omega = dg_2 \wedge \ldots \wedge dg_k$ and we compute :

$$(dg_1 \wedge \omega)(X_1, \ldots, X_k) = \frac{1}{(k-1)!} \sum_{\pi \in S_k} (-1)^\pi (dg_1 \otimes \omega)(X_{\pi(1)}, \ldots, X_{\pi(k)}) \tag{56.168a}$$

$$= \frac{1}{(k-1)!} \sum_{\pi \in S_k} (-1)^\pi dg_1\big(df_x X_{\pi(1)}\big)\omega\big(df_x X_{\pi(2)}, \ldots, df_x X_{\pi(k)}\big) \tag{56.168b}$$

Here we transform

$$dg_1(df_x v) = d(g_1 \circ f)_x v \tag{56.169}$$

and, using the induction hypothesis,

$$\omega\big(df_x X_{\pi(2)}, \ldots, df_x X_{\pi(k)}\big) = d(g_2 \circ f)_x \wedge \ldots \wedge d(g_k \circ f)_x\big(X_{\pi(2)}, \ldots, X_{\pi(k)}\big), \tag{56.170}$$

and we make the computation backward from (10.558b). □

56.3.3 Pull back

PROPooLEYOooJfPtmB

Proposition 56.41 ([832])**.**
Let M and N be smooth manifolds. Let $s\colon M \to N$ be a smooth map. We consider charts $\varphi\colon U \to M$ and $\psi\colon V \to N$ where U and V are small enough to have

$$s\big(\varphi(U)\big) \subset \psi(V). \tag{56.171}$$

If $f \in C^\infty(N)$ we define We define s^*f by

$$(s^*f)(p) = (f \circ s)(p) \tag{56.172}$$

for each $p \in M$. If $\omega \in \Omega^k(N)$ we define

$$(s^*\omega)_p(X_1, \ldots, X_k) = \omega_{s(p)}\big(ds_p(X_1), \ldots, ds_p(X_k)\big) \tag{56.173}$$

EQooCQBJooQpJYUf

for each $p \in M$ and $X_i \in T_pM$.

With these definitions, we have

$$s^*\colon C^\infty(N) \to C^\infty(M) \tag{56.174}$$

and

$$s^*\colon \Omega^k(N) \to \Omega^k(M). \tag{56.175}$$

PROPooHUDYooLCAYZW

Proposition 56.42.
Let $s \in C^{\infty(M,N)}$, let $g \in C^\infty(N)$. We consider

1. *A field of basis $\{e_i\}$ of TM,*
2. *The dual field of basis $\{\alpha_i\}$ of T^*M; that means that for each x, the basis $\{(\alpha_i)_x\}$ of T_x^*M is the dual of the basis $\{(e_i)_x\}$ of T_xM,*
3. *A field of basis $\{e'_i\}$ of TN,*
4. *The dual basis $\{\beta_i\}$ of T^*N.*

We have[12]

$$s^*\big(g\beta_1 \wedge \ldots \wedge \beta_n\big) = (s^*g)\big(\det(ds)\big)\alpha_1 \wedge \ldots \wedge \alpha_n. \tag{56.176}$$

EQooWODAooMrqYCO

where $\det(ds)$ is the determinant of the matrix ds for the basis $\{e_i\}$ and $\{e'_i\}$.

12. As a side note, the fact that $\det(ds)$ is smooth is the proposition 49.102.

Proof. Let $x \in M$. We apply the left hand side of (56.176) to the field of basis $\{e_i\}$:

$$\begin{aligned}
\Big(s^*(g\beta_1 \wedge \ldots \wedge \beta_n)\Big)_x(e_1,\ldots,e_n) & \\
&= (g\beta_1 \wedge \ldots \wedge \beta_n)_{s(x)}\big(ds_x e_1,\ldots,ds_x e_n\big) && \text{def. (56.173)} && (56.177a)\\
&= (g\circ s)(x)(\beta_1 \wedge \ldots \beta_n)_{s(x)}(ds_x e_1,\ldots,ds_x e_n) && && (56.177b)\\
&= (s^*g)(x)\det\big(\beta_i(ds_x e_j)\big) && \text{cf.justif.} && (56.177c)\\
&= (s^*g)(x)\det\Big(\beta_i\big(\sum_k (ds_x)_{kj} e'_k\big)\Big) && && (56.177d)\\
&= (s^*g)(x)\det\Big(\sum_k (ds_x)_{kj}\underbrace{\beta_i(e'_k)}_{=\delta ik}\Big) && && (56.177e)\\
&= (s^*g)(x)\det(ds_x). && && (56.177f)
\end{aligned}$$

Justifications.

— For (56.177c). We used the proposition 10.214. Formally we should write $(\beta_i)_{s(x)}$ and $(e_j)_x$, but we drop the indices.

Now we apply the right hand side at $x \in M$ to $(e_1,\ldots,e_n)$. Using again proposition 10.214, we have

$$(\alpha_1 \wedge \ldots \wedge \alpha_n)(e_1,\ldots,e_n) = \det\big(\alpha_i(e_j)\big) = \det(\delta_{ij}) = 1, \tag{56.178}$$

thus

$$(s^*g)(x)\det(ds_x)(\alpha_1 \wedge \ldots \wedge \alpha_n)(e_1,\ldots,e_n) = (s^*g)(x)\det(ds_x). \tag{56.179}$$

By linearity, the equality still holds pointwise on a general element of $(T_xM)^n$. □

PROPooACWKooYLbIom

Proposition 56.43 ([832])**.**
Let M and N be smooth manifolds. We consider a smooth map $f\colon M \to N$ and $\omega \in \Omega^k(N)$. Then we have

$$f^*(d\omega) = d(f^*\omega), \tag{56.180}$$

EQooWUMCooLdNqmS

or $[f^*, d] = 0$.

Proof. Since d and f^* are linear, we check the equality (56.180) on $\omega = a\partial^*_{\wedge I}$ with $a \in C^\infty(N)$ and I being a k-multiindex. Since $\partial_i^* = dx_1$ (lemma 56.37), we have

$$\begin{aligned}
f^*(d\omega)(X_1,\ldots,X_{k+1}) &= (da \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k})\big(df(X_1),\ldots,df(X_{k+1})\big) && && (56.181a)\\
&= \Big(d(a\circ f) \wedge d(x_{i_1}\circ f) \wedge \ldots \wedge d(x_k \circ f)\Big)(X_1,\ldots,X_{k+1}) && \text{prop. 56.40} && (56.181b)
\end{aligned}$$

SUBEQooOCCWooUpOtqi

Now we search for an expression for $d(f^*\omega)$. For $x \in M$ we have

$$\begin{aligned}
(f^*\omega)_x(X_1,\ldots,X_k) &= \omega_{f(x)}\big(df_x X_1,\ldots,df_x X_k\big) && && (56.182a)\\
&= a\big(f(x)\big)\partial^*_{\wedge I}\big(df_x X_1,\ldots,df_x X_k\big) && && (56.182b)\\
&= (a\circ f)(x)\Big(d(x_{i_1}\circ f)_x \wedge \ldots \wedge d(x_{i_k}\circ f)_x\Big)(X_1,\ldots,X_k) && \text{s.b.} && (56.182c)
\end{aligned}$$

SUBEQooAQVZooNuBcVr

Justification for (56.182c): proposition 56.40.

Now, using the lemma 56.39 we have

$$d(f^*\omega) = d(a\circ f)\wedge d(x_{i_1}\circ f)\wedge\ldots\wedge d(x_{i_k}\circ f) + (a\circ f)\underbrace{d\Big(d(x_{i_1}\circ f)\wedge\ldots\wedge d(x_{i_k}\circ f)\Big)}_{=0} \tag{56.183}$$

This is the same as (56.181b).

□

56.4 Vector valued differential forms

SecVectValFiffFor

Let E be a vector bundle over M. A **E-valued p-form** is a section

$$e \in \Gamma\big(E \otimes \bigwedge^p T^*M\big).$$

We denote by $\Omega(M,E) = \Gamma\big(E \otimes \bigwedge^p T^*M\big)$ the set of E-valued differential forms. An element of $\Omega^1(M,E) = \Gamma\big(E \otimes \bigwedge T^*M\big)$ always reads $\sum_i s_i \otimes \omega_i$ for some sections s_i and usual differential forms ω_i.

A form of $\Omega^p(M,E)$ can be seen as a fiber morphism $\underbrace{TM \otimes \cdots \otimes TM}_{p \text{ times}} \to E$ by associating

$$s \otimes \omega(X_1, \cdots, X_p) = s(x)\omega(X_1, \cdots, X_p) \in E_x$$

to the element $(s \otimes \omega) \in \Omega^p(M,E)$. There exists a wedge product between vector-valued forms. If $e \in \Omega^p(M,E_1)$ and $f \in \Omega^q(M,E_2)$, then we define $e \wedge f \in \Omega^{p+q}(M, E_1 \otimes E_2)$ by

$$(e \wedge f)(v_1, \cdots, v_{p+q}) = \frac{1}{p!q!} \sum_{\pi \in S_{p+q}} (-1)^\pi e(v_{\pi(1)}, \cdots v_{\pi(p)}) \otimes f(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \in E_1 \otimes E_2. \tag{56.184}$$

EqDefwedgevecor

where $(-1)^\pi$ stands for the sign of the permutation π. For example when e, $f \in \Omega^1(M,E)$, we have

$$(e \wedge f)(X,Y) = e(X) \otimes f(Y) - e(Y) \otimes f(X) \in E \otimes E.$$

When M is a differentiable manifold, the **fundamental 1-form** is the element $\theta \in \Omega(M, TM)$ such that

$$\iota(X)\theta = X$$

for every $X \in \Gamma(TM)$.

56.4.1 A digression: $T_Y\mathcal{G}$ and $\mathcal{G}$

subsec:digress

We define two product: $G \times \mathcal{G} \to TG$ and $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$. If $g \in G$ and $X \in \mathcal{G}$, we put

$$gX = \frac{d}{dt} g e^{tX}\Big|_{t=0}, \tag{56.185a}$$

eq_gXdefa

and if X, $Y \in \mathcal{G}$,

$$XY = \frac{d}{dt}\frac{d}{du}\Big[e^{tX}e^{uY}\Big]_{\substack{u=0\\t=0}}. \tag{56.185b}$$

eq:yGyGb

We naturally define the product of a $\mathcal{G}$-valued 1-form A by an element $g \in G$ by $(gA)v = gA(v)$.

Note that gX does not belong to $\mathcal{G}$ but to T_gG. Fortunately, in the expressions which we will meet, there will always be a g^{-1} to save the situation.

Let us give some precisions about derivatives as (56.185b). We consider the expression

$$\frac{d}{du}\left(\frac{d}{dt}c_u(t)\Big|_{t=0}\right)_{u=0},$$

which will be more simply written as:

$$\frac{d}{du}\frac{d}{dt}\Big[c_u(t)\Big]_{\substack{t=0\\u=0}} \tag{56.186}$$

eq:2307e1

with $c_u(t) \in G$ for all u,t; $c_u(0) = e$ for all u and $c_0'(0) = Y \in \mathcal{G}$ where the prime stands for the derivative with respect of t. So $\frac{d}{dt}c_u(t)\big|_{t=0} \in \mathcal{G}$ for each u and (56.186) belongs to $T_Y\mathcal{G}$. But we know that $\mathcal{G}$ is a vector space, then $T_Y\mathcal{G} \simeq \mathcal{G}$, the isomorphism being given by the following idea: if $\{\partial_i\}$ is a basis of $\mathcal{G}$ and $\{e_i\}$ the corresponding basis of $T_Y\mathcal{G}$, we define the action of $A^ie_i \in T_Y\mathcal{G}$ on $f\colon G \to \mathbb{R}$ by $(A^ie_i)f := A^i\partial_i f$.

Lemma 56.44.
Seen as an equality in $\mathcal{G}$, for $f\colon G \to \mathbb{R}$ we have:

$$\frac{d}{du}\frac{d}{dt}\Big[c_u(t)\Big]_{\substack{t=0\\u=0}} f = \frac{d}{du}\frac{d}{dt}\Big[f(c_u(t))\Big]_{\substack{t=0\\u=0}}. \tag{56.187}$$

Proof. Let us consider $X_u = X_u^i\partial_i = c_u'(0)$ and $X_0 = Y$. We naturally have

$$X_u f = \frac{d}{dt} f(c_u(t))\Big|_{t=0}, \qquad \text{and} \qquad \frac{d}{du}X_u\Big|_{u=0} \in T_Y\mathcal{G}. \tag{56.188}$$

Now, we consider a function $h\colon \mathcal{G} \to \mathbb{R}$ and compute:

$$\frac{d}{du}\Big[X_u\Big]_{u=0} h = \frac{d}{du}\Big[h(X_u)\Big]_{u=0} = \frac{d}{du} h\Big(\frac{d}{dt}\Big[c_u(t)\Big]_{t=0}\Big)\Big|_{u=0}.$$

If $\{\partial_i\}$ is a basis of $\mathcal{G}$ and $\{e_i\}$, the corresponding one of $T_Y\mathcal{G}$, thus

$$\frac{d}{du}\Big[X_u\Big]_{u=0} h = \frac{\partial h}{\partial e_i}\Big|_Y \frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}}. \tag{56.189}$$

So, we can write

$$\frac{d}{du}\Big[X_u\Big]_{u=0} = \frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}} \frac{\partial}{\partial e_i}\Big|_Y,$$

and, as element of $\mathcal{G}$, we consider

$$\frac{d}{du}\Big[X_u\Big]_{u=0} = \frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}} \partial_i|_e.$$

Now, we can compute the action of $\frac{d}{du}X_u\big|_{u=0}$ on a function $f\colon G \to \mathbb{R}$ as

$$\begin{aligned}\frac{d}{du}\Big[X_u\Big]_{u=0} f &= \frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}} \frac{\partial f}{\partial x^i}\Big|_e \\ &= \frac{d}{du}\Big[\ \frac{\partial f}{\partial x^i}\Big|_e \frac{d}{dt}c_u^i(t)\Big|_{t=0}\Big]_{u=0} \\ &= \frac{d}{du}\Big[\ \frac{d}{dt} f(c_u(t))\Big|_{t=0}\Big]_{u=0}.\end{aligned} \tag{56.190}$$

¡¡ Avertissement/question à la lectrice !! 56.45
Je ne sais pas pourquoi tout d'un coup la dernière équation était commentée, et donc la phrase n'était pas finie.

□

Let us now see a great consequence of the definition (56.185b)

prop:XY_YX

Proposition 56.46.
The formula

$$XY - YX = [X, Y]. \tag{56.191}$$

links the formal product inside the Lie algebra and the Lie bracket.

Proof. From this, we can precise our definition of XY when X, $Y \in \mathcal{G}$. The product XY acts on $f\colon G \to \mathbb{R}$ by

$$(XY)f = \frac{d}{dt}\frac{d}{du}\Big[f(e^{tX}e^{uY})\Big]_{\substack{u=0\\t=0}}.$$

We can get a more geometric interpretation of this. We know how to build a left invariant vector field $\tilde{Y}$ from any $Y \in \mathcal{G}$: for each $g \in G$ we just have to define

$$\tilde{Y}_g(f) = \frac{d}{ds}\Big[f(gY(s))\Big]_{s=0}.$$

First remark: $\tilde{Y}_g$ is precisely the object that previously named "gY". In order to construct the basis blocks of the formula $XY - YX = [X, Y]$, we turn our attention to $\tilde{X}_e\tilde{Y}$. It is clear that $\tilde{Y}(f)$ is a function from G to $\mathbb{R}$, so we can apply $\tilde{X}_e$ on it. If X_t is a path which gives the vector $\tilde{X}_e$ (for example: $X_t = e^{tX}$), we have

$$\tilde{X}_e(\tilde{Y}(f)) = \frac{d}{dt}\Big[\tilde{Y}(f)_{X_t}\Big]_{t=0} = \frac{d}{du}\frac{d}{dt}\Big[f(X_tY(u))\Big]_{\substack{t=0\\u=0}} = \frac{d}{du}\frac{d}{dt}\Big[f(e^{tX}e^{uY})\Big]_{\substack{t=0\\u=0}}. \tag{56.192}$$

Thus we have: $XY = \tilde{X}_e\tilde{Y}$, but it is clear that $[\tilde{X}, \tilde{Y}]_e = \tilde{X}_e\tilde{Y} - \tilde{Y}_e\tilde{X}$. The claim reads now: $[\tilde{X}, \tilde{Y}]_e = [X, Y]$. We can actually take it as de *definition* of $[X, Y]$. It is done in [813]. The link with the definition in terms of successive derivations of $\mathbf{Ad}_g(x) = gxg^{-1}$ is not trivial but it can be done. □

Now, we can give a powerful definition of the wedge for two $\mathcal{G}$-valued 1-forms. If A, $B \in \Omega^1(M, \mathcal{G})$ and v, $w \in \mathfrak{X}(M)$, we define

$$(A \wedge B)(v, w) = A(v)B(w) - A(w)B(v). \tag{56.193}$$

For A^2, we find back the usual definition:

$$(A \wedge A)(v, w) = A(v)A(w) - A(w)A(v) = [A(v), A(w)].$$

One can see that for any section $\sigma_\alpha : \mathcal{U}_\alpha \to P$, we have

$$\sigma_\alpha^*(A \wedge A) = (\sigma_\alpha^* A) \wedge (\sigma_\alpha^* A). \tag{56.194}$$

eq:1907r2

Lemma 56.47.
If A and B are two $\mathcal{G}$-valued 1-forms, one can make "simplifications" as

$$(Ag) \wedge (g^{-1}B) = A \wedge B. \tag{56.195}$$

lem:simplif

Proof. We just have to prove that for A, $B \in \mathcal{G}$, $(Ag)(g^{-1}B) = AB$ with definitions (56.185a) and (56.185b). Remark that $Ag = \frac{d}{ds}\Big[e^{sA}g\Big]_{s=0}$, so

$$e^{tAg} = \exp(t\,\frac{d}{ds}e^{sA}g\Big|_{s=0}) = \exp(\frac{d}{ds}e^{stA}g\Big|_{s=0}) = e^{tA}g.$$

Therefore

$$(Ag)(g^{-1}B) = \frac{d}{dt}\frac{d}{du}\Big[e^{tAg}e^{ug^{-1}B}\Big]_{\substack{u=0\\t=0}} = \frac{d}{dt}\frac{d}{du}\Big[e^{tA}gg^{-1}e^{uB}\Big]_{\substack{u=0\\t=0}} = AB.$$

□

Lemma 56.48.

$$F_\beta = dA_\beta + A_\beta^2. \tag{56.196}$$

Proof. This is a direct consequence of (56.194) and $[\sigma_\beta^*, d] = 0$. □

56.5 Lie algebra valued differential forms

SecLiaAlgformval

An important particular case of vector valued forms is given by Lie algebra valued forms. That case appears for example in the connection theory over principal bundle [13]. If ω and η are elements of $\Omega^1(M,\mathcal{G})$ for some Lie algebra $\mathcal{G}$, we define

$$(\omega\wedge\eta)(X,Y)=\omega(X)\otimes\eta(Y)-\omega(Y)\otimes\eta(X).$$

Combining with the Lie bracket, we define

EqDefomegawedgebeta
$$[\omega\wedge\eta](X,Y):=[\omega(X),\eta(Y)]-[\omega(Y),\eta(X)]. \tag{56.197}$$

Using the proposition 56.46, we often implicitly transforms the tensor product into a product (56.185b) and put

EqAbuswesgeomom
$$(\omega\wedge\omega)(X,Y)=[\omega(X),\omega(Y)]. \tag{56.198}$$

Let us point out the fact that that kind of formula only holds for a "wedge square", but not for a general product $\omega\wedge\eta$. Remark that for $\omega\in\Omega^1(M,\mathcal{G})$ and $\beta\in\Omega^2(M,\mathcal{G})$, a simple computation of definition (56.184) yields

EqomwedgebetaXYZ
$$(\omega\wedge\beta)(X,Y,Z)=\omega(X)\otimes\beta(Y,Z)-\omega(Y)\otimes\beta(X,Z)+\omega(Z)\otimes\beta(X,Y), \tag{56.199}$$

so that, using the same trick as for equation (56.198), we find

$$(\omega\wedge\beta-\beta\wedge\omega)(X,Y,Z)=[\omega(X),\beta(Y,Z)]-[\omega(Y),\beta(X,Z)]+[\omega(Z),\beta(X,Y)].$$

But that expression is exactly what we find by exchanging the tensor product by Lie bracket in expression (56.199). So we define

EqDefCrochwedgedeux
$$[\omega\wedge\beta]=\omega\wedge\beta-\beta\wedge\omega \tag{56.200}$$

when $\omega\in\Omega^1(M,\mathcal{G})$ and $\beta\in\Omega^2(M,\mathcal{G})$. The reader should remark that this is what one would expect from generalisation of definition (56.197).

56.6 Principal bundle

Let M be a manifold and G, a Lie group whose unit is denoted by e. A G-**principal bundle** on M is a smooth manifold P, a smooth map $\pi\colon P\to M$ and a right action of G on P denoted by $\xi\cdot g$ with $g\in G$ and $\xi\in P$ such that

- $\pi(\xi\cdot g)=\pi(\xi)$,
- $\forall\xi\in\pi^{-1}(x)$, $\pi^{-1}(x)=\{\xi\cdot g \text{ tel que } g\in G\}\simeq G$,
- $\forall x\in M$, there exists a neighbourhood $\mathcal{U}_\alpha$ of x in M, a diffeomorphism $\phi_\alpha\colon\pi^{-1}(\mathcal{U}_\alpha)\to\mathcal{U}_\alpha\times G$ and a diffeomorphism $\phi_{\alpha x}\colon P\to G$ such that
  - $\phi_\alpha(\xi)=(x,\phi_{\alpha x}(\xi))$,
  - $\phi_{\alpha x}(\xi\cdot g)=\phi_{\alpha x}(\xi)\cdot g$.

The group G is often called the **structure group**. We suppose that the action is effective. We will sometimes use the notation $P(G,M)$ to precise that P is a principal bundle over M with structure group G.

lem:phixh

Lemma 56.49.
The map ϕ_α^{-1} fulfills

$$\phi\alpha^{-1}(x,h)\cdot g=\phi_\alpha^{-1}(x,hg).$$

13. So in Maxwell and other gauge field theories.

Proof. From the definition of a principal bundle, any $\xi \in P$ can be written under the form $\xi = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi))$ with ϕ_x satisfying $\phi_x(\xi \cdot h) = \phi_x(\xi)h$ for a certain function $\phi_x \colon P \to G$. We consider in particular $\xi = \phi_\alpha^{-1}(x,h) \cdot g$. Then $\xi \cdot g^{-1} = \phi_\alpha^{-1}(x,h)$. But $\xi \cdot g^{-1} = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi)g^{-1})$, then $h = \phi\alpha_x(\xi)g^{-1}$ and $\phi_{\alpha x}(\xi) = hg$. So we have

$$\xi = \phi_\alpha^{-1}(x,h) \cdot g = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi)) = \phi_\alpha^{-1}(x, hg).$$

□

Let

$$R = \{(x,y) \in P \times P \text{ tel que } x = y \cdot g \text{ for a certain } g \in G\}.$$

Proposition 56.50.
The function $u \colon R \to G$ defined by the condition

$$p \cdot u(p,q) = q.$$

is differentiable.

Proof. Let $\mathcal{U}$ be an open subset of M and $\sigma \colon \mathcal{U} \to P$, a section. We consider a differentiable map $\rho \colon \pi^{-1}(\mathcal{U}) \to G$ such that $\rho(\xi \cdot g) = \rho(\xi) \cdot g$ and $\rho(\sigma(x)) = e$. Such a map is given by

$$\rho(\xi) = \phi_x(\sigma(x))^{-1}\phi_x(\xi)$$

where $x = \pi(\xi)$. We naturally define $R_{\mathcal{U}} = R \cap (\pi^{-1}(\mathcal{U}) \times \pi^{-1}(\mathcal{U}))$ and we pick $(\xi, \eta) \in R_{\mathcal{U}}$. Let $s \in G$ be the one such that $\xi \cdot s = \eta$, so that $\rho(\xi) \cdot s = \rho(\eta)$. Then the restriction of u to $R_{\mathcal{U}}$ is given by $u(\xi,\eta) = \rho(\xi)^{-1}\rho(\eta)$ which makes $u|_{\mathcal{U}}$ differentiable. Since this reasoning can be made on every chart open $\mathcal{U}$, u is differentiable everywhere on P. □

The following is a corollary of Leibniz rule, proposition 49.123.

cor_PrincLeib

Corollary 56.51.
If P is a G-principal bundle and v, a are curves in P and G respectively, we can consider the curve $u(t) = v(t)a(t)$. We have:

$$\frac{d}{dt}u(t)\Big|_{t=0} = \frac{d}{dt}v(t)a(0)\Big|_{t=0} + \frac{d}{dt}v(0)a(t)\Big|_{t=0}.$$

The proof is direct. This result is often written as

$$\dot{u}_t = \dot{v}_t a_t + v_t \dot{a}_t. \tag{56.201}$$

A main application is

$$\frac{d}{dt}\Big[r \cdot h(t)\Big]_{t=0} = \frac{d}{dt}\Big[r \cdot e^{th'(0)}\Big]_{t=0}. \tag{56.202}$$

eq:rdotht

56.6.1 Transition functions

Let $(\mathcal{U}_\alpha, \phi_\alpha)$ be a local trivialization of P. This induces transition functions $g_{\alpha\beta} \colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ defined by

$$\begin{aligned} \phi_\alpha \circ \phi_\beta^{-1} \colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times G &\to \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times G \\ (x,a) &\mapsto (x, g_{\alpha\beta}(x)a). \end{aligned} \tag{56.203}$$

eq:transi_princ

Clearly, $g_{\alpha\alpha} = e$ and $g_{\alpha\beta}g_{\alpha\beta} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. Then the triviality

$$\phi_\alpha \circ \phi_\beta^{-1} \circ \phi_\beta \circ \phi_\gamma^{-1} \circ \phi_\gamma \circ \phi_\alpha^{-1} = \mathrm{Id}$$

implies the compatibility conditions

$$g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = e \tag{56.204}$$

on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$.

There is an inverse construction. Let $\{\mathcal{U}_\alpha \text{ tel que } \alpha \in I\}$ be an open covering of M and $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ a family of functions such that $g_{\alpha\alpha} = e$, $g_{\alpha\beta}g_{\alpha\beta} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and $g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$. Then the following construction gives a G-principal bundle whose transition functions are the $g_{\alpha\beta}$'s.

— $\tilde{P} = \bigsqcup_{\alpha\in I} \mathcal{U}_\alpha \times G$ (disjoint union),

— if $(x, a) \in \mathcal{U}_\alpha \times G$ and $(y, b) \in \mathcal{U}_\beta \times G$, then $(x, a) \sim (y, b)$ if and only if $x = y$ and $b = g_{\alpha\beta}(x)a$,

— $\pi\colon \tilde{P} \to M$ is defined by $\pi[(x, a)] = x$ where $[(x, a)]$ is the class of (x, a) for $\sim$,

— the action is defined by $[(x, a)] \cdot g = [(x, ag)]$.

Theorem 56.52.
Let G be a Lie group; M, a differentiable manifold; $\{\mathcal{U}_\alpha\}_{\alpha\in I}$, an open covering of M and some functions $\varphi_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ such that $\varphi_{\alpha\beta}(x) = \varphi_{\alpha\gamma}(x)\varphi_{\gamma\beta}(x)$. Then there exists a principal bundle P whose transition functions are the φ_α's for the covering $\{\mathcal{U}_\alpha\}_{\alpha\in I}$.

Proof. We consider the topological space

$$E = \bigcup_{\alpha\in I} (G \times \mathcal{U}_\alpha \times I) \tag{56.205}$$

where we put the discrete topology on I. Each $G \times \mathcal{U}_\alpha \times \{\alpha\}$ is a manifold. Thus E has a structure of differentiable manifold induced from the one of $G \times M$. We consider on E the equivalence relation given by the following subset of $E \times E$:

$$R = \big\{\big((g, x, \alpha), (h, y, \beta)\big) \in E \times E \text{ tel que } y = x \text{ and } h = \varphi_{\alpha\beta}(x)g\big\}.$$

We will show that $P = E/R$ has a structure of principal bundle. We begin by defining an action of G on P by

$$[(g, x, \alpha) \cdot h] = [(gh, x, \alpha)].$$

In order to see that this definition is correct, let us consider $[g', x, \beta] = [g, x, \alpha]$. From the definition of the equivalence class, $g' = \varphi_{\alpha\beta}(x)g$. Then $[(g', x, \beta)] \cdot h = [(\varphi_{\alpha\beta}(g)gh, x, \beta)]$, and the form of R shows that this is well $[(gh, x, \alpha)]$. Since the map $(g, h) \to gh$ is differentiable on G, the so defined action is a differentiable action of G on P and G is a transformation group on P [14].

If $[(g, x, \alpha)] = [(gh, x, \alpha)]$, then $gh = \varphi_{\alpha\alpha}g = g$ and $h = e$. So the action is effective.

Now we consider the quotient P/G. A typical element is

$$\overline{(s, x, i)} = \{[s, x, i] \cdot g \text{ tel que } g \in G\}.$$

The projection $\pi\colon P \to M$, $[(s, x, \alpha)] \to x$ is well defined and we can consider $\varphi\colon P/G \to M$, $\varphi\overline{(s, x, \alpha)} = x$. It provides a bijection between P/G and M. So we can identify P/G and M. Now we are going to show that P endowed with the projection $\pi\colon P \to X$ is a principal bundle.

We consider the map

$$h_\alpha\colon G \times \mathcal{U}_\alpha \to P \ (g, x) \quad \mapsto \omega(g, x, \alpha) \tag{56.206}$$

where $\omega\colon E \to P = E/R$ is the canonical projection. Since

$$(\pi \circ h_\alpha)(g, x) = (\pi \circ \omega)(g, x, \alpha) = \pi[(g, x, \alpha)] = x,$$

the map h_α actually is $h_\alpha\colon G \times \mathcal{U}_\alpha \to \pi^{-1}(\mathcal{U}_\alpha)$. In order to see that h_α is surjective on $\pi^{-1}(\mathcal{U}_\alpha)$, let us take a general element of $\pi^{-1}(\mathcal{U}_\alpha)$ under the form $\omega(g, x, \beta)$ with $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$. Then $(g, x, \beta) \in [(\varphi_{\alpha\beta}(x)g, x, \alpha)]$ and therefore $\omega(g, x, \beta) = h_\alpha(\varphi_{\alpha\beta}(x)g, x)$. For the injectivity, remark that $\omega(g, x, \beta) = \omega(h, y, \alpha)$ implies $x = y$ and $h = \varphi_{\beta\beta}(x)g = g$. In particular, $h_\alpha(g, x) = h_\alpha(h, y)$ implies $x = y$ and $g = h$.

14. Faut voir comment ça correspond à la définition de l'autre texte.

Now we will prove that the inverse of h_α is continuous. For this we consider an open set $\Omega \subset G \times \mathcal{U}_\alpha$ and we have to show that $h_\alpha(\Omega)$ is open in $\pi^{-1}(\mathcal{U}_\alpha)$.

We recall the **quotient topology**: if A is a topological space with an equivalence relation $\sim$ and the canonical projection $\varphi\colon A \to A/\sim$, then $V \subset A/\sim$ is open if and only if $\varphi^{-1}(V) \subset A$ is open. So in our case, we have to check the openness of $V = \omega^{-1}(h_\alpha(\Omega))$ in E. We consider the open covering

$$\{G \times \mathcal{U}_\alpha \times \{\alpha\}\}_{\alpha \in I}$$

of E and we will show that the intersection of V with any of these open set is open. We have to show that $\omega^{-1}\big(h_\alpha(\Omega) \cap (G \times \mathcal{U}_\alpha \times \{\beta\})\big)$ is open for any $\beta \in I$. For this, we define a map $\alpha\colon G \times (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \{\beta\} \to G \times \mathcal{U}_\alpha$ by

$$\alpha_\beta(g, x, \beta) = (\varphi_{\alpha\beta}(x)g, x) \tag{56.207}$$

which is continuous. The set $(h_\alpha \circ \alpha_\beta)^{-1}(h_\alpha(\Omega)) = \alpha_\beta^{-1}(\Omega)$ is open because $h_\alpha \circ \alpha_\beta$ is the restriction of ω to $G \times (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \times \{\beta\}$. Then h_α is an homeomorphism from $G \times \mathcal{U}_\alpha$ tp $\pi^{-1}(\mathcal{U}_\alpha)$. Since it is build from differentiable functions, it is moreover a diffeomorphism.

So we have a chart system $\{(h_\alpha, \mathcal{U}_\alpha)\}_{\alpha \in I}$ where h_α fulfils the "good" properties with respect to π. It remains to be proved that the $\varphi_{\alpha\beta}$'s are the transition functions and that $\pi^{-1}(\pi(\xi)) = \xi \cdot G$ for every $\xi \in P$. We begin by the latter. For $\xi = [(g, x, \alpha)]$, $\pi(\xi) = x$ and we have to study the set

$$\pi^{-1}(x) = \{[(h, x, \beta)] \text{ tel que } h \in G, \beta \in I\}.$$

Clearly, $[(h, x, \beta)] \cdot G \subset \pi^{-1}(x)$. The fact that there is nothing else than $[(h, x, \beta)] \cdot G$ in $\pi^{-1}(x)$ is seen by

$$[h, x, \beta] = [\varphi_{\alpha\beta}(x)g, x, \alpha] \in [(h, x, \alpha)] \cdot G.$$

In order to check the change of charts, let us consider $g' = h_{\beta,x}^{-1} \circ h_{\alpha,x}(g)$ where

$$h_{\alpha,x}(g) = h_\alpha(g, x) = \omega(g, x, \alpha). \tag{56.208}$$

The fact that $h_\beta(g', x) = g_\alpha(g, x)$ concludes the proof. To see this fact, remark that $h_{\beta,x}(h_{\beta,x}^{-1} \circ h_{\alpha,x}(g)) = h_{\alpha,x}(g)$, so that $h_\alpha(g', x) = h_\alpha(g, x)$ implies $\omega(g', x, \beta) = \omega(g, x, \alpha)$ which proves that $g' = \varphi_{\alpha\beta}(g)$. □

The **trivial bundle** is simply $P = M \times G$ and $\pi(x, g) = x$ with the action $(x, a) \cdot g = (x, ag)$.

56.6.2 Morphisms and such...

An **homomorphism** between $P(G, M)$ and $P'(G', M')$ is a differentiable map $h\colon P \to P'$ such that $\forall \xi \in P, g \in G$,

$$h(\xi \cdot g) = h(\xi) \cdot h_G(g) \tag{56.209}$$ eq:def_princ_homo

where $h_G\colon G \to G'$ is a Lie group homomorphism. From the definition, h maps a fiber to only one fiber, but it is not specially surjective on any fiber. So h induces a homomorphism $h_M\colon M \to M'$ such that $\pi' \circ h = h_M \circ \pi$.

An **isomorphism** is a homomorphism $g\colon P(G, M) \to P'(G', M')$ such that

— h_P is a diffeomorphism $P \to P'$,

— h_G is a Lie group homomorphism $G \to G'$, and

— h_M is a diffeomorphism $M \to M'$.

A principal bundle is **trivial** if one can find an isomorphism $h\colon G \times M \to P$ such that $\pi \circ h = \mathrm{Id} \circ \mathtt{proj}_2$, i.e. the following diagram commutes:

$$\begin{array}{ccc} G \times M & \xrightarrow{h} & P \\ {\scriptstyle \mathtt{proj}_2}\downarrow & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow[\mathrm{Id}]{} & M \end{array} \tag{56.210}$$ diag:princ_triv

We say that P is **locally trivial** if for every $x \in M$, there exists an open neighbourhood $\mathcal{U}$ in M such that $\pi^{-1}(\mathcal{U})$ endowed with the induced structure of principal bundle is trivial.

56.6.3 Frame bundle: first

pg:frame_bundle

In the ideas, the building of a vector bundle is just to put a vector space on each point of the base manifold. A principal bundle is to put something on which a group acts on each point. If you have a vector bundle on a manifold, you can consider, on each point $x \in M$, the set of all the basis of the fiber E_x over x. The group $GL(r, \mathbb{K})$ naturally acts on this set which becomes a candidate to be a $GL(r, \mathbb{K})$-principal bundle.

More formally, we consider a vector bundle $F \xrightarrow{p} M$, and for each x, the set of the basis of the vector space $F_x = p^{-1}(x)$. We define

$$P = \bigcup_{x \in M} (\text{basis of } F_x).$$

We naturally consider the projection $\pi\colon P \to M$, $\pi(b_x) = x$ if b_x is a basis of F_x.

Let $\phi_\alpha^F\colon p^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times \mathbb{K}^r$ be a local trivialization of F, and $\{\overline{e}_1, \dots, \overline{e}_r\}$, the canonical basis of $\mathbb{K}^r$. We naturally define

$$\overline{S}_{\alpha i}(x) = \phi_\alpha^{F\,-1}(x, \overline{e}_i).$$

The set $\{\overline{S}_{\alpha 1}(x), \dots, \overline{S}_{\alpha r}(x)\}$ is a "reference" basis of F_x with respect to the trivialization ϕ_α. If we choose another basis $\{\overline{v}_1, \dots \overline{v}_r\}$ of F_x, we can find a matrix $A \in GL(r, \mathbb{K})$ such that $\overline{v}_k = A_k^l \overline{S}_{\alpha l}(x)$. This gives a bijection

$$\begin{aligned} \phi_\alpha^P\colon \pi^{-1}(\mathcal{U}_\alpha) &\to \mathcal{U}_\alpha \times GL(r, \mathbb{K}) \\ (\overline{v}_1, \dots, \overline{v}_r) &\mapsto (x, A). \end{aligned} \tag{56.211}$$

One can give to P a $GL(r, \mathbb{K})$-principal bundle structure such that the ϕ_α^P are diffeomorphism.

Let $(\mathcal{U}_\alpha, \phi_\alpha^F)$ be a local trivialization of F and $g_{\alpha\beta}^F\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(r, \mathbb{K})$. In this case, $(\mathcal{U}_\alpha, \phi_\alpha^P)$ is a trivialization of P whose transition function is $g_{\alpha\beta}^P = g_{\alpha\beta}^F$. Indeed

$$\phi_\alpha^P \circ \phi_\beta^{P\,-1}(x, A) = \phi_\alpha^P(\{\overline{v}_1, \dots, \overline{v}_r\})$$

where $\overline{v}_s = (\phi_\beta^F)^{-1}(x, A_s^l \overline{e}_l)$. In order to see it, recall that $\overline{v}_s = A_s^l \overline{S}_{\alpha l}(x)$ and that $\phi_\alpha^{F\,-1}(x, \overline{e}_s) = \overline{S}_{\alpha s}(x)$. Then

$$\overline{v}_s = (\phi_\beta^F)^{-1}(x, A_s^l \overline{e}_l) = A_s^l \overline{S}_{\alpha s}(x).$$

On the other hand, from the definition of ϕ_β^P, the basis $(\phi_\beta^P)^{-1}(x, A)$ is the one obtained by applying A on S. With all this,

$$\begin{aligned} \phi_\alpha^P \circ (\phi_\beta^P)^{-1}(x, A) &= \phi_\alpha^P \{(\phi_\beta^F)^{-1}(x, A_s^l \overline{e}_l)\}_{s=1,\dots r} \\ &= \phi_\alpha^P \{(\phi_\alpha^F)^{-1} \circ (\phi_\alpha^E \circ \phi_\beta^{F\,-1})(x, A_s^l \overline{e}_l)\}_{s=1,\dots r} \\ &= \phi_\alpha^P \{(\phi_\alpha^E)^{-1}(x, g_{\alpha\beta}^F(x)_i^s A_s^l \overline{e}_l)\}_{i=1,\dots r} \\ &= (x, g_{\alpha\beta}^F(x) A). \end{aligned} \tag{56.212}$$

The last product $g_{\alpha\beta}^F(x)A$ is a matricial product.

56.6.4 Frame bundle: second

subsec_frbundle

56.6.4.1 Basis

subsubsecframebundle

If M is a m-dimensional manifold, a **frame** of T_xM is an isomorphism $b\colon \mathbb{R}^m \to T_xM$. In our purpose, we will always deal with (pseudo)Riemannian manifold. So, the tangents spaces T_xM comes with a metric, and we ask a frame to be isometric. In other words, we ask b to be an isometry from $(\mathbb{R}^m, \cdot\,)$ to (T_xM, g_x), where the dot denotes the (pseudo)euclidian product on $\mathbb{R}^m$. Such a frame is given by a base point x of M and a matrix S in $\mathrm{SO}(g_x)$:

$$b(v) = (Sv)^i (\partial_i)_x, \tag{56.213}$$

r1504e1

if the vector v is written as $v = v^i \overline{1}_i$ in the canonical orthogonal frame $\{\overline{1}_i\}$ of $\mathbb{R}^m$ and $\mathrm{SO}(g_x)$ is the set of the $m \times m$ matrix A such that $A^t g_x A = g_x$.

This frame intuitively corresponds to the basis of T_xM (see as a "true" vector space) that we would have written by $\{Se_i\}_x$ if $e_i = \frac{\partial}{\partial x^i}$. In order to follow this idea, we will effectively denote by $\{Se_i\}_x$ the map $b\colon \mathbb{R}^m \to T_xM$ given by (56.213).

We will often write the frame b as $\{be_i\}_x$, making no differences in notation between the b of $\mathrm{SO}(M)$ and the b of $\mathrm{SO}(g_x)$ which implement it.

Remarque 56.53.
One has to distinguish a *frame* and a *basis*: a basis is only a free and generating set while a frame can be interpreted as an ordered basis.

56.6.4.2 Construction

We just saw how to build a frame bundle over a manifold. One can get another expression of the frame bundle when we express a basis of T_xM by means of an isomorphism between $\mathbb{R}^n$ and T_xM. If M is a n-dimensional manifold, a **frame** at x is an ordered basis

$$b = (\overline{b}_1, \dots, \overline{b}_n)$$

of T_xM. It is clear that any frame defines an isomorphism (linear bijective map)

$$\begin{aligned} \tilde{b}\colon \mathbb{R}^n &\to T_xM \\ e_i &\mapsto \overline{e_i} \end{aligned} \tag{56.214}$$

where $\{e_i\}$ is the canonical basis of $\mathbb{R}^n$. It is also clear that any isomorphism gives rise to a frame. Then we see a frame of M at x as an isomorphism $\tilde{b}\colon \mathbb{R}^n \to T_xM$. Let $B(M)_x$ be the of all the frames of M at x; we define

$$B(M) = \bigcup_{x\in M} B(M)_x.$$

For all $b \in B(M)_x$, we define $p_B(b) = x$ and the action $B(M) \times GL(n,\mathbb{R}) \to B(M)$ by $b \cdot g = (\overline{b}'_1, \dots, \overline{b}'_n)$ where

$$\overline{b}'_j = \overline{b}_i g_i{}^j. \tag{56.215}$$

It is easy to see that $\widetilde{b \cdot g} = \tilde{b} \circ g\colon \mathbb{R}^n \to T_xM$. So we can give to

$$\begin{array}{ccc} GL(n,\mathbb{R}) & \rightsquigarrow & B(M) \\ & & \big\downarrow{\scriptstyle p_B} \\ & & M \end{array} \tag{56.216}$$

a structure of principal bundle [15]. If $(\mathcal{U}_\alpha, \varphi_\alpha)$ is a local coordinate chart on M, we define

$$\tilde{\varphi}\colon p_B^{-1}(\mathcal{U}_\alpha) \to \varphi_\alpha(\mathcal{U}_\alpha) \times GL(n,\mathbb{R})\ b \ \mapsto (\varphi_\alpha(x), A(b)) \tag{56.217}$$

where $A(b) \in GL(n,\mathbb{R})$ is defined by the condition $\overline{b}_j = A_j{}^i \partial_i|_x$. The matrix $A(b)$ is the one which transforms the canonical basis (in the trivialization φ_α) into $b \in B(M)_x$. That's for the principal bundle structure.

The manifold structure of $B(M)$ is given by $\Phi_\alpha\colon p_B^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times GL(\mathbb{R})$,

$$\begin{aligned} \Phi(b) &= (\varphi_\alpha^{-1} \times \mathrm{Id}\,|_{GL(n,\mathbb{R})}) \circ \tilde{\varphi}(b) \\ &= (x, A(b)) \\ &= (p_B(b), A(b)). \end{aligned} \tag{56.218}$$

It fulfils $A(b \cdot g) = A(b) \cdot g$. A section $s\colon \mathcal{U}_\alpha \to B(M)$ is sometimes called a **moving frame** over $\mathcal{U}_\alpha$.

Frame bundle over $\mathbb{R}^2$ is given as example in page 3694

15. Much more details and proofs are given in [915].

56.6.5 Sections of principal bundle

A **section** of a G-principal bundle is a smooth map $s\colon M \to P$ such that $s(x) \in \pi^{-1}(x)$ for any $x \in M$. A trivialization ϕ_α^P P on $\mathcal{U}_\alpha$ defines a section of P over $\mathcal{U}_\alpha$ by

$$\sigma_\alpha(x) = (\phi_\alpha^P)^{-1}(x, e)$$

where e is the neutral of the group. In the inverse sense, we have the following:

Proposition 56.54.
If $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ is local section of P over $\mathcal{U}_\alpha \subset M$, then the definition $\phi_\alpha^P(\xi) = (x, a)$ if $\xi = \sigma_\alpha(x) \cdot a$ is a local trivialization.

Proof. The function ϕ_α^P is well defined because $\xi \in \pi^{-1}(\mathcal{U}_\alpha)$ implies the existence of a $x \in \mathcal{U}_\alpha$ such that $\xi \in \pi^{-1}(x) = \{\xi \cdot g\} \simeq G$. For this x, there exists a $g \in G$ such that $\xi = \sigma_\alpha(x) \cdot g$.

Now we prove that the couple (x, a) is unique in the sense that $s_\alpha(x) \cdot a = \sigma_\alpha(y) \cdot b$ implies $(x, a) = (y, b)$. The left hand side belongs to $\pi^{-1}(x)$ while the right one belongs to $\pi^{-1}(y)$. Then $x = y$. The condition $\pi^{-1}(x) \simeq G$ imposes the unicity of the g making $\xi = \eta \cdot g$ for each couple, $\xi, \eta \in \pi^{-1}(x)$. □

If σ and σ' are two sections of the same principal bundle P, then there exists a differentiable map $f\colon M \to G$ such that $\sigma'(x) = \sigma(x) \cdot f(x)$. So all the sections can be deduced from only one and multiplication by such a f.

Theorem 56.55.
If $\pi\colon P(G, M) \to M$ is a principal bundle, then the four following propositions are equivalent: enuymai

(1) P is trivial, enuymaii

(2) P has a global section, enuymaiii

(3) there exists a differentiable map $\gamma\colon P \to G$ such that $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$ for all $\xi \in P$ and $g \in G$, enuymaiv

(4) there exists a differentiable map $\rho\colon P \to G$ such that $\rho(\xi \cdot g) = \rho(\xi)g$. ThoYPrincBTrivSect

Proof. *(1)⇒ (2).* The diagram (56.210) commutes and

$$\tau\colon M \to G \times M \; x \quad \mapsto (e, x) \tag{56.219}$$

is a local section of $G \times M$. From it we build the following global section of P:

$$\sigma\colon M \to P \; x \quad \mapsto h(e, x). \tag{56.220}$$

This is injective because $\pi \circ h = \mathtt{proj}_2$ and differentiable because this is a composition of $x \to (e, x)$ and $(g, x) \to h(g, x)$.
(2)⇒ (1). The principal bundle P admits a global section $\sigma\colon M \to P$. From it, we can build the differentiable map

$$h\colon G \times M \to P \; (g, x) \quad \mapsto \sigma(x) \cdot g \tag{56.221}$$

which satisfies $h(gh, x) = h(g, x) \cdot h$ and $\pi \circ (g, x) = x$. First we show that h is a fiber homomorphism and an isomorphism between P and $G \times M$ so that P is trivial. For this remark that

$$g(gh, x) = g(g, x) \cdot h = \sigma(x) \cdot gh,$$

hence equation (56.209) reduces to $h((g, x) \cdot h) = h(g, x) \cdot h_G(h)$ which is true with $h_G = \mathrm{Id}$. Moreover $h\colon G \times M \to P$ is bijective because $\sigma(\pi(\xi))$ belongs to the fiber of $\xi \in P$, therefore there is one and only one $\gamma(\xi) = u(\xi, \sigma(\pi(\xi)))$ such that $\xi \cdot \gamma(\xi) = (\sigma \circ \pi)\xi$. The inverse map is

$$\theta\colon P \to G \times M \; \xi \quad \mapsto (\gamma(\xi), \pi(\xi)) \tag{56.222}$$

which is differentiable because γ and π are. So far we see that h and h^{-1} are differentiable. Then h is an isomorphism between P and $G \times M$.
(2)⇒ (3). Let σ be the global section and define

$$\gamma\colon P \to G\ \xi \quad \mapsto u(\xi, (\sigma \circ \pi)\xi) \tag{56.223}$$

where $u\colon R \to G$ is the map defined by the condition $\xi \cdot (\xi, \eta) = \eta$. The map γ is differentiable and we have to prove that $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$. Since $\xi \cdot \gamma(\xi) = \sigma \circ \pi(\xi)$,

$$\gamma(\xi \cdot g) = u(\xi \cdot g, (\sigma \circ \pi)(\xi \cdot g)) = u(\xi \cdot g, (\sigma \circ \pi)(\xi)).$$

But $(\xi \cdot g)(g^{-1}\gamma(\xi)) = \xi \cdot \gamma(\xi) = x$. So $\gamma(\xi \cdot g) = u(\xi \cdot g, x)$. Thus $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$.
(3)⇒ (2). The given map γ fulfils $\xi \cdot g\gamma(\xi \cdot g) = \xi \cdot (\xi)$, so

$$\varphi\colon P \to P\ \xi \quad \mapsto \xi \cdot (\xi) \tag{56.224}$$

is just function of the class of ξ, thus we have a section $\sigma'\colon P/G \to P$, but we know that P/G and M are isomorphic.
(3)⇒ (4). Let us define $\rho\colon P \to G$ by $\rho = J \circ \gamma$ with $J(g) = g^{-1}$, thus $\rho(\xi) = \gamma(\xi)^{-1}$ and

$$\rho(\xi \cdot g) = \gamma(\xi \cdot g)^{-1} = (g^{-1}\gamma(\xi))^{-1} = \gamma(\xi)^{-1}g = \rho(\xi)g.$$

(4)⇒ (3). The proof is just the same with $\rho = J \circ \rho$. □

Definition 56.56.
*A section $\psi \in \Gamma(P, TP)$ is G-**equivariant** when*

$$d\tau_g\psi(\xi) = \psi(\xi \cdot g).$$

DefEqVectPrinc

Be careful: this *does not* define equivariant sections of the principal bundle.

56.6.6 Equivalence of principal bundle

Two principal bundles $\pi\colon P \to M$ and $\pi'\colon P' \to M$ are **equivalent** if there exists a diffeomorphism $\varphi\colon P \to P'$ such that

— $\pi' \circ \varphi = \pi$
— $\varphi(\xi \cdot g) = \varphi(\xi) \cdot g$.

If $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is an open covering of M on which we have trivializations ϕ_α of P and ψ_α of P', the diffeomorphism φ induces some functions $\lambda\colon \mathcal{U}_\alpha \to G$ by setting

$$(\phi_\alpha \circ \varphi^{-1} \circ \psi_\alpha^{-1})(x, a) = (x, \lambda_\alpha(x)a).$$

This definition works because from the definitions of principal bundle and equivalence, one sees that $(\phi_\alpha \circ \varphi^{-1} \circ \psi_\alpha^{-1})(x, \cdot\ ) = (x, \cdot\ )$.

56.6.6.1 Transition functions

We have some transition functions for P and P' given by equations

$$(\phi_\alpha \circ \phi_\beta^{-1})(x, g) = (x, g_{\alpha\beta}(x)g)$$
$$(\psi_\alpha \circ \psi_\beta^{-1})(x, g) = (x, g'_{\alpha\beta}(x)g).$$

Now, we want to know what is $g'_{\alpha\beta}$ in function of $g_{\alpha\beta}$. First remark that $(\psi_\alpha \circ \varphi \circ \phi_\alpha^{-1})(x,a) = (x, \lambda_\alpha(x)^{-1})a$, and next, compute

$$\begin{aligned}(x, g_{\alpha\beta}(x)a)a &= (\psi_\alpha \circ \varphi \circ \phi_\beta^{-1} \circ \phi_\beta \circ \varphi^{-1} \circ \psi_\beta^{-1})(x,a)\\ &= (\psi_\alpha \circ \varphi \circ \phi_\beta^{-1})(x, \lambda_\beta(x)a)\\ &= (\psi_\alpha \circ \varphi \circ \phi_\alpha^{-1} \circ \phi_\alpha \circ \phi_\beta^{-1})(x, \lambda_\beta(x)a)\\ &= (x, \lambda_\alpha(x)^{-1} g_{\alpha\beta}(x)\lambda_\beta(x)a).\end{aligned} \tag{56.225}$$

Then

$$g_{\alpha\beta}(x) = \lambda_\alpha(x)^{-1} g_{\alpha\beta}(x)\lambda_\beta. \tag{56.226}$$

One can show that if two principal bundle have transition functions whose fulfill this condition, they are equivalent. A G-principal bundle is **trivial** if it is equivalent to the one given by $\pi_1 \colon M \times G \to M$.

56.6.7 Reduction of the structural group

We say that a principal bundle $P(G,M)$ is **reducible** when there exists a principal bundle $P'(H,M)$ such that

— H is a subgroup of G,

— there exists an homeomorphism $h \colon P' \to P$ such that $h_G \colon H \to G$ is an injective homomorphism.

In this case we say that G is reducible to H and that P' is a reduced principal bundle.

Theorem 56.57.
If P is a principal bundle over M, the structural group G is reducible to the Lie subgroup H if and only if there exists an open covering $\{\mathcal{U}_i\}_{i\in I}$ of M and transition functions φ_{ij} taking their values in H.

Proof. No proof. □

The following comes from [916]. Let us consider the principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \big\downarrow{\scriptstyle \pi_P} \\ & & M \end{array} \tag{56.227}$$

and H, a closed subgroup of G. We denote by $j \colon H \to G$ the inclusion map. The principal bundle

$$\begin{array}{ccc} H & \rightsquigarrow & Q \\ & & \big\downarrow{\scriptstyle \pi_Q} \\ & & M \end{array} \tag{56.228}$$

is a **reduction** of P to the group H if there exists a map $u \colon Q \to P$ such that $\pi_P \circ u = \pi_Q$ and $u(\xi \cdot h) = u(\xi) \cdot j(h)$. In this case, u is an embedding [16] of Q in P and the image is a closed submanifold of P.

Let M be a n-dimensional manifold and $B(M)$ be its frame bundle. This is a $\mathrm{GL}(n,\mathbb{R})$-principal bundle. If G is a closed subgroup [17] of $\mathrm{GL}(n,\mathbb{R})$, a **$G$-structure** is a reduction of $B(M)$ to G.

16. plongement
17. Typically $\mathrm{SO}(p,q)$ or $\mathrm{SO}_0(p,q)$.

56.6.8 Density

A **density** on a d-dimensional manifold M is a section of the principal bundle whose fiber P_x over $x \in M$ is the space of homogeneous non vanishing maps

$$\rho\colon \bigwedge^d T_xM \to \mathbb{R}_+^* \tag{56.229}$$

such that $\rho(\lambda v) = |\lambda|\rho(v)$ for every $\lambda \in \mathbb{R}$ and $v \in \bigwedge^d T_xM$.

56.7 Associated bundle

Let $\pi\colon P \to M$ be a G-principal bundle and $\rho\colon G \to GL(V)$, a representation of G on a vector space V (on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) of dimension r.

The associated bundle $E = P \times_\rho V \xrightarrow{p} M$ is defined as following. On $P \times V$, we consider the equivalence relation

$$(\xi, v) \sim (\xi \cdot g, \rho(g^{-1})v)$$

for $g \in G$, $\xi \in P$ and $v \in V$. Then we define

— $E = P \times_\rho V := P \times V/\sim$,

— $p[(\xi, v)] = \pi(\xi)$

where $[(\xi, v)]$ is the class of (ξ, v) in $P \times V$.

If $\phi_\alpha^P(\xi) = (\pi(\xi), a(\xi))$ is a trivialization of P on $\mathcal{U}_\alpha$, then

$$\phi^E[(\xi, v)] = (\pi(\xi), \rho(a)v) \tag{56.230}$$

eq:triv_P_E

is a trivialization of E.

In order to see that it is a good definition, let us consider $(\eta, w) \sim (\xi, v)$. It immediately gives the existence of a $g \in G$ such that $\eta = \xi \cdot g$ and $w = \rho(g^{-1})v$. Then $\phi^E[(\xi \cdot g, \rho(g^{-1})v)] = (\pi(\xi \cdot g), \rho(b)\rho(g^{-1})v)$. From the definition of ϕ^E, the vector b is given by $\phi^P(\xi \cdot g) = (\pi(\xi \cdot g), b)$, and the definition of a principal bundle gives $b = \phi_{\pi(\xi)}(\xi \cdot g) = \phi_{\pi(\xi)}(\xi) \cdot g = ag$. The fact that ρ is a homomorphism makes $\rho(ag)\rho(g^{-1}) = \rho(a)v$ and ϕ^E is well defined.

Let G be a Lie group, ρ a representation of G on V and M, a manifold. We consider $P = M \times G \xrightarrow{\mathtt{proj}_1} M$, the trivial G-principal bundle on M. Then $E = P \times_\rho V \xrightarrow{p} M$ is **trivial**, i.e. we can build a $\varphi\colon P \times_\rho V \to M \times V$ such that $\mathtt{proj}_1 \circ \varphi = p$. It is rather easy: we define

$$\varphi\big[\big((x, g), v\big)\big] = (x, \rho(g)v).$$

It is easy to see that $(\mathtt{proj}_1 \circ \varphi)[(x, g), v] = x$ and $p[(x, g), v] = \mathtt{proj}_1(x, g) = x$.

56.7.1 Transition functions

Proposition 56.58.
Let $(\mathcal{U}_\alpha, \phi_\alpha^P)$ be a trivialization of $P \xrightarrow{\pi} M$ whose transition functions are $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$. Then $(\mathcal{U}_\alpha, \phi_\alpha^E)$ given by (56.230) is a local trivialization of $E \xrightarrow{p} M$ whose transition functions $g_{\alpha\beta}^E\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(\dim V, \mathbb{K})$ are given by

$$g_{\alpha\beta}^E(x) = \rho(g_{\alpha\beta}^P(x)).$$

Proof. If we write $a := \phi_{\beta x}^E(\pi^{-1}(x))$, we have $\phi_\beta^P(\pi^{-1}(x)) = (x, a)$ and $\phi_\alpha^E \circ (\phi_\beta^E)^{-1}(x, v) = \phi_\alpha^E[(\pi^{-1}(x), \rho(a)^{-1}v)]$. So,

$$\begin{aligned}\phi_\alpha^E[(\pi^{-1}(x), \rho(a)^{-1}v)] &= \Big(x, \rho\big(\phi_{\alpha x}(\pi^{-1}(x))\big)\rho\big(\phi_{\beta x}(\pi^{-1}(x))\big)^{-1}v\Big)\\ &= \Big(x, \rho\big(\phi_{\alpha x}(\pi^{-1}(x))\phi_{\beta x}(\pi^{-1}(x))\big)\Big).\end{aligned} \tag{56.231}$$

Then

$$g_{\alpha\beta}^E = \rho\Big(\phi_{\alpha x}(\pi^{-1}(x))\phi_{\beta x}\big(\pi^{-1}(x)\big)\Big) = \rho(g_{\alpha\beta}^P(x)). \tag{56.232}$$

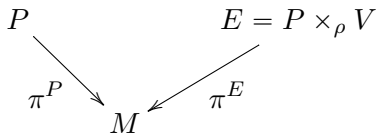
□

56.7.2 Sections on associated bundle

sec_fnequiv

56.7.2.1 Equivariant functions

We consider a bundle $E = P \times_\rho V \xrightarrow{p} M$ associated with the principal bundle $P \xrightarrow{\pi} M$ and a section $\psi\colon M \to E$.



A **section** of E is a map $\psi\colon M \to E$ such that $\pi^E \circ \psi = id_M$. We define the function $\hat{\psi}\colon P \to V$ by

$$\psi(\pi(\xi)) = [\xi, \hat{\psi}(\xi)]. \tag{56.233}$$

eq:equiv_psi

Let us see the condition under which this equation well defines $\hat{\psi}$. First, remark that a ψ defined by this equation is a section because $p[\xi, v] = \pi(\xi)$, so that $(p \circ \psi)(\pi(\xi)) = \pi(\xi)$. Now, consider a η such that $\pi(\eta) = \pi(\xi)$. Then there exists a $g \in G$ for which $\eta \cdot g = \xi$. For any g and for this one in particular,

$$\psi(\pi(\eta)) = [\eta, \hat{\psi}(\eta)] = [\eta \cdot g, \rho(g^{-1})\hat{\psi}(\eta)].$$

Then equation (56.233) defines $\hat{\psi}$ from ψ if and only if

$$\hat{\psi}(\xi \cdot g) = \rho(g^{-1})\hat{\psi}(\xi). \tag{56.234}$$

eq:equiv_psi_b

This condition is called the **equivariance** of $\hat{\psi}$. Reciprocally, any equivariant function $\hat{\psi}$ defines a section of $E = P \times_\rho V$.

If $\eta = \xi \cdot g = \chi \cdot k$, one define a sum

$$[\xi, v] + [\chi, w] = [\eta, \rho(g)v + \rho(k)w]. \tag{56.235}$$

eq:def:som_E

If $\psi, \eta\colon M \to E$ are two sections defined by the equivariant functions $\hat{\psi}, \hat{\eta}\colon P \to V$, then the section $\psi + \eta$ is defined by the equivariant function $\hat{\psi} + \hat{\eta}$.

56.7.2.2 For the endomorphism of sections of E

equivendo

Let us now make a step backward, and take A in $\operatorname{End}\Gamma(E)$. We will now see that A defines (and is defined by) an equivariant function $\hat{A}\colon P \to \operatorname{End} V$. Let $\psi\colon M \to E$ be in $\Gamma(E)$. If $\psi(x) = [\xi, v]$, we define the new section $A\psi$ by

$$(A\psi)(x) = [\xi, \hat{A}(\xi)v] = [\xi, \hat{A}(\xi)\hat{\psi}(\xi)].$$

In order for $A\psi$ to be well defined, the function $\hat{A}$ must satisfy

$$\hat{A}(\xi \cdot g) = \rho(g^{-1})\hat{A}(\xi)\rho(g) \tag{56.236}$$

equivA

for all g in G.

56.7.2.3 Local expressions

We consider a local trivialization $\phi^P_\alpha\colon \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G$ of P on $\mathcal{U}_\alpha$ and the corresponding section $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ given by

$$\sigma_\alpha(x) = (\phi^P_\alpha)^{-1}(x, e).$$

We saw at page 3388 that a trivialization of P gives a trivialization of the associated bundle $E = P \times_\rho V$; the definition is

$$\phi^E_\alpha[(\xi, v)] = (\pi(\xi), \rho(a)v) \tag{56.237}$$

if $\phi^P_\alpha(\xi) = (\pi(\xi), a)$. With $\xi = \sigma_\alpha(x)$, we find

$$\phi^E_\alpha[(\sigma_\alpha(x), v)] = (\pi(\sigma_\alpha(x)), \rho(a)v) = (x, v). \tag{56.238}$$

The section ψ can also be seen with respect to the "reference" sections σ_α by means of the definition

$$\psi(x) = [\sigma_\alpha(x), \psi_{(\alpha)}(x)] \tag{56.239}$$

eq:def:psisa

for a function $\psi_{(\alpha)}: M \to V$.

Lemma 56.59.
Let $\psi: M \to E$ be a section and $\hat{\psi}: P \to V$, the corresponding equivariant function. Then

$$\psi_{(\alpha)}(x) = \hat{\psi}(\sigma_\alpha(x)).$$

Proof. By definition, $\psi(x) = \psi(\pi(\xi)) = [\xi, \hat{\psi}(\xi)]$. Thus if we consider in particular $\xi = \sigma_\alpha(x)$,

$$\phi^E_\alpha(\psi(x)) = \phi^E_\alpha[\xi, \hat{\psi}(\xi)] = \phi^E_\alpha[s_\alpha(x), \hat{\psi}(\sigma_\alpha(x))] = (x, \hat{\psi}(\sigma_\alpha(x))). \tag{56.240}$$

□

Let us anticipate. A **spinor** is a section of an associated bundle $E = P \times_\rho V$ where P is a Lorentz-principal bundle, $V = \mathbb{C}^2$ and ρ is the spinor representation of Lorentz on $\mathbb{C}^2$. So a spinor $\psi: M \to E$ is *locally* described by a function $\psi_{(\alpha)}: M \to \mathbb{C}^2$. The latter is the one that we are used to handle in physics. In this picture, the transformation law of ψ under a Lorentz transformation comes naturally.

Let $\{e_i\}$ be a basis of V; we consider some "reference" sections $\gamma_{\alpha i}$ of the associated bundle $E = P \times_\rho V$ defined by

$$\gamma_{\alpha i}(x) = [\phi^{-1}_\alpha(x, e), e_i]. \tag{56.241}$$

eq:def:gamai

A general section $\psi: M \to E$ is defined by an equivariant function $\hat{\psi}: P \to V$ which can be written as $\hat{\psi}(\xi) = a^i(\xi)e_i$. If $\eta = \phi^{-1}_\alpha(x, e)$ and $\xi = \eta \cdot g(\xi)$,

$$\psi(x) = [\xi, a^i e_i] = a^i[\eta, \rho(g)e_i] = a^i(\xi)\rho(g(\xi))_i{}^j[\eta, e_j] = c^j(\xi)\gamma_{\alpha j}(x). \tag{56.242}$$

Since the left hand side of this equation just depends on x, the functions c^j must actually not depend on the choice of $\xi \in \pi^{-1}(x)$. So we have $c^j: M \to \mathbb{R}$. Indeed, if we choose $\chi \in \pi^{-1}(x)$,

$$\psi(x) = c^j(\xi)\gamma_{\alpha j}(x) \stackrel{!}{=} [\xi, a^i(\chi)e_i] = \ldots = c^j(\chi)\gamma_{\alpha j}(x),$$

so that $c^j(\xi) = c^j(\chi)$. So any section $\psi: M \to E$ can be decomposed (over the open set $\mathcal{U}_\alpha$) as

$$\psi(x) = s^i_\alpha(x)\gamma_{\alpha i}(x). \tag{56.243}$$

56.7.3 Associated and vector bundle

56.7.3.1 General construction

We are going to see that a vector bundle is an associated bundle. For this, we consider a vector bundle $p: F \to M$ with a fiber $F_x = V$ of dimension m. Let $G = GL(V)$, P be the trivial principal bundle $P = M \times G$ and ρ be the definition representation of G on V. We set $E = P \times_\rho V$. Our aim is to put a vector bundle structure on E which is equivalent to the one of F. The bijection $b: F \to E$ will clearly be

$$b(\phi^{-1}(x, v)) = [(x, e), v]. \tag{56.244}$$

We define the projection $q: E \to M$ by

$$q[(x, g), w] = x$$

and we have to show that $q^{-1}(x) = \{[(x, g), w] \text{ tel que } g \in G \text{ and } w \in V\}$ is a vector space isomorphic to V. The following definitions define a vector space structure:

— multiplication by a scalar: $\lambda[(x, g), v] = [(x, g), \lambda v]$,

— addition: $[(x, g), v] + [(x, h), w] = [(x, e), \rho(g)v + \rho(h)w]$.

As local trivialization map, we consider

$$\begin{aligned}\chi\colon q^{-1}(\mathcal{U}) &\to \mathcal{U}\times V\\ [(x,g),v] &\mapsto (x,\rho(g)v).\end{aligned} \tag{56.245}$$

With this structure, the bijection b is an equivalence because $b|_{F_x}$ is a vector space isomorphism and $q\circ b = p$.

56.7.4 Equivariant functions for a vector field

equivvec

In order to define in the same way an equivariant function for a vector field $X\in\mathfrak{X}(M)$, we need to see TM as an associated bundle.

Proposition 56.60.
If M is a n dimensional manifold, we have the following isomorphism:

$$\mathrm{SO}(M)\times_{\rho^M}\mathbb{R}^m \simeq TM$$

where $\rho^M\colon \mathrm{SO}(m)\times\mathbb{R}^m\to\mathbb{R}^m$ is defined by $\rho^M(A)v = Av$.

Proof. Recall that an element $b\in\mathrm{SO}(M)_x$ is a map $b\colon\mathbb{R}^m\to T_xM$. The isomorphism is no difficult. It is $\psi\colon \mathrm{SO}(M)\times_{\rho^M}\mathbb{R}^m\to TM$ defined by

$$\psi[b,v] = b(v).$$

It prove no difficult to see that ψ is well defined, injective and surjective. □

Now, let us consider $X\in\mathfrak{X}(M)$. We can see it as an element of $\Gamma(\mathrm{SO}(M)\times_{\rho^M}\mathbb{R}^m)$, and define an equivariant function $\hat{X}\colon \mathrm{SO}(M)\to\mathbb{R}^m$.

Let us make it more explicit. A vector field $Y\in\mathfrak{X}(M)$ is, for each x in M, the data of a tangent vector $Y_x\in T_xM$. Hence the formula $b(v) = Y_x$ defines an element $[b,v]$ in $\mathrm{SO}(M)\times_{\rho^M}\mathbb{R}^m$, and Y defines a section $\tilde{Y}(x) = [b(x),v(x)]$ of $\mathrm{SO}(M)\times_{\rho^M}\mathbb{R}^m$. The associated equivariant function is given by $\hat{Y}(b) = v$ if $b(v) = Y_x$. In other words, the equivariant function $\hat{Y}\colon\mathrm{SO}(M)\to\mathbb{R}^m$ associated with the vector field $Y\in\mathfrak{X}(M)$ is given by

$$\hat{Y}(b) = b^{-1}(Y_x), \tag{56.246}$$

r1404e1

where $x = \pi(b)$.

56.7.5 Gauge transformations

A **gauge transformation** of the G-principal bundle $\pi\colon P\to M$ is a diffeomorphism $\varphi\colon P\to P$ such that

— $\pi\circ\varphi = \pi$,
— $\varphi(\xi\cdot g) = \varphi(\xi)\cdot g$.

When we consider some local sections on $\sigma_\alpha\colon\mathcal{U}_\alpha\to P$, we can describe a gauge transformation with a function $\tilde{\varphi}_\alpha\colon M\to G$ by requiring

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x)\cdot\tilde{\varphi}_\alpha(x).$$

This formula defines φ from $\tilde{\varphi}$ as well as $\tilde{\varphi}$ from φ.

The group of gauge transformations has a natural action on the space of sections given by

$$(\varphi\cdot\psi)(x) = [\varphi(\xi),v]. \tag{56.247a}$$

if $\psi(x) = [\xi,v] = [\xi,\hat{\psi}(\xi)]$. This law can also be seen on the equivariant function $\hat{\psi}$ which defines ψ. The rule is

$$\widehat{\varphi\cdot\psi}(\xi) = \hat{\psi}(\varphi^{-1}(\xi)). \tag{56.247b}$$

Indeed, in the same way as before we find $(\varphi\cdot\psi)(x) = [\xi,\widehat{\varphi\cdot\psi}(x)] \overset{!}{=} [\varphi(\xi),v] = [\varphi(\xi),\hat{\psi}(\xi)]$. Taking $\xi\to\varphi^{-1}(\xi)$ as representative, $(\varphi\cdot\psi)(x) = [\xi,\hat{\psi}\circ\varphi^{-1}(\xi)]$.

56.8 Adjoint bundle

Let $\pi\colon P \to M$ be a G-principal bundle. The **adjoint bundle** is the associated bundle $\mathrm{Ad}(P) = P \times_{\mathrm{Ad}} \mathcal{G}$. An element of that bundle is an equivalent class given by

$$[\xi, X] = [\xi \cdot g, \mathrm{Ad}(g^{-1})X]$$

for every $g \in G$. Here $\xi \in P$ and $X \in \mathcal{G}$.

56.9 Connection on vector bundle: local description

DEFooIESVooGiNQHd

Definition 56.61.
*A **connection** on the vector bundle $p\colon E \to M$ is a bilinear map*

$$\begin{aligned} \nabla\colon \mathfrak{X}(M) \times \Gamma(E) &\to \Gamma(E) \\ (X, s) &\mapsto \nabla_X s \end{aligned} \tag{56.248}$$

such that

— $\nabla_{fX} s = f\nabla_X s$,
— $\nabla_X(fs) = (X \cdot f)s + f\nabla_X s$

*for all $X \in \mathfrak{X}(M)$, $f \in C^\infty(M)$ and $s \in \Gamma(E)$. The operation ∇ is often called a **covariant derivative**.*

An easy example is given on the trivial bundle $E = \mathtt{proj}_1\colon M \times \mathbb{C} \to M$. For this bundle, $\Gamma(E) = C^\infty(M, \mathbb{C})$ and the common derivation is a covariant derivation: $\nabla_X s = (ds)X$.

PROPooWZYOooEVhgFt

Proposition 56.62.
The value of $(\nabla_X s)(x)$ depends only on X_x and s on a neighbourhood of $x \in M$.

Proof. Let X, $Y \in \mathfrak{X}(M)$ such that $Y_z = f(z)X_z$ with $f(x) = 1$ and $f(z) \neq 1$ everywhere else. Then

$$(\nabla_Y s)(x) - (\nabla_X s)(x) = (f(x) - 1)(\nabla_X s)(x) = 0.$$

Since it is true for any function, the linearity makes that it cannot depend on X_z with $z \neq x$. If we consider now two sections s and s' which are equals on a neighbourhood of x, we can write $s' = fs$ for a certain function f which is 1 on the neighbourhood. Then

$$(\nabla_X s')(x) - (\nabla_X s)(x) = (f(x) - 1)(\nabla_X s)(x) + (Xf)s(x)$$

which zero because on a neighbourhood of x, f is the constant 1. □

This proposition shows that it makes sense to consider only local descriptions of connections. Let $\{e_1, \dots, e_r\}$ be a basis of V and consider the local sections $\overline{S}_{\alpha i}\colon \mathcal{U}_\alpha \to E$,

$$\overline{S}_{\alpha i}(x) = \phi_\alpha^{-1}(x, e_i).$$

A local section $s_\alpha\colon \mathcal{U}_\alpha \to V$ can be decomposed as $s_\alpha(x) = s_\alpha^i(x)e_i$ with respect to this basis (up to an isomorphism between the different V at each point). Then on $\mathcal{U}_\alpha$,

$$s_\alpha^i \overline{S}_{\alpha i}(x) = s_\alpha^i(x)\phi_\alpha^{-1}(x, e_i) = \phi_\alpha^{-1}(x, s_\alpha^i e_i) = \phi_\alpha^{-1}(x, s_\alpha(x)) = s(x). \tag{56.249}$$

The first equality is the definition of the product $\mathbb{R} \times F \to F$.

So any $s \in \Gamma(E)$ can be (locally!) written under the form [18] $s = s_\alpha^i \overline{S}_{\alpha i}$; in particular $\nabla_X(\overline{S}_{\alpha i})$ can. We define the coefficients θ by

$$\nabla_X(\overline{S}_{\alpha i}) = (\theta_\alpha)_i^j(X)\overline{S}_{\alpha j}. \tag{56.250}$$

where, for each i and j, $(\theta_\alpha)_i^j$ is a 1-form on $\mathcal{U}_\alpha$. We can consider θ_α as a matrix-valued 1-form on $\mathcal{U}_\alpha$.

18. be careful on the fact that the "coefficient" s_α^i depends on x: the right way to express this equation is $s(x) = s_\alpha^i(x)\overline{S}_{\alpha i}(x)$.

PropFormnabXthe

Proposition 56.63.
The formula

$$(\nabla_X s)_\alpha = X s_\alpha + \theta_\alpha(X) s_\alpha \tag{56.251}$$

eq:nab_theta

gives a local description of the connection.

Proof. For any $s \in \Gamma(E)$, we have

$$\nabla_X s = \nabla_X\big(\sum_j s^j_\alpha \overline{S}_{\alpha j}\big) = \sum_j \Big((X s^j_\alpha)\overline{S}_{\alpha j} + s^j_\alpha \nabla_X \overline{S}_{\alpha j}\Big) = \sum_i \Big[(X s^i_\alpha) + s^j_\alpha (\theta_\alpha)^i_j(X)\Big]\overline{S}_{\alpha i}. \tag{56.252}$$

□

56.9.1 Connection and transition functions

A connection determines some local matrix-valued 1-forms θ_α on the trivialization $\mathcal{U}_\alpha$. Two natural questions raise. The first is the converse: does a matrix-valued 1-form defines a connection? The second is to know what is θ_α in function of θ_β on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$? The answer to the latter is given by the following proposition:

Proposition 56.64.
The 1-form θ_α relative to the trivialization $(\mathcal{U}_\alpha, \phi_\alpha)$ is related to the 1-form θ_β relative to the trivialization $(\mathcal{U}_\beta, \phi_\beta)$ by

$$\theta_\beta = g^{-1}_{\alpha\beta} d g_{\alpha\beta} + g^{-1}_{\alpha\beta} \theta_\alpha g_{\alpha\beta}. \tag{56.253}$$

eq:theta_g

Proof. We can use equation (56.17) pointwise on $(\nabla_X s)_\alpha$:

$$\begin{aligned}(\nabla_X s)_\alpha &= g_{\alpha\beta}(\nabla_X s)_\beta \\ &= g_{\alpha\beta}\big(X s_\beta + \theta_\beta(X) s_\beta\big) \\ &= g_{\alpha\beta}\big(X(g_{\alpha\beta} s_\alpha) + \theta_\beta(X) g_{\alpha\beta} s_\alpha\big).\end{aligned} \tag{56.254}$$

We have to compare it with equation (56.251). Note that $g_{\alpha\beta}$ and $\theta_\alpha(X)$ are matrices, then one cannot do

$$g_{\alpha\beta}\theta_\beta(X) g_{\alpha\beta} = g_{\alpha\beta} g_{\alpha\beta} \theta_\beta(X) = \theta_\beta(X)$$

by using $g_{\alpha\beta} g_{\alpha\beta} = \mathbb{1}$. Taking carefully subscripts into account, one sees that the correct form is $(g_{\alpha\beta})^i_j \theta_\beta(X)^j_k (g_{\alpha\beta})^k_l$. Applying Leibniz formula $(X(fg) = f(Xg) + (Xf)g)$, and making the simplification $g_{\alpha\beta} g_{\alpha\beta} = \mathbb{1}$ in the first term, we find

$$\theta_\alpha(X) s_\alpha = g_{\alpha\beta}(X g_{\alpha\beta}) s_\alpha + g^{-1}_{\alpha\beta} \theta_\beta(X) g_{\alpha\beta} s_\alpha.$$

The claim follows from the fact that $X g_{\alpha\beta} = d g_{\alpha\beta}(X)$. □

56.65.
The equation (56.253) are related to the equations (56.350) or (56.358) or any physical equation of gauge transformation for the bosons.

56.66.
Notice that formula (56.253) shows in particular that θ_α takes its values in the Lie algebra $\mathfrak{gl}(V)$, see for example subsection 52.1.2.

The inverse is given in the

Propformconnve

Proposition 56.67.
If we choose a family of $\mathfrak{gl}(V)$-valued 1-forms θ_α on $\mathcal{U}_\alpha$ satisfying (56.253),then the formula

$$(\nabla_X s)_\alpha = X s_\alpha + \theta_\alpha(X) s_\alpha$$

defines a connection on E.

prop:thet_conn_F

Proof. Note that θ is $C^\infty(M)$-linear, thus

$$(\nabla_{fX}s)_\alpha = (fX)s_\alpha + \theta_\alpha(fX)s_\alpha = f[Xs_\alpha + \theta_\alpha(X)s_\alpha] = f(\nabla_X s)_\alpha. \tag{56.255}$$

In expressions such that $\theta_\alpha(X)(fs_\alpha)$, the product is a matrix times vector product between $\theta_\alpha(X)$ and s_α; the position of the f is not important. So we can check the second condition:

$$\begin{aligned}(\nabla_X(fs))_\alpha &= X(fs_\alpha) + \theta_\alpha(X)(fs_\alpha)\\ &= X(f)s_\alpha + f(Xs_\alpha) + f\theta_\alpha(X)s_\alpha\\ &= df(X)s_\alpha + f(\nabla_X s)_\alpha.\end{aligned} \tag{56.256}$$

This concludes the proof. □

56.9.2 Torsion and curvature

The map $T^\nabla\colon \mathfrak{X}(X)\times\mathfrak{X}(X)\to\mathfrak{X}(X)$ defined by

$$T^\nabla(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \tag{56.257}$$

deftorsion

is the **torsion** of the connection ∇. When $T^\nabla(X,Y)=0$ for every X and Y in $\mathfrak{X}(X)$, we say that ∇ is a **torsion free** connection. Let X, Y be in $\mathfrak{X}(M)$, and consider the map $R(X,Y)\colon\Gamma(E)\to\Gamma(E)$ defined by

$$\begin{aligned}R(X,Y)\colon \Gamma(E)&\to\Gamma(E)\\ s&\mapsto \nabla_X\nabla_Y s - \nabla_Y\nabla_X s - \nabla_{[X,Y]}s.\end{aligned} \tag{56.258}$$

For each $x\in M$, R can be seen as a bilinear map $R\colon T_xM\times T_xM\to\mathrm{End}(E_x)$. It is called the **curvature** of the connection ∇. For every $f\in C^\infty(M)$, it satisfies

$$R(fX,Y)s = fR(X,Y)s = R(X,Y)fs.$$

In a trivialization $(\mathcal{U}_\alpha,\phi_\alpha)$, we have $(\nabla_X s)_\alpha = Xs_\alpha + \theta_\alpha(X)s_\alpha$. Looking in the expression of $(R(X,Y)s)_\alpha$, the terms coming from the Xs_α part of covariant derivative make

$$XYs_\alpha - YXs_\alpha - [X,Y]s_\alpha = 0.$$

The other terms are no more than matricial product. Hence the formula

$$(R(X,Y)s)_\alpha = \Omega_\alpha(X,Y)s_\alpha \tag{56.259}$$

defines a 2-form Ω_α which takes values in $GL(r,\mathbb{K})$. We can find an expression for Ω in terms of θ:

$$\Omega_\alpha(X,Y) = X\theta_\alpha(Y) - Y\theta_\alpha(X) - \theta_\alpha([X,Y]) + \theta_\alpha(X)\theta_\alpha(Y) - \theta_\alpha(Y)\theta_\alpha(X);$$

it is written as

$$\Omega_\alpha = d\theta_\alpha + \theta_\alpha\wedge\theta_\alpha = d\theta_\alpha + \frac{1}{2}[\theta_\alpha,\theta_\alpha] \tag{56.260}$$

eq:Omega_ttheta

which is a notational shortcut for

$$\Omega_\alpha(X,Y) = d\theta_\alpha(X,Y) + [\theta_\alpha(X),\theta_\alpha(Y)]. \tag{56.261}$$

EaCurvdVVsq

These equations are called **structure equations**. Pointwise, the second term is a matrix commutator; be careful on the fact that, when we will speak about principal bundle, the forms θ's will take their values in a Lie algebra. On $\mathcal{U}_\alpha\cap\mathcal{U}_\beta$, we have

$$\Omega_\beta(X,Y) = g_{\alpha\beta}^{-1}\Omega_\alpha(X,Y)g_{\alpha\beta}.$$

The curvature and the connection fulfill the **Bianchi identities**:

Lemma 56.68.

$$d\Omega_\alpha + [\theta_\alpha \wedge \Omega_\alpha] = 0.$$

Proof. For each matricial entry, θ_α is a 1-form on $\mathcal{U}_\alpha$, then $\theta_\alpha(X)$ is a function which to $x \in M$ assign $\theta_\alpha(x)(X_x) \in \mathbb{R}$. So we can apply d and Leibniz on the product $\theta_\alpha(X)\theta_\alpha(Y)$.

$$d\big(\theta_\alpha(X)\theta_\alpha(Y)\big) = \theta_\alpha(X)d\theta_\alpha(Y) + d\theta_\alpha(X)\theta_\alpha(Y).$$

Differentiating equation (56.260), $d\Omega_\alpha = d\theta_\alpha \wedge \theta_\alpha - \theta_\alpha \wedge d\theta_\alpha$. □

56.10 Connexion on vector bundle: algebraic view

We already defined a connection of the vector bundle $\pi\colon E \to M$ in definition 56.61. As we know from section 54.2.2, this induces a map

$$\nabla\colon \Gamma(E) \to \Gamma(E) \otimes \Omega^1(M) \tag{56.262}$$

EQooBRLHooJgzJyT

Note: there is a non trivial identification $\Gamma(TM)^* = \Gamma(T^*M)$. This ∇ is bilinear and satisfy the Leibniz rule

$$\nabla(\sigma f) = (\nabla\sigma)f + \sigma \otimes df \tag{56.263}$$

for any section $\sigma\colon M \to E$ and function $f\colon M \to \mathbb{C}$. If $\{\sigma_i\}$ is a local basis of E, one can write $\sigma = \sigma_i f^i$ and one defines the **Christoffel symbols** $\Gamma^j_{i\mu}$ in this basis by

$$\nabla\sigma = \nabla(\sigma_i f^i) = (\nabla\sigma_i)f^i + \sigma_i \otimes f(f^i) = f^i\Gamma^j_{i\mu}\sigma_j \otimes dx^\mu + \sigma_i \otimes d(f^i). \tag{56.264}$$

The notations $d\sigma = \sigma_i \otimes d(f^i)$ and $\Gamma\sigma = f^i\Gamma^j_{i\mu}\sigma_j \otimes dx^\mu$ lead us to the compact usual form

$$\nabla\sigma = (d + \Gamma)\sigma.$$

When $E = TM$ over a (pseudo)Riemannian manifold M, we know the Levi-Civita connection which is compatible with the metric:

$$g(\nabla X, Y) + g(X, \nabla Y) = d\big(g(X, Y)\big). \tag{56.265}$$

eq_230605r1

One can see g as acting on $\big(\mathfrak{X}(M) \otimes \Omega^1(M)\big) \times \mathfrak{X}(M)$ with

$$g\big(r^i_\nu \partial_i \otimes dx^\nu, t^j\partial_j\big) := r^i_\nu j^j g(\partial_i, \partial_j)dx^\nu,$$

which at each point is a form. From condition (56.265), we see ∇ as a Levi-Civita connection on the bundle $E = T^*M$ which values in

$$\Gamma^\infty(T^*M) \otimes \Omega^1(M) \simeq \Omega^1(M) \otimes \Omega^1(M).$$

This is defined as follows. A 1-form ω can always be written under the for $\omega = X^\flat := g(X, .)$ for a certain $X \in \mathfrak{X}(M)$. Then (56.265) gives

$$(\nabla X)^\flat Y + \omega(\nabla Y) = d(\omega Y),$$

and we put $\nabla\omega = (\nabla X)^\flat$, i.e

$$(\nabla\omega)Y = d(\omega Y) - \omega(\nabla Y) \tag{56.266}$$

for all $Y \in \mathfrak{X}(M)$. When $\omega = dx^i$ and $Y = \partial_j$, we find

$$(\nabla dx^i)\partial_j = d(dx^i\partial_j) - dx^i(\nabla\partial_j) = d(\delta^i_j) - \Gamma^l_{jk}\partial_l \otimes dx^k = -\Gamma^l_{jk}\delta^i_l \otimes dx^k = -\Gamma^i_{jk}\,dx^k. \tag{56.267}$$

So we get the local formula

$$\nabla dx^i = -\Gamma^i_{jk}\,dx^j \otimes dx^k. \tag{56.268}$$

If the form writes locally $\omega = dx^i f_i$,

$$\nabla\omega = \nabla(dx^i)f_i + dx^i \otimes df_i = -f_i\Gamma^i_{jk}\,dx^j \otimes dx^k + d\omega = (d - \tilde{\Gamma})\omega \tag{56.269}$$

where we taken the notations $d\omega = dx^i \otimes df_i$ and $\tilde{\Gamma}\omega = f_i\Gamma^i_{jk}dx^j \otimes dx^k$.

56.10.1 Exterior derivative

If E is a m-dimensional vector bundle over M and $s\colon M \to E$ is a section, we say that a **exterior derivative** is a map $D\colon \Gamma(E) \to \Gamma(E \otimes \Omega^1 M)$ such that for every $f \in C^\infty(M)$ we have

$$D(fs) = s \otimes df + f(Ds).$$

An exterior derivative can be extended to $D\colon \Gamma(E \otimes \Omega^k M) \to \Gamma(E \otimes \Omega^{k+1} M)$ imposing the condition

$$D(\omega \wedge \alpha) = (D\omega) \wedge \alpha + (-1)^k \omega \wedge d\alpha \tag{56.270}$$

EqExtExtDerrk

for every $\omega \in \Gamma(E \otimes \Omega^k M)$ and $\alpha \in \Gamma(E \otimes \Omega^l M)$. The result is an element of $\Gamma(E \otimes \Omega^{k+l+1} M)$.

Coordinatewise expressions are obtained when one choose a specific section (e_i) of the frame bundle of E. In that case for each i, the derivative e_i is an element of $\Gamma(E \otimes \Omega^1 M)$ and we define $\omega_i^j \in \Omega^1(M)$ by

$$De_i = \sum_{j=1}^{k} e_j \otimes \omega_i^j. \tag{56.271}$$

For each i and j, we have an element $\omega_i^j \in \Omega^1(M)$, so that we say that $\omega \in \Omega^1(M, \mathfrak{gl}(m))$. Now a section can be expressed as $s = s^i e_i$ where s^i are functions, so we have

$$D(s) = D(s^i e_i) = e_i \otimes ds^i + s^i D(e_i) = e_i \otimes ds^i + s^i e_j \otimes \omega_i^j = e_i \otimes ds^i + e_i \otimes s^j \omega_j^i. \tag{56.272}$$

Expressed in component, we find $D(s)^i = ds^i + s^j \omega_j^i$, so that we often write

$$D = d + \omega. \tag{56.273}$$

When a section e is given, we write $s = s^i(e) e_i$, indicating the dependence of the functions s^i in the choice of the frame e:

$$D(s) = e_i \otimes ds^i(e) + e_i \otimes s^j(e) \omega(e)_j^i.$$

When we apply both sides to a vector $X \in \Gamma(TM)$, we find

$$D_X(s) = e_i \otimes \Big(X(s^i) + s^j \omega_j^i(X) \Big). \tag{56.274}$$

By convention we say that, when $f \in C^\infty(M)$, is a function, D_X reduces to the action of the vector field X:

$$D_X(f) = X(f). \tag{56.275}$$

56.10.1.1 Covariant exterior derivative

An important exterior derivative is the covariant exterior derivative. If the vector bundle E is endowed by a covariant derivative ∇, we define the corresponding **covariant exterior derivative** by the following:

(1) for a section $s\colon M \to E$ (i.e. a 0-form) we define

$$(d_\nabla s)(X) = \nabla_X s, \tag{56.276}$$

(2) and on the 1-form $\sum_i (s_i \otimes \omega_i) \in \Gamma(E \otimes T^*M)$,

$$d_\nabla \big(\sum_i s_i \otimes \omega_i \big) = \sum_i (d_\nabla s_i) \wedge \omega_i + \sum_i s_i \otimes d\omega_i. \tag{56.277}$$

The latter relation is the condition (56.270) with $k = 0$.

56.10.2 Divergence, gradient and Laplacian (general)

When we have a connection (definition 56.61) we define the divergence of a vector field. Recall from equation (56.262) that when $x \in \Gamma(TM)$ we can see ∇X as an element of $\Gamma(TM) \otimes \Omega^1(M)$.

DEFooTTSFooDdgiKg

Definition 56.69.
*The **divergence** of a vector field $X \in \Gamma(TM)$, is the function $\nabla \cdot X \in C^\infty(M)$ defined by*

$$(\nabla \cdot X)(x) = \mathrm{Tr}(\nabla X) \tag{56.278}$$

where the trace stands for the contraction of the tensor ∇X with itself.

On an euclidian space, we can describe the divergence in an other way: by proposition 56.62, the vector $(\nabla_X Y)_x \in T_xM$ does depend only on X_x. Thus in fact, when $v \in T_xM$, the value of $\nabla_v X$ is the value of $(\nabla_V X)_x$ where V is any vector field such that $V_x = v$. Thus we are okay speaking about the map

$$\begin{aligned} T_xM &\to T_xM \\ v &\mapsto \nabla_v X. \end{aligned} \tag{56.279}$$

This operation being linear, we can take the trace. Thus the definition could be

$$\nabla \cdot X = \mathrm{Tr}(v \mapsto \nabla_v X). \tag{56.280}$$

EQooEGYQooJAydTQ

The point using the contraction notion is that the operation of tracing is more clear: from equation (56.280) one does not see why the result should depend on the metric (and in particular why does it depend on the *inverse* of the metric tensor in a matricial form.).

56.10.3 Divergence, gradient and Laplacian (Riemannian case)

If we have a (pseudo)Riemannian manifold,we define the gradient of a function.

Definition 56.70.
*We define the **gradient** of a function $f \in C^\infty(M)$, denoted by ∇f as the vector field such that*

$$g(\nabla f, X) = X(f). \tag{56.281}$$

EQooECZSooYfQEYm

Moreover on a Riemannian manifold, we have the Levi-Civita connection (see 56.13.2), and then it is possible to particularize the definition 56.69 to a well defined connection.

The **Laplacian** of the function f is the function Δf given by

$$\Delta f = \nabla \cdot (\nabla f). \tag{56.282}$$

56.10.4 Divergence, gradient and Laplacian (coordinatewise)

Proposition 56.71.
Let $\{e_i\}$ be a field of base for the riemannian manifold (M, g), and a function $f \in C^\infty(M)$. Then the gradient of f is given by the formula

$$\nabla f = \sum_{im} (g^{-1})_{im} \partial_i f e_m. \tag{56.283}$$

The important point is that g^{-1} appears instead of g.

Proof. In coordinates, this is given by the *inverse* of the metric tensor. Indeed equation (56.281) expands to (unwritten dependence to x)

$$\sum_{kl} g_{kl} (\nabla f)_k X_l = X(f) = \sum_l X_l \partial_l f. \tag{56.284}$$

Since that has to be true for every vector field X, we can "simplify" X_l:

$$\sum_k g_{kl}(\nabla f)_k = \partial_l f. \tag{56.285}$$

Multiplying by $(g^{-1})_{lj}$ and making the sum over l:

$$\sum_{kl} g_{kl}(g^{-1})_{lj}(\nabla f)_k = \sum_l (g^{-1})_{lj}\partial_l f, \tag{56.286}$$

and then

$$(\nabla f)_j = \sum_l (g^{-1})_{lj}\partial_l f. \tag{56.287}$$

□

PROPooLIJTooKFTwPY

Proposition 56.72.
Let (M, g) be a (pseudo)Riemannian manifold with constant g. Then

$$\nabla \cdot Y = \sum_{ij} (g^{-1})_{ij}\partial_i Y_j. \tag{56.288}$$

Proof. The fact that g is constant implies that the Christoffel symbols are vanishing and the Levi-Civita connection is given by

$$\nabla_X Y = X(Y) = \sum_{kl} X_l \partial(Y_k)\partial_k. \tag{56.289}$$

As explained around equation 56.262 we see $\nabla Y \in \Gamma(TM) \otimes \Omega^1(M)$ as

$$\nabla Y = \sum_{kl} (\partial_l)(Y_k)\partial_k \otimes dx_l. \tag{56.290}$$

Using formula (54.10) for the contraction,

$$\mathrm{Tr}(\nabla Y) = \sum_{kl} (g^{-1})_{kl}\partial_l(Y_k). \tag{56.291}$$

This is the divergence of Y. □

56.10.5 Soldering form and torsion

Let us particularize to the case where E has the same dimension as the manifold. In that case, we can introduce a **soldering form**, that is an element $\theta \in \Omega^1(M, E)$ such that for every $x \in M$ the map $\theta_x : T_xM \to E_x$ is a vector space isomorphism. When a soldering form θ is given, the **torsion** is the exterior derivative D is

$$T = D\theta. \tag{56.292}$$

Using a local frame e, we have forms $\theta^i(e) \in \Omega^1(M)$ such that

$$\theta(X) = \theta^i(X)e_i.$$

We see θ as an element of $\Gamma(E \otimes \Omega^1(M))$ by identifying $\theta = e_i \otimes \theta^i$. Thus we have

$$D\theta = D(e_i \otimes \theta^i) = De_i \wedge \theta^i(e) + e_i \wedge d\theta^i(e) = (e_j \otimes^j_i) \wedge \theta^i(e) + e_i \wedge d\theta^i(e),$$

or in coordinates:

$$(D\theta)^i = \omega^i_j \wedge \theta^j(e) + d\theta^i(e). \tag{56.293}$$

Notice that it provides the formula

$$T = d_\omega \theta \tag{56.294}$$

for the torsion as exterior covariant derivative of the connection form.

56.10.6 Example: Levi-Civita

We consider the vector bundle $E = TM$ and the local basis $e_i = \partial_i$. An exterior derivative in this case is a map $D\colon \Gamma(TM) \to \Gamma\Big(TM \otimes \Omega^1 M\Big)$. In that particular case, we denote by $\nabla_X Y$ the vector field $D(Y)X$, and it is computed by first writing $D(X)_x = \sum_i Z_x^i \otimes \omega_x^i$ with $Z^i \in \Gamma(TM)$ and $\omega^i \in \Omega^1(M)$. The we have

$$D(X)_x Y_x = \omega + x^i(Y_x) Z_x^i. \tag{56.295}$$

A good choice of soldering form is $\theta_x = \mathrm{Id}$ for every $x \in M$, or $\theta(X) = X$. In coordinates, that soldering form is given by $\theta^i(\partial_j) = \delta^i_j$. The **Christoffel symbols** are defined by

$$\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k, \tag{56.296}$$

and the covariant derivative reads

$$\nabla_X Y = \nabla_{X^i \partial_i}(Y^j \partial_j) = X^i \Big((\partial_i Y^j)\partial_j + Y^j \nabla_{\partial_i} \partial_j \Big) = \Big(X(Y^k) + X^i Y^j \Gamma^k_{ij} \Big) \partial_k. \tag{56.297}$$ EqCovDerGamChr

We can determine the Christoffel symbols in function of the connection form using the fact that on the one hand, $\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k$, and on the other hand,

$$\nabla_{\partial_i} \partial_j = D(\partial_j)(\partial_i) = \partial_k \otimes \omega^k_j(\omega_i),$$

so that

$$\Gamma^k_{ij} = \omega^k_j(\partial_i) \tag{56.298}$$

Now we can get the same result as equation (56.297) using the exterior derivative formalism. First we have $DY = \partial_i \otimes dY^i + \partial_i \otimes X^j \omega^i_j$, so that

$$(DY)X = \partial_i \otimes dY^i(X) = \partial_i \otimes X^j \omega^i_j(X^k \partial_k),$$

in which we use the relation $\omega^i_j(X^k \partial_k) = X^k \omega^i_j(\partial_k) = X^k \Gamma^i_{jk}$ to get

$$(DY)X = \big(X(Y^i) + X^j X^k \Gamma^i_{jk} \big) \partial_i.$$

Notice that the anti-symmetric part of Γ with respect to its two lower indices does not influence the covariant derivative. Let us compute the torsion in terms of Γ. For that remark that $d\theta^i = 0$ because

$$(d\theta^i)(X, Y) = X\theta^i(Y) - Y\theta^i(X) - \theta^i\big([X, Y]\big) = X(Y^i) - Y(X^i) - [X, Y]^i = 0.$$

Thus we have

$$\begin{aligned}
(D\theta)(\partial_k \otimes \partial_l) &= \big((D\partial_i)\partial_k\big)\theta^i(\partial_l) - \big((D\partial_i)\partial_l\big)\theta^i(\partial_k) \\
&= \delta^i_l \Gamma^j_{ik} \partial_j - \delta^i_k \Gamma^j_{il} \partial_j \\
&= (\Gamma^j_{lk} - \Gamma^j_{kl}) \partial_j.
\end{aligned}$$

The connection ∇ is moreover compatible with the metric because

$$\nabla_Z\big(g(X, Y)\big) = Z\big(\eta(eX, eY)\big) = \eta\big(\underbrace{D_Z(eX)}_{=e(\nabla_Z X)}, eY\big) + \eta\big(eX, D_Z(eY)\big) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y).$$

56.11 Connection on principal bundle

56.11.1 First definition: 1-form

pg__connpriic

We consider a G-principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \downarrow{\scriptstyle \pi} \\ & & M \end{array}$$

and $\mathcal{G}$, the Lie algebra of G.

Definition 56.73.
*A **connection** on P is a 1-form $\omega \in \Omega(P, \mathcal{G})$ which fulfills*

pg:def:conne

— $\omega_\xi(A^*_\xi) = A$,
— $(R_g^*\omega)_\xi(\Sigma) = \mathrm{Ad}(g^{-1})(\omega_\xi(\Sigma))$,

for all $A \in \mathcal{G}$, $g \in G$, $\xi \in P$ and $\Sigma \in T_\xi P$. defconnform

Here, R_g is the right action: $R_g\xi = \xi \cdot g$ and A^* stands for the **fundamental field** associated with A for the action of G on P:

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0}, \tag{56.299}$$

defastar

For each $\xi \in P$, we have $\omega_\xi : T_\xi P \to \mathcal{G}$. See section 53.3.

If α is a connection 1-form on P, we say that Σ is an **horizontal** vector field if $\alpha_\xi(\Sigma) = 0$ for all $\xi \in P$. If $X_x \in T_xM$ and $\xi \in \pi^{-1}(x)$, there exists an unique [19] Σ in $T_\xi P$ which is horizontal and such that $\pi_*(\Sigma) = X_x$. This Σ is called the **horizontal lift** of X_x. We can also pointwise construct the horizontal lift of a vector field. The one of X is often denoted by $\overline{X}$; it is an element of $\mathfrak{X}(P)$.

56.11.2 Second definition: horizontal space

For each $\xi \in P$, we define the **vertical space** $V_\xi P$ as the subspace of $T_\xi P$ whose vectors are tangent to the fibers: each $v \in V_\xi P$ fulfills $d\pi v = 0$. Any such vector is given by a path contained in the fiber of ξ. So, $v \in V_\xi P$ if and only if there exists a path $g(t) \in G$ such that $v = \frac{d}{dt}\Big[\xi \cdot g(t)\Big]_{t=0}$.

A **connection** Γ is a choice, for each $\xi \in P$, of an **horizontal space** $H_\xi P$ such that

— $T_\xi P = V_\xi P \oplus H_\xi P$,
— $H_{\xi \cdot g} = (dR_g)_\xi H_\xi$,
— $H_\xi P$ depends on ξ under a differentiable way.

The second condition means that the distribution $\xi \to H_\xi$ is invariant under G. Thanks to the first one, for each $X \in T_\xi P$, there exists only one choice of $Y \in H_\xi P$ and $Z \in V_\xi P$ such that $X = Y + Z$. These are denoted by vX and hX and are naturally named *horizontal* and *vertical components* of X. The third condition means that if X is a differentiable vector field on P, then vX and hX are also differentiable vector fields. We will often write V_ξ and H_ξ instead of $V_\xi P$ and $V_\xi P$.

The word *connection* probably comes from the fact that the horizontal space gives a way to jump from a fiber to the next one. When we consider a connection Γ, we can define a $\mathcal{G}$-valued connection 1-form by

$$\omega(X)^*_\xi = vX_\xi.$$

The existence is explained in section 53.3. It is clear that $\omega(X) = 0$ if and only if X is horizontal. The theorem which connects the two definitions is the following.

Theorem 56.74.
If Γ is a connection on a G-principal bundle, and ω is its 1-form, then

19. See [838], chapter II, proposition 1.2.

enuyai

(1) *for any* $A \in \mathcal{G}$, *we have* $\omega(A^*) = A$,

enuyaii

(2) $(R_g)^*\omega = \mathrm{Ad}(g^{-1})\omega$, *i.e. for any* $X \in T_\xi P$, $g \in G$ *and* $\xi \in M$,

$$\omega((dR_g)_\xi X) = \mathrm{Ad}(g^{-1})\omega_\xi(X)$$

Conversely, if one has a $\mathcal{G}$*-valued* 1*-form on* P *which fulfills these two requirement, then one has one and only one connection on* P *whose associated* 1*-form is* ω.

Proof. (1) The definition of ω is $\omega(X)^*_\xi = vX$. Then $\omega(A^*)^*_\xi = vA^*_\xi = A^*_\xi$ because A^* is vertical. From lemma 53.20, $\omega(A^*) = A$.

(2) Let $X \in \mathfrak{X}(P)$. If X is horizontal, the definition of a connection makes $dR_d X$ also horizontal, then the claim becomes $0 = 0$ which is true. If X is vertical, there exists a $A \in \mathcal{G}$ such that $X = A^*$ and a lemma shows that $dR_g X$ is then the fundamental field of $\mathrm{Ad}(g^{-1})A$. Using the properties of a connection,

$$(R_g^*\omega)_\xi(X) = \omega_{\xi \cdot g}(dR_g X) = \mathrm{Ad}(g^{-1})A = \mathrm{Ad}(g^{-1})\omega_\xi(X). \tag{56.300}$$

Now we turn our attention to the inverse sense: we consider a 1-form which fulfills the two conditions and we define

$$H_\xi = \{X \in T_\xi P \text{ tel que } \omega(X) = 0\}. \tag{56.301}$$

We are going to show that this prescription is a connection. First consider a $X \in V_\xi$, then $X = A^*$ and $\omega(X) = A$. So $H_\xi \cap V_\xi = 0$. Now we consider $X \in T_\xi P$ and we decompose it as

$$X = A^* + (X - A^*)$$

where A^* is the vertical component of X. If $\omega(dR_g X) = 0$ for all $g \in G$, then $\omega(X) = 0$, then a vector $X \in H_\xi$ fulfills at most $\dim G$ independent constraints $\omega(dR_g X) = 0$ and $\dim H_\xi$ is at least $\dim P - \dim G$. On the other hand, $\dim V_\xi = \dim G$; then

$$\dim V_\xi + \dim H_\xi \geqslant \dim G + \dim P - \dim G.$$

Then the equality must holds and $V_\xi \oplus H_\xi = T_\xi P$.

We have now to prove that ω is the connection form of H_ξ, i.e. that $\omega(X)$ is the unique $A \in \mathcal{G}$ such that A^*_ξ is the vertical component of X. Indeed if $X \in T_\xi P$, it can be decomposed as into $A^* \in V_\xi$ and $Y \in H_\xi$ and

$$\omega(X) = \omega(A^* + Y) = \omega(A^*) = A.$$

It remains to be proved that the horizontal space H_ξ of any connection Γ is related to the corresponding 1-form ω by $H_\xi = \{X \in T_\xi P \text{ tel que } \omega_\xi(X) = 0\}$. From the connection Γ, the 1-form is defined by the requirement that $\omega(X)^*_\xi = vX_\xi$. For $X \in H_\xi$, it is clear that $vX = 0$, so that $\omega(X)^* = 0$. This implies $\omega(X) = 0$ because we suppose that the action of G is effective. □

The projection $\pi\colon P \to M$ induces a linear map $d\pi\colon T_\xi P \to T_x M$. We will see that, when a connection is given, it is an isomorphism between H_ξ and $T_x M$ (if $x = \pi(\xi)$). The **horizontal lift** of $X \in \mathfrak{X}(M)$ is the unique horizontal vector field (i.e. it is pointwise horizontal) such that $d\pi(\overline{X}_\xi) = X_{\pi(\xi)}$. The proposition which allows this definition is the following.

Proposition 56.75.
For a given connection on the G*-principal bundle* P *and a vector field* X *on* M, *there exists an unique horizontal lift of* X. *Moreover, for any* $g \in G$, *the horizontal lift is invariant under* dR_g.

The inverse implication is also true: any horizontal field on P *which is invariant under* dR_g *for all* g *is the horizontal lift of a vector field on* M.

This proposition comes from [838], chapter II, proposition 1.2.

Proof. We consider the restriction $d\pi\colon H_\xi \to T_{\pi(\xi)}M$. It is injective because $d\pi(X-Y)$ vanishes only when $X-Y$ is vertical or zero. Then it is zero. It is cleat that $d\pi\colon T_\xi P \to T_{\pi(\xi)}M$ is surjective. But $d\pi X = 0$ if X is vertical, then $d\pi$ is surjective from only H_ξ.

So we have existence and unicity of an horizontal lift. Now we turn our attention to the invariance. The vector $dR_g\overline{X}_\xi$ is a vector at $\xi\cdot g$. From the definition of a connection, $dR_g H_\xi = H_{\xi\cdot g}$, then $dR_g\overline{X}_\xi$ is the unique horizontal vector at $\xi\cdot g$ which is sent to X_x by $d\pi$. Thus it is $\overline{X}_{\xi\cdot g}$.

For the inverse sense, we consider $\overline{X}$, an horizontal invariant vector field on P. If $x\in M$, we choose $\xi\in\pi^{-1}(x)$ and we define $X_x = d\pi(\overline{X}_\xi)$. This construction is independent of the choice of ξ because for $\xi' = \xi\cdot g$, we have

$$d\pi(\overline{X}_{\xi'}) = \pi(dR_g\overline{X}_\xi) = \pi(\overline{X}_\xi).$$

□

An other way to see the invariance is the following formula:

$$\overline{X}_{\xi\cdot g} = (dR_g)_\xi\overline{X}_\xi.$$

By definition, $\overline{X}_{\xi\cdot g}$ is the unique vector of $T_{\xi\cdot g}P$ which fulfils $d\pi\overline{X}_{\xi\cdot g} = X_x$ if $\xi\pi^{-1}(x)$, so the following computation proves the formula:

$$(d\pi)_{\xi\cdot g}((dR_g)_\xi\overline{X}_\xi) = d(\pi\circ R_g)_\xi\overline{X}_\xi = d\pi_\xi\overline{X}_\xi = X_x. \tag{56.302}$$

56.11.3 Curvature

The curvature of a vector or associated bundle satisfies $\Omega_\alpha = d\theta_\alpha + \theta_\alpha\wedge\theta_\alpha$. So we naturally define the **curvature** of the connection ω on a principal bundle as the $\mathcal{G}$-valued 2-form

$$\Omega = d\omega + \omega\wedge\omega. \tag{56.303}$$

When we consider a local section $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ on $\mathcal{U}_\alpha\subset M$, we can express the curvature with a 2-form on M instead of P by the formula

$$F_{(\alpha)} = \sigma_\alpha^*\Omega,$$

or, more explicitly, by $F_{(\alpha)x}(X,Y) = \Omega_{\sigma_\alpha(x)}(d\sigma_\alpha X, d\sigma_\alpha Y))$. Note that if $\mathcal{G}$ is abelian, $\Omega = d\omega$ and $d\Omega = 0$.

56.12 Exterior covariant derivative and Bianchi identity

Let $\omega\in\Omega^1(P,\mathcal{G})$ be a connection 1-form on the G-principal bundle P. Using the operation $[.\wedge.]$ defined in section 56.5, we define the **exterior covariant derivative** by

$$d_\omega\alpha = d\alpha + \frac{1}{2}[\omega\wedge\alpha] \qquad \text{when } \alpha\in\Omega^1(P,\mathcal{G}), \tag{56.304}$$

$$d_\omega\beta = d\beta + [\omega\wedge\beta] \qquad \text{when } \beta\in\Omega^2(P,\mathcal{G}), \tag{56.305}$$

The **curvature** is the 2-form defined by

$$\Omega = d_\omega\omega = d\omega + \omega\wedge\omega \tag{56.306}$$

where d_ω is the exterior covariant derivative associated with the connection form ω, and the wedge has to be understood as in equation (56.198).

Proposition 56.76.
The curvature form satisfies the identity

$$d_\omega\Omega = 0 \tag{56.307}$$

which is the Bianchi identity

Proof. taking the differential of $\Omega = d\omega + \omega \wedge \omega$, we find

$$d\Omega = d^2\omega + d\omega \wedge \omega - \omega \wedge d\omega$$

in which $d^2\omega = 0$ and we replace $d\omega$ by $\Omega - \omega \wedge \omega$, so that

$$d\Omega = \Omega \wedge \omega - \omega \wedge \Omega,$$

which becomes the Bianchi identity using the definition of d_ω and the notation (56.200). □

Remark that the Bianchi identity reads $d_\omega^2\omega = 0$, but that in general d_ω does not square to zero.

56.13 Covariant derivative on associated bundle

Now we consider a general G-principal bundle $\pi\colon P \to M$ and an associated bundle $E = P \times_\rho V$. We define a product $\mathbb{R} \times E \to E$ by

$$\lambda[\xi, v] = [\xi, \lambda v]. \tag{56.308}$$

eq:def:REE

It is clear that the equivariant function $\widehat{\lambda\psi}$ defines the section $\lambda\psi$. A **covariant derivative** is a map

$$\begin{aligned} \nabla\colon \mathfrak{X}(M) \times \Gamma(M,E) &\to \Gamma(M,E) \\ (X,\psi) &\mapsto \nabla_X\psi \end{aligned} \tag{56.309}$$

such that

$$\nabla_{fX}\psi = f\nabla_X\psi, \tag{56.310a}$$

$$\nabla_X(f\psi) = (X \cdot f)\psi + f\nabla_X\psi \tag{56.310b}$$

eq:def:der_covii

where products have to be understood by formula (56.308).

Theorem 56.77.
A connection on a principal bundle gives rise to a covariant derivative on any associated bundle by the formula

$$\widehat{\nabla^E_X\psi}(\xi) = \overline{X}_\xi(\hat{\psi}) \tag{56.311}$$

where $\hat{\psi}\colon P \to V$ is the function associated with the section $\psi\colon M \to E$.

tho_dercovassoequiv

We have to prove that it is a good definition: the function $\widehat{\nabla^E_X\psi}$ must define a section $\nabla^R_X\psi\colon M \to E$ and the association $\psi \to \nabla^E_X\psi$ must be a covariant derivative.

With the discussion of page 3006 about the application of a tangent vector on a map between manifolds, we have $(d\varphi X)f = X(f \circ \varphi)$. By using this equality in the case of $\overline{X}$ with $\hat{\psi}$ and R_g, we find $(dR_g\overline{X})(\hat{\psi}) = \overline{X}(\hat{\psi} \circ R_g)$ and thus

$$\overline{X}_{\xi \cdot g}(\hat{\psi}) = \overline{X}_\xi(dR_g\hat{\psi}).$$

We prove the theorem step by step.

Proposition 56.78.
The function $\widehat{\nabla^E_X\psi}$ defines a section of P.

Proof. We have to see that $\widehat{\nabla^E_X\psi}$ is an equivariant function. The equivariance of $\hat{\psi}$ gives $\hat{\psi} \circ R_g = \rho(g^{-1})\hat{\psi}$, thus

$$\widehat{\nabla^E_X\psi}(\xi \cdot g) = \overline{X}_{\xi \cdot g}(\hat{\psi}) = \big((dR_g)_\xi \overline{X}_\xi\big)(\hat{\psi}) = \overline{X}_\xi(\hat{\psi} \circ R_g) = \overline{X}_\xi(\rho(g^{-1})\hat{\psi}) = \rho(g^{-1})\overline{X}_\xi(\hat{\psi}). \tag{56.312}$$

The last equality comes from the fact that the product $\rho(g^{-1})\hat{\psi}$ is a linear product "matrix times vector" and that $\overline{X}_\xi$ is linear. □

Theorem 56.79.
The definition

$$\widehat{\nabla^E_X\psi}(\xi) = \overline{X}_\xi(\hat{\psi})$$

defines a covariant derivative. tho_nablaE

Proof. We have to check the two conditions given on page 3392.
First condition. By definition, $\widehat{\nabla^E_{fX}\psi}(\xi) = \overline{fX_\xi}(\hat{\psi})$. Now we prove that

$$\overline{fX_\xi}(\hat{\psi}) = (f\circ\pi)(\xi)\overline{X}_\xi(\hat{\psi}). \tag{56.313}$$ eq:fXhpsi

This formula is coherent because $\overline{X}_\xi(\hat{\psi}) \in V$ and $(f\circ\pi)(\xi) \in \mathbb{R}$. By definition of the horizontal lift, $\overline{fX}_\xi$ is the unique vector such that

— $d\pi_\xi(\overline{fX}_\xi) = (fX)_x = f(x)d\pi\overline{X}_\xi = (f\circ\pi)(\xi)d\pi\overline{X}_\xi$,
— $\omega_\xi(\overline{fX}_\xi) = 0$.

We check that $(f\circ\pi)(\xi)\overline{X}_\xi$ also fulfills these two conditions because $d\pi$ and ω are $C^\infty(P)$-linear. Equation (56.313) immediately gives

$$\widehat{\nabla^E_{fX}\psi}(\xi) = (f\circ\pi)(\xi)\widehat{\nabla^E_X\psi}(\xi). \tag{56.314}$$

Now we show that $\widehat{f\nabla^E_X\psi}$ is the same. The section $f\nabla^E_X\psi\colon M\to E$ is given by $(f\nabla^E_X\psi)(x) = f(x)(\nabla^E_X\psi)(x)$, and by definition of the associated equivariant function,

$$f(x)(\nabla^E_X\psi)(x) = [\xi, f(x)\widehat{\nabla^E_X\psi}(\xi)].$$

Then

$$\widehat{f\nabla^E_X\psi}(\xi) = f(x)\widehat{\nabla^E_X\psi}(\xi) = (f\circ\pi)(\xi)\widehat{\nabla^E_X\psi}(\xi). \tag{56.315}$$

All this shows that $\nabla^E_{fX}\psi = f\nabla^E_X\psi$.
Second condition. This is a computation using the Leibniz rule:

$$\begin{aligned}
\widehat{\nabla^E_X(f\psi)}(\xi) &= \overline{X}_\xi(\widehat{f\psi}) \overset{(a)}{=} \overline{X}_\xi((\pi\circ f)\hat{\psi}) \\
&\overset{(b)}{=} \overline{X}_\xi(\pi^*f)\hat{\psi}(\xi) + (\pi^*f)(\xi)\overline{X}_\xi\hat{\psi} = d(f\circ\pi)_\xi\overline{X}_\xi\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) \\
&= df_{\pi(\xi)}d\pi_\xi\overline{X}_\xi\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) = X_x(f)\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) \\
&= \widehat{(Xf)\psi}(\xi) + f\widehat{\nabla^E_X\psi}(\xi)
\end{aligned} \tag{56.316}$$

where (a) is because $\widehat{f\psi} = \pi^*f\hat{\psi}$, and (b) is an application of the Leibniz rule. □

Theorem 56.80.
Using the local coordinates related to the sections $\sigma_\alpha\colon \mathcal{U}_\alpha\to P$, the covariant derivatives reads:

$$(\nabla_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} - \rho_*(\sigma_\alpha^*\omega_x(X))\psi_{(\alpha)}(x) \tag{56.317}$$ eq:nabla_coord

where $\rho_*\colon\mathcal{G}\to\mathrm{End}(V)$ is defined by

$$\rho_*(A) = \frac{d}{dt}\Big[\rho(e^{tA})\Big]_{t=0} \tag{56.318}$$ eq:def_rho_s

Proof. The problem reduces to the search of $\overline{X}$ because

$$(\nabla_X\psi)_{(\alpha)}(x) = \widehat{\nabla_X\psi}(\sigma_\alpha(x)) = \overline{X}_{\sigma_\alpha(x)}(\hat{\psi}).$$

We claim that $\overline{X}_{\sigma_\alpha(x)} = d\sigma_\alpha X_x - \omega(d\sigma_\alpha X_x)^*$. We have to check that $d\pi\overline{X} = X$ and $\omega(\overline{X}) = 0$. The latter comes easily from the fact that $\omega(A^*) = A$. For the first one, remark that s_α is a section, then $d(\pi \circ s_\alpha) = \mathrm{Id}$, and $d\pi(ds_\alpha X_x) = X_x$, while

$$d\pi(A^*_\xi) = d\pi \frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0} = \frac{d}{dt}\Big[\pi(\xi)\Big]_{t=0} = 0. \tag{56.319}$$

Since the horizontal lift is unique, we deduce

$$(\nabla_X\psi)_{(\alpha)}(x) = \big(d\sigma_\alpha X_x - \omega(d\sigma_\alpha X_x)^*\big)\hat{\psi}. \tag{56.320}$$

From the definition of a fundamental vector field,

$$\begin{aligned}
\omega(d\sigma_\alpha X_x)^*_{\sigma_\alpha(x)}\hat{\psi} &= \frac{d}{dt}\Big[\hat{\psi}\big(\sigma_\alpha(x) \cdot e^{-t\omega(d\sigma_\alpha X_x)}\big)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\rho(e^{t\omega(d\sigma_\alpha X_x)})\hat{\psi}(\sigma_\alpha(x))\Big]_{t=0} \qquad \text{from (56.234)} \\
&= (d\rho)_e(\omega \circ d\sigma_\alpha)X_x(\hat{\psi} \circ \sigma_\alpha)(x) \\
&= \rho_*\big((\sigma_\alpha^*\omega)(X_x)\big)\psi_{(\alpha)}(x) \qquad \text{by (56.318)}
\end{aligned} \tag{56.321}$$

□

We can express the covariant derivative by means of some maps $\theta_\alpha\colon \mathfrak{X}(M) \times M \to \mathrm{End}(V)$ given by

$$\nabla_X\gamma_{\alpha i} = \theta_\alpha(X)_i{}^j\gamma_{\alpha j}. \tag{56.322}$$

where the $\gamma_{\alpha i}$'s were given in equation (56.241). By the definition (56.310b),

$$\begin{aligned}
(\nabla_X\psi)(x) &= (X \cdot s^i_\alpha)_x\gamma_{\alpha i}(x) + s^i_\alpha(x)(\nabla_X\gamma_{\alpha i})(x) \\
&= (X \cdot s^i_\alpha)_x\gamma_{\alpha i}(x) + s^i_\alpha(x)\theta_\alpha(X)_i{}^j\gamma_{\alpha j}(x).
\end{aligned}$$

On the othre hand with the notations of equation (56.239), $\gamma_{\alpha j} = e_i$ and $X_x\gamma_{\alpha j} = 0$. Then equation (56.317) gives $\theta_\alpha(X) = \rho_*(\sigma^*_\alpha\omega_x(X))$, or

$$\theta_\alpha = \rho_*(\sigma^*_\alpha\omega_x). \tag{56.323}$$

56.13.1 Curvature on associated bundle

From the definition (56.235), it makes sense to define the curvature 2-form by

$$R(X,Y)\psi = \nabla_X\nabla_Y\psi - \nabla_Y\nabla_X\psi - \nabla_{[X,Y]}\psi.$$

It is also clear that $\psi_{(\alpha)}$ defines a section of the trivial vector bundle $F = M \times V$ by $x \to (x, \psi_{(\alpha)}(x))$, so one can define $\Omega_\alpha(X,Y)\colon \Gamma(M,E) \to \Gamma(M,E)$ by

$$\big(R(X,Y)\psi\big)_{(\alpha)} = \Omega_\alpha(X,Y)\psi_{(\alpha)}$$

and take back all the work around Bianchi because of the relation (56.317) which can be written as $(\nabla_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} + \theta_\alpha(X)\psi_{(\alpha)}(x)$ and which is the same as in proposition 56.67.

56.13.2 Connection on frame bundle

The frame bundle was defined at page 3383. Let $F \xrightarrow{p} M$ be a $\mathbb{K}$-vector bundle with some local trivialization $(\mathcal{U}_\alpha, \phi^E_\alpha)$ and the corresponding transition functions $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(r,\mathbb{K})$. We consider $\pi\colon P \to M$, the frame bundle of F; it is a $GL(r,\mathbb{K})$-principal bundle. Let ∇ be a covariant derivative on F and θ_α, the associated matrices 1-form. The frame bundle is

$$P = \bigcup_{x\in M} (\text{frame of } F_x).$$

A connection is a $\mathcal{G}$-valued 1-form; in our case it is a map

$$\omega_\xi^\alpha \colon T_\xi\big(\pi^{-1}(\mathcal{U}_\alpha)\big) \to \mathfrak{gl}(r,\mathbb{K}).$$

We define our connection by, for $g \in GL(r,\mathbb{K})$, $x \in \mathcal{U}_\alpha$, $X_x \in T_xM$ and $A \in \mathfrak{gl}(r,\mathbb{K})$,

$$\omega^\alpha_{S_\alpha(x)\cdot g}\big(R_{g*}s_\alpha(x)_*X_x + A^*_{S_\alpha(x)\cdot g}\big) := A + \mathrm{Ad}(g^{-1})\theta_\alpha(X_x). \tag{56.324}$$

eq:def_omega_frame

where $S_\alpha \colon \mathcal{U}_\alpha \to P$ is the section defined by the trivialization ϕ^P_α:

$$S_\alpha(x) = \{\overline{v}_\alpha = \phi^{E\,-1}_\alpha(x, e_i)\}_{i=1,\dots,r}.$$

Since $\theta_\alpha(X_x) \in \mathrm{End}(\mathbb{K}^r) \subset \mathfrak{gl}(r,\mathbb{K})$, the second term of (56.324) makes sense. This formula is a good definition of ω because of the following lemma:

Lemma 56.81.
If $\xi = S_\alpha(x)\cdot g$ and $\Sigma \in T_\xi P$, there exists a choice of $A \in \mathcal{G}$, and $X_x \in T_xM$ such that

$$\Sigma = R_{g*}s_\alpha(X)_*X_x + A^*_{S_\alpha(x)\cdot g}. \tag{56.325}$$

eq:geneSigma

Proof. If $\xi \in P$ is a basis of E at y, there exists only one choice of $x \in M$ and $g \in G$ such that $\xi = S_\alpha(x)\cdot g$.

Let us consider a general path $c \colon \mathbb{R} \to P$ under the form $c(t) = s_\alpha(x(t))\cdot g(t)$ where x and g are path in M and $GL(r,\mathbb{K})$. The frame $c(t)$ is the one of $F_{x(t)}$ obtained by the transformation $g(t)$ from $s_\alpha(x(t))$. It is a set of r vectors, and each of them can be written as a combination of the vectors of $s_\alpha(x(t))$, so we write

$$c^i(t) = s^j_\alpha(x(t))g^i_j(t) \tag{56.326}$$

where $s^j_\alpha(x(t)) \in F_{x(t)}$ and $g^i_j(t) \in \mathbb{K}$. We compute $\Sigma = c'(0)$ by using the Leibniz rule and we denote $x'(0) = X_x$, $x(0) = x$ and $g^i_j(0) = g^i_j$ (the matrix of g):

$$\begin{aligned}\Sigma^i &= \frac{d}{dt}\Big[s^j_\alpha(x(t))\Big]_{t=0}g^i_j + s^j_\alpha(x)\frac{d}{dt}\Big[g^i_j(t)\Big]_{t=0}\\ &= (ds^j_\alpha)_x X_x g^i_j + g^{i\prime}_j(0)s^j_\alpha(x).\end{aligned} \tag{56.327}$$

Going to more compact matrix form, it gives

$$\Sigma = (ds_\alpha)_x X_x \cdot g + s_\alpha(x)g'(0).$$

The second term, $s^j_\alpha(x)g'^i_j(0)$, is a general vector tangent to a fiber. So it can be written as a fundamental field A^*_ξ.

□

Lemma 56.82.
On $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$, the form fulfills $\omega^\alpha = \omega^\beta$.

Proof. Let $\gamma \colon \mathbb{R} \to M$ be a path whose derivative is X_x. Then

$$\begin{aligned}(R_g)_*s_\alpha(x)_*X_x &= \frac{d}{dt}\Big[s_\alpha(\gamma_t)\cdot g\Big]_{t=0} = \frac{d}{dt}\Big[s_\beta(\gamma_t)g_{\alpha\beta}(\gamma_t)\cdot g\Big]_{t=0}\\ &= \frac{d}{dt}\Big[s_\alpha(\gamma_t)g_{\alpha\beta}(x)\cdot g\Big]_{t=0} + \frac{d}{dt}\Big[s_\beta(x)\cdot g_{\alpha\beta}(\gamma_t)\cdot g\Big]_{t=0}.\end{aligned} \tag{56.328}$$

What is in the derivative of the first term is $R_{g_{\alpha\beta}(x)g}(s_\beta(\gamma_t))$. Taking the derivative, we find the expected $R_{g_{\alpha\beta}(x)g_*}s_{\beta*}X_x$.

For the second term, we note $r := s_\beta(x) \cdot g_{\alpha\beta}(g)g$, and we have to compute the following, using equation (56.202),

$$\begin{aligned}\frac{d}{dt}\Big[r \cdot \mathrm{Ad}_{g^{-1}}(g_{\alpha\beta}^{-1}(x)g_{\alpha\beta}(\gamma_t))\Big]_{t=0}&\\ = \frac{d}{dt}\Big[r \cdot \exp t\big((d\,\mathbf{Ad}_{g^{-1}})_e(g_{\alpha\beta}^{-1}(x)(dg_{\alpha\beta})_x X_x)\big)\Big]_{t=0}&\\ = \frac{d}{dt}\Big[r \cdot \exp t\big(\mathrm{Ad}_{g^{-1}}\, g_{\alpha\beta}^{-1}(x)dg_{\alpha\beta}\Big]_{t=0}&\\ = \Big(\mathrm{Ad}_{g^{-1}}\, g_{\alpha\beta}^{-1}(x)dg_{\alpha\beta}X_x\Big)_r^*.&\end{aligned} \tag{56.329}$$

Using this, we can perform the computation:

$$\begin{aligned}\omega^\beta_{S_\alpha(x)\cdot g}\big(R_{g*}s_\alpha(x)_*X_x + A^*_{S_\alpha(x)\cdot g}\big) &= \omega^\beta_{S_\beta(x)\cdot g_{\alpha\beta}(x)g}\Big(R_{g_{\alpha\beta}}(x)g_*s_\beta(x)_*X_x\\ &\qquad + (\mathrm{Ad}_{g^{-1}}\, g_{\alpha\beta}^{-1}(x)dg_{\alpha\beta}X_x)_r^* + A^*\Big)\\ &= \mathrm{Ad}_{(g_{\alpha\beta}(x)g)^{-1}}\,\theta_\beta(X_x)\\ &\quad + \mathrm{Ad}_{g^{-1}}\, g_{\alpha\beta}^{-1}(x)dg_{\alpha\beta}(X_x) + A\\ &= \mathrm{Ad}_{g^{-1}}\left((g_{\alpha\beta}^{-1}\theta_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}dg_{\alpha\beta})(X_x)\right) + A\\ &= \omega^\alpha_{S_\alpha(x)g}\big(R_{g*}s_\alpha(x)_*X_x + A^A_{S_\alpha(x)\cdot g}\big).\end{aligned} \tag{56.330}$$

□

Proposition 56.83.
The ω defined by formula (56.324) is a connection 1-form. prop_omconfrom

Proof. The first condition, $\omega(A^*_\xi) = A$, is immediate from the definition. The lemma 53.21 gives the second condition in the case $\Sigma = A^*_\xi$. It remains to be checked that $\omega(dR_g\Sigma) = \mathrm{Ad}(g^{-1})\omega(\Sigma)$ in the case $\Sigma = dR_h ds_\alpha X_x$. This is obtained using the fact that Ad is a homomorphism. □

subsectionLevi-Civita connection subsection_levi

Let (M, g) be a Riemannian manifold. We look at a connection 1-form $\alpha \in \Omega^1(\mathrm{SO}(M), so(\mathbb{R}^m))$ on $\mathrm{SO}(M)$, and we define a covariant derivative $\nabla^\alpha \colon \mathfrak{X}(M) \times T(M) \to T(M)$, where $T(M)$ is the tensor bundle on M by (cf. theorem (56.79))

$$\overline{\nabla}^\alpha_X s = \overline{X}\hat{s}, \tag{56.331}$$

for any $s \in T(M)$. Our purpose now is to prove that an automatic property of this connection is $\nabla^\alpha g = 0$. The unique such connection which is torsion-free is the **Levi-Civita** one.

The metric g is a section of the tensor bundle $T^*M \otimes T^*M$. So we have, in order to find $\hat{g}$ and to use equation (56.331), to see $T^*M \otimes T^*M$ as an associated bundle. As done in 56.7.4, we see that

$$T^*M \otimes T^*M \simeq \mathrm{SO}(M) \times_\rho (V^* \otimes V^*),$$

with the following definitions:

— The isomorphism is given by $\psi[b, \alpha \otimes \beta](X \otimes Y) = \alpha(b^{-1}X)\beta(b^{-1}Y)$,

— $\rho(A)\alpha = \alpha \circ A$,

— $b \cdot A = b \circ A$.

Here, $V = \mathbb{R}^m$; $b \colon V \to T_xM$; $\alpha, \beta \in V^*$; X, $Y \in T_xM$ and $A \in \mathrm{SO}(m)$ is seen as $A \colon V \to V$.

The following shows that ψ is well defined:

$$\begin{aligned}\psi[b \cdot A, \rho(A^{-1})\alpha \otimes \beta](X \otimes Y) &= (\alpha \circ A)(A^{-1} \circ b^{-1}X)(\beta \circ A)(A^{-1} \circ b^{-1}Y)\\ &= \psi[b, \alpha \otimes \beta](X \otimes Y)\end{aligned} \tag{56.332}$$

Proposition 56.84.
The function $\hat{g}$ is given by

$$\hat{g}(b)(v \otimes w) = g_x(b(v) \otimes b(w)) = v \cdot w.$$

Proof. The second equality is just the fact that $b\colon (\mathbb{R}^m, \cdot\,) \to (T_xM, g_x)$ is isometric. On the other hand, if $\hat{g}(b) = \alpha \otimes \beta$, we have:

$$\begin{aligned} g_x(X \otimes Y) &= \psi[b, \alpha \otimes \beta](X \otimes Y) = \alpha(b^{-1}X)\beta(b^{-1}Y) \\ &= \alpha \otimes \beta(b^{-1}X \otimes b^{-1}Y) = \hat{g}(b)(b^{-1}X \otimes b^{-1}Y). \end{aligned} \tag{56.333}$$

Since b is bijective, X and Y can be written as bv and bw respectively for some v, $w \in V$, so that

$$g_x(bv \otimes bw) = \hat{g}(b)v \otimes w.$$

□

It is now easy to see that $\overline{X}\hat{g} = 0$. As $\hat{g}$ takes its values in $V^* \otimes V^*$, $\overline{X}\hat{g}$ belongs to this space and can be applied on $v \otimes w \in V \otimes V$. Let $\overline{X}(t)$ be a path in $\mathrm{SO}(M)$ which defines $\overline{X}$; if $\overline{X} \in T_b\,\mathrm{SO}(M)$, $\overline{X}(0) = b$. We have

$$\overline{X}\hat{g}(v \otimes w) = \frac{d}{dt}\hat{g}(\overline{X}(t))v \otimes w\Big|_{t=0} = \frac{d}{dt}v \cdot w\Big|_{t=0}, \tag{56.334}$$

which is obviously zero.

56.13.3 Holonomy

Let the principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \downarrow{\scriptstyle \pi} \\ & & M \end{array} \tag{56.335}$$

and ω a connection on G. Let $\gamma\colon [0,1] \to M$, a closed curve piecewise smooth; $\gamma(0) = \gamma(1) = x$. For each $p \in \pi^{-1}(x)$, there exists one and only one horizontal lift $\tilde{\gamma}\colon [0,1] \to P$ such that $\tilde{\gamma}(0) = p$. There exists of course an element $g \in G$ such that $\tilde{\gamma}(1) = p \cdot g$.

We define the following equivariance relation on P: we say that $p \sim q$ if and only if p and q can be joined by a piecewise smooth path. The **holonomy group** at the point p is

$$\mathrm{Hol}_p(\omega) = \{g \in G \text{ tel que } p \sim p \cdot g\}.$$

56.13.4 Connection and gauge transformation

Proposition 56.85.
If ω is a connection on a G-principal bundle and φ, a gauge transformation, the form $\beta = \varphi^*\omega$ is a connection 1-form too. prop:vp_conn

Proof. It is rather easy to see that $\varphi_* A^*_\xi = A^*_{\varphi(x)}$:

$$\varphi_* A^*_\xi = \frac{d}{dt}\Big[\varphi(\xi e^{-tA})\Big]_{t=0} = \frac{d}{dt}\Big[\varphi(\xi)e^{-tA}\Big]_{t=0} = A^*_{\varphi(x)}.$$

The same kind of reasoning leads to $\varphi_* R_{g_*} = R_{g_*}\varphi_*$. From here, it is easy to see that

$$(\varphi^*\omega)_\xi(A^*_\xi) = \omega_{\varphi(\xi)}(\varphi_* A^*_\xi) = A,$$

and

$$\big(R^*_g(\varphi^*\omega)_\xi\big)(\Sigma) = (R^*_g\omega)_{\varphi(\xi)}(\varphi_*\Omega) = \mathrm{Ad}(g^{-1})\big((\varphi^*\omega)_\xi(\Sigma)\big).$$

□

So, the "gauge transformed" of a connection is still a connection. It is hopeful in order to define gauge invariants objects (Lagrangian) from connections (electromagnetic fields).

56.13.4.1 Local description

Let $\pi\colon P \to M$ be a G-principal bundle given with some trivializations $\phi_\alpha^P\colon \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G$ over $\mathcal{U}_\alpha \subset M$ and $s_\alpha\colon \mathcal{U}_\alpha \to \pi^{-1}(\mathcal{U}_\alpha)$, a section. In front of that, we consider an associated bundle $p\colon E = P \times_\rho V \to M$ with a trivialization $\phi_\alpha^E\colon E \to \mathcal{U}_\alpha \times V$. One can choose a section s_α compatible with the trivialization in the sense that $\phi_\alpha^P(s_\alpha(x) \cdot g) = (x, g)$; the same can be done with E by choosing $\phi_\alpha^E([s_\alpha(x), v]) = (x, v)$.

A section $\psi\colon M \to E$ is described by a function $\psi_\alpha\colon \mathcal{U}_\alpha \to V$ defined by $\phi_\alpha^E(\psi(x)) = (x, \psi_\alpha(x))$. In the inverse sense, ψ is defined (on $\mathcal{U}_\alpha$) from ψ_α by $\psi(x) = [s_\alpha(x), \psi_\alpha(x)]$. In the same way, a gauge transformation $\varphi\colon P \to P$ is described by functions $\tilde{\varphi}_\alpha\colon \mathcal{U}_\alpha \to G$,

$$\varphi(s_\alpha(x)) = s_\alpha(x) \cdot \tilde{\varphi}_\alpha(x). \tag{56.336}$$

The function $\tilde{\varphi}_\alpha$ also fulfils

$$(\phi_\alpha^P \circ \varphi \circ \phi_\alpha^{P^{-1}})(x, g) = (x, \tilde{\varphi}(x) \cdot g) \tag{56.337}$$

because

$$\begin{aligned}(\phi_\alpha^P \circ \varphi \circ \phi_\alpha^{P^{-1}})(x, g) &= (\phi_\alpha^P \circ \varphi)(s_\alpha(x) \cdot g)\\ &= \phi_\alpha^P(\varphi(s_\alpha(x)) \cdot g)\\ &= \phi_\alpha^P(s_\alpha(c) \cdot \tilde{\varphi}_\alpha(x)g)\\ &= (x, \tilde{\varphi}_\alpha(x)g).\end{aligned} \tag{56.338}$$

We know that a connection on P is given by its 1-form ω. Moreover we have the following:

Proposition 56.86.
A connection on P is completely determined on $\pi^{-1}(\mathcal{U}_\alpha)$ from the data of the $\mathcal{G}$-valued 1-form $\sigma_\alpha^*\omega$ on $\mathcal{U}_\alpha$.

Proof. We consider a 1-form ω which fulfils the two conditions of page 3400. Our purpose is to find back $\omega_\xi(\Sigma)$, $\forall \xi \in P, \Sigma \in T_\xi P$ from the data of $\sigma_\alpha^*\omega$ alone. For any ξ, there exists a g such that $\xi = \sigma_\alpha(x) \cdot g$. We have

$$\mathrm{Ad}_{g^{-1}}(\omega_{\sigma_\alpha(x)\Sigma}) = (R_g^*\omega)_{\sigma_\alpha(x)}(\Sigma) = \omega_{\sigma_\alpha(x)} \cdot {}_g\big((dR_g)_{\sigma_\alpha(x)}\Sigma\big). \tag{56.339}$$

If we know $s_\alpha^*\omega$, then we know $\omega\big((ds_\alpha)_x v\big)$ for any $v \in T_x M$. So

$$\omega_{\sigma_\alpha(x)} \cdot {}_g\big((dR_g)_{\sigma_\alpha(x)}\Sigma\big)$$

is given from $\sigma_\alpha^*\omega$ for every Σ of the form $\Sigma = (d\sigma_\alpha)_x v$. From the form (56.325) of a vector in $T_\xi P$, it just remains to express $\omega_{\sigma_\alpha(x)} \cdot {}_g(A^*_{\sigma_\alpha(x)} \cdot {}_g)$ in terms of s_α^*. The definition of a connection makes that it is simply A.

□

56.13.4.2 Covariant derivative

If we have a connection on P, we can define a covariant derivative on the associated bundle E by

$$(\nabla_X \psi)_{(\alpha)}(x) = X_x(\psi_\alpha) + \rho_*(s_\alpha^*\omega_x(X))\psi_{(\alpha)}(x),$$

the matricial 1-form being given by $\theta_\alpha = \rho_*\sigma_\alpha^*\omega$. The gauge transformation φ acts on the connection ω by defining $\omega^\varphi := \varphi^*\omega$.

Proposition 56.87.
If $\beta = \varphi^*\omega$, then

$$s_\alpha^*(\beta) = \mathrm{Ad}_{\tilde{\varphi}_\alpha(x)^{-1}}\, s_\alpha^*(\omega) + \tilde{\varphi}_\alpha(x)^{-1} d\tilde{\varphi}_\alpha.$$

Proof. Let $\gamma\colon \mathbb{R} \to M$ be a path such that $\gamma(0) = x$ and $\gamma'(0) = X_x$. We have to compute the following:

$$(s_\alpha^*\beta)(X_x) = (s_\alpha^*\varphi^*\omega)(X_x) = \omega_{(\varphi\circ s_\alpha)(x)}\big((\varphi\circ s_\alpha)_* X_x\big). \tag{56.340}$$

eq:ppu

What lies in the derivative is:

$$\begin{aligned}(\varphi\circ s_\alpha)_*(X_x) &= \frac{d}{dt}\Big[(\varphi\circ s_\alpha\circ\gamma)(t)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[s_\alpha(\gamma(t))\cdot\tilde{\varphi}_\alpha(\gamma(t))\Big]_{t=0}\\ &= \frac{d}{dt}\Big[s_\alpha(\gamma(t))\cdot\tilde{\varphi}_\alpha(\gamma(0))\Big]_{t=0} + \frac{d}{dt}\Big[s_\alpha(\gamma(0))\cdot\tilde{\varphi}_\alpha(\gamma(t))\Big]_{t=0}\\ &= R_{\tilde{\varphi}_\alpha(x)_*}s_{\alpha *}X_x + \frac{d}{dt}\Big[s_\alpha(x)\cdot\tilde{\varphi}_\alpha(x)e^{t\tilde{\varphi}_\alpha(x)^{-1}(d\tilde{\varphi}_\alpha)_x\gamma'(0)}\Big]_{t=0}.\end{aligned} \tag{56.341}$$

A justification of the remplacement $\tilde{\varphi}_\alpha(\gamma(t)) \to \tilde{\varphi}_\alpha(x)e^{t\tilde{\varphi}_\alpha(x)^{-1}(d\tilde{\varphi}_\alpha)_x\gamma'(0)}$ is given in the corresponding proof at page 3412. If we put this expression into equation (56.340), the first term becomes

$$\begin{aligned}\omega_{(\varphi\circ s_\alpha)(x)}\big(R_{\tilde{\varphi}_\alpha(x)_*}s_{\alpha *}X_x\big) &= (R^*_{\tilde{\varphi}_\alpha(x)}\omega)_{s_\alpha(x)}(s_{\alpha *}X_x)\\ &= \mathrm{Ad}_{\tilde{\varphi}_\alpha(x)^{-1}}\big(\omega_{s_\alpha(x)}(s_{\alpha *}X_x)\big)\\ &= \mathrm{Ad}_{\tilde{\varphi}_\alpha(x)^{-1}}(s_\alpha^*\omega)(X_x).\end{aligned}$$

The second term is the case of a connection applied to a fundamental vector field.

□

56.14 Product of principal bundle

sec:produit_bundle

In this section, we build a $G_1 \times G_2$-principal bundle from the data of a G_1 and a G_2-principal bundle. The physical motivation is clear: as far as electromagnetism is concerned, particles are sections of $U(1)$-principal bundle while the relativistic invariance must be expressed by means of a $\mathrm{SL}(2,\mathbb{C})$-associated bundle. So the physical fields must be sections of something as the product of the two bundles. See subsection 70.3.

56.14.1 Putting together principal bundle

Let us consider two principal bundle over the same base space

$$G_1 \rightsquigarrow P_1 \xrightarrow{p_1} M,$$

and

$$G_2 \rightsquigarrow P_2 \xrightarrow{p_2} M.$$

First we define the set

$$P_1 \circ P_2 = \{(\xi_1,\xi_2) \in P_1 \times P_2 \text{ tel que } p_1(\xi_1) = p_2(\xi_2)\} \tag{56.342}$$

which will be the total space of our new bundle. The projection $p\colon P_1 \circ P_2 \to M$ is naturally defined by

$$p(\xi_1,\xi_2) = p_1(\xi_1) = p_2(\xi_2),$$

while the right action of $G_1 \times G_2$ on $P_1 \circ P_2$ is given by

$$(\xi_1,\xi_2)\cdot(g_1,g_2) = (\xi_1\cdot g_1, \xi_2\cdot g_2)$$

With all these definitions,

$$\begin{array}{ccc} G_1\times G_2 & \rightsquigarrow & P_1\circ P_2 \\ & & \downarrow p \\ & & M \end{array}$$

is a $G_1 \times G_2$-principal bundle over M. We define the natural projections

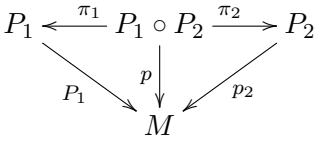
$$\pi_i \colon P_1 \times P_2 \to P_i \ (\xi_1, \xi_2) \ \mapsto \xi_i, \tag{56.343}$$

and if e_i denotes the identity element of G_i, we can identify G_1 to $G_1 \times \{e_2\}$ and G_2 to $G_2 \times \{e_1\}$; in the same way, $\mathcal{G}_1 = \mathcal{G}_1 \times \{0\} \subset \mathcal{G}_1 \times \mathcal{G}_2$. So we get the following principal bundles:

$$G_2 \rightsquigarrow P_1 \circ P_2 \xrightarrow{\pi_1} P_1$$

$$G_1 \rightsquigarrow P_1 \circ P_2 \xrightarrow{\pi_2} P_2.$$

It is clear that the following diagram commutes:



56.14.2 Connections

Let ω_i be a connection on the bundle $p_i \colon P_i \to M$. Using the identifications, $\pi_1^* \omega_1$ is a connection on $\pi_2 \colon P_1 \circ P_2 \to P_2$ (the same is true for $1 \leftrightarrow 2$), and $\pi_1^* \omega_1 \oplus \pi_2^* \omega_2$ is a connection on $p \colon P_1 \circ P_2 \to M$. Let us prove the first claim.

Let $A \in \mathcal{G}_1$. We first have to prove that $\pi_1^* \omega_1(A^*) = A$. For this, remark that $A = (A, 0) \in \mathcal{G}_1 \oplus \mathcal{G}_2$ and

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-t(A,0)}\Big]_{t=0} = \frac{d}{dt}\Big[(\xi_1, \xi_2) \cdot (e^{-tA}, e_2)\Big]_{t=0} = \frac{d}{dt}\Big[(\xi_1 \cdot e^{-tA}, \xi_2)\Big]_{t=0}, \tag{56.344}$$

so $d\pi_1 A^* = \frac{d}{dt}\Big[\pi_1(\ldots)\Big]_{t=0} = \omega_1(A^*) = A$. Let now $\Sigma \in T_{(\xi_1,\xi_2)}(P_1 \circ P_2)$ be given by the path $(\xi_1(t), \xi_2(t))$. In this case we have

$$\begin{aligned}
\big(R^*_{(g,e_2)\pi_1^*\omega_1}\big)_{(\xi_1,\xi_2)}\Sigma &= (\pi_1^*\omega_1)(dR_{(g,e_2)}\Sigma) \\
&= \omega_1\Big(\frac{d}{dt}\Big[\xi_1(t) \cdot g\Big]_{t=0}\Big) \\
&= \omega_1\Big(dR_g \frac{d}{dt}\Big[\xi_1(t)\Big]_{t=0}\Big) \\
&= \mathrm{Ad}(g^{-1})\pi_1^*\omega_1\Big(\frac{d}{dt}\Big[(\xi_1(t), \xi_2(t))\Big]_{t=0}\Big) \\
&= \mathrm{Ad}(g^{-1})\pi_1^*\omega_1\Sigma.
\end{aligned} \tag{56.345}$$

56.14.3 Representations

Let V be a vector space and $\rho_i \colon G_i \to GL(V)$ be some representations such that

$$[\rho_1(g_1), \rho_2(g_2)] = 0 \tag{56.346}$$

eq:cond_reprez

for all $g_1 \in G_1$ and $g_2 \in G_2$ (in the sense of commutators of matrices). In this case, one can define the representation $\rho_1 \times \rho_2 \colon G_1 \times G_2 \to GL(V)$ by

$$(\rho_1 \times \rho_2)(g_1, g_2) = \rho_1(g_1) \circ \rho_2(g_2) = \rho_2(g_2) \circ \rho_1(g_1). \tag{56.347}$$

The relation (56.346) is needed in order for $\rho_1 \times \rho_2$ to be a representation, as one can check by writing down explicitly the requirement

$$(\rho_1 \times \rho_2)\big((g_1, g_2)(g_1', g_2')\big) = (\rho_1 \times \rho_2)(g_1 g_1', g_2 g_2')$$

56.15 Connections

56.15.1 Gauge potentials

Let us consider a **section** σ_α of P over $\mathcal{U}_\alpha$. It is a map $\sigma_\alpha : \mathcal{U}_\alpha \to P$ such that $\pi \circ \sigma_\alpha = \mathrm{Id}$. A **connection** on P is a 1-form $\omega : T_pP \to \mathcal{G} \in \Omega^1(P)$ which satisfies the following two conditions:

$$\omega_p(Y_p^*) = Y, \tag{56.348a}$$
conn_1

$$\omega(dR_g\xi) = g^{-1}\omega(\xi)g. \tag{56.348b}$$
conn_2

The **gauge potential** of ω with respect of the local section σ_α is the 1-form on $\mathcal{U}_\alpha$ given by
PgLocSecConn

$$A_\alpha(x)(v) = (\sigma_\alpha^*\omega)_x(v). \tag{56.349}$$

We will not always explicitly write the dependence of A_α in x. Now we consider another section $\sigma_\beta : \mathcal{U}_\beta \to P$ which is related on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ to σ_α by $\sigma_\beta(x) = \sigma_\alpha(x) \cdot g_{\alpha\beta}(x)$ for a well defined map $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$.

Proposition 56.88.
The gauge potentials A_α and A_β are related by

$$A_\beta = g^{-1}A_\alpha g - g^{-1}dg. \tag{56.350}$$
trans_A

prop:trans_A

Proof. By definition, for $v \in T_x\mathcal{U}_\alpha$,

$$A_\beta(v) = (\sigma_\beta^*\omega)_x(v) = \omega_{\sigma_\alpha(x)\cdot g_{\alpha\beta}(x)}\big((d\sigma_\beta)_x(v)\big).$$

We begin by computing $d\sigma_\beta(v)$. Let us take a path $v(t)$ in $\mathcal{U}_\alpha$ such that $v(0) = x$ and $v'(0) = v$. We have:

$$\begin{aligned}
(d\sigma_\beta)_x(v) &= \frac{d}{dt}\sigma_\beta(v(t))\Big|_{t=0} \\
&= \frac{d}{dt}\sigma_\alpha(v(t)) \cdot g_{\alpha\beta}(v(t))\Big|_{t=0} \\
&= \frac{d}{dt}\Big[\sigma_\alpha(v(t)) \cdot g_{\alpha\beta}(x)\Big]_{t=0} + \frac{d}{dt}\Big[\sigma_\alpha(x) \cdot g_{\alpha\beta}(v(t))\Big]_{t=0} \\
&= dR_{g_{\alpha\beta}(x)}d\sigma_\alpha(v) + \frac{d}{dt}\Big[\sigma_\alpha(x) \cdot g_{\alpha\beta}(x)e^{-ts}\Big]_{t=0} \\
&= dR_{g_{\alpha\beta}(x)}d\sigma_\alpha(v) + s^*_{\sigma_\alpha(x)\cdot g_{\alpha\beta}(x)}
\end{aligned} \tag{56.351}$$
eq:1407r1

where s is defined by the requirement that $g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))$ can be replaced in the derivative by e^{-ts}, so that we can replace $g_{\alpha\beta}(v(t))$ by $g_{\alpha\beta}(x)e^{-ts}$. As far as the derivatives are concerned, $e^{-ts} = g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))$, then
pg:justif_s

$$s = -\frac{d}{dt}g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))\Big|_{t=0} = -g_{\alpha\beta}(x)^{-1}dg_{\alpha\beta}(v),$$

this equality being a notation. Now, properties (56.348a) and (56.348b) make that

$$A_\beta(v) = g_{\alpha\beta}(x)^{-1}\omega_{\sigma_\alpha(x)}(d\sigma_\alpha(v))g_{\alpha\beta}(x) + s.$$

The thesis is just the same, with "reduced" notations (see section 56.4.1). . □

An explicit form for this transformation law is:

$$A_\beta(v) = \frac{d}{dt}\Big[g^{-1}e^{tA_\alpha(v)}g\Big]_{t=0} - \frac{d}{dt}\Big[g^{-1}g_{\alpha\beta}(v(t))\Big]_{t=0}, \tag{56.352}$$

where $g := g_{\alpha\beta}(x)$.

56.15.2 Covariant derivative

When we have a connection on a principal bundle, we can define a covariant derivative on any associated bundle. Let us quickly review it. An associated bundle is the semi-product $E = P \times_\rho V$ where V is a vector space on which acts the representation ρ of G. We denote the canonical projection by $\pi_p \colon E \to M$. The classes are taken with respect to the equivalence relation $(p, v) \sim (p \cdot g, \rho(g^{-1})v)$.

A **section** of E is a map $\psi \colon M \to E$ such that $\pi \circ \psi = \mathrm{Id}$. We denote by $\Gamma(E)$ the set of all the sections of E. A section of E defines (and is defined by) an equivariant function $\hat{\psi} \colon P \to V$ such that

$$\psi(\pi(\xi)) = [\xi, \hat{\psi}(\xi)], \tag{56.353a}$$
$$\hat{\psi}(\xi \cdot g) = \rho(g^{-1})\hat{\psi}(\xi). \tag{56.353b}$$

For a section $\psi \in \Gamma(E)$, we define $\psi_{(\alpha)} \colon \mathcal{U}_\alpha \to V$ by

$$\psi_{(\alpha)}(x) = \hat{\psi}(\sigma(x)).$$

We saw in (56.317) that a covariant derivative on E is given by

$$(D_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} - \rho_*\Big((\sigma_\alpha^*\omega)_x(X_x)\Big)\psi_{(\alpha)}(x). \tag{56.354}$$ 3008e1

Since $(d\psi)(X) = X(\psi)$, we can rewrite this formula in a simpler manner by forgetting the index α and the mention of X:

$$D\psi = d\psi - (\rho_* A_\alpha)\psi.$$

If we note $(\rho_* A_\alpha)\psi$ by $A_\alpha\psi$, we have

$$D\psi = d\psi - A\psi. \tag{56.355}$$

One has to understand that equation as a "notational trick" for (56.354). By the way, remark that $(\rho_* A_\alpha)$ is the only "reasonable" way for A to act on ψ.

56.15.3 Gauge transformation

A **gauge transformation** of a G-principal bundle is a diffeomorphism $\varphi \colon P \to P$ which satisfies

$$\pi \circ \varphi = \pi, \tag{56.356a}$$
$$\varphi(\xi \cdot g) = \varphi(\xi) \cdot g. \tag{56.356b}$$

In local coordinates, it can be expressed in terms of a function $\tilde{\varphi}_\alpha \colon \mathcal{U}_\alpha \to G$ by the requirement that

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x) \cdot \tilde{\varphi}_\alpha(x). \tag{56.357}$$ def_vpt

We have shown in proposition 56.85 that, if ω is a connection 1-form on P, the form $\varphi \cdot \omega := \varphi^*\omega$ is still a connection 1-form on P. Of course, with the same section σ_α than before, we can define a gauge potential $(\varphi \cdot A)_\alpha$ for this new connection. We will see how it is related to A_α. The reader can guess the result (it will be the same as proposition 56.88). We show it.

Proptr_de_A

Proposition 56.89.

$$(\varphi \cdot A) = \tilde{\varphi}^{-1}A\tilde{\varphi} - \tilde{\varphi}^{-1}d\tilde{\varphi}. \tag{56.358}$$ tr_de_A

Proof. Let us consider $x \in \mathcal{U}_\alpha$, and $v \in T_x\mathcal{U}_\alpha$, the vector which is tangent to the curve $v(t) \in \mathcal{U}_\alpha$. We compute

$$\sigma_\alpha^*(\varphi^*\omega)_x(v) = \omega_{(\varphi\circ\sigma_\alpha)(x)}((d\varphi \circ d\sigma_\alpha)(v)),$$

but equation (56.357) makes

$$\begin{aligned}(d\varphi \circ d\sigma_\alpha)(v) &= \frac{d}{dt}\varphi(\sigma_\alpha(v(t)))\Big|_{t=0} \\ &= \frac{d}{dt}\sigma_\alpha(v(t)) \cdot \tilde{\varphi}_\alpha(v(t))\Big|_{t=0}.\end{aligned} \tag{56.359}$$

Now, we are in the same situation as in equation (56.351). □

If $\psi\colon M \to E$ is a section of E, the gauge transformation $\varphi\colon P \to P$ acts on ψ by

$$\widehat{\varphi \cdot \psi}(\xi) = \hat{\psi}(\varphi^{-1}(\xi)). \tag{56.360}$$

On the other hand, φ acts on the covariant derivative (and the potential): $\varphi \cdot D$ is the covariant derivative of the connection $\varphi \cdot \omega$. Of course, we define

$$(\varphi \cdot D)\psi = d\psi - (\varphi \cdot A)\psi. \tag{56.361}$$

Lemma 56.90.
If $\varphi\colon P \to P$ is a gauge transformation, then

(1) φ^{-1} is also a gauge transformation and $(\widetilde{\varphi^{-1}})_\alpha(x) = \tilde{\varphi}_\alpha(x)^{-1}$, lem:i
(2) $(\varphi \cdot \psi)_{(\alpha)}(x) = \rho(\tilde{\varphi}_x^{-1})\psi_{(\alpha)}(x)$. lem:ii

lem:prop_gauge

Proof. The first part is clear while the second is a computation:

$$(\varphi \cdot \psi)_{(\alpha)} = \widehat{\varphi \cdot \psi}(\sigma_\alpha(x)) = \hat{\psi}(\varphi^{-1}(\sigma_\alpha(x))) = \hat{\psi}(\sigma_\alpha(x) \cdot \tilde{\varphi}_\alpha(x)^{-1}) = \rho(\tilde{\varphi}_\alpha(x))\psi_{(\alpha)}(x). \tag{56.362}$$

□

Now, we will proof the main theorem: the one which explains why the covariant derivative is "covariant".

th:covariance

Theorem 56.91.
The covariant derivative D fulfils a "covariant" transformation rule under gauge transformations:

$$(\varphi \cdot D)(\varphi^{-1} \cdot \psi) = \varphi^{-1}(D\psi). \tag{56.363}$$ eq:covariance_math

Proof of theorem 56.91. First, we look at $(\varphi \cdot A)\psi_\alpha$. Using all the notational tricks used to give a sens to $A\psi$, we write:

$$[(\varphi \cdot A)_X\psi]_{(\alpha)}(x) = (\varphi \cdot A)_X\psi_{(\alpha)}(x) = \rho_*(\varphi \cdot A(X))\psi_{(\alpha)}(x).$$

But we know that $\varphi \cdot A = \tilde{\varphi}^{-1}A\tilde{\varphi} - \tilde{\varphi}^{-1}d\tilde{\varphi}$, then

$$\begin{aligned}(\varphi \cdot A)_X\psi_{(\alpha)}(x) &= \rho_*(\tilde{\varphi}^{-1}A(X)\tilde{\varphi})\psi_{(\alpha)}(x) \\ &\quad - \rho_*(\tilde{\varphi}^{-1}d\tilde{\varphi}(X))\psi_{(\alpha)}(x) \\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}e^{tA(X)}\tilde{\varphi})\psi_{(\alpha)}(x)\Big]_{t=0} \\ &\quad - \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t))\psi_{(\alpha)}(x)\Big]_{t=0}\end{aligned} \tag{56.364}$$ eq:en_deux

Now, we have to write this equation with $\varphi^{-1} \cdot \psi$ instead of ψ. Using lemma 56.90, we find:

$$\begin{aligned}(\varphi \cdot A)_X(\varphi^{-1} \cdot \psi)_{(\alpha)}(x) &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}e^{tA(X)}\tilde{\varphi}\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0} \\ &\quad - \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t)\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0}\end{aligned} \tag{56.365}$$

After simplification, the first term is a term of the thesis: $\tilde{\varphi}(x)^{-1}(A\psi)_\alpha(x)$ and we let the second one as it is. Now, we turn our attention to the second term of (56.363); the same argument gives:

$$\begin{aligned} d(\varphi^{-1}\psi_{(\alpha)})_x X &= \frac{d}{dt}\Big[(\varphi^{-1}\cdot\psi)_{(\alpha)}(X_t)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(X_t)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(x)\Big]_{t=0} + \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1})\psi_{(\alpha)}(X_t)\Big]_{t=0}. \end{aligned} \tag{56.366}$$

The second term is $\tilde{\varphi}^{-1}d\psi_\alpha(X)$. In definitive, we need to prove that the two exceeding terms cancel each other:

$$\frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t)\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0} + \frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(x)\Big]_{t=0} \tag{56.367}$$

eq:le_zero

must be zero.

One can find a $g(t)\in G$ such that $\tilde{\varphi}(X_t)=\tilde{\varphi}g(t)$, $g(0)=e$. Then, what we have in the ρ of these two terms is respectively $g(t)\tilde{\varphi}^{-1}$ and $g(t)^{-1}\tilde{\varphi}^{-1}$. As far as the derivative are concerned, $g(t)$ can be written as e^{tZ} for a certain $Z\in\mathcal{G}$. So we see that $g(t)^{-1}=e^{-tZ}$ and the derivative will come with the right sign to makes the sum zero. □

Remarque 56.92.
Let us use more intuitive notations: we write (56.358) under the form $A'=g^{-1}Ag-g^{-1}dg$. If we have two sections ψ and ψ', they are necessarily related by a gauge transformation: $\psi'=g^{-1}\psi$. Then, the theorem tells us that the equation $D\psi=d\psi-A\psi$ becomes $D'\psi'=g^{-1}D\psi$ "under a gauge transformation". This is: $D\psi$ transforms under a gauge transformation as $d\psi$ transforms under a constant linear transformation. This is the reason why D is a *covariant* derivative. The physicist way to write (56.363) is

$$D'\psi'=g^{-1}D\psi \tag{56.368}$$

eq:covariance_phys

rem:intuitif

Remarque 56.93.
If we naively make the computation with the notations of remark 56.92, we replace $\psi'=g^{-1}\psi$ and $A'=g^{-1}Ag-g^{-1}dg$ in

$$D'\psi'=d\psi'-A'\psi',$$

using some intuitive "Leibniz formulas", we find: $D'\psi'=dg^{-1}\psi+g^{-1}d\psi+g^{-1}A\psi+g^{-1}dgg^{-1}\psi$. It is exactly $g^{-1}d\psi+g^{-1}A\psi$ with two additional terms: $dg^{-1}\psi$ and $g^{-1}dgg^{-1}\psi$. One sees that these are precisely the two terms of the expression (56.367). We will give a sens to this "naive" computation in section 56.4.1.

56.15.4 Curvature

From the $\mathcal{G}$-valued connection 1-form ω on P, we may define its **curvature 2-form**:

$$\Omega=d\omega+\omega\wedge\omega. \tag{56.369}$$

As before, we can see Ω as a 2-form on M instead of P. For this, we just need some sections $\sigma_\alpha\colon\mathcal{U}_\alpha\to P$ and define

$$F_\alpha=\sigma_\alpha^*\Omega. \tag{56.370}$$

This F is called the **Yang-Mills field strength**. The question is now to see how does it transform under a change of chart? What is $F_\beta=\sigma_\beta^*\Omega$ in terms of F_α?

First, note that we can't try to find a relation like $d(g\omega)=dg\wedge\omega+g\,d\omega$. Pose $A_x=g(x)\omega_x$:

$$A_x(v)=\frac{d}{dt}g(x)e^{t\omega_x(v)}\Big|_{t=0}.$$

Using

$$(d\alpha)(v,w) = v(\alpha(w)) - w(\alpha(v)) - \alpha([v,w]),$$

we are led to write

$$w(A(v)) = d(A(v))w = \frac{d}{du}A_{w_u}(v)\Big|_{u=0} = \frac{d}{du}\frac{d}{dt}\Big[g(w_u)e^{t\omega_{w_u}(v)}\Big]_{t=0}\Big|_{u=0}. \tag{56.371}$$

But at $t = u = 0$, the expression in the bracket is $g(x)$, and not e. Then the whole expression is not an element of $\mathcal{G}$. In other words, the problem is that for $g\colon M \to G$, we have $dg_x\colon T_xM \to T_{g(x)}G \neq T_eG$.

Now, remark that in our matter, the problem will not arise because in the expressions $A_\beta = g^{-1}A_\alpha g + g^{-1}dg$, each term has a g and a g^{-1}.

Lemma 56.94.

$$d(g^{-1})_x(v) = -g(x)^{-1}dg(v)g(x)^{-1}. \tag{56.372}$$

lem:dgemu

Proof. Let v_t be a path which defines the vector v, and define $Y \in \mathcal{G}$ such that as far as the derivative are concerned, we have $g(v_t) = g(x)e^{tY}$. Then,

$$g(g^{-1})(v) = \frac{d}{dt}\Big[g(v_t)^{-1}\Big]_{t=0} = \frac{d}{dt}\Big[e^{-tY}g(x)^{-1}\Big]_{t=0}.$$

But on the other hand,

$$g^{-1}dg(v)g^{-1} = \frac{d}{dt}\Big[g(x)^{-1}g(v_t)g(x)^{-1}\Big]_{t=0} = \frac{d}{dt}\Big[e^{tY}g(x)^{-1}\Big]_{t=0},$$

thus $d(g^{-1})_x(v) = -g(x)^{-1}dg(v)g(x)^{-1}$, as we want. □

Theorem 56.95.

$$F_\beta = g^{-1}F_\alpha g. \tag{56.373}$$

tho:trans_F

Naive proof. Let us accept $F_\beta = dA_\beta + A_\beta \wedge A_\beta$. With proposition 56.88, we can perform a simple computation with all the intuitive "Leibniz rules":

$$dA_\beta = -g^{-1}dg\,g^{-1} \wedge A_\alpha g + g^{-1}dA_\alpha g + g^{-1}A_\alpha \wedge dg - g^{-1}dg\,g^{-1} \wedge dg,$$

and

$$A_\beta \wedge A_\beta = g^{-1}A_\alpha g \wedge g^{-1}A_\alpha g + g^{-1}A_\alpha g \wedge g^{-1}dg + g^{-1}dg \wedge g^{-1}A_\alpha g + g^{-1}dg \wedge g^{-1}dg.$$

The sum is obviously the announced result. □

This proof seems too beautiful to be false[20]. Here is the full proof.

Ultimate proof of theorem 56.95. First we compute $d(g^{-1}A_\alpha g)$. In order to do this, remark that the 1-form $g^{-1}A_\alpha g$ is explicitly given on $v \in \mathfrak{X}(M)$ by

$$(g^{-1}A_\alpha g)(v)_x = \frac{d}{dt}\Big[g(x)^{-1}e^{tA(v)_x}g(x)\Big]_{t=0}.$$

For all $x \in M$, this expression is an element of $\mathcal{G}$; then we can say that $(g^{-1}A_\alpha g)(v)$ is a map $(g^{-1}A_\alpha g)(v)\colon M \to \mathcal{G}$. So it is unambiguous to write $w((g^{-1}A_\alpha g)(v)) \in \mathcal{G}$ for $w \in T_xM$.

20. More precisely, it is as beautiful as we want it to be true.

We will use the formula

$$d(g^{-1}A_\alpha g)(v,w) = v(g^{-1}A_\alpha g)(w) - w(g^{-1}A_\alpha g)(v) - (g^{-1}A_\alpha g)([v,w]).$$

As $w((g^{-1}A_\alpha g)(v)) = d((g^{-1}A_\alpha g)(v))w$, we have

$$\begin{aligned}
w((g^{-1}A_\alpha g)(v)) &= \frac{d}{du}(g^{-1}A_\alpha g)(v)_{w_u}\Big|_{u=0} \\
&= \frac{d}{du}\frac{d}{dt}\Big[g(w_u)^{-1}e^{tA(v)_{w_u}}g(w_u)\Big]_{t=0}\Big|_{u=0} \\
&= \frac{d}{dt}\frac{d}{du}\Big[g(w_u)^{-1}\Big]_{u=0}e^{tA(v)_x}g(x)\Big|_{t=0} \\
&\quad + \frac{d}{dt}g(x)^{-1}\frac{d}{du}\Big[e^{tA(v)_{w_u}}\Big]_{u=0}g(x)\Big|_{t=0} \\
&\quad + \frac{d}{dt}g(x)^{-1}e^{tA(v)_x}\frac{d}{du}\Big[g(w_u)\Big]_{u=0}\Big|_{t=0} \\
&= d(g^{-1})(w)A(v)_x g(x) \\
&\quad + g(x)^{-1}w_x(A(v))g(x) \\
&\quad + g(x)^{-1}A(v)_x dg(w).
\end{aligned} \tag{56.374}$$

On the other hand, one easily finds that

$$(g^{-1}A_\alpha g)([v,w]) = g(x)^{-1}A([v,w])g(x).$$

Using lemma 56.94, we have

$$\begin{aligned}
d(g^{-1}A_\alpha g)_x(v,w) &= -g(x)^{-1}dg(v)g(x)^{-1}A(w)_x g(x) + g(x)^{-1}v(A(w))g(x) \\
&\quad + g(x)^{-1}A(w)_x dg(v)_x \\
&\quad + g(x)^{-1}dg(w)_x g(x)^{-1}A(v)_x g(x) - g(x)^{-1}w(A(v))g(x) \\
&\quad - g(x)^{-1}A(v)_x dg(w) \\
&\quad - g(x)^{-1}A([v,w])g(x).
\end{aligned} \tag{56.375}$$

We can regroup the terms two by two in order to form dA_α and some wedge; with simpler notations, we can write:

$$d(g^{-1}A_\alpha g) = -(g^{-1}dg\, g)\wedge(A_\alpha g) - (g^{-1}A)\wedge dg + (g^{-1}dAg). \tag{56.376}$$

We compute $d(g^{-1}dg)$ in the same way; the result is

$$(g^{-1}dg)(v)_x = \frac{d}{dt}\Big[g(x)^{-1}g(v_t)\Big]_{t=0} \in \mathcal{G}.$$

For v, $w \in \mathfrak{X}(M)$, we have:

$$\begin{aligned}
w\big((g^{-1}dg)(v)\big) &= \frac{d}{du}(g^{-1}dg)(v)_{w_u}\Big|_{u=0} \\
&= \frac{d}{du}\frac{d}{dt}\Big[g(w_u)^{-1}g(v_{w_u}(t))\Big]_{\substack{t=0\\u=0}} \\
&= \frac{d}{dt}\frac{d}{du}\Big[g(w_u)^{-1}g(v_t)\Big]_{\substack{u=0\\t=0}} + \frac{d}{dt}\frac{d}{du}\Big[g(x)^{-1}g(w_u(t))\Big]_{\substack{u=0\\t=0}} \\
&= d(g^{-1})(w)dg(v) + \frac{d}{du}\Big[g(x)^{-1}dg(v_{w_u})\Big]_{u=0} \\
&= -g^{-1}dg(w)g^{-1}dg(v) + g(x)^{-1}w(dg(v))
\end{aligned} \tag{56.377}$$

where w_u is a path such that $w_0' = w_x$ and $v_{w_u}(t)$ is, with respect of t, a path which gives the vector v_{w_u}. On the another hand, we have

$$(g^{-1}dg)([v,w]) = g^{-1}dg([v,w]).$$

Remark that the term $g(x)^{-1}w(dg(v))$ of $w((g^{-1}dg)(v))$ together with the same with $v \leftrightarrow w$ and $(g^{-1}dg)([v,w])$ which comes from $(g^{-1}dg)([v,w])$ will give $g(x)^{-1}(d^2g)(v,w) = 0$ when we will compute $d(g^{-1}dg)$. Finally,

$$d(g^{-1}dg) = -(g^{-1}dg\, g^{-1} \wedge dg). \tag{56.378}$$

The equations (56.376) and (56.378) allow us to write:

$$\begin{aligned}(dA_\beta) &= d(g^{-1}A_\alpha g) + d(g^{-1}dg)\\ &= -(g^{-1}dg\, g^{-1}) \wedge (A_\alpha g) - (g^{-1}A_\alpha) \wedge dg\\ &\quad + (g^{-1}dA_\alpha g) - (g^{-1}dg\, g^{-1}) \wedge dg.\end{aligned} \tag{56.379}$$

Notice that the term $(g^{-1}dA_\alpha g)$ corresponds to the first one in $F_\beta = g^{-1}(dA_\beta + A_\beta \wedge A_\beta)g$.

For anyone who had understood the whole computations up to here, it is clear that

$$\begin{aligned}[A_\beta(v), A_\beta(w)] = &\frac{d}{dt}\frac{d}{du}\Big[e^{tA_\beta(v)}e^{tA_\beta(w)}\Big]_{\substack{u=0\\t=0}}\\ &- \frac{d}{dt}\frac{d}{du}\Big[e^{tA_\beta(w)}e^{tA_\beta(v)}\Big]_{\substack{u=0\\t=0}},\end{aligned} \tag{56.380}$$

so that

$$\begin{aligned}A_\beta \wedge A_\beta = &g^{-1}A_\alpha g \wedge g^{-1}A_\alpha g + g^{-1}A_\alpha g \wedge g^{-1}dg\\ &+ g^{-1}dg \wedge g^{-1}A_\alpha g + g^{-1}dg \wedge g^{-1}dg.\end{aligned} \tag{56.381}$$

Lemma 56.47 allows us to write it under the form

$$\begin{aligned}A_\beta \wedge A_\beta = &g^{-1}A_\alpha g \wedge g^{-1}A_\alpha g + g^{-1}A_\alpha g \wedge g^{-1}dg\\ &+ g^{-1}dg \wedge g^{-1}A_\alpha g + g^{-1}dg \wedge g^{-1}dg.\end{aligned} \tag{56.382}$$

Here the term $(g^{-1}A_\alpha \wedge A_\alpha g)$ corresponds to the second one in $F_\beta = g^{-1}(dA_\beta + A_\beta \wedge A_\beta)g$. The sum of (56.379) and (56.382) is

$$F_\beta = g^{-1}F_\alpha g.$$

□

56.16 Hodge decomposition theorem and harmonic forms

Among other sources for Hodge decomposition and harmonic forms, we have [917, 918, 919]. Some parts of the wikipedia article Hodge_dual are interesting.

Let E be an oriented Euclidian space of dimension $m = 2n$. We define the operation $*$ by

$$\begin{aligned}*\colon \bigwedge E &\to \bigwedge E\\ e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k} &\mapsto e_{i_{k+1}} \wedge \cdots \wedge e_{i_m}\end{aligned} \tag{56.383}$$

when $\{e_i\}$ is an oriented basis of E and $\{i_k\}$ is an even permutation of $\{1, 2, \cdots, m\}$. If it is impossible to build an even permutation, then we add a minus sign. We have $**\omega = (-1)^{p(m-p)}\omega$ belongs to $\omega \in \bigwedge^p E$.

Exemple 56.96.
If we consider the space $\mathbb{R}^4$ with the coordinates (x, y, z, t),

$$*(dx \wedge dz \wedge dt) = dy \tag{56.384}$$

because (x,z,t,y) is an even permutation of (x,y,z,t). Now, $*dy = dz\wedge dx\wedge dt = -(dx\wedge dz\wedge dt)$ because (y,z,x,t) is an even permutation of (x,y,z,t).

More generally, if we have a differential p-form ω on a m dimensional space, we have

$$*(e_{\sigma(1)}\wedge e_{\sigma(2)}\wedge\ldots\wedge e_{\sigma(p)}) = e_{\sigma(p+1)}\wedge\ldots\wedge e_{\sigma(m)} \tag{56.385}$$

In order to compute $*(e_{\sigma(p+1)}\wedge\ldots\wedge e_{\sigma(m)})$, we need a permutation of $\big(\sigma(1),\ldots,\sigma(m)\big)$ which *begins* by $\sigma(p+1)\ldots\sigma(m)$. This reduces to permute the $m-p$ elements $\sigma(p+1),\ldots,\sigma(m)$ with the p first elements. Thus we have

$$**\omega = (-1)^{p(m-p)}\omega. \tag{56.386}$$

△

Let V be a compact, oriented manifold. Each of the spaces of sections $C^\infty\big(V,\bigwedge^k_{\mathbb{C}}(T^*V)\big)$ is endowed with a 2-form

$$\langle\omega_1,\omega_2\rangle = \int_V \omega_1\wedge *\omega_2. \tag{56.387}$$ EqProdWedgeHOfge

Lemma 56.97.
*The **codifferential** δ defined by*

$$\begin{aligned}\delta\colon \bigwedge^k_{\mathbb{C}}(T^*V) &\to \bigwedge^{k-1}_{\mathbb{C}}(T^*V)\\ \beta &\mapsto (-1)^{mk+m+1}*d*\beta\end{aligned} \tag{56.388}$$

is a formal adjoint of d for the product (56.387).

Proof. If $\beta\in\bigwedge^k_{\mathbb{C}}(T^*V)$ and $\alpha\in\bigwedge^{k-1}_{\mathbb{C}}(T^*V)$, we have

$$\begin{aligned}*\delta\beta &= (-1)^{mk+m+1}**\big(d*\beta\big)\\ &= (-1)^{mk+m+1}(-1)^{(m-k+1)(m-m+k-1)}d*\beta\\ &= (-1)^k d*\beta.\end{aligned} \tag{56.389}$$

Using that formula we find

$$\begin{aligned}\langle d\alpha,\beta\rangle - \langle\alpha,\delta\beta\rangle &= \int_V d\alpha\wedge*\beta - \alpha\wedge*\delta\beta\\ &= \int_V d\alpha\wedge*\beta - (-1)^k\alpha\wedge d*\beta\\ &= \int_V d\alpha\wedge*\beta + (-1)^{k+1}\alpha\wedge d*\beta\\ &= \int_V d(\alpha\wedge*\beta)\\ &= \int_{\partial V}\alpha\wedge*\beta\\ &= 0.\end{aligned} \tag{56.390}$$

This proves that $\langle d\alpha,\beta\rangle = \langle\alpha,\delta\beta\rangle$.

¡¡ Avertissement/question au lecteur !! 56.98
I do not understand why the integral in the boundary is zero.

□

Now we define the **Laplace-Beltrami operator** by

$$\Delta = \delta d + d\delta \tag{56.391}$$

and the space of **harmonic forms**

$$H^k = \{\omega \in \Omega^k \text{ tel que } \Delta\omega = 0\}. \tag{56.392}$$

Lemma 56.99.
If M is a closed manifold, a k-form is harmonic if and only if $d\omega = \delta\omega = 0$.

Proof. No proof. □

Theorem 56.100 (Hodge decomposition theorem)**.**
For every integer $0 \leqslant k \leqslant m$, the space H^p is finite dimensional and $\Omega^k(M)$ has the orthogonal decomposition

$$\Omega^k(M) = H^k \oplus \Delta\big(\Omega^k(M)\big), \tag{56.393}$$

i.e. the space splits into the kernel of Δ and its image.

Theorem 56.101.
Let M be a compact orientable manifold of dimension m. Any exterior differential k-form can be written as a unique sum of an exact form, a coexact form and an harmonic form:

$$\omega = d\alpha + \delta\beta + \gamma. \tag{56.394}$$

with $\omega \in \Omega^k(M)$, $\alpha \in \Omega^{k-1}(M)$, $\beta \in \Omega^{k+1}(M)$ and $\gamma \in \Omega^k(M)$ harmonic.

The operator Δ commutes with the differential d and we have $d\Delta = \Delta d = d\delta d$ since

$$d\Delta\omega = dd\delta\omega + d\delta d\omega = d\delta d\omega, \tag{56.395}$$

because $d^2 = 0$, while

$$\Delta d\omega = d\delta d\omega + \Delta dd\omega = d\delta d\omega. \tag{56.396}$$

Lemma 56.102.
On a close manifold, $\Delta\omega = 0$ if and only if $\delta\omega = d\omega = 0$.

In the case of a closed manifold, a form is harmonic if and only if is belongs to the kernel of $d + \delta$. Moreover, a form in the image of $d + \delta$ is orthogonal to the harmonic forms:

$$\langle d\alpha^{k-1} + \delta\beta^{k+1}, \gamma\rangle = 0 \tag{56.397}$$

whenever γ is harmonic on a closed manifold.

Chapter 57

Integration on manifolds

57.1 Orientation and volume forme

57.1.1 Orientation

PROPooGMOQooSXZEBxy

Proposition-Definition 57.1 ([920]).
Let M be a smooth manifold. Let two charts $\varphi_\alpha\colon U_\alpha \to M$ and $\varphi_\beta\colon U_\beta \to M$. We say that $(U_\alpha, \varphi_\alpha) \sim (U_\beta, \varphi_\beta)$ if

$$\det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_x\right) > 0 \tag{57.1}$$

for every $x \in U_\alpha \cap U_\beta$.

This is an equivalence relation[1].

DEFooQZJPOoOJeNRiAt

Definition 57.2 ([920]).
*An **orientation** on a smooth manifold is an atlas in which all the charts are equivalent*[2].

*If such an atlas exists, we say that the manifold M is **orientable**.*

PROPooMKHKooUTiUjF

Proposition-Definition 57.3 ([920]).
*We say that two orientations $\mathcal{A}$ and $\mathcal{B}$ are **equivalent** if $\mathcal{A} \cup \mathcal{B}$ is an orientation.*

This is an equivalence relation.

LEMooYMIYooUDOhyR

Lemma 57.4 ([1]).
Let $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be charts of the smooth manifold M. For $\alpha, \beta \in \Lambda$, we write $M_{\alpha\beta} = \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$, and we define the map

$$\begin{aligned} f_{\alpha\beta}\colon M_{\alpha\beta} &\to \mathbb{R} \\ x &\mapsto \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{\varphi_\beta^{-1}(x)}\right). \end{aligned} \tag{57.2}$$

ITEMooLHBBooFxUTHJ

(1) If $\alpha, \beta, \sigma \in \Lambda$ and $x \in M_{\alpha\beta} \cap M_{\alpha\sigma} \cap M_{\sigma\beta}$, we have

$$f_{\alpha\beta}(x) = f_{\alpha\sigma}(x) f_{\sigma\beta}(x) \tag{57.3}$$

ITEMooURXTooRWnYsV

(2) For every $x \in M_{\alpha\beta}$, we have

$$f_{\alpha\beta}(x) \neq 0. \tag{57.4}$$

Proof. Several points.

(1) **For (1)**

1. Definition 1.32
2. Equivalence in the sense of definition 57.1.

Using the fact that the differential and the determinant are multiplicative (lemma 49.74 and proposition 8.16(1)), we have

$$f_{\alpha\beta}(x) = \det\left(d(\varphi_\alpha^{-1}\circ\varphi_\beta)_{\varphi_\beta^{-1}(x)}\right) \tag{57.5a}$$
$$= \det\left(d(\varphi_\alpha^{-1}\circ\varphi_\sigma\circ\varphi_\sigma^{-1}\circ\varphi_\beta)_{\varphi_\beta^{-1}(x)}\right) \tag{57.5b}$$
$$= \det\left(d(\varphi_\alpha^{-1}\circ\varphi_\sigma)_{\varphi_\sigma^{-1}\circ\varphi_\beta\circ\varphi_\beta^{-1}(x)}\circ d(\varphi_\sigma^{-1}\circ\varphi_\beta)_{\varphi_\beta^{-1}(x)}\right) \tag{57.5c}$$
$$= \det\left(d(\varphi_\alpha^{-1}\circ\varphi_\sigma)_{\varphi_\sigma^{-1}}\circ d(\varphi_\sigma^{-1}\circ\varphi_\beta)_{\varphi_\beta^{-1}(x)}\right) \tag{57.5d}$$
$$= f_{\alpha\sigma}(x)f_{\sigma\beta}(x). \tag{57.5e}$$

(2) **For (2)** The trick is to compute $f_{\alpha\alpha}$ using the point (1). We have

$$f_{\alpha\alpha}(x) = f_{\alpha s}(x)f_{s\alpha}(x). \tag{57.6}$$

On the other hand, $f_{\alpha\alpha}(x) = \det(\mathrm{Id}) \neq 0$. A product is non vanishing only when the two factors are non vanishing. Thus $f_{\alpha s}(x) \neq 0$.

□

LEMooXMIJooCRspfz

Lemma 57.5 ([1])**.**
Let M be an orientable smooth manifold. Let $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ and $\{(U_s,\phi_s)\}_{s\in S}$ be orientations on M. For every $\alpha,\beta\in\Lambda$ and $s,t\in S$ we have

$$f_{\beta t}(x)f_{\alpha s}(x) > 0. \tag{57.7}$$

Proof. Using both points of lemma 57.4, we have

$$f_{\beta t}(x) = f_{\beta\alpha}(x)f_{\alpha s}(x). \tag{57.8}$$

Since Λ is an orientation, $f_{\beta\alpha}(x) > 0$ and by lemma 57.4(2), we deduce hat $f_{\beta t}(x)$ and $f_{\alpha s}(x)$ have the same sign. □

NORMooPXHIooHTFJoa

57.6.
When we have two charts $(U_\alpha,\varphi_\alpha)$ and (U_β,φ_β), we can write the corresponding «basis» vector fields

$$(\partial_{\alpha,i})(x) = \frac{d}{dt}\left[\varphi_\alpha\left(\varphi_\alpha^{-1}(x) + te_i\right)\right]_{t=0}, \tag{57.9}$$

and

$$(\partial_{\beta,i})(x) = \frac{d}{dt}\left[\varphi_\beta\left(\varphi_\beta^{-1}(x) + te_i\right)\right]_{t=0}. \tag{57.10}$$

57.1.2 Connected atlas

DEFooIQUXooOdVpLS

Definition 57.7 ([1])**.**
Let M be a smooth manifold with an atlas $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$. For $\alpha\in\Lambda$ we write $M_\alpha = \varphi_\alpha(U_\alpha)$. Let $x\in M$. We write

$$S_x = \{y\in M \text{ tel que } \exists\alpha_1,\ldots,\alpha_n\in\Lambda \text{ tel que } \begin{cases} x\in M_{\alpha_1} \\ y\in M_{\alpha_n} \\ M_{\alpha_i}\cap M_{\alpha_{i+1}}\neq\varnothing. \end{cases}\} \tag{57.11}$$

LEMooIFABooVGFYfI

Lemma 57.8 ([1])**.**
Let M be a smooth manifold and $x,y\in M$.

ITEMooCMDVooBobapJ

(1) If $y\in S_x$, then $x\in S_y$.

ITEMooIIBSooNNXIFj

(2) If $y\in S_x$, then $S_y = S_x$.

ITEMooNGSEooIZuiMU

(3) If $S_x \cap S_y \neq \emptyset$ then $S_x = S_y$.

The relation $x \sim y$ if $x \in S_y$ is an equivalence relation.

Proof. Two parts.

(1) **For (1)** Since $y \in S_x$, we have a chain $(\alpha_1, \dots, \alpha_n)$ such that $x \in M_{\alpha_1}$ and $y \in S_{\alpha_n}$ with $M_{\alpha_i} \cap S_{\alpha_{i+1}} \neq \emptyset$. The chain $(\alpha_n, \dots, \alpha_1)$ shows that $x \in S_y$.

(2) **For (2)** Let $y \in S_x$.

(2a) $S_y \subset S_x$ Let $a \in S_y$. We have a chain $\alpha_1, \dots, \alpha_n$ with $y \in M_{\alpha_1}$ and $a \in M_{\alpha_n}$. Since $y \in S_x$ we also have a chain $\beta_1, \dots, \beta_m$ such that $x \in M_{\beta_1}$ and $y \in M_{\beta_m}$. Since $M_{\beta_m} \cap M_{\alpha_1} \neq \emptyset$ we have the chain $(\beta_1, \dots, \beta_m, \alpha_1, \dots, \alpha_n)$ such that $x \in M_{\beta_1}$ and $a \in M_{\alpha_n}$. Thus $a \in S_x$.

(2b) $S_x \subset S_y$ We know that $y \in S_x$. By (1) we also have $x \in S_y$. The inclusion we just proved reads now $S_x \subset S_y$.

(3) **For (3)** Let $a \in S_x \cap S_y$. Since $a \in S_x$ we have $S_a = S_y$, and from $a \in S_y$ we deduce $S_a = S_y$.

The fact that $x \sim y$ is an equivalence relation is now an immediate check. □

PROPooFTFCooYTKaPF

Proposition 57.9 ([1])**.**
Let M be a connected smooth manifold. For every $a \in M$ we have $M = S_a$.

Proof. The set $\{S_x\}_{x \in M}$ is a covering of M by open parts. Let $A = S_a$ and $B = \bigcup_{\substack{x \in M \\ a \notin S_x}} S_x$.

(1) $M = A \cup B$ We prove that $M = A \cup B$. Let $x \in M$. There are two possibilities: $x \in S_a$ or not. If $x \in S_a$, then $x \in A$. Suppose that $x \notin S_a$. Then $a \notin S_x$ (lemma 57.8(1)) and then $x \in B$.

(2) $B = \emptyset$ Suppose that A and B are both non empty. From connectedness of M, we have $A \cap B \neq \emptyset$. Let $y \in A \cap B$. We have $y \in S_a$ and there exists $x \in M$ such that $a \neq S_x$ and $y \in S_x$. We have

$$y \in S_a \Rightarrow a \in S_y \tag{57.12a}$$
$$y \in S_x \Rightarrow x \in S_y. \tag{57.12b}$$

Using lemma 57.8(2) we deduce $S_a = S_x = S_y$. In particular $a \in S_x$ which is a contradiction. Thus A and B are not both non empty. The part A contains a and is then non empty. We conclude that B is empty.

(3) **Conclusion** We have $M = A \cup B = A = S_a$.

□

LEMooSDQUooWEagbY

Lemma 57.10 ([1])**.**
Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in \Lambda}$ be an atlas. For each $\alpha \in \Lambda$ we consider the set $\{U_{\alpha,i}\}_{i \in I_\alpha}$ of its connected components.

Then the set

$$\mathcal{A}_c = \{(U_{\alpha,i}, \varphi_{\alpha,i})\}_{\alpha \in \Lambda, i \in I_\alpha} \tag{57.13}$$

where $\varphi_{\alpha,i} \colon U_{\alpha,i} \to M$ is the restriction of φ_α is an atlas.

PROPooHQSSooDCjlIX

Proposition 57.11 ([1])**.**
Let $\mathcal{A}$ be an orientation. Then the atlas $\mathcal{A}_c$ is an orientation and is the same orientation as $\mathcal{A}$.

Proof. We consider two charts $\varphi_{\alpha,i}$ and $\varphi_{\beta,j}$ of $\mathcal{A}_c$ ($i \in I_\alpha$ and $j \in I_\beta$). Let $x \in M_{\alpha,i} \cap M_{\beta,j}$. Since the connected components are open, x has a neighbourhood in $M_{\alpha,i} \cap M_{\beta,j}$ and we have

$$d(\varphi_{\alpha,i}^{-1} \circ \varphi_{\beta,j})_x = d(\varphi_\alpha^{-1} \circ \varphi_\beta). \tag{57.14}$$

□

57.1.3 Volume form

PROPooBIVHooXOycnS

Proposition 57.12 ([1]).
Let M be a smooth manifold. There exists an orientation $\{(U_\alpha, \varphi_\alpha)_{\alpha\in\Lambda}\}$ and a partition of unity $\{\phi_i\}_{i\in I}$ such that

1. *The orientation $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ is countable and locally finite.*
2. *The partition of unity $\{\phi\}_{i\in I}$ is countable and subordinate to $\mathcal{W} = \{\varphi_\alpha(U_\alpha)\}_{\alpha\in\Lambda}$.*
3. *The supports $\text{supp}(\phi_i)$ are compacts.*

DEFooOBZEooMZauZF

Definition 57.13 ([920]).
*A **volume form** on a smooth n-dimensional manifold is a smooth nowhere vanishing n-differential form.*

LEMooQGVMooSHXUmD

Lemma 57.14 ([1]).
Let ω be a n-differential form on M. Let $\varphi_\alpha : U_\alpha \to M$ be a chart. There exists a unique smooth function $f_\alpha : M_\alpha \to \mathbb{R}$ such that

$$\omega_x = f_\alpha(x)\partial^*_{\alpha,1} \wedge \ldots \wedge \partial^*_{\alpha,n} \tag{57.15}$$

for every $x \in M_\alpha$.

This lemma is a combination of proposition 10.216 for the existence of f and definition 56.20 for the smoothness.

LEMooKJRCooCklXie

Lemma 57.15 ([1]).
If ω is a volume form, then

$$\omega_x(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) \neq 0 \tag{57.16}$$

for every $\alpha \in \Lambda$ and $x \in M_\alpha$.

Proof. From lemma 57.14, we have $a \in \mathbb{R}$ such that $\omega_x = a\partial^*_{\alpha,1} \wedge \ldots \wedge \partial^*_{\alpha,n}$. Notice that $a \neq 0$ because $\omega_x \neq 0$. Proposition 10.214 shows that

$$(\partial^*_{\alpha,1} \wedge \ldots \wedge \partial^*_{\alpha,n})(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) = \det\left(\partial^*_{\alpha,i}(\partial_{\alpha,j})\right) = \det(\delta_{ij}) = 1. \tag{57.17}$$

EQooRRHNooBVgAnF

Thus $\omega_x(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) = a \neq 0$. □

LEMooELOHooYtXSEH

Lemma-Definition 57.16.
Let ω be a volume form on the connected manifold M. Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an orientation on M. If there exists $x \in M$ and $\alpha \in \Lambda$ such that

$$\omega_x(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) > 0, \tag{57.18}$$

then

$$\omega_y(\partial_{\beta,1}, \ldots, \partial_{\beta,n}) > 0 \tag{57.19}$$

for every $\beta \in \Lambda$ and $y \in M_\beta = \varphi_\beta(U_\beta)$.

*In that case we say that ω is **oriented** $\mathcal{A}$-positively.*

Proof. Two parts. In the first part we show that, keeping the same point, we can change the chart, and in the second part we will show how we can change the point.

(1) **Changing chart**
Let $x \in M_\alpha \cap M_\beta$. From lemma 57.14, there exists $a\mathbb{R}$ such that

$$\omega_x = a\partial^*_{\beta,1} \wedge \ldots \wedge \partial^*_{\beta,n}. \tag{57.20}$$

Using lemma 49.95,

$$\omega_x(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) = a(\partial^*_{\beta,1} \wedge \ldots \wedge \partial^*_{\beta,n})(\partial_{\alpha,1}, \ldots, \partial_{\alpha,n}) = a \det\left(d(\varphi_\beta^{-1} \circ \varphi_\alpha)_x\right). \tag{57.21}$$

Since $\mathcal{A}$ is an orientation, the determinant is strictly positive. By hypothesis, the left hand side is strictly positive. We deduce that $a > 0$. Using the same computation as in (57.17), we find

$$\omega_x(\partial_{\beta,1}, \dots, \partial_{\beta,n}) = a \det(\mathrm{Id}) = a > 0. \tag{57.22}$$

(2) **Changing the point** Let $x, y \in M$ We consider the connected orientation $\mathcal{A}_c = \{(U_i, \varphi_i)\}_{i\in I}$ associated with $\mathcal{A}$ (proposition 57.11). Since M is connected, $S_x = M$ (proposition 57.9) and there exists a chain $i_0, \dots, i_n \in I$ and points $a_k \in \varphi_{i_k}(U_k) \cap \varphi_{i_{k+1}}(U_{k+1})$ such that $x \in M_{i_0}$ and $y \in M_{i_n}$.
The map

$$\begin{aligned} f\colon \varphi_{i_0}(U_{i_0}) &\to \mathbb{R} \\ s &\mapsto \omega_s(\partial_{i_0,1}, \dots, \partial_{i_0,n}) \end{aligned} \tag{57.23}$$

satisfy:

(2a) It is continuous

(2b) It's domain is connected,

(2c) It does not vanish (lemma 57.15).

(2d) Let $\beta \in \Lambda$ such that U_{i_0} is one of the connected components of U_β. We have

$$f(x) = \omega_x(\partial_{i_0,1}, \dots, \partial_{i_0,n}) = \omega_x(\partial_{\beta,1}, \dots, \partial_{\beta,n}) > 0. \tag{57.24}$$

The last inequality is the first part of this lemma "Changing charts".

Thus $f(s) > 0$ for every $s \in \varphi_{i_0}(U_{i_0})$ (intermediate value theorem 9.57). In particular $f(a_0) > 0$. Now we change of chat at a_0 to use the chart (U_{i_1}, φ_{i_1}) and we continue [3].

□
THOooQEFUooQTtPDD

Theorem 57.17 ([832]).
A smooth manifold is orientable if and only if it accepts a volume form [4].

Proof. Two parts.

(1) $\Leftarrow$ Let $\omega \in \Omega^n(M)$ be such that $\omega_x \neq 0$ for every $x \in M$. Let $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an atlas of M. We will build an atlas $\{(V_x, \phi_x)\}_{x\in M}$ of M indexed by the elements of M.
Let $x \in M$. There exists $\alpha \in \Lambda$ such that $x \in \varphi_\alpha(U_\alpha)$. With that chart we have (lemma 57.14)

$$\omega_y = f(y)\partial^*_{\alpha,1} \wedge \dots \wedge \partial^*_{\alpha,n} \tag{57.25}$$

for every $y \in M_\alpha$. Since ω is non vanishing, $f(x) \neq 0$. If $f(x) > 0$, we consider $(V_x, \phi_x) = (U_\alpha, \varphi_\alpha)$. If $f(x) < 0$ we have to twist a little bit. We consider the linear map $\tau\colon \mathbb{R}^n \to \mathbb{R}^n$ which permutes e_1 and e_2, that is:

$$\begin{aligned} \tau\colon \mathbb{R}^n &\to \mathbb{R}^n \\ e_i &\mapsto \begin{cases} e_2 & \text{if } i = 1 \\ e_1 & \text{if } i = 2 \\ e_i & \text{otherwise.} \end{cases} \end{aligned} \tag{57.26}$$

Notice that $\tau^{-1} = \tau$. We set $V_x = \tau(U_\alpha)$ and $\phi_x = \varphi_\alpha \circ \tau$. Following the notations of 57.6, we have

$$(\partial_{x,1})(y) = \frac{d}{dt}\left[\phi_x\left(\phi_x^{-1}(y) + te_1\right)\right]_{t=0} \tag{57.27a}$$

$$= \frac{d}{dt}\left[(\varphi_\alpha \circ \tau)\left((\tau \circ \varphi_\alpha^{-1})(y) + te_1\right)\right]_{t=0} \tag{57.27b}$$

$$= \frac{d}{dt}\left[\varphi_\alpha\left(\varphi_\alpha^{-1}(y) + te_2\right)\right]_{t=0} \tag{57.27c}$$

$$= (\partial_\alpha, 2)(y). \tag{57.27d}$$

3. The reader is assumed to be able to write a correct recursion.
4. Volume form, definition 57.13.

In the same way we have $\partial_{x,1} = \partial_{\alpha,1}$ and $\partial_{x,i} = \partial_{\alpha,i}$ for $i > 2$. Thus we have

$$\omega_x = f(x)\partial^*_{\alpha,1} \wedge \partial^*_{\alpha,2} \wedge \ldots \wedge \partial^*_{\alpha,n} = f(x)\partial^*_{x,2} \wedge \partial^*_{x,1} \wedge \ldots \wedge \partial^*_{x,n} = -f(x)\partial^*_{x,1} \wedge \ldots \wedge \partial^*_{x,n}. \quad (57.28)$$

Thus using the atlas $\{(V_x, \phi_x)\}_{x\in M}$, for each $y \in M$, we have a chart (V_y, ϕ_y) and a smooth map $g_y \colon M_y \to \mathbb{R}$ such that

$$\omega_x = g_y(x)\partial^*_{x,1} \wedge \ldots \wedge \partial^*_{x,n} \quad (57.29)$$

with $g_y(x) > 0$ for every $x \in M_y$.
Now we prove that $\{(V_x, \phi_x)\}$ is an orientation of M. Let $z \in M_{xy} = \phi_(V_x) \cap \phi_y(V_y)$. We have

$$\omega_z = g_x(z)\partial^*_{x,1} \wedge \ldots \partial^*_{x,n} = g_y(z)\partial^*_{y,1} \wedge \ldots \wedge \partial^*_{y,n}. \quad (57.30)$$

Applying to $(\partial_{x,1}, \ldots, \partial_{x,n})$ we have

$$\omega_z(\partial_{x,1}, \ldots, \partial_{x,n}) = g_x(z) = g_y(z)(\partial^*_{y,1} \wedge \ldots \wedge \partial^*_{y,n})(\partial_{x,1}, \ldots, \partial_{x,n}). \quad (57.31)$$

Using proposition 10.214 and lemma 49.95 on the right hand side, we get

$$g_x(z) = g_y(z) \det\left(\partial^*_{y,i}(\partial_{x,j})\right) = g_y(z) \det\left(d(\varphi_y^{-1} \circ \varphi_x)_z\right). \quad (57.32)$$

The charts φ_x and φ_y were chosen in such a way that $g_x(z) > 0$ and $g_y(z) > 0$. Thus

$$\det\left(d(\varphi_y^{-1} \circ \varphi_x)_z\right) > 0. \quad (57.33)$$

Thus the maps $\{\varphi_x\}_{x\in M}$ are an atlas and an orientation.

(2) $\Rightarrow$ Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an orientation on M. For each $\alpha \in \Lambda$ we consider the differential form

$$\omega_\alpha = \partial^*_{\alpha,1} \wedge \ldots \wedge \partial^*_{\alpha,n}. \quad (57.34)$$

We consider a partition of unity as in proposition 49.45, and we set $\omega = \sum_{\alpha\in\Lambda} \rho_\alpha\omega_\alpha$. We show that this is nowhere vanishing. Let $m \in M$. Since $\{\mathrm{supp}(\rho_\alpha)\}_{\alpha\in\Lambda}$ is locally finite, the set $\Lambda' = \{\alpha \in \Lambda \text{ tel que } \rho_\alpha(m) \neq 0\}$ is finite. We have

$$\omega(m) = \sum_{\alpha\in\Lambda'} \rho_\alpha(m)\omega_\alpha(m). \quad (57.35)$$

Let $\beta \in \Lambda'$. We have

$$\omega_m(\partial_{\beta,1}, \ldots, \partial_{\beta,n}) = \sum_{\alpha\in\Lambda'} \rho_\alpha(m)\omega_\alpha(\partial_{\beta,1}, \ldots, \partial_{\beta,n}) \quad (57.36a)$$

$$= \sum_{\alpha\in\Lambda'} \rho\alpha(m) \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_m\right). \quad (57.36b)$$

Since $\{\varphi_\alpha\}_{\alpha\in\Lambda}$ is an orientation, we have $\det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_m\right) > 0$ for every $\alpha \in \Lambda'$, and since we have only a finite number of them, we consider the minimum:

$$s = \min_{\alpha\in\Lambda'} \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_m\right). \quad (57.37)$$

Now we have

$$\sum_{\alpha\in\Lambda} \rho_\alpha(m)\omega_\alpha(m)(\partial_{\beta,1}, \ldots, \partial_{\beta,n}) = \sum_{\alpha\in\Lambda'} \rho_\alpha(m) \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_m\right) \quad (57.38a)$$

$$\geqslant s \sum_{\alpha\in\Lambda'} \rho_\alpha(m) \quad (57.38b)$$

$$= s \sum_{\alpha\in\Lambda} \rho_\alpha(m) \quad (57.38c)$$

$$= s > 0. \quad (57.38d)$$

□

LEMooHLYAooAeEnZb

Lemma 57.18 ([1]).
Let $\mathcal{A}$ and $\mathcal{A}'$ be orientations on a connected manifold. Let ω be a $\mathcal{A}$-oriented volume form. The volume form ω is $\mathcal{A}'$-oriented if and only if $\mathcal{A}$ and $\mathcal{A}'$ are equivalent orientations.

Proof. Let $(\varphi_\alpha, U_\alpha) \in \mathcal{A}$ and $(U_\beta, \varphi_\beta) \in \mathcal{A}'$. We can write $\omega_x = f(x)\partial^*_{\alpha,1} \wedge \dots \partial^*_{\alpha,n}$ with $f(x) > 0$, and

$$\begin{aligned} \omega_x(\partial_{\beta,1}, \dots, \partial_{\beta,n}) &= f(x)(\partial^*_{\alpha,1} \wedge \dots \partial_{\alpha,n})(\partial_{\beta,1}, \dots, \partial_{\beta,n}) && \text{(57.39a)} \\ &= f(x) \det\big(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_x\big). \qquad \text{lem. 49.95} && \text{(57.39b)} \end{aligned}$$

Thus we have equivalence between :

(1) ω is $\mathcal{A}'$-oriented,
(2) $\omega_x(\partial_{\beta,1}, \dots, \partial_{\beta,n}) > 0$
(3) $\det\big(d(\varphi_\alpha^{-1} \circ \varphi_\beta)\big) > 0$,
(4) the orientations $\mathcal{A}$ and $\mathcal{A}'$ are equivalent.

□

PROPooNCNJooHFngBW

Proposition 57.19 ([920, 921, 1]).
If a connected manifold accepts one orientation, then it accepts exactly two classes of orientation.

Proof. Let $\mathcal{A}$ and $\mathcal{A}'$ be two different orientations. Let ω be a $\mathcal{A}$-oriented volume form, and $\mathcal{B}$ a third orientation. Since ω is $\mathcal{A}$-oriented and since $\mathcal{A}'$ is not equivalent to $\mathcal{A}$, the volume form ω is $\mathcal{A}'$-negative.

If ω is $\mathcal{B}$-oriented, then $\mathcal{B} \sim \mathcal{A}$ (lemma 57.18). If ω is $\mathcal{B}$-negative, then $\mathcal{B}$ is equivalent to $\mathcal{A}'$.

We have shown that the orientation $\mathcal{B}$ is equivalent to $\mathcal{A}$ or $\mathcal{A}'$. □

57.2 Integral of differential forms

PROPooGVQDooNwnGXs

Proposition 57.20 ([920]).
Let M be a smooth orientable manifold. Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an orientation on M. We consider a differential form ω supported on $S \subset M$. Let $\alpha, \beta \in \Lambda$ such that $S \subset \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$. We define the maps $f\colon M_\alpha \to \mathbb{R}$ and $g\colon M_\beta \to \mathbb{R}$ such that

$$\omega_x = f(x)\partial^*_{\alpha,1} \wedge \dots \wedge \partial^*_{\alpha,n} \tag{57.40a}$$
$$\omega_x = g(x)\partial^*_{\beta,1} \wedge \dots \wedge \partial^*_{\beta,n}. \tag{57.40b}$$

Then we have

$$\int_{U_\alpha} (f \circ \varphi_\alpha) = \int_{U_\beta} g \circ \varphi_\beta. \tag{57.41}$$

Proof. This is a change of variable in the integral: theorem 14.281(3). First we have

$$\int_{U_\alpha} f \circ \varphi_\alpha = \int_{\varphi_\alpha^{-1}(S)} f \circ \varphi_\alpha, \tag{57.42}$$

and since

$$\varphi_\alpha^{-1}(S) = (\varphi_\alpha^{-1} \circ \varphi_\beta)\big(\varphi_\beta^{-1}(S)\big), \tag{57.43}$$

we try the change of variable with

$$\begin{aligned} \phi\colon \varphi_\beta^{-1}(S) &\to \varphi_\alpha(S) \\ \phi &= \varphi_\alpha^{-1} \circ \varphi_\beta. \end{aligned} \tag{57.44}$$

As far as the domain are concerned,

$$\phi^{-1}\big(\varphi_\alpha^{-1}(S)\big) = \varphi_\beta^{-1}\big((\varphi_\alpha \circ \varphi_\alpha^{-1})(S)\big) = \varphi_\beta^{-1}(S). \tag{57.45}$$

For the integral we have

$$\int_{\varphi_\alpha^{-1}(S)} s \circ \varphi_\alpha = \int_{\varphi_\beta^{-1}(S)} (f \circ \varphi_\beta) J_\phi \tag{57.46}$$

where J_ϕ is the function

$$J_\phi(u) = \det\big(d\phi_u\big). \tag{57.47}$$

Now we work out the link between f and g with lemma 49.95:

$$g(x) = \omega_x(\partial_{\beta,1}, \dots, \partial_{\beta,n}) = f(x) \det\big(d\phi_{\varphi_\beta^{-1}(x)}\big). \tag{57.48}$$

Thus we have

$$(g \circ \varphi_\beta)(u) = (f \circ \varphi_\beta)(u) \det\big(d\phi_u\big) = (f \circ \varphi_\beta)(u) J_\phi(u), \tag{57.49}$$

and we have the result. □

57.2.1 Integration over a chart

DEFooIKFYooKxKidt

Definition 57.21 (Integral over a chart[834])**.**
Let M be a smooth oriented manifold and ω, a n-form on M. Let $\varphi_\alpha : U_\alpha \to M$ be a positively oriented chart. We consider the map $f_\alpha : \varphi_\alpha(U_\alpha) \to \mathbb{R}$ of lemma 57.14 and We define

$$\int_\alpha \omega = \int_{U_\alpha} f \circ \varphi_\alpha. \tag{57.50}$$

If φ_α belong to the other orientation, we define

$$\int_\alpha \omega = -\int_{U_\alpha} f \circ \varphi_\alpha. \tag{57.51}$$

PROPooXTVAooVvEiTh

Proposition 57.22 ([1])**.**
Let $\varphi_\alpha : U_\alpha \to M$ and $\varphi_\beta : U_\beta \to M$ be two charts. Let $\omega \in \Omega(M)$. If the support of ω is compact inside $\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$, then

$$\int_\alpha \omega = \int_\beta \omega. \tag{57.52}$$

PROPooDJPYooEnXUAc

Proposition 57.23 (Linearity of the integral[1])**.**
Let ω_1 and ω_2 be two smooth forms on M. Let $\lambda \in \mathbb{R}$. We have

$$\int_\varphi (\omega_1 + \omega_2) = \int_\varphi \omega_1 + \int_\varphi \omega_2 \tag{57.53}$$

and

$$\int_\varphi \lambda\omega_1 = \lambda \int_\varphi \omega_1. \tag{57.54}$$

57.2.2 Integration over a compact part

Let M be an oriented manifold. We denote by $\Omega_c^n(K)$ the set of smooth n-forms on M with compact support $K \subset M$.

PROPooOSOCooUkalzR

Proposition 57.24 ([922]).
There exists one and only one linear form[5] $I\colon \Omega^n_c(M) \to \mathbb{R}$ *such that for every oriented chart* (U,φ) *and every* ω *with compact support in* $\varphi(U)$,

$$I(\omega) = \int_\varphi \omega \tag{57.55}$$

where the integral is the one of definition 57.21.

Proof. Two parts.

(1) **Unicity** Let I be such linear form and $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ be an oriented atlas. We consider a partition of unity $\{\rho_\alpha\}_{\alpha\in\Lambda}$ subordinated to $\{U_\alpha\}_{\alpha\in\Lambda}$.
Let $\omega \in \Omega^n_c(M)$. For each $m \in \operatorname{supp}(\omega)$, we have a open neighbourhood $\mathcal{O}_m$ of m such that $\Lambda_m = \{\alpha \in \Lambda \text{ tel que } \operatorname{supp}(\rho_\alpha) \cap \mathcal{O}_m \neq \varnothing\}$ is finite. These parts $\mathcal{O}_m$ make an open covering of $\operatorname{supp}(\omega)$ which is compact. We extract a finite covering $\{\mathcal{O}_m\}_{m=m_1,\dots,m_l}$. Then if $m \in \operatorname{supp}(\omega)$ we have $\rho_\alpha(m) \neq 0$ only if $\alpha \in \Lambda_\omega = \Lambda_{m_1} \cup \dots \Lambda_{m_l}$ which is finite.
For every $m \in M$ we have

$$\omega_m = \sum_{\alpha\in\Lambda} \rho_\alpha(m)\omega_m = \sum_{\alpha\in\Lambda_\omega} \rho_\alpha(m)\omega_m. \tag{57.56}$$

EQooZXHEooZfiSrw

Notice that $\rho_\alpha\omega$ has its support in $\varphi_\alpha(U_\alpha)$, so that

$$I(\rho_\alpha\omega) = \int_{U_\alpha} \rho_\alpha\big(\varphi_\alpha(x)\big)(f_\alpha\circ\varphi_\alpha)(x)dx \tag{57.57}$$

if $\omega_m = f_\alpha(m)\partial^*_{\alpha,1}\wedge\dots\wedge\partial^*_{\alpha,n}$. Since I is linear and the sum (57.56) is finite we have

$$I(\omega) = \sum_{\alpha\in\Lambda_\omega} I(f_\alpha\omega). \tag{57.58}$$

EQooYOBWooAxciMY

This shows the unicity of I because its value is fixed by the formula (57.58).

(2) **Existence** Let $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ be an oriented atlas. We show that the map

$$\begin{aligned} I\colon \Omega^n_c(M) &\to \mathbb{R} \\ \omega &\mapsto \sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega \end{aligned} \tag{57.59}$$

satisfy the required properties.
Let $\varphi\colon U \to M$ be a chart and ω be compactly supported in $\varphi(U)$. Since $\operatorname{supp}(\omega)$ is compact, there is a finite part $\Lambda_0 \subset \Lambda$ such that $\omega = \sum_{\alpha\in\Lambda_0}\rho_\alpha\omega$. We have

$$\int_\varphi \omega = \int_\varphi \sum_{\alpha\in\Lambda_0} \rho_\alpha\omega \tag{57.60a}$$

$$= \sum_{\alpha\in\Lambda_0}\int_\varphi \rho_\alpha\omega \qquad \text{prop. 57.23} \tag{57.60b}$$

$$= \sum_{\alpha\in\Lambda_0}\int_\alpha \rho_\alpha\omega \qquad \text{prop. 57.22} \tag{57.60c}$$

$$= I(\omega). \tag{57.60d}$$

The map I is linear because the integral is (proposition 57.23) and the sum over Λ can always be reduced to a finite sum in the case of a compactly supported form.

□

5. A linear form is not intended to be continuous.

57.2.3 Non compact

The following definition is inspired from the one of infinite sums, definition 10.112.

PROPooLCYGooLUxEYS

Proposition-Definition 57.25 ([1]).
Let $\omega \in \Omega(M)$. There exists at most one $v \in \mathbb{R}$ such that for every $\epsilon > 0$ there exists a compact L such that

$$\left|I(\omega_{|_K}) - v\right| < \epsilon \tag{57.61}$$

for every compact K such that $L \subset K$.

If such a v exists, we define

$$\int_M \omega \tag{57.62}$$

as being this number.

Proof. Let $v, w \in \mathbb{R}$ with the given property. Let $\epsilon > 0$. There exists compacts L_1 and L_2 in M such that

$$\forall \text{ compact } K \text{ tel que } L_1 \subset K,\ \left|I(\omega_{|_K} - v)\right| < \epsilon \tag{57.63a}$$

$$\forall \text{ compact } K \text{ tel que } L_2 \subset K,\ \left|I(\omega_{|_K} - w)\right| < \epsilon \tag{57.63b}$$

We consider the compact $L = L_1 \cup L_2$. Since $L_1 \subset L$ we have $\left|I(\omega_{|_L}) - v\right| < \epsilon$. Since $L_2 \subset L$ we have $\left|I(\omega_{|_L}) - w\right| < \epsilon$. Thus we have

$$|v - w| \leqslant \left|v - I(\omega_{|_L})\right| + \left|w - I(\omega_{|_L})\right| < 2\epsilon. \tag{57.64}$$

This proves that $v = w$. □

PROPooQTBTooKXJDhx

Proposition 57.26 ([1]).
Let M be a smooth manifold and $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an orientation over M. We consider a partition of the unity $\{\rho_\alpha\}_{\alpha\in\Lambda}$ subordinated to the open covering $\{\varphi_\alpha(U_\alpha)\}$.

(1) If the sum $\sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega$ converges, then

$$\int_M \omega = \sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega. \tag{57.65}$$

Proof. We have to prove that the sum $\sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega$ satisfy the properties of the definition 57.25.

(1) **Some definitions**
Let $\epsilon > 0$. On the one hand, there exists a compact L such that for every compact K containing L,

$$\left|\int_M \omega - I(\omega_{|_K})\right| < \epsilon. \tag{57.66}$$

One the other hand, there exists a finite part $\Lambda' \subset \Lambda$ such that

$$\left|\sum_{\alpha\in\Lambda''}\int_M \rho_\alpha\omega - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| < \epsilon \tag{57.67}$$

for every Λ'' containing Λ'.
Let $\Lambda_1 \subset \Lambda$ be such that

$$L \subset \operatorname{supp}\Big(\sum_{\alpha\in\Lambda_1} \rho_\alpha\omega\Big). \tag{57.68}$$

We set

$$\Lambda_2 = \Lambda_1 \cup \Lambda' \tag{57.69}$$

and

$$K = \bigcup_{\alpha\in\Lambda_2} \operatorname{supp}(\rho_\alpha). \tag{57.70}$$

The part $K \subset M$ is compact as finite union of compact parts.

(2) $L \subset K$ The definition of Λ_1 makes $L \subset \text{supp}\left(\sum_{\alpha\in\Lambda_1} \rho_\alpha\omega\right)$. Let $m \in M$ be such that

$$\sum_{\alpha\in\Lambda_1} \rho_\alpha(m\phi_m) \neq 0. \tag{57.71}$$

There exists $\alpha_0 \in \Lambda_1 \subset \Lambda_2$ such that $\rho_{\alpha_0}(m) \neq 0$. Then we have $m \in \text{supp}(\rho_{\alpha_0}) \subset K$. In other words,

$$\{m \in M \text{ tel que } \sum_{\alpha\in\Lambda_1} \rho_\alpha(m)\omega_m \neq 0\} \subset K. \tag{57.72}$$

Taking the closure on both sides,

$$\text{supp}\left(\sum_{\alpha\in\Lambda_1} \rho_\alpha\omega\right) \subset \bar{K} = K. \tag{57.73}$$

(3) **Definition ofΛ_3** Let

$$\Lambda_3 = \{\alpha \in \Lambda \text{ tel que } \text{supp}(\rho_\alpha) \cap K \neq \varnothing\}. \tag{57.74}$$

This is finite because K is compact and $\{\rho_\alpha\}$ is locally finite. We have

$$\omega_{|_K} = \sum_{\alpha\in\Lambda_3} \rho_\alpha\omega. \tag{57.75}$$

(4) $\Lambda' \subset \Lambda_3$ We prove more: we prove that $\Lambda_2 \subset \Lambda_3$. If $\alpha \in \Lambda_2$, we have $\text{supp}(\rho_\alpha) \subset K$ and then $\alpha \in \Lambda_3$.

(5) **The computation** We have

$$\left|\int_M \omega - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| \leqslant \left|\int_M \omega - I(\omega_{|_K})\right| + \left|I(\omega_{|_K}) - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right|. \tag{57.76}$$

Since $L \subset K$, the first term is lesser than ϵ. For the second term,

$$\left|I(\omega_{|_K}) - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| = \left|\sum_{\alpha\in\Lambda_3} I(\rho_\alpha\omega) - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| \leqslant \epsilon \tag{57.77}$$

because $\Lambda' \subset \Lambda_2 \subset \Lambda_3$ and $I(\rho_\alpha\omega) = \int_M \rho_\alpha\omega$. Finally we have

$$\left|\int_M \omega - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| \leqslant \left|\int_M \omega - I(\omega_{|_K})\right| + \left|I(\omega_{|_K}) - \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega\right| \leqslant 2\epsilon. \tag{57.78}$$

This proves that

$$\int_M \omega = \sum_{\alpha\in\Lambda}\int_M \rho_\alpha\omega. \tag{57.79}$$

□

57.2.4 Non compact (another try)

This subsection contains some more work about integrating non compactly supported forms. Can be skipped.

LEMooEKYAooERjAnc

Lemma 57.27 ([1]).
Let ω be a n-differential form over the smooth manifold M. Let $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an orientation on M. We suppose that[6]

$$\omega_m = f_\alpha(m)\partial^*_{\alpha,1} \wedge \ldots \wedge \partial^*_{\alpha,n} \tag{57.80}$$

with $f_\alpha(m) \geqslant 0$ for every $m \in \varphi_\alpha(U_\alpha)$.

Then if $m \in \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$ we have

$$\omega_m = f_\beta(m)\partial^*_{\beta,1} \wedge \ldots \wedge \partial^*_{\beta,n} \tag{57.81}$$

with $f_\beta(m) \geqslant 0$.

6. Lemma 57.14 for the existence.

Proof. Using lemma 49.95 we have

$$f_\beta(m) = \omega_m(\partial_{\beta,1}, \dots, \partial_{\beta,n}) \tag{57.82a}$$
$$= f_\alpha(m)(\partial_{\alpha,1} \wedge \dots \wedge \partial_{\alpha,n})(\partial_{\beta,1}, \dots, \partial_{\beta,n}) \tag{57.82b}$$
$$= f_\alpha(m) \det\left(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{\varphi_\beta^{-1}(m)}\right). \tag{57.82c}$$

Since $\{\varphi_\alpha\}_{\alpha\in\Lambda}$ is an orientation, the determinant is positive. □

PROPooDYDPooOWNtfD

Proposition 57.28 ([1]).
Let M be an orientable manifold. We consider two oriented atlases with the same orientation

$$\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda} \tag{57.83a}$$
$$\mathcal{B} = \{(V_i, \phi_i)\}_{i\in I}. \tag{57.83b}$$

For each $\alpha \in \Lambda$ we define the map $f_\alpha \colon \varphi_\alpha(U_\alpha) \to \mathbb{R}$ by

$$\omega_m = f_\alpha(m)\partial_{\alpha,1}^* \wedge \dots \wedge \partial_{\alpha,n}. \tag{57.84}$$

We suppose that there exists a $\alpha \in \Lambda$ such that $f_\alpha \geqslant 0$.
Let $\{\rho_\alpha\}_{\alpha\in\Lambda}$ be a partition of the unity subordinated to $\mathcal{A}$ and $\{\sigma_i\}$ be subordinate to $\mathcal{B}$.
We suppose that

$$\sum_{\alpha\in\Lambda} \int_\alpha \rho_\alpha \omega < \infty. \tag{57.85}$$

EQooUOPVooUCTfof

for every $\alpha \in \Lambda$.
Then

ITEMooXJIJooZGNRuU

(1) For every $i \in I$ and $\beta \in \Lambda$ we have $f_i \geqslant 0$ and $f_\alpha \geqslant 0$.

ITEMooLLMCooFGNsYmi

(2) We have

$$\int_\alpha \rho_\alpha \omega = \sum_{i\in I} \int_\alpha \sigma_i \rho_\alpha \omega. \tag{57.86}$$

ITEMooGHRAooYZtFco

(3) We have

$$\int_\alpha \rho_\alpha \sigma_i \omega = \int_i \rho_\alpha \sigma_i \omega. \tag{57.87}$$

ITEMooJOJPooXfziTe

(4) We can commute the sums[7]*:*

$$\sum_{\alpha\in\Lambda} \sum_{i\in I} \int_\alpha \sigma_i \rho_\alpha \omega = \sum_{i\in I} \sum_{\alpha\in\Lambda} \int_\alpha \sigma_i \rho_\alpha \omega. \tag{57.88}$$

ITEMooUTQVooHafpOr

(5) We have the equality

$$\sum_{\alpha\in\Lambda} \int_\alpha \rho_\alpha \omega = \sum_{i\in I} \int_i \sigma_i \omega. \tag{57.89}$$

Proof. Several points.

(1) **For (1)** This is lemma 57.27 and the fact that $\{\varphi_\alpha, \phi_i\}_{\alpha,i}$ is an orientation (definition 57.3).

(2) **For (2)** Let fix $\alpha \in \Lambda$.

$$\int_\alpha \rho_\alpha \omega = \int_\alpha \left(\sum_{i\in I} \sigma_i\right) \rho_\alpha \omega. \tag{57.90}$$

Our first task if to commute the sum over $i \in I$ and the integral. We know that the partition of unity $\{\sigma_i\}$ is locally finite. For each $m \in \text{supp}(\rho_\alpha)$ we consider an open set $\mathcal{O}_m$ and a finite part $I_m \subset I$ such that $\text{supp}(\sigma_i) \cap \mathcal{O}_m \neq \varnothing$ only if $i \in I_m$.

7. I did not seen this part actually done anywhere. Is there a simpler way ? See [923].

These parts $\{\mathcal{O}_m\}_{m\in\text{supp}(\rho_\alpha)}$ make an open covering of the compact $\text{supp}(\rho_\alpha)$. Thus there exist $m_1,\dots,m_k\in\text{supp}(\rho_\alpha)$ such that $\text{supp}(\rho_\alpha)\subset\bigcup_{l=1}^k$. If we set $I_\alpha=\bigcup_{l=1}^k I_{m_l}$ we have

$$\int_\alpha \rho_\alpha\omega=\int_\alpha\sum_{i\in I}\sigma_i\rho_\alpha\omega=\int_\alpha\sum_{i\in I_\alpha}\sigma_i\rho_\alpha\omega=\sum_{i\in I_\alpha}\int_\alpha\sigma_i\rho_\alpha\omega=\sum_{i\in I}\int_\alpha\sigma_i\rho_\alpha\omega \tag{57.91}$$

because the sum over $i\in I_\alpha$ is finite.

(3) **For (3)** Now we prove that $\int_\alpha\sigma_i\rho_\alpha\omega=\int_i\sigma_i\rho_\alpha\omega$. The part $S=\text{supp}(\rho_\alpha)\cap\text{supp}(\sigma_i)$ satisfy $S\subset\varphi_\alpha(U_\alpha)\cap\phi_i(V_i)$. Proposition 57.20 then shows

$$\int_\alpha\rho_\alpha\sigma_i\omega=\int_{U_\alpha}(\rho_\alpha\sigma_i\omega)\circ\varphi_\alpha=\int_{V_i}(\rho_\alpha\sigma_i)\omega\circ\phi_i=\int_i\rho_\alpha\sigma_i\omega. \tag{57.92}$$

(4) **For (4)** Out task now is to commute two sums. We consider the counting measure on I and Λ as well as the function

$$\begin{aligned}g\colon\Lambda\times I&\to\mathbb{R}\\ \alpha,i&\mapsto\int_i\rho_\alpha\sigma_i\omega.\end{aligned} \tag{57.93}$$

We check that $f\in L^1(\Lambda\times I)$ by checking the inequality of corollary 14.303(2) :

$$\begin{aligned}\sum_{\alpha\in\Lambda}\left[\sum_{i\in I}|g(\alpha,i)|\right]&=\sum_{\alpha\in\Lambda}\left[\sum_{i\in I}|\int_{V_i}(\rho_\alpha\sigma_i)\big(\phi_i(x)\big)f_i\big(\phi_i(x)\big)dx|\right] && (57.94a)\\ &=\sum_{\alpha\in\Lambda}\left[\sum_{i\in I}\int_{V_i}(\rho_\alpha\sigma_i)\big(\phi_i(x)\big)f_i\big(\phi_i(x)\big)dx\right] && \text{s.b.}\quad (57.94b)\\ &=\sum_{\alpha\in\Lambda}\left[\sum_{i\in I}\int_i\rho_\alpha\sigma_i\omega\right] && (57.94c)\\ &=\sum_{\alpha\in\Lambda}\left[\sum_{i\in I}\int_\alpha\rho_\alpha\sigma_i\omega\right] && \text{item (3)}\quad (57.94d)\\ &=\sum_{\alpha\in\Lambda}\left[\int_\alpha\sum_{i\in I}\rho_\alpha\sigma_i\omega\right] && \text{s.b.}\quad (57.94e)\\ &=\sum_{\alpha\in\Lambda}\int_\alpha\rho_\alpha\omega && (57.94f)\\ &<\infty. && (57.94g)\end{aligned}$$

Indeed:

— For (57.94b), this is the hypothesis that $f_\alpha\geqslant 0$. By the way the whole point of this proposition and the definition of $\int_M\omega$ in two steps is for this very step [8].

— For (57.94e), we restrict the sum over i to a finite sum as in item (2).

This proves that $g\in L^1(\Lambda\times I)$. The theorem of Fubini 14.304 commutes the sums.

8. I saw this justification nowhere. If you can do simpler, please send me a message. See the question [923].

(5) **For (5)** This is a computation using all the other items:

$$\sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega = \sum_{\alpha\in\Lambda}\sum_{i\in I}\int_\alpha \sigma_i\rho_\alpha\omega \tag{57.95a}$$

$$= \sum_{i\in I}\sum_{\alpha\in\Lambda}\int_\alpha \sigma_i\rho_\alpha\omega \tag{57.95b}$$

$$= \sum_{i\in I}\sum_{\alpha\in\Lambda}\int_i \sigma_i\rho_\alpha\omega \tag{57.95c}$$

$$= \sum_{i\in I}\int_i\sum_{\alpha\in\Lambda} \sigma_i\rho_\alpha\omega \tag{57.95d}$$

$$= \sum_{i\in I}\int_i \sigma_i\omega. \tag{57.95e}$$

□

Proof. In several parts.

(1) **Integration on one chart** Let fix $\alpha\in\Lambda$. We prove that $\int_\alpha \sigma_i\rho_\alpha\omega = \int_{\alpha,i}\sigma_i\rho_\alpha\omega$. If

$$\omega_m = f(m)\partial^*_{\alpha,\wedge\{1,\dots,n\}}, \tag{57.96}$$

we have

$$(\sigma_i\rho_\alpha\omega)_m = \sigma_i(m)\rho_\alpha(m)f(m)\partial^*_{\alpha,1}\wedge\ldots\wedge\partial^*_{\alpha,n}, \tag{57.97}$$

and then

$$\int_\alpha \sigma_i\rho_\alpha\omega = \int_{U_\alpha}\sigma_i\big(\varphi_\alpha(x)\big)\rho_\alpha\big(\varphi_\alpha(x)\big)f\big(\varphi_\alpha(x)\big)dx. \tag{57.98}$$

Since $\sigma_i\big(\varphi_\alpha(x)\big)=0$ if $x\notin(\varphi_\alpha^{-1}\circ\phi_i)(V_i)$, the integration over U_α can be restricted to an integration over

$$U_\alpha\cap\varphi_\alpha^{-1}\big(\phi_i(V_i)\big) = \varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha)\cap\phi_i(V_i)\big) = W_{\alpha,i}. \tag{57.99}$$

Thus we have

$$\int_\alpha \sigma_i\rho_\alpha\omega = \int_{U_\alpha}\sigma_i\big(\varphi_\alpha(x)\big)\rho_\alpha\big(\varphi_\alpha(x)\big)f\big(\varphi_\alpha(x)\big)dx \tag{57.100a}$$

$$= \int_{W_{\alpha,i}}\sigma_i\big(\varphi_\alpha(x)\big)\rho_\alpha\big(\varphi_\alpha(x)\big)f\big(\varphi_\alpha(x)\big)dx \tag{57.100b}$$

$$= \int_{W_{\alpha,i}}\sigma_i\big(\psi_\alpha(x)\big)\rho_\alpha\big(\psi_\alpha(x)\big)f\big(\psi_\alpha(x)\big)dx \tag{57.100c}$$ SUBEQooMIDLooNpqaK

$$= \int_{\alpha,i}\sigma_i\rho_\alpha\omega. \tag{57.100d}$$

For (57.100c), we have $\varphi_\alpha(x)=\psi_{\alpha,i}(x)$ whenever $x\in W_{\alpha,i}$.

(2) **General case**

(2a) **Commute two sums** SPooALWRooAGbyMn

(2b) **Conclusion** We have

$$\sum_{\alpha\in\Lambda}\int_\alpha \rho_\alpha\omega = \sum_{\alpha\in\Lambda}\sum_{i\in I}\int_i \sigma_i\rho_\alpha\omega \tag{57.101a}$$

$$= \sum_{i\in I}\sum_{\alpha\in\Lambda}\int_i \sigma_i\rho_\alpha\omega \qquad \text{pt. (2a)} \tag{57.101b}$$

$$= \sum_{i\in I}\int_i\sum_{\alpha}\rho_\alpha\sigma_i\omega \qquad \text{cf. bellow} \tag{57.101c}$$ SUBEQooRHHNooXrXKCc

$$= \sum_{i\in I}\int_i \sigma_i\omega \tag{57.101d}$$

For (57.101c), we need to redo the work of point (2) with α instead of i.

□

57.3 Manifold with boundary

57.3.1 Some facts about $\mathbb{R}^n_+$

We already proved that $\mathbb{R}^n$ is C^∞-diffeomorphic to $B(0,1)$.
We write

$$\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } x_n \geqslant 0\}. \tag{57.102}$$

We consider the topology induced from $\mathbb{R}^n$, and the differential structure of definition 49.79.

LEMooBICBooFUQyYA

Lemma 57.29 ([1]).
The boundary[9] *of* $\mathbb{R}^n_+$ *is*

$$\partial\mathbb{R}^n_+ = \{x \in \mathbb{R}^n_+ \text{ tel que } x_n = 0\}. \tag{57.103}$$

PROPooPUREooSkeyxs

Proposition 57.30 ([924, 925, 1]).
There are no diffeomorphism[10] *between* $\mathbb{R}^n_+$ *and* $\mathbb{R}^n$.

Proof. We suppose that such a diffeomorphism exists. Let $f\colon \mathbb{R}^n_+ \to \mathbb{R}^n$ be a diffeomorphism. Let $a \in \mathbb{R}^n_+$ be such that $a_n = 0$. There exists open parts U of $\mathbb{R}^n_+$ and V of $\mathbb{R}^n$ and a diffeomorphism $\hat{f}\colon U \to V$ such that $a \in U$. We use the mean value theorem 10.271 on $\hat{f}$:

$$\hat{f}(a+h) = \hat{f}(a) + d\hat{f}_a(h) + \alpha(h). \tag{57.104}$$

EQooTKGQooQmlOhv

Since $U \cap \mathbb{R}^n_+$ is open in $\mathbb{R}^n_+ +$ and since the reciprocal of f is continuous, the part $f(U \cap \mathbb{R}^n_+)$ is open in $\mathbb{R}^n$ (and contains $f(a)$). Let $r > 0$ be such that

$$B\big(\hat{f}(a), r\big) = B\big(f(a), r\big) \subset f\big(U \cap \mathbb{R}^n_+\big). \tag{57.105}$$

From equation (57.104) we get

$$\|\hat{f}(a + \epsilon e_n) - \hat{f}(a)\| \leqslant \epsilon \|df_a\| + \|\alpha(\epsilon e_n)\|. \tag{57.106}$$

If ϵ is small enough we have

$$\hat{f}(a + \epsilon e_n) \in B\big(\hat{f}(a), r\big) \subset f\big(U \cap \mathbb{R}^n_+\big). \tag{57.107}$$

Let $\epsilon < 0$ small enough. There exists $u \in U \cap \mathbb{R}^n_+$ such that $f(u) = \hat{f}(a + \epsilon e_n)$. But on $U \cap \mathbb{R}^n_+$ we have $f = \hat{f}$, so that

$$\hat{f}(u) = \hat{f}(a + \epsilon e_n) \tag{57.108}$$

EQooSZLJooBjICmu

Notice that $u \neq a + \epsilon e_n$ because $u_n \geqslant 0$ while $(a + \epsilon e_n)_n = \epsilon < 0$. Equation (57.108) contradicts that $\hat{f}$ is bijective. Contradiction. Finished. □

LEMooEVSVooKJrgKz

Lemma 57.31 ([1]).
An open part of $\mathbb{R}^n_+$ *containing an element of* $\partial\mathbb{R}^n_+$ *is not open in* $\mathbb{R}^n$.

Proof. Let A be open in $\mathbb{R}^n_+$ and let $a \in A$ be such that $a_n = 0$. We suppose that A is open in $\mathbb{R}^n$. Then there exists $r > 0$ such that $B(a, r) \subset A$. The element $a - re_n$ belongs to $B(a, r)$ but $(a - re_n)_n = a_n - r < 0$. Thus $B(a, r)$ contains an element outside $\mathbb{R}^n_+$. Contradiction. □

LEMooWZOBooGrvNGQ

Lemma 57.32 ([1]).
Let X *be a topological space,* A *be open in* $\mathbb{R}^n_+$*, and* $a \in A$ *satisfying* $a_n > 0$. *We consider an homeomorphism* $f\colon A \to X$. *Then there exists a neighbourhood* U *of* a *in* $\mathbb{R}^n$ *such that* $U \subset A$ *and* $f\colon U \to X$ *is homeomorphic.*

9. Definition 7.125.
10. Definition 10.265. Bijection, differentiable, invertible with differentiable inverse in the sens of definition 49.79.

Proof. Several subsets.

(1) Since A is open in $\mathbb{R}^n_+$, there exists an open U_1 of $\mathbb{R}^n$ such that $A = U \cap \mathbb{R}^n_+$. There exists $r_1 > 0$ such that $B(a, r_1) \subset U_1$.

(2) Since the projection $x \mapsto x_n$ is continuous on $\mathbb{R}^n$, there exists an open neighbourhood U_2 of a in $\mathbb{R}^n$ such that $y_n > 0$ for every $y \in U_2$. Let $r_2 > 0$ be such that $B(a, r_2) \subset U_2$.

(3) The definition of continuity on a closed part (definition 49.79) says that there exists an open U_3 of $\mathbb{R}^n$ containing A and an homeomorphic extension $\tilde{f}\colon U_3 \to X$.

Taking $r = \min\{r_1, r_2\}$ we have

$$B(a, r) \subset U_1 \cap \mathbb{R}^n_+ = A \subset U_3. \tag{57.109}$$

The extension $\tilde{f}$ is thus defined on $B(a, r)$. And, since $B(a, r) \subset U_2$, we have $\tilde{f}(x) = f(x)$ for every $x \in B(a, r)$. Thus the map $f\colon B(a, r) \to X$ is an homeomorphism. □

LEMooLPXFooUBLCqA

Lemma 57.33 ([926]).
Let U and V be open parts of $\mathbb{R}^n_+$. Let $f\colon U \to V$ be a C^1-diffeomorphism[11]*. Then $f(x)_n = 0$ if and only if $x_n = 0$.*

Proof. Since U and V are open in $\mathbb{R}^n_+$, there are open parts A and B of $\mathbb{R}^n$ such that $U = A \cap \mathbb{R}^n_+$ and $V = B \cap \mathbb{R}^n_+$. We also consider the smooth map

$$\begin{aligned} \sigma\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto x_n. \end{aligned} \tag{57.110}$$

Now we prove the two parts.

(1) **Suppose $a_n > 0$** Since $\sigma(a) > 0$, there exists an open B' of $\mathbb{R}^n$ such that $\sigma > 0$ on B. Taking $B = B' \cap A$, we have

$$a \in B \subset A \tag{57.111}$$

where B is open of $\mathbb{R}^n$ and $\sigma|_B > 0$. What about $f(B)$? We use the smooth invariance of the domain theorem 49.80 with the manifolds $M = R^n$ and $N = \mathbb{R}^n$. Since B is open and $f\colon B \to f(B)$ is a C^1-diffeomorphism, the part $f(B)$ is open in $\mathbb{R}^n$.
We show that $f(y)_n > 0$ for every $y \in f(B)$. Suppose that $y \in f(B)$ satisfy $y_n = 0$. Since $f(B)$ is an open of $\mathbb{R}^n$, we have an open S of $\mathbb{R}^n$ satisfying $y \in S \subset f(B)$. Since $y_n = 0$, there are points z of S such that $z_n < 0$. This is a contradiction because $z \in S \subset f(B) \subset V \subset \mathbb{R}^n_+$.

(2) **Suppose $f(a)_n > 0$** Since f^{-1} is a C^1-diffeomorphism as well as f itself, we use the previous point to say that $f^{-1}\big(f(a)\big)_n > 0$. Thus we have $a_n > 0$.

□

LEMooYWBHooDTjCyf

Lemma 57.34 ([1]).
Let $f \in L^1(\mathbb{R}^n_+)$. We have

$$\int_{\mathbb{R}^n_+} f = \int_{\mathbb{R}^{n-1}} \left[\int_0^\infty f(x_1, \dots, x_{n-1}, t) dt \right] dx_1, \dots, dx_{n-1}, \tag{57.112}$$

EQooIYXTooJZpqhl

and

$$\int_{\mathbb{R}^n_+} f = \int_0^\infty \left(\int_{\mathbb{R}^{n-2}} \left[\int_{-\infty}^\infty f(x_1, \dots, x_{n-1}, t) dx_i \right] dx_1 \dots \widehat{dx_i} \dots dx_{n-1} \right) dt. \tag{57.113}$$

EQooSNGAooTSrQKX

LEMooZKWKooBwNywC

Lemma 57.35 ([1]).
Let $f\colon \mathbb{R}^n_+ \to \mathbb{R}$ be smooth with compact support. If $i \neq n$ we have

$$\int_{\mathbb{R}^n_+} (\partial_i f)(x) dx = 0. \tag{57.114}$$

EQooZHSPooDFLaxv

11. Definition 49.79.

In the other case,

$$\int_{\mathbb{R}^n_+} (\partial_n f)(x)dx = -\int_{\mathbb{R}^{n-1}} f(x_1,\dots,x_{n-1},0)dx_1\dots dx_{n-1}. \tag{57.115}$$

EQooZOYBooLyeqSl

Proof. Two parts.

(1) **For (57.114)** We apply the formula (57.113) to $\partial_i f$:

$$\int_{\mathbb{R}^n_+} (\partial_i f)(x)dx = \int_0^\infty \left(\int_{\mathbb{R}^{n-2}} \left[\int_{\mathbb{R}} (\partial_i f)(x_1,\dots,x_i,\dots,x_{n-1},t)dx_i \right] dx_1\dots\widehat{dx_i}\dots dx_{n-1} \right) dt. \tag{57.116}$$

The inner most integral is vanishing as the integral on $\mathbb{R}$ of a derivative of a compactly supported function.

(2) **For (57.115)** We start from (57.112):

$$\int_{\mathbb{R}^n_+} (\partial_n f)(x)dx = \int_{\mathbb{R}^{n-1}} \left(\int_0^\infty (\partial_n f)(x_1,\dots,x_{n-1},t)dt \right) dx_1\dots dx_{n-1}. \tag{57.117}$$

EQooBOIBooPrsqvJ

For fixed $(x_1,\dots,x_{n-1}) \in \mathbb{R}^{n-1}$ we set

$$\begin{aligned} \tilde{f}\colon [0,\infty[&\to \mathbb{R} \\ t &\mapsto f(x_1,\dots,x_{n-1},t). \end{aligned} \tag{57.118}$$

The inner integral of (57.117) is

$$\int_0^\infty \tilde{f}(t)dt = \lim_{R\to\infty} \tilde{f}(R) - \tilde{f}(0) = -\tilde{f}(0) \tag{57.119}$$

because $\tilde{f}$ has compact support. Thus we have

$$\int_{\mathbb{R}^n_+} (\partial_n f)(x)dx = -\int_{\mathbb{R}^{n-1}} f(x_1,\dots,x_{n-1},0)dx_1\dots dx_{n-1}. \tag{57.120}$$

□

57.3.2 Manifold with boundary

DEFooNETXooZaZMcm

Definition 57.36 (Topological manifold with boundary[834])**.**
*A topological space M is a **topological manifold with boundary** if it is Hausdorff and every element of M has a neighbourhood homeomorphic either to $\mathbb{R}^n$ or $\mathbb{R}^n_+$.*

*Let M be a manifold with boundary. The **boundary** of M is set of points of M which have no neighbourhood homeomorphic to $\mathbb{R}^n$. And we denote by ∂M the boundary of M.*

57.37.
The definition of ∂M is not a particular case of the general definition 7.125 of a boundary because every point of a topological space is in the interior of the space itself.

The reason of the definition is that the points of ∂M are the ones who are images of points in $\mathbb{R}^n_+$ that are on $\partial\mathbb{R}^n_+$ as we will see in proposition 57.33.

THOooKKBLooIrITBe

Theorem 57.38 (Smooth manifold with boundary[834])**.**
*Let M be a topological manifold with boundary. A **chart** is a map $\varphi\colon U \to M$ where U is either homeomorphic to $\mathbb{R}^n$ or $\mathbb{R}^n_+$* [12]

A set $\mathcal{A} = \{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ is an atlas if

12. In fact U could be only homeomorphic to $\mathbb{R}^n_+$, but it would complicate everything. In particular the characterisation of ∂M as the points with coordinates $x_n = 0$.

(1) For every $m \in M$, there exists $\alpha \in \Lambda$ such that $m \in \varphi_\alpha(U_\alpha)$,

(2) The maps

$$\varphi_\alpha^{-1} \circ \varphi_\beta \colon \varphi_\beta^{-1}(M_\alpha \cap M_\beta) \to U_\alpha \tag{57.121}$$

are homeomorphic in the sense of definition 49.79.

If these maps are smooth, the manifold M is a ***smooth manifold with boundary****.*

PROPooWQHTooXhkWdZ

Proposition 57.39 ([1])**.**
Let M be a smooth manifold with boundary and let $(U_\alpha, \varphi_\alpha)$ be a chart with $U_\alpha = \mathbb{R}^n_+$. Then $m \in \partial M \cap \varphi_\alpha(U_\alpha)$ if and only if $\varphi_\alpha^{-1}(m)_n = 0$.

Proof. Two parts.

(1) $\Rightarrow$ We suppose that $m \in \partial M \cap \varphi_\alpha(U_\alpha)$. Since $\varphi_\alpha^{-1}(m) \in \mathbb{R}^n_+$, we have $\varphi_\alpha^{-1}(m)_n \geqslant 0$. Suppose that $\varphi_\alpha^{-1}(m) > 0$. Let $a = \varphi_\alpha^{-1}(m)$. Since φ_α^{-1} is continuous, there exists a neighbourhood U of $a \in \mathbb{R}^n$ such that $y_n > 0$ for every $y \in U$. The restriction $\varphi_\alpha \colon U \to M$ is then a chart around m defined on a open part of $\mathbb{R}^n$, which is not possible because $m \in \partial M$.

(2) $\Leftarrow$ Suppose that $\varphi_\alpha^{-1}(m)_n = 0$. We have to prove that $m \in \partial M$, in other words, every chart around m has to be from $\mathbb{R}^n_+$; there exists no charts from an open set of $\mathbb{R}^n$ to a neighbourhood of m in M.

Let $\varphi_\beta \colon U_\beta \to M$ be an other chart around m. We write $M_{\alpha\beta} = \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$. By definition of a chart, the map

$$\varphi_\alpha^{-1} \circ \varphi_\beta \colon \varphi_\beta^{-1}(M_{\alpha\beta}) \to \varphi_\alpha^{-1}(M_{\alpha\beta}) \tag{57.122}$$

is a C^∞-diffeomorphism.

Let $q = \varphi_\beta^{-1}(m)$. We have

$$(\varphi_\alpha^{-1} \circ \varphi_\beta)(q)_n = \varphi_\alpha^{-1}(m)_n = 0, \tag{57.123}$$

so that $q_n = 0$ by lemma 57.33.

There are two possibilities : U_β is an open of $\mathbb{R}^n_+$ or U_β is an open of $\mathbb{R}^n$. These possibilities are not exclusive : $B(10, 1)$ is open in $\mathbb{R}^n$ and in $\mathbb{R}^n_+$.

(2a) **If U_β is open in $\mathbb{R}^n_+$** The part U_β is an open of $\mathbb{R}^n_+$ containing an element of $\partial\mathbb{R}^n_+$ (the point q). By lemma 57.31, the part U_β cannot be an open of $\mathbb{R}^n$. Thus U_β is $\mathbb{R}^n_+$.

(2b) **If U_β is open in $\mathbb{R}^n$** We set again $q = \varphi_\beta^{-1}(m)$. Using the proposition 12.180, we consider a C^∞-diffeomorphism $\psi \colon \mathbb{R}^n \to B(s, r)$ with $s_n > 0$ and $r < \|s\|$ (so $B(s, r)$ keeps in the interior of $\mathbb{R}^n_+$).

Since U_β is open in $\mathbb{R}^n$, there exists $\delta > 0$ such that $B(q, \delta) \subset U_\beta$. We choose δ small enough to have $\varphi_\beta\big(\big(B(q, \delta)\big) \subset \varphi_\alpha(U_\alpha)$. We write

(2b.ii) $A = \psi\big(B(q, \delta)\big)$,

(2b.ii) $U'_\alpha = (\varphi_\alpha^{-1} \circ \varphi_\beta) \circ \psi^{-1}(A)$.

The map

$$\sigma = \varphi_\alpha^{-1} \circ \varphi_\beta \circ \psi^{-1} \colon A \to U'_\alpha \tag{57.124}$$

is a C^∞-diffeomorphism between A which is open in $\mathbb{R}^n_+$ and U'_α which is open in $\mathbb{R}^n$. The point $a = \psi(q)$ belongs to A and satisfy

$$\sigma(a) = \sigma\big(\psi(q)\big) = \varphi_\alpha^{-1}(m). \tag{57.125}$$

Thus $\sigma(a)_n = \varphi_\alpha^{-1}(m)_n = 0$. This is a contradiction because the map ψ takes its values in $B(s, r)$ which does not intersect $\partial\mathbb{R}^n_+$.

□

PROPooTENAooDxIAbf

Proposition 57.40 ([834, 1]).
If M is a smooth manifold with boundary, then the boundary ∂M has a structure of smooth manifold.

Proof. Let $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda}$ be an atlas for M. We consider the "boundary" charts:

$$\Lambda' = \{\alpha \in \Lambda \text{ tel que } U_\alpha = \mathbb{R}^n_+\}, \tag{57.126}$$

and, for each $\alpha \in \Lambda'$, the maps

$$\begin{aligned} \phi_\alpha \colon r(U_\alpha) &\to \partial M \\ (x_1, \dots, x_{n-1}) &\mapsto \varphi_\alpha(x_1, \dots, x_{n-1}, 0) \end{aligned} \tag{57.127}$$

where r is the map

$$\begin{aligned} r \colon \mathbb{R}^n &\to \mathbb{R}^{n-1} \\ (x_1, \dots, x_n) &\mapsto (x_1, \dots, x_{n-1}). \end{aligned} \tag{57.128}$$

Defining

$$\begin{aligned} s \colon \mathbb{R}^{n-1} &\to \mathbb{R}^n \\ x' &\mapsto (x', 0) \end{aligned} \tag{57.129}$$

we have SUBEQSooMASYooFfkvRT

$$\phi_\alpha = \varphi_\alpha \circ s \tag{57.130a}$$
$$\phi_\alpha^{-1} = r \circ \varphi_\alpha^{-1}. \tag{57.130b}$$

We are going to use the theorem 49.10 to show that the maps $\{\big(r(U_\alpha), \phi_\alpha\big)\}_{\alpha\in\Lambda'}$ turn ∂M into a topological manifold.

(1) $r(U_\alpha)$ **is open in** $\mathbb{R}^{n-1}$ Since $U_\alpha = \mathbb{R}^n_+$, we have $r(U_\alpha) = \mathbb{R}^{n-1}$ which is open in $\mathbb{R}^{n-1}$.

(2) **Union is** ∂M Let $m \in \partial M$. There is a $\alpha \in \Lambda'$ such that $m = \varphi_\alpha(q)$ with $q \in U_\alpha$. From proposition 57.39, we have $q_n = 0$ and then $m = \phi_\alpha\big(r(q)\big)$.

(3) ϕ **is injective** Suppose $\phi_\alpha(x) = \phi_\alpha(y)$. This means

$$\varphi_\alpha(x_1, \dots, x_{n-1}, 0) = \varphi_\alpha(y_1, \dots, y_{n-1}, 0) \tag{57.131}$$

and then $(x_1, \dots, x_{n-1}, 0) = (y_1, \dots, y_{n-1}, 0)$ because φ_α is injective.

(4) $\phi_\alpha^{-1}(\dots)$ **is open in** U_α We have to prove that $\phi_\alpha^{-1}\Big(\phi_\alpha\big(r(U_\alpha)\big) \cap \phi_\beta\big(r(U_\beta)\big)\Big)$ is open in $r(U_\alpha) = \mathbb{R}^{n-1}$.
By hypothesis, the part

$$\varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \tag{57.132}$$

is open in U_α. Thus the part

$$A = (r \circ \varphi_\alpha^{-1})\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \tag{57.133}$$

is open in $\mathbb{R}^{n-1}$. Notice that

$$A = \phi_\alpha^{-1}\Big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\Big) = \phi_\alpha^{-1}\Big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta) \cap \partial M\Big) \tag{57.134}$$

because ϕ_α takes its values in ∂M. For $x \in U_\alpha$, there are two possibilities: either $x_n = 0$ or not. In the first case, $\varphi_\alpha(x) \in \partial M$, in the second case, $\varphi_\alpha(x) \notin \partial M$ (proposition 57.39). Defining $U'_\alpha = \{x \in U_\alpha \ tqx_n = 0\}$ and $U'_\beta = \{x \in U_\beta \text{ tel que } x_n = 0\}$ we have

$$A = \phi_\alpha^{-1}\Big(\varphi_\alpha(U'_\alpha) \cap \varphi_\beta(U'_\beta)\Big) \tag{57.135}$$

Let $(x',0)\in U'_\alpha$. We have

$$\varphi_\alpha(x',0)=\phi_\alpha(x')=\phi_\alpha\big(r(x',0)\big),\tag{57.136}$$

and the same with β instead of α. Thus

$$A=\phi_\alpha^{-1}\Big(\varphi_\alpha(U'_\alpha)\cap\varphi_\beta(U'_\beta)\Big)=\phi_\alpha^{-1}\Big(\phi_\alpha\big(r(U_\alpha)\big)\cap\phi_\beta\big(r(U_\beta)\big)\Big).\tag{57.137}$$

EQooWVGBooMzuLYh

Since A is open, the right hand side of (57.137) is open too.

(5) **Continuity and smoothness** Using the relations (57.130) we have

$$\phi_\alpha^{-1}\circ\phi_\beta=r\circ\varphi_\alpha^{-1}\circ\varphi_\beta\circ s.\tag{57.138}$$

Since $\varphi_\alpha^{-1}\circ\varphi_\beta$ is smooth, the whole is smooth.

Now the theorem 49.10 produces a structure of $n-1$-dimensional topological manifold for ∂M. The fact that the transition maps $\phi_\alpha^{-1}\circ\phi_\beta$ is smooth makes ∂M a smooth manifold. □

THOooENBHooOfBTlA

Theorem 57.41 (Induced orientation[834])**.**
Si $\{(U_\alpha,\varphi_\alpha)\}_{\alpha\in\Lambda}$ is an orientation on the n-dimensional smooth manifold M, then the maps $\{r(U_\alpha),\phi_\alpha\}_{\alpha\in\Lambda'}$ is an orientation on ∂M.

*If n is even, then this is the **induced orientation** is that one. If n is odd, the induced orientation is the other one*[13].

Proof. We have to check that

$$\det\big(d(\phi_\alpha^{-1}\circ\phi_\beta)_{\phi_\beta^{-1}(m)}\big)>0\tag{57.139}$$

for every $\mu\in\partial M$ for which it makes sense. We decompose the matrix as

$$d(\varphi_\alpha^{-1}\circ\varphi_\beta)_{(x_1,\dots,x_{n-1},0)}=\begin{pmatrix}\begin{pmatrix}&A&\end{pmatrix}&\begin{pmatrix}\uparrow\\v\\\downarrow\end{pmatrix}\\(\leftarrow\quad w\quad\rightarrow)&r\end{pmatrix}\tag{57.140}$$

where A is $(n-1)\times(n-1)$, v is $(n-1)\times 1$, w is $1\times(n-1)$ and $r\in\mathbb{R}$. We use the proposition 12.256 to study the parts of this matrix.

(1) **Part A** We have

$$A_{ij}=\partial_j(\varphi_\alpha^{-1}\circ\varphi_\beta)(x_1,\dots,x_{n-1},0)_i\tag{57.141}$$

with $i,j\leqslant n-1$. For every t we have

$$(\varphi_\alpha^{-1}\circ\varphi_\beta)\big((x_1,\dots,x_{n-1},0)+te_j\big)=(\phi_\alpha^{-1}\circ\phi_\beta)\big((x_1,\dots,x_{n-1})+te_j\big),\tag{57.142}$$

thus

$$A=d(\phi_\alpha^{-1}\circ\phi_\beta)_{(x_1,\dots,x_{n-1})}.\tag{57.143}$$

By the way, this is the part we really want to study.

(2) **Part v** No need to work out this part.

(3) **Part w** We have

$$w_i=\partial_i(\varphi_\alpha^{-1}\circ\varphi_\beta)(x_1,\dots,x_{n-1},0)_n\tag{57.144}$$

with $i\leqslant n-1$. For every t, we have $\varphi_\beta\big((x_1,\dots,x_{n-1},0)+te_i\big)\in\partial M$, so that

$$(\varphi_\alpha^{-1}\circ\varphi_\beta)\big((x_1,\dots,x_{n-1},0)+te_i\big)_n=0.\tag{57.145}$$

Thus $w_i=0$.

13. Proposition 57.19 says that if an orientation exists, there are exactly two classes of orientation.

(4) **Part r** Since $w = 0$ and the determinant of $d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{x_1,\dots,x_{n-1},0}$ is non zero (because $\{\varphi_\alpha\}_{\alpha\in\Lambda'}$) is an orientation, we must have $r \neq 0$.
We have

$$r = \partial_n(\varphi_\alpha^{-1} \circ \varphi_\beta)(x_1, \dots, x_{n-1}, 0)_n. \tag{57.146}$$

Let

$$\begin{aligned} p\colon [0,\epsilon] &\to \mathbb{R} \\ t &\mapsto (\varphi_\alpha^{-1} \circ \varphi_\beta)(x_1, \dots, x_{n-1}, t). \end{aligned} \tag{57.147}$$

We have $r = p'(0)$. We have $p(0) = 0$ ad $p(t) > 0$ for every $t > 0$. Thus $r = p'(0) \geqslant 0$. Since we already seen that $r \neq 0$, we have $r > 0$.

Putting everything together,

SUBEQooMBRTooViUTtT

$$0 < \det\big(d(\varphi_\alpha^{-1} \circ \varphi_\beta)_{(x_1,\dots,x_{n-1},0)}\big) \tag{57.148a}$$
$$= \det\big(d(\phi_\alpha^{-1} \circ \phi_\beta)_{(x_1,\dots,x_{n-1})}\big) r. \tag{57.148b}$$

The inequality (57.148a) is because φ_α and φ_β have same orientation. Since $r > 0$ we have

$$\det\big(d(\phi_\alpha^{-1} \circ \phi_\beta)_{(x_1,\dots,x_{n-1})}\big) > 0. \tag{57.149}$$

□

57.3.3 Tangent space

The tangent space of a manifold with boundary is still to be done. Maybe in [927].

CORooYCMSooLmSJVN

Corollary 57.42 ([1]).
Let M be a manifold with boundary. We consider a chart $\varphi_\alpha\colon U_\alpha \to M$. We have

$$\phi_\alpha\big(r(U_\alpha)\big) = \varphi_\alpha(U_\alpha) \cap \partial M. \tag{57.150}$$

Proof. Let $x = (x_1, \dots, x_n) \in U_\alpha$. We have $r(x) = (x_1, \dots, x_{n-1})$ and then

$$\phi_\alpha\big(r(x)\big) = \varphi_\alpha(x_1, \dots, x_{n-1}, 0). \tag{57.151}$$

Thus we have $\varphi_\alpha\Big(\phi_\alpha\big(r(x)\big)\Big) = (x_1, \dots, x_{n-1}, 0)$ and the proposition 57.39 makes $\phi_\alpha\big(r(x)\big) \in \partial M$. This proves the inclusion in one sense.

For the other inclusion, let $m \in \varphi_\alpha(U_\alpha) \cap \partial M$. We have $\varphi_\alpha^{-1}(m)_n = 0$. If we write $x = \varphi_\alpha^{-1}(m)$ we have

$$x = (x_1, \dots, x_{n-1}, 0). \tag{57.152}$$

Let $x' = (x_1, \dots, x_{n-1})$. We have

$$\phi_\alpha(x') = \varphi_\alpha(x', 0) = \varphi_\alpha(x) \in \varphi_\alpha(U_\alpha). \tag{57.153}$$

□

PROPooKBKAooKjORhq

Proposition 57.43 ([1]).
Let M be a manifold with boundary. We consider a vector $X \in T_mM$ with $m \in \partial M$. Let $\varphi_\alpha\colon U_\alpha \to M$ be a chart around m. We consider the corresponding chart $\phi_\alpha\colon r(U_\alpha) \to \partial M$. We have

$$\partial^*_{\varphi_\alpha,k}\big(di(X)\big) = \partial^*_{\phi_\alpha,k}(X). \tag{57.154}$$

Proof. We check with $X = \partial_{\phi_\alpha,l}$. Taking $x' = \phi_\alpha^{-1}(m) \in r(U_\alpha)$, we have

$$di(X) = \frac{d}{dt}\big[i\big(\phi_\alpha(x' + te_j)\big)\big]_{t=0} \tag{57.155a}$$
$$= \frac{d}{dt}\big[\varphi_\alpha(x' + te_l, 0)\big]_{t=0} \tag{57.155b}$$
$$= \partial_{\varphi_\alpha,l}. \tag{57.155c}$$

Thus we have

$$\partial_{\varphi_\alpha,k}\big(di(\partial_{\phi_\alpha,l})\big) = \partial^*_{\varphi_\alpha,k}(\partial_{\varphi_\alpha,l}) = \delta_{kl}. \tag{57.156}$$

Thus with $X = \sum_l X_l \partial_l$ we have

$$\partial_{\varphi_\alpha,k}\big(di(X)\big) = \sum_k X_k \delta_{kl} = X_k. \tag{57.157}$$

□

The following proposition is the equivalent of proposition 56.25 in the case of a manifold with boundary.

¡¡ Avertissement/question à la lectrice !! 57.44
I'm far from being sure of proposition 57.45. If you have a better statement and a proof, let me know.

PROPooTSAQooWwDHIf

Proposition 57.45 ([1]).
Let M be a manifold with boundary. If ω is a $n-1$-smooth differential form on M, for each chart $\varphi_\alpha : U_\alpha \to M$, there exists a smooth map such that

$$\omega_m = \sum_{i=1}^{n} f_{\alpha,i}(m)\partial^*_{\alpha,\wedge\{1,\dots,n\}\setminus\{i\}}. \tag{57.158}$$

If $i : \partial M \to M$ is the inclusion map, we have[14]

$$(i^*\omega)_m = f_{\alpha,\{1,\dots,n-1\}}(m)\partial^*_{\phi_\alpha,\wedge\{1,\dots,n-1\}}. \tag{57.159}$$

EQooGFNEooDMSVcI

Proof. I can prove the formula (57.159). Let $X_1, \dots, X_{n-1}$ be vector fields. We have

$$(i^*\omega)_m(X_1,\dots,X_{n-1}) = \sum_I f_{\alpha,I}(m)\partial_{\varphi_\alpha,\wedge I}\big(di(X_1),\dots,di(X_{n-1})\big) \tag{57.160a}$$

$$= \sum_I f_{\alpha,I}(m)\partial_{\phi_\alpha,\wedge I}(X_1,\dots,X_{n-1}) \qquad \text{prop. 57.43.} \tag{57.160b}$$

This is formula (57.159). □

57.4 Stokes theorem

Usually one writes the Stokes theorem as

$$\int_M d\omega = \int_{\partial M} \omega. \tag{57.161}$$

Obviously, the ω on the right hand side cannot be the same as the one on the left hand side since they are not even defined on the same manifold.

THOooFTNKooQzBPIN

Theorem 57.46 (Stokes[834, 1]).
Let M be a compact manifold of dimension n with boundary. Let ω be a smooth $n-1$-differential form on M. We consider the inclusion $i : \partial M \to M$. We have

$$\int_M d\omega = \int_{\partial M} i^*\omega. \tag{57.162}$$

Proof. Several parts.

(1) **If ω is supported in M_α with U_α open in $\mathbb{R}^n$** SPITEMooHFTEooAmNIJQ We suppose that ω is supported in $\varphi_\alpha(U_\alpha)$ where U_α is open in $\mathbb{R}^n$. In other words, the support of ω does not include points of ∂M.

14. The following formula is probably true, but its meaning is not yet clear form me because I do not have a definition of tangent bundle in the case of a manifold with boundary.

(1a) **On ∂M** We prove that $i^*\omega = 0$. For every $m \in \partial M$ we have

$$(i^*\omega)(X_1, \dots, X_{n-1}) = \omega_{i(m)}\big(di(X_1), \dots, di(X_{n-1})\big) = 0 \tag{57.163}$$

because $\omega_{i(m)} = 0$.

(1b) **On M** We prove that $\int_M d\omega = 0$. By proposition 56.25 we have

$$\omega_m = \sum_{I \in C_{n-1}} f_{\alpha,I}(m)\partial^*_{\alpha,\wedge I} \tag{57.164}$$

Where the sum over I is the sum over the multiindices of the form $\{1, \dots, n\}\backslash\{i\}$. The differential is

$$d\omega_m = \sum_I (df_{\alpha,I})_m \wedge \partial^*_{\alpha,\wedge I}. \tag{57.165}$$

Using proposition 49.72 we have

$$\begin{aligned} d\omega_m &= \sum_I \sum_k (\partial_k f_{\alpha,I})(m)\partial^*_{\alpha,k} \wedge \partial^*_{\alpha,\wedge I} && (57.166a)\\ &= \sum_{i=1}^n \sum_{k=1}^n (\partial_k f_{\alpha,\{1,\dots,n\}\backslash\{i\}})(m)\partial^*_{\alpha,k} \wedge \partial^*_{\alpha,\{1,\dots,n\}\backslash\{i\}} && (57.166b)\\ &= \sum_{i=1}^n (\partial_k f_{\alpha,\{1,\dots,n\}\backslash\{i\}})(m)\partial^*_{\alpha,i} \wedge \partial^*_{\alpha,\{1,\dots,n\}\backslash\{i\}} && (57.166c)\\ &= \sum_{i=1}^n (-1)^{i-1}(\partial_i f_{\alpha,\{1,\dots,n\}\backslash\{i\}})(m)\partial^*_{\alpha,\wedge\{1,\dots,n\}}. && (57.166d) \end{aligned}$$

Thus

$$\int_\alpha d\omega = \sum_i (-1)^{i-1} \int_{U_\alpha} (\partial_{\alpha,i} f_{\alpha,i})\big(\varphi_\alpha(x)\big)dx \tag{57.167}$$

where we write $f_{\alpha,i}$ for $f_{\alpha,\{1,\dots,n\}\backslash\{i\}}$. Since M is compact, the support of $f_{\alpha,i}$ is closed in a compact and then is compact. The support of $f_{\alpha,i} \circ \varphi_\alpha$ is compact in the open U_α of $\mathbb{R}^n$. In that case we can extend $f_{\alpha,i}$ by 0 outside U_α and we still have a smooth function. We have (dropping much indices for readability) :

$$(\partial_i f)\big(\varphi_\alpha(x)\big) = \frac{d}{dt}\big[f\big(\varphi_\alpha(x + te_i)\big)\big]_{t=0} = \partial_i(f \circ \varphi_\alpha)(x). \tag{57.168}$$

Notice that the ∂_i on the left hand side is the vector field while the ∂_i in the right hand side is an usual partial derivative. Finally we have

$$\begin{aligned} \int_\alpha d\omega &= \sum_i (-1)^{i-1} \int_{U_\alpha} \partial_i(f_{\alpha,i} \circ \varphi_\alpha)(x)dx && (57.169a)\\ &= 0 && (57.169b) \end{aligned}$$

SUBEQooDLSXooNCOYzN

by proposition 14.312.

(2) **Support in $\varphi_\alpha(U_\alpha)$ with $U_\alpha = \mathbb{R}^n_+$** Two parts.

SPITEMooHDCRooMfAROG

(2a) **On M** The whole computation up to equation (57.169a) is still valid:

$$\int_M d\omega = \sum_i (-1)^{i-1} \int_{U_\alpha} \partial_i(f_{\alpha,i} \circ \varphi_\alpha)(x)dx. \tag{57.170}$$

EQooBOGGooYaNEli

Since now $U_\alpha = \mathbb{R}^n_+$ is not open in $\mathbb{R}^n$, the proposition (14.312) does not apply.
We deal with the terms $i \neq n$ and $i = n$ separately using the formulas of lemma 57.35.
If $i \neq n$, the formula (57.114) produces

$$\int_{U_\alpha} \partial_i(f_{\alpha,i} \circ \varphi_\alpha)(x)dx = 0. \tag{57.171}$$

The sum in (57.170) reduces to the term $i = n$ and, using (57.115) we have

$$\int_{\mathbb{R}^n_+} \partial_n (f_{\alpha,n} \circ \varphi_\alpha)(x) dx = -\int_{\mathbb{R}^{n-1}} (f_{\alpha,n} \circ \varphi_\alpha)(x_1, \dots, x_{n-1}, 0) dx_1 \dots dx_{n)1}, \quad (57.172)$$

so that

$$\int_M d\omega = (-1)^n \int_{\mathbb{R}^{n-1}} (f_{\alpha,n} \circ \varphi_\alpha)(x_1, \dots, x_{n-1}, 0) dx_1 \dots dx_{n-1}. \quad (57.173)$$

EQooKAPUooQewdpN

(2b) **On ∂M** Since the support of ω is contained in $\varphi_\alpha(U_\alpha)$, the formula (57.159) says that the support of $i^*\omega$ is contained in $i^{-1}(M_\alpha)$. Using the fact that $i^{-1}(m) = \varnothing$ when $m \notin \partial M$ and corollary 57.42 we have

$$i^{-1}(M_\alpha) = i^{-1}(M_\alpha \cap \partial M) = M_\alpha \cap \partial M = \phi_\alpha\big(r(U_\alpha)\big). \quad (57.174)$$

Thus we can reduce the integral

$$\int_{\partial M} i^*\omega = (-1)^n \int_\alpha i^*\omega. \quad (57.175)$$

The factor $(-1)^n$ comes from the fact that the orientation of ϕ_α is positive or negative following the parity of n. See the definition 57.41. Using formula (57.159) and dropping some index,

$$\begin{aligned}
\int_{\partial M} \omega &= (-1)^n \int_\alpha i^*\omega && (57.176a)\\
&= (-1)^n \int_{r(U_\alpha)} f_\alpha \circ \phi_\alpha && (57.176b)\\
&= (-1)^n \int_{\mathbb{R}^{n-1}} (f_\alpha \circ \phi_\alpha)(x_1, \dots, x_{n-1}) dx_1 \dots dx_{n-1} && \text{s.b.} \quad (57.176c)\\
&= (-1)^n \int_{\mathbb{R}^{n-1}} (f_\alpha \circ \varphi_\alpha)(x_1, \dots, x_{n-1}, 0) dx_1 \dots dx_{n-1}. && (57.176d)
\end{aligned}$$

SUBEQooCLZZooLyJFLV

For (57.176c), we have $U_\alpha = \mathbb{R}^n_+$ and then $r(U_\alpha) = \mathbb{R}^{n-1}$. We got the same result as in equation (57.173).

(3) **General case** We consider a locally finite partition of unity $\{\rho_\alpha\}_{\alpha\in\Lambda}$. Since M is compact, the partition of unity is not only locally finite, but also finite. Let's see that. For each $m \in M$, we have an open neighbourhood A_m and a finite subset $\Lambda_m \subset \Lambda$ such that $\rho_\alpha \neq 0$ on A_m only if $\alpha \in \Lambda_m$. Since M is compact, we extract a finite subcovering $\{A_{m_i}\}_{i=1,\dots,k}$. If α does not belong to Λ_{m_i} for some i, then $\rho_\alpha = 0$ on M. For that reason, we can commute the integral and the sum over α in the following computation :

$$\begin{aligned}
\int_{\partial M} i^*\omega &= \int_{\partial M} \sum_\alpha \rho_\alpha (i^*\omega) && (57.177a)\\
&= \sum_\alpha \int_{\partial M} \rho_\alpha (i^*\omega) && (57.177b)\\
&= \sum_\alpha \int_M d\big(\rho_\alpha \omega\big) && \text{s.b.} \quad (57.177c)\\
&= \sum_\alpha \int_M d\rho_\alpha \wedge \omega + \sum_\alpha \int_M \rho_\alpha d\omega && (57.177d)\\
&= \sum_\alpha \int_M d\rho_\alpha \wedge \omega + \int_M d\omega && (57.177e)
\end{aligned}$$

SUBEQooOPYCooBSsmHJ

For (57.177c), we used one of the cases (1) or (2) following the nature of U_α. We have:

$$\sum_\alpha \int_M d\rho_\alpha \wedge \omega = \int_M \sum_\alpha (d\rho_\alpha \wedge \omega) \tag{57.178a}$$

$$= \int_M d(\sum_\alpha \rho_\alpha) \wedge \omega \tag{57.178b}$$

$$= 0. \tag{57.178c}$$

□

57.5 Integration of a differential form

I think that this whole section is useless because the matter is already done before.

57.5.1 Open set in $\mathbb{R}^n$

Let U be an open set of $\mathbb{R}^n$. A differential form of degree n over U can always be written under the form

$$\omega_x = f(x) dx_1 \wedge \ldots \wedge dx_n; \tag{57.179}$$

this is proposition 8.9.

DEFooEYRFooRQTmRF

Definition 57.47.
The integral of ω on U is

$$\int_U f\, dx_1 \wedge \ldots \wedge dx_n = \int_U f \tag{57.180}$$

The second integral is the integral of a function on $\mathbb{R}^n$, that is definition 14.167 where the measure is the Lebesgue measure on $\mathbb{R}^n$.

LEMooNCYSooXtnCKq

Lemma 57.48 (Change of variable)**.**
Let $f : V \to U$ be a diffeomorphism of open sets in $\mathbb{R}^n$ and ω be a n-form on U. Then we have

$$\int_U \omega = \int_{f^{-1}(U)} f^*\omega \tag{57.181}$$

if $\det(f) > 0$. A sign change if $\det(df) < 0$.

Proof. Let, for $y \in U$, write the form ω as $\omega_y = h(y) dy_1 \wedge \ldots \wedge dy_n$. Taking $v_i \in \Gamma(TV)$ we have

$$(f^*\omega)_x(v_1, \ldots, v_n) = \omega_{f(x)}\big(df_x v_1, \ldots, df_x v_n\big) \tag{57.182a}$$

$$= h\big(f(x)\big) \det \begin{pmatrix} df_x v_1 \\ \vdots \\ df_x v_n \end{pmatrix} \tag{57.182b}$$

$$= (h \circ f)(x) \det(df_x) \det \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \tag{57.182c}$$

$$= (h \circ f)(x) \det(df_x)(dx_1 \wedge \ldots \wedge dx_n)(v_1, \ldots, v_n). \tag{57.182d}$$

Thus

$$f^*\omega = (h \circ f) \det(df) dx_1 \wedge \ldots \wedge dx_n \tag{57.183}$$

Using the usual change of variable theorem 14.281(3) (and taking a sign if $\det(df) < 0$ because there is an absolute value in around the jacobian in (14.856)) :

$$\int_{f^{-1}(U)} f^*\omega = \int_V (h \circ f) \det(df) = \int_{f(V)} h = \int_U h = \int_U \omega. \tag{57.184}$$

□

That is for integrating a differential form on an open set of $\mathbb{R}^n$. In order to integrate on a manifold we "simply" use a pull-back with a chart system. There will be three complications

— If an atlas is made from more than one chart, what about the intersections ?

— Independence with respect to the choice of the chart.

— Integrating a vector field (that is not obviously a n-form).

57.5.2 One chart on a manifold

We suppose (M, g) to be a n-dimensional Riemannian manifold and S to be a $(n-1)$-dimensional submanifold. We suppose that both are inside only one chart

$$\phi\colon U \subset \mathbb{R}^n \to M \tag{57.185}$$

and

$$\varphi\colon A \subset \mathbb{R}^{n-1} \to S. \tag{57.186}$$

We also consider a differential form $\omega \in \bigwedge^n(T^*M)$ and $\sigma \in \bigwedge^{n-1}(T^*M)$. These are respectively n and $n-1$ differential forms on M. We also consider v, a vector field on M and τ, a 1-form on M.

Let us see what is possible to integrate.

DEFooPDRCooPiBklC

Definition 57.49 ([928])**.**
Let ω be a n-form defined on $\phi(U)$ (vanishing everywhere else). Its integral is :

$$\int_{\phi(U)} \omega = \int_U \phi^*\omega. \tag{57.187}$$

The last integral is an integral of type $\int_U F(x_1, \dots, x_n)dx_1 \wedge \dots \wedge dx_n$ on an open set in $\mathbb{R}^n$. That is definition 57.47.

This definition is nothing if it depend on the parametrisation. The following proposition show slightly more than the independence.

PROPooNJCLooMqeeeX

Proposition 57.50 ([1, 929])**.**
Let be the charts $\phi\colon U \to M$ and $\psi\colon V \to N$ and a map $f\colon M \to N$. The whole is supposed to be minimal :

$$f\big(\phi(U)\big) = \psi(V). \tag{57.188}$$

Then we have the "change of variable" formula :

$$\int_{\phi(U)} \omega = \int_{\psi(V)} (f^{-1})^*\omega. \tag{57.189}$$

Proof. By definition $\int_{\phi(U)} \omega = \int_U \phi^*\omega$ and we have the diffeomorphism

$$\phi^{-1} \circ f^{-1} \circ \psi\colon V \to U, \tag{57.190}$$

so that we can use the result of lemma 57.48 :

$$\int_U \phi^*\omega = \int_{(\phi^{-1}\circ f^{-1}\circ\psi)^{-1}(U)} (\phi^{-1} \circ f^{-1} \circ \psi)^*\phi^*\omega = \int_{(\psi^{-1}\circ f\circ\phi)U} \psi^*(f^{-1})^*\omega = \int_{\psi^{-1}(N)} \psi^*(f^{-1})^*\omega. \tag{57.191}$$

The last integral is the definition of an integral on N :

$$\int_{\psi^{-1}(N)} \psi^*(f^{-1})^*\omega = \int_N (f^{-1})^*\omega. \tag{57.192}$$

□

Here is the lemma that shows the independence of definition 57.49 with respect to the change of chart system.

Lemma 57.51.
Let $\varphi: V \to M$ be a chart such that $\varphi(V) \cap \varphi(U) = N$ is not empty. We define $U' = \phi^{-1}(N)$ and $V' = \varphi^{-1}(N)$. Then

EQooLSZMooPcyMWN
$$\int_{\phi(U')} \omega = \int_{\varphi(V')} \omega. \tag{57.193}$$

This lemma allows us to write $\int_N \omega$ the common value of both sides of (57.193).

Proof. Taking $f = \mathrm{Id}$ and two charts for the same open set in M in proposition 57.50 shows the result. □

57.5.3 On manifold that require a finite atlas

We restrict ourself to manifolds that accept a finite atlas.

DEFooITDTooWwrPPr

Definition 57.52 ([929]).
If ω is a n-form on M and if $\{f_\alpha\}$ is a partition of unity[15] *subordinate to the finite atlas $\{U_\alpha\}$ then*

$$\int_M \omega = \sum_\alpha \int_{\phi_\alpha(U_\alpha)} f_\alpha \omega. \tag{57.194}$$

We show that this definition does not depend on the choice of the partition of unity.

LEMooCMIZooHhHaHV

Lemma 57.53 ([1, 929]).
The definition 57.52 is independent of the choice of atlas and partition of unity.

Proof. Let $\{U_\alpha, \phi_\alpha, f_\alpha\}_{\alpha\in A}$ and $\{V_i, \varphi_i, g_i\}_{i\in I}$ be two choices of atlas, charts and subordinate partition of unity. We have to show that

EQooPVQZooHvbioJ
$$\sum_{\alpha\in A} \int_{\phi_\alpha(U_\alpha)} f_\alpha \omega = \sum_{i\in I} \int_{\varphi_i(V_i)} g_i \omega. \tag{57.195}$$

Since $\{g_i\}$ is a partition of unity,

$$\spadesuit = \sum_\alpha \int_{\phi_\alpha(U_\alpha)} f_\alpha \omega = \sum_\alpha \int_{\phi_\alpha(U_\alpha)} \sum_i g_i f_\alpha \omega. \tag{57.196}$$

Since the atlas are finite, the sums are finite and can be permuted with the integral. Moreover the function $g_i f_\alpha$ is nonzero only on $\phi_\alpha(U_\alpha) \cap \varphi_i(V_i)$ so that the integral can be taken on $\phi_\alpha(U_\alpha)$, $\phi_\alpha(U_\alpha) \cap \varphi_i(V_i)$ or $\varphi_i(V_i)$. We have

$$\begin{aligned}
\spadesuit = \sum_i \sum_\alpha \int_{\phi_\alpha(U_\alpha)} g_i f_\alpha \omega &= \sum_i \sum_\alpha \int_{\phi_\alpha(U_\alpha)\cap\varphi_i(V_i)} g_i f_\alpha \omega && \text{(57.197a)}\\
&= \sum_i \sum_\alpha \int_{\varphi_i(V_i)} g_i f_\alpha \omega && \text{(57.197b)}\\
&= \sum_i \int_{\varphi_i(V_i)} g_i \sum_\alpha f_\alpha \omega && \text{(57.197c)}\\
&= \sum_i \int_{\varphi_i(V_i)} g_i \omega. && \text{(57.197d)}
\end{aligned}$$

□

15. See theorem 48.63.

The common values of both sides of (57.195) is denoted by $\int_M \omega$.

The following is not really a definition, but a particular case of 57.49. The integral of a $n-1$-form on a $(n-1)$-submanifold is

$$\int_S \sigma = \int_A \varphi^* \sigma. \tag{57.198}$$

EQooYPOGooRYOXQe

Once again the last integral is an integral of a $n-1$-form on an open set in $\mathbb{R}^{n-1}$.

DEFooAXFXooWiMLKP

Definition 57.54 ([928])**.**
The integral of a 1-form on a $n-1$ dimensional submanifold is :

$$\int_S \tau = \int_S \star\tau \tag{57.199}$$

where $\star$ is the Hodge dual defined by 49.147.

The last integral is the integral of a $(n-1)$-form on a $(n-1)$-submanifold, given by (57.198).

DEFooAXZGooJairMQ

Definition 57.55.
The integral of a vector field on a $(n-1)$-submanifold is :

$$\int_S v = \int_S v^\flat \tag{57.200}$$

where $v^\flat$ is the 1-form defined by the musical isomorphism (49.351).

The following proposition provides a much more explicit formula for the integral of a vector field.

PROPooETLZooAVsrwy

Proposition 57.56.
Let $\varphi\colon A \subset \mathbb{R}^{n-1} \to \mathbb{R}^n$ be an hypersurface and X be a vector field in $\mathbb{R}^n$. Then

$$\int_S X = \int_A \det\left(X, \frac{\partial\varphi}{\partial y_1}, \dots, \frac{\partial\varphi}{\partial y_{n-1}}\right) \tag{57.201a}$$

SUBEQooWJSPooImJjQN

$$= \int_A X \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & \partial_{y_1}\varphi & \\ & \vdots & \\ & \partial_{y_{n-1}}\varphi & \end{pmatrix} \tag{57.201b}$$

$$= \int_A X \cdot n \tag{57.201c}$$

where $\{y_1, \dots, y_{n-1}\}$ are the coordinates on A and n is the normal vector to the parametrization.

Note : thanks to lemma 57.53, the value of n can depend on the choice of coordinates, but the integral will not depend.

Proof. If $X = \sum_{i=1}^n X_i \partial_i$, then $X^\flat = \sum_i X_i dx_i$ and its Hodge dual is

$$\sum_i (-1)^i dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n \tag{57.202}$$

where the hat denotes a factor that is not present. Using the definitions 57.55, 57.54 and 57.49 it remains to integrate

$$\int_A \sum_i (-1)^i X_i \varphi^*\left(dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n\right). \tag{57.203}$$

If $u_1, \dots, u_{n-1}$ are vectors on A (that is on $T_x A$ where x is the integration variable) we have

$$\varphi^*(dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n)(u_1, \dots, u_{n-1}) = (dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n)(d\varphi u_1, \dots, d\varphi u_{n-1}) \tag{57.204a}$$

$$= \det\left(\tau_i d\varphi u_1, \dots, \tau_i d\varphi u_{n-1}\right) \tag{57.204b}$$

where we used the lemma 49.145.

What lies in the integral is the $(n-1)$ differential form EQooEVAPooSbRfaj

$$(u_1,\dots,u_{n-1}) \mapsto \sum_i (-1)^i X_i \det\left(\tau_i d\varphi u_1,\dots,\tau_i d\varphi u_{n-1}\right) \tag{57.205a}$$

$$= \det\left(X, d\varphi u_1,\dots,d\varphi u_{n-1}\right). \tag{57.205b}$$

Since this is a $(n-1)$ differential form over $\mathbb{R}^{n-1}$, this has to be proportional to $dy_1 \wedge \dots dy_{n-1}$. The proportionality factor is found by applying (57.205) to the basis $\{e_1,\dots,e_n\}$. Since $d\varphi(e_i) = \frac{\partial\varphi}{\partial y_i}$ we have the proportionality factor

$$\det\left(X, \frac{\partial\varphi}{\partial y_1},\dots,\frac{\partial\varphi}{\partial y_n}\right) \tag{57.206}$$

and the integral to be computed is

$$\int_A \det\left(X, \frac{\partial\varphi}{\partial y_1},\dots,\frac{\partial\varphi}{\partial y_n}\right) dy_1 \wedge \dots \wedge dy_{n-1} = \int_A \det\left(X, \frac{\partial\varphi}{\partial y_1},\dots,\frac{\partial\varphi}{\partial y_n}\right). \tag{57.207}$$

The formula (57.201a) is proven. The two others are application of lemma 10.38. □

Exemple 57.57.
Let us make the example with $n = 3$. We have

$$\varphi^*(dx\wedge dy)(v_1,v_2) = (dx\wedge dy)(d\varphi v_1, d\varphi v_2) = \det\begin{pmatrix} d\varphi(v_1)_x & d\varphi(v_2)_x \\ d\varphi(v_1)_y & d\varphi(v_2)_y \end{pmatrix}, \tag{57.208}$$

and then

$$\sum_i (-1)^i X_i \varphi^*(\bigwedge_{k\neq i} dx_k)(v_1,v_2) = \sum_i (-1)^i X_i \begin{pmatrix} d\varphi(v_1)_x & d\varphi(v_2)_x \\ d\varphi(v_1)_y & d\varphi(v_2)_y \end{pmatrix} \tag{57.209a}$$

$$= \det\begin{pmatrix} X_1 & d\varphi(v_1)_x & d\varphi(v_2)_x \\ X_2 & d\varphi(v_1)_y & d\varphi(v_2)_y \\ X_3 & d\varphi(v_1)_z & d\varphi(v_2)_z \end{pmatrix} \tag{57.209b}$$

△

57.5.4 Integrating by part

Proposition 57.58 ([1])**.**
Let Ω be an open set in an manifold M of dimension n and $\varphi\colon A \to \mathbb{R}^n$ be a parametrisation of the boundary $\partial\Omega$ with tangent vector field n (defined on $\partial\Omega$). Let $u, v \in C^\infty(M)$. Then we have

$$\int_{\partial\Omega} uv\, n_j = \int_\Omega \frac{\partial u}{\partial x_j} v + \int_\Omega u \frac{\partial v}{\partial x_j}. \tag{57.210}$$ EQooQSMNooKHwbqp

Proof. We use the Stokes formula (theorem 20.71) on the $(n-1)$-form

$$\omega = uv\, dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n, \tag{57.211}$$

and we know from example 49.148 that $\omega = (-1)^{j+1} \star dx_j$. On the other hand,

$$d\omega = \sum_k \frac{\partial(uv)}{\partial x_k} dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n = \frac{\partial(uv)}{\partial x_j}(-1)^{j+1} dx_1 \wedge \dots \wedge dx_n. \tag{57.212}$$

We can use the Stokes formula :

$$\int_{\partial\Omega} uv dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n = (-1)^{j+1} \int_\Omega \frac{\partial(uv)}{\partial x_j}. \tag{57.213}$$

The left-hand side can be transformed as

$$\int_{\partial\Omega} \star(dx_j) = \int_{\partial\Omega} uv\partial_j = \int_{\partial\Omega} uv\, n_j \tag{57.214}$$

where we used the definition 57.54 and the proposition 57.56.

The coefficients $(-1)^{j+1}$ simplify and the derivation of product produce the result. □

EXooWLUVooNamnKG

Exemple 57.59.
If we integrate by part the function $u\frac{\partial^2 v}{\partial x_j^2}$ we have

$$\int_\omega u\frac{\partial^2}{\partial x_j^2} = -\int_\Omega \frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_j} + \int_{\partial\Omega} u\frac{\partial v}{\partial x_j} n_j. \tag{57.215}$$

Summing over j we have the interesting formula

$$\int_\Omega u\Delta v = -\int_\Omega \nabla u \cdot \nabla v + \int_{\partial\Omega} u\frac{\partial v}{\partial n} \tag{57.216}$$

EQooJLDTooIMtxEX

where $\Delta v = \sum_j \frac{\partial^2 v}{\partial x_j^2}$ and $\frac{\partial v}{\partial n}$ is a notation for $\nabla v \cdot n$. △

Chapter 58

Lie groups of transformations

58.1 Groups of transformations

Definition 58.1.
Let M be an Hausdorff space and G, a topological group. It is a ***topological group of transformations*** *of M if to any element $g \in G$, one associates an homeomorphism $g\colon M \to M$, denoted by dot ($p \to g \cdot p$) such that $\forall p \in M$; $\forall g, h \in G$,*

(1) $(gh) \cdot p = g \cdot (h \cdot p)$,

(2) the map $G \times M \to M$; $(g, p) \to g \cdot p$ is continuous.

DefTopoGpTransfo

If $g \cdot \ = id$ only for $g = e$, then the action is **effective**.

The **isotropy group** of M (with respect to the action of G) is naturally defined by

$$H = \{g \in G \text{ tel que } g \cdot p = p\}. \tag{58.1}$$

The unit sphere is an example. Let's consider the vector $\vec{1}_z$. It is clear that the action of SO(3) on this vector covers the whole sphere. But there is a SO(2) subgroup of rotations which leaves $\vec{1}_z$ unchanged. So the sphere is given by the quotient

$$S^2 = \frac{\mathrm{SO}(3)}{\mathrm{SO}(2)}. \tag{58.2}$$

Corollary 58.2.
Let G and X be two locally compact groups. We suppose G to have countable basis. Then any homomorphism $\psi\colon G \to X$ is open.

Proof. By terminology, when we say that a group has some topological property (as the local compactness here), we suppose that the group is a topological group.

For any $g \in G$, we can build an homeomorphism (see remark 48.50) $\varphi_g\colon X \to X$, $x \to \psi(g)x$, so that G is a transitive topological group of transformation on X. Let us denote by f the identity of X. We know that φ_g is continuous, open and satisfies

$$\varphi_g(f) = \psi(g)f = \psi(g).$$

If we define $H = \{h \in G \text{ tel que } \psi(h) = f\}$, $\psi(gh) = \psi(g)$ for any $h \in H$, so that ψ descent to a map $\psi\colon X/H \to X$. This is precisely the map which theorem 53.100 assure us to be an homeomorphism.

□

58.2 Lie groups of transformations

Definition 58.3.
Let G be a Lie group and M, an analytic manifold. We say that G is a ***Lie group of transformation*** *of M when*

(1) G is a topological group of transformation of M,

(2) the map $G \times M \to M$, $(g,p) \to g \cdot p$ is analytic.

DefLieGpTransfo

In particular, for any $g \in G$, $p \to g \cdot p$ is a diffeomorphism $M \to M$. When G is a Lie transformation group on M, for any $X \in \mathfrak{g}$, we define a **fundamental vector field** $X^\dagger \in \mathfrak{X}(M)$ by [1]

$$X_p^\dagger = \frac{d}{dt}\Big[e^{-tX} \cdot p\Big]_{t=0} \tag{58.3}$$

The existence comes from the fact that $(g,p) \to g \cdot p$ and the exponential are analytic and on a function $f \in C^\infty(M)$, the vector field acts as

$$(X^\dagger f)(p) = \lim_{t\to\infty} \frac{f(e^{-tX} \cdot p) - f(p)}{t}.$$

Theorem 58.4.
Consider G, a Lie transformation group on M and X, $Y \in \mathfrak{g}$. Then

$$[X^\dagger, Y^\dagger] = \lim_{t\to 0} \frac{1}{t}(Y^\dagger - dg_t Y^\dagger) \tag{58.4}$$

where $g(t) = g_t = e^{tX}$.

Proof. We consider $f \in C^\infty(M)$ and $q \in M$. The function F defined by

$$F(f,q) = f(e^{-tX} \cdot q) \tag{58.5}$$

is analytic with respect to t, so that

$$F(t,q) - F(0,q) = t\int_0^1 \left(\frac{\partial F}{\partial t}\right)(st,q)ds = h(t,q)t \tag{58.6}$$

eq:def_F_h

for a certain function $h \in C^\infty(\mathbb{R}\times M)$ which satisfies $h(0,q) = (X^\dagger f)(q)$. Naturally, g_t can be seen as a map $g_t \colon M \to M$ by the action. Then dg_t is a linear map $dg_t \colon T_qM \to T_{g_t \cdot q}M$ (we voluntary omit the index q which was fixed; formally, we speak about $(dg_t)_q$)

$$dg_t v = \frac{d}{du}\Big[g_t \cdot v(u)\Big]_{u=0}.$$

Thus in order to compute $(dg_t Y^\dagger)_p$, we have to consider $Y^\dagger$ at $r = g_t^{-1} \cdot p$. Consider a path $v_r \colon [0,1] \to M$ such that $v_r'(0) = Y_r^\dagger$ and $v_r(0) = r$. So,

$$\begin{aligned}(dg_t \cdot Y^\dagger)_p f &= \frac{d}{du}\Big[f(g_t \cdot v_r(u))\Big]_{u=0}\\ &= \frac{d}{du}\Big[(f\circ g_t \circ v_r)(u)\Big]_{u=0}\\ &= \frac{d}{du}\Big[F(t, v_r(u))\Big]_{u=0}\\ &= \frac{d}{du}\Big[F(0, v_r(u)) + h(t, v_r(u))t\Big]_{u=0}\end{aligned} \tag{58.7}$$

The two terms are computed separately:

$$\frac{d}{du}\Big[F(0,v_r(u))\Big]_{u=0} = \frac{d}{du}\Big[f(v_r(u))\Big]_{u=0} = Y_r^\dagger(f),$$

and

$$\frac{d}{du}\Big[th(t,v_r(u))\Big]_{u=0} = t(Y^\dagger h)_{(t,v_r(0))}.$$

1. Remark, as usual, that some literature (in particular in [813]) gives it without the minus sign.

Finally,

$$(dg_t \cdot Y^\dagger)_p f = Y^\dagger_{g_t^{-1} \cdot p}(p) + t(Y^\dagger h)_{(t, g_t^{-1} \cdot p)}. \tag{58.8}$$

eqnY_h

Now we can compute:

$$\begin{aligned}\lim_{t\to 0} \frac{1}{t}\left((Y^\dagger - dg_t Y^\dagger)_p f\right) &= \lim_{t\to 0} \frac{1}{t}\left\{(Y^\dagger f)p - (Y^\dagger f)(g_t^{-1} \cdot p)\right\} \\ &\quad + \lim_{t\to 0} \frac{1}{t}\left\{(Y^\dagger f)(g_t^{-1} \cdot p) - (Y^\dagger (f \circ g_t))(g_t^{-1} \cdot p)\right\} \\ &= \frac{d}{dt}\Big[(Y^\dagger f)(g_t^{-1} \cdot p)\Big]_{t=0} - \lim_{t\to 0}\left((Y^\dagger h)(t, g_t^{-1} \cdot p)\right).\end{aligned} \tag{58.9}$$

The latter equality comes from (58.8). The first term is computed as following ($Y^\dagger(f)$ is a function):

$$\frac{d}{dt}\Big[Y^\dagger(f)(g_t^{-1} \cdot p)\Big]_{t=0} = \frac{d}{dt}\Big[(Y^\dagger(f) \circ g_t^{-1})(p)\Big]_{t=0} = \big(Y^\dagger(f)\big)_p (g_t^{-1})'(0) = \big(X^\dagger Y^\dagger(f)\big)_p \tag{58.10}$$

In the expression $(Y^\dagger h)(t, g_t^{-1} \cdot p)$, we have to consider the dependence on t as a parameter: the vector $Y^\dagger$ only acts on the "second slot" of h. From definition (58.6) of h,

$$h(t, g_t^{-1} \cdot p) = \frac{1}{t}\left(F(f, g_t^{-1} \cdot p) - F(0, g_t^{-1} \cdot p)\right),$$

but $F(0, q) = f(q)$ and $F(t, g_t^{-1} \cdot p) = f(p)$, then

$$f(t, g_t^{-1} \cdot p) = \frac{1}{t}\left(f(p) - f(e^{-tX} \cdot p)\right)$$

Taking the limit for small t, it becomes

$$\frac{d}{dt}\Big[f(e^{-tX} \cdot p)\Big]_{t=0} = (X^\dagger f)_p$$

□

58.2.1 Action of the Lie algebra

58.5.
We will study the infinitesimal action of the two dimensional conformal group on $\mathbb{C}$ in 71.8 and what follows.

DEFooUYOZooWdcClz

Definition 58.6 ([930]).
If the Lie group G acts on the manifold M, then its Lie algebra $\mathfrak{g}$ has an action

$$\begin{aligned} X\colon M &\to TM \\ x &\mapsto \frac{d}{dt}\Big[\exp^{-tX} \cdot x\Big]_{t=0} \in T_x M. \end{aligned} \tag{58.11}$$

If you want to know how to integrate this to an action of the group, see the differential equation (53.186).

Thus an element of the Lie algebra $\mathfrak{g}$ acts on a function $f \in C^\infty(M)$ as $X(f) \in C^\infty(M)$ by

$$X_x(f) = \frac{d}{dt}\Big[f\big(e^{-tX}x\big)\Big]_{t=0} = df_x\big(-X(x)\big) = -df_x\big(X(x)\big) \tag{58.12}$$

EQooQDLAooZXnWta

where $X_x(f)$ is a notation shortcut for $X(f)(x)$. Here $X(x)$ stands for the map $X\colon M \to M$ defined by $X(x) = e^{tX}x$.

58.3 Cosets and homogeneous spaces

Proposition 53.101 takes its interest in the setting of homogeneous space. An **homogeneous spaces** is a smooth manifolds M which admits a Lie group of transformations. If $p_0 \in M$ and H is the stabilizer of p_0, proposition 53.101 says that H is closed and therefore theorem 53.98 makes G a Lie group of transformations of G/H. Hence the latter becomes a homogeneous space. The map α of proposition 53.101 gives an isomorphism of homogeneous spaces.

This allow us to see a homogeneous space as the quotient of a group by a closed subgroup.

58.4 Isotropy group

If one has a Lie group G and a closed subgroup H, we know from theorem 53.70 that H is a topological Lie subgroup of G. We naturally consider this structure and the analytic structure on G/H given by 53.98. For this structure of the coset space G/H, the group H is the **isotropy group**. We denote by $\tau(x)\colon G/H \to G/H$ $(x \in G)$ the diffeomorphism $\tau(x)[y] = [xy]$. The group H^* of the linear transformations $(d\tau(h))_{\pi(e)}$ $(h \in H)$ is the **linear isotropy group**.

Let N be a Lie subgroup of G. Then the subset $N \cap H$ is closed[2]. Then $N \cap H$ is closed in N and we look at $N/(N \cap H)$. We can exhibit a bijection between this and the orbit of $\pi(e)$ in G/H with respect to the action of N:

$$\{n\pi(e) \text{ tel que } n \in N\} \simeq N/(N \cap H) \tag{58.13}$$

by the map ψ given by $\psi(n\pi(e)) = \overline{n}$. Here the $\overline{x}$ denotes the class of x with respect to $N \cap H$. Note that for $n \in N$, $\overline{n} \neq \overline{e}$ because there are *a priori* elements in $N \backslash H$. The map ψ is well defined because $m[e] = [m] = [n]$ if $m = nh$ for a certain $h \in H$. Then $\psi([nh]) = \overline{nh}$. But in order for nh to belongs to N, one needs $h \in N \cap H$; then $\overline{nh} = \overline{n}$. For the same reason, ψ is injective. The surjectivity is clear.

prop:orbit_N_ss_var

Proposition 58.7.
In this context,

(1) The orbit of e by N in G/H is $N/(N \cap H)$. It is submanifold of G/H.

(2) If N is a topological subgroup of G and if H is compact, then the submanifold $N/(N \cap H)$ is a closed topological submanifold of G/H.

Proof. First item. We denote by $\overline{n}$, the class of n with respect to $N \cap H$ and by $[g]$, the class of g with respect to H. The following diagram is commutative:

exdig_4_4

$$\begin{array}{ccc} N & \xrightarrow{\ i\ } & G \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi} \\ N/(N \cap H) & \xrightarrow{\ I\ } & G/H \end{array} \tag{58.14}$$

where $\pi_1\colon N \to N/(N \cap H)$ and $\pi\colon G \to G/H$ are canonical projections; $i\colon N \to G$ is the inclusion; and $I\colon N/(N \cap H) \to G/H$ is defined by $\overline{n} \to [i(n)]$. Indeed, $\pi(i(n)) = [i(n)] = I(\overline{n}) = I(\pi_1(n))$ for any $n \in N$.

If $\mathfrak{n}$ is the Lie algebra of N and $\mathfrak{h}$ the one of H, $\mathfrak{h}_1 = \mathfrak{h} \cap \mathfrak{n}$ is the Lie algebra of $N \cap H$. We consider $\mathfrak{n}_1$ and $\mathfrak{g}_1$ such that $\mathfrak{n}_1 \oplus \mathfrak{h}_1 = \mathfrak{n}$ and $\mathfrak{g}_1 \oplus (\mathfrak{h} \oplus \mathfrak{n}_1) = \mathfrak{g}$. Let us show why is the sum $\mathfrak{h} \oplus \mathfrak{n}_1$ direct. First remark that $\mathfrak{h} \cap \mathfrak{n}_1$ because $\mathfrak{h} \cap \mathfrak{n}_1 \subset \mathfrak{h} \cap \mathfrak{n} = \mathfrak{h}_1$, but $\mathfrak{h}_1 \cap \mathfrak{n}_1 = \{0\}$. Immediately, the sum $\mathfrak{g} = \mathfrak{g}_1 \oplus (\mathfrak{h} \oplus \mathfrak{n}_1)$ is direct.

Now we apply lemma 53.97 to the decomposition $\mathfrak{n} = \mathfrak{h}_1 \oplus \mathfrak{n}_1$; this give us a submanifold $B_N \subset N$ which contains e and on which π_1 is diffeomorphic to an open neighbourhood of $\pi_1(e)$ in

2. If H' denotes the complementary of H in G (which is open in G), the complementary $N \cap H'$ of $N \cap H$ is open in N for the induced topology of N from G.

$N/(N\cap H)$. The same with $\mathfrak{g}=\mathfrak{h}\oplus(\mathfrak{n}_1+\mathfrak{g}_1)$ gives $B_G\subset G$, a submanifold around e on which π is diffeomorphic to a neighbourhood of $\pi(e)$ in G/H. We can see B_N as a submanifold of B_G.

We denotes $V_1=\pi_1(B_N)$, $V=\pi(B_G)$ and I_{V_1}, the restriction of I to V_1. The Jacobian of I_{V_1} at $\pi(e)$ has a rank equal to $\dim\big(N/(N\cap H)\big)$. Indeed we can write I_{V1} as $I_{V_1}=\pi\circ i\circ\pi^{-1}$ and π_1 is a diffeomorphism, so that π^{-1} don't change the dimension. The fact that the Jacobian at identity is non zero on a neighbourhood makes it regular on this neighbourhood and the analyticity make it regular everywhere. The characterization for a submanifold given by proposition 49.112 gives the first item.

Second item. We know that N is a submanifold of G; the commutative diagram (58.14) shows that I is an homeomorphism because the topologies on G/H and $N/(N\cap H)$ are made in order to π and π_1 are continuous and open. If N is a topological subgroup of G, an open subset of N is written under the form $N\cap\mathcal{O}$ where $\mathcal{O}$ is open in G. If $\overline{n}\in N/(N\cap H)$, $I(\overline{n})=[i(n)]$.

Now we show that $N/(N\cap H)$ is closed. Consider a sequence (p_k) in $N/(N\cap H)$ which converges to $q\in G/H$. The aim is to show that $q\in N/(N\cap H)$. We take a $g\in G$ such that $\pi(g)=q$; we can suppose that the whole sequence (p_k) is in the neighbourhood $g\cdot V$ of q. In order to see that it is a neighbourhood, recall that π is a diffeomorphism from V to an open neighbourhood of e in G/H, thus V is an open neighbourhood of $\pi(e)$ in G/H.

Since π is diffeomorphic, there exists a sequence $g_k\in gB_G$ such that $\pi(g_k)=p_k$. It satisfies $\lim g_k=g$. On the other hand, for each k, $\exists n_k\in N$ such that $\pi_1(n_k)=p_k$; then $\pi(g_k)=p_k=\pi_1(n_k)$, then there exists $h_k\in H$ such that $g_k=n_kh_k$. But H is compact, then h_k is a converging sequence (by eventually passing to a subsequence). Since g_k and h_k converge, n_k also converges. But N is closed in G, then $n^*=\lim n_k\in N$. Finally $\pi_1(n^*)=q$, so that $N/(N\cap H)$ is closed. □

Chapter 59

Classical mechanics

59.1 Some symplectic and Poisson geometry

sec:symple

59.1.1 Symplectic manifold

A **symplectic structure** on a vector space V is a skew-symmetric, nondegenerate bilinear 2-form $\Omega\colon V\times V\to\mathbb{R}$. We define the **symplectic group** $\mathrm{SP}(\Omega)$ as the group of linear operators $A\colon V\to V$ such that $\Omega(Au,Av)=\Omega(u,v)$ for every u, $v\in V$. It is easy to see that elements of $\mathrm{SP}(V)$ satisfy

$$A^t\Omega A=\Omega. \tag{59.1}$$

EqPropodefMtrsym

The Lie algebra of $\mathrm{SP}(\Omega)$ is denoted by $\mathfrak{sp}(\Omega)$. Taking the derivative of equation (59.1) with respect to A, one finds the following condition for $B\in\mathfrak{sp}(\Omega)$:

$$\Omega B+B^t\Omega=0. \tag{59.2}$$

If W is a subspace of the symplectic vector space V, the **symplectic complement** is

$$W^\omega=\{x\in V \text{ tel que } \omega(x,y)=0\ \forall y\in W\}. \tag{59.3}$$

We say that W is a **Lagrangian** of V when $W=W^\omega$. Typically, when one consider the canonical structure $\Omega(p_i,q_j)=\delta_{ij}$ on $\mathbb{R}^{2n}$, the spaces $\mathrm{Span}\{p_i \text{ tel que } i=1,\dots,n\}$ and $\mathrm{Span}\{q_i \text{ tel que } i=1,\cdots,n\}$ are Lagrangian.

A **symplectic manifold** is the data of a smooth manifold M and a symplectic structure ω_x on each tangent space T_xM. The map $x\mapsto\omega_x$ is required to be a smooth section of the 2-tensor bundle.

Definition 59.1.
*A **symplectic Lie algebra** is a Lie algebra $\mathfrak{s}$ endowed with a symplectic structure ω such that $\forall x$, y, $z\in\mathfrak{s}$,*

$$\omega([x,y],z)+\omega([y,z],x)+\omega([z,x],y)=0. \tag{59.4}$$

EqDefAlgSymple

DefSympleAlg

59.1.2 Poisson manifold

Let M be a smooth manifold. A **Poisson structure**, or a *Poisson bracket* on M is a map $\{.,.\}\colon C^\infty(M)\times C^\infty(M)\to C^\infty(M)$ such that

(1) $\big(C^\infty(M),\{.,.\}\big)$ is a Lie algebra,

(2) for each $f\in C^\infty(M)$, the map $\{f,.\}$ is a derivation of the algebra $C^\infty(M)$:

$$\{f,gh\}=\{f,g\}h+g\{f,h\}.$$

59.1.3 Hamiltonian action

app:ham_act

Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. A **symplectomorphism** from M_1 to M_2 is a diffeomorphism $\varphi\colon M_1 \to M_2$ such that $\varphi^*\omega_2 = \omega_1$.

For any function $f \in C^\infty(M)$, we define the **Hamiltonian field** $X_f \in \mathfrak{X}(M)$ associated with f by

$$i(X_f)\omega = df. \tag{59.5}$$

EqDefHamVect

Existence of such a X_f for each f is assured because ω is nondegenerate. A symplectic structure induces a Poisson bracket by defining

$$\{f, g\} = -\omega(X_f, X_g) = -X_g(f) = X_f(g). \tag{59.6}$$

eq:def_Poisson

In local coordinates, one can write $\omega = \frac{1}{2}\omega_{ij}dx^i \wedge dx^j$ and $X_f = \omega^{ij}\partial_i f \partial_j$, where (ω^{ij}) is the inverse matrix of (ω_{ij}). The Poisson tensor defined by

$$\{f, g\} = P^{kl}\partial_k f \partial_l g, \tag{59.7}$$

EqPoisson

is nothing else than $P = \omega^{-1}$.

tho:equiv_Poisson

Theorem 59.2.
If $\varphi\colon M \to M'$ is a diffeomorphism between two Poisson manifolds (M, P) and (M', P'), then the following are equivalent:

ite_equivii

(1) $\varphi_*(X_{f\circ\varphi}) = X'_f$,

(2) $\{u \circ \varphi, v \circ \varphi\} = \{u, v\}' \circ \varphi$,

(3) $\varphi_* P = P'$,

If moreover the Poisson structures P and P' come from symplectic forms ω and ω', $\varphi^*\omega' = \omega$.

Definition 59.3.
*Let $\tau\colon G \times M \to M$ be a symplectic action of a Lie group G on M (i.e. $\tau_g\colon M \to M$ is a symplectic transformation of M for each $g \in G$). That action is **Hamiltonian** if, for every $X \in \mathcal{G}$, there exists a map $\lambda_X \in C^\infty(M, \mathbb{C})$ such that*

eq:act_ham

$$i(X^*)\omega = d\lambda_X, \tag{59.8a}$$

ls17d

$$\{\lambda_X, \lambda_Y\} = \lambda_{[X,Y]} \tag{59.8b}$$

eq:hamil

where X^* is the fundamental vector field associated with the vector X.

Definition 59.4.
*The map $\lambda\colon \mathcal{G} \to C^\infty(M)$ which satisfies (59.8) is the **dual** momentum map while the **momentum map** is $J\colon M \to \mathcal{G}^*$ defined by*

$$\lambda_X(x) = \langle J(x), X\rangle \tag{59.9}$$

eq:defmomm ap

for all $X \in \mathcal{G}$.

def:app_mom_mom_duale

Using the musical isomorphism (see 49.11.3), equation (59.8a) can be written as $(d\lambda_X)^\sharp = X^*$ or $d\lambda_X = (X^*)^\flat$.

Since $i(X_{\lambda_Y})\omega = d\lambda_Y = i(Y^*)\omega$, we have:

$$X_{\lambda_Y} = Y^*. \tag{59.10}$$

eq__XlambdaYs

The **symplectic gradient** is

$$\operatorname{sgrad} u = (du)^\sharp.$$

Remarque 59.5.
If the action only fulfils (59.8a), it is said to be **weakly Hamiltonian**. But one can find some literature in which the first condition is called a *Hamiltonian action* and the second, a *strongly Hamiltonian* one.

Remarque 59.6.
Be careful with some choices of signs. We made at least three choices: the definition of λ_X in (59.8a) can get a minus sign; the definition of the Poisson bracket (59.6) is modulo a sign; and the definition of the fundamental fields that we take is

$$X_x^* = \frac{d}{dt}\Big[\tau_{e^{-tX}}x\Big]_{t=0}. \tag{59.11}$$

EqDefFondsiTHZ

It can be written without the minus sign in the exponential. Each of these choice is free, by condition (59.8b) is not! If one make "wrong" choices, this can become

$$\{\lambda_X, \lambda_Y\} = -\lambda_{[X,Y]}, \tag{59.12}$$

with all the consequences in the computations. For example, if one defines $X_x^* = \frac{d}{dt}\tau_{\exp tX}x\big|_{t=0}$, one finds an extra minus sign in equations (83.39), (83.41) and (83.43).

59.1.4 Coadjoint orbits

sub:coadjoint

Let G be a Lie group and $\mathcal{G}$ its Lie algebra. We know that G acts on the dual $\mathcal{G}^*$ by

$$g \cdot \xi = \xi \circ \mathrm{Ad}(g^{-1}) = \mathrm{Ad}(g)^*\xi \tag{59.13}$$

EqDefActCoadi

for $g \in G$ and $\xi \in \mathcal{G}^*$. The second equality defines the **coadjoint action** $\mathrm{Ad}^* \colon G \times \mathcal{G}^* \to \mathcal{G}^*$. In other words, for all $X \in \mathcal{G}$,

$$(g \cdot \xi)(X) = \langle \xi, \mathrm{Ad}(g^{-1})X\rangle = \langle \mathrm{Ad}(g)^*(\xi), X\rangle.$$

This representation is also called the **contragredient representation**. In this context, the notion of fundamental fields is given bu $X_\xi^* = \xi \circ \mathrm{ad}(X)$. Let $\theta_\xi = \{g \cdot \xi | g \in G\}$, the orbit of ξ in $\mathcal{G}^*$. It can be shown that

$$\tilde{\omega}_x(X_x^*, Y_x^*) = \langle x, [X,Y]\rangle \tag{59.14}$$

eq_omega_Gs

defines a symplectic form on θ_ξ, the coadjoint orbit of ξ.

Proposition 59.7.
The coadjoint action is Hamiltonian.

Proposition 59.8.
The diffeomorphism

$$\begin{aligned} \varphi(g)\colon \theta_\xi &\to \theta_\xi \\ \xi &\mapsto g \cdot \xi \end{aligned} \tag{59.15}$$

fulfils

$$\big(\varphi(g)^*\omega\big)_\xi(X_\xi^*, Y_\xi^*) = \omega_\xi(X_\xi^*, Y_\xi^*).$$

59.1.5 Example of adjoint and coadjoint orbit

Let us consider the following subgroup of $\mathrm{SL}(2,\mathbb{R})$

$$\mathbb{S} = \{\begin{pmatrix} a & n \\ 0 & a^{-1} \end{pmatrix} \text{ where } a > 0 \text{ and } n \in \mathbb{R}\}. \tag{59.16}$$

The Lie algebra is

$$\mathfrak{s} = \{\begin{pmatrix} x & y \\ 0 & -x \end{pmatrix} \text{ where } x, y \in \mathbb{R}\}. \tag{59.17}$$

The following is a basis of $\mathfrak{s}$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{59.18}$$

and the commutation relation is $[H, E] = 2E$. Since we are in a matrix group, the adjoint action is given by $\mathrm{Ad}(g)X = gXg^{-1}$ and we can compute

$$\mathrm{Ad}(a,n)(x,y) = \begin{pmatrix} x & ay^2 - 2nax \\ 0 & -x \end{pmatrix} = xH + (a^2y - 2nax)E \tag{59.19}$$

Thus the adjoint orbits are the straight lines.

For the coadjoint orbits, consider an element $\xi = \xi_H H^* + \xi_E E^* \in \mathfrak{s}^*$. We have

$$\begin{aligned} \langle \mathrm{Ad}^*(g^{-1})\xi, X_0 \rangle &= \langle \xi, \mathrm{Ad}(g)X \rangle \\ &= (\xi_H - 2na\xi_E)x + a^2\xi_E y. \end{aligned} \tag{59.20}$$

If we write $\xi = (\xi_H, \xi_E)$ we have

$$\mathrm{Ad}^*(a,n)^{-1}(\xi_H, \xi_E) = \begin{pmatrix} 1 & -2na \\ 0 & a^2 \end{pmatrix} \begin{pmatrix} \xi_H \\ \xi_E \end{pmatrix}. \tag{59.21}$$

This is the orbit of ξ. Notice that H^* is fix while $E^* \mapsto -2naH^* + a^2E^*$, so that E^* has a quite large orbit. If $\xi_2 \neq 0$, then $\dim \mathrm{Ad}^*(\mathbb{S})\xi = 2$. If $\xi_E = 0$, then the orbit is one single point. Thus we have three kind orbits in the plane (H^*, E^*):

(1) the upper half plane is one two dimensional orbit,

(2) each of the point on the $\xi_E = 0$ is an orbit,

(3) the lower half plane is a second two dimensional orbit.

These orbits are very different from the adjoint orbits.

59.1.6 A second example

Let $\mathfrak{g} = \mathrm{SL}(2, \mathbb{R})$ and the basis $\{H, E, F\}$ with the relations

$$\begin{aligned} [H, E] &= 2E \\ [H, F] &= -2F \\ [E, F] &= H. \end{aligned} \tag{59.22}$$

The Killing form is given by the matrix

$$\beta = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}. \tag{59.23}$$

It is nondegenerate and its signature is $+, +, -$.

Lemma 59.9.
If the Killing form β is nondegenerate, the map

$$\begin{aligned} \flat &\colon \mathfrak{g} \to \mathfrak{g}^* \\ X^\flat(y) &= \beta(X, Y) \end{aligned} \tag{59.24}$$

is a linear Ad*-equivariant isomorphism, that is*

$$\mathrm{Ad}^*(g)(X^\flat) = \big(\mathrm{Ad}(g)X\big)^\flat. \tag{59.25}$$

Proof. We have

$$\begin{aligned} \langle \mathrm{Ad}^*(g)X^\flat, Y \rangle &= \beta\big(X, \mathrm{Ad}(g^{-1})Y\big) \\ &= \beta\big(\mathrm{Ad}(g)X, Y\big) \\ &= \big(\mathrm{Ad}(g)X\big)^\flat Y. \end{aligned} \tag{59.26}$$

□

In that case, the $\flat$ map intertwines the actions and the adjoint orbits are the same as the coadjoint orbits.

59.1.7 Kostant theorem

Let G act transitively on M and on $\mathcal{G}^*$ by the coadjoint action (59.13). We suppose this action to be Hamiltonian. One can prove that J is **equivariant**:

$$J(g \cdot x) = g \cdot J(x) \tag{59.27}$$

where the first dot is the action on M and the second one, the action on $\mathcal{G}^*$. The momentum map of M and $\mathcal{G}^*$ are related by $\lambda_X^M = \lambda_X^{\mathcal{G}^*} \circ J$. Since G is transitive on M, for any $x \in M$, we have $M = G \cdot x$ for any x in M, so

$$J(M) = G \cdot J(x).$$

We conclude that $J(M)$ is only one orbit in $\mathcal{G}^*$.

Proposition 59.10.
Let $\mathcal{O}$ be the coadjoint orbit of a connected Lie group G. Then $\mathcal{O}$ is canonically endowed by a $\mathrm{Ad}^*(G)$*-equivariant symplectic structure given by*

$$\omega_\xi^{\mathcal{O}}(X^*, Y^*) = \langle \xi, [x, Y] \rangle \tag{59.28}$$

EqKKSsymple

where X^* stands for the fundamental vector. Here the equivariance means

$$\mathrm{Ad}^*(g)\omega^{\mathcal{O}} = \omega^{\mathcal{O}} \tag{59.29}$$

for every $g \in G$.

Proof. No proof. □

The formula (59.28) is *Kirillov-Kostant-Souriau*. Notice that this proposition shows that the coadjoint orbits are always even dimensional.

If $\xi \in \mathcal{G}^*$, we know a symplectic structure on its orbit θ_ξ:

$$\omega_\eta^\theta(X_\eta^*, Y_\eta^*) = \eta([X, Y]) \tag{59.30}$$

for all $\eta \in \theta_\xi$. We have

$$\begin{aligned}(J^*\omega^\theta)_x(X^{*M}, Y^{*M}) &= \omega_{J(x)}^\theta(X^*, Y^*) \\ &= J(x)[X, Y] \\ &= \lambda_{[X,Y]}^M(x) \\ &= \{\lambda_X^M, \lambda_Y^M\}(x),\end{aligned} \tag{59.31}$$

and if M is symplectic, it also fulfils $\omega_x^M(X^{*M}, Y^{*M}) = \{\lambda_X^M, \lambda_Y^M\}(x)$, so that

$$J^*\omega^\theta = \omega^M. \tag{59.32}$$

Theorem 59.11 (Kostant theorem)**.**
Let (M, ω) be a symplectic manifold on which the connected Lie group G has a transitive Hamiltonian action. Then the momentum map $J\colon M \to \mathcal{G}^*$ takes his values in only one orbit θ and J is a covering[1] *of G in $\mathcal{G}^*$.*

Proof. Let us fix a $x \in M$, we pose $\xi = J(x)$ and define G_x, the stabilizer of x in G. Formula $\psi(g) = g \cdot x$ defines a bijection $\psi\colon G/G_x \to M$. So we identify $M = G/G_x$. In the same way, $\theta_\xi = G/G_\xi$. So we can write $J(x) = G_\xi$ and

$$J(g \cdot G_x) = g \cdot J(x) = g \cdot G_\xi.$$

Since $g \cdot \xi = g \cdot J(x) = J(g \cdot x) = J(x) = \xi$, we have $G_x \subset G_\xi$. The nondegenerateness of ω^M and ω^θ makes J nondegenerate because $\omega^M = J^*\omega^\theta$. Then dJ is injective and J is an immersion.

We conclude that $\mathcal{G}_x = \mathcal{G}_\xi$ which proves that G_ξ/G_x is discrete and finally that J is a covering. □

1. revêtement, en français.

59.1.8 Central extension

Let $\mathcal{G}$ be a Lie algebra. A **Chevalley coboundary** is a 2-form which reads $\delta\xi$ for a certain $\xi \in \mathcal{G}^*$ with δ defined by

$$(\delta\xi)(A,B) = -\xi([A,B]). \tag{59.33}$$

EqDefCheyCoycl

Let Ω be a 2-cocycle. If it is not a coboundary, we add an element C in $\mathcal{G}$ and we consider $\mathcal{G}' = \mathcal{G} \oplus \mathbb{R}C$ with the Lie algebra structure

$$[A+s, B+t]_{\mathcal{G}'} = [A,B]_{\mathcal{G}} + \Omega(A,B)C. \tag{59.34}$$

This is the **central extension** of $\mathcal{G}$ with respect to the 2-cocycle Ω. The terminology comes from the fact that the extension $\mathbb{R}C$ belongs to the center of $\mathcal{G}'$. The point is that Ω is a coboundary in $\mathcal{G}'$ because

$$(\delta C^*)(A,B) = C^*[A,B]_{\mathcal{G}'} = C^*\big([A,B]_{\mathcal{G}} + \Omega(A,B)C\big) = \Omega(A,B), \tag{59.35}$$

so that $\Omega = \delta C^*$.

Now we suppose that the group G acts on a manifold M. We define the action of the extended group $G' = G \otimes e^{\mathbb{R}C}$ by saying that the "new" part does not act: $(g,s)\cdot x = g\cdot x$. Fundamental fields remains unchanged:

$$(X,s)^* = X^*. \tag{59.36}$$

eq:XseqXss

If the action of G on M is weakly Hamiltonian, we have functions $\mu_x \colon M \to \mathbb{C}$ such that $i(X^*)\omega = d\mu_X$. These functions fulfil $X^* = \{\mu_X, .\}$. We define

$$\lambda_{X,s} = \mu_X + s. \tag{59.37}$$

eq:lamXsmuXs

Proposition 59.12.
The action of G' is (strongly) Hamiltonian for these functions.

Proof. From equation (59.36), we have $\{\mu_X, .\} = \{\mu_{X,s}, .\}$ hence

$$\{\lambda_{(X,s)}, \lambda_{(Y,t)}\} = \{\mu_X, \mu_Y\} = \mu_{[X,Y]} + C_{X,Y} \tag{59.38}$$

for certain constants C_{XY} which satisfy the property $d\big(\{\mu_X,\mu_Y\} - \mu_{[X,Y]}\big) = 0$. Therefore

$$\{\lambda_{(X,s)}, \lambda_{(Y,t)}\} = \lambda_{[X,Y],C_{X,Y}} = \lambda_{[(X,s),(Y,t)]}. \tag{59.39}$$

□

The sense of the whole construction is the following. When the action G is weakly Hamiltonian on M, we have functions μ_X which define Ω by

$$\{\mu_X, \mu_Y\} = \mu_{[X,Y]} + \Omega(X,Y).$$

In this case, the corresponding group extension has a strongly Hamiltonian action with momentum maps given by (59.37).

59.2 Constrained system

59.2.1 Holonomic constraints

References are [931, 932]. An **holonomic constraint** appears when one force a particle to move for an example on a circle. More formally let Q be the configuration space; an holonomic constraint is a submanifold $N \subset Q$, or more generally an integrable subbundle of TQ. From the inclusion $TN \subset TN$, the Lagrangian restricts to $L_N \colon TN \to \mathbb{R}$.

We assume that L_N is a regular Lagrangian (see definition on page 2972) and that N is given by an equation of the form $N = \varphi^{-1}(0)$ for a certain section $\varphi\colon Q \to E^*$ of the dual of a vector bundle over Q. The variational principle is given by

$$\delta \int L_N(q,\dot{q})dt = 0 \tag{59.40}$$

where the function q is constrained by $\varphi\big(q(t)\big) = 0$ for every t. We enforce that condition by the Lagrange multiplier method:

$$\delta \int \Big(L\big(q(t),\dot{q}(t)\big) - \big\langle \lambda\big(q(t),t\big), \varphi\big(q(t)\big)\big\rangle\Big)dt \tag{59.41}$$

where λ is a function of (q,t) with values in E. The δ denotes a variation on the curves q in Q and λ in E. In local coordinates, the Euler-Lagrange equations read

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{\partial L}{\partial q^i} - \lambda\frac{\partial \varphi_a}{\partial q^i}, \qquad \varphi_a = 0. \tag{59.42}$$

So the Euler-Lagrange equations for L_N on N are the sames as the one of L on Q, with the constraint $\varphi = 0$.

59.2.2 Primary constraints

The action is given by

$$S(q) = \int_{t_1}^{t_2} L(q,\dot{q})dt,$$

and we require the action to be stationary with respect to variations δq vanishing on t_1 and t_2. The Euler-Lagrange equations are given by $\frac{\delta S}{\delta q} = \frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0$ or, in coordinates,

$$0 = \frac{\partial L}{\partial q^k} - \dot{q}^l\frac{\partial L}{\partial q^l \partial q^k} - \ddot{q}^l\frac{\partial L}{\partial \dot{q}^l \partial \dot{q}^k}.$$

when the Hessian matrix is invertible, i.e. when

$$\det\left(\frac{\partial L}{\partial \dot{q}^l \partial \dot{q}^k}\right) \neq 0,$$

one can solve these equations in an algebraic way for the accelerations $\ddot{q}$ in functions of the positions q and the velocities $\dot{q}$. When that determinant is vanishing, some of the Euler-Lagrange equations are simple relations between components of q and components of $\dot{q}$, so that the solutions include some arbitrary functions of time.

59.2.2.1 Example: Maxwell

The action is

$$S(A) = -\frac{1}{4}\int F_{\mu\nu}F^{\mu\nu}d^4x$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. That actions has a gauge invariance

$$A_\mu \mapsto A'_\mu = A_\mu + \partial_\mu f$$

for any function f. The equations of motion are $\partial_\nu F^{\mu\nu} = 0$, in particular the equation $\nabla \times \overline{B} = \partial_t \overline{E} = 0$ provides the $\ddot{A}_k$ in terms of the A_l and $\dot{A}_\mu$. Indeed the Hessian

$$\frac{\partial^2 L}{\partial \dot{A}_\alpha \partial \dot{A}_\beta} = -4\big(g^{0\beta}g^{\alpha 0} + g^{\alpha\beta}\big) = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix},$$

has vanishing determinant.

59.2.3 Passage to Hamiltonian formalism

The very principle is to define the momentums $p_k = \frac{\partial L}{\partial \dot{q}^k}$, so that

$$\frac{\partial p_k}{\partial \dot{q}^l} = \frac{\partial^2 L}{\partial \dot{q}^l \partial \dot{q}^k},$$

so that the vanishing determinant condition makes that it is not possible to solve the velocities $\dot{q}^k = \dot{q}^k(q,p)$. That means that the momentums are not independent: there exists relations

$$\phi_m(p,q) = 0 \tag{59.43}$$

for $m = 1, \dots, M$. These relations are the **primary constraints**: they do not rely on the equations of motion. In other terms, we have

$$\phi_m\big(q, p(q,\dot{q})\big) = 0$$

identically, and not only on the solutions.

In the Maxwell example, we pose the momentums $\pi^\alpha = \frac{\partial L}{\partial \dot{A}_\alpha} = F^{0\alpha}$. Notice that $\pi^0 + F^{00} = 0$ and that the remaining is $\pi^k = E^k$ (the electric field).

If one suppose the rank of the Hessian to be constant, the constraints $\phi_m(q,p) = 0$ define a submanifold of the phase space. This is the *primary constraint surface*.

Suppose that the primary constraints can be divided into one set of independent constraints ϕ_m, $m = 1, \dots, M'$ with the rank of $\frac{\partial \phi_{m'}}{\partial X}$ being M' ($X = (q,p)$) and some dependent constraints $\phi_{\bar{m}}$ such that vanishing of all the $\phi_{m'}$ implies the vanishing of all the $\phi_{\bar{m}}$.

In that case, we have two important results.

Theorem 59.13.
If a function G vanishes on the constraint surface $\phi_m = 0$, then there exists functions g^m such that $G = g^m \phi_m$.

Theorem 59.14.
If $\lambda^k \delta q^k + \mu^k \delta p_k = 0$ for every variations δq^k and δp_k tangent to the constraint surface, then

$$\lambda_k = u^m \frac{\partial \phi_m}{\partial q^k} \tag{59.44}$$

$$\mu^k = u^m \frac{\partial \phi_m}{\partial p_k} \tag{59.45}$$

for some functions u^m.

The tangent assumption means that

$$\delta\phi_m = \frac{\partial \phi_m}{\partial q^k}\delta q^k + \frac{\partial \phi_m}{\partial p_k}\delta p_k = 0.$$

Now the Hamilton equations with $\phi_m = 0$ are provided by the action

$$S(q,p,u) = \int_{t_1}^{t_2} \big(p_k \dot{q}^k - H - u^m \phi_m\big) dt \tag{59.46}$$

EaHamSqpuint

where the functions u^m are Lagrange multipliers. The equations of motion for p and u provide expressions $p_k = P_k(q,\dot{q})$ and $u^m = U^m(q,\dot{q})$ in a purely algebraic way, so that one has a reduced action

$$S_L(q) = S(q,P,U)$$

which satisfies

$$\frac{\delta S_L}{\delta q^k} = \frac{\delta S}{\delta q^k},$$

so that the variational principle for q are the same with S_L as with the original S.

One can show that the equations of motion of the action (59.46) read

$$\dot{F} = [F,H] + u^m [F, \phi_m] \tag{59.47}$$

EqdotECrochet

for any function $F = F(q,p)$.

59.2.4 Secondary constraints

For consistency, one can impose that the constraints are preserved in time: $\dot{\phi}_m = 0$. Using the evolution equation (59.47) on the function ϕ_m, one finds, for each m,

$$\dot{\phi}_m = [\phi_m, H] + u^n[\phi_m, \phi_n]$$

which are relations between q and p. In the case where that relation is independent of the ϕ_m, one has a **secondary constraint** $X(p,q) = 0$ which has to be preserved with the time, so that

$$[X, H] + u^n[X, \phi_n] = 0.$$

Once again, this is a relation between p and q.

59.3 Lie groupoids and algebroids

The reference for this section are [933, 934].

Let M be a manifold and $\mathfrak{A}$, a vector bundle over M. Let $[.,.]$ be a Lie algebra structure on $\mathfrak{A}$. Then $\mathfrak{A}$ becomes a **Lie algebroid** when we endow it by a homomorphism $\rho\colon \mathfrak{A} \to TM$ such that DefAloidItemi

(1) it induces a Lie algebra homomorphism on the sections, i.e. a $\rho\colon \Gamma(M,\mathfrak{A}) \to \Gamma(TM)$ such that $\rho[\psi,\varphi] = [\rho(\psi),\rho(\varphi)]$,

(2) for all $f \in C^\infty(M)$ and $\psi, \varphi \in \Gamma(M,\mathfrak{A})$,

$$[f\psi,\varphi] = f[\psi,\varphi] = f[\psi,\varphi] - \big(\rho(\varphi)f\big)\psi. \tag{59.48}$$ DefAlgoidCondc

The map ρ is the **anchor**. In order to be more precise for item (1), the anchor function $\rho\colon \mathfrak{A} \to TM$ induces the homomorphism $\rho'\colon \Gamma(\mathfrak{A}) \to \Gamma(TM)$ by formula

$$\rho'(\psi)(x) = \rho(\psi(x)). \tag{59.49}$$ EqInduitAncra

We will omit the prime and simply write ρ for both functions.

The algebroid $\mathfrak{A}$ admit standard coordinates (q,λ) where q are coordinates on M and λ some coordinates on the fibres (whose are vector spaces). These coordinates depends on a choice of a local base ξ of sections of $\mathfrak{A}$: the set of sections $\{\xi_i(x)\}$ is a basis of the fibre $\mathfrak{A}_x$ for each $x \in M$. IN terms of these coordinates, the Lie algebra structure of $\mathfrak{A}$ reads PgStandCoord

$$[\xi_i, \xi_j] = c_{ijk}\xi_k$$

where the $c_{ijk} \in C^\infty(M)$ are the *structure constant*. In the same way, the anchor is given in terms of functions $a_{ij} \in C^\infty(M)$ by

$$\rho(\xi_i) = a_{ij}\frac{\partial}{\partial q_j}.$$

59.3.1 Example: tangent bundle TM

If we consider $\mathfrak{A} = TM$, the anchor is $\rho = \mathrm{Id}$. Then formula (59.49) is $\rho(\psi)(x) = \psi(x)$, and the condition (59.48), with more familiar notations, reduces to

$$[fX, Y] = f[X,Y] - (Yf)X.$$

59.3.2 Example: a Lie algebra

Any Lie algebra $\mathcal{G}$ can be seen as an algebroid on a manifold $M = \{p\}$ is a manifold containing only one point. In this case, any path in M is constant and the tangent space TM reduces to only the zero vector. The identically vanishing map $\rho\colon \mathcal{G} \to TM$ is an anchor.

59.3.3 Example: gauge algebroid

We consider the following principal bundle:

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \downarrow \pi \\ & & M \end{array}$$

Let's begin to give a structure of vector bundle on the quotient TP/G. What is clear is that TP is a vector bundle $p\colon TP \to M$. The action of G on TP is

$$X_\xi \cdot g = d\tau_g X_\xi$$

where $\tau\colon P \to P$ is the action. So the quotient TP/G is taken with respect to the equivalence $X_\xi \sim d\tau_g X_\xi$. Notice that $X_\xi \in T_\xi P$ and $d\tau_g X_\xi \in T_{\xi \cdot g} P$; in fact in each class $[X_\xi]$, there is one and only one vector at each point of the fibre of ξ.

The differential $d\pi$ of the projection passes to quotient: $d\pi(X_\xi) = d\pi(X_\xi \cdot g)$. It follows that the next definition is correct:

$$\begin{aligned} d\pi &\colon TP/G \to TM \\ d\pi[X_\xi] &= d\pi X_\xi \in T_{\pi(\xi)} M. \end{aligned} \tag{59.50}$$

Now let us study the sections $\psi\colon M \to TP/G$.

Lemma 59.15.
The sections of TP/G are the G-equivariant vector fields on P. (see definition 56.56.)

Proof. We consider the map $L\colon \Gamma(M, TP/G) \to \Gamma(P, TP)$,

$$(L\psi)(\xi) = \psi(\pi(\xi))|_\xi$$

where, when $q \in TP/G$, the symbol $q|_\xi$ denotes the element of $T_\xi P$ which belongs to q. In particular, $[X_\xi]|_\xi = X_\xi$. The section $L\psi$ is G-equivariant, i.e.

$$d\tau_g (L\psi)(\xi) = (L\psi)(\xi \cdot g). \tag{59.51}$$

Indeed, $d\tau_g X_\xi \in [X_\xi]$, thus

$$d\tau_g \big[\psi(\pi\xi)\big]|_\xi \in [\psi(\pi\xi)] \cap T_{\xi \cdot g} P,$$

and there is only one element which is $\big[\psi(\pi\xi)\big]|_{\xi \cdot g} = (L\psi)(\xi \cdot g)$. □

Now, taking $d\pi$ as anchor, this construction gives an algebroid structure to TP/G. In order to prove that we have to prove that

$$[f\psi, \varphi] = f[\psi, \varphi] - \big(\rho(\varphi) f\big)\psi$$

for any choice of sections $\psi, \varphi\colon M \to TP/G$ and of function $f\colon M \to \mathbb{R}$. Here we have

$$\begin{aligned} \rho &\colon \Gamma(M, TP/G) \to \Gamma(M, TM) \\ \rho(\psi)(x) &= d\pi\big(\psi(x)\big). \end{aligned} \tag{59.52}$$

The first think to be remarked is that a good choice of local section $a\colon M \to P$ and vector fields $X, Y \in \Gamma(P, TP)$ on P, one can express ψ and φ under the form

$$\psi(x) = \big[(X \circ a)(x)\big], \qquad \varphi(x) = \big[(Y \circ a)(x)\big] \tag{59.53}$$

EqXcorrapsi

with the same a. On TP/G, we put the Lie bracket inherited from the one of TP:

$$[\psi, \varphi](x) = \Big[[X, Y] \circ a(x)\Big]. \tag{59.54}$$

If X corresponds to ψ by (59.53), we have

$$\begin{aligned}(f\psi)(x) &= \Big[f(x)(X\circ a)(x)\Big]\\ &= \Big[\big((f\circ\psi)X\big)\big(a(x)\big)\Big],\end{aligned}$$

thus $(f\circ\pi)X$ corresponds to $f\psi$ in the sense of

$$\Big((f\circ\pi)X\Big)(\xi) = f(\pi\xi)X(\xi).$$

Then we have

$$\begin{aligned}[f\psi,\varphi](x) &= \Big[\big[(f\circ\pi)X, Y\big]\circ a(x)\Big]\\ &= \Big[\big((f\circ\pi)[X,Y] - Y(f\circ\pi)X\big)\circ a(x)\Big]\\ &= \big[f(x)[X,Y]\circ a(x)\big] - \big[Y(f\circ\pi)X\circ a(x)\big].\end{aligned}$$

What lies in the bracket of the second term reads better under the form:

$$\big(Y(f\circ\pi)X\big)\circ a(x) = Y_{a(x)}(f\circ\pi)X_{a(x)}.$$

We want this thing to be equal to

$$\begin{aligned}\Big(\big(\rho(\varphi)f\big)\psi\Big)(x) &= \big(\rho(\varphi)(x)f\big)\psi(x)\\ &= d\pi\big(\varphi(x)\big)f\big[(X\circ a)(x)\big]\\ &= d\pi\big[(Y\circ a)(x)\big]f\big[(X\circ a)(x)\big].\end{aligned}$$

In the latter expression, $(Y\circ a)(x)\in T_{a(x)}P$ while $d\pi\big[(Y\circ a)(x)\big]\in T_xM$. Since $d\pi\colon TP\to TM$ fulfils $d\pi(X_\xi\cdot g) = d\pi X_\xi$, the differential of π passes to the classes (this is the reason for which we chose it a anchor) and

$$d\pi[(Y\circ a)(x)] = d\pi(Y\circ a)\in TM.$$

It remains to be proved that

$$\Big[\big(Y(f\circ\pi)X\big)\circ a(x)\Big] = \big(d\pi(Y\circ a)(x)\big)f\big[(X\circ a)(x)\big],$$

which is true because

$$\begin{aligned}Y_{a(x)}(f\circ\pi) &= (Y\circ a)(x)\\ &= (df\circ d\pi)(Y\circ a)(x)\\ &= d\pi(Y\circ a)(x)f.\end{aligned}$$

59.3.4 Poisson structure

Let us describe a Poisson structure on the dual $\mathfrak{A}^*$. For this, we put a Poisson bracket on each $\mathfrak{A}^*_x$, $x\in M$. Since the bracket only depends on differential, we just have to define it on affine functions on the fibres. Two remarks: firstly, the functions which are constant on fibres can be seen as functions on M and second, linear functions can be seen as sections of $\mathfrak{A}$ because a linear function on a vector space is equivalent to the data of a single vector. So let $f,g\colon M\to\mathbb{R}$ be two functions and $\psi,\varphi\colon M\to\mathfrak{A}$ be sections. The bracket on $C^\infty(\mathfrak{A}^*)$ is defined as

$$\{f,g\} = 0,\quad \{f,\psi\} = \rho(\psi)f,\quad \{\psi,\varphi\} = [\psi,\varphi]. \tag{59.55}$$

59.4 Lagrangian formalism

Let $\mathfrak{A}$ be a Lie algebroid on a manifold M and a function $L\colon \mathfrak{A} \to \mathbb{R}$ which we will call **Lagrangian**. The **Legendre mapping** is the fibre derivative $FL\colon \mathfrak{A} \to \mathfrak{A}^*$ given by

$$FL(a) \in \mathfrak{A}^*_{\pi(a)}, \quad FL(a)(b) = dL_a(b). \tag{59.56}$$

More precisely, the differential of L at $a \in \mathfrak{A}$ is a map from $T_a\mathfrak{A}$. In order to define $FL(a) \in \mathfrak{A}^*$, we begin to consider the restriction $L|_a$ of L to the fibre of a. The differential of $L|_a\colon \mathfrak{A}_{\pi(a)} \to \mathbb{R}$ is

$$(dL|_a)_a\colon T_a\mathfrak{A}_{\pi(a)} \to \mathbb{R},$$

but the fibre $\mathfrak{A}_{\pi(a)}$ is a vector space which can therefore be identified with its dual space. Then we have

$$(dL|_a)_a \in \mathfrak{A}^*_{\pi(a)} \subset \mathfrak{A}^*.$$

When one has a Lagrangian on $\mathfrak{A}$, one define the **action** as the function $A\colon \mathfrak{A} \to \mathbb{R}$,

$$A(v) = \langle FL(v),\, v\rangle, \tag{59.57}$$

i.e. the action of $FL(v) \in \mathfrak{A}^*_{\pi(v)}$ on $v \in \mathfrak{A}_{\pi(v)}$. The **energy** is the function

$$E = A - L.$$

The Lagrangian L is a **regular Lagrangian** if FL is a local diffeomorphism. In this case, one can bring the Poisson structure of $\mathfrak{A}^*$ on $\mathfrak{A}$. The resulting Poisson structure on $\mathfrak{A}$ is the **Lagrange-Poisson structure**. For this, we have to define $\{\xi, \eta\}$ when ξ, $\eta \in C^\infty(\mathfrak{A})$. The natural definition is

$$\{\xi, \eta\}_{\mathfrak{A}} := \{FL(\xi), FL(\eta)\}_{\mathfrak{A}^*} \tag{59.58}$$

with the following definition of $FL(\xi) \in C^\infty(\mathfrak{A}^*)$:

$$FL(\xi)(\omega) = \xi(FL^{-1}(\omega)). \tag{59.59}$$

The latter definition says that when $\omega \in \mathfrak{A}^*_{\pi(a)}$ the element $FL^{-1}(\omega)$ is the element a such that $(dL|_a)_a = \omega$ (this is an equality in $\mathfrak{A}^*_{\pi(a)}$).

Since we have a Poisson structure, we can consider the Hamiltonian field (see definition (59.5)) corresponding to the energy function $E\colon \mathfrak{A} \to \mathbb{R}$. This is a vector field on $\mathfrak{A}$. This field is the **Lagrangian vector field**

In standard coordinates (see page 3465), the Lagrangian is a function $(q, \lambda) \mapsto L(q, \lambda)$. In order to build FL, we first restrict F to a fibre; so L becomes a function $\lambda \mapsto L(q_0, \lambda)$ and thus the derivatives which appear in FL are

$$\mu_i = \frac{\partial L}{\partial \lambda_i}.$$

Now we try to express the Poisson structure on $\mathfrak{A}$ in the standard coordinates. Fist, q_i is a function on $\mathfrak{A}$ which associates to one point the component i of its projection. If ω and η both belong to $\mathfrak{A}^*_x$, the elements $FL^{-1}(\omega)$ and $FL^{-1}(\eta)$ both belongs to the same fibre $\mathfrak{A}_x$. Thus $FL(q_i)$ is a function that is constant on the fibres. So

$$\{q_i, q_j\} = 0.$$

59.5 Groupoids

A set Γ is a **groupoid** when we consider some maps

— $\alpha, \beta\colon \Gamma \to \Gamma_0 \subset \Gamma$,

— $m\colon \Gamma_2 \to \Gamma$ where $\Gamma_2 = \{(x, y) \in \Gamma \times \Gamma \text{ tel que } \beta(y) = \alpha(x))\}$,

— $i\colon \Gamma \to \Gamma$

such that

(1) $m(x, m(y, z))$ is defined if and only if $m(m(x, y), z)$ is defined, and in this case, they are equal,

(2) $m(\beta(x), x) = m(x, \alpha(x)) = x$,

(3) $m(x, i(x))$ and $m(i(x), x)$ are defined to be respectively equal to $\beta(x)$ and $\alpha(x)$.

Notice that the fact for $m(x, y)$ to be defined means that $\alpha(x) = \beta(x)$. As notations and terminology, we adopt the following conventions. The map i is the *inversion*, m is the *multiplication* and is denoted by a dot: $m(x, y) = x \cdot y$. We also write $i(x) = i^{-1}$.

The first hypothesis is called *associativity*.

Lemma 59.16.
$\forall\, x \in \Gamma$, we have

$$\alpha(\beta(x)) = \beta(x) \tag{59.60a}$$
$$\beta(\alpha(x)) = \alpha(x). \tag{59.60b}$$

Proof. From definition of a groupoid, $\beta(x) \cdot x = x$, so $\alpha(\beta(x)) = \beta(x)$. The other statement comes in the same way.

Notice that the conclusion does not come from the equality $\beta(x) \cdot x = x$, but only from the existence of the product $\beta(x) \cdot x$. □

Lemma 59.17.
If $(x, y) \in \Gamma_2$, we have

$$\beta(x \cdot y) = \beta(x) \tag{59.61a}$$
$$\alpha(x \cdot y) = \alpha(y) \tag{59.61b}$$

Proof. Since $(x, y) \in \Gamma_2$, one can write $x \cdot y$. Using $x = \beta(x) \cdot x$, and associativity we find:

$$x \cdot y = \big(\beta(x) \cdot x\big) \cdot y = \beta(x) \cdot (x \cdot y).$$

Existence of the last product implies $\alpha(\beta(x)) = \beta(x \cdot x)$. Using the equality $\alpha(\beta(x)) = \beta(x)$, we find the second relation. The first one is proven by the same:

$$x \cdot y = x \cdot \big(y \cdot \alpha(y)\big) = (x \cdot y) \cdot \alpha(x).$$

□

Lemma 59.18.
For all $x \in \Gamma$, we have

$$\alpha(\alpha(x)) = \alpha(x) \tag{59.62a}$$
$$\beta(\beta(x)) = \beta(x). \tag{59.62b}$$

Proof. Using formula $x \cdot \alpha(x) = x$ in itself, and using associativity,

$$(x \cdot \alpha(x)) \cdot \alpha(x) = x \cdot \big(\alpha(x) \cdot \alpha(x)\big) = x.$$

The existence of the product $\alpha(x) \cdot \alpha(x)$ and the fact that $\beta \circ \alpha = \alpha$ give the result. □

Lemma 59.19.
Let us mention the following other properties:

$$\alpha(x^{-1}) = \beta(x) \qquad \beta(x^{-1}) = \alpha(x) \tag{59.63}$$
$$\alpha(x) \cdot \alpha(x) = \alpha(x) \qquad \beta(x) \cdot \beta(x) = \beta(x) \tag{59.64}$$

the simplification *rule:*

$$x \cdot y_1 = x \cdot y_2 \Rightarrow y_1 = y_2 \tag{59.65}$$
$$x_1 \cdot y = x_2 \cdot y \Rightarrow x_1 - x_2, \tag{59.66}$$

and the corollary

$$(x^{-1})^{-1} = x. \tag{59.67}$$

Proof. No proof. □

Proposition 59.20.
The set Γ_0 is the set of fixed points of α and β.

Proof. We proof that $\Gamma_0 = \{x \in \Gamma \text{ tel que } \alpha(x) = x\}$; the same is true for β. The definition of Γ_0 is to be the image of α. Let $x \in \Gamma_0$: there exists a $y \in \Gamma$ such that $x = \alpha(y)$. So $\alpha(x) = \alpha(\alpha(y)) = \alpha(y) = x$.

For the reciprocal, let $x = \alpha(x)$. Then $x \in \Gamma_0$ because x belongs to the image of α — for instance, the image of itself. □

59.5.1 Example: when $\Gamma_0 = \{e\}$

Let us prove that in the case when Γ_0 reduces to only one point $\{e\}$, the groupoid Γ is a group whose unit is e.

First, remark that for all $x \in \Gamma$, we have $\alpha(x) = \beta(x) = e$, so that $\Gamma_2 = \Gamma$ and the multiplication is everywhere defined. Associativity is not a problem. The map $x \mapsto i(x)$ is the inverse because $x \cdot i(x) = \beta(x) = e$ and $i(x) \cdot x = \alpha(x) = e$. We also have $x \cdot e = e \cdot x = x$ because $e = \alpha(x)$ and $x \cdot \alpha(x) = x$.

59.5.2 Example: the null groupoid

A **null groupoid** is a groupoid in which $\Gamma_0 = \Gamma$. In this case, since Γ_0 is the set of fix points of α and β, we have

$$\alpha = \beta = \mathrm{Id}\,.$$

In order for $x \cdot y$ to exist, we need $\beta(y) = \alpha(x)$, which in the case of the null groupoid gives $x = y$. So the only products that are defined are

$$x \cdot x = x.$$

59.5.3 The case $\alpha = \beta$

The case $\alpha = \beta$ regroup the two preceding cases. We will prove that for each $u \in \Gamma_0$, the set $\alpha^{-1}(u)$ is a group with u as unit. Indeed let x, $y \in \alpha^{-1}(u)$; it is clear that $x \cdot y$ exists because $\beta(y) = \alpha(y) = x = \alpha(x)$. So the product is defined everywhere. Proof of the fact that u is the unit is easy:

$$x \cdot u = x = x \cdot \alpha(x) = x$$
$$u \cdot x = \beta(x) \cdot x = x,$$

and

$$x \cdot i(x) = \beta(x) = u$$
$$i(x) \cdot x = \alpha(x) = u.$$

59.5.4 An example on a vector bundle

We consider a vector bundle $\pi\colon E \to M$ and

$$\Gamma_0 = \{o_x \text{ tel que } x \in M\},$$

the set of the zero of each fibre. As groupoid law, we choose the addition in the fibres: when $\pi(v) = \pi(w)$, we define $v \cdot w = v + w$, and as map $\alpha = \beta$, we naturally choose

$$\alpha(v) = \beta(v) = o_x$$

if $v \in E_x$. In this case, $\alpha^{-1}(o)$ is the fibre of o which is a group for the addition.

59.5.5 Orbits

Let Γ be a groupoid and $u \in \Gamma_0$. The **isotropy group** of u is

$$\Gamma_u = \alpha^{-1}(u) \cap \beta^{-1}(u)$$

In order to prove that it is a group, remark that if x, $y \in \Gamma_u$,

$$\alpha(x) = \alpha(y) = \beta(x) = \beta(y) = u,$$

in particular, $x \cdot y$ exists and Γ_u has a law. It is easy to prove that u is th unit.

Notice that $\beta(\alpha^{-1}(x)) = \alpha(\beta^{-1}(x))$. Indeed if $y \in \beta(\alpha^{-1}(x))$, there exists a z such that $y = \beta(z)$ and $\alpha(z) = x$. We have to find a z' such that $y = \alpha(z')$ and $\beta(z') = x$. The element $z' = \beta(z)$ works.

The set $\beta(\alpha^{-1}(x) = \alpha(\beta^{-1}(x))$ is the **orbit** of Γ trough x.

Proposition 59.21.
The set of orbits of Γ *is a partition of* Γ_0.

Proof. The proof is as easy as I want not to give you. □

59.5.6 Morphism

If Γ and Γ' are two groupoids, the map $f\colon \Gamma \to \Gamma'$ is a **morphism** if for each existing product $x \cdot y$ in Γ, we have

$$f(x) \cdot f(y) = f(x \cdot y).$$

We see that automatically

$$\beta'(f(x)) \cdot f(x) = f(x) = f(\beta(x) \cdot x) = f(\beta(x)) \cdot f(x);$$

so that the simplification rule gives

$$\beta' \circ f = f \circ \beta \tag{59.68a}$$
$$\alpha' \circ f = f \circ \alpha. \tag{59.68b}$$

59.6 Lie groupoid

A **Lie groupoid** is a (maybe non Hausdorff) manifold endowed with a groupoid structure such that

(1) the set Γ_0 is a Hausdorff manifold,
(2) the maps α and β are differentiable submersions,
(3) the multiplication $m\colon \Gamma_2 \to \Gamma$ is differentiable,
(4) the inversion $x \to x^{-1}$ is a diffeomorphism.

Chapter 60

Hilbert spaces

References for Hilbert spaces are [935, 489].
Do you know what is normed, complete and yellow? Answer in the footnote [1].

60.1 Basis and orthonormal systems

A **sesquilinear map** map on a complex vector space V is a map $(.,.)\colon V \times V \to \mathbb{C}$ such that

$$(x+y, x'+y') = (x,x') + (x,y') + (y,x') + (y,y'),$$
$$(\lambda x, \mu y) = \bar{\lambda}\mu(x,y).$$

DefBanchHilbertpre

Definition 60.1. *(1) A **Banach space** is a complete and normed vector space.*

*(2) A **pre-Hilbert** is a complex vector space with an inner product*

*(3) An **Hilbert space** is a complex Banach space whose norm is induced from an inner product. Equivalently, it is a pre-Hilbert space in which the topology is complete.*

From a pre-Hilbert space, one can construct an Hilbert space by **completion**. The completion of a pre-Hilbert space H_0 is the set of all the Cauchy sequences in H_0. It turns out that this set is an Hilbert space. Points in H_0 are identified with Cauchy sequences that converge in H_0.

Remarque 60.2.
In some literature[936], a pre-Hilbert space is defined as a complex vector space endowed with a sesquilinear positive form. That is a sesquilinear form such that $\langle x,x\rangle \geqslant 0$. In this case the map $x \mapsto \langle x,x\rangle^{1/2}$ is only a seminorm: there could be elements with vanishing norm.

If H_0 is such a space, before to take its completion, we have to take its **separation**. The separation is as follows. Let $I = \{x \in H_0 \text{ tel que } \langle x,x\rangle = 0\}$. The quotient space H/I is then a pre-Hilbert in the sense of definition 60.1.

A subset $\mathcal{S}$ of a pre-Hilbert $\mathcal{P}$ is **total** if 0 is the only element in $\mathcal{P}$ to be orthogonal to each element of $\mathcal{S}$, in other words: $\langle z|s\rangle = 0$ for any $s \in \mathcal{P}$ implies $z = 0$.

PropCconvminiv

Proposition 60.3.
If $C \subset \mathscr{H}$ is a closed convex subset of the Hilbert space $\mathscr{H}$ and if $v \in \mathscr{H}$, there exists one and only one $c_C \in C$ such that

$$\|v - v_C\| = \min_{w \in C} \|v - w\|,$$

i.e. v_C minimizes the distance between v and C.

Proof. No proof. □

1. A bananach space!

Proposition 60.4.
In the same setting that proposition 60.3, with the assumption that C is a vector subspace of $\mathscr{H}$, we have $v - v_C \in C^\perp$.

Proof. Let v_C be the minimizer given by proposition 60.3; by definition for every $w \in C$, the distance between $v_C + tw$ and v is bigger than the one between v_C and v. In particular, the derivative of $\|v_c + tw - v\|^2$ with respect to t vanishes on $t = 0$. A small computation provides

$$\mathrm{Re}\left(\langle v - v_C, w\rangle\right) = 0$$

for every w. Doing the same with iw, we find that the imaginary part of $\langle v - v_C, w\rangle$ vanishes too, so that the proposition is proved. □

We conclude that when C is a convex closed vector subspace of $\mathscr{H}$, the latter accepts the decomposition $\mathscr{H} = C \oplus C^\perp$.

A sequence (x_n) in a Hilbert space $\mathscr{H}$ is an **orthonormal basis** if

— $\langle x_i | x_j\rangle = \delta_{ij}$,

— the sequence (n_n) is total.

An Hilbert space $\mathscr{H}$ is **separable** if it posses a total sequence. The link between this and the topological definition of *separable* is not completely easy. A first step is done in lemma 48.8.

Theorem 60.5.
An Hilbert space is separable if and only if it posses an orthonormal basis.

A classical but powerful theorem about orthonormal basis:

Theorem 60.6.
Let $\mathscr{H}$ be an infinite dimensional Hilbert space and a sequence (x_n) in $\mathscr{H}$. Then the following propositions are equivalent:

(1) (x_n) is an orthonormal basis,

(2) $\sum_{k=1}^{\infty} |\langle x|x_k\rangle|^2 = \|x\|^2$ for every $x \in \mathscr{H}$,

(3) $\sum_{k=1}^{\infty} \langle x_k|x\rangle |x_k = x$ for every $x \in \mathscr{H}$.

We will sometimes denote by $|\psi\rangle$ the vector ψ and $\langle\phi|$ the form $|\psi\rangle \to \langle\phi, \psi\rangle$. This notation is mainly used in the physics literature.

Let $\{v_\alpha\}$ be a maximal orthogonal set in $\mathscr{H}$, then each $v \in \mathscr{H}$ can be written under the form

$$v = \sum_\alpha \langle v_\alpha, v\rangle v_\alpha, \tag{60.1}$$

Eqvsumvalphavmaxorth

and the norm is given by

$$\|v\|^2 = \sum_\alpha |v_\alpha, v|^2. \tag{60.2}$$

Notice that the latter sum is absolutely convergent, while the first one is not. The sum in the right hand side of (60.1) is **unconditionally convergent**. One says that a sum $\sum_{\alpha\in A} X_\alpha = X$ unconditionally in a Banach space if $\forall \epsilon > 0$, there exists a finite subset F of A such that for every finite subset F' containing F,

$$\| \sum_{\alpha\in F'} (X_\alpha - X)\| \leqslant \epsilon.$$

That notion of convergence is the good one in Hilbert space where one does not always have absolute convergence.

ThoRiesz

Theorem 60.7 (Riesz-Fisher)**.**
If $\varphi\colon \mathscr{H} \to \mathbb{C}$ is a continuous linear functional on $\mathscr{H}$, there exists a vector $w \in \mathscr{H}$ such that

$$\varphi(v) = \langle v, w\rangle$$

for every $v \in \mathscr{H}$.

Proof. First, remark that, because of continuity, $\ker\varphi$ is a closed subspace of $\mathscr{H}$, so that $\mathscr{H} = \ker(\varphi)\oplus\ker(\varphi)^\perp$. Let z be any non vanishing element of $\ker(\varphi)^\perp$. In that case, the map $v \mapsto \langle z, v\rangle$ has the same kernel as φ. But we know that two linear maps with the same kernel are related by a simple multiplication by a constant scalar. A rescaling of z by that scalar provides the answer of the theorem.

In order to be complete, notice that the kernel of $v \mapsto \langle z, v\rangle$ has codimension one in $\mathscr{H}$ because the image has dimension one. □

A more complete version of that theorem is 25.18.

60.2 Operators on Hilbert spaces

About spectral theory: [937].

DefVecteurTrace

Definition 60.8.
*A vector $v \in \mathscr{H}$ is a **trace vector** if the functional*

$$T \mapsto \langle v, Tv\rangle \tag{60.3}$$

is a trace (that is $\omega(T^*T) = \omega(TT^*)$).

Definition 60.9.
*An operator $T\colon X \to Y$ between two Banach spaces is **closed** if for every sequence $(x_n) \in D(T)$ such that $x_n \to x \in X$ and $Tx_n \to y \in Y$ we have $x \in D(a)$ and $Ax = y$.*

PropoOpFermableLim

Proposition 60.10.
An operator admits a closure if and only of for every pair of sequences (x_n) and (y_n) in $D(T)$ with $\lim x_n = \lim y_n = x$ and such that Tx_n and Ty_n converge we have $\lim Tx_n = \lim Ty_n$.

In this case the **closure** of T is defined by $Tx = \lim Tx_n$.

60.2.1 Adjoint, unitary and projection operator

Let a bounded operator $T\colon \mathscr{H}_1 \to \mathscr{H}_2$. We define the adjoint operator $T^*\colon \mathscr{H}_2 \to \mathscr{H}_1$ in the following way. Let $v_2 \in \mathscr{H}_2$; we define

$$\begin{aligned}\phi\colon \mathscr{H}_1 &\to \mathbb{C}\\ v_1 &\mapsto \langle Tv_1, v_2\rangle.\end{aligned} \tag{60.4}$$

This is an element of $\mathscr{H}_1'$, so that the Riesz's theorem 60.7 produces an elemen t$y \in \mathscr{H}_1$ such that

$$\phi(v_1) = \langle y, v_1\rangle \tag{60.5}$$

for every $v_1 \in \mathscr{H}_1$. We define T^*v_2 to be that element.

DEFooERIYooIIRLuy

Definition 60.11.
In short, T^* is defined by the formula

$$\langle Tv_1, v_2\rangle = \langle v_1, T^*v_2\rangle \tag{60.6}$$

for every $v_1 \in \mathscr{H}_1$ and $v_2 \in \mathscr{H}_2$.

LemTTzepoT

Lemma 60.12.
If $TT^* = 0$, then $T = 0$.

Proof. The assumption makes that for every $v \in \mathscr{H}$,

$$0 = \langle A^*Av, v\rangle = \langle Av, Av\rangle = \|Av\|^2. \tag{60.7}$$

That proves that $Av = 0$ for every v, or that $A = 0$. □

An operator such that $T^*T = \mathbb{1}$ is an **isometry**, but is not always invertible. An invertible isometry is an **unitary operator** and fulfills $U^*U = \mathbb{1} = UU^*$. A **projection** is an operator P such that $P = P^*$ and $P^2 = P$.

A **partial isometry** is a linear map $W : V_1 \to V_2$ between two vector spaces such that there exists a closed subspace $K_1 \subset V_1$ with $\langle W\psi, W\phi\rangle_2 = \langle \psi, \phi\rangle_1$ for all $\psi, \phi \in K_1$ and $W = 0$ on $K_1^\perp$. The most immediate property is that W is unitary between K_1 and WK_1.

LemPartIsomCstar

Lemma 60.13.
The element $A \in \mathfrak{A}$ is a partial isometry between $\mathfrak{A}$ and itself if and only if A^*A is a projection.

Proof. If we pose $V_1 = V_2 = \mathfrak{A}$ in the definition of a partial isometry, the fact for p to be a projection is the existence of $K_1 \subset \mathfrak{A}$ such that $\langle pA, pB\rangle = \langle A, B\rangle$ for every A, $B \in K_1$. For such a K_1, we have $p^*p = \mathrm{Id}\,|_{K_1}$ and $p|_{K_1^\perp} = 0$. □

60.2.2 Topology on space of continuous endomorphism

subsec_topomL

Let $\mathscr{H}$ be a Hilbert space and $\mathcal{L}$ the space of continuous endomorphism on $\mathscr{H}$. The **uniform topology** is the one of the norm $T \to \|T\|$; the **strong topology** is given by semi-norms $T \to \|T\xi\|$ (one semi-norm for each $\xi \in \mathscr{H}$); the **weak topology** is given by semi-norms $T \to |\langle T\xi, \eta\rangle|$.

For a sequence $A_n \in \mathcal{L}(\mathscr{H})$, we write $A_n \to 0$ in the sense of **strong convergence** in $\mathcal{L}$ if for all neighbourhood V of 0, there exists a $N \in \mathbb{N}$ such that $A_n \in V$ for all $n \geqslant N$. A neighbourhood of 0 is of the form

$$V = \{T \in \mathcal{L}(\mathscr{H}) \text{ tel que } s_\xi(T) < \epsilon\}$$

with s_ξ, the strong semi-norm defined by ξ: $s_\xi(T) = \|T\xi\|$. So we have $A_n \to 0$ in the sense of strong topology in $\mathcal{L}(\mathscr{H})$ if and only if for all $\xi \in \mathscr{H}$, $\|A_n\xi\| \to 0$.

60.2.3 Compact operators

Definition 60.14.
*The **adjoint** A^* of the operator A on a Hilbert space is defined by the property $\langle A\psi, \phi\rangle = \langle \psi, A^*\phi\rangle$. It defines an involution on $\mathfrak{B}(\mathcal{B})$. An element x in an involutive algebra $\mathfrak{A}$ is **hermitian** if $x^* = x$. In the case of $\mathfrak{A} = \mathcal{B}(\mathscr{H})$ –the Banach space of the bounded operators on a Hilbert space $\mathscr{H}$— we say **self-adjoint**.*

As usual notations, $\mathscr{H}$ denotes a Hilbert space, $\mathfrak{K}$ the space of compact operators and $\mathfrak{B}$ the one of bounded operators. Let us recall some properties of compact operators. An operator is compact when it can be norm approximated by operators of finite rank, more precisely, we define the **characteristic values** of the operator T as

$$\mu_n(T) = \inf\{\|T - R\| \text{ where } R \text{ is an operator of range} \leqslant n\} \tag{60.8}$$

Defmuncaharacinfn

Lemma 60.15.
The characteristic values $\mu_n(T)$ are the eigenvalues of the operator $|T| = (T^*T)^{1/2}$ classified in decreasing order with multiplicity.

We define $\sigma_n(t) = \sum_{k=0}^n \mu_k(T)$. In the case of an infinitesimal of order 1, one has to expect a divergence

$$\sigma_n(T) = O(\ln n).$$

Following the lemma, we have $\mu_0(T) \geqslant \mu_1(T) \geqslant \cdots$. One says that the operator T is **compact** if $\lim_{n\to\infty} \mu_n(T) = 0$.

LemAstAcomAcomp

Lemma 60.16.
If A is an operator on $\mathscr{H}$ such that AA^* is compact, then A is compact.

Proof. No proof. □

Proposition 60.17.
Let T be a compact operator on a Hilbert space $\mathscr{H}$. Then

(1) The spectrum $\sigma(T)$ is discrete and has no limit point other than eventually zero,

(2) any non zero element in $\sigma(T)$ is eigenvalue of finite multiplicity.

Let us point out that a compact operator has no specially any eigenvalues.

Proposition 60.18.
Let T be a compact and self-adjoint operator on the Hilbert space $\mathscr{H}$. There exists a complete orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$ of $\mathscr{H}$ such that $T\phi_n = \lambda_n\phi_n$ and $\lambda_n \to 0$ when $n \to \infty$.

Proposition 60.19.
Let T be a compact operator on $\mathscr{H}$. Then we have a (norm) uniform convergent expansion

$$T = \sum_{n\geqslant 0} \mu_n(T)\psi_n\langle\phi_i, .\rangle.$$

where $0 \leqslant \mu_{j+1}(T) \leqslant \mu_j(T)$ and $\{\psi_n\}_{n\in\mathbb{N}}$ is orthogonal to $\{\phi_n\}_{n\in\mathbb{N}}$.

This proposition allows us to decompose T as

$$T = U|T|$$

where $|T| = \sqrt{T^*T}$. Hence the $\mu_n(T)$ are eigenvalues of $|T|$ and

$$\lim_{n\to\infty} \mu_n(T) = 0$$

because $|T|$ is compact and self-adjoint. The ϕ_n are the corresponding eigenvectors and

$$\psi_n = U\phi_n.$$

In this setting, we say that the $\mu_n(T)$ are the **characteristic values** of T. We have $\mu_0(T) = \|T\|$.

pg_char_inv_U

Remark that the characteristic values $\mu_n(T)$ are invariant under $T \to UT$ when U is unitary. Indeed if μ is eigenvalue of T^*T with eigenvector ψ, then $U^*\psi$ is eigenvector of U^*TU with the same eigenvalue μ.

Proposition 60.20.
The operator T is compact if and only if for all $\epsilon > 0$, there exists a finite dimensional subspace $E \subset \mathscr{H}$ such that

$$\|T\|_{E^\perp} \leqslant \epsilon.$$

prop_comp_ini

LemAmtuBcompaBcm

Lemma 60.21.
Let A be a compact operator. If B is an invertible operator such that $(A + \mathbb{1})B$ is compact, then B is compact.

¡¡ Avertissement/question au lecteur !! 60.22
The proof is mine; without guarantee.

Proof. Suppose that B is not compact. There exists a $\sigma > 0$ such that for every finite dimensional subspace G, we have

$$\sup_{\substack{x\in G^\perp \\ \|x\|=1}} \|Bx\| > \sigma. \tag{60.9}$$

Let $\epsilon > 0$. Since A is compact, there exists a finite dimensional subspace F such that $\|A\|_{F^\perp} < \epsilon$. For $x \in F^\perp$, have

$$\|(A + \mathbb{1})x\| \geqslant \Big|\|Ax\| - \|x\|\Big| = 1 - \|Ax\| \geqslant 1 - \epsilon, \tag{60.10}$$

so that

$$\inf_{\substack{x\in F^\perp\\ \|x\|=1}} \|(A+\mathbb{1})x\| \geqslant 1-\epsilon. \tag{60.11}$$

In the same way from compactness of $(A+\mathbb{1})B$, there exists a finite dimensional subspace E such that

$$\sup_{\substack{x\in E^\perp\\ \|x\|=1}} \|(A+\mathbb{1})Bx\| \leqslant \epsilon. \tag{60.12}$$

Using these properties,

$$\begin{aligned}
\epsilon \geqslant \sup_{\substack{x\in E^\perp\\ \|x\|=1}} \|(A+\mathbb{1})x\| &\geqslant \sup_{\substack{x\in E^\perp\\ Bx\in F^\perp\\ \|x\|=1}} \|(A+\mathbb{1})x\| \\
&= \sup_{\substack{y\in F^\perp\\ B^{-1}y\in E^\perp\\ \sigma<\|y\|<\|B_{E^\perp}\|}} \|(A+\mathbb{1})y\| \\
&\geqslant \inf_{\substack{y\in F^\perp\\ B^{-1}y\in E^\perp\\ \sigma<\|y\|<\|B_{E^\perp}\|}} \|(A+\mathbb{1})y\| \\
&\geqslant \inf_{\substack{y\in F^\perp\\ \sigma<\|y\|<\|B_{E^\perp}\|}} \|(A+\mathbb{1})y\| \\
&\geqslant \inf_{\substack{y\in F^\perp\\ \|y\|=\sigma}} \|(A+\mathbb{1})y\| \\
&\geqslant \sigma(1-\epsilon).
\end{aligned} \tag{60.13}$$

It is now sufficient to choose ϵ is such a way that $\frac{\epsilon}{1-\epsilon} < \sigma$ in order to get a contradiction. □

Theorem 60.23 (Spectral theorem)**.**
If T is a compact self-adjoint operator on the Hilbert space $\mathscr{H}$, then there exists an orthogonal basis of $\mathscr{H}$ of eigenvectors of T. The eigenvalues are moreover real and the sequence converges to zero.

Proof. No proof. □

Proposition 60.24.
Two other characterisations of compact operators:

(1) the operator T is compact if and only if it is the limit (for the operator norm) of finite rank operators,

(2) if T is an operator over $L^2(X)$ and if T can be written under the form

$$(Tf)(x) = \int_X k(x,y)f(y)dy,$$

*where k is a square summable function on $X\times X$, then T is compact. Such an operator is said to be **Hilbert-Schmidt**, and every compact operators over $L^2(X)$ are* not *of that form.*

Let T be a bounded operator on $\mathscr{H}$. The **singular values** of T are defined by

$$\mu_j(T) = \inf_{\dim(V)=j} \sup_{v\perp V} \frac{\|Tv\|}{\|v\|}. \tag{60.14}$$

The first singular value gives the operator norm: $\mu_0(T) = \|T\|$.

Proposition 60.25.
The operator T is compact if and only if $\lim_{j\to\infty}\mu_j(T) = 0$.

Lemmulamequ

Lemma 60.26.
If T is a positive[2] *compact operator and if (λ_j) is the sequence of eigenvalues sorted in decreasing order with multiplicity, then*

$$\mu_j(T) = \lambda_j(T) \tag{60.15}$$

to the condition that the eigenvalues are numbered from λ_0 instead of λ_1.

LemIneqscmpborn

Lemma 60.27.
If T_1 and T_2 are two compact operators and if S is a bounded operator, then

$$\begin{aligned}\mu_j(T_1+T_2) &\leqslant \mu_j(T_1)+\mu_j(T_2) \leqslant \mu_{2j}(T_1+T_2)\\ \mu_j(ST) &\leqslant \|S\|\mu_j(T)\\ \mu_j(TS) &\leqslant \|S\|\mu_j(T).\end{aligned}$$

Proof. No proof. □

Let $\mathscr{H}$ be an Hilbert space and $\mathfrak{B}(\mathscr{H})$ be the set of bounded operators over $\mathscr{H}$. The **trace class** operator ideal is

$$\mathscr{L}^1(\mathscr{H}) = \{T \in \mathfrak{B}(\mathscr{H}) \text{ tel que } \sum \mu_j(T) < \infty\}. \tag{60.16}$$

Such an operator is always compact and the inequalities of lemma 60.27 assure that $\mathscr{L}^1$ is an ideal.

An interesting property is that $\mathscr{L}^1$ is not norm-closed, actually its norm closure is the full $\mathfrak{B}(\mathscr{H})$.

From definition of singular values, if $\{v_1, \cdots, v_n\}$ is any orthogonal set in $\mathscr{H}$, then

$$\sum_{j=0}^{N} |\langle v_j, Tv_j\rangle| \leqslant \sum_{j=0}^{N} \mu_j(T). \tag{60.17}$$

Eqineqstrday

So we can give the definition of a trace. If $T \in \mathscr{L}^1$, the **trace** is given by

$$\mathrm{Tr}(T) = \sum_{j=1}^{\infty} \langle v_j, Tv_j\rangle \tag{60.18}$$

where $\{v_j\}$ is any orthonormal basis of $\mathscr{H}$. The relation (60.17) makes the sum absolutely convergent and the independent on the choice of basis, as can see by replacing v_j by $A_{ji}v_i$ and using the fact that $(A^t)_{lj}A_{jk} = \delta_{lk}$.

An interesting property of the trace is

$$\mathrm{Tr}(ST) = \mathrm{Tr}(TS) \tag{60.19}$$

whenever $S \in \mathfrak{B}(\mathscr{H})$ and $T \in \mathscr{L}^1(\mathscr{H})$.

60.2.4 Hilbert-Schmidt operators

If S and T are Hilbert-Schmidt operators, one can show that ST is a trace class operator. An operator T over $L^2(X)$ is Hilbert-Schmidt if and only if $\sum \mu_j(T)^2 < \infty$.

Lemma 60.28.
Let M be a closed manifold endowed with a smooth measure. If $k \in C^\infty(M)$, then the operator defined by

$$(Tf)(x) = \int_M k(x,y)f(y)dy$$

is a trace class operator and we have

$$\mathrm{Tr}(T) = \int_M k(x,x)dx. \tag{60.20}$$

2. Notice that, for a positive operator, $\langle Tv, v\rangle \geqslant 0$, so that T is self-adjoint too.

60.2.5 The Schatten-von Neumann ideal

The **Schatten-von Neumann ideal** is the set

$$\mathscr{L}^p(\mathscr{H}) = \{\text{compact operator } T \text{ tel que } \sum_{n=0}^{\infty} \mu_n(T)^p < \infty\}$$

Interesting properties of this set (including the fact that it is an ideal) are proven by virtue of **Hölder inequality**: when p and q are reals such that $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\sum_k |u_k||v_k| \leqslant \Big(\sum_k |u_k|^p\Big)^{1/p} \Big(\sum_k |v_k|^q\Big)^{1/q}. \tag{60.21}$$

There also exists an integral version:

$$\int |fg| \leqslant \Big(\int |f|^p\Big)^{1/p} \Big(\int |g|^q\Big)^{1/q}. \tag{60.22}$$

¡¡ Avertissement/question au lecteur !! 60.29
We should find a precise statement with precise hypothesis for that inequality.

One can prove that when $\sum_{j=1}^k \frac{1}{p_j} = \frac{1}{q}$, we have

$$\mathscr{L}^{p_1}\mathscr{L}^{p_2}\dots\mathscr{L}^{p_k} \subset \mathscr{L}^q. \tag{60.23}$$ EqPropLLLsyn

Lemma 60.30.
The space $\mathscr{L}^q(\mathscr{H})$ is a left ideal.

Proof. Let $T \in \mathscr{L}^q(\mathscr{H})$: $\sum_n \mu_n(T)^q < \infty$. If a is an other linear operator on $\mathscr{H}$, $\mu_n(aT) = \inf\{\|aT - R\|\,; \mathrm{Rank}(R) \leqslant n\} = \inf\{\|aT - aR\|; \mathrm{Rank}(R) \leqslant n\}$ because $\mathrm{Rank}(aR)$ and $\mathrm{Rank}(a^{-1}R)$ are always lower than $\mathrm{Rank}(T)$. Thus $\mu_n \leqslant \inf\{\|a\|\|T - R\|; \mathrm{Rank}(R) \leqslant n\} = \|a\|\mu_n(T)$. □

The **trace** of an operator is defined on $T \in \mathscr{L}(\mathscr{H})$ by

$$\mathrm{Tr}(T) = \sum_n \langle T\xi_n, \xi_n\rangle \tag{60.24}$$

where ξ_n runs over an orthonormal basis. One can prove that $\mathrm{Tr}(T)$ does not depend on the choice of this basis.

Proposition 60.31.
When T is compact and positive, on has

$$\mathrm{Tr}(T) = \sum_n \mu_n(T).$$

Proof. No proof. □

I found the following definitions in [938].

Definition 60.32.
*A sesquilinear form q on an Hilbert space $\mathscr{H}$ is **symmetric** if $q(u,v) = \overline{q(v,u)}$ or, equivalently, if $q(u,u) \in \mathbb{R}$ for every $u, v \in \mathrm{Dom}(q)$. It is **positive** if $q(u) = q(u,u) \geqslant 0$ for every $u \in \mathrm{Dom}(q)$.*

The space $\mathrm{Dom}(q)$ is endowed by the inner product

$$\langle u, v\rangle_q = \langle u, v\rangle_{\mathscr{H}} + q(u,v). \tag{60.25}$$ EqInnerProdqDomainsq

Most of time, when we speak about topology on $\mathrm{Dom}(q)$, we are speaking about the topology of that norm. The form q is **closed** is $\mathrm{Dom}(q)$ is complete (for the topology of the inner product (60.25)). In that case, $\mathrm{Dom}(q)$ is itself an Hilbert space.

DefFormCoreDomq

Definition 60.33.
*A set $D \subset \mathrm{Dom}(q)$ which is q-norm-dense in $\mathrm{Dom}(q)$ is a **form core** for q.*

Definition 60.34.
Consider the algebra of bounded operators on an Hilbert space $\mathscr{H}$. Let $\mathrm{Spec}(T)$ be the spectrum of T.

*The **point spectrum** of T, $\mathrm{Spec}_P(T)$, is the set of eigenvalues. This is the set of $\lambda \in \mathbb{C}$ such that $T - \lambda\mathbb{1}$ is not injective.*

*The **continuous spectrum**, $\mathrm{Spec}_C(T)$, is the subset of its spectrum given by the values λ such that $T - \lambda\mathbb{1}$ is injective but for which the image of $T - \lambda\mathbb{1}$ is a dense proper subspace of $\mathscr{H}$.*

*The **residual spectrum**, $\mathrm{Spec}_R(T)$, is the part of the spectrum that remains. So $\lambda \in \mathrm{Spec}_R(T)$ if $T - \lambda\mathbb{1}$ is injective and the closure $\overline{\mathrm{Image}(T - \lambda\mathbb{1})}$ is a proper subspace of $\mathscr{H}$.*

60.2.6 Normal operators on Hilbert space

Many properties of normal operators can be found in [937].

DefFQFKZbB

Definition 60.35.
*An operator T on an Hilbert space is said to be **normal** if $T^*T = TT^*$*

In the setting of C^*-algebras we will define the same kind of normal element, see definition 72.19.

Proposition 60.36.
If T is a bounded normal operator then $T - \lambda\mathbb{1}$ is a bounded normal operator for every $\lambda \in \mathbb{C}$.

Proof. An immediate computation shows that $(T - \lambda\mathbb{1})(T^* - \bar{\lambda}\mathbb{1}) = (T^* - \bar{\lambda}\mathbb{1})(T - \lambda\mathbb{1})$, so $T - \lambda\mathbb{1}$ is normal. In order to see that $T - \lambda\mathbb{1}$ is bounded,

$$\|T - \lambda\mathbb{1}\| = \sup_{h\in\mathscr{H}} \frac{\|Th - \lambda h\|}{\|h\|} \tag{60.26a}$$

$$\leqslant \sup_{h\in\mathscr{H}} \frac{\|Th\| + |\lambda|\|h\|}{\|h\|} \tag{60.26b}$$

$$\leqslant \sup_{h\in\mathscr{H}} \frac{\|Th\|}{\|h\|} + |\lambda| \tag{60.26c}$$

$$= \|T\| + |\lambda|. \tag{60.26d}$$

The last line is bounded because T is bounded. □

PropoCartactNormal

Proposition 60.37.
An operator T is normal if and only if $\|Tx\| = \|T^*x\|$ for every $x \in \mathscr{H}$.

Proposition 60.38.
If T is diagonalizable by an unitary operator, then it is normal.

Proof. Let U be unitary and D be diagonal such that

$$UTU^* = D. \tag{60.27}$$

Then we have $(UTU^*)(UTU^*)^* = DD^*$ and thus

$$UTT^*U^* = DD^*. \tag{60.28}$$

In the same time,

$$DD^* = D^*D = UT^*TU^*, \tag{60.29}$$

so we have $UTT^*U^* = UT^*TU^*$ and then $TT^* = T^*T$ because U and U^* are invertible. □

Proposition 60.39.
If T is normal and bounded, the residual spectrum $\mathrm{Spec}_R(T)$ is empty.

Proof. See [937] at page 20. □

Proposition 60.40.
For a normal operator we have

$$\|T\| = r(T) \tag{60.30}$$

where $r(T)$ is the spectral radius of T.

Proof. A proof is available on wikipedia. □

Proposition 60.41.
If T is a normal operator and if λ is an eigenvalue of T, then $\bar{\lambda}$ is an eigenvalue of T^* for the same eigenvector.

Proof. If v is eigenvector of T for the eigenvalue λ, using proposition 60.37 on the normal operator $T - \lambda\mathbb{1}$ we have

$$0 = \|(T - \lambda\mathbb{1})v\| = \|(T - \lambda\mathbb{1})^* v\|. \tag{60.31}$$

Thus $(T^* - \bar{\lambda}\mathbb{1})v = 0$ and v is eigenvector of T^* for the eigenvalue $\bar{\lambda}$. □

60.3 Spectral theory on Banach algebras

Definition 60.42.
*An **involution** on an algebra $\mathfrak{A}$ is a $\mathbb{R}$-linear map $*\colon \mathfrak{A} \to \mathfrak{A}$ which fulfils*

$$A^{**} = A, \tag{60.32a}$$
$$(AB)^* = B^* A^* \tag{60.32b}$$
$$(\lambda A)^* = \bar{\lambda} A^* \tag{60.32c}$$

where λ is any complex number. An algebra endowed with an involution is a $*$-algebra.

Remarque 60.43.
In the setting of Hopf algebras, we do not require $A^{**} = A$. In that we follow [889]; see subsection 54.6.4.

Definition 60.44.
*The **operator norm** of a linear map $A\colon \mathcal{B} \to \mathcal{B}$ on a Banach space is*

$$\|A\| = \sup_{\|v\|=1} \|Av\| \in \mathbb{R}.$$

*The operator A is **bounded** if its norm is finite.* def:normappl

A classical result is

Proposition 60.45.
A linear operator on a Banach space is bounded if and only if it is continuous. prop:cont_born

Proposition 60.46.
The space $\mathfrak{B}(\mathcal{B})$ of bounded operators on the Banach space $\mathcal{B}$ endowed with the norm operator is a Banach space.

The norm of a functional is defined by

$$\|\rho\| := \sup\{|\rho(v)| : v \in \mathcal{B}, \|v\| = 1\}$$

It is the smallest c that can be used in the definition of the continuity.

We denote by $\mathcal{B}^*$ the dual of the Banach space $\mathcal{B}$. It is the set of all the functionals on $\mathcal{B}$ and it is itself a Banach space.

tho:hahnBanach

Theorem 60.47 (Hahn-Banach)**.**
Any functional defined on a linear subspace $\mathcal{B}_0$ of $\mathcal{B}$ can be extended to a functional of same norm defined on the whole $\mathcal{B}$. In particular, if $\rho(v) = 0$ for all $\rho \in \mathcal{B}^*$, then $v = 0$.

The set of self-adjoint operators of a C^*-algebra $\mathfrak{A}$ is written as

$$\mathfrak{A}_{\mathbb{R}} = \{A \in \mathfrak{A} \text{ tel que } A^* = A\}.$$

As notation, we denotes by $G(\mathfrak{A})$ the set of invertible elements of $\mathfrak{A}$:

$$G(\mathfrak{A}) := \{A \in \mathfrak{A} | A^{-1} \text{ exists}\}$$

Lemma 60.48.
A Banach $*$-algebra with $\|A\|^2 \leqslant \|A^*A\|$ for all A is a C^*-algebra.

lem:STARAlC

Proof. The definition is that a C^*-algebra is a Banach $*$-algebra such that $\|A^*A\| = \|A\|^2$, then we have to show that if A belongs to a Banach $*$-algebra, then $\|A\|^2 \leqslant \|A^*A\|$ implies $\|A\|^2 = \|A^*A\|$. In a Banach algebra, $\|A\|^2 \leqslant \|A^*A\| \leqslant \|A^*\|\|A\|$ so that $\|A\| \leqslant \|A^*\|$. The same with A^* instead of A gives the inverse inequality. Then $\|A\| = \|A^*\|$. □

Definition 60.49.
*An **unit** $\mathfrak{A}$ is an element $\mathbb{1}$ such that $\|\mathbb{1}\| = 1$ and $\mathbb{1}A = A\mathbb{1} = A$ for all $A \in \mathfrak{A}$. A Banach algebra which contains an unit is **unital**. When $z \in \mathbb{C}$, we often write z instead of $z\mathbb{1}$.*

Note that in a C^*-algebra, the definition of an unit don't impose the norm because definition of a C^*-algebra applied to $A = \mathbb{1}$ automatically gives $\|\mathbb{1}\| = 1$.

Let $\mathfrak{A}$ be a Banach algebra without unit. We define $\mathfrak{A}_{\mathbb{1}} := \mathfrak{A} \oplus \mathbb{C}$ with the notation $A + \lambda\mathbb{1}$ for $(A, \mathbb{1})$. We enforce an algebra structure on this set following the natural way:

$$(A + \lambda\mathbb{1})(B + \mu\mathbb{1}) := AB + \lambda B + \mu A + \lambda\mu\mathbb{1},$$

in other terms, $1 \in \mathbb{C}$ is assimilated to the $\mathbb{1}$ of $\mathfrak{A}_{\mathbb{1}}$. The norm on $\mathfrak{A}_{\mathbb{1}}$ is defined by

$$\|(A + \lambda\mathbb{1})(B + \mu\mathbb{1})\| \leqslant \|(A + \lambda\mathbb{1})\|\|(B + \mu\mathbb{1})|$$

So $\mathfrak{A}_{\mathbb{1}}$ becomes an unital Banach algebra. We have the following:

Proposition 60.50.
For each Banach algebra without unit, there exists an unital Banach algebra $\mathfrak{A}_{\mathbb{1}}$ and an isometric morphism $\mathfrak{A} \to \mathfrak{A}_{\mathbb{1}}$ such that $\frac{\mathfrak{A}_{\mathbb{1}}}{\mathfrak{A}} \simeq \mathbb{C}$.

Note that such an unitization of a Banach algebra is not unique see page 3745 and [489] page 16 and proposition 2.4.6.

60.4 Spectral theorem and some consequences

pg_spectralthe

60.4.1 Spectrum

It is well know that a norm on a set gives rise to a topology. A normed vector space which is complete in the norm topology a **Banach space**. A **Banach algebra** is a Banach space equipped with an algebra structure such that

$$\|AB\| \leqslant \|A\|B\| \tag{60.33}$$

eq:normBanach

for all A, B in the algebra.

def_banach

def:fonctionelle

Definition 60.51.
*A **functional** on a Banach space $\mathfrak{A}$ is a continuous linear map $\rho\colon \mathfrak{A} \to \mathbb{C}$.*

We say that ρ is **continuous** when there exists $c \in \mathbb{R}$ such that $\forall v \in \mathfrak{A}$, $|\rho(v)| \leqslant c\|v\|$. We recall that, for linear maps, continuity is equivalent to boundedness. In fact we have the following [939].

Proposition 60.52.
For a map $\rho\colon \mathscr{H}_1 \to \mathscr{H}_2$ between two Hilbert space, the following are equivalent:

(1) ρ is continuous at 0,

(2) ρ is continuous,

(3) ρ is bounded.

Definition 60.53.
*Let $\mathfrak{A}$ be an unital Banach algebra. The **resolvent** of an element $A \in \mathfrak{A}$ is*

$$\rho(A) = \{z \in \mathbb{C} \text{ tel que } (A - z\mathbb{1})^{-1} \text{exists as two-sided inverse}\}. \tag{60.34}$$

*The **spectrum** $\sigma(A)$, or $\mathrm{Spec}(A)$, is the complement of $\rho(A)$ in $\mathbb{C}$:*

$$\mathrm{Spec}(A) = \{\lambda \in \mathbb{C} \text{ tel que } (A - \lambda) \text{ is not invertible in the algebra}\}. \tag{60.35}$$

The spectrum of A is sometimes written $\sigma(A)$.

Notice that in an algebra of polynomials, the spectrum of an element is is almost always $\mathbb{C}$. The **spectral radius** of $A \in \mathfrak{A}$ is

$$r(A) = \sup_{z \in \mathrm{Spec}(A)} |z|. \tag{60.36}$$

def:spectre

In the finite dimensional case, the spectrum is equal to the set of eigenvalues.

Remarque 60.54.
When $\mathfrak{A}$ has no unit, the spectrum and the resolvent are defined by considering the unitization $\mathfrak{A}_{\mathbb{1}} := \mathfrak{A} \oplus \mathbb{C}$ instead of $\mathfrak{A}$.

60.4.2 Note about operator algebras

This subsection comes from wikipedia. Read it for more informations.

In this subsection we consider an algebra of operators on H. We say that λ is an **eigenvalue** of A if there exists $h \in H$ such that

$$Ah = \lambda h. \tag{60.37}$$

This implies that λ belongs to the spectrum of A. The contrary is in general not true. If for instance we consider the algebra of *bounded* operators on H, it can happen that $A - \lambda\mathbb{1}$ is invertible but with non bounded inverse. In this case, λ belongs to the spectrum (because $A - \lambda\mathbb{1}$ is not invertible *in the algebra*) but is not an eigenvalue.

The set of eigenvalues is called the **pure point spectrum**. This is in general a part of the spectrum.

An operator $A \in \mathfrak{B}(\mathscr{H})$ is **positive** if $\langle Av, v\rangle \geqslant 0$ for every $v \in \mathscr{H}$. Notice that A^*A is positive for every operator A because $\langle A^*Av, v\rangle = \langle Av, Av\rangle = \|Av\|^2 \geqslant 0$. We will see in theorem 72.31 that, in the case of unital C^*-algebra, all the positive elements are of that form.

60.4.3 Spectrum: next steps

Lemma 60.55.
The spectrum of a self-adjoint operator is a compact subset of $\mathbb{R}^+$.

lem:cv_Ak

Lemma 60.56 ([489]).
Let $\mathfrak{A}$ be an unital Banach algebra, and $A \in \mathfrak{A}$. Then the formula

$$\lim_{N\to\infty} \sum_{k=0}^{N} A^k = (\mathbb{1} - A)^{-1} \tag{60.38}$$

holds if $\|A\| < 1$,. As a consequence, $(A - z\mathbb{1})^{-1}$ exists when $|z| > \|A\|$.

Proof. Since $\mathfrak{A}$ is complete, it is sufficient for convergence to prove the fact that the sequence of partial sums is Cauchy. Suppose $n > m$ and compute:

$$\|\sum_{k=0}^{n} A^k - \sum_{k=0}^{m} A^k\| = \|\sum_{k=m+1}^{n} A^k\| \leqslant \sum_{k=m+1}^{n} \|A^k\| \leqslant \sum_{k=m+1}^{n} \|A\|^k.$$

The last inequality comes from the fact that $\|AB\| \leqslant \|A\|\|B\|$. From the theory of the geometric series, we know that the last sum converges to zero when $n, m \to \infty$ because $\|A\| < 1$. Then $\sum_{k=0}^{\infty} A^k \in \mathfrak{A}$. Remains to check that this is the inverse of $(\mathbb{1} - A)$:

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n}(A^k - A^{k+1}) = \mathbb{1} - A^{n+1},$$

so

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \leqslant \|A\|^{n+1}.$$

Since $\|A\| < 1$, the limit $n \to \infty$ of the right hand side is zero. The conclusion is that $\sum_{k=0}^{\infty} A^k$ is a left inverse of $(\mathbb{1} - A)$. The same shows that is is also a right inverse. This proves the first statement. The second is immediate by working with A/z instead of A.

□

Lemma 60.57.
The set of invertible elements

$$G(\mathfrak{A}) := \{A \in \mathfrak{A} : A^{-1} \text{ exists}\}$$

is open in $\mathfrak{A}$

lem:G_ouvert

Proof. Let us consider a $A \in G(\mathfrak{A})$ and $B \in \mathfrak{A}$ such that $\|B\| < \|A^{-1}\|^{-1}$. By definition 60.4.1, we have $\|A^{-1}B\| \leqslant \|A^{-1}\|\|B\| < 1$. Thus $A + B = A(\mathbb{1} + A^{-1}B)$ has an inverse because A and $(\mathbb{1} + A^{-1}B)$ have both an inverse ($\|A^{-1}B\| < 1$ and lemma 60.56).

Thus, when A has an inverse, the element $A + B$ also has a one when $\|B\|$ is not too big. In other words: any $C \in \mathfrak{A}$ such that $\|A - C\| < \epsilon$ is in $G(\mathfrak{A})$ when $\epsilon \leqslant \|A^{-1}\|^{-1}$. □

Now we can prove a great and fundamental theorem.

ThoSpecBanach

Theorem 60.58.
For any Banach algebra $\mathfrak{A}$ and any element $A \in \mathfrak{A}$, the spectrum $\sigma(A)$ is ItemThoSpecBanachi

(1) a subset of $\{z \in \mathbb{C} : |z| \leqslant \|A\|\}$,

(2) compact,

(3) non empty

tho:prop_sigma

Proof. If $\mathfrak{A}$ has no unit, we can add one and then one can suppose $\mathfrak{A}$ to be unital.

The first point is contained in lemma 60.56. Thus it is bounded and, in order to prove that it is compact, we only need to prove that it is closed.

We consider the map $f\colon \mathbb{C} \to \mathfrak{A}$, $f(z) = z\mathbb{1} - A$. Clearly, $\|f(z+\delta) - f(z)\| = \delta$, then f is continuous [3]. Since $G(\mathfrak{A})$ is open by the lemma 60.4.3, continuity makes $f^{-1}(G(\mathfrak{A}))$ open in $\mathbb{C}$. But this set is exactly the set of $z \in \mathbb{C}$ such that $z - A$ has an inverse: $\rho(A)$. The complement $\sigma(A)$ is thus closed and therefore compact.

Now we show that $\sigma(A)$ is non-empty. For this, we begin by defining $g\colon \rho(A) \to \mathfrak{A}$, $g(z) := (z - A)^{-1}$ (which is well defined by definition of $\rho(A)$). Now, pick a $z_0 \in \rho(A)$ and a $z \in \mathbb{C}$ such that

$$|z - z_0| < \|(A - z_0)^{-1}\|^{-1}.$$

Since $\rho(A)$ is open, we can choose $z \in \rho(A)$. Note that $(z_0 - A)^{-1}$ is a two-sided inverse because $z_0 \in \rho(A)$. Since $\|(z_0 - z)(z_0 - A)^{-1}\| = |z_0 - z|\|(z_0 - A)^{-1}\| < 1$, lemma 60.56 assures the convergence and makes sense to the following computation:

$$\begin{aligned} \frac{1}{z_0 - A}\sum_{k=0}^{\infty}\Big(\frac{z_0 - z}{z_0 - A}\Big)^k &= (z_0 - A)^{-1}\left([(z_0 - A) - (z_0 - z)](z_0 - A)^{-1}\right)^{-1} \\ &= (z - A)^{-1} \\ &= g(z). \end{aligned} \tag{60.39}$$

Then $g(z) = \sum_{k=0}^{\infty}(z_0 - z)^k(z_0 - A)^{k-1}$ is a norm-converging power series with respect to z. Now, assume $z \neq 0$. We have $g(z) = z^{-1}(\mathbb{1} - A/z)^{-1}$, then

$$\|g(z)\| = |z|^{-1}\|(\mathbb{1} - A/z)^{-1}\| \to 0 \tag{60.40}$$

eq:cv_de_g

when $z \to \infty$.

Let us now consider a functional $\rho \in \mathfrak{A}^*$. Recall that, by definition, it is bounded, and define $g_\rho\colon \mathbb{C} \to \mathbb{C}$, $g_\rho(z) := \rho(g(z))$. The limit (60.40) makes

$$\lim_{z\to\infty} |g_\rho(z)| \leqslant \lim_{z\to\infty} c\|g(z)\| = 0$$

where c is the constant of definition 60.51. Then

$$\lim_{z\to\infty} g_\rho(z) = 0. \tag{60.41}$$

eq:limite_g_rho

Because $\lim_{z\to\infty} |\rho_g(z)| \leqslant \lim_{z\to\infty} \|g(z)\| = 0$.

Suppose that $\sigma(A) = \varnothing$, or $\rho(A) = \mathbb{C}$; the function g is then defined on the whole $\mathbb{C}$. More precisely, g_ρ is an analytic complex function whose vanishes at infinity, then Liouville's theorem makes it constant [4]. Due to equation (60.41), for every functional ρ,

$$\rho\big(g(z)\big) = 0 \quad \forall z \in \mathbb{C}.$$

This yields $g(z) = 0$ for any z in $\mathbb{C}$, but it is not possible. Thus we are leads to $\rho(A) \neq \mathbb{C}$ and thus $\sigma(A) \neq \varnothing$. □

Theorem 60.59.
The spectrum of any element in a Banach algebra is

(1) non empty,

(2) compact.

Proof. Let $\mathfrak{A}$ be a Banach algebra and $A \in \mathfrak{A}$, and define $r_\lambda = (A - \lambda)^{-1}$. Formally, we have

$$r_\lambda = -\lambda^{-1}\left(1 - \frac{A}{\lambda}\right)^{-1} = -\lambda^{-1}\left(1 + \frac{A}{\lambda} + \frac{A^2}{\lambda^2} + \dots\right).$$

We have to study in which case does that expansion make sense, and check that it actually is an inverse of $(A - \lambda)$.

3. We consider the topology induced by the metric.
4. As far as I remember, holomorphic equals analytic.

When $|\lambda| > \|A\|$, that series is absolutely convergent thanks to the condition (60.33), and one can check that it is an inverse. We conclude that, when $|\lambda| > \|A\|$, the expression $(A-\lambda)$ is invertible.

Now, the algebraic identity $a^{-1}-b^{-1} = a^{-1}(b-a)b^{-1}$ reads $r_\lambda - r_\mu = r_\lambda(\lambda-\mu)r_\mu$, from which we deduce an expression for r_λ in terms of r_λ and r_μ:

$$r_\lambda = r_\mu + r_\lambda(\lambda-\mu)r_\mu.$$

If one substitutes that expression into itself, we find $r_\lambda = r_\mu + r_\mu(\lambda-\mu)r_\mu + r_\lambda(\lambda-\mu)r_\mu(\lambda-\mu)r_\mu$. Making the same again and again provides the following expansion:

$$r_\lambda = r_\mu + (\lambda-\mu)r_\mu^2 + (\lambda-\mu)^2 r_\mu^3 + \dots \tag{60.42}$$ EqDevrllambu

So if r_μ exists (i.e. $\mu \notin \mathrm{Spec}(A)$), we want to define r_λ by that formula. Now, if $|\lambda-\mu|\|r_\mu\| < 1$, then $\lambda \notin \mathrm{Spec}(A)$ because formula (60.42) actually works. That proves that the set $\mathbb{C}\setminus\mathrm{Spec}(A)$ is open, and then that $\mathrm{Spec}(A)$ is closed. Since it is bounded too, we conclude that, being a part of $\mathbb{C}^2$, the set $\mathrm{Spec}(A)$ is compact.

Let us now prove that the map $\lambda \mapsto r_\lambda$ is continuous. Indeed, consider $r_\lambda - r_\mu j$, and compute the difference using the series (60.42). One sees that it converges to 0 when $\lambda \to \mu$. We know that $(A-\lambda)$ is invertible when $|\lambda| > \|A\|$, so that it makes sense to compute the limit of $\|r_\lambda\|$ when λ goes to infinity. The limit of $\|r_\lambda\|$ when $\lambda \to 0$ is 0. The map $\lambda \mapsto r_\lambda$ is a differentiable map over $\mathbb{C}$ and must then be constant, so that it must be zero everywhere, because its limit is zero.

We deduce that $\mathrm{Spec}(A) \neq \varnothing$.

¡¡ Avertissement/question au lecteur !! 60.60
Some questions to be elucidated:

- *In order to prove that $\|r_\lambda\| \to 0$ when $\lambda \to \infty$, one has to say that the norm of r_λ is the invert of the one of $(A-\lambda)$, and that the latter goes to infinity when λ goes to infinity.*
- *The lase deduction is not clear at all, but it is done in greater details in theorem 60.58. To be merged.*

□

Corollary 60.61 (Theorem of Gelfand-Mazur)**.**
If all elements (except zero) of an unital Banach algebra $\mathfrak{A}$ are invertible, thus $\mathfrak{A} \simeq \mathbb{C}$ as Banach algebras. cor:GelfandMazur

Proof. We just saw that $\sigma(A) \neq \varnothing$, thus for all non-zero element $A \in \mathfrak{A}$, there exists a $z_A \in \mathbb{C}$ such that $A - z_A\mathbb{1}$ is not invertible. From assumption, $A - z_A\mathbb{1} = 0$, so that $A \to z_A$ is the expected isomorphism. Since $\|A\| = \|z_A\mathbb{1}\| = |z_A|$, this isomorphism is isometric. □

As a corollary, we have that

$$r(A) \leqslant \|A\|. \tag{60.43}$$

The following version of the spectral theorem is known as the *continuous functional calculus.*

Theorem 60.62 (First version of the spectral theorem)**.** ThoSpectralTho
Let $\mathscr{H}$ be an Hilbert space, and $T \in \mathfrak{B}(\mathscr{H})$. If T is self-adjoint, there is an algebra isomorphism

$$C^*(T,\mathbb{1}) \xrightarrow{\simeq} C\big(\mathrm{Spec}(T)\big)$$

which maps T to the identity function. That isomorphism is unique and continuous from the norm topology to the topology of uniform convergence. If we denote by $f(T)$ the image of $f \in C\big(\mathrm{Spec}(T)\big)$ by the isomorphism, we have the following properties

(1) $f(1) = \mathrm{Id}$,

(2) $f(T)^* = \overline{f}(T)$,

(3) $\|f(T)\| = \|f\| = \sup\{|f(x)|,\ x \in \mathrm{Spec}(T)\}$,
ItemSpecffSpecThoSpectral

(4) $\mathrm{Spec}\big(f(T)\big) = f\big(\mathrm{Spec}(T)\big)$,

(5) $f(T) \geqslant 0$ *if and only if* $f \geqslant 0$.

The map $f \mapsto f(T)$ is the **continuous functional calculus**.

Proof. We recall that, by theorem 60.58, the set $\mathrm{Spec}(T)$ is non empty and compact.

For the second claim, first suppose that $\dim \mathscr{H}$ is finite. Then there exists an orthonormal basis $\{v_1, \dots, v_n\}$ such that $Tv_j = \lambda_j v_j$, and $\mathrm{Spec}(T) = \{\lambda_1, \dots, \lambda_n\}$. Define $f(T)v_j = f(\lambda_j)v_j$. Properties of that map are easy to prove.

In the infinite dimensional case, the proof is again to build an orthonormal basis of eigenvectors of T. We refer to [935, 489] for a proof. □

Corollary 60.63.
Every positive operator has an unique positive square root.

Proof. Let $A \geqslant 0$ and $B \in C^*(A, \mathbb{1})$ be the element associated with the square root function on $C\big(\mathrm{Spec}(A)\big)$ by the spectral theorem. Then $B^2 = A$ because the correspondence is a morphism and A correspond to the identity. The fact that B is positive comes from the fact that $B = \big(\sqrt[4]{A}\big)^2$. □

Taking the square root of the positive operator T^*T, we define the **absolute value** of the operator T by

$$|T| = \sqrt{T^*T}. \tag{60.44}$$
EqAbsValT

LemkerTkersqrtT

Lemma 60.64.
We have

$$\ker(T) = \ker\big(\sqrt{T^*T}\big),$$

for every $T \in \mathfrak{B}(\mathscr{H})$.

Proof. Taking the adjoint term by term in the development of $f(t) = \sqrt{t}$, we see that

$$\Big((T^*T)^{1/2}\Big)^* = (T^*T)^{1/2},$$

then a vector $v \in \ker(T)$ satisfies

$$0 = \langle Tv, Tv\rangle = \langle v, T^*Tv\rangle = \langle v, \sqrt{T^*T}\sqrt{T^*T}v\rangle = \langle \sqrt{T^*T}v, \sqrt{T^*T}v\rangle,$$

which means that $v \in \ker\big(\sqrt{T^*T}\big)$. □

Lemma 60.65.
Let p *be a polynomial on* $\mathbb{C}$*, if we define*

$$p(\sigma(A)) := \{p(z) : z \in \sigma(1)\},$$

then we have $p(\sigma(A)) = \sigma(p(A))$.
lem:sigma_poly

Proof. Let us consider $z, \alpha \in \mathbb{C}$ and write the factorisations:

$$p(z) - \alpha = c\prod_{i=1}^{n}(z - \beta_i(\alpha)) \tag{60.45a}$$
eq:prod_1

$$p(A) - \alpha\mathbb{1} = c\prod_{i=1}^{n}(A - \beta_i(\alpha)\mathbb{1}), \tag{60.45b}$$
eq:prod_2

where –of course– c and β_i are determined by p and α.

Now, particularise to $\alpha \in \rho(p(A))$: $p(A) - \alpha\mathbb{1}$ is invertible and thus each of $A - \beta_i(\alpha)\mathbb{1}$ is too. In other words (taking the complement) $\alpha \in \sigma(p(A))$ implies that at least one out of $A - \beta_i(\alpha)\mathbb{1}$ is

not invertible: $\beta_i(\alpha) \in \sigma(A)$ for one of the i. On the other hand, by definition $p(\beta_i) - \alpha = 0$, then $\alpha \in p(\sigma(A))$. All in all, we have shown that $\sigma(p(A)) \subset p(\sigma(A))$.

For the inverse inclusion, take $\alpha \in p(\sigma(A))$, i.e. $\alpha = p(z)$ for some $z \in \sigma(A)$. In this case the product (60.45a) is zero and thus z is one of the $\beta_i(z)$. For this i, $\beta_i(\alpha) \in \sigma(A)$, then $A - \beta_i(\alpha)$ is non-invertible. By (60.45b), $p(A) - \alpha\mathbb{1}$ is also non-invertible, and thus $\alpha \in \sigma(p(A))$.

□

LemPolarHilbert

Lemma 60.66 (Polar decomposition)**.**
Every operator $A \in \mathfrak{B}(\mathscr{H})$ has a ***polar decomposition*** *$A = U|A|$ where $|A| = \sqrt{A^*A}$ and U is a partial isometry with the same kernel as A.*

See [489] for a proof. A version of this decomposition in von Neumann algebras is given in proposition 74.10. The operator $\sqrt{A^*A}$ is defined by means of the spectral theorem 60.62, see equation (60.44).

The partial isometry U is sometimes called the **sign** of A because when it is selfadjoint, it is the operator associated with the sign function by the continuous functional calculus. Indeed, let

$$\varphi_A \colon C\big(\operatorname{Spec}(A)\big) \to C^*(A, \mathbb{1}) \tag{60.46}$$

be the isomorphism. We denote by A' the restriction of A to the space $\ker(A)^\perp$ and Sign the sign function. This is a continuous function on $\operatorname{Spec}(A')$, thus we can speak about $\varphi_{A'}(\operatorname{Sign})$. The identity on $\operatorname{Spec}(A') \subset \mathbb{R}$ can be expressed as

$$\operatorname{Id}(x) = \operatorname{Sign}(x)|x| = \operatorname{Sign}(x)\sqrt{x^*x}, \tag{60.47}$$

so that we have $\varphi_{A'}(\operatorname{Id}) = \varphi_{A'}(\operatorname{Sign}) \circ \sqrt{(A')^*A'}$. Thus we have the decomposition

$$A' = \operatorname{Sign}(A') \circ |A'|. \tag{60.48}$$

EqPolarSSKerSign

The operator $\operatorname{Sign}(A')$ is a partial isometry of $\ker(A)^\perp$ because

$$\operatorname{Sign}(A')\operatorname{Sign}(A')^* = \varphi_{A'}(\operatorname{Sign})\varphi_{A'}(\operatorname{Sign}) = \varphi_{A'}(1) = \operatorname{Id}. \tag{60.49}$$

prop:cv_lim_sup

Proposition 60.67.
If (a_k) is a sequence in $\mathbb{R}$ such that there exists a $a \in \mathbb{R}$ for which for any $k \in \mathbb{N}$,

$$\limsup_{n\to\infty} a_n \leqslant a \leqslant a_k$$

then (a_k) admits a limit and $\lim_{n\to\infty} a_k = a$.

Proposition 60.68.
Let $\mathfrak{A}$ be an unital Banach algebra, the spectral radius is given by the formula

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}. \tag{60.50}$$

prop:An_usn

for every $A \in \mathfrak{A}$.

Proof. Lemma 60.56 says that when $|z| > \|A\|$, the operator $(A - z\mathfrak{A})^{-1}$ exists. Let us once again consider the function g. From the lemma,

$$\frac{1}{z}\sum_{k=0}^{\infty}\left(\frac{A}{z}\right)^k = \frac{1}{z}\left(\mathbb{1} - \frac{A}{z}\right)^{-1} = (z - A)^{-1} = g(z) \tag{60.51}$$

On the other hand, for any $z \in \rho(A)$, one can find an element $z_0 \in \rho(A)$ such that

$$g(z) = \sum_{k=0}^{\infty}(z_0 - z)^k(z_0 - A)^{k-1}$$

converges. If $|z| > r(A)$, then $z \in \rho(A)$ and then this series converges. But in the interior of the convergence disk, a power series converges uniformly. Then

$$g(z) = \frac{1}{z} \sum_{k=0}^{\infty} \left(\frac{A}{z} \right)^k$$

uniformly converges with respect to z when $|z| > r(A)$.

On the other hand, classical analysis makes this series norm-convergent only if $\|A\|^n / \|z\|^n$ from a certain large n. Since $\|A^n\| \leqslant \|A\|^n$, one can say:

$\forall |z| > r(A)$, $\exists N$ such that $n > N$ implies

$$\frac{\|A^n\|}{|z|^n} < 1.$$

Of course, the choice of N relies on z. A consequence of this:

$$\limsup_{n\to\infty} \frac{\|A^n\|}{|z|^n} < 1$$

Replacing, $<$ by $\leqslant$, one can replace $|z|$ by $r(A)$; this yields

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leqslant r(A). \tag{60.52}$$

Now, we show that for any n, $r(A) \leqslant \|A^n\|^{1/n}$, so that proposition 60.67 concludes our proof.

Since $\sigma(A)$ is closed (theorem 60.4.3), there exists an element $\alpha \in \sigma(A)$ such that $|\alpha| = r(A)$. Lemma 60.65 makes $\alpha^n \in \sigma(A^n)$, thus $|\alpha^n| \leqslant \|A^n\|$ and finally

$$r(A) = \alpha \leqslant \|A^n\|^{1/n}. \tag{60.53}$$

Now we are in the situation of a real sequence (a_k) such that $\limsup_{n\to\infty} a_n \leqslant a \leqslant a_k$ for all a_k [5]. Let us show that in this situation,

$$\limsup_{n\to\infty} a_n = a. \tag{60.54}$$

eq_limsupnana

First suppose that $\limsup_{n\to\infty} a_k = b < a$. In this case, $\forall \epsilon > 0$, there exists a N such that $|\sup\{a_k \text{ tel que } k \geqslant N\} - b| < \epsilon$; in other words, $\sup\{a_k \text{ tel que } k \geqslant N\} \in B(b, \epsilon)$. The for all $\delta > 0$, there are some a_k with $B(b, \epsilon + \delta)$. If we choose ϵ and δ suitably small, it gives some $a_k < a$. We conclude that

$$\lim_{l\to\infty} \{a_k \text{ tel que } k \geqslant l\} = a \leqslant a_n$$

for all n. Let $\epsilon > 0$ and assume that there exists a $n > N$ with a_n outside $B(a, \epsilon)$, i.e. suppose that $\lim_{n\to\infty} a_n \neq a$ or doesn't exist. So for all l, $\sup\{a_k \text{ tel que } k \geqslant l\} > a + \epsilon$. This is a contradiction with equation (60.54).

Now we have to prove that $r(A) \leqslant \|A^n\|^{1/n}$ for all n. Since $\sigma(A)$ is closed, there exists a $\alpha \in \sigma(A)$ such that $|\alpha| = r(A)$ because the supremum of a bounded closed set is reached [6]. From continuous calculus, $\alpha^n \in \sigma(A^n)$ and therefore $|\alpha^n| \leqslant \|A^n\|$ because $r(1) \leqslant \|A^n\|$. We conclude that

$$r(A) = \alpha \leqslant \|A^n\|^{1/n}.$$

□

5. We recall the definition of supremum limit:

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k \text{ tel que } k \geqslant n\}.$$

6. Anglais ?

60.5 Operators with compact resolvent

The following come from [940] and a remark after the definition of a K-cycle in [941]. Since the Dirac operator of a spectral triple has compact resolvent, we need to know some theory about operators with compact resolvent.

LemResComKerFin

Lemma 60.69.
Let T be an operator with compact resolvent and $\lambda \in \mathbb{C}\backslash \mathrm{Spec}(T)$. Then the kernel of T is an eigenspace of $(T-\lambda\mathbb{1})^{-1}$ for the eigenvalue λ.

Proof. First notice that $\lambda \neq 0$; if not, the kernel would be empty because we choose λ among the values that are *not* eigenvalues of T. If $x \in \ker(T)$, then $x = -\frac{1}{\lambda}(T-\lambda\mathbb{1})x$. Applying $(T-\lambda\mathbb{1})^{-1}$ on both sides,

$$(T-\lambda\mathbb{1})^{-1}x = -\frac{1}{\lambda}x, \tag{60.55}$$

as claimed. □

Since the eigenspaces of a compact operator are finite dimensional, we have in particular

CorRezcomkerfin

Corollary 60.70.
If T is an operator with compact resolvent, then the kernel of T is finite dimensional.

LemResLcmpResLLcmp

Lemma 60.71.
Let $D\colon \mathscr{H} \to \mathscr{H}$ be an operator and consider $R_\lambda = (D-\lambda\mathbb{1})^{-1}$ for each $\lambda \in \mathbb{C}\backslash \mathrm{Spec}(D)$ (i.e. λ is in the resolvent of D). Then

(1) If $\lambda_1, \lambda_2 \in \mathbb{C}\backslash \mathrm{Spec}(D)$, we have

$$R_{\lambda_1} - R_{\lambda_2} = (\lambda_1 - \lambda_2)R_{\lambda_1}R_{\lambda_2}; \tag{60.56}$$

(2) The operator R_{λ_1} is compact if and only if R_{λ_2} is compact. In other words, all the resolvent are compact if only one is compact.

Proof. We have

$$\lambda_1 - \lambda_2 = (D-\lambda_2) - (D-\lambda_1) = (D-\lambda_1)\Big((D-\lambda_1)^{-1} - (D-\lambda_2)^{-1}\Big)(D-\lambda_2), \tag{60.57}$$

so that

$$(D-\lambda_1)^{-1} - (D-\lambda_2)^{-1} = (\lambda_2-\lambda_1)(D-\lambda_1)^{-1}(D-\lambda_2)^{-1}. \tag{60.58}$$

That proves the first claim. In order to prove the second claim, suppose that R_{λ_1} is compact and write

$$R_{\lambda_1} = \big((\lambda_1-\lambda_2)R_{\lambda_1} + \mathbb{1}\big)R_{\lambda_2}. \tag{60.59}$$

We are thus in the case of lemma 60.21 which makes R_{λ_2} compact. □

Proposition 60.72.
Let D be a selfadjoint operator on the Hilbert space $\mathscr{H}$. The operator $(D^2-\mathbb{1})^{-1}$ is compact on $\mathscr{H}$ if and only if $(D-\lambda\mathbb{1})^{-1}$ is compact for some $\lambda \notin \mathrm{Spec}(D)$.

Proof. The operator $(D^2+\mathbb{1})^{-1}$ can be written under the form

$$(D^2+\mathbb{1})^{-1} = \Big((D+i\mathbb{1})(D-i\mathbb{1})\Big)^{-1} = (D-i\mathbb{1})^{-1}(D+i\mathbb{1})^{-1}. \tag{60.60}$$

EqDecDdeuxplusuncmp

Since D is selfadjoint, the latter expression is of the form AA^*. If $(D^2+\mathbb{1})^{-1}$ is compact, this shows that $(D\pm i\mathbb{1})^{-1}$ are compacts by lemma 60.16 while the values $\pm i$ are not in the spectra of D which is real.

If $(D-\lambda\mathbb{1})^{-1}$ is compact for some value $\lambda \notin \mathrm{Spec}(D)$, lemma 60.71 shows that $(D\pm i\mathbb{1})^{-1}$ are compacts because $\pm i$ are outside the spectrum of D. Now the decomposition 60.60 shows that $(D^2+\mathbb{1})$ is compact. □

60.6 Strong, weak and other topologies

We are dealing with separable Hilbert spaces. More details in [942]. The **weak topology** on $\mathfrak{B}(\mathscr{H})$ is the one associated with the following convergence of nets. One says that $T_\alpha \to T$ if and only if for every v and w in $\mathscr{H}$, we have

$$\langle v, T_\alpha w\rangle \to \langle v, Tw\rangle. \tag{60.61}$$

One of the main feature of the weak topology is the **Banach-Alaoglu** theorem.

ThoBanachAlaoglu

Theorem 60.73 (Banach-Alaoglu)**.**
If $\mathscr{H}$ is a Hilbert space, then the unit ball in $\mathfrak{B}(\mathscr{H})$ is weakly compact.

The **strong topology** on $\mathfrak{B}(\mathscr{H})$ is the topology generated by the open sets

$$\mathcal{U}(S, v, \epsilon) = \{T \in \mathfrak{B}(\mathscr{H}) \text{ tel que } \|Tv - Sv\| \leqslant \epsilon\} \tag{60.62}$$

for all $S \in \mathfrak{B}(\mathscr{H})$, $v \in \mathscr{H}$ and $\epsilon > 0$. The associated convergence notion is the one of the pointwise convergence: $T_\alpha \to T$ if and only if

EqDEfimforte

$$\|T_\alpha v - Tv\| \to 0 \tag{60.63}$$

for all $v \in \mathscr{H}$.

The strong topology has more closed and open sets than the weak one. One difference between weak and strong topology is that the weak one is compatible with the involution while the strong is not. More precisely, [

PgStarWeakRespecte

]

$$T_\alpha \xrightarrow{w} T \Rightarrow (T_\alpha)^* \xrightarrow{w} T^* \tag{60.64}$$

while the same is not true for convergence in the strong topology.

60.6.1 Ultraweak topology

subSecUltraWtopol

Definition 60.74.
*The **ultraweak** is the weakest topology (the one with the fewest open sets) on $\mathfrak{B}(\mathscr{H})$ such that the functionals*

$$T \mapsto \sum_{n=1}^{\infty} \langle v_n T, Tw_n\rangle \tag{60.65}$$

are continuous for every choice of sequences (v_n) and (w_n) such that $\sum_n \|v_n\| < \infty$ and $\sum_n \|w_n\| < \infty$.

Proposition 60.75.
Every ultraweakly continuous linear functional on $\mathfrak{B}(\mathscr{H})$ has the form

$$T \mapsto \sum \langle v_n, Tv_n\rangle \tag{60.66}$$

where (v_n) is any sequence of vectors such that $\sum_n \|v_n\| < \infty$.

Proof. No proof. □

An operator $T \in \mathfrak{B}(\mathscr{H})$ is **Hilbert-Schmidt class** if

$$\|T\|_{HS}^2 := \sum_{n,m}^{\infty} |\langle v_n, Tv_m\rangle|^2 < \infty \tag{60.67}$$

for some orthonormal basis $\{v_n\}$ of $\mathscr{H}$. The operator T is **trace class** if

$$\|T\|_1 := \sum_n \langle |T| v_n, v_n\rangle < \infty \tag{60.68}$$

for some orthonormal basis $\{v_n\}$ of $\mathscr{H}$. Notice that, from a simple change of basis formula, the fact to put these conditions for *some* basis of for *every* basis are equivalent. When T is a trace class operator, we define its **trace** by

$$\mathrm{Tr}(T) = \sum_n \langle Tv_n, v_n\rangle \tag{60.69}$$

where $\{v_n\}$ is an orthonormal basis of $\mathscr{H}$. That definition is independent of the choice. We denote by $\mathscr{L}^1(\mathscr{H})$ the set of trace class operators in $\mathscr{H}$.

Proposition 60.76.
One has

$$\|ST\|_1 \leqslant \|S\|_{\mathfrak{B}(\mathscr{H})}\|T\|_1 \tag{60.70}$$

whenever it makes sense. Here, $\|.\|_{\mathfrak{B}(\mathscr{H})}$ denotes the operator norm over $\mathfrak{B}(\mathscr{H})$ and $\|.\|_1$ denotes the trace.

Proof. No proof. □

One consequence of that proposition is that the map

$$S \mapsto \mathrm{Tr}(ST) \tag{60.71}$$

is a bounded linear functional on $\mathfrak{B}(\mathscr{H})$.

Proposition 60.77.
Every trace class operator is a compact linear operator.

Sketch of the proof. First, we know that T is compact if and only if $|T|$ is compact. From spectral theory, the largest spectral value is the largest element in the sum $\sum_n \langle |T|v_n, v_n\rangle$, and the second largest spectral value is the second largest element of that sum and so on. Thus the convergence of the sum implies that $|T|$ is compact. □

Thus every positive trace class operator read

$$Tv = \sum_n \lambda_n \langle v_n, v\rangle v_n, \tag{60.72}$$

where the numbers λ_n are positive reals, while a general trace class operator, being a positive one followed by a partial isometry, read

$$Tv = \sum_n \lambda_n \langle v_n, v\rangle w_n \tag{60.73}$$

EqFormTraceClassGene

where $\{v_n\}$ and $\{w_n\}$ are orthonormal basis of $\mathscr{H}$ related by the partial isometry implied in the decomposition of T. When T is the trace class operator (60.73), the trace reads

$$\mathrm{Tr}(ST) = \sum_n \lambda_n \langle v_n, Sw_n\rangle. \tag{60.74}$$

Since the λ_n are eigenvalues of a compact operator, we have $\sum_n \lambda_n < \infty$, so that one can redefine $v_n \to \sqrt{\lambda_n} v_n$ and $W_n \to \sqrt{\lambda_n} w_n$ that are now summable sequences instead of being actual orthonormal basis. Thus one has

$$\mathrm{Tr}(ST) = \sum_n \langle v_n, Sw_n\rangle, \tag{60.75}$$

which proves that $\mathrm{Tr}(ST)$ is ultraweakly continuous as function of T. In turn, that implies that the space of ultraweakly continuous linear functions identifies with the space $\mathscr{L}(\mathscr{H})$ of trace class operators by the formula

$$\varphi_S(T) = \mathrm{Tr}(ST) \tag{60.76}$$

for $S \in \mathscr{L}^1(\mathscr{H})$. Notice that general theory of trace class operators assures that $\mathrm{Tr}(ST)$ makes sense when $S \in \mathscr{L}^1(\mathscr{H})$, and we have moreover $\|\varphi_S\| = \|S\|_1$. There is also the map

$$\begin{aligned} \mathfrak{B}(\mathscr{H}) &\to \mathscr{L}^1(\mathscr{H}) \\ T &\mapsto \varphi^T \end{aligned} \tag{60.77}$$

where $\varphi^T(S) = \varphi(ST)$.

In the same way as $l^1(\mathbb{N})^* \simeq l^\infty(\mathbb{N})$, we have the following.

Theorem 60.78.
The map $T \mapsto \varphi^T$ is an isometric isomorphism between $\mathfrak{B}(\mathscr{H})$ and $\mathscr{L}^1$.

Proof. No proof. □

Notice that if $x \in l^\infty(\mathbb{N})$, the functions $\varphi_i(x) = x_i$ provide an inclusion of $l^1(\mathbb{N})$ in $l^\infty(\mathbb{N})^*$, the latter space being in fact much bigger.

Theorem 60.79 (Hahn-Banach)**.**
If $M \subseteq \mathfrak{B}(\mathscr{H})$ is a closed subspace in the ultraweak topology, then

$$M \simeq (M_*)* \tag{60.78}$$

where M_* is the space of ultraweakly continuous linear functionals on M. Moreover the weak-$*$ topology on M is equivalent to th ultraweak one.

60.7 Polar cone

Let $\mathscr{H}$ be an Hilbert space and A a part of $\mathscr{H}$. A part $\mathscr{H}^+$ is **self polar** if

$$\mathscr{H}^+ = \{\xi \in \mathscr{H} \text{ tel que } \langle \xi, \eta \rangle \geqslant 0 \forall \eta \in \mathscr{H}^+\}. \tag{60.79}$$

The notation is motivated by the fact mentioned on page 51.12.3 that a cone implies a notion of positivity.

Proposition 60.80.
Let $\mathscr{H}^+$ be a self polar cone in the Hilbert space $\mathscr{H}$. Define

$$\mathscr{H}^J = \{\xi \in \mathscr{H} \text{ tel que } \langle \xi, \eta \rangle \in \mathbb{R} \forall \eta \in \mathscr{H}^+\}. \tag{60.80}$$

Then

1. *We have $\mathscr{H} = \mathscr{H}^J \oplus i\mathscr{H}^J$.*
2. *If we define $J : \mathscr{H} \to \mathscr{H}$, $J(\xi + i\eta) = \xi - i\eta$ for $\xi, \eta \in \mathscr{H}^J$, then J is an anti-unitary involution and $\mathscr{H}^J$ is an ordered vector space.* ItemPropDecompJordanH
3. *An J-real element $\xi \in \mathscr{H}^J$ can be written under the form $\xi = \xi_+ - \xi_-$ where ξ_+ and ξ_- are positive orthogonal elements of $\mathscr{H}^+$.*

The point (3) is called **Jordan decomposition**.

60.8 Representations of Lie groups on Hilbert spaces

Let G be a Lie group. A **linear representation** of G is a pair $(\mathscr{H}, U)$ where $\mathscr{H}$ is an Hilbert space and U is a map $U : G \to \mathrm{GL}(\mathscr{H})$ such that

$$\begin{aligned} U(gg') &= U(g) \circ U(g') \\ U(e) &= \mathrm{Id} \end{aligned} \tag{60.81}$$

Definition 60.81.
*If $\mathscr{H}$ is an Hilbert space, the representation is **unitary** if*

(1) for every element $g \in G$ the operator $U(g)$ is unitary:

$$\langle U(g)\phi, U(g)\psi\rangle = \langle \phi, \psi\rangle \tag{60.82}$$

for every $\phi, \psi \in \mathscr{H}$.

(2) for every $\psi \in \mathscr{H}$, the map

$$\begin{aligned} G &\to \mathscr{H} \\ g &\mapsto U(g)\psi \end{aligned} \tag{60.83}$$

is continuous.

A famous example is the **regular left representation**. The vector space is $C^\infty(G, \mathbb{C})$ and the representation is given by

$$\begin{aligned} L\colon G &\to \mathrm{End}_{\mathbb{C}}\left(C^\infty(G, \mathbb{C})\right) \\ L_g(u) &= L^*_{g^{-1}}u \end{aligned} \tag{60.84}$$

where $L^*_{g^{-1}}u$ is the function defined by

$$\left(L^*_{g^{-1}}u\right)(x) = u(g^{-1}x) \tag{60.85}$$

for $x \in G$.

That is an infinite dimensional representation.

60.8.1 Induced representation

SubSecInducrepresBBGC

Let B be a subgroup of G and let (V, χ) be a representation of B. We are going to build a representation of G from the data of B and χ. First we consider the space of G-**equivariant** functions in G:

$$C^\infty(G, \mathbb{C})^B = \{u \in C^\infty(G, \mathbb{C}) \text{ tel que } u(gb) = \chi(b^{-1})u(g)\}. \tag{60.86}$$

This space is invariant under the regular left representations of G because if $u \in C^\infty(G, \mathbb{C})^B$ and $g \in G$,

$$\begin{aligned} (L_g u)(xb) &= u(g^{-1}xb) \\ &= \chi(b^{-1})u(g^{-1}x) \\ &= \chi(b^{-1})(L_g u)(x). \end{aligned} \tag{60.87}$$

The restriction of the representation L to $C^\infty(G, \mathbb{C})^B$ is the **induced representation**.

Let us now suppose that

(1) the subgroup B is closed

(2) the representation χ is one dimensional.

Let now consider a Lie group G and a closed subgroup B which admits an unitary representation

$$\begin{aligned} \chi\colon B &\to \mathrm{U}(1) \\ b &\mapsto \chi(b). \end{aligned} \tag{60.88}$$

We look at the quotient

$$Q = G/B \tag{60.89}$$

with the natural action

$$\begin{aligned} \tau\colon G \times Q &\to Q \\ (g, [x]) &\mapsto [gx] = \tau_g[x]. \end{aligned} \tag{60.90}$$

In general, $\varphi(\tau_g q)$ is not equal to $g\varphi(q)$, we denote by β the difference:

$$\varphi(\tau_g q) = g\varphi(q)\beta(q, g). \tag{60.91}$$

EqDefBetavptaunotEqual

Lemma 60.82.
The function β does not depend on q: $\beta(g,q) = \beta(g,q')$.

Proof. No proof. □

We suppose that Q has a volume form $dq \in \Omega^{\dim(Q)}(Q)$ which is G-invariant, i.e.

$$\tau_g^*(dq) = dq \tag{60.92}$$

for every $g \in G$.

We also suppose that there exists a global section $\varphi\colon Q \to G$ which provides the global diffeomorphism

$$\begin{aligned} Q \times B &\to G \\ (q,b) &\mapsto \varphi(q)b. \end{aligned} \tag{60.93}$$

In that case we have a linear isomorphism

$$\begin{aligned} C^\infty(G,\mathbb{C})^B &\simeq C^\infty(Q,\mathbb{C}) \\ \hat{f} &\mapsto \tilde{f} \end{aligned} \tag{60.94}$$

EqIsomTildeHat

where $\hat{f}$ is defined from $\tilde{f}$ by the formula

$$\hat{f}\big(\varphi(q)b\big) = \chi(b^{-1})\tilde{f}(q). \tag{60.95}$$

EqIsomtildehat

This definition makes sense because $\varphi(q)b$ is a general element of G. We have to prove that the function $\hat{f}$ is equivariant. For that consider $x = \varphi(g)b \in G$ and $b_0 \in B$ and compute

$$\begin{aligned} \hat{f}\big(\varphi(q)bb_0\big) &= \chi(bb_0)^{-1}\tilde{f}(q) \\ &= \chi(b_0^{-1})\chi(b^{-1})\tilde{f}(q) \\ &= \chi(b_0^{-1})\tilde{f}(x). \end{aligned} \tag{60.96}$$

The map $\tilde{f} \mapsto \hat{f}$ given by (60.95) is moreover surjective. Indeed, if $\hat{f}$ is any element of $C^\infty(G,\mathbb{C})^B$, it satisfy

$$\hat{f}\big(\varphi(q)b\big) = \chi(b^{-1})\tilde{f}(q), \tag{60.97}$$

So that $\hat{f}$ is the image of a function $\tilde{f}$.

We can look at the compact supported functions $\mathscr{D}(Q) = C_c^\infty(Q,\mathbb{C})$ under the isomorphism (60.94). First, we want to know how does the regular representation goes from $C^\infty(G,\mathbb{C})^B$ to $C^\infty(Q,\mathbb{C})$. Let $\tilde{u} \in \mathscr{D}(Q)$ and look at the equivariant function $\hat{u}$ (formula (60.95)). We act with $g \in G$, and we use the function β (equation (60.91)):

$$\begin{aligned} \big(U(g)\hat{u}\big)\big(\varphi(q)\big) &= \hat{u}\big(g^{-1}\varphi(q)\big) \\ &= \hat{u}\big(\varphi(g^{-1}\cdot q)\beta(q,g^{-1})^{-1}\big) \\ &= (\chi\circ\beta)(q,g^{-1})\tilde{u}(g^{-1}\cdot q) \\ &= \underbrace{(\chi\circ\beta)(q,g^{-1})}_{\in \mathrm{U}(1)}\big(\tau_{g^{-1}}^*\tilde{u}\big)(q). \end{aligned} \tag{60.98}$$

Now, since $\tau_{g^{-1}}$ is a diffeomorphism, the function $\tau_{g^{-1}}^*\tilde{u}$ belongs to $\mathscr{D}(Q)$. The fact to multiply by $\chi\big(\beta(q,g^{-1})\big)$ does not changes the compactness of the support.

We define the representation $\tilde{U}$ as

$$\tilde{U}(g)\tilde{u} = \chi\big(\beta(g^{-1})\big)\tau_{g^{-1}}^*\tilde{u}. \tag{60.99}$$

60.8.2 Unitary induced representation

SubSecUnitInducedPrep

Let $\tilde{u}, \tilde{v} \in \mathscr{D}(Q)$, we consider the product in $L^2(Q, dq)$:

$$\langle \tilde{u}, \tilde{v} \rangle = \int_Q \overline{\tilde{u}(q)} \tilde{v}(q) dq. \tag{60.100}$$

If we look at the action of $\tilde{U}(g)$, we have

$$\begin{aligned} \langle \tilde{U}(g)\tilde{u}, \tilde{U}(g)\tilde{v} \rangle &= \int_Q \overline{\chi\big(\beta(q, g^{-1})\big)\tilde{u}(\tau_{g^{-1}}q)} \chi\big(\beta(q, g^{-1})\big)\tilde{v}(\tau_{g^{-1}}q) dq \\ &= \int_Q (\overline{\chi}\chi)(\overline{\tilde{u}} \cdot \tilde{v})(\tau_{g^{-1}}q) dq. \end{aligned} \tag{60.101}$$

The factor $\overline{\chi}\chi$ is equal to one. If we perform the change of variable $q' = \tau_{g^{-1}}q$, we have $\tau^*_{g^{-1}} dq' = dq'$ because we suppose that the measure is invariant. At the end of the day we have

$$\langle \tilde{U}(g)\tilde{u}, \tilde{U}(g)\tilde{v} \rangle = \int_Q (\overline{\tilde{u}}\tilde{v})(q') dq' = \langle \tilde{u}, \tilde{v} \rangle \tag{60.102}$$

Since every unitary operator is continuous and since $\mathscr{D}(Q)$ is dense in $L^2(Q, dq)$, we can extend $\tilde{U}(g)$ into a continuous operator on $L^2(Q, dq)$, so that we get a representation

$$\Big(\mathscr{H}, \tilde{U}\Big). \tag{60.103}$$

where $\mathscr{H}$ stands for $L^2(Q, dq)$. Let $\tilde{u} \in \mathscr{H}$ and consider the map

$$\begin{aligned} \alpha(\tilde{u}) \colon G &\to \mathscr{H} \\ g &\mapsto \tilde{U}(g)\tilde{u}. \end{aligned} \tag{60.104}$$

This is the **coaction**. As composition of continuous maps, the map $g \mapsto \tilde{U}(g)\tilde{u}$ is continuous, so that $\tilde{U}$ is an unitary representation.

In a compact notation, we write

$$\tilde{U} = m_\beta \circ \tau^* \tag{60.105}$$

where m_β is the multiplication by β. That representation is the **unitary induced** defined from B and χ.

Chapter 61

Sobolev spaces

Notations

Some notations about functional spaces:

(1) $\mathscr{E}(X)$ is the set of smooth functions on X also denoted by $C^\infty(X)$,
(2) $\mathscr{D}(X;K)$ is the subspace of $\mathscr{E}(X)$ of functions with support in K,
(3) $\mathscr{D}(X)$ is the union of all $\mathscr{D}(X;K)$ when K runs over the compact subsets of X,
(4) $C_c(X)$ compactly supported functions on X.
(5) $C_c^\infty(X)$ smooth compactly supported functions on X.
(6) $\mathscr{C}_F(E)$ is the set of continuous maps from E to F.
(7) $\mathscr{D}_{\mathbb{R}}^{(0)}(X)$ is the set of continuous real functions with compact support in X,
(8) $\mathscr{D}_p^{(r)}(X;K)$ is the set of the differential p-forms of class C^r on K when K is a compact subset of the manifold X,
(9) $\mathscr{E}_F^{(p)}(A)$ is the set of p times continuously differentiable functions $A \to F$; if p is not mentioned it means ∞,
(10) $\mathscr{S}$, the space of smooth rapidly decreasing Schwartz functions, page 3501.

Some distribution spaces:

(1) $M(X)$ is the space of complex Borel measures on X, page 3505
(2) $M_0(X)$ is the subspace of $M(X)$ of compact supported measures, page 3505.
(3) $R(G)$ is the dual of $C^\infty(G)$. All element of this space are not measures.

When f and g are real function, we write $f = O(g)$ if there exists $C \leqslant \infty$ such that $f(x) \leqslant Cg(x)$ for all x.

61.1 Distributions

sec:Distrib

Matter of this section is taken from [943, 944]

Let X be an open set in $\mathbb{R}^n$. A **distribution** on X is a linear form on $\mathscr{D}(X)$ whose restriction to $\mathscr{D}(X;K)$ is continuous for each compact set $K \subset X$. We denote it by $\mathscr{D}'(X)$. More generally, if $\mathscr{A}$ is a space of function, we denote by $\mathscr{A}'$ the set of linear form on $\mathscr{A}$ whose are continuous on each compact.

If T is a linear form on $\mathscr{D}(X)$. For T to be a distribution, it is necessary and sufficient that for all sequence $(f_k) \in C^\infty(X)$ with $f_k \in \mathscr{D}(X,K)$ such that $f_k \to 0$ in $\mathscr{E}(X)$, the sequence (Tf_k) converges to 0 in $\mathbb{C}$.

Let T be a distribution for which all the restrictions to $\mathscr{D}(X;K)$ are continuous for the induced topology from $\mathscr{D}^{(r)}(X;K)$. In this case, we say that T has **order** lesser than r. The *order* of T is the smaller such r. If it doesn't exist, then we say that T has order infinite.

pg:reforder

Proposition 61.1.
Let T be of order $\leqslant r$. Then T is the restriction to $\mathscr{D}(X)$ of a linear form T' on $\mathscr{D}^{(r)}(X)$ whose restriction to each $\mathscr{D}^{(r)}(X;K)$ is continuous. This T' is unique.

I only give the beginning of the proof[1].

Proof. Let K be a compact in X and K' a compact neighbourhood of K in X. There exists a function $h \in C^\infty(X)$ such that $h = 1$ in a compact neighbourhood of K and $h = 0$ outside K'. Consider a function $g \in \mathscr{D}^{(r)}(X;K)$; there exists a sequence $(f_k \in \mathscr{D}(X)) \to g$ for the topology of $\mathscr{E}^{(r)}$. This is the C^∞ approximation of $C^{(r)}$ functions. Note that in general, g don't belong to $\mathscr{D}(X;K)$.

Now, the sequence (hf_k) which is contained in $\mathscr{D}(X;k')$ converges to g in $\mathscr{D}^{(r)}(X;k')$. Then the closure of $\mathscr{D}(X;K')$ in $\mathscr{E}^{(r)}$ contains $\mathscr{D}^{(r)}(X;K)$ and is contained in $\mathscr{D}^{(r)}(X;K')$. □

61.1.1 Dirac distribution

The distribution $\delta_a\colon \mathscr{D}(\mathbb{R}^d) \to \mathbb{C}$ given by $f \to f(a)$ is the **Dirac distribution** at a. It was already defined in 30.59.

We have a two-dimensional generalisation of that.

LEMooYABKooWPXIXZ

Lemma-Definition 61.2.
Let C be a curve in $\mathbb{R}^2$ and a function $f\colon \mathbb{R}^2 \to \mathbb{C}$ integrable on C. The formula

$$\langle \delta_C^f, \phi \rangle = \int_C f\phi. \tag{61.1}$$

defines and element $\delta_C^f \in \mathscr{D}'(\mathbb{R}^2)$.

Proof. Let $\phi_n \overset{\mathscr{D}(\mathbb{R}^2)}{\longrightarrow} 0$; we have

$$|\langle \delta_C^f, \phi_n \rangle| \leqslant \int_C |f\phi_n| \leqslant \|\phi_n\|_\infty \int_C |f|. \tag{61.2}$$

Since f is integrable on C, we have a majoration $\int_C |f| \leqslant \alpha$ for some $\alpha > 0$ and then

$$|\langle \delta_C^f, \phi_n \rangle| \leqslant \alpha \|\phi_n\|_\infty. \tag{61.3}$$

The definition of convergence in $\mathscr{D}(\mathbb{R}^2)$ implies that $\|\phi_n\| \to 0$. □

The usual Dirac distribution is obtained with $C = \{0\}$ and $f(x) = 1$ on C.

61.1.2 Distribution defined from functions

If $f\colon X \to \mathbb{R}$ is an integrable function, we define the distribution T_f by

$$T_j(g) = \int_X fg.$$

If $\mathcal{U} \subset \mathbb{R}^N$, is open and $K \subset \mathcal{U}$ is compact, for all multi-index ν, the map $f \to D^\nu f$ is continuous from $\mathscr{D}(\mathcal{U};K)$ to $\mathscr{D}(\mathcal{U};K)$. This leads us to define the **derivative** of the distribution $T \in \mathscr{D}'(\mathcal{U})$ by

$$(D^\nu T)f = (-1)^{|\nu|} T(D^\nu f). \tag{61.4}$$ eq:defpartialT

In particular, if $T = T_f$, then

$$(\partial_i T_g)f = -\int g\partial_i f = \int (\partial_i g)f = T_{\partial_i g} f. \tag{61.5}$$ eq:defTpri

This relation explains the sign in definition (61.4). The boundary term which should appears in the integral by part is zero because $f \in \mathscr{D}(\mathcal{U};K)$ has compact support.

1. Il faudra la completer

61.1.3 Fourier transform and Schwartz functions

We consider $\{e_1, \dots, e_n\}$, a basis of $\mathbb{R}^N$ and $\{e'_1, \dots, e'_n\}$ the dual basis defined by $\langle e'_i, e_j\rangle = \delta_{ij}$. If $x = \sum_i x_i e_i$ and $\xi = \sum_j \xi_j e'_j$, we write $\langle x, \xi\rangle = \langle \xi, x\rangle = \sum_i \xi_i x_i$.

The space $\mathscr{S}(\mathbb{R}^N)$ of **Schwartz functions** is the space of C^∞ functions φ such that for all polynomials P on $\mathbb{R}^N$ and all multi-index ν, the quantity

$$|P(x)D_x^\nu \varphi(x)| \tag{61.6}$$

is bounded. The topology is given by seminorms

$$\|\varphi\|_{P,\nu} = \sup_{x\in\mathbb{R}^N} |P(x)D_x^\nu \varphi(x)|. \tag{61.7}$$

It is possible to prove that it is a Fréchet space in which all bounded and closed sets are compact; the topology is independent of the basis. A function in this space has special property that its integral is absolutely convergent:

$$\int_{\mathbb{R}^N} |\varphi(x)| \leqslant \infty.$$

It is proved in [945] that on $\mathscr{S}(\mathbb{R}^n)$, the following families of seminorms are equivalent:

$$p_{\alpha\beta}(f) = \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| \tag{61.8a}$$

$$p'_{\alpha\beta}(f) = \int_{\mathbb{R}^n} |x^\alpha \partial^\beta f(s)|\, ds \tag{61.8b}$$

$$p''_{\alpha\beta}(f) = \left(\int_{\mathbb{R}^n} |x^\alpha \partial^\beta (s)|^2\, ds\right)^{1/2}. \tag{61.8c}$$

Let $\xi \in (\mathbb{R}^N)^*$ and consider the function $x \to e^{-2\pi i\langle x,\xi\rangle}\varphi(x)$ which belongs to $\mathscr{S}(\mathbb{R}^N)$ as long as $\varphi \in \mathscr{S}(\mathbb{R}^N)$. The **Fourier transform** of $\varphi \in \mathscr{S}(\mathbb{R}^N)$ is the function $\hat\varphi\colon (\mathbb{R}^N)^* \to \mathbb{C}$ given by

$$\hat\varphi(\xi) = \int \varphi(x) e^{-2\pi i\langle x,\xi\rangle} dx. \tag{61.9}$$

Classical results are summarized in the following theorem (proof in the case of space $\mathscr{S}$ is given in [943]):

Theorem 61.3.
The Fourier transform is a topological vector space isomorphism from $\mathscr{S}(\mathbb{R}^N)$ into $\mathscr{S}(\mathbb{R}^N)$ and the inverse is given by formula

$$\varphi(x) = \int \hat\varphi(\xi) e^{2\pi i\langle x,\xi\rangle} d\xi.$$

Equalities of Parseval and Plancherel holds:

$$\int \phi\overline{\psi} = \int \hat\phi\overline{\hat\psi} \text{ and } \int |\phi|^2 = \int |\hat\phi|^2 \tag{61.10}$$

where the bar denotes the usual complex conjugation. Left hand side integrals are taken on $\mathbb{R}^N$ and right integrals over $(\mathbb{R}^N)^*$.

Corollary 61.4.
Fourier transform can be extended to an isometry $L^2(\mathbb{R}^N) \to L^2\big((\mathbb{R}^N)^*\big)$.

61.1.4 Support of a distribution

Let $\mathcal{U}$ be an open set in X and K, a compact in $\mathcal{U}$. The map $\mathscr{D}(X;K) \to \mathscr{D}(\mathcal{U};K)$ given by $f \to f_{\mathcal{U}}$ is an isomorphism whose inverse is $f \to f^{\mathcal{U}}$ where $f^{\mathcal{U}}$ is just the prolongation of f with zero outside $\mathcal{U}$. For a distribution T on X, we consider $T|_{\mathcal{U}}\colon \mathscr{D}(\mathcal{U};K) \to \mathbb{C}$,

$$T|_{\mathcal{U}}(f) = T(f^{\mathcal{U}}).$$

This is a distribution on $\mathcal{U}$ called the **induced** from T on $\mathcal{U}$. A distribution on $\mathcal{U}$ is not always the restriction of a distribution on X and when it is, the prolongation is not unique in general. There exists a prolongation theorem in certain cases:

Theorem 61.5.
Let $(\mathcal{U}_\lambda)_{\lambda\in L}$ be an open covering of X and for each $\lambda \in L$, a distribution T_λ on $\mathcal{U}_\lambda$. We suppose that for all $\lambda, \mu \in L$, $T_\lambda|_{\mathcal{U}_\lambda\cap\mathcal{U}_\mu} = T_\mu|_{\mathcal{U}_\lambda\cap\mathcal{U}_\mu}$. Then there exists an unique distribution T on X such that for all $\lambda \in L$, $T|_{\mathcal{U}_\lambda} = T_\lambda$.

Now if the restriction of T to each $\mathcal{U}_\lambda$ is zero, then the restriction to the union is zero too because the null distribution answers the theorem. So the union of all the open set on which T is zero is an open on which T is zero. It is the largest open $V \subset X$ on which T is zero. The complementary $S = \complement V$ is the **support** of the distribution T. We note it $\operatorname{Supp}(T)$.

Let $x \in \operatorname{Supp} T$: there exists no open containing x on which T is zero. With other words, for all neighbourhood V of x, there exists a function $f \in \mathscr{D}(X)$ with support contained in V with $T(f) \neq 0$.

61.1.5 Duality

If L and M are algebras, the set $\operatorname{Hom}(L, M)$ contains all the maps $f\colon L \to M$ such that $f(xy) = f(x)f(y)$. We are mainly interested in $\mathbb{C}$-algebras: algebras L endowed with a product $\mathbb{C} \times L \to L$. Then we are leads to look at $\operatorname{Hom}_{\mathbb{C}}(L, \mathbb{C})$, the subset of $\operatorname{Hom}(L, \mathbb{C})$ of maps which commutes with the latter product: $A \in \operatorname{Hom}_{\mathbb{C}}(L, \mathbb{C})$ when $A \in \operatorname{Hom}(L, \mathbb{C})$ and $A(zl) = zA(l)$. We denote.

$$L^* = \operatorname{Hom}_{\mathbb{C}}(L, \mathbb{C}). \tag{61.11}$$

This is defined when L is a $\mathbb{C}$-module.

We consider now a *topological* vector space: L is a topological group for addition and the map $\mathbb{C} \times L \to L$, $(\lambda, x) \to \lambda x$ is continuous with respect to the two variables. We denote by L' the space of linear *and continuous* maps $L \to \mathbb{C}$; this is the **topological dual** of L.

Elements of L^* have no continuity condition. In the general case, $L' \subset L^*$, and in the finite dimensional case, $L' = L^*$. The space L^* is the **dual module** of L while L' is the **topological conjugate** space of L. In both cases, we speak about **dual space** of L.

Proposition 61.6.
Dual space $\mathscr{E}'(X)$ is the space of distribution with compact support. prop:dualCinfcompact

We can consider the sequence $C_c^\infty \subset \mathscr{S} \subset C^\infty$ of inclusion with dense images. The transposition of this gives the continuous inclusion sequence

$$\mathscr{E}' \subset \mathscr{S}' \subset \mathscr{D}'.$$

We say that an element of $\mathscr{S}'(\mathbb{R}^N)$ is a **tempered distribution**.

Proposition 61.7.
A distribution on $\mathbb{R}^N$ is tempered if and only if it is a finite sum of derivatives of continuous functions which are bounded at infinity by a polynomial. prop_distr_temp_sum

A continuous function is not necessarily derivable. By "derivative of a continuous function" we mean a distribution of the form $(T_f)'$ (derivative in the sense of distributions) which we write $T_{f'}$ by abuse of notation. This notation is motivated by equation (61.5).

Proof of proposition 61.7 . Let us first proof that a distribution T is tempered if and only if the map $\varphi \to T\varphi$ is continuous on C_c^∞ for the topology induced from $\mathscr{S}$. Let $\tilde{T}\colon C_c^\infty \to \mathbb{C}$ be the map equals to T on C_c^∞ and not defined anywhere else. If $\mathcal{O}$ is open in $\mathbb{C}$, then $\tilde{T}^{-1}(\mathcal{O}) = C_c^\infty \cap T^{-1}(\mathcal{O})$. This is open in C_c^∞ (for the induced topology from $\mathscr{S}$) if and only $T^{-1}(\mathcal{O})$ is open in $\mathscr{S}$.

Let f be continuous and $T = T_{f'}$. We want $\varphi \to (T_f\varphi)'$ to be continuous on $\mathscr{S}$. We can write it as

$$(T_f)'\varphi = -T_f(\varphi') = \int_{\mathbb{R}^N} f\varphi'.$$

The integral exists because $\varphi \in \mathscr{S}$, so φ' is decreasing at infinity more rapidly than the inverse of any polynomials, while f is bounded by a polynomial.

Let us now consider, a tempered distribution T. We have to prove that $T = T_{f'}$ for a certain continuous function f bounded by a polynomial. Since T is linear, its continuity is assured by the only continuity at zero. A neighbourhood of zero in $\mathscr{S}$ reads under the form

$$A_{P,Q,\varepsilon} = \{\varphi \in \mathscr{S} \text{ tel que } \sup_{x\in\mathbb{R}^N} |P(X)Q(\partial_x)\varphi(x)| < \varepsilon\}$$

for a choice of polynomials P, Q and a $\varepsilon > 0$. Let $\mathcal{O}$ be a neighbourhood of rayon ε around 0 in $\mathbb{C}$; we have

$$T^{-1}(\mathcal{O}) = \{\varphi \in \mathscr{S} \text{ tel que } |T\varphi| \leqslant \varepsilon\}.$$

In order to be an open set, we have to find two polynomials P and Q (depending on ε but not on φ) such that

$$|T\varphi| \leqslant \sup_{x\in\mathbb{R}^N} |P(x)Q(\partial_x)\varphi(x)|.$$

The continuity of T gives the existence of reals $m, h \geqslant 0$ and $C > 0$ such that

$$|T\varphi| \leqslant \sup_{|p|\leqslant m} \sup_{x\in\mathbb{R}^N} |(1+|x|^2)^h(\partial_x)^p\varphi(x)|.$$

Let us pose $\varphi_h(x) = (1+|x|^2)^h\varphi(x)$. It still belongs to C_c^∞ and moreover, the map $\varphi \to \varphi_g$ is a bijection on C_c^∞. By induction on h, we see that

$$|(\partial_x)^p\varphi(x)| \leqslant C_{p,h}(1+|x|^2)^{-h} \sum_{q\leqslant p} |(\partial_x)^q\varphi_h(x)| \tag{61.12}$$

where $q \leqslant p$ means $q_1 \leqslant p_1, \dots, q_n \leqslant p_n$.[2].

□

The space of **currents** is the dual (with respect to $\mathbb{R}$) space of the space of differential forms. Current is the generalization of differential forms in the same sense that distributions are a generalization of functions. An example of k-current is given by a $(n-k)$-form σ by setting

$$C_\sigma(\omega) = \int_M \sigma \wedge \omega.$$

61.2 Measure, distribution and integral

sec_distrib_mesure

Do you know what is yellow and equivalent to the existence of non measurable functions? Answer in the footnote[3].

Most of links between measure theory and distribution are given in [946]. We will state a lot of results without proof; they can be found in this reference. We always suppose that X is locally compact.

Let $\mathscr{D}^{(0)}(X;K)$ be the set of continuous functions on X whose support is contained in a compact $K \subset X$. A **measure** on X is a linear form $\mu\colon \mathscr{D}^{(0)}(X) \to \mathbb{C}$ such that for all compact $K \subset X$, there exist a real $a_K \geqslant 0$ for which for all $f \in \mathscr{D}^{(0)}(X;K)$

$$|\mu(f)| \leqslant a_K\|f\|. \tag{61.13}$$

2. je ne fais pas le reste de la démonstration
3. The Zorn lemon!

In this case, the restriction $\mu|_{\mathscr{D}^\infty}(X;K)$ is continuous for the induced topology from $\mathscr{D}^{(0)}(X;K)$. Indeed if $\mathcal{O}$ is open in $\mathbb{C}$, then

$$\mu|^{-1}_{\mathscr{D}(X;K)}(\mathcal{O}) = \mathscr{D}(X,K) \cap \mu^{-1}(\mathcal{O}) \tag{61.14}$$

eq:18105r1

but the continuity condition on $\mu\colon \mathscr{D}^{(0)} \to \mathbb{C}$ makes $\mu^{-1}(\mathcal{O})$ open in $\mathscr{D}^{(0)}$. Then expression (61.14) describes an open set in $\mathscr{D}(X;K)$ for the induced topology of $\mathscr{D}^{(0)}(X;K)$.

The measures are exactly the distribution of order zero, *confer* page 3499.

61.2.1 Example: the Lebesgue measure

Let $f \in \mathscr{D}^{(0)}(\mathbb{R})$. For all a, $b \in \mathbb{R}$ such that $\operatorname{Supp} f \subset [a,b]$, the value of $\int_a^b f$ is the same and is denoted by $\int_{\mathbb{R}} f$. The map $f \to \int_{\mathbb{R}} f$ is a linear continuous function on $\mathscr{D}(\mathbb{R})$. This is a measure because if $f \in \mathscr{D}^{(0)}(\mathbb{R},K)$ with $K=[a,b]$, then

$$\left|\int_{\mathbb{R}} f\right| \leqslant (b-a)\|f\|$$

from the mean value theorem. We recognize the usual **Lebesgue measure** on $\mathbb{R}$.

The measure μ is **positive** when for all $f \geqslant 0 \in \mathscr{D}^{(0)}_{\mathbb{R}}$, we have $\mu(f) \geqslant 0$. It is **real** if $\mu(f) \in \mathbb{R}$ whenever $f \in \mathscr{D}^{(0)}_{\mathbb{R}}(X)$. Since any function can be decomposed into $f = f^+ - f^-$, a positive measure is always real.

From now we only consider positive measures on a locally compact set.

Proposition 61.8.
If μ is a linear form on $\mathscr{D}^{(0)}_{\mathbb{R}}(X)$ with $\mu(f) \geqslant 0$ when $f \geqslant 0$, then μ is a measure.

A function $f\colon X \to \overline{\mathbb{R}}$ is **lower semicontinuous** at $x_0 \in X$ when for all $\alpha \in \overline{\mathbb{R}}$ such that $\alpha < f(x_0)$, there exists a neighbourhood of x_0 in which $\alpha < f$. It is **upper semicontinuous** when for all $\alpha' > f(x_0)$, we have $\alpha' > f$ on a neighbourhood.

We consider the set $\mathcal{S}(X)$ of lower semi continuous functions from X to $\overline{\mathbb{R}}$ which are minored by a function of $\mathscr{D}^{(0)}_{\mathbb{R}}(X)$. In particular, if $f \in \mathcal{S}(X)$, one can define

$$\mu^*(f) = \sup_{\substack{g \in \mathscr{D}^{(0)}_{\mathbb{R}}(X) \\ g \leqslant f}} \mu(g) \tag{61.15}$$

which is a well defined number in $\mathbb{R} \cup \{+\infty\}$

Proposition 61.9.
Let $(f_n \in \mathcal{S})$ an increasing sequence and $f = \sup_n f_n$. Then

(1) f exists and $f \in \mathcal{S}$

(2) $\mu^*(f) = \sup_n \mu^*(f_n) = \lim_{n\to\infty} \mu^*(f_n)$

If $f\colon X \to \overline{\mathbb{R}}$ is *any* function, a function $h \in \mathcal{S}(X)$ with $h \geqslant f$ always exists, so we can define

$$\mu^*(f) = \inf_{\substack{h \geqslant f \\ h \in \mathcal{S}(X)}} \mu^*(h). \tag{61.16}$$

This number $\mu^*(f) \in \overline{\mathbb{R}}$ is the **upper integral** of f.

Proposition 61.10.
If $(f_n\colon X \to \overline{\mathbb{R}})$ is an increasing sequence with $\mu^*(f_n) > -\infty$, then

$$\mu^*(\sup_n f_n) = \sup_n \mu^*(f_n) = \lim_{n\to\infty} \mu^*(f_n).$$

Proposition 61.11.
For all sequence of functions $(f_n \geqslant 0)$*, we have*

$$\mu^*\left(\sum_{n=1}^{\infty} f_n\right) \leqslant \left(\sum_{n=1}^{\infty} \mu^*(f_n)\right).$$

Now, for a function $f\colon X \to \overline{\mathbb{R}}$, we put

$$\mu_*(f) = -\mu^*(-f) \tag{61.17}$$

and we call it the **lower integral** of f for the measure μ. It fulfils

$$\mu_*(f) \leqslant \mu^*(f).$$

When the equality are true, we define $\mu(A)$ for a subset $A \subset X$ by

$$\mu(A) = \mu_*(1_A) = \mu_*(1_A). \tag{61.18}$$

We consider $M(X)$, the set of all the **Borel measures** on X: these are measure for which all Borel sets are measurable. When a measure μ is countably additive, we define

$$|\mu|(E) = \sup \sum_{k=1}^{\infty} |\mu(E_k)| \tag{61.19}$$

where the supremum is taken over all decomposition of E in disjoints sets E_k. It induces a norm on $M(G)$ by

$$\|\mu\| = |\mu|(X).$$

The subset $M_0(X)$ contains the Borel regular measures with compact support. A measure is **regular** when for all $B \subset X$ such that $\mu(B)$ exists, we have

$$\begin{aligned}\mu(B) &= \sup\{\mu(K) \text{ tel que } K \subseteq B \text{ is compact}\}\\ &= \inf\{\mu(A) \text{ tel que } B \subseteq A,\ A \text{ is open}\}\end{aligned} \tag{61.20}$$

61.2.2 Integration on more general spaces

Let X be locally compact and μ, a measure on X. We define

$$\|\mu\| = \sup_{\substack{\|f\|\leqslant 1\\ f\in\mathscr{D}^{(0)}(X)}} |\mu(f)| \in \mathbb{R} \cup \{+\infty\} \tag{61.21}$$

We say that the *positive measure* μ is **bounded** when $\mu^*(X) = \mu^*(1_X)$ is a finite real. Properties of $\|\mu\|$ are that

$$\|\mu\| = |\mu|^*(1_X) \tag{61.22}$$

and that $\|\mu\|$ is finite if and only if $|\mu|$ is bounded.

Proposition 61.12.
If μ *is bounded, any bounded and measurable function* $f\colon X \to \mathbb{R}$ *is integrable and*

$$\left|\int_X f\,d\mu\right| \leqslant \|f\|\mu(X) \tag{61.23}$$

When μ is not positive, we say that it is **bounded** if $\|\mu\|$ is finite. Equation (61.23) shows that $f \to \int_X f\,d\mu$ is a continuous linear form on $C^\infty(X)$. All continuous linear form on this space are however not of the form (61.23).

61.2.3 Integration of vector valued functions

Let us consider E, a vector space of *finite* dimension and a function $f\colon X \to E$. The functions f_i are defined by

$$f(x) = \sum_{i=1}^{\infty} f_i(x)e_i$$

where $\{e_i\}$ is a basis of E. We define

$$\int f\,d\mu = \sum[\int f_i]e_i \tag{61.24}$$

when all the integrals of the right hand side make sense. The fact for f to be integrable is equivalent to the fact that $x \to \xi(f(x))$ is integrable for all $\xi \in E'$ because $\xi \circ f$ is a linear combination of all the f_i.

Let I be a set and consider a map $x \to \mathrm{Fun}(I, \mathbb{C})$, $x \to f_x$. We say that this is **scalar integrable** if for all $\alpha \in I$, the map $x \to f_x(\alpha)$ is integrable.

The theorem which treat with infinite dimension is the following:

Theorem 61.13.
Let $\mathcal{F}$ be a Fréchet space and $\mathcal{F}'$ his dual. Let $x \to f_x$ be a map $X \to \mathcal{F}'$ such that for all converging sequence $(a_n \in \mathcal{F})$, there exists a function $g\colon X \to \mathbb{R}$, $g \geqslant 0$ such that $\mu^*(g) \leqslant \infty$ and $|f_x(a_n)| \leqslant g(x)$ for all n. Then there exists one and only one $\xi \in \mathcal{F}'$ such that

$$\xi(z) = \int_X f_x(z)\,d\mu(x). \tag{61.25}$$

61.2.4 Weak integral

Let (X, μ) be a measured space and $f\colon X \to L$ a map from X to a locally convex space L. We say that f is *weakly integrable* if for all $\chi \in L'$, the map $x \to \chi(f(x))$ is integrable for the measure μ, i.e. if $\mu(\chi \circ f)$ makes sense. The **weak integral** of f for with respect to the measure μ is defined by the requirement

$$\chi\left(\int_X f\,d\mu\right) = \int_X (\chi \circ f)\,d\mu \tag{61.26}$$

eq:chbilemirhi

for all $\chi \in L'$. The weak integral $\int_X f$ is an element of $(L')^*$. In fact the left hand side of equation (61.26) should better be noted

$$\left(\int_X f\,d\mu\right)(\chi)$$

We know that $(L')^*$ can be seen as a subset of L. It can be shown that if f is continuous and X compact, then $\int_X f \in L$.

61.3 Distribution on groups

Most of this section comes from [945].

We immediately put the attention to the reader on the fact that $(\mathbb{R}^N, +)$ is a Lie group. Let G be a Lie group, we define $R(G)$ as the dual of $C^\infty(G)$. This is the space of compact supported distributions on G. The support of $T \in R(G)$ is the smallest compact K for which $T(\phi) = 0$ whenever $\phi|_K^{(r)} = 0$ for all r. When K is compact in G, we consider $R(G, K)$, the subspace of $R(G)$ of distributions with support contained in K.

We focus on $R(G, \{e\})$: distributions in this space only depends on values of test functions (and derivatives) at e. It should be possible to reconstruct this space from the only data of the Lie algebra $\mathfrak{g}$ instead of then whole group. We'll see later that it is indeed possible.

61.3.1 Convolution product

The **convolution** of two distributions T_1 and T_2 in $R(G)$ is given by

$$(T_1 \star T_2)(\phi) = T_1(\phi_{T_2}) \tag{61.27}$$

where $\phi_{T_2}\colon G \to \mathbb{C}$ is given by

$$\phi_{T_2}(g) = T_2(L_g^{-1}\phi) \tag{61.28}$$

and $(L_g\phi)(g') = \phi(g^{-1}g')$ or in a more convenient way, $\phi_\nu(g) = \nu(\phi(g\ \cdot\ ))$.

Let us show that one retrieve the well know distribution convolution in the case of $G = (\mathbb{R}^N, +)$. We consider T_f and T_g, the distributions defined from functions f and g on $\mathbb{R}^N$. We have $\phi_{T_g}(t) = T_g(L_t^{-1}\phi)$ where $(L_t^{-1}\phi)(x) = \phi(x+t)$. Then

$$\phi_{T_g}(t) = \int g(x)(L_t^{-1}\phi)(x)\,dx = \int g(x)\phi(x+t)\,dx. \tag{61.29}$$

A change of variable gives

$$(T_g \star T_g)(\phi) = \int f(t)\phi_{t_g}(t)\,dt = \int (f\star g)(u)\phi(u)\,du = T_{f\star g}\phi \tag{61.30}$$

with the usual convolution product between function.

Remarque 61.14.
We see that the convolution operation of functions on $\mathbb{R}^N$, which has remarkable properties in experimental physics, naturally generalises to a convolution product on distribution which will give a homomorphism between these distributions and the enveloping algebra of the group. Wonderful isn't?

Lemma 61.15.
The map $\psi\colon \mathcal{G} \to R(G,\{e\})$ given by $\psi(\tilde{X})f = Xf$ is a homomorphism.

Proof. Let us prove that $[\psi(\tilde{X}) \star \psi(\tilde{Y})]f = (\tilde{X}\tilde{Y}f)_e$. We have

$$\begin{aligned}
[\psi(\tilde{X}) \star \psi(\tilde{Y})] &= \psi(\tilde{X})\big(c(\psi(\tilde{Y}),\phi)\big) \\
&= \frac{d}{dt}\Big[c(\psi(\tilde{Y}),\phi)\tilde{X}_e(t)\Big]_{t=0} \\
&= \frac{d}{dt}\frac{d}{ds}\Big[(L^{-1}_{X(t)}\phi)(Y(s))\Big]_{\substack{s=0\\t=0}} \\
&= \frac{d}{dt}\frac{d}{ds}\Big[\phi(X(t)Y(s)))\Big]_{\substack{s=0\\t=0}}.
\end{aligned} \tag{61.31}$$

Now, $\tilde{Y}\phi$ is a function from G to $\mathbb{C}$ on which we can apply the vector $\tilde{X}_e = X$. We have:

$$\begin{aligned}
(\tilde{X}\tilde{Y})_e\phi &= \tilde{X}_e(\tilde{Y}\phi) \\
&= \frac{d}{dt}\Big[(\tilde{Y}\phi)(\tilde{X}_e(t))\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\tilde{Y}_{X(t)}\phi\Big]_{t=0} \\
&= \frac{d}{dt}\frac{d}{ds}\Big[\phi\big(\tilde{Y}_{X(t)}(s)\big)\Big]_{\substack{s=0\\t=0}} \\
&= \frac{d}{dt}\frac{d}{ds}\Big[\phi\big(X(t)Y(t)\big)\Big]_{\substack{s=0\\t=0}}.
\end{aligned} \tag{61.32}$$

□

61.3.2 Representations

Let G be a locally compact metrisable Lie group and V, an Hausdorff topological vector space on $\mathbb{C}$. We say that $U : G \to \operatorname{End} V$ is a **linear continuous representation** of G on V when

(1) $U(gh) = U(g) \circ U(h)$ for all g, $h \in G$,

(2) for each $v \in V$, then map $g \to U(g)v$ is continuous from G into V.

In most of cases, we want the representation to be **unitary**: for all $g \in G$ and for all v, $w \in V$,

$$\langle U(g)v, U(g)w\rangle = \langle v, w\rangle. \tag{61.33}$$

In particular, $U(g)^* = U(g)^{-1}$. Let v, $w \in V$. The function $G \to \mathbb{C}$ given by

$$g \to \langle U(g)v, w\rangle$$

is continuous and bounded. The boundary comes from Schwartz and the fact that $U(g)$ is an isometry:

$$\langle U(g)v, w\rangle \leqslant \|U(g)v\|\|w\| = \|v\|\|w\|. \tag{61.34}$$

Equation (61.26) then shows that it is an integrable function and that

$$\left|\int_G \langle U(g)v, w\rangle \, d\mu(g)\right| \leqslant \|\mu\|\|\langle U(\,\cdot\,)v, w\rangle\| \leqslant \|\mu\|\|v\|\|w\|. \tag{61.35}$$

eq:Tmuvborn

61.3.3 Representation on a Hilbert space

Let H be a Hilbert space and $v \in H$. The map $\phi_v : H \to \mathbb{C}$,

$$\phi_v(w) = \langle v, w\rangle$$

is linear of norm $\|w\|$ and then is bounded. This is an element of H'. A great theorem allows us to identify H and H' by $v \to \phi_v$.

Theorem 61.16 (Riesz-Fisher)**.**
Let H be a Hilbert space and $\xi \in H'$. Then there exists one and only one $v \in H$ such that

$$\xi(w) = \langle v, w\rangle$$

for all $w \in H$. In particular, $H' \simeq H$.

Let us now consider a continuous, linear and unitary representation $U : G \to \operatorname{End} H$. For given measure μ and vector v, we consider $T_{\mu,v} : H \to \mathbb{C}$,

$$T_{\mu,v}(w) = \int_G \langle U(g)v, w\rangle.$$

Equation (61.35) shows that it is continuous and Riesz-Fischer gives us a vector u such that $T_{\mu,v}(w) = \langle u, w\rangle$. The so defined vector u is written $U(\mu)v$:

$$\int_G \langle U(g)v, w\rangle \, d\mu(g) = \langle U(\mu)v, w\rangle \tag{61.36}$$

is the definition of $U(\mu) \in \operatorname{End} H$. As notation principle, we write

$$U(\mu) = \int_G U(g) \, d\mu(g).$$

Proposition 61.17.
This constructions gives a morphism between algebra of bounded measures (the algebra product is the convolution) and $\operatorname{End} H$*:*

$$U(\mu \star \nu) = U(\mu) \circ U(\nu) \tag{61.37}$$

Proof. We have

$$\langle U(\mu\star\nu)v,w\rangle=\int_G\langle U(g)v,w\rangle d(\mu\star\nu)(g)=(\mu\star\nu)\langle U(\ \cdot\ )v,w\rangle=\mu(f_\nu)\tag{61.38}$$

where $f_\nu(g)=\int_G\langle U(gh)v,w\rangle d\nu(h)$. Then

$$\begin{aligned}\langle U(\mu\star\nu)v,w\rangle&=\int_G\int_G\langle U(h)v,U(g)^*w\rangle\,d\nu(h)d\mu(g)\\&=\int_G\langle U(\nu)v,U(g)^*w\rangle\,d\mu(g)\\&=\langle U(g)\circ U(\nu)v,w\rangle.\end{aligned}\tag{61.39}$$

□

61.3.4 Slightly more general

[4]

Let $\mu\in M_0(G)$ and a continuous representation $T\colon G\to\operatorname{End}V$ where V is locally convex[5]. We can weakly integrate T on G for the measure μ by setting that

$$\int_G T(g)\,d\mu(g)\in\big(\operatorname{End}(V)'\big)^*$$

such that for all $\chi\in\operatorname{End}(v)'$,

$$\Big\langle\int_G T(g)\,d\mu(g),\chi\Big\rangle=\int_G\langle\chi,T(g)\rangle\,d\mu(g).\tag{61.40}$$

Since T is continuous and μ has compact support, one can see that

$$\int_G T(g)\,d\mu(g)\in\operatorname{End}(V)$$

in the sense of the usual embedding $(\operatorname{End}(v)')^*\subset\operatorname{End}(V)$. This element is denoted by $T(\mu)$.

61.3.5 Representation of $M_0(G)$

[6]

Let $T\colon G\to\operatorname{End}(V)$, a representation, μ a compact supported measure on G. Then there exists one and only one $A\in\operatorname{End}(V)$ such that for all $\chi\in\operatorname{End}(V)'$, we have

$$\chi(A)=\int_G\chi(T(g))\,d\mu(g),\tag{61.41}$$

this A is denoted by $T(\mu)$. It fulfils

$$\chi(T(\mu))=\int_g\chi(T(g))\,d\mu(g).\tag{61.42}$$

This way to define $T\colon M_0(G)\to\operatorname{End}(V)$ is a representation. Indeed, $T(\mu\star\nu)$ is defined by

$$\chi(T(\mu\star\nu))=(\mu\star\nu)\langle\chi,T(.)\rangle=\int_G\int_G\langle\chi,T(gh)\rangle\,d\nu(h)\,d\mu(g).\tag{61.43}$$

4. Tout ceci est á préciser.
5. Ce qui implique que $\operatorname{End}(V)$ est locallement convexe pour la topologie forte, il semble que cela joue un rôle.
6. Ceci n'est pas complètement compris non plus

Let us suppose $\chi(T(g)\circ T(h))=\chi(T(g))\chi(T(h))$; in this case Fubini theorem gives

$$\begin{aligned}(\mu\star\nu)\langle\chi,T(.)\rangle&=\int_G\int_G\langle\chi,T(g)\rangle d\mu(g)\langle\chi,T(h)\rangle d\nu(h)\\&=\chi(T(\mu))\chi(T(\nu)).\end{aligned}\tag{61.44}$$

So for all such χ, we have

$$\chi\big(T(\mu\star\nu)\big)=\chi\big(T(\mu)\circ T(\nu)\big).$$

It is not sufficient to conclude that $T(\mu\star\nu)=T(\mu)\circ T(\nu)$ because G is not abelian [7].

61.4 A boundary aware Sobolev space

We want to define $H_0^1(\Omega)$ as the subset of $H^1(\Omega)$ made of the functions vanishing on $\partial\Omega$. But since the elements of $H^1(\Omega)$ are class of functions, and since $\partial\Omega$ is of zero measure, such a definition makes no sense. Instead we define

DEFooFICWooBWCDyO

Definition 61.18.
The space $H_0^1(\Omega)$ is the closure of $\mathscr{D}(\Omega)$ in $H^1(\Omega)$.

Notice that $\mathscr{D}(\Omega)$ is included in $H^1(\Omega)$, so that the definition makes sense.

61.19.
The intuitive setting of this definition is the following. If $v\in H_0^1(\Omega)$, we have a sequence $v_i\in\mathscr{D}(\Omega)$ such that

$$\|v-v_i\|_{H^1(\Omega)}\to 0.\tag{61.45}$$

If one forgets about the classes, for each i one has $v_i(x)=0$ when $x\in\partial\Omega$, so with the limit $v=0$ on $\partial\Omega$.

Due to the class stuff, this is not the truth, but the motivation for the definition 61.18.

Lemma 61.20 ([947])**.**
The formula

$$|u|_{1,\Omega}=\|\nabla u\|_{L^2(\Omega)}=\left(\sum_{i=1}^d\int_\Omega\left(\frac{\partial u}{\partial x_i}\right)^2\right)^{1/2}\tag{61.46}$$

is a norm [8] on $H_0^1(\Omega)$.

Proof. The condition $|\lambda u|_{1,\Omega}=|\lambda||u|_{1,\Omega}$ is immediate. The triangular inequality is shown in much the same way as the one for the euclidian norm, see (10.123).

The point to be proven is that $|u|_{1,\Omega}=0$ only if $u=0$ in the sense of the classes. Let $|u|_{1,\Omega}=0$; since each of the d integrals is positive we have

$$\int_\Omega\left(\frac{\partial u}{\partial x_i}\right)^2=0\tag{61.47}$$

for every $i=1,\dots,d$. Since $(\partial_i u)^2\geqslant 0$ we deduce $\partial_i u=0$ almost everywhere. Now, from a dimensional generalization of proposition 27.190 we get the conclusion [9]. □

So now we consider the metric space

$$\big(H_0^1(\Omega),|.|_{1,\Omega}\big).\tag{61.48}$$

The norm on $H^1(\Omega)$ given by the inner product (31.37) can be written

$$\|u\|_{H^1(\Omega)}=\|u\|_{L^2(\Omega)}+\|\nabla u\|_{L^2(\Omega)}=\|u\|_{L^2(\Omega)}+|u|_{1,\Omega}.\tag{61.49}$$

7. D'où le fait que je ne considère pas ceci comme bien compris.
8. Definition 7.158.
9. If you know a precise statement of the "dimensional generalization", let me know.

LEMooEVQKooYoZmbH

Lemma 61.21.
Let $v_i \in H_0^1(\Omega)$ satisfying $v_i \xrightarrow{H^1(\Omega)} v$. Then

(1) $v_i \xrightarrow{H_0^1(\Omega)} v$

(2) $v_i \xrightarrow{L^2(\Omega)} v$

(3) $\|v_i\|_{L^2} \to \|v\|_{L^2}$

(4) $|v_i|_{1,\Omega} \to |v|_{1,\Omega}$.

Proof. By hypothesis we have

$$\|v_i - v\|_{H^1} = \|v_i - v\|_{L^2} + |v_i - v|_{1,\Omega} \to 0. \tag{61.50}$$

Since the two terms are positive, they separately converge to 0. Thus $\|v_i - v\|_{L^2} \to 0$ and $|v_i - v|_{1,\Omega} \to 0$. This means $v_i \xrightarrow{L^2} v$ and $v_i \xrightarrow{H_0^1} v$.

The two last statement about the norms are immediate consequences of the continuity of the norm. □

THOooMIHQooYShOps

Theorem 61.22 (Poincaré inequality[947])**.**
Let Ω be an open part of $\mathbb{R}^d$ which is bounded in at least one direction. There exists a constant C (which can depend on Ω) such that

$$\|v\|_{L^2(\Omega)} \leqslant C\|v\|_{1,\Omega} \tag{61.51}$$

for every $v \in H_0^1(\Omega)$

Proof. Let us first consider $v \in \mathscr{D}(\Omega)$ and extend v with 0 outside Ω. For the sake of notational simplicity we suppose that Ω is bounded in the direction x_1, that is $\Omega \subset \{x \text{ tel que } x_1 \in]a, b[\}$.

Since v has compact support in Ω, we have $v(a, x') = 0$ for every $x' \in \mathbb{R}^{d-1}$ and we can write

$$v(x_1, x') = \int_a^{x_1} (\partial_1 v)(t, x')dt = \langle 1, \partial_1 v \rangle_{L^2\left(]a,x_1[\right)}, \tag{61.52}$$

so that

$$|v(x_1, x')|^2 \leqslant \langle 1, 1 \rangle \langle \partial_1 v, \partial_1 v \rangle \tag{61.53a}$$

$$= (x_1 - a) \int_a^{x_1} |\partial_1 v|^2 \tag{61.53b}$$

$$\leqslant (b - a) \int_a^b |\partial_1 v|^2. \tag{61.53c}$$

Let us integrate that inequality over $x' \in \mathbb{R}^{d-1}$:

$$\int_{\mathbb{R}^{d-1}} |v(x_1, x')|^2 dx' \leqslant (b - a) \int_{\mathbb{R}^{d-1}} \int_a^b |\partial_1 v|^2 = (b - a) \int_{\mathbb{R}^d} |\partial_1 v|^2 = (b - a)\|\partial_1 v\|^2_{L^2(\mathbb{R}^d)}. \tag{61.54}$$

We used the Fubini theorem 14.304. This is allowed because v is compactly supported, so that $|\partial_1 v|$ is integrable on $\mathbb{R}^{d-1} \times]a, b[$.

We integrate with respect to x_1 over $]a, b[$. On the left we have

$$\int_a^b \int_{\mathbb{R}^{d-1}} |v(x_1, x')|^2 dx' dx_1 = \int_{\mathbb{R}^d} |v(x)|^2 dx, \tag{61.55}$$

And on the right the integration is a simple multiplication by $(b - a)$:

$$\|v\|^2_{L^2(\mathbb{R}^d)} \leqslant (b - a)^2 \|\partial_1 v\|^2_{L^2(\mathbb{R}^d)}. \tag{61.56}$$

Now we have

$$\|v\|_{L^2}^2 \leqslant (b-a)^2\|\partial_1 v\|^2 \leqslant C\sum_{i=1}^{d}\|\partial_1 v\|^2 = C|v|_{1,\Omega} \tag{61.57}$$

where C is a majoration of $(b-a)^2$ and 1.

The result is proved for $v \in \mathscr{D}(\Omega)$.

Let now consider $u \in H_0^1(\Omega)$. By definition, there exists a sequence $v_i \in \mathscr{D}(\Omega)$ such that

$$v_i \xrightarrow{H^1(\Omega)} u. \tag{61.58}$$

For each i we have $\|v_i\|_{L^2(\mathbb{R}^d)}^2 \leqslant C|v_i|_{1,\Omega}$. Taking the limit and using the lemma 61.21 to enter the limits inside the norms we get the result. □

61.5 Smooth diffeomorphisms

We already defined the Sobolev spaces $H^s(\mathbb{R}^d)$ in definition 31.13. Let U and V be open parts of $\mathbb{R}^d$ and consider a C^∞ diffeomorphism $\psi\colon U \to V$ whose derivatives are bounded.

For $u \in C^\infty(U)$ we write

$$\begin{aligned} Tu\colon V &\to \mathbb{C} \\ x &\mapsto (u\circ\psi^{-1})(x). \end{aligned} \tag{61.59}$$

EQooHSVBooIZRxzh

A diffeomorphism $\psi\colon U \to V$ applies to the functions by

$$\begin{aligned} \psi\colon \operatorname{Fun}(U) &\to \operatorname{Fun}(V) \\ u &\mapsto u\circ\psi^{-1}. \end{aligned} \tag{61.60}$$

EQooMZSEooATfSTR

PROPooXAOKooQSBKHg

Proposition-Definition 61.23.
If $T \in \mathscr{D}'(U)$ we define ψT by

$$\langle \psi T, \varphi\rangle_V = \langle T, J\cdot(\varphi\circ\psi)\rangle_U \tag{61.61}$$

for every $\varphi \in \mathscr{D}'(V)$.

ITEMooXHELooYhXNRs

(1) This defines a distribution $\psi T \in \mathscr{D}'(V)$.

ITEMooGDCYooVDFpuy

(2) The so defined map $\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V)$ is bijective.

ITEMooNGSJooEdRgHt

(3) The so defined map $\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V)$ is continuous.

Proof. We have to show that the map $\psi T\colon \mathscr{D}(V) \to \mathbb{C}$ is continuous or, in other words, if $\varphi_n \xrightarrow{\mathscr{D}(V)} 0$ we have to show that $\langle \psi T, \varphi_n\rangle_V \to 0$ in $\mathbb{C}$.

In terms of the seminorms we have to prove that for every $m \in \mathbb{N}$ and every compact $K \subset V$ we have

$$p_{K,m}\big(J\cdot(\varphi_n\circ\psi)\big) \to 0 \tag{61.62}$$

where J is the Jacobian of ψ and

$$p_{K,m}(f) = \sum_{|\alpha|\leqslant m}\|\partial^\alpha f\|_{K,\infty}. \tag{61.63}$$

By definition of ψT and the fact that T is continuous we have equivalence of these three facts:

$$\langle \psi T, \varphi_n\rangle_V \xrightarrow{\mathbb{C}} 0 \tag{61.64a}$$

$$\langle T, J\cdot(\varphi_n\circ\psi)\rangle_U \xrightarrow{\mathbb{C}} 0 \tag{61.64b}$$

$$J\cdot(\varphi_n\circ\psi) \xrightarrow{\mathscr{D}(U)} 0. \tag{61.64c}$$

We are going to prove the last one. The first step is to use Leibniz formula and some obvious notations like $\alpha - \beta$ where α and β are multiindex:

$$\sum_{|\alpha|\leqslant m} \|\partial^\alpha J \cdot (\varphi_n \circ \psi)\|_K = \sum_{|\alpha|\leqslant m} \| \sum_{\beta\leqslant\alpha} (\partial^\beta J) \cdot \partial^{\alpha-\beta}(\varphi_n \circ \psi)\|_K \tag{61.65}$$

Since ψ is a diffeomorphism, J is bounded and from the fact that K is compact, each of the $\partial^\gamma J$ is bounded. Thus we have

$$p_{m,K}(J \cdot (\varphi_n \circ \psi)) \leqslant \sum_{|\alpha|\leqslant m} \|c \sum_{\beta\leqslant\alpha} \partial^\beta(\varphi_n \circ \psi)\|_K \tag{61.66}$$

where the constant c depends on K and m.

We show by induction on $|\beta|$ that for every sequence (φ_n) in $\mathscr{D}(V)$ with $\varphi_n \xrightarrow{\mathscr{D}(V)} 0$ we have $\|\partial^\beta(\varphi_n \circ \psi)\|_{K,\infty} \to 0$. Starting with $|\beta| = 0$ we have

$$\||\varphi_n \circ \psi|\|_K = \|\varphi_n\|_{\psi(K)}. \tag{61.67}$$

Since $\psi(K)$ is compact (proposition 7.220), the latter is one of the seminorms on V, so that $\|\varphi_n\|_\psi(K) \to 0$. For the induction we compute

$$\begin{aligned}
\partial_i\big(\partial^\beta(\varphi_n \circ \psi)\big) &= \partial^\beta\big(\partial_i(\varphi_n \circ \psi)\big) && \text{(61.68a)}\\
&= \partial^\beta\left(x \mapsto \sum_k \frac{\partial\varphi_n}{\partial y_k}\big(\psi(x)\big)\frac{\partial\psi_k}{\partial x_i}(x)\right) && \text{(61.68b)}\\
&\leqslant c\partial^\beta\left(\sum_k \frac{\partial\varphi_n}{\partial y_k} \circ \psi\right) && \text{(61.68c)}
\end{aligned}$$

where we have done the majoration $\frac{\partial\psi_k}{\partial x_i}(x) \leqslant c$ which is legal because the function is continuous on a compact and there is a finite number of couple (k, i), so the bound can be taken uniformly with respect to x, k and i in the same time. Now the function $\frac{\partial\varphi_n}{\partial y_n}$ is an element of $\mathscr{D}(V)$ and by the induction hypothesis,

$$\|\partial^\beta\left(\sum_k \frac{\partial\varphi_n}{\partial y_k} \circ \psi\right)\|_K \to 0. \tag{61.69}$$

This finishes the proof of point (1). We pass to the points (2) and (3).

(1) $\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V)$ **is surjective**

Let $T_2 \in \mathscr{D}'(V)$. We define T_1 on $\mathscr{D}(U)$ by $T_1 = \psi^{-1}T_2$; proposition 61.23 shows that T_1 is a distribution over $\mathscr{D}(U)$. Moreover $\psi T_1 = T_2$. So ψ is surjective.

(2) $\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V)$ **is injective**

Suppose $\psi T_1 = 0$, so

$$\langle \psi T_1, \phi\rangle_V = \langle T_1, J \times (\phi \circ \psi)\rangle_U = 0 \tag{61.70}$$

for every $\phi \in \mathscr{D}(V)$. Since J is itself a C^∞ function, in fact every element $\varphi \in \mathscr{D}(U)$ can be written under the form $J \times (\phi \circ \psi)$ for some $\phi \in \mathscr{D}(V)$. We deduce $T_1 = 0$.

(3) $\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V)$ **is continuous**

Let $T_n \xrightarrow{\mathscr{D}'(U)} O$; we have to show that $\psi T_n \xrightarrow{\mathscr{D}'(V)} 0$. We use the proposition 30.32: for $\phi \in \mathscr{D}(V)$ we have to show that $(\psi T_n)(\phi) \to 0$, that is:

$$\langle \psi T_n, \phi\rangle_V = \langle T_n, J \times (\phi \circ \psi)\rangle_U. \tag{61.71}$$

Since T_n converges to zero in $\mathscr{D}'(U)$, the limit of the right hand side is zero.

□

LEMooJHFUooWdAlar

Lemma 61.24.
Let a smooth diffeomorphism $\psi\colon U \to V$ and a smooth map $u\colon U \to \mathbb{C}$. We look at the composition

$$u \circ \psi^{-1}\colon V \to \mathbb{C}. \tag{61.72}$$

We denote by y the coordinates on V and by x the ones on U. Let α be a y-multiindex. Then

$$\partial^\alpha(u \circ \psi^{-1})(y) = \sum_{|I|\leqslant|\alpha|} c_{I\alpha}(y)(\partial^I u)\big(\psi^{-1}(y)\big) \tag{61.73}$$

for some bounded functions $c_{I\alpha}$.

Written under a compact form,

$$\partial^\alpha(u \circ \psi^{-1}) = \sum_{|I|\leqslant|\alpha|} c_{I\alpha}(\partial^I u) \circ \psi^{-1}. \tag{61.74}$$

Proof. To be clear: β is a x-mutiindex. For α of length 1 we have

$$\frac{\partial}{\partial y_i}\big(u \circ \psi^{-1}\big)(y) = \sum_k \frac{\partial u}{\partial x_k}\big(\psi^{-1}(y)\big)\frac{\partial \psi_k^{-1}}{\partial y_i}(y). \tag{61.75}$$

Then applying the Leibniz rule we will have multi-derivatives of u with respect to x, always taken at the point $\psi^{-1}(y)$ and then derivatives of the components of ψ^{-1} with respect to y taken at the point y. The derivatives on u are derivatives with respect to x, and each derivative with respect to y on $u \circ \psi^{-1}$ will create derivatives on u with respect to all the components of x. Thus the sum is over the x-multiindiex of size lower or equal to the number of y-derivatives on $u \circ \psi^{-1}$. □

Proposition 61.25 ([948])**.**
Let $m \in \mathbb{N}$ and a diffeomorphism $\psi\colon U \to V$ such that all the derivatives of ψ and ψ^{-1} are bounded. We consider the map

$$\begin{aligned}\psi\colon \mathrm{Fun}(U) &\to \mathrm{Fun}(V)\\ u &\mapsto u \circ \psi^{-1}.\end{aligned} \tag{61.76}$$

ITEMooNJZOooOrzQIT

(1) We have

$$\psi\colon L^2(U) \to L^2(V) \tag{61.77}$$

surjectively.

(2) We have the diffeomorphisms

$$\psi\colon H^m(U) \to H^m(V) \tag{61.78a}$$
$$\psi\colon C_0^\infty(U) \to C_0^\infty(V) \tag{61.78b}$$
$$\psi\colon \mathscr{D}'(U) \to \mathscr{D}'(V). \tag{61.78c}$$

The last one is defined by

$$\langle \psi T, \varphi\rangle_V = \langle T, J\ \cdot (\varphi \circ \psi)\rangle_U. \tag{61.79}$$

Proof. The hypothesis on ψ implies in particular that $d\psi$ and $d\psi^{-1}$ are bounded.

(1) $\psi\colon L^2(U) \to L^2(V)$ **is surjective** We consider the operation ψ on the functions as explained in equation (61.60). We know that $x \mapsto |f(x)|^2$ is integrable on U. Thus the change of variable theorem 14.281(3) the function $|f|^2 \circ \psi^{-1} \times |J|$ is integrable on V. That is the function $y \mapsto |f\big(\psi^{-1}(y)\big)|^2|J_\psi(y)|$. Recall that

$$J_\psi(a) = \det(d\psi_a), \tag{61.80}$$

and that by hypothesis, this is bounded. Also by hypothesis, the inverse is bounded. If c is a majoration of $|1/J|$, then

$$|f \circ \psi|^2 = |f \circ \psi|^2 \times |J| \times |1/J| \leqslant c|f \circ \psi|^2 \times |J|. \tag{61.81}$$

The right hand side is integrable, so we deduce that $|f \circ \psi|^2$ is integrable. At the end of the day, the function $f \in L^2(V)$ is the image by ψ of $f \circ \psi^{-1}$ which belongs to $L^2(U)$.

(2) $\psi\colon \mathscr{D}(U)\to\mathscr{D}(V)$ **is bijective and continuous** The bijection is the fact that ψ itself is a bijection of class C^∞, and that it transforms a compact into a compact. The continuity is correct for topology of the norm $\|.\|_\infty$ because ψ turns out to be an isometry.

(3) $\psi\colon \mathscr{D}'(U)\to\mathscr{D}'(V)$ **is bijective and continuous** That is the proposition 61.23.

(4) $\psi\colon H^m(U)\to H^m(V)$
We denote by y the coordinates on V and by x the ones on U. Let $u\in H^m(U)$, let $\phi\in\mathscr{D}(V)$ and a y-multiindex α with $|\alpha|\leqslant m$. By definition of the weak derivative (and the fact that ϕ is compactly supported, so that it vanishes at the boundaries of V) we have

$$\begin{aligned}
\langle u\circ\psi^{-1},\partial^\alpha\phi\rangle_V &= \langle u,|J_\psi|\partial^\alpha\phi\rangle_V && \text{(61.82a)}\\
&= \langle u,\sum_{|I|\leqslant|\alpha|}c_{\alpha I}\partial^I\tilde{\phi}\rangle_U && \text{(61.82b)}\\
&= \sum_{|I|\leqslant|\alpha|}\langle u,c_{\alpha I}\partial^I\tilde{\phi}\rangle_U && \text{(61.82c)}\\
&= \sum_{|I|\leqslant|\alpha|}\langle \bar{c}_{\alpha I}u,\partial^I\tilde{\phi}\rangle_U && \text{(61.82d)}\\
&= \sum_{|I|\leqslant|\alpha|}\langle \partial^I(c_{\alpha I}u),\tilde{\phi}\rangle_U && \text{(61.82e)}
\end{aligned}$$

Justifications:

— For (61.82a): the change of variable formula under the form (14.887).

— For (61.82b), we wrote $\tilde{\phi}=\phi\circ\psi$ and used the lemma 61.24 with $\partial^\alpha(\tilde{\psi}\circ\psi^{-1})$. We have included the function $|J|$ in the coefficients $c_{\alpha I}$.

— **-NoValue-** For (61.82e) we redefined $c_{\alpha I}$ in order to include $(-1)^{|I|}$ and the complex conjugation.

Up to now we have

$$\langle u\circ\psi^{-1},\partial^\alpha\phi\rangle_V=\sum_{|I|\leqslant|\alpha|}\langle\partial^I(c_{\alpha I}u),\tilde{\phi}\rangle_U \tag{61.83}$$

where $\tilde{\phi}=\phi\circ\psi\in\mathscr{D}(U)$ and the coefficients (and all the derivatives) $c_{\alpha I}$ are smooth and bounded because they are combinations of derivatives of the ψ's components.
We can again use the Leibniz rule on the derivative $\partial^I(c_{\alpha I}u)$. We will get all the derivatives $\partial^J u$ (J are sub-indices of I) with coefficients that are again derivatives of $c_{\alpha I}$. Thus we write

$$\partial^I(c_{\alpha I}u)=\sum_{J\leqslant I}c'_{\alpha J}\partial^J u, \tag{61.84}$$

and since J takes any value contained in I and I every values with length lower than $|\alpha|$ we can as well write

$$\langle u\circ\psi^{-1},\partial^\alpha\phi\rangle_V=\sum_{|I|\leqslant|\alpha|}\langle\partial^I(c_{\alpha I}u),\tilde{\phi}\rangle_U=\sum_{|I|\leqslant|\alpha|}\langle c'_{\alpha I}\partial^I u,\tilde{\phi}\rangle_U. \tag{61.85}$$

We have $|I|\leqslant|\alpha|\leqslant m$, so $\partial^I u\in L^2(U)$. Since $c'_{\alpha I}$ is bounded, the product $c'_{\alpha I}\partial^I u$ belongs to $L^2(U)$ too. The integral change of variable formula (14.887) reads in our case:

$$\langle f,g\rangle_U=\langle f\circ\psi^{-1},(g\circ\psi^{-1})|J^{-1}|\rangle_V, \tag{61.86}$$

with $\tilde{\phi}$ in the role of g:

$$\langle f,\tilde{g}\rangle=\langle|J^{-1}|(f\circ\psi^{-1}),\phi\rangle_V \tag{61.87}$$

Now with $\sum_{|I|\leqslant|\alpha|}c'_{\alpha I}\partial^I u$ in the role of f we have $|J^{-1}|(f\circ\psi^{-1})\in L^2(V)$ because of point (1) and the fact that $|J^{-1}|$ is bounded, so that the product makes sense and we have shown that the function

$$f'=|J^{-1}|\big(\sum_{|I|\leqslant|\alpha|}c'_{\alpha I}\partial^I u\big)\circ\psi^{-1} \tag{61.88}$$

is an element of $L^2(V)$ satisfying

$$\langle u\circ\psi^{-1},\partial^\alpha\phi\rangle_V=\langle f',\phi\rangle_V \tag{61.89}$$

for every $\phi\in\mathscr{D}(V)$. This condition is the definition of $(-1)^\alpha\partial^\alpha(u\circ\psi^{-1})$. Thus we have

$$\partial^\alpha(u\circ\psi^{-1})=(-1)^\alpha f'\in L^2(V). \tag{61.90}$$

Notice that the fact that $\partial^\alpha(u\circ\psi^{-1})$ exists is not a big deal: it always weakly exists. The very point we shown is that the derivative is $L^2(V)$.
Since α is amultiindex of any size up to m we shown that $u\circ\psi^{-1}\in H^m(V)$.

(5) $\psi\colon H^m(U)\to H^m(V)$ **is continuous**
Let $u\in H^m(U)$ and a smooth diffeomorphism $\psi\colon U\to V$; we consider $\psi u=u\psi^{-1}$ and we want to prove an estimation like $\|\psi u\|_{H^m(V)}\leqslant C\|u\|_{H^m(U)}$. We use the proposition 10.43 with $\lambda=1$, $V=L^2(V)$ and $A=\mathscr{D}(V)$:

$$\|\partial^\alpha(\psi u)\|_{L^2(V)}=\sup\{|\langle\partial^\alpha(\psi u),\phi\rangle_V|\text{ such that }\phi\in\mathscr{D}(V)\text{ and }\|\phi\|_{L^2(V)}\leqslant 1\}. \tag{61.91}$$

Using lemme 61.24 we have

$$\clubsuit=\langle\partial^\alpha(\psi u),\phi\rangle_V=\sum_{|I|\leqslant|\alpha|}\langle c_{\alpha I}(\partial^I u)\circ\psi^{-1},\phi\rangle_V\leqslant C\sum_{|I|\leqslant|\alpha|}\langle(\partial^I u)\circ\psi^{-1},\phi\rangle_V. \tag{61.92}$$

The first majoration is $C=\sup_{I,x}c_{\alpha I}(x)$. In the next lines, we will often modify C in order to include other majorations. With the change of variable formula (14.887) and including a majoration of $|J_\psi|$ in the constant C,

$$\clubsuit=\sum_{|I|\leqslant|\alpha|}\langle\partial^I u,(\phi\circ\psi)|J_\psi|\rangle_U\leqslant C\sum_{|I|\leqslant|\alpha|}\langle\partial^I u,\phi\circ\psi\rangle_U. \tag{61.93}$$

Now, again using the change of variable formula (but in the reverse sense) we have

$$\|\phi\circ\psi\|^2_{L^2(u)}=\langle\phi\circ\psi,\phi\circ\psi\rangle_U=\langle\phi,\phi|J^{-1}|\rangle_V\leqslant K\|\phi\|^2_{L^2(V)} \tag{61.94}$$

where K is a majoration of $|J^{-1}|$. We can continue the computation:

$$\begin{aligned}\|\partial^\alpha(\psi u)\|_{L^2(V)}&=\sup\{|\clubsuit|\text{ such that }\phi\in\mathscr{D}(V)\text{ and }\|\phi\|_{L^2(V)}\leqslant 1\} &(61.95a)\\
&\leqslant\sup\{C|\sum_{|I|\leqslant|\alpha|}\langle\partial^I u,\phi\circ\psi\rangle_U|\text{ such that }\phi\in\mathscr{D}(V)\text{ and }\|\phi\|_{L^2(V)}\leqslant 1\} &(61.95b)\\
&\leqslant C\sum_{|I|\leqslant|\alpha|}\sup\{|\langle\partial^I u,\varphi\rangle_U|\text{ such that }\varphi\in\mathscr{D}(U)\text{ and }\|\varphi\|_{L^2(U)}\leqslant K\} &(61.95c)\\
&\leqslant C\sum_{|I|\leqslant|\alpha|}\|\partial^I u\|_{L^2(U)}. &(61.95d)\end{aligned}$$

We have:

$$\|\partial^\alpha(\psi u)\|_{L^2(V)}\leqslant C\sum_{|I|\leqslant|\alpha|}\|\partial^I u\|_{L^2(U)}\leqslant C\|u\|_{H^m(U)} \tag{61.96}$$

because of the expression (31.46) of the norm on $H^m(U)$. Now we can conclude:

$$\|\psi u\|_{H^m(V)}=\sum_{|\alpha|\leqslant m}\|\partial^\alpha(\psi u)\|_{L^2(V)}\leqslant C\sum_{|\alpha|\leqslant m}\|u\|_{H^m(V)}=aC\|u\|_{H^m(U)} \tag{61.97}$$

where a is the number of terms in the sum.

(6) $\psi\colon H^m(U) \to H^m(V)$ **is a diffeomorphism**

Replacing ψ by ψ^{-1} shows that $\psi^{-1}\colon H^m(V) \to H^m(U)$ is continuous. Don't forget to show that this is actually the inverse:

$$\psi^{-1}(\psi u) = (\psi u) \circ \psi^{-1} = u \circ \psi \circ \psi^{-1} = u. \tag{61.98}$$

Notice that in the later equation, the symbol "ψ" denotes sometimes the map $\psi\colon U \to V$ and sometimes the map $\psi\colon \mathrm{Fun}(U) \to \mathrm{Fun}(V)$.

□

Remarque 61.26.
The hypothesis on the fact that the derivatives of ψ are bounded is not a weak one. As an example, the map

$$\begin{aligned} \psi\colon]0,1[&\to]0,1[\\ x &\mapsto x^2 \end{aligned} \tag{61.99}$$

is a diffeomorphism with Jacobian given by $J_\psi(x) = 2x$, so that $1/J$ is not bounded on $]0,1[$.

61.6 Restriction operator

Definition 61.27.
Let $j = 1, \dots, d$ and $h \in \mathbb{R}$. We define the ***translation operator***

$$\tau_{j,h}\colon \mathrm{Fun}(\mathbb{R}^d) \to \mathrm{Fun}(\mathbb{R}^d) \tag{61.100}$$

given by the formula

$$\tau_{j,h}(f)(x) = f(x - he_j) \tag{61.101}$$

where e_j is the canonical unit vector.

The adjoint operator (definition 60.11) $\tau_{j,h}^*\colon L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is

$$(\tau_{j,h}^* u)(x) = u(x + he_j). \tag{61.102}$$

Indeed we have

$$\langle \tau(u), v\rangle_{L^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \overline{u(x - he_j)} v(x) dx = \int_{\mathbb{R}^d} \overline{u(y)} v(y + he_j) dy = \langle u, \tau_{j,h}^* v\rangle_{L^2(\mathbb{R}^d)} \tag{61.103}$$

by a simple change of variable $y = x - he_j$. We immediately see that $\tau_{j,h}$ is an unitary operator of $L^2(\mathbb{R}^d)$.

If Ω is an open part of $\mathbb{R}^d$ we define the restriction operator r_Ω by

$$r_\Omega = u|_\Omega. \tag{61.104}$$

We define

$$\mathbb{R}^d_+ = \{x \in \mathbb{R}^d \text{ such that } x_n > 0\}, \tag{61.105}$$

and the same for $\mathbb{R}^d_-$. We denote by $r^\pm$ the restriction operator to $\mathbb{R}^d_\pm$. The extension operator is the extension by 0 outside Ω:

$$(e_\Omega f)(x) = \begin{cases} f(x) & \text{if } x \in \Omega \\ 0 & \text{else.} \end{cases} \tag{61.106}$$

61.28.
Do you believe that there is a difference between $\mathscr{D}\big(\overline{\mathbb{R}^d_+}\big)$ and $\mathscr{D}(\mathbb{R}^d_+)$? Let us see with $d = 1$. Since the support is closed, a function in $\mathscr{D}(\mathbb{R}_+)$ has to be zero on $]-\infty, \delta]$ with $\delta > 0$.

On the other hand, a function in $\mathscr{D}(\overline{\mathbb{R}^d_+})$ can be nonzero immediately after $x = 0$:

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0,1] \\ \text{something} & \text{if } x \in [1,2] \\ 0 & \text{if } x > 2. \end{cases} \tag{61.107}$$

where "something" stands for something smooth.

PROPooCXYRooPTgSLX

Proposition 61.29 ([948]).
Some densities.

(1) The set $\mathscr{D}(\mathbb{R}^d)$ is dense in $H^m(\mathbb{R}^d)$ for every integer $m \geqslant 0$.

(2) The set $r^+\mathscr{D}(\mathbb{R}^d)$ is dense in $H^m(\mathbb{R}^d_+)$ for every $m \geqslant 0$.

Theorem 61.30 ([948]).
There exists a linear continuous operator

$$p_{(m)} \colon H^m(\mathbb{R}^d_+) \to H^m(\mathbb{R}^d) \tag{61.108}$$

such that

$$(r^+ \circ p_{(m)})u = u \tag{61.109}$$

for every $u \in H^m(\mathbb{R}^d_+)$.

Proof. The map r^+ a defined as $r^+ \colon \operatorname{Fun}(\mathbb{R}^d) \to \operatorname{Fun}(\mathbb{R}^d)$. Since the norms $L^2(\mathbb{R}^d_+)$ are majored by the ones $L^2(\mathbb{R}^d)$ we have

$$r^+ \colon H^m(\mathbb{R}^d) \to H^m(\mathbb{R}^d_+). \tag{61.110}$$

We initiate by defining the operator $p_{(m)}$ on a dense subset (see proposition 61.29). Let $u \in r^+\mathscr{D}(\mathbb{R}^d)$. More precisely we choose $u_0 \in \mathscr{D}(\mathbb{R}^d)$ and we write $u = r^+u_0$. We define

$$(p_{(m)}u)(x', x_d) = \begin{cases} u(x', x_d) & \text{if } x_d > 0 \\ u_0(x', 0) & \text{if } x_d = 0 \\ \sum_{k=0}^{m-1} a_k u_0(x', -\lambda_k x_d) & \text{if } x_d < 0. \end{cases} \tag{61.111}$$

Here are some comments ont the definition.

— The value of $p_{(m)}u$ does not depend on the choice of u_0 among all the elements in $\mathscr{D}(\mathbb{R}^d)$ which restrict to u.

— The positivity of λ_k makes that $-\lambda_k x_d > 0$, so that the whole makes sense.

— The numbers $\lambda_0, \dots, \lambda_{m-1}$ are strictly positive different reals that are arbitrarily chosen.

— The coefficients a_k are the solution of the system

$$\sum_{k=0}^{m-1} \lambda_k^j a_k = (-1)^j \tag{61.112}$$

EQooTHYTooUoyXKT

for $j = 0, \dots, m-1$.

— The system (61.112) has an unique solution because its matrix is $A_{kj} = \lambda_k^j$ whose determinant is the non vanishing Vandermonde determinant 8.18.

Let us show that $p_{(m)}(u) \in C^{m-1}(\mathbb{R}^d)$. The derivatives of $p_{(m)}u$ with respect to the variables $x_1, \dots, x_{d-1}$ do not create problems. Only the ones with respect to x_d at $x_d = 0$ can be difficult. Now we show by induction that

$$(\partial_n^l p_{(m)}u)(x', 0) = (\partial_n^l u_0)(x', 0) \tag{61.113}$$

with several steps.

(1) **First:** $(\partial_n p_{(m)}u)(x',0) = (\partial_n u_0)(x',0)$

On the one hand we have

$$\lim_{t\to 0^+} \frac{(p_{(m)}u)(x',t) - (p_(m)u)(x',0)}{t} = \lim_{t\to 0^+} \frac{u_0(x',t) - u_0(x',0)}{t} = (\partial_n u_0)(x',0). \tag{61.114}$$

The last equality comes from the fact that for u_0 we know that the limits with $t \to 0^+$ and $t \to 0^-$ are equal. On the other hand, we know that $\sum_{k=0}^{m-1} a_k = 1$, so that

$$\lim_{t\to 0^-} \frac{(p_{(m)}u)(x',t) - (p_(m)u)(x',0)}{t} = \lim_{t\to 0^+} \frac{\sum_{k=0}^{m-1} a_k u_0(x',-\lambda_k t) - u_0(x',0)}{t} \tag{61.115a}$$

$$= \sum_{k=0}^{m-1} a_k \lim_{t\to 0^-} \frac{u_0(x',-\lambda_k t) - u_0(x',0)}{t} \tag{61.115b}$$

$$= -\sum_{k=0}^{m-1} a_k \lambda_k (\partial_n u_0)(x',0). \tag{61.115c}$$

We made the change of variable $t' = -\lambda_k t$ to compute the limit.

(2) **Second:** $(\partial_n^l p_{(m)}u)(x',x_d)$ **with** $x_d < 0$

In this case the computation is immediate:

$$(\partial_n^l p_{(m)}u)(x',x_d) = \sum_{k=0}^{m-1} a_k (-\lambda_k)^l (\partial_n^l u_0)(x',-\lambda_k x_d). \tag{61.116}$$

(3) **Induction**

We are going to prove by induction over l that

$$(\partial_n^l p_{(m)}u)(x',0) = (\partial_n^l u_0)(x',0). \tag{61.117}$$ EQooVJITooNBgryX

The case $l = 1$ is already done. So we compute with $l+1$ assuming that (61.117) holds. Once again we have to compute separately the limit with $t \to 0^+$ and $t \to 0^-$. We have

$$\lim_{t\to 0^+} \frac{(\partial_n^l p_{(m)}u)(x',t) - (\partial_n^l p_{(m)}u)(x',0)}{t} = \lim_{t\to 0^+} \frac{(\partial_n^l u_0)(x',t) - (\partial_n^l u_0)(x',0)}{t} \tag{61.118a}$$

$$= (\partial_n^{l+1} u_0)(x',0). \tag{61.118b}$$

Now, the limit with negative t. To be computed:

$$\lim_{t\to 0^-} \frac{(\partial_n^l p_{(m)}u)(x',t) - (\partial_n^l u_0)(x',0)}{t}. \tag{61.119}$$ EQooFKHJooPqYWcF

The numerator can be expanded as

$$(-1)^l \sum_{k=0}^{m-1} a_k \lambda_k^l (\partial_n^l u_0)(x',-\lambda_k t) - (\partial_n^l u_0)(x',0). \tag{61.120}$$

Since $(-1)^l \sum_{k=0}^{m-1} a_k \lambda_k^l = 1$ we can multiply the second term by that and enter $(\partial_n^l u_0)(x',0)$ in the sum. What we get is

$$(-1)^l \sum_{k=0}^{m-1} a_k \lambda_k^l \big((\partial_n^l u_0)(x',-\lambda_k t) - (\partial_n^l u_0)(x',0)\big). \tag{61.121}$$

Back to the limit (61.119) we permute the limit and the sum:

$$\lim_{t\to 0^-}\frac{(\partial_n^l p_{(m)}u)(x',t)-(\partial_n^l u_0)(x',0)}{t}=(-1)^l\sum_{k=0}^{m-1}\lambda_k^l\lim_{t\to 0^-}\frac{(\partial_n^l u_0)(x',-\lambda_k t)-(\partial_n^l u_0)(x',0)}{t} \tag{61.122a}$$

$$=(-1)^l\sum_{k=0}^{m-1}\lambda_k^l\lim_{t\to 0^+}\frac{(\partial_n^l u_0)(x',u)-(\partial_n^l u_0)(x',0)}{-u/\lambda_k} \tag{61.122b}$$

$$=(-1)^{l+1}\sum_{k=0}^{m+1}\lambda_k^{l+1}(\partial_n^{l+1}u_0)(x',0) \tag{61.122c}$$

$$=(\partial_n^{l+1}u_0)(x',0). \tag{61.122d}$$

The induction is finished.

We are now able to prove that $p_{(m)}u$ is $C^{m-1}(\mathbb{R}^d)$. The only point is to show that $(\partial_n^l u))$ is continuous at $x_d=0$. The limit with $x_d>0$ does not pone any problem. For $x_d<0$,

$$\lim_{x_d\to 0^-}(\partial_n^l p_{(m)}u)(x',x_d)=\sum_{k=0}^{m-1}a_k(-\lambda_k)^l\lim_{x_d\to 0^-}(\partial_n^l u_0)(x',-\lambda_k x_d) \tag{61.123a}$$

$$=\sum_{k=0}^{m-1}a_k(-\lambda_k)^l(\partial_n^l u_0)(x',0) \tag{61.123b}$$

$$=(\partial_n^l u_0)(x',0) \tag{61.123c}$$

$$=(\partial_n^l p_{(m)}u)(x',0). \tag{61.123d}$$

When one tries to derive more than $m-1$ times, we have a possible discontinuity at $x_d=0$ since we do not have any formula for the sum $\sum_{k=0}^{m-1}k\lambda_k^m$.

The function $p_{(m)}u$ is compactly supported and the derivatives up to the $m-1^{\text{e}}$are continuous. The m^{e}derivatives is continuous on $x_d>0$ and $x_d<0$. The whole makes

$$p_{(m)}u\in H^m(\mathbb{R}^d). \tag{61.124}$$

We still have to prove that $p_{(m)}\colon r^+\mathscr{D}(\mathbb{R}^d)\to H^m(\mathbb{R}^d)$ is continuous and extends to $H^m(\mathbb{R}^d_+)$. We start proving the inequality

$$\|p_{(m)}u\|_{H^m(\mathbb{R}^d)}\leqslant C\|u\|_{H^m(\mathbb{R}^d_+)} \tag{61.125}$$

for some constant C that depend on the λ_i's. By the definition 31.46 we know $\|v\|_{H^m(\mathbb{R}^d)}=\sum_{|\alpha|\leqslant m}\|\partial^\alpha v\|_{L^2(\mathbb{R}^d)}$ and what happens on the hyperplane $x_d=0$ is of no interest. Let's consider a multiindex α containing l times the index x_d; of $x_d<0$ we have

$$(\partial^\alpha p_{(m)}u)(x',x_d)=\sum_{k=0}^{m-1}(-\lambda_k)^l a_k(\partial^\alpha u)(x',-\lambda_k x_d). \tag{61.126}$$

There is a finite number of combinations λ_k^l with $k=0,\dots,d-1$ and $l=0,\dots,m$ (and they are all strictly positive). We majore them all by K. Taking the maximum of the numbers $|a_k|$ and including it in the K we have

$$\big|(\partial^\alpha p_{(m)}u)(x',x_d)\big|\leqslant K\sum_{k=0}^{m-1}|(\partial^\alpha u)(x',-\lambda_k x_d)|. \tag{61.127}$$

As far as the $L^2(\mathbb{R}^d)$ norm is concerned we have

$$\|\partial^\alpha p_{(m)}u\|_{L^2(\mathbb{R}^d_-)} \leqslant K\sum_{k=0}^{m-1}\int_{\mathbb{R}^d_-}\left|(\partial^\alpha u)(x',-\lambda_k x_d)\right|^2 dx'\otimes dx_d \tag{61.128a}$$

$$= K\sum_{k=0}^{m-1}\int_{\mathbb{R}^d_+}\lambda_k\left|(\partial^\alpha u)(x',y)\right|^2 dx'\otimes dy \tag{61.128b}$$

EQooBTCBooHejadM

$$= K\sum_{k=0}^{m-1}\lambda_k\|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)} \tag{61.128c}$$

$$\leqslant K\|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)}. \tag{61.128d}$$

We made the change of variable $y=-\lambda_k x_d$ (this is why a λ_k appeared as Jacobian in (61.128b)). In the last line we included the factor $\sum_{k=0}^{m-1}\lambda_k$ into K.

From the definition we also have

$$\|\partial^\alpha p_{(m)}u\|_{L^2(\mathbb{R}^d_+)} = \|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)}. \tag{61.129}$$

Making the sum (and neglecting the integral over $x_d=0$) and redefining K as $\max\{1,K\}$,

$$\|\partial^\alpha p_{(m)}u\|_{L^2(\mathbb{R}^d)} \leqslant K\|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)} + \|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)} \tag{61.130a}$$

$$= 2K\|\partial^\alpha u\|_{L^2(\mathbb{R}^d_+)}. \tag{61.130b}$$

Taking once again the sum over the $|\alpha|\leqslant m$ we have the majoration

$$\|p_{(m)}u\|_{H^m(\mathbb{R}^d)} \leqslant C\|u\|_{H^m(\mathbb{R}^d_+)}. \tag{61.131}$$

We have continuity of the map

$$p_{(m)}\colon \left(r^+\mathscr{D}(\mathbb{R}^d), \|.\|_{H^m(\mathbb{R}^d_+)}\right) \to \left(H^m(\mathbb{R}^d), \|.\|_{H^m(\mathbb{R}^d)}\right). \tag{61.132}$$

Taking advantage of the fact that the Sobolev space are complete (thus Banach) and the density of $r^+H^m(\mathbb{R}^d_+)$ in $H^m(\mathbb{R}^d_+)$ we use the proposition 17.127 to build a continuous extension

$$p_{(m)}\colon \left(H^m(\mathbb{R}^d_+), \|.\|_{H^m(\mathbb{R}^d_+)}\right) \to \left(H^m(\mathbb{R}^d), \|.\|_{H^m(\mathbb{R}^d)}\right). \tag{61.133}$$

□

61.7 Boundaries

We already made a small step in the world of boundaries and Sobolev space defining $H^1_0(\Omega)$ as the part of $H^1(\Omega)$ made from functions that vanishes on $\partial\Omega$ (definition 61.18). This is far from being the end of the story. The main point in studying Sobolev spaces and boundaries is that the boundary has zero measure, so that the definitions of the usual functional spaces make no sense.

We need an integration theory on the boundary. Let $a>0$ (in $\mathbb{R}$) and define

$$Q_a = \{x\in\mathbb{R}^d \text{ such that } |x_j|<a,\ \forall j\} \tag{61.134a}$$

$$Q'_a = \{x'\in\mathbb{R}^{d-1} \text{ such that } |x'_j|<a,\ \forall j=1,\dots,d-1\} \tag{61.134b}$$

$$Q_a^+ = \{x\in Q_a \text{ such that } |x_n|>0\}. \tag{61.134c}$$

DEFooCDJTooYyibCc

Definition 61.31.
*A **smooth diffeomorphism** $\varphi\colon A\to B$ is an invertible C^∞-map with C^∞ inverse.*

Definition 61.32 ([948]).
*A open set $\Omega \subset \mathbb{R}^d$ is **smooth** if for every $x_0 \in \partial\Omega$ there exists a neighbourhood U of x_0, a $a > 0$ and a smooth diffeomorphism*[10] *$\varphi\colon Q_a \to U$ such that*

(1) $\varphi(Q_a^+) = U \cap \Omega$

(2) $\varphi(Q_a') = U \cap \partial\Omega$

(3) $\varphi(0) = x_0$.

The restriction $\varphi\colon Q_a' \to U \cap \partial\Omega$ will be denoted by λ.

Definition 61.33.
*A triple (φ, U, Q_a) is a **local coordinates system** for Ω around x_0.*

Lemma 61.34.
Let (φ_1, U_1, Q_a) and (φ_2, U_2, Q_a) be local coordinates system around x_1 and x_2 in $\partial\Omega$. The map

$$\varphi_2^{-1} \circ \varphi_1 \colon \varphi_1^{-1}(U_1 \cap U_2) \to \varphi_2^{-1}(U_1 \cap U_2) \tag{61.135}$$

is a smooth diffeomorphism such that if $y \in Q_a$ is an element of $\varphi_1^{-1}(U_1 \cap U_2)$ with $y_n = 0$ we have

$$(\varphi_2^{-1} \circ \varphi_1)(y)_n = 0. \tag{61.136}$$

Proof. The fact that $\varphi_2^{-1} \circ \varphi_1$ is a smooth diffeomorphism is true by composition, and by the careful choice of the domains.

If $y_n = 0$, we have $y \in Q_a'$ and $\varphi_1(y) \in U_1 \cap \partial\Omega$. In the same time, from the hypothesis, $\varphi_1(y) \in U_1 \cap U_2$, so that

$$\varphi_1(y) \in U_1 \cap U_2 \cap \partial\Omega. \tag{61.137}$$

In particular, $\varphi_2^{-1}\big(\varphi_1(y)\big) \in Q_a'$ and $(\varphi_2^{-1} \circ \varphi_1)(y)_n = 0$. □

One can restate the lemma saying that $\varphi_2^{-1} \circ \varphi_1$ is a smooth diffeomorphism preserving the property $y_n = 0$.

61.35.
We suppose that Ω is bounded, so that $\partial\Omega$ is bounded and closed, thus compact (Borel-Lebesgue, theorem 9.24). If one has a local coordinate system (φ_x, U_x, Q_a) around each $x \in \partial\Omega$, the union $\bigcup_{x\in\partial\Omega} U_x$ contains $\partial\Omega$. There exists a finite subcovering.

We conclude that when Ω is bounded, we have a finite family of local coordinates (φ_j, U_j, Q_a) $(j = 1, \dots, \dots, N)$. We also write $U_0 = \Omega$ so that $\bigcup_{j=0}^N U_j$ covers $\bar\Omega$.

As allowed by the corollary 48.66, we consider a partition of the unity for $\bar\Omega$ subordinated to the open sets $\{U_j\}_{j=0,\dots,N}$. For each $j = 0, \dots, N$ we have $\psi_j \in \mathscr{D}(U_j)$ and we have

$$\sum_{j=0}^N \psi_j = 1 \tag{61.138}$$

on a neighbourhood of $\bar\Omega$.

If $u \in Fun(\Omega)$ we have

$$u = \sum_{j=0}^N \psi_j u \tag{61.139}$$

and if $v \in \mathrm{Fun}(\partial\Omega)$ we have

$$b = \sum_{j=1}^N \psi_j v. \tag{61.140}$$

10. Définition 61.31.

The function $\psi_j u\colon U_j \to \mathbb{C}$ can be send to a function on Q_a^+ by a chart because the support of $\psi_j u$ is contained in the interior of U_j while $\varphi_j^{-1}(U_j \cap \Omega) = Q_a^+$.

The diffeomorphism $\varphi_j\colon Q_a \to U_j$ induces a diffeomorphism

$$\lambda_j\colon Q'_a \to U_j \cap \partial\Omega \tag{61.141}$$

This is because by definition of a chart, $\varphi_j\colon U_j \cap \partial\Omega \to Q'_a$ is a bijection.

The following defines $L^p_{loc}(\partial\Omega)$ with respect to a chart system. So this should be written $L^p_{loc}(\partial\Omega, \{\varphi_j\})$.

DEFooFBVPooEeNwuU

Definition 61.36.
Let $p \geqslant 1$. The space $L^p_{loc}(\partial\Omega)$ is the set of functions $u \in \mathrm{Fun}(\partial\Omega)$ such that the map

$$u \circ \lambda_j\colon Q'_a \to \mathbb{C} \tag{61.142}$$

belongs to $L^p_{loc}(Q'_a)$ for every j.

We prove that the set $L^p_{loc}(\partial\Omega, \{\varphi_j\})$ does not depend on $\{\varphi_j\}$.

Lemma 61.37.
The set defined in 61.36 does not depend on the choice of chart system.

Proof. We consider two coordinates charts $(U_j, \varphi_j, Q_a)_{j\in J}$ and $(V_\alpha, \psi_\alpha, Q_a a_{\alpha\in A})$ and we have to show that when u is a L^p_{loc} function with respect to one chart system, it is L^p_{loc} for the other one. We consider the restrictions

$$\lambda_j\colon Q'_a \to U_j \cap \partial\Omega \tag{61.143a}$$
$$\sigma_\alpha\colon Q'_a \to V_\alpha \cap \partial\Omega. \tag{61.143b}$$

So we suppose that $u \circ \lambda_j \in L^p_{loc(Q'_a)}$ for every j and we have to show that $u \circ \sigma_\alpha \in L^p_{loc}(Q'_a)$ for every $\alpha \in A$ (from now on we fix some $\alpha \in A$). Let K be compact un Q'_a and we have to show that $u \circ \sigma_\alpha \in L^p(K)$.

Let $L_j = U_j \cap \sigma_\alpha(K)$; this is a compact part of $\mathbb{R}^d$ and we have $\sigma_\alpha(K) = \bigcup_{j\in J} L_j$. The latter union is however not disjoint. We have

$$\int_K |u \circ \sigma_\alpha|^p = \int_{\sigma_\alpha} |J_{\sigma_\alpha}|^p |u|^p \tag{61.144a}$$
$$\leqslant \sum_{i\in I} \int_{L_j} \underbrace{|J_{\sigma_\alpha}|^p}_{\leqslant C} |u|^p \tag{61.144b}$$
$$\leqslant C \sum_{i\in I} \int_{\lambda_i^{-1}(L_i)} \underbrace{|J_{\lambda_i^{-1}}|}_{\leqslant C'} |u \circ \lambda_i|^p \tag{61.144c}$$

The integrals are finites because the integration domain are compacts while the integrand is $L^p_{loc}(Q'_a)$. □

61.8 Older work

SECooNJLDooFcUzQv

A lot of theory about Sobolev spaces can be found in [949, 950]. The **Fourier transform** of a function φ on $\mathbb{R}^n$ is defined by formulas

$$(F\varphi)(p) = \hat{\varphi}(p) = \int_{\mathbb{R}^N} e^{-2i\pi p \cdot x} \varphi(x)\, dx \tag{61.145a}$$
$$\varphi(x) = \int_{\mathbb{R}^N} e^{2i\pi x \cdot p} \hat{\varphi}(p)\, dp \tag{61.145b}$$

Main properties of Fourier transform are [subeq_prop_Four]

$$(\partial_j F\varphi)(p) = -2i\pi F(x_j\varphi) \tag{61.146a}$$
$$(F\partial_j\varphi) = 2i\pi p_j(F\varphi)(p) \tag{61.146b}$$

The associated formula for the delta Dirac "function" is

$$\int e^{2i\pi k\,\cdot\,x}\,dx = \delta(x). \tag{61.147}$$

Proposition 61.38.
If $\hat{\mathscr{S}}$ denote the set of functions $\hat{\varphi}$ with $\varphi \in \mathscr{S}$, then $\hat{\mathscr{S}} = \mathscr{S}$.

Proof. No proof □

We consider the **Laplace operator**, or Laplacian,

$$\Delta = \sum_{j=1}^{N} \partial_j^2 \tag{61.148}$$

When $k \in \mathbb{N}$, we consider the following norm on $\mathscr{S}(\mathbb{R}^N)$:

$$\|\varphi\|_{H^k}^2 = \int_{\mathbb{R}^N} \overline{\varphi}(x)[1-(2\pi)^{-2}\Delta]^k\varphi(x)\,dx \tag{61.149}$$ [eq_def_norm_Sob]

The **Sobolev space** $H_2^k(\mathbb{R}^N)$ is the completed of $\mathscr{S}$ for this norm.

Proposition 61.39.
These Sobolev spaces are Hilbert spaces

Proof. No proof □

In order to define Sobolev spaces H^k with $k < 0$, we have to find a definition for $[-\Delta+1]^{-l}$. We define

$$\mathscr{S}^{(l)} = \{[1-(2\pi)^{-2}\Delta]^l\varphi \text{ tel que } \varphi \in \mathscr{S}\}$$

Lemma 61.40.
For each $\psi \in \mathscr{S}^{(l)}$, there exists one and only one $\varphi \in \mathscr{S}$ such that $(-\Delta+1)^l\varphi = \psi$.

This unique function φ is naturally denoted by $(-\Delta+1)^{-l}$

Proof. We give the proof with $l = 1$, the other are induction. When $\psi \in \mathscr{S}^{(l)}$, existence is by definition true and only unicity is non trivial. Let φ_1 and φ_2 in $\mathscr{S}$ such that

$$[1-(2\pi)^{-2}\Delta]\varphi_1 = [1-(2\pi)^{-2}\Delta]\varphi_2,$$

the function $f = \varphi_1 - \varphi_2$ fulfils $(-\Delta+1)f = 0$ and thus

$$\int_{\mathbb{R}^N} \overline{f}(x)[1-(2\pi)^{-2}\Delta]f(x)\,dx = 0.$$

Since $f \to 0$ at infinity rapidly, an integration by part of the term containing Δ leads, up to some constants, to

$$\int \overline{f}f - \int \overline{f}\Delta f = \int |f|^2 + \int |\nabla f|^2 = 0.$$

This proves that $f = 0$. □

Now, the norm (61.149) can be used to define the Sobolev space H^{-l}. Elements of H^k $(k < 0)$ which are not functions are distributions. When $m \geqslant 0$, formulas (61.146) give

$$\left(F[1-(2\pi)^{-2}\Delta]^m\psi\right)(p) = \left((1+p^2)^m\hat{\psi}\right)(p). \tag{61.150}$$ [eq_undpi_spi]

Proposition 61.41.
When $\psi \in \mathscr{S}^{(m)}$, equality (61.150) holds even for $m < 0$.

Let us point out that m keep integer; the general real case will be treated later.

Proof. Let $m = -k < 0$; from definition of the space $\mathscr{S}^{(m)}$, the function $\varphi = [1 - (2\pi)^{-2}\Delta]^{-k}\psi$ exists; we have

$$\hat{\psi} = \left(F[1 - (2\pi)^{-2}\Delta]^k \varphi\right)(p) = (p^2 + 1)^k \hat{\varphi}(p), \tag{61.151}$$

therefore $\hat{\varphi}(p) = (p^2 + 1)^{-k}\hat{\psi}(p)$. Replacing φ by its definition,

$$\left(F[1 - (2\pi)^{-2}\Delta]^{-k}\psi\right)(p) = (p^2 + 1)^{-k}\hat{\psi}(p). \tag{61.152}$$

□

Proposition 61.42.
Let $\varphi \in \mathscr{S}$. There exists a $\psi \in \mathscr{S}$ such that

$$\varphi = [1 - (2\pi)^{-2}\Delta]^l \psi$$

Proof. Let us prove it with $l = 1$; other cases are obtained by iteration. We consider the function ψ defined by the condition

$$\hat{\psi}(p) = (p^2 + 1)^{-1}(F\varphi)(p).$$

For this function we have $[1 - (2\pi)^{-2}\Delta]\psi = \varphi$ □

The space $\hat{H}^s(\mathbb{R}^N)$ is the completed of $\mathscr{S}$ for the norm

$$\|\varphi\|^2_{\hat{H}^s} = \int_{\mathbb{R}^N} (p^2 + 1)^2 |\varphi(p)|^2 \, dx \tag{61.153}$$

where s is any positive real.

Theorem 61.43.
For each $\varphi \in \mathscr{S}$ and $k \in \mathbb{N}$, we have

$$\|F\varphi\|_{\hat{H}^k} = \|\varphi\|_{H^k},$$

in other words, the Fourier transform in $\mathscr{S}$ is an isometry $F\colon H^k \to \hat{H}^k$.

Proof. Using Parseval and equality (61.150),

$$\|\varphi\|^2_{H^k} = \int_{\mathbb{R}^N} \overline{\varphi}(x)[1 - (2\pi)^{-2}\Delta]^k \varphi(x) \, dx = \int \overline{F\varphi(p)}(p^2 + 1)^k (F\varphi)(p) \, dp = \|\varphi\|^2_{\hat{H}^k}. \tag{61.154}$$

□

The space $\mathscr{S}$ is dense in H^k and F is a bounded operator on $\mathscr{S}$ (because it is isometric). Therefore it can be extended to an unique isometric homomorphism $F\colon H^k \to \hat{H}^k$ which is evidently called **Fourier transform**.

In the same way, F^{-1} is extended to an inverse of F and finally,

$$F\colon H^k(\mathbb{R}^N) \to \hat{H}^k(\mathbb{R}^N)$$

is an isometric isomorphism. For each $s \in \mathbb{R}$, we define

$$\|\varphi\|_{H^s} := \|F\varphi\|_{\hat{H}^s}, \tag{61.155}$$

and the completed of $\mathscr{S}$ for this norm is the *Sobolev space* $H^s(\mathbb{R}^N)$.

Chapter 62

Analysis

62.1 Fréchet spaces and algebras

Definition 62.1.
*A **Fréchet space** is an Hausdorff locally convex space whose topology can be defined by a* countable *family of seminorm and which is complete for any distance invariant under the translations.*

Somme results without proof:

Lemma 62.2.
An Hausdorff uniformisable space whose topology can be defined by a sequence (d_n) of ecarts is metrisable.

Lemma 62.3.
An Hausdorff locally convex space whose topology can be defined by a countable *family of seminorms is metrisable.*

In particular, a Fréchet space is metrisable.

prop_suiteFk

Lemma 62.4.
Let (F_k) be a decreasing sequence of Fréchet spaces such that the canonical injections $\mathrm{Id}\colon F_{K+1} \to F_k$ are continuous. Let, for each $k \in \mathbb{N}$, be a sequence $(a_{k,q})$ of points in F_k which converges to zero in the sense of F_k. Then there exists an increasing sequence $q_k \to \infty$ such that for all $h \geqslant 0$,

$$\sum_{k \geqslant h} a_{k,q_k}$$

is a converging sum in F_k.

The following result is easy to prove:

Proposition 62.5.
If A is a Fréchet algebra for the seminorms $\|.\|$, the space $C_\flat(V, A)$ is a Fréchet algebra for the seminorms

$$\|f\|_k = \sum_{v \in V} \|f(v)\|_k$$

Definition 62.6.
*A **Fréchet algebra** is a topological algebra whose topology is given by a countable family of sub-multiplicative seminorms:*

$$\|ab\|_k \leqslant \|a\|_k \|b\|_k.$$

Proposition 62.7.
Let A be a Fréchet algebra with seminorms $\|.\|_k$ and α be an action of the vector space V on A.

We put on A^∞ the seminorms

$$\|a\|_{jk} = \frac{1}{\mu!} \sup_{i\leqslant j} \sum_{|\mu|\leqslant k} \|\partial_\mu a\|_i. \tag{62.1}$$

These seminorms are sub-multiplicative.

Proof. The element $\partial_\mu a \in A^\infty$ is defined by the action: for each $X \in V$ we set

$$\partial_X a = \frac{d}{dt}\Big[\alpha_{tX}(a)\Big]_{t=0}$$

and μ is a multi-index. The sum $\sum_{|\mu|\leqslant k} \|\partial_\mu(ab)\|_i$ is made of terms of the form

$$\|\partial^\alpha a \partial^\beta b\|_i \leqslant \|\partial^\alpha a\|_i \|\partial^\beta\|_i.$$

Evidently, $\sup_{i\leqslant j} \|\partial^\alpha a\|_i \leqslant \|a\|_{jk}$. So $\|ab\|_{jk}$ is made of $\mu!$ terms which are all lesser than $\|a\|_{jk}\|b\|_{jk}$. The factor $1/\mu!$ concludes the proof.

□

62.2 Gårding spaces

62.2.1 Strongly continuous action and smooth vector

subsec:actionrn

62.2.1.1 Strongly continuous action

Let $T\colon G \to \mathcal{B}(X)$ be a homomorphism between a topological group G and the space $\mathcal{B}(X)$ of bounded operators on a Banach space, X. In particular, T is an action of G on X. We say that T is **locally bounded** if for all $g \in G$, there exists a neighbourhood $\mathcal{U}$ of g in G and a real $K > 0$ such that $\|T(h)\| \leqslant K$ for all $h \in \mathcal{U}$. This norm is the operator norm on a Banach space 60.44. The action T is **strongly continuous** if for all $x \in X$, the map $g \to T(g)x$ is continuous for the norm topology on X.

Let V be a d-dimensional vector space with an inner-product denoted by $v \cdot w$; we consider $\{X_1, \dots X_d\}$, an orthonormal basis. Let $\{\|.\|_k\}$ be a family of seminorm which defines the topology on a Fréchet space A, and $C_\flat(V, A)$ the set of continuous bounded maps $V \to A$ on which we define a topology from the seminorms

$$\|f\|_k = \sup_{w\in V} \|f(w)\|_k. \tag{62.2}$$

eq:def_semi_n

We consider the action $\tau\colon V \times C_\flat(V, A) \to C_\flat(V, A)$ defined by translation:

$$(\tau_v f)(w) = f(v + w). \tag{62.3}$$

eq:def_act_tau

It is naturally isometric for each of the seminorms (62.2), but it is not specially a strongly continuous action. We write

$$\begin{aligned} \phi_f &\colon V \to C_\flat(V, A) \\ \phi_f(v) &= \tau_v f \end{aligned} \tag{62.4}$$

Remarque 62.8.
The Fréchet space A is not specially $\mathbb{C}$ or something like that. In fact, our intention is to consider as A a space like $C^\infty(R)$ where R is equipped with seminorms of the form $\|f\|_\mu = \sup_{r\in R} |(\partial^\mu f)(r)|$.

Proposition 62.9.
The maximal subset of $C_\flat(V, A)$ on which action τ is strongly continuous is the set $C_u(V, A)$ of uniformly continuous maps.

Proof. We begin by write down the definition of uniform continuity (48.5) of a map $f\colon V \to A$ in the present setting. Let $\mathcal{O}$ be a neighbourhood of zero in A, i.e. a choice of k and ε for which we say that $a \in \mathcal{O}$ if $\|a\|_k < \varepsilon$. Then there exists a neighbourhood $\mathcal{U}_{k\varepsilon}$ of zero in V such that

$$v_1 - v_2 \in \mathcal{U}_{k\varepsilon} \Rightarrow \|f(v_1) - f(v_2)\|_k < \varepsilon. \tag{62.5}$$

The fact for f to be uniformly continuous is the existence of such a $\mathcal{U}_{k\varepsilon}$ for all k and ε.

We begin to prove that if $v \to \tau_v f$ is continuous, then f is uniformly continuous. For this, an open set $\mathcal{R} \subset C_\flat(V, A)$ is given by a choice of k, ε and $g_0 \in C_\flat(V, A)$:

$$\mathcal{R} = \{g \in C_\flat(V, A) \text{ tel que } \|g - g_0\|_k < \varepsilon\}. \tag{62.6}$$

The condition for v to belongs to $\phi_f^{-1}(\mathcal{R})$ is $\|g_0 - \tau_v f\|_k < \varepsilon$ or

$$\sup_{w\in V} \|g_0(w) - f(v + w)\|_k < \varepsilon. \tag{62.7}$$

Let us take a particular $\mathcal{R}$: one which is taken with $g_0 = f$. A necessary condition for ϕ_f to be continuous is that the set of $v \in V$ such that

$$\sup_{w\in V} \|f(w) - f(v + w)\|_k < \varepsilon \tag{62.8}$$

eq:vdansU

is open in V. From hypothesis, we know that it is true (because ϕ_f is continuous). It is also evident that $v = 0$ belongs to this set; let is write it $\mathcal{U}_{k\varepsilon}$ and consider $v_1, v_2 \in V$ such that $v_1 - v_2 \in \mathcal{U}_{k\varepsilon}$. Then

$$\begin{aligned} \|f(v_1) - f(v_2)\|_k &= \|f(v_2) - f(v_2 + (v_1 - v_2))\|_k \\ &\leqslant \sup_{w\in V} \|f(w) - f(w + (v_1 - v_2))\|_k \\ &\leqslant \varepsilon \end{aligned} \tag{62.9}$$

because $v_1 - v_2$ fulfils (62.8).

Now let us prove that if f is uniformly continuous, then ϕ_f is continuous; or, in other words, for all choice of $g_0 \in C_\flat(V, A)$, $\varepsilon > 0$ and k, the set

$$\mathcal{A} = \left\{ v \in V \text{ tel que } \sup_{w\in V} \|g_0(w) - f(v + w)\|_k < \varepsilon \right\} \tag{62.10}$$

is open in V. For, we consider v_1 in this set and we prove that if $v_1 - v_2$ belongs to a suitably small open set around zero, then v_2 belongs to $\mathcal{A}$ too. So let v_2 be such that

$$\|f(v_1) - f(v_2)\| < \varepsilon_1$$

and let us search for an ε_1 such that $v_2 \in \mathcal{A}$:

$$\sup_{w\in V} \|g_0(w) - f(v_2 + w)\| \leqslant \sup_{w\in V} \|g_0(w) - f(v_1 + w)\| + \sup_{w\in V} \|f(v_1 + w) - f(v_1 + w_2)\|. \tag{62.11}$$

The first term is lesser than ε by construction of v_1 while the second is lesser than ε_1 from the uniform continuity of f.

□

62.2.1.2 Smooth vector of an action

Let $\alpha\colon V \times \mathfrak{A}_0 \to \mathfrak{A}_0$ be a strongly continuous action of the topological vector space V on a C^*-algebra $\mathfrak{A}_0$. The space $\mathfrak{A}^\infty$ of the **smooth vector** for the action α is the set

$$\mathfrak{A}^\infty = \{a \in \mathfrak{A}_0 \text{ tel que } x \to \alpha_x(a) \text{ is smooth}\} \tag{62.12}$$

where "smooth" means infinitely continuously derivable.

Exemple 62.10.
The set of smooth vectors of the action by translation of $\mathbb{R}$ on $C_\flat(\mathbb{R}, \mathbb{C})$: $(\tau_x f)y = f(x + y) \in \mathbb{C}$.

We define $\phi_f\colon \mathbb{R} \to C_\flat(\mathbb{R}, \mathbb{C})$ by $\phi_f(x) = \tau_x f$. The derivative is defined as on any vector space by the limit

$$\begin{aligned}(\phi_f)'(x)y &= \lim_{\varepsilon\to 0} \frac{\phi_f(x+\varepsilon) - \phi_f(x)}{\varepsilon} y \\ &= \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon+y) - f(x+y)}{\varepsilon} \\ &= f'(x+y) \\ &= \phi_{f'}(x)y.\end{aligned} \tag{62.13}$$

So $\phi'_f = \phi_{f'}$. First conclusion: f must be smooth. But we just showed that ϕ_f is continuous if and only if f is uniformly continuous. Then $\mathfrak{A}^\infty$ is the set of function from $\mathbb{R}$ to $\mathbb{C}$ whose itself and all their derivatives are uniformly continuous. In particular, these functions and all their derivatives are bounded on any bounded set (not only compact ones).

△

62.2.2 Action on Hilbert space

Let us particularise definition of C^∞ vector to action of a Lie group G on a Hilbert space H. Some proofs will be omitted. They can be found in [859].

For a map $f\colon \mathbb{R}^N \to E$, the **differential** of f at $x_0 \in \mathbb{R}^N$ is the linear map $df_{x_0}\colon \mathbb{R}^N \to E$ defined by

$$(df_{x_0})a = \frac{d}{dt}\Big[f(x_0+at)\Big]_{t=0} = \lim_{t\to 0} \frac{f(x_0+at) - f(x_0)}{t}. \tag{62.14}$$

where the $d/dt(\dots)$ is a notation for the limit in E. It is a particular case of $f\colon G \to E$ where G is a Lie group. In this case, we define

$$df_{g_0}X = \frac{d}{dt}\Big[f(X(t))\Big]_{t=0} = \lim_{t\to 0} \frac{f(X(t) - f(g_0))}{t}. \tag{62.15}$$

One can see $g \to df_g$ as a map $G \to \mathrm{End}(T_gG, E)$. The space $\mathrm{End}(T_gG, E)$ is naturally given with a structure of topological vector space. We say that f is of class C^r at g_0 if df_{g_0} is C^r. We say that f is smooth or C^∞ is it is of class r for all r.

The **partial derivative** of f in the direction of the tangent vector X at x is $(df)_x X$. It comes from the formula $df_x X = X_x(f)$ which gives the usual partial derivative when f is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$.

When T is a representation of the Lie group G on the Hilbert space H, a vector $v \in E$ is a C^∞ vector if the map $g \to T(g)v$ is C^∞ as map $G \to E$. The set of C^∞ vectors is a vector space identified by [1] $C^\infty(T)$.

A **representation** is a map $T\colon G \to \mathrm{End}(E)$ in such a way that the resulting map $G \times E \to E$ is continuous. In particular, $g \to T(g)v$ is continuous for all $v \in E$. The **strong continuity** is the fact that $g \to T(g)v$ is continuous at $g = e$ and $\|T(g)\|$ is uniformly bounded on a neighbourhood of e. A closed subspace $\mathscr{H}_1 \subseteq \mathscr{H}$ is a **subrepresentation** og T is $T(G)\mathscr{H}_1 \subseteq \mathscr{H}_1$.

From now, we denote by f_v the map

$$f_v\colon G \to E \; g \;\; \mapsto T(g)v. \tag{62.16}$$

Let $v \in C^\infty(T)$ and $g_v\colon \mathcal{G} \to E$,

$$g_v(X) = T(e^X)v.$$

As composition of two smooth map, it is a C^∞ map. We define $\varphi(X)\colon E \to E$ by

$$\varphi(X)v = (dg_v)_0(X).$$

1. C'est pas identified with?

Proposition 62.11.
Let $T\colon G \to \mathrm{End}(E)$, a representation of the Lie group G on a Hilbert space H. For each $X \in \mathcal{G}$, we define $\varphi_X\colon C^\infty(T) \to E$ by

$$\varphi_x(v) = \varphi(X)v = \lim_{t\to 0} \frac{T(e^{tX})v - v}{t}. \tag{62.17}$$

eq:250105r1

Then each φ_X leaves $C^\infty(T)$ unchanged and φ is a representation of $\mathcal{G}$ on $C^\infty(T)$. This representation naturally extends to a representation of the universal enveloping algebra $\mathcal{U}(\mathcal{G}^{\mathbb{C}})$ on $C^\infty(T)$ with $\varphi(1) = 1$.

Proof. Let $v \in C^\infty(T)$ and $f_v\colon G \to E$, $f_v(g) = T(g)v$. We apply $T(g)$ to the two sides of equation (62.17):

$$\begin{aligned} T(g)\varphi(X)v &= \lim_{t\to 0} \frac{T(ge^{tX})v - T(g)v}{t} \\ &= \lim_{t\to 0} \frac{f_f(ge^{tX}) - f_v(g)}{t} \\ &= \frac{d}{dt}\Big[f(ge^{tX})\Big]_{t=0} \\ &= \tilde{X}_g f_v \end{aligned} \tag{62.18}$$

Since $g \to f_v(g)$ is C^∞, the map $g \to \tilde{X}_g f_v$ is C^∞ too. Then $g \to T(g)\varphi(X)v$ is C^∞. This proves that $\varphi(X)v \in C^\infty(T)$. So $\varphi_X\colon C^\infty(T) \to C^\infty(T)$.

Proposition 52.20 extends $\varphi\colon \mathcal{G} \to \mathrm{End}(E)$ into a representation of $\mathcal{U}(\mathcal{G}^{\mathbb{C}})$ on $C^\infty(T)$. In order to use this proposition, we have to prove that

$$\varphi[X, Y] = \varphi(X) \circ \varphi(Y) - \varphi(Y) \circ \varphi(X).$$

A proof of this equation is given in [859].

□

Once again, a proposition without proof:

Proposition 62.12.
If U is an unitary representation of G on E, then each $\varphi(X)$ for the associated representation $\varphi\colon \mathcal{G} \to \mathrm{End}(C^\infty(T))$ is skew-hermitian.

Proposition 62.13.
Let T be a representation of G on E. Then $C^\infty(T)$ is stable for each $T(g)$. If φ is the corresponding representation of $\mathcal{U}(\mathcal{G}^{\mathbb{C}})$ on $C^\infty(E)$, then for all $D \in \mathcal{U}(\mathcal{G}^{\mathbb{C}})$ and all $g \in G$, we have

$$T(g)\varphi(D)T(g^{-1}) = \varphi(\mathrm{Ad}(g)D).$$

Proof. The map $g \to T(g)T(g_0)v$ is the composition of the (smooth) product $g \to gg_0$ in G and the smooth map $g \to T(g)v$ which are both C^∞. This proves that for all $g_0 \in G$, the vector $T(g_0)v$ belongs to $C^\infty(T)$.

Let now $X \in \mathcal{G}$, $g \in G$ and $v \in C^\infty(T)$. We have

$$\frac{T(e^{tX}) - \mathbb{1}}{t} T(g)^{-1}v = T(g^{-1}) \left(\frac{T(ge^{tX}g^{-1}) - \mathbb{1}}{t} \right) v \tag{62.19}$$

Let us consider the limit $t \to 0$ of this equation. From (62.17), the left hand side becomes $\varphi(X)T(g)^{-1}v$, while the right hand side is $T(g^{-1})\varphi(\mathrm{Ad}(g)X)v$. Then for all $X \in \mathcal{G}$,

$$T(g)\varphi(X)T(g^{-1}) = \varphi(\mathrm{Ad}(g)X). \tag{62.20}$$

This is the thesis in the particular case $D \in \mathcal{G}$. The conclusion for all $D \in \mathcal{U}(\mathcal{G})$ comes from the discussion of subsection 52.3.1:

$$T(g)\varphi(X \otimes Y)T(g^{-1}) = \varphi(\mathrm{Ad}(g)X)\varphi(\mathrm{Ad}(g)Y) = \varphi(\mathrm{Ad}(g)X \otimes \mathrm{Ad}(g)Y) = \varphi(\mathrm{Ad}(g)(X \otimes Y)). \tag{62.21}$$

□

Corollary 62.14.
If T is a representation of G on E and $D \in \mathcal{Z}(\mathcal{U}(\mathcal{G}))$ – the center of $\mathcal{U}(\mathcal{G})$–, then $\varphi(D)$ commutes with all the $T(g)$.

Proof. Lemma 52.21 make $\mathrm{Ad}(g)D = D$ and the preceding proposition concludes. □

Proposition 62.15.
Let T be a representation of the Lie group G on a Hilbert space E and an open subspace $E \subseteq E$ stable under all operators $\overline{\varphi}(X)$ defined by

$$\overline{\varphi}(X)v = \lim_{t\to 0} \frac{T(e^t X)v - v}{t}. \tag{62.22}$$ eq:ovlphilim

Then $S \subseteq C^\infty(T)$.

Proof. Let $v \in S$ and $f_v \colon G \to H$, $f_v(g) = T(g)v$. All the partial derivatives of f_v exist at $e \in G$ by hypothesis and their value are given by

$$(df_v)_e X = \overline{\varphi}(X)v \tag{62.23}$$ eq:dfeX

Let us apply $T(g)$ to both side of equation (62.22):

$$\begin{aligned} T(g)\overline{\varphi}(X)v &= \lim_{t\to 0} \frac{T(ge^{tX}g^{-1}) - \mathbb{1}}{t} T(g)v \\ &= \lim_{t\to 0} \frac{T(e^{t\,\mathrm{Ad}(g)X} - \mathbb{1}}{t} T(g)v \quad \in S \text{ by assumption} \\ &= \overline{\varphi}(\mathrm{Ad}(g)X)T(g)v. \end{aligned} \tag{62.24}$$

Equation (62.23) only gives partial derivatives of f_v at e. Let us compute, for $X_g \in T_g G$,

$$(df_v)_g X_g = \frac{d}{dt}\Big[f_v(X_g(t)) \Big]_{t=0}.$$

We parametrize $X_g(t) = ge^{tX}$, where $X \in \mathcal{G}$ is chosen in such a way that $X_g = \tilde{X}_g$:

$$\begin{aligned} (df_v)_g X_g &= \lim_{t\to 0} \frac{f_v(ge^{rX}) - f(g)}{t} \\ &= \lim_{t\to 0} \frac{T(ge^{tX})v - T(g)v}{t} \\ &= T(g)\overline{\varphi}(X)v. \end{aligned} \tag{62.25}$$

The continuity of $T(g)\colon E \to E$ justifies the inversion of the limit and $T(g)$. So we have

$$(df_v)_g \tilde{X}_g = T(g)\overline{\varphi}(X)v = \overline{\varphi}(\mathrm{Ad}(g)X)T(g)v.$$

Let us proceed by induction and suppose that for all $v \in S$, the map f_v is C^k. In particular, $f_{\overline{\varphi}(X)v}$ is C^k and thus $g \to T(g)\overline{\varphi}(X)v = (df_v)_g X$ is C^k for each $X \in T_g G$. This is just the condition for f_v to be C^{k+1}.

The induction begins by the fact that f_v is continuous from definition of a representation.

□

Definition 62.16.
*If $T\colon G \to \operatorname{End} E$ is a strongly continuous representation of the Lie group G on the Hilbert space E, the **Gårding space** is the subspace of E generated by vectors of the form*

$$T(f)v = \int_G f(g)T(g)v \, d\mu(g) \tag{62.26}$$

where $\mu$is a left invariant Haar measure on G, $v \in E$ and $f \in C^\infty(G, \mathbb{C})$.

Proposition 62.17.
The Gårding space is stable under $\overline{\varphi}(X)$ and

$$\overline{\varphi}(X)[T(f)v] = -T(\underline{X}f)v \tag{62.27}$$

where $\underline{X}$ is the right invariant vector field on G generated by $X \in \mathcal{G}$.

Proof. For each $t \neq 0$, we have

$$\frac{T(e^{tX}) - \mathbb{1}}{t} T(f)v = \frac{1}{t}[T(e^t X) - \mathbb{1}] \int_G f(g)T(g)v \, d\mu(g), \tag{62.28}$$

let's say that one can permute the operator and the integral, and apply the change of variable $g' = e^{tX}g$ in the first term. The measure being invariant we find

$$= \frac{1}{t} \int_G f(e^{-tX}g'))T(g')v \, d\mu(g') - \frac{1}{t} \int_G f(g)T(g)v \, d\mu(g) \tag{62.29}$$

$$= \int_G \frac{f(e^{-tX}g) - f(g)}{t} T(g)v \, d\mu(g). \tag{62.30}$$

Taking the limit $t \to 0$ (which can be permuted with the integral), it comes

$$\begin{aligned} \overline{\varphi}(X)T(f)v &= \underline{(-X)}_g f T(g)v \\ &= T(-\underline{X}f)v \end{aligned} \tag{62.31}$$

which is the claim.

□

Theorem 62.18.
The Gårding space of $C^\infty(T)$ is dense in E. Then $C^\infty(T)$ itself is dense in E.

Proof. Let $v \in E$ and $\varepsilon > 0$. We are going to find an element of the Gårding space at a distance lower than ε of v. From assumption, T is strongly continuous. We define

$$S = \{g \in G \text{ tel que } |T(g)v - v| \leqslant \varepsilon\}$$

which is open by strong continuity. Let f be a continuous positive function with support contained in S and such that $\int_G f \, d\mu = 1$. With this function,

$$|T(f)v - v| = \left| \int_G f(g)[T(g)v - v] \, d\mu(g) \right|.$$

We can change the integral over G into an integral over S because $f = 0$ outside S and enforce the inequality by entering the norm in the integral. So

$$\begin{aligned} |T(f)v - v| &\leqslant \int_S f(g)|T(g)v - v| \, d\mu(g) \\ &\leqslant \varepsilon \int_S f(g) \, d\mu(g) = \varepsilon. \end{aligned} \tag{62.32}$$

□

62.3 Operator symbol

Operator symbol are natural framework to build oscillatory integral. Most of this theory comes from the book [951] of Dieudonné.

62.3.1 A case without problem

Let X be an open subset of $\mathbb{R}^{N'}$ and $A_\nu \colon X \to M_{N''\times N'}$ with $|\nu| \leqslant m$, some C^∞ maps. A differential operator is a map $P \colon \mathscr{E}_{\mathbb{R}}(X)^{N'} \to \mathscr{E}_{\mathbb{R}}(X)^{N''}$ of the following form:

$$P(u) = \sum_{|\nu|\leqslant m} A_\nu \cdot D^\nu u,$$

where the dot denotes a product matrix times vector. We suppose that the support of u is compact in X, so that it can be extended by 0 in $\mathbb{R}^n \backslash X$ in order to get a function in $\mathscr{D}_{\mathbb{R}}(X)^{N'}$ that we will also denote by u. One can compute the Fourier transform of u: $\mathcal{F}u_k \in \mathscr{C}(\mathbb{R}^n)$, and $\mathcal{F}u \in \mathscr{C}(\mathcal{R}^n)^{N'}$. A main property of Fourier transform is that

$$D^\nu u = \int_{\mathbb{R}^n} e^{2\pi i x \,\cdot\, \xi} (2\pi u\xi)^\nu (\mathcal{F}u)(\xi) d\xi.$$

If we pose

$$A(x,\xi) = \sum_{|\nu|\leqslant m} (2\pi i\xi)^\nu A_\nu(x),$$

$P(u)$ can be written as

$$P(u) = \int_{\mathbb{R}^n} e^{2\pi i x \,\cdot\, \xi} A(x,\xi) \cdot (\mathcal{F}u)(\xi) d\xi.$$

This integral makes only sense because we had carefully chosen allows the regularity conditions: as far as integral over ξ is concerned, the functions $A(x,\xi)$ are polynomial with coefficients in $\mathscr{E}_{\mathbb{R}}(X)^{N'N''}$ while the exponential contains a scalar product. The purpose of oscillatory integral is to generalise these two circumstances.

62.3.2 A problem

Let us consider the integral $\int_1^\infty \frac{1}{x} e^{ix} dx$. How to give a sense to that? Since $e^{ix} = -i\partial_x e^{ix}$, an integration by parts gives

$$\int_1^\infty \frac{1}{x} e^{ix} = -i \int_1^\infty \frac{1}{x} \partial_x(e^{ix}) = ie^{-i} - \int_1^\infty \frac{1}{x^2} e^{ix} \tag{62.33}$$

where the last integral exists in the usual sense.

62.3.3 Basic definitions

Definition 62.19.
*Let X be an open subset of $\mathbb{R}^n$, $a \colon X \times \mathbb{R}^N \to \mathbb{R}$ a C^∞ function and m be any real. We say that a is an **operator symbol** of order m in $X \times \mathbb{R}^N$ when*

$\forall L \subset X$ compact, $\forall$ multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\beta = (\beta_1, \ldots, \beta_n)$,

$\exists$ constant $c_{\alpha,\beta,L} > 0$ such that $\forall (x,\xi) \in L \times \mathbb{R}^N$,

$$|D_x^\alpha D_\xi^\beta a(x,\xi)| \leqslant c_{\alpha,\beta,L}(1+|\xi|)^{m-|\beta|}, \tag{62.34}$$

eq:cond_symbol

where $|\xi|^2 := \sum_{k=1}^N |\xi_k|^2$.

Note that in $X \times \mathbb{R}^N$, $X \subset \mathbb{R}^n$ with n and N not necessarily equals. For short, we will often say "symbol" instead of "operator symbol"; the set of symbols of order m in $X \times \mathbb{R}^N$ is a vector space denoted by $\mathcal{S}^m(X \times \mathbb{R}^N)$. pg:defmS

For terminology issues, we say that a property of the point $(x, \xi) \in X \times \mathbb{R}^N$ is *true when* $|\xi|$ *is large* if $\forall$ compact $L \subset X$, there exists $r_L > 0$ such that the property is true $\forall (x, \xi)$ with $x \in L$ and $|\xi| \geqslant r_L$.

This allows us to re-express the definition of a symbol. We say that $a \in C^\infty(X \times \mathbb{R}^N)$ is a symbol when

$\forall$ multi-indices α, β,

$(1 + |\xi|)^{-m+|\beta|} D_x^\alpha D_\xi^\beta a(x, \xi)$ is bounded when $|\xi|$ is large.

Indeed, the statement that $f(x, \xi)$ is such that $(1 + |\xi|)^{-q} f(x, \xi)$ is bounded when $|\xi|$ is large is the existence of a constant c (which depend on $L \subset X$) such that $f(x, \xi) \leqslant c(1 + |\xi|)^q$.

Proposition 62.20.
A function in $\mathscr{E}(X \times \mathbb{R}^N)$ which is zero for large $|\xi|$ is a symbol for all order.

Proof. The assumption is: for all compact subset $L \subset X$, the restriction of a to $L \times \mathbb{R}^N$ have a compact support. Let us fix a compact L and multi-indices α, β; then on $L \times \mathbb{R}^N$, a and $D_x^\alpha D_\xi^\beta a$ have a compact support. Then it is bounded by continuity. The same makes $|D_x^\alpha D_\xi^\beta a(x, \xi)|(1+|\xi|)^{-m+|\beta|}$ bounded and then it can be majored by a constant $c_{\alpha\beta L}$. □

More generally, for same reason, if $a \in \mathcal{S}^m(X \times \mathbb{R}^N)$ and $b = a$ when $|\xi|$ is large, then $b \in \mathcal{S}^m(X \times \mathbb{R}^N)$.

The function

$$\sigma(\xi) = \frac{1}{1 + \xi^2}$$

is a symbol of order 2. Here, there are no x part and $\xi \in \mathbb{R}$. The problem is to find a c_β for any β. For $\beta = 0$, the condition (62.34) becomes $\sigma(\xi) \leqslant c(1 + \xi)^2$, which is true. For $|\beta| = 1$, $c = 2$ works because

$$\frac{2\xi}{(1 + \xi^2)^2} \leqslant 2(1 + \xi).$$

It is clear that it will always woks because the degree of the denominator becomes bigger a bigger as $|\beta|$ grows.

This is a special case of a more general situation.

Proposition 62.21.
A complex function of $\mathscr{E}(X \times \mathbb{R}^N)$ which is positively homogeneous of degree m when $|\xi|$ is large is a symbol of order m.

prop:23.16.4

Proof. The assumption is that $\forall$ compact $L \subset X$, $\exists r_L > 0$ such that $\forall x \in L$, $|\xi| \geqslant r_L$, and for all $\lambda \geqslant 1$,

$$a(x, \lambda\xi) = \lambda^m a(x, \xi).$$

Let us define $b(x, \xi) = a(x, \lambda\xi)$. We have $(D_\xi b)(x, \xi) = \lambda(D_\xi a)(x, \xi)$, then

$$(D_\xi^\beta b)(x, \xi) = \lambda^{|\beta|}(D_\xi^\beta a)(x, \lambda\xi).$$

Since $b(x, \xi) = a(x, \lambda\xi) = \lambda^m a(x, \xi)$, we have $(D_\xi^\beta b)(x, \xi) = \lambda^m (D_\xi^\beta a)(x, \xi)$. By equalizing both expression of $(D_\xi^\beta b)(x, \xi)$, we find

$$(D_x^\alpha D_\xi^\beta a)(x, \lambda\xi) = \lambda^{m-|\beta|}(D_x^\alpha D_\xi^\beta a)(x, \xi),$$

when $x \in L$, $|\xi| \geqslant r_L$ and $\lambda \geqslant 1$.

On the (compact) sphere $|\xi| = r_L$, $|(D_x^\alpha D_\xi^\beta a)(x,\xi)|$ is bounded. Let c be a majoration, depending only on α, β and L. Any $\eta \in \mathbb{R}^N$ with $|\eta| \geqslant r_L$ can be written as $\eta = \lambda\xi$ with $\lambda \geqslant 1$ and $|\xi| = r_L$. Then

$$|(D_x^\alpha D_\xi^\beta a)(x,\eta)| \leqslant c\lambda^{m-|\beta|}. \tag{62.35}$$ eq:majo

It is cleat that $1 + |\eta| \geqslant |\eta| \geqslant |\eta|/|\xi| = \lambda$. Then, for $m - |\beta| > 0$, the majoration (62.35) keep if we replace $\lambda^{m-|\beta|}$ by $(1 + |\eta|)^{m-|\beta|}$.

If $m - |\beta| < 0$, the replacement of $\lambda^{m-|\beta|}$ by $(1 + |\eta|)^{m-|\beta|}$ need to change the constant c. More precisely, it raises the question to find a constant k such that $|\eta|^{-r} \leqslant k(1 + |\eta|)^{-r}$. It is easy to see that any

$$k > \left(\frac{r_L}{1 + r_L}\right)^{-r}$$

works. Finally,

$$|(D_x^\alpha D_\xi^\beta a)(x,\eta)| \leqslant c\lambda^{m-|\beta|} \leqslant ck(1 + |\eta|)^{m-|\beta|}. \tag{62.36}$$

□

This proposition speaks about a function which is homogeneous when $|\xi|$ is large. There exist functions which are homogeneous although not symbol because of problems of continuity at 0. For example, $a(x,\xi) = |\xi|$ when $N = 1$. It is not C^∞ at zero.

Remarque 62.22.
A function in $\mathscr{E}(X)$ is a C^∞ function on $X \times \mathbb{R}^N$ which is positively homogeneous of degree zero. Then

$$\mathscr{E}(X) \subset \mathcal{S}^0(X \times \mathbb{R}^N).$$

We now give an useful result without proof.

Proposition 62.23.
Let a be a C^∞ function which is only *defined for large $|\xi|$ and such that for all α, β,*

$$(1 + |\xi|)^{-m+|\beta|} D_x^\alpha D_\xi^\beta a(x,\xi)$$

is bounded for large $|\xi|$.

Then there exists a symbol b of order m on $X \times \mathbb{R}^N$ such that $a = b$ when $|\xi|$ is large.

For notational convenience, we define

$$a_\lambda(x,\xi) = a(x,\lambda\xi). \tag{62.37}$$ eq:def_ablambda

Let us clearly compute a derivative of a_λ. The notation $(D_\xi a_\lambda)(x_0,\xi_0)$ has to be read as "The derivative of a_λ with respect to his second argument at point (x_0,ξ_0)". If $M_\lambda(x,\xi) = (x,\lambda\xi)$,

$$\begin{aligned}(D_\xi a_\lambda)(x_0,\xi_0) &= D_\xi(a \circ M_\lambda)(x_0,\xi_0)\\ &= (D_\xi a)(M_\lambda(x_0,\xi_0)) \cdot \frac{dM_\lambda}{d\xi}(x_0,\xi_0)\\ &= \lambda(D_\xi a)(x_0,\lambda\xi_0).\end{aligned} \tag{62.38}$$

tho:dieu23.16.6

Theorem 62.24.
A function $a \in \mathscr{E}(X \times \mathbb{R}^N)$ is a symbol of order m if and only if the set of the restrictions of the functions $\lambda^{-m} a_\lambda$ (with $\lambda \geqslant 1$) to $X \times (\mathbb{R}^N \backslash \{o\})$ is bounded in the sense of definition 48.45 in the Fréchet space $\mathscr{E}(X \times (\mathbb{R}^N \backslash \{o\}))$.

Proof. Let a be a symbol of order m. By definition of a_λ,

$$D_x^\alpha D_\xi^\beta a_\lambda(x,\xi) = \lambda^{|\beta|}(D_x^\alpha D_\xi^\beta a)(x,\lambda\xi),$$

but any compact in $X \times (\mathbb{R}^N \backslash \{o\})$ is contained in a compact which can be written as $L \times G$ with L compact in X and $G = \{\xi \in \mathbb{R}^N \,|\, r \leqslant |\xi| \leqslant R\}$ with $0 < r < R$.

Since a is a symbol, we have

$$|D_x^\alpha D_\xi^\beta a(x, \lambda\xi)| \leqslant c_{\alpha,\beta,L}(1 + \lambda|\xi|)^{m-|\beta|};$$

if we multiply it by $(1/\lambda)^{m-|\beta|}$, we find

$$\begin{aligned} \lambda^{|\beta|-m}|D_x^\alpha D_\xi^\beta a(x, \lambda\xi)| &\leqslant c_{\alpha,\beta,L}\left(\frac{1}{\lambda} + |\xi|\right)^{m-|\beta|} \\ &\leqslant \begin{cases} c_{\alpha,\beta,L}(1+R)^{m-|\beta|} & \text{if } |\beta| \leqslant m \\ c_{\alpha,\beta,L} r^{m-|\beta|} & \text{if } |\beta| > m. \end{cases} \end{aligned} \tag{62.39}$$

Now, we want to know if $\{\lambda^{-m} a_\lambda | \lambda \geqslant 1\}$ is bounded in $\mathscr{E}(X \times (\mathbb{R}^N \backslash \{o\}))$. The computation is

$$\begin{aligned} p_{sj}(\lambda^{-m} a_\lambda) &= \sup_{\substack{(x,\xi)\in K_j \\ \alpha,\beta \text{ st } |\alpha|+|\beta|\leqslant s}} |D_x^\alpha D_\xi^\beta (\lambda^{-m} a_\lambda)(x,\xi)| \\ &= \lambda^{-m} \sup_{\dots} \lambda^{|\beta|} (D_x^\alpha D_\xi^\beta a)(x, \lambda\xi)| \\ &= \lambda^{-m+|\beta|} \sup_{\dots} |(D_x^\alpha D_\xi^\beta a)(x, \lambda\xi)|. \end{aligned} \tag{62.40}$$

Then

$$p_{s,m}(\lambda^{-m} a_\lambda) \leqslant \begin{cases} c_{\alpha,\beta,L}(1+R)^{m-|\beta|} & \text{if } |\beta| \leqslant m \\ c_{\alpha,\beta,L} r^{m-|\beta|} & \text{if } |\beta| > m. \end{cases} \tag{62.41}$$

Now, we prove the reverse sense. Let us suppose that p_{sj} is bounded, *i.e.*

$$|\lambda^{-m+|\beta|}(D_x^\alpha D_\xi^\beta a)(x, \lambda\xi)| \leqslant A$$

with $x \in L$, $r \leqslant |\xi| \leqslant R$ and $\lambda \geqslant 1$. Then, for $x \in L$ and $|\xi| \geqslant r$,

$$|(D_x^\alpha D_\xi^\beta a)(x,\xi)| \leqslant A\left(\frac{|\xi|}{r}\right)^{m-|\beta|}. \tag{62.42}$$ eq:2409r1

Since for $|\xi| \geqslant r$,

$$\frac{r}{1+r} \leqslant \frac{|\xi|}{1+|\xi|} \leqslant 1,$$

one can "forget" the $(1/r)^{m-|\beta|}$ in (62.42): it can be absorbed in a constant which depend on β. By the same as in the proof of proposition 62.3.3, we can also replace $|\xi|^{m-|\beta|}$ by $(1+|\xi|)^{m-|\beta|}$; this proves that a is a symbol of order m.

□

62.3.4 Topology on $\mathcal{S}^m(X \times \mathbb{R}^N)$

We will endow the vector space $\mathcal{S}^m(X \times \mathbb{R}^N)$ with a locally convex topological structure.

Let (p_k) be the family of seminorms defining the topology of $\mathscr{E}(X \times \mathbb{R}^N)$; we consider their restriction to $\mathcal{S}^m(X \times \mathbb{R}^N)$. We also consider (p'_k), the one which defines the topology of $\mathscr{E}(X \times (\mathbb{R}^N \backslash \{o\}))$. Now we pose for $a \in \mathcal{S}^m(X \times (\mathbb{R}^N \backslash \{o\}))$,

$$q_k(a) = \sup_{\lambda \geqslant 1} p'_k(\lambda^{-m} a_\lambda).$$

One can see that these q_k are seminorms. Finally, we put on $\mathcal{S}^m(X \times \mathbb{R}^N)$ the topology defined by the p_k *and* the q_k.

First remark: this topology on $\mathcal{S}^m$ is finer[2] than the one which is induced from $\mathscr{E}(X \times \mathbb{R}^N)$. In particular, $\mathcal{S}(X \times \mathbb{R}^N)$ is metrisabe[3].

2. Anglais ?
3. Anglais ?

Lemma 62.25.
Each seminorm defining the topology of a Fréchet space is continuous. lem:q_cont

Proof. Let q be one of them. By proposition 48.29, it is sufficient to find a neighbourhood of 0 on which q is bounded. Let us show that

$$B(0;q,r) = \{x \in \mathcal{S}^m | q(x) < r\}$$

works. Indeed, it is an open set by definition of the topology, we just have to prove that this contains 0, *i.e.* we have to find a r such that $q(0) < r$.

$$q(0) = sup_{\lambda \geqslant 1} p'(0) = p'(0)$$

where p defines the topology of $\mathscr{E}(X \times (\mathbb{R}^N \backslash \{o\}))$. In proposition 48.44, we see that $p'(f) = sup_{\dots} |(D^{\cdots} f)(x)|$. It is clear that this is zero when $f \equiv 0$.

□

Proposition 62.26.
For this topology, $\mathcal{S}^m(X \times \mathbb{R}^N)$ is a Fréchet space.

Proof. We already know that $\mathcal{S}^m(X \times \mathbb{R}^N)$ is locally convex space because its topology is defined by seminorms. It is also Hausdorff because its topology is finer than the one of $\mathscr{E}$ which is itself separable. Now, we consider $\mathscr{E}(X \times \mathbb{R}^N)$ as an additive group, so that we can use the theory of metrisable groups developed in point 48.4. Thus we can speak of **Cauchy sequences**: a sequence (f_n) in $\mathscr{E}(X \times \mathbb{R}^N)$ is Cauchy when

$\forall\, V \in \mathcal{V}(0)$, $\exists\, n_0$ such that $\forall\, m, n \geqslant n_0$, $f_n - f_m \in V$,

where $0 \colon X \times \mathbb{R}^N \to \mathbb{R}$ is the null function.

From proposition 48.46, we know that $\mathscr{E}(\mathcal{U})$ is a Fréchet space; in particular, it is complete for the distances which defines the Cauchy sequences (the ones whose are invariant under translations).

Let us consider a Cauchy sequence (a^n) in $\mathcal{S}^m(X \times \mathbb{R}^N)$. It converges (in the sense of $\mathscr{E}(X \times \mathbb{R}^N)$) to $b \in \mathscr{E}(X \times \mathbb{R}^N)$. A Cauchy sequence is always contained in a compact set, and by lemma 62.25, q_k is continuous for any k. Thus for each k the sequence (in $\mathbb{R}$) given by $(q_k(a^n))$ is bounded (continuous function on a compact).

So, for each j, the set $\{p'_k(\lambda^{-m} a^n_\lambda)\}$ is bounded when $\lambda, n \geqslant 1$. If we let n goes to infinity, we find that $\{p'_k(\lambda^{-m} b_\lambda) : \lambda > 1\}$ is bounded, and then $b \in \mathcal{S}^m(X \times \mathbb{R}^N)$.

We had just proved that b where the limit of (a^n) in the sense of $\mathscr{E}(X \times \mathbb{R}^n)$; we yet have to see that b is also the limit in the sense of $\mathcal{S}^m(X \times \mathbb{R}^n)$. As (a^n) is a Cauchy sequence, proposition 48.54 assures us that

$\forall\, V \in \mathcal{V}(0)$, $\exists\, n_0$ such that $\forall\, m, n \geqslant 0$, $f_n - f_m \in V$.

In particular, a neighbourhood V of 0 is for example $\{x \in \mathcal{S}^m | q_k(x) < r\}$. Then $\forall \epsilon > 0$, $\exists\, n_0$ such that $i, j > 0$ implies $q_k(a^i - a^j) < \epsilon$. So b is the limit of (a^n) when $n \to \infty$ in the sense of $\mathcal{S}^m(X \times \mathbb{R}^N)$.

□
LemHIUsKABh

Lemma 62.27.
Let B be a bounded set in $\mathscr{E}(\mathcal{U})$ and $A > 0$ such that $f(z) > A$ for all $f \in B$ and for all $z \in \mathcal{U}$. Then the set $\{f^s\}_{f \in B}$ is bounded in $\mathscr{E}(\mathcal{U})$.

Proof. We have to show that $(D^\nu(f^s))(z)$ is bounded when f runs over B and z keeps in a compact subset of $\mathcal{U}$. Let's use an induction on $|\nu|$ in order to prove that

$$D^\nu(f^s) = f^{s-|\nu|} P_\nu\big((d^\rho f_{\rho \leqslant \nu})\big)$$

where P_ν is a polynomial with coefficients independent of f. Indeed, let $\nu = \{\nu_0, \sigma\}$ where ν_0 is a multi-index and σ a single index.

$$\begin{aligned} D_\nu(f^s) = D_\sigma D_{\nu_0}(f^s) &= D^\sigma \left(f^{s-|\nu_0|} P_{\nu_0}\big((D^\rho f)_{\rho \leqslant \nu_0}\big) \right) \\ &= (s - |\nu_0|) f^{s-|\nu_0|-1} (D^\sigma f) P_{\nu_0}\left((D^\rho f)_{\rho \leqslant \nu_0}\right) \\ &\quad + f^{s-|\nu_0|} D^\sigma \left(P_{\nu_0}(D^\rho f)_{\rho \leqslant \nu_0} \right). \end{aligned} \tag{62.43}$$

The quantity $P_\nu\left((D^\rho f)_{\rho\leqslant\nu}\right)$ remains bounded because $|f(z)|^{-1}\leqslant A^{-1}$ and z belongs to a compact set. The function f take bounded values (continuous on a compact set) and while $s-|\nu|>0$, the function $f^{s-|\nu|}$ is bounded too. If $s-|\nu|<0$, there can be a problem, but th condition $f(z)\geqslant A$ avoid this case. □

tho:lenumf

Theorem 62.28.
Multiplication and derivations on symbols behave rather well:

enufi

(1) For all multi-index γ, δ, the map $a\to D_x^\gamma D_\xi^\delta a$ is a continuous linear map from $\mathcal{S}^m$ to $\mathcal{S}^{m-|\delta|}$.

enufii

(2) The map $(a,b)\to ab$ is continuous and bilinear from $\mathcal{S}^m\times\mathcal{S}^{m'}$ into $\mathcal{S}^{m+m'}$.

enufiii

(3) Take $a\in\mathcal{S}^m$. The function a^{-1} is defined and is a symbol of order $-m$ if and only if for all compact $H\subset X$ there exists a constant $c_H>0$ such that, in $H\times\mathbb{R}^n$, the condition

$$|a(x,\xi)|\geqslant c_H(1+|\xi|)^m$$

holds. In this case, for all real number s, we have $|a|^s\in\mathcal{S}^{sm}$.

We only prove the point (3).

Proof. Necessary condition. Let $a\in\mathcal{S}^m$ such that a^{-1} is well defined as symbol of order $-m$. By the definition of a symbol,

$$\left|\frac{1}{a(x,\xi)}\right|\leqslant c_H(1+|\xi|)^{-m} \tag{62.44}$$

for a certain c_H determined by the compact H in which x moves. Then $|a(x,\xi)|\geqslant c_H^{-1}(1+|\xi|)^m$ and the claim follows

Sufficient condition. Let $a\in\mathcal{S}^m$ and for each compact $H\subset X$, there exists a $c_H>0$ such that $\forall(x,\xi)\in H\times\mathbb{R}^N$,

$$|a(x,\xi)|\geqslant c_H(1+|\xi|)^m.$$

By point (2), $|a|^2\in\mathcal{S}^{2m}$ and $a^{-1}=\bar{a}(|a|^3)^{-1}$. Let us suppose that the theorem is proved when $a(x,\xi)>0$ on $X\times\mathbb{R}^N$. So suppose that $a<0$ somewhere. In this case, $a^{-1}=\bar{a}(|a|^2)^{-1}$ where $|a|^2$ is positive in such a manner that our assumption makes it a symbol of order $-2m$; now by (2), a^{-1} is a symbol of order $-m$.

So we can restrict ourself to the case where $a(x,\xi)>0$ in $X\times\mathbb{R}^N$. It is clear that $\lambda^{-ms}(a_\lambda^s)=(\lambda^{-m}a_\lambda)^s$. From theorem 62.24 we just have to prove that $\lambda^{-m}a_\lambda)^s$ is a bounded set when $s=-1$. Lemma 62.27 shows it for all s.

□

Proposition 62.29.
Let Y be an open subset of $\mathbb{R}^p$,

$$\psi\colon Y\times\mathbb{R}^P\to X$$

and

$$\theta\colon Y\times\mathbb{R}^P\to\mathbb{R}^N,$$

two C^∞ functions such that for large $|\xi|$, the function ψ is positively homogeneous of degree zero while θ is of degree 1 with respect to ξ.

Then for all symbol $a\in\mathcal{S}^m(X\times\mathbb{R}^N)$, the function

$$b\colon Y\times\mathbb{R}^P\to\mathbb{R}\ (x',\xi')\ \mapsto a(\psi(x',\xi'),\theta(x',\xi')) \tag{62.45}$$

is a symbol of $\mathcal{S}^m(Y\times\mathbb{R}^P)$ and the linear map $\mathcal{S}^m(X\times\mathbb{R}^N)\to\mathcal{S}^m(Y\times\mathbb{R}^P)$ given by $a\to a\circ(\psi,\theta)$ is continuous.

Proof. We pose $F = (\psi, \theta)$, so that $b = a \circ F$. Homogeneity assumptions make that when $\lambda \geqslant 1$ and when $|\xi|$ is large,

$$\lambda^{-m} b_\lambda = (\lambda^{-m} a_\lambda) \circ F.$$

Then the problem reduces to prove that if B is bounded in $\mathscr{E}(X \times \mathbb{R}^N_0)$, then the set $f \circ F$ is bounded when f runs over B. □

Lemma 62.30.
Let $m < m' \in \mathbb{R}$. Then the canonical injection

$$\mathrm{Id} \colon \mathcal{S}^m \to \mathcal{S}^{m'}$$

is continuous.

Proof. Let $\mathcal{O}$ be an open set in $\mathcal{S}^{m'}$; we have to show that $\mathrm{Id}^{-1}(\mathcal{O}) = \mathcal{O} \cap \mathcal{S}^m$ is open in $\mathcal{S}^m$. Let p_k be the family of seminorm defining the topology on $\mathscr{E}(X \times \mathbb{R}^N)$ and p'_k the one of $\mathscr{E}(X \times \mathbb{R}^N_0)$ and then

$$q_k(a) = \sup_{\lambda \geqslant 1} p'_k(\lambda^{-m} a_\lambda).$$

The topology of $\mathcal{S}$ is defined from p_k and q_k. We define the ball

$$B(a, d, r) = \{x \in E \text{ tel que } d(a, x) < r\}.$$

The property for $\mathcal{O} \subset \mathcal{S}^{m'}$ to be open is that one of these ball is included in $\mathcal{O}$. If a ball build with one of the p_k's is included in $\mathcal{O}$, there are no problems because the seminorms p_k are also in the definition of the topology on $\mathcal{S}^m$. Let

$$B(a, d^{m'}_k, r) = \{x \in \mathcal{S}^{m'} \text{ tel que } q^{m'}_k(a - x) < r\}.$$

The candidate ball of $\mathcal{S}^m$ to be included in $\mathcal{O}$ is

$$\mathcal{O}' = \{x \in \mathcal{S}^m \text{ tel que } q^m_k(a - x) < \overline{r}\}$$

for a certain $\overline{r} < r$. Let us prove that $q^m_k(y) \leqslant q^{m'}_k(y)$. Here, $k = (n, s)$.

$$\begin{aligned} q^m_k(y) &= \sup_{\lambda \geqslant 1} \sup_{\substack{x \in K_n \\ |\nu| < s\|}} |D^\nu(\lambda^{-m} y_\lambda)(x)| \\ &= \sup_{\lambda \geqslant 1} \lambda^{m'-m} \sup_{\substack{x \in K_n \\ |\nu| < s\|}} |D^\nu(\lambda^{-m'} y_\lambda)(x)| \\ &\leqslant \sup_{\lambda \geqslant 1} p'_k(\lambda^{-m'} y_\lambda) \\ &= q^{m'}_k(y). \end{aligned} \tag{62.46}$$

□

Lemma 62.31.
The adherence of $\mathscr{D}(X \times \mathbb{R}^N)$ in $\mathcal{S}^m(X \times \mathbb{R}^N)$ contains $\mathcal{S}^{m'}(X \times \mathbb{R}^N)$ for all $m' < m$. lem:DadhS

One can characterize the topology by another choice of seminorms as the following proposition shows.

Proposition 62.32.
The topology of Fréchet space on $\mathcal{S}(X \times \mathbb{R}^N)$ is defined by the seminorms

$$p_{K,\alpha,\beta}(a) = \sup_{x \in K, \xi \in \mathbb{R}^N} (1 + |\xi|)^{-m+|\beta|} |D^\alpha_x D^\alpha_\xi a(x, \xi)|.$$

prop:topo_alter

62.3.5 Asymptotic expansions

Note that $\mathscr{D}(X\times\mathbb{R}^N)\subset\mathcal{S}^{-\infty}(X\times\mathbb{R}^N)$ because a compact supported function is always a symbol of all order. Let us give without proof the two following results [951].

Lemma 62.33.
Let $a\in\mathcal{S}^{m'}(X\times\mathbb{R}^N)$ a symbol whose support is contained in $H\times\mathbb{R}^N$ where H is a compact subset of X. Let $h\in\mathscr{D}(X\times\mathbb{R}^N)$, a function equals to 1 for $x\in H$ and $|\xi|\leqslant A$ for a certain constant $A>0$. Then for all $m>m'$,

$$\lim_{q\to\infty}(h_{1/q}a)=a$$

in the sense of the convergence in $\mathcal{S}^m$. lem:limha

Proposition 62.34.
The adherence of $\mathscr{D}(X\times\mathbb{R}^N)$ in $\mathcal{S}^m(X\times\mathbb{R}^N)$ contains $S^{m'}(X\times\mathbb{R}^N)$ for all $m'<m$.

The main theorem is

Theorem 62.35.
Consider a strictly increasing sequence in $\mathbb{R}$ (m_j) with $\lim_{j\to\infty}m_j=-\infty$. Let, for each j, $a_j\in\mathcal{S}^{m_j}(X\times\mathbb{R}^N)$. Then there exists $a\in\mathcal{S}^{m_0}$ such that for all k,

$$a-\sum_{j<k}a_j\in\mathcal{S}^{m_k}(X\times\mathbb{R}^N). \tag{62.47}$$ eq:dev_ass

Furthermore if a' shares this property with a, then $a-a'\in\mathcal{S}^{-\infty}(X\times\mathbb{R}^N)$. tho:dev_ass

Proof. The last assertion is easy. Let us subtract equations (62.47) for two such symbols: for all k,

$$a-a'\in\mathcal{S}^{m_k}(X\times\mathbb{R}^N).$$

Then $a-a'\in\mathcal{S}^{X\times\mathbb{R}^N}$.

Now consider a locally finite countable covering $(\mathcal{U}_\alpha)$ of open relatively compact sets and (f_α), a partition of unity for this covering. Let us suppose that, for each α, we know a symbol b_α of order m_0 which is zero outside $\mathcal{U}_\alpha\times\mathbb{R}^N$ and such that for all k,

$$b_\alpha=\sum_{i<k}f_\alpha a_j\in\mathcal{S}^{m_k}(X\times\mathbb{R}^N).$$

In other words, let us suppose that we have the answer for the symbols $f_\alpha a_j$ instead of the symbols a_j. Then the symbol $a=\sum_\alpha b_\alpha$ answer the question for the symbols a_j. Indeed

$$\sum_\alpha b_\alpha-\sum_j a_j=\sum_\alpha(b_\alpha-\sum_j f_\alpha a_j) \tag{62.48}$$

where the sums are pointwise *finite* sums because the covering $(\mathcal{U}_\alpha)$ is locally finite.

So we are left to consider a_j vanishing outside a set $H_j\times\mathbb{R}^N$ where H_j is a compact subset of X. For $j\geqslant 1$, let us define

$$a_{j,q}=(1-h_{j,1/q})a_j\in\mathcal{S}^{m_j} \tag{62.49}$$

where $h_j\in\mathscr{D}(X\times\mathbb{R}^N)$ fulfils $h_j(x,\xi)=1$ when $x\in H_j$ and $0\leqslant|\xi|\leqslant 1$. By $h_{i,1/q}$, we mean $(h_j)_{1/q}$ in the sense of equation (62.37). Since it has a compact support, it is a symbol for all orders. From lemma 62.33,

$$\lim_{q\to\infty}(a_{j,q})=\lim_{q\to\infty}a_j-\lim_{q\to\infty}h_{j,1/q}a_j=0.$$

Proposition 62.4 makes that for all $r\geqslant 0$, the sum $\sum_{j\geqslant r}a_{j,q_j}$ converges in $\mathcal{S}^{m_r}$. We are now going to prove that the symbol $a=\sum_{j\geqslant 0}a_{j,q_j}$ is the one needed by the theorem. First remark that $a_{j,q}$ has support contained in $H_j\times\mathbb{R}^N$. For all $q>0$,

$$a_j-a_{j,q}\in\mathscr{D}(X\times\mathbb{R}^N)\subset\mathcal{S}^{-\infty}(X\times\mathbb{R}^N).$$

Now,

$$\Big(a - \sum_{j<k} a_j\Big) - \Big(a - \sum_{j<k} a_{j,q_j}\Big) = \sum_{j<k} \Big(a_{j,q_j} - a_j\Big) \in \mathcal{S}^{-\infty}(X \times \mathbb{R}^N),$$

then

$$a - \sum_{j<j} a_{j,q_j} = \sum_{l \geqslant k} a_{l,q_l} \in \mathcal{S}^{m_k}$$

from lemma 62.4.

□

When equation (62.47) holds , we say that the sum $\sum_j a_j$ is an **asymptotic expansion** of the symbol a and we write $a \sim \sum_j a_j$. In this case, for all k we have

$$a - \sum_{j<k} a_j \in \mathcal{S}^{m_k}, \tag{62.50}$$

eq:asa

thus for any multi-index ν, the following holds

$$D^\nu a \sim \sum_j D^\nu a_j$$

as can be seen by derivation of equation (62.50) and using point (1) of theorem 62.28.

62.3.6 Tempered functions

A function $f \in \mathscr{E}(X \times \mathbb{R}^N)$ is **tempered with respect to** $|\xi|$ if for all compact $L \subset X$, and for all pair of multi-index (α, β), there exists constants $c(\alpha, \beta, L)$ and $p(\alpha, \beta, L)$ such that

$$|D_x^\alpha D_\xi f(x, \xi)| \leqslant c(\alpha, \beta, L)(1 + |\xi|)^{p(\alpha,\beta,L)}$$

for all $(x, \xi) \in L \times \mathbb{R}^N$.

Lemma 62.36.
Let $c \in \mathscr{E}(X \times \mathbb{R}^N)$ be a tempered function with respect to ξ such that for all compact $L \subset X$ and $q > 0 \in \mathbb{N}$, there exists a constant C_{qL} such that

$$|c(x, \xi)| \leqslant C_{qL}(1 + |\xi|)^{-q}$$

for all $(x, \xi) \in L \times \mathbb{R}^N$. Then $c \in \mathcal{S}^{-\infty}(X \times \mathbb{R}^N)$.

lem:csymbol

The following proposition gives a link between tempered functions and symbols.

Proposition 62.37.
Let X be an open set in $\mathbb{R}^N$ and $(m_j) \in \mathbb{R}$ a strictly decreasing sequence such that $\lim_{j\to\infty} m_j = -\infty$. For each m_j, we consider a symbol $a_j \in \mathcal{S}^{m_j}$. Let $a \in \mathscr{E}(X \times \mathbb{R}^N)$, a tempered function with respect to ξ and suppose that for all compact $L \subset X$, there exists a decreasing sequence $(q_k) \in \mathbb{R}$ with $\lim_{k\to\infty} q_k = -\infty$ and for each k, a constant C_{kL} such that

$$|a(x, \xi) - \sum_{j<k} a_j(x, \xi)| \leqslant C_{kL}(1 + |\xi|)^{q_k} \tag{62.51}$$

for all $(x, \xi) \in L \times \mathbb{R}^N$.

Then a is a symbol of order m_0 and admits the asymptotic expansion

$$a \sim \sum_j a_j.$$

Proof. Theorem 62.35 holds for symbols a_j, then there exists a symbol $b \in \mathcal{S}^{m_0}$ such that $b \sim \sum_j a_j$. So we have to prove that $a - b \in \mathcal{S}^\infty$ because, in this case,

$$a - \sum_{j<k} a_j = (a-b) - (\sum_{j<k} a_j - b) \in \mathcal{S}^{m_k}.$$

On the one hand, the function $a - b$ is tempered (because a symbol is always tempered). On the other hand, for all choice of k, we can write

$$\begin{aligned} |(a-b)(x,\xi)| &= |a - \sum_{j<k} a_j - b + \sum_{j<k} a_j| \\ &\leqslant |a - \sum_{j<k} a_j| + |b - \sum_{j<k} a_j| \\ &\leqslant C_{kL}(1+|\xi|)^{q_k} + D_L(1+|\xi|)^{m_k} \\ &\leqslant C'_{kL}(1+|\xi|)^{q'_k} \end{aligned} \tag{62.52}$$

where $C'_{kL} = \max(C_{kL}, D_L)$ and $q'_k = \max(q_k, m_k)$. Let us now fix $q > 0$; since $q_k, m_k \to -\infty$, there exists k such that $q'_k < -q$. For this one, we have

$$|(a-b)(x,\xi)| \leqslant C''_{qL}(1+|\xi|)^{-q}.$$

Lemma 62.36 concludes the proof.

□

62.4 First construction of oscillatory integrals

62.4.1 Construction

Let us consider X be an open set in $\mathbb{R}^N$, a continuous map $\varphi\colon X \times \mathbb{R}^N \to \mathbb{R}$ and a symbol $a \in \mathcal{S}^m(X \times \mathbb{R}^N)$. For any compact $L \subset X$, the integral

$$\int_{L\times\mathbb{R}^N} e^{i\varphi(x,\xi)} a(x,\xi) dx d\xi$$

converges when $m < -N$. Indeed

$$\int |e^{i\varphi} a| \leqslant \int |e^{i\varphi}||a| \leqslant \int C(1+|\xi|)^m. \tag{62.53}$$

The integral over L is a constant on a compact while the one on $\mathbb{R}^N$ is performed with spherical coordinates, sop we are left with $\int_{\mathbb{R}^+}(1+r)^m r^{N-1}$. Under these assumptions, the form

$$a \to \int_{L\times\mathbb{R}^N} e^{i\varphi} a \tag{62.54}$$

eq:formaprol

is linear. In order to see that is is continuous too, we will prove that there exists a neighbourhood of zero in $S^m(X \times \mathbb{R}^N)$ in which any a fulfils the majoration

$$|a(x,\xi)| \leqslant C(1+|\xi)^m$$

for all $x(x,\xi) \in L \times \mathbb{R}^N$ and a certain constant C. Let us use the seminorms given in proposition 62.32 and consider the neighbourhood

$$\mathcal{O} = \{a \in \mathcal{S}^m \text{ tel que } \sup_{x\in K, \xi\in\mathbb{R}^N} (1+|\xi|)^{-m}|a(x,\xi)| < \varepsilon\}$$

for a certain compact $K \subset X$ and $\varepsilon > 0$. A symbol a in $\mathcal{O}$ fulfils in particular $|a(x,\xi)| < \varepsilon(1+|\xi|)^m$. Then the linear form (62.54) is bounded and thus continuous.

def:phase

Definition 62.38.
*A function $\varphi\colon X\times\mathbb{R}_0^N\to\mathbb{R}$ is a **phase function** if*

(1) $\varphi\in C^\infty(X\times\mathbb{R}_0^N)$,

(2) the $n+N$ first derivatives $D'_j\varphi$ and $D''_k\varphi$ doesn't vanish at same point of $X\times\mathbb{R}_0^N$, in other words, φ is not singular

(3) for all $(x,\xi)\in X\times\mathbb{R}_0^N$ and all $\lambda>0$, the formula $\varphi(x,\lambda\xi)=\lambda\varphi(x,\xi)$ holds.

Here, D'_j denotes the derivative with respect to x_j and D''_k the one with respect to ξ_k. The third point implies that the prolongation $\varphi(X\times\{0\})=0$ in continuous in $X\times\mathbb{R}^N$ but not C^∞ as the example $\varphi(x,\xi)=|\xi|$ shows with $N=1$.

An important example of phase function is given when $N=n$ by $\varphi(x,\xi)=-2\pi\langle x,\xi\rangle$; it gives rise to the Fourier transforms.

An important result will help us to define oscillatory integrals

lem:defL

Lemma 62.39.
Let φ be a phase. There exists symbols $a_k\in\mathcal{S}^0$, $b_j\in\mathcal{S}^{-1}$ and $c\in\mathcal{S}^{-\infty}$ ($1\leqslant j\leqslant N$ and $1\leqslant j\leqslant n$) such that the differential operator

$$L=\sum_{k=1}^N a_kD''_k+\sum_{j=1}^n b_jD'_j+c\mathbb{1}$$

on $X\times\mathbb{R}^N$ fulfil

$$Le^{i\varphi}=e^{i\varphi}$$

in $X\times\mathbb{R}_0^N$. The transposed L^t of L is given by

$$L^t=-\sum_{a_k}D''_k-\sum_{j=1}^n b_jD'_j+c'\mathbb{1}$$

where $c'=c-\sum D''_ka_k-\sum D'_jb_j\in\mathcal{S}^{-1}$.

For all $u\in\mathscr{E}(X\times\mathbb{R}^N)$ and $r\in\mathbb{N}_0^+$,

$$\big((L^t)^ru\big)(x,\xi)=\sum_{|\alpha|+|\beta|\leqslant r}g_{\alpha\beta}(x,\xi)D_x^\alpha D_\xi^\beta u(x,\xi)$$

where each $g_{\alpha\beta}$ is a symbol of order $-r+|\beta|$ on $X\times\mathbb{R}^N$, independent of u. In other words, $(L^t)^r=\sum_{|\alpha|+|\beta|\leqslant r}g_{\alpha\beta}D_x^\alpha D_\xi^\beta$.

In particular $a\to(L^t)ra$ is a continuous linear map from $\mathcal{S}^m$ to $\mathcal{S}^{m-r}$.

We will not prove all assertions.

Proof. We consider the function

$$g(x,\xi)=|\xi|^2\sum_{k=1}^N\big(D''_k\varphi(x,\xi)\big)^2+\sum_{j=1}^n\big(D'_j\varphi(x,\xi)^2.$$

Since $(D''_k\varphi)(x,\lambda\xi)=(D''_k\varphi)(x,\xi)$, the function g is positively homogeneous of degree 2 with respect to ξ and from the fact the φ has no critical points, there is always at least one no zero term the two sums, then $g(x,\xi)\neq0$ on $X\times\mathbb{R}_0^N$. Now we consider a C^∞ function $h\colon\mathbb{R}\to\mathbb{R}$ with compact support (hence a symbol of all order) and equals to 1 in a neighbourhood of zero.

We are going to study the function $\frac{1-h(|\xi|)}{g(x,\xi)}$. In a neighbourhood of zero, the numerator is zero; then we don't mind with the values of $1/g(x,\xi)$ when $|\xi|\to0$. Proposition 62.3.3 shows that $g\mathcal{S}^2(X\times\mathbb{R}^N)$. On the other hand the inequality

$$|g(x,\xi)\geqslant C_H(1+|\xi|)|$$

holds for a certain constant C_H because both side are of degree two. Then point (1) of theorem 62.28 assures that $1/g$ is a symbol of $\mathcal{S}^{-2}$. So $\frac{1-h(|\xi|)}{g(x,\xi)} \in \mathcal{S}^{-2}$. Now we claim that $Le^{i\varphi} = e^{i\varphi}$ when

$$\begin{aligned} a_k(x,\xi) &= |\xi|^2 \frac{1-h(|\xi|)}{ig(x,\xi)} D''_k\varphi(x,\xi) \\ b_j(x,\xi) &= \frac{1-h(|\xi|)}{ig(x,\xi)} D'_j\varphi(x,\xi) \\ c(x,\xi) &= h(|\xi|). \end{aligned} \tag{62.55}$$

A simple counting shows that $a_k \in \mathcal{S}^0$ and $b_j \in \mathcal{S}^{-2}$. The computation is easy:

$$\begin{aligned} (Le^{i\varphi})(x,\xi) &= \sum_k a_k(x,\xi)(D''_k e^{i\varphi})(x,\xi) \\ &\quad + \sum_j (x,\xi)(D'_j e^{i\varphi})(x,\xi) \\ &\quad + c(x,\xi)e^{i\varphi(x,\xi)} \\ &= e^{i\varphi}\frac{1-h(|\xi|)}{ig}\Big(i\sum_k |\xi|^2(D''_k\varphi)^2 + \sum_j i(D'_j\varphi)^2\Big) + h(|\xi|)e^{i\varphi} \\ &= e^{i\varphi}. \end{aligned} \tag{62.56}$$

We don't prove the other assertions.

□

The theorem which defines the oscillatory integrals is the following

Theorem 62.40.
Let φ be a phase on $X \times \mathbb{R}^N$. For all $m \in \mathbb{R}$ and for all compact $H \subset X$, the $\mathbb{C}$-linear form $\mathscr{D}(X \times \mathbb{R}^N) \to \mathbb{C}$

$$a \to \int_{H\times\mathbb{R}^N} e^{i\varphi(x,\xi)} a(x,\xi)dxd\xi \tag{62.57}$$

admits an unique extension into a continuous map $S^m(X) \times \mathbb{R}^N \to \mathbb{C}$.

eskdfml

Proof. As far as integration is concerned, the set $X \times \{0\}$ is negligible, then one can apply the formula $Le^{i\varphi} = e^{i\varphi}$ although id is only true on $X \times \mathbb{R}^N_0$. From construction of L and the definition of the transpose,

$$\langle e^{i\varphi}, w\rangle = \langle Le^{i\varphi}, w\rangle = \langle e^{i\varphi}, L^t w\rangle,$$

then in $\mathscr{D}(X \times \mathbb{R}^N)$, the studied linear form reads

$$a \to \int_{H\times\mathbb{R}^N} e^{i\varphi}(L^t)^r a \tag{62.58}$$

eq:intaprol

for any integer $r > 0$. If one choses $r > m+N$, then $(L^t)^r a \in \mathcal{S}^{m-r}$ if $a \in \mathcal{S}^m$ and the integral exists and is a continuous function of a from discussion below equation (62.54). Then the integral in equation (62.58) defines a prolongation of the map $u \to \int e^{i\varphi}u$ into the whole $\mathcal{S}^m$ when $r > m+N$.

Lemma 62.31 makes that any element in $\mathcal{S}^m X \times \mathbb{R}^N$ is adherent to $\mathscr{D}(X \times \mathbb{R}^N)$, then a continuous prolongation is unique.

□

The (prolonged) linear form

$$a \to \int_{H\times\mathbb{R}^N} e^{i\varphi} a$$

is continuous at all $a \in \mathcal{S}^m$ for all $m \in \mathbb{R}$. If $a \in \mathcal{S}^\infty(X \times \mathbb{R}^N)$, we write it by

$$a \to \widetilde{\iint}_{H\times\mathbb{R}^N} e^{i\varphi(x,\xi)} a(x,\xi)dxd\xi, \tag{62.59}$$

and it is called an **oscillatory integral**. It's value is given by formula

$$\widetilde{\iint}_{H\times\mathbb{R}^N} e^{i\varphi}a = \int_{H\times\mathbb{R}^N} e^{i\varphi}(L^t)^r a$$

with $r > m+N$. Let $h \in \mathscr{D}(X\times\mathbb{R}^N)$ equals to 1 at (x,ξ) if $x\in H$ and $|\xi|<R$; then lemma 62.33 implies $\lim_{q\to\infty} h_{1/q}a = a$ in the sense of $\mathcal{S}^m$ when $a\in\mathcal{S}^{m'}$ and $m>m'$. Since the oscillatory integral is continuous for each $\mathcal{S}^m$, one can commute the limit and the integral:

$$\widetilde{\iint}_{H\times eR^N} e^{i\varphi}a = \lim_{q\to\infty} e^{i\varphi}h_{1/q}a.$$

62.4.2 Parametric oscillatory integral

Let us introduce a parameter in the integral. We suppose the parameter belongs to an open set $\mathcal{U}\subset\mathbb{R}^N$ and we consider a symbol $a\in\mathcal{S}^m(X\times\mathcal{U}\times\mathbb{R}^N)$. For each $z\in\mathcal{U}$, the function $(x,\xi)\to a(x,z,\xi)$ is a symbol of order m in $X\times\mathbb{R}^N$. For all compact sets $H\subset X$ and $K\subset\mathcal{U}$, there exists a constant C_{HK} (independent of $z\in\mathcal{U}$) such that

$$|D_x^\alpha D_\xi^\beta a(x,z,\xi)| \leqslant C_{HK}(1+|\xi|)^{m-|\beta|} \tag{62.60}$$

for all $(x,z,\xi)\in H\times K\times\mathbb{R}^N$. This comes from the fact that all compact in $X\times\mathcal{U}$ is the product of a compact in X by a compact in $\mathcal{U}$.

Theorem 62.41.
Let X be an open set in $\mathbb{R}^N$, $\mathcal{U}$ an open in $\mathbb{R}^q$ and $\varphi\colon X\times\mathcal{U}\times\mathbb{R}^N\to\mathbb{R}$ be a continuous function, C^∞ in $X\times\mathcal{U}\times\mathbb{R}^N_0$ such that for all $z\in\mathcal{U}$, the function $(x,\xi)\to(x,z,\xi)$ is a phase on $X\times\mathbb{R}^N$. Then

(1) for all $a\in\mathcal{S}^m(X\times\mathcal{U}\times\mathbb{R}^N)$ and all compact $H\subset X$, the map

$$z\to F_a(z) = \widetilde{\iint}_{H\times\mathbb{R}^N} e^{i\varphi(x,z,\xi)}a(x,z,\xi)dxd\xi$$

is C^∞ in $\mathcal{U}$ and $a\to F_a$ is continuous from $\mathcal{S}^m(X\times\mathcal{U}\times\mathbb{R}^N)$ to $\mathscr{E}(\mathcal{U})$,

(2) the function φ is a phase on $X\times\mathcal{U}\times\mathbb{R}^N$ and for all compact $K\subset\mathcal{U}$,

$$\int_K F_a(z)dz = \widetilde{\iint}_{H\times K\times\mathbb{R}^N} e^{i\varphi}a. \tag{62.61}$$ eq:intKFa

Proof. Let us begin by proving that φ is a phase on $X\times\mathcal{U}\times\mathbb{R}^N$. From assumptions it is C^∞ on $X\times\mathcal{U}\times\mathbb{R}^N_0$, the positive homogeneity comes from the fact that for each $z\in\mathcal{U}$, the function $(x,\xi)\to(x,z,\xi)$ is positively homogeneous and the partial derivatives cannot vanish at same point because the one with respect to x and ξ yet doesn't.

Now we define $\varphi_z = \varphi(x,z,\xi)$ and the operator L_z given by lemma 62.39. Then for all $u\in\mathscr{E}(X\times\mathcal{U}\times\mathbb{R}^N)$ and for all $r>0$,

$$\big((L_z^t)^r u\big)(x,z,\xi) = \sum_{|\alpha|+|\beta|} g_{\alpha\beta}(x,z,\xi)D_x^\alpha D_\xi^\beta u(x,z,\xi)$$

where $g_{\alpha\beta}$ is a symbol of order $-r+|\beta|$ in $X\times\mathcal{U}\times\mathbb{R}^N$, see proof of lemma 62.39. Since $L_z e^{i\varphi_z} = e^{i\varphi_z}$, the equation

$$F_a(z) = \int_{H\times\mathbb{R}^N} e^{i\varphi}(L_z^t)^r a$$

holds for all $a\in\mathscr{D}(X\times\mathcal{U}\times\mathbb{R}^N)$. If a is just a symbol, the problem is no more complicated because the definition is

$$\widetilde{\iint}_{H\times\mathbb{R}^N} e^{i\varphi}a = \int_{H\times\mathbb{R}^N} e^{i\varphi}(L_z^t)^a,$$

with $r > m + N$. So we have to prove the continuity of

$$z \to \int_{H\times\mathbb{R}^N} e^{i\varphi}(L_z^t)^r a.$$

We do it by using proposition 46.13. First, for all $z \in \mathcal{U}$, the map $(x,\xi) \to e^{i\varphi(x,z,\xi)}(L_z^t)^r a(x,z,\xi)$ is integrable from our choice of r. Second, the map $z \to e^{i\varphi(x,z,\xi)}(L_z^t)^r a(x,z,\xi)$ is continuous on the whole $\mathcal{U}$. The third assumption of proposition 46.13 is the true point.

$$\begin{aligned}
|e^{i\varphi}(L^t)^r a| &\leqslant |L_z^t a| \\
&= \Big|\sum_{|\alpha|+|\beta|\leqslant r} g_{\alpha\beta}(x,z,\xi) D_x^\alpha D_\xi^\beta a(x,z,\xi)\Big| \\
&\leqslant \sum |g_{\alpha\beta}(x,z,\xi)||D_x^\alpha D_\xi^\beta a(x,z,\xi)| \\
&\leqslant C_{\alpha\beta}C'_{\alpha\beta}(1+|\xi|)^{m-r}
\end{aligned} \tag{62.62}$$

which integrable over $H \times \mathbb{R}^N$ because $r > m + N$.

Now we prove equation (62.61). For this, we write the definition of the oscillatory integral of the right hand side and we compute:

$$\begin{aligned}
\widetilde{\iiint}_{H\times K\times\mathbb{R}^N} e^{i\varphi}a &= \iiint_{H\times K\times\mathbb{R}^N} e^{i\varphi}(L^t)^r a \\
&= \iiint_{H\times K\times\mathbb{R}^N} L^r e^{i\varphi_z} a \\
&= \iiint_{H\times K\times\mathbb{R}^N} L_z^r e^{i\varphi_z} \\
&= \iiint_{H\times K\times\mathbb{R}^N} e^{i\varphi_z(x,\xi)}(L_z^t)^r a(x,z,\xi)dxdzd\xi \\
&= \int_K \widetilde{\iint}_{H\times\mathbb{R}^N} e^{i\varphi_z} a\, dz \\
&= \int_K F_a(z)dz.
\end{aligned} \tag{62.63}$$

We don't give the rest of the proof.

□

We give without proof the formula

$$D_z^\alpha F_a(z) = \widetilde{\iint}_{H\times\mathbb{R}^N} D_z^\alpha(e^{i\varphi}a). \tag{62.64}$$

Now let us consider the situation in which $z \in \mathcal{U}$ is the only variable in front of ξ. More precisely, consider a function $\varphi\colon \mathcal{U}\times\mathbb{R}^N \to \mathbb{R}$ which is C^∞ on $\mathcal{U}\times\mathbb{R}_0^N$ and continuous on $\mathcal{U}\times\mathbb{R}^N$ such that for all $z \in \mathcal{U}$, the function $\xi \to \varphi(z,\xi)$ is positively homogeneous of degree 1, without critical points, i.e. φ_z is a phase in the sense of definition 62.38 with $X = \varnothing$. Then for all symbol $a \in \mathcal{S}^m(\mathcal{U}\times\mathbb{R}^N)$, the map

$$z \to F_a(z) = \widetilde{\int}_{\mathbb{R}^N} e^{i\varphi}a$$

is C^∞ on $\mathcal{U}$ and $a \to F_a$ is continuous from $S^m(\mathcal{U}\times\mathbb{R}^N)$ into $\mathscr{E}(\mathcal{U})$. One can moreover commutes the integral and the derivative

$$D_z^\alpha F_a(z) = \widetilde{\int}_{\mathbb{R}^N} D_z^\alpha\big(e^{i\varphi(z,\xi)}a(z,\xi)\big)d\xi$$

The function φ is a phase on $\mathcal{U}\times\mathbb{R}^N$ and for all compact part K of $\mathcal{U}$, one has

$$\int_K F_a = \widetilde{\iint}_{K\times\mathbb{R}^N} e^{i\varphi}a.$$

62.5 Other constructions of oscillatory integral

62.5.1 Derivation definition

This construction is very the same as the first one, and comes from [952].

We consider a strongly continuous action τ of V on $C_u(V,A)$. Let $\mathcal{B}^A(V)$ be the set of smooth vectors for this action; if $W=V\times V$, the set $\mathcal{B}^A(W)$ is well defined and we choose on V a measure such that the unit cube has measure 1. Let $\varphi\colon V\times V\to\mathbb{R}$, $\varphi(u,v)=2\pi u\cdot v$; we want to give a sense to the integral

$$\iint_{V\times V} F(u,v)e^{i\varphi(u,v)}dudv$$

when $F\in\mathcal{B}^A(W)$. As usual, we first consider the case where F is compactly supported; in this case the integral makes sense with usual definitions. An integration by part gives

$$\iint_{V\times V}(\partial_{k,u}F)e^{i\varphi} = -\iint_{V\times V} 2\pi i v_k F e^{i\varphi}.$$

Taking a second derivative and making the same with the variable v,

$$\iint_{V\times V}(\partial_{sv}\partial_{rv}+\partial_{ku}\partial_{lu})Fe^{i\varphi} = -4\pi^2\iint_{V\times V}(u_su_r+v_kv_l)Fe^{i\varphi};$$

taking $s=r$ and $k=l$, and making sum, we find

$$\iint_{V\times V}(\Delta F)e^{i\varphi} = -4\pi^2\iint_{V\times V}(u\cdot u+v\cdot v)Fe^{i\varphi}. \tag{62.65}$$

eq:int_Delta

Now we pose $w=(u,v)\in W$ and $Qw=u\cdot v$; equation (62.65) reads

$$\int_W[(1-\Delta/4\pi^2)F](w)e^{iQw} = \int_W F(w)(1+w\cdot w)e^{iQw}. \tag{62.66}$$

eq:int_unmDelta

This equation only holds for compact supported functions F. In particular, it is also true for $F'=F/(1+w^2)$ where w^2 stands for $w\cdot w$. For notational convenience, we write $K(w)=(1+w^2)^{-1}$ and M_K the operator of pointwise multiplication by $K(w)$. With F' instead of F, equation (62.66) gives $\int_W(1-\Delta/4\pi^2)(M_KF)e^{iQ}=\int_W Fe^{iQ}$. Making k times all the work,

$$\int_W[(1-\Delta/4\pi^2)M_K]^kFe^{iQ} = \int_W Fe^{iQ}. \tag{62.67}$$

Now let us see how to write $[(1-\Delta/4\pi^2)K]^kF$. Direct computation shows that

$$\Delta(KF) = K(\Delta F-2KF+8K^3w^2F) = K\sum_{|\nu|\leqslant 2}B_\nu\partial^\nu F$$

where B_ν are bounded functions. By induction, we see that

$$[(1-\Delta)K]^kF = K^k\sum_{|\nu|\leqslant 2k}B_\nu\partial^\nu F.$$

So one can write

$$\int_W Fe^{iQ} = \int_W K^k\Big(\sum_{|\nu|\leqslant 2k}B_\nu\partial^\nu F\Big)e^{iQ} \tag{62.68}$$

eq:defintosc

for all k and all function F with compact support. But we know that K^k is integrable when $k > d$, and if F is a smooth vector for the action (62.3), it is bounded on any bounded set (and all the derivatives of F too). So the function $K^k \sum_\nu B_\nu \partial^\nu F e^{iQ}$ is the product of an integrable function by a bounded function; it is then integrable. The conclusion is that the right hand side of equation (62.68) makes sense for all $F \in \mathcal{B}^A(W)$. We take it as *definition* of the left hand side.

62.5.1.1 Value estimations for the integral

Let us suppose that F has a compact support $S(F)$ and consider the real numbers $c_k = \max_{|\nu|\leqslant 2k} \nu! \|B_\nu\|_\infty$ where $\|.\|_\infty$ denotes the usual supremum norm.

lem:born_osci_un

Lemma 62.42.
For all $k > 0$, there exists a constant c_k such that for all $F \in \mathcal{B}^A(W)$ with compact support $S(F)$ and all j,

$$\| \int_W F e^{iQ} \|_j \leqslant c_K \|F\|_{j,2k} \int_{S(F)} K^k$$

Proof. From definition (62.68),

$$\| \int_W (w) F e^{iQw} \|_j \leqslant \int_{S(F)} \|K(w)^k\|_j \sum_{|\nu|\leqslant 2k} \|B_\nu(w) \partial^\nu F(w)\|, \tag{62.69}$$

We can majore $\|B_\nu(w)\|_j$ by $\|B_\nu\|_\infty = \sup_{w\in W} \|B_\nu(w)\|_j$ [4] which can at his turn be majored by c_k; the rest of the sum is precisely $\|F\|_{j,2k}$. Since $K(w) > 0$, we can forget the norm around K^k.

□

Note that when $S(F)$ goes to infinity, the integral $\int_W K^k$ goes to zero and so does the integral $\int_W F e^{iQ}$. Intuitively, it comes from the fact that when $w \to \infty$, the function F doesn't makes anything [5] while function e^{iQ} oscillate rapidly and provokes cancellations.

Let us suppose that $S(F)$ is contained in a ball $B(w_0, r)$ of radius r around the point w_0 and suppose that $|w_0| < r$. In this case, K takes its bigger value when $w^2 = (|w_0| - r)^2$. But there exists a constant c' (depending on r) such that $1 + (|w_0| - r)^2 > c'(1 + |w_0|^2)$. Taking volume of $B(w_0, r)$ into account ,we can compute the following majoration

$$\begin{aligned} \int_{B(w_0,r)} K^k(w) &\leqslant vol\big(B(w_0,r)\big) \left(\frac{1}{1 + (|w_0| - r)^2}\right)^k \\ &\leqslant \left(\frac{vol\big(B(w_0,r)\big)^k}{c'}\right) \left(\frac{1}{1 + |w_0|^2}\right)^k \\ &= cK(w_0)^k \end{aligned} \tag{62.70}$$

where c is a new constant which take into account all other constants. If $|w_0| > r$, one can redefine the constant in such a way that this majoration still holds. Using inequality of lemma 62.42 and absorbing once again all constant in c, we find the

Proposition 62.43.
For all $F \in \mathcal{B}^A(W)$ with support in $B(w_0, r)$ and for all j, there exists a constant c depending only on k and r such that

$$\| \int_W F e^{iQ} \|_j \leqslant c \|F\|_{j,2k} K(w_0)^k. \tag{62.71}$$

In particular, c doesn't depends on j.

4. Faut voir si ce sup est prit aussi sur les j, et je crois bien que oui; en tout cas, je l'utilise
5. Because it is a function in A^∞

62.5.2 Lattice way to define oscillatory integrals

The latter proposition holds for a compact supported function; in order to see what happens when the support becomes non compact, choose a basis of W and let L be the lattice of points with integer components in this basis. Consider a function $\varphi_0 \in C_c^\infty(W)$, a smooth compact supported function on W and $\Phi(w) = \sum_{p\in L} \varphi_0(w+p)$. We choose φ_0 in such a way that Φ vanishes nowhere. Let $\varphi = \varphi_0/\Phi$ and $\varphi_p(w) = \varphi(w+p)$ for $p \in L$. By construction, $\sum_{p\in L} \varphi_p = 1$ and each compact in W intersect only finitely many support of φ_p.

Let $F \in \mathcal{B}^A(W)$; if we suppose $S(\varphi) \subset B(0,r)$, then $S(F,\varphi_p) \subset B(p,r)$ and the proposition can be used with $F\varphi_p$. By virtue of the Leibniz rule on $F\varphi_p$, we see that $\|\partial^\nu(F\varphi_p)\|_j$ is domined by a linear combination of terms of the form $\|\partial^\nu F\|_j \|\partial^\rho \varphi_p\|_\infty$ with $|\mu+\rho| \leqslant |\nu|$. But for all $p \in L$ and all multi-index ρ, $\|\partial^\rho\varphi\|_\infty = \|\partial^\rho\varphi\|_\infty$. Then if we fix r and φ, the value of $\|F\varphi_p\|_{j,2k}$ is bounded by a sum from which $\|\partial^\rho\varphi\|_\infty$ get out while the remaining sum is $\|F\|_{j,2k}$. By redefining the constant c_k, we get

$$\| \int_W F\varphi_p e^{iQ} \| \leqslant c_k K(p)^k \|F\|_{j,2k}. \tag{62.72}$$

eq:pr_abs_conv

We can find a k large enough that $\sum_{p\in L} K(p)^k$. For such a k, formula (62.72) proves the absolute convergence of the sum $\sum_{p\in L} \int_W (F\varphi_p)e^{iQ}$.

def:intFeiQ

Definition 62.44.
When $F \in \mathcal{B}^A(W)$ has no compact support, we define

$$\int_W F e^{iQ} = \sum_{p\in L} \int_W (F\varphi_p) e^{iQ}. \tag{62.73}$$

When F has a compact support, only finitely many terms are non zero and we can permute the sum and the integral. Since for all x, $\sum_{p\in L}(F\varphi_p)(x) = F(X)$, we find back the usual integral. For such a k, we can sum equation (62.72) over $p \in L$ and redefine c_k and find

$$\| \int_W F e^{iQ} \| \leqslant c_k \|F\|_{j,2k}.$$

prop:leqd_m

Proposition 62.45.
For all $k > d$ and all positive sequence $(r_i) \to \infty$, there exists a real sequence $(d_i) \to 0$ such that for all $F \in \mathcal{B}^A(W)$ vanishing on $B(0,r_m)$, the inequality

$$\| \int_W F e^{iQ} \|_j \leqslant d_m \|F\|_{j,2k}.$$

holds.

Stated in a more natural way, the statement is that if $F = 0$ in $B(0,r)$, then there exists a constant d such that $\| \int_W F e^{iQ} \|_j \leqslant d\|F\|_{j,2k}$. The dependence of r with respect to r is such that $lim_{r\to\infty} d(r) = 0$. This corresponds to the intuition that, since e^{iQ}, oscillates more and more rapidly with distance to origin, the contribution to the integral is mainly contained in the values of F near zero.

Proof. Let $E_m = \{p \in L \text{ tel que } S(\varphi_p) \nsubseteq B(0,r_m)\}$. If $p \in E_m$, then φ_p takes non zero values outside r_m. In the expression $\int_W F e^{iQ} = \sum_{p\in L} \int -WF\varphi_p e^{iQ}$, we can reduce the sum because F is zero on $B(o,r_m)$: the terms with $p \notin E_m$ are zero. Then, using (62.72),

$$\| \int_W F e^{iQ} \|_j \leqslant \sum_{p\in E_m} \| \int_W \varphi_p F e^{iQ} \|_j \leqslant \sum_{p\in E_m} c_k K(p)^k \|F\|_{j,2k}.$$

For suitably large k (i.e. $k > d$ as in the assumptions), convergence of the sequence

$$d_m = c_k \sum_{p\in E_m} K(p)^k$$

can be studied by the integral $\int_{\complement B(0,r_m)} K(p)^k dp$ which converges to zero when $m \to \infty$. □

62.5.2.1 Independence with respect to choices

We have to show that definition 62.44 don't depend on the choice of L nor φ. For this, we will find a sequence of numbers independent of L and φ which converges to $\int_W Fe^{iQ}$. Let $(\psi_i) \in C_c^\infty$, a sequence of functions with $\psi_i = 1$ on $B(0, r_i)$ for a certain sequence $(r_m \in \mathbb{R}^+) \to \infty$. For each k, we requires that the functions ψ_i are equibounded for the norm $2k$ in the sense that there exists a real b_k such that $\|\psi_i\|_{2k} \leqslant b_k$ for all i. Such a sequence can be build by choosing ψ_1 and defining $\psi_i(w) = \psi_1(w/r_i)$.

Remark that for each i, the sum $F\psi_i = \sim_{p\in L} F\psi_i\varphi_p$ is pointwise a finite because ψ_i has compact support; then we can permute sum and integral and write

$$\int_W \sum_p F\psi_i\varphi_p e^{iQ} = \sum_p \int_W F\psi_i\varphi_i e^{iQ}.$$

This allows us to write down the following majoration

$$\begin{aligned}
\|\int_W Fe^{iQ} - \int_W Fe^{iQ}\psi_i\|_j &= \|\sum_{p\in L}\int_W F\varphi_p(1-\psi_i)e^{iQ}\|_j && \text{Usual integral} \\
&= \|\int_W F(1-\psi_i)e^{iQ}\|_j && \text{Oscillatory integral} \\
&\leqslant d_i\|F(1-\psi_i)\|_{j,2k} \\
&= d_i\|F\|_{j,2k}(1+b_k).
\end{aligned} \tag{62.74}$$

The latter converges to zero when $i \to \infty$. Then

$$\lim_{m\to\infty}\int_W F\psi_m e^{iQ} = \int_W Fe^{iQ} \tag{62.75}$$

where the left hand side don't depend on L and φ.

One can take this equation as *definition* of the right hand side.

¡¡ Avertissement/question au lecteur !! 62.46
Let $f \in L^2(\mathbb{R}^n)$. Does the integral

$$\int e^{i\xi x} f(x)dx \tag{62.76}$$

always exist in the sense of the oscillatory integral?

If yes, does the value of that integral equal the value of the Fourier transform of f as defined by extension from $L^1 \cap L^2$ by 29.32?

62.6 Decomposition into direct sum

Let $W = W_0 \oplus W_1$ be a direct sum decomposition; we denote by $w = w_0 + w_1$ the corresponding decomposition for the vectors $w \in W$. Let B_i be a basis of W_i and L_i be the corresponding lattice. Since $B_0 \cup B_1$ is a basis of W, we can consider the lattice $L_0 \times L_1$ of W. Let $\varphi \in C_x^\infty$ be a positive function on W_1 with $\sum_{p\in L_0}\varphi_p = 0$ and the same for φ' for W_1. The function $\lambda(w) = \varphi(w_0)\varphi'(w_1)$ belongs to $C_c^\infty(W)$ and since the sums are pointwise finite,

$$\sum_{p\in L}\lambda_p(w) = \sum_{p\in L}\varphi(w_0+p_0)\varphi'(w_1+p_1) = 1.$$

Therefore the function λ has the required properties to define the oscillatory integral

$$\begin{aligned}
\int_W Fe^{iQ} &= \sum_{p\in L}\int_W \lambda_p e^{iQ} \\
&= \sum_{\substack{p_0\in L_0 \\ p_1\in L_1}} F(w)\varphi_{p_0}(w_0)\varphi'_{p_1}(w_1)e^{iQw}dw
\end{aligned} \tag{62.77}$$

eq:osc_int_phip

with an absolutely convergent sum, there are no ordering problem in the summation.

Proposition 62.47.
Let $k > d$ and an increasing sequence $(r_i \in \mathbb{R}^+) \to \infty$. Then there exists a sequence $(d_i \in \mathbb{R}) \to 0$ such that, if $F \in \mathcal{B}^A(V \times V)$ and $F(u,v) = 0$ when $u \in B(0, r_m)$, then

$$\| \iint_{V\times V} e^{i\phi(u,v)} \|_j \leqslant d_m \|F\|_{j,2k}$$

Here, $\phi(u,v) = 2\pi u \cdot v$ is the former φ.

Proof. Let $E_m = \{p \text{ tel que } S(\varphi_p) \nsubseteq B(o, r_m)\}$. The computation is performed in rather the same way as in proposition 62.45:

$$\| \iint_{V\times V} Fe^{iQ} \|_j = \| \sum_{p\in E_m} \sum_q F(u,v)\lambda_{p+q}(u+v)e^{i\phi} \| \tag{62.78}$$

where $\lambda_{p+q}(u+v) = \varphi_p(u)\varphi'_q(v)$ because (u,v) and (p,q) are seen as vectors of W. If φ has support contained in $B(r,0)$, the function $F\varphi_p\varphi'_q$ has support contained in the ball centered at $p+q$; the proposition 62.42 gives

$$\| \iint_{V\times V} F\varphi_p\varphi_q \| \leqslant c_k \|F\varphi_p\varphi'_q\| K(p+q)^k.$$

If we suppose that the sum $W = W_0 \oplus W_1$ is orthogonal, $K(p+q) = (1 + p \cdot p + q \cdot q)$, a simple redefinition of c_k, leads to $c_k\|F\varphi_p\varphi'_q\| \leqslant c_k\|F\|$. So we can majore equation (62.78) by

$$\sum_{p\in E_m} \sum_q c_k \|F\|_{j,2k}(1 + p^2 + q^2)^{-k}.$$

Now the definition $d_m = \sum_{p\in E_m} \sum_q c_k(1 + p^2 + q^2)^{-k}$ gives the thesis. □

Proposition 62.48.
Let us consider a sequence $(\psi_i) \in C^\infty(W_0)$ with $\psi_i = 1$ on $B(0, r_i)$ and $\|\psi_i\|_{2k} \leqslant b_k$ for all i. Then for all $F \in \mathcal{B}^A(W)$, we have

$$\int Fe^{iQ} = \lim_{m\to\infty} \int F(w)\psi_m(w_0)e^{iQw}. \tag{62.79}$$

Proof. We consider φ and φ' as in equation (62.77) and we write φ_p instead of $\varphi_p \circ \mathtt{proj}_0$ where $\mathtt{proj}_0$ is the projection into W_0. Then

$$\begin{aligned} \int_W Fe^{iQ} - \int_W (\psi_m \circ \mathtt{proj}_0)e^{iQ} &= \sum_{p,q} \int_W F\varphi_p\varphi'_q e^{iQ} - \sum_{p,q} F(\psi_m \circ \mathtt{proj}_0)\varphi_p\varphi'_q e^{iQ} \\ &= \sum_{p;q} \int_W F(1 - \psi_m \mathtt{proj}_0)\varphi_p\varphi'_q e^{iQ} \end{aligned} \tag{62.80}$$

which converges absolutely. By setting $E_m = \{p \in L_0 \text{ tel que } S(\varphi_p) \nsubseteq B(0, r_m)\}$, we can reduce the sum over p to E_m and majore the norm of the latter expression by $\sum_{p\in E_m} \sum_q K(p+q)^k$. It converges to 0 when $m \to \infty$.

□

Let us now suppose that we have functions (ψ'_i) on W_1 with the usual properties. Writing the latter proposition with $F\psi'_n$ instead of F, we find that, for all b,

$$\lim_{m\to\infty} \int_W F(w)\psi_m(w_0)\psi'_n(w_1)e^{iQw} = \int_W F(w)\psi'_n(w_1)e^{iQw}$$

and inverting W_1 with W_2 in the preceding proposition, we see that the right hand side converges to 0 when $n \to \infty$. So we find

$$\lim_{m,n\to\infty} \int_W F(\psi_m \circ \mathtt{proj}_0)(\psi'_n \circ \mathtt{proj}_1)e^{iQ} = \int_W Fe^{iQ}, \tag{62.81}$$

and we can take the limit of n before m if we want.

62.6.1 Integral over $V \times V$

Let an orthogonal direct sum decomposition $V = V_0 \oplus V_1$ The decomposition $W = V_0 \times V_1 \times V_0 \times V_1$ is naturally orthogonal. Consider an usual function $F\colon V \times V \to A$ such that $F\big((v_0+v_1),(v_0',v_1')\big)$ don't depend on v_0.

Let (ψ_i) and (ψ_j) be sequences of functions such that equation (62.81) gives

$$\iint F(u,v)e^{iQ(u,v)} = \lim_{mnjk} \iint F(u_1,v)\psi_m(u_0)\psi_n'(u_1)\psi_j(v_0)\psi_k'(v_1)e^{u_0 \cdot v_0 + u_1 \cdot v_1}. \tag{62.82}$$

eq:run

We define the **Fourier transform** of ψ_m by

$$\hat{\psi}_m(\xi) = \int_{V_0} \psi_m(u_0)e^{2\pi i u_0 \cdot \xi} du_0.$$

The main property is the inverse transform theorem:

$$\psi_m(v) = \int_{V_0} \hat{\psi}_m(\xi)e^{-2\pi i \xi \cdot v} d\xi = \psi_m(v).$$

Equation (62.82) can be rewritten as

$$\lim_{mnjk} \iint_{V\times V} F(u_1,v) \int_V \hat{\psi}_m(\xi)e^{-2\pi i\xi \cdot u_0}\psi_n'(u_1)\psi_j(v_0)\psi_k'(v_1)$$
$$e^{2\pi i(u_0 \cdot v_0 + u_1 \cdot v_1)} du dv du d\xi$$

$$= \lim_{mnjk} \iiint_{V\times V\times V} F(u_1, v_1+v_0)\hat{\psi}_m(\xi)\psi_n'(u_1)\psi_j(v_0)\psi_k'(v_1)$$
$$e^{2\pi i[u_0 \cdot (v_1-\xi) + u_1 \cdot v_1]} dv du dv d\xi.$$

The integral over u_0 gives a Dirac delta at $\xi - v_0$

$$= \lim_{mnjk} \iint F(u_1, v_1+v_1)\hat{\psi}_m(v_0)\psi_n'(u_1)\psi_j(v_0)\psi_k'(v_1)e^{2\pi i u_1 \cdot v_1} du dv. \tag{62.83}$$

eq:rdeux

We can suppose that the functions ψ_m are chosen in such a way that $(\hat{\psi}_m)$ is an approximation of the delta distribution. Indeed, if we choose ψ_1 equals to 1 in $B(0,r_1)$, the sequence $\psi_m \to 1$ whose Fourier transform is the Dirac delta. In this case, the limit $m \to \infty$ gives $v_0 = 0$ and the limit $j \to \infty$ becomes trivial. Equation (62.83) becomes

$$\begin{aligned} &\lim_{nk} \iint F(u_1,v_1)\hat{\psi}(v_0)\psi_n'(u_1)\psi_k'(v_1)e^{2\pi i u_1 \cdot v_1} \\ &= \iint_{V_1\times V_1} F(u_1,v_1)e^{iQ(u_1,v_1)}. \end{aligned} \tag{62.84}$$

In definitive, we had shown

prop:F_pas_uz

Proposition 62.49.
Let $F \in \mathcal{B}^A(V\times V)$ and an orthogonal decomposition $V = V_0 \oplus V_1$ such that $F(u_0+u_1, v_0+v_1)$ don't depend on u_0. Then

$$\iint_{V\times V} Fe^{iQ} = \iint_{V_1\times V_1} F(u_1, v-1)e^{iQ(u_1,v_1)}.$$

There exists a somewhat degenerate interesting case: $V_1 = \{0\}$ with a measure 1. In this case the assumption is that $F(u,v)$ don't depend on u at all. In this case,

$$\iint\limits_{V_1\times V_1} F(u_1,v_1)e^{2\pi i u_1 \cdot v_1} = \int_{V_1} F(v_1)dv_1 = F(0).$$

The result is that if $F \in \mathcal{B}^A(V\times V)$ don't depend on his first variable (i.e. $F \in \mathcal{B}^A(V)$), then

$$\int_{V\times V} Fe^{iQ} = F(0).$$

62.6.2 Normal operator

Let T be a linear operator on V and a decomposition $V = \ker T \oplus V_1$. In particular, $F(Tu,v)$ doesn't depends on u_O (if we denote $\ker T$ by V_0) and we can apply proposition 62.49:

$$\int F(T_u,v)e^{iQ} = \iint\limits_{V_1\times V_1} F(Tu_1,v_1)e^{iQ(u_1,v_1)}.$$

Note that in general, Tu_1 don't belong to V_1. We now suppose that T is **normal** in the sense that $[T,T^t]=0$. It is invertible on V_1. Then when we have a normal operator in an integral, one can always suppose that it is invertible because the kernel part of the space disappears.

Let T be invertible on V, $F\in\mathcal{B}^A(V\times V)$ and an usual sequence (ψ_m) of functions. We have

$$\begin{aligned}\iint F(Tu,v)r^{iQ(u,v)} &= \lim_{m,n\to\infty}\iint F(Tu,v)\psi_m(u)\psi_n(v)e^{iQ}\\ &= \lim_{m,n\to\infty}(\det T)^{-1}\iint F(u,v)\psi_m(T^{-1}u)\psi_n(v)e^{2\pi i T^{-1}u\cdot v}\\ &= \lim_{m,n\to\infty}\iint F(u,T^tv)\psi_m(T^{-1}u)\psi_n(T^tv)e^{iQ}\end{aligned}$$

but the sequences $(\psi_m\circ T^{-1})$ and $(\psi_n\circ T^t)$ have same fundamental properties as (ψ_m), then taking the limit,

$$= \iint F(u,T^tv)e^{iQ}. \tag{62.85}$$

If T is normal but not invertible, the first integral restricts to V_1 and for all normal operator, the equality [6]

$$\iint F(Tu,v)e^{iQ} = \iint F(u,T^tv)e^{iQ}.$$

hold. From polar decomposition of any operator into two normal operators, we conclude that this equation holds for all linear operator on V.

Proposition 62.50.
If $S\colon A\to C$ is a continuous linear map from A to a Fréchet space C, then $S\circ F\in\mathcal{B}^C(W)$ and

$$S(\int_W Fe^{iQ}) = \int_W (S\circ F)e^{iQ}. \tag{62.86}$$

Proof. The absolute convergence of the defining sum for the left hand side integral allows us to permute linear operator and sum. Since for each term the integral is an usual integral over a compact set, we can permute the continuous operator S and the integral:

$$S\sum_{p\in L}\int F\varphi_p e^{iQ} = \sum S\int F\varphi_p e^{iQ} = \sum\int (S\circ F)\varphi_p e^{iQ}. \tag{62.87}$$

□

6. Je ne comprends pas pourquoi.

62.7 Pseudo-differentiable operators

62.7.1 Differential operator

Let $\pi\colon E \to M$ be a vector bundle of rank k on a compact manifold of dimension n. We denote by $\Gamma(E)$ the $C^\infty(M)$-module of C^∞ sections of E. The module structure is given by $(f \cdot \psi)(x) = f(x)\psi(x)$ where the product in the right hand side is the product $\mathbb{C} \times E_x \to E_x$ which defines the vector space structure on $E_x = \pi^{-1}(x)$.

A **differential operator** of rank m is a linear operator $P\colon \Gamma(E) \to \Gamma(E)$ for which there exists local coordinates $(x_1, \cdots, x_n)$ on M in which P is written as

$$P = \sum_{|\alpha| \leqslant m} A_\alpha(x)(-i)^{|\alpha|} \frac{\partial^{|\alpha|}}{\partial x^\alpha} \tag{62.88}$$

where $\alpha = (\alpha_1, \cdots, \alpha_n)$ is a multi-index with $0 \leqslant \alpha_j \leqslant m$ and $|\alpha| = \sum_{j=1}^n \alpha_j$. Each A_α is a $k \times k$ matrix of C^∞ functions on M and $A_\alpha \neq 0$ for at least one α with $|\alpha| = m$.

Let $\xi \in T_x^*M$ written under the form $\xi = \sum_j \xi_j\, dx_j$ in the previously given local coordinates. The **complete symbol** of P applied to ξ is the following polynomial combination of coordinates of ξ:

$$p^P(x, \xi) = \sum_{j=0}^{m} p^P_{m-j}(x, \xi) \tag{62.89}$$

where

$$p^P_{m-j} = \sum_{|\alpha| \leqslant (m-j)} A_\alpha(x)\xi^\alpha$$

where $\xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$ and $\partial_\alpha^{|\alpha|} = \partial_{x_1}^{\alpha_1} \circ \cdots \circ \partial_{x_n}^{\alpha_n}$. The **principal symbol** of P is the leading term, namely

$$\sigma^P(x, \xi) = p^P_m(x, \xi) = \sum_{|\alpha| = m} A_\alpha(x)\xi^\alpha. \tag{62.90}$$

For each $\xi \in T_x^*M$, the principal symbol gives rise to a map $\sigma_x^P\colon E_x \to E_x$ because $A_\alpha(x)$ is a $k \times k$ matrix:

$$\sigma_x^P(\xi)v = \sum_{|\alpha| = m} \xi^\alpha A_\alpha(x)v.$$

DefGLpDEHy

Definition 62.51.
*The operator P is **elliptic** if for each non zero $\xi \in T^*M$, the map $\sigma_x^P(\xi_x)$ are all invertible.*

If (M, g) is a Riemannian manifold, the fact for P to be elliptic is equivalent to be invertible on the **cosphere**

$$S^*M = \{(x, \xi) \in T^*M \text{ tel que } g_x(\xi, \xi) = 1\} \subset T^*M.$$

62.7.2 Laplace operator

As first example, we see the Laplace operator. Let (M, g) be a Riemannian manifold and its Laplace operator $\Delta\colon C^\infty(L) \to C^\infty(M)$ where $C^\infty(M)$ is seen as $\Gamma(\mathbb{C} \times M)$ is

$$\Delta f = -g^{\mu\nu}\partial^2_{\mu\nu} f + \text{lower order terms}$$

Its principal symbol is $\sigma^\Delta(x, \xi) = g^{\mu\nu}\xi_\mu\xi_\nu = \|\xi\|^2$, which is invertible of all $\xi \neq 0$. Thus Laplace operator is elliptic of order 2.

An other example of pseudo-differential operator is the Dirac operator. We will show that it is an elliptic operator in subsection 68.7.5.

62.8 Invariant differential operators on Lie groups

This section is intended to understand the paper [892]. Let G be a connected Lie group and $\mathfrak{g}$ be its Lie algebra.

An endomorphism $P\colon C^\infty(G,A)\to C^\infty(G,A)$ of the space of functions over G with values in a vector space A is said to be **left invariant** if

$$L_x(P\psi)=P(L_x\psi) \tag{62.91}$$

EqDefLxinvarop

for all $x\in G$ and $\psi\in C^\infty(G,A)$. Here, L is the regular left representation of G on $C^\infty(G,A)$.

Lemma 62.52.
If $X\in\mathfrak{g}$ and if $\tilde{X}$ is the associated left invariant vector field, the operator P defined by

$$(P\psi)(x)=\tilde{X}_x\psi \tag{62.92}$$

is left invariant.

Proof. On the one hand

$$\big(L_x(P\psi)\big)(y)=(P\psi)(xy)=\tilde{X}_{xy}\psi,$$

on the other hand,

$$P(L_x\psi)(y)=\tilde{X}_y(L_x\psi)=\frac{d}{dt}(L_x\psi)(ye^{tX})\Big|_{t=0}=\frac{d}{dt}\psi(xye^{tX})\Big|_{t=0}=\tilde{X}_{xy}\psi. \tag{62.93}$$

□

In fact, if $\mathrm{Diff}^G(G)$ denotes the space of G-left invariant differential operators on G, we have a morphism

$$\begin{aligned}\mathcal{U}(\mathfrak{g})&\to\mathrm{Diff}^G(G)\\ X&\mapsto\tilde{X}.\end{aligned} \tag{62.94}$$

EqMorphismUgDiff

The operator P is **left invariant** when it satisfies

$$(Pf)(x)=P\big(L_x^*f\big)(e). \tag{62.95}$$

As far as notations are concerned, we denote by (L_x^*P) the operator defined by

$$(L_x^*P)(f)=P(L_x^*f). \tag{62.96}$$

Every differential operator $P\colon C^\infty(G)\to C^\infty(G)$ can be written as a sum of products $\sum_i p_i\tilde{X}_i$ of functions $p_i\in C^\infty(G)$ by operators $\tilde{X}_i\in\mathcal{U}(\mathfrak{g})$. The operator P is left invariant if and only if the functions p_i are constant.

As a consequence we have the

CorUisomDiff

Corollary 62.53.
The map $\mathcal{U}(\mathfrak{g})\to\mathrm{Diff}^G(G)$ is an isomorphism.

We denote by $\text{bi-Diff}^G(G)$ the space of G-left invariant **bidifferential operators** on G. These are differential operators

$$\begin{aligned}A\colon C^\infty(G)\times C^\infty(G)&\to C^\infty(G)\\ u\otimes v&\mapsto A(u\otimes v).\end{aligned} \tag{62.97}$$

Since A is a differential operator, the function $A(u,v)$ only depends on the class $u\otimes v$. The left invariance means that

$$L_x\big(A(u\otimes v)\big)=A\big(L_x(u\otimes v)\big) \tag{62.98}$$

EqDefLeftInvarbiDiff

for every $x\in G$.

PropbidiffUU

Proposition 62.54.
The space of left invariant bidifferential operators bi-Diff$^G(G)$ *is canonically isomorphic to the tensor product* $\mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})$.

Proof. Let A be a bidifferential operator and $\{X_i\}$ be a basis of $\mathfrak{g}$. For every multi-index $a = (a_1, \dots, a_n)$ we write $X^a = X^{a_1} \cdots X^{a_n}$ the corresponding element in $\mathcal{U}(\mathfrak{g})$. Thus there exist functions $A_{ab} \in C^\infty(G)$ such that

$$A(u \otimes v) = \sum_{ab} A_{ab} (\tilde{X}^a u)(\tilde{X}^b v) \tag{62.99}$$

EqdefAabdiffopgp

for every $u, v \in C^\infty(G)$.

Now we are supposing that A is left invariant and we will prove that the functions $A_{ab}(x)$ are constant. We evaluate the definition (62.98) at $g \in G$:

$$\begin{aligned} A(u \otimes v)(xg) &= A\big(L_x(u \otimes v)\big)(g) \\ &= A\big((L_x^* u) \otimes (L_x^* v)\big)(g) \\ &= A_{ab}(g) \tilde{X}_g^a (L_x^* u) \tilde{X}_g^b (L_x^* v). \end{aligned} \tag{62.100}$$

If we consider that equation at $g = e$, we have

$$A(u \otimes v)(x) = A_{ab}(e) \tilde{X}_e^a (L_x^* u) \tilde{X}_e^b (L_x^* v) = A_{ab}(e) \tilde{X}_x^a u \tilde{X}_x^b v \tag{62.101}$$

where we used the invariance property (52.23) on each of the operators $X^a \in \mathcal{U}(\mathfrak{g})$. Now, the expression (62.99) evaluated at x gives

$$A(u \otimes v)(x) = A_{ab}(x)(\tilde{X}_x^a u)(\tilde{X}_x^b v), \tag{62.102}$$

so that $A_{ab}(x) = A_{ab}(e)$ when A is left invariant. □

Alternative proof. Using the same notations, we have

$$A(u \otimes v)(x) = A_{ab}(x)\big(\tilde{X}_x^a \otimes \tilde{X}_x^b\big)(u \otimes v) = A_{ab}(x)\big(\tilde{X}_x^a u\big)\big(\tilde{X}_x^b v\big). \tag{62.103}$$

EaAuvDeuxDun

Using the left invariance of A, on the other hand, we have

$$A(u \otimes v)(x) = A\big(L_x^* u \otimes L_x^* v\big)(e) = A_{ab}(e)\big(\tilde{X}_e^a L_x^* u\big) \otimes \big(\tilde{X}_e^b L_x^* v\big), \tag{62.104}$$

EqAuvUnDun

but

$$\tilde{X}_e^a L_x^* u = \frac{d}{dt}\Big[(L_x^* u)(e^{t\tilde{X}^a})\Big]_{t=0} = \frac{d}{dt}\Big[u\big(x e^{t\tilde{X}^a}\big)\Big]_{t=0} = \tilde{X}_x^a u, \tag{62.105}$$

thus the expression (62.104) becomes

$$A_{ab}(e)\big(\tilde{X}_x^a u\big)\big(\tilde{X}_x^a v\big), \tag{62.106}$$

and the comparison with (62.103) shows that $A_{ab}(x) = A_{ab}(e)$, so that the coefficients A_{ab} are constant. □

REMooGIFYooTphiex

Remarque 62.55.
If $\tilde{X}$ is the left invariant differential operator associated with $X \in \mathfrak{g}$, the Leibniz rules reads

$$\tilde{X}(fg) = \widetilde{\Delta(X)}(f \otimes g) \tag{62.107}$$

EqXfgDeltaUnif

where Δ is the coproduct defined in 54.6.5.

Let us now consider the space $\mathrm{Diff}(G\times G)$ of differential operators on $G\times G$. These operators an operator in $\mathrm{Diff}(G\times G)$ acts on the space of functions $C^\infty(G\times G)$. Such an operator reads $X\cdot Y$ with $X,Y\in\mathcal{U}(\mathfrak{g})$. If $f\in C^\infty(G\times G)$,

$$(Pf)(x,y)=(X\cdot Yf)(x,y) \tag{62.108}$$

where X acts on the first variable and Y acts on the second variable. The space $\mathrm{Diff}(G\times G)$ is then naturally isomorphic to $\mathcal{U}(\mathfrak{g})\otimes\mathcal{U}(\mathfrak{g})$ as an operator $P\in\mathrm{Diff}(G\times G)$ can be written as $P=X\otimes Y$. This acts on tensor product of functions by

$$P(u\otimes v)=Xu\otimes Yv. \tag{62.109}$$

Since they are both isomorphic to $\mathcal{U}(\mathfrak{g})\otimes\mathcal{U}(\mathfrak{g})$, the spaces of left invariant operators $\mathrm{Diff}^{G\times G}(G\times G)$ and bi-$\mathrm{Diff}^G(G)$ are isomorphic.

Let us make that isomorphism more explicit. First, an element $P\in\mathrm{Diff}(G\times G)$ provides the element $X\otimes Y\in\mathcal{U}(\mathfrak{g})\otimes\mathcal{U}(\mathfrak{g})$. The latter produces the bidifferential operator $A\in\text{bi-}\mathrm{Diff}(G)$ defined by $A(u\otimes v)(x)=X_xuX_xv$. Thus we define

$$\begin{aligned}\alpha\colon\ &\mathrm{Diff}(G\times G)\to\text{bi-}\mathrm{Diff}(G)\\ &\alpha(P)(u\otimes v)(x)=X_xuY_xv.\end{aligned} \tag{62.110}$$

EqDefAlphaDiffbiDiff

Lemma 62.56.
The map (62.110) is surjective.

Proof. If $A\in\text{bi-}\mathrm{Diff}(G)$ is given by $A(u\otimes v)(x)=X_xuY_xv$, we have $\alpha(P_A)=A$ with $P_A\in\mathrm{Diff}(G\times G)$ given on $f\in C^\infty(G\times G)$ by

$$(P_Af)(x,y)=(XYf)(x,y) \tag{62.111}$$

where X acts on the first variable and Y on the second. □

Proposition 62.57.
If we restrict to the left invariant operators, the map $\alpha\colon\mathrm{Diff}^{G\times G}(G\times G)\to\text{bi-}\mathrm{Diff}^G(G)$ is an isomorphism.

Proof. First, remark that if $A\in\text{bi-}\mathrm{Diff}(G)$ is invariant, then an operator P_A such that $\alpha(P_A)=A$ is also invariant because of proposition 62.54 which states that

$$A(u\otimes v)(x)=A_{ab}(x)(\tilde{X}^a_x\otimes\tilde{X}^b_x)(u\otimes v) \tag{62.112}$$

is invariant only when A_{ab} are constant functions. Thus the map α is surjective from $\mathrm{Diff}^{G\times G}(G\times G)$ to bi-$\mathrm{Diff}^G(G)$.

Now we have to prove that the map α is injective from the subspace of invariant operators. Let $P\in\mathrm{Diff}(G\times G)$ be an operator in the kernel of α, so if P reads

$$(Pf)(x,y)=c(x,y)(\tilde{X}\tilde{Y}f)(x,y), \tag{62.113}$$

we have

$$(\alpha P)(u\otimes v)(x)=c(x,x)\tilde{X}_xu\tilde{Y}_xv=0 \tag{62.114}$$

for every $x\in G$ and every $u,v\in C^\infty(B)$. In that case, $c(x,x)$ has to vanish.

What we proved is that the kernel of α is the set of operators with $c(x,x)=0$. In other words, the kernel is the set of operators P such that

$$(Pf)(x,x)=0 \tag{62.115}$$

EqPfxxzero

for every function $f\in C^\infty(G\times G)$.

Let us now suppose that P is left invariant and compute $(Pg)(x,y)$ using the left invariance. We have

$$(Pg)(x,y)=dL_{x\times y}(Pg)(e,e)=P(L^*_{x\times y}g)(e,e)=0 \tag{62.116}$$

because of (62.115) applied to the function $f=L^*_{x\times y}g$. □

62.9 Pseudo differential operators

62.9.1 Composition

If A and B are pseudo-differential operators with symbols a and b, the symbol of the composition is given by

$$c(x,\xi) \sim \sum_{\alpha\in\mathbb{N}^n} \frac{(-i)^{|\alpha|}}{\alpha!}\partial_\xi^\alpha a \partial_x^\alpha b. \tag{62.117}$$ EqCompPSDSymb

62.9.2 Fourier transform and pseudo-differential operator

Let $\psi\colon M \to E$ be a local section and P a differential operator. One can prove that eq:PspiFour

$$(P\psi)(x) = \frac{1}{(2\pi)^{n/2}}\int e^{i\xi\,\cdot\,x} p(x,\xi)\hat{\psi}(\xi)\,d\xi \tag{62.118a}$$

$$\hat{\psi}(\xi) = \frac{1}{(2\pi)^{n/2}}\int e^{-i\xi\,\cdot\,y}\psi(y)\,dy \tag{62.118b}$$

where $\xi\cdot x = \sum_j \xi_j x_j$ is the usual scalar product. So a pseudo-differential operator has to be written under the form

$$(Au)(x) = (2\pi)^n \iint e^{i(x-y)\,\cdot\,\xi} a(x,\xi)u(y)\,dy\,d\xi \tag{62.119}$$

and a is the **symbol** of A.

Let us see a simple example: $P\psi = \partial_1\psi$. In this case, $p(x,\xi) = \xi_1$ and equations (62.118) give

$$(P\psi)(x) = (2\pi)^{-n}\iint e^{i\xi\,\cdot\,(x-y)}\xi^1\psi(y)\,d\xi\,dy,$$

an integration by part gives

$$(2\pi)^{-n}\iint e^{i\xi\,\cdot\,(x-y)}(-\frac{i}{\xi_1})\xi_1(\partial_{y_1}\psi)(y)\,d\xi\,dy.$$

The integration over ξ produce de Dirac distribution centred at $x-y$, i.e. a factor $(2\pi)^{-n}\delta(x-y)$ and the integral over y leads to $(\partial_1)\psi(x)$.

To define pseudo-differential operator, we begin by only consider the trivial vector bundle over $\mathbb{R}^n$ and thus functions $u\colon \mathbb{R}^n \to \mathbb{C}^k$.

Definition 62.58.
*A **pseudo-differential** operator of order m is an operator P which can be written under the form*

$$(Pu)(x) = (2\pi)^{-n/2}\int e^{i\xi\,\cdot\,x} p(x,\xi)\hat{u}(\xi)\,d\xi \tag{62.120}$$

where

$$\hat{u}(\xi) = (2\pi)^{-n/2}\int e^{-i\xi\,\cdot\,y} u(y)\,dy \tag{62.121}$$

and $p \in \mathcal{S}^m$, the space defined at page 3535.

The set of all pseudo-differential operators is denoted by $\mathcal{P}^m$. If $p\in\mathcal{S}^m$ is the symbol associated with $P\in\mathcal{P}^m$, the **principal symbol** of P is the class $\sigma^P = [p] \in \mathcal{S}^m/\mathcal{S}^{m-1}$. Operators with associated symbol in $\mathcal{S}^\infty$ are call **smoothing operator** and we denote their space by $\mathcal{P}^{-\infty}$.

A pseudo-differential operator Q is a **parametrix** for the pseudo-differential operator P is the operators $PQ-\mathbb{1}$ and $QP-\mathbb{1}$ are compact operators.

62.9.3 Example

pg_exem_psdo

The operator $[1-(2\pi)^{-2}\Delta]^{1/2}$ is the pseudo-differential operator on $\mathbb{R}^N$ with symbol $p(\xi) = (1+\xi^2)^{1/2}$ where ξ^2 stands for $|\xi|^2$. Indeed if P is the pseudo-differential operator with this symbol

$$\begin{aligned}(Pu)(x) &= \int e^{2i\pi\xi\,\cdot\,x}(1+\xi^2)^{1/2}\hat{u}(\xi)\,d\xi \\ &= \int e^{2i\pi\xi\,\cdot\,x}\int e^{-2i\pi\xi\,\cdot\,y}[1-(2\pi)^{-2}\Delta]^{1/2}u(y)\,dy\,d\xi \quad \text{by (61.150)} \\ &= \int e^{2i\pi\xi\,\cdot\,(x-y)} \\ &= \int \delta(x-y)[1-(2\pi)^{-2}\Delta]^{1/2}u(y)\,dy \\ &= [1-(2\pi)^{-2}\Delta]^{1/2}u(x).\end{aligned}$$

62.9.4 Trace operators

subsec_traceop

A pseudo-differential operator P is expressed by its symbol p and the formula

$$(Pu)(x) = \int e^{2i\pi\xi\,\cdot\,x}p(x,\xi)\hat{y}(\xi)\,d\xi.$$

Assume that P can also be written under the form

$$(Pu)(x) = \int K(x,y)u(y)\,dy$$

and let us equalize these two expressions for all u (for example in $H^{1/2}(\mathbb{R}^n)$. When it makes sense,

$$K(x,y) = \int e^{2i\pi\xi\,\cdot\,(x-y)}p(x,\xi)\,d\xi. \tag{62.122}$$

The **trace** of P is defined by

$$\operatorname{Tr} P = \int K(x,y)\,dx = \iint p(x,\xi)\,dx\,d\xi. \tag{62.123}$$

When the latter converges, one says that P is a *trace operator*.

62.9.5 Asymptotic expansions

We say that the pseudo-differential operator A is **classical** and we write $A \in \Psi^p(M)$ if the symbols accepts the asymptotic expansion

$$a(x,\xi) \sim \sum_{j=0}^{\infty} a_{p-j}(x,\xi)$$

where each of the a_r is r-homogeneous with respect to ξ: $a_r(x,t\xi) = t^r a_r(x,\xi)$. Although the whole definitions are made in local coordinates, one can show that the **principal symbol** is a globally defined function over the cotangent bundle T^*M. The operator is elliptic if $a_n(x,\xi)$ is invertible for all $\xi \neq 0$.

Notice that an element of Ψ^p has an asymptotic expansion which begins with order p, so that the spaces Ψ^d fulfil $\Psi^{p-1} \subset \Psi^p$. Now the **algebra of classical pseudo-differential operators** is the quotient

$$\mathcal{P} = \Psi^{-\infty}/\Psi^{\infty}. \tag{62.124}$$

EqDefmPalgQpeClass

62.10 Dirac and Laplace type operators

Let $(M, \partial M)$ be a manifold with boundary and V, a vector bundle over M. A second order differential operator Δ on V is of **Laplace type** if its principal symbol is the metric tensor, i.e. if it has a local expression of the form

$$\Delta = -(g^{ij}\partial^2_{ij} + A^k\partial_k + B) \tag{62.125}$$

where $a \in \Gamma(TM \otimes \mathrm{End}(V))$ and $B \in \Gamma(\mathrm{End}(V))$. A first order operator D on $\Gamma(V)$ is of **Dirac type** if its principal symbol defines a Clifford module over V. Such an operator has a local expression

$$D = \gamma^i\partial_i - r \tag{62.126}$$

EqFormGeneDirac

with $\gamma \in \Gamma(TM \otimes \mathrm{End}(V))$ and $r \in \Gamma(\mathrm{End}(V))$. The condition is that

$$\gamma^i\gamma^j + \gamma^j\gamma^i = 2g^{ij}\,\mathrm{Id}\,|_V. \tag{62.127}$$

Proposition 62.59.
The operator D is of Dirac type if and only if D^2 is of Laplace type.

Proof. If D is of Dirac type, up to lower order terms we have

$$D^2 = \gamma^i\gamma^j\partial^2_{ij} + \ldots = \frac{1}{2}(\gamma^i\gamma^j + \gamma^j\gamma^i)\partial^2_{ij} + \ldots = -g^{ij}\partial^2_{ij} + \ldots$$

□

It is however not true that every Laplace type operator is the square of a Dirac operator.

62.11 Wodzicki residue

Let M be a n-dimensional manifold. In that case the term of order $(-n)$ in the asymptotic expansion

$$a(x,\xi) \sim \sum_{j=0}^{\infty} a_{n-j}(x,\xi)$$

of the principal symbol of the pseudo-differential operator A is specially important. Let $\mathcal{U} \subset M$ be a coordinate open set on which the cotangent bundle is trivial. One can see a_{-n} as a smooth function on $T^*\mathcal{U}^\times$ (the set of sections from which we remove the zero section).

We define

$$\alpha(x,\xi) = a_{-n}(x,\xi)d\xi_1 \wedge \cdots \wedge d\xi_n \wedge dx^1 \wedge \cdots \wedge dx^n.$$

That form is invariant under dilatations $\xi \to t\xi$ because $a_{-n}(x,t\xi) = t^{-n}a_{-n}(x,\xi)$. Let us consider $R = \sum_j \xi_j\partial_{\xi_j}$, the dilatation generator.

From formula $\mathcal{L}_R = \iota_R d + d\iota_R$, we have $d\iota_R\alpha = \mathcal{L}_R\alpha = 0$ where $\mathcal{L}$ denotes the Lie derivative (49.355). If we write $dx = d^n x = dx^1 \wedge \cdots \wedge dx^n$, we have

$$(\iota_R\alpha)(x,\xi) = a_{-n}(x,\xi)\sigma_n \wedge dx$$

where $\sigma_\xi = \sum(-1)^{j-1}\xi_j\, d\xi_1 \wedge \cdots \widehat{d\xi_j} \wedge \cdots \wedge d\xi_n$. For each particular $x \in M$, σ_ξ is a volume form on the unit sphere $|\xi| = 1$ on $T^*_x M$. We can integrate $\iota_R\alpha$ on such a sphere:

$$\int_{|\xi|=1} a_{-n}(x,\xi)\sigma_\xi.$$

One can prove that under the coordinate change $w \to y = \phi(x)$, $\xi \to \eta = \phi'(x)^t\xi$, we have $a_{-n}(x,\xi) \to \tilde{a}_{-n}(y,\eta)$ and

$$\int_{|\eta|=1} \tilde{a}_{-n}(y,\eta)\sigma_\eta = |\det\phi'(x)| \int_{|\xi|=1} a_{-n}(x,\xi)\sigma_\xi.$$

That shows that the quantity

$$\left(\int_{|\xi|=1} a_{-n}(x,\xi)\sigma_\xi\right)dx \tag{62.128}$$

is a 1-density over M that we call $\mathrm{Res}_{\mathrm{W}}(x)A$. The **Wodzicki residue** is the integral of that over M:

$$\mathrm{Res}_{\mathrm{W}} A = \int_M \mathrm{Res}_W(x)A = \int_{S^*M} \iota_R\alpha = \int_{S^*M} a_{-n}(x,\xi)\sigma_\xi dx \tag{62.129}$$

where $S^*M = \{(x,\xi) \in T^*M \text{ tel que } |\xi| = 1\}$. Notice that the latter integral can diverge.

Proposition 62.60.
The operation $\mathrm{Res}_{\mathrm{W}}$ *is a trace on the algebra of classical pseudo-differential operators. That means that* $\mathrm{Res}_W[A,B] = 0$ *whenever* A, B *are classical pseudo-differential operators.*

Proof. If $A \in \Psi^d(M)$ and $B \in \Psi^r(M)$, then the product AB belongs to the space $\Psi^{d+r}(M)$ and the commutator $P = [A,B]$, seen as in $\Psi^{-\infty}$, has symbol (see equation (62.117))

$$p(x,y) \sim \sum_{|\alpha|>0} \frac{(-i)^{|\alpha|}}{\alpha!}\big(\partial_\xi^\alpha a \partial_x^\alpha b - \partial_\xi^\alpha b \partial_x^\alpha a\big).$$

We can suppose that A and B have compact support because the aim is to integrate them with a partition of unity. □

¡¡ Avertissement/question au lecteur !! 62.61
Unfinished proof

Proposition 62.62.
The operation $\mathrm{Res}_{\mathrm{W}}$ *is the unique trace on the algebra* $\mathcal{P}$ *defined in equation 62.124.*

Proof. No proof. □

62.12 Interpolation theory

¡¡ Avertissement/question au lecteur !! 62.63
Il faut citer encore Vá rrily et le second bouquin de Connes et le truc de Landi comme sources

Here is a short review of what is given in the book [953]. Let B_0 and B_1 be two Banach algebras continuously embedded in a topological vector space. We begin to define

$$K(\lambda,x) = \inf\{\|x_0\|_{B_0} + \lambda^{-1}\|x_1\|_{B_1} \text{ tel que } x_0 + x_1 = x,\ x_0 \in B_0,\ x_1 \in B_1\}$$

for all $x \in B_0 + B_1$ and $\lambda \in]0,\infty[$. For a fixed x, we consider the function

$$f(\lambda) = \lambda^\alpha K(\lambda,x),$$

and we define the norm of $x \in_{(\alpha,\beta)}$ by

$$\|x\|_{\alpha,\beta} = \left(\int_{\mathbb{R}_0^+} f(\lambda)^q \frac{d\lambda}{\lambda}\right)^{1/q} \tag{62.130}$$

where $q = 1/\beta$.

Now we consider the special case in with B_0 is the ideal of compact operators on the Hilbert space $\mathscr{H}$ and $B_1 = \mathscr{L}^1(\mathscr{H})$, both seen as subspaces of $\mathscr{L}^1(\mathscr{H})$.

Proposition 62.64.
If $\alpha = 1/p$ and $\beta = 1/q$, and $q < 0$, a compact operator T belongs to $\mathscr{L}^{(p,q)}(\mathscr{H})$ if and only if

$$\sum_{N=1}^{\infty} N^{\alpha-1} q^{-1} \sigma_N(T)^q < \infty$$

where $\sigma_N(T) = \sup\{\|T|_E\|_1$ tel que $\dim E = N\}$. When $q = \infty$, we have $T \in \mathscr{L}^{(p,\infty)}$ if and only if $N^{\alpha-1}\sigma_N(T)$ is a bounded sequence.

Proposition 62.65.
Each space $\mathscr{L}^{(p,q)}$ with $1 < p < \infty$ and $1 \leqslant q \leqslant \infty$ is a two-sided ideal in $\mathscr{L}(\mathscr{H})$ and when $p_1 < p_2$, we have

$$\mathscr{L}^{(p_1,q_1)} \subset \mathscr{L}^{(p_1,q_1)}.$$

When $p_1 = p_2$, this inclusion holds if $q_1 \leqslant q_2$.

Notice that the fact to say that the sequence $N^{(\alpha-1)}\sigma_N(t)$ is bounded is the same as to say that $\sigma_N(T) = O(N^{1-\alpha})$. This in turns is noting else than $\mu_n(T) = O(n^{-\alpha})$ and the norm that we put on $\mathscr{L}^{(p,\infty)}$ is

$$\|T\|_{p,\infty} = \sup_{N \geqslant 1} \frac{1}{N^{1-\alpha}} \sigma_N(T).$$

62.13 Trace class operators

62.13.1 Trace

We say that a bounded linear operator A on the Hilbert space $\mathscr{H}$ is **trace class** if for a certain basis $\{e_k\}$, the sum

$$\sum_k \langle A^* A e_k, e_k\rangle^{1/2} \tag{62.131}$$

converges. In that case the sum $\sum_k \langle Ae_k, e_k\rangle$ converges absolutely (because each term in the sum (62.131) is positive) and is thus independent with respect of the choice of the basis. That number is called the **trace**:

$$\mathrm{Tr}(A) = \sum_k \langle Ae_k, e_k\rangle. \tag{62.132}$$

In the space $\mathscr{L}^1$ of trace class operators, we have that

$$\|T\|_1 = \mathrm{Tr}\,|T| \tag{62.133}$$

is a norm which is not equal to the operator norm $\|T\| = \mu_0(T)$. More generally we have the following lemma.

Lemma 62.66.
We have

$$\sigma_n(T) = \sup\{\|TP_n\| \text{ such that } P_n \text{ is a projector of rank } n\}, \tag{62.134}$$

and each σ_n is a norm on the space $\mathcal{K}(\mathscr{H})$ of compact operators over the Hilbert space $\mathscr{H}$.

From formula (60.8), the sequence of $\mu_k(T)$ is decreasing, so that we get the inequalities

$$\sigma_n(T) \leqslant n\mu_0(T) = n\|T\|.$$

Lemma 62.67.
We have the formula

$$\sigma_n(T) = \inf\{\|R\|_1 + n\|S\| \text{ such that } R \in \mathcal{K}, S \in \mathcal{K}, R + S = T\} \tag{62.135}$$

for every $T \in \mathcal{K}(\mathscr{H})$.

Proof. It $T = R + S$, we have $\sigma_n(T) \leqslant \sigma_n(R) + \sigma_n(S) \leqslant \|R\|_1 + n\|S\|$. Now we have to prove that $\sigma_n(T)$ actually reaches the infimum. We can suppose that T is positive; if not, we can change every signs in (62.135), and nothing is changed. So lemma 60.26 is applicable. Let P_n be the rank n projector over the space spanned by the eigenvectors corresponding to the eigenvalues $\mu_0, \cdots, \mu_{n-1}$ of T. Then we consider $R = (T - \mu_n)P_n$ and $S = \mu_n P_n + T(1 - P_n)$. Then we have $\|R\|_1 = \sum_{k<n}(\mu_k - \mu_n) = \sigma_n(T) - n\mu_n$ and $\|S\| = \mu_n$. We conclude that

$$\sigma_n(T) = \|R\|_1 + n\mu_n = \|R\|_1 + n\|S\|,$$

which concludes the proof. □

That formula only holds for $n \in \mathbb{N}$, but we can *define*

$$\sigma_\lambda = \inf\{\|R\|_1 + \lambda\|S\| \text{ such that } R\} \tag{62.136}$$

for every $\lambda \in \mathbb{R}^+$.

Proposition 62.68.
If $0 \leqslant \lambda \leqslant 1$, we have

$$\sigma_\lambda = \lambda\|T\|,$$

and more generally if $\lambda = n + t$ with $n \in \mathbb{N}$, $0 \leqslant t \leqslant 1$ we have

$$\sigma_\lambda(T) = (1 - T)\sigma_n(T) + t\sigma_{n+1}(T). \tag{62.137}$$

Eqsigunmoinstn

Moreover we have that the function $\lambda \mapsto \sigma_\lambda(T)$ is an increasing piecewise linear concave function.

Proof. No proof. □

The inequality (62.137) makes that σ_λ is a norm that satisfies in particular the triangular inequality.

Proposition 62.69.
The inequality (62.137) can be reinforced into

$$\sigma_\lambda(A) + \sigma_\mu(B) \leqslant \sigma_{\lambda+\mu}(A + B) \tag{62.138}$$

when A and B are positive operators. Combining with the triangular inequality, we find

$$\sigma_\lambda(A + B) \leqslant \sigma_\lambda(A) + \sigma_\lambda(B) \leqslant_\sigma 2\lambda(A + B) \tag{62.139}$$

under the same assumptions.

62.14 Dixmier traces

We follow the approach given in [819, 941].

62.14.1 Banach limit

Let l_∞ be the Banach space of complex-valued bounded sequences. A **Banach limit** is a linear functional $\phi\colon l_\infty \to \mathbb{R}$ such that for every real sequences x and y, we have

— $\phi(\lambda x + \mu y) = \lambda\phi(x) + \mu\lambda(y)$
— if $x \geqslant 0$, then $\phi(x) \geqslant 0$,
— if S is the **shift operator** $(Sx)_i = x_{i+1}$, then $\phi(x) = \phi(Sx)$,
— if x converges, then $\phi(x) = \lim x$.

Such a functional is not unique: if ϕ and φ are two such Banach limits, one can find a sequence x such that $\phi(x) \neq \varphi(x)$. Such an example has to be non-convergent.

62.14.2 Infinitesimal operator

Proposition 60.20 leads us to think to compact operators as infinitesimals because it is almost zero on th major part of the space. Here is the precise definition.

Definition 62.70.
*Let $\alpha \in \mathbb{R}^+$. An **infinitesimal** of order α is a compact operator T such that there exists a $C \leqslant \infty$ for which*

$$\mu_n(T) \leqslant Cn^{-\alpha}.$$

for all $n \geqslant 1$.

When T_1 and T_2 are two compact operators, one can prove that

$$\mu_{n+m}(T_1T_2) \leqslant \mu_n(T_1)\mu_m(T_2)n,$$

hence if T_j is of order α_j, T_1T_2 is of order $\leqslant \alpha_1 + \alpha_2$. Moreover the infinitesimals of order α form a two-sided ideal (non closed) in $\mathfrak{B}(\mathscr{H})$ because for all $T \in \mathfrak{K}(\mathscr{H})$ and $B \in \mathfrak{B}(\mathscr{H})$, we have

$$\begin{aligned} \mu_n(TB) &\leqslant \|B\|\mu_n(T) \\ \mu_n(BT) &\leqslant \|B\|\mu_n(T). \end{aligned} \tag{62.140}$$

For a proof, see bibliography of [819].

Remark that for an infinitesimal of order 1, the characteristic values are bounded by $\mu_n(T) \leqslant 1/n$, so that there are no reason for such a T be belongs to $\mathscr{L}^1$, but the divergence of $\mathrm{Tr}(T)$ is at most logarithmic:

$$\sum_{n=1}^{N-1} \mu_n(T) \leqslant C \ln N.$$

62.14.3 Dixmier trace

We want to build a trace which is non zero on infinitesimals of order 1, but which vanishes on infinitesimals of lager order. The usual trace is defined, for $T \in \mathscr{L}^1$, by

$$\mathrm{Tr}\, T := \sum_n \langle T\xi_n | \xi_n \rangle$$

and is independent of the chosen orthonormal basis $\{\xi_n\}$ of $\mathscr{H}$. When T is positive and compact, we define the trace by

$$\mathrm{Tr}\, T = \sum_{n=1}^{\infty} \mu_n(T).$$

The problem is that infinitesimals of order 1 are not in general in $\mathscr{L}^1$ because on these operators, we do not have a better control that $\mu_n(T) \leqslant \frac{C}{n}$. Hence the sum can diverge. Worse: the space $\mathscr{L}^1$ contains infinitesimals of order lager than 1. However we know that in the case of infinitesimals positive operators of order 1 is at most logarithmic:

$$\sum_{n=0}^{N-1} \mu_n(T) \leqslant C \ln N.$$

We are going to find a way to extract the coefficient of the logarithmic divergence. We denote by $\mathscr{L}^{(1,\infty)}$ the ideal of compact operators which are infinitesimals of order 1. If $T \in \mathscr{L}^{(1,\infty)}$, we want to define the trace by

$$\lim_{N\to\infty} \frac{1}{\ln N} \sum_{n=0}^{N-1} \mu_n(T).$$

This definition has two main problems: it is not specially linear in T and does not converge in general. Let the sums

$$\sigma_N(T) = \sum_{n=1}^{N-1} \mu_n(T) \quad \textit{and} \quad \gamma_N(T) = \frac{\sigma_N(T)}{\ln N}.$$

One can prove that

$$\sigma_N(T_1 + T_2) \leqslant \sigma_N(T_1) + \sigma_N(T_2) \tag{62.141a}$$
$$\sigma_{2N}(T_1 + T_2) \geqslant \sigma_N(T_1) + \sigma_N(T_2); \tag{62.141b}$$

the second relation only holds with $T_1, T_2 \geqslant 0$. Therefore, when $T_1, T_2 > 0$, we have

$$\begin{aligned} \gamma_N(T_1 + T_2) &\leqslant \gamma_N(T_1) + \gamma_N(T_2) \\ &\leqslant \frac{\sigma_{2N}(T_1 + T_2)}{\ln N} \\ &= \frac{\gamma_{2N}(T_1 + T_2)}{\ln N} \ln 2N \\ &\leqslant \gamma_{2N}(T_1 + T_2)\big(1 + \frac{\ln 2}{\ln N}\big) \end{aligned} \tag{62.142}$$

eq_gammaNleq

because for suitably large N,

$$1 + \frac{\ln 2}{\ln N} = \frac{\ln N + \ln 2}{\ln N} \leqslant \frac{\ln 2N}{\ln N} = \ln N.$$

If the sequence γ_N converges, then it is linear because when $N \to \infty$, we have $1 + \ln 2/\ln N \to 1$; hence equalities

$$\gamma_N(T_1 + T_2) \leqslant \gamma_N(T_2) + \gamma_N(T_2) \leqslant \gamma_{2N}(T_2 + T_2)\big(1 + \frac{\ln 2}{\ln N}\big).$$

fix the limit of $\gamma_N(T_1) + \gamma_N(T_2)$ on the one of $\gamma_{2N}(T_1 + T_2)$. This however does not resolve the problem of convergence of γ_n, even when it is bounded.

The trick is to not take the usual limit, but to define a linear form $\lim_\omega$ on the space $l^\infty(\mathbb{N})$ of bounded sequences and to impose to $\lim_\omega$ to fulfil certain conditions.

From remark on page 3477, we know that the values $\mu_n(T)$ are unitary invariant, hence the sequence (γ_N) is also unitary invariant. This leads us to search for an unitary invariant form $\lim_\omega$. The following proposition allows us to only define $\lim_\omega$ in the positive part of $\mathscr{L}^{(1,\infty)}$.

Proposition 62.71.
The space $\mathscr{L}^{(1,\infty)}$ is generated by its positive part.

Proof. No proof. □

Here are the conditions we impose to $\lim_\omega : l^\infty(\mathbb{N}) \to \mathbb{N}$:

(1) it is a linear form,

(2) $\lim_\omega(\gamma_N) \geqslant 0$ is $\gamma_N \geqslant 0$,

(3) $\lim_\omega(\gamma_N) = \lim \gamma_N$ if the usual limit exists, limomiii

(4) $\lim_\omega(\gamma_1, \gamma_1, \gamma_2, \gamma_2, \gamma_3, \gamma_3) = \lim_\omega(\gamma_N)$, limomiv

(5) $\lim_\omega(\gamma_{2N}) = \lim_\omega(\gamma_N)$.

The condition (5) is the *scale invariance*; this property is equivalent to the property (5). Dixmier has found a lot of such form. For each of them, one has a trace

$$\mathrm{Tr}_\omega(T) = \lim_\omega \frac{1}{\ln N} \sum_{n=0}^{N-1} \mu_n(T) \tag{62.143}$$

for positive $T \in \mathscr{L}^{(1,\infty)}$. When T_1 and T_2 are positive, we have linearity:

$$\mathrm{Tr}_\omega(T_1 + T_2) = \mathrm{Tr}_\omega(T_1) + \mathrm{Tr}_\omega(T_2).$$

Since $\mathscr{L}^{(1,\infty)}$ is generated by its positive part, the form Tr_ω —which is initially only defined for positive operators T— extends to the whole $\mathscr{L}^{(1,\infty)}$ with properties

(1) $\mathrm{Tr}_\omega(T) \geqslant 0$ if $T \leqslant 0$,
(2) $\mathrm{Tr}_\omega(\lambda_1 T_2 + \lambda_2 T_2) = \lambda_1 \mathrm{Tr}_\omega(T_1) + \lambda_2 \mathrm{Tr}_\omega(T_2)$,
(3) $\mathrm{Tr}_\omega(BT) = \mathrm{Tr}_\omega(TB)$ for all $B \in \mathfrak{B}(\mathscr{H})$,
(4) $\mathrm{Tr}_\omega(T) = 0$ if T is an infinitesimal of order larger than 1.

item_tromTiv

For a proof, see [819]. For (4), remark that the space of infinitesimals of order larger than 1 form a two-sided ideal whose elements fulfil $n\mu_n(T) \to 0$. Then the sequence (γ_B) converges to zero too and the Dixmier trace vanishes.

62.14.4 Dixmier: second

The set of **infinitesimals of order** 1 is the normed ideal

$$\mathscr{L}^{1+} = \{T \in \mathcal{K} \text{ tel que } \|T\|_{1+} < \infty\} \tag{62.144}$$

where the norm $\|T\|_{1+}$ is defined by

$$\|T\|_{1+} = \sup_{\lambda \geqslant a} \frac{\sigma_\lambda(T)}{\ln \lambda}. \tag{62.145}$$

That idea include the trace class operators. We define

$$\mathscr{L}^p = \{T \in \mathcal{K} \text{ tel que } \mathrm{Tr}\,|T|^p < \infty\}. \tag{62.146}$$

Proposition 62.72.
On the space $\mathscr{L}^p$, we have

$$\sigma_\lambda = O(\lambda^{1-1/p})$$

and $\mathscr{L}^{1+} \subset \mathscr{L}^p$ when $p > 1$.

Notice that, when $T \in \mathscr{L}^{1+}$, the function $\lambda \mapsto \sigma_\lambda(T)/\ln \lambda$ is bounded and continuous on the interval $[e, \infty[$. It belongs thus to the C^*-algebra $C_b\big([e, \infty[\big)$. So we can use the **Cesàro means**:

$$\tau_\lambda(T) = \frac{1}{\ln \lambda} \int_e^\lambda \frac{\sigma_u(T)}{\ln u} \frac{du}{u}, \tag{62.147}$$

EqCearomaen

and the function $\lambda \mapsto \tau_\lambda(T)$ still belongs to $C_B\big([e, \infty[\big)$ with $\|T\|_{1+}$ as upper bound.

Proposition 62.73.
The double inequality

$$0 \leqslant \tau_\lambda(A) + \tau_\lambda(B) - \tau_\lambda(A + B) \leqslant \big(\|A\|_{1+} + \|B\|_{1+}\big) \ln 2 \frac{\ln \ln \lambda}{\ln \lambda}.$$

holds for A, $B \in \mathscr{L}^{1+}$.

Proof. No proof. □

That proves that $\tau\lambda$ becomes additive when λ goes to infinity. We can work on that in order to make it additive. First we consider

$$\mathcal{B} = C_b\big([e, \infty[\big)/C_0\big([e, \infty[\big),$$

and we consider $[\tau(A)]$, the class of $\tau(A)$ (i.e. the function $\lambda \mapsto \tau_\lambda(A)$) with respect to that quotient.

Proposition 62.74.
The map $[\tau]$ *is additive, positive and homogeneous from the positive cone in* $\mathscr{L}^{1+}$ *to* $\mathcal{B}$*. Moreover*

$$[\tau(UAU^{-1})] = [\tau(A)]$$

for every unitary U*.*

That makes that $[\tau]$ extends to a linear map $[\tau]\colon \mathscr{L}^{1+} \to \mathcal{B}$ such that $[\tau](ST) = [\tau](TS)$ for all $T \in \mathscr{L}^{1+}$ and S.

Now if $\omega\colon \mathcal{B} \to \mathbb{C}$ is any state, we define the **Dixmier trace** as

$$\mathrm{Tr}_\omega(T) = \omega\big([\tau](T)\big). \tag{62.148}$$

That definition has a problem: the C^*-algebra $\mathcal{B}$ being non separable, one cannot exhibit a state, so that the formula is in practice unusable.

62.14.5 Noncommutative integral

Let us consider $f \in C_b\big([e,\infty\big)$; the limit $\lim_{\lambda\to\infty} f(\lambda)$ exists if and only if $\omega(f)$ does not depend on ω because of the quotient by $C_0\big([e,\infty[\big)$ which makes that ω can only depend on the behaviour near infinity. We say that the operator $T \in \mathscr{L}^{1+}$ is **measurable** if the function $\lambda \mapsto \tau_\lambda(T)$ converges when $\lambda \to \infty$. In that case, $\mathrm{Tr}_\omega(T)$ equals that limit, and we denote by $⨍ T$ the common value of the Dixmier traces:

$$⨍ T = \lim_{\lambda\to\infty} \tau_\lambda(T) \tag{62.149}$$

if the limit exists. That is the **noncommutative integral** of T.

Proposition 62.75.
It T *is a compact operator and if* $\sigma_n(T)/\ln n$ *converges when* $n \to \infty$*, then the limit* $\lim_{\lambda\to\infty} \tau_\lambda(T)$ *exists and* T *is a measurable operator in* $\mathscr{L}^{1+}$*.*

Proof. Indeed in that case the quantity $\sigma_u(T)/\ln u$ becomes constant when u is large, so that we are left in the definition (62.147) with

$$\lim_{\lambda\to\infty} \frac{C}{\ln\lambda} \int_e^\lambda \frac{du}{u} = C \lim_{\lambda\to\infty} \frac{\ln\lambda - 1}{\ln\lambda}.$$

which exists. □

62.14.6 Residues

Let M be a compact Riemannian spin manifold of dimension n. Let T be a pseudo-differential operator of order $-n$ acting on the sections of a complex vector bundle $E \to M$. Its **residue** is defined by

$$\mathrm{Res}_{\mathrm{W}}\, T = \frac{1}{n(2\pi)^n} \int_{S^*M} \mathrm{Tr}_E\, \sigma_{-n}(T)\, d\mu \tag{62.150}$$

where $\sigma_{-n}\colon T^*M \to \mathrm{End}\, E$ is the principal symbol of T (is as a homogeneous function of degree $-n$) and the integral is taken on the cosphere

$$S^*M = \{(x,\xi) \in T^*M \text{ tel que } \|\xi\| = 1\}$$

with the measure $d\mu = dx\, d\xi$.

Chapter 63

Partial derivative equation

63.1 Basic definitions

We consider PDE's for functions $u\colon \mathbb{R}^d \to \mathbb{R}$. We denote by $D^l u$ the set of partial derivatives of order l:

$$(D^l u)(x) = \{(D^\alpha u)(x)\}_{|\alpha|=l}. \tag{63.1}$$

Some definitions.

Definition 63.1.
*A PDE of order k is **linear** if it is under the form*

$$\sum_{|\alpha|\leqslant k} a_\alpha(x)(\partial^\alpha u)(x) + f(x) = 0. \tag{63.2}$$

*That equation is **homogeneous** if $f = 0$.*

Definition 63.2.
*A PDE is **semi-linear** is it is linear only with respect to the highest derivatives:*

$$\sum_{|\alpha|=k} a_\alpha(\partial^\alpha u)(x) + F\big(x, (Du)(x), \dots, D^{k-1}u(x)\big) = 0 \tag{63.3}$$

where $F\colon \Omega \times \mathbb{R}^{\cdots} \to \mathbb{R}$ is some function. The dots here represent the number of different partial derivative of order 0 to $k-1$ one has to consider.

Definition 63.3.
*A PDE of order k is **quasi-linear** if it is linear with respect to the higher order derivatives, with a coefficient that can depend to the lower order derivatives of the unknown:*

$$\sum_{|\alpha|=k} a_\alpha\big(x, u(x), (Du)(x), \dots, (D^{k-1})u(x)\big)(\partial^\alpha u)(x) + F\big(x, u(x), \dots, (D^{k-1}u)(x)\big) = 0. \tag{63.4}$$

Exemple 63.4.
The equation

$$x(\partial_x u)(x,y) + y(partial_y u)(x,y) + u\sin(xy) = 1 \tag{63.5}$$

is linear of order 1. Indeed the coefficient of the partial derivatives of all orders (0 and 1) depend on x and y, but not of u or the derivatives of u. △

Exemple 63.5.
The equation

$$x(\partial_x u)(x,y) + y(\partial_y u)(x,y) + u^2\sin(xy) = 1 \tag{63.6}$$

is semi-linear of order 1. Indeed the coefficient of the partial derivatives of order 1 (the higher order) depend on x and y but not of u or the derivatives of u. However, the derivatives of lower order (here: 0) are not linear. △

Exemple 63.6.
The heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 1 \tag{63.7}$$

is linear ans non homogeneous. △

Exemple 63.7.
The equation

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(c(u)\frac{\partial u}{\partial x}\right) = 0 \tag{63.8}$$

is an abuse of notations for asking

$$\frac{\partial u}{\partial t}(x,t) - \frac{\partial x}{\partial}\left((c\circ u)\times\frac{\partial u}{\partial x}\right)(x,t). \tag{63.9}$$

for every t and x. This is a second order equation whose highest order term is

$$(c\circ u)\frac{\partial^2 u}{\partial x^2} \tag{63.10}$$

The coefficient of $\partial_x^2 u$ depend on u, but not on the second order derivatives. So this is a quasi-linear equation. △

63.2 Principal symbol

Let the semi-linear equation

$$\sum_{|\alpha|=k} a_\alpha(x)(\partial^\alpha u)(x) + F\big(x,u(x),(Du)(x),\dots,(D^{k-1}u)(x)\big). \tag{63.11}$$

The differential operator associated with that equation is

$$\sum_{|\alpha|=k} a_\alpha \partial^k + F \tag{63.12}$$

EQooNSOMooCfAdxt

where we admit quite a notational shortcut to say that F applies on u by

$$(Fu)(x) = F\big(x,u(x),(Du)(x),\dots,(D^{k-1}u)(x)\big). \tag{63.13}$$

Definition 63.8.
*The **principal symbol** of the operator (63.12) is the function*

$$\begin{aligned}\sigma\colon \mathbb{R}^d\times\mathbb{R}^d &\to \mathbb{R}\\ (x,\xi) &\mapsto \sum_{|\alpha|=k} a_\alpha(x)\xi^\alpha\end{aligned} \tag{63.14}$$

where

$$\xi^\alpha = \xi_1^{\alpha_1}\dots\xi_d^{\alpha_d}. \tag{63.15}$$

Definition 63.9.
The characteristic associated with the symbol σ is the set of surfaces $S\subset\mathbb{R}^d$ of the form

$$S = \{x\in\mathbb{R}^d \text{ such that } \phi(x)=0\} \tag{63.16}$$

where $\phi\colon\mathbb{R}^d\to\mathbb{R}$ satisfies

(1) $\sigma\big(x,(\nabla\phi)(x)\big)=0$ for every $x\in\mathbb{R}^d$,

(2) $(\nabla\phi)(x)\neq 0$ for every $x\in S$.

Exemple 63.10.
Let the differential equation

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + x\frac{\partial^2 u}{\partial y^2} = f \tag{63.17}$$

for the function $u\colon \mathbb{R}^2 \to \mathbb{R}$ and some given functions a, b and c.

The principal symbol of that operator is

$$\sigma(x, y, \xi_1, \xi_2) = a(x, y)\xi_1^2 + b(x, y)\xi_1\xi_2 + c(x, y)\xi_2^2 \tag{63.18}$$

and the characteristic equation reads [1]

$$\sigma(x, y, (\nabla\phi)(x, y)) = a(x, y)\left(\frac{\partial^2\phi}{\partial x}(x, y)\right)^2 + b(x, y)\frac{\partial\phi}{\partial x}(x, y)\frac{\partial\phi}{\partial y}(x, y) + c(x, y)\left(\frac{\partial\phi}{\partial y}(x, y)\right)^2 = 0. \tag{63.19}$$

BQooTMQAooZPdcWT

This is a differential equation for $\phi\colon \mathbb{R}^2 \to \mathbb{R}$.

The condition $\nabla\phi \neq 0$ seems now natural: if $\partial_i\phi = 0$ for every i, the equation trivializes. We suppose that $(\partial_y\phi)(x_0, y_0) \neq 0$ and that this condition holds on a neighborhood of (x_0, y_0). By the implicit function theorem 17.48 there exists a function $x \mapsto y(x)$ such that

$$\phi\big(x, y(x)\big) = 0 \tag{63.20}$$

for every x in a neighborhood of x_0. Thus the function

$$\varphi\colon x \mapsto \phi\big(x, y(x)\big) \tag{63.21}$$

is identically zero. Its derivative is

$$\varphi'(x) = \frac{\partial\phi}{\partial x}\big(x, y(x)\big) + \frac{\partial\phi}{\partial y}\big(x, y(x)\big)y'(x) = 0. \tag{63.22}$$

We substitute $\frac{\partial\phi}{\partial x}$ by $-\frac{\partial\phi}{\partial y}y'$ in the characteristic equation (63.19) written on the point $(x, y(x))$. We simplify the whole by $(\partial_y\phi)(x, y(x))^2$ (which is non vanishing) and obtain

$$ay'(x)^2 - by'(x) + c = 0. \tag{63.23}$$

where a, b and c are taken at $\big(x, y(x)\big)$.

The existence of $y'(x)$ depend on the sign of $b^2 - 4ac$ which is a function of x.

△

Definition 63.11.
Let a partial differential equation.

(1) If the characteristic equation has no real solutions, the equation is ***elliptic***

(2) If the characteristic equation has k distinct real solutions, the equation is ***hyperbolic***

(3) If some solutions to the characteristic equation have a multiplicity, the equation is ***parabolic***

Since the number of solutions of the characteristic equation depend on the coefficients that are themselves functions, an equation can have different type at different places.

Exemple 63.12.
Let the equation

$$\frac{\partial^2 u}{\partial x^2} - (x^2 - y^2)\frac{\partial^2 u}{\partial y^2} = 0. \tag{63.24}$$

The type of this equation is governed by $\delta = 4(x^2 - y^2)$; this can have any sign depending on the point. △

1. We are not going to explicit all dependencies in x and y.

63.2.1 Classification with respect to the boundary conditions

For the next few definitions we suppose that Ω is a sufficiently smooth domain in $\mathbb{R}^d$; in particular we suppose that the boundary $\partial\Omega$ admits a normal outward vector $n(x)$ for every $x \in \partial\Omega$.

Definition 63.13.
*A problem with **Dirichlet** boundary condition is a partial derivative equation with the condition*

$$u(x) = g(x) \tag{63.25}$$

for $x \in \Gamma \subset \partial\Omega$. In other words, we impose the value of u on a part of the boundary.

Definition 63.14.
*A problem with **Neumann** boundary condition is a partial derivative equation with the condition*

$$\frac{\partial u}{\partial n} \cdot x = g(x) \tag{63.26}$$

for $x \in \Gamma \subset \partial\Omega$. In other words, we impose the normal derivative of u on a part of the boundary.

Definition 63.15.
*Let Ω be a domain of $\mathbb{R}^d$ and a differential operator L on a part of $\mathrm{Fun}(\Omega)$. Let g be a function on $\partial\Omega$. A problem with **stationary boundary condition** is a problem of the type: find a function u defined on Ω such that*

$$\begin{cases} L(u) = f & on\ \Omega \\ u = & g on\ \partial\Omega. \end{cases} \tag{63.27}$$

Definition 63.16.
*Let Ω be a domain of $\mathbb{R}^d$ and a differential operator L on a part of $\mathrm{Fun}(\Omega)$. Let g be a function on $\partial\Omega$. A problem with **evolution boundary condition** is a problem of the type: find a function u defined on $]0,\infty[\times \Omega$ such that*

$$\begin{cases} \frac{\partial^m u}{\partial t^m} + l(u) = fu(t,.) = g(t,.) & on\ \partial\Omega \\ u(0,.) = u_0 & on\ \Omega \end{cases} \tag{63.28}$$

where L does not operate on the "t" variable.

Definition 63.17.
*A boundary problem is **well posed** in the sense of Hadamard if*

(1) it accepts an unique solution

(2) this solution is continuous with respect to the parameters.

Here the "parameters" are the constants as well as the functions given in the problem. So one needs to precise the topology on the functional spaces that are implied in the problem.

EXooLTODooOwJtGC

Exemple 63.18 ([954]).
Let the problem: find a function u defined on $\Omega = [0,1]^2$ such that

$$\begin{cases} \Delta u = 0 & \text{on } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{63.29}$$

Let us prove the unicity of the solution (we do not prove existence). Let u_1 and u_2 be two solutions and define $v = u_1 - u_2$. This function satisfies

$$\begin{cases} \Delta v = 0 & \text{on } \Omega \\ v = 0 & \text{on } \partial\Omega. \end{cases} \tag{63.30}$$

Using the by part formula of example 57.59, with $u = v$,

$$\int_\Omega v\Delta v = -\int_\Omega \|\nabla v\|^2 + \int_{\partial\Omega} v\frac{\partial v}{\partial n}. \tag{63.31}$$

Since $\Delta v = 0$ and $v = 0$ on $\partial\Omega$, we are left with

$$\int_\Omega \|\nabla v\|^2 = 0. \tag{63.32}$$

If v is smooth enough, that implies $\nabla v = 0$ on Ω. Since the function v is constant on Ω and is zero on $\partial\Omega$ we deduce that $v = 0$ on Ω. This proves the unicity, up to numerous regularity and topological considerations on the choice of the functional spaces.

Remark that we did not prove existence and even less proved the continuity of the solution with respect to g. For this we should precise a functional space S in which one chooses g and a functional space V in which one chooses u (for sure not the same because g is defined on $\partial\Omega$ while u is defined on Ω) and prove that $g \mapsto u$ is a continuous map $S \to V$. △

From this example one can believe that under some good choices of functional spaces, the Dirichlet problem for the Laplace equation is well posed. We give now different conditions for the same equation, and show that the resulting problem is for sure not well posed.

Exemple 63.19 ([1])**.**
Let $\epsilon > 0$ and the problem for u_ϵ on $\Omega = \mathbb{R} \times]0, \infty[$,

$$\begin{cases} \Delta u_\epsilon = 0 & \text{on } \Omega \\ u_\epsilon(x, 0) = 0 & \text{for } x \in \mathbb{R} \\ \frac{\partial u_\epsilon}{\partial y}(x, 0) = \epsilon \sin(x/\epsilon) & \text{for } x \in \mathbb{R}. \end{cases} \tag{63.33}$$

EQooUBUEooMCDIIZ

We are searching the solutions in

$$\big(C^\infty(\Omega), p_{K,m}\big) \tag{63.34}$$

where $p_{K,m}$ are the semi-norms defined in (30.21). The solution is unique from the same reason as in example 63.18 [2].

One checks that, when $\epsilon \neq 0$, the unique solution is

$$u_\epsilon(x, y) = \epsilon^2 \sin(\frac{1}{\epsilon}x) \sinh(\frac{1}{\epsilon}y). \tag{63.35}$$

The parameters in the equation is $\epsilon \in \mathbb{R}$, and in order to check the well-posedness we consider the map ψ that maps ϵ to the unique corresponding solution:

$$\begin{aligned} \psi\colon \mathbb{R} &\to C^\infty(\Omega) \\ \epsilon &\mapsto \begin{cases} (x, y) \mapsto \epsilon^2 \sin(\frac{x}{\epsilon}) \sinh(\frac{y}{\epsilon}) & \text{if } \epsilon \neq 0 \\ (x, y) \mapsto 0 & \text{if } \epsilon = 0. \end{cases} \end{aligned} \tag{63.36}$$

Our point is that ψ is not continuous at $\epsilon = 0$. Due to proposition 7.344, in order ψ to be continuous we need $p_{K,m}(u_\epsilon) \to 0$ when $\epsilon \to 0$ for every m and every compact $K \subset \Omega$.

Let $m = 0$ and $K = [0, 2\pi] \times [1, 2]$. For every $\epsilon < 1$ we can find $x \in [0, 2\pi]$ such that $\sin(x/\epsilon) = 1$, so that

$$\|u_\epsilon\|_{K,0} \geqslant \epsilon^2 \sinh\left(\frac{2}{\epsilon}\right) = \frac{1}{2}\epsilon^2 |e^{2/\epsilon} - e^{-2/\epsilon}|. \tag{63.37}$$

Now one has

$$\lim_{\epsilon\to 0} \|u_\epsilon\|_{K,0} \geqslant \lim_{\epsilon\to 0} \frac{1}{2}\epsilon^2 |e^{2/\epsilon} - e^{-2/\epsilon}| = \infty. \tag{63.38}$$

Thus one has not $\lim_{\epsilon\to 0} \psi(\epsilon) = \psi(0)$ and ψ is not continuous.

We conclude that the problem (63.33) is not well posed in the sense of Hadamard. △

2. Is that true?

63.2.2 Characteristic speed

Let the PDE for $u\colon \mathbb{R}^2 \to \mathbb{R}$

$$\frac{\partial u}{\partial t} + \lambda \frac{\partial u}{\partial x} = 0. \tag{63.39}$$

EQooEVOBooTabJIg

The principal symbol is

$$\sigma(t, x, \xi_1, \xi_2) = \xi_1 + \lambda \xi_2, \tag{63.40}$$

so that the characteristic equation is the same as the original equation. Thus in the case of a first order equation, the characteristic equation has to be something else.

Suppose u to be a solution of (63.39) and pose $\varphi(t) = u\big(x(t), t\big)$ for some function x. We have

$$\varphi'(t) = \frac{\partial u}{\partial x} x' + \frac{\partial u}{\partial t}. \tag{63.41}$$

If $x'(t) = \lambda$ then the initial equation says that $\varphi'(t) = 0$ for every t. That means that u is constant on the characteristic curves

$$t \mapsto \big(x(t), t\big). \tag{63.42}$$

The data of an initial condition $u(x, 0) = u_0(x)$ fix the value of u on each characteristic curve.

63.2.3 Decoupling a system

Let the system

$$\begin{cases} \rho \dfrac{\partial u}{\partial t} + \dfrac{\partial p}{\partial x} = 0 & \text{(63.43a)} \\ \dfrac{\partial p}{\partial t} + \rho c^2 \dfrac{\partial u}{\partial x} = 0 & \text{(63.43b)} \end{cases}$$

where ρ and c are non vanishing constants. This system can be written under the form

$$\frac{\partial}{\partial t} \begin{pmatrix} u \\ p \end{pmatrix} + \begin{pmatrix} 0 & \rho^{-1} \\ \rho c^2 & 0 \end{pmatrix} \frac{\partial}{\partial x} \begin{pmatrix} u \\ p \end{pmatrix} = 0. \tag{63.44}$$

We set

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u \\ p \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & \rho^{-1} \\ \rho c^2 & 0 \end{pmatrix}, \tag{63.45}$$

so that the equation reads

$$\partial_t \bar{u} + A \partial_x \bar{u} = 0. \tag{63.46}$$

Now we diagonalize the matrix A. Some computations show that the eigenvalues of A are $\pm c$ and that the corresponding eigenvectors are

$$f_1 = \begin{pmatrix} -(\rho c)^{-1} \\ 1 \end{pmatrix}; \tag{63.47a}$$

$$f_2 = \begin{pmatrix} (\rho c)^{-1} \\ 1 \end{pmatrix}. \tag{63.47b}$$

If one consider

$$K = \begin{pmatrix} -(\rho c)^{-1} & (\rho c)^{-1} \\ 1 & 1 \end{pmatrix} \tag{63.48}$$

we have

$$K^{-1} = \frac{1}{2} \begin{pmatrix} -\rho c & 1 \\ \rho c & 1 \end{pmatrix} \tag{63.49}$$

and

$$D = K^{-1} A K = \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix} \tag{63.50}$$

Now we consider the new variables $\bar{v} = K^{-1}\bar{u}$, so that we have

$$\partial_t \bar{u} + A\partial_x \bar{u} = 0 \tag{63.51a}$$
$$\partial_t K\bar{u} + A\partial_x K\bar{u} = 0 \tag{63.51b}$$
$$\partial_t \bar{v} + K^{-1}AK\partial_x \bar{v} = 0 \tag{63.51c}$$
$$\partial_t \bar{v} + D\partial_x \bar{v} = 0. \tag{63.51d}$$

The system is now decoupled:

$$\begin{cases} \dfrac{\partial v_1}{\partial t} - c\dfrac{\partial v_1}{\partial x} = 0 & (63.52a) \\ \dfrac{\partial v_2}{\partial t} + c\dfrac{\partial v_2}{\partial x} = 0 & (63.52b) \end{cases}$$

and we have two equations of the form (63.39) already treated.

63.3 Some examples

Let the equation

$$\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial u}{\partial x} = 0. \tag{63.53}$$

If we let $v = \frac{\partial u}{\partial x}$ we can write the equation as

$$\frac{\partial v}{\partial y} + v = 0. \tag{63.54}$$

Let x_0 be fixed and consider the function $v_0(y) = v(x_0, y)$. It satisfies

$$v_0'(y) = \frac{\partial v}{\partial y}(x_0, y) = -v(x_0, y) = -v_0(y), \tag{63.55}$$

so that

$$v_0(y) = A_0 e^{-y} \tag{63.56}$$

for a constant A_0. The latter can be a function of x. We have:

$$v(x, y) = A(x)e^{-y} \tag{63.57}$$

and we we get back to the original equation $v = \partial_c u$:

$$\frac{\partial u}{\partial x} = A(x)e^{-y}. \tag{63.58}$$

Integrating,

$$u(x, y) = B(x)e^{-y} + C(x) \tag{63.59}$$

where B is a primitive of A.

Notice that B is almost completely arbitrary. The fact that it is a primitive only says that it has to be derivable.

63.4 Principle of superposition

Let L be a linear differential operator and consider the equation

$$Lu = f. \tag{63.60}$$

EQooLAZSooGGfMMT

By linearity of L, if u_1 and u_2 are solution, the difference $v = u_1 - u_2$ satisfies $Lv = 0$. Thus if one has a particular solution u_P of $Lu = f$ and the general solution u_0 of $Lu = 0$, the general solution of (63.60) is

$$u = u_P + u_0. \tag{63.61}$$

Now suppose that we are working on $\mathbb{R}^2$ and that f can be decomposed as $f(x, y) = f_1(x) + f_2(x)$. We should solve separately the equations SUBEQSooNAWRooPayDFv

$$(Lu_1)(x, y) = f_1(x) \tag{63.62a}$$
$$(Lu_2)(x, y) = f_2(y) \tag{63.62b}$$
$$(Lu_0)(x, y) = 0. \tag{63.62c}$$

and then write the solution $u = u_1 + u_2 + u_0$.

This decomposition invites us to solve $(Lu_1)(x, y) = f_1(x)$ for a function u_1 that only depends on x because (if one suppose that L commutes with the derivative with respect to y),

$$L\left(\frac{\partial u_1}{\partial y}\right) = \frac{\partial}{\partial y}\big(L(u_1)\big) = \frac{\partial}{\partial y}\big(f_1(x)\big) = 0. \tag{63.63}$$

This shows that $\partial_y u_1$ is a solution of $Lu = 0$, and thus will be included in u_0.

This is not a proof that searching for u_1 that only depends on x is a good idea. But it is a strong incitation. Anyway, if we solve the equations (63.62) in any way, we are done by the superposition principle.

Exemple 63.20.
Let the equation

$$\frac{\partial^2 u}{\partial t^2} - 4\frac{\partial^2 u}{\partial x^2} = \sin(t) + x^{2017}. \tag{63.64}$$

The differential operator $L = \partial_t^2 - 4\partial_x^2$ is linear and we can use the principle of superposition. So we solve the equations

$$Lu_1 = \sin(t) \tag{63.65a}$$
$$Lu_2 = x^{2017} \tag{63.65b}$$
$$Lu_0 = 0 \tag{63.65c}$$

and we are searching $u_1(t)$, $u_2(x)$ and $u_0(t, x)$.

Since u_1 only depends on t, its equation reduces to

$$\frac{\partial^2 u_1}{\partial t^2} = \sin(t), \tag{63.66}$$

so that

$$u_1(t) = -\sin(t) + Ct + D \tag{63.67}$$

where C and D are constants with respect to t. One could write

$$u_1(t, x) = -\sin(t) + C(x)t + D(x), \tag{63.68}$$

but we are in no way interested in the general solution for u_1. Any solution is sufficient, so that we write the easiest one:

$$u(t, x) = -\sin(t). \tag{63.69}$$

We already know that the "missing" part will be included in u_0.

For u_2 we have

$$u_2(t, x) = \frac{-x^{2019}}{4 \cdot 2018 \cdot 2019}. \tag{63.70}$$

We are left to solve

$$\frac{\partial^2 u_0}{\partial t^2} - 4\frac{\partial^2 u_0}{\partial x^2} = 0, \tag{63.71}$$

which is the wave equation. This is already solved in the subsection 33.3.2. The general solution of the problem is

$$u(t,x) = -\sin(t) - \frac{x^{2019}}{4 \cdot 2018 \cdot 2019} + g_1(x+2t) + g_2(x-2t) \tag{63.72}$$

where g_1 and g_2 are any functions suitable for the functional space in which we are searching for u. △

63.4.1 Separation of variables

Let the equation for $\Omega =]0, l[$

$$\begin{cases} \frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = 0 & \text{on }]0,\infty[\times \Omega \\ u(t,0) = u(t,l) = 0 & \text{for } t \in]0,\infty[\\ u(0,x) = u_0(x) & \text{on } \Omega. \end{cases} \tag{63.73}$$

This is the heat equation in which one has fixed the initial distribution and one maintains the extremities at zero.

We are searching for a solution under the form

$$u(t,x) = \sum_{n\in\mathbb{Z}} f_n(t)g_n(x). \tag{63.74}$$

For this we hope that the superposition principle works for infinite sums. So we have to solve

$$\frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = 0 \tag{63.75}$$

with $u(t,x) = f(t)g(x)$. We have

$$g(x)f'(t) - kf(t)g''(x) = 0, \tag{63.76}$$

so that

$$\frac{f'(t)}{f(t)} = \frac{g''(x)}{g(x)}. \tag{63.77}$$

The left hand side only depends on t while the right hand one only depend on x. So they are both constant: there exists $\lambda \in \mathbb{R}$ such that

$$\frac{f'(t)}{f(t)} = \lambda = \frac{g''(x)}{g(x)}. \tag{63.78}$$

We have the independent equations

$$f'(t) = \lambda f(t) \tag{63.79a}$$
$$g''(x) = \lambda g(x). \tag{63.79b}$$

The first one produces

$$f(t) = Ce^{\lambda t} \tag{63.80}$$

where C is a constant that has to be fixed by the initial and boundary conditions.

For g we have the technique described in subsection 32.5.1. We solve the characteristic equation $r^2 + \lambda = 0$ and we get the general solution

$$g(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}. \tag{63.81}$$

From the condition

$$u(t,0) = u(t,l) = 0 \tag{63.82}$$

we want to deduce that $f(t)g(0) = f(t)g(l) = 0$ so that $g(0) = g(l) = 0$ because $f(t) = 0$ for every t leads to a trivial solution. This is however not true because the product $f(t)g(x)$ does not have to satisfy the boundary conditions; only the final sum $\sum_{n\in\mathbb{Z}} f_n(t)g_n(x)$ has to.

For each $n \in \mathbb{Z}$ we have a free parameter λ_n and the functions

$$f_n(t) = C_n e^{\lambda_n t} \tag{63.83a}$$

$$g_n(x) = A_n e^{\sqrt{\lambda_n}x} + B_n e^{-\sqrt{\lambda_n}x}. \tag{63.83b}$$

where A_n, B_n and C_n are still free parameters. We can suppose that $\lambda_n \neq \lambda_m$ when $n \neq m$; if not we only have to redefine C_n, A_n and B_n.

The boundary condition reads

$$\sum_{n\in\mathbb{Z}} C_n e^{\lambda_n t} g_n(0) = 0. \tag{63.84}$$

Having that for every t implies $g_n(0) = 0$. The same holds for $g_n(l)$, so that one indeed has

$$g_n(0) = g_n(l) = 0. \tag{63.85}$$

The condition $g_n(0) = 0$ implies

$$A_n + B_n = 0, \tag{63.86}$$

and the condition $g_n(l) = 0$ then produces

$$A_n\big(e^{\sqrt{\lambda_n}l} - e^{-\sqrt{\lambda_n}l}\big) = 0. \tag{63.87}$$

If $A_n = 0$, we have $g_n = 0$ which is uninteresting. The possibilities for the parenthesis to be vanishing are not numerous: it needs

$$\sqrt{\lambda_n}l = ki\pi, \tag{63.88}$$

that gives $\lambda_n = -\frac{k^2\pi^2}{l^2}$. Since $\lambda_n \neq \lambda_m$ we can number the λ_n in order to have

$$\lambda_n = -\frac{n^2\pi^2}{l^2}. \tag{63.89}$$

In this case,

$$g_n(x) = A_n\big(e^{in\pi x/l} - e^{-in\pi x/l}\big), \tag{63.90}$$

that is

$$g_n(x) = A_n i \sin\left(\frac{n\pi x}{l}\right). \tag{63.91}$$

If we want a real solution we have to choice A_n to be imaginary and up to redefinition we write

$$g_n(x) = A_n \sin\left(\frac{n\pi x}{l}\right) \tag{63.92}$$

with real A_n. Now we have

$$u_n(t,x) = C_n e^{-n^2\pi^2 t/l} \sin\left(\frac{n\pi x}{l}\right) \tag{63.93}$$

and $n \in \mathbb{Z}$.

As far as negative n are concerned,

$$u_{-n}(t,x) = -C_{-n} e^{-n^2\pi^2 t/l^2} \sin\left(\frac{n\pi x}{l}\right) = -\frac{C_{-n}}{C_n} u_n(t,x). \tag{63.94}$$

Thus one can skip the negative ones by redefining the coefficients. Obviously $n = 0$ produces a vanishing term (because of the sine). At the end of the day we have the following solution for our problem:

$$u(t,x) = \sum_{n\in\mathbb{N}^*} C_n e^{-n^2\pi^2 t/l^2} \sin\left(\frac{n\pi x}{l}\right). \tag{63.95}$$

At this point we have to check that

— this is actually a solution,

— the solution is unique up to choice of C_n.

The first point is an exercise of permutation of sum and differentiation. For the second point, we still have to impose $u(0,x) = u_0(x)$, that is

$$\sum_{n\in\mathbb{N}^*} C_n \sin\left(\frac{n\pi x}{l}\right) = u_0(x). \tag{63.96}$$

So we have to identify $u_0 : [0,l] \to \mathbb{R}$ with its Fourier development. Thanks to the division by l, this problem is equivalent to the Fourier development of a function on $[0,2\pi]$.

From the example 27.135, the function u_0 being real, it can be written as a series of sine ... at least as far as $u_0 \in L^2\big([0,l]\big)$.

Chapter 64

Finite elements

64.1 Lax-Milgram theorem

DEFooGFTZooUQfUdY

Definition 64.1.
*A bilinear form $a\colon V \times V \to \mathbb{R}$ is **elliptic** or **coercive** is there exists a $\alpha > 0$ such that $a(u,u) \geqslant \alpha\|u\|^2$ for every $u \in V$.*

THOooFDJYooCSNnuv

Theorem 64.2 (Lax-Milgram[631])**.**
Let V be ah Hilbert space and

(1) a linear and bounded map $L\colon V \to \mathbb{R}$; we write $\|L\| = C$,

(2) a bilinear map $a\colon V \times V \to \mathbb{R}$ is continuous; we write M a constant such that $|a(u,v)| \leqslant M\|u\|\|v\|$ for all $u, v \in V$,

(3) the bilinear form a is elliptic[1] and we write α a strictly positive constant such that $a(u,u) \geqslant \alpha\|u\|^2$.

Then the problem of finding $u \in V$ such that

$$a(u,v) = L(v) \tag{64.1}$$

for every $v \in V$ has one and only one solution $u \in V$. Moreover this solution satisfies

$$\|u\| \leqslant \frac{M}{\alpha} C. \tag{64.2}$$

The map L is linear and bounded; it is continuous by proposition 10.70. The existence of M is due to the fact that a is bilinear on V and in particular linear (and continuous, then bounded) on $V \times V$. But this is not quite obvious from the definition (7.226). It is shown in [622] that putting on $V \times V$ the product topology [2] that a sesquilinear map is continuous if and only if there exist such a constant. And since the topology of the product norm is the product topology (lemma 7.226), we are safe.

64.2 Variational formulation (not too rigorous)

As mentioned in the title, we are not going to deal with existence of the derivative and the integrals that we will write down.

Let the partial derivative equation EQooZAISooSylvFH

$$\begin{cases} -\Delta u = f & \text{(64.3a)} \\ u|_{\partial\Omega} = 0 & \text{(64.3b)} \end{cases}$$

where $\Delta u = \sum_{j=1}^{n} \frac{\partial^2 v}{\partial x_j}$ on the open bounded part Ω of $\mathbb{R}^n$.

1. Définition 64.1
2. Définition 7.15.

We are searching the solutions in a vector space

$$V = \{v\colon \Omega \to \mathbb{R} \text{ tel que } v|_{\partial\Omega} = 0\}. \tag{64.4}$$

Our aim is to found a bilinear form $a\colon V \times V \to \mathbb{R}$ and a linear map $L\colon V \to \mathbb{R}$ such that the solutions of the original problem (64.3) are solutions of the problem

$$\begin{cases} u \in V & \text{(64.5a)} \\ a(u,v) = L(v)\ \forall v \in V & \text{(64.5b)} \end{cases}$$

The choice of V, a and L is a **variational formulation** of the differential equation [3].

In order to have a variational formulation of the equation (64.3) we multiply $-\Delta u = f$ by a test function $v \in V$ and we integrate over Ω:

$$-\int_\Omega \Delta u\, v = \int_\Omega fv. \tag{64.6}$$

Now if we set

$$a(u,v) = -\int_\Omega (\Delta u)v \tag{64.7a}$$

SUBEQooKUNUooOtKVaP

$$L(v) = \int_\Omega fv \tag{64.7b}$$

we have a variational formulation of our problem. A solution of a variational formulation is a **weak solution** of the partial derivative equation.

Does the form (64.7a) check the hypothesis of the Lax-Milgram theorem 64.2? Obviously not because we did not defined the space V, so nothing has any sense here. But we can say more: the bilinear form a is not *obviously* positive. As we will see it is positive on V because of the boundary condition. We want to write is slightly differently in order to, taking into account the boundary condition, have a bilinear form that is for sure positive.

Using the integration by part of formula (57.216) taking into account the fact that the boundary term vanishes we have

$$\int_\Omega (\Delta u)v = -\int_\Omega \nabla u \cdot \nabla v, \tag{64.8}$$

so that we can as well consider the variational problem

$$a(u,v) = -\int_\Omega \nabla u \cdot \nabla v \tag{64.9a}$$

SUBEQooLFDKooTDKiDA

$$L(v) = \int_\Omega fv \tag{64.9b}$$

In this case, the form a is more clearly positive defined:

$$a(u,u) = \int_\Omega |\nabla u|^2 \geqslant 0. \tag{64.10}$$

Notons que cette formule pour a est symétrique et que nous n'avons pas encore démontré quoi que ce soit pour les hypothèses du théorème de Lax-Milgram. Nous espérons seulement que la forme bilinéaire (64.9a) ait de meilleures propriétés que (64.7a).

The result of this sections is the following.

Proposition 64.3 (Not too rigorous)**.**
A function $u \in V$ is solution of the variational problem if and only if it is solution of the Poisson equation.

3. I said "not too rigorous" in the title, so please don't ask yourself now what space V can be.

Proof. The fact that the solution of the Poisson equation (including the boundary conditions) are solutions of the variational problem is what we just did.

In the other sense we recall the equation:

$$\begin{cases} -\Delta u = f & (64.11a) \\ u|_{\partial\Omega} = 0 & (64.11b) \end{cases}$$

The variational problem is searching a function in V, that is a function that automatically satisfy the boundary condition. If u is solution of the variational problem, then

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv \tag{64.12}$$

We integrate by part the left hand side:

$$\int_\Omega \nabla u \cdot \nabla v = -\int_\Omega (\Delta u)v + \underbrace{\int_{\partial\Omega} \frac{\partial u}{\partial n} v}_{=0}, \tag{64.13}$$

so that

$$\int_\Omega (f - \Delta u)v = 0 \tag{64.14}$$

EQooAJMDooNJTYRm

for every $v \in V$.

If we really know nothing about the space V, we cannot conclude that $f - \Delta u = 0$. We can however do something that will probably work if V is not too strange. If $f - \Delta u \neq 0$ at some point $x \in \Omega$ (suppose $(f - \Delta u)(x) > 0$ in order to fix the ideas), then $f - \Delta u > 0$ on an open set A around x. If v is a positive function that vanishes outside A then, taking $B \subset A$ on which $v > 0$,

$$\int_\Omega (f - \Delta u)v = \int_A (f - \Delta u)v > \int_B (f - \Delta u)v. \tag{64.15}$$

The last integral is for sure strictly positive, which contradicts (64.14). □

64.3 Galerkin's approximation

Let us once again be not too rigorous and deal with the problem

$$\begin{cases} -u'' + u = f & (64.16a) \\ u(0) = u(1) = 0 & (64.16b) \end{cases}$$

on the open interval $]0,1[$. The good functional space seems to be

$$H_0^1\big(]0,1[\big) = \{u \in H^1\big(]0,1[\big) \text{ such that } u(0) = u(1) = 0\}. \tag{64.17}$$

Of course, this definition is not rigorous because the elements in the Sobolev spaces are classes of functions and the boundary values are not defined. Let us go on and see what happens.

Let $v \in H_0^1$. We multiply the equation by v and integrate over the interval $I =]0,1[$:

$$-\langle u'', v\rangle + \langle u, v\rangle = \langle f, v\rangle, \tag{64.18}$$

and an integration by part, taking into account the fact that v vanishes at the border gives

$$\langle u', v'\rangle + \langle u, v\rangle = \langle f, v\rangle, \tag{64.19}$$

If $u \in C^2(I)$, the two formulations are equivalent. If not, we are not sure. The point of the second formulation is that one can build a piecewise affine approximation. We divide the interval I into $N+1$ pieces

$$x_j = \frac{j}{N+1} \tag{64.20}$$

with $j = 0, \dots, N+1$. Let V_N be the set of continuous piecewise affine functions that are vanishing on the border:

$$V_N = \{v \in C^0(I) \text{ such that } v|_{]x_j, x_{j+1}[} \text{ is linear and } v(0) = v(1) = 0\}. \tag{64.21}$$

This is a finite dimensional vector space because the elements are determined by the values on the x_i's. Moreover the space V_N is included in $H_0^1(I)$.

Proposition 64.4.
There exists an unique element $u_N \in V_N$ satisfying the equation

$$\langle u_n', v' \rangle + \langle u_N, v \rangle = \langle f, v \rangle \tag{64.22}$$

EQooOFLCooHmjaOM

*for all $v \in V_N$. This solution is the **Galerkin approximation**.*

Proof. We consider the basis $\{\phi_j\}_{j=1,\dots,N}$ of V_N defined by

$$\phi_j(x_i) = \delta_{ij}. \tag{64.23}$$

We are searching for u_N under the form $u_N = \sum_{j=1}^N a_j \phi_j$. Just by computing on the point x_j we know that

$$a_j = u_N(x_j). \tag{64.24}$$

At this moment, this equality does not help, but we keep it in mind. Since the equality (64.22) has to hold for every $v \in V_N$, it holds in particular for $v = \phi_k$:

$$\langle u_N', \phi_k' \rangle + \langle u_N, \phi_k \rangle = \langle f, \phi_k \rangle \tag{64.25}$$

and by linearity of the inner product,

$$\sum_l a_l \big(\langle \phi_k', \phi_l' \rangle + \langle \phi_l, \phi_k \rangle \big) = \langle f, \phi_k \rangle. \tag{64.26}$$

If we set

$$R_{kl} = \langle \phi_k', \phi_l' \rangle + \langle \phi_k, \phi_l \rangle \tag{64.27a}$$
$$b_k = \langle f, \phi_k \rangle, \tag{64.27b}$$

we have to solve the linear system

$$Ra = b. \tag{64.28}$$

In order to show that this system has an unique solution, we have to get some informations about the matrix R. The matrix R is the matrix of the 2-form

$$R(f, g) = \langle f', g' \rangle + \langle f, g \rangle \tag{64.29}$$

in the basis $\{\phi_j\}$ of V_N.

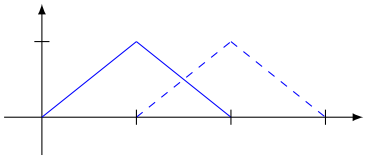
(1) R **is strictly positive defined** We have $R(f, f) \geqslant 0$ and if $R(f, f) = 0$, then $\langle f, f \rangle = 0$ and $\langle f', f' \rangle = 0$. Thus $f = 0$ almost everywhere and since elements of V_N are continuous, $f = 0$.

(2) R **is symmetric** Clear from the definition.

The matrix is thus invertible (in fact we do not use the symmetry to reach this conclusion) and the system has an unique solution in V_N.

□

We can compute the matrix R: the elements are only some inner products and integrals. Here is a graph of ϕ_1 and ϕ_2:



The affines pieces are:

$$f_1(x) = (N+1)x \tag{64.30a}$$
$$f_2(x) = -(N+1)x + 2 \tag{64.30b}$$
$$f_3(x) = (N+1)x - 1 \tag{64.30c}$$
$$f_4(x) = -(N+1)x + 3. \tag{64.30d}$$

And we have to integrate. We make the computations:

```
var('N')
h=1/(N+1)
f1=(N+1)*x
f2=-(N+1)*x+2
f3=(N+1)*x-1
f4=-(N+1)*x+3

inner11=(f1*f1).integrate(x,0,h)+(f2*f2).integrate(x,h,2*h)
inner12=(f2*f3).integrate(x,h,2*h)

print(inner11.simplify_full())
print(inner12.simplify_full())

# For the inner products of the derivatives :

d_inner11=(f1.derivative(x)*f1.derivative(x)).integrate(x,0,h)+(↩
   f2.derivative(x)*f2.derivative(x)).integrate(x,h,2*h)
d_inner12=(f2.derivative(x)*f3.derivative(x)).integrate(x,h,2*h)

print(d_inner11.simplify_full())
print(d_inner12.simplify_full())
```

tex/sage/sageSnip008.sage

returns

```
2/3/(N + 1)
1/6/(N + 1)
2*N + 2
-N - 1
```

which means that

$$\langle \phi_j, \phi_j \rangle = \frac{2}{3(N+1)} \tag{64.31a}$$
$$\langle \phi_j, \phi_{j+1} \rangle = \langle \phi_j, \phi_{j-1} \rangle = \langle f_2, f_3 \rangle = \frac{1}{6(N+1)} \tag{64.31b}$$
$$\langle \phi_j', \phi_j' \rangle = 2N + 2 \tag{64.31c}$$
$$\langle \phi_j', \phi_{j+1}' \rangle = -N - 1 \tag{64.31d}$$

The inner product $\langle \phi_i, \phi_j \rangle$ is zero when $|i-j| \geqslant 2$.

64.4 Gradient on a boundary

Let Ω be an open subset of $\mathbb{R}^3$ and a smooth function $f\colon \mathbb{R}^3 \to \mathbb{R}$ that assumes positives and negatives values in Ω. Let $\Omega_1 = \Omega \cap \{f > 0\}$, $\Omega_2 = \Omega \cap \{f < 0\}$ and $C = \Omega \cap \{f = 0\}$. We have the decomposition

$$\Omega = \Omega_1 \cup \Omega_2 \cup C \tag{64.32}$$

A typical example is $\Omega = B(0,1)$ subdivided into $\Omega_1 = B(0,1) \cap \{z > 0\}$, $\Omega_2 = B(0,1) \cap \{z < 0\}$ and $C = \{(x, y, 0) \text{ tel que } x^2 + y^2 < 1\}$.

Let u be a vector field defined on Ω by

$$u(x) = \begin{cases} u_1(x) & \text{if } x \in \Omega_1 \\ u_2(x) & \text{if } x \in \Omega_2 \end{cases} \tag{64.33}$$

where u_1 and u_2 are defined on Ω. We do not define u on C because it is of measure zero, but we consider the step function

$$s = u_2 - u_1 \tag{64.34}$$

that is defined on Ω. We assume that u is derivable on $\Omega \backslash C$.

Let $T_u \in \mathscr{D}'(\Omega)$ be the distribution associated with u. We compute its gradient:

$$\langle \nabla \cdot T_u, \phi \rangle = -\langle T_u, \nabla \cdot u \rangle = -\int_\Omega u \cdot \nabla\phi = -\int_{\Omega_1} u_1 \cdot \nabla\phi - \int_{\Omega_2} u_2 \cdot \nabla\phi. \tag{64.35}$$

Here we used the fact that $\Omega = \Omega_1 \cup \Omega_2 \cup C$ while the integral on C is zero since we are computing a three-dimensional integral. We use the integration by part (24.55):

$$\int_{\Omega_1} u_1 \cdot \nabla\phi = \int_{\partial\Omega_1} \phi u_1 \cdot n_1 - \int_{\Omega_1} \phi \nabla \cdot u_1. \tag{64.36}$$

EQooTKJDooQopGtW

An element of $\partial\Omega_1$ is in particular the limit of a sequence in $\text{Int}(\Omega_1)$. The limit can *a priori* belong in one of the following: $\text{Int}(\Omega)$, $\text{Int}(\Omega_1)$, $\text{Int}(\Omega_2)$, $\partial\Omega$, $\partial\Omega_1$, $\partial\Omega_2$ or C. Let us review them

— Since we are speaking about an element of $\partial\Omega_1$,it cannot belong to $\text{Int}(\Omega_1)$.

— Since Ω_1 and Ω_2 have no intersection, the limit of a sequence contained in Ω_1 cannot belong to $\text{Int}(\Omega_2)$.

— If the element we are speaking about belong to $\partial\Omega_2$, then it belong to $\partial\Omega_1 \cap \partial\Omega_2 = C$.

— Let $a \in \partial\Omega_1 \cap \text{Int}(\Omega)$. A neighbourhood around a cannot be completely contained in Ω_1 or Ω_2, but it has to have parts in each of these two sets (because C has lower dimension). Thus it has to belong to $\partial\Omega_1 \cap \partial\Omega_2 = C$.

Thus we are left with $\partial\Omega_1 \subset C \cup \partial\Omega$. Then one can write

$$\int_{\partial\Omega_1} \phi u_1 \cdot n_1 \leqslant= \int_C \phi u_1 \cdot n_1 + \int_{\partial\Omega \cap \partial\Omega_1} \phi u_1 \cdot n_1. \tag{64.37}$$

But $\phi \in \mathscr{D}(\Omega)$ and Ω is open, so ϕ vanishes on the boundary of Ω, so that the second term is zero. One can thus reduce the integral over $\partial\Omega_1$ into an integral over C in the equation (64.36):

$$\langle \nabla \cdot T_u, \phi \rangle = \int_{\Omega_1} \phi \nabla \cdot u_1 - \int_C \phi u_1 \cdot n_1 + \int_{\Omega_2} \phi \nabla \cdot u_2 - \int_C \phi u_2 \cdot n_2. \tag{64.38}$$

Since n_i is normal and exterior to Ω_i, we have $n_1 = -n_2$ on C. By convention, we name $n = n_1$ and $s = u_2 - u_1$. So we have

$$u_1 \cdot n_1 + u_2 \cdot n_2 = u_1 \cdot n - u_2 \cdot n = -s \cdot n \tag{64.39}$$

and we can write

$$\langle \nabla \cdot T_u, \phi \rangle = \sum_{i=1}^{2} \int_{\Omega_i} \phi \nabla \cdot u_i + \int_C (s \cdot n)\phi. \tag{64.40}$$

The whole concludes in

$$\nabla \cdot T_u = T_{\nabla \cdot u} + \delta_C^{s \cdot n} \tag{64.41}$$

where we used the Dirac "generalization" of 61.2.

64.5 Locally integrable functions

Definition 64.5.
*A function $f\colon \Omega \to \mathbb{C}$ is **locally integrable** if $f \in L^1(K)$ for every compact K in Ω. The set of locally integrable functions on Ω is denoted by $L^1_{loc}(\Omega)$*

Proposition 64.6.
We have the inclusion $L^2(\Omega) \subset L^1_{loc}(\Omega)$.

Proof. Let K be compact in Ω in $\mathbb{R}^d$ and $f \in L^2(\Omega)$. We have to show $f \in L^1(K)$. Using the Cauchy-Schwarz inequality 10.1 in $L^2(K)$ we have

$$\int_K |f| = \langle f, 1\rangle_{L^2(K)} \leqslant \|f\|_{L^2(K)}\|1\|_{L^2(K)} \leqslant \|f\|_{L^2(\Omega)} \operatorname{Vol}(K) < \infty. \tag{64.42}$$

Since $f \in L^2(\Omega)$, □

64.6 An approximation result

Theorem 64.7 ([947])**.**
Let Ω be an open set in $\mathbb{R}^3$ subdivided into a finite number of subdomains Ω_i. We suppose that for each i, the polynomials of degree n are part of $H^1(\Omega_i)$.

We consider a function $u\colon \Omega \to \Omega$ which is a polynomial of degree n on each of the subdomains Ω_i and we suppose it to be continuous on the boundaries of Ω_i.

(1) Then $u \in H^1(\Omega)$

(2) If we suppose that u is differentiable on each boundary, then $u \in H^2(\Omega)$.

Proof. First, the function u belongs to $L^2(\Omega)$ because the integral of $|u|^2$ on Ω reduces to the integral over the interior of each Ω_i (because the boundaries have zero measure). Since u belongs to $L^2(\Omega_i)$ for each i, we have $u \in L^2(\Omega)$.

Now we have to prove that the partial derivatives (in the weak sense) belong to $L^2(\Omega)$ too. Since u is a polynomial on the interiors we can consider the function

$$f_\alpha(x) = \frac{\partial u}{\partial x_\alpha} \tag{64.43}$$

on $\operatorname{Int}(\Omega_i)$ (for each i). Here the partial derivative is not in the weak sense. We do not define the function f_α on the boundaries. Let $\varphi \in \mathscr{D}(\Omega_i)$ and perform an integration by part:

$$\int_{\Omega_i} f_\alpha \varphi = \int_{\Omega_i} \frac{\partial u}{\partial x_\alpha}\varphi = -\int_{\Omega_i} u\frac{\partial \varphi}{\partial x_\alpha} + \int_{\partial\Omega_i} u\varphi(n_i \cdot e_\alpha) \tag{64.44}$$

EQooMGESooSkBybZ

where n_i is the normal vector field to $\partial\Omega_i$. In order to make sense, the integrals over Ω_i are in fact integrals over $\operatorname{Int}(\Omega_i)$ because f_α and $\partial_\alpha u$ are only well defined on the interior.

We sum (64.44) over i:

$$\int_\Omega f_\alpha\varphi = \sum_i \int_{\Omega_i} f_\alpha\varphi = -\int_\Omega u\frac{\partial\varphi}{\partial x_\alpha} + \sum_i \int_{\partial\Omega_i} u\varphi(n_i \cdot e_\alpha). \tag{64.45}$$

EQooKACAooGlBMaQ

The value of u on $\partial\Omega_i$ is univoque since we assume that u is continuous on the boundaries.

The set $\bigcup_i \partial\Omega_i$ can be subdivided into two parts. Some points are in $\partial\Omega$ and the other ones are on the intersections $\partial\Omega_i \cap \partial\Omega_j$.

The function φ vanishes on $\partial\Omega$. And the contribution of $\partial\Omega$ in $\sum_i \int_{\partial\Omega_i} u\varphi(n_i \cdot e_\alpha)$ is zero.

The contribution of the intersections in this integral is double: the intersection $\partial\Omega_i \cap \partial\Omega_j$ comes in the integral over $\partial\Omega_i$ and in the one over $\partial\Omega_j$. Since on the intersection we have $n_i = -n_j$

(because they are outwards), the sum vanishes and the whole sum of integrals over $\partial\Omega_i$ in (64.45) disappear. Here we also use the fact that u is continuous on the intersections. We are left with

$$\int_\Omega f_\alpha \varphi = -\int_\Omega u \frac{\partial \varphi}{\partial x_\alpha}. \tag{64.46}$$

We proved that setting

$$f_\alpha(x) = \begin{cases} \frac{\partial u}{\partial x_\alpha} & \text{if } x \in \operatorname{Int}(\Omega_i) \text{for some } i \\ \text{whatever} & \text{otherwise} \end{cases} \tag{64.47}$$

we get

$$\langle f_\alpha, \varphi \rangle = -\langle u, \partial_\alpha \varphi \rangle \tag{64.48}$$

for every $\varphi \in \mathscr{D}(\Omega)$. This shows that f_α is the weak derivative of u.

Moreover f_α is a sum of polynomials of degree $n-1$ on Ω and is thus integrable. So by hypothesis $f_\alpha \in H^1(\Omega) \subset L^2(\Omega)$. □

64.7 Lax-Milgram with a boundary condition

We want to give an example of use of the Sobolev space $H^1_0(\Omega)$ defined in 61.18.

Exemple 64.8.
Let the differential equation for $u\colon [0,1] \to \mathbb{R}$

$$\begin{cases} -u''(x) = f(x) & \text{(64.49a)} \\ u(0) = u(1) = 0 & \text{(64.49b)} \end{cases}$$

with f in a not yet well precise functional space. The most obvious prescription for the functional spaces if to ask $f \in L^2$ and then $u \in H^2$, so that u has L^2 second derivative. Instead of that, we will look at the variational form of the problem. We consider $w \in \mathscr{D}\big([0,1]\big)$ satisfying $w(0) = w(1) = 0$ and compute the inner product between v and the equation:

$$-\langle u'', w \rangle = \langle f, w \rangle. \tag{64.50}$$

An integration by part produces the equation

$$\int_0^1 u'w' = \int_0^1 fw. \tag{64.51}$$

EQooRTXVooYmoAJM

Since no precision is provided about the functional spaces, the equation (64.51) is by no means related to initial equation. We only hope that the solution of the variational problem will be a solution of the initial problem, and that we will be able to furnish a functional setting in which the integral by part makes sense.

Since the variational formulation only needs first derivative of u and v we are lead to consider $H^1\big(]0,1[\big)$. Let us write down the variational problem.

Let $f \in L^2\big(]0,1[\big)$ and consider the space $V = H^1_0\big(]0,1[\big)$. We define the functionals defined by

$$\begin{aligned} a\colon V \times V &\to \mathbb{C} \\ (u,w) &\mapsto \int_0^1 u'w' \end{aligned} \tag{64.52}$$

and

$$\begin{aligned} l\colon V &\to \mathbb{C} \\ w &\mapsto \int_0^1 fw. \end{aligned} \tag{64.53}$$

We are searching for $u \in V$ such that $a(u, \cdot\,) = l$, that is such that

$$a(u,w) = l(w) \tag{64.54}$$

for every $w \in V$.

Let us check the hypothesis of the Lax-Milgram theorem 64.2.

(1) $l\colon H_0^1 \to \mathbb{R}$ **is continuous**

Let $w_i \xrightarrow{H_0^1} 0$. Using the Cauchy-Schwarz inequality 10.1 we have

$$|l(w_i)|^2 = |\langle w_i, f\rangle|^2 \leqslant \|w_i\|_{L^2}\|f\|_{L^2} \leqslant C|w_i|_{1,\Omega}\|f\|_{L^2} \tag{64.55}$$

where we used the Poincaré inequality of theorem 61.22. By hypothesis we have $w_i \xrightarrow{H^1(\Omega)} 0$ and in particular $|w_i|_{1,\Omega} \to 0$, so that

$$|l(w_i)|^2 \to 0, \tag{64.56}$$

and l is continuous.

(2) $a\colon H_0^1 \times H_0^1 \to \mathbb{R}$ **is continuous**

By definition

$$a(u,v) = \langle u', w'\rangle_{L^2\left(]0,1[\right)}. \tag{64.57}$$

If $(u_i, w_i) \xrightarrow{H_0^1\times H_0^1} (0,0)$ we have

$$|a(u_i, w_i)| \leqslant \|u'\|_{L^2}\|w'\|_{L^2} = |u_i|_{1,\Omega}|w_i|_{1,\Omega} \to 0. \tag{64.58}$$

<++>

<++>

△

Chapter 65

Chain complexes

65.1 Chain complexes

65.1.1 Homotopy groups

The reference for (co)chain complexes is [955].

Let $\mathbb{A}$ be a ring and $(M,+)$ be a commutative group. The triple $(M,+,\ \cdot\ )$ is a $\mathbb{A}$-left module if the operation $\ \cdot\ : \mathbb{A}\times M\to M$ fulfils

(1) $a\cdot(x+y)=a\cdot x+a\cdot y$,

(2) $(a+b)\cdot x=a\cdot x+b\cdot x$,

(3) $(ab)\cdot x=a\cdot(b\cdot x)$,

(4) $1\cdot x=x$

for every a, $b\in\mathbb{A}$ and x, $y\in M$. When M is a field, then we have the notion of vector space where $\mathbb{A}$ is the set of scalars. Let now $\mathbb{A}$ be an unital commutative ring. All modules are now taken over $\mathbb{A}$.

Proposition 65.1.
Every abelian group is a $\mathbb{Z}$-module.

Proof. Let $(M,+)$ be a commutative group. The structure of $\mathbb{Z}$-module is given by

$$n\ \cdot\ x=\begin{cases}0 & \text{if } n=0\\ \underbrace{x+\cdots+x}_{n\text{ terms}} & \text{if } n>0\\ -(\underbrace{x+\cdots+x}_{n\text{ terms}}) & \text{if} n<0.\end{cases}$$

□

This is moreover the only way to turn $(M,+)$ into a $\mathbb{Z}$-module.

A **chain complex** $C=C(C_*,\partial_*)$ is a sequence of modules $(C_n)_{n\in\mathbb{N}}$ and of morphisms of modules $\partial_{n+1}\colon C_{n+1}\to C_n$ such that $\partial_*^2=0$. By convention, we put $C_{-1}=0$ and $\partial_0=0$. The morphisms ∂_n are the **boundary morphisms** of C.

If C and D are chain complexes, a **morphism** $\phi\colon C\to D$ is a sequence of modules morphisms $\phi_n\colon C_n\to D_n$ such that $\phi_n\circ\partial_{n+1}=\partial_{n+1}\circ\phi_{n+1}$. The diagram

$$\begin{array}{ccccccccccc} C_0 & \xleftarrow{\partial_0} & C_1 & \xleftarrow{\partial_1} & C_2 & \longleftarrow & \dots & \longleftarrow & C_{n-1} & \xleftarrow{\partial_n} & C_n & \longleftarrow & \dots \\ \downarrow{\scriptstyle\phi_0} & & \downarrow{\scriptstyle\phi_1} & & & & & & \downarrow{\scriptstyle\phi_{n-1}} & & \downarrow{\scriptstyle\phi_n} & & \\ D_0 & \xleftarrow[\partial_0]{} & D_1 & \xleftarrow[\partial_1]{} & D_2 & \longleftarrow & \dots & \longleftarrow & D_{n-1} & \xleftarrow[\partial_n]{} & D_n & \longleftarrow & \dots \end{array} \tag{65.1}$$

commutes.

The submodule

$$Z_n(C) = \ker(\partial\colon C_n \to C_{n-1}) \tag{65.2}$$

of C_n is the space of n-**cycle**. The submodule

$$B_n(C) = \text{Image}(\partial\colon C_{n+1} \to C_n) \tag{65.3}$$

of C_n is the space of n-**boundary**. The nth **homology group** of C is the quotient

$$H_n(C) = Z_n(C)/B_n(C). \tag{65.4}$$

The homology group $H_n(C)$ has a $\mathbb{A}$-module structure. The direct sum $\bigoplus H_n(C)$ is denoted by $H_*(C)$. If $c \in Z_n(C)$, we denote by $[c]$ the image of c in $H_n(C)$:

$$[c] = \{c + \partial d \text{ tel que } d \in C_{n+1}\}. \tag{65.5}$$

Since a chain morphism preserves the boundary operator ∂, it induces a module morphism

$$\phi_*\colon H_n(C) \to H_n(D).$$

If $\phi, \psi\colon C \to D$ are two morphisms of chain complexes, an **homotopy** is a sequence of module morphisms $K_n\colon C_n \to D_{n+1}$ such that

$$\phi_n - \psi_n = \partial_{n+1} \circ K_n + K_{n-1} \circ \partial_n. \tag{65.6}$$ EqDefmorch

Proposition 65.2.
Let ϕ and ψ be two morphisms between the chain C and D. Then we have

$$\phi_* = \psi_*\colon H_*(C) \to H_*(D).$$

when they are related by an homotopy.

Proof. An element of $H_*(C)$ reads $[c]$ with $c \in Z_*(C)$. From $\partial c = 0$ and (65.6), we have $(\phi_n - \psi_n)(c) = \partial_{n+1}\big(K_n(c)\big)$, so that $\phi_n(c) = \psi_n(c) + \partial_{n+1}\big(K_n(c)\big)$, which means that $[\phi_n(c)] = [\psi_n(c)]$. □

65.1.2 Exact sequences

A **short exact sequence** of chain complexes is three complexes C, D and E with morphisms $g\colon D \to E$ and $f\colon C \to D$ such that for every n, the sequence

$$0 \longrightarrow C_n \xrightarrow{f_n} D_n \xrightarrow{g_n} E_n \longrightarrow 0$$

is exact. The short exact sequence is logically denoted by

$$0 \longrightarrow C \xrightarrow{f} D \xrightarrow{g} E \longrightarrow 0.$$

A **morphism of short exact sequences** of chain complexes is a triple (α, β, γ) of morphisms of chain complexes such that the following diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & C & \xrightarrow{f} & D & \xrightarrow{g} & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} & & \\ 0 & \longrightarrow & C' & \xrightarrow[f']{} & D' & \xrightarrow[g']{} & E' & \longrightarrow & 0 \end{array} \tag{65.7}$$ EqmorexactseqM

commutes for each degree.

Theorem 65.3.
If $0 \longrightarrow C \xrightarrow{f} D \xrightarrow{g} E \longrightarrow 0.$ *is a short exact sequence of chain complexes, there exists a long exact sequence of modules*

$$\dots \longrightarrow H_{n+1}(E) \xrightarrow{\delta} H_n(C) \xrightarrow{f_*} H_n(D) \xrightarrow{g_*} H_n(E) \xrightarrow{\delta} H_{n-1}(C) \longrightarrow \dots$$

such that if (α, β, γ) *is as in* (65.7)*, then the diagram*

$$\begin{array}{ccccccccc} \dots & \longrightarrow & H_n(C) & \xrightarrow{f_*} & H_n(D) & \xrightarrow{g_*} & H_n(E) & \xrightarrow{\delta} & \dots \\ & & \downarrow{\scriptstyle \alpha_*} & & \downarrow{\scriptstyle \beta_*} & & \downarrow{\scriptstyle \gamma_*} & & \\ \dots & \longrightarrow & H_n(C) & \xrightarrow{f_*} & H_n(D) & \xrightarrow{g_*} & H_n(E) & \xrightarrow{\delta} & \dots \end{array} \tag{65.8}$$

is commutative.

Proof. No proof. □

65.1.3 Euler-Poincaré

Let $\mathbb{A}$ be a field and C, a chain complex on $\mathbb{A}$. We denote by β_k the dimension of $H_k(C)$ as vector space on $\mathbb{A}$. This is the kth **Betty number** of C. If they are all finite and if they vanish for sufficiently large k, we pose

$$\chi(C) = \sum_{j \in \mathbb{N}} (-1)^k \beta_k, \tag{65.9}$$

and we name it the **Euler-Poincaré characteristic** of C.

65.2 Cochain complexes

A **cochain complex** is $C = (C^*, d^*)$ where $(C^n)_{n\in\mathbb{N}}$ is a sequence of modules with morphisms $d^n \colon C^n \to C^{n+1}$ such that $d^{n+1} \circ d^n = 0$. The latter is regularly written under the more compact form $d^2 = 0$. The morphisms d^* are the **coboundary morphisms**. By convention, $C^{-1} = 0$ and $d^{-1} = 0$.

A **morphism of cochain complex** is a sequence $f \colon C \to D$ of morphisms of modules $f^n \colon C^n \to D^n$. The other concepts are defined as before.

Theorem 65.4.
If

$$0 \longrightarrow C \xrightarrow{f} D \xrightarrow{g} E \longrightarrow 0. \tag{65.10}$$

EqSeqThoo

is a short exact sequence of cochain complexes, there exists a long exact sequence of modules

$$\dots \longrightarrow H^{n-1}(E) \xrightarrow{\delta} H^n(C) \xrightarrow{f^*} H^n(D) \xrightarrow{g^*} H^n(E) \xrightarrow{\delta} H^{n+1}(C) \longrightarrow \dots$$

such that if (α, β, γ) *is a morphism of exact sequence of cochain between* (65.10) *and*

$$0 \longrightarrow C' \xrightarrow{f'} D' \xrightarrow{g'} E' \longrightarrow 0, \tag{65.11}$$

then the diagram

$$\begin{array}{ccccccccc} \dots & \longrightarrow & H^n(C) & \xrightarrow{f^*} & H^n(D) & \xrightarrow{g^*} & H^n(E) & \xrightarrow{\delta} & H^{n+1}(C) \longrightarrow \dots \\ & & \downarrow{\scriptstyle \alpha^*} & & \downarrow{\scriptstyle \beta^*} & & \downarrow{\scriptstyle \gamma^*} & & \downarrow{\scriptstyle \alpha^*} \\ \dots & \longrightarrow & H^n(C') & \xrightarrow[f'^*]{} & H^n(D') & \xrightarrow[g'^*]{} & H^n(E') & \xrightarrow[\delta]{} & H^{n+1}(C') \longrightarrow \dots \end{array} \tag{65.12}$$

is commutative.

Proof. No proof. □

65.2.1 Singular chains

Let $p \in \mathbb{N}$. The ordered **standard p-simplex** Δ_p is the affine convex envelop in $\mathbb{R}^{p+1}$ of the canonical basis $\{e_0, \dots, e_p\}$:

$$\Delta_p = \{(t_0, \dots, t_p) \in \mathbb{R}^{p+1} \text{ tel que } t_i \geqslant 0, \text{ and } \sum_i t_i = 1\}. \tag{65.13}$$

When X is a topological space, a **singular p-simplex** in X is a continuous map $\sigma\colon \Delta_p \to X$. The 0-simplexes in X are the points of X while the 1-simplexes are path in X as can be seen on the simple explicit expression

$$\Delta_1 = \{(\lambda, 1-\lambda) \text{ tel que } \lambda \in [0,1]\}.$$

We denote by $C_p(X, \mathbb{A})$ the free module over $\mathbb{A}$ whose basis is the set of p-simplexes in X. A general element of this set reads

$$\sum_{i=1}^{k} n_i \sigma_i \tag{65.14}$$ EqExpchainesigma

with $n_i \in \mathbb{A}$ and where σ_i is a singular p-simplex. These elements are called **singular p-chain**. The **face** i of the singular p-simplex σ, denoted by $\partial_i\sigma$ is the $(p-1)$-singular simplex given by

$$(\partial_i\sigma)(t_0, \dots, t_{p-1}) = \sigma(t_0, \dots, t_{i-1}, 0, t_i, \dots, t_{p-1}).$$

The **boundary** of the singular p-chain 65.14 is

$$\partial\sigma = \sum_{i=0}^{p} (-1)^i \partial_i\sigma \in C_{p-1}(X, \mathbb{A}). \tag{65.15}$$

By linearity, we define the morphism

$$\partial\colon C_p(X, \mathbb{A}) \to C_{p-1}(X, \mathbb{A}), \tag{65.16}$$

and by convention we pose $C_{-1}(X) = \{0\}$ and $\partial\colon C_0(X) \to C_{-1}(X)$ to be the zero map.

Lemma 65.5.
The boundary morphism ∂ fulfils $\partial \circ \partial = 0$.

Proof. No proof. □

Thus we can consider the following chain complex:

$$C_0(X) \xleftarrow{\partial} C_1(X) \xleftarrow{\partial} \dots \xleftarrow{\partial} C_{n-1}(X) \xleftarrow{\partial} C_n(X) \xleftarrow{\partial} \dots \tag{65.17}$$

which is the **complex of singular chains** whose n-cycles are the **singular n-cycles** with coefficients in $\mathbb{A}$. The definitions of $Z_n(X, \mathbb{A})$ and $B_n(X, \mathbb{A})$ are as usual and the **singular homology** with coefficients in $\mathbb{A}$ is

$$H_n(X, \mathbb{A}) = Z_n(X, \mathbb{A})/B_n(X, \mathbb{A}). \tag{65.18}$$

65.3 Singular homology

Let $\mathbb{A}$ be a unitary commutative ring. In this section, all modules are taken over $\mathbb{A}$. First we fix a module M that we call the **module of coefficients**. When N is a module, we denote by $\mathrm{Hom}_{\mathbb{A}}(N, M)$ the module of morphisms from N to M. In particular, the module $\mathrm{Hom}_{\mathbb{A}}(N, \mathbb{A})$ is the **dual module** of N. When $f\colon N \to N'$ is a morphism of module, we define $f^t\colon \mathrm{Hom}(N', M) \to \mathrm{Hom}(N, M)$, the module morphism defined by

$$f^t(\phi) = \phi \circ f. \tag{65.19}$$

It automatically satisfies

$$\mathrm{Id}^t = \mathrm{Id} \qquad \text{and} \qquad (f \circ g)^t = g^t \circ f^t. \tag{65.20}$$

If we denote by $\mathscr{C}$ the category of modules over $\mathbb{A}$, what we just did is to define $\mathrm{Hom}_{\mathbb{A}}(.,M)$, a contravariant functor from $\mathscr{C}$ to itself. Indeed, let $N \in \mathscr{C}$, of course we have $\mathrm{Hom}_{\mathbb{A}}(N,M) \in \mathscr{C}$. Now if N and N' are objects of $\mathscr{C}$, then for every $f \in \hom(N,N')$, we have to define the image of the arrow f by the functor $\mathrm{Hom}_{\mathbb{A}}(.,M)$. This is f^t which has the right properties given on page 2915.

Let now take two modules N and N' and a linear map $\phi\colon N \times N' \to M$. It induces a morphism of module from N to $\mathrm{Hom}_{\mathbb{A}}(N',M)$ by

$$x \mapsto \phi(x,.). \tag{65.21}$$

Proposition 65.6.
If

$$C \xrightarrow{f} C' \xrightarrow{g} C'' \longrightarrow 0$$

is an exact sequence of modules, then

$$\mathrm{Hom}_{\mathbb{A}}(C,M) \xleftarrow{f^t} \mathrm{Hom}_{\mathbb{A}}(C',M) \xleftarrow{g^t} \mathrm{Hom}_{\mathbb{A}}(C'',M) \longleftarrow 0$$

is an exact sequence of modules. If f is injective, and if C'' is a free module, then f^t is surjective.

Proof. No proof. □

Chapter 66

Homogeneous and symmetric spaces

Most of the material of this section can be found in a more general framework in the references [813, 956, 838, 957].

66.1 Action of groups on sets

Recall that the action of a group is **transitive** when it has only one orbit (i.e. each point can reach anyone other point). An action is **free** if the fact that $g \cdot x = x$ for all $x \in M$ implies $g = e$. In other words, the action is free when e is the only element to be represented by the identity. The action is **simply transitive** when it has only one orbit and the stabilizer of one point is reduced to identity, in other words when $\forall\, (x, y) \in M^2$, $\exists!\, g \in G$ such that $x \cdot g = y$.

LemCompactSurFermeFerme

Lemma 66.1.
Let G be a Lie group acting on a manifold M. Consider K, a compact subgroup of G and F, a closed set in M. The set $K \cdot F$ is closed in M.

Proof. We will prove that any sequence in $K \cdot F$ which converges in M converges in $K \cdot F$. Let $\{k_n\} \in K$ and $\{\xi_n\} \in F$ and suppose that the sequence $\phi_n = k_n \cdot \xi_n$ converges to $\phi \in M$.

Since K is compact, the sequence $\{k_n\}$ has a converging subsequence. Thus, without loss of generality, we can suppose that $k_n \to k \in K$ and $k_n \cdot \xi_n \to \phi \in M$. Since we are considering the action of a group, and since K is a subgroup, we also have $k_n^{-1} \cdot \phi_n = \xi_n$. The action being continuous on M, we have

$$k_n^{-1} \cdot \phi_n \to k^{-1} \cdot \phi, \tag{66.1}$$

so that $\xi_n \to k^{-1} \cdot \phi$. But $\{\xi_n\}$ is a sequence in the closed space F. Thus its limit must belong to F: we have $\phi \in F$. Thus $k^{-1} \cdot \phi \in F$ and finally $\phi \in K \cdot F$. □

66.1.1 Fundamental and invariant fields

Subsec_Funda_conv

Let G be a Lie group with Lie algebra $\mathfrak{g}$. For each element of $\mathfrak{g}$, there are two distinguished vector fields on G, the **left invariant** and the **right invariant** one:

$$\tilde{X}_g = \frac{d}{dt}\Big[ge^{tX}\Big]_{t=0} \qquad \underline{X}_g = \frac{d}{dt}\Big[e^{tX}g\Big]_{t=0} \tag{66.2}$$

$$dL_h\tilde{X}_g = \tilde{X}_{hg} \qquad dR_h\underline{X}_g = \underline{X}_{gh}. \tag{66.3}$$

When G is a Lie group with an action on the manifold M denoted by

$$\begin{aligned} \tau\colon G \times M &\to M \\ (g, x) &\mapsto \tau_g(x), \end{aligned} \tag{66.4}$$

we define the **fundamental vector field** associated with $X \in \mathfrak{g}$ on the point $x \in M$ by

$$X_x^* = \frac{d}{dt}\Big[\tau_{e^{-tX}}(x)\Big]_{t=0}. \tag{66.5}$$

EqDefChmpFonfOff

An usual case is the one of a Lie group acting on itself for which we have

$$X_g^* = \frac{d}{dt}\Big[e^{-tX}g\Big]_{t=0}. \tag{66.6}$$

66.2 Rough introduction to homogeneous spaces

pg:esp_homo

An **homogeneous space** is a differentiable manifold with a transitive diffeomorphism group.

An important class of homogeneous space is given by the coset spaces. When we have a topological group G and a closed subgroup H, the coset space G/H has a structure of homogeneous space. Theorem 53.100 shows that almost every homogeneous space is of this class. For this, we use classes on right:

$$[g] = \{gh : h \in H\}.$$

The canonic projection is $\pi\colon G \to M$ and we denote $\vartheta = [e]$. The following construction shows that (almost [1]) every homogeneous space are of this kind.

Let M be a homogeneous space; $\vartheta \in M$, a point; G, a group which acts transitively on M (in particular, $G\vartheta = M$); and H, the stabilizer of ϑ in G. Then, the map $[g] \mapsto g\vartheta$ is a homogeneous space isomorphism between M and G/H. This thesis only deals with this kind of homogeneous spaces. The Lie algebras of G and H are denoted by $\mathfrak{g}$ and $\mathfrak{h}$ respectively.

Let $\pi\colon G \to M = G/H$ be the canonical projection. We denote $\vartheta := [e]$. It is clear that $d\pi_e\colon \mathfrak{g} \to T_\vartheta M$ is surjective.

Proposition 66.2.
The kernel of the differential of the projection is given by

$$\ker(d\pi_e) = \mathfrak{h}. \tag{66.7}$$

Proof. It is easy to see that $\mathfrak{h} \subset \ker(d\pi_e)$ but it turns out to be non trivial to prove the inverse. Our demonstration follows a part of the one of the proposition 4.3 of [813] (cf proposition 53.101).

Let X be in $\ker(d\pi_e)$. Since $d\pi_e X \in T_\vartheta M$, it can be applied on a function $f\colon M \to \mathbb{R}$. As $d\pi_e X = 0$ on any function and $X = \frac{d}{dt}\exp tX\big|_{t=0}$, we have

$$0 = (d\pi_e X)f = \frac{d}{dt}f(\pi\circ X)(t)\Big|_{t=0} = \frac{d}{dt}f([\exp tX])\Big|_{t=0} \tag{66.8}$$

But proposition 53.98 makes $\exp sX \in G$ acting on M. As a f, we can consider $g(q) = f(\exp sX \cdot q)$. Replacing f by g in (66.8), we get:

$$0 = (d\pi_e X)g = \frac{d}{dt}g([\exp tX])\Big|_{t=0} = \frac{d}{dt}f([\exp(s+t)X])\Big|_{t=0} \tag{66.9}$$

Then for any function f, the number $f([\exp tX]) \in \mathbb{R}$ doesn't depend on t, but for $t = 0$, $[\exp tX] = \vartheta$. Then $\forall t \in \mathbb{R}$, $\exp tX \in H$, and therefore, $X \in \mathfrak{h}$. □

Lemma 66.3.
We have

$$d\pi_g \circ dL_g = d\tau_g \circ d\pi_e. \tag{66.10}$$

Proof. Let $X \in \mathfrak{g}$ be the tangent vector to the curve $X(t)$ in G. We have

$$(d\pi_g \circ dL_g)(X) = \frac{d}{dt}\Big[\pi\big(gX(t)\big)\Big]_{t=0} = \frac{d}{dt}\Big[\tau_g\pi\big(X(t)\big)\Big]_{t=0} = (d\tau_g \circ d\pi_e)(X) \tag{66.11}$$

where we used the fact that, by definition, the action is given by $\tau_g\pi(g') = \pi(gg')$. □

1. Problems are possible with topology choices and differentiability of certain maps.

Definition 66.4.
*The homogeneous space $M = G/H$ is said to be **reductive** if there exists a subspace $\mathfrak{m}$ of $\mathfrak{g}$ such that*

— $\mathfrak{m} \oplus \mathfrak{h} = \mathfrak{g}$,

— $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}$.

Because of the second condition, such a $\mathfrak{m}$ is said to be H**-invariant**.

LemdpiisomMTM

Lemma 66.5.
If $\mathfrak{m}$ is reductive, the restriction $d\pi_e \colon \mathfrak{m} \to T_\vartheta M$ is an isomorphism between $\mathfrak{m}$ and $T_\vartheta M$.

Proof. As $d\pi_e \colon \mathfrak{g} \to T_\vartheta M$ is surjective, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ and $\mathfrak{h}$ is the kernel, $d\pi_e \colon \mathfrak{m} \to T_\vartheta M$ must be surjective. On the other hand, if we have $d\pi_e m = d\pi_e n$ for n, $m \in \mathfrak{m}$, $(m - n) \in Ker(d\pi_e) = \mathfrak{h}$ which is impossible because $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ is a direct sum. □

We can generalize this proposition by considering the space $\mathfrak{q}_g = dL_g\mathfrak{q}$.

PropDiffCosetBijTgGH

Proposition 66.6.
The differential $d\pi \colon \mathfrak{m}_g \to T_{[g]}M$ of the canonical projection provides an isomorphism between $\mathfrak{m}_g$ and $T_{[g]}M$.

Proof. In order to prove injectivity, take a $X \in \mathfrak{m}_g$ (i.e. $X = dL_g X'$ for a certain $X' \in \mathfrak{m}$) such that $d\pi X = 0$. If $X' = X'_h + X'_m$, we have

$$0 = \frac{d}{dt}\Big[\pi\big(ge^{tX'_h + tX'_m}\big)\Big]_{t=0} = \frac{d}{dt}\Big[\pi\big(ge^{tX'_m}e^{tX'_h}\big)\Big]_{t=0} = \frac{d}{dt}\Big[\pi\big(ge^{tX'_m}\big)\Big]_{t=0} = d\tau_g d\pi X'_m$$

where $d\pi X'_m \neq 0$ by definition of the quotient. Now, $d\tau_g \colon T_{[e]}M \to T_{[g]}M$ is a surjective linear map between two vector spaces of same dimension. Thus $d\tau_g$ is bijective and $d\pi X'_m = 0$, which proves that $X'_m = 0$ by lemma 66.5. □

The homogeneous space G/H is endowed with its **natural topology** which is defined by the requirement that the projection π is continuous and open. We refer to [813] for the properties of that topology.

66.2.1 Killing induced product

SubsecKillHomo

The product will be described with more details in point 66.2.4.2.
Since the Killing form B is an Ad_H-invariant product on $\mathfrak{q}$, we can define

$$B_g(X, Y) = B_e(dL_{g^{-1}}X, dL_{g^{-1}}Y) \tag{66.12}$$

which descent (see [958] for properties) to a homogeneous metric on $T_{[g]}M$:

$$B_{[g]}(d\pi X, d\pi Y) = B_g(\mathtt{proj}\, X, \mathtt{proj}\, Y) \tag{66.13}$$

EqDefMetrHomo

where $\mathtt{proj} \colon T_gG \to dL_g\mathfrak{q}$ is the canonical projection. An useful property of that projection is $\mathtt{proj}(dL_g X) = dL_g X_Q$ when $X = X_Q + X_H$. Using that property, we can write the product under the more manageable form

$$B_{[g]}(d\mu_g X, d\mu_g Y) = B_e(\mathtt{proj}\, X, \mathtt{proj}\, Y)$$

for all X, $Y \in \mathfrak{g}$ where we wrote $\mu_g = \tau_g \circ \pi$.

Although equation (66.6) looks like (66.5), we find a major difference here: the norm of $q_i^*[g]$ is not a constant. One should expect that it was a constant because (66.5) expresses a left translation while the Killing form is invariant under left translations. But the metric (66.13) is a composition

of the Killing form with a projection. Let us study this case in details in computing the product of two vectors of the form

$$X^*_{[g]} = d\pi \frac{d}{dt}\Big[e^{-tX}g\Big]_{t=0},$$

with $X \in \mathfrak{q}$:

$$\begin{aligned} B_{[g]}(X^*, Y^*) &= B_g\big(\texttt{proj}\,\frac{d}{dt}\Big[e^{-tX}g\Big]_{t=0}, \texttt{proj}\,\frac{d}{dt}\Big[e^{-tY}g\Big]_{t=0}\big) \\ &= B_g\big(dL_g\,\texttt{proj}\,\mathrm{Ad}(g^{-1})X, dL_g\,\texttt{proj}\,\mathrm{Ad}(g^{-1})Y\big) \\ &= B_e\Big(\big(\mathrm{Ad}(g^{-1})X\big)_{\mathfrak{q}}, \big(\mathrm{Ad}(g^{-1})Y\big)_{\mathfrak{q}}\Big) \\ &\neq B_e\Big(\mathrm{Ad}(g^{-1})X_{\mathfrak{q}}, \mathrm{Ad}(g^{-1})Y_{\mathfrak{q}}\Big) \\ &= B_e(X, Y) \end{aligned}$$

where the symbol $\neq$ has to be understood as "not equal in general" because equality holds of course for certain particular vectors such as zero.

66.2.2 Homogeneous space

Proposition 66.7.
Let $M = G/K$ be a homogeneous space where the Lie algebra $\mathfrak{g}$ has the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Then

(1) $T_{[\mathbb{1}]}M = \mathfrak{p}$

(2) $TM = G \times_{\mathrm{Ad}(K)} \mathfrak{p}$

Proof. The first part is already know. For the second, an element of $G \times_{\mathrm{Ad}(K)} \mathfrak{p}$ is of the form $[g, X]$ where $g \in G$, $X \in \mathfrak{p}$ and the equivalence relation $(g, X) \sim (gk, \mathrm{Ad}(k)X)$ for all $k \in K$.

Let us define $\psi\colon G \times_{\mathrm{Ad}(K)} \mathfrak{p} \to TM$ by

$$\psi[g, X] = [dL_g X]$$

where $[Y] = d\pi_g Y$ when $Y \in T_gG$. We are going to prove that ψ is injective and surjective. Suppose $\psi[g, X] = \psi[h, Y]$. Since $[dL_gX] = [dL_hY]$, we have $T_{[g]}M = T_{[h]}M$ and there exists a $k \in K$ such that $h = gk$. We have to prove that $\mathrm{Ad}(k)X = Y$. We have

$$d\pi_g(dL_gX) = d\pi_{gk}(dL_{gk}Y),$$

but

$$d\pi_{gk}(dL_{gk}Y) = \frac{d}{dt}\Big[\pi(gkY(t))\Big]_{t=0} = \frac{d}{dt}\Big[\pi(gkY(t)k^{-1})\Big]_{t=0} = d\pi_g dL_g\,\mathrm{Ad}(k)Y.$$

This proves injectivity of ψ. For surjectivity, take $X_{[g]} \in T_{[g]}M$: there exists a $\tilde{X}_g \in t_gG$ such that $X_{[g]} = d\pi_g \tilde{X}_g$. So, for a certain $X \in \mathfrak{g}$,

$$X_{[g]} = d\pi_g dL_g X = [dL_g X]$$

It remains to be proved that one can choose $X \in \mathfrak{p}$. Let us decompose $X = X_p + X_k$; it gives

$$[dL_g(X_p + X_k)] = [dL_gX_p] + [dL_gX_k],$$

but the latter is

$$\frac{d}{dt}\Big[\pi(gX_k(t))\Big]_{t=0} = \frac{d}{dt}\Big[\pi(g)\Big]_{t=0} = 0$$

because $X_k(t) \in K$ by definition. □

Let us consider G/H, a homogeneous space. If the Lie algebra $\mathfrak{h}$ is moreover the set of points fixed by an involution $\theta\colon \mathfrak{g} \to \mathfrak{g}$, the quotient G/H is said to be a **symmetric space**.

Let us point out that the Iwasawa decomposition naturally gives rise to a symmetric space: $AN = G/K$.

Remarque 66.8.
This is not our final definition of a symmetric space. A more precise definition will be given later, see section 66.3.

66.2.2.1 Frame bundle over reductive homogeneous spaces

PgFrameHomo

Let us consider an homogeneous space of the for $M = G/H$ with $G = \mathrm{SO}_0(p,q)$, and $\mathcal{G} = \mathfrak{h} \oplus \mathfrak{q}$. We have $T_\vartheta(G/H) = \mathfrak{q}$ and $T_{[g]}(G/H) = dL_g\mathfrak{q} = \mathfrak{q}_g$. Let $V = \mathbb{R}^{p,q}$ on which G acts by definition. Let $B(\vartheta)$ the set of orthonormal frames of $\mathfrak{q}$: the linear isometries $b\colon V \to \mathfrak{q}$. The we consider

$$B\big([G]\big) = \{[v \mapsto dL_g b(v)] \text{ tel que } b \in B(\vartheta)\}. \tag{66.14}$$

The frame bundle of G/H is

$$\begin{array}{ccc} \mathrm{SO}(p,q) & \rightsquigarrow & B\big([G]\big) \\ & & \big\downarrow \pi \\ & & G/H, \end{array} \tag{66.15}$$

where the action of $\mathrm{SO}(p,q)$ is given by

$$(b \cdot g)(v) = b(gv). \tag{66.16}$$

66.2.3 Invariant metric on homogeneous space (first)

Let G/H be a reductive homogeneous space, i.e. $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ with $[\mathfrak{h}, \mathfrak{m}] \subseteq \mathfrak{m}$. We denote by $T_A, T_B, \dots$ the generators of $\mathfrak{g}$ while $T_i, T_j, \dots$ particularise the generators of $\mathfrak{h}$ and $T_a, T_b, \dots$ the ones of $\mathfrak{m}$. The reducibility condition reads

$$[T_i, T_a] = C_{ia}^A T_A = C_{ia}^b T_b. \tag{66.17}$$

ProbAvecCorwell

¡¡ Avertissement/question à la lectrice !! 66.9
Vas voir dans Cornwell (que tu dois ajouter à la biblio) comment on fait pour montrer que C_{AB}^C est complètement antisymétrique dans tout les sens. Avec ça, je devrais pouvoir dire que C_{ij}^A sont nuls.

An element of G can be locally parametrized with $\dim G$ real numbers; for example

$$g(y^a, x^i) = e^{y^a T_a} e^{x^i T_i}$$

in a neighbourhood of identity. The classes $[g] \in G/H$ are given by only $\dim G - \dim H = \dim \mathfrak{m} = m$ real numbers and we can consider a choice of a representative of each class, i.e. a map $L\colon \mathbb{R}^m \to G$ with $L(y) \in [g]$ if $g = e^{y^a T_a} h$. For example, a possible choice is

$$L(y) = e^{y^a T_a}.$$

If we multiply at left $L(y)$ by $g \in G$, we obtain an element of another class whose representative is $L(y')$. Then

$$gL(y) = L(y')h \tag{66.18}$$

eq:gLyL

where $y' \in \mathbb{R}^m$ and $h \in H$ both depend on g and y (and the choice of the representative L). We consider the $\mathfrak{g}$-valued 1-form on $\mathbb{R}^m$ defined at $y \in \mathbb{R}^m$ by

$$V(y) = L(y)^{-1} dL_y. \tag{66.19}$$

If $v(t)$ is a path in $\mathbb{R}^m$ with $v(0) = y$, it defines a vector $v \in T_y\mathbb{R}^m$ and

$$V(y)v = \frac{d}{dt}\Big[L(y)^{-1} L(v(t))\Big]_{t=0} \in \mathfrak{g}.$$

So we can develop it with respect to a basis of $\mathfrak{g}$:

$$V(y) = T_a V^a(y) + T_i \Omega^i(y).$$

Now we are going to write $V(y')$ when y' is given by relation (66.18). First, it is clear that $L(y')^{-1} = hL(y)^{-1}g^{-1}$. Now if $y' = f(y)$, we have

$$d(l \circ f)_y = dL_{y'} \circ df_y,$$

then

$$V(y') = hL(y)^{-1}g^{-1}d(L \circ f)_y \circ (df^{-1})_{y'}. \tag{66.20}$$

We will forget the $(df^{-1})_{y'}$, keeping in mind that if $V(y')$ is applied to a vector of $T_y\mathbb{R}^m$, we have to transport the vector with f. Now let us explicit the expression (66.20). For this, remark that $L \circ f = gL(\ \cdot\ )h^{-1}$ where h is a map from $\mathbb{R}^m$ to H. Then we have tu use the Leibniz formula; let $v \in T_y\mathbb{R}^m$,

$$d(L \circ f)_y v = \frac{d}{dt}\Big[gL(v(t))h^{-1}(v(t))\Big]_{t=0} = g(dLv)h^{-1}(y) + gL(y)(dh^{-1}v). \tag{66.21}$$

So,

$$\begin{aligned} V(y') &= hL(y)^{-1}g^{-1}\big(g(dL_y)h^{-1} + gL(y)dh_y^{-1}\big) \\ &= hV(y)h^{-1} + hdh_y^{-1}. \end{aligned} \tag{66.22}$$

Then the transformation rule of V under an action of G is given by

$$V(y') = hV(y)h^{-1} + hdh_y^{-1}. \tag{66.23}$$

In particular this induces a transformation rule for V^a by

$$V^a(y') = \big(hV(y)h^{-1}\big)^a = \big(\mathrm{Ad}(h)V(y)\big)^a = V^A(y)D(h^{-1})_A{}^a \tag{66.24}$$

where D is defined by $\mathrm{Ad}(g^{-1})T_A = D(g)_A{}^B T_B$.

66.2.3.1 Infinitesimal expressions

Now we want to write the equation (66.18) in the case where g is close to the identity. We start by considering g under the form $g = e^{\epsilon^A T_A}$ with small ϵ. If we write h under the form $h = e^{R^i T_i}$, R^i is a function of y and ϵ. If we suppose that ϵ is very small (our intention is to make a derivation with respect to epsilon at $\epsilon = 0$), we can suppose that R is linear with respect to epsilon. Then $h = e^{\epsilon^A W_A^i(y)T_i}$. For the same reason, y' is linear with respect to epsilon: $y'^\alpha = y^\alpha + \epsilon^A K_A^\alpha(y)$. So we write

$$e^{\epsilon^A T_A}L(y) = L(y^\alpha + \epsilon^A K_A^\alpha(y))e^{\epsilon^A W_A^i(y)T_i}$$

which an equality in G. Let us derive it with respect to ϵ^A at $\epsilon = 0$. Note that by $L(y^\alpha)$, we mean $L(y^\alpha e_\alpha)$ where $\{e_\alpha\}$ is the canonical basis of $\mathbb{R}^m$. Then we find

$$T_A L(y) = dL_y(K_A^\alpha(y)\partial_\alpha) + L(y)W_A^i(y)T_i.$$

If one multiply it at left by $L(y)^{-1}$,

$$D(L(y))_A{}^B T_B = V(y)\big(K_A^\alpha(y)\partial_\alpha\big) + W_A^i(y)T_i. \tag{66.25}$$

Remark that $K_A^\alpha(y)$ is just a real number, then it can get out the form $V(y)$. From notational convenience, we write $V(y)\partial_\alpha = V_\alpha(y)$. We write separately the $\mathfrak{h}$ and $\mathfrak{m}$ components in equation (66.25):

$$D(L(y))_A{}^i T_i = K_A^\alpha(y)(\Omega^i(y)T_i)(\partial_\alpha) + W_A^i(y)T_i \tag{66.26a}$$

$$D(L(y))_A{}^b T_b = K_A^\alpha(y)(V^b(y)T_b)(\partial_\alpha). \tag{66.26b}$$

Be careful on one fact: the expression $V^b(y)T_b(\partial_\alpha)$ means $V^b(y)(\partial_\alpha)T_b$ where which is the product of the vector $T_b \in \mathfrak{g}$ by the real $V^b(y)T_b(\partial_\alpha)$. So we can "simplify" the T_A's in equations (66.26) to find

$$W^i_A(y) = D(l(y)_A{}^i - K^\alpha_A(y)\Omega^i_\alpha(y) \tag{66.27a}$$
$$D(l(y)_A{}^b = K^\alpha_A(y)V^b_\alpha(y) \tag{66.27b}$$

whose are equalities in $\mathbb{R}$.

Let us find a form for $D(h^{-1})_A{}^b$ when h is given by equation (66.18) with $y'^\alpha = y^\alpha + \epsilon^A K^\alpha_A(y)$ and for small ϵ. The matrix $D(g)$ is given by

$$D(g)_A{}^B T_B = g^{-1}T_A g = \frac{d}{dt}\Big[\,\mathbf{Ad}_{g^{-1}}\, e^{tT_A}\Big]_{t=0} = \mathrm{Ad}(g^{-1})T_A.$$

So in our case,

$$D(e^{\epsilon^B W^i_B(y)T_i})_A{}^C T_C = \mathrm{Ad}\left(\exp(\epsilon^B W^i_B(y)T_i)\right)T_A. \tag{66.28}$$

If we derive it with respect to ϵ^B at $\epsilon = 0$, we find

$$\begin{aligned}\frac{d}{d\epsilon^B}\Big[D(\exp(\epsilon^B W^i_B(y)T_i))_A{}^C T_C\Big]_{\epsilon=0} &= \mathrm{ad}(W^i_B(y)T_i)T_A \\ &= W^i_B(y)C_{iA}{}^D T_D,\end{aligned} \tag{66.29}$$

so that we can power expand $D(h^{-1})_A{}^a$ with respect to ϵ around $\epsilon = 0$:

$$D(h^{-1})_A{}^a = \delta^a_A + \epsilon^B W^i_B(y)C_{iA}{}^a + \dots \tag{66.30}$$ eq:Dinfin

Then equation (66.24) reads $V^a(y') - V^a(y) = V^a(y)\epsilon^B W^i_B(y)C_{iA}{}^a$, but remarking that the reducibility makes $C^a_{ij} = 0$,

$$V^a(y+\delta y) - V^a(y) = \epsilon^B W^i_B(y)C_{ib}{}^a V^b(y). \tag{66.31}$$

The fact that $C_{ib}{}^a$ can be made skews-symmetric shows that this equation describe the infinitesimal action of G on $V(y)$ by the action of $so(n)$. It allows us to state the following theorem.

66.2.3.2 Invariant metric

Theorem 66.10.
The metric

$$g_{\alpha\beta} = g(\partial_\alpha, \partial_\beta) = \delta_{ab}V^a_\alpha V^b_\beta \tag{66.32}$$ eq:metric_GH

is invariant with respect to the left action of G.

¡¡ Avertissement/question à la lectrice !! 66.11
Regardes s'il faut semi-simple pour obtenir l'antisymétrie des constantes de structure.

An other way to write this metric is

$$g = \delta_{ab}(V^a \otimes V^b).$$

Proof. We have to show that $g_{y'}(\partial_\alpha, \partial_\beta) = g_y(\partial_\alpha, \partial_\beta)$. For this we will show that the derivative of $g_{\alpha\beta}(y)$ with respect to y is zero. So we write $y'^\alpha + \epsilon^A K^\alpha_A(y)$ and

$$g_{\alpha\beta}(y') = \delta_{ab}V^a_\alpha(y)V^b_\beta(y)D(h^{-1})_A{}^a D(h^{-1})_B{}^b.$$

The computation is performed using (66.30) which gives $D(h^{-1})_A{}^a\big|_{\epsilon=0} = \delta^a_A$ and

$$\begin{aligned}\frac{d}{d\epsilon^C}\Big[g_{\alpha\beta}(y')\Big]_{\epsilon=0} &= \delta_{ab}V^A_\alpha(y)V^B_\beta\left\{\frac{d}{d\epsilon^C}\Big[D(h^{-1})_A{}^a\Big]_{\epsilon=0} D(h^{-1})_B{}^b\Big|_{\epsilon=0}\right. \\ &\qquad \left. D(h^{-1})_A{}^a\big|_{\epsilon=0}\frac{d}{d\epsilon^C}\Big[D(h^{-1})_B{}^b\Big]_{\epsilon=0}\right\} \\ &= \delta_{ab}V^a_\alpha(y)V^B_\beta(y)\Big[\delta^b_B W^i_C(y)C_{iA}{}^a + \delta^a_A W^i_C(y)C_{iB}{}^b\Big] \\ &= \sum_a V^A_\alpha(y)V^a_\beta(y)W^i_C(y)C_{iA}{}^a + \sum_b V^B_\alpha(y)V^B_\beta(y)W^i_C(y)C_{iB}{}^b\end{aligned} \tag{66.33}$$

Taking into account the fact that $C_{ij}{}^{a} = 0$, one can reduce some summations like $V_\alpha^A(y)C_{iA}{}^{a} = V_\alpha^b(y)C_{ib}{}^{a}$. Using the antisymmetry of $C_{ib}{}^{a}$ with respect to a, b, we find that the sum is zero. □

There is an other invariant metric:

$$g_{\alpha\beta} = B_{ab}V_\alpha^a V_\beta^b \tag{66.34}$$

where B is the matrix of the Killing form. Following the same proof of the invariance than the previous one, one finds

$$\frac{d}{d\epsilon^C}\Big[g_{\alpha\beta}(y')\Big]_{\epsilon=0} = V_\alpha^a(y)V_\beta^b(y)W_C^i(B_{cb}C_{ia}{}^{c} + B_{ac}C_{ib}{}^{c}).$$

This is zero because of the formula $B((\operatorname{ad} X)Y, Z) = -B(Y, (\operatorname{ad} X)Z)$.

¡¡ Avertissement/question à la lectrice !! 66.12
Apparemment cette métrique est invariante indépendamment d'hypothèse de semi-simplicité.

66.2.3.3 The choice of $L(y)$

Let us see what happens if we had done the work with $L'(y) = L(y)h(y)$ instead of $L(y)$. In this case,

$$V'(y) = L'(y)^{-1}dL'_y = h(y)^{-1}L(y)^{-1}dL'_y,$$

but using Leibniz formula, we find

$$dL'_y v = \frac{d}{dt}\Big[L(v(t))h(v(t))\Big]_{t=0} = (dL_y v)h(y) + L(y)dh_y v, \tag{66.35}$$

so that $V'(y) = V(y)$ up to a renaming $h \leftrightarrow h^{-1}$. The conclusion is that $\delta_{ab}V_\alpha^a V_\beta^b$ and $B_{ab}V_\alpha^a V_\beta^b$ are independent of the choice of L.

66.2.4 Homogeneous metric on homogeneous spaces

66.2.4.1 One way to obtain it

Let $M = G/H$ be a homogeneous space. A Riemannian metric $\langle\ \cdot\ ,\ \cdot\ \rangle$ on M is **homogeneous** when

$$\langle dL_g v, dL_g w\rangle_{g[x]} = \langle v, w\rangle_{[x]} \tag{66.36}$$ eq:def:homo_metric

for all $g \in G$, $[x] \in M$. Note that this formula cannot define an inner product on each $T_{[x]}M$ from the data of an inner product on $T_\vartheta M$ because –unless certain conditions– it is not well-defined.

From the definition of the homogeneous structure of G/H, all element of H fixes $\vartheta = [e]$ (by the left action).Then dL_h is an automorphism of $T_\vartheta M$ and we can define the **isotropic representation** $\rho\colon H \to \operatorname{Aut}(T_\vartheta)M$ by

$$\rho(h)X = dL_h X$$

with $X \in T_\vartheta M$.

Now let $\langle\ \cdot\ ,\ \cdot\ \rangle$ be an inner product on $T_\vartheta M$ (for example the Killing form on the $\mathfrak{m}$ part of $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$ in the reductive case). We can try to export this product at $[g]$ by the formula

$$\langle v, w\rangle_{[g]} = \langle dL_{g^{-1}}v, dL_{g^{-1}}w\rangle_\vartheta. \tag{66.37}$$ eq:scal_gdee

Proposition 66.13.
The product $\langle\ \cdot\ ,\ \cdot\ \rangle_{[g]}$ defined by formula (66.37) is well defined if and only if $\langle\ \cdot\ ,\ \cdot\ \rangle_\vartheta$ is invariant under the isotropic representation.

Proof. Let us proof the necessary condition; the sufficient one is just the same written backward. The assumption makes

$$\begin{aligned}\langle v, w\rangle_{[gh]} &= \langle dL_{h^{-1}g^{-1}}v, , dL_{h^{-1}g^{-1}}w\rangle_\vartheta \\ &\stackrel{!}{=} \langle dL_{g^{-1}}v, dL_{g^{-1}}w\rangle_\vartheta\end{aligned} \tag{66.38}$$

for every v, $w \in T_{[g]}M$, $g \in G$, $h \in H$. In particular,

$$\langle dL_h X, dL_h Y\rangle_\vartheta = \langle X, Y\rangle_\vartheta$$

for all X, $Y \in T_\vartheta M$. □

Two remarkable properties of this inner product are the fact that it is a homogeneous Riemannian structure and that *all* the homogeneous metric are such. In order to see the first claim, just remark that if $[x] \in M$,

$$\begin{aligned}\langle dL_g v, dL_g w\rangle_{g[x]} &= \langle dL_{x^{-1}g^{-1}}dL_g v, dL_{x^{-1}g^{-1}}dL_g w\rangle_\vartheta \\ &= \langle dL_{x^{-1}}v, dL_{x^{-1}}w\rangle_\vartheta \\ &= \langle v, w\rangle_\vartheta.\end{aligned} \tag{66.39}$$

The second claim comes from the choice $g = x^{-1}$ in the definition (66.36).

66.2.4.2 One other way to obtain it

SubSubSecTheKillingHomo

Let us consider a metric on $\mathfrak{g}$ and see in which case it can be extended to gives rise to a well defined homogeneous metric on the quotient $M = G/H$. Let $\mathfrak{g} = T_eG$, $\mathfrak{h} = T_eH$ and $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{h}$. Using dL, we can propagate the space $\mathfrak{m}$ to the point $g \in G$ by defining

$$\mathfrak{m}_g = dL_g\mathfrak{m}.$$

We saw in proposition 66.6 that $\mathfrak{m}_g$ was isomorphic to $T_{[g]}M$.

Let $\langle ., . \rangle$ be a product on $\mathfrak{g}$ which is Ad_H-invariant on $\mathfrak{m}$. We claim that the following construction gives a well defined and homogeneous product on $\mathfrak{g}$. First, the product on $\mathfrak{g}$ extends to a product on T_gG for every g by

$$\langle X, Y\rangle_g = \langle dL_{g^{-1}}X, dL_{g^{-1}}Y\rangle;$$

this induces the following inner product on $T_{[g]}(G/H)$ that will reveal to be well defined under the current assumptions:

$$\langle d\pi_g X, d\pi_g Y\rangle_{[g]} = \langle X, Y\rangle_g \tag{66.40}$$

eq:scal_TgM

where $X, Y \in \mathfrak{m}_g = dL_g\mathfrak{m}$. Indeed, the map $d\pi\colon \mathfrak{m}_g \to T_{[g]}M$ is an isomorphism, hence for all $v \in T_{[g]}M$, there exists one and only one $X \in \mathfrak{m}_g$ such that $d\pi X = v$. Since $d\pi\colon \mathfrak{m}_{gh} \to T_{[g]}M$ is also an isomorphism, the condition for (66.40) to be a good definition, we must have

$$\langle X, Y\rangle_g = \langle dL_{g^{-1}}X, dL_{g^{-1}}Y\rangle_e = \langle X', Y'\rangle_{gh} \tag{66.41}$$

where $X' = dR_h X$. It is easy to remark that this condition is the Ad_H-invariance of the inner product $\langle \cdot\ ,\ \cdot \rangle_e$ defined on $\mathfrak{g}$.

The reader should remark that all the conditions are satisfied by the Killing inner product.

Now, if X is any element of $\mathfrak{g}$, we define successively

$$\begin{aligned}\langle d\pi_g dL_g X, d\pi_g dL_g Y\rangle_{[g]} &= \langle d\pi_g dL_g X_\mathfrak{m}, d\pi_g dL_g Y_\mathfrak{m}\rangle_{[g]} \\ &= \langle dL_g X_\mathfrak{m}, dL_g Y_\mathfrak{m}\rangle_g \\ &= \langle \mathtt{proj}_\mathfrak{m} X, \mathtt{proj}_\mathfrak{m} Y\rangle_e.\end{aligned} \tag{66.42}$$

EqDefProdGsurH

The last line is the usual Killing form on $\mathfrak{g}$, or any other inner product which has the right properties.

Let us prove that the first line is well defined. First, notice that $d\pi_g\colon \mathfrak{m}_g \to T_{[g]}M$ is an isomorphism, thus there exists one and only one $\tilde{X} \in dL_g\mathfrak{m}$ such that $d\pi_g\tilde{X} = d\pi_g dL_g X$. Since

$$d\pi_g\big(dL_g X_{\mathfrak{m}}\big) = \frac{d}{dt}\Big[\pi\big(ge^{tX_{\mathfrak{m}}}\big)\Big]_{t=0} = 0, \tag{66.43}$$

we know that

$$d\pi_g dL_g X_{\mathfrak{q}} = d\pi_g dL_g X \tag{66.44}$$

for every $X \in \mathfrak{g}$.

66.3 Symmetric spaces

sec:symm

This section is mainly taken from [956, 959, 960, 946].

66.3.1 Basic facts

Definition 66.14.
*A **symmetric space** is a manifold M and an analytic "multiplication" $\mu\colon M \times M \to M$ –written $s_x(y)$ as $\mu(x,y)$– such that*

(1) $\forall x \in M$, s_x is an involutive diffeomorphism of M called "the symmetry at x ",

(2) $\forall x \in M$, x is an isolated fixed point of s_x,

(3) $\forall x, y \in M$, $s_x \circ s_y \circ s_x = s_{s_x(y)}$.

def:esp_sym

Definition 66.15.
*An **homomorphism** of symmetric space (M,s) and (M',S) is an analytic map $\varphi\colon M \to M'$ which satisfies*

$$\varphi(s_x(y)) = S_{\varphi(x)}\varphi(y).$$

Immediately, for any z in M, the symmetry s_z is an automorphism of M (as symmetric space). Indeed,

$$s_{s_z(x)}(s_z(y)) = s_z \circ s_x \circ s_z \circ s_z(y) = (s_z \circ s_x)(y) \tag{66.45}$$

The group generated by all the $s_x \circ s_y$ (x, $y \in M$) is the **displacement group** and is denoted by $G(M)$.

Lemma 66.16.
The displacement group is a normal subgroup of $\mathrm{Aut}(M)$.

Proof. If φ is an automorphism of M, we have

$$\varphi \circ s_x \circ s_y \circ \varphi^{-1} = s_{\varphi(x)} \circ \varphi \circ s_y \circ \varphi^{-1} = s_{\varphi(x)} \circ s_{\varphi(y)} \tag{66.46}$$

because $\varphi \circ s_x = s_{\varphi(x)} \circ \varphi$. □

We define $Q\colon M \to G(M)$ by $Q(x) = s_x s_e$. This is the **quadratic representation** of M. Since $Q(x)Q(y)^{-1} = s_x s_y$, $Q(M)$ generate $G(M)$.

ThSymSymGd

Theorem 66.17.
The space M is symmetric for the structure

$$s_{[x]}[y] = [x\sigma(x)^{-1}\sigma(y)], \tag{66.47}$$

eq:sym_M

while L_σ is a symmetric space for

$$s_x(y) = xy^{-1}x. \tag{66.48}$$

The map $q\colon M \to L$, $q([x]) = x\sigma(x^{-1})$ is a homomorphism from M to L_σ and L/L^σ is isomorphic to L_σ by q.

Moreover $\tau\colon L \to \mathrm{Aut}(M)$ is a homomorphism and the displacement group $G(M)$ is the subgroup of $\tau(L)$ generated by $\tau(L_\sigma)$.

Proof. Symmetric structure on M. First we prove that M is symmetric. The symmetry (66.47) is well defined: if $k, k' \in K$,

$$s_{[xk]}[yk'] = [xk\sigma(xk)^{-1}\sigma(yk')] = [x\sigma(x)^{-1}\sigma(y)]. \tag{66.49}$$

It is clear that $s_{[x]} \circ s_{[x]} = id$ because

$$\begin{aligned} s_{[x]} \circ s_{[x]}[y] &= s_{[x]}([x\sigma(x)^{-1}\sigma(y)]) = [x\sigma(x)^{-1}\sigma(x)\sigma(\sigma(x)^{-1})(\sigma\circ\sigma)(y)] \\ &= [xx^{-1}y] = [y]. \end{aligned} \tag{66.50}$$

Since L acts transitively by automorphism on M (this is: $\tau(x)$ is an automorphism of M and we can always find a $x \in L$ such that $\tau(x)[y] = [z]$ for given y, $z \in M$), we just have to prove the property of isolated fixed point for $[e] \in M = L/K$. So we consider s_ϑ ($\vartheta = [e]$) on a neighbourhood of ϑ in M.

Identification $T_\vartheta M = \mathfrak{L}_-$.

Now we show how to identify (as vector spaces) $T_\vartheta M$ with

$$\mathfrak{L}_- = \{X \in \mathfrak{L} \text{ tel que } \sigma(X) = -X\}.$$

where $\mathfrak{L}$ is the Lie algebra of L. For this, we will show that $\psi\colon \mathfrak{L}_- \to T_\vartheta M$,

$$\psi(X) = \frac{d}{dt}\Big[[X(t)]\Big]_{t=0} \tag{66.51}$$

if $X(t)$ is a path in L whose derivative is X. Any vector in $T_\vartheta M$ comes from a path $[Y(t)]$ where $Y(t) \in L$ can be written as $Y(t) = c(t)k(t)$ where $k(t) \in K$ has no continuity property, and c is the "main" part of the path. Then

$$\psi(c'(0)) = \frac{d}{dt}\Big[[Y(t)]\Big]_{t=0}$$

and ψ is surjective. In order to see the injectivity, remark that in a neighbourhood of e,

$$\sigma(e^{tX}) = e^{td\sigma X} = e^{-tX}$$

because $X \in \mathfrak{L}_-$. With other words, if $X \in \mathfrak{L}_-$,

$$\sigma(X(t)) = X(t)^{-1}$$

when t is small. But e is an isolated fixed point of the inversion. Then $\psi(X) = 0$ let only one possibility: $[X(t)] = cst$. Thus (for small t) $X(t)$ can be written as $X(t) = gk(t)$ with $\sigma(k(t)) = k(t) \in L$. Since $X(0) = e$, $k(0) = g^{-1}$ and $\sigma(g) = k(0)^{-1} = g$. Then

$$\sigma(X(t)) = X(t).$$

But on the other hand, $X \in \mathfrak{L}_-$ implies $\sigma(X(t)) = X(t)^{-1}$ and finally $X(t) = X(t)^{-1}$, so that $X(t) = e$. See eventually the error 88.21.pg:X_t

Now we can see that 0 is an isolated fixed point of s_ϑ. We looks at $(ds_\vartheta)_\vartheta X$ with $X \in \mathfrak{L}_- \equiv T_\vartheta M$.

$$\begin{aligned} (ds_\vartheta)X = (ds_\vartheta \circ \psi)(X) &= ds_\vartheta \frac{d}{dt}\Big[[X(t)]\Big]_{t=0} = \frac{d}{dt}\Big[[\sigma(X(t))]\Big]_{t=0} \\ &= d\sigma\frac{d}{dt}\Big[[X(t)]\Big]_{t=0} = \sigma(X) = -X. \end{aligned} \tag{66.52}$$

With the notation $X^* = \sigma(x)^{-1}$,

$$\begin{aligned} q\left(s_{[x]}[y]\right) &= q([xx^*\sigma(y)]) \\ &= xx^*\sigma(y)\sigma^2(y)^{-1}\sigma(x^*)^{-1}\sigma(x)^{-1} \\ &= (xx^*)(\sigma(y)y^{-1})(xx^*) \\ &= q(x)q(y)^{-1}q(y), \end{aligned} \tag{66.53}$$

then

$$q(s_{[x]}[y]) = s_{q(x)}q(y)$$

and q is a homomorphism between M and L for they respective symmetric spaces structure. It is contained in the definition of q that

$$q(M) = L_\sigma = \{x\sigma(x)^{-1} \text{ tel que } x \in L\}$$

On the other hand, $q([x]) = q([y])$ if and only if $xx^* = yy^*$ which is equivalent to $x^{-1}y \in L^\sigma$. But $x^{-1}y \in L^\sigma$ implies $\overline{x} = \overline{y}$ where the bar stands for the classes with respect to L_σ. In definitive, $q([x]) = q([y])$ if and only if $\overline{x} = \overline{y}$. Hence, q is an isomorphism of symmetric spaces between L_σ and L/L^σ.

We recall the definition $\tau(x)\colon M \to M$, $\tau(x)[y] = [xy]$, and we use the quadratic representation of M:

$$Q([x])[y] = s_{[x]}s_\vartheta[y] = s_{[x]}([\sigma(y)]) = \tau(xx^*)[y]. \tag{66.54}$$

Then $G(M)$ is generated by $Q(M) = \tau(q(M)) = \tau(L_\sigma)$. □

66.3.2 Choice of a Cartan involution

Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the Cartan decomposition of $\mathfrak{g}$ and B, the Killing form on $\mathfrak{g}$. We know that the linear transformation of $\mathfrak{g}$ defined by

$$\theta(X) = \begin{cases} X & \text{if } X \in \mathfrak{k} \\ -X & \text{if } X \in \mathfrak{p} \end{cases}$$

is an involutive automorphism of $\mathfrak{g}$; and the bilinear form

$$(X, Y) \to \langle X, Y\rangle := -B(X, \theta Y)$$

is positive definite on $\mathfrak{g}$.

Theorem 66.18.
Let $\mathfrak{g}$ be a real semisimple Lie algebra, σ an involutive automorphism and θ a Cartan involution. Then

(1) there exists a Cartan involution θ_1 such that $[\sigma, \theta_1] = 0$,

(2) if θ_1 and θ_2 are two such involutions then they are conjugated by an automorphism of $\mathfrak{g}$ of the form $e^{\operatorname{ad} X}$ with $\sigma(X) = X$.

tho:sigma_theta

The first point is contained in theorem 53.125 and the second one is exactly the corollary 53.124

66.3.3 Affine Symmetric spaces

The matter may be found in chapter XI of [957]

Let M be a n-dimensional manifold endowed with a connection ∇. The **symmetry** at $x \in M$, denoted by s_x, is defined on a normal neighbourhood of x by $\exp_x X \to \exp_x(-X)$. Properties of the exponential and normal neighbourhood make it a well defined diffeomorphism because it doesn't depends on the choice of the normal neighbourhood.

It clearly fulfils $s_x^2 = id$ and x is an isolated fixed point of s_x.

¡¡ Avertissement/question au lecteur !! 66.19
je ne vois pas comment démontrer que $s_{s_x y} = s_x \circ s_y \circ s_x$.

If we consider the normal coordinates in a neighbourhood around x, it is clear that

$$s_x(x_1, \dots, x_n) = (-x_1, \dots, -x_n).$$

Then $(ds_x)_x = -I_x$ where $I_x\colon T_xM \to T_xM$ is the identity.

If for all $x \in M$, the map s_x is an affine transformation, we say that M is a locally affine symmetric space [2].

Lemma 66.20.
On an affine locally symmetric space, an odd tensor invariant under s_x is zero at x.

Proof. From $(ds_x)_x = -I_x$, the transformation s_x transforms a tensor K of degree p into $(-1)^p K$. □

Theorem 66.21.
Let M and M' be two manifolds with $\nabla T = \nabla R = \nabla T' = \nabla R' = 0$ and a linear endomorphism $F\colon T_{x_0}M \to T_{y_0}M$ such that $FT_{x_0} = T'_{y_0}$ and $FR_{x_0} = R'_{y_0}$.

Then there locally exists an isometry [3] $f\colon M \to M'$ such that $f(x_0) = y_0$ and $(df)_{x_0} = F$.

Proposition 66.22.
A manifold M with an affine connection is affine locally symmetric if and only if

$$T = 0 \text{ and } \nabla R = 0.$$

Proof. Since s_x is affine, it preserves T and ∇R whose are tensor of degree 3 and 5. From lemma, they are zero.

For the converse, $-I_x$ preserves R_x because R is a tensor of degree 4. In this context, the theorem gives a $f\colon M \to M$ such that $f(x) = x$ and $df_x = -I_x$. But f is also an affine transformation, then

$$f(\exp X) = \exp(df X) = \exp(-I_x X) = \exp(-X). \tag{66.55}$$

This gives $f = s_x$. □

Two results without proof.

Theorem 66.23 ([838])**.**
If M is a differentiable manifold with a linear connection such that $\nabla T = 0$ and $\nabla R = 0$, then the atlas of normal coordinates gives to M a structure of analytic manifold and the connection is analytic.

Proposition 66.24 ([838])**.**
Let M be a connected, simply connected and complete manifold with a linear connection such that $\nabla T = 0 = \nabla R$. Let $F\colon T_x M \to T_y M$ a linear isomorphism such that $T_x \to T_y$ and $R_x \to R_y$.

Then there exist one and only one affine transformation f of M such that $f(x) = y$ and $df_x = F$.

In particular the group $\mathcal{A}(M)$ of the affine transformations of M is transitive on M.

A manifold M with an affine connection is an **affine symmetric space** if for all $x \in M$, the symmetry s_x can be globally extended to an affine transformation of M. Thanks to the latter proposition, an affine locally symmetric complete and simply connected space is affine.

Proposition 66.25.
An affine symmetric space is complete.

Proof. Let γ be a geodesic from x to y, i.e. $\gamma(0) = x$ and $\gamma(1) = y$. Let us pose $\gamma(1 + t) = s_y(\gamma(1 - t-)$ for $0 \leqslant t \leqslant a$. It extends γ beyond y. Let us prove that the extension still is a geodesic. For a certain $Y \in T_x M$, we have $y = \exp_x Y$, so for t between 0 and 1, $\gamma(t) = \exp_x(tY)$.

2. A differentiable map $f\colon M \to M'$ is an **affine** if $df\colon TM \to TM'$ transforms all horizontal curves to an horizontal curve. An affine transformation automatically fulfils

$$f(\exp X) = \exp(df X)$$

for all $X \in T_x M$.

3. relis pour voir si c'est bien ça.

Let Y_t be the parallel vector field along γ with $Y_0 = Y$; for example, $x = \exp_y(-Y_1)$ and $\exp_x(tY) = \exp_y(t-1)Y_1$.

$$\gamma(1+t) = s_y(\exp_y(-tY_1)) = \exp_y(tY_1). \tag{66.56}$$

□

Proposition 66.26.
The group of affine transformations of an affine symmetric space is transitive.

Proof. Let x and y be two points in M. There exists a sequence of convex normal neighbourhood $\mathcal{U}_1, \dots, \mathcal{U}_k$ such that $x \in \mathcal{U}_1$, $y \in \mathcal{U}_k$, $\mathcal{U}_i \cap \mathcal{U}_{i+1} \neq \varnothing$. So one can reach y from x with geodesic segments. This construction is just the fact that, if $c\colon [a,b] \to M$, is a path from x to y, the set $c([a,b])$ is compact in M. On each point of $c([a,b])$ we consider a convex neighbourhood which gives an open covering of a compact set.

It remains to be proved that if x and y are reachable by a geodesic curve, then they can be reached by an affine transformation. If $y = \exp_x Y$ and $z = \exp_x(\frac{1}{2}Y)$, then $s_z(x) = y$.

□

One can prove that the group $\mathcal{A}(M)$ is a Lie group. We denote by G its identity component. It is clear that is a group acts transitively on a manifold, then its identity component also acts transitively. Then G acts transitively on M and one has a homogeneous space structure which allows us to write $M = G/H$.

More precisely, we have the

ThoGplugdSymssgpAff

Theorem 66.27.
Let G be the largest connected group of affine transformation of an affine symmetric space M and H, the isotropy group of a fixed point $o \in M$, so that $M = G/H$.

Let s_o the symmetry of M at o and σ the automorphism of G defined by

$$\sigma(g) = s_o \circ g \circ s_o^{-1}.$$

Let G_σ the closed subgroup of G which fixes σ. Then $G_\sigma^o \subset H \subset G_\sigma$.

Proof. Let $h \in H$ and $\sigma(h) = s_o \circ h \circ s_o^{-1}$. We know that $(ds_o)_o = -I_o$, then $(d\sigma(h))_o = dh_o$. But general theory about affine transformations says that if two affine transformations has same differential at one point then they are equals. In our case, it gives $\sigma(h) = h$; therefore $H \subset G_\sigma$.

Let now g_t be a one parameter subgroup of G_σ. From the definition of σ, $s_o \circ g_t = \sigma(g_t) \circ s_o$, then $s_o \circ g_t(o) = g_t \circ s_o(o) = g_t(o)$. Then the orbit $g_t(o)$ is fixed by s_o. But o is an isolated fixed point of s_o, then $g_t(o) = o$ for all t and $g_y \in H$.

From general theory of Lie groups, a connected Lie group is generated by its one parameter subgroups. Then G_σ^o is generated by elements which fix σ. So $G_\sigma^o \subset H$. □

66.3.4 Symmetric pair

Let G be a connected Lie group and H, a closed subgroup.

Definition 66.28.
*We say that (G,H) is a **symmetric pair** if there exists an analytic involutive automorphism $\sigma\colon \mathfrak{g} \to \mathfrak{g}$ such that $(H_\sigma) \subset H \subset H_\sigma$ where H_σ is the set of fixed points by σ. If the group $\mathrm{Ad}_G(H)$ is compact, the pair is **Riemannian**.*

Note: by $\mathrm{Ad}_G(H)$ we mean the Lie subgroup of $\mathrm{Ad}_G(G)$ which is the image of H by Ad_G.

Proposition 66.29.
Let (G,K) be a Riemannian symmetric pair and $\mathfrak{k}$ the Lie algebra of K. We denote by $\mathcal{Z}$ the center of the Lie algebra $\mathfrak{g}$. If $\mathfrak{k} \cap \mathcal{Z} = \{0\}$, then there exists one and only one analytic involutive automorphism σ of G such that $(K_\sigma)_0 \subset K \subset K_\sigma$.

Proof. The point is the unicity: the existence is contained in the definition of a symmetric pair. Let us consider two such automorphism σ_1 and σ_2. As far as the Lie algebras are concerned, the identity component only is relevant. Since $(K_{\sigma_1})_0 = (K_{\sigma_1})_0$; thus $\mathfrak{k}_1 = \mathfrak{k}_2$. We consider the respective decompositions of $\mathfrak{g}$ for σ_1 and σ_2:

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}_1 \tag{66.57a}$$
$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}_2. \tag{66.57b}$$

where $\mathfrak{p}_i$ is the eigenspace with eigenvalue -1 for the automorphism $d\sigma_i$ of $\mathfrak{g}$. Since the Killing form B of $\mathfrak{g}$ is invariant under σ_i, $\mathfrak{k}$ is B-orthogonal to $\mathfrak{p}_i$. Indeed $B(k,p) = B(d\sigma_i k, d\sigma_i p) = -B(k,p)$; then $B(k,p) = 0$. Consider $X_1 \in \mathfrak{p}_1$ and $T \in \mathfrak{k}$. We have a $X_2 \in \mathfrak{p}_2$ such that $X_1 = T + X_2$. Since $\mathfrak{p}_i \perp \mathfrak{k}$,

$$0 = B(k, X_1) = B(k,T) + B(k, X_2),$$

then $B(k,T) = 0$ and $T \perp \mathfrak{k}$. In particular, $B(T,T) = 0$. From proposition 53.93, B is strictly negative definite on $\mathfrak{k}$; then $T = 0$ so that $\mathfrak{p}_1 = \mathfrak{p}_2$ and $\sigma_1 = \sigma_2$.

Now we have to see that the $\mathfrak{k}$ here is actually the $\mathfrak{k}$ of the proposition 53.93 in order to see that it is applicable. Since the pair is Riemannian, $\mathrm{Ad}(K)$ –which is the analytic Lie subgroup of $\mathrm{Int}(\mathfrak{g})$ image of K by Ad– is compact. The Lie algebra of $\mathrm{Ad}(K)$ is given by thinks of the form

$$\frac{d}{dt}\Big[\,\mathrm{Ad}(k(t))\Big]_{t=0} = d\,\mathrm{Ad}_e(k'(0)) = \mathrm{ad}\,k'(0) \tag{66.58}$$

then the Lie algebra of $\mathrm{Ad}(\mathfrak{k})$ is $\mathrm{ad}(\mathfrak{k})$. Thus the fact that $\mathrm{Ad}(K)$ is compact is equivalent than the fact that $\mathfrak{k}$ is compactly embedded in $\mathfrak{g}$.

□

We can build a symmetric pair from an involutive automorphism σ of G. Take $H = (G_\sigma)_0$ and the pair (G,H); it is clear that it is a symmetric pair. However it is not automatically a Riemannian one.

Proposition 66.30.
Consider a Lie algebra $\mathfrak{g}$ and a direct decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. Then the map $\sigma = id_{\mathfrak{h}} \oplus (-id)_{\mathfrak{m}}$ is an automorphism of $\mathfrak{g}$ if and only if

$$[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{h} \tag{66.59a}$$
$$[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m} \tag{66.59b}$$

prop:invol_ssi_comm

Proof. We just have to compute $\sigma[h+m, h'+m']$ and $[\sigma(h+m), \sigma(h'+m')]$ and see under which conditions it is equal. □

66.3.4.1 Example: Lie group

Let L be a Lie group endowed with the structure

$$s_x y = xy^{-1}x. \tag{66.60}$$

It is immediate to check that $\forall x, y \in L$, $(s_x \circ s_x)(y) = y$ and $s_x \circ s_y \circ s_x = s_{s_x(y)}$. In order to see that x is an isolated fixed point of s, first remark that

$$x(s_y(z)) = (xy)(xz)^{-1}(xy) = s_{xy}(xz),$$

so that one just needs to check the property on s_e because the left translation is analytic. Since $s_e(y) = y^{-1}$, the property follows from the fact that e is an isolated fixed point for the inversion in a topological group.

66.3.4.2 Example: homogeneous spaces

Let L be a connected Lie group with an involutive automorphism, and L^σ the set of fixed points by σ. We consider a subgroup K such that $L_0^\sigma \subset K \subset L^\sigma$. The space L^σ is closed because it is defined by some equalities. The theorem 53.113 assure us that as topological Lie subgroup of L, K is also closed.

Now we consider $M = L/K$ and for $x \in L$, we define the translations $\tau(x)\colon M \to M$, $\tau(x)[y] = [xy]$ where the classes are defined with respect to K: $[x] = [xk]$ for any $k \in K$. We also define

$$L_\sigma = \{x\sigma(x)^{-1} : x \in L\}; \tag{66.61}$$

this is the space of the **symmetric elements** of L.

66.3.5 Symmetric spaces as quotient

We saw in the previous subsection that an affine symmetric space gives rise to a homogeneous space G/H and an involutive automorphism σ of G. From now we define a **symmetric space** as a triple (G, H, σ) where

— G is a Lie group,

— H is a closed subgroup of G,

— σ is an involutive automorphism of G such that $G_\sigma^0 \subset H \subset G_\sigma$

where $G_\sigma = \{g \in G \text{ tel que } \sigma(g) = g\}$. The space is **effective** if the largest normal subgroup N of G contained in H reduces to the identity. As N is normal in G and contained in H, the quotients G/N and H/N admits a canonical group structure.

Here, we will suppose that G is connected, but it is not an important issue.

Proposition 66.31.
If σ' is the involutive automorphism on G/N induced from σ and (G, H, σ) is a symmetric space, then $(G/N, H/N, \sigma')$ is an effective symmetric space.

Remark that σ' is well defined because, from definition, $\sigma'([g]) = [\sigma(g)]$, then for $h \in H$

$$\sigma'[gh] = [\sigma(g)\sigma(h)] = [\sigma(g)h] = [\sigma(g)] \tag{66.62}$$

because $H \subset G_\sigma$ implies $\sigma(h) = h$.

Proof. Let S be a normal subgroup of G/N contained in G/H. From the definitions, $S = \{id\}$. Indeed $S \subset H/N$; let $[a] \in S$ and $[g] \in G/N$. The first point is that $[gag^{-1}] \in S$. On the other hand if $[a] \in S$, then $a \in H$ because $S \subset H/N$ and $N \subset H$. Then for all representative g of $[g]$, $gag^{-1} \in H$. In particular for all $g \in G$, $gag^{-1} \in H$ and $a \in H$. From this, the thesis is immediate. □

Following definition 66.14 of a symmetric space, we should define good symmetries on G/H from the data of (G, H, σ). At $\vartheta \in G/H$, we define $s_\vartheta = \sigma'$. Let $g \cdot \vartheta = [g]$ be a fixed point of s_ϑ for a certain $g \in G$. Hence $\sigma'([g]) = [g]$, but from the definition of σ', we also have $\sigma'(g \cdot \vartheta) = [\sigma(g)]$. Then $\sigma(g) \in [g]$. Let $h = g^{-1}\sigma(g) \in H$. Since $\sigma(h) = h$, $h^2 = h\sigma(h)$, but $\sigma(h) = \sigma(g^{-1}\sigma(g))$, then $h^2 = e$. Since σ is an automorphism, if g is near the identity, then h will be too and $h^2 = e$. So g is near the identity and invariant under σ, then $g \in G_\sigma^0 \subset H$ and $g \cdot \vartheta = \vartheta$.

Let $x = g \cdot \vartheta$. We set

$$s_x = g \circ s_\vartheta \circ g^{-1}.$$

As a first remark, the choice of $g \in G$ such that $x = g \cdot \vartheta$. Indeed consider a $k \in G$ such that $gk \cdot \vartheta = x = g \cdot \vartheta$; we must show that $s_x = gks_\vartheta k^{-1}g^{-1}$. Since $k \in H$, it is sufficient to prove that for all $h \in H$, $h \circ s_\vartheta = s_\vartheta$. For this, let $[g] \in G/H$.

$$s_\vartheta[h^{-1}g] = [\sigma(h^{-1})\sigma(g)] = [h^{-1}\sigma(g)] = h^{-1}s_\vartheta[g], \tag{66.63}$$

so that $(hs_\vartheta h^{-1}) = s_{\vartheta}[g]$.

The **transvection group** is the subgroup of $\mathrm{Aut}(M, \omega, s)$ spanned by

$$\{s_x \circ s_y \text{ tel que } x, y \in M\}.$$

The definition of s_x also fulfils $s_x \circ s_y \circ s_x^{-1} = s_{s_x(y)}$. In order to see it, let us consider $x = g \cdot \vartheta$ and $y = k \cdot \vartheta$.

$$s_x \circ s_y \circ s_x^{-1} = g s_\vartheta g^{-1} k s_\vartheta k^{-1} g s_\vartheta^{-1} g^{-1} = s_{(gs_\vartheta g^{-1}k)\,\cdot\,\vartheta} = s_{s_x y}. \tag{66.64}$$

66.3.6 Symmetric Lie algebras

A **symmetric Lie algebra** is a triple $(\mathfrak{g}, \mathfrak{h}, \sigma)$ with

— $\mathfrak{g}$: a Lie algebra,

— $\mathfrak{h}$: a Lie subalgebra of $\mathfrak{g}$,

— σ: an involutive automorphism of $\mathfrak{g}$ whose $\mathfrak{h}$ is the set of fixed points.

Proposition 66.32.
Every symmetric space (G, H, σ) gives rise to a symmetric Lie algebra $(\mathfrak{g}, \mathfrak{h}, \sigma')$ with $\mathfrak{g}$ and $\mathfrak{h}$ being the Lie algebras of G and H while $\sigma' = d\sigma_e$.

Proof. If $X \in \mathfrak{h}$, one has $d\sigma_e X = \frac{d}{dt}\Big[\sigma e^{tX}\Big]_{t=0}$. Since $\mathfrak{h}$ is the Lie algebra of a Lie subgroup of G, equation (53.13) makes

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ tel que } t \to e^{tX} \text{ is a path in } H\}.$$

Then $e^{tX} \in H$ for all t and $\sigma(e^{tX}) = e^{tX} \in H$ because $H \subset G_\sigma$. This proves that the elements of $\mathfrak{h}$ are fixed by σ'.

Let us see the converse. If X is fixed by σ'

$$d\sigma_e X = \frac{d}{dt}\Big[\sigma(e^{tX})\Big]_{t=0} \overset{!}{=} X = \frac{d}{dt}\Big[e^{tX}\Big]_{t=0}.$$

This equation shows that e^{tX} and $\sigma e^{tX} = e^{t\sigma' X}$ are two exponential path whose start at the same point with the same tangent vector. They are equals on a neighbourhood of e. In this case, $\sigma(e^{tX}) = e^{tX}$ and $e^{tX} \in H$. This gives $X \in \mathfrak{h}$.

□

The association of a symmetric space to a symmetric Lie algebra is less automatic. Let $(\mathfrak{g}, \mathfrak{h}, \sigma')$ be a symmetric Lie algebra. We first have to find a connected, simply connected Lie group G whose Lie algebra is $\mathfrak{g}$. From this we define $\sigma\colon G \to G$ by

$$\sigma(e^X) = e^{\sigma' X}.$$

This is a local definition. Under analyticity hypothesis, one can extend σ into the whole G. Now one can take any subgroup H of G such that $G_\sigma^0 \subset H \subset G_\sigma$ to complete the symmetric space (G, H, σ).

¡¡ Avertissement/question à la lectrice !! 66.33
Il est dit que H est fermé parce qu'il est inclu à G_σ qui l'est.

66.3.6.1 Symmetric and reductive Lie algebras

Let $(\mathfrak{g}, \mathfrak{h}, \sigma)$ be a symmetric Lie algebra. As linear transformation of the vector space, σ has eigenvalues 1 and -1 (because it is involutive) and then induces a decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m} \tag{66.65}$$

where $\mathfrak{h}$ is the $+1$ eigenspace and $\mathfrak{m}$ the -1 eigenspace. This is the **canonical decomposition**. It is easy to see that this decomposition fulfils

$$[\mathfrak{h},\mathfrak{h}]\subset\mathfrak{h},\quad [\mathfrak{h},\mathfrak{m}]\subset\mathfrak{m},\quad [\mathfrak{m},\mathfrak{m}]\subset\mathfrak{h}. \tag{66.66}$$

eq:propreduc

On the one hand, if it exists, the homogeneous space G/H is automatically reductive. On the other hand if we have a Lie algebra $\mathfrak{g}$ and a decomposition $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{m}$ which fulfils (66.66), then definition $\sigma=\mathrm{Id}_{\mathfrak{h}}\oplus(-\mathrm{Id})_{\mathfrak{m}}$ gives a symmetric Lie algebra $(\mathfrak{g},\mathfrak{h},\sigma)$.

A homogeneous space is symmetric if and only if it is reductive.

Proposition 66.34.
Let (G,H,σ) a symmetric space, $(\mathfrak{g},\mathfrak{h},\sigma')$ the corresponding symmetric Lie algebra and $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{m}$ its canonical decomposition. Then

$$\mathrm{Ad}(H)\mathfrak{m}\subset\mathfrak{m}.$$

Proof. If $X\in\mathfrak{m}$ and $h=e^Y\in H$, then

$$\sigma'(\mathrm{Ad}(e^Y)X)=\mathrm{Ad}(e^{\sigma' Y})(\sigma' X)=\mathrm{Ad}(\sigma h)(\sigma' X)=\mathrm{Ad}(h)(-X)=-\mathrm{Ad}(h)X$$

because $\sigma h=h$ and $\sigma' X=-X$. So $\mathrm{Ad}(h)X\in\mathfrak{m}$ because it has eigenvalue -1 for σ. □

66.3.6.2 An affine example

Let M be an affine locally symmetric n-dimensional space. We consider $x\in M$, $\mathfrak{m}=T_xM$ and the curvature tensor R_x. Let $\mathfrak{h}$ be the set of linear endomorphism $U\colon T_xM\to T_xM$ which sends R_x on zero. More precisely an endomorphism of T_xM extends to a derivation of the tensor algebra with the definition

$$(U\cdot R_x)(X,Y)=U\big(R_x(X,Y)\big)-R_x(UX,Y)-R_x(X,UY)-R_x(X,Y)\circ U. \tag{66.67}$$

eq:defHa

In order to understand the last term, let us recall ourself that the curvature is given, from the connection $\nabla\colon\mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M)$ by formula

$$R(X,U)Z=\nabla_X\nabla_Y Z-\nabla_Y\nabla_X Z-\nabla_{[X,Y]}Z.$$

So one can see pointwise $R_x\colon T_xM\times T_xM\to\mathrm{End}(T_xM)$. Now we define $\mathfrak{h}$ by the condition $(U\cdot R_x)(X,Y)=0$ for all X, $Y\in\mathfrak{m}$. One can prove that $\mathfrak{h}$ is a Lie algebra for the usual bracket.

Remark that for all X, $Y\in\mathfrak{m}$, the endomorphism $R_x(X,Y)$ belongs to $\mathfrak{h}$ because $R(X,Y)=[\nabla_X,\nabla_Y]-\nabla_{[X,Y]}$ and $\nabla R=0$. We consider the direct sum $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ on which we put a Lie algebra structure by defining

$$[X,Y]=-R(X,Y) \qquad X,Y\in\mathfrak{m} \tag{66.68a}$$
$$[U,X]=UX \qquad U\in\mathfrak{h},X\in\mathfrak{m} \tag{66.68b}$$
$$[U,V]=[U,V] \qquad U,V\in\mathfrak{h}. \tag{66.68c}$$

We have to check the Jacobi identities. The first case is $X,Y,Z\in\mathfrak{m}$. It gives $[X,Y]=-R(X,Y)\in\mathfrak{h}$, then $[[X,Y],Z]=-R(X,Y)Z$ and the cyclic sum is zero from Bianchi. If X, $Y\in\mathfrak{m}$ and $U\in\mathfrak{h}$, then

$$[[X,Y],U]=-[R(X,Y),U]=-R(X,Y)\circ U+U\circ R(X,Y) \tag{66.69a}$$
$$[[Y,U],X]=[-UY,X]=R(UY,X) \tag{66.69b}$$
$$[[U,X]Y]=[UX,Y]=R(UX,Y). \tag{66.69c}$$

From definition (66.67), the sum of these three terms is zero. The last case, U, $V\in\mathfrak{h}$, $X\in\mathfrak{m}$, is easy.

With all that, the algebra $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}$ becomes a Lie algebra satisfying (66.66). Then it gives rise to a symmetric Lie algebra with $\sigma=\mathrm{Id}_{\mathfrak{m}}\oplus(-\mathrm{Id})_{\mathfrak{h}}$

66.3.7 Connection on symmetric spaces

We use theory from [956]. First, we extend the notion of tangent bundle. Consider a smooth curve $c\colon \mathbb{R} \to M$. Its **acceleration** $\ddot{c}(0)$ at the point $c(0)$ is defined by its action on a function $f\colon M \to \mathbb{R}$:

$$\ddot{c}(0)j = \frac{d^2}{dt^2}\big(f(c(t))\big) \tag{66.70}$$

with usual abuse of notation. The set of such accelerations at x is denoted by $T^2_x M$, and we naturally define the bundle T^2M with a suitable manifold structure. The set of sections of T^2M is logically denoted by $\mathfrak{X}^2(M)$. If X, $Y \in \mathfrak{X}(M)$, we define the **symmetric product** by

$$X \ltimes Y = \frac{1}{2}(X \otimes Y + Y \otimes X), \tag{66.71}$$

and the "composition product"

$$X \bullet Y = \mathtt{proj}_{T^2M}\, XY, \tag{66.72}$$

or in local coordinates

$$(X \bullet Y)_x f = X^i(x) Y^j(x)\, \frac{\partial^2 f}{\partial x_i \partial x_j}\bigg|_x .$$

We finally define, for X, $Y \in \mathfrak{X}(M)$

$$\begin{cases} P_2(X,Y) = \dfrac{1}{2}(X \otimes Y + Y \otimes X) \\ P_2(X) = 0. \end{cases} \tag{66.73}$$

Lemma 66.35.
For each connection on TM, there exists one and only one connection form $\Gamma\colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to T^2M$ such that

$$P_2(\Gamma(X,Y)) + X \ltimes Y = 0. \tag{66.74}$$

The correspondence is given by

$$\nabla_X Y = XY + \Gamma(X,Y) \tag{66.75}$$

Sketch of proof. Let us just give the link between equation (66.74) and our general culture about Christoffel symbols. The general form of a $\Gamma(X,Y) \in T^2M$ which is bilinear with respect to X and T is

$$\Gamma(X,Y) = X^k Y^l \Gamma^{ij}_{kl} \partial^2_{ij} + X^k Y^l \Gamma^i_{kl} \partial_i.$$

If we look at the coefficient of $\partial_i \otimes \partial_j$ when we impose the condition

$$X^k Y^l \Gamma^{ij}_{kl} \frac{1}{2}(\partial_i \otimes \partial_j + \partial_j \otimes \partial_i) + \frac{1}{2}(X^i Y^j \partial_i \otimes \partial_j + X^i Y^j \partial_j \otimes \partial_i) = 0,$$

we find

$$X^k Y^l (\Gamma^{ij}_{kl} + \Gamma^{ji}_{kl}) = -X^i Y^j.$$

If we suppose that Γ^{ij}_{kl} is symmetric with respect to ij, we find $\Gamma^{ij}_{kl} = -\frac{1}{2}\delta^i_j \delta^j_l$, so

$$\Gamma(X,Y) = -\frac{1}{2} X^i Y^j \partial^2_{ij} + X^k Y^l \Gamma^i_{kl} \partial_i. \tag{66.76}$$

□

Let us now consider a manifold M with a product $\mu(x,y) = x \cdot y$ which is $s_x y$ in the case of a symmetric space. It induces a product on TM by the following formula:

$$(X \cdot Y) f = (X \otimes Y)(f \circ \mu). \tag{66.77}$$

More explicitly, the function $f \circ \mu \colon M \times M \to \mathbb{R}$ has two entries; the product $X \otimes Y$ apply with X on the first entry and Y on the second one:

$$(X \cdot Y)_x f = \frac{d}{dt}\frac{d}{ds}\Big[(f \circ \mu)(X_x(t), Y_x(s))\Big]_{\substack{s=0\\t=0}}$$

where $X_x, Y_x \colon \mathbb{R} \to M$ are path defining X_x and $Y_x \in T_xM$. We can extend pointwise this product to a product between vector fields: $(X \cdot Y)_x = X_x \cdot Y_x$ when X, $Y \in \mathfrak{X}(M)$.

An easy adaptation of equation (66.77) defines $X \cdot x$ when $X \in T_pM$ and $x \in M$:

$$(X \cdot p)f = (u \otimes p)(f \circ \mu) = \frac{d}{dt}\Big[f\big(u(t) \cdot p\big)\Big]_{t=0} \tag{66.78}$$

because the expression $(u \otimes \mu)(f \circ \mu)$ suggests to put u in the first entry of μ and p in the second one. It defines a $X \cdot p \in T_pM$. From $v \in T_oM$, we can build $\tilde{v} \in \mathfrak{X}(M)$ by

$$\tilde{v}_p = \frac{1}{2} v \cdot (o \cdot p), \tag{66.79}$$

explicitly:

$$\tilde{v}_p f = \frac{1}{2}\frac{d}{dt}\Big[f\big(v(t) \cdot (o \cdot p)\big)\Big]_{t=0}$$

Theorem 66.36.
When X, $Y \in \mathfrak{X}(M)$, formula

$$\Gamma(X, Y) = \frac{1}{2} X \cdot Y \tag{66.80}$$

defines a connection on M.

Proof. Since $XY = X(Y^i\partial_i + X^iY^j\partial_{ij}$, we immediately see that $P_2(XY) = X \ltimes T$; it remains to be proved that $-v\tilde{u} = \frac{1}{2} u \cdot v$ for all u, $v \in T_oM$. This is a computation using the definitions:

$$\begin{aligned}
(v\tilde{u})f = v(\tilde{u}f) &= \frac{d}{ds}\Big[(\tilde{u}f)_{v(s)}\Big]_{s=0} \\
&= \frac{d}{ds}\Big[\frac{1}{2}\frac{d}{dt}\Big[f\big(u(t) \cdot (o \cdot v(s))\big)\Big]_{t=0}\Big]_{s=0} \\
&= \frac{1}{2}\frac{d}{dt}\Big[(df \circ d\mu_{u(t)} \circ \underbrace{d\mu_o}_{=-\mathbb{1}})v\Big]_{t=0} \\
&= -\frac{1}{2}\frac{d}{dt}\Big[df \circ d\mu_{u(t)} v\Big]_{t=0} \\
&= -\frac{1}{2}\frac{d}{dt}\frac{d}{ds}\Big[f\big(u(t) \cdot v(s)\big)\Big]_{\substack{s=0\\t=0}} \\
&= -\frac{1}{2}(u \cdot v)f.
\end{aligned} \tag{66.81}$$

□

66.3.8 Canonical connection and covariant derivative

subsecCanConCovDer

Let G be a Lie group, H a closed Lie subgroup and let us consider the principal bundle

$$\begin{array}{ccc} H & \rightsquigarrow & G \\ & & \downarrow{\scriptstyle \pi} \\ & & G/H \end{array} \tag{66.82}$$

with the action of H on G being defined by $g \cdot h = gh$ and π being the canonical projection. We have a canonical identification $T_{[e]}(G/H) = \mathcal{G}/\mathcal{H}$. We suppose that G is connected and that

(G,H) is a symmetric pair: we have an involutive automorphism $\sigma\colon G\to G$ for which H is the set of fixed points. We suppose moreover that H does not contain non trivial normal subgroups. Let $\mathcal{Q}$ be the space of vector such that $d\sigma(X)=-X$. By the canonical projection parallel to $\mathcal{H}$, we have an identification $\mathcal{Q}=\mathcal{G}/\mathcal{H}$.

When $g\in G$, we define

$$\begin{aligned} r(g)\colon \mathcal{Q}&\to T_{[g]}(G/H)\\ X&\mapsto d\pi dL_g X. \end{aligned} \tag{66.83}$$

We have $r(g)=r(g')$ when there exists a $h\in H$ such that $g'=gh$ and $\rho(h)=\mathrm{Id}$ where ρ is defined by

$$\begin{aligned} \rho(t)\colon \mathcal{Q}&\to \mathcal{Q}\\ X&\mapsto dL_t(X) \end{aligned} \tag{66.84}$$

for all $t\in H$. This definition works because of the identification $\mathcal{Q}=\mathcal{G}/\mathcal{H}$.

66.4 The group SO(2,1) and its algebra

subsec_IwSOdu

The condition for a matrix A to belong to $\mathfrak{so}(2,1)$ is $A^tgA=g$ with $g=diag(1,1,-1)$. Hence considering the Cartan involution pg:Cartan_SO $\theta(X)=-X^t$, we find

$$\mathfrak{g}=\mathfrak{so}(2,1)\hookrightarrow\begin{pmatrix}0&a&u_1\\-a&0&u_2\\u_1&u_2&0\end{pmatrix},\quad \mathfrak{k}\hookrightarrow\begin{pmatrix}0&1&\\-1&0&\\&&0\end{pmatrix},\quad \mathfrak{p}\hookrightarrow\begin{pmatrix}&&u_1\\&&u_2\\u_1&u_2&0\end{pmatrix}. \tag{66.85}$$ EqMtrsDansSOdu

As generator of $\mathfrak{a}$, we choose

$$J=\begin{pmatrix}&0&\\0&0&1\\&1&\end{pmatrix}.$$

The condition $[J,X]=X$ which defines $\mathcal{N}$ is satisfied by

$$L=\begin{pmatrix}0&1&1\\-1&&\\1&&\end{pmatrix}.$$

If we denote by X the general matrix of $\mathfrak{so}(2,1)$ given in equation (66.85), we have

$$\mathrm{ad}(J)X=\begin{pmatrix}0&-u_1&-a\\u_1&0&0\\-a&0&0\end{pmatrix}$$

which leads to the following root spaces

$$S_0=\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix},\quad S_1=\begin{pmatrix}0&1&-1\\-1&0&1\\-1&1&0\end{pmatrix},\quad S_{-1}=\begin{pmatrix}0&1&1\\-1&0&1\\1&1&0\end{pmatrix} \tag{66.86}$$

with the relations

$$[S_0,S_1]=S_1 \tag{66.87a}$$
$$[S_0,S_{-1}]=-S_{-1} \tag{66.87b}$$
$$[S_1,S_{-1}]=-2S_0. \tag{66.87c}$$

Notice that the comparison with (53.329) shows that $SL(2,\mathbb{R})$ and $\mathrm{SO}(2,1)$ *are not* isomorphic.

66.4.1 Root spaces for $\mathfrak{so}(2,1)$

Proposition 66.37.
The matrices of $\mathrm{SO}(p,q)$ *satisfy the relation*

$$A^{-1} = \eta A^t \eta$$

where η *is the diagonal matrix with* p *times* 1 *and* q *times* -1*. That relation is often taken as the definition of* $\mathrm{SO}(p,q)$*.*

The algebra $\mathfrak{so}(2,1)$ is made up from 3×3 matrices such that $X^t\eta + \eta X = 0$ with vanishing trace. If we choice $\eta = diag(-,-,+)$, we find matrices of the form

$$\mathfrak{so}(2,1) \rightsquigarrow \begin{pmatrix} a & u^t \\ u & 0 \end{pmatrix}$$

where a is an antisymmetric 2×2 matrix and u is any 1×2 matrix. We find the Cartan decomposition

$$\mathcal{K} \rightsquigarrow \begin{pmatrix} \mathfrak{so}(2) & \\ & 0 \end{pmatrix}, \qquad \mathcal{P} \rightsquigarrow \begin{pmatrix} 0 & 0 & u_1 \\ 0 & 0 & u_2 \\ u_1 & u_2 & 0 \end{pmatrix}.$$

If one chooses

$$J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & & \\ 1 & & \end{pmatrix}$$

as generator for the abelian subalgebra of $\mathcal{P}$, one finds

$$V_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad V_{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

as eigenvectors for $\mathrm{ad}(J)$ with eigenvalues 1 and -1. The root space decomposition commutator table of $\mathfrak{so}(2,1)$ is thus given by

$$[V_0, V_1] = V_1 \tag{66.88a}$$
$$[V_0, V_{-1}] = -V_{-1} \tag{66.88b}$$
$$[V_1, V_{-1}] = -2V_0. \tag{66.88c}$$

Notice that the map $\phi(A_0) = 2V_0$, $\phi(A_2) = V_1$, $\phi(A_{-2}) = -V_{-1}$ provides an isomorphism between this table and the one of $\mathfrak{sl}(2,\mathbb{R})$, equations (53.329). This fact assures a Lie algebra isomorphism $\mathfrak{sl}(2,\mathbb{R}) \simeq \mathfrak{so}(2,1)$. We actually have a stronger result:

Proposition 66.38.
The group $\mathrm{SL}(2,\mathbb{R})$ *is a double-covering of the identity component* $\mathrm{SO}_0(1,2)$*.*

66.4.2 Position of $\mathrm{SO}(1,n)$ inside $\mathrm{SO}(2,n)$

The quotient $\mathrm{SO}(2,n)/\mathrm{SO}(1,n)$ has a particular importance in section 79.5. Hence, we will work out the Iwasawa decompositions of these groups imposing certain compatibility conditions. We already build the Cartan involution θ in such a way that $[\sigma,\theta] = 0$. In this section, $\mathcal{H}$ is $\mathrm{SO}(1,n)$ and we are searching for an Iwasawa decomposition $\mathcal{A}_\mathcal{H} \oplus \mathcal{N}_\mathcal{H} \oplus \mathcal{K}_\mathcal{H}$.

As we want compatibility between the Cartan decomposition of $\mathrm{SO}(1,n)$, the Cartan decomposition of $\mathrm{SO}(2,n)$ and the reductive decomposition of $\mathrm{SO}(2,n)$ we start by giving the "position" of $\mathrm{SO}(1,n)$ inside $\mathrm{SO}(2,n)$. The in subsection 66.4.3 we will proceed with the Iwasawa decomposition of $\mathrm{SO}(1,n)$.

lem:SO_pq_ss

Lemma 66.39.
The Lie group $\mathrm{SO}(p,q)$ *is semisimple.*

Sketch of proof. We are going to give some ideas of the proof. Remark that $\mathrm{SO}(p,q)$ essentially contains "rotations" $J_{ij}(\lambda)$ of angle λ in planes (i,j). (If the coordinates i and j are not of the same type[4], then it is not a true rotation but a boost and λ is no more an angle but any real.) If one try to build an ideal $\mathcal{I}$ containing J_{ij}, then the element $J_{kl}J_{ij}$ must also be in $\mathcal{I}$. If we want $\mathcal{I}$ to be abelian, we must have $[J_{ij}, J_{ij}J_{kl}]$, or

$$J_{kl}J_{ij} = J_{ij}J_{kl}.$$

It is really easy to find a J_{kl} for which it is false. Thus $\mathrm{SO}(p,q)$ doesn't contain any abelian ideal, so that it is semisimple by lemma 51.56. □

¡¡ Avertissement/question au lecteur !! 66.40
J'ai l'impression que c'est la simplicité de $\mathrm{SO}(p,q)$ *que je démontre. À vérifier.*

The Lie algebra $\mathcal{G} = \mathfrak{so}(2,n)$ is the set

$$\left\{X \in M_{(2+n)\times(2+n)} \text{ such that } X^t\eta + \eta X = 0 \text{ and } \operatorname{Tr} X = 0\right\} \tag{66.89}$$

def_sodn

where η is the diagonal metric $\eta = diag(-,-,+,\dots,+)$. An element of $\mathfrak{so}(2,n)$ can be written as $X = \begin{pmatrix} a & u^t \\ v & B \end{pmatrix}$ with the matrices $a \in M_{2\times 2}$, $u \in M_{n\times 2}$, $v \in M_{n\times 2}$, and $B \in M_{n\times n}$. The conditions in (66.89) give: $a = -a^t$, $u = v$, and $B = -B^t$. Hence, a general matrix of $\mathfrak{so}(2,n)$ is given by

$$X = \begin{pmatrix} a & u^t \\ u & B \end{pmatrix} \tag{66.90}$$

eq:gene_sodn

where a, B are skew-symmetric.

A compatible reductive decomposition is build as follows. The matrices of $\mathfrak{so}(1,n)$ have to be seen as matrices of $\mathfrak{so}(2,n)$ with the condition $Y^t\sigma + \sigma Y = 0$ for the "metric" $\sigma = diag(0,-,+,\dots,+)$. Hence,

$$\mathcal{H} = \mathfrak{so}(1,n) \rightsquigarrow \begin{pmatrix} 0 & 0 & \begin{pmatrix} \cdots 0 \cdots \\ \leftarrow v^t \rightarrow \end{pmatrix} \\ 0 & 0 & \\ \begin{pmatrix} \vdots & \uparrow \\ 0 & v \\ \vdots & \downarrow \end{pmatrix} & & B \end{pmatrix} \tag{66.91}$$

eq:gene_H

where $v \in M_{n\times 1}$ and $B \in M_{n\times n}$ is skew-symmetric. Comparing this with the general form (66.90) of a matrix of $\mathfrak{so}(2,n)$ matrix, one immediately finds that, with the choice

$$\mathcal{Q} \rightsquigarrow \begin{pmatrix} 0 & a & \begin{pmatrix} \leftarrow w^t \rightarrow \\ \cdots 0 \cdots \end{pmatrix} \\ -a & 0 & \\ \begin{pmatrix} \uparrow & \vdots \\ w & 0 \\ \downarrow & \vdots \end{pmatrix} & & 0 \end{pmatrix}, \tag{66.92}$$

EqGeneRedQ

the decomposition $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$ is reductive:

$$[\mathcal{H}, \mathcal{Q}] \subseteq \mathcal{Q}, \qquad [\mathcal{Q}, \mathcal{Q}] \subseteq \mathcal{H}, \tag{66.93}$$

EqDefRedHQ

and $B(\mathcal{H}, \mathcal{Q}) = 0$. In the sequel, we will use the basis of $\mathcal{Q}$ defined by

$$q_0 = E_{12} - E_{21}, \qquad q_i = E_{1,(i-2)} + E_{(i-2),1}. \tag{66.94}$$

EqDefBaseQi

Notice, for later use that $q_1 = J_2$ in the Iwasawa decomposition of $\mathrm{SO}(2,n)$.

4. i.e. if the metric in the plane (i,j) has signature $(+,-)$.

66.4.3 Iwasawa decomposition for $\mathrm{SO}(1,n)$

SubSecIwaSOunn

We want our Cartan involution $\theta\colon \mathcal{G}\to\mathcal{G}$, $\theta(X)=-X^t$ and to be such that $[\theta,\sigma]=0$ with $\sigma\colon\mathcal{G}\to\mathcal{G}$ given by $\sigma=id|_{\mathcal{H}}\oplus(-id)|_{\mathcal{Q}}$, see theorem 66.18.pg:calcul_sigma_theta

Consider a general matrix h in $\mathcal{H}$ (from (66.91)); we have

$$(\sigma\theta)(h)=-\sigma\begin{pmatrix}0&0&0\\0&0&v^t\\0&v&-B\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&-v^t\\0&-v&B\end{pmatrix},$$

while

$$\theta\sigma(h)=\theta(h)=-\begin{pmatrix}0&0&0\\0&0&v^t\\0&v&-B\end{pmatrix},$$

then $[\sigma,\theta]|_{\mathcal{H}}=0$. The same computation with a matrix in $\mathcal{Q}$ (see (66.92)) gives $[\sigma,\theta]|_{\mathcal{Q}}=0$.

Now we show that θ descent to a Cartan involution on H. It is clear that the restriction of θ is an involutive automorphism of H. Lemma 51.60 assures us that the restriction of the Killing form of G to H is the Killing form of H, so that the condition of positivity of B_θ holds on H as well as on G. Last, θ leaves $\mathcal{H}$ invariant. Indeed suppose that $\theta X_{\mathcal{H}}=X'_{\mathcal{H}}+X_{\mathcal{Q}}\in\mathcal{H}\oplus\mathcal{Q}$. Then $\sigma\theta X_{\mathcal{H}}=h'-q$ and $\theta\sigma h=h'+q$; since $[\theta,\sigma]=0$, we have $q=0$.

The whole proves that we can use the same Cartan involution on G as well as on H. Since $\theta=id|_{\mathcal{K}}\oplus(-id)|_{\mathcal{P}}$, it is clear that

$$\mathcal{K}_{\mathcal{H}}=\mathcal{K}\cap\mathcal{H},\qquad\qquad \mathcal{P}_{\mathcal{H}}=\mathcal{P}\cap\mathcal{H}\tag{66.95}$$

is the Cartan decomposition of $\mathcal{H}$. We can write explicit matrices as

$$\mathcal{K}_{\mathcal{H}}=\mathfrak{so}(n)\rightsquigarrow\begin{pmatrix}0&0&\cdots\\0&0&\cdots\\\vdots&\vdots&B\end{pmatrix},\tag{66.96}$$

where B is skew-symmetric, and

$$\mathcal{P}_{\mathcal{H}}\rightsquigarrow\begin{pmatrix}0&0&\cdots\\0&0&u^t\\\vdots&u&0\end{pmatrix}\tag{66.97}$$

where $u\in\mathbb{M}_{n\times1}(\mathbb{R})$ is a line. One remark that there are no two-dimensional subalgebra of $\mathcal{P}_{\mathcal{H}}$. So $\mathcal{A}_{\mathcal{H}}$ reduces to the choice of any element $J_1\in\mathcal{P}_{\mathcal{H}}$. A positivity notion is easy to find: the form $\omega\in{\mathcal{A}_{\mathcal{H}}}^*$ such that $\omega(J_1)=1$ is positive while $-\omega$ is negative. We choose

$$J_1=\begin{pmatrix}&0&&\\0&0&0&1\\&0&&\\&1&&\end{pmatrix}.$$

The computation of $\mathcal{N}_{\mathcal{H}}=\{X\in\mathcal{H}\text{ tel que }(\operatorname{ad}J_1)X=X\}$ yields the following:

$$\mathcal{N}_{\mathcal{H}}\rightsquigarrow\begin{pmatrix}\begin{pmatrix}&0&0\\&a&0\\0&a&0&-a\\0&0&a&0\end{pmatrix}&\begin{pmatrix}\cdots&0&\cdots\\\leftarrow&\overline{v}&\rightarrow\\\cdots&0&\cdots\\\leftarrow&\overline{v}&\rightarrow\end{pmatrix}\\\begin{pmatrix}\vdots&\uparrow&\vdots&\uparrow\\0&\overline{v}&0&-\overline{v}\\\vdots&\downarrow&\vdots&\downarrow\end{pmatrix}&0\end{pmatrix}.\tag{66.98}$$

eq:re_N_H

Finally, we consider the algebra $\mathcal{R}_{\mathcal{H}}=\mathcal{A}_{\mathcal{H}}\oplus\mathcal{N}_{\mathcal{H}}$.

66.5 Iwasawa decomposition for $\mathrm{SO}(2,n)$

subsecIwasawa_un

As seen in the general construction and in previous examples, the Iwasawa decomposition of a group or an algebra depends on several choices. We will study two out of them in the case of $\mathrm{SO}(2,n)$ and see that some "compatibility conditions" with the decomposition of $\mathrm{SO}(1,n)$ and the symmetric space structure of AdS (see section 79.5) fix most of choices. When the two decompositions will be used together, the unadapted one will be tilded (see subsection 79.7.4).

66.5.1 Cartan decomposition and compatible reductive decomposition

SubSecCartandeuxN

The Cartan decomposition of $\mathfrak{so}(2,n)$ associated with the Cartan involution $\theta(X) = -X^t$ is

$$\mathcal{K} \rightsquigarrow \begin{pmatrix} \mathfrak{so}(2) & \\ & \mathfrak{so}(n) \end{pmatrix}, \qquad \mathcal{P} \rightsquigarrow \begin{pmatrix} 0 & u^t \\ u & 0 \end{pmatrix}. \tag{66.99}$$

K_et_P

Elements of $\mathrm{SO}(2)$ are represented by

$$\begin{pmatrix} \cos\mu & \sin\mu \\ -\sin\mu & \cos\mu \end{pmatrix}.$$

A common abuse of notation in the text will be to identify the angle μ with the element of $\mathrm{SO}(2)$ itself. In the same spirit, when we speak about a matrix of $A \in \mathrm{SO}(2)$, we mean a matrix whose upper left corner is A and the rest is the unit matrix. For example, for AdS_3, the matrix $-\mathbb{1} \in \mathrm{SO}(2)$ is

$$\begin{pmatrix} -1 & 0 & & \\ 0 & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

That matrix will be denoted [5] by k_θ and has the property that

$$\theta = \mathbf{Ad}(k_\theta). \tag{66.100}$$

Indeed, we have $k_\theta = \mathbb{1}$, so that $\mathbf{Ad}(k_\theta)^2 = \mathrm{Id}$. Since k_θ commutes with every element of K, $\mathbf{Ad}(k_\theta)|_K = \mathrm{Id}$, and a simple computation shows that $\mathrm{Ad}(k_\theta)|_{\mathcal{P}} = -\mathrm{Id}$.

Remark that $\mathcal{K}$, the compact part of $\mathcal{G}$ is made up from "true" rotations while $\mathcal{P}$ contains boosts. This remark allows us to guess a right choice of maximal abelian subalgebra in $\mathcal{P}$. Indeed elements of $\mathcal{A}$ must be boosts and the fact that there are only two time-like directions restricts $\mathcal{A}$ to a two dimensional algebra. Up to reparametrization, it is thus generated by $t\partial_x + x\partial_t$ and $u\partial_y + y\partial_t$. Hence the following choice seems to be logical:

$$J_1 = \begin{pmatrix} & 0 & & \\ 0 & 0 & 0 & 1 \\ & 0 & & \\ & 1 & & \end{pmatrix} \in \mathcal{H}, \qquad J_2 = q_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & & & \\ 1 & & & \\ 0 & & & \end{pmatrix} \in \mathcal{Q}. \tag{66.101}$$

EqDeyineA

66.5.2 Maximal abelian subalgebra

We have to find an abelian subalgebra $\mathcal{A} \subset \mathcal{P}$. If we use the convention that the indices a, b range from 1 to 2 and i, j from 3 to $n+2$, we can write

$$u = \begin{pmatrix} 0 & u^t \\ u & 0 \end{pmatrix} = u^{ai}(E_{ai} + E_{ia}), \tag{66.102}$$

5. See also proposition 79.87.

where E_{ij} denotes the matrix whose only non zero component is a 1 in the place ij. Of course we use an abuse of notation between the $(n+2)\times(n+2)$ matrix of $\mathcal{P}$ and $n\times 2$ matrix u which defines it. We compute the commutator of two such matrices:

$$\begin{aligned}[u,v] = u^{ai}v^{bj}(&\underline{\delta_{ib}E_{aj}} - \underline{\delta_{aj}E_{bi}} + \delta_{ij}E_{ab} - \delta_{ab}E_{ji}\\ &+ \delta_{ab}E_{ij} - \delta_{ij}E_{ba} + \underline{\delta_{aj}E_{ib}} - \underline{\delta_{ib}E_{ja}})\end{aligned} \tag{66.103}$$

Since $[\mathcal{P},\mathcal{P}]\subseteq\mathcal{K}$, the latter commutator takes the form $\begin{pmatrix} a & 0\\ 0 & B\end{pmatrix}$, thus the underlined terms must vanish. So we are left with:

$$[u,v] = u^{ai}v^{bj}(\delta_{ij}(E_{ab}-E_{ba}) + \delta_{ab}(E_{ij}-E_{ji})) \tag{66.104}$$

$$= u^{ai}v^{bj}(E_{ab}-E_{ba}) + u^{ai}v^{aj}(E_{ij}-E_{ji}). \tag{66.105}$$

Finally, $[u,v]=0$ if and only if

$$\sum_i u^{ai}v^{bi} = 0 \quad\text{and}\quad \sum_a u^{ai}v^{aj} = 0 \tag{66.106}$$

EqCondSOdnsumia

with $a\neq b$, and $i\neq j$.

In order to build a maximal abelian subalgebra $\mathcal{A}$ of $\mathcal{P}$, we take one element in $\mathcal{P}$ and then we extend it using the relations (66.106). We choose $u^{ai}=\delta_{a1}E_{i3}$.

If this u lies in $\mathcal{A}$, an other v in $\mathcal{A}$ must satisfy $v\in\mathcal{P}$, $v^{23}=0$ and $v^{1j}=0$ for $j>3$. We choose the one with $u^{24}=1$. It is no difficult to see that these two are maximal. A basis of $\mathcal{A}$ is thus $\tilde{H}_1\equiv u^{24}=1$ and $\tilde{H}_2\equiv u^{13}=1$. We immediately perform a change of basis by defining:

$$H_1 = \tilde{H}_1 - \tilde{H}_2 \tag{66.107a}$$

$$H_2 = \tilde{H}_1 + \tilde{H}_2, \tag{66.107b}$$

and finally: $\mathcal{A}=span\{H_1,H_2\}$. The matrices are given by

$$H_p = \left(\begin{array}{c|c} N_p & 0\\ \hline 0 & 0\end{array}\right) \tag{66.108}$$

with

$$N_p = \begin{pmatrix} & & 1 & 0\\ & & 0 & (-1)^p\\ 1 & 0 & & \\ 0 & (-1)^p & & \end{pmatrix}. \tag{66.109}$$

Up to a change of basis, they are exactly the ones guessed in equation (66.101). That matrix can be written under the more compact form

The generators of $\mathcal{A}$ that we choose are the following linear combination of J_1 and J_2:

$$H_p = E_{13}+E_{31}+(-1)^p(E_{24}+E_{42}). \tag{66.110}$$

mtr_A

66.5.3 Compatibility conditions between $\mathcal{H}$ and $\mathcal{G}$

The decomposition of $\mathrm{SO}(1,n)$ of section 66.4.3 incites us to search for a decomposition of $\mathrm{SO}(2,n)$ which satisfies certain compatibility conditions. If you want some explicit matrix, read subsection 66.5.4 and 66.5.6.

We go back to the choice of $\mathcal{A}$. Remark that $\mathcal{P}_{\mathcal{H}}\subset\mathcal{P}$. This suggests us to try to get $\mathcal{A}_{\mathcal{H}}\subset\mathcal{A}$. So we will choose J as first matrix in $\mathcal{A}$, and then extend it to a maximal abelian subalgebra of $\mathcal{P}$. Note in particular that this don't affect the choice of the Cartan involution. Now we consider Φ and Ψ, the restricted roots for G and H; for recall:

$$\begin{aligned}\mathcal{H}_\alpha &= \{X\in\mathcal{H} \text{ tel que } \forall H'\in\mathcal{A}_{\mathcal{H}}, (\mathrm{ad}\,H')X=\alpha(H')X\},\\ \mathcal{G}_\lambda &= \{X\in\mathcal{G} \text{ tel que } \forall H\in\mathcal{A}, (\mathrm{ad}\,H)X=\lambda(H)X\},\\ \Phi &= \{\lambda\in\mathcal{A}^* \text{ tel que } \lambda\neq 0, \mathcal{G}_\lambda\neq 0\},\\ \Psi &= \{\lambda\in\mathcal{A}^*_{\mathcal{H}} \text{ tel que } \lambda\neq 0, \mathcal{H}_\lambda\neq 0\}.\end{aligned} \tag{66.111}$$

Lemma 66.41.
The root spaces of G and H are related by

$$\Phi|_{\mathcal{A}_\mathcal{H}} = \Psi,$$

in other words, elements of Ψ are the restriction to $\mathcal{A}_\mathcal{H}$ of the ones of Φ

Proof. Let us consider a $\alpha \in \Phi$: there exists a non-zero subspace $\mathcal{G}_\alpha$ of $\mathcal{G}$ such that for any $H \in \mathcal{A}$, $[H, X] = \alpha(H)X$. Clearly, α also fulfils $[H', X] = \alpha(H')X = \alpha|_{\mathcal{A}_\mathcal{H}}(H')X$ for any $H' \in \mathcal{A}_\mathcal{H} \subset \mathcal{A}$. Then $\alpha|_{\mathcal{A}_\mathcal{H}} \in \Psi$.

In order to prove the inverse sense, consider $\beta \in \Psi$ and $X \in \mathcal{H}_\beta$. We have $[H', X] = \beta(H')X$ for every $H' \in \mathcal{A}_\mathcal{H}$. Let us consider $H \in \mathcal{A}$:

$$[H', [H, X]] = [H, [H', X]] = \beta(H')[H, X]. \tag{66.112}$$

Thus for any $H \in \mathcal{A}$ and $X \in \mathcal{H}_\beta$, $[H, X] \in \mathcal{H}_\beta$. In other words,

$$[\mathcal{A}, \mathcal{H}_\beta] \subset \mathcal{H}_\beta.$$

Now, $\operatorname{ad}\mathcal{A}$ is a set of semisimple[6] commuting operators on $\mathcal{H}_\beta$; hence there exists a basis of $\mathcal{H}_\beta$ of common eigenvectors. Thus we can write a decomposition of $\mathcal{H}_\beta$ under the form

$$\mathcal{H}_\beta = \bigoplus_{\alpha\in\Phi} \mathcal{H}_\beta^\alpha, \tag{66.113}$$

eq:decomp_sH

with $\mathcal{H}_\beta^\alpha = \mathcal{H}_\beta \cap \mathcal{G}_\alpha$. In order to see this last point, consider an eigenvector $h_\beta \in \mathcal{H}_\beta$ of all the elements of $\operatorname{ad}\mathcal{A}$: there exists $\alpha \in \mathcal{A}^*$ such that $\forall A \in \mathcal{A}$, $[A, h_\beta] = \alpha(A)h_\beta$. In other words, $h_\beta \in \mathcal{G}_\alpha$.

This decomposition allows us to write $X \in \mathcal{H}_\beta$ as

$$X = \bigoplus_{\alpha\in\Phi} X^\alpha.$$

On $X \in \mathcal{H}_\beta^\alpha$, $\operatorname{ad} A$ acts as $(\operatorname{ad} A)X = \alpha(A)X$. So, for $X \in \mathcal{H}_\beta$, $H' \in \mathcal{A}_\mathcal{H}$,

$$[H', X] = \beta(H')X = \bigoplus_{\alpha\in\Phi} \alpha(H')X^\alpha.$$

Then $\alpha(H') = \beta(H')$, i.e. $\alpha|_{\mathcal{A}_\mathcal{H}} = \beta$.

□

Now, we choose a clever positivity notion on Φ. The one of Ψ is clear; we extend it to Φ, so that $\Psi^+ \subset \Phi^+$ and $\mathcal{N}_\mathcal{H} \subset \mathcal{N}$. Finally for our new Iwasawa decomposition of $\mathrm{SO}(2, n)$,

$$\mathcal{R}_\mathcal{H} \subset \mathcal{R}.$$

In the decomposition (66.113), it is possible that, for certain α and β, the set $\mathcal{H}_\beta^\alpha$ reduces to $\{0\}$. We define

$$\phi(\beta) = \{\alpha \in \Phi \text{ tel que } \mathcal{H}_\beta^\alpha \neq \{0\}\}; \tag{66.114}$$

it defines a map ϕ from Ψ to the parts of Φ. Moreover, $\phi(\Psi)$ gives a partition of Φ because $\phi(\beta) \cap \phi(\beta') = \varnothing$ if $\beta \neq \beta'$. Indeed, $\phi(\beta) \cap \phi(\beta')$ is the set of all the $\alpha \in \Phi$ such that $\mathcal{H}_\beta \cap \mathcal{G}_\alpha \neq \{0\}$ and $\mathcal{H}_{\beta'} \cap \mathcal{G}_\alpha \neq \{0\}$. But for $\alpha \in \phi(\beta)$, $\alpha_{\mathcal{A}_\mathcal{H}} = \beta$, then $\alpha(H') = \beta(H')$; for the same reason, $\alpha(H') = \beta'(H')$, so that $\beta = \beta'$.

We know the ± 1 eigenspaces decompositions $\mathcal{G} \overset{\sigma}{=} \mathcal{H} \oplus \mathcal{Q} \overset{\theta}{=} \mathcal{K} \oplus \mathcal{P}$ with $[\sigma, \theta] = 0$, and the Iwasawa decomposition $\mathcal{A}_\mathcal{H} \oplus \mathcal{N}_\mathcal{H} \oplus \mathcal{K}_\mathcal{H}$ of $\mathcal{H}$.

6. Il faudra encore voir ça...

For compatibility and simplicity purposes, we want the Iwasawa decompositions $\mathcal{G} = \mathcal{A}\oplus\mathcal{N}\oplus\mathcal{K}$ of $\mathcal{G}$ in such a way that $\mathcal{A}_\mathcal{H} \subset \mathcal{A}$ and $\mathcal{N}_\mathcal{H} \subset \mathcal{N}$. We denote by A, N, K, A_H, N_H and K_H the analytic connected subgroups of G whose Lie algebras are $\mathcal{A}$, $\mathcal{N}$, $\mathcal{K}$, $\mathcal{A}_\mathcal{H}$, $\mathcal{N}_\mathcal{H}$, and $\mathcal{K}_\mathcal{H}$ respectively.

For the $\mathcal{A}$-part, we just perform a change of basis

$$J_1 = \frac{1}{2}(H_2 - H_1) \tag{66.115a}$$

$$J_2 = \frac{1}{2}(H_1 + H_2) \tag{66.115b}$$

in order to have $J_1 \in \mathcal{P} \cap \mathcal{H}$ and $J_2 \in \mathcal{P} \cap \mathcal{Q}$. The involution σ has a simpler expression in this basis.

In order to get $\mathcal{A} \subset \mathcal{A}_\mathcal{H}$, we take the element of $\mathcal{A}_\mathcal{H}$ as first basis element of $\mathcal{A}$:

$$J_1 = \begin{pmatrix} & 0 & & \\ 0 & 0 & 0 & 1 \\ & 0 & & \\ & 1 & & \end{pmatrix}. \tag{66.116}$$ EqCeueleJun

One can check from from (66.92) and (66.99) that J_1 belongs to $\mathcal{P} \cap \mathcal{H}$. Now we want to extend it to a maximal abelian subalgebra of $\mathcal{P}$. In order to build an Iwasawa decomposition of $\mathrm{SO}(2,n)$ compatible with $\mathcal{Q}$, we search a $J_2 \in \mathcal{P} \cap \mathcal{Q}$. We know that

$$\mathcal{P} \cap \mathcal{Q} \rightsquigarrow \begin{pmatrix} 0 & 0 & u^t \\ 0 & 0 & \\ u & & \end{pmatrix},$$

so that the matrix

$$J_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & & & \\ 1 & & & \\ 0 & & & \end{pmatrix} \tag{66.117}$$ EqgueueleJdeux

belongs to $\mathcal{P} \cap \mathcal{Q}$ and commutes with J_1.

Remarque 66.42.
The new $\mathcal{A}$ for $\mathrm{SO}(2,n)$ is the abelian Lie algebra spanned by J_1 and J_2. This is just a change of basis in $\mathcal{A}$: $J_1 = \frac{1}{2}(H_2 - H_1)$, $J_2 = \frac{1}{2}(H_1 + H_2)$.

Note that $J_1 \in \mathcal{H}$ and $J_2 \in \mathcal{Q}$; then the expression of σ on $\mathcal{A}$ should be rather simple. Let us first show the action of σ on the root spaces.

If $\alpha \in \mathcal{A}^*$, the notation $\sigma^*\alpha$ means $f \circ \alpha$, and $\mathcal{G}_\alpha$ denotes the root space associated with the form α.

LemSigmaThetaRootSpaces

Lemma 66.43.
The involutions σ and θ act on the root spaces by

$$\sigma\mathcal{G}_\varphi = \mathcal{G}_{\sigma^*\varphi}, \qquad \theta\mathcal{G}_\varphi = \mathcal{G}_{-\varphi} \tag{66.118}$$ eq:sig_G_alpha

for all $\varphi \in \mathcal{A}^*$,

Proof. Let $H \in \mathcal{A}$ and $X \in \mathcal{G}_\varphi$; by definition: $[H,X] = \varphi(H)X$. We have

$$\varphi(H)\sigma X = \sigma[H,X] = [\sigma H, \sigma X] = \varphi(\sigma H)\sigma X.$$

We conclude that $\sigma X \in \mathcal{G}_{\sigma^*\varphi}$. For the second equality, we take $X \in \mathcal{G}_\varphi$ and

$$[H,\theta X] = \theta[\theta H, X] = \theta\big(\varphi(\theta H)X\big) = -\varphi(H)\theta(X).$$

□

LemSigmaChangeDeux

Lemma 66.44.
The involution σ changes the sign of the J_2^*-part of the root spaces:

$$\sigma\mathcal{G}_{(x,y)} = \mathcal{G}_{(x,-y)}$$

where (x, y) denote the coordinates of a root in the basis $\{J_1^*, J_2^*\}$ of $\mathcal{A}$.

Proof. Since σ is an involutive automorphism, it satisfies $[\sigma X, Y] = \sigma[X, \sigma Y]$. So when $X \in \mathcal{G}_{(x,y)}$, we have $(\mathrm{ad}(J_1))(\sigma X) = x\sigma X$ and $\mathrm{ad}(J_2)(\sigma X) = -y\sigma X$.

□

It is also clear that $\sigma^*\alpha = 0$ implies $\alpha(J_2) = 0$. Indeed, $(\sigma^*\alpha)(J_2) = \alpha(-J_2) = -\alpha(J_2)$, because $J_2 \in \mathcal{Q}$. Now we can give a precision about the decomposition $\mathcal{H}_\beta = \oplus_{\alpha\in\Phi'}\mathcal{H}_\beta^\alpha$ given by (66.113). In fact, $\Phi' = \Phi_\sigma = \{\alpha \text{ tel que } \sigma^*\alpha = \alpha\}$. Indeed σ fixes $\mathcal{H}$ and thus fixes $\mathcal{H}_\beta^\alpha$. If $h_\beta^\alpha \in \mathcal{H}_\beta^\alpha$, it satisfies $\sigma h_\beta^\alpha = h_\beta^\alpha$. But from equation (66.118), $\sigma h_\beta^\alpha \in \mathcal{G}_{\sigma^*\alpha}$.

Proposition 66.45.
The spaces $\mathcal{P}$ and $\mathcal{K}$ are Killing-orthogonal. The spaces $\mathcal{H}$ and $\mathcal{Q}$ are Killing-orthogonal.

Proof. Let consider $X \in \mathcal{H}$ and $Y \in \mathcal{Q}$. Since σ is an automorphism of the algebra $\mathcal{G}$, the Killing form is σ-invariant and we have

$$B(X, Y) = B(\sigma X, \sigma Y) = B(X, -Y) = -B(X, Y), \tag{66.119}$$

which proves that $B(X, Y) = 0$. Exactly the same holds for $X \in \mathcal{K}$ and $Y \in \mathcal{P}$, using θ instead of σ. □

66.5.4 First choice: AN

subsecIwasawa_deux

We turn now our attention to the Iwasawa decomposition. As before we are searching for matrices under the form $E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with dimensions $4+(n-2)$. The condition which determines the root spaces reads

$$(\mathrm{ad}\, J_i)E = \begin{pmatrix} [j_i, A] & j_i B \\ -Cj_i & 0 \end{pmatrix} \stackrel{!}{=} \lambda_i \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{66.120}$$

The computations are rather the same as the ones of the first time. The result is pg:root_n

$$\mathcal{G}_{(0,0)} \rightsquigarrow \begin{pmatrix} & & x & 0 & \\ & & 0 & y & \\ x & 0 & & & \\ 0 & y & & & \\ & & & & D \end{pmatrix}, \tag{66.121}$$

where $D \in M_{(n-2)\times(n-2)}$ is skew-symmetric. Notice that all but the D-part of this space is spanned by q_1 and J_1, so the $\mathcal{Q}$-component of that matrix is a multiple of q_1. Other root spaces are given by

$$\mathcal{G}_{(1,0)} \rightsquigarrow W_i = E_{2i} + E_{4i} + E_{i2} - E_{i4} \in \mathcal{H}, \tag{66.122a}$$
$$\mathcal{G}_{(-1,0)} \rightsquigarrow Y_i = -E_{2i} + E_{4i} - E_{i2} - E_{i4}, \tag{66.122b}$$
$$\mathcal{G}_{(0,1)} \rightsquigarrow V_i = E_{1i} + E_{3i} + E_{i1} - E_{i3}, \tag{66.122c}$$
$$\mathcal{G}_{(0,-1)} \rightsquigarrow X_i = -E_{1i} + E_{3i} - E_{i1} - E_{i3} \tag{66.122d}$$

with[7] $i\colon 5 \to n+2$. For example,

$$\mathcal{G}_{(1,0)} \rightsquigarrow W_5 = \begin{pmatrix} & & & & 0 \\ & & & & 1 \\ & & & & 0 \\ & & & & 1 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix} \in \mathcal{H} \qquad \mathcal{G}_{(-1,0)} \rightsquigarrow Y_5 = \begin{pmatrix} & & & & 0 \\ & & & & -1 \\ & & & & 0 \\ & & & & 1 \\ 0 & -1 & 0 & -1 & 0 \end{pmatrix} \in \mathcal{H}$$

$$\mathcal{G}_{(0,1)} \rightsquigarrow V_5 = \begin{pmatrix} & & & & 1 \\ & & & & 0 \\ & & & & 1 \\ & & & & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix} \qquad \mathcal{G}_{(0,-1)} \rightsquigarrow X_5 = \begin{pmatrix} & & & & -1 \\ & & & & 0 \\ & & & & 1 \\ & & & & 0 \\ -1 & 0 & -1 & 0 & 0 \end{pmatrix} \tag{66.123}$$

EqGeueuleVWXY

$$\mathcal{G}_{(1,1)} \rightsquigarrow M = \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{pmatrix} \qquad \mathcal{G}_{(-1,-1)} \rightsquigarrow F = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathcal{G}_{(1,-1)} \rightsquigarrow L = \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix} \qquad \mathcal{G}_{(-1,1)} \rightsquigarrow N = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

These are the same spaces as the previous ones, but the reasoning at page 3953 shows that there must be a difference. The subtlety is that we will choose an other notion of positivity, so that the space $\mathcal{N}$ will be different.

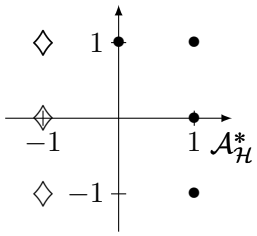


Figure 66.1: The root space

LabelFigHNxitLj

Let us recall the aim of our new decomposition: we want to have $\mathcal{R}_{\mathcal{H}} \subset \mathcal{R}$. For this purpose, the equation (66.98) gives us a constraint on the choice of the positivity notion on $\mathcal{A}^*$. First we must have $\mathcal{G}_{(1,0)} \subset \mathcal{N}$. The upper left 4×4 corner of $\mathcal{N}_{\mathcal{H}}$ is spanned by $\mathcal{G}_{(1,1)} - \mathcal{G}_{(1,-1)}$. We complete our choice of $\mathcal{N}$ with $\mathcal{G}_{(0,1)}$. The underlying notion of positivity is that the element $\alpha(a,b)$ is positive in $\mathcal{A}^*$ when $(a>0) \vee (a=0 \wedge b>0)$.

The difference between decomposition and the previous one is the replacement of N by $L \in \mathcal{G}_{(1,-1)}$. Now

PgTableNEnDimAlg

$$\mathcal{N} = \{W_i, V_j, M, L\} = \{X_{1,0}, X_{0,1}, X_{1,1}, X_{1,-1}\} \tag{66.124a}$$

$$\mathcal{A} = \{J_1, J_2\}, \tag{66.124b}$$

with the commutator table

EqTableSOIwa

$$[V_i, W_j] = \delta_{ij} M \qquad [V_j, L] = 2W_j \tag{66.125a}$$

$$[J_1, W_j] = W_j \qquad [J_2, V_i] = V_i \tag{66.125b}$$

$$[J_1, L] = L \qquad [J_2, L] = -L \tag{66.125c}$$

$$[J_1, M] = M \qquad [J_2, M] = M. \tag{66.125d}$$

7. Let us remember that we are dealing with $\mathrm{SO}(2,n)$ and that AdS_l is a quotient of $\mathrm{SO}(2,l-1)$, so in the case of AdS_l the index j runs from 5 to $l+1$. The first anti de Sitter space which contains such root spaces is AdS_4. More generally, remark that the table (66.125) of $\mathfrak{so}(2,n)$ gives the feeling that if something works with AdS_4, it will work for $AdS_{l\geqslant 4}$.

It is important to note that W_i, $J_1 \in \mathcal{H}$ and $J_2 \in \mathcal{Q}$. On the other hand, the vectors V_i begin to appear in $\mathrm{SO}(2,3)$. The structure of $\mathrm{SO}(2,3)$ is thus slightly different to the one of $\mathrm{SO}(2,2)$, while the structures of $\mathrm{SO}(2,n)$ with $n \geqslant 3$ are more or less all the same. In the terminology of chapter 79, theses special vectors arrive with AdS_4.

The following change of basis in $\mathcal{A}$ reveals to be useful in some circumstances: EqChmHJ

$$H_1 = J_1 - J_2 \qquad H_2 = J_1 + J_2 \tag{66.126a}$$

$$J_1 = \frac{1}{2}(H_1 + H_2) \qquad J_2 = -\frac{1}{2}(H_1 - H_2) \tag{66.126b}$$

which leads to the table TableSeconde

$$[V_i, W_j] = \delta_{ij} M \qquad [V_j, L] = 2W_j \tag{66.127a}$$

$$[H_1, V_i] = -V_i \qquad [H_2, V_i] = V_i \tag{66.127b}$$

$$[H_1, W_i] = W_i \qquad [H_2, W_i] = W_i \tag{66.127c}$$

$$[H_1, L] = 2L \qquad [H_2, M] = 2M. \tag{66.127d}$$

66.5.5 A companion: $A\bar{N}$

That Iwasawa immediately induces a new one by conjugation by the Cartan automorphism θ. That new decomposition will play a central role when we will be interested in black holes on $\mathrm{SO}(2,n)/\mathrm{SO}(1,n)$. See chapter 79, and more precisely the subsection 79.2.3 and the definition 79.17.

The algebras $\mathcal{K}$ and $\mathcal{A}$ are not affected by θ, but the algebra $\mathcal{N}$ changes into a new algebra that we denote by $\bar{\mathcal{N}} = \theta(\mathcal{N})$. The maximal parabolic algebra in that new Iwasawa decomposition is thus given by

$$\begin{aligned} \mathcal{A} &= \{J_1, J_2\} \\ \bar{\mathcal{N}} &= \{X_{-1,0}, X_{0,-1}, X_{-1,1}, X_{-1,-1}\} = \{Y_i, X_i, N, F\}. \end{aligned} \tag{66.128}$$

Indeed, using lemma 66.43, we have

$$\begin{aligned} \theta(L) \in \theta\big(\mathcal{G}_{(1,-1)}\big) \in \mathcal{G}_{(-1,1)} &\rightsquigarrow N, \\ \theta(M) \in \theta\big(\mathcal{G}_{(1,1)}\big) \in \mathcal{G}_{(-1,-1)} &\rightsquigarrow F, \\ \theta(V_i) \in \theta\big(\mathcal{G}_{(0,1)}\big) \in \mathcal{G}_{(0,-1)} &\rightsquigarrow X_i, \\ \theta(W_j) \in \theta\big(\mathcal{G}_{(1,0)}\big) \in \mathcal{G}_{(-1,0)} &\rightsquigarrow Y_j. \end{aligned} \tag{66.129}$$

66.5.6 Second choice

SubSecANbarIwa

Here, we show an other set of choices that can be done in order to get an Iwasawa decomposition. The one that we will create will not be compatible with the Iwasawa decomposition of $\mathrm{SO}(1,n)$.

66.5.6.1 Nilpotent part

We search the eigenvectors and eigenvalues of $\mathrm{ad}(H_p)$ under the form $E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with $A \in M_{4\times 4}$, $B \in M_{(n-2)\times 4}$, $C \in M_{4\times(n-2)}$, $D \in M_{(n-2)\times(n-2)}$. Remark that, thanks to (66.99), the matrix C is completely determined by B:

$$\begin{cases} C^{i1} = B^{1i}, & C^{i2} = B^{2i}, \\ C^{i3} = -B^{3i}, & C^{i4} = -B^{4i}. \end{cases} \tag{66.130}$$ C_de_B

The equation to be solved is

$$(\mathrm{ad}\, H_p)E = \begin{pmatrix} [N_p, A] & N_p B \\ -CN_p & 0 \end{pmatrix} \stackrel{!}{=} \lambda_p E, \tag{66.131}$$ eq_gene

with the notation $\lambda_p = \lambda(H_p)$.

66.5.6.2 Search for two non zero eigenvalues

By "two non zero eigenvalues", we mean a $\lambda \in \mathcal{A}^*$ such that $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$. We immediately find $D = 0$.

The next step is to determine B by the condition $N_p B = \lambda_p B$. We change the range of the indices. Now, a, $b : 1 \to 4$, and i, $j : 5 \to n+2$ and a few computation give $\sum_a N_p^{ca} B^{ai} = \lambda_p B^{ci}$ (with sum over a). Taking successively $c = 1, 2, 3, 4$ and taking into account $\lambda_p \neq 0$, we find: pour_B

$$B^{3i} = \lambda_p B^{1i} \tag{66.132a}$$
$$(-1)^p B^{4i} = \lambda_p B^{2i} \tag{66.132b}$$
$$\lambda_p = \pm 1. \tag{66.132c}$$

We can check that the equations obtained by $-CN_p = \lambda_p C$ are exactly the one that we can find directly using (66.130) and (66.132).

Now, we determine A by the condition $[N_p, A] = \lambda_p A$. We know that A and N_p are 4×4 matrices. Again, we redefine the range of the indices: $a = 1, 2$ and $i = 3, 4$. Symmetry properties of A are $A^{ai} = A^{ia}$, $A^{ij} = -A^{ji}$ and $A^{ab} = -A^{ba}$, so that

$$A = A^{12}(E_{12} - E_{21}) + A^{ai}(E_{ai} + E_{ia}) + A^{34}(E_{34} - E_{43})$$

Using equation (66.110), a quite tedious (but direct) computation give the following for $[N_p, A]$:

$$\begin{pmatrix} 0 & A^{23} - (-1)^p A^{14} & 0 & A^{34} - (-1)^p A^{12} \\ -A^{23} + (-1)^p A^{14} & 0 & A^{12} - (-1)^p A^{34} & 0 \\ 0 & A^{12} - (-1)^p A^{34} & 0 & A^{14} - (-1)^p A^{23} \\ A^{34} - (-1)^p A^{12} & 0 & A^{14} - (-1)^p A^{23} & 0 \end{pmatrix} \tag{66.133}$$ grosse_A

which has to be equated to $\lambda_p A$. We immediately have $A^{13} = A^{24} = A^{31} = A^{42} = 0$. The others conditions are: pour_A

$$\lambda_p A^{12} = A^{23} - (-1)^p A^{14} \tag{66.134a}$$
$$\lambda_p A^{14} = A^{34} - (-1)^p A^{12} \tag{66.134b}$$

Since we are in the case $\lambda_p \neq 0$, using the fact that $\lambda_p = \pm 1$, we easily find $A = 0$. Now, we define $\lambda_{\pm\pm} \in \mathcal{A}^*$ by

$$\begin{aligned} \lambda_{++}(H_1) &= 1 \quad & \lambda_{++}(H_2) &= 1, \\ \lambda_{+-}(H_1) &= 1 \quad & \lambda_{+-}(H_2) &= -1, \\ \lambda_{-+}(H_1) &= -1 \quad & \lambda_{-+}(H_2) &= 1, \\ \lambda_{--}(H_1) &= -1 \quad & \lambda_{--}(H_2) &= -1. \end{aligned} \tag{66.135}$$

Root spaces are (with $i : 5 \to n+2$):

$$\begin{aligned} \lambda_{++} &\rightsquigarrow V_i = E_{3i} + E_{1i} + E_{i1} - E_{i3}, \\ \lambda_{-+} &\rightsquigarrow W_i = E_{4i} + E_{2i} - E_{i4} + E_{i2}, \\ \lambda_{--} &\rightsquigarrow X_i = E_{3i} - E_{1i} - E_{i1} - E_{i3}, \\ \lambda_{+-} &\rightsquigarrow Y_i = E_{4i} - E_{2i} - E_{i4} - E_{i2}. \end{aligned} \tag{66.136}$$

66.5.6.3 Search for eigenvalues with one zero

We denote by $\lambda(a, b)$ the element of $\mathcal{A}^*$ defined by $\lambda(x_1, x_2)(H_i) = x_i$. Equations (66.133) and (66.134) with for example $a = 0$ and $b \neq 0$ give $\lambda_2 = \pm 2$. Serious computation give: EqsTableRacinesSOdeuxn

$$\lambda(0,-2) \rightsquigarrow F = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \lambda(2,0) \rightsquigarrow N = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \tag{66.137a}$$

$$\lambda(0,2) \rightsquigarrow M = \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{pmatrix}, \quad \lambda(-2,0) \rightsquigarrow L = \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix}. \tag{66.137b}$$

In these expressions, we only wrote the upper left part of the matrices which are zero everywhere else.

66.5.6.4 Two vanishing eigenvalues

We now search for a matrix E of $\mathfrak{so}(2,n)$ such that $(\operatorname{ad} H_p)E = 0$ for $p = 1, 2$. Taking a look at (66.131), we see that D has no more constraints (apart the usual symmetries). The equation (66.133) gives us $A^{12} = A^{23} = A^{14} = A^{34} = 0$, but A^{13}, A^{24},A^{31},A^{42} are free. Therefore we can write:

$$\mathcal{G}_{\lambda(0,0)} = \{x(E_{13} + E_{31}) + y(E_{42} + E_{24}) + \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & D \end{array}\right)\}. \tag{66.138}$$

Remark that $\mathcal{G}_{\lambda(0,0)} \cap \mathcal{P}$ is spanned by matrices of the form

$$\begin{pmatrix} 0 & 0 & x & 0 \\ 0 & 0 & 0 & y \\ x & 0 & 0 & 0 \\ 0 & y & 0 & 0 \end{pmatrix},$$

so that $\mathcal{G}_{\lambda(0,0)} \cap \mathcal{P} = \mathcal{A}$. This is a simple consequence of the very definition of $\mathcal{A}$ as maximal abelian subalgebra of $\mathcal{P}$.

66.5.6.5 Choice of positivity

We are now able to write down the component $\mathcal{N}$. We just have to "elect" some $\mathcal{G}_{\lambda_{(a,b)}}$ with a notion of positivity. Our choice is:

$$\mathcal{N} = \{V_i, W_i, M, N\}, \tag{66.139}$$

with $i,\ j : 5 \to n+2$. A basis of $\mathcal{K}$ is given by $K_{rs} = E_{rs} - E_{sr}$, and $K_a = E_{12} - E_{21}$ with $r, s \colon 3 \to n+2$. The commutator with $\mathcal{A}$ are:

$$\begin{aligned} [H_p, K_a] &= -(-1)^p(E_{41} + E_{14}) + E_{32} + E_{23} \\ [H_p, K_{rs}] &= \delta_{r3}(E_{1s} + E_{s1}) - \delta_{s3}(E_{1r} + E_{r1}) \\ &\quad + (-1)^p(\delta_{4r}(E_{2s} + E_{s2}) - \delta_{4s}(E_{2r} + E_{r2})) \end{aligned} \tag{66.140}$$

In the decomposition $\mathcal{G} = \mathcal{A} \oplus \mathcal{N} \oplus \mathcal{K}$, the matrix $E_{32} + E_{23}$ is obtained by

$$\frac{1}{2}(\lambda(0,2) + \lambda(2,0)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where the upper left corner $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ can be removed by use of K_a:

$$E_{32} + E_{23} = \frac{1}{2}(\lambda(0,2) + \lambda(2,0)) - K_a$$

In the same way,

$$E_{41} + E_{14} = \frac{1}{2}(\lambda(2,0) - \lambda(0,2)) - K_{34}$$

We also have:

$$\begin{aligned} E_{1i} - E_{i1} &= V_i - K_{3i} \\ E_{2i} + E_{i2} &= W_i - K_{4i}. \end{aligned} \tag{66.141}$$

Let us summarize the result obtained. The Iwasawa decomposition of $\mathrm{SO}(2,n)$ is given by eq:Iwasawa_explicite

$$\mathcal{A} = \{H_1, H_2\} \tag{66.142a}$$
$$\mathcal{N} = \{V_i, W_j, M, N\} \tag{66.142b}$$
$$\mathcal{K} = \{\begin{pmatrix} a & 0 \\ 0 & B \end{pmatrix}\}. \tag{66.142c}$$

with $i, j : 5 \to n+2$, $a \in M_{2\times 2}$, $B \in M_{n\times n}$ skew-symmetric, and $H_p = E_{13} + E_{31} + (-1)^p(E_{24} + E_{42})$. pg:exp_AN

We are going to write down a general element of AN (forgetting for the moment W_i and V_j). Since $[H_1, H_2] = [M, N] = 0$, the only product to be considered is $e^{tH_1}e^{uH_2}e^{aM}e^{bM}$. Our first task is to exponentiate the matrices of $\mathcal{A}$ and $\mathcal{N}$. The first is quite easy: since $H_p^2 = \mathbb{1}$,

$$e^{tH_p} = \sum_k \frac{t^k}{k!}\,(\mathbb{1}_{\text{if } k\text{is even}}, H_{p\text{ if } k\text{ is odd}}) = (\cosh t)\mathbb{1} + (\sinh t)H_p;$$

then

$$e^{tH_1}e^{uH_2} = \begin{pmatrix} \cosh\xi & 0 & \sinh\xi & 0 \\ 0 & \cosh\eta & 0 & \sinh\eta \\ \sinh\xi & 0 & \cosh\xi & 0 \\ 0 & \sinh\eta & 0 & \cosh\eta \end{pmatrix} \tag{66.143}$$

with the change of variable eq:chm_xi_eta

$$\xi = t + u, \tag{66.144a}$$
$$\eta = u - t. \tag{66.144b}$$

For the second, we remark that $e^{aM} = \mathbb{1} + aM$ and $e^{bN} = \mathbb{1} + bN$ because $M^2 = N^2 = 0$. The result is:

$$e^{aM}e^{bN} = \begin{pmatrix} 1-2ab & b+a & 2ab & b-a \\ -b-a & 1 & b+a & 0 \\ -2ab & b+a & 1+2ab & b-a \\ b-a & 0 & a-b & 1 \end{pmatrix}. \tag{66.145}$$

Remark that, by construction, matrices of $\mathcal{N}$ are nilpotent. But a sum of nilpotent matrices is nilpotent and exponential on nilpotent matrices is easy to compute. Hence there are no obstructions to compute an explicit general matrix of AN, this is done at the page 3950.

Now, we compute some commutators of the different matrices that we had seen.

$$[V_i, W_j] = \delta_{ij}A_{(0,2)} \tag{66.146a}$$
$$[V_i, X_j] = 2\delta_{ij}A_{(0,0)}(-1,0) + 2(E_{ij} - E_{ji}) \tag{66.146b}$$
$$[X_i, Y_j] = \delta_{ij}\lambda(0,-2) \tag{66.146c}$$
$$[W_i, X_j] = \delta_{ij}\lambda(-2,0) \tag{66.146d}$$
$$[V_i, Y_j] = \delta_{ij}\lambda(2,0) \tag{66.146e}$$
$$[W_i, Y_j] = -2\delta_{ij}A_{(0,0)}(0,1) \tag{66.146f}$$
$$[V_i, V_j] = 0 \tag{66.146g}$$
$$[\lambda(0,2), \lambda(0,-2)] = 4A_{(0,0)}(-1,-1) \tag{66.146h}$$
$$[\lambda(2,0), \lambda(-2,0)] = 4A_{(0,0)}(-1,1) \tag{66.146i}$$

$$\begin{aligned} [H_p, K_{rs}] = {} & \delta_{r3}(V_s - K_{3s}) - \delta_{s3}(V_r - K_{3r}) \\ & + (-1)^p[\delta_{4r}(W_s - K_{4s}) - \delta_{4s}(W_r - K_{4r})], \end{aligned} \tag{66.147a}$$

$$[H_p, K_a] = (-1)^{p+1}(\frac{1}{2}(N - M) - K_{34}) + \frac{1}{2}(N + M) - K_a. \tag{66.147b}$$

The non-zero commutators in $\mathcal{A} \oplus \mathcal{N}$ are: TabelPrem

$$[V_i, W_j] = \delta_{ij} M \qquad [W_i, N] = -2V_i \tag{66.148a}$$
$$[H_1, V_i] = -V_i \qquad [H_2, V_i] = V_i \tag{66.148b}$$
$$[H_1, N] = 2N \qquad [H_2, M] = 2M \tag{66.148c}$$
$$[H_1, W_i] = W_i \qquad [H_2, W_i] = W_i. \tag{66.148d}$$

66.5.7 Application: an isomorphism of Lie algebra

sssIsomsoslplussl

Root spaces are a powerful tool in proving isomorphisms of Lie algebras. Comparing the tables of $\mathfrak{sl}(2, \mathbb{R})$ and $\mathfrak{so}(2, 2)$ given in 53.329 and 66.137, it is easy to prove the following.

Proposition 66.46.
We have

$$\mathfrak{so}(2, 2) = \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{sl}(2, \mathbb{R}) \tag{66.149}$$

as isomorphism of Lie algebras.

Proof. The algebra $\mathfrak{so}(2, 2)$ has the following root spaces:

$$\lambda(0, 2) \qquad \lambda(2, 0)$$
$$\lambda(0, -2) \qquad \lambda(-2, 0)$$

with $\mathcal{A} = \{H_1, H_2\}$. So, by construction the vector space $\mathfrak{g}_1$ generated by $\{H_1, \lambda(2, 0), \lambda(-2, 0)\}$ is a Lie subalgebra of $\mathfrak{so}(2, 2)$ with roots equal to 2 and -2; the corresponding root spaces being one dimensional. Looking on the root spaces of $\mathfrak{sl}(2, \mathbb{R})$, we see that they are exactly the same. Thus $\mathfrak{g}_1 = \mathfrak{sl}(2, \mathbb{R})$. The same is true for $\mathfrak{g}_2 = \{H_2, \lambda(0, 2), \lambda(0, -2)\}$ and we also have $[\mathfrak{g}_1, \mathfrak{g}_2] = 0$ because of property (2) of proposition 53.129. Thus we have

$$\mathfrak{so}(2, 2) = \mathfrak{g}_1 \oplus \mathfrak{g}_2 = \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{sl}(2, \mathbb{R}),$$

as announced. □

66.6 Openness of orbits in homogeneous spaces

pSecDNqdsJOp

Proposition 66.47.
The orbits of AN are submanifolds of G/H.

Proof. Indeed proposition 58.7 makes $R/(R \cap H)$ the orbit of $\pi(e)$ by R and assure us that it is a submanifold of G/H. That proves the proposition for the orbit of e.

For the other orbits, we consider the group $R_z = \mathbf{Ad}(z^{-1})R$ which is also a Lie subgroup of G. The space $R_z/(R_z \cap H)$ is isomorphic to the orbit of $\pi(e)$ under the action of R_z. Therefore $zR[z^{-1}]$ is a submanifold of G/H and the very definition of a Lie group makes that $R[z^{-1}]$ is a submanifold too.

□

tho:pr_ouvert

Theorem 66.48.
If R is a subgroup of G with Lie algebra $\mathcal{R}$, then the orbit $R \cdot \vartheta$ is open in G/H if and only if the projection $\mathtt{proj}\colon \mathcal{R} \to \mathcal{Q}$ parallel to $\mathcal{H}$ is surjective.

The projection is defined by $\mathtt{proj}(X) = X_{\mathcal{Q}}$ if $X = X_{\mathcal{Q}} + X_{\mathcal{H}}$ is the decomposition of $X \in \mathcal{G}$ with respect to the decomposition $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$. We need two lemmas before to prove the theorem.

Lemma 66.49.
The orbit $R \cdot \vartheta$ is open if and only if

$$\mathrm{Span}\{X^*_\vartheta | X \in \mathcal{R}\} = T_\vartheta M$$

where X^* is the fundamental field defined by equation (66.5). lem:equiv_1

Proof. From general theory of fundamental fields (lemma 53.19) we know that

$$\mathrm{Span}\{X^*_\vartheta | X \in \mathcal{G}\} = T_\vartheta M.$$

The game is now to prove that one can replace $\mathcal{G}$ by $\mathcal{R}$ if and only if $R \cdot \vartheta$ is open.
Necessary condition. If $R \cdot \vartheta$ is open, we have a neighbourhood of ϑ which is contained in $R \cdot \vartheta$. Then for any $X \in \mathcal{G}$, and for a small enough t, the element $e^{-tX} \cdot \vartheta$ belongs to $R \cdot \vartheta$. Hence we have a path $r_X(t)$ in R such that $e^{-tX} \cdot \vartheta = r_X(t) \cdot \vartheta$:

$$\frac{d}{dt}\Big[e^{-tX} \cdot \vartheta\Big]_{t=0} = \frac{d}{dt}\Big[r_X(t) \cdot \vartheta\Big]_{t=0}.$$

Since $r_X(t)$ is a path in R, we can replace it by a e^{-tY} with a $Y \in \mathcal{R}$ in the derivative. For this Y, we have $X^*_\vartheta = Y^*_\vartheta$.
Sufficient condition. We have $\dim(R \cdot \vartheta) = \dim \mathrm{Span}\{X^*_\vartheta \text{ tel que } X \in \mathcal{R}\} = \dim T_\vartheta M$, so $R \cdot \vartheta$ has the same dimension as M. The conclusion follows from the fact that a submanifold is open if and only if it has maximal dimension.

□

Lemma 66.50.
The canonical projection is surjective from $\mathcal{R}$ to the tangent space to identity:

$$\mathrm{Span}\{X^*_\vartheta | X \in \mathcal{R}\} = d\pi_e(\mathcal{R}). \tag{66.150}$$ eq:equiv_2

XsdpiR

Proof. Consider the following computation when $X \in \mathcal{R} = T_eR$ is given by the path $X(t) = e^{tX}$:

$$d\pi_e X = \frac{d}{dt}\Big[[X(t)]\Big]_{t=0} = \frac{d}{dt}\Big[e^{tX}\vartheta\Big]_{t=0} = Y^*_\vartheta \tag{66.151}$$

with $Y = -X$. Reading these lines from left to right shows that $d\pi_e(\mathcal{R}) \subseteq \{X^*_\vartheta : X \in \mathcal{R}\}$ while reading it from right to left shows the inverse inclusion. □

We are now able to prove the theorem.

Proof of theorem 66.48. From lemma 66.49 and lemma 66.50, the orbit $R \cdot \vartheta$ is open if and only if $d\pi_e : \mathcal{R} \to T_\vartheta M$ is surjective. On the one hand any $X \in \mathcal{R}$ can uniquely be written as $X = X_\mathcal{H} + X_\mathcal{Q}$ with $X_\mathcal{H} \in \mathcal{H}$ and $X_\mathcal{Q} \in \mathcal{Q}$. On the other hand it is clear that $d\pi_e X_\mathcal{H} = 0$, thus $R \cdot \vartheta$ is open if and only if $d\pi_e : \mathtt{proj}_\mathcal{Q}\,\mathcal{R} \to T_\vartheta M$ is surjective.

Now, recall that $d\pi_e$ is surjective from $\mathcal{G}$, hence it is surjective from $\mathcal{Q}$. The first conclusion is that if $\mathtt{proj} : \mathcal{R} \to \mathcal{Q}$ is surjective, then $R \cdot \vartheta$ is open. The inverse implication remains to be proved.

We know that openness $R \cdot \vartheta$ implies that $d\pi_e : \mathtt{proj}_\mathcal{Q}\,\mathcal{R} \to T_\vartheta M$ is bijective (surjective because $R \cdot \vartheta$ is open and injective because $d\pi_e : \mathcal{Q} \to T_\vartheta M$ is injective by lemma 66.5). From all that, one concludes that $\mathtt{proj}_\mathcal{Q}\,\mathcal{R} = \mathcal{Q}$. Indeed, suppose that $X_\mathcal{Q} \in \mathcal{Q}$ and $X_\mathcal{Q} \notin \mathtt{proj}_\mathcal{Q}\,\mathcal{R}$. Since $d\pi_e : \mathtt{proj}_\mathcal{Q}\,\mathcal{R} \to T_\vartheta M$ is surjective, there exists a $X'_\mathcal{Q} \in \mathtt{proj}_\mathcal{Q}\,\mathcal{R}$ such that $d\pi_e X'_\mathcal{Q} = d\pi_e X'_\mathcal{Q}$. This is impossible because $d\pi_e$ is injective from the whole $\mathcal{Q}$.

□

66.7 Symplectic symmetric spaces

Definition 66.51.
*A **symplectic symmetric** space is a triple (M, s, ω) where (M, s) is a symmetric space, (M, ω) is a symplectic space such that $s_x^*\omega = \omega$ for every $x \in M$.*

Remarque 66.52.
We can weaken the symplectic condition in the definition and only ask for ω to be non degenerate because the condition $s_x^*\omega = \omega$ implies $d\omega = 0$.

66.7.1 Example

Let $G = \mathrm{SL}(2, \mathbb{R})$ and look at the coadjoint action $\mathrm{Ad}^*\colon G \to \mathrm{GL}(\mathfrak{g}^*)$. We consider the element $Z = E - F$ and the orbit

$$\mathcal{O} = \mathrm{Ad}^*(G)(Z^\flat). \tag{66.152}$$

The space $\mathfrak{g}^*$ has the metric

$$\langle X^\flat, Y^\flat \rangle = \beta(X, Y) \tag{66.153}$$

Let us consider $\vartheta = Z^\flat \in \mathfrak{g}^*$ and consider the stabilizer:

$$\mathrm{Stab}_\vartheta(\mathcal{O}) = \{g \in G \text{ tel que } \mathrm{Ad}^*(g)Z^\flat = Z^\flat\}. \tag{66.154}$$

The Lie algebra is given by

$$\mathfrak{Stab}_\vartheta(\mathcal{O}) = \{X \in \mathfrak{g} \text{ tel que } Z^\flat \circ \mathrm{ad}(X) = 0\}. \tag{66.155}$$

The condition of the Lie algebra reads

$$0 = \langle Z^\flat, [X, Y] \rangle = \beta(Z, [X, Y]) = -\beta\big([X, Z], Y\big) \tag{66.156}$$

for every Y, which implies $[X, Z] = 0$ because β is nondegenerate. Now, in $\mathfrak{sl}(2, \mathbb{R})$, the only possibility is that X is proportional to Z. Thus the Lie algebra reduces to $\mathbb{R}Z$ in fact.

66.7.2 Algebraic setting

We want now to encode the symplectic space structure in an algebraic data. What we are going to discover is the notion of symplectic triple that will be developed in section 66.7.3.

Let (G, σ) be an involutive Lie group and H a closed subgroup of G such that

$$G_0^\sigma \subset H \subset G^\sigma. \tag{66.157}$$

Let π be the projection $\pi\colon G \to M = G/H$.

The symmetry on the quotient G/H is given by the theorem 66.17:

$$s_{[g]}[g'] = \big[\sigma(g^{-1}g')\big] \tag{66.158}$$ EaSymGH

Now, if we denote by σ the differential $d\sigma_e$, we can decompose the Lie algebra $\mathcal{G}$ into $\mathcal{G} = \mathcal{H} \oplus \mathcal{P}$ and we have the isomorphism (see lemma 66.5)

$$d\pi_e|_{\mathcal{P}}\colon \mathcal{P} \to T_\vartheta(M) \tag{66.159}$$

where $\vartheta = [e]$. Thus we can see the form ω_ϑ on $\mathcal{P}$ by

$$\Omega = \big(d\pi_e|_{\mathcal{P}}\big)^*\omega_\vartheta \tag{66.160}$$

and the space $(\mathcal{P}, \Omega)$ becomes a symplectic vector space.

Lemma 66.53.
We have

(1) the space $\mathcal{K} = [\mathcal{P}, \mathcal{P}]$ is a Lie subalgebra of $\mathcal{H}$,

(2) the adjoint action of $\mathcal{K}$ over $\mathcal{P}$ preserves the symplectic form, i.e.

$$\Omega\big([Z, X], Y\big) + \Omega\big(X, [Z, Y]\big) = 0 \tag{66.161}$$

Proof. Sketch of the proof.

Let $x_j, y_j \in \mathcal{P}$. Using the Jacobi identity on the nested commutator $\big[[x_1, y_1], [x_2, y_2]\big]$ and the facts that $[\mathcal{P}, \mathcal{P}] \subset \mathcal{H}$ and $[\mathcal{H}, \mathcal{P}] \subset \mathcal{P}$, we find the commutator of two elements of $[\mathcal{P}, \mathcal{P}]$ belongs to $[\mathcal{P}, \mathcal{P}]$.

First we consider $\mathcal{G}^{(M)} = \mathcal{K} \oplus \mathcal{P}$ and $G(M)$, the associated Lie group. Then we have

$$M \simeq G(M)/K. \tag{66.162}$$

Now one can see that the group $G(M)$ is generated by the products $\{s_\vartheta s_x\}$ with $x \in M$. Indeed let $X \in \mathcal{P}$ and look at $\exp(\mathcal{P})$ as map on M. Using the symmetry (66.158) and the fact that $\sigma e^{X/2} = e^{-X/2}$, we have

$$s_{\exp(X/2)} \cdot {}_\vartheta\vartheta = e^{X/2}\big[\sigma[e^{-X/2}]\big] = [e^X] = e^X \cdot \vartheta. \tag{66.163}$$

If we act on an other point than ϑ, we have

$$\begin{aligned} s_{[e^{X/2}]}[g] &= e^{X/2}\big[\sigma(e^{-X/2}g)\big] \\ &= e^{X/2}\big[e^{X/2}\sigma(g)\big] \\ &= e^X s_\vartheta[g] \end{aligned} \tag{66.164}$$

because $\big[\sigma(g)\big] = s_\vartheta[g]$.

Now, using the lemma 53.75, the fact that the elements e^X with $X \in \mathcal{P}$ generate $e^{\mathcal{P}}$ in G implies that it also generate the elements of the form $e^{[\mathcal{P},\mathcal{P}]}$ and then the whole $G(M)$. Since the elements $e^{\mathcal{P}}$ are of the form $s_x s_\vartheta$, we conclude that $G(M)$ is generated by the products $s_x s_\vartheta$.

Thus we have $g^*\omega = \omega$ for every $g \in G(M)$ because ω is preserved by all the symmetries. □

66.7.3 Symplectic triple

SubSecTripleSylple

A **symplectic triple** is the data of the triple $(\mathcal{G}, \sigma, \Omega)$ where $(\mathcal{G}, \sigma)$ is an involutive Lie algebra and Ω is a $\mathcal{K}$-invariant nondegenerate 2-form $\Omega \in \Lambda^2(\mathcal{P}^*)$. The $\mathcal{K}$ invariance means that for every $Z \in \mathcal{K}$ and $X, Y \in \mathcal{P}$,

$$\Omega\big([Z, X], Y\big) + \Omega\big(X, [Z, Y]\big) = 0. \tag{66.165}$$

A symplectic triple is the infinitesimal version of a symplectic symmetric space. The following more abstract version of the definition comes from [961]:

Definition 66.54.
*The triple $(\mathcal{G}, \sigma, \Omega)$ is a **symplectic triple** when $\Omega \in \Lambda^2\mathcal{G}$ and*

(1) If $\mathcal{G} = \mathcal{K} \oplus \mathcal{P}$ is the decomposition of $\mathcal{G}$ into eigenspaces of σ, then $[\mathcal{P}, \mathcal{P}] = \mathcal{K}$ and the adjoint representation of $\mathcal{K}$ on $\mathcal{P}$ is faithful. ($\mathcal{K}$ is the eigenspaces with eigenvalue $+1$ of σ while $\mathcal{P}$ is the one of -1)

(2) The 2-form Ω is a Chevalley 2-cocycle for the trivial representation of $\mathcal{G}$ on $\mathbb{R}$.

(3) $i(\mathcal{K})\Omega = 0$ and $\Omega|_{\mathcal{P}\times\mathcal{P}}$ is nondegenerate.

Notice that $[\mathcal{P}, \mathcal{P}] \subset \mathcal{K}$ is automatic from the definition of $\mathcal{P}$ and $\mathcal{K}$ as eigenspaces of σ; the hypothesis is the equality.

Let us now see how one build a symplectic symmetric space from the data of the symplectic triple $(\mathcal{G}, \sigma, \Omega)$. First we consider G, the group associated with $\mathcal{G}$ and $M = G/K$ with the left invariant form ω build on Ω.

66.7.4 Example on the Heisenberg algebra

Let $\mathscr{H} = V \oplus \mathbb{R}E$ be the Heisenberg algebra of (V, Ω^0), and consider the derivation

$$D = \mathrm{Id}\,|_V \oplus (2\,\mathrm{Id})|_{\mathbb{R}E}. \tag{66.166}$$

If we consider the algebra $\mathcal{A} = \mathbb{R}H$, we build the semi direct product

$$\mathcal{S} = \mathcal{A} \ltimes_D \mathscr{H} \tag{66.167}$$

with the definition $[H, x] = D(x)$ when $x \in \mathscr{H}$.

An other split extension that can be done is

$$\mathcal{G}_0 = \mathcal{A} \ltimes_\rho (\mathscr{H} \oplus \mathscr{H}) \tag{66.168}$$

with $\rho = D \oplus (-D)$. The algebra $\mathcal{G}_0$ is to be endowed with a symplectic triple structure. We define $\sigma_0 \colon \mathcal{G}_0 \to \mathcal{G}_0$

$$\begin{aligned} \sigma_0(x, y) &= (y, x) \quad \in \mathscr{H} \oplus \mathscr{H} \\ \sigma_0(H) &= -H \end{aligned} \tag{66.169}$$

and $(\mathcal{G}_0, \sigma_0)$ is an involutive automorphism. Indeed, we have

$$\sigma[H, x \oplus y] = \sigma\big(Dx \oplus (-Dy)\big) = -Dy \oplus Dx, \tag{66.170}$$

while

$$\big[\sigma H, \sigma(x \oplus y)\big] = [-H, y \oplus x] = -Dy \oplus Dx. \tag{66.171}$$

Let us take the notation

$$W_\pm = \{(w, \pm w)\}_{w \in W}. \tag{66.172}$$

EqDefNitWpm

If we decompose $\mathcal{G}_0 = \mathcal{K}_0 \oplus \mathcal{P}_0$, we have

$$\begin{aligned} \mathcal{K}_0 &= \mathscr{H}_+ \\ \mathcal{P}_0 &= \mathcal{A} \oplus \mathscr{H}_-. \end{aligned} \tag{66.173}$$

We have $H \in \mathcal{P}$ $x \oplus (-x) \in \mathcal{P}$ and $x \oplus x \in \mathcal{K}$.

In fact we have an identification between $\mathcal{S}$ and $\mathcal{P}_0$ by

$$\begin{aligned} \mathcal{S} &\to \mathcal{P}_0 \\ a + x &\mapsto a + x_- \end{aligned} \tag{66.174}$$

where $x_\pm = \frac{1}{2}(x, \pm x)$. Using the notation (66.172), we write

$$\begin{aligned} \mathcal{K} &= \mathscr{H}_+ \\ \mathcal{P} &= \mathcal{A} \oplus \mathscr{H}_-. \end{aligned} \tag{66.175}$$

Under that identification we have le following lemma.

Lemma 66.55.
We have

$$\Lambda^2(\mathcal{P}_0^*) \simeq \Lambda^2(\mathcal{S}^*) \tag{66.176}$$

and if we define

$$\Omega^{\mathcal{S}}(a + x, a' + x') = \Omega(a + x_-, a' + x'_-), \tag{66.177}$$

we have $\Omega \in \Lambda^2(\mathcal{P}_0^*)$ and it is $\mathcal{K}_0$ invariant if and only if the two conditions

$$\begin{aligned} \Omega^{\mathcal{S}}(E, \mathscr{H}) &= 0 \\ \Omega^{\mathcal{S}}|_{V \times V} &= \frac{1}{2}\Omega^{\mathcal{S}}(H, E)\Omega^0. \end{aligned} \tag{66.178}$$

hold.

Proof. Let us write down the condition of $\mathcal{K}$-invariance of the symplectic form

$$\Omega\big([x_+, a+y_-], a'+y'_-\big) + \Omega\big(a+y_-, [x_+, a'+y'_-]\big) = 0. \tag{66.179}$$

If we develop x_+ and y_-, the commutator in the first term becomes

$$\big[\frac{1}{2}(x,x), a+\frac{1}{2}(y,-y)\big] = -\frac{a}{2}(Dx,-Dx) + \frac{1}{4}\big([x,y],-[x,y]\big). \tag{66.180}$$

The first term is rewritten as $[x,a]_-$, while the second term is $\frac{1}{2}[x,y]_-$. The sum is then $[x, a+\frac{1}{2}y]_-$. Looking at the definition of $\Omega^{\mathcal{S}}$, the invariance condition reads

$$\begin{aligned} 0 &= \Omega\big([x, a+\frac{1}{2}y], a'+y'\big) + \Omega^{\mathcal{S}}\big(a+y, [x, a'+\frac{1}{2}y']\big) \\ &= \Omega^{\mathcal{S}}\big(-aDx + \frac{1}{2}\Omega_0(x,y)E, a'+y'\big) + \Omega^{\mathcal{S}}\big(a+y, -a'Dx + \frac{1}{2}\Omega_0(x_V, y'_V)E\big). \end{aligned} \tag{66.181}$$

If we look at that condition with $a'=0$, $a=1$, $x_V=0$ and $y=0$ and taking into account $Dx = x_V + 2x_E E$ we find

$$\Omega^{\mathcal{S}}(2x_E E, y') = 0, \tag{66.182}$$

so that $\Omega^{\mathcal{S}}(E, y')$. We conclude that a necessary condition for the invariance is

$$\Omega^{\mathcal{Z}}(E, \mathscr{H}) = 0. \tag{66.183}$$

Now if we consider $y' \in V$, $x_E = 0$ and $x_V \neq 0$, we find

$$\Omega^{\mathcal{S}}|_{V\times V} = \frac{1}{2}\Omega^{\mathcal{S}}(H,E)\Omega_0. \tag{66.184}$$

One can check that these two conditions insure the $\mathcal{K}$-invariance of $\Omega^{\mathcal{S}}$. □

To each non degenerate form satisfying these two conditions corresponds a symplectic triple $(\mathcal{G}_0, \sigma_0, \Omega)$.

66.7.5 Realization as coadjoint orbit

Let $(M = G/H, s, \omega)$ be a symmetric symplectic space. We are going to study under which conditions we can realise M as a coadjoint orbit, i.e. we want the two conditions

(1) there exists a $\xi_0 \in \mathcal{G}^*$ such that $\mathrm{Stab}_G(\xi_0) = H$ where Stab stands for the stabilizer for the coadjoint action of G. Let

$$\begin{aligned} \Phi\colon M &\to \mathrm{Ad}^*(G)\xi_0 = \mathcal{O} \\ [g] &\mapsto \mathrm{Ad}^*(g)\xi_0 \end{aligned} \tag{66.185}$$

be the identification between M and the coadjoint orbit $\mathcal{O}$.

(2) The identification Φ fits the symplectic structures:

$$\Phi^*\omega^{\mathcal{O}} = \omega \tag{66.186}$$

where $\omega^{\mathcal{O}}$ is the canonical symplectic structure on the coadjoint orbit given by (59.14).

Definition 66.56.
*A **good polarization** associated to ξ_0 is a Lie subalgebra $\mathcal{B}$ of $\mathcal{G}$ which is maximal for the property $\delta\xi_0|_{\mathcal{B}\times\mathcal{B}} \equiv 0$ where the alternate bilinear 2-form $\delta\xi_0$ on $\mathcal{G}$ is defined by*

$$\delta\xi_0 = \langle \xi_0, [.,.] \rangle. \tag{66.187}$$

If $\mathcal{B}$ is a good polarization, we consider $B = \exp(\mathcal{B})$ and we have a representation $\chi\colon \mathcal{B} \to \mathrm{U}(1)$ given by

$$\begin{aligned} \chi\colon \mathcal{B} &\to \mathrm{U}(1) \\ \exp(y) &\mapsto e^{i\langle \xi_0, y\rangle}. \end{aligned} \tag{66.188}$$

It turns out that χ is a representation even when $\mathcal{B}$ is non abelian. Indeed, if $x, y \in \mathcal{B}$, we have

$$\chi(e^x e^y) = \chi(e^{x+y+W}) = e^{i\langle \xi_0, x+y+W\rangle} = e^{i\langle \xi_0, x+y\rangle} = \chi(e^x)\chi(e^y) \tag{66.189}$$

where W is a combination if commutators of x and y (Campbell-Baker-Hausdorff[8]) so that by definition of $\mathcal{B}$, $\langle \xi_0, W\rangle = 0$.

Since we are in the hypothesis (see subsection 60.8.2), we can define the induced unitary representation

$$U\colon G \to \mathrm{U}(\mathscr{H}_\chi) \tag{66.190}$$

where $\mathscr{H}_\chi = L^2(Q, dq)$. Let dg be the left invariant Haar measure on G. To each $u \in L^1(G, dg)$, we make correspond an operator $U(u)$ on $\mathscr{H}_\chi$ given by

$$\langle U(u)\varphi, \psi\rangle = \int_G u(g)\langle U(g)\varphi, \psi\rangle dg \tag{66.191}$$ EqDefUudansHb

for every $\varphi, \psi \in \mathscr{H}_\chi$. Let us prove that this integral exists. We have

$$|\langle U(g)\varphi, \psi\rangle| \leqslant \int_G |u(g)||\langle U(g)\varphi, \psi\rangle| dg, \tag{66.192}$$ EqIntdefUgGdgvvpsi

but the Cauchy-Schwarz inequality shows that $|\langle U(g)\varphi, \psi\rangle| \leqslant |U(g)\varphi||\psi| = |\varphi||\psi|$, so that the integral in (66.192) is smaller than

$$|\varphi||\psi| \int_G |u(g)| dg \tag{66.193}$$

which exists because we supposed $u \in L^1(G, dg)$.

¡¡ Avertissement/question à la lectrice !! 66.57
The following paragraph can be more precise.

We can rewrite the definition (66.191) using the measure theory given around section 61.2. Indeed the space $\mathfrak{B}(\mathscr{H})$ of bounded operators[9] on the Hilbert space $\mathscr{H}$ is endowed with the operator norm for which $\mathfrak{B}(\mathscr{H})$ becomes a normed algebra ($\|AB\|_{op} \leqslant \|A\|_{op}\|B\|_{op}$). The unitary group $\mathrm{U}(\mathscr{H})$ is a subalgebra (because it is closed for the composition), so that one can consider, for each function u, the function

$$\begin{aligned} G &\to \mathfrak{B}(\mathscr{H}) \\ g &\mapsto u(g)U(g) \end{aligned} \tag{66.194}$$

and its integral

$$\int_G u(g)U(g)dg \tag{66.195}$$

which is an element in $\mathfrak{B}(\mathscr{H})$. This integral is well defined in $\mathfrak{B}(\mathscr{H}_\chi)$ because

$$\| \int_G u(g)U(g)dg\|_{op} \leqslant \int_G |u(g)| \cdot \|U(g)\|_{op} dg = \|u\|_{L^1}. \tag{66.196}$$

Using the measure theory, one can prove that

$$\langle U(u)\varphi, \psi\rangle = \langle \big(\int_G u(g)U(g)dg\big)\varphi, \psi\rangle. \tag{66.197}$$

8.
9. For linear operators on Hilbert spaces, the fact to be bounded is equivalent to continuity.

What we build up to here is a map

$$\begin{aligned} U\colon L^1(G,dg) &\to \mathfrak{B}(\mathscr{H}_\chi) \\ u &\mapsto U(u) \end{aligned} \tag{66.198}$$

given by

$$U(u) = \int_G u(g)U(g)dg. \tag{66.199}$$

This map is linear and continuous because $\|U(u)\|_{op} \leqslant \|u\|_{L^1}$.

We are now going to use the symmetry on M in order to descend U from $L^1(G,dg)$ to $L^1(M)$. Let us take a look at the two projections from G:

$$\begin{array}{ccc} G & \xrightarrow{\pi^M} & G/H \\ {\scriptstyle \pi^Q}\downarrow & & \\ G/B & & \end{array} \tag{66.200}$$

If $X, Y \in \mathcal{H}$, we recall that the definition of $\mathrm{ad}(X)^*$ is

$$\langle \xi_0, [X,Y]\rangle = -\langle \mathrm{ad}(X)^*\xi_0, Y\rangle, \tag{66.201}$$

but, since $e^{tX} \in \mathrm{Stab}(\xi_0)$, we have $\frac{d}{dt}\Big[\,\mathrm{Ad}(\exp(tX))^*\xi_0\Big]_{t=0} = 0$, thus

$$\langle \mathrm{ad}(X)^*\xi_0, Y\rangle = 0 \tag{66.202}$$

and we can suppose that the good polarization $\mathcal{B}$ contains $\mathcal{H}$. In that case we have the well defined map

$$\begin{aligned} \tilde{\pi}\colon G/H &\to G/B \\ gH &\mapsto gB \end{aligned} \tag{66.203}$$

This is well defined because, since $H \subset B$, we have

$$\tilde{\pi}(ghH) = ghB = gB \tag{66.204}$$

for every $h \in H$.

Lemma 66.58.
The map $\tilde{\pi}$ is a submersion.

Proof. No proof. □

The following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{\pi^M} & G/H \\ {\scriptstyle \pi^Q}\downarrow & \swarrow_{\tilde{\pi}} & \\ G/B & & \end{array} \tag{66.205}$$

Still two assumptions about σ:

(1) we suppose that B is stable under σ,

(2) we suppose that ξ_0 is σ-invariant, that is $\xi_0(\sigma X) = \xi_0(X)$.

The second assumption is easy to fulfill. If ξ_0 is not σ-invariant, we consider

$$\xi_0' = \frac{1}{2}(\xi_0 + \sigma^*(\xi_0)) \tag{66.206}$$

instead.

Now, the symmetry

$$\begin{aligned} \sigma_H \colon M &\to M \\ gH &\mapsto \sigma(g)H \end{aligned} \tag{66.207}$$

descends to Q as

$$\begin{aligned} \underline{\sigma} \colon Q &\to Q \\ gB &\mapsto \tilde{\pi}\big(\sigma_H(gH)\big) = \sigma(g)B. \end{aligned} \tag{66.208}$$

The so defined map $\underline{\sigma}$ is well defined because

$$\underline{\sigma}(gbB) = \sigma(gb)B = \sigma(g)\sigma(b)B = \sigma(g)B = \underline{\sigma}(gB). \tag{66.209}$$

Lemma 66.59.
Using the hypothesis of σ-invariance of ξ_0, we have that

$$\sigma^* \colon C^\infty(G,\mathbb{C})^B \to C^\infty(G,\mathbb{C})^B, \tag{66.210}$$

the image of a B-equivariant function on G by σ^* is still B-equivariant.

Proof. Let $\hat{\varphi} \in C^\infty(G,\mathbb{C})^B$, then we have

$$\begin{aligned} (\sigma^*\hat{\varphi})(gb) &= \hat{\varphi}\big(\sigma(g)\sigma(b)\big) \\ &= \chi\big(\sigma(b)^{-1}\big)\hat{\varphi}\big(\sigma(g)\big) \\ &= e^{-i\langle \xi_0, \sigma\log(b)\rangle}\hat{\varphi}(\sigma g) \\ &= e^{-i\langle \xi_0, \log(b)\rangle}(\sigma^*\hat{\varphi})(g) \quad \text{because } \xi_0(\sigma X) = \xi_0(X) \\ &= \chi(b^{-1})(\sigma^*\hat{\varphi})(g). \end{aligned} \tag{66.211}$$

□

Since the measure dq is σ^*-invariant by hypothesis, we have

$$\int_Q \overline{(\underline{\sigma}^* u)}(q)(\underline{\sigma}^* v)(q)dq = \int_Q \overline{u(q')}v(q)\underline{\sigma}^* dq = \int_Q \overline{u(q')}v(q)dq \tag{66.212}$$

where we used the change of variable $q' = \sigma q$. A consequence is that $\underline{\sigma}^*$ is an involution

$$\underline{\sigma}^* \colon L^2(Q,dq) \to L^2(Q,dq). \tag{66.213}$$

Since, in an abstract way, we denoted $L^2(Q,dq)$ by $\mathscr{H}_\chi$, we denote by Σ the involution $\underline{\sigma}^*$ on $\mathscr{H}_\chi$. Now we consider the function

$$\begin{aligned} \Omega \colon G &\to \mathrm{U}(\mathscr{H}_\chi) \\ g &\mapsto U(g)\Sigma U(g^{-1}) \end{aligned} \tag{66.214}$$

which is a composition of unitary maps. This is not a representation of the group G, but we have

$$\Omega(gh) = \Omega(g) \tag{66.215}$$

for every $h \in H$ and $g \in G$. Indeed let us compute $\widehat{\Omega(gh)\varphi}$ for $\varphi \in \mathscr{D}(Q)$. We have

$$\widehat{\Omega(gh)\varphi} = \hat{U}(gh)\sigma^*\hat{U}(h^{-1}g^{-1})\hat{\varphi} = \hat{U}(h)\hat{U}(h)\sigma^*\hat{U}(h^{-1})\hat{U}(g^{-1})\hat{\varphi}. \tag{66.216}$$

The element h only appears in the combination $\hat{U}(h)\sigma^*\hat{U}(h^{-1})$, so let us see how it acts on an equivariant function $\hat{\varphi}$. If we evaluate it on g_0 we find

$$\big(\hat{U}(h)\sigma^*\hat{U}(h^{-1})\hat{\varphi}\big)(g_0) = \big(\sigma^*\hat{U}(h^{-1})\hat{\varphi}\big)(h^{-1}g_0) = \hat{\varphi}\big(h\sigma(h^{-1}g_0)\big). \tag{66.217}$$

Let us recall that we are in the context [10] of subsection 60.8.1: the representation U on $\mathscr{D}(Q)$ comes from the regular left representation $\hat{U}$ on $C^\infty(G,\mathbb{C})^B$. Thus we have $\big(\hat{U}(g)\hat{\varphi}\big)(g_0) = \hat{\varphi}(g^{-1}g_0)$. Equation (66.217) is thus equal to

$$\hat{\varphi}\big(h\sigma(h^{-1}g_0)\big) = \hat{\varphi}\big(h\sigma(h^{-1})\sigma(g_0)\big) = \hat{\varphi}\big(\sigma(g_0)\big) = (\sigma^*\hat{\varphi})(g_0), \tag{66.218}$$

so that equation (66.216) does not depend on h, which proves that

$$\Omega(gh) = U(g)\Sigma U(g^{-1}) = \Omega(g). \tag{66.219}$$

One consequence of this circumstance is that Ω is a function which pass to the quotient $G \to G/H$. Thus we consider the map

$$\Omega\colon M = G/H \to \mathrm{U}(\mathscr{H}_\chi) \subset \mathfrak{B}(\mathscr{H}). \tag{66.220}$$

Since M is a symplectic manifold, we have a natural volume form

$$dx = \frac{1}{n!}\omega^n \tag{66.221}$$

where $n = \frac{1}{2}\dim M$. This measure allows us to consider the map

$$\begin{aligned} \Omega\colon L^1(M,dx) &\to \mathfrak{B}(\mathscr{H}) \\ u &\mapsto \Omega(u) \end{aligned} \tag{66.222}$$

defined by

$$\Omega(u) = \int_M u(x)\Omega(x)dx \tag{66.223}$$

which is a continuous linear map. This is not a representation (even on G the initial Ω was not a representation and M is not a group), but it is an unitary representation of M in the following sense.

Definition 66.60.
*An **unitary representation** is a map $\Omega\colon M \to \mathrm{U}(\mathscr{H})$ of the symmetric space M when it satisfies to the properties*

(1) $\Omega(x)\Omega(y)\Omega(x) = \Omega(s_x y)$

(2) $\Omega(x)^2 = \mathrm{Id}\,|_{\mathscr{H}}$.

for every $x, y \in M$.

66.8 Hermitian and symplectic spaces

SecHermEtSymplecticSpaces

If ANK is the Iwasawa decomposition of a Lie group [11], one can consider the manifold $M = G/K$. There is a natural identification

$$\mathcal{P} = T_K M$$

where $\mathcal{P}$ comes from the Cartan decomposition $\mathcal{G} = \mathcal{P} \oplus \mathcal{K}$. Indeed, the Iwasawa theorem says that $M \simeq AN$ so that a path in it reads $g(t) = a(t)n(t)$ with $g(0) = e$. But one has a diffeomorphism $A \times N \times K \to G$, so that $g(0) = e$ implies $a(0) = n(0) = e$. Thus Leibniz makes $g'(0) = a'(0) + n'(0)$ and $g'(0) \in \mathcal{A} \oplus \mathcal{N} = \mathcal{P}$.

Let us recall that when G is a Lie group, and H a closed connected subgroup of G, G/H is a manifold on which G acts. This structure is an **homogeneous space**. If moreover H is the set of the fixed points of an involution on G, G/H is says to be a **symmetric space**.

More precisely, the involution is a $\theta\colon \mathcal{G} \to \mathcal{G}$ which let fixed $\mathcal{H} \subset \mathcal{G}$; then H is the connected Lie group whose Lie algebra is $\mathcal{H}$. All this makes that the G/K from Iwasawa is a symmetric space.

10. With many notational incoherences.
11. c'est pas mal de dire quel genre de groupes : simple ? semi ? compact ?

DefCLtjFtD

Definition 66.61.
*Here, M denotes a connected smooth manifold. An **almost complex structure** on M is a $(1,1)$ tensor field J such that $\forall X \in \mathfrak{X}(M)$,*

$$(J \circ J)X = -X. \tag{66.224}$$

*The tensor field J is a **complex structure** when moreover it satisfies the* integrability condition*: $\forall X, Y \in \mathfrak{X}(M)$,*

$$N(X,Y) := [X,Y] + J[JX,Y] + J[X,JY] - [JX,JY] = 0. \tag{66.225}$$

DefComplStruct

We already spoke about complex structure in order to define the signed curvature of planar curve around the definition 21.72.

DefSymHermMGKalg

Definition 66.62.
*The symmetric space $M = G/K$ is **hermitian** if there exists an endomorphism $J \in \operatorname{End}\mathcal{P}$, $J\colon \mathcal{P} \to \mathcal{P}$ such that*

$$J^2 = -id_{\mathcal{P}}, \tag{66.226a}$$ eq:herm_1

$$B(JX,JY) = B(X,Y) \qquad \forall\, X, Y \in \mathcal{P}, \tag{66.226b}$$ eq:herm_2

$$\operatorname{ad}(k) \circ J = J \circ \operatorname{ad}(k) \qquad \forall\, k \in \mathcal{K}. \tag{66.226c}$$ eq:herm_3

def:hermitien

Since one has the identification $\mathcal{P} = T_K M$, J is only defined on $T_{[e]}M$. The following proposition extends the definition.

prop:ext_J

Proposition 66.63.
The hermitian structure J can be extended to a complex structure $\overline{J}$ on the whole TM.

Proof. For $X \in T_{[g]}M$, we define

$$\overline{J}(X) := dL_g \circ J \circ dL_{g^{-1}} X. \tag{66.227}$$

where dL is the differential of $L_g\colon G/K \to G/K$, $L_g[h] = [gh]$. From this, $\overline{J}^2(X) = -X$ because

$$(\overline{J} \circ \overline{J})X = (dL_g J dL_{g^{-1}}) \circ (dL_g J dL_{g^{-1}})X = dL_g J^2 dL_{g^{-1}} X = -X. \tag{66.228}$$

On the other hand, J satisfies $\operatorname{ad}(k) \circ J = J \circ \operatorname{ad}(k)$ and we want the same for $\overline{J}$:

$$\operatorname{ad}(X) \circ \overline{J} = \overline{J} \circ \operatorname{ad}(X)$$

for $X \in T_{[g]}M$. Note that it is true for $[g] = [e]$ because $T_{[e]}M = \mathcal{K}$. Let us consider $X \in T_{[g]}M$, and let us see what is $\big((\operatorname{ad} X) \circ J\big)Y$ for a $Y \in T_{[g]}M$. Consider x, $y \in T_{[e]}M$ such that $X = dL_g x$ and $Y = dL_g y$. Suppose one has

$$\operatorname{ad}(dL_g x)Y = dL_g \circ \operatorname{ad}(x)(dL_{g^{-1}} Y); \tag{66.229}$$ eq:suppose

then one can compute

$$\begin{aligned}
\big(\operatorname{ad}(X) \circ \overline{J}\big)Y &= \operatorname{ad}(dL_g x) \circ dL_g \circ J \circ dL_{g^{-1}} Y && \text{(66.230a)}\\
&= dL_g \operatorname{ad}(x) \circ J \circ dL_{g^{-1}} Y && \text{(66.230b)}\\
&= dL_g \circ J \circ \operatorname{ad}(x) \circ dL_{g^{-1}} Y && \text{(66.230c)}\\
&= (dL_g \circ J \circ dL_{g^{-1}}) \circ (dL_g \circ \operatorname{ad}(x) \circ dL_{g^{-1}}) && \text{(66.230d)}\\
&= \overline{J} \circ \operatorname{ad}(X)Y. && \text{(66.230e)}
\end{aligned}$$

subEqEMyROwA

The line (66.230b) comes from $\mathrm{ad}(x)\circ J = J\circ\mathrm{ad}(x)$ because $x\in T_{[e]}M$.

Now, we prove equation (66.229) which is rewritten in a more convenient way as $[dL_g x, Y] = dL_g[x, dL_{g^{-1}}Y]$. Thus one has to see if for any x, $y\in T_{[e]}M$,

$$dL_g[x,y] = [dL_g x, dL_g y].$$

This is true because of [813], proposition 3.3, page 34. Now we know that $\forall X\in T_{[g]}M$ we have $\overline{J}\circ\mathrm{ad}(X) = \mathrm{ad}(X)\circ\overline{J}$ and $\overline{J}^2X = -X$. In order to have a complex structure, one also need to check condition (66.225), which is true because

$$\begin{aligned} J[JX,Y] &= -\operatorname{ad}Y\circ JJX = \mathrm{ad}(Y)X = [X,Y],\\ J[X,JY] &= J\circ\mathrm{ad}(X)JY = -[X,Y],\\ -[JX,JY] &= -(\operatorname{ad}JX\circ J)Y = -J(\operatorname{ad}JX)Y = -[Y,X]. \end{aligned} \tag{66.231}$$

□

If $X\in T_{[g]}M$,

$$\begin{aligned} (\overline{J}\circ dL_h)X &= dL_{hg}\circ J\circ dL_{(hg)^{-1}}dL_h X\\ &= dL_{hg}\circ J\circ dL_{g^{-1}}X\\ &= dL_h\circ dL_g\circ J\circ dL_{g^{-1}}X\\ &= (dL_h\circ\overline{J})X. \end{aligned}$$

so we have an important property:

$$\overline{J}\circ dL_h = dL_h\circ\overline{J}. \tag{66.232}$$ eq:J_o_dL

From now, it is clear that we will often forget the bar on $\overline{J}$. In the same way that J extends to M,

Proposition 66.64.
For X, $Y\in T_{[g]}M$, the formula

$$\overline{B}(X,Y) := B(dL_{g^{-1}}X, dL_{g^{-1}}Y) \tag{66.233}$$ eq:BdL

defines a Riemannian metric on M.

Proof. One has to see that it is nondegenerate. Say that $Z\in T_{[g]}M$ is such that for any X, $\overline{B}(Z,X) = 0$. Then $B(dL_{g^{-1}}Z, dL_{g^{-1}}X) = 0$. But dL_g is a vector space isomorphism because $dL_g(o) = \frac{d}{dt}\Big[L_g(X_t)\Big]_{t=0}$ with X_t, a constant path at $[e]\in M$.

But since B is nondegenerate, the definition (66.233) says us $dL_{g^{-1}}Z = 0$, and then $Z = 0$. □

Now, one knows [12] that

$$\omega^M_x(X,Y) = g_x(JX,Y) \tag{66.234}$$

defines a G-invariant symplectic structure on M.

In order to see it, one has to show that (M,g,J) is a Kähler structure. The G-invariance comes from the extension of Killing form that we had chosen: $\forall X,Y\in T_{[g]}M$, $B_{[g]}(X,Y) = B(dL_{g^{-1}}X, dL_{g^{-1}}Y)$. It is clear that

$$B_{[hg]}(dL_h X, dL_h Y) = B_{[g]}(X,Y). \tag{66.235}$$

From this and equation (66.232), one can see the G-invariance of ω^M:

$$\begin{aligned} \omega^M_{[hg]}\big((dL_h)_{[g]}X, (dL_h)_{[g]}Y\big) &= B_{[hg]}(JdL_hX, dL_hY) = B_{[hg]}(dL_hJX, dL_hY)\\ &= B_{[g]}(JX,Y) = \omega^M_{[g]}(X,Y). \end{aligned} \tag{66.236}$$

The formulation of the G-invariance is

$$\omega^M_{[hg]}\Big((dL_h)_{[g]}X, (dL_h)_{[g]}Y\Big) = \omega^M_{[g]}(X,Y). \tag{66.237}$$

12. Cf cours de géométrie symplectique

66.8.1 The Chevalley cohomology

Let $\mathcal{G}$ be a Lie algebra (maybe infinite dimensional) and (V,ρ) a representation of $\mathcal{G}$ on the vector space V. The **Chevalley cohomology** of $\mathcal{G}$ associated with the representation ρ is given by the following definitions:

A p-cochain is a map $C\colon \underbrace{\mathcal{G}\times\ldots\times\mathcal{G}}_{p \text{ times}} \to V$ which is multi-linear and skew-symmetric. In particular, a 1-cochain is a linear map $\xi\colon \mathcal{G}\to V$. In the case of the trivial representation on $\mathbb{R}$, a 1-cochain is an element of $\mathcal{G}^*$. The coboundary of a p cochain is the $p+1$-cochain given by

$$\begin{aligned}(\delta C)(X_0,\ldots,X_p) = &\sum_{i=0}^{p}(-1)^i\rho(X_i)C(X_0\ldots,\hat{X}_i,\ldots,X_p)\\ &+\sum_{i<j}(-1)^{i+j}C\big([X_i,X_j],\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots,X_p\big).\end{aligned} \tag{66.238}$$

The main property is $\delta^2=0$. The others definitions are as usual: a p-cocycle is a p-cochain C such that $\delta C=0$, a p-coboundary is a p-cochain which can be written as δB for some $(p-1)$-cochain B. Finally, the cohomology classes are:

$$H^p_{(V,\rho)} = \frac{p\text{-cocycles}}{p\text{cochain}} = H^p_\rho(\mathcal{G},V). \tag{66.239}$$

When one consider the trivial representation, i.e. $\rho(X)=0$, a 1-cochain is $\xi\in\mathcal{G}^*$ and

$$(\delta\xi)(X,Y) = -\xi([X,Y]). \tag{66.240}$$ EqDefcochaintrivC

¡¡ Avertissement/question à la lectrice !! 66.65
Au cas où ça t'intéresserait, je te dis que le signe moins, tu ne l'as ajouté qu'en février 2007. T'étonnes pas si y'a des signes qui foirent plus bas.

Now, on the symmetric hermitian space $M=G/K$, one defines a $\Omega\in\Lambda^2(\mathcal{G}^*)$ by

$$\Omega(X,Y) = B(JX,Y) \qquad \text{for } X,Y\in\mathcal{P} \tag{66.241a}$$ eq:def_Omega_1

$$\Omega(\mathcal{K},\mathcal{G}) = 0. \tag{66.241b}$$ eq:def_Omega_2

A great property of this definition is that Ω is a 2-cocycle for the trivial representation of $\mathcal{G}$ on $\mathbb{R}$:

$$\Omega([X,Y],Z)+\Omega([Y,Z],X)+\Omega([Z,X],Y)=0.$$

Indeed, if X ,Y, $Z\in\mathcal{P}$, the commutators are in $\mathcal{K}$, so that (66.241b) makes the whole null. The second case is X, $Y\in\mathcal{P}$ and $Z\in\mathcal{K}$; for this, we are led to consider the quantity $-B([Y,Z],JX)-B([Z,X],JY)$. The first term can be transformed as:

$$\begin{aligned}B([Y,Z],JX) &= -B(J[Y,Z],X) && \text{by def. (66.226b)}\\ &= B([Z,JY],X) && \text{by def. (66.226c)}\\ &= -B(JY,[Z,X]) && \text{ad}-\text{invariance of } B\\ &= -B([Z,X],JY).\end{aligned}$$

So it is zero.

Lemma 66.66 (Whitehead's lemma)**.**
If $\mathcal{G}$ is a finite dimensional semisimple Lie algebra and ρ a non trivial[13] *representation of $\mathcal{G}$ on V, then $\forall\, q\geqslant 0$,*

$$H^q(\mathcal{G},V)=0.$$

13. Ce qui n'est pas le cas ici

This gives us the existence of a $\xi_0 \in \mathcal{G}^*$ (an Chevalley 1-cochain) such that

$$\delta\xi_0 = \Omega.$$

$$\xi_0 = B(Z_0, .). \tag{66.242}$$

Definition 66.67.
*The **center** of the Lie algebra $\mathcal{G}$ is the set $\mathcal{Z}(\mathcal{G}) \subset \mathcal{G}$ of elements Z such that $[Z, X] = 0$ for every $X \in \mathcal{G}$. See also the definition of a centralizer on page 3070.*

Proposition 66.68.
The Z_0 defined in (66.242) and the J of proposition 66.63 satisfy

$$Z_0 \in \mathcal{Z}(\mathcal{K}), \tag{66.243a}$$

$$J = \pm \operatorname{ad}(Z_0)|_{\mathcal{P}}. \tag{66.243b}$$

Proof. First, we see that for any $K \in \mathcal{K}$, $[Z_0, K] = 0$. We know from (66.241b) that $\forall\, G \in \mathcal{G}$, $K \in \mathcal{K}$, $\Omega(K, G) = 0$, or

$$0 = \delta B(Z_0, .)(K, G) = -B(Z_0, .)([K, G]) = B([K, Z_0], G), \tag{66.244}$$

thus $[K, Z_0] = 0$ because B is nondegenerate. We will see below that $Z_0 \in \mathcal{P}$ is not possible. On the other hand, the condition (66.241a) gives us

$$B([X, Z_0], Y) = B(JX, Y)$$

for any $Y \in \mathcal{P}$. Thus $[X, Z_0] = JX$ and the second claim follows. Let us now see that $Z_0 \in \mathcal{P}$ is not possible (and so we finish the proof of the first claim). We know that $J^2X = [Z_0, [Z_0, X]]$, but for Z_0, $X \in \mathcal{P}$, $[Z_0, X] \in \mathcal{K}$ and so $J^2X = 0$ which is not possible.

□

Lemma 66.69.
A symmetric space G/K is hermitian if and only if $\mathcal{Z}(\mathcal{K}) \neq 0$.

Proof. If the space is hermitian, we just said that the J can be written under the form $J = -\operatorname{ad}(Z_0)|_{\mathcal{P}}$ for a $Z_0 \in \mathcal{Z}(\mathcal{K})$. For the sufficient condition, we define $J = -\operatorname{ad}(Z_0)$ for a certain $Z_0 \in \mathcal{Z}(\mathcal{K})$. As a first point for all $k \in \mathcal{K}$ and $p \in \mathcal{P}$,

$$(J \circ \operatorname{ad} k)p = [[k, p], Z_0] = -[[p, Z_0], k] - [[Z_0, k], p] = [k, [p, Z_0]] = (\operatorname{ad} k \circ J)p \tag{66.245}$$

The two other points are

$$J^2 = (-\operatorname{ad} Z_0)[X, Z_0] = -[[Z_0, X], Z_0] \tag{66.246a}$$

and

$$B([Z_0, X], [Z_0, Y]) = -B(X, [Z_0, [Z_0, Y]]) = B(X, Y) \tag{66.246b}$$

These are true if $[[Z_0, X], Z_0] = X$ [14]

Let us prove that $\mathcal{M} := \ker(\operatorname{ad} Z_0) = 0$. For remark that $(\mathcal{G}, B)$ is a Riemannian space and let W be the orthogonal complement of $\mathcal{M}$ in $\mathcal{P}$. We begin to prove that $\mathcal{M}$ is $(\operatorname{ad}\mathcal{K})$-invariant.

If $x \in |Z\rangle_0$ and $k \in \mathcal{K}$, then

$$[Z_0, [k, x]] = -[k, [x, Z_0]] - [x, [Z_0, k]] = 0$$

because $[x, Z_0] = 0 = [Z_0, k]$. Now if $\mathcal{M}$ is $(\operatorname{ad}\mathcal{K})$-invariant, then $W = \mathcal{M}^\perp$ is too because

$$B([k, x], m) = -B(x, [k, m]) = 0$$

14. Mais je ne vois pas comment obtenir ça. Si $\operatorname{ad} Z_0$ est un automorphisme de $\mathcal{P}$, alors je suis d'accord.

since $[k,x] \in \mathcal{M}$. So we have the orthogonal direct sum $\mathcal{P} = \mathcal{M} \oplus W$. We are now going to see that $[\mathcal{M}, W] = 0$. Let $X, X' \in \mathcal{P}$;

$$B\big([[m,w],X],X'\big) = B\big([m,w],[X,X']\big) = B\big(w,[[X,X'],m]\big) = 0 \tag{66.247}$$

since $[X,X'] \in \mathcal{K}$ and $[[X,X'],m] \in \mathcal{M}$ from the $\mathcal{K}$invariance of $\mathcal{M}$. If we define $A = [m,w]$, the endomorphism $\mathrm{ad}(A)|_{\mathcal{P}}$ is zero.

We know [15] that $[\mathcal{K},\mathcal{K}] = \mathcal{P}$, and then that

$$[A,k] = [A,[p,p']] = -[p,[p',A]] - [p',[A,p]] = 0.$$

So $[A,\mathcal{G}] = 0$ and $A = 0$ because $\mathcal{G}$ is semisimple. If we write $\mathcal{G} = [\mathcal{P},\mathcal{P}] \oplus \mathcal{P}$, we find

$$\mathcal{G} = \big([\mathcal{M},\mathcal{M}] \oplus \mathcal{M}\big) \oplus \big([W,W] \oplus W\big),$$

where the two brackets commute. It furnish a decomposition of $\mathcal{G}$ into ideals which impossible from the semi-simplicity assumption. We conclude that $\mathcal{M} =$ and that $\mathrm{ad}\, Z_0$ is bijective on $\mathcal{P}$.

□

Lemma 66.70.
Let $\mathcal{G}$ a simple Lie algebra with Iwasawa decomposition $\mathcal{G} = \mathcal{K} \oplus \mathcal{A} \oplus \mathcal{N}$. We suppose that $\mathcal{Z}(\mathcal{K}) \neq 0$. Then $\dim \mathcal{A} \geqslant \dim \mathcal{Z}(\mathcal{N})$.

Proof. Let $i\colon \mathcal{R} \to \mathcal{G}$ be the canonical projection and $\xi_0 \in \mathcal{G}^*$ such that $\delta\xi_0 = \Omega$. Since $\mathcal{K} \cap \mathcal{R} = \{0\}$, the radical of $\delta(i^*\xi_0)$ is trivial. Indeed, when $X \in \mathcal{R}$, we have $(i^*\xi_0)X = \xi_0 X$ and equation $\delta(i^*\xi_0)(X,Y) = 0$ for all X, $Y \in \mathcal{R}$ gives $B(JX,Y) = 0$ because $\mathcal{K} \cap \mathcal{R} = \{0\}$. Then $JX = 0$ and $X = 0$. [16].

Let V be the radical of Ω in $\mathcal{N}$; if $z \in \mathcal{Z}(\mathcal{N})$, then $\Omega(\mathcal{N},z) = (\delta\xi_0)(z,\mathcal{N}) = \xi_0[z,\mathcal{N}] = 0$. Then $\mathcal{Z}(\mathcal{N}) \subset V$. Now let us consider the map $\psi\colon V \to \mathcal{A}^*$,

$$\psi(v) = \Omega(v,.)|_{\mathcal{A}}.$$

Let us prove that ψ is injective. For, we consider a $v \in V$ such that $\Omega(v,\mathcal{A}) = 0$. Since $v \in V$, we have $[v,\mathcal{N}] = 0$ and then $\Omega(v,\mathcal{N}) = 0$. So,

$$0 = \Omega(v, \mathcal{A} \oplus \mathcal{N}) = \delta(i^*\xi_0)(v,\mathcal{R}),$$

and then $v = 0$ because the radical of $i^*\xi_0$ in $\mathcal{R}$ is only zero. Consequently,

$$\dim \mathcal{A} = \dim \mathcal{A}^* \geqslant \dim V \geqslant \dim \mathcal{Z}(\mathcal{N})$$

because there exists an injection from V into $\mathcal{A}^*$ and $\mathcal{Z}(\mathcal{N}) \subset V$.

□

Lemma 66.71.
Let us suppose that $\dim \mathcal{A} = 1$ and $\dim \mathcal{G} \geqslant 3$. Then

(1) The root system is $\Phi = \{\pm\alpha, \pm 2\alpha\}$,

(2) $\mathcal{N} = \mathcal{G}_\alpha \oplus \mathcal{G}_{2\alpha}$ and $\mathcal{G}_{2\alpha} = \mathcal{Z}(\mathcal{N})$

(3) $\dim \mathcal{Z}(\mathcal{N}) = \dim \mathcal{A} = 1$

(4) There exists a $E \in \mathcal{Z}(\mathcal{N})$ such that $[x,y] = \Omega(x,y)E$ for all x, $y \in \mathcal{N}$. The subspaces $\mathcal{A} \oplus \mathcal{N}$ and $\mathcal{G}_\alpha$ are symplectic and orthogonal in $(\mathcal{R},\Omega)$. In particular, $\mathcal{N}$ is an Heisenberg algebra.

Proof. No proof. □

15. Il faut encore voir d'où sort ce truc.
16. Ça demande que B soit non dégénérée sur $\mathcal{R}$, et je ne vois pas trop pourquoi ce serait vrai.

This lemma allows us to parametrize $\mathcal{R}$ as

$$r = aA + x + zE$$

with $x \in \mathcal{G}_\alpha$ and $a \in \mathcal{A}$ because $\mathcal{Z}(\mathcal{N})$ is spanned by the unique element E. Now if we consider a function $u \in C^\infty(\mathcal{R})$, we can define a partial Fourier transform

$$F(u)(a,x,\xi) = \hat{u}(a,x,\xi) = \int_{\mathcal{Z}(\mathcal{N})} e^{-i\xi z} u(aA + x + zE)dz.$$

Theorem 66.72.
Let consider the diffeomorphism $\phi_\hbar \colon \mathcal{R} \to \mathcal{R}$ given by

$$\phi_\hbar(a,x,\xi) = \left(a, \frac{1}{\cosh(\frac{\hbar\xi}{2})}x, \frac{\sinh(\hbar\xi)}{\hbar}\right).$$

Then

(1) $\phi_\hbar^* \mathscr{S}(\mathcal{R}) \subset \mathscr{S}(\mathcal{R})$,
(2) $(\phi_\hbar^{-1})^* \mathscr{S}(\mathcal{R}) \subset \mathscr{S}'(\mathcal{R})$.
where $\mathscr{S}$ and its dual $\mathscr{S}'$ are defined in section 61.1.

We recall the notation for functions: $\varphi^* f = f \circ \varphi$.

Proof. For sake of simplicity, we forget about variable a, we pose $y = \hbar\xi$ and we look at the function $\phi \colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\phi(x,y) = (\operatorname{sech}(\frac{y}{2})x, \sinh(y)).$$

Formula

$$\frac{\sqrt{2}}{2}(1+\sqrt{1+y^2})^{1/2} = \cosh\left(\frac{\operatorname{arcsinh}(y)}{2}\right),$$

allows us to write

$$\phi^{-1}(x,y) = \left(\cosh\left(\frac{\operatorname{arcsinh}(y)}{2}\right)x, \operatorname{arcsinh}(y)\right). \tag{66.248}$$

We pose $p_{nm}(x,y) = x^n y^m$ and we are going to study

$$(p_{nm} \circ \phi^{-1})(x,y).$$

It has a polynomial grown because, for large y, $\sinh(\ln y) = \frac{1}{2}y$. Hence $\operatorname{arcsinh}(y) \simeq \ln(2y)$. The matrix of $d\phi_{(x,y)}$ is given by

$$d\phi_{(x,y)} = \begin{pmatrix} \operatorname{sech}(\frac{y}{2}) & -\frac{x}{2}\tanh(\frac{y}{2})\operatorname{sech}(\frac{y}{2}) \\ 0 & \cosh(y) \end{pmatrix}. \tag{66.249}$$

Since ϕ is a diffeomorphism, and then is bijective,

$$\begin{aligned} \sup_{a\in\mathbb{R}^2} |p_{nm}(a)(u\circ\phi)(a)| &= \sup_{a\in\mathbb{R}^2} |p_{nm}(\phi^{-1}(a))u(a)| \\ &\leqslant \sup_{a\in\mathbb{R}^2} |P_{MN}(a)u(a)| \end{aligned} \tag{66.250}$$

for a choice of N, $M \in \mathbb{N}$. In order to check the derivatives, we need the asymptotic behaviour of $(u\circ\phi)'(x,y)$ given components of [17]

$$\partial_a(u\circ\phi)(x,y) = \sum_i (\partial_i u)_{\phi(x,y)}(\partial_a\phi_i)(x,y).$$

17. Pour moi, la composante $a = i = 2$ ne fonctionne pas parce que c'est $(\partial_2)_{\phi(x,y)}\cosh(y)$.

The derivatives of $\phi^* u$ are the quantities

$$\partial_a(u\circ\phi^{-1})(x,y) = (\partial_i u)(\phi^{-1}(x,y))\partial_a(\phi_i^{-1})(x,y).$$

This has a polynomial behaviour. One can see recursively that the same is true for second derivatives $\partial^2_{ab}(u\circ\phi^{-1})$ and higher. This proves that $\phi^*\mathscr{S}(\mathbb{R}^2)\subset\mathscr{S}(\mathbb{R}^2)$.

In order to see that $(\phi^{-1})^* u\in\mathscr{S}'$ when $u\in\mathscr{S}$, we have to prove that

$$\int_{\mathcal{U}} |x^{-N}y^{-M}(\phi^{-1})^* u(x,y)dxdy| < \infty \tag{66.251}$$

eq:r1181205

for a choice of N, M. Here, $\mathcal{U}$ is the complement in $\mathbb{R}^2$ of a compact neighbourhood of the origin. Indeed, the fact for f to belongs to $\mathscr{S}'$ is the *distribution* T_f to belongs to $\mathscr{S}'$. In other words, the condition $f\in\mathscr{S}'$ is the continuity of $\varphi\to\int_X f\varphi$ when $\varphi\in\mathscr{S}$. The essentially resides in the existence of the integral.

In general – here, f take the role of $\phi^* u$ – we have

$$|\int_X f\varphi|\int|f\varphi|\leqslant\int|fp_{-N,-M}|$$

for all $N, M\geqslant 0$ because φ decrease more rapidly than any polynomial. If we find M and N such that $\int|fp_{-N,-M}|<\infty$, then we prove that the distribution belongs to $\mathscr{S}'$. In our case more precisely, we know that φ is smooth. Then it can be majored in any compact set. This is the reason why we write an integral over $\mathcal{U}$ instead of the whole $\mathbb{R}^2$.

In equation (66.251), we perform the change of variable $a'=\phi^{-1}(a)$:

$$\begin{aligned}\int_{\mathcal{U}}|x^{-N}y^{-M}(u\circ\phi^{-1})(a)|\,da &= \int_{\mathcal{U}'}|\frac{1}{p_{MN}(\phi(a'))}u(a')||J_\phi(a')|\,da'\\ &= \int_{\mathcal{U}'}\frac{|u(a)|}{\left|\left(\frac{x}{\cosh(\frac{y}{2})}\right)^N\sinh(y)^M\right|}\left|\frac{\cosh(y)\cosh(\frac{y}{2})}{}\right|da\\ &= \int_{\mathcal{U}'}\left|\frac{1}{x^N}2^{1-M}\cosh(\frac{y}{2})^{N-M}\sinh(\frac{y}{2})^{1-M}\right||u(a)|\,da.\end{aligned} \tag{66.252}$$

The latter integral is finite when $M\geqslant 1$ and $M>N$.[18]

The same kind of upper bound[19] holds for the derivatives of $u\circ\phi^{-1}$ for which we have to study the behaviour of the inverse matrix $(d\phi)^{-1}$. All this proves that $(\phi^{-1})^*\mathscr{S}(\mathbb{R}^2)\subset\mathscr{S}(\mathbb{R}^2)$.

□

66.8.2 Involutive symmetric Lie algebras

Definition 66.73.
*An **involutive Lie algebra** is a doublet $(\mathcal{G},\sigma)$ where $\mathcal{G}$ is a real finite dimensional Lie algebra and $\sigma:\mathcal{G}\to\mathcal{G}$ is an involutive automorphism of $\mathcal{G}$.*

There are three types of triples $(\mathcal{G},\sigma,\Omega)$:

(1) the symplectic triple,

(2) the exact triple,

(3) the elementary solvable exact triple (ESET).

In these three types, $(\mathcal{G},\sigma)$ is an involutive Lie algebra. The following definitions can be found in [961].

Symplectic triples were already defined in section 66.7.3.

18. Pierre trouve d'autres choses, mais sa conlusion est la même; comme s'il utilisait un autre formulaire de trigono hyperbolique que moi.

19. Traduction de «majoration»

Definition 66.74.
*An **exact triple** is a triple $(\mathcal{G}, \sigma, \Omega)$ such that*

(1) $\mathcal{G} \stackrel{\sigma}{=} \mathcal{K} \oplus \mathcal{P}$ and $[\mathcal{P}, \mathcal{P}] = \mathcal{K}$,

(2) Ω is a Chevalley 2-coboundary such that $i(\mathcal{K})\Omega = 0$ and $\Omega|_{\mathcal{P}\times\mathcal{P}}$ is a symplectic structure on $\mathcal{P}$.

The exact triple has the following differences compared to the symplectic one:

— Ω is a coboundary instead as a cocycle,

— $\Omega|_{\mathcal{P}\times\mathcal{P}}$ is not only nondegenerate, but also symplectic.

From definition of a coboundary, in an exact triple, there exists a $\xi \in \mathcal{G}^*$ such that $\Omega = \delta\xi$.

Definition 66.75.
*An **elementary solvable exact triple** (ESET) is an exact triple $(\mathcal{G}, \sigma, \Omega)$ such that*

(1) The Lie algebra $\mathcal{G}$ is a split extension of abelian algebras:

$$\mathcal{G} = \mathcal{A} \oplus_\rho \mathcal{B}., \tag{66.253}$$ EqSplitmGABab

(2) the automorphism σ preserves the vector space decomposition $\mathcal{G} = \mathcal{A} \oplus \mathcal{B}$.

Remarque 66.76.
When one writes $\mathcal{G} = \mathcal{A} \oplus_\pi \mathcal{B}$, one has $\pi\colon \mathcal{A} \to \mathtt{Der}(\mathcal{B})$. This is the inverse convention of the one chosen in the article [961].

In the case of an ESET, we have $\mathcal{A} \cap \mathcal{K} \subset \mathcal{A} \cap [\mathcal{G}, \mathcal{G}]$ because $\mathcal{K}$ is equal to $[\mathcal{P}, \mathcal{P}]$ and is thus included in $[\mathcal{G}, \mathcal{G}]$. But $[\mathcal{G}, \mathcal{G}]$ can be $[a, b]$, $[a, b]$ or $[b, b]$. The two latter are zero (because $\mathcal{A}$ and $\mathcal{B}$ is abelian) and, by definition of the split extension, $[a, b] = \rho(a)b \in \mathcal{B}$. So $\mathcal{A} \cap [\mathcal{G}, \mathcal{G}] = 0$. Therefore,

$$\mathcal{A} \cap \mathcal{K} = 0;$$

we deduce that $\mathcal{A} \subset \mathcal{P}$ and $\mathcal{K} \subset \mathcal{B}$. Since $\mathcal{K} \subset \mathcal{B}$, we define $\mathcal{L}$ as the complement:

$$\mathcal{B} = \mathcal{K} \oplus \mathcal{L}.$$

In particular, $\mathcal{K}$ and $\mathcal{L}$ are abelian.

The dimension of a triple is the dimension of $\mathcal{P}$ and two triples $(\mathcal{G}_i, \sigma_i, \Omega_i)$ are **isomorphic** if there exists a Lie algebra isomorphism $\psi\colon \mathcal{G}_1 \to \mathcal{G}_2$ such that $\psi \circ \sigma_1 = \sigma_2 \circ \psi$ and $\psi^*\Omega_2 = \Omega_1$.

66.8.3 Symplectic symmetric spaces and involutive Lie algebra

Let (M, ω, s) be a symplectic symmetric space; we associate an involutive Lie algebra $(\mathfrak{g}, \sigma)$ in the following way (we omit some non trivial proofs). Let $o \in M$, G the transvection group and H, the stabiliser of o in $\mathrm{Aut}(M, \omega, s)$ and $K = G \cap H$. We consider the map

$$\begin{aligned} \tilde{\sigma}\colon\ \mathrm{Aut}(M, \omega, s) &\to \mathrm{Aut}(M, \omega, s) \\ \tilde{\sigma}(g) &= s_o \circ g \circ s_o. \end{aligned} \tag{66.254}$$

Let $\mathcal{G}$ be the Lie algebra of the group G and $\sigma\colon \mathcal{G} \to \mathcal{G}$ the induced involutive automorphism from $\tilde{\sigma}$. Now, $(\mathcal{G}, \sigma)$ is an involutive Lie algebra. We have a natural projection $\pi\colon G \to M$ because H stabilises o, so that $K = G \cap H$ is the stabiliser of o in $\mathrm{Aut}(M, \omega, s)$ which is transitive on M. Then $M = G/K$ as homogeneous spaces. One can see that $(\mathcal{G}, \sigma, \pi^*(\omega_o))$ is a symplectic triple.

The precise proposition is the following.

Proposition 66.77.
There exists a bijection between symplectic simply connected symmetric spaces and symplectic triples. This bijection is given up to isomorphism.

66.8.4 Symmetric spaces and coadjoint orbits

We are now going to describe (M,ω,s) as coadjoint orbit on $\mathcal{G}^*$. When a Lie group G of symplectomorphism acts on a symplectic manifold (M,ω), we say that the action is weakly Hamiltonian if there exists $\mu_X\colon M\to\mathbb{C}$ such that $i(X^*)\omega = d\mu_X$. If $\mu\colon \mathcal{G}\to C^\infty(M)$ is a Lie algebra homomorphism, we say that the action is Hamiltonian and we usually write λ instead of μ.

Proposition 66.78.
Let $(\mathcal{G},\sigma,\Omega)$ be a simple triple, (M,ω,s) the associated symmetric simply connected symplectic space and G, the transvection group. Then

(1) The action of the transvection group on M is Hamiltonian if and only if there exists a $\xi\in\mathcal{G}^*$ with $\Omega=\delta\xi$ for the Chevalley cohomology.

(2) In this case, (M,ω,s) is a G-equivariant symplectic covering of the coadjoint orbit of ξ in $\mathcal{G}^*$.

Definition 66.79.
*A **symmetric symplectic space** is a triple (M,ω,s) where*

— *M is a connected smooth (C^∞) manifold,*

— *ω is a symplectic form on M,*

— *$s\colon M\times M\to M$ is a smooth map which we write with the notation $s_x(y):=s(x,y)$.*

These elements must satisfy the following conditions:

*(1) $\forall x\in M$, s_x is an involutive symplectic diffeomorphism of (M,ω) which is called the **symmetry** at x,*

(2) $\forall x\in M$, x is an isolated fixed point of s_x,

(3) $\forall x,y\in M$, $s_xs_ys_x = s_{s_x(y)}$.

Definition 66.80.
*Two symplectic symmetric spaces (M,ω,s) and (M',ω',s') are **isomorphic** if there exists a symplectic diffeomorphism $\varphi\colon(M,\omega)\to(M',\omega')$ such that*

$$\varphi\circ s_x = s'_{\varphi(x)}\circ\varphi. \tag{66.255}$$

Definition 66.81.
*An **exact triple** is a triple $(\mathcal{G},\sigma,\Omega)$ such that*

(1) $(\mathcal{G},\sigma)$ is an involutive Lie algebra with $[\mathcal{P},\mathcal{P}]=\mathcal{K}$ if $\mathcal{G}=\mathcal{K}\oplus\mathcal{P}$ is the decomposition of $\mathcal{G}$ with respect to σ.

(2) Ω is a Chevalley 2-coboundary such that $i(\mathcal{K})\Omega=0$ and $\Omega_{\mathcal{P}\times\mathcal{P}}$ is symplectic.

From definition, there exists a $\xi\in\mathcal{G}^*$ for which $\Omega=\delta\xi$. We can choose it in such a way that $\xi(\mathcal{P})=0$; in this case we say that $\xi\in\mathcal{K}^*$ by abuse of notation. Indeed, put $\Omega=\delta\xi$ with $\xi=\xi'+\eta'$ where $\xi'\in\mathcal{K}^*$ and $\eta'\in\mathcal{P}^*$. If we consider $k\in\mathcal{K}$ and $B=B_k+B_p\in\mathcal{G}$, using $i(\mathcal{K})\Omega=0$, we find

$$0=\Omega(k,B)=-\xi'[k,B_k]-\eta'[k,B_p].$$

Taking $B_k=0$, we find $\eta'[\mathcal{K},\mathcal{P}]=0$ while with $B_p=0$, we find $\eta'[\mathcal{K},\mathcal{K}]=0$. Moreover $\eta'[\mathcal{P},\mathcal{P}]=\eta'\mathcal{K}=0$. Then an acceptable η' must satisfy $\eta'[\mathcal{G},\mathcal{G}]=0$, so that

$$\Omega[A,B]=-\xi'[A,B]$$

which proves that the η' part of ξ has no importance; we can choose it as zero.

Let $(\mathcal{G},\sigma)$ be an involutive Lie algebra associated with a triple (M,ω,s) with transvection group G. If $(\mathcal{G},\sigma,\Omega)$ is exact, $\mathcal{Z}(\mathcal{G})\subset\mathcal{K}$ because $[Z,p]=0$ for all $p\in\mathcal{P}$ whenever $Z\in\mathcal{P}\cap\mathcal{Z}(\mathcal{G})$. Then $\Omega(Z,p)=0$ for all $p\in\mathcal{P}$ which is not possible from non degeneracy of Ω.

66.8.5 Elementary solvable symmetric spaces

Let (M, ω, s) a symmetric space with associated triple $(\mathcal{G}, \sigma, \Omega)$. The space M is **elementary solvable** if

(1) $\mathcal{G}$ is a split extension (see subsection 51.15.1) of two abelian algebras $\mathcal{A}$ and $\mathcal{B}$,

(2) the automorphism σ preserves the decomposition $\mathcal{G} = \mathcal{B} \oplus \mathcal{A}$.

Since $\mathcal{K} = [\mathcal{P}, \mathcal{P}]$, we have

$$\mathcal{A} \cap \mathcal{K} \subset \mathcal{A} \cap [\mathcal{G}, \mathcal{G}] = 0.$$

Indeed, let $\rho \colon \mathcal{A} \to \mathtt{Der}\,\mathcal{B}$ be the split homomorphism; the commutator on the split extension is defined by

$$[A, B] = \rho(A)B \in \mathcal{B}.$$

Then $[\mathcal{G}, \mathcal{G}] \subset \mathcal{B}$. All this shows that $\mathcal{A} \subset \mathcal{P}$. So there exists a $\mathcal{L} \subset \mathcal{P}$ such that $\mathcal{B} = \mathcal{K} \oplus \mathcal{L}$. Let us show that $\mathcal{L}$ is abelian.

$$0 = [\mathcal{B}, \mathcal{B}] = [\mathcal{K}, \mathcal{K}] + [\mathcal{K}, \mathcal{L}] + [\mathcal{L}, \mathcal{K}] + [\mathcal{L}, \mathcal{L}].$$

The three first terms are in $\mathcal{P}$ while the last one is included in $\mathcal{K}$. The identical annihilation of the sum imposes $[\mathcal{L}, \mathcal{L}] = 0$.

66.8.6 Mid-point map

Let us now take an ESET $(\mathcal{G}, \sigma, \Omega)$ and its corresponding ESSS (M, ω, s). There exists a $\xi \in \mathcal{G}^*$ such that $\Omega = \delta\xi$ and we define

$$\begin{aligned} \zeta \colon \mathcal{A} \times \mathcal{L} &\to \mathbb{R} \\ (a, l) &\mapsto \xi(\sinh(a)l) \end{aligned} \tag{66.256}$$

where $\sinh(a)l$ has to be understood as $\frac{1}{2}(e^{\rho(a)} - e^{-\rho(a)})l$ with $\rho(a) \in \mathrm{End}(\mathcal{B})$ being the splitting homomorphism of (66.253).

Proposition 66.82.
*Let (M, ω, s) be a ESSS and $\omega = \Omega = \delta\xi$ the symplectic form of the corresponding ESET. The **mid-point map** $M \to M$, $x \mapsto x/2$ defined by*

$$s_{x/2}o = x$$

is globally defined if and only if ϕ is a diffeomorphism.

Notice that the affirmation $\omega = \Omega = \delta\xi$ means that one has a symplectic form $\omega \colon \mathcal{A} \times \mathcal{L} \to \mathbb{R}$,

$$\omega(a, l) = \Omega(a, l) = -\xi\big([a, l]\big).$$

66.8.7 Kähler structures

DefKONtphK

Definition 66.83.
If M posses an almost complex structure[20] *J and a Riemannian metric g, we say that the metric is **hermitian** when*

$$g(JX, JY) = g(X, Y). \tag{66.257}$$

Notice that a symmetric space must be hermitian (definition 66.62), hence equation (66.226c), implies the integrability condition.

Definition 66.84.
*If one has an almost complex structure with an hermitian metric such that $\nabla J = 0$, then (M, J, g) is a **Kähler manifold**.*

20. Definition 66.61.

Remarque 66.85.
$\nabla J = 0$ reads $\forall X, Y \in M$,

$$(\nabla_X J)(Y) = \nabla_X(JY) - J(\nabla_X Y) = 0.$$

Remarque 66.86.
By "∇" we mean the Levi-Civita connection for g. In particular it is torsion free:

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

Lemma 66.87.
If (M, g, J) is a Kähler manifold, then J is integrable.

Proof. From the formula $\nabla_Z(JY) = (\nabla_Z J)Y + J\nabla_Z Y$ and the fact that $\nabla J = 0$, we know that

$$\nabla_Z(JY) = J(\nabla_Z Y), \tag{66.258}$$

while the torsion-free condition for ∇ gives

$$\nabla_X Y - \nabla_Y X = [X, Y]. \tag{66.259}$$

With these two, we find $\nabla_Z(JY) - \nabla_Y(JZ) = J\nabla_Z Y - J\nabla_Y Z = J[Z, Y]$. Writing it with JZ instead of Z,

$$\nabla_{JZ}(JY) + \nabla_Y(Z) = J[JZ, Y].$$

The anti-symmetric part of this equation gives

$$\nabla_{JZ}(JY) + \nabla_Y Z - J[JZ, Y] - \nabla_{JY}(JZ) - \nabla_Z Y + J[JY, Z].$$

Using (66.258) and (66.259), one finds the thesis.

□

When (M, g, J) is a Kähler manifold, one defines the **Kähler 2-form** by

$$\omega(X, Y) := g(X, JY). \tag{66.260}$$

Proposition 66.88.
The Kähler 2-form is a symplectic structure on M.

Proof. Since g is nondegenerate and $JX = 0$ implies $X = 0$, it is clear that ω is nondegenerate. The antisymmetry of ω is because the metric is hermitian. The only point is to see that $d\omega = 0$.

From (46.5) which gives $d\omega$ in terms of $\nabla\omega$, we see that we just have to prove that $\nabla\omega = 0$. By definition,

$$\begin{aligned}(\nabla_Z\omega)(X, Y) &= Z(\omega(X, Y)) - \omega(\nabla_Z X, Y) - \omega(X, \nabla_Z Y) \\ &= Zg(X, JY) - g(\nabla_Z X, JY) - g(X, J\nabla_Z Y). \\ &= (\nabla_Z g)(X, JY) = 0\end{aligned}$$

because the vanishing of ∇J implies that $J(\nabla_Z Y) = \nabla_Z(JY)$. □

66.8.8 Symplectic structure on the Iwasawa component

The Iwasawa theorem gives us a global diffeomorphism between $R = AN$ and $M = G/K$ by $\varphi\colon R \to G/K$, $\varphi(an) = [an]$. But one has a symplectic form on M: $\omega^M_x(X, Y) = g_x(JX, Y)$. So, R has also a symplectic form defined by, $\forall\, X, Y \in T_{an}R$,

$$\omega^R = \varphi^*\omega^M, \tag{66.261}$$

or more explicitly:

$$\omega^R_{an}(X, Y) = \omega^M_{\varphi(an)}(d\varphi_{an}X, d\varphi_{an}Y).$$

Proposition 66.89.
This symplectic form is R-invariant under the left action; in other words, $\forall r \in R$,

$$\omega^R_{ran}\Big((dL_r)_{an}X, (dL_r)_{an}Y\Big) = \omega^R_{an}(X,Y). \tag{66.262}$$

Proof. For a $r \in R$, we want to looks at

$$\omega^R_{ran}(dL_rX, dL_rY) = \omega^M_{[ran]}(d\varphi_{ran}dL_rX, d\varphi_{ran}dL_rY) \tag{66.263}$$

But we know the invariance of ω^M:

$$\omega^M_{[hg]}(dL_hX, dL_hY) = \omega^M_{[g]}(X,Y),$$

Now, let us show that $d\varphi_{ran}dL_rX = dL_rd\varphi_{an}X$. For this, we consider a path which gives $X \in T_{an}R$: $X(t) \in R$, $X(0) = an$. So,

$$d\varphi_{ran}(dL_r)_{an}X = \frac{d}{dt}\Big[[rX(t)]\Big]_{t=0} = \frac{d}{dt}\Big[L_r[X(t)]\Big]_{t=0} = (dL_r)_{[an]}d\varphi_{an}X. \tag{66.264}$$

Finally,

$$\begin{aligned}\omega^R_{ran}(dL_rX, dL_rY) &= \omega^M_{[ran]}(d\varphi_{ran}(dL_r)_{an}X, d\varphi_{ran}(dL_r)_{an}Y)\\ &= \omega^M_{[ran]}((dL_r)_{[an]}d\varphi_{an}X, (dL_r)_{[an]}d\varphi_{an}Y)\\ &= \omega^M_{[an]}(d\varphi_{an}X, d\varphi_{an}Y)\\ &= \omega^R_{an}(X,Y).\end{aligned} \tag{66.265}$$

□

66.8.9 Iwasawa coordinates

We consider $M = G/K$, an hermitian symmetric space (irreducible of non-compact type [21]). Let us consider a $Z \in \mathcal{Z}(\mathcal{K})$ as before: $\delta B(Z,.)|_{\mathcal{P}\times\mathcal{P}}$ is a $\mathcal{K}$-invariant 2-form on $\mathcal{P}$. There are some remarkable spaces: $\mathcal{R} = \mathcal{A}\oplus\mathcal{N}$, the Lie algebra of $R = AN$; $\mathcal{O} = \mathrm{Ad}(G)Z \subset \mathcal{G}$. We consider the following diffeomorphism:

$$\mathcal{I}\colon \mathcal{A}\oplus\mathcal{N} \to R, \qquad \mathcal{I}(a,n) = e^ae^n, \tag{66.266a}$$
$$\varphi\colon R \to M, \qquad \varphi(an) = [an], \tag{66.266b}$$
$$\phi\colon \mathcal{R} \to \mathcal{O}, \qquad \phi(r) = \mathrm{Ad}(\mathcal{I}(r))Z, \tag{66.266c}$$
$$\lambda\colon \mathcal{O} \to M, \qquad \lambda(\mathrm{Ad}(g)Z) = [g]. \tag{66.266d}$$

Note that $\mathcal{R} = T_eR = \mathcal{A}\oplus\mathcal{N} = T_r\mathcal{R}$ where $\mathcal{R} = T_r\mathcal{R}$ is a standard identification of vector spaces. The symplectic forms on these spaces are naturally defined by

$$\omega^R = \varphi^{*-1}\omega^M \tag{66.267a}$$
$$\omega^{\mathcal{R}} = \mathcal{I}^*\omega^R \tag{66.267b}$$
$$\omega^{\mathcal{O}} = \lambda^*\omega^R \tag{66.267c}$$

By the way, the diffeomorphism $\mathcal{I}$ is called the **Iwasawa coordinates**.

Proposition 66.90.
The map ϕ is bijective.

21. je ne sais pas ce que ça veut dire, mais je ne sais pas non plus où on l'utilise. (p. 301 d'Helgason)

Proof. For the surjective condition, we have to obtain $\mathrm{Ad}(ank)Z$ under the form $\mathrm{Ad}(\mathcal{I}(A,N))Z$. For this, remark that one can find $K \in \mathcal{K}$, $A \in \mathcal{A}$, $N \in \mathcal{N}$ such that $k = e^K$, $a = e^A$,$n = e^N$, then

$$\mathrm{Ad}(e^A e^N)Z = \mathrm{Ad}(e^A e^N e^K)Z = \mathrm{Ad}(ank)Z.$$

In order to see the injective condition, let us consider $r, r' \in \mathcal{A} \oplus \mathcal{N}$ such that

$$\mathrm{Ad}(\mathcal{I}(r))Z = \mathrm{Ad}(\mathcal{I}(r'))Z.$$

Then, $\mathrm{Ad}(\mathcal{I}(r'))^{-1} \circ \mathrm{Ad}(\mathcal{I}(r)) = \mathrm{Id}$. This makes $\mathrm{Ad}(e^{-N'}e^{-A'}e^A e^N) = id$, so that

$$e^{-N'}e^{-A'} = \left(e^A e^N\right)^{-1},$$

but exp is a diffeomorphism, then $(A,N) = (A',N')$.

□

lem:om_O_om_R

Lemma 66.91.
The symplectic forms $\omega^{\mathcal{R}}$ and $\omega^{\mathcal{O}}$ are related by

$$\omega^{\mathcal{O}} = (\phi^{-1})^* \omega^{\mathcal{R}}. \tag{66.268}$$

Proof. The definitions make that

$$(\phi^{-1})^*\omega^{\mathcal{R}} = (\phi^{-1})^*\mathcal{I}^*\omega^R = (\phi^{-1})^*\mathcal{I}^*\varphi^*\omega^M, \tag{66.269}$$

so that we just need to see that $\varphi \circ \mathcal{I} \circ \phi^{-1} = \lambda$. This is true because for any $g \in K$,

$$\varphi \circ \mathcal{I} \circ \phi^{-1}(\mathrm{Ad}(g)Z) = \varphi \circ \mathcal{I}(\mathcal{I}^{-1}(g)) = \varphi(g) = [g].$$

□

Now, consider $u \in T_r\mathcal{R}$, with $r = a + n \in \mathcal{A} \oplus \mathcal{N}$, and (just for fun) let us compute $d\phi_r(u)$. In the following computation, u_A and u_N denotes the unit vectors in the direction of $\mathcal{A}$ and $\mathcal{N}$.

$$\begin{aligned}
d\phi_r(u) &= \frac{d}{dt}\Big[\,\mathrm{Ad}(e^{a+tu_A}e^{n+tu_N})Z\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\,\mathrm{Ad}(e^{tu_A})\,\mathrm{Ad}(e^{an})Z\Big]_{t=0} \\
&\quad + \mathrm{Ad}(e^a)\frac{d}{dt}\Big[\,\mathrm{Ad}(e^n)\,\mathrm{Ad}(e^{-n})\,\mathrm{Ad}(e^{n+tu_N})Z\Big]_{t=0} \\
&= -(u_A^*)_{\phi(r)} \\
&\quad + \mathrm{Ad}(e^a e^n)\frac{d}{dt}\Big[\,\mathrm{Ad}(e^{-n})\,\mathrm{Ad}(e^{n+tu_N})Z\Big]_{t=0} \\
&= -(u_A^*)_{\phi(r)} + \mathrm{Ad}(\mathcal{I}(r))\frac{d}{dt}\Big[\,\mathrm{Ad}(e^{CBH(-n,n+tu_N)})Z\Big]_{t=0},
\end{aligned} \tag{66.270}$$

where CBH denote the Campbell-Baker-Hausdorff function defined by

$$e^x e^y = e^{CBH(x,y)}.$$

One maybe knows the formula

$$\frac{d}{dt}\Big[CBH(-n, n+tu_N)\Big]_{t=0} = F(\mathrm{ad}(n))u_N, \tag{66.271}$$

where $F(\mathrm{ad}(n))$ is defined by the expansion of

$$F(z) = \frac{1-e^{-z}}{z}$$

for $z \in \mathbb{C}$ [22]. Finally,

$$d\phi_r(u) = -(u_A^*)_{\phi(r)} - (\mathrm{Ad}(\mathcal{I}(r))F(\mathrm{ad}(n))u_N)^*_{\phi(r)}\,. \tag{66.272}$$

Now, remark that $\mathrm{Ad}(e^a)|_{\mathcal{A}} = id|_{\mathcal{A}}$ because $\mathrm{ad}\, a|_{\mathcal{A}} = 0$ ($\mathcal{A}$ is abelian) and the famous lemma 53.76.

We have [23], $\omega_x^{\mathcal{O}}(X^*, Y^*) = -B(x, [X, Y])$ for $x \in \mathcal{O}$, X, $Y \in \mathcal{G}$. The lemma 66.91 gives us immediately

$$(\mathcal{I}^*\omega^R)_r(u, v) = (\phi^*\omega^{\mathcal{O}})_r(u, v).$$

66.8.10 Summary of the construction

We pick [24] $Z \in \mathcal{Z}(\mathcal{K})$. Then one defines

$$J = \mathrm{ad}(Z)|_{\mathcal{P}}$$

and

$$\omega^M(X, y) = \begin{cases} B(JX, Y) & \text{If } X, Y \in \mathcal{P} \\ 0 & \text{if } X \text{ or } Y \text{ belong to } \mathcal{K}. \end{cases}$$

The maps $\mathcal{I}$, φ, ϕ and λ between spaces R, $\mathcal{A} \oplus \mathcal{N}$, M, $\mathcal{R}$ and $\mathcal{O}$ are designed to propagate the symplectic form from ω^M to $\omega^{\mathcal{R}}$, ω^R, $\omega^{\mathcal{O}}$. The group R acts on each of these spaces and in particular on $\mathcal{O}$ by the adjoint action. One can prove that $\omega^{\mathcal{O}} := \lambda^*(\varphi^{-1})^*\omega^M$ is

$$\omega_x^{\mathcal{O}}(X^*, Y^*) = -B(x, [X, Y])$$

for all X, $Y \in \mathcal{R}$. In the whole construction, σ is the Cartan involution which gives the decomposition

$$\sigma = \mathrm{Id}\,|_{\mathcal{K}} \oplus (-\mathrm{Id})|_{\mathcal{P}}.$$

Therefore $\sigma E = -E$ because $E \in \mathcal{N} \subset \mathcal{P}$.

The Lie algebra $\mathcal{G}$ possesses two roots: α and 2α, so we decompose it as [25]

$$\mathcal{G} = \mathcal{A} \oplus \mathcal{G}_\alpha \oplus \mathcal{G}_{2\alpha}.$$

We pick $A \in \mathcal{A}$, $y \in \mathcal{G}_\alpha$ and $E \in \mathcal{G}_{2\alpha}$. For example, if $B \in \mathcal{A}$, $[B, y] = \alpha(A)y$.

66.8.11 Continuation

Proposition 66.92.
Let $M = G/K$ be an hermitian irreducible symmetric space of non compact type. We suppose that $\dim \mathcal{P} \geqslant 4$. We consider the action $\tau\colon \mathcal{R} \times \mathcal{R} \to \mathcal{R}$,

$$\tau_g(X) = \mathcal{I}^{-1}(g\mathcal{I}(X)).$$

This action is Hamiltonian for the constant symplectic structure Ω on $\mathcal{R}$ and the dual momentum maps are given by

$$\lambda_A(X) = 2\alpha(A)B(\sigma A, E)n_E \qquad (A \in \mathcal{A} \tag{66.273a}$$

$$\lambda_y(X) = e^{-\alpha(a)}\Omega(n, y) \qquad (y \in \mathcal{G}_\alpha \tag{66.273b}$$

$$\lambda_E(X) = e^{-Z\alpha(a)}B(\sigma E, E) \tag{66.273c}$$

where $X = (a, n)$ and $n = n_\alpha + n_E E$ for the decomposition $\mathcal{N} = \mathcal{G}_\alpha \oplus \mathbb{R}E$.

As a consequence, the Moyal star product is R-covariant.

22. Il faut encore aller voir dans Duitsermaat les tenants et aboutissants de ce truc.
23. From proposition 1.1 page 5 in BM
24. L'existence de ce $\mathcal{K}$ contre-dit ce que je dis quand une autre question à propos du type non-compact
25. Le fait d'être de type non compact est peut-être l'absence de composante K pour l'Iwasawa, qu'en penses-tu ?

Proof. From equation (59.10), we have to prove the identities

$$\{\lambda_X, \lambda_Y\} = X^*(\lambda_Y).$$

We begin by proving the identity

$$\{\lambda_A, \lambda_y\}(L) = A_L^*(\lambda_y)$$

where $L = (a', n') \in \mathcal{R}$. In these coordinate, we suppose without loss of generality that $A = (1, 0)$. As usual, we will use some abuse of notation as $\mathcal{I}(L) = e^{a'}e^{n'} = e^{a'A}e^{n'}$;

$$\begin{aligned} A_L^*(\lambda_y) &= \frac{d}{dt}\Big[\lambda_y(\tau_{e^{-tA}}L)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[(\lambda_y \circ \mathcal{I}^{-1})e^{(a'-t)A}e^{n'}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\lambda_y\big((a'-t), n'\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[e^{-\alpha(a'-t)}\Omega(n', y)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[e^{(t-a')\alpha(A)}\Big]_{t=0}\Omega(n', y) \\ &= \alpha(A)e^{-\alpha(a')\Omega(n',y)}. \end{aligned} \tag{66.274}$$

In this computation, we used the fact that $\alpha(a' - t) = (a' - t)\alpha(A)$. On the other hand,

$$\lambda_{[A,y]}(L) = \alpha(A)\lambda_y(L) = \alpha(A)e^{-\alpha(a')\Omega(n',y)}.$$

This concludes the first check. The check that $\{\lambda_A, \lambda_E\} = \lambda_{[A,E]}$ is the same, using the fact that $E \in \mathcal{G}_{2\alpha}$ and thus that $[A, E] = 2\alpha(A)E$. For the third, $[y, E] = 0$ therefore, we have to prove that $\|\lambda_y, \lambda_E\| = 0$. We have

$$\|\lambda_y, \lambda_E\|(L) = y_L^*(\lambda_E) = \frac{d}{dt}\Big[(\lambda_E \circ \mathcal{I}^{-1})\big(e^{-ty}e^{a'}e^{n'}\big)\Big]_{t=0}.$$

The problem is to commute e^{-ty} with $e^{a'}$. Since the t will always stands in front of y and λ_E doesn't depends on y, the derivative is zero [26]. cs

$$\begin{aligned} A_o^* &= \frac{d}{dt}\Big[e^{-tA} \cdot (0, 0)\Big]_{t=0} \\ &= -A. \end{aligned}$$

Since $d\mathcal{I} = \mathrm{Id}$,

$$\begin{aligned} \omega^{\mathcal{R}}(A^*, X) =& \omega^{\mathcal{R}}(A, x_y y + x_E E) \\ =& -\omega^R(A, x_y y + x_E E), \end{aligned}$$

but for any element in $\mathcal{A} \oplus \mathcal{N}$, via the identification $\mathcal{R} = [\mathcal{A} \oplus \mathcal{N}]$ (the additive class),

$$\begin{aligned} d\lambda^{-1}A &= \frac{d}{dt}\Big[\lambda^{-1}[e^{tA}]\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\operatorname{Ad}(e^{tA})Z\Big]_{t=0} \\ &= -A^*. \end{aligned}$$

Thus

$$\begin{aligned} \omega^{\mathcal{R}}(A^*, X) &= -\omega^{\mathcal{O}}(d\lambda^{-1}A, d\lambda^{-1}(x_y y + x_E E)) \\ &= -\omega^{\mathcal{O}}(A^*, x_y y^* + x_E E^*) \end{aligned} \tag{66.275}$$

eq_omeQmQ

26. Je ne crois pas que cette justification soit juste.

where A^*, y^* and E^* are taken in the sense of the adjoint action of R on $\mathcal{O}$.

Now we prove that λ_A is well a dual momentum map. For this, we choose $X = x_A A + x_y y + x_E E \in \mathcal{R}$ and we check the identity $d\lambda_A X = \omega^{\mathcal{R}}(A^*, X)$ where A^* stands for the given action of R on $\mathcal{R}$.

A question arise: at which point is taken $\omega^{\mathcal{O}}$ in equation (66.275)? Since we compute $\omega^{\mathcal{R}}$ at identity, we compute $\omega^{\mathcal{O}}$ at Z. So

$$\begin{aligned}\omega^{\mathcal{R}}(A^*, X) &= -\omega^{\mathcal{O}}_Z(A^*, x_y y^* + x_E E^*)\\ &= -B(Z, [A, y + E])\\ &= -\alpha(A)B(Z, y) - 2\alpha(A)B(Z, E).\end{aligned}$$

Here, we have to remark that it is not zero because $\mathcal{N}$ is not included in $\mathcal{P}$, but is transverse.

□

66.9 Elementary normal symplectic spaces

SecElemNormSymplSpace

This section is closely related to the Pyatetskii-Shapiro theory treated in section 66.10. See [892] as reference.

Let (V, Ω) be a symplectic real vector space of dimension $2n$. We build the **Heisenberg algebra** by

$$\mathscr{H} = V \oplus \mathbb{R}E \tag{66.276}$$

with the relation $[v, v'] = \Omega(v, v')E$. Now we consider a new element H and $\mathfrak{a} = \mathbb{R}H$ and the split extension

$$0 \to \mathscr{H} \to \mathfrak{s} \to \mathfrak{a} \to 0 \tag{66.277}$$

where $\mathfrak{s} = \mathfrak{a} \oplus_\rho \mathscr{H}$ and $\rho\colon \mathfrak{a} \to \mathtt{Der}(\mathscr{H})$ is given by

$$\rho(H)(v \oplus tE) = [H, v \oplus tE] = v \oplus 2tE. \tag{66.278}$$

We denote by (a, v, t) an element of $\mathfrak{s}$, that is

$$(a, v, t) = aH + v + tE \tag{66.279}$$

with $a, t \in \mathbb{R}$ and $v \in V$. We consider the 2-form

$$\omega^{\mathcal{S}} = 2da \wedge dt + \Omega, \tag{66.280}$$

the pair $(\mathfrak{s}, \omega^{\mathfrak{s}})$ is said to be a **normal elementary symplectic algebra**. We denote by $(\mathbb{S}, \omega)$ the associated connected simply connected Lie group.

Prop2807DescSMdarboux

Proposition 66.93.
With the previous notations we have

(1) The map

$$\begin{aligned}(\mathfrak{s}, \omega^{\mathfrak{s}}) &\to (\mathbb{S}, \omega)\\ (a, v, t) &\mapsto e^{aH}e^{v+tE}\end{aligned} \tag{66.281}$$

is a global Darboux chart (in particular it is a global diffeomorphism).

By this diffeomorphism we identify $\mathfrak{s}$ and $\mathbb{S}$, i.e. we will denote by (a, v, t) the element $e^{aH}e^{v+tE} \in \mathbb{S}$ as well as the element $aH + v + tE \in \mathfrak{s}$.

(2) Within the coordinates (a, v, t) the group law is given by

$$(a, v, t) \cdot (a', v', t') = (a + a', e^{-a'}v + v', e^{-2a,}t + t' + \frac{1}{2}e^{-a'}\Omega(v, v')). \tag{66.282}$$

(3) If we define

$$s_{(a,v,t)}(a',v',t') = \big(2a-a', 2\cosh(a-a')v - v', 2\cosh(2(a-a'))t + \Omega(v,v')\sinh(a-a') - t'\big), \tag{66.283}$$

Eq1807StuctSymM

the space $\mathbb{M} = (\mathbb{S}, \omega, \mathfrak{s})$ *becomes a symplectic symmetric space.*

(4) The structure of symplectic symmetric space is preserved by the left translations. In other words, for every $x \in \mathbb{S}$ *we have* $L_x \in \mathrm{Aut}(\mathbb{M})$*. And the subgroup* $\{L_x \text{ tel que } x \in \mathbb{S}\}$ *acts simply transitively on* $\mathbb{M}$*.*

(5) We have

$$\mathrm{SP}(V,\Omega) \subset \mathrm{Aut}(\mathbb{M}) \tag{66.284}$$

if we define $g \cdot (a,v,t) = (a, g\cdot v, t)$ *for every* $g \in \mathrm{SP}(V,\Omega)$*.*

66.10 Pyatetskii-Shapiro structure theorem

SecPyateskiiShapiro

Definition 66.94.
A ***normal*** j***-algebra*** *is a triple* $(\mathfrak{s}, \alpha, j)$ *where*

(1) the Lie algebra $\mathfrak{s}$ *is solvable and such that* $\mathrm{ad}(X)$ *has only real eigenvalues for every* $X \in \mathfrak{s}$*,*

(2) the map $j\colon \mathfrak{s} \to \mathfrak{s}$ *is an endomorphism of* $\mathfrak{s}$ *such that* $j^2 = -1$ *and*

$$[X,Y] + j[jX,Y] + j[X,jY] - [jX,jY] = 0 \tag{66.285}$$

for every $X, Y \in \mathfrak{s}$*,*

(3) α *is is a linear form on* $\mathfrak{s}$ *such that*

(3a) $\alpha\big([jX,X]\big) > 0$ *if* $X \neq 0$*,*

(3b) $\alpha\big([jX,jY]\big) = \alpha\big([X,Y]\big)$*.*

If $\mathfrak{s}'$ is a subalgebra of $\mathfrak{s}$ which is invariant under j, then the triple $(\mathfrak{s}', \alpha|_{\mathfrak{s}'}, j|_{\mathfrak{s}'})$ is a also normal j-algebra and is said to be a **normal j-subalgebra** of $\mathfrak{s}$.

A normal j-algebra has a real inner product defined by the formula

$$g(X,Y) = \alpha\big([jX,Y]\big). \tag{66.286}$$

If $\mathfrak{g}$ is an Hermitian Lie algebra [27], we can build a normal j-algebra out of $\mathfrak{g}$ in the following way. First, we choose an Iwasawa decomposition

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{n} \oplus \mathfrak{k}, \tag{66.287}$$

EqDecIwalGi

and we pick $\mathfrak{s} = \mathfrak{a} \oplus \mathfrak{n}$. Let $G = ANK$ be the group associated with the Iwasawa decomposition (66.287). The manifold $M = G/K$ is an Hermitian symmetric space, and we have a global diffeomorphism

$$\begin{aligned} R = AN &\to G/K \\ g &\mapsto gK \end{aligned} \tag{66.288}$$

which endows the group R with an exact left invariant symplectic structure and a compatible complex structure, see section 66.8. We define α by $\Omega_e = d\alpha$ (Ω is exact) and j is the complex structure evaluated at identity.

A normal j-algebra build from an Hermitian symmetric space of rank 1 (i.e. $\dim \mathfrak{a} = 1$.) is **elementary**. Elementary normal j-algebra are well understood by the following proposition.

27. i.e. the center of its maximal compact is one dimensional.

PropStructNormalElementaireJalg

Proposition 66.95.
An elementary normal j-algebra is a split extension

$$\mathfrak{s}_{el} = \mathfrak{a}_1 \oplus_{\text{ad}} \mathfrak{n}_1 = \mathfrak{a}_1 \oplus_{\text{ad}} \big(V \oplus \mathfrak{z}_1\big) \tag{66.289}$$

EqDecoEleJal

where $\mathfrak{n}_1$ is an Heisenberg algebra $\mathfrak{n}_1 = V \oplus \mathfrak{z}_1$ and $\mathfrak{a}_1$ is one dimensional. Moreover, V is a symplectic vector space and one can choose $H \in \mathfrak{a}_1$ and $E \in \mathfrak{z}_1$ in such a way that

$$\begin{aligned} [H, v] &= v, \\ [v, v'] &= \Omega(v, v')E, \\ [H, E] &= 2E. \end{aligned} \tag{66.290}$$

EqRelColNormaliAlg

Any normal j-algebra is build from elementary normal j-algebras by mean of the following lemma.

PropStructNormalJalg

Proposition 66.96.
Let $(\mathfrak{s}, \alpha, j)$, a normal j-algebra and $\mathfrak{z}_1$, a one dimensional ideal of $\mathfrak{s}$.

(1) There exists a vector space V such that

$$\mathfrak{s}_1 = j\mathfrak{z}_1 + V + \mathfrak{z}_1 \tag{66.291}$$

is an elementary normal j-algebra, and such that $\mathfrak{s}$ is a split extension

$$\mathfrak{s} = \mathfrak{s}' \oplus_{\text{ad}} \mathfrak{s}_1 \tag{66.292}$$

EqDecNormale

where $\mathfrak{s}'$ is, itself, a normal j-algebra.

(2) If $\mathfrak{s}_1 = \mathfrak{a}_1 \oplus_{\text{ad}} (V \oplus \mathfrak{z}_1)$, then

$$j\mathfrak{z}_1 + \mathfrak{z}_1 = \mathfrak{a}_1 \oplus \mathfrak{z}_1 \tag{66.293}$$

and

$$\begin{aligned} [\mathfrak{s}', \mathfrak{a}_1 \oplus \mathfrak{z}_1] &= 0, \\ [\mathfrak{s}', V] &\subset V. \end{aligned} \tag{66.294}$$

(3) Such an ideal $\mathfrak{z}_1$ exists in every normal j-algebra.

Let us see what are the possibilities for j. If $jE = aH + b_i v_i + cE$, then

$$[jE, E] = 2aE. \tag{66.295}$$

We can prove that $a \neq 0$. Indeed, if $a = 0$, then $jE = cE$ and $-E = j^2E = cjE = c^2E$.

Now, we use the following Jacobi identity on $[H, [jE, v]]$ and the commutation relations, we find $b_i = 0$. Now, suppose that $jH = a'H + b'_i + c'E$. In that case,

$$-E = j^2E = j(aH + cE) = aa'H + ab'_i v_i + ac'E + caH + c^2E. \tag{66.296}$$

Since $a \neq 0$, we have $b'_i = 0$. So we have

$$\begin{aligned} jE &= aH + cE \\ jH &= a'H + c'E. \end{aligned} \tag{66.297}$$

Expressing that $j^2E = -E$ and $j^2H = -H$, we find the following constrains on the coefficients:

$$\begin{aligned} aa' + ca &= 0 \\ ac' + c^2 &= -1 \\ c'^2 + c'a &= -1 \\ c'c + c'c &= 0. \end{aligned} \tag{66.298}$$

We check that $a \neq 0$, $c' \neq 0$ and $a' = -c$. The remaining relation is $c^2 + c'a = -1$. Thus in the basis $\{H, E\}$, the endomorphism j reads

$$j = \begin{pmatrix} -c & a \\ c' & c \end{pmatrix} \tag{66.299}$$

with $\det j = 1$.

Lemma 66.97.
An elementary normal j-algebra has no proper j-ideal.

Proof. Let $\mathfrak{i}$ be a j-ideal of the elementary normal j-algebra $\mathfrak{s}_{el}$. Let $\mathfrak{s}_{el} = \mathfrak{a} \oplus_{\text{ad}} (V \oplus \mathfrak{z})$. We denote by H and E the elements of $\mathfrak{a}$ and $\mathfrak{z}$ (which are one dimensional) who fulfill the standard relations (66.290). If $X = aH + b_i v_i + cE \in \mathfrak{i}$, then $\big[[X, v], v\big] \in \mathfrak{i}$. Using the relations, we conclude that $\mathfrak{z} \subset \mathfrak{i}$. By j-invariance of $\mathfrak{i}$, we have $j\mathfrak{z} \subset \mathfrak{i}$. Now, the fact that $[jE, v] = av$ implies that $\mathfrak{i} = \mathfrak{s}_{el}$. □

The structure of a normal j-algebra $\mathfrak{s}$ is thus as follows. We have the decomposition

$$\mathfrak{s} = \mathfrak{s}' \oplus_{\text{ad}} \Big(\mathfrak{a}_1 \oplus_{\text{ad}} (V_1 \oplus \mathfrak{z}_1) \Big) \tag{66.300}$$

where $\mathfrak{s}'$ is again a normal j-algebra. Furthermore, $\dim \mathfrak{a}_1 = \dim \mathfrak{z}_1 = 1$ and we can choose a basis $H \in \mathfrak{a}_1$, $E \in \mathfrak{z}_1$ such that

$$\begin{aligned} [H, v] &= v \\ [H, E] &= 2E \\ [v, v'] &= \Omega(v, v')E \\ [\mathfrak{s}', V] &\subset V \\ [\mathfrak{s}', \mathfrak{a}_1 \oplus \mathfrak{z}_1] &= 0. \end{aligned} \tag{66.301}$$

for all $v, v' \in V_1$. The algebra $V_1 \oplus \mathfrak{z}_1$ is an Heisenberg algebra.

The algebra $\mathfrak{s}'$ can be decomposed in the same way again and again up to end up with a sequence of elementary normal j-algebra.

Chapter 67

Heat kernel expansions

67.1 Mellin transform

If f is a function, we define its **Mellin transform** by

$$(Mf)(s) = \int_0^\infty x^s f(x)\frac{dx}{x} = \phi_f(s), \tag{67.1}$$

and the inverse transform is given by

$$(M^{-1}\phi_f)(x) = f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s}\phi_f(s)ds \tag{67.2}$$

where the integral is taken on a vertical line of the complex plane.

67.2 General setting

Matter about heat kernel expansions, boundary conditions and related physical consequences can be found in [962, 963, 964, 965]. Let M be a compact Riemannian manifold without boundary. A **heat kernel** is a function $K \in C^\infty\big((0,\infty)\times M\times M\big)$ which satisfies the equation

$$(\partial_t - \Delta_x)K(t,x,y) = 0 \tag{67.3}$$

with the initial condition

$$\lim_{t\to 0^+} K(t,x,y) = \delta_y(x) \tag{67.4}$$

EqCondLimoplusdelta

where Δ_x is the **Laplace-Beltrami** operator acting on the variable x. That operator is defined, in local coordinates, by the expression

$$\Delta = \frac{1}{\sqrt{g}}\partial_i\big(\sqrt{g}g^{ij}\partial_j\big)$$

The limit in the left hand side of condition (67.4) means that for every smooth function f on M, the function

$$u(t,x) = \int_M K(t,x,y)f(y)\sqrt{g}dy$$

is everywhere (eventually not at $t=0$) and satisfies $u(0,x) = f(x)$.

Theorem 67.1.
Heat kernel have the following properties:

(1) There is one and only one heat kernel on a given manifold,

(2) for each $x \in M$ there is an asymptotic expansion

$$K(t,x,x) \sim t^{-m/2}\big(a_0(x) + a_1(x)t + a_2(x)t^2 + \cdots\big)$$

when $t \to 0$,

(3) the coefficients a_j are smooth functions over M and $a_j(x)$ is determined by the metric and its derivatives at x.

67.2.1 Mellin

Let D be any operator of order q an (x_i), local coordinates on M. We write $D = \sum_I D_I \partial_I$ where I runs over multi-indices. If D has order 2, we have $D(x_j\partial_j f) = 2Df + D_{ik}x_j\partial_{ijk}f$, and in general one can check that

$$\sum_i [D, x_j]\partial_j = qD + R_1$$

where q is the order of D and R_1 is of order less or equal to $q-1$. One also can show that $(n+q)D = \sum_j [\partial_j D, x_j] + R_2$ where the order of R_2 is less than q and $n = \dim M$. Since the trace of a commutator is always zero, we find

$$(n+q)\,\mathrm{Tr}(D) = \mathrm{Tr}(R_2). \tag{67.5}$$

Let now D_z be a pseudo differential operator of order $\mathrm{Re}(z)$. Then we have

$$\mathrm{Tr}(D_z) = \frac{1}{(z+n)}\,\mathrm{Tr}(R_z)$$

with R_z being an operator of order less or equal to $\mathrm{Re}(z)-1$. Notice that $\mathrm{Tr}(D_z)$ has a pole on $z=-n$. If one continues, one sees that $\mathrm{Tr}(D_z)$ has poles on every negative integer.

67.2.2 Residues and zeta function

The main assumption we make about the operator D is that we have the following asymptotic expansion when t goes to zero $\mathrm{Tr}\left(e^{-tD^2}\right) \sim \sum a_\alpha t^\alpha$. The zeta function is defined by

$$\zeta_D(s) = \mathrm{Tr}\left(|D|^{-s}\right) = \mathrm{Tr}\left(\Delta^{-s/2}\right) \tag{67.6}$$

where $\Delta = D^2$. We suppose that D is invertible.

LemaalphaGammaCo

Lemma 67.2.
Let us suppose that we have an asymptotic expansion of the form

$$\mathrm{Tr}\left(e^{-tD^2}\right) \sim \sum_\alpha a_\alpha t^\alpha \tag{67.7}$$

EqDevHeatDsquare

when $t \to 0$. Then

(1) if $a_\alpha \neq 0$ with $\alpha < 0$, then ζ_D has pole at -2α with residue given by

$$\mathrm{Res}_{s=-2\alpha}\,\zeta_D(s) = \frac{2a_\alpha}{\Gamma(-\alpha)}. \tag{67.8}$$

(2) The absence of term proportional to the logarithm of t provides regularity for ζ_D with

$$\zeta_D(0) + \dim(\ker D) = a_0. \tag{67.9}$$

LemiizetaDzero

¡¡ Avertissement/question au lecteur !! 67.3
Je te signale que la preuve n'est pas complete.

Proof. We know that the complex power of an operator can be written under the form

$$|D|^{-s} = \Delta^{-s/2} = \frac{1}{\Gamma(s/2)}\int_0^\infty e^{-t\Delta}t^{(s/2)-1}dt. \tag{67.10}$$

Taking the trace of both sides ,

$$\begin{aligned}\zeta_D(s) = \mathrm{Tr}\left(|D|^{-s}\right) &= \frac{1}{\Gamma(s/2)}\int_0^\infty \mathrm{Tr}\left(e^{-t\Delta}\right)t^{(s/2)-1}dt \\ &= \sum_\alpha \frac{1}{\Gamma(s/2)}\int_0^\infty a_\alpha t^\alpha t^{(s/2)-1}dt.\end{aligned}$$

We can compute the integral using the formula

$$\int_0^1 t^{\alpha+\frac{s}{2}-1}dt = \Big(\alpha+\frac{s}{2}\Big)^{-1}$$

and find

$$\zeta_D(s) = \sum_\alpha \frac{1}{\Gamma(s/2)}\Big(\alpha+\frac{s}{2}\Big)^{-1} a_\alpha. \tag{67.11}$$

Now we have to compute

$$\mathrm{Res}_{s=-2\alpha}\left(\frac{a_\alpha}{\Gamma\left(\frac{s}{2}\right)\left(\alpha+\frac{s}{2}\right)}\right);$$

using the formula $\mathrm{Res}_{z=a} f(z) = \lim_{z\to a}(z-a)f(z)$ we easily find that it is $2a_\alpha = \Gamma(-\alpha)$. □

67.2.3 Boundary conditions for heat kernel expansions

Let $(M, \partial M)$ be a m dimensional manifold with boundary and N be the inward vector field normal to ∂M. The geodesic flow of N allows to identify $\mathcal{C} = \partial M \times [0, \epsilon[$ to a neighborhood of ∂M in M. If $(y^1, \dots, y^{m-1})$ is a coordinate system on ∂M, we define the coordinate x^m saying that (y, x^m) is the point at geodesic distance x^m of $y \in \partial M$ following the geodesic flow of N. If $\{e_1, \dots, e_{m-1}\}$ is an orthonormal basis of $T(\partial M)$, we denote by $e_m = N = \partial_m$ the last vector of the orthonormal basis $\{e_1, \dots, e_m\}$ of $TM|_{\partial M}$.

Let Δ be a Laplace type operator on $\Gamma(V)$. For the moment, we do not suppose it to be the square of nay Dirac operator. Let χ be an endomorphism of V defined on ∂M such that $\chi^2 = 1$. One can extend χ to $\mathcal{C}$ with the condition $\nabla'\chi = 0$ where ∇' is the unique connection on V and E the unique endomorphism of V such that

$$\Delta = -\big(g^{ij}\nabla'_i\nabla'_j + E\big).$$

We denote by $V_\pm$ the two complementary bundle of V (eigenspaces of χ) and we consider the operators $\Pi_\pm = \frac{1}{2}(1 \pm \chi)$. Let $S \in \mathrm{End}(V_+)$ be an auxiliary endomorphism. We consider the boundary condition for Δ as the operator $\mathcal{B} \in \mathrm{End}\big(C^\infty(V)\big)$,

$$\mathcal{B}s = \Pi_-(s)|_{\partial M} \oplus \Pi_+(\nabla'_m + S)\Pi_+(s)|_{\partial M}. \tag{67.12}$$

The operator we will study is $\Delta_\mathcal{B}$ which is Δ restricted to the domain

$$\{s \in \Gamma(v) \text{ tel que } \mathcal{B}(s) = 0\}.$$

In that case we have an expression

$$\mathrm{Tr}_{L^2}\left(fe^{-t\Delta_\mathcal{B}}\right) \sim \sum_{n=0}^\infty t^{(n-m)/2} a_n(f, \Delta, \mathcal{B}) \tag{67.13}$$

when $t \to 0^+$. If dx and dy are the Riemann measures on M and ∂M, there exists local invariants $a_n(x, \Delta)$ and $a_{n,\nu}(y, \Delta, \mathcal{B})$ such that

$$a_n(f, \Delta, \mathcal{B}) = \int_M a_n(x, \Delta)dx + \int_{\partial M} \sum_{\nu \leqslant n} N^\nu(f) a_{n,\nu}(y, \Delta, \mathcal{B})dy. \tag{67.14}$$

67.2.4 Boundary and Dirac

We use boundary conditions of the form $\mathcal{B}\phi = 0$ where the most frequent operators are

$\mathcal{B}^-\phi = \phi\|_{\partial M}$	Dirichtlet
$\mathcal{B}^+\phi = (\nabla_n\phi + S\phi)$	Neumann

where S is a matrix valued function and ∇_n denotes the covariant derivative in the direction of n in the sense of the Levi-Civita connection on M. Let $\Pi_\pm$ be projections operators on complementary bundle over $V|_{\partial M}$. We have $\Pi_\pm^2 = \Pi_\pm$ because they are projections and $\Pi_+ \oplus \Pi_- = \mathrm{Id}$ because the projectors are complementary. Thus we have the decomposition

$$V_{\partial M} = V_+ \oplus V_-$$

where $V_\pm = \Pi_\pm V|_{\partial M}$. A section of V reads on the boundary $\phi = \phi^+ \oplus \phi^-$. The boundary condition thus reads

$$\mathcal{B}\phi = \phi^- \oplus (\nabla_n \phi^+ + S\phi^+)|_{\partial M} \tag{67.15}$$

where S acts on V_+, in particular $S = \Pi_+ S = S\Pi_+$. These conditions are said to be **mixed boundary conditions**. One often meets the notations V_N and V_D instead of V_+ and V_-.

Let $D = \gamma^i \partial_i - r$ be a Dirac type operator. We want operator χ which appears in the boundary condition to satisfy $\chi^2 = \mathrm{Id}$, $\chi\gamma_m = -\gamma_m\chi$ and $\chi\gamma_a = \gamma_a\chi$ for every $1 \leqslant a \leqslant m-1$. Remark that γ_m is invertible because, from the definition of a Dirac operator, $2\gamma^m\gamma^m = -2g^{mm}\,\mathrm{Id}$, so up to a (non constant) multiple, $\gamma_m^{-1} = \gamma_m$. Then we have $\gamma_m V = V$ and

$$\gamma_m V_+ = \frac{1}{2}(\gamma_m + \gamma_m\chi)V = \frac{1}{2}(\gamma_m - \chi\gamma_m)V = \frac{1}{2}(\mathrm{Id} - \chi)\gamma_m V$$

proves that $\gamma_m V_+ = V_-$. In particular $\dim V_+ = \dim V_-$.

We want to study the boundary problem given by the operator $b_\chi s = \Pi_- s|_{\partial M}$, and $D_\chi = D|_{\{s\in\Gamma(V) \text{ tel que } b_\chi s = 0\}}$. As far as the Laplace type operator $\Delta_\chi = D_\chi^2$ is concerned, we look at the domain $\mathrm{dom}(\Delta_\chi) = \{s \in \Gamma(V) \text{ tel que } \mathcal{B}_\chi s = 0\}$ where

$$\mathcal{B}_\chi s = \Pi_- s|_{\partial M} \oplus \Pi_- Ds|_{\partial M}. \tag{67.16}$$

PropCondBordphiform

Proposition 67.4.
The boundary condition

$$\mathcal{B}_\chi s = \Pi_- s|_{\partial M} \oplus Ds|_{\partial M}$$

has the form

$$\mathcal{B}s = \Pi_- s = \Pi_- s|_{\partial M} \oplus \Pi_+(\nabla'_m + S)\Pi_+ s|_{\partial M}$$

with $S = \Pi_+\big(\gamma_m\Phi - \frac{1}{2}\gamma_m\gamma_a\nabla'_a\chi\big)\Pi_+$ *where* $\Phi \in C^\infty\big(\mathrm{End}(V)\big)$ *is defined by* $D = \gamma_i\nabla'_i - \Phi$.

A proof of this proposition can be found in the lemma 7 of [964].

Here is the main result. Let $D = \gamma^i\nabla_i - \Phi$ with $\nabla_i = \partial_i + \omega_i$, the connection ω being the spin connection torsion free. Let E, ∇' and A be defined by $\Delta = D^2 = -(g^{ij}\nabla'_i\nabla'_j + E)$ with $\nabla'_i = \partial_i + \omega'_i$ and $\omega'_i = \frac{1}{2}g_{ij}\big(A^j + g^{kl}\Gamma^j_{kl}\big)$. In that case, proposition 67.4 applies and we have

$$\nabla'_a\chi = \partial_a\chi + [\omega'_a, \chi] = K_{ab}\chi\gamma^n\gamma^b + [\theta_a, \chi]$$

where $\theta_a = \omega'_a - \omega_a$. In that case, we have the following formulae for the first three Seeley-de Witt coefficients of $\Delta = D^2$:

$$\begin{aligned} a_0(\Delta,\chi) &= \frac{1}{16\pi^2}\int_M \mathrm{Tr}(1)\sqrt{g}d^4x \\ a_1(\Delta,\chi) &= 0 \\ a_2(\Delta,\chi) &= \frac{1}{96\pi^2}\left(\int_M \mathrm{Tr}(6E+R)\sqrt{g}d^4x + \int_{\partial M}\mathrm{Tr}(2K+12S)\sqrt{h}d^3x\right). \end{aligned} \tag{67.17}$$

EqSeeleydeWitt

An example from [965] is given in section 86.5.

Chapter 68

From Clifford algebras to Dirac operator

Bibliography for Clifford algebras, spin group and related topics are [966, 967, 968, 969, 970]. More algebraic point of view can be found in [971, 972]. More details about "square rooting" second order differential operators are in [973]. For physical concerns, the reader should refer to [974, 975, 976].

68.1 Invitation: Clifford algebra in quantum field theory

Secqft

68.1.1 Schrödinger, Klein-Gordon and Dirac

The origin of the Klein-Gordon equation is almost the same as the one of the Schrödinger: one replace physical functions by operators. For a free particle, the correspondence are

$$\begin{array}{cc} \text{energy} & E \to i\hbar\frac{\partial}{\partial t}, \\ \text{momentum} & \overline{p} \to -i\hbar\overline{\nabla}. \end{array}$$

The Schrödinger equation (which is the non relativistic quantum wave equation) comes from replacement in the non non relativistic expression of the Hamiltonian

$$E = \frac{\overline{p}^2}{2m} \longrightarrow \left(\partial_t - \frac{i\hbar}{2m}\nabla^2\right)\psi = 0,$$

while the Klein-Gordon one (which is the relativistic quantum wave equation) comes from the relativistic corresponding equation:

$$E^2 = \overline{p}^2c^2 + m^2c^4 \longrightarrow \left(\partial^\mu\partial_\mu + (\frac{mc}{\hbar})^2\right)\psi = 0.$$

This is a second order differential equation; there are however no "law of nature" which forbid a first order equation. We try

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\frac{\hbar c}{i}\alpha^k\partial_k + \beta mc^2\right)\psi \equiv \hat{H}\psi.$$

There are some physical constraints on the coefficients α^k and β. We will study one of them: we want the components of ψ to satisfy the Klein-Gordon equation, so that the plane waves fulfill the fundamental relation $E^2 = p^2c^2 + m^2c^4$.

In order to see the implications of this constraint on the coefficients, we apply two times the operator $\hat{H}$, and we compare the result with the Klein-Gordon equation. We find:

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij}\mathbb{1}, \tag{68.1a}$$

$$\alpha^i\beta + \beta\alpha^i = 0, \tag{68.1b}$$

$$(\alpha^i)^2 = \beta^2 = \mathbb{1}. \tag{68.1c}$$

If we define $\gamma^0 = \beta$ and $\gamma^i = \beta\alpha^i$, we find that the matrices γ^μ have to give a representation of the Clifford algebra [1]:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1}. \tag{68.2}$$

The Dirac equation reads

$$\left(-i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0.$$

If we want to perform some computation with the quantum field theory, we need an explicit form for the γ's; that's the reason why we study representations of the Clifford algebra. The **Dirac operator** $\mathcal{D}$ is the operator which lies in the Dirac equation:

$$\mathcal{D} = \sum_{\mu=0}^{3} \gamma^\mu \frac{\partial}{\partial x^\mu}. \tag{68.3}$$

Definition 68.1.
*Let V be a (finite dimensional) vector space and q, a bilinear quadratic form over V. The **Clifford algebra** $\mathrm{Cl}(V,q)$ is the unital associative algebra generated by V subject to the relation*

$$v \cdot v = q(v) \tag{68.4}$$

for all v in $\mathrm{Cl}(V,q)$. Here the dot denotes the algebra product and $q(v)$ means $q(v,v)$.

Theorem proves 51.62 proves unicity of such an algebra, so that it makes sense.

Remarque 68.2.
The relation (68.4) is no more a restriction for the elements in $\mathrm{Cl}(V,q)$ than a restriction on the choice of the algebra product.

An explicit construction of $\mathrm{Cl}(V,q)$ can be achieved in the following way. On the tensor algebra [2] $T(V)$, we consider the two-sided ideal $\mathcal{I}$ generated by elements of the form $v\otimes v - q(v)1$. The **Clifford algebra** for (V,q) is given by

$$\mathrm{Cl}(p,q) := T(V)/\mathcal{I} \tag{68.5}$$

in which product of $\mathrm{Cl}(V,q)$ is naturally defined by $[a]\otimes[b] = [a\otimes b]$ if $[a]$ is the class of $a \in T(V)$.

Let us now fix some notations more adapted to what we want to do. Let $V = \mathbb{R}^{p,q}$ the vector space $\mathbb{R}^{p+q}$ endowed with a diagonal metric which contains p plus sign and q minus signs. For v, $w \in V$, the inner product with respect to the metric η of v by w will be denoted by $\eta(v,w)$. The norm on V will be defined by $\|v\|^2 = -\eta(v,v)$. It is neither positive defined, nor negative defined. The explanation of the minus sign will come soon. The Clifford algebra is the quotient $\mathrm{Cl}(p,q) := T(V)/\mathcal{I}$ of the tensor algebra by the two-sided ideal $\mathcal{I}$ generated by elements of the form

$$(v\otimes w)\oplus(w\otimes v)\oplus 2\eta(v,w)1$$

for v, w in V. Depending on the context, we will often use the notations $\mathrm{Cl}(\eta)$ or $\mathrm{Cl}(V)$ or $\mathrm{Cl}(p,q)$. The algebra product is $[x]\cdot[y] = [x\otimes y]$, x, $y \in T(V)$. As long as $z \in V \subset \mathrm{Cl}(p,q)$, the expression $\eta(z,z)$ is meaningful. The definition of Cl is such that $z\cdot z = -\eta(z,z)$. This leads to the somewhat surprising formula $z^2 = \|z\|^2 = -\eta(z,z)$.

68.1.2 First representation

Let (V,g) be a metric vector space and $\mathrm{Cl}(V,g)$ its Clifford algebra. For each $v \in V$, we define the two following elements of $\mathrm{End}_{\mathbb{R}}(\bigwedge V)$:

$$\epsilon(v)\big(u_1\wedge\cdots\wedge u_k\big) = v\wedge u_1\wedge\cdots\wedge u_k \tag{68.6a}$$

$$\iota(v)\big(u_1\wedge\cdots\wedge u_k\big) = \sum_{j=1}^{k}(-1)^{j-1}g(u,u_j)u_1\wedge\cdots\wedge\hat{u}_j\wedge\cdots\wedge u_j. \tag{68.6b}$$

1. Don't be afraid with the extra minus sign: the quantum field theory is most written with the metric $(+,-,-,-)$ instead of $(-,+,+,+)$.
2. Definition 49.138.

One has $\epsilon(v)^2 = 0$ and $\iota(v)^2 = 0$ because $v \wedge v = 0$. In order to understand the latter, we wonder what are the terms with $g(v, u_i)g(v, u_j)$ are in

$$\iota(v)^2\big(u_1 \wedge \cdots \wedge u_k\big) = \sum_{l=1}^{k}(-1)^{j-1}g(v, u_j)\sum_{l=1}^{k-1}(-1)^{l-1}g(v, u_l)u_1 \wedge \hat{u}_l \wedge \hat{u}_j \wedge \cdots \wedge u_k.$$

Let's suppose $i < j$. The first term comes when the first $\iota(v)$ acts on u_j, its sign is given by $(-1)^{j-1}(-1)^{i-1}$. The second term has the same $(-1)^{i-1}$, but in this term, u_j is on the position $j-1$ because u_i has disappeared.

Now we use $c(v) = \epsilon(v) + \iota(v)$ which fulfils for all u, $v \in V$:

$$\begin{aligned} c(v)^2 &= g(v, v,)1 \\ c(u)v(v) + c(v)c(u) &= 2g(u, v)1. \end{aligned}$$

Therefore c can be extended to a representation $c\colon \mathrm{Cl}(V, g) \to \mathrm{End}(\bigwedge V)$. If $\{e_0, \cdots e_n\}$ is an orthonormal basis of V (i.e. $g(e_i, e_j) = \eta_{ij}$); in this case the $c(e_j)$ are anticommuting and a basis of $\mathrm{Cl}(V, g)$ is given by

$$\{c(e_{k_1}) \cdots c(e_{k_r}) \text{ tel que } 1 \leqslant k_1 < \cdots < k_r \leqslant n\}. \tag{68.7}$$

68.1.3 Some consequences of the universal property

The map $-\mathrm{Id}\,|_V$ extends to $\alpha \in \mathrm{Aut}\big(\mathrm{Cl}(V)\big)$,

$$\alpha(v_1 \cdots v_r) = (-1)^r v_1 \cdots v_r$$

($v_i \in V$) and provides a graduation

$$\mathrm{Cl}(V) = \mathrm{Cl}^0(V) \oplus \mathrm{Cl}^1(V).$$

The map $\tau\colon \mathrm{Cl}(V) \to \mathrm{Cl}(V)$ extends $\mathrm{Id}\,|_V$ to an anti-homomorphism:

$$\tau(v_1 \cdots v_r) = v_r \cdots v_1. \tag{68.8}$$

The **complexification** of $\mathrm{Cl}(V, g)$ is

$$\mathrm{Cl}^{\mathbb{C}}(V, g) := \mathrm{Cl}(V, g) \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathrm{Cl}(V^{\mathbb{C}}, g^{\mathbb{C}}),$$

the isomorphism being a $\mathbb{C}$-algebra isomorphism. The $\mathbb{R}$-linear operator $v \mapsto \overline{v}$ in $V^{\mathbb{C}}$ of complex conjugation extends to a $\mathbb{R}$-linear automorphism $a \mapsto \overline{a}$. We define the **adjoint** by

$$a^* = \tau(\overline{a}) \tag{68.9}$$

68.1.4 Trace

Theorem 68.3.
There exists one an only one trace $\mathrm{Tr}\colon \mathrm{Cl}^{\mathbb{C}}(V) \to \mathbb{C}$ *such that*

(1) $\mathrm{Tr}(1) = 1$,

(2) $\mathrm{Tr}(a) = 0$ *when* a *is odd.*

Proof. Let $\{e_1, \cdots, e_n\}$ be an orthonormal basis of (V, g) and $a \in \mathrm{Cl}^{\mathbb{C}}(V)$. When one decomposes a into the basis of e_i, one finds a lot of terms of each order. Since Tr is a trace, when the k_i are all different,

$$\mathrm{Tr}(e_{k_1} \cdots e_{k_{2r}}) = \mathrm{Tr}(-e_{k_2} \cdots e_{k_{2r}} e_{k_1} = \mathrm{Tr}(-e_{k_1} \cdots e_{k_{2r}})$$

So the trace of any even element is zero. We decompose a into

$$a = \sum_K a_K \prod_{i \in K} e_i$$

where the sum is taken on the subsets of $\{1,\dots,n\}$. A trace which fulfils the conditions must vanishes on even (but non zero) elements as well as on odd elements, so the only possible form is

$$\operatorname{Tr} a = a_\varnothing.$$

Notice that in order to get this precise form, we used $\operatorname{Tr}(1) = 1$ and linearity. This proves unicity and existence. Now we have to prove that this is a good definition in the sense that an other choice of basis gives the same result. So we take a new orthonormal basis

$$e'_j = \sum_{k=1}^{n} H_{jk} e_k$$

with $H^t H = \mathbb{1}_{n\times n}$. Now we have

$$a = \sum_K a_K \prod_{i\in K} e_i = \sum_K a'_K \prod_{i\in K} e'_i,$$

and we will prove that $a_\varnothing = a'_\varnothing$. Let's compute a lot:

$$\begin{aligned} e'_i e'_j &= \sum_k \sum_l H_{ik} H_{jl} e_k e_l \\ &= \sum_{k=l} H_{ik} H_{jl} e_k e_l + \sum_{k\neq l} H_{ik} H_{jl} e_k e_l \\ &= \sum_k H_{ik} H_{jk} 1 + \sum_{k\neq l} H_{ik} H_{jl} e_k e_l \\ &= (HH^t)_{ij} 1 + \sum_{k\neq l} H_{ik} H_{jl} e_k e_l. \end{aligned}$$

The sense of this formula is that when $i \neq j$, the product $e'_i e'_j$ has no term of order zero. In other terms, as long as we only have terms of order zero, one and two, a change $e \to e'$ does not change the term of order zero. We are now going to an induction proof: we want to prove that $e'_{j_1} \dots e'_{j_{2r}} e'_l e'_k$ has no scalar term assuming that no even combination has scalar terms up to $2(r-1)$. It reads

$$\sum_{K \text{ even}} a_K \prod_{i\in K} e_i e'_l e'_k,$$

therefore we just have to look at terms of the form

$$e_{j_1} \dots e_{j_{2r}} \Big((HH)^t_{kl} 1 - \sum_{i\neq j} C^{ij}_{kl} e_i e_j \Big)$$

where the e_{j_l} are all different. The first term cannot produce a scalar term. In order to find a scalar term in $e'_{j_1} \dots e'_{j_{2r}} e_k e_l$, we begin to look at terms whose decomposition of $e'_{j_1} \dots e'_{j_{2r}}$ ends by $e_l e_k$, i.e.

$$H_{j_{2r-2}l} H_{j_{2r-1}k} e'_{j_1} \dots e'_{2r-3} e_l e_k e_k e_l.$$

The induction assumption says that there are no scalar term in $e'_{2r-3} e_l e_k e_k e_l$.

□

One can prove that $\mathrm{Cl}^{\mathbb{C}}(C)$ is a Hilbert space with the scalar product

$$\langle a|b\rangle = \operatorname{Tr}(a^* b). \tag{68.10}$$

Let $v \in V$ with $g(v,v) = 1$ (thus in $\mathrm{Cl}(V)$, we have $v^2 = 1$); since $v = \overline{v}$, we have

$$a^* v = v v^* = v^2 = 1.$$

Lemma 68.4.
The maps $a \mapsto ua$ and $a \mapsto au$ are unitary if and only if $uu^* = u^*u = 1$.

Proof. We pick $\lambda \in U(1)$ and $w = \lambda v \in V^{\mathbb{C}}$ which fulfils $w^*w = 1$. This is the most general element such that $ww^* = w^*w = 1$. Now for an arbitrary a, $b \in \mathrm{Cl}^{\mathbb{C}}(V)$, we compute the two followings:

$$\langle wa|wb\rangle = \mathrm{Tr}\left((wa)^*wb\right) = \mathrm{Tr}\left(a^*w^*wb\right) = \mathrm{Tr}(a^*b) = \langle a|b\rangle,$$

and

$$\langle aw|bw\rangle = \mathrm{Tr}\left(w^*a^*bw\right) = \mathrm{Tr}(ww^*a^*b) = \mathrm{Tr}(a^*b) = \langle a|b\rangle.$$

This proves that $a \mapsto wa$ and $a \mapsto aw$ are two unitary operators on the Hilbert space $\mathrm{Cl}^{\mathbb{C}}(V)$.

For the converse, we impose for all a, $b \in \mathrm{Cl}^{\mathbb{C}}(V)$:

$$\langle ua|ub\rangle = \mathrm{Tr}(ba^*u^*u) \stackrel{!}{=} \mathrm{Tr}(ba^*).$$

In particular with $a^*b = 1$, $\mathrm{Tr}(u^*u) = \mathrm{Tr}(1) = 1$, thus the scalar part of u^*u is 1. So we write $u^*u = 1 + f$ where f is non scalar, and for any $x \in \mathrm{Cl}^{\mathbb{C}}(V)$, we have

$$\mathrm{Tr}(x) = \mathrm{Tr}(xu^*u) = \mathrm{Tr}(x) + \mathrm{Tr}(xf).$$

We conclude that $\mathrm{Tr}(xf) = 0$, and therefore that $f = 0$. □

68.2 Spinor representation

For the spinor representation, we restrict ourself to the even case $p + q = 2n$.

The aim of this subsection is to find some faithful [3] representations of the complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(p, q)$. In order to achieve this, we first consider $V^{\mathbb{C}}$, the complex vector space of V with an orthonormal basis $\{e_1, \cdots, e_{p-1}, e_p, \cdots, e_{p+q}\}$. The metric is $\eta(e_k, e_k) = 1$ and $\eta(e_{p+k}, e_{p+k}) = -1$ for $k = 0, \cdots, p-1$. We use the following basis:

$$f_k = \frac{1}{2}(e_k + e_{p+k}), \qquad g_k = \frac{1}{2}(e_k - e_{p+k}), \tag{68.11}$$

SUBEQooZLYUooUfZQFL

$$f_{p+s} = \frac{1}{2}(e_{2p+2s} + ie_{2p+2s+1}), \qquad g_{p+s} = \frac{1}{2}(e_{2p+2s} - ie_{2p+2s}) \tag{68.12}$$

for $k = 0, \cdots, p-1$. We note that $\{f_0, g_0\}$ spans a $\mathbb{C}^2$-space which is η-orthogonal to the one which is spanned by $\{f_1, g_1\}$. The following two spaces will prove to be useful:

SUBEQSooHDHKooMWSPxg

$$W = \mathrm{Span}_{\mathbb{C}}\{f_0, f_1\} \simeq \mathbb{C}^2, \tag{68.13a}$$

$$\underline{W} = \mathrm{Span}_{\mathbb{C}}\{g_0, g_1\} \simeq \mathbb{C}^2. \tag{68.13b}$$

It is easy to compute the various products; among others we find

$$\eta(f_0, f_0) = 0, \quad \eta(f_1, f_0) = 0, \quad \eta(f_1, f_1) = 0; \tag{68.14}$$

so that for any $w \in W$, we have $\langle w, w\rangle = 0$; for this reason, we say that W is a **completely isotropic** subspace of $(V^{\mathbb{C}}, \eta^{\mathbb{C}})$. The space $\underline{W}$ has the same property.

Proposition 68.5.
We have

$$\underline{W} \simeq W^*, \tag{68.15}$$

where W^* is the dual space of W. By $\simeq$ we mean that there exists a linear bijective map $\psi\colon \underline{W} \to W^*$.

3. Definition 4.135.

Proof. For each $\underline{w} \in \underline{W}$, we define $\psi(\underline{w})\colon W \to \mathbb{C}$ by

$$\psi(\underline{w})(w) = \eta(w, \underline{w}).$$

We first show that the map ψ is injective. Let $\underline{w} \in \underline{W}$ be so that $\psi(\underline{w}) = 0$. Thus for all $v \in W$, we have

$$\psi(\underline{w})v = \eta(\underline{w}, v) = 0. \tag{68.16}$$

3101r1

By decomposing $\underline{w} = ag_0 + bg_1$ and taking successively $v = f_0$ and $v = f_1$, we see that $a = b = 0$.

The next step is to see that the map ψ is surjective. We know that $dim_{\mathbb{C}}\underline{W} = dim_{\mathbb{C}}W^* = 2$ and that $\psi(g_0) \neq 0$. Let's prove that $\{\psi(g_0), \psi(g_1)\}$ is a basis of W^*. It is clear by linearity that $\{\psi(ag_0) : a \in \mathbb{C}\} = \text{Span}\{\psi(g_0)\}$. The fact that ψ is injective imposes that $\psi(g_1)$ doesn't belong to $\text{Span}\{\psi(g_0)\}$. So $\{\psi(g_0), \psi(g_1)\}$ is a two-dimensional free subset of W^*, and therefore is a basis of W^*. □

We turn our attention to the exterior algebra $\Lambda W = \mathbb{C} \oplus W \oplus (W \wedge W) \oplus \cdots \oplus \wedge^{p+q} W$ of W.

DEFooURQDooMoZVbH

Definition 68.6.
We define the homomorphism $\tilde{\rho}\colon V^{\mathbb{C}} \to \text{End}(\Lambda W)$ *by*

$$\begin{aligned} \tilde{\rho}(f_i)\alpha &= f_i \wedge \alpha, \\ \tilde{\rho}(g_i)\alpha &= -\iota(g_i)\alpha \end{aligned} \tag{68.17}$$

($v \in V^{\mathbb{C}}$, $\alpha \in \Lambda W$) where ι denotes the interior product defined in page 3019.

defrt

More explicitly, for all $z \in \mathbb{C}$ and for all $w, w' \in W$, we have

$$\tilde{\rho}(f_i)z = zf_i, \qquad \tilde{\rho}(g_i)z = 0, \tag{68.18a}$$

$$\tilde{\rho}(f_i)w = f_i \wedge w, \qquad \tilde{\rho}(g_i)w = -\eta(g_i, w)1, \tag{68.18b}$$

$$\tilde{\rho}(f_i)(w \wedge w') = 0, \qquad \tilde{\rho}(g_i)(w \wedge w') = -\eta(g_i, w)w' + \eta(g_i, w')w. \tag{68.18c}$$

We will see that, *via* some manipulations, $\tilde{\rho}$ provides a faithful representation of the Clifford algebra, the **spinor representation**.

Remarque 68.7.
By "endomorphism of ΛW", we mean an endomorphism for the *linear* structure of ΛW. We obviously not have $\tilde{\rho}(x)(\alpha \wedge \beta) = \tilde{\rho}(x)\alpha \wedge \tilde{\rho}(x)\beta$.

Proposition 68.8.
The map $\tilde{\rho}$ is injective.

Proof. We have to show that $\tilde{\rho}(v) = 0$ (v in $V^{\mathbb{C}}$) implies $v = 0$. Any $v \in V^{\mathbb{C}}$ can be written as $v = a^i f_i + b^i g_i$ with a sum over i. We first have that

$$\tilde{\rho}(a^i f_i + b^i g_i)z = za^i f_i = 0,$$

but the f_i are independents and then $a^i = 0$. We can also write

$$\tilde{\rho}(b^0 g_0 + b^1 g_1)f_1 = -b^0\eta(g_0, f_1) - b^1\eta(g_1, f_1) = -\frac{b^1}{2} = 0,$$

then $b^1 = 0$. The same with f_0 proves that $b^0 = 0$. □

The homomorphism $\tilde{\rho}$ extends to the whole the tensor algebra of $V^{\mathbb{C}}$ by the following definitions:

$$\tilde{\rho}(1) = \text{Id}_{\Lambda W}, \tag{68.19a}$$

$$\tilde{\rho}(e_k) = \tilde{\rho}(e_k), \tag{68.19b}$$

$$\tilde{\rho}(e_{k_1} \otimes \ldots \otimes e_{k_r}) = \tilde{\rho}(e_{k_1}) \circ \ldots \circ \tilde{\rho}(e_{k_r}). \tag{68.19c}$$

eq:3101r2

So we get $\tilde{\rho}\colon T(V^{\mathbb{C}}) \to \text{End}(\Lambda W)$. The following proposition will allow us to descent $\tilde{\rho}$ to a representation of the Clifford algebra.

Proposition 68.9.
The homomorphism $\tilde{\rho}$ maps $\mathcal{I}$ to 0: $\tilde{\rho}(\mathcal{I}) = 0$.

¡¡ Avertissement/question au lecteur !! 68.10
This proposition is wrong: there is a double covering.

Moreover, there is a sign problem in the proof: the sign in the first lines is not the one used in the definition of the Clifford algebra.

Proof. We have to check the following:

$$\tilde{\rho}(v \otimes w \oplus w \otimes v - 2\eta(v,w)1) = 0$$

for any choice of v, w in $\{e_0, e_1, e_2, e_3\}$. Here we will just check it explicitly for $v = e_0$ and $w = e_1$. The computation uses the definition (68.19c):

$$\begin{aligned} \tilde{\rho}(e_0 \otimes e_1 \oplus e_1 \otimes e_0 - 2\eta(e_0, e_1) &= \tilde{\rho}(e_0) \circ \tilde{\rho}(e_1) + \tilde{\rho}(e_1) \circ \tilde{\rho}(e_0) \\ &= 2\left[\tilde{\rho}(f_0)^2 - \tilde{\rho}(g_0)^2\right]. \end{aligned} \tag{68.20}$$

It is easy to see that $\tilde{\rho}(f_0)^2 = 0$:

$$\tilde{\rho}(f_0)^2 \left[z \oplus w \oplus w_1 \wedge w_2\right] = \tilde{\rho}(f_0)[zf_0 \oplus f_0 \wedge w] = zf_0 \wedge f_0, = 0. \tag{68.21}$$

The proof that $\tilde{\rho}(g_0)^2 = 0$ is almost the same:

$$\tilde{\rho}(g_0)^2 \left[z \oplus w \oplus w_1 \wedge w_2\right] = \tilde{\rho}(g_0)[-\eta(g_0, w)1 \oplus -\eta(g_0, w_1)w_2 \oplus \eta(g_0, w_2)w_1].$$

□

We can now see $\tilde{\rho}$ as a map $\tilde{\rho}\colon \mathrm{Cl}^{\mathbb{C}}(p,q) \to \mathrm{End}(\Lambda W)$. By construction, it is a homomorphism and, thus, is a representation of $\mathrm{Cl}^{\mathbb{C}}(p,q)$ on ΛW. For compactness, we use the notation

$$\gamma_a := \sqrt{2}\tilde{\rho}(e_a). \tag{68.22}$$ defgamma

Lemma 68.11.
The γ's operators satisfy the following relation:

$$\gamma_a\gamma_b + \gamma_b\gamma_a = -2\eta_{ab}\mathbb{1}. \tag{68.23}$$ 3101r3

310111

Proof. We have to check this equality on any element of ΛW. If we choose $w_1 = af_0 + bf_1$ and $w_2 = a'f_0 + b'f_1$, we find $w_1 \wedge w_2 = (ab' - ba')f_0 \wedge f_1$.

For example, we will explicitly check (68.23) with $a = b = 0$, i.e. $\tilde{\rho}(e_0) \circ \tilde{\rho}(e_0) = \frac{1}{2}\,\mathrm{Id}$, which proves that $\gamma_0 \circ \gamma_0 = \mathrm{Id}$.

$$\begin{aligned} \tilde{\rho}(e_0)^2[z \oplus w \oplus (ab' - ba')f_0 \wedge f_1] &= \tilde{\rho}(f_0 + g_0)^2[z \oplus w \oplus (ab' - ba')f_0 \wedge f_1] \\ &= \tilde{\rho}(f_0 + g_0)\Big[zf_0 \oplus f_0 \wedge w \oplus -\eta(g_0, w)1 \\ &\qquad - (ab' - ba')\eta(g_0, f_0)f_1 \\ &\qquad + (ab' - ba')\eta(g_0, f_1)f_0\Big] \\ &= \frac{1}{2}(z \oplus w \oplus (ab' - ba')f_0 \wedge f_1). \end{aligned} \tag{68.24}$$

□

Lemma 68.12.
For any sequence $i_0, \dots i_3$ of 0 and 1 (with at least one of them equals to 1), we have

$$\mathrm{Tr}(\gamma_0^{i_0} \cdots \gamma_{2n-1}^{i_{2n-1}}) = 0. \tag{68.25}$$

We take the convention that $\gamma_a^0 = \mathbb{1}$. 310112

Proof. If the number of nonzero i_k is even (say $2m$), we have:

$$\mathrm{Tr}(\gamma_{a_1}\dots\gamma_{a_{2m}}) = \mathrm{Tr}(\gamma_{a_{2n}}\gamma_{a_1}\dots\gamma_{a_{2m-1}})$$

because the trace is invariant under cyclic permutations. But we can also permute $\gamma_{a_{2m}}$ with the $2m-1$ other γ's. $\mathrm{Tr}(\gamma_{a_1}\dots\gamma_{a_{2m}}) = (-1)^{2n-1}\,\mathrm{Tr}(\gamma_{a_{2m}}\gamma_{a_1}\dots\gamma_{a_{2m-1}})$ because each permutation gives an extra minus sign (lemma 68.11). Then the trace is zero.

If the number of nonzero i_k is odd (say $2m-1$). Let $i_a = 0$ (we restrict ourself to the even dimensional case). We have $\mathrm{Tr}(A) = -\eta_{aa}\,\mathrm{Tr}(A\gamma_a\gamma_a)$. Using once again the cyclic invariance of the trace, $\mathrm{Tr}(\gamma_{a_1}\dots\gamma_{a_{2m-1}}\gamma_a\gamma_a) = \mathrm{Tr}(\gamma_a\gamma_{a_1}\dots\gamma_{a_{2m-1}}\gamma_a)$. But, if we permute the *first* γ_a with the $2m-1$ first γ's, we find $\mathrm{Tr}(\gamma_{a_1}\dots\gamma_{a_{2m-1}}\gamma_a\gamma_a) = -\,\mathrm{Tr}(\gamma_a\gamma_{a_1}\dots\gamma_{a_{2m-1}}\gamma_a)$, and the trace is zero again. □

Proposition 68.13.
The subset

$$\{\mathbb{1}, \gamma_a\gamma_b\,(a<b), \gamma_a\gamma_b\gamma_c\,(a<b<c), \cdots, \gamma_0\cdots\gamma_{2n}\}$$

is free in $\mathrm{End}(\Lambda W)$.

Proof. We consider a general linear combination of these operators:

$$E = \lambda\mathbb{1} + \sum_a \lambda_a\gamma_a + \sum_{a<b}\lambda_{ab}\gamma_a\gamma_b + \cdots + \sum_{a<b<c<d}\lambda_{abcd}\gamma_a\gamma_b\gamma_c\gamma_d.$$

The claim is that if $E=0$, then all the coefficients $\lambda_{(\dots)}$ must be zero. First note that $Tr(E) = 0 = \lambda$ by lemma 68.12. It is also clear that $Tr(\gamma_i E) = 0 = \lambda_i$. In order to see that $\lambda_{ij} = 0$, we compute $Tr(\gamma_j\gamma_i E) = 0 = \lambda_{ij}$. And so on. □

How many operators does we have in this free system? Any operators in this system can be written as $\gamma_0^{i_0},\cdots\gamma_{2n-1}^{i_{2n-1}}$ with i_k equal to zero or one. Thus we have 2^{2n} operators. On the other hand, we know that $dim_{\mathbb{C}}\Lambda W = 2p+2$, and then that $dim_{\mathbb{C}}\,\mathrm{End}(\Lambda W) = 4^2 = 16$. The result is that $\{\gamma_0^{i_0},\cdots\gamma_{2n-1}^{i_{2n-1}}$ tel que $i_k = 0\,or\,1\}$ is a basis of $\mathrm{End}(\Lambda W)$. In other words (if we suppose a suitable ordering), the image by $\tilde\rho$ of

$$B = \{1, e_a, e_a\otimes e_b, e_a\otimes e_b\otimes e_c, e_a\otimes e_b\otimes e_c\otimes e_d\}$$

is a basis of $\mathrm{End}(\Lambda W)$.

If B is a basis of $C^{\mathbb{C}}_{(p,q)}$, then $\tilde\rho$ is bijective and thus isomorphic. Therefore, we expect $\tilde\rho\colon C^{\mathbb{C}}_{(p,q)} \to \mathrm{End}(\Lambda W)$ to be a faithful representation. It is not difficult to see that B is indeed a basis thanks to the equivalence relation.

68.2.1 Explicit representation

First, we choose a basis for ΛW:

$$\mathbb{1} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad f_0 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad f_1 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad f_0\wedge f_1 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{68.26}$$

Here is the explicit computation for the matrix γ_0 in this basis. First remark that $\tilde\rho(e_0)1 = f_0$, $\tilde\rho(e_0)f_0 = \frac{1}{2}$, $\tilde\rho(e_0)f_1 = f_0\wedge f_1$, $\tilde\rho(e_0)(f_0\wedge f_1) = \frac{1}{2}f_1$. Then

$$\begin{aligned}
&\gamma_0\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \sqrt{2}\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \gamma_0\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \sqrt{2}\begin{pmatrix}\frac{1}{2}\\0\\0\\0\end{pmatrix},\\
&\gamma_0\begin{pmatrix}0\\0\\1\\0\end{pmatrix} = \sqrt{2}\begin{pmatrix}0\\0\\0\\1\end{pmatrix},\quad \gamma_0\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \sqrt{2}\begin{pmatrix}0\\0\\\frac{1}{2}\\0\end{pmatrix}.
\end{aligned} \tag{68.27}$$

This allows us to write down γ_0; the same computation gives the other matrices.

$$\begin{aligned}
\gamma_0 &= \sqrt{2}\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, & \gamma_1 &= \sqrt{2}\begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\
\gamma_2 &= \sqrt{2}\begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, & \gamma_3 &= \sqrt{2}\begin{pmatrix} 0 & 0 & -\frac{i}{2} & 0 \\ 0 & 0 & 0 & \frac{i}{2} \\ -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{68.28}$$

It is easy to check that these matrices satisfies (68.23).

Notice that, up to a suitable change of basis in ΛW, these are the usual Dirac matrices. Indeed we actually solved the physical problem to find a representation of the algebra (68.2). We understand by the way why do physicists work with 4-components spinors: the γ's are operators on the four-dimensional space ΛW; hence the Dirac operator will naturally acts on four-components objects.

The main result of this section is an explicit faithful representation of $\mathrm{Cl}^{\mathbb{C}}(p,q)$. This allows us to write a **Dirac operator** which solve (see the invitation 68.1 and [973]) the problem to find a "square root" of the d'Alembert operator: the differential operator $\mathcal{D} = \gamma^\mu \partial_\mu$ satisfies $\mathcal{D}^2 = \Box$.

68.2.2 A remark

Let us compare the two faithful representations

$$\begin{aligned} c\colon\ & \mathrm{Cl}(V) \to \mathrm{End}_{\mathbb{R}}(\wedge V) \\ \tilde{\rho}\colon\ & \mathrm{Cl}^{\mathbb{C}} \to \mathrm{End}_{\mathbb{R}}(\wedge W). \end{aligned}$$

They obviously comes from the same ideas. One common point is that

$$c(e_1)(e_1 \wedge e_2) = 2\tilde{\rho}(e_1)(e_1 \wedge e_2) = e_2,$$

but they are different:

$$\begin{aligned} \tilde{\rho}(e_3)(e_0 \wedge e_2) &= 0 \\ c(e_3)(e_0 \wedge e_2) &= e_3 \wedge e_0 \wedge e_1. \end{aligned}$$

68.2.3 General two dimensional Clifford algebra

The Clifford algebra for the metric

$$g = \begin{pmatrix} \alpha & \delta \\ \delta & \beta \end{pmatrix}$$

is realised by matrices

$$\gamma_1 = \epsilon \begin{pmatrix} \sqrt{\alpha} & \\ & -\sqrt{\alpha} \end{pmatrix}, \quad \gamma_2 = \epsilon \begin{pmatrix} \delta/\sqrt{\alpha} & \beta - \delta^2/|\alpha| \\ 1 & -\delta/\sqrt{\alpha} \end{pmatrix}$$

where $\epsilon = \pm 1$ is chosen in such a way that $\epsilon|\alpha| = \alpha$.

68.3 Spin group

We will not immediately go on with Dirac operators on Riemannian manifolds because we still have to build some theory about the Clifford algebra itself. In particular, we have to define the spin group which will play a central role in the definition of the Dirac operator. Almost all –and

(too?) much more– the concepts we will introduce in this section can be found in [972]; a more physical oriented but useful approach can be found in [977].

Let define the map $\chi\colon \Gamma(p,q) \to GL(\mathbb{R}^{1,3})$ by

$$\chi(x)y = \alpha(x) \cdot y \cdot x^{-1}. \tag{68.29}$$

Let

$$\Gamma(p,q) = \{x \in \mathrm{Cl}(p,q) \text{ tel que } x \text{ is invertible and } \chi(x)y \in V \text{ for all } y \in V\}.$$

It should be remarked that this definition comes back to the real Clifford algebra. The Clifford algebra product gives this subset a group structure which is called the **Clifford group**. Any $x \in V$ is invertible since $x \cdot x = -\eta(x,x)1$, the inverse of x is given by $x^{-1} = x/\|x\|^2$.

The subset $\mathrm{Cl}(p,q)^+$ (resp. $\mathrm{Cl}(p,q)^-$) of $\mathrm{Cl}(p,q)$ is the image of even (resp. odd) tensors of $T(V)$ by the canonical projection $T(V) \to \mathrm{Cl}(p,q)$. With these definitions, we have a natural grading of Cl:

$$\mathrm{Cl}(p,q) = \mathrm{Cl}(p,q)^+ \oplus \mathrm{Cl}(p,q)^-, \tag{68.30}$$ directC

and the subgroups

$$\Gamma(p,q)^+ = \Gamma(p,q) \cap \mathrm{Cl}(p,q)^+, \qquad \Gamma(p,q)^- = \Gamma(p,q) \cap \mathrm{Cl}(p,q)^-. \tag{68.31}$$ defgplus

For $x_1, \dots, x_n \in V$, we have $\tau(x_1 \cdots x_n) = x_n \cdots x_1$. The **spin group** is

$$\mathrm{Spin}(p,q) = \{x \in \Gamma(p,q)^+ | \tau(x) = x^{-1}\} \tag{68.32}$$ defSpinpq

while the **spin norm** is the map $N\colon \Gamma(p,q) \to \Gamma(p,q)$ defined by

$$N(x) = x\tau(\alpha(x)).$$ proppourN

Proposition 68.14.
The map N takes values in $\mathbb{R}$ and the formula

$$N(x \cdot y) = N(x)N(y), \tag{68.33}$$

holds for all x, $y \in \Gamma(p,q)$.

Proof. We write as usual $x \in \Gamma(p,q)$ as $x = cv_1 \cdots v_r$. So,

$$N(x) = cv_1 \cdots v_r \tau(\alpha(cv_1 \cdots v_r)) = (-1)^r c^2 v_1 \cdots v_r \cdot v_r \cdots v_1. \tag{68.34}$$

The first equality comes from the fact that $\alpha(cv_1 \cdots v_r) = (-1)^r cv_1 \cdots v_r$. Now $N(x) \in \mathbb{R}$ because $v_i \cdot v_i = -\langle v_i, v_i \rangle \in \mathbb{R}$ for all i. Hence the following hold:

$$\begin{aligned} N(x \cdot y) &= v \cdot y \cdot \tau(\alpha(v \cdot y)) \\ &= v \cdot y \cdot \tau(\alpha(y)) \cdot \tau(\alpha(v)) \\ &= v \cdot N(y)\tau(\alpha(v)) \\ &= N(y)N(x). \end{aligned} \tag{68.35}$$

This is the claim. □

Therefore $N\colon \Gamma(p,q) \to \mathbb{R}$ is an homomorphism.

Remarque 68.15.
The elements of $\mathrm{Spin}(p,q)$ are spin-normed at 1. Indeed, take a s in $\mathrm{Spin}(p,q)$. We have $N(s) = s \cdot \tau(s) = 1$ because $\alpha(s) = s$ and $\tau(s) = s^{-1}$. In particular $\mathrm{Spin}(p,q) \cap \mathbb{R} = \mathbb{Z}_2$. rem:spin_norm_u

68.3.1 Studying the group structure

Proposition 68.16.
The set $\Gamma(p,q)$ admits a Lie group structure.

Proof. During this proof, μ denotes the Clifford multiplication: $\mu(x,y) = x \cdot y$. We know that $\mathrm{Cl}^{\mathbb{C}}(p,q)$ is isomorphic to $\mathrm{End}(\Lambda W)$ in which the multiplication is a continuous map. Thus μ is continuous on $\mathrm{Cl}^{\mathbb{C}}(p,q)$. But $\mathrm{Cl}(p,q)$ is a closed subset of $\mathrm{Cl}^{\mathbb{C}}(p,q)$, so μ is a continuous map in $\mathrm{Cl}(p,q)$. This proves that χ seen as a map from $\Gamma(p,q) \times V$ to V is a continuous map.

The space V is closed in $\mathrm{Cl}(p,q)$, thus $\sigma^{-1}(V)$ is also closed. But $\sigma^{-1}(V) = \Gamma(p,q) \times \mathrm{Cl}(p,q)$. So $\Gamma(p,q)$ is closed in $\mathrm{Cl}(p,q)$.

Now the result is just a consequence of theorems 46.6 and 53.98. Indeed, let us study the subset $\mathcal{I}$ which appears in the definitions of the Clifford algebra. It makes no difficult to convince ourself that it is a closed subgroup of $T(V)$. The theorem 53.98 thus makes $\mathrm{Cl}(p,q) = T(V)/\mathcal{I}$ a Lie group. But we just say that $\Gamma(p,q)$ is closed in $\mathrm{Cl}(p,q)$, and the fact that $\Gamma(p,q)$ is a subgroup of $\mathrm{Cl}(p,q)$ is clear. By theorem 46.6 we conclude that there exists a Lie group structure on $\Gamma(p,q)$. □

Lemma 68.17.
The map χ is a homomorphism, in other words χ is a representation of $\Gamma(p,q)$.

Proof. The following computation uses the fact that α is a homomorphism:

$$\begin{aligned}\chi(a \cdot b)y &= \alpha(a \cdot b) \cdot y \cdot (a \cdot b)^{-1} = \alpha(a) \cdot \alpha(b)y \cdot b^{-1} \cdot a^{-1} \\ &= \alpha(a) \cdot \chi(b)y \cdot a^{-1} = \chi(a)\chi(b)y.\end{aligned}$$

□

Let $y \in \Gamma(p,q)^-$ and $v \in V$. Where is $y \cdot v$? First note that $(y \cdot v)^{-1} = v^{-1} \cdot y^{-1}$, so that

$$\begin{aligned}\alpha(y \cdot v) \cdot w \cdot (y \cdot v)^{-1} &= -\alpha(y) \cdot v \cdot w \cdot v^{-1} \cdot y^{-1} \\ &= -\alpha(y)\big(2\eta(v,w) - w \cdot v\big) \cdot v^{-1} \cdot y^{-1} \\ &= -2\eta(v,w)\alpha(y) \cdot v^{-1} \cdot y + \alpha(y) \cdot w \cdot y^{-1}\end{aligned} \tag{68.36}$$

which belongs to V because $y \in \Gamma(p,q)$. This reasoning shows that (apart for 0), $y \cdot v \in \Gamma(p,q)^+$ if and only if $y \in \Gamma(p,q)^-$.

Lemma 68.18.
If $x \in V$ is non-isotropic (i.e. $\eta(x,x) \neq 0$), the automorphism $\chi(x)$ is the orthogonal symmetry with respect to $x^\perp$.

We recall that

$$x^\perp = \{y \in V \text{ tel que } \eta(x,y) = 0\}.$$

We will denote by σ^x the orthogonal symmetry with respect to $x^\perp$.

Proof. When the operator σ^x acts on y, it just change the sign of the "x-part" of y. So we can write $\sigma^x y = y - 2\eta(x,y)1_x$, where $1_x := x/\|x\|$. It should be checked if $\chi(x)y = \alpha(x) \cdot y \cdot x^{-1}$ is equal to $y - 2\eta(x,y)1_x$ or not. We know that $x \cdot x = \eta(x,x)1 = -\|x\|$. It follows that

$$x \cdot y + y \cdot x = 2\eta(x,y)\frac{x \cdot x}{\|x\|}.$$

If we multiply this at right by x^{-1}, using the fact that $\alpha(x) = -x$, we find

$$-\alpha(x) \cdot y \cdot x^{-1} = -y + 2\eta(x,y)1_x,$$

which is precisely the identity we wanted to check. □

The following result will help us to identify subgroups of Clifford group as isometry groups.

Theorem 68.19 (Cartan-Dieudonné theorem)**.**
Each σ in $O(1,3)$ can be written as $\sigma = \tau_1 \circ \ldots \circ \tau_m$, where the τ's are orthogonal symmetries with respect to hyperplanes which are orthogonal to non-isotropic vectors. CartanDieu

Proposition 68.20.

$$\chi(\Gamma(p,q)) = O(p,q).$$

prop1001t1

Proof. In order to show that $\chi(\Gamma(p,q)) \subset O(p,q)$ take $z \in V$ and $x \in \Gamma(p,q)$. Since $\alpha(x) \cdot z \cdot x^{-1}$ lies in V, we can write:

$$\alpha(x) \cdot z \cdot x^{-1} = -\alpha\left(\alpha(x) \cdot z \cdot x^{-1}\right) = -x \cdot \alpha(z) \cdot \alpha(x^{-1}) = x \cdot z \cdot \alpha(x^{-1}).$$

In order to see that $\chi(x) \in O(p,q)$, we have to prove that $\|\chi(x)y\|^2_{(p,q)} = \|y\|^2_{(p,q)}$. This is achieved by the following computation:

$$\begin{aligned} \|\chi(x)y\|^2_{(p,q)} &= -\left(\alpha(x) \cdot y \cdot x^{-1}\right)^2 = \left(\alpha(x) \cdot y \cdot x^{-1}\right)\left(x \cdot y \cdot \alpha(x^{-1})\right) \\ &= -\alpha(x) \cdot y^2 \cdot \alpha(x^{-1}) = \|y\|^2_{(p,q)}. \end{aligned} \tag{68.37}$$

The last step is simply the fact that $y^2 \in \mathbb{R}$ and therefore commutes with anything. We now know that $\chi(x) \in O(p,q)$ for all $x \in \Gamma(p,q)$. Thus $\chi(\Gamma(p,q)) \subset O(p,q)$.

For the second part, let σ be in $O(p,q)$. The Cartan-Dieudonné theorem(theorem 68.19) says that $\sigma = \sigma^{x_1} \circ \ldots \circ \sigma^{x_r}$ for some $x_1, \ldots, x_r$ in V. Thus $\sigma = \chi(x_1 \cdots x_r)$, and $O(p,q) \subset \chi(\Gamma(p,q))$. □

Proposition 68.21.

$$\ker \chi = \mathbb{R}^\times \tag{68.38}$$

where the right hand side is the set of invertible elements of $\mathbb{R}$. prop1001p1

Proof. Before beginning the proof, we want to insist on the fact that $x \in \ker \chi$ does not mean that $\chi(x)y = 0$ for all y in V. The "zero" of an algebra is the element e which satisfies $e \cdot y = y \cdot e = y$ for all y in the algebra. In other words, x is in the kernel of χ if and only if $\chi(x) = \mathrm{Id}$.

First we show that $\mathbb{R}_0 \subset \ker \chi$. If $x \in \mathbb{R}$, then $\alpha(x) = x$. Therefore, when $x \neq 0$,

$$\chi(x)y = \alpha(x) \cdot y \cdot x^{-1} = y,$$

because the algebra product $\cdot$ between an element of $\mathrm{Cl}(p,q)$ and a real is commutative. Note that this does not work with $x = 0$.

We are now going to show that $\ker \chi \subset \mathbb{R}$. Let $z \in \ker \chi$. We decompose (definitions (68.31)) it into his odd and even part: $z = z^+ + z^-$, with $z^\pm \in \Gamma(p,q)^\pm$. These two can be written as $z^+ = e_{j_1} \cdots e_{j_{2r}}$ and $z^- = e_{i_1} \cdots e_{i_{2r-1}}$ with no two i_k or j_k equals. This is almost the general form of elements in even and odd part of $\Gamma(p,q)$: the only other possibility is z in $\mathbb{R}$. Obviously $\alpha(z^\pm) = \pm z^\pm$. Multiplying the condition $\chi(z)y = y$ at right by $(z^+ + z^-)$, we find

$$(z^+ - z^-)y = y(z^+ + z^-).$$

Thanks to equation (68.30), we can split this condition into even and odd parts:

$$z^+ y = y z^+, \qquad z^- y = -y z^-. \tag{68.39}$$

The first equation with $y = e_{j_1}$ gives $e_{j_1} \cdots e_{j_{2r}} \cdot e_{j_1} = e_{j_1} e_{j_1} \cdots e_{j_{2r}}$. In the left hand side, permute the last e_{j_1} from last to second position. So we find the right hand side, with an extra minus sign. This means that $z^+ = 0$. In the same way, the second equation gives $z^- = 0$. We are left with the last possibility: $z \in \mathbb{R}$. □

Corollary 68.22.
For any $s \in \Gamma(p,q)$, there exists some non-isotropic vectors $x_1, \dots, x_r$, and $c \in \mathbb{R}$ such that $s = cx_1 \cdots x_r$. 602c1

Proof. Let us take a $s \in \Gamma(p,q)$; we just saw (theorem 68.20) that $\chi(s)$ is an element of $O(p,q)$. It can be written $\chi(s) = \sigma_1 \circ \ldots \circ \sigma_m$. But we had shown that $\sigma_i = \chi(x_i)$ for any x_i normal to the hyperplane defining σ_i. We thus have

$$\chi(s) = \chi(x_1 \cdots x_m),$$

where s belongs to $\Gamma(p,q)$ and is therefore invertible. This leads us to write $\mathrm{Id} = \chi(s^{-1} \cdot x_1 \cdots x_m)$. But the kernel of χ is $\mathbb{R}$ (proposition 68.21); so one can find a $r \in \mathbb{R}$ such that $s^{-1} \cdot x_1 \cdots x_m = r$. The claim follows. □

Lemma 68.23.
If $v \in V$,

$$\det \chi(v) = -1. \tag{68.40}$$

Proof. We already know that $det\chi(v) = \pm 1$. To check that the right sign is plus, take the following basis of V: $\{v, v_i^\perp\}$ where $\{v_i^\perp\}$ is a basis of $v^\perp$. Calculating the action of $\chi(v)$ on this basis, we find:

$$\begin{aligned} \chi(v)v &= -v \cdot v \cdot v^{-1} = -v, \\ \chi(v)v_i^\perp &= -v \cdot v_i^\perp \cdot v^{-1} = v_i^\perp \cdot v \cdot v^{-1} = v_i^\perp. \end{aligned} \tag{68.41}$$

In this computation, we used the relation $v \cdot w = -w \cdot v - 2\langle v, w\rangle$ which is true for all v, w in V. The action of $\chi(v)$ on this basis is thus to let unchanged three vectors and to change the sign of the fourth. This proves the claim. □

Theorem 68.24.

$$\chi(\Gamma(p,q)^+) = \mathrm{SO}(p,q). \tag{68.42}$$

2102p1

Proof. From corollary 68.22, and definition 68.31, an element $s \in \Gamma(p,q)^+$ reads $s = cv_1 \cdots v_{2r}$. Thus

$$\det \chi(s) = \det \chi(v_1 \cdots v_{2r}) = \det\left[\chi(v_1) \ldots \chi(v_{2r})\right]. \tag{68.43}$$

But we know that, for all v_i in V, $det\chi(v_i) = -1$. So $\det \chi(s) = 1$ and $\chi(\Gamma(p,q)^+) \subseteq \mathrm{SO}(p,q)$. As set,

$$\Gamma(p,q) = \Gamma(p,q)^+ \cup \Gamma(p,q)^-,$$

but the lemma shows that $\det \chi(\Gamma(p,q)^-) = -1$ so, from theorem 68.20, $\chi(\Gamma(p,q)^+)$ must be the whole $\mathrm{SO}(p,q)$. □

Theorem 68.25.
We have the following isomorphism of groups

$$\mathrm{Spin}(p,q) = \mathrm{SO}_0(p,q).$$

provided by the map χ.

¡¡ Avertissement/question au lecteur !! 68.26
This result is wrong because of a double covering issue. The real proposition is the next one. I should try to merge the proofs.

Proof. Let $\{e_1, \cdots, e_p, f_1, \cdots, f_p\}$ be a basis of $\mathbb{R}^{p+q}$ where the e_i's are time-like and the f_j's are space-like. Following the discussion at page 3946, we have

$$\mathrm{SO}(p,q) = \mathrm{SO}_0(p,q) \cup \xi\, \mathrm{SO}_0(p,q)$$

where ξ is defined as follows: $\xi e_1 = -e_1$, $\xi f_1 = -f_1$ and $\xi e_k = e_k$, $\xi f_k = f_k$ for $k \neq 1$. This element can be implemented as $\xi = \chi(g)$ for $g = e_1 f_1$. It is easy to see that $g^{-1} = -f_1 e_1$ and that $\tau(g) = f_1 e_1$, so that $g \notin \mathrm{Spin}(p,q)$.

Is it possible to find another $h \in \Gamma(p,q)$ such that $\chi(h) = \xi$? If $\chi(a) = \chi(b)$ for a, $b \in \Gamma(p,q)$, then $a = rb$ for a certain $r \in \mathbb{R}$. So we find that $h = g^{-1}/r$ is the general form of an element in $\Gamma(p,q)$ such that $\chi(h) = \xi$. This is an element of $\mathrm{Spin}(p,q)$ if and only if $\tau(h) = h^{-1}$, or $-e_1 f_1/r = r e_1 f_1$ which has no solutions. We conclude that no element of $\mathrm{Spin}(p,q)$ is send on ξ by χ. So

$$\chi\big(\mathrm{Spin}(p,q)\big) \subset SO_0(p,q).$$

¡¡ Avertissement/question au lecteur !! 68.27
I still have to prove the surjectivity of χ from $\mathrm{Spin}(p,q)$ to $\mathrm{SO}(p,q)$.

□

Theorem 68.28.

$$\chi(\mathrm{Spin}(p,q)) = \mathrm{SO}_0(p,q) \tag{68.44}$$ EqchiSpinSO

where the index 0 means the identity component.

Proof. Proposition 68.21, theorem 68.24 and remark 68.15 show that the map $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}(p,q)$ is a homomorphism with $\mathbb{Z}_2$ as kernel. We begin to prove that $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$ is surjective. On the one hand, elements of $\mathrm{Spin}(p,q)$ satisfy one more condition than the ones of $\Gamma(p,q)^+$. Thus the algebra $\mathrm{Spin}(p,q)$ has codimension one in $\Gamma(p,q)^+$.

On the other hand, we know that $\mathrm{SO}(p,q) = \mathrm{SO}_0(p,q) \cup h\, \mathrm{SO}_0(p,q)$ where h is the matrix such that $he_i = -e_i$ for $i = 0, \ldots, 3$. Since $\mathrm{Spin}(p,q)$ has codimension one in $\Gamma(p,q)^+$, there is at most one more generator in $\chi(\Gamma(p,q)^+)$ than in $\chi(\mathrm{Spin}(p,q))$ (because χ is a homomorphism). In order to prove this theorem, we just need to show that elements of $\chi(\Gamma(p,q)^+)$ which do not belong to $\chi(\mathrm{Spin}(p,q))$ is h.

Is is no difficult to see that $\chi(e_0 \cdot e_1 \cdot e_2 \cdot e_3)e_i = -e_i$ for $i = 0 \ldots 3$: just write

$$\chi(e_0 \cdot e_1 \cdot e_2 \cdot e_3)e_i = e_0 \cdot e_1 \cdot e_2 \cdot e_3 \cdot e_i \cdot e_3^{-1} \cdot e_2^{-1} \cdot e_1^{-1} \cdot e_0^{-1} \tag{68.45}$$

and use the commutation relations. An easy computation gives $N(e_0 \cdot e_1 \cdot e_2 \cdot e_3) = -1$; then this is not in $\mathrm{Spin}(p,q)$ by remark 68.15. □

We write it by the exact sequence

$$\mathbb{Z}_2 \hookrightarrow \mathrm{Spin}(p,q) \xrightarrow{\chi} \mathrm{SO}_0(p,q) \tag{68.46}$$

we say that the group $\mathrm{Spin}(p,q)$ is a **double covering** of $\mathrm{SO}_0(p,q)$.

Lemma 68.29.
If $\pi\colon \tilde{X} \to X$ is a covering which satisfies

(1) X is path connected,

(2) $\forall x \in X$, $\tilde{X}_x := \pi^{-1}(x)$ is path connected in $\tilde{X}$ i.e. *for all a, $b \in \tilde{X}$, there exist a path in $\tilde{X}$ which joins a and b,*

then $\tilde{X}$ is path connected. lem_cov_path_con

Proof. If $\tilde{x}$ and $\tilde{y}$ are in $\tilde{X}$, we can suppose that $\pi(\tilde{x}) \neq \pi(\tilde{y})$ (because if $\pi(\tilde{x}) = \pi(\tilde{y})$, the second assumption gives the thesis). We define x and y as their projections: $x = \pi(\tilde{x})$ and $y = \pi(\tilde{y})$. Let γ be a path such that $\gamma(0) = x$ and $\gamma(1) = y$, and $\tilde{\gamma}$ be the lift of γ in $\tilde{X}$ which contains $\tilde{x}$: $\tilde{\gamma}(0) = \tilde{x}$ and $\pi(\tilde{\gamma}(1)) = \gamma(1) = y$. Then $\tilde{\gamma}(1)$ lies in $\tilde{X}_y$. Therefore, we can consider γ' which joins $\tilde{\gamma}(1)$ and $\tilde{y}$.

So, $\gamma' \circ \tilde{\gamma}$ is a path which contains $\tilde{x}$ and $\tilde{y}$. □

Proposition 68.30.
The group $\mathrm{Spin}(p,q)$ *is connected.*

Proof. We will prove that the covering $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$ fulfils lemma 68.29. We just have to show that $\mathrm{Spin}(p,q)$ fulfills the second assumption of the lemma. First note that $\chi(\tilde{x}) = \chi(\tilde{y})$ implies $\chi(\tilde{x}\tilde{y}^{-1}) = e$, and then $\tilde{x} = \pm\tilde{y}$ because of proposition 68.21. Since the other case is trivial, we can suppose $\tilde{x} = -\tilde{y}$.

It remains to prove that for every $g \in \mathrm{Spin}(p,q)$, there is a path in $\mathrm{Spin}(p,q)$ which joins g and $-g$. The answer is given by the path $t \mapsto \gamma(t)g$ where

$$\gamma(t) = \exp(te_1 \cdot e_2) = \cos(t)(-1) + \sin(t)e_1 \cdot e_2$$

which satisfies $\gamma(0) = 1$ and $\gamma(\pi) = -1$. □

Proposition 68.31.
The homomorphism $\tilde{\rho}$ *restricts to a homomorphism* $\tilde{\rho}\colon \mathrm{Spin}(p,q) \to \mathrm{GL}(\Lambda^+W)$.

Proof. An element in $\mathrm{Spin}(p,q)$ reads $s = cv_1 \cdots v_{2r}$ and its image by $\tilde{\rho}$ is

$$\tilde{\rho}(s) = c\tilde{\rho}(v_1) \circ \cdots \circ \tilde{\rho}(v_{2r}).$$

When one applies $\tilde{\rho}(v_1)$ to an element $\alpha \in \Lambda^k W$, one obtains a linear combination of an element of $\Lambda^{k-1}W$ and one of $\Lambda^{k+1}W$. The element $\tilde{\rho}(s)$ being an even composition of such maps, its transforms an element of Λ^+W into an element of Λ^+W. □

Notice that an element of V —no $V^{\mathbb{C}}$— is represented on Λ^+W by complex matrices. This is not a problem. In the case of $\mathbb{R}^{1,3}$, we have $\dim \Lambda^+W = 2$ and thus

$$\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) \subset \mathrm{GL}(2,\mathbb{C}).$$

The following is the lemma 8.5 (page 57) of [967].

Lemma 68.32.
Let $\rho\colon \mathrm{Cl}(p,q) \to \mathrm{Hom}_{\mathbb{C}}(E,E)$ *be a representation of the Clifford algebra on a vector space* E*. If* $p+q \geqslant 2$*, then for all* $s \in \mathrm{Spin}(p-1,q) \subset \mathrm{Cl}(p,q)$*,*

$$\det{}_{\mathbb{C}}\big(\rho(s)\big) = \pm 1.$$

Proof. No proof. □

Theorem 68.33.
The representation $\tilde{\rho}$ *provides a group isomorphism*

$$\mathrm{Spin}(1,3) \simeq \mathrm{SL}(2,\mathbb{C})$$

Proof. In the case $p = 2$, $q = 3$, the lemma assures us that for each s in the spin group, $\det \tilde{\rho}(s) = 1$. Since $\mathrm{Spin}(1,3)$ is connected and the determinant function is continuous, we deduce that $\det \tilde{\rho}(s) \equiv 1$. This proves that $\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) \subset \mathrm{SL}(2,\mathbb{C})$. The proposition 52.5 thus implies that

$$\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) = \mathrm{SL}(2,\mathbb{C}),$$

but from $\mathrm{Cl}(1,3)$, the representation $\tilde{\rho}$ is yet injective. *A forciori*, the representation $\tilde{\rho}$ is injective from $\mathrm{Spin}(1,3)$. This finishes the proof. □

68.3.2 Redefinition of $\mathrm{Spin}(V)$

As it, this new definition only holds when g is positive defined.

¡¡ Avertissement/question au lecteur !! 68.34
When we work with a signature (p,q), maybe we only get the connected part. To be checked.

Let us take v, $x \in V$ with $g(v,v) = 1$. We have

$$-vxv^{-1} = -vxv = -2g(x,v)v + xv^2 = x - 2g(x,v)v \in V.$$

The effect was to reverse the v component of x; the map $x \mapsto -vxv^{-1}$ is σ^v. Now, when $\lambda \in U(1)$ and $w = \lambda v$, we also have that $x \mapsto -wxw^{-1}$ is σ^v. Now we look at $\chi(a)\colon x \mapsto \alpha(a)xa^{-1}$ with $a = w_1 \dots w_r$, a product of unitary vectors in $V^{\mathbb{C}}$. Explicitly,

$$\chi(a)x = (-1)^r w_1 \dots w_r x w_r^{-1} \dots w_1^{-1},$$

a composition of reflexions in V. When r is even, it is a rotation. We conclude that when a is an even product of unitary vectors in $V^{\mathbb{C}}$, then $\chi(a) \in \mathrm{SO}(V)$. Theorem 68.19 states that any rotation of V is a composition of reflexions. So we define

$$\mathrm{Spin}^c(V) = \{w_1 \dots w_{2k} \text{ tel que } w_j \in V^{\mathbb{C}},\ w_j^* w_j = 1\} \subset \mathrm{Cl}^{\mathbb{C}\,0}(V), \tag{68.47}$$

and $\chi\colon \mathrm{Spin}^c(V) \to \mathrm{SO}(V)$ is a surjective group homomorphism. The inverse in $\mathrm{Spin}^c(V)$ is given by

$$(w_1 \dots w_{2k})^{-1} = w_{2k}^* \dots w_1^* = \overline{w_{2k}} \dots \overline{w_1}.$$

In the real case, proposition 68.21 says that $\ker\chi = \mathbb{R}^\times$. In the complex case we get $\ker\chi = \mathbb{C}^\times$ and, when we look at $\ker\chi|_{\mathrm{Spin}^c(V)}$, we find

$$\ker\chi = U(1). \tag{68.48}$$

Then we find the short exact sequence

$$1 \xrightarrow{\mathrm{Id}} U(1) \xrightarrow{\mathrm{Id}} \mathrm{Spin}^c(V) \xrightarrow{\chi} \mathrm{SO}(V) \xrightarrow{\mathrm{Id}} 1. \tag{68.49}$$

Let $u = w_1 \dots w_{2k} \in \mathrm{Spin}^c(V)$ with $w_j = \lambda_j v_j$ and $\lambda_j \in V$, so $\tau(u) = w_{2k} \dots w_1$ and

$$\tau(u)u = w_{2k} \dots w_1 w_1 \dots w_{2k} = \lambda_1^2 \dots \lambda_{2k}^2 \in U(1).$$

This proves that $\tau(u)u$ is central in $\mathrm{Spin}^c(V)$. We define the homomorphism

$$\begin{aligned} \nu\colon \mathrm{Spin}^c(V) &\to U(1) \\ u &\mapsto \tau(u)u. \end{aligned} \tag{68.50}$$

This is a homomorphism because

$$\begin{aligned} \nu(u_1u_2) = \tau(u_1u_2)u_1u_2 = \tau(u_2)\underbrace{\tau(u_1)u_1}_{\text{central}}u_2 = \tau(u_2)u_2\tau(u_1)u_1 \\ = \nu(u_2)\nu(u_1) = \nu(u_1)\nu(u_2). \end{aligned}$$

The map ν naturally restricts to $U(1)$ as

$$\nu(\lambda) = \lambda^2.$$

The combined map $(\chi,\nu)\colon \mathrm{Spin}^c(V) \to \mathrm{SO}(V) \times U(1)$ has kernel $\{\pm 1\}$. We define

$$\mathrm{Spin}(V) = \ker\nu|_{\mathrm{Spin}^c(V)}. \tag{68.51}$$

eq_defSpindeux

Lemma 68.35.
This group is the same as the one defined in equation (68.32).

Proof. Let $u \in \mathrm{Spin}(V)$ (in the sense of equation (68.51)). The fact for u to belongs to $\mathrm{Spin}(V)$ implies the two following:

(1) $u \in \mathrm{Spin}^c(V) \Rightarrow u^* u = 1$,

(2) $u \in \ker \nu \Rightarrow \tau(u)u = 1$.

The second point says that $u^{-1} = \tau(u)$, which is a first good point to fit the first definition of $\mathrm{Spin}(V)$. Now we have to prove that $u \in \Gamma^+(V)$: u must be invertible and $\chi(u)x$ must belongs to V for all $x \in V$. These two points are contained in the definition of $\mathrm{Spin}^c(V)$. □

Let us see in the new definition how is $\chi\colon \mathrm{Spin}(V) \to \mathrm{SO}(V)$. On $\mathrm{Spin}^c(V)$, we have $\ker \chi = U(1)$, but on $\mathrm{Spin}(V)$ we require moreover $\tau(u)u = 1$, thus an element of $\ker \chi$ in $\mathrm{Spin}(V)$ fulfils $\tau(\lambda)\lambda = 1$, so that $\lambda = \{\pm 1\}$. We conclude that $\ker \chi|_{\mathrm{Spin}(V)} = \{\pm 1\}$, and then that $\mathrm{Spin}(V)$ is a double covering of $\mathrm{SO}(V)$.

68.3.3 A few about Lie algebra

Proposition 68.36.
We have an isomorphism

$$\mathfrak{spin}(p,q) \simeq \mathfrak{so}(p,q)$$

between the Lie algebras of $\mathrm{Spin}(p,q)$ *and* $\mathrm{SO}(p,q)$. prop:spin_so

Proof. Using the second part of lemma 46.7, with the map $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}(p,q)$, we find that $d\chi_e(\mathfrak{spin}(p,q)) = \mathfrak{so}(p,q)$. Then we know (lemma 46.8) that

$$\mathfrak{so}(p,q) = \mathfrak{spin}(p,q)/\ker d\chi_e.$$

On the other hand, the first part of the same lemma gives us that $\chi^{-1}(e)$ is a Lie subgroup of $\mathrm{Spin}(p,q)$ whose Lie algebra is $\ker d\chi_e$. But $\chi^{-1}(e) = \mathbb{Z}_2$, so $\ker d\chi_e = \{0\}$. □

Let us now shortly speak about the Lie algebra of $\Gamma(p,q)^+$. A basis of $\mathrm{Cl}(p,q)^+$ is

$$\{1, \gamma_0 \cdot \gamma_1, \gamma_0 \cdot \gamma_1, \gamma_0 \cdot \gamma_3, \gamma_0 \cdot \gamma_1 \cdot \gamma_2 \cdot \gamma_3\}.$$

Thanks to the anticommutation relations, we don't need $\gamma_1 \cdot \gamma_2$ in the basis.

Remember that $\Gamma(p,q)^+$ is the set of the $x \in \mathrm{Cl}^+(p,q)$ such that $x \cdot v \cdot \alpha(x^{-1})$ lies in V for all $v \in V$. Let $x(t)$ be a path in $\Gamma(p,q)^+$ such that $x(0) = e$ and $\dot{x}(0) = X$. Differentiating the definition relation, we find

$$\dot{x} \cdot v \cdot \alpha(x^{-1})|_0 + x \cdot v \cdot (-)\alpha(\dot{x})|_0 = X \cdot v - v \cdot X,$$

therefore

$$\mathfrak{Lie}(\Gamma(p,q)^+) = \left\{X \in \mathrm{Cl}^+(p,q) \text{ such that } X \cdot v - v \cdot X \in V, \forall v \in V\right\}.$$

It is clear that $\mathbb{C}$ is a subset of $\mathfrak{Lie}(\Gamma(p,q)^+)$, and that V is not. The following computation shows that $V \cdot V$ is a subset $\mathfrak{Lie}(\Gamma(p,q)^+)$:

$$a \cdot b \cdot v - v \cdot a \cdot b = 2\eta(v,a)b - 2\eta(v,b)a.$$

We can also check that $V \cdot V \cdot V \cdot V \cap \mathfrak{Lie}(\Gamma(p,q)^+) = \varnothing$. A basis of $\mathfrak{Lie}(\Gamma(p,q)^+)$ is

$$\{1, e_\alpha \cdot e_\beta \text{ tel que } \alpha < \beta\}$$

We know that $\ker[\chi\colon \Gamma(p,q)^+ \to \mathrm{SO}(p,q)] = \mathbb{R}_0$. So the kernel of the restriction of $d\chi_e$ to $\mathfrak{Lie}(\Gamma(p,q)^+)$ is the Lie algebra of $\mathbb{R}_0$ (see lemma 46.7), which is $\mathbb{R}$. Therefore, a basis of $\mathfrak{spin}(p,q)$ is

$$\{e_\alpha \cdot e_\beta \text{ tel que } \alpha < \beta\}.$$

68.3.4 Grading ΛW

We already know that $\Lambda W = \mathbb{C} \oplus W \oplus \Lambda^2 W$. This space can be written as

$$\Lambda W = \Lambda W^+ \oplus \Lambda W^-,$$

with $\Lambda W^+ = W$ and $\Lambda W^- = \mathbb{C} \oplus \Lambda^2 W$. The interest of such a decomposition lies in the definition of an action of $\mathrm{Cl}^+(p,q)$ on ΛW. This action will be defined by $\bullet$: $\mathrm{Cl}^+(p,q) \times \Lambda W \to \Lambda W$,

$$x \bullet \alpha = \tilde{\rho}(x)\alpha$$

for any x in $\mathrm{Cl}^+(p,q)$ and any α in ΛW (see definition 68.6).

Proposition 68.37.
This action preserves the grading of ΛW:

$$\begin{aligned} \mathrm{Cl}^+(p,q) \bullet \Lambda W^+ &= \Lambda W^+ \\ \mathrm{Cl}^+(p,q) \bullet \Lambda W^- &= \Lambda W^-. \end{aligned} \tag{68.52}$$

Proof. For $x \in \mathbb{C}$, theses equalities are obvious. We have to check it for $x = e_i \cdot e_j$. Here, we will just check that $(e_1 \cdot e_0) \bullet (v \wedge w) \in \Lambda W^+$. This follows from a simple computation:

$$\begin{aligned} \tilde{\rho}(e_1)\tilde{\rho}(f_0 + g_0)(v \wedge w) &= \tilde{\rho}(f_1 + g_1)\left[-\eta(g_0, v)w + \eta(g_0, w)v\right] \\ &= -\eta(g_0, v) f_1 \wedge w + \eta(g_0, w) f_1 \wedge v \\ &\quad + \eta(g_0, v)\eta(g_1, w) - \eta(g_0, w)\eta(g_1, v). \end{aligned} \tag{68.53}$$

□

Since $\mathrm{Spin}(p,q)$ is a subgroup of $\mathrm{Cl}^+(p,q)$, the following two are representations of $\mathrm{Spin}(p,q)$: $\rho^\pm$: $\mathrm{Spin}(p,q) \times \Lambda W^\pm \to \Lambda W^\pm$,

$$\begin{aligned} \rho^-(s)w^- &= \tilde{\rho}(s)w^-, \\ \rho^+(s)w^+ &= \tilde{\rho}(s)w^+, \end{aligned} \tag{68.54}$$

for $w^\pm$ in $\Lambda W^\pm$. This is no more than the fact that $\tilde{\rho}$ is reducible and that two invariant subspaces are ΛW^+ and ΛW^-.

68.3.5 Clifford algebra for $V = \mathbb{R}^2$

cliffR2

68.3.5.1 General definitions

The whole construction can also be applied to $V = \mathbb{R}^2$ with the Euclidean metric. This is our business now. We take the complex vector space $V^{\mathbb{C}}$ and an orthonormal basis $\{e_1, e_2\}$. As before, we define

$$f_1 = \frac{1}{2}(e_1 + ie_2), \qquad g_1 = \frac{1}{2}(e_1 - ie_2).$$

There are no difficulties to see that $Span(f_1)$ is a completely isotropic subspace of $V^{\mathbb{C}}$. Thus we define $W = \mathbb{C}f_1$, $\Lambda W = \mathbb{C} \oplus W$, $\Lambda W^+ = \mathbb{C}$, and $\Lambda W^- = W$. The homomorphism $\tilde{\rho}$: $V^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ in ΛW is defined by

$$\begin{aligned} \tilde{\rho}(f_1)\alpha &= f_1 \wedge \alpha, \\ \tilde{\rho}(g_1)\alpha &= -i(g_1)\alpha, \end{aligned} \tag{68.55}$$

where α is any element of ΛW. In the basis $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we easily find that

$$\tilde{\rho}(e_1) = \begin{pmatrix} 0 & -\frac{1}{2} \\ 1 & 0 \end{pmatrix}, \quad \tilde{\rho}(e_2) = \begin{pmatrix} 0 & -\frac{i}{2} \\ -i & 0 \end{pmatrix}.$$

For $c \in \mathbb{R}$ we also have $\tilde{\rho}(c)f_1 = cf_1$ and $\tilde{\rho}(c)1 = c$, thus we assign the matrix $\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$ to the number c.

As before, we define $\gamma_i = \sqrt{2}\tilde{\rho}(e_i)$. We immediately have $\gamma_1\gamma_2 + \gamma_2\gamma_1 = 0$ and $\gamma_i\gamma_i = -2\mathbb{1}$, so that the γ's satisfy the Clifford algebra for the euclidian metric.

For notational conveniences, it proves useful to make a change of basis so that we get

$$\gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_2 = -\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{68.56}$$ gammaR2

The algebra Cl(2) is isomorphic to the algebra which is generated by direct sum $\mathrm{Cl}(2) \simeq \mathbb{R} \oplus \gamma_1 \oplus \gamma_2 \oplus \mathbb{R}\gamma_1\gamma_2$. A general element of Cl(2) can be written as $x\gamma_1 + y\gamma_2 + x'\mathbb{R} + y'\gamma_1\gamma_2$. In the representation of $\tilde{\rho}$, a general element of Cl(2) is therefore

$$\begin{pmatrix} x' + iy' & x + iy \\ -x + iy & x' - iy' \end{pmatrix},$$

so that we can write the Clifford algebra of $\mathbb{R}^2$ as

$$\mathrm{Cl}(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\}.$$

The following four matrices provide a basis:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad i = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \qquad j = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad k = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{68.57}$$ pauli

We can check that these matrices satisfies the quaternionic algebra:

$$\begin{aligned} i^2 = j^2 = k^2 = -1 \\ ij = -ji = k, \\ jk = -kj = i, \\ ki = -ik = j. \end{aligned} \tag{68.58}$$

The algebra $\mathrm{Cl}(2) = \mathbb{H}$ is represented by $\tilde{\rho}$ on $\mathbb{C}^2$ by the **Pauli matrices** $1, i, j, k$ which are given by (68.57).

68.3.5.2 The maps α and τ

What are the matrices which represent V? These are $\tilde{\rho}(e_1)$ and $\tilde{\rho}(e_2)$. Thus we can write $V = \mathrm{Span}_{\mathbb{R}}\{\gamma_1, \gamma_2\} = \mathrm{Span}_{\mathbb{R}}\{j, k\}$, or

$$V = \left\{ \begin{pmatrix} 0 & \xi \\ -\overline{\xi} & 0 \end{pmatrix} : \xi \in \mathbb{C} \right\}.$$

As before, α is the unique homomorphic extension to Cl(2) of $-\mathrm{Id}$ on V. From the definitions, we get $\alpha(j) = -j$, $\alpha(k) = -k$. The extension present no difficult. For example: $\alpha(i) = \alpha(jk) = \alpha(j)\alpha(k) = jk = i$, but $\alpha(jk) = \alpha(i)$; then $\alpha(i) = i$. The same gives $\alpha(1) = 1$.

The case of τ is treated in similar way. We find: $\tau(j) = j$, $\tau(k) = k$, $\tau(i) = -i$, $\tau(1) = 1$.

Now, we can find the group $\Gamma_{(2)}$. The condition for $x \in \mathrm{Cl}(2)$ to be in $\Gamma_{(2)}$ is $\alpha(x)yx^{-1}$ to belongs to V for all $y \in V$. We put

$$x = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, \qquad \alpha(x) = \begin{pmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix}.$$

A typical y in V is

$$y = \begin{pmatrix} 0 & \eta \\ -\overline{\eta} & 0 \end{pmatrix}.$$

A few computation gives:

$$\alpha(x)yx^{-1} = \frac{1}{|\alpha|^2+|\beta|^2}\begin{pmatrix} \alpha\eta\overline{\beta}+\beta\overline{\eta\alpha} & \alpha\alpha\eta-\beta\beta\overline{\eta} \\ \overline{\beta}\overline{\beta}\eta-\overline{\alpha}\overline{\alpha}\eta & \eta\alpha\overline{\beta}+\overline{\alpha\eta}\beta \end{pmatrix}.$$

If we impose it to be of the form $\begin{pmatrix} 0 & \xi \\ -\overline{\xi} & 0 \end{pmatrix}$ for all $\eta \in \mathbb{C}$, we get, for all $\eta \in \mathbb{C}$, $\mathrm{Re}(\overline{\alpha}\beta\overline{\eta}) = 0$, which implies $\overline{\alpha}\beta = 0$. So we conclude:

$$\Gamma_{(2)} = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix}, \begin{pmatrix} 0 & \beta \\ -\overline{\beta} & 0 \end{pmatrix} : \alpha, \beta \in \mathbb{C} \text{ not both equals zero} \right\}.$$

Be careful on a point: $\Gamma_{(2)}$ is the *multiplicative* group generated by these two matrices, not the additive one.

68.3.5.3 The spin group

It present no difficult to find that

$$\Gamma^+_{(2)} = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix} : \alpha \neq 0 \right\}. \tag{68.59}$$

The **spin group** is made of elements of $\Gamma^+_{(2)}$ which satisfy $\tau(x) = x^{-1}$. We know that $\tau\begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix} = \begin{pmatrix} \overline{\alpha} & 0 \\ 0 & \alpha \end{pmatrix}$ and that $\begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix}^{-1} = \frac{1}{\alpha\overline{\alpha}}\begin{pmatrix} \overline{\alpha} & 0 \\ 0 & \alpha \end{pmatrix}$. Thus the condition $\tau(x) = x^{-1}$ becomes $|\alpha|^2 = 1$. The first conclusion is that

$$\mathrm{Spin}(2) = U(1). \tag{68.60}$$

A typical s in Spin(2) is

$$s = e^{i\theta} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

The next point is to see the action of Spin(2) on V. The action of $s \in \mathrm{Spin}(2)$ on a vector $v \in V$ is still defined by $s \bullet v = \chi(s)v = \alpha(s) \cdot v \cdot s^{-1}$. More explicitly:

$$\chi(s)v = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}\begin{pmatrix} 0 & z \\ -\overline{z} & 0 \end{pmatrix}\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} = \begin{pmatrix} 0 & e^{2i\theta}z \\ -e^{-2i\theta}\overline{z} & 0 \end{pmatrix}, \tag{68.61}$$

where the matrix $\begin{pmatrix} 0 & z \\ \overline{z} & 0 \end{pmatrix}$ denotes the representation of the vector v of V. This equality can be written $e^{i\theta} \cdot v = e^{2i\theta}v$. If we note $v = v_1 + iv_2 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, we get

$$e^{2i\theta} \bullet v = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Therefore, we can write

$$\chi(e^{i\theta}) = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}.$$

So χ projects $U(1)$ into SO(2) with a kernel $\mathbb{Z}_2$, for this reason, we say that $U(1)$ is a **double covering** of SO(2). We note it

$$\mathbb{Z}_2 \to U(1) \xrightarrow{\chi} \mathrm{SO}(2). \tag{68.62}$$

68.4 Clifford modules

susec_Cliffmodule

References: [970, 969].

Let M be a manifold. We denote by $\mathrm{Cl}^{\mathbb{C}}(M)$ the bundle whose fibre at $x \in M$ is the complex Clifford algebra of the metric g_x: $\mathrm{Cl}^{\mathbb{C}}(M)_x = \mathrm{Cl}^{\mathbb{C}}(g_x)$. We define the important map

$$\begin{aligned} \gamma\colon \Gamma(M, \mathrm{Cl}^{\mathbb{C}}(M)) &\to \mathfrak{B}(\mathscr{H}) \\ \gamma(dx^\mu) &\mapsto \gamma^\mu(x) \end{aligned} \tag{68.63}$$

which can be extended to the whole Clifford algebra.

Let V be a vector space endowed with a bilinear symmetric form. We consider $\mathrm{Cl}(V)$, the corresponding Clifford algebra. A **Clifford module** is a real vector space E with a $\mathbb{Z}_2$-graduation and a morphism

$$\rho_E\colon \mathrm{Cl}(V) \to \mathrm{End}(E)$$

of $\mathbb{Z}_2$-graded vector spaces. It is defined by a linear map $\rho_E\colon V \to \mathrm{End}(V)$ such that

$$\rho_E(v)\rho_E(w) + \rho_E(w)\rho_E(v) = B(v,w)\,\mathrm{Id} \tag{68.64}$$

for every v, $w \in E$. The element $\rho_E(x)v$ will often be denoted by $x \cdot v$ and the operation ρ_E is the **Clifford multiplication**. The **dual module** E^* is defined by $\rho_{E^*}(x) = \rho_E(x^t)^*$, i.e.

$$\langle \rho_{E^*}(x)\psi, v\rangle = (-1)^{|\psi||x|}\langle \psi, \rho_E\big(\tau(x)\big)v\rangle \tag{68.65}$$

for every $\psi \in E^*$ and $v \in E$. Here

Let $\mathfrak{A}$ be a $\mathbb{Z}_2$-graded subalgebra of $\mathrm{Cl}(V)$ and E_1, a $\mathfrak{A}$-module. Then the space

$$E = \mathrm{Ind}_{\mathfrak{A}}^{\mathrm{Cl}(V)}(E_1) = \mathrm{Cl}(V) \otimes_{\mathfrak{A}} E_1$$

has a structure of Clifford module, the **induced module**. The tensor product $\otimes_{\mathfrak{A}}$ is the usual one modulo the subspace spanned by elements of the form

$$x \otimes a \cdot y - xa \otimes y$$

for every x, $a \in \mathrm{Cl}(V)$ and $y \in E_1$. In a similar way, if E is a complex vector space we have a notion of $\mathrm{Cl}^{\mathbb{C}}(V)$-module.

Let $x \in \mathrm{Cl}(V)$ be such that $x^2 = 1$. In that case the Clifford multiplication $\rho_E(x)$ decomposes E in eigenspaces

$$E^{\pm} = \frac{1}{2}\big(1 \pm \rho_E(x)\big)E.$$

If V is a n-dimensional vector space with an oriented orthonormal basis $\{e_1, \dots, e_n\}$, the algebra $\mathrm{Cl}(V)$ has a **volume element** $\omega = e_1 e_2 \dots e_n$ which does not depend on the choice of the basis. The volume element squares to

$$\omega^2 = (-1)^{n(n+1)/2}. \tag{68.66}$$

In the complex case, we consider the complex vector space $V^{\mathbb{C}}$ and the complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(V) = \mathrm{Cl}(V) \otimes_{\mathbb{R}} \mathbb{C}$, and the volume element is defined as

$$\omega_{\mathbb{C}} = i^{[(n+1)/2]}\omega. \tag{68.67}$$

where $[x]$ is denotes the integer part of x. Performing a separate computation for n even or odd, it is easy to see that in both case,

$$\omega_{\mathbb{C}}^2 = 1. \tag{68.68}$$

So in the complex case we always have an element in $\mathrm{Cl}(V)$ which squares to 1, and a $\mathrm{Cl}^{\mathbb{C}}(V)$-module W always accepts a decomposition as $W^{\pm} = \frac{1}{2}(1 + \omega_{\mathbb{C}})W$.

One says that a representation ρ of $\mathrm{Cl}(V)$ on W is **reducible** if there exists a splitting $W = W_1 \oplus W_2$ such that $\rho(\mathrm{Cl}(V))W_i \subset W_i$. If the representation is not reducible, it is said to be irreducible. Two representations $\rho_j\colon \mathrm{Cl}(V) \to \mathrm{End}(W_j)$ are **equivalent** if there exists a linear isomorphism $F\colon W_1 \to W_2$ such that $F \circ \rho_1(x) \circ F^{-1} = \rho_2(x)$ for every $x \in \mathrm{Cl}(V)$.

Proposition 68.38.
The real Clifford algebra has

$$\begin{cases} 2 & \text{if } n+1=0 \mod 4 \\ 1 & \text{otherwise} \end{cases}$$

inequivalent irreducible representations. The complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(V)$ *has*

$$\begin{cases} 2 & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{is even} \end{cases}$$

inequivalent irreducible representations.

Proof. No proof. □

If M is a manifold, we denote by $\mathrm{Cl}(M) = \mathrm{Cl}(TM)$ the bundle whose fiber at x is the Clifford algebras of T_xM. We consider an orthonormal basis $\{e_i\}$ and if Σ is a multi-index $\{1 \leqslant \sigma_1, \dots, \leqslant \sigma_t \leqslant m\}$, we pose $e_\Sigma = e_{\sigma_1} \dots e_{\sigma_t} \in \mathrm{Cl}(M)$. By convention, $e_\varnothing = 1$. Since the elements e_i are ordered, they provide an orientation:

$$d\,\mathrm{Vol} = e_1 \wedge \dots \wedge e_m \in \bigwedge^m (M). \tag{68.69}$$

Since the map $e_{\sigma_1} \wedge \dots \wedge e_{\sigma_t} \mapsto e_{\sigma_1 \dots e_{\sigma_t}}$ is an isomorphism between $\mathrm{Cl}(M)$ and $\bigwedge(M)$, we say that $d\,\mathrm{Vol} \in \mathrm{Cl}(M)$. Now we define

$$\kappa = i^{-[(m+1)/2]} d\,\mathrm{Vol},$$

which is nothing else that the volume form normalised in such a way that $\kappa^2 = 1$. If m is even, it anti-commutes with TM, and if m is odd, it commutes with TM.

Let V be a m-dimensional real vector space, and $\mathrm{Cl}^{\mathbb{C}}(V)$, the corresponding complex Clifford algebra.

Lemma 68.39.
Every $\mathrm{Cl}^{\mathbb{C}}(V)$*-module accepts an unique decomposition as sum of irreducible representations as follows*

(1) if $m = 2n$*, there exists one and only one irreducible* $\mathrm{Cl}^{\mathbb{C}}(V)$*-module* Δ *and* $\dim(\Delta) = 2n$*,*

(2) if $m = 2n+1$*, we have two inequivalent irreducible modules* $\Delta_\pm$ *with* $\gamma(\kappa) = \pm 1$ *on* $\Delta_\pm$ *and* $\dim(\Delta_\pm) = 2^n$*.*

Proof. No proof. □

Let V be a vector bundle over M. A structure of $\mathrm{Cl}(M)$-module over V is a morphism of unital algebra $\gamma\colon \mathrm{Cl}(M) \to \mathrm{End}(V)$. When one has a basis $\{e_i\}$ of V, we pose $\gamma_i = \gamma(e_i)$. The following lemma is the lemma 1.2 of [970].

LemGammaBaseConstant

Lemma 68.40.
Let V *be a* $\mathrm{Cl}(V)$*-module and* $\{e_i\}$*, an orthonormal basis for* TM *on a contractible open set* V*. Then there exists a local frame for* V *such that the matrices* $\gamma(e_i)$ *are constant.*

We also define $\gamma^i = \gamma(dx^i) = g^{ij}\gamma_j$. One easily proves that

$$\gamma^i\gamma^j + \gamma^j\gamma^i = -2g^{ij} \tag{68.70}$$

where (g^{ij}) is the inverse matrix of (g_{ij}). If the endomorphisms γ_i are constant in the basis $\{e_i\}$, then the endomorphisms γ^i are constant in the basis $\{f_i = g_{ki}e_k\}$.

68.5 Spin structure

sec:spin_str

We consider a (pseudo-)Riemannian manifold (M, g) with metric signature (p, q), and $\mathrm{SO}(M)$, its frame bundle; it admits a $\mathrm{SO}(p, q)$-principal fiber bundle structure which is well defined by the metric g (see 56.6.4.1).

Definition 68.41.
*We say that (M, g) is a **spin manifold** if there exists a $\mathrm{Spin}(p, q)$-principal bundle P over M and a principal bundle homomorphism $\varphi\colon P \to \mathrm{SO}(M)$ which induced covering for the structure groups is χ, i.e. $\varphi(\xi \cdot s) = \varphi(\xi) \cdot \chi(s)$. A choice of P and φ is a **spin structure** on M.* defvarspin

$$\begin{array}{ccccc} \mathrm{Spin}(p,q) & \rightsquigarrow & P & \xrightarrow{\varphi} & \mathrm{SO}(M) \\ & & \searrow^{\pi} & & \swarrow_{p} \\ & & & M & \end{array} \qquad \mathrm{SO}(p,q) \leftsquigarrow \mathrm{SO}(M)$$

The wavy arrows mean "structural group of".

Remarque 68.42.
When we will use the concept of spin structure in the physical oriented chapters, we will naturally use $\mathrm{SL}(2, \mathbb{C})$ as group instead of $\mathrm{Spin}(p, q)$. The isomorphism $\mathrm{SL}(2, \mathbb{C}) \simeq \mathrm{Spin}(1, 3)$ is proved in [967]. A physical motivation of such a structure is given at page 3720.

68.5.1 Example: spin structure on the sphere S^2

It is no difficult to see that $\mathrm{SO}(S^2) \simeq \mathrm{SO}(3)$. Indeed, each element of $\mathrm{SO}(S^2)$ is described by three orthonormal vectors: one which point to an element x of S^2 and two which gives a basis of $T_x S^2$. The action $\mathrm{SO}(3) \times S^2 \to S^2$ is transitive, and the stabilizer of any element is $\mathrm{SO}(2)$.

We define $\alpha\colon \mathrm{SO}(3)/\mathrm{SO}(2) \to S^2$ by $\alpha(g\,\mathrm{SO}(2)) = g$. Proposition 53.101 shows that α is a diffeomorphism. Then

$$S^2 = \frac{\mathrm{SO}(3)}{\mathrm{SO}(2)}.$$

On the other hand, we know that

$$T_e SU(2) = su(2) = \left\{ \begin{pmatrix} ix & \xi \\ -\bar{\xi} & -ix \end{pmatrix} : \xi \in \mathbb{C}, x \in \mathbb{R} \right\}. \tag{68.71}$$

explsu2

It is a classical result that $\mathfrak{su}(2) \simeq \mathbb{R}^3$ not only as set but also as metric space with the identification

$$\langle X, Y \rangle = -\frac{1}{2} \mathrm{Tr}(XY),$$

for all X, $Y \in su(2)$. As we are in matrix groups, we know (see [872] to get more details) that $Ad_x Y = xYx^{-1}$. In our case, this gives the formula

$$\langle Ad(g)X, Ad(g)Y \rangle = \langle X, Y \rangle.$$

We can now state the result for S^2.

Proposition 68.43.
The manifold S^2 with the usual metric induced from $\mathbb{R}^3$ admits the following spin structure:

$$\begin{array}{ccccc} \mathrm{Spin}(2) & \rightsquigarrow & SU(2) & \xrightarrow{\varphi = Ad} & \mathrm{SO}(3) \\ & & \searrow^{\pi} & & \swarrow_{p} \\ & & U(1) & & \mathrm{SO}(2) \\ & & & S^2 & \end{array}, \tag{68.72}$$

spins2

where the arrow $X \xrightarrow[G]{f} Y$ means that G is the kernel of the map $f\colon X \to Y$.

Proof. First, let us precise the concept of frame bundle for S^2, and how it is well described by SO(3). Let $\{e_1, e_2, e_3\}$ be the canonical basis of $\mathbb{R}^3$. To $A \in \mathrm{SO}(3)$, we make correspond the basis $\{Ae_2, Ae_3\}$ at the point Ae_1 of S^2. The projection $p\colon \mathrm{SO}(3) \to S^2$ is then defined by $p(A) = Ae_1$. It is clear that we will define the map $\pi\colon SU(2) \to S^2$ in the same way: $\pi(U) = p(Ad(U))$.

For the rest of the demonstration, we will use the "$su(2)$ description" of $\mathbb{R}^3$ given by (68.71) with $\xi = y + iz$.

Now, let us show that $\pi\colon SU(2) \to S^2$ is a Spin(2)-principal bundle. Since we had already shown that $\mathrm{Spin}(2) \simeq U(1)$, we define the right action of Spin(2) on $SU(2)$ by right multiplication: $U \cdot s = Us$ with $s = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$. It is clear that $\pi(Us) = \pi(U)$:

$$Ad(Us)e_1 = (Us)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} s^{-1}U^{-1} = Us\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} s^{-1}U^{-1}, \tag{68.73}$$

because $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ is the vector e_1 in the "$su(2)$ description" of $\mathbb{R}^3$.

In order for $\pi\colon SU(2) \to S^2$ to be a Spin(2)-principal bundle, we still need to show that for all $x \in S^2$,

$$\pi^{-1}(x) = \left\{\xi \cdot g \text{ tel que } g \in \mathrm{Spin}(2)\ \forall \xi \in \pi^{-1}(x)\right\}.$$

Take A, $B \in \pi^{-1}(x)$, i.e. $Ae_1 = Be_1 = x$. We need to find a $s \in \mathrm{Spin}(2)$ such that

$$A = B \cdot s. \tag{68.74}$$

The matrices A and B are such that

$$B^{-1}A\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} A^{-1}B = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{68.75}$$

This implies that $B^{-1}A \in \mathrm{Spin}(2)$. As Ad is surjective from $SU(2)$ into SO(3), a general C in SO(3) which acts on e_1 can be written Ue_1U^{-1} for $U \in SU(2)$ such that $Ad(U) = C$. The condition (68.75) becomes

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} \overline{\alpha} & -\beta \\ \overline{\beta} & \alpha \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

which implies $\alpha = e^{i\theta}$, $\beta = 0$. Then $B^{-1}A$ belongs to Spin(2), and $s = B^{-1}A$ fulfills the condition (68.74).

What about the induced covering for the structural groups? The group Spin(2) is the the structural group of $\pi\colon SU(2) \to S^2$, while SO(2) is the one of $p\colon \mathrm{SO}(3) \to S^2$. Indeed, for each $x \in S^2$, SO(2) acts on T_xS^2, leaving x unchanged. We have the following associations:

$$U \in SU(2) \xrightarrow{\varphi} A \in \mathrm{SO}(3),$$

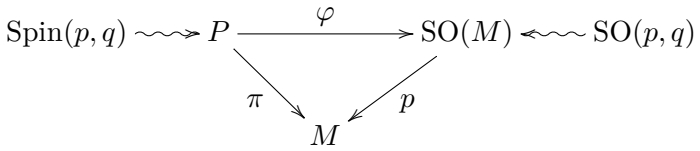
the matrix A being defined by $A \cdot X = UXU^{-1}$. For $s \in \mathrm{Spin}(2)$ we of course also have

$$Us \in SU(2) \xrightarrow{\varphi} As \in \mathrm{SO}(3),$$

with $As \cdot X = UsXs^{-1}U^{-1}$. As we act by Spin(2) on $SU(2)$, in the fibres of SO(3), the action of Spin(2) is –via φ– the composition with $X \to sXs^{-1}$. But this is exactly $\chi(s)X$ because $\alpha(s) = s$, since $s \in \mathrm{Spin}(2)$. □

68.5.2 Spinor bundle

Let us take once again the spin structure on the (pseudo-)Riemannian manifold (M, g):



with $\varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$.

Let us define $S = \Lambda W$, and $\mathcal{S} = P \times_\rho S$. Take $\rho\colon \mathrm{Spin}(p,q) \times \mathcal{S} \to \mathcal{S}$, $\rho(g,s) = \tilde{\rho}(g)s$, where $\tilde{\rho}$ is the spinor representation of $\mathrm{Spin}(p,q)$ on S. We also have $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$, $\chi(g)v = \alpha(g) \cdot v \cdot g^{-1}$, with $\alpha(g) = g$ for $g \in \mathrm{Spin}(p,q)$.

The **spinor bundle** is the associated bundle

$$\mathcal{S} = P \times_\rho S \to M \tag{68.76}$$

A **spinor field** is an element of $\Gamma(\mathcal{S})$, the space of section of the spinor bundle.

On $\mathrm{SO}(M)$, we look at a connection 1-form $\alpha \in \Omega^1(\mathrm{SO}(M), so(\mathbb{R}^m))$, and, if $T(M)$ is the tensor bundle over M, we define a covariant derivative $\nabla^\alpha\colon \mathfrak{X}(M) \times T(M) \to T(M)$ by

$$\widehat{\nabla^\alpha_X s} = \overline{X}\hat{s},$$

for any $s \in T(M)$. See theorem 56.79, and the fact that $T(M)$ can be see as an associated bundle; it is explicitly done for $\mathfrak{X}(M)$ at page 3391.

As seen in 56.13.2, an automatic property of this connection is $\nabla^\alpha g = 0$ if g is the metric of M. The **Levi-Civita connection** is the unique[4] such connection which is torsion-free: $T^{\nabla^\alpha} = 0$.

Proposition 68.44.
The 1-form $\tilde{\alpha} = \varphi^*\alpha \in \Omega^1(P, so(\mathbb{R}^m))$ defines a connection on P. See definition 56.73 and theorem 56.79.

Proof. Let us denote by R_g the right action of $g \in \mathrm{Spin}(p,q)$ on P (*id est* $R_g\xi = \xi \cdot g$), and by $R_u^{\mathrm{SO}(M)}$ the right action of $u \in \mathrm{SO}(p,q)$ on $\mathrm{SO}(M)$. We have to check the usual two conditions of a connection.

First condition. The first one is:

$$(R_g^*\tilde{\alpha})_\xi(\Sigma) = Ad(g^{-1})(\tilde{\alpha}_\xi(\Sigma)),$$

for all $\xi \in P$, and $\Sigma \in T_\xi P$. In order to check this, we first remark that $\varphi \circ R_g = R^{\mathrm{SO}(M)}_{\chi(g)} \circ \varphi$. Indeed, for all $\xi \in P$, definition 68.41 gives us $\varphi(R_g\xi) = \varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$. With this, we can make the following computation:

$$\begin{aligned} R_g^*\tilde{\alpha} = R_g^*\varphi^*\alpha = (\varphi \circ R_g)^*\alpha = (R^{\mathrm{SO}(M)}_{\chi(g)} \circ \varphi)^*\alpha \\ = \varphi^* R^{\mathrm{SO}(M)*}_{\chi(g)}\alpha = \varphi^*(Ad(\chi(g)^{-1}) \circ \alpha). \end{aligned} \tag{68.77}$$

1603r4

The last equality comes from the fact that α is a connection 1-form. As we are in matrix groups, we have $Ad(g)x = gxg^{-1}$, so

$$[Ad(\chi(g))x]v = [\chi(g)x\chi(g)^{-1}]v = \chi(g)[xg^{-1}vg] = gxg^{-1}. \tag{68.78}$$

In the first line, the product is the usual matrix product which can be seen as operator composition.

But $(Ad(g)x)v = gxg^{-1}v$. Then $Ad(g) = Ad(\chi(g))$, if we identify $\mathfrak{spin}(p,q) \simeq \mathfrak{so}(p,q)$ by proposition 68.36. Moreover, the action of Ad is linear, so it commutes with φ^*. With these remarks, we can continue the computation (68.77):

$$\varphi^*(Ad(\chi(g)^{-1}) \circ \alpha) = \varphi^*(Ad(g^{-1}) \circ \alpha) = Ad(g^{-1}) \circ \varphi^*\alpha = Ad(g^{-1}) \circ \tilde{\alpha}. \tag{68.79}$$

This proves the first condition.

Second condition. The second one is $\tilde{\alpha}(A^*_\xi) = -A$ with the definition (56.299). This is also a computation. First remark

$$\tilde{\alpha}_\xi(A^*_\xi) = (\varphi^*\alpha)_\xi(A^*_\xi) = \alpha_{\varphi(\xi)}(\varphi_{*\xi}A^*_\xi).$$

4. We will not prove unicity.

We compute $\varphi_{*\xi}A^*$ with lemma 53.38:

$$\begin{aligned}\varphi_{*\xi}A^* &= \frac{d}{dt}\varphi(\xi \cdot \exp -tA)\Big|_{t=0} = \frac{d}{dt}(R^{\mathrm{SO}(M)}_{\chi(\exp -tA)} \circ \varphi)(\xi)\Big|_{t=0} \\ &= \frac{d}{dt}\varphi(\xi) \cdot \chi(\exp -tA)\Big|_{t=0} = \frac{d}{dt}\varphi(\xi) \cdot \exp(-td\chi_e A)\Big|_{t=0} = (d\chi_e A)^*_{\varphi(\xi)}.\end{aligned} \tag{68.80}$$

But $d\chi_e = \mathrm{Id}_{so(p,q)}$, thus $\varphi_{*\xi}A^* = A^*_{\varphi(\xi)}$. The whole makes that:

$$\tilde{\alpha}_\xi(A^*_\xi) = \alpha_{\varphi(\xi)}(\varphi_{*\xi}A^*_\xi) = \alpha_{\varphi(\xi)}(A^*_{\varphi(\xi)}) = -A.$$

This completes the proof. □

Definition 68.45.
This connection 1*-form on* P *is called the* ***spinor connection****. It gives us a covariant derivative on any associated bundle and in particular on the spinor bundle,* $\tilde{\nabla}\colon \mathfrak{X}(M) \times \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$. spinconn

Proposition 68.46.
If X, $Y \in \mathfrak{X}(M)$ *are such that* $X_x = Y_x$, *then for all* $s \in \Gamma(\mathcal{S})$,

$$(\tilde{\nabla}_X s)(x) = (\tilde{\nabla}_Y s)(x).$$

2303p1

Proof. We just have to show that for all vector field Z such that $Z_x = 0$, $(\tilde{\nabla}_Z s)(x) = 0$. Such a Z can be written as $Z = fZ'$ for a function f on M which satisfies $f(x) = 0$. We have:

$$\tilde{\nabla}_Z s = \tilde{\nabla}_{fZ's} = f\tilde{\nabla}_{Z's},$$

which is obviously zero at x. □

Let $x \in M$ and $\{e_{\alpha\, x}\}$ be an orthonormal basis of T_xM. We can extend it to $\{e_\alpha\}$, a local basis field around x such that e_α is a section of the frame bundle (in other words, we ask the extension to be smooth). The claim of proposition 68.46 is that $\tilde{\nabla}_{e_\alpha}(x)$ is an element of $\mathcal{S}_x$ which doesn't depend on the extension.

68.6 Dirac operator

applgamma

68.6.1 Preliminary definition

Let M be a m-dimensional (pseudo)Riemannian manifold with its spin structure

$$\begin{array}{ccccccc} \mathrm{Spin}(p,q) & \rightsquigarrow & P & \xrightarrow{\varphi} & \mathrm{SO}(M) & \rightsquigarrow & \mathrm{SO}(p,q) \\ & & & \searrow^{\pi} \quad \swarrow_{p} & & & \\ & & & M & & & \end{array}$$

where φ satisfies $\varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$.

Recall that for any vector space, one can see $\mathrm{End}\, V = V^* \otimes V$ with the definition $(v^* \otimes v)w = (v^*w)v$. This allows us to define an action of $\mathrm{Spin}(p,q)$ on $\mathrm{End}\, S$ by defining an action of $\mathrm{Spin}(p,q)$ on S and S^* separately. We know the action

$$\begin{aligned}\mathrm{Spin}(p,q) \times S &\to S \\ (g, v) &\mapsto \tilde{\rho}(g)v,\end{aligned} \tag{68.81}$$

and as action on S^*, we take the dual one

$$\begin{aligned}\mathrm{Spin}(p,q) \times S^* &\to S^* \\ g \cdot \alpha &= \alpha \circ \tilde{\rho}(g^{-1})\end{aligned} \tag{68.82}$$

for all $g \in \mathrm{Spin}(p,q)$ and $\alpha \in S^*$.

Now we can make the following computation with $g \in \mathrm{Spin}(p,q)$, $\alpha \in S^*$ and $v \in S$, using the fact that $\tilde{\rho}$ is linear:

$$\begin{aligned}[g \cdot (\alpha \otimes v)]w &= [(\alpha \circ \tilde{\rho}(g^{-1}))w]\tilde{\rho}(g)v \\ &= \tilde{\rho}\left([(\alpha \circ \tilde{\rho}(g^{-1}))w]g\right) v \\ &= \left[\tilde{\rho}(g) \circ (\alpha \otimes v) \circ \tilde{\rho}(g^{-1})\right]w.\end{aligned} \tag{68.83}$$

Then we write the action of $\mathrm{Spin}(p,q)$ on $\mathrm{End}\, S$ by ($A \in \mathrm{End}\, S$)

$$g \cdot A = \tilde{\rho}(g) \circ A \circ \tilde{\rho}(g^{-1}). \tag{68.84}$$ actspin

Notice that this definition is the one required in condition (56.236).

The tangent bundle T_xM is given with a metric g_x. As usual, we build $S_x = \Lambda W_x$, a completely isotropic subspace of T_xM with respect to the metric g_x, and a representation

$$\tilde{\rho}_x : T_xM \to \mathrm{End}(\Lambda W_x)$$

The first step in the definition of $\gamma(X)$ is to build $\hat{a}_X : P \to \mathrm{End}(\Lambda W)$ setting [5] $\hat{a}_X(p) = \tilde{\rho}(\hat{X}_{\varphi(p)})$.

Lemma 68.47.
The function $\hat{a}$ is equivariant, i.e. it satisfies

$$\hat{a}_X(p \cdot g) = g^{-1} \cdot \hat{a}_X(p) \tag{68.85}$$ equivaX

for all $g \in \mathrm{Spin}(p,q)$.

Proof. It is no more than a simple computation using the equivariance of $\hat{X}$. Indeed:

$$\begin{aligned}\hat{a}_X(p \cdot g) &= \tilde{\rho}(\hat{X}_{\varphi(p \cdot g)}) = \tilde{\rho}(\hat{X}_{\varphi(p)\chi(g)}) = \tilde{\rho}(\chi(g^{-1}) \cdot \hat{X}_{\varphi(p)}) \\ &= \tilde{\rho}(g^{-1} \cdot \hat{X}_{\varphi(p)} \cdot g) = \tilde{\rho}(g^{-1}) \circ \tilde{\rho}(\hat{X}_{\varphi(p)}) \circ \tilde{\rho}(g) \\ &= g^{-1} \cdot \hat{a}_X(p).\end{aligned} \tag{68.86}$$

In the fourth line, the dots mean the Clifford product, and the last equality comes from the definition of the action (68.84) of $\mathrm{Spin}(p,q)$ on $\mathrm{End}\, S$. □

From the discussion of section 56.7.2, the function $\hat{a}_X : P \to \mathrm{End}\, S$ defines a section $a_X : M \to \mathrm{End}\, \mathcal{S}$. We define $\gamma : \mathfrak{X}(M) \to \mathrm{End}\, \Gamma(\mathcal{S})$ by

$$\gamma(X) = a_X. \tag{68.87}$$ EqDefgammax

We immediately have

$$\widehat{\gamma(X)}(p) = \tilde{\rho}(\hat{X}_{\varphi(p)})$$

for any $p \in P$. If we define

$$\widehat{\gamma \cdot a_X}(p) = \widehat{\gamma(X)}(p), \tag{68.88}$$ 3103r1

the map γ can be seen as an action on the section of $\mathcal{S}$. Indeed, $\widehat{\gamma \cdot s}_X$ is an equivariant function:

$$\begin{aligned}\hat{\gamma}(p \cdot g)(\hat{a}_X(p \cdot g)) &= \rho(g)^{-1}\hat{\gamma}(p)\rho(g)\rho(g^{-1})\hat{a}_X(p) \\ &= \rho(g)^{-1}\hat{\gamma}(p)\hat{a}_X(p) \\ &= \rho(g^{-1})\widehat{\gamma \cdot a_X}(p),\end{aligned} \tag{68.89}$$

so that

$$\widehat{\gamma \cdot a_X}(p) = \rho(g^{-1})\widehat{\gamma \cdot a_X}(p).$$

The map $\widehat{\gamma \cdot a_X} : P \to \mathrm{End}\, \Lambda W$ defined by (68.88) is equivariant, and thus defines a section $\gamma \cdot a_X \in \Gamma(\mathcal{S})$, as seen in the section 56.7.2.

5. See subsection 56.7.4 for the definition of $\hat{X}$.

68.6.2 Definition of Dirac

If we consider a basis $\{e_\alpha\}$ of TM, *i.e.* m sections $e_\alpha\colon M \to TM$ such that for all x in M, the set $\{e_{\alpha x}\}$ is a basis of T_xM, we note $\gamma^\alpha := \gamma(e_\alpha) \in \mathrm{End}(\mathcal{S})$.

Remarque 68.48.
This is not always globally possible. The example of the sphere is given in subsection 54.3.2. rem_secnoglobal

For any $s \in \Gamma(\mathcal{S})$, we consider the local[6] section ψ of $\mathcal{S}$ given by

$$\psi(x) = \sum_{\alpha\beta} g_x(e_\alpha, e_\beta)\gamma_x^\beta(\widetilde{\nabla}_{e_\alpha}s)(x).$$

For each $x \in M$, take a A_x in[7] $\mathrm{SO}(g_x)$, and consider the new basis $e'_\alpha = A_\alpha{}^\beta e_\beta$. As A is an isometry, $g_x(e'_\alpha, e'_\beta) = g_x(e_\alpha, e_\beta)$; and since $\tilde\rho$ is linear, $\gamma_x'^\alpha = \tilde\rho_x(e'_{\alpha x}) = A_\alpha{}^\beta\tilde\rho(e_{\beta x}) = A_\alpha{}^\beta\gamma_x^\beta$. In the new basis, the section reads:

$$\begin{aligned}\psi(x) &= \sum_{\alpha\beta\eta\sigma} g_x(e_\alpha, e_\beta)A_\beta{}^\sigma\gamma_x^\sigma(\widetilde{\nabla}_{A_\alpha{}^\eta e_\eta}s)(x)\\ &= \sum_{\alpha\beta\eta\sigma} (A^t)^\eta{}_\alpha g_{\alpha\beta}(x)A_\beta{}^\sigma\gamma_x^\sigma(\widetilde{\nabla}_{e_\eta}s)(x)\\ &= \sum_{\eta\sigma} g_x(e_\eta, e_\sigma)\gamma_x^\sigma(\widetilde{\nabla}_{e_\eta}s)(x),\end{aligned} \tag{68.90}$$

where we used the fact that $A^tgA = g$ and that all the $A_\alpha{}^\beta$ are C^∞ functions on M, so that $\widetilde{\nabla}_{A_\alpha{}^\beta X} = A_\alpha{}^\beta\widetilde{\nabla}_X$. This shows that $\psi(x)$ doesn't depend on the choice of the basis, so it defines a section from the data of s alone.

NORMooUCULooJVYkzO

68.49.
Let's summarize how we build the map $\tilde\rho$.

(1) We have a metric vector space (V, g), and a g-orthonormal (g has signature (p, q)) basis $\{e_1, \dots, e_{p+q}\}$ of V.
(2) We consider the vectors $f_0 = (e_0 + e_p)/2$ and $f_1 = (e_1 + e_{p+1})/2$ (see equations (68.11)).
(3) We set $W = \mathrm{Span}_{\mathbb{C}}\{f_0, f_1\}$.
(4) The map $\tilde\rho\colon V^{\mathbb{C}} \to \mathrm{End}(\wedge W)$ is defined in (68.6).

The map $\tilde\rho$ and the space W depend on the choice of the basis $\{e_i\}$. But I am ready to bet that, if we consider the isomorphism $\phi\colon W \to \mathbb{C}^2$ given by $\phi(f_0) = (1, 0)$ and $\phi(f_1) = (0, 1)$ (and extended to $\wedge W$ in some way), the map

$$\begin{aligned}\tilde\rho_0\colon V^{\mathbb{C}} &\to \mathrm{End}(\wedge\mathbb{C}^2)\\ \tilde\rho_0(v) &= \phi\circ\tilde\rho(v)\circ\phi^{-1}\end{aligned} \tag{68.91}$$

does not depend on the choices[8]. Better : the map $\tilde\rho_0$ only depends on the signature of g.

In the case of a pseudo Riemannian manifold (M, g) and a field of orthonormal basis basis $\{e_\alpha\}$, we make the whole construction for each T_xM, g_x, using, at each $x \in M$ the orthonormal basis $\{(e_\alpha)_x\}_{\alpha=1,\dots,p+q}$.

Up to an abuse of notation writing $\tilde\rho$ instead of $\tilde\rho_0$, we say that the matrix of $\tilde\rho_x$ does not depend on x.

DEFooEGABooZPccEp

Definition 68.50.
*The **Dirac operator** $\mathcal{D}\colon \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$ acting on a spinor field is defined by*

$$(\mathcal{D}s)(x) = g_x(e_\alpha, e_\beta)\gamma_x^\beta(\widetilde{\nabla}_{e_\alpha}s)(x). \tag{68.92}$$ dirac

6. Extensions of e_α do not always globally exist, see remark 68.48.
7. By $\mathrm{SO}(g_x)$, we mean the set of all the matrix A such that $A^tg_xA = g$; A_x is an isometry of (T_xM, g_x). In other words, we consider A as a section of what we could call the "isometry bundle".
8. If you are able to prove that, please send me the proof. I am happy with a picture of a hand written proof.

PROPooUCECooOmdvmn

Proposition 68.51.
If the field of basis $e_\alpha \in \mathfrak{X}(M)$ is everywhere an orthonormal basis, the Dirac operator reads

$$(\mathcal{D}s)(x) = g_{\alpha\beta}\gamma^\alpha(\tilde{\nabla}_{e_\beta}s)(x) \tag{68.93}$$

where γ^α is a constant numeric matrix acting on ΛW.

Proof. The building of the Dirac operator begins by considering the vector space T_xM endowed with the metric g_x; then the spinor representation $\tilde{\rho}_x\colon T_xM \to \mathrm{End}(\Lambda W_x)$ where ΛW_x is build from isotropic vectors of T_xM is defined. If the vector fields $e_\alpha \in \mathfrak{X}(M)$ are everywhere orthonormal for the metric g, then we have the matrix equality [9]

EQooQMKEooWwYvus
$$\tilde{\rho}_x\big((e_\alpha)_x\big)_{ij} = \tilde{\rho}(v_\alpha)_{ij} \tag{68.94}$$

where the left hand side describe the matrix component of a linear operator acting on ΛW_x while in the right hand side we have the matrix component of a linear operator acting on ΛW and v_α is a basis on $\mathbb{R}^n$ with respect to which the metric is the same as the metric g_x in the basis $(e_\alpha)_x$. Let $\hat{\psi}\colon P \to \Lambda W$ be an equivariant function; from definition (68.87) of γ we have

$$\big(\gamma(e_\alpha\hat{\psi})\big)(\xi) = (a_\alpha\hat{\psi})(\xi)$$

where $a_\alpha(\xi) = \tilde{\rho}\Big(\tilde{e}_\alpha\big(\phi(\xi)\big)\Big)$. In this expression, $\tilde{e}_\alpha$ is the equivariant function associated with the vector field $e_\alpha \in \mathfrak{X}(M)$. It is defined in subsection 56.7.4 as

$$\begin{aligned}\tilde{e}_\alpha\colon \mathrm{SO}(M) &\to \mathbb{R}^m\\ b &\mapsto b^{-1}\big((e_\alpha)_{\pi(b)}\big).\end{aligned} \tag{68.95}$$

So we have $\hat{a}_\alpha\colon P \to \mathrm{End}(\Lambda W)$ defined by

$$\hat{a}_\alpha(\xi) = \tilde{\rho}\big(\varphi(\xi)^{-1}e_\alpha(x)\big)$$

with $x = \pi(\xi)$. Now if ξ is any element of $\pi^{-1}(x)$, we have

$$\big(\gamma(e_\alpha)\psi\big)(x) = (a_\alpha\psi)(X) = \big[\xi, \hat{a}_\alpha(\xi)\hat{\psi}(\xi)\big] = \big[\xi, \tilde{\rho}\big(\varphi(\xi)^{-1}e_\alpha(x)\big)\hat{\psi}(\xi)\big].$$

There exists a $g \in \mathrm{Spin}(p,q)$ such that $\varphi(\xi \cdot g) = \mathbb{1}$; taking this element and using equivariance of the latter expression,

$$\big(\gamma(e_\alpha)\psi\big)(x) = \big[\xi \cdot g, \tilde{\rho}\big(e_\alpha(x)\big)\hat{\psi}(\xi \cdot g)\big] = \big[\xi \cdot g, \gamma^\alpha\hat{\psi}(\xi)\big] = [\xi, \gamma^\alpha\hat{\psi}(\xi)]. \tag{68.96}$$

What we proved is that $\big(\gamma e_\alpha\psi\big)(x) = \gamma^\alpha\psi(x)$ is the sense that

$$\widehat{\gamma(e_\alpha)\psi} = \gamma^\alpha\hat{\psi}. \tag{68.97}$$

Hence the Dirac operator reads

$$(\mathcal{D}s)(x) = g_{\alpha\beta}\gamma^\alpha\big(\tilde{\nabla}_{e_\beta}s\big)(x)$$

in the sense that

$$\widehat{\mathcal{D}s} = g_{\alpha\beta}\gamma^\alpha\,\widetilde{\nabla_{e_\beta}s}. \tag{68.98}$$

□

9. See the more or less justification in 68.49.

An often more convenient way to write the Dirac operator is to consider an orthonormal basis (so that the metric g and the matrices γ are constant) and to consider the equivariant functions:

$$\widehat{\mathcal{D}\psi} = g_{\alpha\beta}\gamma^\alpha \widehat{\nabla_{e_\alpha}\psi}.$$

This formulation is typically used when one search for Dirac operator on Lie groups. In this case, we choose left invariant vector fields generated by an orthonormal basis of the Lie algebra. The resulting field of basis is everywhere Killing-orthonormal.

Acting on a function $f\colon M \to \mathbb{R}$, it is defined by $\mathcal{D}\colon C^\infty(M) \to C^\infty(M)$,

$$(\mathcal{D}f)(x) = g_x(e_\alpha, e_\beta)\gamma_x^\beta(e_{\alpha x} \cdot f). \tag{68.99}$$

eq_defDirac_f

With these definitions, one has

$$(\mathcal{D}(fs))(x) = (f\mathcal{D}s)(x) + (\mathcal{D}f)(x).$$

Indeed,

$$\begin{aligned}(\mathcal{D}(fs))(x) &= g_{\alpha\beta}\gamma_x^\beta(\widetilde{\nabla}_{e_\alpha} fs)(s)\\ &= g_{\alpha\beta}\Big((e_\alpha \cdot f)s(x) + f(x)(\widetilde{\nabla}_{e_\alpha}s)(x)\Big)\\ &= f(x)(\mathcal{D}s)(x) + g_{\alpha\beta}\gamma_x^\beta(e_{\alpha x}\cdot f)\\ &= (f\mathcal{D}s)(x) + (\mathcal{D}f)(x).\end{aligned} \tag{68.100}$$

With that definition, the Dirac operator becomes a derivation of the spinor bundle.

68.7 Example: Dirac operator on $\mathbb{R}^2$ with the euclidian metric

Pg_exempleRdeux

Since the frame bundle $B(M)$ is a principal bundle (see subsection 56.6.4), one can consider some associated bundles on it. We are now going to see that the one given by the definition representation $\rho\colon GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ on $\mathbb{R}^n$ is the tangent bundle. So we study $B(M) \times_\rho \mathbb{R}^n$. By choosing a basis on each point of M, we identify each T_xM to $\mathbb{R}^n$. An element of $B(M) \times \mathbb{R}^n$ is a pair (b, v) with $b = (\overline{b}_1, \dots, \overline{b}_n)$ and $v = (v^1, \dots, v^n)$. We can identify v to the element of T_xM given by $v = v^i\overline{b}_i$.

In order to build the associated bundle, we make the identifications

$$(b, v) \cdot g \sim (b \cdot g, g^{-1}v).$$

Here, by gv we mean the vector whose components are given by $(gv)^i = v^j g_j{}^i$. The tangent vector given by $(b \cdot g, g^{-1}v)$ is $(g^{-1}v)^i(b \cdot g)_i = v^j(g^{-1})_j{}^i g_i{}^k\overline{b}_k = v^k\overline{b}_k$ So the identification map $\psi\colon B(M) \times_\rho \mathbb{R}^n \to TM$ given by

$$\psi([b, v]) = v^i\overline{b}_i$$

is well defined.

The following step is to consider the following spin structure:

$$\mathrm{Spin}(2) \rightsquigarrow \mathbb{R}^2 \times \mathrm{SO}(2) \xrightarrow{\varphi} \mathrm{SO}(\mathbb{R}^2) \leftsquigarrow \mathrm{SO}(2)$$

We have to define the two actions and φ. One of the main result of section 68.3.5 is that $\chi\colon \mathrm{Spin}(2) = U(1) \to \mathrm{SO}(2)$ is surjective. So, we can define the action of $\mathrm{Spin}(2)$ on P by

$$(x, b) \cdot s = (x, \chi(s)^{-1}b).$$

On the other hand, an element A in $\mathrm{SO}(\mathbb{R}^2)$ can be written as $A = \{ae_i\}_x$ where e_i is the canonical basis of $T_x\mathbb{R}^2$, and a is a matrix of $\mathrm{SO}(2)$. See subsection 56.6.4. For $g \in \mathrm{SO}(2)$, we define

$$A \cdot g = \{g^{-1}ae_i\}_x. \tag{68.101}$$

r1504d2

and $\varphi\colon \mathbb{R}^2 \times \mathrm{SO}(2) \to \mathrm{SO}(\mathbb{R}^2)$ by

$$\varphi(x, b) = \{be_i\}_x.$$

The following shows that these definitions give a spin structure:

$$\varphi((x, b) \cdot s) = \varphi(x, \chi(s)^{-1}b) = \{\chi(s)^{-1}be_i\}_x = \{be_i\}_x \cdot \chi(s) = \varphi(x, b) \cdot \chi(s). \tag{68.102}$$

68.7.1 Connection on $\mathrm{SO}(\mathbb{R}^2)$

We are searching for a torsion-free connection on the simplest metric space: the euclidian $\mathbb{R}^2$. Thus we will try the simplest choice of horizontal space: we want an horizontal vector to be tangent to a curve of the form $X(t) = \{be_i\}_{x(t)}$. For this reason, we want to define the connection 1-form by $\omega(X) = b'(0)$. For technical reasons which will soon be apparent, we will not exactly proceed in this manner. For $X(t) = \{be_i\}_{x(t)}$, we define

$$\omega(X) = -(b(t)b(0)^{-1})'(0). \tag{68.103}$$

We of course have $\omega(X) = 0$ if and only if $b'(0) = 0$: this choice of ω follows our first idea. In order for ω to be a connection form, we have to verify the two conditions of definition 56.73.

Proposition 68.52.
The 1*-form defined by*

$$\omega(X) = -(b(t)b(0)^{-1})'(0)$$

for $X = \dfrac{d}{dt}\{b(t)e_i\}_{x(t)}\Big|_{t=0}$ *is a connection* 1*-form.*

Proof. Let $A \in \mathrm{SO}(2)$. If $u = \{be_i\}_x$, equation (68.101) gives:

$$A^*_u = \frac{d}{dt}\{e^{-tA}be_i\}_x\Big|_{t=0},$$

so that $\omega(A^*_u) = -(e^{-tA}bb^{-1})'(0) = A$. This checks the first condition. For the second, one remarks that the path in $\mathrm{SO}(\mathbb{R}^2)$ which defines the vector $R_{g*}X$ is $(R_{g*}X)(t) = \{g^{-1}b(t)e_i\}_x$. It follows that

$$\begin{aligned}
\omega(R_{g*}X) &= -(g^{-1}b(t)b(0)^{-1}g)'(0) \\
&= -\left(\mathbf{Ad}_{g^{-1}}(b(t)b(0)^{-1})\right)'(0) \\
&= -Ad_{g^{-1}}(b(t)b(0)^{-1})'(0) \\
&= Ad_{g^{-1}}\omega(X).
\end{aligned} \tag{68.104}$$

□

Proposition 68.53.
The covariant derivative induced on M *by this connection is*

$$\nabla_X Y = X(Y). \tag{68.105}$$

derrcovexplicite

Proof. In this demonstration, we will use the equivariant functions defined in 56.7.4. In order to compute $(\nabla_X Y)_x$, we have to use the definition of theorem 56.79. We first have to compute the horizontal lift of X. It is no difficult to see that $\overline{X}_{\{be_i\}_x}$ is given by the path

$$\overline{X}(t) = \{be_i\}_{X(t)}$$

if the vector field X is given by the path $X(t)$ in M. Indeed, it is trivial that $\omega(\overline{X}) = 0$, and

$$d\pi_*\overline{X} = \frac{d}{dt}\pi\{be_i\}_{X(t)}\Big|_{t=0} = \frac{d}{dt}X(t)\Big|_{t=0} = X.$$

Now, we compute $(\overline{X}\hat{s})(b)$ for $b = \{Se_i\}_x$. We begin using the basic definitions and notations:

$$(\overline{X}\hat{s})(b) = \overline{X}_b\hat{s} = \frac{d}{dt}\hat{s}(\overline{X}_b(t))\bigg|_{t=0} = \frac{d}{dt}\hat{s}(\{Se_i\}_{X(t)})\bigg|_{t=0}.$$

We can rewrite it with $\hat{Y}$ instead of $\hat{s}$. By construction (see (56.246)), if $b = \{Se_i\}_x$, $\hat{Y}(b) = S^{-1}(Y_x)$. Thus

$$(\overline{X}\hat{Y})(b) = \frac{d}{dt}S^{-1}(Y_{X(t)})\bigg|_{t=0},$$

where, if $\{\overline{1}_i\}$ is a basis of $\mathbb{R}^m$, then S is

$$\begin{aligned} S\colon \mathbb{R}^m &\to T_{X(t)}M \\ v^i\overline{1}_i &\mapsto S^i_j v^j(\partial_j)_{X(t)} \end{aligned} \tag{68.106}$$

So if we write $Y_x = Y^i(x)\partial_i$, we have

$$S^{-1}(Y_{X(t)}) = (S^{-1})^i_j Y^j(X(t))\overline{1}_i$$

and

$$\frac{d}{dt}S^{-1}(Y_{X(t)})\bigg|_{t=0} = (S^{-1})^i_j\,\frac{d}{dt}Y^j(X(t))\bigg|_{t=0}\overline{1}_i = (S^{-1})^i_j X(Y^j)\overline{1}_i.$$

Since b is an isomorphism, we can apply b on both side of $\hat{X}(b) = b^{-1}(X_x)$, and take $\nabla_X Y$ instead of X:

$$(\nabla_X Y)(x) = b\big((S^{-1})^i_j X(Y^j)\overline{1}_i\big) = S^k_i(S^{-1})^i_j X(Y^j)(\partial_k)_x = X(Y^j)(\partial_j)_x = X(Y)_x. \tag{68.107}$$

□

From this and definition 56.257, we immediately conclude that our connection is torsion-free. In a certain manner, one can say that our covariant derivative is the usual one.

68.7.2 Construction of γ

Now, we construct the map γ of subsection 68.6. The first step is to define $\hat{a}_X\colon P \to \mathrm{End}\,(\Lambda W)$ by

$$\hat{a}_X(p) = \tilde{\rho}(\hat{X}_{\varphi(p)}).$$

Here, ΛW is the completely isotropic subspace of $(\mathbb{R}^2)^{\mathbb{C}}$ with euclidian metric; thus we can use the result of section 68.3.5. In particular, we know the representation $\tilde{\rho}$.

To see it more explicitly, we need the expression of $\hat{X}$. It is given in subsection 56.7.4: if b is the basis $\{be_i\}_x$, $\hat{Y}(b) = b^{-1}(Y_x)$. As $\varphi(b,x) = \{be_i\}_x$, we have

$$\hat{a}_X(b,x) = \tilde{\rho}(b^{-1}(X_x)).$$

The subsection 56.7.2.2 explains how to explicitly get $\gamma(X)$ with the definition $\gamma(X) = a_X$. If ψ is a section of $\mathcal{S}$ and $\psi(x) = [\xi, v]$, the general definition gives us $(a_X\psi)(x) = [\xi, \hat{a}_X(\xi)v]$ and in our particular case, if $\xi = (b,x)$, we get:

$$(\gamma(X)\psi)(x) = [\xi, \tilde{\rho}(b^{-1}(X_x))v]. \tag{68.108}$$

68.7.3 Covariant derivative on $\Gamma(\mathcal{S})$

Remember the spin structure of $\mathrm{SO}(\mathbb{R}^2)$: $\varphi(x,S)=\{Se_i\}_x$. We now construct the connection on $P=\mathbb{R}^2\times\mathrm{SO}(2)$. It is defined by the 1-form $\tilde\omega=\varphi^*\omega$. If v is a vector of P, it is described by a path $v(t)=(x(t),b(t))$, then the path of $d\varphi(v)$ is $\{b(t)e_i\}_{x(t)}$ and $\tilde\omega(v)=\omega(d\varphi(v))=-(b(t)b(0)^{-1})'(0)$.

The next step defining the Dirac operator is to find out an explicit form for the map $\widetilde\nabla\colon\mathfrak{X}(M)\times\Gamma(\mathcal{S})\to\Gamma(\mathcal{S})$. A section $s\in\Gamma(\mathcal{S})$ is a map $s\colon M\to\mathcal{S}=(\mathbb{R}^2\times\mathrm{SO}(2))\times_\rho\Lambda W$; it is defined by an equivariant function $\hat s\colon P\to\Lambda W$. In order to find the value of $(\widetilde\nabla_X s)(x)$ for $X\in\mathfrak{X}(M)$, we use the definition

$$\widehat{\widetilde\nabla_X s}(\xi)=\overline{X}_\xi(\hat s)$$

where $\overline{X}$ is the horizontal lift in the sense of $\tilde\omega$. For the same reason as in the proof of proposition 68.105, $\overline{X}_{(b,x)}$ is given by the path $\overline{X}(t)=(b,X(t))$ where $X(t)$ is the path which defines X. So we have

$$\widehat{\widetilde\nabla_X s}(\xi)=\overline{X}_{(b,x)}(\hat s)=\frac{d}{dt}\hat s(b,X(t))\bigg|_{t=0}.$$

Remark that ΛW is a vector space; so for every $\alpha\in\Lambda W$, the identification $T_\alpha\Lambda W=\Lambda W$ is correct.

Our first form of $\widetilde\nabla$ is

$$(\widetilde\nabla_X s)(x)=\Big[\xi,\frac{d}{dt}\hat s(b,X(t))\Big|_{t=0}\Big],$$

but we can modify this in order to get simpler expressions. Remark that we have an equivalence class, so that we can always choose the element of the class such that $\xi=(\mathbb{1},x)$. We define $\overline{s}\colon\mathbb{R}^2\to\Lambda W$, $\overline{s}(v)=\hat s(\mathbb{1},v)$. Our second and final form for $\widetilde\nabla$ is:

$$(\widetilde\nabla_X s)(x)=\Big[(\mathbb{1},x),\frac{d}{dt}\overline{s}(X(t))\Big|_{t=0}\Big] \tag{68.109a}$$

$$=[(\mathbb{1},x),X(\overline{s})], \tag{68.109b}$$

where $X(\overline{s})$ is well defined because $\overline{s}$ is a map from $\mathbb{R}^2$ into a vector space (namely: ΛW).

68.7.4 Dirac operator on the euclidian $\mathbb{R}^2$

We continue to write explicitly the definition (68.92). Putting together (68.108) and (68.109b), one finds

$$\gamma^\alpha_x(\widetilde\nabla_{e_\beta}s)(x)=\gamma(e_{\alpha x})[\xi,e_\beta(\overline{s})]=[\xi,\tilde\rho(b^{-1}(e_{\alpha x}))e_\beta(\overline{s})]. \tag{68.110}$$

Here, $e_\beta=\partial_\beta$ and $b=\mathbb{1}$, then

$$\gamma^\alpha_x(\widetilde\nabla_{e_\beta}s)(x)=[(\mathbb{1},x),\tilde\rho(e_\alpha)\partial_\beta\overline{s}].$$

Now, the Dirac operator reads

$$(\mathcal{D}s)(x)=[(\mathbb{1},x),\gamma^\alpha\partial_\alpha\overline{s}].$$

We can obtain a more compact expression by defining "Ys" and "As" when $s\in\Gamma(\mathcal{S})$, $Y\in\mathfrak{X}(\mathbb{R}^2)$ and $A\in\operatorname{End}\Lambda W$. The definitions are

$$(Ys)(x)=[(\mathbb{1},x),(Y\overline{s})(x)],$$
$$(As)(x)=[(\mathbb{1},x),A\overline{s}(x)].$$

With these conventions, one writes:

$$(\mathcal{D}s)(x)=\gamma^\alpha(\partial_\alpha s)(x).$$

This justifies the expression (68.3): $\mathcal{D}=\gamma^\alpha\partial_\alpha$ on flat spaces. With a good choice of basis of ΛW, the matrices γ^α are given by (68.56), and

$$\gamma^\alpha\partial_\alpha=\begin{pmatrix}0&-1\\1&0\end{pmatrix}\partial_x-\begin{pmatrix}0&i\\i&0\end{pmatrix}\partial_y.$$

If we identify $\mathbb{R}^2$ with $\mathbb{C}$ we have the following definitions:

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \qquad \partial_{\overline{z}} = \frac{1}{2}(\partial_x + i\partial_y),$$

so that

$$\mathcal{D} = \begin{pmatrix} 0 & -\partial_{\overline{z}} \\ \partial_z & 0 \end{pmatrix}.$$

68.7.5 Dirac operator as elliptic pseudo-differential operator

subSecREctBOh

Let (M, g) be a spin manifold and D, its Dirac operator which is locally written under the form $D = \gamma(dx^\mu)\partial_\mu$. So $A_\mu(x) = \gamma(dx^\mu)$, and the principal symbol is

$$\xi^\mu A_\mu(x) = \gamma(\xi).$$

Let us point out that $\gamma(\xi)$ is not a real number, but an endomorphism of the spinor bundle. Using relation (68.23), we find that

$$\gamma(\xi)^2 = -\|\xi\|^2 \,\mathrm{Id},$$

which is invertible when $\xi \neq 0$. We conclude that Dirac is an elliptic operator [10].

68.8 Clifford algebras and Morita equivalence

Let $\mathfrak{A}$ be an algebra. An algebra $\mathcal{B}$ is said to be **Morita equivalent** to $\mathfrak{A}$ if $\mathcal{B} = \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ for some finite projective module $\mathcal{E}$ over $\mathfrak{A}$. The algebra $\mathfrak{A}$ is Morita equivalent to itself taking the trivial module $\mathcal{E} = \mathfrak{A}$.

PgMoritaEq

We consider a manifold M of dimension $n = 2m$.

¡¡ Avertissement/question à la lectrice !! 68.54
The two following statements are imprecise.

Proposition 68.55.
A module which implement a Morita equivalence between two C^*-algebras is finite projective.

Theorem 68.56 (Serre-Swan)**.**
If one of the two Morita equivalent is the continuous function space over a manifold $\mathfrak{A} = C(M)$, then the module which gives the Morita equivalence is the section of continuous sections of a vector bundle over M, $\mathcal{E} = \Gamma(E)$.

Furthermore, if $\mathfrak{A} = C(M)$ and $\mathcal{B} = \Gamma(\mathrm{Cl}(M))$, we have $\mathrm{End}\, E \simeq \mathrm{Cl}(M)$ as isomorphism of vector bundle. Since $\mathrm{Cl}\, M$ is of rank 2^n, $\mathrm{End}\, E$ has same rank and E_x has dimension $\sqrt{2^n} = 2^m$. So it is possible to choose the Clifford action in such a way that $\Gamma(E)$ is an irreducible Clifford module.

We often look at an anti-linear map $J\colon \Gamma(E) \to \Gamma(E)$ such that for all $\psi \in \Gamma(E)$

(1) $J(\psi f) = (J\psi)\overline{f}$ for all $f \in C(M)$,

(2) $J(a\psi) = \epsilon(a) a J\psi$ for all $a \in \Gamma^\infty(\mathrm{Cl}\, M)$.

How to define $a\psi$? We consider $\mathfrak{A} = C(M)$, $\mathcal{B} = \Gamma(\mathrm{Cl}\, M)$ and we define $\Gamma(E)$ is such a way that it implements a Morita equivalence between $\mathfrak{A}$ and $\mathcal{B}$; hence $\Gamma(E)$ is a $C(M)$-module. From dimensional considerations, we can define on $\Gamma(E)$ a Clifford module structure, i.e. a $C(M)$-linear

$$c\colon \Gamma(\mathrm{Cl}\, M) \to \mathrm{End}(\Gamma E), \tag{68.111}$$

hence $a\psi$ makes sense for any $a \in \Gamma^\infty(\mathrm{Cl}\, M)$ and $\psi \in \Gamma(E)$ with definition

$$(a\psi)(x) = \big(c(a)\psi\big)(x) = c(a(x))\psi(x) \tag{68.112}$$

10. Definition 62.51.

Theorem 68.57.
Let (M, S, J) be a spin manifold of dimension n. There exists an unique connection

$$\nabla^S \colon \Gamma^\infty(S) \to \Gamma^\infty(S) \otimes \Omega^1(S)$$

such that

(1) $(\nabla^S \psi | \phi) + (\psi | \nabla^S \phi) = d(\psi|\phi)$,

(2) $[\nabla^S, J] = 0$,

(3) $\nabla^S\big(c(a)\psi\big) = c\big(\nabla a\big)\psi + c(a)\nabla^S\psi$ *for all* $a \in \mathrm{Cl}(M)$ *and* $\psi \in \Gamma^\infty(S)$.

In the latter, the action of $\Gamma^\infty(\mathrm{Cl}\, M)$ *on* $\Gamma^\infty(S)$ *is induced from the action* $c\colon \mathrm{Cl}(T_x^* M) \to \mathrm{End}\, S$. *The* ∇ *which acts on* a *is the connection extended to* $\Gamma^\infty(\mathrm{Cl}\, M)$ *by virtue of Leibniz rule* $\nabla(uv) = \nabla(u)v + u\nabla(c)$.

Proof. No proof □

In this setting, we define

$$\begin{aligned} \hat{c}\colon \Gamma^\infty(S) \otimes \Gamma^\infty(\mathrm{Cl}\, M) &\to \Gamma^\infty(S) \\ \psi \otimes a &\mapsto c(a)\psi. \end{aligned} \tag{68.113}$$

Then we define the **Dirac operator** $\mathcal{D}\colon \Gamma^\infty(S) \to \Gamma^\infty(S)$,

$$\mathcal{D} = -i(\hat{c} \circ \nabla^S). \tag{68.114}$$

68.8.1 Example: quantum field theory

Let us show how does this operator gives back the usual Dirac operator of quantum field theory. Let M be a manifold and with two local basis $\{\partial_u\}$ and $\{\partial_\alpha\}$ of $T_x M$. The first one is the "natural" basis: $g(\partial_u, \partial_v) = g_{uv}$ has no particular properties while the second one is orthonormal $g(\partial_\alpha, \partial_\beta) = \delta_{\alpha\beta}$. The first dual basis is defined by $dx^\alpha \partial_\beta = \delta^\alpha_\beta$.

We write $\partial_\alpha = e^u_\alpha \partial_u$ and for the dual basis, $dx^\alpha = e^\alpha_u dx^u$. In order these definition to be coherent, we impose $dx^\alpha \partial_\beta = \delta^\alpha_\beta$:

$$dx^\alpha \partial_\beta = e^\alpha_u dx^u\big(e^v_\beta \partial_v\big) = e^\alpha_u e^v_\beta \delta^u_v = e^\alpha_u e^u_\beta. \tag{68.115}$$

We conclude that the **vielbein** (e^α_u) is the inverse of (e^v_β): $e^\alpha_u e^u_\beta = \delta^\alpha_\beta$. The vielbein are eventually complexes.

68.8.2 An other definition of the Dirac operator

Let us consider an orthonormal basis $\{e_a\}$ of M, i.e. on each $x \in M$,

$$g_x(e_a(x), e_b(x)) = \eta_{ab}.$$

This basis is related to a "natural" basis $\{\partial_\mu\}$ by

$$e_a = e^\mu_a \partial_\mu \tag{68.116}$$

where e^μ_a is called **vielbein** (here, they are more precisely n-beins). As far as metric is concerned we have

$$g^{\mu\nu} = e^\mu_a e^\nu_b \eta_{ab} \tag{68.117a}$$

$$\eta_{ab} = e^\mu_a e^\nu_b g_{\mu\nu}. \tag{68.117b}$$

If ∇ is the covariant derivative associated with g, we define the coefficients $\omega^b_{\mu a}$ by

$$\nabla_\mu e_a = \omega^b_{\mu a} e_b. \tag{68.118}$$

On the other hand, ∇ is related to the Christoffel symbols by

$$\nabla_\mu \partial_\nu = \Gamma^\sigma_{\mu\nu} \partial_\sigma. \tag{68.119}$$

Let $\mathrm{Cl}(M)$ be the Clifford module whose fibre is the Clifford complex algebra $\mathrm{Cl}(T^*_x M)^{\mathbb{C}}$. We consider $\Gamma(\mathrm{Cl}(M))$, the module of corresponding sections. It gives an algebra morphism

$$\begin{aligned} \gamma \colon \Gamma(\mathrm{Cl}(M)) &\to \mathfrak{B}(\mathscr{H}) \\ dx^\mu &\mapsto \gamma^\mu(x) = \gamma^a e^\mu_a \end{aligned} \tag{68.120}$$

which can be extended to the whole Clifford algebra. One can choose matrices $\gamma^\mu(x)$ and γ^a to be hermitian; they satisfy

$$\gamma^\mu(x)\gamma^\nu(x) + \gamma^\nu(x)\gamma^\mu(x) = -2g(dx^\mu, dx^\nu) = -2g^{\mu\nu} \tag{68.121a}$$

$$\gamma^a\gamma^b + \gamma^b\gamma^a = -2\eta^{ab}. \tag{68.121b}$$

All this allow us to lift the Levi-Civita connection from the tangent bundle to the spinor bundle by defining

$$\nabla^S_\mu = \partial_\mu + \omega^S_\mu = \partial_\mu + \frac{1}{2}\omega_{\mu ab}\gamma^a\gamma^b. \tag{68.122}$$

The **Dirac operator** is then given by

$$\mathcal{D} = \gamma \circ \nabla$$

and can locally be written under the form

$$\mathcal{D} = \gamma^\mu(x)(\partial_\mu + \omega^S_\mu) = \gamma^a e^\mu_a (\partial_\mu + \omega^S_\mu). \tag{68.123}$$

eq_Dirac_deux

Chapter 69

Relativistic fields and group theory

69.1 Mathematical framework of field theory

This is a short review; the aim is to see why the quantum theory of fields needs representations of the Poincaré group. It will be mostly physics oriented. References dealing with field theory including gauge theory and representations are [900, 977, 915, 978, 852, 974, 883].

69.1.1 Axioms of the (quantum) relativistic field theory

The quantum mechanics is based on a few number of axioms: pg:axiomes

(1) We have a Hilbert space $\mathscr{H}$. A physical state is given by a **ray** in $\mathscr{H}$, i.e. a set

$$\mathcal{R} = \{\xi\psi : |\xi| = 1\}$$

for a certain $\psi \in \mathscr{H}$ with $\langle\psi|\psi\rangle = 1$. In other words, the set of physical sates is the quotient of the set of unital vectors in $\mathscr{H}$ by the relation $\psi \sim \psi'$ if and only if $\psi = \xi\psi'$ for some unimodular complex number ξ. We denote by $\operatorname{Ray}\mathscr{H}$ the set of all rays in $\mathscr{H}$. ax:vaps

(2) The observables are represented by hermitian linear operators on $\mathscr{H}$. A state $\mathcal{R}$ has value α for the observable A if $A\mathcal{R} = \alpha\mathcal{R}$, where the action of A on the ray is obvious (and well defined because A is linear).

(3) If one has a system described by a state $\mathcal{R}$, and if one want to measure if it is in one of the state $\mathcal{R}_1, \dots, \mathcal{R}_n$ (orthogonal rays), the answer will be $\mathcal{R}_i$ with probability

$$P(\mathcal{R} \to \mathcal{R}_i) = |\langle\mathcal{R}|\mathcal{R}_i\rangle|^2.$$

If the $\mathcal{R}_n$ form a complete system, one has a theorem which states that

$$\sum_i P(\mathcal{R} \to \mathcal{R}_i) = 1.$$

ax:reprez

(4) The rays of $\mathscr{H}$ furnish a representation of the (identity component of) Poincaré group.

This last point can look strange; we will see later (page 3710) how it comes. It is the expression of a relativistic theory. That axiom is the reason why one make intensive use of representation theory in relativistic (quantum) field theory ... or maybe the intensive use of representation theory is the reason of that axiom. However, we will make an intensive use of representation theory developed in chapter 55.

69.1.2 Symmetries and Wigner's theorem

Consider the following situation: someone observes a system in a state $\mathcal{R}$, and makes measures $P(\mathcal{R} \to \mathcal{R}_i)$. An other person observes the same system which is, for him, in a state $\mathcal{R}'$ and observes $P(\mathcal{R}' \to \mathcal{R}'_i)$.

If two observers are related by a transformation of the Hilbert state which induces $\mathcal{R} \to \mathcal{R}'$ and $\mathcal{R}_i \to \mathcal{R}'_i$, there are said **equivalent** if

$$P(\mathcal{R} \to \mathcal{R}_i) = P(\mathcal{R}' \to \mathcal{R}'_i). \tag{69.1}$$

eq:sym_isom

Let us say it more precisely from a mathematical point of view. A **symmetry** is an invertible operator $T\colon \operatorname{Ray}\mathscr{H} \to \operatorname{Ray}\mathscr{H}$ such that for any $\phi_i \in \mathcal{R}_i$, $\phi'_i \in T\mathcal{R}_i$ and $\phi''_i \in T^{-1}\mathcal{R}_i$,

$$|\langle\phi'_1|\phi'_2\rangle|^2 = |\langle\phi_1|\phi_2\rangle|^2 = |\langle\phi''_1|\phi''_2\rangle|^2 \tag{69.2}$$

eq:

Remarque 69.1.
Here, neither $\mathcal{R}$ nor $\mathcal{R}'$ are measurable: the P's only are measurable.

The following can be found in [974] p.91, [900] p.354.

tho:Wigner

Theorem 69.2 (Wigner)**.**
Any symmetry T is induced by an operator U on $\mathscr{H}$ such that $\psi \in \mathcal{R}$ implies $U\psi \in \mathcal{R}'$. This operator is either unitary and linear, either anti-unitary and antilinear.

So, the symmetry operator must satisfy

$$\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle \tag{69.3a}$$
$$U(\xi\psi + \eta\psi) = \xi U\psi + \eta U\phi, \tag{69.3b}$$

or

$$\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle^* \tag{69.4a}$$
$$U(\xi\psi + \eta\psi) = \xi^* U\psi + \eta^* U\phi. \tag{69.4b}$$

In the anti-linear case operator, we do not define $U^\dagger$ by $\langle\phi|U^\dagger\psi\rangle = \langle U\phi|\psi\rangle$ because the left-hand side should be anti-linear with respect to ψ while the right-had should be linear. In place, for an antilinear operator A, we define $A^\dagger$ by

$$\langle\phi|A^\dagger\psi\rangle = \langle A\phi|\psi\rangle^* = \langle\psi|A\phi\rangle. \tag{69.5}$$

In this way, the definitions of unitary and anti-unitary in term of dagger are the same: $U^\dagger = U^{-1}$.

For any transformation $T\colon \operatorname{Ray}\mathscr{H} \to \operatorname{Ray}\mathscr{H}$, the Wigner's theorem provides an operator $U(T)\colon \mathscr{H} \to \mathscr{H}$ which induces T on Ray. If the operator T depends on a parameter θ, the operator $U(T(\theta))$ depends on θ. If T depends continuously on the parameter then the family $U(T(\theta))$ only contains unitary/linear operators or only antiunitary/antilinear operators.

In physical cases, $T(\theta)$ is mostly a Poincaré transformation: $\theta = (\Lambda, p)$. But $T(\mathbb{1}, 0)$ is the identity which is represented by $U(\mathbb{1}, 0) = \mathbb{1}$. Then all the (connected to identity) Poincaré transformations are represented by linear and unitary operators on $\mathscr{H}$.

We will follow the proof given in [974]. An other form of the proof can be found in [900]. The latter use a slightly different formalism in the axioms of the quantum mechanics; this is explained in appendix 88.1. It is now time to prove the theorem.

Proof of Wigner's theorem. We consider an orthonormal basis $\{\psi_k\}$ of $\mathscr{H}$ with $\psi_k \in \mathcal{R}_k$, and a choice of $\psi'_k \in T\mathcal{R}_k$. From this and the assumptions, we have

$$|\langle\psi'_k|\psi'_l\rangle|^2 = |\langle\psi_k|\psi_l\rangle|^2 = \delta_{kl}.$$

Then $\langle\psi'_k|\psi'_k\rangle = 0$ whenever $k \neq l$ and, since $\langle\psi'_k|\psi'_k\rangle$ is real and positive, $\langle\psi'_k|\psi'_k\rangle = 1$. So $\langle\psi'_k|\psi'_l\rangle = \delta_{kl}$.

The set ψ'_k is also complete in $\mathscr{H}$. Indeed suppose that we have a vector $\psi' \in \mathscr{H}$ such that $\langle\psi'|\psi'_k\rangle = 0$ for all k. If $\psi' \in \mathcal{R}$, we consider a $\psi'' \in T^{-1}\mathcal{R}$ and we have

$$|\langle\psi''|\psi_k\rangle|^2 = |\langle\psi'|\psi'_k\rangle|^2 = 0,$$

which contradicts the fact that the ψ_k's form a complete set. Now we have to fix a phase convention for the ψ_k. Since there are no canonical choice of phase, we fix with respect to an arbitrary one of the ψ_k, say ψ_1. We put

$$\gamma_k = \frac{1}{\sqrt{2}}(\psi_1 + \psi_k) \in \mathcal{C}_k \tag{69.6}$$

for $k \neq 1$. Any $\gamma'_k \in T\mathcal{C}_k$ can be written in the basis $\{\psi'_k\}$:

$$\gamma'_k = \sum_l c_{kl}\psi'_l. \tag{69.7}$$

From assumption (69.1) and the fact that $|c_{kl}|^2 = |\langle\gamma'_k|\psi'_l\rangle|^2$, we find, for $k, l \neq 1$

$$|c_{kl}|^2 = \frac{1}{2}\delta_{kl}.$$

We can choose the phase of γ'_k and ψ'_k in order to get $c_{kk} = c_{k1} = 1/\sqrt{2}$. For this, we begin to fix γ'_k in such a manner to get $c_{k1} = 1/\sqrt{2}$ (from $|c_{k1}| = |\langle\gamma'_k|\psi'_1\rangle|$), and next we fix ψ'_k for the c_{kk}. From now on, the so chosen γ'_k and ψ'_k are denoted by $U\gamma_k$ and $U\psi_k$.

What we did until now is to take a basis $\{\psi_k\}$ of $\mathscr{H}$ and define $\gamma_k = 1/\sqrt{2}(\psi_1 + \psi_k)$. Next we had chosen the phases of $\psi'_k \in T\mathcal{R}_k$ and $\gamma'_k \in T\mathcal{C}_k$ in order to have

$$\begin{aligned} c_{kk} = c_{k1} = 1/\sqrt{2} \quad & \forall k, \\ c_{kl} = 0 \qquad & \text{if } l \neq k, l \neq 1. \end{aligned} \tag{69.8}$$

This allows us to check a certain linearity for the operator U:

$$\begin{aligned} U\left(\frac{1}{\sqrt{2}}(\psi_k + \psi 1)\right) &= U\gamma_k \\ &= \gamma'_k \\ &= \frac{1}{\sqrt{2}}\psi'_1 + \frac{1}{\sqrt{2}}\psi'_k \qquad \text{from (69.7) and (69.8)} \\ &= \frac{1}{\sqrt{2}}\big(U\psi_1 + U\psi_k\big). \end{aligned} \tag{69.9}$$

Now we have to build U on a general vector $\psi = \sum_k \psi_k \in \mathcal{R}$. Any vector $\psi' \in T\mathcal{R}$ can be decomposed with respect to the basis $\{\psi'_k = U\psi_k\}$:

$$\psi' = \sum_k C'_k U\psi_k. \tag{69.10}$$

From the conservation of probability $|\langle\psi_k|\psi\rangle|^2 = |\langle U\psi_k|\psi'\rangle|^2$ and $|\langle\gamma_k|\psi\rangle|^2 = |\langle U\gamma_k|\psi'\rangle|^2$, we find

$$|C_k|^2 = |C'_k|^2, \tag{69.11a}$$

$$|C_k + C_1|^2 = |C'_k + C'_1|^2. \tag{69.11b}$$

If one writes $C_k = a_k + ib_k$, one finds $\mathrm{Re}(C_k/C_1) = (a_k a_1 + b_k b_1)/|C_1|^2$. By doing the same with C'_k and using (69.11),

$$\mathrm{Re}(C_k/C_1) = \mathrm{Re}(C'_k/C'_1). \tag{69.12}$$

Equation (69.11a) also imposes

$$|C_k/C_1|^2 = |C'_k/C'_1|^2, \tag{69.13}$$

while compatibility between (69.13) and (69.12) requires

$$\mathrm{Im}(C_k/C_1) = \pm\,\mathrm{Im}(C'_k/C'_1). \tag{69.14}$$

Equations (69.12) and (69.14) show that C_k and C'_k must satisfy

$$C_k/C_1 = C'_k/C'_1 \tag{69.15a}$$

xor

$$C_k/C_1 = (C'_k/C'_1)^*. \tag{69.15b}$$

For a given ψ we have to show that the choice must be the same for all the C_k [1]. Let $l \neq k$ and suppose that $C_k/C_1 = C'_k/C'_1$ and $C_l/C_1 = (C'_l/C'_1)^*$; we will show that in this case, one of the two ratios is real. So we can suppose $k \neq 1 \neq l$. We consider the vector $\Phi = \frac{1}{\sqrt{3}}(\psi_1 + \psi_k + \psi_l)$,

$$\Phi' = \frac{\alpha}{\sqrt{3}}\big(U\psi_1 + U\psi_k + U\psi_l\big)$$

where $\alpha \in \mathbb{C}$ satisfies $|\alpha| = 1$. The conservation of probability $|\langle\Phi|\psi\rangle|^2 = |\langle\Phi'|\psi'\rangle|^2$ gives $|C_1 + C_k + C_l|^2 = |C'_1 + C'_k + C'_l|^2$. Since $|C_1|^2 = |C'_1|^2$, we can divide the left hand side by $|C_1|^2$ and the right one by $|C'_1|^2$. We find

$$\left|1 + \frac{C_k}{C_1} + \frac{C_l}{C_1}\right|^2 = \left|1 + \frac{C'_k}{C'_1} + \frac{C'_l}{C'_1}\right|^2.$$

Using the assumption $C_k/C_1 = C'_k/C'_1$ and $C_l/C_1 = (C'_l/C'_1)^*$, we are in a case of an equation of the form $|u + v|^2 = |u + v^*|^2$ with u, $v \in \mathbb{C}$. If we write $u = a + bi$ and $v = x + iy$, we find $b + y = \pm(b - y)$, so that it leaves the choice $y = 0$ or $b = 0$ which corresponds to $(C_k/C_1) \in \mathbb{R}$ or $(C_l/C_1) \in \mathbb{R}$. So the coefficients C'_k $(k \neq 1)$ in the expansion (69.10) must satisfy

$$C_k/C_1 = C'_k/C'_1 \quad \forall k \tag{69.16a}$$

xor

$$C_k/C_1 = (C'_k/C'_1)^* \quad \forall k. \tag{69.16b}$$

Note that the phase of C_1 is not yet fixed. We naturally choice $C_1 = C'_1$ or $C_1 = C'^*_1$ following the case. We define $U : \mathscr{H} \to \mathscr{H}$ by

$$U\left(\sum_k C_k\psi_k\right) = \sum_k C_k U\psi_k \qquad \text{if (69.16a)}, \tag{69.17a}$$

xor

$$U\left(\sum_k C_k\psi_k\right) = \sum_k C^*_k U\psi_k \qquad \text{if (69.16b)}. \tag{69.17b}$$

This preserves the probability because $|\langle\psi|\psi_k\rangle|^2 = |C_k|^2$ while $|\langle U\psi|U\psi_k\rangle|$ is equal to $|C_k|^2$ or $|C^*_k|^2$ (which are the same) following the case (69.17a) or (69.17b).

Now we have to prove that the choice (69.17a) or (69.17b) is fixed by the data of T and must be the same for all the $\psi \in \mathscr{H}$. For, let us consider two vectors $\phi = \sum A_k\psi_k$, $\varphi = \sum B_k\psi_k$ and suppose that

$$U\phi = \sum_k A_k U\psi_k \text{ but } U\varphi = \sum_k B^*_k U\psi_k.$$

In order to see that it is impossible, looks at the conservation of probability $|\sum_k A_k B^*_k|^2 = |\sum_k A_k B_k|^2$, then

$$\sum_{kl}\left(B^*_l B_k A_l A^*_k - B^*_l B_k A^*_l A_k\right) = \sum_{kl} B^*_l B_k \operatorname{Im}(A_l A^*_k) = 0. \tag{69.18}$$

Since $A_l A^* l \in \mathbb{R}$, we can regroup each term (k, l) with the corresponding term (l, k). We get

$$0 = \sum_{kl}\operatorname{Im}(A_l A^*_k)(B^*_l B_k - B^*_k B_l) = \sum_{kl}\operatorname{Im}(A^*_k A_l)\operatorname{Im}(B^*_k B_l). \tag{69.19}$$

1. We will show later that for a given T, the choice must be the same for all the ψ.

We can find a vector $\sum_k C_k\psi_k$ such that eq:choix_C

$$\sum_{kl} \mathrm{Im}(C_k^* C_l)\,\mathrm{Im}(A_k^* A_l) \neq 0 \tag{69.20a}$$

eq:choix_C_a

and

$$\sum_{kl} \mathrm{Im}(C_k^* C_l)\,\mathrm{Im}(B_k^* B_l) \neq 0. \tag{69.20b}$$

In order to see how to find such a vector, let us show that there always exists a choice (i,j) such that $B_i^* B_j$ is not real. Let us say $B_1 = x + iy$ and $B_k = a_k + bi$. If $y \neq 0$, the condition $\mathrm{Im}(B_1^* B_k) = 0$ gives $B_k = \frac{b_k}{y} B_1$. It is always possible to find a sequence (b_k) which gives 1 as norm for $\sum B_k\psi_k$; the problem is not there. The problem is that $B_k/B_1 \in \mathbb{R}$, so that the choice (69.17) is not a true choice. For the same reason, all the $B_i^* B_k$ can't be pure imaginary.

Now we can find the vector which satisfy (69.20). There are several cases. If there is a pair (k,l) such that $A_k^* A_l$ and $B_k^* B_l$ are both complex, we can take all C_i's zero for $k \neq i \neq l$ and choose C_k and C_l in such a way that $C_k^* C_l$ is not real. If there is a pair (k,l) with $A_k^* A_l$ complex and $B_k^* B_l$ real, we consider a pair (m,n) such that $B_m^* B_n$ is complex. If $A_m^* A_n$ is complex, we take all the C_i's zero except C_m and C_n such that $\mathrm{Im}(C_m^* C_n) \neq 0$. If $A_m^* A_n$ is real, we take all the C_i's zero except C_k, C_l, C_m, C_n which we choose in such a way that $\mathrm{Im}(C_m^* C_n) \neq 0$ and $\mathrm{Im}(C_k^* C_k) \neq 0$.

Equation (69.20a) makes that the same choice must be made for $\sum A_k\psi_k$ and $\sum C_k\psi_k$ (if it was not the case, we would have an equation of the form of (69.19)). For the same reason, the same choice must be made for $\sum B_k\psi_k$ and $\sum C_k\psi_k$. So we conclude that the data of T fixes the choice between (69.17a) and (69.17b) and that this choice must be the same for all the vectors of $\mathscr{H}$.

We have to show that the possibility (69.17a) makes U linear and unitary while the possibility (69.17b) makes U antilinear and antiunitary. For we consider $\psi = \sum_k A_k\psi_k$ and $\phi = \sum_k B_k\psi_k$. If (69.17a) works,

$$\begin{aligned} U(\alpha\psi + \beta\phi) &= U\Big(\sum_k (\alpha A_k + \beta B_k)\psi_k\Big) \\ &= \sum_k (\alpha A_k + \beta B_k) U\psi_k \\ &= \alpha U\psi + \beta U\phi, \end{aligned} \tag{69.21}$$

and

$$\langle U\psi | U\phi\rangle = \sum_{kl} A_k^* B_l \langle U\psi_k | U\psi_l\rangle = \sum_k A_k^* B_k, \tag{69.22}$$

so that $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle$. Thus in this case U is linear and unitary. In the case where (69.17a) works, the computations are almost the same:

$$\begin{aligned} U(\alpha\psi + \beta\phi) &= U\Big(\sum_k (\alpha A_k + \beta B_k)\psi_k\Big) \\ &= \sum_k (\alpha^* A_k^* + \beta^* B_k^*) U\psi_k \\ &= \alpha^* U\psi + \beta^* U\phi, \end{aligned} \tag{69.23}$$

and

$$\langle U\psi | U\phi\rangle = \sum_{kl} A_k B_l^* \langle U\psi_k | U\psi_l\rangle = \sum_k A_k B_k^*, \tag{69.24}$$

so that $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle^*$. In this case, U is antilinear and antiunitary.

□

69.1.3 Projective representations

If $T_1(\mathcal{R}_n) = \mathcal{R}'_n$ and $\psi_n \in \mathcal{R}_n$, then $U(T_1)\psi_n \in \mathcal{R}'_n$. If $T_2(\mathcal{R}') = \mathcal{R}''$, then $U(T_2)U(T_1)\psi_n \in \mathcal{R}''_n$. But $U(T_2T_1)\psi_n$ also belongs to $\mathcal{R}''_n$. Then there exists a $\phi_n(T_2, T_1) \in \mathbb{R}$ such that

$$U(T_2)U(T_1)\psi_n = e^{i\phi_n(T_2,T_1)}U(T_2T_1)\psi_n.$$

Note that for any $\psi \in \mathscr{H}$, there exists a $\lambda \in \mathbb{R}$ such that $\|\lambda\psi\| = 1$. Since a real can be sent out the $U(T)$'s, for *any* $\psi \in \mathscr{H}$, there exists a ϕ which only depends on $\psi/\|\psi\|$ such that

$$U(T_2)U(T_1)\psi = e^{i\phi(T_2,T_1)}U(T_2T_1)\psi \tag{69.25}$$

Proposition 69.3.
The ϕ doesn't depend at all on the ψ:

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2T_1). \tag{69.26}$$

Proof. Let us consider a ψ_A and a ψ_B which are not proportional each other. One has a $\phi_{AB}(T_2, T_1)$ such that

$$\begin{aligned} e^{i\phi_{AB}(T_2,T_1)}U(T_2T_1)(\psi_A + \psi_B) &= U(T_2)U(T_1)(\psi_A + \psi_B) \\ &= e^{i\phi_A(T_2,T_1)}U(T_2T_1)\psi_A \\ &\quad + e^{i\phi_B(T_2,T_1)}U(T_2T_1)\psi_B. \end{aligned} \tag{69.27}$$

Now, we apply $U(T_2T_1)^{-1}$ to both sides. If it is unitary, the $e^{i\phi}$ get out without problems; else is get out as $e^{-i\phi}$:

$$e^{\pm i\phi_{AB}}(\psi_A + \psi_B) = e^{\pm i\phi_A}\psi_A + e^{\pm i\phi_B}\psi_B. \tag{69.28}$$

Since ψ_A and ψ_B are linearly independent, the only solution is $e^{i\phi_{AB}} = e^{i\phi_A} = e^{i\phi_B}$.

□

Since the operators $U(T)$ must only fulfil

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2, T_1), \tag{69.29}$$ eq:projectif

these form a **projective representation** of the symmetry group on the physical Hilbert space $\mathscr{H}$.

Remarque 69.4.
In order to have some physical relevance, this demonstration supposes that a state $\psi_A + \psi_B$ exists in nature. If one can divide the particles in several "incompatibles" classes labeled by a, b such that $\psi_a + \psi_b$ doesn't exist, then equation (69.29) is false and one has to write

$$U(T_2)U(T_1)\psi_a = e^{i\phi_a(T_2,T_1)}U(T_2T_1)\psi_a$$

because we can't show that $\phi_a = \phi_b$ from the simple fact that $\psi_a + \psi_b$ doesn't exist!

For example, physicists think that there are no superposition of state of integer and semi-integer spin.

Remarque 69.5.
If the group satisfies some requirements, one can choose $\phi = 0$. From now we suppose that we are in this case: we work with "true" representations.

69.1.4 Representations and power expansions

Let G be an arc connected Lie group whose elements are denoted by $T(\theta)$ with θ, a continuous family of parameters (from a local chart). The multiplication law is given by a function $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$:

$$T(\theta')T(\theta) = T\big(f(\theta',\theta)\big). \tag{69.30}$$

If $\theta = 0$ is the coordinate of the identity,

$$f(0,\theta) = f(\theta,0) = \theta. \tag{69.31}$$

We suppose that G acts on the rays of a Hilbert space $\mathscr{H}$, so that there are represented on $\mathscr{H}$ by unitary operators $U\big(T(\theta)\big)$. We denote by W the group of transformations of $\mathscr{H}$; roughly speaking,

$$W = U(G).$$

Now, we are going to cheat a little. We know that there exists a normal neighbourhood of e in W. In simple words, the map $\exp\colon \mathcal{W} \to W$ is a diffeomorphism between the elements of $\mathcal{W}$ "close" to 0 and the ones of W close to e. By *close to*, we mean that the components of θ are small enough. If $\{it_a\}$ is a basis of $\mathcal{W}$, we define

$$U(T(\theta)) = e^{i\theta^a t_a}. \tag{69.32}$$

In other words, one considers the exponential map for a neighbourhood of identity.

The cheat is the fact that $U(T(\theta))$ is actually defined by Wigner's theorem from the data of the group G. So equation (69.32) should be seen as a requirement in the choice of the basis $\{t_a\}$.

Remarque 69.6.
The i in the exponential in (69.32) and in the definition of the basis $\{it_a\}$ is a convention in order the t_a's to be hermitian. Indeed, the Lie algebra of a group of unitary matrices is made of *anti*hermitian matrices.

With all that,

$$U(T(\theta)) = \mathbb{1} + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \dots \tag{69.33}$$

where t_{bc} is defined (among other requirements) to absorb the "intuitive" minus sign in the third term.

Now we are going to explore some consequences of equation (69.30). Equation (69.31) makes the expansion of f as

$$f^a(\theta',\theta) = \theta^a + \theta'^a + f^a_{bc}\theta'^b\theta^c + \dots \tag{69.34}$$

From expansions (69.34) and (69.33) of f and $U(T(\theta))$, "group structure" equation (69.30) gives (at order two):

$$t_{bc} = -t_b t_c - i f^a_{bc} t_a \tag{69.35}$$

and nothing for the first order. Then, providing that one knows the group structure (the f), one knows the second order of the representation from the first one. From equation (69.32), one finds the value of t_{ab}:

$$e^{i\theta^a t_a} = 1 + i\theta^a t_a + \frac{1}{2}(i)^2(\theta^a t_a)(\theta^b t_b),$$

up to constant coefficients, one can choose t_{ab} to be symmetric with respect to a and b:

$$t_{ab} = \frac{1}{2}(t_a t_b + t_b t_a).$$

Taking this convention and computing $t_{bc} - t_{cb}$ from (69.35), we find

$$[t_a, t_b] = i C^c_{ab} t_c \tag{69.36}$$

with $C^c_{ab} = f^c_{ab} - f^c_{ba}$.

On the other hand, one knows that if a group is abelian, its algebra is also abelian; we can see it here by considering that if G is abelian, $f(\theta,\theta') = f(\theta',\theta)$, then f^c_{ab} is symmetric and $[t_a, t_b] = 0$. We can say more about f Since the t_a commute, equations (69.30) and (69.32) make that

$$e^{if(\theta,\theta')^a t_a} = e^{i\theta^a t_a} e^{i\theta'^b t_b} = e^{i(\theta^a+\theta'^a)t_a}, \tag{69.37}$$

so that

$$f(\theta,\theta') = \theta + \theta'.$$

69.2 The symmetry group of nature

69.2.1 Spin and double covering

subsec:sym_nature

Some of literature carry an ambiguity in the choice of the right space-time symmetry group in the quantum field theory. A very good and deep discussion about the choice of the space-time symmetry group of nature is given in the book [915] which will be used here. An other enlightening review can be found in [895].

From a relativistic point of view, the group is the Poincaré group of all the maps $\mathbb{R}^4 \to \mathbb{R}^4$ which leaves invariant the quantity $s^2 = -t^2 + x^2 + y^2 + z^2$. At this point we can already make an important remark: the so defined quantity s is in fact *not* a relativistic invariant. Indeed if I follow a (spatially) closed path, I will measure $\Delta t \neq 0$ and $\Delta x = \Delta y = \Delta z = 0$ because in *my* frame, my displacement is zero. A guy who keeps at my starting point will measure (between the beginning and the end of my travel) $\Delta' t \neq 0$ and also $\Delta x = \Delta y = \Delta z = 0$. If $s = s'$, then $\Delta t = \Delta t'$.

So the relativistic invariance is only local: $ds^2 = ds'^2$, and as far as relativity is concerned, one can work with infinitesimal transformations only. In this case, the distinction between the *groups* $L^\uparrow_+$ and $\mathrm{SL}(2,\mathbb{C})$ is no relevant. Intuitively, we choose $L^\uparrow_+$ to be the space-time symmetry group. As we will see the difference will reveal to be crucial in relativistic field theory because $L^\uparrow_+$ has no half-integer spin representations.

This group naturally splits into two parts: the translations and the rotations (and boost). As far as I know, the translation part makes no difficulties. For the other one, there are some difficulties to find the *minimal* group of symmetry. First, one often want to separate the space-time inversions P and T from the remaining: the group then becomes the homogeneous orthochrone Lorentz group $L^\uparrow_+$. An other often presented group is $\ldots\,\mathrm{SL}(2,\mathbb{C})$. This is our choice here. The physical reason of this choice is all but immediate. As we will see during the following pages, an elementary particle is an irreducible representation of the symmetry group.

For massive particles, the relevant subgroup of $\mathrm{SL}(2,\mathbb{C})$ reveals to be $SU(2)$. If we had chosen the most intuitive $L^\uparrow_+$, we would have found $\mathrm{SO}(3)$. There is an important difference between $SU(2)$ and $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$: the first one admits representations of any integer and half-integer spin while the second only posses the integer spin representations (see 55.74).

Let us now be more precise about the relation between $L^\uparrow_+$ and $\mathrm{SL}(2,\mathbb{C})$. A know result is

$$L^\uparrow_+ = \frac{\mathrm{SL}(2,\mathbb{C})}{\mathbb{Z}_2}.$$

Let $\mathrm{Spin}\colon \mathrm{SL}(2,\mathbb{C}) \to L^\uparrow_+$ be the surjective homomorphism with kernel $\pm\mathbb{1}_{2\times 2}$ giving this relation. We will not give a complete proof, but we will explain how $\mathrm{SL}(2,\mathbb{C})$ acts by isometries on $\mathbb{R}^4$. First, we remark that there exists a bijection between $\mathbb{R}^4$ and the 2×2 complex hermitian matrices:

$$v = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}. \tag{69.38}$$

If $\lambda \in \mathrm{SL}(2,\mathbb{C})$, the matrix $\lambda v \lambda^\dagger$ is also hermitian and $\|v\|^2 = \det v$. Thus

$$\begin{aligned} \Lambda(\lambda)\colon \mathbb{R}^4 &\to \mathbb{R}^4 \\ v &\mapsto \lambda v \lambda^\dagger \end{aligned} \tag{69.39}$$

is a Lorentz transformation if and only if $|\det\lambda| = 1$. Moreover,

$$\Lambda(\lambda\lambda') = \Lambda(\lambda)\Lambda(\lambda').$$

If $\lambda' = e^{\phi}\lambda$, then $\Lambda(\lambda') = \Lambda(\lambda)$, thus it is natural to impose $\det v = 1$ and to consider $\mathrm{SL}(2,\mathbb{C})$ instead of $L(2,\mathbb{C})$ to fit $L_+^\uparrow$. Now, $\Lambda(\lambda) = \Lambda(-\lambda)$, and we wish to consider $\mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$.

I think the problem is the following: as far as the action of the "nature group" on the space-time is concerned, it is sufficient to consider $L_+^\uparrow$. But the group which acts on the state space is wider: it must be $SL(2,\mathbb{C})$.

From now, when we say "Poincaré group", we mean $\mathrm{SL}(2,\mathbb{C}) \times \mathbb{R}^4$ while "Lorentz" means $\mathrm{SL}(2,\mathbb{C})$ acting on $\mathbb{R}^4$ by $\Lambda(\lambda)v = \lambda v \lambda^\dagger$.

We know by lemma 55.73 a link between the representations of SU(2) and SO(3). A know result is the fact that the map Spin restricts to a surjective homomorphism $\mathrm{Spin}\colon SU(2) \to \mathrm{SO}(3)$ with kernel $\pm\mathbb{1}$ giving the relation $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$. If one considers a representation $\rho\colon \mathrm{SO}(3) \to GL(V)$, then $\tilde{\rho} = \rho \circ \mathrm{Spin}$ is a representation of $SU(2)$ on V. So every representation of SO(3) comes from a representation of $SU(2)$.

As far as the transformation rule of a (quantum mechanical) wave function under a rotation $R \in \mathrm{SO}(3)$ is concerned, one can see (it is done in [915]) that the try

$$\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \to T(R)\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$$

doesn't works if $T(R)$ is a representation of SO(3) on $\mathbb{C}^2$. If one allows T to be a representation of $SU(2)$, then our choice —for an electron— should naturally be the spin one half representation $T = D^{(1/2)}$. Let us do it. The remaining problem is the following. Let's consider that in a certain frame, an electron is described by the wave function $\begin{pmatrix}\psi_1 & \psi_2\end{pmatrix}$, the question is to know the wave function observed by a guy which use another frame linked to the first frame by $R \in \mathrm{SO}(3)$. We always have exactly two elements in SU(2) projected to R by Spin; namely $\mathrm{Spin}(\pm g) = R$; so how to choose between

$$D^{(1/2)}(g)\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \quad \text{and} \quad D^{(1/2)}(-g)\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} ?$$

The trick is to remark that a change of frame is not the mathematical process described by a single element R of SO(3), but a physical *continuous* process which begins at the identity and stops at R. In other word, we have to ask ourself *how to go from a frame to another*? Taking as example the rotations around the x axis, we can look at two different path in SO(3) from $\mathbb{1}$ to $\mathbb{1}$ given by the same expression

$$R_1(t) = R_2(t) = \begin{pmatrix}1 & 0 & 0\\ 0 & \cos t & \sin t\\ 0 & -\sin t & \cos t\end{pmatrix},$$

but considering $t\colon 0 \to 2\pi$ for R_1 and $t\colon 0 \to 4\pi$ for R_2. The covering map $\mathrm{Spin}\colon SU(2) \to \mathrm{SO}(3)$ allows us to lift any path in SO(3) to a path in $SU(2)$ in an unique way providing a starting point. In other words, if $\mathrm{Spin}(g) = R$,

$$\exists!\, \tilde{R}(t) \in SU(2) \text{ such that } \mathrm{Spin}\circ\tilde{R} = R \text{ and } \tilde{R}(0) = \mathbb{1},$$
$$\exists!\, \tilde{R}(t) \in SU(2) \text{ such that } \mathrm{Spin}\circ\tilde{R} = R \text{ and } \tilde{R}(0) = -\mathbb{1}.$$

The question is now: how to choose the right path among these two? The answer comes from the homotopy of SO(3): the path R_1 and R_2 belongs to two different classes.

Considering the "change of frame" as a continuous process, the initial point is naturally chosen to be $\mathbb{1}$. With this choice, the lift of R_1 and R_2 are given by

$$g_1(t) = g_2(t) = \begin{pmatrix}\cos\frac{t}{2} & -i\sin\frac{t}{2}\\ -i\sin\frac{t}{2} & \cos\frac{t}{2}\end{pmatrix}$$

with $t\colon 0 \to 2\pi$ for g_1 and $t\colon 0 \to 4\pi$ for g_2. In $SU(2)$, the ending point of g_1 is $-\mathbb{1}$ while the one of g_2 is $\mathbb{1}$.

It is still possible to say a lot of interesting thinks about the space-time symmetry group of nature; let's just conclude saying that $SU(2)$ is more adapted to the rotations of non zero spin than SO(3). (it is not intuitive!)

69.2.2 How to implement the Poincaré group

pg:poincare_act

We are not making physics here, but differential geometry and group theory; so we will not discuss the physical relevance of the Poincaré group from a "speed of light" point of view. We consider the **Poincaré group** as the group of all the affine isometries of metric $\eta = diag(-1,1,1,1)$ and the **Lorentz group** as the subgroup of rotations and boost.

A Poincaré transformation of $\mathbb{R}^4$ is given by (Λ, a) with Λ a 4×4 matrix and $a \in \mathbb{R}^4$, a translation vector. The composition of (Λ, a) with (Λ', a') is given by $(\Lambda'\Lambda, \Lambda' a + a')$, the inverse is $(\Lambda^{-1}, -\Lambda^{-1}a)$, the neutral is $(\mathbb{1}, 0)$, and $(\det \Lambda)^2 = 1$.

The axiom (4) at page 3701 gives us a group of transformation of the rays in $\mathscr{H}$ parametrised by (Λ, a) such that

$$T(\Lambda', a')T(\Lambda, a) = T(\Lambda'\Lambda, \Lambda' a + a'), \tag{69.40}$$

$T(\Lambda, a)\colon \mathrm{Ray}\,\mathscr{H} \to \mathrm{Ray}\,\mathscr{H}$. Then Wigner's theorem defines a representation of the Poincaré group on $\mathscr{H}$ by unitary matrices:

$$\psi \to U(\Lambda, a)\psi.$$

Remarque 69.7.
Wigner only ensure existence of *projective* representations. Here we suppose that our symmetry group (maybe slightly different that Poincaré) is such that any projective representations can be turn into a classical representation. We will therefore use the composition law

eq:composition_U

$$U(\Lambda', a')U(\Lambda) = U(\Lambda'\Lambda, \Lambda' a + a') \tag{69.41}$$

instead of $U(\Lambda', a')U(\Lambda, a') = e^{i\phi(\Lambda, a, \Lambda', a')}U(\Lambda'\Lambda, \Lambda' a + a')$.

By axiom, the (connected) Poincaré group acts on rays of $\mathscr{H}$, and we have the representation U which form a group acting on $\mathscr{H}$. The Lie algebra acts also:

$$u\psi = \frac{d}{dt}\Big[U(t)\Big]_{t=0}\psi := \frac{d}{dt}\Big[U(t)\psi\Big]_{t=0}. \tag{69.42}$$

This definition is natural because $\mathscr{H}$ is a vector space: it can be identified with its tangent space: $U(t)\psi$ is a path in $\mathscr{H}$ and its derivative at $t=0$ is still a well defined element in $\mathscr{H}$. Now recall that the operators U are unitary, so that the corresponding operators u are hermitian (therefore diagonalisable).

Let us consider an abelian subgroup A of Poincaré with Lie algebra $\mathfrak{a}$. One can find a basis of $\mathscr{H}$ made of common eigenvectors of a basis of $\mathfrak{a}$. In other words, one can find a basis of $\mathscr{H}$ which simultaneously diagonalises all $\mathfrak{a}$. If $\{a_i\}$ is a basis of $\mathfrak{a}$, one can find a basis $\{|\psi_\lambda\rangle\}$ (here λ labels a basis of $\mathscr{H}$: it might take continuous values) such that

$$a_i|\psi_\lambda\rangle = \lambda_i|\psi_\lambda\rangle. \tag{69.43}$$

69.2.3 Momentum operator

Of course, there exists an abelian subgroup of Poincaré: the pure translations, $A = \{U(\mathbb{1}, a)\}$. A basis of the Lie algebra is given by four vectors labeled as P^μ and defined by

$$P^\mu = \frac{d}{dt}\Big[U(\mathbb{1}, te^\mu)\Big]_{t=0}$$

where e^μ is the unit vector following the direction μ (for $\mu = 0$, $e^0 = (1,0,0,0)$). One can consider a basis which diagonalises the P^μ's:

$$P^\mu|p,\sigma\rangle = p^\mu|p,\sigma\rangle \tag{69.44}$$

where by definition,

$$P^\mu|p,\sigma\rangle = \frac{d}{dt}\Big[U(\mathbb{1},te^\mu)|p,\sigma\rangle\Big]_{t=0}. \tag{69.45}$$

eq:def_P

Remarque 69.8.
Be careful on a point: we don't say anything about the symbol "p" in the ket. The only property is that it labels a Hilbert space $\mathscr{H}$. But nothing is already imposed to $\mathscr{H}$: it must just carry a representation of the Poincaré group on its rays. In particular, it is *a priori* false to say that p is a "momentum 4-vector" and that p^μ is a component of p. Naturally, our notations are adapted to think that! Maybe it is a pedagogical mistake; I don't know.

This remark can be disturbing: why is generally $|p,\sigma\rangle$ called "a state of momentum p"? Since $U(\mathbb{1},a)$ is unitary, P^μ is hermitian; the p^μ are eigenvalues for an hermitian operator, so by axiom (2) (page 3701) they are candidate to be physical values. But equation (69.45) shows that P^μ is what a physicist should call an "infinitesimal translation", so that Noether suggests us to interpret the eigenvalue as momentum. We are safe!

The parameters σ are not yet defined neither. It will come later. For the moment, we include into the definition of a **one particle state** that σ takes discrete values.

Since $U(\mathbb{1},a) = e^{a_\mu P^\mu}$,

$$U(\mathbb{1},a)|p,\sigma\rangle = e^{ia_\mu p^\mu}|p,\sigma\rangle.$$

Now we are interested in the determination of $U(\Lambda,a)|p,\sigma\rangle$.

Proposition 69.9.
The operators P^μ are subject to the "transformation law"

$$U(\Lambda,a)P^\mu U(\Lambda,a)^{-1} = \Lambda^\mu_\nu P^\nu. \tag{69.46}$$

Proof. Since operators $U(\Lambda,a)$ are linear, they can be putted in the derivative which defines P^μ. Using the composition law (69.41) we find:

$$\begin{aligned} U(\Lambda,a)P^\mu U(\Lambda,a)^{-1} &= \frac{d}{dt}\Big[U(\Lambda,a)U(\mathbb{1},te^\mu)U(\Lambda,a)^{-1}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[U(\mathbb{1},t\Lambda e^\mu)\Big]_{t=0}. \end{aligned} \tag{69.47}$$

The Λ can be putted out of derivative; let us see it for a sum of two terms (here it is four):

$$\begin{aligned} \frac{d}{dt}\Big[U(\mathbb{1},t(e^\mu+e^\nu))\Big]_{t=0} &= \frac{d}{dt}\Big[U(\mathbb{1},te^\mu)U(\mathbb{1},te^\nu)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[U(\mathbb{1},te^\mu)U(\mathbb{1},0)\Big]_{t=0} + \frac{d}{dt}\Big[U(\mathbb{1},0)U(\mathbb{1},te^\nu)\Big]_{t=0} \\ &= P^\mu + P^\nu. \end{aligned} \tag{69.48}$$

Thus

$$\frac{d}{dt}\Big[U(\mathbb{1},\Lambda^\mu_\nu e^\nu)\Big]_{t=0} = \Lambda^\mu_\nu \frac{d}{dt}\Big[U(\mathbb{1},te^\nu)\Big]_{t=0} = \Lambda^\mu_\nu P^\nu. \tag{69.49}$$

□

69.2.4 Pure Lorentz transformation

Now we consider a pure Lorentz transformation $U(\Lambda) \equiv U(\Lambda,0)$, and we want to look at $U(\Lambda)|p,\sigma\rangle$. In order to see its decomposition into others $|k,\sigma'\rangle$, we apply a P^μ:

$$\begin{aligned} P^\mu U(\Lambda)|p,\sigma\rangle &= U(\Lambda)\Big(U(\Lambda)^{-1}P^\mu U(\Lambda)\Big)|p,\sigma\rangle \\ &= U(\Lambda)(\Lambda^{-1})^\mu_\nu P^\nu|p,\sigma\rangle \\ &= (\Lambda^{-1})^\mu_\nu p^\nu U(\Lambda)|p,\sigma\rangle. \end{aligned} \tag{69.50}$$

Thus the vector $U(\Lambda)|p,\sigma\rangle \in \mathscr{H}$ has $(\Lambda^{-1})^\mu_\nu p^\nu$ as eigenvalue for P^μ. If the p^μ's are seen as components of a 4-vector p, one can write

$$P^\mu U(\Lambda)|p,\sigma\rangle = (\Lambda p)^\mu U(\Lambda)|p,\sigma\rangle;$$

thus we naturally write

$$U(\Lambda)|p,\sigma\rangle = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle. \tag{69.51}$$

Note that we had not yet given anything about the nature of the p in the ket $|p,\sigma\rangle$ so we can *define* the product Λp by the fact that the ket $|\Lambda p,\sigma\rangle$ has eigenvalue $(\Lambda^{-1})^\mu_\nu p^\nu$ for the operator P^μ. So it is one of the $|p',\sigma'\rangle$.

69.2.5 Rebuilding of a basis for $\mathscr{H}$

From general considerations about the Lorentz group (many physicists had written very better books than me about) anyone knows that the only functions of the p^μ's which are invariant under all the Lorentz transformations are $p^2 = \eta_\mu nup^\mu p^\nu$ and the sign of p^0 when $p^2 < 0$.

For any value of p^2 and sign of p^0, one consider a "standard vector" k. For example:

$$k = (1,0,0,1) \qquad \text{for } p^2 = 0, \tag{69.52a}$$

$$k = (1,0,0,0) \qquad \text{for } p^2 < 0, p^0 > 0 \tag{69.52b}$$

$$k = (-1,0,0,0) \qquad \text{for } p^2 < 0, p^0 < 0. \tag{69.52c}$$

With this convention, p can be written as $p = L(p)k$ for a suitable Lorentz transformation $L(p)$. The vector $U(L(p))|k,\sigma\rangle$ has eigenvalue $L(p)k$ for the operator P, thus it is a linear combination of some $|p,\sigma'\rangle$.

Now we will cheat and redefine our basis of the Hilbert space $\mathscr{H}$. First, we consider a fixed k; in other words, we build the state space for a given particle which has given momentum p. The basis vectors must be eigenvectors for the fours operators P^μ. As far as we say no more, any eigenvalue is possible. Thus our basis must be labelled by at least an element p of $\mathbb{R}^4$ with only one constraint: the value of p^2 (plus eventually the sign of p^0). So we define the $|k,\sigma\rangle$ to be such that

$$P^\mu|k,\sigma\rangle = k^\mu|k,\sigma\rangle.$$

Since we know that with this definition of $|k,\sigma\rangle$, the eigenvalue of $U(L(p))|k,\sigma\rangle$ for P^μ is p^μ, we *define* $|p,\sigma\rangle$ as

$$|p,\sigma\rangle = N(p)U(L(p))|k,\sigma\rangle. \tag{69.53}$$

where $N(p)$ is a normalization to be discussed later. With this construction, we have an eigenvector for any possible eigenvalue for P^μ. We have to show that these vectors are linearly independent.

The set of the $|p,\sigma\rangle$ with different p is free in $\mathscr{H}$ because they are eigenvectors for different eigenvalue of an hermitian operator [2]. There are no reason to think that the set of operators P^μ is complete; in other words, it remains not clear that there exist only one way to diagonalise the all the P^μ. The function of the extra label σ is to label different linearly independent vectors with same eigenvalue for P.

From now, we are interested in $|k,\sigma\rangle$ and $N(p)$.

69.2.6 Little group

We have:

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= N(p)U(\Lambda L(p))|k,\sigma\rangle \\ &= N(p)U(L(\Lambda p))U\big(L(\Lambda p)^{-1}\Lambda L(p)\big)|k,\sigma\rangle, \end{aligned} \tag{69.54}$$

2. I did not checked that it is sufficient

So we will try to understand the operation $L(\Lambda p)^{-1}\Lambda L(p)$. First remark that

$$U(L(\Lambda p)^{-1})|\Lambda p,\sigma\rangle = N(\Lambda p)|k,\sigma\rangle,$$

and then compute:

$$\begin{aligned} U(L(\Lambda p)^{-1}\Lambda L(p))N(p)|k,0\rangle &= U(L(\Lambda p)^{-1}\Lambda)|p,\sigma\rangle \\ &= U(L(\Lambda p)^{-1})\sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle \\ &= \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)N(\Lambda p)|k,\sigma'\rangle. \end{aligned} \tag{69.55}$$

The **little group** is the subgroup of the Lorentz transformations which leaves the chosen standard vector k invariant: $Wk = k$. For any W in the little group,

$$U(W)|k,\sigma\rangle = \sum_{\sigma'} D_{\sigma'\sigma}(W)|k,\sigma'\rangle$$

With this definition, the D's form a representation of the little group. Indeed for any V, W in the little group,

$$\begin{aligned} \sum_{\sigma'} D_{\sigma'\sigma}(VW)|k,\sigma'\rangle &= U(VW)|k,\sigma\rangle \\ &= U(V)\sum_{\sigma''} D_{\sigma''\sigma}(W)|k,\sigma''\rangle \\ &= \sum_{\sigma'\sigma''} D_{\sigma'\sigma''}(V)D_{\sigma''\sigma}(W)|k,\sigma'\rangle. \end{aligned} \tag{69.56}$$

Since we want the $|p,\sigma\rangle$ with different p and σ to form a basis of $\mathscr{H}$, they are linearly independent, then we can get rid of the sum over the σ' and keep the equation

$$D_{\sigma'\sigma}(VW) = \sum_{\sigma''} D_{\sigma'\sigma''}(VW)D_{\sigma''\sigma}(VW);$$

if we adopt a more "matricial" notation,

$$D(VW) = D(V)D(W). \tag{69.57}$$

We are now able to perform a step in the study of the vector $U(\Lambda)|p,\sigma\rangle$. We naturally define $W(\Lambda,p) = L(\Lambda p)^{-1}\Lambda L(p)$. This belongs to the little group [3]. Then,

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= N(p)U(L(\Lambda p))U(W(\Lambda,p))|k,\sigma\rangle \\ &= N(p)\sum_{\sigma'} D_{\sigma'\sigma}(W)U(L(\Lambda p))|k,\sigma'\rangle \\ &= \frac{N(p)}{N(\Lambda p)}\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))|\Lambda p,\sigma'\rangle. \end{aligned} \tag{69.58}$$

But we have no constraint on the D's: it must just form a representation of the little group. Consequently, we are at a point in which our axioms are no more sufficient to continue the building of quantum field theory: we will get as many theories as representations of the little group.

The physical interpretation is the following: each type of particle has its own representation. When we consider a Hilbert space on which $U(\Lambda)$ acts via one given representation of the little group, we consider the Hilbert space which describes the corresponding particle. Note that the little group depends on the choice of k, and therefore depends on the particle which is studied (massive or not).

pg:phyz_reprez

In this sense, a particle is a representation of the Poincaré group [4]. In particular, the nature of the index σ can change from the one representation to the other.

3. Pay attention that $L(p)$ depends implicitly on the choice of k.

4. I think that the irreducibility of a representation is related to *elementary* particles.

Remarque 69.10.
As far as normalization is concerned, we will pose

$$N(p) = \sqrt{k^0/p^0}.$$

There are some good reasons to take it; but it is irrelevant from our group point of view of the theory.

69.2.7 Positive mass

This is the easy case. The choice of standard momentum is $k = \begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix}$. One could believe that the little group is SO(3). It would be the case if we had chosen $L_+^\uparrow$ instead of $\mathrm{SL}(2,\mathbb{C})$ –see point 69.2.1. In our hermitian representation of $\mathbb{R}^4$, $k = \mathbb{1}$. Then a matrix of $\mathrm{SL}(2,\mathbb{C})$ which leaves it invariant fulfills

$$\lambda k \lambda^\dagger = \lambda\lambda^\dagger = \mathbb{1},$$

this is $\lambda \in SU(2)$. By the way, note that $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$.

The celebrated "law of transformation" of a massive particle of spin j (integer or half integer) under the Lorentz transformation Λ is

$$U(\Lambda)|p,\sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D^{(j)}_{\sigma'\sigma}(W(\Lambda,p))|\Lambda p,\sigma'\rangle \tag{69.59}$$

where σ runs from $-j$ to j by step of 1.

69.2.8 Null mass

In the case of a null mass, the standard vector is $k = (1,0,0,1)$ and an element of the little group fulfils $Wk = k$. As the little group is part of the Lorentz group, this is an isometry, so

$$\langle Wt|Wk\rangle = \langle t|k\rangle \tag{69.60a}$$
$$\langle Wt|Wt\rangle = \langle t|t\rangle, \tag{69.60b}$$

for any $t \in \mathbb{R}^4$. Taking in particular $t = (1,0,0,0)$,

$$(Wt)^\mu k_\mu = t^\mu k_\mu = -1 \tag{69.61a}$$
$$(Wt)^\mu (Wt)_\mu = t^\mu t_\mu = -1. \tag{69.61b}$$

If we write $Wt = (a,b,c,d)$, the first relation gives $d = a-1$, so that $Wt = (1+\xi,\alpha,\beta,\xi)$, while the second one gives $\xi = (\alpha^2+\beta^2)/2$. The conclusion is that W acts on t as a certain Lorentz transformation $S(\alpha,\beta)$:

$$Wt = \begin{pmatrix}1+\xi\\ \alpha\\ \beta\\ \xi\end{pmatrix} = \begin{pmatrix}1+\xi & -\xi & \alpha & \beta\\ \alpha & -\alpha & 1 & 0\\ \beta & -\beta & 0 & 1\\ \xi & (1+\xi) & \alpha & \beta.\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}. \tag{69.62}$$

Be careful: it doesn't means that $W = S$, but $Wt = St$. However it is an information: $S(\alpha,\beta)^{-1}W$ is a Lorentz transformation which leaves t invariant. Then it is a spatial rotation. More precisely, since W and S conserve $(1,0,0,1)$, it is a rotation around the z axis: $S(\alpha,\beta)^{-1}W = R(\theta)$, and

$$W(\theta,\alpha,\beta) = S(\alpha,\beta)R(\theta) \tag{69.63}$$

is the most general element of the non massive little group.

Chapter 70

Relativistic fields and fiber bundle formalism

This chapter actually don't deal with *quantum* field theory in the sense that our wave functions aren't operators which acting on a Fock space. So this is just relativistic field theory.

70.1 Example: electromagnetism

Let us consider the electromagnetism as the simplest example of a gauge invariant physical theory. We first discuss the theory of free electromagnetic field (this is: without taking into account the interactions with particles) from Maxwell's equations [851, 979]. The electric field $\overline{E}$ and the magnetic field $\overline{B}$ are subject to following relations:

$$\nabla \cdot \overline{E} = \rho, \tag{70.1a}$$
$$\nabla \cdot \overline{B} = 0, \tag{70.1b}$$
$$\nabla \times \overline{E} + \partial_t \overline{B} = 0, \tag{70.1c}$$
$$\nabla \times \overline{B} - \partial_t \overline{E} = \overline{j}. \tag{70.1d}$$

Comparing (70.1a) and (70.1b), we see that Maxwell's theory does not incorporate magnetic monopoles. Suppose that we can use the Poincaré lemma. Equation (70.1b) gives a vector field $\overline{A}$ such that $\overline{B} = \nabla \times \overline{A}$, so that (70.1c) becomes $\nabla \times (\overline{E} + \partial_t \overline{A}) = 0$ which gives a scalar field ϕ such that $-\nabla \cdot \phi = \overline{E} + \partial_t \overline{A}$.

Now the equations (70.1a)–(70.1d) are equations for the potentials $\overline{A}$ and ϕ, and we find back the "physical" field by

$$\overline{B} = \nabla \times \overline{A}, \tag{70.2a}$$
$$\overline{E} = -\nabla\phi - \partial_t \overline{A}. \tag{70.2b}$$

One can easily see that there are several choice of potentials which describe the same electromagnetic field. Indeed, if

$$\overline{A}' = \overline{A} + \nabla\lambda, \tag{70.3a}$$
$$\phi' = \phi - \partial_t \lambda, \tag{70.3b}$$

the electromagnetic field given (via (70.2)) by $\{\phi', \overline{A}'\}$ is the same as the one given by $\{\phi, \overline{A}\}$

The Maxwell's equations can be written in a more "covariant" way by defining

$$F = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ \cdot & 0 & -B_z & \cdot \\ \cdot & \cdot & 0 & -B_x \\ \cdot & -B_y & \cdot & 0 \end{pmatrix}, \tag{70.4}$$

$F^{\mu\nu} = -F^{\nu\mu}$ and

$$J = \begin{pmatrix} c\rho & j_x & j_y & j_z \end{pmatrix}.$$

We also define $\star F^{\alpha\beta} = \frac{1}{2} e^{\alpha\beta\lambda\mu} F_{\lambda\mu}$. With all that, Maxwell's equations read:

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= \mu_0 J^\nu, \\ \partial_\alpha \star F^{\alpha\beta} &= 0. \end{aligned} \tag{70.5}$$

If we define

$$A = \begin{pmatrix} \frac{\phi}{c} & -A_x & -A_y & -A_z \end{pmatrix}, \tag{70.6}$$ def_A

the physical fields are given by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

The **gauge invariance** of this theory is the fact that

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = F_{\mu\nu} \tag{70.7a}$$

when

$$A'_\mu(x) = A_\mu(x) + \partial_\mu f(x) \tag{70.7b}$$

for any scalar *function* f (to be compared with (70.3)).

This is: in the picture of the world in which we see the A as fundamental field of physics, several (as much as you have functions in $C^\infty(\mathbb{R}^4)$) fields A, A',... describe the *same* physical situation because the fields $\overline{E}$ and $\overline{B}$ which acts on the particle are the same for A and A'.

Now, we turn our attentions to the interacting field theory of electromagnetism. As far as we know, the electron makes interactions with the electromagnetic field via a term $\overline{\psi} A_\mu \psi$ in the Lagrangian. The free Lagrangian for an electron is

$$\mathcal{L} = \overline{\psi}(\gamma^\mu \partial_\mu + m)\psi. \tag{70.8}$$ eq:freeL

The easiest way to include a $\overline{\psi} A \psi$ term is to change ∂_μ to $\partial_\mu + A_\mu$. But we want to preserve the powerful gauge invariance of classical electrodynamics, then we want the new Lagrangian to keep unchanged if we do

$$A_\mu \to A'_\mu = A_\mu - i\partial_\mu \phi. \tag{70.9}$$ eq__jaugeA_em

In order to achieve it, we remark that the ψ must be transformed *simultaneously* into

$$\psi'(x) = e^{i\phi(x)} \psi(x). \tag{70.10}$$ eq:jaugepsi_em

The conclusion is that if one want to write down a Lagrangian for QED, one must find a Lagrangian which remains unchanged under certain transformation $A \to A'$ and $\psi \to \psi'$. In other words the set $\{\psi, A\}$ of fields which describe the world of an electron in an electromagnetic field is not well defined from data of the physical situation alone: it is defined up to a certain invariance which is naturally called a **gauge invariance**.

70.2 Connections for the gauge invariance

In the physics books, the matter is presented in a slightly different way. We observe that the Lagrangian (70.8) is invariant under

$$\psi(x) \to \psi'(x) = e^{i\alpha} \psi(x) \tag{70.11}$$ eq:globale

for any *constant* α. One can see that the associated conserved current (Noether) is closely related to the electric current. The idea (of Yang-Mills) is to upgrade this symmetry. Since the symmetry (70.11) depends only on a constant, we say it a **global** symmetry; we will simultaneously add a new field A_μ and upgrade (70.11) to a **local** symmetry:

$$\psi(x) \to \psi'(x) = e^{i\phi(x)} \psi(x). \tag{70.12}$$ eq:locale

Then, we deduce the transformation law of A_μ.

Because of the form of (70.10), we say that the electromagnetism is a $U(1)$-gauge theory. The fact that this is an abelian group has a deep physical meaning and many consequences.

70.2.1 Little more general, slightly more formal

The aim of this text is to interpret the field A as a gauge potential for a connection. But equation (70.9) is not exactly the expected one which is (56.358). The point is that equation (70.9) concerns a theory in which the gauge transformation of the field was a simple multiplication by a scalar field, so that simplifications as $e^{-i\phi(x)}A_\mu(x)e^{i\phi(x)} = A_\mu(x)$ are allowed.

Now, we consider a vector space V, a manifold M and a function $\psi\colon M \to V$ which "equation of motion" is

$$L^i(\partial_i + m_i)\psi = 0$$

Where we imply an unit matrix behind ∂ and m; the indices i, j are the (local) coordinates in M and a, b, the coordinates in V. Let G be a matrix group which acts on V. If ψ is a solution, $\Lambda^{-1}\psi$ is also a solution as far as Λ is a constant –does not depend on $x \in M$– matrix of G. In other words, $L^i(\partial_i + m_i)\psi_a = 0$ for all a implies $L^i(\partial_i + m_i)((\Lambda^{-1})^b_a\psi_b) = 0$.

The function, $\psi'(x) = \Lambda(x)^{-1}\psi(x)$ is no more a solution. If we want it to be solution of the same equation as ψ, we have to change the equation and consider

$$L^i(\partial_i + A_i + m_i)\psi = 0.$$

This equation is preserved under the *simultaneous* change

$$\begin{cases} \psi'_a = (\Lambda^{-1})^b_a\psi_b \\ (A'_i)^a_b = (\Lambda^{-1})^c_b(A_i)^d_c(\Lambda^a_d) - (\partial_i\Lambda^{-1})^d_b\Lambda^a_d. \end{cases} \tag{70.13}$$ eq:jaugeG

The second line show that the formalism in which A is a connection is the good one to write down covariant equations. This has to be compared with (56.350). Logically, a theory which includes an invariance under transformations as (70.13) is called a G-gauge theory.

70.2.2 A "final" formalism

Now, we work with fields which are sections of some fiber bundle build over M, the physical space. More precisely, let G be a matrix group.

¡¡ Avertissement/question à la lectrice !! 70.1
For sure, it also works for a much lager class of groups. Which one?

We search for a theory which is "locally invariant under G". In order to achieve it, we consider a G-principal bundle P over M and the associated bundle $E = P \times_\rho V$ for a certain vector space V, and a representation ρ of G on V. Typically, V is $\mathbb{C}$ or the vector space on which the spinor representation acts.

The physical fields are sections $\psi\colon M \to E$. If we choose some reference sections $\sigma_\alpha\colon M \to P$, they can be expressed by $\psi_{(\alpha)}(x) = \hat{\psi}(\sigma_\alpha(x))$. We translate the idea of a local invariance under G by requiring an invariance under

$$\psi'_{(\alpha)}(x) = \rho(g(x))\psi_{(\alpha)}(x)$$

for every $g\colon M \to G$. By (2) of lemma 56.90, we see that $\psi'_{(\alpha)}(x) = (\varphi^{-1} \cdot \psi)_{(\alpha)}(x)$, where $\varphi\colon P \to P$ is the gauge transformation given by

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x) \cdot g(x).$$

We want ψ and ψ' to "describe the same physics". From a mathematical point of view, we want ψ and ψ' to *satisfy the same equation*. It is clear that equation $d\psi = 0$ will not work.

The trick is to consider any connection ω on P and the gauge potential A of ω. In this case the equation

$$(d - A)\psi = 0 \qquad \text{or} \qquad D\psi = 0 \tag{70.14}$$ eq:Dpsi

is preserved under

$$A \to \varphi \cdot A,$$
$$\psi \to \varphi^{-1} \cdot \psi.$$

Theorem 56.91 powa!

In this sense, we say that equation (70.14) is gauge invariant, and is thus taken by physicists to build some theories when they need a "local G-covariance". This gives rise to the famous Yang-Mills theories.

In this picture the matter field ψ and the bosonic field A are both defined from a $U(1)$-principal bundle. When physicists say

ψ transforms as "blahblah" under a $U(1)$ transformation,

they mean that ψ is a section of an $U(1)$-associated bundle; when they say

A transforms as "blahblah" under a $U(1)$ transformation,

they mean that A is the gauge potential of a connection on a $U(1)$-principal bundle. In each case, the "blahblah" denotes an irreducible [1] representation of $U(1)$.

Remarque 70.2.
The mathematics of equation (70.14) only requires a $\mathcal{G}$-valued connection on P. There are several physical constraints on the choice of the connection. These give rise to interaction terms between the gauge bosons. We will not discuss it at all. This a matter of books about quantum field theories.

The most used Yang-Mills groups in physics are $U(1)$ for the QED, $SU(2)$ for the weak interactions and $SU(3)$ for chromodynamic.

70.2.3 The electromagnetic field F

Now, we are able to interpret the field F introduced in equation (70.4). We follow [977]. From now, we use the usual Minkowski metric $g = diag(-,+,+,+)$. From the vector given by (70.6), we define a (local) potential 1-form

$$A = A_\mu dx^\mu = -\phi dt + A_x dx + A_y dy + A_z dz.$$

The field strength is $F = dA$. We easily find that

$$\begin{aligned} F = &(dt \wedge dx)(\partial_x \phi + \partial_t A_x) + \dots \\ &+ (dx \wedge dy)(-\partial_z A_x + \partial_x A_y) + \dots \end{aligned} \tag{70.15}$$

But the fields $\overline{B}$ and $\overline{E}$ are defined from $\overline{A}$ and ϕ by (70.2), so

$$\begin{aligned} F = &-E_x(dt \wedge dx) - E_y(dt \wedge dy) - E_z(dt \wedge dz) \\ &+ B_x(dy \wedge dz) + B_y(dz \wedge dx) + B_z(dx \wedge dy). \end{aligned} \tag{70.16}$$

We naturally have $dF = d^2A = 0$. But conversely, $dF = 0$ ensures the existence of a 1-form A such that $F = dA$. If we define [2] $\overline{B} = \nabla \times \overline{A}$ and $\overline{E} = -\nabla\phi - \partial_t \overline{A}$, equations (70.1b) and (70.1c) are obviously satisfied. So in the connection formalism, the equations "without sources" are written by

$$dF = 0. \tag{70.17}$$ M23

In order to write the two others, we introduce the current 1-form:

$$j = j_\mu dx^\mu = -\rho dt + j_x dx + j_y dy + j_z dz.$$

1. Irreducibility is for elementary particles
2. *i.e.* we consider F as the main physical field while $\overline{E}$ and $\overline{B}$ are "derived" fields.

One sees that

$$\begin{aligned}\delta F := \star d \star F = &-dt(\nabla \cdot \overline{E}) \\ &+ dx(-\partial_t \overline{E}_x + (\nabla \times \overline{B})_x) \\ &+ dy(-\partial_t \overline{E}_y + (\nabla \times \overline{B})_y) \\ &+ dz(-\partial_t \overline{E}_z + (\nabla \times \overline{B})_z),\end{aligned} \tag{70.18}$$

so that equation $\delta F = j$ gives equations (70.1a) and (70.1d). Now, the complete set of Maxwell's equations is:

$$dF = 0 \tag{70.19a}$$ SM1

$$\delta F = j \tag{70.19b}$$ SM2

with

$$j = -\rho dt + j_x dx + j_y dy + j_z dz, \tag{70.20a}$$

$$\overline{B} = \nabla \times \overline{A} \tag{70.20b}$$

$$\overline{E} = -\nabla\phi - \partial_t \overline{A} \tag{70.20c}$$

where A is a 1-form such that $F = dA$ whose existence is given by (70.19a).

70.3 Spin manifold for Lorentz invariance

subsec:incl_Lorentz

Up to now we had seen how to express the *gauge* invariance of a physical theory. In particle physics, a really funny field theory must be invariant under the Lorentz group; it is rather clear that, from the bundle point of view, this feature will be implemented by a Lorentz-principal bundle and some associated bundles. A spinor will be a section of an associated bundle for spin one half representation of the Lorentz group on $\mathbb{C}^4$. In order to describe non-zero spin particle interacting with an electromagnetic field (represented by a connection on a $U(1)$-principal bundle), we have to build a correct $\mathrm{SL}(2,\mathbb{C}) \times U(1)$-principal bundle. We are going to use the ideas of 69.2.1.

A **space-time** is a differentiable **pseudo-Riemannian** 4-dimensional manifold. The pseudo-Riemannian structure is a 2-form $g \in \Omega^2(M)$ for which we can find at each point $x \in M$ a basis $b = (\overline{b}_0, \dots, \overline{b}_3)$ which fulfils

$$g_x(\overline{b}_i, \overline{b}_j) = \eta_{ij}.$$

When we use an adapted coordinates, the metric reads $g = \eta_{ij} dx^i \otimes dx^j$.

One says that M is **time orientable** if one can find a vector field $T \in \mathfrak{X}(M)$ such that $g_x(T_x, T_x) > 0$ for all $x \in M$. A **time orientation** is a choice of such a vector field. A vector $v \in T_x M$ is **future directed** if $g_x(T_x, v) > 0$.

The Lorentz group L acts on the orthogonal basis of each $T_x M$, but you may note that L don't act on M; it's just when the metric is flat that one can identify the whole manifold with a tangent space and consider that L is the space-times isometry group. In the case of a curved metric, the Lorentz group have to be introduced pointwise and the building of a frame bundle is natural.

Now, we are mainly interested in the frame related each other by a transformation of $L^{\uparrow}_{+}$. An arising question is to know if one can make a choice of some basis of each $T_x M$ in such a manner that

(1) pointwise, the chosen frames are related by a transformation of $L^{\uparrow}_{+}$,

(2) the choice is globally well defined.

The first point is trivial to fulfil from the definition of a space-time. For the second, it turns out that a good choice can be performed if and only if there exists a vector field $V \in \mathfrak{X}(M)$ such that $g_x(V_x, V_x) > 0$ for all $x \in M$. We suppose that it is the case [3].

3. That condition is rather restrictive because we cannot, for example, find an everywhere non zero vector field on the sphere S^n with n even.

So our first principal bundle attempt to describe the space-time symmetry is the $L^{\uparrow}_{+}$-principal bundle of orthonormal oriented frame on M:

$$\begin{array}{ccc} L^{\uparrow}_{+} & \rightsquigarrow & L(M) \\ & & \downarrow{\scriptstyle p_L} \\ & & M \end{array} \tag{70.21}$$

bun:LpF

The notion of "**relativistic invariance**" has to be understood in the sense of associated bundle to this one. The next step is to recall ourself (see subsection 69.2.1) that the physical fields doesn't transform under representation of the group $L^{\uparrow}_{+}$ but rather under representations of $\mathrm{SL}(2,\mathbb{C})$. So we build a $\mathrm{SL}(2,\mathbb{C})$-principal bundle

$$\begin{array}{ccc} \mathrm{SL}(2,\mathbb{C}) & \rightsquigarrow & S(M) \\ & & \downarrow{\scriptstyle p_S} \\ & & M \end{array}$$

In order this bundle to "fit" as close as possible the bundle (70.21), we impose the existence of a map $\lambda\colon S(M) \to L(M)$ such that

(1) $p_B(\lambda(\xi)) = p_S(\xi)$ for all $\xi \in S(M)$ and

(2) $\lambda(\xi \cdot g) = \lambda(\xi) \cdot \mathrm{Spin}(g)$ for all $g \in \mathrm{SL}(2,\mathbb{C})$.

You can recognize the definition of a **spin structure**. Notice that the existence of a spin structure on a given manifold is a non trivial issue.

pg.spinenphyz

Now a physical field is given by a section of the associated bundle $E = S(M) \times_\rho V$ where ρ is a representations of $\mathrm{SL}(2,\mathbb{C})$ on V. For an electron, it is $V = \mathbb{C}^4$ and $\rho = D^{(1/2,0)} \oplus D^{(0,1/2)}$. That describes a *free* electron is the sense that it doesn't interacts with a gauge field. So in order to write down the formalism in which lives a non zero spin particle, we have to build a $U(1) \times \mathrm{SL}(2,\mathbb{C})$-principal bundle. For this, we follow the procedure given in section 56.14

70.4 Interactions

70.4.1 Spin zero

The general framework is the following:

$$\begin{array}{ccccc} U(1) & \rightsquigarrow & P & & E = P \times_\rho V \\ & & \pi \downarrow & \nearrow^{\sigma_\alpha} & \nearrow_{\phi} \\ & & M & \hookleftarrow & \mathcal{U}_\alpha \end{array}$$

a $U(1)$-principal bundle over a manifold M (as far as topological subtleties are concerned, we suppose $M = \mathbb{R}^4$) and a section ϕ of an associated bundle for a representation ρ of $U(1)$ on V. We consider M with the Lorentzian metric but, since we are intended to treat with scalar (spin zero) fields, we still don't include the Lorentz (or $\mathrm{SL}(2,\mathbb{C})$) group in the picture. We also consider local sections $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$, a connection ω on P and Ω its curvature. We define $A_\alpha = \sigma_\alpha^* \omega$.

Now we particularize ourself to the target space $V = \mathbb{C}$ on which we put the scalar product

$$\langle z_1, z_2 \rangle = \frac{1}{2}(z_1 \overline{z}_2 + z_2 \overline{z}_1), \tag{70.22}$$

EqProdScalVeCU

and the representation [4] $\rho_n\colon U(1) \to GL(\mathbb{C})$,

$$\rho_n(g)z = g \cdot z = g^n z$$

4. We know the list of irreducible representations of $U(1)$ by the proposition 55.32.

where we identify $U(1)$ to the unit circle in $\mathbb{C}$ in order to compute the product. A property of the product (70.22) is to make ρ_n an isometry: for all $g \in U(1)$, $z_1, z_2 \in \mathbb{C}$,

$$\langle \rho_n(g)z_1, \rho_n(g)z_2\rangle = \langle z_1, z_2\rangle.$$

Our first aim is to write the covariant derivative of ϕ with respect to the connection ω. For this we work on the section ϕ under the form $\phi_{(\alpha)}\colon M \to V$ and we use formula (56.354):

$$(D_X\phi)_{(\alpha)}(x) = X_x\phi_{(\alpha)} - \rho_*\big((\sigma_\alpha^*\omega)_x X_x\big)\phi_{(\alpha)}(x). \tag{70.23}$$

Let us study this formula. We know that $(\sigma_\alpha^*\omega)_x = A_\alpha(x) : T_x\mathcal{U}_\alpha \xrightarrow{\sigma} T_{\sigma_\alpha(x)}P \xrightarrow{\omega} u(1)$. Thus $A_\alpha(x)X_x$ is given by a path in $U(1)$; it is this path which is taken by ρ_*. Therefore (we forget some dependences in x)

$$\begin{aligned}\rho_*\big(A_\alpha(x)X_x\big)\phi_{(\alpha)}(x) &= \frac{d}{dt}\Big[\rho_n\big((A_\alpha X)(t)\big)\phi_{(\alpha)}(x)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[(A_\alpha X)(t)^n\Big]_{t=0}\phi_{(\alpha)}(x)\\ &= n\frac{d}{dt}\Big[(A_\alpha X)(t)\Big]_{t=0}\phi_{(\alpha)}(x)\\ &= nA_\alpha(X)\phi_{(\alpha)}(x).\end{aligned} \tag{70.24}$$

Thus the covariant derivative is given by

$$(D_X\phi)_{(\alpha)}(x) = X_x\phi_{(\alpha)} - nA_\alpha(x)(X_x)\phi_{(\alpha)}(x). \tag{70.25}$$

One can guess an electromagnetic coupling for a particle of electric charge n. If this reveals to be physically relevant, it shows that the "electromagnetic identity card" of a particle is given by a representation of $U(1)$. This has to be seen in relation to the discussion on page 3713 where the "type of particle" was closely related to representations of the Lorentz group. It is a remarkable piece of quantum field theory: the properties of a particle are encoded in representations of some symmetry groups.

Now we are going to prove that $\|D\phi\|^2$ is a gauche invariant quantity. The first step is to give a sense to this norm. We consider X_i $(i = 0, 1, 2, 3)$, an orthonormal basis of T_xM and we naturally denote $D_i = D_{X_i}$, $\partial_i = X_i$ and $A_{\alpha i} = A_\alpha(\partial_i)$. Remark that

$$A_\alpha(x)X_x = (\sigma_\alpha^*)_x X_x = \omega(d\sigma_\alpha X_x) = \omega\frac{d}{dt}\Big[\sigma_\alpha(X(t))\Big]_{t=0} \in u(1), \tag{70.26}$$

so this is given by a path in $U(1)$ which can be taken by ρ. Let $c(t)$ be this path, then

$$A_\alpha\phi_{(\alpha)}(x) = \frac{d}{dt}\Big[e^{ic(t)}\phi_{(\alpha)}(x)\Big]_{t=0},$$

so that under the conjugation, $\overline{A_\alpha\phi_{(\alpha)}(x)} = -A_\alpha\overline{\phi}_{(\alpha)}(x)$. Now our definition of $\|D\phi\|^2$ is a composition of the norm on V and the one on T_xM:

$$\|D\phi\|^2 = \eta^{ij}\langle D_i\phi_{(\alpha)}, D_j\phi_{(\alpha)}\rangle \tag{70.27}$$

Using the notation in which the upper indices are contractions with η^{ij}, we have

$$\|D\phi\|^2 = \Big((\partial_i\phi_{(\alpha)})(x) - nA_{\alpha i}\phi_{(\alpha)}(x)\Big)\Big((\partial^i\overline{\phi}_{(\alpha)})(x) + nA_\alpha^i\overline{\phi}_{(\alpha)}(x)\Big).$$

70.4.1.1 Gauge transformation law

A gauge transformation φ is given by an equivariant function $\tilde{\varphi}_\alpha\colon \mathcal{U}_\alpha \to U(1)$ which can be written under the form

$$\tilde{\varphi}_\alpha(x) = e^{i\Lambda(x)}$$

for a certain function $\Lambda\colon \mathcal{U}_\alpha \to \mathbb{R}$. From the general formula (2) of lemma 56.90,

$$(\varphi \cdot \phi)_{(\alpha)}(x) = \rho_n(e^{-i\Lambda(x)})\phi_{(\alpha)}(x) = e^{-ni\Lambda(x)}\phi_{(\alpha)}(x). \tag{70.28}$$

The transformation of the gauche field A is given by equation (56.358). Let us see the meaning of the term $d\tilde{\varphi}$. For $v \in T_x\mathcal{U}_\alpha$,

$$(d\tilde{\varphi}_\alpha)_x v = \frac{d}{dt}\Big[\tilde{\varphi}_\alpha(v(t))\Big]_{t=0} = \frac{d}{dt}\Big[e^{i\Lambda(v(t))}\Big]_{t=0} = i\frac{d}{dt}\Big[\Lambda(v(t))\Big]_{t=0} e^{i\Lambda(v(0))} = i(d\Lambda)_x v e^{i\Lambda(x)}. \tag{70.29}$$

Thus $\tilde{\varphi}_\alpha^{-1}(x)(d\tilde{\varphi}_\alpha)_x = i(d\Lambda)_x$. Since $U(1)$ is abelian, $\tilde{\varphi}^{-1}A\tilde{\varphi} = A$. Finally,

$$(\varphi \cdot A)_\alpha(x) = A_\alpha(x) + i(d\Lambda)_x. \tag{70.30}$$

Now we are able to prove the invariance of $\|D\phi\|^2$. First,

$$(\varphi \cdot A)_{i\alpha}(x) = (\varphi \cdot A)_\alpha(\partial_i) = A_{i_\alpha}(x) + i(\partial_i\Lambda)(x); \tag{70.31}$$

second,

$$\partial_i\left(e^{-ni\Lambda(x)}\phi_{(\alpha)}(x)\right) = -ni(\partial_i\Lambda)(x)\phi_{(\alpha)}(x) + e^{-in\Lambda(x)}(\partial_i\phi_{(\alpha)})(x). \tag{70.32}$$

With these two results,

$$\partial_i(\varphi \cdot \phi)_{(\alpha)}(x) + n(\varphi \cdot A)_{\alpha i}(\varphi \cdot \phi)_{(\alpha)}(x) = e^{-in\Lambda(x)}(nA_{\alpha i}(x) + \partial_i\phi_{(\alpha)}(x)). \tag{70.33}$$

The Yang-Mills **field strength** is given by $F_{(\alpha)} = \sigma_\alpha^*\Omega$ (cf. page 3402). Since $U(1)$ is abelian, $dF_{(\alpha)} = 0$, so that the second pair of Maxwell's equations is complete without any Lagrangian assumptions.

The full Yang-Mills action is written as

$$S(\omega,\phi) = \int_M \left[-\frac{1}{4}F_{(\alpha)ij}F_{(\alpha)}{}^{ij} + \frac{1}{2}\|D\phi\|^2 + \frac{1}{2}m\phi_{(\alpha)}\overline{\phi_{(\alpha)}}\right].$$

The Euler-Lagrange equations are

$$(\partial_i - inA_{\alpha i})(\partial^i - inA_\alpha^i)\phi_\alpha + m^2\phi_\alpha = 0 \tag{70.34a}$$

$$\partial_i F_{(\alpha)}{}^{ij} = 0. \tag{70.34b}$$

So the Yang-Mills Lagrangian only gives the first pair of Maxwell's equations while the second one is given by the geometric nature of fields.

As explained in [980], the topology of the physical space has deep implications on the physics of Yang-Mills equations. The absence of magnetic monopoles for example is ultimately linked to the (simple) connectedness of $\mathbb{R}^4$. When one consider the $U(1)$ Yang-Mills on a sphere, some topological charges appear and magnetic monopoles naturally arise.

70.4.2 Non zero spin formalism

The formalism for a non zero spin particle in an electromagnetic field is described in section 56.14. We consider the spinor bundle

$$\begin{array}{ccc} \mathrm{SL}(2,\mathbb{C}) & \rightsquigarrow & S(M) \\ & & \downarrow p_S \\ & & M \end{array}$$

with the spinor connection on $S(M)$, and ρ_1, a representation of $\mathrm{SL}(2,\mathbb{C})$ on V. For an electron, it is $V = \mathbb{C}^4$ and $\rho_1 = D^{(1/2,0)} \oplus D^{(0,1/2)}$, so for $g_1 \in \mathrm{SL}(2,\mathbb{C})$,

$$\rho_1(g_1)\begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix} = \begin{pmatrix} g_1 & \\ & (\overline{g_1}^t)^{-1} \end{pmatrix}\begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix}. \tag{70.35}$$

On the other hand, we consider the principal bundle

$$\begin{array}{ccc} U(1) & \rightsquigarrow & P \\ & & \downarrow{\scriptstyle p_U} \\ & & M \end{array}$$

with a connection ω_2 which describes the electromagnetic field. As representation $\rho_2 \colon U(1) \to GL(\mathbb{C}^4)$ we choose the coordinatewise multiplication:

$$\rho_2(g_2) \begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix} = \begin{pmatrix} g_2 z_1 \\ \vdots \\ g_2 z_4 \end{pmatrix}. \tag{70.36}$$

The physical picture of the electron is now the principal bundle

$$\begin{array}{ccc} \mathrm{SL}(2,\mathbb{C}) \times U(1) & \rightsquigarrow & S(M) \circ P \\ & & \downarrow{\scriptstyle p} \\ & & M, \end{array}$$

and the field is a section of the associated bundle $(S(M) \circ P) \times_\rho \mathbb{C}^4$.

70.5 Singletons as sections of bundle

SecUKPhZVd

If one wants to see singletons as a field theory, we have to follow the path of section 70.3. So we consider the frame bundle over $M = AdS_4$, and an associated bundle for the scalar singleton representation:

$$\begin{array}{ccc} \mathrm{SO}(2,3) & \rightsquigarrow & L(G/H) \\ & & \pi \downarrow \\ & & M \end{array} \qquad E = L(G/H) \times_\rho V \tag{70.37}$$

where (V, ρ) is the scalar singleton representation of $\mathrm{SO}(2,3)$, and $L(G/H)$ is the frame bundle over AdS_4, as build page 3601.

Chapter 71

Conformal fields theory

71.1 The conformal group

71.1.1 Preliminary discussion

If one looks at [981, 982], one sees that a conformal transformation is a transformation of a (pseudo)riemannian manifold that leaves the metric unchanged up to a positive scalar function: $g'_{\mu\nu}(x') = \Omega(x)g_{\mu\nu}(x)$. Thus we are looking for the maps $\phi\colon M \to M$ that realise that condition.

Since our objective is to do differential geometry, we cannot follow the computations in [981, 982] because there are too much "infinitesimal" in that and we don't understand anything [1].

In order to determine the Poincaré group that leaves the metric invariant, we were fortunate because of theorem 8.157 that ensured linearity of the map ϕ. Our search for the Poincaré group [2] was thus simplified by two circumstances:

— ϕ and $d\phi$ are the same.

— the condition $g_{\mu\nu} = g'_{\mu\nu}$ does not involve a specific point, so that we had not to ask ourself questions about the "base point" of the vectors.

Here we are in a more complicated situation.

Definition 71.1.
Let (M, g) be a (pseudo)riemannian manifold. A conformal map will be $\phi\colon M \to M$ such that

$$g_{\phi(x)}\big(d\phi_x v, d\phi_x w\big) = \Omega(x)g_x(v, w) \tag{71.1}$$

EQooCCMMooXVbTAd

for a function $\Omega \in C^\infty(M)$. The condition (71.1) has to hold for every $x \in M$ and $v, w \in T_xM$.

We particularise ourself to the case where the manifold M is a vector space V of dimension d with constant metric η. If $\{e_i\}_{i=1,\dots,d}$ is a basis of V, we consider the same basis on each T_xV and for $v, w \in T_xV$ we have

$$v \cdot w = \sum_{ij} \eta_{ij} v_i w_j. \tag{71.2}$$

All the sums are intended from 1 to d.

Definition 71.2.
*A **conformal map** is a C^∞ map $\phi\colon V \to V$ for which there exists a function $\Omega \in C^\infty(V)$ satisfying*

$$v \cdot w = \Omega(x) d\phi_x(v) \cdot d\phi_x(w) \tag{71.3}$$

EQooOZDUooCDaIrb

DEFooVKNBooFBWQQM

for every $x \in V$ and every $v, w \in V_x$.

1. These books are very okay, but as far as we are interested in bundles, there are some works to do. As far as the *calculations* are concerned, we will follow them.
2. By the way given by theorem 18.79.

NorooVEVOooRBpvXF

71.3.
This is not the group of "angles preserving" transformations, but rather the group of *locally* preserving the angles.

In order to understand that, let us imagine a deforming material. A vector at point A is an arrow joining point A to a point B. The image of the vector $\overrightarrow{AB}$ has to be $\overrightarrow{\phi(A)\phi(B)}$ when the material is deformed. Thus instead of moving the vector by $d\phi$, we should move both extremities by ϕ itself. We could speak about affine spaces as described around definition 11.7, but instead we will describe our subject with a vector bundle.

We will discuss this point in 71.1.4.

Remarque 71.4.
We are going not to determine all the conformal transformations [3], but the (connected to the identity) Lie group of conformal transformations. That is: we are searching for a Lie group G of diffeomorphisms $M \to M$ satisfying the condition (71.1).

We will use the strategy presented in 53.71, as in the proof of proposition 53.73

Since G is a Lie group, each element can be written under the form $g = e^X$ for some $X \in \mathfrak{g}$ (here $\mathfrak{g}$ is the Lie algebra, that is T_eG). Following the definition 58.6 we define $X : V \to V$ by

$$X(v) = \frac{d}{dt}\Big[\exp(-tX)(v)\Big]_{t=0} \tag{71.4}$$

The exponential here is the one from the Lie algebra to the Lie group. We also define

$$\phi_t(x) = \frac{d}{dt}\Big[e^{-tX}x\Big]_{t=0} \tag{71.5}$$

for $t \in \mathbb{R}$ and $x \in V$ (t being restricted to an open set around 0).

71.1.2 Generators

We consider the case in which the metric is flat an $g_x = \eta$ for every x. The starting point is the condition (71.3). We write it for the map

$$\phi_t(x) = e^{-tX}(x) \tag{71.6}$$

and we search the function $X : V \to V$. We derive with respect to t the equality (for fixed $x \in V$)

$$(d\phi_t)_x(v) \cdot (d\phi_t)_x(w) = \Omega_t(x)v \cdot w. \tag{71.7}$$

EQooJZDTooVJEUvo

Permuting two derivatives, we have

$$\begin{aligned}
\frac{d}{dt}\Big[(d\phi_t)_x(v)\Big]_{t=0} &= \frac{d}{dt}\Big[\frac{d}{du}\Big[\phi_t(x+uv)\Big]_{u=0}\Big]_{t=0} && (71.8a)\\
&= \frac{d}{du}\Big[\frac{d}{dt}\Big[\phi_t(x+uv)\Big]_{t=0}\Big]_{u=0} && (71.8b)\\
&= \frac{d}{du}\Big[\frac{d}{dt}\Big[e^{-tX}(x+uv)\Big]_{t=0}\Big]_{u=0} && (71.8c)\\
&= \frac{d}{du}\Big[-X(x+uv)\Big]_{u=0} && (71.8d)\\
&= -dX_x(v). && (71.8e)
\end{aligned}$$

If $t = 0$ we also have

$$(d\phi_0)_x(w) = d(\mathrm{Id})_x(w) = w. \tag{71.9}$$

3. Can you answer the question http://math.stackexchange.com/questions/1549670/rigorous-definition-of-a-generator-for-a-transformation-group?

Thus the derivative of (71.7) produces the following differential equation for X:

$$dX_x(v) \cdot w + dX_x(w) \cdot v = \omega(x) v \cdot w. \tag{71.10}$$

Let $\{e_i\}$ be a basis of V. We have $dX_x(e_i) = \frac{\partial X}{\partial x_i}(x)$ and

$$dX_x(e_i) \cdot e_j = \sum_{kl} \big(dX_x(e_i)\big)(e_j)_l \eta_{kl} = \sum_k \eta_{kj} \frac{\partial X_k}{\partial x_i} = \partial_i \tilde{X}_j \tag{71.11}$$

where $\tilde{X} = \eta X$, this is $\tilde{X}_i = \sum_k \eta_{ik} X_k$. The equation for X, if $v = e_i$ and $w = e_j$ is

$$\frac{\partial \tilde{X}_j}{\partial x_i} + \frac{\partial \tilde{X}_i}{\partial x_j} = \omega(x) \eta_{ij}. \tag{71.12}$$

EQooAKZWooQVIdur

Lemma 71.5 ([982])**.**
Let V be a vector space of dimension $d \geqslant 3$ endowed with the constant metric η. The solutions of the differential equation

$$\partial_i f_j + \partial_j f_i = \omega(x) \eta_{ij} \tag{71.13}$$

EQooVBIMooQBKAKQ

for a function $f : V \to V$ of class C^∞ are

$$f(x) = a + \sum_{km} b_{km} x_k e_m + \sum_{klm} q_{klm} x_k x_l e_m \tag{71.14}$$

EQooEFOMooUhcgfT

where

(1) $a \in V$ is any vector.

(2) the antisymmetric part of b is free

(3) the symmetric part of b is proportional to η.

(4) $q_{ijk} = -d_j \eta_{ik} + d_k \eta_{ij} - d_i \eta_{jk}$ and d is arbitrary.

In this case, ω is given by $\omega(x) = \frac{2}{d}(\nabla \cdot f)(x)$ where the divergence is the one of proposition 56.72.

Proof. We can initiate our work by expressing ω in terms of f. For that, multiply both sides by $(\eta^{-1})_{ij}$ and make the sum over i and j. On the right hand side we have

$$\omega(x) \sum_i j (\eta^{-1})_{ij} \eta_{ij} = \omega(x) \sum_j (\eta^{-1}\eta)_{jj} = \omega(x) d. \tag{71.15}$$

On the left hand side we have (reaming the summation indices $i \leftrightarrow j$ in one of the two terms):

$$2 \sum_{ij} (\eta^{-1})_{ij} \frac{\partial f_j}{\partial x_i} = 2 \nabla \cdot f. \tag{71.16}$$

For the divergence, see the proposition 56.72. Thus $\omega(x) = \frac{2}{d} \nabla \cdot f$ and the equation is

$$\partial_i f_j + \partial_j f_i = \frac{2}{d} (\nabla \cdot f) \eta_{ij}. \tag{71.17}$$

EQooAPOPooBdKskD

We apply on both sides the operator [4]

$$\sum_{ijkl} (\eta^{-1})_{ik} (\eta^{-1})_{jl} \partial_k \partial_l \tag{71.18}$$

which is nothing else than the contraction with $\partial^i \partial^j$ when one uses the uppper-lower index notation. The first term is

$$\sum_{ijkl} (\eta^{-1})_{ik} (\eta^{-1})_{jl} \partial_k \partial_l \partial_i f_j = \sum_{jl} (\eta^{-1})_{jl} \partial_l (\Box f_j) = \nabla \cdot (\Box f) \tag{71.19}$$

4. Since the function f is supposed to be of class C^∞, we derivatives commute.

where we introduced the operator

$$\Box = \sum_{kl} (\eta^{-1})_{kl} \partial_k \partial_l. \tag{71.20}$$

The second term gives the same result. On the right hand side,

$$\frac{2}{d} \sum_{ijkl} (\eta^{-1})_{ik} (\eta^{-1})_{jl} \partial_k \partial_l \eta_{ij} \nabla \cdot f = \frac{2}{d} \sum_{kl} (\eta^{-1})_{kl} \partial_k \partial_l \nabla \cdot f = \frac{2}{d} \nabla \cdot \Box f. \tag{71.21}$$

If $d \neq 1$ we are left with [5]

$$\Box(\nabla \cdot f) = 0. \tag{71.22}$$

Now we go back to equation (71.17).

$$\underbrace{\partial_i f_j}_{A} + \underbrace{\partial_j f_i}_{B} = \underbrace{\frac{2}{d} (\nabla \cdot f) \eta_{ij}}_{C}. \tag{71.23}$$

We apply the operator

$$\sum_{im} \partial_l (\eta^{-1})_{im} \partial_m, \tag{71.24}$$

which is the contraction with $\partial_l \partial^i$.

On A, we recognize the $\Box$ operator:

$$\partial_l \Box f_j. \tag{71.25}$$

On B we recognize the divergence operator:

$$\partial_j \partial_l \nabla \cdot f. \tag{71.26}$$

On C we have $\eta^{-1}\eta = \mathbb{1}$ and

$$\frac{2}{d} \partial_l \partial_j \nabla \cdot f. \tag{71.27}$$

Putting the whole together,

$$\partial_l \Box f_j + \big(1 - \frac{2}{d}\big) \partial_j \partial_l \nabla \cdot f = 0. \tag{71.28}$$

EQooCHJAooVzpeos

Let us write the same equation with $l \leftrightarrow j$:

$$\partial_j \Box f_l + \big(1 - \frac{2}{d}\big) \partial_l \partial_j \nabla \cdot f = 0. \tag{71.29}$$

EQooSMQVooBqKydQ

We sum (71.28) with (71.29) and use the equation (71.17):

$$\Box\big(\underbrace{\partial_l f_j + \partial_j f_l}_{=\frac{2}{d}\eta_{lj} \nabla \cdot f}\big) + 2\big(1 - \frac{2}{d}\big) \partial_l \partial_j \nabla \cdot f = 0. \tag{71.30}$$

Using the fact that $\Box \nabla \cdot f = 0$, the first term vanishes. If $d \neq 2$ we have

$$\partial_l \partial æ \nabla \cdot f = 0. \tag{71.31}$$

Applying the operation $\partial_k \partial_l$ on the equation (71.17),

$$\partial_i \partial_k \partial_l \nabla \cdot f_j + \partial_k \partial_l \partial_j f_i = \frac{2}{d} \eta_{ij} \underbrace{\partial_k \partial_l \nabla \cdot f}_{=0} = 0 \tag{71.32}$$

If we write $\Delta_{ijkl} = \partial_i \partial_k \partial_k f_l$ we have

$$\Delta_{ijkl} = -\Delta_{ijlk} \tag{71.33}$$

EQooFBXSooLncZtv

5. By the way, if $d = 1$, the question about the angle preserving diffeomorphisms does not make much sense.

while Δ is symmetric with respect to its first three indices. Starting from Δ_{ijlk} we permute the two inner indices and apply (71.33); the whole two times:

$$\Delta_{i\,jk\,l} = \Delta_{i\,kj\,l} = -\Delta_{ikl\,j} = -\Delta_{il\,kj} = \Delta_{il\,jk} = \Delta_{ijlk}. \tag{71.34}$$

Thus $\Delta_{ijlk} = 0$, and

$$\partial_i\partial_j\partial_k f_l = 0 \tag{71.35}$$

for every i, j, k, l. Thus each component of the function f is a polynomial of degree 2 with respect to x. This leaves us with

$$f(x) = a + \sum_{km} b_{km} x_k e_m + \sum_{klm} q_{klm} x_k x_l e_m. \tag{71.36}$$

We get constraints on a, b, q putting that solution into the equation (71.13) with $\omega(x) = \frac{2}{d}\nabla \cdot f$. Using the fact that $\partial_i e_m = \delta_{im}$ we get

$$\partial_i f_j = \sum_k b_{kj}\delta_{ki} + \sum_{kl} q_{klj}(\delta_{ki}x_l + x_k\delta_{li}) = b_{ij} + 2\sum_k q_{kij}x_k, \tag{71.37}$$

and

$$\nabla \cdot f = \sum_{ij}(\eta^{-1})_{ij}b_{ij} + 2\sum_{ijk} q_{kij}(\eta^{-1})_{ij}x_k = \operatorname{Tr}(b\eta^{-1}) + 2\sum_{ijk} q_{kij}(\eta^{-1})_{ij}x_k. \tag{71.38}$$

Putting all together we have the condition

$$b_{ij} + b_{ji} + 2\sum_k q_{kij}x_k + 2\sum_k q_{kji}x_k = \frac{2}{d}\eta_{ij}\big(\operatorname{Tr}(b\eta^{-1}) + 2\sum_{klm} q_{klm}(\eta^{-1})_{lm}x_k\big). \tag{71.39}$$

Since that condition has to hold for every choice of x we can separate the part with and without x.

(1) **Part without x** We have

$$b_{ij} + b_{ji} = \frac{2}{d}\eta_{ij}\operatorname{Tr}(b\eta^{-1}). \tag{71.40}$$

The antisymmetric part of b gives zero on left, but also zero on right because η^{-1} is symmetric. Let us be more specific; if A is antisymmetric and S be symmetric. Then $(AS)^t = -SA$ and $\operatorname{Tr}(AS) = \operatorname{Tr}\big((AS)^t\big) = -\operatorname{Tr}(SA) = -\operatorname{Tr}(AS)$ by the cyclic invariance of the trace. We conclude that the antisymmetric part of b is free.

The symmetric part of b is proportional to η, and the coefficient of proportionality is $\frac{2}{d}\operatorname{Tr}(b\eta^{-1})$. We can check that this is not a new constraint by putting directly $b = \alpha\eta$ into the equation:

$$\alpha\eta_{ij} + \alpha\eta_{ji} = \frac{2}{d}\eta_{ij}\operatorname{Tr}(\alpha\eta\eta^{-1}), \tag{71.41}$$

after simplifications and taking into account $\operatorname{Tr}(\alpha) = \alpha d$, we see that it is valid for every α.

(2) **Path with x**

We can factorize x_k and write the condition as

$$2q_{kij} + 2q_{kji} = \frac{4}{d}\eta_{ij}\sum_{lm}(\eta^{-1})_{lm}q_{klm}. \tag{71.42}$$

Thus there exists a vector d such that $q_{ijk} = -q_{ikj} - 2d_i\eta_{jk}$ (the coefficient 2 is for later convenience). Maybe each vector d will not produce a solution; we'll have to check it later. Using three times the symmetry of q with respect to the first two indices and the latter formula,

$$\begin{aligned}
q_{kij} &= -q_{kji} - 2d_k\eta_{ij} && (71.43a)\\
&= -q_{jki} - 2d_k\eta_{ij} && (71.43b)\\
&= q_{jik} + 2d_j\eta_{ik} - 2d_k\eta_{ij} && (71.43c)\\
&= q_{ijk} + 2d_j\eta_{ik} - 2d_k\eta_{ij} && (71.43d)\\
&= -q_{ikj} - 2d_i\eta_{kj} + 2d_j\eta_{ik} - 2d_k\eta_{ij}. && (71.43e)
\end{aligned}$$

Thus

$$q_{ijk} = -d_j\eta_{ik} + d_k\eta_{ij} - d_i\eta_{jk}. \tag{71.44}$$

EQooKZVTooWSMmyM

At this point we proved that every solution of (71.13) are of the given form (71.14) with the constrains. We still have to check that for every choice of antisymmetric b and every choice of d (giving q by the formula (71.44)), the corresponding function actually is a solution.

(1) **Antisymmetric** Let us check for $f(x) = \sum_{km} b_{km}x_k e_m$ when b is antisymmetric. We have

$$\partial_i f_j = \sum_k b_{kj} \underbrace{\partial_i x_k}_{=\delta_{ik}} = b_{ij} \tag{71.45}$$

and

$$\nabla \cdot f = \sum_{kl} (\eta^{-1})_{kl}\partial_k f_l = \sum_{kl} (\eta^{-1})_{kl} b_{kl} = 0. \tag{71.46}$$

The lase zero is because of the contraction of a symmetric matrix (η^{-1}) with an antisymmetric matrix. Now $\partial_i f_j + \partial_j f_i = b_{ij} + b_{ji} = 0$ and the equation is satisfied.

(2) **The quadratic part** We have to check that the function

$$f(x) = \sum_{klm} \big(- d_l\eta_{km} + d_m\eta_{kl} - d_k\eta_{lm}\big)x_k x_l e m \tag{71.47}$$

is a solution of $\partial_i f_j + \partial_j f_i = \frac{2}{d}\eta_{ij}\nabla \cdot f$. Few computations provide on the one hand

$$\partial_i f = 2\sum_k \big(- d_i\eta_{kj} + d_j\eta_{ki} - d_k\eta_{ij}\big)x_k, \tag{71.48}$$

and on the other hand,

$$\nabla \cdot f = -2d\sum_k d_k x_k. \tag{71.49}$$

A few more computations show that the equation is satisfied for every choice of vector d.

□

71.1.3 Exponentiation

So we have $X(x) = \eta^{-1}\tilde{X}(x)$ with

$$\tilde{X}(x) = a + \sum_{km} b_{km}x_k e_m + \sum_{klm} q_{klm}x_k x_l e_m. \tag{71.50}$$

We are now going to exponentiate each part of that "infinitesimal" transformation.

Translation For the constant part we have $X(x) = \eta^{-1}a$ and up to redefinition of a we suppose $X(x) = a$. Thus

$$e^{tX}(x) = \sum_{k=0}^{\infty} = x + \sum_{k=1}^{\infty} \frac{t^k a}{k!} = x + (e^t - 1)a. \tag{71.51}$$

One can check that with $\phi_t(x) = x + (e^t - 1)a$ we have $\frac{d}{dt}\Big[\phi_t(x)\Big]_{t=0} = a$.

The corresponding global transformation is the translation ϕ_1, that is $\phi(x) = x + (e - 1)a$. This is an arbitrary translation.

Dilatation For the symmetric part of b we have

$$\tilde{X} = \sum_{km} (\alpha\eta)_{km}x_k e_m \tag{71.52}$$

and then

$$X(x) = \alpha\sum_{kml} \eta_{km}x_k(\eta^{-1})_{ml}e_l = \alpha\sum_l x_l e_l = \alpha x. \tag{71.53}$$

Thus $X(x) = \alpha x$ for every k (included $k = 0$) and

$$e^{tX}(x) = \sum_{k=0}^{\infty} \frac{t^k \alpha^k}{k!} x = e^{t\alpha} x. \tag{71.54}$$

The corresponding global transformation is

$$\phi(x) = e^{\alpha} x, \tag{71.55}$$

that is a dilatation with positive coefficient.

Rotation For the antisymmetric part of b we have $\tilde{X}(x) = bx$, that is

$$X(x) = \eta^{-1} b x. \tag{71.56}$$

We check that e^{tX} is an η-isometry, that is

$$e^{tX}(x) \cdot e^{tX}(y) = x \cdot y. \tag{71.57}$$

That equality is obviously true with $t = 0$. For checking other values of t, we take the derivative with respect to t and we check that this is zero:

$$X(x) \cdot y + x \cdot X(y) \stackrel{?}{=} 0, \tag{71.58}$$

that is

$$\eta^{-1} b x \cdot y \stackrel{?}{=} -x \cdot \eta^{-1} b y. \tag{71.59}$$

EQooGKSSooBxzPzx

Before to make the (simple) computation, we could want to use something like $\langle Ax, y\rangle = \langle x, A^t y\rangle$. Since the product here is given by the metric η, the notion of "antisymmetric" has to be given by $\eta^{-1}A$ where A is a regular antisymmetric matrix, and search for some notion of η-transposition. But ... We know from 8.188 that the notion of transpose of a linear map is badly defined. So let's make the computation by brute force.

Computing the two sides of (71.59) with the full indices techniques we find

$$\eta^{-1} b x \cdot y = \sum_{ij} b_{ij} x_j y_i \tag{71.60a}$$

$$x \cdot \eta^{-1} b y = \sum_{ij} b_{ij} x_i y_j, \tag{71.60b}$$

and the yes we have

$$X(x) \cdot y + x \cdot X(y) = 0. \tag{71.61}$$

In the global conformal group, the infinitesimal transformation

$$X(x) = \eta^{-1} b x \tag{71.62}$$

with an antisymmetric matrix b correspond to the global isometries of η.

Special conformal transformations We work on the exponentiation of the transformation

$$\tilde{X}(x) = \sum_{ijk} q_{ijk} x_i x_j e_k = \sum_{ijk} (-d_j \eta_{ik} + d_k \eta_{ij} - d_j \eta_{jk}) x_i x_j e_k. \tag{71.63}$$

The first work is to write down $X(x) = \eta^{-1}\tilde{X}(x)$, that is to replace e_k by $\sum_l (\eta^{-1})_{lk} e_k$. We have :

$$X(x) = -2 \sum_{ij} d_i x_i x_j e_j + \sum_{kl} d_k (\eta^{-1})_{kl} e_l (x \cdot x). \tag{71.64}$$

Now we define $d' = \eta^{-1} d$ and we write

$$X(x) = -2(d' \cdot x)x + x^2 d' \tag{71.65}$$

where x^2 stands for $x \cdot x$ (which is not specially positive since the metric η is not positive defined). Notice that the vector d' is arbitrary in the sens that, since d can take any value, d' can also take any value. For that reason, we will drop the prime.

In order to determine $e^X(x)$ we will solve the differential equation for $x(t) = e^{tX}(x_0)$ (with $x(0) = x_0$). We have:

$$x'(s) = \frac{d}{dt}\Big[e^{tX}(x_0)\Big]_{t=s} = \frac{d}{dt}\Big[e^{(t-s)X}e^{sX(x_0)}\Big]_{t=s} \tag{71.66a}$$

$$= \frac{d}{dt}\Big[e^{(t-s)X}x(s)\Big]_{t=s} \tag{71.66b}$$

$$= \frac{d}{du}\Big[e^{uX}x(s)\Big]_{u=0} \tag{71.66c}$$

$$= X\big(x(s)\big). \tag{71.66d}$$

Thus for every t we have

$$x'(t) = X\big(x(t)\big) \tag{71.67}$$

and dropping the dependence in t on x and x' the differential equation to solve reads[983]

$$x' = -2(d \cdot x)x + (x \cdot x)d. \tag{71.68}$$

In order to solve it we pose $y(t) = \frac{x(t)}{x(t) \cdot x(t)}$. We have

$$y' = \frac{x'x^2 - x(2x \cdot x')}{(x \cdot x)^2} \tag{71.69}$$

Then we substitute x' by $-2d \cdot x + x^2 d$ in the right hand side and after some computations we get

$$y'(t) = d. \tag{71.70}$$

Thus $y'(t) = y_0 + td$ and knowing that $y_0 = y(0) = \frac{x_0}{x_0^2}$ we can write

$$\frac{x}{x^2} = y = \frac{x_0}{x_0^2} + td. \tag{71.71}$$

We are left with the algebraic equation

$$\frac{x}{x^2} = A \tag{71.72}$$

with $x, A \in (V, \eta)$. Since x^2 in only a numerical coefficient we have $x = \lambda A$ for some $\lambda \in \mathbb{C}$. Replacing:

$$\frac{A}{A^2} = \lambda A \tag{71.73}$$

and thus $\lambda = \frac{1}{A^2}$, so that

$$x = \frac{A}{A^2}. \tag{71.74}$$

In our case, $A = \frac{x_0}{x_0^2 + td}$, and few computations provide

$$x(t) = \frac{x_0 + tx_0^2 d}{1 + tx_0 \cdot d + t^2 x_0^2 d^2}. \tag{71.75}$$

As fas as $x_0^2 \neq 0$, this is the integral curve of the special conformal transformation of parameter d. It always exists on an open set around $t = 0$.

The generic form of the special conformal transformation is got with $t = 1$, which provides

$$e^X(x_0) = \frac{x_0 + x_0^2 d}{1 + x_0 \cdot d + x_0^2 d^2}. \tag{71.76}$$

EQooPMIGooVLXpYt

The whole derivation of the special conformal transformation (71.76) raises the question of the well-definiteness. Let us provide an answer in the euclidian case. We suppose that η is strictly positive defined. Let (x_0, d) such that the denominator in (71.76) is negative. Taking a rotation R that brings x_0 to e_1 (in the canonical basis). The value of $x_0 \cdot d$ and $x_0^2 d^2$ are not changed if we substitute x_0 by Rx_0 and d by Rd. Thus we have a vector d such that with $(x_0, d) = (e_1, d)$, the denominator is zero.o

We have $1 + x_0 \cdot d + x_0^2 d^2 = 1 + d_1 + d^2$, with $b^2 = b_1^2 + \dots b_n^2 > |b_1|$. Thus the denominator is not only positive but larger than 1. This shows (by contradiction) that in the euclidian case, the expression (71.76) is always well defined.

71.1.4 What about really preserving the angles?

sebsecooCBKEooQOWqFo

We continue the discussion of 71.3. If one wants a map $\phi\colon M \to M$ that preserves the angles in the sense that for any three points $A, B, C \in M$ the angle formed by $\overrightarrow{AB}$ and $\overrightarrow{AC}$ is the same as the angle formed by the vectors $\overrightarrow{\phi(A)\phi(B)}$ and $\overrightarrow{\phi(A)\phi(C)}$, we need more structure, since there are no notion of "vector from one point to another" in a general manifold. So we particularize ourself to the case in which M is a vector space. Thus we can write differences like $\phi(x) - \phi(y)$.

Let V be a finite dimensional vector space and $E = V \times V$ be the trivial vector bundle with fibre V. We denote $V_x = \{(x, v) \text{ tel que } v \in V\}$ and for each $x \in V$ we have a non-degenerate bilinear form $g_x\colon V \to V$. We define

$$g\big((x, v), (x, w)\big) = g_x(v, w) \tag{71.77}$$

and we will often directly write $g_x(v, w)$ or $v \cdot w$ when v, w belong to V_x instead of V.

A map $\phi\colon V \to V$ also acts on the vector bundle as

$$\phi(x, v) = \big(\phi(x), \phi(x + v) - \phi(x)\big). \tag{71.78}$$ EQooFICEooQACFoU

This way to act translates the fact that for a vector, we displace the ending point as well as the starting point with ϕ. This is not the same as displacing the vector by $d\phi_x$.

Remarque 71.6.
We consider a vector bundle with fibre V in order to make sense to the definition (71.78). One can also working on the more intuitive setting of the tangent bundle TM, but in this case, on has to deal with a linear bijection between T_xM and V.

With no abuse of notations, the condition (71.3) reads, for $v, w \in V$:

$$g_x(v, w) = \Omega(x) g_{\phi(x)}\big(\phi(x + v) - \phi(x), \phi(x + w) - \phi(x)\big). \tag{71.79}$$ EQooFZUFooTGWpBa

71.2 Two dimensional conformal group

Let us consider the euclidian two dimensional space, that is $d = 2$ and $\eta = \delta$. The equations (71.12) reduce to

$$\begin{cases} \partial_1 X_1 = \partial_2 X_2 & \text{(71.80a)} \\ \partial_1 X_2 = -\partial_2 X_1 & \text{(71.80b)} \end{cases}$$

that are the Cauchy-Riemann equations (26.2). Theorem 26.2 says that the map $Z\colon \mathbb{C} \to \mathbb{C}$ defined by

$$Z(x + iy) = X_1(x, y) + iX_2(x, y) \tag{71.81}$$

is holomorphic.

For the sake of generality, let us note that Z has to be holomorphic on the domain where we want the global transformation to be conformal: a differential equation is something local. Thus, if we want to check also for transformations that are only conformal on an open set, we want to

study maps Z that are holomorphic on an open set. We can suppose Z to be meromorphic outside that open set. Thus we say that the infinitesimal generators of the conformal transformations are meromorphic maps and can be written as [6]

$$Z(z) = \sum_{k\in\mathbb{Z}} a_k z^k. \tag{71.82}$$

Remarque 71.7.
The map Z has to be seen as the "component" maps of the vector field that is tangent to a global transformation. That is the vector field which is everywhere tangent to the integral curves.

NORMooHDLPooQBfEif

71.8.
The map Z being seen as an element of the Lie algebra of the conformal group [7], it acts on an holomorphic [8] function $f\colon \mathbb{C}\to\mathbb{C}$ by the definition 58.6 and its formula (58.12):

$$Z_{z_0}(f) = -df_{z_0}\big(Z(z_0)\big) = -Z(z_0)(\partial_z f)(z_0) \tag{71.83}$$

where we used the proposition 26.4.

Thus for each meromorphic function Z we have a generator of the conformal group that is written as $Z(z)\partial_z$. But we know a basis of the meromorphic functions (at least the ones that have only one pole at zero): these are the functions $l_n = z^n$ with $n\in\mathbb{Z}$. So we set

$$l_n = -z^{n+1}\partial_z. \tag{71.84}$$

These operators are a basis of the Lie algebra of conformal transformations that are defined on an open set $B(0,r)\backslash\{0\}$. The commutation relations are easy to compute:

$$[l_m, l_n] = (m-n)l_{m+n}. \tag{71.85}$$

Let us determine what are the operators l_n which are globally defined[984]. First the operator $z^k\partial_z$ has a pole at $z=0$ for every $k<0$. In order to determine the values of k for which $z^k\partial_z$ has a pole at $z=\infty$ we follow the definition 48.60 [9] and perform the change of variable $w=1/z$.

Let U be a neighbourhood of 0 in $\mathbb{C}$ and consider the chart

$$\begin{aligned} z\colon U &\to \hat{\mathbb{C}} \\ w &\mapsto \frac{1}{w}. \end{aligned} \tag{71.86}$$

This is a chart for a neighbourhood of ∞. The definition of ∂_w is $\partial_w f\big(z(w_0)\big) = \frac{d}{dt}\Big[f\big(z(w_0+t)\big)\Big]_{t=0}$. Here the limit is with $t\in U$. Let us compute, using the fact that the differential of an holomorphic function f is the multiplication by ∂_z (proposition 26.4) EQooOCHQooFVHtYV

$$\frac{\partial f}{\partial w}\big(z(w_0)\big) = \frac{d}{dt}\Big[f\big(z(w_0+t)\big)\Big]_{t=0} \tag{71.87a}$$
$$= df_{z(w_0)}z'(w_0) \tag{71.87b}$$
$$= (\partial_z f)(\frac{1}{w_0})\frac{-1}{w_0^2} \tag{71.87c}$$

Thus

$$\frac{\partial f}{\partial z}\big(z(w_0)\big) = -w_0^2\frac{\partial f}{\partial w}\big(z(w_0)\big). \tag{71.88}$$

In that sense we write

$$\partial_w = -z^2\partial_z. \tag{71.89}$$

6. Theorem 28.26.
7. The extended conformal group in the sense that some of the transformations are not globally well defined.
8. In the context of complex manifold we are dealing with holomorphic functions like we are dealing with C^∞ function in real differential geometry?
9. We follow in fact the *idea* of that definition, but we should have a more precise definition of an holomorphic vector field and a pole of a vector field.

Remarque 71.9.
This computation (71.87) is nothing else than the usual chain formula

$$\frac{\partial f}{\partial w} = \frac{\partial f}{\partial w}\frac{\partial w}{\partial z}. \tag{71.90}$$

We have

$$z^k(\partial_z f) = z^k \frac{-1}{z^2}\partial_w f \tag{71.91a}$$
$$= -z^{k-2}\partial_w f \tag{71.91b}$$
$$= -w^{2-k}\partial_w f. \tag{71.91c}$$

For the limit $w \to 0$ to exist we need $k \leqslant 2$.

At the end of the day, $z^k partial_z$ has no pole at 0 and ∞ when $k = 0, 1, 2$. The operators l_n are thus well defined with

$$n = -1, 0, 1. \tag{71.92}$$

Chapter 72

Banach and C^*-algebras

The main references for this chapter are [959, 489]. In this chapter, all algebras are over $\mathbb{C}$, or $\mathbb{R}$ when it is mentioned. Definition and spectral properties of Banach algebras are given in chapter 60.3.

72.1 Commutative Banach algebra

We suppose the Banach algebra $\mathfrak{A}$ to be commutative.

72.1.1 Structure space

DefStructureSpaceDel

Definition 72.1.
*The **structure space** $\Delta(\mathfrak{A})$ of a commutative algebra is the set of the nonzero linear maps $\omega\colon \mathfrak{A} \to \mathbb{C}$ such that $\forall A, B \in \mathfrak{A}$,*

$$\omega(AB) = \omega(A)\omega(B).$$

*We say that an element of this space is a **character**, or a **multiplicative** map of $\mathfrak{A}$.*

Proposition 72.2.
Let $\mathfrak{A}$ be an unital commutative Banach algebra. Then for any $\omega \in \Delta(\mathfrak{A})$,

(1) $\omega(\mathbb{1}) = 1$.

(2) the character ω is bounded (and then continuous from 60.45) with norm $\|\omega\| = 1$ and for all $A \in \mathfrak{A}$,

$$\|\omega(A)\| \leqslant \|A\|. \tag{72.1}$$

eq:omAleqnA

Proof. The first claim is obvious because $\omega(A) = \omega(\mathbb{1}A) = \omega(\mathbb{1})\omega(A)$. For the second one, we know from lemma 60.56 that $(A - z)$ is invertible when $|z| > \|A\|$. By linearity,

$$\omega(A - z) = \omega(A) - z \neq 0$$

because ω in a homomorphism. Now remark that $A - z$ is invertible implies $|\omega(A)| \neq |z|$. Indeed let us suppose the opposite, then there exists a $\alpha \in \mathbb{R}$ such that $\omega(A) = e^{i\alpha}z$, but $|e^{i\alpha}z| = |z|$. Conclusion: if $|z| > \|A\|$, then $|\omega(A)| \neq |z|$. This immediately yields $|\omega(A)| \leqslant \|A\|$.

From there, it is clear that $\|\omega\| = 1$ because the norm is the supremum of $|\omega(A)|$ with $\|A\| = 1$. Since $\omega(\mathbb{1}) = 1$, $\|\omega\| \geqslant 1$, but what we just showed implies $\|\omega\| \leqslant 1$.

□

Theorem 72.3.
Let $\mathfrak{A}$ be an unital commutative Banach algebra. Then we have a bijection between $\Delta(\mathfrak{A})$ and the set of maximal ideals in $\mathfrak{A}$. More precisely,

(1) $\ker(\omega)$ *is an ideal,* enuei

(2) $\omega_1 = \omega_2$ *if and only if* $\ker\omega_1 = \ker\omega_2$, enueii

(3) each maximal ideal is the kernel of an element in $\Delta(\mathfrak{A})$. enueiii

tho:ideal_kernel

Proof. (1) Since ω is continuous, the set $\ker(\omega)$ is closed. It is also clear that is $Z \in \ker(\omega)$, then $AZ \in \ker(\omega)$ for all $Z \in \mathfrak{A}$ because ω is multiplicative. Then $\ker(\omega)$ is an ideal. In order to see that it is a maximal ideal, remark that $\omega(X) = 0$ is a linear equation which describe a vector subspace of $\mathfrak{A}$ of codimension 1. pg_codimun

(2) In any vector space, $\ker\omega_1 = \ker\omega_2$ implies that ω_1 and ω_2 are multiples each others. In the case of $\Delta(\mathfrak{A})$, this in turn implies the equality.

(3) Let $\mathfrak{I} \neq \mathfrak{A}$ be a maximal ideal and $B \neq 0$ outside $\mathfrak{I}$. Consider

$$\mathfrak{I}_B = \{BA + J \text{ tel que } A \in \mathfrak{A} \text{ and } J \in \mathfrak{I}\}.$$

By construction it is a left-ideal and by commutativity of $\mathfrak{A}$, it is an ideal. We have $\mathfrak{I} \subsetneq \mathfrak{I}_B$. Since $\mathfrak{I}$ is maximal, the conclusion is $\mathfrak{I}_B = \mathfrak{A}$. In particular $\mathbb{1} = BA + J$ for a suitable choice of $A \in \mathfrak{A}$ and $J \in \mathfrak{I}$. For these,

$$\tau(\mathbb{1}) = \tau(BA) = \tau(B)\tau(A),$$

but B is arbitrary. Then any element of $\mathfrak{A}/\mathfrak{I}$ is invertible and the Gelfand-Mazur theorem (corollary 60.61) concludes $\mathfrak{A}/\mathfrak{I} \simeq \mathbb{C}$. Let $\psi\colon \mathfrak{A}/\mathfrak{I} \to \mathbb{C}$ be the isomorphism. We consider

$$\omega\colon \mathfrak{A} \to \mathbb{C} \; A \;\; \mapsto \psi(\tau(A)). \tag{72.2}$$

It is clearly linear (because ψ and τ are) and $\omega(A)\omega(B) = \omega(AB)$. Furthermore $\omega(B) \neq 0$ and $\omega(\mathbb{1}) = 1$ are two good reasons to conclude that $\omega \neq 0$. Then $\omega \in \Delta(\mathfrak{A})$. It remains to be proved that $\mathfrak{I} = \ker\omega$. First, $\mathfrak{I} = \ker\tau$, then $\mathfrak{I} \subseteq \ker\omega$. But when $B \notin \mathfrak{I}$, we have $\omega(B) \neq 0$, then $\mathfrak{I} = \ker\omega$. This finish the proof. □

Theorem 72.4 (Banach-Alaoglu)**.**
If X is a closed normed vector space, then the unit closed ball in the dual X^* is compact for the x^*-topology. tho:Banach_Alaoglu

Proposition 72.5.
When $\mathfrak{A}$ is an unital commutative Banach algebra, the space $\Delta(\mathfrak{A})$ is compact and Hausdorff for the Gelfand topology.

prop:DcA_comp_Hauss

Proof. We first prove that $\Delta(\mathfrak{A})$ is closed by showing that it contains all limits of converging sequences [1]. Let us take a sequence $\omega_n \to \omega$ with $\omega_n \in \Delta(\mathfrak{A})$. We will show that $\omega \in \Delta(\mathfrak{A})$:

$$|\omega(AB) - \omega(A)\omega(B)| \leqslant |\omega(AB) - \omega_n(AB)| + |\omega_n(A)\omega_n(B) - \omega(A)\omega(B)|,$$

but

$$\begin{aligned}\omega_n(A)\omega_n(B) - \omega(A)\omega(B) &= [\omega_n(A) - \omega(A)]\omega_n(B) + \omega(A)[\omega_n(B) - \omega(B)] \\ &\leqslant |\omega_n(A) - \omega(A)|\|B\| + \|A\|\omega_n(B) - \omega(B)|\end{aligned}$$

because $\omega_n(B) \leqslant \|B\|$. Taking the limit $n \to \infty$, we find

$$\begin{aligned}|\omega(AB) - \omega(A)\omega(B)| \leqslant\; & |\omega(AB) - \omega_n(AB)| \\ & + |\omega_n(A) - \omega(A)|\|B\| \\ & + \|A\||\omega_n(B) - \omega(B)| \to 0.\end{aligned}$$

This proves that $\omega \in \Delta(\mathfrak{A})$ and therefore that $\Delta(\mathfrak{A})$ is closed. Since $\|\omega\| = 1$ for all ω, we have $\Delta(\mathfrak{A}) \subset \mathfrak{A}_1^*$, the unit ball in $\mathfrak{A}^*$. Theorem 72.4 claims that $\mathfrak{A}_1^*$ is compact in the Gelfand topology. So $\Delta(\mathfrak{A})$ is closed in a compact. This makes $\Delta(\mathfrak{A})$ compact by lemma 7.96(1).

Now, we check that it is also Hausdorff. If $\omega \neq \eta \in \Delta(\mathfrak{A})$, there exists a $A \in \mathfrak{A}$ such that $\omega(A) \neq \eta(A)$. We thus consider $\mathcal{O}$ and $\mathcal{O}'$, two disjoints open subsets of $\mathbb{C}$ around $\omega(A)$ and $\eta(A)$ respectively. With these definition, it is easy to see that $\hat{A}^{-1}(\mathcal{O})$ and $\hat{A}^{-1}(\mathcal{O}')$ are disjoints neighbourhoods of ω and η. □

1. It is no related to *complete* spaces in which any Cauchy sequence converge

72.1.2 Topology on $\Delta(\mathfrak{A})$

subsec:topo_Delta

We begin to put the **w^*-weak topology** on $\mathfrak{A}^*$ which defined by the convergence notion $\omega_n \to \omega$ if and only if $\omega_n(A) \to \omega(A)$ for all $A \in \mathfrak{A}$.

The **Gelfand topology** is the induced topology from $\mathfrak{A}^*$ on $\Delta(\mathfrak{A})$. Let us define the **Gelfand transform**

$$\begin{aligned} \hat{A}\colon \Delta(\mathfrak{A}) &\to \mathbb{C} \\ \hat{A}(\omega) &\mapsto \omega(A). \end{aligned} \tag{72.3}$$

General theory of functional analysis shows that the w^*-weak topology is the weakest in which all linear functional are continuous, so a basis of this topology is given by sets of the form $\{f \in \mathfrak{A}^* \text{ tel que } f(A) \in \mathcal{O}\}$ where $\mathcal{O}$ is an open in $\mathbb{C}$ and $A \in \mathfrak{A}$.

A basis of the Gelfand topology is the intersection of these set with $\Delta(\mathfrak{A})$:

$$\hat{A}^{-1}(\mathcal{O}) = \{\omega \in \Delta(\mathfrak{A}) \text{ tel que } \omega(A) \in \mathcal{O}\}. \tag{72.4}$$

Lemma 72.6.
An element $A \in \mathfrak{A}$ is invertible if and only if $\omega(A) \neq 0$ for all $\omega \in \Delta(\mathfrak{A})$.

Proof. Let A be an invertible element in A and $\omega \in \Delta(\mathfrak{A})$ such that $\omega(A) = 0$. Then

$$1 = \omega(\mathbb{1}) = \omega(A)\omega(A^{-1}) = 0.$$

Let us take now a $A \notin G(\mathfrak{A})$, then the ideal $\mathfrak{I}_A := \{AB \text{ tel que } B \in \mathfrak{A}\}$ don't contain $\mathbb{1}$ and is not a proper ideal. From choice axiom, $\mathfrak{I}_A$ is contained in a maximal ideal $\mathfrak{I}$. From (3) of 72.3, there exists a $\omega \in \Delta(\mathfrak{A})$ whose kernel is $\mathfrak{I}$. In particular, $\omega(A) = 0$. □

Theorem 72.7.
Let $\mathfrak{A}$ be an unital commutative Banach algebra. Then

(1) The Gelfand transform is a homomorphism $\mathfrak{A} \to C(\Delta(\mathfrak{A}))$. enugi

(2) The image of $\mathfrak{A}$ under the Gelfand transform separates the points in $\Delta(\mathfrak{A})$, see definition 48.25. enugii

enugiii

(3) The spectrum of $A \in \mathfrak{A}$ is

$$\sigma(A) = \sigma(\hat{A}) = \{\hat{A}(\omega) : \omega \in \Delta(\mathfrak{A})\}.$$

enugiv

*(4) The Gelfand transform is a **contraction**: $\|\hat{A}\|_\infty \leqslant \|A\|$.*

tho:unital_comm

Proof. Item (1) is easy: $\widehat{AB}(\omega) = \omega(AB) = \omega(A)\omega(B) = \hat{A}(\omega)\hat{B}(\omega)$. Point (2) is immediate too: let $\omega_1 \neq \omega_2 \in \Delta(\mathfrak{A})$. We need a $A \in \mathfrak{A}$ such that $\hat{A}(\omega_1) \neq \hat{A}(\omega_2)$. But the definition of the inequality $\omega_1 \neq \omega_2$ is the existence of a $A \in \mathfrak{A}$ such that $\omega_1(A) \neq \omega_2(A)$.

For (3), recall that

$$\rho(A) = \{z \in \mathbb{C} \text{ tel que } (A - z)^{-1} \text{ exists}\}.$$

From the lemma the existence of $(A-z)^{-1}$ makes that $\forall \omega \in \Delta(\mathfrak{A})$, $\omega(A) \neq z$. So the complementary is

$$\begin{aligned} \sigma(A) &= \{z \in \mathbb{C} \text{ tel que } \exists \omega \in \Delta(\mathfrak{A}) \text{ such that } \omega(A) = z\} \\ &= \{\omega(A) \text{ tel que } \omega \in \Delta(\mathfrak{A})\} \\ &= \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A})\}. \end{aligned} \tag{72.5}$$

The fifth point comes from definition 48.26 and the fact that, because of the third point, $r(A) = sup\{\hat{A}(\omega) : \omega \in \Delta(\mathfrak{A})\}$. Therefore

$$\|\hat{A}\|_\infty = \sup_{\omega \in \Delta(\mathfrak{A})} |\hat{A}(\omega)| = r(A) \leqslant \|A\|.$$

□

When a Banach algebra is non unital, one can extend it to $\mathfrak{A}_{\mathbb{1}}$ and a character $\omega \in \Delta(\mathfrak{A})$ can be extended too as $\tilde{\omega} \in \Delta(\mathfrak{A}_{\mathbb{1}})$ by

$$\tilde{\omega}(A + \lambda \mathbb{1}) = \omega(\mathbb{1}) + \lambda.$$

The fact that it is multiplicative is a simple computation.

Theorem 72.8.
For every element A of a commutative Banach algebra $\mathfrak{A}$, we have $\mathrm{Spec}(A) = \mathrm{Spec}(\hat{A})$.

Proof. We want to prove that when $\lambda \in \mathrm{Spec}(A)$, there exists a φ such that $\varphi(A) = \lambda$. The ideal generated by $(A - \lambda)$ is a proper ideal which is thus contained in a maximum ideal M by Zorn's lemma. This maximal ideal is closed (if not, the closure would be bigger ideal). Consider an element x in the quotient $\mathfrak{A}/M$. Since $\mathrm{Spec}(x) \neq \varnothing$, the element $(x - \lambda)$ is not invertible for some λ. That provides an isomorphism $\mathfrak{A}/M \simeq \mathbb{C}$, and we define φ as the composition of that isomorphism by the projection of $\mathfrak{A}$ into $\mathfrak{A}/M$. For this φ, we have $\varphi(A) = \lambda$.

¡¡ Avertissement/question au lecteur !! 72.9
Faudrait creuser pourquoi on a un isomorphisme $\mathfrak{A}/M \simeq \mathbb{C}$.

□

72.1.3 An example

Let $\mathfrak{A} = L^1(\mathbb{R})$ with the norm

$$\|f\|_1 = \int_{\mathbb{R}} |f(x)| dx,$$

and the convolution product

$$(f \star g)(x) = \int_{\mathbb{R}} f(x - y) g(y) dy.$$

We don't take care to analysis subtleties as completion and precise convergence of integrals. For example, we will use and abuse of Fubini's theorem and often say "for all" when "for almost all" should be preferable. From the fact that $|f(x - y)g(y)| \leqslant |f(x - y)||g(y)|$ and $\int_{\mathbb{R}} f(x - y)dx = \int_{\mathbb{R}} f(x)dx$, we find that

$$\|f \star g\|_1 \leqslant \|f\|_1 \|g\|_1$$

as needed to prove that $(L^1(\mathbb{R}), \star)$ is a Banach algebra. This is a non unital Banach space because the unit should be the Dirac delta. From analysis, one knows that the dual space of $L^1(\mathbb{R})$ is $L^\infty(\mathbb{R})$ with, for $u \in L^\infty(\mathbb{R})$,

$$u(f) = \int_{\mathbb{R}} f(x) u(x) dx.$$

Since $\Delta(\mathfrak{A})$ is a subset of $L^\infty(\mathbb{R})$, there exists, for each $\omega \in \Delta(L^1(\mathbb{R}))$, a $\hat{\omega}$ such that $\omega(f) = \int_{\mathbb{R}} f(x)\hat{\omega}(x)dx$. With an easy change of variable, the multiplicative condition $\omega(f \star g) = \omega(f)\omega(g)$ gives

$$\int_{\mathbb{R}^2} f(t)g(y)\hat{\omega}(t + y)dtdy = \int_{\mathbb{R}^2} f(x)g(y)\hat{\omega}(x)\hat{\omega}(y)dxdy.$$

We can conclude that $\hat{\omega}(x + y) = \hat{\omega}(x)\hat{\omega}(y)$, in such a manner that

$$\hat{\omega}(x) = e^{ipx}$$

for a certain $p \in \mathbb{C}$. For $\hat{\omega}$ to belongs to $L^\infty(\mathbb{R})$, we must have $p \in \mathbb{R}$. So we get a bijection $\Delta(\mathfrak{A}) \simeq \mathbb{R}$. By this identification, we denote by p the element of $\Delta(L^1(\mathbb{R}))$ given by $\hat{\omega}(x) = e^{ipx}$. With theses notations, the Gelfand $\hat{A}(\omega) = \omega(A)$ transform reads

$$\hat{f}(p) = \omega(f) = \int_{\mathbb{R}} f(x)\hat{\omega}(x) = \int_{\mathbb{R}} f(x)e^{ipx}dx. \tag{72.6}$$

This is nothing else than the Fourier transform! We know that Fourier transform changes the convolution product into the pointwise usual product of functions:

$$\widehat{f \star g}(p) = \hat{f}(p)\hat{g}(p) = (\hat{f}\hat{g})(p).$$

This express the fact that the Gelfand transform is a homomorphism between $\mathfrak{A}$ —i.e. the product $\star$— and $C(\Delta(\mathfrak{A}))$ —i.e. the pointwise product. It is precisely the claim (1) of theorem 72.7.

Theorem 72.10.
Let $\mathfrak{A}$ be a non unital commutative Banach algebra. Then

enuhi
(1) *The space $\Delta(\mathfrak{A})$ is Hausdorff locally compact for the Gelfand topology,*

enuhii
(2) *$\Delta(\mathfrak{A}_{\mathbb{1}})$ is the one point compactification of $\Delta(\mathfrak{A})$,*

enuhiii
(3) *the Gelfand transformation is a homomorphism $\mathfrak{A} \to C_0(\Delta(\mathfrak{A}))$,*

enuhiv
(4) *the spectrum of $A \in \mathfrak{A}$ is*

$$\mathrm{Spec}(A) = \sigma(A) = \{0\} \cup \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A})\}.$$

enuhv
(5) *The image of $\mathfrak{A}$ by the Gelfand transform separates points in $\Delta(\mathfrak{A})$,*

enuhvi
(6) *Gelfand transform is a contraction:*

$$\|\hat{A}\|_\infty \leqslant \|A\|.$$

For one point compactification issues, see section 48.5.

Proof. (1) We add an unity to $\mathfrak{A}$ and we remark that

$$\Delta(\mathfrak{A}_{\mathbb{1}}) = \Delta(\mathfrak{A}) \cup \infty$$

where ∞ is defined by $\infty(A + \lambda\mathbb{1}) = \lambda$. Indeed let $\psi \in \Delta(\mathfrak{A})$ and let us ask ourself how to extend it to a multiplicative functional in $\varphi \in \Delta(\mathfrak{A}_{\mathbb{1}})$. For, let $B \in \mathfrak{A}$ such that $\psi(A) \neq 0$ remark that multiplicative condition imposes $\varphi((\lambda\mathbb{1})(B)) = \varphi(\lambda)\varphi(B)$ while the linearity gives $\varphi(\lambda B) = \lambda\varphi(B)$. Thus $\varphi(\lambda\mathbb{1}) = \lambda$ and the unique possibility to extends ψ is

$$\varphi(A + \lambda\mathbb{1}) = \varphi(A) + \lambda$$

and we note ∞ the new functional

$$\infty(A + \lambda\mathbb{1}) = \lambda.$$

Since $\mathfrak{A}_{\mathbb{1}}$ is unital, the character space $\Delta(\mathfrak{A}_{\mathbb{1}})$ is Hausdorff and compact for its Gelfand topology. As set

$$\Delta(\mathfrak{A}) = \Delta(\mathfrak{A}_{\mathbb{1}})\backslash\{\infty\}.$$

We should prove that the induced topology on $\Delta(\mathfrak{A})$ from the Gelfand of $\Delta(\mathfrak{A}_{\mathbb{1}})$ is precisely the own Gelfand topology of $\Delta(\mathfrak{A})$. In this case, properties of compactification shall gives local compactness.

A basis of the topology of $\Delta(\mathfrak{A}_{\mathbb{1}})$ is given by $\hat{A}^{-1} = \{\omega \in \Delta(\mathfrak{A}_{\mathbb{1}}) \text{ tel que } \omega(A) \in \mathcal{O}\}$. Then any open set of $\Delta(\mathfrak{A})$ is open for the induced topology because

$$\{\omega \in \Delta(\mathfrak{A}) \text{ tel que } \omega(A) \in \mathcal{O}\} = \{\eta \in \Delta(\mathfrak{A}_{\mathbb{1}}) \text{ tel que } \eta(A) \in \mathcal{O}\} \cap \Delta(\mathfrak{A}).$$

For the converse, an open set for the induced topology is given by

$$\begin{aligned}&\{\omega \in \Delta(\mathfrak{A}_{\mathbb{1}}) \text{ tel que } \exists A \in \mathfrak{A}, \lambda \in \mathbb{C} : \omega(A + \lambda\mathbb{1}) \in \mathcal{O}\}\backslash\{\infty\}\\ &= \{\omega \in \Delta(\mathfrak{A}) \text{ tel que } \exists A \in \mathfrak{A}, \lambda \in \mathbb{C} : \omega(A) \in \mathcal{O} - \lambda\}\end{aligned}$$

where $\mathcal{O} - \lambda$ is as open as $\mathcal{O}$. This proves (1) and (2).

For (3), the point is not to prove that Gelfand transform is a homomorphism (that is trivial), but rather that it takes values in $C_0(\Delta(\mathfrak{A}))$.

The complementary of a compact set K in $\Delta(\mathfrak{A}_\mathbb{1})$ is an open set which contains ∞. Since $\hat{A}(\infty) = 0$, the values of $\hat{A}$ in the complementary of K are as small as we want when K becomes larger and larger.

In order to prove (4), recall that, by definition, $\sigma_\mathfrak{A}(A) = \sigma_{\mathfrak{A}_\mathbb{1}}(A)$. Then

$$\begin{aligned}\sigma_{\mathfrak{A}_\mathbb{1}} &= \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A}_\mathbb{1})\} && (72.7a)\\ &= \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A})\} \cup \hat{A}(\infty) && (72.7b)\\ &= \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A})\} \cup \{0\}. && (72.7c)\end{aligned}$$

Since (5) and (6) are true for $\mathfrak{A}_\mathbb{1}$, they are true for $\mathfrak{A}$.

□

72.2 Commutative C^*-algebras

Definition 72.11.
A C^****-algebra*** *is an involutive (complex) Banach algebra such that for all* $A, B \in \mathfrak{A}$,

(1) $\|AB\| \leqslant \|A\|\|B\|$,

(2) $\|A^*A\| = \|A\|^2$.

One immediately has $\|A\|^2 = \|A^*A\| \leqslant \|A^*\|\|A\|$, *then*

$$\|A\| = \|A^*\| \tag{72.8}$$

for all element A *in a* C^**-algebra.*

LemFiniCSestVNa

Lemma 72.12.
Every finite dimensional C^**-algebra is a von Neumann algebra.*

Lemma 72.13 (Stone-Weierstrass theorem)**.**
Let X *be a compact and Hausdorff space. Any* C^**-subalgebra of* $C(X)$ *containing* 1_X *and separating points in* X *is exactly* $C(X)$ *seen as* C^**-algebra.*

lem:Stone_W

Here, 1_X denotes the constant function 1 on X.

PropcomCstarDelCeqX

Proposition 72.14.
Let X *be a compact Hausdorff space and see* $C(X)$ *as a commutative* C^**-algebra. Then* $\Delta(C(X))$ *is homeomorphic to* X.

prop:comHauffhomeo

Proof. For $x \in X$, one defines

$$\begin{aligned}\omega_x \colon C(X) &\to \mathbb{C}\\ f &\mapsto f(x).\end{aligned} \tag{72.9}$$

It is clearly non zero and multiplicative. Then $\omega_x \in \Delta(C(X))$. We denote by $E \colon X \to \Delta(C(X))$ the map which makes the correspondence between x and ω_x

$$E(x)f = f(x).$$

Urysohn lemma 48.23 applied to the compact Hausdorff space $X = \Delta(C(X))$ makes that if $x \neq y$, then there exists a function $f \in C(X)$ such that $f(x) \neq f(y)$. This proves that E is injective.

From theorem 72.3, we know that

$$\mathfrak{I}_x = \ker \omega_x = \{f \in C(X) \text{ tel que } f(x) = 0\}$$

is an ideal in $C(X)$. Suppose that E is not surjective. Then there exists some $\omega \in \Delta(C(X)))$ which don't come from a $x \in X$; for such a ω, we pose

$$\mathfrak{I}_\omega = \ker \omega = \{f \in C(X) \text{ tel que } \omega(f) = 0\}.$$

This $\mathfrak{I}_\omega$ can't contains any $\mathfrak{I}_x$ because they are maximal ideals. Then for all $x \in X$, there exists a $f \in C(X)$ such that $f(x) = 0$ with $f \notin \mathfrak{I}_\omega$. If E is not surjective, then there exists a maximum ideal $\mathfrak{I}$, kernel of a character which is not in the image of E. In order this ideal to be included in none of the $\mathfrak{I}_x$, one needs that for all $x \in X$, there exists $f_x \in \mathfrak{I}$ such that $f_x(x) \neq 0$. Let $\mathcal{O}_x$ be an open set on which $f_x \neq 0$. Since X is compact, one can extract a finite subcovering $X = \bigcup_i \mathcal{O}_{x_i}$. Now we build

$$g := \sum_{i=1}^{n} |f_{x_i}|^2.$$

This is a strictly positive function, then $1/g \in C(X)$, and then $\mathfrak{I} = C(X)$ and $\mathfrak{I}$ should be the kernel of a zero character. This is impossible, then E is surjective and it is a bijection.

In order to prove that E is an homeomorphism, we will use the lemmas 48.20 and 48.21. Let X_0 be the space X endowed with its initial topology and X_G the same space with the topology induced from E^{-1}, i.e. that an open set in X_G is always the image by E^{-1} of an open set in $\Delta(C(X))$. From definition, E is continuous for the topology X_G. We are going to prove that $X_0 = X_G$. Definitions give for all $f \in C(X)$,

$$(\hat{f} \circ E)(x) = \hat{f}(\omega_x) = \omega_x(f) = f(x).$$

But Gelfand topology is the weakest topology for which all f are continuous. On the other hand, f is continuous because it belongs to $C(X_0)$. Then the topology of X_G is weaker than the one of X_0. Indeed, let $\mathcal{O}$ be an open set for X_G and let us prove that it contains an open set of X_0. From definition, $\mathcal{O} = E^{-1}(\mathcal{O}')$ for a certain open set $\mathcal{O}'$ of $\Delta(C(X))$, i.e. $\mathcal{O}' = \hat{f}^{-1}(A)$ for an open A in $\mathbb{C}$. The topology X_G is the minimal one for which $E^{-1} \circ \hat{f}^{-1}(A)$ is open. But $E^{-1} \circ \hat{j}^{-1} = f^{-1}$, then $(E^{-1} \circ \hat{f}^{-1})(A) = f^{-1}(A)$ is open in X_0. □

Theorem 72.15 (Gelfand theorem)**.**
For any commutative unital C-algebra $\mathfrak{A}$, there exists an unique (up to isomorphism) compact and Hausdorff space X such that $\mathfrak{A}$ is isomorphic to $C(X)$.* thoGelfand

Proof. We immediately give the answer: $X = \Delta(\mathfrak{A})$ and the isomorphism is

$$\begin{aligned} \varphi\colon \mathfrak{A} &\to C(\Delta(\mathfrak{A})) \\ A &\mapsto \hat{A}. \end{aligned} \tag{72.10}$$

We first have to prove that $\Delta(\mathfrak{A})$ is compact and Hausdorff. Then it should be proved that φ is an isometric C^*-algebra isomorphism and finally that this is the only possibility.

The space $\Delta(\mathfrak{A})$ is compact and Hausdorff.

Proposition 72.1.1 gives it.

The map φ takes values in $C(\Delta(\mathfrak{A}))$.

From discussion at top of subsection 72.1.2, the functional $\hat{A}$ is continuous on $X = \Delta(\mathfrak{A})$.

The map φ is a morphism.

Linearity of φ is clear. Property $\varphi(AB) = \varphi(A)\varphi(B)$ comes from point (1) of proposition 72.1.1. So we are left to prove that $\varphi(A^*) = \varphi(A)^*$. It is sufficient to prove that, if $A = A^*$, then $\varphi(A)$ takes his values in $\mathbb{R}$. So let $A \in \mathfrak{A}_{\mathbb{R}}$ and write $\omega(A) = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$. If we define $B = A - \alpha\mathbb{1}$, then $\omega(B) = i\beta$ because $\omega(\mathbb{1}) = 1$. Furthermore $B = B^*$. Let $t \in \mathbb{R}$; we have

$$|\omega(B + it\mathbb{1})|^2 = |\omega(B) + it|^2 = \beta^2 + 2\beta t + t^2. \tag{72.11}$$ eq:rcinq

Using formulas $|\omega(A)| \leqslant \|A\|$ and $\|AA^*\| = \|A\|^2$, we find

$$|\omega(B + it\mathbb{1})|^2 \leqslant \|B + it\mathbb{1}\|^2 = \|B^2 + t^2\| \leqslant \|B\|^2 + t^2. \tag{72.12}$$

Then net result is that for all $t \in \mathbb{R}$, $\beta^2 + 2t\beta \leqslant \|B\|^2$. It is only possible when $\beta = 0$. Then $\omega(A) \in \mathbb{R}$ as soon as $A = A^*$.

The map φ is isometric.

Let us begin with $A = A^*$. So $\|A^2\| = \|A\|^2$ and $\|A^{2^m}\| = \|A\|^{2^m}$. Using proposition 60.68, we find

$$r(A) = \|A\|.$$

On the other hand the definition of the supremum norm on the Hausdorff space $\Delta(\mathfrak{A})$ reads

$$\|\hat{A}\|_\infty = \sup_{\omega \in \Delta(\mathfrak{A})} |\hat{A}(\omega)| = r(A) = \|A\|. \tag{72.13}$$

eq:AinfA

Then φ is isometric when $A = A^*$. Now, A^*A is selfadjoint and $\|A^*A\| = \|A\|^2$, then

$$\|A\|^2 = \|A^*A\| = \|\widehat{A^*A}\|_\infty = \|\hat{A}^*\hat{A}\|_\infty = \|\hat{A}\|_\infty^2.$$

The map φ is injective.

If $\varphi(A) = \varphi(B)$, then $\varphi(A - B) = 0$. The only way for φ to be an isometry is $A - B = 0$.

The map φ is surjective.

Since φ is an isometry, it sends a closed set into a closed set, but $\mathfrak{A}$ is closed because it is a Banach space. Point (2) of theorem 72.7 says that $\varphi(\mathfrak{A})$ separates points in $\Delta(\mathfrak{A})$ and we just proved the φ preserves the adjoint, so $\varphi(\mathfrak{A})$ is a C^*-subalgebra of $C(\Delta(\mathfrak{A}))$. Finally, it is clear that $\hat{\mathbb{1}} = 1_X$. Lemma 72.2 concludes $\varphi(\mathfrak{A}) = C(\Delta(\mathfrak{A}))$.

Now proposition 72.14 makes φ and homeomorphism between $\mathfrak{A}$ and $\Delta(\mathfrak{A})$. So the topological structure of $\mathfrak{A}$ is encoded in the algebraic (Banach) structure of $C(\Delta(\mathfrak{A}))$. So if $C(Y) \simeq \mathfrak{A} \simeq C(X)$ as C^*-algebras, then $X \simeq Y$ as topological space. This proves the unicity part and concludes the Gelfand theorem.

□

As far as notations are concerned, let us recall that the Gelfand transform is $A \mapsto \hat{A}$ with

$$\begin{aligned} \hat{A} \colon \Delta(\mathfrak{A}) &\to \mathbb{C} \\ \omega &\mapsto \omega(A). \end{aligned} \tag{72.14}$$

One particular class of elements in $\Delta\big(C(X)\big)$ is the ones of the form ω_x for $x \in X$. These are defined by

$$\begin{aligned} \omega_x \colon C(X) &\to \mathbb{C} \\ f &\mapsto f(x). \end{aligned} \tag{72.15}$$

The Gelfand theorem says that every element of $\Delta\big(C(X)\big)$ reads ω_x for a certain $x \in X$.

Lemma 72.16.
Let $\mathfrak{A}$ be a C^*-algebra, and $\mathfrak{B}(\mathfrak{A})$, the set of bounded operators on $\mathfrak{A}$. Then

(1) The map

$$\begin{aligned} \rho \colon \mathfrak{A} &\to \mathfrak{B}(\mathfrak{A}) \\ \rho(A)B &\mapsto AB \end{aligned} \tag{72.16}$$

is a diffeomorphism between $\rho(\mathfrak{A})$ and $\rho(\mathfrak{A}) \subset \mathfrak{B}(\mathfrak{A})$.

(2) If $\mathfrak{A}$ has no unit, one can define a norm on $\mathfrak{A}_\mathbb{1}$ by

$$\|A + \lambda\mathbb{1}\| = \|\rho(A) + \lambda\mathbb{1}\| \tag{72.17}$$

eq:normCAu

where the right hand side norm is the one in $\mathfrak{B}(\mathfrak{A})$, see 60.44. With the usual multiplication and the involution

$$(A + \lambda\mathbb{1})^* = A^* + \overline{\lambda}\mathbb{1}, \tag{72.18}$$

the set $\mathfrak{A}_\mathbb{1}$ becomes an unital C^*-algebra.

lem:unitariz_C

Proof. Since $\mathfrak{A}$ is a C^*-algebra, $\|\rho(A)B\| \leqslant \|A\|\|B\|$, then for all $A \in \mathfrak{A}$, one has $\|\rho(A)\| \leqslant \|A\|$. On the other hand, we know that $\|A^*A\| = \|A\|^2$ and $\|A^*\| = \|A\|$, then

$$\|A\| = \frac{\|AA^*\|}{\|A\|} = \left\|\rho(A)\frac{A^*}{\|A\|}\right\| \leqslant \|\rho(A)\|$$

from definition of the sup norm. Then $\|\rho(A)\| = \|A\|$ and ρ is an isometry and then is injective because it is linear. It is clearly a homomorphism too. The map $A + \lambda\mathbb{1} \to \rho(A) + \lambda\mathbb{1}$ is a C^*-algebra-morphism if we define[2] $\rho(A)^* = \rho(A^*)$. Since the sup norm fulfils condition (60.33), the norm (72.17) fulfils the same. So $\mathfrak{A}_\mathbb{1}$ becomes a Banach $*$-algebra and lemma 60.48 will help us to conclude that it is a C^*-algebra.

The formula $\|A\|^2 - \varepsilon \leqslant \|Av\|^2$ holds for an operator A on a general Banach algebra and an arbitrary vector v with norm 1. In our present case, if $\|B\| = 1$,

$$\begin{aligned}\|\rho(A) + \lambda\mathbb{1}\|^2 - \varepsilon &\leqslant \|(\rho(A) + \lambda\mathbb{1})B\|^2 \\ &= \|AB + \lambda B\|^2 \\ &= \|(AB + \lambda B)^*(AB + \lambda B)\| \\ &= \|\rho(B^*)\rho(A^* + \overline{\lambda}\mathbb{1})\rho(A + \lambda\mathbb{1})B\| \\ &\leqslant \|\rho(B^*)\|\|(\rho(A) + \lambda\mathbb{1})^*(\rho(A) + \lambda\mathbb{1})\|\|B\|,\end{aligned} \tag{72.19}$$

but we also know that $\|\rho(B^*)\| = \|B^*\| = \|B\| = 1$. Letting $\varepsilon \to 0$, we find $\|A\|^2 \leqslant \|A^*A\|$ in the Banach $*$-algebra $\mathfrak{A}_\mathbb{1}$.

□

pg:unit_nonunic

This lemma gives us an unitization of a C^*-algebra which is not the one previously given for a Banach algebra. This shows that unitization of Banach algebra is not unique. For a C^*-algebra, however, we have an unicity result:

Proposition 72.17.
For every C^*-algebra without unit, there exists an unique unital C^*-algebra $\mathfrak{A}_\mathbb{1}$ and an isometric morphism (hence injective) $\mathfrak{A} \to \mathfrak{A}_\mathbb{1}$ such that $\mathfrak{A}/\mathfrak{A}_\mathbb{1} = \mathbb{C}$. prop_unitariz_csa

Proposition 72.18.
If $\mathfrak{A}$ is a commutative C^*-algebra, any character is hermitian.

Proof. When χ is a character, $\chi(A) \in \sigma(A)$ for all $A \in \mathfrak{A}$ and when $A = A^*$, we have $\sigma(A) \subset \mathbb{R}$. For any A, we have a decomposition $A = A_1 + iA_2$ and

$$\chi(A^*) = \chi(A_1 - iA_2) = \underbrace{\chi(A_1)}_{\in\mathbb{R}} - i\underbrace{\chi(A_2)}_{\in\mathbb{R}} = \overline{\chi(A)}.$$

□

72.3 Functional calculus in unital C^*-algebras

From now, the C^*-algebra $\mathfrak{A}$ is no more assumed to be commutative, but it is unital.

DefElemNormal

Definition 72.19.
*An element A in an involutive algebra is said **normal** when $[A, A^*] = 0$.*

This is a direct generalisation of the concept of normal operator in the Hilbert space setting (definition 60.35).

2. We know a definition of $*$ when we look at $\mathfrak{B}(H)$ where H is a Hilbert space, but we are here with $\mathfrak{B}(\mathfrak{A})$ where $\mathfrak{A}$ is no more than a Banach space; hence we do not have a definition of $*$.

If $\mathfrak{A}$ is a C^*-algebra and A, $B \in \mathfrak{A}$ we denotes by $C^*(A_1, \dots, A_n)$ the C^*-algebra generated by the A_i. This is the closure of every finite products of the form $Z_1 \cdots Z_k$ where each Z_j is one of the A_i.

For any A in a C^*-algebra we know that $\|A^*A\| = \|A\|^2$.

If A is normal, then $C^*(A, \mathbb{1})$ is commutative. Indeed any element of the form $A_A \dots A_n$ with $A_i = A$ or A^* can be written under the form $A \dots AA^* \dots A^*$.

Proposition 72.20.

$$\|A\| = \sqrt{r(A^*A)} \tag{72.20}$$ eq:ray_norme

Proof. Let A be in $\mathfrak{A}$ and consider a $z \in \rho(A)$. By definition, $(A - z)^{-1}$ exists in $\mathfrak{A}$; since φ is a morphism, $\varphi(A - z)$ is also invertible: it is clear that $\varphi((A - z)^{-1})$ is a two-sided inverse of $\varphi(A - z)$. Hence $\rho(A) \subseteq \rho(\varphi(A))$ and thus $\sigma(\varphi(A)) \subseteq \sigma(A)$. Definition (60.4.1) of the spectral radius makes $r(\varphi(A)) \leqslant r(A)$ and equation (72.20) gives the thesis. □

Theorem 72.21.
Consider an unital C^*-algebra $\mathfrak{A}$ and a $A \in \mathfrak{A}$ such that $A^* = A$. Then

(1) The spectrum $\sigma_{\mathfrak{A}}(A)$ is the same as $\sigma_{C^*(A,\mathbb{1})}(A)$, so that one can speak about $\sigma(A)$ without ambiguities.

(2) $\sigma(A) \subset \mathbb{R}$. enukiii

(3) $\Delta(C^*(A, \mathbb{1}))$ is homeomorphic to $\sigma(A)$ and $C^*(A, \mathbb{1})$ is isomorphic to $C(\sigma(A))$. Under this isomorphism, the Gelfand transformed $\hat{A}\colon \sigma(A) \to \mathbb{R}$ is the identity $id_{\sigma(A)}\colon t \to t$. tho:l_2.5.1

Proof. We first consider a normal $B \in G(\mathfrak{A})$, and the C^*-algebra $C^*(B, B^{-1}, \mathbb{1})$ generated by B, B^{-1} and $\mathbb{1}$. Since $(B^{-1})^* = (B^*)^{-1}$ and $BB^* = N^*B$, $[B^{-1}, B^{*-1}] = 0$.

Now, we are going to show that $[B^{*-1}, B] = 0$. First remark that $B^{*-1}B = (B^{-1}B^*)^{-1}$. We have to show that $B^{-1}B^*BB^{*-1} = \mathbb{1}$ and $BB^{*-1}B^{-1}B^* = \mathbb{1}$. These two equalities comes from $[B, B^*] = 0$ and $[B^{-1}, B^{*-1}] = 0$. The same makes that $[B^*, B^{-1}] = 0$.

The result is that $C^*(B, B^{-1}, \mathbb{1})$ is a commutative C^*-algebra So one can simply say that it is the closure of the polynomials in B, B^*, B^{-1}, and B^{*-1}.

By the Gelfand theorem, $C^*(B, B^{-1}, \mathbb{1})$ is then isomorphic to a $C(X)$ for some compact Hausdorff space X. Since B is invertible and the Gelfand transform is an isomorphism, $\hat{B}$ is invertible. Then $\forall\, x \in X$, $\hat{B}(x) \neq 0$. Indeed, the X is (up to an isomorphism) $\Delta(C^*(B, B^{-1}, \mathbb{1}))$. If for an $\omega \in \Delta(C^*(B, B^{-1}, \mathbb{1}))$, $\hat{B}(\omega)$, then $\omega(B) = 0$ and thus $\omega \equiv 0$. But in the definition of $\Delta(\mathfrak{A})$, we have explicitly excluded the null form.

On the other hand let us consider $f \in C(X)$ everywhere non zero. Since (pointwise) $0 < \|f\|_\infty^{-2} ff^* \leqslant 1$,

$$0 \leqslant 1_X - \|f\|_\infty^{-2} ff^* < 1. \tag{72.21}$$ eq:ff

But if $\|A\| < 1$, then

$$\sum_{k=0}^{n} A^k \to (\mathbb{1} - A)^{-1}.$$

As far as f is concerned for the sup norm, equation (72.21) makes $1_X - ff^*/\|f\|_\infty^2$ satisfy this convergence. Then

$$\left(\frac{ff^*}{\|f\|_\infty^2}\right)^{-1} = \sum_{k=0}^{\infty} \left(\mathbb{1} - \frac{ff^*}{\|f\|_\infty^2}\right)^k,$$

so that

$$\frac{1}{f} = \frac{f^*}{\|f\|_\infty^2} \sum_{k=0}^{\infty} \left(\mathbb{1} - \frac{ff^*}{\|f\|_\infty^2}\right)^k. \tag{72.22}$$

This is true for any f such that $f(x) \neq 0\ \forall x \in X$; in particular, it is true for $\hat{B}$. Thus $\hat{B}^{-1}$ is a limit of polynomials in $\hat{B}$ and $\hat{B}^*$. By the inverse Gelfand transform (which is obviously an isomorphism), B^{-1} is a limit of polynomials in B and B^*. This is:

$$C^*(B, B^{-1}, \mathbb{1}) = C^*(B, \mathbb{1}). \tag{72.23}$$

Now, we take our A from the hypothesis: $A = A^*$. Clearly, A is normal and $A - z$ too. If we take $z \in \rho(A)$, our work about B applies to $A - z$. Then $(A - z)^{-1}$ can be written as polynomials in $(A - z)$ and $\mathbb{1}$. Thus $(A - z)^{-1} \in C^*(A - z, \mathbb{1})$ and $z \in \rho_{C^*(A-z,\mathbb{1})}(A)$, but it is clear that $C^*(A - z, \mathbb{1}) = C^*(A, \mathbb{1})$. Finally:

$$\rho_{\mathfrak{A}}(A) = \rho_{C^*(A,\mathbb{1})}(A).$$

The set σ being nothing else than the complement of ρ, the first point of the theorem is finish.

In the course of the demonstration of the Gelfand theorem, we had shown that since $A = A^*$, $\forall \omega \in \Delta(\mathfrak{A})$, $\hat{A}(\omega) \in \mathbb{R}$. But

$$\sigma(A) = \{\hat{A}(\omega) : \omega \in \Delta(\mathfrak{A})\}.$$

Then $\sigma(A) \subset \mathbb{R}$.

The proof that $\hat{A}$ is a bijection and that it is continuous is not done here. Here we will just prove the continuity of $\hat{A}^{-1}$. From theorem 72.7 and what we just did, we know that

$$\sigma(A) = \{\hat{A}(\omega) : \omega \in \Delta(\mathfrak{A})\} \subset \mathbb{R},$$

but $\hat{A}^{-1}(z) = \omega$ when $\hat{A}(\omega) = z$, or $\omega(A) = z$. Then $\hat{A}^{-1}$ is defined on $\sigma(A)$. So from now, one can only consider $z \in \sigma(A)$ and $\hat{A}^{-1}(z)(A) = z$. By induction, $\hat{A}^{-1}(z)(A^n) = z^n$. An element in $\Delta(C^*(A, \mathbb{1}))$ is completely determined by its value on A. Then an open set therein has the general form

$$\mathcal{R} = \{\omega \in \Delta(C^*(A, \mathbb{1})) \text{ tel que } \hat{A}(\omega) \in \mathcal{O}\}$$

where $\mathcal{O}$ is any open set in $\mathbb{C}$. From definition, $\hat{A}(\mathcal{R}) = \mathcal{O}$. So $\hat{A}^{-1}$ is continuous.

□

¡¡ Avertissement/question au lecteur !! 72.22
There is a notational clash: what is written $\sigma(A)$ is the spectrum of A. I want to write it $\mathrm{Spec}(A)$ *instead.*

The following proposition is the **continuous functional calculus**.

ThoContFuncCalculus

Theorem 72.23 (Continuous functional calculus)**.**
Let $A \in \mathfrak{A}$ be self-adjoint and $f \in C(\mathrm{Spec}(A))$. One can define a map $\tilde{f} : \mathfrak{A} \to \mathfrak{A}$ in such a way that when f is a polynomial, $\tilde{f} = f$ and in other cases, it is the uniform approximation of f by polynomials. This map $\tilde{f}$ which will be denoted by f fulfills

(1) $\mathrm{Spec}(f(A)) = f(\mathrm{Spec}(A))$, [enuji]

(2) $\|f(A)\| = \|f\|_\infty$.

prop:cont_calc

¡¡ Avertissement/question au lecteur !! 72.24
The proof has to be reordered.

Proof. We know from theorem 72.21 that $\Delta(C^*(A, \mathbb{1}))$ is homeomorphic to $\sigma(A)$ and $C^*(A, \mathbb{1})$ to $C(\sigma(A))$.

The Gelfand theorem says that if one has a commutative unital C^*-algebra then one has an unique (up to homeomorphism) X such that $\mathfrak{A}$ is isomorphic to $C(X)$. Moreover, this isomorphism is an isometry [3]. But we just showed that $C^*(A, \mathbb{1})$ where isomorphic to $C(\sigma(A))$, then one has

$$\|\varphi(B)\| = \|B\|, \tag{72.24}$$

eq:norm_vp_B

3. Is is correct?

the first norm is taken in $C(\sigma(A))$ and the second one in $C^*(A,\mathbb{1})$. But when X is Hausdorff, we had adopted the $\|.\|_\infty$ norm, so that $\|\varphi(B)\| = \|f\|_\infty$ and equation (72.24) reads:

$$\|f\|_\infty = \|f(A)\|. \tag{72.25}$$

Now remark that $f(\sigma(A))$ is the set of values that f takes on $\sigma(A)$, but we know[4] that

$$\sigma(A) = \sigma(\hat{A}) = \{\hat{A}(\omega) : \omega \in \Delta(C^*(A,\mathbb{1}))\}.$$

It is now times to give a sense to $f(\hat{A})$. Since f is continuous on $\sigma(A)$, there exists a converging infinite sum such that $f(t) = \sum_{k=0}^{\infty} c_k t^k$ for any $t \in \sigma(A)$. In particular, $\forall \omega \in \Delta(C^*(A,\mathbb{1}))$, $\hat{A}(\omega) \in \sigma(A)$; thus $\sum c_k[\hat{A}(\omega)]^k$ converges everywhere we want. This sum will be denoted by $f(\hat{A})(\omega)$:

$$f(\hat{A})(\omega) = \sum_{k=0}^{\infty} c_k[\hat{A}(\omega)]^k. \tag{72.26}$$

In other words, $f(\hat{A})(\omega) = f(\hat{A}(\omega))$. We have

$$f(\sigma(A)) = \{f(\hat{A}(\omega)) : \omega \in \Delta(C^*(A,\mathbb{1}))\} = \{f(\hat{A})(\omega) : \omega \in \Delta(C^*(A,\mathbb{1}))\} = \sigma(f(\hat{A})). \tag{72.27}$$

It remains to be proved that $\sigma(f(\hat{A})) = \sigma(f(A))$. We already know that $\sigma(A) = \sigma(\hat{A})$, so we just have to prove that $f(\hat{A}) = \widehat{f(A)}$. On the one hand,

$$f(\hat{A})\omega = \sum c_k[\hat{A}(\omega)]^k = \sum c_k[\omega(A)]^k = f\big(\omega(A)\big),$$

on the other hand,

$$\widehat{f(A)}\omega = \omega\big(f(A)\big) = \omega\big[\sum c_k A^k\big].$$

On the other hand, one already know that $\sigma(A) = \sigma(A)$, thus we just have to see that $f(\hat{A}) = \widehat{f(A)}$ when $f\colon \sigma(A) \to \mathbb{C}$ is continuous. The problem is a permutation of ω and a limit:

$$f(\hat{A})\omega = \sum_{k=0}^{\infty} c_k\,\omega(A^k), \quad \widehat{f(A)}\omega = \omega\left(\sum_{k=0}^{\infty} c_k A^k\right).$$

What theorem 72.21 says is that $C^*(A,\mathbb{1})$ is isomorphic to $C(\sigma(A))$ with

$$\sum c_k A^K \mapsto f(x) = \sum c_k x^k.$$

In this isomorphism, the map $\hat{A}\colon \sigma(A) \to \mathbb{R}$ corresponds to the identity map. More precisely, the isomorphism $\varphi\colon C^*(A,\mathbb{1}) \to C(\sigma(A))$ is the following:

$$\varphi(B)(t) = a + \sum c_k t^k$$

when $B = a + \sum c_k A^k$. We know in general that

$$\sigma(A) = \sigma(\hat{A}) = \{f\big(\hat{A}(\omega)\big) \text{ tel que } \omega \in \Delta(\mathfrak{A})\}.$$

In the present case, we are working with $\mathfrak{A} = C^*(A,\mathbb{1})$, therefore

$$f\big(\sigma(A)\big) = \{f\big(\hat{A}(\omega)\big) \text{ tel que } \omega \in \Delta\big(C^*(A,\mathbb{1})\big)\}.$$

We have to prove that $f\big(\hat{A}(\omega)\big) = f(\hat{A})\omega$. Since f is continuous on $\sigma(A)$, the sum $f(t) = \sum c_k t^k$ converges for all $t \in \sigma(A)$. In particular for $\hat{A}(\omega) \in \sigma(A)$, the sum $\sum_{k=0}^{\infty}\big[\hat{A}(\omega)\big]^k$ converges.

It is now time to give a sense to $f(\hat{A})$. We know from theorem 72.21 that $\hat{A}\colon \Delta\big(C^*(A,\mathbb{1})\big) \to \sigma(A)$ is an isomorphism. As definition we set

$$f(\hat{A})(\omega) = \sum c_k\big[\hat{A}(\omega)\big]^k \tag{72.28}$$

4. Vas voir si on know ça vraiment

everywhere it converges. But, since $\hat{A}(\omega) \in \sigma(A)$, it converges everywhere it is interesting for us.

By definition, $\sum_{k=0}^{\infty} = \lim_{n\to\infty} \sum_{k=0}^{n}$, but the proposition 48.35, which gives link between convergence and continuity, assures us that one can permute the sum and ω because it is a continuous function on $C^*(A, \mathbb{1})$ which is by definition the closure of all polynomials in A:

$$\begin{aligned}\omega(\sum_k B^k) = \omega(\lim_{n\to\infty} \sum_k^n B^k) &= \lim_{n\to\infty} \omega(\sum_k^n B^k) \\ &= \lim_{n\to\infty} \sum_k^n \omega(B^k) = \sum_{k=0}^{\infty} \omega(B^k).\end{aligned} \tag{72.29}$$

We now turn our attention to the second point: $\|f(A)\|_{C^*(A,\mathbb{1})} = \|f(A)\|_{\mathfrak{A}}$. It uses proposition 2.26, chapter 4 of [985].

□

72.3.1 The isomorphism $C^*(A, \mathbb{1}) \leftrightarrow C(\sigma(A))$

By definition an element $B \in C^*(A, \mathbb{1})$ can be written as $B = f(A)$ where f is a sum of $\mathbb{1}$, A, A^2, A^*, $(A^*)^2$,... In the setting of continuous functional calculus, we suppose that A is selfadjoint, i.e. $A = A^*$, so that $f(A)$ is polynomial (eventually infinite) in A with an independent term $\mathbb{1}$. The isomorphism that we consider is

$$\begin{aligned}\varphi : C^*(A, \mathbb{1}) &\to C(\sigma(A)) \\ \varphi(B) &= f \in C(\sigma(A))\end{aligned} \tag{72.30}$$

where f is the "definition" function of B in $C^*(A, \mathbb{1})$.

RemExpansionSqrtConCal

Remarque 72.25.
The map φ depends on A. It could be better written φ_A. As an example, if $A = A^*$, the element $A^{1/2}$ is computed as follows. First, we know the expansion

$$\sqrt{t} = \sum_k a_k t^k. \tag{72.31}$$

EqExpanSqrtt

Then we define $\sqrt{A} = \sum_k a_k A^k$ as element of $C^*(A, \mathbb{1})$.

¡¡ Avertissement/question au lecteur !! 72.26
An expansion (72.31) is only possible when t is close to 1. Maybe the definition of $\sqrt{A}$ has to first look at $B = \lambda A$ with λ such that the norm of B is close to 1. Then we write $\sqrt{A} = \frac{1}{\sqrt{\lambda}}\sqrt{B}$. The square root of λ is well defined as a square root in $\mathbb{R}^+$.

In order to show that it is actually an isomorphism, we have to show the following points:

(1) it is linear;
(2) bijective;
(3) $\varphi(CD) = \varphi(C)\varphi(D)$;
(4) $\varphi(B^*) = \varphi(B)^*$.

Here are the proofs.

(1) The linearity is clear.
(2) Suppose $\varphi(B) = \varphi(C)$. Definition of φ gives $B = \varphi(B)(A)$ and $C = \varphi(C)(A)$. For the surjectivity, note that $C(\sigma(A))$ is given by continuous functions whose can be uniformly approximated by polynomials; then for each $f \in C(\sigma(A))$, there corresponds a $B = \varphi(A) \in C^*(A, \mathbb{1})$.
(3) Consider $C = f(A)$, $D = g(A)$; thus $CD = (fg)(A)$ and $\varphi(CD) = fg = \varphi(C)\varphi(D)$.

(4) The last point comes from the fact that $A = A^*$. Indeed, consider $B = f(A) = \sum_k c_k A^k$. Then
$$B^* = \sum_k c_k^* (A^*)^k = \sum_k c_k^* A^k = f^*(A).$$

We have shown that $\varphi(B) = f$ when $B = f(A)$ is an isomorphism between $C^*(A, \mathbb{1})$ and $C(\sigma(A))$ if A is selfadjoint: $A = A^*$.

Corollary 72.27.
For each C^*-algebra, there exists an unique unital C^*-algebra $\mathfrak{A}_\mathbb{1}$ and an isometric morphism $\mathfrak{A} \to \mathfrak{A}_\mathbb{1}$ such that $\mathfrak{A}_\mathbb{1}/\mathfrak{A} \simeq \mathbb{C}$.

cor_csa_unit

Proof. We yet defined $\mathfrak{A}_\mathbb{1}$ in lemma 72.16 and we just prove that the norm was unique. Since all elements in $\mathfrak{A}_\mathbb{1}$ are given under the form $A + \lambda\mathbb{1}$ with $A \in \mathfrak{A}$, it is obvious that $\mathfrak{A}_\mathbb{1}/\mathfrak{A} \simeq \mathbb{C}$. The canonical injection $\varphi(A) = A$ is a morphism. □

Lemma 72.28.
If $\varphi\colon \mathfrak{A} \to \mathcal{B}$ is a morphism of C^*-algebra and if $A = A^*$, then
$$f(\varphi(A)) = \varphi(f(A)).$$
for all $f \in C(\sigma(A))$.

lem:fvpvpf

Proof. Since $\sigma(\varphi(A)) \subseteq \sigma(A)$, the function f is well defined on $\varphi(A)$. If f is a polynomial, the result comes from the fact that $\varphi(AB) = \varphi(A)\varphi(B)$. If f is a general continuous function, it can be approximated by polynomials. Taking partial sums, $s_n = \sum_{k=1}^n \varphi(c_k A^k)$ and $v_n = \varphi(\sum_{k=1} c_k A^k)$, the linearity of *finite sums* gives the result. □

Where in the proof did we use the assumptions? The definition of $f(A)$ when $f\colon \mathbb{R} \to \mathbb{R}$ was given in 72.23 in order to get formulas $\sigma \circ f = f \circ \sigma$ and $\|f(A)\| = \|f\|$.

72.4 Positivity

Let $\mathfrak{A}$ be a C^*-algebra. We say that $A \in \mathfrak{A}$ is **positive** when

(1) $A = A^*$

(2) $\mathrm{Spec}(A) \subset \mathbb{R}^+$.

In this case, we write $A \geqslant 0$ or $A \in \mathfrak{A}^+$,
$$\mathfrak{A}^+ = \{A \in \mathfrak{A}_\mathbb{R} \text{ tel que } \sigma(A) \subset \mathbb{R}^+\}.$$
A set of particular importance is the set of selfadjoint elements:
$$\mathfrak{A}_\mathbb{R} = \{A \in \mathfrak{A} \text{ tel que } A = A^*\}. \tag{72.32}$$
These elements have real spectrum. We will see in theorem 72.31 that the set of positive elements in $\mathfrak{A}$ is given by
$$\mathfrak{A}^+ = \{A^2 \text{ tel que } A \in \mathfrak{A}_\mathbb{R}\} = \{B^*B \text{ tel que } B \in \mathfrak{A}\}. \tag{72.33}$$

Lemma 72.29.
For all A such that $A = A^*$, we have a decomposition
$$A = A_+ + A_-$$
where $A_+, A_- \in \mathfrak{A}^+$ and $A_+A_- = 0$. Furthermore
$$\|A_\pm\| \leqslant \|A\|.$$

lem:AsAdecm

Proof. We apply the continuous functional calculus with $f = \in_{\sigma(A)} = f_+ + f_-$ where

$$\begin{aligned} \mathrm{Id}_{\sigma(A)(t)} &= \max 0, t \qquad \text{because } \sigma(A) \subset \mathbb{R}^+ \\ f_+(t) &= \max\{0, t\} \\ f_-(t) &= \max\{-t, 0\}. \end{aligned} \tag{72.34}$$

eq:rdeux

Recall that when $A = A^*$, the spectral radius is given by $r(A) = \|A\|$. Then $\|f_\pm\|_\infty \leqslant r(A) = \|A\|$.

Let us prove that $f_+(A) \in \mathfrak{A}^+$. From the continuous calculus and the fact that $f_+(A)^* = f_+(A)$, we find that $\sigma(f_+(A)) \subset \mathbb{R}^+$. Since $A \in \mathfrak{A}_\mathbb{R}$, we know that $\sigma(A) \subset \mathbb{R}$ and thus that $f_+(\sigma(A)) \subset \mathbb{R}^+$. From equation part (1) of the continuous functional calculus, theorem 72.23, we conclude that $\sigma(f_+(A)) \subset \mathbb{R}^+$ and then that $f_+(t)f_-(t) = 0$. □

lem:rtrois

Lemma 72.30.
If $-C^*C \in \mathfrak{A}^+$ for $C \in \mathfrak{A}$, then $C = 0$.

Proof. We can decompose $C = D + iE$ with D, $E \in \mathfrak{A}_\mathbb{R}$; then

$$C^*C = 2D^2 + 2E^2 - CC^*. \tag{72.35}$$

eq:rquare

If $z \neq 0$ and $AB - z$ is invertible, then $BA - z^{-1}\mathbb{1}$ is invertible and $(BA - z)^{-1} = B(AB - z)^{-1}A - z^{-1}\mathbb{1}$. Then $\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}$ and $\sigma(C^*C) \subset \mathbb{R}^-$ imply $\sigma(-CC^*) \subset \mathbb{R}^+$. Now all terms of the right hand side of (72.35) are in $\mathfrak{A}^+$ and $C^*C \in \mathfrak{A}^+$. Since the assumption is $-C^*C \in \mathfrak{A}^+$, we conclude that $\sigma(C^*C) = 0$ and $C = 0$. □

ThoElsPositifsBBstar

Theorem 72.31.
The set of positive elements in $\mathfrak{A}$ is given by

$$\mathfrak{A}^+ = \{A^2 \text{ tel que } A \in \mathfrak{A}_\mathbb{R}\} = \{B^*B \text{ tel que } B \in \mathfrak{A}\} \tag{72.36}$$

when $\mathfrak{A}$ is an unital C^*-algebra.

Proof. If $A \in \mathfrak{A}^+$, one can define $\sqrt{A} \in \mathfrak{A}_\mathbb{R}$ in the same way as in proposition 72.23 with $f = \sqrt{\cdot}$. With this definition we have $(\sqrt{A})^2 = A$, so that $\mathfrak{A}^+ \subset \{A^2 \text{ tel que } A \in \mathfrak{A}_\mathbb{R}\}$.

Using the linearity of the involution term by term in the formula $\sqrt{A} = \sum_k c_k A^k$ shows that $\sqrt{A} \in \mathfrak{A}_\mathbb{R}$ when $A = A^*$.

For the inverse inclusion, consider $A \in \mathfrak{A}_\mathbb{R}$. Since $A = A^*$, we have $\sigma(A) \subset \mathbb{R}$. Using formula $\sigma(f(A)) = f(\sigma(A))$ with $f(t) = t^2$, we find $\sigma(A^2) = \sigma(A)^2 \subset \mathbb{R}^+$. The first equality is proved.

For the second equality, we begin by applying lemma 72.29 to B^*B, let $B^*B = A_+ - A_-$. From equations (72.34) we see that $A_+ - A_- = -A_-$. Then $(A_-)^3 = -A_-(A_+ - A_-)A_- = -A_-B^*BA_- = -(BA_-)^*BA_-$. Since AA_- is positive, $\sigma(A_-) \subset \mathbb{R}^+$. Using the continuous calculus with $f(t) = t^3$, it proves that $(A_-)^3 \geqslant 0$ and thus that $-(BA_-)^*BA_- \geqslant 0$. Lemma 72.30 shows that $BA_- = 0$.

This proves that $(A_-)^3 = 0$. From the continuous functional calculus with $f(t) = t^{1/3}$, it proves that $A_- = 0$ and then $B^*B = A_+ \in \mathfrak{A}^+$. □

Corollary 72.32.
When $A_1, A_2 \in \mathfrak{A}_\mathbb{R}$ and $B \in \mathfrak{A}$, if $A_1 \leqslant A_2$, then $B^*A_1B \leqslant B^*A_2B$.

Proof. The assumption is $A_2 - A_1 \geqslant 0$, but from theorem 72.31, there exists $A_3 \in \mathfrak{A}$ such that $A_2 - A_1 = A_3^*A_3$. The same property shows that $(A_3B)^*A_3B \geqslant 0$. This gives the corollary. □

Corollary 72.33.
For all A, $B \in \mathfrak{A}$, we have

$$B^*A^*AB \leqslant \|A\|^2 B^*B.$$

cor:BeAAeB

Proof. The inequality $-\|A\|\mathbb{1} \leqslant A \leqslant \|A\|\mathbb{1}$ holds when $A = A^*$. Let us write it for A^*A and recall that $\|A^*A\| = \|A\|^2$ in all C^*-algebra. Then $A^*A \leqslant \|A\|^2\mathbb{1}$ and by applying the previous corollary, we find $B^*A^*AB \leqslant \|A\|^2 B^*B$ □

An element $A \in \mathfrak{A}$ is a **projection** if $A = A^*$ and $A^2 = A$. In the case of a C^*-algebra of linear operators acting on a vector space, if x is an eigenvector of the projection A with the eigenvalue λ, then $Ax = \lambda x$ and $A^2x = \lambda^2 x = \lambda x$. Thus 1 is the only eigenvalue of a projection (or zero, which is the kernel). In particular a projection is positive and reads $A = B^*B$ for some $B \in \mathfrak{A}$ by theorem 72.31. PgProjPositif

¡¡ Avertissement/question à la lectrice !! 72.34
I think that the notation $\mathfrak{A}_{\mathbb{R}}$ stand for the elements with real spectrum. I have to check it and add to the notation index.

Proposition 72.35.
An element $A \in \mathfrak{A}_{\mathbb{R}}$ is positive if and only if the Gelfand transform $\hat{A}$ is pointwise positive in $C(\sigma(A))$.

Proof. Necessary condition. We know from theorem 72.7, (3) that

$$\sigma(A) = \sigma(\hat{A}) = \{\hat{A}(\omega) \text{ tel que } \omega \in \Delta(\mathfrak{A})\},$$

but $\sigma(A) \subset \mathbb{R}^+$ if A is positive.
Sufficient condition. From hypothesis, $A \in \mathfrak{A}_{\mathbb{R}}$ and $A^* = A$. We have to see that positivity of $\hat{A}$ implies $\sigma(A) \subset \mathbb{R}^+$. From point (3) of theorem 72.21, the function $\hat{A}\colon \sigma(A) \to \mathbb{R}$ is identity and positive, $\sigma(A) \subset \mathbb{R}^+$.

□

PropAplusConvexCone

Proposition 72.36.
The set $\mathfrak{A}^+$ of positive elements of the C^*-algebra $\mathfrak{A}$ is a convex cone (see definition 51.187).

Note that the C^*-algebra $\mathfrak{A}$ has to be commutative in order the Gelfand transform to be defined. It is supposed unital too.

Proof. We have to check the 3 points of definition 51.187.

(1) We know that if $A = A^*$ and $f \in C(\sigma(A))$, the commutator $[\sigma, f]$ is zero; as a particular case $\sigma(tA) = t\sigma(A)$. Then for $t > 0$, the element tA is positive.

(2) The fact that $\sigma(A) \subset [0, r(A)]$ implies that for all $t \in \sigma(A)$ and for all $c \geqslant r(A)$, $|c - t| \leqslant c$. Now we study the quantity

$$\sup_{t\in\sigma(A)} |c1_{\sigma(A)} - \hat{A}|.$$

The function $1_{\sigma(A)}$ is 0 or 1 following the argument belongs to $\sigma(A)$ or not while $\hat{A}(t) = t$ in $\sigma(A)$. Then

$$\sup_{t\in\sigma(A)} |c1_{\sigma(A)}(t) - \hat{A}(t)| = \sup_{t\in\sigma(A)} |c - t| \leqslant c.$$

This shows that

$$\|c1_{\sigma(A)} - \hat{A}\|_\infty \leqslant c \tag{72.37}$$ eq:cunhatA

for all $c > r(A)$ and then for all $c > \|A\|$. Since $\mathfrak{A}$ is commutative and $A = A^*$, we know that $\|\hat{A}\|_\infty = \|A\|$ from (72.13). Taking the inverse Gelfand transform of equation (72.37), we find

$$\|c\mathbb{1} - A\| \leqslant c \tag{72.38}$$ eq:normcin

for all $c \geqslant \|A\|$. Be careful on a point: the inverse Gelfand transform is not taken into $\mathfrak{A}$, but into $C^*(A, \mathbb{1})$ which is commutative and unital and then fulfills $\|\hat{A}\|_\infty = \|A\|$.

We know that the norm of $f(A)$ in $\mathfrak{A}$ and in $C^*(A,\mathbb{1})$ are the same, namely equation (72.38) is a relation for the norm of $c\mathbb{1}-A$ in $C^*(A,\mathbb{1})$. Until now we had proved that if $\sigma(A)\subset\mathbb{R}^+$, then $\|c\mathbb{1}-A\|\leqslant c$ for all $c\geqslant\|A\|$.

Taking the inverse argument, we can say that if $\|c\mathbb{1}-A\|\leqslant c$ for a certain $c\geqslant\|A\|$, then $\sigma(A)\subset\mathbb{R}^+$. Indeed the Gelfand transform of the assumption gives $\|cA_{\sigma(A)}-\hat{A}\|_\infty\leqslant c$, i.e. $\sup_{t\in\sigma(A)}|c1_{\sigma(A)}A-\hat{A}|\leqslant c$. As $\hat{A}$ is identity on $\sigma(A)$, for all $t\in\sigma(A)$, we have $|c-t|\leqslant c$. This shows that $t>0$ for all $t\in\sigma(A)$. Thus $\sigma(A)\subset\mathbb{R}^+$.

Let us now take $A+B$ instead of A and $c=\|A\|+\|B\|$. Remark that $c\geqslant\|A+B\|$. We have

$$\|c\mathbb{1}-(A+B)\|\leqslant\|(\|A\|-A)\|+\|(\|B\|-B)\|$$

where $\|A\|-A=r\mathbb{1}-A$ with $r=\|A\|$. On the other hand, $\|r\mathbb{1}-A\|\leqslant r$ for all $r\geqslant\|A\|$, then we can apply the first result to get

$$\|c\mathbb{1}-(A+B)\|\leqslant\|A\|+\|B\|=c$$

with $c\geqslant\|A+B\|$. Then the inverse argument gives $\sigma(A+B)\subset\mathbb{R}^+$ and $A+B\in\mathfrak{A}^+$.

(3) If $A\in\mathfrak{A}^+\cup(-\mathfrak{A}^+)$. Then $\sigma(A)\subset\mathbb{R}^+$ and $\sigma(A)\subset\mathbb{R}^-$; we conclude that $\sigma(A)=\{0\}$. Since $\|A\|=r(A)$, this gives $\|A\|=0$.

□

Proposition 72.37.
Let E be a real locally convex space and C a closed convex cone with top on 0 [5] *and $x\in E$, $x\notin C$. Then there exists a continuous linear function $f\colon E\to\mathbb{R}$ such that*

— *$f\geqslant 0$ on C,*

— *$f(x)<0$.*

Proof. It is possible to find a continuous linear form f and a real α such that $f(y)\geqslant\alpha$ on C and $f(x)<\alpha$ (see Urysohn lemma 48.23). We have $0=f(0)\geqslant\alpha$, so $f(x)<0$. If $f(y)<0$ fora $y\in C$, we find $f(\lambda y)<\alpha$ for a large enough λ. This is absurd and we conclude that $f\geqslant 0$ on C. □

72.5 Ordering relation

We know from proposition 72.36 that the set $\mathfrak{A}^+$ of positive elements in the C^*-algebra $\mathfrak{A}$ is a convex cone in $\mathfrak{A}$. We saw in subsection51.12.3 that in a real vector space a convex cone is equivalent to a notion of positivity and to a partial ordering. Here we consider the real vector space $\mathfrak{A}_\mathbb{R}$ and we say that $A\leqslant B$ if when $B-A\in\mathfrak{A}^+$. Thus we write

$$\mathfrak{A}^+=\{A\in\mathfrak{A}_\mathbb{R}\text{ tel que }A\geqslant 0\}. \tag{72.39}$$

PropAAsmAuAAu

Proposition 72.38.
If $A=A^*$, then

$$-\|A\|\mathbb{1}\leqslant A\leqslant\|A\|\mathbb{1}.$$

Proof. Since $A=A^*$ and $\|A\|\in\mathbb{R}^+$, we know that $\|A\|\mathbb{1}-A\in\mathfrak{A}_\mathbb{R}$. We have to show that $\|A\|\mathbb{1}-1$ is positive. In other words we have to show that its Gelfand transform is pointwise positive in $C(\sigma(A))$. We consider $\|A\|\mathbb{1}-A$ as an element of $C^*(A,\mathbb{1})$ and the Gelfand transform gives

$$\|A\|1_{\sigma(A)}-\hat{A}.$$

We will explain just after the proof why we write $1_{\sigma(A)}$ instead of 1. If one apply this on an element of $\sigma(A)$, one has to remember that $\hat{A}(t)=t$, so $\hat{A}(t)$ is at most $r(A)$ because $\sigma(A)\subset\mathbb{R}^+$. Since $r(A)\leqslant\|A\|$, one sees that

$$\big(\|A\|1_{\sigma(A)}-\hat{A}\big)(t)=\|A\|-t, \tag{72.40}$$

but $t\in\sigma(A)$ implies $t\leqslant r(A)$. Then the latter expression is positive and $\|A\|\mathbb{1}-A\in\mathfrak{A}^+_\mathbb{R}$.

□

5. Par top je veux dire le sommet du cône je ne sais pas comment dire en anglais.

Let us now see why $\hat{\mathbb{1}} = 1_{\sigma(A)}$ in $C^*(A, \mathbb{1})$. First, we consider $\mathbb{1} \in C^*(A, \mathbb{1})$; from properties of characters, $\hat{\mathbb{1}}(\omega) = \omega(\mathbb{1}) = 1$. So considering $\mathbb{1} \in C^*(A, \mathbb{1})$, the character $\hat{\mathbb{1}}$ is defined by $\hat{\mathbb{1}}(\omega) = 1 \in \mathbb{R}$ for all $\omega \in \Delta(C^*(A, \mathbb{1}))$. Now, we know that there exists an homeomorphism between $\Delta(C^*(A, \mathbb{1}))$ and $\sigma(A)$. Under this homeomorphism, $\mathbb{1}(t) = t$ for all $t \in \sigma(A)$.

Since there is a bijection between $\sigma(A)$ and $\Delta(C^*(A, \mathbb{1}))$, it should make no sense to define $\hat{\mathbb{1}}$ outside of $\sigma(A)$.

prop:mBABineq

Proposition 72.39.
If $-B \leqslant A \leqslant B$, then $\|A\| \leqslant \|B\|$.

Proof. The relations $-B \leqslant A \leqslant B$ and $-\|B\|\mathbb{1} \leqslant B \leqslant \|B\|\mathbb{1}$ give $-B \leqslant A \leqslant B \leqslant \|B\|\mathbb{1}$. But $B \leqslant \|B\|\mathbb{1}$ implies $-\|B\|\mathbb{1} \leqslant -B$ from definition of a linear partial order. Then $-\|B\|\mathbb{1} \leqslant A \leqslant \|B\|\mathbb{1}$. All this make that $\sigma(A) \subseteq [-\|B\|, \|B\|]$ and then that $\|A\| \leqslant \|B\|$ because $r(A) = \|A\|$ when $A = A^*$ and $r(A) = \sup\{|z| \text{ tel que } z \in \sigma(A)\}$. □

Proposition 72.40.
Let A, $B \in \mathfrak{A}^*$ such that $\|A + B\| \leqslant k$. Then $\|A\| \leqslant k$.

Proof. From assumptions, $A = A^*$, the $A \leqslant \|A\|\mathbb{1}$. Taking this relation with $A + B$ instead of A, we get $A + B \leqslant \|A + B\|\mathbb{1} \leqslant k\mathbb{1}$. From linearity of the partial order, this implies $0 \leqslant A \leqslant k\mathbb{1} - B$. Since $k \geqslant 0$, $-k\mathbb{1} \leqslant 0$ and $k\mathbb{1} - B \leqslant k\mathbb{1}$ because $0 \leqslant B$. Finally, proposition 72.39 makes that $\|A\| \leqslant k\|\mathbb{1}\|$ implies $\|A\| \leqslant k$. □

PropMDfqcUs

Proposition 72.41 ([986])**.**
Let $\mathfrak{A}$ be an involutive Banach algebra, $\mathcal{B}$ a C-algebra and $\pi\colon \mathfrak{A} \to \mathcal{B}$ a morphism. Then $\forall A \in \mathfrak{A}$,*

$$\|\pi(A)\| \leqslant \|A\| \tag{72.41}$$

eq_morleqpi

and π is continuous.

Proof. If $B \in \mathcal{B}$ is hermitian, $\|B^2\| = \|B^*B\| = \|B\|^2$. An induction shows that $\|B^{2n}\|^{-2n} = \|B\|$. With $n \to \infty$, the left hand side goes to the spectral radius (60.68). Then

$$r(B) = \|B\| \tag{72.42}$$

when B is hermitian.

Let us now take $A \in \mathfrak{A}$. We have $\sigma_{\mathcal{B}}(\pi(A)) \subset \sigma_{\mathfrak{A}}(A)$ because $(\pi(A) - \lambda\mathbb{1})v = \mathbb{1}$ for a $v \neq 0$ in $\mathcal{B}$ implies $(A - \lambda\mathbb{1})\pi^{-1}(v) = \mathbb{1}$ with $\pi^{-1}(v) \neq 0$ (because $\pi(\mathbb{1}) = \mathbb{1}$).

Remark that $(A - \lambda\mathbb{1})(a) = 0$ with $a \neq 0$

The continuity of π is now a consequence of lemma 48.4. □

prop:vp_geq

Proposition 72.42 ([986])**.**
Let $\mathfrak{A}$ be a C-algebra $\mathcal{B}$ an involutive normed algebra and $\varphi\colon \mathfrak{A} \to \mathcal{B}$ an injective morphism. Then $\forall A \in \mathfrak{A}$,*

$$\|\varphi(A)\| \geqslant \|A\|. \tag{72.43}$$

Proof. Consider $A \in \mathfrak{A}$, and suppose that $\|\varphi(A^*A)\| \geqslant \|A^*A\|$. Then

$$\|A^2\| = \|A^*A\| \leqslant \|\varphi(A^*A)\| = \|\varphi(A)^*\varphi(A)\| \leqslant \|\varphi(A)\|^2,$$

so that we can only consider the case where A is hermitian. One can also consider only the restriction of φ to the subC*-algebra generated by A and then suppose that $\mathfrak{A}$ is commutative. In the same way, we consider only $\varphi(\mathfrak{A})$ instead of $\mathcal{B}$ so that $\mathcal{B}$ can also be considered as commutative. In other words, we consider $C^*(A, \mathfrak{A})$ and $C^*(\varphi(A), \mathbb{1})$ instead of $\mathfrak{A}$ and $\mathcal{B}$. But from now to the end of the proof, we will often write it as $\mathfrak{A}$ and $\mathcal{B}$.

Now, we consider the completion of $\mathcal{B}$ and we add an unit to $\mathfrak{A}$ and $\mathcal{B}$. Then the structure spaces $\mathcal{S} := \Delta(C^*(A, \mathbb{1}))$ and $\mathcal{T} := \Delta(C^*(\varphi(A), \mathbb{1}))$ are compacts in the weak topology.

If $\chi \in \mathcal{T}$, it is clear that $\chi \circ \varphi$ is a character because these are both multiplicative. Then $\chi \circ \varphi \in \mathcal{S}$, and we denotes it by $\varphi'(\chi)$. It is easy to prove that $\varphi' \colon \mathcal{T} \to \mathcal{S}$ is a continuous map. Indeed, an open set in $\mathcal{S} = \Delta(C^*(A, \mathbb{1}))$ is of the form

$$\hat{B}^{-1}(\mathcal{O}) = \{\omega \in \mathcal{S} : \omega(B) \in \mathcal{O}\}$$

for $B \in C^*(B, \mathbb{1})$. But

$$\begin{aligned} \varphi'^{-1}(\hat{B}^{-1}(\mathcal{O})) &= \{\chi \in \Delta(\mathcal{B}) : \varphi'(\chi) \in \hat{B}^{-1}(\mathcal{O})\} \\ &= \{\chi \in \Delta(\mathcal{B}) : (\chi \circ \varphi)(B) \in \mathcal{O}\} \\ &= \widehat{\varphi(B)}^{-1}(\mathcal{O}), \end{aligned} \tag{72.44}$$

which is an open subset of $\mathcal{T} = \Delta(\mathcal{B})$.

Thus $\varphi'(\mathcal{T})$ is a compact subset of $\mathcal{S}$.

Suppose that $\varphi'(\mathcal{T}) \neq \mathcal{S}$, and consider two functions f, g on $\mathcal{S}$ such that $fg = 0$, but $f \neq 0$ and $g = 1$ on $\varphi'(\mathcal{T})$. The Urysohn lemma gives a function f which works for it: $f = 0$ on $\varphi'(\mathcal{T})$ and $f \neq 0$ out of $\varphi'(\mathcal{T})$. Now, f, $g \in C(\Delta(\mathfrak{A}))$ but we have gives a homomorphism between $\mathfrak{A}$ and $C(\Delta(\mathfrak{A}))$. Then we have A, $B \in \mathfrak{A}$ such that $\hat{A} = f$, $\hat{B} = g$ with $AB = 0$, $x \neq 0$, $\chi(\varphi(y)) = 1$ for any $\chi \in \mathcal{T}$. From this last equality one conclude that $\varphi(B)$ is invertible in $\mathcal{B}$. But $\varphi(A)\varphi(B) = 0$ and $\varphi(A) \neq 0$, which contradict.

Thus $\varphi'(\mathcal{T}) = \mathcal{S}$, so that $\forall A \in \mathfrak{A}$, $\forall \xi \in \mathcal{S}$, there exists $\chi \in \mathcal{T}$ such that $\xi(A) = \varphi'(\chi)(A)$. Finally, for any $x \in \mathfrak{A}$,

$$\|A\| = \sup_{\xi \in \mathcal{S}} |\xi(A)| = \sup_{\chi \in \mathcal{T}} |\varphi'(\chi)(A)| = \sup_{\chi \in \mathcal{T}} |\chi \circ \varphi(A)| \leqslant \|\varphi(A)\|. \tag{72.45}$$

□

Proposition 72.43.
Every C^*-algebra isomorphism is isometric.

Proof. This is a combination of propositions 72.41 and 72.42. □

lem:injmorpisom

Lemma 72.44.
An injective C^*-algebra morphism is is always an isometry and in particular, the image is closed.

Proof. Let φ be the morphism and suppose that there exists a $B \in \mathfrak{A}$ such that $\|\varphi(B)\| \neq \|B\|$. Then

$$\|\varphi(B^*B)\| = \|\varphi(B)\|^2 \neq \|B\|^2 = \|B^*B\|.$$

We pose $A = B^*B$ and we know that $\|A\| = \sqrt{r(A^*A)}$ because $A^* = A$. By definition, $r(B) = \sup\{|z| \text{ tel que } z \in \sigma(B))\}$. If we apply $\sigma(f(A)) = f(\sigma(A))$ to $f(t) = t^2$, we conclude that $\sigma(A) \neq \sigma(\varphi(A))$. We saw, during proof of 72.41, that $\sigma(\varphi(A)) \subseteq \sigma(A)$. Then we have a strict inclusion

$$\sigma(\varphi(A)) \subset \sigma(A).$$

Thus there exists a continuous function $f \neq 0$ on $\sigma(A)$ such that $f(x) = 0$ when $x \in \sigma(\varphi(A))$. In particular $f(\varphi(A)) = 0$ and by 72.3.1, we see that $\varphi(f(A)) = 0$. This is in contradiction with the injectivity of φ.

□

Corollary 72.45.
The image of a C^*-algebra morphism $\varphi \colon \mathfrak{A} \to \mathcal{B}$ is closed. In particular $\varphi(\mathfrak{A})$ is a C^*-subalgebra of $\mathcal{B}$.

Proof. We define $\psi \colon \mathfrak{A}/\ker\varphi \to \mathcal{B}$ by $\psi([a]) = \varphi(a)$. It is a (bijective) vector space isomorphism, and $\varphi = \psi \circ \tau$. Since φ is a morphism, $|\varphi\rangle$ is an ideal and then $\mathfrak{A}/\ker\varphi$ is a C^*-algebra. Now $\varphi = \psi \circ \tau$ is an injective morphism of C^*-algebra and its image is closed: $\psi(\mathfrak{A}/|\varphi\rangle)$ is closed in $\mathcal{B}$. Since φ is a morphism, its range is a $*$-algebra in $\mathcal{B}$ which is closed for the norm in $\mathcal{B}$. With induced operation from $\mathcal{B}$, $\varphi(\mathfrak{A})$ becomes a C^*-algebra. □

72.6 Approximate unit

Definition 72.46.
*A **approximate unit** of a normed algebra $\mathfrak{A}$ is a family $(u_i)_{i\in I}$ (where I is an increasing filtered set, in order to makes sense to $i \to \infty$) such that*

(1) $\forall i$, $\|u_i\| \leqslant 1$,

(2) $\forall A \in \mathfrak{A}$, $\|u_i A - A\| \to 0$ *and* $\|Au_i - A\| \to 0$. [enuoii]

def:app_unit

It is clear that if $\mathfrak{A}$ has an unit $\mathbb{1}$, the choice $u_i = \mathbb{1}\ \forall i$ is an approximate unit.

Example

Let $C_0(\mathbb{R})$, the C^*-algebra of continuous functions on $\mathbb{R}$ such that for all $\varepsilon > 0$, there exists a compact K outside of which $|f(x)| < \varepsilon$. This has no unit, but we can build an approximate unit as following. Let $n \in \mathbb{N}$ and define

$$\mathbb{1}_n = \begin{cases} 1 & \text{on } [-n, n] \\ 0 & \text{when } |x| > n+1 \\ \text{continuous} & \text{otherwhise} \end{cases}$$

It should be noted that the limit $\lim_{n\to\infty} \mathbb{1}_n$ in the norm $\|.\|_\infty$ is *not* $1_{\mathbb{R}}$.

Proposition 72.47.
All non unital C^*-algebra admits an approximate unit. If $\mathfrak{A}$ is separable, then Λ can be chosen countable.

Proof. Let Λ be the set of finite subsets of $\mathfrak{A}$ endowed with the partial ordering given by inclusion. To each $\lambda = \{A_1, \dots, A_n\} \in \Lambda$, we define $B_\lambda = \sum_{i=1}^n A_i^* A_i$. From construction, $B_\lambda^* = B_\lambda$ and $B_\lambda \in \mathfrak{A}^+$. The latter point gives $\sigma(B_\lambda) \subset \mathbb{R}^+$. Si for all $z \in \mathbb{R}^-$, the element $A - z\mathbb{1}$ is invertible (in $\mathfrak{A}_\mathbb{1}$) and $n^{-1}\mathbb{1} + B_\lambda$ is invertible when $n < 0$. It allows us to define

$$\mathbb{1}_\lambda = B_\lambda (n^{-1}\mathbb{1} + B_\lambda)^{-1}.$$

Since $B_\lambda = B_\lambda^*$ and B_λ commutes with all functions of itself (as $(n^{-1}\mathbb{1} + B_\lambda)^{-1}$), it leads to $\mathbb{1}_\lambda = \mathbb{1}_\lambda^*$. The element $(n^{-1}\mathbb{1} + B_\lambda)^{-1}$ is computed in $\mathfrak{A}_\mathbb{1}$, the it can be written as $C + \mu\mathbb{1}$ for certain $C \in \mathfrak{A}$ and $\mu \in \mathbb{C}$. However, $\mathbb{1}_\lambda \in \mathfrak{A}$. Indeed

$$\mathbb{1}_\lambda = B_\lambda(C + \mu\mathbb{1}) = B_\lambda C + \mu B_\lambda \in \mathfrak{A}.$$

Now we use the continuous functional calculus on B_λ with $f(f) = \frac{t}{n^{-1}+t}$; more precisely, we will use formula $\sigma(f(B_\lambda)) = f(\sigma(B_\lambda))$ where we know that $\sigma(B_\lambda) \subset \mathbb{R}^+$. We find $f(B_\lambda) = B_\lambda(n^{-1}\mathbb{1} + B_\lambda)^{-1} = \mathbb{1}_\lambda$. Then

$$f(\sigma(B_\lambda)) \subset f(\mathbb{R}^+) = [0, 1].$$

In particular, $\sigma(\mathbb{1}_\lambda) \subset [0, 1]$.

In order to prove the second condition of definition 72.46, we pose $C_i = \mathbb{1}_\lambda A_i - A_i$. One can prove that [6]

$$\sum_{i=1}^n C_i C_i^* = n^{-2} B_\lambda (n^{-1} + t)^{-2}.$$

We consider the function $f(t) = n^{-2} t (n^{-1} + t)^{-2}$ on $\sigma(B_\lambda)$, then $f \geqslant 0$ and in particular

$$\sup_{t\in\mathbb{R}^+} |f(t)| = \frac{1}{4n}$$

6. Je ne vois pas comment.

and $f(t)$ takes its maximum at $t = 1/n$. Since f is defined on $\sigma(B_\lambda)$ which is a part of $\mathbb{R}^+$, we have $\|f\|_\infty \leqslant 1/4n$. Then $\|n^{-2}B_\lambda(n^{-1}\mathbb{1} + B_\lambda)^{-2}\| \leqslant 1/4n$ and

$$\|\sum_i C_i C_i^*\| \leqslant \frac{1}{4n}.$$

In particular, for each i, we have $\|C_i C_i^*\| \leqslant \frac{1}{4n}$.

Now consider $A \in \mathfrak{A}$. Then A belongs to a $\lambda \in \Lambda$ and we can build a directed subset of Λ in which A is always present. For $n \to \infty$,

$$\begin{aligned} \lim_{\lambda\to\infty} \|\mathbb{1}_\lambda - A\|^2 &= \lim_{\lambda\to\infty} \|(\mathbb{1}_\lambda A - A)^*(\mathbb{1}_\lambda A - A)\| \\ &= \lim_{\lambda\to\infty} \|C_i C_i^*\| \\ &= 0 \end{aligned} \tag{72.46}$$

because $\lambda \to \infty$ needs $n \to \infty$ and $\|C_i C_i^*\| \leqslant 1/4n$.

If $\mathfrak{A}$ is separable, then we can find a countable set of A_i dense in $\mathfrak{A}$. We build Λ from them and the proof works because when $\lambda \to \infty$, any $A \in \mathfrak{A}$ is reached from density.

□

lem:taulimA

Lemma 72.48.
Let $\{\mathbb{1}_\lambda\}$ be an approximate unit in the ideal $\mathfrak{I}$ and $A \in \mathfrak{A}$. Then

$$\|\tau(A)\| = \lim_{\lambda\to\infty} \|A - A\mathbb{1}_\lambda\|. \tag{72.47}$$

Proof. First remark that property (2) of definition of an approximate unit don't imply that the above limit is zero because it holds when $\{\mathbb{1}_\lambda\}$ is an approximate unit in $\mathfrak{A}$.

Definition $\|\tau(A)\| := \inf_{J\in\mathfrak{I}} \|A + J\|$ immediately gives

eq:ri512_2

$$\|\tau(A)\| \leqslant \|A - A\mathbb{1}_\lambda\|. \tag{72.48}$$

In order to prove the inverse inequality, we add (if necessary) an unit to $\mathfrak{A}$ and we consider a J in $\mathfrak{I}$. Then

eq:ri512

$$\begin{aligned} \|A - A\mathbb{1}_\lambda\| &= \|(A + J)(\mathbb{1} - \mathbb{1}_\lambda) + J(\mathbb{1}_\lambda - \mathbb{1})\| \\ &\leqslant \|A + J\|\|\mathbb{1} - \mathbb{1}_\lambda\| + \|J\mathbb{1}_\lambda - J\|. \end{aligned} \tag{72.49}$$

Remember equation (72.38): for all $c \geqslant \|A\|$, we have $\|c\mathbb{1} - A\| \leqslant c$. Let us write it with $\mathbb{1}_\lambda$ instead of A and $c = 1$:

$$\|\mathbb{1} - \mathbb{1}_\lambda\| \leqslant 1.$$

On the other hand, equation (72.49) with $\lambda \to \infty$ (and then with $\|J\mathbb{1}_\lambda - J\| \to 0$) gives

$$\lim_{\lambda\to\infty} \|A - A\mathbb{1}_\lambda\| \leqslant \|A + J\|.$$

From definition of the norm on $\mathfrak{A}/\mathfrak{I}$ and of an infimum, for all $\varepsilon > 0$, there exists a $J \in \mathfrak{I}$ such that $\|\tau(A)\| + \varepsilon \geqslant \|A + J\|$. Let us fix an ε and such a J. Then, using equation (72.48) ,

$$\|A + J\| - \varepsilon \leqslant \|\tau(A)\| \leqslant \|A - A\mathbb{1}_\lambda\|,$$

and then

$$\lim_{\lambda\to\infty} \|A - A\mathbb{1}_\lambda\| - \varepsilon \leqslant \|\tau(A)\| \leqslant \|A - A\mathbb{1}_\lambda\|,$$

letting $\varepsilon \to 0$, it gives the lemma:

$$\|\tau(A)\| \geqslant \lim_{\lambda\to\infty} \|A - A\mathbb{1}_\lambda\|.$$

□

prop:ideal_Banach

Proposition 72.49.
If $\mathfrak{I}$ is an ideal in a Banach algebra $\mathfrak{A}$, then the quotient $\mathfrak{A}/\mathfrak{I}$ becomes a Banach algebra with the norm

$$\|\tau(A)\| = \inf_{J\in\mathfrak{I}} \|A+J\| \tag{72.50}$$

eq:norm_ideal

and the multiplication rule

$$\tau(A)\tau(B) = \tau(A)\tau(B). \tag{72.51}$$

eq:prod_ideal

Proof. The fact that it is a Banach space is non trivial[987]. We begin to prove that (72.51) is a well defined product:

$$\tau(A+J_1)\tau(A+J_2) = \tau(AB + AJ_2 + J_1B + J_1J_2) = \tau(AB) \tag{72.52}$$

because $AJ_2 + J_1B + J_1J_2 \in \mathfrak{I}$.

Next we have to prove that $\|\tau(A)\tau(B)\| \leqslant \|\tau(A)\|\|\tau(B)\|$. Remark that

$$\|\tau(A)\| \leqslant \|A\| \tag{72.53}$$

eq:tauAA

because

$$\|\tau(A)\| = \inf\|A+J\| \leqslant \inf(\|A\| + \|J\|) = \|A\| + \inf\|J\| = \|A\|.$$

Then $\forall\varepsilon > 0$, there exists a $J \in \mathfrak{I}$ such that

$$\|\tau(A)\| + \varepsilon \geqslant \|A+J\| \tag{72.54}$$

eq:tauAeps

It also gives $\|\tau(A)\| = \|\tau(A+J)\| \leqslant \|A+J\|$. Take A, $B \in \mathfrak{A}$ and a ε for which (72.54) holds for both A and B. Then

$$\begin{aligned}\|\tau(A)\tau(B)\| &= \|\tau\big((A+J_1)(A+J_2)\big)\| \\ &\leqslant \|(A+J_A)(B+J_2)\| \\ &\leqslant \|A+J_1\|\|B+J_2\| \\ &\leqslant (\|\tau(A)\|+\varepsilon)(\|\tau(B)\|+\varepsilon),\end{aligned} \tag{72.55}$$

which gives the result as $\varepsilon \to 0$.

□

tho_idautadjquo

Theorem 72.50.
Let $\mathfrak{I}$ be an ideal in the C-algebra $\mathfrak{A}$. Then*

1. *the ideal $\mathfrak{I}$ is selfadjoint; in other words, if $A \in \mathfrak{I}$, then $A^* \in \mathfrak{I}$,* enuni
2. *the quotient $\mathfrak{A}/\mathfrak{I}$ is a C*-algebra for the norm* enunii

$$\|\tau(A)\| := \inf_{J\in\mathfrak{I}} \|A+J\|,$$

the multiplication

$$\tau(A)\tau(B) := \tau(AB),$$

and the convolution

$$\tau(A)^* = \tau(A^*)$$

where $\tau\colon \mathfrak{A} \to \mathfrak{A}/\mathfrak{I}$ is the canonical projection.

Proof. It is necessary to prove (1) before (2) because the point (1) proves the well definiteness of the involution. Let $\mathfrak{I}^* = \{A^* \text{ tel que } A \in \mathfrak{I}\}$ and $J \in \mathfrak{I}$. Then $J^*J \in \mathfrak{I} \cap \mathfrak{I}^*$ because an ideal is left and right ideal (it is easy to see that $\mathfrak{I}^*$ is an ideal).

We are going to prove that $\mathfrak{I} \cap \mathfrak{I}^*$ is a C^*-subalgebra of $\mathfrak{A}$. If $J \in \mathfrak{I} \cap \mathfrak{I}^*$,then $J^* \in \mathfrak{I} \cap \mathfrak{I}^*$ too. On the other hand, $\mathfrak{I} \cap \mathfrak{I}^*$ is closed linear subspaces of $\mathfrak{A}$ because $\mathfrak{I}$ and $\mathfrak{I}^*$ are such from

definition of an ideal. Ideals are Banach space and it is clear that the conditions about norm and involutions are true in $\mathfrak{I} \cap \mathfrak{I}^*$.

So $\mathfrak{I} \cap \mathfrak{I}^*$ is a non unital C^*-algebra and posses an approximate unit $\{\mathbb{1}_\lambda\}$. Let $J \in \mathfrak{I}$; the following upper bound holds:

$$\begin{aligned}\|J^* - J^*\mathbb{1}_\lambda\|^2 &= \|(J - \mathbb{1}_\lambda J)(J^* - J^*\mathbb{1}_\lambda)\| \\ &= \|(JJ^* - JJ^*\mathbb{1}_\lambda) - \mathbb{1}_\lambda(JJ^* - JJ^*\mathbb{1}_\lambda)\| \\ &\leqslant \|JJ^* - JJ^*\mathbb{1}_\lambda\| + \|\mathbb{1}_\lambda(JJ^* - JJ^*\mathbb{1}_\lambda)\| \\ &\leqslant \|JJ^* - JJ^*\mathbb{1}_\lambda\| + \|\mathbb{1}_\lambda\|\|JJ^* - JJ^*\mathbb{1}_\lambda\|.\end{aligned} \tag{72.56}$$

Since JJ^* belongs to $\mathfrak{I} \cap \mathfrak{I}^*$ which is a C^*-algebra an [7] which $\mathbb{1}_\lambda$ is build, we recognize in the latter two terms something of the form $\|A\mathbb{1}_\lambda - A\|$ whose limit is zero. Then $\lim_{\lambda\to\infty} \|J^* - J^*\mathbb{1}_\lambda\| = 0$. We have proved that $J^*\mathbb{1}_\lambda \in \mathfrak{I} \cap \mathfrak{I}^*$ and in particular that $J^*\mathbb{1}_\lambda \in \mathfrak{I}$ for all λ. Then J^* is the limit of a sequence in $\mathfrak{I}$, but the latter is closed, then $J^* \in \mathfrak{I}$.

Now we prove (2). From proposition 72.49, the quotient space $\mathfrak{A}/\mathfrak{I}$ is a Banach algebra in the chosen product and norm. We just have to prove that $\tau(A)^* = \tau(A^*)$ gives a well behaved involution on $\mathfrak{A}/\mathfrak{I}$, i.e. we have to show that $\|A^*A\| = \|A\|^2$ for $A \in \mathfrak{A}/\mathfrak{I}$, or $\|\tau(A)^*\tau(A)\| = \|\tau(A)\|^2$ for $A \in \mathfrak{A}$. Using lemma 72.48, we can compute

$$\begin{aligned}\|\tau(A)^2\| &= \lim_{\lambda\to\infty} \|A - A\mathbb{1}_\lambda\|^2 && \text{The lemma} \\ &= \lim_{\lambda\to\infty} \|(A - A\mathbb{1}_\lambda)^*(A - A\mathbb{1}_\lambda) && \text{from } \|A\|^2 = \|A^*A\| \text{ in } \mathfrak{A}_\mathbb{1} \\ &= \lim_{\lambda\to\infty} \|(\mathbb{1} - \mathbb{1}_\lambda)A^*A(\mathbb{1} - \mathbb{1}_\lambda)\| \\ &\leqslant \lim_{\lambda\to\infty} \|\mathbb{1} - \mathbb{1}_\lambda\|\|A^*A(\mathbb{1} - \mathbb{1}_\lambda)\| \\ &\leqslant \lim \lambda \to \infty \|A^*A(\mathbb{1} - \mathbb{1}_\lambda)\| && \text{because } \|\mathbb{1} - \mathbb{1}_\lambda\| \leqslant 1 \\ &= \lim_{\lambda\to\infty} \|A^*A - A^*\mathbb{1}_\lambda\| \\ &= \|\tau(A^*A)\| && \text{definition of the norm} \\ &= \|\tau(A)^*\tau(A)\|.\end{aligned} \tag{72.57}$$

Now lemma 60.48 makes the Banach $\mathfrak{A}/\mathfrak{I}$ a C^*-algebra.

□

Corollary 72.51.
Morphism of C^*-algebras have following properties related to ideals:

(1) The kernel of a morphism between two C^*-algebras is an ideal. enupi

(2) Any ideal in a C^*-algebra is the kernel of a morphism. enupii

(3) Then any morphism has norm 1. enupiii

Proof. (1) If $A \in \ker\varphi$, then $AB \in \ker\varphi$ because $\varphi(AB) = \varphi(A)\varphi(B) = 0$. From proposition 72.41, a morphism is continuous. The kernel of a continuous map is always continuous by the "intermediate value" property.

(2) Let $\mathfrak{I}$ be an ideal in the C^*-algebra $\mathfrak{A}$ and $\tau\colon \mathfrak{A} \to \mathfrak{A}/\mathfrak{I}$, the canonical projection. We know that $\mathfrak{A}/\mathfrak{I}$ is a C^*-algebra and that τ is a morphism. Let us show that $\mathfrak{I} = \ker\tau$.

On the one hand, if $J \in \mathfrak{I}$, then $\tau(J) = \tau(0)$ which is the zero of $\mathfrak{A}/\mathfrak{I}$. Let, on the other hand, $J \in \ker\tau$. Then $\tau(J) = 0 = \tau(0)$. The fact that $\tau(J) = \tau(0)$ shows that the difference between 0 and J is an element of $\mathfrak{I}$. In other words: $J \in \mathfrak{I}$.

This and the fact that $\|\tau\| = 1$ give (3). [8]

□

7. Cette phase me semble avoir un problème.
8. Je ne vois cependant pas pourquoi.

Proposition 72.52.
Let $\mathfrak{A}$ be a C-algebra and M a two-sided ideal in $\mathfrak{A}$ everywhere dense in $\mathfrak{A}$. There exists an increasing filtering approximate unit of $\mathfrak{A}$ contained in M.*

If $\mathfrak{A}$ is separable, we can ask the approximate unit to be indexed by $\mathbb{N}$.

9

Proof. Let $\mathfrak{A}_{\mathbb{1}}$ be the algebra obtained by adding an unit to $\mathfrak{A}$, and Λ the set of finite parts of M ordered by inclusion. For $\lambda = \{A_1, \dots, A_n\}$, we pose

$$v_\lambda = A_1 A_1^* + \cdots + A_n A_n^* \in M$$

and

$$u_\lambda = v_\lambda \big(\frac{1}{n} + v_\lambda\big)^{-1}$$

with $A_i \in M$ and $\lambda \in \Lambda$. The fact that $A^* \in M$ when $A \in M$ comes from point (1) of theorem 72.50.

¡¡ Avertissement/question au lecteur !! 72.53
Cette démonstration n'est pas du tout finie.

□

The same kind of proof gives the following

Proposition 72.54.
Let $\mathfrak{A}$ be a C-algebra and $\mathcal{I}$ a right ideal in $\mathfrak{A}$. There exists, in $\mathcal{I} \cap \mathfrak{A}^+$, a family (u_λ) with λ in a filtered set such that*

(1) $\|u_\lambda\| \leqslant 1$,

(2) $\lambda \leqslant \mu$ *implies* $u_\lambda \leqslant u_\mu$,

(3) $\forall A \in \overline{\mathcal{I}}$, $\|u_\lambda A - A\| \to 0$.

The definition of "$A \leqslant B$" for A, B in a C^*-algebra comes from the positivity notion.

72.7 Ideal in Banach algebras

Definition 72.55.
*An **ideal** in a Banach algebra $\mathfrak{A}$ is a closed vector subspace $\mathfrak{I} \subseteq \mathfrak{A}$ such that $\forall I \in \mathfrak{I}$, $\forall A \in \mathfrak{A}$, $AI \in \mathfrak{I}$ and $IA \in \mathfrak{I}$.*

*A **left ideal** has just $AI \in \mathfrak{I}$.*

*A **maximum ideal** is an ideal $\mathfrak{I} \neq \mathfrak{A}$ for which there exists no strict intermediary ideal between $\mathfrak{I}$ and $\mathfrak{A}$, i.e. no ideal $\tilde{\mathfrak{I}} \neq \mathfrak{I}$ with $\tilde{\mathfrak{I}} \neq \mathfrak{A}$ and $\mathfrak{I} \subset \tilde{\mathfrak{I}}$.*

An ideal is a Banach algebra, and the only ideal in $\mathfrak{A}$ which contains an invertible element is $\mathfrak{A}$ itself.

If $\mathfrak{A}$ is unital, then $\mathfrak{A}/\mathfrak{I}$ is unital and his unit is given by $\tau(\mathbb{1})$. Let us prove that $\|\tau(\mathbb{1})\| = 1$. From equation (72.53), $\|\tau(\mathbb{1})\| \leqslant \|\mathbb{1}\| = 1$. On the other hand, from equation $\|\tau(A)\tau(B)\| \leqslant \|\tau(A)\|\|\tau(B)\|$ with $B = \mathbb{1}$, we find $1 \leqslant \|\tau(\mathbb{1})\|$.

Corollary 72.56.
Let $\mathfrak{A}$ and $\mathcal{B}$ be C-algebras, $\varphi\colon \mathfrak{A} \to \mathcal{B}$ a morphism and $\mathfrak{I}$ the kernel of φ. We consider the canonical decomposition of φ into*

$$\mathfrak{A} \xrightarrow{\tau} \mathfrak{A}/\mathfrak{I} \xrightarrow{\psi} \varphi(\mathfrak{A}) \longrightarrow \mathcal{B}. \tag{72.58}$$

Then $\mathfrak{I}$ is closed in $\mathfrak{A}$, $\varphi(\mathfrak{A})$ is closed in $\mathcal{B}$ and ψ is an isometric isomorphism.

9. Ici, y'a un point pas clair dans Landsman 2.7.2. De toutes facons, c'est une démonstration à refaire

Proof. A morphism of C^*-algebras is always continuous,then $\mathfrak{I} = \ker\varphi = \varphi^{-1}(\{0\})$ is closed because $\{0\}$ is closed. The map $\psi\colon \mathfrak{A}/\mathfrak{I} \to \varphi(\mathfrak{A})$ is injective because when $\psi([A]) = \psi([B])$, we have $\varphi(A) = \varphi(B)$ and $A = B + I$ for a certain $I \in \mathfrak{I}$, hence $[A] = [B]$. Proposition 72.42 gives $\|\varphi(A)\| \geqslant \|A\|$ while proposition 72.41 gives $\|\varphi(A)\| \leqslant \|A\|$. So φ is isometric.

Hence $\varphi(\mathfrak{A})$ is complete and therefore closed in $\mathcal{B}$. □

72.8 States

Theorem 72.57 (Riesz representation theorem)**.**
Let X be a locally compact Hausdorff space and $\Lambda\colon C_0(X) \to \mathbb{C}$ be a positive linear functional. Then there exists a σ-algebra $\mathcal{M}$ in X, which contains all Borel sets, and an unique measure μ on $\mathcal{M}$ such that

(1) $\Lambda f = \int_X f d\mu$,

(2) $\mu(K) < \infty$ *for all compact* $K \subset X$,

(3) If $E \in \mathcal{M}$, *then*

$$\mu(E) = \inf\{\mu(V) \text{ tel que } E \subset V,\ V \text{ open}\}$$

(4) If E is open or if $E \in \mathcal{M}$ with $\mu(E) < \infty$, then

$$\mu(E) = \sup\{\mu(K) \text{ tel que } K \subset E,\ K \text{ compact}\},$$

(5) The space $(X, \mathcal{M}, \mu)$ is a complete measure space. That is, if $E \in \mathcal{M}$, $A \subset E$ and $\mu(E) = 0$, then $A \in \mathcal{M}$.

72.9 States on unital C^*-algebras

DefStateUnital

Definition 72.58.
*A **state** on an* unital *C^*-algebra $\mathfrak{A}$ is a linear map $\omega\colon \mathfrak{A} \to \mathbb{C}$ which is*

— *positive:* $\omega(A) \geqslant 0$ *for every* $A \in \mathfrak{A}^+$,

— *normalised:* $\omega(\mathbb{1}) = 1$.

*We say that the state ω is **faithful** if $\omega(A^*A) = 0$ implies $A = 0$. We denote by $\mathcal{S}(\mathfrak{A})$ the space of all states on $\mathfrak{A}$.*

As an example, consider $\mathfrak{A} \subset \mathfrak{B}(\mathscr{H})$, a C^*-subalgebra of the bounded operators on the Hilbert space $\mathscr{H}$. Then any $\psi \in \mathscr{H}$ such that $\|\psi\| = 1$ defines a state on $\mathfrak{A}$ by $\psi(A) = \langle\psi, A\psi\rangle$. Indeed if $A \in \mathfrak{A}^+$, there exists a B such that $A = B^*B$ and from definition of the involution on $\mathfrak{B}(\mathscr{H})$, we have $\psi(B^*B) = \|B\psi\|^2 \geqslant 0$.

When a state is not faithful, the set on which $\omega(a^*a) = 0$ is a vector space, and one usually takes the quotient of $\mathfrak{A}$ by that subspace.

As an example, let us prove the following.

Propstaretattraces

Proposition 72.59.
Every state in the $*$-algebra $\mathbb{M}_n(\mathbb{C})$ of $n \times n$ matrices with complex coefficients read

$$\varphi(a) = \mathrm{Tr}(as) \tag{72.59}$$

for a certain positive matrix $s \in \mathbb{M}_n(\mathbb{C})$ with $\mathrm{Tr}(s) = 1$. Moreover, every functional of that form is a state.

Proof. Let us first prove the second claim. Let s be a positive matrix with trace equal to 1, we have

$$\varphi(a^*a) = \mathrm{Tr}(a^*as) = \mathrm{Tr}(a^*as^{1/2}s^{1/2}) = \mathrm{Tr}(s^{1/2}a^*as^{1/2})$$

which is positive. Here we used the fact that every positive element reads a^*a and positivity of s to define $s^{1/2}$.

By linearity, a state on $\mathbb{M}_n(\mathbb{C})$ must read $\varphi(a) = \sum_i \sum_j \varphi_{ij} a_{ij}$ for some coefficients φ_{ij}, so that $\varphi(a) = \mathrm{Tr}(sa)$ where s is the matrix of (φ_{ij}). The condition $\varphi(\mathbb{1}) = 1$ immediately imposes $\mathrm{Tr}(s) = 1$. Now we study $\varphi(a^*a) = \mathrm{Tr}(sa^*a)$. In the basis which diagonalises a^*a, we have $(a^*a)_{ij} = \delta_{ij} n_i$ (no sum) with $n_i > 0$. Thus

$$\varphi(a) = \sum_k (sa^*a)_{kk} = \sum_k s_{kk} n_k$$

which must be positive for every choice of positive n_k. That imposes for all s_{kk} to be positive, so that s is a positive matrix. □

Notice that s being positive, it is diagonalisable, so that it is nothing else than a list of positive numbers.

Theorem 72.60.
The space of states of $\mathfrak{A} = C(X)$ is the set of probability measures on X. We suppose that X is compact and Hausdorff.

Proof. If X is compact and Hausdorff, we know from a long time that $C(X)$ is an unital C^*-algebra with norm

$$\|f\|_\infty := \sup_{x \in X} |f(x)|.$$

We can apply the Riesz representation theorem: we have on X a σ-algebra $\mathcal{M}$ which contains all Borel sets and an (unique) measure μ on $\mathcal{M}$ such that, among other properties,

$$\omega(f) = \int_X f d\mu. \tag{72.60}$$

Since X is compact, it fulfills $\mu(X) < \infty$. We have $\omega(\mathbb{1}) = \int_X \mathbb{1} d\mu = \mu(X) = 1$. Then $\mu(X) = 1$ and it is a probability measure.

On the other hand, it is clear that a probability is a state.

□

Let $\omega \in \mathcal{S}$ be positive. Then the definition[PgStateInn]

$$(A, B)_\omega := \omega(A^*B) \tag{72.61}$$

[eq:defprodetat]

gives a **pre-inner product**[pgdef_preinned], i.e. an inner product without the condition $v^2 = 0 \Rightarrow v = 0$. Indeed, $(A, A)_\omega = \omega(A^*A)$, but $A^*A \in \mathfrak{A}^+$, then $\omega(A^*A) \geqslant 0$ because ω is a state. In the case of a faithful state, we get an inner product.

From Cauchy-Schwarz inequality,

$$|\omega(A^*B)|^2 \leqslant \omega(A^*A)\omega(B^*B). \tag{72.62}$$

[eq:omABleq]

Moreover, $\omega(A^*) = (A, \mathbb{1})_\omega$, then

$$\omega(A^*) = \overline{\omega(A)}. \tag{72.63}$$

[eq:omABlegs]

Proposition 72.61.
Let $\omega : \mathfrak{A} \to \mathbb{C}$, a linear map on an unital C^*-algebra $\mathfrak{A}$. It is positive if and only if ω is bounded and $\|\omega\| = \omega(\mathbb{1})$. In particular

(1) A state in an unital C^*-algebra is bounded and has norm 1, [71125ai]

(2) an element $\omega \in \mathfrak{A}^*$ such that $\|\omega\| = \omega(\mathbb{1}) = 1$ is a state on $\mathfrak{A}$. [7125aii]

[prop:linposboun]

Proof. Direct sense. Let ω be positive and begin by $A = A^*$. Then inequality $-\|A\|\mathbb{1} \leqslant A \leqslant \|A\|\mathbb{1}$ gives $\omega(A) \leqslant \|A\|\omega(1)$. Indeed

$$\begin{aligned} A \leqslant \|A\|\mathbb{1} &\Rightarrow A - \|A\|\mathbb{1} \leqslant 0 \\ &\Rightarrow \omega(\|A\|\mathbb{1} - A) \geqslant 0 \\ &\Rightarrow \|A\|\omega(\mathbb{1}) \geqslant \omega(A). \end{aligned} \tag{72.64}$$

The same can be done with $-\|A\|\mathbb{1} \leqslant A$ as starting point, then for $A = A^*$, we have $\|\omega(A)\| \leqslant \omega(\mathbb{1})\|A\|$.

Let us now consider any A. We know that $|\omega(A^*B)|^2 \leqslant \omega(A^*A)\omega(B^*B)$. Writing it for $A = \mathbb{1}$ and using $\|A^*A\| = \|A\|^2$, we find that $|\omega(B)|^2 \leqslant \omega(\mathbb{1})^2\|B\|^2$, and from the definition of the functional norm, it leads to

$$\|\omega\| \leqslant \omega(\mathbb{1}). \tag{72.65}$$

A C^*-algebra is a normed vector space, then for all $A \in \mathfrak{A}$, there exists a $R \in \mathfrak{A}$ such that $A = \|A\|R$ and $\|R\| = 1$. For this R, we have

$$\frac{|\omega(A)|}{\|A\|} = |\omega(R)|,$$

and the definition of $\|\omega\|$ can be rewritten as

$$\|\omega\| = \sup\{\frac{|\omega(A)|}{\|A\|} \text{ tel que } A \in \mathfrak{A}\}.$$

Now equation (72.65) gives $\|\omega\| \leqslant \omega(A)$. From definition of $\|\omega\|$, we also know that $\|\omega\| \geqslant \omega(\mathbb{1})$. Finally, $\|\omega\| = \omega(\mathbb{1})$.

Inverse sense.

We know that ω is bounded and that $\|\omega\| = \omega(\mathbb{1})$ and we want to prove that ω is positive. Let $A \in \mathfrak{A}_{\mathbb{R}}$ and let us decompose $\omega(A) = \alpha + i\beta$ where $\alpha, \beta \in \mathbb{R}$. We can adapt the reasoning that around equation (72.11) to prove that $\beta = 0$. Namely,

$$|\omega(B + it\mathbb{1})|^1 \leqslant \omega(\mathbb{1})^2\|B + it\mathbb{1}\| \leqslant \omega(\mathbb{1})^2(\|B\|^2 + t^2), \tag{72.66}$$

and for the self-adjoint element $B := \omega(\mathbb{1})A - \alpha\mathbb{1}$,

$$|\omega(B + it\mathbb{1})|^2 = \beta^2 + 2\beta t\omega(\mathbb{1}) + t^2\omega(\mathbb{1})^2.$$

Combining the two equations, and taking into account the fact that $\omega(\mathbb{1}) = \|\omega\| > 0$,

$$\beta^2 + \beta t\omega(\mathbb{1}) \leqslant \omega(\mathbb{1})^2\|B\|^2.$$

Since the right hand side don't depend on t, we conclude that $\beta = 0$. This shows that ω is real on $\mathfrak{A}_{\mathbb{R}}$. Now we are going to prove that $\omega(A) \geqslant 0$ when $A \geqslant 0$. Let $s > 0$ be so small that $\|\mathbb{1} - sA\| \leqslant 1$. Then

$$1 \geqslant \|\mathbb{1} - sA\| = \underbrace{\frac{\|\omega\|}{\omega(\mathbb{1})}}_{=1}\|\mathbb{1} - sA\| \geqslant \frac{|\omega(\mathbb{1} - sA)|}{\omega(\mathbb{1})}.$$

We conclude that $|\omega(\mathbb{1}) - s\omega(A)| \leqslant \omega(\mathbb{1})$, which is only possible if $\omega(A) \geqslant 0$. It proves that ω is positive.

Let us now check that points (1) and (2) are effectively obtained. We proved that for a state $\|\omega\| = \omega(\mathbb{1})$ and in the definition, we impose $\omega(\mathbb{1}) = 1$.

□

72.9.1 States on non unital C^*-algebras

We now relax the unital hypothesis.

DefApplPositive

Definition 72.62.
*A linear map $q\colon \mathfrak{A} \to \mathcal{B}$ between two C^*-algebra is **positive** when $q(A) \geqslant 0$ in $\mathcal{B}$ whenever $A \geqslant 0$ in $\mathfrak{A}$.*

Proposition 72.63.
A positive map is bounded. prop:posborn

Since for linear map, to be bounded is equivalent to continuous, all positive map is continuous.

Proof. We first prove that a bounded map on $\mathfrak{A}^+$ is bounded on $\mathfrak{A}$. Let us decompose $A \in \mathfrak{A}$ into $A = A' + iA''$ with $A', A'' \in \mathfrak{A}_{\mathbb{R}}$ and use lemma 72.29 to write $A = A'_+ - A'_- + iA''_+ - iA''_-$.We have $\|A'\| \leqslant \|A\|$ and $\|A''\| \leqslant \|A\|$. If B is one of $A'_\mp$ or $A''_\pm$, lemma also says that $\|B\| \leqslant \|A\|$. Now let us suppose that $\|q(B)\| \leqslant C\|N\|$ for all $B \in \mathfrak{A}^+$ and a certain $C > 0$. Then $\|q(A)\| \leqslant 4C\|A\|$ and q is then bounded.

We are now going to prove that a positive map q is bounded. Suppose that q is unbounded. In particular, it is not bounded in $\mathfrak{A}^+$ (if it were, it should be bounded everywhere) and there exists a sequence $(A_n) \in \mathfrak{A}_1^+$ such that $\|q(A_n)\| \geqslant n^3$ for all n. Here, the symbol $\mathfrak{A}_A^+$ denote the elements $A \in \mathfrak{A}^+$ for which $\|A\| \leqslant 1$.

Let us consider the series $\sum_{n=0}^{\infty} n^{-2}A_n$. Since $\|A_n\| \leqslant 1$, this converges to an element $A \in \mathfrak{A}^+$. If q is positive, then $q(A) \geqslant n^{-2}q(A_n)$ and from property $-B \leqslant A \leqslant B \Rightarrow \|A\| \leqslant \|B\|$, we see that

$$\|q(A)\| \geqslant n^{-2}\|q(A_n)\| \geqslant n$$

for all $n \in \mathbb{N}$. Indeed, q positive implies $q(A_n) \geqslant 0$ because $A_n \in \mathfrak{A}^+$. Now,

$$\begin{aligned} &- q(A) \leqslant n^{-2}q(A_n) \leqslant q(A) \\ &\Rightarrow \|n^{-2}q(A_n)\| \leqslant \|q(A)\| \\ &\Rightarrow \|q(A)\| \geqslant n^{-2}\|q(A_n)\| \geqslant n^3. \end{aligned} \tag{72.67}$$

This proves that $\|q(A)\|$ is greater than n^3 for all n; this is impossible (because this is a finite number). Then q is bounded on $\mathfrak{A}^+$ and on the whole $\mathfrak{A}$.

□

If we consider $\mathcal{B} = \mathbb{C}$, states are particular cases of positive maps and then states on unital C^*-algebra are bounded and have norm 1.

In order to define a state on a non-unital C^*-algebra, we can't take the normalization $\omega(\mathbb{1}) = 1$. But we just prove a condition to get an unit norm in the unital case. So we define

Proposition 72.64.
Let $\mathfrak{A}$ be an involutive Banach algebra with unit $\mathbb{1}$ such that $\|\mathbb{1}\| = 1$. If f is a positive linear form, then it is continuous and $\|f\| = f(\mathbb{1})$. prop_Dix214

Proof. This statement is nothing else than propositions 72.61 and 72.63. □

Proposition 72.65.
Let $\mathfrak{A}$ be an involutive Banach algebra with approximate unit and $\mathfrak{A}_{\mathbb{1}}$ the involutive algebra deduced from $\mathfrak{A}$ by adding an unit. We consider f, a linear positive and continuous form on $\mathfrak{A}$. Then

(1) $\forall A \in \mathfrak{A}$, $f(A^*) = \overline{f(A)}$ *and* itemi_prop_invaddunit

$$|f(A)|^2 \leqslant \|f\| f(A^*A),$$

(2) $|f(B^*AB)| \leqslant \|A\| f(B^*B)$, itemii_prop_invaddunit

(3) $\|f\| = \sup_{\substack{A \in \mathfrak{A} \\ \|A\| \leqslant 1}} f(A^*A)$, itemiii_prop_invaddunit

itemiv_prop_invaddunit

(4) *Let $(A_i)_{i\in I}$ be elements of $\mathfrak{A}$ indexed by the filtering set I such that $\|A_i\| \leqslant 1$ and $f(A_i) \to \|f\|$. Then $f(A_i^*A_i) \to \|f\|$,*

itemv_prop_invaddunit

(5) *If $(u_j)_{j\in J}$ is an approximate unit of $\mathfrak{A}$, then $f(u_j) \to \|f\|$ and $f(u_j^*u_j) \to \|f\|$,*

itemvi_prop_invaddunit

(6) *The function f can be extended to an unique positive form $\tilde{f}$ on $\mathfrak{A}_\mathbb{1}$ in such a way that $\tilde{f}(\mathbb{1}) = \|f\|$. If a positive form on $\mathfrak{A}_\mathbb{1}$ extends f, it is everywhere bigger than $\tilde{f}$.*

itemvii_prop_invaddunit

(7) *With the same $(A_i)_{i\in I}$ that in (4), we have $A_i \to \mathbb{1}$ for the prehilbert structure defined by $\tilde{f}$ on $\mathfrak{A}_\mathbb{1}$. Then $\mathfrak{A}$ is everywhere dense in $\mathfrak{A}_\mathbb{1}$ for this structure.*

prop_invaddunit

Proof. The definition $(A, B) := f(A^*B)$ gives a pre-inner product (see page 3762) because

$$(\lambda A, B) = f(\overline{\lambda}A^*B) = \overline{\lambda}(A, B).$$

So we want to write

$$f(A^*) = f(A^*\mathbb{1}) = (A, \mathbb{1}) = \overline{(\mathbb{1}, A)} = \overline{f(A)},$$

but we do not have unit. We therefore have to use continuity of f instead:

$$f(A^*) = \lim(A^*u_j) = \lim \overline{f(u_j^*A)} = \lim \overline{f\big((A^*u_j)^*\big)} = \overline{f(A^{**})} = \overline{f(A)}.$$

The second equality is not a particular case of $f(A^*) = \overline{f(A)}$, but the fact that $f(B^*A) = \overline{f(A^*B)}$ because f defines a pre-inner product [10].

Since we have a pre-inner product, we have Cauchy-Schwarz:

$$\lim |f(Au_j)|^2 \leqslant \lim f(A^*A) f(u_j^*u_j),$$

hence

$$|f(A)|^2 = \lim |f(Au_j)|^2 \leqslant f(A^*A) \lim f(u_j^*u_j). \tag{72.68}$$

An involutive Banach algebra fulfils $\|AB\| \leqslant \|A\|\,\|B\|$ for all A, $B \in \mathfrak{A}$, so $\|u_j^*u_j\| \leqslant 1$ because definition of an approximate unit gives $\|u_j\| \leqslant 1$. We have

$$\|f\| = \sup_{\|A\|=1} |f(A)|$$

Let us consider $\lambda \in \mathbb{R}$ such that $\|\lambda u_j^*u_j\| = 1$; we necessarily have $\lambda \geqslant 1$. We have

$$\|f\| \geqslant f(\lambda u_j^*u_j) = \lambda f(u_j^*u_j) \geqslant f(u_j^*u_j)$$

because f is positive and $u_j^*u_j \in \mathfrak{A}_\mathbb{R}$. We conclude that

$$f(A^*A) \lim f(u_j^*u_j) \leqslant \|f\| f(A^*A)$$

and this finally gives the first point:

$$|f(A)|^2 \leqslant \|f\| f(A^*A). \tag{72.69}$$

(7)⇒(7). Indeed, (1) implies $\|f\|^2 \leqslant \|f\| \lim f(A_i^*A_i)$ which in turn gives $\lim f(A_i^*A_i) \geqslant \|f\|$. But $\|A_i\| \leqslant 1$, so $\|f\| \geqslant \lim f(A_i^*A_i)$ and we conclude that $\|f\| = \lim f(A_i^*A_i)$.

(4)⇒(3). On the one hand, for all $\|A\| \leqslant 1$, we have $f(A^*A) \leqslant \|f\|$; on the other hand we have a sequence A_i such that $f(A_i^*A_i) \to \|f\|$. So $\|f\|$ is an upper bound for the set of $f(A^*A)$ and in the same time, it is in the adherence of this set. So $\|f\|$ is the supremum of this set.

10. C'est un problème pcq sans ça, je ne vois pas comment démontrer que ce f donne bien un pré produit scalaire.

Proof of (6). Unicity is clear because the prescription $\tilde{f}(\mathbb{1}) = \|f\|$ gives $\tilde{f}$ on a basis of $\mathfrak{A}_\mathbb{1}$. For existence, we pick $(\lambda, A) = \lambda + A \in \mathfrak{A}_\mathbb{1}$ with $\lambda \in \mathbb{C}$ and $A \in \mathfrak{A}$ and we pose $\tilde{f}(\lambda + A) = \lambda\|f\| + f(A)$. It is linear, extends f and $\tilde{f}(\mathbb{1}) = \|f\|$. To get positivity,

$$\begin{aligned}\tilde{f}\big((\lambda + A)^*(\lambda + A)\big) &= f(A^*A - \overline{\lambda}A + \lambda A^*) + |\lambda|^2\|f\| \\ &= f(A^*A) + 2\operatorname{Re}\overline{\lambda}f(A) + |\lambda|^2\|f\| \\ &\geqslant f(A^*A) - 2|\lambda|\|f\|^{\frac{1}{2}} f(A^*A)^{\frac{1}{2}} + |\lambda|^2\|f\| \\ &= \big[f(A^*A)^{\frac{1}{2}} - |\lambda|\,\|f\|^{\frac{1}{2}}\big]^{\frac{1}{2}} \\ &\geqslant 0.\end{aligned} \tag{72.70}$$

Let g be a positive form on $\mathfrak{A}_\mathbb{1}$ which extends f. The involutive Banach algebra of $\mathfrak{A}$ is extended to $\mathfrak{A}_\mathbb{1}$ by setting $\mathbb{1} = 1$. Proposition 72.64 and the fact that g extends f make $\|f\| \leqslant \|g\| = g(\mathbb{1})$. Hence

$$g\big((\lambda + A)^*(\lambda + A)\big) \geqslant f\big((\lambda + A)^*(\lambda + A)\big),$$

and finally $g \geqslant \tilde{f}$.

Now we take $B \in \mathfrak{A}$ and we look at the form $g(A) = \tilde{f}(B^*AB)$. It is a positive form because

$$g(A^*A) = \tilde{f}(B^*A^*AB) = \tilde{f}\big((AB)^*(AB)\big) \geqslant 0.$$

For the norm of g, we have $\|g\| = g(\mathbb{1}) = \tilde{f}(B^*B)$. Since $\|A^*A\|\mathbb{1} \geqslant A^*A$, we find $\|A^*A\|f(\mathbb{1}) \geqslant f(A^*A)$ and therefore

$$\|f\|\,\|A^*A\| \geqslant f(A^*A).$$

We find

$$|f(B^*AB)|^2 \leqslant f(B^*B)\|g\|\,\|A^*A\|,$$

but $\|g\| = f(B^*B)$ and $A^*A \leqslant \|A\|^2$, so

$$|f(B^*AB)| \leqslant f(B^*B)\|A\|. \tag{72.71}$$

Verification of (7). The pre-Hilbert structure defined by $\tilde{f}$ is the convergence notion $A_i \to A$ when $f(A_i) \to f(A)$. The first part of statement (7) is thus just the fact that

$$\tilde{f}\big((A_i - \mathbb{1})^*(A_i - \mathbb{1})\big) \to 0.$$

¡¡ Avertissement/question à la lectrice !! 72.66
Pas fait le reste de la preuve. C'est dans [959].

□

We consider an involutive Banach algebra $\mathfrak{A}$; f, a positive linear continuous form on $\mathfrak{A}$ and $\tilde{f}$ its extension to $\mathfrak{A}_\mathbb{1}$ by proposition 72.65. We claim that

$$\{A \in \mathfrak{A} \text{ tel que } \tilde{f}(A^*A) = 0\}$$

is a left ideal. Indeed let us show that BA belongs to this set when A does. We use (2) of proposition 72.65 to compute the norm on both side of the identity

$$\tilde{f}\big((BA)^*(BA)\big) = \tilde{f}(A^*B^*BA).$$

We find

$$|\tilde{f}\big(A^*(B^*B)A\big)| \leqslant \|B^*B\|f(A^*A) = 0.$$

Thus $\tilde{f}\big(A^*(B^*B)A\big) = 0$.

DefStateCSA

Definition 72.67.
*A **state** on a C*-algebra is a linear map $\omega\colon \mathfrak{A} \to \mathbb{C}$ which is positive and has norm 1.* def:etatnon

Proposition 72.63 shows that a state is bounded; then one can impose the condition $\|\omega\| = 1$. From proposition 72.61, a state in the sense of definition 72.67 on an unital C^*-algebra is a state in the sense of definition 72.58 because $1 = \|\omega\| = \omega(\mathbb{1})$.

Proposition 72.68 ([489]).
A state on a non unital C^*-algebra has an unique extension on the unitized algebra. prop_st_unit_ext

Proof. Let ω be a state on the C^*-algebra $\mathfrak{A}$ and let us define the extension

$$\omega_{\mathbb{1}}(A + \lambda\mathbb{1}) := \omega(A) + \lambda$$

on $\mathfrak{A}_{cun}$. It is clear that $\omega_{cun} = 1$. We have now to show that $\omega_{cun}(A) \geqslant 0$ whenever $A \in \mathfrak{A}^+$. On $\mathfrak{A}$, the form ω is a state and the is bounded; let $(\mathbb{1}_\lambda)$ be an approximate unit in $\mathfrak{A}$. Since $\lim_{\lambda\to\infty} \|A - \mathbb{1}_\lambda A\| = 0$, we have $|\omega(A - A\mathbb{1}_\lambda)| \to 0$. Let us use (72.62) and (72.63) in the particular case $A = B$:

$$|\overline{\omega(A)}\omega(A)|^2 \leqslant \omega(A^*A)^2,$$

but since ω is positive, the right hand side is a positive number and we can remove the square in both sides; using the fact that for any complex number, $|z\overline{z}| = |z|^2$, we find

$$|\omega(A)^2| \leqslant \omega(A^*A).$$

Then, with $\omega_{\mathbb{1}}$, we find

$$\omega_{\mathbb{1}}\big((A + \lambda\mathbb{1})^*(A + \lambda\mathbb{1})\big) \geqslant |\omega(A + \lambda\mathbb{1})|^2 = |\omega(A) + \lambda|^2,$$

but $(A + \lambda\mathbb{1})^*(A + \lambda\mathbb{1})$ is a general element of $\mathfrak{A}_{\mathbb{1}}$. It proves that ω is positive.

□

The following result gives a lot of examples of states.

Lemma 72.69.
For all $A \in \mathfrak{A}$ and $a \in \sigma(A)$, there exists a state ω_a on $\mathfrak{A}$ such that $\omega_a(A) = a$. When $A = A^*$, we can find a state ω such that $|\omega(A)| = \|A\|$. lem:omAenomA

Proof. If $\mathfrak{A}$ has no unit, we add one. We define $\tilde{\omega}\colon \mathbb{C}A + \mathbb{C}\mathbb{1} \to \mathbb{C}$ by $\tilde{\omega}(\lambda A + \mu\mathbb{1}) := \lambda a + \mu$. Since $a \in \sigma(A)$, we know that $\lambda a + \mu \in \sigma(\lambda A + \mu\mathbb{1})$ and then that $\lambda A + \mu\mathbb{1} - (\lambda a + \mu)\mathbb{1}$ has no inverse.

When $\mathfrak{A}$ has an unit, we know that $\omega \in \Delta(\mathfrak{A}) \Rightarrow |\omega(A)| \leqslant \|A\|$. We want to get the same for our $\tilde{\omega}$. Looking at the proof of (72.1), we see that we have to prove that $\tilde{\omega}(x) \neq 0$ whenever x is invertible. Be careful on a point: the question is posed in the C^*-algebra $\mathbb{C}A + \mathbb{C}\mathbb{1}$, not in $\mathfrak{A}$ or $\mathfrak{A}_{\mathbb{1}}$. So we take an invertible element $\lambda A + \mu\mathbb{1}$. Then $-\mu/\lambda \neq a$ because it is not in $\sigma(A)$. So $\tilde{\omega}_a(\lambda A + \mu\mathbb{1}) = \lambda a + \mu \neq 0$ and we can affirm that

$$|\tilde{\omega}_a(\lambda A + \mu\mathbb{1})| \leqslant \|(\lambda A + \mu\mathbb{1})\|. \tag{72.72}$$

Since $\tilde{\omega}_a(\mathbb{1}) = 1$, we know that $\|\tilde{\omega}_a\| \geqslant 1$, but the equation above shows that $\|\tilde{\omega}_a\| \leqslant 1$. We conclude that $\|\tilde{\omega}_a\| = 1$.

Hahn-Banach theorem 60.47 gives an extension ω_a of $\tilde{\omega}_a$ to the whole $\mathfrak{A}$ with norm 1. Point (2) of proposition 72.61 shows that ω_a is a state on $\mathfrak{A}_{\mathbb{1}}$ with $\omega_a(A) = \tilde{\omega}_a(A) = a$.

For the second assertion, we know that $\sigma(A)$ is a closed set, then there exists a $a \in \sigma(A)$ such that $r(A) = |A|$. For this a, $|\omega(A)| = |a| = r(a) = \|A\|$ because $\|A\| = r(A)$ when $A = A^*$.

Now if $\mathfrak{A}$ has no unit, we consider as ω_a, the restriction to $\mathfrak{A}$ of the ω_a that we build on $\mathfrak{A}_{\mathbb{1}}$.

□

72.10 Uniqueness of the norm

prop:unicitenormcsa

Proposition 72.70.
The norm on a C-algebra is unique in the sense that if $\mathfrak{A}$ is a C*-algebra, one cannot find an other norm on $\mathfrak{A}$ (as Banach algebra) for which $\mathfrak{A}$ is a C*-algebra.*

Proof. Let us consider an element such that $A = A^*$. Point (4) of theorem 60.62 applied to $f = \mathrm{Id}_{\sigma(A)}$, and the fact that $\|f\|_\infty = r$ imply that $\|A\| = r(A)$.

Considering this argument with a general A^*A instead of a particular A, we find that

$$\|A\| = \sqrt{r(A^*A)}. \tag{72.73}$$

Since the spectral radius is determined by the Banach algebra structure only, this formula shows that the norm on a general element A is unique.

□

Remarque 72.71.
The way to complete a $*$-algebra is not unique. Thus one can construct several C^*-algebra from a given $*$-algebra.

72.10.1 Convexity

A **convex** subspace C of a vector space V is a subset of V such that for all v, $w \in C$ and $\lambda \in [0,1]$, we have $\lambda v + (1-\lambda)w \in C$.

If $p_i > 0$, $\sum_i p_i = 1$ and $v_i \in C$, then $\sum_i v_i \in C$.

Lemma 72.72.
If $\mathfrak{A}$ is an unital C-algebra, then the space $\mathcal{S}(\mathfrak{A})$ of all states on $\mathfrak{A}$ is convex.*

Proof. If ω and η are states, then they are positive and $\omega(\mathbb{1}) = \eta(\mathbb{1}) = 1$. Then $\lambda\omega + (1-\lambda)\eta$ is also positive and $\lambda\omega(\mathbb{1}) + (1-\lambda)\eta(1) = 1$. □

Proposition 72.73.
In the non unital case, $\mathcal{S}$ is convex too.

Proof. Let ω and η be two states on a non unital C^*-algebra $\mathfrak{A}$ and ω_{cun}, η_{cun} their extensions to $\mathfrak{A}_\mathbb{1}$. From lemma, $\xi_{cun} = \lambda\omega_{cun} + (1-\lambda)\eta_{cun}$ is a state on $\mathfrak{A}_\mathbb{1}$. We want the restriction $\xi = \lambda\omega + (1-\lambda)\eta$ to be a state on $\mathfrak{A}$.

Since $\xi_{cun}(A + \mu\mathbb{1}) = \lambda\omega(A) + (1-\lambda)\eta(A) + \mu$, the restriction reads

$$\xi(A) = \lambda\omega(A) + (1-\lambda)\eta(A).$$

Then the unique extension of ξ with same norm is precisely ξ_{cun}. It proves that $\|\xi\| = 1$. □

The dual $\mathcal{B}^*$ of a Banach space $\mathcal{B}$ is the space of functional $\rho\colon \mathcal{B} \to \mathbb{C}$ and a functional is by definition linear and continuous. If $\mathfrak{A}$ is unital, proposition 72.61 ensures that $\mathcal{S}(\mathcal{A}) \subset \mathfrak{A}^*$. Let us recall the w^*-limit in $\mathfrak{A}^*$. We say that $\omega_n \to \omega$ when for all $v \in \mathfrak{A}$, $\omega_n(A) \to \omega(A)$. It is clear that the w^*-limit preserves the positivity: if $\omega_n \geqslant 0$ and $\omega_n \to \omega$, then $\omega \geqslant 0$. This proves that $\mathcal{S}(\mathfrak{A})$ is closed in $\mathfrak{A}^*$ for the w^*-topology. From normalization $\|\omega\| = 1$, we conclude that $\mathcal{S}(\mathfrak{A})$ is a closed set in the unit ball of $\mathfrak{A}^*$. From Banach-Alaoglu theorem, $\mathcal{S}(\mathfrak{A})$ is compact. What is proved is

Proposition 72.74.
The space of states of an unital C-algebra is convex and compact.*

The simplest example is $\mathfrak{A} = \mathbb{C}$. From $\omega(1) = 1$ and linearity, we deduce $\omega(i) = i\omega(1) = i$. Then $\omega(a + bi) = a + bi$ is the only state on $\mathbb{C}$. An isolated point is compact and convex. Let us now consider a less trivial example.

We now consider $\mathfrak{A} = \mathbb{C}^2$. We see $\mathbb{C}^2$ as $\mathbb{C} \oplus \mathbb{C}$ and we write $z \dot{+} z'$ the element $(z, z') \in \mathbb{C}^2$. We define the product by

$$(\lambda_1 \dot{+} \mu_1)(\lambda_2 \dot{+} \mu_2) = \lambda_1\lambda_2 \dot{+} \mu_1\mu_2.$$

From characterization $\mathfrak{A}^+ = \{B^*B \text{ tel que } B \in \mathfrak{A}\}$, a generic element in $\mathfrak{A}^+$ reads

$$(\bar{\lambda} \dot{+} \bar{\mu})(\lambda \dot{+} \mu) = \bar{\lambda}\lambda \dot{+} \bar{\mu}\mu.$$

So positive elements in $\mathbb{C}^2$ are (λ, μ) with $\lambda, \mu \geqslant 0$. In order for ω to be positive, we need $\omega(\lambda, \mu) = c\lambda + d\nu \geqslant 0$ for all $\lambda, \mu \geqslant 0$. For normalization $\omega(1, 1) = 1$, we also need $c + d = 1$. Then $\mathcal{S}(\mathbb{C}^2)$ can be identified with $[0, 1]$ which is convex and compact.

72.11 Representation

72.11.1 Representation of involutive algebra

Let $\mathfrak{A}$ be an involutive algebra and H a Hilbert space. A **representation** of $\mathfrak{A}$ in H is a map $\pi : \mathfrak{A} \to \mathcal{L}(H)$ such that

$$\begin{aligned} \pi(A + B) &= \pi(A) + \pi(B), \quad \pi(\lambda A) = \lambda\pi(A), \\ \pi(AB) &= \pi(A) \circ \pi(B), \quad \pi(A^*) = \pi(A)^*. \end{aligned} \tag{72.74}$$

A linear form $f : \mathfrak{A} \to \mathbb{R}$ on the involutive algebra $\mathfrak{A}$ is **positive** if $f(A) \geqslant 0$ whenever $A > 0$.

Lemma 72.75.
If $\rho : \mathbb{M}_n(\mathbb{C}) \to \mathrm{End}(V)$ is a representation of the matrix algebra $\mathbb{M}_n(\mathbb{C})$ on the finite dimensional space V, then there exists an isomorphism $V \to \mathbb{C}^n \oplus \ldots \oplus \mathbb{C}^n$ which intertwines ρ to a multiple of the standard representation of the matrices on $\mathbb{C}^n$.

Proposition 72.76.
Let $\mathfrak{A}$ be an involutive algebra. We have

itemi_prop_invalgrepr

(1) If π is a representation of $\mathfrak{A}$ in H and if $\xi \in H$, then $A \mapsto \langle \pi(A)\xi, \xi \rangle$ is a positive form on $\mathfrak{A}$.

itemii_prop_invalgrepr

(2) Let π and π' be two representations of $\mathfrak{A}$ in H and H' respectively, and ξ, ξ', two corresponding totalizing vectors. If $\langle \pi(A)\xi, \xi \rangle = \langle \pi'(A)\xi', \xi' \rangle$ for all $A \in \mathfrak{A}$, then there exists an unique isometry $H \to H'$ which transforms π into π' and ξ into ξ', i.e.

$$\begin{aligned} U\pi(A)U^{-1} &= \pi'(A) \\ U\xi &= \xi' \end{aligned}$$

prop_invalgrepr

Proof. For the first point, it is easy:

$$\langle \pi(A^*A)\xi, \xi \rangle = \langle \pi(A^*)\pi(A)\xi, \xi \rangle = \langle \pi(A)\xi, \pi(A)\xi \rangle = \|\pi(A)\xi\|^2 \geqslant 0.$$

We used property $\pi(A^*) = \pi(A)^*$.

For the second point, $\overline{\pi(\mathfrak{A})\xi} = H$ because ξ is totalizing. We have

$$\langle \pi(A)\xi, \pi(B)\xi \rangle = \langle \pi(B^*A)\xi, \xi \rangle = \langle \pi'(B^*A)\xi', \xi' \rangle = \langle \pi'(A)\xi, \pi'(B)\xi' \rangle, \tag{72.75}$$

eq_piAxipreis

but vectors of the form $\pi(A)\xi$ are everywhere dense in H (and $\pi'(A)\xi'$ in H'), so we have an isomorphism

$$\begin{aligned} U &: H \to H' \\ \pi(A)\xi &\mapsto \pi'(A)\xi' \end{aligned} \tag{72.76}$$

Equation (72.75) shows that U is an isometry; it is linear because of linearity of π. The map U is surjective because ξ' is totalizing and injective because if $\pi'(A)\xi' = \pi'(B)\xi'$, $\pi'(A - B) = 0$ which proves that $A = B$ because π' is linear.

Now we check that this U is the searched map. For all A, $B \in \mathfrak{A}$,

$$\begin{aligned}\big[U \circ \pi(A)\big](\pi(B)\xi) &= U\pi(AB)\xi = \pi'(AB)\xi' \\ &= \pi'(A)\pi'(B)\xi' = \big[\pi'(A) \circ U\big](\pi(B)\xi),\end{aligned} \tag{72.77}$$

so U transforms π into π'. In order to prove that $U\xi = \xi'$, we will use the fact that U is an isometry:

$$\langle \xi', \pi'(A)\xi' \rangle = \langle \xi, \pi(A)\xi \rangle = \langle U\xi, U\pi(A)\xi \rangle = \langle U\xi, \pi'(A)\xi' \rangle.$$

For unicity, notice that equation $U\big(\pi(A)\xi\big) = \pi'(A)\xi'$ defines U on a dense subspace of H. Continuity finishes to fix U.

□

With the same notations, the map $A \to \langle \pi(A), A \rangle$ is the form **associated** with representation π and the vector ξ. Let $\mathcal{B}$ be an involutive subalgebra of $\mathcal{L}(H)$ and ξ, any element in H. We denote by ω_ξ the form on $\mathcal{B}$ defined by

$$\omega_\xi(A) = \langle A\xi, \xi \rangle \tag{72.78}$$

for all $A \in \mathcal{B}$. A positive form η on $\mathcal{B}$ is a **vector** form if there exists a $\xi \in H$ such that $\eta = \omega_\xi$.

Proposition 72.77.
Let $\mathfrak{A}$ be an involutive Banach algebra with approximate unit (u_i),π a nondegenerate representation of $\mathfrak{A}$ in H, $\xi \in H$ and f the positive form defined from π and ξ. We have

$$\|f\| = \langle \xi, \xi \rangle. \tag{72.79}$$

Proof. Point (5) of proposition 72.65 states that if $(u_i)_{i \in J}$ is an approximate unit in $\mathfrak{A}$, then

$$f(u_j) \to \|f\| \quad \text{and} \quad f(u_j^* u_j) \to \|f\|.$$

Hence $\|f\| = \lim f(u_j) = \lim \langle \pi(u_j)\xi, \xi \rangle$.

□

Proposition 72.78.
Let $\mathfrak{A}_\mathbb{1}$ be the involutive algebra deduced from $\mathfrak{A}$ by adding an unit and π, a representation of $\mathfrak{A}$ in H. There exists one and only one way to extend π to a representation $\pi_{(\mathbb{1})}$ of $\mathfrak{A}_\mathbb{1}$ in such a way that $\pi_{(\mathbb{1})}(\mathbb{1}) = \mathrm{Id}$.

The representation $\pi_{(\mathbb{1})}$ is the **canonical extension** of π. When π is a representation of $\mathfrak{A}$ in H, the set

$$K = \{\pi(A)\xi \text{ tel que } A \in \mathfrak{A}, \xi \in H\} \tag{72.80}$$

is a closed vector subspace of H. We say that K is the **essential subspace** of π. The representation is **nondegenerate** if $K = H$.

Proposition 72.79.
Let $\mathfrak{A}$ be an involutive Banach algebra and π, a nondegenerate representation of $\mathfrak{A}$ on H. If (u_i) is an approximate unit in $\mathfrak{A}$, then $\pi(u_i)$ strongly converges to Id *(see subsection 60.2.2).*

Proof. We have to prove that $\|\pi(u_i)\xi - \xi\| \to 0$ for any $\xi \in H$. From non degeneracy, the set $\{\pi(B)\xi\,, B \in \mathfrak{A}\}$ is total in H, so it is sufficient to prove the convergence for ξ of the form $\pi(B)\xi$. We have

$$\|\pi(u_i B) - \pi(B)\| \leqslant \|u_i B - B\| \to 0$$

because u_i is an approximate unit and relation (72.41).

The fact that π is nondegenerate makes

$$\{\pi(A)\xi \text{ tel que } A \in \mathfrak{A},\ \xi \in H\} = H.$$

If $\mathcal{U}$ is an open set in $\mathcal{L}(H)$ around Id, we have to prove that $\pi(u_i) \in \mathcal{U}$ for all $i \geqslant i_0$. Open sets are taken in the sense of seminorms $s_\xi = \|A\xi\|$. The balls —which are not open— are of the form

$$\begin{aligned} B(A;(\xi_j),(r_j)) &= \{X \in \mathcal{L}(H) \text{ tel que } s_{\xi_i}(X-A) < r_j \,\forall j\} \\ &= \{X \in \mathcal{L}(H) \text{ tel que } \|X\xi_j - A\xi_j\| < r_j \,\forall j\}. \end{aligned}$$

If a sequence fall into the balls $B(A;(\xi_j),(r_j))$ for a fixed A and if we consider an open set around this A, the latter open set will contain at least one of the $B(A;(\xi_j),(r_j))$, hence the sequence will fall into this open set too. So we have to prove that for all *finite* sequence (ξ_j) and r_j ($\xi_j \in H$ and $r_j \in \mathbb{R}^+$),

$$\pi(u_i) \in B(\mathrm{Id};(\xi_j),(r_j))$$

when i is large enough.

Let $B \in \mathfrak{A}$ and $\xi \in H$, we have already proved that $\|\pi(u_i B) - \pi(B)\| \leqslant \|u_i B - B\| \to 0$. We have to prove that $\|\pi(u_i)\xi_j - \xi_j\| < r_j$. Since $\pi(C)\zeta$ is a total set, by redefinition of ξ_j, we can write ξ_j under the form $\pi(B_j)\xi_j$. So

$$\|\pi(u_i)\pi(B_j)\xi_j - \pi(B_j)\xi_j\| = \|\pi(u_i B_j)\xi_j - \pi(B_j)\xi_j\| = \|[\pi(u_i B_j) - \pi(B_j)]\xi_j\|,$$

but we know that when A is a bounded operator on a Hilbert space, $\|Av\| \leqslant \|A\|\,\|v\|$. Thus we have

$$\|\pi(u_i)\pi(B_j)\xi_j - \pi(B_i)\xi_j\| = \|[\pi(u_i B_j) - \pi(B_i)]\xi_j\| = \|\pi(u_i B_j) - \pi(B_i)\|\,\|\xi_j\|.$$

Since the sequence (ξ_j) is finite, one can bound $\|\xi_j\|$ by a certain M, hence

$$\|\pi(u_i)\pi(B_j)\xi_j - \pi(B_i)\xi_j\| \leqslant M\|\pi(u_i B_j) - \pi(B_i)\|.$$

When i is large, the latter is as small as we want and in particular it can become smaller than all the r_j of the sequence. This proves that $\pi(u_i)$ strongly converges to Id in $\mathcal{L}(H)$. □

PropGNSanother

Proposition 72.80 (Another version of GNS construction)**.**
Let $\mathfrak{A}$ be an involutive Banach algebra with an approximate unit and the following elements:

- *$\mathfrak{A}_\mathbb{1}$ the involutive algebra obtained by adding an unit to $\mathfrak{A}$,*
- *f a positive continuous form on $\mathfrak{A}$,*
- *$\tilde{f}$ its canonical extension to $\mathfrak{A}_\mathbb{1}$,*
- *N the left ideal of $\mathfrak{A}_\mathbb{1}$ defined by $N = \{A \in \mathfrak{A}_\mathbb{1} \text{ tel que } \tilde{f}(A^*A) = 0\}$,*
- *$\mathfrak{A}'_f$ the pre-Hilbert space $\mathfrak{A}_\mathbb{1}/N$,*
- *$\mathfrak{A}_f$ the Hilbert space obtained by completion of the previous one.*

For each $A \in \mathfrak{A}_\mathbb{1}$, let

- *$\pi'(A)$, the operator in $\mathfrak{A}_\mathbb{1}/N$ obtained from quotient of the left multiplication by A in $\mathfrak{A}_\mathbb{1}$,*
- *ξ, the canonical image of $\mathbb{1}$ in $\mathfrak{A}'_f$.*

In this setting we have

itemi_prop_DixGNS
(1) Each $\pi'(A)$ extends in one and only one way to a linear continuous representation of $\mathfrak{A}$ on $\mathfrak{A}_f$,

itemii_prop_DixGNS
(2) the map $A \to \pi(A)$ with $A \in \mathfrak{A}$ is a representation of $\mathfrak{A}$ in $\mathfrak{A}_f$,

itemiii_prop_DixGNS
(3) the vector ξ is totalizing for $\pi(\mathfrak{A})$,

itemiv_prop_DixGNS
(4) $\forall A \in \mathfrak{A}$, we have $f(A) = \langle \pi(A)\xi, \xi\rangle$.

prop_DixGNS

In this context, the vector ξ, being the representative of the identity, is sometimes called the **vacuum** of the GNS representation.

Proof. Let $\eta \in \mathfrak{A}_{\mathbb{1}}/N$ and let's say $\eta = [B]$ for $B \in \mathfrak{A}_{\mathbb{1}}$, $\pi'(A)\eta = [AB]$ and $\xi = [\mathbb{1}] = \{\mathbb{1} + n \text{ tel que } n \in N\}$. If $A \in N$, we have $\tilde{f}(A^*A) = 0$, thus for all $B \in N$,

$$\left|\tilde{f}\big((BA)^*(BA)\big)\right| = \left|\tilde{f}\big(A^*(B^*B)A\big)\right| \leqslant \|B^*B\| f(A^*A) = 0.$$

This proves that N is an ideal. Now $\pi'(B)\xi = \pi'(B)[\mathbb{1}] = [B]$, so

$$\langle \pi'(A)\pi'(B)\xi, \pi'(A)\pi'(B)\xi\rangle = \langle [AB], [AB]\rangle = \langle [B^*A^*AB], [\mathbb{1}]\rangle$$

where this product is defined by

$$\langle [A]\,,\, [B]\rangle := \tilde{f}(AB). \tag{72.81}$$

So

$$\begin{aligned}\langle \pi'(A)\pi'(B)\xi, \pi'(A)\pi'(B)\xi\rangle &= \tilde{f}(B^*A^*AB)\\ &\leqslant \|A^*A\|\tilde{f}(B^*B)\\ &= \|A^*A\|\langle \pi'(Bj\xi), \pi'(B)\xi\rangle.\end{aligned}$$

It gives $\|\pi'(A)[B]\| \leqslant \|A^*A\|\,\|[B]\|$, and finally

$$\|\pi'(A)\| \leqslant \|A^*A\|, \tag{72.82}$$

which proves that π' is continuous, and therefore bounded. But a bounded operator on a part of a Hilbert space may be extended to the whole space. This finish the proof of (1).

Now we prove that π is a representation for the structure of involutive algebra. The algebra structure is clear. For involution,

$$\langle \pi(A)\pi(B)\xi, \pi(C)\xi\rangle = \tilde{f}(C^*AB) = \tilde{f}\big((A^*C)^*B\big) = \langle \pi(B)\xi, \pi(A^*)\pi(C)\xi\rangle,$$

so $\pi(A)^* = \pi(A^*)$. Notice that there are no argument as "$\pi(B)\xi$ is dense"; we just use the fact that $\pi(B)\xi = [B]$ is the most general element in $\mathfrak{A}_{\mathbb{1}}/N$. Ok for point (2).

On the one hand, $\mathfrak{A}_{\mathbb{1}}$ is everywhere dense in $\mathfrak{A}$. On the other hand, $\mathfrak{A}'_f$ is everywhere dense in $\mathfrak{A}_f$ from the definition of a completion. But $\pi(\mathfrak{A})\xi$ is the image of $\mathfrak{A}$ in $\mathfrak{A}'_f$, so $\pi(\mathfrak{A})\xi = \mathfrak{A}'_f$ is everywhere dense in $\mathfrak{A}_f$. This proves that ξ is totalizing for $\pi(\mathfrak{A})$. This proves (3).

Finally, for all $A \in \mathfrak{A}_{\mathbb{1}}$,

$$\langle \pi(A)\xi, \xi\rangle = \tilde{f}(\mathbb{1}^*A\mathbb{1}) = \tilde{f}(A).$$

□

We say that the representation π and the vector ξ are defined from f. So we often write π_f and ξ_f.

72.11.2 Cyclic representations of C^*-algebra

Definition 72.81.
*A **representation** of the C^*-algebra $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$ is a linear map $\pi\colon \mathfrak{A} \to \mathfrak{B}(\mathscr{H})$ such that*

(1) $\pi(AB) = \pi(A) \circ \pi(B)$,
(2) $\pi(A^*) = \pi(A)^*$

for all A, $B \in \mathfrak{A}$.

Most of time, we will denote a representation by the pair $(\pi, \mathscr{H})$. When the represented C^*-algebra is ambiguous, we write $(\mathfrak{A}, \pi, \mathscr{H})$.

Lemrepresnormpresou

Lemma 72.82.
A representation π is continuous and fulfills

$$\|\pi(A)\| \leqslant \|A\|, \tag{72.83}$$

moreover when the representation is faithful, we have $\|\pi(A)\| = \|A\|$.

Proof. The spaces $\mathfrak{A}$ and $\mathfrak{B}(\mathscr{H})$ are C^*-algebra and π is a morphism proposition 72.41 concludes. The second claims follows from lemma 72.44. □

Let $(\pi_1, \mathscr{H}_1)$ and $(\pi_2, \mathscr{H}_2)$ be two representations. They are **equivalent** when there exists an unitary isomorphism $U : \mathscr{H}_1 \to \mathscr{H}_2$ such that

$$U\pi_1(A)U^* = \pi_2(A) \tag{72.84}$$

for all $A \in \mathfrak{A}$.

A representation $(\mathfrak{A}, \pi, \mathscr{H})$ is **nondegenerate** if 0 is the only vector to be cancelled by all $\pi(A)$. It is **cyclic** if there exists a **cyclic vector** $\Omega \in \mathscr{H}$, i.e. the closure of $\pi(\mathfrak{A})\Omega = \mathscr{H}$.

Lemma 72.83.
If π is an irreducible representation on $\mathscr{H}$, then any non zero vector is cyclic.

Proof. Let $v \neq 0$; if it were not cyclic, then $\overline{\pi(\mathfrak{A})v}$ should be a proper invariant subspace of $\mathscr{H}$. □

72.11.3 Primitive spectrum

For this short note about primitive spectrum, we follow [819].

When $\mathfrak{A}$ is any (not specially commutative) BC^*-algebra, the **primitive spectrum** of $\mathfrak{A}$ is the set $\operatorname{Prim}\mathfrak{A}$ of kernels of $*$ irreducible representations. An element in $\operatorname{Prim}\mathfrak{A}$ is a two-sided ideal.

There exists a suitable topology on this space, the **hull-kernel topology** or **Jacobson topology**. This is given by means of closure. If $W \subset \operatorname{Prim}\mathfrak{A}$, then we define the closure of W by

$$\overline{W} = \{\mathcal{I} \in \operatorname{Prim}\mathfrak{A} \text{ tel que } \cap W \subseteq \mathcal{I}\}. \tag{72.85}$$

where $\cap W = \bigcap_{\mathcal{J}\in W} \mathcal{J}$. The inclusion $\cap W \subseteq \mathcal{I}$ is an inclusion of subsets of $\mathfrak{A}$.

Proposition 72.84.
This definition defines a topology. Namely, it fulfils the Kuratowsky axioms:

(1) $\overline{\varnothing} = \varnothing$, enu802i

(2) $W \subseteq \overline{W}$, enu802ii

(3) $\overline{\overline{W}} = \overline{W}$, enu802iii

(4) $\overline{W_1 \cup W_2} = \overline{W_1} \cup \overline{W_2}$. enu802iv

Proof. Points (1) and (2) are trivial. For (3), remark that $\cap W = \cap\overline{W}$ (equality as subsets of $\mathfrak{A}$). Indeed, consider $A \in \cap W$. Then any $\mathcal{J} \in \overline{W}$ contains A and then $\cap\overline{W}$ contains A. Now if $A \in \cap\overline{W}$, then any $\mathcal{I}$ such that $\cap W \subseteq \mathcal{I}$ contains A and then $A \in \cap W$.

The proof of (4) is more complicated. If $V \subset W$, then $(\cap W) \subseteq (\cap V)$ and then $\overline{V} \subseteq \overline{W}$. Then $\overline{W_i} \subseteq \overline{W_1 \cup W_2}$ for each of $i = 1, 2$.

The inverse inclusion is as follows. Let $\mathcal{I}$ be an ideal kernel of the irreducible representation π of $\mathfrak{A}$ on the Hilbert space $\mathscr{H}$. Suppose that $\mathcal{I} \notin \overline{W_1} \cup \overline{W_2}$. Then we will prove that $\mathcal{I} \notin \overline{W_1 \cup W_2}$. There exists $A \in W_1$ and $B \in W_2$ such that $\pi(A) \neq 0$ and $\pi(B) \neq 0$. Let $\xi \in \mathscr{H}$ such that $\pi(A)\xi \neq 0$. Since π is irreducible, then $\pi(A)\xi$ is cyclic. Since $\pi(B) \neq 0$, then there exists a $\psi \in \mathscr{H}$ such that $\pi(B)\psi \neq 0$. Cyclicity of $\pi(A)\xi$ shows that there exists a C such that $\pi(C)\pi(A)\xi$ is sufficiently close to ψ to satisfy

$$\pi(B)\big(\pi(C)\pi(A)\xi\big) \neq 0.$$

Then $BCA \notin \ker\pi = \mathcal{I}$. Since W_i are ideals,

$$BCA \in (\cap W_1) \cap (\cap W_2) = \cap(W_1 \cup W_2).$$

Then $BCA \in W_1$ and $\cap(W_1 \cup W_2) \nsubseteq \mathcal{I}$. Consequently, $\mathcal{I} \notin \overline{W_1 \cap W_2}$.

□

72.11.4 GNS construction

Lemma 72.85.
Let $\mathfrak{M}$ be a $*$-algebra in $\mathfrak{B}(\mathscr{H})$, $\psi \in \mathscr{H}$ and p, the projection into the closure of $\mathfrak{M}\psi$. Then $p \in \mathfrak{M}'$, i.e. $[p, A] = 0$ for all $A \in \mathfrak{M}$.
lem_preGNS

Proof. Let $A \in \mathfrak{M}$; by definition of p, we have $Ap\mathscr{H} \subseteq p\mathscr{H}$. For a $A \in \mathfrak{M}$, we have

$$Ap\mathscr{H} = \{AB\psi \text{ tel que } B \in \mathfrak{M}\}, \tag{72.86}$$

but $\mathfrak{M}$ is an algebra, then $AB \in \mathfrak{M}$ and $Ap\mathscr{H} \subseteq \mathfrak{M}\psi \subseteq p\mathscr{H}$. If we define $p^\perp = \mathbb{1} - p$, we find $p^\perp Ap = 0$.

Indeed $(\mathbb{1} - p)Ap = Ap - pAp$ and $Apx - pApx$ can be computed by setting $x = B\psi$ for a certain B in the closure of $\mathfrak{M}$ (the part of x "outside" $\mathfrak{M}$ has no importance). Then

$$\begin{aligned} ApB\psi - pAB\psi &= AB\psi - pAB\psi \quad \text{because } B\psi \in \mathfrak{M} \\ &= AB\psi - AB\psi \quad \text{because } AB\psi \in \mathfrak{M}\psi \\ &= 0. \end{aligned} \tag{72.87}$$

It shows that $ApB\psi - pAB\psi = 0$ and then that $p^\perp Ap = 0$.

From this, we see that $Ap = pAp$. Let us now consider $A = A^*$ (for a general element in $\mathfrak{A}$, use the decomposition). We have $(Ap)^* = p^*A^* = pA$, but $(Ap)^* = (pAp)^* = pAp = Ap$. Then $[A, p] = 0$.

□

Proposition 72.86.
Any nondegenerate representation is direct sum of cyclic representations.

Proof. We apply the lemma with $\mathfrak{M} = \pi(\mathfrak{A})$. The non degeneracy of π makes p non zero: $\mathfrak{M}\psi$ is never zero. Now we consider the map $\rho\colon \mathfrak{A} \to \mathfrak{B}(\mathscr{H})$, $\rho(A) = p\pi(A)$. This is a representations because

$$\rho(AB) = p\pi(A)\pi(B) = p\pi(A)p\pi(B) = \rho(A)\rho(B), \tag{72.88}$$

and

$$\begin{aligned} \rho(A^*) &= p\pi(A^*) = \pi(A^*)p \\ &= \pi(A)^*p = (p\pi(A))^* = \rho(A)^*. \end{aligned} \tag{72.89}$$

More precisely, ρ is a representation on $p\mathscr{H}$ and when $\pi(A) \in p\mathscr{H}$, we have $\rho(A) = \pi(A)$. The representation ρ is constructed in such a way that ψ is a cyclic vector:

$$\rho(A)\psi = p\mathscr{H}.$$

The same construction with $\psi_2 \in p^\perp\mathscr{H}$ gives and going on gives the thesis. □

Let $(\mathfrak{A}, \pi, \mathscr{H})$ be a nondegenerate representation and $\Psi \in \mathscr{H}$ a vector with norm 1. The state ψ given by formula

$$\psi(A) = \langle \Psi, \pi(A)\Psi \rangle$$

is the **vector state** of Ψ relative to π.

ThoGNScontruction

Theorem 72.87 (GNS construction)**.**
Let $\mathfrak{A}$ be an unital C-algebra and $\omega\colon \mathfrak{A} \to \mathbb{C}$, a state*
GNSi

(1) There exists a Hilbert space $\mathscr{H}$, a representation $\pi_\omega\colon \mathfrak{A} \to \mathfrak{B}(\mathscr{H}_\omega)$ and a cyclic unit vector Ω_ω such that

$$\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle \tag{72.90}$$

for all $A \in \mathfrak{A}$.

GNSii

(2) *The triple* $(\mathscr{H}_\omega, \pi_\omega, \Omega_\omega)$ *is unique up to isomorphism in the following sense. Let* $\mathscr{H}$ *be a Hilbert space,* $\pi\colon \mathfrak{A} \to \mathfrak{B}(\mathfrak{A})$ *a representation and* $\Omega \in \mathscr{H}$ *an unit cyclic vector such that* $\omega(A) = \langle \Omega, \pi(A)\Omega\rangle$ *for all* $A \in \mathfrak{A}$*; then there exists an unitary isomorphism* $u\colon \mathscr{H}_\omega \to \mathscr{H}$ *such that* $u(\Omega_\omega) = \omega$ *and*

$$\pi(A) = u\pi_\omega(A)u^*$$

for all $A \in \mathfrak{A}$.

tho:GNS

Proof. From (72.62) we know that, if $A \in \mathcal{N}_\omega$ and $B \in \mathfrak{A}$, then $\omega(B^*A) = 0$. So we can define $\mathcal{N}$ in the two equivalent ways:

$$\begin{aligned}\mathcal{N}_\omega &= \{A \in \mathfrak{A} \text{ tel que } \omega(A^*A) = 0\}\\ &= \{A \in \mathfrak{A} \text{ tel que } \omega(B^*A) = 0 \text{ for all } B \in \mathfrak{A}\}.\end{aligned} \tag{72.91}$$

From the second line, we see that $\mathcal{N}_\omega$ is an ideal in $\mathfrak{A}$. The set $\mathcal{N}_\omega$ is closed from continuity of ω. We use the product defined by equation (72.61): $(A, B) := \omega(A^*B)$. It defines a sesquilinear form on the quotient $\mathfrak{A}/\mathcal{N}_\omega$

$$\langle VA, VB\rangle = \omega(A^*B) \tag{72.92}$$

where $V\colon \mathfrak{A} \to \mathfrak{A}/\mathcal{N}_\omega$ is the canonical projection $VA = A + \mathcal{N}_\omega$. So $\mathfrak{A}/\mathcal{N}_\omega$ is a pre-Hilbert space from which we build $\mathscr{H}_\omega$ by completion. We define the cyclic vector $\Omega_\omega = V\mathbb{1} \in \mathscr{H}_\omega$.

For each $A \in \mathfrak{A}$, we define $L_A\colon \mathfrak{A}/\mathcal{N}_\omega \to \mathfrak{A}/\mathcal{N}_\omega$ by

$$L_AVB = V(AB). \tag{72.93}$$

eq:defpiomega

We have

$$\begin{aligned}\|VAB\|^2 &= \langle VAB, VAB\rangle^2\\ &= \omega(B^*A^*AB)^2\\ &\leqslant \|A\|^4\omega(B^*B)^2 \quad \text{corollary 72.33}\\ &= \|A\|^4\|VB\|^2,\end{aligned} \tag{72.94}$$

then for all $\psi \in \mathfrak{A}/\mathcal{N}_\omega$, $\|L_A\psi\| \leqslant \|A\|^2\|\psi\|$. We conclude that

$$\|L_A\| \leqslant \|A\|^2. \tag{72.95}$$

Then L_A can be extended to a bounded operator $\pi_\omega(A)$ on the whole $\mathscr{H}_\omega$. The map $\pi_\omega\colon \mathfrak{A} \to \mathfrak{B}(\mathscr{H}_\omega)$ is a representation such that for all $A \in \mathfrak{A}$,

$$\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega\rangle, \tag{72.96a}$$

$$\overline{\pi_\omega(A)\Omega_\omega} = \overline{\mathfrak{A}/\mathcal{N}_\omega} = \mathscr{H}_\omega. \tag{72.96b}$$

It proves point (1); we now turn our attention to (2). For all A, $B \in \mathfrak{A}$, we have

$$\langle \pi_\omega(B)\Omega_\omega, \pi_\omega(A)\Omega_\omega\rangle = \langle \Omega_\omega, \pi_\omega(B^*A)\Omega_\omega\rangle = \langle V\mathbb{1}, V(B^*A)\rangle = \omega(B^*A), \tag{72.97}$$

eq:r1903r4

but from assumptions, $\omega(A) = \langle \Omega, \pi(A)\Omega\rangle$; then

$$\omega(B^*A) = \langle \Omega, \pi(B^*A)\Omega\rangle = \langle \pi(B)\Omega, \pi(A)\Omega\rangle. \tag{72.98}$$

eq:BesAOmBeA

We conclude that

$$\langle \pi_\omega(B)\Omega_\omega, \pi_\omega(A)\Omega_\omega\rangle = \langle \pi(B)\Omega, \pi(A)\Omega\rangle. \tag{72.99}$$

From definition of a cyclic vector, $\pi_\omega(\mathfrak{A})\Omega_\omega$ is dense in $\mathscr{H}_\omega$ and $\pi(\mathfrak{A})$ in $\mathscr{H}$. This allows us to define $u\colon \mathscr{H}_\omega \to \mathscr{H}$ by the condition

$$u\pi_\omega(A)\Omega_\omega = \pi(A)\Omega.$$

It is a well defined Hilbert space isomorphism because if $\pi_\omega(A)\Omega_\omega = \pi_\omega(B)\Omega_\omega$, then equation (true for all $D \in \mathfrak{A}$) $\langle \pi_\omega(B)\Omega_\omega, \pi_\omega(D)\Omega_\omega\rangle = \langle \pi(B)\Omega, \pi(D)\Omega\rangle$ gives an equation of the form $\langle x, d\rangle = \langle y, d\rangle$ for all d in a dense subset. This equation implies that $x = y$, or $\pi(A)\Omega = \pi(B)\Omega$. We know from general Hilbert space theory that a surjective isometry is unitary; this is the case of u.

□

For later use, we mention that equation (72.98) gives

$$\|\pi_\omega(A)\Omega_\omega\| = \omega(A^*A) \tag{72.100}$$ eq:pionomaesm

when $A = B$.

Corollary 72.88.
Let $(\pi_i, \mathscr{H}_i)$ $(i = 1, 2)$ be two cyclic representations with cyclic vectors Ω_i. If for all $A \in \mathfrak{A}$,

$$\omega_1(A) := \langle \Omega_1, \pi_1(A)\Omega_1\rangle = \langle \Omega_2, \pi_2(A)\Omega_2\rangle =: \omega_2(A),$$

then they are equivalent representations.

Proof. The representation π_i in $\mathscr{H}_i$ is cyclic and induces a GNS representation π_{ω_i}. Since $\omega_1 = \omega_2$, these two GNS representations are the same and π_1 and π_2 are thus both equivalent to the same representation.

□

The following is just a restatement of (2) of theorem 72.87.

Proposition 72.89.
If $(\mathfrak{A}, \pi, \mathscr{H})$ is a cyclic representation, then all the GNS constructions build from a vector state are unitary equivalent to π. prop:cyclequivGNS

72.11.5 Universal representation

Theorem 72.90.
Any C-algebra accepts an isometric representation on a Hilbert space.*

Proof. Let $\mathfrak{A}_h$ be the real Banach space build from hermitian elements of $\mathfrak{A}$ and choose a non zero $A \in \mathfrak{A}$. The element $-A^*A$ does not belong to $\mathfrak{A}^+$. Since $\mathfrak{A}^+$ is a convex closed cone, there exists a linear continuous form f_A on $\mathfrak{A}_h$ such that $f_A(B) \geqslant 0$ for all $B \in \mathfrak{A}^+$ and $f_A(-A^*A) < 0$. One can identify f_A to an hermitian form on $\mathfrak{A}$ because B is decomposed as $B = A_1 + iA_2$ with $A_1, A_2 \in \mathfrak{A}_h$; we define $f_A(B) = f_A(B_1) + f_A(B_2)$. The function f_A is also a positive form on $\mathfrak{A}$ because on any positive element B^*B, we have $f_A(B^*B) > 0$.

Now we look at π_A, the representation defined by f_A. We have $f_A(B) = \langle \pi_A(B)\xi, \xi\rangle$, thus

$$f_A(A^*A) = \langle \pi_A(A^*)\pi_A(A)\xi, \xi\rangle,$$

which is zero if $\pi_A(A) = 0$. Then $\pi_A(A)\xi \neq 0$ and we conclude that $\pi_A(A) \neq 0$.

We consider π, the direct sum of all the representations π_A for all $A \in \mathfrak{A}$, $A \neq 0$. First we prove that π is injective. Indeed if $\pi(B) = 0$, we have $\pi_A(B) = 0$ for all A; in particular

$$0 = \langle \pi_A(B)\xi, \pi_A(B)\rangle = \langle \pi_A(B^*B)\xi, \xi\rangle = f_A(B^*B) \tag{72.101}$$

by (4) of proposition 72.80. So if π is not invertible, there exists a $B \neq 0$ such that for all A, $f_A(B^*B) = 0$; this implies $\|B^*B\| = 0$ and $B = 0$. Contradiction.

The map π is isometric as injective morphism: $\|\pi(A)\| = \|A\|$. □

The **universal representation** π_u of a C^*-algebra $\mathfrak{A}$ is the direct sum of all the GNS representations π_ω with $\omega \in \mathcal{S}(\mathfrak{A})$. The representation space is

$$\mathscr{H}_u = \bigoplus_{\omega \in \mathcal{S}(\mathfrak{A})} \mathscr{H}_\omega.$$

Theorem 72.91 (Gelfand-Neumark)**.**
A C-algebra is isomorphic to a subalgebra of $\mathfrak{B}(\mathscr{H})$ for a certain Hilbert space $\mathscr{H}$.*

Proof. Let's show that $\mathscr{H} = \mathscr{H}_u$ and the isomorphism π_u answer the question.
Injective. Let $A \in \mathfrak{A}$ such that $\pi_u(A) = 0$; from definition of the direct sum and of universal representation, $\pi_\omega(A) = 0$ for all $\omega \in \mathcal{S}(\mathfrak{A})$. Using equation (72.100) we find that for such a A, we have $0 = \|\pi_\omega(A)\Omega_\omega\|^2 = \omega(A^*A)^2$. It is true for all $\omega \in \mathcal{S}(\mathfrak{A})$. Lemma 72.69 then shows that $\|A^*A\| = 0$ and then that $A = 0$ because of definition of a C^*-algebra.
Surjective. The representation π_u is not specially surjective on $\mathfrak{B}(\mathscr{H}_u)$, but it is surjective on the subalgebra $\pi_u(\mathfrak{A})$ which is enough for the present purpose.
Morphism. The map $\pi_u : \mathfrak{A} \to \mathfrak{B}(\mathscr{H}_u)$ is a morphism because it is a representation.
Isometry. Lemma 72.44 says that an injective morphism of C^*-algebra is isometric.

□

The universal representation is trivially faithful, but it is very huge. For example the smallest faithful representation of $\mathfrak{B}(\mathscr{H})$ is the definition representation on $\mathscr{H}$.

Corollary 72.92.
An operator A is positive if and only if $\pi(A) \geqslant 0$ for all cyclic representation π.

Proof. Since A is positive, then $\sigma(A) \subset \mathbb{R}^+$ and $A^* = A$. In order to prove that $\pi_u(A)$ is positive, we have to show that $\sigma(\pi_u(A)) \subset \sigma(A)$. Let $z \in \sigma(\pi_u(A))$: the operator $\pi_u(A) - z\mathbb{1}_u$ where $\mathbb{1}_u$ is the unit operator on $\mathscr{H}_u$ is not invertible. From equation (72.93), we have $\mathbb{1}_u = \pi_u(\mathbb{1}) = \sum_{\omega\in\mathcal{S}(\mathfrak{A})} \pi_\omega(\mathbb{1})$.

Let $z \notin \sigma(A)$ and let us see that $z \notin \sigma(\pi_u(A))$. From assumption on z, there exists a $B \in \mathfrak{A}$ such that $B(A - z\mathbb{1}) = (A - z\mathbb{1}) = \mathbb{1}$. Then $\pi_u(B)$ is the inverse of $\pi_u(A) - z\pi_u(\mathbb{1})$. It proves that $z \notin \sigma(\pi_u(A))$ and so that $\pi_u(A)$ is positive.

We now prove that $\pi_\omega(A)$ is positive for all GNS representation π_ω. Since π_u acts separately on each space $\mathscr{H}_\omega$, all what we said about the invertibility about π_u can be said for each π_ω.

But we know that all cyclic representation is equivalent to a GNS representation from proposition 72.89. We just have to prove that positivity is conserved by equivalence. Suppose that it is not the case. Let $z \in \sigma(\pi_\omega(A))$ and $B \in \mathfrak{A}$ such that $B\big(\pi_\omega(A) - z\mathbb{1}_\omega\big) = \mathbb{1}_\omega$. Then

$$UBU^*\big(\pi(A) - z\mathbb{1}_\omega\big) = \mathbb{1}_\omega$$

and then UBU^* is the inverse of $\pi(A) - z\mathbb{1}_\omega$ and $z \notin \sigma(\pi(A))$. Thus

$$\sigma(\pi(A)) \subset \sigma(\pi_\omega(A)) \subset \mathbb{R}^+$$

which proves that $\pi(A)$ is positive.

We now prove the second sense of the corollary. All GNS representation is cyclic, then $\pi_u(A)$ is positive as sum of positive representation. We want to deduce that $A \geqslant 0$. Since $\pi_u(A)$ is positive, $\pi_u(A) = \pi_u(A)^* = \pi_u(A^*)$. This implies that $A = A^*$ because π_u is injective. The positivity of $\pi_u(A)$ gives the existence of a B such that $\pi_u(A) = \pi_u(B)^*\pi_u(B) = \pi_u(B^*B)$. The injectivity then shows that A is positive.

□

The GNS construction is done for unital algebras with states. Since there exists a notion of state on non unital algebras, ones raises the question to generalization of the GNS construction to non unital algebras.

72.12 Spaces of matrices

Let $\mathfrak{A}$ be a C^*-algebra and $n \in \mathbb{N}$. The C^*-algebra $\mathfrak{M}^n(\mathfrak{A})$ is the space on $n \times n$ matrices with entries in $\mathfrak{A}$.The multiplication is defined by

$$(MN)_{ij} = \sum_k M_{ik}N_{kj} \tag{72.102}$$

where the product in the right hand side is the (in general noncommutative) one in $\mathfrak{A}$. The involution is naturally given by

$$(M^*)_{ij} = M^*_{ij}. \tag{72.103}$$

One can identify $\mathfrak{M}^n(\mathfrak{A})$ to $\mathfrak{A}\otimes\mathfrak{M}^n(\mathbb{C})$ by identifying[11] $E_{ij}\in\mathfrak{M}^n(\mathfrak{A})$ to $A\otimes E_{ij}\in\mathfrak{A}\otimes\mathfrak{M}^n(\mathbb{C})$.

The Gelfand-Neumark gives the existence of a faithful representation π of $\mathfrak{A}$ on $\mathscr{H}$. If one sees elements of $\mathscr{H}\otimes\mathbb{C}^n$ as n-uples $(v_1,\dots,v_n)$ where each $v_i\in\mathscr{H}$, we can define π_n on $\mathscr{H}\otimes\mathbb{C}^n$ by linear extension of

$$\big[\pi_n(M)v\big]_i := \pi(M_{ij})v_j. \tag{72.104}$$

eq:defreprezmfM

The norm $\|M\|$ is defined as the norm of $\pi_n(M)$. Since $n<\infty$, $\pi_n(\mathfrak{M}^n(\mathfrak{A}))$ is a closed $*$-algebra in $\mathfrak{B}(\mathscr{H}\otimes\mathbb{C}^n)$ and then $\mathfrak{M}^n(\mathfrak{A})$ is a C^*-algebra for this norm. Proposition 72.70 states that the norm is unique and then that the norm $\|M\|=\pi_n(M)$ is independent of the choice of π.

DefComplPositive

Definition 72.93.
*A linear map $q\colon\mathfrak{A}\to\mathcal{B}$ is **completely positive** if for all $n\in\mathbb{N}$, the map $q_n\colon\mathfrak{M}^n(\mathfrak{A})\to\mathfrak{M}^n(\mathcal{B})$ defined by*

$$(q_n(M))_{ij}=q(M_{ij})$$

is positive.

As an example, a morphism φ is always completely positive because if $a=b^*b$ in $\mathfrak{M}^n(\mathfrak{A})$, then $\varphi(a)=\varphi(b)^*\varphi(b)$ which is positive in $\mathfrak{M}^n(\mathcal{B})$.

72.13 Stinespring theorem

Let us state a classical result about Hilbert space

Lemma 72.94.
If K is closed in a Hilbert space $\mathscr{H}$ and if the linear operator $T\colon\mathscr{H}\to\mathscr{H}$ fulfils

- *$\langle Tx,y\rangle=\langle x,Ty\rangle$ for all x, $y\in\mathscr{H}$,*
- *$Ty=y$ for all $y\in K$*
- *$Tz=0$ for all $z\in K^\perp$,*

then T is the projection on K.

This lemma allows us to check that WW^* is the projection to the image of W. Let us prove that W^*W is the projection on K_1. The first condition is clear. The third is satisfied by definition of W: if $z\in K_1^\perp$, then $W^*Wz=0$. For the second one, remark that if $y\in K_1$, $\langle W^*Wx,y\rangle=\langle x,y\rangle$ and if $y\in K_1^\perp$, then $\langle W^*Wx,y\rangle=0=\langle x,y\rangle$. In both cases $y\in K_1$ and $y\in K_1^\perp$, we have $\langle W^*Wx,y\rangle=\langle x,y\rangle$. It is sufficient to conclude that $W^*Wx=x$ because $\mathscr{H}=K_1\oplus K_1^\perp$.

Theorem 72.95.
Let $q\colon\mathfrak{A}\to\mathcal{B}$ be a completely positive map between unital C^*-algebra such that $q(\mathbb{1})=\mathbb{1}$. We suppose that $\mathcal{B}$ is given with a faithful representation $\mathcal{B}\simeq\pi_\chi(\mathcal{B})\subseteq\mathfrak{B}(\mathscr{H}_\chi)$ for a certain Hilbert space $\mathscr{H}_\chi$. Then there exists a Hilbert space $\mathscr{H}^\chi$, a representation π^χ of $\mathfrak{A}$ on $\mathscr{H}^\chi$ and a partial isometry $W\colon\mathscr{H}_\chi\to\mathscr{H}^\chi$ with $W^*W=\mathbb{1}$ and

$$\pi_\chi(q(A))=W^*\pi^\chi(A)W \tag{72.105}$$

eq:stinun

for all $A\in\mathfrak{A}$.

Stated in an equivalent way, if we define $P=WW^*$ and $\tilde{\mathscr{H}_\chi}=P\mathscr{H}^\chi\subset\mathscr{H}^\chi$, and $U\colon\mathscr{H}_\chi\to\tilde{\mathscr{H}_\chi}$ as the restriction of W (in such a way that U is unitary because W is a partial isometry), then we have

$$U\pi_\chi(q(A))U^{-1}=P\pi^\chi(A)P. \tag{72.106}$$

eq:stindeux

tho:stinespring

11. The matrix E_{ij} is the matrix full of zero except a 1 at position ij.

Proof. On $\mathscr{H}_\chi$ we have the scalar product $\langle .,. \rangle_\chi$ and we define the sesquilinear form $\langle .,. \rangle_0^\chi$ on $\mathfrak{A} \otimes \mathscr{H}_\chi$ by sesquilinear extension of

$$(A \otimes v, B \otimes w)_0^\chi = \langle v, \pi_\chi(q(A^*B))w \rangle_\chi.$$

This form is semi positive definite. Let us compute

$$\sum_{ij} (A_i \otimes v_i, A_j \otimes v_j)_0^\chi = \sum_{ij} \langle v_i, \pi(q(A_j^* A_j))v_j \rangle_\chi. \tag{72.107}$$

For this, we consider $a \in \mathfrak{M}^n(\mathfrak{A})$ with elements $a_{ij} = A_i^* A_j$. Let π_n be the faithful representation (72.104) of $\mathfrak{M}^n(\mathfrak{A})$ on $\mathscr{H} \otimes \mathbb{C}^n$ defined by

$$[\pi_n(M)v]_i = \sum_j \pi(M_{ij})v_j \in \mathscr{H}.$$

where π is a faithful representation of $\mathfrak{A}$. If $z \in \mathbb{C}^n \otimes \mathscr{H}$, we can define $(az) \in \mathbb{C}^n \otimes \mathscr{H}$ by

$$(az)_i = \big(\pi_n(a)z\big)_i.$$

So we have

$$\langle z, az \rangle = \sum_{ij} \langle z_i, \pi(a_{ij})z_j \rangle = \sum_{ij} \langle \pi(A_i)z_i, \pi(A_j)z_j \rangle = \|Az\|^2 \geqslant 0 \tag{72.108}$$

where we use the notation $Az = \sum_i \pi(A_i)z_i \in \mathscr{H}$. The conclusion is that $a \geqslant 0$. Since q is completely positive, b is positive if we define $b_{ij} = q(A_i^* A_j)$. Positivity of b gives rise to an element $c \in \mathfrak{M}^n(\mathcal{B})$ such that $n = c^*c$ where $(c^*)_{ij} = (c_{ji})^*$. From the representation π_χ of $\mathcal{B}$, we can build the faithful representation π'_n of $\mathfrak{M}^n(\mathcal{B})$ on $\mathbb{C}^n \otimes \mathscr{H}_\chi$: $[\pi'_n(b)v]_i = \sum_j \pi_\chi(b_{ij})v_j$ where each v_i now belongs to $\mathscr{H}_\chi$.

We are now able to prove that $\langle .,. \rangle_0^\chi$ is a positive form. Indeed

$$\begin{aligned} \sum_{ij} (A_i \otimes v_i, A_j \otimes v_j)_0^\chi &= \sum_{ij} \langle v_i, \pi_\chi\big(q(A_i^* A_j)\big)v_j \rangle_\chi = \sum_{ij} \langle v_i, \pi_\chi(b_{ij})v_j \rangle_\chi \\ &= \sum_{ijk} \langle \pi_\chi(c_{ki})v_i, \pi_\chi(c_{kj})v_j \rangle_\chi \geqslant 0 \end{aligned} \tag{72.109}$$

where the last equality comes from the fact that $(c^*c)_{ij} = \sum_k (c^*)_{ik} c_{kj} = \sum_k (c_{ki})^* c_{kj}$. This proves that $(.,.)_0^\chi$ is positive semi definite.

Definition of π^χ..

Now we denote by $\mathcal{N}_\chi$ the null space of $(.,.)_0^\chi$. Let $V_\chi : \mathfrak{A} \otimes \mathscr{H}_\chi \to \mathfrak{A} \otimes \mathscr{H}_\chi / \mathcal{N}_\chi$ be the canonical projection. We define

$$\langle V_\chi(A \otimes v), V_\chi(B \otimes w) \rangle^\chi := (A \otimes v, B \otimes w)_0^\chi \tag{72.110}$$

and we denote by $\mathscr{H}^\chi$ the closure of $\mathfrak{A} \otimes \mathscr{H}_\chi / \mathcal{N}_\chi$ with respect to this scalar product. We can now define π^χ, a representation of $\mathfrak{A}$ on $\mathfrak{A} \otimes \mathscr{H}_\chi / \mathcal{N}_\chi$ by linear extension of

$$\pi^\chi(A) V_\chi(B \otimes w) = V_\chi(AB \otimes w) \tag{72.111}$$

which is well defined because $\pi^\chi(A)\mathcal{N}_\chi \subseteq \mathcal{N}_\chi$.

Let us now prove that $\|\pi^\chi(A)\| \leqslant \|A\|$. From equation (72.33) used in $\mathfrak{M}^n(\mathfrak{A})$ with $B = \mathbb{1}_n$,we know that

$$0 \leqslant A^* A \mathbb{1}_n \leqslant \|A\|^2 \mathbb{1}_n. \tag{72.112}$$

esr502061

Now we consider any $B_1, \cdots, B_n \in \mathfrak{A}$ and we build the matrix

$$b = \begin{pmatrix} B_1 & \cdots & B_n \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \quad b^* = \begin{pmatrix} B_1^* & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ B_n^* & 0 & \dots & 0 \end{pmatrix}.$$

We conjugate (72.112) with b:

$$0 \leqslant b^*A^*Ab \leqslant \|A\|^2 b^*b,$$

but q is completely positive, then it respects the inequality:

$$q_n(b^*A^*Ab) \leqslant \|A\|^2 q_n(b^*b)$$

where $q_n\colon \mathfrak{M}^n(\mathfrak{A}) \to \mathfrak{M}^n(\mathcal{B})$ is defined by $\big(q_n(M)\big)_{ij} = q(M_ij)$. The definition of complete positivity of q is precisely positivity of q_n. We consider now the representation π_χ of $\mathcal{B}$ on $\mathscr{H}_\chi$. Since $a \geqslant 0$, we can find a $e \in \mathfrak{M}^n(\mathcal{B})$ such that $q_n(a) = e^*e$; then

$$\begin{aligned}\sum_{ij}\langle v_i, \pi_\chi\big(q(a_{ij})v_j\big)\rangle &= \sum_{ij}\langle v_i, \pi_\chi(e^*e)_{ij}v_j\rangle \\ &= \sum_{ijk}\langle v_i, \pi_\chi(e^*_{ki})\pi_\chi(e_{kj}\rangle \\ &= \sum_{ijk}\langle \pi_\chi(e)_{ki}v_i, \pi_\chi(e_{ij})v_j\rangle \geqslant 0.\end{aligned} \tag{72.113}$$

Let us consider $\Psi = \sum_i V_\chi B_i \otimes v_i$ and compute

$$\begin{aligned}\|\pi^\chi(A)\Psi\|^2 &= \sum_{ij}(AB_i \otimes v_i, AB_j \otimes v_j)^\chi_0 \\ &= \sum_{ij}\langle v_i, \pi_\chi\big(q(B_i^*A^*AB_j)\big)v_j\rangle_\chi \\ &\leqslant \|A\|^2 \sum_{ij}\langle v_i, \pi_\chi\big(q(B_i^*B_j)\big)v_j\rangle_\chi \\ &= \|A\|^2 \sum_{ij}\langle B_i \otimes v_i, B_j \otimes v_j\rangle_\chi \\ &= \|A\|^2 (V_\chi \sum_i B_i \otimes v_i, V_\chi \sum_j B_j \otimes v_j)^\chi \\ &= \|A\|^2 \|\Psi\|^2.\end{aligned} \tag{72.114}$$

Then $\|\pi^\chi(A)\Psi\| \leqslant \|A\|^2\|\Psi\|$ which proves that

$$\|\pi^\chi(A)\| \leqslant \|A\|^2. \tag{72.115}$$

The formula

$$\pi^\chi(A)V_\chi(B \otimes w) = V_\chi(AB \otimes w) \tag{72.116}$$

defines a continuous representation π^χ on $\mathfrak{A} \otimes \mathscr{H}_\chi/\mathcal{N}_\chi$ which can be extended to a continuous representation on the whole $\mathscr{H}^\chi$. This extension fulfils $\pi^\chi(A^*) = \pi^\chi(A)^*$.

We define $W\colon \mathscr{H}_\chi \to \mathscr{H}^\chi$ by

$$Wv = V_\chi \mathbb{1} \otimes v. \tag{72.117}$$

The map W is a partial isometry.

In order to prove that W is a partial isometry, just compute

$$\begin{aligned}(Wv, Ww)^\chi &= (V_\chi \mathbb{1} \otimes v, V_\chi \mathbb{1} \otimes w)^\chi \\ &= (\mathbb{1} \otimes v, \mathbb{1} \otimes w)^\chi_0 \\ &= \langle v, w\rangle_\chi.\end{aligned} \tag{72.118}$$

Adjoint of W.

We claim that $W^*\colon \mathscr{H}^\chi \to \mathscr{H}_\chi$, $W^*V_\chi A \otimes v = \pi_\chi(q(A))v$ is the adjoint of W. Recall that the definition of the adjoint requires that

$$\langle w, W^*\psi\rangle_\chi = (Ww, \psi)^\chi$$

for all $w \in \mathscr{H}_\chi$ and all $\psi \in \mathscr{H}^\chi = \overline{A \otimes \mathscr{H}_\chi / \mathcal{B}_\chi}$. A $\psi \in \mathscr{H}^\chi$ can be written under the form $V_\chi A \otimes v$ with $A \in \mathfrak{A}$ and $v \in \mathscr{H}_\chi$; then

$$\begin{aligned} \langle w, W^* V_\chi A \otimes v \rangle_\chi &= \langle w, \pi_\chi(q(A)) v \rangle_\chi \\ &= (\mathbb{1} \otimes w, A \otimes v)_0^\chi \\ &= (V_\chi \mathbb{1} \otimes w, V_\chi A \otimes v)^\chi \\ &= \langle W w, \psi \rangle \end{aligned} \tag{72.119}$$

as expected. One can check that $W^*W = \mathbb{1}$ and that $W^* \pi^\chi(A) W = \pi_\chi(q(A))$ because

$$W^* \pi^\chi(A) W v = W^* \pi^\chi(A) V_\chi(\mathbb{1} \otimes v) = W^* V_\chi(A \otimes v) = \pi_\chi(q(A)). \tag{72.120}$$

Last point: (72.105)⇒(72.106).

Since W is a partial isometry, $P = WW^*$ is the projection on the image of W and W^*W $(= \mathbb{1})$ is the projector on the subspace of $\mathscr{H}_\chi$ on which W is isometric; this space is $\mathscr{H}_\chi$ itself. The $\tilde{\mathscr{H}_\chi} = P\mathscr{H}^\chi = \mathscr{H}^\chi$ and $U = W$. Then

$$U \pi_\chi\big(q(A)\big) U^{-1} = W \pi_\chi\big(q(A)\big) W^* = W W^* \pi^\chi(A) W W^* = P \pi^\chi(A) P \tag{72.121}$$

This concludes the proof of theorem 72.95.

□

Remarque 72.96.
If on the one hand $q(\mathbb{1})$ is not $\mathbb{1}$, then the construction works, but W is no more a partial isometry and we have

$$\|W\|^2 = \|q(\mathbb{1})\|,$$

so $\mathscr{H}_\chi$ can not be seen as a subspace of $\mathscr{H}^\chi$ by the map $W\colon \mathscr{H}_\chi \to \mathscr{H}^\chi$.

If on the other hand $\mathfrak{A}$ or $\mathcal{B}$ is not unital, then works if q can be extended (keeping positive) to the unitization of $\mathfrak{A}$ in such a way that it conserves the identity in (the unitization of) $\mathcal{B}$.

Proposition 72.97.
Any positive map $q\colon \mathfrak{A} \to \mathcal{B}$ from a commutative unital C^*-algebra $\mathfrak{A}$ is completely positive.

The proof will be decomposed into several propositions. Let us begin by a remark: from theorem 72.15, we can write $\mathfrak{A} = C(X)$ for a certain locally compact Hausdorff space X. So we can identify $\mathfrak{M}^n(C(X))$ with $C(X, \mathfrak{M}^n(\mathbb{C}))$ because to each $a \in \mathfrak{M}^n(C(X))$, $(a_{ij}$ is a map $a_{ij}\colon X \to \mathbb{C})$ we make correspond the map $\eta\colon X \to \mathfrak{M}^n(\mathbb{C})$ defined by $\eta(x)_{ij} = a_{ij}(x)$.

Proposition 72.98.
The set of finite linear combinations of elements F of the form

$$F(x) = \sum_i f_i(x) M_i$$

with $f_i \in C(X)$ and $M_i \in \mathfrak{M}^n(\mathbb{C})$ is dense in $C(X, \mathfrak{M}^n(\mathbb{C}))$. prop:lencombpart

Proof. Let $G \in C(X, \mathfrak{M}^n(\mathbb{C}))$ and $\varepsilon > 0$. Continuity of G makes the set

$$\mathcal{O}_x^\varepsilon = \{y \in X \text{ tel que } \|G(x) - G(y)\| \leqslant \varepsilon\}$$

open for all $x \in X$. These set give an open covering of the compact space X, then we can extract a finite subcovering and build an unity partition φ_i. We define $F_l \in C\big(X, \mathfrak{M}^n(\mathbb{C})\big)$ by

$$F_l(x) = \sum_{i=1}^{l} \varphi_i(x) G(x_i). \tag{72.122}$$ eq:Fllim

We have

$$\|F_l(x) - G(x)\| = \|\sum_{i=1}^{l} \varphi(x)(G(x_i) - G(x))\| \leqslant \sum \varphi_i(x)\|G(x_i) - G(x)\| \leqslant \sum \varphi_i \varepsilon = \varepsilon \qquad (72.123)$$

Then $\|F_j - G\| = \sum_{x\in X} \|F_l(x) - G(x)\| \leqslant \varepsilon$ and the sequence F_l converges to G. It proves the density. □

Proposition 72.99.
When $\{M_i\}$ is a basis of $\mathfrak{M}^n(\mathbb{C})$ composed with positive elements, an element $F \in C(X, \mathfrak{M}^n(\mathbb{C}))$ of the form $F(x) = \sum_i f_i(x)M_i$ is positive if and only if it each of f_i is positive.

Proof. We know that $F \in C(X, \mathfrak{M}^n(\mathbb{C}))$ is positive when $F(x)$ is positive in $\mathfrak{M}^n(\mathbb{C})$ for all $x \in X$. We have $F(x) = \sum_i f_i(x)M_i$, but positivity of $F(x)$ requires $F(x) = F(x)^*$ and then $f_i(x) = f_i(x)^*$ because M_i is positive.

□

¡¡ Avertissement/question à la lectrice !! 72.100
[489] states a stronger result that seems wrong to me because $-3 + 7$ is positive.

Proposition 72.101.
If $G \in C(X, \mathfrak{M}^n(\mathbb{C}))$ is positive, then there exists a sequence $F_k \geqslant 0$ such that $\lim_{k\to\infty} F_k = G$

Proof. Each element F_l (72.122) is positive because $G(x_i)$ is positive for all x_i. □

¡¡ Avertissement/question à la lectrice !! 72.102
There are too much unclear thinks in my mind; I do not finish the proof.

The main result is the following proposition.

Proposition 72.103.
If $\mathfrak{A}$ is a commutative unital C-algebra, then any positive map $q\colon \mathfrak{A} \to \mathcal{B}$ is completely positive.*

72.14 Representations

As notational convention, when H is a Hilbert space, we denote by $\mathcal{L}(H)$ the set of the continuous endomorphism of H. Topology on $\mathcal{L}(H)$ is discussed in subsection 60.2.2.

When π is a representation of $\mathfrak{A}$ in H and $\xi \in H$, the space $\overline{\pi(\mathfrak{A})\xi}$ is a closed subspace of H stable under $\pi(\mathfrak{A})$.

Definition 72.104.
*A vector $\xi \in H$ is said **totalizing** for a representation π of $\mathfrak{A}$ if $\overline{\pi(\mathfrak{A})\xi} = H$.*

Proposition 72.105.
Let $\mathfrak{A}$ be an involutive algebra, H an hermitian space and π a representation of $\mathfrak{A}$ in H. Then the following facts are equivalent :

(1) The only closed subspace in H which are stable for $\pi(\mathfrak{A})$ are $\{o\}$ and H.

(2) The subset of $\mathcal{L}(H)$ which commutes with $\pi(\mathfrak{A})$ is reduced to $\mathcal{C}$.

(3) Any non zero vector in H is totalizing for π, or π have dimension 1.

prop:reprez_topo

Proof. We begin proving that (ii) implies (iii). Let $\xi \in H$, $\xi \neq 0$. If $\pi(\mathfrak{A})\xi$ is not everywhere dense in H, (i) makes $\pi(\mathfrak{A})\xi = 0$. Then $\mathbb{C}\xi$ is stable under $\pi(\mathfrak{A})$. But $\mathbb{C}\xi \neq \{o\}$, then $\mathbb{C}\xi = H$. Thus H has dimension 1 and π is the null representation.

Now, we prove that (iii) implies (i). Let $K \neq \{o\}$ be a closed vector subspace of H stable under $\pi(\mathfrak{A})$. We have to show that $K = H$. If $\dim H = 1$, it is obvious. Let us consider a non zero $\xi \in K$. Since K is stable, $\pi(\mathfrak{A})\xi \subset K$, but (iii) implies $\overline{\pi(\mathfrak{A})\xi} = H$. Thus $K = H$.

We turn our attention to the equivalence between (i) and (ii). First (ii) implies (i). We consider K, a vector subspace of H stable under $\pi(\mathfrak{A})$. We want $K = \{o\}$ or $K = H$. Let us consider $p_K : H \to K$, the orthogonal projection. Since K is a vector subspace, it makes sense to write $H \ominus K$.

Let us consider $\xi \in K$ and $\eta \in H \ominus K$. For any $A \in \mathfrak{A}$, $\pi(A^*)\xi \in K$ because $\pi(A)\xi \in K$. This yields

$$\langle \pi(A)\eta|\xi\rangle = \langle \eta|\pi(A^*)\xi\rangle = 0,$$

then $\pi(A)\eta \in H \ominus K$. We can conclude that

$$[p_K, \pi(\mathfrak{A})] = 0$$

because $\pi(\mathfrak{A})p_K\xi = \pi(\mathfrak{A})\xi = p_K\pi(\mathfrak{A})\xi$, $\pi(\mathfrak{A})p_K\eta = 0$ and $p_K\pi(\mathfrak{A})\eta = 0$ from $\langle \pi(\mathfrak{A})\eta|\xi\rangle = 0$.

From (ii), p_K is then a scalar operator: $p_K = 0$ or $p_K = id$, *i.e.* $K = \{o\}$ or $K = H$.

Finally, we show the implication from (i) to (ii). Let T be an element of $\mathcal{L}(H)$ which commute with the whole $\pi(\mathfrak{A})$; we have to show that T is scalar. Since it is clear that $T + T^*$ and $T - T^*$ also commute with $\pi(\mathfrak{A})$, we can suppose $T = T^*$.

We had shown that a projector p_K commutes with $\pi(\mathfrak{A})$ if K is stable under $\pi(\mathfrak{A})$, closed and a sub vector space of H. This is the case of the eigenspaces because $(T - \lambda\mathbb{1})v = 0$, then $(T - \lambda\mathbb{1})\pi(\mathfrak{A})v = 0$ because $[T, \pi(\mathfrak{A})] = 0$.

The spectral projector of T commute with $\pi(\mathfrak{A})$. Thus, the eigenspaces H_λ are stable under $\pi(\mathfrak{A})$, thus (by (i)) these are only $\{o\}$ and H. In other words $(T - \lambda\mathbb{1})v = 0$ has solutions which are on two subspaces whose projectors are 0 and 1. On the space on which $p_\lambda = 1$, $T = \lambda\mathbb{1}$ and the one where $p_\lambda = 0$ is $\{0\}$ then $T = 0$.

¡¡ Avertissement/question au lecteur !! 72.106
We have to show that every v belong to one of these two spaces. Is it because T is hermitian, or do we need the compact assumption?

□

Definition 72.107.
*Let $\mathfrak{A}$ be an involutive algebra, H a hermitian space and π a representation of $\mathfrak{A}$ in H. We say that π is **topologically irreducible** if it fulfils proposition 72.105. The representation π is **algebraically irreducible** if the only stable vector subspaces of H under $\pi(\mathfrak{A})$ are $\{0\}$ and H.*

When $\dim H = \infty$, the second notion is stronger because it excludes the case where one has an *open* stable subspace. Point 2.8.4 in [959] shows that in the case of C^*-algebras, a topologically irreducible representation is automatically algebraically irreducible, so one can simply speak about irreducible representations.

72.15 Pure states

A subset C of a vector space is **convex** if for all $\lambda \in [0, 1]$ and $v, w \in C$, the element $\lambda v + (1-\lambda)w$ belongs to C.

An **extreme point** of a convex set K is an element $\omega \in K$ which can de written under the form $\omega = \lambda\omega_1 + (1 - \lambda)\omega_2$ with $\lambda \in [0, 1]$ only for $\omega_1 = \omega_2 = \omega$.

Definition 72.108.
*An extreme point of the state space $\mathcal{S}(\mathfrak{A})$ is a **pure state** and states that are not pure are **mixed states**.*

The set of extreme points of the convex set K is denoted by $\partial_c K$ and is called the **extreme boundary** of K. An extreme point in the state space $\mathcal{S}(\mathfrak{A})$ of a C^*-algebra is a *pure state* and other states are **mixed states**. As notation, $\partial_c(\mathcal{S}(\mathfrak{A}))$ is denoted by $\mathcal{P}(\mathfrak{A})$ or simply $\mathcal{P}$ when there are no ambiguity.

Example $\mathfrak{A} = \mathbb{C} \oplus \mathbb{C}$

Points are given by (λ, μ). Let's consider the convex set of states $r \in [0,1]$ defined by $r(\lambda,\mu) = (1-r)\lambda - r\mu$. Extreme points are given by 0 and 1: $0(\lambda,\mu) = \lambda$ and $1(\lambda,\mu) = \mu$.

Example: $\mathfrak{A} = \mathfrak{M}^2(\mathbb{C})$

We can identify $\mathfrak{M}^2(\mathbb{C})$ with its dual in the following way. A linear form ω on $\mathfrak{A}$ can always be written as

$$\omega \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \omega_{11}A_{11} + \omega_{12}A_{21} + \omega_{21}A_{12} + \omega_{22}A_{22}$$

So to each form ω, one can associate the matrix of ω_{ij} and the following holds:

$$\omega(A) = \mathrm{Tr}(\omega A)$$

The identification between $\mathfrak{M}^2(\mathbb{C})$ and its dual is then well given by $\omega \simeq (\omega_{ij})$. Let us see in terms of this identification the set $\mathcal{S}(\mathfrak{A})$. The condition $\omega(A) \geqslant 0$ imposes to the matrix (ω_{ij}) to be positive and the condition $\omega(\mathbb{1}) = 1$ imposes $\mathrm{Tr}\,\omega = 1$. Then $\mathcal{S}(\mathfrak{A})$ is parametrized by

$$\rho = \frac{1}{2}\begin{pmatrix} 1+x & y+iz \\ y-iz & 1-x \end{pmatrix}$$

where x, y, $z \in \mathbb{R}$. The pure states are such matrices ρ with $x^2+y^2+z^2 = 1$.

If M is a $*$-algebra in $\mathfrak{B}(\mathscr{H})$, the **commutant** M' is

$$M' = \{A \in \mathfrak{B}(\mathscr{H}) \text{ tel que } [A,m] = 0 \forall\, m \in M\}.$$

Proposition 72.109.
The following properties are equivalent:

enumgz

(1) The representation $\pi(\mathfrak{A})$ is irreducible in $\mathscr{H}$.

enumgi

(2) The commutant of $\pi(\mathfrak{A})$ in $\mathfrak{B}(\mathscr{H})$ is

$$\pi(\mathfrak{A})' = \{\lambda\mathbb{1} \text{ tel que } \lambda \in \mathbb{C}\}.$$

enumgii

(3) $\pi(\mathfrak{A})'' = \mathfrak{B}(\mathscr{H})$.

enumgiii

(4) Each vector $\Omega \in \mathscr{H}$ is cyclic for $\pi(\mathfrak{A})$.

prop_equiv_rep_irred

Proof. (3)⇒(2) We have to see that an operator which commutes with the whole $\mathfrak{B}(\mathscr{H})$, then it is a multiple of identity. If $[a,A] = 0$ for all $A \in \mathfrak{B}(\mathscr{H})$, then it commutes in particular with an operator A which leaves the basis vector e_β (and only this basis vector). In this case, $ae_\beta = \lambda_\beta e_\beta$. We conclude that a must be diagonal. Since a must also commute with an operator which leaves $e_\alpha + e_\beta$ unchanged, we conclude that $\lambda_\alpha = \lambda_\beta$, so that $a = \lambda\mathbb{1}$.

(1)⇒(2) Will be done later. (2)⇒(1) Let us suppose that $\pi(\mathfrak{A})' = \mathbb{C}\mathbb{1}$ and that π is irreducible; we will find out a contradiction. We have a non trivial subspace of $\mathcal{H}$ stable under $\pi(\mathfrak{A})$. The projection operator on this space commutes with the whole $\pi(\mathfrak{A})$ although it is not a multiple of identity.

(1)⇒(4) We proceed by contradiction once again. Let $\psi \in \mathscr{H}$ such that $\pi(\mathfrak{A})\psi$ is not dense in $\mathscr{H}$ and P be the projection on the closure of $\pi(\mathfrak{A})\psi$. Lemma 72.85 assures that $P \in \pi(\mathfrak{A})'$. Consequently $\pi(\mathfrak{A})' \neq \mathbb{C}\mathbb{1}$ and $\pi(\mathfrak{A})$ is not irreducible by (2).

(4)⇒(1) If ψ belongs to a (non trivial) invariant subspace, then $\pi(\mathfrak{A})\psi$ cannot be dense because it is a proper subspace of $\mathscr{H}$.

□

¡¡ Avertissement/question à la lectrice !! 72.110
There are still unfinished points in that proof.

Let us point out the following part of the proposition:

Lemma 72.111 (Schur's lemma)**.**
The representation π on $\mathfrak{A}$ is irreducible if and only if the commutant of $\pi(\mathfrak{A})$ in $\mathfrak{B}(\mathscr{H})$ is

$$\pi(\mathfrak{A})' = \{\lambda \mathbb{1}\}_{\lambda\in\mathbb{C}}.$$

Lemma 72.112.
Let $\hat{Q}$ be a bounded quadratic form on a Hilbert space $\mathscr{H}$. There exists a bounded operator Q on $\mathscr{H}$ such that for each $\psi, \phi \in \mathscr{H}$ we have

$$\hat{Q}(\psi, \phi) = \langle \psi, Q\phi \rangle$$

and $\|Q\| \leqslant C$ where C is the "bounding constant": $|\hat{Q}(\psi, \phi)| \leqslant \|\psi\|\|\phi\|$. Moreover if

$$\hat{Q}(\phi, \psi) = \overline{\hat{Q}(\psi, \phi)} \tag{72.124}$$

eq_r19032

the operator Q will be selfadjoint. lem_r19031

Proof. Let us fix a $\psi \in \mathscr{H}$ and look at the map $\phi \mapsto \hat{Q}(\psi, \phi)$. It is a bounded form, so Riesz theorem gives the existence of a $\Omega \in \mathscr{H}$ such that $\hat{Q}(\psi, \phi) = \langle \Omega, \phi \rangle$. We can define Q by $Q\psi = \Omega$. It is clear that is is self-adjoint if equation (72.124) is satisfied.

From equalities $\hat{Q}(\psi, \phi) = \langle \Omega, \phi \rangle = \langle Q\psi, \phi \rangle$, we find

$$\begin{aligned} \|Q\psi\|^1 &= |\langle Q\psi, Q\psi \rangle| \\ &= |\hat{Q}(Q\psi, \psi)| \\ \leqslant \|Q\psi\|\|\psi\| & \\ &= \|Q\|\|\psi\|^2. \end{aligned} \tag{72.125}$$

Taking supremum on $\|\psi\| = 1$, we find $\|Q\|^2 \leqslant C\|Q\|$ and then

$$\|Q\| \leqslant C.$$

□

Theorem 72.113.
The GNS representation π_ω of a state $\omega \in \mathcal{S}(\mathfrak{A})$ is irreducible if and only if ω is pure. tho_GNS_irred_pure

Proof. Let us begin to suppose that ω is a pure state and that π_ω is reducible. Then the projection P onto the invariant subspace K of $\pi(\mathfrak{A})$ belongs to $\pi(\mathfrak{A})'$ (see the proof of Schur's lemma). Let Ω_ω be the cyclic vector of π_ω.

If $P\Omega_\omega = 0$, then for each $A \in \mathfrak{A}$ we have

$$0 = \pi_\omega P\Omega_\omega = P\pi_\omega(A)\Omega_\omega.$$

Since Ω_ω is cyclic, it proves that $P = 0$ which is impossible if π_ω is reducible. For the same reason, $P^\perp\Omega_\omega$ is neither not possible because it should implies that $P = \mathbb{1}$.

We define the two following states on $\mathfrak{A}$:

$$\psi(A) = k\langle P\Omega_\omega, \pi(A)P\Omega_\omega \rangle \tag{72.126a}$$

$$\psi^\perp(A) = l\langle P^\perp\Omega_\omega, \pi_\omega(A)P^\perp\Omega_\omega \rangle \tag{72.126b}$$

From definition of a projection, we have $\langle Px, Py \rangle = \langle P^*Px, y \rangle = \langle Px, y \rangle$. Taking any k, $\lambda = 1/k$ and $l = 1/(1 - 1/k)$, using the relation $\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle$ we find

$$\omega = \lambda\psi + (1 - \lambda)\psi^\perp,$$

which is in contradiction with the fact that ω is pure.

We now suppose that π_ω is irreducible, and that ω reads

$$\omega = \lambda\omega_1 + (1-\lambda)\omega_2$$

with $\lambda \in [0,1]$ and $\omega_1, \omega_2 \in \mathcal{S}(\mathfrak{A})$. We will prove that ω_1 is proportional to ω, so that ω is pure. Since elements of $\mathcal{S}(\mathfrak{A})$ are positive and $(1-\lambda) \geqslant 0$, the form $\omega - \lambda\omega_1 = (1-\lambda)\omega_2$ is positive. Therefore for all $A \in \mathfrak{A}$, we have $\lambda\omega_1(A^*A) \leqslant \omega(A^*A)$. From equation (72.62), we find

$$\begin{aligned} |\lambda\omega_1(A^*B)| &\leqslant \lambda^2\omega_1(A^*A)\omega_1(B^*B) \\ &\leqslant \omega(A^*A)\omega(B^*B). \end{aligned} \tag{72.127}$$

eq_r19031

It allows us to define a quadratic form $\hat{Q}$ on $\pi_\omega(\mathfrak{A})\Omega_\omega$ by

$$\hat{Q}\big(\pi_\omega(A)\Omega_\omega, \pi_\omega(B)\Omega_\omega\big) := \lambda\omega_1(A^*B). \tag{72.128}$$

$\hat{Q}$ is well defined.

We have to prove that $\pi_\omega(A_1) = \pi_\omega(A_2)\Omega_\omega$ implies

$$\hat{Q}(\pi_\omega(A_1)\Omega_\omega, \ \cdot\ ) = \hat{Q}(\pi_\omega(A_2)\Omega_\omega, \ \cdot\ ).$$

Equation (72.100) gives

$$\|\pi_\omega(A)\Omega_\omega\|^2 = \omega(A^*A)$$

and makes that $\omega\big((A_1 - A_2)^*(A_1 - A_2)\big) = 0$. Thus, using (72.127), we have

$$\begin{aligned} &\Big|\hat{Q}(\pi_\omega(A_1)\Omega_\omega, \pi_\omega(B)\Omega_\omega) - \hat{Q}(\pi_\omega(A_2)\Omega_\omega, \pi_\omega(B)\Omega_\omega)\Big|^2 \\ &= \Big|\lambda\omega_1(A_1^*B) - \lambda\omega_1(A_2^*B)\Big|^2 \\ &= \Big|\lambda\omega_1\big((A_1 - A_2)^*B\big)\Big|^2 \\ &\leqslant 0. \end{aligned} \tag{72.129}$$

The whole is finally zero.

$\hat{Q}$ is bounded.

Equation (72.100) together with the equality $|\lambda\omega_1(A^*B)|^2 \leqslant \omega(A^*A)\omega(B^*B)$ give

$$\begin{aligned} \Big|\hat{Q}(\pi_\omega(A)\Omega_\omega, \pi_\omega(B)\Omega_\omega)\Big|^2 &= |\lambda\omega_1(A^*B)|^2 \\ &\leqslant \omega(A^*A)\omega(B^*B) \\ &= \|\pi_\omega(A)\Omega_\omega\|^2\|\pi_\omega(B)\Omega_\omega\|^2. \end{aligned} \tag{72.130}$$

Therefore $|\hat{Q}(\psi,\phi)|^2 \leqslant \|\psi\|\|\phi\|$ and $\hat{Q}$ is bounded.

The quadratic form $\hat{Q}$ can be continuously extended to the whole $\mathscr{H}_\omega$. From general property $\omega(A^*) = \overline{\omega(A)}$ (when ω is a state),

$$\hat{Q}(\phi,\psi) = \overline{\hat{Q}(\psi,\phi)}.$$

Lemma 72.112 immediately applies to our $\hat{Q}$, so we have an operator Q such that $\langle\psi, Q\phi\rangle = \hat{Q}(\psi,\phi)$. Therefore

$$\langle\pi_\omega(A)\Omega_\omega, Q\pi_\omega(B)\Omega_\omega\rangle = \lambda\omega_1(A^*B) \tag{72.131}$$

eq_1903r3

Since π is a representation, we have for all A, $B \in \mathfrak{A}$:

$$\begin{aligned} \langle\pi_\omega(A)\Omega_\omega, Q\pi_\omega(B)\Omega_\omega\rangle &= \hat{Q}\big(\pi_\omega(B^*A)\Omega_\omega, \Omega_\omega\big) \\ &= \langle\pi_\omega(B^*A)\Omega_\omega, Q\Omega_\omega\rangle \\ &= \langle\pi_\omega(A)\Omega_\omega, \pi_\omega(B)Q\Omega_\omega\rangle. \end{aligned} \tag{72.132}$$

This proves that $[Q, \pi_\omega(C)] = 0$ for each $C \in \mathfrak{A}$. Therefore $Q \in \pi_\omega(\mathfrak{A})'$ and there exists a $t \in \mathbb{R}$ such that $Q = t\mathbb{1}$. Using equation (72.131) and (72.97), we find

$$\begin{aligned} t\omega(A^*B) &= t\langle \pi_\omega(A)\Omega_\omega, \pi_\omega(B)\Omega_\omega\rangle \\ &= \lambda\omega_1(A^*B) \\ &= \hat{Q}\langle \pi_\omega(A)\Omega_\omega, \pi_\omega(B)\Omega_\omega\rangle \end{aligned} \tag{72.133}$$

So $t\omega(A^*B) = \lambda\omega_1(A^*B)$; thus ω and ω_1 are proportional and the decomposition

$$\omega = \lambda\omega_1 + (1-\lambda)\omega_2$$

is only possible for $\omega_1 = \omega_2 = \omega$. This proves that ω is a pure stare.

□

Corollary 72.114.
If the representation $\big(\pi(\mathfrak{A}), \mathscr{H}\big)$ is irreducible, then the GNS representation $\big(\pi_\omega(\mathfrak{A}), \mathscr{H}_\omega\big)$ build from any vector state (corresponding to $\Psi \in \mathscr{H}$ such that $\|\Psi\| = 1$) is unitary equivalent to $(\pi(\mathfrak{A}), \mathscr{H})$. cor_GNSirredst

Proof. Since π is irreducible, any vector in $\mathscr{H}$ is cyclic from proposition 72.109. So π is cyclic and proposition 72.89 concludes.

□

Corollary 72.115.
All irreducible representation of a C-algebra is (up to an equivalence) the GNS construction from a pure state.*

Proof. We know that an irreducible representation is unitary equivalent to a GNS representation, but the GNS representation will only be irreducible when ω is pure state. □

Proposition 72.116.
A state ω is pure if and only if for each positive functional ρ such that $0 \leqslant \rho \leqslant \omega$, there exists a $t \in \mathbb{R}^+$ such that $\rho = t\omega$. prop_pureiff

Proof. From corollary 72.3.1 and proposition 72.68 we can suppose that $\mathfrak{A}$ is unital because if not, the notion of positivity is defined from unitization. When $\rho = 0$ or $\rho = \omega$, the result is trivial. Now we suppose that $0 \neq \rho \neq \omega$.
Direct sense.

We know that ω is pure and $0 \leqslant \rho \leqslant \omega$, $0 < \rho(\mathbb{1}) < 1$ because $\omega - \rho$ is positive. Therefore

$$\|\omega - \rho\| = \omega(\mathbb{1}) - \rho(\mathbb{1}) = 1 - \rho(\mathbb{1}),$$

thus $\rho(\mathbb{1}) = 1$ should implies $\|\omega - \rho\| = 0$ and then $\omega = \rho$. The possibility $\rho(\mathbb{1}) = 1$ is also not possible. Thus $\rho(\mathbb{1})$ is between 0 and 1.

This allows us to consider the states

$$\frac{\omega - \rho}{1 - \rho(\mathbb{1})}, \quad \text{and} \quad \frac{\rho}{\rho(\mathbb{1})}.$$

For $\lambda = 1 - \rho(\mathbb{1})$,

$$\omega = \lambda\frac{\omega - \rho}{1 - \rho(\mathbb{1})} - (1-\lambda)\frac{\rho}{\rho(\mathbb{1})}.$$

Since ω is pure, it implies $\frac{\omega-\rho}{1-\rho(\mathbb{1})} = \frac{\rho}{\rho(\mathbb{1})}$. So, $\rho = \rho(\mathbb{1})\omega$.
Inverse sense.

Let us consider a decomposition $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ of ω. In the proof of theorem 72.113, we find that $0 \leqslant \lambda\omega_1 < \omega$. Then the assumption says that $\lambda\omega_1 = \omega = \omega_2$ and the normalization makes automatically $\omega_1 = \omega = \omega_2$ which proves that ω is pure.

□

Lemma 72.117.
Let $\mathfrak{A}$ be a C-algebra and ρ, a positive form on $\mathfrak{A}$. If we pose $M = \ker(\rho)$, $N = \{A \in \mathfrak{A}$ tel que $\rho(A^*A) = 0\}$, then*

$$N + N^* \subseteq M$$

and if ρ is pure, $N + N^* = M$.

Proof. The functional ρ being positive, equation (72.61) holds. With $A = \mathbb{1}$ we find

$$|\rho(B)^2| \leqslant \rho(\mathbb{1})\rho(B^*B) \tag{72.134}$$

eq_nos_Bdtho

and thus $\rho(A^*A) = 0$ implies $\rho(A) = 0$.

No proof for the second part. □

Theorem 72.118.
The space of pure states of the (commutative) C-algebra $C_0(X)$ (the space of functions which are decreasing to zero at infinity) endowed with the relative w*-topology is homeomorphic to X.*

¡¡ Avertissement/question au lecteur !! 72.119
The following proof is buggy and very unsure.

Proof. The C*-algebra $\mathfrak{A} = C_0(X)$ is not specially unital. If it is not, Gelfand theorem 72.15 says that there exists a locally compact and Hausdorff space Y such that $\mathfrak{A}$ is isomorphic to $C_0(Y)$. If $\mathfrak{A} = C_0(X)$ with a non compact X, we begin to prove that the unitization is $\mathfrak{A}_\mathbb{1} = C(\tilde{X})$ where $\tilde{X}$ is the one point compactification of X. From proposition 72.17, we have an unique unital C*-algebra $\mathfrak{A}_\mathbb{1}$ with an isometric morphism $\mathfrak{A} \to \mathfrak{A}_\mathbb{1}$ such that $\mathfrak{A}_\mathbb{1}/\mathfrak{A} \simeq \mathbb{C}$.

Therefore, we have to prove that $C(\tilde{X})$ is a suitable $\mathfrak{A}_\mathbb{1}$ and so it will be the unique unitization of $C_0(X)$. The identity map $\mathrm{Id}\colon C_0(X) \to C(\tilde{X})$ works. We have to check that $C(\tilde{X})/C_0(X) \simeq \mathbb{C}$. By identifying all functions of $C(\tilde{X})$ which only differ by a function of $C_0(X)$, there are in fact only one function for each complex number z: the which is constant (or another which is z at ∞).

We have proved that if $\mathfrak{A} = C_0(X)$, then $\mathfrak{A}_\mathbb{1} = C(\tilde{X})$. From proposition 72.68 pure states on $C_0(X)$ uniquely extend to a pure state on $C(\tilde{X})$. So if X is non compact, we do not loss anything by considering $C(\tilde{X})$ instead of $C_0(X)$. We still have to prove that taking $C(\tilde{X})$ instead of $C(X)$ does not *gain* anything: we must have $\mathcal{S}\big(C_0(X)\big) = \mathcal{S}\big(C(\tilde{X})\big)$. In the case of unital C*-algebras, pure states are linear functionals. The way to extend linear functionals from $\mathfrak{A}$ to $\mathfrak{A}_\mathbb{1}$ is the same as the one to extend states.

If X is compact, then $C_0(X) = C(X)$. Hence we are in both case (X compact or not) reduced to prove the theorem for $C(X)$ with compact X.

Following proposition 72.14, $\Delta(C(X))$ is homeomorphic to X because the latter is compact and Hausdorff. We have now to prove that there exists an homeomorphism between $\Delta(C(X))$ and set of pure states on $C(X)$. We are going to prove a bijection. Let on the one hand $\omega_x \in \Delta(C(X))$ be defined by

$$\begin{aligned} \omega_x &: C(X) \to \mathbb{C} \\ &\omega_x(f) = f(x), \end{aligned} \tag{72.135}$$

and on the other hand a functional ρ such that $0 \leqslant \rho \leqslant \omega_x$. We have $\ker(\omega_x) \subset \ker(\rho)$. When f is positive, $0 \leqslant \rho(f) \leqslant \omega_x(f)$, so for any function,

$$o \leqslant \rho(f^*f) \leqslant \omega_x(f^*f),$$

but ω_x is multiplicative, therefore $\rho(f^*f) \leqslant \omega_x(f^*)\omega_x(f)$. If $f \in \ker(\omega_x)$, we have $\rho(f^*f) = 0$. Lemma concludes $\rho(f) = 0$. So if $f \in \ker\omega_x$, we have $f(x) = 0$. From theorem 72.3, $\ker\omega_x$ is a maximal ideal while $\ker(\rho)$ is an ideal. So $\ker(\omega_x)$ is a maximal ideal contained in an ideal, therefore if $\rho \neq 0$, $\ker(\omega_x) = \ker(\rho)$ which implies that $\rho = \lambda\omega_x$. Proposition 72.116 concludes that ω_x is pure.

¡¡ Avertissement/question à la lectrice !! 72.120
Why is $\ker(\rho)$ an ideal?

Now we suppose that ω is pure and we take $g \in C(X)$ such that $0 \leqslant g \leqslant 1_X$ on $C(X)$, we define

$$\begin{aligned} \omega_g \ : \ C(X) &\to \mathbb{C} \\ f &\mapsto \omega(fg) \end{aligned} \tag{72.136}$$

So $\omega(f) - \omega_g(f) = \omega\big(f(1-g)\big)$ and if f is positive, $\omega(f) - \omega_g(f) \geqslant O$. So for a certain $t \in \mathbb{R}^+$, we have

$$\omega_g = t\omega$$

because $0 \leqslant \omega_g \leqslant \omega$ and proposition 72.116. In particular, $\ker(\omega_g) = \ker(\omega)$, hence when $f \in \ker(\omega)$, for any $g \in C(X)$, $fg \in \ker(\omega)$ because any function in $C(X)$ is a linear combination of functions g with $0 \leqslant g \leqslant 1_X$. This proves that $\ker(\omega)$ is an ideal. On the other hand, $\ker(\omega)$ is a maximal ideal because the kernel of any functional on a vector space has codimension 1, see page 3738. Theorem 72.3 shows that ω is multiplicative. So $\omega \in \Delta\big(C(X)\big)$.

□

72.15.1 Existence of pure states

It is possible for $\mathcal{S}$ to do not contain pure states. It is the case when $\mathcal{S}$ is a convex cone. Such a C^*-algebra has no irreducible representations. We are going to prove that this case is not possible. We define the convex hull of the part A of a vector space by

$$co(A) = \{\lambda + (1-\lambda)w \text{ with } w \in A, \lambda \in [0,1]\}.$$

Theorem 72.121.
A compact connected set K in a locally convex vector space is the closure of the convex hull of its extreme points. In other words:

$$K = \overline{co(\partial_e K)}.$$

Proof. No proof □

Lemma 72.122.
Let $\mathfrak{A}$ be an unital C^*-algebra and $\mathcal{B}$ an self-adjoint vector subspace of $\mathfrak{A}$ with $\mathbb{1} \in \mathcal{B}$. Let F be the set of linear forms g on B such that

- *$g(A^*) = \overline{g(A)}$ for all $A \in \mathcal{B}$,*
- *$g(A) \geqslant 0$ for all $A \in \mathcal{B} \cap \mathfrak{A}^+$,*
- *$g(\mathbb{1}) = 1$.*

Any element of F can be extended into a state on $\mathfrak{A}$. lem_DixcBprol

Theorem 72.123.
For all $A \in \mathfrak{A}$ and a $a \in \sigma(A)$, we have a pure state ω_a on $\mathfrak{A}$ such that $\omega_a(A) = a$. There also exists a pure state ω such that $|\omega(A)| = \|A\|$. tho_existsetat

Proof. Let us take an intermediary result in proof of lemma 72.68:

$$\begin{aligned} \tilde{\omega}, : \ \mathbb{C}A + \mathbb{C}\mathbb{1} &\to \mathbb{C} \\ (\lambda A + \mu\mathbb{1}) &\mapsto \lambda a + \mu \end{aligned} \tag{72.137}$$

We extend this state by continuity and multiplicatively to $C^*(A, \mathbb{1})$ with formulas as $\tilde{\omega}_a(A^n) = a^n$. We have to check that this extension is pure: $\tilde{\omega}_a$ is positive and belongs to $\Delta(C^*(A, \mathbb{1}))$.

Positivity comes from assumption that $A \in \mathfrak{A}_{\mathbb{R}}$. Indeed, $a \in \sigma(A) \subset \mathbb{R}^+$. So positives elements in $C^*(A, \mathbb{1})$ are even power (and completion) of A. So the images are even powers of $a \in \mathbb{R}$ and are therefore positives. The fact that $\tilde{\omega}_a$ is multiplicative on $C^*(A, \mathbb{1})$ comes from the fact that it is the same, in expressions as $\tilde{\omega}_a(xy)$, to distribute inside the $\tilde{\omega}_a$ and push out terms a^n by linearity, or write $\tilde{\omega}_a(x)\tilde{\omega}_a(y)$ and distribute outside. So $\tilde{\omega}$ is a pure state in $C^*(A, \mathbb{1})$.

We consider the set K_a of extensions of $\tilde{\omega}_a$ which are states on $\mathfrak{A}$. The fact that K_a is non empty comes from the lemma 72.122.

Let us now prove that K_a is convex. For, we take ω_1 and ω_2, two extensions of $\tilde{\omega}_a$ and we will prove that $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ is an extension too. By linearity, ω is a state. It is an extension of $\tilde{\omega}_a$ because

$$\begin{aligned}\omega(\mathbb{1} + A^2) &= \lambda\omega_1(\mathbb{1} + A^2) + (A-\lambda)\omega_1(\mathbb{1} + A^2)\\ &= \lambda(1+a^2) + (1-\lambda)(1+a^2)\\ &= 1 + a^2.\end{aligned}$$

In order to prove that K_a is closed, we prove that its complement is open. Let $\omega \notin K_a$. We will prove that there exists ε such that for all η with $\|\omega - \eta\| \leqslant \varepsilon$. Let λ be such that $\|\lambda A\| = 1$ in such a manner that $\tilde{\omega}_a(\lambda A) = \lambda A$. Let $\omega(A) = s$ ($\omega \notin K_a$). We have

$$\begin{aligned}\|\omega - \eta\| &= \sup\{|\omega - \eta| \text{ st } \|B\| = 1\}\\ &\geqslant |(\omega - \eta)(\lambda A)|\\ &= |\omega(\lambda A) - \eta(\lambda A)|,\end{aligned}$$

but

$$|\omega(\lambda A) - \eta(\lambda A)| \leqslant \|\omega - \eta\| \leqslant \varepsilon.$$

So $|s - \eta(\lambda A)| \leqslant \varepsilon$, and when ε is small (for example when $\varepsilon < |s - \lambda a|$), $\eta(\lambda A)$ is close to s which is different of λA. This proves that K_a is closed.

We know that K_a is convex and closed. So it has at least one extreme point. Let ω_a be one of them. We are going to prove that it is also extreme in $\mathcal{S}$. If not it can be decomposed as $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$. Taking the latter at A shows that, on $C^*(A, \mathbb{1})$, the functionals ω, ω_1 and ω_2 are equals. So ω_a cannot be an extreme point in K_a. This concludes the first part of the proof.

Theorem 60.4.3 says us that in a Banach algebra, $\sigma(A)$ is compact for all A. So there exists a $a \in \sigma(A)$ such that $r(A) = |a|$. With this a,

$$|\omega_a(A)| = |A| = r(A) = \|A\|.$$

because $A = A^*$.

□

The Gelfand Neumark theorem was proved using lemma 72.69. Now we have at hand a refining of this lemma. Hence we can use

$$\pi_r = \bigoplus_{\omega \in \mathcal{P}(\mathfrak{A})} \pi_\omega$$

instead of the universal representation. The justification of this claim is that when A is such that $\omega(A^*A) = 0$ for all pure states, $\|A^*A\| = 0$. Indeed A^*A is positive, so there exists a pure state ω_a such that $\omega_a(A^*A) = \|A^*A\|$. Then $\|A^*A\| = 0$.

Now we say that two states are **equivalent** if their GNS representations are equivalent. Gelfand Neumark theorem says that

$$\mathfrak{A} \simeq \pi_r(\mathfrak{A}) := \bigoplus_{\omega \in [\mathcal{P}(\mathfrak{A})]} \pi_\omega(\mathfrak{A}).$$

where $[\mathcal{P}(\mathfrak{A})]$ stands for the set of equivalences classes in $\mathcal{P}(\mathfrak{A})$.

Proposition 72.124.
Any finite dimensional C-algebra is isomorphic to a direct sum of matricial algebras.*

Proof. No proof. □

72.16 Generalization of matrix C^*-algebra

If we want to generalize the C^*-algebra $\mathfrak{M}^n(\mathbb{C})$ to infinite dimensional Hilbert spaces, we first try to use the C^*-algebra of bounded operators. It does not work because such a C^*-algebra possesses many non equivalent representations on non separable Hilbert spaces.

72.16.1 Example

Let us consider the C^*-algebra $\mathfrak{B}(\mathscr{H})$ and its definition representation which is obviously irreducible. From GNS construction and corollary 72.114, we know that all the GNS representations build from a vector state are equivalent to the definition one.

On the other hand, there exists some self-adjoint bounded operators with non empty continuous spectrum. The take $A \in \mathfrak{B}(\mathscr{H})$ and $a \in \sigma(A)$ such that there are no $\psi_a \in \mathscr{H}$ such that $A\psi_a = a\psi_a$. Theorem 72.123 gives a pure state ω_a such that $\omega_a(A) = a$, and hence a GNS representation π_a on $\mathscr{H}$. This representation is irreducible because ω_a is pure. In π_a, we have a cyclic vector Ω_a such that

$$\langle \Omega_a, \pi_a(A)\Omega_a \rangle = \omega_a(A) = a.$$

72.17 Tensor product

SecTensProdCSA

If $\mathfrak{A}$ and $\mathcal{B}$ are C^*-algebra, the **tensor product** is the completion of the space generated by the finite sums of the form $\sum_{i=1}^n A_i \otimes B_i$ with $A_i \in \mathfrak{A}$ and $B_i \in \mathcal{B}$.

As an example of the importance of the completion, consider a compact group G and $\mathfrak{A} = C(G)$ the C^*-algebra of continuous functions on G. We can build the map $\Delta\colon \mathfrak{A} \to \mathfrak{A} \otimes \mathfrak{A}$ by

$$\Delta(f)(x, y) = f(xy) \tag{72.138}$$

for every x and y in G and $f \in C(G)$. The well-definiteness of Δ is due to the fact that $C(G) \otimes C(G) \simeq C(G \times G)$ by completion. This trick is used whenever we define a coproduct on a space of functions on a group, see for example definition 72.126 and section 81.3 around equation (81.53).

If A and B are manifolds, we have

$$C^\infty(A) \otimes C^\infty(B) \simeq C^\infty(A \times B) \tag{72.139}$$

by the map

$$\begin{aligned} \varphi\colon C^\infty(A) \otimes C^\infty(B) &\to C^\infty(A \times B) \\ \sum_i a_i \otimes b_i &\mapsto \Big[(x, y) \mapsto \sum_i a_i(x) b_i(y)\Big]. \end{aligned} \tag{72.140}$$

EqIsoCABCACBCstar

The image by φ of the *algebraic* tensor product $C^\infty(A) \otimes C^\infty(B)$ is dense in $C^\infty(A \times B)$ as for example the polynomials are contained in the image. Indeed let $f(x, y) = \sum_{ij} f_{ij} x^i y^j$. The function f is the image by φ of

$$\sum_{kl} f_{kl} a_k \otimes b_l \tag{72.141}$$

EqDecompffklCABCACB

where $a_k(x) = x^k$ and $b_k(y) = y^k$. Thus, if we consider the C^*-algebraic tensor product, we have the equality.

See also [988] for the sequel about tensor products. Let $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be C^*-algebra and $\mathfrak{A}_1 \odot \mathfrak{A}_2$ be their algebraic tensor product. There are at least two ways to define a C^*-norm on the $*$-algebra $\mathfrak{A}_1 \odot \mathfrak{A}_2$.

(1) The **maximal norm** of $A \in \mathfrak{A}_1 \odot$ is defined by

$$\|A\|_{max} = \sup_\pi \|\pi(A)\| \tag{72.142}$$

where the supremum is taken over all the representations [12] of $\mathfrak{A}_1 \odot \mathfrak{A}_2$ over some Hilbert space $\mathscr{H}$. The **maximal tensor product** is the completion of $\mathfrak{A}_1 \odot \mathfrak{A}_2$ for that norm.

(2) The **minimal norm** is obtained by taking the supremum only over the representations of the for $\pi_1 \otimes \pi_2$:

$$\|x\|_{min} = \sup_{\pi_1, \pi_2} \|(\pi_1 \otimes \pi_2)(x)\| \tag{72.143}$$

where π_i is a representation of $\mathfrak{A}_i$ on an Hilbert space $\mathscr{H}_i$.

12. i.e. all the homomorphisms $\pi\colon \mathfrak{A}_1 \odot \mathfrak{A}_2 \to \mathfrak{B}(\mathscr{H})$.

Lemma 72.125.
If $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are sub-C-algebra of $\mathfrak{B}(\mathscr{H}_1)$ and $\mathfrak{B}(\mathscr{H}_2)$, then $\mathfrak{A}_1 \otimes_{min} \mathfrak{A}_2$ is the closure of $\mathfrak{A}_1 \odot \mathfrak{A}_2$ seen as subalgebra of $\mathfrak{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$.*

72.17.1 Examples: Hopf algebra of functions

DefHopfsurCG

Definition 72.126.
Let $C(G)$ be the set of continuous functions on a topological group G. If we denote by 1 the function $1\colon G \to \mathbb{C}$, $1(x) = 1$ for all $x \in G$, the following produces a structure of Hopf algebra[13] *on $C(G)$:*

(1) $(f \cdot g)(x) = f(x)g(x)$,
(2) $\eta(\lambda) = \lambda 1$, *the unit,*
ItemHopfCGiii
(3) $(\Delta f)(x \otimes y) = f(xy)$,
ItemHopfCGiv
(4) $\epsilon(f) = f(e)$,
(5) $S(f)(x) = f(x^{-1})$.

The item (3) deserves some comments. If $f \in C(G)$, the element $\Delta(f) \in C(G) \otimes C(G)$ is given by $\Delta(f) = f_{(1)} \otimes f_{(2)}$ with the requirement that

$$f_{(1)}(x)f_{(2)}(y) = f(xy). \tag{72.144}$$

Such choice is possible since by density of the functions of the form $f(x)g(y)$ in $C(G \times G)$, see section 72.17.

Let us check that $(\mathrm{Id} \otimes \epsilon)\Delta = \mathrm{Id}$. First $(\mathrm{Id} \otimes \epsilon)\big(f_{(1)} \otimes f_{(2)}\big) = f_{(1)} \otimes f_{(2)}(e)$. Now we use the identification between $C(G) \otimes \mathbb{C}$ and $C(G)$ $((f \otimes z)(g) = zf(g))$ in order to get

$$\Big((\mathrm{Id} \otimes \epsilon)\big(f_{(1)} \otimes f_{(2)}\big)\Big)(g) = \big(f_{(1)} \otimes f_{(2)}(e)\big)(g) = f_{(1)}(g)f_{(2)}(e) = f(ge) = f(g). \tag{72.145}$$

¡¡ Avertissement/question au lecteur !! 72.127
Unicity of $f_{(1)}$ and $f_{(2)}$ seems doubtful to me.

72.18 Traces over C^*-algebra

SecTraceCstar

Source: [986]. For tracial functionals on von Neumann algebras, see section 74.11.

Definition 72.128.
Let $\mathfrak{A}$ be a C-algebra. A **trace** on $\mathfrak{A}^+$ is a function $\tau\colon A^+ \to [0, \infty]$ such that*

(1) if A and B are in $\mathfrak{A}^+$, $\tau(A + B) = \tau(A) + \tau(B)$,
(2) If $A \in \mathfrak{A}^+$ and $\lambda \in \mathbb{R}^+$, then $\tau(\lambda A) = \lambda\tau(A)$. Here if $\lambda = 0$, we pose $0 \times \infty = 0$;
(3) If $Z \in \mathfrak{A}$, we have $\tau(ZZ^*) = \tau(Z^*Z)$.

We say that τ is semifinite if for every $A \in \mathfrak{A}^+$,

$$\tau(A) = \sup\{\tau(B) \text{ tel que } B \in \mathfrak{A}^+, B \leqslant A, \tau(B) < \infty\}. \tag{72.146}$$

We recall that for every $Z \in \mathfrak{A}$, we have $ZZ^* \in \mathfrak{A}^+$.
The following lemma is the lemma 3 in [986].

LemTraceAplusextmlmn

Lemma 72.129.
Let $\mathfrak{A}$ be a C-algebra and τ a trace on $\mathfrak{A}^+$. Then the following hold.*

(1) The set

$$\mathcal{L} = \{A \in \mathfrak{A} \text{ tel que } \tau(AA^*) < \infty\} \tag{72.147}$$

is a bilateral ideal in $\mathfrak{A}$.

13. Definition 54.27.

(2) The set $\mathcal{N} = \langle \mathcal{L}^2 \rangle$ is the set of complex linear combinations of $\mathcal{N}^+$: $\mathcal{N} = \langle \mathcal{N}^+ \rangle$.

(3) We have

$$\mathcal{N}^+ = \{A \in \mathfrak{A}^+ \text{ tel que } \tau(A) < \infty\}. \tag{72.148}$$

(4) There exists one and only one linear form f on $\mathcal{N}$ which coincides with τ on $\mathcal{N}^+$.

(5) The linear form f satisfies

(5a) $f(A^*) = \overline{f(A)}$;

(5b) $f(AB) = f(BA)$ for every u and v in $\mathcal{L}$;

(5c) $f(ZA) = f(AZ)$ for every $A \in \mathcal{N}$ and $Z \in \mathfrak{A}$.

Proof. Let us begin by pointing out the fact that if $A \leqslant B$, then $\tau(A) \leqslant \tau(B)$ because of linearity: $\tau(B) = \tau(A) + \tau(B - A) \geqslant \tau(A)$ since $B - A \in \mathfrak{A}^+$.

(1) If $A \in \mathcal{L}$, then $A^* \in \mathcal{L}$ because $\tau(A^*A) = \tau(AA^*)$ (by definition of a trace). If $A, B \in \mathcal{L}$, we have

$$(A + B)(A + B)^* \leqslant 2(AA^* + BB^*), \tag{72.149}$$

so that

$$\tau\big((A + B)(A + B)^*\big) < 2\tau(AA^*) + 2\tau(BB^*) < \infty. \tag{72.150}$$

This proves that $A + B \in \mathcal{L}$.

Let now $A \in \mathcal{L}$ and $Z \in \mathfrak{A}$. Since $AZZ^*X^* \leqslant \|ZZ^*\|AA^*$, we have

$$\tau\big(AZ(AZ)^*\big) < \infty. \tag{72.151}$$

So $\mathcal{L}$ is a right ideal in $\mathfrak{A}$. This is also a left ideal because $ZA = (A^*Z^*)^*$, but the fact that $A^* \in \mathcal{L}$ implies $A^*Z^* \in \mathcal{L}$, so that $ZA \in \mathcal{L}$.

(2) An element X in $\mathcal{N}$ reads

$$X = \sum_{j=1}^{n} A_j B_j^* \tag{72.152}$$

with $A_j, B_j \in \mathcal{L}$. The **polarization** relation reads

$$\begin{aligned} 4X = &\sum_j (A_j + B_j)(A_j + B_j)^* \\ &+ \sum_j (A_j - B_j)(A_j - B_j)^* \\ &+ \sum_j (A_j + iB_j)(A_j + iB_j)^* \\ &+ \sum_j (A_j - iB_j)(A_j - iB_j)^*. \end{aligned} \tag{72.153}$$

EqPolaXAB

So X is a linear combination of elements in $\mathcal{N}^+$ (that is elements of the form AA^* with $A \in \mathcal{L}$). Notice that $A \in \mathcal{L}$ implies $iA \in \mathcal{L}$ because

$$\tau\big(iA(iA)^*\big) = -\tau(iAiA^*) = \tau(AA^*) < \infty. \tag{72.154}$$

This proves that $\mathcal{N} \subset \langle \mathcal{N}^+ \rangle$. The fact that $\langle \mathcal{N}^+ \rangle$ is a subset of $\mathcal{N}$ is by construction.

(3) An element X in $\langle \mathcal{N}^+ \rangle$ reads $X = \sum_j A_j B_j^*$ where, for each j, we have $A_j B_j^* \in \mathcal{N}^+$, in particular $A_j B_j^* = (A_j B_j^*)^* = B_j A_j^*$. We can still write down the polarization identity, but now the last two terms of (72.153) give $2i(B_j A_j^* - A_j B_j^*) = 0$.

Thus we have

$$\begin{aligned} 4X &= \sum_j (A_j + B_j)(A_j + B_j)^* - \sum_j (A_j - B_j)(A_j - B_j)^* \\ &\leqslant \sum_j (A_j + B_j)(A_j + B_j)^*, \end{aligned} \tag{72.155}$$

but we already know that $\tau\big(\sum_j (A_j+B_j)(A_j+B_j)^*\big) < \infty$. Thus we have $\tau(X) < \infty$. So we proved that

$$\mathcal{N}^+ \subset \{X \in \mathfrak{A}^+ \text{ tel que } \tau(X) < \infty\}. \tag{72.156}$$

Let now $X \in \mathfrak{A}^+$ be such that $\tau(X) < \infty$. In order to prove that $X \in \mathcal{N}^+$, it is sufficient to prove that $X \in \mathcal{N}$. Since $X = X^*$, we can use the continuous functional calculus (see theorem 72.23 and remark 72.25) in order to define $X^{1/2}$. We have

$$\tau\big(X^{1/2}(X^{1/2})^*\big) = \tau\big(X^{1/2}X^{1/2}\big) < \infty, \tag{72.157}$$

so that $X^{1/2} \in \mathcal{L}$.

(4) Since $\mathcal{N}$ is generated by $\mathcal{N}^+$, the functional τ on $\mathcal{N}^+$ there is one and only one extension of τ to a linear functional f on $\mathcal{N}$.

(5) (5a) An element of $\mathcal{N}$ is a linear combination of elements of $\mathcal{N}^+$: $X = \sum_j \lambda_i A_i$ with $A_i \in \mathcal{N}^+$. Thus using the linearity of f and the properties of τ, we have

$$\begin{aligned} f(X^*) &= f\big(\sum_j \overline{\lambda_j} A_j^*\big) \\ &= \sum_j \overline{\lambda_j} f(A_j) \\ &= \sum_j \overline{\lambda_j}\, \underbrace{\tau(A_j)}_{\in\mathbb{R}^+} \\ &= \sum_j \overline{\lambda_j \tau(A_j)} \\ &= \overline{f(X)}. \end{aligned} \tag{72.158}$$

This is the first property we had to check.

(5b) If $A \in \mathcal{L}$, we have $AA^* \in \mathcal{N}$ and by definition of a trace,

$$f(AA^*) = \tau(AA^*) = \tau(A^*A) = f(A^*A). \tag{72.159}$$

If A and B belong to $\mathcal{L}$, we use the polarization identity:

$$4AB^* = (A+B)(A+B)^* - (A-B)(A-B)^* + i(A+iB)(A+iB)^* - i(A-iB)(A-iB)^*, \tag{72.160}$$

and we do the same computation in order to get $f(AB^*) = f(B^*A)$ whenever A and B belong to $\mathcal{L}$.

(5c) Let $Z \in \mathfrak{A}$. An element in $\mathcal{N}$ reads $X = \sum_j A_j B_j$ with A_j and B_j in $\mathcal{L}$. Since $\mathcal{L}$ is an bilateral ideal in $\mathfrak{A}$ we have $ZA_j \in \mathcal{L}$ and $B_j Z \in \mathcal{L}$, thus we can make the following computation:

$$\begin{aligned} f\big(Z\sum_j A_j B_j\big) &= \sum_j f\big((ZA_j)B_j\big) \\ &= \sum_j f\big(B_j(ZA_j)\big) \\ &= \sum_j f\big((B_jZ)A_j\big) \\ &= \sum_j f\big(A_j(B_jZ)\big) \\ &= f\big((\sum_j A_jB_j)Z\big) \\ &= f(XZ). \end{aligned} \tag{72.161}$$

This concludes the proof of the lemma.

□

Chapter 73

Compact quantum groups

73.1 Definitions

Literature: [989, 990, 991, 992] and Wikipedia: quantum group and compact quantum group. For introductions about Hopf algebra and some related topics such as quantum groups and deformation quantization of (co)-Poisson structures, see [890, 993].

Definition 73.1.
*A **compact quantum group** is a pair $G = (A, \Phi)$ where A is an unital separable C^*-algebra and $\Phi\colon A \to A \otimes A$ is a $*$-homomorphism such that*

(1) $(\mathrm{Id} \otimes \mathrm{Id})\Phi = (\mathrm{Id} \otimes \mathbb{1})\Phi$;

(2) The sets

$$\begin{array}{c} \{(b \otimes \mathbb{1})\Phi(c) \text{ tel que } b, c \in A\}, \\ \{(\mathbb{1} \otimes b)\Phi(c) \text{ tel que } b, c \in A\} \end{array} \tag{73.1}$$

are linearly dense in $A \otimes A$.

73.1.1 Example: representation of groups

If u is a finite-dimensional representation of a group G, we can build the $*$-subalgebra of $C(G)$ generated by the elements u_{ij} and $\alpha(u_{ij})$ where α is defined by

$$\alpha(u_{ij})(g) = u_{ij}(g^{-1}). \tag{73.2}$$

Using the definition 72.126, we turn this algebra into a Hopf algebra [1]. Let us particularize the definitions. For the coproduct we have

$$\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj} \tag{73.3}$$

because if g and h are elements of G, the coproduct on $C(G)$ is equal to

$$\Delta(u_{ij})(g,h) = u_{ij}(gh) = \big(u(gh)\big)_{ij} = \big(u(g)u(h)\big)_{ij} = \sum_k u(g)_{ik}u(h)_{kj} = \Big(\sum_k u_{ik} \otimes u_{kj}\Big)(g,h) \tag{73.4}$$

where we used the fact that u is a representation and the fact that, by definition, $u_{ij}(g) = u(g)_{ij}$. The counit is given by

$$\epsilon(u_{ij}) = \delta_{ij}. \tag{73.5}$$

since the neutral element e is always represented by the unit matrix. The antipode is

$$\alpha(u_{ij})(g) = u_{ij}(g^{-1}). \tag{73.6}$$

1. This is not trivial since nothing guarantee *a priori* that the right hand sides of the definitions are elements of our small algebra.

The unit is given by the constant function 1. Fortunately, it turns out that this function is a combination of the functions u_{ij}:

$$1 = \sum_k u_{1k}\alpha(u_{k1}) = \sum_k \alpha(u_{1k})u_{k1}. \tag{73.7}$$

Indeed,

$$\begin{aligned}\sum_k u_{1k}(g)\alpha(u_{k1})(g) &= \sum_k u_{1k}(g)u_{k1}(g^{-1})\\ &= \Big(u(g)u(g^{-1})\Big)_{11}\\ &= u(e)_{11}\\ &= 1.\end{aligned} \tag{73.8}$$

Let us check one of the diagrams (54.44):

$$\begin{array}{ccc}\mathbb{C}\otimes C(G) & \xleftarrow{\epsilon\otimes\mathrm{Id}} & C(G)\otimes C(G)\\ & \nwarrow^{\psi} & \uparrow{\scriptstyle \Delta}\\ & & C(G)\end{array} \tag{73.9}$$

We have

$$\begin{aligned}(\epsilon\otimes\mathrm{Id})\Delta(u_{ij}) &= (\epsilon\otimes\mathrm{Id})\sum_k u_{ik}\otimes u_{kj}\\ &= \sum_k \delta_{ik}\otimes u_{kj}\\ &= \sum_k 1\otimes\delta_{ik}u_{kj}\\ &= 1\otimes u_{ij}\\ &= \psi(u_{ij}).\end{aligned} \tag{73.10}$$

73.1.2 Matrix quantum group

DefQuantumMatrixGroup

Definition 73.2.
*A **matrix quantum group** is a pair (C, u) where C is a C^*-algebra and $u = (u_{ij})$ is a matrix with entries in C such that*

(1) The $*$-subalgebra C_0 generated by the elements u_{ij} is dense in C.

(2) There exists a C^*-algebra homomorphism $\Delta\colon C \to C\otimes C$ such that

$$\Delta(u_{ij}) = \sum_k u_{ik}\otimes u_{kj}. \tag{73.11}$$

DefQuantumMatrixGroupItemiii

(3) There exists an antimultiplicative map $\alpha\colon C_0 \to C_0$ called coinverse *such that*

(3a) $\alpha\big(\alpha(v^*)^*\big) = v$ for any $v \in C_0$,

(3b) $\sum_k \alpha(u_{ik})u_{kj} = \sum_k u_{ik}\alpha(u_{kj}) = \delta_{ij}\mathbb{1}$ where $\mathbb{1}$ is the unit in C.

73.1.2.1 Example: $\mathrm{SU}_q(2)$

As an example, we give the quantum matrix group $\mathrm{SU}_q(2)$. Let q be a positive real number. As C^*-algebra, $\mathrm{SU}_q(2)$ is generated by the elements a and b and the relations [2]

$$\begin{aligned}ba &= qab\\ b^*a &= qab^*\\ bb^* &= b^*b\\ a^*a + q^2b^*b &= 1\\ aa^* + bb^* &= 1\end{aligned} \tag{73.12}$$

EqDefAlfSUab

2. We follow the notations of [994].

We build a compact matrix quantum group $\Big(C\big(\mathrm{SU}_q(2)\big), u\Big)$ where

$$u = \begin{pmatrix} a & qb \\ -b^* & a^* \end{pmatrix}. \tag{73.13}$$

The Hopf-$*$-algebra structure on $C\big(\mathrm{SU}_q(2)\big)$ is given by the relations in the definition 73.2. We have to check that these are compatible.

We have

$$\Delta(a) = \Delta(u_{11}) = \sum_k u_{1k} \otimes u_{k1} = u_{11} \otimes u_{11} + u_{12} \otimes u_{21} = a \otimes a - qb \otimes b^*, \tag{73.14}$$

and

$$\Delta(a^*) = \Delta(u_{22}) = \sum_k u_{2k} \otimes u_{k2} = u_{21} \otimes u_{12} + u_{22} \otimes u_{22} = -qb^* \otimes b + a^* \otimes a^*, \tag{73.15}$$

which is compatible with (73.14). Doing the same with $b = \frac{1}{q}u_{12}$ (and checking the compatibility with b^*), we find the coproduct

$$\begin{aligned} \Delta(a) &= a \otimes a - qb \otimes b^* \\ \Delta(b) &= b \otimes a^* + a \otimes b. \end{aligned} \tag{73.16}$$

The counit $\epsilon(u_{ij}) = \delta_{ij}$ provides

$$\begin{aligned} \epsilon(a) &= 1 \\ \epsilon(b) &= 0. \end{aligned} \tag{73.17}$$

Using the relations (73.12), we can check that the condition (3) of definition 73.2 are satisfied by the antipode

$$\begin{aligned} \alpha(a) &= a^*, \quad & \alpha(a^*) &= a, \\ \alpha(b) &= -qb, \quad & \alpha(b^*) &= -q^{-1}b^*. \end{aligned} \tag{73.18}$$

As an example,

$$\alpha\big(\alpha(b^*)^*\big) = \alpha\big((-\frac{1}{q}b^*)^*\big) = -\frac{1}{q}\alpha(b) = b \tag{73.19}$$

and $\sum_k \alpha(u_{ik})u_{k2} = \delta_{12} = 0$ since

$$\alpha(u_{11})u_{12} + \alpha(u_{12})u_{22} = \alpha(a)qb + q\alpha(b)a^* = qa^*b - q^2ba^* = 0 \tag{73.20}$$

because $a^*b = (b^*a)^* = (qab^*)^* = qba^*$.

All this structure provides a Hopf $*$-algebra named $C\big(\mathrm{SU}_q(2)\big)$.

73.1.2.2 Corepresentation

A **corepresentation** of the quantum group (A, Δ) is an element $u \in \mathbb{M}_n(A)$ such that

$$\Delta(u_{kl}) = \sum_{j=1}^{n} u_{kj} \otimes u_{jl} \tag{73.21}$$

for every $k, l \in \{1, \dots n\}$.

73.1.3 Haar measure on compact quantum groups

Let A be the C^*-algebra $C(G)$ where G is a compact group. The **Haar measure** μ on G is the unique regular Borel measure on G such that $\mu(G) = 1$ and

$$\int_G f(st)d\mu(s) = \int_G f(ts)d\mu(s) = \int_G f(s)d\mu(s). \tag{73.22}$$

We can see the Haar measure μ as a positive functional $\varphi\colon A \to \mathbb{C}$,

$$\varphi(f) = \int_G f(s)d\mu(s). \tag{73.23}$$

If $h \in A$ is such that $\Delta(h) = h_{(1)} \otimes h_{(2)}$ (sum implied), then we have

$$\Delta(h)(s,t) = h(st) = \big(h_{(1)} \otimes h_{(2)}\big)(s,t) = h_{(1)}(s)h_{(2)}(t). \tag{73.24}$$

Thus we have

$$\int_G \Delta(h)(s,t)d\mu(s) = \int_G h_{(1)}(s)h_{(2)}(t)d\mu(s) = (\varphi \otimes \mathrm{Id})\big(h_{(1)} \otimes h_{(2)}\big)(t) = (\varphi \otimes \mathrm{Id})\Delta(h). \tag{73.25}$$

Using the invariance of the Haar measure, this is also equal to

$$\begin{aligned}\int_G \Delta(h)(s,t)d\mu(s) &= \int_G h(st)d\mu(s)\\ &= \int_G h(ts)d\mu(s)\\ &= \int_G h_{(1)}(t)h_{(2)}(s)d\mu(s)\\ &= (\mathrm{Id} \otimes \varphi)\big(h_{(1)} \otimes h_{(2)}\big)(t)\\ &= (\mathrm{Id} \otimes \varphi)\Delta(h).\end{aligned} \tag{73.26}$$

Since $C(G) \otimes CC(G)$ is dense in $C(G \times G)$, we have

$$(\mathrm{Id} \otimes \varphi)\Delta(h) = (\varphi \otimes \mathrm{Id})\Delta(h) = \varphi(h)1 \tag{73.27}$$

for every $h \in A$.

The generalisation of that result to an arbitrary compact quantum group is the following theorem.

Theorem 73.3.
If (A, Φ) is a compact quantum group, there exists one and only one state $\phi\colon A \to \mathbb{C}$ such that

$$(\mathrm{Id} \otimes \phi)\Phi(a) = \phi(a)1 = (\phi \otimes \mathrm{Id})\Phi(a) \tag{73.28}$$

for every $a \in A$.

The unique state guaranteed by that theorem is called the **Haar state** of A. If ω and ω' are linear functionals on A, we define the **convolution** by

$$\begin{aligned}\omega * a &= (\mathrm{Id} \otimes \omega)\Phi(a)\\ a * \omega &= (\omega \otimes \mathrm{Id})\Phi(a)\\ \omega * \omega' &= (\omega \otimes \omega')\Phi(a).\end{aligned} \tag{73.29}$$

DefProdConvCQG

The first two are elements of A when the lase one is a linear functional on A. Remember that here the map Φ plays the role of a coproduct; when the context is about a coalgebra, we need to adapt the notation and write Δ instead of Φ in the definitions 73.29.

The Haar state is the state ϕ such that

$$\phi * a = a * \phi = \phi(a)1. \tag{73.30}$$

73.2 Construction of $\mathrm{SU}_q(n)$

SecGeneratorsonSUQn

In order to build the quantum group $\mathrm{SU}_q(n)$, we will proceed the following step[991]

(1) we consider the n^2 elements $U = (u_{ij})$ and we name $\mathcal{A}_q$ the algebra generated by these elements;
(2) we impose some relations;
(3) we define the determinant D;
(4) we impose $D = 1$;
(5) we define the involution $u_{ij} \mapsto u_{ij}^*$;
(6) we impose $UU^* = U^*U = 1$.

ItemQGSUqnC

Notice that the point (5) will not introduce new elements u_{ij}^*. It will define u_{ij}^* as combination of the u_{ij}'s.

The first relations that we impose are

$$\sum_{kl} R_{ij}^{kl} u_{km} u_{lp} = \sum_{kl} R_{kl}^{mp} u_{ik} u_{jl} \tag{73.31}$$

EqRelsSUqnAvecR

where $R \in \mathbb{M}_{n^2}(\mathbb{R})$ is the matrix given by

$$\begin{aligned} R_{ii}^{ii} &= q^{-1} & R_{ij}^{ji} &= 1 \text{ if } i \neq j \\ R_{ij}^{ij} &= q^{-1} - q \text{ if } i > j \\ R_{ij}^{kl} &= 0 \text{ otherwise.} \end{aligned} \tag{73.32}$$

Remarque 73.4.
In [995], there is an alternative compact form for the same constraints.

SUBEquuijcondiv

Proposition 73.5.

$$u_{ij}u_{il} = qu_{il}u_{ij} \quad \text{if } j < l \tag{73.33a}$$
$$u_{ij}u_{kj} = qu_{kj}u_{ij} \quad \text{if } i < k \tag{73.33b}$$
$$u_{ij}u_{kl} = u_{kl}u_{ij} \quad \text{if } i > k, j < l \tag{73.33c}$$
$$u_{ij}u_{kl} - u_{kl}u_{ij} = (q - q^{-1})u_{il}u_{kj} \quad \text{if } i < k, j < l. \tag{73.33d}$$

subEquuijcondi subEquuijcondii subEquuijcondiii subEquuijcondiv

Proof. Let us check the relation (73.33c). If we consider i, j, m, p with $i < j$ and $m < p$, the only non vanishing R_{ij}^{kl} is $k = j$, $l = i$ and the only non vanishing R_{kl}^{mp} is $k = p$, $l = m$. Thus we immediately get $u_{jm}u_{ip} = u_{ip}u_{jm}$. □

A particular consequence of (73.33c) is that

$$u_{ij}u_{ji} = u_{ji}u_{ij}; \tag{73.34}$$

an element commutes with its "transposed".

lestmunijdiff

Lemma 73.6.
If $l = \min(i, j)$ with $i \neq j$ then

$$u_{ij}u_{ll} = q^{-1}u_{ll}u_{ij}. \tag{73.35}$$

Proof. If $l = i$ then $l < j$ and the relation (73.33a) provides $u_{lj}u_{ll} = q^{-1}u_{ll}u_{lj}$. Now if $l = j$, then $l < i$ and the relation (73.33b) provides $u_{il}u_{ll} = q^{-1}u_{ll}u_{il}$. □

lestmaxjdiff

Lemma 73.7.
If $l = \max(i, j)$ with $i \neq j$ then

$$u_{ij}u_{ll} = qu_{kk}u_{ij}. \tag{73.36}$$

Proof. If $l = i$, we have $j < i$ and the relation (73.33a) leads to $u_{li}u_{ll} = qu_{ll}u_{li}$. If $l = j$, then $i < j$ and the relation (73.33b) shows $u_{il}u_{ll} = qu_{ll}u_{il}$. □

The R matrix is a Yang-Baxter operator:

$$R_{(12)}R_{(23)}R_{(12)} = R_{(23)}R_{(12)}R_{(23)} \tag{73.37}$$

where the operators $R_{(12)}, R_{(23)} : \mathbb{C} \otimes \mathbb{C} \otimes \mathbb{C} \to \mathbb{C} \otimes \mathbb{C} \otimes \mathbb{C}$ are defined by

$$\begin{aligned} R_{(12)} &= R \otimes \mathrm{Id} \\ R_{(23)} &= \mathrm{Id} \otimes R. \end{aligned} \tag{73.38}$$

Now we define the homomorphisms

$$\begin{aligned} &\Delta : \mathcal{A}_q \to \mathcal{A}_q \otimes \mathcal{A}_q \\ &\Delta(u_{ij}) \mapsto \sum_{k=1}^{n} u_{ik} \otimes u_{kj} \end{aligned} \tag{73.39}$$

and

$$\begin{aligned} &\epsilon : \mathcal{A}_q \to \mathbb{C} \\ &\epsilon(u_{ij}) = \delta_{ij}. \end{aligned} \tag{73.40}$$

We also set $\epsilon(1) = 1$ as required for a counit (definition 54.24). At this point $\mathcal{A}_q$ is a bialgebra.

73.2.1 Determinant

We introduce the **quantum determinant**

$$D = \sum_{\sigma \in S_n} (-q)^{|\sigma|} u_{1\sigma(1)} \dots u_{n\sigma(n)} \tag{73.41}$$

where $|\sigma|$ is the length of the permutation. We also introduce the **quantum minor** D^{ij}

$$D^{ij} = \sum_{\sigma : \{1,\dots,\hat{i},\dots,n\} \to \{1,\dots,\hat{j},\dots,n\}} (-q)^{|\sigma|} u_{1\sigma(1)} \dots \widehat{u_{i\sigma(i)}} \dots u_{n\sigma(n)}. \tag{73.42}$$ Eqdefdijmineurquan

Let I and J be subsets of size t of $\{1, \dots, n\}$. We introduce the **quantum minor**

$$[I|J] = \sum_{\sigma \in S_t} (-q)^{|\sigma|} u_{i_{\sigma(1),j_1}} \dots u_{i_{\sigma(t),j_t}}. \tag{73.43}$$

The coproduct of that minor has the simple expression[996]

LemMineurQuantique

Proposition 73.8.
We have

$$\Delta[I|J] = \sum_{|K|=|I|} [I|K] \otimes [K|I] \tag{73.44}$$

where the sum is over all the subset $K \subseteq \{1, \dots, n\}$ of size equal to the size of I.

Sketch of the proof. We prove by induction on the size of I. Using the fact that Δ is an homomorphism,

$$\begin{aligned} \Delta[I|J] &= \sum_{\sigma \in S_t} (-q)^{|\sigma|} \sum_{k=1}^{n} u_{i_{\sigma(1)}k} \otimes u_{kj_1} \Delta\big(u_{i_{\sigma(2)}j_2} \dots u_{i_{\sigma(n)}j_n}\big) \\ &= \sum_{l=1}^{t} \sum_{\sigma \in S_t(l)} (-q)^{|\sigma|} \sum_{k=1}^{n} u_{lk} \otimes u_{kj_1} \Delta\big(u_{i_{\sigma(2)}j_2} \dots u_{i_{\sigma(n)}j_n}\big) \\ &= \sum_{l=1}^{t} \sum_{\sigma \in S_t(l)} (-q)^{|\sigma|} \sum_{k=1}^{n} u_{lk} \otimes u_{kj_1} \sum_{|K|=|I|-1} [I\backslash\{l\}|K][K|J\backslash\{j_1\}] \end{aligned} \tag{73.45}$$

where $S_t(l)$ is the subset of S_t of permutations such that $\sigma(1) = l$.

¡¡ Avertissement/question au lecteur !! 73.9
This proof is not finished. I think that there remain a few combinatorial work in order to get the result. At least it looks very like everything is going to sum up correctly.

□

Using lemma 73.8, we prove that

$$\Delta(D) = D \otimes D. \tag{73.46}$$

We also have $\epsilon(D) = 1$.

Now we add to $\mathcal{A}_q$ an element called D^{-1} on which we impose the conditions

$$DD^{-1} = D^{-1}D = 1. \tag{73.47}$$

We extend Δ and ϵ by

$$\begin{aligned} \Delta(D^{-1}) &= D^{-1} \otimes D^{-1} \\ \epsilon(D^{-1}) &= 1. \end{aligned} \tag{73.48}$$

The extension of $\mathcal{A}_q$ by D^{-1} is $\mathrm{Pol}\big(\mathrm{U}_q(n)\big)$, or $\mathrm{Pol}\big(\mathrm{SU}_q(n)\big)$ if we impose the extra relation $D = 1$.

For the involution we define κ by

$$\begin{aligned} \kappa(u_{ij}) &= (-q)^{i-j} D^{ji} D^{-1} \\ \kappa(D^{-1}) &= D. \end{aligned} \tag{73.49}$$

EqDefInvolutionSSUqn

This extends in an unique way to $\kappa\colon \mathrm{Pol}\big(\mathrm{U}_q(n)\big) \to \mathrm{Pol}\big(\mathrm{U}_q(n)\big)$ in such a way that $\mathrm{Pol}\big(\mathrm{U}_q(n)\big)$ becomes a Hopf algebra with κ as antipode.

We introduce the involution by

$$\begin{aligned} u_{ij}^* &= \kappa(u_{ji}) \\ (D^{-1})^* &= D. \end{aligned} \tag{73.50}$$

We have $DD^* = 1 = D^*D$.

Lemma 73.10.
We have $\epsilon(u_{ij}^*) = \delta_{ij}$.

Proof. If we apply ϵ on the definition (73.42) we get a non vanishing contribution only from the term $\sigma = \mathrm{Id}$. Thus we see that $\epsilon(D^{ij}) = \delta_{ij}$ and the result follows immediately. □

Notice that this result is part of the fact that $\mathrm{Pol}(U_q(n))$ is an Hopf algebra with involution (see lemma 54.36).

Using the formula

$$\delta_{ij} D = \sum_{k=1}^{n} (-q)^{k-j} u_{ik} D^{jk}, \tag{73.51}$$

one proves that

$$\sum_{k=1}^{n} u_{ik} u_{jk}^* = \delta_{ij} 1 = \sum_{k=1}^{n} u_{ki}^* u_{kj}. \tag{73.52}$$

Equustrunsugn

Lemma 73.11.
In $\mathrm{SU}_q(n)$ we have the relations[991, 995]

$$u_{kl}^* u_{ij} = u_{ij} u_{kl}^* \qquad \text{if } k \neq i \tag{73.53a}$$

eqRElsusikl

$$u_{ij}^* u_{kj} = q^{-1} u_{kj} u_{ij}^* + (q^{-1} - q) \sum_{s<j} u_{ks} u_{is}^* \qquad \text{if } i \neq k \tag{73.53b}$$

subequkluijeqkneqiusst

$$u_{ij}^* u_{ik} = q u_{ik} u_{ij}^* + (q^2 - 1) \sum_{s>i} u_{sj}^* u_{sk} \qquad \text{if } j \neq k \tag{73.53c}$$

$$u_{ij}^* u_{ij} = u_{ij} u_{ij}^* + (q^2 - 1) \sum_{s>i} u_{sj}^* u_{sj} + (1 - q^2) \sum_{s<j} u_{is} u_{is}^* \tag{73.53d}$$

EqLemsumsumstarpar

where $i,j,k,l \in \{1,2,\dots,n\}$ and the sums are running up to n.

Proof. We are starting from equation (73.31) that we multiply on the left by u^*_{ir}, on the right by u^*_{sp} and sum over i and p:

$$\sum_{pikl} R^{kl}_{ij} u^*_{ir} u_{km} u_{lp} u^*_{sp} = \sum_{irkl} R^{mp}_{kl} u^*_{ir} u_{ik} u_{jl} u^*_{sp}. \tag{73.54}$$

Using (73.52), the sum over p is easy to perform on the left hand side while the sum over i is easy to perform on the right hand side:

$$\sum_{ikl} R^{kl}_{ij} u^*_{ir} u_{km} \delta_{ls} = \sum_{pkl} R^{mp}_{kl} \delta_{rk} u_{jl} u^*_{sp}. \tag{73.55}$$

For the sake of unifying the notations, we rename the summation variable p to i in the right hand side:

$$\sum_{ik} R^{ks}_{ij} u^*_{ir} u_{km} = \sum_{il} R^{mi}_{rl} u_{jl} u^*_{si}. \tag{73.56}$$

At this point we have four possibilities following $s = j$ or $s \neq j$ and $m = r$ or not. We look at the case $s = j$, $m = r$. In the expression

$$\sum_{il} R^{mi}_{ml} u_{sl} u^*_{si}, \tag{73.57}$$

only the term $i = l$ is not zero. Thus the equality reduces to

$$\sum_{k} R^{ks}_{ks} u^*_{km} u_{km} = \sum_{i} R^{mi}_{mi} u_{si} u^*_{si}. \tag{73.58}$$

We divide the sum over k into $k = s$ and $k \neq s$ and the sum over i into $i = m$ and $i \neq m$:

$$R^{ss}_{ss} u^*_{sm} u_{sm} + \sum_{j>s} R^{ks}_{ks} u^*_{km} u_{km} = R^{mm}_{mm} u_{sm} u^*_{sm} + \sum_{i<m} R^{mi}_{mi} u_{si} u^*_{si}. \tag{73.59}$$

Changing the names $s \to i$ and $m \to j$ we have

$$q^{-1} u^*_{ij} u_{ij} = u_{ij} u^*_{ij} + (q^2 - 1) \sum_{k>i} u^*_{kj} u_{kj} + (1 - q^2) \sum_{k<j} u_{ik} u^* u_{ik}. \tag{73.60}$$

This proves the equality (73.53d). □

¡¡ Avertissement/question à la lectrice !! 73.12
For the other ones, I think that one has to work with the other possibilities about $s = j$ and $m = r$. To be checked.

The quantum group $\mathrm{SU}_q(n)$ is the unital C^*-algebra obtained from the completion of $\mathcal{A}_q$ on which the map Δ extends to a C^*-homomorphism

$$\Delta\colon \mathrm{SU}_q(n) \to \mathrm{SU}_q(n) \otimes \mathrm{SU}_q(n). \tag{73.61}$$

¡¡ Avertissement/question au lecteur !! 73.13
Is the completion with that requirement unique?

73.2.2 Norm

Propqpiideinve

Proposition 73.14.
If π is a representation of $\mathcal{A}_q$ on D, then $\|\pi(u_{ij})\|^2 \leqslant 1$.

Proof. Using relation (73.52) we have for every $f \in D$ that

$$\begin{aligned}\|f\|^2 &= \langle f, f\rangle \\ &= \langle \sum_k \pi(u_{ik})^*\pi(u_{ik})f, f\rangle \\ &= \sum_k \langle \pi(u_{ik})f, \pi(u_{ik})f\rangle \\ &= \sum_k \|\pi(u_{ik})f\|^2.\end{aligned} \tag{73.62}$$

If $\|f\| = 1$, we thus have $\|\pi(u_{ik})\| \leqslant 1$ and

$$\|\pi(u_{ik})\|^2 = \sup_{\|f\|=1} \|\pi(u_{ik})f\|^2 \leqslant 1. \tag{73.63}$$

□

CorOpOdquijInverti

Corollary 73.15.
The operator $\mathrm{Id} - q\pi(u_{ij})$ *is invertible.*

Proof. Since $q < 1$ and $\|\pi(u_{ij})\| \leqslant 1$, it is impossible to have $\big[\mathrm{Id} - q\pi(u_{ij})\big]f = 0$ without violating $\|\pi(u_{ij})\| \leqslant 1$. □

73.3 Representations of $\mathrm{SU}_q(3)$

We are following [995].

Lemxijxstijsuqt

Lemma 73.16.
In $\mathrm{SU}_q(3)$ *we have the relations*

subequustuuuutroisi
$$u_{11}u_{11}^* = (1-q^2) + q^2u_{11}^*u_{11} \tag{73.64a}$$
subequustuuuutroisii
$$u_{12}u_{12}^* = q^2(1-q^2) + q^2u_{12}^*u_{12} - q^2(1-q^2)u_{11}^*u_{11} \tag{73.64b}$$
$$u_{21}u_{21}^* = (1-q^2) + q^2u_{21}^*u_{21} - (1-q^2)u_{11}^*u_{11} \tag{73.64c}$$
$$u_{22}u_{22}^* = (1-q^2) + q^2u_{22}^*u_{22} - (1-q^2)(u_{21}^*u_{21} + u_{12}^*u_{12}) - (1-q^2)^2u_{31}^*u_{31}. \tag{73.64d}$$

Proof. Let us set $i = j = 1$ in (73.53d). The second sum vanishes while we reform the first sum by adding and substracting $(q^2-1)u_{11}^*u_{11}$. What we obtain is

$$u_{11}^*u_{11} = u_{11}u_{11}^* + (q^2-1)\underbrace{\sum_k u_{k1}^*u_{k1}}_{=1} - (q^2-1)u_{11}^*u_{11}. \tag{73.65}$$

So

$$q^2u_{11}^*u_{11} = u_{11}u_{11}^* + (q^2-1) \tag{73.66}$$

which is the relation (73.64a).

In order to prove the relation (73.64b) we start from

$$u_{12}u_{12}^* = u_{12}^*u_{12} - (q^2-1)(\underbrace{u_{22}^*u_{22} + u_{32}^*u_{32}}_{=1-u_{12}^*u_{12}}) - (1-q^2)u_{11}u_{11}^* \tag{73.67}$$

in which we substitute the value of $u_{11}u_{11}^*$ given by (73.64a).

¡¡ Avertissement/question à la lectrice !! 73.17
I guess this is the same for the other relations, setting other values for i and j.

□

Lemuijustijnstarnsubeqijuk

Lemma 73.18.
In $\mathrm{SU}_q(3)$ *we have the relations*

subeqlemuijuklstarnj

$$u_{11}(u_{11}^*)^n = (1-q^{2n})(u_{11}^*)^{n-1} + q^{2n}(u_{11}^*)^n u_{11} \tag{73.68a}$$
$$u_{12}(u_{12}^*)^n = q^2(1-q^{2n})(u_{12}^*)^{n-1} + q^{2n}(u_{12}^*)^n u_{12} - q^2(1-q^{2n})(u_{12}^*)^{n-1}u_{11}^*u_{11} \tag{73.68b}$$
$$u_{21}(u_{21}^*)^n = (1-q^{2n})(u_{21}^*)^{n-1} + q^{2n}(u_{21}^*)^n u_{21} - (1-q^{2n})(u_{21}^*)^{n-1}u_{11}^*u_{11} \tag{73.68c}$$

subeqlemuijuklstarnivd

$$u_{22}(u_{12}^*)^n = q^n(u_{12}^*)^n u_{22} - q^2\frac{1-q^{2n}}{q^n}(u_{12}^*)^{n-1}u_{11}^*u_{21} \tag{73.68d}$$
$$u_{22}(u_{21}^*)^n = q^n(u_{21}^*)^n u_{22} - \frac{1-q^{2n}}{q^n}(u_{21}^*)^{n-1}u_{11}^*u_{12} \tag{73.68e}$$
$$u_{23}(u_{21}^*)^n = q^n(u_{21}^*)^n u_{23} - \frac{1-q^{2n}}{q^n}(u_{21}^*)^{n-1}u_{11}^*u_{13}. \tag{73.68f}$$

Proof. Equation (73.68a) is checked by setting $i = j = 1$ in equation (73.53d).

Let us now check the equation (73.68d). First, writing (73.53b) with $i = j = l = 1$, $k = 2$ and taking the adjoint produces

Equduuuqstarlem

$$u_{21}u_{11}^* = qu_{11}^*u_{21}. \tag{73.69}$$

Now if we write equation (73.53b) with $k = 1$ and $l = i = j = 2$ we have

equdduusstarlem

$$u_{22}u_{12}^* = qu_{12}^*u_{22} - (1-q^2)u_{21}u_{11}^*. \tag{73.70}$$

Substituting (73.69) into (73.70) we get

$$u_{22}u_{12}^* = qu_{12}^*u_{22} - q(1-q^2)u_{11}^*u_{21} \tag{73.71}$$

which is (73.68d) with $n = 1$. We proceed now by induction over n. Using the relations

$$u_{22}u_{12}^* = qu_{12}^*u_{22} - (1-q^2)u_{21}u_{11}^* \tag{73.72a}$$
$$u_{21}u_{12}^* = u_{12}^*u_{21} \tag{73.72b}$$

we find

$$\begin{aligned} u_{22}(u_{12}^*)^{n+1} &= \left[q^n(u_{12}^*)^n u_{22} - q^{2-n}(1-q^{2n})(u_{12}^*)^{n-1}u_{11}^*u_{21}\right]u_{12}^* \\ &= q^{n+1}(u_{12}^*)^{n+1}u_{22} - q^n(1-q^2)(u_{12}^*)^n u_{21}u_{11}^* - q^{2-n}(1-q^{2n})(u_{12}^*)^{n-1}u_{11}^*u_{12}u_{21}. \end{aligned} \tag{73.73}$$

In the second term we substitute $u_{21}u_{11}^* = qu_{11}^*u_{21}$ and in the third one we substitute $u_{11}^*u_{12}^* = q^{-1}u_{12}^*u_{11}^*$. What we get is

$$u_{22}(u_{12}^*)^{n+1} = q^{n+1}(u_{12}^*)^{n+1}u_{22} - \left[q^{n+1}(1-q^2) + q^{1-n}(1-q^{2n})\right](u_{12}^*)^n u_{11}^*u_{21} \tag{73.74a}$$
$$= q^{n+1}(u_{12}^*)^{n+1}u_{22} - (q^{3+n} - q^{1-n})(u_{12}^*)^n u_{11}^*u_{21} \tag{73.74b}$$
$$= q^{n+1}(u_{12}^*)^{n+1}u_{22} - q^2\frac{1-q^{2(n+1)}}{q^{n+1}}(u_{12}^*)^n u_{11}^*u_{21}. \tag{73.74c}$$

This is equation (73.68d).

¡¡ Avertissement/question au lecteur !! 73.19
I guess the other ones are checked similarly.

□

By equation (73.33c) we have

$$u_{13}u_{31} = u_{31}u_{13}. \tag{73.75}$$

Lemma 73.20.
The elements u_{13} *and* u_{31} *are normal.*

Proof. What we have to show is $u_{13}^* u_{13} = u_{13} u_{13}^*$ (definition 72.19). The relations (73.53d) and (73.52) provide

$$u_{13}^* u_{13} = u_{13} u_{13}^* + (q^2-1)(\underbrace{u_{23}^* u_{23} + u_{33}^* u_{33}}_{=1-u_{13}^* u_{13}}) + (1-q^2)(\underbrace{u_{11} u_{11}^* + u_{12} u_{12}^*}_{=1-u_{13} u_{13}^*}) \tag{73.76a}$$

$$= u_{13} u_{13}^* + (q^2-1)(u_{13} u_{13}^* - u_{13}^* u_{13}). \tag{73.76b}$$

Thus

$$u_{13}^* u_{13} - u_{13} u_{13}^* = (q^2-1)(u_{13} u_{13}^* - u_{13}^* u_{13}) \tag{73.77}$$

and the combination $u_{13}^* u_{13} - u_{13} u_{13}^*$ vanishes. □

Let $\pi\colon SU_q(3) \to \mathfrak{B}(H)$ be a representation. We write $x_{ij} = \pi(u_{ij})$.

¡¡ Avertissement/question à la lectrice !! 73.21
Check if x_{1n} and x_{n1} are normal for every n and that $x_{1n}x_{n1} = x_{n1}x_{1n}$.

¡¡ Avertissement/question au lecteur !! 73.22
I think that looking very hard to [937] can help to justify the fact that x_{13} and x_{31} have a basis of eigenvectors. Since they are commuting, we have a common basis of eigenvectors.

In the remaining I suppose that x_{13} and x_{31} have an unique basis of common eigenvectors.

In [995], they ask to see at [997].

Propxuteigenuud

Proposition 73.23.
Let $h \in H\backslash\{0\}$ be a common eigenvector of x_{13} and x_{31}, namely suppose $x_{13}h = \lambda h$ and $x_{31}h = \rho h$. Then

$$x_{13}(x_{11}h) = \frac{\lambda}{q}(x_{11}h) \qquad x_{13}(x_{12}h) = \frac{\lambda}{q}(x_{12}h) \tag{73.78a}$$

$$x_{31}(x_{11}h) = \frac{\rho}{q}(x_{11}h) \qquad x_{31}(x_{21}h) = \frac{\rho}{q}(x_{21}h). \tag{73.78b}$$

Proof. This is nothing else than the relations

$$\begin{aligned} q u_{13} u_{11} &= u_{11} u_{13} \\ u_{12} u_{13} &= q u_{13} u_{12} \\ u_{11} u_{31} &= q u_{31} u_{11}. \end{aligned} \tag{73.79}$$

□

Let us suppose that $\rho \neq 0 \neq \lambda$ and iterate the equations of proposition (73.23):

$$x_{31}(x_{21}^n)h = \frac{\rho}{q^n} x_{21}^n h. \tag{73.80}$$

So if not vanishing, the vector $x_{21}^n h$ is an eigenvector for x_{31} with eigenvalue ρ/q^n. Since $q < 1$, we must have $k \in \mathbb{N}$ such that $x_{21}^k h \neq 0$ and $x_{21}^{k+1} h = 0$, if not we would have arbitrary large eigenvalues for x_{31}, which is impossible because $\|x_{13}\| \leqslant 1$ by proposition 73.14.

Doing the same with x_{11} and the eigenvector $x_{21}^k h$ of x_{31} we find $l \in \mathbb{N}$ such that

$$x_{11}^l x_{12}^k \neq 0 \tag{73.81}$$

and

$$x_{11}^{l+1} x_{12}^k = 0. \tag{73.82}$$

We name h_1 that vector: $h_1 = x_{11}^l x_{21}^k h$. It satisfies

$$x_{11} h_1 = 0 \tag{73.83a}$$

$$x_{21} h_1 = 0 \tag{73.83b}$$

$$x_{31} h_1 = \frac{\rho}{q^{l+k}} h_1. \tag{73.83c}$$

subeqxtuhurhoqlk

Indeed by construction $x_{11}h_1 = x_{11}^{l+1}x_{21}^k h = 0$ and

$$x_{21}h_1 = x_{21}x_{11}^l x_{21}^k h = \frac{1}{q^l}x_{11}^l x_{21}^{k+1}h = 0 \tag{73.84}$$

because (73.33b) implies $qx_{21}x_{11} = x_{11}x_{21}$.

Proposition 73.24.
We have $\|x_{31}\| = 1$ and every eigenvalue of x_{31} is of the form q^n for some $n \in \mathbb{N}$.

Proof. Using the same way as in the proof of proposition 73.14, we have, for every $f \in \mathscr{H}$,

$$\|f\|^2 = \langle \sum_k x_{ki}^* x_{ki} f, f\rangle = \sum_k \|x_{ki}f\|^2. \tag{73.85}$$

Taking $f = h_1$ and $i = 1$ we find

$$\|h_1\|^2 = \|x_{11}h_1\|^2 + \|x_{21}h_1\|^2 + \|x_{31}h_1\|^2 = \|x_{31}h_1\|^2. \tag{73.86}$$

Thus

$$\|x_{31}\| = \sup_{v\in\mathscr{H}} \frac{x_{13}v}{\|v\|} \geqslant \frac{\|x_{31}h_1\|}{h_1} = 1. \tag{73.87}$$

From proposition 73.14 we have $\|x_{31}\| \leqslant 1$ and thus $\|x_{31}\| = 1$.

Comparing (73.86) with (73.83c) we have

$$\|h_1\| = \left|\frac{\rho}{q^{l+k}}\right| \|h_1\|, \tag{73.88}$$

so that $|\rho/q^{l+k}| = 1$. □

Let h_2 be an eigenvector of x_{31} with eigenvalue ρ_2 with $|\rho_2| = 1$. We have

$$x_{31}x_{11}h_2 = \frac{\rho_2}{q}x_{11}h_2, \tag{73.89}$$

so that $x_{11}h_2$ has to be an eigenvector of x_{13} with eigenvalue ρ/q. Since $|\rho/q| > 1$ we have $x_{11}h_2 = 0$. In the same way we find $x_{21}h_2 = 0$.

Let $m \in \mathbb{N}$ be such that $x_{12}^m h_2 \neq 0$ and $x_{12}^{m+1}h_2 = 0$ and set

$$h_0 = x_{12}^m h_2. \tag{73.90}$$

Using the relations

$$u_{11}u_{12} = qu_{12}u_{11} \tag{73.91a}$$
$$u_{21}u_{12} = u_{12}u_{21} \tag{73.91b}$$
$$u_{31}u_{12} = u_{12}u_{31} \tag{73.91c}$$

we have

$$x_{11}h_0 = x_{21}h_0 = x_{12}h_0 = 0 \tag{73.92}$$

and, if we set $\rho_0 = \rho_2$,

$$x_{31}h_0 = \rho_0 h_0 \tag{73.93}$$

with $|\rho_0| = 1$.

In addition, if $x_{13}h_0 = \lambda_0 h_0$, we have $|\lambda_0| = q^2$. In order to prove this claim, we need some more relations. Applying (73.64a) to h_0 we have

$$x_{11}x_{11}^* h_0 = (1 - q^2)h_0. \tag{73.94}$$

Now we apply to h_0 the equality

$$x_{11}x_{11}^* + x_{12}x_{12}^* + x_{13}x_{13}^* = 1. \tag{73.95}$$

If $x_{13}h_0 = \lambda_0 h_0$ we have $x_{13}x_{13}^* h_0 = \overline{\lambda_0}\lambda_0 h_0 = |\lambda_0|^2 h_0$. Thus

$$x_{12}x_{12}^* h_0 = (q^2 - |\lambda_0|^2)h_0. \tag{73.96}$$

On the other hand applying (73.64b) to h_0 we have

$$x_{12}x_{12}^* h_0 = q^2(1 - q^2)h_0. \tag{73.97}$$

Thus we have $|\lambda_0| = q^2$.

What we have proved up to here is the following.

Propeignxutxtunonzeroeigvvap

Proposition 73.25.
If there exists $h \in \mathscr{H}$ such that $x_{13}h = \lambda h$ and $x_{31}h = \rho h$ with $\lambda \neq 0 \neq \rho$, then there exists $h_0 \in \mathscr{H}$ and $\psi, \varphi \in S^1$ such that $\|h_0\| = 1$,

$$x_{11}h_0 = x_{12}h_0 = x_{21}h_0 = 0 \tag{73.98}$$

and

$$x_{13}h_0 = \lambda_0 h_0 \tag{73.99a}$$
$$x_{31}h_0 = \rho_0 h_0, \tag{73.99b}$$

with $\lambda_0 = q^2\psi$ and $\rho_0 = \varphi$. The eigenvalues of x_{31} are of the form q^n.

¡¡ Avertissement/question au lecteur !! 73.26
Question: do we have $\|x_{13}\| = q^2$?

We already know that the representation space $\mathscr{H}$ has a basis of common eigenvectors of $x_{13} = \pi(u_{13})$ and $x_{31} = \pi(u_{31})$. The representations of $\mathrm{SU}_q(3)$ are now divided in families that depend on the existence of such eigenvectors with simultaneously non vanishing eigenvalues. The proposition 73.25 describes the situation where one has a common eigenvector with non vanishing eigenvalue of x_{13} and x_{31}. The following theorem describes the representations in that case.

ThoRepresSUqtnz

Theorem 73.27.
Let π be a representation of $\mathrm{SU}_q(3)$ on $\mathscr{H}$. Suppose that it contains a common eigenvector h_0 of x_{13} and x_{31} with simultaneously non vanishing eigenvalues. Let us write $x_{13}h_0 = \lambda_0 h_0$, $x_{31}h_0 = \rho_0 h_0$ with $\lambda_0 \neq 0 \neq \rho_0$. Then there exist $\psi, \varphi \in S^1$ such that $\lambda_0 = q^2\psi$ and $\rho_0 = \varphi$.

Moreover the representation is described in the following way: let

$$|0,0,0\rangle = h_0 \tag{73.100a}$$
$$|N,0,0\rangle = \frac{1}{A_N}(x_{11}^*)^N |0,0,0\rangle \tag{73.100b}$$
$$|N,M,0\rangle = \frac{1}{q^{M(N+1)}}(x_{12}^*)^M |N,0,0\rangle \tag{73.100c}$$
$$|N,M,L\rangle = \frac{1}{q^{NL}A_L}(x_{21}^*)^L |N,M,0\rangle \tag{73.100d}$$

where $N, M, L \in \mathbb{N}$ *and* $A_N = \Big(\prod_{j=1}^{N}(1-q^{2j})\Big)^{1/2}$. *Then the action of* $\mathrm{SU}_q(3)$ *on* $\mathscr{H}$ *is given by*

subeqxiiNMLa

$$x_{11}|N,M,L\rangle = (1-q^{2N})^{1/2}|N-1,M,L\rangle \tag{73.101a}$$
$$x_{12}|N,M,L\rangle = q^{N+1}(1-q^{2M})^{1/2}|N,M-1,L\rangle \tag{73.101b}$$
$$x_{13}|N,M,L\rangle = q^{2+N+M}\psi|N,M-1,L\rangle \tag{73.101c}$$
$$x_{21}|N,M,L\rangle = q^{N}(1-q^{2L})^{1/2}|N,M,L-1\rangle \tag{73.101d}$$
$$x_{22}|N,M,L\rangle = -q^{L+M+1}\bar{\psi}\bar{\varphi}|N,M,L\rangle \tag{73.101e}$$
$$-\Big((1-q^{2L})(1-q^{2M})(1-q^{2(N+1)})\Big)^{1/2}|N+1,M-1,L-1\rangle \tag{73.101f}$$
$$x_{23}|N,M,L\rangle = q^{L+1}(1-q^{2(M+1)})^{1/2}\bar{\varphi}|N,M+1,L\rangle \tag{73.101g}$$
$$-q^{M+1}\Big((1-q^{2L})(1-q^{2(N+1)})\Big)^{1/2}\psi|N+1,M,L-1\rangle \tag{73.101h}$$
$$x_{31}|N,M,L\rangle = q^{N+L}\varphi|N,M,L\rangle \tag{73.101i}$$
$$x_{32}|N,M,L\rangle = q^{M}(1-q^{2(L+1)})^{1/2}\bar{\psi}|N,M,L+1\rangle \tag{73.101j}$$
$$-q^{L}\Big(1-q^{2(N+1)}(1-q^{2M})\Big)\varphi|N+1,M-1,L\rangle \tag{73.101k}$$
$$x_{33}|N,M,L\rangle = -q^{L+M+1}(1-q^{2(N+1)})\psi\varphi|N+1,M,L\rangle \tag{73.101l}$$
$$-\Big((1-q^{2(L+1)})(1-q^{2(M+1)})\Big)^{1/2}|N,M+1,L+1\rangle, \tag{73.101m}$$

and

$$x_{11}^*|N,M,L\rangle = (1-q^{2(N+1)})^{1/2}|N+1,M,L\rangle \tag{73.102a}$$
$$x_{12}^*|N,M,L\rangle = q^{N+1}(1-q^{2(M+1)})^{1/2}|N,M+1,L\rangle \tag{73.102b}$$
$$x_{13}^*|N,M,L\rangle = q^{2+M+N}\bar{\psi}|N,M,L\rangle \tag{73.102c}$$
$$x_{21}^*|N,M,L\rangle = q^{N}(1-q^{2(L+1)})^{1/2}|N,M,L+1\rangle \tag{73.102d}$$
$$x_{22}^*|N,M,L\rangle = -q^{L+M+1}\varphi|\psi|^2|N,M,L\rangle \tag{73.102e}$$
$$-\Big((1-q^{2(L+1)})(1-q^{2(M+1)})(1-q^{2N})\Big)^{1/2}|N-1,M+1,L+1\rangle \tag{73.102f}$$
$$x_{23}^*|N,M,L\rangle = q^{L+1}(1-q^{2M})^{1/2}\varphi|N,M-1,L\rangle \tag{73.102g}$$
$$-q^{M+1}\Big((1-q^{2N})(1-q^{2(M+1)})\Big)^{1/2}\bar{\psi}|N-1,M+1,L\rangle \tag{73.102h}$$
$$x_{31}^*|N,M,L\rangle = q^{N+L}(1-q^{2L})^{1/2}\psi|N,M,L-1\rangle \tag{73.102i}$$
$$q^{L}\Big((1-q^{2N})(1-q^{2(M+1)})\Big)^{1/2}\bar{\varphi}|N-1,M+1,L\rangle \tag{73.102j}$$
$$x_{33}^*|N,M,L\rangle = q^{L+M+1}(1-q^{2N})^{1/2}\bar{\psi}\bar{\varphi}|N-1,M,L\rangle \tag{73.102k}$$
$$-\Big((1-q^{2L})(1-q^{2M})\Big)^{1/2}|N,M-1,L-1\rangle. \tag{73.102l}$$

The basis $\{|N,M,L\rangle\}$ *of* $\mathscr{H}$ *is orthonormal, that is*

$$\langle L,M,N|A,B,C\rangle = \delta_{AN}\delta_{BM}\delta_{CL}. \tag{73.103}$$

Before to pass to the proof, we give some formulas that are used in the proof.

LemTechrepresSUqtnez

Lemma 73.28.
Under the assumptions of theorem 73.27 we have

ItemLemTechrepresSUqtnez

(1) $u_{11}(u_{12}^*)^M h_0 = 0$.

Proof. We proof (1) by induction over M. For $M = 1$ we apply the relation

$$u_{11}u_{12}^* = \frac{1}{q}u_{12}^*u_{11} + (q^{-1} - q)(u_{22}^*u_{21} + u_{32}u_{31}) \tag{73.104}$$

to h_0. By construction $x_{11}h_0 = x_{21}h_0 = 0$. Since $qu_{31}u_{32}^* = u_{32}^*u_{31}$, the vector $u_{32}^*h_0$ is an eigenvector of x_{31} with eigenvalue ρ_0/q. Thus

$$u_{32}^*h_0 = 0 \tag{73.105}$$

because $|\rho_0/q| > 1$. Item (1) is proved.

□

proof of theorem 73.27. Existence of h_0, ψ and φ are part of proposition 73.25. The action of $\mathrm{SU}_q(3)$ on the basis $|N, M, L\rangle$ is computed using the commutation relations and the lemmas 73.16 and 73.18. That computation proves in the same time that the vectors $|N, M, L\rangle$ actually form a basis.

For sake of simplicity in the computations we introduce the vectors

$$h'_{N,M,L} = (x_{21}^*)^L(x_{12}^*)^M(x_{11}^*)^N h_0 \tag{73.106}$$

or

$$h'_{N,M,L} = A_N A_L q^{M(N+1)}q^N L|N, M, L\rangle. \tag{73.107}$$

Let us perform the computation for (73.101a). First we compute $x_{11}h'_{N,M,0}$:

$$x_{11}h'_{N,M,0} = x_{11}(X_{12}^*)^M(x_{11}^*)^N = q^{NM}x_{11}(x_{11}^*)^N(x_{12}^*)^M h_0. \tag{73.108}$$

We used $x_{12}^*x_{11}^* = qx_{11}^*x_{12}^*$. Using lemma 73.18 we get

$$x_{11}h'_{N,M,0} = q^{MN}\Big((1 - q^{2N})(u_{11}^*)^{N-1} + q^{2N}(u_{11}^*)^N u_{11}\Big)(u_{12}^*)^M h_0 \tag{73.109}$$

Using lemma 73.28 it remains

$$\begin{aligned} x_{11}h'_{N,M,0} &= q^{MN}(1 - q^{2N})(x_{11}^*)^{N-1}(x_{12}^*)^M h_0 && (73.110a)\\ &= q^{MN}(1 - q^{2N})\frac{1}{q^{M(N-1)}}(u_{12}^*)^M(u_{11}^*)^{N-1} && (73.110b)\\ &= q^M(1 - q^{2N})h'_{N-1,M,0}. && (73.110c) \end{aligned}$$

Using now the relation $u_{11}u_{21}^* = qu_{21}^*x_{11}$ we have

$$\begin{aligned} x_{11}h'_{N,M,L} &= x_{11}(x_{21}^*)^L h'_{N,M,0} && (73.111a)\\ &= q^L(x_{21}^*)^L x_{11}h'_{N,L,0} && (73.111b)\\ &= q^{L+M}(1 - q^2N)h'_{N-1,M,N}. && (73.111c) \end{aligned}$$

Using the conversion factor (73.107) we get the desired action:

$$x_{11}|N, M, L\rangle = (1 - q^{2N})^{1/2}|N - 1, M, L\rangle. \tag{73.112}$$

□

73.4 Quantized universal algebras

We follow [889].

Let $\mathfrak{g}$ be a complex simple Lie algebra of rank n with its standard Lie bialgebra structure (proposition 54.57). Let $q \in \mathbb{C}\backslash\{0\}$ a complex number which is not a root of unity [3].

We denote by $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$, by $\alpha_1, \dots, \alpha_n$ the simple roots and $A_{ij} = \frac{2(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}$ the Cartan matrix [4]. We also introduce the following notations:

$$\begin{aligned} q_i &= q^{(\alpha_i,\alpha_i)/2} \\ (a;t)_k &= (1-a)(1-at)\dots(1-at^{k-1}) \\ \binom{m}{n}_t &= \frac{(t;t)_m}{(t;t)_n(t;t)_{m-n}}. \end{aligned} \tag{73.113}$$

The latter are the t-binomial coefficients also called **Gauss polynomial**.

Definition 73.29.
Let $\mathfrak{g}$ be a simple complex Lie algebra with its Cartan matrix

$$A_{ij} = 2\frac{(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}. \tag{73.114}$$

The Hopf algebra $U_h\mathfrak{g}$ on $\mathbb{C}[\![h]\!]$ is the algebra generated by X_i, Y_i, H_i $(i = 1,\dots,n)$ which is complete for the h-adic topology and subject to the relations

$$\begin{aligned} [H_i, H_j] &= 0 \\ [H_i, X_j^\pm] &= \pm(\alpha_i,\alpha_j)X_j^\pm \\ [X_i^+, X_j^-] &= \delta_{ij}\frac{\sinh\left(\frac{h}{2}H_i\right)}{\sinh\left(\frac{h}{2}\right)} \end{aligned} \tag{73.115}$$

$$\sum_{k=0}^{1-A_{ij}} (-1)^k \binom{1-A_{ij}}{k} (X_i^\pm)^k X_j^\pm (X_i^\pm)^{1-A_{ij}-k} = 0 \qquad \text{if } i \neq j.$$

Proposition 73.30.
The following defined a structure of Hopf algebra on $U_h\mathfrak{g}$:

$$\begin{aligned} &\Delta(H_i) = H_i \otimes 1 + 1 \otimes H_i, && \Delta(X_i^\pm) = q^{-H_i/2} \otimes X_i^\pm + X_i^\pm \otimes q^{H_i/2}, \\ &S(H_i) = -H_i, && S(X_i^\pm) = -q_i^{\pm 1} X_i^\pm, \\ &\epsilon(H_i) = 0, && \epsilon(X_i^\pm) = 0. \end{aligned} \tag{73.116}$$

Lemma 73.31.
The algebra $U_h\mathfrak{g}$ accepts the set of generators $\{H_i, E_i, F_i\}$ with

$$\begin{cases} E_i = X_i^+ q^{-H_i/2} & \text{(73.117a)} \\ F_i = X_i^- q^{H_i/2}. & \text{(73.117b)} \end{cases}$$

The coproduct is given by

$$\Delta(E_i) = E_i \otimes 1 + q^{-H_i} \otimes E_i \tag{73.118a}$$
$$\Delta(F_i) = 1 \otimes F_i + F_i \otimes q^{H_i}, \tag{73.118b}$$

the antipode is

$$S(E_i) = -q^{H_i} E_i \tag{73.119a}$$
$$S(f_i) = -F_i q^{-H_i}, \tag{73.119b}$$

3. There are no $n \in \mathbb{N}$ such that $q^n = 1$.
4. This is a different convention from [889] in which $A_{ij} = 2(\alpha_i,\alpha_i)/(\alpha_j,\alpha_j)$.

and the counit is

$$\epsilon(E_i) = 0 \tag{73.120a}$$
$$\epsilon(F_i) = 0. \tag{73.120b}$$

Proof. In order to see how it works, we show the formula for $\Delta(E_i)$.

¡¡ Avertissement/question au lecteur !! 73.32
One has to check the others.

We consider the expansion $q^x = \sum_k a_k x^k$. We have

$$(a \otimes b)q^{X\otimes 1} = \sum_k a_k (a \otimes b)(X \otimes 1)^k \tag{73.121a}$$
$$= \sum_k a_k a X^k \otimes b \tag{73.121b}$$
$$= a q^X \otimes b. \tag{73.121c}$$

In the same way $(a \otimes b)q^{1\otimes X} = a \otimes b q^X$. Using these formulas we find

$$\Delta(E_i) = \Delta(X_i)\Delta(q^{-H_i/2}) = \big(q^{-H_i/2} \otimes X_i^{\pm} + X_i^{\pm} \otimes q^{H_i/2}\big) q^{-\frac{1}{2}(H_i\otimes 1)} q^{-\frac{1}{2}(1\otimes H_i)} \tag{73.122a}$$
$$= q^{-H_i/2}q^{-H_i/2} \otimes X_i^{\pm} q^{-H_i/2} + X_i^{\pm} q^{-H_i/2} \otimes q^{H_i/2} q^{-H_i/2} \tag{73.122b}$$
$$= q^{-H_i} \otimes E_i + E_i \otimes 1. \tag{73.122c}$$

□

If V_1 and V_2 are $\mathbb{K}[\![h]\!]$-modules we define $V_1 \hat{\otimes} V_2$ as the limit

$$V_1 \hat{\otimes} V_2 = \lim_{\leftarrow n} V_1/h^n V_1 \otimes V_2/h^n V_2. \tag{73.123}$$

This is the completion of $V_1 \otimes V_2$ for the h-adic topology.

¡¡ Avertissement/question à la lectrice !! 73.33
I've to know if that limit is the one defined in section 47.2.

ThoIsomUhGUg

Theorem 73.34.
Let $\mathfrak{g}$ be a simple complex finite dimensional Lie algebra with its standard Lie bialgebra structure.

(1) There exists an algebra isomorphism

$$U_h\mathfrak{g} \to (U\mathfrak{g})[\![h]\!] \tag{73.124}$$

which is the identity modulo $h\mathbb{C}[\![h]\!]$.

(2) That isomorphism can be chosen in such a way to be the identity on $\mathfrak{h}$.

(3) The center of $U_h\mathfrak{g}$ is isomorphic to $Z[\![h]\!]$ if Z is the center of $U\mathfrak{g}$.

For a proof, see [889], page 60.

¡¡ Avertissement/question à la lectrice !! 73.35
If you understand the proof, please write me an email.

A $U_h\mathfrak{g}$-module V is **finite dimensional** if it is a $\mathbb{K}$-module of finite type. The corresponding homomorphism $\pi : U_h\mathfrak{g} \to \mathrm{End}(V)$ is a finite dimensional **representation**.

This theorem allows to bring the representation theory of the classical case ($U\mathfrak{g}$) to the quantum case. More precisely if $\mathcal{M}_h$ is the category of finite dimensional $U_h\mathfrak{g}$-modules and if $\mathcal{M}_\mathfrak{g}$ is the one of finite dimensional $\mathfrak{g}$-modules, we can define a functor

$$\begin{aligned} F : \mathcal{M}_\mathfrak{g} &\to \mathcal{M}_h \\ V &\mapsto V[\![h]\!] \end{aligned} \tag{73.125}$$

and the action of $U_h\mathfrak{g}$ on $V[\![h]\!]$ is the composition of the action of $(U\mathfrak{g})[\![h]\!]$ and ψ where

$$\psi\colon U_h\mathfrak{g} \to (U\mathfrak{g})[\![h]\!] \tag{73.126}$$

is the isomorphism of theorem 73.34.

We also consider the functor

$$\begin{aligned} G\colon \mathcal{M}_h &\to \mathcal{M}_{\mathfrak{g}} \\ G(W) &= W/hW. \end{aligned} \tag{73.127}$$

The action of $U\mathfrak{g}$ on W/hW is given by

$$X \cdot [v] = \left[\psi^{-1}(X) \cdot v\right] \tag{73.128}$$

where $v \in W$, $[v]$ is the class modulo $[\![h]\!]$ and ψ is the isomorphism of theorem 73.34. Notice that the action is well defined since

$$\psi^{-1}(X) \cdot (v + hw) = \psi^{-1}(X) \cdot v + h\psi^{-1}(X) \cdot w. \tag{73.129}$$

The last term belongs to hW.

Proposition 73.36.
The functors F and G are a bijections between the isomorphisms classes of objects of $\mathcal{M}_{\mathfrak{g}}$ and $\mathcal{M}_h$. The simple modules in $\mathcal{M}_{\mathfrak{g}}$ correspond to the indecomposable modules in $\mathcal{M}_h$.

The reference [889] says that the proof is straightforward. I didn't tried [5].

Let P_+ be the set of dominant weights of $\mathfrak{g}$. We denote by $L_0(\Lambda)$ the $U\mathfrak{g}$-module of highest weight $\Lambda \in P_+$. Then we denote by $L(\Lambda)$ such that

$$L_0(\Lambda) = L(\Lambda)/hL(\Lambda). \tag{73.130}$$

The corresponding representation is written π_Λ. If $\Lambda = \omega_i$ is the ith fundamental weight (i.e. $\omega_i(H_j) = \delta_{ij}$) then we say that π_{ω_i} is the ith **fundamental representation** of $U_h\mathfrak{g}$.

Lemma 73.37.
The module $L(\Lambda)$ defined by $L_0(\Lambda) = L(\Lambda)/hL(\Lambda)$ is equivalently given by

$$L(\Lambda) = L_0(\Lambda)[\![h]\!]. \tag{73.131}$$

Proof. The quotient $L(\Lambda)/hL(\Lambda)$ consist in removing all the terms of nonzero order in h. Thus

$$L(\Lambda) = \frac{L(\Lambda)}{hL(\Lambda)}[\![h]\!]. \tag{73.132}$$

□

Proposition 73.38.
Let Λ be a dominant weight and $L_0(\Lambda)$ be the $U\mathfrak{g}$-module of highest weight Λ. We consider the $U_h\mathfrak{g}$-module $L(\Lambda) = L_0(\Lambda)[\![h]\!]$. We have the weight decomposition

$$L_0(\Lambda) = \bigoplus_{\lambda \in P(\lambda)} L_0(\Lambda)_\lambda \tag{73.133}$$

where $P(\Lambda)$ is the set of weights of $L_0(\Lambda)$ and

$$L_0(\Lambda)_\lambda = \{v \in L_0(\Lambda) \text{ tel que } av = \lambda(a)v\, \forall a \in \mathfrak{h}\}. \tag{73.134}$$

Correspondingly we have the decomposition

$$L(\Lambda) = \bigoplus_{\lambda \in P(\Lambda)} L(\Lambda)_\lambda \tag{73.135}$$

where

$$L(\Lambda)_\lambda = L_0(\Lambda)_\lambda[\![h]\!] = \{v \in L(\Lambda) \text{ tel que } av = \lambda(a)v\, \forall a \in \mathfrak{h} \subset U_h\mathfrak{g}\}. \tag{73.136}$$

5. Let me know if it is straightforward;)

Proposition 73.39.
The finite dimensional indecomposable $U_h\mathfrak{g}$-module $L(\Lambda)$ is generated by a vector v_Λ such that

$$\begin{cases} X_i^+ v_\Lambda = 0 & (73.137a) \\ H_i v_\Lambda = \Lambda(H_i) v_\Lambda. & (73.137b) \end{cases}$$

*This vector is the **highest weight vector** and Λ is the highest weight of $L(\Lambda)$.*

In the same way, the module $L(\Lambda)$ is generated by a vector $v_{\tilde{\Lambda}}$ such that

$$\begin{cases} X_i^- v_{\tilde{\Lambda}} = 0 & (73.138a) \\ H_i v_{\tilde{\Lambda}} = \tilde{\Lambda}(H_i) v_{\tilde{\Lambda}} & (73.138b) \end{cases}$$

*where $\tilde{\Lambda} = w_0\Lambda$ and w_0 is an element of the Weyl group of longest length. The vector $v_{\tilde{\Lambda}}$ is a **lowest weight vector** of $L(\Lambda)$.*

The vectors v_Λ and $v_{\tilde{\Lambda}}$ are unique up to scalar multiple.

73.4.1 Example with $\mathfrak{sl}(2,\mathbb{C})$

The classical representations of $\mathfrak{sl}(2,\mathbb{C})$ are given in section 53.18 and equations (53.339). In the case of $U_h\,\mathfrak{sl}(2,\mathbb{C})$, the representation space is given by $V_m[\![h]\!]$ and if $v \in V_m[\![h]\!]$, $Z \in U_h\,\mathfrak{sl}(2,\mathbb{C})$ we define the corresponding representation

$$Z \cdot v = \psi(Z) \cdot v \tag{73.139}$$

where $\psi(Z)$ $U\,\mathfrak{sl}(2,\mathbb{C})[\![h]\!]$ has a well defined action on $v \in V_m[\![h]\!]$. In order to determine the representation, we have to write the isomorphism ψ.

73.5 Quantum universal enveloping algebra

One speak about the representations of $U_q\mathfrak{g}$ in the chapter 10 of [998].

DefUqlG

Definition 73.40.
*The **quantum universal enveloping algebra** $U_q\mathfrak{g}$ is the complex unital algebra with generators X_i^+, X_i^-, K_i and K_i^{-1} $(i = 1,\dots,n)$ and the relations*

(1) $K_iK_i^{-1} = K_i^{-1}K_i = 1$;

(2) $[K_i, K_j] = 0$;

(3) $[X_i^+, X_j^-] = \delta_{ij}\frac{K_i^2 - K_i^{-2}}{q_i - q_i^{-1}}$;

EqUqlGdefiv

(4) $K_iX_j^\pm = q^{\pm(\alpha_i,\alpha_j)}X_j^\pm K_i$;

(5)

$$\sum_{k=0}^{1-A_{ij}} (-1)^k \begin{pmatrix} 1-A_{ij} \\ k \end{pmatrix}_{q_i} (X_i^\pm)^k X_j^\pm (X_i^\pm)^{1-A_{ij}-k} = 0 \tag{73.140}$$

where $q_i = q^{(\alpha_i,\alpha_i)/2}$.

We will sometimes write X_i for X_i^+ and Y_i for X_i^-.

Remarque 73.41.
In the literature we find other conventions. In [881] the algebra $U_q\mathfrak{g}$ has the relations

$$[E_i, F_j] = \delta_{ij}\frac{K_i - K_i^{-1}}{q - q^{-1}} \tag{73.141}$$

and

$$K_iE_jK_i^{-1} = q^{A_{ij}}E_j. \tag{73.142}$$

This correspond to an other choice of q and $K \to K^2$.

Proposition 73.42.
The algebra $U_q\mathfrak{g}$ becomes a Hopf algebra with the definitions

$$\Delta(K_i) = K_i \otimes K_i \qquad \Delta(X_i^\pm) = X_i^\pm \otimes K_i + K_i^{-1} \otimes X_i^\pm \tag{73.143a}$$
$$S(K_i) = K_i^{-1} \qquad S(X_i^\pm) = -q_i^{\pm 1} X_i^\pm \tag{73.143b}$$
$$\epsilon(K_i) = 1 \qquad \epsilon(X_i^\pm) = 0 \tag{73.143c}$$

subEgantpUqGl

where $q_i = q^{(\alpha_i,\alpha_i)/2}$.

Proof. We have to check the relations of definition 54.27 on the generators. Let us begin by $(\mathrm{Id}\otimes\epsilon)\Delta = \mathrm{Id}$:

$$(\mathrm{Id}\otimes\epsilon)\Delta K_i = (\mathrm{Id}\otimes\epsilon)(K_i \otimes K_i) = K_i \otimes 1 = K_i \tag{73.144a}$$
$$(\mathrm{Id}\otimes\epsilon)\Delta X_i^+ = (\mathrm{Id}\otimes\epsilon)(X_i^+ \otimes K + K^{-1} \otimes X_i^+) \tag{73.144b}$$
$$= X_i^+ \otimes 1 + K^{-1} \otimes 0 \tag{73.144c}$$
$$= X_i^+. \tag{73.144d}$$

We also have

$$\mu(\mathrm{Id}\otimes S)\Delta X_i^+ = \mu(X_i^+ \otimes SK + K^{-1} \otimes SX_i^+) \tag{73.145a}$$
$$= X_i^+ K^{-1} - q_i K^{-1} X_i^+ \tag{73.145b}$$
$$= 0 \tag{73.145c}$$
$$= \eta\epsilon X_i^+. \tag{73.145d}$$

We used the relation ((4)) of definition 73.40. □

We denote by $U_q\mathfrak{b}_+$ the Hopf subalgebra of $U_q\mathfrak{g}$ generated by $\{X_i^+, K_i, K_i^{-1}\}_{i=1,\dots,n}$ and by $U_q\mathfrak{b}_-$ the one generated by $\{X^-, K_i, K_i^{-1}\}_{i=1,\dots,n}$. The Hopf subalgebra generated by $\{K_i, K_i^{-1}\}_i$ will be denoted by $U_q\mathfrak{h}$.

We denote by $U_q\mathfrak{n}_+$ the subalgebra generated by $\{X_i^+\}_i$ and by $U_q\mathfrak{n}^-$ the one generated by $\{X_i^-\}$. These two subalgebras are not Hopf subalgebras since the coproduct of $X_i^\pm$ involves K_i. For each $\nu \in \mathbb{N}$ we will denote by $(U_q\mathfrak{n}_+)_\nu$ the subspace of $U_q\mathfrak{n}_+$ generated by the monomials of length ν:

$$X_{i_1}^+ X_{i_2}^+ \dots X_{i_\nu}^+. \tag{73.146}$$

73.5.1 Admissible modules

If V is a finite dimensional $U_q\mathfrak{g}$-module and if $\lambda \in \mathfrak{h}^*$ we define

$$V_\lambda = \{v \in V \text{ tel que } K_i v = q^{(\lambda,\alpha_i)} v\}. \tag{73.147}$$

Definition 73.43.
*A finite dimensional $U_q\mathfrak{g}$-module is **admissible** if it accepts the decomposition*

$$V = \bigoplus_{\lambda\in\mathfrak{h}^*} V_\lambda. \tag{73.148}$$

If V and W are admissible $U_q\mathfrak{g}$-modules we define

$$\begin{aligned} \psi_{V,W} : V \otimes W &\to V \otimes W \\ v \otimes w &\mapsto q^{(\lambda,\mu)} v \otimes w \end{aligned} \tag{73.149}$$

if $v \in V_\lambda$ and $w \in W_\mu$.

Theorem 73.44.
There exists an element $\Theta \in U_q\mathfrak{g}\hat{\otimes}U_q\mathfrak{g}$ such that

(1) Θ reads as a sum $\Theta = \sum_{\nu=0}^{\infty} \Theta_\nu$ with $\Theta_0 = 1 \otimes 1$ and

$$\Theta_\nu \in (U_q\mathfrak{n}_+)_\nu(U_q\mathfrak{h}) \otimes (U_q\mathfrak{n}_-)_\nu(U_q\mathfrak{h}); \tag{73.150}$$

(2) for every pair of modules V and W we have

$$\Theta\psi_{V,W}\Delta(a) = \Delta'(a)\Theta\psi_{V,W} \tag{73.151}$$

where $\Delta' = \sigma \circ \Delta$.

¡¡ Avertissement/question à la lectrice !! 73.45
Write me if you know a proof of that.

Let an admissible module

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda \tag{73.152}$$

with

$$V_\lambda = \{v \text{ tel que } K_i v = q^{(\lambda,\alpha_i)}v\}. \tag{73.153}$$

Let $v \in V_\lambda$, using the commutation relations we have

$$\begin{aligned} K_i X_j^+ v &= q^{(\alpha_i,\alpha_j)} X_j^+ K_i v && \text{(73.154a)}\\ &= q^{(\alpha_i,\alpha_j)} q^{(\lambda,\alpha_i)} X_j^+ v && \text{(73.154b)}\\ &= q^{(\alpha_j+\lambda,\alpha_i)} X_j^+ v. && \text{(73.154c)} \end{aligned}$$

Thus we have

$$X_j^+ V_\lambda \subset V_{\alpha_j+\lambda}. \tag{73.155}$$

73.5.2 Example on $U_q\,\mathfrak{sl}(2,\mathbb{C})$

For notational simplicity we write

$$[n] = \frac{q^n - q^{-n}}{q - q^{-1}} \tag{73.156}$$

and

$$[n]! = [1][2]\dots[n]. \tag{73.157}$$

We follow [881] and we consider the algebra $U_q\,\mathfrak{sl}(2,\mathbb{C})$ defined by the generators E, F, K, K^{-1} and the relations SubEqsDefUrsldc

$$\begin{aligned} KK^{-1} &= K^{-1}K = 1 && \text{(73.158a)}\\ KEK^{-1} &= q^2E && \text{(73.158b)} \quad \text{EqUqslKEK}\\ KFK^{-1} &= q^{-2}F && \text{(73.158c)} \quad \text{EqsiKFKqdGF}\\ [E,F] &= \frac{K - K^{-1}}{q - q^{-1}}. && \text{(73.158d)} \end{aligned}$$

The first point is to show that this family of algebra is the same as the one defined in definition 73.40 in which we pose $(\alpha,\alpha) = \frac{1}{2}$ because of equation (51.156). For this purpose we temporally rewrite the equations (73.158) as

$$\begin{aligned} HEH^{-1} &= r^2E && \text{(73.159a)}\\ HFH^{-1} &= r^{-2}F && \text{(73.159b)}\\ [E,F] &= \frac{H - H^{-1}}{r - r^{-1}} && \text{(73.159c)} \end{aligned}$$

and we consider the map $\psi(H) = K^2$, $\psi(E) = aX$, $\psi(F) = bY$. We have

$$\psi(HEH^{-1}) = aK^2XK^{-1} = aKq^{1/2}XK^{-1} = aqX, \tag{73.160}$$

and on the other hand $r^2\psi(E) = ar^2X$, so that we must have $q = r^2$. As for the commutator we have

$$[\psi(E), \psi(F)] = ab[X, Y] = ab\frac{K^2 - K^{-2}}{q^{1/4} - q^{-1/4}}, \tag{73.161}$$

so we need to fix a and b in such a way that

$$\frac{ab}{q^{1/4} - q^{-1/4}} = \frac{1}{r - r^{-1}}. \tag{73.162}$$

The algebras (73.158) is then the same as the algebra (73.40) with $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$.

Lemma 73.46.
There is an unique automorphism of $U_q\,\mathfrak{sl}(2, \mathbb{C})$ such that $\omega(E) = F$, $\omega(F) = E$ and $\omega(K) = K^{-1}$.

Proof. Unicity is automatic since ω is defined on the generators. The point is only to see that it extends as an automorphism. As an example we have $\omega(KEK^{-1}) = K^{-1}FK$ while $q^2\omega(E) = q^2F$, but equation (73.158c) says that $K^{-1}FK = q^2F$.

The other relations are checked in the same way. □

Lemma 73.47.
Let $m \geqslant 0$ and $n \in \mathbb{Z}$. We have the following relations:

$$E^mK^n = q^{-2mn}K^nE^m \tag{73.163a}$$
$$F^mK^n = q^{2mn}K^nF^m, \tag{73.163b}$$

and

$$[E, F^m] = [m]F^{m-1}\frac{q^{-(m-1)}K - q^{m-1}K^{-1}}{q - q^{-1}} \tag{73.164a}$$
$$= [m]\frac{q^{m-1}K - q^{-(m-1)}K^{-1}}{q - q^{-1}}F^{m-1} \tag{73.164b}$$

and

$$[E^m, F] = [m]\frac{q^{-(m-1)}K - q^{m-1}K^{-1}}{q - q^{-1}}E^{m-1} \tag{73.165a}$$
$$= [m]E^{m-1}\frac{q^{m-1}K - q^{-(m-1)}K^{-1}}{q - q^{-1}}. \tag{73.165b}$$

Proof. We check the relations (73.163a) and (73.164a). The other can be deduced applying ω. For the first one, we use the commutator (73.158b) under the form $EK = q^{-2}KE$. We have $E^mK = q^{-2m}KE^m$.

With $m = 1$ the relation (73.164a) reduces to the definition of $U_q\,\mathfrak{sl}(2, \mathbb{C})$. We proceed by induction on m: assuming that the result is true with $m - 1$ we have

$$[E, F^m] = [E, F^{m-1}]F + F^{m-1}[E, F] \tag{73.166}$$

in which we substitute

$$[E, F^{m-1}]F = [m - 1]F^{m-2}\frac{q^{-(m-2)}K - q^{m-2}K^{-1}}{q - q^{-1}}F. \tag{73.167}$$

Grouping the terms we find

$$[E, F^m] = \frac{F^{m-1}}{q - q^{-1}}\Big([m - 1](q^{-m}K - q^mK^{-1}) + K + K^{-1}\Big) \tag{73.168}$$

The coefficient with K in the parenthesis is

$$[m-1]q^{-m}+1=\frac{q^{m-1}-q^{-m+1}}{q-q^{-1}}q^{-m}+1. \tag{73.169a}$$

Taking the common denominator we find

$$q^{-m+1}\frac{q^m-q^{-m}}{q-q^{-1}}=q^{-m+1}[m]. \tag{73.170}$$

Putting all together we find the result. □

CorvKEnvKvaep

Corollary 73.48.
If v is an eigenvector of K with eigenvalue α, and if $v_n = E^n v$ is non vanishing, then it is an eigenvector with eigenvalue $q^{2n}\alpha$. In particular the vectors v_n are distinct.

Proof. This is a computations using the relation (73.163a):

$$KE^n v = q^{2n}E^n Kv = q^{2n}\alpha E^n v. \tag{73.171}$$

□

Since $\mathfrak{g}=\mathfrak{sl}(2,\mathbb{C})$, the Cartan algebra is one dimensional and the roots are just numbers. Let V be a $U_q\,\mathfrak{sl}(2,\mathbb{C})$-module. We consider, for $\lambda\neq 0$,

$$V_\lambda=\{v\in V \text{ tel que } Kv=\lambda v\}. \tag{73.172}$$

We say that λ is a **weight** is $V_\lambda\neq\{0\}$.

Let V be a $U_q\,\mathfrak{sl}(2,\mathbb{C})$-module and $\lambda\in\mathbb{C}$. One say that $v\neq 0$ is a **height weight vector** of weight λ if $Ev=0$ and $Kv=\lambda v$. The module V is said to be a **highest weight** if it is generated by a highest weight vector.

Lemma 73.49.
We have $EV_\lambda\subset V_{q^2\lambda}$ and $FV_\lambda\subset V_{q^{-2}\lambda}$.

Proof. If $v\in V_\lambda$ the relations (73.158b) and (73.158c) imply

$$K(Ev)=q^2EKv=q^2\lambda(Ev) \tag{73.173}$$

and

$$K(Fv)=q^{-2}FKv=q^{-2}\lambda(Fv). \tag{73.174}$$

□

PropFDmodulehashiweightvec

Proposition 73.50.
Every finite dimensional $U_q\,\mathfrak{sl}(2,\mathbb{C})$-module has a highest weight vector.

Proof. Since $\mathbb{C}$ is algebraically closed, the operator K has an eigenvector. Let $Kw=\alpha v$ with $\alpha\neq 0$. If $Ew=0$, this is a highest weight vector. If not we consider $w_n=E^n w$ ($n\in\mathbb{N}$). By corollary 73.48, the vectors w_n are eigenvectors of K with distinct eigenvalues; thus there exists a n such that $w_n\neq 0$ and $w_{n+1}=0$. □

Lemma 73.51.
If V is a finite dimensional $U_q\,\mathfrak{sl}(2,\mathbb{C})$-module, the elements E and F are nilpotent as endomorphisms of V.

Proof. Let U be an unitary matrix such that U^*EU is upper diagonal (proposition 46.3). The eigenvalue of E and U^*EU are the same and are on the diagonal of U^*EU. If we prove that the eigenvalue of E are zero, then U^*EU will be nilpotent and thus E itself will be nilpotent.

Let $Ev=\lambda v$. Since $EKv=\lambda q^{-2}Kv$, the vectors $K^n v$ are eigenvectors of E with distinct eigenvalues λq^{-2n}. This is impossible in a finite dimensional space, so $\lambda=0$. □

Proposition 73.52.
Let V be a $U_q\mathfrak{sl}(2,\mathbb{C})$-module and v a highest weight vector of weight λ. We consider the sequence $v_0 = v$

$$v_p = \frac{1}{[p]!}F^p v. \tag{73.175}$$

Then each time v_p is nonzero we have subeqsActionUqsldcVp

$$Kv_p = \lambda q^{-2p} v_p \tag{73.176a}$$ EqKvpqdpvp
$$K^{-1}v_p = \lambda^{-1} q^{2p} v_p \tag{73.176b}$$
$$Ev_p = \frac{q^{-(p-1)}\lambda - q^{p-1}\lambda^{-1}}{q - q^{-1}} v_{p-1} \tag{73.176c}$$
$$Fv_{p-1} = [p]v_p. \tag{73.176d}$$

Proof. For the first one, using the commutation relations,

$$Kv_p = \frac{1}{[p]!}KF^p v \tag{73.177a}$$
$$= \frac{1}{[p]!}q^{2p}F^p Kv \tag{73.177b}$$
$$= \frac{1}{[p]!}\lambda q^{2p}F^p v \tag{73.177c}$$
$$= \lambda q^{2p} v_p. \tag{73.177d}$$

The action of K^{-1} is immediately deduced from that one applying K on both sides. For the last one,

$$Fv_{p-1} = \frac{1}{[p-1]!}F^p v = \frac{1}{[p-1]!}[p]!v_p = [p]v_p. \tag{73.178}$$

And for the second one we write $EF^p v = [E, F^p]v + \underbrace{F^p Ev}_{=0}$ in which we substitute (73.164b). We find

$$Ev_p = \frac{q^{p-1}K - q^{-(p-1)}K^{-1}}{q - q^{-1}} v_{p-1}. \tag{73.179}$$

We already know the action of K and K^{-1} on v_{p-1}. □

Corollary 73.53.
The vectors v_p are eigenvectors of K with distinct eigenvalues.

Proof. This is equation (73.176a) and the fact that q is not a root of unity. □

The structure of the simple $U_q\mathfrak{sl}(2,\mathbb{C})$-modules is described by the theorems 73.54 and 73.55. ThoVfintemofdsldcun

Theorem 73.54.
Let V be a finite dimensional $U_q\mathfrak{sl}(2,\mathbb{C})$-module generated by a highest weight vector v of weight λ. Then ItemThoVintmoddcuni

(1) The weight λ reads $\lambda = \epsilon q^n$ where $\epsilon = \pm 1$ and $n = \dim(V) - 1$. ItemThoVintmoddcunii

(2) Set $v_0 = v$ and $v_p = \frac{1}{[p]!}F^p v$. We have $v_p = 0$ with $p > n$ and the set

$$\{v_0, v_1, \dots, v_n\} \tag{73.180}$$

is a basis of V. ItemThoVintmoddcuniii

(3) The action of K on V is diagonalisable and has the $m+1$ distinct eigenvalues

$$\{\epsilon q^n, \epsilon q^{n-2}, \cdots, \epsilon q^{-n+2}, \epsilon q^{-n}\}. \tag{73.181}$$

ItemThoVintmoddcuniv

(4) *Any highest weight vector in V is a multiple of v.*

ThoVfintemofdsldeux

Theorem 73.55.
A $U_q\,\mathfrak{sl}(2,\mathbb{C})$-module is simple if and only if it is generated by an highest weight. Moreover if two modules are generated by a highest weight vector of same weight, they are isomorphic as $U_q\,\mathfrak{sl}(2,\mathbb{C})$-modules.

Proof of theorem 73.54. We know that the vectors v_p are eigenvector of K with distinct eigenvalues. Thus there exits $n \in \mathbb{N}$ such that $v_n = 0$ and $v_{n+1} = 0$. For that value of n we have $v_m = 0\ \forall m > n$ and $v_m \neq 0$ when $m \leqslant n$. We also have

$$0 = Ev_{n+1} = \frac{q^{-n}\lambda - q^n\lambda^{-1}}{q - q^{-1}}v_n, \tag{73.182}$$

so that $q^{-n}\lambda - q^n\lambda^{-1} = 0$. This implies $\lambda = \pm q^n$.

Let us now prove that $\{v_0, \dots, v_n\}$ is a basis of V. This will show that $\dim(V) = n + 1$. First we know that the vectors v_i are eigenvectors of K for distinct eigenvalues, so that the set $\{v_i\}$ is free. By hypothesis, the vector space V is generated by v_0. From the relations (73.176), we see that the action of $U_q\,\mathfrak{sl}(2,\mathbb{C})$ on the vectors v_i will only generate linear combinations of the vectors v_i. Thus the set of v_i's is generating. The points (1) and (2) are proved.

For item (3), the operator K is diagonalisable because the vectors v_p form a basis of eigenvectors of K. The eigenvalues are given by the relation $Kv_p = \lambda q^{-2p}v_p$. Since $\lambda = \epsilon q^n$ we have the eigenvalues

$$\{\epsilon q^n, \epsilon q^{n-2}, \dots, \epsilon q^{-n}\} \tag{73.183}$$

corresponding to $p = 0, \dots, n$.

In order to prove point (4) let v' be an other highest weight vector. By definition $Ev' = 0$ and there exists an $\lambda' \in \mathbb{C}$ such that $Kv' = \lambda' v'$. Since v' is eigenvector of V, this is up to scalar multiple one of the vectors v_i. From the constraint $Ev_i = 0$ we see that $i = 0$ so that v' is multiple of $v_0 = v$.

□

Proof of theorem 73.55. First we suppose that V is generated by an highest weight v and we prove that it is simple. LetV' be a submodule of V. By proposition 73.50 V' has an highest weight vector v'. The vector v' is highest weight for V also, thus v' is a multiple of v. Since $v \in V'$ we have $V \subset V'$ and consequently $V = V'$.

Let V be a simple module, v a highest weight vector in V and V' the subspace of V generated by v. The space V' is a submodule of V while V is simple, so $V' = V$ and v generated V.

Finally let V be generated by v of weight λ and V' be generated by v' of same weight λ. Since

$$\lambda = \epsilon q^n = \epsilon' q^{n'} \tag{73.184}$$

we have $\epsilon = \epsilon'$ and[6] $n = n'$, so the modules V and V' have same dimension $n + 1$. The set $\{v_p = F^p v/[p]!\}$ is a basis of V while the set $\{v'_p = F^p v'/[p]!\}$ is a basis of V'. One checks that the map $\psi\colon V \to V'$, $\psi(v_i) = v'_i$ is an isomorphism of $U_q\,\mathfrak{sl}(2,\mathbb{C})$-modules. □

The conclusion is that the $U_q\,\mathfrak{sl}(2,\mathbb{C})$-modules are classified up to isomorphisms by their dimension $n + 1$ and $\epsilon = \pm 1$. We denote by $V_{\epsilon,n}$ the module of dimension $n + 1$ with highest weight $\lambda = \epsilon q^n$. The corresponding representation $\rho_{\epsilon,n}\colon U_q\,\mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V_{\epsilon,n})$ is given by

$$\begin{aligned} Kv_p &= \epsilon q^{n-2p}v_p \\ Ev_p &= \epsilon[n - p + 1]v_{p-1} \\ Fv_{p-1} &= [p]v_p \end{aligned} \tag{73.185}$$

with $p = 0, \dots, n$.

6. Once again we use the fact that q is not a root of unity.

73.6 Quantized function algebra

Definition 73.56.
*Let G be the connected simply connected Lie group with Lie algebra $\mathfrak{g}$. The **quantized algebra of regular functions** on the group G is the Hopf $*$-subalgebra of matrix elements of the unitarizable finite dimensional $U_q\mathfrak{g}$-modules. We denote it by $\mathbb{C}[G]_q$*

The multiplication law in $\mathbb{C}[G]_q$ is given by formula (54.94a):

$$c^{\Lambda'}_{l',v'}c^{\Lambda}_{l,v}(x) = (l'\otimes l)\rho_{\Lambda'\otimes\Lambda}(x)(v'\otimes v) = (l'\otimes l)(\Delta x)(v'\otimes v), \tag{73.186}$$

so

$$c^{\Lambda'}_{l',v'}c^{\Lambda}_{l,v} = c^{\Lambda'\otimes\Lambda}_{l'\otimes l,v'\otimes v} \tag{73.187}$$

where $\Lambda'\otimes\Lambda$ stands for the representation of $U_q\mathfrak{g}\otimes U_q\mathfrak{g}$ on $L(\Lambda')\otimes L(\Lambda)$. We can be more explicit if we write $\Delta x = \sum_i a_i\otimes b_i$, we have

$$c^{\Lambda}_{l,v}c^{\Lambda'}_{l',v'}(x) = \sum_i c^{\Lambda}_{l,v}(a_i)c^{\Lambda'}_{l',v'}(b_i) = \sum_i l\big(\rho_\Lambda(a_i)v\big)l'\big(\rho_{\Lambda'}(b_i)v'\big) \tag{73.188}$$

where ρ_Λ stands for the representation on $L(\Lambda)$. At the end the product $c^{\Lambda'}_{l',v'}c^{\Lambda}_{l,v}$ is still an element in $\mathbb{C}[G]_q$ and so a linear form on $U_q\mathfrak{g}$.

Remarque 73.57.
In the setting of Hopf algebra we consider the notion of involution of the definition 54.40 in which we do not require $a^{**} = a$.

Proposition 73.58.
If $q\in\mathbb{R}_0$ the map $\omega\colon U_q\mathfrak{g}\to U_q\mathfrak{g}$ defined on the generators by

$$\omega(X_i^{\pm}) = X_i^{\mp} \tag{73.189a}$$
$$\omega(K_i) = K_i. \tag{73.189b}$$

One also speaks about that structure in [999].

Since $U_q\mathfrak{g}$ is a Hopf $*$-algebra, the dual also becomes a Hopf $*$-algebra by proposition 54.41. The algebra $(U_q\mathfrak{g})^*$ contains in particular the matrix elements of $U_q\mathfrak{g}$-modules, so the space $\mathbb{C}[G]_q$ is a Hopf $*$-algebra.

Let $V = L(\Lambda)$ be the simple admissible $U_q\mathfrak{g}$-module of highest weight $\Lambda\in P_+$ where P_+ is the set of dominant weights in $\mathfrak{h}^*$. Since it is admissible we have the decomposition

$$L(\Lambda) = \bigoplus_{\lambda\in\mathfrak{h}^*} L(\Lambda)_\lambda \tag{73.190}$$

with

$$L(\Lambda)_\lambda = \{v\in L(\Lambda) \text{ tel que } K_iv = q^{(\lambda,\alpha_i)}v\}. \tag{73.191}$$

By general theory (see subsection 54.6.6), the dual vector space $L(\Lambda)^*$ is also a $U_q\mathfrak{g}$-module. We chose on $L(\Lambda)^*$ the left $U_q\mathfrak{g}$-module structure $(a\cdot l) = \mathcal{L}(a)l$ in the sense of definition (54.119), that means

$$(a\cdot l)(v) = l\big(S^{-1}(a)v\big) \tag{73.192}$$

for any $a\in U_q\mathfrak{g}$, $l\in L(\Lambda)^*$ and $v\in L(\Lambda)$.

Lemma 73.59.
The module $L(\Lambda)^*$ accepts the decomposition

$$L(\Lambda)^* = \bigoplus_{\lambda\in\mathfrak{h}^*} L(\Lambda)^*_\lambda \tag{73.193}$$

EqLLamstdecompllsma

where $L(\Lambda)^*_\lambda = \big(L(\Lambda)_{-\lambda}\big)^*$.

Proof. As vector space the decomposition (73.193) is nothing else that $(A \oplus B)^* = A^* \oplus B^*$. Let $\alpha \in \big(L(\Lambda)_\lambda\big)^*$. By definition (54.119) and using the antipode (73.143b) if $v \in L(\Lambda)_\lambda$ we have

$$(K_i\alpha)v = \alpha\big(S^{-1}(K_i)v\big) = \alpha(K_i^{-1}v) = \alpha\big(q^{-(\lambda,\alpha_i)}v\big) = q^{-(\lambda,\alpha_i)}\alpha(v). \tag{73.194}$$

Thus

$$K_i\alpha = q^{-(\lambda,\alpha_i)}\alpha. \tag{73.195}$$

This proves that

$$\big(L(\Lambda)_\lambda\big)^* = \big(L(\Lambda)^*\big)_{-\lambda}. \tag{73.196}$$

□

From a notational point of view, what we write $L(\Lambda)^*_\lambda$ is

$$L(\Lambda)^*_\lambda = \big(L(\Lambda)^*\big)_\lambda = \big(L(\Lambda)_{-\lambda}\big)^* \tag{73.197}$$

We can consider the regular left and right representations of $U_q\mathfrak{g}$ on $\mathbb{C}[G]_q$, since the latter is a part of the dual of $U_q\mathfrak{g}$. The space $\mathbb{C}[G]_q$ becomes a $U_q\mathfrak{g} \otimes U_q\mathfrak{g}$-module with the representation $\mathcal{L} \otimes \mathcal{R}$ described around equation (54.125).

Proposition 73.60.
The map

$$\begin{aligned} \psi\colon \mathbb{C}[G]_q &\to \bigoplus_{\Lambda\in P_+} L(\Lambda)^* \otimes L(\Lambda) \\ c^{L(\Lambda)}_{l,v} &\mapsto l \otimes v \end{aligned} \tag{73.198}$$

EqhomoCGqLLstar

is an isomorphism of $U_q\mathfrak{g} \otimes U_q\mathfrak{g}$-modules. Here P_+ is the set of dominant weights of $\mathfrak{g}$.

Proof. The fact that ψ is bijective is contained in the definition of $\mathbb{C}[G]_q$. The point is to check that this is a morphism. Let $a \otimes b \in U_q\mathfrak{g} \otimes U_q\mathfrak{g}$. The same computation as in (54.126) shows that

$$\big((a \otimes b) \cdot c^\Lambda_{l,v}\big)(x) = \big(\mathcal{L}(a)\mathcal{R}(b)c^\Lambda_{l,v}\big)(x) = c^\Lambda_{l,v}\big(S^{-1}(a)xb\big) = l\big(S^{-1}(a)xbv\big) = c^\Lambda_{\mathcal{L}(a)l,bv}(x), \tag{73.199}$$

so we have

$$\psi(a \otimes b) \cdot c^\Lambda_{l,v} = \mathcal{L}(a)l \otimes bv. \tag{73.200}$$

On the other hand we have

$$(a \otimes b)\psi c^\Lambda_{l,v} = (a \otimes b)l \otimes v = \mathcal{L}(a)l \otimes bv. \tag{73.201}$$

The map ψ is then an homomorphism.

The fact that the sum in (73.198) only runs over the dominant weight comes from subsection 53.21.3. □

Proposition 73.61.
Let w_0 be the element of the Weyl group which sends positive roots to negative ones[7]*. Then we have*

$$c^\Lambda_{-\Lambda,\lambda}c^{\Lambda'}_{-w_0\Lambda',\Lambda'} = q^{-2\langle\Lambda,\lambda\rangle}c^{\Lambda'}_{-w_0\Lambda',\Lambda'}c^\Lambda_{-\Lambda,\lambda}. \tag{73.202}$$

EqPropcclvlvppp

¡¡ Avertissement/question au lecteur !! 73.62
Formula (73.202) is not the one given in [889], equation (2.1.4) page 98. I don't know where I got wrong in the computation of the proof.

7. See theorem 51.175

Proof. Consider the elements $c^{\Lambda}_{l,v}$ and $c^{\Lambda'}_{l',v'}$ in $c^{\Lambda}_{-\Lambda,\lambda}$ and $c^{\Lambda'}_{-w_0\Lambda',\Lambda'}$ with

$$\begin{aligned} l &\in L(\Lambda)^*_{-\Lambda}, & v &\in L(\Lambda)_\lambda \\ l' &\in L(\Lambda')^*_{-w_0\Lambda'}, & v' &\in L(\Lambda')_{\Lambda'}. \end{aligned} \tag{73.203}$$

First notice that $l'(v') = 0$ because $l' \in \big(L(\Lambda')_{w_0\Lambda'}\big)^*$ while $v' \in L(\Lambda')_{\Lambda'}$ and $w_0\Lambda' \neq \Lambda'$. In the same way $\rho_{\Lambda'}(X_i^+)v' = 0$ since v' is a highest weight vector in $L(\Lambda')$.

The product $c^{\Lambda}_{l,v}c^{\Lambda'}_{l',v'}$ is defined by the equation (54.94a). If we apply it to K_i we find

$$\begin{aligned} (c^{\Lambda}_{l,v}c^{\Lambda'}_{l',v'})(K_i) &= (c^{\Lambda}_{l,v} \otimes c^{\Lambda'}_{l',v'})(\Delta K_i) \\ &= l\big(\rho_\Lambda(K_i)v\big)l'\big(\rho_{\Lambda'}(K_i)v'\big) \\ &= 0 \end{aligned} \tag{73.204}$$

because $\rho_{\Lambda'}(K_i)v'$ is a multiple of v'. The same shows that $(c^{\Lambda'}_{l',v'}c^{\Lambda}_{l,v})(K_i) = 0$.

We still have to check the equality on $X_i^\pm$. On the left hand side of (73.202) we get

$$(c^{\Lambda}_{l,v}c^{\Lambda'}_{l',v'})(X_i^\pm) = (c^{\Lambda}_{l,v} \otimes c^{\Lambda'}_{l',v'})(X_i^\pm \otimes K_i + K_i^{-1} \otimes X_i^\pm) \tag{73.205a}$$

$$= l\big(\rho_\Lambda(X_i^\pm)v\big)\underbrace{l'\big(\rho_{\Lambda'}(K_i)v'\big)}_{=0} + l\big(\rho_\Lambda(K_i^{-1})v\big)l\big(\rho_{\Lambda'}(X_i)^\pm v'\big) \tag{73.205b}$$

$$= q^{-\langle\Lambda,\lambda\rangle}l(v)l'\big(\rho_{\Lambda'}(X_i^\pm)v'\big). \tag{73.205c}$$

On the right hand side we have

$$(c^{\Lambda'}_{l',v'}c^{\Lambda}_{l,v})(X_i^\pm) = l'\big(\rho_{\Lambda'}(X_i^\pm)v'\big)l\big(\rho_{K_i}v\big) + \underbrace{l'\big(\rho_{\Lambda'}(K_i^{-1})v'\big)}_{=0}\, l\big(\rho_\Lambda(X_i^\pm)v\big) \tag{73.206a}$$

$$= q^{\langle\Lambda,\lambda\rangle}l'\big(\rho_{\Lambda'}(X_i^\pm)v'\big)l(v).j \tag{73.206b}$$

□

Chapter 74

von Neumann algebras

74.1 Functional, representation and automorphism

Let A be an algebra and consider a linear functional $\varphi\colon A \to \mathbb{C}$. That induces a GNS representation $\pi\colon A \to \mathrm{End}(V)$ with

$$V = \frac{A}{\{b \text{ tel que } \varphi(ab) = 0\, \forall a\}}. \tag{74.1}$$

If φ is nondegenerate, the latter ideal reduces to $\{0\}$. Let us assume a sort of Riesz theorem: for every linear functional $\psi\colon A \to \mathbb{C}$, there exists a $b \in A$ such that $\psi(a) = \varphi(ab)$, $\forall a$. If one fixes b, one can see $\varphi(ba)$ as a functional for a, and thus define $\sigma(b)$ by $\varphi(ba) = \varphi\big(a\sigma(b)\big)$.

One can check that this σ is an automorphism of A.

74.2 Commutant

Let $M \subset \mathfrak{B}(\mathscr{H})$ be a collection of bounded operators on $\mathscr{H}$. The **commutant** of M in $\mathfrak{B}(\mathscr{H})$ is

$$M' = \{S \in \mathfrak{B}(\mathscr{H}) \text{ tel que } TS = ST\, \forall T \in M\}. \tag{74.2}$$

From now let M be a self-adjoint algebra of operators on an Hilbert space $\mathscr{H}$ which contains $\mathbb{1}$. We have three lemmas.

LemUnVN

Lemma 74.1.
If $v \in \mathscr{H}$ and if P is the projection onto $\overline{Mv} \subseteq \mathscr{H}$, then P commutes with all operators in M.

LemDeuxVN

Lemma 74.2.
If $S \in M''$, $v \in \mathscr{H}$ and $\epsilon > 0$, there exists a $T \in M$ such that $\|Sv - Tv\| \leqslant \epsilon$. Moreover, in the finite dimensional case, there exists a $T \in M$ such that $Sv = Tv$.

Proof. Consider P as in lemma 74.1, then $P \in M'$, so that $XP = PX$. Now, the fact that M is unital gives $v \in \overline{Mv}$ from which we deduce $Pv = v$. Thus we have

$$Xv = XPv = PXv \in \overline{Mv}$$

because $Xv = P(Xv)$. The lemma now result from the fact that an element of the closure of Mv is as close as we want from Mv. □

LemTroisVN

Lemma 74.3.
If $T \in M''$, $v_1, \dots v_n \in \mathscr{H}$, and $\epsilon > 0$, there exists a $T \in M$ such that $\|Sv_i - Tv_i\| \leqslant \epsilon$ for all i. Moreover if the finite dimensional case, there exists a T such that $Sv_i = Tv_i$.

Proof. Apply lemma 74.2 to the set

$$N = \{T \oplus T \oplus \dots \oplus T \in \mathfrak{B}(\mathscr{H} \oplus \dots \oplus \mathscr{H}) \text{ tel que } T \in M\}.$$

□

Corollary 74.4.
In the finite dimensional case, we have $M'' = M$.

Proof. Take a basis of $\mathscr{H}$ in the lemma 74.3. □

ThoDoubleCommutant

Theorem 74.5 (Double commutant)**.**
Let $M \subset \mathfrak{B}(\mathscr{H})$ be a set of bounded operators on the Hilbert space $\mathscr{H}$. The double commutant M'' is the strong closure of M, i.e. the set of strong limits of nets.

Proof. A first evidence is that $M \subseteq M''$. Secondly, every commutant is strongly closed, so that M'' in particular is strongly closed. So it remains to be proved that for each $S \in M''$ is the limit of some net in M in the strong topology. For that, consider the directed set of finite subsets of $\mathscr{H}$ and consider the directed set

$$A = \{\text{finite subsets of } \mathscr{H}\} \times]0, \infty[$$

on which we say $(F, \epsilon) \geqslant (F', \epsilon')$ if and only if $F' \subseteq F$ and $\epsilon < \epsilon'$.

For each $a = (F, \epsilon) \in A$, we define $T_{(F,\epsilon)}$ as in lemma 74.3, so we have, for all $v \in F$,

$$\|T_{(F,\epsilon)} v - Sv\| \leqslant \epsilon,$$

which proves that $T_{(F,\epsilon)}$ converges to S in the strong topology. □

Notice that the operator S is in general *not* a limit of a sequence.

Definition 74.6.
*A **von Neumann algebra** is an unital $*$-subalgebra of $\mathfrak{B}(\mathscr{H})$ which is equal to its double commutant.*

Let G be a group and π, an unitary representation of G and M be the commutant of $\pi(G)$. Since $\pi(g)^* = \pi(g)^{-1} = \pi(g^{-1})$, we have that $\pi(G)$ is self-adjoint and M is a von Neumann algebra.

We say that a von Neumann algebra is a **factor** if its center is trivial, i.e. reduces to $\mathbb{C}$ Id.

PropprojrepresVN

Proposition 74.7.
Let π be a unitary representation of G on $\mathscr{H}$ and $\mathscr{H}_1$, an invariant subspace of $\pi(G)$. We have

- *the orthogonal projection on $\mathscr{H}_1$ belongs to M,*
- *if $P \in M$ is a projection, then $P\mathscr{H}$ is a subrepresentation of π, i.e. the closed subspace $\pi(G)P\mathscr{H}$.*

So, if π is irreducible then M is made of multiples of identity.

LeminvarMprime

Lemma 74.8 ([935])**.**
Let $\mathscr{S}$ be a selfadjoint part of $\mathfrak{B}(\mathscr{H})$. A closed subspace $\mathscr{H}_1$ of $\mathscr{H}$ is $\mathscr{S}$-invariant if and only if the orthogonal projection on $\mathscr{H}_1$ belongs to $\mathscr{S}'$.

Proof. First, suppose that $\mathscr{H}_1$ is $\mathscr{S}$-invariant. Thus for every $v \in \mathscr{H}_1$ and $w \in \mathscr{H}_1^\perp$, we have

$$0 = \langle Sv, w, \rangle = \langle v, S^* w \rangle$$

where $S^* \in \mathscr{S}$ by assumption. We conclude that $S^* w \perp \mathscr{H}_1$ for every $S \in \mathscr{S}$. Since $\mathscr{S}^* = \mathscr{S}$, we have that $\mathscr{H}_1^\perp$ is $\mathscr{S}$-invariant. Now if P is the orthogonal projector on $\mathscr{H}_1$, we decompose $x \in \mathscr{H}$ as

$$x = Px \oplus (\mathrm{Id} - P)x \in \mathscr{H}_1 \oplus \mathscr{H}_1^\perp. \tag{74.3}$$

EqDecomxPxSP

since $\mathscr{H}_1^\perp$ is $\mathscr{S}$-invariant, or every $S \in \mathscr{S}$, we have $S(\mathrm{Id} - P)x \in \mathscr{H}_1^\perp$, so that $PS(\mathrm{Id} - P)x = 0$. Thus, using the decomposition (74.3), $PSx = PSPx \oplus 0 = PSPx$. But $Px \in \mathscr{H}_1$, then $SPx \in \mathscr{H}_1$ and $P\big(SPx\big) = SPx$. We have proven that $PS = SP$, it is $P \in \mathscr{S}'$.

For the second part, assume that $P \in \mathscr{S}'$, then for every $v \in \mathscr{H}_1$, we have $v = Pv$ and

$$Sv = SPv = PSv \in \mathscr{H}_1,$$

which proves that $\mathscr{H}_1$ is $\mathscr{S}$-invariant. □

Proposition 74.9.
If $\mathscr{H}_1$ and $\mathscr{H}_2$ are equivalent subrepresentations of $\mathscr{H}$, then the intertwining operator $W : \mathscr{H}_1 \to \mathscr{H}_2$ determines on $\mathscr{H}$ a partial isometry such that $P_1 = W^*W$ and $P_2 = WW^*$.

The following decomposition is the **polar decomposition**.

PropPolarvNA

Proposition 74.10.
There exists a partial isometry $V : \overline{\text{Image}\left(|T|\right)} \to \overline{\text{Image}(T^*)}$ such that $T = V|T|$.

In other words, any element of a von Neumann algebra is the product of a positive operator by a projection. A version of this decomposition for operators in Hilbert spaces is given in lemme 60.66.

Proof. The operator V must satisfy $V\big(|T|v\big) = Tv$ for every $v \in \mathscr{H}$. That operator is a partial isometry because

$$\big\||T|v\big\|^2 = \langle |T|v, |T|v\rangle = \langle |T|^2 v, v\rangle = \langle T^*Tv, v\rangle = \|Tv\|^2.$$

It remains to be proved that $V \in M$. □

LemVNCommunit

Lemma 74.11.
Every von Neumann algebra is the commutant of an unitary representation.

Proof. Let M be a von Neumann algebra and consider that group $G = U(M')$, the group of unitary operators on M'. Let $T \in M'$, it reads as the sum of self-adjoint operators by

$$T = \frac{1}{2}(T + T^*) - \frac{i}{2}(iT - iT^*).$$

Thus we can restrict ourself to self-adjoint operators. Let $S \in M'$, it can be written as a combination of unitary operators:

$$S = \frac{1}{2}\left(S + i\sqrt{1 - S^2}\right) + \frac{1}{2}\left(S - i\sqrt{1 - S^2}\right).$$

Notice that the square root makes sense because $1 - S^2$ is positive. So M' is spanned by $U(M')$ and then $M = M'' = \big(U(M')\big)'$. □

74.3 Examples of von Neumann algebras

The very first example of von Neumann algebra is $\mathfrak{B}(\mathscr{H})$ itself which is the commutant of the identity.

74.3.1 Algebra $L^\infty(X)$

For the second example, consider an σ-finite measure space (X, μ), and then define $\mathscr{H} = L^2(X, \mu)$ and $M = L^\infty(X)$, the set of measurable bounded functions on X. The algebra M acts on $\mathscr{H}$ by pointwise multiplication and is therefore a $*$-subalgebra of $\mathfrak{B}(\mathscr{H})$. In order to prove that $L^\infty(X)$ is a von Neumann algebra, we prove that $L^\infty(X)' = L^\infty(X)$.

Assume for simplicity that $\mu(X) < \infty$ for simplicity. Now consider $T \in L^\infty(X)'$ and $f_)(x) = 1 \in L^\infty(X)$ because the measure is finite. For each $g \in L^\infty(X)$, we have

$$Tg = T(gf_0) = gTf_0 = gh$$

where $h = Tf_0$. So an element of $L^\infty(X)'$ reveals to be a multiplication by a function. We have

$$\|gh\|_2 = \|Tg\|_2 \leqslant \|T\| \cdot \|g\|_2$$

That proves that $\|h\|_\infty \leqslant \|T\|$ because if $\|h\|_\infty = \|T\| + \delta$, then we have a set (of non vanishing measure) on which h is bigger than $\|T\|$. We conclude that $h \in L^\infty(X)$.

74.3.2 Countable direct sum of Hilbert spaces

Consider $\mathscr{H}_\infty = \mathscr{H} \oplus \mathscr{H} \oplus \ldots = \{f\colon \mathbb{N} \to \mathscr{H} \text{ tel que } \sum \|f(n)\|^2 < \infty\}$. One can show that it is an Hilbert space. Let

$$M = \{T_\infty \text{ tel que } T \in \mathfrak{B}(\mathscr{H})\}$$

where $(T_\infty f)(n) = T\big(f(n)\big)$. One claims that this is a von Neumann algebra.

Let us see that of a direct sum of two copies of $\mathscr{H}$. In that case an element of M reads $\begin{pmatrix} T & 0 \\ 0 & T \end{pmatrix}$ with $T \in \mathfrak{B}(\mathscr{H})$. An element of M' must be of the form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with $[A,T] = [B,T] = [C,T] = [D,T] = 0$ for every $T \in \mathfrak{B}(\mathscr{H})$. Thus we have

$$M' = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ tel que } A,B,C,D \in \mathbb{C}\mathbb{1} \right\}. \tag{74.4}$$

In turn, one can see that the commutant of the right hand side is M itself.

Let us now go back with the case of $\mathscr{H}_\infty$. The map $T \mapsto T_\infty$ provides an isomorphism $M \simeq \mathfrak{B}(\mathscr{H})$ as $*$-algebras. But it is not sufficient to conclude that M is a von Neumann algebra because the topologies do not correspond. A net T_α strongly converges in M when $T_\alpha f$ converges for every $f \in M$. But f is an infinite list of vectors in $\mathscr{H}$, so that the strong topology in M is something like an infinite collection of strong topology on $\mathfrak{B}(\mathscr{H})$. Every strongly open set in $\mathfrak{B}(\mathscr{H})$ is strongly open in M, while the reciprocal is not true.

74.3.3 Direct limit

Let us take the example of direct limit of vector spaces given on page 2916. We are considering the matrix algebras $M_k = \mathbb{M}_{n_k}(\mathbb{C})$, and $A_n = M_1 \otimes_{\mathbb{C}} \ldots \otimes_{\mathbb{C}} M_n$, together with the maps

$$\begin{aligned} \sigma_n\colon A_n &\to A_{n+1} \\ T_1 \otimes \ldots \otimes T_n &\mapsto T_1 \otimes \ldots \otimes T_n \otimes \mathbb{1}. \end{aligned} \tag{74.5}$$

We pose $A = \lim_{\to} A_n$.

Lemma 74.12.
If π_1 and π_2 are $*$-homomorphisms from A to $\mathfrak{B}(\mathscr{H})$, then

$$\|\pi_1(a)\| = \|\pi_2(a)\| \tag{74.6}$$

for every $a \in A$. In other words, there is an unique way to close A in the norm topology.

¡¡ Avertissement/question au lecteur !! 74.13
If one does not ask π_i to be faithful, I can take $\pi_1(A) = 0$ as counter-example. Thus I think that I have to add the faithful assumption.

Proof. One has $\|\pi_i(a)\|^2 = \|\pi_i(a^*a)\|$, so that we only have to prove (74.6) in the case of self-adjoint elements of A. Using what is said in the proof of proposition 72.70, we have (with obvious notations) $\|\pi_i(b)\| = r_{\mathfrak{B}(\mathscr{H})}\big(\pi_i(b)\big)$. Since b belongs to one of the $A_k = \mathbb{M}_{n_k}(\mathbb{C})$ and $\dim A_k < \infty$. □

Now we ask the question of an actual way to represent A on an Hilbert space. First, as matrix algebra, $M_k = \mathfrak{B}(\mathscr{H}_k)$ for a certain finite dimensional Hilbert space $\mathscr{H}_k$. We form the Hilbert space

$$\mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_k$$

with the inner product

$$\langle v_1 \otimes \ldots \otimes v_k, w_1 \otimes \ldots \otimes w_k \rangle = \langle v_1, w_1 \rangle \ldots \langle v_k, w_k \rangle. \tag{74.7}$$

Now pick unit vectors $v_k \in \mathscr{H}_k$ and define the maps

$$\begin{aligned} \mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_{k-1} &\to \mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_k \\ w_1 \otimes \ldots \otimes w_{k-1} &\mapsto w_1 \otimes \ldots \otimes w_{k-1} \otimes v_\otimes v_k k \end{aligned} \tag{74.8}$$

that can be shown to be isometries. Finally, we consider the Hilbert space

$$H = \bigotimes_1^\infty (\mathscr{H}_k, v_k) = \text{completion of } \lim_{\to}(\mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_k). \tag{74.9}$$

Notice the dependence in the vectors v_k. Now we define the map $\pi\colon A \to \mathfrak{B}(H)$ by

$$\pi(T_1 \otimes \ldots \otimes T_k)(w_1 \otimes w_2 \otimes \ldots) = T_1 w_1 \otimes \ldots \otimes T_k w_k \otimes w_k \otimes w_{k+1} \otimes \ldots \tag{74.10}$$

More intrinsically, $\pi\colon \bigotimes M_k \to \mathfrak{B}\big(\bigotimes(\mathscr{H}_k, v_k)\big)$,

$$\pi\big(\otimes T_k\big)(\otimes w_k) = \otimes T_k w_k. \tag{74.11}$$

One can show that the strong closure of $\pi(A)$ is $\mathfrak{B}(H)$. The strong topology on $\mathfrak{B}(\mathscr{H})$ is generated by the open sets

$$\mathcal{U}(S, v, \epsilon) = \{T \in \mathfrak{B}(H) \text{ tel que } \|Tv - Sv\| \leqslant \epsilon\}$$

with $S \in \mathfrak{B}(\mathscr{H})$, $v \in H$ and $\epsilon > 0$. The direct limit $\mathscr{H} = \lim_{\to}(\mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_k)$, is given by a vector space $\mathscr{H}$ and maps φ_k such that

$$\begin{array}{ccc} \mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_k & \xrightarrow{\sigma} & \mathscr{H}_1 \otimes \ldots \otimes \mathscr{H}_{k+1} \\ & {\scriptstyle\varphi}\searrow \quad \swarrow{\scriptstyle\varphi_{k+1}} & \\ & \mathscr{H} & \end{array}$$

commutes where $\sigma(w_1 \otimes \ldots \otimes w_k) = w_1 \otimes \ldots \otimes w_k \otimes v_{k+1}$. The space $\mathscr{H}$ is the free vector space generated by symbols of the form $w_1 \otimes \ldots \otimes w_k \otimes v_{k+1} \otimes \ldots$ and we pose $\varphi_i(w_1 \otimes \ldots \otimes w_i) = w_1 \otimes \ldots \otimes w_i \otimes v_{i+1} \otimes \ldots$ and the Hilbert space H is the completion of $\mathscr{H}$.

The definition of A proceeds in the same way: M_i are matrix algebras and we pose $A_n = M_1 \otimes \ldots \otimes M_n$ with the map $\sigma(T_1 \otimes \ldots \otimes T_n) = T_1 \otimes \ldots \otimes T_n \otimes \mathbb{1}$, and

$$\begin{array}{ccc} A_n & \xrightarrow{\sigma} & A_{k+1} \\ & {\scriptstyle\varphi_n}\searrow \quad \swarrow{\scriptstyle\varphi_{k+1}} & \\ & A & \end{array}$$

with $\varphi_i(T_1 \otimes \ldots \otimes T_i) = T_1 \otimes \ldots \otimes T_i \otimes \mathbb{1} \otimes \ldots$, the space A being the free vector space generated by the symbols $T_1 \otimes \ldots \otimes T_j \otimes \mathbb{1} \otimes \ldots$

Now an element of H reads

$$\sum_{i=1}^\infty w_1^i \otimes \ldots \otimes w_{k_i}^i \otimes v_{k_i+1}^i \otimes \ldots$$

with $w_l^i \in \mathscr{H}_l$, and we act on it by an element of A by

$$\begin{aligned} \pi(T_1 \otimes \ldots \otimes T_k \otimes \mathbb{1} \otimes \ldots)\big(\sum_{i=1}^\infty w_1^i \otimes \ldots \otimes w_{k_i}^i \otimes v_{k_i+1}^i \otimes \ldots\big) \\ = \sum_{i=1}^\infty T_1 w_1^i \otimes \ldots \otimes T_{k_i} w_{k_i}^i \otimes v_{k_i+1}^i \otimes \ldots \end{aligned} \tag{74.12}$$

where some of the T_j are subject to actually be $\mathbb{1}$. We are now going to prove that the strong closure of $\pi(A)$ is $\mathfrak{B}(H)$. Let E be the set of finites sets of elements of the form $w_1 \otimes \ldots \otimes w_k \otimes v_{k+1} \otimes \ldots$ and the directed set $I = E \times]0, \infty[$.

¡¡ Avertissement/question au lecteur !! 74.14
I'm not sure of the next affirmation.

Let $T \in \mathfrak{B}(H)$. For each $(F,\epsilon) \in E$, there is a $T_{(F,\epsilon)} \in \pi(A)$

$$\|T_{(F,\epsilon)}X_k - TX_k\| < \epsilon \tag{74.13}$$

for all $X_k \in F$. In that case, the limit of the net $(F,\epsilon) \to T_{(F,\epsilon)}$ is T, which shows that the strong closure of $\pi(A)$ is $\mathfrak{B}(H)$.

So that example does not provide new example of von Neumann algebra.

Proposition 74.15.
Let $V_k = \mathbb{C}^{n_k}$ endowed with the standard representation of $\mathbb{M}_{n_k}(\mathbb{C})$. Then the representation $\pi\colon \bigotimes_k \mathbb{M}_{n_k} \to \mathfrak{B}\big(\bigotimes_k (V_k, v_k)\big)$ fulfils

$$\big(\pi(M)\big)'' = \mathfrak{B}\big(\bigotimes_k (V_k, v_k)\big)$$

where $M = \bigotimes_k \mathbb{M}_{n_k}$.

Proof. No proof. □

Notice that $\bigotimes_1^\infty \mathbb{M}_{n_k}$ can differ from $\bigotimes_1^\infty \mathbb{M}_{n'_k}$ when the n_k and n'_k do not agree. In fact, we have

$$\bigotimes_1^\infty \mathbb{M}_{n_k} = \bigotimes_1^\infty \mathbb{M}_{n'_k}$$

if and only if $\prod_k n_k = \prod_k n'_k$ in the sense of generalised products: each prime factor arise the same number of time in both side. For example, $\mathbb{M}_2 \otimes \mathbb{M}_2 \otimes \ldots = \mathbb{M}_4 \otimes \mathbb{M}_4 \otimes \ldots \neq \mathbb{M}_3 \otimes \mathbb{M}_3 \otimes \ldots$.

Take for example $V_k = \mathbb{C}^2$ for every k and then $v_k = v$, $v'_k = v'$. In this case, the representations π and π' are equivalent if and only if $v = \lambda v'$. This provides an Hilbert sphere of inequivalent irreducible representations.

74.4 Continuous dimensions

Let $\mathfrak{A}$ be a $*$-algebra with a faithful state φ defining the inner product $\langle a,b\rangle = \varphi(a^*b)$. We define a representation ρ of A on itself by $\rho(a)b = ab$.

Let us take the situation and the notations of proposition 72.59, and for $n = 2$, consider the choice

$$s_\lambda = \begin{pmatrix} \lambda/(1+\lambda) & \\ & 1/(1+\lambda) \end{pmatrix}$$

with $0 \leqslant \lambda \leqslant 1$.

Let $M = \bigotimes_1^\infty \mathbb{M}_2(\mathbb{C})$ and consider the positive continuous form φ_λ on M defined by

$$\varphi_\lambda(a_1 \otimes a_2 \otimes \ldots \otimes a_k \otimes \mathbb{1} \otimes \ldots) = \varphi_\lambda(a_1)\varphi_\lambda(a_2)\ldots = \operatorname{Tr}(a_1 s_\lambda)\operatorname{Tr}(a_2 s_\lambda) \tag{74.14}$$

Notice that the product is finite because from a certain point, $a_i = \mathbb{1}$. It is on the other hand not difficult to see that the $\varphi_\lambda(a) = 0$ only if $a_i = 0$ for every i. The purpose now is to follow the GNS construction of proposition 72.80. The remark we just made says that there is no ideal to quotient with in order to have the Hilbert space of representation.

The representation of $\mathfrak{A}$ we get is the simple $\rho(a)b = ab$ of $\mathfrak{A}$ on its completion. Let M_λ be the double commutant in this representation.

Theorem 74.16 (Powers)**.**
The von Neumann algebra M_λ are all distinct for different values of $0 \leqslant \lambda \leqslant 1$.

Proof. No proof. □

Let us see an example of that result. When $\lambda = 0$, we get the standard representation of matrices on $\mathbb{C}^n$, so that $M_0 = \mathfrak{B}(\mathscr{H})$. When $\lambda = 1$, we can show that

$$\mathrm{Tr}(T) = \langle V, TV \rangle$$

where $V = v \otimes v \otimes \ldots \otimes v \otimes \ldots$ is a trace on M_1 while there does not exist any trace on $\mathfrak{B}(\mathscr{H})$. We conclude that $M_1 \neq M_0$.

74.5 Cantor

Let $M = \bigoplus_{k=1}^{\infty} M_k$ with $M_k = \mathbb{M}_{n_k}(\mathbb{C})$ that can be seen as $\mathrm{Span}\{T_1 \otimes \ldots\}$ with $T_k = \mathbb{1}$ for sufficiently large k. Now take the case $n_k = 2$ for all k. The algebra M naturally acts on the space of locally constant functions on the Cantor set.

More generally, for arbitrary n_k, one can think of M as an algebra of endomorphisms of the space of locally constant functions on $\prod \mathbb{Z}_{n_k}$.

¡¡ Avertissement/question à la lectrice !! 74.17
Still to be developed.

74.6 More general state

Let φ_i be state on M_i. We define

$$\begin{aligned} \varphi \colon M &\to \mathbb{C} \\ T_1 \otimes T_2 \otimes \ldots &\mapsto \varphi_1(T_1)\varphi_2(T_2)\ldots \end{aligned} \tag{74.15}$$

where the product is finite because $T_k = \mathbb{1}$ for sufficiently large k. One can prove that φ fulfills

(1) $\varphi(\mathbb{1}) = 1$,

(2) $\varphi(T^*T) \geqslant 0$,

enuitemvarpsdex

(3) $\varphi(T^*S^*ST) = \|S\|^2\varphi(T^*T)$,

so that φ is in particular a state on M. The norm $\|S\|$ is the following. We know that $S \in \bigotimes_1^N M_k \subseteq M$ but, by construction, $\bigotimes_1^N M_k = \mathrm{End}(\mathbb{C}^{n_1} \otimes \ldots \otimes \mathbb{C}^{n_N})$. The norm of S is taken as the operator norm in the sense of that endomorphism space.

The property (3) shows that the multiplication by S is a bounded operator. We can build the GNS representation and define M_φ to be M'' is that representation.

Definition 74.18.
*A **factor of type** II_1 is an infinite dimensional factor M which accepts a non vanishing linear functional* $\mathrm{Tr}\colon M \to \mathbb{C}$ *such that*

— $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$,

— $\mathrm{Tr}(T^*T) \geqslant 0$,

— *the function* Tr *is continuous for the ultraweak topology.*

74.7 Group measure space construction

Let (X, μ) be a measured space that we assume to be σ-finite, and G, a discrete countable group acting on X in such a way that for every $g \in G$,

(1) if $E \subset X$ is measurable then gE is measurable,

(2) if $\mu(E) = 0$, then $\mu(gE) = 0$.

We do not impose the action to preserve the measure. As an example we take $G \subset \mathrm{SL}(2, \mathbb{R}) = \left\{ \frac{ax+b}{cx+d} \right\}$ acting on $\mathbb{R} \cup \{\infty\} = \mathbb{R}P^1$.

74.7.1 First attempt

Take $L^2(X,\mu)$ as Hilbert space and, to $f \in L^\infty(X,\mu)$, we associate the pointwise multiplication operator $M_f\colon L^2(X,\mu) \to L^2(X,\mu)$. We also introduce the new measure $g\mu$ by

$$(g\mu)(E) = \mu(g^{-1}E), \tag{74.16}$$

and the action of G on the functions by

$$(gf)(x) = f(g^{-1}x). \tag{74.17}$$

So we have

$$\int_X (gf)(x)d(g\mu)(x) = \int_X f(x)d\mu(x). \tag{74.18}$$

Notice that if $f \in L^\infty(X,\mu)$, we have $gf \in L^\infty(X,\mu)$, but when $f \in L^2$, there are no guarantee that $gf \in L^2$.

Let μ_1 and μ_2 be two measures on the set X. One says that μ_2 is **absolutely continuous** with respect to μ_1 if every μ_1 null set is μ_2 null.

ThoRadonNikodym

Theorem 74.19 (Radon-Nikodým)**.**
Let μ_1 and μ_2 be two σ-finite measures on X. The measure μ_2 is absolutely continuous with respect to μ_1 if and only if there exists a measurable positive function f such that $\mu_2 = f\mu_1$.

The useful statement in our case is:

Proposition 74.20.
There exists an unique function

$$\frac{df\mu}{d\mu}\colon X \to]0,\infty[$$

such that

$$\int (gf)(x)\left(\frac{dg\mu}{d\mu}\right)(x)d\mu(x) = \int (gf)(x)d(g\mu)(x) = \int f(x)d\mu(x). \tag{74.19}$$

EqDefRadonNiko

Applying g^{-1} to the function $(gf)(x)\left(\frac{dg\mu}{d\mu}\right)(x)$ and applying the theorem, we see that for every function f, we have

$$\int f(x)\left(\frac{dg\mu}{d\mu}\right)(gx)\left(\frac{dg^{-1}\mu}{d\mu}\right)(x) = \int f(f)d\mu(x),$$

so that

$$\left(\frac{dg\mu}{d\mu}\right)(gx)\left(\frac{dg^{-1}\mu}{d\mu}\right)(x) = 1. \tag{74.20}$$

Using the function provided by that theorem, we define

$$\mathcal{U}_g f = (gf)\cdot\left(\frac{dg\mu}{d\mu}\right)^{\frac{1}{2}}, \tag{74.21}$$

so that $\|\mathcal{U}_g f\|_{L^2(X,\mu)} = \|f\|_{L^2(X,\mu)}$. Moreover we have the following two important relations

(1) $\mathcal{U}_g\mathcal{U}_h = \mathcal{U}_{gh}$,

(2) $\mathcal{U}_g M_f \mathcal{U}_g^{-1} = M_{gf}$

where M_f stands for the operator of pointwise product with f. The second relation implies $\mathcal{U}_g M_f = M_{gf}\mathcal{U}_g$, thus we have

$$M := \{M_f, \mathcal{U}_g\}'' = \{\sum_{i=1}^n M_{f_i}\mathcal{U}_{g_i}\} \tag{74.22}$$

(because the double commutant is the strong closure) which is a $*$-algebra of operators. Notice that $\mathcal{U}_g\mathcal{U}_g^* = 1$ because $\mathcal{U}_g$ is an isometry.

We say that an action $G \times M \to M$ is **ergodic** when $gf = f$ for every $g \in G$ implies that f is constant almost everywhere.

LemergoBLCmu

Lemma 74.21.
If the action $G \times X \to X$ is ergodic, then $M = \mathfrak{B}\big(L^2(X,\mu)\big)$.

Proof. Let us first study the commutant $\{M_f, \mathcal{U}_g\}'$ which is of course contained in $\{M_f\}'$. But we know that the commutant of $L^\infty(X,\mu)$ is $L^\infty(X,\mu)$, so that

$$\{M_f\}' = \{M_f \text{ tel que } f \in L^\infty(X,\mu)\}.$$

Is there an element in $L^\infty(X,\mu)$ which commutes with all the elements of the form $\mathcal{U}_g$, or in other words, is there a $f \in L^\infty(X,\mu)$ such that $M_f = \mathcal{U}_g M_f \mathcal{U}_g^* = M_{gf}$? The only element f such that $M_f = M_{gf}$ for every $g \in G$ is $f = 1$, since the action is ergodic. □

That lemma shows that we didn't construct any interesting von Neumann algebras in the ergodic case.

74.7.2 Second attempt

Let $H = L^2(X \times G, \mu)$. Notice that $G \times X$ is nothing else than a countable number of copies of X, on which each of them we consider the measure μ. The multiplication operator is now replaced by

$$\begin{aligned} M_f &\colon H \to H \\ (M_f\varphi)(x,h) &= f(x)\varphi(x,h), \end{aligned} \tag{74.23}$$

and the unitary operator $\mathcal{U}_g$ is replaced by

$$\begin{aligned} \mathcal{U}_g &\colon H \to H \\ (\mathcal{U}_g\varphi)(x,h) &= \varphi(g^{-1}x, g^{-1}h)\left(\frac{dg\mu}{d\mu}\right)^{\frac{1}{2}}(x) \end{aligned} \tag{74.24}$$

The introduction of the Radon-Nikodým function serves to preserve the norm. The so defined operators have the following properties:

(1) $\mathcal{U}_g M_f \mathcal{U}_g^* = M_{gf}$,
(2) $\mathcal{U}_g \mathcal{U}_h = \mathcal{U}_{gh}$,
(3) $\mathcal{U}_g^{-1} = \mathcal{U}_{g^{-1}} = \mathcal{U}_g^*$.

Now we define

$$M(G,X) = \{\mathcal{U}_g, M_f\}'' = \{\sum_{i=1}^{n} M_{f_i}\mathcal{U}_{g_i}\}, \tag{74.25}$$

and we will show (later) that

Theorem 74.22.
If the action is ergodic, then $M(G,X)$ is a factor.

Proof. No proof up to now. □

We emphasize the progress: lemma 74.21 says that the commutant is trivial while now the center only is trivial.

As an example, take $X = S^1 \subset \mathbb{C}$ and $G = \mathbb{Z}$, the action being $g \cdot z = e^{2\pi i\theta g}z$ for some irrational θ. One can prove, using Fourier transform, that this equation is ergodic.

Let us now define

$$\begin{aligned} \varphi &\colon M(G,X) \to \mathbb{C} \\ T &\mapsto \langle f_0, T f_0\rangle \end{aligned} \tag{74.26}$$

where $f_0 \in L^2(X \times G, \mu)$ is defined as follows

$$f_0(g,x) = \begin{cases} \mu(X)^{-2} & \text{if } g = e \\ 0 & \text{otherwise.} \end{cases}$$

We show that the so defined φ is a trace over $M(G,X)$, i.e. it satisfies $\varphi(TS)=\varphi(ST)$ for every S, $T\in M(G,X)$. We know that $M(G,X)$ is generated by expressions of the form $L_f\mathcal{U}_g$. When $g\neq e$, the functions f_0 and $\mathcal{U}_g f_0$ have disjoint support, so that $\varphi(M_f\mathcal{U}_g)=0$. If $g=e$, the computation is easy and we finally find

$$\varphi(M_f\mathcal{U}_g)=\begin{cases}\frac{1}{\mu(X)}\int_X f(x)d\mu(x) & \text{if } g=e\\ 0 & \text{if } g\neq e.\end{cases} \tag{74.27}$$

It is easy to check that this expression is a trace on $\{\sum_i M_{f_i}\mathcal{U}_{g_i}\}$. Since $\mathfrak{B}(H)$ has no trace, we know that $M(G,X)$ is a non trivial von Neumann algebra. Stated in a different way, what we just proved is that, provided that μ is a G-invariant **probability measure** (i.e. $\mu(X)=1$), the formula

$$\varphi_\mu\Big(\sum_{h\in G} M_{f_h}\mathcal{U}_h\Big)=\int_X f_e(x)d\mu(x). \tag{74.28}$$

Eqvpmupickid

extends to a trace state on $M(G,X)$.

Proposition 74.23.
If μ is not invariant (but still $\mu(X)=1$), then φ_μ is still a state, but no more a trace.

Proof. By definition, $\varphi_\mu(T)=\langle f_0,Tf_0\rangle$ where $f_0(e)=1$ and $f_0(g)=0$ otherwise. We have

$$\begin{aligned}\varphi_\mu\Big(\big(\sum M_{f_h}\mathcal{U}_h\big)^*\big(\sum M_{f_h}\mathcal{U}_h\big)\Big)&=\varphi_\mu\Big(\sum_{h_1,h_2}\mathcal{U}_{h_1}^*M_{f_{h_1}}^*M_{f_{h_2}}\mathcal{U}_{h_2}\Big)\\&=\varphi_\mu\Big(\sum_{h_1,h_2}M_{f_{h_1}}^*M_{f_{h_2}}\mathcal{U}_{h_1^{-1}}\mathcal{U}_{h_2}\Big)\end{aligned}$$

because $\mathcal{U}_g^*=\mathcal{U}_g^{-1}=\mathcal{U}_{g^{-1}}$ commutes with the M_f. Now, according to the expression (74.28), the function φ_μ pick up the identity component and integrates. So we have

$$\begin{aligned}\varphi_\mu\Big(\big(\sum M_{f_h}\mathcal{U}_h\big)^*\big(\sum M_{f_h}\mathcal{U}_h\big)\Big)&=\sum_h\int_X h^{-1}(f_h^*f_h)d\mu\\&=\sum+h\int_X|f_h|^2(h^{-1}x)d\mu(x)\geqslant 0.\end{aligned}$$

□

74.7.3 First generalisation

Let us replace the space $L^\infty(X,\mu)$ by any von Neumann algebra N of operators on the Hilbert space $\mathscr{H}$. As Hilbert H space we take the completion of $\mathscr{H}\times l^2(G)$ and we assume to have an action of G over N. We know that the metric of a C^*-algebra is determined by its algebra structure, so that the action must be isometric. We assume the action to be *via* strongly continuous automorphisms. We have $H\otimes l^2(G)=l^2(G,H)$. For $t\in N$, we define

$$(M_T\varphi)(g)=g^{-1}(T)\varphi(g)$$

where $\varphi\in l^2(GmH)$. We also define

$$(\mathcal{U}_h\varphi)(g)=\varphi(h^{-1}g).$$

Notice that the group does not act on H while in the previous constructions, it did act on X. We still have the relation

$$\mathcal{U}_hM_T\mathcal{U}_h^*=M_{hT}. \tag{74.29}$$

Now, we form the von Neumann algebra

$$M(G,N)=\{\mathcal{U}_h,M_T\}''. \tag{74.30}$$

74.7.4 Second generalisation

Now we replace the discrete group G by a second countable locally compact topological group. Typical examples are $(\mathbb{R},+)$ or the group "$ax+b$" generated by matrices of the form $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$ with $a>0$ and $b\in\mathbb{R}$, acting on the real line by affine transformations. Each such group has an unique (up to constant multiple) Borel measure m such that each compact set has finite measure and which is in the same time the Haar measure: for every compactly supported functions on G,

$$\int_G f(hg)dm(g)=\int_G f(g)dm(g).$$

The measure on "$ax+b$" is $dm=\frac{1}{a^2}da\,db$.

We consider on $G\times N$ the product topology and we assume the action $G\times N\to N$ to be strongly continuous. Now, we proceed as before: we take the Hilbert space H the completion of

$$\mathscr{H}\otimes L^2(G,m)=L^2(G,\mathscr{H}),$$

and M_T is defined by the same formulas as before. The von Neumann algebra that we obtain is denoted by $M(G,N)$.

74.7.5 One particular case

sssOnePartCaseMG

Take a discrete group G, so that $M(G)$ is generated by the operators $\mathcal{U}_g\in\mathfrak{B}\big(l^2(G)\big)$ who are defined by

$$(\mathcal{U}_h\varphi)(g)=\varphi(h^{-1}g). \tag{74.31}$$

Consider the function

$$f_e(g)=\begin{cases}1 & \text{if } g=e\\ 0 & \text{otherwise.}\end{cases} \tag{74.32}$$

We have

$$(\mathcal{U}_hf_e)(g):=f_h(g)=\begin{cases}1 & \text{if } g=h\\ 0 & \text{otherwise,}\end{cases}$$

so that $\{f_h\}$ is an orthonormal basis of $l^2(G)$ and $\mathcal{U}_{h_1}f_{h_2}=f_{h_1h_2}$. We conclude that $\mathcal{U}_h$ is a permutation of the basis vectors.

ProplDeuxGFGP

Proposition 74.24.
The module $l^2(G)$ over $M(G)$ is projective and finitely generated.

Proof. The fact that $l^2(G)$ as module over $M(G)$ is finitely generated comes from the fact that f_e by itself generates the basis $\{f_h\}$ as we just said.

In order to prove that the module is projective, we will prove the condition (2) of proposition 54.14. Let $\mathfrak{M}$ be a $M(G)$-module and $\rho\colon\mathfrak{M}\to l^2(G)$ be a surjective module map. Consider any $\xi\in\mathfrak{M}$ such that $\rho(\xi)=f_e$, define $s(f_e)=\xi$ and extend by linearity and action of $M(G)$. The so defined map s obviously fulfils $\rho\circ s=\mathrm{Id}\,|_{l^2(G)}$. □

¡¡ Avertissement/question à la lectrice !! 74.25
That statement and the proof are correct uhm? I use them on page 3875.

For each $S\in M(G)'$, we define $f_S=S(f_e)\in l^2(G)$. Now if $S\in M(G)\cap M(G)'$, we have

(1) $S\mathcal{U}_h=\mathcal{U}_hS$,

(2) $S\mathcal{U}_hf_e=\mathcal{U}_hSf_e$.

The operators W_k defined by

$$(W_kf)(g)=f(gk)$$

commute with $\mathcal{U}_h$. So they commute with $M(G)$ and with S. Therefore we have

$$W_k S f_h = S W_k f_h = S f_{hk^{-1}},$$

and taking $h = k$, we find

$$f_S(hg) = f_S(gh) \tag{74.33}$$

for every g, $h \in G$. That means that f_S is constant on the conjugacy classes because g and hgh^{-1} are two elements of the form $g_1 g_2$ and $g_2 g_1$ with $g_1 = h$ and $g_2 = gh^{-1}$.

If G has no finite conjugacy class (except $\{e\}$), then f_S has to be a multiple of f_e because its norm would contains infinitely many constant non zero terms. But when S_1 and S_2 belongs to $M'(G)$ with $f_{S_1} = f_{S_2}$, then $S_1 = S_2$ because $S_1 f_h = S_1 \mathcal{U}_h f_e = \mathcal{U}_1 S_1 f_e = \mathcal{U}_h f_{S_1}$, while the same computation with S_2 gives $S_2 f_h = \mathcal{U}_h f_{S_2}$. So S_1 and S_2 agree on a basis.

For this reason, if we assume that G has infinite conjugacy class, we have $M(G)' \cap M(G) = \mathbb{C}\mathbb{1}$. The most famous example of such a group is $G = F_k$, the group of formal words of $1 \dots k$ with $k > 1$.

Now fix an element of finite order $h \in G$, and the natural homomorphism $\mathbb{Z} \to G$ given by $n \mapsto h^n$. The group G has a natural equivalence relation $g_1 \sim g_2$ if and only if there exists a $n \in \mathbb{Z}$ such that $g_1 = g_2^n$. We denote by $G/\mathbb{Z}$ the quotient of G by this relation. One class of this space is

$$\mathbb{Z}_g = \{g^n \text{ tel que } n \in \mathbb{Z}\},$$

and we have

$$l^2(G) = \bigoplus_{g \in G/\mathbb{Z}} l^2(\mathbb{Z}_g). \tag{74.34}$$

The operator $\mathcal{U}_g$ acts on $l^3(\mathbb{Z}_g)$ by translation: $\mathcal{U}_g g^n = g^{n+1}$. Now we fix a $g \in G$ and we make the identifications

$$l^2(\mathbb{Z}_g) \simeq l^2(\mathbb{Z}) \simeq L^2(S^1), \tag{74.35}$$

the latter identification being given by the Fourier series of a function on the circle (seen as a periodic function on $\mathbb{R}$). In that framework, $\mathcal{U}_g$ acts by translation of 1 on $l^2(\mathbb{Z})$:

$$\mathcal{U}_g(x_i)_{i \in \mathbb{N}} = (x_i + 1)_{i \in \mathbb{N}}.$$

A function $f \in C(S^1)$ is bounded (since S^1 is compact), to that if $s \in L^2(S^1)$, the product function (fs) still belongs to $L^1(S^1)$ and then corresponds to an element of $l^2(S^1)$. Thus we have a map

$$C(S^1) \to \mathfrak{B}\big(l^2(\mathbb{Z}_g)\big).$$

If one acts separately on each of the "fixed" g, we obtain an action

$$C(S^1) \to \mathfrak{B}\Big(\bigoplus_{g \in G/\mathbb{Z}} l^2(\mathbb{Z}_g)\Big).$$

We define the operator $\mathcal{U} \in \bigoplus_{g \in G/\mathbb{Z}} l^2(\mathbb{Z}_g)$ as acting on $l^2(\mathbb{Z}_g)$ with $\mathcal{U}_g$. That allows us to consider

$$\begin{aligned} C(S^1) &\to M(G) \\ z^n &\mapsto \mathcal{U}^n \end{aligned} \tag{74.36}$$

that can be composed with the trace φ to give

$$f \mapsto \int_{S^1} f(z) dm(z)$$

for each $f \in C(S^1)$.

74.8 More about projections

Let M be a von Neumann algebra and p, q, two projections in $\mathfrak{B}(\mathscr{H})$. We define

$$p \vee q = \text{projection onto } \overline{\text{Image}(p) + \text{Image}(q)} \tag{74.37}$$

$$p \wedge q = \text{projection onto } \text{Image}(p) \cap \text{Image}(q) \tag{74.38}$$

A projection is determined by the closed space of its range, so one can order the projection by ordering the closed subspace of $\mathscr{H}$.

LemDimSupDeuxProjs

Lemma 74.26.
We have

$$\dim(R_1 \vee R_2) = \dim R_1 + \dim R_2 - \dim(R_1 \wedge R_2) \tag{74.39}$$

for any two projections R_1 and R_2 in M.

Proof. Left as an exercise. □

For two projections p and q, we write that $p \preceq q$ if there exists a partial isometry V such that

(1) $V^*V = p$,
(2) $VV^* \leqslant q$.

One proves that this determines a partial ordering on the projections.

Proposition 74.27.
If p and q belong to M, then the projections $p \vee q$ and $p \wedge q$ belong to M too.

Proof. There exists an unitary representation $\pi\colon G \to \text{End}(\mathscr{H})$ whose M is the commutant: $M = \pi(G)'$ by lemma 74.11. Since p is a projection, proposition 74.7 says that $p\mathscr{H}$ is an invariant subspace of $\pi(G)$. The same being true for q, we have that the spaces $\overline{\text{Image}(p) + \text{Image}(q)}$ and $\text{Image}(p) \cap \text{Image}(q)$ are invariant too. Now the proposition 74.7 assures that the corresponding projections (namely $p \vee q$ and $p \wedge q$) are part of M. □

Before to define the infinite case $p_1 \vee p_2 \vee \ldots$, we need a lemma.

Lemma 74.28.
If $\{T_\alpha\}$ is a net of selfadjoint bounded operators such that

(1) $\sup_\alpha \|T_\alpha\| < \infty$,
(2) if $\alpha > \beta$, then $\langle v, T_\alpha v\rangle > \langle t, T_\beta v\rangle$ for every $v \in \mathscr{H}$.

Then $\{T_\alpha\}$ strongly converges to an operator T, see condition (60.63).

Proof. We want the limit T to fulfil $\langle v, Tv\rangle = \lim_{\alpha\to\infty}\langle v, T_\alpha v\rangle$. From weak-compactness of the unit ball, that formula defines the quadratic form $Q(v) = \lim_{\alpha\to\infty}\langle v, T_\alpha v\rangle$ with $\|v\| = 1$. The form Q is then defined on the whole space $\mathscr{H}$ by homogeneity: $Q(\lambda v) = \lambda^2 Q(v)$. The polarization identity

$$4\,\text{Re}\langle v, T_\alpha w\rangle = \langle (v+w), T_\alpha(v+w)\rangle - \langle (v-w), T_\alpha(v-w)\rangle$$

defines $Q(v, w)$ and then defines the value of $\langle v, Tw\rangle$ for every v and w. This is the candidate to be the strong limit of T_α. The question is now to know if this is an actual strong limit.

Since the net is increasing, we have $\langle v, (T - T_\alpha)v\rangle \geqslant 0$, so that $T - T_\alpha$ is positive which implies that $(T - T_\alpha)^{1/2}$ is well defined. We have

$$\|(T - T_\alpha)^{1/2} v\| = \langle v, (T - T_\alpha)v\rangle \to 0,$$

so that the strong limit of $(T - T_\alpha)^{1/2}$ is zero.

It is not true in general that, in the strong topology, $a_\alpha \to a$ and $b_\alpha \to b$ imply $a_\alpha b_\alpha \to ab$. But it is true when a_α and b_α are contained in a ball. Since $\sup \|T - T_\alpha\|^{1/2} < \infty$, we can thus make

$$\text{s}\lim(T - T_\alpha) = \text{s}\lim\big((T - T_\alpha)^{1/2}\big)^2 = 0. \quad \text{s}\lim_{a\to 0} f(a)$$

□

In short, that lemma claims that M contains the limits of all bounded increasing nets of selfadjoint operators. In particular, for a net of projections $P_\alpha \in M$ with $\alpha \in X$,

$$\bigvee_\alpha P_\alpha \in M, \qquad \bigwedge_\alpha P_\alpha \in M.$$

Proposition 74.29.
Let $\{P_\alpha\}$ and $\{Q_\alpha\}$ to be two nets of projectors on the same directed set, and suppose that we have operators $T_\alpha\colon \mathrm{Image}(P_\alpha) \to \mathrm{Image}(Q_\alpha)$ such that

(1) $\sup_\alpha \|T_\alpha\| < \infty$,

(2) $T_\alpha|_{\mathrm{Image}(P_\beta)} = T_\beta$,

then there exists an operator $T\colon \mathrm{Image}\left(\bigvee_\alpha P_\alpha\right) \to \mathrm{Image}\left(\bigvee_\alpha Q_\alpha\right)$ with

$$T|_{\mathrm{Image}(P_\alpha)} = T_\alpha.$$

If moreover P_α, Q_α and T_α belong to M, then $T \in M$.

Proof. No proof. □

¡¡ Avertissement/question à la lectrice !! 74.30
I think that $p_1 \vee p_2 \vee \dots$ is now well defined when p_i are projections. Indeed, the projections $P_k = p_1 \vee \dots p_k$ form a net and, by the proposition, the limit belongs to M. More generally, if A is a set of projections, for $\bigvee_{p\in A} p$, one considers the set of parts of A (which is partially ordered) and then the net $p_\alpha = \bigvee_{p\in\alpha} p$ where α is a part of A.

Now, one has to understand why the result is still a projection.

LemPTQnnzero

Lemma 74.31.
Let M be a factor. If P and Q are nonzero projections in M, then there is a $T \in M$ such that $PTQ \neq 0$.

Proof. Let $U(M)$ be the unitary group of M. We saw that if T commutes with $U(M)$, then it commutes with all M. Assume that $PTQ = 0$ for every $T \in M$, then in particular it holds for $T = U \in U(M)$ and we have $UPU^*Q = 0$.

Consider the operator $R = \bigvee_{U\in U(M)} UPU^*$ which is a projection in M. Since the set $\{UPU^*$ tel que $U \in U(M)\}$ is invariant under the adjoint action of $U(M)$, we have $URU^* = R$, or $UR = RU$ for every $U \in U(M)$. That proves that $R \in M'$. So $R \in M \cap M'$ and is thus a multiple of identity by the fact that M is a factor: $R = \mathrm{Id}$ (the multiple has to be 1 because R is a projection).

We said, on the other hand, that $UPU^*Q = 0$, which means that the range of UPU^* is orthogonal to the one of Q for every $U \in U(M)$, so that the range of R has to be orthogonal too. That contradicts the fact that R is a multiple of identity. □

LemVVPVVQfactreu

Lemma 74.32.
Let M be a factor and P, Q two nonzero projections in M, then there exists a nonzero partial isometry $V \in M$ such that $V^*V \leqslant P$ and $VV^* \leqslant Q$.

Proof. Let T be such that $PTQ \neq 0$ (by lemma 74.31), and V be the partial isometry part of QTP. The image of V^*V is at most the one of QTP which is smaller (or equal) to the image of Q, so $V^*V \leqslant Q$. For the same reason, the image of VV^* is contained in $P^*T^*Q^*\mathscr{H} = PY^*Q\mathscr{H} \subseteq \mathrm{Image}(P)$. Then $VV^* \leqslant P$. □

Proposition 74.33.
Let M be a factor and P, Q two nonzero projections in M, then there exists a nonzero partial isometry $V \in M$ such that

(1) $V^*V \leqslant P$,

(2) $VV^* \leqslant Q$

and either $V^*V = P$, or $VV^* = Q$ or both.

Proof. Let X be the set of partial isometries V such that $V^*V \leqslant P$ and $VV^* \leqslant Q$. We write $P_V = V^*V$ and $Q_V = VV^*$. We define a partial ordering on X by $V_1 \leqslant V_2$ if $P_{V_1} \leqslant P_{V_2}$, $Q_{V_1} \leqslant Q_{V_2}$ and $V_2 P_{V_1} = V_1$. In that case, V_2 is some kind of extension of V_1.

If one considers an increasing sequence $V_1 \leqslant V_2 \leqslant \dots$, the Zorn's lemma assures the existence of a V bigger than all the elements of the sequence. We claim that this V is the solution of the proposition. Indeed, suppose $V^*V \neq P$ and $VV^* \neq Q$. Then the operators $P' = P - V^*V$ and $Q' = Q - VV^*$ are nonzero projections. Now, lemma 74.32 provides a partial isometry which contradicts maximality of V. □

The **dimension** of the von Neumann algebra M is the set of equivalence class of projections in M. This is a linearly ordered set.

Let $M = M(G)$ where G is a group with only infinite conjugacy classes (but the one of identity). We saw that there is a trace $\tau\colon M \to \mathbb{C}$, and that $\forall t \in [0,1]$, there exists a projection whose trace is t.

Proposition 74.34.
If $T \geqslant 0$ is an element of M such that $\tau(T) = 0$, then $T = 0$.

Proof. We know the injection $M \to l^2(G)$ defined by $T \mapsto f_T = Tf_e$. Since the right translation commutes with every T_1 (because $T_1(gf_e) = g(T_1 f_e)$), we have $T_1 f_g = 0$ for every g if $T_1 f_e = 0$. On the other hand,

$$\tau(T_1^* T_1) := \langle f_e, T_1^* T_1 f_e \rangle = \|T f_e\|^2$$

which is positive. So if $\tau(T_1^* T_1) = 0$, then $T_1 f_e = 0$, and so $T_1 f_g = 0$ which proves that $T_1 = 0$. This concludes the proof that $T \geqslant 0$ and $\tau(T) = 0$ imply $T = 0$ because every positive element reads as a product $T_1^* T_1$. □

Now let define $\tau'\colon \dim M \to [0,1]$ by

$$\tau'\big([P]\big) = \tau(P). \tag{74.40}$$

It is well defined because if $P \sim Q$, then we have a u such that $P = uu^*$ and $Q = u^*u$, so that $\tau(P) = \tau(uu^*) = \tau(u^*u) = \tau(Q)$ by cyclic invariance of the trace. The map τ' preserves the order because $P - Q$ is positive when $P \geqslant Q$.

Proposition 74.35.
The map τ' is a bijection

Proof. What we have to prove is that $P \sim Q$ if $\tau(P) = \tau(Q)$. Let $V^*V = P$ and $VV^* \leqslant Q$ and compute

$$\tau(Q - VV^*) = \tau(Q) - \tau(VV^*) = \tau(Q) - \tau(V^*V) = \tau(Q) - \tau(P) = 0.$$

The fact that $\tau(Q - VV^*) = 0$ implies $Q - VV^* = 0$, which in turn proves that V implements the equivalence $P \sim Q$. □

74.8.1 Comparison of projections

In this section we follow [935]. Let M be a von Neumann algebra and p, q, two projectors. **Equivalence of projectors** is given by $p \sim q$ if there exists a partial isometry $u \in M$ such that $u^*u = p$ and $uu^* = q$. If p is equivalent to a subprojection of q, we write $p \prec q$.

Proposition 74.36.
If $p = \oplus_i p_i$ and $q = \oplus_i q_i$ are orthogonal sums of projectors with $p_i \sim q_i$, then $p \sim q_i$. If moreover $p_i \prec q_i$ for every i, then $p \prec q$.

Proof. No proof. □

Let $T \in \mathfrak{B}(\mathscr{H})$. The closed spaces $\overline{\text{Image}(T)}$ and $\overline{\text{Image}(T^*)}$ are the **left support** and **right support** of T.

Theorem 74.37.
Let T be an invertible operator, then it reads under the form

$$T = |T|U$$

where U is unitary. This decomposition is the unique one of the form $T = SU$ with S positive and U unitary.

Proof. Let $T = PU$, then $TT^* = PUU^*P = P^2$, then $P = \sqrt{TT^*}$ and then $U = P^{-1}T$. That proves unicity of the decomposition. For existence, we pose $U = (TT^*)^{-1/2}T$ and $P = (TT^*)^{1/2}$. Then we check

$$UU^* = (TT^*)^{-1/2}TT^*(TT^*)^{-1/2} = \mathbb{1}. \tag{74.41}$$

Now, the operator U is the product of two invertible operators, so that it is invertible. Thus the fact that $UU^* = \mathbb{1}$ forces $U^*U = \mathbb{1}$. □

When T is not invertible, we have a weaker result of decomposition.

PropdecmbTbV

Proposition 74.38.
Every $T \in \mathfrak{B}(M)$ can be written under the form

$$T = |T|V$$

where $V\colon \overline{\text{Image}(T^*)} \to \overline{\text{Image}(|T|)}$ is a partial isometry whose final projection is the support of $|T| = \sqrt{TT^*}$.

Proposition 74.39.
The left and right support of any operator in $\mathfrak{B}(\mathscr{H})$ are equivalent projections.

Proof. Let $T = |T|V$ be the decomposition of T by proposition 74.38. By definition, $e = V^*V$ is a projection and we have $(1-e)V^*V(1-e) = 0$ as can be checked by developing the expression and using $e^2 = e$. Since $e = e^*$, the latter equation rewrites

$$\big(V(1-e)\big)^*\big(V(1-e)\big) = 0,$$

so that $V(1-e) = 0$ and $Ve = V$ (we used lemma 60.12). Now look at $f = VV^*$, we have

$$fV = VV^*V = Ve = V,$$

so f is a projection such that $fV = V$. We deduce that $V\colon e\mathscr{H} \to f\mathscr{H}$ is an isometry. The projectors e and f are the **initial** and **final projections** of V. □

¡¡ Avertissement/question à la lectrice !! 74.40
This proof is not finished.

Using lemma 60.64, we have in particular

$$\overline{\text{Image}(T^*)} = (\ker T)^\perp = \big(\ker(T^*T)^{1/2}\big)^\perp = \text{Image}(T^*T)^{1/2}. \tag{74.42}$$

74.9 Type I factor and factorization

LemPMPPMPprime

Lemma 74.41.
Let M be a von Neumann algebra and $P \in M$, a projection. Then

$$(PM'P)' = PMP, \tag{74.43}$$

and

$$(PMP)' = PM'P \tag{74.44}$$

EqLemPMPPMPprimedeux

where we see PMP as an algebra of operators on $P\mathscr{H} \subseteq \mathscr{H}$.

Let us give an intuitive argument with matrices before to give a proof. For the first claim, consider M as the set of two by two matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In that case, $PMP = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ while $M' = \{\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}\}$ with $[\alpha, a] = 0$. Thus we have $PM'P = \begin{pmatrix} \alpha & 0 \\ ) & 0 \end{pmatrix}$.

For the second claim, suppose $Q \in (PNP)'$. We are looking for a $R \in M'$ such that $Q = PRP$. Notice that Q itself does not belong to M' in general. Take for example

$$M = \begin{pmatrix} \star & 0 & \star \\ 0 & \star & 0 \\ \star & 0 & \star \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}.$$

we have $M' = \begin{pmatrix} a & & \\ & b & \\ & & a \end{pmatrix}$, so that Q does not lie in M', but

$$Q = P \begin{pmatrix} 1 & & \\ & 0 & \\ & & 1 \end{pmatrix} P$$

anyway. We see on that example that the matrix R to be chosen is bigger than Q.

Proof of lemma 74.41. Let $Q \in (PMP)'$ and R be the projection onto the closure of $MQ\mathscr{H}$. The latter space being invariant under M, the projection R belong to M' by lemma 74.8. We are now going to prove that $Q = PRP$. Since $P = P^2$ commutes with R, the operator PRP is the projection onto the intersection of the target spaces of R and P, because $PRP = PR = RP$. Therefore Q is smaller than P, so that $MQ\mathscr{H} = MPQ\mathscr{H}$. That proves that PRP is the projection on

$$PMQ\mathscr{H} = PMPQ\mathscr{H} = QPMP\mathscr{H} = QP\mathscr{H} = Q\mathscr{H}$$

where we used the fact that $Q \in (PMP)'$. That concludes the proof that $Q = PRP$. □

LemMapipsomPMmP

Lemma 74.42.
If M is a factor and P, a nonzero projection in M, then the map

$$\begin{aligned} M' &\to PM'P \\ S &\mapsto PSP \end{aligned} \tag{74.45}$$

is a $*$-algebra isomorphism.

Proof. The fact that the map is multiplicative and surjective is clear. The only point to prove is injectivity. So we will prove that $S = 0$ under the assumption $SP = 0$. We have $0 = MSP = SMP$, so that S vanishes on $MP\mathscr{H}$.

The latter subspace is obviously invariant under M, but also under M' because for every $S_1 \in M'$, we have $S_1 MP\mathscr{H} = MPS_1\mathscr{H} \subseteq MP\mathscr{H}$. We conclude that the projection onto $MP\mathscr{H}$ lies in $M' \cap M$. Hence that projection must be a multiple of the identity or zero. Since that space contains at least $P\mathscr{H}$, the zero possibility is ruled out. Thus the projection onto $MP\mathscr{H}$ is the identity and $MP\mathscr{H} = \mathscr{H}$. Now, $SMP = 0$ implies $S = 0$ which concludes the proof of the lemma. □

A projection $P \in M$ is a **minimal projection** P if if $0 < Q \leqslant P$ implies $Q = P$. In other words, the projection P is minimal if $PMP = \mathbb{C}P$ because the projection onto PMP is of course smaller or equal to P. A minimal projection is always finite, indeed, a projection Q such that $Q \sim P$ and $Q \leqslant P$ (which exists when P is infinite) contradicts minimality of P.

Definition 74.43.
Let M be a factor. It is

— *of* ***type*** *I if M contains a non vanishing minimal projection,*
— *of* ***type*** *II if M contains a non vanishing finite projection (and is not of type I)*
— *of* ***type*** *III if no projection in M is finite (but the vanishing one).*

A type II factor is of **type** II_1 if it is finite and of **type** II_∞ if it is not finite.

PartypeIonavu Let us now examine what are the type I factors. Suppose that M is a factor of type I and $P \in M$ is a minimal projection. Then $\mathbb{C}P = PMP \subseteq \mathfrak{B}(P\mathscr{H})$, and formula (74.44) in lemma 74.41 makes $(PMP)' = PM'P = \mathfrak{B}(P\mathscr{H})$. Now the map of lemma 74.42 is an isomorphism, then

$$M' \simeq \mathfrak{B}(P\mathscr{H})$$

as $*$-algebra. Thus M' is a factor. This factor is moreover of type I because $\mathfrak{B}(P\mathscr{H})$ has minimal projections, namely projections on one dimensional spaces.

If we repeat the same argument with M' instead of M, we obtain that

EqMPHtypeIBh
$$M \simeq \mathfrak{B}(P'\mathscr{H}) \tag{74.46}$$

where P' is a minimal projection in M'.

What we proved is

Proposition 74.44.
If M is a factor of type I, there exists a separable Hilbert space $\mathscr{H}$ such that

$$M = \mathfrak{B}(\mathscr{H}). \tag{74.47}$$

When $M = \mathfrak{B}(\mathscr{H})$ is of type I says that M is of **type** I_n if $\dim \mathscr{H} = n$ where $1 \leqslant n \leqslant \infty$.

Proposition 74.45.
If M is of type I_n with $1 \leqslant n < \infty$, then $M \simeq \mathbb{M}_n(\mathbb{C})$ and the unique trace is the ususal matrix trace up to renormalization:

$$\mathrm{Tr}(T) = \frac{1}{n}\sum_{i=1}^{n} T_{ii}. \tag{74.48}$$

Proof. No proof. □

In particular we have that the composition of the two minimal projection $PP' = P'P$ is a rank one projection in $\mathfrak{B}(\mathscr{H})$. Now consider $Q = PP'$ and v be the unital vector in the target space of Q, i.e. $Qv = v$ and $\|v\| = 1$.

Proposition 74.46.
Let M be a factor of type I and P, a minimal projection. In the same way, let P' be a minimal projection in M'. Let $Q = PP'$ and $v \in Q\mathscr{H}$ such that $\|v\| = 1$. Then

$$\begin{aligned} M'Q &\to P\mathscr{H} \\ SQ &\mapsto Sv \end{aligned} \tag{74.49}$$

is an isomorphism.

Proof. Since Q projects on the space spanned by v, the fact that $S_1 v = S_2 v$ implies $S_1 = S_2$ and the injectivity is proved. For surjectivity, we know that, M being a factor of type I and P a minimal projection, we have $\mathbb{C}P = PMP \subseteq \mathfrak{B}(\mathscr{H})$ where the last equality has to be understood in the sense of that $T \in PMP$ restricts to an operator on $P\mathscr{H}$, and that this restrictions completely defines T. But as operators on $P\mathscr{H}$, we have $(PMP)' = PM'P$ (lemma 74.41), while lemma 74.42 assures that the latter algebra is M'. Thus $M' = \mathfrak{B}(P\mathscr{H})$ and $M'v = P\mathscr{H}$. □

Proposition 74.47.
In the same way, we have that

$$\begin{aligned} MQ &\to P'\mathscr{H} \\ TQ &\mapsto Tv \end{aligned} \tag{74.50}$$

is an isomorphism

Proof. No proof. □

Proposition 74.48.
Let M be a factor $S \in M$ and $T \in M'$, then if $ST = 0$, then $S = 0$ or $T = 0$.

Proof. Let P be the projection on the right support of S, i.e. onto $\text{Image}(S^*) = (\ker S)^\perp$. This operator is equal to the limit [1]

$$P = \lim_{n\to\infty} (S^*S)^{1/n}.$$

So P is a strong limit, and then belongs to M. Since, by assumption, $ST = 0$, we have $S^*ST = 0$, and then $PT = 0$. Since M is a factor, it implies that $P = 0$ or $T = 0$. □

Proposition 74.49.
Let $M \subseteq \mathfrak{B}(\mathscr{H})$ be a factor. Then the map

$$\begin{aligned} M \otimes_{\mathbb{C}} M' &\to \mathfrak{B}(\mathscr{H}) \\ S \otimes T &\mapsto ST \end{aligned} \tag{74.51}$$

is an injective $*$-homomorphism. Its image is strongly dense in $\mathfrak{B}(\mathscr{H})$.

Proof. First, as M is a factor, we have

$$\{ST \text{ tel que } S \in M \text{ and } T \in M'\}' = \mathbb{C}\,\mathrm{Id},$$

so that $\{ST\}'' = \mathfrak{B}(\mathscr{H})$, which proves that $\{ST\}$ is strongly dense in $\mathfrak{B}(\mathscr{H})$.

We have elements $S_i \in M$ and $T_i \in M'$ $(i = 1, \dots, n)$ such that

$$\sum_{i=1}^{n} S_i T_i = 0, \tag{74.52}$$

EqDecSTiMMprime

and we are going to prove that the set $\{S_i\}$ is not linearly independent, so that all the S_i, or all the T_i vanish. An element of $M \otimes M'$ can be decomposed under the form (74.52) in several ways.

Let Q be the projection on the closure of the space

$$\big\{(TT_1v, \dots TT_nv) \in \mathscr{H} \oplus \dots \oplus \mathscr{H} \text{ tel que } T \in M', v \in \mathscr{H}\big\}. \tag{74.53}$$

EqEspaceTTHHinv

This space is invariant under the algebra

$$M'^{(n)} = \left\{ \begin{pmatrix} T & & 0 \\ & \ddots & \\ 0 & & T \end{pmatrix} \text{ tel que } T \in M' \right\}, \tag{74.54}$$

so that $Q \in \left(M'^{(n)}\right)'$. Since $\{T_1Tv \text{ tel que } v \in \mathscr{H}, T \in M\} = \{T_1TSv \text{ tel que } v \in \mathscr{H}, T \in M, S \in M'\}$ the space (74.53) is also invariant under

$$\left\{ \begin{pmatrix} S & & 0 \\ & \ddots & \\ 0 & & S \end{pmatrix} \text{ tel que } S \in M \right\}, \tag{74.55}$$

EqSSSinvMBig

On the other hand, the operators which commute with all $M'^{(n)}$ are element of $\mathbb{M}_n(M)$, then

$$Q \in \left(M'^{(n)}\right)' = \mathbb{M}_n(n),$$

and for the same reason, the operators which commute with all (74.55) are elements of $\mathbb{M}_n(M')$, so that

$$Q \in \mathbb{M}_n(M \cap M') = \mathbb{M}_n(\mathbb{C}).$$

1. To be proven.

consider now the operator

$$R = \begin{pmatrix} S_1 & \dots & S_n \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}.$$

We havea

$$\begin{pmatrix} S_1 & \dots & S_n \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \cdot \begin{pmatrix} TT_1v \\ TT_2v \\ \vdots \\ TT_nv \end{pmatrix} = \begin{pmatrix} S_1TT_1v + \cdots + S_nTT_nv \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The first component is

$$T\big(\sum_i S_iT_iv\big)$$

which vanishes by assumption, so that $RQ = 0$. Let us write $Q = (q_{ij}\,\mathrm{Id})$ as element of $\mathbb{M}_n(\mathbb{C})$. The relation $RQ = 0$ says that

$$\sum_{i=1}^{n} S_iq_{ij} = 0 \tag{74.56}$$

for every j. Since Q is a non vanishing projector, at least one of these relations is nontrivial. That nontrivial relation shos that the S_i's are vanishing. If not, we have $Q = 0$, which means that $T_i = 0$. □

74.9.1 Tensor product of von Neumann algebras

Let $M_1 \subseteq \mathfrak{B}(\mathscr{H}_1)$ and $M_2 \subseteq \mathfrak{B}(\mathscr{H}_2)$ be two von Neumann algebras. We define the **tensor product**

$$M_1 \bar{\otimes} M_2 = \{S_1 \otimes S_2 \text{ tel que } S_i \in M_i\}'' \subseteq \mathfrak{B}(\mathscr{H}_1 \otimes \mathscr{H}_2) \tag{74.57}$$

74.10 Dimensions

74.10.1 Finite and infinite projections

A projection in a factor is **infinite** if it is equivalent to a proper subprojection; it is **finite** is it is not infinite. The notion of finite and infinite projections descents to the equivalence classes. A factor is **finite** when the identity is finite.

Proposition 74.50.
If P and Q are projections with $P \leqslant Q$ and Q is finite, then P is finite.

Proof. Assume that P is infinite, that is there exists a projection $P' < P$ with $P \sim P'$. So we have $P' < P \leqslant Q$ and there exists a $v \in \mathscr{H}$ such that $Qv = v$, $Pv = v$ and $P'v = 0$. Thus we have

$$Q = P + (Q - P) \sim P' + (Q - P) < Q,$$

the last inequality being assured by the vector v. What we proved is that Q is equivalent to a subprojection of itself, which contradicts the fact that Q is finite. □

Corollary 74.51.
Let α, $\beta \in \dim M$. If α is finite and β infinite, then $\alpha \leqslant \beta$.

If $\alpha = [P]$ and $\beta = [Q]$ with $PQ = 0$, we define $\alpha + \beta = [P + Q]$. We denote by $n\alpha$ the sum

$$n\alpha = \underbrace{\alpha + \cdots + \alpha}_{n \text{ times}}, \tag{74.58}$$

$n \in \mathbb{N}$. If $n\alpha$ exists for every positive integer n, we write

$$\infty\alpha = \bigvee_{n=1}^{\infty} n\alpha.$$

Proposition 74.52.
The sum $\alpha + \beta$ fulfils

(1) If $\alpha + \beta$ exists, then it is independent of the representative P and Q in α and β,

(2) It is commutative and associative: $\alpha + \beta = \beta + \alpha$ and $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$,

(3) zero is the neutral: $\alpha + [0] = \alpha$.

Proof. No proof. □

Lemma 74.53.
If α and β are finite, then $\alpha + \beta$ is finite when it exists.

Proof. No proof. □

Lemma 74.54.
If α is finite and β is infinite, then

(1) $\infty\alpha$ exists,

(2) $\infty\alpha = \beta$.

In particular, all the infinite projections are equivalent each other.

Proof. Since β is infinite and α is finite, the projection β contains a subprojection equivalent to α. What remains in β is still an infinite projection, and thus still contains a subprojection equivalent to α. Let $\alpha = [P]$ and $\beta = [Q]$. We have

$$Q = \big(\sum_n P_n\big) + R$$

where for each n, the projection P_n is equivalent to P, and R is a finite projection which does not contain a subprojection equivalent to P. Thus P contains a subprojection equivalent to R.

Notice that β can only contain a countable number of copies of α because one only has a countable number of basis vectors in a separable Hilbert space.

Let us consider the projection

$$\beta' = [Q - R] = [\sum P_n] = \infty\alpha. \tag{74.59}$$

One can construct a partial isometry implementing an equivalence between β and β' using an Hilbert hotel argument.

¡¡ Avertissement/question à la lectrice !! 74.55
This proof has to be finished.

□

LemVstarVPP

Lemma 74.56.
If P is a projection and if $V^*V = P$, then V vanishes on the target space of $P^\perp$.

Proof. It is evident that V^* does not vanish on the range of V (apart on zero) because

$$\langle v, V^*Vv\rangle = \langle Vv, Vv\rangle \neq 0.$$

Thus, in order V^*V to vanish on the range of $P^\perp$, one requires V to vanish on Image$(P^\perp)$. □

LemfinifactisemVUP

Lemma 74.57.
Let M be a finite factor, the every partial isometry V reads

$$V = UP \tag{74.60}$$

where U is unitary and P is a projection.

Proof. Consider the equivalent projections $P = V^*V$ and $Q = VV^*$ and their complement $P^\perp$, $Q^\perp$. First, we remark that $P^\perp$ and $Q^\perp$ are equivalent. Indeed, if they are not equivalent let say that $Q^\perp$ is equivalent to a subprojection of $P^\perp$, in this case, the identity $\mathbb{1} = P + P^\perp = Q + Q^\perp$ provides an equivalence between $\mathbb{1}$ and a subprojection of $\mathbb{1}$, which is in contradiction with the assumption that M is a finite factor.

Thus $P^\perp \sim Q^\perp$ and wet set $W^*W = P^\perp$ and $WW^* = Q^\perp$ and then $U = V + W$ is the answer because

$$(V + W)P = VP + WP,$$

but P is the projection on the target space of V, so $VP = V$ and, by lemma 74.56, we have $WP = 0$. □

Notice that the decomposition $V = UP$ is a special feature of the finite factor case. In the general infinite case, we have for example the operator which shifts the basis vectors

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

It has $S^*S = \mathbb{1}$ while $SS^* \neq \mathbb{1}$.

PropDecoTUTabsfinifacteur

Proposition 74.58.
Every element T of a finite factor reads

$$T = U|T| \tag{74.61}$$

EqDecoTUTabsfinifacteur

with U unitary.

Proof. One guess the form of U by the decomposition (74.61): $U = T(T^*T)^{-1/2}$. One checks that $U^*U = \mathbb{1}$. Since we are in a finite factor, the unit cannot be equivalent to something else than itself, so $UU^* = \mathbb{1}$ is forced, and U is thus unitary. □

Proposition 74.59.
If N is a finite factor, then so is $\mathbb{M}_2(N)$.

Proof. Let us suppose that we have a partial isometry $V \in \mathbb{M}_2(N)$ such that $VV^* = \mathbb{1}_2$. We have to prove that $V^*V = \mathbb{1}_2$; if not, the identity would be equivalent to a subprojection. In other words, we have to prove that V is invertible. It will be done by constructing an invertible operator W such that V^*W is invertible. In that case, V^* is invertible and so is V.

If we set $V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with a, b, c, $d \in N$, the relation $V^*V = \mathbb{1}_2$ imposes among others relations

SubEassEqaaccumVV

$$a^*a + c^*c = \mathrm{Id} \tag{74.62a}$$

ssEqaaccumVV

$$b^*a + d^*c = 0 \tag{74.62b}$$

ssEqbadczVV

Relation (74.62a) says that $|a|^2 + |c|^2 = 1$, so that in the sense of the continuous functional calculus, we have $|a| = \sqrt{1 - |c|^2}$. We deduce that $|a|$ is a limit of polynomials in $1 - |c|^2$ and that $\big[|a|, |c|\big] = 0$.

Since a and c belong to a finite factor, proposition 74.58 provide unitary elements u and v of N such that $a = u|a|$ and $c = v|c|$. Now we consider the unitary element

$$W = \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix} \begin{pmatrix} |a| & -|c| \\ |c| & |a| \end{pmatrix}$$

Using relations (74.62), we find

$$V^*W = \begin{pmatrix} 1 & -a^*u|c| + c^*v|a| \\ 0 & x \end{pmatrix},$$

but $a^*u|c| = |a|u^*u|c| = |a||c| = |c||a|$, so that the upper-right element is actually zero. We are left with

$$V^*W = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$$

in which we want to prove that x is invertible. Using the fact that W is unitary, we have $(V^*W)(V^*W)^* = V^*V = \mathbb{1}$, then

$$\begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & x^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & xx^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus we have $xx^* = \mathbb{1}$ in N, which in turn imposes $x^*u = \mathbb{1}$ because N is a finite factor. We have finished to prove that V^*W is invertible. □

Lemma 74.60.
If P and Q are projections with $PQ = 0$, then P is equivalent to a subprojection of Q, or Q is equivalent to a subprojection of P.

Proof. No proof. □

PropnnminSTperpssR

Proposition 74.61.
If R is not a minimal projection, then there exist subprojections S, T of R such that $S \sim T$ and $S \perp T$.

Proof. Since R is not minimal, we have a non vanishing projection $S < R$. Let consider $S_1 = R - S$. One checks that $SS_1 = 0$. Now S or S_1 is equivalent to a subprojection of the other. Let $S \sim T \leqslant S_1$. Now, we have $S \sim T$ and $ST = 0$ which means $S \perp T$. □

Let P and Q be two projection in M such that $PQ = 0$. In $\mathbb{M}_2(M)$, we have

$$\begin{pmatrix} P+Q & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}$$

where R is a subprojection of P which is equivalent to Q. Define the partial isometry V by $V^*V = Q$ and $VV^* = R$. We have

$$\begin{pmatrix} P & V^* \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & 0 \\ V & 0 \end{pmatrix} = \begin{pmatrix} P+Q & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\begin{pmatrix} P & 0 \\ V & 0 \end{pmatrix} \begin{pmatrix} P^* & V^* \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} P & PV^* \\ VP^* & R \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & R \end{pmatrix}$$

because $\text{Image}(V^*) = \text{Image}(Q)$ implies $PV^* = VP^* = VP = 0$. Now, $\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}$ is a subprojection of $\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}$ because R is a subprojection of P. Thus we have

$$\begin{pmatrix} P+Q & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} P & 0 \\ 0 & E \end{pmatrix} \leqslant \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} = \text{finite}. \tag{74.63}$$

So we deduce that $P+Q$ is finite.

A factor is **semifinite** if there are projections P_α such that $\bigvee P_\alpha = \mathrm{Id}$. Notice that if $\{P_1,\dots,P_n\}$ is a finite set of finite projections, then $\bigvee P_i$ is a finite projection. A factor is **purely infinite** if 0 is the only finite projection.

Proposition 74.62.
If P is any finite projection, one can find a sequence of equivalent projections $\{P_i\}$ such that $\mathbb{1} = \sum P_i$.

Proof. No proof. □

A consequence of that proposition is that any factor is semifinite or purely infinite, while a finite factor is always semifinite.

Remarque 74.63.
A factor of type I is semifinite because a minimal projection is finite. Indeed, when P is infinite, the projection Q such that $Q \sim P$ and $Q \leqslant P$ contradicts minimality of P.

Proposition 74.64.
A factor is semifinite if and only if it reads (finite)$\bar{\otimes}$(type I).

Proof. If a factor reads (finite)$\bar{\otimes}$(type I), the type I part is a $\mathfrak{B}(\mathscr{H})$ in which one can take as approximation of the identity the sequence of projections P_i given by

$$P_i e_j = \begin{cases} e_j & \text{if } j \leqslant i \\ 0 & \text{otherwise} \end{cases} \tag{74.64}$$

where $\{e_i\}$ is an orthonormal basis of $\mathscr{H}$.

Now if M is semifinite, let P_i be a sequence of finite projections with $P_iP_j = 0$ and $P_i \sim P_j$ such that $VP_i = \mathrm{Id}$. Then we have $\mathscr{H} = \oplus_i P_i\mathscr{H}$, but each of the P_i is equivalent to $P_1\mathscr{H}$, so that $\mathscr{H} = \oplus_i P_i\mathscr{H} = \oplus P_1\mathscr{H} = P_1\mathscr{H} \otimes l^2(\mathbb{N})$.

Under that isomorphism, an operator $T \in M$ acts on $P_1\mathscr{H}$ by P_1TP_1, so that

$$M = P_1MP_1\bar{\otimes}\mathfrak{B}\big(l^2(\mathbb{N})\big). \tag{74.65}$$

□

Thus, in order to understand the semifinite factors, it is sufficient to understand the finite factors.

74.10.2 Finite factor

LemfassminPPperp

Lemma 74.65.
Every factor without minimal projection contains a projection P such that $P \sim P^\perp$.

Proof. Consider the set Z of triples (P,Q,V) with

— P and Q are projections,

— $P \perp Q$,

— V is a partial isometry such that $V^*V = P$ and $VV^* = Q$.

The set Z is endowed with a partial order given by $(P_1,Q_1,V_1) \leqslant (P_2,Q_2,V_2)$ when $P_1 \leqslant P_2$, $Q_1 \leqslant Q_2$ and $V_2P_1 = V_1$. The Zorn's lemma provides a maximal element that we denote by (P,Q,V). Let us suppose that $P+Q < \mathbb{1}$ (strictly). Then we write $R = \mathbb{1} - P - Q$ and by proposition 74.61 we have subprojection S and T of R such that $S \sim T$ and $S \perp T$.

We put $W^*W = S$ and $WW^* = T$, then $(P+S, Q+T, V+W)$ contradicts maximality of (P,Q,V), and $P+Q = \mathrm{Id}$ with $P \perp Q$ and $P \sim Q$. □

LemfinfacPPQQPsimQ

Lemma 74.66.
Let M be a finite factor. If $P \sim P^\perp$ and $Q \sim Q^\perp$, then $P \sim Q$.

This lemma means that there is only one way tu cut $\mathscr{H}$ in two equal subspaces.

Proof. Let us suppose that P is not equivalent to Q, thus it is equivalent to a subprojection of Q.

¡¡ Avertissement/question à la lectrice !! 74.67
What is the exact statement which says that a projection is equivalent to a subprojection of the other, or the contrary?

Since $P \sim sub\, Q$, we have $P^\perp \sim sub\, Q^\perp$ and $P + P^\perp \sim sub(Q + Q^\perp)$. That means that Id is equivalent to a subprojection of itself, which is impossible because of the assumption of finite factor. We deduce that $P \sim Q$. □

74.10.3 Rational and real dimensions

SubSecRationalRealDim

We consider a factor of finite type M without minimal projection. Lemma 74.65 ensures the existence of a projection P such that $P \sim P^\perp$ while lemma 74.66 says that every such projection lie in the class $[P]$. We define $\frac{1}{2} \in \dim M$ by

$$\frac{1}{2} = [P] \in \dim M \tag{74.66}$$

where $P \sim P^\perp$. Now we look at the algebra PMP, which posses its own $\frac{1}{2}$ that is called $\frac{1}{4}$ in M. So we set the definition

$$\frac{1}{4} = [Q] \in \dim M \tag{74.67}$$

where $Q \leqslant P$ and $Q \sim P - Q$. Notice that $P - Q = Q^\perp$ in $P\mathscr{H}$. In the same way, we define $\frac{3}{4}$ by

$$\frac{3}{4} = [P + R] \in \dim M \tag{74.68}$$

with $R \leqslant R^\perp$ and $R \sim P^\perp - R$. Continuing the process, we can define

$$\frac{r}{2^n} \text{ for } r \in \{0, \dots, 2^n\},$$

and we define $\tau\colon \dim M \to [0,1]$ by

$$\tau(P) = \sup\left\{ \frac{r}{2^n} \text{ tel que } \frac{r}{2^n} \leqslant [P] \right\} \in [0,1]. \tag{74.69}$$

ThobijzudimM

Theorem 74.68.
The map $\tau\colon \dim M \to [0,1]$ is a bijection for every finite factor without minimal projection.

Proof. No proof. □

74.10.4 Summary

When M is a finite factor, it has two possibilities: it contains or not a minimal projection.

74.10.4.1 Finite factor with minimal projection

We saw at page 3840, and more precisely on equation (74.46) that if P' is the minimal projection of M', then $M \simeq \mathfrak{B}(\mathscr{H})$, so that the possible dimensions are

$$\dim M = \{0, 1, \dots, N\} \tag{74.70}$$

where N can eventually take the value ∞, as shows the example $M = \mathbb{C}\,\mathrm{Id}$.

74.10.4.2 Finite factor without minimal projection

In the case of finite factor without minimal projection, we saw by theorem 74.68 that

$$\dim M = [0,1]. \tag{74.71}$$

74.10.4.3 Semifinite factor

If the factor is not finite, there are two possibilities: it is semifinite or purely infinite. In the semifinite case

74.10.5 Purely infinite factor

One still has to see what are the dimensions for a purely infinite factor.

LemORSneqzeroini

Lemma 74.69.
Let P be an infinite projection in a factor and $P = R \oplus S$ (orthogonal sum) with $R \sim P$ and $S \neq 0$. Then $[P] = \infty[S]$.

Proof. Let $R \sim P$ and $P = R + S$ with $S \neq 0$. Since $R \sim P$, we can decompose R in the same way as P is decomposed, namely

$$R = R_1 + S_1$$

with $R_1 \sim R \sim P$ and $S_1 \sim S$. Explicitly, if $V^*V = P$ and $VV^* = R$, we put $R_1 = VRV^*$ and $S_1 = VSV^*$. Proceeding, we decompose $R_1 = R_2 + S_2$, $R_2 = R_3 + S_3$, and so on. Since $R \perp S$, we have $S_1 \perp R_1$, and S_2 being a subprojection of R_1 which is orthonormal to S_1, we have $S_1 \perp S_2$.

What we get is finally is a set $\mathcal{S} = \{S, S_i\}_{i\in\mathbb{N}}$ of projections two by two orthogonal and equivalent. Let $\mathcal{Z}$ be the set of sets of projections two by two orthogonal and equivalent which contains $\mathcal{S}$. This is the set of extensions of $\mathcal{S}$. The Zorn's lemma provides a maximal element in $\mathcal{Z}$ that we name $\mathcal{Q} = \{Q_n\}$. By very definition, $\mathcal{S} \subseteq \mathcal{Q}$.

We want to prove that $\sum_n Q_n = P$, so that $P \leqslant \infty[S]$ because $\infty[S] \leqslant \sum_n Q_n$. Let $E = P - \sum_n Q_n$. By maximality, it does not contain a subprojection equivalent to S, so we have $E \preceq S$. We are going to prove that $E + \sum_n Q_n \sim \sum Q_n$. The argument is once again an Hilbert hotel construction.

Let $Q_i = A_i + E_i$ with $A_i \sim A_j$ and $E_i \sim E$. That decomposition is nothing else that the fact that E is equivalent to a subprojection of Q_i. Now consider the following equivalences:

$$\begin{aligned} E &\sim E_1 \\ A_1 &\sim A_1 \\ E_1 &\sim E_2 \\ A_2 &\sim A_2 \\ E_2 &\sim E_3 \\ &\vdots \end{aligned}$$

On the left hand side we have $E + \sum_n Q_n$ while on the right hand side we have $\sum_n Q_n$. Thus we have $E + \sum_n Q_n \sim \sum_n Q_n$ and then $P \sim \sum_n Q_n$. □

Corollary 74.70.
If M is a purely infinite factor, then $\dim M = \{0, \infty\}$

Proof. In a purely infinite factor, zero is the only finite projection, thus any projection reads $P = R + S$ with $P \sim R$ and $S \neq 0$. The aim is to prove that $[P] = \mathrm{Id}$, in such a way that there exists only one infinite.

We have $\mathrm{Id} = Q + S$ for a certain $Q \sim \mathrm{Id}$. The S here can be chosen the same as the S of P. Indeed $\mathrm{Id} = P^\perp + P = P^\perp + R + S$. Using the equivalences $P^\perp \sim P^\perp$ and $P \sim R$, we see that $\mathrm{Id} = P^\perp + P \sim P^\perp + R$, so that $\mathrm{Id} = Q + S$ with $Q = P^\perp + R$. Using lemma 74.69, we

have $[P] = \infty[S]$, and then using the lemma again on $[P]$, we get $[P] = \infty[S]$, which proves that $[P] = [\mathrm{Id}]$. □

74.11 Tracial functional

SecTracevonNeuman

References: [986, 1000]. This section is related to section 72.18 about trace on C^*-algebra.

A linear functional $\tau : M \to \mathbb{C}$ is

a state when $\tau(A^*A) \geqslant 0$ and $\tau(1) = 1$;

tracial when $\tau(A^*A) = \tau(AA^*)$;

faithful when $\tau(A^*A) = 0$ if and only if $A = 0$;

normal when if $A = \sum_n A_n$ exists for a sequence $\{A_n\}$ of positive operators, then $\tau(A) = \sum_n \tau(A_n)$;

semifinite if for every non zero $A \in M^+$ majorizes some non zero $B \geqslant 0$ with $\tau(B) < \infty$.

Lemma 74.71.
A linear tracial functional on a von Neumann algebra is a trace.

Proof. Let U be unitary, using the tracial property we find $\tau(U^*A^*AU) = \tau(AA^*)$, so that

$$\tau(U^*BU) = \tau(B) \tag{74.72}$$

EqtauUBUtauB

for every positive B. Since any element of M reads as composition of four positive, the relation (74.72) holds in fact for every B in M. Since every element $C \in M$ can be written as $C = U^*B$, we have

$$\tau(CU) = \tau(U^*BU) = \tau(B) = \tau(UC),$$

and we have $\tau(CU) = \tau(UC)$ for every C and U unitary. Now, every element is a combination of four unitary, so that $\tau(CD) = \tau(DC)$ for every C and D, which means that τ is a trace. □

Notice that the existence of a faithful normal trace on M implies that M is finite because if $V^*V = 1$,

$$\tau(1 - VV^*) = 1 - \tau(VV^*) = 1 - \tau(V^*V) = 1 - \tau(1) = 0,$$

which, from faithfulness, implies that $VV^* = 1$.

Let $\epsilon > 0$, an **ϵ-trace** is a normal state $\varphi : M \to \mathbb{C}$ such that

$$\varphi(A^*A) \leqslant (1 + \epsilon)\varphi(AA^*). \tag{74.73}$$

Lemma 74.72.
A normal state on a finite factor is an ϵ-state if and only if $P \sim Q$ implies $\varphi(P) \leqslant (1 + \epsilon)\varphi(Q)$.

Proof. No proof. □

Lemma 74.73.
If φ is an ϵ-trace, then we have

$$\frac{1}{(1 + \epsilon)}\varphi(P) \leqslant \frac{\dim(P)}{\dim(\mathbb{1})} \leqslant (1 + \epsilon)\varphi(P) \tag{74.74}$$

for every projection P.

Proof. Let us first suppose that

$$\dim(P) = \dim(\mathbb{1})/n, \tag{74.75}$$

EqdimPsimmmtun

and decompose $\mathbb{1} = P_1 + \cdots + P_n$ with $P_iP_j = 0$ $(i \neq j)$ and $P_i \sim P_j$ for every i. So we have $V_i^*V_i = P$ and $V_iV_i^* = P_i$. In that case

$$\begin{aligned} n\varphi(P) = \varphi(V_1^*V_1 + \cdots + V_n^*V_n) &\leqslant (1 + \epsilon)\varphi(V_1V_1^* + \cdots + V_nV_n^*) \\ &= (1 + \epsilon)\varphi(1) = 1 + \epsilon, \end{aligned} \tag{74.76}$$

thus

$$\varphi(P) \leqslant \frac{1+\epsilon}{n} = (1+\epsilon)\frac{\dim(P)}{\dim(\mathbb{1})}. \tag{74.77}$$

The other sense is obtained by the same computation:

$$1 = \varphi(V_1V_1^* + \cdots + V_nV_n^*) \leqslant (1+\epsilon)\varphi(V_1^*V_1 + \cdots + V_n^*V_n) = (1+\epsilon)n\varphi(P).$$

The lemma is now proved with the assumption (74.75). If that assumption does not hold, we can build projections P_i with $P_iP_j = 0$, $\dim(P_i) = \dim(\mathbb{1})/n_i$ and $P = \sum_i P_i$. Then we can use the fact that φ is normal and carry on the same computation as before. □

Corollary 74.74.
If φ_1 is an ϵ_1-trace and φ_2 is an ϵ_2-trace, then

$$\|\varphi_1 - \varphi_2\| \to 0$$

when $\epsilon_1, \epsilon_2 \to 0$.

Proof. No proof. □

The interest of this lemma resides in the fact that if φ_k is an ϵ_k-trace with $\epsilon_k \to 0$, we have a Cauchy sequence, so that $\lim_{k\to\infty}\varphi_k$ is an actual trace. We are thus not obliged to build traces, but ϵ-trace are sufficient.

LemPMPnormalfaithstate

Lemma 74.75.
Let φ be a faithful normal state. If $\epsilon > 0$, there exists a non vanishing projection $P \in M$ such that

$$\varphi(A^*A) \leqslant (1+\epsilon)\varphi(AA^*)$$

if $A \in PMP$. In other words, φ is an ϵ-trace on PMP.

Proof. Let $Q = \sum_\alpha Q_\alpha$ and $Q' = \sum_\alpha Q'_\alpha$ where $\{(Q_\alpha, Q'_\alpha)\}$ is a maximal set of projections pairs such that

(1) $Q_\alpha \perp Q_\beta$ and $Q'_\alpha \perp Q'_\beta$ if $\alpha \neq \beta$,
(2) $Q_\alpha \sim Q'_\alpha$,
(3) $\varphi(Q_\alpha) < \varphi(Q'_\alpha)$.

From normality of φ, the third condition implies that $\varphi(Q) = \varphi(Q')$. Since $Q \neq \mathbb{1}$, we have $\varphi(Q') \leqslant 1$ and thus

$$\varphi(Q) < \varphi(Q') \leqslant 1.$$

Let us now choose a pair (R, R') such that $R \perp Q$, $R' \perp Q'$ and $R \sim R'$. By maximality, that pair cannot fulfil the third condition of $\{Q_\alpha, Q'_\alpha\}$:

$$\varphi(R) \geqslant \varphi(R').$$

Now we consider the number

$$m = \min\{k \text{ tel que } k\varphi(R) \geqslant \varphi(R') \text{ for every } (R, R')\}. \tag{74.78}$$

We just saw that 1 belongs to the set, so $m \leqslant 1$. Let us suppose that $m = 0$. One typical projection orthogonal to R is $\mathbb{1} - R$, but $\varphi(\mathbb{1} - R) \neq 0$ because φ is faithful and $\mathbb{1} - R \neq \mathbb{1}$. So we conclude that m cannot be zero and

$$m \in]0, 1].$$

Now, take $\epsilon > 0$ and choose a pair (R, R') which fulfils the two conditions and such that

EqcondpaireRRprime
$$m\varphi(R) < (1+\epsilon)\varphi(R'). \tag{74.79}$$

Existence of such a pair is ensured by minimality of m.

We build the projections $S = \sum_\alpha S_\alpha$ and $S' = \sum_\alpha S'_\alpha$ where $\{S_\alpha, S'_\alpha\}$ is maximal for the properties

(1) $S_\alpha \leqslant R$ and $S'_\alpha \leqslant R'$,
(2) $S_\alpha \perp S_\beta$, $S'_\alpha \perp S'_\beta$ if $\alpha \neq \beta$,
(3) $S_\alpha \sim S'_\alpha$,
(4) $m\varphi(S_\alpha) \geqslant (1+\epsilon)\varphi(S'_\alpha)$.

Using normality, the last condition sums to

$$m\varphi(S) \geqslant (1+\epsilon)\varphi(S'). \tag{74.80}$$

We also have $S \sim S'$ and $S \neq R$. Indeed, suppose that $S = R$. Then $S \sim R$, but $S \sim S'$, so $R \sim R'$. Since $S'_\alpha \leqslant R$ for every α, we have $S' \leqslant R'$. We conclude that $S' = R'$ because M is a finite factor: in this case, $S' \leqslant R'$ and $S' \sim R'$ imply $S' = R'$. Now the fourth condition over S_α becomes $m\varphi(R) \geqslant (1+\epsilon)\varphi(R')$ which contradicts the condition (74.79) that we made on the pair (R, R'). Thus we conclude that $S \neq R$.

Let take $P = R - S$ and choose projections E and E' such that $E \leqslant P$, $E' \leqslant P'$ with $E \sim E'$. In this case we have

$$m\varphi(E) < (1+\epsilon)\varphi(E'),$$

if not (E, E') contradicts maximality of $\{S_\alpha, S'_\alpha\}$. On the other hand, minimality of m provides $m\varphi(E) \geqslant \varphi(E')$, so that we are left with

$$\varphi(E') \leqslant m\varphi(E) < (1+\epsilon)\varphi(E'), \tag{74.81}$$

which shows that φ is an ϵ-trace and that P' is the answer to the lemma.

□

An example of faithful normal state is

$$\varphi(T) = \sum_n 2^{-n}\langle v_n, Tv_n\rangle \tag{74.82}$$

where $\{v_i\}$ is an orthonormal basis of $\mathscr{H}$.

Proposition 74.76.
There is an ϵ-trace on M.

Proof. We just saw that normal states do exist, so we apply lemma 74.75 and we find a projection P for which we suppose $\dim(P) = 1/n$, and we choose $P_1, \dots P_n$ such that $P_iP_j = 0$ $(i \neq j)$ and $P_i \sim P$. We define the partial isometries V_i by

$$V_i^*V_i = P, \qquad\qquad V_iV_i^* = P_i.$$

Now we define

$$\psi(A) = \sum_i \varphi(V_i^*AV_i)$$

where φ is the normal faithful state on PMP. Thus we have

$$\psi(A) = \sum \varphi(V_i^*AV_i) = \sum \psi(V_i^*AV_i) = \sum \psi(AV_iV_i^*) = \sum \psi(AP_i),$$

and we find

$$\begin{aligned}
\psi(A^*A) &= \sum_i \varphi(V_i^*A^*AV_i) \\
&= \sum_{ij} \varphi(\underbrace{V_i^*A^*V_j}_{\in PMP}\underbrace{V_j^*AV_i}_{\in PMP}) \\
&\leqslant (1+\epsilon)\sum_{ij} \varphi(V_j^*AV_iV_i^*A^*V_j) \\
&\leqslant (1+\epsilon)\psi(AA^*),
\end{aligned} \tag{74.83}$$

which means that ψ is a trace. □

PropExistenceTrace

Proposition 74.77.
Every finite factor accepts an unique faithful, normal and tracial state. Moreover, if P and Q are projections and τ is the trace, then

$$P \sim Q \Leftrightarrow \tau(P) = \tau(Q). \tag{74.84}$$

PropFactIIunttedim

Proposition 74.78.
If M is a factor of type II_1, then

$$\{\tau(P) \text{ tel que } P \in \mathscr{P}(M)\} = [0,1]. \tag{74.85}$$

where τ is the trace given by proposition 74.77.

Proof. No proof. □

74.12 Modules over von Neumann algebras

SecOverModVNalgDim

Let X be a topological space and denote by π the group $\pi_1(X)$. We denote by $\tilde{X}$ the universal covering of X, and the group π acts on $\tilde{X}$ by the monodromy action defined in subsection 48.10.2. We denote by $C_p(X)$ the space of p-chains over X and $C_p(\tilde{X})$ the one of $\tilde{X}$ from which we build the complexes

$$\ldots \xrightarrow{b} C_p(X) \xrightarrow{b} C_{p-1}(X) \xrightarrow{b} \ldots$$

and

$$\ldots \xrightarrow{b} C_p(\tilde{X}) \xrightarrow{b} C_{p-1}(\tilde{X}) \xrightarrow{b} \ldots$$

with $b \circ b = 0$. The chain space $C_p(\tilde{X})$ is not only a group, but a module over $\mathbb{Z}[\pi]$.

We are now going to look at the tensor product von Neumann algebra

$$M_p(\tilde{X}) = M(\pi) \otimes_{\mathbb{Z}[\pi]} C_p(\tilde{X}). \tag{74.86}$$

In the right hand side, $M(\pi)$ is the von Neumann algebra described in subsection 74.7.5, which is generated by the operators $\mathcal{U}_g \in \mathfrak{B}\big(l^2(\pi)\big)$ defined by $\mathcal{U}_g(\varphi)(g) = \varphi(h^{-1}g)$, and the tensor product is the one defined in equation (54.12).

We can then look at the complex of $M(\pi)$-modules

$$\ldots \xrightarrow{b} M_{p+1}(\tilde{X}) \xrightarrow{b} M_p(\tilde{X}) \xrightarrow{b} M_{p-1}(\tilde{X}) \xrightarrow{b} \ldots$$

and look at the corresponding Betty numbers which are integers numbers associated with any topological space.

74.12.1 Modular conjugation and factor of type II_1

The results presented here partially come from [935], section 9 and [1001], proposition 2.2.6.

A **module** over a von Neumann algebra M is an Hilbert space[2] $\mathcal{E}$ for which there exists a weakly continuous map $\pi_r : M \to \mathfrak{B}(\mathcal{E})$ such that $\pi_r(T) = \big(\pi_r(T)\big)^*$ and $\pi_r(ST) = \pi_r(S)\pi_r(T)$. If $\mathcal{E}$ and $\mathcal{F}$ are two modules over M, a map $a : \mathcal{E} \to \mathcal{F}$ is **linear** if

$$a(S\xi) = Sa(\xi) \tag{74.87}$$

for every $S \in M$ and $\xi \in \mathcal{E}$. We denote by[3] $\mathscr{L}_M(\mathcal{E}, \mathcal{F})$ the set of M-linear maps from $\mathcal{E}$ to $\mathcal{F}$. When $\mathcal{E} = \mathcal{F}$, we write $\mathscr{L}_M(\mathcal{E})$. By definition of linearity, we have $\mathscr{L}_M(\mathcal{E}) = \pi(M)'$. The same

2. Hilbert space are generally denoted by $\mathscr{H}$, while modules are denoted by $\mathcal{E}$ or $\mathcal{F}$. Here, the notation will depend on what aspect we are focusing on.

3. This is what [1001] denote by ${}_M\mathcal{L}(\mathscr{H})$

notion exists for linearity at right, and in the general case, we denote by $\mathscr{L}_{M,N}(\mathcal{E},\mathcal{F})$, the set of M-left-linear and N-right-linear maps from $\mathcal{E}$ to $\mathcal{F}$.

We say that the modules $\mathcal{E}$ and $\mathcal{F}$ are **isomorphic** if there exists an unitary M-linear map between them.

When M is a type II_1 factor, we denote $\mathscr{H}_1 = L^2(M)$ and $\mathscr{H}_\infty = \mathscr{H}_1 \otimes l^2$.

PropTypeIIProjhInfty

Proposition 74.79.
Let M be a type II_1-factor and $\mathcal{E}$ be a separable M-module. There exists a projection $P \in \mathbb{M}_\infty(M)$ such that $\mathcal{E} \simeq \mathscr{H}_\infty P$. All such projections are Murray-von Neumann equivalent.

Proof. No proof, explanations can be found in [1001], theorem 2.2.2. □

When $\mathcal{E}$ is a separable M-module (M is a type II_1 factor), we define

$$\dim_M \mathcal{E} = \operatorname{Tr} p \tag{74.88}$$

where $p \in \mathscr{P}\big(\mathbb{M}_\infty(M)\big)$ is the projection such that $\big(L^2(M)\otimes l^2\big)p \simeq \mathcal{E}$ and whose existence is given by proposition 74.79.

Let M be a factor of type II_1 with the unique trace τ given by proposition 74.77. We denote by $L^2(M,\tau)$ the Hilbert space of its GNS representation, and π_τ the representation. Since τ is unique, we will simply denote them by $L^2(M)$ and π. We denote by $\Omega \in L^2(M)$ a cyclic vector. If $\pi(T) = 0$, we have

$$\tau(T^*T) = \|\pi(T)\Omega\|^2 = 0,$$

so that $T = 0$ and the representation is faithful. We can thus identify $T \in M$ with $\pi(T) \in \mathfrak{B}\big(L^2(M)\big)$ and we have $M \subset \mathfrak{B}\big(L^2(M)\big)$. The basics properties of the GNS construction say also that $\overline{M\Omega} = L^2(M)$ and $\tau(T) = \langle T\Omega, \Omega\rangle$ for every T in M.

If $T\Omega = 0$, then $0 = \|T\Omega\|^2 = \tau(T^*T)$, so that $T = 0$. That means that Ω is separating for M.

Lemma 74.80.
If $M \subseteq \mathfrak{B}(\mathscr{H})$ is a von Neumann algebra, a vector in $\mathscr{H}$ is cyclic for M if and only if it is separating for M'.

Proof. A proof is given in [1001]. □

Now, for every $\xi \in L^2(M)$, we define the operators π_l and π_r by

$$\pi_l(\xi)(T'\Omega) = T'\xi \tag{74.89a}$$
$$\pi_r(\xi)(T\Omega) = T\xi \tag{74.89b}$$

for $T' \in M'$ and $T \in M$. The domains are $\operatorname{dom}\big(\pi_l(\xi)\big) = M'\Omega$ and $\operatorname{dom}\big(\pi_r(\xi)\big) = M\Omega$. These operations are well defined because Ω is cyclic and separating for M and M'.

If $\pi_l(\xi)$ extends to a bounded operator on $L^2(M)$, one says that ξ is **left bounded**., and the (necessarily unique) extension is still denoted by $\pi_l(\xi)$. We do the same for π_r. Let now the map

$$\begin{aligned} J\colon M\Omega &\to M\Omega \\ T\Omega &\mapsto T^*\Omega, \end{aligned} \tag{74.90}$$

this is a conjugate linear isometry, so that it extends to an anti-unitary involution.

$$J\colon L^2(M) \to L^2(M) \tag{74.91}$$

That map J is the **modular conjugation** operator for the factor M of type II_1.

The algebra JMJ acts on $L^2(M)$ as well as on $L^2(JMJ)$. If $\xi \in L^2(M)$, the action is

$$(JSJ)\xi = \xi S^*. \tag{74.92}$$

EqActJMJLdM

That one is inspired by the fact that $JSJT = J(ST^*) = TS^*$. Since $J^2 = 1$, the action of JMJ on $L^2(JMJ)$ is given by

$$(JSJ)(JTJ) = J(ST)J. \tag{74.93}$$ EqActJMJLdJMJ

Let M be a type II_1 factor. We denote $\mathscr{H}_1(M) = L^2(M)$, or simply by $\mathscr{H}_1$ when no confusion is possible, and $\mathscr{H}_\infty = L^2(M) \otimes l^2$. The algebra M acts on $\mathscr{H}_1$ by

$$\pi_1(S)\xi = S\xi, \tag{74.94}$$

and M acts on $\mathscr{H}_\infty$ by

$$\pi_\infty(S) = \pi_1(s) \otimes \mathrm{Id}_{l^2}\,. \tag{74.95}$$

We also note $\mathbb{M}_\infty(M) = M \otimes \mathfrak{B}(l^2)$, that has to be understood as infinite matrices with elements in M. It is a type II_∞ factor, and the trace is a follows. If $p \in \mathbb{M}_\infty(M) = (p_{ij})$, then

$$\mathrm{Tr}(p) = \sum_{i=1}^{\infty} \mathrm{Tr}_M(p_{ii}). \tag{74.96}$$ EqTraceMinfinuM

With that formula, when $q \in \mathscr{P}(l^2)$ has rank 1, then $\mathrm{Tr}(1_m \otimes q) = 1$. Let $q \in \mathbb{M}_n(M)$ and consider the element of $\mathbb{M}_\infty(M)$ given by the following:

$$\begin{pmatrix} q & 0 \\ 0 & 0 \end{pmatrix}. \tag{74.97}$$

It's trace, in $\mathbb{M}_\infty(M)$, is $\sum_{i=1}^n \mathrm{Tr}(q_{ii})$, but the trace of q in $\mathbb{M}_n(M)$ is $\frac{1}{n}\sum_{i-1}^n \mathrm{Tr}(q_{ii})$. The normalization is not the same in $\mathbb{M}_n(M)$ and in $\mathbb{M}_\infty(M)$.

We denote by e_{11}, the element of $\mathbb{M}_n(M)$ given by

$$e_{11} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}. \tag{74.98}$$

We have $\mathrm{Tr}(e_{11}) = 1/n$, and $\pi_e(e_{11}) \in \mathscr{P}\Big(\pi_r\big(\mathbb{M}_n(M)\big)\Big)$.

Lemma 74.81.
The $\mathscr{L}_M\big(L^2(M)\big)$-modules $\mathscr{L}_M\big(L^2(M)\big)$ and $L^2(M)$ are isomorphic.

Proof. The result comes from the fact that an element of $\mathscr{L}_M\big(L^2(M)\big)$ is uniquely defined by its value at $1 \in M$. □

As a consequence of that lemma, we have

$$\dim_{\mathscr{L}_M\left(L^2(M)\right)} L^2(M) = 1, \tag{74.99}$$

because

$$\Big(\mathscr{L}_M\big(L^2(M)\big) \otimes l^2\Big)p = L^2(M) \tag{74.100}$$

when $p = \mathrm{Id} \otimes e_{11}$, whose trace is [4] 1.

LemLJMJequalLM

Lemma 74.82.
We have the isomorphism

$$L^2(JMJ) \simeq L^2(M) \tag{74.101}$$

as JMJ-module.

4. it is not $1/n$ or anything like that because of the remark we did bellow the definition (74.96).

Proof. Let us prove that the map which extend the following is an isomorphism:

$$\begin{aligned} \psi\colon L^2(JMJ) &\to L^2(M) \\ JSJ &\mapsto S^* \end{aligned} \tag{74.102}$$

So we have to prove that for every $a \in JMJ$ and $\xi \in L^2(JMJ)$, we have $\psi(a\xi) = a\psi(\xi)$. Using the actions (74.92) and (74.93), if $a = JSJ$ and $\xi = JTJ$, we find

$$\psi(a\xi) = \psi(JSJJTJ) = \psi(JSTJ) = T^*S^*, \tag{74.103}$$

while

$$a\psi(\xi) = (JSJ)\psi(JTJ) = (JSJ)T^* = T^*S^*. \tag{74.104}$$

That prove that $L^2(M) \simeq L^2(JMJ)$ as JMJ-modules. □

Lemma 74.83.
We have

$$\mathscr{L}_M\big(L^2(M)\big) = \mathscr{L}_{M,\pi_r(e_{11})}(\mathcal{E}_n) \tag{74.105}$$

EqoLLdpireunnmodE

where $\mathcal{E}_n = L^2(M) \oplus \cdots \oplus L^2(M)$ *(n terms).*

Proof. An element in the right hand side of (74.105) is a right-linear map with respect to $\pi_r(e_{11})$, i.e. a map $f\colon \mathcal{E} \to \mathcal{E}$,

$$f\big((\xi_1, \dots, \xi_n)\pi_r(e_{11})\big) = f(\xi_1, \dots, \xi_n)\pi_r(e_{11}). \tag{74.106}$$

Since the left hand side only depends on ξ_1, the right hand side shows that $f(\xi_1, \cdots, \xi_n)$ only depends on ξ_1. Now, the right hand side takes its values in $L^2(M)$, so that the left hand side shows that f takes its values in $L^2(M)$. □

The following is the proposition 2.2.6 in [1001].

PropDimIIun

Proposition 74.84.
Let $\mathscr{H}$ be a separable Hilbert space and $M \subseteq \mathfrak{B}(\mathscr{H})$, a factor of type II_1. We have

ItemiPropDimIIun

(1) for every $d \in [0,1]$, there exists a M-module $\mathcal{E}_d$ such that $\dim_M \mathcal{E}_d = d$.

ItemiiPropDimIIun

(2) There exists an unique $d \in [0,\infty]$ such that $\mathscr{H} \simeq \mathcal{E}_d$ as M-module.

ItemiiiPropDimIIun

(3) The algebra M' is a factor of type II_1 if and only if $\dim_M \mathscr{H} < \infty$.

ItemivPropDimIIun

(4) $\dim_M L^2(M) = 1$.

ItemvPropDimIIun

(5) If $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ is a set of separable M-modules, then

$$\dim_M\big(\oplus_n \mathcal{F}_n\big) = \sum_M \dim_M \mathcal{F}_n. \tag{74.107}$$

ItemviPropDimIIun

(6) If $\dim_M \mathscr{H} < \infty$ and if $P' \in \mathscr{P}(M')$, then

$$\dim_{PMP}(P\mathscr{H}) = \big(\operatorname{Tr}_M(P)\big)^{-1} \dim_M \mathscr{H}. \tag{74.108}$$

ItemviiPropDimIIun

(7) If P is a projection in M, then

$$\dim_{PMP}(P\mathscr{H}) = \big(\operatorname{Tr}_M(P)\big)^{-1} \dim_M \mathscr{H}. \tag{74.109}$$

ItemviiiPropDimIIun

(8) We have

$$\dim_{M'} \mathscr{H} = (\dim_M \mathscr{H})^{-1} \tag{74.110}$$

EqDimMMprimeprodun

if M' is a factor of type II_1

Proof. For (1), let begin with $d = n \in \mathbb{N}$, and define

$$\mathcal{E}_n = \underbrace{L^2(M) \oplus \ldots \oplus L^2(M)}_{n \text{ factors}} \tag{74.111}$$

which can be seen as an element of $\mathbb{M}_{1\times n}\big(L^2(M)\big)$. This is a M-$\mathbb{M}_n(M)$-bimodule, for the matrix multiplication. What we have is to compute $\dim_M \mathscr{H}_n = \operatorname{Tr} p$ where $p \in \mathscr{P}\big(\mathbb{M}_\infty(M)\big)$ is the projection such that $\big(L^2(M)\otimes l^2\big)P = \mathscr{H}_n$.

Intuitively, P is the projection onto the first n component of the space of infinite vertical matrices. Indeed, the picture is that $\mathscr{H}_n = \mathbb{M}_{1\times n}\big(L^2(M)\big)$ while $\mathscr{H}_\infty = \mathbb{M}_{1\times\infty}\big(L^2(M)\big)$. We are searching for $p \in \mathscr{P}(\mathbb{M}_\infty(M))$ such that

$$\big(L^2(M)\otimes l^2\big)p = L^2(M)\oplus\ldots\oplus L^2(M). \tag{74.112}$$

The answer is $p = \mathrm{Id}_M \otimes \mathtt{proj}_n$ where $\mathtt{proj}_n$ stands for the projection onto the first n components. Using formula (74.96), we conclude that $\dim_M \mathcal{E}_n = n$.

Let us now consider $d \in [0,\infty[$, and an integer $n \geqslant d$. Since $\mathbb{M}_n(M)$ is a factor of type II_1, proposition 74.78 provides a projection $q \in \mathscr{P}\big(\mathbb{M}_m(M)\big)$ such that $\operatorname{Tr}_{\mathbb{M}_n(M)}(q) = d/n$. Now we look at

$$\mathcal{E}_d = \mathcal{E}_n q. \tag{74.113}$$

In order to compute $\dim_M(\mathcal{E}_d)$, we have to find $r \in \mathscr{P}\big(\mathbb{M}_\infty(M)\big)$ such that

$$\big(L^2(M)\otimes l^2\big)r \simeq \big(L^2(M)\oplus\ldots\oplus L^2(M)\big)q. \tag{74.114}$$

The answer is

$$r = \begin{pmatrix} q & 0 \\ 0 & 0 \end{pmatrix} \tag{74.115}$$

whose trace is

$$\operatorname{Tr}_{\mathbb{M}_\infty(M)}(r) = \sum_{i=1}^{\infty} \operatorname{Tr}_M(r_{ii}) = \sum_{i=1}^{n} \operatorname{Tr}_M(q_{ii}) = n \operatorname{Tr}_{\mathbb{M}_n(M)}(q), \tag{74.116}$$

because of the discussion bellow the definition (74.96). This concludes the proof of (1).

We pass to the proof of (8). Since $\dim_M \mathscr{H} < \infty$, we know that $\mathscr{H}$ is a factor of type II_1 by (3). We can thus suppose that $\mathscr{H} = \mathcal{E}_d$ for some d. One knows that $M' = JMJ$, so that $\dim_{M'} \mathcal{E}_d = \dim_{JMJ} \mathcal{E}_d$.

Let us first work with $d = 1$. In order to compute $\dim_{JMJ}\big(L^2(M)\big)$, we are searching for $p \in \mathscr{P}\big(\mathbb{M}_\infty(JMJ)\big)$ such that $\big(L^2(JMJ)\otimes l^2\big)p \simeq L^2(M)$ as JMJ-modules. From lemma 74.82, $P = e_{11}$ works.

Let us now study the case $d = n \in \mathbb{N}$.

□

Corollary 74.85.
When M is a factor of type II_1, two M-modules are isomorphic if and only if they have same dimension.

This is a direct consequence of point (2) in proposition 74.84.

74.12.2 Dimension

We are going to associate any $M = M(\pi)$-module with a dimension in $[0,1]$ such that for every projection $P \in M$, the dimension of the submodule $M(\pi)P$ is equal to $\dim\big(M(\pi)P\big) = \operatorname{Tr}(P)$. We will in fact give a dimension to every module over a von Neumann algebra accepting a trace. For that, we use theory developed in section 54.4.

ThofgsurMFSubEEClSplits

Theorem 74.86.
If M is a finite von Neumann algebra and if E is a finitely generated module over M, and if F is any submodule, then the quotient map $E \to E/\operatorname{Cl}_E(F)$ splits and furthermore the quotient space $E/\bar{E}$ is projective and finitely generated.

¡¡ Avertissement/question au lecteur !! 74.87
Has one to add the assumption that F is projective?

Proof. Later. □

As consequence of the this theorem, $F \simeq \bar{E} \oplus$ (finite projective module). In particular,

$$F \simeq \mathrm{Cl}_F(0) \oplus (\text{finitely generated projective module}) \tag{74.117}$$

ThoPropDimiM

Theorem 74.88.
Let M be a finite von Neumann algebra with a normal faithful tracial state φ. There is an unique function

$$\dim_\varphi \colon \{M\} \to [0, \infty]$$

such that

(1) $F_1 \simeq F_2$ implies $\dim_\varphi F_1 = \dim_\varphi F_2$,

(2) (normalisation) if F is a finitely generated projective module, then the value of $\dim_\varphi(F)$ coincides with the definition (54.41), in particular when $E \simeq M^n P$ is finitely generated and projective, we have $\dim E = \sum_i \varphi(P_{ii})$,

(3) (continuity) if E is any submodule on a finitely generated module F, then $\dim_\varphi(E) = \dim_\varphi\big(\mathrm{Cl}_F(E)\big)$ [5],

ItemCofinalitySuiteExact

(4) (cofinality) given an exact sequence

$$0 \longrightarrow F_1 \longrightarrow F \longrightarrow F_2 \longrightarrow 0,$$

then $\dim_\varphi(F) = \dim_\varphi(F_1) + \dim_\varphi(F_2)$.

Proof. later. □

Notice that, since $P = P^*P$ for every projection, we have $\dim E \geqslant 0$.
We will prove the following

PropDimClEgalDim

Proposition 74.89.
If E is a finitely generated projective module over M, we have

$$\mathrm{Dim}(F) = \mathrm{Dim}\big(\mathrm{Cl}_E(F)\big) \tag{74.118}$$

if F is any submodule of E.

¡¡ Avertissement/question au lecteur !! 74.90
One has to check that is it actually proved somewhere in the next pages.

We will not often explicitly write the trace φ and simply write $\dim(E)$. However, we will sometimes precise the ring over which we consider the module and denote by $\dim_M(E)$ the dimension of E as module over M.

74.12.2.1 Examples

subsubsecExemDimMMMod

Let M be a von Neumann algebra with a normal faithful tracial state φ, and let us give some examples of modules over M.

First, M itself is a free module of dimension 1 [6] over M. If P is a projection, then MP is a projective module and $\dim_\varphi(MP) = \varphi(P)$. The fact that MP is a projective module comes from the direct sum decomposition $M = MP \oplus M(\mathbb{1} - P)$ and the third characterisation if projective module in proposition 54.14. Notice that M is a free module over M because $M = M\mathbb{1}$.

PgMPprojModule

5. From an analytic point of view, that condition is not usual, as can be seen on the example $\dim \mathbb{Q} = 0$ while $\dim \mathbb{R} = 1$.

6. In fact that dimension is $\mathrm{Tr}(\mathbb{1})$ which is usually normalised at 1, see definition (54.40) and bellow.

PropMTprojpourtoutT

Proposition 74.91.
For every element $T \in M$, the set MT is a projective M-module.

Proof. Let P be the projection onto $\overline{\text{Image}(T)}$, so that $P = \lim_{n\to\infty}(TT^*)^{1/n}$ as already seen. We are going to prove that $MT \simeq MP$ as M-module. For we define $\Phi\colon MT \to MP$ by $\Phi(ST) = SP$. In order to see that this is a good definition, suppose $S_1T = S_2T$, then $(S_1 - S_2)T = 0$. But we have

$$ST = 0 \quad \Rightarrow \quad S|_{\text{Image}(T)} = 0 \quad \Rightarrow \quad S|_{\overline{\text{Image}(T)}} = 0 \quad \Rightarrow \quad S|_{\text{Image}(P)} = 0, \tag{74.119}$$

so that $SP = 0$. It shows that Φ is a well defined M-module isomorphism between MT and MP. □

By polar decomposition, we have $MT = M|T|$, so we can suppose $T \geqslant 0$ without loss of generality. We have the exact sequence

$$0 \longrightarrow MT \xrightarrow{\Phi} MP \longrightarrow MP/MT \longrightarrow 0,$$

so that by property (4) of theorem 74.88, we have $\dim_\varphi(MT) = \dim_\varphi(MP) + \dim_\varphi(MP/MT)$. Since the modules MP and MT are isomorphic, that means that

$$\dim_\varphi(MP/MT) = 0.$$

As another example, take projections P_m with $\varphi(P_m) = 2^{-m}$, and let F be the infinite algebraic direct sum $MP_1 \oplus MP_2 \oplus \ldots$. By cofinality, it has at least the dimension of each of its submodule. In other words, we have

$$\dim F = \sup_n \dim(MP_1 \oplus \ldots \oplus MP_n) = 1.$$

Since we can embed F into M, we also get $\dim(M/F) = 0$.

74.12.2.2 Summary

We want to define, for each module, a dimension such that

— if E is a projective left module with $E \simeq M^nP$ (see proposition 54.15) and $P = P^2 \in \mathbb{M}_n(M)$,

$$\dim(E) = \sum \varphi(P_{ii}),$$

and does not depend on the choices.

— If E is not a finitely generated projective module, then

$$\text{Dim}(E) = \sup\big\{\dim(F) \text{ with } F\big\}. \tag{74.120}$$

DefDimAvecGrandD

74.13 Position of submodules

Proposition 74.92.
Let M be a von Neumann algebra (not specially with trace). Every finitely generated submodule of a finite generated projective module is projective.

Proof. Passing to the matrix algebra (lemmas 54.19 and 54.20), one can assume that the module and the submodule are in fact singly generated.

A singly generated module over M has the form $E = MT \subseteq M$ for some element $T \in M$. We saw during the proof of proposition 74.91 that $MT \simeq MP$ where P is the projection onto $\overline{\text{Image}(T)}$. The fact that MP is a projective module was already argued on page 3857. □

PropEfgpFssmodQuotProj

Proposition 74.93.
If E is a finite projective module over M and F is a submodule of E, then $E/\operatorname{Cl}_E(F)$ is projective.

Proof. Once again, using the matrix trick, one can suppose that E is singly generated by an element e. A left module map $\varphi\colon E \to M$ such that $\varphi(F) = 0$ is determined by the value of $\varphi(e) \in M$. The element $\varphi(e)$ is an operator on $\mathscr{H}$ and we define P_φ as the projection onto $\overline{\text{Image}\,(\varphi(e))}$.

Now we define the module map

$$\begin{aligned} \tilde{\varphi}\colon E &\to M \\ \tilde{\varphi}(Te) &\mapsto TP_\varphi. \end{aligned} \tag{74.121}$$

This is well defined. Indeed if $Se = 0$, then $\varphi(Se) = S\varphi(e) = 0$, which implies that $SP_\varphi = 0$. For the same reason, $\ker(\varphi) = \ker(\tilde{\varphi})$. We define

$$Q = \bigvee_\varphi P_\varphi, \tag{74.122}$$

and

$$\begin{aligned} \psi\colon E &\to M \\ \psi(Te) &\mapsto TQ \end{aligned} \tag{74.123}$$

which is well defined because $TQ = 0$ if and only if $TP\varphi = 0$ for every φ, since the operator T is bounded [7].

¡¡ Avertissement/question au lecteur !! 74.94
The explanation in the footnote is unclear; it has to be expressed in terms of nets.

By construction, $\ker(\psi) = \operatorname{Cl}_E(F)$, while $\text{Image}(\psi) = MQ$, so that

$$E/\ker(\psi) \simeq MQ, \tag{74.124}$$

while the latter is projective. □

CorEfgpFssIsom

Corollary 74.95.
When E is a finite projective module and F a submodule, we have an isomorphism

$$E = \operatorname{Cl}_E(F) \oplus E/\operatorname{Cl}_E(F)$$

as direct sum of modules.

Proof. If $\xi \in E$, the class of ξ in $E/\operatorname{Cl}_E(F)$ is

$$[\xi] = \{\xi + \eta \text{ tel que } \eta \in \operatorname{Cl}_E(F)\}. \tag{74.125}$$

Let us choose a representative $[\xi]_0$ in each of the classes [8]. The following is the module isomorphism we are searching for

$$\begin{aligned} \psi\colon E &\to \operatorname{Cl}_E(F) \oplus E/\operatorname{Cl}_E(F) \\ \xi &\mapsto \eta \oplus [\xi]_0 \end{aligned} \tag{74.126}$$

where η is the unique element of $\operatorname{Cl}_E(F)$ such that $[\xi]_0 + \eta = \xi$. □

When E is a finite projective module over M, we say that $\operatorname{Cl}_E(0)$ is the **torsion submodule** of E.

Let us see an example. Let

$$E = M/MT \tag{74.127}$$

7. If T is not bounded, it can get bigger and bigger on the range of φ_k when k goes to infinity, so that the limit of $T\bigvee_{\varphi\in\alpha}$ is not zero when α gets bigger.

8. This uses the famous axiom; it would be possible to do otherwise, isn't?

with $T \geqslant 0$ and $\overline{\text{Image}(T)} = \mathscr{H}$. We claim that the torsion submodule of E is E itself. A module map $\varphi\colon E \to M$ is a module map $M \to M$ composed with a projection, or in other words a module map $\varphi\colon M \to M$ such that $\varphi(T) = 0$. Since φ is a module map, its fulfils

$$\varphi(S) = \varphi(S\mathbb{1}) = S\varphi(\mathbb{1}) = SX \tag{74.128}$$

for a certain $X \in M$ such that $TX = 0$. Thus we have $\overline{\text{Image}(X)} \subset \ker(T)$, but since $T \geqslant 0$ (in particular T is self-adjoint) and $\overline{\text{Image}(T)} = \mathscr{H}$, we have $\ker(T) = 0$, and we conclude that $X = 0$, so that $\varphi = 0$. The question is to know if $E = M/MT$ has a finitely generated submodule or not. Let P be a projective submodule of M/MT; by the lifting property (54.34) we have a map $\tilde{\lambda}$ such that the following commutes

$$\begin{array}{ccc} & & M \\ & \overset{\tilde{\lambda}}{\nearrow} & \downarrow \lambda \\ P & \longrightarrow & M/MT \end{array} \tag{74.129}$$

where the arrow from P to M/MT is injective. The image of P by $\tilde{\lambda}$ is a submodule MX that has to be injectively [9] projected in M/MT, so that $MX \cap MT = \varnothing$. Notice that it is not possible when $\mathscr{H}$ is finite dimensional because T is invertible (from the fact that the closure of its image is the whole space), so that $MT = M$.

Now we suppose that X is positive. This is done without loss of generality because from polar decomposition, for every $X \in M$, there exists a positive Y such that $MX = MY$.

Since T is positive, we can consider the spectral theorem 60.62 and the isomorphism

$$\theta\colon C^*(T, \mathbb{1}) \to C\big(\text{Spec}(T)\big). \tag{74.130}$$

We define the following function on $\text{Spec}(T)$

$$f_\epsilon = \begin{cases} 0 & \text{if } x < \epsilon \\ \frac{1}{(\theta T)(x)} & \text{if } x \geqslant \epsilon \end{cases} \tag{74.131}$$

and then one defines the operator $S_\epsilon = \theta^{-1}(f_\epsilon) \in C^*(T, \mathbb{1})$. Let us prove that $P_\epsilon = S\epsilon T$ is a projection. We have $P_\epsilon^2 = S_\epsilon T S_\epsilon T$. Take an orthonormal basis of $\mathscr{H}$ of eigenvectors of T and let's call λ_i the eigenvalues: $Te_i = \lambda_i e_i$. If $\lambda_k < \epsilon$, then $S_\epsilon e_i = 0$ and of course $S_\epsilon T S_\epsilon e_i = S_\epsilon e_i$. Otherwise, we have

$$S_\epsilon T S_\epsilon e_i = \frac{1}{\lambda_i} S_\epsilon T e_i = S_\epsilon e_i. \tag{74.132}$$

That proves that $P_\epsilon^2 = P_\epsilon$, and so that this is a projection. We obviously also have $P_\epsilon \to \mathbb{1}$ in MT.

Similarly one can define Q, the projection onto $\overline{\text{Image}(X)}$ and we have projections $Q_\epsilon \to Q$ with $Q_\epsilon = Y_\epsilon X$ for some $Y\epsilon \in M$ and $Q_\epsilon \in MX$. Now take ϵ_1 and ϵ_2 and look at the projection

$$Q_{\epsilon_1} \vee P_{\epsilon_2} \tag{74.133}$$

onto $\text{Image}(Q_{\epsilon_1}) \cap \text{Image}(P_{\epsilon_2})$. The latter intersection is in fact 0 because $A_{\epsilon_1} \vee P_{\epsilon_2} \in MQ_{\epsilon_1} \cap MP_{\epsilon_2} = MX \cap MT = 0$. Using lemma 74.26, we get

$$\dim \mathbb{1} \geqslant \dim(Q_{\epsilon_1} \vee P_{\epsilon_2}) = \dim(Q_{\epsilon_1}) + \dim(P_{\epsilon_2}). \tag{74.134}$$

Recall that the trace used to define the dimension has to be normal, so that the dimension function is continuous in such a way that taking the limit $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ provides the expected result

$$\dim(\mathbb{1}) \geqslant \dim(Q) + \dim(\mathbb{1}), \tag{74.135}$$

from which one deduce that $\dim Q = 0$ and therefore $X = 0$. This finish the proof that the module $E = M/MT$ with a positive T and $\overline{\text{Image}(T)} = \mathscr{H}$ has no projective submodules.

9. How to say the fact to be injective in English? Does the word *injectively* exist?

Proposition 74.96.
If E is a finitely generated module over M, then the torsion submodule does not contains non zero projective finite submodules.

Proof. Later. □

LemFClosEF

Lemma 74.97.
If E is a finitely generated submodule of a finitely generated projective module E, then F is projective and $F \simeq \mathrm{Cl}_E(F)$.

Proof. We assume as usual that E and F are singly generated. The singly generated projective module E reads $E = MP$ while the general form of a singly generated submodule is $F = MT$ for a positive T which vanishes on $P^\perp\mathscr{H}$ and $\mathrm{Image}(T) \subseteq \mathrm{Image}(P)$. We already proved that $MT \simeq MQ$ where Q is the projection over $\mathrm{Image}(T)$.

We have $\mathrm{Cl}_{MP}(MT) = MQ$ because of the direct sum decomposition $MP = MQ \oplus M(P - Q)$ from which we can build an homomorphism $MP \to M$ which vanishes on MT, namely the projection because $MT \subseteq MQ$.

¡¡ Avertissement/question au lecteur !! 74.98
I do not understand one single word about the latter justification : (

□

Corfgfgdilleqdim

Corollary 74.99.
If F is a finitely generated submodule of a finitely generated projective module E, then $\dim(F) \leqslant \dim(E)$.

Proof. By lemma 74.97, we have $\dim(F) \simeq \dim\big(\mathrm{Cl}_E(F)\big)$ while we know that $E = \mathrm{Cl}_E(F) \oplus E/\mathrm{Cl}_E(F)$. The latter makes that $\dim E$ is given by the trace of two projections:

$$\dim E = \mathrm{Tr}\, P_{\mathrm{Cl}_E(F)} + \mathrm{Tr}\, P_{E/\mathrm{Cl}_E(F)}. \tag{74.136}$$

□

Corollary 74.100.
If E is a finitely projective module over a von Neumann algebra with a trace, then $\dim(E) = \mathrm{Dim}(E)$.

Proof. By very definition, $\mathrm{Dim}(E) \geqslant \dim(E)$ because E itself belongs to the set on which the supremum is taken in the definition (74.120) of Dim. The corollary 74.99 provides the opposite inequality. □

PropProjFiniDimCldim

Proposition 74.101.
If E is projective finitely generated and if $F \subseteq E$, then $\dim\big(\mathrm{Cl}_E(F)\big) = \mathrm{Dim}(F)$.

Proof. The case where F is finitely generated is already done by lemma 74.97. For the general case, suppose that the proposition does not hold. In this case, the lemma 74.97 yields

$$\sup\{\dim\big(\mathrm{Cl}_E(H)\big) \text{ tel que } H \text{ is finite projective } \} < \dim\big(\mathrm{Cl}_E(F)\big). \tag{74.137}$$

On the one hand, by corollary 74.95, the module $\mathrm{Cl}_E(H)$ is a direct summand of $\mathrm{Cl}_E(F)$ which is itself (by the same result) a direct summand of E. On the other hand, being a finitely generated module, it reads $E = M^n P$ for some projection P by proposition 54.15. Combining both, we have a projection $P_H \leqslant P$ such that $\mathrm{Cl}_E(H) = M^n P_H$.

Now the finitely generated projective submodules of F form a directed system. For every finitely generated projective submodules H_1 and H_2 of F, there exists at least $H_3 = H_1 \oplus H_2$ which contains H_1 and H_2. Thus the set of P_H is directed too and one can look at $\bigvee_H P_H$, the smalest projection whose range contains the range of all of the P_H.

We have $\bigvee_H P_H < P$ because if not, using normality of the trace,

$$\dim\Big(\bigvee_H P_H\Big) = \sup_H \dim(P_H) < P. \tag{74.138}$$

Now let $Q = P - \bigvee_H P_H$. If F is some submodule of E, one has

$$\begin{aligned} F &= \bigcup\{H \subseteq F \text{ tel que } H \text{ is finitely generated}\} \\ &= \bigcup\{H \subseteq F \text{ tel que } H \text{ is finitely generated and projective}\} \end{aligned} \tag{74.139}$$

because $H \subseteq F \subseteq E$ which is finitely generated projective. Such a H has the form $M^n P_H$, sp $F \subseteq M^n\big(\bigvee_H P_H\big)$. Notice that this F lies in the kernel of the non zero map

$$\begin{aligned} \mathrm{Cl}_E(F) = M^n P &\to M^n Q \\ V &\mapsto VQ. \end{aligned} \tag{74.140}$$

Indeed, since $\bigvee_H P_H < P$, we have $\big(\bigvee_H P_H\big)P = \bigvee_H P_H$, so that for every $T \in M^n$ we have $T\bigvee_H P_H\big(P - \bigvee_H P_H\big) = 0$. By looking at the complement of VQ, one has a nonzero homomorphism $\mathrm{Cl}_E(F) \to M$ which vanishes on F. That contradicts the definition of the closure.

□

The three point in this demonstration that use the von Neumann algebra background (and not only general module theory) are the following.

— First we used normality of the trace to commute the dimension with the supremum,

— and second, we used continuity of $\bigvee$ with respect to the dimension.

74.13.1 Summary

We have two dimension functions dim and Dim such that

(1) $\dim E = \mathrm{Dim}\, E$ whenever E is a finitely generated projective module,

(2) for every finitely generated module E and every submodule F, the module $E/\mathrm{Cl}_E(F)$ is finitely generated and projective,

¡¡ Avertissement/question à la lectrice !! 74.102
Check if one does not need the assumption that E is projective too.

(3) If $F \subseteq E$ and if E is a finitely generated projective module, then $\mathrm{Dim}(F) = \mathrm{Dim}\big(\mathrm{Cl}_E(F)\big)$.

From now we do no more use the "von Neumann algebra" assumption. Instead we suppose to have a ring R and two dimensions functions satisfying these three properties.

74.13.2 Properties of the dimension function

Proposition 74.103.
If E is the union of a directed system of submodules E_α, then $\mathrm{Dim}(E) = \sup \mathrm{Dim}\, E_\alpha$.

Proof. A finitely generated projective submodule in E is generated by n elements, each of them being contained in some E_{α_1}, $E_{\alpha_2}, \ldots$ By definition of a directed set, the union of all the so defined E_{α_i} is contained in a E_β. Thus every finitely generated projective submodule in E is of the form E_β. □

LemHinjectifHdimdim

Lemma 74.104.
If one has a projective module map $\rho\colon H_1 \to H_2$, then $\mathrm{Dim}(H_1) \geqslant \mathrm{Dim}(H_2)$.

Proof. Let F be any projective module for which there exists an inclusion $\iota\colon F \to H_2$. That map can be lifted because F is projective. So among all the submodules of H_1, there is the one which is the image of F be the lifted map. That one of course contains H_2 itself. Thus H_2 is a submodule of H_1 and the supremum defining the dimension in H_1 is automatically bigger or equal to the one defining the dimension of H_2. □

Proposition 74.105.
If

$$0 \longrightarrow E_0 \longrightarrow E_1 \xrightarrow{p} E_2 \longrightarrow 0 \tag{74.141}$$

is a short exact sequence of modules, then $\dim(E_1) = \dim(E_0) + \dim(E_2)$.

Proof. Using last proposition, we can assume that all of E_0, E_1 and E_2 are finitely generated. Indeed when a module is not finitely generated, it is still the union of the directed system of all its finitely generated submodules.

Let F be a finitely generated projective submodule of E_2, we have the exact sequence

$$0 \longrightarrow E_0 \longrightarrow p^{-1}(F) \longrightarrow F \longrightarrow 0. \tag{74.142}$$

The module F being projective, we have $p^{-1}(F) \simeq F \oplus E_0$. We do not know if the dimension function is additive with respect to direct sum, but by definition of a supremum, we have the inequality $\mathrm{Dim}(E_0) + \mathrm{Dim}(F) \leqslant \mathrm{Dim}\left(p^{-1}(F)\right)$, and the chain

$$\mathrm{Dim}(E_0) + \mathrm{Dim}(F) \leqslant \mathrm{Dim}\left(p^{-1}(F)\right) \leqslant \mathrm{Dim}(E_1) \tag{74.143}$$

which in turn provides the inequalities

$$\mathrm{Dim}(E_0) + \mathrm{Dim}(E_2) \leqslant \mathrm{Dim}(E_1). \tag{74.144}$$

For the reverse inequality, let F be a finitely generated projective submodule of E_1, and consider the following exact sequence of finitely generated projective module

$$0 \longrightarrow \mathrm{Cl}_F(F \cap E_0) \longrightarrow F \longrightarrow F/\mathrm{Cl}_F(F \cap E_0) \longrightarrow 0 \tag{74.145}$$

The fact that $F/\mathrm{Cl}_F(F \cap E_0)$ is projective is proposition 74.93. Since $F/\mathrm{Cl}_F(F \cap E_0)$ is at most a subset of F which is finitely generated, it has to be finitely generated too. We also know by corollary 74.95 that F splits into

$$F \simeq \mathrm{Cl}_F(F \cap E_0) \oplus F/\mathrm{Cl}_F(F \cap E_0) \tag{74.146}$$

which is a direct sum of finitely generated projective module, so that we can use the definition of dimension with traces (which sums up over direct sum) instead of the one with supremum. Thus we have

$$\dim(E) = \dim\left(\mathrm{Cl}_F(F \cap E_0)\right) + \dim\left(F/\mathrm{Cl}_F(F \cap E_0)\right). \tag{74.147}$$

The module F being projective and finitely generated, proposition 74.101 allows us to replace $\dim\left(\mathrm{Cl}_F(F \cap E_0)\right)$ by $\mathrm{Dim}(F \cap E_0)$ and write

$$\dim(E) = \mathrm{Dim}(F \cap E_0) + \mathrm{Dim}\left(F/\mathrm{Cl}_F(F \cap E_0)\right) \leqslant \mathrm{Dim}(E_0) + \mathrm{Dim}(F/(F \cap E_0)) \tag{74.148}$$

where we also used lemma 74.104. Since $F/(F \cap E_0)$ is a quotient of E_2, we have $\mathrm{Dim}(F/(F \cap E_0)) \leqslant \mathrm{Dim}(E_2)$, and taking the supremum over all suitable F, we find the result

$$\dim(E_1) \leqslant \mathrm{Dim}(E_0) + \mathrm{Dim}(E_2). \tag{74.149}$$

□

Proposition 74.106.
If E is a finitely generated module and $F \subseteq E$, then $\mathrm{Dim}(F) = \mathrm{Dim}\left(\mathrm{Cl}_E(F)\right)$.

Proof. Hint: by the proposition, one can assume that E is actually projective. □

¡¡ Avertissement/question à la lectrice !! 74.107
That has to be completed.

74.14 Decomposition of operators and representations

74.14.1 Motivation

Let G be a compact topological group and consider the space $L^2(G)$ (with respect to the Haar measure), and M be the von Neumann algebra generated by the left regular representation

$$M = \{U_g \text{ tel que } g \in G\}'' \tag{74.150}$$

with $(U_g f)(f) = g(g^{-1}h)$. Each U_g is an unitary operator in H. In this case, the commutant M' turns out to be the von Neumann algebra generated by the regular right representation.

¡¡ Avertissement/question à la lectrice !! 74.108
The operator R given by

$$\big(R(f)\big)(x) = \frac{f(x)}{2} \tag{74.151}$$

commutes with all the regular left representation, but is not part of the regular right representation. Do we impose unitarity conditions?

When $\sigma\colon G \to \mathrm{End}(V)$ is a finite dimensional irreducible representation of the group G, the **character** of σ is the function $\chi_\sigma\colon G \to \mathbb{C}$ defined by the formula

$$\chi_\sigma(g) = \mathrm{Tr}\big(\sigma(g)\big). \tag{74.152}$$

The representation σ thus defines the operator P_σ on $H = L^2(G)$,

$$(P_\sigma)(f) = \frac{\dim(\sigma)}{\mathrm{vol}(G)} \int \chi_\sigma(g^{-1})(U_g f)\, dg. \tag{74.153}$$

ThoPeterWeyl

Theorem 74.109 (Peter-Weyl)**.**
The operator P_σ is a projection which belongs to $M \cap M'$ and for which the following hold

ItemPeterWeyli

(1) If $\sigma_1 \not\simeq \sigma_2$, then $P_{\sigma_1} P_{\sigma_2} = 0$,

ItemPeterWeylii

(2) in the strong topology, $\sum_\sigma P_\sigma = \mathbb{1}$,

(3) the algebra $M \cap M'$ is generated by $\{P_\sigma\}$ and each of the P_σ is minimal in the center,

(4) $P_\sigma M P_\sigma$ is a factor on H_σ and we have $(P_\sigma M P_\sigma)' = P_\sigma M' P_\sigma$,

(5) $P_\sigma M P_\sigma$ is finite dimensional.

Notice that, because of point (1) and (2), if H_σ denotes the range of P_σ, then

$$H = \bigoplus_\sigma H_\sigma. \tag{74.154}$$

Proof. No proof. □

74.14.2 Conventions and definitions

From now, all Hilbert space will be separable and measure spaces (X, μ) will be as follows. The topological space X will be an Hausdorff, secondly countable metrisable space and μ will be the completion of a Borel measure which is finite on compact sets on X. By **completion**, we mean that all subset of a null set are measurable of measure zero.

Now let X be a set. A **field of Hilbert space** over X is a set of Hilbert spaces H_x for each $x \in X$. A **section** of the field is a function $V\colon C \to \bigsqcup_{x\in X} H_x$ such that $V(x) \in H_x$. We often write V_x instead of $V(x)$. The set of sections is denoted by $\Gamma\{H_x\}$.

Let H be a (separable) Hilbert space. A **direct integral decomposition** of H is

— a measure space (X, μ),

— a field of Hilbert spaces over X,

— a function from H to the space of sections of the field that we denote by v_x, the value at x of the section associated with $v \in H$

such that

(1) for every v and w in H, the function $x \mapsto \langle v_x, w_x \rangle$ is integrable and

$$\int_X \langle v_x, w_x \rangle d\mu(x) = \langle v, w \rangle, \tag{74.155}$$

(2) if u is any section of the field such that $x \mapsto \langle u_x, w_x \rangle$ is integrable for all $w \in H$, then u is almost everywhere equal to the section associated with a vector of H,

(3) the set of all decompositions of vectors in H is maximal.

¡¡ Avertissement/question au lecteur !! 74.110
I do not understand the precise signification if the maximality, cf problem 74.115.

Notice that the set of sections that are equal almost everywhere to sections obtained by decomposition of vectors in H form a module over $L^\infty(X, \mu)$.

An technical fact that will be used in much of proofs is the following.

LemdensHdensHx

Lemma 74.111.
If $\{v_1, v_2, \ldots\} \subseteq H$ has a dense span in H, then for almost every $x \in X$, the set $\mathrm{Span}\{v_{1x}, v_{2x}, \ldots\}$ *is dense in H_x.*

74.14.3 Examples

74.14.3.1 First example

Let H_x with $x \in X$ be a field of Hilbert space over a countable set X. Then define $H = \bigoplus_{x \in X} H_x$. That Hilbert space decomposes by (X, μ) with μ being the counting measure, and to $v \in H$ corresponds a component v_x in each H_x that are taken as the section associated to the vector v.

74.14.3.2 Second example

Let (X, μ) be the counting measure over the countable set X, and consider the Hilbert space $H = L^2(X, \mu)$. A decomposition of that Hilbert space is given as follows: first pose $H_x = \mathbb{C}$ for each $x \in X$. Then, for each vector $v \in H$, we associate some choice of function f in the equivalence class that defines v. Then we define the section associated with v is $v_x = f(x)$.

Let us check that this construction fulfils the axioms of a direct integral decomposition of Hilbert space. First the function $x \mapsto \langle v_x, w_x \rangle$ has to be integrable. If f is the function associated with v and g the one associated with w, we have $\langle v_x, w_x \rangle = f(x)\overline{g(x)}$ which is integrable by Cauchy-Schwarz, and by the definition of the product in L^2, we have

$$\langle v, w \rangle = \int_X f(x)\overline{g(x)}\, d\mu(x). \tag{74.156}$$

Second, suppose that u is a section such that the map $x \mapsto u_x\overline{f(x)}$ is integrable for every f associated with a vector in H. The measured space (X, μ) being countable, X is a union of finite measure sets $\{X_i\}$ on each of them one can consider the characteristic function 1_i. Then for every i, the function $x \mapsto u_x 1_i$ is measurable on X_i. That proves that u is measurable. We conclude that it is square integrable and thus defines a function which is almost everywhere equal to the section associated with a vector $v \in L^2(X, \mu)$, namely the element of $L^2(X, \mu)$ which is the class of u.

74.14.3.3 Third example

Let $H = L^2(X, \mu_1) \oplus L^2(X, \mu_2)$ and $\mu = \mu_1 + \mu_2$ which is still a complete measure. By Radon-Nikodým theorem 74.19, there exists positive measurable functions f_1 and f_2 such that $\mu_i = f_i\mu$. We can decompose H as the direct integral

$$H_x = l^2\Big(\{x\} \cap \{f_1 > 0\} \sqcup \{x\} \cap \{f_2 > 0\}\Big). \tag{74.157}$$

Depending on the x, the set $\{x\} \cap \{f_1 > 0\} \sqcup \{x\} \cap \{f_2 > 0\}$ has zero, one or two elements. Now if $v = (v_1, v_2) \in H$ with $v_! = [g_1]$ and $v_2 = [g_2]$, we define the section associated with v as

$$v_x = \big(f_1(x)^{1/2} g_1(x), f_2(x)^{1/2} g_2(x)\big). \tag{74.158}$$

As notation, if $f_1(x) = 0$, we identify this with the element $f_2(x)^{1/2} g_2(x)$ instead of the couple $(0, f_2(x)^{1/2} g_2(x))$.

74.14.4 Decompositions of operators

Suppose that the separable Hilbert space H is provided with a decomposition over (X, μ). A **decomposition** of the operator $T \in \mathfrak{B}(H)$ is a family of bounded operators $T_x \colon H_x \to H_x$ such that for every $v \in H$, the condition

$$(Tv)_x = T_x v_x \tag{74.159}$$

EqCondTvDecompOps

holds for almost every $x \in X$.

Remarque 74.112.
The conditions 74.159 are in fact uncountably many conditions (one for each $v \in H$) each of them having the possibility to be wrong on a null set. One has thus to be prudent because an uncountable union of null set *might* be a set of non zero measure.

Lemma 74.113.
The set of decomposable operators is a $*$-algebra.

Proof. If the operator T is decomposable, the rule $(T^*)_x = (T_x)^*$ provides a decomposition of T^*. □

Lemma 74.114.
Let T be a decomposable operator. We have $T \geqslant 0$ if and only if $T_x \geqslant 0$ almost everywhere.

Proof. In order to prove that T is positive, we have to evaluate $\langle Tv, v\rangle$ from the decomposition of T and v. We suppose that $T_x \geqslant 0$ for almost every $X \in X$ and we compute

$$\langle Tv, v\rangle = \int_X \langle (Tv)_x, v_x\rangle d\mu(x) = \int_X \langle T_x v_x, v_x\rangle d\mu(x) \geqslant 0. \tag{74.160}$$

where we were allowed to change $(Tv)_x$ by $T_x v_x$ because the equality holds almost everywhere.

For proving the contrary, let pick a dense, countable and rational vector subspace $H_0 \subseteq H$, and assume that $T \geqslant 0$. If $v \in H$ and if $E \subseteq X$ is measurable, then by maximality of the Hilbert space decomposition there is some vector $v_E \in H$ such that

$$(v_E)_x = \begin{cases} v_x & \text{if } x \in E \\ 0 & \text{otherwise.} \end{cases} \tag{74.161}$$

EqCondvEyx

ProbMaximvE

¡¡ Avertissement/question au lecteur !! 74.115
I do not see why to use the maximality. Indeed, one considers the section

$$s_x = \begin{cases} v_x & \textit{if } x \in E \\ 0 & \textit{otherwise.} \end{cases} \tag{74.162}$$

By definition of decomposition of v, the integral

$$\int_X \langle v_x, w_x \rangle d\mu(x) \tag{74.163}$$

exists for every $w \in H$. Now we have that

$$\int_X \langle s_x, w_x \rangle d\mu(x) = \int_E \langle v_x, w_x \rangle d\mu(x) \tag{74.164}$$

whose existence is assured by existence of (74.163), isn't?

Existence of integral (74.164) assures the existence of a vector $v_E \in H$ such that $(v_E)_x = s_x$ almost everywhere. So we have the v_E without maximality axiom. Do we really want the condition (74.161) to hold everywhere and not only almost everywhere? Since we only use it in integrals, I think that one does not care about a violation of condition (74.161) on a null set.

We have that

$$\int_E \langle T_x v_x, v_x \rangle d\mu(x) = \int_X \langle T_x v_{Ex}, v_{Ex} \rangle = \langle T v_E, v_E \rangle \geqslant 0 \tag{74.165}$$

because $T \geqslant 0$ by assumption. So we proved that $x \mapsto \langle T_x v_x, v_x \rangle$ is an integrable function on X for which the integral over any subset is a positive real number. Then the function is real positive almost everywhere. Thus we have

$$\langle T_x v_x, v_x \rangle \geqslant 0 \tag{74.166}$$

almost everywhere. Therefore there exists a null set $N \subseteq X$ such that $\langle T_x v_x, v_x \rangle \geqslant 0$ for every $v \in H_0$ whenever $x \in X \backslash N$. If $x \in X \backslash N$, then $\mathrm{Span}\{v_x \text{ tel que } v \in H_0\}$ is dense in H_x because of lemma 74.111. Notice that H_0 is countable and there is one null set for each element of H_0 on which we are unsure of the sign of $\langle T_x v_x, v_x \rangle$.

By enlarging the set N we can assume that

$$(a_1 v_1 + a_2 v_2) = a_1 (v_1)_x + a_2 (v_2)_x \tag{74.167}$$

for every $x \in X \backslash N$ and $a_i \in \mathbb{Q}[i]$. Indeed the equalities are an equation for each $a_i \in \mathbb{Q}[i]$ and $v_i \in H_0$. That is thus a countable set of equations; thus the set where the equations are not satisfied is a countable union of null sets.

What we have now is that every vector in H_x (with $x in X \backslash N$) can be written as a limit if vectors in H_0, and then $\langle T_x v_x, v \rangle \geqslant 0$, so that $T_x \geqslant 0$.

□

Lemma 74.116.
We have

$$\|T\| = \text{ess-sup}\, \|T_x\| = \min\{c \text{ tel que } \|T_x\| \leqslant c \text{ almost everywhere}\}. \tag{74.168}$$

Proof. The inequality $\|T\| \leqslant \text{ess-sup}\, \|T_x\|$ comes from the fact that

$$\|Tv\| = \int_X \langle T_x v_x, T_x v_x \rangle d\mu(x) \tag{74.169}$$

which gives the bound. For the inverse inequality, remark that $\|T\|^2 = \|T^* T\|$ and $(T^* T)_x = (T_x)^* T_x$ almost everywhere, so that one can replace T by $T^* T$, or more simply we can assume that $T \geqslant 0$. Now consider the positive and self-adjoint operator $\|T\| \mathbb{1} - T$. The "component" operator is positive almost everywhere: $\|T\| \mathbb{1} = T_x \geqslant 0$ for almost every $x \in X$. We know that the maximum of the spectrum of T_x is $\|T_x\|$, so that the number $\|TY\| - \|T_x\|$ is almost everywhere positive because it is an element of the spectrum of $\|T\| \mathbb{1} - T_x$ which is positive. □

74.14.5 Diagonal and decomposable operators

An operator S is **diagonal** if it is decomposable and if S_x is multiple of identity for almost every $x \in X$.

As motivation, consider the example $H = \mathbb{C}^{n_1+n_2} = \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2}$, the measure space being two points with the counting measure. A diagonal operator looks like

$$\begin{pmatrix} \lambda_1 \mathbb{1} & \\ & \lambda_2 \mathbb{1} \end{pmatrix}, \tag{74.170}$$

while decomposable operators are

$$\begin{pmatrix} T_1 & \\ & T_2 \end{pmatrix}. \tag{74.171}$$

One sees that these two sets are mutually commutant. We will see in theorem 74.119 that this is a general fact that decomposable operators and diagonal operators are mutually commutant.

ThoKaplanskyDensity

Theorem 74.117 (Kaplansky density theorem)**.**
If A is a $*$-algebra of operators on H, then every strong limit of net in A is a strong limit of a bounded net of elements in A. In other words, if $T = \lim_{\rightarrow} T_\alpha$, then there exist a net S_β with $\|S_\beta\| \leqslant M$ (fixed M) such that $T = \lim_{\rightarrow} T_\beta$.

A complete proof is given in [942].

Sketch of proof. First, a convex subset of $\mathfrak{B}(\mathscr{H})$ is strongly closed if and only if it is weakly closed. So, using what is said on page 3492 about the fact that weak topology is compatible with the involution, we know that if T lies in the strong closure of A, and if $T = T^*$, then T is the limit of a net of self-adjoint elements in A. That allows us to focus the proof of the theorem on self-adjoint elements of A.

¡¡ Avertissement/question à la lectrice !! 74.118
For me, the fact that this reasoning actually allows to restrict ourself to self-adjoint elements. But I suppose that it would become more clear when one understands the link between the statement here and the statement of Kaplansky's theorem in [942], page 32.

Consider the function $f(x) = 2x/(1+x^2)$ that provides an homeomorphism from $[-1,1]$ to $[-1,1]$. Thus by continuous functional calculus, if $T = T^*$ and $\|T\| \leqslant 1$, we have $T = f(S)$ where $S = S^*$ and S belongs to the norm closure of the algebra generated by T.

Now if T is the strong limit of the net T_α of self-adjoint operators, then S is the strong limit of S_α with $S_\alpha \in A$ and $S_\alpha^* = S_\alpha$. Indeed one can prove that $S_\alpha \to S$ implies [10] $f(S_\alpha) \to f(S)$.

Since $\|f(S_\alpha)\| \leqslant 1$, thus $\|T\| \leqslant 1$. □

ThoDecopDiagCommE

Theorem 74.119.
The algebra of decomposable and diagonal operators are mutually commutant.

Proof. The fact that decomposable operators are in the commutant of diagonal operators and vice versa. The strategy to prove the theorem will be to first prove that decomposable and diagonal operators are von Neumann algebras, and then show that every operator in the commutant of diagonal operators is decomposable. From there, the conclusion yields from the double commutant theorem.

Suppose that T is a strong limit of decomposable operators, i.e

$$T = \mathrm{s}\lim T_\alpha \tag{74.172}$$

where T_α is decomposable and $\sup\{|T_\alpha|\} < \infty$. We want to show that T is decomposable. For, let H_0 be a dense countable rational subspace of H. If T is limit of a uniformly bounded *net*, the fact

10. In the weak topology, that implication is the continuous functional calculus, but in the strong topology (the one we are considering here), the validity of this assertion relies on the particular form of f.

that H_0 is countable makes that there is a *sequence* of uniformly bounded decomposable operators T_n such that $T_n v \to Tv$ for every $v \in H_0$.

We have

$$\lim_{n\to\infty} \int_X \|T_{nx}v_x - (Tv)_x\|^2 d\mu(x) = 0 \tag{74.173}$$

for almost every $v \in H_0$. A general fact about L^2 spaces is that when a sequence goes to zero, then the functions (representative of the elements of L^2) themselves goes to zero on a subsequence. So for every given $v \in H_0$ there is a subsequence $\{n_j\}$ such that

$$\|T_{n_j x}v_x - (Tv)_x\| \to 0 \tag{74.174}$$

for almost every $x \in X$. The subsequence might depends on the v, bu a diagonal argument provides a subsequence $\{n_k\}$ of $\{n_j\}$ such that for every $v \in H_0$,

$$\|T_{n_k x}v_x - (Tv)_x\| \to 0 \tag{74.175}$$

for almost every $x \in X$. These relations are a countable number of convergence almost everywhere. Then there exists a null set $N \subseteq X$ such that

$$\|T_{n_k x}v_x - (Tv)_x\| \to 0 \tag{74.176}$$

EqTnkxTvxHorsN

for every $v \in H_0$ and $x \in X\backslash N$.

By enlarging again the null set N, we define

$$H_{0x} = \{v_x \text{ tel que } v \in H_0\} \tag{74.177}$$

for $x \in X\backslash N$. This is a dense subspace of H_x by lemma 74.111. Since the essential supremum of $\|T_{nx}\|$ is the norm of T_n, we have a "double uniformly bounded" relation

$$\sup_n \sup_x \|T_{nx}\| < \infty. \tag{74.178}$$

Now for $x \in X\backslash N$ we define T_x on H_{0x} by

$$T_x v_x = \lim_{k\to\infty} T_{n_k x}v_x \tag{74.179}$$

for each $v \in H_0$. That limit exists by construction.

¡¡ Avertissement/question au lecteur !! 74.120

I suppose that this definition together with property (74.176) *means that for every* $w \in H_0$ *and every* $x \in X\backslash N$,

$$(Tw)_x - T_x w_x = 0. \tag{74.180}$$

EqSurHzeroTwTxwx

It extends by continuity from H_{0x} to H_x. The resulting operator is bounded because the sequence $T_{n_k x}$ is uniformly bounded.

Now we enlarge once again the set N by the null set appearing in the definition of the essential supremum we find that T_x is bounded for every $x \in X\backslash N$. All the work make $\{T_x\}_{x\in X}$ a candidate decomposition of T.

In order to check that this actually is a decomposition of T, we have to prove that $(Tv)_x = T_x v_v$ for every $v \in H$ and almost every $x \in X$. We will prove that property by proving that the integral

$$\int_X \|(Tv)_x - T_x v_x\| \, d\mu(x) \tag{74.181}$$

vanishes. For that, we choose $w \in H_0$ such that $\|w - v\| \leqslant \epsilon$ and, by virtue of (74.180), we add $T_x w_x - (Tw)_w$ in the integrand:

$$\begin{aligned}
\int_X \|(Tv)_x - T_x v_x\|^2 \, d\mu(x) &= \int_X \|(Tv)_x - (Tw)_x + T_x w_x - T_x v_x\|^2 \, d\mu(x) \\
&\leqslant 2\int_X \|(Tv)_x - (Tw)_x\|^2 \, d\mu(x) + \int_X \|T_x w_x - T_x v_x\|^2 \, d\mu(x) \\
&= 2\|T(v-w)\|^2 + 2\|\tilde{T}(v-w)\|^2 \\
&= 2\|T\| \, \|v-w\|^2 + 2\|\tilde{T}\| \, \|v-w\|^2 \\
&\leqslant \text{constant} \cdot \epsilon
\end{aligned} \tag{74.182}$$

where the operator $\tilde{T}$ is defined by $(\tilde{T}w)_x = T_x w_x$.

Thus the set of decomposable operators is strongly closed, so that it is a von Neumann algebra. The same argument holds to prove that diagonal operators form a von Neumann algebra too. We want now to prove that the commutant of diagonal operators is the set of decomposable. For that, it is sufficient to prove that every operator in the commutant of diagonals is decomposable. Since a von Neumann algebra is generated by its projections , we have only to prove that every projection in the commutant of the diagonal is decomposable.

Let P be such a projection and enlarge H_0 in order to have $PH_0 \subset H_0$. That remains a countable set because such an enlargement is $H_0' = H_0 + PH_0$ for example. We choose a null set $N \subseteq X$ such that if $x \in X\backslash N$ then

$$(a_1 v_1 + a_2 v_2)_x = a_1 v_{1x} + a_2 v_{2x} \tag{74.183}$$

for every v_1 and v_2 in H_0 and $a_i \in \mathbb{Q}[i]$. The set $H_{0x} = \{v_x \text{ tel que } v \in H_0\}$ is dense in H_x. For $f \in L^\infty(X, \mu)$, we define the multiplication operator

$$(M_f v)_x = f(x) v_x, \tag{74.184}$$

which is, by definition, a diagonal operator. Let $v \in PH_0$ and $w \in P^\perp H_0$, we have

$$\int_X f(x)\langle v_x, w_x\rangle \, d\mu(x) = \int_X \langle f(x)v_x, w_x\rangle \, d\mu(x) = \langle M_f v, w\rangle. \tag{74.185}$$

Since P is in the commutant of diagonals, it commutes with M_f, so that $M_f v \in PH$ while $w \in P^\perp H_0$ and the latter product is thus zero: $\langle M_f v, w\rangle = 0$. That proves that $x \mapsto \langle v_x, w_x\rangle$ is a function that integrates to zero when multiplicated [11] by any L^∞ function. That means in turn that $\langle v_x, w_x\rangle = 0$ almost everywhere. We enlarge N in such a way that

EqPerpPHPperpH
$$\{v_x \text{ tel que } v \in PH_0\} \perp \{w_x \text{ tel que } w \in P^\perp H_0\} \tag{74.186}$$

for every $x \in X\backslash N$. Of course, $PH_0 \oplus P^\perp H_0 = H_0$, so that the direct sum of the two spaces of (74.186) is H_{0x}. Taking the strong closure,

$$\overline{\{v_x \text{ tel que } v \in PH_0\}} \oplus \overline{\{w_x \text{ tel que } w \in P^\perp H_0\}} = H_x \tag{74.187}$$

where the double bar denotes the strong closure. Now, for $x \in C\backslash N$, define $P_x \colon H_x \to H_x$, the projection onto $\overline{\{v_x \text{ tel que } v \in PH_0\}}$. By the same reasoning as for T before, we prove that $\{P_x\}$ decomposes P. □

LemMabelAstrDense

Lemma 74.121.
If M is an abelian von Neumann algebra on a separable Hilbert space H, then M has a strongly dense C^*-algebra A

Proof. No proof. □

ThoVNableHDiag

Theorem 74.122.
If M is an abelian von Neumann algebra on a (as usual separable) Hilbert space H, then there is a decompositon of H such that M is the set of diagonal operators.

Proof. Pick a $A \subseteq M$ as in the lemma 74.121, and add an unit is needed. Then we have $A \simeq C(X)$ for some metrisable compact space X by the Gelfand theorem (see [489]). We can write

$$H = \bigoplus_{k=1}^{\infty} H_k \tag{74.188}$$

where H_k is invariant under the action of A and has a cyclic vector. Indeed, pick a vector v_1 in H and define $H_1 = Av_1$, then pick v_2 in the complement and continue. We have an isomorphism $H_k \simeq L^2(X, \mu_k)$ where μ_k is some probability measure by $v_k \mapsto 1$ and $Tv_k \mapsto \hat{T}$ where the hat denotes the Gelfand isomorphism.

11. My spelling corrector does not know that word. Bad word or bad corrector?

¡¡ Avertissement/question au lecteur !! 74.123
That gives an element of $L^2(X,\mu_k)$ for each vector inside H_k. But I do not see the isomorphism. For example a simple step function belongs to $L^2(X,\mu_k)$ and corresponds to which element of H_k?

The map

$$\begin{aligned} \psi\colon Av_1 &\to L^2(X,\mu) \\ Tv_1 &\mapsto \hat{T} \end{aligned} \tag{74.189}$$

(whose image is included in $L^2(X,\mu)$ because X is compact, so that every continuous function is square integrable with respect to a probability measure) is injective because two different continuous functions cannot belong to the same class in $L^2(X,\mu)$.

For me, surjectivity is not clear (and even wrong) because there exists many elements in $L^2(X,\mu)$ who have no continuous representative. Do we have to game with limits and exploit the fact that A is strongly closed?

So we can assume that $A = C(X)$ and $H = \bigoplus_k L^2(X,\mu)$. Let

$$\mu = \sum_{k=1}^{\infty} 2^{-k}\mu_k \tag{74.190}$$

which is still a probability measure, and define $H_x = l^2(\mathbb{N},\mu_x)$ where μ_x is the measure on $\mathbb{N}$ defined by

$$\mu_x\big(\{k\}\big) = \frac{d\mu_k}{d\mu}(x) \tag{74.191}$$

by the Radon-Nikedým theorem. For recall, $d\mu_k/d\mu$ is the function on X defined by the fact that $d\mu_k = (d\mu_k/d\mu)d\mu$ in the sense that

$$\int_X f d\mu_k = \int_X f \frac{d\mu_k}{d\mu} d\mu \tag{74.192}$$

for every functions f on X. Now for each $\varphi = (\varphi_1, \varphi_2, \ldots) \in H$, we define $\varphi_x \in H_x$ by

$$\begin{aligned} \varphi_x\colon \mathbb{N} &\to \mathbb{C} \\ k &\mapsto \varphi_k(x). \end{aligned} \tag{74.193}$$

One has to show that the so defined φ_x is actually an element of $l^2(\mathbb{N},\mu_x)$, that is

$$\sum_k |\varphi_k(x)|^2 \frac{d\mu_k}{d\mu}(x) \tag{74.194}$$

EqSumVpNK

has to be finite. For, we have the computation

$$\begin{aligned} \int_X \sum_k |\varphi_k(x)|^2 \frac{d\mu_k}{d\mu}(x)\, d\mu(x) &= \sum_k \int_X |\varphi_k(x)|^2 \frac{d\mu_k}{d\mu}(x)\, d\mu(x) \\ &= \sum_k \int_X |\varphi_k|^2\, d\mu_k(x) \\ &= \sum_k \|\varphi_j\|^2 \\ &= \|\varphi\|^2. \end{aligned} \tag{74.195}$$

EqIntXsumnormVarPhi

where, in the first line we permuted the sum and the integral using the monotone convergence theorem. The fact that integral (74.195) is finite proves that the function (74.194) has finite values almost everywhere, or

$$\|\varphi_x\|^2_{l^2(\mathbb{N},\mu_x)} < \infty \tag{74.196}$$

almost everywhere. We redefine now φ_x as zero for the x for which the former definition gives $\|\varphi_x\|^2_{l^2(\mathbb{N},\mu_x)}$. This is a redefinition over a null set.

One can check that this construction provides a decomposition of H for which M is the algebra of diagonal operators.

¡¡ Avertissement/question au lecteur !! 74.124
To be done...

□

Let M be an abelian von Neumann algebra of operators on H and suppose that there are two decompositions $\{H_x\}$ and $\{H_y\}$ of H using M by theorem 74.122. Let now

$$M_2 = \left\{ \begin{pmatrix} T & \\ & T \end{pmatrix} \right\} \text{ tel que } T \in M \tag{74.197}$$

That von Neumann algebra decomposes $H \oplus H$ into $\{H_x\} \oplus \{H_y\}$. The matrix

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{74.198}$$

is decomposable because it belongs to the commutant of M_2. Thus it provides an unitary isomorphism of H that applies H_x on H_y.

PropNprimexxNprime

Proposition 74.125.
If N lies in the commutant of M, there are von Neumann algebras $N_x \subseteq \mathfrak{B}(H_x)$ for almost every x such that

(1) $N_x = \{T_x \text{ tel que } T \in N\}''$,

(2) $(N_x)' = (N')_x$ almost everywhere.

Proof. No proof. □

74.14.6 Decompositions of representations

Let A be a separable C^*-algebra and $\pi\colon A \to \mathfrak{B}(H)$ be a representation of A on a separable Hilbert space. We do not assume that π is faithful.

LemHDecPipixxpi

Lemma 74.126.
If H is provided with a decomposition $\{H_x\}$ and if $\pi(A)$ consists of decomposable operators, then there are representations $\pi_x\colon A \to \mathfrak{B}(H_x)$ such that $\pi(a)_x = \pi_x(a)$.

Proof. No proof. □

ThoRepDecSepIrrep

Theorem 74.127.
If $\pi\colon A \to \mathfrak{B}(H)$ is any representation of a separable C^*-algebra, then there is a decomposition of H such that $\pi(A)$ consists of decomposable operators. There also exists a decomposition $\{\pi_x\}$ of π such that each π_x is irreducible.

That theorem says that every representation decomposes into irreducible.

Proof. We do not prove the first part, and suppose then that $\{\pi_x\}$ is a decomposition given by lemma 74.126, so that $\pi_x(a) = \pi(a)_x$.

Let M be a maximal (Zorn's lemma) abelian von Neumann subalgebra of $\big(\pi(A)\big)'$, and decompose H so that M is the algebra of diagonal operators. Now we consider N, the von Neumann algebra generated by M and $\pi(A)$. An element in N' belongs to $\pi(A)'$ and commutes with M which is maximal. Thus $N' = M$.

By the property of the decomposition π_x, an element of $\pi_x(A)'$ has to commute with all M, so that

$$\pi_x(A)' = (N_x)' = (N')_x = M_x = \mathbb{C}\mathbb{1}, \tag{74.199}$$

and by Schur's lemma, the representation π_x is irreducible almost everywhere. □

As an example, take $A = C^*(F_2)$ where F_2 is the free group with two generators and A is formed by taking all the linear combination and then the closure as C^*-algebra. One can show that

(1) There are two decompositions of the regular representation π of A on $H = l^2(F_2)$ that we denote by $\{H_x\}$ and $\{H_y\}$. These decompositions are *a priori* built on different measured spaces.

(2) All the representations in $\{\pi_x\} \sqcup \{\pi'_y\}$ are irreducible and mutually inequivalent.

So the decomposition in irreducible representations is not unique at all, in contrast to the group representation case. The point is that the regular representation of $C^*(F_2)$ is a factor, while the following theorem only assures unicity of decomposition into factors.

Theorem 74.128.
Let $\pi\colon A \to \mathfrak{B}(H)$ be a representation of the C^*-algebra A. There is an essentially unique decomposition of H such that

(1) the representation $\pi(A)$ consists of decomposable operators,

(2) the representations π_x are factors representations for almost every x.

By "essentially unique", one means that if

$$\big(X, \mu, \{H_x\}, H \to \Gamma\{H_x\}\big) \text{ and } \qquad \big(X', \mu', \{H'_x\}, H \to \Gamma\{H'_x\}\big) \tag{74.200}$$

are two decompositions, there exists a map $X \to X'$ which is almost everywhere an equivalence of measurable spaces. That is it maps μ to a measure which is mutually absolutely continuous with μ', and an unitary isomorphism $H_x \to H'_x$ defined for almost every X which maps v_x to v'_x where v'_x is the decomposition of v with respect to the decomposition $\{H'_x\}$.

Proof. Let us prove the existence part of the theorem For we proceed as in the proof of theorem 74.127, but we take $M = \pi(A)'' \cap \pi(A)'$, this is the center of the von Neumann algebra generated by $\pi(A)$. Using proposition 74.125, we find

$$\pi_x(A)'' \cap \pi_x(A)' = \big(\pi(A)'' \cap \pi(A)'\big)_x = \mathbb{C}\mathbb{1}. \tag{74.201}$$

That shows that $\pi_x(A)$ is a factor. □

A C^*-algebra is **liminal** if for every irreducible representation $\pi\colon A \to \mathfrak{B}(H)$, we have $\pi(A) = \mathcal{K}(H)$, the space of compact operators on H.

Theorem 74.129.
If G is a semisimple Lie group, then $C^*(G)$ is liminal.

Proof. No proof. □

Theorem 74.130.
If A is a liminal C^*-algebra, then every factor representation is of type I.

Theorem 74.131.
If π is any representation of a liminal C^*-algebra on a separable Hilbert space, then there is a decomposition of H such that

(1) the space H_x carries a representation which is a factor of type I,

(2) $H_x = H_{\pi_x} \otimes L_x$ where H_{π_x} is an irreducible representation of A, all mutually inequivalent and L_x is some Hilbert space which says the "multiplicity" of the representation π_x inside π.

That decomposition is essentially unique.

Proof. No proof. □

Lemma 74.132.
Every irreducible representation of $A_1 \otimes A_2$ with both A_i being liminal is a tensor product of irreducible representations of A_1 and A_2.

Now take an unimodular group G such that $C^*(G)$ is liminal. One knows that the von Neumann algebras generated by the left and right regular representations are mutually commutant. We consider the bi-regular representation $G \times G$ on $L^2(G)$. The commutant of that representation is equal to the center of the von Neumann algebra generated by the two regular representation and is in particular abelian.

Using the lemma, the space $L^2(G)$ decomposes into representations $H_{\pi\tau} = H_\pi \otimes H_\tau$ of $G \times G$ where π and τ are irreducible representations. From the unimodular assumption, the symmetry $f(g) \to f(g^{-1})$ is an isometry that intertwines left and right regular representations. Notice that a function can be seen as a vector on $L^2(G)$ as well as as an operator over $L^2(G)$ by the convolution. Thus we turn $H_{\pi\tau}$ into a $*$-algebra by the multiplication $H_{\pi\tau} \otimes H_{\pi\tau} \tau H_{\pi\tau}$.

We denote by $\hat{G}$ the set of irreducible representations of G

Theorem 74.133.
*Let G be an unimodular locally compact liminal group. There is an unique measure on $\hat{G}$, the **Plancherel measure**, denoted by μ such that for every $f \in L^1(G) \cap L^2(G)$ (as vector in $L^2(G)$ or as operator by involution),*

$$\|f\|^2_{L^2(G)} = \int_{\hat{G}} \|\pi(f)\|^2_{HS} \, d\mu(\pi) \tag{74.202}$$

where $\|.\|_{HS}$ denotes the Hilbert-Schmidt operator norm. Moreover we have

$$\pi(f) = \int_G f(g)\pi(g) \, dg. \tag{74.203}$$

That theorem has to be compared to the Fourier theory for abelian groups like $\mathbb{R}$ or S^1, in particular the Parseval equality. This has also to be compared with the following theorem.

Theorem 74.134 (Peter-Weyl)**.**
If G is a compact group, then

$$\|f\|^2_{L^2(G)} = \sum_{\hat{G}} \frac{\dim(\pi)}{\operatorname{vol}(G)} \|\pi(f)\|^2_{HS} \tag{74.204}$$

where the ration $\dim(\pi)/\operatorname{vol}(G)$ is the Plancherel measure made explicit in the compact case.

74.15 Index theory

One speaks about index and subfactors in [942, 953, 1001].

74.15.1 Introduction

Let M be a type II_1 factor, which, thus, possesses an unique normal, tracial, normalized faithful state. We can assign a dimension to every modules over M by theorem 74.88. In particular, M being an algebra of operators on the Hilbert space $\mathscr{H}$, the space $\mathscr{H}$ is a M-module and we can consider the number

$$\dim_M(\mathscr{H}). \tag{74.205}$$

Following the cases, that can be any number in $]0, \infty[$. Indeed, taking any $v \in \mathscr{H}$, we have the M-module map

$$\begin{aligned} \rho \colon M &\to \mathscr{H} \\ T &\mapsto Tv, \end{aligned} \tag{74.206}$$

so that by lemma 74.104, we have

$$\dim_M(\mathscr{H}) \geqslant \dim_M M. \tag{74.207}$$

But we saw on page 3857 that $\dim_M M = \mathrm{Tr}(\mathbb{1})$, see definition (54.40) and bellow. We deduce that $\dim_M \mathscr{H} \geqslant \mathrm{Tr}(\mathbb{1})$, which can be any non vanishing positive real number.

Let $N \subseteq M \subseteq \mathfrak{B}(\mathscr{H})$ be a subfactor, i.e. that N is a factor in $\mathfrak{B}(\mathscr{H})$ and $N \subseteq M$. For the same reason as before, $\dim_N(\mathscr{H}) \in]0, \infty[$. We consider the ratio

$$[M : N] = \frac{\dim_N(\mathscr{H})}{\dim_M(\mathscr{H})} \tag{74.208}$$

that is called the **index**.

The theorem that we are intend to prove is the following.

Theorem 74.135 (Jones)**.**
If $[M : N] < 4$*, then* $[M : N] = 4\cos^2(\pi/n)$ *for some* $n < 2$*.*

Proof. No proof. □

n	$4\cos^2(\pi/n)$	
2	0	
3	1	
4	2	
5	2.618	the golden ratio
6	3	
⋮	⋮	

In the dots, we have an increasing sequence of numbers bigger than 3 that converges to 4.

74.15.2 Example

Let G be a discrete group with infinite conjugacy classes and H, a subgroup which has infinite conjugacy classes for its own right. Consider $M = M(G) \subseteq \mathfrak{B}\big(l^2(G)\big)$ and $N = M(H) \subseteq \mathfrak{B}\big(l^2(H)\big)$. By simple inclusion of H in G, we have $M(H) \subseteq \mathfrak{B}\big(l^2(G)\big)$ too.

As defined in subsection 74.7.5, the von Neumann algebra $M(H)$ is generated by the operators $\mathcal{U}_h$ with $h \in H$. We already proved in proposition 74.24 PgIDeuxGEGDutiliseIci that $l^2(G)$ is a finitely generated projective module over $M(G)$, so that one can address the question of $\dim_{M(G)}\big(l^2(G)\big)$ using the decomposition given by corollary 74.95:

$$l^2(G) = \mathrm{Cl}_{l^2(G)}\big(M(G)\big) \oplus l^2(G)/M(G). \tag{74.209}$$

On the one hand, by proposition 74.89, we have $\dim\Big(\mathrm{Cl}_{l^2(G)}\big(M(G)\big)\Big) = \dim\big(M(G)\big)$, and on the other hand, one can prove that $l^2(G)/M(G)$ is a torsion, so that

$$\dim_{M(G)}\big(l^2(G)\big) = \dim_{M(G)}\big(M(G)\big) = 1 \tag{74.210}$$ EqdimmGlDeuxGbig

when a correct choice of normalisation is done.

¡¡ Avertissement/question au lecteur !! 74.136
Is it correct that the latter dimension is in fact the trace of $\mathbb{1}$ *on* $M(G)$ *which has to be normalised?*

Let us now consider $g_1, \ldots, g_n$ be representatives of the classes of G/H and look at the decomposition

$$l^2(G) = l^2(Hg_1) \oplus \ldots \oplus l^2(Hg_n). \tag{74.211}$$

Since $l^2(Hg_i)$ is a submodule of $l^2(H)$, each of $\dim_n\big(l^2(G)\big)$ lies in the same case as equation (74.210), so that

$$\dim_N\big(l^2(G)\big) = n. \tag{74.212}$$

Thus $[M : N] = n$.

One can explicitly construct examples of index for every real bigger than 4.

Let $M \subseteq \mathfrak{B}(\mathscr{H})$ be a von Neumann algebra and form $M^{(\infty)} \subseteq \mathfrak{B}(\mathscr{H} \oplus \ldots \oplus \mathscr{H} \oplus \ldots)$,

$$M^{(\infty)} = \big\{(T, T, \ldots) \text{ tel que } T \in M\big\}. \tag{74.213}$$

Since $M^{(\infty)}$ does not contains more information than M itself, we want to be able to say that they are the same von Neumann algebra. The interest of ultraweak topology described in subsection 60.6.1 is that if we consider both M and $M^{(\infty)}$ with the ultraweak topology, then they are homeomorphic in a natural way.

74.15.3 Isomorphisms of abstract von Neumann algebras

If $M \subseteq \mathfrak{B}(\mathscr{H})$ is an ultraweakly closed subspace (what a von Neumann algebra always is), then the Hahn-Banach theorem applies and in particular, every von Neumann algebra is the dual of a space.

A positive linear map Φ between von Neumann algebras is **normal** if

$$\Phi(\bigvee_\alpha P_\alpha) = \bigvee_\alpha \Phi(P_\alpha) \tag{74.214}$$

for every increasing net of projections. A net of projections being said **increasing** when $\alpha > \beta$ implies $P_\alpha > P_\beta$. An example of a normal map is the trace of a type II_1 factor.

ThoDixLinVNanormifffuwc

Theorem 74.137.
A positive linear functional on a von Neumann algebra is normal if and only if it us ultraweakly continuous.

A proof can be found in [1002].

¡¡ Avertissement/question à la lectrice !! 74.138
It is the right citationm, isn't?

ThoPhiEquivuwcvna

Theorem 74.139.
Let $\Phi\colon M \to \mathfrak{B}(\mathscr{H})$ be a $*$-homomorphism from a von Neumann algebra into $\mathfrak{B}(\mathscr{H})$. The the following are equivalent:

ItemPhiEquivuwcvnai
(1) Φ is ultraweakly continuous,

ItemPhiEquivuwcvnaii
(2) Φ is normal,

ItemPhiEquivuwcvnaiii
(3) $\Phi(M)$ is a von Neumann algebra.

Notice that we do not suppose M to be a von Neumann algebra acting on the Hilbert space $\mathscr{H}$.

Proof. The implication (3) $\Rightarrow$ (2) is proved by using the theorem 74.137. The implication (2) $\Rightarrow$ (1) is only the fact that ultraweak continuity is defined in terms of linear functionals.

The difficult part is to prove that (1) implies (3). Since $\Phi(M)$ is a $*$-algebra, proving that it is strongly closed proves that it is a von Neumann algebra. So, let us suppose T to be in the strong closure of $\Phi(M)$. By Kaplansky density theorem 74.117, there exists a net $T_\alpha \in \Phi(M)$ with $\sup \|T_\alpha\| < \infty$ and

$$T = \text{s}\lim T_\alpha = \text{s}\lim \Phi(S_\alpha) \tag{74.215}$$

for some S_α. Since Φ is isometric, the set $\{S_\alpha\}$ is uniformly bounded: $\|S_\alpha\| < \infty$. Thus, using the Banach-Alaogu theorem 60.73, and taking a subnet, we can assume that there is a $S \in M$ such that

$$S = \text{uw}\lim S_\alpha. \tag{74.216}$$

Since Φ is ultraweakly continuous by assumption, we have $\Phi(S) = \text{uw}\lim \Phi(S_\alpha)$. Since the limit $T = \text{s}\lim T_\alpha$ exists, the same exists in the weak topology (and is equal), thus $T = \Phi(S)$, which proves that $\Phi(M)$ is strongly closed. □

So if a von Neumann algebra is algebraically realised as two different algebras of operators on two different Hilbert spaces, the two realisations are homeomorphic. For short, we say that $*$-isomorphisms of von Neumann algebras correspond to weakly continuous maps between realisations of them.

Remark that a map Φ which fulfils the theorem 74.139 is automatically isometric because $\|T\|$ is the spectral radius of $|T|$ while the spectral radius (which is a purely algebraic notion) does not change when we consider T as an operator acting on one Hilbert space or an other. So, the map Φ preserves the C^*-algebra structure.

Let M be a finite factor. We denote by $L^2(M)$ the completion of M in the norm derived from the inner product PgLdM

$$\langle T_1, T_2 \rangle = \tau(T_1 T_2^*). \tag{74.217}$$

This is the GNS construction applied to the tracial state τ. The **standard form** of the finite factor M is the realisation of M as left multiplication operators on $L^2(M)$. The image is a von Neumann algebra because the representation is normal and faithful.

74.15.4 Index of finite subfactors

We say that the vector $v \in L^2(M)$ is **bounded** by the positive real K if for every projection $P \in M$, the condition

$$\|Pv\|^2_{L^2(M)} \leqslant K\tau(P). \tag{74.218}$$

Lemma 74.140.
An element $v \in L^2(M)$ belongs to M if and only if it is bounded.

Proof. No proof. □

We consider the map

$$\begin{aligned} J\colon L^2(M) &\to L^2(M) \\ T &\mapsto T^*. \end{aligned} \tag{74.219}$$

A priori, that is only defined on M, but it is an isometry because

$$\|T\| = \mathrm{Tr}(TT^*) = \mathrm{Tr}(T^*T) = \|T^*\|. \tag{74.220}$$

So it extends to the completion $L^2(M)$.

Lemma 74.141.
We have

$$M' = JMJ \tag{74.221}$$

in the standard form.

Notice that operators in JMJ are operators of right multiplication on M because $(JTJ)S = JT(S^*) = J(TS^*) = ST^*$ while the elements of M are left multiplications, so that it is clear that $JMJ \subseteq M'$.

Proof. No proof. □

Let M_0 be a subfactor of a finite factor M_1. We define the **index**

$$[M_1 : M_0] = \dim_{M_0}\big(L^2(M_1)\big). \tag{74.222}$$

Theorem 74.142.
One has

$$[M_1 : M_0] = \dim_{M_0}(M_1). \tag{74.223}$$

and moreover

$$[M_1 : M_0] = \frac{\dim_{M_0}(\mathscr{H})}{\dim_{M_1}(\mathscr{H})} \tag{74.224}$$

if M_1 is represented on the Hilbert space $\mathscr{H}$.

Proof. No proof. □

Suppose given M_1 and a finite subfactor M_0 and suppose $[M_1 : M_0] < \infty$. Since M_0 is a subfactor of M_1, we can see M_0' as subset of $\mathfrak{B}\big(L^2(M_1)\big)$. If M_1' is not finite, then there is a part of $L^2(M_1)$ which is identified with $L^2(M_1)$ as M_0-module. That part has obviously the same dimension as the whole $L^2(M_1)$ (because the module structure is the same). Thus in this case, $L^2(M_1)$ contains infinitely many submodules with all the same dimension, so that $\dim_{M_0}\big(L^2(M_1)\big) = \infty$, which contradicts the assumption. We deduce that

$$M_0' \subseteq \mathfrak{B}\big(L^2(M_1)\big) \tag{74.225}$$

is a finite von Neumann algebra. We consider

$$P_1 \colon L^2(M_1) \to L^2(M_1) \tag{74.226}$$

be the orthogonal projection onto $L^2(M_0) \subseteq L^2(M_1)$.

LemTMunPTTPiffTMzero

Lemma 74.143.
If $T \in M_1$, then $TP_1 = P_1T$ if and only if $T \in M_0$.

Proof. It $T \in M_0$, we have

$$TP_1 = P_1TP_1. \tag{74.227}$$

EqTPPTPTProj

Taking the adjoint of that equation, $P_1T^* = P_1T^*P_1$, and using the relation (74.227) for T^*, we find $P_1T^* = T^*P_1$ which holds for every $T \in M_0$. Thus we have

$$TP_1 = P_1T \tag{74.228}$$

for every $T \in M_0$.

For the reverse sense, equality $TP_1 = P_1T$ is an equality of operators on M_1. Let us apply it to $\mathbb{1} \in M_1$. Since $\mathbb{1}$ belongs to M_0 too (every von Neumann algebra contains the unit), $P_1(\mathbb{1}) = \mathbb{1}$ and we stay with $T = P_1T$, so that $T \in M_0$. □

Let now M_2 be a von Neumann algebra generated by M_1 and P_1

$$M_2 = \{M_1, P_1\}''. \tag{74.229}$$

The properties of M_1, P_1 and M_2 are summarized in the following theorem.

Theorem 74.144.
If $[M_1 : M_0] < \infty$, then

(1) M_2 is a finite factor, in particular it has a unique normal faithful trace τ,

(2) $[M_2 : M_1] = [M_1 : M_0]$,

(3) $\tau(P_1) = [M_1 : M_0]^{-1}$.

Proof. No proof. □

The von Neumann algebra M_2 was created from M_0 and M_1, while the theorem shows that M_1 has the same properties in M_2 than M_0 in M_1, so that one can redo the construction starting from M_2 and M_1 instead of M_0 and M_1 to build M_3 and P_2. The index does not change and the projection P_2 still has the same trace.

74.15.5 Subfactors

¡¡ Avertissement/question au lecteur !! 74.145
Je pense qu'à partir d'ici, j'ai inversé M_0 et M_1 partout jusqu'à la subsection 74.15.7.

Let $M_1 \subseteq M_0$ be a subfactor of the finite factor M_0, and $P\colon L^2(M_0) \to L^2(M_0)$ be the projection onto $L^2(M_1)$. Notice that $L^2(M_1) \subseteq L^2(M_0)$ because the trace on M_1 is the restriction to the one of M_0 from unicity. Now we consider $\langle M_0, P\rangle$, the von Neumann algebra generated by M_0 and P.

LemPMNinclu

Lemma 74.146.
We have $P(M_0) \subseteq M_1$.

Proof. No proof. □

PropPropMappMN

Proposition 74.147.
The map

$$\begin{aligned} p\colon M_0 &\to M_1 \\ T &\mapsto P(T). \end{aligned} \tag{74.230}$$

satisfies

ItemPropMappMNi
(1) $p(T^*) = p(T)^*$,

ItemPropMappMNii
(2) $p(T^*T) \geqslant 0$,

ItemPropMappMNiii
(3) $p(T^*T) = 0$ *if and only if* $T = 0$,

ItemPropMappMNiv
(4) $p(S_1TS_2) = S_1p(T)S_2$ *for every* S_1 *and* S_2 *in* M_1.

for every $T \in M_0$.

Proof. The point (1) is the fact that M_1 is closed under the involution. The point (2) comes form lemma 74.146 and the positivity property in M_1. For the point (3), remark that $\mathrm{Tr}\big(p(T)\big) = \mathrm{Tr}(T)$ because $p(\mathbb{1}) = \mathbb{1}$, while by unicity up to normalisation, the trace is determined by its value on $\mathbb{1}$. □

A basic implication if (1) is that

$$PJ = JP. \tag{74.231}$$

LemPTPpTPopLdeux

Lemma 74.148.
If $T \in M_0$, then

$$PTP = p(T)P \tag{74.232}$$

as operators on $L^2(M_0)$.

Proof. No proof. □

LemNMpPNcup

Lemma 74.149.
If $M_1 \subseteq M_0$ is a subfactor of the finite factor M_0 and if P is the projection onto $L^2(M_1)$ (as operator in $L^2(M_0)$), then

$$M_1 = \{M_0' \cup P\}', \tag{74.233}$$

and

$$M_1' = \{M_0' \cup P\}'' \tag{74.234}$$

as consequence.

Proof. First, every element of $\{M_0' \cup P\}'$ has to commute with M_0' and then to belongs to $M_0'' = M_0$. But we proved in lemma 74.143 that when $T \in M_0$ and $TP = PT$, then $T \in M_1$. That proves that $\{M_0' \cup P\}' = M_1$. □

As other consequence of lemma 74.143, we have that, when $T \in M_1$, the operator $PT\colon S \mapsto P(TS) \in L^2(M_0)$ is equal to the operator $S\colon \mapsto TP(S)$, in other words,

$$P(TS) = TP(S) \tag{74.235}$$

EqPTSeqalTPS

for every $T \in M_1$ and $S \in M_0$.

LemMOJNJequal

Lemma 74.150.
We have

$$\langle M_0, P\rangle = JM_1'J, \tag{74.236}$$

EqLemMPJNJequal

and as consequence, $\langle M_0, P\rangle$ is a finite factor.

Proof. First note that the commutant of a factor is a factor, so that M_1' is a factor and $JM_1'J$ is a factor because it is algebraically isomorphic to the factor M_1'. Thus the lemma proves that the left hand side of (74.236) is in particular a finite factor.

Let us now prove the equality (74.236). One needs to show that $J\langle M_0, P\rangle J = M_1'$. The commutant of the left multiplication action is the right one, while the left action of JTJ is the right action of T, so one has $JM_0J = M_0'$. Since $JPJ = P$, it is sufficient to show that

$$\langle M_0', P\rangle = M_1'. \tag{74.237}$$

The von Neumann algebra generated by M_0' and P is $\{M_0', P\}''$ that is equal to M_1' by lemma 74.149. □

So $\langle M_0, P\rangle$ is a factor and it is finite when M_1' is finite, which happens when $[M_0 : M_1] < \infty$. Assuming that finiteness, we thus have a trace $\mathrm{Tr}\colon \langle M_0, P\rangle \to \mathbb{C}$ which extends the trace on M_0 (by uniqueness).

Notice that, as particular case of equation (74.236) when $M_1 = M_0$, we have $P = \mathrm{Id}$ and

$$M_0 = JM_0'J. \tag{74.238}$$

That gives us an action of M_0' on $L^2(M_0)$ by

$$S \cdot T = (JSJ)T \tag{74.239}$$

EqSdotTTSJT

when $S \in M_0'$ and $T \in M_0$.

PropdimMQhHQDIMMhH

Proposition 74.151.
If M_0 is any finite factor acting on $\mathscr{H}$ and if $Q \in M_0'$, then $Q\mathscr{H}$ is invariant by M_0, and

$$\dim_{M_0}(Q\mathscr{H}) = \mathrm{Tr}(Q)\dim_{M_0}(\mathscr{H}) \tag{74.240}$$

where $Q\mathscr{H}$ is seen as module over M_0 (because of its invariance).

Proof. First, notice that a finite factor always has an unique trace. If $Q = 0$ or $Q = \mathbb{1}$, the claim is obvious. Suppose now that $\mathrm{Tr}(Q) = r/s$, that is a rational number.

In that case, we have

$$\underbrace{Q\mathscr{H} \oplus \ldots \oplus Q\mathscr{H}}_{s \text{ times}} = \underbrace{\mathscr{H} \oplus \ldots \oplus \mathscr{H}}_{r \text{ times}} \tag{74.241}$$

EqQhQhHHeqMmod

¡¡ Avertissement/question à la lectrice !! 74.152
I can understad that in the setting of the fractional dimensions as described in subsection 74.10.3, but there, we have the assumption of no minimal projection.

First, notice that the trace of the identity over $Q\mathscr{H}$ is the trace of Q on $\mathscr{H}$. Indeed, let $\{v_i\}$ be an Hilbertian basis of $Q\mathscr{H}$ and $\{w_k\}$ be a one of $Q^\perp\mathscr{H}$. Then

$$\mathrm{Tr}(Q) = \sum_i \langle v_i, Qv_i\rangle + \sum_k \langle w_k, Qw_k\rangle = \sum_i \langle v_i, v_i\rangle = \mathrm{Tr}(\mathbb{1}_{Q\mathscr{H}}). \tag{74.242}$$

Thus, by additivity,

$$\operatorname{Tr}(\mathbb{1}_{Q\mathscr{H}\oplus\dots\oplus Q\mathscr{H}}) = s\operatorname{Tr}(Q) = r. \tag{74.243}$$

On the other hand, $\operatorname{Tr}(\mathbb{1}_{\mathscr{H}\oplus\dots\oplus\mathscr{H}}) = r\operatorname{Tr}(\mathbb{1}_{\mathscr{H}}) = r$. So the identity over these two M_0-modules are the same.

¡¡ Avertissement/question à la lectrice !! 74.153
Is that enough to deduce that these are isomorphic as M_0-modules?

Taking the dimension of both sides of (74.241), we find $s\dim_{M_0}(Q\mathscr{H}) = r\dim_{M_0}(\mathscr{H})$, which proves the statement in the case of $\operatorname{Tr}(Q) \in \mathbb{Q}$.

Now, the functions $\dim_{M_0}(Q\mathscr{H})$ and $\operatorname{Tr}(Q)\dim_{M_0}(\mathscr{H})$ are increasing functions of $\operatorname{Tr}(Q)$ that coincide on $\mathbb{Q}$. Thus they are equal on $\mathbb{R}$.

□

ProptrPTMNtrT

Proposition 74.154.
We have the formula

$$\operatorname{Tr}(PT) = \frac{1}{[M_0 : M_1]}\operatorname{Tr}(T) \tag{74.244}$$

EqClaimLemPTfracMNtrT

for every $T \in M_0$.

Proof. We have $\operatorname{Tr}(PT) = \operatorname{Tr}(PPT) = \operatorname{Tr}(PTP) = \operatorname{Tr}\big(p(T)P\big)$ because of lemma 74.148. It is sufficient to prove the claim (74.244) for $T \in M_1$ because when $S \in M_0$, we would have $\operatorname{Tr}(PS) = \operatorname{Tr}\big(p(S)P\big) = \operatorname{Tr}(S)/[M_0 : M_1]$ because $p(S) \in M_1$. Let thus $T \in M_1$ and consider the map

$$T \mapsto \operatorname{Tr}(PT) \tag{74.245}$$

where PT is considered as operator on $L^2(M_0)$. Thanks to (74.235), this is a (normal faithful) trace on M_0. So that must be a multiple of the trace on M_0, namely there exists a $\tau \in \mathbb{C}$ such that $\operatorname{Tr}(PT) = \tau\operatorname{Tr}(T)$. Taking that equality with $T = \mathbb{1}$ shows that $\tau = \operatorname{Tr}(P)$. Using proposition 74.151, we have thus

$$\dim_{M_1}\big(P \cdot L^2(M_0)\big) = \operatorname{Tr}(P)\dim_{M_0}\big(L^2(M_0)\big), \tag{74.246}$$

but, as M_0-module, we have $P \cdot L^2(M_0) = L^2(M_1)$ by definition of P, so we have

$$1 = \operatorname{Tr}(P) \cdot \dim_{M_1}\big(L^2(M_0)\big). \tag{74.247}$$

□

The set of expressions

$$T_0 + \sum_{j=1}^{n} T_{j_1} P T_{j_2} \tag{74.248}$$

with T_0, T_{j_1} and T_{j_2} in M_0 is a $*$-algebra. Indeed, in the multiplication of two such expressions, we get monomials of the form

$$T_1 \underbrace{PT_2P}_{=p(T_2)P} T_3 = \underbrace{T_1p(T_2)}_{\in M_0} PT_3. \tag{74.249}$$

By the double commutant theorem, the von Neumann algebra $\langle M_0, P\rangle$ is the strong closure of such expressions, and

$$\operatorname{Tr}(T_1PT_2) = \operatorname{Tr}(PT_2T_1) = \frac{1}{[M_0 : M_1]}\operatorname{Tr}(T_1T_2) \tag{74.250}$$

by proposition 74.154. Thus we have an explicit formula for the trace on $\langle M_0, P\rangle$ from the trace over M_0.

lemPhHPNPfrac

Lemma 74.155.
If $P \in M_1$, then we have

$$\dim_{PM_1P}(P\mathscr{H}) = \frac{1}{\mathrm{Tr}(P)} \dim_{M_1}(\mathscr{H}), \tag{74.251}$$

but $P\mathscr{H}$ is no more a M_1-module.

Proof. Let $\mathcal{E}$ be any M_1-module. Since $P\mathcal{E}$ is a PM_1P-module, we can define $\dim^{(P)}(\mathcal{E})$ by

$$\dim^{(P)}(\mathcal{E}) = \dim_{PM_1P}(P\mathcal{E}). \tag{74.252}$$

That defines a dimension function on the M_1-modules, which is thus a multiple of the standard one. We know that, as M_1-module, M_1P as dimension $\mathrm{Tr}(P)$, so that the proportionality factor can be fixed on M_1P.

¡¡ Avertissement/question à la lectrice !! 74.156
That proof is not complete.

□

Let take $M_1 = \mathbb{M}_3(\mathbb{C})$ and $\mathscr{H} = \mathbb{C}^3$ as example. Each column of $a \in M_1$ is an element of $\mathbb{C}^3$, so M_1 is three copies of $\mathscr{H}$ and, as left module, we have $\mathscr{H} \oplus \mathscr{H} \oplus \mathscr{H} = N$, and $\dim_{M_1}(\mathscr{H}) = 1/3$. Let

$$P = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}.$$

We have $P\mathscr{H} \simeq \mathbb{C}$, and $PM_1P = \mathbb{C}$ (the upper left element of the matrix). Thus if we have $\mathrm{Tr}(P) = 1/3$, we conclude by the lemma 74.155 that $\dim_{PM_1P}(P\mathscr{H}) = 1$.

Lemma 74.157.
We have

$$\dim_{M_1}(\mathscr{H}) \dim_{M_1'}(\mathscr{H}) = 1 \tag{74.253}$$

if M_1 and M_1' are finite factors.

Proof. Let us first check for $\mathscr{H} = L^2(M_1)$. The action of M_1' on $L^2(M_1)$ is given by (74.239), so we have $\dim_{M_1'}\big(L^2(M_1)\big) = \dim_{JM_1J}\big(L^2(M_1)\big)$, so that the action of M_1' on $L^2(M_1)$ is conjugate of the one of M_1. That proves the lemma in the case where $\mathscr{H} = L^2(M_0)$.

¡¡ Avertissement/question à la lectrice !! 74.158
Pourquoi ? Cela est fait dans la proposition 74.84. Et je crois que du côté de cette proposition, on trouve pas mal de réponses à pas mal de questions ici.

□

Corollary 74.159.
We have

$$[M_0 : M_1] = [M_1' : M_0'] \tag{74.254}$$

when $M_0 \in \mathfrak{B}(\mathscr{H})$ is a factor of type II_1 and M_1 is a subfactor of M_0.

Proof. No proof. □

Theorem 74.160.
We have

$$[\langle M_0, P\rangle : M_0] = [M_0 : M_1] \tag{74.255}$$

under the same assumptions.

Proof. If $\mathcal{E}$ is any representation of $\langle M_0, P\rangle$, we have the formula

$$[\langle M_0, P\rangle : M_0] = \frac{\dim_{M_0} \mathcal{E}}{\dim_{\langle M_0,P\rangle} \mathcal{E}}. \tag{74.256}$$

EqlangleMOmodEME

We know by formula (74.110) that

$$\Big(\dim_{\langle M_0,P\rangle} \mathcal{E}\Big) \cdot \Big(\dim_{\langle M_0,P\rangle'} \mathcal{E}\Big) = 1. \tag{74.257}$$

On the other hand, since $\langle M_0, P\rangle = JM_1'J$ (lemma 74.150), we have

$$\dim_{\langle M_0,P\rangle'} \mathcal{E} = \dim_{JM_1J} \mathcal{E} = \dim_{M_1} \mathcal{E} = [M_0 : M_1]. \tag{74.258}$$

Since formula (74.256) holds for every choice of $\mathcal{E}$, we compute the right hand side in the case where $\mathcal{E} = L^2(M_0)$. In that particular case, $\dim_{M_0} \mathcal{E} = 1$, and the claim follows. □

74.15.6 Example of index bigger than 4

Consider R, the hyperfinite factor of type II_1 given by [12]

$$R = \bigotimes_1^\infty \big(\mathbb{M}_2(\mathbb{C}), \mathrm{Tr}\,\big) \tag{74.259}$$

It turns out that, for that factor, we have $R \simeq PRP$ for every projection $P \in R$. Choose isomorphism $\alpha\colon R \to PRP$ and $\beta\colon R \to P^\perp RP^\perp$, and form the algebra

$$S = \{\alpha(T) + \beta(T) \text{ tel que } T \in R\}. \tag{74.260}$$

That algebra is algebraically isomorphic to R, and the isomorphism is ultraweakly continuous, so that S is a von Neumann algebra. The index of S in R is given by

$$[R : S] = \mathrm{Tr}(P)^{-1} + \mathrm{Tr}(P^\perp)^{-1} \tag{74.261}$$

where $\mathrm{Tr}(P)$ can take any value between 0 and 1, and $\mathrm{Tr}(P^\perp) = 1 - \mathrm{Tr}(P)$. The possible values of the index are then given by the range of the function

$$f(x) = \frac{1}{x} + \frac{1}{1-x} \tag{74.262}$$

when x runs over $[0, 1]$. One easily checks that that range is $[4, \infty]$.

74.15.7 Properties of the sequence of M_i, P_i

SubSecPropSeqMO

Let M_0 be a subfactor of M_1. We define

$$[M_1 : M_0] = \dim_{M_0} \big(L^2(M_1)\big) \tag{74.263}$$

and we consider $P_1\colon L^2(M_1) \to L^2(M_1)$, the projection onto $L^2(M_0)$. Then we consider the new factor

$$M_2 = \langle M_1, P_1\rangle. \tag{74.264}$$

For that definition, we see M_1 as subalgebra of $\mathfrak{B}\big(L^2(M_1)\big)$. Now, M_2 is also a subalgebra of $\mathfrak{B}\big(L^2(M_2)\big)$, so that, defining $P_2\colon L^2(M_2) \to L^2(M_2)$ as the projection onto $L^2(M_1)$, allows to define

$$M_3 = \langle M_2, P_2\rangle \subseteq \mathfrak{B}\big(L^2(M_2)\big). \tag{74.265}$$

Using that construction again and again, we get a sequence

$$M_0 \subseteq M_1 \subseteq M_2 \subseteq \dots \tag{74.266}$$

12. Recall that one needs a functional in order to define the infinite tensor product.

of factors defined by $M_{n+1} = \langle M_n, P_n \rangle$ where $P_n \colon L^2(M_n) \to L^2(M_n)$ is the orthogonal projection onto $L^2(M_{n-1})$. One can prove that

$$P_n(M_n) \subseteq M_{n-1}. \tag{74.267}$$

We thus consider the maps

$$\begin{aligned} p_n \colon M_n &\to M_{n-1} \\ p_n(T_n) &\mapsto P_n(T_n), \end{aligned} \tag{74.268}$$

and we have the sequence

$$P_{n-1} \in M_n \xrightarrow{p_n} M_{n-1} \xrightarrow{p_{n-1}} M_{n-2}. \tag{74.269}$$

The following is the proposition 3.3.2 in [1001].

PropropsindexMarkov

Proposition 74.161.
We have

(1) $[M_n : M_{n+1}] = \lambda^{-1}$ *does not depend on* n,
(2) $\operatorname{Tr}(P_n) = \lambda$,
(3) $\operatorname{Tr}(P_n T_n) = \lambda \operatorname{Tr}(T_n)$ *for every* $T_n \in M_n$.

The last property is the **Markov property**.

Proof. No proof. □

LemPnPllamcun

Lemma 74.162.
We have

$$p_n(P_{n-1}) = \lambda \mathbb{1} \tag{74.270}$$

where $\lambda = \operatorname{Tr}(P_n)$ *does not depends on* n *by the proposition 74.161.*

Proof. Since the map $(S,T) \mapsto \operatorname{Tr}(ST)$ is a nondegenerate bilinear functional on M_{n-1}, it is sufficient to prove that $\operatorname{Tr}\big(Sp_n(P_{n-1})\big) = \lambda \operatorname{Tr}(S)$ for every $S \in M_{n-1}$. Using point (4) of proposition 74.147, we have $Sp_n(P_{n-1}) = p_n(SP_{n-1})$, so that

$$\operatorname{Tr}\big(Sp_n(P_{n-1})\big) = \operatorname{Tr}(SP_{n-1}) = \lambda \operatorname{Tr}(S), \tag{74.271}$$

where the last equality is the Markov property. □

PropAlgPPPKoi

Proposition 74.163.
The projections P_i *fulfil the algebra* SubeqPnPalgPPI

$$P_n P_{n+1} P_n = \lambda P_n \tag{74.272a}$$ EqLoiPPun
$$P_n P_{n-1} P_n = \lambda P_n \tag{74.272b}$$ EqLoiPPdeux
$$P_n P_m = P_m P_n \tag{74.272c}$$ EqLoiPPtrois

when $|n-m| \geqslant 2$.

Proof. Let us prove the second one. Using lemma 74.162, we find

$$P_n P_{n-1} P_n = p_n(P_{n-1}) P_n = \lambda P_n, \tag{74.273}$$

which is the claim. □

Corollary 74.164.
The algebra generated by

$$\{\mathbb{1}, P_1, \dots, P_n\} \tag{74.274}$$

is a finite dimensional C^**-algebra in which there exists an unique tracial state such that for every* k,

$$\operatorname{Tr}(P_k T) = \lambda \operatorname{Tr}(T) \tag{74.275}$$

whenever T *belongs to the algebra generated by* $\mathcal{S}_{k-1} = \{\mathbb{1}, P_1, \dots, P_{k-1}\}$.

Sketch of the proof. Let a *reduced word* in $P_1, \dots, P_k$ be a word which is as small as possible using the three rules (74.272). We prove by induction that such a reduced word contains P_k at most once. For beginning, the only reduced word in P_1 is P_1 itself.

Now, let a word containing twice P_k. What is between two successive occurrences of P_k is a word of $\mathcal{S}_{k-1}$. That word is reduced, and thus contains P_{k-1} only once by induction hypothesis. From rule (74.272c), the operator P_k commutes with $\mathcal{S}_{k-2}$, and then the rule (74.272b) reduces the word (because P_{k-1} appears only once).

Now, a general reduced word of $\mathcal{S}_k$ has only one P_k and thus has the form $m_1 P_k m_2$ where m_1 and m_2 are reduces words of $\mathcal{S}_{k-1}$. Of course, m_1 them self decomposes in $n_1 P_{k-1} n_2$ where n_i are reduces words of $\mathcal{S}_{k-2}$, and the process continues.

Every reduced word in $\mathcal{S}_9$ look like

$$(P_5P_4P_3)(P_8P_7P_6P_5P_4)(P_9P_8P_7P_6). \tag{74.276}$$

Some comments

— The sub words are made of consecutive projections. For example the sub word (P_5P_3) can be rearranged as the two sub words $(P_3)(P_5)$ by rule (74.272c).

— The sub words begin by P_5, P_8 and P_9. Notice that $5 < 8 < 9$. It will always be like that. If not, a rearrangement is possible.

Now fix a n and consider the set of all the reduced words of $\mathcal{S}_n$; there are only finitely many of them which will be labelled as W_i. We look at the matrix whose element ij is given by

$$\mathrm{Tr}(W_i^* W_j). \tag{74.277}$$

EqMtrWWr

Each of $W_i^* W_j$ is a long string of elements of $\mathcal{S}_n$ and the cyclic property of the trace allows us to do

$$\mathrm{Tr}\big((\dots)(\dots)(P_nP_{n-1}\dots)\big) = \mathrm{Tr}\big((P_nP_{n-1}\dots)(\dots)(\dots)\big). \tag{74.278}$$

Then, using the Markov property of proposition 74.161, we can extract P_n. That matrix is positive semidefinite because it is formed of scalar products. If not, that would means that the assumption $M_0 \subseteq M_1$ is wrong.

□

74.15.8 Some interesting functions

Consider the polynomials defined by the induction

$$f_0(x) = 0 \tag{74.279a}$$
$$f_1(x) = 1 \tag{74.279b}$$
$$f_n(x) = f_{n-1}(x) - x f_{n-2}(x). \tag{74.279c}$$

Since $f_n(0) = f_{n-1}(0)$, all these polynomials pass by the point $(0, 1)$. By induction, one can see that

$$f_n\left(\frac{1}{4\cos^2\theta}\right) = \frac{\sin(n\theta)}{2^{n-1}\cos^{n-1}(\theta)\sin(\theta)}, \tag{74.280}$$

so that the smallest positive root of f_n is

$$\frac{1}{4\cos^2\left(\frac{\pi}{n}\right)}, \tag{74.281}$$

and f_n is negative on the interval between $1/4\cos^2(\pi/n)$ and $1/4\cos^2(\pi/(n-1))$.

Suppose $\lambda^{-1} = [M : M_0]$ and $\lambda > 1/4$. Now, assume that

$$\lambda \neq \frac{1}{4\cos^2(\pi/n)} \tag{74.282}$$

for any value of $n \in \mathbb{N}$. In this case, we will found a negative diagonal entry of the matrix (74.277), which is impossible because a semipositive definite matrix has no negative diagonal element.

Let

$$Q_k = \mathbb{1} - P_1 \vee \ldots \vee P_k. \tag{74.283}$$

From lemma 72.12, the operator Q_k is in fact a projection in a finite dimensional algebra, and is thus a polynomial in the P_i's.

Lemma 74.165.
We have

$$(Q_{k-1}P_kQ_{k-1})^2 = \frac{f_k(\lambda)}{f_{k+1}(\lambda)}Q_{k-1}P + kQ_{k-1}, \tag{74.284}$$

and

$$Q_{k+1} = Q_k - \frac{f_{k+1}(\lambda)}{f_k(\lambda)}Q_{k-1}Q_{k-1}, \tag{74.285}$$

if $f_{k+1}(\lambda), \ldots, f_1(\lambda) \neq 0$.

Notice in particular that, up to a factor, the operator of equation (74.284) is a projection.

Proof. No proof. □

Corollary 74.166.
The operators Q_{k-1}, P_k and Q_{k+1} are projection, and the product is a positive operator.

That corollary shows that

$$\frac{f_k(\lambda)}{f_{k+1}(\lambda)} > 0. \tag{74.286}$$

But a simple study of the polynomials shows that this ratio is in fact negative. We conclude that operator satisfying the algebra of proposition (74.163) are only possible when $\lambda = 1/4\cos(\pi/n)$.

74.16 Modular theory

74.16.1 Modular isometry

Let M be a von Neumann algebra provided with φ, a faithful normal state. We consider $L^2_\varphi(M)$, the completion of M for the norm associated with the inner product

$$\langle T_1, T_2 \rangle = \varphi(T_1^* T_2). \tag{74.287}$$

That product is conjugate-linear in his first argument and linear in the second. The algebra M acts at left by multiplication on $L^2_\varphi(M)$, by extending

$$T \cdot S = TS. \tag{74.288}$$

The space $L^2_\varphi(M)$ has a distinguished vector: the one represented by $\mathbb{1} \in M$ that we call v. It is a cyclic vector for M, i.e. the space Mv is dense in $L^2_\varphi(M)$. It is also separating, because $Tv = 0$ with $T \in M$ implies $T = 0$.

Conversely, if $\mathscr{H}$ is an Hilbert space and $v \in \mathscr{H}$ is cyclic and separating for a von Neumann algebra, then the expression

$$\varphi(T) = \langle v, Tv \rangle \tag{74.289}$$

defines a faithful (because separating) normal state on the von Neumann algebra. Thus, in fact $\mathscr{H} \simeq L^2_\varphi(M)$ via an unitary isomorphism compatible with M.

A natural question arising is to know the commutant of M acting by left multiplication on $L^2_\varphi(M)$. Let us suppose that φ is also a trace, so we drot the index and we write $L^2(M)$ instead of $L^2_\varphi(M)$. We define

$$\begin{aligned} J \colon L^2(M) &\to L^2(M) \\ Tv &\mapsto T^* v. \end{aligned} \tag{74.290}$$

That definition is correct because $\{Tv \text{ tel que } T \in M\}$ is dense in $L^2(M)$. Notice that $T_1 v = T_2 v$ implies $T_1 = T_2$, so that, when $\xi \in L^2(M)$ can be represented as Tv, that $T \in M$ is unique.

The following lemma provides the main properties of J.

LemPropJ

Lemma 74.167.
The map J has the following properties.

(1) J is conjugate-linear. ItemPropJii

(2) It is an isometry in the sense of conjugate linear operators.

(3) $J^2 = \mathbb{1}$. ItemPropJiv

(4) $JMJ \subseteq M'$. ItemPropJv

(5) $\langle Sv, JTv\rangle = \langle JS^*v, T^*v\rangle$ for every S and T in M.

Proof. The proof of (2) is a computation using the definition of J, the definition of the inner product and the fact that φ is a state:

$$\langle JT_1v, JT_2v\rangle = \langle T_1^*v, T_2^*v\rangle\varphi(T_1T_2^*) = \overline{\varphi(T_2T_1^*)} = \overline{\varphi(T_1^*T_2)} = \overline{\langle T_1v, T_2v\rangle}. \tag{74.291}$$

For (4), remark that

$$(JTJ)(Sv) = JT(S^*v) + J(TS^*v) = ST^*v. \tag{74.292}$$

□

Definition 74.168.
*A **modular isometry** for M with respect to the cyclic vector v is a map $J\colon L^2(M) \to L^2(M)$ which satisfies*

(1) $Jv = v$.,

(2) J is a conjugate linear isometry,

(3) $J^2 = 1$,

(4) $JMJ \subseteq M'$,

(5) $\langle JSv, Tv\rangle = \langle S^*v, JT^*v\rangle$ for every S, $T \in M'$.

Notice the difference between the claim (5) of lemma 74.167.

Lemma 74.169.
If v is cyclic for M, then it is separating for M', and if it is separating for M, then it is cyclic for M'.

Proposition 74.170.
If J is a modular isometry, then $JMJ = M'$. In particular, M and M' are of same type.

Proof. We have to prove that $T \in JMJ$ whenever $T \in M'$. Let S and T be in M', and let us prove that JSJ and T commute. For that, we are going to prove that

$$\langle JSJTv, v\rangle = \langle TJSJv, v\rangle, \tag{74.293}$$

which is sufficient because the product is nondegenerate. For all S and T in M', one successively has

$$\begin{aligned}\langle JSJTv, v\rangle = \langle JSJTv, J^2v\rangle = \langle Jv, SJTv\rangle = \langle v, SJTv\rangle \\ = \langle S^*v, JTv\rangle = \langle JSv, T^*v\rangle = \langle TJSJv, v\rangle,\end{aligned} \tag{74.294}$$

where we used the extra property $Jv = v$ which defines a modular isometry.

Now we repeat the same with $S_1 = A^*S$ and $T_1 = TB$ where $A \in JMJ \subseteq M'$ and $B \in JMJ$. The operators S_1 and T_1 belong to M' because S and T belong to M'. We find

$$\langle JSJT(Bv), (JAv)\rangle = \langle TSJS(Bv), (JAv)\rangle. \tag{74.295}$$

Using the fact that v is cyclic, the vectors Bv and JAv separately span dense subspaces, so that, as operators, $TSJT = TJSJ$ because they are equal on a dense subspace. □

Proposition 74.171.
If J is a modular isometry for M (with respect to v), then it is a modular isometry for M' (with respect to v) too.

Proof. The fact that $JM'J = M$ is obtained by application of J at left and right of both sides of $JMJ = M'$. For the last one, compute

$$\langle Sw, Tw\rangle = \langle w, S^*Tw\rangle = \langle w, TS^*w\rangle = \langle T^*w, S^*\rangle \tag{74.296}$$

for every $S \in M$ and $T \in M'$. □

Proposition 74.172.
If $\varphi(T) = \langle v, Tv\rangle$ is a trace, then $JSv = S^*v$ for every $S \in M'$.

Proof. When T runs over M, the element Tv spans a dense space, and we have, for $T \in M$ and $S \in M'$:

$$\langle JSv, Tv\rangle = \langle JTv, Sv\rangle = \langle T^*v, Sv\rangle = \langle v, TSv\rangle = \langle v, STv\rangle = \langle S^*v, Tv\rangle. \tag{74.297}$$

So the inner product of JSv and S^*v on any Tv are the same. That proves that $JSv = S^*v$. □

Theorem 74.173.
For every von Neumann algebra M and every cyclic and separating vector v, there exists a modular isometry.

Proof. No proof. □

Corollary 74.174.
The algebras M and M' are always algebraically isomorphic.

74.16.2 Example

Let $M(G)$ where G is a non unimodular locally compact group. Thus $M(G) = C_c(G)''$. A function $f \in C_c(G)$ defines an operator on $L^2(G)$ by

$$(f\xi)(g) = \int_G f(h)\xi(h^{-1}g)\, d\mu(h) \tag{74.298}$$

where μ is a left Haar measure. We consider the functional φ on $M(G)$

$$\varphi(f) = f(e). \tag{74.299}$$

If G was unimodular, this would be a tracial weight, but in the non unimodular case, φ is not even a trace on $C_c(G)$.

74.16.3 Other example

Let G be a discrete group with counting measure. Suppose that G_0 is an **almost normal** of G, that means that each double coset G_0gG_0 is a finite union of left cosets [13]. We define

$$A = \mathbb{C}[G, G_0] = \{f \colon G_0\backslash G/G_0 \to \mathbb{C} \text{ such that } f \text{ is finitely supported}\}, \tag{74.300}$$

and we introduce the product

$$(f_1 \star f_2)(g) = \sum_{h \in G_0/G} f_1(gh^{-1})f_2(h), \tag{74.301}$$

13. In the case of a normal subgroup, the double cosets are made of only one left coset.

and the involution $f^*(g) = \overline{f(g^{-1})}$. The map

$$\begin{aligned} \varphi\colon \mathbb{C}[G, G_0] &\to \mathbb{C} \\ f &\mapsto f(e) \end{aligned} \tag{74.302}$$

is not a trace in the general case. Since $\varphi(f^*f) \geqslant 0$ and $\varphi(f^*f) = 0$ implies $f = 0$, we conclude that φ is a state.

In the case where G_0 is a normal subgroup, however, it is a trace. We define

$$\begin{aligned} \sigma_z\colon \mathbb{C}[G, G_0] &\to \mathbb{C}[G, G_0] \\ \big(\sigma_z(f)\big)(g) &= \Delta(g)^z f(g) \end{aligned} \tag{74.303}$$

where $\Delta(g) = L(g)/R(g)$ where $L(g)$ is the number of left cosets in $G_0 g G_0$ and $R(g)$ is the number of left cosets in $G_0 g G_0$.

Lemma 74.175.
We have

$$\varphi(f_1 f_2) = \varphi\big(f_2 \sigma(f_1)\big) \tag{74.304}$$

where $\sigma = \sigma_1$. The map σ_z is, moreover, an automorphism.

That lemma says that the non triviality of σ measures at what extend φ is not a trace.

Proof. No proof. □

If φ is a trace, then the class of identity in $\mathbb{C}[G, G_0]$ is cyclic and one can make a GNS construction. Notice that σ_z is not a $*$-automorphism because

$$\sigma_z(f^*) = \sigma_{-\bar{z}}(f)^*. \tag{74.305}$$

If z is purely imaginary, σ_z is a $*$-automorphism.

Let $\mathscr{H}$ be the completion of A with respect to the inner product defined by the state φ. The algebra A acts on $\mathscr{H}$ by bounded operators of left multiplication [14]. Let $v \in \mathscr{H}$ be defined by

$$v(g) = \begin{cases} 1 & \text{if } g \in G_0 \\ 0 & \text{else,} \end{cases} \tag{74.306}$$

and let $\mathscr{H}_0 = A \cdot v \subseteq \mathscr{H}$.

Lemma 74.176.
We have

$$\|\sigma_{1/2}(f)^* v\| = \|fv\|, \tag{74.307}$$

in other words, $f \mapsto \sigma_{1/2}(f^*)$ is a conjugate linear isometry.

Proof. First, notice that for every $h \in A$ and $z \in \mathbb{C}$,

$$\varphi\big(\sigma_z(h)\big) = \varphi(h). \tag{74.308}$$

Eqsigmazvarhib

We have

$$\|\sigma_{1/2}(f^*)v\|^2 = \varphi\big(\sigma_{1/2}(f^*)^* \sigma_{1/2}(f^*)\big) = \varphi\big(\sigma_{-1/2}(f)\sigma_{1/2}(f^*)\big). \tag{74.309}$$

Now we apply the relation (74.308) with $z = -1/2$, and we find that the latter line is equal to

$$\varphi\big(\sigma_{-1}(f)\sigma_0(f^*)\big) = \varphi\big(\sigma_{-1}(f) f^*\big) = \varphi\Big(f^* \sigma_1\big(\sigma_{-1}(f)\big)\Big) = \varphi(f^* f) = \|fv\|^2. \tag{74.310}$$

□

14. Since A is not a C^*-algebra, we cannot invoke a GNS construction here.

Let $M \subseteq \mathfrak{B}(\mathscr{H})$ be the von Neumann algebra generated by A.

Proposition 74.177.
The formula

$$J(fv) = \sigma_{1/2}(f^*)v \tag{74.311}$$

defines a modular isometry for M with respect to v.

Proof. No proof. □

74.16.4 Still an other example

Let $A = \mathbb{M}_n(\mathbb{C})$ and

$$\varphi\colon S \mapsto \operatorname{Tr}(ST) \tag{74.312}$$

where T is invertible and chosen in such a way that $T \geqslant 0$, $\operatorname{Tr}(T) = 1$, so that φ is a positive faithful state. Let now $\mathscr{H}$ be the GNS Hilbert space and consider the map

$$\sigma_z = T^z S T^{-z}. \tag{74.313}$$

We have

$$\varphi(S_1S_2) = \operatorname{Tr}(S_1S_2T) = \operatorname{Tr}(S_2TS_1) = \operatorname{Tr}(S_2TS_1T^{-1}T) = \operatorname{Tr}\big(S_2\sigma(S_1)T\big) = \varphi\big(S_2\sigma(S_1)\big), \tag{74.314}$$

thus σ_1 has the same property as before to measure the lack of tracial property of φ.

Once again, we pick v, the vacuum of the GNS representation, i.e. the representative of the identity. The map

$$J\colon Sv \mapsto \sigma_{1/2}(S^*)v \tag{74.315}$$

is a modular isometry for M with respect to v.

74.16.5 Tomita's theorem

Lemma 74.178.
The map

$$T\xi \mapsto T^*\xi \tag{74.316}$$

EqApplQuiDoitEtreIsom

is isometric if and only if ξ is a trace vector.

Proof. The fact for the map (74.316) to be an isometry means that $\langle T\xi, T\xi\rangle = \langle T^*\xi, T^*\xi\rangle$. This is equivalent to

$$\langle \xi, T^*T\xi\rangle = \langle \xi, TT^*\xi\rangle, \tag{74.317}$$

which means that ξ is a trace vector. □

Theorem 74.179.
If v is a cyclic and separating vector for $M \subseteq \mathfrak{B}(\mathscr{H})$, and if

$$\varphi(T) = \langle v, Tv\rangle, \tag{74.318}$$

then there is a strongly dense $*$-subalgebra $A \subseteq M$ and a one parameter group of automorphisms $\sigma_z\colon A \to A$ ($z \in \mathbb{C}$) such that

(1) The map $z \mapsto \sigma_z(S)$ is holomorphic as map from $\mathbb{C}$ into $\mathfrak{B}(\mathscr{H})$ in the sense that it accepts a power expansion $\sigma_z = T_0 + zT_1 + z^2T_2 + \dots$

ItemTomitaii

(2) If $z \in i\mathbb{R}$, then σ_z extends to a strongly continuous one parameter group of automorphisms on M which only depends on φ when the other conditions are satisfied.

(3) The maps φ and σ are related by

$$\varphi(ST) = \varphi\big(T\sigma(S)\big) \tag{74.319}$$

for every S, $T \in A$, where $\sigma = \sigma_1$.

Moreover the formula

$$J\colon Sv \mapsto \sigma_{1/2}(S^*)v \tag{74.320}$$

defines a modular isometry for M with respect to v.

Notice that the point (2) does not depend on A.

Idea of the proof. We ignore the analytical issues. We define

$$\begin{aligned} \mathcal{S}\colon \mathscr{H} &\to \mathscr{H} \\ Tv &\mapsto T^*v. \end{aligned} \tag{74.321}$$

For a general v, this is not an isometry (this is even unbounded), but we can try to have in idea of what is $\mathcal{S}^*$. Since $\mathcal{S}$ is conjugate linear, the equation to solve is

$$\langle \mathcal{S}T_1v, T_2v\rangle = \langle \mathcal{S}^*T_2v, T_1v\rangle. \tag{74.322}$$

We have $\langle T_1^*v, T_2v\rangle = \langle \mathcal{S}^*T_2v, T_1v\rangle$. Since the Hilbert space $\mathscr{H}$ is build from a GNS construction, the inner product is in fact given in terms of φ:

$$\langle T_1^*v, T_2v\rangle = \varphi(T_1T_2) = \varphi\big(T_2\sigma(T_1)\big) = \varphi\big(\sigma_{-1}(T_2)T_1\big) = \langle \sigma_{-1}(T_2)^*v, T_1v\rangle = \langle \sigma(T_2^*)v, T_1c\rangle, \tag{74.323}$$

so that

$$\mathcal{S}^*(Tv) = \sigma(T^*)v, \tag{74.324}$$

which is an equation for $\mathcal{S}^*$ in the same time as for σ. After computations, one shows that

$$\mathcal{S}^*\mathcal{S}(Tv) = \sigma_{-1}(T)v. \tag{74.325}$$

Conclusion: the automorphism σ_{-1} (which is the inverse of σ_1) can be deduced from $\mathcal{S}$. Moreover, one can prove that $\mathcal{S}^*\mathcal{S}$ is a linear (no more conjugate linear) positive, selfadjoint operator and

$$(\mathcal{S}^*\mathcal{S})^z Sv = \sigma_z(S). \tag{74.326}$$ eqMsstarMssigmaz

Thus, in fact, the whole set $\{\sigma_z\}_{z\in\mathbb{C}}$ can be recovered from $\mathcal{S}$ and thus from φ.

Notice that equation (74.326) makes no sense in the infinite dimensional case. More analytic work is needed to be right.

□

Two one parameter families of automorphisms $\alpha_t, \beta_t\colon M \to M$ ($\in \mathbb{R}$) are **outer equivalent** if there is a one parameter family of automorphisms $\gamma_t\colon \mathbb{M}_2(M) \to \mathbb{M}_2(M)$ such that

$$\gamma_t \begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix} = \begin{pmatrix} \alpha_t(S) & 0 \\ 0 & \beta_t(T) \end{pmatrix}. \tag{74.327}$$

This is an equivalence relation.

Since $\begin{pmatrix} 0 & \mathbb{1} \\ 0 & 0 \end{pmatrix}$ is a partial isometry, the automorphism γ_t has to send it to another partial isometry, i.e.

$$\gamma_t \begin{pmatrix} 0 & \mathbb{1} \\ 0 & \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{U}_t \\ 0 & 0 \end{pmatrix} \tag{74.328}$$

where $\mathcal{U}_t$ is unitary. I we apply γ_t to the equality

$$\begin{pmatrix} 0 & 0 \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix}, \tag{74.329}$$

we find

$$\gamma_t \begin{pmatrix} 0 & 0 \\ \mathbb{1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \mathcal{U}_t^* & 0 \end{pmatrix}, \tag{74.330}$$

so that

$$\mathcal{U}_t^* \alpha_t(S) \mathcal{U}_t = \beta_t(S). \tag{74.331}$$

The conclusion is that α_t and β_t only differ by an inner automorphism.

When $M = \mathfrak{B}(\mathscr{H})$, all automorphisms are inner and all the one parameter families of automorphisms are equivalent to the identity.

If α and β are outer equivalent, they define the same homomorphism

$$\mathbb{R} \to \mathrm{Out}(M) = \frac{\mathrm{Aut}(M)}{\mathrm{Inn}(M)}. \tag{74.332}$$

We recall that if M is a von Neumann algebra of operators of the Hilbert space $\mathscr{H}$, a vector ξ is **cyclic** if $\mathscr{H} = \overline{M\xi}$ and it is **separating** for M is $T\xi = 0$ if and only if $T = 0$.

Theorem 74.180 (Connes' theorem)**.**
The modular group of different cyclic and separating vectors are all outer equivalent.

Proof. Given two states φ_1 and φ_2, we define

$$\begin{aligned} &\varphi\colon \mathbb{M}_2(M) \to \mathbb{C} \\ &\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \mapsto \frac{1}{2}\big(\varphi_1(T_{11}) + \varphi_2(T_{22})\big), \end{aligned} \tag{74.333}$$

which is a faithful normal state. Then apply Tomita's theorem. □

74.16.6 Modular group

This subsection is based on [1003].

If ξ is cyclic and separating and if $T \mapsto T^*$ is an isometric involution on M, the antilinear map $T\xi \mapsto T^*\xi$ is densely defined on $M\xi \subset \mathscr{H}$. We denote by

$$S_\xi\colon \mathscr{H} \to \mathscr{H} \tag{74.334}$$

the closure. Thus everywhere it makes sense we have $S_\xi(T\xi) = T^*\xi$. The operator S_ξ has the polar decomposition

$$S_\xi = J_\xi \Delta_\xi^{1/2}. \tag{74.335}$$

The antilinear part J_ξ is said the **modular conjugation** while the operator $\Delta_\xi^{1/2}$ is the **modular operator** associated with the pair (M, ξ). We have $\Delta_\xi = S_\xi^* S_\xi$.

Now, instead of referring to the cyclic separating vector, we can consider a faithful normal state ω and think about its associated cyclic separating vector in the GNS cyclic representation $(\pi_\omega, \mathscr{H}_\omega, \xi_\omega)$. In this setting, the **modular group of automorphism** of the pair (M, ω) is σ^ω defined by

$$\sigma_t^\omega(T) = \pi_\omega^{-1}\big(\Delta_\omega^{it} \pi_\omega(T) \Delta_\omega^{-it}\big) \tag{74.336}$$

for $T \in M$ and $t \in \mathbb{R}$. Here Δ_ω is the modular operator associated with the pair $\big(\pi_\omega(M), \xi_\omega\big)$.

Theorem 74.181.
*The modular group $\sigma^\omega\colon \mathbb{R} \to \mathrm{Aut}(M)$ associated with the pair (M, ω) satisfies the following conditions (the **modular condition**):*

(1) ω is σ^ω-invariant, that means that $\omega = \omega \circ \sigma_t^\omega$.

(2) for every $x, y \in M$, there exists a bounded function F_{xy} on the strip

$$\bar{D} = \{z \in \mathbb{C} \text{ tel que } 0 \leqslant \mathrm{Im}(z) \leqslant 1\} \tag{74.337}$$

which is holomorphic on the interior of $\bar{D}$ and such that

$$F_{xy}(t) = \omega\big(\sigma_t^\omega(x) y\big) \tag{74.338a}$$

$$F_{xy}(t + i) = \omega\big(y \sigma_t^\omega(x)\big). \tag{74.338b}$$

Moreover, the map σ^ω is defined in an unique way by these two properties.

Let $\mathscr{H}$ be an Hilbert space and an acting von Neumann algebra M of compact operators on $\mathscr{H}$. Let ξ be a cyclic separating vector. We define

$$\begin{aligned} j_\xi \colon M &\to M' \\ x &\mapsto J_\xi x J_\xi \end{aligned} \tag{74.339}$$

Tomita's theorem shows that this map is well defined and that it takes values in M'.

If ξ is a trace vector (definition 60.8), the map j_ξ is the involution:

$$j_\xi(x) = x^* \tag{74.340}$$

74.16.7 Standard positive cone

The **standard positive cone** $\mathscr{H}_\xi^+$ associated wit the pair (M, ξ) is the set

$$\overline{\{x j_\xi(x)\xi \in \mathscr{H} \text{ tel que } x \in M\}}. \tag{74.341}$$

A **standard form** of a von Neumann algebra M acting on $\mathscr{H}$ is a self dual cone $\mathscr{H}^+$ and an antilinear involution J such that $(M, \mathscr{H}, \mathscr{H}^+, J)$ satisfies

(1) $JMJ = M'$;

(2) $JxJ = x^*$ for every $x \in M \cap M'$;

(3) $J\eta = \eta$ for every $\eta \in \mathscr{H}^+$;

(4) $(xJxJ)\mathscr{H}^+ \subset \mathscr{H}^+$ for every $x \in M$.

LaFin

Chapter 75

Dirichlet forms

75.1 KMS states and CAR algebras

75.1.1 KMS states

We refer to [1003] for missing proofs and more details. The document [1004] seems interesting too.

Definition 75.1.
*Let $\{\alpha_t\}_{t\in\mathbb{R}}$ be a strongly continuous one parameter semigroup of automorphisms of a C^*-algebra A and $\beta\in\mathbb{R}$. A state ω is a $(\alpha,\beta)-KMS$-state if is fulfils the **KMS condition**:*

$$\omega\big(a\alpha_{i\beta}(b)\big)=\omega(ba) \tag{75.1}$$

EqKMScondPourOmega

for every a and b in a norm dense and α-invariant $*$-subalgebra.

Note that for $\beta=0$, the state ω is a trace.

A map $\Phi\colon A\to A$ is **positive** when $\Phi(A_+)\subset A_+$ where A_+ is the set of positive elements in A (definition 72.62). It is **completely positive** if for every $n\in\mathbb{N}_0$, the extension

$$\begin{aligned}\Phi\otimes I_n\colon \mathbb{M}_(A)&\to\mathbb{M}_n(A)\\ (\Phi\otimes I_n)[a_{ij}]&=\big[\Phi(a_{ij})\big]\end{aligned} \tag{75.2}$$

is positive on the C^*-algebra $\mathbb{M}_n(A)=A\otimes\mathbb{M}_n(\mathbb{C})$, see definition 72.93. The map Φ is **Markovian** is it is positive and if $\Phi(a)\leqslant 1_A$ for every $a\in A$ such that $a=a*$ and $a\leqslant 1_A$. The map is **completely Markovian** if the map $\Phi\otimes I_n$ is Markovian on $\mathbb{M}_n(A)$.

A one parameter semigroup $\{\Phi_t\}_{t\geqslant 0}$ of maps $\Phi_t\colon A\to A$ has one of these properties when each of Φ_t has.

75.1.2 CAR algebras

Let V be a vector space with a bilinear non degenerated symmetric form. The $*$-algebra generated by the elements of V and subject to the relations

$$\begin{cases} vw+wv=\langle v,w\rangle & \text{(75.3a)}\\ v^*=v & \text{(75.3b)}\end{cases}$$

is the **CAR algebra** over V. Formally it is the quotient of the tensor algebra over V by the relations $v\otimes w+w\otimes v=\langle v,w\rangle$.

Let $\mathscr{H}_0$ be a pre-Hilbert space and $\mathscr{H}$ be its completion. We have a CAR algebra $U(\mathscr{H})$ given by the antilinear map $a\colon\mathscr{H}_0\to\mathfrak{B}(\mathscr{H})$ and the relations

$$\begin{cases} a(f)a(g)+a(g)a(f)=0 & \text{(75.4a)}\\ a(f)a(g)^*+a(g)^*a(f)=\langle f,g\rangle & \text{(75.4b)}\end{cases}$$

subEqafagssproddCAR

for every f and g in $\mathscr{H}_0$[1005].

If H is a selfadjoint on $\mathscr{H}$, we can define the automorphism group $\alpha\colon\mathbb{R}\to\mathrm{Aut}\big(U(\mathscr{H})\big)$ by

(1) $\alpha_t\big(a(f)\big) = a\big(e^{itH}f\big)$,
(2) $\alpha_t\big(a^*(f)\big) = a^*\big(e^{itH}f\big)$

for every $f \in \mathscr{H}_0$. This is a strongly continuous group of automorphisms of the CAR algebra called the **Bogoliubov transform**. If $\beta \in \mathbb{R}$, the condition for a state ω to be (α,β)-KMS requires among others

$$\omega\big(a(g)a^*(e^{-\beta H}f)\big) = \omega\Big(a(g)\alpha_{i\beta}\big(a^*(f)\big)\Big) = \omega\big(a^*(f)a(g)\big). \tag{75.5}$$

The CAR relation (75.4b) provides

$$a(g)a^*\big(e^{-\beta H}f\big) = \langle g, e^{-\beta H}f\rangle \mathrm{Id} - a^*(e^{-\beta H}f)a(g). \tag{75.6}$$

Writing the relation (75.5) with $(\mathrm{Id}+e^{-\beta H})^{-1}f$ instead of f, we get

$$\omega\Big(a(g)a^*\big(e^{-\beta H}(\mathrm{Id}+e^{-\beta H})^{-1}f\big)\Big) = \omega\Big(a^*\big((\mathrm{Id}+e^{-\beta H})^{-1}fa(g)\big)\Big), \tag{75.7}$$

but using the (anti)linearity of a^* and ω we have

$$\begin{aligned}\omega\Big(a^*\big((\mathrm{Id}+e^{-\beta H})f\big)a(g)\Big) &= \omega\Big(\big(a^*(f)+a^*(e^{-\beta H}f)\big)a(g)\Big)\\ &= \omega\big(a^*(f)a(g)\big) + \omega\big(a^*(e^{\beta H}f)a(g)\big)\\ &= \omega\big(a^*(f)a(g)\big) + \langle g, e^{-\beta H}f\rangle - \omega\big(a(g)a^*(e^{-\beta H}f)\big)\end{aligned} \tag{75.8}$$

where we used the fact that $a^*(e^{-\beta H}f)a(g) = \langle g, e^{-\beta H}f\rangle \mathrm{Id} - a(g)a^*(e^{-\beta H}f)$ and the fact that $\omega(\mathrm{Id}) = 1$ since ω is a state. In the last line, the first and last terms sum to zero because of the KMS condition. What we obtained is

$$\omega\Big(a^*\big([\mathrm{Id}+e^{-\beta H}]f\big)a(g)\Big) = \langle g, e^{-\beta H}f\rangle. \tag{75.9}$$

If we write this with $[\mathrm{Id}+e^{-\beta H}]^{-1}f$ instead of f, we get

$$\omega\big(a^*(f)a(g)\big) = \langle g, e^{-\beta H}[\mathrm{Id}+e^{-\beta H}]^{-1}f\rangle. \tag{75.10}$$

It turns out that, using the CAR relations, that formula uniquely defines ω. This is then the unique (α,β)-KMS state on $U(\mathscr{H})$ and is denoted by ω_β.

75.1.3 KMS symmetric maps

Definition 75.2.
Let $\{\alpha_t\}_{t\in\mathbb{R}}$ be a strongly continuous group of automorphisms of the C-algebra A and ω a (α,β)-KMS-state for some $\beta \in \mathbb{R}$. A bounded map $\Phi\colon A \to A$ is said to be (α,β)-**KMS symmetric** with respect to ω if*

$$\omega\big(b\Phi(a)\big) = \omega\Big(\alpha_{-\frac{i\beta}{2}}(a)\Phi\big(\alpha_{\frac{i\beta}{2}}(b)\big)\Big) \tag{75.11}$$

for every a and b in a norm dense and α-invariant $*$-algebra B of A_α.

We already mentioned that, when $\beta = 0$, the state ω is a trace; what happens with the KMS symmetry condition on Φ in the case $\beta = 0$ is

$$\omega\big(b\Phi(a)\big) = \omega\big(a\Phi(b)\big). \tag{75.12}$$

Let us suppose that, for a generic β, the map Φ commutes with the action. The symmetry condition becomes

$$\begin{aligned}\omega(b\Phi(a)) &= \omega\Big(\alpha_{-\frac{i\beta}{2}}(a)\alpha_{\frac{i\beta}{2}}\big(\Phi(b)\big)\Big)\\ &= \omega\big(\Phi(b)\alpha_{-\frac{i\beta}{2}}(a)\big)\\ &= \omega\big(a\Phi(b)\big)\end{aligned} \tag{75.13}$$

where we used twice the KMS condition for ω.

Lemma 75.3.
If ω is a (α,β)-KMS state, the map Φ is KMS-symmetric with respect to ω if and only if

$$\omega\big(\Phi(a)\alpha_{\frac{i\beta}{2}}(b)\big) = \omega\big(\alpha_{\frac{-i\beta}{2}}(a)\Phi(b)\big). \tag{75.14}$$

Proof. We write the usual KMS condition (75.1) for ω with the replacements $b \to \alpha_{\frac{-i\beta}{2}}(b)$ and $a \to \Phi(a)$:

$$\omega\big(\Phi(a)\alpha_{\frac{i\beta}{2}}(b)\big) = \omega\big(\alpha_{-\frac{i\beta}{2}}(b)\Phi(a)\big), \tag{75.15}$$

then we use the KMS condition (75.11) for Φ with the replacements $a \to b$ and $b \to \alpha_{\frac{-i\beta}{2}}(a)$ in order to get $\omega\big(\alpha_{-\frac{i\beta}{2}}(a)\Phi(b)\big)$. □

We recall that an **entire function** is a function holomorphic over $\mathbb{C}$. We define the set D_β as

$$\begin{aligned} D_\beta &= \{z \in \mathbb{C} \text{ tel que } 0 < \operatorname{Im} z < \beta\} \quad \text{if } \beta \geqslant 0 \\ D_\beta &= \{z \in \mathbb{C} \text{ tel que } \beta < \operatorname{Im} z < 0\} \quad \text{if } \beta \leqslant 0 \end{aligned} \tag{75.16}$$

Proposition 75.4.
Let $\alpha = \{\alpha_t\}_{t\in\mathbb{R}}$ be a strongly continuous group of automorphisms of the C^*-algebra A, $\beta \in \mathbb{R}$ and ω be a state on A. The following two conditions are equivalent:

(1) ω is a (α,β)-KMS state;

(2) for every pair a, b in A, there exists an analytic map $F_{a,b}\colon \overline{D_\beta} \to A$ such that EqDEfFabi

$$F_{a,b}(t) = \omega(a\alpha_t(b)) \tag{75.17a}$$

$$F_{a,b}(t+i\beta) = \omega(\alpha_t(b)a) \tag{75.17b}$$

for $t \in \mathbb{R}$.

If these conditions are satisfied, we have $|F_{a,b}(z)| \leqslant \|a\| \cdot \|b\|$ for every $z \in \overline{D}_\beta$.

Moreover, if $a \in A$ and $b \in A_\alpha$, the function $F_{a,b}$ is the restriction to $\bar{D}_\beta$ of the entire function

$$z \mapsto \omega\big(a\alpha_z(b)\big). \tag{75.18}$$

The point in the proposition is that the relations (75.17) define only $F_{a,b}$ on the boundary of D_β.

We have the same kind of proposition for a KMS symmetry.

PropFabcomegaKSM

Proposition 75.5.
Let $\alpha = \{\alpha_t\}_{t\in\mathbb{R}}$ be a strongly continuous automorphism group of the C^*-algebra A and ω, a (α,β)-KMS state for some $\beta \in \mathbb{R}$. Then for a bounded map $\Phi\colon A \to A$, the following conditions are equivalent:

(1) Φ is (α,β)-KMS symmetric with respect to ω;

(2) pour every pair a, b in A, there exists a continuous bounded analytic map $F_{a,b}\colon \bar{D}_\beta \to A$ such that

$$F_{a,b}(t) = \omega\Big(\alpha_{-\frac{t}{2}}(a)\Phi\big(\alpha_{\frac{t}{2}}(b)\big)\Big) \tag{75.19a}$$

$$F_{a,b}(t+i\beta) = \omega\Big(\alpha_{\frac{t}{2}}(b)\Phi\big(\alpha_{\frac{t}{2}}(a)\big)\Big) \tag{75.19b}$$

subEqFabitbhi

for every $t \in \mathbb{R}$.

If these conditions are satisfied, the function $F_{a,b}$ is bounded on $\bar{D}_\beta$ and

$$|F_{a,b}(z)| \leqslant \|\Phi\| \cdot \|a\| \cdot \|b\| \tag{75.20}$$

for every $z \in \bar{D}_\beta$. If $a \in A$ and $b \in A_\alpha$, the function $F_{a,b}$ is the restriction to $\bar{D}_\beta$ of the entire function

$$G_{a,b}(z) = \omega\Big(\alpha_{-\frac{z}{2}}(a)\Phi\big(\alpha_{\frac{z}{2}}(b)\big)\Big) \tag{75.21}$$

on $\mathbb{C}$.

Let us briefly recall the GNS construction of a C^*-algebra A (see theorem 72.87). Let ω be a state on A. First we consider the space

$$\mathcal{N}_\omega = \{a \in A \text{ tel que } \omega(a^*a) = 0\} = \{a \in A \text{ tel que } \omega(b^*a) = 0\, \forall b \in A\}, \tag{75.22}$$

and we define the Hilbert space $\mathscr{H}_\omega$ as the completion of the quotient $A/\mathcal{N}_\omega$. The representation π_ω on $\mathscr{H}_\omega$ is then defined as

$$\pi_\omega(a)[b] = [ab] \tag{75.23}$$

where the bracket $[a]$ denotes the class of $a \in A$ for the quotient by $\mathcal{N}_\omega$.

The kernel of the representation is given by the elements a such that

$$[ab] = 0 \tag{75.24}$$

for every $b \in A$. This means that $a \in \mathcal{N}_\omega$. Thus the kernel of the GNS representation is given by the set of elements $a \in A$ such that $\omega(a^*a) = 0$.

Suppose that we have an semigroup α_t of automorphisms of A and a α-invariant state ω. Then if $a \in \ker(\pi_\omega)$ we have

$$\begin{aligned} \omega\big(\alpha_t(a^*)\alpha_t(a)\big) &= \omega\big(\alpha_t(a^*a)\big) \\ &= \omega(a^*a) \\ &= 0, \end{aligned} \tag{75.25}$$

so that $\alpha_t(a)$ belongs to $\ker(\pi_\omega)$ too.

Proposition 75.6.
Let $\{\alpha_t\}_{t\in\mathbb{R}}$ be a strongly continuous group of automorphisms of the C^*-algebra A and ω be a fixed (α,β)-KMS state for some $\beta \in \mathbb{R}$. Then a map $\Phi\colon A \to A$ which is (α,β)-KMS symmetric with respect to ω leaves globally invariant the kernel $\ker(\pi_\omega)$ of the GNS representation of ω.

Proof. Let $a \in \ker(\pi_\omega)$. Since α_t is an automorphism of A, we also have $\alpha_t(a) \in \ker(\pi_\omega)$. Let $F_{a,b}$ is the map guaranteed by proposition 75.5; since ω is positive (it is a state, definition 72.67), it fulfils the condition (72.62), so that

$$|F_{a,b}(t)|^2 = |\omega\Big(\alpha_{-\frac{t}{2}}(a)\Phi\big(\alpha_{\frac{t}{2}}(b)\big)\Big)|^2 \leqslant \omega\Big(\alpha_{-\frac{t}{2}}(a)\big(\alpha_{-\frac{t}{2}}(a)\big)^*\Big)\omega\Big(\big(\Phi\alpha_{\frac{t}{2}}(b)\big)^*\Phi\big(\alpha_{\frac{t}{2}}\big)(b)\Big). \tag{75.26}$$

Since α_t is an automorphism and since ω is α_t-invariant, the first factor becomes

$$\omega(\alpha_{-\frac{t}{2}}(aa^*)) = \omega(aa^*) = 0. \tag{75.27}$$

What we proved is that $|F_{a,b}(t)| = 0$ when t is real. Being analytic, the function $F_{a,b}$ vanishes everywhere. By property (75.19b),

$$0 = F_{a,b}(i\beta) = \omega\big(b\Phi(a)\big). \tag{75.28}$$

Taking $b = \Phi(a)^*$, we get $\Phi(a) \in \ker(\pi_\omega)$. □

75.2 Dirichlet and traces

Other source: [1006, 1007].

Let A be a C^*-algebra and $\tau\colon A^+ \to [0,\infty]$ be a faithful, densely defined, semifinite and lower semicontinuous trace on A^+ (see section 72.18). We consider $\mathcal{L}$, the set of **square integrable** elements in A, that is:

$$\mathcal{L} = \{a \in A \text{ tel que } \tau(aa^*) < \infty\}. \tag{75.29}$$

This set is an ideal in A and τ can be uniquely extended to a linear functional on $\mathcal{N} = \langle \mathcal{L}^*\mathcal{L}\rangle$ (lemma 72.129). We still denote by τ the extension. In particular the formula

$$(a,b)_\tau = \tau(a^*b) \tag{75.30}$$

defines a sesquilinear form on $\mathcal{L}$. We denote by $L^2(A,\tau)$ the Hilbert space obtained by completion of $\mathcal{L}$ with respect to the sesquilinear form $(.,.)_\tau$. Such a construction is similar to the one of $L^2(M)$ on page 3877. Since the trace is faithful, we have $\omega(a^*a)=0$ only when $a=0$. The space $L^2(A,\tau)$ is thus also the GNS construction described in theorem 72.87.

We denote by $\eta_\tau\colon \mathcal{L}\to L^2(A,\tau)$ the natural injection whose image is, by construction, dense. The space $\eta_\tau(\mathcal{L})$ has the product

$$\langle \eta_\tau(a),\eta_\tau(b)\rangle = (a,b)_\tau = \tau(a^*b). \tag{75.31}$$

If, for $a\in A$ and $b\in\mathcal{L}$, we define

$$\pi_\tau\eta_\tau(a) = \eta_\tau(ab), \tag{75.32}$$

we have

$$\langle \eta_\tau(c),\pi_\tau(a)\eta_\tau(b)\rangle = \tau(c^*ab). \tag{75.33}$$

The map $\pi_\tau\colon A\to\mathfrak{B}\big(L^2(A,\tau)\big)$ is a representation of A on $L^2(A,\tau)$. We consider the von Neumann algebra $L^\infty(A,\tau)\mathfrak{M}=\pi_\tau(A)''$. Since τ is defined on $\pi_\tau(A)$ and that $\mathfrak{M}$ is a strong closure of $\pi_\tau(A)$ (theorem 74.5), the form τ can be extended to a normal functional on $\mathfrak{M}$. From now we write a the element $\eta_\tau(a)$ in $L^2(A,\tau)$.

Let

$$C=\overline{\{a\in\mathcal{L}\text{ tel que }a=a^*\leqslant 1_{\mathfrak{M}}\}}\subset L^2(A,\tau). \tag{75.34}$$

This set is convex. Indeed, let $a,b\in C$ and consider $(1-t)a+tb$. We have

$$(1-t)a+tb\leqslant(1-t)1_{\mathfrak{M}}+t1_{\mathfrak{M}}=1_{\mathfrak{M}}. \tag{75.35}$$

We write $a\wedge 1$ the projection of a onto C.

The notation is inspired from the fact that if we consider the real function $f(t)=t\wedge 1=\min(t,1)$, we have $a\wedge 1=f(a)$ in the case $a=a^*\in\mathcal{L}$ by the continuous functional calculus, theorem 72.23. If $b\in C$, we have $\mathrm{Spec}(b)\subset\,]-\infty,1]$, so that $f(t)=t$ and $f(b)=b$. If a is outside C, we have

$$\mathrm{Spec}(a\wedge 1)=f\big(\mathrm{Spec}(a)\big)\subset\,]-\infty,1]. \tag{75.36}$$

Then we have $\mathrm{Spec}(1-a\wedge 1)\subset[0,\infty]$, which proves that $1-a\wedge 1\geqslant 0$.

A quadratic functional

$$\mathcal{E}\colon L^2(A,\tau)\to\,]-\infty,\infty] \tag{75.37}$$

can be extended to $\mathbb{M}_n(A)=A\otimes\mathbb{M}_n(\mathbb{C})$ in the following way. First we consider the trace $\tau_n=\tau\otimes\mathrm{Tr}_n$ on $\mathbb{M}_n(A)$ then we define $\mathcal{E}_n\colon L^2\big(\mathbb{M}_n(A),\tau_n\big)\to\mathbb{C}$ by

$$\mathcal{E}_n[a]=\sum_{ij}\mathcal{E}[a_{ij}]. \tag{75.38}$$

The domain of $\mathcal{E}$ is the set

$$\mathcal{F}=\{\xi\in L^2(A,\tau)\text{ tel que }\mathcal{E}[\xi]<\infty\}. \tag{75.39}$$

The functional $\mathcal{E}$ is said to be

(1) **J-real** if $\mathcal{E}[J\xi]=\mathcal{E}[\xi]$ for every $\xi\in L^2(A,\tau)$;

(2) **Markovian** if $\mathcal{E}[\xi\wedge 1]\leqslant\mathcal{E}[\xi]$ for every $\xi\in L^2(A,\tau)$ such that $\xi=J\xi$;

(3) **Dirichlet form** if it is Markovian and lower semicontinuous;

(4) **completely Dirichlet** if the extensions $\mathcal{E}_n$ are Dirichlet for all $n\geqslant 1$;

(5) **regular** if the subspace $\mathcal{B}=A\cap\mathcal{F}$ is dense-norm in A and a form core for $(\mathcal{E},\mathcal{F})$ (definition 60.33);

(6) **C^*-Dirichlet form** if it is regular and completely Dirichlet.

LemELfermEinter

Lemma 75.7.
Let $E = \{a \in \mathcal{L} \text{ tel que } a = a^* \leqslant 1_{\mathfrak{M}}\}$ and $C = \overline{E}$. We recall that $L^2_+(A,\tau) = \overline{\mathcal{L}_+}$. Then

$$C \cap L^2_+(A,\tau) = \overline{\{a \in \mathcal{L} \text{ tel que } 0 \leqslant a \leqslant 1_{\mathfrak{M}}\}}. \tag{75.40}$$

Proof. First, it is obvious that $\{a \in \mathcal{L} \text{ tel que } 0 \leqslant a \leqslant 1_{\mathfrak{M}}\} = E \cap \mathcal{L}_+$. Thus we have to show that $\overline{E} \cap \overline{\mathcal{L}_+} = \overline{E \cap \mathcal{L}_+}$. The inclusion

$$\overline{E \cap \mathcal{L}_+} \subset \overline{E} \cap \overline{\mathcal{L}_+} \tag{75.41}$$

is only topology.

The delicate part is the reverse inclusion. Let $x \in \overline{E} \cap \overline{\mathcal{L}_+}$. In a first time we suppose that $\|x\| = 1$. By hypothesis we can find a sequence $x_i \to x$ with $x_i \in \mathcal{L}_+$. Since $\|x_i\| \to 1$, we consider the sequence

$$y_i = \frac{x_i}{\|x_i\|} \to x. \tag{75.42}$$

The element y_i still belongs to $\mathcal{L}_+$ from proposition 72.36. Moreover $y_i \leqslant \mathbb{1}$ because proposition 72.38 shows that $x_i \leqslant \|x_i\|\mathbb{1}$. So we have $y_i \in E \cap \mathcal{L}_+$ and $x \in \overline{E \cap \mathcal{L}_+}$.

Let us now consider any $x \in \overline{E} \cap \overline{\mathcal{L}_+}$. By the preceding point we build a sequence $y_i \in E \cap \mathcal{L}_+$ such that

$$\begin{cases} y_i \to \dfrac{x}{\|x\|} & \text{(75.43a)} \\ \|y_i\| = 1. & \text{(75.43b)} \end{cases}$$

Then we consider the sequence $z_i = \|x\|y_i$. This converges to x and $z_i \in \mathcal{L}_+$. We still have to show that $z_i \in E$.

First we prove that $\|x\| \leqslant \mathbb{1}$. For that we consider $a_i \to x$ with $a_i \in E$. Since $x \in \overline{\mathcal{L}_+}$ for every ϵ we have

$$\mathrm{Spec}(a_i) \subset [-\epsilon, \|a_i\|]. \tag{75.44}$$

We conclude that $a_i \geqslant 1$. Indeed the spectrum of $a_i + 1$ is made of λ such that $a_i - (\lambda - 1)\mathbb{1}$ is not invertible. Thus $\lambda \in \mathrm{Spec}(a_i + 1)$ if and only if $\lambda - 1 \in \mathrm{Spec}(a_i)$ and $\lambda - 1 \geqslant \epsilon$, so that $\lambda \geqslant 1 - \epsilon > 0$. The spectrum of $a_i + \mathbb{1}$ being positive we have $a_i \geqslant -\mathbb{1}$. Now we have $-\mathbb{1} \leqslant a_i \leqslant \mathbb{1}$ and proposition 72.39 concludes that $\|a_i\| \leqslant 1$. Since $a_i \to x$ we conclude that $\|x\| \leqslant 1$.

¡¡ Avertissement/question à la lectrice !! 75.8
Here I'm using the fact that the norm is continuous. This is due to the fact that we are precisely working on the completion with respect to the norm.

By the definition $z_i = \|x\|y_i$ and $\|y_i\| = 1$, we have $\|z_i\| \leqslant 1$. Once again by proposition 72.38 we have $z_i \leqslant \|z_i\|\mathbb{1} \leqslant \mathbb{1}$.

□

Definition 75.9.
Let $\{T_t\}_{t\geqslant 0}$ be a J-real, symmetric, strongly continuous semigroup acting on $L^2(A,\tau)$. That semigroup is said to be

(1) ***Markovian*** *if it leaves globally invariant the set (see lemma 75.7)*

$$C \cap L^2(A,\tau) = \overline{\{a \in \mathcal{L} \text{ tel que } 0 \leqslant a \leqslant 1_{\mathfrak{M}}\}}; \tag{75.45}$$

(2) ***completely positive*** *if its extensions to $L^2\big(\mathbb{M}_n(A),\tau_n\big)$ are positive;*

(3) ***completely Markovian*** *if its extensions to $L^2\big(\mathbb{M}_n(A),\tau_n\big)$ are Markovian.*

75.3 Modules over C^*-algebra

Definition 75.10.
*Let A be a C^*-algebra. A A-**bimodule** is an Hilbert space $\mathscr{H}$ with commuting left and right representations. The bimodule is said to be **symmetric** if there is an isometric antilinear involution $J\colon \mathscr{H} \to \mathscr{H}$ exchanging the left and right actions:*

$$J(a\xi b) = b^* J(\xi) a^* \tag{75.46}$$

for every $a, b \in A$ and $\xi \in \mathscr{H}$.

Let $a \in A$ and $\xi \in \mathscr{H}$. If π is a representation of A on $\mathscr{H}$ we have $\|\pi(a)\xi\| \leqslant \|\pi(a)\|\|\xi\|$, and lemma 72.82 shows that

$$\|\pi(a)\xi\| \leqslant \|a\|\|\xi\|. \tag{75.47}$$ Eqpiaxilequpormxi

In the case of a A-module we do not explicitly write the representation π and we write $\|a\xi b\| \leqslant \|a\|\|b\|\|\xi\|$.

When $\mathscr{H}$ is a bimodule over A, a **derivation** is a linear map $\partial\colon D(\partial) \subset A \to \mathscr{H}$ such that

$$\partial(ab) = \partial(a)b + a\partial(b) \tag{75.48}$$

for every $a, b \in D(\partial)$. Notice that one always has $\partial(1) = 0$. A derivation is **symmetric** if its domain $D(\partial)$ is a selfadjoint subalgebra and if the A-bimodule $(\mathscr{H}, J)$ is symmetric with $\partial(a^*) = J(\partial a)$.

75.3.1 Example: gradient of a function

Let M be a Riemannian manifold and consider $A = C_b(M)$, the C^*-algebra of continuous bounded function on M; $\mathscr{H} = L^2(TM)$ the space of square summable vector fields on M. The gradient

$$\nabla\colon C_b^\infty(M) \to L^2(TM) \tag{75.49}$$

is a derivation and $L^2(TM)$ is a $C_b(M)$-bimodule with the pointwise multiplication.

75.3.2 Left and right representation

If $a \in A$ is a selfadjoint element, we can represent the C^*-algebra $C\big(\operatorname{Spec}(a)\big)$ on $\mathscr{H}$ by

$$L_a(f)\xi = \begin{cases} f(a)\xi & \text{if } f(0) = 0 \\ \xi & \text{if } f \equiv 1 \end{cases} \tag{75.50}$$

where $f \in C\big(\operatorname{Spec}(a)\big)$ and $\xi \in \mathscr{H}$. This is a good definition when f is a polynomial. The action with non polynomial continuous functions is defined by density. The right action is defined by

$$R_a(f)\xi = \begin{cases} \xi f(a) & \text{if } f(0) = 0 \\ \xi & \text{if } f \equiv 1. \end{cases} \tag{75.51}$$

Let I be an interval in $\mathbb{R}$ and $f \in C^1(I)$. The **quantum derivative** of f is the function $\tilde{f} \in C(I \times I)$ defined by

$$\tilde{f}(s,t) = \begin{cases} \frac{f(s)-f(t)}{s-t} & \text{if } s \neq t \\ f'(s) & \text{if } s = t. \end{cases} \tag{75.52}$$

Lemma 75.11.
Let $\mathscr{H}, J$ a symmetric Hilbert A-bimodule and $(\partial, D(\partial))$ be a symmetric derivation defined on an involutive subalgebra of A. Let $a = a^* \in D(\partial)$. Then

(1) *For every polynomials f, we have $f(a) \in D(\partial)$ and the* chain rule

$$\partial\big(f(a)\big) = (L_a \otimes R_a)(\tilde{f})\partial(a). \tag{75.53}$$

EqLeibnizDerrAbiomdule

holds. Moreover we have the bound

$$\|\partial\big(f(a)\big)\| \leqslant \|f'\|_{C(\mathrm{Spec}(a))}\|\partial(a)\|. \tag{75.54}$$

EqBoundpartialfanfnnpa

(2) *If $(\partial, D(\partial))$ is closable as operator $A \to \mathscr{H}$, then the closure is a derivation.*

(3) *If $(\partial, D(\partial))$ is a closed derivative $A \to \mathscr{H}$, then the formula (75.53) holds for every function $f \in C^1\big(\mathrm{Spec}(a)\big)$ such that $f(0) = 0$. In particular the domain of a closed derivative is closed for the C^1 functional calculus.*

Remarque 75.12.
The equation (75.53) is a short notation for the following. First we introduce the following representation of $A \otimes A^{op}$ on $\mathscr{H}$:

$$(a \otimes b)\sharp\xi = a\xi b. \tag{75.55}$$

Then $(L_a \otimes R_a)(\tilde{f})\xi$ with $\tilde{f}(0) = 0$ stands for

$$(L_a \otimes R_a)(\tilde{f}_1 \otimes \tilde{f}_2)\sharp\xi = \Big(L_a(\tilde{f}_1) \otimes R_a(\tilde{f}_2)\Big)\sharp\xi = \tilde{f}_1(a)\xi\tilde{f}_2(a). \tag{75.56}$$

EqLaRattfsurxi

Here the functions $\tilde{f}_i$ are associated with $\tilde{f}$ by the map (72.140).

Proof. (1) If $a = a^* \in D(\partial)$, then $f(a) \in D(\partial)$ because the domain of ∂ is an algebra.

Let us begin with the constant function $f = 1$. In this case $\tilde{f}(s,t) = 0$ and all terms are vanishing in the decomposition (72.141), so $\tilde{f} = 0 \otimes 0$. Formula (75.56) then produces $\partial(1) = 0$.

If $f(t) = t$, then $\tilde{f}(s,t) = 1$ and we get $\partial(a) = \partial(1)$.

The general statement comes from the formula

$$\frac{x^n - y^n}{x - y} = \sum_{k=0}^{n-1} x^k y^{n-1-k}. \tag{75.57}$$

Then, for $f(t) = \sum_n \alpha_n t^n$, one defines

$$Df_a = \sum_n \alpha_n \sum_{k=0}^{n-1} a^k \otimes a^{n-1-k} \tag{75.58}$$

and the formula to be checked becomes $\partial\big(f(a)\big) = Df_a\sharp\partial(a)$ that can be checked by induction. Let us now prove the bound (75.54). If $f(t) = t^n$, then

$$Df_a\sharp\xi = \sum_{k=0}^{n-1} a^k\xi a^{n-1-k}. \tag{75.59}$$

Using the inequality (75.47) we found

$$\begin{aligned} \|Df_a\sharp\xi\| &\leqslant \sum_k \|a^k\xi a^{n-1-k}\| \\ &\leqslant \sum_k \|a^k\|\|a^{n-1-k}\|\|\xi\| \\ &\leqslant \sum_{k=0}^{n-1} \|a\|^{n-1}\|\xi\| \\ &= n\|a\|^{n-1}\|\xi\|. \end{aligned} \tag{75.60}$$

In this computation we also used the fact that $\|ab\| \leqslant \|a\|\|b\|$. Since $\|a\| \leqslant \sup_{t\in\mathrm{Spec}(a)} |t|$ by proposition 60.58(1), we have

$$n\|a\|^{n-1} \leqslant \sup_{t\in\mathrm{Spec}(a)} |f'(t)| = \|f'\|_{C(\mathrm{Spec}(a))}. \tag{75.61}$$

(2) Let $a_n \to a$ and $b_n \to b$ and consider the Leibniz formula

$$\partial(a_n b_n) = a_n(\partial b_n) + (\partial a_n)b_n \tag{75.62}$$ eqPartialanbnLeibniz

Since the operator ∂ is supposed to be closable (see proposition 60.10), we can define $\partial(a) = \lim \partial(a_n)$. By continuity of the left and right actions, we also have

$$\lim a_n \partial b_n = a\partial b. \tag{75.63}$$

In the same spirit, $\lim(a_n b_n) = ab$ and we have $\lim \partial(a_n b_n) = \partial(ab)$. Thus taking the limit in both side of equation (75.62), we found $\partial(ab) = a\partial b + (\partial a)b$.

(3) Now we suppose that $(\partial, D(\partial))$ is closed and we consider $f \in C^1\big(\operatorname{Spec}(a)\big)$. In particular f' is continuous and we can consider polynomials f_n such that $f_n' \to f'$ and $f_n \to f$ uniformly on $\operatorname{Spec}(a)$. Indeed on each connected component K of $\operatorname{Spec}(a)$ we have

$$\|f_n(t) - f(t)\| \leqslant \sup_{t\in K} |f_n'(t) - f'(t)|\ell = \ell\|f_n' - f'\| \leqslant 2\|a\|\|f_n' - f'\| \tag{75.64}$$

where ℓ is the length of K. We also used the fact that the length of $\operatorname{Spec}(a)$ is smaller or equal to $2\|a\|$ by Since $f_n' \to f_n$ uniformly, the right-hand side can be made smaller than ϵ by a suitable choice of n. We suppose that $f_n(0) = 0$ by choosing the suitable primitive of f_n'.

For each n we have

$$\|\partial\big(f_n(a)\big)\| \leqslant \|f_n'\|\|\partial(a)\|. \tag{75.65}$$

We prove now that $\partial\big(f_k(a)\big)$ is a Cauchy sequence. By linearity, $\partial\big(f_k(a)\big) - \partial\big(f_l(a)\big) = \partial\big((f_k - f_l)(a)\big)$. By the bound (75.54) we have

$$\|\partial\big((f_k - f_l)(a)\big)\| \leqslant \|f_k' - f_l'\|\|\partial(a)\|. \tag{75.66}$$ EqpartfkfllmeqCauchy

Since $f_k' \to f'$ uniformly, the right hand side of (75.66) can be set smaller than ϵ by choosing suitably large k and l. This shows that $f(a) \in D(\partial)$ and, by definition, $\partial\big(f(a)\big) = \lim \partial\big(f_n(a)\big)$.

As far as $\tilde{f}$ is concerned, we have the uniform limit $\tilde{f}_{n,i} \to \tilde{f}_i$ and then

$$\partial\big(f_n(a)\big) = \tilde{f}_{n,1}(a)\partial(a)\tilde{f}_{n,2}(a). \tag{75.67}$$

By definition the limit of the left hand side is $\partial\big(f(a)\big)$ while the limit of the right hand side is, by continuity of the actions, $\tilde{f}_1(a)\partial(a)\tilde{f}_2(a)$.

□

Theorem 75.13.
Let $\big(\partial, D(\partial)\big)$ be a symmetric derivation on A defined on a dense subset with values in a symmetric Hilbert A-bimodule $(\mathscr{H}, J)$. We suppose that $D(\partial)$ is dense in $L^2(A,\tau)$ and that $\big(\partial, D(\partial)\big)$ is a closable form from $L^2(A,\tau)$ to $\mathscr{H}$. The form $(\mathcal{E}, \mathcal{F})$ defined by

$$\begin{aligned} \mathcal{E}\colon L^2(A,\tau) &\to [0,\infty[\\ a &\mapsto \|\partial a\|^2_{\mathscr{H}} \end{aligned} \tag{75.68}$$

is a C-Dirichlet form with $\mathcal{F} = D(\partial)$.*

Moreover, if Δ is the generator of the completely Markovian semigroup associated to $\mathcal{E}$, we have $\Delta = \partial^* \circ \partial$.

Proposition 75.14.
Let $\big(\mathcal{E}, D(\mathcal{E})\big)$ be a completely Dirichlet form on $L^2(A,\tau)$. Then $\mathcal{B} = A \cap D(\mathcal{E})$ is an involutive subalgebra of $\mathfrak{M}$.

If $\big(\mathcal{E}, D(\mathcal{E})\big)$ is a C-Dirichlet form, the subalgebra $\mathcal{B}$ is an involutive subalgebra of A and a core form for $\big(\mathcal{E}, D(\mathcal{E})\big)$.*

The subalgebra $\mathcal{B} = A \cap D(\mathcal{E})$ is the **Dirichlet algebra** of the form $\big(\mathcal{E}, D(\mathcal{E})\big)$. We also consider the hermitian sesquilinear form obtained from $\mathcal{E}$ by polarization:

$$\begin{aligned} \mathcal{E}\colon D(\mathcal{E}) \times D(\mathcal{E}) &\to \mathbb{C} \\ (\xi, \eta) &\mapsto \frac{1}{4}\Big(\mathcal{E}[\xi+\eta] - \mathcal{E}[\xi-\eta] + i\mathcal{E}[\xi+i\eta] - i\mathcal{E}[\xi-i\eta]\Big). \end{aligned} \tag{75.69}$$

It allows to recover $\mathcal{E}$.

PropSesquidEmbmb

Proposition 75.15.
The sesquilinear form on the algebraic tensor product $\mathcal{B} \otimes \mathcal{B}$ given by

$$(c \otimes d, a \otimes b) \mapsto \frac{1}{2}\big(\mathcal{E}(c, abd^*) + \mathcal{E}(cdb^*, a) - \mathcal{E}(db^*, c^*a)\big) \tag{75.70}$$

is definite positive.

We denote by $\mathscr{H}_0$ the Hilbert space obtained after separation[936] and completion of $\mathcal{B}\otimes\mathcal{B}$ with respect to the sesquilinear form given by proposition 75.15. The inner product will be denoted by $(.|.)_{\mathscr{H}_0}$ and if $a \otimes b \in \mathcal{B} \otimes \mathcal{B}$, we denote by $a \otimes_{\mathcal{E}} b$ the canonical image in $\mathscr{H}_0$.

Theorem 75.16.
Si $\big(\mathcal{E}, D(\mathcal{E})\big)$ is a C^*-Dirichlet form on $L^2(A, \tau)$, there exists a A-bimodule structure on $\mathscr{H}_0$ characterised by

(1) $a(b \otimes_{\mathcal{E}} c) = ab \otimes_{\mathcal{E}} c - a \otimes_{\mathcal{E}} bc$

(2) $(b \otimes_{\mathcal{E}} c)a = b \otimes_{\mathcal{E}} ca$

for every a, b, c in the Dirichlet algebra $\mathcal{B}$.

Chapter 76

K-theory

Elements of K-theory are taken from [909, 819].

76.1 The group K_0

Let $\mathfrak{A}$ be an unital C^*-algebra and $M_N(\mathfrak{A})$, the C^*-algebra of $N \times N$ matrices with coefficients in $\mathfrak{A}$. We have

$$M_N(\mathfrak{A}) \simeq \mathfrak{A} \otimes_{\mathbb{C}} M_N(\mathbb{C}). \tag{76.1}$$

The isomorphism being given by

$$a \mapsto \sum_{ij} a_{ij} \otimes_{\mathbb{C}} E_{ij}$$

if $a_{ij} \in \mathfrak{A}$ are the coefficients of $a \in M_N(\mathfrak{A})$. Two projectors $p, q \in M_N(\mathfrak{A})$ are **equivalent in the sense of Murray-von Neumann** if there exists a matrix $u \in M_N(\mathfrak{A})$ such that $p = u^*u$ and $q = uu^*$. In that case, the matrix u is a partial isometry by lemma 60.13. Notice that this notion of equivalence was already developed on page 3837 when speaking of projectors in von Neumann algebras.

We consider

$$M_\infty(\mathfrak{A}) = \bigcup_{N=1}^{\infty} M_N(\mathfrak{A}),$$

with the inclusion map

$$\begin{aligned} \phi\colon M_N(\mathfrak{A}) &\to M_{N+1}(\mathfrak{A}) \\ a &\mapsto \phi(a) = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{76.2}$$

We consider the equivalence $p \sim q$ in $M_\infty(\mathfrak{A})$ if and only if there exists $u \in M_\infty(\mathfrak{A})$ such that $p = u^*u$ and $q = uu^*$. The class of q is denoted by $[q]$ and the set of classes is denoted by $V(\mathfrak{A})$. The set $V(\mathfrak{A})$ is a semigroup for the addition

$$[p] + [q] = \left[\begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \right].$$

Some elements are not invertible.

¡¡ Avertissement/question à la lectrice !! 76.1
Pour bien faire, il faut encore prouver que cette somme est bien définie.

Proposition 76.2.
The semigroup $V(\mathfrak{A})$ is abelian.

Proof. If $p = uu^*$ and $q = vv^*$ and $a = \begin{pmatrix} u & \\ & v \end{pmatrix}$, we have

$$[p] + [q] = \left[\begin{pmatrix} uu^* & \\ & vv^* \end{pmatrix} \right] = \left[\begin{pmatrix} u & \\ & v \end{pmatrix} \begin{pmatrix} u^* & \\ & v^* \end{pmatrix} \right] = [aa^*] = [a^*a],$$

but

$$a^*a = \begin{pmatrix} & v^* \\ u^* & \end{pmatrix} \begin{pmatrix} & u \\ v & \end{pmatrix} = \begin{pmatrix} v^*v & \\ & u^*u \end{pmatrix} = \begin{pmatrix} q & \\ & p \end{pmatrix}.$$

□

We pose, on $V(\mathfrak{A}) \times V(\mathfrak{A})$, the equivalence relation

$$\big([p],[q]\big) \sim \big([p'],[q']\big)$$

if and only if there exists $[r] \in V(\mathfrak{A})$ such that

$$[p] + [q'] + [r] = [p'] + [q] + [r]. \tag{76.3}$$

EqConfEquivpqrK

Now the group $K_0(\mathfrak{A})$ is defined as

$$K_0(\mathfrak{A}) = \frac{\big(V(\mathfrak{A}) \times V(\mathfrak{A})\big)}{\sim}.$$

That construction is not a particularity of K-theory. When S is a semigroup, the group $K(S) = S \times S/\sim$ where $(s,t) \sim (u,v)$ if and only if there exists a $r \in S$ such that $s + v + r = t + u + r$ is the **Grothendiek group** of the semigroup S. The group $K(S)$ is always abelian. For example, $K(\mathbb{N},+) = \mathbb{Z}$ and $K(\mathbb{Z},\ \cdot\ ) = \mathbb{Q}$.

An addition in $K_0(\mathfrak{A})$ is defined by

$$\Big[[p],[q]\Big] + \Big[[p'],[q']\Big] = \Big[[p]+[p'],[q]+[q']\Big].$$

From definition we have $\Big([p],[p]\Big) \sim \Big([q],[q]\Big)$ and $\Big[[p],[p]\Big]$ is the neutral for addition in $K_0(\mathfrak{A})$. The invert is given by

$$-\Big[[p],[q]\Big] = \Big[[q],[p]\Big].$$

We have a natural homomorphism

$$\begin{aligned} \kappa_{\mathfrak{A}} \colon V(\mathfrak{A}) &\to K_0(\mathfrak{A}) \\ [p] &\mapsto \Big[[p],[0]\Big]. \end{aligned} \tag{76.4}$$

We say that $V(\mathfrak{A})$ has the **simplification property** if

$$[p] + [r] = [q] + [r] \ \Rightarrow\ [p] = [q].$$

Proposition 76.3.
This map is injective if and only if $V(\mathfrak{A})$ has the simplification property.

Proof. First, we suppose that $\kappa_{\mathfrak{A}}$ is injective and we suppose the equality $[p] + [r] = [q] + [r]$ for a certain $[r] \in V(\mathfrak{A})$. This implies $\Big[[p],[0]\Big] = \Big[[q],[0]\Big]$ and then injectivity of $\kappa_{\mathfrak{A}}$ gives $[p] = [q]$ and the simplification property is proved.

Now we suppose that $\mathfrak{A}$ is such that $V(\mathfrak{A})$ fulfils the simplification property. Let $[p]$ and $[q]$ such that $\kappa_{\mathfrak{A}}([p]) = \kappa_{\mathfrak{A}}([q])$, or $\Big[[p],[0]\Big] = \Big[[q],[0]\Big]$ which means that $\Big([p],[0]\Big) \sim \Big([q],[0]\Big)$. From the definition of the classes, it provides the existence of a $[r] \in V(\mathfrak{A})$ such that $[p] + 0 + [r] = [q] + [r]$ which in turn proves that $[p] = [q]$ by the simplification rule. □

Notice that every element in $K_0(\mathfrak{A})$ can be written under the form of a difference

$$\kappa_{\mathfrak{A}}([p]) - \kappa_{\mathfrak{A}}([g]),$$

namely,

$$\Big[[p],[q]\Big] = \Big[[p],[0]\Big] + \Big[[0],[q]\Big] = \kappa_{\mathfrak{A}}[p] - \Big[[q],[0]\Big] = \kappa_{\mathfrak{A}}[p] - \kappa_{\mathfrak{A}}[q].$$

76.1.1 Example

In the case $\mathfrak{A} = \mathbb{C}$, the space $V(\mathbb{C})$ is determined by the simple following result.

Proposition 76.4.
Two projectors of $p, q \in M_N(\mathbb{C})$ are equivalent if and only if the dimensions of their image are equal.

Proof. From what was said about projectors on page 3752 , there exists a $u \in M_N(\mathbb{C})$ such that $p = uu^*$. Moreover p and q are both diagonalisable, and from the dimension assumption, their diagonal form are the same. In particular, $\mathrm{Tr}(p) = \mathrm{Tr}(q)$ is their common dimension of image space.

If $p = uu^*$, then p can be written as $u_A u_A^*$ with $u_A = uA$ for every matrix A such that $AA^* = \mathbb{1}$. We are going to find such a matrix A for which $q = u_A^* u_A$, i.e. $q = A^*(u^*u)A$. Notice that u^*u and A are diagonalisable; the proposition should be proved if their diagonal form are the same: in that case A is the unitary matrix which diagonalises u^*u in the basis in which q is diagonal.

We have $(u^*u)^2 = u^*uu^*u = u^*pu$, and using the fact that $p^2 = p$, we find $(u^*u)^3 = (u^*u)^2$, so that the eigenvalues of u^*u are 0 and 1. From that and the assumptions, we deduce

$$\dim\big(\mathrm{Image}(u^*u)\big) = \mathrm{Tr}(u^*u) = \mathrm{Tr}(uu^*) = \mathrm{Tr}(p) = \dim\big(\mathrm{Image}(p)\big) = \dim\big(\mathrm{Image}(q)\big),$$

so that $\dim\big(\mathrm{Image}(q)\big) = \dim\big(\mathrm{Image}(u^*u)\big)$ and they have the same diagonal form. □

Since two projectors are equivalent if and only if the dimensions of their respective image are equal, the set of equivalence classes is $V(\mathbb{C}) = \mathbb{N}$. The Grothendieck group associated with the semigroup $\mathbb{N}$ is $\mathbb{Z}$, so that $K_0(\mathbb{C}) = \mathbb{Z}$.

76.1.2 Universality

Proposition 76.5.
Let G be an abelian group and $\Phi\colon V(\mathfrak{A}) \to G$, a semigroup homomorphism such that $\Phi\big(V(\mathfrak{A})\big)$ is invertible in G. Then Φ extends in an unique way to an homomorphism $\Psi\colon K_0(\mathfrak{A}) \to G$ such that $\Psi \circ \kappa_{\mathfrak{A}} = \Phi$.

In other words, one can invent the map Ψ is such a way the following diagram commutes:

$$\begin{array}{ccc} K_0(\mathfrak{A}) & & \\ {\scriptstyle \kappa_{\mathfrak{A}}}\downarrow & \overset{\Psi}{\searrow} & \\ V(\mathfrak{A}) & \xrightarrow[\Phi]{} & G. \end{array}$$

Proof. For unicity, consider $\Psi_1, \Psi_2\colon K_0(\mathfrak{A}) \to G$, two maps extending the map Φ, so

$$\Psi_1\big([[p],[q]]\big) = \Psi_1\big(\kappa_{\mathfrak{A}}[p] - \kappa_{\mathfrak{A}}[q]\big) = \Phi[p] - \Phi[q] = \Psi_2\big([[p],[q]]\big), \tag{76.5}$$

where we used the assumption $\Psi_1 \circ \kappa_{\mathfrak{A}} = \Phi$.

For existence, we define $\Psi\colon K_0(\mathfrak{A}) \to G$ by the formula

$$\Psi\big([[q],[p]]\big) = \Phi[q] - \Phi[p]. \tag{76.6}$$

Let us prove that it is well defined. Suppose that $\big([p],[q]\big) \sim \big([p'],[q']\big)$, and consider a $[r]$ such that condition (76.3) holds. Since $\Phi\big(V(\mathfrak{A})\big)$ is invertible in G, it makes sense to compute

$$\begin{aligned} \Psi\big([[p'],[q']]\big) - \Psi\big([[p],[q]]\big) &= \Phi[p'] - \Phi[q'] - \Phi[p] + \Phi[q] \\ &= -\big(\Phi[p'] + \Phi[q] + \Phi[r]\big) + \Phi[p'] + \Phi[q] + \Phi[r] \\ &= \Phi[0] \\ &= 0 \end{aligned} \tag{76.7}$$

where we used the fact that G is abelian in order to rearrange terms and the fact that Φ is an homomorphism. We have $\Psi \circ \kappa_{\mathfrak{A}} = \Phi$ because

$$\Psi\big(\kappa_{\mathfrak{A}}[p]\big) = \Psi\left(\big[[p],[0]\big]\right) = \Phi[p].$$

□

Proposition 76.6.
If $\alpha\colon \mathfrak{A} \to \mathcal{B}$ is an homomorphism of C^*-algebra, then the map

$$\begin{aligned} \alpha_*\colon V(\mathfrak{A}) &\to V(\mathcal{B}) \\ [a] &\mapsto [\alpha(a)] \end{aligned} \tag{76.8}$$

is a well defined group homomorphism which extends to a group homomorphism $\alpha_*\colon K_0(\mathfrak{A}) \to K_0(\mathcal{B})$.

Proof. Let consider a projector $(a_{ij}) \in M_\infty(\mathfrak{A})$. Since α is an homomorphism, the matrix $\alpha(a_{ij})$ is a projector of $M_\infty(\mathcal{B})$. Since α preserves the involution (by definition of a C^*-algebra homomorphism), we have that $a \sim b$ implies $\alpha(a) \sim \alpha(b)$. Thus the induced map $\alpha_*\colon V(\mathfrak{A}) \to V(\mathcal{B})$ is an homomorphism.

In order to get the extension, we use the universality property. We know that α is a semigroup homomorphism and $\Phi = \kappa_{\mathcal{B}} \circ \alpha_*$ is a semigroup homomorphism from the semigroup $V(\mathfrak{A})$ and the group $K_0(\mathcal{B})$. Then universality provides $\Psi\colon K_0(\mathfrak{A}) \to K_0(\mathcal{B})$. □

Proposition 76.7.
If $\mathfrak{A}$ is a C^*-algebra obtained as the inductive limit

$$\mathfrak{A} = \varinjlim\{\mathfrak{A}_i, \Phi_{ij}\}_{i,j\in\mathbb{N}},$$

then $\{K_0(\mathfrak{A}_i), (\Phi_{ij})_*\}$ is an inductive system of groups and

$$K_0(\mathfrak{A}) = K_0\big(\varinjlim \mathfrak{A}_i\big) = \varinjlim K_0(\mathfrak{A}_i). \tag{76.9}$$

76.2 Vector bundle and K-theory

We follow [909] in which one finds the proofs that are omitted here. In this section, we only deal with complex vector bundles over compact Hausdorff base. We accept that, when the basis is non connected, the vector bundle has different dimensions on different components.

We denote by $\epsilon^n \to X$ the trivial bundle of dimension n over X and we say that two vector bundles E_1 and E_2 are **stably isomorphic** if there exists a n such that $E_1 \oplus \epsilon^n \simeq E_2 \oplus \epsilon^n$. We write it by $E_1 \simeq_s E_2$.

We also define the relation $E_1 \sim E_2$ if and only if $E_1 \oplus \epsilon^n \simeq E_2 \oplus \epsilon^m$ for some n and m. One can prove that $\simeq_s$ and $\sim$ are equivalence relations.

Proposition 76.8.
If X is compact and Hausdorff, the set of equivalence classes for $\sim$ is an abelian group for the direct sum.

Proof. No proof. □

That group is denoted by $\tilde{K}(X)$

The set of equivalence classes for $\simeq_s$ cannot be a group because the inverse does not exists. Indeed if $E \oplus E' \simeq_s \epsilon^0$, then we have a n such that $E \oplus E' \oplus \epsilon^n = \epsilon^n$, which implies that E and E' are zero dimensional.

Proposition 76.9.
The set of equivalence classes with respect to $\simeq_s$ has the simplification property

$$F \oplus E_1 \simeq_s F \oplus E_2 \Rightarrow E_1 \simeq_s E_2 \tag{76.10}$$

on every compact of X.

Proof. Using proposition 56.6, we can add on both sides of $F \oplus E_1 \simeq_s F \oplus E_2$ a vector bundle F' such that $F \oplus F' = \epsilon^n$. Thus we get $E_1 \oplus \epsilon^n \simeq_s E_2 \oplus \epsilon^n$, which means that $E_1 \simeq_s E_2$. □

We know that the positive rational numbers $\mathbb{Q}^+$ are build from the integers by taking the pairs $(a, b) \in \mathbb{N}^2$ and the quotient by

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc.$$

We perform the same construction with vectors bundle and we define $K(X) = \{(E, E')\}/\sim$ where

$$(E_1, E_1') \sim (E_2, E_2') \Leftrightarrow E_1 \oplus E_2' \simeq_s E_2 \oplus E_1'. \tag{76.11}$$

Let us prove that that relation is transitive. For this, we suppose $(E_1, E_1') \sim (E_2, E_2')$, and $(E_2, E_2') \sim (E_/, E_3')$, i.e. there exists integers n and q such that

$$E_1 \oplus E_2' \oplus \epsilon^n \oplus E_3' \oplus \epsilon^q = E_2 \oplus E_1' \oplus \epsilon^n \oplus E_3' \oplus \epsilon^q.$$

Now we use the relation $E_2 \oplus E_3' \oplus \epsilon^q = E_3 \oplus E_2' \oplus \epsilon^q$ in the right hand side and we use the simplification property by E_2', we get

$$E_1 \oplus E_3' \oplus \epsilon^{n+q} = E_3 \oplus E_1' \oplus \epsilon^{n+q},$$

which means that $(E_1, E_1') \sim (E_3, E_3')$. Notice that the capacity assumption was used when we made the simplification by E_2. The simplification would have added a ϵ^m which was implicitly absorbed in the n or the q.

In that context, the class of the pair (E, E') is often denoted by $E - E'$. The set $K(X)$ becomes a group for the addition

$$(E_1 - E_1') + (E_2 - E_2') = (E_1 \oplus E_2) - (E_1' \oplus E_2'). \tag{76.12}$$

The zero of the group $K(X)$ is $E - E$ for any E and the inverse of $E - E'$ is $E' - E$. Since $E \oplus (E' \oplus F) = E' \oplus (E \oplus F)$, we have

$$E - E' = (E \oplus F) - (E' \oplus F). \tag{76.13}$$

Taking an inverse of E' as F, we find that $E - E' = E \oplus F - \epsilon^n$, so that every element of $K(X)$ can be written as a difference

$$E - \epsilon^n.$$

PropvphomKtK

Proposition 76.10.
The map

$$\begin{aligned} \varphi\colon K(X) &\to \tilde{K}(X) \\ E - \epsilon^n &\mapsto [E]_\sim \end{aligned} \tag{76.14}$$

is a surjective homomorphism.

Proof. We prove first that it is well defined. Suppose that $E - \epsilon^n = E' - \epsilon^m$, so $E \oplus \epsilon^m = E' \oplus \epsilon^m$, so that $E' \sim E$. Surjectivity is clear. □

The kernel of that map of made from the differences $E - \epsilon^n$ with $E \sim \epsilon^0$, or in other words, $E \simeq_s \epsilon^m$ for a certain m,

$$\ker(\varphi) = \{\epsilon^m - \epsilon^n\} \subset K(X). \tag{76.15}$$

This is isomorphic to $\mathbb{Z}$.

76.2.1 Reduced group

Let us fix a base point $x_0 \in X$. We have a natural homomorphism $K(X) \to K(x_0) = \mathbb{Z}$. By the proposition 76.10, this homomorphism restricts to an isomorphism $\{\epsilon^m - \epsilon^n\} \to \mathbb{Z}$. Thus we have a decomposition

$$K(X) = \tilde{K}(X) \oplus \mathbb{Z} \tag{76.16}$$

which depends on the choice of the base point x_0. For this reason, we says that $\tilde{K}(X)$ is **reduced** of $K(X)$.

Since $K(X)$ is an additive group, we can turn it into a ring by the following multiplication:

$$(E_1 - E_1')(E_2 - E_2') = E_1 \otimes E_2 - E_1 \otimes E_2' - E_1' \otimes E_2 + E_1' \otimes E_2'. \tag{76.17}$$

One can check that $K(X)$ becomes a ring with ϵ^1 as identity.

76.2.2 Functorial description

If $f\colon X \to Y$ is a map between two compact Hausdorff topological spaces, we have an induces map

$$\begin{aligned} f^*\colon K(Y) &\to K(X) \\ E - E' &\mapsto f^*(E) - f^*(E') \end{aligned} \tag{76.18}$$

which satisfies

(1) $f^*(E_1 \oplus E_2) \simeq f^*(E_1) \oplus f^*(E_2)$,

(2) $f^*(E_1 \otimes E_2) = f^*(E_1) \otimes f^*(E_2)$,

and is then a ring homomorphism. We are then lead to see K as a functor between

— the category of compact Hausdorff topological spaces with continuous maps as arrows, $\mathscr{C}$, and

— the category of rings, $\mathscr{D}$.

The functor K makes the correspondence $X \to K(X)$ for the objects and for the arrows,

$$\begin{aligned} Kf\colon K(X) &\to K(Y) \\ KF &\mapsto (f^{-1})^*. \end{aligned} \tag{76.19}$$

Since $K(\mathrm{Id}_X) = \mathrm{Id}_{K(X)}$ and $K(g \circ f) = Kg \circ Kf$ for every $f\colon X \to Y$ and $g\colon Y \to Z$, the operation K is a functor.

76.2.3 External product

One has the external product

$$\begin{aligned} \mu\colon K(X) \otimes K(Y) &\to K(X \times Y) \\ \mu(a \otimes b) &= \mathtt{proj}_X^*(a)\, \mathtt{proj}_Y^*(b) \end{aligned} \tag{76.20}$$

where $\mathtt{proj}_X$ and $\mathtt{proj}_Y$ are the projections of $X \times Y$ onto X and Y. So we have $\mathtt{proj}_X^*\colon K(X) \to K(X \times Y)$ and the product $\mathtt{proj}_X^*(a)\, \mathtt{proj}_Y^*(b)$ is the internal product in $K(X \times Y)$. It is a general fact that tensor product of rings is a ring with the rule

$$(a \otimes b)(c \otimes d) = ac \otimes bd. \tag{76.21}$$

In our case, $K(X) \otimes K(Y)$ is a ring. We are going to prove that μ is a ring homomorphism. Let us denote by $f^*(E)$ the function given by proposition 56.7. We have

$$\mu\big((a \otimes b)(c \otimes d)\big) = \mu(ac \otimes bd) = \mathtt{proj}_X^*(ac)\, \mathtt{proj}_Y^*(bd).$$

Now suppose that $a = E_a - E'_a$ and $c = E_c - E'_c$. We have $ab = E_a \otimes E_c - E_a \otimes E'_c = E'_a \otimes E_c + E'_a \otimes E'_c$, and using the properties (56.15), we find

$$\begin{aligned}\mu\big((a\otimes b)(c\otimes d)\big) &= \mathtt{proj}_X^*(a)\,\mathtt{proj}_X^*(c)\,\mathtt{proj}_Y^*(b)\,\mathtt{proj}_Y^*(d)\\ &= \mathtt{proj}_X^*(a)\,\mathtt{proj}_Y^*(b)\,\mathtt{proj}_X^*(c)\,\mathtt{proj}_Y^*(d)\\ &= \mu(a\otimes b)\mu(c\otimes d),\end{aligned}$$

and μ is an homomorphism. We have in particular an homomorphism

$$\mu\colon K(X)\otimes K(S^2)\to K(X\otimes S^2). \tag{76.22}$$

EamuhomSdeuxX

The main content of **Bott periodicity** is to prove that (76.22) is in fact an isomorphism.

76.2.4 Clutching function

Let $p\colon E\to X$ be a vector bundle and $f\colon E\times S^1\to E\times S^1$ be an homomorphism where $E\times S^1$ denotes the bundle $(p\times \mathrm{Id})\colon E\times S^1\to X\times S^1$. For each $x\in X$ and $z\in S^1$, the map f produces an isomorphism $f(x,z)\colon p^{-1}(x)\to p^{-1}(x)$. Notice that, the S^1 part being trivial, we immediately restrict to X. If not, we would have written $(p\times \mathrm{Id})^{-1}(x,z)$.

If we identify the boundary of the two copies of D^2 in $D^2\cup D^2$ (i.e. we identify the two copies of S^1), what we obtain is the sphere S^2. So we take two copies of $E\times D^2$ and we identify the boundaries S^1 with f. What we get is a vector bundle over $X\times S^1$ that we name $[e,f]$. The function f is the **clutching** for $[E,f]$.

Let $f_t\colon E\times S^1\to E\times S^1$, be an homotopy of clutching functions. It allows us to build a vector bundle on $X\times S^2\times I$ (I is the interval $[0,1]$ in which t varies) which restricts to $[E,f_0]$ and $[E,f_1]$ on $X\times S^2\times\{0\}$ and $X\times S^2\times\{1\}$. So we have $[E,f_0]\simeq[E,f_1]$.

Chapter 77

BF theory

77.1 BF theory and relativity

References for BF theory, Palatini formalism and spin foam models are [1008, 1009, 1010, 1011]. Let M be a smooth n-dimensional manifold and consider the principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \downarrow \\ & & M \end{array} \tag{77.1}$$

with G being a connected semisimple Lie group with Lie algebra $\mathcal{G}$. We consider the following as the basics fields of the theory:

— a connection ω on P,

— a $\mathrm{Ad}(P)$-valued $(n-2)$-form E on M.

In order to form the wedge product $E \wedge F$, we express both fields as $\mathcal{G}$-valued forms on M using a local coordinate chart $\sigma_\alpha : \mathcal{U}_\alpha \to P$ (see pages 3402 and 3412). We pose $A_\alpha = \sigma_\alpha^* \omega$ and $F_\alpha = \sigma_\alpha^* \Omega$ where Ω is the curvature of ω.

There exists a $\eta \in \bigwedge^{n-2} T_x^* M$ such that $E(x) = [\xi, X] \otimes \eta$; using the same section σ_α, we find an element $X_\alpha \in \mathcal{G}$ by the condition

$$E(x) = [\sigma_\alpha(x), X_\alpha] \otimes \eta.$$

Hence we see E as a $(n-2)$ form with values in $\mathcal{G}$ as

$$E_\alpha(X_1, \cdots, X_{n-2}) = \eta(X_1, \cdots, X_{n-2}) X_\alpha(x). \tag{77.2}$$

We are now able, using the wedge product of subsection 56.4, to see $E_\alpha \wedge F_\alpha$ as a n-form with values in $\mathcal{G} \otimes \mathcal{G}$. The use of the Killing form $\mathcal{G} \otimes \mathcal{G} \to \mathbb{R}$ is what we write Tr. From now we drop all indices of local coordinates.

When we write $E \wedge F$, we mean the 4-form defined by

$$\begin{aligned} (E \wedge F)(v_1, v_2, v_3, v_4) = {} & E(v_1, v_2) \otimes F(v_3, v_4) - E(v_1, v_3) \otimes F(v_2, v_4) \\ & + E(v_1, v_4) \otimes F(v_2, v_3) + E(v_2, v_3) \otimes F(v_1, v_4) \\ & - E(v_2, v_4) \otimes F(v_1, v_2) + E(v_3, v_4) \otimes F(v_1, v_2). \end{aligned}$$

Obviously $E \wedge F \neq F \wedge E$, but taking the trace (i.e changing the tensor product by the Killing form),

$$\mathrm{Tr}(E \wedge F + F \wedge E) = 2\,\mathrm{Tr}(E \wedge F). \tag{77.3}$$

The action that we want to study first is

$$S(E, \omega) = \int_M \mathrm{Tr}(E \wedge F) \tag{77.4}$$

where F is the curvature associated with the connection ω. When we vary that action with respect to E, i.e. when we impose

$$\frac{d}{dt}\Big[S(E+tE',\omega)\Big]_{t=0}=0$$

for every $(n-2)$ form E', we find $F=0$ as equation of motion. A small computation shows that

$$\frac{d}{dt}\Big[F(\omega+t\omega')\Big]_{t=0}=dA'+A\wedge A'=d_AA',$$

where d_A denotes the exterior covariant derivative associated with the connection ω. Now the action principle imposes that for every connection ω',

$$0=\frac{d}{dt}\Big[S(E,\omega+t\omega')\Big]_{t=0}=\int_M \mathrm{Tr}(E\wedge d_AA')=\int_M \mathrm{Tr}(d_AE\wedge A')$$

where we used an integral by part. So the equation of motion are

$$F=0 \qquad \text{and} \qquad d_AE=0. \tag{77.5}$$

The inclusion of a cosmological constant in the game is done by modifying the action as

$$S_{BF}(A,E)=\int_M \mathrm{Tr}\left(E\wedge F+\frac{\Lambda}{12}E\wedge E\right). \tag{77.6}$$

Using the remark above, the equations of motion for the BF action are

$$d_AE=0 \qquad \text{and} \qquad F+\frac{\Lambda}{6}E=0. \tag{77.7}$$

The BF action with vanishing cosmological constant, $S_{BF}(\omega,E)=\int_M \mathrm{Tr}(E\wedge F)$ has an interesting symmetry. Let η be a $(n-3)$-$\mathrm{Ad}(P)$-valued form, and consider the transformation

$$A\mapsto A \qquad \text{and} \qquad E\mapsto E+d_A\eta. \tag{77.8}$$

EqSymPalaAEdAeta

We have

$$\int_M \mathrm{Tr}\left(E\wedge F+d_A\eta\wedge F\right)=\int_M \mathrm{Tr}\left(E\wedge F+(-1)^{n-3+1}\eta\wedge d_AF\right)$$

where we used an integral by part. The Bianchi identity $d_AF=0$ makes the last term vanishes, so that transformation (77.8) actually is a symmetry of the Palatini action without cosmological constant.

77.2 Palatini formalism

Let M be a smooth four dimensional manifold on which we consider an oriented vector bundle $\mathcal{T}$ which is isomorphic to TM (without being *canonically* isomorphic) endowed with a nondegenerate metric η with fixed signature. We say that $\mathcal{T}$ is the **internal space** and that η is the **internal metric**.

The basic fields of the theory are

— a $\mathcal{T}$-valued 1-form e on M,

— and a metric-preserving connection ∇ on $\mathcal{T}$.

One can define a metric on M from the internal metric by the formula

$$g(X,Y)=\eta\big(e(X),e(Y)\big). \tag{77.9}$$

In order to do that, we have to see e as a map $e\colon TM\to\mathcal{T}$. We know that $e\in\Gamma(\mathcal{T}\otimes T^*M)$, so that for each $x\in M$, we can write $e(x)=\sum_i v_i\otimes\omega_i$, so that we define

$$e(x)X=\sum_i \omega_i(X)v_i$$

for all $X \in T_xM$. Notice that g is nondegenerate if and only if e is an isomorphism.

We are now going to express the basic fields as forms with values in the exterior algebra bundle $\bigwedge \mathcal{T}$. For e, it is easy: if $e(x) = v_i \otimes \omega_i$, we consider $e(x) = v_i^* \otimes \omega_i$.

We know from propositions 56.63 and 56.67 that, with a choice of a local trivialization, the connection can be expressed in terms of a $\mathfrak{gl}(V)$-valued 1-form A_α on $\mathcal{U}_\alpha \subset M$ where V is the vector space on which $\mathcal{T}$ is modeled. Here, the connection form takes its values in $\mathfrak{so}(\eta)$ because of the assumption of metric preserving.

Passing to the curvature, we have $F \in \Gamma(\mathfrak{so}(\eta) \otimes \bigwedge^2 T^*M)$. Now an element of $\mathcal{G} = \mathfrak{so}(\eta)$ can be seen as an element of $\bigwedge^2 \mathcal{T}$ by the following formula:

$$S(v, w) = \eta(v, Sw) - \eta(Sv, w) \tag{77.10}$$

where $S \in \mathfrak{so}(\eta)$ and v, $w \in \mathcal{T}_x$. In fact, the knowledge of that 2-form allows to rebuild the element S because when v and w run over a basis $\{e_i\}$ of V, we find the numbers

$$(Se_j)_i - (Se_i)_i = S_{ij} - S_{ji} = 2S_{ij}$$

because S has to be skew symmetric.

The orientation of M provides an everywhere non vanishing volume form, this is an element of $\bigwedge^4 \mathcal{T}$. Now if β is another element of $\bigwedge^4 \mathcal{T}$, it has to be a multiple of μ. When H is a $\bigwedge^4 \mathcal{T}$-valued n-form on M, we define $\mathrm{Tr}(H)$ as the usual n-form on M defined by

$$H(X_1, \cdots, X_n) = \mathrm{Tr}(H)(X_1, \cdots, X_n)\mu.$$

Now the following action, which is the **Palatini action**, makes sense as integral of a n-form on a n-dimensional manifold:

$$S_{Pal}(\nabla, e) = \int_M \mathrm{Tr}\left(e \wedge e \wedge F + \frac{\Lambda}{12} e \wedge e \wedge e \wedge e\right). \tag{77.11}$$

Here the wedge product is a composition of wedge product as forms *with values in* $\bigwedge \mathcal{T}$ and the wedge product as *elements of* $\bigwedge \mathcal{T}$, so that

$$(e \wedge f)(X, Y) = e(X) \wedge f(y) - e(Y) \wedge f(X)$$

where the wedge of the right hand side is taken in the sense of $\bigwedge \mathcal{T}$, so that we actually have $e \wedge f = f \wedge e$. The variation with respect to ∇ provides the equation of motion

$$d_A(e \wedge e) = 0, \tag{77.12}$$

while the variation with respect to e provides

$$0 = \frac{d}{dt}\Big[S_{Pal}(\nabla, e + te')\Big]_{t=0} = \int_M \mathrm{Tr}\left(2e' \wedge e \wedge F + \frac{\Lambda}{12} 4e' \wedge e \wedge e \wedge e\right),$$

or

$$e \wedge \left(F + \frac{\Lambda}{3} e \wedge e\right) = 0. \tag{77.13}$$

77.2.1 Other point of view

Let M be the four dimensional physical space and E, a four dimensional vector bundle. We suppose to know a local trivialization $E = M \times \mathbb{R}^4$ and we consider the Lorentzian metric η on each of the $\mathbb{R}^4$. Finally we consider a soldering form e: for each point x, the map $e_x \colon T_xM \to E_x$ is an isomorphism. If ∂_i is the local basis of TM, we use the local basis $\xi_i = e(\partial_i)$ as local coordinates on E.

These data allows us to define a metric on M by

$$g(x, Y) = \eta(eX, eY). \tag{77.14}$$

The second dynamical variable will be an exterior derivative D on E. Any section $s \in \Gamma(E)$ reads $s(x) = s^i \xi_i$ and the action of the exterior derivative reads

$$D_X s = \big(X(s^i) + s^j \omega_j^i(X)\big)\xi_i. \tag{77.15}$$

At this point, we cannot define a torsion for D because we do not have a canonical soldering form on E. However we pose the hypothesis that D preserves the metric in the sense that

$$d_X\big(\eta(s,t)\big) = \eta(D_X s, t) + \eta(s, D_X t) \tag{77.16}$$

EqDefDXpreserveMetr

for all s, $t \in \Gamma(E)$. Since $\eta(s,t)$ is a function on M, we have $D_X\big(\eta(s,t)\big) = X\big(\eta(s,t)\big)$. So we have on the one hand $\eta(D_X s, t) = \eta_{ij}\Big(X(s^i) + s^k \omega_k^i\Big)t^j$ and on the other hand,

$$D_X\big(\eta(s,t)\big) = X^i \eta_{kl}\big(\partial_i(s^k)t^l + s^k \partial_i t^l\big).$$

Putting these two expressions into (77.16), we find

$$\eta_{ij}\omega_k^i = -\eta_{ki}\omega_j^i, \tag{77.17}$$

so that $\omega \in \Omega^1(M, \mathfrak{so}(1,3))$. Now we use the isomorphism e to build a connection on TM:

$$\nabla_X Y = e^{-1}\big(D_X(eY)\big) \tag{77.18}$$

where we see eY as a section of E by $(eY)(x) = e(Y_x)$. The Christoffel symbols of ∇ and D are defined as usual by $\nabla_{\partial_i}(\partial_j) = \tilde{\Gamma}_{ij}^k \partial_k$, and $\omega_j^k(\partial_i) = \Gamma_{ij}^k$. It is simple to see that numerically, $\Gamma_{ij}^k = \tilde{\Gamma}_{ij}^k$. Indeed

$$D_{\partial_i}(\xi_j) = \xi_k \otimes \big(\partial_i(\delta_j^k) + \delta - j^l \omega_l^k(\partial_i)\big) = \xi_k \otimes \omega + j^k(\partial_i) = \xi_k \otimes \Gamma_{ij}^k,$$

so that $\nabla_{\partial_j}(\partial_j) = e^{-1}\big(\xi_k \otimes \Gamma_{ij}^k\big) = \Gamma_{ij}^k \partial_k$.

The connection ∇ is compatible with the metric because

$$\begin{aligned}
\nabla_Z\big(g(X,Y)\big) &= Z\big(\eta(eX, eY)\big) \\
&= \eta\big(\underbrace{D_Z(eX)}_{=e(\nabla - ZX)}, eY\big) + \eta\big(eX, D_Z(eY)\big) \\
&= g(\nabla_Z X, Y) + g(X, \nabla_Z Y).
\end{aligned}$$

Chapter 78

BTZ black hole from identifications

78.1 Group realization of low dimensional anti de Sitter spaces

78.1.1 Two dimensional anti de Sitter

SubsecTwoDimAdSAdGH

Using notations and conventions of section 53.17, the set $\mathrm{Ad}(G)H$ is a subset of $\mathfrak{sl}(2,\mathbb{R})$ made of elements of norm 8. One can show by brute force computation or using the commutation relations that

$$\begin{aligned}\mathrm{Ad}(e^{x_K T}e^{x_N E})H = &\big(-\sin(2x_K)x_N+\cos(2x_K)\big)H\\ &+\big(-\cos(2x_K)x_N-\sin(2x_K)\big)(E+F)\\ &-x_N T,\end{aligned}$$

which is a, following the Killing form (53.328), general element of norm 8 in $\mathfrak{sl}(2,\mathbb{R})$. Thus as sets, AdS_2 is $\mathrm{Ad}(KN)H$. Since A is the stabilizer of H for the adjoint action of G on H and $G=ANK=KNA$, we also have

$$AdS_2 = G/A = \mathrm{Ad}(KN)H = \mathrm{Ad}(G)H.$$

That provides isomorphisms

$$\begin{aligned}\phi\colon [0,\pi[\times\mathbb{R} &\to AdS_2\\ (x_K,x_B) &\mapsto \mathrm{Ad}(e^{xKT}e^{xNE})H,\end{aligned} \tag{78.1}$$

or

$$\begin{aligned}\phi\colon Cyl &\to AdS_2\\ (\theta,h) &\mapsto \mathrm{Ad}(e^{\frac{\theta}{2}T}e^{hE})H.\end{aligned} \tag{78.2}$$

EqCylAdSDeux

where Cyl is the usual cylinder in $\mathbb{R}^3$.

78.1.2 Three dimensional anti de Sitter

SubsecGpAdsDeux

The space AdS_3 is the hyperboloid

$$u^2+t^2-x^2-y^2=1 \tag{78.3}$$

hyperboloide

embedded in $\mathbb{R}^{2,2}$. There exists a bijection between $\mathbb{R}^{2,2}$ and the two by two real matrices given by

$$v=\begin{pmatrix}u\\t\\x\\y\end{pmatrix}\mapsto g(v)=\begin{pmatrix}u+x & y+t\\ y-t & u-x\end{pmatrix}.$$

As far as norm is concerned, we have $\|v\| = \det g(v)$. Among these matrices, the ones of $\mathrm{SL}(2,\mathbb{R})$ are given by the condition $\det g = 1$, which is precisely the equation of the hyperboloid in $\mathbb{R}^4$. That shows that $AdS_3 = \mathrm{SL}(2,\mathbb{R})$.

From (78.3), the isometry group of AdS_3, denoted by $\mathrm{Iso}(G)$, is the group $O(2,2)$. It is locally isomorphic to $G \times G$, through the action

$$(G \times G) \times G \longrightarrow G : ((g_L, g_R), z) \to g_l\, z g_R^{-1}, \tag{78.4}$$

which corresponds to the identity component of $Iso(G)$ (from the bi-invariance of the Killing metric), and because of the Lie algebra isomorphism

$$\Phi : \mathcal{G} \times \mathcal{G} \to iso(G) : (X, Y) \to \overline{X} - \underline{Y}, \tag{78.5}$$

Iso

where $\overline{X}$ (resp. $\underline{Y}$) denotes the right invariant (resp. left invariant) vector field on G associated with the element X (resp. Y) of its Lie algebra.

78.2 Simple example on AdS_2

sec_AdSdeux

As is subsection 78.1.1, we see AdS_2 as $\mathrm{Ad}(G)H$. By definition the singularity is composed of closed orbits of AN and $A\bar{N}$ for the adjoint action on AdS_2, and the notion of fundamental field is

$$H_x^* = \frac{d}{dt}\Big[\,\mathrm{Ad}(e^{-tH})x\Big]_{t=0} = -[H, x]. \tag{78.6}$$

A basis of the Lie algebra $\mathcal{A} \oplus \mathcal{N}$ is given by $\{E, H\}$. So x will belong to a closed orbit if and only if $E_x^* \wedge H_x^* = 0$. If we put $x = x_H H + x_E E + x_F F$, the computation is

$$E_x^* \wedge H_x^* = [E, x] \wedge [H, x] = 4x_H x_F (E \wedge F) + 2x_E x_F (H \wedge E) - 2x_F^2 (H \wedge F).$$

It is zero if and only if $x_F = 0$. The closed orbit of $A\bar{N}$ is given by the same computation with $H_x^* \wedge F_x^*$. The part of these orbits contained in AdS_2 is the one with norm 8:

$$B(x, x) = 8(x_H^2 + x_E x_F) \stackrel{!}{=} 8.$$

In both cases, it imposes $x_H = \pm 1$, and the closed orbits in AdS_2 are given by

EqQuatreLInesDeux

$$\mathscr{S}_{AN} = \pm H + \lambda E \tag{78.7a}$$
$$\mathscr{S}_{A\bar{N}} = \pm H + \lambda F, \tag{78.7b}$$

with $\lambda \in \mathbb{R}$. The singularity is then the union $\mathscr{S}_{AN} \cup \mathscr{S}_{A\bar{N}}$ of four lines in the hyperboloid.

PropAdSDeuxJannule

Proposition 78.1.
The singularity of AdS_2 can equivalently be described as

$$\mathscr{S} = \{x \in \mathrm{Ad}(G)H \ tel\ que\ \|H_x^*\| = 0\}, \tag{78.8}$$

condHH

This result is also proved for AdS_3 in corollary 79.20.

Proof. The condition (78.8) on x reads

$$B([H, x], [H, x]) = 0. \tag{78.9}$$

eqBHxHx

The most general [1] element x in $\mathfrak{sl}(2, \mathbb{R})$ is $x = x_A H + x_N E + x_F F$. We have $[x, H] = -2x_N E + 2x_F F$, so that the condition (78.9) becomes $x_N x_F = 0$. The two possibilities are $x = x_A H + x_N E$ and $x = x_A H + x_F F$. The singularity in $\mathfrak{sl}(2, \mathbb{R})$ is composed of the planes (H, F) and (H, E). The intersection between the plane (H, F) and the hyperboloid is given by the equation

$$B(aH + bF, aH + bF) = 8$$

whose solutions are $a = \pm 1$. The same is also true for the plane (H, E). So we find that the set (78.8) is exactly the four lines (78.7).

□

1. It is actually *more* than the most general element to be considered because our space is $\mathrm{Ad}(G)H$, which is only a part of $\mathfrak{sl}(2, \mathbb{R})$.

One can check that light cone of a given point of the hyperboloid is given by the two straight lines trough the point; so it automatically intersects the singularity. As conclusion, every point of AdS_2 belong to the black hole. For this reason we say that there is no black hole in the two dimensional case because the black hole should be the whole space while one ask the singularity, the black hole and the complete space to be different. See also the condition (79.49).

78.3 Still AdS_2

If h is the generator of $\mathrm{SO}(1,1)$ in the quotient $AdS_2 = \mathrm{SO}(2,1)/\mathrm{SO}(1,1)$, then we have

$$AdS_2 = \frac{\mathrm{SO}(2,1)}{\mathrm{SO}(1,1)} = \mathrm{Ad}(G)h \tag{78.10}$$

by the isomorphism

$$\psi[g] = \mathrm{Ad}(g)h. \tag{78.11}$$

The multiplication table of the algebra $\mathfrak{so}(2,1)$ is

$$\begin{aligned} [q_0, q_1] &= h \\ [h, q_0] &= q_1 \\ [h, q_1] &= q_0. \end{aligned} \tag{78.12}$$

Knowing the adjoint map, we can compute the Killing for. In particular we have, in the basis q_0, q_1, h,

$$B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{78.13}$$

Thus the norm on $\mathfrak{so}(2,1)$ is

$$\|aq_0 + bq_1 + ch\| = 2(-a^2 + b^2 + c^2), \tag{78.14}$$

and

$$AdS_2 = \{aq_0 + bq_1 + ch \text{ tel que } -a^2 + b^2 + c^2 = 2\}. \tag{78.15}$$

If $x = \mathrm{Ad}(g)h$, we consider the adjoint action of G over x and the fundamental vector fields

$$X_x^* = \frac{d}{dt}\Big[e^{-tX} \cdot \mathrm{Ad}(g)h\Big]_{t=0} = \frac{d}{dt}\Big[e^{-\mathrm{ad}(tX)}\,\mathrm{Ad}(g)h\Big]_{t=0} = -\,\mathrm{ad}(X)x, \tag{78.16}$$

so we have

$$X^*_{\mathrm{Ad}(g)h} = -\,\mathrm{ad}(X)\,\mathrm{Ad}(g)h. \tag{78.17}$$

For the metric, the first choice is to set

$$B(X_x^*, Y_x^*) = B(\mathrm{Ad}_{g^{-1}} X, \mathrm{Ad}_{g^{-1}} Y), \tag{78.18}$$

but that definition would not take the quotient into account. Indeed, we have $B(h,h) \neq 0$ while

$$h_h^* = -\,\mathrm{ad}(h)h = 0. \tag{78.19}$$

Thus we consider the metric

$$B(X_x^*, Y_y^*) = B\big(\mathtt{proj}_{\mathcal{Q}}(\mathrm{Ad}_{g^{-1}} X), \mathtt{proj}_{\mathcal{Q}}(\mathrm{Ad}_{g^{-1}} Y),\big). \tag{78.20}$$

We have in particular

$$\|X_h^*\| = \|(aq_0 + bq_1 + ch)_h^*\| = (aq_0 + bq_1) \cdot (aq_0 + bq_1) = -a^2 + b^2. \tag{78.21}$$

So the time-like direction at h is q_0^*. We can ask ourself in which direction it goes:

$$(q_0)_h^* = \frac{d}{dt}\Big[\operatorname{Ad}(e^{-tq_0})h\Big]_{t=0} = q_1. \tag{78.22}$$

That provides the orientation of time at the point $\operatorname{Ad}(e)h$. Now we would like to know the orientation of time on other points, like at the point

$$\operatorname{Ad}(e^{bq_1})h = \cosh(b)h - \sinh(b)q_0. \tag{78.23}$$

Since $\operatorname{Ad}(e^{-bq_1})q_1 = q_1$, we have

$$\|q_1^*\|_{\operatorname{Ad}(e^{bq_1})h} = \|q_1\| = 1, \tag{78.24}$$

the vector q_1^* is not time-like at the point $\operatorname{Ad}(e^{bq_1})h$. On the other hand, we have

$$\operatorname{Ad}(e^{-bq_1})q_0 = \cosh(b)q_0 + \sinh(b)h, \tag{78.25}$$

so

$$\|q_0^*\|_{\operatorname{Ad}(e^{bq_1})h} = \|\,\mathtt{proj}_{\mathcal{Q}}\big(\cosh(b)q_0 + \sinh(b)h\big)\| = -2\cosh(b), \tag{78.26}$$

which is time-like.

78.4 Causal structure of AdS_2

SecAdS2

As far as I know, results about AdS_2 comes from discussions with Pierre Bieliavsky. The approach that I followed for AdS_3 is the one of [1012, 1013] although results are already know from [1014, 1015].

The two dimensional case does not present black hole structure, but the geometric setting is very clear as it is no more that intersections of lines in an hyperboloid.

We consider the following situation: $G = SL(2,\mathbb{R})$, $H = \mathrm{SO}(1,1) = A$ (the A of Iwasawa for G), $AdS_2 \simeq G/A$. The generators of $SL(2,\mathbb{R})$ are given as usual by E, F, H and the generators of the groups A, N, K are H, E, and $E - F$. We set

$$\mathcal{S} = R \cdot A \cup \overline{R} \cdot A = \pi(N) \cup \pi(\overline{N}).$$

where the bar denotes the conjugation by the Cartan involution θ. The light cone of $\pi(g)$ in AdS_2 is given by

$$C^+_{\pi(g)} = \{\pi(ge^{tE}) \cup \pi(ge^{tF})\}_{t\in\mathbb{R}^+} \tag{78.27}$$

The study of the causal structure on this space is the following question: which are the $\pi(g) \in AdS_2$ such that $C^+_{\pi(g)} \cap \mathcal{S} \neq \varnothing$? One can write $g = kna$ and forget the a-part because all the expressions appears in a π.

Let us consider the ray $\pi(ge^{tE}) = \pi(kne^{tE})$. The aim is to see what choice of $k \in K$ and $n \in N$ makes $\pi(kne^{tE}) \in \mathcal{S}$ for a certain t. If we set

$$n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \qquad k = \begin{pmatrix} \cos k & \sin k \\ -\sin k & \cos k \end{pmatrix},$$

we find a general element of the ray (before the projections by π) under the form

$$kne^{tE} = \begin{pmatrix} \cos k & (t+n)\cos k + \sin k \\ -\sin k & -(t+n)\sin k + \cos k \end{pmatrix}$$

Taking the quotient into account, we have to see which are the $k \in [0, 2\pi]$, $n \in \mathbb{R}$ such that

$$\pi(kne^{tE}) = \begin{pmatrix} \cos k & (t+n)\cos k + \sin k \\ -\sin k & -(t+n)\sin k + \cos k \end{pmatrix} \begin{pmatrix} e^a & 0 \\ 0 & e^{-a} \end{pmatrix} \tag{78.28}$$

eq:chem_tE

can be fitted on $\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$ by a suitable choice of $t \in \mathbb{R}^+$ and $a \in \mathbb{R}$. The equations are

$$e^a \cos k = 1 \tag{78.29a}$$
$$-e^a \sin k = 0 \tag{78.29b}$$
$$e^{-a}\big((t+n)\cos k + \sin k\big) = l \tag{78.29c}$$
$$e^{-a}\big(-(t+n)\sin k + \cos k\big) = 1. \tag{78.29d}$$

Equation (78.29b) yields $\sin k = 0$, so that (78.29a) imposes the choice $a = 0$ and $\cos k = 1$. This makes $k = \mathbb{1}$, so that $\pi(g) \in \mathcal{S}$. Then as far as the singularity $A \cdot N$ is concerned, the direction e^{tE} is safe.

Now we study the possibility to fall into $A \cdot \overline{N}$ by walking in the direction e^{tE}, so we have to fit equation (78.28) with $\begin{pmatrix} 1 & 0 \\ f & 1 \end{pmatrix}$:

$$e^a \cos k = 1 \tag{78.30a}$$
$$-e^a \sin k = f \tag{78.30b}$$
$$e^{-a}\big((t+n)\cos k + \sin k\big) = l \tag{78.30c}$$
$$e^{-a}\big(-(t+n)\sin k + \cos k\big) = 1. \tag{78.30d}$$

Condition (78.30d) makes $\cos k < 0$ completely safe. Then we suppose $\cos k > 0$ and $e^a = 1/\cos k$. On the other hand, (78.30b) is always true because f can take any real value. The two last equations give the system

$$\begin{cases} (t+n)\cos k + \sin k = 0 \\ -(t+n)\sin k + \cos k = 0 \end{cases} \tag{78.31}$$

The first gives $t = -\frac{\sin k + n\cos k}{\cos k}$. When substituted in the second equation, we find $\sin^2 k + \cos^2 k = 0$.

Then the directions e^{tE} is completely secure.

The second possible direction in the light cone is $e^{tF} = \begin{pmatrix} 1 & 0 \\ f & 1 \end{pmatrix}$. It is rather easy to find

$$kne^{tF} = \begin{pmatrix} (1+nt)\cos k + t\sin k & n\cos k + \sin k \\ -(1+nt)\sin k + t\cos k & -n\sin k + \cos k \end{pmatrix}$$

The equations which explain us how to to fall into $\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$ are

$$t(\cos k - n\sin k) = \sin k \tag{78.32a}$$
$$e^a\big((1+nt)\cos k + t\sin k\big) = 1 \tag{78.32b}$$
$$e^{-a}\big(n\cos k + \sin k\big) = l \tag{78.32c}$$
$$e^{-a}(-n\sin k + \cos k) = 1. \tag{78.32d}$$

Substituting $(\cos k - n\sin k)$ from the first to the fourth equation, we find $e^a = \sin k/t$. The third equations is always true because l can take any value (note that $\sin k$ must be nonzero). If we suppose $\cos k - n\sin k \neq 0$, we find that the direction e^{tF} falls into $\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$ after a time $t = \frac{\sin k}{\cos k - n\sin k}$. Examining the case of $\cos k - n\sin k = 0$, we see that it is safe with respect to the hole $\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$ in the direction e^{tF}.

The last point is to check what are the points which fall into $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$ in the direction e^{tF}. The

equations are

$$e^a\big(t\cos k-(1+nt)\sin k\big)=c \tag{78.33a}$$
$$e^a\big((1+nt)\cos k+t\sin k\big)=1 \tag{78.33b}$$
$$e^{-a}\big(n\cos k+\sin k\big)=0 \tag{78.33c}$$
$$e^{-a}(-n\sin k+\cos k)=1. \tag{78.33d}$$

We immediately find that $n\cos k+\sin k\neq 0$ is safe and that $e^a=1/\cos k$. Substituting, we find $-n\sin k\cos k=1-\cos^2 k$ and $\sin k\neq 0$ is obligatory. Then there are no hope to fall down in this part of the back hole.

As conclusion, the only way for a light ray to fall into the black hole is to begin at $\pi(kn)$ with $\cos k-n\sin k\neq 0$ and to follow the direction e^{tF} during a time

$$t=\frac{\sin k}{\cos k-n\sin k}.$$

78.4.1 Second point of view

Proposition 78.2.
Any point in the physical space can be written as $\mathrm{Ad}(ak)H$*, with* $k\in]0,\pi/2[$. [prop:AdAK]

Proof. The physical space contains the curve $\cos\beta H+\sin\beta(E+F)$ with $\beta\in]0,\pi[$, which is exactly [2] $\mathrm{Ad}(k)H$ for $k\in]0,\pi/2[$. It is also the intersection of AdS_2 and the part of $\mathfrak{sl}(2,\mathbb{R})$ contained between the planes (E,H) and (F,H). If we use the coordinates x,y,z on $\mathfrak{sl}(2,\mathbb{R})$ (a point is given by $\overline{r}=(x,y,z)=xH+yE+zF$), the physical space is given by the relations

$$\begin{cases} x^2+yz=1 & (78.34)\\ y>0 & (78.35)\\ z>0. & (78.36)\end{cases}$$

The first equation gives a β such that $x=\cos^2\beta$, $yz=\sin^2\beta$. It is always possible to define a $a\in\mathbb{R}$ such that $y=e^{2a}\sin\beta$ and $z=e^{-2a}\sin\beta$. Finally, the physical space is parameterized by

$$(\beta,a)\mapsto\cos\beta H+\sin\beta(e^{2a}E+e^{-2a}F). \tag{78.37}$$

From commutation relations in $\mathfrak{sl}(2,\mathbb{R})$, one finds

$$\mathrm{Ad}(e^{aH})E=e^{2a}E, \tag{78.38a}$$
$$\mathrm{Ad}(e^{aH})F=e^{-2a}F, \tag{78.38b}$$

and then

$$\begin{aligned}\mathrm{Ad}(ak)H&=\mathrm{Ad}(e^{aH})(\cos\beta H+\sin\beta(E+F))\\&=\cos\beta H+\sin\beta(e^{2a}E+e^{-2a}F).\end{aligned} \tag{78.39}$$

□

78.4.2 Light cone

The light-like vectors of $\mathfrak{sl}(2,\mathbb{R})$ are E and F, so the light cone of the point $\mathrm{Ad}(g)H$ has two parts:

$$\mathrm{Ad}(g)\,\mathrm{Ad}(e^{tE})H\ \text{ and }\ \mathrm{Ad}(g)\,\mathrm{Ad}(e^{tF})H$$

whose are better written under the more compact form

$$C^+_{\mathrm{Ad}(g)H}=\{\mathrm{Ad}(g)\,\mathrm{Ad}(e^{t\epsilon E})H\}_{\substack{t>0\\ \epsilon=\mathrm{Id},\theta}} \tag{78.40}$$

2. Remark that the whole closed curve with β running from 0 to 2π is double covered by K.

where ϵ is the identity or the Cartan involution ($\theta(X) = -X^t$). Since $H_x^* = -[H, x]$, the equation of the cone which falls into the hole is

$$\|[H, \mathrm{Ad}(g)\,\mathrm{Ad}(e^{t\epsilon E})H]\| = 0. \tag{78.41}$$

It is somewhat easy to remark that for all X, Y in a Lie algebra $\mathfrak{g}$ and all automorphism φ the formula, $\varphi(\mathrm{Ad}(e^X)Y) = \mathrm{Ad}(e^{\varphi X})(\varphi Y)$ holds. Then

$$\mathrm{Ad}(e^{t\epsilon E})H = s(\epsilon)\epsilon(\mathrm{Ad}(e^{tE})H) \tag{78.42}$$

with

$$s(\epsilon) = \begin{cases} 1 & \text{if } \epsilon = \mathrm{Id}, \\ -1 & \text{if } \epsilon = \theta. \end{cases}$$

In order to see if the cone intersects the singularity, we have to solve –with respect to t– the equation

$$B\left(\left[\mathrm{Ad}(g)s(\epsilon)\epsilon(\mathrm{Ad}(e^{tE})H), H\right], \dots\right) = 0 \tag{78.43}$$

where the dots means that the same lies in the second entry of the Killing form. Using the invariance of the Killing form, the fact that $\epsilon^{-1} = \epsilon$ and equation (53.322d) that $\mathrm{Ad}(e^{tE})H = H - 2tE$, we rewrite the condition under the form

$$B\big([H - 2tE, s(\epsilon)\epsilon^{-1}\,\mathrm{Ad}(g^{-1})H], \dots\big) = 0. \tag{78.44}$$

If we pose $g = a^{-1}n^{-1}k^{-1}$ and $x = s(\epsilon)\epsilon kn \cdot H = hH + lE + fF$, the condition becomes

$$B\big([H - 2tE, x], [H - 2tE, x]\big) = 32f^2t^2 - 64fht - 32lf \tag{78.45}$$

The conclusion is that the point $\mathrm{Ad}(g)H$ with $g = a^{-1}n^{-1}k^{-1}$ falls into the hole in the direction ϵ if and only if

$$s(\epsilon)\epsilon kn \cdot H = hH + lE + fF$$

with $l + h^2 < 0$.

¡¡ Avertissement/question à la lectrice !! 78.3
Il faudrait démontrer que lesdits points ne sont jamais dans la lamelle.

78.4.3 The graphical way

¡¡ Avertissement/question au lecteur !! 78.4
C'est quoi la motivation pour dire que cette partie de l'espace est physique ?

Proposition 78.2 fix the form of a light cone in the physical space:

$$C_x^+ = \mathrm{Ad}(a)C_c^+, \tag{78.46}$$

which is of course also a line. A line can intersect an other line at only one point. Then if the light cone of a point intersect the singularity at two points, then it intersects the *two* lines which delimits the physical space.

$$\sharp(C_x^+ \cap \mathscr{S}) = \sharp\,\mathrm{Ad}(a)(C_c^+ \cap \mathscr{S}) = \sharp(C_c^+ \cap \mathscr{S}) = 2. \tag{78.47}$$

Then every point in the physical space begins and finishes their live in the singularity.

78.4.4 The black hole

The light cone of the point $\mathrm{Ad}(ak)H$ –which is a general point of the physical part– is given by $\mathrm{Ad}(ak)s(\epsilon)\epsilon(\mathrm{Ad}(e^{tE})H)$. The computation of $\mathrm{Ad}(ak)(H-2tE)$ and $-\mathrm{Ad}(ak)(-H+2tF)$ gives eq:lin_light

$$(\cos(2k)-t\sin(2k))H-e^{2a}(\sin(2k)+2t\cos^2 k)E-e^{-2a}(\sin(2k)-2t\sin^2 k)F \tag{78.48a}$$ eq:lin_light_a

and

$$(\cos(2k)-t\sin(2k))H-e^{2a}(\sin(2k)-2t\sin^2 k)E-e^{-2a}(\sin(2k)+2t\cos^2 k)F \tag{78.48b}$$

With respect to t, these are two straight lines, so they the intersection of AdS_2 and the tangent plane to AdS_2 at $\mathrm{Ad}(ak)H$.

It allows us immediately to infer the non-existence of a black hole structure for this choice of singularity. The light cone at $x\in AdS_2$ is given by the tangent plane C of AdS_2 at x. The part of the singularity passing by H is given by a vertical plane S. The intersection of these two planes is a line, and the intersection of a line with AdS_2 is two points. Then each of the two lines of $C\cap AdS_2$ intersect one of the two lines of $S\cap AdS_2$. The same is true for the other part of the singularity.

The conclusion is that both two lines of the light cone intersect the singularity passing by H *and* the one passing by $-H$. So any point comes from the singularity and returns to the singularity; no point is connected to the infinity.

The AdS_2-black hole is doomed.

78.4.5 Closed orbits

Let us show that the singularity are the closed orbits of AN and $A\bar{N}$ for the adjoint action on $AdS_2=\mathrm{Ad}(G)H$. A basis of AN is given by $\{E,H\}$. So x will belong to a closed orbit if and only if $E_x^*\wedge H_x^*=0$. If we put $x=x_HH+x_EE+x_FF$, the computation is

$$\begin{aligned}E_x^*\wedge H_x^* &= [E,x]\wedge[H,x]\\ &= 4x_Hx_FE\wedge F+2x_Ex_FH\wedge E-2x_F^2H\wedge F.\end{aligned} \tag{78.49}$$

It is zero if and only if $x_F=0$. The closed orbit of $A\bar{N}$ is given by the same computation with $H_x^*\wedge F_x^*$. The part of these orbits contained in AdS_2 is the one with norm 8:

$$B(x,x)=8(x_H^2+x_Ex_F). \tag{78.50}$$

In both cases, it gives $x_H=\pm1$, and the closed orbits in AdS_2 are given by

$$\pm H+\lambda F \tag{78.51a}$$
$$\pm H+\lambda E, \tag{78.51b}$$

whose just are the singularity previously given in (78.7).

78.4.6 Time orientation

The tangent vectors to the two lines of equations (78.48) give a basis of the tangent space $T_{\mathrm{Ad}(ak)H}AdS_2$:

$$-\sin(2k)H+2e^{2a}\sin^2 kE+2e^{-2a}\cos^2 kF \tag{78.52a}$$

and

$$-\sin(2k)H+2e^{2a}\sin^2 kE-2e^{-2a}\cos^2 kF. \tag{78.52b}$$ eq:deux_vec_tg

The norm of the first one is $8\sin^2 2k+32\sin^2 k\cos^2 k>0$ while the one of the second is $8\sin^2 2k(1-2\cos^2 k)$ whose sign is not well defined. We use the first one to define a **time-orientation** on AdS_2:

$$T=-\sin(2k)H+2e^{2a}\sin^2 kE+2e^{-2a}\cos^2 kF. \tag{78.53}$$

A vector $X\in T_xAdS_2$ is **future directed** when $B(T,X)>0$. Note that the vector (78.52b) is also future directed. Then the given light cone is well the future light cone.

Remarque 78.5.
We had identified $(\partial_H,\partial_E,\partial_F)$ with (H,E,F) as in any vector space.

At H, the future oriented light-like vectors are E and F because $T_H=2F$. It is important to remark that the ones at $-H$ are the same E and F. The light cone at $-H$ "go back" to H! It can seem counter-intuitive, but it is necessary because, in order to be a physical space, the part $k\colon \frac{\pi}{2}\to\pi$ must contains its own future. So the time orientation of the AdS_2-light cone in $\mathfrak{sl}(2,\mathbb{R})$ should be a little complicated...and at least uneasy to guess on a picture.

Let's see for example the destiny of a guy walking at $E+F$. Particularising (78.48), we find the two lines

$$a\equiv tH+(1-t)E+(1+t)F \tag{78.54a}$$
$$b\equiv tH+(1+t)E+(1-t)F. \tag{78.54b}$$

The time orientation is given by $H+E+F$ and one can check that a and b are future-directed. When $t=1$, the light cone intersect $H+2F$ and $H+2E$ and when $t=-1$, it intersects $-HG+2E$ and $-H+2F$.

In the general case, with $k\colon \frac{\pi}{2}\to\pi$, the line (78.48a) does

— intersect $-H+lE$ at $l>0$ and $t=\frac{\sin 2k}{2\sin^2 k}<0$,
— intersect $H+fF$ at $f>0$ and $t=-\frac{\sin k}{\cos k}>0$,
— don't intersect $-H+fE$,
— don't intersect $H+lE$.

78.4.7 Singularity and physical space

The two dimensional case is very special because it doesn't present a black hole structure. The particular structure directly appears in the groupal formalism [3].

We look at $M=\mathrm{Ad}(G)H$ on which $G=SL(2,\mathbb{R})$ acts by the adjoint action. The set $A=e^{\mathbb{R}H}$, which is the abelian part of G with respect to the Iwasawa decomposition, is the stabilizer of H. Hence the theorem 53.100 makes, up to a double covering,

$$AdS_2=G/A=\mathrm{Ad}(G)H. \tag{78.55}$$

The double covering is the fact that $SL(2,\mathbb{R})=\mathrm{SO}(2,2)/\mathbb{Z}_2$, so that $\mathrm{Ad}(G)H$ is a double covering of AdS_2. This is the reason why proposition 78.2 works with $k\in]0,\pi/2[$ instead of $]0,\pi[$.

Since the Killing form is Ad-invariant, G/A is a sphere for the metric B.

In the basis $\{H,E,F\}$ of $SL(2,\mathbb{R})$, the matrix of the Killing form is given by

$$B=\begin{pmatrix}8&&\\&&4\\&4&\end{pmatrix} \tag{78.56}$$

while the basis $\{H,E+F,E-F\}$ gives

$$B=\begin{pmatrix}8&&\\&8&\\&&-8\end{pmatrix},$$

3. See section 78.6 for notations related to $SL(2,\mathbb{R})$.

so that we have the following isometry, $(\mathfrak{sl}(2,\mathbb{R}), B) \sim (\mathbb{R}^3, \eta_{1,2})$. The latter form of the metric allows us to visualize AdS_2 as an hyperboloid in $\mathbb{R}^3$.

We will use the Cartan involution $\theta(X) = -X^t$, so that $\bar{N}$ denotes $\theta(N) = \{e^{fF}\}$.

Remarque 78.6.
The vector field H^* is a Killing because it is the tangent vector of the flow of the one-parameter isometry group $\psi_t \colon AdS_2 \to AdS_2$ given by $\psi_t(x) = \mathrm{Ad}(e^{tH})x$.

Another way to express the singularity is

$$\mathrm{Ad}(e^{nE})(\pm H) \text{ and } \mathrm{Ad}(e^{fF})(\pm H), \tag{78.57}$$

which shows that these are orbits of AN and $A\bar{N}$. Indeed, as $\mathrm{Ad}(a)$ fixes H, we can write $\mathrm{Ad}(an)H = \mathrm{Ad}(ana^{-1})H$. Using the CBH formula we find

$$ana^{-1} = e^{nE+2anE+\dots} = e^{ne^{2a}E} = n' \in N.$$

The same can be done with f. So $\mathrm{Ad}(an)H = \mathrm{Ad}(n')H$ and $\mathrm{Ad}(af)H = \mathrm{Ad}(f')H$. This shows that for all $n \in N$ and $a \in A$, there exists $n' \in N$ such that [eq:singuAdd]

$$\mathrm{Ad}(an)H = \mathrm{Ad}(n')H \tag{78.58a}$$

The same is true with f:

$$\mathrm{Ad}(af)H = \mathrm{Ad}(f')H. \tag{78.58b}$$

In the basis E, F, H the singularity is made up from four lines with angle 45° trough H and $-H$. They divide the space AdS_2 into four pieces. We define the **physical space** as the part of AdS_2 contained between $H+\lambda E$ and $-H+\lambda E$. The K part of $SL(2,\mathbb{R})$ gives a double covering of this curve. The part contained between the parts $H+\lambda F$ and $-H+\lambda F$ of the singularity should be another choice of physical space.

78.4.8 Light cone

The light-like vectors of $\mathfrak{sl}(2,\mathbb{R})$ are E and F, so the light cone of point $\mathrm{Ad}(g)H$ consists in two parts:

$$\mathrm{Ad}(g)\,\mathrm{Ad}(e^{tE})H,$$
$$\mathrm{Ad}(g)\,\mathrm{Ad}(e^{tF})H.$$

It is best rewritten in the compact form

$$C^+_{\mathrm{Ad}(g)H} = \{\mathrm{Ad}(g)\,\mathrm{Ad}(e^{t\epsilon E})H\}_{\substack{t>0\\ \epsilon=\mathrm{Id},\theta}} \tag{78.59}$$

where ϵ is the identity or the Cartan involution.

It is somewhat easy to remark that for all X, Y in a Lie algebra and all automorphism φ, the formula $\varphi(\mathrm{Ad}(e^X)Y) = \mathrm{Ad}(e^{\varphi X})(\varphi Y)$ holds. Then

$$\mathrm{Ad}(e^{t\epsilon E})H = s(\epsilon)\epsilon(\mathrm{Ad}(e^{tE})H) \tag{78.60}$$

with

$$s(\epsilon) = \begin{cases} 1 & \text{if } \epsilon = \mathrm{Id}, \\ -1 & \text{if } \epsilon = \theta. \end{cases}$$

Since $H^*_x = -[H, x]$, the intersection of the light cone with the singularity is expressed, using proposition 78.1, as

$$\|[H, \mathrm{Ad}(g)\,\mathrm{Ad}(e^{t\epsilon E})H]\|^2 = 0. \tag{78.61}$$

78.5 Causal structure of AdS_3

78.5.1 Isometries of AdS_3

We consider the group

$$G := \mathrm{SL}(2,R) \simeq AdS_3 = \{A \in GL(2,\mathbb{R}) : \det A = 1\}$$

on which we put the Killing metric in order to use the fact that $\mathrm{Aut}(G) \subset \mathrm{Iso}(G)$ (equation (52.10)). For notational convenience, we write $\underline{G} := G \times G$.

Lemma 78.7.
The group $\mathrm{SL}(2,R) \times \mathrm{SL}(2,R)$ *is locally isomorph to* $O(2,2) \subseteq \mathrm{Iso}(G)$.

Proof. We consider the action $\psi\colon \underline{G} \times G \to G$ defined by

$$\psi(g_1, g_2, x) = \psi(g_1, g_2)x = g_1 x g_2^{-1}.$$

Since $\psi(g_1, g_2) = L_{g_1} \circ R_{g_2^{-1}}$, the Killing form is invariant under $\psi(g_1, g_2)$ (cf. theorem 52.14). Thus $\psi(g_1, g_2) \in \mathrm{Iso}(G)$. Now we define $\tilde{X}$: the left invariant vector field generated by $X \in \mathcal{G}$ and $\underline{X}$, the right invariant one.

$$dR\underline{X} = \underline{X}, \quad dL\tilde{X} = \tilde{X}.$$

We proof that the differential of ψ is $d\psi(X,Y) = \underline{X} - \tilde{X}$. Consider the path X_t and Y_t in G. One has $\psi(X_t, Y_t, x) = X_t x Y_t^{-1}$. Taking the d/dt of this,

$$\begin{aligned} d\psi_{(e,e,x)}(X,Y,0) &= \frac{d}{dt}\Big[\psi(X_t, Y_t, x)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[X_t x\Big]_{t=0} + \frac{d}{dt}\Big[xY_t^{-1}\Big]_{t=0} \\ &= \underline{X}_x - \tilde{Y}_x. \end{aligned} \tag{78.62}$$

From this to $d\psi_e(X,Y) = \underline{X} - \tilde{Y}$, there is just a notational game.

On the other hand, $\dim(\mathcal{G}\times\mathcal{G}) = 6$ because a basis of $\mathcal{G}$ is for example $\{E, H, F\}$ (cf. page 3221). But the dimension of $\mathfrak{Lie}(O(2,2))$ is also 6. Indeed, if we denotes by (t,u,x,y) the coordinates in the 4-dimensional space on which $O(2,2)$ acts, one has the rotations in the planes (x,y), (x,t), (x,u), (y,t), (y,u), (t,u), and some reversals. But the reversals are not connected to identity (because determinant is -1), then they don't appear in the Lie algebra. The conclusion is [4]

$$\mathcal{G} \times \mathcal{G} \simeq \mathfrak{Lie}(\mathrm{Iso}(G)).$$

□

78.5.2 Orbits in AdS_3

We know that $\underline{G} := \mathrm{SL}(2,R) \times \mathrm{SL}(2,R)$ acts by isometries on $G = \mathrm{SL}(2,R)$. The question is now to see what are the orbits of this actions and are they open? We will use the techniques developed in section 66.6. Let us take some notations: we denote the action by $\tau\colon \underline{G} \times G \to G$, $\tau((g,h),x) = gxh^{-1}$. This is the same as the former ψ. Clearly, $\underline{AN}$ denotes $AN \times AN$ with respect to the decomposition ANK of $\mathrm{SL}(2,R)$ (see 53.17) The orbits of $\underline{AN}$ in G are

$$\tau_{(an,(a'n')^{-1})}(x) = anxa'n'.$$

4. Il faut encore conclure, et ça se fait dans le papier des proceedings de Stephane. Mais de toutes façons pour déformer, tout ce dont nous avons besoin c'est d'avoir une action par isométries de $\underline{G}$. Et ça, on l'a.

78.5.2.1 Two closed orbits

Let us first consider the simplest case: $x \in AN$. Since AN is a group, the orbit $\tau_{\underline{G}}x$ is AN itself.

The second easy case is $x \in \mathcal{Z}(AN)$. In this case, the orbit is ANx. But it is rather clear that AN is closed in G because it is the exponential of $\mathcal{A} \oplus \mathcal{N}$ which is closed in $\mathfrak{sl}(2, \mathbb{R})$, as any subspace of a vector space.

Now, let us find "natural" representatives of these orbits. For the first one, the choice is clear, it is $\mathbb{1} \in AN$. For the second one, we have to find a matrix in G which commutes with the whole G. Equation (53.320) gives the explicit matrices of $\mathrm{SL}(2, R)$. It is not hard to see that only $\pm\mathbb{1}$ works:

$$\mathcal{Z}(G) = \{\pm\mathbb{1}\}.$$

Finally, our two first orbits are $\pm AN$. These are closed in G.

78.5.2.2 Two open orbits

Consider a $k \in K$. We well see at page 3631 that $\tau_{\underline{G}}k$ is a submanifold of G, thus it will be open if and only if it has dimension three. If $\underline{X} \in (\mathcal{A}\oplus\mathcal{N}) \times (\mathcal{A}\oplus\mathcal{N})$, we define its **fundamental vector field** $\underline{X}^*_x$ by

$$\underline{X}^*_x = \frac{d}{dt}\Big[\tau_{e^{t\underline{X}}}(x)\Big]_{t=0}$$

for any $x \in G$, and we take the notation

$$\mathcal{O} = \tau_{\underline{G}}(k).$$

On the other hand, we know that the tangent space of an orbit for a group action is spanned by the fundamentals fields for this action. Hence, the action is generated by the left and right actions, and the fundamentals fields are the corresponding invariant fields:

$$T_k(\mathcal{O}) = \mathrm{Span}\{\underline{H}_k, \tilde{H}_k, \underline{E}_k, \tilde{E}_k\}, \tag{78.63}$$

where we recall that $\mathcal{G} = \mathrm{SL}(2, R)$ and $\mathcal{A} = \mathbb{R}H$, $\mathcal{N} = \mathbb{R}E$.

Remarque 78.8.
Be careful on this fact: in the expression "$T_k(\mathcal{O})$", the k appears two times: the first is the index k to T and the second is in the definition of $\mathcal{O}$. We could work with any $T_s(\mathcal{O})$, but k is the only element of G which *certainly* is in $\mathcal{O}$. So we will work with $T_k(\mathcal{O})$.

As last notation, we define $X^\flat(Y) = B(X, Y)$. We consider the 3-form

$$\nu = \tilde{E}^\flat \wedge \tilde{H}^\flat \wedge \tilde{F}^\flat.$$

Since E, H and F are linearly independent, this form is nondegenerate. Thus $\nu_k(A, B, C) = 0$ if and only if A, B and C are linearly independent in T_kG. Let us compute:

$$\nu_k(\underline{H}_k, \tilde{H}_k, \tilde{E}_k) = \nu_e(dL_{k^{-1}}\underline{H}_k, H, E)$$

because $\tilde{H}$ and $\tilde{E}$ are left invariants. But $\underline{H}$ is right invariant, so that $dL_{k^{-1}}\underline{H} = \mathrm{Ad}(k)H$. Finally,

$$\nu_k(\underline{H}, \tilde{H}, \tilde{E}) = \nu_e(\mathrm{Ad}(k)H, H, E). \tag{78.64}$$

As we know that [5] $\mathrm{Ad}(k)H = \alpha H + \beta(E + F)$ (for $k \in K$), we begin to restrict, *with* loss of generality, ourself to a k_0 such that $\mathrm{Ad}(k_0)H \neq \alpha H$. In this case,

$$\nu_{k_0}(\underline{H}, tH, tE) = \nu_e(\beta E, H, F) \neq 0. \tag{78.65}$$

The conclusion is that the orbit of k_0 is open because it is three-dimensional.

If on the other hand we chose k_1 such that $\mathrm{Ad}(k_1)H = -H$,

$$\nu_{k_0}(\tilde{E}, \underline{E}, \tilde{H}) = \nu_e(E, \mathrm{Ad}(k_1)E, H) \neq 0.$$

Then, such a k_1 gives an other open orbit.

5. In view of (53.320), it is possible to determine the exact form of α and β; it will be done soon.

78.5.2.3 Do we have all the orbits?

Loosely said, we will show that the orbit of k_0 contains at least the half of K, and then the half of $G = ANK$. It will then be clear that the orbit of k_1 takes the remaining part.

An easy computation shows that if $k \in K$ is as in equation (53.320),

$$\mathrm{Ad}(k)H = \alpha H + \beta(E+F)$$

with $\alpha = \cos^2 t - \sin^2 t$ and $\beta = -2\sin t \cos t$. On the one hand, the condition $\mathrm{Ad}(k_0)H \neq \alpha H$ is fulfilled for example by

$$k_0 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

On the other hand, a general element of AN can be written as $\begin{pmatrix} e^a & le^a \\ 0 & e^{-a} \end{pmatrix}$, so that a general element of ANk_0AN is

$$\frac{\sqrt{2}}{2}\begin{pmatrix} e^{a+b}(l+1) & me^{a+b}(l+1) + e^{a-b}(l-1) \\ e^{b-a} & me^{b-a} + e^{-(a+b)} \end{pmatrix}.$$

With a suitable choice of a and b, one can give to e^{a+b} and e^{a-b} any positive value. Thus for any m, $l \in \mathbb{R}$ and $x, y > 0$, one can find

$$\frac{\sqrt{2}}{2}\begin{pmatrix} x(l+1) & mx(l+1) + y^{-1}(l-1) \\ y & my + x^{-1} \end{pmatrix} \tag{78.66}$$

as general element of $\tau_{\underline{AN}}(k_0)$. Taking

$$\begin{aligned} y &= -\frac{2}{\sqrt{2}}\sin t \\ x &= \frac{\cos t}{\sqrt{2}} \\ m &= \tan t \\ l &= 1, \end{aligned} \tag{78.67}$$

one finds the matrix $\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$ for any $\sin t > 0$. In this sense, we say that one has at least *the half* of K in $\tau_{\underline{G}}(k_0)$. Then, by re-acting with AN, one finds the half of $G = ANK$.

78.5.3 The metric

We compute the metric in the basis $\{\partial_\tau, \partial_u, \partial_\phi\}$ at the point $z = e^{\phi H}e^{eE}e^{\tau T}e^{\alpha\phi H}$:

$$\begin{aligned} dL_{z^{-1}}\partial_\tau &= \frac{d}{dt}\Big[e^{-\alpha\phi H}e^{-\tau T}e^{-uE}e^{-\phi H}e^{\phi H}e^{uE}e^{(\tau+t)T}e^{\alpha\phi H}\Big]_{t=0} \\ &= \mathrm{Ad}(e^{-\alpha\phi H})T. \end{aligned} \tag{78.68}$$

In the same way,

$$\begin{aligned} dL_{z^{-1}}\partial_u &= \mathrm{Ad}(e^{-\alpha\phi H}e^{-\tau T})E \\ dL_{z^{-1}}\partial_\phi &= \mathrm{Ad}(z^{-1})H + \alpha H.s \end{aligned} \tag{78.69}$$

In the computation of $dL_{z^{-1}}\partial_\phi$, we used the Leibniz rule. The rest is a rather mechanical computation using the Ad-invariance of B and the fact that, by definition if B_z, $B_z(X,Y) = B_e(dL_{z^{-1}}X, dL_{z^{-1}}Y)$. The following relations will be useful:

$$\mathrm{Ad}(e^{uE})T = T - nH + n^2E \tag{78.70a}$$

$$\mathrm{Ad}(e^{uE})H = H + 2nE \tag{78.70b}$$

$$\mathrm{Ad}(e^{-\tau T})E = -\sin\tau\cos\tau H + \cos^2\tau E - \sin^2\tau F \tag{78.70c}$$

$$
\begin{aligned}
B(\partial_\tau, \partial_\tau) &= B(T,T) = -8 && (78.71)\\
B(\partial_\tau, \partial_u) &= B(T, \mathrm{Ad}(e^{-\tau T})E) \\
&= B(\mathrm{Ad}(e^{\tau T})T, E) && (78.72)\\
&= B(E-F, E) = 4, \\
B(\partial_\tau, \partial_\phi) &= B(\mathrm{Ad}(e^{-\alpha\phi H})T, \alpha H) \\
&\quad + B(T, \mathrm{Ad}(e^{-\tau T}e^{-uE}e^{-\phi H})H) \\
&= \underbrace{B(T, \alpha H)}_{=0} + B(\mathrm{Ad}(e^{uE})T, \mathrm{Ad}(e^{-\phi H})H) && (78.73)\\
&= B(\mathrm{Ad}(e^{uE})T, H) \\
&= -4n, \\
B(\partial_u, \partial_u) &= B(E,E) = 0, && (78.74)\\
B(\partial_u, \partial_\phi) &= B(E, \mathrm{Ad}(e^{-uE}e^{-\phi H})H) \\
&\quad + B(\mathrm{Ad}(e^{-\tau T})E, \alpha H) \\
&= B(-\sin\tau\cos\tau H + \cos^2\tau E - \sin^2\tau F, H) && (78.75)\\
&= -2\alpha\cos(2\tau).
\end{aligned}
$$

78.6 Identifications in AdS_3

secBTZ

In this section, we recall for the reader convenience the definition and construction of the non-rotating BTZ black hole [1014, 1015], emphasizing on some geometrical properties put forwards in [1012, 1016, 1017, 1018]. To lighten the presentation, the proofs will essentially be omitted and referred to the existing literature.

This situation will serve us as a guideline in defining black holes in general anti de Sitter spaces (see section 79.8).

Bañados, Henneaux, Teitelboim and Zanelli observed that taking the quotient of (a part of) the three-dimensional anti de Sitter space (AdS_3) under the action of well-chosen discrete subgroups of its isometry group gives rise to solutions which correspond to axially symmetric and static black hole solutions of (2+1)-dimensional Einstein gravity with negative cosmological constant, characterized by their mass M and angular momentum J.

The space AdS_3 is defined as the (universal covering of the) simple Lie group $\mathrm{SL}(2,R)$

$$AdS_3 \cong \mathrm{SL}(2,R) = \{g \in GL(2,\mathbb{R}) \,|\, \det g = 1\} := G \qquad (78.76)$$

AdS3

endowed with its Killing metric $B\colon \mathcal{G}\times\mathcal{G} \to \mathbb{R}$, which can be extended to the whole group by

$$B_g(X,Y) = B(dL_{g^{-1}}X, dL_{g^{-1}}Y). \qquad (78.77)$$

Here, $\mathcal{G}$ stands for the Lie algebra of $\mathrm{SL}(2,R)$. See section 53.17 for a review of $\mathrm{SL}(2,\mathbb{R})$ and its Lie algebra. We define the following one-parameter subgroups of $\mathrm{SL}(2,R)$:

$$A = \exp(\mathbb{R}H), \quad N = \exp(\mathbb{R}E), \quad \bar{N} = \exp(\mathbb{R}F), \quad K = \exp(\mathbb{R}T), \qquad (78.78)$$

IwasawaSubgroups

with $T = E - F$. They are the building blocks of the Iwasawa decomposition

$$K\times A\times N \longrightarrow \mathrm{SL}(2,R) : (k,a,n) \longrightarrow kan \quad \text{or} \quad ank. \qquad (78.79)$$

The $\mathrm{SL}(2,R)$ subgroups AN and $A\bar{N}$ are called *Iwasawa subgroups*; they are minimal parabolic subgroups.

We have now dispose of all necessary ingredients to make the definition of BTZ black holes more precise.

Definition 78.9.
The one-parameter subgroup of $Iso(G)$ defined by

$$\psi_t(g) = \exp(t\,a\,H)\,g\exp(-t\,a\,H), \quad a \in \mathbb{R}_0, \quad g \in G \tag{78.80}$$

*is called the **BHTZ subgroup**. Its generator $\Xi = a(H, H)$ is called the **identification vector**. The **BHTZ action** associated with Ξ is $\psi_{\mathbb{Z}} : G \to G$. It is an action of $\mathbb{Z}$ on G.* defBHTZ

Definition 78.10.
*A **safe region** in AdS_3 is defined as an open connected domain in which*

$$\|\Xi\|^2 := \beta_z(\Xi, \Xi) > 0 \quad . \tag{78.81}$$

safe

Definition 78.11.
*A **non-rotating massive BTZ black hole** is defined as the quotient of a safe region in AdS_3 under the BHTZ action.* NRMassive

This definition deserves some comments. First, the restriction to a safe region in AdS_3 ensures that the resulting quotient space be free of closed time-like curves. This means that the other parts of AdS_3 have to be "cut out" from the original space. Furthermore, due to the identifications, one may restrict to a fundamental domain of the BHTZ action. Secondly, the **black hole singularity** $\mathscr{S}$ is defined as the surfaces where the identification vector becomes light-like:

$$\mathscr{S} = \{z \in AdS_3 \text{ tel que } \beta_z(\Xi, \Xi) = 0\}. \tag{78.82}$$

Singularities

Thus, the BTZ singularity represents singularities in the causal structure, not curvature ones. The resulting space is causally inextensible, i.e. trying to extend it would produce closed time-like curves. Finally, the BTZ space-time exhibits all characteristic features of a black hole. Namely, it has *event horizons*, that is, surfaces hiding a region (*the interior region*, see hereafter) causally disconnected from spatial infinity.

Note that it is the choice of identification vector which dictates the nature (rotating, extremal, vacuum or non-rotating massive) of the resulting black hole. Moreover, not all choices give rise to black holes.

The reason why we here focus on the non-rotating massive case lies in the peculiar geometrical properties of its horizons and singularity.

We are now ready to define the horizons. The following definition is adapted to the present case, but cannot be used in general, see 79.8 for more general definitions:

Definition 78.12.
*A point g will be said to lie in the **future interior region**, denoted by $\mathcal{M}^{int,+}$, if all future-directed light rays issued from g necessarily fall into the black hole singularity, that is*

$$g \in \mathcal{M}^{int,+} \Leftrightarrow \forall k \in K, \exists s \in \mathbb{R}^+ s.t. \|\underline{H} - \overline{H}\|^2_{l^k_g(s)} = 0. \tag{78.83}$$

interiorreg

*The **future horizon** $\mathscr{H}^+$ is defined as the boundary of $\mathcal{M}^{int,+}$.* interior-horizons

Equation (78.83) simply expresses that any future-directed causal signal necessarily falls into the black hole singularity and cannot escape it. The **past interior region** and **past horizon** are defined in a similar way.

Using the embedding (78.3) of AdS_3 into $\mathbb{R}^{2,2}$, one finds, from (78.82) and (78.83), that

$$\begin{aligned} \mathscr{S} &\equiv t^2 - y^2 = 0, \\ \mathscr{H} &\equiv u^2 - x^2 = 0, \end{aligned} \tag{78.84}$$

BTZSingHor

where $\mathscr{H} = \mathscr{H}^+ \cup \mathscr{H}^-$.

As pointed out in [1017], these results can be stated more intrinsically as follows:

Proposition 78.13.
In $G = AdS_3$, the non-rotating BTZ black hole singularity is given by a union of minimal parabolic subgroups of G:

$$\mathscr{S} = Z(G)AN \cup Z(G)A\overline{N}, \tag{78.85}$$

where $Z(G) = \{e, -e\}$ denotes the center of $G = \mathrm{SL}(2, R)$. BTZSing

PropLatClassANSLdeuxR

Proposition 78.14.
In $G = AdS_3$, the non-rotating BTZ black hole horizon corresponds to a union of lateral classes of minimal parabolic subgroups of G:

$$\mathscr{H} = Z(G)ANJ \cup Z(G)A\overline{N}J \quad , \tag{78.86}$$

EqHorClassLatdeux

where $J = \exp(\frac{3\pi}{2}T) \in K$ satisfies $J^2 = e$. BTZHor

These two propositions follow directly from (78.84), using the parametrization (78.3). Note that the small difference between the formula given here and the one in the paper [1017] is simply due to a different choice of parametrization (78.3).

They show that the black hole structure is closely related to the minimal parabolic subgroup of $\mathrm{SL}(2, R)$. Of course, this construction cannot be generalized in a straightforward way to higher-dimensional anti de Sitter spaces, because of the peculiar nature of the three-dimensional case, being the only to enjoy a group manifold structure. Rather, we will reconsider in the next section the case treated here in a more general framework, putting on an equal footing all anti de Sitter spaces. Again, a minimal parabolic subgroup will reveal crucial in the construction.

78.6.1 Global description of the black hole

In this appendix, we use results and techniques of [1012, 1013, 1017, 1018] to derive the equation of the non-rotating BTZ black holes horizons. We will begin by stating some results which will be useful in describing the global geometry of the black hole.

Proposition 78.15.
Let σ be the unique exterior automorphism of G fixing pointwise the Cartan subgroup A and consider the following twisted *action of G on itself:*

$$\tau : G \times G \longrightarrow G : (g, x) \to \tau_g(x) := g\, x\, \sigma(g^{-1}). \tag{78.87}$$

Then, the BHTZ action can be rewritten as

$$\psi_n = \tau_{\exp(n\sqrt{M}H)}, \quad n \in \mathbb{Z}. \tag{78.88}$$

The proof follows from the fact that σ fixes the generator H. Using the action τ, one finds the following global decomposition of G:

Proposition 78.16.
The map

$$\begin{aligned} \phi\colon A \times G/A &\to G \\ (a, [g]) &\mapsto \tau_g(a) \end{aligned} \tag{78.89}$$

is well-defined as a global diffeomorphism.

This follows from the observation that the application

$$\phi : K \times A \times N \to G : (k, a, n) \to \phi(k, a, n) = \tau_{kn}(a) \tag{78.90}$$

twistedI

is a global diffeomorphism on G ("twisted Iwasawa decomposition"). As a consequence, the space G appears as the total space of a trivial fibration over $A = \mathrm{SO}(1, 1) \simeq \mathbb{R}$ whose fibers are the

τ_G-orbits, i.e. the $\sigma-$twisted conjugacy classes. As a homogeneous $G-$space, every fiber is isomorphic to $G/A = AdS_2$. Moreover, the BHTZ action is fiberwise, because

$$\tau_h(\phi(a,[g])) = \phi(a, h.[g]) = \phi(a,[hg]). \tag{78.91}$$ fiberwise

The Killing metric on G turns out to be globally diagonal with respect to the twisted Iwasawa decomposition [1012]:

$$ds^2_G = da^2_A - \frac{1}{4}\cosh^2(a) ds^2_{G/A}, \tag{78.92}$$

where $ds^2_{G/A}$ denotes the canonical projected AdS_2-metric on G/A. The study of the quotient space $G/\mathbb{Z}$ therefore reduces to the study of $(G/A)/\mathbb{Z}$.

The space G/A can be realized as the $G-$equivariant universal covering space of the adjoint orbit $\mathcal{O} := \mathrm{Ad}(G)H$ in $\mathfrak{sl}(2,\mathbb{R})$, where it corresponds to a one sheet hyperboloid. In this picture, we may identify the part of the hyperboloid corresponding to a safe region (see definition 78.10) in G.

Lemma 78.17.
In $\mathcal{O}$, a connected region where the orbits of the BHTZ action are space-like is given by

$$\{X = x^H H + x^E E + x^F F \in \mathcal{O} \mid \ -1 < x^H < 1\}. \tag{78.93}$$

Furthermore, it can be parameterized as

$$X = Ad\left(\exp(\frac{\theta}{2}H)\,\exp(-\frac{\tau}{2}T)\right) H,\ 0 < \tau < \pi\,,\ -\infty < \theta < +\infty. \tag{78.94}$$

This has been proven in [1012]. From this and the preceding proposition, we find a global description of a safe region in G well adapted to the BHTZ identifications.

Proposition 78.18.
A global description of a safe region in G is given by

$$z(\rho,\theta,\tau) = \tau_{\exp(\frac{\theta}{2}H)\,\exp(-\frac{\tau}{2}T)}(\exp(\rho H)). \tag{78.95}$$ CoordGlob

Furthermore, the action of the BHTZ subgroup reads in these coordinates

$$(\tau,\rho,\theta) \to (\tau,\rho,\theta + 2na). \tag{78.96}$$

78.6.2 Derivation of the horizons

Now we have to study the equation of (78.83). Using the bi-invariance of the Killing metric and the Ad-invariance of the Killing form, it reduces to

$$B(H,H) - B(H, Ad(\mathrm{e}^{-sAd(k)E})Ad(x)H) = 0. \tag{78.97}$$ EqHoriz

Lemma 78.19.
$\mathcal{M}^{int,+}$ is A bi-invariant. BiInv

Proof. This equation is clearly invariant under $x \to x \cdot a$ for each $a \in A$. In order to see the invariance under $x \to a \cdot x$, one uses the cyclicity of the trace to bring the second term to

$$B(H, Ad(Ad(a^{-1})\mathrm{e}^{uAd(k)E})Ad(x)H).$$

But $Ad(a^{-1})\mathrm{e}^{-sAd(k)E} = \mathrm{e}^{-\tilde{s}Ad(\tilde{k})E}$, with $\tilde{s} = s(\mathrm{e}^{-2a}\cos^2\theta + \mathrm{e}^{2a}\sin^2\theta)$ and $\cot t = \mathrm{e}^{-2a}\cot\theta$, where $k = \mathrm{e}^{\theta T}$ and $\tilde{k} = \mathrm{e}^{tT}$. The net result is thus simply a relabelling of the parameters (note that s and $\tilde{s}$ have the same signs!) □

Let us now consider a light ray starting from a safe region in G. Because of (78.95) and lemma 78.19, we may restrict our study to

$$z = e^{-\tau/2T} e^{\rho H} \sigma(e^{\tau/2T}) \quad (78.98)$$
$$= e^{-\tau/2T} e^{\rho H} e^{-\tau/2T}. \quad (78.99)$$

The equation to study reduces to

$$B(H,H) - B(H, Ad(e^{-sAd(k)E}) Ad(e^{-\tau/2T} e^{\rho H} e^{-\tau/2T}) H) = 0 \quad (78.100)$$

with $\tau \in]0,\pi[$ and $\rho \in \mathbb{R}$.

Let us focus on the points in $Ad(G)H$ corresponding to

$$\mathcal{B} := Ad(e^{-\tau/2T} e^{\rho H} e^{-\tau/2T}) H,$$

with $\tau \in]0,\pi[$, $\rho \in \mathbb{R}$. First note that $Ad(e^{\rho H} e^{-\tau/2H})H$ precisely corresponds to a safe region on the hyperboloid. Thus $\mathcal{B}$ is the region swept out by the a safe region when rotating it counter-clockwise around the T-axis with an angle π.

It can be seen that the domain $\mathcal{B}$ can be decomposed into three regions:

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3, \quad (78.101)$$

with

$$\mathcal{B}_1 = Ad(A) Ad(e^{-\beta/2T}) H \qquad \beta \in]0, 2\pi[, \quad (78.102a)$$
$$\mathcal{B}_2 = Ad(A) Ad(e^{t(E+F)}) H \qquad t \in \mathbb{R}, \quad (78.102b)$$
$$\mathcal{B}_3 = Ad(A)(-H \pm E) \text{ or } Ad(A)(-H \pm F). \quad (78.102c)$$

Thanks to the A bi-invariance, we may forget about the $Ad(A)$ in the above equations. We are thus led to analyze the existence of solutions of (78.100) with $X \in \mathcal{B}$ of the form $X_1 = Ad(e^{-\beta/2T})H$, $X_2 = Ad(e^{t(E+F)})H$ and $X_3 = -H \pm E$, $-H \pm F$.

Consider the first case. With $Ad(e^{-\tau/2T} e^{\rho H} e^{-\tau/2T})H$ of the form $Ad(e^{-\beta/2T})H$, equation (78.100) becomes

$$\frac{1}{4} s^2 (\cos\beta - \cos(\beta + 4\theta)) + s \sin\beta + 2\sin^2\beta = 0. \quad (78.103)$$

We are looking for the values of β for which this equation admits a solution for $s > 0$, for all $\theta \in [0,\pi]$ –this range for θ originates from the fact that G/A is a $\mathbb{Z}_2$ covering of $\mathrm{Ad}(G)H$. By considering the particular case $\theta = 0$, we find $s = -\tan\frac{\beta}{2}$, thus the allowed values of β have to lie in the range $]\pi, 2\pi[$. Let us look at the constrains imposed by other values of θ. If we denote by s_1 and s_2 the two roots of (78.103), we have

$$s_1 \cdot s_2 = \frac{4\sin^2\beta/2}{\sin 2\theta \sin(\beta + 2\theta)}, \quad (78.104)$$
$$s_1 + s_2 = \frac{-2\sin\beta}{\sin 2\theta \sin(\beta + 2\theta)}. \quad (78.105)$$

First note that, $\forall \beta \in]0, 2\pi[$, the quantity $\sin 2\theta \sin(\beta + 2\theta)$ may be positive or negative as θ varies in the range $[0,\pi]$. If $\sin\beta < 0$, then there are two positive roots when $\sin 2\theta \sin(\beta + 2\theta) > 0$, and one positive and one negative when $\sin 2\theta \sin(\beta + 2\theta) < 0$. Thus there always exist a positive solution for u, for any θ. If $\sin\beta > 0$, there are two negative roots when $\sin 2\theta \sin(\beta + 2\theta) > 0$. Consequently, the interior region will correspond to points $X_1 = Ad(e^{-\beta/2T})H$, $\beta \in]\pi, 2\pi[$ on the adjoint orbit.

¡¡ Avertissement/question à la lectrice !! 78.20
On parle d'une solution pour tout u. Quel u ?

For the second case, $X_2 = Ad(\mathrm{e}^{t(E+F)})H$, the equation we get is

$$\frac{1}{4}s^2(\cosh 2t - \cos 4\theta \cosh 2t + 2\sin 2\theta \sinh 2t) + s\cos 2\theta \sinh 2t + (1 - \cosh 2t) = 0. \tag{78.106}$$

By considering two special cases, it is easy to see that this equation does not admit a positive solution in u for all θ. Indeed, for $\theta = \pi/2$, one finds $s = -\tanh t$, while for $\theta = 0$, one gets $s = \tanh t$. Thus there is no $t \neq 0$ satisfying both conditions. The last case yields no positive solution for all θ neither.

As a conclusion we find that the interior region is given by

$$x \in \mathcal{M}^{\mathrm{int},+} \Leftrightarrow Ad(x)H = Ad(A)Ad(\mathrm{e}^{-\beta/2T})H, \quad \text{with } \beta \in]\pi, 2\pi[. \tag{78.107}$$

The boundaries of the corresponding region in $Ad(G)H$ are given by $-H + r^2E$ and $-H + r^2F$ or

$$Ad(N^-)(-H) \cup Ad(\bar{N}^+)(-H), \tag{78.108}$$

with $N^- = \{\mathrm{e}^{tE}\}_{t\leqslant 0}$ and $\bar{N}^+ = \{\mathrm{e}^{tF}\}_{t\geqslant 0}$.

The horizons can be deduced as

$$x \in \mathcal{H}^+ \Leftrightarrow Ad(x)H = Ad(N^-)(-H) \text{ or } Ad(x)H = Ad(\bar{N}^+)(-H). \tag{78.109}$$

Because of the A-invariance, we may write $x = \tau_{\mathrm{e}^{-\frac{\tau}{2}T}}(\mathrm{e}^{\rho H})$ and look for the relation between τ and ρ such that

$$Ad(\tau_{\mathrm{e}^{-\tau/2T}}(\mathrm{e}^{\rho H}))H = Ad(N^-)Ad(\mathrm{e}^{\pi/2T})H. \tag{78.110}$$

This amounts to require that

$$\left(\mathrm{e}^{-\tau/2T}\mathrm{e}^{\rho H}\mathrm{e}^{-\tau/2T}\right)^{-1} (\mathrm{e}^{-t^2E}\mathrm{e}^{\pi/2T}) \in A \cup Z(G). \tag{78.111}$$

CondHor

This condition gives $\cos\tau = \tanh\rho$, $\rho < 0$, $\tau \in]\pi/2, \pi[$. By replacing e^{-t^2E} with e^{t^2F}, one gets $\cos\tau = -\tanh\rho$, $\rho > 0$, $\tau \in]\pi/2, \pi[$.

The domain $\mathcal{M}^{\mathrm{int},-}$ is of course defined as

$$x \in \mathcal{M}^{\mathrm{int},-} \Leftrightarrow \forall k \in K, \exists u \in \mathbb{R}^- \text{s.t.} \|\underline{H} - \overline{H}\|^2_{l^k_x(u)} = 0. \tag{78.112}$$

interior

The *past horizon* $\mathcal{H}^-$ is defined as the boundary of $\mathcal{M}^{\mathrm{int},-}$. By proceeding the same way, we find that

$$x \in \mathcal{M}^{\mathrm{int},-} \Leftrightarrow Ad(x)H = Ad(A)Ad(\mathrm{e}^{-\beta/2T})H, \quad \text{with } \beta \in]\pi, 2\pi[, \tag{78.113}$$

and

$$x \in \mathcal{H}^- \Leftrightarrow Ad(x)H = Ad(N^+)(-H) \text{ or } Ad(x)H = Ad(\bar{N}^-)(-H), \tag{78.114}$$

or in coordinates: $\tau \in]0, \pi/2[$, $\cos\tau = \tanh\rho$ for $\rho > 0$ and $\cos\tau = -\tanh\rho$ for $\rho < 0$.

We thus established the following

Proposition 78.21.
In a safe region in G parameterized by

$$z(\rho, \theta, \tau) = \tau_{\exp(\frac{\theta}{2}H)\,\exp(-\frac{\tau}{2}T)}(\exp(\rho H)),$$

the horizons $\mathcal{H} := \mathcal{H}^+ \cup \mathcal{H}^-$ of the non-rotating BTZ black hole are given by

$$\cos\tau = \pm\tanh\rho. \tag{78.115}$$

As a direct consequence, we have the

CorHorClassLat

Corollary 78.22.
In terms of the embedding coordinates (78.3) of G in $\mathbb{R}^{2,2}$, the horizons of the non-rotating BTZ black hole are

$$\mathcal{H} \equiv u^2 - x^2 = 0. \tag{78.116}$$

78.6.3 General considerations

78.6.3.1 Construction of the black hole

Let Ξ be a Killing vector field on $G = \mathrm{SL}(2,R)$, and $\psi_t\colon \mathrm{SL}(2,R) \to \mathrm{SL}(2,R)$ be its flow, i.e

$$\frac{d}{dt}\Big[\psi_t(g)\Big]_{t=s} = \Xi_{\psi_s(g)}$$

for all $g \in \mathrm{SL}(2,R)$. We will discuss later why we need Ξ to be a Killing vector field, and the reason why we will choice [6]

$$\Xi = \underline{H} + \alpha\overline{H}. \tag{78.117}$$

eq:def_Xi

Theorem 78.23.
There is a diffeomorphism $A \times N \times K \to \mathrm{SL}(2,R)$ given by

$$(a,n,k) \to anka^{\alpha}$$

if $|\alpha| < 1$.

Proof. No proof. □

This theorem suggests us a coordinate system (θ,n,τ) given by $a = e^{\theta H}$, $n = e^{nE}$, $k = e^{\tau T}$ where $T + E - F$:

$$g(\theta,\tau,u) = e^{\theta H}e^{uE}e^{\tau T}e^{-\theta\alpha H}.$$

We consider the space $BTZ = AdS_3/\psi_n = AdS_3/\sim$, where $\sim$ is the equivalence relation on $\mathrm{SL}(2,R)$ given by $g \sim g'$ if and only if there exists a $n \in \mathbb{Z}$ such that $\psi_n(g) = g'$.

If $\psi_t(g)$ is a time-like curve, then the identification creates some closed time-like curves; so the region of AdS_3 where $\|\Xi\|^2 < 0$ must be banned of our theory. More precisely, we put

$$\mathscr{S} = \{z \in \mathrm{SL}(2,R) \text{ tel que } \|\Xi_z\|^2 = 0\},$$

and we cut the space by $\mathscr{S}$: $AdS_3' = \mathrm{SL}(2,R)\backslash\{z \in \mathrm{SL}(2,R) \text{ tel que } \|\Xi_z\|^2 \leqslant 0\}$. So it is valuable to choice Ξ in order to be not everywhere time-like or everywhere space-like. The choice (78.117) corresponds to

$$\psi_t(g) = e^{tH}ge^{-\alpha tH}, \tag{78.118}$$

and it is easy to see that $\psi_t g(\theta,\tau,u) = g(\theta + t,\tau,u)$. In these coordinates, $\Xi = \partial_\theta$.

78.6.3.2 Light rays and causal structure

Let us point out that a geodesic $exp_x(t\Xi)$ has a constant velocity (definition of a geodesic); then the norm of Ξ along an identification curve is constant, figure 78.1. The part of space in which this is positive is the physical part and the part in which it is negative is the singular part.

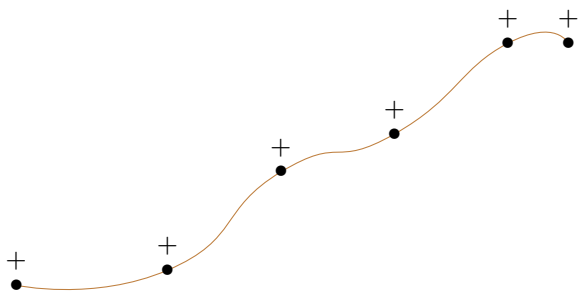


Figure 78.1: The norm (in particular its sing) of Ξ is contant along a geodesic.

LabelFigNEtAchr

6. Here, $\underline{H}$ is the left invariant vector field of H and $\overline{H}$ the right one: $\underline{H}_z = (dL_z)_e H$ and $\overline{H}_z = (dR_z)_e H$, and α is a real constant related to some underlying physics.

A light ray in AdS_3 passing through the point g is given by curves of the form

$$g(s) = ge^{s\,\mathrm{Ad}(e^{\kappa T})E}. \tag{78.119}$$

The light ray passing through z is the direction κ is written $\ell_g^\kappa(s)$. We say that g_0 is in the **black hole** if for any κ, there exists a $s < \infty$ such that $\ell_{g_0}^\kappa(s) \in \mathscr{S}$. The **horizon** is the boundary of the black hole. In order to describe it, we have to solve

$$B_{\ell_{g_0}^\kappa(s)}(\Xi, \Xi) = 0. \tag{78.120}$$

78.6.4 The metric

We compute the metric in the basis $\{\partial_\tau, \partial_u, \partial_\phi\}$ at the point $z = e^{\phi H}e^{eE}e^{\tau T}e^{\alpha\phi H}$:

$$\begin{aligned} dL_{z^{-1}}\partial_\tau &= \frac{d}{dt}\Big[e^{-\alpha\phi H}e^{-\tau T}e^{-uE}e^{-\phi H}e^{\phi H}e^{uE}e^{(\tau+t)T}e^{\alpha\phi H}\Big]_{t=0} \\ &= \mathrm{Ad}(e^{-\alpha\phi H})T. \end{aligned} \tag{78.121a}$$

In the same way,

$$dL_{z^{-1}}\partial_u = \mathrm{Ad}(e^{-\alpha\phi H}e^{-\tau T})E \tag{78.121b}$$

$$dL_{z^{-1}}\partial_\phi = \mathrm{Ad}(z^{-1})H + \alpha H.s \tag{78.121c}$$

In the computation of $dL_{z^{-1}}\partial_\phi$, we used the Leibniz rule. The rest is a rather mechanical computation using the Ad-invariance of B and the fact that, by definition if B_z, $B_z(X,Y) = B_e(dL_{z^{-1}}X, dL_{z^{-1}}Y)$. The following relations will be useful:

$$\mathrm{Ad}(e^{uE})T = T - nH + n^2E \tag{78.122a}$$

$$\mathrm{Ad}(e^{uE})H = H + 2nE \tag{78.122b}$$

$$\mathrm{Ad}(e^{-\tau T})E = -\sin\tau\cos\tau H + \cos^2\tau E - \sin^2\tau F \tag{78.122c}$$

$$B(\partial_\tau, \partial_\tau) = B(T,T) = -8 \tag{78.123}$$

$$\begin{aligned} B(\partial_\tau, \partial_u) &= B(T, \mathrm{Ad}(e^{-\tau T})E) \\ &= B(\mathrm{Ad}(e^{\tau T})T, E) \\ &= B(E - F, E) = 4, \end{aligned} \tag{78.124}$$

$$\begin{aligned} B(\partial_\tau, \partial_\phi) &= B(\mathrm{Ad}(e^{-\alpha\phi H})T, \alpha H) \\ &\quad + B(T, \mathrm{Ad}(e^{-\tau T}e^{-uE}e^{-\phi H})H) \\ &= \underbrace{B(T, \alpha H)}_{=0} + B(\mathrm{Ad}(e^{uE})T, \mathrm{Ad}(e^{-\phi H})H) \\ &= B(\mathrm{Ad}(e^{uE})T, H) \\ &= -4n, \end{aligned} \tag{78.125}$$

$$B(\partial_u, \partial_u) = B(E,E) = 0, \tag{78.126}$$

$$\begin{aligned} B(\partial_u, \partial_\phi) &= B(E, \mathrm{Ad}(e^{-uE}e^{-\phi H})H) \\ &\quad + B(\mathrm{Ad}(e^{-\tau T})E, \alpha H) \\ &= B(-\sin\tau\cos\tau H + \cos^2\tau E - \sin^2\tau F, H) \\ &= -2\alpha\cos(2\tau). \end{aligned} \tag{78.127}$$

Chapter 79

BTZ black holes in anti de Sitter spaces

ChapBHinAdS

79.1 Introduction

The question arises when one watches movies such as Star Trek: what is a black hole? One knows from special relativity that light speed cannot be exceeded. So, as a first attempt to define the notion of black hole, we just say that it is a region of the space from which even light cannot escape. Such an object causes a scientific problem because it is by assumption not observable. This fact allows science-fiction writers to invent whatever with no risk of contradiction. That is a Star Trek black hole.

Physical black holes are much more interesting because they are the signal of a general relativity failure.

The Newtonian gravitational field is given by a potential which increases as $1/r$ when you get closer to a massive object. At $r = 0$, this potential makes no sense and physics is in trouble. One can avoid the problem by postulating that there exist no pointwise masses and that particles cannot penetrate each other. From these assumptions, the fact that Newtonian mechanics does not impose any limit speed makes the divergence at $r = 0$ unimportant.

In general relativity, the divergence at small distances is much more problematic because there is a limit speed; hence a pointwise mass always creates a whole region from which nothing (not even light) can escape. Worse: even a homogeneous ball produces a divergence in the metric when it is too dense, and such objects may exist in the real world. Stated in a more mathematical way: solutions of Einstein's equations for the real world may be singular. From this point of view, black holes are nothing else than a feature in the mathematical framework of relativity which indicates that this is not a final theory. That is the notion of black hole in general relativity and in cosmology.

The transfer of concept from physics to mathematics always consists in taking the key features of the mathematical framework of a physical theory and posing them as definition of a new mathematical object. What are the main mathematical points in the concept of black hole in general relativity? First, we retain the notion of pseudo-Riemannian manifold. The sign of the norm of a vector is the crucial property which allows one to define causality (the light cone).

The second main feature that we extract from the physical situation is the fact that a general relativity black hole has a non empty interior. We saw that this is the key difference between the Newtonian case in which all points are equivalent except the unique point where the mass lies, and the general relativistic case in which a whole region was causally disconnected from the rest of the space.

More precisely, as mathematicians, we ask a black hole to separate the pseudo-Riemannian manifold into two causally disconnected parts in the sense that no light-like geodesics can reach the second region from the first one. Notice that we do not include metric singularity in our mathematical black hole notion. In cosmology, in contrast, black holes always take root in a divergence of some metric invariant such as the curvature.

The anti de Sitter space is a solution of Einstein's equations with constant negative curvature. We consider this space as our framework. First, we define as *singular* the closed orbits of the action of some subgroup of the isometry group $\mathrm{SO}(2,n)$ of anti de Sitter. This is done in such a way to generalize to any dimensions the celebrated BTZ black hole. Then we prove that the resulting structure is a black hole in the sense that it cuts the space into two parts: an interior region from which every light-like geodesic intersects the singularity and an exterior region in which every point accepts at least one light-like geodesics which does not intersect the singularity. Notice that our black hole does not present any curvature singularity.

The second theme of this thesis is deformation quantization. The key ingredient of quantum mechanics is the noncommutativity of quantum observables. When one tries to measure the velocity and the position of a classical particle (such as a tennis ball or a planet), one can choose the order of measurement. It does not matter which of velocity or position is measured first. Quantum mechanics (the mechanics which governs subatomic particles) is very different. If you measure the position of an electron and then you measure its velocity, you do not get the same result as if you had measured the velocity first and then the position. That noncommutativity in measurements is the very foundation of the quantum mechanics. In the usual mathematical framework, it is implemented by describing each measurable quantity by an operator acting on a Hilbert space. The eigenvalues of these operators correspond to physical measurements. The position and momentum operators for example are respectively $f(x) \mapsto xf(x)$ and $f(x) \mapsto -i\hbar(\partial_x f)(x)$. These two operators obviously do not commute.

In a more abstract way, we say that noncommutativity of quantum mechanics is implemented by considering some noncommutative algebra of operators acting on a Hilbert space, while the classical mechanics deals with observables that are usual functions that form a commutative algebra. The procedure to pass from commutative function algebras to noncommutative operator algebras is the so-called *quantization* in physics.

In our sense [1], *deforming* a manifold is simply putting a one-parameter family of new noncommutative products on the set of functions on this manifold. We impose that it reduces to the usual commutative product when the parameter goes to zero. In order to speak of *quantization*, we ask the first order term in the expansion with respect to the parameter to somehow "contain" the symplectic structure given on the original manifold.

Questions that arise in this context are: is it possible to study causality in a noncommutative framework? does it apply to real physics?

The main result of the present work is not to directly address these large questions, but to build a concrete example in which one can work. Namely, we consider the anti de Sitter space — that is the simplest non trivial solution of Einstein's equations with constant negative curvature — that we endow with a black hole structure defined from the action of a subgroup of the isometry group. Then we select the physical part of the space — the one which is causally connected to infinity — and we perform a deformation of that part.

The work is divided into three main parts. In a first time (chapter 79) we define a "BTZ" black hole in anti de Sitter space in any dimension. That will be done by means of group theoretical and symmetric spaces considerations. A physical "good domain" is identified as an open orbit of a subgroup of the isometry group of anti de Sitter.

Then (chapter 84) we show that the open orbit is in fact isomorphic to a group (we introduce the notion of *globally group type* manifold) for which a quantization exists. The quantization of the black hole is performed and its Dirac operator is computed.

The chapter 81 is given in a pedagogical purpose: it exposes generalities about deformation quantization and careful examples with $\mathrm{SL}(2,\mathbb{R})$ and split extensions of Heisenberg algebras. Explicit decompositions are given for every algebra that will be used in the thesis in chapter 55. It serves to make the whole text more self contained and to fix notations. Basics of quantization by group action are given in appendix 82.6.

One more chapter is inserted (chapter 85). It contains two small results which have no true

1. Quantization is a very large field of mathematics; as far I know, the idea of noncommutativity is always present, but precise notion of "to quantize something" may vary from one subject to another

interest by themselves but which raise questions and call for further development. We discuss a product on the half-plane (or, equivalently, on the Iwasawa subgroup of $\mathrm{SL}(2,\mathbb{R})$) due to A. Unterberger. We show that the *quantization by group action* machinery can be applied to this product in order to deform the dual of the Lie algebra of that Iwasawa subgroup. Although this result seems promising, we show by two examples that the product is not universal in the sense that even the product of compactly supported functions cannot be defined on AdS_2 by the quantization induced by Unterberger's product.

Then we show that the Iwasawa subgroup of $\mathrm{SO}(2,n)$ (i.e. the group which defines the singularity) is a symplectic split extension of the Iwasawa subgroup of $SU(1,1)$ by the Iwasawa subgroup of $SU(1,n)$. A quantization of the two latter groups being known, a quantization of $SO(2,n)$ is in principle possible using an extension lemma (section 83.3). Properties of this product and the resulting quantization of AdS_l were not investigated because we found a more economical way to quantize AdS_4.

Abstract

This chapter deals with black holes in anti de Sitter spaces. The latter are the simplest non flat solutions to Einstein's equations with constant negative cosmological constant; they are in particular pseudo-Riemannian manifolds that carry a causal structure, physically due to the finiteness of speed of light. That physical restriction is mathematically encoded by the existence of three types of geodesics: the space-, time- and light-like ones, existence which is in turn implied by the non positivity of the metric. A causal structure is introduced by defining two points as *causally connected* when there exists a time- or light-like path connecting them.

The originality of our approach is that the l-dimensional space AdS_l is seen as a quotient of groups $\mathrm{SO}(2,l-1)/\mathrm{SO}(1,l-1) = G/H$, and that the special causal black hole structure is described in terms of orbits of the action of a subgroup of the isometry group of the manifold.

Using symmetric spaces techniques, we show that closed orbits of the Iwasawa subgroup of $\mathrm{SO}(2,l-1)$ naturally define a causal black hole singularity in anti de Sitter spaces in $l \geqslant 3$ dimensions. In particular, we recover for $l = 3$ the non-rotating massive BTZ black hole. The method presented here is very simple and in principle generalizable to any semisimple symmetric space.

The main references for this part are [1019, 1020, 1021].

79.2 Introduction

79.2.1 Physics and mathematics of black holes

SubSecGeneBH

79.2.1.1 Notion of Causality

This subsection is devoted to introduce the mathematical definition of a black hole from the intuitive physical notions of causality and maximality of the speed of light. Let us pose the origin of time and space respectively now and here. So we are at $(0,0)$. If we denote by c the seed of light, we cannot reach the moon before time $340000\,\mathrm{km}/c$. More generally we cannot reach a point at spacial distance d within a time inferior to d/c. Then the space is thus divided into three very different regions with respect to causality: the points that we can reach traveling slower than light, the points that only light can reach and points that we cannot reach at all.

It is convenient to introduce the function $s(t,x) = c^2t^2 - x^2$ which basically says you which points are accepted and which points are unaccepted. The mathematical way to implement these ideas is to consider a pseudo-Riemannian manifold (M,g). The **norm** of a vector $X \in T_xM$ is defined as $\|X\|^2 = g_x(X,X)$. There are three possibilities:

PgDefsGenre

— if $\|X\|^2 > 0$, we say that X is **time-like**,

— if $\|X\|^2 < 0$, we say that X is **space-like**,

— if $\|X\|^2 = 0$, we say that X is **light-like**.

A path $c\colon \mathbb{R} \to M$ is time, space or light-like when its tangent vector is everywhere time, space or light-like. The manifold M is **time orientable** if it accepts an everywhere time-like vector field. A *time orientation* is the choice of such a vector field. If T is a time orientation on M, we say that

a vector $X_x \in T_xM$ is **future directed** if $g_x(T_x, X) > 0$. From now we suppose that a choice of time orientation is possible and done.

The concept of causality is now easy to determine. If x and y belong to M, the point x has a **causal influence** on y if there exists a future directed path $c\colon [0,1] \to M$ such that $c(0) = x$ and $c(1) = y$. One has to notice that the relation *has a causal influence on* is not symmetric in general, but there exist some examples in which it is symmetric.

As example consider the space $M = \mathbb{R}^2$ endowed with the constant pseudo-Riemannian structure $g = \begin{pmatrix} c^2 & 0 \\ 0 & 1 \end{pmatrix}$. That space is the **Minkowski space**. The relation of causality is given by the previously mentioned function s; this relation is *never* symmetric and there exist pairs of point who have no causal effect on each other. If one takes the quotient by the relation $t \sim t+1$, we get a space in which the causality is everywhere symmetric.

79.2.1.2 Notion of singularity and black hole

Up to the choice of a time orientation, a pseudo-Riemannian manifold is comes with a canonical notion of causality. In order to have a black hole in our causal space we need an extra structure: the singularity. We take here a very conservative point of view and we say that a *singularity* in M is any strict subset of M. In the literature one often add conditions on the singularity such like to be a submanifold, time-like, connected,... of course most of "real live" singularities fulfil that kind of conditions.

The singularity defines two types of points in the space: the ones from which every time-like path intersect the singularity (from a physical point of view, these points correspond to observers who will fall in the singularity without doubt) and the points from which at least one time-like path does not intersect the singularity. We define the black hole associated with the singularity $\mathscr{S}$ as

$$BH = \big\{x \in M \text{ tel que } \forall \text{ future directed time-like path } c \text{ with } c(0) = x, \tag{79.1}$$
$$\exists t \geqslant 0 \text{ such that } c(t) \in \mathscr{S}\big\}. \tag{79.2}$$

The easiest example is given by defining a small line as singular in the Minkowski space as shown in figure 79.1.

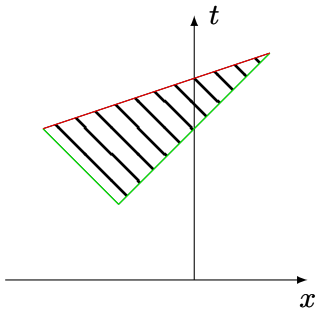


Figure 79.1: The red line is the singularity and the green zone is the black hole associated with. LabelFigEJRsWXw

In order the construction to be non trivial, we ask the black hole to be bigger than the singularity (that is of course part of the black hole), but smaller that the full space.

From a geometric point of view, a black hole is the data of a causal manifold M together with a subset $\mathscr{S} \subset M$ called *singularity* such that the whole manifold is divided into two parts: the *interior* and the *exterior* of the black hole. A point is said to be *interior* if all future light-like geodesics through the point have a non empty intersection with the singularity. A point is *exterior* if it is not interior. An important subset of the space is the *event horizon*: the boundary between these two subsets.

79.2.2 BTZ black hole

The BTZ black hole introduced and developed by Bañados, Teitelbaum, Zannelli and Henneaux in [1014, 1015] is an example of a black hole whose singularity is not motivated by metric divergences [2]. The construction is roughly as follows. We consider the anti de Sitter space in which we pick up a Killing vector field whose sign of norm is not constant. Then we perform a *discrete* quotient along the integral curves of this vector field. Of course we obtain a lot of closed geodesics. The point is that, in the region where the Killing vector field is space-like, these closed curves are space-like. That violates the physical principle of causality. For that reason, we decree that this region is singular or, equivalently, that the boundary of this region is singular. The BTZ singularity is then the loci where the chosen Killing vector field has a vanishing norm. Since discrete quotients do not affect local structures, the resulting space remains a solution of the $(2+1)$-dimensional general relativity with negative cosmological constant [3]. In this context one can define pertinent notions of *mass* and *angular momentum* which depend on the chosen Killing vector field.

¡¡ Avertissement/question à la lectrice !! 79.1
Il faut trouver une référence pour dire que la constante cosmologique est positive.

In the case of the *non-rotating massive* BTZ black hole, the structure of the singularity and the horizon are closely related to the action of a minimal parabolic (Iwasawa) subgroup of the isometry group of anti de Sitter, see [1013, 1017]. The whole work on the BTZ black hole and the fact that it belongs to the class of causal symmetric spaces (for definitions and some examples, see [1022]) motivate the following definition:

Definition 79.2.
*A **causal solvable symmetric black hole** is a causal symmetric space where the closed orbits of minimal parabolic subgroups of its isometry group define a black hole singularity. See section 79.8 for definitions of causality and singularity in the AdS case.* ⌈Def1

79.2.3 Generalization and group setting

SubSecGEneBHGrop

The original BTZ black hole was constructed in dimension three, but we will see in this chapter that, exploiting their group theoretical description, they can easily be generalized to any dimension, as pointed out in [1016, 1019]. Notice that higher-dimensional generalizations of the BTZ construction have been studied in the physics literature, by classifying the one-parameter isometry subgroups of $\mathrm{Iso}(AdS_l) = \mathrm{SO}(2, l-1)$, see [1023, 1024, 1025, 1026, 1027, 1028], but these approaches do not exploit the symmetric space structure of anti de Sitter.

The structure that will be described with full details in next pages may be summarized as follows. Take $G = \mathrm{SO}(2, l-1)$, fix a Cartan involution θ and a θ-commuting involutive automorphism σ of G such that the subgroup H of G of the elements fixed by σ is locally isomorphic to $\mathrm{SO}(1, l-1)$. The quotient space $M = G/H$ is a l-dimensional Lorentzian symmetric space, the *anti de Sitter space-time*. We denote by $\mathcal{G}$ and $\mathcal{H}$ the Lie algebras of G and H. We have the decomposition $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$ into the ± 1-eigenspace of the differential at e of σ that we denote again by σ. We also consider $\mathcal{G} = \mathcal{K} \oplus \mathcal{P}$, the Cartan decomposition induced by θ; and $\mathcal{A}$, a σ-stable maximally abelian subalgebra of $\mathcal{P}$. A positive system of roots is chosen and let $\mathcal{N}$ be the corresponding nilpotent subalgebra (see Iwasawa decomposition, theorem 51.188). Set $\overline{\mathcal{N}} = \theta(\mathcal{N})$, $\mathcal{R} = \mathcal{A} \oplus \mathcal{N}$ and $\overline{\mathcal{R}} = \mathcal{A} \oplus \overline{\mathcal{N}}$. Finally denote by $R = AN$ and $\overline{R} = A\overline{N}$ the corresponding analytic subgroups of G. One then has

Theorem 79.3.

2. It turns out that general relativity accepts a lot of solutions presenting metric divergences; or more precisely, there are a lot of *physical situations* from which Einstein's equations lead to divergences of some metric invariant such as the curvature.

3. For honesty, we have to warn the reader that the real world's cosmological constant has been measured very small but positive. We also have to point out that the four dimensional anti de Sitter space is a solution of general relativity *without masses*. From a physical point of view, this thesis has to be seen as a toy model.

The l-dimensional anti de Sitter space with $l \geqslant 3$, seen as the symmetric space $\mathrm{SO}(2,l-1)/\mathrm{SO}(1,l-1)$, becomes a causal solvable symmetric black hole, as defined above, when the closed orbits of the Iwasawa subgroup R of $\mathrm{SO}(2,l-1)$ and its Cartan conjugated $\overline{R}$ are said to be singular. There exists in particular a non empty event horizon. The group R has exactly two such closed orbits. ThoLeBut

This chapter intends to prove this theorem, and for the sake of completeness, we also analyze in some detail in section 78.2 the two-dimensional case, for which the construction does not yield a black hole structure.

The black hole causal structure is thus completely determined by the action of a solvable group. This observation gives prominence to potential embeddings of these spaces in the framework of noncommutative geometry, in defining noncommutative causal black holes (see also [1016]) through the existence of universal deformation formulae for solvable groups actions which have been obtained in the context of WKB-quantization of symplectic symmetric spaces [961, 1029]. These issues are investigated in chapter 84 and in [1020].

79.2.4 Some notations

We are going to use the following notations. We denotes the **free part** of the space AdS_l by $\mathscr{F}_l$; this is the subset of AdS_l for which there exists a light-like direction which escapes the singularity. We denote by BH_l the black hole in AdS_l; this is the set of points from which all the light-like geodesics intersect the singularity.

79.3 Visite guidée

79.3.1 En termes de BTZ

Nous travaillons dans $AdS_l = \mathrm{SO}(2,l-1)/\mathrm{SO}(1,l-1) = G/H$. Nous définissons les orbites fermées de AN et $A\bar{N}$ (le groupe d'Iwasawa de G et son conjugué par une involution de Cartan) comme *singulières*.

Il a été prouvé il y a déjà bien longtemps que cette définition donne lieu à une structure de trou noir. Cette structure est par ailleurs la même, en dimension 3, que celle du trou noir BTZ connu de la physique. J'ai récemment poussé un peu plus loin et donné la structure de l'horizon en dimension 4 en termes de celle en dimension 3. Il se fait que (théorème 79.49)

Theorem 79.4.
L'horizon de AdS_4 est donné par

$$\mathscr{H}_4 = G_V \cdot \iota(\mathscr{H}_3) \cup G_X \cdot \iota(\mathscr{H}_3), \tag{79.3}$$

où $\iota\colon \mathbb{R}^4 \to \mathbb{R}^5$ est l'inclusion de AdS_3 dans AdS_4 et où les groupes G_V et G_X sont donnés par

$$G_V = \{e^{\alpha V} \text{ tel que } \alpha \in \mathbb{R}\}, \tag{79.4}$$

le vecteur V étant l'élément de base de l'espace de racine $\mathcal{G}_{(0,1)}$ de $\mathrm{SO}(2,3)$, et X est l'élément de base de $\mathcal{G}_{(0,-1)}$. Ces espaces de racines sont vides dans le cas de AdS_3.

L'inclusion ι peut également être vue comme l'inclusion du groupe $\mathrm{SO}(2,3)$ dans $\mathrm{SO}(2,4)$. La preuve est faite avec du calcul matriciel explicite très peu généralisable à d'autres espaces symétriques.

L'énoncé de ce théorème est la seule chose élégante de la section 79.11. Le reste est du calcul matriciel. Ce théorème donne, cependant, une bonne idée de ce vers quoi on va : il semble possible que les horizons en dimension supérieure s'obtiennent par récurrence. Le groupe qui générerait la singularité en dimension l serait le groupe généré par les espaces de racines $(1,0)$ et $(0,1)$ de $\mathrm{SO}(2,l-1)$.

Afin de trouver des preuves plus intrinsèques, on commence par bien définir les différents éléments de l'algèbre, et en particulier la base de $\mathcal{Q}$ en termes des espaces de racines. Cela se passe

à la section 79.16. Je définit par exemple

$$\begin{aligned} q_0 &= (X_{++})_{\mathcal{QK}} \\ q_2 &= (X_{++})_{\mathcal{QP}}, \end{aligned} \tag{79.5}$$

et je montre que le premier est de norme (de Killing) positive et le second de norme négative, mais qu'en valeur absolue, ils ont la même norme. Je choisit donc X_{++} de telle façon que q_0 et q_1 soient normés à 1. Les vecteurs $q_0 \pm q_1$ sont donc de genre lumière.

Toute une série de propriétés sont ensuite prouvées. Le but est évidement de construire, de façon intrinsèque, une base de $\mathcal{A}$, $\mathcal{N}$, $\mathcal{K}$ et de $\mathcal{Q}$ de telle façon à avoir toutes les propriétés agréables que les matrices explicites avaient.

Cette partie est destinée à être remplacée par une application du théorème de structure de Pyatetskii-Shapiro. Un petit changement de base sera toutefois indispensable parce que l'élément dont l'annulation de la norme du champ de vecteur fondamental donne la singularité n'est pas dans la base donnée par Pyatetskii-Shapiro.

Tant que l'on travaillait avec des matrices et qu'on utilisait explicitement le fait que AdS était un sous-ensemble de $\mathbb{R}^n$, nous utilisions la caractérisation suivante de la singularité :

$$\mathscr{S} \equiv t^2 - y^2 = 0. \tag{79.6}$$

Maintenant, il est bon d'utiliser une caractérisation de la singularité qui ne fait pas appel aux coordonnées. Une telle caractérisation existe : si J_1 est un élément de $\mathcal{A} \cap \mathcal{H}$ (qui est de dimension 1), alors la singularité est donnée par

$$\mathscr{S} \equiv \|J_1^*\| = 0 \tag{79.7}$$

où J_1^* est le champ de vecteur fondamental de l'action de G sur G/H associé au vecteur J_1. Cette caractérisation fait qu'un point $[g] \in G/H$ est dans la singularité si et seulement si le vecteur

$$\texttt{proj}_{\mathcal{Q}}\left(\mathrm{Ad}(g^{-1})J_1\right) \tag{79.8}$$

VisiteEqprQcaract

a une norme nulle. Ici, $\texttt{proj}_{\mathcal{Q}}$ est la projection sur $\mathcal{Q}$.

À part une foule de petits détail encore à vérifier, il est maintenant prouvé, en utilisant la caractérisation (79.8), qu'un point $[kan]$ est dans la singularité si et seulement s'il appartient à $[AN]$, $[A\bar{N}]$, $[-\mathbb{1}_{\mathrm{SO}(2)}AN]$ ou $[-\mathbb{1}_{\mathrm{SO}(2)}A\bar{N}]$, c'est-à-dire à une des orbites fermées de AN ou de $A\bar{N}$.

Tout cela est dans le chapitre 79.

79.3.2 En termes de généralisations

Afin de se mettre dans une perspective de généralisation, l'idée suivante est proposée.

(1) On considère un espace homogène symétrique G/H où G a 1000 décompositions d'Iwasawa possibles.

(2) On sait par des arguments d'hermiticité et de $\mathcal{Z}(\mathcal{K})$ non nul que la composante d'Iwasawa de G est une j-algèbre.

(3) Il y a sûrement un argument pour dire qu'il existe des choix d'Iwasawa qui font que les racines positives et les éléments correspondants de $\mathcal{A} \oplus \mathcal{N}$ tombent exactement dans les A ,Z et V de la décomposition en j-algèbres élémentaires.

(4) On choisit cette décomposition particulière d'Iwasawa comme décomposition "de référence".

(5) On définit la singularité sur G/H par $\|H1 + H2\| = 0$. Ici, c'est la première fois que le H apparaît dans la construction.

(6) On considère l'Iwasawa qui change de base dans $\mathcal{A}$ pour choisir $J1 = H1 + H2$ et $J2 = H1 - H2$. Cela devrait être fait sans changer de décomposition $\mathcal{G} = \mathcal{K} \oplus \mathcal{P}$.

ItemVGDern

(7) On prouve que la singularité est les orbites fermées de AN et $A\theta(N)$ pour cette nouvelle Iwasawa.

Le point (7) est là uniquement pour montrer que l'ensemble de la construction redonne le BTZ déjà connu.

79.4 Connectedness of groups and anti de Sitter spaces

PgDisGeoConnSO

Let us give some detail on the geometric nature of the two connected components of $\mathrm{SO}(p,q)$ [4]. What is proved in [894] is that $\mathrm{SO}(p,q)$ is homeomorphic to the topological product

$$\mathrm{SO}(p,q) = \mathrm{SO}(p,q) \cap \mathrm{SU}(p+q) \times \mathbb{R}^d = \mathrm{SO}(p,q) \cap \mathrm{SO}(p+q) \times \mathbb{R}^d$$

for some $d \in \mathbb{N}$. Hence an element of $\mathrm{SO}(p,q)$ reads

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \times v$$

where $v \in \mathbb{R}^d$, $A \in \mathrm{O}(p)$, $B \in \mathrm{O}(q)$ are such that $\det A \det B = 1$. The v part corresponds to boost while A and B correspond to pure temporal and pure spatial rotations. An element of $\mathrm{O}(n)$ has always determinant equals to ± 1. Therefore one can decompose the rotation part as $(\det A = \det B = 1) \otimes (\det A = \det B = -1)$. Both parts are connected.

Hence the first connected component contains $\mathbb{1}$ while the second one contains the element that simultaneously changes the sign of one spacial and one time direction.

79.4.1 The quotient for anti de Sitter

Homogeneous space considerations (see section 79.5) will naturally lead us to define the anti de Sitter space as the quotient $G/H = \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1)$ while the black hole definition (section 79.8) needs to consider Iwasawa decompositions of G. So we face the problem that the Iwasawa theorem 51.188 only works with connected groups. In order to prevent any problems of this type, we prove now that, if G_0 and H_0 denote the identity component of $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$ respectively, then $G/H = G_0/H_0$.

The groups that are considered here have only two connected components G_0 and G_1. We can chose $i_1 \in G_1 \cap H$ such that $i_1^2 = \mathbb{1}$. Using lemma 55.1, it easy to prove that

— $G_0 G_0 = G_0$,
— $G_0 G_1 = G_1$,
— $G_1 G_1 = G_0$.

For the last one, take g and g' in G_1. Then consider g_0 and g_0' in G_0 such that $g = g_0 i_1$ and $g' = g_0' i_1$. If $g_0(t)$ and $g_0'(t)$ are path from $\mathbb{1}$ to g_0 and g_0', then the path $g_0(t) i_1 g_0'(t) i_1$ is a path from $\mathbb{1}$ to gg'.

PropGHconn

Proposition 79.5.
The map

$$\begin{aligned} \psi \colon G/H &\to G_0/H_0 \\ [g] &\mapsto \overline{g_0} \end{aligned} \tag{79.9}$$

where we define $g_0 = g$ when $g \in G_0$ or $g_0 = g i_1$ when $g \in G_1$ is a diffeomorphism. The classes are $[g] = \{gh \text{ tel que } h \in H\}$ and $\overline{g} = \{gh_0 \text{ tel que } h_0 \in H_0\}$.

Proof. First we prove that ψ is well defined. For that we suppose that $[g] = [g']$. There are three cases:

(1) The elements g and g' both belong to G_0. In this case, $g' = gh_0$ with $h_0 \in H_0$ and $\overline{gh} = \overline{g}$.
(2) The element g belongs to G_0 while g' belongs to G_1. In this case, $g' = gh$ with $h = h_0 i_1$ and $h_0 \in H_0$. Then $\psi[g] = \overline{g}$ and $\psi[g'] = \overline{(gh_0 i_1)_0} = \overline{gh_0 i_1 i_1} = \overline{gh_0} = \overline{g}$.
(3) The case with g and g' in G_1 is similar.

The fact that the map ψ is surjective is clear. For injectivity, let $\psi[g] = \psi[g']$, i.e. there exists a h_0 in H_0 such that $g_0' = g_0 h_0$. Thus we have $g' i_1^k = g i_1^l h_0$ with $k, l = 0, 1$ following the cases. Then $g' = g i_1^l h_0 i_1^k$ in which $i_1^l h_0 i_1^k$ belongs to H, so that $[g'] = [g]$.

□

4. See lemma 55.2. A discussion about the physics is in [851].

79.5 Symmetric space structure on anti de Sitter

SecSymeStructAdS

The l-dimensional anti de Sitter space AdS_l can be described as set of points $(u, t, x_1, \dots, x_{l-1}) \in \mathbb{R}^{2,l-1}$ such that $u^2 + t^2 - x_1^2 - \dots - x_{l-1}^2 = 1$. The next few pages are devoted to describe the homogeneous and symmetric space structures on AdS_l induced by the transitive an isometric action of $\mathrm{SO}(2, l-1)$. We suppose that the groups $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$ are parametrized in such a way that the second, seen as subgroup of the first one, leaves unchanged the vector $(1, 0, \dots, 0)$. In this case, proposition 4.3 of chapter II in [813] provides the homogeneous space isomorphism

$$\begin{aligned} \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1) &\to AdS_l \\ [g] &\mapsto g \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix} \end{aligned} \tag{79.10}$$

where the dot denotes the usual "matrix times vector" action of the representative $g \in [g]$ in the defining representation of $\mathrm{SO}(2, l-1)$ on $\mathbb{R}^{2,l-1}$. As far as notations are concerned, the classes are taken from the right: $[g] = \{gh \text{ tel que } h \in H\}$; in particular the class of the identity e is denoted by ϑ; the groups $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$ are denoted by G and H respectively and their Lie algebras by $\mathcal{G}$ and $\mathcal{H}$. Following proposition 79.5, we can in fact only consider the identity components of G and H. We denote by τ the natural action of G on G/H:

$$\begin{aligned} \tau &: G \times AdS_l \to AdS_l \\ \tau_r[g] &= [rg] \end{aligned} \tag{79.11}$$

As far as dimensions are concerned, a candidate $R \subset G$ such that $R \cdot \vartheta$ is open must satisfy

$$\dim \mathcal{R} \geqslant \dim M. \tag{79.12}$$

cond_dim

The case that interest us is $G = \mathrm{SO}(2, n)$ and $H = \mathrm{SO}(1, n)$:

$$M = AdS_{n+1} = \frac{\mathrm{SO}(2, n)}{\mathrm{SO}(1, n)},$$

so that we have to consider the action of $\mathrm{SO}(2, n)$ on AdS_n. If ANK is the Iwasawa decomposition of $\mathrm{SO}(2, n)$, we can consider more particularly the action of $R = AN$, and ask us if the orbit $R \cdot \vartheta$ is open or not. It is easy to see that the condition (79.12) is satisfied. Indeed,

$$\dim \mathfrak{g} = \frac{n(n-1)}{2} + 2n + 1, \qquad \dim \mathfrak{k} = \frac{n(n-1)}{2} + 1,$$

so that $\dim(\mathcal{A} \oplus \mathcal{N}) = 2n$, but $\dim AdS_n = n$. The Iwasawa subgroup AN is a candidate for $AN \cdot \vartheta$ to be open in AdS_n.

Proposition 79.6.
The homogeneous space AdS_l is reductive. PropAdSreduct

Proof. The proof relies on lemma 51.60 and the fact that $\mathrm{SO}(2, n)$ is semisimple. From the Killing form of G, one defines

$$\mathcal{Q} = \mathcal{H}^\perp = \{X \in \mathcal{G} : B(X, H) = 0 \, \forall H \in \mathcal{H}\}.$$

Let H, $H' \in \mathcal{H}$ and $Y \in \mathcal{Q}$. From ad-invariance of the Killing form, we have $B([H, Y], H') = 0$. Hence $(\mathrm{ad}(\mathcal{H})\mathcal{Q}) \subset \mathcal{Q}$ and the claim is proved.

□

Matrices of $\mathrm{SO}(2, n)$ are $(2+n) \times (2+n)$ matrices while the n-dimensional anti de Sitter space is a quotient of $\mathrm{SO}(2, n-1)$. In order to avoid confusions, we will reserve the letter n to the study of the group $\mathrm{SO}(2, n)$ and the letter l will denote the dimension of the anti de Sitter space which will thus be AdS_l.

We define the involutive automorphism $\sigma = \mathrm{Id}|_{\mathcal{H}} \oplus (-\mathrm{Id})|_{\mathcal{Q}}$. The vector space $\mathcal{Q}$ can be identified with the tangent space $T_{[e]}AdS_l$, and that identification can be extended by defining $\mathcal{Q}_g = dL_g\mathcal{Q}$. In this case $d\pi\colon \mathcal{Q}_g \to T_{[g]}AdS_l$ is a vector space isomorphism. PgdpibaseQTgM An homogeneous metric on $T_{[g]}AdS_l$ is defined as in subsection 66.2.1.

Cartan decomposition of $\mathrm{SO}(2, l-1)$ are of crucial importance in chapter 79, so that we want to use a Cartan involution θ such that $[\sigma, \theta] = 0$ (see [956] page 153, theorem 2.1). One can show that $X \mapsto -X^t$ has that property. The corresponding Cartan decomposition is described in subsection 66.5.1.

As a consequence of relations (66.93),

$$d\pi\,\mathrm{Ad}(h) = \mathrm{Ad}(h)d\pi \tag{79.13}$$ EqdpiAdpi

because, if $X \in \mathcal{Q}$, $d\pi^{-1}(X) = \{X + Y \text{ tel que } Y \in \mathcal{H}\}$, so $\mathrm{Ad}(h)Y \in \mathcal{H}$ and $\mathrm{Ad}(h)X \in \mathcal{Q}$.

79.5.1 Anti de Sitter as symmetric space

pg:AdS_n_syme

We know the decomposition $\mathfrak{so}(2, n) = \mathcal{Q} \oplus \mathcal{H}$. From equation (66.93) one can find an involutive automorphism σ of $\mathcal{G}$ which leaves $\mathcal{H}$ invariant.

There exists a neighbourhood U of 0 in $\mathfrak{so}(2, n)$ on which exp is diffeomorphic to a neighbourhood V of e in $\mathrm{SO}(2, n)$. We define $\sigma_G\colon V \to V$ by $\sigma_G(e^X) = e^{\sigma X}$. Now, this σ_G can be extended to the whole G. From now we will denote by σ this map or its differential (i.e. an abuse of notation between σ and $d\sigma_e$).

All this make $(\mathrm{SO}(2, n), \mathrm{SO}(1, n))$ a symmetric pair. Since $H = \mathrm{SO}(1, n)$ is connected and fixed by σ, $H = H_\sigma = (H_\sigma)_0$. Thus theorem 66.18 gives us a Cartan involution θ on $\mathcal{G}$ such that $[\sigma, \theta] = 0$ and theorem 66.17 gives a symmetric structure to $M = G/H$. Now we understand the computations of page 3620.

79.6 Open and closed orbits

79.6.1 Open orbits in anti de Sitter spaces

Let us start by computing the closed orbits of the actions of AN and $A\bar{N}$ on AdS_l. In order to see if $[g] \in AdS_l$ belongs to a closed orbit of AN, we "compare" the space spanned by the basis $\{d\pi dL_g q_i\}$ of $T_{[g]}AdS_l$ and the space spanned by the fundamental vectors of the action. If these two spaces are equal, then $[g]$ belongs to an open orbit (because a submanifold is open if and only if it has same dimension as the main manifold). That idea is precisely contained in the following theorem which holds for any homogeneous space $M = G/H$.

¡¡ Avertissement/question au lecteur !! 79.7
Il faut trouver une référence pour ce théorème.

The strategy is to to check openness of the R-orbit of $[g]$ by checking openness of the $\mathbf{Ad}(g^{-1})R$-orbit of ϑ using the theorem 66.48.

The problem is simplified by the following remark. We know that matrices of K and H are given by

$$K \rightsquigarrow \begin{pmatrix} \mathrm{SO}(2) & \\ & \mathrm{SO}(n) \end{pmatrix}, \quad H \rightsquigarrow \begin{pmatrix} 1 & \\ & \mathrm{SO}(1, n) \end{pmatrix}, \tag{79.14}$$ eq:K_H_SO

so we obviously have

$$\bigcup_{s\in\mathrm{SO}(2)} \tau_{AN}([s]) = \bigcup_{\substack{s\in\mathrm{SO}(2)\\ h\in\mathrm{SO}(n)}} [ANsh] = \bigcup_{k\in K} [ANk] = [G].$$

This is nothing else than the fact that the AN-orbits are AN-invariant. So the K part of $[g] = ank$ alone fixes the orbit which contains $[g]$ and we have at most one orbit for each element in $\mathrm{SO}(2)$.

Computations using theorem 66.48 show that the R-orbits of $[\mu]$ with

$$\mu = \begin{pmatrix} \cos\mu & \sin\mu & \\ -\sin\mu & \cos\mu & \\ & & \mathbb{1} \end{pmatrix}$$

is not open if and only if $\sin\mu = 0$. We will see later that they are actually closed (page 3963), so that the singularity is described as

$$\mathscr{S} = [AN(\pm\mathbb{1}_{\mathrm{SO}(2)})] \bigcup [A\bar{N}(\pm\mathbb{1}_{\mathrm{SO}(2)})]. \tag{79.15}$$

Because of AN-invariance of the AN-orbits, the equation of the AN-closed orbits can be expressed as

$$\sin\mu = 0. \tag{79.16}$$

Let us recall that $-\mathbb{1}_{\mathrm{SO}(2)} = k_\theta = e^{\pi q_0}$. With these notations, we have that the closed orbits of AN are

$$[AN] \quad \text{and} \quad [ANk_\theta] = [k_\theta A\bar{N}], \tag{79.17}$$

while the closed orbits of $A\bar{N}$ are given by

$$[A\bar{N}] \quad \text{and} \quad [A\bar{N}k_\theta] = [k_\theta AN]. \tag{79.18}$$

Notice that there are some differences between the two choices of Iwasawa decompositions of equations (66.148) and (66.127) in the determination of open and closed orbits. In the AN Iwasawa decomposition, up to matrices of $\mathcal{H}$ (given in equation (66.91)), a general matrix of $\mathcal{R}$ is $jJ_1 + mM + lL + kJ_2$. If we note $x = m + l$,

$$\mathcal{R} \rightsquigarrow \begin{pmatrix} 0 & x & k & -x \\ -x & & & \\ k & & & \\ -x & & & \end{pmatrix} \tag{79.19}$$

and it is obvious that the matrix q_0 can't be obtained by combinations of such matrices. So the R-orbit of ϑ is not open.

We can do the same computation with the Iwasawa group $\bar{\mathcal{R}} = \mathcal{A} \oplus \bar{\mathcal{N}}$. A general element of this is of the form $jJ_1 + kJ_2 + nN + fF$. If we write $a = n + f$ and $b = n - f$, we get

$$\begin{pmatrix} 0 & a & k & a & 0 \\ -a & 0 & b & j & 0 \\ k & b & 0 & b & 0 \\ a & j & -b & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{79.20}$$

Looking at the positions of the a, we see that it is impossible to put the element q_0 under that form. We deduce that the $\bar{R}$-orbits of ϑ are not open neither.

That situation is, however, not generic. If we use for example the other Iwasawa decomposition, the one of subsection 66.5.6, the result is completely different. We have

$$q_0 = \mathtt{proj}\left(\frac{N+M}{2}\right), \qquad q_1 = \mathtt{proj}\, H_2, \qquad q_2 = \mathtt{proj}\left(N - \frac{N+M}{2}\right), \tag{79.21}$$

and other elements of $\mathcal{Q}$ are projections of the matrices V_i's. So we see that the map $\mathtt{proj}\colon \mathcal{R} \to \mathcal{Q}$ is surjective and the orbit $R \cdot \vartheta$ is open.

Here is some explicit matricial computation.

$$M + N = 2\underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{\in \mathcal{Q}} + \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{\in \mathcal{H}},$$

thus $\mathtt{proj}(\frac{M+N}{2})$ is yet a part of $\mathcal{Q}$. An other:

$$\underbrace{\begin{pmatrix}0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&0\end{pmatrix}}_{=m_2}=\underbrace{\begin{pmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{pmatrix}}_{=H_1}+\underbrace{\begin{pmatrix}0&0&0&0\\0&0&0&1\\0&0&0&0\\0&1&0&0\end{pmatrix}}_{=h\in\mathcal{H}},$$

so that $\mathtt{proj}\, H_1 = \mathtt{proj}(m_2 - h) = m_2$. Third,

$$\underbrace{\begin{pmatrix}0&0&0&1\\0&&&\\0&&&\\1&&&\end{pmatrix}}_{=m_3}=\underbrace{N-\frac{M+N}{2}}_{\in\mathcal{R}}-\underbrace{\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}}_{\in\mathcal{H}},$$

thus $\mathtt{proj}(N - \frac{M+N}{2}) = m_3$. The last possibility in $\mathcal{Q}$ is $m_i = E_{1i} + E_{i1}$ $(i \geqslant 5)$, but

$$\underbrace{V_i}_{\in\mathcal{R}}=\underbrace{E_{1i}+E_{i1}}_{=m_i}+\underbrace{E_{3i}-E_{i3}}_{\in\mathcal{H}}.$$

79.7 Some complements

79.7.1 A first brute force computation

Let us use the "old" Iwasawa decomposition for a little demonstrative and inessential computation. The exponentiations from $\mathcal{A}\oplus\mathcal{N}$ to AN is given at page 3630. Remark that a matrix of $\mathrm{SO}(1,n)$ leave unchanged the first column of $\mathrm{SO}(2,n)$:

$$\begin{pmatrix}a&.&.&.\\b&.&.&.\\c&.&.&.\\d&.&.&.\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&.&.&.\\0&.&.&.\\0&.&.&.\end{pmatrix}=\begin{pmatrix}a&.&.&.\\b&.&.&.\\c&.&.&.\\d&.&.&.\end{pmatrix}.$$

Thus, in order to compute the orbit $[AN]$ of $[\mathbb{1}]$, one can begin to compute a general matrix of AN and impose conditions on the first column (it will not be affected by the equivalence classes). For example, in order to see if $[-\mathbb{1}] \in [\mathrm{SO}(2)]$ belongs to the orbit $[AN]$, we compute: PgExplAN

$$\begin{pmatrix}\cosh\xi&0&\sinh\xi&0\\0&\cosh\eta&0&\sinh\eta\\\sinh\xi&0&\cosh\xi&0\\0&\sinh\eta&0&\cosh\eta\end{pmatrix}\begin{pmatrix}1-2ab&b+a&2ab&b-a\\-b-a&1&b+a&0\\-2ab&b+a&1+2ab&b-a\\b-a&0&a-b&1\end{pmatrix} \tag{79.22}$$

eq:gene_R

and we impose the first column to be $(-1,0,0,0)$: eq:S_14

$$(1-2ab)\cosh\xi-(2ab)\sinh\xi=-1 \tag{79.23a}$$
eq:S1

$$(-b-a)\cosh\eta+(b-a)\sinh\eta=0 \tag{79.23b}$$
eq:S2

$$(1-2ab)\sinh\xi-(2ab)\cosh\xi=0 \tag{79.23c}$$
eq:S3

$$(-b-a)\sinh\eta+(b-a)\cosh\eta=0. \tag{79.23d}$$
eq:S4

It is easy to see that these equations doesn't accept any solutions. Indeed, the sum of equations (79.23b) and (79.23d) gives

$$-2a(\cosh\eta+\sinh\eta)=0.$$

The possibility $a = 0$ gives $\cosh\xi = -1$ in (79.23a); but the hyperbolic cosine is always bigger than one. Then $\cosh\eta = -\sinh\eta$. In this case, (79.23d) gives $2b\cosh\eta = 0$; since $\cosh\eta = 0$ is not possible, b must be zero. But with $b = 0$, (79.23a) gives once again $\cosh\xi = -1$.

Now we have to see that the matrices e^{V_i} and e^{W_j} doesn't change the result. Since $W_j \in \mathcal{H}$, it is clear that they will not change any result. If we compute e^{cV_5} for example ($V_5^3 = 0$), we find

pg influence_V_W

$$e^{cV_5} = \begin{pmatrix} 1+c^2/2 & . & -c^2/2 & . & c \\ . & 1 & . & . & . \\ c^2/2 & . & 1-c^2/2 & . & c \\ . & . & . & 1 & . \\ c & . & -c & . & 1 \end{pmatrix} \tag{79.24}$$

EqExpDeV

By multiply it by matrices of AN, we find

$$\begin{pmatrix} . & . & . & 0 \\ . & . & . & 0 \\ . & . & . & 0 \\ . & . & . & 0 \\ . & . & . & 1 \end{pmatrix} \begin{pmatrix} 1+c^2/2 & . & . & . & . \\ 0 & . & . & . & . \\ c^2/2 & . & . & . & . \\ 0 & . & . & . & . \\ c & . & . & . & . \end{pmatrix} = \begin{pmatrix} . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ c & . & . & . & . \end{pmatrix}.$$

Then $c = 0$ if we want it to be equal to $\pm\mathbb{1}$. As far as the matrices W_j are concerned, it is even simpler: $W_j \in \mathcal{H}$; then it will not affect the classes.

All that suppose that N can *globally* be written under the form

$$e^{aM} e^{bN} \prod_{i=5}^{n} e^{c_i V_i} \prod_{j=5}^{n} e^{c_j W_j}$$

It is locally true from lemma 53.61.

79.7.2 Search for $Z(K)$

Now we are going to find the center of K. From the explicit form (79.14), we see that $Z(K) = \mathrm{SO}(2)$. Let us show it more abstractly. We consider the decomposition $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$ of $\mathfrak{so}(2,n)$ with respect to the involution σ:

$$\mathcal{K} = \mathcal{K}_{\mathcal{Q}} \oplus \mathcal{K}_{\mathcal{H}};$$

This is an expression of the fact that $\mathrm{SO}(2,n)/\mathrm{SO}(1,n)$ is a symmetric space. Since $\dim Z(K) = 1$, it is a subset of $\mathcal{K}_{\mathcal{Q}}$ or of $\mathcal{K}_{\mathcal{H}}$ because $Z(K)$ is σ

On the other hand, we know [5] that a Cartan involution can be written as $\theta = \mathrm{Ad}(\exp Z)$ for a $Z \in \mathcal{Z}(\mathcal{K})$. Then

$$\begin{aligned} 0 = [\sigma, \theta](X) &= \sigma \,\mathrm{Ad}(e^Z)X - \mathrm{Ad}(e^Z)\sigma X \\ &= \sigma e^{\mathrm{ad}\, Z} X - \mathrm{Ad}(e^Z)\sigma X \\ &= \mathrm{Ad}(e^{\sigma Z})\sigma X - \mathrm{Ad}(e^Z)X. \end{aligned} \tag{79.25}$$

Since G has $\pm\mathbb{1}$ as center, this implies $e^{\sigma Z} = e^Z$. Then $e^Z \in H$. We can't however conclude that $z \in \mathcal{H}$.

79.7.3 The same with the "old" Iwasawa decomposition

A general matrix of $\widetilde{\mathcal{R}}$ is $mM + nN + tH_1 + uH_2$; with the change of variable $x = n + m$, $y = n - m$, $a = t + u$ and $b = u - t$, it is

$$r = \begin{pmatrix} 0 & x & a & y \\ -x & 0 & x & b \\ a & x & 0 & y \\ y & b & -y & 0 \end{pmatrix}. \tag{79.26}$$

5. Faudra un peu voir pourquoi

Now we want to explicitly compute $\mathcal{R}_z$ for a general matrix z in the "SO(2)" part of SO$(2,n)$, i.e. for

$$z = \begin{pmatrix} \cos u & \sin u & & \\ -\sin u & \cos u & & \\ & & 1 & 0 \\ & & 0 & 1 \end{pmatrix}.$$

The result is

$$\mathrm{Ad}(z)\mathcal{R} = \begin{pmatrix} 0 & x & a\cos u + x\sin u & y\cos u + b\sin u \\ -x & 0 & x\cos u - a\sin u & b\cos u - y\sin u \\ a\cos u + x\sin u & x\cos u - a\sin u & 0 & y \\ y\cos u + b\sin u & b\cos u - y\sin u & -y & 0 \end{pmatrix}. \tag{79.27}$$

We are interested in the projection of these matrices with respect to $\mathcal{H}$ (see equation (66.91)); then in order to see if $\mathcal{R}_z = \mathcal{R}$, we have to compare

$$\begin{pmatrix} 0 & x & a\cos u + x\sin u & y\cos u + b\sin u \\ -x & & & \\ a\cos u + x\sin u & & & \\ y\cos u + b\sin u & & & \end{pmatrix}$$

with

$$\begin{pmatrix} 0 & x' & a' & y' \\ -x' & & & \\ a' & & & \\ y' & & & \end{pmatrix}$$

By working on a, b, x and y, we can easily fit the first matrix on the second one if $c \neq 0$. In other words, $\mathcal{R}_z \neq \mathcal{R}$ only if the SO(2) part of z is the rotation of an angle of $\pm\pi/2$. All that we can say now is that $\cos u = 0$ is not in the $\widetilde{R}$-orbit of ϑ.

Up to here we have not taken the matrices V_i and W_j into account. It is rather easy to see that they don't change anything because (taking $i = j = 5$ for sake of notational simplicity)

$$V_i + W_i \simeq \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \\ 1 & & & & \end{pmatrix}$$

but

$$\mathrm{Ad}(z)(V_i + W_i) \simeq \begin{pmatrix} 0 & 0 & 0 & 0 & \sin u + \cos u \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \\ \sin u + \cos u & & & & \end{pmatrix}.$$

79.7.4 A non-open orbit and a precision

subsec:precision

We consider the situation where $[\sigma, \theta] = 0$ and the Iwasawa decomposition of SO$(2,n)$ for which R_H is a subgroup of R (R_H is the "AN" of SO$(1,n)$). We denote by ϑ the identity class: $\vartheta = [e] = eH$. Since $\mathcal{H} = \mathcal{R}_H \oplus \mathcal{K}_\mathcal{H}$,

$$T_\vartheta(R\vartheta) = \mathtt{proj}_\mathcal{H}(\mathcal{R}) = (\mathcal{R} + \mathcal{H})/\mathcal{H}$$

where the sum $\mathcal{R} + \mathcal{H}$ is not a direct sum. But $\mathcal{R} + \mathcal{H} = \mathcal{R} + (\mathcal{R}_H \oplus \mathcal{K}_\mathcal{H}) = \mathcal{R} \oplus \mathcal{K}_H$ because $\mathcal{K}_H \subset \mathcal{K}$. Thus

$$T_\vartheta(R\vartheta) = (\mathcal{R} \oplus \mathcal{K}_H)/(\mathcal{R}_H \oplus \mathcal{K}_H) = \mathcal{R}/\mathcal{R}_H.$$

pg:subt_tilde
In the case AdS_3, $\dim(R\vartheta) = 6-3 = 3 < 4$; so that the identity orbit is not open. It is important to note that it contradicts the result of page 3949. The reason is that the latter was obtained with the "old" Iwasawa decomposition.

From now the "old" decomposition will be denoted by a tilde:

$$\mathfrak{g} = \widetilde{\mathcal{A}} \oplus \widetilde{\mathcal{N}} \oplus \mathcal{K}.$$

In this case, there exists a $k_0 \in K$ such that $\mathrm{Ad}(k_0^{-1})\widetilde{\mathcal{A}} = \mathcal{A}$. Moreover with this k_0, we have $\widetilde{A} = k_0 A k_0^{-1}$ and $\widetilde{N} = k_0 N k_0^{-1}$. This result can be found in [813] and maybe [6] in [850] [7].

With these precisions, the previous computations are still relevant because

$$\widetilde{R}[g] = k_0 R k_0^{-1}[g], \tag{79.28}$$

which assures that the $\widetilde{R}$-orbit of $[g]$ is the R-orbit of $[k_0^{-1}g]$ translated by k_0. Since the translation is a diffeomorphism, the openness is not affected by the translation. So the open $\widetilde{R}$-orbit of ϑ stated at page 3949 is now the open R-orbit of k_0^{-1}.

Now let us find out this famous k_0 matrix. Since the orbits are defined by the elements of the center of K (i.e. a bloc-diagonal matrix in $\mathrm{SO}(2,n)$ with a $\mathrm{SO}(2)$ matrix in the upper left corner and $\mathbb{1}$ anywhere else), we guess that k_0 is such a matrix. Before to compute it, we take a change of basis in $\widetilde{\mathcal{A}}$: we consider $\frac{1}{2}(H_1 + H_2)$ and $\frac{1}{2}(H_2 - H_1)$ instead of H_1 and H_2. If we consider $d = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, the problem is to find a matrix $z \in \mathrm{SO}(2)$ such that

$$\begin{pmatrix} z & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} 0 & d \\ d & 0 \end{pmatrix} \begin{pmatrix} z^{-1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can easily find that the only solution is $z = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then

$$k_0 = \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & 1 & 0 \\ & & 0 & 1 \end{pmatrix}. \tag{79.29}$$

79.7.5 The same with the "old" Iwasawa decomposition

This useless point shows how to get the same conclusion with a bad choice of Iwasawa decomposition. If we consider some z in the centrer of K ($\mathrm{SO}(2)$) and $\widetilde{R}_z = z\widetilde{R}z^{-1}$, then the $\widetilde{R}$-orbit of ϑ is the $\widetilde{R}$-orbit of $[z^{-1}]$ translated by z and the R-orbit of $k_0 z^{-1}$ translated by zk_0^{-1}:

$$\widetilde{R}_z = z\widetilde{R}z^{-1} = zk_0 R k_0^{-1} z^{-1}. \tag{79.30}$$

Thus the study of the openness of the $\widetilde{R}_z$-orbit of ϑ is the study of the openness of the R-orbit of $k_0 z^{-1}$. Proposition 66.48 allow us to perform this study by means of the surjectivity of $\mathtt{proj}_{\mathcal{H}} \colon \tilde{\mathcal{R}}_z \to \mathcal{Q}$.

We begin by compute the matrices of $\mathcal{R}_z$: zH_1z^{-1}, zH_2z^{-1}, zMz^{-1}, zNz^{-1}, zV_iz^{-1}, zW_jz^{-1}. We make the following change of basis in $\widetilde{\mathcal{R}}$:

$$H_1, H_2 \to \frac{1}{2}(H_1 + H_2), \frac{1}{2}(H_2 - H_1), \tag{79.31a}$$

$$M, N \to \frac{1}{2}(M + N), \frac{1}{2}(M - N). \tag{79.31b}$$

6. Faut que tu vérifies.
7. Quand tu auras refusioné, il faudra référentier ta transcription de Wisser.

The result is that with

$$z = \begin{pmatrix} S & 0 \\ 0 & \mathbb{1} \end{pmatrix} \text{ where } S = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, \tag{79.32}$$

and $c^2 + s^2 = 1$, we have

$$H_{1z} = \begin{pmatrix} & & c & 0 \\ & & -s & 0 \\ c & -s & & \\ 0 & 0 & & \end{pmatrix}, \quad H_{2z} = \begin{pmatrix} & & & s \\ & & & c \\ & & & 0 \\ s & c & 0 & 0 \end{pmatrix}; \tag{79.33}$$

$$M_z = \begin{pmatrix} 0 & 1 & s & 0 \\ -1 & 0 & c & 0 \\ s & c & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N_z = \begin{pmatrix} & & & -c \\ & & & s \\ & & & -1 \\ -c & s & 1 & 0 \end{pmatrix}; \tag{79.34}$$

$$V_{iz} = c(E_{1i} + E_{i1}) - s(E_{2i} + E_{i2}) + E_{3i} - E_{i3}, \tag{79.35a}$$
$$W_{jz} = s(E_{1j} + E_{j1}) + c(E_{2j} + E_{j2}) + E_{4j} - E_{j4}. \tag{79.35b}$$

The matrices we *really* need are the projections of these one with respect to $\mathcal{H}$. We denote it by symbols with a line:

$$\overline{H}_{1z} = \begin{pmatrix} & & c & 0 \\ & & 0 & 0 \\ c & 0 & & \\ 0 & 0 & & \end{pmatrix}, \qquad \overline{H}_{2z} = \begin{pmatrix} & & 0 & s \\ & & 0 & 0 \\ 0 & 0 & & \\ s & 0 & & \end{pmatrix}, \tag{79.36a}$$

$$\overline{M}_z = \begin{pmatrix} 0 & 1 & s & 0 \\ -1 & 0 & 0 & 0 \\ s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \overline{N}_z = \begin{pmatrix} & & & -c \\ & & & 0 \\ & & & 0 \\ -c & 0 & 0 & 0 \end{pmatrix} \tag{79.36b}$$

$$\overline{V}_{iz} = c(E_{1i} + E_{i1}) \qquad \overline{W}_{jz} = s(E_{1j} + E_{j1}). \tag{79.36c}$$

An explicit study shows that $\mathtt{proj}_{\mathcal{H}} \colon \widetilde{\mathcal{R}}_z \to \mathcal{Q}$ is surjective if and only if $c \neq 0$. Then the open R-orbits are the ones of $k_0^{-1} z^{-1}$ with $c \neq 0$:

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c & -s & 0 & 0 \\ s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -s & -c & 0 & 0 \\ c & -s & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{79.37}$$

Consequently, the not open R-orbits are the ones of $[\pm \mathbb{1}_{\mathrm{SO}(2)}]$.

79.7.6 Orbits as homogeneous spaces

A homogeneous space is a space with a transitive homeomorphism group. We consider the orbit $\mathcal{O} = R[z]$ ($z \in \mathrm{SO}(2)$) which is a homogeneous space because the action of G is trivially transitive. Theorem 53.100 assures us that $\mathcal{O}$ can be written as R/S where S is the group which fixes some point in $\mathcal{O}$.

We will firstly work out the homogeneous structure of the open $\widetilde{\mathcal{R}}$-orbits; the matrices of $\widetilde{\mathcal{R}}$ which leave invariant the point $[\mathbb{1}]$ of G/H are the ones of the product (79.22) which satisfy some equations that are almost the same as (79.23):

$$(1 - 2ab)\cosh\xi - (2ab)\sinh\xi = 1 \tag{79.38a}$$
$$(-b - a)\cosh\eta + (b - a)\sinh\eta = 0 \tag{79.38b}$$
$$(1 - 2ab)\sinh\xi - (2ab)\cosh\xi = 0 \tag{79.38c}$$
$$(-b - a)\sinh\eta + (b - a)\cosh\eta = 0. \tag{79.38d}$$

It is easy to see that the solutions are given by $a = b = \xi = 0$ and no constraint on η. The set S is thus given by

$$S \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh\eta & 0 & \sinh\eta \\ 0 & 0 & 1 & 0 \\ 0 & \sinh\eta & 0 & \cosh\eta \end{pmatrix}. \tag{79.39}$$

which corresponds [8] to $e^{t(H_1-H_2)}$ via the change of variable (66.144). Since this is a matrix of H, this leaves $[\mathbb{1}]$ unchanged. As far as the matrices V_i and W_j are concerned, the reasoning of page 3951 still holds.

So the part of $\widetilde{\mathcal{R}}$ which leaves $\mathbb{1}$ unchanged is $\widetilde{\mathcal{R}} \cap H$ (this is not really amazing). Now recall that $\widetilde{\mathcal{R}} = k_0 R k_0^{-1}$. If $\tilde{r} \in \widetilde{\mathcal{R}}$ fixes $\mathbb{1}$, then $k_0^{-1}\tilde{r}k_0 \in R$ fixes k_0^{-1}, and if $r \in R$ fixes k_0^{-1}, then $k_0 r k_0^{-1} \in \widetilde{\mathcal{R}}$ fixes $\mathbb{1}$, so that $k_0 r k_0^{-1} \in \widetilde{R} \cap H$.

Then $r \in R$ fixes $[k_0^{-1}]$ if and only if $r \in k_0^{-1}(\widetilde{\mathcal{R}} \cap H)k_0$. On the other hand, for the closed orbits the stabilizer in R is $R \cap H$: $r \in R$ fixes $[\mathbb{1}]$ if $r[\mathbb{1}] = [\mathbb{1}]$, i.e. $[r] = [\mathbb{1}]$, which needs a $h \in H$ such that $r = h$.

79.8 Causality, light cone and related topics on anti de Sitter

SecCausal

We particularize the general definitions of subsection 79.2.1 to the case of the anti de Sitter space. We consider the l-dimensional [9] anti de Sitter space

$$AdS_l = \frac{\mathrm{SO}(2, l-1)}{\mathrm{SO}(1, l-1)} (\equiv u^2 + t^2 - x_1^2 - \cdots - x_{l-1}^2 = 1). \tag{79.40}$$

eqdefAdS

According to proposition 79.5, we can only consider the identity component of $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$ instead of full groups [10]. The metric that we put on AdS_l is the one induced from the Killing form of $\mathrm{SO}(2, l-1)$ by formula (66.13). This metric has a Minkowskian signature, so that we have natural notions of time-, space- and light-like vectors. From now we denote by G and H the groups $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$.

An other beautiful way to see that the metric on AdS as one and only one time-like direction is the following. The tangent space of AdS at the point $(u, t, x_1, \cdots, x_{l-1})$ is the orthogonal complement (in $\mathbb{R}^{2,l-1}$) of that vector. From the very definition of AdS, the given vector is time-like (its norm is 1), so that it remains one and only one time-like vector in the tangent space.

We write ANK the Iwasawa decomposition of the connected group $\mathrm{SO}_0(2, l-1)$ (see the theorem 51.188). Let $A\bar{N}$ be the θ-conjugate [11] group of AN where θ is the Cartan involution of subsection 66.5.1. We will see that the actions of AN and $A\bar{N}$ have closed and open orbits. The closed ones are denoted by $\mathscr{S}_{AN}$ and $\mathscr{S}_{A\bar{N}}$. The following definition is motivated all previously existing work about BTZ black hole.

Remarque 79.8.
Here, we consider $\mathrm{SO}(1, n)$ as a subgroup of $\mathrm{SO}(2, n)$. Thus the matrices of $\mathrm{SO}(1, n)$ are $(n+2) \times$

8. More precisely, because of the quotient by H, we had derived a necessary characterization of S, not a sufficient one. But we will soon see that in facts this is also a sufficient condition.

9. The symbol n denotes the number of space-like directions of the underlying space of the matricial group $\mathrm{SO}(2, n)$; this space has dimension $n + 2$ while AdS is a quotient by (something like) one time-like direction. In order to avoid confusions, the symbol l denotes the dimension of the AdS space. This is the reason for which we write $\mathrm{SO}(2, n)$ and AdS_l. So equation (79.40) is best written as

$$AdS_l = \frac{\mathrm{SO}(2, n)}{\mathrm{SO}(1, n)}.$$

10. Since we are about to consider Iwasawa decompositions of these groups, actually we *have to* use the identity components.

11. Roughly speaking, it corresponds to different choices in the Iwasawa decomposition of $\mathrm{SO}(2, l-1)$.

$(n+2)$ of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & \boxed{\mathrm{M}} \end{pmatrix}$$

where M is a $(n+1)\times(n+1)$ matrix of the "true" $\mathrm{SO}(1,n)$. From this, one can believe the closeness of $\mathrm{SO}(1,n)$ in $\mathrm{SO}(2,n)$.

In order to get a full definition of the black hole and its structure, we need to define and characterise the notions of light ray and light cone. These notions are of course directly issued from physics of relativity.

Definition 79.9.
*A **light ray** is a geodesic whose tangent vector is everywhere light-like.* lightraycone

The **causal structure** of a general pseudo-Riemannian manifold M is the fact that two points are said to be *causally connected* when there exists a light ray which passes by both points. More precisely, we say that x has a **causal effect** on y if there exists a future oriented time-like path $c\colon [0,1]\to M$ such that $c(0)=x$ and $c(1)=y$.

A light ray trough ϑ is given by a vector of $\mathcal{Q}$ with vanishing norm. So let us study these vectors. Let $E_1 = q_0+q_1$ and k, a general element of $\mathrm{SO}(n)$ which reads $k=e^K$ with $K = a^{ij}(E_{ij}-E_{ji})$, $i,j\geqslant 3$ and $a^{ij}=-a^{ji}$. If we pose $A_j = E_{1j}+E_{j1}$, we have $[K,E_1]=(2a)^{j3}A_j$ and $[K,A_k]=a^{jk}A_j$. Hence,

$$\mathrm{ad}(K)^n E_1 = \left((2a)^n\right)^{k3} A_k,$$

and

$$\begin{aligned}\mathrm{Ad}(k)E_1 = e^{\mathrm{ad}\,K}E_1 &= E_1 + \sum_{n\geqslant 1}\left((2a)^n\right)^{k3}A_k \\ &= E_1 + \sum_{n=0}^{\infty}\left((2a)^n\right)^{k3}A_k - \delta^{j3}A_j \\ &= E_1 - E_{31} - E_{13} + \left(e^{2a}\right)^{j3}A_j \\ &= q_0 + \sum_{j=1}^{l-1} w_j q_j\end{aligned} \tag{79.41}$$

where $w_i = \left(e^{2a}\right)^{i3}$. Under an explicit form, we have

$$\mathrm{Ad}(k)E_1 = \begin{pmatrix} 0 & 1 & w_1 & w_2 & \dots \\ -1 & & & & \\ w_1 & & & & \\ w_2 & & & & \\ \vdots & & & & \end{pmatrix} \tag{79.42}$$

The exponential e^{2a} being an element of $\mathrm{SO}(n)$, the parameters w_i are restricted by the condition $\sum_k w_k^2 = 1$. Remark moreover that *every* matrix of $\mathrm{SO}(2)$ can be written under the form e^{2a} for a good choice of $a\in\mathfrak{so}(2)$. The light cone is therefore given by the set of vectors of the form $(1,w_i)$ with $\|w\|^2=1$. If we consider the metric $diag(+--\cdots)$ on $\mathcal{Q}$ with respect to the basis $\{q_i\}$, we have

$$\|\,\mathrm{Ad}(k)E_1\|^2 = 0.$$

This is coherent with the intuitive notion of light cone. On the one hand *every* light-like vector of $\mathcal{Q}$ reads $\mathrm{Ad}(k)E_1$ for some $k\in\mathrm{SO}(n)$. On the other hand every nilpotent element of $\mathcal{Q}$ is light-like because trace of nilpotent matrix is zero (using Engel's theorem). In definitive, we proved the following:

PropNormZeroEQnil

Proposition 79.10.
When E is any nilpotent element of $\mathcal{Q}$, the set of light-like vectors of $\mathcal{Q}$ is parametrized by $\lambda \operatorname{Ad}(k)E$ with $k \in \mathrm{SO}(n)$ and $\lambda \in \mathbb{R}$. PropToutVectLumQ

CorNormZeroEQnil

Corollary 79.11.
An element of $\mathcal{Q}$ has a vanishing norm if and only if it is nilpotent.

Proof. We know that, when E, is any nilpotent in $\mathcal{Q}$, the set of vanishing norm vectors in $\mathcal{Q}$ are given by $\{\lambda \operatorname{Ad}(k)E\}$, but all these vectors are nilpotent. □

Let us point out the fact that only the first column of the "direction" $k \in \mathrm{SO}(n)$ has an importance in causality issues. So the word "directions" will often be used to refer to the vector w. It is not a particular feature of our particular matrix representation choice. Indeed the element k only appears in the combination $\operatorname{Ad}(k)E$ which is a light-like vector in $\mathcal{Q}$, i.e. $\operatorname{Ad}(k)E = tq'_0 + \sum_i x_i q'_i$ with $t^2 - \sum_i x_i^2 = 0$ for any orthonormal basis $\{q'_i\}$ of $\mathcal{Q}$. As far as causality is concerned, a rescaling $\operatorname{Ad}(k)E$ to $\lambda \operatorname{Ad}(k)E$ has no importance, so one can choice $t = 1$ and find back $\sum_i x_i^2 = 1$. We see that it is a natural feature that the light-like rays are parametrized by unital vectors of $\mathbb{R}^n$.

LemGeodGenreLumiere

Lemma 79.12.
Let E be a nilpotent element in $\mathcal{Q}$, and $\pi : G \to G/H$, the canonical projection. A light ray through $[g] \in AdS_l$ has the form

$$l^k_{[g]}(s) = \pi\left(ge^{-s \operatorname{Ad}(k)E}\right) \tag{79.43}$$

for a certain $k \in K_H = K \cap H = \mathrm{SO}(n)$. lem:AdkEcone

Proof. General theory of symmetric spaces (see [957], pages 230–233, particularly theorem 3.2) proves that a light ray through $\vartheta = [e]$ has the form

$$l(s) = \pi\left(e^{sX}\right).$$

¡¡ Avertissement/question à la lectrice !! 79.13
Il me semble que ce qui est de cette forme, ce sont les géodésiques, et non les rayons de lumière. Relire Kobayashi-Nomizu.

In our context, we have the additional request for the tangent vector to be light-like. Proposition 79.10 thus imposes X to be of the form $\operatorname{Ad}(k)E$. That proves the claim for geodesics trough ϑ.

The fact that $d\tau_g$ is an nondegenerate isometry then extends the result to all points.

□

CorNilLightQ

Corollary 79.14.
If E is nilpotent in $\mathcal{Q}$, then $\{\operatorname{Ad}(k)E\}_{k\in K_H}$ is the set of light-like vectors in $T_{[\vartheta]}AdS_l \simeq \mathcal{Q}$. Therefore

$$\exp_\vartheta(t \operatorname{Ad}(k)E) = \exp(t \operatorname{Ad}(k)E) \cdot \vartheta. \tag{79.44}$$

is the light cone of ϑ in AdS_l. Note that in this equation, the first exp *is the one defined from the AdS_l-connection while the second is the exponential from a Lie algebra to the Lie group. It comes from the fact that in a symmetric space, $\exp_o v = e^z \cdot \vartheta$.*

In order to fix ideas, we will always use the element E_1 as choice of nilpotent element in $\mathcal{Q}$ in order to parametrize light-cone. Since $\mathrm{SO}(2, l-1)$ acts on AdS_l by isometries, the **light cone** at $\pi(g)$ is given by a translation of the one at ϑ:

$$C^+_{\pi(g)} = g \cdot C_\vartheta = \{\pi\left(ge^{t \operatorname{Ad}(k)E_1}\right)\}_{\substack{t\in\mathbb{R}^+ \\ k\in K_H}}. \tag{79.45}$$ eq:defcone

The product being taken at left while the quotient being taken at right, one can fear a problem of well definiteness in this expression. The following proposition shows that all is right.

PropDefConeIndepRepre

Proposition 79.15.
Definition (79.45) is independent of the representative g in the class $\pi(g)$. In other words,

$$\{\mathrm{Ad}(hk)E_1\}_{k\in K_H} = \{\mathrm{Ad}(k)E_1\}_{k\in K_H} \tag{79.46}$$

eq__statdefcone

for all $h \in H$.

Proof. The metric on $\mathcal{Q}$ is the restriction of the Killing form of $\mathcal{G}$ (notice that $\mathcal{Q}$ has no own Killing form for the simple reason that it is not a Lie algebra). From Ad-invariance, we have in particular

$$B\big(\mathrm{Ad}(h)X, \mathrm{Ad}(h)Y\big) = B(X,Y)$$

for all $h \in \mathrm{SO}(1,l-1)$. The point is that reducibility makes $\mathrm{Ad}(h)X \in \mathcal{Q}$ when $X \in \mathcal{Q}$. The element $\mathrm{Ad}(hk)E_1$ in the left hand side of equation (79.46) being zero-normed in $\mathcal{Q}$, it reads $\mathrm{Ad}(k')E_1$ for some $k' \in K_H$. That proves the inclusion in one sense. For the second inclusion, we have to find a $k' \in K_H$ such that $\mathrm{Ad}(hk')E_1 = \mathrm{Ad}(k)E_1$. Existence of such a k' follows from the fact that $\mathrm{Ad}(h^{-1}k)E_1$ is a light-like vector of $\mathcal{Q}$. □

RemGedNonInvarChoix

Remarque 79.16.
Although the *set* of geodesics $\{\pi(ge^{s\,\mathrm{Ad}(k)E_2})\}$ is equal to the *set* of geodesics $\pi(ghe^{s\,\mathrm{Ad}(k)E_1})$, each geodesic are not independent in the choice of the representative g: $\pi(ge^{\mathrm{Ad}(k)E_1}) \neq \pi(ghe^{\mathrm{Ad}(k)E_1})$ in general.

In particular, in the setting of the anti de Sitter black hole, the property "intersect the singularity" for the geodesic $\pi(ge^{s\,\mathrm{Ad}(k)E_1})$ is not invariant under the choice of the representative g in the class $[g]$.

It is also possible to prove result of independence 79.15 with a lot of matricial computations: let us decompose $h = a_h n_h k_h$; the part k_h is just a redefinition of k in equation (79.46), so we forget it. We begin by proving that (79.46) holds whenever $\mathrm{Ad}(h) \in \mathrm{SO}(\mathcal{Q})$. Consider $\mathrm{Ad}(k')E_1 = X \in \mathcal{Q}$. If $\mathrm{Ad}(h) \in \mathrm{SO}(\mathcal{Q})$, then $\mathrm{Ad}(h^{-1}) \in \mathrm{SO}(\mathcal{Q})$ too and we consider $Y = \mathrm{Ad}(h^{-1})X$ which is a vector of norm zero in $\mathcal{Q}$. There exists $\bar{k} \in K_H$ such that $\mathrm{Ad}(\bar{k})E_1 X = Y$. Now,

$$\mathrm{Ad}(h\bar{k}k')E_1 = \mathrm{Ad}(h\bar{k})X = \mathrm{Ad}(h)Y = X. \tag{79.47}$$

In order to prove that $\mathrm{Ad}(a_h) \in \mathrm{SO}(\mathcal{Q})$, we compute

$$\mathrm{ad}(J_1)\begin{pmatrix} 0 & z & w_1 & w_2 & w3 \\ -z & & & & \\ w_1 & & & & \\ w_2 & & & & \\ w_3 & & & & \end{pmatrix} = \mathrm{ad}(J_1)(zq_0 + w_i q_i).$$

In the basis $\{q_0, q_i\}$, we see that

$$\mathrm{ad}(J_1) = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & & & \\ -1 & & & \\ 0 & & & \end{pmatrix} \in \mathfrak{so}(1,3),$$

so $\mathrm{Ad}(J_1) \in \mathrm{SO}(\mathcal{Q})$. On the other hand, a general element of $\mathcal{N}_{\mathcal{H}}$ is

$$A = \begin{pmatrix} \cdot & & & & \\ & \cdot & a & \cdot & v \\ & a & \cdot & -a & \cdot \\ & \cdot & a & \cdot & v \\ & v & \cdot & -v & \cdot \end{pmatrix},$$

and simple computations shows that on $\mathcal{Q}$,

$$\mathrm{ad}(A) = \begin{pmatrix} \cdot & -a & \cdot & -v \\ -a & \cdot & -a & \cdot \\ \cdot & a & \cdot & v \\ -v & \cdot & v & \cdot \end{pmatrix} \in \mathfrak{so}(1,3).$$

79.8.1 Time orientation

A **time orientation** on $\mathcal{Q}$ is the choice of a vector T such that $\langle T, T\rangle > 0$. When such a choice is made, a vector v is **future directed** when $\langle v, T\rangle > 0$. In our case, the choice is the intuitive one: the vector q_0 defines the time orientation on $\mathcal{Q}$ and $v = (v^0, v^1, v^2, v^3)$ is future directed if and only if $v^0 > 0$. So a light-like future directed vector is always –up to a positive multiple– of the form $(1, \overline{v})$ with $\|\overline{v}\| = 1$. For this reason, the set

$$\{t\,\mathrm{Ad}(k)E_1\}_{\substack{t>0\\ k\in \mathrm{SO}(3)}} \tag{79.48}$$

EqTousVecLumTy

is exactly the set of light-like future-directed vectors of $\mathcal{Q}$.

We are now able to define causality as follows. A point $[g] \in AdS_l$ belongs to the **interior region** if for every direction $k \in K_H$, the future light ray $l^k_{[g]}$ intersects the singularity within a *finite* time. In other words, it is interior when the whole light cone ends up in the singularity. A point which is not interior is said to be **exterior**. A particularly important set is the **event horizon**, or simply *horizon*, defined as the boundary of the interior. When a space contains a non trivial causal structure (i.e. when there exists a non empty horizon), we say that the definition of singularity gives rise to a **black hole**. By extension, the term "black hole" often refers to the set of interior points.

79.8.1.1 Singularity

Singular

Definition 79.17.
*The **singularity** in AdS_l is the set*

$$\mathscr{S} = \mathit{singularity} = \mathscr{S}_{AN} \cup \mathscr{S}_{A\bar{N}},$$

*so that a point is **singular** when it belongs to a closed orbit of AN or $A\bar{N}$. The **black hole** is defined as*

$$BH = \{x \in AdS_l \ st\ \forall\ \textit{time-like vector}\ k \in T_x AdS_l,\ l^k_x \cap \mathcal{S} \neq \varnothing\}$$

where l^k_x is the (future directed) geodesic in the direction k starting at x (see equation (79.48) and the discussion above).

The aim of this chapter is to prove that the so-defined black hole is non trivial in the sense that the following inclusions are strict:

$$\mathscr{S} \subsetneq BH \subsetneq AdS_l. \tag{79.49}$$

EqhSsubBH

79.8.2 Action of H and $\mathrm{Ad}(\mathcal{Q})$

Remember that we decree closed orbits to be *singular*. Now the fact for a point $\pi(g) \in AdS_l$ to be *exterior* is that there exists an non empty set $\mathcal{O}$ of K_H such that $\forall k \in \mathcal{O}$,

$$\pi\big(ge^{t\,\mathrm{Ad}(k)E_1}\big) \cap \mathcal{S} = \varnothing.$$

The restriction of the Killing form to $\mathcal{Q}$ reads

$$B(q_0, q_0) = \mathrm{Tr}(q_0 q_0) = -2, \tag{79.50a}$$
$$B(q_i, q_i) = \mathrm{Tr}(q_i, q_i) = 2 \qquad \text{for } i \geqslant 1. \tag{79.50b}$$

So the norm on $\mathcal{Q}$ is $\|X\| = -\frac{1}{2}B(X,X)$. The bi-invariance of the Killing form and the fact that the decomposition $\mathcal{G} = \mathcal{Q} \oplus \mathcal{H}$ is reductive imply $\|\operatorname{Ad}(h)X\| = \|X\|$, hence

$$\operatorname{Ad}(H)|_{\mathcal{Q}} \subset \operatorname{SO}(\mathcal{Q}). \tag{79.51}$$

EqInclAdHSOq

A question is to know the kernel of this inclusion: which $h \in H$ fulfill $\operatorname{Ad}(h)q_i = q_i$ for all i? The equation $Aq_iA^{-1} = q_i$ can be simplified (from a computational point of view) using the relation $A^{-1} = \eta A^t \eta$ which defines $\operatorname{SO}(1,n)$. It is a somewhat long but easy computation to prove that $A = \pm\mathbb{1}$ are the only two solutions in $SO(1,n)$ to the system $A(q_i\eta)A^t = q_i\eta$.

One can go further than inclusion (79.51) and prove the following

Proposition 79.18.
Let $h \in H_0$ seen as a matrix acting on $\mathbb{R}^{1,l-1}$ and let see $\operatorname{Ad}(h)$ as a matrix acting on $\mathcal{Q}$. In this case we have $\operatorname{Ad}(h)_{ij} = h_{ij}$. In particular

$$\operatorname{Ad}(H_0) = \operatorname{SO}_0(\mathcal{Q}) \tag{79.52}$$

where the index zero denotes the identity component.

PropSOADHequal

Proof. We will prove that for each unital vector $X \in \mathcal{Q}$, the element $\operatorname{Ad}(h)X$ is a general element of norm 1 in $\mathcal{Q}$ when h runs over H_0. Explicit matrix computation will show by the way the equality $\operatorname{Ad}(h)_{ij} = h_{ij}$. The general product to be computed is

$$\operatorname{Ad}(h)X = \begin{pmatrix} 1 & 0 \\ 0 & \begin{pmatrix} & \\ & h^{-1} \end{pmatrix} \end{pmatrix} \begin{pmatrix} 0 & -w_0 & w_1 & \cdots \\ w_0 & & & \\ w_1 & & & \\ \vdots & & & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \begin{pmatrix} & \\ & h \end{pmatrix} \end{pmatrix}.$$

But we know that the result is a matrix of $\mathcal{Q}$, so it is sufficient to compute the first line. If we denote by c_i the columns of h, we find

$$\operatorname{Ad}(h)X = \sum_{i=0}^{l-1} (w \cdot c_i) q_i$$

where the dot denotes the inner product of $\mathbb{R}^{1,l-1}$. Since $\{c_i\}$ is a general orthonormal basis of $\mathbb{R}^{1,l-1}$, the latter expression is a general vector of norm 1 in $\mathcal{Q}$. □

79.8.3 Two other characterizations of the singularity

SubSecTwoCharSing

In this short section, we first give a coordinatewise characterization of the singularity (which allows some brute force computations), and then we point out that the vector field J_1^* has vanishing norm on the singularity (see also proposition 78.1). That should make the connection with the quotient construction of the original BTZ black hole. Notice that we do not classify all vectors from which vanishing of the norm define a singularity. The point is that one can make our black hole "causally inextensible" by making a discrete quotient of AdS_l along the integral curves of J_1^*.

Proptcarrycarr

Proposition 79.19.
In term of the embedding of AdS_l in $\mathbb{R}^{2,l-1}$, the closed orbits of $AN \subset \operatorname{SO}(2,l-1)$ are located at $y - t = 0$. Similarly, the closed orbits of $A\bar{N}$ correspond to $y + t = 0$. In other words, the equation

$$t^2 - y^2 = 0 \tag{79.53}$$

tcarrycarr

describes the singularity $\mathscr{S} = \mathscr{S}_{AN} \cup \mathscr{S}_{A\bar{N}}$.

More precisely, a point belongs to a closed orbit of AN if and only if $t - y = 0$ and to a closed orbit of $A\bar{N}$ if and only if $t + y = 0$.

Proof. The different fundamental vector fields of the AN action can be computed with the matricial relation $X^*_{[g]} = -Xg \cdot \vartheta$. For example, in AdS_3,

$$M^*_{[g]} = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} = \begin{pmatrix} -t+y \\ u-x \\ -t+y \\ u-x \end{pmatrix}$$
$$= (y-t)\partial_u + (u-x)\partial_t + (y-t)\partial_x + (u-x)\partial_y.$$

Full results are [Gen]

EqNormeJun
$$J_1^* = -y\partial_t - t\partial_y \tag{79.54a}$$
eq:Ids
$$J_2^* = -x\partial_u - u\partial_x \tag{79.54b}$$
$$M^* = (y-t)\partial_u + (u-x)\partial_t + (y-t)\partial_x + (u-x)\partial_y \tag{79.54c}$$
$$L^* = (y-t)\partial_u + (u+x)\partial_t + (t-y)\partial_x + (u+x)\partial_y \tag{79.54d}$$
$$W_i^* = -x_i\partial_t - x_i\partial_y + (y-t)\partial_i \tag{79.54e}$$
eq:Vis
$$V_j^* = -x_j\partial_u - x_j\partial_x + (x-u)\partial_j, \tag{79.54f}$$

with $i, j = 3, \dots, l-1$. First consider points satisfying $t-y=0$. It is clear that, at these points, the l vectors J_1^*, M^*, L^* and W_i^* only span the direction $\partial_t + \partial_y$. Thus, there are at most $l-1$ linearly independent vectors amongst the $2(l-1)$ vectors (79.54). We conclude that a point satisfying $t-y=0$ belongs to a closed orbit of AN.

Now we show that a point with $t-y \neq 0$ belongs to an open orbit of AN. It is easy to see that J_1^*, M^* and L^* are three linearly independent vectors. The vectors V_i^* gives us $l-3$ more. Then they span a l-dimensional space.

The same can be done with the closed orbits of $A\bar{N}$. We have

$$N^* = -(y+t)\partial_u + (u-x)\partial_t - (y+t)\partial_x + (x-u)\partial_y \tag{79.55a}$$
$$F^* = -(y+t)\partial_u + (x+y)\partial_t + (y+t)\partial_x - (x+u)\partial_y \tag{79.55b}$$
$$X_i^* = -x_i\partial_u + x_i\partial_x - (x+u)\partial_i \tag{79.55c}$$
$$Y_j^* = z\partial_t - z\partial_y + (y+t)\partial_i. \tag{79.55d}$$

When $t+y=0$, the vectors J_1^*, N^*, F^* and Y_j^* only span the direction $\partial_t - \partial_y$. On the other hand, if $t+y \neq 0$, we look at the vectors J_1^*, N^* and F^*. The vector J_1^* is linearly independent of N^* and F^* because is does not contain a ∂_u component. Now, the vector N^* contains a component $\partial_u + \partial_x$ while F^* contains $\partial_u - \partial_x$. We conclude that the vectors J_1^*, N^* and F^* span three linearly independent vectors. Thus a point with $t+y \neq 0$ belongs to an open orbit of $A\bar{N}$.

The result is that a point belongs to a closed orbit of $A\bar{N}$ if and only if $t+y=0$. □

This shows that in the three dimensional case, our black hole reduces to the previously existing one.

The following corollary shows that a discrete quotient of AdS_l along the orbits of J_1^* gives a direct higher-dimensional generalization of the non-rotating BTZ black hole [12].

CorJannsingul

Corollary 79.20.
The singularity coincides with the set of points in AdS_l where $\|J_1^*\|^2 = 0$ for the metric induced from the ambient space $\mathbb{R}^{2,l-1}$.

Proof. The expression (79.54a) shows that the norm of J_1^* is $y^2 - t^2$ which vanishes on the singularity. □

In the three-dimensional case, it was shown in [1015, 1012] that the non-rotating BTZ black hole singularity is precisely given by equation (79.53). Hence, the following is a particular case of theorem 79.3:

12. It was also proved for AdS_2 in proposition 78.1.

Corollary 79.21.
The non-rotating BTZ black hole is a causal symmetric solvable black hole.

79.8.4 A criterion with the tangent spaces

subsec:R_z

Since G acts transitively on G/H, the tangent spaces of G/H at different points are not really independents: it is possible to guess global structure from consideration about tangent spaces. If $\mathcal{O}$ denotes the orbit of $[z]$ under R, we have

$$T_{[z]}\mathcal{O} = \mathrm{Span}\{X^*_{[z]} \text{ tel que } X \in \mathcal{R}\}.$$

¡¡ Avertissement/question à la lectrice !! 79.22
C'est le lemme 66.49 pris en un autre point. Il faut trouver un argument pour voir que c'est correct.

We can work out the structure of the fundamentals vector fields:

$$X^*_{[z]} = \frac{d}{dt}\Big[\pi(e^{-tX}z)\Big]_{t=0} = (d\pi)_z(dR_z)_e\frac{d}{dt}\Big[e^{-tX}\Big]_{t=0} = -(d\pi)_z\underline{X}_z \tag{79.56}$$

where $\underline{X}$ denotes the right invariant vector field of $X \in \mathcal{G}$.

¡¡ Avertissement/question au lecteur !! 79.23
Il y a presque certainement une faute de notation entre le tilde au dessus et celui en dessous.

If τ is the action of G on G/H, we can try to bring the expression of $X^*_{[z]}$ in $T_{[e]}\mathcal{O}$ in the following sense:

$$\begin{aligned}(d\tau_{z^{-1}})_{[z]}X^*_{[z]} &= (d\tau_{z^{-1}})_{\pi(z)}(d\pi)_z\underline{X}_z = d(\tau_{z^{-1}}\circ\pi)_z\underline{X}_z \\ &= \frac{d}{dt}\Big[\pi(z^{-1}e^{-tX}z)\Big]_{t=0} = (d\pi)_e\,\mathrm{Ad}(z^{-1})X.\end{aligned} \tag{79.57}$$

Now we define the space

$$\mathcal{R}_z = (d\pi)_e\,\mathrm{Ad}(z^{-1})\mathcal{R}, \tag{79.58}$$

and we can state a necessary condition for two points to belongs to the same orbit.

Proposition 79.24.
If the elements z and z' of G are related by $r \in R$ (i.e. $z' = rz$), then $\mathcal{R}_{z'} = \mathcal{R}_z$.

Proof. If is just a computation. Let $z' = rz$; we have

$$\mathcal{R}_{z'} = (d\pi)_e\,\mathrm{Ad}(z'^{-1})\mathcal{R} = (d\pi)_e\,\mathrm{Ad}(z^{-1}r^{-1})\mathcal{R} = (d\pi)_e\,\mathrm{Ad}(z^{-1})\,\mathrm{Ad}(r^{-1})\mathcal{R} = \mathcal{R}_z \tag{79.59}$$

because $\mathrm{Ad}(r^{-1})\mathcal{R} = \mathcal{R}$. □

Taking the general form (79.19) of an element in $\mathcal{R}$, we compute

$$\begin{aligned}\mathrm{Ad}(z)\mathcal{R} &= \begin{pmatrix}\cos\mu & \sin\mu & \\ -\sin\mu & \cos\mu & \\ & & \mathbb{1}\end{pmatrix}\begin{pmatrix}0 & x & k & -x\\ -x & 0 & 0 & j\\ k & 0 & 0 & 0\\ -x & j & 0 & 0\end{pmatrix}\begin{pmatrix}\cos\mu & -\sin\mu & \\ \sin\mu & \cos\mu & \\ & & \mathbb{1}\end{pmatrix} \\ &\simeq \begin{pmatrix}0 & x & k\cos\mu & -x\cos\mu + j\sin\mu\\ -x & & & \\ k\cos\mu & & & \\ -c\cos\mu + j\sin\mu & & & \end{pmatrix}\end{aligned} \tag{79.60}$$

where $\simeq$ stand for "equals up to a matrix of $\mathcal{H}$". If $\cos\mu = 0$, then $\mathcal{R}_z \neq \mathcal{R}$. This shows that

$$\begin{pmatrix}0 & 1 & & \\ -1 & 0 & & \\ & & 1 & \\ & & & 1\end{pmatrix} \text{ and } \begin{pmatrix}0 & -1 & & \\ 1 & 0 & & \\ & & 1 & \\ & & & 1\end{pmatrix} \tag{79.61}$$

does not belong to the orbit of ϑ. We also see that $\cos\mu = -1$ is either not in the orbit of ϑ.

79.8.4.1 Search for open orbits

We consider the group $R_z = \mathbf{Ad}(z)R = zRz^{-1}$ for some $z \in \mathrm{SO}(2)$. The openness of the R_z-orbit of ϑ is the same as the one of the R-orbit of $[z^{-1}]$. Matrices of $\mathcal{R}_z = z\mathcal{R}z^{-1}$ are easy to find. Here are the projections on $\mathcal{Q}$:

$$\mathtt{proj}_{\mathcal{Q}} M_z = \begin{pmatrix} 0 & 1 & \sin\mu & -\cos\mu \\ -1 & & & \\ \sin\mu & & & \\ -\cos\mu & & & \end{pmatrix} \qquad \mathtt{proj}_{\mathcal{Q}} L_z = \begin{pmatrix} 0 & 1 & -\sin\mu & -\cos\mu \\ -1 & & & \\ -\sin\mu & & & \\ -\cos\mu & & & \end{pmatrix}$$

(79.62a)

$$\mathtt{proj}_{\mathcal{Q}} J_{1z} = \begin{pmatrix} 0 & 0 & 0 & \sin\mu \\ 0 & & & \\ 0 & & & \\ \sin\mu & & & \end{pmatrix} \qquad \mathtt{proj}_{\mathcal{Q}} J_{2z} = \begin{pmatrix} 0 & 0 & \cos\mu & 0 \\ 0 & & & \\ \cos\mu & & & \\ 0 & & & \end{pmatrix} \tag{79.62b}$$

$$\mathtt{proj}_{\mathcal{Q}} W_{iz} = \sin\mu(E_{i1} + E_{1i}) \qquad \mathtt{proj}_{\mathcal{Q}} V_{iz} = \cos\mu(E_{i1} + E_{1i}) \tag{79.62c}$$

When $\sin\mu$ and $\cos\mu$ are non-zero, we have

$$q_0 = \frac{1}{2}\big(M_z + L_z + \frac{\cos\mu}{\sin\mu} J_{1z}\big) \qquad q_1 = \frac{1}{\cos\mu} J_{2z} \tag{79.63a}$$

$$q_2 = \frac{1}{\sin\mu} J_{1z} \qquad q_i = \frac{1}{\sin\mu} W_{iz} = \frac{1}{\cos\mu} V_{iz} \tag{79.63b}$$

So when $\sin\mu = 0$, the element q_0 does not belong to $\mathtt{proj}_{\mathcal{Q}} \mathcal{R}_z$. Hence the R_z-orbit of ϑ is non open if and only if $\sin\mu = 0$.

79.8.5 Orbits and topology

PgTopoOrb

Let $D^{\pm} = AN\,\mathrm{SO}(n)\,\mathrm{SO}(2)^{\pm}$ where $\mathrm{SO}(2)^{\pm}$ are the subgroups of $\mathrm{SO}(2) \subset \mathrm{SO}(2,n)$ with strictly positive (negative) cosine. We see $\mathrm{SO}(2)$ and $\mathrm{SO}(n)$ as subgroups of $\mathrm{SO}(2,n)$ in the way indicated by equation (79.14). Notice that the parts $\mathrm{SO}(2)$ and $\mathrm{SO}(n)$ are commuting and that $\mathrm{SO}(n) \subset H$. The notation $-\mathbb{1}_{\mathrm{SO}(2)}$ refers to the element of $\mathrm{SO}(2,n)$ which the identity as AN-component and $-\mathbb{1}$ as $\mathrm{SO}(2)$-component.

A continuous path from $[D^+]$ to $[D^-]$ must pass trough an element of the form $[AN\mathbb{1}_{\mathrm{SO}(2)}]$. We saw that the AN-orbit of such an element is not open while the AN-orbit of an element of $[D^+]$ is open. So we deduce that an orbit passing trough $[D^+]$ does not intersect $[D^-]$.

The set $[D^+]$ is connected in G/H and D^+ being open in G, the set $[D^+] = \pi(D^+)$ is also open in G/H from the definition of the topology (see theorem 53.98). Now, the orbits of AN in $[D^+]$ (who are all open) furnish an open partition of $[D^+]$. Such a partition is impossible for an open connected set. We deduce that $[D^+]$ is only one orbit of AN in G/H. The same can be done with $[D^-]$.

We are left with the sets $[AN]$ and $[AN(-\mathbb{1}_{\mathrm{SO}(2)})]$ whose union is closed because we just saw that the complement is open. Now we prove that these two sets are disjoint, in such a way that they have to be separately closed. Existence of an intersection point between $[AN]$ and $[AN(-\mathbb{1}_{\mathrm{SO}(2)})]$ would lead to the existence of a $h \in H$ such that $an\mathbb{1}_{\mathrm{SO}(2)} = (-\mathbb{1}_{\mathrm{SO}(2)})h$, or

$$h = (-\mathbb{1}_{\mathrm{SO}(2)})an,$$

that is a non trivial K-component to h in the decomposition KAN, but the only K-component in H is $\mathrm{SO}(n)$. Hence such a h does not exist and $R[\mathbb{1}] \cap R[-\mathbb{1}_{\mathrm{SO}(2)}] = \varnothing$.

The conclusion is that the Iwasawa group AN has only four orbits:

$$[D^+], \qquad [D^-], \qquad [AN\mathbb{1}_{\mathrm{SO}(2)}], \qquad [AN(-\mathbb{1}_{\mathrm{SO}(2)})]. \tag{79.64}$$

The two first are open and the other two are closed. Remark that an element of $[K]$ does not belong to a closed orbit of AN or $A\bar{N}$.

PgNoticeKpassung

79.8.6 The volume form method

subsecVolumeForm

Let us give an alternative to proposition 66.48 to study the openness of an AN-orbit. We explain the method for $\mathscr{S}_{AN}$, but the same with trivial adaptations is true for $\mathscr{S}_{A\bar{N}}$.

If $x \in M$ belongs to $\mathscr{S}_{AN}$, the tangent space of its AN-orbit has lower dimension than the tangent space of M. In this case the volume spanned by the fundamental vectors at x is zero. The idea is to build the volume form ν_x of T_xM and then apply it on a basis of the fundamental fields. If the result is zero, then x belongs to the $\mathscr{S}_{AN}$. More precisely, the action is given by

$$\tau\colon AN \times M \to M\ (an, [g]) \ \mapsto [ang]. \tag{79.65}$$

If $X \in \mathcal{A} \oplus \mathcal{N}$ and $[g] \in M$, then

$$X^*_{[g]} = -d(\pi \circ R_g)X. \tag{79.66}$$

As mentioned in corollary 66.6, if $\{q_i\}$ is a basis of $\mathcal{Q}$ then a basis of $T_{[g]}M$ is given by $\{d\pi dL_g q_i\}$. We define

$$\nu = q_0^\flat \wedge q_1^\flat \wedge \ldots \wedge q_{l-1}^\flat$$

where $q^\flat_{i[g]} = B_{[g]}(d\pi dL_g q_i, \ \cdot\ )$. The condition for $[g]$ to belongs to $\mathscr{S}_{AN}$ reads

eq:nusurN

$$\nu_{[g]}(N^*_1{}_{[g]}, N^*_2{}_{[g]}, \ldots, N^*_l{}_{[g]}) = 0 \tag{79.67}$$

for every choices of N_j in a basis of $\mathcal{A} \oplus \mathcal{N}$. It corresponds to the vanishing of $l \times l$ determinants. Our purpose is now to compute the products

$$\begin{aligned} B_{[g]}(d\pi dL_g q_i, N^*_j{}_{[h]}) &= -B_g(\mathtt{proj}\, dL_g q_i, \mathtt{proj}\, dR_g N_j) \\ &= -B_g(dL_g q_i, dR_g N_j) \\ &= -B_e(q_i, \mathrm{Ad}(g^{-1})N_j). \end{aligned}$$

where $\mathtt{proj}\colon T_gM \to dL_g\mathcal{Q}$ is the projection. The step from the first to the second line is as follows. First, $\mathtt{proj}\, dL_g q_i = dL_g q_i$ by definition. For the second, let us write $dR_g X = dL_g X_h + dL_g X_q$ with $X_h \in \mathcal{H}$ and $X_q \in \mathcal{Q}$. From equations (66.93), we see that $B(\mathcal{Q}, \mathcal{H}) = 0$, so $B(dL_g q_i, dL_g X_h + dL_g X_q) = B(dL_g q_i, dL_g X_q)$. Remark that one cannot do it computing $\|J_i^*\|$.

We consider the quantity

$$\Delta_{ij}([g]) = B(q_i, \mathrm{Ad}(g^{-1})N_j)$$

where N_j runs over a basis of $\mathcal{A}\oplus\mathcal{N}$ and q_i a one of $\mathcal{Q}$. Our problem of light cone (see explanations in section 79.8) leads us to compute

eq:elemtr

$$\Delta_{ij}(\pi(ge^{-tk\ \cdot\ E})) = B(\mathrm{Ad}(e^{-tk\ \cdot\ E})q_i, \mathrm{Ad}(g^{-1})N_j) \tag{79.68}$$

where $k \cdot E$ is a notation for $\mathrm{Ad}(k)E$.

A way to proceed is, following proposition 53.129, to express all our elements of $\mathfrak{so}(2,n)$ in the root space decomposition

$$\mathcal{G} = \mathcal{G}_{(0,0)} \bigoplus_{\lambda\in\Sigma} \mathcal{G}_\lambda.$$

The purpose of that resides in the fact that the Killing form $B(X,Y)$ is easier to compute when X and Y belongs to some root spaces.

An important computational remark is the fact that E is nilpotent, so $\mathrm{Ad}(k)E$ also is and $\mathrm{Ad}(e^{-t\,\mathrm{Ad}(k)E})X = e^{-t\,\mathrm{ad}(k)E}X$ only gives second order expressions with respect to t. These computations are nevertheless heavy, but can fortunately be circumvented by a simple counting of dimensions, as we describe in proposition 79.19.

Let us make some computations now. In a first time, we restrict ourself to elements in K: we put $g = e^{uR}$ with

$$R = \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \in \mathcal{K}.$$

On the other hand, an useful way to express $k \cdot E_1$ is the following (cf. equations (79.41)):

$$\mathrm{Ad}(k)E_1 = \begin{pmatrix} 0 & 1 & w_1 & w_2 & w_3 \\ -1 & & & & \\ w_1 & & & & \\ w_2 & & & & \\ w_3 & & & & \end{pmatrix} = q_0 + w_1 q_1 + w_2 q_2 + w_3 q_3 \in \mathcal{Q}.$$

It should be noted that by choosing k, all the vectors $\begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix}$ with $\|w\|^2 = 1$ are possible.

Let us begin by systematically computing the elements $[k \cdot E_1, q_i]$ and $[k \cdot E_1, [k \cdot E_1, q_i]]$; the others $\mathrm{ad}(k \cdot E_1)^n q_i$ are zero because one can see that $\mathrm{ad}(E_1)^3 X_\alpha = 0$ for all X_α in the root spaces. All computations can be performed by decomposing the q_i's in the root space basis and using the known commutations relations between root spaces. The way we choose here is to directly use the huge formula

$$\begin{aligned} [k \cdot E_1, [k \cdot E_1]] = {} & w_1^2 [q_1, [q_1, q_i]] \\ & + w_1 w_2 \big([q_1, [q_2, q_i]] + [q_2, [q_1, q_i]] \big) \\ & + w_1 w_3 \big([q_1, [q_3, q_i]] + [q_3, [q_1, q_i]] \big) \\ & + w_2^2 [q_2, [q_2, q_i]] \\ & + w_2 w_3 \big([q_2, [q_3, q_i]] + [q_3, [q_2, q_i]] \big) \\ & + w_3^2 [q_3, [q_3, q_i]] \end{aligned} \tag{79.69}$$

and use the commutations relations between the q_i's. The results are

$$\begin{aligned} \mathrm{ad}(k \cdot E_1) q_0 &= \frac{w_1}{4}(M + N - L - F) + w_2 J_1 + \frac{w_3}{2}(W - Y) \\ \mathrm{ad}(k \cdot E_1) q_1 &= \frac{1}{4}(L + F - M - N) + \frac{w_2}{4}(F + M - L - M) - \frac{w_3}{2}(V + X) \\ \mathrm{ad}(k \cdot E_1) q_2 &= J_1 + \frac{w_1}{4}(L + N - F - M) - \frac{w_3}{2}(W + Y) \\ \mathrm{ad}(k \cdot E_1) q_3 &= \frac{1}{2}(Y - W) + \frac{w_1}{2}(V + X) + \frac{w_2}{2}(W + Y) \end{aligned} \tag{79.70}$$

and

$$\begin{aligned} \mathrm{ad}(k \cdot E_1)^2 q_0 &= k \cdot E_1 \\ \mathrm{ad}(k \cdot E_1)^2 q_1 &= -w_1 q_0 + (w_2^2 + w_3^2 - 1) q_1 - w_1 w_2 q_2 - w_1 w_3 q_3 \\ \mathrm{ad}(k \cdot E_1)^2 q_2 &= -w_2 q_0 - w_1 w_2 q_1 + (w_1^2 + w_3^2 - 1) q_2 - w_2 w_3 q_3 \\ \mathrm{ad}(k \cdot E_1)^2 q_3 &= -w_3 q_0 - w_1 w_3 q_1 - w_2 w_3 q_2 + (w_1^2 + w_2^2 - 1) q_3 \end{aligned} \tag{79.71}$$

It is rather easy to check that $\mathrm{ad}(k \cdot E_1)^3 q_i = 0$ by virtue of $\|w\|^2 = 0$. All these expressions have to be extended in the basis of the root spaces.

$$\begin{aligned} \mathrm{ad}(k \cdot E_1)^2 q_0 = {} & \frac{1}{4}(M + N + L + F) + \frac{w_2}{4}(N + F - M - L) \\ & + \frac{w_3}{2}(V - X) + w_1 q_1 \\ \mathrm{ad}(k \cdot E_1)^2 q_1 = {} & -\frac{w_1}{4}(M + N + L + F) + \frac{w_1 w_2}{4}(M + L - N - F) \\ & + \frac{w_1 w_3}{2}(X - V) + (w_2^2 + w_3^2 - 1) q_1 \\ \mathrm{ad}(k \cdot E_1)^2 q_2 = {} & -\frac{w_2}{4}(M + N + L + F) + \frac{w_1^2 + w_3^2 - 1}{4}(N + F - M - L) \\ & + \frac{w_2 w_3}{2}(X - V) - w_1 w_2 q_1 \\ \mathrm{ad}(k \cdot E_1)^2 q_3 = {} & -\frac{w_3}{4}(M + N + L + F) + \frac{w_2 w_3}{4}(M + L - N - F) \\ & + \frac{w_1^2 + w_2^2 - 1}{2}(V - X) - w_1 w_3 q_1 \end{aligned} \tag{79.72}$$

79.8.6.1 The column of V

An explicit computation shows that

$$\begin{aligned}\mathrm{Ad}(e^{uR})V &= \begin{pmatrix} & & & & \cos u \\ & & & & -\sin u \\ & & & & 1 \\ & & & & 0 \\ \cos u & -\sin u & -1 & 0 & 0 \end{pmatrix} \\ &= \frac{1}{2}(1-\cos u)X + \frac{1}{2}(\sin u)Y \\ &\quad + \frac{1}{2}(1+\cos u)V - \frac{1}{2}(\sin u)W.\end{aligned} \tag{79.73}$$

Remarque 79.25.
Because of the invert in (79.68), we are looking at the destiny of the point $[e^{-uR}]$, not the one of $[e^{uR}]$.

Thanks to the properties of the root space decomposition, we know that the only non zero Killing form containing X, Y, V, W are $B(W, Y)$ and $B(V, X)$. So in the expression

$$\mathrm{Ad}(k \cdot E_1)q_0 = q_0 + \frac{tw_1}{4}(N+M+L+F) + tw_2J_1 + \frac{tw_3}{2}(W-Y) + \frac{t^2w_3}{2}(V-X),$$

we can forget the three first terms when we compute $\Delta_{q_0,V}$. The result is

$$\boxed{\Delta_{q_0,V} = B(W,Y)\frac{tw_3}{2}\sin u - B(V,X)\frac{t^2w_3}{4}\cos u} \tag{79.74}$$

In the same way,

$$\boxed{\Delta_{q_1,V} = -B(V,X)\left(\frac{tw_3}{2} + \frac{t^2w_2w_3}{4}\right)}, \tag{79.75}$$

$$\boxed{\Delta_{q_2,V} = B(X,V)\frac{t^2w_2w_3}{4}\cos u}, \tag{79.76}$$

$$\boxed{\Delta_{q_3,V} = -B(V,X)\frac{1}{2}\big(\cos u - tw_1 + \frac{t^2}{2}(w_1^2+w_2^2-1)\cos u\big) - B(W,Y)\frac{t}{2}\sin u} \tag{79.77}$$

Remark that the only term in this column which doesn't vanishes when $t = 0$ contains $\cos u$.

79.8.6.2 The column of J_1

¡¡ Avertissement/question à la lectrice !! 79.26
C'est justement un de ceux que tu soupçonnes de ne servir à rien.

A direct computation shows that

$$\begin{aligned}\mathrm{Ad}(e^{uR})J_1 &= \sin(u)q_2 + \cos(u)J_1 \\ &= \frac{1}{4}\sin(u)(N+F-M-L) + \cos(u)J_1.\end{aligned} \tag{79.78}$$

We only have to consider the non zero Killing form $B(J_1, J_1)$, $B(W, Y)$, $B(V, X)$, $B(N, L)$, $B(M, F)$.

$$\boxed{\Delta_{q_0,J_1} = 6t^2w_2\sin u + 6tw_2\cos u} \tag{79.79}$$

79.8.6.3 The column of J_2

For the computation of $\mathrm{Ad}(e^{uR})J_2$, we recall that $R = q_0$ and $J_2 = q_1$. It is easy to see that $[q_0, q_1] = \frac{1}{4}(L + F - M - N)$ and $[q_0, [q_0, q_1]] = -q_1$, so that the exponential series looks good and gives

$$\mathrm{Ad}(e^{uR})q_1 = \cos(u)q_1 + \frac{\sin u}{4}(L + F - M - N).$$

A lot of computation gives

$$\boxed{\Delta_{q_0,J_2} = 3t^2 w_1 \cos u - 6t w_1 \sin u} \tag{79.80}$$

$$\boxed{\Delta_{q_1,J_2} = -3t^2 w_1 \cos u + 6t \sin u + 6\cos u} \tag{79.81}$$

$$\boxed{\Delta_{q_2,J_2} = -3t^2 w_1 w_2 \cos u} \tag{79.82}$$

$$\boxed{\Delta_{q_3,J_2} = -3t^2 w_1 w_3 \cos u} \tag{79.83}$$

79.8.6.4 The column of M

The first computation is

$$\mathrm{Ad}(e^{uR})M = \frac{1 - \cos u}{2}(F - M) + \sin(u)(q_1 + J_1) + M.$$

$$\boxed{\Delta_{q_0,M} = 6(1 + w_1 \sin u) + 6t(w_1(1 - \cos u) + w_2 \sin u) + 3t^2(1 + w_2 \cos u).} \tag{79.84}$$

$$\begin{aligned}
\Delta_{q_1,M} = B\Big(& q_1 + \frac{t}{4}(F - M) + \frac{t w_2}{4}(F - M) && (79.85)\\
&+ \frac{t^2}{2}\big[- \frac{w_1}{4}(F + M) + \frac{w_1 w_2}{4}(M - F) - w_1^2 q_1\big], && (79.86)\\
&\frac{1}{2}(1 - \cos u)(F - M) + \sin u(q_1 + J_1) + M\Big). && (79.87)
\end{aligned}$$

Collecting the terms and using the following relations,

$$B(M + F, F - M) = 0 \qquad B(M - F, F - M) = 3 \cdot 16 \tag{79.88a}$$

$$B(F - M, M) = 3 \cdot 8 \qquad B(F + M, M) = 3 \cdot 8. \tag{79.88b}$$

we find

$$\boxed{\Delta_{q_1,M} = 6\sin u - 6t(2 - \cos u)(1 + w_2) + 3t^2(w_1 + w_1 w_2 \cos u - w_2^2 \sin u)} \tag{79.89}$$

$$\boxed{\begin{aligned}\Delta_{q_2,M} = &-6(2 - \cos u) + 6t(\sin u + w_1)\\ &+ 3t^2\big(- w_2 + \frac{w_2^2}{2}(1 - \cos u) - w_1 w_2 \sin u\big)\end{aligned}} \tag{79.90}$$

$$\boxed{\Delta_{q_3,M} = 3t^2 w_3(1 + w_2(2 - \cos u) - w_1 \sin u).} \tag{79.91}$$

79.8.6.5 Existence for AdS_3

From computer computations, the (non identically zero) volume determinants are given by

$$-32(tw_2 + t\cos u - \sin u)^2\big(\cos u + t(\sin u - w_1)\big) \tag{79.92a}$$

$$-32(tw_2 + t\cos u - \sin u)^9 \tag{79.92b}$$

$$\begin{aligned}16t^2w_3\Big(&w_2(w_1 - \sin u) + \cos u(w_1 - \sin u) - w_2\cos u\\ &-\cos^2 u + \sin u(-w_1 + \sin u)\Big) + 16tw_3\cos u\sin u\end{aligned} \tag{79.92c}$$

$$16tw_3\big(-\cos u + t(w_1 - \sin u)\big)(tw_2 + t\cos u - \sin u) \tag{79.92d}$$

One can deduce the existence of an horizon. Indeed the vanishing of all the determinants for a point in [SO(2)] with respect to the AN singularity only requires [eq:annul_trois]

$$t_{AN} = \frac{\sin u}{\cos u - \sin k} \tag{79.93a}$$

while the same for $A\overline{N}$ requires

$$t_{A\overline{N}} = \frac{\sin u}{\sin k + \cos u} \tag{79.93b}$$

The (class of the) point u belongs to the black hole if for all $k \in \mathrm{SO}(2)$, $t_{AN} > 0$ or $t_{A\overline{N}} > 0$. In this case, all light-like geodesic from the point u fall into the hole after a positive time. There are two possibilities:

$$\begin{aligned}\sin u &< 0\\ \cos u &< 0\end{aligned} \tag{79.94a}$$

or

$$\begin{aligned}\sin u &> 0\\ \cos u &> 0\end{aligned} \tag{79.94b}$$

Let us insist to the fact that the points $u = 0$ and $u = \pi$ are not in the horizon although they separate black points and free points. These two points belongs to the singularity. In fact the spaces $\sin u \geqslant 0$ and $\sin u \leqslant 0$ are two completely separated spaces.

So in the space $\sin u \geqslant 0$, the point $u = \pi/2$ is part of the horizon. This proves the existence of an horizon and gives one point of it. The determination of the horizon is not likely easy.

79.8.6.6 Existence for AdS_4

One can parametrize $\mathrm{Ad}(k)E_1$ as

$$\mathrm{Ad}(k)E_1 = \begin{pmatrix} 0 & 1 & w_1 & w_2 & w_3\\ -1 & & & & \\ w_1 & & & & \\ w_2 & & & & \\ w_3 & & & & \end{pmatrix}. \tag{79.95}$$

The volume forms for the AN and $A\overline{N}$ orbits are respectively annihilated by

$$t_{AN} = \frac{\sin u}{\cos u + w_2}, \text{ and } t_{A\overline{N}} = \frac{\sin u}{\cos u - w_2}. \tag{79.96}$$

These are the same as (79.93). Once again the doomed part of the space is given by

$$\begin{aligned}\sin u &< 0\\ \cos u &< 0\end{aligned} \tag{79.97a}$$

or

$$\begin{aligned}\sin u &> 0\\ \cos u &> 0\end{aligned} \tag{79.97b}$$ eq:possdeux

For example in the case (79.97b), the directions with $\cos u < w_2 < -\cos u$ escape the singularity.

79.9 Existence of a non trivial horizon

SecExistenceHor

We are now able to prove that definition 79.17 provides a non empty horizon satisfying condition (79.49). First we consider points of the form $\mathrm{SO}(2)\cdot\vartheta$, which are parametrized by an angle μ. By lemma 79.12, up to the choice of this parametrization, a light-like geodesic trough μ is given by

$$K\cdot \mathrm{e}^{-s\,\mathrm{Ad}(k)E_1}\cdot\vartheta \tag{79.98}$$

with $k\in\mathrm{SO}(l-1)$ and $s\in\mathbb{R}$. Using the isomorphism $[g]\mapsto g\cdot\vartheta$ between G/H and AdS_l, we find

$$l^k_{[u]}(s)=\pi\big(ue^{s\,\mathrm{Ad}(k)E_1}\big)=\begin{pmatrix}\cos\mu & \sin\mu & & & & \\ -\sin\mu & \cos\mu & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & \ddots\end{pmatrix}e^{s\,\mathrm{Ad}(k)E_1}\begin{pmatrix}1\\0\\0\\0\\0\\\vdots\end{pmatrix}=\begin{pmatrix}u_k(s)\\t_k(s)\\x_k(s)\\y_k(s)\\z_k(s)\\\vdots\end{pmatrix} \tag{79.99}$$ EqIotaInCondHOrExpl

According to proposition 79.19, this geodesic reaches the singularity in the future if $t_k(s)^2-y_k(s)^2=0$ for a certain positive s. Since $\mathrm{Ad}(k)E_1$ is nilpotent, the computation of $e^{s\,\mathrm{Ad}(k)E_1}$ is simple and we only need the first column because it only acts on the first basis vector. A short computation shows that

$$l^k_{[\mu]}(s)=\begin{pmatrix}\cos\mu-s\sin\mu\\-\sin\mu-s\cos\mu\\sw_1\\sw_2\\\vdots\end{pmatrix}. \tag{79.100}$$ EqGedComp

We used the computation

$$e^{s\,\mathrm{Ad}(k)E_1}=\mathbb{1}+s\begin{pmatrix}0&1&w_1&w_2&w_3&\cdots\\-1&&&&&\\w_1&&&&&\\w_2&&&&&\\w_3&&&&&\\\vdots&&&&&\end{pmatrix}+\frac{s^2}{2}\begin{pmatrix}0&0&0&0&0&\cdots\\0&-1&-w_1&-w_2&-w_3&\cdots\\0&w_1&w_1w_1&w_1w_2&w_1w_3&\cdots\\0&w_2&w_2w_1&w_2w_2&w_2w_3&\cdots\\\vdots&\vdots&\vdots&\vdots&\vdots&\end{pmatrix}+\cdots$$

Notice that the sum if finite because E_1 is nilpotent. However, the first power of E_1 which vanishes depends on the dimension.

We conclude that the geodesic reaches $\mathscr{S}_{AN}$ and $\mathscr{S}_{A\bar{N}}$ for values s_{AN} and $s_{A\bar{N}}$ of the affine parameter, given by

$$s_{AN}=\frac{\sin\mu}{\cos\mu-w_2}\qquad\qquad s_{A\bar{N}}=\frac{\sin\mu}{\cos\mu+w_2} \tag{79.101}$$ eq:tempssingul

where w_2 is the second component of the first column of k, see equation (79.42); in particular $-1\leqslant w_2\leqslant 1$.

Since the part $\sin\mu=0$ is precisely $\mathscr{S}_{AN}$, we may restrict ourselves to the open connected domain of AdS_l given by $\sin\mu>0$. More precisely, $\sin\mu=0$ is the equation of $\mathscr{S}_{AN}$ in the ANK decomposition. In the same way, $\mathscr{S}_{A\bar{N}}$ is given by $\sin\mu'=0$ in the $A\bar{N}K$ decomposition. In order

to escape the singularity, the point $[\mu]$ needs both s_{AN} and $s_{A\bar{N}}$ to be strictly positive. It is only possible to find directions (i.e. a parameter w_2) which respects this condition when $\cos\mu > 0$. So the point

$$u \equiv \cos\mu = 0 \tag{79.102}$$

EqUnPtHoriz

is one point of the horizon. Theorem 79.3 is now proved. Remark that the two-dimensional case here appears as degenerate. Therefore, it is treated later in section 78.4, where we show that *no black hole arises from this construction in AdS_2*.

The following proposition contains some physical intuition about the nature of the horizon.

Proposition 79.27.
A light-like geodesic which escapes the singularity (i.e. which does not intersect $\mathscr{S}$) and which passes trough a point of the horizon is contained in the horizon.

Proof. Let $x = [g]$ be a point of the horizon and $\pi(ge^{tAd(k)E_1})$, a light-like geodesic escaping the singularity. Near from x, there exists a point $y = [g']$ in the black hole. From definition of a black hole, for all $k \in \mathrm{SO}(3)$ and $t_0 \in \mathbb{R}^+$, points of the form $\pi(g'e^{t_0 Ad(k)E_1})$ also belong to the black hole. From continuity, in each neighbourhood of $\pi(ge^{t_0 Ad(k)E_1})$, there is such a $\pi(g'e^{t_0 Ad(k)E_1})$. This proves that $\pi(ge^{t_0 Ad(k)E_1})$ belongs to the closure of the black hole. But it is not in the interior of the black hole because (by assumption) the given geodesic escapes the singularity, so every point of the form $\pi\big(ge^{t_0\,\mathrm{Ad}(k)E_1}\big)$ belongs to the horizon. □

PropTNFerme

Proposition 79.28.
The set $BH_l\backslash\mathscr{S}_l$ is open.

Proof. A point $v \in AdS_l$ belongs to $BH_l\backslash\mathscr{S}_l$ if and only if all the solutions in s of the equation

$$(T \pm Y)\big(ve^{s\,\mathrm{Ad}(k)E_1}\big) \in \mathscr{S}_l \tag{79.103}$$

are strictly positive (and non infinite). The *strict* is due to the fact that we excluded $\mathscr{S}_l$ itself. Let $s_{\pm(v,k)}$ be these solutions for the point $v \in AdS_l$ and the direction $k \in S^l$. Let now consider $v_0 \in BH_l\backslash\mathscr{S}_l$. The function $s_\pm(v_0,.)\colon S^l \to \mathbb{R}$ is a continuous function on the compact set S^l, thus its image is a compact subset of $\mathbb{R}_0^+$, because the function reach its extrema.

The function $v \mapsto s_\pm(v,k)$ is also continuous, so that, if ϵ is small enough, and if $v \in B(v_0,\epsilon)$, the image of $s_\pm(v,.)$ is still a compact subset of $\mathbb{R}_0^+$. That means that, from the point v, every light-like geodesic intersect the singularity within a finite strictly positive time, this is the fact that $v \in BH_l\backslash\mathscr{S}_l$. □

CorTNFermeHorEchape

Corollary 79.29.
The set of free points in AdS_l is closed and the points on the horizon do have at least one direction which escape the singularity.

Let us consider the point of the horizon that we know (the one given by (79.102)), and see how can that point hope to escape the singularity. Equations (79.101) which give the time needed to fall into the singularity become

$$t_{AN} = \frac{1}{w_2} \qquad\qquad t_{A\bar{N}} = -\frac{1}{w_2}. \tag{79.104}$$

So for every $w_2 \neq 0$, this point reaches the singularity within a finite time. Taking the direction $w_2 = 0$ the point is able to reject his fall to infinity. This agrees to physical intuition which is that the horizon corresponds to points that fall into the singularity within an infinite time.

Up to a reparametrization of $\mathrm{SO}(n)$, the safe directions are given by (equation (79.42) with $w_2 = 0$)

$$\mathrm{Ad}(k)E_1 = \begin{pmatrix} 0 & 1 & \cos a & 0 & \sin a \\ -1 & & & & \\ \cos a & & & & \\ 0 & & & & \\ \sin a & & & & \end{pmatrix}.$$

A direct computation of equation (79.100) shows that the points of the horizon that are joined by this way are given by $\begin{pmatrix} -1 \\ 0 \\ \cos a \\ 0 \\ \sin a \end{pmatrix}$.

79.10 Characterization by angles in $SO(l-1)$

Let $D[g]$ be the set of light-like directions (vectors in $\mathrm{SO}(n)$) for which the point $[g]$ falls into $\mathscr{S}_{AN}$. Similarly, the set $\overline{D}[g]$ is the one of directions which fall into $\mathscr{S}_{A\bar{N}}$. One can express $\overline{D}$ in terms of D:

$$\begin{aligned}\overline{D}[g] &= \{k \in \mathrm{SO}(n) \text{ tel que } \exists t \text{ for which } \pi\big(ge^{t\,\mathrm{Ad}(k)E_1}\big) \in \mathscr{S}_{A\bar{N}}\} \\ &= \{k \in \mathrm{SO}(n) \text{ tel que } \exists t \text{ for which } \pi\big(\theta(g)\theta(e^{tAd(k)E_1})\big) \in \mathscr{S}_{AN}\} \\ &= \{k \in \mathrm{SO}(n) \text{ tel que } \pi(k) \in D\big(\theta[g]\big)\} \\ &= \{k \in \mathrm{SO}(n) \text{ tel que } k \in \big(D(\theta[g])\big)_\theta\},\end{aligned}$$

So

$$\overline{D}[g] = (D\theta[g])_\theta \tag{79.105}$$ eq:DbarD

where by definition, $k_\theta = Jk$ with J being defined by $\theta = \mathrm{Ad}(J)$ (θ is the Cartan involution). It is easy to see that θ changes the sign of the spacial part of k, i.e. changes $w_i \to -w_i$.

¡¡ Avertissement/question au lecteur !! 79.30
C'est la même chose qu'un autre problème que de voir l'involution de Cartan comme un automorphisme interne. propCrtadeux

A main property of k_θ is

$$\theta(\mathrm{Ad}(k)E_1) = \mathrm{Ad}(k_\theta)E_1.$$

Since k_θ only appears in the expression $\mathrm{Ad}(k)E_1$, that property is actually a sufficient characterization of k_θ for our purpose. In particular, $k_{\theta\theta} \neq k$, but $\mathrm{Ad}(k_{\theta\theta})E_1 = \mathrm{Ad}(k)E_1$.

How to express the condition $g \in \mathscr{H}$ in terms of $D[g]$? The condition to belong to the black hole is $D[g] \cup \overline{D}[g] = \mathrm{SO}(n)$. If the complementary of $D[g] \cup \overline{D}[g]$ has an interior (i.e. if it contains an open subset), then by continuity the complementary $D[g'] \cup \overline{D}[g']$ has also an interior for all $[g']$ near from $[g]$. In this case, $[g]$ cannot belong to the horizon. So a characterization of $\mathscr{H}$ is the fact that the boundary of $D[g]$ and $\overline{D}[g]$ coincide. Equation (79.105) expresses this condition under the form

$$\mathrm{Fr}\, D[g] = \mathrm{Fr}\,\big(D(\theta[g])\big)_\theta, \tag{79.106}$$

from which one immediately deduces that $\mathscr{H}$ is θ-invariant.

We have an expression of $D[\mu]$ for $\mu \in \mathrm{SO}(2)$ by examining equations (79.101). The set $D[\mu]$ is the set of $w_2 \in [-1, 1]$ such that $\cos\mu + w_2 > 0$:

$$D[\mu] =]-\cos\mu, 1[. \tag{79.107}$$

So in order for μ to belong to $\mathscr{H}$, the point $[\mu]$ must satisfy

$$\overline{D}[\mu] = D[\theta\mu]_\theta =]-1, -\cos\mu[.$$

Consequently, if μ' is the K-component of $\theta\mu$ in the ANK decomposition, we impose $]-\cos\mu', 1[= D[\theta\mu] \overset{!}{=}]-\cos\mu', 1[$, and we can describe the horizon by

$$\cos\mu = -\cos\mu' \tag{79.108}$$ eq:caractcos

where μ' is the K-component of μ in the $A\bar{N}K$ decomposition.

79.10.1 Another (useless) characterisation

A way to express our characterization (79.108) is $ank = a'\overline{n}k'$ with $k' = e^{i\pi}k^{-1}$. We know[13] that $NA\bar{N}$ is dense in G. Let $k_0 \in \mathrm{SO}(2)$ and $m = k_0^2 e^{i\pi}$. We define $n, n' \in N$, $a \in A$ such that $m = n^{-1}a\theta(n')$.[14]. For this n, the point $[k_0 n]$ belongs to the horizon because

$$nk_0 = a\theta(n')m^{-1}k_0 = a\theta(n')e^{-i\pi}k_0^{-1}. \tag{79.109}$$

Then this nk reads in decomposition $A\bar{N}K$ with $k' = e^{-i\pi}k^{-1}$. Then (almost) all element in $\mathrm{SO}(2)$ give rise to an element in $\mathscr{H}$.

79.11 Characterisation as orbit of group (by the equation)

SecHOrOrbEquation

This section proves that, if we embed AdS_3 in AdS_4, one can express the horizon in AdS_4 as the result of the action of a one dimensional group on the horizon of AdS_3 (seen in AdS_4), theorem 79.49.

79.11.1 The old three dimensional case

As mentioned in [1017], the singularity of the three dimensional black hole in AdS_3 (seen as the group $\mathrm{SL}(2,\mathbb{R})$) accepts a nice description as lateral classes of AN and $A\bar{N}$. That description is recalled in the proposition 78.14. We want here to provide a similar description for the dimensional generalization $AdS_l = \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1)$.

Let us first make a simple remark. A lateral class in the description of proposition 78.14 is not guaranteed to be a lateral class in the description $AdS = G/H$. Moreover the "AN" of equation (78.86) is not the "AN" of $\mathrm{SO}(2,2)$, but the one of $\mathrm{SL}(2,\mathbb{R})$. The results from the description $AdS_3 = \mathrm{SL}(2,\mathbb{R})$ cannot be that simply translated into results in the description of $AdS_3 = \mathrm{SO}(2,2)/\mathrm{SO}(1,2)$.

Let us begin by finding a group description of the horizon in AdS_3 in the description $AdS_3 = \mathrm{SO}(2,2)/SO(1,2)$. The matricial expression of ANJ in $AdS_3 = \mathrm{SL}(2,\mathbb{R})$ is

$$\begin{pmatrix} e^a & le^a \\ 0 & e^{-a} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -le^a & e^a \\ -e^{-a} & 0 \end{pmatrix} \tag{79.110}$$

EqProSLJANexp

The part of the hyperboloid described by these matrices is obtained by equating (79.110) with the matrix

$$\begin{pmatrix} u+x & y+t \\ y-t & u-x \end{pmatrix}. \tag{79.111}$$

EqIdentMatriSLAdS

The result is the vectors of the form

$$\psi\big(Z(G)ANJ\big) \rightsquigarrow \begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} = \pm \begin{pmatrix} -\frac{1}{2}e^a l \\ \cosh(a) \\ -\frac{1}{2}e^a l \\ \sinh(a) \end{pmatrix} = \pm \begin{pmatrix} \alpha \\ \cosh(a) \\ \alpha \\ \sinh(a) \end{pmatrix} = \pm r_{AN} \tag{79.112}$$

EqVectoPotementSingAN

with α, $a \in \mathbb{R}$. This is a (almost[15]) general vector of AdS_3 with $u^2 - x^2 = 0$, which is coherent with the description (78.84).

The same computation, using (53.325), shows that the other part of the horizon in AdS_3 is given by

$$\psi\big(Z(G)A\bar{N}J\big) = \pm\psi \begin{pmatrix} 0 & e^a \\ -e^{-a} & le^{-a} \end{pmatrix} \rightsquigarrow \begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} = \pm \begin{pmatrix} \frac{1}{2}e^{-a} l \\ \cosh(a) \\ -\frac{1}{2}e^{-a} l \\ \sinh(a) \end{pmatrix} = \pm \begin{pmatrix} \alpha \\ \cosh(a) \\ -\alpha \\ \sinh(a) \end{pmatrix} = \pm r_{A\bar{N}} \tag{79.113}$$

EqVectoPotementSingANbar

13. Mais faudra lire Helgason hein.
14. Il faudra voir si le coup de la densité fait quelque chose dans cette histoire
15. We did not compute the $A\bar{N}$ part of the horizon in $\mathrm{SL}(2,\mathbb{R})$.

where a and α are running over $\mathbb{R}$.

From the equations (79.112) and (79.113), we are able to express the horizon in AdS_3 as union of lateral classes of the element

$$b = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{79.114}$$

because $G_{X_{(-1,1)},J_1} \cdot b = G_{X_{(1,1)},J_1} \cdot b$ and $G_{J_1,X_{(1,-1)},J_1} \cdot b = G_{J_1,X_{(-1,-1)},J_1} \cdot b$. We can express the horizon $\mathscr{H}_3$ in the following way:

$$\begin{aligned} \mathscr{H}_3 &= \pm G_{X_{(-1,1)},J_1} \cdot b \cup \pm G_{X_{(1,-1)},J_1} \cdot b \\ &= \pm G_{\{J_1,X_{(1,1)}\}} \cdot b \cup \pm G_{\{J_1,X_{(-1,-1)}\}} \cdot b, \end{aligned} \tag{79.115}$$

and the two other combinations. Here, $G_{X,Y}$ is the group generated by X and Y.

79.11.2 Characterization by induction on the dimension

From a computational point of view, it reveals to be more or less impossible to directly check that (79.112) belongs to the singularity using the method of equation (79.99), not even in dimension 4. Here is the strategy to compute the horizon in higher dimension:

(1) The map $\psi\colon \mathrm{SL}(2,\mathbb{R}) \to AdS_3$ given by (79.111) is an isometry which maps the singularity into the singularity. Thus it has to map the horizon to the horizon. If $\mathscr{H}_{\mathrm{SL}(2,\mathbb{R})}$ denotes the horizon in $\mathrm{SL}(2,\mathbb{R})$, then the set $\psi\big(\mathscr{H}_{\mathrm{SL}(2,\mathbb{R})}\big)$ is the horizon in $AdS_3 = \mathrm{SO}(2,2)/SO(2,1)$.

(2) We consider the inclusion $\iota\colon \mathrm{SO}(2,n) \to \mathrm{SO}(2,n+1)$ given by $g \mapsto \begin{pmatrix} g & 0 \\ 0 & 1 \end{pmatrix}$ and its differential $d\iota\colon \mathfrak{so}(2,n) \to \mathfrak{so}(2,n+1)$, $X \mapsto \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}$. Now, we are going to build the horizons of AdS_l by induction over l, starting on $l = 3$.

We denote by $\mathscr{H}_l$ and $\mathscr{S}_l$ the horizon and the singularity in AdS_l. The structure of the algebras (equations (66.124) and (66.125)) show immediately that

$$(\mathcal{A} \oplus \mathcal{N})_{\mathfrak{so}(2,n+1)} = \mathrm{Span}\left\{ d\iota(\mathcal{A} \oplus \mathcal{N})_{\mathfrak{so}(2,n)}, V_{n+2}, W_{n+2} \right\}, \tag{79.116}$$

so that the structure of one dimension is defined from the structure of the previous one by adding the two new vectors V and W. The same holds for $\mathcal{A} \oplus \bar{\mathcal{N}}$.

Now, the work is to find what is *added* to the horizon when one passes from one dimension to the higher one. From that point of view, the matrix V_i has a wonderful property: e^V does not change the t and y component of the vector on which it acts. Thus we have the following.

LemHorpigeVDdeux

Lemma 79.31.
We have

$$\pi(g e^{-s\,\mathrm{Ad}(k)E_1}) \in \mathscr{S} \tag{79.117}$$

if and only if

$$\pi(e^V g e^{-s\,\mathrm{Ad}(k)E_1}) \in \mathscr{S}. \tag{79.118}$$

The same holds replacing V by X.

Proof. The exponential of the matrix V_5 is given in equation (79.24). The second and fourth column being the identity, $e^V 1_t = 1_y$ and $e^V 1_y = 1_y$. Thus the characterisation $t^2 - y^2 = 0$ of the singularity is satisfied for one point $x \in AdS$ if and only if it is satisfied by the point $e^V x$. □

We consider the following points in the horizon:

$$r(a,\alpha,w) = e^{wW} r_{AN} = \frac{1}{2}\begin{pmatrix} 2\alpha \\ e^{-a}w^2 + 2\cosh(a) \\ 2\alpha \\ e^{-a}w^2 + 2\sinh(a) \\ 2e^{-a}w \end{pmatrix}, \tag{79.119}$$

EqPartewWrAN

$$\bar{r}(a,\alpha,w) = e^{wW} r_{A\bar{N}} = \frac{1}{2}\begin{pmatrix} 2\alpha \\ e^{-a}w^2 + 2\cosh(a) \\ -2\alpha \\ e^{-a}w^2 + 2\sinh(a) \\ 2e^{-a}w \end{pmatrix}.$$

The tangent vectors of that surface are given by

$$(\partial_a r)(a,\alpha,w) = \begin{pmatrix} 0 \\ \frac{-e^{-a}w^2+2\sinh(a)}{2} \\ 0 \\ \frac{-e^{-a}w^2+2\cosh(a)}{2} \\ -e^{-a}w \end{pmatrix}, \quad (\partial_\alpha r)(a,\alpha,w) = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad (\partial_w r)(a,\alpha,w) = \begin{pmatrix} 0 \\ e^{-a}w \\ 0 \\ e^{-a}w \\ e^{-a} \end{pmatrix}. \tag{79.120}$$

Notice that these three vectors are nowhere vanishing. It is immediate that the vector $\partial_\alpha r$ is linearly independent of $\partial_a r$ and of $\partial_w r$. It is also immediately apparent that $\partial_a r = -w\partial_w r$ is the worse possible situation. It is, however, not possible because it would imply that

$$-w^2 e^{-a} = \frac{-e^{-a}w^2 + 2\sinh(a)}{2} \quad \text{and} \quad -w^2 e^{-a} = \frac{-e^{-a}w^2 + 2\cosh(a)}{2}, \tag{79.121}$$

which is only possible when $\cosh(a) = \sinh(a)$, in other words: never. Thus, the part of AdS_4 described by (79.119) has dimension 3.

Proposition 79.32.
We have

$$G_W \cdot \iota(\mathscr{H}_3) = \{r(a,\alpha,w) \cup \bar{r}(a,\alpha,w)\}_{a,\alpha,w\in\mathbb{R}}. \tag{79.122}$$

Proof. The facts that $G_W \cdot \iota(\mathscr{H}_3) = \{r(a,\alpha,w) \cup \bar{r}(a,\alpha,w)\}_{a,\alpha,w\in\mathbb{R}}$ and that all the elements of that set are subject to $u^2 - x^2 = 0$ are by construction.

We still have to prove that $\{u^2 - x^2 = 0\} \subseteq G_W \cdot \iota(\mathscr{H}_3)$.

Let $v = (y,t,x,y,z)$ be a vector which satisfies $u^2 - x^2 = 0$. Following the signs of u and t, we are searching v under the form $\pm r(\alpha,a,w)$ or $\pm\bar{r}(\alpha,a,w)$. In any case, the value of u and x fix α and we are left with the condition

$$\pm\frac{1}{2}\begin{pmatrix} e^{-a}w^2 + 2\cosh(a) \\ e^{-a}w^2 + 2\sinh(a) \\ 2e^{-a}w \end{pmatrix} = \begin{pmatrix} t \\ y \\ z \end{pmatrix} \tag{79.123}$$

with $t^2 - y^2 - z^2 = 1$. If $t - y > 0$, we choose the sign $+$ and the value of $t - y$ fixes a because $t - y = e^{-a}$. In that situation, w is given by $w = e^a\big(2y - 2\sinh(a)\big)$. If $t - y < 0$, we have $t - y = -e^{-a}$ and the same argument holds.

□

In the sequel, we will use the following notations:

$$\begin{aligned} G_W &= \{e^{wW} \text{ tel que } w \in \mathbb{R}\} \\ G_V &= \{e^{\alpha V} \text{ tel que } \alpha \in \mathbb{R}\} \\ G_X &= \{e^{\beta X} \text{ tel que } \beta \in \mathbb{R}\} \\ G_Y &= \{e^{yY} \text{ tel que } y \in \mathbb{R}\} \end{aligned} \tag{79.124}$$

These are one parameter subgroups of $SO(2,3)$.

PropInclusionsTroisQuatreWVXY

Proposition 79.33.
If $v \in AdS_4$ satisfies $u - t \neq 0$, then $v = e^{wW} v'$ for a certain $v' \in AdS_3$. In other words,

$$\{y - t \neq 0\}_4 \subset G_W \cdot \iota(AdS_3). \tag{79.125}$$

In particular, every points outside the singularity $\mathscr{S}_4$ are obtained by action of G_W on a point of AdS_3. We also have

$$\begin{aligned} \{x - u \neq 0\}_4 &\subseteq G_V \cdot \iota(AdS_3) \\ \{x + u \neq 0\}_4 &\subseteq G_X \cdot \iota(AdS_3). \end{aligned} \tag{79.126}$$

Proof. We have

$$e^{wW} \begin{pmatrix} u' \\ t' \\ x' \\ y' \\ 0 \end{pmatrix} = \begin{pmatrix} u' \\ \left(1 + \frac{w^2}{2}\right) t' - \frac{w^2}{2} y' \\ x' \\ \frac{w^2}{2} t' + \left(1 - \frac{w^2}{2}\right) y' \\ w(t' - y') \end{pmatrix} = \begin{pmatrix} u \\ t \\ x \\ y \\ z \end{pmatrix} \tag{79.127}$$

when

$$u' = u, \quad t' = \frac{z^2 + 2ty - 2t^2}{2(y - t)}, \quad x' = x, \quad y' = \frac{z^2 - 2ty + 2y^2}{2(y - t)}, \quad w = -\frac{z}{y - t}. \tag{79.128}$$

In the same way, the equation

$$e^{\alpha V} \begin{pmatrix} u' \\ t' \\ x' \\ y' \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ t \\ x \\ y \\ z \end{pmatrix} \tag{79.129}$$

is solved by

$$u' = \frac{z^2 + 2ux - 2u^2}{2(x - u)}, \quad x' = \frac{z^2 - 2ux + 2x^2}{2(x - u)}, \quad \alpha = -\frac{z}{x - u}. \tag{79.130}$$

Thus, $\{x - u \neq 0\}_4 \subseteq G_V \cdot \iota(AdS_3)$. And, finally, the equation

$$e^{\beta X} \begin{pmatrix} u' \\ t' \\ x' \\ y' \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ t \\ x \\ y \\ z \end{pmatrix} \tag{79.131}$$

is solved by

$$u' = \frac{z^2 - 2ux - 2u^2}{2(x + u)}, \quad x' = \frac{z^2 + 2ux + 2x^2}{2(x + u)}, \quad \beta = -\frac{z}{x + u}. \tag{79.132}$$

Thus, $\{x + u \neq 0\}_4 \subseteq G_X \cdot \iota(AdS_3)$.

□

One interest of that proposition resides in the fact that every element of AdS_4 outside the singularity is the image of an element of AdS_3 by G_W.

PropSingQTiV

Proposition 79.34.
We have

$$\mathscr{S}_4 = G_V \cdot \iota(\mathscr{S}_3) \tag{79.133}$$

where $G_V = \{e^{vV} \text{ tel que } v \in \mathbb{R}\}$ is the group generated by V.

Proof. A point of $\iota(\mathscr{S}_3)$ is of the form $\begin{pmatrix} u \\ \alpha \\ x \\ \epsilon\alpha \\ 0 \end{pmatrix}$, while an element of $\mathscr{S}_4$ is of the form $\begin{pmatrix} u' \\ \alpha \\ x' \\ \epsilon\alpha \\ z' \end{pmatrix}$ where $\epsilon = \pm 1$. So we have to solve the equation

$$e^{vV}\begin{pmatrix} u \\ \alpha \\ x \\ \epsilon\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} u\left(\frac{v^2}{2}+1\right) + \frac{v^2}{2}x \\ \alpha \\ \left(1-\frac{v^2}{2}\right)x + \frac{v^2}{2}u \\ \epsilon\alpha \\ v(u-x) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} u' \\ \alpha \\ x' \\ \epsilon\alpha \\ z' \end{pmatrix} \tag{79.134}$$

with respect to v, u and x. A solution is given by

$$u = \frac{z'^2 + 2u'x' - 2u'^2}{2(x'-u')}, \quad x = \frac{z'^2 - 2u'x' + 2x'^2}{2(x'-u')}, \quad v = -\frac{z'}{x'-u'}. \tag{79.135}$$

The condition $u'^2 - x'^2 - z'^2 = 1$ imposes $x' \neq u'$, so that that solution always makes sense: a point of $\mathscr{S}_4$ is always obtained as the result of the action of an element of G_V on an element of $\mathscr{S}_3$.

Since the operator e^{vV} does not touch the variables t and y, it is obvious that $G_V \cdot \mathscr{S}_3 \subseteq \mathscr{S}_4$. □

LemTNTroisIneq

Lemma 79.35.
In AdS_3, the black hole is given by $u^2 - x^2 > 0$

Proof. The black hole is the set of point from which every light ray intersect the singularity. The boundary of that set is given by the horizon (this is the definition of the horizon), and we already proved that $\mathscr{H}_3 \equiv u^2 - x^2 = 0$. Thus the black hole is $u^2 - x^2 > 0$, or $u^2 - x^2 < 0$. Since the singularity (which is part of the black hole) is given by $t^2 - y^2 = 0$, the singularity satisfies $u^2 - x^2 = 1$, and is thus in the part $u^2 - x^2 > 0$. □

Let $TN[g]$ be the subset of $\{\mathrm{Ad}(k)E_1\}_{k\in \mathrm{SO}(3)}$ of elements for which there exists a $s > 0$ such that

$$\pi(g e^{s\,\mathrm{Ad}(k)E_1}) \in \mathscr{S}. \tag{79.136}$$

In other words, $TN[g]$ is the set of directions along which $[g]$ falls in the singularity. If the complementary $TN[g]^c$ has a non empty interior, the by continuity, the complementary $TN[g']$ will have an interior as well for every $[g']$ close enough from $[g]$. In that case, $[g]$ does not belongs to the horizon. So a point belongs to the horizon when the set of safe direction has no interior.

Lemma 79.36.
We have

$$G_V \cdot \iota(\mathscr{H}_3) \equiv u^2 - x^2 - z^2 = 0, \tag{79.137}$$

so that it is the good candidate to be the horizon.

Proof. An element of $\iota(\mathscr{H}_3)$ has the form $r = \begin{pmatrix} u' \\ t' \\ x' \\ \pm\sqrt{t'^2-1} \\ 0 \end{pmatrix}$, so that we have to solve the equation

$$e^{vV}r = \begin{pmatrix} \left(\frac{v^2}{2}+1\right)u' - \frac{v^2}{2}x' \\ t' \\ \left(1-\frac{v^2}{2}\right)x' + \frac{v^2}{2}u' \\ \pm\sqrt{t'^2-1} \\ v(u'-x') \end{pmatrix} = \begin{pmatrix} u \\ t \\ x \\ \pm\sqrt{t^2-1} \\ z \end{pmatrix}. \tag{79.138}$$

The solution is

$$u' = \frac{z^2 + 2ux - 2u^2}{2(x-u)}, \quad x' = \frac{z^2 - 2ux + 2x^2}{2(x-u)}, \quad v = -\frac{z}{x-u}. \tag{79.139}$$

Since $u^2 - x^2 - z^2 = 1$, we have $x - u \neq 0$, so that these solutions always make sense. □

Lemma 79.37.

If $[g] = \begin{pmatrix} u \\ t \\ x \\ y \\ z \end{pmatrix} \in AdS_4$ *with* u *and* x *not both vanishing, then*

$$[g] \in G_V \cdot \iota(AdS_3) \cup G_X \cdot \iota(AdS_3). \tag{79.140}$$

Notice that the union is not disjoint.

Proof. The proof is a simple computation. Following proposition 79.33, we have $\{x - u \neq 0\}_4 \subseteq G_V \cdot \iota(AdS_3)$ and $\{x + u \neq 0\}_4 \subseteq G_X \cdot \iota(AdS_3)$.

So the only part of AdS_4 which is not included in $G_V \cdot \iota(AdS_3) \cup G_X \cdot \iota(AdS_3)$ is the part where $x + u = x - u = 0$. □

Now, we want to study the horizon, that means the boundary of BH_4. If $v \in \partial\big(\operatorname{Adh}(BH_4)\big)$, there exists, in any neighbourhood of v, an element $\bar{v}$ and a direction following which the geodesic from $\bar{v}$ escapes the singularity.

Up to now, we studied the way AdS_3 embed in AdS_4. In particular, we proved that the horizon of AdS_3 is included in the horizon of AdS_4. We can propagate the results by G_V and G_X because, given a $v \in AdS_3$, the existence of a α such that $e^{\alpha V} v \in \iota(AdS_3)$ or $e^{\alpha X} v \in \iota(AdS_3)$ is related to the fact that $u^2 - x^2 \neq 0$, while that condition holds in a neighbourhood of v.

79.12 Organization of the next few pages

Abstract

This paper is a sequel of *Solvable symmetric black hole in anti de Sitter spaces* [1019]. In the latter, we described the BTZ black hole in every dimension by defining the singularity as the closed orbits of the Iwasawa subgroup of $\mathrm{SO}(2,n)$. In this article, we study the horizon of the black hole and we show that it is expressed as lateral classes of one point of the space. The computation is given in the four-dimensional case, but it makes no doubt that it can be generalized to any dimension.

The main idea is to define an "inclusion map" from AdS_3 into AdS_4 and to show that all the relevant structure pass trough the inclusion. We prove, for example, that the inclusion of the three dimensional horizon into AdS_4 belongs to the four dimensional horizon: $\iota(\mathscr{H}_3) \subseteq \mathscr{H}_4$ and then we deduce the expression of the horizon in AdS_4.

In section 79.13, we describe some old results about BTZ black hole.

In subsection 79.13.1, we recall how we proved the existence of the black hole structure in [1019] and how the horizon was described in the three dimensional case in [1017]. We adapt the latter result in our homogeneous space setting.

The subsection 79.13.2 gives some topological remarks about the black hole and the horizon. We point out that there are some light-like geodesics that are intersecting the singularity *and then* the free part later in the future. We explain why that circumstance is very different from the situation of the most famous black holes in physics like the Schwarzschild's one.

Section 79.14 is devoted to the proof of our main result: the horizon of the BTZ black hole in AdS_4 is given by

$$\mathscr{H}_4 = G_{X_{0+}} \cdot \iota(\mathscr{H}_3) \cup G_{X_{0-}} \iota(\mathscr{H}_3). \tag{79.141}$$

where ι is the inclusion of AdS_3 in AdS_4 and $\mathscr{H}_3$ is the horizon of the BTZ black hole in AdS_3.

79.13 Some old results

SecOldResults

From the results of section 79.9, we know that a non trivial horizon exists. However, the question of the structure of the horizon was not yet addressed. This is what we are going to do now.

79.13.1 Horizon in the three dimensional case

SubSecHorInThreeDimensionOld

The structure of the horizon of AdS_3 was described in [1017] in the setting of $AdS_3 = \mathrm{SL}(2,\mathbb{R})$. Our first job is to translate that result into the language of quotient of groups. This is done by the identification

$$\begin{aligned}\psi\colon\ \mathrm{SL}(2,\mathbb{R}) &\to AdS_3\\ \begin{pmatrix} u+x & y+t\\ y-t & u-x\end{pmatrix} &\mapsto \begin{pmatrix} u\\ t\\ x\\ y\end{pmatrix}.\end{aligned} \tag{79.142}$$

We see that the points of the horizon are given by

$$\pm\begin{pmatrix}\alpha\\ \cosh(a)\\ \alpha\\ \sinh(a)\end{pmatrix}\quad\text{and}\quad\pm\begin{pmatrix}\alpha\\ \cosh(a)\\ -\alpha\\ \sinh(a)\end{pmatrix}, \tag{79.143}$$

EqHOrAdSTroisVecteur

which correspond to the points (u,t,x,y) such that $u^2-x^2=0$. One should notice that these points can be expressed as lateral classes of the point $b=(0,1,0,0)$:

$$\mathscr{H}_3 = \pm G_{\{J_1,X_{++}\}}b \cup \pm G_{\{J_1,X_{--}\}}b \tag{79.144}$$

where $G_{\{X,Y\}}$ is the group of elements of the form $\exp(aX+bY)$. Notice that $G_{\{J_1,X_{++}\}}b = G_{\{J_1,X_{-+}\}}b$ and $G_{\{J_1,X_{--}\}}b = G_{\{J_1,X_{+-}\}}b$. For example,

$$e^{aJ_2}e^{\alpha X_{++}}b = \begin{pmatrix}\alpha\\ \cosh(a)\\ \alpha\\ \sinh(a)\end{pmatrix}. \tag{79.145}$$

We are now intended to extend that result and express the horizon in AdS_4 as lateral classes of the horizon in AdS_3. Before to complete that work, we have to make a few remarks about the topology.

79.13.2 Topology and horizon

subSecTopoHor

The definition given in the previous sections produces a paradox. Let $x \in AdS$ and $l(s)$ be a light like geodesic trough x which only intersects the singularity in past. We suppose that $l(0)=x$ and that $s_0<0$ is the biggest value of s such that $l(s_0)\in\mathscr{S}$. Thus, all points of the form $l(s)$ with $s_0<s<0$ are free. That form a sequence of free points which converges to the singularity, and then $l(s_0)$ belongs to the horizon.

This is however not possible in AdS_3 because the equation of the singularity is $t^2-y^2=0$ while the equation of the horizon is $u^2-x^2=0$. These two parts are really separated.

The situation here is really different from the situation in the Schwarzschild's case. In the latter the singularity is well inside the horizon, and there are no geodesics reaching the infinity which have intersected the singularity in the past.

In our case, however, such geodesics do exist. The reason of such a difference resides in the fact that the causal structure (geodesics) are defined by the metric while, in our BTZ black hole, the singularity is not defined from metric considerations. There are thus no reasons to expect some compatibility relations like the fact to have a non naked singularity.

In order to correctly define the horizon, we have to introduce the space $BTZ = AdS\backslash\mathscr{S}$ which in endowed with the induced topology. Then we define

$$BH = \{v \in BTZ \text{ tel que } \forall k \in \mathrm{SO}(n),\ l_v^k(s) \in \mathscr{S} \text{ has a solution with } s > 0\}. \tag{79.146}$$

Let us point out that the singularity itself is not part of the black hole, because it is not even part of BTZ. We define the free part of BTZ as the set of points from which there exists a light-like geodesics which does not intersects the singularity in the future:

$$\mathscr{F} = \{v \in BTZ \text{ tel que } \exists k \in \mathrm{SO}(n),\ l_v^k(s) \in \mathscr{S} \Rightarrow s < 0\}. \tag{79.147}$$

The first definition makes that the black hole part is open by continuity and compactness of $\mathrm{SO}(n)$: the minimum and the maximum of time to reach the singularity from one point of the black hole are both strictly positive numbers, and then can be maintained strictly positive in a neighborhood of the point.

PropBHouvertLibreFerme

Proposition 79.38.
The set of points in the black hole is open and set of free points is closed. In particular, the horizon is contained in the free set.

Proof. The first point is the remark above. Now, the free part is closed in BTZ as complementary of an open set. □

The following theorem says that if the set of directions escaping the singularity from a point in BTZ has an interior, then that point does not lies in the horizon.

PropvFOsvghorvec

Proposition 79.39.
A point $v \in \mathscr{F}_l$ such that there is an open set $\mathcal{O} \subset S^{l-1}$ of directions for which $l_v^w(s) \in \mathscr{S}$ has no solutions for $s \in \mathbb{R}_0^+$ belongs to $\mathrm{Int}(\mathscr{F})$.

Proof. Using the matricial representation (79.42), we see that a point $v = [g]$ belongs to the singularity if the vector

$$g \cdot \begin{pmatrix} 1 \\ -s \\ s\bar{w} \end{pmatrix} \tag{79.148}$$

satisfies $t^2 - y^2 = 0$. That equation is a second order polynomial in s whose coefficients cannot be a constant for an open set with respect to $\bar{w} \in S^{l-1}$. From the assumptions, all the roots of that polynomial belong to $\mathbb{C}\backslash\mathbb{R}_0^+$. The latter being open, the roots of $l_{v'}^w(s) \in \mathscr{S}$ are still in $\mathbb{C}\backslash\mathbb{R}_0^+$ when v' runs over a small enough open set around v.

We conclude that v is in the interior of the free zone rather than on the horizon. □

An important characterisation of the horizon, pointed out in [1017], is the following.

ThoHorIntDansS

Theorem 79.40.
A point belongs to the horizon if and only if the set of light-like directions for which the geodesics does not intersects the singularity has no interior in S^{l-1}.

79.14 The horizon of the BTZ black hole

SecNewWithMatrices

In this section, we show, that the horizon of the horizon of AdS_4 can be obtained using the action of a very simple group on the horizon of AdS_3, which is, itself, the orbit of one point under a known group. The result opens the possibility of describing the horizon in AdS_l by induction on the dimension, and the possibility to compute the group which generates the horizon. We define

the inclusion map

$$\begin{aligned} \iota\colon AdS_3 &\to AdS_4 \\ \begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} u \\ t \\ x \\ y \\ 0 \end{pmatrix}. \end{aligned} \tag{79.149}$$

At the matrix level, it corresponds to add a line and a column of zeros. We will denote by $\mathscr{F}_l$ the free part of AdS_l. By definition, if $v \in \mathscr{F}_l$, there exists a light like geodesic trough v which does not intersect the singularity in the future. We also denote by BH_l the set of elements of AdS_l from which all the light-like geodesics intersect the singularity in the future.

Notice that BH_l is open while $\mathscr{F}_l$ is closed, as explained in proposition 79.38.

LemOouversttq

Lemma 79.41.
Let $v \in AdS_4$ and $g \in \mathrm{SO}(2,3)$ be a representative of v. If the set

$$\{\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in S^2 \text{ tel que } \pi g \begin{pmatrix} 1 \\ -s \\ s\bar{w} \\ 0 \end{pmatrix} \cap \mathscr{S}_4 = \varnothing \text{ with } s > 0\} \tag{79.150}$$

has an interior in S^1, then the set

$$\{\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in S^2 \text{ tel que } \pi g \begin{pmatrix} 1 \\ -s \\ s\bar{w} \end{pmatrix} \cap \mathscr{S}_4 = \varnothing \text{ with } s > 0\} \tag{79.151}$$

has an interior in S^2.

Proof. The matrix g in $\mathrm{SO}(2,3)$ representing the point v has the form

$$g = \begin{pmatrix} u & . & . & . & . \\ t & a & b & c & d \\ x & . & . & . & . \\ y & a' & b' & c' & d' \\ z & . & . & . & . \end{pmatrix} \tag{79.152}$$

where the numbers $a, b, c, d, a', b', c', d'$ are not uniquely determined. We choose the representative in such a way to have $b \neq \pm b'$, which is always possible.

The assumption is that there exists an open set (with respect to $(w_1, w_2) \in S^1$) around $(w_1, w_2, 0)$ such that the path

$$\pi(ge^{s\,\mathrm{Ad}(k)E_1}) = \begin{pmatrix} U \\ T \\ X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} u & . & . & . & . \\ t & a & b & c & d \\ x & . & . & . & . \\ y & a' & b' & c' & d' \\ z & . & . & . & . \end{pmatrix} \begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \\ 0 \end{pmatrix} \tag{79.153}$$

EqPathgexpUTXYZ

does not intersects the singularity in the future. In other words, we have $T \pm Y = 0$ only with $s \leqslant 0$. Let

$$\begin{aligned} T(w_1, w_2) &= t + s(bw_1 + cw_2 - a) \\ Y(w_1, w_2) &= y + s(b'w_1 + c'w_2 - a') \\ A_+(w_1, w_2) &= (b + b')w_1 + (c + c')w_2 - (a + a') \\ A_-(w_1, w_2) &= (b - b')w_1 + (c - c')w_2 - (a - a'). \end{aligned} \tag{79.154}$$

We also denote by $\sigma_\pm$ the sign of $t \pm y$.

A simple computation shows that $T + Y = 0$ when

$$s = s_+ = -\frac{t+y}{A_+(w_1, w_2)}, \tag{79.155}$$

and $T - Y = 0$ when

$$s = s_- = -\frac{t-y}{A_-(w_1, w_2)}, \tag{79.156}$$

The assumption is that the direction $(w_1, w_2, 0)$ (and an open set in S^1 with respect to (w_1, w_2)) escapes the singularity, so that for every (w_1', w_2') in a neighborhood of (w_1, w_2), we have

$$\sigma_\pm A_\pm(w_1', w_2') \geqslant 0, \tag{79.157}$$

which assures that the values of s which annihilate $T + Y$ and $T - Y$ are negative or non existing. Since we choose $b \neq \pm b'$, the functions $A_\pm$ are nowhere constant, so we can find a direction (w_1, w_2) such that $\sigma_\pm A_\pm(w_1, w_2) > 0$. Notice that, by continuity, there exists a neighbourhood of (w_1, w_2) in S^1 which escapes the singularity.

We are now studying what happens when one looks at a neighbourhood of $(w_1, w_2, 0)$ in S^3. The path (79.153) is replaced by

$$\pi(g e^{s\,\mathrm{Ad}(k)E_1}) = \begin{pmatrix} u & . & . & . & . \\ t & a & b & c & d \\ x & . & . & . & . \\ y & a' & b' & c' & d' \\ z & . & . & . & . \end{pmatrix} \begin{pmatrix} 1 \\ -s \\ s(w_1 + \epsilon_1) \\ s(w_2 + \epsilon_2) \\ \epsilon_3 \end{pmatrix}, \tag{79.158}$$

and we consider

$$\begin{aligned} T(w_1, w_2, \bar\epsilon) &= t + s\big(b(w_1 + \epsilon_1) + c(w_2 + \epsilon_2) + d\epsilon_3 - a\big) \\ Y(w_1, w_2, \bar\epsilon) &= y + s\big(b'(w_1 + \epsilon_1) + c'(w_2 + \epsilon_2) + d'\epsilon_3 - a'\big) \end{aligned} \tag{79.159}$$

where $\bar\epsilon$ stands for ϵ_1, ϵ_2 and ϵ_3. The same computations as before shows that $T + Y = 0$ when

$$s = s_+ = -\frac{t+y}{A_+(w_1, w_2) + (b+b')\epsilon_1 + (c+c')\epsilon_2 + (d+d')\epsilon_3}, \tag{79.160}$$

Since $\sigma_+ A(w_1, w_2) > 0$, there exists a δ such that s_+ remains negative for every choice of $\bar\epsilon < \delta$. The same holds with $T - Y$ which is zero when

$$s = s_- = -\frac{t-y}{A_-(w_1, w_2) + (b-b')\epsilon_1 + (c-c')\epsilon_2 + (d-d')\epsilon_3}. \tag{79.161}$$

Since $\sigma_- A_-(w_1, w_2) > 0$, one can find a $\delta > 0$ such that $\bar\epsilon < \delta$ implies that this fraction remains negative.

Thus, there exists a neighbourhood of $(w_1, w_2, 0)$ in S^2 of directions escaping the singularity from the point v. □

LemIntTroisQueatr

Lemma 79.42.
With the notations defined before, we have

$$\iota\big(\mathrm{Int}(\mathscr{F}_3)\big) \subseteq \mathrm{Int}\,(\mathscr{F}_4) \tag{79.162}$$

where Int *stands for the interior. In other words,*

$$\mathrm{Adh}(BH_4) \cap \iota(AdS_3) \subset \iota\big(\mathrm{Adh}(BH_3)\big). \tag{79.163}$$

Proof. Let $v = \iota(v') \notin \iota\big(\mathrm{Adh}(BH_3)\big)$, we also consider g' a representative of v' and $g = \iota(g')$, which is a representative of v. The element v' is in the interior of the free zone: there exists an open set of directions which do not intersect the singularity of AdS_3 by theorem 79.40. In other words, the set

$$\{\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in S^1 \text{ tel que } \pi g' \begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \end{pmatrix} \cap \mathscr{S}_3 = \varnothing\} \tag{79.164}$$

EqwwswswUn

contains an open set of S^1. On the other hand, the z-component of the latter vector is obviously zero because $g = \iota(g')$ has the form

$$g = \begin{pmatrix} . & . & . & . & 0 \\ . & . & . & . & 0 \\ . & . & . & . & 0 \\ . & . & . & . & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{79.165}$$

thus equation (79.164) can be "extended" and there exists an open set in S^1 such that

$$\pi g \begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \\ 0 \end{pmatrix} \cap \iota(\mathscr{S}_3) = \varnothing. \tag{79.166}$$

Now, lemma 79.41 shows that the set

$$\{\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in S^2 \text{ tel que } \pi g \begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \\ sw_3 \end{pmatrix} \cap \mathscr{S}_4 = \varnothing\} \tag{79.167}$$

contains an open subset of S^2. That means that $\pi(g) = v$ belongs to the interior of $\mathscr{F}_4$. □

PropFqTroisFt

Proposition 79.43.
We have $\mathscr{F}_4 \cap \iota(AdS_3) \subset \iota(\mathscr{F}_3)$.

Proof. Let $v \in \mathscr{F}_4 \cap \iota(AdS_3)$. With the same notations as above, we have

$$\iota(g') = \begin{pmatrix} u & . & . & . & 0 \\ t & a & b & c & 0 \\ x & . & . & . & 0 \\ y & a' & b' & c' & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{79.168}$$

EqRepresSQiotag

The assumption is that, for every representative g' of v', there exists a direction $(w_1, w_2, w_3) \in S^2$ such that the path

$$\pi\iota(g') \begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \\ sw_3 \end{pmatrix} \tag{79.169}$$

EqGedgpudt

only intersects the singularity fore negative values of s. The values of s that annihilate $t^2 - y^2$ in the geodesic (79.169) are

$$\begin{aligned} s_+ &= -\frac{t+y}{-(a+a')+(b+b')w_1+(c+c')w_2} \\ s_- &= -\frac{t-y}{-(a-a')+(b-b')w_1+(c-c')w_2}, \end{aligned} \tag{79.170}$$

and these two values are either negative either non existing (vanishing denominator).

The work is now to find a direction $(w_1', w_2') \in S^1$ such that the geodesic

$$\pi\big(g' \begin{pmatrix} 1 \\ -s \\ sw_1' \\ sw_2' \end{pmatrix}\big) \tag{79.171}$$

does not intersect the singularity. The values of s for which the latter geodesics intersects the singularity are

$$\begin{aligned} s'_+ &= -\frac{t+y}{-(a+a')+(b+b')w_1'+(c+c')w_2'} \\ s'_- &= -\frac{t-y}{-(a-a')+(b-b')w_1'+(c-c')w_2'}. \end{aligned} \tag{79.172}$$

If $w_3 = 0$, the proposition is true because one can choose $(w_1', w_2') = (w_1, w_2)$. If $w_3 \neq 0$, the vector (w_1, w_2) does not belong to S^1, and we have to find something else.

Let us consider the following two cases.

(1) there exists a representative (79.168) with $a = a' = 0$,

(2) there exists a representative (79.168) with $c = c' = 0$.

In the first case, we have

$$s'_\pm = -\frac{t \pm y}{(b \pm b')w_1' + (c \pm c')w_2'}, \tag{79.173}$$

EqDenoAAmulersnn

and we can choose $(w_1', w_2') = N(w_1, w_2)$ with $N \in \mathbb{R}$ fixed in such a way that $(w_1', w_2') \in S^1$. Thus we have $s'_\pm = \frac{1}{N} s_\pm$ and it is sufficient to choose $N > 0$ in order to leave the denominators of (79.173) of the right sign or zero.

In the second case, we have

$$s'_\pm = -\frac{t \pm y}{-(a \pm a') + (b \pm b')w_1'}, \tag{79.174}$$

thus one has to choose $w_1' = w_1$ and $w_2' = \sqrt{1 - w_1^2}$.

Let us now discuss the values of u, t, x and y for which the first or the second cases are enforced. In order to be in the first case, we need to build a matrix of $\mathrm{SO}(2,2)$ of the form

$$g' = \begin{pmatrix} u & \alpha & . & . \\ t & 0 & . & . \\ x & \beta & . & . \\ y & 0 & . & . \end{pmatrix}. \tag{79.175}$$

That requires $\alpha^2 - \beta^2 = 1$ and $u\alpha - x\beta = 0$, while, for the second case, we need to build a matrix of $\mathrm{SO}(2,2)$ of the form

$$g' = \begin{pmatrix} u & . & \alpha & . \\ t & . & 0 & . \\ x & . & \beta & . \\ y & . & 0 & . \end{pmatrix}. \tag{79.176}$$

That requires $\alpha^2 - \beta^2 = -1$ and $u\alpha - x\beta = 0$.

In both cases, we have $\beta = \frac{u}{x}\alpha$ and $\alpha^2 - \beta^2 = \alpha^2\left(1 - \frac{u^2}{x^2}\right)$. If $|u| > |x|$, we can solve $\alpha^2 - \beta^2 = -1$, and if $|u| < |x|$, then we can solve $\alpha^2 - \beta^2 = 1$.

The last possible situation is $u = \pm x$. A point of AdS_3 in that situation belongs to the horizon by equation (79.143), while one knows that point of horizon do have some directions which escape the singularity by corollary 79.38. Notice that in the latter situation, we do not use the assumption that $\iota(v')$ is free in AdS_4. □

CorBHBHHHHH

Corollary 79.44.
We have $\iota(BH_3) \subset BH_4$ and $\iota(\mathscr{H}_3) \subset \mathscr{H}_4$.

Proof. If $\iota(v) \notin BH_4$, we have $\iota(v) \in \mathscr{F}_4 \cap \iota(AdS_3) \subset \iota(\mathscr{F}_3)$, which is not possible if $v \in BH_3$.

For the second part, we consider $v \in \mathscr{H}_3 \subset \mathscr{F}_3$ (proposition 79.38). There is a direction $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in S^1$ which escapes the singularity from v in AdS_3. Of course, the direction $\begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} \in S^2$ escapes the singularity from $\iota(v)$ in AdS_4. Thus $\iota(v) \in \mathscr{F}_4$.

In every neighborhood of v, there exists a $\bar{v} \in BH_3$, and thus $\iota(\bar{v}) \in BH_4$. In other words, in every neighborhood of $\iota(v)$, there is that $\iota(\bar{v})$ which belongs to BH_4. That proves that $\iota(v)$ belongs to $\mathscr{H}_4$. □

LemHinteridansH

Lemma 79.45.
We have $\mathscr{H}_4 \cap \iota(AdS_3) \subset \iota(\mathscr{H}_3)$.

Proof. Let $v \in \mathscr{H}_4 \cap \iota(AdS_3)$. Since $\mathscr{H}_4 \subset \mathscr{F}_4$, we have $v \in \mathscr{F}_4 \cap \iota(AdS_3) \subset \iota(\mathscr{F}_3)$ (proposition 79.43), and then there exists a $v' \in \mathscr{F}_3$ such that $v = \iota(v')$. Now, we have to prove that $v' \in \mathscr{H}_3$. If v' belongs to the interior of $\mathscr{F}_3$, lemma 79.42 implies that

$$v = \iota(v') \in \iota\big(\operatorname{Int}(\mathscr{F}_3)\big) \subset \operatorname{Int}(\mathscr{F}_4), \tag{79.177}$$

which disagrees with the fact that $v \in \mathscr{H}_4$. □

LemPresqueHOrQadp

Lemma 79.46.
Let $v \in \mathscr{H}_4$ such that u and x are not both vanishing. In that case, $v \in G_V \cdot \iota(\mathscr{H}_3) \cup G_X \cdot \iota(\mathscr{H}_3)$.

Proof. The assumption on u and x make that $v \in G_V \cdot (AdS_3) \cup G_X\iota(AdS_3)$. In order to fix ideas, let us suppose that $v = e^{\alpha V}\iota(v')$ with $v' \in AdS_3$. Since the set of directions $(w_1, w_2, w_3) \in S^2$ which save the points v, $e^{\alpha V}v$ and $e^{\beta X}v$ are the same, the assumption that $v \in \mathscr{H}_4$ implies that $\iota(v') \in \mathscr{H}_4$, which in turn proves that $v' \in \mathscr{H}_3$ by lemma 79.45. Thus $v \in G_V \cdot \iota(\mathscr{H}_3)$.

The same being true with X instead of V, the lemma is proved. □

PropovHhnonXYzero

Proposition 79.47.
Let $v' = (u', t', x', y', z') \in \mathscr{H}_4$ with u' and x' not both vanishing. Then

$$v' \in G_{X_{0+}} \cdot \iota(\mathscr{H}_3) \cup G_{X_{0-}} \cdot \iota(\mathscr{H}_3). \tag{79.178}$$

Proof. As a first step, we want to solve the equation

$$e^{\alpha X_{0+}} \begin{pmatrix} u \\ t \\ x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\alpha^2(u-x)}{2} + u \\ t \\ \frac{\alpha^2(u-x)}{2} + x \\ y \\ -\alpha(x-u) \end{pmatrix} = \begin{pmatrix} u' \\ t' \\ x' \\ y' \\ z' \end{pmatrix} \tag{79.179}$$

with respect to u, t, x, y and α. The result is $t = t'$, $y = y'$ and

$$\alpha = \frac{z'}{u' - x'}, \quad u = u' - \frac{z'^2}{2(u' - x')}, \quad x = \frac{z'^2}{2(u' - x')} - x'. \tag{79.180}$$

We conclude that, as long as $u' - x' \neq 0$, the point v' belongs to $G_{X_{0+}} \cdot \iota(AdS_3)$. The same computation shows that $v' \in G_{X_{0-}} \cdot \iota(AdS_3)$ as long as $x' + u' \neq 0$. Let us observe that the actions of the matrices $e^{\alpha X_{0+}}$ and $e^{\beta X_{0-}}$ do not change the t and y component of a vector in $\mathbb{R}^{2,l-1}$, so that the set of directions for which v falls in the singularity is exactly the same as the set of directions for which $e^{\alpha X_{0+}}v$ and $e^{\beta X_{0-}}v$ fall in the singularity.

Now, let us suppose that $v = e^{\alpha X_{0+}}\iota(v') \in \mathscr{H}_4$ with $v' \in AdS_3$. We want to prove that $\iota(v') \in \mathscr{H}_4$ (i.e. there is an element in the black hole in each neighbourhood of $\iota(v')$) because in that case, lemma 79.45 would conclude that $v' \in \mathscr{H}_3$.

Let $\mathcal{O}$ be a neighbourhood of $\iota(v')$. The set $e^{\alpha X_{0+}}\mathcal{O}$ is a neighborhood of v, and thus there exists an element $\bar{v} \in e^{\alpha X_{0+}}\mathcal{O} \cap BH_4$. Now the element $e^{-\alpha X_{0+}}\bar{v}$ belongs to $\mathcal{O} \cap BH_4$, so that $\iota(v')$ belongs to $\mathscr{H}_4$. □

LemPasLEsDecomposAQ

Lemma 79.48.

The points of AdS_4 of the form $v = \begin{pmatrix} 0 \\ t \\ 0 \\ y \\ z \end{pmatrix}$ do not belong to the horizon.

Proof. Since the horizon is A-invariant, we can reduce the lemma to the case of any element of the form $e^{\eta J_1}v$. We have

$$e^{\eta J_1}\begin{pmatrix} 0 \\ t \\ 0 \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cosh(\eta) & 0 & \sinh(\eta) & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \sinh(\eta) & 0 & \cosh(\eta) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ t \\ 0 \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ \cosh(\eta)t + \sinh(\eta)y \\ 0 \\ \sinh(\eta)t + \cosh(\eta)y \\ z \end{pmatrix} \tag{79.181}$$

We annihilate the y component by choosing $\eta = \ln\left(\frac{t-y}{t+y}\right)$. Notice that $t^2 - y^2 > 0$, thus we have $|t| > |y|$ and the expression in the logarithm is always positive.

A representative of $(0, t, 0, 0, z)$ in $\mathrm{SO}(2,2)$ is easy to find, and the geodesic in the direction $\bar{w} \in S^2$ is given by

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ t & 0 & 0 & 0 & -z \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ z & 0 & 0 & 0 & -t \end{pmatrix}\begin{pmatrix} 1 \\ -s \\ sw_1 \\ sw_2 \\ sw_3 \end{pmatrix} = \begin{pmatrix} . \\ t - szw_3 \\ . \\ sw_2 \\ . \end{pmatrix}. \tag{79.182}$$

It belongs to the singularity when s takes one of the values

$$s_\pm = \frac{t}{w_3 z \pm w_2}. \tag{79.183}$$

As long as $|w_2| < |w_3 z|$, the two values $s_\pm$ have the same sign, which can be decided by making w_3 positive or negative. That provides an open set in S^2 of directions which escape the singularity, so that $v \notin \mathscr{H}_4$. □

ThoHorQuatreIndEsHoHQTrois

Theorem 79.49.

The horizon of AdS_4 is given by

$$\mathscr{H}_4 = G_{X_{0+}} \cdot \iota(\mathscr{H}_3) \cup G_{X_{0-}}\iota(\mathscr{H}_3). \tag{79.184}$$

EqEqHOrGVGXQuatr

i.e. an union of lateral classes of the horizon of AdS_3 by one dimensional subgroups of N and $\bar{N}$.

The equation in the ambient $\mathbb{R}^5$ is $\mathscr{H}_4 \equiv u^2 - x^2 - z^2 = 0$.

Proof. We begin by the direct inclusion. If $v = (u, t, x, y, z) \in \mathscr{H}_4$ with $u \neq 0$ or $x \neq 0$, we proved in proposition 79.47 that v has the form (79.184). Now, if $u = x = 0$, the lemma 79.48 shows that v does not belongs to the horizon.

For the reverse inclusion, we know that elements of $\iota(\mathscr{H}_3)$ belong to $\mathscr{H}_4$ by corollary 79.44. If v belong to $\mathscr{H}_4$, then $e^{\alpha X_{0+}}v$ and $e^{\beta X_{0-}}v$ also belong to the horizon. □

79.15 Conclusion

The horizon of the BTZ black hole in AdS_3 was already expressed in [1017] as lateral classes of one point under the action of the Iwasawa component of the isometry group of AdS_3.

We proved that the simple inclusion map $\iota\colon AdS_3 \to AdS_4$ transports the causal structure (free zone, black hole, horizon) from AdS_3 to AdS_4. We studied in particular the way the horizon changes when ones jumps from dimension 3 to dimension 4 and we obtained that the horizon in AdS_4 is expressed as lateral classes of the inclusion of the horizon of AdS_3 in AdS_4. In the same time, we obtained a simple equation for the horizon seen as a subset of $\mathbb{R}^5$.

Although the results are quite satisfying, the method used here to prove them is quite unsatisfactory because we didn't used all the wealth structure of $\mathfrak{so}(2,3)$ and of its reductive decompositions $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q} = \mathcal{K} \oplus \mathcal{K}$. We plan, in a future work, to get a much deeper understanding of the structure of $\mathcal{G}$ and $\mathcal{Q}$, in such a way to provide simpler proofs, in the same time as a dimensional generalization of the result of theorem 79.49. We would also like to define a class of homogeneous spaces G/H which accept a BTZ-like black hole.

79.16 The algebras without matrices

SecRebuildStructRoot

We have two decompositions

$$\begin{aligned} \mathcal{G} &= \mathcal{K} \overset{\theta}{=} \mathcal{P} \\ \mathcal{G} &= \mathcal{H} \overset{\sigma}{=} \mathcal{Q} \end{aligned} \tag{79.185}$$

of $\mathcal{G} = \mathfrak{so}(2,n)$. From there, we will build the basis elements of $\mathcal{A}$, $\mathcal{N}$, $\bar{\mathcal{N}}$ with all the properties we used so far. The explicit matrices (66.123) and (66.92) consist in a concrete realisation of what we are about to do.

79.16.1 The structure theorem by Pyatetskii-Shapiro

We are going to use the Pyatetskii-Shapiro's decompositions of normal j-algebra (66.292) and (66.289).

Lemma 79.50.
We have

$$\|(X_{\alpha\beta})_{\mathcal{K}}\| = \|(X_{\alpha\beta})_{\mathcal{P}}\|. \tag{79.186}$$

Proof. We use the invariance of the Killing form:

$$\begin{aligned} B\big((X_{\alpha\beta})_{\mathcal{K}}, (X_{\alpha\beta})_{\mathcal{K}}\big) &= \frac{1}{\alpha} B\big((X_{\alpha\beta})_{\mathcal{K}}, \mathrm{ad}(J_1)(X_{\alpha\beta})_{\mathcal{P}}\big) \\ &= -\frac{1}{\alpha} B\big(\mathrm{ad}(J_1)(X_{\alpha\beta})_{\mathcal{K}}, (X_{\alpha\beta})_{\mathcal{P}}\big) \\ &= -B\big((X_{\alpha\beta})_{\mathcal{P}}, (X_{\alpha\beta})_{\mathcal{P}}\big). \end{aligned} \tag{79.187}$$

□

Lemma 79.51.
we have

$$\|(X_{\alpha\beta})_{\mathcal{K}}\| = \|(X_{\alpha,-\beta})_{\mathcal{K}}\|. \tag{79.188}$$

Proof. First, remark that $X_{\alpha,-\beta} = \sigma X_{\alpha\beta}$. We also know that $[\mathtt{proj}_{\mathcal{K}}, \sigma] = 0$ because $[\sigma, \theta] = 0$. The conclusion now comes from the fact that σ is an isometry. □

79.17 Characterisation of the horizon (vanishing norm)

SecVanNormChar

In order to get the theorem 79.49, we used the equation of the singularity, $\mathscr{S} \equiv t^2 - y^2 = 0$, which was proved in proposition 79.19. But subsection 79.8.3 provides an other characterisation of the singularity, namely the loci of points $[g]$ such that $\|J_1^*\| = 0$.

79.18 Conclusions and perspectives

SecConcPerspAd

Higher-dimensional generalizations of the BTZ construction have been studied in the physics' literature, by classifying the one-parameter subgroups of $\mathrm{Iso}(AdS_l) = \mathrm{SO}(2, l-1)$, see [1023, 1024, 1025, 1026, 1027, 1028]. Nevertheless, the approach we adopt here is conceptually different. We first reinterpret the non-rotating BTZ black hole solution using symmetric spaces techniques and present an alternative way to express its singularity. We saw the latter as the union of the closed orbits of Iwasawa subgroups of the isometry group. As shown, this construction extends straightforwardly to higher dimensional cases, allowing to build a non trivial black hole on anti de Sitter spaces of arbitrary dimension $l \geqslant 3$. From this point of view, all anti de Sitter spaces of dimension $l \geqslant 3$ appear on an equal footing.

A natural question arising from this analysis is the following: *given a semisimple symmetric space, when does the set of closed orbits of the Iwasawa subgroups of the isometry group, seen as singularity, define a non-trivial causal structure?* We answered this question in the case of anti de Sitter spaces, using techniques allowing in principle for generalization to any semisimple symmetric space.

We also proved that performing a discrete quotient along the orbits of J_1 makes the resulting space causally inextensible (closed space-like curves appear in the singular part of the space), but we did not address questions like: are there other vector fields defining singularities (in the three dimensional case, we know that the answer is positive)? Can we identify a mass and an angular momentum from these hypothetic vectors? Are *all* BTZ black holes obtainable in this way in higher dimensions?

79.19 BTZ from the structure of $\mathfrak{so}(2, n)$

Abstract

In this section, we study the relevant structure of the algebra $\mathfrak{so}(2, n)$ which makes the BTZ black hole possible in the anti de Sitter space $AdS = SO(2, n)/SO(1, n)$. We pay a particular attention to the reductive Lie algebra structures of $\mathfrak{so}(2, n)$ and we study how this structure evolves when one increases the dimension.

We define the singularity as the closed orbits of the Iwasawa subgroup of the isometry group of anti de Sitter, but we insist on an alternative (closely related to the original conception of the BTZ black hole) way to describe the singularity as the loci where the norm of fundamental vector field vanishes. We provide a manageable Lie algebra oriented formula which describes the singularity and we use it in order to derive the existence of a black hole and to give a geometric description of the horizon.

79.20 Introduction

LONGSecSumStructExist

79.20.1 Anti de Sitter space and the BTZ black hole

The anti de Sitter space (hereafter abbreviated by AdS or AdS_l) is a static solution to the Einstein's equations that describes a universe without mass. It is widely studied in different context in mathematics as well as in physics.

The BTZ black hole in AdS, initially introduced in [1014, 1015] and then described and extended in various ways [1028, 1027, 1025], is an example of black hole structure which does not derives from a metric singularity, but from a causal issue.

The point of view we consider here grown from the papers [1013, 1017] in the case of AdS_3 and insists on the homogeneous space structure and the action of Iwasawa groups, in particular the singularity was described by looking at the closed orbits of the action of the Iwasawa group of $SO(2,2)$ on AdS_3. A dimensional generalization was then performed in [1019] (see also [1021] and the references therein for for a longer review).

One of the motivation in going that way is to embed the study of BTZ black hole into the noncommutative geometry and singleton physics [1030, 1020] since the deformation and the black hole are defined in a compatible way. Indeed the deformation is performed using the action of the same group as the groups which produce the singularity.

79.20.2 The way we describe the BTZ black hole

The anti de Sitter space AdS_l is the surface in $\mathbb{R}^{l+1}$ described by the equation

$$AdS_l \equiv t^2 + u^2 - x_1^2 - \ldots - x_{l-1}^2 = 1. \tag{79.189}$$

This space can be investigated in the framework of homogeneous spaces since it reads as the following quotient of groups:

$$AdS_l = \frac{\mathrm{SO}(2,l-1)}{\mathrm{SO}(1,l-1)} = G/H. \tag{79.190}$$

We denote by $\mathcal{G} = \mathfrak{so}(2,l-1)$ and $\mathcal{H} = \mathfrak{so}(1,l-1)$ the Lie algebras and by π the projection $G \to G/H$. The class of g will be written $[g]$ or $\pi(g)$. We choose an involutive automorphism $\sigma : \mathcal{G} \to \mathcal{G}$ which fixes elements of $\mathcal{H}$, and we call $\mathcal{Q}$ the eigenspace of eigenvalue -1 of σ. Thus we have the reductive decomposition

$$\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}. \tag{79.191}$$

LONGEqIntroRedDecompHQLieAlg

The compact part of $\mathrm{SO}(2,l-1)$ decomposes into $K = \mathrm{SO}(2) \times \mathrm{SO}(l-1)$. We denote by $\mathcal{K}$ the Lie algebra of K. Let θ be a Cartan involution which commutes with σ, and consider the corresponding Cartan decomposition

$$\mathcal{G} = \mathcal{K} \oplus \mathcal{P}, \tag{79.192}$$

LONGEqIntroRedDecompKPLieAlg

where $\mathcal{K}$ is the $+1$ eigenspace of θ and $\mathcal{P}$ is the -1 eigenspace. A maximal abelian algebra $\mathcal{A}$ in $\mathcal{P}$ has dimension two and one can choose a basis $\{J_1, J_2\}$ of $\mathcal{A}$ in such a way that $J_1 \in \mathcal{H}$ and $J_2 \in \mathcal{Q}$.

Now we consider an Iwasawa decomposition

$$\mathcal{G} = \mathcal{K} \oplus \mathcal{A} \oplus \mathcal{N}, \tag{79.193}$$

where $\mathcal{K}$ and $\mathcal{A}$ are the compact and abelian parts while $\mathcal{N}$ is the nilpotent part corresponding to a choice of positivity on the set of roots. We denote by $\mathcal{R}$ the Iwasawa component $\mathcal{R} = \mathcal{A} \oplus \mathcal{N}$. We are also going to use the algebra $\bar{\mathcal{N}} = \theta\mathcal{N}$ and the corresponding Iwasawa component $\bar{\mathcal{R}} = \mathcal{A} \oplus \bar{\mathcal{N}}$.

The Iwasawa groups $R = AN$ and $\bar{R} = A\bar{N}$ are naturally acting on anti de Sitter by $r[g] = [rg]$. It turns out that each of these two action has exactly two closed orbits, regardless to the dimension we are looking at. The first one is the orbit of the identity and the second one is the orbit of $[k_\theta]$ where k_θ is the element which generates the Cartan involution at the group level: $k_\theta g k_\theta^{-1} = \theta(g)$ ($g \in G$). In a suitable choice of matrix representation, the element k_θ is the block-diagonal element which is $-\mathbb{1}$ on $\mathrm{SO}(2)$ and $\mathbb{1}$ on $\mathrm{SO}(l-1)$. The $A\bar{N}$-orbits of $\mathbb{1}$ and k_θ are also closed. Moreover we have

$$\begin{aligned} [A\bar{N}k_\theta] &= [k_\theta AN] \\ [ANk_\theta] &= [k_\theta A\bar{N}] \end{aligned} \tag{79.194}$$

because A is invariant under the adjoint action of k_θ and, by definition, $\theta(N) = \bar{N}$. We define as **singular** the points of the closed orbits of AN and $A\bar{N}$ in AdS.

The Killing form of $\mathrm{SO}(2,l-1)$ induces a Lorentzian metric on AdS. The sign of the squared norm of a vector thus divides the vectors into three classes:

$$\begin{aligned} \|X\|^2 > 0 &\quad\to\quad \text{time like}, \\ \|X\|^2 < 0 &\quad\to\quad \text{space like}, \\ \|X\|^2 = 0 &\quad\to\quad \text{light like}. \end{aligned} \tag{79.195}$$

A geodesic is time (resp. space, light) like if its tangent vector is time like (resp. space, light).

If E_1 is a nilpotent element in $\mathcal{Q}$, then every nilpotent in $\mathcal{Q}$ are given by $\{\mathrm{Ad}(k)E_1\}_{k\in \mathrm{SO}(l-1)}$. These elements are also all the light like vectors at the base point. A light like geodesic trough the point $\pi(g)$ in the direction $\mathrm{Ad}(k)E_1$ is given by

$$\pi(ge^{s\,\mathrm{Ad}(k)E_1}). \tag{79.196}$$

One says that points with $s > 0$ are in the **future** of $\pi(g)$ while points with $s < 0$ are in the **past** of $\pi(g)$.

We say that a point in AdS_l belongs to the **black hole** if every light like geodesics trough that point intersect the singularity in the future. We call **horizon** the boundary of the set of points in the black hole. One says that there is a (non trivial) black hole structure when the horizon is non empty or, equivalently, when there are some points in the black hole, and some outside.

All these properties can be easily checked using the matrices given in [1021, 1019]. In this optic, I wrote a program using Sage[902] which checks all the properties that are shown in this paper. It will be published soon.

As far as notations are concerned, we denote by $X_{\alpha\beta}$ the basis of $\mathcal{N}$ and $\bar{\mathcal{N}}$ corresponding to our choice of Iwasawa decomposition. That basis is chosen in such a way that $\mathrm{ad}(J_1)X_{\alpha\beta} = \alpha X_{\alpha\beta}$ and $\mathrm{ad}(J_2)X_{\alpha\beta} = \beta X_{\alpha\beta}$.

79.20.3 Organization of the paper

The main goal of this paper is to reorganize all this structure in a coherent way. Then we use it efficiently in order to define the singularity of the BTZ black hole, to prove that one has a genuine black hole in every dimension, and to determine the horizon.

In section 79.21, we list the commutators of $\mathfrak{so}(2,n)$ with respect to its root spaces and we organize them in such a way to get a clear idea about the evolution of the structure when the dimension increases. We prove that, when one passes from $\mathfrak{so}(2,n)$ to $\mathfrak{so}(2,n+1)$, one gets four new vectors in the root spaces and that these are Killing-orthogonal to the vectors existing in $\mathfrak{so}(2,n)$ (this is the "dimensional slice" described in subsection 79.21.3).

We give in subsection 79.21.4 an original way to describe the space $\mathcal{Q}$ without reference to $\mathcal{H}$. The space $\mathcal{Q}$ is usually described as a complementary of $\mathcal{H}$. Here we show that it can be described by means of the root spaces and the Cartan involution θ. The space $\mathcal{H}$ is then described as $\mathcal{H} = [\mathcal{Q}, \mathcal{Q}]$. In some sense, we describe the quotient space $AdS = G/H$ directly by its tangent space $\mathcal{Q}$ without passing trough the definition of H. Of course, the knowledge of $\mathcal{H}$ will be of crucial importance later.

The subsection 79.21.5 is devoted to the proof of many properties of the decompositions $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$ and $\mathcal{G} = \mathcal{K} \oplus \mathcal{P}$.

The first important result is the proposition 79.59 that shows that the elements of $\mathcal{Q}$ are ad-conjugate to each others: there exist elements of the adjoint group which are intertwining the elements of $\mathcal{Q}$.

Using that property, we compute the norm of these elements and we identify the nilpotent vectors in $\mathcal{Q}$ (these are the light-like vectors). In the same time, we prove that the space G/H is Lorentzian and we provide an orthogonal basis of $\mathcal{Q}$.

The second central result is the fact that nilpotent elements in $\mathcal{Q}$ are of order two: if $E \in \mathcal{Q}$ is nilpotent, then $\mathrm{ad}(E)^3 = 0$. That result will be used in a crucial way in the proof of the black hole existence, as well as in the study of its properties.

In section 79.22, we define and study the structure of the BTZ black hole in the anti de Sitter space. First we identify the closed orbits of the Iwasawa group (theorem 79.92) and we define them as singular. In a second time, we provide an alternative description the singularity: theorem 79.95 shows that the singularity can be described as the loci of points at which a fundamental vector field has vanishing norm. We also provide in lemma 79.96 a convenient way to compute that norm on arbitrary point of the space.

We prove, in section 79.22.3, that our definition of singularity gives rise to a genuine black hole in the sense that there exists points from which some geodesics escape the singularity in the future

and there exists some points from which all the geodesics are intersecting the singularity in the future.

In section 79.24, we provide a geometric description of the horizon (theorem 79.110). We show that, if we see AdS_l as a subset of AdS_{l+1}, the space AdS_{l+1} is generated by the action on AdS_l by a one parameter group which leaves the singularity invariant. This action thus leaves invariant the whole causal structure and we are able to express the horizon in any dimension from the well known horizon in AdS_3.

79.21 Structure of the algebra

LONGSecProgressRidMatrices

79.21.1 The Iwasawa component

Our study of $AdS_l = \mathrm{SO}(2,l-1)/\mathrm{SO}(1,l-1)$ will be based on the properties of the algebra $\mathcal{G} = \mathfrak{so}(2,l-1)$ endowed with a Cartan involution θ and an Iwasawa decomposition $\mathcal{G} = \mathcal{A}\oplus\mathcal{N}\oplus\mathcal{K}$. In this section we want to underline the most relevant facts for our purpose. The part we are mainly interested in is the Iwasawa component $\mathcal{A}\oplus\mathcal{N}$ where LONGEqLeANEnDimAlg

$$\mathcal{N} = \{X^k_{+0}, X^k_{0+}, X_{++}, X_{+-}\} \tag{79.197a}$$
$$\mathcal{A} = \{J_1, J_2\}, \tag{79.197b}$$

where k runs [16] from 3 to $l-1$. The commutator table is LONGEqTableSOIwa

$$[X^k_{0+}, X^{k'}_{+0}] = \delta_{kk'}X_{++} \qquad [X^k_{0+}, X_{+-}] = 2X^k_{+0} \tag{79.198a}$$
$$[J_1, X^k_{+0}] = X^k_{+0} \qquad [J_2, X^k_{0+}] = X^k_{0+} \tag{79.198b}$$
$$[J_1, X_{+-}] = X_{+-} \qquad [J_2, X_{+-}] = -X_{+-} \tag{79.198c}$$
$$[J_1, X_{++}] = X_{++} \qquad [J_2, X_{++}] = X_{++}. \tag{79.198d}$$

We see that the Iwasawa algebra belongs to the class of j-algebras whose Pyatetskii-Shapiro decomposition is

$$\mathcal{A}\oplus\mathcal{N} = (\mathcal{A}_1 \oplus_{\mathrm{ad}} \mathcal{Z}_1) \oplus_{\mathrm{ad}} \big(\mathcal{A}_2 \oplus_{\mathrm{ad}} (V \oplus \mathcal{Z}_2)\big), \tag{79.199}$$

with

$$\mathcal{A}_1 = \langle H_1\rangle \qquad \mathcal{A}_2 = \langle H_2\rangle \tag{79.200a}$$
$$\mathcal{Z}_1 = \langle X_{+-}\rangle \qquad \mathcal{Z}_2 = \langle X_{++}\rangle \tag{79.200b}$$
$$V = \langle X^k_{0+}, X^k_{+0}\rangle_{k\geqslant 3} \tag{79.200c}$$

where

$$\begin{aligned} H_1 &= J_1 - J_2 \\ H_2 &= J_1 + J_2. \end{aligned} \tag{79.201}$$

The general commutators of such an algebra are LONGsubEqsGenPySO

$$[H_1, X_{+-}] = 2X_{+-} \qquad [H_2, X^k_{0+}] = X^k_{0+} \qquad [H_1, V] \subset V \tag{79.202a}$$
$$[H_2, X^k_{+0}] = X^k_{+0} \qquad [X_{+-}, V] \subset V \tag{79.202b}$$
$$[H_2, X_{++}] = 2X_{++} \tag{79.202c}$$
$$[X^k_{0+}, X^l_{+0}] = \Omega(X^k_{0+}, X^l_{+0})X_{++} \tag{79.202d}$$

In the case of $\mathfrak{so}(2,n)$, we have the following more precise relations: LONGSubEqsPlusPresPySO

$$[H_1, X^k_{0+}] = -X^k_{0+} \tag{79.203a}$$
$$[X_{+-}, X^k_{0+}] = -2X^k_{+0} \tag{79.203b}$$

16. The "new" vectors which appear in AdS_l with respect to AdS_{l-1} are $X^{l-1}_{0\pm}$ and $X^{l-1}_{\pm 0}$. Such an element appears for the first time in AdS_4 and is not present when one study AdS_3.

and the link between $\mathcal{N}$ and $\tilde{\mathcal{N}}$ is given by LONGSubEqsThethaPySO

$$
\begin{aligned}
[\theta X^k_{+0}, X_{++}] &= 2X^k_{0+} && (79.204a)\\
[\theta X^k_{0+}, X^k_{0+}] &= 2J_2 && (79.204b)\\
[\theta X_{++}, X_{++}] &= 4H_2 = 4(J_1 + J_2) && (79.204c)\\
[\theta X_{++}, X^k_{0+}] &= 2X^k_{-0} && (79.204d)\\
[\theta X_{+-}, X_{+-}] &= 4H_1 = 4(J_1 - J_2) && (79.204e)\\
[\theta X_{+-}, X^k_{+0}] &= 2X^k_{0+}. && (79.204f)
\end{aligned}
$$

The relations between the higher dimensional root spaces are

$$
\begin{aligned}
[X^i_{0+}, X^j_{-0}] &= -\delta_{ij} X_{-+}\\
[X^i_{0+}, X^j_{+0}] &= \delta_{ij} X_{++}\\
[X^i_{+0}, X^j_{0+}] &= -\delta_{ij} X_{++}\\
[X^i_{+0}, X^j_{0-}] &= \delta_{ij} X_{+-}.
\end{aligned}
\tag{79.205}
$$

The space $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ of the elements of $\mathcal{K}$ which commute with all the elements of $\mathcal{A}$ is given by the elements $r_{ij} = \frac{1}{2}[X^i_{0+}, X^j_{0+}]$ for every $i, j \geqslant 3$, $i \neq j$ realise the $\mathfrak{so}(l-1)$ algebra. We will say more about them in subsection 79.21.2.

We deduce the following relations that will prove useful later

$$
\begin{aligned}
[\theta X^k_{0+}, X_{++}] &= -2X^k_{+0}\\
[\theta X^k_{+0}, X_{+-}] &= -2X^k_{0-}\\
[\theta X_{++}, X^k_{+0}] &= -2X^k_{0-}\\
[X_{-+}, X^k_{0-}] &= -2X^k_{-0}.
\end{aligned}
\tag{79.206}
$$

79.21.2 The compact part

LONGSubSec_Thecompactpart

The compact part of $\mathfrak{so}(2, l-1)$, the algebra $\mathfrak{so}(2) \oplus \mathfrak{so}(l-1)$ is well known. What is interesting from our point of view is to write the commutation relations between the elements of $\mathfrak{so}(l-1)$ and the roots.

We define the following elements that are non vanishing:

$$
[X^i_{0+}, X^j_{0-}] = [X^i_{0-}, X^j_{0+}] = 2r_{ij}. \tag{79.207}
$$

One immediately has $\theta r_{ij} = r_{ij}$, so that $r_{ij} \in \mathcal{K}$. We also have $r_{ij} \in \mathcal{G}_0$ so that $r_{ij} \in \mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ and they act on the root spaces. The action is given by

$$
\begin{aligned}
[r_{ij}, X^k_{+0}] &= 0 && \text{if } i, j, k \text{ are different} && (79.208a)\\
[r_{ij}, X^k_{0+}] &= 0 && \text{if } i, j, k \text{ are different} && (79.208b)\\
[r_{ij}, X^j_{+0}] &= X^i_{+0} && \text{if } i \neq j && (79.208c)\\
[r_{ij}, X^j_{0+}] &= -X^i_{0+} && \text{if } i \neq j && (79.208d)\\
[r_{ij}, X_{\pm\pm}] &= 0. && && (79.208e)
\end{aligned}
$$

LONGEqComsRRN

LONGsubegrXpmpm

The elements r_{ij} satisfy the algebra of $\mathfrak{so}(n)$.

Remarque 79.52.
If σ is an involutive automorphism which commutes with θ and such that $\sigma J_1 = J_1$, $\sigma J_2 = -J_2$, then one has $\sigma r_{ij} = r_{ij}$. We will see later that this fact makes $r_{ij} \in \mathcal{H}$.

We know [17] that $\mathcal{G} = \mathcal{G}_0 \oplus \mathcal{N} \oplus \bar{\mathcal{N}}$ where $\mathcal{G}_0 = \mathcal{A} \oplus \mathcal{Z}_{\mathcal{K}}(\mathcal{A})$. Let us perform a dimension count in order to be sure that the vectors r_{ij} generate $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$. When we are working with AdS_l, we have

$$\begin{aligned} \dim(\mathcal{A}) &= 2 \\ \dim(\tilde{\mathcal{N}}_2) &= 4 \\ \dim(\bigoplus_k \tilde{\mathcal{N}}_k) &= 4(l-3) \\ \dim(\langle r_{ij} \rangle) &= \frac{1}{2}(l-4)(l-3). \end{aligned} \tag{79.209}$$

The last line comes from the fact that we have the elements r_{34}, r_{35},...,r_{45},... The first such element appears in AdS_5. Making the sum, we obtain $\frac{l(l+1)}{2}$, which is the dimension of $\mathfrak{so}(2, l-1)$. Thus we have

$$\mathcal{G} = \mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \oplus \mathcal{A} \oplus \tilde{\mathcal{N}}_2 \oplus \bigoplus_k \tilde{\mathcal{N}}_k \tag{79.210}$$

where $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ is generated by the elements r_{ij}.

79.21.3 Dimensional slices

LONGSubSecDimensionalSlices

Since the set $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ of elements in $\mathcal{K}$ that commute with elements of $\mathcal{A}$ is a part of $\mathcal{H}$ (see later), they will have almost no importance in the remaining [18]. The most important part of $\mathcal{G}$ is

$$\mathcal{A} \oplus \mathcal{N} \oplus \bar{\mathcal{N}} = \underbrace{\langle J_1, J_2, X_{\pm,\pm} \rangle}_{\text{for every dimension}} \oplus \underbrace{\langle X^4_{0\pm}, X^4_{\pm 0} \rangle}_{\text{for } \mathfrak{so}(2,\geqslant 3)} \oplus \ldots \oplus \underbrace{\langle X^l_{0\pm}, X^l_{\pm 0} \rangle}_{\text{for } \mathfrak{so}(2,l-1)}. \tag{79.211}$$

LONGEqDecompDimQuim

We use the following notations in order to make more clear how does the algebra evolve when one increases the dimension:

$$\begin{aligned} \mathcal{N}_2 &= \langle X_{+-}, X_{++} \rangle, & \mathcal{N}_k &= \langle X^k_{0+}, X^k_{+0} \rangle \\ \bar{\mathcal{N}}_2 &= \langle X_{-+}, X_{--} \rangle, & \bar{\mathcal{N}}_k &= \langle X^k_{0-}, X^k_{-0} \rangle \\ \tilde{\mathcal{N}}_2 &= \langle \mathcal{N}_2, \bar{\mathcal{N}}_2 \rangle, & \tilde{\mathcal{N}}_k &= \langle \mathcal{N}_k, \bar{\mathcal{N}}_k \rangle \end{aligned} \tag{79.212}$$

for $k \geqslant 3$. The relations are

$$\begin{aligned} [\tilde{\mathcal{N}}_2, \tilde{\mathcal{N}}_2] &\subseteq \mathcal{A} & [\tilde{\mathcal{N}}_2, \tilde{\mathcal{N}}_k] &\subseteq \tilde{\mathcal{N}}_k \\ [\tilde{\mathcal{N}}_k, \tilde{\mathcal{N}}_k] &\subseteq \mathcal{A} \oplus \tilde{\mathcal{N}}_2 & [\tilde{\mathcal{N}}_k, \tilde{\mathcal{N}}_{k'}] &\subset \mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \end{aligned} \tag{79.213}$$

LONGEqsCommWithtsNDeuxkA

As a consequence of the splitting and the commutation relations, we have many Killing-orthogonal subspaces in $\mathfrak{so}(2, l-1)$:

$$\tilde{\mathcal{N}}_k \perp \tilde{\mathcal{N}}_{k'} \tag{79.214a}$$

LONGEqtsnkperprtsnkp

$$\mathcal{A} \perp \tilde{\mathcal{N}}_2 \tag{79.214b}$$

LONGEqAperpNTrois

$$\mathcal{A} \perp \tilde{\mathcal{N}}_k. \tag{79.214c}$$

LONGEqAperpNk

$$\tilde{\mathcal{N}}_2 \perp \tilde{\mathcal{N}}_k \tag{79.214d}$$

LONGEqNTtroisperpNk

In order to check there relations, look at the trace of action of $\mathrm{ad}(X) \circ \mathrm{ad}(Y)$ on the various spaces:

$$\mathrm{ad}(\mathcal{A}) \circ \mathrm{ad}(\tilde{\mathcal{N}}_2) : \begin{cases} \mathcal{A} \to \tilde{\mathcal{N}}_2 \to \tilde{\mathcal{N}}_2 \\ \tilde{\mathcal{N}}_2 \to \mathcal{A} \to 0 \\ \tilde{\mathcal{N}}_k \to \tilde{\mathcal{N}}_k \to \tilde{\mathcal{N}}_k \\ \mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \to \tilde{\mathcal{N}}_2 \to \tilde{\mathcal{N}}_2, \end{cases} \tag{79.215}$$

17. See [854] for example.
18. We will however need them in the computation of the coefficients (79.445).

and

$$\mathrm{ad}(\mathcal{A})\circ\mathrm{ad}(\tilde{\mathcal{N}}_k)\colon \begin{cases} \mathcal{A}\to\tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k \\ \tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k \\ \tilde{\mathcal{N}}_k\to\mathcal{A}\oplus\tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_2 \\ \langle r_{kl}\rangle\to\tilde{\mathcal{N}}_l\to\tilde{\mathcal{N}}_l \end{cases} \tag{79.216}$$

and

$$\mathrm{ad}(\tilde{\mathcal{N}}_2)\circ\mathrm{ad}(\tilde{\mathcal{N}}_k)\colon \begin{cases} \mathcal{A}\to\tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k \\ \tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k \\ \tilde{\mathcal{N}}_k\to\mathcal{A}\oplus\tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_2\oplus\mathcal{A} \\ \langle r_{kl}\rangle\to\tilde{\mathcal{N}}_l\to\tilde{\mathcal{N}}_l\oplus\mathcal{A} \end{cases} \tag{79.217}$$

and

$$\mathrm{ad}(J_1)^2|_{\tilde{\mathcal{N}}_2}=\mathrm{ad}(J_2)^2|_{\tilde{\mathcal{N}}_2}=\mathrm{Id}\,|_{\tilde{\mathcal{N}}_2}. \tag{79.218}$$

As a consequence,

$$\mathcal{G}=\mathcal{Z}_{\mathcal{K}}(\mathcal{A})\oplus\mathcal{A}\oplus\tilde{\mathcal{N}}_2\bigoplus_{k\geqslant 3}\tilde{\mathcal{N}}_k \tag{79.219}$$

LONGEqDecompGPourKillOrtho

is a Killing-orthogonal decomposition of $\mathfrak{so}(2,l-1)$.

79.21.4 Reductive decomposition

LONGSubSecReductiveDecompQ

Related to the decompositions (79.191) and (79.192) we introduce the following notations: $X_{\mathcal{K}}$, $X_{\mathcal{P}}$ will denote the $\mathcal{K}$ and $\mathcal{P}$-components of the vector X with respect to the decomposition $\mathcal{G}=\mathcal{K}\oplus\mathcal{P}$. We define $X_{\mathcal{H}}$ and $X_{\mathcal{Q}}$ similarly with respect to the decomposition $\mathcal{G}=\mathcal{H}\oplus\mathcal{Q}$. In the same way we define $X_{\mathcal{A}}$, $X_{\mathcal{Z}(\mathcal{K})}$ and so on by the Iwasawa decomposition

$$\mathcal{G}=\mathcal{A}\oplus\mathcal{N}\oplus\mathcal{K}=\mathcal{A}\oplus\mathcal{N}\oplus\mathcal{Z}(\mathcal{K})\oplus\mathfrak{so}(l-1). \tag{79.220}$$

Let $\mathcal{Q}$ be the following vector subspace of $\mathcal{G}$:

$$\mathcal{Q}=\left\langle\mathcal{Z}(\mathcal{K}),J_2,[\mathcal{Z}(\mathcal{K}),J_1],(X^k_{0+})_{\mathcal{P}}\right\rangle_{k\geqslant 3}. \tag{79.221}$$

LONGEqDecQEspacesCools

Then we choose a subalgebra $\mathcal{H}$ of $\mathcal{G}$ which, as vector space, is a complementary of $\mathcal{Q}$. In that choice, we require that there exists an involutive automorphism $\sigma\colon\mathcal{G}\to\mathcal{G}$ such that

$$\sigma=(\mathrm{Id})_{\mathcal{H}}\oplus(-\mathrm{Id})_{\mathcal{Q}}. \tag{79.222}$$

In that case the decomposition $\mathcal{G}=\mathcal{H}\oplus\mathcal{Q}$ is reductive, i.e. $[\mathcal{Q},\mathcal{Q}]\subset\mathcal{H}$ and $[\mathcal{Q},\mathcal{H}]\subset\mathcal{Q}$.

From definition (79.221), it is immediately apparent that one has a basis of $\mathcal{Q}$ made of elements in $\mathcal{K}$ and $\mathcal{P}$, so that one immediately has

$$[\sigma,\theta]=0. \tag{79.223}$$

If $X\in\mathcal{G}$, the projections are given by

$$\begin{aligned} X_{\mathcal{H}}&=\frac{1}{2}(X+\sigma X), & X_{\mathcal{K}}&=\frac{1}{2}(X+\theta X), \\ X_{\mathcal{Q}}&=\frac{1}{2}(X-\sigma X), & X_{\mathcal{P}}&=\frac{1}{2}(X-\theta X). \end{aligned} \tag{79.224}$$

LONGEqProjHQPKsigmatheta

In particular $\theta\mathcal{H}\subset\mathcal{H}$ since θ and σ commute.

We introduce the following elements of $\mathcal{Q}$:

$$\begin{aligned} q_0&=(X_{++})_{\mathcal{K}\mathcal{Q}} \\ q_1&=J_2 \\ q_2&=-[J_1,q_0]=-(X_{++})_{\mathcal{P}\mathcal{Q}} \\ q_k&=(X^k_{0+})_{\mathcal{P}}. \end{aligned} \tag{79.225}$$

LONGAlignPRedDefQQ

In order to see that $[J_1, q_0] = (X_{++})_{\mathcal{PQ}}$, just write the $\mathcal{PQ}$-component of the equality $[J_1, X_{++}] = X_{++}$. We will prove later that this is a basis and that each of these elements correspond to one of the spaces listed in (79.221).

Since $X_{\mathcal{P}} = (\theta X - X)/2$, we have

$$[q_i, q_j] = -\frac{1}{4}\big([X^i_{0+}, X^j_{0-}] + [X^i_{0-}, X^j_{0+}]\big) = r_{ij}. \tag{79.226}$$

From equations (79.214b) and (79.214c), we have $q_1 \perp q_2$ and $q_2 \perp q_k$. Using the other perpendicularity relations $\mathcal{K} \perp \mathcal{P}$ and (79.214a), (79.214b), (79.214c), we see that the set $\{q_i\}_{0 \leqslant i \leqslant l-1}$ is orthogonal.

The space $\mathcal{H}$ is defined as generated by the elements

$$\begin{array}{ll} J_1 & r_k = [J_2, q_k] \\ p_1 = [q_0, q_1] & p_k = [q_0, q_k] \\ s_1 = [J_1, p_1] & s_k = [J_1, p_k]. \end{array} \tag{79.227}$$

LONGAlignPremDefHH

Elements (79.225) and (79.227) will be studied in great details later.

79.21.4.1 Remark on the compact part

LONGSubSubSecRemCompPart

Elements of $\mathcal{K}$ are elements of the form $X + \theta X$. A part of the elements inside $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$, these elements are of two kinds:

$$X_{++} + X_{--} \tag{79.228a}$$
$$X_{+-} + X_{-+} \tag{79.228b}$$

on the one hand, and

$$X^k_{0+} + X^k_{0-} \tag{79.229a}$$
$$X^k_{+0} + X^k_{-0} \tag{79.229b}$$

on the other hand. The first two are commuting, so that $\mathcal{Z}(\mathcal{K})$ is two dimensional when one studies AdS_3. That correspond to the well known fact that the compact part of $\mathfrak{so}(2,2)$ is $\mathfrak{so}(2) \oplus \mathfrak{so}(2)$ which is abelian. These elements, however, do not commute with the two other. For example, the combination

$$\big(X_{++} + X_{--}\big) - \big(X_{+-} + X_{-+}\big) \tag{79.230}$$

does not commute with the elements of the second type. Now, one checks that the combination

$$\big(X_{++} + X_{--}\big) + \big(X_{+-} + X_{-+}\big) \tag{79.231}$$

LONGEqZKChoixblabl

commutes with all the other, so that it is the generator of $\mathcal{Z}(\mathcal{K})$ for $AdS_{\geqslant 4}$. This corresponds to the fact that the compact part of $\mathfrak{so}(2,n)$ is $\mathfrak{so}(2) \oplus \mathfrak{so}(n)$. In other terms,

$$\mathcal{Z}(\mathcal{K}) = \langle X_{++} + X_{--} + X_{+-} + X_{-+} \rangle \oplus \underbrace{\langle X_{++} + X_{--} - X_{+-} - X_{-+} \rangle}_{\text{only for } AdS_3}. \tag{79.232}$$

LONGEqDecompZKenDeuxSuivantDim

Notice that, for $AdS_{\geqslant 4}$, we can define $q_0 = (X_{++})_{\mathcal{Z}(\mathcal{K})}$ as $\mathcal{K} = \mathfrak{so}(2) \oplus \mathfrak{so}(l-2)$ for AdS_l. The case of AdS_3 is particular because $\mathcal{Z}(\mathcal{K})$ is of dimension two and we have to set by hand what part of $\mathcal{Z}(\mathcal{K})$ belongs to $\mathcal{Q}$ (the other part belongs to $\mathcal{H}$). From what is said around equation (79.231), we know that q_0 is a multiple of $X_{++} + X_{--} + X_{+-} + X_{-+}$.

Dimension counting shows that $\dim \mathcal{Q} = l$ and general theory of homogeneous spaces shows that $\mathcal{Q}$ has to be seen as the tangent space of the manifold G/H.

79.21.5 Properties of the reductive decompositions

LONGSubSecPropRedDecompQH

We are considering the two reductive decompositions $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q} = \mathcal{K} \oplus \mathcal{P}$ in the same time as the root space decomposition (79.219). We are now giving some properties of them.

We know that $\mathcal{K} \cap \mathcal{Q} = \langle q_0 \rangle$ belongs to $\tilde{\mathcal{N}}_2$. As a consequence, the elements $X^k_{\alpha 0}$ and $X^k_{0\alpha}$ have no $\mathcal{KQ}$-components and

$$\begin{aligned} [J_2, \tilde{\mathcal{N}}_k]_{\mathcal{PH}} &= 0 \\ [J_2, \tilde{\mathcal{N}}_k]_{\mathcal{PQ}} &= 0 \\ \mathtt{proj}_{\mathcal{PQ}} X^k_{\alpha 0} &= 0 \\ \mathtt{proj}_{\mathcal{PH}} X^k_{0\alpha} &= 0. \end{aligned} \tag{79.233}$$

Since $X^k_{\alpha 0}$ and $X^k_{0\alpha}$ are not eigenvectors of θ, they have a non vanishing $\mathcal{P}$-component. We deduce that

$$\begin{aligned} \mathtt{proj}_{\mathcal{PH}} X^k_{\alpha 0} &\neq 0 \\ \mathtt{proj}_{\mathcal{PQ}} X^k_{0\alpha} &\neq 0. \end{aligned} \tag{79.234}$$

LONGEqXalphazeroaduPH

As a consequence of compatibility between θ and σ, we have

$$\begin{aligned} [J_1, (X_{\alpha\beta})_{\mathcal{H}}] &= \alpha (X_{\alpha\beta})_{\mathcal{H}} \\ [J_1, (X_{\alpha\beta})_{\mathcal{Q}}] &= \beta (X_{\alpha\beta})_{\mathcal{Q}} \end{aligned} \tag{79.235}$$

and

$$\begin{aligned} [J_2, (X_{\alpha\beta})_{\mathcal{H}}] &= \beta (X_{\alpha\beta})_{\mathcal{Q}} \\ [J_2, (X_{\alpha\beta})_{\mathcal{Q}}] &= \alpha (X_{\alpha\beta})_{\mathcal{H}}. \end{aligned} \tag{79.236}$$

LONGSubEqsJdeuxXalphaneta

So $X_{\mathcal{Q}}$ itself is an eigenvector of $\mathrm{ad}(J_1)$. In the same way, we prove that

$$\begin{aligned} [J_1, (X_{\alpha\beta})_{\mathcal{P}}] &= \alpha (X_{\alpha\beta})_{\mathcal{K}} \\ [J_1, (X_{\alpha\beta})_{\mathcal{K}}] &= \alpha (X_{\alpha\beta})_{\mathcal{P}} \end{aligned} \tag{79.237}$$

LONGEqJUnXabPK

because $J_1 \in \mathcal{P}$.

LONGCorHPHKQPQKXuu

Corollary 79.53.
The vector X_{++} has non vanishing components in $\mathcal{H} \cap \mathcal{P}$, $\mathcal{H} \cap \mathcal{K}$, $\mathcal{Q} \cap \mathcal{P}$ and $\mathcal{Q} \cap \mathcal{K}$.

Proof. Since $\mathrm{ad}(J_2)$ inverts the $\mathcal{H}$ and $\mathcal{Q}$-components of X_{++} (equation (79.236)), they must be both non zero. In the same way $\mathrm{ad}(J_1)$ inverts the components $\mathcal{P}$ and $\mathcal{K}$ of vectors of $\mathcal{H}$ and $\mathcal{Q}$ (equations (79.237)). □

LONGLEmDesZPP

Lemma 79.54.
We have $(X^k_{0+})_{\mathcal{KQ}} = (X^k_{0+})_{\mathcal{PH}} = 0$ and consequently, $(X^k_{0+})_{\mathcal{P}} = (X^k_{0+})_{\mathcal{Q}}$.

Proof. Consider the decomposition of the equality $[J_1, X^k_{0+}] = 0$ into components $\mathcal{PQ}$, $\mathcal{PH}$, $\mathcal{KQ}$, $\mathcal{KH}$. Since $J_1 \in \mathcal{P} \cap \mathcal{H}$, the $\mathcal{KH}$ and $\mathcal{PQ}$ components are

$$[J_1, (X^k_{0+})_{\mathcal{PH}}] = 0 \tag{79.238a}$$

$$[J_1, (X^k_{0+})_{\mathcal{KQ}}] = 0. \tag{79.238b}$$

In the same way, using the fact that $J_2 \in \mathcal{P} \cap \mathcal{Q}$, we have

LONGEqDeuxJUDzpKq

$$[J_2, (X^k_{0+})_{\mathcal{PH}}] = (X^k_{0+})_{\mathcal{KQ}} \tag{79.239a}$$

$$[J_2, (X^k_{0+})_{\mathcal{KQ}}] = (X^k_{0+})_{\mathcal{PH}}. \tag{79.239b}$$

LONGsubEqDeuxJUDzpKqb

Since $\dim(\mathcal{K} \cap \mathcal{Q}) = 1$, the component $(X^k_{0+})_{\mathcal{KQ}}$ has to be a multiple of $(X_{++})_{\mathcal{KQ}}$. Thus we have

$$0 = [J_1, (X^k_{0+})_{\mathcal{KQ}}] = \lambda [J_1, (X_{++})_{\mathcal{KQ}}] = \lambda (X_{++})_{\mathcal{PQ}}, \tag{79.240}$$

but $(X_{++})_{\mathcal{PQ}} \neq 0$, thus $\lambda = 0$ and we conclude that $(X^k_{0+})_{\mathcal{KQ}} = 0$. Now, equation (79.239b) shows that $(X^k_{0+})_{\mathcal{PH}} = 0$. □

LONGLemSigmaXzpBien

Lemma 79.55.
We have $\sigma X^k_{0+} = X^k_{0-}$.

Proof. Since we know that $\sigma\mathcal{G}_{\alpha\beta} \subset \mathcal{G}_{\alpha,-\beta}$, the work is to fix the sign in

LONGEqsigmaXzppmthetaXzp
$$\sigma X^k_{0+} = \pm X^k_{0-} = \pm\theta X^k_{0+}. \tag{79.241}$$

Lemma 79.54 states that $(X^k_{0+})_{\mathcal{P}} = (X^k_{0+})_{\mathcal{Q}}$. Thus the $\mathcal{Q}$-component of θX^k_{0+} is $-(X^k_{0+})_{\mathcal{Q}}$, which is also equal to the $\mathcal{Q}$-component of $\sigma(X^k_{0+})$. That fixes the choice of sign in equation (79.241). □

The following is an immediate corollary of lemma 79.55 and the fact that θ fixes $\mathcal{P}$ and $\mathcal{K}$ while σ fixes $\mathcal{H}$ and $\mathcal{Q}$.

LONGCorXzpHQPKXzm

Corollary 79.56.
We have

$$(X^k_{0+})_{\mathcal{H}} = (X^k_{0-})_{\mathcal{H}} \tag{79.242a}$$
$$(X^k_{0+})_{\mathcal{Q}} = -(X^k_{0-})_{\mathcal{Q}} \tag{79.242b}$$
$$(X^k_{0+})_{\mathcal{P}} = -(X^k_{0-})_{\mathcal{P}} \tag{79.242c}$$
$$(X^k_{0+})_{\mathcal{K}} = (X^k_{0-})_{\mathcal{K}}. \tag{79.242d}$$

Proof. Since σ acts as the identity on $\mathcal{H}$ and changes the sign on $\mathcal{Q}$, we have

$$\sigma X^k_{0+} = \sigma\big((X^k_{0+})_{\mathcal{H}} + (X^k_{0+})_{\mathcal{Q}}\big) = (X^k_{0+})_{\mathcal{H}} - (X^k_{0+})_{\mathcal{Q}}, \tag{79.243}$$

but lemma 79.55 states that $\sigma X^k_{0+} = X^k_{0-} = (X^k_{0-})_{\mathcal{H}} + (X^k_{0-})_{\mathcal{Q}}$. Equating the $\mathcal{H}$ and $\mathcal{Q}$-component of these two expressions of σX^k_{0+} brings the two first equalities.

The two other are proven the same way. We know that $\theta X^k_{0+} = X^k_{0-}$, but

$$\theta X^k_{0+} = \theta\big((X^k_{0+})_{\mathcal{P}} + (X^k_{0+})_{\mathcal{K}}\big) = -(X^k_{0+})_{\mathcal{P}} + (X^k_{0+})_{\mathcal{K}}. \tag{79.244}$$

The two last relations follow. □

79.21.5.1 An interesting basis of $\mathcal{Q}$

LONGSubSubSecInterestingBasisQ

Being the tangent space of AdS, the space $\mathcal{Q}$ is of a particular importance. Let us now have a closer look at the vectors that we already mentioned in equations (79.225):

LONGEqBasQQzi
$$q_0 = (X_{++})_{\mathcal{KQ}} \tag{79.245a}$$
$$q_1 = J_2 \tag{79.245b}$$
$$q_2 = -[J_1, q_0] = -(X_{++})_{\mathcal{PQ}} \tag{79.245c}$$
$$q_k = (X^k_{0+})_{\mathcal{Q}} \qquad \text{lemma 79.54.} \tag{79.245d}$$

By lemma 79.54, and the discussion about $\mathcal{Z}(\mathcal{K})$ (equation (79.232)), we can express the elements q_i without explicit references to $\mathcal{Q}$ itself and each element corresponds to a particular space (once again, the choice of q_0 is not that simple in AdS_3):

LONGEqBasQQziPlusMieux
$$q_0 = (X_{++})_{\mathcal{Z}(\mathcal{K})} \qquad \in \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 \tag{79.246a}$$
$$q_1 = J_2 \qquad \in \mathcal{Q} \cap \mathcal{A} \tag{79.246b}$$
$$q_2 = -[J_1, q_0] \qquad \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 \tag{79.246c}$$
LONGsubEqJUnQZQDeux
$$q_k = (X^k_{0+})_{\mathcal{P}} \qquad \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k. \tag{79.246d}$$

These elements correspond to the expression (79.221). The compact part isomorphic to $\mathfrak{so}(2, l-1)$ is then generated by the elements

Remarque 79.57.
The space $\mathcal{Z}(\mathcal{K})$ is given by the structure of the compact part of $\mathfrak{so}(2,n)$, the elements $(X_{0+}^k)_{\mathcal{P}}$ are defined from the root space structure of $\mathfrak{so}(2,n)$ and the Cartan involution. The elements J_1 and J_2 are a basis of $\mathcal{A}$. However, we need to know $\mathcal{H}$ in order to distinguish J_1 from J_2 that are respectively generators of $\mathcal{A}_{\mathcal{H}}$ and $\mathcal{A}_{\mathcal{Q}}$.

Thus the basis (79.246) is given in a way almost independent of the choice of $\mathcal{H}$.

LONGCorQdansPetK

Corollary 79.58.
We have $q_0 \in \mathcal{K}$ and $q_i \in \mathcal{P}$ if $i \neq 0$. The set $\{q_0, q_1, \dots, q_l\}$ is a basis of $\mathcal{Q}$. Moreover we have $\mathcal{Q} \cap \tilde{\mathcal{N}}_k = \langle q_k \rangle$.

Proof. The first claim is a direct consequence of the expressions (79.246). Linear independence is a direct consequence of the spaces to which each vector belongs. A dimensional counting shows that it is a basis of $\mathcal{Q}$. □

79.21.5.2 Magic intertwining elements

The vectors q_i have the property to be intertwined by some elements of $\mathcal{H}$. Namely, the adjoint action of the elements LONGAlignDefMagicIntert

$$X_1 = p_1 = -[J_2, q_0] \qquad \in \mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 \tag{79.247a}$$
$$X_2 = s_1 = [J_1, X_1] \qquad \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 \tag{79.247b}$$
$$X_k = -r_k = -[J_2, q_k] \qquad \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k \tag{79.247c}$$ LONGEqDefXkCommeComm

intertwines the q_i's in the sense of the following proposition.

LONGXUnALaTwistingSuperCool

Proposition 79.59 (Intertwining properties)**.**
The elements defined by equation (79.247) satisfy

LONGEqCalculBBBJUnUnNirme

$$\mathrm{ad}(J_1)q_0 = -q_2 \tag{79.248a}$$ LONGEqCalculBBBJUnUnNirmeA
$$\mathrm{ad}(J_1)q_2 = -q_0. \tag{79.248b}$$ LONGEqCalculBBBJUnUnNirmeB

$$\mathrm{ad}(X_2)q_2 = q_1 \tag{79.250a}$$ LONGSubEqXdeuxQdeuxa
$$\mathrm{ad}(X_2)q_1 = -q_2 \tag{79.250b}$$ LONGSubEqXdeuxQun

LONGEqSubEqbXUnqZero

$$\mathrm{ad}(X_1)q_1 = q_0 \tag{79.249a}$$ LONGSubEqbXZeroqUn
$$\mathrm{ad}(X_1)q_0 = q_1, \tag{79.249b}$$ LONGSubEqbXUnqZero

LONGEqSubEqbXkqZero

$$\mathrm{ad}(X_k)q_k = -q_1. \tag{79.251a}$$ LONGSubEqbXUmoinsqZero
$$\mathrm{ad}(X_k)q_1 = q_k \tag{79.251b}$$ LONGSubEqbXkQunQk

Proof. Equation (79.248a) is by definition while equation (79.248b) follows from the first one and the fact that $\mathrm{ad}(J_1)^2$ acts as the identity on $\tilde{\mathcal{N}}_2$.

The equality (79.249a) is a direct consequence of the fact that $\mathrm{ad}(J_2)^2$ is the identity on $\tilde{\mathcal{N}}_2$, so that

$$[X_1, q_1] = -\big[[J_2, q_0], q_1\big] = \mathrm{ad}(J_2)^2 q_0 = q_0. \tag{79.252}$$

For the relation (79.249b), first remark that, since $q_0 = (X_{++})_{\mathcal{KQ}}$, we have

$$X_1 = -(X_{++})_{\mathcal{PH}} \tag{79.253}$$

and we have to compute

$$[X_1, q_0] = -\big[(X_{++})_{\mathcal{PH}}, (X_{++})_{\mathcal{KQ}}\big] \tag{79.254}$$

Using the projections (79.224), we have

$$\begin{aligned}(X_{++})_{\mathcal{PH}} &= \frac{1}{4}(X_{++} + \sigma X_{++} - \theta X_{++} - \sigma\theta X_{++}) \\ (X_{++})_{\mathcal{KQ}} &= \frac{1}{4}(X_{++} - \sigma X_{++} + \theta X_{++} - \sigma\theta X_{++})\end{aligned} \tag{79.255}$$

We compute the commutator taking into account the facts that σ is an automorphism and that, for example, $[X_{++}, \sigma X_{++}] = 0$ because $\sigma X_{++} \in \mathcal{G}_{(+-)}$. What we find is

$$\big[(X_{++})_{\mathcal{PH}}, (X_{++})_{\mathcal{KP}}\big] = \frac{1}{4}\frac{1}{2}\Big([X_{++}, \theta X_{++}] - \sigma[X_{++}, \theta X_{++}]\Big) = \frac{1}{4}[X_{++}, \theta X_{++}]_{\mathcal{Q}}. \tag{79.256}$$

Since $[X_{++}, X_{--}] = -4(J_1 + J_2)$, we have $[X_1, q_0] = J_2 = q_1$ as expected.

For equation (79.250a) we use the Jacobi relation and the relation (79.248b).

$$\begin{aligned}[q_2, X_2] &= \big[q_2, [J_1, p_1]\big] \\ &= -\big[J_1, [p_1, q_2]\big] - \big[p_1, [q_2, J_1]\big] \\ &= -[p_1, q_0] \\ &= -q_1\end{aligned} \tag{79.257}$$

For equation (79.250b), we use the definition of X_2, the Jacobi identity and the facts that $[p_1, J_2] = q_0$ and $[J_1, q_0] = -q_2$.

We pass now to the fourth pair of intertwining relations. By definition, $q_k = (X^k_{0+})_{\mathcal{P}}$, but taking into account the fact that $J_2 \in \mathcal{P}$ we can decompose the relation $[J_2, X_{0+}] = X_{0+}$ into

$$[J_2, (X_{0+})_{\mathcal{P}}] = (X_{0+})_{\mathcal{K}} \tag{79.258a}$$

$$[J_2, (X_{0+})_{\mathcal{K}}] = (X_{0+})_{\mathcal{P}}. \tag{79.258b}$$

LONGEqJdeuxXzpsPsK

Equation (79.258a) tolds us that

$$X_k = -(X_{0+})_{\mathcal{K}}. \tag{79.259}$$

Now we have to compute $[X_k, q_k] = -\big[(X_{0+})_{\mathcal{K}}, (X_{0+})_{\mathcal{P}}\big]$. We know that $[X_{0+}, X_{0-}] = -2J_2 \in \mathcal{P}$. Thus corollary 79.56 brings

$$-2J_2 = \big[(X_{0+})_{\mathcal{K}}, (X_{0-})_{\mathcal{P}}\big] + \big[(X_{0+})_{\mathcal{P}}, (X_{0-})_{\mathcal{K}}\big] = -2\big[(X_{0+})_{\mathcal{K}}, (X_{0+})_{\mathcal{P}}\big] = 2[X_k, q_k], \tag{79.260}$$

and the result follows.

For equation (79.251b), we have to compute $[X_k, q_1] = [J_2, (X^k_{+0})_{\mathcal{K}}]$. The $\mathcal{P}$-component of $[J_2, X^k_{0+}] = X^k_{0+}$ is exactly

$$[J_2, (X^k_{0+})_{\mathcal{K}}] = (X^k_{0+})_{\mathcal{P}} = q_k. \tag{79.261}$$

□

These intertwining relations will be widely used in computing the norm of the vectors q_i in proposition 79.62 as well as in some other occasions.

Let us now give a few words about the existence and unicity of these elements. The fact that there exists an element X_1 such that $\mathrm{ad}(J_2)X_1 = q_0$ comes from the decomposition (79.276) and the fact that each $X_{\pm\pm}$ is an eigenvector of $\mathrm{ad}(J_2)$. It is thus sufficient to adapt the signs in order to manage a combination of X_{++}, X_{+-}, X_{-+} and X_{--} on which the adjoint action of J_2 creates q_0. However, the fact that this element has in the same time the "symmetric" property $\mathrm{ad}(X_1)q_0 = q_1$ could seem a miracle. See theorem 79.86.

Lemma 79.60.
An element X_1 such that $\mathrm{ad}(X_1)q_1 = q_0$ can be chosen in $\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2$. Moreover, this choice is unique up to normalisation.

Proof. Unicity is nothing else than the fact that $\dim(\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2) = 1$. Indeed, since $\mathcal{G} = \mathcal{A} \oplus \tilde{\mathcal{N}} \oplus \mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ and $\mathcal{A} \subset \mathcal{P}$, we have $\mathcal{K} \subset \tilde{\mathcal{N}}$. Dimension counting shows that $\dim(\tilde{\mathcal{N}}_2 \cap \mathcal{H}) = 2$ (because $\dim(\tilde{\mathcal{N}}_2) = 4$ and $q_0, q_2 \in \mathcal{Q} \cap \tilde{\mathcal{N}}_2$). As we are looking in $\tilde{\mathcal{N}}_2$, we are limited to elements in $\mathfrak{so}(2,2)$ (not the higher dimensional slices), so that we can consider $\mathcal{K} = \mathfrak{so}(2) \oplus \mathfrak{so}(2)$. One of these two $\mathfrak{so}(2)$ factors belongs to $\mathcal{H}$, so that $\dim(\mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2) = 1$ and finally $\dim(\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2) = 1$.

Let now X_1 be such that $[X_1, q_1] = q_0$. If X_1 has a component in $\mathcal{Q}$, that component has to commute with q_1 (if not, the commutator $[X_1, q_1]$ would have a $\mathcal{H}$-component). So we can redefine X_1 in order to have $X_1 \in \mathcal{H}$.

In the same way, a $\mathcal{A}$-component has to be J_1 (because $J_2 \in \mathcal{Q}$) which commutes with q_1. We redefine X_1 in order to remove its J_1-component. We remove a component in $\tilde{\mathcal{N}}_k$ because $[\tilde{\mathcal{N}}_2, \tilde{\mathcal{N}}_k] \subset \tilde{\mathcal{N}}_k$, and a $\mathcal{K}$-component can also be removed since its commutator with q_1 would produce a $\mathcal{P}$-component. We showed that $X_1 \in \mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2$. □

LONGLemChoixDeXk

Lemma 79.61.
An element X_k such that $\mathrm{ad}(X_k)q_1 = q_k$ can be chosen in $\mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k$.

Proof. The proof is elementary in tree steps using the fact that $q_1 \in \mathcal{P} \cap \mathcal{Q} \cap \mathcal{A}$:

(1) A $\mathcal{P}$-component can be annihilated because $[\mathcal{P}, \mathcal{P}] \subset \mathcal{K}$ while $q_k \in \mathcal{P}$,

(2) a $\mathcal{Q}$-component can be annihilated because $[\mathcal{Q}, \mathcal{Q}] \subset \mathcal{H}$ while $q_k \in \mathcal{Q}$,

(3) if $k' \neq k$, a $\tilde{\mathcal{N}}_{k'}$-component can be annihilated because $[\tilde{\mathcal{N}}_{k'}, \mathcal{A}] \subset \tilde{\mathcal{N}}_{k'}$ while $q_k \in \tilde{\mathcal{N}}_k$.

□

79.21.5.3 Killing form and orthogonality

We *define* the norm of an element in $\mathcal{G} = \mathfrak{so}(2, n)$ as

$$\|X\| = -\frac{1}{2n} B(X, X). \tag{79.262}$$

LONGEqDefNormeKillingSix

Notice that q_0 belongs to the compact part of $\mathcal{G}$, so that its Killing norm is negative, so that $\|q_0\|$ is positive.

LONGPropBaseQOrtho

Proposition 79.62.
We have $\|q_0\| = 1$ and $\|q_i\| = -1$ ($i \neq 0$). As a consequence, the space G/H is Lorentzian.

Proof. We begin by computing the norm of $q_1 = J_2$. The Killing form $B(J_2, J_2) = \mathrm{Tr}\big(\mathrm{ad}(J_2) \circ \mathrm{ad}(J_2)\big)$ is the easiest to compute in the basis $\mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \oplus \mathcal{A} \oplus \mathcal{N} \oplus \bar{\mathcal{N}}$ of eigenvectors of J_2. If we look at the matrix of $\mathrm{ad}(J_2) \circ \mathrm{ad}(J_2)$, we have one 1 at each of the positions of X_{++}, X_{+-}, X_{-+} and X_{--}. Moreover, for each higher dimensional slice, we get additional 2 because of X^k_{0+} and X^k_{+0}. When one looks at $\mathfrak{so}(2, n)$ we have $n - 2$ higher dimensional slices, so that

$$B(J_2, J_2) = 4 + 2(n - 2) = 2n. \tag{79.263}$$

The result is that $B(q_1, q_1) = 6$, so that $\|q_1\| = -1$.

We are going to propagate that result to other elements of the basis, using the"magic" intertwining elements X_1, X_k and J_1.

Using left invariance of the Killing form, we find

$$B(q_0, q_0) = B\big(q_0, -\mathrm{ad}(J_1)q_2\big) = B\big(\mathrm{ad}(J_1)q_0, q_2\big) = -B(q_2, q_2), \tag{79.264}$$

LONGEqCalculBBBqZeroqUnNirme

so that $\|q_0\| = -\|q_2\|$.

Now, the same computation with X_1 and X_k instead of J_1 show that $\|q_0\| = -\|q_1\|$ and $\|q_1\| = \|q_k\|$. □

LONGRemBProdScal

Remarque 79.63.
Using the fact that the basis $\{q_i\}$ is orthonormal, we can decompose an element of $\mathcal{Q}$ by the Killing form. One only has to be careful on the sign: if $X = aq_0 + \sum_{i>0} b_i q_i$, we have

$$\begin{aligned} a &= B(X, q_0) \\ b_i &= -B(X, q_i). \end{aligned} \tag{79.265}$$

LONGEqsabKillProjComp

LONGRemOrdreNilpotentQ

Remarque 79.64.
As a consequence, a light like direction reads, up to normalization, $E(w) = q_0 + \sum_{i=1}^{l-1} w_i q_i$ with $w \in S^{l-2}$.

79.21.5.4 Other properties

LONGLemXZUAHetQ

Lemma 79.65.
We have $\sigma X_{\alpha\beta} \in \mathcal{G}_{(\alpha,-\beta)}$. In particular, X^k_{0+} has non vanishing components in $\mathcal{H}$ and in $\mathcal{Q}$.

Proof. If one applies σ to the equality $[J_2, X_{\alpha\beta}] = \beta X_{\alpha\beta}$, we see that $\sigma X_{\alpha\beta}$ is an eigenvector of $\mathrm{ad}(J_2)$ with eigenvalue $-\beta$. The same with $\mathrm{ad}(J_1)$ shows that $\sigma X_{\alpha\beta}$ has $+1$ as eigenvalue. Thus $\sigma X_{\alpha\beta} \in \mathcal{G}_{(\alpha,-\beta)}$.

In particular, $\sigma X^k_{0+} \neq \pm X^k_{0+}$ so that it does not belongs to $\mathcal{H}$ nor to $\mathcal{Q}$. □

Notice that, as corollary, we have

$$\sigma X_{\alpha,\beta} = \pm X_{\alpha,-\beta}. \tag{79.266}$$

LONGLemSigmaXppEgalXPm

Lemma 79.66.
We have $(X_{++})_{\mathcal{Q}} = (X_{+-})_{\mathcal{Q}}$ or, equivalently, $\sigma X_{++} = -X_{+-}$.

Proof. Since $q_1 = J_2 \in \mathcal{A}$ and $q_k \in \tilde{\mathcal{N}}_k$, the $\mathcal{Q}$-component of X_{++} and X_{+-} are only made of q_0 and q_2. We are going to prove the following three equalities.

LONGItemBpmqDeux

(1) $B(X_{+-}, q_2) = B(X_{+-}, q_0)$
(2) $B(X_{++}, q_2) = B(X_{++}, q_0)$
(3) $B(X_{++}, q_0) = B(X_{+-}, q_0)$

The first point is proved using the fact that $q_2 = [q_0, J_1]$ and the ad-invariance of the Killing form:

$$B(X_{+-}, q_2) = -B\big(X_{+-}, \mathrm{ad}(J_1)q_0\big) = B\big(\mathrm{ad}(J_1)X_{+-}, q_0\big) = B(X_{+-}, q_0). \tag{79.267}$$

One checks the second point in the same way. For the third equality, we know from decomposition (79.232) that q_0 is a multiple of $X_{++} + X_{--} + X_{+-} + X_{-+}$. If the multiple is λ, $B(X_{++}, q_0) = \lambda B(X_{++}, X_{--})$ and $B(X_{+-}, q_0) = \lambda B(X_{+-}, X_{-+})$. Thus we have to prove that the traces of the operators

$$\begin{aligned} \gamma_1 &= \mathrm{ad}(X_{++}) \circ \mathrm{ad}(X_{--}) \\ \gamma_2 &= \mathrm{ad}(X_{+-}) \circ \mathrm{ad}(X_{-+}) \end{aligned} \tag{79.268}$$

are the same. That trace is straightforward to compute on the natural basis of $\mathcal{G} = \mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \oplus \mathcal{A} \oplus \mathcal{N} \oplus \bar{\mathcal{N}}$. The only elements on which $\mathrm{ad}(X_{--})$ is not zero are $\mathcal{A}$, X^k_{0+}, X^k_{+0} and X^k_{++}, while for $\mathrm{ad}(X_{-+})$, the only non vanishing elements are $\mathcal{A}$, X^k_{0-}, X^k_{+0} and X_{+-}. From equation (79.208e), we have $\gamma_1(r_{ij}) = \gamma_2(r_{ij}) = 0$. Using the commutation relations, we find

$$\gamma_1 J_1 = [X_{++}, X_{--}] = -4(J_1 + J_2) \tag{79.269a}$$
$$\gamma_1 J_2 = [X_{++}, X_{--}] = -4(J_1 + J_2) \tag{79.269b}$$
$$\gamma_1 X^k_{0+} = 2[X_{++}, X^k_{-0}] = -4X^k_{0+} \tag{79.269c}$$
$$\gamma_1 X^k_{+0} = -2[X_{++}, X^k_{0-}] = -4X^k_{+0} \tag{79.269d}$$
$$\gamma_1 X_{++} = [X_{++}, 4(J_1 + J_2)] = -8X_{++}. \tag{79.269e}$$

Thus $\mathrm{Tr}(\gamma_1) = -24$. The same computations bring

$$\gamma_2 J_1 = [X_{+-}, X_{-+}] = -4(J_1 - J_2) \tag{79.270a}$$
$$\gamma_2 J_2 = [X_{+-}, X_{-+}] = 4(J_1 - J_2) \tag{79.270b}$$
$$\gamma_2 X^k_{0-} = -2[X_{+-}, X^k_{-0}] = -4X^k_{0-} \tag{79.270c}$$
$$\gamma_2 X_{+-} = [X_{+-}, 4(J_1 - J_2)] = -8X_{+-} \tag{79.270d}$$
$$\gamma_2 X^k_{+0} = 2[X_{+-}, X^k_{0+}] = -2X^k_{+0}, \tag{79.270e}$$

and $\mathrm{Tr}(\gamma_2) = -24$. Thus we have

$$\mathtt{proj}_{\mathcal{Z}(\mathcal{K})}(X_{++}) = \mathtt{proj}_{\mathcal{Z}(\mathcal{K})}(X_{+-}). \tag{79.271}$$

□

Notice that the lemma is trivial if we consider that $X_{++} - X_{+-}$ belongs to $\mathcal{H}$ by definition of $\mathcal{H}$. From a AdS point of view, we define $AdS = G/H$ and we have to define H, so from that point of view, lemma 79.66 is by definition. However, the direction we have in mind is to use the more generic tools as possible. From that point of view, the fact to set $\mathcal{Z}(\mathcal{K}) \subset \mathcal{Q}$ is more intrinsic than to set $X_{++} - X_{+-} \in \mathcal{H}$.

LONGPropXmpXppqq

Proposition 79.67.
We have $(X_{++})_{\mathcal{Q}} = (X_{+-})_{\mathcal{Q}} = q_0 - q_2$.

Proof. Using the remark 79.63, the three Killing forms computed in the proof of lemma 79.66 are expressed under the form

$$(X_{+-})_{q_0} = -(X_{+-})_{q_2} \tag{79.272a}$$
$$(X_{++})_{q_0} = -(X_{++})_{q_2} \tag{79.272b}$$
$$(X_{++})_{q_0} = (X_{+-})_{q_2}. \tag{79.272c}$$

Consequently, we have $(X_{++})_{\mathcal{Q}} = \lambda(q_0 - q_2)$ and $(X_{+-})_{\mathcal{Q}} = \lambda(q_0 - q_2)$ for a constant λ to be fixed. It is fixed to be 1 by the facts that, by definition, $q_0 = (X_{++})_{\mathcal{K}\mathcal{Q}}$ and $q_2 \in \mathcal{P}$. □

LONGLemComJDeuxQ

Lemma 79.68.
We have

$$\begin{aligned}
[J_2, q_0] &= (X_{++})_{\mathcal{H}\mathcal{P}} \neq 0 \\
[J_2, q_1] &= 0 \\
[J_2, q_2] &= (X_{++})_{\mathcal{H}\mathcal{K}} \neq 0 \\
[J_2, q_k] &= (X^k_{0+})_{\mathcal{H}} \neq 0 \\
[J_1, p_1] &= -(X_{++})_{\mathcal{K}\mathcal{H}} \neq 0
\end{aligned} \tag{79.273}$$

where $k \geqslant 3$.

Proof. Using the fact that $J_2 \in \mathcal{Q} \cap \mathcal{P}$ and that X_{++} has non vanishing components "everywhere" (corollary 79.53), we have

$$\begin{aligned}
[J_2, q_0] &= [J_2, (X_{++})_{\mathcal{K}\mathcal{Q}}] = (X_{++})_{\mathcal{P}\mathcal{H}} \neq 0 \\
[J_2, q_2] &= [J_2, (X_{++})_{\mathcal{P}\mathcal{Q}}] = (X_{++})_{\mathcal{K}\mathcal{H}} \neq 0 \\
[J_1, p_1] &= [J_1, (X_{++})_{\mathcal{P}\mathcal{H}}] = -(X_{++})_{\mathcal{K}\mathcal{H}} \neq 0 \\
[J_2, q_k] &= [J_2, (X^k_{0+})_{\mathcal{Q}}] = (X^k_{0+})_{\mathcal{H}} \neq 0 \qquad \text{lemma 79.65}
\end{aligned} \tag{79.274}$$

□

LONGLemNonHXaz

Lemma 79.69.
If $\alpha \neq 0$, *then* $X^k_{\alpha 0} \in \mathcal{H}$.

Proof. The element $\mathtt{proj}_{\mathcal{Q}} X^k_{\alpha 0}$ is a combination of q_i. Since $\mathrm{ad}(J_2)\,\mathtt{proj}_{\mathcal{Q}} X^k_{\alpha 0} = 0$, we must have $(X^k_{\alpha 0})_{\mathcal{Q}} = \lambda J_2$ by lemma 79.68. Using the fact that $J_1 \in \mathcal{H}$, the $\mathcal{Q}$-component of the equality $[J_1, X^k_{\alpha 0}] = \alpha X^k_{\alpha 0}$ becomes

$$[J_1, \lambda J_2] = \alpha\lambda J_2. \tag{79.275}$$

The left-hand side is obviously zero, so that $\lambda = 0$ which proves that $X^k_{\alpha 0} \in \mathcal{H}$. □

Applying successively the projections (79.224), and lemma 79.66, we write the basis elements of $\mathcal{Q}$ in the decomposition $\mathcal{G} = \mathcal{G}_0 \oplus \mathcal{N} \oplus \bar{\mathcal{N}}$: LONGEqsDecopmQXpmpm

$$q_0 = \frac{1}{4}(X_{++} + X_{+-} + X_{-+} + X_{--}), \tag{79.276a}$$
$$q_1 = J_2, \tag{79.276b}$$
$$q_2 = \frac{1}{4}(-X_{++} - X_{+-} + X_{-+} + X_{--}), \tag{79.276c}$$
$$q_k = \frac{1}{2}(X^k_{0+} - X^k_{0-}) \tag{79.276d}$$

with $k \geqslant 3$. Notice that none of them has component in $\mathcal{Z}_{\mathcal{K}(\mathcal{A})}$.

These decompositions allow us to compute the commutators $[q_i, q_j]$ and $[q_i, J_p]$. Instead of listing here every commutation relations, we will only write the ones we use when we need them.

LONGLemQzQdeuxJun

Lemma 79.70.
We have $[q_0, q_2] = -J_1$.

Proof. The proof is exactly the same as the one of equation (79.249b) in lemma 79.59. Here we use

$$(X_{++})_{\mathcal{PQ}} = \frac{1}{4}\big(X_{++} - \sigma X_{++} - \theta X_{++} + \sigma\theta X_{++}\big) \tag{79.277}$$

and we find

$$[q_0, q_2] = -\big[(X_{++})_{\mathcal{KQ}}, (X_{++})_{\mathcal{PQ}}\big] = -\frac{1}{4}[\theta X_{++}, X_{++}]_{\mathcal{H}} = -J_1. \tag{79.278}$$

□

Lemma 79.71.
We have

$$[X_1, q_2] = [X_1, q_k] = 0 \tag{79.279}$$

LONGEqXunQdeuxcommutent

for $k \geqslant 3$.

Proof. The proof is elementary:

$$\begin{aligned} [X_1, q_2] &\in [\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2, \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2] \subset \mathcal{K} \cap \mathcal{Q} \cap \mathcal{A} = \{0\} \\ [X_1, q_k] &\in [\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2, \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k] \subset \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \{0\}. \end{aligned} \tag{79.280}$$

□

The following is a first step in the proof of theorem 79.76.

LONGCorAdQUncarreqi

Corollary 79.72.
We have $\mathrm{ad}(J_1)|^2_{\tilde{\mathcal{N}}_2} = \mathrm{ad}(J_2)|^2_{\tilde{\mathcal{N}}_2} = \mathrm{Id}$ *and* $\mathrm{ad}(q_1)^2 q_i = q_i$.

Proof. The action of $\mathrm{ad}(q_1)^2$ is to change two times the sign of the components $X_{\alpha -}$. Thus $\mathrm{ad}(q_1)^2 = \mathrm{Id}$ on $\tilde{\mathcal{N}}_2$. The result is now proved for $i = 0, 1, 2$. For the higher dimensions, we use the fact that $J_2 = q_1$ and we find

$$q_k = [X_k, q_1] = -\big[[q_1, X_k], q_1\big] = \mathrm{ad}(q_1)^2 q_k \tag{79.281}$$

as claimed.

Moreover, the elements of $\tilde{\mathcal{N}}_2$ are build of elements of the form $X_{\pm\pm}$, so that $\mathrm{ad}(J_1)^2$ changes at most twice the sign.

□

LONGLemXkqzero

Lemma 79.73.
We have

$$[X_k, q_0] = [X_k, J_1] = [X_k, q_2] = 0 \tag{79.282a}$$

$$[J_1, q_k] = 0. \tag{79.282b}$$

LONGEqJUnqkzero

Proof. The first claim is proved in a very standard way:

$$[X_k, q_0] \in [\mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k, \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2] \subset \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \{0\}. \tag{79.283}$$

For the second commutator, we use the Jacobi identity and the definition $X_k = -[J_2, q_k]$:

$$\big[J_1, [J_2, q_k]\big] = -\big[J_2, [q_k, J_1]\big] - \big[q_k, \underbrace{[J_1, J_2]}_{=0}\big], \tag{79.284}$$

LONGEqJ1J2qkzero

while

$$[q_k, J_1] \in [\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k, \mathcal{P} \cap \mathcal{H} \cap \mathcal{A}] \subset \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \{0\}. \tag{79.285}$$

That proves (79.282b) in the same time.

For the third commutator, remark that, since $q_2 = [q_0, J_1]$, we have

$$[X_k, q_2] = -\big[q_0, [J_1, X_k]\big] - \big[J_1, [X_k, q_0]\big]. \tag{79.286}$$

which is zero by the two first claims. □

LONGEtOrdreDeux

Proposition 79.74.
If E is nilpotent in $\mathcal{Q}$, then $\mathrm{ad}(E)^3 = 0$.

Proof. If E_1 is a nilpotent element of $\mathcal{Q}$, then every nilpotent elements in $\mathcal{Q}$ are of the form $\lambda\,\mathrm{Ad}(k)E_1$ for some $k \in K$ and $\lambda \in \mathbb{R}$[1021]. It is then sufficient to prove that one of them is of order two. The element

LONGEqDecqzmoinsqDeux

$$q_0 - q_2 = \frac{1}{2}(X_{++} + X_{+-}), \tag{79.287}$$

is obviously of order two because the eigenvalue for $\mathrm{ad}(J_1)$ increases by one unit at each iteration of $\mathrm{ad}(q_0 - q_2)$. □

Lemma 79.75.
We have $[J_1, q_k] = 0$.

Proof. The proof is standard:

$$[J_1, q_k] \in [\mathcal{H} \cap \mathcal{P} \cap \mathcal{A}, \mathcal{Q} \cap \mathcal{P} \cap \tilde{\mathcal{N}}_k] \subset \mathcal{Q} \cap \mathcal{K} \cap \tilde{\mathcal{N}}_k = \{0\}. \tag{79.288}$$

□

The following theorem, which relies on the preceding lemmas, will be central in computing the Killing form which appears in the characterization of theorem 79.95.

LONGThoAdESqqq

Theorem 79.76.
We have

$$\mathrm{ad}(q_i)^2 q_j = q_j \tag{79.289}$$

if $i \neq j$ and $i \neq 0$. If $i = 0$, we have

$$\mathrm{ad}(q_0)^2 q_j = -q_j. \tag{79.290}$$

Proof. The case $i = 1$ is already done in corollary 79.72.

We propagate that result to the other $\mathrm{ad}(q_i)^2$ with the intertwining elements J_1, X_1 and X_k.

Let us compute $\mathrm{ad}(q_0)^2 q_i = \mathrm{ad}\big([X_1, q_1]\big)^2 q_i$ using twice the Jacobi identity and the properties of X_1 (in order to be more readable, we write XY for $[X, Y]$)

LONGEqAdqZsqqi

$$\begin{aligned}
\mathrm{ad}(q_0)^2 q_i &= (X_1 q_1)\Big((X_1 q_1) q_i\Big) \\
&= -(X_1 q_1)\Big((q_1 q_I) X_1 + (q_i X_1) q_1\Big) \\
&= (q_1 q_i)\big(X_1 (X_1 q_1)\big) + (q_i X_1)\big(q_1 (X_1 q_1)\big) \\
&\quad + X_1\big((X_1 q_1)(q_1 q_i)\big) + q_1\big((X_1 q_1)(q_i X_1)\big) \\
&= (q_1 q_i) q_1 - \mathrm{ad}(X_1)^2 q_i + X_1\big(q_0 (q_1 q_i)\big) + q_1\big(q_0 (q_i X_1)\big).
\end{aligned} \tag{79.291}$$

If $i = 1$, the only non vanishing term is $-\mathrm{ad}(X_1)^2 q_1 = -q_1$. Thus $\mathrm{ad}(q_0)^2 q_1 = -q_1$.

If $i = 2$, the relation (79.279) annihilates the second and fourth terms while $[q_1, q_2]$ commutes with q_0 because $q_0 \in \mathcal{Z}(\mathcal{K})$. We are thus left with the term $-q_2$, and $\mathrm{ad}(q_0)^2 q_2 = -q_2$.

If $i = k \geqslant 3$, we find

$$\mathrm{ad}(q_0)^2 q_k = -\mathrm{ad}(q_1)^2 q_k - \mathrm{ad}(X_1)^2 q_k + X_1\big(q_0(q_1 q_k)\big) + q_1\big(q_0(q_k X_1)\big). \tag{79.292}$$

Since $[q_1, q_k] \in \mathcal{K}$, it commutes with q_0. Using the fact that $[X_1, q_k] = 0$, we get $\mathrm{ad}(q_0)^2 q_k = -q_k$.

Let us perform the same computations as in (79.291) with q_k ($k \geqslant 3$) instead of q_0 and X_k (equations (79.251)) instead of X_1. What we get is

$$\mathrm{ad}(q_k)^2 q_i = \mathrm{ad}(q_1)^2 q_i - \mathrm{ad}(X_k)^2 q_i + X_k\big(q_k(q_1 q_i)\big) + q_1\big(q_k(q_i X_k)\big). \tag{79.293}$$

If we set $i = 0$, taking into account the commutator $[X_k, q_0] = 0$, we have

$$\mathrm{ad}(q_k)^2 q_0 = \mathrm{ad}(q_1)^2 q_0 + X_k\big(q_k(q_1 q_0)\big). \tag{79.294}$$

LONGEqkdeuxqzerointer

As already proved, the first term is q_0. Now,

$$\big[q_k, [q_1, q_0]\big] \in \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \{0\}, \tag{79.295}$$

so that the second term in (79.294) is zero. Thus we proved that $\mathrm{ad}(q_k)^2 q_0 = q_0$.

If we set $i = 1$, taking into account the relations (79.251), we find

$$\mathrm{ad}(q_k)^2 q_1 = -\mathrm{ad}(X_k)^2 q_1 + q_1\big(q_k(q_1 X_k)\big) = q_1. \tag{79.296}$$

If we set $i = 2$ and using the fact that $[X_k, q_2] = 0$, we find

$$\mathrm{ad}(q_k)^2 q_2 = q_2 - q_k\big(X_k(q_1 q_2)\big). \tag{79.297}$$

Using once again the Jacobi identity inside the big parenthesis, we find $2q_2 - \mathrm{ad}(q_k)^2 q_2$. This proves that $\mathrm{ad}(q_k)^2 q_2 = q_2$.

We turn now our attention to $\mathrm{ad}(q_2)^2 q_i$. We perform the same computation, using the intertwining property (79.248) of J_1. What we get is

$$\mathrm{ad}(q_2)^2 q_i = (J_1 q_i)(q_0 q_2) - \mathrm{ad}(q_0)^2 q_i + q_0\big(q_2(J_1 q_i)\big) + J_1\big(q_2(q_i q_0)\big). \tag{79.298}$$

LONGEqadqqqDeuxIGene

If we set $i = 1$, we use the already proved property $\mathrm{ad}(q_0)^2 q_1 = -q_1$, and we obtain

$$\mathrm{ad}(q_2)^2 q_1 = (J_1 q_1)(q_0 q_2) + q_1 + q_0\big(q_2(J_1 q_1)\big) + J_1\big(q_2(q_1 q_0)\big). \tag{79.299}$$

We claim that all of these terms are zero except of q_1. First, $\big[q_2, [q_1, k_k]\big] \in \big[\tilde{\mathcal{N}}_2, [\mathcal{A}, \tilde{\mathcal{N}}_2]\big] \subset \mathcal{A}$. Thus the last term vanishes .The commutator $[J_1, q_1]$ vanishes because $q_1 = J_2$. We are done with $\mathrm{ad}(q_2)^2 q_1 = q_1$.

If we set $i = k$ ($k \geqslant 3$) in (79.298), we use $\mathrm{ad}(q_0)^2 q_k = -q_k$ and what we find is

$$\mathrm{ad}(q_2)^2 q_k = (J_1 q_k)(q_0 q_2) + q_k + q_0\big(q_2(J_1 q_k)\big) + J_1\big(q_2(q_k q_0)\big). \tag{79.300}$$

LONGEqadqqqdeuxkk

We already know that $[J_1, q_k] = 0$. We have $\big[q_2, [q_k, q_0]\big] = 0$ because

$$\begin{aligned}\big[q_2, [q_k, q_0]\big] &\in \big[\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2, [\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k, \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2]\big] \\ &\subset [\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2, \mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k] \\ &\subset \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \{0\}.\end{aligned} \tag{79.301}$$

The remaining terms in (79.300) are $\mathrm{ad}(q_2)^2 q_k = q_k$.

In order to compute $\mathrm{ad}(q_2)^2 q_0$, we write $q_0 = \mathrm{ad}(X_1) q_1$. Using twice the Jacobi identity, we get

$$\mathrm{ad}(q_2)^2 q_0 = X_1\big((q_1 q_2) q_2\big) + q_1\big((X_1 q_2) q_2\big) + (q_1 q_2)(X_1 q_2) + (X_1 q_2)(q_2 q_1). \tag{79.302}$$

Using the fact that $[X_1, q_2] = 0$, we are left with

$$\mathrm{ad}(q_2)^2 q_0 = X_1\big(\mathrm{ad}(q_2)^2 q_1\big) = [X_1, q_1] = q_0 \tag{79.303}$$

as desired. □

79.21.6 A convenient basis for the root spaces and computations

LONGSubSecMOreConvBasisBlbla

The most natural basis of $\tilde{\mathcal{N}}_2$ is

$$\tilde{\mathcal{N}}_2 = \langle X_{++}, X_{+-}, X_{-+}, X_{--} \rangle, \tag{79.304}$$

but the multiple commutators of these elements with q_0 reveal to require some work.

We provide in this section an other basis for $\tilde{\mathcal{N}}$ that corresponds to the decomposition $\mathcal{K} \oplus \mathcal{P}$. Since q_0 is central in $\mathcal{K}$, the exponential $e^{xq_0}X$ is trivial when $X \in \mathcal{K}$ and, since $q_0 \in \mathcal{K}$, the commutator $[q_0, X]$ remains in $\mathcal{P}$ when $X \in \mathcal{P}$.

Here is the new basis:

$$\mathbb{B} = \{J_1, J_2, q_0, q_2, p_1, s_1, q_k, p_k, r_k, s_k\}_{k=3,\dots,l-1}. \tag{79.305}$$

LONGEqSuperBaseeB

where LONGEqDEqsBasqSrkXsNTrois

$$q_0 = \frac{1}{4}(X_{++} + X_{+-} + X_{-+} + X_{--}) \in \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 \tag{79.306a}$$

$$q_2 = [q_0, J_1] = \frac{1}{4}(-X_{++} - X_{+-} + X_{-+} + X_{--}) \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 \tag{79.306b}$$

$$p_1 = [q_0, q_1] = \frac{1}{4}(-X_{++} + X_{+-} - X_{-+} + X_{--}) \in \mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 \tag{79.306c}$$

LONGEqqzqupuDef

$$s_1 = [J_1, p_1] = \frac{1}{4}(-X_{++} + X_{+-} + X_{-+} - X_{--}) \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 \tag{79.306d}$$

$$q_k = \frac{1}{2}(X^k_{0+} - X^k_{0-}) \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k \tag{79.306e}$$

$$p_k = [q_0, q_k] = \frac{1}{2}(X^k_{-0} - X^k_{+0}) \in \mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k \tag{79.306f}$$

$$r_k = [J_2, q_k] = \frac{1}{2}(X^k_{0+} + X^k_{0-}) \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k \tag{79.306g}$$

$$s_k = [J_1, p_k] = -\frac{1}{2}(X^k_{-0} + X^k_{+0}) \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k \tag{79.306h}$$

$$r_{ij} = [q_i, q_j] \in \mathcal{K} \cap \mathcal{H} \cap \mathcal{G}_0. \tag{79.306i}$$

and

$$J_1 \in \mathcal{P} \cap \mathcal{H} \cap \mathcal{A} \tag{79.306j}$$

$$q_1 = J_2 \in \mathcal{P} \cap \mathcal{Q} \cap \mathcal{A} \tag{79.306k}$$

Notice that the elements p_1 and s_1 are non vanishing by lemma 79.68. We have

$$\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 = \langle q_2 \rangle \qquad \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2 = \langle q_0 \rangle \tag{79.307a}$$

$$\mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \langle q_k \rangle \qquad \mathcal{K} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \varnothing \tag{79.307b}$$

$$\mathcal{P} \cap \mathcal{Q} \cap \mathcal{A} = \langle J_2 \rangle \qquad \mathcal{K} \cap \mathcal{Q} \cap \mathcal{A} = \varnothing \tag{79.307c}$$

$$\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 = \langle p_1 \rangle \qquad \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_2 = \langle s_1 \rangle \tag{79.307d}$$

$$\mathcal{P} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k = \langle p_k \rangle \qquad \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k = \langle r_k, s_k \rangle \tag{79.307e}$$

$$\mathcal{P} \cap \mathcal{H} \cap \mathcal{A} = \langle J_1 \rangle \qquad \mathcal{K} \cap \mathcal{H} \cap \mathcal{A} = \varnothing \tag{79.307f}$$

The decomposition of $\tilde{\mathcal{N}}_2$ into $\mathcal{K} \oplus \mathcal{P}$ is

$$\tilde{\mathcal{N}}_2 = \langle q_0, s_1 \rangle \oplus \langle q_1, p_1 \rangle. \tag{79.308}$$

LONGEqDecomptsNTroisKP

The decomposition of $\tilde{\mathcal{N}}_k$ into $\mathcal{K} \oplus \mathcal{P}$ is

$$\tilde{\mathcal{N}}_k = \langle r_k, s_k, r_{ij} \rangle_{k,i,j \geqslant 3} \oplus \langle q_k, p_k \rangle. \tag{79.309}$$

LONGEqDecomptsNkKP

We are now going to compute all the Killing form and commutators in this basis.

LONGPropBprsk

Proposition 79.77.
We have

$$\begin{aligned} B(p_k, p_k) &= -B(q_0, q_0) \\ B(r_k, r_k) &= B(q_0, q_0) \\ B(s_k, s_k) &= B(q_0, q_0), \end{aligned} \tag{79.310}$$

and then

$$-\|p_k\|^2 = \|r_k\|^2 = \|s_k\|^2 = 1. \tag{79.311}$$

LONGGeqNormInHigherDimensionalSlices

Proof. Using the definitions (79.306) and the theorem 79.76, we have

LONGSubEqsBpprrssk

$$B(p_k, p_k) = B\big(\operatorname{ad}(q_k)q_0, \operatorname{ad}(q_k)q_0\big) = -B\big(\operatorname{ad}(q_k)^2 q_0, q_0\big) = -B(q_0, q_0), \tag{79.312a}$$

and

$$B(r_k, r_k) = B\big(\operatorname{ad}(q_1)q_k, \operatorname{ad}(q_1)q_k\big) = -B(q_k, q_k) = B(q_0, q_0). \tag{79.312b}$$

and

$$B(s_k, s_k) = -B\big(\operatorname{ad}(J_1)^2 p_k, p_k\big) = -B(p_k, p_k) = B(q_0, q_0). \tag{79.312c}$$

□

Notice that we are not surprised by the positivity of the norms of r_k and s_k because they belong to the compact part of the algebra.

Proposition 79.78.
The Killing norm in the space $\mathcal{Z}_{\mathcal{K}}(\mathcal{A})$ are given by

$$B(r_{ij}, r_{kl}) = \begin{cases} B(q_0, q_0) & \text{if } \{i, j\} = \{k, l\} \\ 0 & \text{otherwise.} \end{cases} \tag{79.313}$$

Proof. First, we have

$$\begin{aligned} B(r_{ij}, r_{ij}) &= B\big([q_i, q_j], [q_i, q_j]\big) \\ &= -B\big(\operatorname{ad}(q_i)^2 q_j, q_j\big) \\ &= -B(q_j, q_j) \\ &= B(q_0, q_0). \end{aligned} \tag{79.314}$$

For the mixed case we have

$$B(r_{ij}, r_{ik}) = -B\big(\operatorname{ad}(q_i)^2 q_j, q_k\big) = 0. \tag{79.315}$$

We suppose now that i, j, k and l are four different numbers. The action of $\operatorname{ad}(r_{kl})$ on $\mathcal{A}$ is zero because $r_{kl} \in \mathcal{Z}(\mathcal{A})$. From (79.208a), the action of $\operatorname{ad}(r_{ij}) \circ \operatorname{ad}(r_{kl})$ on $\mathcal{N}$ is zero. Since the elements r_{ij} satisfy the algebra of $\mathfrak{so}(n)$, we have $\operatorname{ad}(r_{ij}) \circ \operatorname{ad}(r_{kl}) r_{mn} = 0$ when i, j, k and l are four different numbers. Finally we have

$$B(r_{ij}, r_{kl}) = 0. \tag{79.316}$$

□

LONGLempunjdeuxqzero

Lemma 79.79.
We have $[p_1, J_2] = q_0$.

Proof. The lemma comes from theorem 79.76 because

$$[p_1, J_2] = -[q_1, p_1] = \operatorname{ad}(q_1)^2 q_0 = q_0. \tag{79.317}$$

□

LONGlemJDeuxqDeuxsUn

Lemma 79.80.
We have $s_1 = [J_2, q_2]$.

Proof. We use the definition $p_1 = [q_0, q_1]$ and the Jacobi identity:

$$[J_1, p_1] = \big[J_1, [q_0, q_1]\big] = -\big[q_0, [q_1, J_1]\big] - \big[q_1, [J_1, q_0]\big]. \tag{79.318}$$

The first terms vanishes because $q_1 \in \mathcal{A}$ while $[J_1, q_0] = -q_2$ by definition. □

LONGPropBJpsun

Proposition 79.81.
We have

$$\begin{aligned} B(J_1, J_1) &= -B(q_0, q_0) \\ B(p_1, p_1) &= -B(q_0, q_0) \\ B(s_1, s_1) &= B(q_0, q_0), \end{aligned} \tag{79.319}$$

and then

$$-\|J_1\|^2 = -\|p_1\|^2 = \|s_1\|^2 = 1. \tag{79.320}$$

Proof. Using $J_1 = [q_0, q_2]$ (lemma 79.70) we find

$$B(J_1, J_1) = B\big(\operatorname{ad}(q_2)q_0, \operatorname{ad}(q_2)q_0\big) = -B\big(\operatorname{ad}(q_2)^2 q_0, q_0\big) = -B(q_0, q_0). \tag{79.321}$$

In much the same way, using the definition of p_1 and $s_1 = [q_1, q_2]$ (lemma 79.80), we find $B(p_1, p_1) = B(q_1, q_1)$ and $B(s_1, s_1) = -B(q_2, q_2)$.

□

LONGLemJunrkzero

Lemma 79.82.
We have $[J_1, r_k] = 0$.

Proof. Using the definition of r_k and the Jacobi identity,

$$[J_1, r_k] = \big[J_1, [J_2, q_k]\big] = -\big[J_2, [q_k, J_1]\big] - \big[q_k, [J_1, J_2]\big] = 0 \tag{79.322}$$

because of equation (79.282b) and the fact that $\mathcal{A}$ is abelian. □

Now, the Killing norms of the basis $\mathbb{B}$ can be computed.

LONGThoBaisXXorthoigher

Theorem 79.83.
The basis

$$\mathbb{B} = \{q_0, q_2, p_1, s_1, q_k, p_k, r_k, s_k, J_1, J_2\}. \tag{79.323}$$

given by the definitions (79.305) *is orthonormal and*

$$\begin{aligned} \|J_1\|^2 = \|q_1\|^2 = \|q_2\|^2 = \|p_1\|^2 = \|q_k\|^2 = \|p_k\|^2 = -1 \\ \|r_{ij}\| = \|q_0\|^2 = \|s_1\|^2 = \|r_k\|^2 = \|s_k\|^2 = 1 \end{aligned} \tag{79.324}$$

Proof. The norms are given by the propositions 79.77, 79.81, and 79.62.

For the orthogonality, we know that $\mathcal{P} \perp \mathcal{K}$, $\mathcal{Q} \perp \mathcal{H}$ (from general theory) as well as $\tilde{\mathcal{N}}_2 \perp \tilde{\mathcal{N}}_k$, $\tilde{\mathcal{N}}_2 \perp \mathcal{A}$ and $\tilde{\mathcal{N}}_k \perp \mathcal{A}$ (equations (79.214a), (79.214b), (79.214c) and (79.214d)). Thus, among the elements of the basis $\mathbb{B}$, the two only ones that could not orthogonal are r_k and s_k. However, we have $B(r_k, s_k) = 0$ because

$$B(r_k, s_k) = B\big(r_k, \operatorname{ad}(J_1)p_k\big) = -B\big(\operatorname{ad}(J_1)r_k, p_k\big) = 0 \tag{79.325}$$

by lemma 79.82.

□

It turns out that we are able to compute all the commutators using the following techniques

(1) the orthonormality of the basis, theorem 79.83 among with the ad-invariance of the Killing form
(2) the Jacobi identity

(3) the theorem 79.76 and the commutators of lemmas 79.79, 79.80 and 79.82.

Computing all the commutators that way is quite long and very few interesting. The interesting point is that it is possible. You can immediately jump to subsection 79.21.8.

(1) $\mathrm{ad}(q_0)J_1 = q_2$. Definition.

(2) $\mathrm{ad}(q_0)J_2 = p_1$. Definition.

(3) $\mathrm{ad}(p_1)J_1 = -s_1$. Definition.

(4) $\mathrm{ad}(q_0)q_k = p_k$. Definition. LONGItemComqzpk

(5) $\mathrm{ad}(q_k)J_2 = -r_k$. Definition.

(6) $\mathrm{ad}(p_k)J_1 = -s_k$. Definition. LONGItemCompkJun

(7) $\mathrm{ad}(p_1)J_2 = q_0$. Lemma 79.79. LONGItemCommunJdeux

(8) $\mathrm{ad}(q_2)J_2 = -s_1$. Lemma 79.80.

(9) $\mathrm{ad}(J_1)r_k = 0$. Lemma 79.82. LONGItemComJunrk

(10) $\mathrm{ad}(q_0)p_1 = -J_2$. We have $\mathrm{ad}(q_0)p_1 = \mathrm{ad}(q_0)^2 q_1 = -q_1$. LONGItemComqzpun

(11) $\mathrm{ad}(q_0)s_1 = 0$. By the usual techniques, we get $[q_0, s_1] \in \mathcal{K} \cap \mathcal{Q} \cap \mathcal{A} = \{0\}$. The following few are obtained in the same way.

(12) $\mathrm{ad}(q_0)r_k = 0$. LONGItemComqzrk

(13) $\mathrm{ad}(q_0)s_k = 0$.

(14) $\mathrm{ad}(q_2)p_1 = 0$.

(15) $\mathrm{ad}(q_k)J_1 = 0$.

(16) $\mathrm{ad}(q_2)p_k = 0$.

(17) $\mathrm{ad}(q_2)p_k = 0$.

(18) $\mathrm{ad}(p_1)q_k = 0$. LONGItemCompunqk

(19) $\mathrm{ad}(J_2)p_k = 0$.

(20) $\mathrm{ad}(q_2)s_1 = -J_2$. We know that $[q_2, s_1] \in \mathcal{P} \cap \mathcal{Q} \cap \mathcal{A} = \langle J_2 \rangle$. Thus $[q_0, s_1]$ is a multiple of J_2. The coefficient is given by $B\big([q_2, s_1], J_2\big)/B(J_2, J_2)$. Using the ad-invariance of the Killing form, we are left to compute $B(s_1, \mathrm{ad}(q_2)J_2)$. Lemma 79.80 shows then that the coefficient we are searching if $B(s_1, s_1)/B(J_2, J_2) = 1$. LONGItemComqdeuxqk

(21) $\mathrm{ad}(q_2)q_k = -s_k$. The spaces show that $[q_2, q_k] \in \mathcal{K} \cap \mathcal{H} \cap \tilde{\mathcal{N}}_k = \langle r_k, s_k \rangle$. Thus we have to check the two possible components. First, $B\big([q_2, q_k], r_k\big) = -B\big(q_2, \mathrm{ad}(q_k)^2 J_2\big) = 0$ by theorem 79.76. For the second, we use the definition of s_k and the Jacobi identity:

$$\begin{aligned} B\big([q_2, q_k], s_k\big) &= B\big([q_2, q_k], \mathrm{ad}(J_1)[q_0, q_k]\big) \\ &= B\big([q_2, q_k], -\mathrm{ad}(q_0)\underbrace{[q_k, J_1]}_{=0} - \mathrm{ad}(q_k)\underbrace{[J_1, q_0]}_{=-q_2}\big) \\ &= -B\big(\mathrm{ad}(q_2)q_k, \mathrm{ad}(q_2)q_k\big) \\ &= B\big(q_k, \mathrm{ad}(q_2)^2 q_k\big) \\ &= B(q_k, q_k). \end{aligned} \tag{79.326}$$

Finally, what we have is

$$[q_2, q_k] = \frac{B(q_2, q_2)}{B(s_k, s_k)} s_k = -s_k. \tag{79.327}$$

(22) $\mathrm{ad}(q_2)r_k = 0$. From the spaces, $[q_2, r_k] \in \langle q_k \rangle$, but

$$B\big(\mathrm{ad}(q_2)r_k, q_k\big) = -B\big(r_k, [q_2, q_k]\big) = B(r_k, s_k) = 0. \tag{79.328}$$

(23) $\mathrm{ad}(q_2)s_k = -q_k$. From the spaces, $[q_2, s_k] \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \langle q_k \rangle$, so we compute

$$B\big([q_2, s_k], q_k\big) = -B(s_k, [q_2, s_k]) = B(s_k, s_k), \tag{79.329}$$

and

$$[q_2, s_k] = \frac{B(s_k, s_k)}{B(q_k, q_k)} q_k = -q_k. \tag{79.330}$$

(24) $\mathrm{ad}(q_2)r_k = 0$. From the spaces, $[q_k, q_2] \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k = \langle q_k \rangle$. Thus we compute

$$B\big([q_2, r_k], q_k\big) = -B\big(r_k, [q_2, q_k]\big) = B(r_k, s_k) = 0 \tag{79.331}$$

where we used the item (21).

LONGItemComjunpun

(25) $\mathrm{ad}(J_1)p_1 = s_1$. From the spaces, $[J_1, p_1] \in \langle s_1 \rangle$. We have

$$B\big([J_1, p_1], s_1\big) = -B\big(p_1, [J_1, s_1]\big) = -B\big(p_1, \mathrm{ad}(J_1)^2 p_1\big) = -B(p_1, p_1), \tag{79.332}$$

thus

$$[J_1, p_1] = -\frac{B(p_1, p_1)}{B(s_1, s_1)} s_1 = s_1. \tag{79.333}$$

LONGItemComskqk

(26) $\mathrm{ad}(s_k)q_k = -q_2$. From the spaces, $[s_k, q_k] \in \langle J_2, q_2 \rangle$. Thus we have two Killing forms to compute. The first is

$$B\big([s_k, q_k], J_2\big) = B\big(s_k, [q_k, J_2]\big) = B(s_k, r_k) = 0 \tag{79.334}$$

were we used the definition of r_k. The second is

$$B\big([s_k, q_k], q_2\big) = B\big(s_k, [q_k, q_2]\big) = B(s_k, s_k). \tag{79.335}$$

where we used item (21). Thus we have

$$[s_k, q_k] = \frac{B(s_k, s_k)}{B(q_2, q_2)} q_2 = -q_2 \tag{79.336}$$

(27) $\mathrm{ad}(s_k)J_2 = 0$. From the spaces, $[s_k, J_2] \in \langle q_k \rangle$, but

$$B\big(\mathrm{ad}(s_k)J_2, q_k\big) = -B\big(J_2, [s_k, q_k]\big) = -B(J_2, q_2) = 0 \tag{79.337}$$

where we used item (26).

(28) $\mathrm{ad}(q_2)s_k = -q_k$. From the spaces, $[q_2, s_k] \in \langle q_k \rangle$. We have

$$B\big(\mathrm{ad}(q_2)s_k, q_k\big) = -B\big(s_k, \mathrm{ad}(q_2)q_k\big) = B(s_k, s_k) \tag{79.338}$$

where we used item (21). Thus

$$[q_2, s_k] = \frac{B(s_k, s_k)}{B(q_k, q_k)} q_k = -q_k. \tag{79.339}$$

(29) $\mathrm{ad}(p_1)s_1 = -J_1$. From the spaces, $[p_1, s_1] \in \langle J_1 \rangle$. We have

$$B\big([p_1, s_1], J_1\big) = -B\big(s_1, [p_1, J_1]\big) = B(s_1, s_1), \tag{79.340}$$

where we used item (25). Thus,

$$[p_1, s_1] = \frac{B(s_1, s_1)}{B(J_1, J_1)} J_1 = -J_1. \tag{79.341}$$

(30) $\mathrm{ad}(p_1)r_k = p_k$. LONGItemCompunrk We use the definition of r_k and Jacobi:

$$[p_1, r_k] = \big[p_1, [J_2, q_k]\big] = -\big[J_2, \underbrace{[q_k, p_1]}_{=0}\big] - \big[q_k, \underbrace{[p_1, J_2]}_{=q_0}\big] = -[q_k, q_0] = p_k. \tag{79.342}$$

where we used items (18), (7) and (4).

(31) $\mathrm{ad}(q_k)J_2 = -r_k$. LONGItemComkJdeux From the spaces, $[q_k, J_2] \in \langle r_k, s_k\rangle$. First, we have

$$B\big([q_k, q_2], s_k\big) = -B\big(J_2, \underbrace{[q_k, s_k]}_{=-q_2}\big) = 0 \tag{79.343}$$

where we used item (26). Secondly we have

$$B\big([q_k, J_2], r_k\big) = B\big(q_k, \underbrace{[J_2, r_k]}_{\mathrm{ad}(J_2)^2 q_k}\big) = B(q_k, q_k) \tag{79.344}$$

Thus

$$[q_k, J_2] = \frac{B(q_k, q_k)}{B(r_k, r_k)} r_k = -r_k. \tag{79.345}$$

(32) $\mathrm{ad}(p_1)p_k = r_k$. LONGItemCompunpk We use the definition of p_k and the Jacobi identity:

$$\begin{aligned}
[p_1, p_k] &= -\big[q_0, \underbrace{[q_k, p_1]}_{=0}\big] - \big[q_k, \underbrace{[p_1, q_0]}_{=J_2}\big] \\
&= -[q_k, J_2] \\
&= r_k
\end{aligned} \tag{79.346}$$

where we used items (18), (10) and (31).

(33) $\mathrm{ad}(p_1)s_k = 0$. From the spaces, $[p_1, s_k] \in \langle p_k\rangle$. We have

$$B\big([p_1, s_k], p_k\big) = -B\big(s_k, \underbrace{[p_1, p_k]}_{=0}\big) = 0 \tag{79.347}$$

where we used the item (32).

(34) $\mathrm{ad}(s_1)q_k = 0$. From the spaces, $[s_1, q_k] \in \langle q_k\rangle$. We have

$$B\big([s_1, q_k], q_k\big) = B\big(s_1, [q_k, q_k]\big) = 0. \tag{79.348}$$

(35) $\mathrm{ad}(s_1)p_k = 0$. LONGItemComsunpk From the spaces, $[s_1, p_k] \in \langle p_k\rangle$. We have

$$B\big([s_1, p_k], p_k\big) = B\big(s_1, [p_k, p_k]\big) = 0. \tag{79.349}$$

(36) $\mathrm{ad}(s_1)r_k = s_k$. From the spaces, $[s_1, q_k] \in \langle r_k, s_k\rangle$. Using Jacobi,

$$\begin{aligned}
[s_1, r_k] &= \big[[J_1, p_1], r_k\big] \\
&= -\big[\underbrace{[p_1, r_k]}_{=p_k}, J_1\big] - \big[\underbrace{[r_k, J_1]}_{=0}, p_1\big] \\
&= -[p_k, J_1] \\
&= s_k
\end{aligned} \tag{79.350}$$

where we used items (30), (9) and the definition of s_k.

(37) $\mathrm{ad}(s_1)J_1 = -p_1$. LONGItemComsunJun We have $[s_1, J_1] = -\mathrm{ad}(J_1)^2 p_1 = -p_1$ by theorem 79.76.

(38) $\mathrm{ad}(s_1)s_k = -r_k$. We use the definition of s_k and Jacobi:

$$\begin{aligned}[s_1, s_k] &= \big[s_1, [J_1, p_k]\big] \\ &= -\big[J_1, \underbrace{[p_k, s_1]}_{=0}\big] - \big[p_k, \underbrace{[s_1, J_1]}_{=-p_1}\big] \\ &= [p_k, p_1] \\ &= -r_k \end{aligned} \tag{79.351}$$

where we used items (35), (37) and (32).

(39) $\mathrm{ad}(J_2)s_k = 0$. From the spaces, $[J_2, s_k] \in \langle q_k \rangle$. We have

$$B\big([J_2, s_k], q_k\big) = B\big(J_2, \underbrace{[s_k, q_k]}_{=-q_2}\big) = 0 \tag{79.352}$$

where we used item (26).

(40) $\mathrm{ad}(q_k)p_k = -q_0$. Using the definition of p_k and the theorem 79.76,

$$[q_k, p_k] = \big[q_k[q_0, q_k]\big] = -\,\mathrm{ad}(q_k)^2 q_0 = -q_0. \tag{79.353}$$

(41) $\mathrm{ad}(q_k)r_k = -J_2$. Using the definition of r_k and theorem 79.76, we have

$$\big[q_k, [J_2, q_k]\big] = -\,\mathrm{ad}(q_k)^2 J_2 = -J_2. \tag{79.354}$$

(42) $\mathrm{ad}(p_k)r_k = -p_1$. Using the definition of r_k and Jacobi,

$$\begin{aligned}[p_k, r_k] &= -\big[pk, [q_0, q_k]\big] \\ &= \big[q_0, \underbrace{[q_k, r_k]}_{=-J_2}\big] + \big[q_k, \underbrace{[r_k, q_0]}_{=0}\big] \\ &= [J_2, q_0] \\ &= -p_1 \end{aligned} \tag{79.355}$$

where we used the items (41), (12) and the definition of p_1.

(43) $\mathrm{ad}(p_k)s_k = -J_1$. The spaces show that $[p_k, s_k] \in \langle J_1, p_1 \rangle$. We have

$$B\big([p_k, s_k], J_1\big) = -B\big(s_k, \underbrace{[p_k, s_1]}_{=-s_k}\big) = B(s_k, s_k) \tag{79.356}$$

and

$$B\big([p_k, s_k], p_1\big) = -B\big(s_k, \underbrace{[p_k, p_1]}_{=-r_k}\big) = 0 \tag{79.357}$$

where we used the items (6) and (32). Thus

$$[p_k, s_k] = \frac{B(s_k, s_k)}{B(J_1, J_1)} J_1 = -J_1. \tag{79.358}$$

(44) $\mathrm{ad}(r_k)s_k = s_1$. From the spaces, $[r_k, s_k] \in \langle s_1 \rangle$. We have

$$B\big([r_k, s_k], s_1\big) = B\big(r_k, \underbrace{[s_k, s_1]}_{=r_k}\big) = B(r_k, r_k) \tag{79.359}$$

where we used item (38). Thus

$$[r_k, s_k] = \frac{B(r_k, r_k)}{B(s_1, s_1)} s_1 = s_1. \tag{79.360}$$

(45) $\mathrm{ad}(J_1)q_2 = -q_0$. Using the definition of q_2 and theorem 79.76,

$$[J_1, q_2] = -\mathrm{ad}(J_1)^2 q_0 = -q_0. \tag{79.361}$$

(46) $\mathrm{ad}(J_1)s_k = p_k$. From the spaces, $[J_1, s_k] \in \langle p_k \rangle$. We have

$$B\big([J_1, s_k], p_k\big) = -B\big(s_k, \underbrace{[J_1, p_k]}_{=s_k}\big) = -B(s_k, s_k) \tag{79.362}$$

where we used the definition of s_k.

(47) $\mathrm{ad}(J_2)p_1 = -q_0$. LONGItemComJdeuxpun Using the definition of p_1 and theorem 79.76, $[J_2, p_1] = -\mathrm{ad}(q_1)^2 q_0 = -q_0$.

(48) $\mathrm{ad}(J_1)s_1 = q_2$. We use the definition of s_1 and Jacobi:

$$\begin{aligned}
[J_2, s_1] &= \big[J_2, [J_1, p_1]\big] \\
&= -\big[J_1, \underbrace{[p_1, J_2]}_{=q_0}\big] - \big[p_1, \underbrace{[J_2, q_1]}_{=0}\big] \\
&= -[J_1, q_0] \\
&= q_2
\end{aligned} \tag{79.363}$$

where we used item (47) and the definition of q_2.

(49) $\mathrm{ad}(J_2)r_k = q_k$. Using the definition of r_k and theorem 79.76, $[J_2, r_k] = \mathrm{ad}(q_1)^2 q_k = q_k$.

Proposition 79.84.
We have $\mathcal{H} = [\mathcal{Q}, \mathcal{Q}]$.

Proof. The inclusion $[\mathcal{Q}, \mathcal{Q}] \subset \mathcal{H}$ is by construction. Now every elements in the basis (79.227) can be expressed in terms of commutators in $\mathcal{Q}$ because

$$J_1 = [q_0, q_2] \tag{79.364a}$$
$$s_k = [q_k, q_2] \tag{79.364b}$$
$$s_1 = [J_2, q_2] \tag{79.364c}$$

LONGEqJdodsu

□

79.21.7 Properties of the basis

The basis $\mathbb{B}$ is motivated by the fact that $\mathrm{ad}(q_0)^2 q_k = -q_k$, so that $e^{\mathrm{ad}(xq_0)}$ is easy to compute on q_k and p_k. Moreover, r_k and s_k belong to $\mathcal{K}$, so that $[q_0, r_k] = [q_0, s_k] = 0$. The drawback of that decomposition is that the basis elements do not belong to $\mathcal{N}$ or $\bar{\mathcal{N}}$ while it will be useful to have basis elements in $\mathcal{N}$ and $\bar{\mathcal{N}}$, among other for theorem 79.92.

At a certain point, we are going to compute the exponentials $e^{\mathrm{ad}(xq_0)}X$ when X runs over $\mathcal{N}$ and $\bar{\mathcal{N}}$. We are going to extensively use the commutation relations listed in (79.202), (79.203) and (79.204). A particular attention will be devoted to the projection over $\mathcal{Q}$ which will be central in determining the open and closed orbits of AN in G/H.

LONGLemDecomptsNDanseB

Lemma 79.85.
The decomposition of $\tilde{\mathcal{N}}_2$ with respect to $\mathbb{B}$ is LONGSubeqsDecompXqps

$$X_{++} = q_0 - q_2 - p_1 - s_1 \tag{79.365a}$$
$$X_{+-} = q_0 - q_2 + p_1 + s_1 \tag{79.365b}$$

and the decomposition of $\tilde{\mathcal{N}}_k$ with respect to $\mathbb{B}$ is

$$X^k_{0+} = q_k + r_k \qquad X^k_{0-} = -q_k + r_k \tag{79.366a}$$
$$X^k_{+0} = -p_k - s_k \qquad X^k_{-0} = p_k - s_k \tag{79.366b}$$

LONGsubEqqrkXzpsgal

Proof. Using known commutator and the fact that $[\mathrm{ad}(J_1), \mathrm{ad}(J_2)] = 0$ on $\tilde{\mathcal{N}}_2$, we find the following commutators:

LONGsubEqJunqzJdeuxqzmoinspun

$$[J_1, q_0] = -q_2 \qquad [J_2, q_0] = -p_1 \tag{79.367a}$$

LONGsubEqJunqzJdeuxqzmoinspdeux

$$[J_1, q_2] = -q_0 \qquad [J_2, q_2] = s_1 \tag{79.367b}$$

$$[J_1, p_1] = s_1 \qquad [J_2, p_1] = -q_0 \tag{79.367c}$$

$$[J_1, s_1] = p_1 \qquad [J_2, s_1] = q_2. \tag{79.367d}$$

From these properties, we deduce that $q_0 - q_2 - p_1 - s_1$ is proportional to X_{++}. Since, by definition, q_0 is the $\mathcal{KQ}$-component of X_{++}, the proportionality factor is 1. We also know that X_{+-} is proportional to $q_0 - q_2 + p_1 + s1$. Since $q_0 - q_2 = (X_{++})_{\mathcal{Q}} = (X_{+-})_{\mathcal{Q}}$ (proposition 79.67), the proportionality coefficient is 1. The relations (79.365) are now proved.

For the elements of $\tilde{\mathcal{N}}_k$, the commutation relations give

$$[J_1, q_k + r_k] = 0 \qquad [J_1, p_k - s_k] = s_k - p_k \tag{79.368a}$$

$$[J_2, q_k + r_k] = q_k + r_k \qquad [J_2, p_k - s_k] = 0 \tag{79.368b}$$

so that

$$\begin{aligned} q_k + r_k &\propto X^k_{0+} \in \mathcal{N} \\ s_k + p_k &\propto X^k_{+0} \in \mathcal{N} \\ p_k - s_k &\propto X^k_{-0} \in \bar{\mathcal{N}} \end{aligned} \tag{79.369}$$

We have $r_k = [J_2, q_k] \in \mathcal{K} \cap \mathcal{H}$, so that the $\mathcal{P}$-component of $q_k + r_k$ is q_k. But $q_k = (X^k_{0+})_{\mathcal{P}}$ is the $\mathcal{P}$-component of X^k_{0+}. The proportionality between $q_k + r_k$ and X^k_{0+} together with the equality of their $\mathcal{P}$-component provide the equality (79.366a).

For the two other, let us suppose that

$$X^k_{+0} = a(p_k + s_k) \tag{79.370a}$$

$$X^k_{-0} = b(p_k - s_k). \tag{79.370b}$$

In this case, we have

$$(X^k_{+0})_{\mathcal{P}} = \frac{1}{2}(X^k_{+0} - \theta X^k_{+0}) = \frac{1}{2}\big((a-b)p_k + (a+b)s_k\big), \tag{79.371}$$

so that $a = -b$ because $s_k \in \mathcal{K}$. Now let us look at the $\mathcal{KQ}$-component of the equality $[X^k_{+0}, X^k_{0+}] = -X_{++}$ taking into account the fact that $X^k_{+0} \in \mathcal{H}$ and $(X^k_{0+})_{\mathcal{KQ}} = 0$. What we have is

$$\big[(X^k_{+0})_{\mathcal{PH}}, (X^k_{0+})_{\mathcal{PQ}}\big] = -q_0, \tag{79.372}$$

but $(X^k_{0+})_{\mathcal{P}} = q_k$ and $(X^k_{+0})_{\mathcal{P}} = ap_k$, so that $[ap_k, q_k] = -q_0$. If we replace p_k by its definition $[q_0, q_k]$, we get

$$a\big[[q_0, q_k], q_k\big] = a\,\mathrm{ad}(q_k)^2 q_0 = -q_0, \tag{79.373}$$

so that $a = -1$.

The last point comes from $X^k_{0-} = \theta X^k_{0+}$.

□

Notice that this result was already obvious from the decompositions given in (79.306).

79.21.8 Some exponentials

LONGSubSecSomeExpo

It will be important to compute the element $e^{\mathrm{ad}(xq_0)}X$ when X runs over the vectors of $\mathbb{B}$. The action of e^{xq_0} on $\mathcal{A}$ is LONGSubEqsAdxqzJJ

$$e^{\mathrm{ad}(xq_0)}J_1 = \cos(x)J_1 + \sin(x)q_2 \tag{79.374a}$$

$$e^{\mathrm{ad}(xq_0)}J_2 = \sin(x)p_1 + \cos(x)q_1, \tag{79.374b}$$

on $\tilde{\mathcal{N}}_2$ we have LONGSubEqsexpxzqpsdzuu

$$e^{\mathrm{ad}(xq_0)}q_2 = \cos(x)q_2 - \sin(x)J_1 \tag{79.375a}$$
$$e^{\mathrm{ad}(xq_0)}q_0 = q_0 \tag{79.375b}$$
$$e^{\mathrm{ad}(xq_0)}p_1 = \cos(x)p_1 - \sin(x)q_1 \tag{79.375c}$$
LONGEqAdqzpUp
$$e^{\mathrm{ad}(xq_0)}s_1 = s_1, \tag{79.375d}$$

and the action on $\tilde{\mathcal{N}}_k$ is LONGEqExpAdqkpk

$$e^{\mathrm{ad}(xq_0)}q_k = \cos(x)q_k + \sin(x)p_k \tag{79.376a}$$
$$e^{\mathrm{ad}(xq_0)}p_k = \cos(x)p_k - \sin(x)q_k \tag{79.376b}$$
$$e^{\mathrm{ad}(xq_0)}p_k = \cos(x)p_k - \sin(x)q_k. \tag{79.376c}$$

Combining with lemma 79.85, LONGEqExpoQzSurNk

$$e^{\mathrm{ad}(xq_0)}X^k_{0+} = e^{\mathrm{ad}(xq_0)}(q_k + r_k) = r_k + \cos(x)q_k + \sin(x)p_k \tag{79.377a}$$
LONGEqExpoQzSurNka
$$e^{\mathrm{ad}(xq_0)}X^k_{+0} = -e^{\mathrm{ad}(xq_0)}(s_k + p_k) = -s_k - \cos(x)p_k + \sin(x)q_k. \tag{79.377b}$$
LONGEqExpoQzSurNkb

The same way, LONGEqExpoQzN

$$e^{\mathrm{ad}(xq_0)}X_{++} = q_0 + \sin(x)q_1 - \cos(x)q_2 + \sin(x)J_1 - \cos(x)p_1 \tag{79.378a}$$
$$e^{\mathrm{ad}(xq_0)}X_{+-} = q_0 - \sin(x)q_1 - \cos(x)q_2 + \sin(x)J_1 + \cos(x)p_1. \tag{79.378b}$$

The projections on $\mathcal{Q}$ of all these combinations are immediate.

79.21.9 Classification of the basis by the spaces

The basis (79.306) allows to generalize the theorem 79.76.

By very definition, we have $\mathrm{ad}(\mathcal{P})^2\mathcal{P} \subset \mathcal{P}$, $\mathrm{ad}(\mathcal{P})^2\mathcal{K} \subset \mathcal{K}$ and the same for the couple $(\mathcal{H}, \mathcal{Q})$. The relations (79.213) say that the same is true with the triple $(\tilde{\mathcal{N}}_2, \tilde{\mathcal{N}}_k, \mathcal{A})$, i.e. $\mathrm{ad}(\mathcal{A})^2\tilde{\mathcal{N}}_k \subset \tilde{\mathcal{N}}_k$ and $\tilde{\mathcal{N}}_k(\tilde{\mathcal{N}}_2) \subset \tilde{\mathcal{N}}_2$.

LONGThoAdSqIouZero

Theorem 79.86.
The basis $\mathbb{B}$ has the property to be stable under the commutators: $[X, Y] \in \{0, \pm\mathbb{B}\}$ when $X, Y \in \mathbb{B}$. Moreover, we have

$$\mathrm{ad}(X)^2Y = \begin{cases} 0 & \text{if } \mathrm{ad}(X)Y = 0 \\ Y & \text{if } X \in \mathcal{P} \\ -Y & \text{if } X \in \mathcal{K}. \end{cases} \tag{79.379}$$
LONGEqadXsqYzYmY

Proof. This theorem can be immediately checked using the commutators. It is however instructive to see that equation (79.379) can be checked from few considerations. First, remark that, if we look at the decompositions $\mathcal{G} = \mathcal{Q} \oplus \mathcal{H} = \mathcal{P} \oplus \mathcal{K} = \mathcal{Z}_{\mathcal{K}}(\mathcal{A}) \oplus \mathcal{A} \oplus \tilde{\mathcal{N}}_2 \oplus \tilde{\mathcal{N}}_k$, the commutation relations (79.213), we find

$$\mathrm{ad}(\tilde{\mathcal{N}}_2) \circ \mathrm{ad}(\tilde{\mathcal{N}}_2) : \begin{cases} \tilde{\mathcal{N}}_2 \to \mathcal{A} \to \tilde{\mathcal{N}}_2 \\ \tilde{\mathcal{N}}_k \to \tilde{\mathcal{N}}_k \to \tilde{\mathcal{N}}_k \\ \mathcal{A} \to \tilde{\mathcal{N}}_2 \to \mathcal{A}, \end{cases} \tag{79.380}$$

$$\mathrm{ad}(\tilde{\mathcal{N}}_k) \circ \mathrm{ad}(\tilde{\mathcal{N}}_k) : \begin{cases} \tilde{\mathcal{N}}_2 \to \tilde{\mathcal{N}}_k \to \mathcal{A} \oplus \tilde{\mathcal{N}}_2 \\ \tilde{\mathcal{N}}_k \to (\mathcal{A} \oplus \tilde{\mathcal{N}}_2) \to \mathcal{A} \oplus \tilde{\mathcal{N}}_k \\ \mathcal{A} \to \tilde{\mathcal{N}}_k \to \mathcal{A} \oplus \tilde{\mathcal{N}}_2, \end{cases} \tag{79.381}$$

$$\operatorname{ad}(\mathcal{A})\circ\operatorname{ad}(\mathcal{A})\colon\begin{cases}\tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_2\to\tilde{\mathcal{N}}_2\\ \tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k\to\tilde{\mathcal{N}}_k\\ \mathcal{A}\to 0.\end{cases}\tag{79.382}$$

Thus we have $\operatorname{ad}(X)^2Y=\lambda Y$ whenever we are not in the cases $(X,Y)\in(\tilde{\mathcal{N}}_k,\tilde{\mathcal{N}}_2)$ and $(X,Y)\in(\tilde{\mathcal{N}}_k,\tilde{\mathcal{N}}_k)$. We should check that these cases cannot bring a $\mathcal{A}$-component. We should also check that, since $\mathcal{K}\cap\mathcal{H}\cap\tilde{\mathcal{N}}_k$, is two dimensional, $\operatorname{ad}(X)^2r_k$ has no s_k-component as well as $\operatorname{ad}(X)^2s_k$ has no r_k-component.

Let us suppose that the spaces fit. We have $\operatorname{ad}(X)^2Y=\lambda Y$ and we still have to check the values of λ. Since $\mathbb{B}$ is orthonormal, we have

$$\lambda=\frac{B\big(\operatorname{ad}(X)^2Y,Y\big)}{B(Y,Y)}=-\frac{B\big(\operatorname{ad}(X)Y,\operatorname{ad}(X)Y\big)}{B(Y,Y)}.\tag{79.383}$$

First, remark that this is zero if and only if $\operatorname{ad}(X)Y=0$. Now, there are 4 possibilities following that X,Y belong to $\mathcal{P}$ or $\mathcal{K}$ because we know that $B|_{\mathcal{K}}<0$ and $B|_{\mathcal{P}}>0$. The result is that λ is positive when $X\in\mathcal{P}$ and negative when $X\in\mathcal{K}$. Now, the fact that $\mathbb{B}$ is stable under the commutators implies that $\operatorname{ad}(X)^2Y\in\{0,\pm\mathbb{B}\}$, so that $\lambda\in\{0,1,-1\}$. □

79.22 The causally singular structure

LONGSecBlacHole

79.22.1 Closed orbits

The singularity in AdS_l is defined as the closed orbits of AN and $A\bar{N}$ in G/H. This subsection is intended to identify them.

LONGPropCartanExtExpo

Proposition 79.87.
The Cartan involution $\theta\colon\mathcal{G}\to\mathcal{G}$ is an inner automorphism, namely it is given by $\theta=\operatorname{Ad}(k_\theta)$ where $k_\theta=e^{\pi q_0}$.

Proof. The operator $\operatorname{Ad}(k_\theta)$ acts as the identity on $\mathcal{K}$ because q_0 is central in $\mathcal{K}$ by definition. Looking at the decompositions (79.309) and (79.308), and taking into account that the result is already guaranteed on $\mathcal{K}$, we have to check the action of $\operatorname{Ad}(k_\theta)$ on J_1, J_2, q_k, p_k and p_1. It is done in setting $x=\pi$ in equations (79.374), (79.376) and (79.375c). What we get is that $\operatorname{Ad}(k_\theta)$ changes the sign on $\mathcal{P}$.

□

LONGPropKanUnicAbarN

Proposition 79.88.
For each $an\in AN$, there exists one and only one $k\in K$ such that $kan\in A\bar{N}$. There also exists one and only one $k\in K$ such that $ank\in A\bar{N}$.

Proof. For unicity, let $an\in AN$ and suppose that $k_1^{-1}an$ and $k_2^{-1}an$ both belong to $A\bar{N}$. Then there exist a_1, a_2, $\bar{n}_1$ and $\bar{n}_2$ such that $k_1^{-1}an=a_1\bar{n}_1$ and $k_2^{-1}an=a_2\bar{n}_2$ and we have

$$an=k_1a_1\bar{n}_1=k_2a_2\bar{n}_2.\tag{79.384}$$

By unicity of the decomposition $KA\bar{N}$, we conclude that $k_1=k_2$.

For the existence, let $an\in AN$ and consider the KAN decomposition $\theta(an)=ka'n'$. We claim that k^{-1} answers the question. Indeed, θ is the identity on K, so that $an=k\theta(a'n')$, and then

$$k^{-1}an=\theta(a'n')\in A\bar{N}.\tag{79.385}$$

One checks the statement about $ank\in A\bar{N}$ in much the same way. □

In the following results, we use the fact that the group K splits into the commuting product $K=\mathrm{SO}(2)\times\mathrm{SO}(l-1)$.

LONGLemExistxTqansAbarN

Lemma 79.89.
For every $an \in AN$, there exists $x \in [0, 2\pi[$ such that $[ane^{xq_0}] \in [A\bar{N}]$.

Proof. Let $k \in K$ such that $ank \in A\bar{N}$. The element k decomposes into $k = st$ with $s = e^{xq_0} \in \mathrm{SO}(2)$ and $t \in \mathrm{SO}(n) \subset H$. Thus $[ans] \in [A\bar{N}]$. □

LONGLemansse

Lemma 79.90.
If $[an] = [s]$ with $s \in \mathrm{SO}(2)$, then $s = e$.

Proof. The assumption implies that there exists a $h \in H$ such that $an = sh$. Using the KAN decomposition of H, such a h can be written under the form $h = ta'n'$ with $t \in \mathrm{SO}(n)$. Thus we have $an = sta'n'$. By unicity of the decomposition kan, we must have $st = e$, and then $s = e$. □

LONGLemANksk

Lemma 79.91.
If $an \in AN$ and if $[ank_\theta] = [s]$ with $s \in \mathrm{SO}(2)$, then $s = k_\theta$.

The hypothessis implies that there exists a $h \in H$ such that $ank_\theta = sh$. The element h can be decomposed as

$$h = h_K h_{AN} k_\theta \tag{79.386}$$

with $h_K \in K$ and $h_{AN} \in AN$ but with no warranty that h_K or h_{AN} belong to H. We have

$$an = \underbrace{sh_K}_{\in K} h_{AN}, \tag{79.387}$$

so that $sh_K = e$ by unicity of the KAN decomposition. In particular $h_K \in \mathrm{SO}(2)$.

Now we want to prove that $h_K = k_\theta$ because it would implies $s = k_\theta$ by the relation $sh_K = e$. If $h_K = e^{yq_0}k_\theta$ we have

$$h = h_K h_{AN} k_\theta = e^{yq_0} k_\theta h_{AN} k_\theta = e^{yq_0} h_{A\bar{N}}. \tag{79.388}$$

where $h_{A\bar{N}} = \mathbf{Ad}(k_\theta)h_{AN}$. By unicity of the $KA\bar{N}$ decomposition in H, the latter relation shows that e^{yq_0} has to be the compact part of H and then belong to K_H which is only possible when $y = 0$.

LONGThoOrbitesOuverttes

Theorem 79.92.
The closed orbits of AN in AdS_l are $[AN]$ and $[ANk_\theta]$ where k_θ is the element of K such that $\theta = \mathrm{Ad}(k_\theta)$. The closed orbits of $A\bar{N}$ are $[A\bar{N}]$ and $[A\bar{N}k_\theta]$. The other orbits are open.

Proof. Let us deal with the AN-orbits in order to fix the ideas. First, remark that each orbit of AN pass trough $[SO(2)]$. Indeed, each $[ank]$ is in the same orbit as $[k]$ with $k \in K = \mathrm{SO}(2) \otimes \mathrm{SO}(n)$. Since $\mathrm{SO}(n) \subset H$, we have $[k] = [s]$ for some $s \in \mathrm{SO}(2)$.

We are thus going to study openness of the AN-orbit of elements of the form $[e^{xq_0}]$ because these elements are "classifying" the orbits. Using the isomorphism $dL_{g^{-1}} : T_{[g]}(G/H) \to \mathcal{Q}$, we know that a set $\{X_1, \dots X_l\}$ of vectors in $T_{[e^{xq_0}]}AdS_l$ is a basis if and only if the set $\{dL_{e^{-xq_0}}X_i\}_{i=1,\dots l}$ is a basis of $\mathcal{Q}$. We are thus going to study the elements

$$\begin{aligned} dL_{e^{-xq_0}} X^*_{[e^{xq_0}]} &= dL_{e^{-xq_0}} \frac{d}{dt}\Big[\pi\big(e^{-tX}e^{xq_0}\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\pi\big(\,\mathbf{Ad}(e^{-xq_0})e^{-tX}\big)\Big]_{t=0} \\ &= -\,\mathtt{proj}_{\mathcal{Q}}\, e^{-\,\mathrm{ad}(xq_0)} X \end{aligned} \tag{79.389}$$

when X runs over the elements of $\mathcal{A} \oplus \mathcal{N}$ and $\mathtt{proj}_{\mathcal{Q}}$ stands for the projection on $\mathcal{Q}$ parallel to $\mathcal{H}$.

The projections on $\mathcal{Q}$ of equations (79.374), (79.377) and (79.378) are LONGSubEqsExpAdxqDivers

$$\mathtt{proj}_{\mathcal{Q}}\left(e^{\mathrm{ad}(xq_0)}J_1\right) = \sin(x)q_2 \tag{79.390a}$$
$$\mathtt{proj}_{\mathcal{Q}}\left(e^{\mathrm{ad}(xq_0)}J_2\right) = \cos(x)q_1 \tag{79.390b}$$
$$\mathtt{proj}_{\mathcal{Q}}\left(e^{xq_0}X_{++}\right) = q_0 + \sin(x)q_1 - \cos(x)q_2 \tag{79.390c}$$
$$\mathtt{proj}_{\mathcal{Q}}\left(e^{\mathrm{ad}(xq_0)}X_{+-}\right) = q_0 - \sin(x)q_1 - \cos(x)q_2 \tag{79.390d}$$
$$\mathtt{proj}_{\mathcal{Q}}\left(e^{\mathrm{ad}(xq_0)}(s_k - p_k)\right) = \sin(x)q_k \tag{79.390e}$$
$$\mathtt{proj}_{\mathcal{Q}}\left(e^{\mathrm{ad}(xq_0)}(q_k + r_k)\right) = \cos(x)q_k. \tag{79.390f}$$

It is immediately visible that an orbit trough $[e^{xq_0}]$ is open if and only if $\sin(x) \neq 0$. It remains to study the orbits of $[e^{\pi q_0}]$ and $[e]$. Lemma 79.90 shows that these two orbits are disjoint.

Let us now prove that $[AN]$ is closed. A point outside $\pi(AN)$ reads $\pi(ans)$ where s is an elements of $\mathrm{SO}(2)$ which is not the identity. Let $\mathcal{O}$ be an open neighborhood of ans in G such that every element of $\mathcal{O}$ read $a'n's't'$ with $s' \neq e$. The set $\pi(\mathcal{O})$ is then an open neighborhood of $\pi(ans)$ which does not intersect $[AN]$. This proves that the complementary of $[AN]$ is open. The same holds for the orbit $[A\bar{N}]$.

The orbit $[ANk_\theta]$ and $[A\bar{N}k_\theta]$ are closed too because $ANk_\theta = k_\theta A\bar{N}$.

□

Lemma 79.93.
We have $[AN] \cap [ANk_\theta] = \varnothing$.

Proof. A representative of an element in $[AN] \cap [ANk_\theta]$ can be written $an = a'n'k_\theta h$, and then we have $k_\theta h \in AN$. Decomposing h into its components $h_K \in K$ and $h_{AN} \in AN$, we see that $k_\theta h_K \in AN$, which is impossible because $k_\theta \in \mathrm{SO}(2)$ while $K_H = \mathrm{SO}(l-1)$. □

LONGPropUniquexxxxANANbarktheta

Proposition 79.94.
Let $an \in AN$. *Then there exist unique* x_0, x_1, x_2, x_3 *in* $[0, 2\pi[$ *such that*

(1) $[ane^{x_0q_0}] \in [AN]$, *and* $ane^{x_0q_0} = an$,

(2) $[ane^{x_1q_0}] \in [ANk_\theta]$, *and* $ane^{x_1q_0} = ank_\theta$,

(3) $[ane^{x_2q_0}] \in [A\bar{N}]$, *and* $ane^{x_2q_0} = a'\bar{n}h_K$ *(lemma 79.89)*,

(4) $[ane^{x_3q_0}] \in [A\bar{N}k_\theta]$, *and* $ane^{x_3q_0} = a'\bar{n}k_\theta h_K$.

These numbers satisfy $x_0 = 0$, $x_1 = \pi$ *and* $x_3 = x_2 + \pi$ *modulo* 2π.

Moreover if P *does not belong to* $[AN] \cap [A\bar{N}]$ *nor to* $[AN] \cap [A\bar{N}k_\theta]$, *these are four different numbers.*

Proof. Existence comes from lemma 79.89. Now we discuss the unicity.

(1) Let x_0' such that $[ane^{x_0'q_0}] \in [AN]$. In that case, there exist $a'n' \in AN$ and $h \in H$ such that $e^{x_0'q_0} = a'n'$. If we decompose $h = h_{AN}h_K$, unicity of the ANK decomposition show that $e^{x_0'q_0} = h_K$, which is only possible when $x_0' = 0$ and $h_K = e$.

(2) If $ane^{x_1'q_0} = a'n'k_\theta h$, then there exist $a''n''$ such that $e^{x_1'q_0} = a''n''k_\theta h$. Taking the class, $[a''n''k_\theta] = [e^{x_1'q_0}] \in [\mathrm{SO}(2)]$. Then $e^{x_1'q_0} = k_\theta$ by lemma 79.91.

(3) Let x_2 such that $ane^{x_2q_0}t = a'\bar{n}$ with $t \in \mathrm{SO}(l-1)$ (proposition 79.88), and suppose that we have an other x_2' such that $[ane^{x_2'}] \in [A\bar{N}]$. We have the system

$$\begin{cases} ane^{x_2q_0}t = a'\bar{n} & \text{(79.391a)} \\ ane^{x_2'q_0} = a''\bar{n}'h. & \text{(79.391b)} \end{cases}$$

We extract an from the first equation and we put the result in the second one. Taking into account the fact that t commutes with e^{xq_0} and that $t \in H$ and renaming $h \to ht$, if we decompose $h = h_{A\bar{N}}h_K$ we find

$$a'\bar{n}e^{(x_2'-x_2)q_0} = a''\bar{n}'h_{A\bar{N}}h_K. \tag{79.392}$$

By unicity of the decomposition $A\bar{N}K$ we find $h_K = e^{(x_2'-x_2)q_0}$. The left hand side belongs to $\mathrm{SO}(l-1)$ and the right hand side belongs to $\mathrm{SO}(2)$, so that one has to have $x_2' = x_2$.

(4) If we consider $x_3 = x_2 + \pi$ we have $ane^{x_3q_0}t = a'\bar{n}k_\theta$ (with $t \in \mathrm{SO}(l-1)$) and if we consider an other x_3' such that $[ane^{x_3'q_0}] \in [A\bar{N}k_\theta]$, we have the system

$$\begin{cases} ane^{(x_2+\pi)q_0}t = a'\bar{n}k_\theta & \text{(79.393a)} \\ ane^{x_3'q_0} = a''\bar{n}'k_\theta h. & \text{(79.393b)} \end{cases}$$

Some manipulations including a redefinition of h to include t yield

$$e^{(x_3'-x_3)q_0} = k_\theta a_0\bar{n}_0k_\theta h, \tag{79.394}$$

but $k_\theta A\bar{N}k_\theta \subset AN$, thus

$$e^{(x_3'-x_3)q_0} = a_1n_1h_{AN}h_K \tag{79.395}$$

so that unicity of the decomposition ANK implies, $e^{(x_3'-x_3)q_0=h_K}$, so that $x_3' = x_3$.

Let us now prove the second part. We suppose that $[P]$ does not belong to $[AN] \cap [A\bar{N}]$. If $x_0 = x_2 = 0$, we have $[P] \in [AN] \cap [A\bar{N}]$. If $x_0 = x_3$, then we have $an = a'\bar{n}k_\theta h_K$, and then $[an] \in [AN] \cap [A\bar{N}k_\theta]$ since h_K commutes with k_θ.

If $x_1 = x_2$, then $an = a'\bar{n}k_\theta h_K$, so that $[an] \in [AN] \cap [A\bar{N}k_\theta]$. If $x_1 = x_3$, then $ank_\theta = a'\bar{n}k_\theta h_K$, so that $an \in [AN] \cap [A\bar{N}]$. □

79.22.2 Vanishing norm criterion

In the preceding section, we defined the singularity by means of the action of an Iwasawa group. We are now going to give an alternative way of describing the singularity, by means of the norm of a fundamental vector of the action. This "new" way of describing the singularity is, in fact, much more similar to the original BTZ black hole where the singularity was created by identifications along the integral curves of a Killing vector field[1031]. The vector J_1 in theorem 79.95 plays here the role of that "old" Killing vector field.

Discrete identifications along the integral curves of J_1 would produce the causally singular space which is at the basis of our black hole.

What we will prove is the

LONGThosSequivJzero

Theorem 79.95.
We have $\mathcal{S} \equiv \|J_1^*\|_{[g]} = \|\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})J_1\| = 0$.

Thanks to this theorem, our strategy will be to compute $\|\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})J_1\|$ in order to determine if $[g]$ belongs to the singularity or not. The proof will be decomposed in three steps. The first step is to obtain a manageable expression for $\|J_1^*\|$.

LONGLemExpressionCoolNormJUn

Lemma 79.96.
We have $\|(J_1^*)_{[g]}\| = \|\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})J_1\|$ *for every* $[g] \in AdS_l$.

Proof. By definition,

$$(J_1^*)_{[g]} = \frac{d}{dt}\Big[\pi(e^{-tJ_1}g)\Big]_{t=0} = -d\pi dR_gJ_1. \tag{79.396}$$

The norm of this vector is the norm induced from the Killing form on $\mathcal{G}$[958]. First we have to put dR_gJ_1 under the form dL_gX with $X \in \mathfrak{g}$. One obviously has $dR_gJ_1 = dL_g\,\mathrm{Ad}(g^{-1})J_1$, and the

norm to be computed is

$$\begin{aligned}\|J_1^*\|_{[g]} = \|d\pi_g dL_g \operatorname{Ad}(g^{-1})J_1\|_{[g]} &= \|d\pi_g dL_g \,\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1\|_{[g]} \\ &= \|dL_g \,\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1\|_g \\ &= \|\,\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1\|_e\end{aligned} \tag{79.397}$$

□

LONGPropPtpsSjzero

Proposition 79.97.
If $p \in \mathscr{S}$, then $\|J_1^*\|_p = 0$.

Proof. We are going to prove that $\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1$ is a light like vector in $\mathcal{Q}$ when g belongs to $[AN]$ or $[A\bar{N}]$. A general element of AN reads $g = a^{-1}n^{-1}$ with $a \in A$ and $n \in N$. Since $\operatorname{Ad}(a)J_1 = J_1$, we have $\operatorname{Ad}(g^{-1})J_1 = \operatorname{Ad}(n)J_1$. We are going to study the development

$$\operatorname{Ad}(e^Z)J_1 = e^{\operatorname{ad}(Z)}J_1 = J_1 + \operatorname{ad}(Z)J_1 + \frac{1}{2}\operatorname{ad}(Z)^2 J_1 + \dots \tag{79.398}$$

where $Z = \ln(n) \in \mathcal{N}$. The series is finite because Z is nilpotent (see theorem 79.74 for more informations) and begins by J_1 while all other terms belong to $\mathcal{N}$. Notice that the same remains true if one replaces $\mathcal{N}$ by $\bar{\mathcal{N}}$ everywhere.

Moreover, $\operatorname{Ad}(e^Z)J_1$ has no X_{0+}-component (no X_{0-}-component in the case of $Z \in \bar{\mathcal{N}}$) because $[X_{0+}, J_1] = 0$, so that the term $[Z, J_1]$ is a combination of X_{+0}, X_{++} and X_{+-}. Since the action of $\operatorname{ad}(X_{+\pm})$ on such a combination is always zero, the next terms are produced by action of $\operatorname{ad}(X_{0+})$ on a combination of X_{+0}, X_{++} and X_{+-}. Thus we have

$$\operatorname{Ad}(e^Z)J_1 = J_1 + aX_{++} + bX_{+-} + c_k X_{+0}^k \tag{79.399}$$

LONGEqAdanJUnabck

for some [19] constants a, b and c_k.

The projection of $\operatorname{Ad}(e^Z)J_1$ on $\mathcal{Q}$ is made of a combination of the projections of X_{+0}, X_{++} and X_{+-}. From the definitions (79.245), we have $\mathtt{proj}_{\mathcal{Q}} X_{++} = q_0 + q_2$, lemma 79.69 implies $\mathtt{proj}_{\mathcal{Q}} X_{+0} = 0$ and lemma 79.66 yields $\mathtt{proj}_{\mathcal{Q}} X_{+-} = -\sigma\,\mathtt{proj}_{\mathcal{Q}} X_{++} = q_0 + q_2$. The conclusion is that $\mathtt{proj}_{\mathcal{Q}}\left(e^{\operatorname{ad}(Z)}J_1\right)$ is a multiple of $q_0 + q_2$, which is light like. The conclusion still holds with $\bar{\mathcal{N}}$, but we get a multiple of $q_0 - q_2$ instead of $q_0 + q_2$.

Now we have $\operatorname{Ad}(k_\theta)J_1 = J_1$ and $\operatorname{Ad}(k_\theta)(q_0 \pm q_2) = -(q_0 \pm q_2)$, so that the same proof holds for the closed orbits $[ANk_\theta]$ and $[A\bar{N}k_\theta]$. □

LONGRemANANbarYapas

Remarque 79.98.
The coefficients a, b and c_k in equation (79.399) are continuous functions of the starting point $an \in AN$. More precisely, they are polynomials in the coefficients of X_{++}, X_{+-}, X_{0+} and X_{0+} in Z. The vector $\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1 = (a+b)(q_0+q_1)$ is thus a continuous function of the point $[g] \in [AN]$.

If $[g] \in [AN] \cap [A\bar{N}]$, then $\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(g^{-1})J_1$ has to vanish as it is a multiple of $q_0 + q_1$ and of $q_0 - q_1$ in the same time. We conclude that in each neighborhood in $[AN]$ of an element of $[AN]$, there is an element which does not belong to $[A\bar{N}]$.

Proposition 79.99.
If $\|J_1^*\|_p = 0$, then $p \in \mathcal{S}$.

Proof. As before we are looking at a point $[g] = [(an)^{-1}s^{-1}]$ with $s = e^{xq_0}$. The norm $\|J_1^*\|$ vanishes if

$$\|\,\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}(e^{xq_0}) \operatorname{Ad}(an)J_1\| = 0. \tag{79.400}$$

19. One can show that every combinations of these elements are possible, but that point is of no importance here.

We already argued in the proof of proposition 79.97 that $\mathrm{Ad}(an)J_1$ is equal to J_1 plus a linear combination of X_{++}, X_{+-} and X_{+0}. Using the relations (79.390), we see that

$$\begin{aligned}\mathtt{proj}_{\mathcal{Q}}\, e^{\mathrm{ad}(xq_0)}(J_1 + aX_{++} + bX_{+-} + \sum_k c_k X_{+0}^k) \\ = (a+b)q_0 + (a-b)\sin(x)q_1 + \big(\sin(x) - (a+b)\cos(x)\big)q_2 + \sum_k c_k \sin(x) q_k.\end{aligned} \tag{79.401}$$

LONGEqprQexpxqzXanroots

The norm of this vector, as function of x, is given by

$$n(x) = (a+b)\sin(2x) + (4ab - C^2 - 1)\big(1 - \cos(2x)\big), \tag{79.402}$$

where $C^2 = \sum_k c_k^2$. Using the variables $u = a + b$ and $v = (1 + c^2 - 4ab)/2$,

$$n(x) = u\sin(2x) + v\cos(2x) - v. \tag{79.403}$$

Following $u = 0$ or $u \neq 0$, the graph of that function vanishes two or four times between 0 and 2π, see figure 79.2. Points of $[AN]$ are divided into two parts: the *red points* which correspond to $u \neq 0$, and the *blue points* which correspond to $u = 0$. By continuity, the red part is open.

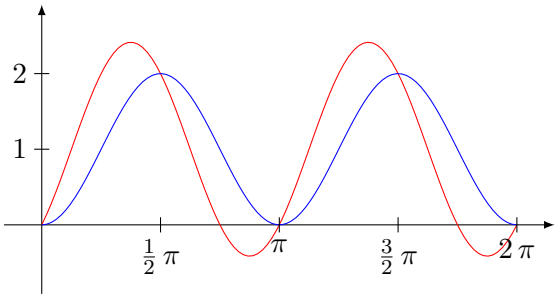


Figure 79.2: In red, the function $n(x)$ with $u \neq 0$ and in blue, the function with $u = 0$.

LabelFigRLuqsrr

Let $P = an \in AN$. By proposition 79.94, we consider the unique x_0, x_1, x_2 and x_3 in $[0, 2\pi[$ such that

$$[Pe^{x_0 q_0}] \in [AN] \tag{79.404a}$$
$$[Pe^{x_1 q_0}] \in [ANk_\theta] \tag{79.404b}$$
$$[Pe^{x_2 q_0}] \in [A\bar{N}] \tag{79.404c}$$
$$[Pe^{x_3 q_0}] \in [A\bar{N}k_\theta]. \tag{79.404d}$$

They satisfy $x_0 = 0$, $x_1 = \pi$ and $x_3 = x_2 + \pi$ modulo 2π. Now, we divide $[AN]$ into two parts. The elements of $[AN] \cap [A\bar{N}]$ and $[AN] \cap [A\bar{N}k_\theta]$ are said to be of *type I*, while the other are said to be of *type II*. We are going to prove that type I points are exactly blue points, while type II points are the red ones.

If P is a point of type II, we know that the x_i are four different numbers so that the norm function $n_P(x)$ vanishes *at least* four times on the interval $[0, 2\pi[$, each of them corresponding to a point in the singularity. But our division of $[AN]$ into red and blue points shows that $n_P(x)$ can vanish *at most* four times. We conclude that a point of type II is automatically red, and that the four roots of $n_P(x)$ correspond to the four values x_i for which $Pe^{x_i q_0}$ belongs to the singularity. The proposition is thus proved for points of type II.

Let now P be of type I (say $P \in [AN] \cap [A\bar{N}]$) and let us show that P is blue. We consider a sequence of points P_k of type II which converges to P (see remark 79.98). We already argued that P_k is red, so that $x_0(P_k) \neq x_2(P_k)$ and $x_1(P_k) \neq x_3(P_k)$, but

$$x_0(P_k) - x_2(P_k) \to 0 \tag{79.405a}$$
$$x_1(P_k) - x_3(P_k) \to 0. \tag{79.405b}$$

The continuity of $n_Q(x)$ with respect to both $x \in [0, 2\pi[$ and $Q \in [AN]$ implies that P has to be blue, and then $n_P(x)$ vanishes for exactly two values of x which correspond to the points Pe^{xq_0} in the singularity.

Let us now prove that everything is done. We begin by points of type I. Let P be of type I and say $P \in [AN] \cap [A\bar{N}]$. The curve $n_P(x)$ vanishes exactly two times in $[0, 2\pi[$. Now, if $Pe^{x_1 q_0} \in [ANk_\theta]$, thus $x_1 = \pi$ and we also have $Pe^{x_1 q_0} \in [A\bar{N}k_\theta]$, but P does not belong to $[ANk_\theta]$, which proves that $n_P(x)$ vanishes *at least* two times which correspond to the points Pe^{xq_0} that are in the singularity. Since the curve vanishes in fact exactly two times, we conclude that $n_P(x)$ vanishes if and only if Pe^{xq_0} belongs to the singularity.

If we consider a point P of type II, we know that the values of x_i are four different numbers, so that the curve $n_P(x)$ vanishes *at least* four times, corresponding to the points Pe^{xq_0} in the singularity. Since the curve is in fact red, it vanishes *exactly* four times in $[0, 2\pi[$ and we conclude that the curve $n_P(x)$ vanishes if and only if Pe^{xq_0} belongs to the singularity.

The conclusion follows from the fact that

$$AdS_l = \Big\{ [Pe^{xq_0}] \text{ tel que } P \text{ is of type I or II and } x \in [0, 2\pi[\Big\}. \tag{79.406}$$

□

Proof of theorem 79.95 is now complete.

79.22.3 Existence of the black hole

LONGSubSecExistenceTrouNoir

We know that the geodesic trough $[g]$ in the direction X is given by

$$\pi\big(ge^{sX}\big) \tag{79.407}$$

and that a geodesics is light-like when the **direction** X is given by a nilpotent element in $\mathcal{Q}$[1019].

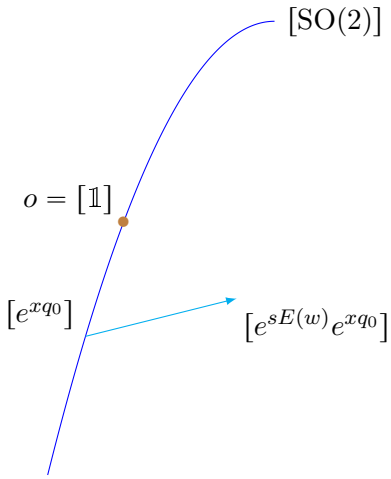


Figure 79.3: We are looking at a geodesics issued from one point of the line $[\mathrm{SO}(2)] = \{e^{xq_0}\}_{x\in[0,2\pi[}$. Here, $E(w) = q_0 + w_1q_1 + w_2q_2 + \sum_{k\geqslant 3} w_kq_k$ with $\sum_k w_k^2 = 1$.

LabelFigDTIYKkP

Let us study the geodesics issued from the point $[e^{-xq_0}]$, see figure 79.3. They are given by

$$l_x^w(s) = \pi\big(e^{-xq_0}e^{sE(w)}\big) \tag{79.408}$$

where $E(w) = q_0 + \sum_i w_iq_i$ with $\|w\| = 1$. According to our previous work, the point $l_x^w(s)$ belongs to the singularity if and only if

$$n_x^w(s) = \Big\| \mathtt{proj}_{\mathcal{Q}}\, e^{-\mathrm{ad}(sE(w))} e^{\mathrm{ad}(xq_0)} J_1 \Big\|^2 = 0. \tag{79.409}$$

LONGEqNormAFaireZeroOuPas

We already computed that $e^{\mathrm{ad}(xq_0)}J_1 = \cos(x)J_1 + \sin(x)q_2$. By construction, $E(w)$ is nilpotent and $\mathrm{ad}(E)^3 = 0$ by proposition 79.74. Using the fact that $[\mathcal{Q},\mathcal{H}] \subset \mathcal{Q}$ and $[\mathcal{Q},\mathcal{Q}] \subset \mathcal{H}$, we collect the terms in $\mathcal{Q}$ in the development of the exponential. The $\mathcal{Q}$ component of

$$e^{-s\,\mathrm{ad}(E)}\big(\cos(x)J_1 + \sin(x)q_2\big) \tag{79.410}$$

is

$$\ell = \frac{s^2}{2}\sin(x)\,\mathrm{ad}(E)^2q_2 - s\cos(x)\,\mathrm{ad}(E)J_1 + \sin(x)q_2. \tag{79.411}$$

The square norm of that expression is *a priori* a polynomial of order 4. Hopefully, the coefficient of s^4 contains

$$B\big(\mathrm{ad}(E)^2q_2, \mathrm{ad}(E)^2q_2\big), \tag{79.412}$$

while the coefficient of s^3 is given by

$$B\big(\mathrm{ad}(E)J_2, \mathrm{ad}(E)^2q_2\big). \tag{79.413}$$

Both of these two expressions are zero because the ad-invariance of the Killing form makes appear $\mathrm{ad}(E)^3$. Equation (79.409) is thus the second order polynomial given by

$$\begin{aligned} n_x^w(s) = {} & s^2\sin^2(x)B\big(\mathrm{ad}(E)^2q_2, q_2\big) \\ & + s^2\cos^2(x)B\big(\mathrm{ad}(E)J_1, \mathrm{ad}(E)J_1\big) \\ & - 2s\cos(x)\sin(x)B\big(\mathrm{ad}(E)J_1, q_2\big) \\ & + \sin^2(x)B(q_2,q_2). \end{aligned} \tag{79.414}$$

LONGEqnwxBBB

The problem now reduces to the evaluation of the three Killing products in this expression. Let us begin with $B\big(\mathrm{ad}(E)^2q_2, q_2\big)$. For this one, we need to know the q_2-component of $\mathrm{ad}(E)^2q_2$. We have to review all the possibilities $\mathrm{ad}(q_i)\,\mathrm{ad}(q_j)q_2$ and determine which one(s) have a q_2-component.

In this optic, let us recall that q_2 is characterised by

$$q_2 \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2. \tag{79.415}$$

All the combinations $\mathrm{ad}(q_i)^2$ work. Since $\mathcal{G} = \mathcal{K} \oplus \mathcal{P}$ is reductive and since q_0 is the only basis element of $\mathcal{Q}$ to belong to $\mathcal{K}$, none of the combinations $\mathrm{ad}(q_i)\,\mathrm{ad}(q_0)$ or $\mathrm{ad}(q_0)\,\mathrm{ad}(q_i)$ work. Using the relations

$$\begin{aligned} &q_1 \in \mathcal{P} \cap \mathcal{Q} \cap \mathcal{A}, \quad && [\tilde{\mathcal{N}}_k, \tilde{\mathcal{N}}_2] \subset \tilde{\mathcal{N}}_k \\ &q_2 \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_2, \quad && [\tilde{\mathcal{N}}_2, \tilde{\mathcal{N}}_2] \subset \mathcal{A} \\ &q_k \in \mathcal{P} \cap \mathcal{Q} \cap \tilde{\mathcal{N}}_k, \quad && [\tilde{\mathcal{N}}_k, \tilde{\mathcal{N}}_k] \subset \mathcal{A} \oplus \tilde{\mathcal{N}}_2, \end{aligned} \tag{79.416}$$

we check that the only working combinations are $\mathrm{ad}(q_i)^2$.

Thus, the only elements $\mathrm{ad}(q_i)\,\mathrm{ad}(q_j)q_2$ which have a q_2-component are $\mathrm{ad}(q_i)^2q_2$, while theorem 79.76 says that this component is q_2 for $2 \neq i \neq 0$ and $-q_2$ for $i = 0$. Therefore, the q_2-component of $\mathrm{ad}(E)^2q_2$ is

$$\mathrm{ad}(q_0)^2q_2 + w_1^2\,\mathrm{ad}(q_1)^2q_2 + \sum_{k\geqslant 3} w_k^2\,\mathrm{ad}(q_k)^2q_2 = -w_2^2q_2 \tag{79.417}$$

where we used the fact that $\sum_i w_i^2 = 1$. Thus we have

$$B\big(\mathrm{ad}(E)^2q_2, q_2\big) = -w_2^2B(q_2,q_2). \tag{79.418}$$

LONGEqBeDeuxqqwDeux

Let us now search for the q_2-component of $\mathrm{ad}(E)J_1$. We have $[q_1, J_1] \in [\mathcal{A},\mathcal{A}] = 0$, $[q_k, J_1] = 0$ (equation (79.282b)), and $[q_2, J_1] = -q_0$, $[q_0, J_1] = -q_2$ (equation (79.248)). Then, we have

$$\mathrm{ad}(E)J_1 = w_2q_0 + q_2. \tag{79.419}$$

That implies

$$B\big(\mathrm{ad}(E)J_1, q_2\big) = B(q_2,q_2), \tag{79.420}$$

and

$$B\big(\operatorname{ad}(E)J_1, \operatorname{ad}(E)J_1\big) = B(q_2, q_2) + w_2^2 B(q_0, q_0). \tag{79.421}$$

Equation (79.414) now reads

$$\frac{n_x^w(s)}{B(q_2, q_2)} = \big(\cos^2(x) - w_2^2\big)s^2 - 2\cos(x)\sin(x)s + \sin^2(x). \tag{79.422}$$

We have $n_x^w(s) = 0$ when s equals

$$s_\pm = \frac{\cos(x)\sin(x) \pm |w_2 \sin(x)|}{\cos^2(x) - w_2^2}. \tag{79.423}$$

If $w_2 \sin(x) \geqslant 0$, we have [20]

$$s_+ = \frac{\sin(x)}{\cos(x) - w_2} \quad \text{and} \quad s_- = \frac{\sin(x)}{\cos(x) + w_2}, \tag{79.424}$$

LONGEqRacineselITN

and if $w_2 \sin(x) < 0$, we have to exchange s_+ with s_-.

If we consider a point e^{xq_0} with $\sin(x) > 0$ and $\cos(x) < 0$, the directions w with $|w_2| < |\cos(x)|$ escape the singularity as the two roots (79.424) are simultaneously negative. Such a point does not belong to the black hole. That proves that the black hole is not the whole space.

If we consider a point e^{xq_0} with $\sin(x) > 0$ and $\cos(x) > 0$, we see that for every w_2, we have $s_+ > 0$ or $s_- > 0$ (or both). That shows that for such a point, every direction intersect the singularity. Thus the black hole is actually larger than only the singularity itself.

The two points with $\sin(x) = 0$ belong to the singularity. At the points $\cos(x) = 0$, $\sin(x) = \pm 1$, we have $s_+ = -1/w_2$ and $s_- = 1/w_2$. A direction w escapes the singularity only if $w_2 = 0$ (which is a closed set in the set of $\|w\| = 1$).

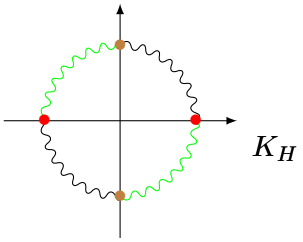


Figure 79.4: Points in $\pi(K)$ are classified by their angle in SO(2). Red points are part of the singularity, points in the black zone belong to the black hole and points in the green zone are free. The upper and lower boundaries belong to the horizon.

LabelFigSFdgHdO

79.23 Some more computations

LONGSecMoreComputations

As we saw, the use of theorem 79.95 leads us to study the function

$$n_{[g]}^w(s) = \|\,\mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}\big(e^{-sE(w)}\big)X\|^2 = 0 \tag{79.425}$$

LONGEqnrSqbignor

where $E(w) = q_0 + w_1 q_1 + \cdots + w_{l-1} q_{l-1}$ ($w \in S^{l-2}$) and (see equation (79.399))

$$X = e^{\operatorname{ad}(xq_0)} \operatorname{Ad}(na) J_1 = e^{\operatorname{ad}(xq_0)}\big(J_1 + aX_{++} + bX_{+-} + \sum_{k=3}^{l-1} c_k X_{+0}^k\big). \tag{79.426}$$

From proposition 79.74 we have $\operatorname{ad}(E)^3 = 0$ and the exponential in equation (79.425) contains only three terms. Using the fact that $\mathtt{proj}_{\mathcal{Q}} \operatorname{ad}(E)X = \operatorname{ad}(E)X_{\mathcal{H}}$ and we are then lead to study the norm of

$$X_{\mathcal{Q}} - s\operatorname{ad}(E)X_{\mathcal{H}} + \frac{s^2}{2}\operatorname{ad}(E)^2 X_{\mathcal{Q}}. \tag{79.427}$$

20. The solutions (79.424) were already deduced in [1019] in a quite different way; these are equations (79.101).

Notice that, since $\mathcal{Q}$ is Killing-orthogonal to $\mathcal{H}$, we have $B(X_{\mathcal{Q}}, Y_{\mathcal{Q}}) = B(X, Y_{\mathcal{Q}})$. Thus we have

$$n^w_{[g]}(s) = \| \mathtt{proj}_{\mathcal{Q}} \operatorname{Ad}\big(e^{-sE(w)}\big)X\|^2 = a(E)s^2 + b(E)s + c \tag{79.428}$$

where LONGSubEqsabcEBBB

$$a(E) = -B\big(\operatorname{ad}(E)X, \sigma \operatorname{ad}(E)X\big) \tag{79.429a}$$
LONGEqCosfaEBX
$$b(E) = -2B\big(X_{\mathcal{Q}}, \operatorname{ad}(E)X_{\mathcal{H}}\big) \tag{79.429b}$$
$$c = B(X_{\mathcal{Q}}, X_{\mathcal{Q}}). \tag{79.429c}$$

It is convenient to decompose X into $X = X_2 + \sum_k c_k X_k$ with $X_2 \in \tilde{\mathcal{N}}_2$ and $X_k \in \tilde{\mathcal{N}}_k$ as well as $E = E_2 + \sum_k E_k$. Using the decompositions (79.365), the first part is

$$\begin{aligned} X_2 &= \operatorname{Ad}(e^{xq_0})\big(a(q_0 - q_2) + b(p_1 + s_1)\big) \\ X_k &= \operatorname{Ad}(e^{xq_0})(p_k + s_k) \end{aligned} \tag{79.430}$$
LONGEqXsiAdSTrois

where we have renamed a, b and c_k. Using the exponentials (79.375) and (79.377b) we have

$$X_2 = aq_0 - b\sin(x)q_1 - a\cos(x)q_2 + a\sin(x)J_1 + bs_1 + b\cos(x)p_1 \tag{79.431a}$$
LONGsubEqXtroisdonne
$$X_k = -s_k - \cos(x)p_k + \sin(x)q_k \tag{79.431b}$$
LONGsubEqXkdonne

Let us first look at the case we encounter in AdS_3, i.e. with $X_k = 0$ and $E_k = 0$. In this case, a general direction is given by

$$E_2(w) = E(\theta) = q_0 + \cos(\theta)q_1 + \sin(\theta)q_2. \tag{79.432}$$

We achieve the computation of $e^{\operatorname{ad}(E(\theta))}X_2$ using the known commutators: LONGSubEqsEXtrois

$$\begin{aligned} \operatorname{ad}(E_2)X_{\mathcal{Q}} = {} & J_1\big(a\sin(\theta) + a\cos(x)\big) \\ & + p_1\big(-a\cos(\theta) - b\sin(x)\big) \\ & + s_1\big(b\sin(x)\sin(\theta) - a\cos(x)\cos(\theta)\big). \end{aligned} \tag{79.433a}$$

and

$$\begin{aligned} \operatorname{ad}(E_2)X_{\mathcal{H}} = {} & q_0\big(a\sin(x)\sin(\theta) - b\cos(x)\cos(\theta)\big) \\ & - q_1 b\big(\sin(\theta) + \cos(x)\big) \\ & + q_2\big(a\sin(x) + b\cos(\theta)\big). \end{aligned} \tag{79.433b}$$

Putting together and writing $\big(\cos(\theta), \sin(\theta)\big) = (w_1, w_2)$,

$$\begin{aligned} \operatorname{ad}(E_2)X_2 = {} & J_1\big(a\cos(x) + aw_2\big) \\ & p_1\big(-b\sin(x) - aw_1\big) \\ & s_1\big(-a\cos(x)w_1 + b\sin(x)w_2\big) \\ & q_0\big(-b\cos(x)w_1 + a\sin(x)w_2\big) \\ & q_1\big(-bw_2 - b\cos(x)\big) \\ & q_2\big(bw_1 + a\sin(x)\big) \end{aligned} \tag{79.434}$$
LONGEqadEtroissurXtrois

Using the expression (79.434) among with the norms and collecting the terms with respect to the dependence in θ, we have

$$\begin{aligned} \frac{B\big(\operatorname{ad}(E)X_{\mathcal{Q}}, \operatorname{ad}(E)X_{\mathcal{Q}}\big)}{B(q_0, q_0)} = {} & a^2\cos^2(x) - a^2 \\ & + \sin(\theta)\big(-2a^2\cos(x)\big) \\ & + \cos(\theta)\big(-2ab\sin(x)\big) \\ & + \cos^2(\theta)\big(a^2\cos^2(x) - b^2\sin^2(x)\big) \\ & + \sin(\theta)\cos(\theta)\big(-2ab\sin(x)\cos(x)\big), \end{aligned} \tag{79.435}$$
LONGEqKillingEXQEXQ

and

$$\begin{aligned}\frac{B\big(\mathrm{ad}(E)X_{\mathcal{H}},\mathrm{ad}(E)X_{\mathcal{H}}\big)}{B(q_0,q_0)} &= -b^2(1+\cos^2(x))\\ &\quad+\sin(\theta)\big(-2b^2\cos(x)\big)\\ &\quad+\cos(\theta)\big(-2ab\sin(x)\big)\\ &\quad+\cos^2(\theta)\big(b^2\cos^2(x)-a^2\sin^2(x)\big)\\ &\quad+\sin(\theta)\cos(\theta)\big(-2ab\sin(x)\cos(x)\big)\end{aligned} \tag{79.436}$$

and finally, in the case of AdS_3 we have

$$\begin{aligned}a(E) &= B\big(\mathrm{ad}(E)X_{\mathcal{H}},\mathrm{ad}(E)X_{\mathcal{H}}\big)-B\big(\mathrm{ad}(E)X_{\mathcal{Q}},\mathrm{ad}(E)X_{\mathcal{Q}}\big)\\ &= (a^2-b^2)\big(\cos^2(x)+\sin(\theta)\cos(x)+\sin^2(\theta)\big).\end{aligned} \tag{79.437}$$

LONGEqaEdansAdsTrois

For investigating the higher dimensional cases, we decompose $\mathrm{ad}(E)X$ into the four parts $\mathrm{ad}(E_2)X_2$, $\mathrm{ad}(E_2)X_k$, $\mathrm{ad}(X_k)X_2$ and $\mathrm{ad}(E_k)X_k$.

The action of E_2 on $\tilde{\mathcal{N}}_k$ is given by

$$\mathrm{ad}(E_2)r_k = w_1q_k \tag{79.438a}$$
$$\mathrm{ad}(E_2)p_k = -q_k \tag{79.438b}$$
$$\mathrm{ad}(E_2)q_k = p_k+w_1r_k-w_2s_k \tag{79.438c}$$
$$\mathrm{ad}(E_2)s_k = -w_2q_k. \tag{79.438d}$$

Thus

$$\mathrm{ad}(E_2)X_k = q_k\big(w_2+\cos(x)\big)+p_k\big(\sin(x)\big)+r_k\big(w_1\sin(x)\big)+s_k\big(-w_2\sin(x)\big) \tag{79.439}$$

LONGadEsurXki

The same way we find

$$\mathrm{ad}(E_k)X_2 = \sum_{k=3}^{l-1}\Big[p_k(-aw_k)+r_k\big(b\sin(x)w_k\big)+s_k\big(-a\cos(x)w_k\big)\Big]. \tag{79.440}$$

Using the relations

$$\mathrm{ad}(q_k)X_k = -q_2+\cos(x)q_0. \tag{79.441}$$

and

$$\mathrm{ad}(E)q_0 = -w_1p_1+w_2J_1-\sum_{k=1}^{l}w_kp_k \tag{79.442a}$$
$$\mathrm{ad}(E)q_1 = p_1-w_2s_1-\sum_{k=3}^{l}w_kr_k \tag{79.442b}$$
$$\mathrm{ad}(E)q_2 = -J_1+w_1s_1+\sum_{k=3}^{l}s_k \tag{79.442c}$$
$$\mathrm{ad}(E)q_k = p_k+w_1r_k-w_2s_k, \tag{79.442d}$$

we find

$$\mathrm{ad}(E_k)X_{k'} = \begin{cases}c_{k'}w_k\sin(x) & \text{if } k\neq k'\\ -c_kw_k\cos(q_0)-c_kw_kq_2 & \text{if } k=k'\end{cases} \tag{79.443}$$

Putting all together, we find

$$\begin{aligned}\mathrm{ad}(E)X = {} & J_1\big(a\cos(x)+aw_2\big)\\ & + p_1\big(-b\sin(x)-aw_1\big)\\ & + s_1\big(-a\cos(x)w_1+bw_2\sin(x)\big)\\ & + q_0\Big(-bw_1\cos(x)+aw_2\sin(x)+\sum_k c_kw_k\cos(x)\Big)\\ & + q_1\big(-bw_2-b\cos(x)\big)\\ & + q_2\Big(bw_1+a\sin(x)-\sum_k c_kw_k\Big)\\ & + \sum_k q_kc_k\big(w_2+\cos(x)\big)\\ & + \sum_k p_k\big(c_k\sin(x)-aw_k\big)\\ & + \sum_k r_k\big(c_kw_1+bw_k\big)\sin(x)\\ & + \sum s_k\big(-c_kw_2\sin(x)-aw_k\cos(x)\big)\\ & + \sum_{k\geqslant 3}\sum_{k'>k} r_{kk'}\big(c_kw_{k'}-c_{k'}w_k\big)\sin(x)\end{aligned} \tag{79.444}$$

It is quite easy but long to compute $a(E)$, $b(E)$ and c from that expression. The results are LONGEqCoefsabcBE

$$\frac{a(E)}{B(q_0,q_0)} = M\Big(w_2^2+\cos(x)w_2+\cos^2(x)\Big) \tag{79.445a}$$ LONGEqCoeffaE

$$\frac{b(E)}{B(q_0,q_0)} = -2M\sin(x)\big(w_2+\cos(x)\big)\big) \tag{79.445b}$$

$$\frac{c}{B(q_0,q_0)} = M\sin^2(x) \tag{79.445c}$$

where $M=\big(a^2-b^2-\sum_k c_k^2\big)$. The important point to notice is that these expressions only depend on the w_2-component of the direction. Notice that $c=0$ if and only if the point $[g]$ belongs to the singularity because $s=0$ is a solution of (79.425) only in the case $[g]\in\mathscr{S}$.

We can avoid the computation of a certain number of terms by exploiting the properties of the decomposition $\mathcal{G}=\mathcal{Z}_{\mathcal{K}}(\mathcal{A})\oplus\tilde{\mathcal{N}}_2\oplus\tilde{\mathcal{N}}_k\oplus\mathcal{A}$. The dependence in w_1^2 of $a(E)$ is given by the term

$$B\big(\mathrm{ad}(q_1)X,\sigma\,\mathrm{ad}(q_1)X\big) = B\big(\mathrm{ad}(q_1)^2X,\sigma X\big). \tag{79.446}$$

This is easily computed using the theorems 79.86 and 79.83. The result is that the coefficient of w_1^2 in $a(E)/B(q_0,q_0)$ is

$$-\sin^2(x)(a^2-b^2-C^2) \tag{79.447}$$

where $C^2=\sum_{k\geqslant 3}c_k^2$.

The term which does not depend on w is

$$B\big(\mathrm{ad}(q_0)X,\sigma\,\mathrm{ad}(q_0)X\big) = B\big(\mathrm{ad}(q_0)^2X,\sigma X\big). \tag{79.448}$$

The result is that the independent term in $a(E)/B(q_0,q_0)$ is

$$\big(1-2\cos^2(x)\big)(a^2-b^2-C^2). \tag{79.449}$$

In the same way, the coefficient of w_2^2 is $B\big(\mathrm{ad}(q_2)^2X,\sigma X\big)$ and we find

$$-\big(\sin^2(x)+1\big)(a^2-b^2-C^2). \tag{79.450}$$

LONGRemImapoabcE

Remarque 79.100.
Importance of the coefficients (79.445). If $v \in \mathscr{F}_l$, there is a direction E_0 in AdS_l which escapes the singularity from v. Thus the polynomial $a(E_0)s^2 + b(E_0)s + c$ has only non positive roots. From the expressions (79.445), we see that the polynomial corresponding to $\iota(v)$ is the same, so that the direction E_0 escapes the singularity from $\iota(v)$ as well. This is the main ingredient of the next section.

79.24 Description of the horizon

LONGSecHorizonSansMatrices

79.24.1 Induction on the dimension

The horizon in AdS_3 is already well understood [1017]. We are not going to discuss it again. We will study how does the causal structure (black hole, free part, horizon) of AdS_l includes itself in AdS_{l+1} by the inclusion map[1032]

$$\iota\colon AdS_l \to AdS_{l+1}. \tag{79.451}$$

We will use the following notations. The symbols $\mathscr{S}_l$, BH_l, $\mathscr{F}_l$ and $\mathscr{H}_l$ respectively denote the singularity, the black hole, the free part and the horizon in AdS_l. If $D \subset AdS_l$, then $\mathrm{Int}(D)$ and $\mathrm{Adh}(D)$ are the interior and the closure of D.

LONGLemMemeQueLemQuatre

Lemma 79.101.
Let $[g] \in \iota(AdS_3) \subset AdS_l$ be outside the singularity. We suppose that there is an open set $\mathcal{O}$ in S^1 of directions escaping the singularity from $[g]$. Then there exists an open set $\mathcal{O}'$ in S^{l-2} of directions escaping the singularity.

Proof. The proof is a consideration about the coefficients (79.429). The hypothesis means that the points

$$\pi\left(ge^{sE(\theta)}\right) \tag{79.452}$$

do not belong to $\mathscr{S}$ for $s \geqslant 0$ when $E(\theta) = q_0 + \cos(\theta)q_1 + \sin(\theta)q_2$ and θ belongs to the given open set $\mathcal{O} \subset [0, 2\pi]$. If $a(E_0) \neq 0$ for some $E_0 \in \mathcal{O}$, the solutions are given by

$$s_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{79.453}$$

In such a direction, there are two values, both outside[21] of $\mathbb{R}^+$, of s such that $[ge^{sE_0}] \in \mathscr{S}$. By continuity, we can find a neighborhood of E_0 in S^{l-2} such that $[ge^{sE}]$ belongs to the singularity only for non positive numbers.

A problem arises when $a(E) = 0$ for every direction E in the open set $\mathcal{O}$. In that case the equation (79.425) has only one solution which is negative by hypothesis. But it could appear that in every neighborhood of E, a second solution, positive, appears. If we write $X = \mathrm{Ad}(g^{-1})J_1$, what we have to prove is that the quantity

$$a(E) = B\big(\mathrm{ad}(E)X, \sigma\,\mathrm{ad}(E)X\big) \tag{79.454}$$

is not constant when E runs over $\mathcal{O}$, in particular, there exists a direction $\theta_0 \in \mathcal{O}$ such that $a(\theta_0) \neq 0$. We supposed that $[g] \in \iota(AdS_3)$, so that $X = \mathrm{Ad}(an)J_1$ is given by X_3 of equation (79.431a).

The function $a(E)$ (79.437) is analytic with respect to θ, thus if it vanishes on an open set $\mathcal{O}$, it has to vanish everywhere. This can only be achieved with $a = \pm b$. Now, simple computation show that

$$c = a^2 - b^2\sin^2(x) - a^2\cos^2(x) = (a^2 - b^2)\sin^2(x) \tag{79.455}$$

21. When we say "outside" of $\mathbb{R}^+$, we include the case of complex solutions.

which vanishes when $a = \pm b$, so that $a(E)$ can only be constant with respect to E on the singularity. Thus we conclude that $a(E)$ is not constant with respect to $E \in S^1$ outside the singularity.

This concludes the proof of lemma 79.101. □

Lemma 79.102.
The direction E_0 in AdS_l escapes the singularity from $v \in AdS_l$ if and only if it escapes the singularity from $\iota(v)$ in AdS_{l+1}.

Proof. The fact for v to escape the singularity in the direction E_0 means that the equation

$$a_v(E_0)s^2 + b_v(E_0) + c_v = 0 \tag{79.456}$$

where the coefficients are given by (79.445) has no positive solutions with respect to s. Since these coefficients are the same for v and $\iota(v)$, the equation for $\iota(v)$ is in fact the same and has the same solutions. □

As a warm up, let us prove the following, which is a particular case of lemma 79.104.

Lemma 79.103.
Let $[g] \in AdS_l$ be such that there exists an open set $\mathcal{O} \in S^1$ of directions that escape the singularity. Then there is an open set in S^{l-1} that escapes the singularity from $i[g] \in AdS_{l+1}$.

Proof. With the notations of section 79.23, we only have to compute $a(E)$ when $E = E_2$ and X is general. So we pick the expression (79.445a) and we put $w_2 = \cos(\theta)$ while $w_k = 0$ for every $k \geqslant 3$. What we have is

$$a(E) = M\Big(\cos^2(\theta) + \cos(x)\cos(\theta) + \cos^2(x)\Big) \tag{79.457}$$

If $a(E) = 0$ for every $\theta \in \mathcal{O}$, then $M = 0$ which is impossible since we suppose that the starting point does not belong to the singularity. □

LONGLemDueiINtlIntlpu

Lemma 79.104.
We have

$$\iota\big(\operatorname{Int}(\mathscr{F}_l)\big) \subset \operatorname{Int}(\mathscr{F}_{l+1}) \tag{79.458}$$

or, equivalently,

$$\operatorname{Adh}(BH_{l+1}) \cap \iota(AdS_l) \subset \iota\big(\operatorname{Adh}(BH_l)\big). \tag{79.459}$$

Proof. Let $v' \in \operatorname{Int}(\mathscr{F}_l)$ and $\mathcal{O}$, an open set of directions in AdS_l that escape the singularity. The coefficient $a_l(E)$ is not constant on $\mathcal{O}$ because the coefficient $M = a^2 - b^2 - C^2$ is only zero on the singularity (see equation (79.445a)). Thus we can choose $E_0 \in \mathcal{O}$ such that $a_l(E_0) \neq 0$. We consider $a_{l+1}(E_0)$, the coefficient of s^2 for the point $\iota(v')$ in the direction E_0. From the expression (79.445a) we know that $a_{l+1}(E_0) = a_l(E_0)$. The coefficients $b(E_0)$ and c are also the same for v' and $\iota(v')$.

Since $a(E_0) \neq 0$ and $v' \in \operatorname{Int}(\mathscr{F}_l)$, we have two solutions to the equation $a(E_0)s^2 + b(E_0)s + c = 0$ and both of these are outside $\mathbb{R}_0^+$. This conclusion is valid for $v' \in AdS_l$ as well as for $\iota(v') \in AdS_{l+1}$. Then there is a neighborhood of $\iota(v')$ on which the two solutions keep outside $\mathbb{R}_0^+$. That proves that $\iota(v') \in \operatorname{Int}(\mathscr{F}_{l+1})$.

For the second line, suppose that $v \in \iota(AdS_l)$ does not belong to $\iota\big(\operatorname{Adh}(BH_l)\big)$, thus $v \in \iota\big(\operatorname{Int}(\mathscr{F}_l)\big) \subset \operatorname{Int}(\mathscr{F}_{l+1})$. In that case v does not belong to $\operatorname{Adh}(BH_{l+1})$.

□

LONGProphFdanshF

Proposition 79.105.
We have

$$\mathscr{F}_{l+1} \cap \iota(AdS_l) \subset \iota(\mathscr{F}_l) \tag{79.460}$$

Proof. If $v = \iota(v') \in \mathscr{F}_{l+1}$, there is a direction E_0 in AdS_{l+1} which escape the singularity from v. That direction is given by a vector $(w_1, \cdots v_l) \in S^l$. Since the coefficients $a(E)$, $b(E)$ and c do only depend on w_2, a direction $(w'_1, \cdots, w'_{l-1}, 0)$ with $w'_2 = w_2$ escapes the singularity from v'. This proves that $v' \in \mathscr{F}_l$. □

LONGLemHiH

Lemma 79.106.
We have

$$\mathscr{H}_{l+1} \cap \iota(AdS_l) \subset \iota(\mathscr{H}_l). \tag{79.461}$$

Proof. First,

$$v \in \mathscr{H}_{l+1} \cap \iota(AdS_l) \subset \mathscr{F}_{l+1} \cap \iota(AdS_l) \subset \iota(\mathscr{F}_l) \tag{79.462}$$

from proposition 79.105. Now, let's take $v' \in \mathscr{F}_l$ such that $v = \iota(v')$. We have to prove that $v' \in \mathscr{H}_l$. Let us suppose that $v' \in \mathrm{Int}(\mathscr{F}_l)$, so $v \in \mathrm{Int}(\mathscr{F}_{l+1})$ because of lemma 79.104. This is in contradiction with the fact that $v \in \mathscr{H}_{l+1}$. □

LONGCorDeuxTrucsBHhH

Corollary 79.107.
We have

(1) $\iota(\mathscr{S}_l) \subset \mathscr{S}_{l+1}$,
(2) $\iota(\mathscr{F}_l) \subset \mathscr{F}_{l+1}$,
(3) $\iota(BH_l) \subset BH_{l+1}$,
(4) $\iota(\mathscr{H}_l) \subset \mathscr{H}_{l+1}$.

Proof. We have $\mathrm{Ad}\big(\iota(g^{-1})\big)J_1 = \mathrm{Ad}(g^{-1})J_1$, so that the condition of theorem (79.95) is invariant under ι. Thus one immediately has $\iota(\mathscr{S}_l) \subset \mathscr{S}_{l+1}$ and $\iota(\mathscr{F}_l) \subset \mathscr{F}_{l+1}$.

An element v which does not belong to BH_{l+1} belongs to $\mathscr{F}_{l+1}$, but if v belongs to $\iota(AdS_l) \cap \mathscr{F}_{l+1}$, it belongs to $\iota(\mathscr{F}_l)$ by proposition 79.105 and then does not belong to $\iota(BH_l)$. Thus $\iota(BH_l) \subset BH_{l+1}$.

Now if $v' \in \mathscr{H}_l$, let us consider $\mathcal{O}$, a neighborhood of $v = \iota(v')$ in AdS_{l+1}. The set $\iota^{-1}\big(\mathcal{O} \cap \iota(AdS_l)\big)$ contains a neighborhood $\mathcal{O}'$ of v' in AdS_l. Since $v' \in \mathscr{H}_l$, there is $\bar{v} \in \mathcal{O}'$ such that $\bar{v} \in BH_l$. Thus $\iota(\bar{v}) \in \mathcal{O}$ belongs to BH_{l+1} by the first item.

□

Let $X \in \mathcal{G}_{l+1}$ such that $[X, J_1] = 0$, and let R be the group generated by X. The following results are intended to show that such a group can be used in order to transport the causal structure from AdS_l to AdS_{l+1}

The key ingredient will be the fact that, since X commutes with J_1, we have

$$\mathrm{Ad}\left((ge^{sE})^{-1}\right)J_1 = \mathrm{Ad}\left((e^{\alpha X}ge^{sE})^{-1}\right)J_1. \tag{79.463}$$

LONGEqAdOkSurJun

LONGLemRSsubsetS

Lemma 79.108.
A group R as described above preserves the causal structure in the sense that

(1) $R \cdot \mathscr{S} \subset \mathscr{S}$
(2) $R \cdot BH \subset BH$
(3) $R \cdot \mathscr{F} \subset \mathscr{F}$
(4) $R \cdot \mathscr{H} \subset \mathscr{H}$.

Proof. From equation (79.463), we deduce that a direction E_0 will escape the singularity from the point $[g]$ is and only if it escapes the singularity from the points $r[g]$ for every $r \in R$. The first three points follow.

Now let $v \in \mathscr{H}$ and $r \in R$ and let us prove that $r \cdot v \in BH$. By the third point, $r \cdot v \in \mathscr{F}$. Let now $\mathcal{O}$ be a neighborhood of $r \cdot v$. The set $r^{-1} \cdot \mathcal{O}$ is a neighborhood of v and we can consider $\bar{v} \in BH \cap r^{-1} \cdot \mathcal{O}$. By the second point, $r \cdot \bar{v}$ is a point of the black hole in $\mathcal{O}$. □

LONGRemdqnqRSlsubsetSlpu

Remarque 79.109.
Combining corollary 79.107 and lemma 79.108 we have

(1) $R \cdot \iota(\mathscr{S}_l) \subset \mathscr{S}_{l+1}$
(2) $R \cdot \iota(\mathscr{F}_l) \subset \mathscr{F}_{l+1}$
(3) $R \cdot \iota(BH_l) \subset BH_{l+1}$
(4) $R \cdot \iota(\mathscr{H}_l) \subset \mathscr{H}_{l+1}$

where the dot stands for the action.

LONGThoCausalPasseParR

Theorem 79.110.
If moreover the one parameter group R has the property to generate AdS_{l+1} (in the sense that $R \cdot \iota(AdS_l) = AdS_{l+1}$), then we have

(1) $R \cdot \iota(\mathscr{S}_l) = \mathscr{S}_{l+1}$
(2) $R \cdot \iota(\mathscr{F}_l) = \mathscr{F}_{l+1}$
(3) $R \cdot \iota(BH_l) = BH_{l+1}$
(4) $R \cdot \iota(\mathscr{H}_l) = \mathscr{H}_{l+1}$.

LONGItemStrucalpb

Proof. The inclusions in the direct sense are already done in the remark 79.109.

Let $r = e^{\alpha X}$ be an element of R. Since, by assumption, we have $[X, J_1] = 0$, the action of r leaves invariant the condition of theorem (79.95):

$$\mathrm{Ad}\left((ge^{sE})^{-1}\right)J_1 = \mathrm{Ad}\left((e^{\alpha X}ge^{sE})^{-1}\right)J_1. \tag{79.464}$$

LONGEqalpharinagitpas

(1) Let $[g] \in \mathscr{S}_{l+1}$, there exists a $r \in R$ such that $r[g] \in \iota(AdS_l)$. There exists an element $g' \in \mathrm{SO}(2, l-1)$ such that $rg = \iota(g')$. Now $[g'] \in \mathscr{S}_l$ because

$$\mathrm{Ad}(g'^{-1})J_1 = \mathrm{Ad}(\iota(g'^{-1}))J_1 = \mathrm{Ad}\left((rg)^{-1}\right)J_1 = \mathrm{Ad}(g^{-1})J_1, \tag{79.465}$$

but by assumption the norm of the projection on $\mathcal{Q}$ of the right hand side is zero.

(2) If v is free in AdS_{l+1}, there is a direction E_0 escaping the singularity from v and an element $r \in R$ such that $v' = r \cdot v \in \iota(AdS_l)$. The point v' is also free in AdS_{l+1} as the direction E_0 works for $r \cdot v$ as well as for v. Thus by proposition 79.105 we have

$$v' \in \mathscr{F}_{l+1} \cap \iota(AdS_l) \subset \iota(\mathscr{F}_l) \tag{79.466}$$

and $v \in R \cdot \iota(\mathscr{F}_l)$.

(3) If $v \in BH_{l+1}$, the point $r \cdot v \in \iota(AdS_l)$ also belongs to BH_{l+1}. If $r \cdot v = \iota(v')$, then $v' \in BH_l$ from if $v' \in \mathscr{F}_l$, then $\iota(v') \in \mathscr{F}_{l+1}$.

(4) If $v \in \mathscr{H}_{l+1}$, there exists a $r \in R$ such that $v' = r \cdot v \in \iota(AdS_l)$ and, moreover, v' belongs to the horizon in AdS_{l+1} since the horizon is invariant under R. Thus v' belongs to $\iota(AdS_l) \cap \mathscr{H}_{l+1} \subset \iota(\mathscr{H}_l)$ by lemma 79.106. Now, $v \in R \cdot \iota(\mathscr{H}_l)$.

□

79.24.2 Examples of surjective groups

Theorem 79.110 describes the causal structure in AdS_l by induction on the dimension provided that one knows a group R such that $AdS_{l+1} = R \cdot \iota(AdS_l)$. Can one provide examples of such groups? The following proposition provides a one.

LONGPropSurjectif

Proposition 79.111.
If R is the one parameter subgroup of $\mathrm{SO}(2, l)$ generated by $r_{l,l+1}$, then we have

$$R \cdot \iota(AdS_l) = AdS_{l+1}. \tag{79.467}$$

Proof. If one realises AdS_n as the set of vectors of length 1 in $\mathbb{R}^{2,n-1}$, AdS_l is included in AdS_{l+1} as the set of vectors with vanishing last component and the element $r_{l,l+1}$ is the rotation in the plane of the two last coordinates. In that case, we have to solve

$$e^{\alpha r_{l,l+1}} \begin{pmatrix} u' \\ t' \\ x'_1 \\ \vdots \\ x'_{l-2} \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ t \\ x_1 \\ \vdots \\ x_{l-2} \\ x_{l-1} \end{pmatrix} \tag{79.468}$$

with respect to α, u', t' and x_i'. There are of course exactly two solution if $x_{l-2}^2 + x_{l-1}^2 \neq 0$. □

In fact, many others groups are available, as the one showed at the end of [1032]. In fact, since, in the embedding of AdS in $\mathbb{R}^{2,n}$, the singularity is given by $t^2 - y^2 = 0$, almost every group which leaves invariant the combination $t^2 - y^2$ can be used to propagate the causal structure. One can found lot of them for example by looking at the matrices given in [1021].

79.24.3 Backward induction

Using proposition 79.105, lemma 79.106, corollary 79.107 and the fact that the norm of J_1^* is the same in AdS_l as in $\iota(AdS_l) \subset AdS_{l+1}$, we have

(1) $\iota(\mathscr{F}_l) = \mathscr{F}_{l+1} \cap \iota(AdS_l)$,
(2) $\iota(\mathscr{H}_l) = \mathscr{H}_{l+1} \cap \iota(AdS_l)$,
(3) $\iota(\mathscr{S}_l) = \mathscr{S}_{l+1} \cap \iota(AdS_l)$.

These equalities hold for $l \geqslant 3$. For $l = 2$ we can take the latter as a definition and set

$$\mathscr{S}_2 = \{v \in AdS_2 \text{ tel que } \iota(v) \in \mathscr{S}_3\}. \tag{79.469}$$

The Iwasawa decomposition of $\mathrm{SO}(2,1)$ is given by $\mathcal{A} = \langle J_2 \rangle$, $\mathcal{N} = \langle X_+ \rangle$, $\mathcal{K} = \langle q_0 \rangle$ where $X_+ = p_1 - q_0$. Notice that we *do not* have $\iota(X_+) = X_{++}$. Instead we have $\iota(X_+) = \frac{1}{2}(X_{++} + X_{-+})$. Thus $\mathscr{S}_2$ is not given by the closed orbits of AN in AdS_2.

The light-like directions are given by the two vectors $E = q_0 \pm q_1$. In order to determine if the point $[e^{-\alpha J_2} e^{-aX_+} e^{-xq_0}]$ belongs to the black hole, we follow the same way as in section 79.22.3: we compute the norm

$$\| \mathtt{proj}_{\mathcal{Q}}\, e^{-s\,\mathrm{ad}(E)} e^{x\,\mathrm{ad}(q_0)} e^{a\,\mathrm{ad}(X_+)} e^{\alpha\,\mathrm{ad}(J_2)} J_1 \| \tag{79.470}$$

and we see under which conditions it vanishes.

$$e^{a\,\mathrm{ad}(X_+)} J_1 = J_1 + a(q_2 + s_1), \tag{79.471}$$

and then

$$\begin{aligned} X &= e^{x\,\mathrm{ad}(q_0)}\big(J_1 + a(q_2 + s_1)\big) \\ &= J_1\big(\cos(x) - a\sin(x)\big) + q_2\big(\sin(x) + a\cos(x)\big) + as_1. \end{aligned} \tag{79.472}$$

With $E = q_0 + q_1$, we have

$$\begin{aligned} e^{s\,\mathrm{ad}(E)} J_1 &= \big(-\frac{1}{2}s^2 + 1\big)J_1 + sq_2 + \frac{1}{2}s^2 s_1 \\ e^{s\,\mathrm{ad}(E)} q_2 &= -sJ_1 + q_2 + ss_1 \\ e^{s\,\mathrm{ad}(E)} s_1 &= -\frac{1}{2}s^2 J_1 + sq_2 + (\frac{1}{2}s^2 + 1)s_1. \end{aligned} \tag{79.473}$$

Thus

$$\mathtt{proj}_{\mathcal{Q}}\, e^{s\,\mathrm{ad}(E)} X = \Big(\big(\cos(x) - a\sin(x) + a\big)s + \big(\sin(x) + a\cos(x)\big)\Big)q_2. \tag{79.474}$$

Its norm vanishes for the value of s given by

$$s^+ = \frac{a\cos(x) + \sin(x)}{\big(\sin(x) - 1\big)a - \cos(x)}. \tag{79.475}$$

The same computation with $E = q_0 - q_1$ provides the value

$$s^- = \frac{a\cos(x) + \sin(x)}{\big(\sin(x) + 1\big)a - \cos(x)}. \tag{79.476}$$

For small enough a, the signs of s^+ and s^- are both given the sign of $-\tan(x)$ that can be either positive or negative. Thus there is an open set of points in AdS_2 which intersect the singularity in every direction and an open set of points which escape the singularity.

As a side note, the singularity $\mathscr{S}_2$ described here is not given by the closed orbits of AN or $A\bar{N}$. Indeed, we show that

$$\| \mathtt{proj}_{\mathcal{Q}}\, \mathrm{Ad}(e^{a\iota(X_+)}) J_1 \|^2 = \| \mathtt{proj}_{\mathcal{Q}}(J_1 + aq_2 + as_1) \| = -4a^2 \neq 0. \tag{79.477}$$

Thus the points of N are not part of the singularity.

79.25 Dirac operator on AdS_2

Why to compute Dirac operator on anti de Sitter spaces? Let $M = AdS_2$ and $R = AN$ acts on M. Let $\mathcal{O}$ be an open orbit of $R \times M \to M$. In the specific case of AdS_2, we have $R = \mathcal{O} = R \cdot \vartheta$. In larger dimensions, there is a $\mathrm{SO}(1,n)$ which causes that the orbit is not exactly the acting group. It is

$$\mathcal{O} = \frac{R}{R \cap \mathrm{SO}(1,n)}.$$

79.25.1 Clifford algebra and spin group

As definition, we retain

$$\begin{aligned} AdS_2 &\equiv t^2 + u^2 - x^2 = 1 \\ &= \frac{\mathrm{SO}(2,1)}{\mathrm{SO}(1,1)}. \end{aligned} \tag{79.478}$$

Let $V = \mathbb{R}^{1,1}$ and e_0, e_1 an orthonormal basis. We pose

$$f_0 = \frac{1}{2}(e_0 + e_1) \quad g_0 = \frac{1}{2}(e_0 - e_1)$$

and we define $\tilde{\rho}$ by

$$\tilde{\rho}(f_0)\alpha = f_0 \wedge \alpha \tag{79.479a}$$
$$\tilde{\rho}(g_0)\alpha = -i(g_0)\alpha \tag{79.479b}$$

where $\alpha \in \Lambda W$, W being the space spanned by f_0. More explicitly we have:

$$\tilde{\rho}(f_0)1 = f_0 \qquad \tilde{\rho}(f_0)f_0 = 0 \tag{79.480a}$$
$$\tilde{\rho}(g_0)1 = 0 \qquad \tilde{\rho}(g_0)f_0 = -\eta(f_0, g_0). \tag{79.480b}$$

As element of ΛW, f_0 stands for $\eta(f_0, .)$. If we choose the basis

$$1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

the matrices of $\tilde{\rho}$ are given by

$$\tilde{\rho}(e_0) = \begin{pmatrix} 0 & -1/2 \\ 1 & 0 \end{pmatrix}, \quad \tilde{\rho}(e_1) = \begin{pmatrix} 0 & 1/2 \\ 1 & 0 \end{pmatrix}.$$

Up to a change of basis,

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \gamma_0\gamma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and a general element of $Cl_{(1,1)}$ reads

$$x\gamma_0 + y\gamma_1 + u\mathbb{R} + v\gamma_0\gamma_1 = \begin{pmatrix} u+v & x-y \\ x+y & u-v \end{pmatrix}.$$

With the change of basis $e_1 \to ie_1$, we write it under a more simple form:

$$Cl_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{79.481}$$

with $\alpha, \beta \in \mathbb{C}$. In particular, an element of V, i.e. a combination of γ_0 and γ_1 is

$$V \rightsquigarrow \begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}. \tag{79.482}$$

Let

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

Let us now determine α, the extension of $-\operatorname{Id}|_V$ into an automorphism and τ, the extension of $\operatorname{Id}|_V$ into an anti-automorphism. We have $\alpha(b) = -b$, $\alpha(c) = -c$, $\tau(b) = b$ and $\tau(c) = c$. We find the others by virtue of relations $bc = -a$ and $b^2 = 1$. Finally

$$\begin{aligned} \alpha(1) &= 1 \quad & \alpha(a) &= a \\ \alpha(b) &= -b \quad & \alpha(c) &= -c \end{aligned} \tag{79.483}$$

and

$$\begin{aligned} \tau(1) &= 1 \quad & \tau(a) &= -a \\ \tau(b) &= b \quad & \tau(c) &= c. \end{aligned} \tag{79.484}$$

The condition for $s \in Cl_{(1,1)}$ to belongs to $\Gamma_{(1,1)}$ is that $\alpha(s)vs^{-1} \in V$ for all $v \in V$. If we consider $s = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}$, we have

$$\alpha(s) = \begin{pmatrix} \alpha & -\beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \text{ and } s^{-1} = \frac{1}{|\alpha|^2 - |\beta|^2} \begin{pmatrix} \bar{\alpha} & -\beta \\ -\bar{\beta} & \alpha \end{pmatrix}.$$

If we impose $\alpha(s)vs^{-1}$ to be of the form $\begin{pmatrix} 0 & \eta \\ \bar{\eta} & 0 \end{pmatrix}$ for all v of the form $\begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}$, we find $\operatorname{Re}(\bar{\alpha}\beta\bar{\xi}) = 0$ and the $\bar{\alpha}\beta = 0$. So generators of $\Gamma_{(1,1)}$ are

$$\Gamma_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}, \quad \begin{pmatrix} 0 & \beta \\ \bar{\beta} & 0 \end{pmatrix}. \tag{79.485}$$

Elements of $Spin_{(1,1)}$ are elements of $\Gamma^+_{(1,1)}$ such that $\tau(s) = s^{-1}$. So

$$Spin_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}, \text{ such that } |\alpha|^2 = 1. \tag{79.486}$$

We recognize $Spin_{(1,1)} = U(1)$.

¡¡ Avertissement/question au lecteur !! 79.112
This is wrong: in fact $\mathrm{Spin}(1,1) \neq U(1)$.

79.25.2 Relation between $SU(1,1)$ and $\mathrm{SO}(2,1)$

A general matrix of $SU(1,1)$ is

$$U = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad U^{-1} = \begin{pmatrix} \bar{\alpha} & -\beta \\ -\bar{\beta} & \alpha \end{pmatrix}$$

with $|\alpha|^2 - |\beta|^2 = 1$. They are matrices which fulfil $\det U = 1$ and $U^+ g = gU^{-1}$. If we denote by V the space of matrices of the form $(r,z) = \begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}$ with $r \in \mathbb{R}$ and $z \in \mathbb{C}$, we have a bijection $\psi\colon \mathbb{R}^{1,1} \to V$ given by

$$\begin{pmatrix} u \\ t \\ x \end{pmatrix} \mapsto \begin{pmatrix} x & t - iu \\ t + iu & x \end{pmatrix}.$$

It becomes an isometry if we pose $\|(r,z)\| = z\bar{z} - r^2 = -\det(r,z)$. The group $SU(1,1)$ has an isometric action on V given by

$$Uv = UvU^\dagger.$$

We immediately remark that $Uv = (-U)v$. We define

$$\begin{aligned} T\colon\, & SU(1,1) \to \mathrm{SO}(2,1) \\ & T(U)\begin{pmatrix} u \\ t \\ x \end{pmatrix} = \psi^{-1}\big(U\psi\begin{pmatrix} u \\ t \\ x \end{pmatrix}U^\dagger\big). \end{aligned} \tag{79.487}$$

Now we want to know when $T(U) = T(\tilde{U})$. Using the fact that $U^{-1} = gU^\dagger g$ in the condition $UvU^\dagger = \tilde{U}v\tilde{U}^\dagger$, we find

$$VvV^\dagger = v$$

with $V = \tilde{U}^{-1}U$. Then imposing

$$\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}\begin{pmatrix} \bar{\alpha} & \beta \\ \bar{\beta} & \alpha \end{pmatrix},$$

we find $T(U) = T(\tilde{U}) \Leftrightarrow \tilde{U} = \pm U$. We have

$$T\begin{pmatrix} i & \\ & -i \end{pmatrix} = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}.$$

The map $T\colon SU(1,1) \to \mathrm{SO}(2,1)$ is a double covering.

We are now going to explicitly compute the map T. First:

$$\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}\begin{pmatrix} \bar{\alpha} & \beta \\ \bar{\beta} & \alpha \end{pmatrix} = \begin{pmatrix} \bar{\alpha}(\alpha r + \beta z) + \bar{\beta}(\alpha\bar{z} + \beta r) & \beta(\alpha r + \beta z) + \alpha(\alpha\bar{z} + \beta r) \\ \bar{\alpha}(\bar{\beta} + \bar{\alpha}z) + \bar{\beta}(\bar{\beta}\bar{z} + \bar{\alpha}r) & \beta(\bar{\beta}r + \bar{\alpha}z) + \alpha(\bar{\beta}\bar{z} + \bar{\alpha}r) \end{pmatrix}.$$

When we pose $z = 0$ and $r = 1$, i.e., when we look at $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, we find

$$\begin{pmatrix} \bar{\alpha}\alpha + \bar{\beta}\beta & 2\beta\alpha \\ 2\bar{\alpha}\bar{\beta} & \beta\bar{\beta} + \alpha\bar{\alpha} \end{pmatrix}$$

which corresponds to $x = \bar{\alpha}\alpha$, $t - iu = 2\alpha\beta$ and $t + iu = 2\bar{\alpha}\bar{\beta}$. We conclude that

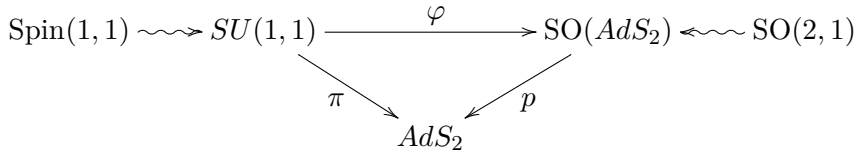
$$T\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} . & . & i(\alpha\beta - \bar{\alpha}\bar{\beta}) \\ . & . & \alpha\beta + \bar{\alpha}\bar{\beta} \\ . & . & \alpha\bar{\alpha} + \beta\bar{\beta} \end{pmatrix}.$$

Similar computations lead to

$$T\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} \frac{\bar{\alpha}^2 + \alpha^2 - \beta^2 - \bar{\beta}^2}{2} & \frac{i}{2}(\alpha^2 - \bar{\alpha}^2 + \beta^2 - \bar{\beta}^2) & i(\alpha\beta - \bar{\alpha}\bar{\beta}) \\ \frac{i}{2}(\beta^2 - \bar{\beta}^2 + \bar{\alpha}^2 - \alpha^2) & \frac{1}{2}(\alpha^2 + \bar{\alpha}^2 + \beta^2 + \bar{\beta}^2) & \alpha\beta + \bar{\alpha}\bar{\beta} \\ i(\bar{\alpha}\beta - \bar{\beta}\alpha) & \bar{\alpha}\beta + \bar{\beta}\alpha & \alpha\bar{\alpha} + \beta\bar{\beta} \end{pmatrix} \tag{79.488}$$

79.25.3 Spin structure on AdS_2

We are going to build elements of the following spin structure:



First let $\{e_t, e_u, e_x\}$ be a basis of $\mathbb{R}^{1,1}$ with $e_t \in AdS_2$, $e_t \cdot e_t = e_u \cdot e_u = -e_x \cdot e_x = -1$. We suppose that e_u and e_x span tangent space at e_t. Let T be a representation of $\mathrm{SO}(1,1)$ on $\mathbb{R}^{2,1}$ which leaves e_t unchanged: $T(A)e_t = e_t$ for all $A \in \mathrm{SO}(1,1)$. To each element $B \in \mathrm{SO}(AdS_2)$, one can associate an element of $B' \in \mathrm{SO}(AdS_2)$ such that B has the form

$$B = \{B'e_u, B'e_x\}_{B'e_t}. \tag{79.489}$$

eq_blaseAdS

We define

$$p(B) = B'e_t.$$

Now the action of $A \in \mathrm{SO}(1,1)$ on $B \in \mathrm{SO}(AdS_2)$ is defined, if B has the form (79.489), by

$$B \cdot A = \{T(A)B'e_u, T(A)B'e_x\}_{B'e_t}. \tag{79.490}$$

The map $\varphi\colon SU(1,1) \to \mathrm{SO}(AdS_2)$ is given by

$$\big(\varphi(U)\big)' = (T \circ S)(U), \tag{79.491}$$

and the projection $\pi\colon SU(1,1) \to AdS_2$, $\pi = p \circ \varphi$.

The group $\mathrm{Spin}(1,1)$ must act on $SU(1,1)$; we define

$$U \cdot s = \varphi^{-1}\big(\varphi(U) \cdot \chi(s)\big). \tag{79.492}$$

We have $\pi(U \cdot s) = \pi(U)$ because

$$\pi(U \cdot s) = p\big(\varphi(U) \cdot \chi(s)\big), = \big[\varphi(U) \cdot \chi(s)\big]'e_t = \varphi(U) \circ T(\chi(s))e_t = \varphi(U)'e_t = \pi(U).$$

We have used the fact that $\chi(s) \in \mathrm{SO}(1,1)$ and that, therefore, $(T \circ \chi)(s)e_t = e_t$.

79.25.4 Spinor bundle and connection

We define $S = \Lambda W$ where W is the (one dimensional) space spanned by f_0 and we define

$$\mathcal{S} = SU(1,1) \times_\rho S \tag{79.493}$$

eq_mSSUrho

where $\rho\colon Spin_{(1,1)} \times \Lambda W \to \Lambda W$ is the representation of $Spin_{(1,1)}$ on $SU(1,1)$ given by

$$\rho(s, \alpha) = \tilde{\rho}(s)\alpha. \tag{79.494}$$

Recall that α is either a scalar either a multiple of f_0. The equivalence relation which arises in equation (79.493) is

$$(U, \alpha) \sim (U \cdot s, \rho(s^{-1})\alpha). \tag{79.495}$$

The projection is

$$\pi_{\mathcal{S}}[(U, \alpha)] = \pi(U).$$

For the connection on $\mathrm{SO}(AdS_2)$, we want that horizontal vector are tangent vectors to curves formed by parallel transport. In other word, a path

$$B(s) = \{B'(s)e_u, B'(s)e_x\}_{B'(s)e_t}$$

has horizontal tangent vector if $B'(s)e_i$ $(i = u, x)$ is a parallel transport of $B'(0)e_i$ along the curve $B'(s)e_e$ on AdS_2. Here, $B'(s)$ denotes the matrix of $\mathrm{SO}(2,1)$ associated with the basis $B(s)$: the prime doesn't denotes a derivation. Let us define the $so(1,1)$ valued connection 1-form which corresponds to this intuition. We consider $b_i(s)$ the parallel transported along the curve $B'(s)e_t$ of $B'(0)e_i$, and $A(s)$, the matrix of $\mathrm{SO}(1,1)$ such that $A(s)B'(s)e_i = b_i(s)$ $(i = u, x)$. The definition is

$$\omega(\dot{B}) = \frac{d}{ds}\Big[A(s)\Big]_{s=0}.$$

Proposition 79.113.
It is a connection 1*-form.*

Proof. First we consider a fundamental vector field

$$X_B^* = \frac{d}{dt}\Big[B \cdot e^{-tX}\Big]_{t=0} = \frac{d}{dt}\Big[\{T(e^{-tX})B'e_u, T(e^{-tX})B'e_x\}_{B'e_t}\Big]_{t=0}.$$

The path in AdS_2 on which this path in $\mathrm{SO}(AdS_2)$ is build is constant: it is $B'e_t$. So the parallel transport is constant and the path $A(s)$ is given by

$$A(s)T(e^{-tX})B'e_u = B'e_u$$

and $\omega(X_B^*) = X$.

It remains to be proved that for all $B \in \mathrm{SO}(AdS_2)$, $g \in \mathrm{SO}(1,1)$ and $X \in T_B\,\mathrm{SO}(AdS_2)$,

$$\omega\big((dR_g)_B X\big) = \mathrm{Ad}(g^{-1})\omega_B(X). \tag{79.496}$$

We give X by the path

$$X(s) = \{B'(s)e_u, B'(s)e_x\}_{B'(s)e_t}.$$

The differential dR_g gives rise to the new path

$$(dR_g X)(s) = \{gB'(s)e_u, gB'(s)e_x\}_{B'(s)e_t}.$$

Let $b_i(s)$ be the parallel transport of $B'(0)e_i$ $(i = u, x)$ along the path $B'(s)e_t$. We have to compute $\omega_B(X)$ with $A(s)$ defined by $A(s)B'(s)e_i = b_i$. The parallel transport of $gB'(0)e_i$ is given by gb_i. Therefore $\omega(dR_g X)$ is given by the path $A^g(s)$ which satisfies $A^g(s)gB'(s)e_i = gA(s)B'(s)e_i$. So

$$A^g(s) = gA(s)g^{-1}$$

and

$$\frac{d}{ds}\Big[A^g(s)\Big]_{s=0} = \mathrm{Ad}(g)\omega(X).$$

□

79.25.5 Clifford algebra $(1,1)$

We consider the left invariant vector fields

$$\tilde{e}_J(r_0) = \frac{d}{ds}\Big[r_0 e^{-sJ}\Big]_{s=0} = -r_0 J \tag{79.497a}$$

$$\tilde{e}_L(r_0) = \frac{d}{ds}\Big[r_0 e^{-sL}\Big]_{s=0} = -r_0 L. \tag{79.497b}$$

More precisely, we consider the vectors given by action of these matrices on the "base point" $\begin{pmatrix}0\\1\\0\end{pmatrix}$.

Hence

$$\tilde{e}_J(r_0) = -r_0 \begin{pmatrix}0\\0\\1\end{pmatrix}, \quad \tilde{e}_L(r_0) = -r_0 \begin{pmatrix}1\\0\\0\end{pmatrix} \tag{79.498}$$

and

$$g = \begin{pmatrix}1 & \\ & -1\end{pmatrix}.$$

Remark that this metric is constant (it does not depend on r_0) because r_0 is an isometry. For this reason, we now turn our attention to Clifford algebra and spin group for $V = \mathbb{R}^{1,1}$. Following matrices fulfill relation (68.23)

$$\gamma_J = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \gamma_L = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}.$$

The complete Clifford algebra has the following matrices too:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma_{JL} = \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}.$$

The Clifford algebra is nothing else than $GL(2,\mathbb{R})$, the set of all real 2×2 matrices. From definitions, one can check that

$$\begin{aligned} \alpha(J) &= -J & \tau(J) &= J \\ \alpha(L) &= -L & \tau(L) &= L \\ \alpha(JL) &= JL & \tau(JL) &= -JL \\ \alpha(1) &= 1 & \tau(1) &= 1 \end{aligned}$$

Inverse and α of general element in $\mathrm{Cl}(1,1)$ are given by

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \frac{1}{ps-qr}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}, \quad \alpha\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} p & -q \\ -r & s \end{pmatrix}.$$

A general element in $\mathbb{R}^{1,1}$ is $\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$ with $\alpha, \beta \in \mathbb{R}$, so the condition to belongs to $\Gamma(1,1)$ is that

$$\frac{1}{ps-qr}\begin{pmatrix} p & -q \\ -r & s \end{pmatrix}\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$$

belongs to $\mathbb{R}^{1,1}$ for all α and β. It requires, among others, that $qs\beta - rp\alpha = 0$ for all α and β. Hence $qs = rp = 0$, but the alternatives $p = r = 0$ and $q = s = 0$ are ruled out because we want the determinant $ps - qr$ to be non zero. Therefore, $\Gamma(1,1)$ is generated by

$$\Gamma(1,1) \rightsquigarrow \begin{pmatrix} p & 0 \\ 0 & s \end{pmatrix}, \begin{pmatrix} 0 & q \\ r & 0 \end{pmatrix}.$$

The latter belongs to $\mathbb{R}^{1,1}$, so

$$\Gamma^+(1,1) \rightsquigarrow \begin{pmatrix} z+c & 0 \\ 0 & z-c \end{pmatrix} = z\mathbb{1} + c\gamma_{JL}.$$

From

$$\tau(z\mathbb{1} + c\gamma_{JL}) = z\mathbb{1} - c\gamma_{JL},$$

elements in $\mathrm{Spin}(1,1)$ are subject to the relation

$$\begin{pmatrix} s & 0 \\ 0 & p \end{pmatrix} = \frac{1}{ps}\begin{pmatrix} s & 0 \\ 0 & p \end{pmatrix}.$$

As consequence, we find

$$\mathrm{Spin}(1,1) = \mathbb{R}_0 \rightsquigarrow \begin{pmatrix} 1/p & \\ & p \end{pmatrix}. \tag{79.499}$$

If we put (see decomposition (53.320))

$$A = \begin{pmatrix} e^a & \\ & e^{-a} \end{pmatrix},$$

we have $\mathrm{Spin}(1,1) = A \times \mathbb{Z}_2$. Let us check that $\mathrm{Spin}(1,1)$ is a double covering of $\mathrm{SO}_0(1,1)$. We know that

$$\mathrm{SO}(1,1) = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} \times \mathbb{Z}_2$$

while $\mathrm{SO}_0(1,1)$ is

$$\mathrm{SO}_0(1,1) = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} = \mathbb{R}.$$

This structure of $\mathrm{SO}(1,1)$ comes from the fact (true for all $\mathrm{SO}(1,n)$) that $|\Lambda_0^0| \geqslant 1$ when Λ is a Lorentz transformation. So $\mathbb{1}$ and $-\mathbb{1}$ cannot belong to the same connected component. Note that $\cosh\xi \geqslant 1$. We see intuitively how to cover two times $\mathbb{R}$ with $\mathbb{R}_0$. Let us see how the map χ does that. From definition, $\chi(x)y = \alpha(x)yx^{-1}$, so it is easy to see that

$$\chi(1) = \chi(-1) = \mathrm{Id}\,|_{\mathbb{R}^{1,1}}$$

79.25.6 Parallel transport

We have a connection on the frame bundle of AdS_2 and we wan to lift the vectors $\tilde{e}_J$ and $\tilde{e}_L$, i.e. we consider a point

$$\xi_0 = (r_0, v_1, v_2) \in \mathrm{SO}(AdS_2)$$

where v_1 and v_2 form an orthonormal (in the sense of g) basis of $T_{r_0}AdS_2$. Then we have to find a path $s \to \xi(s)$ in $\mathrm{SO}(AdS_2)$ such that $\xi(0) = \xi_0$, $\omega(\xi'(0)) = 0$ and $dp\xi'(0) = \tilde{e}_a$. The latter condition allows us to compute $r(s)$ in the expression

$$\xi(s) = (r(s), v_1(s), v_2(s)),$$

namely, $r(s)$ is the path of $\tilde{e}_a$. The condition to be horizontal imposes that vectors $v_i(s)$ are parallel transport of v_i along $\tilde{e}_a$. So we have to compute the different $T_a(\tilde{e}_b)(s)$ which is the parallel transported of $\tilde{e}_b$ along the path of $\tilde{e}_a$ at a distance s; this is an element of $T_{\tilde{e}_a(s)}AdS_2$. It will be decomposed in the basis

$$\begin{aligned} \tilde{e}_J\big(\tilde{e}_a(s)\big) &= -r_0 e^{-sa} J \\ \tilde{e}_L\big(\tilde{e}_a(s)\big) &= -r_0 e^{-sJ} L. \end{aligned}$$

where we imply the action on the base point $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. For notational simplicity, from now we write $a(s)$ instead of $\tilde{e}_a(s)$. Various products are easy to compute; for example

$$\tilde{e}_J(J(s)) \cdot \tilde{e}_L(J(s)) = r_0 e^{-sJ} J \cdot r_0 e^{-sJ} L = J \cdot L = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0$$

because $r_0 e^{-sJ}$ is an isometry. In general:

$$\tilde{e}_a\big(c(s)\big) \cdot \tilde{e}_b\big(c(s)\big) = a \cdot b$$

Now we pose in general

$$T_a(\tilde{e}_b)(s) = \alpha(s)\tilde{e}_b\big(a(s)\big) + \beta(s)\tilde{e}_L\big(a(s)\big),$$

and we want to find the (real valued) functions α and β. Parallel transport fulfils two conditions: the norm and the angle with the path are constant. This leads us to two conditions:

$$T_a\big(\tilde{e}_b(s)\big) \cdot T_a\big(\tilde{e}_b(s)\big) = b \cdot b \tag{79.500a}$$

$$T_a\big(\tilde{e}_b(s)\big) \cdot \tilde{e}_a\big(a(s)\big) = b \cdot a. \tag{79.500b}$$

These equations extends to

$$\beta(s)^2 - \alpha(s)^2 = b \cdot b \tag{79.501a}$$

$$\alpha(s) J \cdot a + \beta(s) L \cdot a = b \cdot a. \tag{79.501b}$$

There are four cases to be considered following that $a = J, L$ and $b = J, L$. The result is that

$$T_a\big(\tilde{e}_b\big) = \tilde{e}_b, \tag{79.502}$$

in other terms, the vectors $\tilde{e}_J$ and $\tilde{e}_L$ are not only parallel vector fields, but each is parallel along the path of the other.

79.25.7 Covariant derivative

We will give the horizontal lift of $\tilde{e}_a$ at point

$$\xi(0) = \{B_1^b \tilde{e}_b, B_2^c \tilde{e}_c\}_{r_0 e_t}$$

under the form of the path

$$\xi(s) = \{B_1^b \tilde{e}_b\big(a(s)\big), B_2^c \tilde{e}_c\big(a(s)\big)\}_{\tilde{e}_a(s)}.$$

We create a connection on the spinor bundle from the connexion via the formula

$$\widehat{\nabla_X^E \psi}(\xi) = \overline{X}_\xi(\hat{\psi}).$$

In our case, we take $\psi\colon M \to \mathcal{S}$, or $\hat{\psi}\colon SU(1,1) \to \Lambda W$ such that

$$\hat{\psi}(U \cdot g) = \rho(g^{-1})\hat{\psi}(U).$$

Since $\tilde{\omega} = \varphi^* \omega$, we have $\tilde{\omega}(X) = \omega(d\varphi X)$ and

$$\overline{e}_{a\xi_0} = \varphi^{-1}\big(\tilde{e}_a(s), \dots\big).$$

Therefore

$$\widehat{\nabla_a \psi}(\xi_0) = \frac{d}{ds}\Big[(\hat{\psi} \circ \varphi^{-1})\{B_i^c \tilde{e}_c\big(a(s)\big)\}_{\tilde{e}_a(s)}\Big]_{s=0} \tag{79.503}$$

eq_whidpsinabla

where φ is defined by

$$\varphi(U) = \{U\tilde{e}_J, U\tilde{e}_L\}_{Ur_0 e_t}.$$

We have to find

$$\varphi^{-1}\{B_i^c \tilde{e}_c\big(a(s)\big)\}_{\tilde{e}_a}. \tag{79.504}$$

eq_varpBic

Before to write down the inverse of φ, let us perform some computations.

$$J = \begin{pmatrix} & 0 & \\ 0 & 0 & 1 \\ & 1 & \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 1 & 1 \\ -1 & & \\ 1 & & \end{pmatrix},$$

and as far as we only wants to compute derivatives, we can write the exponentials as

$$e^{sJ} = \mathbb{1} + sJ = \begin{pmatrix} 1 & & \\ & 1 & s \\ & s & 1 \end{pmatrix} \tag{79.505}$$

$$e^{sL} = \mathbb{1} + sL = \begin{pmatrix} 1 & s & s \\ -s & 1 & 0 \\ s & 0 & 1 \end{pmatrix}. \tag{79.506}$$

The path are given by

$$\tilde{e}_a(s) = r_0 e^{-sa} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{79.507}$$

in particular

$$\tilde{e}_J(s) = r_0 \begin{pmatrix} 0 \\ 1 \\ -s \end{pmatrix}, \qquad \tilde{e}_L(s) = r_0 \begin{pmatrix} -s \\ 1 \\ 0 \end{pmatrix}. \tag{79.508}$$

For the various $\tilde{e}_b\big(a(s)\big)$, we have

$$\tilde{e}_b\big(a(s)\big) = \frac{d}{dt}\Big[a(s)e^{-tb}\Big]_{t=0} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{d}{dt}\Big[r_0 e^{-sa} e^{-tb}\Big]_{t=0} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = -r_0 e^{-sa} b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{79.509}$$

Results are

$$\tilde{e}_J\big(J(s)\big) = -r_0 \begin{pmatrix} 0 \\ -s \\ 1 \end{pmatrix} \qquad \tilde{e}_J\big(L(s)\big) = -r_0 \begin{pmatrix} -s \\ 0 \\ 1 \end{pmatrix} \tag{79.510a}$$

$$\tilde{e}_L\big(J(s)\big) = -r_0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \tilde{e}_L\big(L(s)\big) = -r_0 \begin{pmatrix} 1 \\ s \\ -s \end{pmatrix}. \tag{79.510b}$$

We finally have to know that

$$B^c\tilde{e}_c\big(J(s)\big) = -r_0 \begin{pmatrix} B^L \\ -sB^J \\ B^J \end{pmatrix}, \quad B^c\tilde{e}_c\big(L(s)\big) = -r_0 \begin{pmatrix} -sB^J + B^L \\ sB^L \\ B^J - sB^L \end{pmatrix}.$$

Following equation (79.504), in order to write down $\widehat{\nabla_a\psi}$, we have to find $U(s) \in SU(1,1)$ such that

(1) $Ur_0e_t = \tilde{e}_a(s)$,

(2) $U\tilde{e}_J = B_1^c\tilde{e}_c\big(a(s)\big)$,

(3) $U\tilde{e}_L = B_2^c\tilde{e}_c\big(a(s)\big)$.

If $\overline{f}$ and $\overline{g}$ are vectors, solutions in U of equation $Ur_0\overline{f} = r_0\overline{g}$ are $U = \mathbf{Ad}(r_0)B$ where B fulfils $B\overline{f} = \overline{g}$. In the case of $a = J$, the three conditions successively give

$$U = \mathbf{Ad}(r_0) \begin{pmatrix} . & 0 & . \\ . & 1 & . \\ . & -s & . \end{pmatrix} \tag{79.511a}$$

$$U = \mathbf{Ad}(r_0) \begin{pmatrix} . & . & B_1^L \\ . & . & -sB_1^J \\ . & . & B_1^J \end{pmatrix} \tag{79.511b}$$

$$U = \mathbf{Ad}(r_0) \begin{pmatrix} -B_2^L & . & . \\ sB_2^J & . & . \\ B_2^J & . & . \end{pmatrix}. \tag{79.511c}$$

Putting all together in equation (79.503) we find

$$\begin{aligned} \widehat{\nabla_J\psi}(\xi_0) &= \frac{d}{ds}\hat{\psi}\,\mathbf{Ad}(r_0) \begin{pmatrix} -B_2^L & 0 & B_1^L \\ sB_2^J & 1 & -sB_1^J \\ B_2^J & -s & B_1^J \end{pmatrix} \\ &= d\hat{\psi}\,\mathbf{Ad}(r_0) \begin{pmatrix} 0 & 0 & 0 \\ B_2^J & 0 & -B_1^J \\ 0 & -1 & 0 \end{pmatrix}. \end{aligned} \tag{79.512}$$

eq_nablJmoi

The same with L instead of J leads to

$$\widehat{\nabla_L\psi}(\xi_0) = d\hat{\psi}\,\mathbf{Ad}(r_0) \begin{pmatrix} -B_2^J & -1 & B_1^J \\ B_2^L & 0 & B_1^L \\ -B_2^L & 0 & -B_1^L \end{pmatrix}. \tag{79.513}$$

eq_nablLmoi

However it should be shocking to get 3×3 matrices in $SU(1,1)$: we had abused between $\mathrm{SO}(2,1)$ and $SU(1,1)$.

79.25.8 Another way to write a section (wrong way to do)

The equivariant function $\hat{\psi}\colon SU(1,1) \to \Lambda W$ fulfills

$$\hat{\psi}(U \cdot g) = \rho(g^{-1})\hat{\psi}(U)$$

for all $g \in \mathrm{Spin}(1,1)$; in particular with $g = -\mathbb{1}$,

$$\hat{\psi}(-U) = -\hat{\psi}(U). \tag{79.514}$$

This gives the idea that it is not impossible to define $\hat{\psi}$ from its projection on $\mathrm{SO}(2,1)$: we want to get $\tilde{\psi}\colon \mathrm{SO}(2,1) \to \Lambda W$ and define

$$\hat{\psi}(U) = \tilde{\psi}\big(T(U)\big).$$

More precisely, we parametrize $SU(1,1)$ by α and β such that $|\alpha|^2 - |\beta|^2 = 1$. Then we divide $SU(1,1)$ into two parts: $\alpha = x + iy$ is green when $x > 0$ and when $x = 0$, $y < 0$; α is red when $x < 0$ and when $x = 0$, $y > 0$. When $\alpha = 0$, we classify following β in the same way. The result is that U is green if and only if $-U$ is red. For a map $\tilde{\psi}\colon \mathrm{SO}(2,1) \to \Lambda W$, we define

$$\hat{\psi}(U) = \begin{cases} \tilde{\psi}\big(T(U)\big) & \text{if } U \text{ is green} \\ -\tilde{\psi}\big(T(U)\big) & \text{if } U \text{ is red} \end{cases} \tag{79.515}$$

We define $T^{-1}\colon \mathrm{SO}(2,1) \to SU(1,1)$ as follows: $T^{-1}(A)$ is the green element of $SU(1,1)$ whose image by T is A. In any cases we have

$$\hat{\psi} \circ T^{-1} = \tilde{\psi}.$$

The meaning of equations (79.512) and (79.513) is that $\mathbf{Ad}(r_0)$ is a matrix whose inverse image by T should be given to $\hat{\psi}$; the difficulty is to know which of the two. When U_0 is green,

$$\begin{aligned} \widehat{\nabla_a \psi}(U_0) &= \frac{d}{ds}\Big[(\hat{\psi} \circ T^{-1})\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0} \\ &= \frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0}, \end{aligned}$$

while when U_0 is red,

$$\widehat{\nabla_a \psi}(U_0) = -\frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0}.$$

These two show that

$$\widetilde{\nabla_a \psi}\big(T(U_0)\big) = \frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0} \tag{79.516}$$

All this is only proved in the interior of the green and red regions so that the path $U(s)$ keeps on only one region.

79.25.9 Once again

pg_DiracADsdeux

We see AdS_2 as[22] $\mathcal{O} = \mathrm{Ad}(G)H$ and we consider a base point $o = \mathrm{Ad}(k_0)H$ with $G = SL(2,\mathbb{R}) = ANK$. Let the principal bundle

$$\begin{array}{ccc} A & \overset{R}{\rightsquigarrow} & G \\ & & \downarrow \pi \\ & & \mathcal{O} \end{array}$$

with A acting on G by $(a, g) \mapsto ga$ and the projection

$$\pi(rk_0a) = \mathrm{Ad}(rk_0a)H. \tag{79.517}$$

22. Here, $G = SL(2,\mathbb{R})$

where $r \in R$ and $a \in A$. More precisely, the principal bundle we look at is

$$\begin{array}{ccc} A & \overset{R}{\rightsquigarrow} & \mathcal{U}_G \\ & & \downarrow \pi \\ & & \mathcal{U} \end{array} \tag{79.518}$$

where $\mathcal{U}_G = Rk_0A$ and $\mathcal{U} = \pi(\mathcal{U}_G) = \mathrm{Ad}(Rk_0A)H = \mathrm{Ad}(Rk_0)H = \mathrm{Ad}(R)o$. The $\mathcal{U}_G$ is so defined in order to be the π^{-1} of an orbit $\mathcal{U} = \mathrm{Ad}(R)o$.

We have a manifold isomorphism $R \simeq \mathcal{U}$ given by

$$\phi\colon r \to \mathrm{Ad}(r)o.$$

How to see a left invariant vector field on R *via* this identification?

$$d\phi\tilde{X}_r = d\phi\frac{d}{dt}\Big[re^{tX}\Big]_{t=0} = \frac{d}{dt}\Big[\,\mathrm{Ad}(r)\,\mathrm{Ad}(e^{tX})o\Big]_{t=0}.$$

This leads us to consider the following field for $X \in \mathcal{R}$. We define $\xi_X(rk_0a) \in T_{rk_0a}\mathcal{U}_G$,

$$\xi_X(rk_0a) = \frac{d}{dt}\Big[re^{tX}k_0a\Big]_{t=0}. \tag{79.519}$$

Let's see the projection:

$$\begin{aligned} d\pi_{rk_0a}\xi_X(rk_0a) &= \frac{d}{dt}\Big[\pi(re^{tX}k_0a)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\,\mathrm{Ad}(re^{tX}k_0a)H\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\,\mathrm{Ad}(re^{tX})o\Big]_{t=0}. \end{aligned}$$

This gives us the idea to define $X^\sharp \in T_{\mathrm{Ad}(rk_0a)H}\mathcal{U} = T_{\pi(rk_0a)}\mathcal{U}$ by

$$X^\sharp_{rk_0a} = \frac{d}{dt}\Big[\,\mathrm{Ad}(re^{tX})o\Big]_{t=0}, \tag{79.520}$$

which is a good definition because $\pi(rk_0a) = \pi(r'k_0a')$ only when $r = r'$. We put the following connection on $\mathcal{U}_G$:

$$\alpha_{rk_0a}(\Sigma) = \Big[\big(dL^{-1}_{rk_0a}\big)_{rk_0a}\Sigma\Big]_{\mathcal{A}}. \tag{79.521}$$

We hope ξ_X to be the horizontal lift [23] of $X^\sharp$; by construction $d\pi\xi_X = X^\sharp$. We have

$$\begin{aligned} \alpha_{rk_0a}(\xi_X) &= \big[dL_{(rk_0a)^{-1}}\xi_X(rk_0a)\big]_{\mathcal{A}} \\ &= \frac{d}{dt}\Big[a^{-1}k_0^{-1}r^{-1}re^{tX}k_0a\Big]^{\mathcal{A}}_{t=0} \\ &= \frac{d}{dt}\Big[a^{-1}\,\mathbf{Ad}(k_0^{-1})e^{tX}a\Big]^{\mathcal{A}}_{t=0} \\ &= \big[\mathrm{Ad}(a^{-1}k_0^{-1})X\big]_{\mathcal{A}}. \end{aligned}$$

One can, by brute force computation [24], show that the difference $\mathrm{Ad}(ak_0)X - \mathrm{Ad}(k_0)X$ is skew-diagonal when

$$X = \begin{pmatrix} a' & n \\ 0 & -a' \end{pmatrix}, \quad k_0 = \begin{pmatrix} \cos a & \sin a \\ \sin a & \cos a \end{pmatrix}, \quad a = \begin{pmatrix} a & 0 \\ 0 & 1/p \end{pmatrix}.$$

So $\mathrm{Ad}(a)$ does not change the $\mathcal{A}$-component of $\mathrm{Ad}(k_0)X$. We conclude that

$$\alpha(\xi_X) = \big[\mathrm{Ad}(k_0^{-1})X\big]_{\mathcal{A}}. \tag{79.522}$$

23. We will see in proposition 79.114 that it is not the case, but for the moment, we hope it.
24. Or by remarking that $\mathcal{A}$ is abelian.

When $X \in \mathcal{R}$, we consider $\tilde{X}_g = (dL_g)_e X$;

eq_defXtilde

$$\tilde{X}_{rk_0a} = dL_{rk_0a} X, \tag{79.523}$$

in particular, $\tilde{X}_r = \frac{d}{dt}\Big[re^{tX}\Big]_{t=0}$. We denote by τ the action

$$\begin{aligned} \tau_g &: \mathcal{O} \to \mathcal{O} \\ \tau_g \operatorname{Ad}(r)H &= \operatorname{Ad}(gr)H \end{aligned} \tag{79.524}$$

In particular

$$\begin{aligned} d\pi_g dL_g Y &= \frac{d}{dt}\Big[\pi(ge^{tY})\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\operatorname{ad}(ge^{tY}H)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\tau_g \operatorname{Ad}(e^{tY})H\Big]_{t=0} \\ &= (d\tau_g)_H d\pi_e Y, \end{aligned}$$

thus

$$d\pi \circ dL = d\tau \circ d\pi. \tag{79.525}$$

With definition (79.523), we have $\alpha(\tilde{X}) = X_{\mathcal{A}}$ because

$$\alpha\big(dL_{rk_0a}X\big) = \big[dL_{(rk_0a)^{-1}} dL_{rk_0a} X\big]_{\mathcal{A}} = X_{\mathcal{A}}.$$

We are now able to find some horizontal lift.

prop_horliftXdiz

Proposition 79.114.
The horizontal lift of $X^\sharp$ is

$$\overline{X^\sharp} = \xi_X - [\widetilde{\operatorname{Ad}(k_0^{-1})X}]_{\mathcal{A}}.$$

Proof. First, we have

$$\begin{aligned} d\pi \overline{X^\sharp}|_{rk_0a} &= d\pi \xi_X - d\pi (dL_{rk_0a})_e [\operatorname{Ad}(k_0^{-1})X]_{\mathcal{A}} \\ &= \frac{d}{dt}\Big[\operatorname{Ad}(re^{tX}o)\Big]_{t=0} - d\tau\, d\pi [\operatorname{Ad}(k_0^{-1})X]_{\mathcal{A}}. \end{aligned}$$

The first term is $X^\sharp$ while the second is zero because if $A \in \mathcal{A}$,

$$\begin{aligned} d\pi A &= \frac{d}{dt}\Big[\pi(e^{tA})\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\operatorname{Ad}(e^{tA})H\Big]_{t=0} \\ &= 0. \end{aligned}$$

On the other hand,

$$\alpha(\overline{X^\sharp}) = \alpha(\xi_X) - [\operatorname{Ad}(k_0)^{-1}X]_{\mathcal{A}} = 0.$$

□

Now we prove that the function $\overline{X^\sharp} \cdot \hat{\psi}$ is equivariant, and therefore that the definition

$$\widehat{\nabla_{X^\sharp}\psi} = \overline{X^\sharp} \cdot \hat{\psi}$$

works. Using equivariance of $\hat{\psi}$,

$$\begin{aligned}
\overline{X^\sharp} \cdot \hat{\psi}(ga_1) &= \frac{d}{dt}\Big[\hat{\psi}(\xi_X(t)\Big]_{t=0} - \hat{\psi}\big(dL_{ga_1}[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big) \\
&= \frac{d}{dt}\Big[\hat{\psi}\big(re^{tX}k_0aa_1\big)\Big]_{t=0} - \frac{d}{dt}\Big[\hat{\psi}\big(ga_1e^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[\rho(a_1)\hat{\psi}\big(re^{tX}k_0a\big)\Big]_{t=0} - \frac{d}{dt}\Big[\rho(a_1)\hat{\psi}\big(ge^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0}, \qquad (79.526) \\
&= \rho(a_1)\hat{\psi}(\xi_X) - \frac{d}{dt}\Big[\rho(a_1)\hat{\psi}\big([\widetilde{\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}|_g\big)\Big]_{t=0} \\
&= \rho(a_1)\hat{\psi}(\xi_X) - \rho(a_1)[\widetilde{\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\hat{\psi}(g) \\
&= \rho(a_1)(\overline{X^\sharp} \cdot \hat{\psi})(g).
\end{aligned}$$

for the third line, we used the fact that $\mathcal{A}$ is abelian

79.25.9.1 Clifford algebra for AdS_2

Our basis of $\mathcal{A} \oplus \mathcal{N}$ is

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and we choose

$$o = \mathrm{Ad}(k_0)H = \cos(2k_0)H + \sin(2k_0)(E+F).$$

Since (at first order in t) $\mathrm{Ad}(e^{tH})o = \cos(2k_0)\mathbb{1} + \sin(2k_0)\big(E + 2tE + F - 2tF\big)$,

$$H^\sharp_{rk_0a} = 2\sin(2k_0)\,\mathrm{Ad}(r)(E-F),$$

and

$$E^\sharp_{rk_0a} = \mathrm{Ad}(r)\big(-2\cos(2k_0)E + \sin(2k_0)H\big).$$

We have to compute the metric matrix for this basis; we know from equation (52.11), the Killing form is Ad-invariant and $(\mathfrak{sl}(2,\mathbb{R}), B) \simeq (\mathbb{R}^3, \eta_{21})$. So the $\mathrm{Ad}(r)$ disappears in the computation of $B(X^\sharp, Y^\sharp)$. We get

$$\begin{aligned}
B(H^\sharp, H^\sharp) &= 4\sin^2(2k_0)B(E-F, E-F) \\
&= -32\sin^2(2k_0) \\
B(E^\sharp, E^\sharp) &= \sin^2(2k_0)B(H,H) \\
&= 8\sin^2(2k_0) \\
B(E^\sharp, H^\sharp) &= -4\sin(2k_0)\cos(2k_0)B(E, E-F) + 2\sin^2(2k_0)B(H, E-F) \\
&= 16\sin^2(2k_0)\cos(2k_0).
\end{aligned}$$

So the metric is in the basis $\{H^\sharp, E^\sharp\}$

$$g = \begin{pmatrix} -32\sin^2(2k_0) & 16\sin(2k_0)\cos(2k_0) \\ 16\sin(2k_0)\cos(2k_0) & 8\sin^2(2k_0) \end{pmatrix}. \qquad (79.527)$$

When we consider the orbit of $E+F$, we choose $o = E+F$, i.e. $\cos(2k_0) = 0$, $\sin(2k_0) = 1$ so that

$$H^\sharp_{rk_0a} = 2\,\mathrm{Ad}(r)(E-F), \quad E^\sharp_{rk_0a} = \mathrm{Ad}(r)H, \qquad (79.528)$$

and

$$g = \begin{pmatrix} -32 & 0 \\ 0 & 8 \end{pmatrix};$$

in the case of the orbit of $-(E+F)$, we get the same. The negative vector is $H^\sharp$ and the positive one is $E^\sharp$.

79.25.9.2 Identification $\mathcal{Q} \leftrightarrow \Lambda W$

We want a linear bijection $\phi\colon \mathcal{Q} \to \Lambda W$ such that

$$\rho(s)\phi(X) = \phi\big(\rho(s)X\big)$$

where the left hand side action of Spin is the usual on ΛW while the right hand side one remains to be defined. The implementation of this is easy: we can take any bijection between $\mathcal{Q}$ and ΛW and define

$$\rho(s)X = \phi^{-1}\big(\rho(s)\phi(X)\big). \tag{79.529}$$

Spinors on AdS_2 are given by equivariant functions $\hat{\psi}\colon \mathcal{U}_G \to \Lambda W$ which are now replaced by $\tilde{\psi}\colon R \to \mathcal{Q} \simeq \Lambda W$ by

$$\hat{\psi}(rk_0a) = \rho(a^{-1})\tilde{\psi}(r).$$

So the set of sections of the spinor bundle over $\mathcal{U}$ is

$$\Gamma_{\mathcal{U}} \simeq C^\infty(R, \Lambda W).$$

79.25.9.3 Covariant derivative

The aim is now to compute

$$\begin{aligned}
\widetilde{\nabla_{X^\sharp}\psi}(r) &= \widehat{\nabla_{X^\sharp}\psi}(rk_0) \\
&= \overline{X^\sharp} \cdot \hat{\psi}|_{rk_0} \\
&= \big(\xi_X - [\widetilde{\mathrm{Ad}(k_0^{-1})X}]_{\mathcal{A}}\big) \cdot \hat{\psi}|_{rk_0} \\
&= \frac{d}{dt}\Big[\tilde{\psi}(re^{tX})\Big]_{t=0} - \frac{d}{dt}\Big[\rho\big(e^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0}\tilde{\psi}(r) \\
&= \tilde{X}_r\tilde{\psi}(r) - d\rho_e\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big)\tilde{\psi}(r).
\end{aligned}$$

Our final formula for the covariant derivative is

$$\widetilde{\nabla_{X^\sharp}\psi}(r) = \tilde{X}_r\tilde{\psi} - d\rho\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big)\tilde{\psi}. \tag{79.530}$$

The Dirac operator will be a linear combination of vectors of the form

$$\tilde{X} + d\rho\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big).$$

Notice that $\tilde{X}$ is left invariant and the second term is even independent of the point, so the whole is left invariant.

79.26 Dirac operator on AdS_3

PgDiracAdSTrois

The definition is

$$AdS_3 = \frac{\mathrm{SO}(2,2)}{\mathrm{SO}(1,2)},$$

and the group which acts is the AN of $\mathrm{SO}(2,2)$. The Lie algebra is given by

$$\begin{aligned}
\mathcal{A} &= \{J_1, J_2\} \\
\mathcal{N} &= \{M, L\}
\end{aligned}$$

which has dimension 4. So there is a stabiliser. One can prove that for the open orbit of $u = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, the stabiliser is $\{e^{aJ_2}\}$, i.e.

$$[e^{aJ_2}u] = [u]. \tag{79.531}$$

For the spin group, we find

$$\mathrm{Spin}(2,1) \simeq SL_2^*(\mathbb{R}),$$

the group of 2×2 matrices with determinant equals to ± 1 (cf [967]). Let us recall that the isomorphism $AdS_3 \simeq SL(2,\mathbb{R})$ is given by

$$SL(2,\mathbb{R}) = \begin{pmatrix} t+x & y-u \\ y+u & t-x \end{pmatrix}$$

with $u^2 + t^2 - x^2 - y^2 = 1$. For sake of simplicity, we denote $SL(2,\mathbb{R})$ by G. It is explained in [1018] that the map

$$\begin{aligned} \psi\colon (G \times G) \times AdS_3 &\to AdS_3 \\ (g_1, g_2)x &= g_1 x g_2^{-1} \end{aligned} \tag{79.532}$$

provides a local isomorphism $G \times G \simeq O(2,2)$. Moreover we have locally:

$$\frac{G \times G}{\mathbb{Z}_2} \simeq \mathrm{SO}(2,2).$$

At each point $x \in AdS_3$, we have an isomorphism

$$\mathrm{SO}(2,2)_x \simeq \mathrm{SO}(2,1)$$

where $\mathrm{SO}(2,2)_x$ is the stabiliser of x in $\mathrm{SO}(2,2)$. So we define the isomorphism

$$\chi_x\colon \mathrm{Spin}(2,1) \to \mathrm{SO}(2,2)_x$$

which is a double covering. If $d\psi\colon \mathcal{G} \oplus \mathcal{G} \to \mathfrak{so}(2,2)$ is the isomorphism of [1018], we define $\psi\colon G \times G \to \mathrm{SO}(2,2)$ by

$$\psi(e^X) = e^{d\psi X},$$

which is a good definition because the exponential is surjective on $G \times G$. For each $x \in AdS_3$, we consider the isomorphism

$$\phi_x\colon \mathrm{SO}(2,1) \to \mathrm{SO}(2,2)_x$$

such that $\phi_x\big(\mathrm{SO}(2,1)\big) = \mathrm{SO}(2,2)_x$.

We define $\chi(s)_i\colon \mathrm{Spin}(2,1) \to \mathrm{SO}(2,2)$ by

$$\chi(s) = \chi(s)_1 v \chi(s)_2.$$

The choice of $\chi(s)_i$ is not unique. So we define the action of $\mathrm{Spin}(2,1)$ on $G \times G$ by

$$(g,h) \cdot s = \big(\chi(s)_1 g, \chi(s)_2^{-1} h\big). \tag{79.533}$$

Therefore we have

$$\begin{aligned} \psi\big((g,h) \cdot s\big)x &= \chi(s)_1 g x h^{-1} \chi(s)_2 \\ &= \chi(s)\big(g x h^{-1}\big) \\ &= \chi(s)\big(\psi(g,h)x\big). \end{aligned}$$

79.26.1 Spin structure on AdS_3

From previous considerations, the first choice should be

$$P = \frac{AN}{S} \times \mathrm{Spin}(2,1),$$

but it is easy to remark that $\mathcal{R}' = \{J_1, M, L\}$ is a Lie algebra. So we use the corresponding Lie group R instead of the homogeneous space AN/S (these two are isomorphic). Thus the choice is

$$P = R' \times \mathrm{Spin}(2,1), \tag{79.534}$$

with the projection $\pi\colon P \to AdS_3$,

$$\pi\big(r', s\big) = \left[r' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]$$

We consider

$$\begin{aligned} \theta\colon R' &\to \mathcal{U} = Ro \\ r' &\mapsto ro = \left[r' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]. \end{aligned} \tag{79.535}$$

The projection $\pi\colon P \to \mathcal{U}$ reads $\pi = \theta \circ \mathtt{proj}_1$,

$$\pi\big(r', s\big) = [ro].$$

This definition works because for all a, there exists a $h \in H$ such that

$$e^{aJ_2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} h,$$

from construction of S. Then we look at the following:

$$\begin{array}{ccccccc} \mathrm{Spin}(2,1) & \rightsquigarrow & R' \times \mathrm{Spin}(2,1) & \xrightarrow{\varphi} & \mathrm{SO}\left(\mathcal{U}\right) & \rightsquigarrow & \mathrm{SO}(2,1) \\ & & & \searrow^{\pi} \quad \swarrow & & & \\ & & & \mathcal{U} & & & \end{array}$$

The action of $\mathrm{Spin}(2,1)$ on P is

$$\big(r', s'\big) \cdot s = \big(r', s's\big).$$

In order to define φ, we consider the isomorphism

$$\phi_x\colon \mathrm{SO}(2,1) \to \mathrm{SO}(2,2)_x$$

between $\mathrm{SO}(2,1)$ and the stabiliser of x in $\mathrm{SO}(2,2)$. This extends to an automorphism

$$\phi_x\colon \mathrm{SO}(2,2) \to \mathrm{SO}(2,2),$$

¡¡ Avertissement/question à la lectrice !! 79.115
I'm not sure of that extension, but we do not use it here.

and we define the action of $\mathrm{SO}(2,1)$ on $\mathrm{SO}\left(AdS_3\right)$ by

$$\{b_i\}_x \cdot g = \{\phi_x(g) b_i\}_x. \tag{79.536}$$

Then we define

$$\varphi\big(r', s\big) = \{\phi_{\pi r'}\big(\chi(s)\big) b_i\}_{\pi r'} \tag{79.537}$$

if $\{b_i\}$ is a reference basis at $\pi[r]$. So this construction implies the choice of a section of $\mathrm{SO}\left(AdS_3\right)$. Now, using the fact that both ϕ_x and χ are morphisms, we find

$$\begin{aligned} \varphi\big((r', s) \cdot s'\big) &= \big\{\phi_{\pi r'}\big(\chi(ss')\big)\big\}_{\pi r'} \\ &= \big\{\phi_{\pi r'}\big(\chi(s)\big) b_i\big\} \cdot \chi(s) \\ &= \varphi(r', s) \cdot \chi(s'). \end{aligned} \tag{79.538}$$

This proves that the construction gives a spin structure.

79.26.2 Connection on the spinor bundle

A left invariant vector on $\mathcal{U}$ is of the form

$$\tilde{X}_{xo} = \frac{d}{dt}\Big[xe^{tX}o\Big]_{t=0} = \frac{d}{dt}\Big[\pi\big(xe^{tX},s\big)\Big]_{t=0}$$

for any $s \in \mathrm{Spin}(2,1)$. On AdS_3 (in fact on $\mathcal{U}$) we consider the left invariant vector field

$$X^\sharp_{[x]} = \frac{d}{dt}\Big[[xe^{tX}]\Big]_{t=0} \tag{79.539}$$

which leads us to consider the following field on P:

$$\xi_X\big(r',s\big) = \frac{d}{dt}\Big[r'e^{tX},s\Big]_{t=0} \in T_{(r',s)}P. \tag{79.540}$$

This defines a field which projects to the left invariant field on $\mathcal{U}$:

$$d\pi\xi_X(r',s) = X^\sharp_{r'}. \tag{79.541}$$

Lemma 79.116.
On the general vector

$$\Sigma = \frac{d}{dt}\Big[r'(t),s(t)\Big]_{t=0}, \tag{79.542}$$

the formula

$$\alpha_{(r',s_0)}\Sigma = -\frac{d}{dt}\Big[s_0^{-1}s(t)\Big]_{t=0} \in \mathfrak{spin}(2,1) \tag{79.543}$$

where $s_0 = s(0)$ defines a connection form.

Proof. First let $A \in \mathfrak{spin}(2,1)$ and

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0}.$$

We have

$$\alpha\big(A^*_{(r'),s_0}\big) = \alpha\frac{d}{dt}\Big[(r',s_0)\cdot e^{-tA}\Big]_{t=0} = \alpha\frac{d}{dt}\Big[(r',s_0e^{-tA})\Big]_{t=0} = -\frac{d}{dt}\Big[s_0^{-1}s_0e^{-tA}\Big]_{t=0} = A.$$

Now we take back the vector Σ of equation (79.542), an element $a \in \mathrm{Spin}(2,1)$ and we compute

$$\begin{aligned}
(dR_a\alpha)\Sigma &= \alpha\frac{d}{dt}\Big[\big(r'(t),s(t)\big)\cdot a\Big]_{t=0}\\
&= \alpha\frac{d}{dt}\Big[\big(r'(t),s(t)a\big)\Big]_{t=0}\\
&= -\frac{d}{dt}\Big[a^{-1}s(0)^{-1}s(t)a\Big]_{t=0}\\
&= -\,\mathrm{Ad}(a^{-1})\frac{d}{dt}\Big[s(0)^{-1}s(t)\Big]_{t=0}\\
&= \mathrm{Ad}(a^{-1})\alpha(\Sigma).
\end{aligned}$$

□

Thus that is a connection. This is however not the spin connection. Let β be the Levi-Civita connection on the frame bundle $\mathrm{SO}(AdS_3)$. If

$$\Sigma = \frac{d}{dt}\Big[r'(t),s(t)\Big]_{t=0},$$

we have

$$\beta d\phi\Sigma = \phi_{r'}\big(\chi(s_0)\big)^{-1}\frac{d}{dt}\Big[\phi_{r'(t)}\big(\chi(s_t)\big)\Big]_{t=0}\Big|_{\mathcal{H}}. \tag{79.544}$$

If we note $\phi_{r'(t)}\big(\chi(s_t)\big) = \phi\big(r'(t), \chi(s_f)\big)$, the derivative in (79.544) with respect to t reads

$$\frac{d}{dt}\Big[\phi\big(r'(t), \chi(s_0)\big)\Big]_{t=0} + \frac{d}{dt}\Big[\phi\big(r', \chi(s_t)\big)\Big]_{t=0}. \tag{79.545}$$

The second term of $\beta d\varphi\Sigma$ is

$$\begin{aligned}\frac{d}{dt}\Big[\phi_{r'}(\chi(s_0))^{-1}\phi_{r'}(\chi(s_t))\Big]_{t=0}\Big|_{\mathcal{H}} &= \frac{d}{dt}\Big[\phi_{r'}\big(\chi(s_0^{-1}s_t)\big)\Big]_{t=0}\Big|_{\mathcal{H}} \\ &= d\phi_{r'}d\chi(s_0^{-1}s'(0))\big|_{\mathcal{H}}.\end{aligned}$$

From all that we want to define

$$\alpha^S_{(r',s_0)}\Sigma = d\phi d\chi\big(s_0^{-1}s'(0)\big)\big|_{\mathcal{H}} + \phi_{r'}\big(\chi(s_0)\big)^{-1}\frac{d}{dt}\Big[\phi_{r'(t)}\chi(s_0)\Big]_{t=0}\Big|_{\mathcal{H}}, \tag{79.546}$$

and we would not have $\alpha^S(\xi_X) = 0$.

79.26.3 Horizontal lift

Since the spin component of the path of ξ_X is constant, we have $\alpha(\xi_X) = 0$, so equation (79.541) says that

$$\overline{X^\sharp} = \xi_X. \tag{79.547}$$

Let us recall that an equivariant function (which defined a section of an associated bundle) is

$$\begin{aligned}\hat{\psi}&\colon P \to V \\ \hat{\psi}(\xi \cdot g) &= \rho(g^{-1})\hat{\psi}(\xi).\end{aligned} \tag{79.548}$$

General definition of an equivariant derivative (theorem 56.77) leads to

$$\widehat{\nabla_{X^\sharp}\psi} = \overline{X^\sharp} \cdot \hat{\psi} = \xi_X \cdot \hat{\psi}.$$

In our setting, the equivariance of $\hat{\psi}$ reads, for all $a \in \mathrm{Spin}(2,1)$,

$$\hat{\psi}\big(([r], s) \cdot a\big) = \hat{\psi}\big([r], sa\big) \overset{!}{=} \rho(a^{-1})\hat{\psi}\big([r], s\big).$$

We check the equivariance of $\widehat{\nabla_{X^\sharp}\psi}$ by the following computation:

$$\begin{aligned}\widehat{\nabla_{X^\sharp}\psi}\big(([r], s) \cdot a\big) &= \widehat{\nabla_{X^\sharp}\psi}([r], sa) \\ &= (\xi_X \cdot \hat{\psi})([r], sa) \\ &= \frac{d}{dt}\Big[\hat{\psi}\big([re^{tX}], sa\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\rho(a^{-1})\hat{\psi}\big([re^{tX}], s\big)\Big]_{t=0} \\ &= \rho(a^{-1})(\xi_X \cdot \hat{\psi})\big([r], s\big) \\ &= \rho(a^{-1})\widehat{\nabla_{X^\sharp}\psi}\big([r], s\big).\end{aligned}$$

We define $\tilde{\psi}\colon AN/S \to \Lambda W$ by

$$\tilde{\psi}([r]) = \hat{\psi}([r], e),$$

so that

$$\hat{\psi}([r], s) = \rho(s^{-1})\tilde{\psi}([r]). \tag{79.549}$$

We can conclude

$$\begin{aligned}\widetilde{\nabla_{X^\sharp}\psi}([r]) &= \widehat{\nabla_{X^\sharp}\psi}([r],e)\\ &= \xi_X\hat{\psi}([r],e)\\ &= \frac{d}{dt}\Big[\hat{\psi}\big([re^{tX}],e\big)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\tilde{\psi}\big([re^{tX}]\big)\Big]_{t=0}\\ &= \tilde{X}_{[r]}\tilde{\psi}([r]).\end{aligned}$$

So

$$\widetilde{\nabla_{X^\sharp}\psi} = \tilde{X}_{[r]}\tilde{\psi}. \tag{79.550}$$

79.26.4 Spin structure on AdS_3

79.26.4.1 Spin structure on the whole AdS_3

¡¡ Avertissement/question au lecteur !! 79.117
The following seems to contradict what I find in Michelson-Donaldson

The central fact is that

$$\mathrm{Spin}(2,1) \simeq \Delta \simeq \mathrm{SL}(2,\mathbb{R})$$

where $\Delta = \{(g,g) \text{ tel que } g \in \mathrm{SL}(2,\mathbb{R})\} \subset G_0$. We take as notations: $G_0 = \mathrm{SL}(2,\mathbb{R})$ and $\overline{G} = G_0 \times G_0$.

Lemma 79.118.
We have the following homogeneous space isomorphism:

$$\overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}).$$

Proof. We have an action $\overline{G} \times AdS_3 \to AdS_3$,

$$(g,h)x = gxh^{-1} \tag{79.551}$$

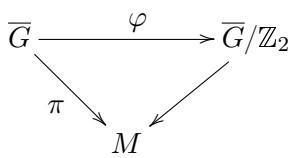
where $x \in \mathrm{SL}(2,\mathbb{R})$ is seen as in AdS_3 by the usual isomorphism. Moreover we consider the isomorphism

$$\overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}) \tag{79.552}$$

$$[x_1,x_2] \mapsto x_1x_2^{-1} \tag{79.553}$$

which is well defined because $[x_1g,x_2g] \mapsto x_1gg^{-1}x_2^{-1} = x_1x_2^{-1}$. In particular, $[g,g] \mapsto e \in \mathrm{SL}(2,\mathbb{R})$. So $\overline{G}$ acts on $\mathrm{SL}(2,\mathbb{R})$ and the elements which fix e are the one of Δ. It proves the lemma. □

We are going to take the following structure:



(79.554)

where M is G_0 seen as $M = \overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}) \simeq AdS_3$, and the projection $\pi\colon \overline{G} \to M$ is given by $\pi(g,h) = gh^{-1}$. The action of $\Delta \simeq \mathrm{Spin}(2,1)$ on $\overline{G}$ is given by formula $(xg,g)\cdot(a,a) = (xga,ga)$. First, let us prove the following.

Proposition 79.119.
The frame bundle over AdS_3 can be seen as

$$\mathrm{SO}(AdS_3) \simeq \overline{G}/\mathbb{Z}_2$$

where $\overline{G} = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$.

Proof. In the fiber bundle $\pi\colon \overline{G} \to M$, the fibre over $x \in \mathrm{SL}(2,\mathbb{R})$ is the set of (g,h) such that $gh^{-1} = x$, or

$$\overline{G}_x = \{(xg, g)\} \subset \overline{G}.$$

We will give a surjective map $\overline{G}_x \to \mathrm{SO}(M)_x$, the fibre of the frame bundle over $x \in AdS_3$. For this, we see a basis of AdS_3 as an isometric map $b\colon \mathcal{G}_0 \to T_xM$ where $\mathcal{G}_0 = \mathfrak{sl}(2,\mathbb{R})$, and we define

$$\begin{aligned} \psi_x &\colon \overline{G}_x \to \mathrm{SO}(M)_x \\ \psi_x(xg, g)(X) &= (dL_x)_e\big(\mathrm{Ad}(g^{-1})X\big) \end{aligned} \tag{79.555}$$

for all $X \in \mathcal{G}_0$. Let us study the kernel of this map, i.e. elements such that $\psi(xg_1, g_1) = \psi(xg_2, g_2)$. It needs, for all $X \in \mathfrak{sl}(2,\mathbb{R})$,

$$\mathrm{Ad}(g_1^{-1})X = \mathrm{Ad}(g_2^{-1})X,$$

but we know that the requirement $\mathrm{Ad}(g)X = X$ is the fact the g is in the center of the group. In our case, it results that $g_2^{-1}g_1 = \pm\,\mathrm{Id}$, so

$$\psi(xg_1, g_1) = \psi(\pm xg_1, \pm g_1)$$

where the same $\pm$ has to be taken in both appearances of the right hand side. Now we put all the ψ_x together to get $\psi\colon \overline{G} \to \mathrm{SO}(M)$. Once again we look in which cases $\psi(g_1,h_1) = \psi(g_2,h_2)$. We put this condition under the form

$$\psi(g_1h_1^{-1}h_1, h_1) = \psi(g_2h_2^{-1}h_2, h_2)$$

which immediately gives $h_1 = \pm h_2$. But on the other hand the base point of $\psi(g_ih_i^{-1}, h_i)$ is $g_ih_i^{-1}$, so that the condition also ask $g_1h_1^{-1} = g_2h_2^{-1}$ which in turn gives $g_1 = \pm g_2$ with the same $\pm$ as in $h_1 = \pm h_2$. We conclude that $\mathbb{Z}_2$ is the problem for the inverse of ψ. This proves the proposition. □

We will usually use the same notation, ψ, to denote the map from $\overline{G}$ and the one from $\overline{G}/\mathbb{Z}_2$. The following lemma will prove useful to study the actions of the structure groups in the picture (79.554).

Lemma 79.120.
The map

$$\begin{aligned} \mathrm{SL}(2,\mathbb{R}) &\to \mathrm{SO}_0(1,2) \\ g &\mapsto \mathrm{Ad}(g). \end{aligned} \tag{79.556}$$

is a double covering.

Proof. No proof. □

The action of $a \in \mathrm{SO}_0(1,2)$ on $(xg, g) \in \overline{G}/\mathbb{Z}_2$ is defined by

$$\psi\big((xg,g)\cdot a\big) = (dL_x)_e\,\mathrm{Ad}(a^{-1}g^{-1}). \tag{79.557}$$

On the other hand, let us see how does $(a,a) \in \Delta \simeq \mathrm{Spin}(2,1)$ acts on $\overline{G}$ and how does it reflects on the ψ level. Since $(xg,g)\cdot(a,a) = (xga, ga)$, we have

$$\psi\big([xg,g]\cdot a\big) = \psi\big((xg,g)\cdot(a,a)\big),$$

and then

$$\varphi\big((xg,g)\cdot a\big) = (xg,g)\cdot(a,a).$$

This proves that our structure is a spin structure.

79.26.4.2 Reduction to one open orbit

We will use this isomorphism between AdS_3 and $\mathrm{SL}(2,\mathbb{R})$:

$$\begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u+x & y-t \\ y+t & u-x \end{pmatrix}.$$

Then the famous point $[u] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in AdS_3$ corresponds to the element $J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$. This is our base point of the open orbit. We could also take

$$k_0 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \in K_0$$

where K_0 is the "K" of $\mathrm{SL}(2,\mathbb{R})$.

¡¡ Avertissement/question au lecteur !! 79.121
I think that J is also a complex structure. To be checked.

We have $J = k_0^2$ and following the action (79.551), we have $J = (k_0, k_0^{-1})e$. The subgroup $\overline{R} \subset \overline{G}$ acts on AdS_3, and we want to know the stabilizer of J. The condition is $(r, r') \cdot J = J$, or

$$r = \mathbf{Ad}(J)r',$$

but $\mathbf{Ad}(J) = \theta$ (the Cartan involution). So an element $(r, r') \in \overline{R}$ stabilises J if it is of the form $(r, \theta r)$, thus

$$\mathfrak{s} = \text{Lie algebra of the stabiliser of } J = \{(X, \theta X) \text{ tel que } X \in \mathcal{R}_0\} \cap \mathcal{R},$$

where the intersection with $\mathcal{R}$ is important because θ can send out of $\mathcal{R}_0$. Note that when X has a $\mathcal{N}$ component, then θX has a $\overline{\mathcal{N}}$ component, so $(X, \theta X) \in (\mathcal{A} \oplus \mathcal{N}, -\mathcal{A} \oplus \overline{\mathcal{N}})$ where the minus sign comes from the fact that $\theta(\mathcal{A}) = -\mathcal{A}$. Then X cannot have a $\mathcal{N}$ component and finally,

$$\mathfrak{s} = \mathbb{R}(H, -H) \in \mathcal{Q}.$$

The group R' is

$$R' = e^{\mathcal{R}'} = \{(an, an') \text{ tel que } n, n' \in N_0\} \tag{79.558}$$

because $\mathcal{R}'$ is $\mathcal{R}$ minus the stabiliser, i.e. $\mathcal{R}' = \mathbb{R}(H, H) \oplus \mathcal{N}$. We have the identification $r' \mapsto r' \cdot J$ between R' and the open orbit $\mathcal{U}$. As usual, the action is $(g, h) \cdot x = gxh^{-1}$ if $r' = (g, h)$. Notice in particular that $R' \neq R'_0 \times R'_0$.

Up to now we studied the fiber $\overline{G} \to M$; we are now able to restrict it to $\overline{G}|_{\mathcal{U}} \to \mathcal{U}$ and to establish an isomorphism with the trivial bundle $R' \times G_0 \to R'$. The fiber over $x \in \mathcal{U}$ is

$$\overline{G}_x = \{(xg, g)\}.$$

We define the isomorphism as follows:

$$\begin{aligned} \tau \colon R' \times G_0 &\to \overline{G}|_{\mathcal{U}} \\ (r', g) &\mapsto (r' \cdot Jg, g) \end{aligned} \tag{79.559}$$

and we have the following picture:

$$\begin{array}{ccc} R' \times G_0 & \xrightarrow{\tau} & \overline{G}|_{\mathcal{U}} \\ \downarrow & & \downarrow \pi \\ R' & \overset{\tau}{\dashrightarrow} & \mathcal{U} \end{array}$$

in which are defines by

$$\begin{array}{ccc} (r',g) & \xrightarrow{\tau} & (r'\cdot Jg,g) \\ \downarrow & & \downarrow \pi \\ r' & \xrightarrow{\tau} & r'\cdot J \end{array}$$

where the dotted line denotes the induced map from τ, which is denoted by the same symbol. The map $\tau\colon R'\to\mathcal{U}$ is just the restriction of the original τ to $g=e$. Notice that this τ provides a diffeomorphism of the basis spaces R' and $\mathcal{U}$.

79.26.4.3 Spin connection

The spin connection on $\overline{G}|_{\mathcal{U}}$ is given by

$$\alpha^S_{(g,h)}\Sigma = \big[dL_{(g,h)^{-1}}\Sigma\big]_{\mathcal{H}}\,, \tag{79.560}$$

EqDefConnAdS3

or

$$\alpha^S_{(g,h)} = \mathtt{proj}_{\mathcal{H}}\circ\big(dL_{(g,h)^{-1}}\big)_{(g,h)}. \tag{79.561}$$

Notice that when we write $\mathcal{H}$, we think about Δ: the group by which quotient $\overline{G}$ in order to get $\mathrm{SL}(2,\mathbb{R})\simeq AdS_3$. Our task now is to transfer this connection to $R'\times G_0$ by defining $\alpha'=\tau^*\alpha^S$. If $\Sigma\in T_{(r',g)}(R'\times G_0)$, we define

$$\alpha'_{(r',g)}\Sigma = \alpha^S(d\tau\Sigma). \tag{79.562}$$

Let us take $X\in\mathcal{G}_0$ and $0\in\mathcal{R}'$ and let us compute $d\tau(0,X)$. More precisely, we consider

$$\begin{aligned} d\tau(0\oplus\tilde{X}_g)_{(r',g)} &= d\tau\frac{d}{dt}\Big[r',ge^{tX}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[r'\cdot Jge^{tX},ge^{tX}\Big]_{t=0} \\ &= \big(\tilde{X}_{(r'\cdot Jg)},\tilde{X}_g\big). \end{aligned}$$

The next step is to compute $d\tau\Sigma$ in the case where $\Sigma=(\underline{Y}\oplus -1)_{(r',g)}$ with $Y\in\mathcal{R}'\subset\mathcal{R}_0\oplus\mathcal{R}_0$. We have

$$d\tau\Sigma = \frac{d}{dt}\Big[\tau\big(e^{tY}r',g\big)\Big]_{t=0} \tag{79.563}$$

$$= \frac{d}{dt}\Big[(e^{tY}r'\cdot J)g,g\Big]_{t=0} \tag{79.564}$$

where, if $r'=(r_1,r_2)$, we consider $Y=\big((\underline{Y}_1)_{r_1},(\underline{Y}_2)_{r_2}\big)$. This appears to be difficult to be computed. This reflects the fact that the connection should be complicated in the trivial bundle $R'\times G_0$.

But there are no fate. We remember that τ furnish a diffeomorphism between the basis spaces, so one can consider the bundle

$$\begin{array}{c} \overline{G}|_{\mathcal{U}} \\ \downarrow \tau^{-1}\circ\pi \\ R' \end{array}$$

Vectors of $\mathcal{H}$ are of the form (X,X) with $X\in\mathfrak{sl}(2,\mathbb{R})$, thus $A\in T_{(xg,g)}\overline{G}|_{\mathcal{U}}$ fulfils $\alpha^S(A)=0$ if and only if

$$dL_{(xg,g)^{-1}}(A) = (X,-X)$$

for a certain $X\in\mathfrak{sl}(2,\mathbb{R})$. All this makes that the horizontal space over (xg,g) is given by

$$\mathrm{hor}(xg,g) = \big\{(\tilde{X}_{xg},-\tilde{X}_g)\text{ tel que }X\in\mathcal{G}_0=\mathfrak{sl}(2,\mathbb{R})\big\}. \tag{79.565}$$

The strategy now is to project that on R' and express Dirac operator in terms of the result. Let us make this simple computation:

$$\begin{aligned} d\pi(\tilde{X}_{xg}, \tilde{X}_g) &= \frac{d}{dt}\Big[\pi\big(xge^{tX}, ge^{-tX}\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[xge^{tX}e^{tX}g^{-1}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[xe^{2t\,\mathrm{Ad}(g)X}\Big]_{t=0} \\ &= 2(dL_x)_e\,\mathrm{Ad}(g)X. \end{aligned}$$

This result has to be brought from $\mathcal{U}$ to R' by τ^{-1}. Now we take a $\tilde{Y} \in \mathfrak{X}(R')$ and we want to know which is the corresponding X, i.e. the $X \in \mathfrak{sl}(2,\mathbb{R})$ such that

$$d\tau^{-1}d\pi(\tilde{X}_{xg}, -\tilde{X}_g) = \tilde{Y}.$$

From the previous computation, $\tilde{Y} = 2d\tau^{-1}dL_x\,\mathrm{Ad}(g)X$, so

$$X = \frac{1}{2}\,\mathrm{Ad}(g^{-1})dL_{x^{-1}}d\tau\tilde{Y}. \tag{79.566}$$

EqXfracAdY

We now precise our idea:

$$\tilde{Y}_{(r_1,r_2)} = \big((\tilde{Y}_1)_{r_1}, (\tilde{Y}_2)_{r_2}\big) = \frac{d}{dt}\Big[r_1e^{tY_1}, r_2e^{tY_2}\Big]_{t=0} \tag{79.567}$$

EqtildeYrunrdeux

for $Y_i \in \mathcal{R}_0'$ and r_1, $r_2 \in R_0$. In this case, the "x" in equation (79.566) is $(r' \cdot J)^{-1}$. Let us begin by taking $s' \in R'$ and compute $L_{(r'\,\cdot\,J)^{-1}\tau}(s')$. Remember that $r' \cdot J = r_1Jr_2^{-1}$ from the general action (79.551), so if $r' = (r_1, r_2)$,

$$\begin{aligned} dL_{(r'\,\cdot\,J)^{-1}\tau}(s') &= (r' \cdot J)^{-1}s_1Js_2^{-1} \\ &= (r_1Jr_2^{-1})^{-1}s_1Js_2^{-1} \\ &= -r_2Jr_2^{-1}s_1Js_2^{-1}. \end{aligned}$$

Now, we apply that result on computation of (79.566) with (79.567):

$$\begin{aligned} dL_{(r'\,\cdot\,J)^{-1}}d\tau\tilde{Y} &= \frac{d}{dt}\Big[-r_2Jr_1^{-1}r_1e^{tY_1}Je^{-tY_2}r_2^{-1}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathbf{Ad}(r_2)\big(-Je^{tY_1}Je^{-tY_2}\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathbf{Ad}(r_2)e^{-tY_2}\Big]_{t=0} + \mathrm{Ad}(r_2)\,\mathrm{Ad}(J)Y_1 \\ &= -\mathrm{Ad}(r_2)Y_2 + \mathrm{Ad}(r_2)\theta(Y_1), \end{aligned}$$

and finally,

$$\begin{aligned} X &= \frac{1}{2}\,\mathrm{Ad}(g^{-1})dL_{(r'\,\cdot\,J)^{-1}}d\tau\tilde{Y} \\ &= \frac{1}{2}\,\mathrm{Ad}(g^{-1})\big[\mathrm{Ad}(r_2)\theta(Y_1) - \mathrm{Ad}(r_2)Y_2\big]. \end{aligned} \tag{79.568}$$

EqValeurXAdtheta

For this X, the horizontal lift of $\tilde{Y} \in \mathfrak{X}(R')$ is $(X, -X) \in T\overline{G}|_{\mathcal{U}}$.

79.26.5 Left invariance of Dirac

Sections of the spin bundle over the open orbit $\mathcal{U}$ are given by equivariant functions $\hat{\psi}\colon \overline{G}|_{\mathcal{U}} \to \mathbb{R}^2$. The action of $\Delta \simeq \mathrm{Spin}(2,1)$ on $\overline{G}$ is

$$(g,h) \cdot (a,a) = (ga, ha).$$

We define $\tilde{\psi}$ by

$$\tilde{\psi}(g) = \hat{\psi}(g, e) \tag{79.569}$$

for $g \in \mathcal{U}$. We get back the original $\hat{\psi}$ by formula

$$\hat{\psi}(g, h) = \rho(h, h)^{-1}\tilde{\psi}(gh^{-1}). \tag{79.570}$$

Our intention is now to compute $\widehat{\nabla_Z \psi}(\xi) = \overline{Z}_\xi(\hat{\psi})$ with $\xi = (xg, g) \in \overline{G}|_{\mathcal{U}}$ (hence $x \in \mathcal{U}$) and $Z \in \mathfrak{X}(R')$. For instance we choose a left invariant $Z = \tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2)$ for $Y_1, Y_2 \in \mathcal{R}'_0$. Recall that $\tilde{Y}$ is given by equation (79.567). From definition of the covariant derivative associated with the connection,

$$\widehat{\nabla_{\tilde{Y}}\psi}(\xi) = \overline{\tilde{Y}}_\xi(\hat{\psi}) = \overline{\tilde{Y}}_{(xg,g)}(\hat{\psi})$$

where $\overline{\tilde{Y}}_{(xg,g)}$ is an horizontal vector at (xg, g) whose projection is $\tilde{Y}$. From our previous work,

$$\overline{\tilde{Y}}_{xg,g} = (\tilde{X}_{xg}, -\tilde{X}_g)$$

with $X = \frac{1}{2}\operatorname{Ad}(g^{-1})\big(\operatorname{Ad}(r_2)\theta Y_1 - \operatorname{Ad}(r_2)Y_2\big)$. Let us understand the link between (r_1, r_2) and g, x. The vector $(\tilde{X}_{xg}, \tilde{X}_g)$ actually projects to a vector at $\tau^{-1} \circ \pi(xg, g) = \tau^{-1}(x)$. The fact that $x \in \mathcal{U}$ guarantees existence and uniqueness of $(r_1, r_2) \in R'$ such that $r_1 J r_2^{-1} = x$. We have

$$\begin{aligned}
\widehat{\nabla_{\tilde{Y}}\psi}(x) &= \widehat{\nabla_{\tilde{Y}}\psi}(x, e) \\
&= (\tilde{X}_x, -\tilde{X}_e)\hat{\psi} \\
&= \frac{d}{dt}\hat{\psi}\Big(xe^{tX}, e^{-tX}\Big)\Big|_{t=0} \\
&= \frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(xe^{2tX})\Big|_{t=0}.
\end{aligned}$$

The first term of the derivation (the one with $t = 0$ in the ρ) gives $2\tilde{X}_x\tilde{\psi}$. This is left invariant. The second is

$$\frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(x)\Big|_{t=0}.$$

We want to test the condition (62.91) on this term. Let us pose

$$(E\tilde{\psi})(x) = (\tilde{X}_x, \tilde{X}_e)\hat{\psi} = \frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(x)\Big|_{t=0}$$

with X given by equation (79.568). On the one hand,

$$L_y(E\tilde{\psi})(x) = (E\tilde{\psi})(yx) = \frac{d}{dt}\rho(e^{tX_a}, e^{tX_a})\tilde{\psi}(yx)\Big|_{t=0} \tag{79.571a}$$

with

$$X_a = \frac{1}{2}\Big(\operatorname{Ad}(r_2)\theta Y_1 - \operatorname{Ad}(r_2)\Big)Y_2 \tag{79.571b}$$

where (r_1, r_2) is given by yx. On the other hand,

$$E(L_y\tilde{\psi})(x) = \frac{d}{dt}\rho(e^{tX_b}, e^{tX_b})\tilde{\psi}(yx)\Big|_{t=0} \tag{79.572a}$$

with

$$X_b = \frac{1}{2}\big(\operatorname{Ad}(s_2)\theta Y_1 - \operatorname{Ad}(s_2)Y_2\big) \tag{79.572b}$$

where (s_1, s_2) is given by x.

The problem is that the choice of y is arbitrary, so that X_a and X_b could be too different. Ok. That's the proof that Dirac is not invariant. Here is the proof that Dirac is invariant.

Following equation (79.560), the spin connection form is

$$\alpha^S_{(g,h)}\Sigma = \big(dL_{(g,h)^{-1}}\Sigma\big)_{\mathcal{H}}\,.$$

If $L_{(x,y)}$ is the left translation by (x,y) we have

$$\big(L^*_{(x,y)}\alpha\big)_{(g,h)}\Sigma = \alpha_{(xg,yh)}\big(dL_{(x,y)}\Sigma\big) = \big(dL_{(g,h)^{-1}}\Sigma\big)_{\mathcal{H}}\,.$$

Thus we have $L^*_{(x,y)}\alpha^S = \alpha^S$. Now we consider the formula $\widehat{\nabla_X\psi}(\xi) = (\tilde{X}_{xg}, -\tilde{X}_g)\hat{\psi}$, and we will check that

$$\big(L_\eta\widehat{\nabla_Z\psi}\big)(\xi) = \nabla_{\widehat{Z(L_\eta\psi)}}(\xi). \tag{79.573}$$

with $\xi = (xg, g)$ and $\eta = (a, b)$. On the one hand,

$$\begin{aligned}\big(L_{(a,b)}\widehat{\nabla_Z\psi}\big)(xg,g) &= \widehat{\nabla_Z\psi}(axg,bg)\\ &= \widehat{\nabla_Z\psi}(axgg^{-1}b^{-1}bg,bg)\\ &= \big(\tilde{X}_{(axb^{-1})bg}, -\tilde{X}_{bg}\big)\hat{\psi}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\nabla_{\widehat{Z(L_{(a,b)}\psi)}}(xg,g) &= (\tilde{X}_{xg}, -\tilde{X}_g)\widehat{L_{(a,b)}\psi}\\ &= \frac{d}{dt}\widehat{L_{(a,b)}\psi}\big(xge^{tX}, ge^{-tX}\big)\Big|_{t=0}\\ &= \frac{d}{dt}\hat{\psi}\big(axge^{tX}, bge^{-tX}\big)\Big|_{t=0}\\ &= (\tilde{X}_{axg}, -\tilde{X}_{bg})\hat{\psi}\\ &= (\tilde{X}_{(axb^{-1})bg}, -\tilde{X}_{bg})\hat{\psi}.\end{aligned}$$

79.27 Spin structure and Dirac operator on AdS_l

SecDirADs

Construction of the frame bundle is a straightforward adaptation of theorem 2.2 (chapter II) in [1033], while connection issues are adapted from proposition 1.3 (chapter III). According proposition 79.5, notations G and H stand for the identity components of $\mathrm{SO}(2, l-1)$ and $\mathrm{SO}(1, l-1)$.

79.27.1 Frame bundle and spin structure

An element of the frame bundle is a map from $\mathcal{Q}$ to $T(G/H)$ of the form[25] $d\mu_g \circ A$ where $g \in G$ and $A \in \mathrm{SO}_0(\mathcal{Q})$. By proposition 79.18, there exists a $h \in H$ for which $A = \mathrm{Ad}(h)$ for every $A \in \mathrm{SO}_0(\mathcal{Q})$ so we have

$$d\mu_g \circ A = d\pi \circ dL_g \circ \mathrm{Ad}(h) = d\pi \circ dL_g \circ dL_h \circ dR_h = d\pi \circ dL_g \circ dL_h = d\mu_{gh}$$

hence in fact every element in the frame bundle reads $d\mu_g$ for some $g \in G$. We conclude that the fibre $B_{[g]}$ over $[g]$ is made of maps of the form $d\tau_k$ with $k \in [g]$. The action of H on the frame bundle is given by

$$(d\mu_g) \cdot h = d\mu_g \circ \mathrm{Ad}(h).$$

Proposition 79.122.
The map

$$\begin{aligned}\beta\colon G &\to B\\ g &\mapsto d\mu_g\end{aligned} \tag{79.574}$$

25. See 66.2 for notations.

is a principal bundle isomorphism between the frame bundle and the principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & H \\ \downarrow{\scriptstyle \pi} & & \\ G/H & & \end{array} \tag{79.575}$$

PrincHGGH

where π is the natural projection, the action of H is the right one and the wavy line means "acts on".

Proof. Surjectivity of β is clear. For injectivity, suppose $d\mu_g = d\mu_{g'}$. In order for the two target spaces to be equal, one needs $g' = gh$ for a $h \in H$. Now we have, for all $q_j \in \mathcal{Q}$,

$$d\mu_g q_j = d\mu_{gh} q_j = d\pi dR_{h^{-1}} dL_g dL_h q_j = d\pi dL_g\big(\operatorname{Ad}(h) q_j\big), \tag{79.576}$$

but $d\pi$ is an isomorphism from $\mathcal{Q}_g$, so we deduce that $q_j = \operatorname{Ad}(h) q_j$. Since we are using the connected component of $\mathrm{SO}(\mathcal{Q})$, that implies that $h = e$, and thus that $g = g'$. The following proves that β is a morphism:

$$\beta(gh) = d\pi dL_g dL_h = d\pi dL_g dL_h dR_{h^{-1}} = d\pi dL_g \operatorname{Ad}(h) = \beta(g) \cdot h.$$

□

The following lemma provides a convenient way to express the tangent bundle over G/H as an associated bundle to the principal bundle (79.575). We denote by $G \times_\rho \mathcal{Q}$ the quotient of $G \times \mathcal{Q}$ by the equivalence relation $(g, X) \sim (gh, \operatorname{Ad}(h^{-1})X)$ for all $h \in H$.

Lemma 79.123.
The map

$$\begin{aligned} \beta \colon G \times_\rho \mathcal{Q} &\to TM \\ [g, X] &\mapsto d\tau_g d\pi X \end{aligned} \tag{79.577}$$

with $\rho(h)X = \operatorname{Ad}(h^{-1})X$ is diffeomorphic.

LemBazHGGH

Proof. In order to check that β is well defined, first compute

$$\beta[gh, \operatorname{Ad}(h^{-1})X] = d\tau_{gh} d\pi \operatorname{Ad}(h^{-1})X = d\pi dL_{gh} \operatorname{Ad}(h^{-1})X,$$

and then using the fact that $d\pi dR_h = d\pi$, the latter line reduces to $d\pi dL_g X = \beta(g, X)$. For injectivity, let $\beta[g, X] = \beta[g', X']$. In order for these two to be vectors on the same point, there must exists a $h \in H$ such that $g' = gh$. The equality becomes $d\pi dL_g dL_h X' = d\pi dL_g X$. Commuting $d\pi$ with dL_g and using the fact that $d\tau_g$ is an isomorphism, we are left with the condition $d\pi dL_h X' = d\pi X$.

An element of $\mathcal{G}/\mathcal{H}$ is an equivalence class which contains exactly one element of $\mathcal{Q}$. In the right hand side of the condition, this element is X while the element of $\mathcal{Q}$ in the class $d\pi dL_h X$ is $\operatorname{Ad}(h)X'$. Equating these two elements, we find the condition $X' = \operatorname{Ad}(h^{-1})X$, which proves that $[g, X] = [g', X']$ and concludes the proof of the injectivity of β. □

The following proposition will prove useful in order to identity the spin structure over AdS_4.

Proposition 79.124.
If G is a connected Lie group and if Z is the center of G, then

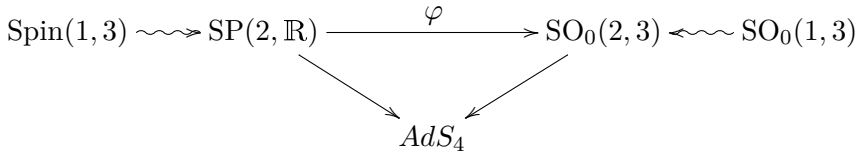
(1) Ad_G is an analytic homomorphism from G to $\operatorname{Int}(G)$, with kernel Z,

(2) the map $[g] \to \operatorname{Ad}_G(g)$ is an analytic isomorphism from G/Z to $\operatorname{Int}(\mathfrak{g})$ (the class $[g]$ is taken with respect to Z).

On the one hand that proposition together with the fact that $Z\big(\mathrm{SP}(2,\mathbb{R})\big) = \mathbb{Z}_2$ proves that the quotient $\mathrm{SP}(2,\mathbb{R})/\mathbb{Z}_2$ is isomorphic to $\mathrm{Int}\big(\mathfrak{sp}(2,\mathbb{R})\big)$. On the other hand one knows that $\mathrm{SO}_0(2,3)$ has no center, so that $\mathrm{SO}_0(2,3) \simeq \mathrm{Int}(\mathfrak{so}(2,3))$. But the subsection 55.8.2 provides an isomorphism between $\mathfrak{so}(2,3)$ and $\mathfrak{sp}(2,\mathbb{R})$. Thus we have

$$\mathrm{SP}(2,\mathbb{R})/\mathbb{Z}_2 \simeq \mathrm{SO}_0(2,3). \tag{79.578}$$

We denote by $\varphi\colon \mathrm{SP}(2,\mathbb{R}) \to \mathrm{SO}_0(2,3)$ the corresponding homomorphism with kernel $\mathbb{Z}_2$. In particular the restriction $\varphi|_{\mathrm{SL}(2,\mathbb{C})}$ is a double covering of $\mathrm{SO}_0(1,3)$. But χ is the same kind of double covering, so universality of $\mathrm{SL}(2,\mathbb{R})$ on $\mathrm{SO}_0(1,3)$ provides an automorphism $f\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(2,\mathbb{C})$ such that $\varphi = \chi \circ f$. The spin structure to be considered on AdS_4 is



where the action of $\mathrm{Spin}(1,3)$ on $\mathrm{SP}(2,\mathbb{R})$ is given by $a \cdot s = af^{-1}(s)$ where we identified $\mathrm{Spin}(1,3)$ with $\mathrm{SL}(2,\mathbb{C})$ as subgroup of $\mathrm{SP}(2,\mathbb{R})$. One immediately has $\varphi(a \cdot s) = \varphi(a)\chi(s)$.

79.27.2 Connection

There are a lot of ways to express a vector field $X\colon G/H \to T(G/H)$. From the identification $T(G/H) = G \times_\rho \mathcal{Q}$, one has $X\colon G/H \to G \times_\rho \mathcal{Q}$. As section of an associated bundle, X can be expressed by an equivariant function $\hat{X}\colon G \to \mathcal{Q}$ such that $X_{[g]} = [g, \hat{X}(g)]$. The H-equivariance of $\hat{X}$ means that $\hat{X}(gh) = \mathrm{Ad}(h^{-1})\hat{X}(g)$. Let $X \in \mathcal{G}$ and consider the function

$$\begin{aligned} \hat{A}_X\colon G &\to \mathcal{Q} \\ g &\mapsto \big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}} \end{aligned} \tag{79.579}$$

EqDefhatAcol

which is equivariant because the decomposition $\mathcal{G} = \mathcal{H} \oplus \mathcal{Q}$ is reductive. The corresponding vector field is

$$A_X[g] = \big[g, \big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}}\big];$$

or

$$A_X[g] = d\tau_g d\pi\big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}} = d\pi dL_g\big(\mathrm{Ad}(g^{-1})X\big)$$

because $d\pi X_{\mathcal{Q}} = d\pi X$. It is easy to check that the form

$$\omega_g(X) = -\big(dL_{g^{-1}}X\big)_{\mathcal{H}}$$

is a connection form on the principal bundle (79.575). We are going to determine the associated covariant derivative of this connection on the tangent space, and prove that it is torsion free. The horizontal lift of $A_X[g]$ is

$$\overline{A}_X(g) = dL_g\big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}} = \frac{d}{dt}\Big[ge^{t\,\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})X}\Big]_{t=0}. \tag{79.580}$$

EqovlAprQhor

The equivariant function associated with the covariant derivative of A_Y in the direction of A_X is given by $(\overline{A}_X)_g\hat{A}_Y$. Using expressions (79.579) and (79.580) of $\hat{A}_Y(g)$ and $\overline{A}_X(g)$, we have

$$\begin{aligned} (\bar{A}_X)_g\hat{A}_Y &= \frac{d}{dt}\Big[\hat{A}_Y\big(ge^{t\,\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})X}\big)_{\mathcal{Q}}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\Big(\mathrm{Ad}\big(e^{-t\,\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})X}g^{-1}\big)Y\Big)_{\mathcal{Q}}\Big]_{t=0} \\ &= \Big(\mathrm{ad}\big(-\mathtt{proj}_{\mathcal{Q}}\,\mathrm{Ad}(g^{-1})X\big)\,\mathrm{Ad}(g^{-1})Y\Big)_{\mathcal{Q}} \\ &= -\Big[\big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}}, \mathrm{Ad}(g^{-1})Y\Big]_{\mathcal{Q}}. \end{aligned}$$

This commutator is an expression of the form $[Z_{\mathcal{Q}}, Z'_{\mathcal{Q}} + Z'_{\mathcal{H}}]_{\mathcal{Q}}$. Using reducibility we find

$$(\overline{A}_X)_g \hat{A}_Y = -\Big[\big(\mathrm{Ad}(g^{-1})X\big)_{\mathcal{Q}}, \big(\mathrm{Ad}(g^{-1})Y\big)_{\mathcal{H}}\Big]. \tag{79.581}$$

The commutator produces

$$(\overline{A}_X)_g \hat{A}_Y - (\overline{A}_Y) g \hat{A}_X = -\hat{A}_{[X,Y]}(g),$$

which by construction the equivariant function associated with the vector field $\nabla_{A_X} A_Y - \nabla_{A_Y} A_X$; so on the one hand we have

$$(\nabla_{A_X} A_Y - \nabla_{A_Y} A_X)[g] = -d\tau_g d\pi \hat{A}_{[X,Y]}(g) = -d\tau_g d\pi \big(\mathrm{Ad}(g^{-1})[X,Y]\big)_{\mathcal{Q}} = -d\pi dR_g[X,Y].$$

On the other hand,

$$[A_X, A_Y][g] = d\pi[dR_g X, dR_g Y] = -d\pi dR_g[X,Y],$$

which proves that the connection is torsion free.

We are now going to study the horizontal vector fields on $\mathrm{SP}(2,\mathbb{R})$ with this connection and the homomorphism h^{-1} of equation (55.123). We have to study for which elements $\Sigma_a \in \mathrm{SP}(2,\mathbb{R})$ the expression

$$\omega_a(\Sigma_a) = \omega_{h^{-1}(a)}\big((dh^{-1})_a \Sigma_a\big) = -\Big(dL_{h^{-1}(a)^{-1}} dh^{-1} \Sigma_a\Big)_{\mathcal{H}} \tag{79.582}$$

vanishes. Every such element can of course be written under the form $\Sigma_a = dL_a \psi X$ for some $X \in \mathfrak{so}(2,3)$. So we are lead to consider the expression

$$(dh^{-1})_a (dL_a)_e \psi X. \tag{79.583}$$

It is easy to deal with that expression in the case of $a = e$:

$$(dh^{-1})_e \psi(X) = \psi^{-1}\psi X = X.$$

In particular, if $\Sigma \in \mathcal{T}$, then $dh^{-1}\Sigma \in \mathcal{Q}$ and when $\Sigma \in \mathcal{I}$, we have $dh^{-1}\Sigma \in \mathcal{H}$. This result propagates to other elements $a \in \mathrm{SP}(2,\mathbb{R})$ using the general result

$$df \circ dL_g = \big(dl_{f(g)}\big) \circ df$$

which holds for any group homomorphism f. Using that property with h^{-1} on the point $a \in \mathrm{SP}(2,\mathbb{R})$, we find $(dh^{-1})_a \circ (dL_a)_e = \big(dL_{h^{-1}(a)}\big) \circ (dh^{-1})_e$, and the expression (79.582) becomes

$$\omega_a(dL_a \psi X) = \Big(dL_{\left(h^{-1}(a)\right)^{-1}} dh^{-1} dL_a \psi X\Big)_{\mathcal{H}} = X_{\mathcal{H}}.$$

It is zero if and only if $X \in \mathcal{Q}$, so that the horizontal vectors on a are exactly the ones of $dL_a \psi \mathcal{Q} = \mathcal{T}_a$.

79.27.3 Dirac operator

When $\hat{s}\colon \mathrm{SP}(2,\mathbb{R}) \to \Lambda W$ is the equivariant function associated with a spinor, the Dirac operator reads

$$\widehat{Ds}(a) = g_{ij}\gamma^j \widehat{\nabla_{t_i} s}(a) = g_{ij}\gamma^j \bar{t}_i(a)\hat{s} = g_{ij}\gamma^j \tilde{t}_i(a)\hat{s} \tag{79.584}$$

where the metric g is the usual four-dimensional Minkowskian metric and the matrices γ are the associated 4×4 Dirac matrices. The elements $\tilde{t}_i(a) = dL_a t_i = dL_a \psi(q_i)$ span the natural basis of $\mathcal{T}_a$, see appendix 55.8.3. The matrices γ^i are the usual 4×4 Dirac matrices for the 4-dimensional Minkowskian metric.

¡¡ Avertissement/question à la lectrice !! 79.125
Et il serait aussi pas mal de préciser un peu une fois tout de suite ce qu'est l'espace ΛW.

One can find a change of basis which express the Dirac operator in terms of vectors on R_1. For that, let $\{X_i\}$ be a basis of $\mathcal{R}_1$. We have

$$X_i^*[u] = \frac{d}{dt}\Big[[e^{-tX_i}u]\Big]_{t=0} = -d\pi dR_u X_i$$

that is necessarily decomposable by corollary 66.6 as combinations of vectors of the form $d\pi dL_u q_i$ because $[u]$ belongs to an open orbit of the action of R_1. That defines a matrix B by

$$d\pi dL_u q_i = B_{ij} d\pi dR_u X_j,$$

and then a vector $Y \in \mathcal{H}$ by

$$q_i = B_{ij}\,\mathrm{Ad}(u^{-1})X_j + Y. \tag{79.585}$$

Now we have $\tilde{t}_i(a) = dL_a\psi\big(\mathrm{Ad}(u^{-1})B_{ij}X_j + Y\big)$. We can go further using the fact that

$$\psi\Big(\mathrm{Ad}\big(h^{-1}(a)\big)X\Big) = \mathrm{Ad}(a)\psi(X) \tag{79.586}$$

for every $a \in \mathrm{SP}(2,\mathbb{R})$ and $X \in \mathcal{G}$. Defining the vectors $s_i = \mathrm{Ad}\big(h(^{-1})\big)\psi X_i$ we find

$$\tilde{t}_i(a) = B_{ij}\tilde{s}_i(a) + \widetilde{\psi(Y)}(a). \tag{79.587}$$

79.27.4 Frame bundle

Construction of the frame bundle and the spin structure is a straightforward adaptation of theorem 2.2 (chapter???) in [1033], while Dirac operator and connection issues are adapted from proposition 1.3 (chapter III)

A **basis** of a m dimensional vector space V is a free and generating part; it only has the structure of a set. A frame of the vector space V is a nondegenerate map $b\colon \mathbb{R}^m \to V$. Let us give an example in three dimensions the difference. If $\{v_1, v_2, v_3\}$ is a basis of V, of course $\{v_2, v_1, v_3\}$ is the same basis. Order has no importance. But if $\{e_1, e_2, e_3\}$ is the canonical basis of $\mathbb{R}^3$, the *frames* $b(e_1) = v_1$, $b(e_2) = v_2$, $b(e_3) = v_3$ and $c(e_1) = v_2$, $c(e_2) = v_1$, $c(e_3) = v_3$ are not the same.

Now we consider $AdS_l = G/H = \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1)$, the Lie algebra $\mathcal{G}$ has a reductive homogeneous space decomposition $\mathcal{G} = \mathcal{Q} \oplus \mathcal{H}$ and we consider the canonical projection $\pi\colon G \to AdS_l$.

Let the map (see relation (79.51))

$$\begin{aligned} \alpha\colon H &\to \mathrm{SO}(\mathcal{Q}) \\ h &\mapsto \mathrm{Ad}(h)|_{\mathcal{Q}}. \end{aligned} \tag{79.588}$$

We consider, on $G \times \mathrm{SO}(\mathcal{Q})$, the equivalence relation $(g, A) \sim (g', A')$ if and only if there exists $h \in H$ such that $g' = gh$ and $A' = \alpha(h^{-1})A$. We denote by $G \times_\alpha \mathrm{SO}(\mathcal{Q})$ the set of equivalence classes. Now we have a principal bundle

$$\begin{array}{ccc} \mathrm{SO}(\mathcal{Q}) & \rightsquigarrow & G \times_\alpha \mathrm{SO}(\mathcal{Q}) \\ & & \big\downarrow p \\ & & G/H \end{array} \tag{79.589}$$

EqPrincPreR

where $p[g, A] = [g]$ and the action is given by $[g, A] \cdot B = [g, A]$. The fact that the projection fulfils $p\big([g, A] \cdot B\big) = p[g, A]$ is evident, and the fact that the action is well defined is a simple computation: if $[g', A'] = [g, A]$, we have a $h \in H$ such that

$$[g', A'] \cdot B = [g', A'B] = [gh, \alpha(h^{-1})AB] = [g, AB] = [g, A] \cdot B.$$

Proposition 79.126.
Let $\tau(g)\colon AdS_l \to AdS_l$ be the action of $g \in G$ on AdS_l: $\tau(g)[g'] = [gg']$, and B be the frame bundle. We also consider the map $\sigma\colon \mathbb{R}^{1,l-1} \to \mathcal{Q}$ the isometry which sends the canonical basis of $\mathbb{R}^{1,l-1}$ to the usual basis $\{q_0, q_1, \dots, q_{l-1}\}$ of $\mathcal{Q}$. The map

$$\begin{aligned} \beta\colon G \times_\alpha \mathrm{SO}(\mathcal{Q}) &\to B \\ [g, A] &\mapsto d\tau(g)_\vartheta A \circ \sigma \end{aligned} \tag{79.590}$$

provides a principal bundle isomorphism between the principal bundle (79.589) and the frame bundle over AdS_l.

By abuse of notation, we will not always write the σ.

Proof. We have to prove first that the map $\beta\colon G \times \mathrm{SO}(\mathcal{Q}) \to B$ respects the classes. For that, consider $(g, A) \sim (g', A')$ and remark that

$$\begin{aligned} \beta(gh, \alpha(h^{-1})) &= d\tau(gh)_\vartheta \alpha(h^{-1})A = d\tau(g)d\tau(h)d\pi \,\mathrm{Ad}(h^{-1})d\pi^{-1}A \\ &= d\tau(g)d\tau(h)d\pi \,\mathrm{Ad}(h^{-1})d\pi^{-1}A = d\tau(g)d\pi dR_h d\pi^{-1}A \\ &= d\tau(g)_\vartheta A = \beta(g, A). \end{aligned}$$

where we used equation (79.13) and the fact that $\pi \circ L_g = \tau(g) \circ \pi$. The frame bundle is

$$\begin{array}{ccc} \mathrm{SO}(1, l-1) & \rightsquigarrow & B \\ & & \downarrow p \\ & & G/H \end{array} \tag{79.591}$$

EqPrincB

where the fibre $B_{[g]}$ in B over $[g]$ is the set of isometric maps $\mathbb{R}^{1,l-1} \to T_{[g]}(AdS_l)$. So an element of B is of the form $\big([g], \tilde{f} \circ \sigma\big)$ where $g \in G$ and $\tilde{f}\colon \mathcal{Q} \to T_{[g]}(AdS_l)$ contains the main information while σ is the previously explained isometry. The action of $h \in \mathrm{SO}(1, l-1)$ on $\big([g], \tilde{f} \circ \sigma\big)$ is defined by means of any fixed isomorphism $\varphi_0\colon \mathrm{SO}(1, l-1) \to \mathrm{SO}(\mathcal{Q})$ by

$$\big([g], \tilde{f} \circ \sigma\big) \cdot h = \big([g], \tilde{f} \circ \varphi_0(h) \circ \sigma\big). \tag{79.592}$$

The map β is a morphism of principal bundle because

$$\beta[g, A] \cdot \varphi_0^{-1}(B) = \big([g], d\tau(g)A \circ \sigma\big) \cdot \varphi_0^{-1}(B) = \big([g], d\tau(g)A \circ B \circ \sigma\big) = \beta\big([g, A] \cdot B\big).$$

It remains to be proved that β is a bijection. Surjectivity is natural: since $d\tau(g)$ is an isometry, $d\tau(g)A$ runs over the whole $\mathrm{SO}\left(T_{[g]}(AdS_l)\right)$ when A runs over $\mathrm{SO}(\mathcal{Q})$. Injectivity is as follows; let's suppose $\beta[g, A] = \beta[g', A']$. It is immediate that in this case, $\exists h \in H$ such that $g' = gh$. Using the fact that $d\pi \circ dR_{h^{-1}} \circ d\pi^{-1} = \mathrm{Id}$ and $d\tau(h)d\pi = d\pi dL_h$, we have

$$d\tau(g)_\vartheta A = d\tau(g)d\tau(h)A' = d\tau(g)d\tau(h)d\pi dR_{h^{-1}} d\pi^{-1}A' = d\pi \,\mathrm{Ad}(h)d\pi^{-1}A' = \alpha(h)A'.$$

□

From now on, we identify $G \times_\alpha \mathrm{SO}(\mathcal{Q})$ with the frame bundle over AdS_l.

79.27.5 Spin structure

We consider the principal bundle

$$\begin{array}{ccc} \mathrm{Spin}(1, l-1) & \rightsquigarrow & G \times_{\tilde{\alpha}} \mathrm{Spin}(1, l-1) \\ & & \downarrow p \\ & & G/H \end{array} \tag{79.593}$$

where $\times_{\tilde{\alpha}}$ is the following equivalence relation on $G \times \mathrm{Spin}(1, l-1)$. We say that $(g, s) \sim (g', s')$ if and only if there exists a $h \in H$ such that

(1) $g' = gh$,

(2) $\chi(s') = \mathrm{Ad}(h^{-1})\chi(s)$.

Notice that the second condition implies that $\mathrm{Ad}(h) \in \mathrm{SO}_0(\mathcal{Q})$. It is easy to prove that the given structure is well defined and is a principal bundle. Now we consider the spin structure as follows:

$$\begin{array}{ccccccc} \mathrm{Spin}(1,l-1) & \rightsquigarrow & G\times_{\tilde{\alpha}}\mathrm{Spin}(1,l-1) & \xrightarrow{\varphi} & G\times_{\alpha}\mathrm{SO}(\mathcal{Q}) & \rightsquigarrow & \mathrm{SO}(\mathcal{Q}) \\ & & & \searrow \quad \swarrow & & & \\ & & & G/H & & & \end{array} \tag{79.594}$$

where $\varphi[g,s] = [g,\chi(s)]$. It is well defined since when $[g,s] = [g',s']$, there exists a $h \in H$ with $\chi(s') = \mathrm{Ad}(h^{-1})\chi(s)$ such that $\varphi[g',s'] = \varphi[gh,s'] = [gh,\chi(s')] = [gh,\mathrm{Ad}(h^{-1})\chi(s)] = [g,\chi(s)] = \varphi[g,s]$.

79.27.6 Reduction of the structural group

The case of AdS_l can be seen in the setting of subsection 66.3.8. Let us show now that the bundle

$$\begin{array}{ccc} H_0 & \rightsquigarrow & G \\ & & \downarrow{\scriptstyle \pi} \\ & & G/H \end{array} \tag{79.595}$$

EqPrincHzGM

is a reduction to H_0 (the identity component of $\mathrm{SO}(\mathcal{Q})$) of

$$\begin{array}{ccc} G & \rightsquigarrow & r(G) \\ & & \downarrow{\scriptstyle \pi} \\ & & G/H. \end{array} \tag{79.596}$$

Indeed, $u\colon G \to r(G)$ given by $u(g) = r(g)$ provides the reduction homomorphism: $r(gh)X = d\pi dL_{gh}X$ while $\big(r(g)\cdot h\big)X$ is the same.

Lemma 79.127.
The tangent space $T(G/H)$ is an associated bundle of $r(G)$ trough the identification

$$\begin{aligned} \beta'\colon r(G)\times_{\rho}\mathcal{Q} &\to T(G/H) \\ [r(g),X] &\mapsto r(g)X \end{aligned} \tag{79.597}$$

where $\rho(h)X = \mathrm{Ad}(h)X$, so that the quotient is given by $[g,X] = [gh,\mathrm{Ad}(h^{-1})X]$.

Proof. The proof is entirely similar to the one of lemma 79.123. □

79.28 Conclusion

In a first time we defined a black hole in anti de Sitter space. This construction is not related to any metric divergence but is a dimensional generalization of a causal black hole whose singularity is dictated by causal issues. The originality of our approach lies in the fact that our method uses essentially group theoretical and symmetric spaces techniques. That result should be generalisable to any semisimple symmetric space.

Then we proved that the physical domain of the black hole (the non singular part) is equivalent to a group in the sense that there exists a group which acts freely and transitively by diffeomorphisms. So we identify the group with the manifold and it is easy to prove that the latter group is a split extension of an Heisenberg group which happens to be quantizable by a twisted pull-back of a previously known quantization of $\mathrm{SU}(1,n)$.

We also proved two somewhat out of subject small results. The first one is the fact that a deformation of the half-plane by Unterberger can be transported to a deformation of the Iwasawa subgroup of $\mathrm{SL}(2,\mathbb{R})$ which can in turn deform (by the group action method) the dual of its Lie algebra. We showed however that that deformation is not universal; indeed we pointed out two different actions of the Iwasawa subgroup of $\mathrm{SL}(2,\mathbb{R})$ on AdS_2 for which the deformation by group action method reveals to be unable to even multiply two compactly supported functions. An interesting question is to know the precise point in the construction of Unterberger which makes his product non universal.

The second small result is a proof of concept for quantization of the Iwasawa subgroup of $\mathrm{SO}(2,n)$ by the method of the extension lemma. We wrote $\mathrm{SO}(2,n)$ as a symplectic split extension of $\mathrm{SU}(1,n)$ by $\mathrm{SU}(1,1)$. The extension lemma then provided a kernel on $\mathrm{SO}(2,n)$ because kernels were known on $\mathrm{SU}(1,1)$ and $\mathrm{SU}(1,n)$. Is that quantization equivalent in some sense to the one that we performed in the main line of the black hole deformation? That question still has to be solved.

As a final remark, I want to point out that the major challenge of this century is not quantization, but global warming.

Chapter 80

General non commutative geometry

My main references for noncommutative geometry are [819, 953, 962, 941].

80.1 Non commutative differential forms

80.1.1 Universal differential forms

Let $\mathfrak{A}$ be an associative unital algebra on $\mathbb{C}$. We are going to define step by step the **universal algebra** of differential forms

$$\Omega\mathfrak{A} = \bigoplus_p \Omega^p\mathfrak{A}.$$

Existence the universal algebra will be proved by explicitly construction later. For unicity, we will prove an universality property.

First of all, $\Omega^0\mathfrak{A} = \mathfrak{A}$. Next, $\Omega^1\mathfrak{A}$ is the left $\mathfrak{A}$-module generated by the symbols δa with $a \in \mathfrak{A}$ and relations

$$\delta(ab) = (\delta a)b + a\delta b \tag{80.1a}$$
seq, deltaabi

$$\delta(\alpha a + \beta b) = \alpha\delta a + \beta\delta b \tag{80.1b}$$

for all $a, b \in \mathfrak{A}$ and $\alpha, \beta \in \mathbb{C}$. A general element in $\Omega^1\mathfrak{A}$ is of the form $\sum_i a_i\delta b_i$, with a_i and b_i in $\mathfrak{A}$.

Equation (80.1a) with $a = 1$ gives $(\delta 1)b = 0$ for all $b \in \mathfrak{A}$, hence $\delta(1) = 0$ and $\delta(\mathbb{C}) = 0$. So $\Omega^1\mathfrak{A}$ is a left $\mathfrak{A}$-module; we can give a structure of right $\mathfrak{A}$-module by defining

$$(\sum_i a_i\delta b_i)c = \sum_i a_i(\delta b_i)c,$$

but equation (80.1a) gives $(\delta a)c = \delta(ac) - a\delta c$, therefore if $\omega = \sum_i a_i\delta b_i$,

$$\omega c = \sum_i a_i\delta(b_i c) - \sum a_i b_i \delta c. \tag{80.2}$$

The rule (80.1a) is Leibniz for the map $\delta\colon \mathfrak{A} \to \Omega^1\mathfrak{A}$, so we see δ as a derivation of $\mathfrak{A}$ with values in the bimodule $\Omega^1\mathfrak{A}$.

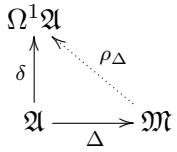
80.1.1.1 Universal properties

The following proposition gives an universal property of $\Omega^1\mathfrak{A}$; in a certain sense, it is unique.

Proposition 80.1.
Let $\mathfrak{M}$ be a $\mathfrak{A}$-bimodule and $\Delta\colon \mathfrak{A} \to \mathfrak{M}$ a derivation, i.e.

$$\Delta(ab) = (\Delta a)b + a\Delta b.$$

There exists one and only one bimodule morphism $\rho_\Delta\colon \Omega^1\mathfrak{A} \to \mathfrak{M}$ *such that* $\Delta = \rho_\Delta \circ \delta$*, i.e. the following diagram commutes:*



prop_modMununique

Proof. First, remark that any bimodule morphism $\rho\colon \Omega^1\mathfrak{A} \to \mathfrak{M}$ makes $\rho \circ \delta$ a derivation with values in $\mathfrak{M}$. Indeed

$$(\rho \circ \delta)(ab) = \rho\big((\delta a)b + a\delta b\big) = (\rho \circ \delta)(a)b + a(\rho \circ \delta)b$$

Let us now prove the inverse: let $\Delta\colon \mathfrak{A} \to \mathfrak{M}$ be a derivation and $\rho_\Delta\colon \Omega^1\mathfrak{A} \to \mathfrak{M}$ be such that $\Delta = \rho_\Delta \circ \delta$. We want to prove unicity of this derivation. First, definition of Δ makes

eq_rhoDelta

$$\rho_\Delta(\delta a) = \Delta a. \tag{80.3}$$

This completely defines ρ_Δ from Δ because δ generates the whole $\Omega^1\mathfrak{A}$ as left $\mathfrak{A}$-module. Indeed the only way to extends ρ_Δ from (80.3) as a morphism on $\Omega^1\mathfrak{A}$ is

$$\rho_\Delta\big(\sum_i a_i\delta b_i\big) = \sum_i a_i\Delta b_i.$$

Now it is sufficient to prove that (80.3) is a bimodule morphism:

$$\begin{aligned}\rho_\Delta\big(f(\textstyle\sum a_i\delta b_i)g\big) &= \rho_\Delta\big(\sum_i fa_i[\delta(b_ig) - b_i\delta g]\big)\\ &= \textstyle\sum fa_i[\Delta(b_ig) - b_i\Delta g]\\ &= \textstyle\sum fa_i(\Delta b_i)g\\ &= f\big(\textstyle\sum ai\Delta b_i\big)g.\end{aligned}$$

□

The space $\Omega^p\mathfrak{A}$ is defined by

$$\Omega^p\mathfrak{A} := \underbrace{\Omega^1\mathfrak{A}\dots\Omega^1\mathfrak{A}}_{p\text{times}}$$

with multiplication rule

$$(a_0\delta a_1)(b_0\delta b_1) := a_0(\delta a_1)b_0\delta b_1 = a_0\delta(a_1b_0)\delta b_1 - a_0a_1\delta b_0\delta b_1. \tag{80.4}$$

This rule serves to show how to write elements of $\Omega^2\mathfrak{A}$ under the form $a\delta b_1\delta b_2$. Elements of $\Omega^p\mathfrak{A}$ are linear combination of elements of the form

$$\omega = a_0\delta a_1\delta a_2\dots\delta a_p$$

with $a_k \in \mathfrak{A}$, and the product $\Omega^p\mathfrak{A} \times \Omega^q\mathfrak{A} \to \Omega^{p+q}\mathfrak{A}$ is a juxtaposition and a rearrangement:

$$\begin{aligned}&(a_1\delta a_1\dots\delta a_p)(a_{p+1}\delta a_{p+2}\dots\delta a_{p+q})\\ &\qquad := a_0\delta a_1\dots(\delta a_p)a_{p+1}\delta a_{p+2}\dots\delta a_{p+q}\\ &\qquad = a_0\delta a_1\dots\big[\delta(a_pa_{p+1}) - a_p(\delta a_{p+1})\big]\delta a_{p+2}\dots\delta a_{p+q}\\ &\qquad = a_0\delta a_1\dots\delta(a_pa_{p+1})\delta a_{p+2}\dots\delta a_{p+q}\\ &\qquad\quad - a_0\delta a_1\dots\big[\delta(a_{p-1}a_p) - a_{p-1}\delta a_p\big]\delta a_{p+1}\dots\delta a_{p+q}.\end{aligned}$$

With $p = q = 3$ for example, we have

$$\begin{aligned}(a_0\delta a_1\delta a_2)(a_3\delta a_4\delta a_5) &= a_0\delta a_1(\delta a_2)a_3\delta a_4\delta a_5\\ &= a_0\delta a_1\big[\delta(a_2a_4) - a_2\delta a_3\big]\delta a_4\delta a_5\\ &= a_0\delta a_1\delta(a_2a_3)\delta a_4\delta a_5\\ &\quad - a_0\big[\delta(a_1a_2) - a_1\delta a_2\big]\delta a_3\dots\delta a_5\\ &= a_0\delta a_1\delta(a_2a_3)\delta a_4\delta a_5\\ &\quad - a_0\delta(a_1a_2)\delta a_3\dots\delta a_5\\ &\quad + a_0a_1\delta a_2\dots\delta a_5,\end{aligned}$$

and in general,

$$\begin{aligned}&(a_0\delta a_1\dots\delta a_p)(a_{p+1}\delta a_{p+2}\dots\delta a_{p+q})\\ &\qquad = (-1)^p a_0a_1\delta a_2\dots\delta a_{p+q}\\ &\qquad\quad + \sum_{i=1}^{p}(-1)^{p-i}a_0\delta a_1\dots\delta a_{i-1}\delta(a_ia_{i+1})\delta a_{i+2}\dots\delta a_{p+q}.\end{aligned} \tag{80.5}$$

Eq_decmProdConnForDelta

So $\Omega\mathfrak{A}$ is a left $\mathfrak{A}$-module. We turn it into $\mathfrak{A}$-bimodule by defining

$$\begin{aligned}(a_0\delta a_1\dots\delta a_p)b &= a_0\delta a_1\dots(\delta a_p)b\\ &= (-1)^p a_0a_1\delta a_2\dots\delta a_pb\\ &\quad + \sum_{i=1}^{p-1}(-1)^{p-i}a_0\delta a_1\dots\delta a_{i-1}\delta(a_ia_{i+1})\delta_{i+1}\dots\delta a_p\delta b\\ &\quad + a_0\delta a_1\dots\delta a_{p-1}\delta(a_pb).\end{aligned} \tag{80.6}$$

Now we put a differential algebra structure on the $\mathfrak{A}$-bimodule $\Omega\mathfrak{A}$ by extending δ to

$$\begin{aligned}&\delta\colon \Omega^p\mathfrak{A}\to\Omega^{p+1}\mathfrak{A}\\ &\delta(a_0\delta a_1\dots\delta a_p) := \delta a_0\delta a_1\dots\delta a_p.\end{aligned} \tag{80.7}$$

One checks that

$$\delta^2 = 0$$

and

$$\delta(\omega_1\omega_2) = (\delta\omega_1)\omega_2 + (-1)^p\omega_1\delta\omega_2 \tag{80.8}$$

for any $\omega_1\in\Omega^p\mathfrak{A}$ and $\omega_2\in\Omega\mathfrak{A}$.

If $\mathcal{E}$ is a $\mathfrak{A}$-module, an element $\alpha\in \mathrm{End}_{\mathfrak{A}}(\mathcal{E})\otimes\Omega^1\mathfrak{A}$ acts on an element of $\mathcal{E}\otimes_{\mathfrak{A}}\Omega^p\mathfrak{A}$ in the following way. If $\alpha = \sum_i\big(A_i\otimes_{\mathfrak{A}}a_0^i\delta a_1^i\big)$, we define

$$\alpha\big(\sum_j\xi_j\otimes_{\mathfrak{A}}\omega_j\big) = \sum_{ij}\big(A_i\xi_j\otimes_{\mathfrak{A}}a_0^i\delta a_1^i\omega_j\big)\in\mathcal{E}\otimes_{\mathfrak{A}}\Omega^{p+1}\mathfrak{A} \tag{80.9}$$

EqActallphaEOAp

where $A_i\in\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$, $a_j^i\in\mathfrak{A}$, $\xi_j\in\mathcal{E}$ and $\omega_j\in\Omega^p\mathfrak{A}$. The element α acts in particular on $\mathcal{E}$ *via* the identification $\xi\in\mathcal{E}\leftrightarrow\xi\otimes_{\mathfrak{A}}1\in\mathcal{E}\otimes_{\mathfrak{A}}\Omega^0\mathfrak{A}$. In this case the result is denoted by $\alpha(\xi)$.

The following proposition gives the same type of result as proposition 80.1.

Proposition 80.2.
Let (Γ,Δ) a differential graded algebra and $\rho\colon\mathfrak{A}\to\Gamma^0$, a morphism of unital algebras. There exists one and only one extension of ρ into a differential graded algebra morphism $\tilde\rho\colon\Omega\mathfrak{A}\to\Gamma$ with $\tilde\rho\circ\delta = \Delta\circ\tilde\rho$:

$$\begin{array}{ccc}\Gamma^p & \xrightarrow{\Delta} & \Gamma^{p+1}\\ \downarrow{\scriptstyle\tilde\rho} & & \downarrow{\scriptstyle\tilde\rho}\\ \Omega^p\mathfrak{A} & \xrightarrow[\delta]{} & \Omega^{p+1}\mathfrak{A}\end{array} \tag{80.10}$$

eq_diagrhotrho

Proof. The map $\rho\colon \mathfrak{A} \to \Gamma^0$ being given, we define

$$\begin{aligned} \tilde{\rho}\colon \Omega^p\mathfrak{A} &\to \Gamma^p \\ \tilde{\rho}(a_0\delta a_1 \dots \delta a_p) &:= \rho(a_0)\Delta(\rho(a_1)) \dots \Delta(\rho(a_p)). \end{aligned} \tag{80.11}$$

It well defines $\tilde{\rho}$ since $\Omega^p\mathfrak{A}$ is generated by $a_0\delta a_1 \dots \delta a_p$. It sends products on products; for example with $p = 2$,

$$\begin{aligned} \tilde{\rho}\big[(a_0\delta a_1)(b_0\delta b_1)\big] &= \rho(a_0)\Delta\rho(a_1b_0)\Delta\rho b_1 \\ &\quad - \rho(a_0a_1)\Delta\rho b_0\Delta\rho b_1 \\ &= \big[\rho(a_0)\Delta\rho(a_1)\big]\big[\rho(b_0)\Delta\rho b_1\big]. \end{aligned}$$

Commutativity of diagram (80.10) is as follows:

$$\begin{aligned} (\tilde{\rho}\circ\Delta)(a_0\delta a_1 \dots \delta a_p) &= \tilde{\rho}(\delta a_0 \dots \delta a_p) \\ &= \Delta\rho a_0 \dots \Delta\rho a_p \\ &= \Delta\big(\rho(a_0)\Delta\rho a_1 \dots \Delta\rho a_p\big) \\ &= (\Delta\circ\tilde{\rho})(a_0\delta a_1 \dots \delta a_p). \end{aligned}$$

□

80.1.1.2 Cohomology

From definition,

$$\delta(a_0\delta a_1 \dots \delta a_p) = \delta a_0 \dots \delta a_p.$$

If $\omega = a_0\delta a_1 \dots \delta a_p$, the only way to have $\delta\omega = 0$ is $a_0 = 1$, i.e. $\omega = \delta(a_1\delta a_2 \dots \delta a_p)$. Then

$$H^p(\Omega\mathfrak{A}) = 0$$

when $p \neq 0$ because any closed form is exact. In the case $p = 0$,

$$H^0(\Omega\mathfrak{A}) = \mathbb{C}$$

because $\delta(z) = 0$ for all $z \in \mathbb{C}$ while z is not exact.

80.1.1.3 Isomorphisms

Lemma 80.3.
As bimodule, $\Omega^1(\mathfrak{A})$ is isomorphic to $\ker m$ where $m\colon \mathfrak{A}\otimes_{\mathbb{C}}\mathfrak{A} \to \mathfrak{A}$ is the multiplication map. The isomorphism is given by

$$\begin{aligned} \varphi\colon \ker m &\to \Omega^1(\mathfrak{A}) \\ \sum_j a_i \otimes b_i &\mapsto \sum_j a_i db_i. \end{aligned} \tag{80.12}$$

EqDEfphirov

Proof. First remark that $\ker m$ is generated by elements of the form $1\otimes_{\mathbb{C}} a - a\otimes_{\mathbb{C}} 1$. Indeed if $\sum_j a_jb_j = m\big(\sum_i a_j\otimes b_j\big) = 0$, we have $\sum_j a_j\otimes_{\mathbb{C}} b_j = \sum_j a_j(1\otimes_{\mathbb{C}} b_j - b_j\otimes_{\mathbb{C}} b_j)$.

Now consider the map

$$\begin{aligned} \Delta\colon \mathfrak{A} &\to \ker m \\ \Delta a &= 1\otimes_{\mathbb{C}} a - a\otimes_{\mathbb{C}} 1 \end{aligned} \tag{80.13}$$

This map satisfies $\Delta(ab) = (\Delta a)b + a\Delta b$. Now we prove that the equation (80.12) is surjective: $\sum_j a_jdb_j = \varphi\big(\sum_j a_j(1\otimes b_j - b_j\otimes 1)\big)$ where indeed $m\big(\sum_j a_j(1\otimes b_j - b_j\otimes 1)\big) = 0$.

The injectivity is evident because $\varphi\big(\sum_j a_j\otimes b_j\big) = \varphi\big(\sum_j a'_j\otimes b'_j\big)$ implies $\sum_j a_jdb_j = \sum_j a'_jdb'_j$.

The fact that φ provides an isomorphism of bimodule is

$$\varphi\big(c(a_i\otimes b_i)y\big) = \varphi\big((xa_i)\otimes(b_i)\big) = xa_id(b_iy) = xa_i(db_i)y + x\underbrace{a_ib_i}_{=0}dy = x\varphi(a_i\otimes b_i)y.$$

□

80.1.1.4 Involution

If the algebra $\mathfrak{A}$ has an involution (which is often the complex conjugation when $\mathfrak{A}$ is a function algebra), the universal algebra $\Omega\mathfrak{A}$ becomes an involutive algebra with the definition

$$(\delta a)^* = -\delta(a^*),$$
$$(a_0\delta a_1\cdots\delta a_p)^* = (\delta a_p)^*\cdots(\delta a_1)^* a_0^*.$$

80.1.2 Connes differential forms

Proposition 80.4.
The formula

$$\pi(a^0\delta a^1\cdots\delta a^n) = a^0[D,a^1]\cdots[D,a^n]$$

defines a $*$-representation of the reduced universal algebra on $\mathscr{H}$.

Proof. No proof. □

Proposition 80.5.
Let $J_0 = \ker\pi$ and $J_0^{(k)} = \{\omega \in \Omega^k(\mathfrak{A})$ tel que $\pi(\omega) = 0\}$. In this case, $J := J_0 + \delta J_0$ is a graded differential two-sided ideal of $\Omega^*(\mathfrak{A})$.

Proof. The fact that J is differential comes from the fact that $d^2 = 0$, but one has to remark that J_0 is not differential by itself because there exists some **junk form** ω such that $\pi(\omega) = 0$ and $\pi(d\omega) \neq 0$. In order to see that J is a left ideal, consider $\omega \in J^{(k)}$:

$$\omega = \omega_1 + \delta\omega_2$$

with $\omega_1 \in J_0 \cap \Omega^k$ and $\omega_2 \in J_0 \cap \Omega^{k-1}$. If $\omega' \in \Omega^l$, we have

$$\omega\omega' = \big(\omega_1\omega' + (-1)^k\omega_2\delta\omega'\big) + \delta(\omega_2\omega') \in J^{(k+l)}.$$

The right side is proven in the same way. □

Since J is an ideal, one can define

$$\Omega_D^*(\mathfrak{A}) = \Omega^*(\mathfrak{A})/J.$$

One can prove that for all $k \in \mathbb{N}$, $\Omega_D^k(\mathfrak{A}) \simeq \pi(\Omega^k(\mathfrak{A}))/\pi\big(d(J_0 \cap \Omega^{k-1})\big)$. We consider the product

$$\langle T_1, T_2\rangle = \mathrm{Tr}_\omega(T_2^* T_1 |D|^{-}d) \tag{80.14}$$

on $\pi(\Omega^k)$, and we define $\mathscr{H}_k$ as the completion of $\pi(\Omega^k)$ for this product. This provides a Hilbert space in which $\pi\big(d(J_0 \cap \Omega^{k-1})\big)$ is a subspace. We denote by P the projection parallel to this space.

Proposition 80.6.
For all $\omega_i \in \pi\big(\Omega^k(\mathfrak{A})\big)$,

$$\langle P\omega_1, \omega_2\rangle = \langle P\omega_1, P\omega_2\rangle.$$

Proof. No proof. □

We denote by Λ^k the completion of Ω_D^k with respect to this inner product. In fact, the proposition shows that the inner product can be used on $\pi(\Omega^k)$.

Let $(\mathfrak{A}, \mathscr{H}, D)$ be a spectral triple, and $\Omega\mathfrak{A}$, the universal algebra of $\mathfrak{A}$. We consider the map

$$\begin{aligned}\pi\colon \Omega\mathfrak{A} &\to \mathfrak{B}(\mathscr{H})\\ a_0\delta a_1\ldots\delta a_p &\mapsto a_0\circ[D,a_1]\circ\ldots\circ[D,a_p]\end{aligned} \tag{80.15}$$

for $a_i \in \mathfrak{A}$. Since the operations δ and D both are derivations, one can expect that π will be a homomorphism.

Proposition 80.7.
The operation π is a homomorphism.

Proof. We will just check it in the case of 1-forms. First, remark that

$$\begin{aligned}(a_0\delta a_1)(b_0\delta b_1) &= a_0\big(\delta(a_1b_0) - a_1\delta b_0\big)\delta b_1\\ &= a_0\delta(a_1b_0)\delta b_1 - a_0a_1\delta b_0\delta b_1.\end{aligned}$$

Thus we have

$$\begin{aligned}\pi\big((a_0\delta a_1)(b_0\delta b_1)\big) &= a_0[D,a_1b_0][D,b_1] - a_0a_1[D,b_0][d,b_1]\\ &= a_0[D,a_1]b_0[D,b_1]\\ &= \pi(a_0\delta a_1)\pi(b_0\delta b_1)\end{aligned}$$

□

80.1.2.1 Junk forms

One cannot define $\pi(\Omega\mathfrak{A})$ as differential forms because there exists some $\omega \in \Omega\mathfrak{A}$ such that $\pi(\omega) = 0$ and $\pi(\delta\omega) \neq 0$. Such a form is said to be **junk**. We define

$$J_0^p = \{\omega \in \Omega^p\mathfrak{A} \text{ tel que } \pi(\omega) = 0\} \tag{80.16}$$

eq__defJConnes

and $J = J_0 + \delta J_0$. Then we define the **Connes differential forms** as

$$\Omega_D\mathfrak{A} = \Omega\mathfrak{A}/J. \tag{80.17}$$

Proposition 80.8.
Let $J_0 = \bigoplus_p J_0^p$, the two-sided graded ideal of $\Omega\mathfrak{A}$ generated by (80.16). The set $J = J_0/\delta J$ is a two-sided graded ideal of $\Omega\mathfrak{A}$.

Proof. We begin by proving that J is a differential algebra for the same δ as for $\mathfrak{A}$; the fact that $\delta^2 = 0$ is not a question. We have to prove that δ is internal in J. A general element of J is $\omega = \alpha + \delta\beta$ with $\pi(\alpha) = \pi(\beta) = 0$. In this case, $\delta\omega = \delta\alpha$ is of the same form. Now we want to prove that J is a two-sided ideal. Let $\omega = \omega_1 + \delta\omega_2 \in J^p$ with $\omega_1 \in J_0^p$ and $\omega_2 \in J_0^{p-1}$ and consider $\eta \in \Omega^q\mathfrak{A}$. We have

$$\omega\eta = \omega_1\eta + (\delta\omega_2)\eta = \omega_1\eta + \delta(\omega_2\eta) - (-1)^{p-1}\omega_2\delta\eta.$$

The first term fulfil $\pi(\omega_1\eta) = \pi(\omega_1)\pi(\eta) = 0$, the second one is the δ of something whose π is zero while the π of the third one is zero. So $\omega\eta \in J$. In the same manner, we conclude that $\eta\omega \in J$ too. □

Lemma 80.9.
We have

$$\Omega\mathfrak{A}/J \simeq \pi(\Omega\mathfrak{A})/\pi(\delta J_0). \tag{80.18}$$

lem__OCAisomppiOA

Proof. On element of $\Omega\mathfrak{A}$ is of the form $[\omega] \sim [\omega + \omega_1 + \delta\omega_2]$ with $\pi(\omega_1) = \pi(\omega_2) = 0$. We define

$$\begin{aligned}\psi\colon \Omega\mathfrak{A}/J &\to \pi(\Omega\mathfrak{A})/\pi(\delta J_0)\\ [\omega] &\mapsto [\pi(\omega)]_B\end{aligned} \tag{80.19}$$

where the class $[\cdots]_B$ is modulo $\pi(\delta J_0)$, in other words when $\pi(\omega) = 0$, we have $[A]_B = [A+\pi(\delta\omega)]$. The map ψ is well defined because, when $\pi(\omega_1) = \pi(\omega_2) = 0$,

$$\psi[\omega + \omega_1 + \delta\omega_2] = [\pi(\omega + \omega_1 + \delta\omega_2)]_B = [\pi(\omega)]_B + [\pi(\delta\omega_2)]_B = [\pi(\omega)] = \psi[\omega].$$

First suppose that $\psi[\omega] = 0$, i.e $\psi[\omega] = [\pi(\delta\eta)]_B$ with $\pi(\eta) = 0$. Then $\omega = \delta\eta + \sigma$ with $\sigma = 0$. This is the definition of $[\omega] = 0$. This proves that ψ is injective. For surjectivity, consider $[\omega]_B \in \pi(\Omega\mathfrak{A})/\pi(\delta J_0)$ and $\pi(\omega)$ a representative in $\pi(\Omega\mathfrak{A})$. For this ω, we have $\psi[\omega] = [\omega]_B$. □

80.1.2.2 Grading the Connes differential forms

LemOmpmdDp

Lemma 80.10.
The isomorphism of lemma 80.9 induces a grading

$$\Omega^p_D\mathfrak{A} \simeq \Omega^p\mathfrak{A}/J^p, \tag{80.20}$$

eq_OpDAsimeqOpA

and the differential

$$\begin{aligned} d\colon \Omega^p_D\mathfrak{A} &\to \Omega^{p+1}_D\mathfrak{A} \\ [\omega] &\mapsto [\delta\omega] \end{aligned} \tag{80.21}$$

is well defined. lem_isomgraODA

Proof. The well definiteness of the differential is because when $\pi(\omega_1) = \pi(\omega_2)$,

$$d[\omega + \omega_1 + \delta\omega_2] = [\delta\omega + \delta\omega_1 + \delta^2\omega_2] = [\delta\omega].$$

The isomorphism (80.20) is given by $\psi[\omega^p] = [\pi(\omega^p) + \pi(\delta\eta)]$ with $\pi(\eta) = 0$ when $[\omega^p] \in \Omega^p\mathfrak{A}/J^p$, i.e. when $\omega^p \in \Omega^p\mathfrak{A}$. □

80.1.2.3 0-forms

We have $J^0 = J \cap \Omega^0\mathfrak{A} = J \cap \mathfrak{A}$ and $J^0 = \{a \in \mathfrak{A} \text{ tel que } \pi(a) = 0\}$. As operators on $\mathscr{H}$, $J^0 = \{0\}$. Therefore $\Omega^0_D\mathfrak{A} = \Omega^0\mathfrak{A} = \mathfrak{A}$.

Let us now briefly study the spaces of low degree forms.

80.1.2.4 1-forms

From the isomorphism of lemma 80.10, we begin to study δJ^0_0

$$J^0_0 = \{\omega \in \mathfrak{A} \text{ tel que } \pi(\omega) = 0\}.$$

The crucial point is that $\pi(a) = 0$ implies $a = 0$ when $a \in \mathfrak{A}$ (is is not true for any $\omega \in \Omega\mathfrak{A}$!), so

$$\Omega^1\mathfrak{A} = \pi(\Omega^1\mathfrak{A}) \tag{80.22}$$

and Connes 1-forms are of the form

$$\omega_1 = \sum_j a^j_0[D, a^j_1]$$

with $a^j_i \in \mathfrak{A}$.

80.1.2.5 Example on the canonical triple

Now we will use the γ defined in subsection 68.4. When we consider the canonical triple $(\mathfrak{A}, \mathscr{H}, D)$ on a manifold M, $\mathfrak{A} \subset \mathrm{Fun}(M)$. We know that $\mathfrak{A}$ acts on $\mathscr{H}$ by $(f\psi)(x) = f(x)\psi(x)$, so that $[D, f]\psi = (\gamma^\mu\partial_\mu f)\psi$. So we say that

$$[D, f] = \gamma^\mu\partial_\mu f = \gamma(df).$$

The Dirac operator act on functions as follows (see equation (68.99)):

$$(Df)(x) = g_{\alpha\beta}(x)\gamma^\beta_x(e_\alpha \cdot f),$$

this definition is intended to get a Leibniz rule for $D(f\psi)$. We have:

$$[D, f]\psi(x) = D(f\psi)(x) - f(x)D\psi(x) = (Df)(x)\psi(x),$$

so $[D,f]\psi = (Df)\psi$ and as operator on $\mathscr{H}$, $[D,f]$ is the multiplicative operator by Df. When we consider a local orthonormal basis e_α, we have

$$(Df)(x) = g_{\alpha\beta}(x)\gamma_x^\beta(e_\alpha \cdot f) = \gamma^\mu \partial_\mu f(x).$$

From all that we conclude that

$$[D,f] = \gamma^\mu \partial_\mu f,$$

and therefore that, on the canonical triple,

$$\pi(\delta f) = [D,f] = \gamma^\mu \partial_\mu f. \tag{80.23}$$

But $\gamma(df) = \partial_\mu f \gamma(dx^\mu) = \gamma^\mu \partial_\mu f = \pi(\delta f)$. We will soon define the differential d of $\Omega_D\mathfrak{A}$, so from now we denote by d_M the usual differential of M and we write $\pi(\delta f) = \gamma(d_M f)$ and finally,

$$\pi(f_0 \delta f_1 \dots \delta f_p) = f_0 \gamma(d_M f_1) \dots \gamma(d_M f_p). \tag{80.24}$$

Note that $d_M \in \Gamma(M, \mathrm{Cl}(M))$ and, since γ is a morphism,

$$\pi(f_0 \delta f_1 \dots \delta f_p) = f_0 \gamma(d_M f_1 \cdot \dots \cdot d_M f_p) \tag{80.25}$$

Eqpildotsdeltagamd M

where $\cdot$ denotes the Clifford product.

80.1.2.6 Differential 0-forms

Lemma 80.11.

$$\gamma(\Lambda^1(M)) \simeq \Omega^1_D\mathfrak{A}$$

Proof. A general 1-form has the form $\sum_j f_0^j d_M f_1^j$. Since $\gamma\colon \Gamma(M, \mathrm{Cl}(M)) \to \mathfrak{B}(\mathscr{H})$, we claim that the isomorphism is given by

$$\psi\big(f_0[D,f_1]\big) = \gamma\big(f_0 d_M f_1\big).$$

Surjectivity poses no problems because $f_0 d_M f_1$ is the general form of an element of $\Lambda^1(M)$. Now suppose that $\psi(f_0[D,f_1]) = \gamma(f_0 d_M f_1) = 0$. From linearity of γ,

$$f_0 \gamma^\mu \partial_\mu f_1 = 0.$$

At each point, either $f_0 = 0$ or $[D,f_1] = 0$, so globally $f_0[D,f_1] = 0$. □

80.1.2.7 Differential 1-forms

Let $f \in \mathfrak{A}$ and $\alpha = \frac{1}{2}\big(f\delta f - (\delta f)f\big)$. We have $(f\delta f)(x,y,z) = f(x)(\delta f)(y,z) = f(x)\big(f(y) - f(z)\big)$, while

$$(\delta f)f(x,y,z) = \big(f(x) - f(y)\big)f(z).$$

¡¡ Avertissement/question à la lectrice !! 80.12
When we will speak about two points spaces, we will see that $(f\delta g)(x,y) = f(x)\big(g(y) - g(x)\big)$, and more or less the same for $(\delta f)g$. Thus what is done here is wrong and the correct result is

$$2\alpha(x,y) = 2f(x)f(y) - f(x)^2 - f(y)^2.$$

This does not change the conclusion, but it asks for a precise definition of $f\delta g$. Notice that, f being a zero-form, and δg a 1-form, the product should be a 1-form, and not a 2-form.

This proves that $\alpha \neq 0$. The following computation uses the fact that the Leibniz rule for δ is graded

$$\delta\alpha = \frac{1}{2}\big(\delta f \delta f + f\delta^2 f - (\delta^2 f)f + \delta f \delta f\big) = \delta f \delta f,$$

so, on the one hand,

$$\pi(\delta f) = \gamma^\mu \partial_\mu f \gamma^\nu \partial_\nu f = \frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)\partial_\mu f \partial_\nu f = -g^{\mu\nu}\partial_\mu f \partial_\nu f \mathbb{1}_{2^{[N/2]}}$$

where $\mathbb{1}_{2^{[N/2]}}$ is the unit in the Clifford algebra of $\mathbb{R}^n$. On the other hand,

$$\pi(\alpha) = \frac{1}{2}\big(f\gamma^\mu\partial_\mu f - (\gamma^\mu\partial_\mu f)f\big) = 0.$$

This proves that α is a junk 1-form.

¡¡ Avertissement/question à la lectrice !! 80.13
It is also said that this is the general form of a junk, but I didn't succeed to prove it.

80.1.2.8 Differential 2-forms

Let now take $f_1,\ f_2 \in \mathfrak{A}$ and look at

$$\begin{aligned}\gamma\big(d_M f_1 \cdot d_M f_2\big) &= \gamma^\mu\gamma^\nu\partial_\mu f_{1j\partial_\nu} f_2 \\ &= \frac{1}{2}\big(\gamma^\mu\gamma^\nu + \gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu + \gamma^\nu\gamma^\mu\big)\partial_\mu f_1 \partial_\nu f_2 \\ &= \frac{1}{2}\big(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\big)\partial_\mu f_1\partial_\nu f_2 + \frac{1}{2}\big(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\big)\partial_\mu f_1\partial_\nu f_2.\end{aligned}$$

The first term is

$$-g^{\mu\nu}\mathbb{1}\partial_\mu f_1\partial_\nu f_2 = -g(dx^\mu, dx^\nu)\partial_\mu f_1\partial_\nu f_2\mathbb{1} = -g(d_M f_1, d_M f_2)\mathbb{1}.$$

For the second term, first recall that

$$d_M f_1 \wedge d_M f_2 = d_M f_1 \otimes d_M f_2 - d_M f_2 \otimes d_M f_1$$

where the $\otimes$ is, up to equivalence class, the product in Clifford. So the first term is

$$\frac{1}{2}\gamma(d_M f_1 \cdot d_M f_2 - d_M f_2 \cdot d_M f_1) = \gamma(d_M f_1 \wedge d_M f_2).$$

Finally we have

$$\gamma(d_M f_1 \cdot d_M f_2) = -g(d_M f_1, d_M f_2)\mathbb{1} + \gamma(d_M f_1 \wedge d_M f_2). \tag{80.26}$$

EqGamgGam

On the other hand, a general element of $\wedge^2(M)$ (the skew-symmetric differential 2-forms on M) is

$$\sum_j f_0^j\, d_M f_1^j \wedge f_2^j$$

with $f_0^j,\ f_1^j,\ f_2^j \in \mathfrak{A}$.

Lemma 80.14.

$$\Omega^2_D\mathfrak{A} \simeq \gamma\big(\wedge^2(M)\big)$$

Proof. We use the isomorphism

$$\Omega^2_D\mathfrak{A} \simeq \pi\big(\Omega^2\mathfrak{A}\big)/\pi\big(\delta(J_0 \cap \Omega^1\mathfrak{A})\big).$$

where the elements of $J_0 \cap \Omega^1\mathfrak{A}$ are of the form $\alpha_f = \frac{1}{2}\big(f\delta f - (\delta f)f\big)$. A general element of $\Omega^2_D\mathfrak{A}$ is a class of (sum of)

$$\pi(f_0\delta f_1\delta f_2),$$

so from equation (80.25), the idea is to define the candidate isomorphism ψ by

$$\psi\Big(\big[\gamma(f_0 d_M f_1 \cdot d_M f_2)\big]\Big) = \gamma\big(d_M f_1 \wedge d_M f_2\big), \tag{80.27}$$

and its linear extension. Let us compute $\psi[0]$ or $\psi[\pi\delta(\alpha_f)]$. We have $\delta\alpha_f = \delta f \delta f$, so

$$\pi\delta(\alpha_f) = \gamma(d_M f \cdot d_M f).$$

Thus

$$\psi\big([\delta\pi\alpha_f]\big) = \psi\Big(\big[\gamma(d_M f \cdot d_M f)\big]\Big) = \gamma(d_M f \wedge d_M f) = 0.$$

We conclude that ψ is well defined and injective. Surjectivity is clear. □

One can also prove the following generalization.

Lemma 80.15.

$$\Omega^p_D\mathfrak{A} \simeq \wedge^p(M). \tag{80.28}$$

Proof. No proof. □

80.1.3 Example: two points space

SubSecTripleDeuxPoints

Let $Y = \{1, 2\}$, a space containing only two points. The space of continuous functions is $\mathfrak{A} = \mathbb{C} \oplus \mathbb{C}$ and a continuous function is of the form $f = (f_1, f_2)$ with $f_i = f(i) \in \mathbb{C}$. We can build an even spectral triple of dimension zero $(\mathfrak{A}, \mathscr{H}, D, \Gamma)$ as follows. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be two finite dimensional Hilbert space and $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$. We define the action of $f \in \mathfrak{A}$ on $\psi \in \mathscr{H}$ by

$$f\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}f_1\psi_1\\ f_2\psi_2\end{pmatrix}$$

if $\psi_i \in \mathscr{H}_i$. This operator f is clearly bounded on $\mathscr{H}$. Let $M\colon \mathscr{H}_1 \to \mathscr{H}_2$ be a linear map and define the operator D as

$$D = \begin{pmatrix}0 & M^*\\ M & 0\end{pmatrix}.$$

We want $[D, f]$ to be bounded, but

$$[D, f]\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = D\begin{pmatrix}f_1\psi_1\\ f_2\psi_2\end{pmatrix} - \begin{pmatrix}f_1(D\psi_1)\\ f_2(D\psi)_2\end{pmatrix}.$$

The component $(D\psi_1)_1$ does not affect the commutator; it is the reason why we had chosen an anti-diagonal operator D.

As parity map $\Gamma\colon \mathscr{H} \to \mathscr{H}$, we choose

$$\Gamma = \begin{pmatrix}\mathrm{Id}\,|_{\mathscr{H}_1} & \\ & -\mathrm{Id}\,|_{\mathscr{H}_2}\end{pmatrix}.$$

Now consider $f \in \mathfrak{A}$ and compute the commutator

$$\begin{aligned}[D, f]\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} &= \begin{pmatrix}f_2M^*\psi_2\\ f_1M\psi_1\end{pmatrix} - \begin{pmatrix}f_1M^*\psi_2\\ f_2M\psi_1\end{pmatrix}\\ &= (f_2 - f_1)\begin{pmatrix}M^*\psi_2\\ -M\psi_1\end{pmatrix}\\ &= (f_1 - f_2)\begin{pmatrix}0 & -M^*\\ M & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}.\end{aligned}$$

So

$$\|[D, f]\| = |f_1 - f_2|\lambda$$

where λ is the larger eigenvalue of $\sqrt{MM^*}$. Hence the noncommutative distance between 1 and 2 is

$$d(1, 2) = \sup\{|f_1 - f_2| \text{ tel que } \|[D, f]\| \leqslant 1\} = \frac{1}{\lambda}.$$

As real structure, one can take

$$J\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}\overline{\psi_2}\\ \overline{\psi_1}\end{pmatrix}.$$

Let $Y = \{1, 2\}$, its triple $(\mathfrak{A}, \mathscr{H}, D)$ with $\mathfrak{A} = \mathbb{C} \oplus \mathbb{C}$. We are going to study $\Omega^1\mathfrak{A}$. We define δf as being the map

$$(\delta f)(x, y) = f(x) - f(y). \tag{80.29}$$

The space $\Omega^1\mathfrak{A}$ is a left $\mathfrak{A}$-module by the definition

$$(f\delta g)(x, y) = f(x)\delta g(x, y). \tag{80.30a}$$

Now the Leibniz rule imposes the following structure of right module:

$$(\delta f)g(x, y) = (\delta f)(x, y)g(y). \tag{80.30b}$$

Indeed $(\delta f)g = \delta(fg) - f\delta g$, so that

$$\begin{aligned}(\delta f)g(x, y) &= \delta(fg)(x, y) - f(x)(\delta g)(x, y)\\ &= f(x)g(x) - f(y)g(y) - f(x)g(x) + f(x)g(y)\\ &= \big(f(x) - f(y)\big)g(y)\\ &= (\delta f)(x, y)g(y).\end{aligned}$$

¡¡ Avertissement/question au lecteur !! 80.16
I do not understand why things are like that, but if we look at an usual 1*-form on* $\mathbb{R}^N$*, we need two vectors in order to get a number. For the* 1*-form* ω *we need* x *and* X *in order to get* $\omega_x(X) \in \mathbb{R}$*. As far as the multiplication by a function is concerned we write*

$$(f\omega)_x(X) = f(x)\omega_x(X).$$

Thus we have something like $(f\omega)(x, X) = f(x)\omega(x, X)$*. This is more or less the philosophy of* (80.30a)*. For* (80.30b)*, the fact that* g *takes the* y *instead of the* x *is difficult to understand.*

It gives

$$(f\delta g)(x, y) = f(x)\big(g(y) - g(x)\big) \tag{80.31a}$$
$$(\delta f)g(x, y) = \big(f(y) - f(x)\big)g(y), \tag{80.31b}$$

and thus

$$\begin{aligned}\delta(fg)(x, y) &= (\delta f)g(x, y) + f(\delta g)(x, y)\\ &= f(y)g(y) - f(x)g(x)\\ &= (fg)(y) - (fg)(x),\end{aligned}$$

which is coherent.

The 1-forms are functions of two variables which are zero on the diagonal. In the case of our two point space, they takes non zero values only at $(1, 2)$ and $(2, 1)$, so a basis of $\Omega^1\mathfrak{A}$ is given by ω and η with

$$\omega(1, 2) = 1 \qquad \eta(1, 2) = 0 \tag{80.32}$$
$$\omega(2, 1) = 0 \qquad \eta(2, 1) = 1. \tag{80.33}$$

Such a basis can be constructed by defining $e(0) = 0$, $e(1) = 1$ and considering $e\delta e$ and $(1-e)\delta(1-e)$.

80.1.4 Example: manifold

Let M be a compact spin Riemannian manifold and $\mathfrak{A}$ the algebra of (continuous or more) functions on M. We also consider D, the Dirac operator on $\mathscr{H}$, the space of the square integrable spinors over M. The algebra $\mathfrak{A}$ acts on $\mathscr{H}$ by multiplication.

Let C be the vector bundle over M whose fibre are given by $C_x = \mathrm{Cl}^{\mathbb{C}}(T_x^*M)$. It is possible to define the notion of bounded measurable section of C. Let $\rho\colon M \to C$ one of them.

¡¡ Avertissement/question à la lectrice !! 80.17
What is a bounded measurable section of C?

Since $f \in \mathfrak{A}$ is a function on M, the element df is a section of T^*M and can, up to the quotient (68.5), be seen as a section of C. When df is seen in this way, it is denoted by $d_c f$ and we have

$$\pi(f^0 df^1) = f^0[D, f^1] = i^{-1}\gamma(f^0 d_c f^1).$$

¡¡ Avertissement/question à la lectrice !! 80.18
It is true that

$$[D, f] = c(df)$$

when c is a Clifford action on a vector bundle. I should try to understand it better.

80.2 Fredholm modules

Most of theory here and related topics is taken from [953, 819].

If X and Y are Banach spaces, an operator $T\colon X \to Y$ is a **Fredholm operator** if there exists a bounded linear operator $S\colon Y \to X$ such that the operators

$$\mathrm{Id}_X - ST \tag{80.34a}$$
$$\mathrm{Id}_X - TS \tag{80.34b}$$

are compact.

80.2.1 Introductory example

Let M be a compact manifold and $\mathfrak{A} = C(M)$ the C^*-algebra of continuous functions on M. We consider $E^{\pm}$, two Hermitian complex vector bundles on M and an elliptic pseudo-differential operator of order 0, $P\colon C^\infty(M, E^+) \to C^\infty(M, E^-)$. Such an operator can be extended to an operator

$$P\colon L^2(M, E^+) \to L^2(M, E^-) \tag{80.35}$$

which has a parametrix Q. Consequently, P is a Fredholm operator (in fact, wikipedia says that all elliptic operators can be extended to Fredholm operator.)

The algebra $C(M)$ is naturally represented on $L^2(M, E^{\pm})$ by $\pi^{\pm}(f)(x)\xi = f(x)\xi(x)$ (pointwise multiplication of ξ by f). Let us now consider the Hilbert space

$$\mathscr{H} = \mathscr{H}^+ \oplus \mathscr{H}^- = L^2(M, E^+) \oplus L^2(M, E^-) \tag{80.36}$$

and its $\mathbb{Z}_2$-graduation

$$\gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{80.37}$$

We represent $C(M)$ on $\mathscr{H}$ by

$$\pi(f) = \begin{pmatrix} \pi^+(f) & 0 \\ 0 & \pi^-(f) \end{pmatrix}, \tag{80.38}$$

and we pose

$$F = \begin{pmatrix} 0 & Q \\ P & 0 \end{pmatrix}. \tag{80.39}$$

In this case, the operators $[F,\pi(f)]$ and $F^2-\mathbb{1}$ are compact for every $f\in C^\infty(M)$ because the operators

$$\pi^- P - P\pi^+ \tag{80.40a}$$
$$\pi^+ Q - Q\pi^- \tag{80.40b}$$

are compact [1].

80.2.2 Definition

Let $\mathfrak{A}$ be an involutive algebra over $\mathbb{C}$. A **odd Fredholm module** over $\mathfrak{A}$ is

(1) an involutive representation π of $\mathfrak{A}$ on an Hilbert space $\mathscr{H}$,

(2) an operator $F\colon \mathscr{H}\to\mathscr{H}$ such which satisfies

— $F=F^*$, $F^2=\mathbb{1}$,

— $[F,\pi(a)]$ is a compact operator for each $a\in\mathfrak{A}$

An **even Fredholm modules** is an odd Fredholm module with a $\mathbb{Z}/2$ grading $\gamma\colon\mathscr{H}\to\mathscr{H}$ such that

— $\gamma=\gamma^*$, $\gamma^2=\mathbb{1}$,

— $[\gamma,\pi(a)]=0$ for all $a\in\mathfrak{A}$,

— $\gamma F=-F\gamma$.

We will almost always write $a\xi$ instead of $\pi(a)\xi$ when the underlying representation is clear. The Fredholm module $(\mathscr{H},F)$ is said to be **p-summable** when $\forall a\in\mathfrak{A}$,

$$[F,a]\in\mathscr{L}^p(\mathscr{H}).$$

One says that the Fredholm module $(\mathscr{H},F)$ is θ-summable when $[F,a]\in J^{1/2}$ for all $a\in\mathfrak{A}$. The set $J^{1/2}$ is the two-sided ideal of compact operators T such that

$$\mu_n(T)=O\big((\ln n)^{-1/2}\big).$$

80.2.3 Cycle associated with Fredholm module

Let $(\mathscr{H},F)$ be a Fredholm module. We are going to build a cycle in the sense of section 54.12 associated with $(\mathscr{H},F)$. For the graded algebra $\Omega=\oplus_k\Omega^k$, we begin by $\Omega^0=\mathfrak{A}$ and for $k>0$, we define Ω^k as the vector space spanned by operators of the form

$$a^0[F,a^1]\cdots[F,a^k]$$

with $a^j\in\mathfrak{A}$. Except from a^0, this is a product of k elements of $\mathscr{L}^{n+1}$ which belongs to $\mathscr{L}^{(n+1)/k}$ by equation (60.23). The fact that $\mathscr{L}^q$ is an ideal makes that

$$\Omega^k\subset\mathscr{L}^{(n+1)/k}(\mathscr{H}).$$

The product in Ω is defined as the usual operator product.

Lemma 80.19.
If $\omega\in\Omega^k$ and $\omega'\in\Omega^{k'}$, $\omega\omega'\in\Omega^{k+k'}$.

Proof. The fact that $[F,.]$ is a derivation on $\mathfrak{A}$ makes that

$$\begin{aligned}a^0[F,a^1]\cdots[F,a^k]a^{k+1}=&\sum_{j=1}^{k}(-1)^{k-j}a^0[F,a^1]\cdots[F,a^ja^{j+1}]\cdots[F,a^{k+1}]\\&+(-1)^k a^0a^1[F,a^2]\cdots[F,a^{k+1}].\end{aligned}$$

1. I do not know why.

This is the same computation as in equation (80.5). Since each term of this sum has the form $r^0[F,r^1]\cdots[F,r^k]$, thus the product $\omega\omega'$ reads

$$\underbrace{a^0[F,a^1]\cdots[F,a^k]b^0}_{\text{Sum of } a^0[F,r^1]\cdots[F,r^k]}[F,b^1]\cdots[F,b^k]$$

which belongs to $\Omega^{k+k'}$. □

From here we have a graded algebra Ω^* with a product $\Omega^j\times\Omega^{k'}\to\Omega^{k+k'}$. As differential, we choose

$$d\omega = F\omega - (-1)^k\omega F = [F,a^0][F,a^1]\cdots[F,a^k]. \tag{80.41}$$

The second equality can be checked by virtue of $F[F,a] = -[F,a]F$. This differential is a **graded differential**, i.e.

$$d(\omega_1\omega_2) = (d\omega_1)\omega_2 + (-1)^{k_1}\omega_1 d\omega_2 \tag{80.42}$$

for all $\omega_1\in\Omega^{k_1}$. Indeed,

$$\begin{aligned} d(\omega_1\omega_2) &= F\omega_1\omega_2 - (-1)^{k_1+k_2}\omega_1\omega_2 F\\ &= F\omega_1\omega_2 - (-1)^{k_1+k_2}\omega_1 b^0[F,b^1]\cdots[F,b^{k_2}]F\\ &= F\omega_1\omega_2 - (-1)^{k_1}\omega_1\big(-[F,b^0]+Fb^0\big)[F,b^1]\cdots[F,b^{k_2}]\\ &= F\omega_1\omega_2 + (-1)^{k_1}\omega_1 d\omega_2 - (-1)^{k_1}\omega_1 F\omega_2\\ &= (d\omega_1)\omega_2 + (-1)^{k_1}\omega_1 d\omega_2. \end{aligned}$$

We also check that $d^2=0$ in the following way:

$$\begin{aligned} d^2\omega &= d(F\omega - (-1)^k\omega F)\\ &= F\big(F\omega - (-1)^k\omega F\big) + (-1)^k\big(F\omega - (-1)^k\omega F\big)\\ &= \omega - (-1)^k F\omega F + (-1)^k F\omega F - \omega\\ &= 0 \end{aligned} \tag{80.43}$$

where we used the fact that $F^2=\mathbb{1}$.

The pair (Ω^*,d) is a graded differential algebra. We have to find a graded closed trace $\mathrm{Tr}_s\colon\Omega^n\to\mathbb{C}$. Let T be an operator on $\mathscr{H}$ such that $FT+TF\in\mathscr{L}^1(\mathscr{H})$. We begin to define

$$\mathrm{Tr}'(T) = \frac{1}{2}\,\mathrm{Tr}\big(F(FT+TF)\big).$$

When $T\in\mathscr{L}^1$, it makes sense to distribute the F in the trace, in such a way that we obtain $\mathrm{Tr}'(T)=\mathrm{Tr}(T)$. In this case, we have $\mathrm{Tr}'(T) = \frac{1}{2}\,\mathrm{Tr}\big(F(FT+TF)\big) = \frac{1}{2}\,\mathrm{Tr}(T+FTF) = \mathrm{Tr}(T)$. Now de define the trace $\mathrm{Tr}_s\colon\Omega^n\to\mathbb{C}$,

$$\mathrm{Tr}_s\,\omega = \begin{cases}\mathrm{Tr}'(\omega) & \text{if } n \text{ is odd}\\ \mathrm{Tr}'(\gamma\omega) & \text{if } n\text{is even.}\end{cases}$$

It makes sense because when $\omega\in\Omega^n$, it fulfills $F\omega+\omega F = d\omega\in\Omega^{n+1}\subset\mathscr{L}^{(n+1)/(n+1)}=\mathscr{L}^1$, so that we can use the usual trace. By the way, remark that the trace Tr_s reads $\frac{1}{2}\,\mathrm{Tr}(Fd\omega)$ when n is odd and $\frac{1}{2}\,\mathrm{Tr}(F\gamma d\omega)$ when n is even.

Proposition 80.20.
The triple (Ω,d,Tr_s) is a n-dimensional cycle over $\mathfrak{A}$ (see definition 54.60).

Proof. Most of the work is already done; it just remains to prove that Tr_s is a graded closed trace. First, we know that $d^2=0$, so the fact that Tr_s only depends to $d\omega$ gives $\mathrm{Tr}_s(d\omega)=0$. The form is thus closed.

Now we have to prove that it is a graded trace. If $\omega \in \Omega^k$ and $\omega' \in \Omega^{k'}$ with $k + k' = n$ (let us assume n odd), we have

$$\mathrm{Tr}_s(\omega\omega') = \frac{1}{2}\,\mathrm{Tr}\left(F(d\omega)\omega' + (-1)^k F\omega d\omega'\right) = \frac{1}{2}\,\mathrm{Tr}\left((-1)^{k+1}(d\omega)F\omega' + (-1)^k(F\omega)d\omega'\right),$$

but the usual trace has the property that, when $T_j \in \mathscr{L}^j$ with $\frac{1}{p_1} + \frac{1}{p_2} = 1$, $\mathrm{Tr}(T_1T_2) = \mathrm{Tr}(T_2T_1)$. So we have $\mathrm{Tr}(F\omega d\omega') = \mathrm{Tr}(d\omega' F\omega)$ because $F\omega \in \mathscr{L}^{(n+1)/k}$ and $d\omega' \in \Omega^{k'+1} \subset \mathscr{L}^{(n+1)/(k'+1)}$. Thus we have

$$\mathrm{Tr}_s(\omega\omega') = \frac{1}{2}\big((-1)^{k+1}d\omega F\omega' + (-1)^k d\omega' F\omega\big).$$

This expression is symmetric or anti-symmetric with repsect to the inversion $\omega \leftrightarrow \omega'$ following that $(-1)^{kk'}$ equals 1 or -1. We conclude that (at least when n is odd) t

$$\mathrm{Tr}_s(\omega\omega') = (-1)^{kk'}\,\mathrm{Tr}_s(\omega'\omega).$$

□

The **character** is the cyclic cocycle

$$\tau_n(a^0, \cdots, a^n) = \begin{cases} \mathrm{Tr}'\left(a^0[F, a^1]\cdots[F, a^n]\right) & \text{if } n\text{is odd} \\ \mathrm{Tr}'\left(\gamma a^0[F, a^1]\cdots[F, a^n]\right) & \text{if } n\text{is aven.} \end{cases} \tag{80.44}$$

80.3 Hochschild cohomology

Proposition 80.21.
Let $(\mathscr{H}, F)$ be a Fredholm module $(n+1)$-summable over $\mathfrak{A}$. We suppose that this module has the same parity as n. Then the characters τ_{n+2q} satisfy

$$\tau_{m+2} = -\frac{2}{m+2} S\tau_m \in HC^{m+2}(\mathfrak{A}) \tag{80.45}$$

when $m = n + 2q$, with $q \geqslant 0$.

Proof. In order to see that $\tau \in C^n_\lambda(\mathfrak{A})$, we have to check that the equality $\tau_n(a^1, \cdots, a^n, a^0) = (-1)^n\tau_n(a^0, \cdots, a^n)$ holds for all $a^i \in \mathfrak{A}$. We have

$$\tau_n(a^0, \cdots, a^n) = \mathrm{Tr}'(a^0 da^1 \cdots da^n) = \mathrm{Tr}'\left(\big(d(a^0a^1) - da^0a^1\big)da^2 \cdots da^n\right)$$

in which the first term vanishes because $d^2 = 0$. We go on commuting da^0 and we finally get $(-1)^n\tau_n(a^1, \cdots, a^n, a^0)$.

¡¡ Avertissement/question au lecteur !! 80.22
This proof is not finished.

□

80.4 Fredholm module and conformal structure

Let V be a compact, oriented even dimensional manifold endowed with a **conformal structure**. That is an equivalence class of metrics where $g \sim h$ if and only if there exists a positive smooth function λ such that $g = \lambda^2 h$.

Let $\mathscr{H}_0 = L^2\big(V, \bigwedge^n_{\mathbb{C}}(T^*V)\big)$ with the product (56.387). This becomes a $C^\infty(V)$-module when we define

$$(f\omega)(p) = f(p)\omega(p) \tag{80.46}$$

for every $\omega \in \mathscr{H}_0$, $f \in C^\infty(V)$ and $p \in V$. We can extend the graduation (56.383) to the Hilbert space $\mathscr{H}_0$ by

$$(\gamma\omega)(p) = \gamma\big(\omega(p)\big). \tag{80.47}$$

Consider $\mathscr{H}_0 = L^2\big(V, \bigwedge^n_{\mathbb{C}}(T^*V)\big)$ be the space of square integrable sections of the bundle $\bigwedge^n_{\mathbb{C}}(T^*V)$ for the product

We can consider the complex space $\wedge^n_{\mathbb{C}}E$ and the operator $\gamma\colon \bigwedge^n E \to \bigwedge^n E$,

$$\gamma = (-1)^{n(n-1)/2} i^n * . \tag{80.48}$$

This operator squares to $\mathbb{1}$, so that it creates a $\mathbb{Z}/2$-graduation of $\bigwedge^n_{\mathbb{C}} E$.

We have

$$\frac{1+\gamma}{2} d(*\alpha^{n+1}) = \frac{1}{2}\Big(d(*\alpha^{n+1}) + (-1)^s \delta\alpha^{n+1}\Big) \tag{80.49}$$

where s is a sign. We used the fact that $\delta = - * d *$. Thus the elements of $\mathscr{H}_0$ of the form $\frac{1+\gamma}{2} d\alpha$ are orthogonal to the harmonic forms.

80.5 Fredholm modules and K-cycles

Definition 80.23.
*A **K-cycle** $(\mathscr{H}, D)$ over an involutive algebra $(\mathfrak{A}, *)$ is*

(1) a $*$-representation of $\mathfrak{A}$ on $\mathscr{H}$,

(2) a selfadjoint non bounded operator D with compact resolvent and such that $[D, a]$ is bounded for each $a \in \mathfrak{A}$.

Remarque 80.24.
The condition "compact resolvent" means that the operators $(D - \lambda\mathbb{1})^{-1}$ are compact for every λ in $\mathbb{C}\backslash \mathrm{Spec}(D)$ (lemma 60.71). In particular, the kernel of D is finite dimensional (corollary 60.70).

From a K-cycle on $\mathfrak{A}$ we canonically build a Fredholm module $(\mathscr{H}', F)$, the Fredholm module **associated** with the K-cycle $(\mathscr{H}, D)$, in the following way.

(1) $\mathscr{H}' = \mathscr{H} \oplus \ker(D) = \ker(D)^\perp \oplus \ker(D) \oplus \ker(D)$;

(2) $a(\xi, \eta) = (a\xi, 0)$ for every $\xi \in \mathscr{H}$ and $\eta \in \ker(D)$;

(3) $F = \mathrm{Sign}(D) \oplus F_1$.

The definition of F deserves some comments. First, $\mathrm{Sign}(D)$ is the sign of D, that is the partial isometry in the polar decomposition $D = V|D|$ of D. That acts on $\ker(D)^\perp$. The operator F_1 is the operator which exchanges the two copies of $\ker(D)$. More explicitly, if $\xi \in \mathscr{H}$ and $\eta \in \ker(D)$,

$$F(\xi, \eta) = F(\xi_1 + \xi_0, \eta) = (V\xi_1 + \eta, \xi_0) \tag{80.50}$$

where $\xi = \xi_1 + \xi_0$ is the decomposition of ξ with respect to $\mathscr{H} = \ker(D) \oplus \ker(D)^\perp$. In order to prove that $F^2 = \mathbb{1}$, we have to show that $V^2 = \mathbb{1}$ and that $V\xi_1 \in \ker(D)^\perp$.

Since the part $\mathrm{Sign}(D)$ only acts on $\ker(D)^\perp$, we can see this operator as in equation (60.48), that is the sign of the operator D restricted to the space $\ker(D)^\perp$. This is an operator on $\ker(D)^\perp$, so that $V\xi_1 \in \ker(D)^\perp$ and $V^2 = \mathrm{Id}_{\ker(D)^\perp}$.

80.6 Spectral triple

80.6.1 General spectral triple

A **spectral triple** is a triple $(\mathfrak{A}, \mathscr{H}, D)$ where

— $\mathscr{H}$ is a Hilbert space,

— $\mathfrak{A}$ is an involutive algebra of bounded operators on $\mathscr{H}$,

— D is a self-adjoint ($D = D^*$) operator on $\mathscr{H}$ such that

(1) the resolvent $(D-\lambda)^{-1}$, $\lambda \notin \mathbb{R}$ is a compact operator on $\mathscr{H}$,

(2) $[D,a] := D \circ a - a \circ D$ is a bounded operator for all a. item_DaDcirii

In general condition (2) can only be imposed on a dense subalgebra of $\mathfrak{A}$.

DefDimSpec

Definition 80.25.
*The **dimension spectrum** of the spectral triple $(\mathfrak{A}, \mathscr{H}, D)$ is the set Π of complex numbers z such that $\mathrm{Re}(z) \geqslant 0$ and z is a singularity of the analytic function $\zeta_b(z)$ for $b \in \mathcal{B}$ with positive real part. Here $\mathcal{B}$ is the operator algebra generated by $\delta^k(a)$ and $\delta^k[D,a]$ with $a \in \mathfrak{A}$ and $\delta T = [|D|, T]$. When $b \in \mathcal{B}$, the function ζ is given by*

$$\zeta_b(z) = \mathrm{Tr}(b|D|^{-z})$$

which is well defined when $\mathrm{Re}(z) > m$ where m is the crude dimension of the triple.

We say that the dimension of the triple is **simple** id the poles of the functions ζ_b are at most simple.

We say that the triple is **even** if there exists an operator Γ on $\mathscr{H}$ such that

(1) $\Gamma = \Gamma^*$,

(2) $\Gamma^2 = 1$,

(3) $[\Gamma, D] = 0$ and $[\Gamma, a] = 0$ for all $a \in \mathfrak{A}$.

If the triple is not even, it is **odd**. The triple $(\mathfrak{A}, \mathscr{H}, D)$ is of **dimension** $n > 0$ if $|D|^{-1}$ is an infinitesimal of order $1/n$, in other words, if $|D| - n$ is infinitesimal of order 1. A n-dimensional spectral triple is sometimes said to be n-summable. A **real structure** on the spectral triple $(\mathfrak{A}, \mathscr{H}, D)$ is an antilinear isometry $J : \mathscr{H} \to \mathscr{H}$ such that

$$\begin{aligned} J^2 &= \epsilon(n)\mathbb{1} & [a, b^0] &= 0 \\ JD &= \epsilon'(n)DJ & \big[[D,a], b^0\big] &= 0 \\ J\Gamma &= (i)^n \Gamma J \end{aligned}$$

where $b^0 = Jb^*J^*$ and Γ is the $\mathbb{Z}_2$ graduation if the triple is even; if the triple is odd, then the corresponding condition is removed. The functions ϵ and ϵ' are periodic with period 8 and

$$\begin{aligned} \epsilon(n) &= (1,1,-1,-1,-1,-1,1,1) \\ \epsilon'(n) &= (1,-1,1,1,1,-1,1,1). \end{aligned}$$

Notice that as direct consequence of the properties, we also have $\big[[D,b^0]\big] = 0$. When we consider a real spectral triple, we can endow $\mathscr{H}$ with a bimodule structure over $\mathfrak{A}$ by

$$a\xi b = \pi(a)J\pi(b^*)J^*\xi.$$

The left module structure is the usual one while the right is well defined because $J^*J = \mathbb{1}$, so that

$$\xi(an) = Jb^*a^*J^*\xi = Jb^*J^*Ja^*J^*\xi = (\xi a)b.$$

Two spectral triples $(\mathfrak{A}_i, \mathscr{H}_i, \pi_i, D_i)$ are **equivalent** when there exists an unitary operator $U : \mathscr{H}_1 \to \mathscr{H}_2$ such that $U\pi_1(a)U^* = \pi_2(a)$ for every $a \in \mathfrak{A}$ and $UD_1U^* = D_2$. If the triple is even or real, we ask moreover $U\Gamma_1U^* = \Gamma_2$ and $UJ + 1U^* = J_2$.

80.6.2 Commutative real triple

When $\mathfrak{A}$ is commutative, the right action of a is equivalent to the left action of Ja^*J^* in the sense that

$$\xi(ab) = (Ja^*J^*)(Jb^*J^*)\xi$$

80.6.3 Analysis on a spectral triple

The following is a direct computation.

Lemma 80.26.
The operator $[D, \ \cdot\ ]$ *is a derivation of* $\mathfrak{A}$.

Lemma 80.27.

$$[D,a]^* = -[D,a] \tag{80.51}$$

Proof. It is nothing else than the fact that $D = D^*$:

$$[D,a]^* = (D\circ a)^* - (a\circ D)^* = a^*D - Da^* = [a,D] = -[D,a].$$

□

From definition of the spectral triple, the operator $(D - z\mathbb{1})$ exists for all non real z, so the spectrum (definition 60.4.1) of D is real:

$$\sigma(D) \subset \mathbb{R}.$$

¡¡ Avertissement/question à la lectrice !! 80.28
The following statements need more theory about operators with compact resolvent.

The spectrum of D is discrete and the elements $\{\lambda_n\}$ are eigenvalues of finite multiplicity. Moreover characteristic values $\mu_n\big((D-1)^{-1}\big) \to 0$ when $n \to \infty$ and so $|\lambda_n| = \mu_n(|D|) \to 0$.

When $[D,a]$ is bounded we say that $a \in \mathfrak{A}$ is **Lipschitz**. Let δ be the derivation on $\mathfrak{B}(\mathscr{H})$ defined (on a dense subspace) by

$$\delta(T) = [|D|, T].$$

This generates a one parameter group of automorphism of $\mathfrak{B}(\mathscr{H})$ defined by

$$\alpha_s(T) = e^{is|D|}Te^{-is|D|}. \tag{80.52}$$

We say that $a \in \mathfrak{A}$ is **smooth** and we write $a \in C^\infty$ if the map

$$s \to \alpha_s(a)$$

is smooth. The element a is of class C^k when $s \to \alpha_s(a)$ is C^k.

Proposition 80.29.
An element $a \in \mathfrak{A}$ *is smooth if and only if* $a \in \bigcap_{n\in\mathbb{N}} \mathrm{dom}(\delta^n)$.

Proof. If a is smooth, the existence of the derivative of $s \to e^{is|D|}ae^{-is|D|}$ makes that $a \in \mathrm{dom}\,\delta$ because the derivative of this map is precisely δ. A few computation shows that the second derivative of this map is δ^2.

If, on the other hand, $a \in \bigcap_{n\in\mathbb{N}} \mathrm{dom}(\delta^n)$, we have existence of all the derivatives of $s \to \alpha_s(a)$ and continuity is given by derivability

¡¡ Avertissement/question à la lectrice !! 80.30
Is that justification correct?

□

80.6.4 Spectral triple over a manifold

Let (M,g) be a Riemannian spin manifold of dimension n. The **canonical triple** on M is

(1) $\mathfrak{A} = C^\infty(M)$,
(2) $\mathscr{H} = L^2(M,S)$, the bundle of square integrable spinors on M.

(3) D is the Dirac operator associated with the Levi-Civita connection of g.

The rank of the spinor bundle is $2^{[n/2]}$ and the scalar product on $\mathscr{H}$ is given by

$$(\psi, \phi) = \int \overline{\psi(x)}\phi(x)\, d\mu(g)$$

where the bar denotes the complex conjugation, $d\mu(g)$ is the measure associated with g and $\overline{\phi(x)}\phi(x)$ is the usual product in $\mathbb{C}^{2[n/2]}$. The space $C^\infty(M)$ is an algebra of operators on $\mathscr{H}$ by simple multiplication:

$$(f\psi)(x) := f(x)p\psi(x) \tag{80.53}$$

for all $f \in C^\infty(M) = \mathfrak{A}$ and $\psi \in \mathscr{H}$.

Most of the interest in spectral triples over manifold comes from the following theorem.

Theorem 80.31.
Let $(\mathfrak{A}, \mathscr{H}, D)$ be the canonical triple on a manifold M. Then

(1) M is the structure space of the algebra $\overline{\mathfrak{A}} = C(M)$, the norm closure of $\mathfrak{A} = C^\infty(M)$,

(2) the geodesic distance between p, $q \in M$ is given by

$$d(p, g) = \sup\{|f(p) - f(q)| \text{ tel que } f \in \mathfrak{A} \text{ and } \|[D, f]\| \leqslant 1\}, \tag{80.54}$$

eq_defdtriple

(3) the Riemannian measure on M is given by

$$\int_M f = c(n)\, \mathrm{Tr}_\omega(f|D|^{-n}) \tag{80.55}$$

where $c(n) = 2^{n-[n/2]-1}\pi^{n/2}n\Gamma(\frac{n}{2})$.

See [819] for more complete proof and reference for even more complete proof.

Proof. The algebra $\overline{\mathfrak{A}}$ is an unital commutative C^*-algebra. So Gelfand theorem 72.15 says that $\overline{\mathfrak{A}} \simeq C\big(\Delta(\overline{\mathfrak{A}})\big)$, but by definition $\overline{\mathfrak{A}} = C(M)$, hence $M = \Delta(\overline{\mathfrak{A}})$. This proves the first point.

For the second point, we use the form (68.123) of Dirac operator, so

$$\begin{aligned}
[D, f]\psi(x) &= D(f\psi)(x) - f(x)D\psi(x) \\
&= \gamma^\mu(x)\big((\partial_\mu f)\psi + f\partial_\mu\psi + f\omega_\mu^S\psi\big) \\
&\quad - f(x)\gamma^\mu(x)\big(\partial_\mu\psi + \omega_\mu^S\psi\big) \\
&= (\gamma^\mu\partial_\mu f)\psi.
\end{aligned} \tag{80.56}$$

Hence, as multiplicative operator on $\mathscr{H}$, we have $[D, f] = \gamma^\mu\partial_\mu f = \gamma(df)$ for all $f \in \mathfrak{A}$. The norm of this operator is

$$\|[D, f]\| = \sup |(\gamma^\mu\partial_\mu f)(\gamma^\nu\partial_\nu f)^*|^{\frac{1}{2}}.$$

One can prove (cf [819] for a reference) that this is the Lipschitz norm of f:

$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d_\gamma(x, y)}$$

where d_γ is the usual geodesic distance that we want to prove to be equals to d. The condition $\|[D, f] \leqslant 1\|$ in the definition (80.54) retrains us to only looks at f such that

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{d_\gamma(x, y)} \leqslant 1.$$

If we fix $x = p$ and $y = q$, this condition becomes $|f(p) - f(q)| \leqslant d_\gamma(p, q)$; hence $d(p, q) \leqslant d_\gamma(p, q)$.

We have to work out the inverse inequality. For, we fix a point q and consider $f_{\gamma,q} = d_\gamma(x, q)$. This function fulfils $\|[D, f_{\gamma,q}]\| \leqslant 1$ and using this function as lower bound for the supremum which defines $d(p, q_0)$, we find

$$d(p, q) \geqslant |f_{\gamma,q}(p) - f_{\gamma,q}(q)| = d_\gamma(p, q)$$

because $f_{\gamma,q} = d_\gamma(p, q)$ and $f_{\gamma,q}(q) = 0$.

□

¡¡ Avertissement/question au lecteur !! 80.32
The use of Gelfand's theorem at the beginning of the proof requires M to be compact?

80.6.5 Distance over general triple

We will now show that formula (80.54) generalises to a distance formula between states, see definition 72.58. The **distance** is the distance on $\mathcal{S}(\overline{\mathfrak{A}})$ given by

$$d(\omega,\chi)=\sup_{a\in\mathfrak{A}}\{|\omega(a)-\chi(a)|\text{ tel que }\|[D,a]\|\leqslant 1\}. \tag{80.57}$$

When $(\mathfrak{A},\mathscr{H},D)$ is a triple of dimension n, we define the **integral** of $a\in\mathfrak{A}$ by

$$\int a:=\frac{1}{V}\operatorname{Tr}_\omega(a|D|^{-1}) \tag{80.58}$$

where V is a constant defined by $\mu_j\leqslant Vj^{-1}$ when $j\to\infty$. Here, (μ_j) is the sequence of characteristic values of $|D|^{-1}$. Why does $|D|^{-n}$ appears? The operator a is just bounded on $\mathscr{H}$, hence the trace $\operatorname{Tr}_\omega a$ makes no sense. The multiplication by $|D|^{-n}$ gives rise to an infinitesimal of order 1 and Dixmier trace makes sense. On the other hand, the integral is normalised in the following sense:

$$\int\mathbb{1}=\frac{1}{V}\operatorname{Tr}_\omega|D|^{-n}=\frac{1}{V}\lim_{N\to\infty}\sum_{j=1}^{N-1}\mu_j\big(|D|^{-n}\big)\leqslant\lim_{N\to\infty}\sum_{j=1}^{N-1}\frac{1}{j}=1$$

because $\frac{1}{V}\mu_j(|D|^{-n})\leqslant j^{-1}$.

¡¡ Avertissement/question au lecteur !! 80.33
This only proves that $\int\mathbb{1}\leqslant 1$, but the second line is probably

$$\frac{1}{V}\lim_{\omega}\sum_{j=1}^{N-1}\mu_j(|D|^{-n}),$$

and we should define $\lim_\omega$ *in such a way that*

$$\frac{1}{V}\lim_{N\to\infty}\sum_{j=1}^{N-1}Vj^{-1}.$$

After that, we still have to define $\lim_{\omega(V)}$ *and prove that* $\int a$ *does not depend on its choice, on the choice of* V *and of* $\lim_{\omega(V)}$.

80.6.6 Real triple

Let $(\mathfrak{A},\mathscr{H},D)$ be a spectral triple of dimension n. A **real structure** is an anti-linear isometry $J\colon\mathscr{H}\to\mathscr{H}$ such that

(1) $J^2=\epsilon(n)\,\mathrm{Id}$,
(2) $JD=\epsilon'(n)DJ$,
(3) $G\Gamma=i^n\Gamma J$ if n is even with the grading Γ,
(4) $[a,b^0]=0$,
(5) $[[D,a],b^0]=0$ where $b^0=Jb^*J^*$.

The functions ϵ and ϵ' are defined modulo 8:

$$\epsilon(n)=1,1,-1,-1,-1,-1,1,1$$
$$\epsilon'(n)=1,-1,1,1,1,-1,1,1.$$

80.6.7 Example: compact manifold

Let M be a compact spin Riemannian manifold. We consider the triple $(\mathfrak{A}, \mathscr{H}, D)$ where

— $\mathscr{H}$ is the Hilbert space of L^2 spinors on M,

— D is the Dirac operator on $\mathscr{H}$,

— $\mathfrak{A}$ is the abelian algebra of bounded measurable functions on M with the multiplicative action on $\mathscr{H}$.

Via the representation, space $C(M)$ of continuous functions on M is seen as a subspace of the space of linear operators on $\mathscr{H}$. Hence, up to a closure, we have $M = \Delta(\mathfrak{A})$ and Gel'fand theorem 72.15 says that the compact topological space structure of M is given by $\mathfrak{A}$. A point of M is associated with an element of $\Delta(\mathfrak{A})$, i.e. a homomorphism $\mathfrak{A} \to \mathbb{C}$.

Proposition 80.34.
Let $a \in \mathfrak{A}$. The operator $[D, a]$ is

(1) defined on a dense subspace of $\mathscr{H}$,

(2) bounded if and only if a is almost everywhere equals to a Lipschitz function.

Proof. No proof. □

A function f on the manifold M is **Lipschitz** if for all p, $q \in M$,

$$|f(p) - f(q)| \leqslant Cd(p, q)$$

where C is a constant and d denotes the geodesic distance on M. Any Lipschitz function is continuous and the space of Lipschitz functions is dense in $C(M)$. Then $C(M)$ is the closure of $\mathfrak{A}$ in $\mathscr{L}(\mathscr{H})$.

80.7 Non commutative vector bundle

80.7.1 The category of complex vector bundles

In order to define the concept of noncommutative vector bundle, we have to get a more precise comprehension of a commutative vector bundle. Let M be a compact manifold and $\mathscr{C}$ be the category of complex vector bundle over M in which the arrows are the vector bundle isomorphisms, i.e. maps of the form $\tau\colon E \to E'$ such that $\pi' \circ \tau = \pi$ and $\tau_x\colon E_x \to E'_x$ being a linear map with obvious notations. One usually denote $\Gamma(E) = C^\infty(M, E)$.

If $\tau\colon E \to E'$ is an isomorphism of vector bundle, we denote by $\Gamma\tau$ the map

$$\begin{aligned} \Gamma\tau\colon \Gamma(E) &\to \Gamma(E) \\ (\Gamma\tau)s &\mapsto= \tau \circ s. \end{aligned} \tag{80.59}$$

Let $\mathfrak{A}$ be a (not yet determined) dense subalgebra of $C^\infty(M)$. For every $a \in \mathfrak{A}$ and $x \in M$, the map $\Gamma\tau$ satisfies $(\Gamma\tau)(sa) = (\Gamma\tau)(s)a$, so that $\Gamma\tau$ is a morphism of right $\mathfrak{A}$-module.

Let $\mathscr{C}$ be the category of vector bundle over M in which the arrows are isomorphisms of vector bundle, and $\mathscr{D}$, the category of right $\mathfrak{A}$-modules. We define the functor Γ from $\mathscr{C}$ to $\mathscr{D}$ which to each vector bundle make correspond the right $\mathfrak{A}$-module of its sections. An arrow $\tau\colon E \to E'$ is transformed into the arrow $(\Gamma\tau)(s) = \tau \circ s$.

That functor respects the tensor product in the sense that $\Gamma(E \otimes E') = \Gamma(E) \otimes_{\mathfrak{A}} \Gamma(E')$ where $\otimes_{\mathfrak{A}}$ denotes the tensor product with the relation $sa \otimes s' = s \otimes as'$ for every $s \in \Gamma(E)$, $s' \in \Gamma(E')$ and $a \in \mathfrak{A}$.

Proposition 80.35.
Any $\mathfrak{A}$-linear map from $\Gamma(E)$ to $\Gamma(E')$ is of the form $\Gamma\tau$ for an unique map $\tau\colon E \to E'$.

Proof. No proof. □

Let us now study the image of the functor Γ: what are the $\mathfrak{A}$-modules of the form $\Gamma(E)$? First, remark that, for the trivial bundle $E = M \times \mathbb{C}^r$, then the image is the free $\mathfrak{A}$-module $\Gamma(E) = \mathfrak{A}^r$. Let us now take a more general vector bundle E. Compactness of M provide us partition of unity $\psi_i : \mathcal{U}_i \to \mathbb{R}$ with $\psi_i > 0$ and $\psi_1^2 + \cdots + \psi_q^2 = 1$. For sake of notational simplicity, we denote $\mathcal{U}_{ij} = \mathcal{U}_i \cap \mathcal{U}_j$. We suppose without loss of generality that E is trivial over each of the $\mathcal{U}_i$. The transition functions $f_{ij} : \mathcal{U}_{ij} \to \mathrm{GL}(r, \mathbb{C})$ are subject to the relations $f_{ik} f_{kj} = f_{ij}$ on $\mathcal{U}_{ijk}$. Let us pose

$$p_{ij} = \begin{cases} \psi_i f_{ij} \psi_j & \text{in } \mathcal{U}_{ij} \\ 0 & \text{outside} \end{cases} \tag{80.60}$$

One easily finds that

$$\sum_k p_{ij} p_{kj} = p_{ij},$$

so that p is an element of $M_{qr}(\mathfrak{A})$ with the property that $p^2 = p$. Now a section of E is given by the local functions $s_j : \mathcal{U}_j \to \mathbb{C}^r$ given by $s_i = f_{ij} s_j$ on $\mathcal{U}_{ij}$ and can be seen as a column

$$s = \begin{pmatrix} \psi_1 s_1 \\ \vdots \\ \psi_q s_q \end{pmatrix} \in C^\infty(M)^{qr}. \tag{80.61}$$

Let us compute ps:

$$\begin{aligned} (ps)_i &= \sum_k p_{ik} s_k = \sum_k p_{ik} \psi_k s_k \\ &= \sum_k \psi_i f_{ik} \psi_k \psi_k s_k = \sum_k \psi_i f_{ik} \psi_k^2 f_{ki} s_i \\ &= \psi_i s_i \sum_k \psi_k^2 = \psi_i s_i = s_i, \end{aligned}$$

so that $ps = s$. Thus we identify $\Gamma(E)$ with $p\mathfrak{A}^{qr}$ with r being the dimension of the vector bundle and q, the number of open sets needed to have a partition of unity in the same time as a trivialization of E. The Serre-Swan theorem provides the result in the inverse sense.

Theorem 80.36 (Serre-Swan)**.**
Every right $\mathfrak{A}$-module of the form $p\mathfrak{A}^m$ with p, an idempotent element of $M_m(\mathfrak{A})$ is of the form

$$\Gamma(E) = C^\infty\big(M(\mathfrak{A}), E\big)$$

where the fibre over $\mu \in M(\mathfrak{A})$ is the vector space $p\mathfrak{A}^m \otimes_{\mathfrak{A}} (\mathfrak{A}/\ker \mu)$ whose dimension is the trace of the matrix $\mu(p) \in M_m(\mathbb{C})$.

So Γ is a functor between the category of complex vector bundle over M and the category of finite projective $C^\infty(M)$-modules. The content of the Serre-Swan theorem is that that functor has an inverse, so that the two categories are equivalent.

That motivates the following definition

Definition 80.37.
*A **noncommutative vector bundle** is a right finite projective module over an algebra which is not specially commutative.*

Most of time we study modules overs algebras whose are dense subalgebra of a C^*-algebra.

80.8 Hermitian structure and compatible connection

80.8.1 Hermitian structures

The motivation of the following definitions is the fact that when $\mathfrak{A}$ is an involutive algebra of functions on the manifold M, any vector bundle is characterised by its $\mathfrak{A}$-module $\mathcal{E}$ of sections

which have the same regularity as the functions in $\mathfrak{A}$. If E is locally trivial and has finite dimension, $\mathcal{E}$ is a direct sum of free modules $\mathfrak{A}^N$ and $\mathcal{E}$ is a finite projective bundle on $\mathfrak{A}$.

The data of an inner product $\langle ., . \rangle_x$ on each fibre E_x provides a sesquilinear form

$$\begin{aligned} \langle ., . \rangle &: \mathcal{E} \times \mathcal{E} \to \mathfrak{A} \\ \langle \xi, \eta \rangle(x) &= \langle \xi(x), \eta(x) \rangle_x. \end{aligned} \tag{80.62}$$

This product fulfils the following relations for every $\xi, \eta \in \mathcal{E}$ and for all $a, b \in \mathfrak{A}$:

(1) $\langle \xi a, \eta b \rangle = a^* \langle \xi, \eta \rangle b$,

(2) $\langle \xi, \xi \rangle \geqslant 0$,

(3) $\mathcal{E}$ is self-dual for the product.

The first comes from the following simple computation:

$$\langle \xi a, \eta b \rangle(x) = \langle \xi(x)a(x), \eta(x)b(x) \rangle_x = a(x)^* b(x) \langle \xi(x), \eta(x) \rangle_x = \big((a^* b) \langle \xi, \eta \rangle \big)(x)$$

In the present case, the algebra $\mathfrak{A}$ is commutative and $a^* b \langle \xi, \eta \rangle = a^* \langle \xi, \eta \rangle b$, but since we will soon pass to a more general case, the right hand side form is more natural.

Definition 80.38.
*Let $\mathfrak{A}$ be an unital involutive algebra and $\mathcal{E}$, a finite projective module over $\mathfrak{A}$. An **hermitian structure** on $\mathcal{E}$ is a sesquilinear map $\langle ., . \rangle : \mathcal{E} \times \mathcal{E} \to \mathfrak{A}$ which fulfils the three conditions above.*

Notice that an hermitian structure is not a map $\mathcal{E} \times \mathcal{E} \to \mathbb{C}$ because, in the noncommutative picture, elements of $\mathcal{E}$ correspond to sections of E, not to *elements* of E. But on the sections we have $\langle s, s' \rangle(x) = \langle s(x), s'(x) \rangle \in \mathbb{C}$, so that the pairing of two sections is a functions, which corresponds to an element of $\mathfrak{A}$.

A **pre-C^*-module** structure on a dense subalgebra $\mathfrak{A}$ of the C^*-algebra A is a right module $\mathcal{E}$ over $\mathfrak{A}$ endowed with an Hermitian structure. If $\|.\|_A$ is the C^*-norm of A, we can put the norm

$$\|s\|_{\mathcal{E}} = \sqrt{\|(s|s)\|_A} \tag{80.63}$$

on $\mathcal{E}$. The completion of $\mathcal{E}$ for that norm gives rise to a C^*-module.

The typical example of that is a vector bundle $E \to M$ endowed with an hermitian structure, i.e. a product $\langle ., . \rangle_x : E_x \times E_x \to \mathbb{C}$ on each fibre. One consider the $C^\infty(M)$-module $\mathcal{E} = \Gamma(E, M)$ of sections and the product

$$\begin{aligned} \langle ., . \rangle &: \mathcal{E} \times \mathcal{E} \to C^\infty(M) \\ \langle \eta_1, \eta_2 \rangle(x) &= \langle \eta_1(c), \eta_2(c) \rangle_x. \end{aligned} \tag{80.64}$$

One immediately checks that it fulfils the three conditions

(1) $\langle \eta_1 a, \eta_2 b \rangle = a^* \langle \eta_1, \eta_2 \rangle b$,

(2) $\langle \eta_1, \eta_2 \rangle^* = \langle \eta_2, \eta_1 \rangle$,

(3) $\langle \eta, \eta \rangle \geqslant 0$ and $\langle \eta, \eta \rangle = 0$ if and only if $\eta = 0$.

The **dual module** of the right module $\mathcal{E}$ is the module

$$\mathcal{E}' = \{ \phi : \mathcal{E} \to \mathfrak{A} \text{ tel que } \phi(\eta a) = \phi(\eta) a \}.$$

The right module structure being defined by $\phi a = a^* \phi \in \mathcal{E}'$. One says that the Hermitian structure is **nondegenerate** when the map $\eta \mapsto \langle \eta, . \rangle$ is an isomorphism between $\mathcal{E}$ and $\mathcal{E}'$. We denote by $U_N(\mathfrak{A}) = U(\mathfrak{A}^N)$ the space of unitary endomorphisms of the free module $\mathfrak{A}^N$.

The free $\mathfrak{A}$-module $\mathfrak{A}^N$ accepts the structure of pre-C^*-module by

$$(r|s) = \sum_{j=1}^{N} r_j^* s_j. \tag{80.65}$$ EqHermCanN

Lemqqstarherm

Lemma 80.39.
If $q \in \mathbb{M}_N(\mathfrak{A})$, the product (80.65) fulfills

$$(q^*\xi|\eta) = (\xi|q\eta)$$

where q^* is the matrix defined by $(q^*)_{ij} = (q_{ji})^*$ where, in the right hand side, the star denotes the involution on $\mathfrak{A}$.

Proof. The proof is a simple computation:

$$(p^*\xi|\eta) = \sum_j (p^*\xi)_j^* = \sum_{jk} \big((p^*)_{jk}\xi_k\big)^*\eta_j = \sum_{jk} \xi_k^*(p^*)_{jk}^*\eta_j = \sum_{jk} \xi_k^* p_{kj}\eta_j = \sum_k \xi_k^*(p\eta)_k = (\xi|p\eta).$$

□

We suppose that the $\mathfrak{A}$-module $\mathcal{E}$ can be written as $\mathcal{E} = p\mathfrak{A}^N$ for a certain $N \in \mathbb{N}$ and an idempotent element [2] $p \in M_N(\mathfrak{A})$. Since $\mathcal{E}$ is self-dual for the hermitian structure, the data of $\langle \xi, \eta\rangle$ for all $\eta \in \mathcal{E}$ defines ξ. So we can define an involution $T \mapsto T^*$ on $\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ by the formula

$$\langle T^*\xi, \eta\rangle = \langle \xi, T\eta\rangle$$

for all ξ,$\eta \in \mathcal{E}$.

What we want to do is to put the restriction of the canonical structure (80.65) to $p\mathfrak{A}^N$ (with $p^2 = p$) as Hermitian structure on $\mathcal{E}$.

Proposition 80.40.
The canonical structure (80.65) restricts to $p\mathfrak{A}^N$ only if $p = p^*$.

Proof. We pose $\mathcal{E} = p\mathfrak{A}^N$ and

$$\mathcal{E}^\perp = \{u \in \mathfrak{A}^N \text{ tel que } (u|\eta) = 0 \text{ for all } \eta \in \mathcal{E}\}.$$

If $u \in \mathcal{E}^\perp$, we have $(ua|\eta) = a^*(u|\eta) = 0$, so that $\mathcal{E}^\perp$ is a right $\mathfrak{A}$-module. Take $u \in \mathfrak{A}^N$ and $\eta \in \mathcal{E}$, we have

$$\big((1-p^*)u|\eta\big) = \big(u|(1-p)\eta\big) = 0,$$

so that $(1-p^*)\mathfrak{A}^N \subseteq \mathcal{E}^\perp$. Let us now try to express $\mathcal{E}^\perp$ as $q\mathfrak{A}^N$. For every ξ and η in $\mathfrak{A}^N$, we must have

$$0 = (q\xi|p\eta) = (p^*q\xi|\eta),$$

which proves that $p^*q = 0$ because we suppose the hermitian structure to be nondegenerate. The image of q is contained in the kernel of p^*, so that $q = r(1-p^*)$ for a certain operator $r\colon \ker(p^*) \to \ker(p^*)$. Since $(1-p^*)\mathfrak{A}^N \subseteq q\mathfrak{A}^N$, the operator r must have an empty kenrel. So as set, $r(1-p^*)\mathfrak{A}^N = (1-p^*)\mathfrak{A}^N$. What we proved up to now is that

$$\mathcal{E}^\perp = (p\mathfrak{A}^N)^\perp = (1-p^*)\mathfrak{A}^N.$$

But the free module $\mathfrak{A}^N$ decomposes as $\mathfrak{A}^N = p\mathfrak{A}^N \oplus (1-p)\mathfrak{A}^N$. The canonical hermitian structire on $\mathfrak{A}^N$ restricts to $p\mathfrak{A}^N$ only if that decomposition is orthogonal with respect to the Hermitian product, so the condition is $(1-p)\mathfrak{A}^N = (p\mathfrak{A}^N)^\perp = (1-p^*)\mathfrak{A}^N$, or $p = p^*$. □

From now we suppose the module $\mathcal{E} = p\mathfrak{A}^N$ to fullfil the condition $p = p^2 = p^*$.

2. Notice that $p\mathfrak{A}^N \neq \mathfrak{A}^N$ in general because $a\mathfrak{A} \neq \mathfrak{A}$ in general. This situation is different of a group situation where $gG = G$ for any element $g \in G$.

80.8.2 Morita equivalence

If $\mathcal{E}$ is a right $\mathfrak{A}$-module , we define its **conjugate module** as

$$\bar{\mathcal{E}} = \{\bar{s} \text{ tel que } s \in \mathcal{E}\} \tag{80.66}$$

with a left $\mathfrak{A}$-module structure $(a\bar{s}) = \overline{sa^*}$. When $\mathcal{E} = p\mathfrak{A}^N$, we have $\bar{\mathcal{E}} = \overline{\mathfrak{A}^N}p$ and its elements have to be seen as row-vectors.

Let $\mathcal{E}$ be a projective $\mathfrak{A}$-module of finite type. An operator on $\mathcal{E}$ is a **ket-bra** if it is of the form

$$\begin{aligned} |\eta_1\rangle\langle\eta_2| \colon \mathcal{E} &\to \mathcal{E} \\ \xi &\mapsto \eta_1(\eta_2|\xi) \end{aligned} \tag{80.67}$$

Since $\eta_1(\eta_2|\xi a) = \eta_1(\eta_2|\xi)a$ for every ξ, η_1, $\eta_2 \in \mathcal{E}$ and $a \in \mathfrak{A}$, the ket-bra commutes with the right action of $\mathfrak{A}$ on $\mathcal{E}$, so that one can write

$$|\eta_1\rangle\langle\eta_2|\xi a$$

without ambiguities. The composition of ket-bra is still a ket-bra:

$$\Big(|\xi_1\rangle\langle\xi_2|\Big)\big(|\eta_1\rangle\langle\eta_2|\big)\sigma = |\xi_1\rangle\langle(\eta_1|\xi_2)\eta_2|\sigma. \tag{80.68}$$

The algebra of finite sums of ket-bras is an algebra that we denote by $\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$.

If V is a vector space, it is a known fact that $\mathrm{End}(V) = V \otimes V^*$. The same kind of identification holds for modules.

Proposition 80.41.
The map

$$\begin{aligned} \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) &\to \mathcal{E} \otimes_{\mathfrak{A}} \bar{\mathcal{E}} \\ |\xi_1\rangle\langle\xi_2| &\mapsto \xi_1 \otimes \bar{\xi}_2 \end{aligned} \tag{80.69}$$

is an isomorphism.

Proof. Any element of $\mathcal{E} \otimes_{\mathfrak{A}} \bar{\mathcal{E}}$ can be written under the form $(\xi_1 a) \otimes_{\mathfrak{A}} \bar{\xi}_2$, with eventually $a = 1$. We are going to prove that if

$$(\xi_1 a) \otimes_{\mathfrak{A}} \bar{\xi}_2 = \eta_1 \otimes \bar{\eta}_2, \tag{80.70}$$

then $|\xi_1 a\rangle\langle\xi_2| = |\eta_1\rangle\langle\eta_2|$. If (80.70) is true, then $\eta_1 = \xi_1$ and $\bar{\eta}_2 = a\bar{\xi}_2$, or $\xi_1 a = \eta_1$ and $\xi_2 = \eta_2$. The conclusion is immediate in the second case. In the first case, we have

$$|\xi_1 a\rangle\langle\xi_2|\sigma = \xi_1 a(\xi_1|\sigma) = \xi_1(\xi_2 a^*|\sigma) = \eta_1(\eta_2|\sigma) = |\eta_1\rangle\langle\eta_2|\sigma.$$

□

Now we pose $\mathcal{B} = \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$, and $\mathcal{E}$ becomes a left $\mathcal{B}$-module, so that one says that $\mathcal{E}$ is a $\mathcal{B}$-$\mathfrak{A}$-bimodule. We have $\mathcal{B} = p\mathbb{M}_N(\mathfrak{A})p$, and we denote an element of $\mathcal{B}$ by $\varphi_A = pAp$. We have $(\varphi_A)^* = \varphi_{A^*} = pA^*p$ because we assume $p = p^*$. As particular case of lemme 80.39, we have $(\xi|\varphi_A\eta) = (\varphi_A^*\xi|\eta)$.

The algebra $\mathcal{B} = \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ acts at right on $\bar{\mathcal{E}}$ by

$$\bar{\eta}\varphi_A = \overline{\varphi_A^*\eta}. \tag{80.71}$$

We are now able to consider the tensor product $\bar{\mathcal{E}} \otimes_{\mathcal{B}} \mathcal{E}$ in which

$$\bar{\eta} \otimes_{\mathcal{B}} (\varphi_A\xi) = \overline{\varphi_A^*\eta} \otimes_{\mathcal{B}} \eta. \tag{80.72}$$

Proposition 80.42.
The map

$$\begin{aligned} \psi \colon \bar{\mathcal{E}} \otimes_{\mathcal{B}} \mathcal{E} &\to \mathfrak{A} \\ \bar{\eta} \otimes_{\mathcal{B}} \xi &\mapsto (\eta|\xi) \end{aligned} \tag{80.73}$$

is an isomorphism of $\mathfrak{A}$-bimodule.

Proof. The map ψ is well defined because of relation (80.72). For the surjectivity, we use the assumption of non degenerate Hermitian structure. We have to prove that $\{(\xi|\eta) \text{ tel que } \xi\,\eta \in \mathcal{E}\} = \mathfrak{A}$. Pick any $\xi_0 \in \mathcal{E}$ and define a map $\phi\colon \mathcal{E} \to \mathfrak{A}$ such that $\phi(\xi_0) = a$ and $\phi(\eta b) = \phi(\eta)b$. In this case, there exists a $\eta \in \mathcal{E}$ such that $\phi(\xi) = (\eta|\xi)$. For this η we have $(\eta|\xi_0) = a$.

¡¡ Avertissement/question au lecteur !! 80.43
I'm not able to prove injectivity.

□

80.8.3 Differential over a module

We have now to extend the definition of δ from $\mathfrak{A}$ to $\mathcal{E} = p\mathfrak{A}^N$. In that purpose we begin to extend λ and p of page 3235 as

$$\begin{aligned} p\colon \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^p\mathfrak{A} &\to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p\mathfrak{A} \\ \sum_i e_i \otimes_{\mathbb{C}} \omega_i &\mapsto \sum_i p(f_i) \otimes_{\mathfrak{A}} \omega_i \end{aligned} \tag{80.74}$$

and

$$\begin{aligned} \lambda\colon \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p\mathfrak{A} &\to \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^p\mathfrak{A} \\ \lambda(\xi \otimes_{\mathfrak{A}} \omega) &= \sum_i e_i \otimes_{\mathbb{C}} \xi^i\omega \end{aligned} \tag{80.75}$$

if $\lambda(\xi) = \sum_i e_i \otimes_{\mathbb{C}} \xi^i$. Now we define $p\delta\colon \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p\mathfrak{A} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^{p+1}\mathfrak{A}$ by

$$p\delta = p \circ (\mathbb{1} \otimes \delta) \circ \lambda.$$

In order to see what is going on, let us apply that on a general element of $\mathcal{E} \otimes_{\mathfrak{A}} \Omega^p\mathfrak{A}$:

$$\begin{aligned} \big(p \circ (\mathbb{1} \otimes \delta) \circ \lambda\big)p\big(\sum_i e_i \otimes_{\mathbb{C}} \omega_i\big) &= p \circ (\mathbb{1} \otimes \delta)\big(\sum_i e_i \otimes_{\mathbb{C}} \omega_i\big) \\ &= p\big(\sum_i e_i \otimes_{\mathbb{C}} \delta\omega_i\big) \\ &= \sum_i p(f_i) \otimes_{\mathfrak{A}} \delta\omega_i. \end{aligned}$$

The definition of $pd\xi$ is in the same way: if $\xi = p\big(\sum_i e_i \otimes_{\mathbb{C}} \omega_i\big)$, we pose

$$pd\xi = \sum_i p(f_i) \otimes_{\mathfrak{A}} d\omega_i. \tag{80.76}$$

80.8.4 Connections over Hermitian modules

Definition 80.44.
Let $\mathcal{E}$ be a finite projective hermitian module on $\mathfrak{A}$. A ***connection*** *is a linear map $\nabla\colon \mathcal{E} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1_D(\mathfrak{A})$ such that*

$$\nabla(\xi a) = (\nabla\xi)a + \xi \otimes da \tag{80.77}$$

for all $\xi \in \mathcal{E}$ and $a \in \mathfrak{A}$. Recall that da means $[D, a]$ in the operator representation. The connection is said to be ***compatible with the hermitian structure*** *if*

$$\langle \xi, \nabla\eta\rangle - \langle \nabla\xi, \eta\rangle = d\langle \xi, \eta\rangle. \tag{80.78}$$

We denote by $\mathcal{C}(\mathcal{E})$ the space of compatible connections.

The latter deserve some comments. First, $\langle \xi,\, \eta\rangle \in \mathfrak{A}$, so that $d\langle \xi,\, \eta\rangle$ makes sense. Second, we have to define $\langle .,\, .\rangle$ when there are a component in $\Omega^1(\mathfrak{A})$. We naturally define

$$\langle \xi,\, \varphi \otimes_{\mathfrak{A}} \omega\rangle = \underbrace{\langle \xi,\, \varphi\rangle}_{\in\mathfrak{A}}\omega.$$

which is a well defined element in $\Omega^1_D(\mathfrak{A})$. There are no kind of left Leibniz rule in the definition of a connection. There always exists a compatible connection. For example the **Grassmann connection** defined by

$$\nabla_0\xi = p\eta \tag{80.79}$$

if $\eta_j = d\xi_j$. That connection is compatible because

$$\begin{aligned} d\langle \eta,\, \xi\rangle &= d\Big(\sum_{i=1}^N \eta_i^*\xi_i\Big) = \sum d\eta_i^*\xi_i + \sum \eta_i^* d\xi_i = \sum(-)(d\eta_i)^*\xi_i + \sum \eta_i^* d\xi_i \\ &= -\langle d\eta,\, p\xi\rangle + \langle p\eta,\, d\xi\rangle = -\langle pd\eta,\, \xi\rangle + \langle \eta,\, pd\xi\rangle = -\langle \nabla_0\eta,\, \xi\rangle + \langle \eta,\, \nabla_0\xi\rangle \end{aligned} \tag{80.80}$$

where we used the fact that $p = p^*$.

We denote by $C(\mathcal{E})$ the space of compatibles connections on $\mathcal{E}$, and the **unitary group** of $\mathcal{E}$ is

$$\mathrm{U}(\mathcal{E}) = \{u \in \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \text{ tel que } uu^* = u^*u = 1\}.$$

This group acts by conjugation on $C(\mathcal{E})$ following the formula $\gamma_u(\nabla) = u\nabla u^*$, or more explicitly:

$$(\gamma_u\nabla)\xi = u\nabla(u^*\xi). \tag{80.81}$$

One easily checks that $\gamma_u(\nabla)$ is a compatible connection when ∇ is a compatible connection:

$$\begin{aligned} (\gamma_u\nabla)(\xi a) &= u\nabla(u^*\xi a) = u\big(\nabla(u^*\xi)a + u^*\xi\otimes da\big) \\ &= u\nabla(u^*\xi)a + \xi\otimes da = \big(\gamma_u(\nabla)\xi\big)a + \xi\otimes da, \end{aligned}$$

and, using the sesquilinearity,

$$\begin{aligned} \langle \xi,\, (\gamma_u\nabla)\eta\rangle - \langle (\gamma_u\nabla)\xi,\, \eta\rangle &= \langle \xi,\, u\nabla(u^*\eta)\rangle - \langle u\nabla(u^*\xi),\, \eta\rangle \\ &= \langle u^*\xi,\, \nabla(u^*\eta)\rangle - \langle \nabla(u^*\xi),\, u^*\eta\rangle, \end{aligned}$$

which gives the compatibility using the compatibility of ∇.

80.8.5 Universal compatible connection

An **universal compatible connection** on a finite projective Hermitian module $\mathcal{E}$ is a map $\nabla\colon \mathcal{E} \to \mathcal{E}\otimes_{\mathfrak{A}}\Omega^1(\mathfrak{A})$ such that

(1) $\nabla(\xi a) = (\nabla\xi)a + \xi\otimes da$,

(2) $\langle \xi, \nabla\eta\rangle - \langle \nabla\xi, \eta\rangle = d\langle \xi, \eta\rangle$

for every $\xi,\ \eta \in \mathcal{E}$ and $a \in \mathfrak{A}$. A map which only fulfils the first condition is an **universal connection**.

We denote by $\mathcal{CC}(\mathcal{E})$ the space of universal compatible connections. An universal compatible connection extends in an unique way to a map

$$\begin{aligned} \tilde\nabla\colon \mathcal{E}\otimes_{\mathfrak{A}}\Omega^p\mathfrak{A} &\to \mathcal{E}\otimes_{\mathfrak{A}}\Omega^{p+1}\mathfrak{A} \\ \xi\otimes\omega_0 &\mapsto (\nabla\xi)\omega_0 + \xi\otimes d\omega_0 \end{aligned} \tag{80.82}$$

with $\xi \in \mathcal{E}$ and $\omega \in \Omega^*_D$. If $\nabla\xi = \eta\otimes\omega_0$, the expression $(\nabla\xi)\omega$ means $(\nabla\xi)\omega = \eta\otimes(\omega_0\omega)$. The operation $\tilde\nabla$ fulfils the Leibniz rule

$$\tilde\nabla(\eta\omega) = (\tilde\nabla\eta) + (-1)^\eta\eta d\omega$$

if η is homogeneous and $\omega \in \Omega^*_D$. Indeed the fact that η is homogeneous means that it reads $\xi \otimes \omega_0$ with $\omega_0 \in \Omega^k_D$ for a certain k. Indeed,

$$\begin{aligned}\tilde{\nabla}(\eta\omega) &= \tilde{\nabla}(\xi \otimes \omega_0\omega) = (\nabla\xi)\omega_0\omega + \xi \otimes \big((d\omega_0)\omega + (-1)^{\omega_0}\omega_0 d\omega\big)\\ &= \big((\nabla\xi)\omega_0 + (d\omega_0)\big)\omega + (-1)^{\omega_0}\xi \otimes \omega_0 d\omega = \tilde{\nabla}(\xi \otimes \omega_0)\omega + (-1)^{\eta}\xi \otimes \omega_0 d\omega\\ &= (\tilde{\nabla}\eta)\omega + (-1)^{\eta}\eta d\omega.\end{aligned}$$

Most of time, we will make no difference between ∇ and $\tilde{\nabla}$.

Lemma 80.45.
Every universal connection $\nabla \in \mathcal{CC}(\mathcal{E})$ reads

$$\nabla = p\delta + \alpha \tag{80.83}$$

for a certain $\alpha \in p\mathbb{M}_{N\times N}(\mathfrak{A})p \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$.

Proof. We can express the module as $\mathcal{E} = p\mathfrak{A}^N$. If ∇_1 and ∇_2 are two universal connections, we have

$$(\nabla_1 - \nabla_2)(\eta a) = (\nabla_1\eta)a + \eta(\delta a) - (\nabla_2\eta)a - \eta(\delta a) = \big((\nabla_1 - \nabla_2)\eta\big)a,$$

so that $\nabla_1 - \nabla_2 \in \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A} \simeq p\mathbb{M}_{N\times N}(\mathfrak{A})p \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$. Taking in particular the Grassmannian connection ∇_0 as ∇_1, we find that for any universal connection ∇ fulfils $\nabla - p\delta = \alpha$ with $\alpha \in \mathbb{M}_{N\times N}(\mathfrak{A}) \otimes \Omega^1(\mathfrak{A})$ satisfying $\alpha = p\alpha = \alpha p = p\alpha p$. □

When ∇ is given, the corresponding α is the **gauge potential** of ∇.

Proposition 80.46.
The representation $\pi\colon \Omega(\mathfrak{A}) \to \mathfrak{B}(\mathscr{H})$ given in equation (80.15) extends to

$$\begin{aligned}(\mathbb{1} \otimes \pi)\colon \mathcal{CC}(\mathcal{E}) &\to \mathcal{C}(\mathcal{E})\\ \nabla &\mapsto (\mathbb{1} \otimes \pi) \circ \nabla\end{aligned} \tag{80.84}$$

which is surjective.

As a particular case of the proposition, let us prove that the Grassmann connection belongs to $\pi\big(\mathcal{CC}(\mathcal{E})\big)$. When one defines the Grassmann connection as $(\nabla_0\xi) = p\eta$, the right hand side has to be understood as $p\eta \otimes_{\mathfrak{A}} 1$ with the 1 of $\Omega^1_D(\mathfrak{A})$. So the Grassmann connection is obtained by applying π to the connection $\nabla'_0(\xi) = p\eta \otimes_{\mathfrak{A}} 1$ with the 1 of $\Omega^1(\mathfrak{A})$.

In particular existence of that connection shows that every projective module accepts a connection.

¡¡ Avertissement/question au lecteur !! 80.47
The following proof is finished?

Proof. By construction, the map

$$(\mathbb{1} \otimes \pi)\colon \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \otimes \Omega^1(\mathfrak{A}) \to \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \otimes_{\mathfrak{A}} \Omega^1_D(\mathfrak{A})$$

is surjective.

□

Proposition 80.48.
A right module has a connection if an only if it is projective.

Proof. See proposition 7.4 in [819]. □

80.8.6 Curvature

Now we define the **curvature** of ∇ by $\theta(\nabla) = \tilde{\nabla}^2$.

Proposition 80.49.
This curvature is an endomorphism of $\tilde{\mathcal{E}}$ for its structure of right $\Omega^*_D(\mathfrak{A})$-module.

Proof. If we consider $\xi \otimes \omega_0 \in \tilde{\mathcal{E}}$ and $\omega \in \Omega^*_D$ with $\nabla\xi = \eta \otimes \omega'$, it is just a computation to prove that

$$\tilde{\nabla}^2(\xi \otimes \omega_0\omega) = \big((\nabla\eta)\omega'\omega_0 + \eta \otimes d\omega'\omega_0\big)\omega,$$

and then to see that $\theta\big((\xi \otimes \omega_0)\omega\big) = \theta(\xi \otimes \omega_0)\omega$. □

From the formula

$$\tilde{\nabla}^2(\xi \otimes \omega_0) = (\nabla\eta)\omega'\omega_0 + \eta \otimes d\omega'\omega_0, \tag{80.85}$$

we see that $\theta = \tilde{\nabla}^2$ is completely determined by its restriction to $\mathcal{E}$.

Let us compute $\nabla^2(\xi \otimes_{\mathfrak{A}} \omega_0)$ with $\omega_0 \in \Omega^p\mathfrak{A}$ and $\xi \in \mathcal{E}$. If $\nabla\xi = \eta \otimes_{\mathfrak{A}} \omega'$,

$$\begin{aligned}
\nabla^2(\xi \otimes_{\mathfrak{A}} \omega_0) &= \nabla\big((\nabla\xi)\omega_0 + \xi\delta\omega_0\big) \\
&= (\nabla\eta)\omega'\omega_0 + (-1)^\eta\eta \otimes_{\mathfrak{A}} \delta(\omega'\omega_0) + (\nabla\xi)\delta\omega_0 + \xi\delta^2\omega_0 \\
&= (\nabla\eta)\omega'\omega_0 + (-1)^\eta\eta(\delta\omega')\omega_0 + (-1)^\eta\eta(-1)^{\omega'}\delta\omega_0 + (\eta \otimes_{\mathfrak{A}} \omega')\delta\omega_0 \\
&= \big((\nabla\eta)\omega' + \eta\delta\omega'\big)\omega + (\eta \otimes_{\mathfrak{A}} \omega' - \eta \otimes \omega').
\end{aligned}$$

We finally have

$$\nabla^2(\xi \otimes_{\mathfrak{A}} \omega_0) = \big((\nabla\eta)\omega' + \eta \otimes_{\mathfrak{A}} \delta\omega'\big)\omega_0 \tag{80.86}$$

if $\nabla\xi = \eta \otimes \omega'$. That shows that

$$\theta\big((\xi \otimes_{\mathfrak{A}} \omega_0)\omega\big) = \theta(\xi \otimes_{\mathfrak{A}} \omega_0)\omega. \tag{80.87}$$ EqThetaConhomega

In terms of the gauge potential α, the curvature reads

$$\theta(\xi) = (p\delta p\delta p + p(\delta\alpha) + \alpha^2)\xi \tag{80.88}$$ EqConnPotalpha

because $\alpha = p\alpha = \alpha p = p\alpha p$ and $p\xi = \xi$.

80.8.7 Transformation law under the unitary group

The unitary group (subsection 54.3.7) acts on an element $\nabla \in \mathcal{CC}(\mathcal{E})$ by

$$(u \cdot \nabla)(\xi \otimes_{\mathfrak{A}} \omega_0) = u\nabla\big(u^*(\xi \otimes_{\mathfrak{A}} \omega_0)\big), \tag{80.89}$$

or in a more compact way $u \cdot \nabla = \nabla^u = u\nabla u^*$. We want to know how does the curvature change under such a transformation of the connection: we want to know $\theta(\nabla^u)$ in terms of ∇ and u. We immediately have

$$\theta^u(\Sigma) = u\nabla\Big(u^*\big(u\nabla u^*(\xi \otimes \omega_0)\big)\Big) = u\nabla u^*u\nabla\big(u^*(\xi \otimes_{\mathfrak{A}} \omega_0)\big) = u\theta u^*(\Sigma)$$

with $\Sigma \in \mathcal{E} \otimes_{\mathfrak{A}} \Omega\mathfrak{A}$. So we have the covariant transformation law

$$\theta^u = u\theta u^*. \tag{80.90}$$

Any connection can be written under the form (80.83): $\nabla = p\delta + \alpha$. It is a simple computation to find α^u defined by $\nabla^u = p\delta + \alpha^u$; using the notation $\Sigma = \xi \otimes_{\mathfrak{A}} \omega_0$, we have

$$\begin{aligned}
\nabla^u(\xi \otimes_{\mathfrak{A}} \omega_0) &= u(p\delta + \alpha)u^*(\xi \otimes_{\mathfrak{A}} \omega_0) \\
&= up\delta(u^*\Sigma) + u\alpha(u^*\sigma) \\
&= up(\delta u^*)\Sigma + upu^*\delta\Sigma + u\alpha u^*\Sigma \\
&= (p\delta + pu\delta u^* + u\alpha u^*)\Sigma,
\end{aligned}$$

so the transformation law of the gauge potential reads

$$\alpha^u = pu\delta u^* + u\alpha u^*. \tag{80.91}$$

80.8.8 Connection on bimodule

Let $\mathcal{E}$ be a left and right-projective $\mathfrak{A}$-bimodule with a right connection $\nabla\colon \mathcal{E} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$. If $\sigma\colon \Omega^1 \otimes_{\mathfrak{A}} \mathcal{E} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$ is a bimodule isomorphism, we say that the couple (∇, σ) is **compatible** if we have the Leibniz rule

$$\nabla(a\xi) = (\nabla\xi)a + \sigma(\delta a \otimes_{\mathfrak{A}} \xi). \tag{80.92}$$

80.8.9 Connection on dual module

The pairing $(.,.)\colon \mathcal{E}' \times \mathcal{E} \to \mathfrak{A}$ defined by $(\phi, \xi) = \phi(\xi)$ can be extended by two ways:

(1)

$$\begin{aligned}(.,.)\colon \big(\mathcal{E}' \otimes_{\mathfrak{A}} \Omega\mathfrak{A}\big) \times \mathcal{E} &\to \Omega\mathfrak{A}\\ (\phi \otimes_{\mathfrak{A}} \omega, \xi) &= \omega^*(\phi, \xi)\end{aligned} \tag{80.93}$$

(2)

$$\begin{aligned}(.,.)\colon \mathcal{E}' \times \big(\mathcal{E} \otimes_{\mathfrak{A}} \Omega\mathfrak{A}\big) &\to \Omega\mathfrak{A}\\ (\phi, \xi \otimes_{\mathfrak{A}} \omega) &= (\phi, \xi)\omega.\end{aligned} \tag{80.94}$$

When one has a connection over $\mathcal{E}$, one defines the **dual connection** $\nabla'\colon \mathcal{E}' \to \mathcal{E}' \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$ by

$$(\nabla'\phi, \xi) = (\phi, \nabla\xi) - \delta(\phi, \xi). \tag{80.95}$$

The dual connection satisfies the Leibniz rule

$$\nabla'(\phi a) = \phi \otimes_{\mathfrak{A}} \delta a + (\nabla'\phi)a. \tag{80.96}$$

Indeed,

$$\begin{aligned}\big(\nabla'(\phi a), \xi\big) &= (\phi a, \nabla\xi) - \delta\underbrace{(\phi a, \xi)}_{=a^*(\phi,\xi)}\\ &= (\phi a, \nabla\xi) + (\delta a)^*(\phi, \xi) - a^*\big((\phi, \nabla\xi) - (\nabla'\phi, \xi)\big)\\ &= (\delta a)^*(\phi, \xi) + a^*(\nabla'\phi, \xi)\\ &= (\phi \otimes_{\mathfrak{A}} \delta a, \xi) + \big((\nabla'\phi)a, \xi\big).\end{aligned}$$

80.8.10 Connections and Connes' calculus

Let $\mathcal{E}$ be a projective $\mathfrak{A}$-module and let us consider the Connes' calculus $(\Omega_D\mathfrak{A}, d)$. In this framework, a **connection** is a $\mathbb{C}$-linear map $\nabla\colon \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p_D\mathfrak{A} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^{p+1}_D\mathfrak{A}$ such that

$$\nabla(\Sigma\omega) = (\nabla\Sigma)\omega + (-1)^p \Sigma\, d\omega \tag{80.97}$$

for every $\Sigma \in \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p_D\mathfrak{A}$ and $\omega \in \Omega_D\mathfrak{A}$. The composition $\nabla^2 = \nabla \circ \nabla$ is $\Omega_D\mathfrak{A}$-linear and the restriction to $\mathcal{E}$ is the **curvature** denoted by $F(\nabla)$. An interesting property of the curvature is its $\mathbb{C}$-linearity

$$F(\xi a) = F(\xi)a, \tag{80.98}$$

indeed

$$\begin{aligned}F(\xi a) &= \nabla\big((\nabla\xi)a + \xi da\big) = (\nabla^2\xi)a - (\nabla\xi)da + \nabla(\xi da)\\ &= F(\xi)a - (\nabla\xi)da + (\nabla\xi)\, da + \xi\, d^2a.\end{aligned}$$

More generally we have

$$\nabla^2(\xi \otimes_{\mathfrak{A}} \rho) = F(\xi)\rho. \tag{80.99}$$

The operator ∇^2 can be seen as an element of $\mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \otimes_{\mathfrak{A}} \Omega^2_D\mathfrak{A}$ and we have the **Bianchi identities**

$$[\nabla, F] = 0 \tag{80.100}$$

because $\nabla \circ \nabla^2 = \nabla^2 \circ \nabla$.

In the same way as equation (80.88), we have

$$F = p(dA) + A^2 + p\,dp\,p\,dp \tag{80.101}$$

We say that the connection $\nabla\colon \mathcal{E} \otimes_{\mathfrak{A}} \Omega^p_D\mathfrak{A} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^{p+1}_D\mathfrak{A}$ is **compatible** with the hermitian structure if for all ξ, $\eta \in \mathcal{E}$,

$$-\langle \nabla\eta, \xi\rangle + \langle \eta, \nabla\xi\rangle = d\langle \eta, \xi\rangle. \tag{80.102}$$

That definition relies on the following two extensions of the hermitian structure:

$$\begin{aligned} \langle ., .\rangle\colon \left(\mathcal{E} \otimes_{\mathfrak{A}} \Omega^1_D\mathfrak{A}\right) \times \mathcal{E} &\to \Omega^1_D\mathfrak{A} \\ \langle \xi \otimes_{\mathfrak{A}} \omega, \eta\rangle &= \omega^*\langle \xi, \eta\rangle \end{aligned} \tag{80.103}$$

and

$$\begin{aligned} \langle ., .\rangle\colon \mathcal{E} \times \left(\mathcal{E} \otimes_{\mathfrak{A}} \Omega^1_D\mathfrak{A}\right) &\to \Omega^1_D\mathfrak{A} \\ \langle \eta, \xi \otimes_{\mathfrak{A}} \omega\rangle &= \langle \eta, \xi\rangle\omega. \end{aligned} \tag{80.104}$$

Proposition 80.50.
The connection $\nabla = pd + A$ is compatible with the canonical hermitian structure $\langle \xi, \eta\rangle = \sum_{j=1}^N \xi_j\eta_j$ if and only if $A = A^*$.

80.8.10.1 The simplest example in the world

ConnEequalAsime

We will build the connection theory in the case of $\mathcal{E} = \mathfrak{A}$ with the hermitian structure $\langle a, b\rangle = a^*b$. A connection on that module is a map $\nabla\colon \mathfrak{A} \to \mathfrak{A} \otimes_{\mathfrak{A}} \Omega^1_D\mathfrak{A} = \Omega^1_D\mathfrak{A}$ such that $\nabla(ab) = (\nabla a)b + a\,db$.

A connection has the form $\nabla = d + A$ with $A \in \mathbb{M}_{1\times 1}(\mathfrak{A}) \otimes_{\mathfrak{A}} \Omega^1_D\mathfrak{A} = \Omega^1_D\mathfrak{A}$ and the compatibility reads $A = A^*$. The curvature is given by $F = dA + A^2$.

In this particular case, one often write V instead of A and θ instead of F.

80.8.10.2 Action of the gauge group

PgSimplestwordl

If $u \in U(\mathcal{E})$, we define the action of u on the connection ∇ by

$$\nabla^u = u\nabla u^* \tag{80.105}$$

in the sense that $\nabla^u(\xi) = (u \circ \nabla)(u^*\xi)$. In order to find the transformation law for the vector potential, we consider $\nabla = pd + A$ and we compute

$$\nabla^u\xi = u\nabla(u^*\xi) = upd(u^*\xi) + uAu^*\xi = up(du^*)\xi + upu^*d\xi + uAu^*\xi = (updu^* + pd + uAu^*)\xi,$$

so that we have the transformation law

$$A^u = updu^* + uAu^*. \tag{80.106}$$

EqTransAjauge

80.8.11 Scalar product on $\pi(\Omega^p\mathfrak{A})$ and Hilbert structure

A spectral triple $(\mathfrak{A}, \mathscr{H}, D)$ is **tame** if for all $T \in \pi(\Omega\mathfrak{A})$ and $S \in \mathfrak{B}(\mathscr{H})$, we have

$$\mathrm{Tr}_\omega(ST|D|^{-n}) = \mathrm{Tr}_\omega(S|D|^{-n}T)$$

where Tr_ω is the Dixmier trace. In this case, we have

$$\mathrm{Tr}_\omega(T_1^*T_2|D|^{-n}) = \mathrm{Tr}_\omega(T_1^*|D|^{-n}T_2) = \mathrm{Tr}_\omega(T_2|D|^{-n}T_1^*) := \langle T_1, T_2\rangle. \tag{80.107}$$

Second equality is the cyclic invariance of trace while the last one is a definition. Tameness assures the symmetry: $\langle T_1, T_2\rangle = \langle T_2, T_1\rangle$.

Note that when $T \in \pi(\Omega^p\mathfrak{A})$, we have $T \in \mathfrak{B}(\mathscr{H})$. Indeed the definition of π is $\pi(a_0\delta a_1) = a_0[D, a_1]$ and the very definition of a spectral triple says that $[D, a_1]$ and a_0 are bounded.

Notice that the definition

$$\langle T_1, T_2\rangle_p = \mathrm{Tr}_\omega(T_1^*|D|^{-n}T_2) \tag{80.108}$$

only holds for $T_1, T_2 \in \pi(\Omega^p\mathfrak{A})$. When $T_1 \in \pi(\Omega^p\mathfrak{A})$ and $T_2 \in \pi(\Omega^q\mathfrak{A})$ with $p \neq q$, we define $\langle T_1, T_2\rangle$.

Let us now take $\tilde{\mathscr{H}_p}$, the completion of $\pi(\Omega^p\mathfrak{A})$ for $\langle ., .\rangle_p$. When $a \in \mathfrak{A}$ and $T_1, T_2 \in \pi(\Omega^p\mathfrak{A})$, using tameness and cyclic invariance, we find [EqUnitRepHtilde]

$$\langle aT_1, aT_2\rangle_p = \mathrm{Tr}_\omega(T_1^*a^*|D|^{-n}aT_2) = \mathrm{Tr}_\omega(aT_2|D|^{-n}T_1^*a^*) = \mathrm{Tr}_\omega(T_2|D|^{-n}T_1^*a^*a), \tag{80.109a}$$

and with the same computation

$$\langle T_1a, T_2a\rangle_p = \mathrm{Tr}_\omega(aa^*T_1^*|D|^{-n}T_2). \tag{80.109b}$$

We define the **unitary group** $U(\mathfrak{A})$ by

$$U(\mathfrak{A}) = \{u \in \mathfrak{A} \text{ tel que } u^*u = uu^* = 1\}. \tag{80.110}$$

This group can be represented on $\tilde{\mathscr{H}_p}$ by

$$L(a)T = aT \quad \text{and} \quad R(a)T = Ta.$$

Equations (80.109) show that R and L are unitary representations:

$$\langle L(a)T_1, L(a)T_2\rangle = \langle R(a)T_1, R(a)T_2\rangle = \langle T_1, T_2\rangle.$$

Lemma 80.51.
The set

$$\pi\big(\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})\big)$$

is a two-sided submodule of $\pi(\Omega^p\mathfrak{A})$.

Proof. An element of $\pi\delta(J_0 \cap \Omega^p\mathfrak{A})$ has the form $\pi\delta(\omega)$ with $\omega \in \Omega^p\mathfrak{A}$ and $\omega \in J_0$, i.e. $\pi(\omega) = 0$. Let $a \in \mathfrak{A}$; from definition of π, we have $a\pi\delta(\omega) = \pi(a\delta\omega)$. Then

$$\pi(a\delta\omega) = \pi\big(\delta(a\omega) - \delta(a)\omega\big)$$

in which the second term is $\pi(\delta a)\pi(\omega) = 0$. The conclusion is that $a\pi\delta(\omega) = \pi(\delta(a\omega))$. In order for $\pi\delta(J_0 \cap \Omega^p\mathfrak{A})$ to be a left sub-module of $\pi(\Omega^{p+1}\mathfrak{A})$, it remains to be proved that $a\omega \in J_0 \cap \Omega^p\mathfrak{A}$ when $\omega \in J_0 \cap \Omega^p\mathfrak{A}$. It is the case because $\pi(a\omega) = \pi(a)\pi(\omega) = 0$.

The right sub-module structure is left as an exercise.

□

Let us see that the closure of $\pi\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})$ in $\tilde{\mathscr{H}}$ is invariant under the representations L and R of $U(\mathfrak{A})$. Let $u \in U(\mathfrak{A})$ and $\omega \in \overline{\pi\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})}$, we want $R(u)\pi\delta\omega \in \overline{\pi\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})}$ where the bar denotes the closure. We have $R(u)\pi\delta\omega = (\pi\delta\omega)u$, but we just proved that $\pi\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})$ is a right module, so $(\pi\delta\omega)u$ still belongs to this set.

Proposition 80.52.
Let P_p *be the projection parallel to* $\pi\delta(J_0 \cap \Omega^{p-1}\mathfrak{A})$ *in* $\tilde{\mathscr{H}_p}$. *Then* $[L(a), P_p] = 0$.

Proof. Let $x = \pi\delta\omega + y \in \tilde{\mathscr{H}_p}$ with $\omega \in J_0 \cap \Omega^{p-1}\mathfrak{A}$ and y in the complement. We have

$$P_pL(a)x = P_p(ay)$$

because $a\pi\delta\omega \in \pi\delta J_0^{p-1}$ from the fact that $\pi\delta J_0^{p-1}$ is a sub-module. On the other hand, $L(a)P_px = ay$. It remains to be proved that $P_p(ay) = ay$ when, by construction, $P_py = y$. Indeed if ay has a component in J_0^{p-1}, the y has too because $y = a^{-1}(ay)$ and J_0^{p-1} is a submodule. □

80.9 Connections on modules and Yang-Mills functional

80.9.1 Potential vector and field strength

Let $(\mathfrak{A}, \mathscr{H}, D)$ be a tame spectral triple of dimension b on which we build $\Omega_D\mathfrak{A} = \oplus_p \Omega^p_D\mathfrak{A}$. A **potential vector**, or a **gauge potential** is a self-adjoint element of $\Omega^1_D\mathfrak{A}$. The **field strength** corresponding to the potential vector V is

$$\theta = dV + V^2,$$

to be compared with the expression (56.261) for the curvature.

Proposition 80.53.
The field strength is self-adjoint.

Proof. What does mean, at the class level, the fact to be self-adjoint? Remember that

$$\Omega_D\mathfrak{A} \simeq \Omega\mathfrak{A}/J \simeq \pi(\Omega\mathfrak{A})/\pi(\delta J_0).$$

An element $V \in \Omega^1_D\mathfrak{A}$ is $V = [\omega]$ with $\omega \in \Omega^1\mathfrak{A}$ and the class taken as $[\omega] = [\omega + \eta_1 + \delta\eta_2]$ for all $\eta_1 \in \Omega^1\mathfrak{A}$ and $\eta_2 \in \mathfrak{A}$ such that $\pi(\eta_1) = \pi(\eta_2) = 0$. The condition $V = V^*$ means $[\omega] = [\omega^*]$ which in turn *doesn't* implies that $\omega = \omega^*$.

Now let $\omega = a\delta b$, we have

$$dV - dV^* = [\delta a \delta b] - [d(\delta b^*\, a^*)] = [\delta a \delta b] + [\delta b^* \delta a^*] \neq 0.$$

So we have $[\delta(\omega - \omega^*)] \neq 0$ although $[\omega - \omega^*] = 0$. That proves that $\omega - \omega^*$ is junk.

From definition, η is junk when $\pi(\eta) = 0$ while $\pi(\delta\eta) \neq 0$. The condition $V = V^*$ means $[\omega] = [\omega^*]$, and then $[\omega - \omega^*] = 0$. Therefore we can write $\omega - \omega^* = \eta$ with $\pi(\eta) = 0$ and $\pi(\delta\eta) \neq 0$. In conclusion we have

$$\pi(\omega - \omega^*) = 0 \tag{80.111a}$$
$$\pi(\delta\omega - \delta\omega^*) \neq 0. \tag{80.111b}$$

Since $dV = [\delta\omega]$ and $(dV)^* = [\delta\omega^*]$,

$$dV - (dV)^* = [\delta\omega - \delta\omega^*] = [\delta(\omega - \omega^*)].$$

The claim $dV = (dV)^*$ needs to show that $[\delta(\omega - \omega^*)] = 0$. It is the case because it is of the form $[\delta\eta]$ with $\pi(\eta) = 0$.

□

80.9.2 Action of the gauge group

The element $u \in U(\mathfrak{A})$ acts on the potential vector V by

$$V \mapsto V^u := uVu^* + u[D, u^*], \tag{80.112}$$ EqDefActGaugePotVec

to be compared with the expression given in proposition (56.261) of the transformation rule of a connection under a gauge transformation.

Lemma 80.54.
When V is a potential vector, the element V^u is still a potential vector.

Proof. Our work is to prove that $V^u - (V^u)^* = 0$ in the sense of classes. If $V = [\omega]$, the definition is $[\omega]^u = [u\omega u^*] + [u\delta u^*]$, and thus

$$(V^u)^* = [u\omega u^*] + [\delta u u^*].$$

So we have to prove that $u\delta u^* - \delta u u^* \in J$. We have $\delta(1) = \delta(uu^*) = \delta u u^* - u\delta u^*$, but $\pi(\delta(1)) = [D, 1] = 0$, so $\delta(1) \in J$. It proves that $V^u = (V^u)^*$. □

Lemma 80.55.
The transformation law of the field strength is

$$\theta^u = u\theta u^*. \tag{80.113}$$

LemTrFieldStrm

Proof. No proof. □

80.9.3 Yang-Mills

The **Yang-Mills functional** is defined by

$$YM(V) := \langle dV + V^2, dV + V^2 \rangle. \tag{80.114}$$

The product is *a priori* only defined on $\Omega^2\mathfrak{A}$, not on the classes. Let us prove that the Yang-Mills functional is well defined. We have to prove that

Theorem 80.56.
The Yang-Mills functional

(1) is positive,

(2) is gauge invariant:

$$YM(V) = YM(V^u), \tag{80.115}$$

(3) and satisfies a minimal principle

$$YM(V) = \inf\{I(\alpha) \text{ tel que } \pi(\alpha) = V\} \tag{80.116}$$

where

$$I(\alpha) = \mathrm{Tr}_\omega \left(\pi(\delta\alpha + \alpha^2)^2 |D|^{-n}\right)$$

for $\alpha \in \Omega^1\mathfrak{A}$.

Proof. From property of the Dixmier trace, $\mathrm{Tr}_\omega(T) \geqslant 0$ when $T \geqslant 0$. In our case we are looking on $T = \theta^*\theta|D|^{-n}$

¡¡ Avertissement/question au lecteur !! 80.57
This is for sure positive in some sense, but I have to formalise that concept.

The invariance of Yang-Mills under $U(\mathfrak{A})$ is easy:

$$YM(V^u) = YM\big(uVu^* + u[D, u^*]\big) = \langle u\theta u^*, u\theta u^* \rangle = \mathrm{Tr}_\omega(\theta^2|D|^{-n}) = YM(V). \tag{80.117}$$

For the third claim, we begin by proving that I is positive and invariant under the transformation

$$\alpha \mapsto \alpha^u := u\alpha u^* + u\delta u^*$$

in the space $\{\alpha \in \Omega^1\mathfrak{A} \text{ tel que } \alpha = \alpha^*\}$. Positivity comes from Dixmier trace while for invariance first notice that lemma 80.113 says that under $\alpha \mapsto u\alpha u^* + u\delta u^*$, we have the transformation $\delta\alpha + \alpha^2 \mapsto u(\delta\alpha + \alpha^2)u^*$. Thus

$$\begin{aligned} I(\alpha^u) &= \mathrm{Tr}_\omega \left[\pi\big(u(\delta\alpha + \alpha^2)u^*\big)^2 |D|^{-n}\right] \\ &= \mathrm{Tr}_\omega \left[\pi\big(u(\delta\alpha + \alpha^2)^2 u^*\big)|D|^{-n}\right] \\ &= \mathrm{Tr}_\omega \left[\pi(\delta\alpha + \alpha^2)^2 |D|^{-n}\right] \\ &= I(\alpha). \end{aligned}$$

Now we see $\Omega_D\mathfrak{A} \simeq \pi(\Omega\mathfrak{A})/\pi(\delta J_0)$, and we take a $\alpha \in \Omega^1\mathfrak{A}$ such that $[\pi(\alpha)] = V$ where the class is taken modulo $\pi(\delta J_0)$. We are going to prove that

$$dV + V^2 = [P\pi(\delta\alpha + \alpha^2)]. \tag{80.118}$$

EqDyVppi

First, $dV = [\pi(\delta\alpha)]$, so $dV + V^2 = [\pi(\delta\alpha + \alpha^2)]$. From definition, P projects $\pi(\delta J_0)$ on zero, so from definition of the classes that precisely are modulo δJ_0, we have for any ω, $[\pi(\omega)] = [P\pi(\omega)]$. In particular, for all $\alpha \in \Omega^1\mathfrak{A}$ such that $[\pi(\alpha)] = V$, we have equation (80.118).

□

80.9.4 Fermionic action

Let us first work out the case of the $\mathfrak{A}$-module $\mathcal{E} = \mathfrak{A}$ described in page 4095.

Theorem 80.58.
If $V \in \Omega^1_D\mathfrak{A}$ is a gauge potential (i.e. if $V = V^*$), the quantity

$$I_{Dir}(V, \psi) = \langle \psi, (D + \pi(V))\psi \rangle \tag{80.119}$$

*with $\psi \in \mathrm{dom}(D) \subset \mathscr{H}$ is invariant under the action of $U(\mathfrak{A})$. That quantity is the **fermionic action**.*

Proof. First we have

$$\pi(V^u)(u\psi) = (u[D, u^*] + uVu^*)(u\psi) = uD(u^*u\psi) - uu^*D(u\psi) + u\pi(V)u^*u\psi;$$

using the fact that $uu^* = u^*u = 1$, we find

$$\big(D + \pi(V^u)\big)(u\psi) = D(u\psi) + uD\psi - D(u\psi) + u\pi(V)\psi = u\big(D\psi + \pi(V)\psi\big).$$

The claim now results from the fact that u is an isometry for the inner product. □

Let $\mathfrak{A}$ and $\mathcal{B}$ be two Morita equivalent algebras. If $\mathfrak{A}$ acts on the Hilbert space $\mathscr{H}$, then we define

$$\mathscr{H}' = \mathcal{E} \otimes_{\mathfrak{A}} \mathscr{H}$$

where $\mathcal{E}$ is the module that implement the Morita equivalence. The algebra naturally acts on $\mathscr{H}'$. If $\mathcal{E}$ has an hermitian structure with values in $\mathfrak{A}$ which satisfies the conditions of an inner product (including positivity), then

$$\big(\xi_1 \otimes \psi_1, \xi_2 \otimes \psi_2\big) = \langle \psi_1, \langle \xi_1, \xi_2 \rangle \psi_2 \rangle \tag{80.120}$$

defines an Hilbert space structure on $\mathscr{H}'$. Since $\mathcal{B}$ acts on $\mathscr{H}'$ in the same way as $\mathfrak{A}$ acts on $\mathscr{H}$, we want to define an associated Dirac operator. The most immediate choice is $D'(\xi \otimes_{\mathfrak{A}} \psi) = \xi \otimes_{\mathfrak{A}} D\psi$, but it is not well defined because of the tensor product:

$$\xi a \otimes_{\mathfrak{A}} \psi = \xi \otimes_{\mathfrak{A}} \pi(a)\psi,$$

so that $D'(\xi a \otimes_{\mathfrak{A}} \psi) = \xi \otimes_{\mathfrak{A}} aD\psi$ while $D'(\xi \otimes_{\mathfrak{A}} a\psi) = \xi \otimes_{\mathfrak{A}} D(a\psi)$, but in general $aD\psi \neq D(a\psi)$. In order to build a suitable extension of D to $\mathscr{H}'$, we consider a connection $\nabla \colon \mathcal{E} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1_D(\mathfrak{A})$ and we define

$$D_\nabla(\xi \otimes_{\mathfrak{A}} \psi) = \xi \otimes_{\mathfrak{A}} D\psi + (\nabla\xi)\psi$$

where the second terms has to be understood as $(\nabla\xi)\psi = (\eta \otimes_{\mathfrak{A}} \omega)\psi = \eta \times \pi(\omega)\psi$. In the case where $\mathcal{E} = \mathfrak{A}$, we know from *the simplest example in the world* (page 4095) that a connection reads $d + A$, so that the fluctuation of the Dirac operator reads $D \mapsto D + V$ where V is any potential vector.

Let us now pass to a general hermitian projective finitely generated module $\mathcal{E}$ on $\mathfrak{A}$. For that, we consider the space of **gauged spinors** $\mathcal{E} \otimes_{\mathfrak{A}} \mathscr{H}$. That space is endowed with the inner product

$$(\xi_1 \otimes_{\mathfrak{A}} \psi_1, \xi_2 \otimes_{\mathfrak{A}} \psi_2) = (\psi_1, \langle \xi_1, \xi_2 \rangle \psi_2). \tag{80.121}$$

The action of an element of $\mathrm{End}_{\mathfrak{A}}(\mathcal{E})$ extends to gauges spinors by

$$\phi(\xi \otimes_{\mathfrak{A}} \psi) = \phi(\xi) \otimes_{\mathfrak{A}} \psi$$

for $\phi \in \mathrm{End}_{\mathfrak{A}}(\mathcal{E})$. In particular the action of $U(\mathcal{E})$ is unitary: $\big(u(\xi_1 \otimes \psi_1), u(\xi_2 \otimes \psi_2)\big) = (\xi_2 \otimes \psi_1, \xi_2 \otimes \psi_2)$.

If $\nabla\colon \mathcal{E} \to \mathcal{E} \otimes_{\mathfrak{A}} \Omega^1_D\mathfrak{A}$ is a compatible connection, we define the **gauged Dirac operator**

$$\begin{aligned} D_\nabla\colon \mathcal{E} \otimes_{\mathfrak{A}} \mathscr{H} &\to \mathcal{E} \otimes_{\mathfrak{A}} \mathscr{H} \\ D_\nabla(\xi \otimes \psi) &= \xi \otimes D\psi + \big((\mathbb{1} \otimes \pi)\nabla_{un}\xi\big)\psi \end{aligned} \tag{80.122}$$

where ∇_{un} is any universal connection projecting on ∇, $\pi(\nabla_{un}) = \nabla$. Let $\nabla_{un} = p\delta + \alpha$ with $\alpha \in \mathrm{End}_{\mathfrak{A}}(\mathcal{E}) \otimes_{\mathfrak{A}} \Omega^1\mathfrak{A}$. Using the action (80.9), the second term in the gauged Dirac operator reads

$$\begin{aligned} (\mathbb{1} \otimes \pi)\nabla_{un} p\big(\sum_i e_i \otimes \xi^i\big) &= (\mathbb{1} \otimes \pi)(p\delta + \alpha)\sum_i p(f_i)\xi^i \\ &= (\mathbb{1} \otimes \pi)\big(\sum_i p(f_i) \otimes_{\mathfrak{A}} \delta\xi^i + \alpha(\xi)\big) \\ &= \sum_i p(f_i) \otimes_{\mathfrak{A}} [D, \xi^i] + \pi(\alpha)\xi \\ &= [D, \xi] + \pi(\alpha)\xi \\ &= d\xi + \pi(\alpha)\xi \end{aligned}$$

in which the three last lines are notations. If we apply D_∇ on $\Psi = \sum_j \xi_j \otimes \psi_j$, we have

$$\sum_j \sum_i p(f_i) \otimes_{\mathfrak{A}} [D, \xi^i]\psi_j = \sum_j [D, \xi]\psi_j \in \mathcal{E} \otimes_{\mathfrak{A}} \mathscr{H}$$

and

$$\begin{aligned} \pi(\alpha)\Psi &= \sum_k (\mathbb{1} \otimes \pi)\big(\alpha(\xi)\big)\psi_k \\ &= \sum_{ijk} (\mathbb{1} \otimes \pi)\big(A_i\xi_j \otimes_{\mathfrak{A}} a_0^i \delta a_1^i\big)\psi_k \\ &= \sum_{ijk} A_i\xi_j \otimes_{\mathfrak{A}} a_0^i [D, a_0^i]\psi_k \in \mathcal{E} \otimes \mathscr{H}. \end{aligned}$$

The whole gives

$$\begin{aligned} D_\nabla\Psi &= \xi \otimes D\psi + \big([D, \xi] + \pi(\alpha)\xi\big)\psi \\ &= \xi \otimes D\psi + (d\xi)\psi + \pi(\alpha)\psi \\ &= \big(pD + \pi(\alpha)\big)\Psi \end{aligned}$$

where $D(\xi \otimes \psi) = \xi \otimes D\psi + d\xi\psi$. Remark that D_∇ only depend on $\pi(\alpha)$ and not of the whole choice of ∇_{un}. Since $\pi(\nabla^1_{un}) = \pi(\nabla^2_{un})$ only when $\pi(\alpha^1) = \pi(\alpha^2)$, the gauged operator D_∇ in fact does not depend on the choice of ∇_{un} in $\pi^{-1}(\nabla)$.

Proposition 80.59.
*The **gauged Dirac action** is*

$$I_{Dir}(\nabla, \Psi) = \langle \Psi, D_\nabla\Psi \rangle. \tag{80.123}$$

That action is invariant under the unitary group:

$$I_{Dir}(\nabla^u, u\Psi) = I_{Dir}(\nabla, \Psi)$$

for every $\Psi \in \mathcal{E} \otimes_{\mathfrak{A}} \mathrm{dom}\, D$ and $\nabla \in \mathcal{C}(\mathcal{E})$.

Proof. We know that $\nabla = pd + A$ and $\nabla^u = pd + A^u$. Using the fact that $(1 \otimes \pi)$ commutes with u and $(1 \otimes \pi)\alpha = A$,

$$\begin{aligned} (1 \otimes \pi)\alpha^u &= (1 \otimes \pi)(up\delta u^* + u\alpha u^*) \\ &= u(\mathbb{1} \otimes \pi)(p\delta)u^* + u(\mathbb{1} \otimes \pi)\alpha u^* \\ &= updu^* + uAu^* \\ &= A^u. \end{aligned}$$

Thus, using $up : pu$, we have

$$\big(pD + \pi(\alpha^u)\big)(u\Psi) = \big(pD + \pi(up\delta u^* + u\alpha u^*)\big)(u\Psi) \tag{80.124}$$
$$= (pD)(u\Psi) + updu^*(u\Psi) + u\pi(\alpha)u^*(u\Psi) \tag{80.125}$$
$$= (pD)(u\Psi) + up[D, u^*](u\Psi) + u\pi(\alpha)(\Psi) \tag{80.126}$$
$$= upD(u^*u\Psi) + u\pi(\alpha)\Psi \tag{80.127}$$
$$= u\big(pD + \pi(\alpha)\big)\Psi. \tag{80.128}$$

□

80.10 Forms and physics formalism

80.10.1 Unitary group

We consider a K-cycle $(\mathscr{H}, D)$ over an involutive unital algebra $\mathfrak{A}$. Its **unitary group** is

$$U(\mathfrak{A}) = \{u \in \mathfrak{A} \text{ tel que } u^*u = uu^* = 1\}.$$

For such a u, we have $uDu^* = D + u[D, u^*]$. This shows that D is not invariant under $U(\mathfrak{A})$, but that its variation is of the form $u[D, u^*]$. When one looks at equation (70.9), we want the variation of D to be a vector potential. Hence we take the following definitions. A k**-form** on $\mathfrak{A}$ is an operator on $\mathscr{H}$ of the form

$$\omega^k = \sum_j a_0^j [D, a_1^j] \cdots [D, a_k^j],$$

and a **vector potential** is a self-adjoint 1-form of the form

$$V = \sum_j a_0^j [D, a_1^j] \tag{80.129}$$

with $a_k^j \in \mathfrak{A}$. The action of $u \in U(\mathfrak{A})$ on the operator $D + V$ is

$$D + V \mapsto u(D + V)u^* = uDu^* + uVu^* = D + u[D, u^*] + uVu^*.$$

For that reason, we define the action of u on a potential vector V by

$$V \mapsto \gamma_u(V) = u[D, u^*] + uVu^*,$$

in such a way that $D + V \mapsto D + \gamma_u(V)$. The definition of V is tricky. Because of the ambiguity of the way to write V under the form (80.129). From this form, in order to make the same as in the Fredholm modules associated with a cycle (see equation (80.41)), we want to define $dV = \sum_j [D, a_0^j][D, a_1^j]$. We defer a few the discussion about the ill definiteness of this definition. Let us go on and study the consequence of that definition on the curvature $\theta(V) = dV/V^2$.

If we replace V by $\gamma_u(V)$, we have to compute $\theta\big(\gamma_u(V)\big) = d\big(\gamma_u(V)\big) + \gamma_u(V)^2$. Since $[D, .]$ is a derivation we find

$$\gamma_u(V) = u[D, u^*] + \sum_j ua_0^j[D, a_1^j]u^* = u[D, u^*] + \sum_j ua_0^j\big([D, a_1^j u^*] - a_1^j[D, u^*]\big),$$

so

$$d\big(\gamma_u(V)\big) = [D, u][D, u^*] + \sum_j [D, ua_0^j][D, a_1^j u^*] - \sum_j [D, ua_0^j a_1^j][D, u^*],$$

Proposition 80.60.
We have the equality

$$d\big(\gamma_u(V)\big) + \gamma_u(V)^2 = u(dV + V^2)u^*$$

as operator on $\mathscr{H}$.

Proof. No proof. □

¡¡ Avertissement/question à la lectrice !! 80.61
What is a reduced *algebra?*

As far as we are considering a K-cycle over $\mathfrak{A}$, we can consider the graded universal reduced algebra $\Omega^*(\mathfrak{A})$. In this algebra, we have $\Omega^0(\mathfrak{A}) = \mathfrak{A}$ while $\Omega^1(\mathfrak{A})$ is generated by the symbols da (with $a \in \mathfrak{A}$) and the rules

— $d(ab) = (da)b + a(db)$,
— $d1 = 0$.

If $\alpha \in \mathbb{C}$, these two requirements imply that $d(\alpha a) = (d\alpha)a + \alpha(da) = \alpha(da)$. Hence if we suppose the linearity, we find

$$d(\alpha a + \beta b) = \alpha da + \beta db.$$

Recall that the involution of $\mathfrak{A}$ extends to $\Omega^*(\mathfrak{A})$ by $(da)^* = -da^*$, and the differential is defined by $d(a^0 da^1 \cdots da^n) = da^0 da^1 \cdots da^n$. It fulfills $d^2 = 0$ and $d(\omega_1\omega_2) = (d\omega_1)\omega_2 + (-1)^{\omega_1}\omega_1 d\omega_2$.

80.11 Yang-Mills on the two point space

Let $X = \{a, b\}$ be a two point space; the space of functions on X is identified with $\mathbb{C} \oplus \mathbb{C}$. An element $f \in \mathfrak{A}$ is given by (f_a, f_b). We are going to define a K-cycle on X. First we consider a finite dimensional Hilbert space $\mathscr{H}$ and we consider the following representation of $\mathfrak{A}$ on $\mathscr{H}$:

$$\pi(f) = \begin{pmatrix} f_a \mathbb{1}_a & \\ & f_b \mathbb{1}_b \end{pmatrix}$$

with respect to a direct space decomposition $\mathscr{H} = \mathscr{H}_a \oplus \mathscr{H}_b$. The operator D is constrained by commutators with elements of $\mathfrak{A}$. Since a diagonal part of D would vanish in commutators with $\mathfrak{A}$, we can choose an off-diagonal operator D under the form

$$D = \begin{pmatrix} 0 & M^* \\ M & 0 \end{pmatrix} \tag{80.130}$$

with $M : \mathscr{H}_1 \to \mathscr{H}_2$. For the graduation, one can choose

$$\Gamma = \begin{pmatrix} \mathbb{1}_{\mathscr{H}_1} & \\ & -\mathbb{1}_{\mathscr{H}_2} \end{pmatrix}.$$

The commutator $[D, \mathfrak{A}]$ reads

$$[D, f] = (f_b - f_a) \begin{pmatrix} 0 & M^* \\ -M & 0 \end{pmatrix} \tag{80.131}$$

The fact that D is self-adjoint on a finite dimensional space makes that there exists a basis of eigenvectors and that the biggest eigenvalue of D is the square root of the one of D^2. So if λ is the biggest eigenvalue of $|M| = \sqrt{M^*M}$,

$$\|[D, f]\| = \lambda |f_b - f_a|.$$

Therefore the noncommutative distance between a and b is given by

$$d(a, b) = \sup\{|f_b - f_a| \text{ such that } \lambda|f_b - f_a| \leqslant 1\} = 1/\lambda. \tag{80.132}$$

One can show that the map $\pi : \Omega^1\mathfrak{A} \to \mathcal{L}(\mathscr{H})$ is injective, so that $\Omega^1_D\mathfrak{A} = \Omega^1\mathfrak{A}$ by lemma 80.10. A general element of $\Omega^1_D\mathfrak{A}$ has the form $\lambda ede + \mu(1 - e)de$. Under the representation π we have

$$\pi(e) = \begin{pmatrix} \mathbb{1} & \\ & 0 \end{pmatrix} \qquad \pi(de) = [D, e] = \begin{pmatrix} 0 & -M^* \\ M & 0 \end{pmatrix}$$

$$\pi(ede) = \begin{pmatrix} 0 & -M^* \\ 0 & 0 \end{pmatrix} \qquad \pi\big((1 - e)de\big) = \begin{pmatrix} 0 & 0 \\ M & 0 \end{pmatrix},$$

so that

$$\pi\big(\lambda ede + \mu(1-e)de\big) = \begin{pmatrix} 0 & -\lambda M^* \\ \mu M & 0 \end{pmatrix} \tag{80.133}$$

We are now going to study two examples of modules on $\mathfrak{A}$ and write down the corresponding Yang-Mills theory.

80.11.1 The simplest module

As first example of module, we consider $\mathcal{E} = \mathfrak{A}$. We know a spectral triple and the space of 1-forms $\Omega^1\mathfrak{A}$ on the two point space from 80.1.3. We know from 80.8.10.1 that a potential vector is a self-adjoint element of $\Omega^1_D\mathfrak{A}$; using expression 80.133 we find

$$V = -\overline{\Phi}ede + \Phi(1-e)de \tag{80.134}$$

$$\pi(V) = \begin{pmatrix} 0 & \overline{\Phi}M^* \\ \Phi M & \end{pmatrix}, \tag{80.135}$$

where Φ is any complex number. The curvature is given by $\theta = dV + V^2$ and is computed using the relations

$$ede(1-e) = ede \qquad e(de)e = 0 \qquad (1-e)de(1-e) = 0$$
$$e^2 = e \qquad d(1-e) = -de. \qquad (de)e = (1-e)de$$
$$e(1-e) = 0$$

We find the result

$$\theta(V) = -(\overline{\Phi} + \Phi)dede - \overline{\Phi}\Phi dede \tag{80.136}$$

The general formula $YM(\nabla) = \mathrm{Tr}_\omega\left(\pi(\theta)^2|D|^{-d}\right)$ reduces to

$$YM(V) = \mathrm{Tr}\left(\pi(\theta)^2\right)$$

because the spectral triple is zero-dimensional and the Dixmier trace is the usual trace.

¡¡ Avertissement/question au lecteur !! 80.62
The usual trace is used as Dixmier trace because the Hilbert space is finite dimensional?

We have

$$\pi(dede)^2 = \begin{pmatrix} -M^*M & 0 \\ 0 & -MM^* \end{pmatrix},$$

so that $\mathrm{Tr}\left(\pi(dede)^2\right) = 2\,\mathrm{Tr}(MM^*)^2$, and

$$\pi(\theta) = -\big((\overline{\Phi} + \Phi) + \Phi\overline{\Phi}\big)\begin{pmatrix} 0 & M^* \\ M & 0 \end{pmatrix}.$$

What is interesting is the dependence in Φ in the Yang-Mills action:

$$YM(V) = 2\big(|\Phi + 1|^2 - 1\big)\,\mathrm{Tr}\left((M^*M)^2\right). \tag{80.137}$$

If one sees that as a field equation for Φ, one concludes that the right field is $\Phi + 1$ instead of Φ. For that reason we perform the change of basis $\phi = 1 + \Phi$.

In our simple case $\mathcal{E} = \mathfrak{A}$, an element of $U(\mathcal{E})$ is an element $u \in \mathfrak{A}$ such that $uu^* = u^*u = 1$. If we write $u = u_a e + u_b(1-e)$ with $u_i \in \mathbb{C}$, it is easy to see that the condition is $u_a u_a^* = u_a^* u_a = u_b u_b^* = u_b^* u_b = 1$, so that $U(\mathfrak{A}) = U(1) \times U(A)$.

Let us now apply formula (80.106) on the gauge potential $A = V = (1-\overline{\phi})ede + (\phi - 1)(1-e)de$ and $u = u_a e + u_b(1-e)$. A simple computation shows that

$$V^u = ede - (1-e)de - u_a\overline{u}_b\overline{\phi}ede + u_b\overline{u}_a\phi(1-e)de.$$

Comparing the expression of V with the one of V^u we see that the action of u is to replace $\phi \to u_b\overline{u}_a\phi$.

Chapter 81

Deformations: formal aspects

ChapDefo

81.1 Twists of module (co)algebras

SecTheoryTwist

We are going to follow [891] and use the notions of subsection 54.8.1.

DefTwist

Definition 81.1.
*An element $F \in B \otimes B$ is a **twisting element** based on B if it satisfies the two following conditions*

ItemTwistUn

(1) $(\epsilon_B \otimes \mathrm{Id})F = (\mathrm{Id} \otimes \epsilon_B)F = 1 \otimes 1$

ItemTwistDeux

(2) $\big[(\Delta_B \otimes \mathrm{Id})(F)\big](F \otimes 1) = \big[(\mathrm{Id} \otimes \Delta_B)(F)\big](1 \otimes F)$.

If F is invertible, we have

$$\big[(\Delta_B \otimes \mathrm{Id})F\big]^{-1} = (\Delta_B \otimes \mathrm{Id})F^{-1}. \tag{81.1}$$

EqDelidFinve

In order to see that, first notice that the third component in the tensor product is F_2^{-1} in both sides of (81.1) while the two first components are given by $\Delta_B(F)^{-1}$ in the left hand side and by $\Delta_B(F^{-1})$ in the right hand side. Using the coalgebra properties, we have

$$\Delta_B(F)\Delta_B(F^{-1}) = \Delta_B(FF^{-1}) = \Delta_B(1) = 1 \otimes 1. \tag{81.2}$$

Thus if we take the inverse of the property (2) in the definition of a twist, we get

$$(F^{-1} \otimes 1)\big[(\Delta_B \otimes \mathrm{Id})F^{-1}\big] = (1 \otimes F^{-1})\big[(\mathrm{Id} \otimes \Delta_B)F^{-1}\big]^{-1} \tag{81.3}$$

EqFemuVaLeLem

81.1.1 Twisting the module algebras

Let $\mathbb{A}$ be a B-module algebra with its multiplication $\mu_{\mathbb{A}}$. If F is a twist based on B, we can define the new multiplication

$$\begin{aligned} \mu_F = \mu_{\mathbb{A}} \circ F_l \colon \mathbb{A} \otimes \mathbb{A} &\to \mathbb{A} \\ a \otimes a' &\mapsto \mu_{\mathbb{A}}(ba \otimes b'a')^a \end{aligned} \tag{81.4}$$

if $F = b \otimes b'$. In the same way, if C is a B-module coalgebra with comultiplication $\Delta_C \colon C \to C \otimes C$, we can deform it by

$$\Delta_F = F_r \circ \Delta_C \colon C \to C \otimes C. \tag{81.5}$$

ThoTwistAlgEtCoalg

Theorem 81.2.
Let B be a bialgebra.

(1) If $\mathbb{A}$ is a left B-module algebra, then $\mathbb{A}_F = (\mu_{\mathbb{A}} \circ F_l, 1_{\mathbb{A}})$ is an associative algebra on $\mathbb{K}$.

(2) If C is a B-module coalgebra, then $C_F = (F_r \circ \Delta_C, \epsilon_C)$ is a coassociative $\mathbb{K}$-coalgebra.

Proof. Coassociativity of $\mathbb{A}_F$ means that for every a, b, c in $\mathbb{A}$,

$$(\mu_{\mathbb{A}} \circ F_l)\big(a \otimes (\mu_{\mathbb{A}} \circ F_l)(b \otimes c)\big) = (\mu_{\mathbb{A}} \circ F_l)\big((\mu_{\mathbb{A}} \circ F_l)(a \otimes b) \otimes c\big). \tag{81.6}$$

In other terms,

$$(\mu_{\mathbb{A}} \circ F_l) \circ \big(\mathrm{Id} \otimes (\mu_{\mathbb{A}} \circ F_l)\big) = (\mu_{\mathbb{A}} \circ F_l) \circ (\mu_{\mathbb{A}} \circ F_l) \otimes \mathrm{Id} \tag{81.7}$$

as map from $\mathbb{A} \otimes \mathbb{A} \otimes \mathbb{A}$ to $\mathbb{A}$. If we use the decomposition

$$\mathrm{Id} \otimes (\mu_{\mathbb{A}} \otimes F_l) = (\mathrm{Id} \otimes \mu_{\mathbb{A}}) \circ (\mathrm{Id} \otimes F_l), \tag{81.8}$$

we see that we have to prove the commutativity of the diagram

$$\begin{array}{ccccc}
\mathbb{A} \otimes \mathbb{A} \otimes \mathbb{A} & \xrightarrow{F_l \otimes \mathrm{Id}} & \mathbb{A} \otimes \mathbb{A} \otimes \mathbb{A} & \xrightarrow{\mu_{\mathbb{A}} \circ \mathrm{Id}} & \mathbb{A} \otimes \mathbb{A} \\
\downarrow{\scriptstyle \mathrm{Id} \otimes F_l} & & & & \downarrow{\scriptstyle F_l} \\
\mathbb{A} \otimes \mathbb{A} \otimes \mathbb{A} & & & & \mathbb{A} \otimes \mathbb{A} \\
\downarrow{\scriptstyle \mathrm{Id} \otimes F_l} & & & & \downarrow{\scriptstyle \mu_{\mathbb{A}}} \\
\mathbb{A} \otimes \mathbb{A} & \xrightarrow[F_l]{} & \mathbb{A} \otimes \mathbb{A} & \xrightarrow[\mu_{\mathbb{A}}]{} & \mathbb{A}
\end{array} \tag{81.9}$$

We refer to [891] for the remaining of the proof. □

LemRemakAecP

Lemma 81.3.
If $P \in B \otimes B$ satisfies

$$(P \otimes 1)\big[(\Delta_B \otimes \mathrm{Id})P\big] = (1 \otimes P)\big[(\mathrm{Id} \otimes \Delta_B)(P)\big], \tag{81.10}$$

then P twists the left B-modules coalgebras.

Proof. No proof. □

Notice that, by equation (81.3), F^{-1} satisfies lemma 81.3 when it is invertible.

81.1.2 Twisting the bialgebra itself

Let F be an invertible twist base on the bialgebra B. We can deform B by

$$\Delta'_B = F_l^{-1} \circ F_r \circ \Delta_B. \tag{81.11}$$

Theorem 81.4.
The structure

$$B_F = (\mu_B, \Delta'_B, 1_B, \epsilon_B) \tag{81.12}$$

is a $\mathbb{K}$-bialgebra.

Proof. The second point of theorem 81.2 makes $F_r \circ \Delta_B$ coassociative since F is a twist. In order to check that this is still a B-module coalgebra, we rewrite the diagram (54.137) with B instead of C and $\Delta_B(b)$ by $\Delta_B(b)F$:

$$\begin{array}{ccc}
B \otimes B & \xleftarrow{F_r \circ \Delta} & B \\
\uparrow{\scriptstyle F_r \circ \Delta(b)_r} & & \uparrow{\scriptstyle b_r} \\
B \otimes B & \xleftarrow{F_r \circ \Delta} & B
\end{array} \tag{81.13}$$

We see that the result of this diagram is the one of the non twisted one multiplied by F. For example, the vertical left arrow is the mapping

$$c \otimes c' \mapsto (c \otimes c') \cdot \Delta(b)F. \tag{81.14}$$

Thus the coalgebra B endowed with $\Delta = F_r \circ \Delta_B$ is still a B-module coalgebra. Now, using the lemma 81.3 and the fact that F^{-1} satisfies that lemma, the structure $\Delta'_B = F_l^{-1} \circ F_r \circ \Delta_B$ is coassociative. We check that this structure is compatible with the multiplication:

$$\begin{aligned}\Delta'_B(bb') &= F^{-1}\Delta_B(bb')F\\ &= F^{-1}\Delta_B(b)\Delta_B(b')F\\ &= F^{-1}\Delta_B(b)FF^{-1}\Delta_B(b')F\\ &= \Delta'_B(b)\Delta'_B(b').\end{aligned} \tag{81.15}$$

□

81.1.3 Twisting the left module algebra

Theorem 81.5.
Let $\mathbb{A}$ be a left B-module algebra and F, an invertible twist based on B. Then

$$A_F = (\mathbb{A}, \mu_{\mathbb{A}} \circ F_l, 1_{\mathbb{A}}) \tag{81.16}$$

is a left B-module algebra.

Proof. We already proved that the product in $\mathbb{A}_F$ is associative. Let us now prove that $\mathbb{A}_F$ is a B_F-module. The action of B_F on $\mathbb{A}_F$ is unchanged: $b \cdot a$ if $b \in B_F$ and $a \in \mathbb{A}_F$. The first axiom of definition 54.51, $b \cdot A_{\mathbb{A}} = \epsilon(b) \cdot 1_{\mathbb{A}}$, remains because we didn't change $1_{\mathbb{A}}$ neither ϵ. In the present case, the second condition reads

EqDiagCommdeThobmodalgtwist

$$\begin{array}{ccc} \mathbb{A}\otimes\mathbb{A} & \xrightarrow{\mu_{\mathbb{A}}\circ F_l} & \mathbb{A} \\ {\scriptstyle \left(F^{-1}\Delta_B(b)F\right)_l}\downarrow & & \downarrow{\scriptstyle b_l} \\ \mathbb{A}\otimes\mathbb{A} & \xrightarrow[\mu_{\mathbb{A}}\circ F_l]{} & \mathbb{A} \end{array} \tag{81.17}$$

Let us consider $a \otimes a' \in \mathbb{A} \otimes \mathbb{A}$ and follow its evolution when we apply $b_l \circ F_l \circ \mu_{\mathbb{A}}$.

$$\begin{aligned} a\otimes a' \overset{F_l}{\to} F_1a\otimes F_2a' \overset{\Delta_B(b)}{\to} \sum b_{(1)}F_1a\otimes b_{(2)}F_2a' \overset{F^{-1}}{\to} \sum F_1^{-1}b_{(1)}F_1a\otimes F_2^{-1}b_{(2)}F_2a' \\ \overset{F_l}{\to} \sum b_{(1)}F_1a\otimes b_{(2)}F_1a' \overset{\mu_{\mathbb{A}}}{\to} \mu_{\mathbb{A}}\big(\overset{F_l}{\to}\sum b_{(1)}F_1a\otimes b_{(2)}F_1a'\big). \end{aligned} \tag{81.18}$$

If we follow the other arrows, we find

$$a\otimes a' \overset{F_l}{\to} F_1a\otimes F_2a' \overset{\mu_{\mathbb{A}}}{\to} \mu_{\mathbb{A}}\big(F_1a\otimes F_2a'\big) \overset{b_l}{\to} b\mu_{\mathbb{A}}\big(F_1a\otimes F_2a'\big). \tag{81.19}$$

Now, the commutativity of the untwisted diagram states that

$$b\mu_{\mathbb{A}}(a\otimes a) = \mu_{\mathbb{A}}\big(\sum b_{(1)}a\otimes b_{(2)}a'\big). \tag{81.20}$$

Thus the commutativity of diagram (81.17) is nothing else than the commutativity of the untwisted diagram taken with $F_1a\otimes F_2a'$ instead of $a\otimes a'$. □

81.2 Twist and star product

Let $\mathbb{B}$ be a connected Lie group with Lie algebra $\mathfrak{b}$ and universal enveloping algebra $\mathcal{U}(\mathfrak{b})$.

Definition 81.6.
If Δ and ϵ are the usual co-product and co-unit[1] *on $\mathcal{U}(\mathfrak{b})$, then a **twist based on** based on $\mathcal{U}(\mathfrak{b})$ is an element $F \in \mathcal{U}(\mathfrak{b})[[h]] \otimes \mathcal{U}(\mathfrak{b})[[h]]$ such that*

subEqHopfun

$$(\epsilon\otimes \mathrm{Id})F = 1\otimes 1 = (\mathrm{Id}\otimes\epsilon)F \tag{81.21a}$$

subEqHopfcocs

$$\big[(\Delta\otimes\mathrm{Id})(F)\big](F\otimes 1) = \big[(\mathrm{Id}\otimes\Delta)(F)\big](1\otimes F). \tag{81.21b}$$

1. See definition 72.126.

Remark in that definition the difference between "Id" which is the identity on $\mathcal{U}(\mathfrak{b})$ and "1" which is the constant function, and then the identity on $C^\infty(\mathbb{B})$.

Definition 81.7.
A formal ***universal deformation formula*** *based on $\mathcal{U}(\mathfrak{g})$ is a twisting element F based on $\mathcal{U}(\mathfrak{b})[[h]]$ which reads*

$$F = 1\otimes 1 + hF_1 + h^2F_2 + \cdots + h^nF_n + \dots \tag{81.22}$$

where $F_i \in \big(\mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b})\big)[[h]]$.

Theorem 81.8.
The data of a formal universal deformation formula on $\mathcal{U}(\mathfrak{b})[[h]]$ is equivalent to the data of a left invariant star product $\star$ on $\mathbb{B}$.

The correspondence is as follows. Let $\star = \sum_k h^k C_k$ where C_k are left invariant bidifferential operators on $\mathbb{B}$. So

$$F = \sum_k h^k C_k \tag{81.23}$$

is an element of $\mathcal{U}(\mathfrak{b})[[h]]\otimes\mathcal{U}(\mathfrak{b})[[h]]$.

In that setting, F is a Drinfel'd twist and every Drinfel'd twists are produced in that way.

Proof. Let us make the correspondence more explicit. From proposition 62.54, we have

$$\text{bi-Diff}^{\mathbb{B}}(\mathbb{B}) \simeq \mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b}), \tag{81.24}$$

the elements C_k can be seen in $\mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b})$ and we define the F_k by $C_k = F_k^L$ where $T^L \in \text{bi-Diff}^{\mathbb{B}}(\mathbb{B})$ is the left invariant operator associated with $T \in \mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b})$. Then one defines

$$F = \sum_k t^k F_k \in \big(\mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b})\big)[[t]]. \tag{81.25}$$

What we have to proof is that the so defined F is a twist based on $\mathcal{U}(\mathfrak{b})$. We are going to prove that the associativity of $\star$ is equivalent to the condition (2) while the condition $f\star 1 = 1\star f = f$ is equivalent to the condition (1) in definition 81.1.

Consider $T \in \mathcal{U}(\mathfrak{b})\otimes\mathcal{U}(\mathfrak{b})$ as $T = \sum_i X_i\otimes Y_i$ where $X_i, Y_i \in \mathcal{U}(\mathfrak{b})$ and the sum is finite. Associativity of the product means that

$$\tilde{T}\circ(\tilde{T}\otimes \mathrm{Id}) = \tilde{T}\circ(\mathrm{Id}\otimes\tilde{T}). \tag{81.26}$$

In our computations we are going to use the following rules:

ItemOptimRuleDeux

(1) $(X\otimes Y)(f\otimes g) = Xf\otimes Yg$,

(2) $\widetilde{(X\otimes Y)}(f\otimes g) = (\tilde{X}g)(\tilde{Y}g) \in C^\infty(\mathbb{B})$,

(3) $(\tilde{X}\otimes\tilde{Y})(f\otimes g) = \tilde{X}f\otimes\tilde{Y}g \in C^\infty(\mathbb{B})\otimes C^\infty(\mathbb{B})$

(4) $\widetilde{(X\cdot Y)}f = \tilde{X}(\tilde{Y}f)$ where the dot denotes the product in $\mathcal{U}(\mathfrak{b})$.

with, as usual, the definition $(\tilde{X}f)(x) = \tilde{X}_x f \in \mathbb{R}$. As far as the product in $\mathcal{U}(\mathfrak{b})$ is concerned, we have

$$\begin{aligned}\big[(X_1\otimes Y_1\otimes Z_1)\cdot(X_2\otimes Y_2\otimes Z_2)\big]\tilde{}\,(f\otimes g\otimes h) &= \big[X_1\cdot X_2\otimes Y_1\cdot Y_2\otimes Z_1\cdot Z_2\big]\tilde{}\,(f\otimes g\otimes h)\\ &= (X_1X_2)\tilde{}\,(f)(Y_1Y_2)\tilde{}\,(g)(Z_1Z_2)\tilde{}\,(h)\end{aligned} \tag{81.27}$$

where the dot $\cdot$ denotes the product in $\mathcal{U}(\mathfrak{b})$.

Using these rules we have

$$\begin{aligned}\tilde{T}\circ(\tilde{T}\otimes 1)(f\otimes g\otimes h) &= \sum_{ij}(\tilde{X}_i\otimes\tilde{Y}_j)\circ\Big((\tilde{X}_i\otimes\tilde{Y}_j)\otimes 1\Big)(f\otimes g\otimes h)\\ &= \sum_{ij}(\tilde{X}_i\otimes\tilde{Y}_i)\Big((\tilde{X}_j\otimes\tilde{Y}_j)(f\otimes g)\otimes h\Big)\\ &= \sum_{ij}(\tilde{X}_i\otimes\tilde{Y}_i)\Big((\tilde{X}_j f)(\tilde{Y}_j g)\otimes h\Big)\\ &= \sum_{ij}\tilde{X}_i\Big((\tilde{X}_j f)(\tilde{Y}_j g)\Big)\tilde{Y}_i h.\end{aligned} \tag{81.28}$$

EqUneCyclTcondass

Using the fact that Δ is the usual coproduct on $\mathcal{U}(\mathfrak{b})$ and the formula (62.107), the last line equals

$$\begin{aligned}\sum_{ij}\widetilde{\Delta(x_i)}\big((\tilde{X}_j f)\otimes(\tilde{Y}_j g)\big)\tilde{Y}_i h &= \sum_{ij}\big(\widetilde{\Delta(X_i)}\otimes\tilde{Y}_i\big)\Big((\tilde{X}_j f\otimes\tilde{Y}_j g)\otimes h\Big)\\ &= \sum_{ij}\Big[(\Delta\otimes\mathrm{Id})(X_i\otimes Y_i)\Big]^{\sim}\Big((\tilde{X}_j f\otimes\tilde{Y}_j g)\otimes h\Big)\\ &= \sum_{j}\Big[(\Delta\otimes\mathrm{Id})T\Big]^{\sim}\circ\Big[X_j\otimes Y_j\otimes 1\Big]^{\sim}(f\otimes g\otimes h)\\ &= \Big[(\Delta\otimes\mathrm{Id})T\Big]^{\sim}\circ\Big[T\otimes 1\Big]^{\sim}(f\otimes g\otimes h)\\ &= \Big[(\Delta\otimes\mathrm{Id})(T)\cdot(T\otimes 1)\Big]^{\sim}(f\otimes g\otimes h).\end{aligned} \tag{81.29}$$

EqExpiiPqsDex

Equating the last line with the left hand side of (81.28), we get

$$\tilde{T}\circ(\tilde{T}\otimes 1) = \Big[(\Delta\otimes I)\cdot(T\otimes 1)\Big]^{\sim} \tag{81.30}$$

Doing the same for each power of t, we get the same equation for F instead of T. The same way, we also get

$$F^L\circ(I\otimes F^L) = \big[(I\otimes\Delta)(F)\cdot(1\otimes F)\big]^L, \tag{81.31}$$

so that the associativity of $\star$ is equivalent to the Hopf cocycle condition (81.21b).

Now, we prove that the fact that 1 is an unit for the star product is equivalent to the condition (81.21a). Associativity means that for every function f we have

$$F^L(f\otimes 1) = F^L(1\otimes f) = f. \tag{81.32}$$

Using the counit on $\mathcal{U}(\mathfrak{b})$ (given by item (4) on page 3253), we have

$$\begin{aligned}T^L(f\otimes 1) &= \sum_i(X_i^L f)(Y_i^L 1)\\ &= \sum_i(X_i^L f)\epsilon(Y_i)\\ &= (I\otimes\epsilon)(T)^L f.\end{aligned} \tag{81.33}$$

By the same computation,

$$T^L(1\otimes f) = (\epsilon\otimes I)(T)^L f. \tag{81.34}$$

Thus the equality $1\star f = f\star 1$ is equivalent to $(I\otimes\epsilon)T = (\epsilon\otimes I)(T)$. Now if we want $\sum_i\epsilon(X_i)Y_i^L f$ to be equal to f for every f, we need $Y_i = 1$ whenever $X_i = 1$. Thus T has to be of the form

$$T = 1\otimes 1 + \sum_i X_i\otimes Y_i \tag{81.35}$$

where none of the X_i and Y_i are 1. In other words, the only term containing 1 is the term $1\otimes 1$. □

81.3 Formal Extension lemma

SecExtenLemK

The *extension lemme* was already presented in [1021, 1020]. A non-formal version is given in section 83.3; here we follow the presentation of [892] and we give here more details from the formal twist point of view.

Let $\mathbb{B}_1$ and $\mathbb{B}_2$ be two Lie groups and let us consider the direct product $\mathbb{B} = \mathbb{B}_1 \times_R \mathbb{B}_2$. In particular, $\mathbb{B} \simeq \mathbb{B}_1 \times \mathbb{B}_2$ as manifold, $\mathbb{B}_1$ is normal in $\mathbb{B}$ and $\mathbb{B}_1 \cap \mathbb{B}_2 = \{e\}$. The extension map R is given by

$$R_x(y) = xyx^{-1} \tag{81.36}$$

for every $x \in \mathbb{B}_1$ and $y \in \mathbb{B}_2$.

Let now $\mathfrak{b}$, $\mathfrak{b}_1$ and $\mathfrak{b}_2$ be the respective Lie algebras and consider

$$\rho\colon \mathfrak{b}_2 \to \mathtt{Der}(\mathfrak{b}_1) \tag{81.37}$$

be the differential of R. At the Lie algebra level we have $\rho_X(Y) = [X, Y]$, but it can be extended to the universal enveloping algebra by the formula

$$\rho_X(Y) = X^{(1)} \cdot Y \cdot S(X^{(2)}) \tag{81.38}$$

where S is the antipode in $\mathcal{U}(\mathfrak{b}_2)$ and $\Delta(X) = X^{(1)} \otimes X^{(2)}$.

¡¡ Avertissement/question au lecteur !! 81.9
Could be great to have a confirmation of that.

We have an action

$$\begin{aligned} r\colon \mathbb{B}_2 &\to \mathrm{Aut}(\mathfrak{b}_1) \\ r_h(X) &\mapsto \frac{d}{dt}\Big[R_h(e^{tX})\Big]_{t=0} \end{aligned} \tag{81.39}$$

for every $h \in \mathbb{B}_2$ and $X \in \mathfrak{b}_1$. This action extends to $\mathcal{U}(\mathfrak{b}_1)$.

Lemma 81.10.
The action R can be retrieved from r by the formula

$$R_h(e^X) = e^{r_h(X)}. \tag{81.40}$$

Proof. By definition, $e^{r_h(X)} = e^{(dR_h)_e X} = R_h(e^X)$. □

The map R_h can be extended to $\mathcal{U}(\mathfrak{b}_1)$ by

$$r_h(X \cdot Y) = \frac{d}{dt}\frac{d}{ds}\Big[R_h(e^{tX})R_h(e^{sY})\Big]_{\substack{s=0\\t=0}} = \frac{d}{dt}\frac{d}{ds}\Big[he^{tX}e^{sY}h^{-1}\Big]_{\substack{s=0\\t=0}} \tag{81.41}$$

Let now denote by $\mathcal{U}(\mathfrak{b}_1)^{\mathbb{B}_2}$ the set of the elements of $\mathcal{U}(\mathfrak{b}_2)$ that are fixed by the action r, that is the elements $X \in \mathcal{U}(\mathfrak{b}_1)$ such that $r_h(X) = X$ for every $h \in \mathbb{B}_2$.

Lemma 81.11.
Let $X \in \mathcal{U}(\mathfrak{b}_1)$ such that $r_h(X) = X$ and $Y \in \mathfrak{b}_2$. Thus $X \cdot Y = Y \cdot X$.

Proof. The condition $r_h(X) = X$ means that $\frac{d}{dt}\Big[he^{tX}h^{-1}\Big]_{t=0} = X$ for every $h \in \mathbb{B}_2$. Let us take a path $Y(s)$ in $\mathbb{B}_2$ and derive the equation $r_{Y(s)}X = X$ with respect to s:

$$0 = \frac{d}{dt}\frac{d}{ds}\Big[Y(s)e^{tX}Y(s)^{-1}\Big]_{\substack{s=0\\t=0}} = \frac{d}{dt}\frac{d}{ds}\Big[\mathbf{Ad}\left(Y(s)\right)e^{tX}\Big]_{\substack{s=0\\t=0}} = [Y, X] \tag{81.42}$$

□

Lemma 81.12.
We have $\mathrm{Diff}^{\mathbb{B}_1\times\mathbb{B}_1,\mathbb{B}_2\times\mathbb{B}_2}(\mathbb{B}_1 \times \mathbb{B}_1) \simeq \text{bi-Diff}^{\mathbb{B}_1,\mathbb{B}_2}(\mathbb{B}_1)$.

Proof. For the proof we show that both of the two sets can be identified with $\mathcal{U}(\mathfrak{b}_1)^{\mathfrak{b}_2} \otimes \mathcal{U}(\mathfrak{b}_1)^{\mathfrak{b}_2}$.

The $\mathbb{B}_2 \times \mathbb{B}_2$-invariance of an operator $P \in \mathrm{Diff}^{\mathbb{B}_1 \times \mathbb{B}_1}(\mathbb{B}_1 \times \mathbb{B}_1)$ means that for every $u \in C^\infty(\mathbb{B}_1 \times \mathbb{B}_1)$ and $(b_2, b_2') \in \mathbb{B}_2 \times \mathbb{B}_2$ we have

$$(Pu)(b_2 b_1, b_2' b_1') = P\big(L_{(b_2, b_2')} u\big)(b_1, b_1'). \tag{81.43}$$

Thus for each $t, s \in \mathbb{R}$, $Z_2, Z_2' \in \mathfrak{b}_2$ we have

$$(Pu)\big(e^{tZ_2} b_1, e^{sZ_2'} b_1'\big) = P\big(L_{(e^{tZ_2}, e^{sZ_2'})} u\big)(b_1, b_1') \tag{81.44}$$

If P corresponds to $X \otimes Y \in \mathcal{U}(\mathfrak{b}_1) \otimes \mathcal{U}(\mathfrak{b}_1)$, the derivative with respect to t and s yields

$$\big[(Z_2 \otimes Z_2') \cdot (X \otimes Y) u\big](b_1, b_2) = \big[(X \otimes Y) \cdot (Z_2 \otimes Z_2') u\big](b_1, b_2), \tag{81.45}$$

that is the fact that Z_2 and Z_2' respectively commute with X and Y. Thus the elements X and Y have to belong to $\mathcal{U}(\mathfrak{b}_1)^{\mathfrak{b}_2}$.

Let now $c \in \text{bi-Diff}^{\mathbb{B}_1}(\mathbb{B}_1)$ be given by

$$c(u \otimes v)(b_1) = \sum_{ab} \tilde{X}^a_{b_1}(u) \tilde{X}^b_{b_1}(v) \tag{81.46}$$

with $X^a, X^b \in \mathcal{U}(\mathfrak{b}_1)$. Now we want to impose it to be $\mathbb{B}_2$-invariant for the action $R_y(x) = yxy^{-1}$. So

$$(R_y^* c)(u \otimes v) = c\big(R_y^* u \otimes R_y^* v\big), \tag{81.47}$$

or more explicitly

$$c(u \otimes v)(yxy^{-1}) = \sum_{ab} \big(\tilde{X}^a_{yxy^{-1}} u\big)\big(\tilde{X}^b_{yxy^{-1}} v\big) \stackrel{!}{=} \sum_{ab} \tilde{X}^a_x (R_y^* u) \tilde{X}^b_x (R_y^* v). \tag{81.48}$$

EqCuvexpliny

On the one hand

$$\begin{aligned} \tilde{X}^a_x (R_y^* u) &= \frac{d}{dt}\Big[(R_y^* u)(x e^{tX})\Big]_{t=0} \\ &= \frac{d}{dt}\Big[u\big(y x e^{tX} y^{-1}\big)\Big]_{t=0}, \end{aligned} \tag{81.49}$$

and on the other hand

$$\tilde{X}^a_{yxy^{-1}} u = \frac{d}{dt}\Big[u\big(yxy^{-1} e^{tX}\big)\Big]_{t=0}. \tag{81.50}$$

If we consider the relation (81.48) at $x = e$ we find the condition $e^{tX} y = y e^{tX}$ for every y. Taking $y = e^{sY}$ and taking the derivative with respect to t and s we find the relation $[X, Y] = 0$. □

Consider the extension

$$\mathfrak{b} = \mathfrak{b}_2 \times_\rho \mathfrak{b}_1 \tag{81.51}$$

and the associated group extension

$$\mathbb{B} = \mathbb{B}_2 \times_R \mathbb{R}_1 \tag{81.52}$$

with $R \colon \mathbb{B}_2 \to \mathrm{Aut}(\mathbb{B}_1)$.

When $\mathfrak{A}$ and $\mathcal{B}$ are C^*-algebra, the elements of $\mathfrak{A} \otimes \mathcal{B}$ are limits of sums of the form $\sum_i a_i \otimes b_i$ with $a_i \in \mathfrak{A}$ and $b_i \in \mathcal{B}$. In the case with $\mathfrak{A} = C^\infty(A)$ and $\mathcal{B} = C^\infty(B)$, we have

$$C^\infty(A) \otimes C^\infty(B) \simeq C^\infty(A \times B) \tag{81.53}$$

EqCABsimeqCACB

from section 72.17.

The product on $\mathbb{B}$ is given by

$$(b_2, b_1) \cdot (b_2', b_1') = \big(b_2 b_2', b_1 \cdot R(b_2) b' 1\big). \tag{81.54}$$

At the level of the regular left representation, we have

$$L^*_{(b_2, b_1)} = L^*_{b_2} \otimes L^*_{b_1} \circ R(b_2). \tag{81.55}$$

Eq1507Lbdb1LoL

For sake of compactness we write

$$A = L_{b_1} \circ R(b_2). \tag{81.56}$$

The identification (81.53) allows us to give a sense to the action of $L^*_{(b_2,b_1)}$ on elements like $(a \otimes u) \otimes (b \otimes v)$. First, the expression $(a \otimes u) \otimes (b \otimes v)$ is associated to $(a \cdot b) \otimes (u \cdot v)$ and (81.55) allows us to write

$$\begin{aligned} L^*_{(b_2,b_1)}\big(a \cdot b \otimes u \cdot v\big) &= L^*_{b_2}(a \cdot b) \otimes \big(L_{b_1} \circ R(b_2)\big)^*(u \cdot v) \\ &= (L^*_{b_2}a) \cdot (L^*_{b_2}b) \otimes (A^*u) \cdot (A^*v) \\ &= \big(L^*_{b_2}a \otimes A^*u\big) \otimes \big(L^*_{b_2}b \otimes A^*v\big), \end{aligned} \tag{81.57}$$

where the last equality is the identification (81.53) in the reverse sense. Thus we *define* that $L^*_{(b_2,b_1)}$ acts on $\Big(C^\infty(\mathbb{B}_2) \otimes C^\infty(\mathbb{B}_1)\Big) \otimes \Big(C^\infty(\mathbb{B}_2) \otimes C^\infty(\mathbb{B}_1)\Big)$ as

$$L^*_{(b_2,b_1)}(a \otimes u) \otimes (b \otimes v) = \big(L^*_{b_2}a \otimes A^*u\big) \otimes \big(L^*_{b_2}b \otimes A^*v\big) \tag{81.58}$$

Eq1507LsucBBBB

where $A = L_{b_1} \circ R(b_2)$.

Consider the multiplications μ^1 on $C^\infty(\mathbb{B}_1)$, μ^2 in $C^\infty(\mathbb{B}_2)$ and μ^{12} on $C^\infty(\mathbb{B}_2 \times \mathbb{B}_1)$.

Lem1607mualalmu

Lemma 81.13.
We have $\mu^{21} \circ (L_{b_2} \otimes A)^* \otimes (L_{b_2} \otimes A)^* = (L_{b_2} \otimes A)^* \circ \mu^{21}$.

Proof. The result comes essentially that A and $L^*_{b_2}$ commute with the pointwise product of functions:

$$\begin{aligned} A^*(u \cdot v) &= A^*u \cdot A^*v \\ L^*_{b_2}(a \cdot b) &= (L^*_{b_2}a) \cdot (L^*_{b_2}b). \end{aligned} \tag{81.59}$$

We have

$$\begin{aligned} (L_{b_2} \otimes A)^*\mu^{21}(a \otimes u) \otimes (b \otimes v) &= (L_{b_2} \otimes A)^*(a \cdot b \otimes u \cdot v) \\ &= \mu^{21}(L^*_{b_2}a \otimes A^*u) \otimes (L^*_{b_2}b \otimes A^*v) \\ &= \mu^{21} \circ \Big((L_{b_2 j \otimes A})^* \otimes (L_{b_2} \otimes A)^*\Big)(a \otimes u) \otimes (b \otimes v). \end{aligned} \tag{81.60}$$

□

The same kind of result.

Lem1607LmumuLotimes

Lemma 81.14.
We have

$$L^*_{(b_2,b_1)} \circ \mu^{21} = \mu^{21} \circ \big(L^*_{(b_2,b_1)} \otimes L^*_{(b_2,b_1)}\big). \tag{81.61}$$

Proof. If $f \in C^\infty(\mathbb{B}_2)$ and $g \in C^\infty(\mathbb{B}_1)$, seen as functions on $£B_2 \times \mathbb{B}_1$ which depend of only one variable, we have

$$\begin{aligned} L^*_{(b_2,b_1)} \circ \mu^{21}(f \otimes g)(b'_2, b'_1) &= L^*_{(b_2,b_1)}(f \cdot g)(b'_2, b'_1) \\ &= (L^*_{(b_2,b_1)}f)(b'_2, b'_1)(L^*_{(b_2,b_1)}g)(b'_2, b'_1) \\ &= \mu^{21}\big(L^*_{(b_2,b_1)} \otimes L^*_{(b_2,b_1)}g\big)(b'_2, b'_1) \\ &= \mu^{21} \circ (L^*_{(b_2,b_1)} \otimes L^*_{(b_2,b_1)})(f \otimes g)(b'_2, b'_1). \end{aligned} \tag{81.62}$$

□

A bidifferential operator on $C^\infty(\mathbb{B}_1)$ reads

$$\tilde{P} = \mu^1 \circ \sum_i \tilde{P}'_i \otimes \tilde{P}''_i \tag{81.63}$$

with $P_i', P_i'' \in \mathcal{U}(\mathfrak{b}_1)$. From $\tilde{P}$ we define a bidifferential operator on $C^\infty(\mathbb{B}_2)\otimes C^\infty(\mathbb{B}_1)$ by

$$\tilde{P}^{(1)} = \mu^{21}\circ(\mathrm{Id}\otimes\tilde{P}')\otimes(\mathrm{Id}\otimes\tilde{P}''). \tag{81.64}$$

That operator acts on (sums of) elements of the type $(u_1\otimes v_1)\otimes(u_2\otimes v_2)$ with $u_i\in C^\infty(\mathbb{B}_2)$ and $v_i\in C^\infty(\mathbb{B}_1)$ with

$$(\mathrm{Id}\otimes\tilde{P}'')(u_2\otimes v_2) = u_2\otimes\tilde{P}''v_2 \in C^\infty(\mathbb{B}_2)\otimes C^\infty(\mathbb{B}_1). \tag{81.65}$$

Thus

$$P^{(1)}(u_1\otimes v_1)\otimes(u_2\otimes v_2) = \mu^{21}(u_1\otimes\tilde{P}v_1)\otimes(u_2\otimes\tilde{P}''v_2) = u_1u_2\otimes\tilde{P}'v_1\tilde{P}''v_2. \tag{81.66}$$

In the same way, to every bidifferential operator $\tilde{L}$ we associate the bidifferential operator $\tilde{L}^{(2)}$ on $C^\infty(\mathbb{B}_2\otimes\mathbb{B}_1)$ defined by

$$\tilde{L}^{(2)} = \mu^{21}\circ(\tilde{L}'\otimes\mathrm{Id})\otimes(\tilde{L}''\otimes\mathrm{Id}). \tag{81.67}$$

where a summation is understood.

The following proposition (as everything here) comes from [892].

PropPBBPtildeComm

Proposition 81.15.
Let $P\in\mathcal{U}(\mathfrak{b}_1)\otimes\mathcal{U}(\mathfrak{b}_1)$ and $\tilde{P}$ the associated left invariant bidifferential operator on $\mathbb{B}_1$.

ItemPropPBBPtildeCommi

(1) For every $g\in\mathbb{B}_2$ we have $R(g)\circ\tilde{P} = \tilde{P}\circ R(g)\otimes R(g)$ if and only if $P\in\big(\mathcal{U}(\mathfrak{b}_1)\otimes\mathcal{U}(\mathfrak{b}_1)\big)^{\mathfrak{b}_2}$. The latter subalgebra of $\mathcal{U}(\mathfrak{b}_1)\otimes\mathcal{U}(\mathfrak{b}_1)$ being formed by the elements $X_1\otimes Y_1$ such that $\rho(Z)(X_1\otimes Y_1)=0$ for every $Z\in\mathfrak{b}_2$.

ItemPropPBBPtildeCommii

(2) For every $P\in\mathcal{U}(\mathfrak{b}_1)^{\mathfrak{b}_2}\otimes\mathcal{U}(\mathfrak{b}_1)^{\mathfrak{b}_2}$, the associated left invariant operator $P^{(1)}$ on $C^\infty(\mathbb{B}_2\times\mathbb{B}_1)$ is $\mathbb{B}$-invariant. In other words we have

$$L^*_{b_2,b_1}\circ\tilde{P}^{(1)} = \tilde{P}^{(1)}\circ L^*_{b_2,b_1} \tag{81.68}$$

in the sense of the identification (81.53).

ItemPropPBBPtildeCommiii

(3) If $\tilde{L}$ is a left invariant bidifferential operator on $C^\infty(\mathbb{B}_2)$, the associated bidifferential operator $\tilde{L}^{(2)}$ on $C^\infty(\mathbb{B})$ is $\mathbb{B}$-invariant, that is

$$L^*_{(b_2,b_1)}\circ\tilde{L}^{(2)} = \tilde{L}^{(2)}\circ L^*_{(b_2,b_1)}. \tag{81.69}$$

Proof. Let $P = X\otimes Y$ with $X,Y\in\mathfrak{b}_1$. First we have $R(g^{-1})\otimes R(g^{-1}) = R(g^{-1})X\otimes R(g^{-1})Y$, but by definition we have

$$R(g)X = e^{\rho(Z)}X \tag{81.70}$$

and $e^{\rho(Z)}X = X$ if and only if $\rho(Z)X = 0$.

¡¡ Avertissement/question au lecteur !! 81.16
Because $e^{tZ}X$ is a polynomial in t which can only be equal to X when the coefficients of all the powers of t are zero??

Let us now prove that

$$R(g)\circ\tilde{P} = \big[R(g^{-1})\otimes R(g^{-1})P\big]^{\tilde{}}\circ R(g)\otimes R(g), \tag{81.71}$$

which is equivalent to the property (1) because of the property we just mentioned. We begin with $P = X_1\otimes X_2\in\mathfrak{b}_1\otimes\mathfrak{b}_2$. We write $X = X_1\otimes X_2$.

Using the definitions,

$$\begin{aligned}\big(R(g)\circ\tilde{P}\big)(u\otimes v)(x) &= \tilde{X}(u\otimes v)\big(R(g)x\big)\\ &= \frac{d}{dt}\Big[u\big(R(g)(x)e^{tX_1}\big)\Big]_{t=0}\frac{d}{dt}\Big[v\big(R(g)(x)e^{tX_2}\big)\Big]_{t=0}.\end{aligned} \tag{81.72}$$

EqRgxtXunRgpT

Since $R_g \in \mathrm{Aut}(\mathbb{B}_1)$, we have $(R_g x)y = R_g(xR_{g^{-1}}y)$, so that the right hand side of (81.72) becomes

$$\begin{aligned}\frac{d}{dt}\Big[u\big(R_g(xR_{g^{-1}}e^{tX_1})\big)\Big]_{t=0}\frac{d}{dt}\Big[\big(R_g(xR_{g^{-1}}e^{tX_2})\big)\Big]_{t=0} \\ = \frac{d}{dt}\Big[(R_g^*u)\big(xR_{g^{-1}}e^{tX_1}\big)\Big]_{t=0}\frac{d}{dt}\Big[(R_g^*v)\big(xR_{g^{-1}}e^{tX_2}\big)\Big]_{t=0}.\end{aligned} \tag{81.73}$$

EqDsdduRgXuvgv

For sake of shortness, we write $R_gY = \frac{d}{dt}\Big[R_ge^{tY}\Big]_{t=0}$, so that we have

$$\frac{d}{dt}\Big[w\big(xR_{g^{-1}}e^{tX}\big)\Big]_{t=0} = (\widetilde{R_{g^{-1}}X})_x(w) = (\widetilde{R_{g^{-1}}X})(w)(x), \tag{81.74}$$

and (81.73) becomes

$$\begin{aligned}(\widetilde{R_{g^{-1}}X_1})(R_g^*u)(x)(\widetilde{R_{g^{-1}}X_2})(R_g^*v)(x) &= \Big[(R_{g^{-1}}\otimes R_{g^{-1}})(X_1\otimes X_2)\Big]^{\sim}\big(R_g^*u\otimes R_g^*v\big)(x) \\ &= \Big[(R_{g^{-1}}\otimes R_{g^{-1}})X\Big]^{\sim}\circ(R_g\otimes R_g)(u\otimes v)(x).\end{aligned} \tag{81.75}$$

This concludes the proof of the point (1).

For point (2), we take $a_i, b_j \in C^\infty(\mathbb{B}_2)$ and $u_i, v_j \in C^\infty(\mathbb{B}_1)$ and we apply $L^*_{(b_2,b_1)}\circ\tilde{P}^{(1)}$ to the combination $\big(\sum_i a_i\otimes u_i\big)\otimes\big(\sum_j b_j\otimes v_j\big)$. By linearity, we restrict ourself to only one term and we denote by a single dot the pointwise function product:

$$\begin{aligned}\big(L^*_{(b_2,b_1)}\circ\tilde{P}^{(1)}\big)\big((a\otimes u)\otimes(b\otimes v)\big) = \\ L^*_{(b_2,b_1)}\circ\mu^{21}\big((a\otimes\tilde{P}'u)\otimes(b\otimes\tilde{P}''v)\big) \\ = L^*_{(b_2,b_1)}\big(a\cdot b\otimes\tilde{P}'u\cdot\tilde{P}''v\big) \\ = L^*_{b_2}(a\cdot b)\otimes\big(L_{b_1}\circ R(b_2)\big)^*(\tilde{P}'u\cdot\tilde{P}''v) \\ = (L^*_{b_2}a)\cdot(L^*_{b_2}b)\otimes\big(L_{b_1}\circ R(b_2)\big)^*\tilde{P}'u\cdot\big(L_{b_1}\circ R(b_2)\big)^*\tilde{P}''v \\ = \mu^{12}\Big[(L^*_{b_2}a)\otimes\big(L_{b_1}\circ R(b_2)\big)^*\tilde{P}'u\Big]\otimes\Big[L^*_{b_2}b\otimes(L_{b_1}\circ R(b_2))^*\tilde{P}''v\Big].\end{aligned} \tag{81.76}$$

EqLnnmuLBPP1507

What lies in the first bracket is

$$\Big(L^*_{b_2}\otimes\big(L_{b_1}\circ R(b_2)\big)^*\Big)(a\otimes\tilde{P}'u) = L^*_{(b_2,b_1)}(a\otimes\tilde{P}'u). \tag{81.77}$$

Thus the last line of (81.76) provides

$$\begin{aligned}\big(L^*_{(b_2,b_1)}\circ\tilde{P}^{(1)}\big)\big((a\otimes u)\otimes(b\otimes v)\big) \\ = \mu^{12}\circ\big(L^*_{(b_2,b_1)}\otimes L^*_{(b_2,b_1)}\big)(a\otimes\tilde{P}'u)\otimes(b\otimes\tilde{P}''v) \\ = \mu^{12}\circ\big(L^*_{(b_2,b_1)}\otimes L^*_{(b_2,b_1)}\big)\circ\big((\mathrm{Id}\otimes\tilde{P}')\otimes(\mathrm{Id}\otimes\tilde{P}'')\big)(a\otimes u)\otimes(b\otimes v)\end{aligned} \tag{81.78}$$

Dropping the explicit reference to the functions a, b, u and v and using the formula (81.55), we have

$$\begin{aligned}\big(L^*_{(b_2,b_1)}\circ\tilde{P}^{(1)}\big) &= \mu^{21}\circ\Big[\,\mathrm{Id}\circ L^*_{b_2}\otimes\tilde{P}'\circ A\Big]\otimes\Big[\,\mathrm{Id}\circ L^*_{b_2}\otimes\tilde{P}'\circ A\Big] \\ &= \mu^{21}\circ\Big[(\mathrm{Id}\otimes\tilde{P}')\circ\big(L^*_{b_2}\otimes A^*\big)\Big]\otimes\Big[(\mathrm{Id}\otimes\tilde{P}'')\circ\big(L^*_{b_2}\otimes A^*\big)\Big] \\ &= \mu^{21}\circ(\mathrm{Id}\otimes\tilde{P}')\otimes(\mathrm{Id}\otimes\tilde{P}'')\circ(L^*_{b_2}\otimes A)^*\otimes(L^*_{b_2}\otimes A)^* \\ &= \tilde{P}^{(1)}\circ L^*_{(b_2,b_1)}\end{aligned} \tag{81.79}$$

where we used the expression (81.58). This concludes the proof of (2).

For (3), the hypothesis of invariance is that for every $a, b \in C^\infty(\mathbb{B}_2)$,

$$\tilde{L}(L^*_{b_2}a\otimes L^*_{b_2}b) = (L^*_{b_2}\circ\tilde{L})(a\otimes b). \tag{81.80}$$

From linearity and the density of $C^\infty(\mathbb{B}_2)\otimes C^\infty(\mathbb{B}_1)$ in $C^\infty(\mathbb{B}_2\times\mathbb{B}_1)$, it is sufficient to study

$$L^*_{(b_2,b_1)}\tilde{L}^{(2)}\Big((a\otimes u)\otimes(b\otimes v)\Big). \tag{81.81}$$

We have

$$\begin{aligned}
L^*_{(b_2,b_1)}\tilde{L}^{(2)}\Big((a\otimes u)\otimes(b\otimes v)\Big) &= L^*_{(b_2,b_1)}\circ\mu^{21}(\tilde{L}'a\otimes u)\otimes(\tilde{L}''b\otimes v)\\
&= L^*_{(b_2,b_1)}(\tilde{L}'a\cdot\tilde{L}''b)\otimes u\cdot v\\
&= L^*_{b_2}(\tilde{P}'a\cdot\tilde{L}''b)\otimes A^*(u\cdot v)\\
&= \mu^{21}\Big[L^*_{b_2}\tilde{P}'a\otimes A^*u\Big]\otimes\Big[L^*_{b_2}\tilde{L}''b\otimes A^*v\Big]\\
&= \mu^{21}\circ\big((L_{b_2}\otimes A)^*\otimes(L_{b_2}\otimes A)^*\big)\\
&\quad\circ\big((\tilde{L}'\otimes\mathrm{Id})\otimes(\tilde{L}''\otimes\mathrm{Id})\big)(a\otimes u)\otimes(b\otimes v)
\end{aligned} \tag{81.82}$$

The result now comes from the fact that

$$\mu^{21}\circ(L_{b_2}\otimes A)^*\otimes(L_{b_2}\otimes A)^* = (L_{b_2}\otimes A)^*\circ\mu^{21}. \tag{81.83}$$

This is lemma 81.13. □

Theorem 81.17 (formal extension lemma).
Let $\star^j=\sum_k h^kC^j_k$ be left invariant products on $C^\infty(\mathbb{B}_j)[[h]]$, and suppose that $\star^2$ is invariant under $R(g)$ for every $g\in\mathbb{B}_2$. Then we define the product $\star$ on $C^\infty(\mathbb{B})[[h]]$ by

$$\star=\star^2\otimes\star^2=\sum_k h^kC_k \tag{81.84}$$

with

$$C_k(a\otimes u,b\otimes v)=\sum_{k_1+k_2=k}C^2_{k_1}(a,b)C^1_{k_2}(u,v). \tag{81.85}$$

This product is left invariant under the action of $\mathbb{B}$.

Proof. What we have to study is

$$L^*_{(b_2,b_1)}\circ C_k\big((a\otimes u),(b\otimes v)\big)=\sum_{k_1+k_2=k}L^*_{(b_2,b_1)}C^2_{k_1}(a,b)C^1_{k_2}(u,v), \tag{81.86}$$

and by linearity, we can restrict ourself to study only one term in the sum:

$$\begin{aligned}
L^*_{(b_2,b_1)}C^2_k(a,b)C^1_l(u,v) &= L^*_{(b_2,b_1)}\circ\mu^{21}C^2_k(a,b)\otimes C^1_l(u,v)\\
&= \mu^{21}\Big[L^*_{(b_2,b_1)}C^2_k(a,b)\Big]\otimes\Big[L^*_{(b_2,b_1)}C^1_l(u,v)\Big].
\end{aligned} \tag{81.87}$$

Eq1607LCCLCLC

because here, μ^{21} has to be understood as in lemma 81.14, in particular it commutes with $L^*_{(b_2,b_1)}$. We are going to study separately the content of the two brackets. For the first one we have

$$L^*_{(b_2,b_1)}C^2_k(a,b)(b'_2,b'_1)=C^2_k(a,b)(b_2b'_2,b_1R_{b_2}b'_1), \tag{81.88}$$

but a and b do not depend on the $\mathbb{B}_1$ component of the variable while C^2 is $\mathbb{B}_2$-invariant, thus

$$C^2_k(a,b)(b_2b'_2,b_1R_{b_2}b'_1)=C^2_k(a,b)(b_2b'_2)=C^2_k(L^*_{b_2}a,L^*_{b_2}b). \tag{81.89}$$

Eq1607BrackUnMu

For the second bracket of (81.87) we have

$$\begin{aligned}
L^*_{(b_2,b_1)}C^1_l(u,v)(b'_2,b'_2) &= C^1_l(u,v)\big(b_2b'_2,b_1R(b_2)b'_1\big)\\
&= (L^*_{b_1}C^1_l)(u,v)\big(.,R(b_2)b'_1\big)\\
&= C^1_l\big(L^*_{b_1}u,L^*_{b_1}v\big)(.,R(b_2)b'_1)\\
&= C^1_l\big(R(b_2)^*\circ L^*_{b_1}u,A^*v\big)
\end{aligned} \tag{81.90}$$

1607BrackDeuxMu

Putting (81.89) and (81.90) into the brackets of (81.87) we find

$$\begin{aligned} L^*_{(b_2,b_1)} C^2_k(a,b) C^1_l(u,v) &= \mu^{21} C^2(L^*_{b_2} a, L^*_{b_2} b) \otimes C^1_l(A^* u, A^* v) \\ &= \mu^{21} \circ (C^2_k \otimes C^1_l)(L^*_{b_2} a \otimes L^*_{b_2} b) \otimes (A^* u \otimes A^* v) \end{aligned} \tag{81.91}$$

If we make the sum $\sum_{k_1+k_2=k}$ of the latter equation, we find

$$\sum_{k_1+k_2=k} C_k\Big((L^*_{b_2} a \otimes A^* u), (L^*_{b_2} b \otimes A^* v)\Big) = \sum_{k_1+k_2=k} C_k\Big(L^*_{(b_2,b_1)}(a \otimes u), L^*_{(b_2,b_1)}(b \otimes v)\Big). \tag{81.92}$$

□

This was the extension lemma from the point of view of the formal star product. From the point of view of the Drinfel'd twist, we have the following.

Proposition 81.18.
Let $F^{(j)}$ be formal twists based on $\mathcal{U}(\mathfrak{b}_j)[[h]]$ with

$$F^{(1)} \in \Big(\mathcal{U}(\mathfrak{b}_1)[[h]] \otimes \mathcal{U}(\mathfrak{b}_1)[[h]]\Big)^{\mathfrak{b}_2}. \tag{81.93}$$

Then we define $F \in \mathcal{U}(\mathfrak{b})[[h]] \otimes \mathcal{U}(\mathfrak{b})[[h]]$ by

$$F = F^{(2)} \otimes F^{(1)}. \tag{81.94}$$

This is a Drinfel'd twist based on $\mathcal{U}(\mathfrak{b})[[h]]$.

The article about deformation philosophy by Flato is [1034].

Chapter 82

Deformations: non-formal aspects

82.1 Introduction to Moyal star product

app:Moyal

82.1.1 General definition of a star product

The symbol $A[[\nu]]$ denotes the set of formal series of power of ν with coefficients in A, *i.e.*

$$A[[\nu]] = \{\sum_{k=0}^{\infty} a_k \nu^k \text{ with } a_i \in A\}.$$

For such formal series, the equality $\sum_{k=0}^{\infty} a_k \nu^k = \sum_{k=0}^{\infty} b_k \nu^k$ means $a_k = b_k$ for all k.

Let (M, P) be a Poisson manifold and $C^\infty(M)$, the algebra of differentiable functions on M. We know that it is associative and that it is a Lie algebra for the Poisson bracket.

Definition 82.1.
*A **formal star product** on M is a bilinear map $*\colon C^\infty(M) \times C^\infty(M) \to C^\infty(M)[[\nu]]$ which can be written as*

$$u * v = \sum_{k=0}^{\infty} \nu^k C_k(u, v) \tag{82.1}$$

where the $C_k\colon C^\infty(M) \to C^\infty(M)$ are bilinear maps such that

(1) the $\mathbb{R}[[\nu]]$-linear extension $*$ to $C^\infty(M)[[\nu]] \times C^\infty(M)[[\nu]]$ is associative:

$$u * (v * w) = (u * v) * w, \tag{82.2}$$

(2) the maps C_0 and C_1 are subject to

$$\begin{aligned} C_0(u, v) &= uv \qquad \text{usual product} \\ C_1(u, v) - C_1(v, u) &= \{u, v\} \quad \text{Poisson bracket.} \end{aligned} \tag{82.3}$$

(3) $1 * u = u * 1 = u$.

*If moreover the C_r for $r \geqslant 1$ are bidifferential operators on M, we say that $*$ is a **differential star product**.*

82.1.2 Definition of the Moyal star product

82.1.2.1 On $\mathbb{R}^n$ with constant P

We consider the Poisson manifold $(\mathbb{R}^{2m}, P)$ with $P = \frac{1}{2}\sum_{ij} P^{ij}\partial_i \wedge \partial_j$ where P^{ij} are constant. By definition, $\{u, v\} = P(u, v)$.

We want to define a star product $*$ with $\exp(P^{ij}\partial_i \wedge \partial_j)$, but $\exp P \circ \exp P$ makes no sense because $\exp P\colon C^\infty(M) \times C^\infty(M) \to C^\infty(M)$. So we write

eq:Moyal

$$\begin{aligned} (u *_M v)(x) &= \exp \frac{\nu}{2}(P^{ij}\partial_{y^i} \wedge \partial_{z^j})(u(y)v(z))|_{y=z=x} \\ &= u(x)v(x) + \frac{\nu}{2}\{u, v\}(x) + \dots \end{aligned} \tag{82.4}$$

The Moyal product reads

$$u *_M v = uv + \nu\{u,v\} + \sum_{k=2}^{\infty} \frac{\nu^k}{k!} \sum_{\substack{i_1 \dots i_k \\ j_1 \dots j_k}} \Omega^{i_1 j_1} \dots \Omega^{i_k j_k} \partial_{i_1} \dots \partial_{i_k} u \, \partial_{j_1} \dots \partial_{j_k} v. \tag{82.5}$$ EqDevStarMoyal

Let us consider the canonical form of P and a Darboux system of chart:

$$P = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} x_p \\ x_q \end{pmatrix}.$$

We define a **twisted Fourier transform** on $C^\infty(M)$ as follows. For $u \in C^\infty(M)$, $\hat{u} = F(u) \in C^\infty(M)$ is defined by:

$$\hat{u}(x) = F(u)(y) = \int e^{-iP(x,y)} u(x) dx. \tag{82.6}$$

This satisfies some properties which looks like the true Fourier transform; *inter alia*, this posses an associated "convolution" product:

$$(\hat{u} \times \hat{v})(y) = \int e^{-iP(\xi,c)} u(c-b) v(b) e^{iP(y,b)} d\xi \, dc \, db = \int e^{iP(y,b)} u(-b) v(b) \, db = F(\underline{u} v) \tag{82.7}$$ 1906r1

where $\underline{u}(x) = u(-x)$.

Lemma 82.2.
The twisted Fourier transform has the following property with respect to the derivative:

$$\widehat{\partial_i v} = i P^{il} y^l \hat{v}.$$

Proof. First notice that the integration by part usually gives a "boundary" term, but in a dense subset of $C^\infty(M)$, the following computation is true:

$$\begin{aligned} \int (\partial_i v)(x) e^{-iP(x,y)} dx &= -\int v(x) \partial_i \big(e^{-iP^{kl} x^k y^l}\big) dx \\ &= i P^{il} y^l \int v(x) e^{-P(x,y)} dx \\ &= i P^{il} y^l \hat{v}(y). \end{aligned} \tag{82.8}$$

□

We state the following lemma without proof.

Lemma 82.3.

$$\underline{\partial_i u} = -\partial_i \underline{u}$$

Theorem 82.4.
The Moyal star product is associative.

Proof. Equation (82.7) allows us to write the first term of $F(u * v)$; the two lemmas will help us to deal with the others terms.

With an obvious multi-index notation, we can write:

$$\begin{aligned} &F(P^{i_1 j_1} \dots P^{i_k j_k} \partial_{x^{i_1} \dots x^{i_k}} u \partial_{x^{j_1} \dots x^{j_k}} v) \\ &\quad = P^{IJ} (\widehat{\partial_I \underline{u}}) \times \widehat{\partial_J v} \\ &\quad = P^{IJ} (-1)^k \widehat{\partial_I \underline{u}} \times \widehat{\partial_J v} \\ &\quad = P^{IJ} (-1)^k (i)^k P^{IL} [\hat{\underline{u}} \cdot y^L] \times P^{JM} (i)^k [\hat{v} \cdot y^M] \\ &\quad = (-1)^k P^{ML} (\hat{\underline{u}} y^L) \times (\hat{v} y^M) \\ &\quad = (-1)^k \int (z-y)^L P^{ML} z^M \underline{\hat{u}}(z-y) \hat{v}(z) dz \\ &\quad = (-1)^k \int P(z-y,z)^k \underline{\hat{u}}(z-y) \hat{v}(z) dz. \end{aligned} \tag{82.9}$$

The last equality comes from the fact that $a^I P^{IJ} b^J = P(a,b)^k$. It is the way to compute the twisted Fourier transform of $u * v$. If we put $\nu = i\lambda$ The whole series gives:

$$\begin{aligned} F(u*v)(y) &= \int \exp -\frac{\nu}{2}(P(z-y,z))\underline{\hat{u}}(z-y)\hat{v}(z)dz \\ &= \int e^{i(\lambda/2)P(z,y)}\hat{u}(y-z)\hat{v}(z)dz \\ &= (\hat{u} \times_\lambda \hat{v})(y). \end{aligned} \tag{82.10}$$

The operation $\times_\lambda$ is called the **twisted convolution**. We can end the proof:

$$\begin{aligned} (\hat{u} \times_\lambda (\hat{v} \times_\lambda \hat{w}))(y) &= \int \exp i\frac{\lambda}{2}P(z,y)\hat{u}(u-z)\exp i\frac{\lambda}{2}P(z',z)\hat{v}(z-z')\hat{w}(z') \\ &= \exp i\frac{\lambda}{2}P(z,y-z')\hat{u}(y-z)\hat{v}(z-z')\hat{w}(z'). \end{aligned} \tag{82.11}$$

Computing $(\hat{u} \times_\lambda \hat{v}) \times_\lambda \hat{w}$, we find the same.

□

82.1.2.2 General Moyal product M

The true Moyal star product is nothing but to take a Darboux system of charts on M and to put the $\mathbb{R}^n$ Moyal product on each.

Remarque 82.5.
The Moyal product of polynomials leads to *finite* sums because the multi-derivation always ends with a zero.

82.2 Rieffel's deformation by action of $\mathbb{R}^d$

Let A be a Fréchet $*$-algebra and a strongly continuous action by $*$-homomorphisms α of $\mathbb{R}^d$ on A. We suppose that A has a system of seminorms $\|.\|_k$ defining the topology of A and such that α_v is an isometry for each $\|.\|_k$. The space $A^\infty \subset A$ of smooth vectors of the action is a Fréchet subalgebra of A with the seminorms

$$\|a\|_{k,\mu} = \sup_{|\beta|\leqslant\mu} \|\partial_\beta a\|_k \tag{82.12}$$

where $\partial_\beta a$ is defined by

$$\partial_\beta a = \frac{d}{dt}\Big[\alpha_{te_\beta}(a)\Big]_{t=0} \tag{82.13}$$

when $|\beta| = 1$. Here e_β is the vector of the canonical basis.

In this setting, Rieffel showed that if θ is an antisymmetric bilinear non degenerate on $\mathbb{R}^d$ and if $\hbar > 0$, then the formula

$$a \star_\hbar b = \frac{1}{(\pi\hbar)^{2n}} \int_{V\times V} \alpha_u(a)\alpha_v(b)e^{\frac{2i}{\hbar}\theta(u,v)}dudv \tag{82.14}$$

yields a well defined and continuous product on A^∞.

82.3 Weyl product

Let (V,Ω) be a symplectic vector space of dimension $2n$. For u, $v \in C_c^\infty(V)$, we define the **Weyl product** $u \star_\hbar^W v$ by the formula

$$(u \star_\hbar^W v)(x) = \hbar^{-2n} \int_{V\times V} e^{\frac{2i}{\hbar}S^0(x,y,z)}u(y)v(z)\,dy\,dz \tag{82.15}$$

eq_def_Weyl

where $S^0(x,y,z) = \Omega(x,y) + \Omega(y,z) + \Omega(z,x)$.

In order to prove stability of Schwartz space $\mathscr{S}(V)$ under the Weyl product, we follow [1035]. We begin by defining the following on $\mathscr{S}(V)$:

$$\begin{aligned} (\mu_j f)(u) &= u_j f(u) \qquad & \partial_i f &= \frac{\partial f}{\partial u_i} \\ (\tau_s f)(u) &= f(u-s) & (\epsilon_s f)(u) &= e^{i\Omega(s,u)} f(u). \end{aligned} \tag{82.16}$$

We also define

$$\hat{\partial}_j f = \begin{cases} \partial_{j+N} f & \text{if } 1 \leqslant j \leqslant N \\ -\partial_{j-N} f & \text{if } N < j \leqslant 2N. \end{cases} \tag{82.17a}$$

In other words, $\hat{\partial}_i$ is a derivative with respect to the momentum associated with the position i. If we denote by $\bar{\imath}$ the conhjugated variable of i and ϵ_i the corresponding sign, we can rewrite this definition under the form

$$\hat{\partial}_i = \epsilon_i \partial_{\bar{\imath}}. \tag{82.17b}$$

The change of variable $s = u - v$ and $t = w - u$ gives

$$(f \star_\hbar^W g)(u) = \iint f(u-s) g(t+u) e^{-i\Omega(s,t)}\, ds\, dt.$$

We define the **twisted convolution**

$$(f \diamond g)(u) = \int f(u-t) g(t) e^{-i\Omega(u,t)}\, dt \tag{82.18}$$

Proposition 82.6.
We have the formula

$$f \star_\hbar^W g = F^{-1}(Ff \diamond Fg)$$

where F denotes the Fourier transform.

Proof. The computation is as follows:

$$\begin{aligned} F^{-1}(Ff \diamond Fg)(u) &= \int e^{i\alpha u} (Ff \diamond Fg)(\alpha)\, d\alpha \\ &= \int e^{i\alpha u} e^{-i\Omega(\alpha,t)} e^{-\beta(\alpha-t)} e^{-\sigma t} f(\beta) g(\sigma)\, d\beta\, d\sigma\, dt\, d\alpha. \end{aligned}$$

If J is the symplectic matrix, $\Omega(\alpha,t) = \alpha \cdot Jt$; we use this fact to make appears Dirac delta's with the change of variable $t' = Jt$

$$\begin{aligned} &= \int \delta(u-\beta-t') e^{-iJ(\beta-\sigma)\,\cdot\, t'} f(\beta) g(\sigma)\, d\sigma\, d\beta\, dt' \\ &= \int e^{-iJ(\beta-\sigma)\,\cdot\,(u-\beta)} f(\beta) g(\sigma)\, d\sigma\, d\beta \\ &= (f \star_\hbar^W g)(u). \end{aligned} \tag{82.19}$$

□

Theorem 82.7.
The Schwartz space is stable for the Weyl product: if f, $g \in \mathscr{S}$, then $f \star_\hbar^W g \in \mathscr{S}$. Moreover the map $\mathscr{S} \times \mathscr{S} \to \mathscr{S}$, $(f,g) \mapsto f \star_\hbar^W g$ is a bilinear continuous map. The following equalities hold:

$$\partial_j(f \star_\hbar^W g) = \partial_i j \star_\hbar^W g + f \star_\hbar^W \partial_i g \tag{82.20a}$$

$$\begin{aligned} \mu_i(f \star_\hbar^W g) &= f \star_\hbar^W \mu_i g + i\hat{\partial}_i(f \star_\hbar^W g) \\ &= \mu_i(f \star_\hbar^W g) - if \star_\hbar^W \hat{\partial}_i g. \end{aligned} \tag{82.20b}$$

Proof. In order to prove equation (82.20a), we remark that the integral in the right hand side of (82.15) converges uniformly with respect to u. So one can permute the derivative and the integral:

$$\begin{aligned}\frac{\partial}{\partial u_i}\big((f \star_\hbar^W g)(u)\big) &= \int f(v)g(w)\frac{\partial}{\partial u_i}\big(e^{i[\Omega(u,v)+\Omega(v,w)+\omega(w,u)]}\big)\,dv\,dw \\ &= \int f(v)g(w)e^{i[\Omega(u,v)+\Omega(v,w)+\omega(w,u)]}(\epsilon_i v_{\bar\imath} - \epsilon_i w_{\bar\imath})\end{aligned}$$

where ϵ_i is a sign and $\bar\imath$ is the conjugated variable of i. An integration by part leads to (82.20a).

Definitions give

$$(f \star_\hbar^W \mu_i g)(u) = \int f(u-s)(t_j+u_j)g(t+u)e^{-i\Omega(s,t)}\,ds\,dt \tag{82.21a}$$

$$(\hat\partial_j f \star_\hbar^W g)(u) = \int \epsilon_j(\partial_{\bar\imath} f)(u-s)g(t+u)e^{-i\Omega(s,t)}\,ds\,dt \tag{82.21b}$$

we integrate the second integral by part using formula

$$\frac{\partial}{\partial s_{\bar\jmath}}\big(e^{-i\Omega(s,t)}\big) = i\epsilon_j t_j e^{-i\Omega(s,t)}.$$

Then, combining with the first, we find equation (82.20b).

An induction argument using uniform convergence of the integral and Leibniz shows that $f\star_\hbar^W g$ is smooth. The limit $\lim_{u\to\infty}(f\star_\hbar^W g)(u)$ is zero because [1]

$$\lim_{u\to\infty}\int f(u-u)e^{it}\,dt = 0.$$

An induction using formula (82.20b) shows that $f\star_\hbar^W g \in \mathscr{S}$.

In order to prove continuity, we consider the upper bound

$$\begin{aligned}\|f\star_\hbar^W g\|_\infty &= \sup_u |(f\star_\hbar^W g)(u)| \\ &\leqslant \int |f(u-s)|\,ds \int |g(t+u)|\,dt \\ &= \|f\|_1\|g\|_1\end{aligned} \tag{82.22}$$

The topologies given by seminorms [2]

$$q_{\alpha\gamma}(f) = \|\mu^\alpha\partial^\gamma f\|_1$$

is equivalent to the one defined by the seminorms

$$p_{\alpha\gamma}(f) = \|\mu^\alpha\partial^\gamma f\|_\infty.$$

So the equality

$$\mu^\alpha\partial^\gamma(f\star_\hbar^W g) = \sum_{\beta<\alpha}\sum_{\epsilon<\gamma}(-i)^{|\beta|}\binom{\alpha}{\beta}\binom{\gamma}{\epsilon}\mu^{\alpha-\beta}\partial^{\gamma-\epsilon}f\star_\hbar^W\hat\partial^\beta\partial^\epsilon g.$$

imply

$$p_{\alpha\gamma}(f\star_\hbar^W g) \leqslant \sum_\beta\sum_\epsilon\binom{\alpha}{\beta}\binom{\gamma}{\epsilon}q_{\alpha-\beta,\gamma-\epsilon}(f)q_{0,\eta+\epsilon}(g).$$

because $\|f\star_\hbar^W g\|_\infty \leqslant \|f\|_1\|g\|_1$. This proves that $(f,g)\mapsto f\star_\hbar^W g$ is continuous.

□

1. J'espère que c'est vrai !
2. Il faut vérifier ce point

82.4 The twisting map

We consider the elementary normal symplectic symmetric space $\mathbb{S}$ defined in section 66.9 and its description of proposition 66.93. We follow the paper [892].

Let $\tilde{\mathbb{S}} = \{(a, v, \xi\} = \mathbb{R} \times \mathbb{R}^{2n} \times \mathbb{R}$. The **twisting map** is the smooth family (with respect to θ) of diffeomorphisms

$$\begin{aligned} \phi_\theta &: \tilde{\mathbb{S}} \to \tilde{\mathbb{S}} \\ (a, v, \xi) &\mapsto \left(a, \cosh\left(\frac{\theta}{4}\xi\right)^{-1} v, \frac{2}{\theta}\sinh\left(\frac{\theta}{2}\xi\right)\right) \end{aligned} \tag{82.23}$$

with $\theta \in \mathbb{R}$. We denote by $\varphi_\theta \in \mathrm{Diff}(\mathbb{R})$ the partial map

$$\varphi_\theta(\xi) = \frac{2}{\theta}\sinh(\frac{\theta}{2}\xi). \tag{82.24}$$

We denote by $\mathcal{F}$ the partial Fourier transform

$$(\mathcal{F}u)(a, v, \xi) = \hat{u}(a, v, \xi) = \int e^{-i\xi t} u(a, v, t) dt. \tag{82.25}$$

The space Schwartz functions in the variables (a, v, t) is denoted by $\mathscr{S}$ and $\tilde{\mathscr{S}}$ in the variables (a, v, ξ). By general theory, the Fourier transform is an isomorphism

$$\mathcal{F} : \mathscr{S} \to \tilde{\mathscr{S}} \tag{82.26}$$

Let us now consider the space of function

$$\mathcal{O}_C = \{f \in C^\infty(\mathbb{R}^m) \text{ tel que } \exists r > 0 \text{ tel que } \forall \alpha \in \mathbb{N}^m, |\partial^\alpha f(x)| < C_\alpha(1 + |x|)^r\}. \tag{82.27}$$

Definition 82.8.
The space Θ is the space of functions $\theta : \mathbb{R} \to C^\infty(\mathbb{R})$ denoted by $\tau : \theta \to \tau_\theta$ such that

(1) For every θ, the functions $x \mapsto e^{\pm\tau_\theta(x)}$ belong to $\mathcal{O}_C(\mathbb{R})$.

(2) We have $(\varphi_\theta^* \tau_\theta)|_{\theta=0} \equiv 0$. More explicitly, for every ξ we have

$$\lim_{\theta \to 0} (\varphi_\theta \circ \tau_\theta)(\xi) = 0 \tag{82.28}$$

82.5 Wigner function

The Weyl quantization formula is

$$\big(\mathrm{Op}(a)u\big)(x) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} a\left(\frac{x+y}{2}, \eta\right) e^{2\pi i\langle x-y, \eta\rangle} u(y)\, dy\, d\eta \tag{82.29}$$

when $a \in \mathscr{S}'(\mathbb{R}^n \times \mathbb{R}^n)$. It defines a continuous map $\mathrm{Op}(a) : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$. Let φ, the usual gaussian function on $\mathbb{R}^n$:

$$\varphi(t) = 2^{n/4} e^{-\pi|t|^2},$$

and for all $X = (x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\varphi_X(t) = \varphi(t - x) e^{2\pi i\langle t - \frac{x}{2}, \xi\rangle}.$$

When $u, v \in L^2(\mathbb{R}^n)$, the scalar product is defined by

$$(u, v) = \int u(t)\overline{v}(t)\, dt \tag{82.30}$$

and Parseval equality leads to

$$(u, v) = \int_{\mathbb{R}^n} (u, \varphi_X)(\varphi_X, v)\, dX \tag{82.31}$$

eq_intuvpXb

because the integral of $e^{2\pi i\langle z-y, \xi\rangle}$ over ξ gives rise to a Dirac delta $\delta(z - y)$. Equation (82.31) still holds when $u \in \mathscr{S}(\mathbb{R}^n)$ and $v \in \mathscr{S}'(\mathbb{R}^n)$.

Without proof:

¡¡ Avertissement/question au lecteur !! 82.9
Il manque le reste de l'égalité. Énoncé imcomplet donc. À compléter.

Proposition 82.10.
If $u \in \mathscr{S}'(\mathbb{R}^n)$ and for all $k \in \mathbb{N}$, the equality

$$\int_{\mathbb{R}^{2n}} (1+|X|)^2 |(u, \varphi_X)|\, dX$$

holds, then $u \in \mathscr{S}(\mathbb{R}^n)$.

The **Wigner function** $W_{XX'}$ is defined by

$$\big(\operatorname{Op}(a)\varphi_X, \varphi_{X'}\big) = \int a(Y) W_{XX'}(Y)\, dY. \tag{82.32}$$

It can be developed under the form [3]

$$W_{XX'}(Y) = 2^n e^{-i\pi\Omega(X,X')} e^{2\pi i\Omega(Y,X'-X)} e^{-2\pi\left|Y-\frac{X+X'}{2}\right|^2} \tag{82.33}$$

where Ω is the symplectic form on $\mathbb{R}^n \times \mathbb{R}^n$ defined by

$$\Omega\big((x,\xi),(x',\xi')\big) = -\langle x, \xi'\rangle + \langle x', \xi\rangle.$$

A **weight function** on $\mathbb{R}^{2n}$ is a positive function m for which there exists c_1 and n_1 such that

$$m(X) \leqslant c_1 m(X')\big[1+|X-X'|^2\big]^{n_1}$$

for all $X, X' \in \mathbb{R}^{2n}$. A function $a \in C^\infty(\mathbb{R}^{2n})$ is a **symbol of weight** m if for all multi-index $\alpha \in \mathbb{N}^{2n}$, the function $m^{-1}D^\alpha a$ is bounded, i.e. if

$$\sup_{(x,\xi)} \Big|\frac{D^\alpha a(x,\xi)}{m(x,\xi)}\Big| \leqslant \infty.$$

Theorem 82.11.
Let $a \in \mathscr{S}'(\mathbb{R}^{2n})$. This is a symbol of weight m if and only if for all $k \in \mathbb{N}$, there exists a $c > 0$ such that

$$\big|(\operatorname{Op}(a)\varphi_X, \varphi_{X'})\big| \leqslant c\big(1+|X-X'|^2\big)^{-k} m\Big(\frac{X+X'}{2}\Big)$$

for all $X, X' \in \mathbb{R}^{2n}$.

82.6 Deformation by action of group

SecDefAction

The procedure of deformation by group action is described in [1036]. Let G be a Lie group. We suppose to know a subset A^G of $\operatorname{Fun}(G, \mathbb{C})$ such that

(1) A^G is invariant under the left regular representation of G on itself,
(2) A^G is provided with a G-invariant product $\star^G$ such that $(A^G, \star^G)$ is an associative algebra. The G-invariance means that $\forall a, b \in A^G$,

$$(L_g^* a) \star^G (L_g^* b) = L_g^*(a \star^G b).$$

Notice that we do not impose any regularity condition on this product. The reason is that the deformation by group action is a formal procedure which allows to guess a product on a manifold. The "true" work to prove convergences and invariances has to be done on the level of the deformed manifold.

3. Que je ne suis pas très bien parvenu à démonter.

Now, let X be a manifold endowed with a right action $\tau\colon G\times X\to X$ of G. For $u\in \mathrm{Fun}(X)$, $x\in X$ and $g\in G$, we consider $\alpha^x(u)\in\mathrm{Fun}(G)$ and $\alpha_g(u)\in\mathrm{Fun}(X)$ defined by

$$\alpha^x(u)(g)=\alpha_g(u)(x)=u(\tau_{g^{-1}}(x)), \tag{82.34}$$

EqDefalphaxu

and the following functional space on X:

$$A^X=\{u\in\mathrm{Fun}(X)|\alpha^x(u)\in A^G\,\forall x\in X\}.$$

For example, the A^X corresponding to $A^G=\mathrm{Fun}(G)$ is the whole $\mathrm{Fun}(X)$. For u, $v\in A^X$, we define $\star^X\colon A^X\times A^X\to\mathrm{Fun}(X)$

$$(u\star^X v)(x)=\big(\alpha^x(u)\star^G\alpha^x(v)\big)(e) \tag{82.35}$$

eq:def_stG

where e is the identity of G.

Theorem 82.12.
The product (82.35) obtained by action of the group G on the manifold X fulfils the following properties:

itemthostG

(1) The operation α^x intertwines the products $\star^X$ and $\star^G$:

$$\alpha^x(u\star^X v)=(\alpha^x u)\star^G(\alpha^x v).$$

(2) A^X is stable under $\star^X$,

(3) $(A^X,\star^X)$ is an associative algebra.

Proof. First remark that $\alpha^{\tau_{g^{-1}}(x)}u=L_g^*\alpha^x u$ because

$$(\alpha^{\tau_{g^{-1}}(x)}u)(h)=u\big(\tau_{(gh)^{-1}}(x)\big)=(\alpha^x u)(gh)=\big(L_g^*\alpha^x u\big)(h),$$

It follows that

$$\begin{aligned}\alpha^x(u\star^X v)(g)&=(u\star^X v)(\tau_{g^{-1}}(x))=(\alpha^{\tau_{g^{-1}}(x)}u\star^G\alpha^{\tau_{g^{-1}}(x)}v)(e)\\&=\big[L_g^*(\alpha^x u\star^G\alpha^x v)\big](e)=(\alpha^x u\star^G\alpha^x v)(g).\end{aligned} \tag{82.36}$$

The first point is proved.

Using the first point, we see that $\alpha^{\tau_{g^{-1}}(x)}u$ belongs to A^G because $\alpha^x u\in G^G$ and A^G is stable under L_g. So we have the second point. For the third one,

$$\begin{aligned}[(u\star^X v)\star^X w](x)&=\big(\alpha^x(u\star^X v)\star^G\alpha^x(w)\big)(e)\\&=\big((\alpha^x u\star^G\alpha^x v)\star^G\alpha^x(w)\big)(e).\end{aligned}$$

The conclusion follows from associativity of $\star^G$. □

Let us summarize what was done up to now. When G acts on X, and when we have a "good" product on $A^G\subset\mathrm{Fun}(G)$, we are able to build an associative product on $A^X\subset\mathrm{Fun}(X)$. The space A^X is defined by A and the action. So a deformation of a group gives rise to a deformation of any manifold on which the group acts. This is why we call it an "universal" deformation. That universal construction is the motivation to deform groups.

Lemma 82.13.
A function u belongs to A^X if and only if there exists one y such that $\alpha^y(u)\in A^G$ in each g-orbit in X.

LemUnPtParOrbite

Proof. The necessary condition is direct because, when $u\in A^X$, the function $\alpha^x(u)$ belongs to A^G for every x. For the sufficient condition, suppose $\alpha^y(u)\in A^G$, then $\alpha^{g\cdot y}(u)=L_{g^{-1}}^*(\alpha^y u)\in A^G$ for all g because A^G is left invariant. If it holds for a y in each G-orbit, then $\alpha^x u\in A^G$ for all $x\in X$. □

The content of this lemma is that if one wants to check if a given function u belongs to A^X, one only has to check is $\alpha^y u \in A^G$ for one y in each G-orbit.

The functions $\alpha^x(u)$ are not "gentle" functions, even when u is. Let us give two examples of pathology that can occur in $\alpha^x(u)$ without to be present in u. Firstly, if the action is the identity, the support of $\alpha^x(u)$ is the whole G which can be non compact. So, even when u is compactly supported, there are no guarantee with respect to the support of $\alpha^x(u)$.

Secondly, the function $\alpha^x(u)$ is of course bounded; but the derivatives are not specially such. Indeed, in order to fix ideas, suppose that the group G is a two parameter group and that the manifold X is a two dimensional manifold. In this case, one can write

$$f(a,l) = \alpha^x(u)(a,l) = u\big(z_1(a,l), z_2(a,l)\big) \tag{82.37}$$

EqDefziDefA

where x is a parameter in the functions z_i. Depending on the action, the function z can be very odd. In particular, the derivatives

$$(\partial_a f)(a,l) = (\partial_1 u)(z_1,z_2)(\partial_a z_1)(a,l) + (\partial_2 u)(z_1,z_2)(\partial_a z_2)(a,l)$$

in which $\partial_a z_i$ can be divergent. Even worse, the degree of the divergence can increase with the degree of the derivation. Two examples of such a hill behaviour are given in section 85.1.

82.7 A first analysis question

The big work will be done in section 83.2,

82.7.1 A simple toy model

We consider the multiplicative action of $\mathbb{R}_0$ (i.e. $\mathbb{R}$ minus zero) on $\mathbb{R}$ and the induced one on $C(\mathbb{R})$:

$$(r \cdot f)(x) = f(rx). \tag{82.38}$$

We are searching for the functions $f \in C(\mathbb{R})$ such that the map $\phi_f \colon \mathbb{R} \to C(\mathbb{R})$, $\phi_f(r)x = f(rx)$ is continuous. For, we consider the topology on $C(\mathbb{R})$ given by a choice of a fundamental sequence of compact sets (K_m) and the open sets

$$\mathcal{O} = \{g \in C(\mathbb{R}) \text{ tel que } \sup_{x\in K_m} |f_0(x) - g(x)| < \varepsilon\}.$$

for any choice of f_0 and ε. We want a f such that $\phi_f^{-1}(\mathcal{O})$ is open in $\mathbb{R}_0$. We have

$$\phi_f^{-1} = \{z \in \mathbb{R}_0 \text{ tel que } \sup_{x\in K_m} |f_0(x) - f(zx)| < \varepsilon\}. \tag{82.39}$$

Let us prove that $f(x) = e^x$ makes $\phi_f^{-1}(\mathcal{O}) = \phi_{\exp}^{-1}(\mathcal{O})$ open. Let $z_0 \in \phi_{\exp}^{-1}$, i.e.

$$\sum_{x\in K_m} |f_0(x) - e^{z_0 x}| < \varepsilon.$$

It just a little game to see that, by choosing $|\delta|$ small enough, we have $z_0 + \delta \in \phi_{\exp}^{-1}$.

The important point to remark is that the chosen topology will accept [4] all functions which are uniformly continuous on all compact. In particular, it will admit $f(x) = e^{x^2}$. This is *not* the same topology as the one used by Rieffel.

4. Est-ce le bon mot ?

82.7.2 The case of interest

Let us now address a problem. Let $R = AN \subset SL(2,\mathbb{R})$ and the coadjoint action of R on the dual of its Lie algebra $R \times \mathcal{R}^* \to \mathcal{R}^*$:

$$(r \cdot \xi)X = \big(\mathrm{Ad}^*(r)\xi\big)(X) = \xi\big(\mathrm{Ad}(r^{-1})X\big). \tag{82.40}$$

for $X \in \mathcal{R}$, $\xi \in \mathcal{R}^*$ and $r \in R$. We consider the following decomposition of R:

$$(a,l) = \begin{pmatrix} e^a & e^a l \\ 0 & e^{-a} \end{pmatrix} \tag{82.41}$$

eq:defeasal

which induces the group law

$$(a,l)(a',l') = (a+a', l' + e^{-2a'}l) \tag{82.42}$$

and the inverse

$$(a,l)^{-1} = (-a, -e^{2a}l). \tag{82.43}$$

Among other relations, we point out $(a,0) = e^{aH}$, $(0,l) = e^{lE}$, $\mathbb{1} = (0,0)$ and $(a,l) = (a,0)(0,l)$.

Let us compute explicitly the action of $r = (a,l)$ on $\xi = y_H H^* + y_E E^*$. For this, we consider $X = x_H H + x_E E$ and we compute

$$\begin{aligned}(r \cdot \xi)X &= [(a,l) \cdot (y_H H^* + y_E E^*)](x_H H + x_E E) \\ &= (y_H H^* + y_E E^*)(\mathrm{Ad}(a,l)^{-1}(x_H H + x_E E)) \\ &= (y_H H^* + y_E E^*)(x_H H + e^{-2a}(x_E + 2e^{2a} l x_H) Z E) \\ &= (y_H + 2y_E l)x_H + y_E e^{-2a} x_E.\end{aligned} \tag{82.44}$$

So

$$r \cdot \xi = (y_H + 2y_E l)H^* + y_E e^{-2a} E^*. \tag{82.45}$$

From a matrix point of view, the matrix of (a,l) on the basis $\begin{pmatrix} H^* \\ E^* \end{pmatrix}$, we have

$$r = (a,l) \rightsquigarrow \begin{pmatrix} 1 & 2l \\ 0 & e^{-2a} \end{pmatrix}.$$

In particular, the determinant of this matrix is non zero (so the action of R on $\mathcal{R}^*$ is always bijective), the space spanned by H^* is fixed [5], and the sign of the E^* component doesn't changes under the action. Then the action has three orbits in $\mathcal{R}^*$: two half plane of dimension 2 and one linear subspace of dimension 1.

82.7.3 Smooth vectors of the action

For $u \in \mathrm{Fun}(\mathcal{R}^*)$, the actions we are looking at is

$$\alpha^\xi(u)(r) = \alpha_r(u)\xi = u(\alpha_{r^{-1}}(\xi))$$

and we consider the map

$$f_u \colon R \to \mathrm{Fun}\,\mathcal{R}^* \quad r \mapsto u(\alpha_{r^{-1}}(\xi)). \tag{82.46}$$

The question is: for which $u \in \mathrm{Fun}(\mathcal{R})$, this map is smooth? As first question, we have to precise the sense of a smooth map from R into $\mathrm{Fun}(\mathcal{R}^*)$. The derivative of $f_u \colon R \to \mathrm{Fun}(\mathcal{R}^*)$ is defined by means of the map $\tilde{f} = f \circ \varphi \colon \mathbb{R}^2 \to \mathrm{Fun}\,\mathcal{R}^*$ where $\varphi \colon \mathbb{R}^2 \to R$ is the chart (82.41). We naturally forget the tilde. The quantity which have to be defined is

$$f_u'(x) = \lim_{\varepsilon \to 0} \frac{f_u(x+\varepsilon) - f_u(x)}{\varepsilon}.$$

It is natural to apply it to a $\xi \in \mathcal{R}^*$ and define

$$f_u'(\xi) := \lim_{\varepsilon \to 0} \frac{f_u(X+\varepsilon)\xi - f_u(x)\xi}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{u\big(\alpha_{(x+\varepsilon)^{-1}}(\xi)\big) - u\big(\alpha_{x^{-1}}(\xi)\big)}{\varepsilon}. \tag{82.47}$$

5. *each* point of this space is fixed.

82.7.4 Oscillatory definition on $\mathrm{SL}(2,R)$

subsec:Oscdefsl

The integral we have to study now is

$$\int_{R\times R} F(x_1,x_2)e^{2iS(x_0,x_1,x_2)}dx_1dx_2.$$

Where $x_i=(a_i,l_i)$. As usual, we begin to suppose that F has compact support so that integral makes sense and integration by part does not produce boundary term.

We consider the phase function as

$$S(x_0,x_1,x_2)=l_2\sinh(a_0-a_1)+l_0\sinh(a_1-a_2)+l_1\sinh(a_2-a_0). \tag{82.48}$$

We have

$$\partial_{a_1}S=-l_2\cosh(a_0-a_1)+l_0\cosh(a_1-a_2) \tag{82.49a}$$
$$\partial^2_{a_1}S=l_2\sinh(a_0-a_1)+l_0\sinh(a_1-a_2) \tag{82.49b}$$
$$\partial_{a_2}S=-l_0\cosh(a_1-a_2)+l_1\cosh(a_2-a_0) \tag{82.49c}$$
$$\partial^2_{a_2}S=l_0\sinh(a_1-a_2)+l_1\sinh(a_2-a_0) \tag{82.49d}$$
$$\partial_{l_1}S=\sinh(a_2-a_0) \tag{82.49e}$$
$$\partial_{l_2}S=\sinh(a_0-a_1) \tag{82.49f}$$

It is easy to compute that, for each t in $\{a_1,a_2,l_1,l_2\}$

$$\int(\partial_t^2F)e^{iS}=\int F[i\partial_t^2S-(\partial_tS)^2]e^{iS}$$

that we write under the form

$$\int(\Delta F)e^{iS}=\int FLe^{iS} \tag{82.50}$$

with

$$L=\sum_t[i\partial_t^2S-(\partial_tS)^2].$$

Let us write it for $F'=F/(1-L)$:

$$\int(1-\Delta)\left(\frac{F}{1-L}\right)e^{iS}=\int Fe^{iS}.$$

Making an induction, one sees that

$$\int Fe^{iS}=\int[(1-\Delta)\circ M_K]^kFe^{iS} \tag{82.51}$$

where $M_KF=F/(1-L)$ and $K=1/(1-L)$.

Let us suppose that F is a polynomial with respect to $\sinh(a_i)$ and $\cosh(a_i)$. We write $P_{a,b}$, a polynomial of degree a in $\cosh(a_1)$, $\sinh(a_1)$ and of degree b in $\cosh(a_2)$, $\sinh(a_2)$. For example,

$$P_{a,b}=\cosh(a_1)^a\sinh(a_2)^{b-1}\cosh(a_2).$$

We will use any abuse of notation as

$$\partial_{a_1}(P_{a,b})=P_{a,b}$$

or $P_{1,7}+P_{4,2}=P_{4,7}$ and so on. We have $L=iP_{1,1}+P_{2,2}$ with real coefficients. One can see that

$$\begin{aligned}\Delta\left(K\frac{P_{a,b}}{P_{c,d}}\right)&=K^3\left(\frac{P_{a+4,b+4}}{P_{c,d}}\right)\\&+K^2\left(\frac{P_{a+2,b+2}}{P_{c,d}}+\frac{P_{a+c+2,b+d+2}}{P_{2c,2d}}+\frac{P_{a+2c+2,b+2d+2}}{P_{3c,3d}}\right)\\&+K\left(\frac{P_{a+2c,b+2d}}{P_{3c,3d}}+\frac{P_{a+5c,b+5d}}{P_{6c,6d}}\right)\end{aligned} \tag{82.52}$$

which is for all terms two degrees lower than $\frac{P_{a,b}}{P_{c,d}}$.

So by virtue of operator $[(1-\Delta)\circ M_K]^k$, we can get

$$\int Fe^{iS} = \int \frac{P_{a,b}}{P_{c,d}} e^{iS} \tag{82.53}$$

with $c-a$ and $d-b$ as large as we want.

82.7.4.1 First example

Let us perform the computation to prove the integrability of $F=\cosh(a_1-a_2)$. First,

$$[(1-\Delta)\circ M_K]\cosh(a_1-a_2) = \frac{\cosh(a_1-a_2)}{1-L} - \Delta\left(\frac{\cosh(a_1-a_2)}{1-L}\right). \tag{82.54}$$

eq:umDelfracch

As far as the first term is concerned, we write

$$\left\|\frac{\cosh(a_1-a_2)}{1-L}\right\| \leqslant \frac{\|\cosh(a_1-a_2)\|}{\|1-L\|}.$$

We find an upper bound of it by forgetting the imaginary part of L and the three first terms of

$$\mathrm{Re}(1-L) = 1 + [\partial_{a_1} S]^2 + [\partial_{a_2} S]^2 + \sinh^2(a_2-a_0) + \sinh^2(a_0-a_1).$$

Then

$$\left\|\frac{\cosh(a_1-a_2)}{1-L}\right\| \leqslant \frac{\cosh(a_1-a_2)}{\sinh^2(a_2-a_0)+\sinh^2(a_0-a_1)}.$$

If one looks at it on the line $a_1=a_2=x$, then one finds $\cosh(x)/2\sinh(2x)$ whose limit is $1/2$. So we have to take a second iteration of operator $[(1-\Delta)\circ M_K]$. So let us continue computation (82.54).

Since $\nabla F=0$, we find $\Delta F = 2\cosh(a_1-a_2)$ and

$$\begin{aligned}\Delta\left(\frac{\cosh(a_1-a_2)}{1-L}\right) &= \Delta K\cosh(a_1-a_2) + 2K\cosh(a_1-a_2)\\ &= \cosh(a_1-a_2)(\Delta K + 2K).\end{aligned} \tag{82.55}$$

Simple computation shows that

$$\begin{aligned}[(1-\Delta)\circ M_K]^2\cosh(a_1-a_2) &= \cosh(a_1-a_2)(3K^2 - K\Delta K)\\ &\quad - \Delta(\cosh(a_1-a_2)(3K^2-K\Delta K)).\end{aligned} \tag{82.56}$$

Lemma 82.14.
If the function F becomes larger than any exponential, the integral $\int Fe^{iS}$ makes no sense.

Proof. The definition of $\int Fe^{iS}$ is given by iteration of operator $[(1-\Delta)\circ M_K]$, but

$$\begin{aligned}\int Fe^{iS} &= \int [(1-\Delta)\circ M_K] Fe^{iS}\\ &= \int M_K Fe^{iS} - \int \Delta\left(\frac{F}{1_L}\right)e^{iS}\end{aligned} \tag{82.57}$$

where the function KF in the first integral has same fundamental property as F. □

82.7.4.2 Functional spaces

From theory developed in 82.6, we have to study the set

$$A^{\mathcal{R}^*} = \{u \in \mathrm{Fun}(\mathcal{R}^*) \text{ tel que } \alpha^\xi \in A^R \, \forall \xi \in \mathcal{R}^*\}$$

where, by definition, $(\alpha^\xi u)(r) = u(r^{-1} \cdot \xi)$. From our previous work in 82.7.4, we define A^R as the space of functions $f : R \to \mathbb{C}$ for which there exists a m, $n \in \mathbb{N}$ such that

$$(\alpha^\xi u)(a,l)^{-1} = u(y_H + 2y_E l, y_E e^{-2a}). \tag{82.58}$$

eq:alpxial

In order the map $(a,l)^{-1} \to u(y_H + 2y_E l, y_E e^{-2a})$ to belongs to A^R, we need

$$|u(x_H, x_E)| < |x_E|^n |x_H|^m. \tag{82.59}$$

The point is the fact that a comes in (82.58) in an exponential.

82.8 Method improvement

82.8.1 Change of point of view for the integral

The product of two functions u, $v \in \mathrm{Fun}(R)$ reads

$$(u \star^R v)(x) = \int_{R\times R} K(x,y,z)u(y)v(z)\,dy\,dz \tag{82.60}$$

where

— K is left invariant,

— the measure on R is left invariant,

— the action $R \times \mathbb{A} \to \mathbb{A}$ is the regular right action: $\alpha_g(u)x = u(xg)$.

Using these properties, we can transform the integral:

$$\begin{aligned}
(u \star^R v)(x) &= \int K(x,y,z)u(y)v(z)\,dy\,dz \\
&= \int K(e,x^{-1}y,x^{-1}z)u(y)v(z)\,dy\,dz && \text{invariance of } K \\
&= \int K(e,\eta,\zeta)u(x\eta)v(x\zeta)\,d\eta\,d\zeta && \text{measure invariance} \\
&= \int K(e,\eta,\zeta)(R^*_\eta u)(x)(R^*_\zeta v)(x)\,d\eta\,d\zeta
\end{aligned}$$

Then one can write

$$u \star^R v = \int_{R\times R} K(e,\eta,\zeta)R^*_\eta u R^*_\zeta v\,d\eta\,d\zeta \tag{82.61}$$

eq:intadefinir

as a $\mathbb{A}$-valued integral of a function $R \times R \to \mathbb{A}$. The problem now is to find a subset $\mathbb{A} \subset \mathrm{Fun}(R)$ in which it is possible to give a sense to integral (82.61).

82.8.2 Adapted vector field and by part integration

In 82.7.4, we used derivation $\partial_{a_1}, \dots$ to define the oscillatory integral. It was meaningful in the case of the additive action of $\mathbb{R}^n$ in 62.2.1 and more generally in the case of Rieffel [952]. However, these vector fields on R are not natural vector fields in this sense that it "fits" no structure. In particular, it does not take account of the action. More natural vector fields are the invariant vector fields for the action:

$$(\tilde{X} \cdot u)(x) = \frac{d}{dt}\Big[\alpha_{e^{tX}} u\Big]_{t=0}(x) = \frac{d}{dt}\Big[u(xe^{tX})\Big]_{t=0}. \tag{82.62}$$

The middle expression shows that $\tilde{X}$ is a derivation of the action. Let us compute the corresponding by part integration formula.

$$\begin{aligned}\int (\tilde{X} \cdot u)v &= \frac{d}{dt}\int (\alpha_{e^{tX}} u)(x)v(x)\,dx\Big|_{t=0} \\ &= \frac{d}{dt}\int u(xe^{tX})v(x)\,dx\Big|_{t=0}.\end{aligned} \tag{82.63}$$

Now the non right invariance of the measure is painful because we want to effect a change of variable $y = xe^{tX}$. We introduce the **modular function** Δ by

$$d(xg) = \Delta(g, x)\,dx. \tag{82.64}$$

So $x = ye^{-tX}$ implies $dx = \Delta(e^{-tX}, y)\,dy$ and the integral becomes

$$\begin{aligned}\int (\tilde{X} \cdot u)v &= \frac{d}{dt}\int u(y)v(ye^{-tX})\Delta(e^{-tX}, y)\,dy\Big|_{t=0} \\ &= \int u(y)(-\tilde{X})v(y)\,dy + \int u(y)v(y)\frac{d}{dt}\Big[\Delta(e^{-tX}, y)\,dy\Big]_{t=0}\end{aligned} \tag{82.65}$$

eq:euler

and we finally write

$$\int (\tilde{X} \cdot u)v = -\int u\tilde{X} \cdot v + \int uv\mu_X \tag{82.66}$$

where μ_X is defined as the last derivative. Note that in the whole computation, we would had written $\int \frac{d}{dt}$ instead of $\int \frac{d}{dt}$. Then there are in fact no inversion problem.

82.8.3 New way to express the problem

We are now able to give a better statement of our analysis problem. Let $\mathbb{A}$ be a topological algebra (presently, a subspace of $\mathrm{Fun}(R)$) and a map $F \in C^\infty(R \times R, \mathbb{A})$, presently

$$F(x, y) = \cosh(x_1 - y_1)\alpha_x u \alpha_y v. \tag{82.67}$$

eq:Fchxuv

The purpose is to give a sense to

$$\int e^{i\sigma(x,y)} F(x, y)\,dx\,dy \tag{82.68}$$

as element in $\mathbb{A}$.

Remarque 82.15.
In order the function (82.67) to be smooth, one needs in particular u and v to be smooth vector of the action.

We denote by $\tilde{X}^{(1)}F$ the action of $\tilde{X}$ on the first variable of F. Then

$$\begin{aligned}\int \tilde{X}^{(1)}Fe^{i\sigma} &= -\int F\tilde{X} \cdot (e^{i\sigma}) + \int F\mu_X^{(1)}e^{i\sigma} \\ &= -i\int F(\tilde{X}^{(1)} \cdot \sigma)e^{i\sigma} + \int F\mu_X^{(1)}e^{i\sigma} \\ &= \int [\mu_X^{(1)} - i\tilde{X}^{(1)} \cdot \sigma]Fe^{i\sigma}\end{aligned} \tag{82.69}$$

where

$$(\tilde{X}^{(1)} \cdot \sigma)(x, y) = \frac{d}{dt}\Big[\sigma(xe^{tX}, y)\Big]_{t=0}. \tag{82.70}$$

When u, $v \in \mathbb{A}$, we look at

$$\begin{aligned}\Phi\colon R \times R &\to \mathbb{A} \\ (x, y) &\mapsto K(e, x, y)(\alpha_x u)(\alpha_y v)e^{i\sigma(x,y)}\end{aligned} \tag{82.71}$$

and we ask ourself if the integral

$$\int_{R\times R} \Phi$$

exists or not.

82.8.4 Topology choice

In order to define a Fréchet space, we want to put seminorms on $\mathbb{A}$ in such a way that the action is isometric for each seminorm. For example, if K is a compact set in R,

$$p_K(u) = \sup_{x\in K} |u(x)|$$

doesn't works because

$$p_K(\alpha_g u) = \sup_{x\in K} |u(xg)| = -p_{gK}(u) \neq p_K(u).$$

We now suppose that $\mathbb{A}$ is a Banach algebra for the supremum norm and we are searching for a topological structure on the subspace $\mathbb{A}^\infty$. So we can try the following seminorm:

$$p'_X(u) = \|X \cdot u\|_{\mathbb{A}}, \tag{82.72}$$

but it does not work because

$$\begin{aligned}
(X \cdot \alpha_g u)(x) &= \frac{d}{dt}\Big[(\alpha_g u)(xe^{tX})\Big]_{t=0} \\
&= \frac{d}{dt}\Big[u(xe^{tX}g)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[(\alpha_{e^{tX}g} u)(x)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[(\alpha_{g\,\mathbf{Ad}(g^{-1})e^{tX}} u)(x)\Big]_{t=0} \\
&= \frac{d}{dt}\Big[(\alpha_{ge^{t\,\mathrm{Ad}(g^{-1})X}} u)(x)\Big]_{t=0} \\
&= \alpha_g\,\mathrm{Ad}(g^{-1}) \cdot u.
\end{aligned} \tag{82.73}$$

Then

$$p'_X(\alpha_g u) = \|\alpha_g\,\mathrm{Ad}(g^{-1})X \cdot u\|_{\mathbb{A}},$$

but the action is isometric with respect to $\|.\|_{\mathbb{A}}$ because it is a supremum norm. It leads to

$$p'_X(\alpha_g u) = p_{\mathrm{Ad}(g^{-1})X} u$$

and there are no reason for $\|\,\mathrm{Ad}(g^{-1})X \cdot u\|_{\mathbb{A}}$ to be equal to $\|X \cdot u\|_{\mathbb{A}}$.

Semi-norms in which the action is isometric are given by

$$p_X(y) = \sup_{g\in R} \|X \cdot \alpha_g(u)\|_{\mathbb{A}}. \tag{82.74}$$

We put on $\mathbb{A}^\infty$ the topology of these seminorms when $X \in \mathcal{U}(\mathbb{R})$, the enveloping algebra of $\mathbb{R}$, see 52.3. This is the quotient of the tensor algebra $T(\mathbb{R})$ by relation $X \otimes Y - Y \otimes X \sim [X, Y]$.

It is immediate to see that $\mathbb{A}^\infty$ is Hausdorff with this topology by taking $X = 1 \in \mathcal{U}(\mathbb{R})$:

$$p_1(w) = \sum_{g\in R} \|\alpha_g(w)\|_{\mathbb{A}},$$

but

$$\|\alpha_g w\|_{\mathbb{A}} = \sum_{x\in R} |\alpha_g w(x)|_{\mathbb{C}} = \sup_{x\in R} |w(xg)|_{\mathbb{C}} \tag{82.75}$$

which is zero only if w is identically zero.

82.8.5 Convergence condition

Our strategy is to transform the integral over $R \times R$ into an integral over $\mathbb{R}^2 \times \mathbb{R}^2$ and to reduce the problem to a simple application of Rieffel's method described in [952].

Let α be the regular right action of R on $\mathrm{Fun}(R)$ and the integral

$$\iint_{R\times R} \cosh(a_1 - a_2)\alpha_{x_1}(u)\alpha_{x_2}(v)e^{i[\sinh(a_1)l_2]-\sinh(a_2)l_1} \tag{82.76}$$

in which we perform the change of variable $\xi_j = \sinh(a_j)$, $da_j = \frac{d\xi_j}{\sqrt{1+\xi_j^2}}$. The phase becomes $\xi_j l_2 - \xi_2 l_1 = \omega(p_1, p_2)$ where ω is the usual symplectic form. Since $\cosh(a_1 - a_2) = \sqrt{1+\xi_1^2}\sqrt{1+\xi_2^2}$, the integral to be computed is

$$\int_{\mathbb{R}}\int_{\mathbb{R}^2} \frac{\sqrt{1+\xi_1^2}\sqrt{1+\xi_2^2} - \xi_1\xi_1}{\sqrt{1+\xi_1^2}\sqrt{1+\xi_2^2}} e^{i\omega(p_1,p_2)}\tilde{\alpha}_{p_1}(u)\tilde{\alpha}_{p_2}(v)$$

where

$$\big(\tilde{\alpha}_p(u)\big)(x,y) = \big(R^*_{(\operatorname{arcsinh}\xi,l)}u\big)(x,y) = u(x + \operatorname{arcsinh}\xi, l + e^{-2\operatorname{arcsinh}\xi}y). \tag{82.77}$$

This is not a bounded operator on $C^\infty(R)$. It is easy to see that the fraction is a bounded function and that all the derivatives are also bounded by studying the function $\beta(\xi) = \xi/\sqrt{1+\xi^2}$.

Let us see more precisely $\tilde{\alpha}(u)$:

$$\begin{aligned}(\tilde{\alpha}_p y)(x,y) &= u\big((x,y)(\operatorname{arcsinh}\xi, l)\big)\\ &= u\big(x + \operatorname{arcsinh}\xi, l + e^{-2\operatorname{arcsinh}\xi}y\big)\\ &= u\big(x + \operatorname{arcsinh}\xi, l + y(\sqrt{\xi^2+1} - \xi)^2\big).\end{aligned} \tag{82.78}$$

Then

$$\partial_\xi(\tilde{\alpha}_p u)(x,y) = (\partial_x u)(\ldots)\frac{1}{\sqrt{1+\xi^2}} + y(\partial_y u)(\ldots)2(\sqrt{\xi^2-1})(\frac{\xi}{\sqrt{\xi^2+1}-2}) \tag{82.79}$$

The conclusion is that

$$\begin{aligned}\partial_\xi(\tilde{\alpha}_p u)(x,y) &= \tilde{\alpha}_p(\partial_x u)(x,y)f_1(\xi)\\ &\quad + \tilde{\alpha}_p(y\partial_y u)(x,y)f_2(\xi).\end{aligned} \tag{82.80}$$

where f_1 and f_2 are bounded functions with all derivative bounded, i.e. functions in the space of Rieffel. In the same way,

$$(\partial_l\tilde{\alpha}u)(x,y) = \tilde{\alpha}_p(\partial_y u). \tag{82.81}$$

If we put $X = \partial_x$, $Y = y\partial_y$ and $Z = \partial_y$

We are studying the integral

$$\int_{\mathbb{R}^2\times\mathbb{R}^2} B(\xi_1,\xi_2)\tilde{\alpha}_{p_1}(u)\tilde{\alpha}_{p_2}(v)e^{i\omega(p_1,p_2)}\,dp_1\,dp_2.$$

Since the phase function has the same form as the one of Rieffel, the condition that we have to impose on u and v must be such that $\tilde{\alpha}_p(u)$ fulfils Rieffel's condition, namely $\partial_\xi(\tilde{\alpha}u)$, $\partial_l(\tilde{\alpha}u)$ and so on must be bounded. This leads us to only consider functions u on R such that $\partial_x^k u$ and $y^n\partial_y^m u$ are bounded for all k and $m \geqslant n$.

Let us now more precisely the Fréchet algebra issues. For u, $v \in \mathrm{Fun}(R)$, we pose

$$\begin{aligned}u \star^R v &= \int_{R\times R} \cosh(a_1 - a_2)\alpha_{x_1}(u)\alpha_{x_2}(v)e^{[\cdots]}\,dx_1\,dx_2\\ &= \int_{\mathbb{R}^2\times\mathbb{R}^2} B(\xi_1,\xi_2)\tilde{\alpha}_{p_1}(u)\tilde{\alpha}_{p_2}(v)e^{i\omega(p_1,p_2)}\,dp_1\,dp_2.\end{aligned}$$

We want the function $F\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathrm{Fun}(R)$ given by

$$F(p_1, p_2) = \tilde{\alpha}_{p_1}(u)\tilde{\alpha}_{p_2}(v)$$

We put the following seminorms on $C^\infty(R)$:

$$\|u\|_k = \sup_{(a,l)\in R} |l^k u(a,l)|; \tag{82.82}$$

this is an increasing family. We say that A is the set of $C^\infty(R)$ functions which are bounded for each of these seminorms. It is Fréchet because

$$\begin{aligned}\|uv\|_k &= \sup |l^k u(a,l)v(a,l)| \\ &\leqslant \sup |l^k u(a,l)| \sup |l^k v(a,l)| \\ &= \|u\|_k \|v\|_k.\end{aligned}$$

We have

$$\begin{aligned}(\partial_X u)(x,y) &= \frac{d}{dt}\Big[\big(\tilde{\alpha}_{tX}(u)\big)(x,y)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[u\big(x + \operatorname{arcsinh}(tX_x), tX_y + e^{-2\operatorname{arcsinh} tX_x} y\big)\Big]_{t=0} \\ &= \partial_x u + (\partial_y u)\frac{d}{dt}\Big[yX_y + e^{-2\operatorname{arcsinh} tX_x} y\Big]_{t=0} \\ &= \partial_x u + X_y(\partial_y u) - 2y(\partial_y u).\end{aligned} \tag{82.83}$$

82.8.6 Some inequalities in $C^\infty(R)$

Lemma 82.16.
For each $u \in A^\infty$, we have

(1) $\|u\|_{jk} \leqslant \|u\|_{(j+1)k}$,

(2) $\|\partial_t u\|_{jk} \leqslant c_k \|u\|_{j(k+1)}$,

(3) $\|yu\|_{jk} \leqslant \|u\|_{(j+1)k}$,

where t denotes x or y, and yu denotes the function $(x,y) \mapsto yu(x,y)$. In particular

$$\|y\partial_y u\|_{jk} \leqslant c_k \|u\|_{(j+1)(k+1)} \tag{82.84}$$

Proof. The first claim is immediate. For the second, remark that each term in $\|u\|_{j(k+1)}$ has a corresponding term in $\|\partial_t u\|_{jk}$ with a different coefficient. For example, the term $\sup_{i\leqslant j} \|\partial_x^2 u\|_i$ in $\|\partial_x u\|_{jk}$ corresponds to $\frac{1}{2}\sup_{i\leqslant j} \|\partial_x^2 u\|_i$ in $\|u\|_{j(k+1)}$. These coefficients are depend only on k and can be majored by a constant c_k. This proves the second inequality.

For the next inequality, computes

$$\begin{aligned}\|yu\|_{jk} &= \sup_{i\leqslant j} \sum_{|\mu|\leqslant k} \frac{1}{\mu}\|\partial_\mu(yu)\|_i \\ &\leqslant \sup_{i\leqslant j} \sum_{\mu} \frac{1}{\mu}\|y\partial_\mu\|_i \\ &= \sup_{i\leqslant j} \sum_{\mu} \frac{1}{\mu}\|\partial_\mu u\|_{i+1}.\end{aligned} \tag{82.85}$$

The set on which the supremum is taken for $\|u\|_{(j+1)k}$ has the extra term $\sum_\mu \frac{1}{\mu}\|\partial_\mu\|_0$. Then $\|yu\|_{jk} \leqslant \|u\|_{(j+1)k}$.

□

Proposition 82.17.
For each $X \in V$ and $u \in C^\infty(R)$, we have

$$\|\tilde{\alpha}_X u\|_{jk} \leqslant c_{kX} \|u\|_{(j+1)(k+1)} \tag{82.86}$$

where c_{kX} is a constant which only depend on X and k.

Proof. From definitions,

$$\begin{aligned}(\tilde{\alpha}_X u)(x,y) &= \frac{d}{dt}\Big[\tilde{\alpha}_{tX} u(x,y)\Big]_{t=0} \\ &= X_\xi(\partial_x u)(x,y) + [X_l - 2X_\xi y](\partial_y u)(x,y).\end{aligned} \tag{82.87}$$

Then, using the lemma,

$$\begin{aligned}\|\tilde{\alpha}_X u\|_{jk} &\leqslant |X_\xi| \|\partial_x u\|_{jk} + |X_l| \|\partial_y u\|_{jk} + 2|X_\xi| \|y \partial_y u\|_{jk} \\ &\leqslant c_k \|u\|_{(j+1)(k+1)}.\end{aligned} \tag{82.88}$$

□

Proposition 82.18.
Let $u \in A^\infty$ then $F_u \in C_u(V, A^\infty)$ defined by

$$F_u(p) = \tilde{\alpha}_p(u).$$

is a smooth vector for the action τ of V on $C_u(V, A^\infty)$ by translation:

$$(\tau_r F)(p) = \tilde{\alpha}_{p+r}(u).$$

Proof. Let

$$\begin{aligned} f \colon V &\to C_u(V, A^\infty) \\ r &\mapsto \tau_r F_u. \end{aligned} \tag{82.89}$$

and

$$\begin{aligned} g \colon V &\to AA^\infty \\ r &\mapsto \tilde{\alpha}_r(u). \end{aligned} \tag{82.90}$$

The map g is smooth by definition of A^∞. One can write

$$(\tau_r F_u)(p) = \tilde{\alpha}_{p+r}(u) = g(p+r) = (g \circ (+r))(p), \tag{82.91}$$

so that

$$f = g \circ (+),$$

where $(+) \colon V \to \operatorname{Trans} V$ is the map which to $v \in V$ assign the translation by v. When we write $g \circ (+)$, we identify $\operatorname{Trans} V$ with V itself. So by $f = g \circ (+)$, we mean the chain

$$V \overset{(+)}{\to} \operatorname{Trans} V \overset{\mathrm{Id}}{\to} V \overset{g}{\to} C_u(C, A^\infty).$$

This is a composition of smooth maps.

□

82.8.7 Compactly supported function

When $u \in C_c^\infty(R)$, F_u is bounded:

$$\|F_u(p)\|_{jk} = \|\tilde{\alpha}_p(u)\| = \sup_{i \leqslant j} \sum_{\mu \leqslant k} \frac{1}{\mu!} \|\partial_\mu(\tilde{\alpha}_p(u))\|_i$$

which is bounded because u is compact supported.

Proposition 82.19.
The smooth compact supported functions on R are smooth vectors of the action $\tilde{\alpha}$.

Proof. We are looking at

$$f(r) = \tilde{\alpha}_r(u) = u\big(\alpha(x,r),\beta(x,r)\big)$$

for certain functions α and β. Let us, for an induction argument a map $g\colon V \to C^\infty(R)$,

$$g(r)(x,y) = a\big(\gamma(x,y),\delta(y,r)\big) \tag{82.92}$$

eq. formfind

where a is smooth and compactly supported and γ and δ are smooth but not special compactly supported and neither bounded. We suppose however that γ and δ are bounded on every bounded set. We have

$$\begin{aligned}(\partial_X g)(r)(x,y) &= \frac{d}{dt}\Big[g(r+tX)(x,y)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[a\big(\gamma(x,r+tX),\delta(y,r+tX)\big)\Big]_{t=0}\\ &= (\partial_x a)\big(\gamma(x,r),\delta(y,r)\big)\big(X_\xi(\partial_\xi\gamma)(x,r) + X_l(\partial_l\gamma)(x,r)\big)\\ &\quad + (\partial_y a)\big(\gamma(x,r),\delta(y,r)\big)\big(X_\xi(\partial_\xi\delta)(y,r) + X_l(\partial_l\delta)(y,r)\big).\end{aligned} \tag{82.93}$$

The functions $(\partial_x a)$ and $\partial_y a$ are compact supported and the whole expression is bounded. Then the element $(\partial_X g)(r)$ is a sum and product of functions of the type of g itself. By induction, any derivative of f will be of the same type.

We have now to prove that $g\colon V \to C^\infty(R)$ is continuous. Let $\mathcal{O}$ be an open subset of $C^\infty(R)$:

$$\mathcal{O} = \{h \in C^\infty(R) \text{ tel que } \sup_{(x,y)\in R} |y^j h(x,y)| \leqslant \varepsilon\ \forall j \geqslant j_0\}.$$

Let $g(r) \in \mathcal{O}$. We want to prove that $g(r+\varepsilon')$ belongs to $\mathcal{O}$ too when ε' is small enough. From continuity, taking a suitable $t \in V$, we can get

$$\big|y^j a\big(\gamma(x,r),\delta(y,r)\big) - y^j a\big(\gamma(x,r+t),\delta(y,r+t)\big)\big| \leqslant \varepsilon'$$

Then

$$\begin{aligned}\sup_{(x,y)\in R}\Big|y^j a\big(\gamma(x,r+t),\delta(y,r+t)\big)\Big| &= \sup_{(x,y)\in R}\Big|y^j a\big(\gamma(x,r+t),\delta(y,r+t)\big)\\ &\qquad - y^j a\big(\gamma(x,r),\delta(y,r)\big)\\ &\qquad + y^j a\big(\gamma(x,r),\delta(y,r)\big)\Big|\\ &\leqslant \sup_{(x,y)\in R}\Big|y^j a\big(\gamma(x,r+t),\delta(y,r+t)\big)\\ &\qquad - y^j a\big(\gamma(x,r),\delta(y,r)\Big|\\ &\qquad + \sup_{(x,y)\in R}\Big|y^j a\big(\gamma(x,r),\delta(y,r)\Big|\\ &\leqslant \varepsilon' + \varepsilon.\end{aligned} \tag{82.94}$$

We conclude that a map of the form (82.92) is continuous and the induction proves that any compact supported smooth function on R is a smooth vector for the action $\tilde{\alpha}$.

□

82.9 The construction

We consider the Fréchet algebra $A = C^\infty_{XYZ}(R)$, the space of smooth function which are bounded for the following seminorms:

$$\|u\|_j = \sup_{(x,y)\in R} |y^j u(x,y)|. \tag{82.95}$$

We consider the action $\tilde{\alpha}$ of $V = \mathbb{R}^2$ on A defined by

$$\tilde{\alpha}_{(\xi,l)}(u)(x,y) = u\big(x + \operatorname{arcsinh}\xi, l + e^{-2\operatorname{arcsinh}\xi} y\big) \tag{82.96}$$

Lemma 82.20.
This action is intern on A, i.e. for all $p \in V$ and $u \in A$, $\tilde{\alpha}_p(u) \in A$.

Proof. We have to prove that, for each j, the quantity

$$\|\tilde{\alpha}_p(u)\|_j = \sup_{(x,y)} |y^j u(x + C, B + Ay)|$$

is bounded for arbitrary constants[6] A, B, $C \in \mathbb{R}$ with $A \neq 0$. We can completely forget C because of the sup on x, but we have to take care about A and B because y appears one time under the form $Ay + B$ and one under the form y^j. So we have to perform the change of variable $z = Ay + B$ in the sup:

$$\begin{aligned}
\|\tilde{\alpha}_p(u)\|_j &= \sup_{(x,z)} \left| \Big(\frac{z - B}{A}\Big)^j u(x,z) \right| = \Big|\frac{1}{A}\Big|^j \sup_{(x,z)} |(z - B)^j u(x,z)| \\
&= \Big|\frac{1}{A}\Big|^j \sup_{(x,z)} |[z^j - jBz^{j-1} \pm \cdots] u(x,z)| \\
&\leqslant \Big|\frac{1}{A}\Big|^j \left[\sup_{(x,z)} |z^j u| - Bj \sup_{(x,z)} |z^{j-1} u| \pm \cdots \right] \\
&= \Big|\frac{1}{A}\Big|^j \big[\|u\|_j - Bj \|u\|_{j-1} \pm \cdots \big]
\end{aligned} \tag{82.97}$$

where each term is finite because $u \in C^\infty_{XYZ}(R)$.

□

Now we consider A^∞, the Fréchet algebra of smooth vectors of the action $\tilde{\alpha}$ endowed with the seminorms

$$\|u\|_{jk} = \sum_{i \leqslant j} \sum_{|\alpha| \leqslant k} \frac{1}{\alpha!} \|\partial_\alpha u\|_i \tag{82.98}$$

where

$$\partial_X u = \frac{d}{dt}\Big[\tilde{\alpha}_{tX}(u)\Big]_{t=0}.$$

Lemma 82.21.
Let suppose that the map $\tilde{\alpha} \colon A \to A$ is continuous for each $p \in V$. If $u \in A^\infty$, then $\partial_X u \in A^\infty$.

Proof. Let $g \colon V \to A$ given by $g(r) = \tilde{\alpha}_r(\partial_X u)$. Our purpose is to prove that g is smooth. The derivative $\partial_Y g$ is given by

$$\begin{aligned}
(\partial_Y g)(r) &= \frac{d}{dt}\Big[g(r + tY)\Big]_{t=0} = \frac{d}{dt}\Big[\tilde{\alpha}_{r+tY} \frac{d}{ds}\Big[\tilde{\alpha}_{sX} u\Big]_{s=0}\Big]_{t=0} \\
&= \frac{d}{dt}\frac{d}{ds}\Big[\tilde{\alpha}_{r+tY+sX} u\Big]_{\substack{s=0 \\ t=0}} = (\partial_{XY} f)(r)
\end{aligned}$$

where f is smooth because it is the function $f(p) = \tilde{\alpha}_p u$. The continuity of $\tilde{\alpha}_p$ we used when we inverted $\tilde{\alpha}_{r+tY}$ and d/ds.

□

6. In fact, A and B are related but it doesn't matter here.

82.9.1 The problem with Rieffel

Let $\mathbb{A}$ be a C^*-algebra on which R has a strongly continuous action $\tilde{\alpha}$. The integral that we consider is

$$\int_{\mathbb{R}^2\times\mathbb{R}^2} e^{i\omega(p,q)} B(p,q)\tilde{\alpha}_p(a)\tilde{\alpha}_q(b)\,dp\,dq$$

where $a \in \mathbb{A}$. The theorem of convergence of oscillatory integrals says that

Theorem 82.22.
Let $F \in C^\infty(\mathbb{R}^N, \mathbb{A})$ such that there exists a $m \in \mathbb{R}$ with property that for all multi-index α, there exists a C_α such that

$$|\partial_\eta^\alpha F(\eta)| \leqslant C_\alpha(1+|\eta|)^{m-|\alpha|}, \tag{82.99}$$

then the integral

$$\int_{\mathbb{R}^N\times\mathbb{R}^N} e^{ip\ \cdot\ q} F(p,q)\,dp\,dq$$

makes sense.

When F checks condition (82.99), we say that $F \in \mathscr{S}(\mathbb{R}^N, \mathbb{A})$.

Definition 82.23.
*We say that $a \in \mathbb{A}$ is a **symbolic vector** of order m is the map $x \mapsto \tilde{\alpha}_x(a)$ belongs to $\mathscr{S}^m(\mathbb{R}^2, \mathbb{A})$.*

We denote by $\mathbb{A}^{(m)}$ the set of all symbolic vectors of order m on $\mathbb{A}$. The question is to know if $A^\infty := \bigcup_{m\in\mathbb{R}} \mathbb{A}^{(m)}$ is at least dense in $\mathbb{A}$.

Let us consider the case $\mathbb{A} = C_0(R)$: the functions which decrease to zero at infinity with the right regular action

$$(\alpha_r a)(r') = a(r'r).$$

Since the group law is

$$(a,l)(a',l') = (a+a', l+e^{-2a'}l),$$

we find

$$\big(\tilde{\alpha}_{(\xi,l)}a\big)(x,y) = a\big(x+\operatorname{arcsinh}\xi, l+e^{-2\operatorname{arcsinh}\xi}y\big). \tag{82.100}$$

When we want to perform $\int Fe^{i\omega}$ with methods from Rieffel, we are led to integrate derivatives and among others, we have to integrate over (ξ, l) the expression

$$\begin{aligned}
\partial_\xi\big(\tilde{\alpha}_{(\xi,l)}a\big)(x,y) &= \partial_\xi a\big(x+\operatorname{arcsinh}\xi, l+e^{-2\operatorname{arcsinh}\xi}y\big)\\
&= \frac{1}{\sqrt{1+\xi^2}}(\partial_x a)(x+\operatorname{arcsinh}\xi, l+e^{-2\operatorname{arcsinh}\xi}y)\\
&\quad + 2y\left(\frac{-\xi}{\sqrt{\xi^1-1}}\right)\big(\sqrt{\xi^2+1}-\xi\big)\times\\
&\qquad \times(\partial_y a)(x+\operatorname{arcsinh}\xi, l+e^{-2\operatorname{arcsinh}\xi}y)\\
&= \tilde{\alpha}(\partial_x a)(x,y)f_1(\xi) + y\tilde{\alpha}_p(\partial_y a)(x,y)f_2(\xi).
\end{aligned} \tag{82.101}$$

Since we want to integrate it with respect to p as function $R \to \mathbb{A}$, we have to regroup the dependence in x and y in order to show some bound in the supremum norm. Using formula

$$e^{-2\operatorname{arcsinh}x} = \big(\sqrt{x^2+1}-x\big)^2,$$

we find $\tilde\alpha_{(\xi,l)}y = l + y(\sqrt{\xi^2} - x)^2$ and then

$$\partial_\xi\big(\tilde\alpha_{(\xi,l)}a\big)(x,y) = (\partial_x a)(x + \operatorname{arcsinh}\xi, l + e^{-2\operatorname{arcsinh}\xi}y)f_1(\xi) \tag{82.102a}$$

$$+ (\partial_y a)(x + \operatorname{arcsinh}\xi, l + e^{-2\operatorname{arcsinh}\xi}y)\times \tag{82.102b}$$

$$\times\, 2y\big(\sqrt{\xi^2+1} - \xi\big)\left(\frac{1}{2}\frac{2\xi}{\sqrt{\xi^2+1}} - 1\right) \tag{82.102c}$$

$$= \frac{-2}{\sqrt{\xi^2+1}}\Big[\tilde\alpha_{(\xi,l)}(\partial_y a)\big(\tilde\alpha_{(\xi,l)}y\big) - \tilde\alpha_p(\partial_y a)l\Big] \tag{82.102d}$$

Using the fact that $\tilde\alpha$ is a representation,

$$\partial_\xi(\tilde\alpha_{(\xi,l)}a) = \tilde\alpha(\partial_x a)f_1(\xi) - \frac{2}{\sqrt{\xi^2+1}}\Big[\tilde\alpha_p(y\partial_y a) - \tilde\alpha_p(\partial_y a)l\Big]$$

whatever the (semi)norms in $\mathbb{A}$ are, one cannot hope a better upper bound that something like

$$|\partial_\xi\tilde\alpha_p a| \leqslant |f_1(\xi)|\|\tilde\alpha(\partial_x a)\|_k + |f_2(\xi)|\|\tilde\alpha_p(y\partial_y a)\|_k - |l||f_2(\xi)|\|\tilde\alpha\partial_y a\|_k$$

The presence of the l factor dooms any attempt to perform the integral with Rieffel methods.

$$\mathcal{FIN}$$

82.10 Klein-Gordon calculus

82.10.1 Affine space primer

Let V be a vector space. The **affine space** $A(V)$ is the set V endowed with the notion of line (inherited from V as lateral classes of one dimensional vector subspace), parallel lines (inherited from V) and of the section ratio, i.e. a function $\rho\colon A(V)^3 \to \mathbb{R}$ such that if a, b and c are aligned,

$$c - a = \rho(a,b,c)(b-a).$$

A localised vector in $A(V)$ is a pair $(a,b) \in A(V)^2$. We put the following equivalence relation on the set of localised vectors: $(a,b) \sim (c,d)$ if and only if $b - a = d - c$. Although expressions $b-a$ and $d-c$ need the vector space structure of V, one can prove that this relation can be expressed only in terms of the affine structure. An equivalence class is a **free vector** (in the sense of freedom). We denote by $\overrightarrow{ab}$ the class of (a,b). For each $\overrightarrow{ab}$ and $c \in A(V)$, there exists one and only one $d \in A(V)$ such that $\overrightarrow{ab} = \overrightarrow{cd}$. So a free vector defines a permutation of points in $\mathbb{M}$; we say that it is a **translation**.

If A is a linear transformation of V, we define the action of A on a free vector by

$$A(\overrightarrow{ab}) = \overrightarrow{(Aa)(Ab)}$$

This definition doesn't depends on the choice of a and b because linearity of A gives $\tilde Pb' - \tilde Pa' = \tilde Pb - \tilde Pa$ and thus $\tilde P(\overrightarrow{a'b'}) = \overrightarrow{\tilde Pa'\tilde Pb'} = \overrightarrow{\tilde Pa\tilde Pb}$.

82.10.2 Space-time structure

Let $n \geqslant 1$. The **Minkowski space** or **space-time** $\mathbb{M}$ is a real affine space of dimension $n+1$ endowed with a quadratic form on the free vector space that, from a physical usage, we denote by ds which has signature $-,+,\dots,+$. The quantity $ds(v,v)$ is usually written as $ds^2(v)$.

The **Poincaré group** $\mathscr{P}$ is the group of isometries whose identity component is denoted by $\mathscr{P}_0$. The vector space of translations in $\mathbb{M}$ is $\tilde{\mathbb{M}}$. When $P \in \mathscr{P}$, $\tilde P$ is the linear part of P. The **Lorentz group** is

$$\mathscr{L} = \{\tilde P \text{ tel que } P \in \mathscr{P}\}.$$

An unrestricted observer is a triple $\omega = (x, T, E)$ where

— $x \in \mathbb{M}$,

— T and E are vector subspaces endowed with an orientation with $\dim T = 1$ and $\dim E = n$,

— $ds|_T$ is positive definite and $ds|_E$ is negative definite,

— $ds(T, E) = 0$.

For the two last items, we have to define $ds(P, Q)$ when P and Q are translations in $\mathbb{M}$, i.e. are free vectors. For, we choose $a \in \mathbb{M}$ and then P_a, Q_a such that $P = \overrightarrow{aP_a}$ and $Q = \overrightarrow{aQ_a}$ and we define $ds(P, Q) = ds(P_a, Q_a)$. In order to prove that it doesn't depends on the choice of a, we remark that $P_b = P_a - a + b$ and that ds, being an affine notions, is invariant under translations. So

$$ds(P_a - a + b, Q_a - a + b) = ds(\overrightarrow{ba}(P_a - a + b), \overrightarrow{ba}(Q_a - a + b)) = ds(P_a, Q_a).$$

Note that the orientations are part of the definition of an observer. So it is possible to have $(x, T_1, E_1) \neq (x, T_2, E_2)$ as observers but $T_1 = T_2$ and $E_1 = E_2$ as (non oriented) vector spaces.

The Poincaré group acts on the space of unrestricted observers by

$$P \cdot (x, T, E) = (Px, \tilde{P}T, \tilde{P}E) \tag{82.103}$$

where we endow $\tilde{P}T$ and $\tilde{P}E$ with orientations such that $\tilde{P}$ is positive. This action is transitive and the stabilizer of (x, T, E) is canonically isomorphic to the group $O^+(E)$ of rotations in E[7].

Let $\omega = (x, T, E)$. The **spacial symmetry** P_ω of the observer ω is the unique affine transformation of $\mathbb{M}$ such that

— $P_\omega x = x$,

— $\tilde{P}_\omega z = z$ if $z \in T$ and $\tilde{P}_\omega z = -z$ if $z \in E$.

Let us prove the unicity. For, we pick a $y \in \mathbb{M}$ and we write $P = \overrightarrow{ab} \circ \tilde{P}$ where $\tilde{P}$ is Lorentz and $P\overrightarrow{ab}$ is a translation. A $z \in E$ is a translation and can be written under the form $z = \overrightarrow{xz}$ for a certain $z \in \mathbb{M}$. From definitions,

$$\tilde{P}_\omega(\overrightarrow{xz}) = \overrightarrow{\tilde{P}_\omega x \tilde{P}_\omega z} = \overrightarrow{x\tilde{P}_\omega z}$$

Therefore $\tilde{P}_\omega(\overrightarrow{xz}) = \overrightarrow{x\tilde{P}_\omega z} = \pm\overrightarrow{xz}$ with the minus sign when $z \in T$ and the plus sign when $z \in E$. A general $y \in \mathbb{M}$ can be decomposed as $y = y_T + y_E$, and finally, we find

$$P_\omega y = x + \mathtt{proj}_T(\overrightarrow{xy}) - \mathtt{proj}_E(\overrightarrow{xy}) \tag{82.104}$$ eq_PomegaprT

where $\mathtt{proj}_T(\overrightarrow{xy})$ is a translation in $\mathbb{M}$ identified with a point in $\mathbb{M}$ by $\overrightarrow{ab} \simeq r$ if $\overrightarrow{ab} = \overrightarrow{xr}$.

If n is odd, the spacial symmetry P_ω has a negative determinant and therefore does not belong to $\mathscr{P}_0$.

82.10.3 Observer

We now fix a reference unrestricted observer $\omega_0 = (x_0, T_0, E_0)$. Then an **observer** is an unrestricted observer of the form $\omega = P \cdot \omega_0$ with $P \in \mathscr{P}_0$. We denote by Ω the space of all observers. By restricting ds to T_0 and $-ds$ to E_0, we build euclidian spaces T_0 and E_0. Let us write $\|.\|_0$ the norm on E_0.

We denote by ϵ_0 the positive unit vector in T_0. Let $P \in \mathscr{P}$; we can write $\tilde{P}\epsilon_0 = \alpha\epsilon_0 + w$ with $\alpha > 0$ (because $\mathscr{P}_0$ is orthochrone) and $\alpha^2 - \|w\|_0^2 = 1$. Indeed

$$\begin{aligned} ds^2(\tilde{P}\epsilon_0) &= ds(\alpha\epsilon_0, +w, \alpha\epsilon_0 + w) \\ &= ds^2(\alpha\epsilon_0) + ds^2(w) \qquad \text{from } ds(E_0, T_0) = 0 \\ &= \alpha^2 - \|w\|_0^2 \end{aligned}$$

because $ds^2(\epsilon_0) = 1$ and $\|.\|_0$ is the restriction of $-ds$ to E_0.

7. Moi je crois que ce O^+, c'est SO(3) parce que c'est des rotations qui doivent garder l'orientation.

We define $v_{\tilde{P}} = \alpha^{-1}w$ which is an element of E_0 of norm $\leqslant 1$ because

$$\|v_{\tilde{P}}\|_0^2 = \alpha^{-2}\|w\|_0^2 = \alpha^{-2}(\alpha^2 - 1) = 1 - \frac{1}{\alpha^2}.$$

When P runs over $\mathscr{P}$, $v_{\tilde{P}}$ spans the unit open ball $B_{E_0}(o, 1)$. We say that $v_{\tilde{P}}$ is the velocity of the observer $P \cdot \omega_0$ with respect to ω_0 in the frame of ω_0.

Proposition 82.24.
Let $\omega_0 = (x_0, T_0, E_0) \in \Omega$. The map

$$\Omega = \mathscr{P}_0/O^+(E_0) \to \mathbb{M} \times B_{E_0}(o, 1)$$

given by the quotient of $P \mapsto (Px_0 - x_0, v_{\tilde{P}})$ is bijective.

Proof. The action of $\mathscr{P}_0$ on Ω is transitive and the stabilizer of ω_0 is O^+, so as homogeneous spaces, $\Omega = \mathscr{P}_0/O^+(E_0)$. For the remaining, let us choose a positive oriented orthonormal basis which allows us to identify $\mathbb{M}$ and $\tilde{\mathbb{M}}$. For this, we choose $o = x_0$, $ds^2 = dt^2 - d\overline{x}^2$ in such a way that $T_O = \mathbb{R} \times \{0\}$ and $E_0 = \{0\} \times \mathbb{R}^n$. The spatial symmetry P_{ω_0} is given by

$$J(t, \overline{(}x)) = (t, -\overline{(}x))$$

and the Lorentz group is made up with matrices M which satisfies

$$M^t J M = J.$$

The vector ϵ_0 is the first vector of the chosen basis $(\epsilon_i)_{i\geqslant 0}$ of $\mathbb{R}^{n+1}$. The group $O^+(E_0) = \{1\} \times \mathrm{SO}(n)$ is the subgroup of $\mathscr{P}_0$ which leaves ϵ_0 unchanged. In fact, $\mathscr{P}_0$ is the semi-direct product

$$\mathscr{P}_0 = \mathbb{R}^{n+1} \times_\rho \mathscr{L}_0 \tag{82.105}$$

with $\rho_M(a') = Ma'$.

A **boost** is an selfadjoint positive definite matrix of $\mathscr{L}$. This is automatically in $\mathscr{L}_0$ and antisymmetric (because we consider real spaces). These are "rotations" in mixed space-times planes. One can show that any element M in $\mathscr{L}$ can be decomposed into $M = AK$ where A is a boost and K a matrix of $\mathrm{SO}(n)$. Now we ask ourself which elements of $\mathscr{P}$ define the same observer. For a given $(a, M) \in \mathscr{P}$, which are the (a', M') such that $(a, M) \cdot \omega_0 = (a', M') \cdot \omega_0$? In order to find $E = E'$ and $T = T'$, we have to put $a = a'$. Then we need $M\epsilon_0 = M'\epsilon_0$ and $ME_0 = M'E_0$ which in turn imposes M and M' to differ only by a matrix of $\mathrm{SO}(n)$: there exists $S \in \mathrm{SO}(n)$ such that $M = M'S$. It proves that Ω is parametrized by translations and boosts. It remains to prove that $\mathbb{M} \times B(E_0)$ can too.

From now, we take the convention to put a bar over vectors in $\mathbb{R}^n$ in order to distinguish them from vectors in $\mathbb{R}^{n+1}$. A vector $\xi \in \mathbb{R}^{n+1}$ is thus written $\begin{pmatrix} \xi_0 \\ \overline{\xi} \end{pmatrix}$ with $\xi_0 \in \mathbb{R}$ and $\overline{\xi} \in \mathbb{R}^n$.

To a $\overline{v} \in \mathbb{R}^n$ with $|\overline{v}| < 1$ we associate

$$(1 - |\overline{v}|^2)^{1/2} \begin{pmatrix} 1 \\ v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^{n+1}.$$

Now we prove that formula

$$A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (1 - |\overline{v}|^2)^{-1/2} \begin{pmatrix} 1 \\ v_1 \\ \vdots \\ v_n \end{pmatrix} \tag{82.106}$$

defines a bijection between the boosts A and the $|\overline{v}| < 1$. It is first clear that this formula gives only one $\overline{v}$ for each boost A. Now let $K_0 \in \mathrm{SO}(n)$ and $\mu \in \mathbb{R}$ such that $\overline{v} = \tanh\mu\, K_0\epsilon_1$, or

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (\tanh\mu) K_0 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It is possible because one can always find a matrix in $\mathrm{SO}(n)$ which places ϵ_1 in the right direction and since tanh is surjective on $]-1,1[$, it is a suitable function to rescale $K_0\epsilon_1$. Let

$$K = \begin{pmatrix} 1 & 0 \\ 0 & K_0 \end{pmatrix},$$

we have

$$K^{-1}AK\epsilon_0 = K^{-1}A\epsilon_0 = K^{-1}(1-|\overline{v}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v} \end{pmatrix}$$

and $(1-|\overline{v}|^2)^{-1/2} = \cosh\mu$. It leads to

$$\begin{aligned} K^{-1}AK\epsilon_0 &= \cosh\mu(K^{-1}\epsilon_0) + \cosh\mu(K^{-1}v) \\ &= \cosh\mu\epsilon_0 + \sinh\mu\epsilon_1. \end{aligned}$$

Therefore the first column of $K^{-1}AK$ is $\begin{pmatrix} \cosh\mu \\ \sinh\mu \\ 0 \\ \vdots \end{pmatrix}$ and finally the form of $K^{-1}AK$ is fixed by

$$K^{-1}AK = \begin{pmatrix} \cosh\mu & \sinh\mu & \\ \sinh\mu & \cosh\mu & \\ & & \mathbb{1} \end{pmatrix}.$$

This concludes the proof.

□

The latter matrix is called *boost in direction of* ∂_1. It represent a "rotation" in the plane (t, x_1).

82.10.4 Action of $\mathscr{P}_0$ on Ω

We want now to describe the action of $\mathscr{P}_0$ on Ω in the realization $\Omega = \mathbb{R}^{n+1} \times B_n$ where $B_n = \{v \in \mathbb{R}^n \text{ tel que } |v| < 1\}$. Let $\omega = (N, x) \cdot \omega_0$ and suppose that it is described by (x, v) (recall that $x_o = o$, thus (N, x) must be of the form $(x, .)$). Our aim is to express v in terms of ω_0 and N. If ϵ_T is the positive unit vector in T,

$$\epsilon_T = N\epsilon_0 = AS\epsilon_0 = A\epsilon_0 = (1-|\overline{v}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v} \end{pmatrix} \tag{82.107}$$

eq_epsTv

from definition of $\overline{v}$. Now, take $\omega = (x, T, E)$, $(M, a) \in \mathscr{P}_0$ and $\omega' = (M, a) \cdot \omega$ corresponds to (x', v'). We have $x' = Mx + a$ and

$$\epsilon_{T'} = M\epsilon_T = M(1-|\overline{v}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v} \end{pmatrix}.$$

On the other hand,

$$\epsilon_{T'} = (1-|\overline{v'}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v'} \end{pmatrix}.$$

We want to establish a link between v and v' with M. Suppose in a first time that $M = \begin{pmatrix} 1 & 0 \\ 0 & K_0 \end{pmatrix}$ with $K_0 \in \mathrm{SO}(n)$. In this case,

$$(1-|\overline{v}|^2)^{-1/2}\begin{pmatrix} A \\ K_0\overline{v}\end{pmatrix} = (1-|\overline{v'}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v'}\end{pmatrix}.$$

The first component of this equation gives $|\overline{v}| = |\overline{v'}|$ while the others leads to $\overline{v'} = K_0\overline{v}$. Now suppose that M is a boost of velocity $w = \tanh\mu$ in the direction ∂_1. It gives

$$(1-|\overline{v}|^2)^{-1/2}\begin{pmatrix} \cosh\mu & \sinh\mu & \\ \sinh\mu & \cosh\mu & \\ & & \mathbb{1}\end{pmatrix}\begin{pmatrix} 1 \\ v_1 \\ \vdots \\ v_n\end{pmatrix}(1-|\overline{v'}|^2)^{-1/2}\begin{pmatrix} 1 \\ \overline{v'}\end{pmatrix}. \tag{82.108}$$

eq. sysumy2

The components 0 and 1 of this equation give the system

$$(1-|\overline{v}'|^2)^{-1/2}\overline{v}'_1 = (1-|\overline{v}|^2)^{-1/2}(\sinh\mu + v_1\cosh\mu) \tag{82.109a}$$

$$(1-|\overline{v}'|^2)^{-1/2} = (1-|\overline{v}'|^2)^{-1/2}(\cosh\mu + v_1\sinh\mu) \tag{82.109b}$$

Substituting the $(1-|\overline{v}'|^2)^{-1/2}$ of the second equation into the left hand side of the first, we find

$$\overline{v}'_1 = \frac{\sinh\mu + v_1\cosh\mu}{\cosh\mu + v_1\sinh\mu}.$$

The kth component of (82.108) with $x = \tanh\mu$ leads[8] to

$$\overline{v}'_1 = \frac{w+v_1}{1+wv_1} \text{ and } v'_k = \frac{v_k\sqrt{1-w^2}}{1+v_1w}.$$

82.10.5 The spacial symmetry P_ω

We want to express P_ω in terms of (x,v). Let us recall that $z \in E$ if and only if $ds(z,\epsilon_T) = 0$, but $ds(x,y) = \langle x, Jy\rangle$ where $\langle .,.\rangle$ is the usual euclidian inner product on $\mathbb{R}^{n+1}$; the space E is characterised by equation $\langle z, J\epsilon_T\rangle = 0$. We begin to express the linear part $\tilde{P}_\omega$ of P_ω. We have $\tilde{P}_\omega z = z$ on T and $\tilde{P}_\omega = -z$ on E. Then

$$\tilde{P}_\omega = -z + 2\langle z, J\epsilon_T\rangle\epsilon_T, \tag{82.110}$$

or

$$\mathtt{proj}_T\, z = \langle z, J\epsilon_T\rangle\epsilon_T \tag{82.111a}$$

$$\mathtt{proj}_E\, z = z - \langle z, J\epsilon_T\rangle\epsilon_T. \tag{82.111b}$$

Lemma 82.25.
Suppose $\omega = (x,T,E)$ and $\omega' = (x',T,E)$, i.e. among others, $\epsilon_T = \epsilon_{T'}$ or ϵ'_T. Then $P_\omega = P'_\omega$ is and only if $x - x'$ is a multiple of ϵ_T.

Proof. Equating formula (82.104) for P_ω and P'_ω,

$$x + \mathtt{proj}_T(y-x) - \mathtt{proj}_E(y-x) = x' + \mathtt{proj}_T(y-x') - \mathtt{proj}_E(y-x)$$

which is true if and only if

$$x - x' = \mathtt{proj}_T(x-x') - \mathtt{proj}_E(x-x'). \tag{82.112}$$

The projection of this equation on T gives nothing while the projection on E leads to $2\,\mathtt{proj}_E(x-x') = 0$, which proves the proposition. □

8. We use formula

$$\cosh\mu = 1/\sqrt{1-\tanh^2\mu}$$

Proposition 82.26.
If $\omega = (x, v) \cdot \omega_0$ and $\omega' = (x', v') \cdot \omega_0$, then $P_\omega = P'_\omega$ if and only if

$$\overline{x} - t\overline{v} = \overline{x}' - t\overline{v}', \tag{82.113a}$$
$$\overline{v} = \overline{v}'. \tag{82.113b}$$

In other words, P_ω depends only on $(\overline{x} - t\overline{v}; \overline{v})$.

Proof. We begin to explicit the condition $P_\omega = P_{\omega_0}$. For this, we need $\epsilon_T = \epsilon_0$ and thus $\overline{v} = 0$. Since $x_0 = o$, the condition $x - x_0 \sim \epsilon_0$ is equivalent to $\overline{x} = o$.

For the same reason, it is cleat that $\overline{v} = \overline{v}'$ is a condition in order to get $P_\omega = P'_\omega$. Now the condition $x - x' \sim \epsilon_T$ is equivalent to the existence of a $\lambda \in \mathbb{R}$ such that

$$\begin{pmatrix} t \\ \overline{x} \end{pmatrix} = \lambda(1 - |\overline{v}'|^2)^{-1/2} \begin{pmatrix} 1 \\ \overline{v}' \end{pmatrix} + \begin{pmatrix} t' \\ \overline{x}' \end{pmatrix}.$$

Taking $\overline{v} = \overline{v}'$ into account, we find the system

$$t = \lambda(1 - |\overline{v}|^2)^{-1/2} + t' \tag{82.114a}$$
$$\overline{x} = \lambda(1 - |\overline{v}|^2)^{-1/2}\overline{v}' + \overline{x}. \tag{82.114b}$$

Taking the difference, we find $\overline{x} - t\overline{v} = \overline{x}' = t'\overline{v}'$. The converse is also true: if $\overline{x} - t\overline{v} = \overline{x}' - t\overline{v}'$, the $\lambda = (1 - |\overline{v}|^2)^{-1/2}(t - t')$ is a solution.

□

82.10.6 Momentum

Let $\tilde{\mathbb{M}}^*$ be the dual of $\tilde{\mathbb{M}}$ on which we consider the dual coordinates of $\tilde{\mathbb{M}}$. We call it the **energy-momentum space** and we write the elements as line matrices: $p = \big(p_0, \overline{p}\big)$. The **mass hyperboloid** is the set

$$\mathscr{M} = \{p \in \tilde{\mathbb{M}}^* \text{ tel que } p_0 > 0 \text{ and } (1 - |\overline{p}|^2)^{1/2}\}. \tag{82.115}$$

There is a bijection $f\colon \mathscr{M} \to B_n$ by the formula

$$p = (1 - |f(p)|^2)^{-1/2} \begin{pmatrix} 1 & -f(p)^t \end{pmatrix}. \tag{82.116a}$$

From now we denote this $f(p)^t$ by $\overline{v}$,and the latter equations is best written

$$p = (1 - |\overline{v}|^2)^{-1/2} \begin{pmatrix} 1 \\ -\overline{v} \end{pmatrix}. \tag{82.116b}$$

Comparing this equation with expression (82.107) of ϵ_T, one sees that

$$\epsilon_T = Jp. \tag{82.117}$$

We have an action of $\mathscr{L}_0$ on B_n given by $\sigma_M\colon B_n \to B_n$,

$$(1 - |\overline{v}|^2)^{-1/2} M \begin{pmatrix} 1 \\ \overline{v} \end{pmatrix} = (1 - |\sigma_M(\overline{v})|^2)^{-1/2} \begin{pmatrix} 1 \\ \sigma_M(\overline{v}) \end{pmatrix}$$

Let us study how this σ acts on $\mathscr{M}$: let $M \in \mathscr{L}_0$, $p \in \mathscr{M}$ and compute $(f^{-1} \circ \sigma_M \circ f)(p)$. As usual, we denote $f(\overline{p})$ by $\overline{v}$: they are related by

$$(1 - |\overline{v}|^2)^{-1/2} \begin{pmatrix} 1 \\ -\overline{v} \end{pmatrix} = \begin{pmatrix} p_0 \\ \overline{p} \end{pmatrix}.$$

The first component of this equality is $(1-|\overline{v}|^2)^{-1/2} = p_0$ while the other gives $\overline{v}^t = -\overline{p}(1-|\overline{v}|^2)^{-1/2}$.

$$f(p) = -\overline{p}/p_0 \tag{82.118}$$

Now we compute $\overline{v}' = \sigma_M(f(p))$ whose definition is

$$(1-|f(p)|^2)^{-1/2} M \begin{pmatrix} 1 \\ f(p) \end{pmatrix} = (1-|\overline{v}'|^2)^{-1/2} \begin{pmatrix} 1 \\ \overline{v}' \end{pmatrix}.$$

Using definition of p_0 in terms of $\overline{p}$ and the fact that $p_0 > 0$, we find

$$p_0 M \begin{pmatrix} 1 \\ f(p) \end{pmatrix} = (1-|\overline{v}'|^2)^{-1/2} \begin{pmatrix} 1 \\ \overline{v}' \end{pmatrix},$$

but

$$\begin{pmatrix} 1 \\ f(p) \end{pmatrix} = \begin{pmatrix} 1 \\ -\overline{p}/p_0 \end{pmatrix} = \frac{1}{p_0} \begin{pmatrix} p_0 \\ -\overline{p} \end{pmatrix} = \frac{1}{p_0} Jp,$$

therefore

$$JMJp = (1-|\overline{v}'|^2)^{-1/2} \begin{pmatrix} 1 \\ -\overline{v}' \end{pmatrix}. \tag{82.119}$$

The latter proves that $f^{-1}(\overline{v}')JMJp$. Finally the action of $\mathscr{L}_0$ on $\mathscr{M}$ is given by

$$p \mapsto M^{t^{-1}} p.$$

The observer space Ω is now renamed **phase space** and is seen as

$$\Omega = \mathbb{R}^{n+1} \times \mathscr{M}. \tag{82.120}$$

The action of $\mathscr{P}_0$ on Ω is given by

$$(M,a) \cdot (x;p) = (Mx+a; JMJp). \tag{82.121}$$

eq_Maxp

lem_boost_inverse

Lemma 82.27.
The boost associated with the momentum $-\overline{p}$ is the inverse or the one associated with $\overline{p}$.

Proof. It is first clear that if $\overline{v}$ is associated with $\overline{p}$, then the boost of $-\overline{p}$ is the one of $-\overline{v}$. So we have

$$A\epsilon_0 = (1-|\overline{v}|^2)^{-1/2} \begin{pmatrix} A \\ \overline{v} \end{pmatrix} \tag{82.122a}$$

and

$$A'\epsilon_0 = (1-|\overline{v}|^2)^{-1/2} \begin{pmatrix} 1 \\ -\overline{v} \end{pmatrix}. \tag{82.122b}$$

If K_0 and μ are such that $\overline{v} = \tanh \mu K_0 \epsilon_0$, then, as far as upper left corner is concerned,

$$K^{-1}AK = \begin{pmatrix} \cosh\mu & \sinh\mu \\ \sinh\mu & \cosh\mu \end{pmatrix},$$

but $-\overline{v} = \tanh(-\mu) K_0 \epsilon_0$, then

$$K^{-1}A'K = \begin{pmatrix} \cosh\mu & -\sinh\mu \\ -\sinh\mu & \cosh\mu \end{pmatrix}$$

and $K^{-1}AKK^{-1}A'K = \mathbb{1}$. This proves that $AA' = \mathbb{1}$. □

82.10.7 Invariant measure

General measure theory is given in section 61.2. A more complete discussion and construction of the present measure on $\mathscr{M}$ is given in the book [1037] We will prove that $p_0^{-1}d\overline{p}$ is an invariant measure on $\mathscr{M}$, but we begin to give a sense to expressions like

$$\int_{\mathscr{M}} f(p)p_0^{-1}d\overline{p}.$$

First, $\mathscr{M}$ is parametrized by

$$\overline{p} \to \begin{pmatrix} \sqrt{1+\overline{p}^2} \\ \overline{p} \end{pmatrix},$$

and we know that dp is an invariant measure on $\mathbb{R}^4$ under the group $\mathscr{L}_0$ because the determinant is 1. Let $f \in C^\infty(\mathscr{M})$, there exists (at least everywhere locally) an extension $\hat{f}$ of f on an open set around $\mathscr{M}$. We parametrize $\mathbb{R}^4$ by $(\overline{p}, y)$ where $y = \|y\|^2$, $dp_0 = \frac{1}{2}p_0^{-1}\,dy$ and we are leads to consider

$$\int_{\mathbb{R}^4} \hat{f}(\overline{p}, y)\frac{dy}{2p_0}\,d\overline{p}$$

where $\frac{dy}{2p_0}d\overline{p}$ is an invariant measure. The mass hyperboloid is given by equation $y = m^2$ (we usually set $m = 1$). The variable y is invariant under LoL_0, so we define "$p_0^{-1}d\overline{p}$" by the expression

$$\int_{\mathscr{M}} f(p)p_0^{-1}d\overline{p} = \int_{\mathbb{R}^4} \hat{f}(\overline{p}, y)\delta(y - m^2)\frac{dy\,d\overline{p}}{\sqrt{m^2+\overline{p}^2}} \tag{82.123}$$

where the δ has to be taken in the sense of distributions. When we integrate this expression with respect to y, we find

$$\int_{\mathbb{R}^3} \hat{f}(\overline{p}, m^2)p_0^{-1}\,d\overline{p}$$

and $\hat{f}(\overline{p}, m^2)$ is exactly what one wants to call $f(\overline{p})$.

82.10.8 Sobolev setting, Bargmann and Wigner representation

We consider the following action of $\mathscr{P}_0$ on $\mathscr{S}'(\mathbb{R}^{n+1})$; if $T \in \mathscr{S}'(\mathbb{R}^{n+1})$ and $(M, a) \in \mathscr{P}_0$, the distribution $U(M, a)$ acts on $\varphi \in \mathscr{S}(\mathbb{R}^{n+1})$ as

$$\langle U(M, a)T, \varphi\rangle = \langle T, \varphi \circ (M, a)\rangle \tag{82.124}$$

where in the right hand side (M, a) is seen as a map $\mathbb{R}^{n+1} \to \mathbb{R}^{n+A}$. Let $\mathscr{P}^\uparrow$ be the maximal subgroup of $\mathscr{P}$ containing $\mathscr{P}_0$ and not the temporal symmetry $(t, \overline{x}) \to (-t, \overline{x})$. This is the **orthochoneous Poincaré group**. The action U extends by the same to $\mathscr{P}^\uparrow$ and becomes the so-called **Wigner-Bargmann representation**. We define in particular

$$\sigma_\omega = U(P_\omega) \tag{82.125}$$

We are here interested in the Sobolev space $H^{1/2}$ defined in 31.13, and more precisely, in the pseudo differential operator of example at page 3560. Using all that, here is some ways to write the scalar product in $H^{1/2}$

$$\begin{aligned}(u, v)_{1/2} &= (Fu, Fv)_{\hat{H}^{1/2}} \\ &= \int (\overline{xi}^2 + 1)^{1/2}(Fu)(\xi)\overline{Fv(\xi)}\,d\overline{\xi} \\ &= \int_{\mathscr{M}} \hat{u}(\xi)\overline{\hat{v}(\xi)}\xi_0\,d\overline{\xi}.\end{aligned} \tag{82.126}$$

We define $\mathcal{G}\colon H^{1/2}(\mathbb{R}^n) \to L^2(\mathscr{M})$ by

$$(\mathcal{G}u)(p) = p_0\hat{u}(\overline{p}); \tag{82.127}$$

this is isometric because

$$\begin{aligned}\|u\|_{1/2}^2 &= \int_{\mathscr{M}} \hat{u}(\overline{xi})\overline{\hat{u}(\overline{\xi})}\xi_0\, d\xi \\ &= \int_{\mathscr{M}} \frac{1}{\xi_0}(\mathcal{G}u)(\xi)\frac{1}{\xi_0}\overline{\mathcal{G}u(\xi)}\xi_0\, d\xi \\ &= \int_{\mathscr{M}} |\mathcal{G}u(\xi)|^2\xi_0^{-1}\, d\overline{\xi} \\ &= \|\mathcal{G}u\|_{L^2}(\mathscr{M}).\end{aligned}$$

Now, for $u \in H^{1/2}(\mathbb{R}^n)$, we pose

$$\tilde{u}(x) = \int_{\mathscr{M}} e^{2i\pi x\,\cdot\,\langle x,p\rangle}(\mathcal{G}u)(p)p_0^{-1}\, d\overline{p} = \int_{\mathscr{M}} e^{2i\pi\langle x,p\rangle}\hat{u}(\overline{p})\, d\overline{p}. \tag{82.128}$$

We have $\tilde{u}(0,\overline{x}) = u(x)$. An other important notion is the **d'Alembert operator**

$$\Box = \partial_t^2 - \Delta$$

on $\mathbb{R}^{n+1}$.

Proposition 82.28.
The prolongation $\tilde{u}$ of $u \in H^{1/2}(\mathbb{R}^n)$ is solution of the Klein-Gordon equation:

$$\Box\tilde{u} = -4\pi^2\tilde{u}$$

Proof. One can invert derivatives and integrals because $H^{1/2}$ is a completion of $\mathscr{S}$, so

$$\begin{aligned}(\Box\tilde{u})(x) &= \int_{\mathscr{M}} \Box\left(e^{2i\pi\langle x,p\rangle}(\mathcal{G}u)(p)p_0^{-1}\right)\, d\overline{p} \\ &= \int_{\mathscr{M}} e^{2i\pi\langle x,p\rangle}(2i\pi)^2(p_0^2 - \overline{p}^2)\mathcal{G}u(p)p_0^{-1}\, d\overline{p} \\ &= -4\pi^2\int_{\mathscr{M}} e^{2i\pi\langle x,p\rangle}p_0\hat{u}(\overline{p})p_0^{-1}\, d\overline{p} \\ &= -4\pi^2\tilde{u}(x).\end{aligned}$$

□

An other way to state the same equation is to write [9]

$$\begin{aligned}(2i\pi)^{-1}\partial_t\tilde{u}(x) &= \int_{\mathscr{M}} \partial_t(e^{2i\pi\langle x,p\rangle})\hat{u}(\overline{p})\, d\overline{p} \\ &= \int_{\mathscr{M}} 2i\pi p_0 e^{2i\pi\langle x,p\rangle}\hat{u}(\overline{p})\, d\overline{p}\end{aligned}$$

using $p_0 = (1 - |\overline{p}|^2)^{-1/2}$ on $\mathscr{M}$ and equation (61.150),

$$\begin{aligned}&= \int_{\mathscr{M}} 2i\pi e^{2i\pi\langle x,p\rangle}(1 - |\overline{p}|^2)^{1/2}\hat{u}(\overline{p})\, d\overline{p} \\ &= 2i\pi[1 - \widetilde{(2\pi)^2\Delta}]^{1/2}u(x)\end{aligned}$$

Proposition 82.29.
Let $\tilde{u}$ and $\tilde{v}$ be prolongations of the elements u and v of $H^{1/2}(\mathbb{R}^n)$. If $\tilde{v} = U(M,a)\tilde{u}$, then

$$(\mathcal{G}v)(p) = e^{-2i\pi\langle a,p\rangle}(\mathcal{G}u)(M^t). \tag{82.129}$$

9. On n'obtient pas le résultat obtenu: le tilde est beaucoup trop large.

Proof. The condition $\tilde{v} = U(L,a)\tilde{u}$ is to be taken in the sense of distributions, i.e., for each $\varphi \in \mathscr{S}(\mathbb{R}^{n+1})$,

$$\int_{\mathbb{R}^{n+1}} \tilde{v}(x)\varphi(x)\,dx = \int_{\mathbb{R}^{n+1}} \tilde{u}(x)\varphi(Mx+a).$$

It leads us to the equality

$$\begin{aligned}\iint\limits_{\mathbb{R}^{n+1}\times\mathscr{M}} & e^{2i\pi\langle x,p\rangle}(\mathcal{G}v)(p)\varphi(x)p_0^{-1}\,d\overline{p}\,dx \\ &= \iint\limits_{\mathbb{R}^{n+1}\times\mathscr{M}} e^{2i\pi\langle x,p\rangle}\mathcal{G}u(p)p_0^{-1}\varphi(Mx+a)\,d\overline{p}\,dx \\ &= \iint\limits_{\mathbb{R}^{n+1}\times\mathscr{M}} e^{2i\pi\langle x,p\rangle}e^{-2i\pi\langle a,p\rangle}\mathcal{G}u(M^{-1}p)\varphi(x)\,dx\,p_0^{-1}d\overline{p}.\end{aligned}$$

It is an equality of the form

$$\int e^{2i\pi\langle x,p\rangle}f(p)\varphi(x) = \int e^{2i\pi\langle x,p\rangle}g(p)\varphi(x)\,dx$$

that must be satisfied for all φ. It implies $f(p) = g(p)$ and the claim.

□

The correspondence $u \mapsto \tilde{u}$ allows us to identify $H^{1/2}(\mathbb{R}^n)$ to a subspace of $\mathscr{S}'(\mathbb{R}^{n+1})$. The representation U can be restricted to the subspace of $\mathscr{S}'$ made up with elements of the form $\tilde{u}$ for a $u \in H^{1/2}$. In other words, for all $u \in H^{1/2}$, there exists a $v \in H^{1/2}$ such that $U(M,a)\tilde{u} = \tilde{v}$. Such a v is obtained by formula

$$\hat{v}(\overline{p}) = \frac{(M^{-1}p)_0}{p_0}\hat{u}(M^{-1}\overline{p})e^{-2i\pi\langle a,p\rangle}.$$

How does $U(P)$ acts on $u \in H^{1/2}(\mathbb{R}^n)$? This action is defined *via* the extension $\tilde{u} \in \mathscr{S}(\mathbb{R}^{n+1})$ and the diagram

$$\begin{array}{ccc} u & \longrightarrow & \tilde{u} \\ \vdots & & \downarrow \\ v := U(P)u & \longleftarrow & U(P)\tilde{u} =: \tilde{v} \end{array}$$

Since $v(\overline{x}) = \tilde{v}(0,\overline{x})$, we have $U(P)u(\overline{x}) = U(P)\tilde{u}(0,\overline{x})$. When functions make sense, the distribution $U(P)\overline{u}$ applied to φ gives

$$(U(P)\tilde{u})\varphi = \int_{\mathbb{R}^{n+1}} \tilde{u}(x)\varphi(Px)\,dx = \int_{\mathbb{R}^{n+1}} \tilde{u}(P^{-1}y)\varphi(y)\,dy;$$

then $(U(P)\tilde{u})(x) = \tilde{u}(P^{-1}x)$. Finally,

$$U(P)u(\overline{x}) = \tilde{u}\left(P^{-1}\begin{pmatrix}0\\ \overline{x}\end{pmatrix}\right). \tag{82.130}$$

eq. Upxuxb

82.10.9 Back to operator P_ω

We consider an observer $\omega = (x;p)$ and its spatial symmetry P_ω. We know from (82.104) that $P_\omega y = x + \mathtt{proj}_T(y-x) - \mathtt{proj}_E(y-x)$ and that the linear part of P_ω is given by $\tilde{P}_\omega = J_v$ where v is the velocity associated with the momentum p. Let us search for the non linear part of P_ω. We have

$$\begin{aligned} P_\omega o &= x - \mathtt{proj}_T\, x + \mathtt{proj}_E\, x \\ &= x - J_v x. \end{aligned}$$

Finally, P_ω can be written under the form (M,a) as

$$P_\omega = (J_v, x - J_v x).$$

Proposition 82.30.

$$(\mathcal{G}\sigma_\omega u)(p') = \mathcal{G}u(S_p p')e^{2i\pi x\ \cdot\ (S_p p' - p')} \tag{82.131}$$

where S_p is the transposed matrix of J_v.

Proof. We know that $\sigma_\omega = U(P_\omega)$ acts on $u \in H^{1/2}(\mathbb{R}^n)$ with formula (82.130). From definition of $\mathcal{G}$, we have

$$\mathcal{G}(\sigma_\omega u)(p') = p_0' \widehat{\sigma_\omega u}(\overline{p}').$$

Since $\tilde{v} = U(M,a)\tilde{u}$ and more precisely $\tilde{v} = \sigma_\omega \tilde{u}$, we see that

$$\mathcal{G}(\sigma_\omega u)p' = e^{-2i\pi a\ \cdot\ p'}\mathcal{G}u(\tilde{P}_\omega^{-1}p') \tag{82.132}$$

where $(a, \tilde{P}_\omega)$ is the representation of P_ω under the form (a, M), i.e. $\tilde{P}_\omega^{-1} = J_v^{-1}$ and $a = x - J_v x$. This a fulfils

$$a \cdot p' = x \cdot p' - J_v x \cdot p' = x \cdot (p' - S_p p');$$

it allows us to rewrite (82.132) as

$$(\mathcal{G}\sigma_\omega u)(p') = e^{2i\pi x\ \cdot\ (p' - S_p p')}\mathcal{G}u(S_p p').$$

□

82.10.10 Passive symbol

Let A be a trace operator (see subsection 62.9.4) on $H^{1/2}(\mathbb{R}^n)$ its **passive symbol** in the sense of Klein-Gordon is the function $g\colon \Omega \to \mathbb{R}$,

$$g(\omega) = 2^n \operatorname{Tr}(A\sigma_\omega) \tag{82.133}$$

This is an **admissible** function: its value at $(t, \overline{x}; \overline{v})$ only depends on $(\overline{x} - t\overline{v}; v)$. In the $(x; p)$ parametrization of Ω, it means that it only depends on $(\overline{x} + x_0 p_0^{-1}\overline{p}; p)$.

Proposition 82.31.
The Klein-Gordon calculus is covariant under the action of $\mathscr{P}_0$, i.e. for all $P \in \mathscr{P}_0$ and trace operator A,

$$U(P)AU(P^{-1}) = g \circ P^{-1}$$

Proof. From now, J_p denotes J_v with the corresponding $p \leftrightarrow v$. We know that $\epsilon_T = Jp$ and that $J_v z = -z + 2\langle z, J\epsilon_T\rangle\epsilon_T$; therefore

$$J_p z = -z + 2\langle z, p\rangle Jp$$

and if $M \in \mathscr{L}_0$, we have

$$M^{-1}J_p M = J_{M^t p}. \tag{82.134}$$

Indeed we know that

$$J_{M^t p}z = -z + 2\langle z, M^t p\rangle JM^t p,$$

but $J = M^{-1}JM^{t-1}$ because $M \in \mathscr{L}$ and on the other hand,

$$\begin{aligned} M^{-1}J_p Mz &= M^{-1}\big(-Mz + 2\langle Mz, p\rangle Jp\big) \\ &= -z + 2\langle Mz, p\rangle M^{-1}Jp. \end{aligned}$$

Since $\langle ., . \rangle$ denotes the euclidian scalar product on $\mathbb{R}^{n+1}$, we have $\langle M., .\rangle = \langle ., M^t.\rangle$, then

$$M^{-1}J_p Mz = -z + 2\langle z, M^t p\rangle M^{-1}Jp.$$

Relation (82.134) follows because $JM^t p = M^{-1}Jp$ which comes from the fundamental relation $M^t J = JM^{-1}$.

Let now consider $P_\omega = (J_p, x - J_p x)$, (M, a) in $\mathscr{P}_0$ and $y \in \mathbb{R}^{n+1}$. First, one can check that

$$(M,a)^{-1} = M^{-1}\big(J_p(My+a) + x - J_p x - a\big),$$

then We have

$$\begin{aligned}(M,a)^{-1}P_\omega(M,a)y &= (M,a)^{-1}\big(J_p(My+a) + x - J_p x\big)\\ &= J_{M^t p}y + M^{-1}(x-a) + M^{-1}J_p(a-x)\\ &= J_{M^t p}y + M^{-1}(x-a) - J_{M^t p}M^{-1}(x-a).\end{aligned}$$

On the other hand, if $\omega = (x;p)$, we know that $P_\omega = (J_p, x - J_p x)$. Finally,

$$(M,a)^{-1}P_\omega(M,a)y = P_{(M^{-1}(x-a),M^t p)}y; \tag{82.135}$$

This observer, $(M^{-1}(x-a), M^t p)$, can be written under a more elegant way using formula (82.121). We find

$$(M,a)^{-1}P_\omega(M,a) = P_{(M,a)^{-1}\,\cdot\,\omega}.$$

Since $\sigma_\omega = U(P_\omega)$, we conclude that

$$U(M,a)^{-1}\sigma_\omega U(M,a) = \sigma_{(M,a)^{-1}\,\cdot\,\omega}. \tag{82.136}$$

□

Remarque 82.32.
If (M, a) don't belong to LoP_0, $(M,a)^{-1}\,\cdot\,\omega$ is not an observer because it is not linked to ω_0 by a transformation of $\mathscr{P}_0$.

82.10.11 The spacial bundle

The **spacial bundle** $\mathbb{E}$ is the following part of Ω:

$$\mathbb{E} = \{(x;p) \in \Omega \text{ tel que } \langle x, p\rangle = 0\}$$

where, we insist once again, the product $\langle x, p\rangle$ is the euclidian one on $\mathbb{R}^{n+1}$. When $p \in \mathscr{M}$ is fixed, E_p denote the hyperplane $\langle x, p\rangle = 0$. In particular,

$$E_0 = \{x \in \mathbb{R}^{n+1} \text{ tel que } x_0 = 0\} \simeq \mathbb{R}^n.$$

Proposition 82.33.
The space E_p is the spacial component of the observer whose velocity is $\overline{v} = -p_0^{-1}\overline{p}$ with respect to the reference observer ω_0.

Proof. The spacial component is given by $x \cdot \epsilon_T$ (Minkowskian product) and $\epsilon_T = Jp$ where p is related to $\overline{v}$ by $\overline{v} = -p_0^{-1}\overline{p}$. But equation $\langle x, p\rangle = 0$ is precisely $x \cdot Jp = x \cdot \epsilon_t = 0$. □

The group $\mathscr{L}_0$ acts on $\mathbb{E}$: if $x \in E_p$, then $Mx \in E_{M^{t-1}p}$.

Proposition 82.34.
The bundle $\mathbb{E}$ admits a canonical trivialization as $\mathbb{R}^3 \times \mathbb{R}^3$: an element of $\mathbb{E}$ is parametrized by an element of E_0 and a boots which is itself given by $\overline{p}$ or $\overline{v}$.

Proof. We will prove that, as x runs over E_0, $y = \Lambda_{\overline{p}}x$ runs over E_p if $\Lambda_{\overline{p}}$ is the boost associated with $\overline{p}$. Let Λ be a boost and consider p such that

$$Jp = \Lambda e_0. \tag{82.137}$$

eq_condJpEp

For each $x \in E_p$, we have

$$0 = \langle x, p\rangle = x \cdot Jp = x \cdot \Lambda e_0 = \Lambda^{-1}x \cdot e_0.$$

So $x \in E_p$ if and only if $(\Lambda^{-1}x)_0 = 0$. This shows that $E_p = \Lambda E_0$.

On the other hand, if we pick a p, then equation (82.137) gives as Λ the map $x \mapsto y$ defined by eq_transofl

$$y_0 = p_0 x_0 - \langle \overline{p}, \overline{x} \rangle \tag{82.138a}$$

$$\overline{y} = \overline{x} - x_0 \overline{p} + (1 - |\overline{p}|^2)^{-1} \langle \overline{p}, \overline{x} \rangle \overline{p}. \tag{82.138b}$$

It is easy to see that this transformation effectively gives $\Lambda e_0 = Jp = \begin{pmatrix} p_0 \\ -\overline{p} \end{pmatrix}$. To prove that this is a boost is nos as easy. First, check that $y_0^2 - \overline{y}^2 = x_0^2 - \overline{x}^2$ by developing the left hand side and use the fact that $\overline{p}^2 = (p_0 + 1)(p_0 - 1)$. So (82.138) defines a Lorentz transformation. One can check that it is selfadjoint from its explicit matricial representation. In order to check that transformation (82.138) is positive defined, one checks that for all $x \in \mathbb{R}^{n+1}$, $\langle x, Ax \rangle > 0$ where A is the transformation. For this, consider $\langle x, Ax \rangle$ as a polynomial with respect to x_0. Since $p_0 > 1$, it goes to infinity when x_0 goes to $\pm\infty$. It remains to be proved that it doesn't vanishes. For this, solve a second degree polynomial as usual using Cauchy-Schwarz $|\langle \overline{p}, \overline{x} \rangle|^2 \leqslant \overline{p}^2 \overline{x}^2$ and finds for the determinant a second degree polynomial with respect to p_0 which admits no roots with $p_0 > 1$.

From the representation (82.138) of Λ in function of p, we see that when $x_0 = 0$, y runs over E_p.

□

82.10.12 Measures on spacial bundle

From system (82.138), one easy see that

$$\frac{\partial y_j}{\partial x_k} = \delta_{jk} + (1 + p_0)^{-1} p_j p_k,$$

from which we derive

$$\begin{aligned} d\overline{y} &= d\overline{x} + (1 + p_0)^{-1} \overline{p}^2 \, d\overline{x} \\ &= p_0 \, d\overline{x} \end{aligned} \tag{82.139}$$

¡¡ Avertissement/question au lecteur !! 82.35
Dans cette dérivation ainsi que les autres mesures qui vont suivre, je ne comprends pas le raisonement d'Unterberger.

There is another way to parametrize E_p with E_0. For this, we projects E_0 on E, the spacial component of the observer ω. This projection is given by $\mathtt{proj}_E x = x - \langle x, J\epsilon_T \rangle \epsilon_T$, or

$$z = \mathtt{proj}_E x = x - \langle x, p \rangle Jp. \tag{82.140}$$ eq_xtozzeq

This gives a bijection $x \mapsto z$ between E_0 and E_p. Indeed if $x \in E_0$, then $\langle z, p \rangle = \langle x, p \rangle - \langle x, p \rangle - \langle Jp, p \rangle = 0$. This is injective because $x - \langle x, p \rangle Jp = x' - \langle x', p \rangle Jp$ implies $x - x' = \langle x - x', p \rangle Jp$. But $x, x' \in E_0$, then $x_0 = x'_0 = 0$ and $(Jp)_0 = 0$ which should implies $p_0 = 0$ which is impossible.

Let us now prove that (82.140) is surjective: for all $z \in \mathbb{R}^{n+1}$ such that $\langle z, p \rangle = 0$, there exists a $x \in E_0$ such that $z = x - \langle x, p \rangle Jp$. First, we remark that $x = z + aJp$ works for all $a \in \mathbb{R}$ as far as condition $\mathtt{proj}_E x = z$ is concerned. It is therefore easy to fix a in order to get $x_0 = 0$: it is $a = -p_0/z_0$ where we know that $z_0 \neq 0$ because (82.140) should imply $p_0 = 0$.

If $d\overline{x}$ is the measure on E_0, which is the corresponding measure $d\overline{z}$ on E_p?

¡¡ Avertissement/question à la lectrice !! 82.36
Toujours le même truc d'Unterbeger que je ne comprends pas sur sa façon de trouver les $d\overline{z}$ en fonction des $d\overline{x}$.

The answer is

$$d\overline{z} = (1 + \overline{p}^2) \, d\overline{x} = p_0^2 \, d\overline{x}. \tag{82.141}$$

The measure $p_0^{-2}\,d\overline{x}\,d\overline{p}$ is invariant under pure spacial rotations. Now equality $\Lambda(x;p)=(y;q)$ means that $\Lambda x = y$ and $\Lambda^{-1}p = q$ because momentum transform with $M^{t^{-1}}$ and here $\Lambda^t = \Lambda$ because Λ is a boost. So we look at a boost $\Lambda_{\overline{p'}}$ related to the momentum p' and we will prove that $p_0^{-2}\,d\overline{x}\,d\overline{p}$ is invariant under this boost. In other words, we want $q_0^{-2}\,d\overline{y}\,d\overline{q} = p_0^{-2}\,d\overline{x}\,d\overline{p}$.

By lemma 82.27 and system (82.138) , we find $q_0 = (\Lambda^{-1})_0 = p'_0p_0 + \langle\overline{p}',\overline{p}\rangle$. On the other hand, $x \in E_p$, so $x_0 = -p_0^{-1}\langle\overline{x},\overline{p}\rangle$ and formula

$$\overline{y} = \overline{x} - x_0\overline{p}' + (1+p'_0)^{-1}\langle\overline{p}',\overline{x}\rangle\overline{p}'$$

gives $d\overline{y}/d\overline{x} = p_0^{-1}$ which proves that $d\overline{y} = p_0^{-1}q_0\,d\overline{x}$. But we know from a long time that measure $p_0^{-1}\,d\overline{p}$ is invariant, then $d\overline{q} = q_0p_0^{-1}\,d\overline{p}$.

Lemma 82.37.
The function $f(x;p)$ defined on Ω depends only on $(\overline{x}+x_0p_0^{-1}\overline{p};p)$ if and only if fulfils the differential equation

$$p_0\frac{\partial f}{\partial x_0}f = \sum_j \frac{\partial f}{\partial x_j}. \tag{82.142}$$

Proof. When one computes $\partial_0 f$, one remark that $f(x_0+t,\overline{x};p)$ only contains x_0 under the form $\overline{x}+(x_0+t)p_0^{-1}\overline{p} = \overline{x}+x_0p_0^{-1}\overline{p}+tp_0^{-1}\overline{p}$, then

$$f(x_0+t,\overline{x};p) = f(x_0,\overline{x}+tp_0^{-1}\overline{p};p).$$

This leads to the conclusion. □

82.10.13 Active symbol

Let f be an admissible function on Ω which fulfils

$$\int_{\mathbb{E}} |f(x;p)|p_0^{-2}\,d\overline{x}\,\overline{p} < \infty.$$

The operator $\mathrm{Op}(f)$ of **active symbol** f is the operator on $H^{1/2}(\mathbb{R}^n)$ defined by

$$\mathrm{Op}(f) = 2^n\int_{\mathbb{E}} f(x;p)\sigma_{(x;p)}p_0^{-2}\,d\overline{x}\,d\overline{p}. \tag{82.143}$$

One can prove that it is a bounded operator [10].

prop_symbadj

Proposition 82.38.
The active or passive symbol of the adjoint of an operator is the complex conjugated of the corresponding symbol.

We begin by proving a lemma:

Lemma 82.39.
The operator σ_ω is involutive.

Proof of the lemma. Formula (82.131) with σ_ω^2 gives

$$\begin{aligned}(\mathcal{G}\sigma_\omega^2u)(q) &= (\mathcal{G}\sigma_\omega u)(S_pq)e^{2i\pi x\,\cdot\,(S_pq-q)}\\ &= \mathcal{G}u(q),\end{aligned}$$

taking into account $S_p^2 = \mathbb{1}$. □

Remark that an unitary involutive operator is selfadjoint.

10. Saut que moi, je n'y parviens pas.

Proof of proposition 82.38 . Let A be a trace operator; its passive symbol is the function $g\colon \Omega \to \mathbb{C}$,

$$g_A(\omega) = 2^n \operatorname{Tr}(A\sigma_\omega),$$

therefore

$$\begin{aligned}\overline{g_{A^*}(\omega)} &= 2^n \overline{\operatorname{Tr}(\overline{A}\sigma_\omega^t)} \\ &= 2^n \operatorname{Tr}(A\sigma_\omega^*) \\ &= 2^n \operatorname{Tr}(A\sigma_\omega) \\ &= g_A(\omega).\end{aligned}$$

We now turn our attention to ative symbol. We prove that

$$\big(\operatorname{Op}(f)u, v\big)_{1/2} = \big(u, \operatorname{Op}(\overline{f})v\big)_{1/2}.$$

by the following computation:

$$\begin{aligned}\big(\operatorname{Op}(f)u, v\big)_{1/2} &= \int_{\mathscr{M}} \xi_0 \widehat{\operatorname{Op}(f)u}(\overline{\xi})\overline{\hat{v}(\overline{\xi})}\, d\overline{\xi} \\ &= 2^n \int_{\mathscr{M}} \int_{\mathbb{R}^n} \xi_0 \int_{\mathbb{E}} e^{-2i\pi\langle\overline{\xi},\overline{y}\rangle} f(x;p)\sigma_{(x;p)}u(\overline{y})\overline{\hat{v}(\overline{\xi})} p_0^{-1} \\ &\qquad d\overline{x}\, d\overline{y}\, d\overline{\xi}\, d\overline{p},\end{aligned}$$

but integral over $\overline{y}$ leads to the Fourier transform of $\sigma_{(x;p)}u$ at $\overline{\xi}$, so

$$\begin{aligned}&= 2^n \int_{\mathbb{E}} \big(\sigma_{(x;p)}u, v\big)_{1/2} f(x;p) p_0^{-1}\, d\overline{x}\, d\overline{p} \\ &= 2^n \int_{\mathbb{E}} \big(u, \overline{f}(x;p)\sigma_{(x;p)}v\big)_{1/2} p_0^{-1}\, d\overline{x}\, d\overline{p} \\ &= \big(u, \operatorname{Op}(\overline{f})v\big)_{1/2}.\end{aligned}$$

□

82.10.14 The operator ∇^λ

Let f be a continuous function on Ω which grows slowly with respect to the variable x. We denote by $F_1 f$ the Fourier transform of f with respect to this variable and we suppose that the support of $(F_1 f)(\xi, p)$ is space-like: for all $\xi \in \operatorname{Supp} F_1 f$, we have $\overline{\xi}^2 - \xi_0^2 \geqslant 0$. In the same way that we defined $(1+\Delta)^s$ in section 61.8, we define

$$\nabla^s = \big(1 + (4\pi)^{-2}\Box\big)^{s/2}$$

by formula

$$F_1(\nabla^s f)(\xi, p) = \big(1 + \frac{1}{4}(\overline{\xi}^2 - \xi_0^2)\big)^{s/2}(F_1 f)(\xi, p). \tag{82.144}$$

When $s = 2$, we find back the d'Alembert operator $\Box = \partial_t^2 - \sum_j \partial_j^2$. If f is admissible, it fulfils equation $\langle Jp, \partial_{\overline{x}} f\rangle = 0$ where we defined $(\partial f/\partial_{\overline{x}})_i = \partial f/\partial x_i$. Let us prove that $\langle Jp, \xi\rangle = 0$ on the support of $F_1 f$. Standard result (61.146) leads to

$$\langle Jp, \partial_{\overline{x}} f\rangle = 2i\pi F_1^{-1}\big(\langle Jp, \xi\rangle F_1 f\big)(x) = 0,$$

which imposes $\langle Jp, \xi\rangle = 0$ on the support of $F_1 f$. So $p_0\xi_0 = \langle \overline{p}, \overline{\xi}\rangle \leqslant |\overline{p}||\overline{\xi}|$, and thus on the support of $F_1 f$, we have $\xi_0^2 \leqslant p_0^{-2}\overline{p}^2\overline{\xi}^2$, but $1 - p_0^{-2}\overline{p}^2 = p_0^{-2}(p_0^2 - \overline{p}^2) = p_0^{-2}$ and finally,

$$\overline{\xi}^2 - \xi_0 \geqslant p_0^{-2}\overline{\xi}^2 \geqslant \frac{1}{2}p_0^{-2}\overline{\xi}^2.$$

Chapter 83

WKB quantization

Abstract

Deformation is a main theme of research in the present work. We begin here to describe WKB quantization and a general method to guess deformations of function algebras. The role of Darboux charts and momentum maps appears clearly. A careful example is given by the deformation of $\mathrm{SL}(2,\mathbb{R})$.

We prove a useful result (from [1020]), the extension lemma, which allows to deform a split extension when one knows a deformation of the two components of the extension. The kernel is simply the product of the two kernels.

Then we see the principle of deformation by action of group: when a Lie group is deformable, one can find a deformation of any manifold on which the group acts. Universal formulas exist in some cases. This is why deformations of groups are studied. An application of that extension lemma to the Iwasawa subgroup of $\mathrm{SO}(2,n)$ is given in chapter 85.

83.1 WKB quantization

subsec:WKB

More details can be found in the article [1038]. A manifold M is given with its usual commutative and associative algebra $(C^\infty(M),\ \cdot\ )$ of smooth functions. A **deformation**, or a *quantization* [1], of M is the data of a new product $\star_\hbar^M$ on a functional space over M.

Let G be a Lie group acting on a manifold M. We consider $\mathrm{Fun}(M,\mathbb{C})$, the space of all the maps from M to $\mathbb{C}$, without any regularity conditions. The **regular left representation** of G on M is the representation of G on $\mathrm{Fun}(M)$ given by

$$[L_g^*(a)](h) = a(gh) \tag{83.1}$$

for all $a \in \mathrm{Fun}(M)$, g, $h \in G$.

A G-invariant WKB quantization of M is a product on a space of functions A^M on M of the form

$$(u \star_\hbar^M v)(x) = \int_{M\times M} a_\hbar(x_1,x_2,x) e^{\frac{i}{\hbar}S(x_1,x_2,x)} u(x_1)v(x_2)\,dx_1\,dx_2$$

for which we require, among other conditions, (see complete definition 83.2)

— $A^M \subset \mathrm{Fun}(M)$ is invariant under the regular left representation of G and contains at least the smooth compactly supported functions,

— the pair $(A^M, \star_\hbar^M)$ is an associative algebra,

— the functions $a_\hbar$ and S are invariant under the left regular action of G,

1. In fact, we make a difference between these two words. A *deformation* is only the fact to find a new product from an old one; the new product depends on a parameter and has to reduce to the old one when the parameter goes to zero. A *quantization* is a deformation in which the first order term (whatever it means) of the new product contains the symplectic structure as in condition (83.4) below.

— $\forall\, x \in M$ and $\forall\, u, v \in A^M$ the product accepts an asymptotic expansion compatible with the symplectic structure in the following sense:

$$(u *_\hbar v)(x) \sim u(x)v(x) + \frac{\hbar}{i} c_1(u,v)(x) + o(\hbar^2)$$

where c_1 satisfies $c_1(u,v) - c_1(v,u) = 2\{u,v\}$.

The main property of this product is its G-invariance:

$$L_g(u \star_\hbar^M v) = (L_g u) \star_\hbar^M (L_g v).$$

83.1.1 Definitions and general setting

Let (M, ω, ∇) be an affine symplectic manifold, i.e. a $2n$-dimensional symplectic manifold (M,ω) endowed with a torsion-free connection ∇ such that $\nabla\omega = 0$. The **automorphism** group $\mathrm{Aut}(M,\omega,\nabla)$ is defined as

$$\mathrm{Aut}(M,\omega,\nabla) = \mathrm{Aff}(\nabla) \cap \mathrm{Symp}(\omega)$$

where $\mathrm{Aff}(\nabla)$ is the group of affine transformations of the affine manifold (M,∇) and $\mathrm{Symp}(\omega)$ is the group of symplectomorphisms of (M,ω).

¡¡ Avertissement/question au lecteur !! 83.1
Non mais; où intervient ∇ dans cette définition ? D'après Pierre, il est contenu dans les troisième ordre, mais il faudrait une référence. ProbNablades

Let R be a subgroup of $\mathrm{Aut}(M,\omega,\nabla)$. The following definition of a R-invariant WKB quantization can be found in [961].

Definition 83.2.
*A R-invariant **WKB quantization** of (M,ω,∇) is the data of a product*

$$(u \star_\theta v)(x) = \frac{1}{\theta^{2n}} \int_{M\times M} a_\theta(x,y,z) e^{\frac{i}{\theta} S(x,y,z)} u(y) v(z)\, dy\, dz \tag{83.2}$$ EqFormeWKBdel

(where $dy\,dz$ is the Liouville measure $\omega^n/n!$) with the following constrains:

(1) For each θ, we have a space A_θ containing the space $C_c^\infty(M)$ of compactly supported smooth functions. The product $\star_\theta$ extends to A_θ in such a way that $(A_\theta, \star_\theta)$ becomes a one-parameter family of associative $*$-algebras.

(2) The product $*_0$ on A_0 is the usual pointwise product and $(A_0, *_0)$ is a Poisson subalgebra of $C^\infty(M)$ for the induced Poisson structure from the symplectic form ω.

(3) $\forall \theta \geqslant 0$, the space A_θ is a $*$-vector subspace of $C^\infty(M)$ such that

$$C_c^\infty(M) \subset A_0 \subset A_\theta$$

where the involution $*$ on $C^\infty(M)$ is the usual complex conjugation.

(4) S is a real valued smooth function $S\colon M\times M\times M \to \mathbb{R}$ such that for all $x_0 \in M$, the function $S(x_0,.,.) \in C^\infty(M\times M)$ has a nondegenerate critical point at (x_0,x_0).

(5) The functions a_θ are positive real-valued:

$$a_\theta\colon M\times M\times M \to \mathbb{R}^+.$$

(6) The functions S and a_θ are invariant under the diagonal action of R on $M\times M\times M$.

(7) $\forall\, x \in M$ and $\forall\, u, v \in C_c^\infty(M)$ with support in a suitably small neighbourhood of x, a stationary phase method yields the extension

$$(u \star_\theta v)(x) \sim u(x)v(x) + \frac{\theta}{i} c_1(u,v)(x) + o(\theta^2) \tag{83.3}$$

where c_1 satisfies

$$c_1(u,v) - c_1(v,u) = 2\{u,v\}. \tag{83.4}$$ EqExigSympleFormOrd

DefWKBCompl

We emphasize the fact that the functional space A^M is stable under $\star_\theta$: this is a *strict* quantization in contrast to a *formal* star product which only stabilises the space of formal power series of θ.

An example of WKB quantization is the Weyl product which is nothing but an integral reformulation of the Moyal star product:

$$(f \star_\hbar^W g)(x) = \frac{1}{\hbar^{2n}} \int_{\mathbb{R}^{2n}\times\mathbb{R}^{2n}} e^{\frac{2i}{\hbar}S^0(x,y,z)} f(y)g(z)\,dy\,dz$$

where $S^0(x,y,z) = \Omega(x,y) + \Omega(y,z) + \Omega(z,x)$, and Ω denotes the usual symplectic form on $\mathbb{R}^{2n}$.

¡¡ Avertissement/question à la lectrice !! 83.3
Ce serait bien d'avoir une référence pour cette affirmation. Et aussi de savoir s'il faut un $1/\hbar^{2n}$ devant l'intégrale. ProbWeylMoy

The function $K = a_\theta e^{\frac{i}{\theta}S}$ is the **kernel** of the product $\star_\theta$. The **associativity** of $\star_\theta$ on the functional space A_θ is the fact that the equality

$$\big((u \star_\theta v) \star_\theta r\big)(x) = \big(u \star_\theta (v \star_\theta r)\big)(x)$$

holds for every u, v, $r \in A_\theta$ and $x \in M$. That condition translates under an integral form to the following relation

$$\begin{aligned}&\int_{M\times M} K(x,y,z)\left[\int_{M\times M} K(y,t,s)u(t)v(s)\mu_M(t,s)\right] r(z)\mu_M(y,z)\\ &= \int_{M\times M} K(x,y,z)u(y)\left[\int_{M\times M} K(z,t,s)v(t)r(s)\mu_M(t,s)\right]\mu_M(y,z)\end{aligned} \tag{83.5}$$

EqCondAssoc

where $\mu_M(y,z) = \mu_M(y)\mu_M(z)$ is the Liouville measure on M. Performing formal manipulations (such as a Fubini theorem), one can express this condition as

$$\int_M K(x,y,t)K(t,p,q)\mu(t) = \int_M K(x,\tau,q)K(\tau,y,p)\,\mu(\tau). \tag{83.6}$$

EqAssosssens

That form is easier to handle and to check, but it is meaningless in general.

The fact to have a **left invariant kernel** on a group G means that the kernel $K\colon G\times G\times G \to \mathbb{C}$ has the property $L_g^* K = K$, or

$$K(gh_1, gh_2, gh_3) = K(h_1,h_2,h_3) \tag{83.7}$$

for every $g \in G$. The following lemma allows us to use group isomorphisms to push forward a kernel from a group to another.

Lemma 83.4.
Let G_1 and G_2 be two symplectic Lie groups and K_1, a left invariant kernel on G_1 which provides an associative product on the functional space A_1. Let $\phi\colon G_2 \to G_1$ be a symplectic Lie group isomorphism. Then the kernel $K_2 = \phi^* K_1$ is invariant and gives rise to an associative product on $A_2 = \phi^* A_1$. LemKerINvarIsom

Proof. By definition,

$$(\phi^* K_1)(h_1,h_2,h_3) = K_1\big(\phi(h_1),\phi(h_2),\phi(h_3)\big).$$

Therefore, using the left invariance of K_1, we have

$$L_g^*\phi^* K_2 = (\phi\circ L_g)^* K_2 = (L_g\circ\phi)^* K_2 = \phi^* L_{\phi(g)}^* K_1 = \phi^* K_1.$$

That proves left invariance of ϕ^*K_1 on G_2. Now we prove the associativity of K_2, this is to check condition (83.5). We have

$$\begin{aligned}&\int_{G_2\times G_2} K_2(x,y,z)\Big[\int_{G_2\times G_2} K_2(y,t,s)(\phi^*u)(t)(\phi^*v)(s)\mu_2(t,s)\Big]\\&\qquad(\phi^*r)(z)\mu_2(y,z)\\&=\int_{G_2\times G_2} K_1(\phi x,\phi y,\phi z)\Big[\int_{G_2\times G_2} K_1(\phi y,\phi t,\phi s)u(\phi t)v(\phi s)\mu_2(t,s)\Big]\\&\qquad r(\phi z)\mu_2(y,z).\end{aligned}$$

We perform in this integral the change of variables $\tau_y=\phi y$, $\tau_t=\phi t$, $\tau_z=\phi z$ and $\tau_s=\phi s$. This does not affect the measure because ϕ is a symplectomorphism and μ_i are the Liouville measures on G_i, so that for example, $\mu_2(t)=\mu_2(\phi^{-1}\tau_t)=\mu_1(\tau_t)$. The previous integral becomes

$$\begin{aligned}&\int_{G_1\times G_1} K_1(\phi x,\tau_y,\tau_z)\Big[\int_{G_1\times G_1} K_1(\tau_y,\tau_t,\tau_s)u(\tau_t)v(\tau_s)\mu_1(\tau_t,\tau_s)\Big]\\&\qquad r(\tau_z)\mu_1(\tau_y,\tau_z).\end{aligned}$$

Using now the associativity of K_1 on G_1 and performing the inverse change of variables, we find

$$\begin{aligned}&\int_{G_2\times G_2} K_2(x,y,z)(\phi^*u)(y)\Big[\int_{G_2\times G_2} K_2(z,t,s)(\phi^*v)(t)(\phi^*r)(s)\mu_2(t,s)\Big]\\&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\mu_2(y,z),\end{aligned}$$

which proves the associativity of K_2 on ϕ^*A_1.

Notice that condition (83.6) can be checked in much the same way.

□

The proposition 88.8 gives an improved form of this lemma. Unfortunately, this generalization revealed to be a mistake. It is worth noticing that lemma 83.4 needs a group isomorphism while one often only has a Lie algebra isomorphism. Due to Campbell-Backer-Hausdorff formula, it may be very difficult to find a group isomorphism from an algebra one. An example of this difficulty is in subsection 83.5.6.

Remarque 83.5.
Most of the time, the symplectic condition (83.4) does not have to be checked because we just define the symplectic form ω_2 on G_2 as $\omega_2=\phi^*\omega_1$ where ω_1 is the symplectic form on G_1.

Definition 83.6.
*When $\alpha\colon G\times A\to A$ is an action of a Lie group G on a vector space A, one says that the element $a\in A$ is a **differentiable vector** of α if the map $g\mapsto\alpha_g(a)$ is a differentiable map from G into A.*

We are now interested in the regular left representation $L\colon R\times A_\theta\to A_\theta$ defined as usual by $\big(L_r(u)\big)(x)=u(r\cdot x)$. A function $u\in A_\theta$ is a differentiable vector of L when the map

$$\begin{aligned}\alpha_u\colon R&\to A_\theta\\ r&\mapsto L_r(u)\end{aligned}\tag{83.8}$$

is differentiable. The differential of α_u is what we will denote by dL in the next few pages: $dL(X)u=(d\alpha_u)_eX$. By definition,

$$(d\alpha_u)_eX=\frac{d}{dt}\Big[\alpha_u(e^{tX})\Big]_{t=0}=\frac{d}{dt}\Big[L_{e^{tX}}(u)\Big]_{t=0},$$

and the element $(d\alpha_u)_e X \in A_\theta$ applied to $x \in M$ is

$$\big(dL(X)u\big)(x) = \Big((d\alpha_u)X\Big)(x) = \frac{d}{dt}\Big[L_{e^{tX}}(u)x\Big]_{t=0} = \frac{d}{dt}\Big[u(e^{tX}\cdot x)\Big]_{t=0}. \tag{83.9}$$

We denote by A_θ^∞ the space of differentiable vectors of the representation L.

If one particularises to $A_\theta \subset C^\infty(R)$ (the manifold M being R itself), the vector fields of R naturally act on A_θ. In particular, if $u\colon R \to \mathbb{C}$ and $X \in \mathcal{R}$ we have

$$\big(X^*(u)\big)(r) = X_r^*(u) = \frac{d}{dt}\Big[u\big(e^{-tX}r\big)\Big]_{t=0} = \big(dL(-X)u\big)(r),$$

so that

$$dL(X) = -X^* \tag{83.10}$$

holds on the space of differentiable vectors A_θ^∞.

Definition 83.7.
*A formal star product $*_G\colon C^\infty(M)[[\nu]] \times C^\infty(M)[[\nu]] \to C^\infty(M)[[\nu]]$ is said to be **$\mathcal{G}$-covariant** if for all X, $Y \in \mathcal{G}$,*

$$[\lambda_X, \lambda_Y]_{*_G} = 2\nu\{\lambda_X, \lambda_Y\} \tag{83.11}$$

where $[\lambda_X, \lambda_Y]_{*_G} := \lambda_X *_G \lambda_Y - \lambda_Y *_G \lambda_X$. In other words the start product is $\mathcal{G}$-covariant when the expected terms of higher order in the right hand side are zero.

A crucial use of $\mathcal{G}$-covariance will be done in proposition 83.9 in order to build a map ρ_ν that fulfils the following proposition (instead of dL itself).

Proposition 83.8.
In the setting of definition 83.2, the map dL is a representation by derivation of $\mathcal{R}$ on A_θ^∞. prop:dL_reprez

Proof. We will not pay attention on the domain A_θ. Its definition will come later. First, we prove that $dL\colon \mathcal{R} \to \operatorname{End} A_\theta^\infty$ is a representation. Indeed,

$$\begin{aligned} dL([X,Y])u = \frac{d}{dt}\Big(L^*_{\exp(-t[X,Y])}u\Big)_{t=0} &= \frac{d}{dt}\Big([L^*_{\exp(-tX)}, L^*_{\exp(-tY)}]u\Big)_{t=0} \\ &= \big[dL(X), dL(Y)\big]u. \end{aligned} \tag{83.12}$$

Next, L_R-invariance of $*_\theta$ yields

$$\big(L^*_{\exp -tX}u\big) *_\theta \big(L^*_{\exp -tX}v\big) = L^*_{\exp -tX}(u *_\theta v).$$

If we derive this equality with respect to t at $t = 0$, we find

$$dL(X)u *_\theta v + u *_\theta dL(X)v = dL(X)(u *_\theta v).$$

□

83.1.2 Deformation of Iwasawa subgroups

The motivation in deforming (or quantizing) groups resides in the method of deformation by group action (section 82.6) which states that if one can deform a group, one can write a formula for a deformed product on any manifold on which the group acts.

Let first describe the next few steps in the construction of WKB quantizations of groups. Let G be a semisimple Lie group with its Iwasawa decomposition $G = ANK$. The group $R = AN$ is solvable and can be seen as the homogeneous space $R = G/K$. We consider the canonical multiplicative action $\tau\colon G \times R \to R$ which we restrict to $\tau\colon R \times R \to R$. We are interested in a R-invariant quantization of R. Here is a summary of the notations that will be used.

— $*_M$ is the Moyal star product on $\mathbb{R}^n$ endowed with its canonical symplectic form,

— $\star_\theta^R$ is the product we are searching for. It has to be defined at least on $C_c^\infty(R)$ and should be extended to $C^\infty(R)$,

— $A^R \subset \mathrm{Fun}(R,\mathbb{C})$ must contain $C_c^\infty(R)$. The purpose is $(A^R, \star_\theta^R)$ to be an associative algebra and A^R to be invariant under the left regular representation of R,

— $\mathbb{A}_\nu = C^\infty(R)[[\nu]]$ is an intermediary space which serves to guess $\star_\theta^R$ and perform formal manipulations with ρ_ν and dL,

— $*_M^R$ is the pull-back of Moyal to $\mathbb{A}_\nu$. It serves to formal manipulations in order to guess the twist that defines $*_\nu^R$,

— $*_\nu^R$ is the product on $\mathbb{A}_\nu$. The problem of determining that product is formal. When this problem is solved, we have to prove that in a well chosen A^R, taking $*_\nu^R \to \star_\theta^R$ yields a solution to the problem. As previously noticed, in order to make sense, one has to apply dL on the subspace $\mathbb{A}_\nu^\infty$ of differentiable vector of the regular left representation. We will however not take care of this issue in the formal manipulations.

The main steps are the following:

(1) In the case of a WKB product we saw in proposition 83.8 that dL is a representation of $\mathcal{R}$ on $\mathbb{A}_\nu^\infty$. Hence we will try to build a formal product for which dL is a representation by derivation. From this point of view, the manipulation with ρ_ν is only a trick designed to guess a product formula.

(2) We suppose that the group R —the one that we are trying to quantize— has a symplectic structure ω and we consider $\phi\colon \mathbb{R}^{2n} \to R$, a Darboux chart; i.e. $\omega = \phi^*\Omega$ where Ω is the canonic symplectic form on $\mathbb{R}^{2n}$.

(3) We suppose that the left action of R on itself is strongly hamiltonian and we denote by λ_X the momentum maps. We suppose that the Moyal product is $\mathcal{G}$-covariant [2].

(4) We pose $\rho_\nu(X) = \frac{1}{2\nu}\,\mathrm{ad}_{*_M^R}(\lambda_X)$. The $\mathcal{R}$-covariance of $*_M^R$ is used in order to prove that ρ_ν is a representation by derivations of $\mathcal{R}$ on $(\mathbb{A}_\nu, *_M^R)$.

(5) If one can find an intertwining operator between dL and ρ_ν (i.e. if they are equivalent representations), we define $*_\nu^R$ as the pull-back of $*_M^R$ by this intertwining operator. In this case, we prove that dL is a representation by derivations of the product $*_\nu^R$.

It is time to read appendix 82.1 about the Moyal star-product.

We try now to find a formal product $*_\nu^R$ on $\mathbb{A}_\nu^\infty$ such that dL is a representation by derivations. For this purpose we suppose R to accept a symplectic structure ω and $\phi\colon \mathbb{R}^{2n} \to R$ to be a Darboux chart, i.e. $\omega = \phi^*\Omega$ where Ω denotes the canonical symplectic form on $\mathbb{R}^{2n}$. Then we bring the Moyal product of $\mathbb{R}^{2n}$ to R by the usual formula

$$(u *_M^R v) = (u\circ\phi *_M v\circ\phi)\circ\phi^{-1}. \tag{83.13}$$

We suppose that product to be $\mathcal{G}$-covariant [3]:

$$[\lambda_X,\lambda_Y]_{*_M^R} = 2\nu\{\lambda_X,\lambda_Y\}_R. \tag{83.14}$$

Now we consider the left action of R on itself and we suppose that this is an Hamiltonian action for the symplectic structure $\omega = \phi^*\Omega$ with dual momentum maps $\lambda_X\colon R \to \mathbb{C}$. We define, for each $X \in \mathcal{R}$, a linear map, $\rho_\nu(X)\colon \mathbb{A}_\nu \to \mathbb{A}_\nu$ by

$$\begin{aligned} \rho_\nu\colon \mathcal{R} &\to \mathrm{End}\,\mathbb{A}_\nu \\ X &\mapsto \frac{1}{2\nu}\,\mathrm{ad}_{*_M^R}(\lambda_X) \end{aligned} \tag{83.15}$$

Notice that the formal series of $[\lambda_X, u]_{*_M^R}$ begins with order one, so the division by ν make sense in the space of formal series. The main interest of ρ_ν is to be as we want dL to be. So it will be used to guess how to twist the product in order to make dL work as ρ_ν.

2. In fact, we only need the $\mathcal{R}$-covariance.
3. Only the $\mathcal{R}$-covariance will be actually used.

Proposition 83.9.
The map ρ_ν is a representation of $\mathcal{R}$ on $\mathbb{A}_\nu$, and $\rho_\nu(X)$ is a derivation of $(\mathbb{A}_\nu, *_M^R)$ for each $X \in \mathcal{R}$. Proprhonureprez

Proof. The proof that ρ_ν is a representation is only to check that the relation $[\rho_\nu(X), \rho_\nu(Y)]f = \rho_\nu([X,Y])f$ holds for any X, $Y \in \mathcal{R}$ and $f \in \mathbb{A}_\nu$. Using the $\mathcal{G}$-covariance and the Jacobi identity,

$$\begin{aligned}\rho_\nu([X,Y])f &= \frac{1}{4\nu^2}\operatorname{ad}_{*_M^R}(2\nu\lambda_{[X,Y]})f = \frac{1}{4\nu^2}\operatorname{ad}_{*_M^R}([\lambda_X,\lambda_Y]_{*_M^R})f \\ &= \frac{1}{4\nu^2}[[\lambda_X,\lambda_Y]_{*_M^R}, f]_{*_M^R} \\ &= \frac{1}{4\nu^2}(\operatorname{ad}_{*_M^R}\lambda_X \circ \operatorname{ad}_{*_M^R}\lambda_Y - \operatorname{ad}_{*_M^R}\lambda_Y \circ \operatorname{ad}_{*_M^R}\lambda_X)f \\ &= [\rho_\nu(X), \rho_\nu(Y)]f.\end{aligned} \tag{83.16}$$

It remains to check that $\rho_\nu(X)(u *_M^R v) = \rho_\nu(X)u *_M^R v + u *_M^R \rho_\nu(X)v$ for every $X \in \mathcal{R}$. This is once again just a computation.

$$\begin{aligned}\rho_\nu(X)u *_M^R v + u *_M^R \rho_\nu(X)v &= \frac{1}{2\nu}(\lambda_X *_M^R u - u *_M^R \lambda_X) *_M^R v \\ &\quad + \frac{1}{2\nu}u *_M^R (\lambda_X *_M^R v - v *_M^R \lambda_X) \\ &= \frac{1}{2\nu}\operatorname{ad}_{*_M^R}\lambda_X(u *_M^R v).\end{aligned} \tag{83.17}$$

□

Notice that the $\mathcal{G}$-covariance of $*_M^R$ was used to prove that ρ_ν is a representation. Now, if we could show that $\rho_\nu = dL$, then the answer to our deformation problem would be $A_\theta = \mathbb{A}_\nu^\infty$ and $\star_\theta = *_M^R$. But instead of that we have $\rho_\nu = dL + o(\nu)$ because

$$\begin{aligned}\rho_\nu(X)u &= \frac{1}{2\nu}[\lambda_X, u]_{*_M^R} = \frac{1}{2\nu}2\nu\{\lambda_X, u\} + o(\nu) = X^*(u) + o(\nu) \\ &= -dL(X)u + o(\nu)\end{aligned} \tag{83.18}$$

where the notion of fundamental field X^* is taken for the regular left representation (which is Hamiltonian). That shows that ρ_ν is something like a deformation of dL. As a consequence, one has $dL(X) = X_{\lambda_X}$, or

$$dL(x)u = X_{\lambda_X}(u) = \{\lambda_X, u\} \tag{83.19}$$ eq:dL_et_Poisson

(see subsection 59.1.3).

Since ρ_ν is not dL, the hope is to see if ρ_ν and dL should be *equivalent* representations. As next proposition shows, the fact to find an equivalence between ρ_ν and dL actually solves the problem to find a product for which dL is a representation by derivation.

Proposition 83.10.
Let $\mathcal{T} : \mathbb{A}_\nu \to \mathbb{A}_\nu$ be an intertwining operator between dL and ρ_ν:

$$\mathcal{T}\rho_\nu(X)\mathcal{T}^{-1} = dL(X). \tag{83.20}$$ eq:TrnT

If we define the star product $*_\nu^R$ by

$$u *_\nu^R v = \mathcal{T}_\nu(\mathcal{T}_\nu^{-1}u *_M^R \mathcal{T}_\nu^{-1}v), \tag{83.21}$$ Eq_candprodANSL

dL becomes a derivation of $*_\nu^R$. prop:def_stn

Proof. If we develop the expression of $dL(X)(u *_M^R v)$, we find $\mathcal{T}\rho_\nu(X)(\mathcal{T}^{-1}u *_M^R \mathcal{T}^{-1}v)$, using the fact that ρ_ν is a derivation of $*_M^R$, one easily finds $dL(X)u *_M^R v + u *_\nu^R dL(X)v$. □

83.2 Deformation of $\mathrm{SL}(2,\mathbb{R})$

sec:unifsl

Abstract

This section shows in some detail an instructive example of deformation of an Iwasawa subgroup: the Iwasawa subgroup of $\mathrm{SL}(2,\mathbb{R})$. In this section we will use the parametrization (53.321) of $\mathrm{SL}(2,\mathbb{R})$, as well as the notations $G = \mathrm{SL}(2,\mathbb{R})$ and $\mathcal{G} = \mathfrak{sl}(2,\mathbb{R})$. Here are the main steps that will be performed:

(1) The Iwasawa component $R = AN = G/K$ provides a double covering onto $\mathcal{O} = \mathrm{Ad}(G)Z$ where Z is any element of $\mathcal{K}$ (which is one dimensional). The adjoint orbit $\mathcal{O}$ being endowed with a canonical symplectic form described in subsection 59.1.4, we consider on R the corresponding symplectic structure.

(2) The map $(a,l) \mapsto \mathrm{Ad}(e^{aH}e^{lE})Z$ turns out to be a global Darboux chart and induces the diffeomorphism
$$R \simeq \mathcal{O} \simeq \mathbb{R}^2.$$
Under these identifications, the adjoint action of R on $\mathcal{O}$ becomes the simple multiplication of R in itself, which is strongly hamiltonian.

(3) The Moyal product is $\mathfrak{sl}(2,\mathbb{R})$-covariant for the action of $\mathrm{SL}(2,\mathbb{R})$ on $\mathbb{R}^2$.

(4) We explicitly build the intertwining operator between ρ_ν and dL and we write down a product (see proposition 83.10).

(5) A theorem is stated in which we list the properties of the so constructed product.

83.2.1 Actions and Symplectic structure

For our purpose, we consider the Ad^*-invariant 2-form $\xi_0 \in \mathcal{G}^*$ and $\widetilde{\mathcal{O}} = \mathrm{Ad}^*(G)\xi_0$. As seen in 59.1.4, the orbit $\widetilde{\mathcal{O}}$ is a symplectic manifold with

$$\widetilde{\omega}_\xi(X^*, Y^*) = \langle \xi, [X,Y] \rangle \tag{83.22}$$ EqAStrucSymmCoAdj

and the dual momentum maps are $\lambda_X(\xi) = \langle \xi, X \rangle$.

In the present framework, we can work with adjoint orbits instead of the coadjoint ones because the group $G = \mathrm{SL}(2,R)$ is semisimple. Indeed, in this case, the Killing form $B\colon \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ gives a $\mathrm{Ad}(G)$-equivariant isomorphism between $\mathcal{G}$ and $\mathcal{G}^*$. In order to see that, recall that a basic property of the Killing form is

$$B\big((\mathrm{ad}\,X)Y, Z\big) = -B\big(Y, (\mathrm{ad}\,X)Z\big), \tag{83.23}$$ eq:B_ad_invar

and when the group is semisimple, B is nondegenerate. The isomorphism is given by $B'\colon \mathcal{G} \to \mathcal{G}^*$, $B'(X)Y = B(X,Y)$. The fact that B is nondegenerate makes B' an isomorphism, and the property (83.23) gives the $\mathrm{ad}(G)$-equivariance of B':

$$B'\big((\mathrm{ad}\,X)Y\big)Z = -B'(Y)\big((\mathrm{ad}\,X)Z\big).$$

Here, in contrast with the case studied in 59.1.4, we are working with adjoint orbits (and not the coadjoint orbits), so the subalgebra to be studied is no more $\widetilde{\mathcal{O}}$ but

$$\mathcal{O} = \mathrm{Ad}(G)Z,$$

where Z is the generator of $\mathcal{K}$ and the symplectic form is not exactly (59.14), but

$$\omega_X(A^*, B^*) = B(X, [A,B]). \tag{83.24}$$ eq:omega_G

The action of G on $\mathcal{O}$ is $g \cdot X = \mathrm{Ad}(g)X$. The corresponding notion of fundamental field is given by

$$X^*_{\phi(a,l)} = \frac{d}{dt}\Big[\,\mathrm{Ad}(e^{-tX})\phi(a,l)\Big]_{t=0}.$$

The Iwasawa theorem 51.188 claims that $G/K = AN$ and that we have global diffeomorphism $\mathcal{A}\oplus\mathcal{N} \to AN$, $(a,n) \to e^a e^n$; $\mathcal{A} \to A$, $a \to e^a$; $\mathcal{N} \to N$, $n \to e^n$. We define $\mathcal{R} = \mathcal{A}\oplus\mathcal{N}$ and the global diffeomorphism

$$\begin{aligned} \phi\colon \mathcal{A}\oplus\mathcal{N} &\to \mathcal{O} \\ aH + lE &\mapsto \operatorname{Ad}(e^{aH}e^{lE})Z. \end{aligned} \tag{83.25}$$

EqDefphiHlEZ

That map can also be seen as

$$\begin{aligned} \phi\colon \mathbb{R}^2 &\to \mathcal{O} \\ (a,l) &\mapsto \operatorname{Ad}(e^{aH}e^{lE})Z. \end{aligned} \tag{83.26}$$

In this way, we identify $\mathcal{A}\oplus\mathcal{N}$ and $\mathbb{R}^2$ as two dimensional space.

Proposition 83.11.
As homogeneous space, there is a double covering

$$\begin{aligned} \psi\colon G/K &\to \mathcal{O} \\ [g] &\mapsto \operatorname{Ad}(g)Z. \end{aligned} \tag{83.27}$$

Proof. The map ψ is well defined and injective (up to the double covering) because the stabilizer of $\mathcal{K}$ is K from theorem 51.191. The surjective condition is clear. The *double* covering is expressed by the fact that $\psi([g]) = \psi([g'])$ if and only if $g = \pm g'$. □

The symplectic 2-form ω on $\mathcal{O}$ induces a symplectic form

$$\Omega = \phi^*\omega$$

on $\mathcal{A}\oplus\mathcal{N} \simeq \mathbb{R}^2$.

Proposition 83.12.
The 2-form $\phi^*\omega$ is constant and its value is

$$\Omega := \phi^*\omega = -2B(F,E)da\wedge dl = \beta da\wedge dl;$$

in other words, ϕ is a global *Darboux chart for $\mathcal{O}$.*

prop:Omega

Proof. We have to compute

$$\Omega_{(a,l)}(\partial_a,\partial_l) = \omega_{\phi(a,l)}\big((d\phi)_{(a,l)}\partial_a, (d\phi)_{(a,l)}\partial_l\big).$$

First, we show that $d\phi(\partial_a) = -H^*_\phi$:

$$\begin{aligned} d\phi_{(a,l)}\partial_a &= \frac{d}{dt}\Big[\phi(a+t,l)\Big]_{t=0} = \frac{d}{dt}\Big[\operatorname{Ad}(e^{(a+t)H}e^{lE})Z\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\operatorname{Ad}(e^{tH}e^{aH}e^{lE})Z\Big]_{t=0} = \frac{d}{dt}\Big[\operatorname{Ad}(e^{tH})\phi(a,l)\Big]_{t=0} \\ &= -H^*_{\phi(a,l)}. \end{aligned}$$

In the same way, we find $d\phi(\partial_l) = \big(\operatorname{Ad}(e^{aH})E\big)^*_\phi$:

$$\begin{aligned} d\phi_{(a,l)}\partial_l &= \frac{d}{dt}\Big[\operatorname{Ad}(e^{aH}e^{l+tE})Z\Big]_{t=0} = \frac{d}{dt}\Big[\operatorname{Ad}(e^{aH}e^{tE}e^{-aH}e^{aH}e^{lE})Z\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\operatorname{Ad}(e^{aH}e^{tE}e^{-aH})\phi(a,l)\Big]_{t=0} = \frac{d}{dt}\Big[\operatorname{Ad}(e^{t\operatorname{Ad}(e^{aH})E})\phi(a,l)\Big]_{t=0} \\ &= -\big(\operatorname{Ad}(e^{aH})E\big)^*_{\phi(a,l)}. \end{aligned}$$

Using formula (83.24) for the symplectic form,

$$\begin{aligned} \Omega_{(a,l)}(\partial_a,\partial_l) &= B\big(\phi(a,l),[-H,-\operatorname{Ad}(e^{aH})E]\big) \\ &= B\big(\operatorname{Ad}(e^{aH})\operatorname{Ad}(e^{lE})Z, \operatorname{Ad}(e^{aH})[H,E]\big) \\ &= 2B\big(Z,\operatorname{Ad}(e^{-lE})E\big) \\ &= 2B(Z,E). \end{aligned} \tag{83.28}$$

Defining $\beta = -2B(E,F)$ we write it as

$$\Omega = \phi^*\omega = -2B(F,E)da \wedge dl = \beta da \wedge dl. \tag{83.29}$$

□

So, as symplectic manifold, $(\mathcal{O},\omega)$ is nothing but $(\mathbb{R}^2, da \wedge dl)$, the diffeomorphism being ϕ. The symplectic structure Ω induces a Poisson structure P given by equation (59.6). In the present case, it reads

$$(\Omega_{ij}) = \beta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad (P) = \beta^{-1} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{83.30}$$

and

$$\{f,g\} = \beta^{-1}(\partial_l f \partial_a g - \partial_a f \partial_l g). \tag{83.31}$$ eq:Poisson

The action of G on $\mathcal{O}$ can be turned into an action on $\mathbb{R}^2$ using the chart ϕ. It is done by defining $\tau\colon G \times \mathbb{R}^2 \to \mathbb{R}^2$,

$$\tau = \phi^{-1} \circ \mathrm{Ad} \circ \phi, \tag{83.32}$$

or $\tau_g(a,l) = \phi^{-1}\big(\mathrm{Ad}(g)\phi(a,l)\big)$. The notion of fundamental field at $x = (a,l) \in \mathbb{R}^2$ is thus given by

$$X^*_x = \frac{d}{dt}\Big[e^{-tX} \cdot x\Big]_{t=0} = \frac{d}{dt}\Big[\phi^{-1}\big(\mathrm{Ad}(e^{-tX})\phi(a,l)\big)\Big]_{t=0}, \tag{83.33}$$

for which we will often use the path representation

$$X^*_x(t) = \phi^{-1}\big(\mathrm{Ad}(e^{-tX})\phi(a,l)\big).$$

From Ad-invariance of ω,

$$\tau^*\Omega = \tau^*\phi^*\omega = (\phi \circ \phi^{-1} \circ \mathrm{Ad} \circ \phi)^*\omega = \phi^*(\mathrm{Ad})^*\omega = \phi^*\omega = \Omega.$$

Thus the symplectic form is G-invariant:

$$\tau^*\Omega = \Omega, \tag{83.34}$$ eq:tau_s_Omega

That implies in particular that τ satisfies theorem 59.2.

Proposition 83.13.
The action τ of G on the symplectic space $(\mathbb{R}^2,\Omega)$ is Hamiltonian and the dual momentum maps $\lambda'_X\colon \mathbb{R}^2 \to \mathbb{R}$ are given by (cf .59.4)

$$\lambda'_X(a,l) = -B\big(X, \phi(a,l)\big) \tag{83.35}$$

for each $X \in \mathcal{G}$. prop:lambda_X

Proof. We have first to check the identity $i(X^*)\Omega = i(X^*)(\phi^*\omega) = d\lambda'_X$. Let us apply both sides on the vector [4] A^*_x, with $A \in \mathcal{G}$ and $x = (a,l) \in \mathbb{R}^2$. On the one hand

$$i(X^*_x)\Omega_x(A^*_x) = \omega_{\phi(x)}\big(d\phi_x X^*_x, d\phi_x A^*_x\big),$$

but

$$d\phi_x X^*_x = \frac{d}{dt}\Big[\phi(X^*_x(t))\Big]_{t=0} = \frac{d}{dt}\Big[\mathrm{Ad}(e^{-tX})\phi(aH, lE)\Big]_{t=0} = -X^*_{\phi(a,l)}. \tag{83.36}$$

The same being true for A,

$$i(X^*_x)\Omega_x(A^*_x) = \omega_{\phi(x)}(X^*_{\phi(x)}, A^*_{\phi(x)}) = B(\phi(x), [X,A]).$$

4. Existence comes from lemma 53.19.

On the other hand,

$$\begin{aligned}(d\lambda'_X)_x(A^*_x) &= \frac{d}{dt}\Big[(\lambda'_X\circ\phi^{-1})\Big(\operatorname{Ad}(e^{tA})\phi(a,l)\Big)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[B\Big(X,\operatorname{Ad}(e^{tA})\phi(a,l)\Big)\Big]_{t=0}\\ &= B\Big(\frac{d}{dt}\Big[\operatorname{Ad}(e^{tA})\phi(x)\Big]_{t=0},X\Big) \qquad B \text{ is linear}\\ &= B\Big((\operatorname{ad}A)\phi(x),X\Big)\\ &= -B\big(\phi(x),(\operatorname{ad}A)X\big) \qquad B \text{ is } \operatorname{Ad}-\text{invariant}\\ &= B(\phi(x),[X,A]).\end{aligned} \tag{83.37}$$

That proves that $i(X^*)\Omega = d\lambda'_X$. The second part of the proof is to see that condition (59.8b) holds. Using the fact that $X_{\lambda'_Y} = Y^*$, we find

$$\begin{aligned}\{\lambda'_X,\lambda'_Y\}(a,l) &= -\Omega(X_{\lambda'_X},X_{\lambda'_Y}) = -\Omega_{(a,l)}(X^*,Y^*)\\ &= -\omega_{\phi(a,l)}(X^*,Y^*) = -B([X,Y],\phi(a,l))\\ &= \lambda'_{[X,Y]}(a,l)\end{aligned}$$

where the star refers to the action on $\mathcal{O}$. Explicit computations of Poisson bracket between λ'_X's at page 4164 will confirm that result.

□

We are now able to furnish explicit formulas for λ'_H, λ'_E and λ'_F by virtue of the latter proposition. The first computation is:

$$\begin{aligned}\lambda'_H(a,l) &= -B(H,\operatorname{Ad}(e^{lE})Z) = -B(\operatorname{Ad}(e^{-lE})H,Z)\\ &= -B(H+[-lE,H]+\dots,Z) = -B(H,Z)+B([-lE,H],Z)\\ &= -2lB(E,F),\end{aligned} \tag{83.38}$$

so

$$\lambda'_H(a,l) = -\beta l. \tag{83.39}$$

Second,

$$\lambda'_E(a,l) = -B(\operatorname{Ad}(e^{-aH})E,\operatorname{Ad}(e^{lE})) = -e^{-2a}B(\operatorname{Ad}(e^{-lE})E,Z) = -\frac{\beta}{2}e^{-2a}. \tag{83.40}$$

Then,

$$\lambda'_E(a,l) = -\frac{\beta}{2}e^{-2a}. \tag{83.41}$$

The last one is

$$\begin{aligned}\lambda'_F(a,l) &= -B\big(\operatorname{Ad}(e^{lE})Z,e^{-aH}F\big) = -e^{2a}B\big(Z,\operatorname{Ad}(e^{-lE})F\big)\\ &= -e^{2a}B\big(Z,F-l[E,F]+\frac{l^2}{2}[E,[E,F]]+\dots\big)\\ &= -e^{2a}\left[B(Z,F)-lB(Z,H)-\frac{l^2}{2}B(Z,2E)\right]\\ &= -e^{2a}\big(B(Z,F)+l^2B(F,E)\big)\\ &= -e^{2a}\big(-\frac{\beta}{2}-l^2\frac{\beta}{2}\big) = e^{2a}\frac{\beta}{2}(l^2+1).\end{aligned} \tag{83.42}$$

Finally,

$$\lambda'_F(a,l) = \frac{\beta}{2}e^{2a}(l^2+1). \tag{83.43}$$

Using formula (59.7) for the Poisson bracket, one can check that the required relations (59.8b) are satisfied: pg:explic_com_lamb

$$\{\lambda'_H, \lambda'_E\} = 2\lambda'_E \tag{83.44a}$$
$$\{\lambda'_H, \lambda'_F\} = -2\lambda'_F \tag{83.44b}$$
$$\{\lambda'_E, \lambda'_F\} = \lambda'_H. \tag{83.44c}$$

This confirms the fact that our action of $\mathrm{SL}(2,\mathbb{R})$ on AN is Hamiltonian.

Using the global diffeomorphism (83.25), and the map

$$\begin{aligned} j\colon AN &\to \mathcal{O} \\ r &\mapsto \mathrm{Ad}(r)Z \end{aligned} \tag{83.45}$$

we identify

$$R \simeq \mathcal{O} \simeq \mathbb{R}^2.$$

The action of R on itself induced from the adjoint action of R on $\mathcal{O}$ is

$$r \cdot s = j^{-1}\big(r \cdot j(s)\big) = j^{-1}\big(\mathrm{Ad}(rs)Z\big) = rs.$$

It is the left multiplicative action required in definition 83.2. The Lie group R is endowed with the symplectic form

$$\omega^R = j^* {\phi^{-1}}^* \Omega.$$

The notion of fundamental vector for the action of R on itself is given by

$$X^*_r = \frac{d}{dt}\Big[e^{-tX} \cdot r\Big]_{t=0} = \frac{d}{dt}\Big[j^{-1}\big(e^{-tX} \cdot j(r)\big)\Big]_{t=0} = dj^{-1} X^*_{j(r)}, \tag{83.46}$$

but we know that

$$e^{-tX} \cdot j(r) = \mathrm{Ad}(e^{-tX_r})Z = [\phi \circ \tau(e^{-tX} r) \circ \phi^{-1}]Z,$$

then

$$X^*_r = dj^{-1} \circ d\phi X^*_{r \cdot \phi^{-1}(Z)}.$$

If $r = e^{aH}e^{lE}$, then $r \cdot \phi^{-1}(Z) = (a,l)$ and

$$X^*_r = (dj^{-1} \circ d\phi) X^*_{(a,l)} \tag{83.47}$$

where the fundamental field of the right hand side is taken in the sense of the action of R on $\mathbb{R}^2$.

The following proposition shows that the explicit form of λ and λ' are the same up to natural identifications.

Proposition 83.14.
The left multiplicative action of R on itself is Hamiltonian and the dual momentum maps are given by $\lambda_X\colon R \to \mathbb{C}$,

$$\lambda_X = \lambda'_X \circ \phi^{-1} \circ j. \tag{83.48}$$

for each $X \in \mathcal{R}$. PropMomslR

Proof. Once again, the proof is just a verification of the two properties of a momentum map. The first one is

$$\begin{aligned} i(X^*_r)\omega^R Y &= \omega^R_r(dj^{-1}d\phi X^*_{(a,l)}, Y) = \Omega_{(\phi^{-1}\circ j)(r)}\big(X^*_{(a,l)}, d\phi^{-1} dj_r Y\big) \\ &= (\lambda'_X \circ d\phi^{-1} \circ dj)Y = d\lambda_X Y. \end{aligned} \tag{83.49}$$

For the second condition, we consider $r = e^{aH}e^{lE}$ and

$$\begin{aligned} \{\lambda_X, \lambda_Y\}(r) = X^*_r(\lambda_Y) &= (dj^{-1}d\phi X^*_{(a,l)})(\lambda'_Y \circ \phi^{-1} \circ j) \\ &= X^*_{(a,l)}(\lambda'_Y) = \lambda'_{[X,Y]}(a,l) \in \mathbb{C} \end{aligned} \tag{83.50}$$

while

$$\lambda_{[X,Y]}(r) = \lambda'_{[X,Y]} \circ \phi^{-1} \circ j(r) = \lambda'_{[X,Y]}(a,l).$$

□

83.2.2 Guessing the star product

The Moyal star product is invariant under the action of $\mathbb{R}^2$ on itself $L_x y = x + y$ in the sense that if we pose $(L_y^* f)(x) = f(x + y)$ it is clear that

$$\begin{aligned}(L_s^* f *_M L_s^* g)(x) &= \exp\left[\frac{\nu}{2} P^{ij}(\partial_{y^i} \wedge \partial_{z^j})\right] f(y+s)g(z+s)|_{y=z=x} \\ &= L_s^*(f *_M g)(x).\end{aligned} \tag{83.51}$$

We are however not interested by that action on $\mathbb{R}^2$. The action which we look at is the one of $\mathrm{SL}(2, R)$.

Proposition 83.15.
The product $*_M$ is $\mathfrak{sl}(2, \mathbb{R})$-invariant at order 0 and 1.

Proof. The invariance at order zero is given with some concise notations by

$$(gu)(gv)(x) = u(gx)v(gx) = (uv)(gx),$$

The action τ_g of an element $g \in G$ satisfies $\tau_g^* \Omega = \Omega$ (equation (83.34)), so theorem 59.2 gives $\{u \circ \tau_g, v \circ \tau_g\} = \{u, v\} \circ \tau_g$. Since Poisson bracket is the first term of the Moyal product, at first order

$$\tau_g^*(u *_M v) = \tau_g^* u *_M \tau_g^* v.$$

□

Proposition 83.16.
The product $*_M$ is $\mathfrak{sl}(2, \mathbb{R})$-covariant for the homomorphism given by proposition 83.13 or equivalently by equations (83.39), (83.41), and (83.43).

Proof. The Moyal star product can be written as

$$u *_M v = \sum \frac{\nu^k}{k!} P_k(u, v)$$

with $P_k(u, v) = \Omega^{IJ} \partial_I u \partial_J v$ where I and J are summed over k-uple of 0 and 1, including a sum over k itself ($x^0 = a$, $x^1 = l$). For a given I, there is only one J such that $\Omega^{IJ} \neq 0$. There are $\binom{k}{m}$ multi-indices I providing the term $\partial_I = \partial_0^m \partial_1^n$ with $n + m = k$. For each of them, $\Omega^{IJ} = (-1)^n$. Therefore

$$P_k(u, v) = \sum_{m=0}^{k} (-1)^{k-m} \binom{k}{m} \partial_0^m \partial_1^n u \, \partial_0^n \partial_1^m v. \tag{83.52}$$ eq:P_k

For example,

$$P_1(u, v) = -\partial_1 u \partial_0 v + \partial_0 u \partial_1 v = \{u, v\}.$$

If k is even, the expression (83.52) is symmetric with respect of u and v, so that these terms will not contribute in the computation of the commutators $[u, v]_{*_M}$. We are left with

$$[\lambda'_X, u]_{*_M} = 2 \sum_{k=0}^{\infty} \frac{\nu^{2k+1}}{(2k+1)!} P_{2k+1}(\lambda'_X, u). \tag{83.53}$$ eq:comm_lambda_X

First we compute $[\lambda'_H, u]_{*_M}$:

$$P_{2k+1}(\lambda'_H, u) = \delta_{k0}(-\partial_1 \lambda'_H \partial_0 u + \partial_0 \lambda'_H \partial_1 u) = \delta_{k0}\{\lambda', u\}, \tag{83.54}$$

thus

$$[\lambda'_H, u]_{*_M} = 2\nu P_1(\lambda'_H, u) = 2\nu\{\lambda'_H, u\} = 2\nu\beta\partial_a u. \tag{83.55}$$

By the way, we point out the relation

$$\mathrm{ad}_{*_M}\lambda'_H = 2\nu\beta\partial_a.$$

Now, we turn our attention to the commutator $[\lambda'_E, u]_{*_M}$:

$$\begin{aligned} P_{2k+1}(\lambda'_E, u) &= -\sum_{n=0}^{k}(-1)^m\binom{2k+1}{m}\binom{2k+1}{m}(\partial_0^m\partial_1^n\lambda'_E)\,(\partial_0^n\partial_1^m u) \\ &= \partial_a^{2k+1}\big(-\frac{\beta}{2}e^{-2a}\big)\partial_l^{2k+1}u = \beta 2^{2k}e^{-2a}\partial_l^{2k+1}u, \end{aligned} \tag{83.56}$$

thus

$$[\lambda'_E, u]_{*_M} = 2\sum_{k=0}^{\infty}\frac{\nu^{2k+1}}{(2k+1)!}\beta 2^{2k}e^{-2a}\partial_l^{2k+1}u = \beta e^{-2a}\sinh(2\nu\partial_l)u, \tag{83.57}$$

eq:comm_lambda_E

so that

$$\mathrm{ad}_{*_M}\lambda'_E = \beta e^{-2a}\sinh(2\nu\partial_l).$$

Last we check $[\lambda'_E, \lambda'_F]_{*_M} = 2\nu\{\lambda'_E, \lambda'_F\}$. When $u = 0$, the only non vanishing term in the sum (83.57) is $k = 0$. Since $\partial_l^3\lambda'_F = 0$,

$$[\lambda'_E, \lambda'_F]_{*_M} = 2\nu\beta e^{-2a}\partial_l\lambda'_F,$$

but

$$2\nu\{\lambda'_E, \lambda'_F\} = 2\nu(\partial_a\lambda'_E\partial_l\lambda'_F - \partial_l\lambda'_E\partial_a\lambda'_F) = 2\nu\beta e^{-2a}\partial_l\lambda'_F. \tag{83.58}$$

□

Before to go on, let us compute the operator $\mathrm{ad}_{*_M}\lambda'_F$ in order to complete our collection. We take once again the formula (83.52), with λ'_F and u:

$$P_{2k+1}(\lambda'_F, u) = -\sum_{m=0}^{2k+1}(-1)^m\binom{2k+1}{m}\partial_a^m\partial_l^n\lambda'_F\partial_a^n\partial_l^m u. \tag{83.59}$$

It is clear that λ'_F can be derived only two times with respect of l and as much as we want with respect of a. Then possible n are $n = 0, 1, 2$, whose corresponding m are $2k-1$, $2k$, and $2k+1$. Some computations lead to

$$\begin{aligned} P_{2k+1}(\lambda'_F, u) = &-k(2k+1)\beta 2^{2k-1}e^{2a}\partial_a^2\partial_l^{2k-1}u \\ &+ (2k+1)\beta 2^{2k}l\partial_a\partial_l^{2k}u \\ &- \beta 2^{2k}(1+l^2)e^{2a}\partial_l^{2k+1}u. \end{aligned} \tag{83.60}$$

Replacing into the series (83.53), we find

$$\begin{aligned} [\lambda'_F, u]_{*_M} = e^{2a}\Big\{&\sum\frac{\nu^{2k+1}}{(2k+1)!}\beta(-k)(2k+1)\frac{2^{2k+1}}{2}\partial_a^2\partial_l^{2k-1}u \\ &+\sum\frac{\nu^{2k+1}}{(2k+1)!}(2k+1)2^{2k+1}l\partial_a\partial_l^{2k}u \\ &\beta\sum\frac{\nu^{2k+1}}{(2k+1)!}2^{2k+1}(1+l^2)\partial_l^{2k+1}u\Big\} \\ = &-\beta e^{2a}\partial_a^2\sum_{k=1}^{\infty}\frac{(2\nu)^{2k}}{(2k)!}k\nu\partial_l^{2k-1} \\ &+ 2\beta\nu e^{2a}\partial_a\circ\cosh(2\nu\partial_l) \\ &- \beta e^{2a}(1+l^2)\sinh(2\nu\partial_l). \end{aligned}$$

Finally,

$$\begin{aligned}\mathrm{ad}_{*_M}\lambda'_F = &-\nu^2\beta e^{2a}\partial_a^2\circ\sinh(2\nu\partial_l)\\ &+2\nu\beta e^{2a}l\partial_a\circ\cosh(2\nu\partial_l)\\ &-e^{2a}(1+l^2)\sinh(2\nu\partial_l).\end{aligned}\tag{83.61}$$

Corollary 83.17.
The star product $*_M^R$ on R defined for $u, v \in C^\infty(R)$ by

$$(u *_M^R v)(r) = (u\circ T^{-1} *_M v\circ T^{-1})T(r)\tag{83.62}$$

where $T = \phi^{-1}\circ j$ is covariant for the functions λ of proposition 83.14.

Remark that from general theory of star products, the so-defined $*_R$ is a formal star product on R.

Proof. From definition of $*_M^R$, on the one hand

$$(\lambda_X *_M^R \lambda_Y)(r) - X\leftrightarrow Y = (\lambda'_X *_M \lambda'_Y)T(r) - X\leftrightarrow Y = 2\nu\{\lambda'_X,\lambda'_Y\}_{\mathbb{R}^2}T(r),$$

while on the other hand, $\omega^R = T^*\Omega$, so that point (2) of theorem 59.2 gives

$$\{\lambda'_X,\lambda'_Y\}_{\mathbb{R}^2}\circ T = \{\lambda'_X\circ T,\lambda'_Y\circ T\}_R = \{\lambda_X,\lambda_Y\}_R.$$

□

All that makes the theory developed earlier (in particular proposition 83.10) valid here. So we pose

$$\begin{aligned}\rho_\nu\colon \mathcal{R} &\to \mathrm{End}\left(C^\infty(R)[[\nu]]\right)\\ X &\mapsto \frac{1}{2\nu}\mathrm{ad}_{*_M^R}(\lambda_X);\end{aligned}\tag{83.63}$$

using the explicit expressions of $ad_{*_M}(\lambda'_X)$, we find

$$\rho_\nu(H) = \beta\partial_a,\qquad \rho_\nu(E) = \frac{\beta}{2\nu}e^{-2a}\sinh(2\nu\partial_l).\tag{83.64}$$

Using (83.19) with $\lambda_H = -\beta l$, it is clear that $dL(H) = -\beta\{l,u\} = \beta\partial_a u$. Therefore

$$\rho_\nu(H) = dL(H),\tag{83.65}$$

but the requested identity $\rho_\nu(E) = dL(E)$ will not hold. The problem is that $dL(X) = X_{\lambda_X}$ is a vector field, while $\rho_\nu(E)$ comes with (infinitely) multiple derivatives, hence this is not a vector field. Conclusion: the operator $\mathcal{T}$ of equation (83.20) must not act on the variable a.

First we consider a partial Fourier transform $\mathcal{F}$:

$$(\mathcal{F}u)(a,\alpha) = \hat{u}(a,\alpha) := \frac{1}{\sqrt{2\pi}}\int e^{-i\alpha l}u(a,l)dl,\tag{83.66}$$

the inverse being given by

$$(\mathcal{F}^{-1}\hat{u})(a,l) = \frac{1}{\sqrt{2\pi}}\int e^{il\alpha}\hat{u}(a,\alpha)d\alpha.\tag{83.67}$$

It is clear that $\mathcal{F}\rho_\nu(H)\mathcal{F}^{-1} = \rho_\nu(H)$, but $\mathcal{F}\rho_\nu(E)\mathcal{F}^{-1} = \frac{\beta}{2\nu}e^{-2a}\sinh(2i\nu\alpha)$. Indeed, if we define $\hat{v}(a,\alpha) = \sinh(2i\nu\alpha)\hat{u}(a,\alpha)$,

$$\begin{aligned}(\rho_\nu(E)\mathcal{F}^{-1}\hat{u})(a,l) &= \frac{\beta}{2\nu}e^{-2a}\frac{1}{\sqrt{2\pi}}\sinh(2\nu\partial_l)\int e^{il\alpha}\hat{u}(a,\alpha)d\alpha\\ &= \frac{\beta}{2\nu}e^{-2a}\frac{1}{\sqrt{2\pi}}\int e^{il\alpha}\sinh(2i\alpha\nu)\hat{u}(a,\alpha)d\alpha\\ &= \frac{\beta}{2\nu}e^{-2a}(\mathcal{F}^{-1}\hat{v})(a,l).\end{aligned}\tag{83.68}$$

This is nothing but the fact that the Fourier transform turns a derivation into a multiplication.

As can be seen on an asymptotic expansion, the deformation ν parameter is necessarily purely imaginary, then we can here pose $\nu = i\theta$ with $\theta \in \mathbb{R}$, so that

$$\mathcal{F}\rho_\nu(E)\mathcal{F}^{-1} = \frac{\beta i}{2\theta} e^{-2a} \sinh(2\alpha\theta). \tag{83.69}$$

eq:FrnEF

Using (83.19), we find

$$dL(E) = \beta e^{-2a}\partial_l. \tag{83.70}$$

eq:dLE

Comparing it with the expression of $\mathcal{F}\rho_\nu(E)\mathcal{F}^{-1}$, we see that (up to constant factor) we have to act in such a way that $\sinh(2\alpha\theta)$ is converted into a derivation. This is done by a Fourier transform. We pose $\xi = \sinh(2\theta\alpha)$ and

$$\tilde{f}(a,\xi) = \frac{1}{\sqrt{2\pi}} \int e^{i\xi p} f(a,p) dp.$$

As usual,

$$\widetilde{\partial_\alpha f} = -i\xi \tilde{f}.$$

This suggests us to consider the change of variable

$$\phi_\theta(a,\alpha) = (a, \frac{1}{2\theta}\sinh(2\theta\alpha)),$$

and finally,

$$\mathcal{T}_\theta := \mathcal{F}^{-1} \circ \phi_\theta^* \circ \mathcal{F}, \tag{83.71}$$

where ϕ_θ^* is defined by $(\phi_\theta^* u)(a,\alpha) = u(a, \frac{1}{2\theta}\sinh(2\theta\alpha))$. The result of our construction is the following which proves that we are in the situation of proposition 83.10.

Theorem 83.18.

$$\mathcal{T}_\theta \circ \rho_\nu(E) \circ \mathcal{T}_\theta^{-1} = \beta e^{-2a}\partial_l = dL(E).$$

Proof. Notice that $(\phi_\theta^* \mathcal{F} u)(a,\alpha) = \hat{u}(a, \sinh(2\theta\alpha))$, and then define $\hat{v}(a,\alpha) = \hat{u}(a, \sinh(2\theta\alpha))$; equation (83.69) is

$$(\mathcal{F}\rho_\nu(E)\mathcal{F}^{-1}\hat{v})(a,\alpha) = \frac{\beta i}{2\theta} e^{-2a} \sinh(2\theta\alpha)\hat{v}(a,\alpha).$$

Applying $(\phi^*)^{-1}$ on the right hand side, we find $\frac{\beta i}{2\theta} e^{-2a} 2\theta\alpha\hat{u}(a,\alpha)$. This allows us to compute

$$\begin{aligned}(\mathcal{T}_\theta \rho_\nu(E) \mathcal{T}_\theta^{-1}) u(a,l) &= \beta i e^{-2a} \mathcal{F}^{-1}(\alpha\hat{u})(a,l) \\ &= \beta i e^{-2a} \frac{1}{\sqrt{2\pi}} \int \hat{u}(a,\alpha)(-i)\partial_l e^{il\alpha} d\alpha \\ &= \beta e^{-2a}(\partial_l u)(a,l).\end{aligned}$$

□

83.2.3 Formula for the product

The fact the $\mathcal{T}_\theta$ intertwines ρ_ν and dL makes that the candidate to be a product on the AN of $\mathrm{SL}(2,\mathbb{R})$ can be computed using formula (83.21). Computations are rather long and done in the articles [961] and [1039] (see particularly point 4), so we will not give them here. We will also not precise the functional space of convergence for the resulting product.

¡¡ Avertissement/question au lecteur !! 83.19
Il faut voir si ça entrelace effectivement ces deux cocos, ou bien juste $\rho_\nu(E)$ et $dL(E)$. Probintertw

We are searching for a function $\rho\colon R \to \mathbb{R}$ such that

$$\int_R f(r)\rho(r)dr = \int_R f(br)\rho(r)dr, \tag{83.72}$$

where $dr = da \wedge dl$. We consider the multiplication map $\phi\colon R \times R \to R$, $\phi(r, r') = rr'$, and its partial derivatives $A^i_j = \frac{\partial \phi^i}{\partial r^j}$. We will perform the change of variable $s = br$ in (83.72). So, $s^i = \phi^i(b, r)$, and if we pose , we have $dr^k = (A^{-1})^k_i ds^i$. We find

$$\int_R f(r)\rho(r)dr = \int_R f(s)\rho(b^{-1}s)\det(A^{-1})ds.$$

In the right-hand side, we rename s to r and we impose the equality to be correct for all functions $f\colon R \to \mathbb{C}$. Then for all r, $k \in R$,

$$\rho(r) = \rho(k^{-1}r)\det(A^{-1}).$$

In order to compute the value of $\det(A)$, we have to write $\big(dL_{(a,b)}\big)_{(a',l')}$ in a matrix form. If we consider $v = (a'(t), l'(t))$, we have

$$\big(dL_{(a,b)}\big)_{(a',l')} v = \frac{d}{dt}\begin{pmatrix} e^{a+a'(t)} & e^{a+a'(t)}\big(l'(t) + e^{-2a'(t)l}\big) \\ 0 & e^{-a-a'(t)} \end{pmatrix}_{t=0}, \tag{83.73}$$

but we know that v can be written as

$$v = \begin{pmatrix} \frac{da'}{dt}e^{a'} & \frac{da'}{dt}e^{a'}l' + e^{a'}\frac{dl'}{dt} \\ 0 & -\frac{da'}{dt}e^{a'} \end{pmatrix}$$

Now, if we want to write $\big(dL_{(a,b)}\big)_{(a',l')}$ as $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, we obtain: $A = e^a$, $C = 0$, $D = e^{-a}$, and thus $\det(A) = 1$.

So we can use $\rho(r) = 1$ for all $r \in R$, and the integral of $f\colon R \to \mathbb{C}$ over R can be written as

$$\int_R f(r)dr$$

with $dr = da \wedge dl$.

In the parametrization

$$(a, l) = \begin{pmatrix} e^a & e^a l \\ 0 & e^{-a} \end{pmatrix},$$

of $R = AN$ the form $da \wedge dl$ is a left invariant measure, so the integral of the function $f\colon R \to \mathbb{R}$ on R is given by

$$\int_R f = \int_{\mathbb{R}^2} f(a, l)da\, dl.$$

Remark that $da\, dl$ is the Liouville measure by proposition 83.12. It is important for definition 83.2.

We consider a subset $\mathbb{A} \subset \mathrm{Fun}(R)$, and we define the product $\star^R_\theta$ on $\mathbb{A}$ by

$$(a \star^R tb)(a_0, l_0) = \int_{R\times R} K^R_\theta\big((a_0, l_0), (a_1, l_1), (a_2, l_2)\big) \\ a(a_1, l_1)b(a_2, l_2)da_1 dl_1 da_2 dl_2. \tag{83.74}$$

where

$$K^A_\theta(g_0, g_1, g_2) = \frac{1}{\theta^2}\mathcal{A}^R(g_0, g_1, g_2)e^{i\theta\mathcal{S}^R(g_0,g_1,g_2)}$$

with

$$\mathcal{A}^R(g_0, g_1, g_2) = \bigoplus_{0,1,2} \cosh(a_1 - a_2) \tag{83.75a}$$

$$\mathcal{S}^R(g_0, g_1, g_2) = \bigoplus_{0,1,2} \sinh(2(a_0 - a_1))l_2. \tag{83.75b}$$

Here, the symbol $\bigoplus_{0,1,2}$ stands for a cyclic sum over the indices $0, 1, 2$. These functions are invariant under the $\mathrm{SL}(2, R)$ left action: when x, y, z, $g \in \mathrm{SL}(2, R)$,

$$\mathcal{A}^R(gx, gy, gz) = \mathcal{A}^R(x, y, z) \quad \text{and} \quad \mathcal{S}^R(gx, gy, gz) = \mathcal{S}^R(x, y, z). \tag{83.76}$$

The first equality is clear; let us show the second. If $x = (a_x, l_x)$ (the same for y and z) and $g = (a, l)$, the computation of $\mathcal{S}^R(gx, gy, gz)$ is the one of $\mathcal{S}^R(x, y, z)$ with the replacements

$$a_x \to a_x + l \tag{83.77a}$$
$$l_x \to l_x + e^{-2a_x} l, \tag{83.77b}$$

and the same for y and z. Using the formula $\sinh t = \frac{1}{2}(e^t - e^{-t})$, one finds the right cancellations.

¡¡ Avertissement/question à la lectrice !! 83.20
Il faut trouver l'article où ce résultat se trouve, et citer le théorème exact. ProbEnonSLdef

83.2.3.1 Remark on (formal) star product

subsec:rem_on_sp

Since proposition 83.10, we know that the star product to be used is

$$(u \star_\theta v) = T_\theta(T_\theta^{-1} u *_M T_\theta^{-1} v)$$

with $T_\theta = \mathcal{F}^{-1} \circ \phi_\theta^* \circ \mathcal{F}$ and $\phi_\theta(a, \alpha) = (a, \frac{1}{2\theta}\sinh(2\alpha\theta))$. One easily finds that

$$\begin{aligned}(T_\theta^{-1} u)(a, l) &= (\mathcal{F}^{-1} \circ \phi_\theta^* \circ \mathcal{F} u)(a, l) \\ &= \frac{1}{2\pi} \int e^{i\alpha l} d\alpha \int e^{-im \frac{1}{2\theta} \sinh(2\alpha\theta)} u(a, m) dm \\ &= \frac{1}{2\pi} \int e^{i[\alpha l - \frac{1}{2\pi} \sinh(2\theta\alpha) l_1]} u(a, l_1) dl_1 d\alpha.\end{aligned} \tag{83.78}$$

From this expression, we have to find an integral expression for the product $\star_\theta$ and see that it is a WKB product. In order to do it, we have to write the Moyal star product into an integral form. So we have to develop a theory in which one can write oscillatory integrals and power expansions which give the usual Moyal expansion (82.4). This is our subject of concern now.

An other way to state the same problem is to consider (V, Ω), a symplectic vector space of dimension $2n$. The **Weyl product** of u, $v \in C_c^\infty(V)$ at $x \in V$ is defined by

$$(u \star^W v)(x) = \theta^{-2n} \int_{V \times V} e^{\frac{2i}{\theta} S^0(x,y,z)} u(y) v(z) dy z$$

where $S^0(x, y, z) = \Omega(x, y) + \Omega(y, z) + \Omega(z, x)$. With the definition $\nu = \theta/2i$, it can be asymptotically developed by

$$u \star_\theta^W v \sim uv + \nu\{u, v\} + \sum_{k=2}^{\infty} \frac{\nu^k}{k!} \Omega^{IJ} \partial_I u \partial_J v.$$

How can we give a sense to this expansion?

One can find a definition of an asymptotic expansion for oscillating integrals in [951] under the form

$$I_\lambda = \int e^{(i/\lambda) S(x)} \phi(x) \sim \sum_n \lambda^n c_n.$$

It can be shown that such a expansion used on (83.74) gives rise of a formal star product:

$$(a \star^R b)(g) \sim a(g) b(g) + \frac{\theta}{2i}\{a, b\}(g) + o(\theta^2). \tag{83.79}$$ EqDevFedFor

83.3 Non formal Extension lemma

SecExtLem

Let $(\mathfrak{s}_i, \Omega_i)_{i=1,2}$ be symplectic Lie algebras and (S_i, ω_i) the respective Lie groups with left invariant symplectic forms: $(\omega_i)_g = (L_g)^*\Omega_i$. We suppose to know a homomorphism $\rho\colon \mathfrak{s}_1 \to \mathtt{Der}(\mathfrak{s}_2) \cap \mathfrak{sp}(\Omega_2)$ and a Darboux chart $\phi_i\colon \mathfrak{s}_i \to S_i$ for each of the two symplectic Lie groups. Our first purpose is to build a Darboux chart on the split extension

$$\mathfrak{s} := \mathfrak{s}_1 \oplus_\rho \mathfrak{s}_2.$$

Remarque 83.21.
Most of the time we are in the inverse situation: we have an algebra $\mathfrak{s}$ which turn out to be a split extension $\mathfrak{s}_1 \oplus_{\mathrm{ad}} \mathfrak{s}_2$ for which we have to check that $\mathrm{ad}(\mathfrak{s}_1)$ is a symplectic action of $\mathfrak{s}_1$ on $(\mathfrak{s}_2, \Omega_2)$. See the example of section 85.2.

Proposition 83.22.
In this setting, the map $\phi\colon \mathfrak{s} \to S$,

$$\phi(X_1, X_2) = \phi_2(X_2)\phi_1(X_1) \tag{83.80}$$

is a Darboux chart. prop:Darboux

Proof. An element $X \in T_{\phi^{-1}(g)}(\mathfrak{s}_1 \oplus \mathfrak{s}_2) = \mathfrak{s}_1 \oplus \mathfrak{s}_2$ is denoted by $X = (X_1, X_2)$ with $X_i \in \mathfrak{s}_i$, and the symplectic form on $\mathfrak{s}_1 \oplus \mathfrak{s}_2$ is given by

$$\Omega\big((A_1, A_2), (B_1, B_2)\big) = \Omega_1(A_1, B_1) + \Omega_2(A_2, B_2) \tag{83.81}$$ eq:Omega

where we identify $\mathfrak{s}_i$ and $T_e\mathfrak{s}_i$. Let A and B belongs to $T_{\phi^{-1}(g)}(\mathfrak{s}_1 \oplus \mathfrak{s}_2)$. We have to show that the quantity

$$\begin{aligned}\omega_g\Big((d\phi)_{\phi^{-1}(g)}A, (d\phi)_{\phi^{-1}(g)}B\Big)& \\ = \omega_e\Big((dL_{g^{-1}})_g(d\phi)_{\phi^{-1}(g)}A, (dL_{g^{-1}})_g(d\phi)_{\phi^{-1}(g)}B\Big)&\end{aligned} \tag{83.82}$$ eq:omega_g_et_omega_e

does not depend on g.

The vector A is represented by a path $A(t) = (A_1(t), A_2(t))$ with $A_i(t) \in \mathfrak{s}_i$. In order to characterise that path, we want first to know precisely what is $A_i(0)$. Since $A \in T_{\phi^{-1}(g)}\mathfrak{s}$, the path must fulfil $\phi(A_1(0), A_2(0)) = g$, or

$$\phi_2(A_2(0))\phi_1(A_1(0)) = g. \tag{83.83}$$

We denote $A_i(0) = G_i \in \mathfrak{s}_i$ and $\phi_i(G_i) = g_i$. The relation between g_1 and g_2 is $g_2g_1 = g$. In particular, it is wrong to say "$A_1(0) = \phi^{-1}(g)$, thus $\phi_1(A_1(0)) = g$". This point being clear,

$$(d\phi)_{\phi^{-1}(g)}A = \frac{d}{dt}\Big[\phi(A(t))\Big]_{t=0} = \frac{d}{dt}\Big[\phi_2(A_2(t))\phi_1(A_1(t))\Big]_{t=0}. \tag{83.84}$$

If one particularises to the case $A \in \mathfrak{s}_2$, that is $A_1(t) = cst = G_1$,

$$(d\phi)_{\phi^{-1}(g)}A = (dR_{g_1})_{g_2}(d\phi_2)_{G_2}A_2. \tag{83.85}$$

Since $g = g_1g_2$, we have $L_{g^{-1}} = L_{g_1^{-1}} \circ L_{g_2^{-1}}$, and the first argument of ω_e in equation (83.82) is

$$(dL_{g_1^{-1}})_{g_1}(dL_{g_2^{-1}})_{g_2g_1}(dR_{g_1})_{g_2}(d\phi_2)_{G_2}A_2.$$

If we write that in terms of the derivative of the path $A_2(t)$, what we get in the derivative is

$$g_1^{-1}g_2^{-1}\phi_2(A_2(t))g_1 = \mathbf{Ad}_{g_1^{-1}}\Big(g_2^{-1}\phi_2(A_2(t))\Big). \tag{83.86}$$

Since $g_2^{-1}\phi\big(A_2(0)\big) = g_2^{-1}\phi_2(G_2) = e$, the derivative of that term is

$$(dL_{g^{-1}})_g(d\phi)_{\phi^{-1}(g)}A = \mathrm{Ad}_{g_1^{-1}}\Big((dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}A\Big) \tag{83.87}$$

with some abuse between $A \in \mathfrak{s}$ and $A_2 \in \mathfrak{s}_2$. Doing the same computation with $B \in \mathfrak{s}_1$ (so that $B_2(t) = cst = G_2$), we find

$$(d\phi)_{\phi^{-1}(g)}B = \frac{d}{dt}\Big[g_2\phi_1(B_1(t))\Big]_{t=0} = (dL_{g_2})_{g_1}(d\phi_1)_{G_1}B_1, \tag{83.88}$$

and what appears in ω_e reads

$$(dL_{g^{-1}})_g(dL_{g_2})_{g_1}(d\phi_1)_{G_1}B = (dL_{g_1^{-1}})_{g_1}(d\phi_1)_{G_1}B. \tag{83.89}$$

Finally, for $A \in \mathfrak{s}_2$ and $B \in \mathfrak{s}_1$,

$$\begin{aligned}\omega_g\big((d\phi)_{\phi^{-1}(g)}A, (d\phi)_{\phi^{-1}(g)}B\big)&\\ = \omega_e\Big(\mathrm{Ad}_{g_1^{-1}}\big[(dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}A\big], (dL_{g_1^{-1}})_{g_1}(d\phi_1)_{G_1}B\Big).&\end{aligned} \tag{83.90}$$

eq:gros_omega_e

The first argument belongs to $T\mathfrak{s}_2$ (because $g_2 \in \mathfrak{s}_2$) while the second belongs to $T\mathfrak{s}_1$. Hence definition (83.81) makes the right hand side vanishing.

If we want to compute equation (83.90) with A, $B \in \mathfrak{s}_2$,

$$\begin{aligned}\omega_e\Big(\mathrm{Ad}_{g_1^{-1}}\big[(dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}A\big], \mathrm{Ad}_{g_1^{-1}}\big[(dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}B\big]\Big)&\\ = \Omega(\dots) = \underbrace{\Omega_1(\dots)}_{=0} + \Omega_2(\dots)&\\ = \Big(\mathrm{Ad}^*_{g_1^{-1}}\Omega_2\Big)\Big((dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}A, \dots B\Big)&\end{aligned} \tag{83.91}$$

At this point, notice that $\mathrm{Ad}^*_{g_1}\Omega_2 = \Omega_2$. Indeed the exponential $\exp\colon \mathfrak{s}_1 \to S_1$ being surjective, there exists a $X_1 \in \mathfrak{s}_1$ such that $\mathrm{Ad}(g_1) = e^{\mathrm{ad}(X_1)}$. Now, $\mathrm{ad}(X_1) \in \mathfrak{sp}(\Omega_2)$ by assumption, so that $\mathrm{Ad}(g_1) \in \mathrm{SP}(\Omega_2)$. The previous expression becomes

$$\begin{aligned}\Omega_2((dL_{g_2^{-1}})_{g_2}(d\phi_2)_{G_2}A, \dots B) &= (\omega_2)_{g_2}((d\phi_2)_{g_2}A, \dots B)\\ &= (\phi_2^*\omega_2)_{G_2}(A, B)\\ &= \Omega_2(A, B).\end{aligned} \tag{83.92}$$

The last line is the fact that ϕ_2 is a Darboux chart: $\phi_2^*\omega_2 = \Omega_2$. The case with A, $B \in \mathfrak{s}_1$ yields to compute

$$\omega_e\Big((dL_{g_1^{-1}})_{g_1}(d\phi_1)_{G_1}A, (dL_{g_1^{-1}})_{g_1}(d\phi_1)_{G_1}B\Big).$$

It is done by the same way as the previous cases. □

A direct computation shows the following **extension lemma**.

EXT

Lemma 83.23 (Extension lemma)**.**
Let $K_i \in \mathrm{Fun}(S_i^3)$ be a left invariant three point kernel on S_i ($i = 1, 2$). Assume that $K_2 \otimes 1 \in \mathrm{Fun}(S^3)$ is invariant under conjugation by elements of S_1. Then $K := K_1 \otimes K_2 \in \mathrm{Fun}(S^3)$ is left invariant (under S).

Proof. Every element of S can be written under the form g_1g_2 with $g_i \in S_i$. The multiplication is given by $(g_1g_2)(a_1a_2) = (g_1a_1)(g_2a_2)$. Using this rule, the definition of the tensor product, and the left invariance of both K_i,

$$\begin{aligned}\big(L_{g_1g_2}(K_1 \otimes K_2)\big)(a_1a_2, b_1b_2, c_1c_2)&\\ = (K_1 \otimes K_2)\big((g_1g_2)(a_1a_2), (g_1g_2)(b_1b_2), (g_1, g_2)(c_1c_2)\big)&\\ = K_1(g_1a_1, g_1b_1, g_1c_1)K_2(g_2a_2, g_2b_2, g_2c_2)&\\ = K_1(a_1, b_1, c_1)K_2(a_2, b_2, c_2)&\\ = (K_1 \otimes K_2)(a_1a_2, b_1b_2, c_1c_2).&\end{aligned}$$

□

This lemma shows that if one has kernels on S_1 and S_2 satisfying the above hypotheses, their tensor product provides a kernel for an associative left invariant kernel on $S = S_1 \otimes_\rho S_2$. Proposition 83.22 allows us to hope that the product on S will satisfy the same kind of symplectic compatibility as the products on S_i; in particular when the latter were constructed using Darboux chart in the same way as the product described in section 83.2.

83.4 Deformation of SU(1, n)

SecDefSURme

Before going on with the construction of a deformation of one dimensional split extensions of Heisenberg algebras, we have to recall a result on deformation in $\mathrm{SU}(1,n)$. The product on the extension of Heisenberg algebra will be nothing else than a transport of this one.

The article [1039] provides a formal universal deformation formula for the actions of the Iwasawa component $\mathrm{R}_0 := \mathrm{A}_0\mathrm{N}_0$ of $\mathrm{SU}(1,n)$ under an oscillatory integral form. It turns out (see [1038]) that this deformation formula is in fact non-formal for proper actions on topological spaces.

Here is the precise result. The Iwasawa decomposition of $\mathrm{SU}(1,n)$ induces the identification $\mathrm{R}_0 = \mathrm{SU}(1,n)/U(n)$. The group R_0 is endowed with a (family of) left invariant symplectic structure(s) [5] ω. If we denote by $\mathcal{R}_0 = \mathcal{A}_0 \oplus \mathcal{N}_0$ the Lie algebra of R_0, the map

$$\begin{aligned} \phi_0 \colon \mathcal{R}_0 &\to \mathrm{R}_0 \\ (a,n) &\mapsto \exp(a)\exp(n) \end{aligned} \tag{83.93}$$

DARBOUX

reveals to be a global Darboux chart for (R_0, ω). The nilpotent component appears to accept a decomposition $\mathcal{N}_0 = V \times \mathbb{R}Z$ in which the Lie bracket reads

$$[(x,z)\,,\,(x',z')] = \Omega_V(x,x')\,Z;$$

the full Iwasawa component is now parametrized by $\mathcal{R}_0 = \{(a,v,z)\,|\,, a,z \in \mathbb{R}; x \in V\}$. The interest of this situation resides in the fact that the algebra $\mathcal{R}_0$ turns out to be a one dimensional split extension of an Heisenberg algebra; namely,

$$\mathcal{R}_0 = \mathcal{F}(\mathbb{1}, 0, 2).$$

The deformation result is the following.

ThoDefHeizsansB

Theorem 83.24.
For every non-zero $\theta \in \mathbb{R}$, there exists a Fréchet function space $\mathscr{E}_\theta$ satisfying the inclusions $C_c^\infty(\mathrm{R}_0) \subset \mathscr{E}_\theta \subset C^\infty(\mathrm{R}_0)$, such that, defining for all $u, v \in C_c^\infty(\mathrm{R}_0)$

$$\begin{aligned} (u \star_\theta v)(a_0,x_0,z_0) := \frac{1}{\theta^{\dim \mathrm{R}_0}} \int_{\mathrm{R}_0 \times \mathrm{R}_0} & \cosh(2(a_1-a_2)) \\ & [\cosh(a_2-a_0)\cosh(a_0-a_1)\,]^{\dim \mathrm{R}_0 - 2} \\ & \times \exp\Big(\frac{2i}{\theta}\varphi(r_0,r_1,r_2)\Big) \\ & \times u(a_1,x_1,z_1)\,v(a_2,x_2,z_2)\,da\,dx\,dz; \end{aligned} \tag{83.94}$$

PRODUCT

where

$$\begin{aligned} \varphi(r_0,r_1,r_2) =& S_V\big(\cosh(a_1-a_2)x_0, \cosh(a_2-a_0)x_1, \cosh(a_0-a_1)x_2\big) \\ & - \bigoplus_{0,1,2} \sinh(2(a_0-a_1))z_2 \end{aligned}$$

with $S_V(x_0,x_1,x_2) := \Omega_V(x_0,x_1) + \Omega_V(x_1,x_2) + \Omega_V(x_2,x_0)$ is the phase for the Weyl product on $C_c^\infty(V)$ and $\bigoplus_{0,1,2}$ stands for cyclic summation, one has:

5. This is done using the hermitian symmetric structure, cf proposition 1.1 in [1039].

tBMi

(1) *$u \star_\theta v$ is smooth and the map $C^\infty_c(\mathrm{R}_0) \times C^\infty_c(\mathrm{R}_0) \to C^\infty(\mathrm{R}_0)$ extends to an associative product on $\mathscr{E}_\theta$. The pair $(\mathscr{E}_\theta, \star_\theta)$ is a (pre-$C^\star$) Fréchet algebra.*

tBMii

(2) *In coordinates (a, x, z) the group multiplication law reads*

$$L_{(a,x,z)}(a', x', z') = \left(a + a', e^{-a'}x + x', e^{-2a'}z + z' + \frac{1}{2}\Omega_V(x, x')e^{-a'}\right).$$

The phase and amplitude occurring in formula (83.94) are both invariant under the left action $L : \mathrm{R}_0 \times \mathrm{R}_0 \to \mathrm{R}_0$.

tBMiii

(3) *Formula (83.94) admits a formal asymptotic expansion of the form:*

$$u \star_\theta v \sim uv + \frac{\theta}{2i}\{u, v\} + O(\theta^2);$$

where $\{\,,\,\}$ denotes the symplectic Poisson bracket on $C^\infty(\mathrm{R}_0)$ associated with ω. The full series yields an associative formal star product on (R_0, ω) denoted by $\tilde{\star}_\theta$.

The setting and (1) and (2) may be found in [1039], while (3) is a straightforward adaptation to R_0 of [1038].

83.5 One dimensional split extensions of Heisenberg algebras

SecExtHeiz

83.5.1 Introduction

The one dimensional extensions of Heisenberg algebras are classified by triples $(\overline{X}, \mu, d)$. The quantization in the case $(\mathrm{Id}, 0, \mu)$ reveals to be a particular case of the one studied in [1039] (see also section 83.4),while quantization of other extensions can be found using symmetries of the kernel. Here we are reporting results of [1020] and most of proofs (in particular the trick of subsection 83.5.5 which allows to extend the known product to every one dimensional split extensions) are due to Y. Voglaire. It is to be published in his future PhD thesis.

The kernel of the quantization of [1039] will be denoted by K. Then we will give a way to twist K in order to obtain a kernel K' on any extension of the form $(\overline{X}, 0, 2)$. Quantizations of other extensions can be obtained by composing with Lie group isomorphisms. The kernel for an arbitrary extension is denoted by $K_0(\overline{X}, \mu, d)$, or simply K_0 when there are no possible ambiguity.

When we will deal with the anti de Sitter situation, our starting point will be this K_0 that we will have to adapt to another symplectic form that δE^* invoking lemma 83.35.

83.5.2 General definitions

Let $\mathscr{H}_n = V \oplus \mathbb{R}E$ be the Heisenberg algebra of dimension $2n + 1$, with a natural symplectic structure defined from the Heisenberg algebra structure:

$$[v, w] = \Omega(v, w)E$$

for all v, $w \in V$. Now we consider a one dimensional algebra $\mathcal{A} = \mathbb{R}A$ generated by an element A, and we build the split extension of $\mathscr{H}_n$ by $\mathcal{A}$:

$$\mathcal{F}(\rho) = \mathcal{A} \oplus_\rho \mathscr{H}_n \tag{83.95}$$

where the split homomorphism is an action by derivation $\rho\colon \mathcal{A} \to \mathtt{Der}(\mathscr{H}_n)$. The so obtained algebra is what we call a **one dimensional extension of Heisenberg algebra**. Let us study the possibilities for $\rho(A)$. From linearity, its general form is

$$\rho(A)(v, z) = \rho(A)(v, 0) + \rho(A)(0, z) = (\overline{X}v, \mu(v)) + (zv_0, 2dz)$$

with $\overline{X} \in \mathrm{End}(V)$, $\mu \in V^*$, $v_0 \in V$ and $d \in \mathbb{R}$. Since $\mathbb{R}E = [\mathscr{H}_n, \mathscr{H}_n]$, the fact that $\rho(A)$ is a derivation of $\mathscr{H}_n$, implies that $v_0 = 0$ because

$$\rho(A)\mathbb{R}E = \rho(A)[\mathscr{H}_n, \mathscr{H}_n] = [\rho(A)\mathscr{H}_n, \mathscr{H}_n] + [\mathscr{H}_n, \rho(A)\mathscr{H}_n] \subset \mathbb{R}E. \tag{83.96}$$

Thus we have

$$\rho(A)(v, z) = (\overline{X}v, \mu(v) + 2dz). \tag{83.97}$$ EqRhoBmudz

All that is summarized on the short exact sequence

$$0 \longrightarrow \mathscr{H}_n \xrightarrow{\ i\ } \mathcal{R} \underset{i}{\overset{r}{\rightleftarrows}} \mathcal{A} \longrightarrow 0$$

where the two i are injections and r is the projection. Of course $r \circ i = \mathrm{Id}$. From commutation relations in $\mathscr{H}_n$, we easily find

$$[(v, z), (v', z')] = [v, v'] = \Omega(v, v')E.$$

Applying $\rho(A)$ to this equality, and using the fact that this is a derivation, we find

$$\Omega(\overline{X}, v')E + \Omega(v, \overline{X}')E = \rho(A)\Omega(v, v')E = 2d\Omega(v, v')E$$

which can be rewritten as

$$\Omega\big((\overline{X} - d\,\mathbb{1})v, v'\big) + \Omega\big(v, (\overline{X} - d\,\mathbb{1})v'\big) = 0. \tag{83.98}$$

In conclusion, the endomorphism $\rho(A)$ is given by a triple $(\overline{X}, \mu, 2d)$ with $(\overline{X} - d\,\mathbb{1}) \in \mathfrak{sp}(V, \Omega)$, $\mu \in V^*$ and $d \in \mathbb{R}$. Using this result, we write the general commutator on $\mathcal{R} = \mathcal{A} \oplus_\rho \mathscr{H}_n$ under the form

$$\big[(a, v, z), (a', v', z')\big] = \big(0, \overline{X}(av' - a'v), \mu(av' - a'v) + 2d(az' - a'z) + \Omega(v, v')\big) \tag{83.99}$$ EqGeneExtHeizCom

where we adopted the notation

$$(a, v, z) = aA + v + zE. \tag{83.100}$$

83.5.3 Symplectic structure

The following proposition gives a symplectic structure on $\mathcal{F}$.

Proposition 83.25.
The algebra $(\overline{X}, \mu, d)$ endowed with

$$\Omega^{\mathcal{F}} = -\delta E^* = E^*([.,.]) \tag{83.101}$$

where the star denotes the Chevalley cocycle defined by (59.33) is symplectic if and only if $d \neq 0$. PropSymplestarE

Proof. It is evident that $\Omega^{\mathcal{F}}$ is closed because it is exact. For non-degeneracy, we compute

$$\begin{aligned}\Omega^{\mathcal{F}} &= E^*[.,.] = a\mu(v') - a'\mu(v) + 2d(az' - a'z) + \Omega(v, v') \\ &= \begin{pmatrix} 0 & \mu & 2d \\ -\mu^t & \Omega & 0 \\ -2d & 0 & 0 \end{pmatrix}\end{aligned}$$

whose determinant is $\det \Omega^{\mathcal{F}} = -4d^2 \det \Omega$ which is non vanishing if and only if $d \neq 0$.

□

This symplectic algebra is denoted by $\mathcal{F}_\Omega(\overline{X}, \mu, d)$, or simply $\mathcal{F}$ when there are no possible confusions.

Since we are only interested in symplectic algebras, we suppose $d \neq 0$ and we look at extensions of type $(d\overline{X}, d\mu, 2d)$ with $\overline{X} - d\mathbb{1} \in \mathfrak{sp}(V, \Omega)$. The bracket is given by

$$\big[(a, v, z), (a', v', z')\big] = \big(0, d\overline{X}(a'v - a'v), d\mu(av' - a'v) + 2d(az' - a'z) + \Omega(v, v')\big). \quad (83.102)$$ EqCommGenF

Then we consider F, the corresponding group and the left invariant symplectic form

$$\omega_g(X_g, Y_g) = \Omega(dL_{g^{-1}}X_g, dL_{g^{-1}}Y_g).$$

By construction, $L_g^*\omega = \omega$. One can prove that the Iwasawa coordinates

$$\begin{aligned} I &: \mathcal{F} \to F \\ (a, n) &\mapsto e^{aA}e^n \end{aligned} \quad (83.103)$$

is not a Darboux chart, but one can find a twist.

Proposition 83.26.
The chart

$$\begin{aligned} \tilde{I} &: \mathcal{F} \to F \\ \tilde{I}(a, v, z) &= e^{aA}e^v e^{\left(z + \frac{1}{4}\Omega(v\overline{X})\right)E} \end{aligned} \quad (83.104)$$

is Darboux for the symplectic group (F, ω). In other words,

$$\tilde{I}^*\omega = \Omega.$$

Proof. No proof.

¡¡ Avertissement/question au lecteur !! 83.27
The proof of this is in the Yannick's master thesis. ProbMemYan

□

This chart provides an action of F on $\mathcal{F}$ by

$$\tau_g(X) = (\tilde{I}^{-1} \circ L_g \circ \tilde{I})(X)$$

for $X \in \mathcal{F}$, $g \in F$.

Proposition 83.28.
The action τ is symplectic and strongly hamiltonian. The momentum maps are EqAppMomHamActSPHm

$$\lambda_A(X) = d\mu(v) + 2dz, \quad (83.105a)$$

$$\lambda_w(X) = \mu\left(\frac{e^{-2da} - e^{da\overline{X}}}{d(2 - \overline{X})}w\right) - \Omega(v, e^{-da\overline{X}}w), \quad (83.105b)$$

$$\lambda_E(X) = e^{-2da}. \quad (83.105c)$$

Proof. The action τ is symplectic because $I^*\omega = \Omega$ and $L_g^*\omega = \omega$. □

Proposition 83.29.
The Moyal product on $(\mathcal{F}, \Omega^{\mathcal{F}})$ is $\mathcal{F}$-covariant.

Proof. Using the expansion (82.5) of the Moyal product and the explicit form (83.105) of the momentum maps, we remark that the higher order terms in $\lambda_X *_M \lambda_Y$ are vanishing. □

83.5.4 Isomorphisms

SubsecIsomsdX

The extension obtained by the derivation $D = (\overline{X}, \mu, d)$ is *a priori* not the same as the one obtained by $D' = (\overline{X}', \mu', d')$. Two extensions are isomorphic when there exists a linear bijection $dL\colon \mathcal{F}_D \to \mathcal{F}_{D'}$ such that [6]

$$dL\big([X,Y]_{D'}\big) = \big[dL(X), dL(Y)\big]_{D'} \tag{83.106a}$$

$$(dL)^*\Omega^{D'} = \Omega^D. \tag{83.106b}$$

We find the following isomorphisms: SubEqsIsommud

— $\mathcal{F}(d\overline{X}, d\mu, 2d) \simeq \mathcal{F}(\overline{X}, \mu, d)$ by

$$dL(a,v,z) = (da, v, z), \tag{83.107a}$$

— $\mathcal{F}(\overline{X}, \mu, 2) \simeq \mathcal{F}(\overline{X}, 0, 2)$ by

$$dL(a,v,z) = (a, v + au, z), \tag{83.107b}$$

where u is the vector of V satisfying $i(u)\Omega = \mu$,

— $\mathcal{F}(\overline{X}, 0, 2) \simeq \mathcal{F}(\overline{X}', 0, 2)$ by

$$dL(a,v,z) = (a, M(v), z) \tag{83.107c}$$

where $M \in \mathrm{SP}(V,\Omega)$ fulfills $M\overline{X}M^{-1} = \overline{X}'$ or, equivalently,

$$M(\overline{X} - \mathbb{1})M^{-1} = \overline{X}' - \mathbb{1}.$$

The third isomorphism only gives the equivalence between $\overline{X} - \mathbb{1}$ and $\overline{X}' - \mathbb{1}$ when they belongs to the same orbit of the adjoint action of $\mathrm{SP}(V,\Omega)$. In particular, there are no isomorphisms between the identity and anything else.

Let us prove the second isomorphism. If $D = (\overline{X}, \mu, 2)$, $D' = (\overline{X}, 0, 2)$, $Y = (a, v, z) \in \mathcal{F}$ and $Y' = (a', v', z') \in \mathcal{F}$, we have

$$\begin{aligned} dL\big([Y,Y']_D\big) &= dL\big(0, \overline{X}(av' - a'v), \mu(av' - a'v) + 2(az' - a'z) + \Omega(v,v')\big) \\ &= \big(0, \overline{X}(av' - a'v), \mu(av' - a'v) + 2(az' - a'z) + \Omega(v,v')\big), \end{aligned} \tag{83.108}$$

EqPrIsoDDtilde

while

$$\big[dL(Y), dL(Y')\big]_{D'} = \Big(0, \overline{X}\big(a(v' + a'u) - a'(v + au)\big), 2(az' + a'z) + \Omega(v + au, v' + a'u)\Big),$$

but

$$\Omega(v + au, v' + a'u) = \Omega(v,v') + \mu(av' - a'v),$$

thus we find the same as in equations (83.108).

The theorem 83.24 together with the isomorphisms given in 83.5.4 only provide a product on extensions of type $(d\mathbb{1}, 0, 2d)$. But we saw that the extensions $(\overline{X}, 0, 2d)$ with $\overline{X} \neq \mathbb{1}$ are different. Hence the generalization of this result to other extensions is not straightforward. We address now this question.

83.5.5 Extensions with non trivial $\overline{X}$

subsecTrick

The group $F(\mathbb{1}, 0, 2)$ is provided with a kernel $K\colon F \times F \times F \to \mathbb{C}$ by theorem 83.24. The symplectic group $\mathrm{SP}(V,\Omega)$ acts on F by

$$\begin{aligned} \Phi\colon\ & \mathrm{SP}(V,\Omega) \times F \to F \\ & (M, I(a,v,z)) \mapsto \Phi_M(I(a,v,z)) := I(a, M(v), z) \end{aligned} \tag{83.109}$$

where

$$\begin{aligned} I\colon\ & \mathcal{F} \to F \\ & (a,n) \mapsto e^{aA}e^n \end{aligned} \tag{83.110}$$

is the Iwasawa coordinate on F.

6. the reason why we write dL instead of L comes from the fact that we will be interested in the corresponding group isomorphism later.

Proposition 83.30.
The kernel K is invariant under this action: $\Phi_M^* K = K$. PropkernelinvarSp

Proof. We are looking on the kernel in expression (83.94). The amplitude of K, i.e. all what lies outside the exponential, and the cyclic sum in the phase only depend on the a_i's. So Φ_M does not act on them. As far as S_0 is concerned, up to coefficients which only depend on the a_i's, it is a sum of elements of the form $\Omega(Mv_i, Mv_j) = \Omega(v_i, v_j)$. □

Let $\overline{X}$ be a matrix such that $\bar{\overline{X}} = \overline{X} - \mathbb{1} \in \mathfrak{sp}(V,\Omega)$ and $\mathcal{F}' = \mathcal{F}'(\overline{X}, 0, 2)$. We consider $\mathcal{S}$, the one dimensional subalgebra of $\mathfrak{sp}(V,\Omega)$ generated by $\bar{\overline{X}}$ and we define

$$\mathcal{G} = \mathcal{S} \oplus_\rho \mathcal{F} \tag{83.111}$$

with

$$\rho(\bar{\overline{X}})(a, v, z) = (0, \bar{\overline{X}}v, 0).$$

We denote by G and S the corresponding groups. We have in particular $F \simeq G/S$. An element of $\mathcal{G}$ has the form

$$(k\bar{\overline{X}}, a, v, z) = k\bar{\overline{X}} + aA + v + zE. \tag{83.112}$$ EqDefkavzG

Proposition 83.31.
The group F' is a subgroup of G.

Proof. We will prove that $\mathcal{F}'$ is isomorphic to a subalgebra of $\mathcal{G}$, namely, the subalgebra $\mathcal{L} \subset \mathcal{S} \oplus_\rho \mathcal{F}$,

$$\mathcal{L} = \mathbb{R}(A + \bar{\overline{X}}) \oplus_\sigma (V + \mathbb{R}E)$$

where σ is the splitting homomorphism (83.97) of $\mathcal{F}$, which in the present case reads $\sigma(A + \bar{\overline{X}})(0, v, z) = (0, Xv, 2z)$. In other words, the algebra $\mathcal{L}$ is made of elements of the form (83.112) with $k = a$. The isomorphism is

$$\begin{aligned} \phi\colon \mathcal{L} &\to \mathcal{F}'(\overline{X}, 0, 2) \\ a(A + \bar{\overline{X}}) + v + zE &\mapsto aA + v + zE. \end{aligned} \tag{83.113}$$

Indeed, using formula (83.102) with $d = 1$ and $\mu = 0$, we find

$$\begin{aligned} \Big[\phi\big(a(A + \bar{\overline{X}}) + v + zE\big), \phi\big(a'(A + \bar{\overline{X}}) + v' + z'E\big)\Big]_{(\overline{X},0,2)} \\ &= \big[aA + v + zE, a'A, v' + z'E\big] \\ &= \overline{X}(av' - a'v) + \big(2(az' - a'z) + \Omega(v, v')\big)E \\ &= \phi\big(X(av' - a'v), 2(az' - a'z + \Omega(v, v'))\big) \\ &= \phi\big[a(A + \bar{\overline{X}}) + v + zE, a'(A + \bar{\overline{X}}) + v' + z'E\big]. \end{aligned}$$

□

From now on, we identify $\mathcal{F}'$ with $\mathcal{L}$ by the isomorphism ϕ which will no longer be explicitly written. Image of F' in G by the isomorphism are elements of the form

$$g' = e^{a(A+\bar{\overline{X}})} e^{v+zE}.$$

Since the elements e^A and $e^{\overline{X}}$ commute in G, we can decompose an element $\phi^{-1}(g')$ as

$$\underbrace{e^{a\bar{\overline{X}}}}_{\in S} \underbrace{e^{aA} e^{v+zE}}_{\in F}.$$

The element $a(A + \bar{\overline{X}}) + v + zE$ seen in $\mathcal{S} \oplus_\rho \mathcal{F}$ will be denoted by (a, v, z) as well

$$(a, v, z) = \phi^{-1}(aA + v + zE).$$

We consider the following coordinate on F':

$$\begin{aligned} J\colon \mathcal{F}' &\to F' \\ (a,v,z) &\mapsto e^{a(A+\bar{\bar{X}})}e^{v+zE}. \end{aligned} \tag{83.114}$$

Proposition 83.32.
The group F' is diffeomorphic to the homogeneous space $F \simeq G/S$.

Proof. We will prove that F' acts simply transitively on G/S. Let us look at

$$g' = J(a,v,z) = \underbrace{e^{a(A+\bar{\bar{X}})}}_{g'_S}\,\underbrace{e^{v+zE}}_{g'_F}. \tag{83.115}$$

Noticing that $e^{a\bar{\bar{X}}}[e] = [e^{a\bar{\bar{X}}}] = [e]$ we find

$$g'[e] = e^{a\bar{\bar{X}}}e^{aA}e^{v+zE}e^{-a\bar{\bar{X}}}e^{a\bar{\bar{X}}}[e] = \mathbf{Ad}(e^{a\bar{\bar{X}}})\big(e^{aA}e^{v+zE}\big)[e].$$

In $\mathcal{G} = \mathcal{S} \oplus_\rho \mathcal{F}$, by definition of ρ, we successively have

$$\operatorname{ad}(a\bar{\bar{X}})(a,v,z) = (0, a\bar{\bar{X}}v, 0) \tag{83.116a}$$

$$\operatorname{Ad}(e^{a\bar{\bar{X}}})(a,v,z) = (a, e^{a\bar{\bar{X}}}v, z) \tag{83.116b}$$

$$\mathbf{Ad}(e^{a\bar{\bar{X}}})\big(e^{aA}e^{v+zE}\big) = e^{aA}e^{e^{a\bar{\bar{X}}}v+zE}, \tag{83.116c}$$

thus $g'[e] = e^{aA}e^{e^{a\bar{\bar{X}}}v+zE}[e] = [I(a, e^{a\bar{\bar{X}}}v, z)]$. So, in order to get the element $[I(a,v,z)] \in G/S$, we have to act on $[e]$ with the element $g' = J(a, e^{-a\bar{\bar{X}}}v, z)$. All that proves that the map

$$\begin{aligned} H\colon F' &\to G/S \\ (a,v,z) &\mapsto \big[I(a, e^{a\bar{\bar{X}}}v, z)\big] \end{aligned} \tag{83.117}$$

is a diffeomorphism. □

The work done up to now provides a diffeomorphism

$$\begin{aligned} \varphi\colon F' &\to F \\ \varphi\big(J(a,v,z)\big) &= I(a, e^{a\bar{\bar{X}}}v, z) \end{aligned} \tag{83.118}$$

which has suitable properties listed in the proposition below.

Proposition 83.33.
This map $\varphi\colon F' \to F$ has the following properties:

(1) if $g' = J(a,v,z) = g'_S g'_F$ in the sense of decomposition (83.115),

$$\varphi \circ L_{g'} = \boldsymbol{Ad}(g'_S) \circ L_{g'_F} \circ \varphi = \Phi_{e^{a\bar{\bar{X}}}} \circ L_{g'_F} \circ \varphi = \Phi_{g'_S} \circ L_{g'_F} \circ \varphi, \tag{83.119}$$

(2) the differential fulfils

$$d(\varphi \circ J)_{(0,0,0)} = dI_{(0,0,0)}, \tag{83.120}$$

(3) if ω is the left invariant symplectic form on F and ω' the one on F', we have

$$\varphi^*\omega = \omega',$$

in other words, φ is a symplectomorphism.

Proof. The first point is a computation:

$$\begin{aligned}\varphi\big(L_{g'_Sg'_F}(g_sg_E)\big) &= \varphi\big(g'_Sg_S\,\mathbf{Ad}(g_S^{-1})(g'_F)g_F\big)\\ &= \mathbf{Ad}(g'_Sg_S)\big(\mathbf{Ad}(g_S^{-1})(g'_F)g_F\big)\\ &= \mathbf{Ad}(g'_S)\big(g'_F\,\mathbf{Ad}(g_S)(g_F)\big)\\ &= \big(\mathbf{Ad}(g'_S)\circ L_{g'_F}\big)\big(\varphi(g_Sg_F)\big).\end{aligned}$$

When $g' = g'_Sg'_F = J(a,v,z)$, we have $g'_S = \exp(a\overline{X})$ and $g'_F = I(a,v,z)$, so the result is given by

$$\mathbf{Ad}(g'_S)(g'_F) = e^{\mathrm{Ad}(a\overline{X})}I(a,v,z) = I(a,e^{a\overline{X}}v,z) = \Phi_{e^{a\overline{X}}}I(a,v,z).$$

That concludes the proof of the first point. For the second statement, we have $(\varphi\circ J)(a,v,z) = \Phi_{e^{a\overline{X}}}I(a,v,z)$, so

$$\begin{aligned}d(\varphi\circ J)_{(0,0,0)}(Y_a,Y_v,Y_z) &= \frac{d}{dt}\Big[\Phi_{e^{tY_a}}I(tY_a,tY_v,tY_z)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[I\big(tY_a,e^{tY_a\overline{X}}tY_v,tY_z\big)\Big]_{t=0}\\ &= dI_{(0,0,0)}(Y_a,Y_v,Y_z).\end{aligned}\tag{83.121}$$

For the third point, we denote by e and e' the neutral of F and F'. On the one hand,

$$(\varphi^*\omega)_{g'} = \omega_{\varphi(g')}\circ d\varphi_{g'} = \omega_e\circ d\big(L_{\varphi(g')^{-1}}\circ\varphi\big)_{g'};$$

on the other hand, $\omega'_{g'} = \omega'_{e'}\circ d\big(L_{(g')^{-1}}\big)_{g'}$. Hence, in order to have $\varphi^*\omega = \omega'$, it is necessary that

$$\omega'_{e'}\circ dJ_{(0,0,0)} = \omega_e\circ d\big(L_{\varphi(g')^{-1}}\circ\varphi\circ L_{g'}\big)_{e'}\circ dJ_{(0,0,0)}.$$

But, for $g' = g'_Sg'_F$, we have

$$\begin{aligned}L_{\varphi(g')^{-1}}\circ\varphi\circ L_{g'}(g) &= \varphi(g')^{-1}\varphi(g'g)\\ &= \varphi(g')^{-1}\,\mathbf{Ad}(g'_S)\big(g'_F\varphi(g)\big)\\ &= \mathbf{Ad}\big((g'_F)^{-1}\big)\,\mathbf{Ad}(g'_S)\big(g'_F\varphi(g)\big)\\ &= \big(\mathbf{Ad}(g'_S)\circ\varphi\big)(g).\end{aligned}$$

The first property yields

$$d\big(L_{\varphi(g')^{-1}}\circ\varphi\circ J\big)_{(0,0,0)} = \mathrm{Ad}(g'_S)\circ dI_{(0,0,0)} = d(\Phi_{e^{a\overline{X}}})_e\circ dI_{(0,0,0)}.$$

Since ω_e is invariant under $\Phi_{e^{a\overline{X}}}$, it remains to be proved that $\omega'_{e'}\circ dJ_{(0,0,0)} = \omega_e\circ dI_{(0,0,0)}$. This is true because, in these coordinates, both sides applied on vectors (Y_a,Y_v,Y_z) and (Z_a,Z_v,Z_z) give

$$2(Y_aZ_z - Z_aY_z) + \Omega(Y_v,Z_v),$$

so φ is a symplectomorphism.

□

Now, if K is the kernel on F, we define the kernel on F' by

$$\begin{aligned}K'\colon F'\times F'\times F' &\to \mathbb{C}\\ K' &= \varphi^*K.\end{aligned}\tag{83.122}$$

EqKerRprime

Theorem 83.34.
The kernel K' is

- *left invariant under F',*
- *associative on F'.*

ThoDefoHeizAvecB

Proof. For left invariance, let $g' = J(a,v,z)$. We have

$$L^*_{g'}K' = \big(\varphi\circ L_{J(a,v,z)}\big)^*K = \big(\Phi_{e^{a\overline{X}}}\circ L_{g'_F}\circ\varphi\big)^*K = \varphi^*L^*_{g'_F}\Phi^*_{e^{a\overline{X}}}K = K',$$

because of left invariance of K under F and its invariance under Φ. Associativity can be checked in much the same way as in lemma 83.4. □

83.5.6 Get the μ back

SubSecRemetreMu

Equation (83.122) provides a kernel on F', the group generated by the algebra

$$\mathcal{F}' = \mathbb{R}A \oplus_{(\overline{X},0,2)} \mathscr{H}_n.$$

Now we try to find a kernel for the group corresponding to the algebra

$$\mathcal{F}_1 = \mathbb{R}A \oplus_{(d\overline{X},d\mu,2d)} \mathscr{H}_n.$$

The algebra isomorphism is know from equations (83.107), it is

$$\begin{aligned} dL\colon \mathbb{R}A \oplus_{(d\overline{X},d\mu,2d)} \mathscr{H}_n &\to \mathbb{R}A \oplus_{(\overline{X},0,2)} \mathscr{H}_n \\ dL(aA + v + zE) &= daA + (v + au) + zE \end{aligned} \tag{83.123}$$

where u is characterised by

$$i(u)\Omega = \mu.$$

The problem is to find the group isomorphism L. Since L must be an isomorphism, we have

$$L(e^{aA}e^{v+zE}) = L(e^{aA})L(e^{v+zE}) = e^{daA+au}e^{v+zE}$$

that we want to write under the form $e^{a'A}e^{v'+z'E}$. So we write

$$e^{daA+au} = e^{daA}\underbrace{e^{-daA}e^{daA+au}}_{\text{to be traeated}}.$$

Let $F = F(\mathbb{1},0,2)$ and $F' = F'(\overline{X},0,2)$. By proposition 83.30, the kernel K on F is invariant under $\mathrm{SP}(V,\Omega)$, i.e. $\Phi_M^* K = K$ for all $M \in \mathrm{SP}(V,\Omega)$. The action of $\mathrm{SP}(V,\Omega)$ on F is given by

$$\Phi_M\big(I(a,v,z)\big) = I(a,Mv,z).$$

Define the map $\Phi'_M\colon F' \to F'$,

$$\Phi'_M\big(J(a,v,z)\big) = J\big(a, e^{-a\bar{\overline{X}}}Me^{a\bar{\overline{X}}}v, z\big) \tag{83.124}$$

EqDefPhiprimeM

which fulfils

$$\Phi_M \circ \varphi = \varphi \circ \Phi'_M.$$

Thus, using the $\mathrm{SP}(V,\Omega)$-invariance of K, we have

$$\phi'^*_M K' = (\varphi \circ \phi'_M)^* K = (\phi_M \circ \phi)^* K = \varphi^* K = K'.$$

This proves that K' is also invariant under $\mathrm{SP}(V,\Omega)$ too.

83.5.7 Jump from one kernel to another

We know the deformation of $\mathrm{SU}(1,n)$ described in section 83.4.

We have a kernel for the extensions $F_{\delta E^*}(d\mathbb{1},0,2d)$ and $F_{\delta E^*}(\overline{X},0,2)$. We can consider the isomorphism $L\colon F(\overline{X},0,2) \to F(d\overline{X},d\mu,2d)$ which is the lift of

$$\begin{aligned} dL\colon \mathcal{F}(\overline{X},0,2) &\to \mathcal{F}(d\overline{X},d\mu,2d) \\ (a,v,z) &\mapsto (da, v+au, z). \end{aligned} \tag{83.125}$$

If K' is a kernel on $F_{\delta E^*}(d\overline{X},d\mu,2d)$, then

$$K_0 = L^* K'$$

is a kernel on $F_{\delta E^*}(\overline{X},0,2)$.

An action $\Phi_0(M)\colon F(d\overline{X},d\mu,2d) \to F(d\overline{X},d\mu,2d)$ is given by

$$\Phi_0(M) = L^{-1} \circ \Phi'(M) \circ L \tag{83.126}$$

where $\Phi'(M)\colon F(\overline{X},0,2) \to F(\overline{X},0,2)$ is given by equation (83.124). By lemma 83.4, the kernel K_0 is left invariant under the action of F and invariant under the following action of $\mathrm{SP}(V,\Omega)$:

$$\Phi_0(M)^* K_0 = K_0.$$

LemJumpCoadOrb

Lemma 83.35.
Let $\delta\eta^*$ and $\delta\xi^*$ be two exact forms on $\mathcal{F}$ such that ξ^* and η^* belong to the same coadjoint orbit: there exists a $g \in F$ such that

$$\xi^* \circ \mathrm{Ad}(g) = \eta^*. \tag{83.127}$$

A solution of the problem to find an automorphism $\sigma\colon F \to F$ such that

EqHypLemSigmaAD

$$\delta\eta^*_{\sigma(h)}\big(d\sigma_h X_h, d\sigma_h Y_h\big) = \delta\xi^*_h(X_h, Y_h) \tag{83.128}$$

for all $h \in F$ and X_h, $Y_h \in T_hF$ is given by $\sigma = \boldsymbol{Ad}(g^{-1})$.

Proof. Transported to the identity, the condition (83.128) becomes:

$$\begin{aligned}\delta\eta^*\big(dL_{\sigma(h)^{-1}}d\sigma_h X_h, dL_{\sigma(h)^{-1}}d\sigma_h Y_h\big) &\stackrel{!}{=} \delta\xi^*\big(dL_{h^{-1}}X_h, dL_{h^{-1}}Y_h\big)\\ &= \delta\eta^*\big(\mathrm{Ad}(g^{-1})dL_{h^{-1}}X_h, \mathrm{Ad}(g^{-1})dL_{h^{-1}}Y_h\big).\end{aligned}$$

If $X_h = \frac{d}{dt}X_h(t)\big|_{t=0}$, we are searching for a σ such that

$$\frac{d}{dt}\Big[\sigma(h)^{-1}\sigma\big(X_h(t)\big)\Big]_{t=0} = \frac{d}{dt}\Big[\,\mathbf{Ad}(g^{-1})\big(h^{-1}X_h(t)\big)\Big]_{t=0}.$$

Since σ is a group isomorphism, $\sigma(h)^{-1} = \sigma(h^{-1})$ and the constraint on σ becomes

$$\sigma\big(h^{-1}X_h(t)\big) = g^{-1}\big(h^{-1}X_h(t)\big)g.$$

A solution is therefore

$$\sigma = \mathbf{Ad}(g^{-1}). \tag{83.129}$$

□

Chapter 84

Deformation of anti de Sitter spaces

ChDefoBH

Abstract

We are now going to apply deformation theory to the physical part of the AdS_4 black hole. The first idea was to deform the AN of $\mathrm{SO}(2,l-1)$ and to deform AdS_l by action of this group (see section 82.6 for an introduction to deformation by group action). We show in section 85.2 that this procedure is possible.

Instead of that, we will only deform an open orbit of AN in the four dimensional case [1]. There are two reasons for that. First a physical domain of the black hole is contained in an open orbit of AN; and second it reveals possible to deform such a domain by action of a four dimensional group. Deforming a four dimensional space by a four dimensional group instead of a six dimensional one is a matter of "no waste" of dimensions.

The main lines of the construction are the following:

— We pick an open orbit $\mathcal{U}$ of AN in AdS_4, and we select a point $[u] \in \mathcal{U}$.

— We compute the stabilizer S of $[u]$ in AN, in such a way that, as homogeneous space, $\mathcal{U} = AN/S$. We consider the "remaining group" R' of R when one removes S from R.

— We prove that R' acts freely and transitively on $\mathcal{U}$, so that $\mathcal{U}$ is globally of group type (definition 84.1).

— It turns out that R' does not accept any symplectic structure; hence we will search for other groups acting transitively on $\mathcal{U}$, and show that one and only one of them accepts a symplectic structure.

— The latter group turns out to be a split extension of an Heisenberg group (see appendix 83.5) for which we know a deformation.

This part is published in [1020].

84.1 Group structure on the open orbits

SecGpStructOuvertOrb

84.1.1 Global structure

The following definition formalises the idea for a manifold to be "like a group". We will prove that the physical domain of our anti de Sitter black hole is of this type.

Definition 84.1.
*A m-dimensional homogeneous space M is **locally of (symplectic) group type** if there exists a Lie (symplectic) subgroup R of the group of automorphisms of M which acts freely on one of its orbits in M. The homogeneous space is **globally of group type** if R has only one orbit. In this case, for every choice of a base point ϑ in M, the map $R \to M : g \mapsto g.\vartheta$ is a diffeomorphism.* DefGlobGpType

Lie groups are themselves examples of symmetric spaces (globally) of group type. In the symplectic situation, however, a symplectic symmetric Lie group must be abelian ([1040], page 12). We will see in what follows other non-abelian examples.

1. But the structure of algebra (66.125), promises easy higher dimension generalisation.

In the context of our anti de Sitter black hole (in particular when one has causal issues in mind), it is not important to deform the whole space but it is sufficient [2] to only deform one open orbit of AN. Indeed, if an observer begins his life somewhere in the physical space (hence in an open orbit of AN), he will never exit the orbit because one open orbit of AN is bounded by closed orbit of AN which are singular.

Let us recall that the solvable part of the Iwasawa decomposition of $\mathfrak{so}(2,3)$ may be realized with a nilpotent part $\mathcal{N}$ and an abelian one $\mathcal{A}$ with elements

$$\mathcal{A} = \{J_1, J_2\} \qquad \mathcal{N} = \{W, V, M, L\} \tag{84.1a}$$

and the commutator table [EqTableRappelSO]

$$[V, W] = M \qquad [V, L] = 2W \tag{84.2a}$$
$$[J_1, W] = W \qquad [J_2, V] = V \tag{84.2b}$$
$$[J_1, L] = L \qquad [J_2, L] = -L \tag{84.2c}$$
$$[J_1, M] = M \qquad [J_2, M] = M, \tag{84.2d}$$

where we know that that W, $J_1 \in \mathcal{H}$, and $J_2 \in \mathcal{Q}$. We pick the point

$$[u] = \left[\begin{pmatrix} 0 & 1 & \\ -1 & 0 & \\ & & \mathbb{1}_{3\times 3} \end{pmatrix} \right].$$

This is an element of K which, as already mentioned in page 3963, therefore does not belong to a closed orbit of AN, neither to a one of $A\bar{N}$. Hence $[u]$ lies in the physical part of AdS_4. We denote by $\mathcal{U}$ the AN-orbit of $[u]$.

Elements of the stabilizer of $[u]$ in $\mathrm{SO}(2,3)$ are elements r such that $r \cdot [u] = [u]$, i.e. elements for which there exists h in H such that $ru = uh$. It is easy to see that J_2 once again belongs to the Lie algebra of the stabilizer. The new element for AdS_4 is V. Indeed $V^3 = 0$, so the exponential is a three term sum:

$$e^{aV} = \begin{pmatrix} b+1 & 0 & -b & 0 & a \\ 0 & 1 & 0 & 0 & 0 \\ b & 0 & 1-b & 0 & a \\ 0 & 0 & 0 & 1 & 0 \\ a & 0 & -a & 0 & 1 \end{pmatrix}$$

where $b = a^2/2$. We hope that $[e^{aV}u] = [u]$, i.e. that

$$e^{aV}u = uh$$

for a certain $h \in H$. A direct computation of $h = u^{-1}e^{aV}u$ gives

$$h = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & b+1 & -b & 0 & a \\ 0 & b & 1-b & 0 & a \\ 0 & 0 & 0 & 1 & 0 \\ 0 & a & -a & 0 & 1 \end{pmatrix}$$

which is a matrix of H because it leaves unchanged the first basis vector.

The stabilizer cannot contain more than two generators because an open orbit must be four dimensional. The stabilizer of $[u]$ in $G = \mathrm{SO}(2,3)$ is thus the group generated by e^{aJ_2} and e^{bV} plus eventually a discrete set making S non connected. The group S is in fact connected because

$$S = \{r \in R \text{ tel que } r \cdot [u] = [u]\} = \{r \in R \text{ tel que } \mathbf{Ad}(u^{-1})r \in H\}. \tag{84.3}$$

2. We will however point out in the perspectives (section 79.18) that a quantization of the whole space has a real interest.

Since R is an exponential group, we have $S = \exp \mathcal{S}$ where

$$\mathcal{S} = \{X \in \mathcal{R} \text{ tel que } \mathrm{Ad}(u^{-1})X \in \mathcal{H}\} = \mathrm{Ad}(u)\mathcal{H}.$$

The set $\mathcal{S}$ being connected (because it is the image by a continuous map of the connected set $\mathcal{H}$), S is connected too.

The open orbit that we are studying is thus realised as the homogeneous space $\mathcal{U} = AN/S$. An important result is the fact that what we obtain by simply removing J_2 and V from the table (84.2) is still an algebra. The orbit $\mathcal{U}$ is therefore isomorphic to the group R' generated by the Lie algebra $\mathcal{R}' = \{J_1, W, M, L\}$. The table of $\mathcal{R}'$ is

$$[J_1, W] = -W \tag{84.4a}$$
$$[J_1, L] = L \tag{84.4b}$$
$$[J_1, M] = M. \tag{84.4c}$$

From construction, $R' \cap S = \{e\}$. Unfortunately, using the conditions (59.4), we find that in order to be compatible with the Lie algebra structure, the form ω of the algebra must satisfy $\omega(W, M) = \omega(W, L) = \omega(M, L) = 0$, so that it is degenerate. The action of R' on $\mathcal{U}$ enjoys however some remarkable properties.

Proposition 84.2.
The action

$$\begin{aligned} \tau\colon R' \times \mathcal{U} &\to \mathcal{U} \\ r[r_0 u] &= [r r_0 u] \end{aligned} \tag{84.5}$$

is free and simply transitive. PropURsimptra

Proof. First, we prove that the action of R' is transitive. As an algebra, $\mathcal{R}$ is a split extension $\mathcal{R} = \mathcal{S} \oplus_{\mathrm{ad}} \mathcal{R}'$. Hence, as group, $R = SR'$, or equivalently $R = R'S$. That proves that the action is transitive.

If the action is not simply transitive, there exists $x \in \mathcal{U}$ and r, $r' \in R'$ such that $\tau_r x = \tau_{r'} x$. Since the action of R' is transitive, we have a $r_1 \in R'$ such that $x = r_1[u]$. In this case, the element $r_1^{-1} r^{-1} r'$ of R' fixes $[u]$, but $R' \cap S = \{e\}$. Then one deduces that $r' = r r_1$, so that $[r r_1 u] = [r' r_1 u] = [r r_1^2 u]$. It follows that r_1 fixes $[u]$, and thus that $r_1 = e$, so that $r = r'$.

For freeness remark that, in a neighbourhood of e, the neutral e itself is the only element trivially acting on $[u]$.

□

As corollary, the orbit $\mathcal{U}$ is locally of group type R'.

PropmUsimpl

Proposition 84.3.
The orbit $\mathcal{U}$ is simply connected.

Proof. The fibration $S \to R \to \mathcal{U}$ induces the long exact sequence of cohomology groups

$$H^0(\mathcal{U}) \to H^0(R) \to H^0(S) \to H^1(\mathcal{U}) \to H^1(R) \to \dots$$

The group R being connected and simply connected, the sequence shows that $H^0(S) \simeq H^1(\mathcal{U})$, but we already mentioned that S is connected, so $H^1(\mathcal{U}) = 0$. □

¡¡ Avertissement/question au lecteur !! 84.4
Il faut une citation pour le coup de la suite qui découle de la fibration. ProbFibra

Corollary 84.5.
The open orbit $\mathcal{U}$ is globally of group type. CormUgloGppasSym

Proof. It is immediately apparent from proof of proposition 84.2. Since

$$R'[u] = R'S[u] = R[u] = \mathcal{U},$$

the group R' acts freely on $\mathcal{U}$ and has only one orbit. □

Remark that it remains to be proved that $\mathcal{U}$ is globally of *symplectic* group type. For that, there should be a symplectic form on R'. Exploiting the fact that $\mathrm{Span}\{W, M, L\}$ is a three-dimensional abelian subalgebra of $\mathcal{R}'$, it is easy to see that $\mathcal{R}'$ does not accept a symplectic form. Hence corollary 84.5 does not prove that $\mathcal{U}$ is globally of symplectic group type. The lack of symplectic form on the algebra reflects on $\mathcal{U}$ as manifold by the following lemma, and motivates the search for other four-dimensional groups than R' acting transitively on $\mathcal{U}$.

Lemma 84.6.
The open orbit $\mathcal{U} = R \cdot [u]$ does not admit any R-invariant symplectic form.

Proof. Let $\omega^{\mathcal{U}}$ be such an invariant symplectic form and $\omega^{R'}$ be the pull-back of $\omega^{\mathcal{U}}$ by the action: $\omega^{R'} = \tau^*\omega^{\mathcal{U}}$. We have $d\tau \circ dL_{r'} = dL_{r'} \circ d\tau$ because $\tau(r'X(t)) = [r'X(t)u] = r'\tau(X(t))$, thus

$$L_{r'}^*\omega^{R'} = (\tau \circ L_{r'})^*\omega^{\mathcal{U}} = (L_{r'} \circ \tau)^*\omega(\mathcal{U}) = \omega^{R'},$$

so that $\omega^{R'}$ is a R'-invariant symplectic form on R'. But we saw that such a form does not exist. □

Proposition 84.7.
The R-homogeneous space $\mathcal{U}$ admits a unique structure of globally group type symplectic symmetric space. The latter is isomorphic to $(\mathrm{R}_0, \omega, s)$ described in appendix 83.4. GT

The next few pages are dedicated to prove this proposition and to give explicit algebra whose group gives the answer. We are searching for 4-dimensional groups $\tilde{R}$ which

— has a free and simply transitive action on $\mathcal{U}$, i.e. $\tilde{R}[u] = R[u]$,

— admits a symplectic structure,

and we want it to be unique. As already mentioned, the algebra $\mathcal{R}'$ fails to fulfil the symplectic condition. The algebra $\tilde{\mathcal{R}} = \mathrm{Span}\{A, B, C, D\}$ of a group which fulfils the first condition must at least act transitively on a small neighbourhood of $[u]$ and thus be of the form EqGeneAlgabcd

$$A = J_1 + aJ_2 + a'V \tag{84.6a}$$
$$B = W + bJ_2 + b'V \tag{84.6b}$$
$$C = M + cJ_2 + c'V \tag{84.6c}$$
$$D = L + dJ_2 + d'V. \tag{84.6d}$$

Indeed, in a first attempt, we choose an algebra for which each of A, B, C and D contains a combination of J_1, W, M and L. We consider the matrix of coefficients of J_1, W, M and L in A, B, C and D. If the determinant of this matrix is zero, then one of the lines can be written as combination of the three others. In this case the action can even not be locally transitive because the algebra only spans three directions actually acting (J_2 and V have no importance here). So the determinant is non vanishing. In this case, the inverse of this matrix is a change of basis which puts A, B, C and D under the form (84.6).

The problem is now to fix the parameters $a, a', b, b', c, c', d, d'$ in such a way that the space $\mathrm{Span}\{A, B, C, D\}$ becomes a Lie algebra (i.e. it closes under the Lie bracket) which admits a symplectic structure and whose group acts transitively on $\mathcal{U}$. We will begin by proving that the surjectivity condition imposes $b = c = d = 0$. Then the remaining conditions for $\tilde{\mathcal{R}}$ to be an algebra are easy to solve by hand.

First, remark that A acts on the algebra $\mathrm{Span}\{B, C, D\}$ because J_1 does not appear in $[\mathcal{R}, \mathcal{R}]$. Hence we can write $\tilde{\mathcal{R}} = \mathbb{R}A \oplus_{\mathrm{ad}} \mathrm{Span}\{B, C, D\}$ and, a subalgebra of a solvable exponential Lie algebra being a solvable exponential algebra, a general element of the group $\tilde{R}$ reads $\tilde{r}(\alpha, \beta, \gamma, \delta) =$

$e^{\alpha A}e^{\beta B+\gamma C+\delta D}$. Our strategy will be to split this expression in order to get a product SR' (which is equivalent to a product $R'S$). As Lie algebra, $\mathrm{Span}\{B,C,D\} \subset \mathbb{R}J_2 \oplus_{\mathrm{ad}} \{W,M,L,V\}$. Hence there exist functions w, m, l, v and x of $(\alpha,\beta,\gamma,\delta)$ such that

$$e^{\beta B+\gamma C+\delta D} = e^{xJ_2}e^{wW+mM+lL+vV}. \tag{84.7}$$

EqGeneRi

We are now going to determine $l(\alpha,\beta,\gamma,\delta)$ and study the conditions needed in order for l to be surjective on $\mathbb{R}$. Since J_2 does not appear in any commutator, the Campbell-Baker-Hausdorff formula yields $x = \beta b + \gamma c + \delta d$. From the fact that $[J_2, L] = -L$, we see that the coefficient of L in the left hand side of (84.7) is $-l(1-e^{-x})/x$. The V-component in the exponential can also get out without changing the coefficient of L. We are left with $\tilde{r}(\alpha,\beta,\gamma,\delta) = e^{\alpha A}e^{xJ_2}e^{yV}e^{w'W+m'M+lL}$ where w' and m' are complicated functions of (β,γ,δ) and l is given by

$$l(\beta,\gamma,\delta) = \frac{-\delta(\beta b+\gamma c+\delta d)}{1-e^{-\beta b-\gamma c-\delta d}}, \tag{84.8}$$

which is only surjective when $b = c = d = 0$. Taking the inverse, a general element of $\tilde{R}[u]$ reads $\left[e^{-wW-mM-lM}e^{j_1J_1}u\right]$, where the range of l is not the whole $\mathbb{R}$. Since the action of R' is *simply* transitive, $\tilde{R}$ is not surjective on $R[u]$ when $l(\alpha,\gamma,\delta)$ is not surjective on $\mathbb{R}$.

When $b = c = d = 0$, the conditions for (84.6) to be an algebra are easy to solve, leaving only two *a priori* possible two-parameter families of algebras.

(1) The first one is the following:

PgAlgUn

$$\begin{aligned} A &= J_1 + \frac{1}{2}J_2 + sV & [A,B] &= B + sC \\ B &= W & [A,C] &= \frac{3}{2}C \\ C &= M & [A,D] &= 2sB + \frac{1}{2}D \\ D &= L + rV & [B,D] &= -rC. \end{aligned}$$

with $r \neq 0$. The general symplectic form on that algebra is given by

$$\omega_1 = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & \frac{2\beta r}{3} \\ \beta & 0 & 0 & 0 \\ \gamma & -\frac{2\beta r}{3} & 0 & 0 \end{pmatrix}, \tag{84.10}$$

$$\det\omega = \left(\frac{2\beta r}{3}\right)^2,$$

Conditions: $\beta \neq 0$, $r \neq 0$. That algebra will be denoted by $\mathcal{R}_1$. The analytic subgroup of R whose Lie algebra is $\mathcal{R}_1$ is denoted by R_1. One can eliminate the two parameters in algebra $\mathcal{R}_1$ by the isomorphism

$$\phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4s \\ 0 & 2sr & 1/r & 4s^2/r \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{84.11}$$

EqIsomRUnrs

which fixes $s = 0$ and $r = 1$. The algebra $\mathcal{R}_1$ is thus isomorphic to

$$\begin{aligned} [A,B] &= B & [A,C] &= \frac{3}{2}C \\ [A,D] &= \frac{1}{2}D & [B,D] &= -C. \end{aligned} \tag{84.12}$$

Comparing with equation (83.97), one recognizes the one-dimensional extension of Heisenberg algebra with parameters $d = 3/4$, $\mu = 0$ and $\overline{X} = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$. Hence R_1 is isomorphic to R_0 and, by the way, we have a product on that group (see appendix 83.5).

(2) The second algebra whose group acts simply transitively on $\mathcal{U}$ is:

$$\begin{aligned} A &= J_1 + rJ_2 + sV & [A,B] &= B + sC \\ B &= W & [A,C] &= (r+1)C \\ C &= M & [A,D] &= 2sB + (1-r)D. \\ D &= L \end{aligned}$$

There is no way to get a nondegenerate symplectic form on that algebra.

From proposition 83.25, the symplectic structure to be chosen on $\mathcal{R}_1$ is δC^* and lemma 83.35 shows that we are able to quantize[3] $\mathcal{R}_1$ with any symplectic form in the coadjoint orbit $\delta\big(C^* \circ \mathrm{Ad}(g)\big)$ with $g \in R_1$. The coadjoint adjoint action of R_1 on $\mathcal{R}_1$ can be computed using the fact that $\mathcal{R}_1$ splits into four parts; the non trivial results are

$$\begin{aligned} \mathrm{Ad}(e^{dD})A &= A - \frac{d}{2}D & \mathrm{Ad}(e^{aA})B &= e^a B \\ \mathrm{Ad}(e^{cC})A &= A - \frac{3c}{2}C & \mathrm{Ad}(e^{aA})C &= e^{3a/2}C \\ \mathrm{Ad}(e^{bB})D &= D - bC & \mathrm{Ad}(e^{aA})D &= e^{a/2}D \\ \mathrm{Ad}(e^{bB})D &= D - bC. \end{aligned}$$

Direct computations show that

$$\begin{aligned} \mathrm{Ad}\left(e^{aA}e^{bB}e^{cC}e^{dD}\right)&(x_A A + x_B B + x_C C + x_D D) \\ &= x_A A + e^a(x_B - x_A b)B \\ &\quad + e^{3a/2}\Big(x_C - \frac{3x_A c}{2} - bx_D + \frac{bdx_A}{2}\Big)C \\ &\quad + e^{a/2}\Big(x_D - \frac{dx_A}{2}\Big)D, \end{aligned} \tag{84.13}$$

so that, with more compact notations,

$$\big(C^* \circ \mathrm{Ad}(g)\big)(X) = \big(x_C - \frac{3x_A c - bx_D + \frac{bdx_A}{2}}{2}\big)e^{3a/2}, \tag{84.14}$$

and the symplectic forms that we are able to deform are given by $\delta\big(C^* \circ \mathrm{Ad}(g)\big)$. It provides a two-parameter familly of symplectic forms

$$\omega_1^g = \begin{pmatrix} 0 & 0 & \beta & \gamma \\ 0 & 0 & 0 & -2\beta/3 \\ -\beta & 0 & 0 & 0 \\ -\gamma & 2\beta/3 & 0 & 0 \end{pmatrix}, \tag{84.15}$$

$$\det \omega_1^g = \frac{4\beta^4}{9}.$$

It turns out that the action of the group R_1 has good properties that are given in the following proposition.

Proposition 84.8.
The action of R_1 on $\mathcal{U}$ is free and simply transitive. PropCRunXXX

Proof. First remark that the algebra $\mathcal{R}_1$ can be written as a split extension:

$$\mathcal{R}_1 = \mathbb{R}A \oplus_{\mathrm{ad}} \mathbb{R}D \oplus_{\mathrm{ad}} \mathrm{Span}\{B, C\},$$

3. by opposition to *deform*: there are no symplectic condition in deformation.

hence a general element of R_1 reads

$$r_1(a,b,c,d) = e^{aA}e^{dD}e^{bW}e^{cM}. \tag{84.16}$$

The work is now to expand it by replacing A, B, C, D by their values in function of J_1, W, M, L, J_2 and V, and then to try to put all elements of $\mathcal{S}$ on the left. This is done by virtue of Campbell-Baker-Hausdorff formula. The fact that $\mathrm{Span}\{W,M,N,V\}$ is nilpotent dramatically reduces the difficulty. We have

$$\ln(e^{drV}e^{dL}e^{wW}e^{mM}) = drV + dL + (d^2r + w)W + (\frac{1}{6}d^2r^2 + m + \frac{1}{2}drw)M.$$

One can find m and w (functions of d) such that the right hand side reduces to $drV + dL$. Hence we have, for some auxiliary functions w and m,

$$e^{drV+dL} = e^{drV}e^{dL}e^{w(d)W}e^{m(d)M}$$

and a general element of R_1 reads

$$e^{asV+\frac{a}{2}J_2}e^{drV}e^{aJ_1}e^{dL}e^{\left(w(d)+b\right)W}e^{\left(m(d)+c\right)M} = s(a,d)r'(a,b,c,d) \tag{84.17}$$

with $s \in S$ and $r' \in R'$ which defines a bijective map $r_1 \mapsto r'$ from R_1 to R'. This proves the transitivity of the action of R_1.

For freeness, just remark that in a neighbourhood of e, no element of R_1 (but e) leaves $[u]$ unchanged. □

The conclusion is that R_1 is the group $\tilde{R}$ that we were searching for and that it is unique (up to the two-parameter isomorphism (84.11)) as symplectic subgroup of AN acting transitively on $\mathcal{U}$. It concludes the proof of proposition 84.7.

84.1.2 Alternative more intrinsic proofs

subSecAltreintr

Proposition 84.9.
Let $J \in Z(K)$ whose associated conjugation coincides with the Cartan involution: $\boldsymbol{Ad}(J) = \theta$ and $u \in \mathrm{SO}(2, l-1)$ such that $u^2 = J$ and $u \in e^{\mathcal{Q}} \cap K$. Then the AN-orbit of $[u]$ is open.

¡¡ Avertissement/question au lecteur !! 84.10
Il faut prouver cela. Pour cela, je crois que ce qu'il faut faire, c'est dire que Helgason ou Loos assure l'existence d'un $J \in Z(K)$ tel que $\theta = \boldsymbol{Ad}(J)$ pour le groupe, et $\theta = \mathrm{Ad}(J)$ pour l'algebre. Ensuite, vu que les orbites fermées de AN et $A\bar{N}$ sont $\pm AN$ et $\pm A\bar{N}$, évidement un élément d'une orbite fermée ne peut pas avoir son carré dans K. ProbAdJthetaj

Proof. Let us find the Lie algebra $\mathcal{S}$ of the stabilizer S of $[u]$. First, the Cartan involution $X \mapsto -X^t$ is implemented as $\mathbf{Ad}(J)$ with

$$J = \begin{pmatrix} -\mathbb{1}_{2\times 2} & \\ & \mathbb{1}_{3\times 3} \end{pmatrix}$$

which satisfies $u^2 = J$ and $\sigma(u) = u^{-1}$ because $u \in \mathcal{Q}$. Now, $\mathbf{Ad}(u^{-1})r \in H$ if and only if $\sigma\Big(\mathbf{Ad}(u^{-1})r\Big) = \mathbf{Ad}(u^{-1})r$. Using the fact that σ is an involutive automorphism, we see that this condition is equivalent to

$$\theta\sigma r = r. \tag{84.18}$$

On the one hand the Cartan involution θ restricts on $\mathcal{A}$ to $\theta|_{\mathcal{A}} = -\mathrm{Id}$ because $\mathcal{A} \subset \mathcal{P}$; and on the other hand, $\sigma(\mathcal{A}) = \mathcal{A}$ because $J_1 \in \mathcal{H}$ and $J_2 \in \mathcal{Q}$. So σ splits $\mathcal{A}$ into two parts: $\mathcal{A} = \mathcal{A}^+ \oplus \mathcal{A}^-$ with $\mathcal{A}^+ = \mathcal{A} \cap \mathcal{H} = \mathbb{R}J_1$ and $\mathcal{A}^- = \mathcal{A} \cap \mathcal{Q} = \mathbb{R}J_2$. Let $\beta_1, \beta_2 \in \mathcal{A}^*$ be the dual basis: $\beta_i(J_j) = \delta_{ij}$. We know that $W \in \mathcal{G}_{\beta_1}$, $V \in \mathcal{G}_{\beta_2}$, $L \in \mathcal{G}_{\beta_1-\beta_2}$, and $M \in \mathcal{G}_{\beta_1+\beta_2}$. The set of simple roots is given by

$$\Delta = \{\alpha = \beta_1 - \beta_2,\ \beta = \beta_2\},$$

and the positive roots are

$$\Phi^+ = \{\alpha, \beta, \alpha+\beta, \alpha+2\beta\},$$

in terms of whose, the space $\mathcal{N}$ is given by

$$\begin{array}{ll} W \in \mathcal{G}_{\alpha+\beta} & V \in \mathcal{G}_\beta \\ L \in \mathcal{G}_\alpha & M \in \mathcal{G}_{\alpha+2\beta}. \end{array}$$

Since $(\sigma^*\beta)(h_1J_1 + h_2J_2) = \beta_2(h_1J_1 - h_2J_2) = -h_2$, we find we find

$$\begin{aligned} \sigma^*\beta &= -\beta \\ \sigma^*\alpha &= \alpha + 2\beta \\ \sigma^*(\alpha+\beta) &= \alpha + \beta \\ \sigma^*(\alpha+2\beta) &= \alpha. \end{aligned}$$

We are now able to identify the set $\mathcal{S} = \{X \in \mathcal{A}\oplus\mathcal{N} \text{ tel que } \sigma\theta X = X\}$. Let us take $X \in \mathcal{R} = \mathcal{A}\oplus\mathcal{N}$ and apply $\sigma\theta$:

$$\begin{aligned} X &= X^+ + X^- + X_\alpha + X_\beta + X_{\alpha+\beta} + X_{\alpha+2\beta}, \\ \theta X &= -X^+ - X^- + Y_{-\alpha} + Y_{-\beta} + Y_{-\beta} + Y_{-(\alpha+\beta)} + Y_{-(\alpha+2\beta)}, \\ \sigma\theta X &= -X^+ + X^- + Z_{-(\alpha+2\beta)} + Z_\beta + Z_{-(\alpha+\beta)} + Z_{-\alpha} \end{aligned} \tag{84.19}$$

where X_φ, Y_φ and Z_φ denote elements of $\mathcal{G}_\varphi$, and $X^\pm$ denote the component $\mathcal{A}^\pm$ of X.

It is immediately apparent that $\sigma\theta X^- = X^-$, so that $X^- \in \mathcal{S}$. The only other component common to X and $\sigma\theta X$ is in $\mathcal{G}_\beta$, but it is *a priori* not clear that $X_\beta = Z_\beta$. We know however that $\sigma\theta V = \alpha V$ because $\mathcal{G}_\beta$ has only one dimension. Using the fact that σ and θ are commuting involutions, it is apparent that $\alpha = \pm 1$. Decomposing V into $V = V_\mathcal{H} + V_\mathcal{Q}$, we have $\theta\sigma V = \theta(V_\mathcal{H} + V_\mathcal{Q})$ which has to be equal to V or $-V$. Thus there are only two possibilities

$$\begin{array}{lll} \theta V_\mathcal{H} = V_\mathcal{H} & \text{or} & \theta V_\mathcal{H} = -V_\mathcal{H} \\ \theta V_\mathcal{Q} = -V_\mathcal{Q} & & \theta V_\mathcal{Q} = V_\mathcal{Q}. \end{array}$$

If one compares the commutator table of $\mathrm{SO}(2,3)$ with the one of $\mathrm{SO}(2,2)$, one sees that V is not present in $\mathrm{SO}(2,2)$. Since $\mathcal{H}$ possesses every purely spatial rotation generators, the orthogonal complement $\mathcal{Q}$ contains the time-time rotation as only rotations. Other components of $\mathcal{Q}$ are boost. In particular, $V_\mathcal{Q}$ is zero or a boost generator. In the latter case, $\theta V_\mathcal{Q} = -V_\mathcal{Q}$, and the conclusion is that $\sigma\theta V = V$. In the other case, $V_\mathcal{Q} = 0$ implies that $V \in \mathcal{H}$ which is impossible because $[J_2, V] = 0$ while $J_2 \in \mathcal{Q}$ and $[\mathcal{Q}, \mathcal{H}] \subset \mathcal{Q}$.

The stabilizer S is thus generated by J_2 and $\mathcal{G}_\beta = \mathbb{R}V$, i.e.

$$\mathcal{S} = \mathrm{Span}\{J_2, V\}. \tag{84.20}$$

The stabilizer of $[u]$ being two-dimensional, the orbit of $[u]$ is four-dimensional and is then open in AdS_4.

□

Notice that in contrast to the first way to find $\mathcal{S}$, this time we have no eventually double covering problems.

Let $\tilde{R}$ be a subgroup of R whose Lie algebra is a complement of $\mathcal{S}$ in $\mathcal{R}$, i.e. $\tilde{\mathcal{R}} \oplus \mathcal{S} = \mathcal{R}$. This group does not act transitively on $\mathcal{U}$ if and only if the boundary of $\tilde{R}[u]$ is non empty. Let $x_0 = \tau_{r'_0}[u]$ belong to that boundary with $r'_0 \in R'$. On that point, the fundamental fields of $\tilde{R}$ are not surjective on the tangent space of $\mathcal{U}$:

$$\begin{aligned} \ker\left[\tilde{\mathcal{R}} \to T_{x_0}\mathcal{U}\right] &\neq \{0\} \\ Y &\mapsto Y^*_{x_0}. \end{aligned}$$

Let $Y \in \tilde{R}$ belongs to this kernel: $Y^*_{x_0} = 0$. Since the linear map $\big(d\tau_{r_0'^{-1}}\big)_{x_0}$ is nondegenerate, $Y^*_{x_0}$ vanishes if and only if $\big(d\tau_{r_0'^{-1}}\big)_{x_0}(Y^*_{x_0}) = 0$, but

$$\begin{aligned}\big(d\tau_{r_0'^{-1}}\big)_{x_0}(Y^*_{x_0}) = -\frac{d}{dt}\Big[r_0'^{-1}e^{tY}r_0'[u]\Big]_{t=0} &= \big(\operatorname{Ad}(r_0'^{-1})Y\big)^*_{[u]}\\ &= \mathtt{proj}_{\tilde{\mathcal{R}}}\big(\operatorname{Ad}(r_0'^{-1})Y\big)\end{aligned} \tag{84.21}$$

because, on the point $[u]$, the action to take the fundamental field is nothing else than the projection parallel to $\mathcal{S}$. Hence the group $\tilde{R}$ is not surjective if and only if

$$\big(\operatorname{Ad}(R')\mathcal{S}\big) \cap \tilde{\mathcal{R}} \neq \{0\}.$$

We are now going to determine $\operatorname{Ad}(R')\mathcal{S}$. Let $X = X^- + X_\beta \in \mathcal{S}$ and act with an element of $R' = \exp\big(\mathcal{A}^+ \oplus \mathcal{G}_\alpha \oplus \mathcal{G}_{\alpha+\beta} \oplus \mathcal{G}_{\alpha+2\beta}\big)$:

$$\begin{aligned}\operatorname{Ad}(e^{H^+ + Y_\alpha + Y_{\alpha+\beta} + Y_{\alpha+2\beta}})(X^- + X_\beta) = X^- + X_\beta&\\ + \underbrace{\big[H^+ + Y_\alpha + Y_{\alpha+\beta} + Y_{\alpha+2\beta}, X^- + X_\beta\big]}_{N'}&\\ + \frac{1}{2}\big[H^+ + Y_\alpha + Y_{\alpha+\beta} + Y_{\alpha+2\beta}, N'\big]&\\ + \dots&\end{aligned}$$

The computation of N' is as follows:

$$\begin{aligned}[H^+, X^-] &= 0 & [H^+, X_\beta] &= 0\\ [Y_\alpha, X^-] &= -\alpha(X^-)Y_\alpha & [Y_\alpha, X_\beta] &= Z_{\alpha+\beta}\\ [Y_{\alpha+\beta}, X^-] &= -(\alpha+\beta)(X^-)Y_{\alpha+\beta} = 0 & [Y_{\alpha+\beta}, X_\beta] &= Z_{\alpha+2\beta}\\ [Y_{\alpha+2\beta}, X^-] &= -(\alpha+2\beta)(X^-)Z'_{\alpha+2\beta} & [Y_{\alpha+2\beta}, X_\beta] &= 0,\end{aligned}$$

so $N' = -\alpha(X^-)Y_\alpha + Z_{\alpha+\beta} - Z_{\alpha+2\beta} - (\alpha+2\beta)(X^-)Z'_{\alpha+2\beta}$. Since $\beta(H^+) = 0$, the computation of $[H^+, N']$, produces terms like $[H^+, X_{\alpha+\beta}] = (\alpha+\beta)(H^+)X_{\alpha+\beta} = \alpha(H^+)X_{\alpha+\beta}$. Therefore, $[H^+, N'] = \alpha(H^+)N'$ and

$$\begin{aligned}\operatorname{Ad}(e^{H^+ + Y})(X^- + X_\beta) &= X + N' + \sum_{k\geqslant 1}\frac{1}{(k+1)!}\alpha(H^+)^k N'\\ &= X + \frac{e^{\alpha(H^+)} - 1}{\alpha(H^+)}N'\end{aligned} \tag{84.22}$$

EqElsInterditsSu

What we have proven is the following result.

Proposition 84.11.
The group $\tilde{R}$ acts transitively on $\mathcal{U}$ if and only if the Lie algebra $\tilde{\mathcal{R}}$ does not contain elements of the form

$$X + \frac{e^{\alpha(H^+} - 1}{\alpha(H^+)}N'$$

with $X \in \mathcal{S}$ and $Y \in \mathcal{R}'$; the element N' being given by

$$N' = -\alpha(X^-)Y_\alpha + Z_{\alpha+\beta} - Z_{\alpha+2\beta} - (\alpha+2\beta)(X^-)Z'_{\alpha+2\beta}$$

where $X = X^- + X_\beta$ is the decomposition of an element of $\mathcal{S}$ and Z_φ are elements of their respective root spaces $\mathcal{G}_\varphi$.

One can distinguish three case: the first is $X = X^- \in \mathcal{A}^-$, the second is $X = X_\beta \in \mathcal{G}_\beta$ and the last one is $X = X^- + X_\beta$ ($X^- \neq 0 \neq X_\beta$).

In the first case, formula (84.22) forbids $\tilde{\mathcal{R}}$ to contains elements of the form

$$X^- + \mathcal{G}_\alpha \oplus_{\alpha+2\beta} . \tag{84.23}$$

The second case forbids elements of the form

$$X_\beta + \mathcal{G}_{\alpha+\beta} \oplus \mathcal{G}_{\alpha+2\beta}, \tag{84.24}$$

and the third case forbids

$$X^- + X_\beta \oplus \mathcal{G}_\alpha \oplus \mathcal{G}_{\alpha+\beta} \oplus \mathcal{G}_{\alpha+2\beta}. \tag{84.25}$$

We can extract constraints on the coefficients of algebra (84.6) from that analysis. The third interdiction makes that a linear combination of J_2 and V in an element of $\tilde{\mathcal{R}}$ can only occur in A, so that

$$bb' = cc' = dd' = 0.$$

The second interdiction says that B and C cannot contain V alone, so $b' = c' = 0$. Finally, the first condition imposes $c = d = 0$ because C and D cannot contain J_2 alone. The remaining constraints for (84.6) to be an algebra are easy to solve by hand. The results are the same two algebras as previously found.

84.1.3 Local group structure

We saw in proposition 84.8 that the open orbit $\mathcal{U}$ can be identified with the group generated by the algebra $\{J_1, W, M, L\}$.

We want to find an algebra (whose group is) acting transitively on a neighbourhood of $[u]$ in the AN orbit of $[u]$ and which admits a symplectic form. Let A, B, C, D be a basis of this algebra. For local transitivity, each of them must contains at least one of J_1, W, M and L. As in the previous case, the most general algebra to be studied is [EqAlgGEnennsy]

$$A = J_1 + aJ_2 + a'V \tag{84.26a}$$
$$B = W + bJ_2 + b'V \tag{84.26b}$$
$$C = M + cJ_2 + c'V \tag{84.26c}$$
$$D = M + dJ_2 + d'V. \tag{84.26d}$$

Among such algebras, we will have to check surjectivity of the action and the possibility to endow with a symplectic form.

If we impose that $\mathrm{Span}\{A, B, C, D\}$ is a subalgebra for the bracket inherited from $\mathfrak{so}(2,3)$, we find a lot of conditions on the coefficients a, b, c, d, a', b', c' and d'. If, for example, we look at $[A, B]$, we find

$$[A, B] = W + a'M + ab'V - a'bV.$$

In this combination, the coefficient of W is 1 and the one of M is a', so the only possibility is $[A, B] = B + a'C$. This leads to the following conditions (equating coefficients of J_2 and V): [SubEqSystemeAlgTN]

$$b + a'c = 0 \tag{84.27a}$$
$$b' + a'c' = ab' - a'b. \tag{84.27b}$$

Proceeding in a similar way for the six different commutators, we find:

For $[A, C]$

$$ac' - a'c = (a+1)c' \tag{84.27c}$$
$$(a+1)c = 0, \tag{84.27d}$$

for $[A, D]$

$$(1-a)d + 2a'b = 0 \tag{84.27e}$$
$$ad' - a'd = 2a'b' + (1-a)d', \tag{84.27f}$$

for $[B,C]$

$$(b-c')c = 0 \tag{84.27g}$$
$$(b-c')c' = bc' - b'c, \tag{84.27h}$$

for $[B,D]$

$$-d'c + 2b'b - bd = 0 \tag{84.27i}$$
$$-d'c' - bd' + 2(b')^2 = bd' - b'd, \tag{84.27j}$$

for $[C,D]$

$$cd' = c'b' \tag{84.27k}$$
$$cd = c'b. \tag{84.27l}$$

Solutions of these equations [4], parametrized by reals r and s and the corresponding commutators are listed below.

The next step is to determine which of these algebras admit a compatible symplectic structure in the sense of definition 59.1. For this, we just have to consider a general skew-symmetric matrix

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & -\delta & -\sigma \\ \beta & \delta & 0 & -\epsilon \\ \gamma & \sigma & \epsilon & 0 \end{pmatrix}$$

and, for each algebra, solve the four constrains. In the first algebra (see below), we find for example

$$\omega_1([A,B],C) + \omega_1([B,C],A) + \omega_1([C,A],B) = -\frac{-5\omega_1(C,B)}{2} \stackrel{!}{=} 0,$$

so that $\omega_1(C,B) = 0$. Full results are listed below (the symplectic matrices are written in the basis $\{A,B,C,D\}$). We see in particular that most of the solutions reduce to the *canonical algebra*, $\mathcal{R}_c$ given by

$$[a,b] = b \qquad [a,c] = 2c \qquad [c,d] = c.$$

Let us compute the metric on these groups too. Let G_L be one of these groups (the groups corresponding to algebras one to seven, see below). We have a diffeomorphism between G_L and the orbit $\mathcal{U}$ of $[u] \in AdS_4$,

$$\begin{aligned} \phi\colon G_L &\to \mathcal{U} \\ r &\mapsto [ru]. \end{aligned} \tag{84.28}$$

We define the metric on G_L at e by

$$g_e(A,B) := B_{\phi(e)}\big(d\phi(A), d\phi(B)\big) \tag{84.29}$$

EqgeABBphi

where

$$d\phi(A) = \frac{d}{dt}\Big[\phi\big(A(t)\big)\Big]_{t=0} = \frac{d}{dt}\Big[\big(\pi \circ R_u\big)\big(A(t)\big)\Big]_{t=0} = d\pi\, dR_u A.$$

By definition of the metric on the cosets (cf equation (66.40)) we have

$$\begin{aligned} B_{[g]}(d\pi X, d\pi Y) &= B_g(\mathtt{proj}\, X, \mathtt{proj}\, Y) \\ &= B_e(dL_{g^{-1}}\, \mathtt{proj}\, X, dL_{g^{-1}}\, \mathtt{proj}\, Y) \end{aligned}$$

Using that in equation (84.29), we find (e is the neutral of G_L):

$$\begin{aligned} B_e(A,B) &= B_{[u]}(d\pi dR_u A, d\pi dR_u B) \\ &= B_u(\mathtt{proj}\, dR_u A, \mathtt{proj}\, dR_u B) \\ &= B_e(dL_{u^{-1}}\, \mathtt{proj}_{\mathcal{Q}_u}\, dR_u A, dL_{u^{-1}}\, \mathtt{proj}_{\mathcal{Q}_u}\, dR_u B) \end{aligned}$$

where $\mathtt{proj}$ is the projection on $\mathcal{Q}_u = dL_u \mathcal{Q}$. The latter formula allows to easily compute the metric on the various groups.

4. from now until the determination of symplectic forms, all results are computed by Maxima [1041].

(1) As first solution, we find of course the same algebra $\mathcal{R}_1$ as the one of page 4187.

(2) The second solution is also the same as the previously found one.

(3) The third solution is

$$\begin{aligned} A &= J_1 + J_2 + sV & [A,B] &= B + sC \\ B &= W - \frac{r}{2}V & [A,C] &= 2C \\ C &= M & [A,D] &= 2sB \\ D &= L + rJ_2 & [B,D] &= -rB \\ & & [C,D] &= -rC, \end{aligned}$$

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & \frac{\beta rs - 2\alpha r}{2} \\ \beta & 0 & 0 & \frac{\beta r}{2} \\ \gamma & -\frac{\beta rs - 2\alpha r}{2} & -\frac{\beta r}{2} & 0 \end{pmatrix}, \tag{84.30}$$

$$\det \omega = \frac{\beta^4 r^2 s^2}{4} - \frac{\alpha\beta^3 r^2 s}{2} + \frac{\alpha^2\beta^2 r^2}{4},$$

Conditions: $r \neq 0$, $\beta \neq 0$ and $\alpha \neq \beta r$. The map $\phi_3 \colon \mathcal{R}_3 \to \mathcal{R}_c$

$$\phi_3 = \begin{pmatrix} 1 & 0 & 0 & r \\ 0 & 1/2s & 0 & 1 \\ 0 & s & 1 & s^2 \\ 0 & 0 & 0 & r \end{pmatrix}.$$

$(\det \phi_3 = r/2s)$ provides an isomorphism between $\mathcal{R}_3$ and the canonical algebra.

(4) The fourth solution is

$$\begin{aligned} A &= J_1 + J_2 + sV & [A,B] &= B + sC \\ B &= W & [A,C] &= 2C \\ C &= M & [A,D] &= 2sB \\ D &= L + rJ_2 + rsV & [B,D] &= -rsC \\ & & [C,D] &= -rC, \end{aligned}$$

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & \frac{\beta rs}{2} \\ \beta & 0 & 0 & \frac{\beta r}{2} \\ \gamma & -\frac{\beta rs}{2} & -\frac{\beta r}{2} & 0 \end{pmatrix}, \tag{84.31}$$

$$\det \omega = \frac{\beta^4 r^2 s^2}{4} - \frac{\alpha\beta^3 r^2 s}{2} + \frac{\alpha^2\beta^2 r^2}{4}.$$

Conditions: $r \neq 0$, $\beta \neq 0$ and $\alpha \neq \beta s$. The map $\phi_4 \colon \mathcal{R}_4 \to \mathcal{R}_c$

$$\phi_4 = \begin{pmatrix} 2 & 0 & 0 & -r \\ 0 & 1 & 1/s & rs \\ 0 & 1 & 0 & 2rs \\ 3 & 0 & 0 & -r \end{pmatrix}$$

$(\det \phi_4 = -r/s)$ provides an isomorphism between $\mathcal{R}_4$ and the canonical algebra.

(5) The fifth solution is

$$\begin{aligned} A &= J_1 - J_2 + sV & [A,B] &= B + sC \\ B &= W - rsJ_2 + rs^2V & [A,D] &= 2sB + 2D \\ C &= M + rJ_2 - rsV & [B,D] &= 2rs^2B + rs^3C + rsD \\ D &= L + rs^2J_2 - rs^3V & [C,D] &= -2rsB - rs^2C - rD, \end{aligned}$$

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & \frac{\beta r s^2 + 2\alpha r s + 2\epsilon}{r} \\ \alpha & 0 & 0 & \epsilon s \\ \beta & 0 & 0 & -\epsilon \\ -\frac{\beta r s^2 + 2\alpha r s + 2\epsilon}{r} & -\epsilon s & \epsilon & 0 \end{pmatrix} \tag{84.32}$$

$$\det \omega = \beta^2 \epsilon^2 s^2 + 2\alpha\beta\epsilon^2 s + \alpha^2 \epsilon^2, \tag{84.33}$$

Conditions: $r \neq 0$, $\epsilon \neq 0$, $\alpha \neq -\beta s$. The map $\phi_5 \colon \mathcal{R}_5 \to \mathcal{R}_c$

$$\phi_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2s \\ 0 & 0 & 0 & -1 \\ 0 & -rs & r & rs^2 \end{pmatrix} \tag{84.34}$$

$(\det \phi_5 = r)$ provides an isomorphism between $\mathcal{R}_5$ and the canonical algebra.

(6) The sixth solution is

$$\begin{aligned} A &= J_1 + J_2 & [A, B] &= B \\ B &= W & [A, C] &= 2C \\ C &= M & [C, D] &= -rC, \\ D &= L + rJ_2 \end{aligned}$$

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & 0 \\ \beta & 0 & 0 & \frac{\beta r}{2} \\ \gamma & 0 & -\frac{\beta r}{2} & 0 \end{pmatrix}, \tag{84.35}$$

$$\det \omega = \left(\frac{\alpha\beta r}{2} \right)^2, \tag{84.36}$$

Conditions: $\alpha \neq 0$, $\beta \neq 0$ and $r \neq 0$. This algebra is isomorphic to the next one.

(7) The seventh solution is

$$\begin{aligned} A &= J_1 - J_2 & [A, B] &= B \\ B &= W & [A, D] &= 2D \\ C &= M + rJ_2 & [C, D] &= -rD, \\ D &= L \end{aligned}$$

$$\omega = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & 0 \\ \beta & 0 & 0 & \frac{\gamma r}{2} \\ \gamma & 0 & -\frac{\gamma r}{2} & 0 \end{pmatrix}, \tag{84.37}$$

$$\det \omega = \pm \left(\frac{\alpha\gamma r}{2} \right), \tag{84.38}$$

and the conditions are $\alpha \neq 0$, $\gamma \neq 0$, $r \neq 0$. The map $\phi_7 \colon \mathcal{R}_7 \to \mathcal{R}_c$

$$\phi_7 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & r & 0 \end{pmatrix}$$

$(\det \phi_7 = -r)$ provides an isomorphism between $\mathcal{R}_7$ and the canonical algebra.

All these algebras are solvable of order two (the commutators of commutators vanish) — but not nilpotent.

Proposition 84.12.
Surjectivity of the action of the group R_3.

Proof. The Lie algebra $\mathcal{R}_3$ can be written as a sequence of split extensions

$$\mathcal{R}_3 = \mathbb{R}A \oplus_{\text{ad}} \mathbb{R}D \oplus_{\text{ad}} \text{Span}\{B, C\},$$

thus a general element of R_3 reads

$$r_3(a, b, c, d) = e^{aJ_1+aJ_2+asV} e^{dL+drJ_2} e^{bW-\frac{br}{2}V} e^{cM}. \tag{84.39}$$

Eqelgener31

When $[X, Y] = sY$, one has the formula

$$\ln(e^X e^Y) = X + \frac{1 - e^s}{s} Y. \tag{84.40}$$

This allows us to write

$$\ln(e^{xJ_2} e^{yL}) = xJ_2 - \frac{1 - e^{-x}}{x} yL,$$

and to conclude that we can make the splitting

$$e^{dL+drJ_2} = e^{drJ_2} e^{f(d)L}$$

where $f(d) = d^2 r/(e^{-dr} - 1)$. The general element (84.39) becomes

$$r_3(a, b, c, d) = e^{aJ_2+asV} e^{drJ_2} e^{aJ_1} e^{f(d)L} e^{bW-brV/2} e^{cM}.$$

Up to redefinition of c, we can split $e^{bW-brV/2}$ into $e^{bW} e^{-brV/2}$, and a new redefinition allows us to commute $e^{-brV/2}$ and e^{bW}. Up to new redefinitions we are left with

$$r_3(a, b, c, d) = s(a, b, d) e^{aJ_1} e^{f(d)L} e^{bW} e^{cM}$$

where $s(a, b, d)$ is an element of S which depends on a, b and d. So a general element of $R_3[u]$ is

$$[e^{cM} e^{bW} e^{-f(d)L} e^{aJ_1} u] \tag{84.41}$$

where f fails to be surjective. □

84.2 Deformation of AdS_4

Abstract

Ceci contient les rebuts de démonstrations et de choses non retenues pour la ligne droite de ma thèse.

A description of the black hole construction is given in section 79.8. As far as Iwasawa decomposition is concerned, we recall that we fix choices in such a way that we have

$$\mathcal{N} = \{W_i, V_j, M, L\} \qquad i = 5, \dots, l + 1 \tag{84.42a}$$
$$\mathcal{A} = \{J_1, J_2\} \tag{84.42b}$$

with the commutator table

$$[V_i, W_j] = \delta_{ij} M \qquad [V_j, L] = 2W_j \tag{84.43a}$$
$$[J_1, W_j] = W_j \qquad [J_2, V_i] = V_i \tag{84.43b}$$
$$[J_1, L] = L \qquad [J_2, L] = -L \tag{84.43c}$$
$$[J_1, M] = M \qquad [J_2, M] = M \tag{84.43d}$$

where $\mathcal{N}$ is the nilpotent part of $\text{SO}(2, l - 1)$ and $\mathcal{A}$ the abelian one. Notice that $W, J_1 \in \mathcal{H}$, and $J_2 \in \mathcal{Q}$.

PropRsurSglobgroup

Proposition 84.13.
The homogeneous space $\mathcal{U} = R/S$ is globally of group type.

Proof. The fact that J_2 and V do not act on $[u]$ makes that *locally* the group corresponding to the algebra $\text{Span}\{J_1, W, M, L\}$ is diffeomorphic to $\mathcal{U}$. It is however not clear that the action will be *globally* transitive.

We know that $u = \exp q_0$ and

$$[q_0, [q_0, J_2]] = \left(\frac{\pi}{2}\right)^2 J_2, \tag{84.44a}$$

EqqqJ2a

$$\texttt{proj}_{\mathcal{Q}}[q_0, [q_0, V]] = \left(\frac{\pi}{2}\right)^2 \texttt{proj}_{\mathcal{Q}} V, \tag{84.44b}$$

EqqqVb

Now, r belongs to S when there exists a $h \in H$ such that $ru = uh$, or when $\mathbf{Ad}(u)r \in H$. Thus we are lead to compute thinks like

$$\mathbf{Ad}(e^{q_0})r$$

and check for which r this belongs to H. We will prove that this condition is satisfied for e^{aJ_2} and e^{bV}. We have to study the $\mathcal{Q}$-component of $e^{\text{ad}(q_0)X}$, and e^X will belong to S when this component vanishes. Since $J_2 \in \mathcal{Q}$, we see that, in the expansion of $e^{\text{ad}(q_0)}J_2$, the even terms are in $\mathcal{Q}$ while the odd ones are in $\mathcal{H}$. Now equation (84.44a) makes that the odd terms provide the expansion of $\cos(\pi/2) = 0$.

For V, we remark that it has a component $V_{\mathcal{Q}}$ and $V_{\mathcal{H}}$, the fact that $[q_0, V] \in \mathcal{H}$ makes that once again the odd terms does not pose any problem. Equation (84.44b) gives the expansion of $\cos(\pi/2)$ for the even terms.

So the stabilizer of $[u]$ is the group generated by e^{aJ_2} and e^{bV}. We denote by $\mathcal{S}$ the Lie algebra:

$$\mathcal{S} = \text{Span}\{J_2, V\}.$$

The stabilizer cannot be larger because R has 6 dimensions while $\mathcal{U}$ has 4 dimensions. Now we consider

$$\mathcal{R}' = \text{Span}\{J_1, W, M, L\},$$

and the first important remark is that it is an algebra. So it is immediately clear that locally, $\mathcal{U}$ has the structure of the corresponding group R'. Let us prove that the action of R' on $\mathcal{U}$ is globally transitive, i.e. $R'[u] = R[u]$. For that, just remark that $\mathcal{S}$ acts on $\mathcal{R}'$, so

$$\mathcal{R} = \mathcal{S} \oplus_{\text{ad}} \mathcal{R}'$$

and, as far as groups are concerned, we conclude that $R = SR'$ and hence that $R = R'S$. This proves that R' acts transitively on $\mathcal{U}$. We will see later that this group is not satisfactory because it does not posse a symplectic form.

We know that R' acts on $\mathcal{U} = R/S$ freely because none of the elements of R' acts as identity on $[u]$. This proves that $\mathcal{U}$ is locally of group type. It is also globally of group type because $\mathcal{U} = R'[u]$ is only the orbit of $[u]$.

□

We try to deform the open orbit of AdS_4 seen as the group of the first algebra (see page 4187). We rewrite this algebra, that we name $\mathcal{R}_1$, as

SubEqTablealgun

$$[p, q] = -rE \qquad r \neq 0 \tag{84.45a}$$

$$[A, p] = p + sE \tag{84.45b}$$

$$[A, E] = \frac{3}{2}E \tag{84.45c}$$

$$[A, p] = 2sp + \frac{1}{2}q, \tag{84.45d}$$

because we are going to see it as a split extension of the Heisenberg algebra spanned by $\{p, q, E\}$. The first work is to identity the parameters $(d\overline{X}, d\mu, 2d)$ of the derivation which extends the Heisenberg algebra to $\mathcal{R}_1$. The condition is that the commutator

$$[aA + v_1p + v_2q + zE, a'A + v_1'p + v_2'q + z'E]_{(d\overline{X},d\mu,2d)}$$

seen as split extension (equation (83.99))

$$\begin{aligned} d\overline{X}\big(a(v_1'p - v_2'q) - a'(v_1p - v_2q)\big) + d\mu\big(a(v_1'p + v_2'q) - a'(v_1p + v_2q)\big)E \\ + 2d(az' - a'z)E \\ + \Omega(v_1p + v_2q, v_1'p + v_2'q)E \end{aligned}$$

equals the same commutator given by the table (84.45):

$$\begin{aligned} p\big(av_1' + 2sav_2' - a'v_1 - 2a'v_2s\big) + q\big(\frac{a}{2}v_2' - \frac{a'}{2}v_2\big) \\ + E\big(sav_1' + \frac{3a}{2}z' - a'v_1s - v_1v_2'r + v_2v_1'r - \frac{3a'}{2}z\big). \end{aligned}$$

Equating the terms with z in the coefficient of E, we find

$$d = 3/4. \tag{84.46}$$

The terms with p and q give the matrix $\overline{X}$ by

$$d\overline{X}\begin{pmatrix} av_1' - a'v_1 \\ av_2' - a'v_2 \end{pmatrix} = \begin{pmatrix} av_1' - a'v_1 + 2s(av_2' - a'v_2) \\ \frac{1}{2}(av_2' - a'v_2) \end{pmatrix},$$

from what it follows that

$$\overline{X} = \begin{pmatrix} 4/3 & 8s/3 \\ 0 & 2/3 \end{pmatrix}. \tag{84.47}$$

When we pose $v_2 = v_2'$, we find

$$\mu(p) = 4s/3 \tag{84.48}$$

and with $v_1 = v_1' = 0$, we find

$$\mu(q) = 0. \tag{84.49}$$

The symplectic form on $V = \mathrm{Span}\{p, q\}$ is

$$\Omega = \begin{pmatrix} 0 & -r \\ r & 0 \end{pmatrix}. \tag{84.50}$$

One can check that the condition $\Omega(\mathrm{Id} - \overline{X}) + (\mathrm{Id} - \overline{X})^t\Omega = 0$ is satisfied, so that $\overline{X} \in \mathfrak{sp}(V, \Omega)$. Therefore, the algebra $\mathcal{R}_1$ is a split extension of $\mathscr{H}_n$ in the sense of section 83.5. The symplectic form on $\mathcal{R}_1$ is given by proposition (83.25):

$$\omega_1 = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ \alpha & 0 & 0 & \frac{2\beta r}{3} \\ \beta & 0 & 0 & 0 \\ \gamma & -\frac{2\beta r}{3} & 0 & 0 \end{pmatrix}, \tag{84.51}$$

and the symplectic form on $\mathcal{R}_1$ seen as extension of Heisenberg algebra is given by equation (83.25) which in the present case reads, in the basis $\{A, p, E, q\}$,

$$-\delta E^* = \begin{pmatrix} 0 & s & 3/2 & 0 \\ -s & 0 & 0 & -r \\ -3/2 & 0 & 0 & 0 \\ 0 & r & 0 & 0 \end{pmatrix}. \tag{84.52}$$

We see that the latter is a particular case of the general one with $\beta = -3/2$, $\alpha = -s$ and $\gamma = 0$. In order to prove that ω_1 is a Chevalley coboundary for the trivial representation, we search for an element $\xi \in \mathcal{R}_1$ such that $\delta\xi^* = \omega_1$, see equation (66.240). We pose

$$\xi = \xi_A A + \xi_p p + \xi_E E + \xi_q q$$

¡¡ Avertissement/question à la lectrice !! 84.14
. . . we search for an element. . . *est-ce qu'il faut vraiment le* for*??*

and we impose that

$$\delta\xi^*(X,Y) = -\xi^*([X,Y]) = \omega_1(X,Y). \tag{84.53}$$

A resolution with Maxima (see appendix 88.2) gives the result

$$\xi_p = \frac{2\beta s - 3\alpha}{3}, \quad \xi_E = -\frac{2\beta}{3}, \quad \xi_q = -\frac{6\gamma + 8\beta s^2 - 12\alpha s}{3}, \tag{84.54}$$ EqSolxialgun

while ξ_A has no constrains [5]. This proves that ω_1 is a Chevalley 2-coboundary.

A problem is now to adapt the results of theorem 83.24 and 83.34 to the new symplectic form which is not the one of proposition 83.25.

First, we fix the element $g \in \mathcal{R}_1$ such that PgAdgXEbbekl

$$(\mathrm{Ad}(g)^*\delta\xi^*)(X) = E^*(X), \quad \text{or} \quad \xi^*(\mathrm{Ad}(g)X) = E^*(X). \tag{84.55}$$

We search g under the form

$$g = e^{aA}e^{v_1 p + v_2 q + zE}$$

Let us begin with $X = E$: $\xi^*(\mathrm{Ad}(g)E) = 1$. Using the commutation relations, $Ad(g)E = e^{\mathrm{ad}(aA)}E = e^{3a/2}E$. It fixes the value of a:

$$a = -\frac{2}{3}\ln\big(-\frac{2\beta}{3}\big), \tag{84.56}$$ Eqleadegxistar

if $\beta < 0$. The parameters v_1 and v_2 are more difficult:

$$\begin{aligned}\mathrm{Ad}(g)p &= e^{\mathrm{ad}(aA)}(p + rE)\\ &= e^{\mathrm{ad}(aA)}p + rv_2 e^{3a/2}E\end{aligned}$$

The first terms of $e^{\mathrm{ad}(aA)}p$ are EqPremtermsAdgp

$$p \tag{84.57a}$$
$$a(p + sE) \tag{84.57b}$$
$$\frac{a^2}{2}(p + sE + s\frac{3}{2}E) = \frac{a^2}{2}p + \frac{a^2}{2}\big(1 + \frac{3}{2}\big)sE \tag{84.57c}$$
$$\frac{a^3}{3!}(p + sE + s\frac{3}{2}E) + \frac{a^3}{3!}s\frac{3}{2}E = \frac{a^3}{3!}p + \frac{a^3}{3}\big(1 + \frac{3}{2} + \frac{3}{2}\frac{3}{2}\big)sE. \tag{84.57d}$$

At each step, the first term gives the expansion of $e^a p$ and the term with sE of $[A,p]$ can be regrouped with the other to give, in the term of a^n,

$$\big(1 + \frac{3}{2} + \cdots + \big(\frac{3}{2}\big)^{n-1}\big)sE.$$

Using the well know sum formula $S_n(x) = 1 + x + \cdots + x^n = \frac{x^{n+1} - 1}{x - 1}$, the term with E in $e^{\mathrm{ad}(aA)}p$ is

$$\frac{a^n}{n!}S_{n-1}\big(\frac{3}{2}\big)sE = \frac{a^n - 1}{x - 1}$$

where $x = 3/2$. Putting all terms together,

$$\begin{aligned}e^{\mathrm{ad}(aA)}p &= e^a p + 2\sum_{n=0}^{\infty}\frac{a^n}{n!}x^n sE - 2\sum_{n=0}^{\infty}\frac{a^n}{n!}sE\\ &= e^a p + 2e^{3a/2}sE - 2e^a sE.\end{aligned}$$

5. This is clear because A does not appear in the commutators.

Computations with Maxima give:

$$\mathrm{Ad}(g)p = gpg^{-1} = e^a p + 2e^{3a/2}sE - 2e^a sE + v_2 r e^{3a/2}E,$$

which must satisfy $\xi^*\big(\mathrm{Ad}(g)p\big) = E^*(p) = 0$. So v_2 is given by the equation

$$e^a\xi_p + 2s\xi_E\big(e^{3a/2} - e^a\big) + \xi_E v_2 r e^{3a/2} = 0 \tag{84.58}$$

where a is given by equation (84.56). For $\mathrm{Ad}(g)q$, we find that v_1 is solution of

$$4s\xi_p(e^a - e^{a/2}) + \xi_E\big(-e^{3a/2}rv_1 + 4e^{3a/2}s^2 - 8e^a s^2 + 4e^{a/2}s^2\big) + \xi_q e^{a/2} = 0. \tag{84.59}$$

84.2.1 Isomorphism with $\mathcal{R}$

Let $\mathscr{H}_1 = V \oplus \mathbb{R}E$ be the Heisenberg algebra and its extension $\mathcal{F}(\mathbb{1}, 0, 2)$, i.e.

$$\mathrm{ad}(A) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix}$$

on $\mathscr{H}_1$. We extend $\mathcal{F}$ by a two dimensional algebra $\mathcal{E} = \mathrm{Span}\{E_1, E_2\}$ with

$$[E_1, E_2] = 2E_2$$

and

$$\mathrm{ad}(E_1)\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} p \\ -q \end{pmatrix}, \quad \mathrm{ad}(E_2)\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} q \\ 0 \end{pmatrix}. \tag{84.60}$$

All this produces a new algebra

$$\mathcal{E} \oplus_\rho \mathcal{F} = \mathrm{Span}\{E_1, E_2, A, p, q, E\}$$

whose full commutator table is

$$\begin{array}{lll} [E_1, E_2] = 2E_2 & [A, p] = p & [E_2, q] = p \\ [E_1, p] = p & [A, p] = p & [p, q] = E \\ [E_1, q] = -q & [A, E] = 2E & \end{array}$$

One can check that the following is an isomorphism $\alpha\colon \mathcal{E} \oplus_\rho \mathcal{F} \to \mathcal{R}$: [EqDefalphaisomRF]

$$\alpha(E_1) = J_1 - J_2 \qquad \alpha(E_2) = L/2 \tag{84.61a}$$
$$\alpha(A) = J_1 + J_2 \qquad \alpha(p) = W \tag{84.61b}$$
$$\alpha(q) = -V \qquad \alpha(E) = M. \tag{84.61c}$$

Trough this isomorphism, we say that $\mathcal{R} = \mathcal{E} \oplus_\rho \mathcal{F}$. The projection $\mathcal{R} \to \mathcal{F}$ is given by $E_1 \to 0$, $E_2 \to 0$ or, trough α, $L \to 0$ and $J_1 - J_2 \to 0$. After projection, the table (66.125) becomes

$$[V, W] = M \qquad [H_2, W] = W \tag{84.62a}$$
$$[H_2, V] = V \qquad [H_2, M] = 2M, \tag{84.62b}$$

which is the table of $\mathcal{F}(\mathbb{1}, 0, 2)$. Projections of algebras $\mathcal{R}_j$ are:

84.2.1.1 Algebra 1

$$\begin{aligned} A &\to \frac{3}{4}H_2 + rV \\ B &\to W \\ C &\to M \\ D &\to sV. \end{aligned}$$

Not surjective on $\mathcal{F}$ when $s = 0$.

84.2.1.2 Algebra 3

$$
\begin{aligned}
A &\to H_2 + sV \\
B &\to W - \frac{r}{2}V \\
C &\to M \\
D &\to \frac{r}{2}H_2.
\end{aligned}
$$

Not surjective on $\mathcal{F}$ when $s = 0$.

84.2.1.3 Algebra 4

$$
\begin{aligned}
A &\to H_2 + sV \\
B &\to W \\
C &\to M \\
D &\to \frac{r}{2}H_2 + rsV.
\end{aligned}
$$

Not surjective on $\mathcal{F}$ when $s = 0$.

84.2.1.4 Algebra 5

$$
\begin{aligned}
A &\to sV \\
B &\to W - \frac{rs}{2}H_2 + rs^2V \\
C &\to M + \frac{r}{2} - rsV \\
D &\to \frac{rs^2}{2}H_2 - rs^3V.
\end{aligned}
$$

Not surjective on $\mathcal{F}$ when $s = 0$.

The pathologic cases are $\mathcal{R}_3$, $\mathcal{R}_4$, $\mathcal{R}_5$ and $\mathcal{R}_6$ with $s = 0$. They are each isomorphic to the algebra generated by $\{A, B, C, D\}$ and subject to the relations

$$
\begin{aligned}
[A, B] &= B \\
[A, C] &= 2C \\
[C, D] &= -rC.
\end{aligned}
$$

We consider the map

$$
\begin{aligned}
\varphi_j &\colon R_j \to F(\mathbb{1}, 0, 2) \\
\varphi_j(r_j) &= \mathbf{Ad}(r_{j1})r_{j0}
\end{aligned}
\tag{84.63}
$$

where r_{j0} and r_{j1} are defined by the condition $\alpha(r_{j1}r_{j0}) = r_j$. This map actually takes its values in F because, by construction, $\mathrm{ad}(\mathcal{E})\mathcal{F} \subset \mathcal{F}$.

Theorem 84.15.
The map φ_j is bijective.

Proof. Let us first assume that $\varphi_j(r_j) = \varphi_j(s_j)$. In this case, we have

$$
r_{j1}r_{j0}r_{j1}^{-1} = s_{j1}s_{j0}s_{j1}^{-1},
$$

but $r_{j1}r_{j0} = \alpha^{-1}(r_j)$ and $s_{j1}s_{j0} = \alpha^{-1}(s_j)$, thus

$$\alpha^{-1}(s_j^{-1}r_j) = s_{j1}^{-1}r_{j1},$$

or

$$r_j s_j^{-1} = \alpha(s_{j1}^{-1}r_{j1})$$

The left-hand side of the latter equation is an element of R_j while the right-hand side is an element of $\alpha(E)$, i.e. of the group generated by $J_1 - J_2$ and L. One can check that the intersection between $\mathcal{R}_j$ and $\mathcal{E}$ reduces to $\{0\}$, so $r_j = s_j$. This proves injectivity. □

84.2.2 The symplectic issue

When one pull back the kernel from $F(\mathbb{1}, 0, 2)$ to R_j, there is no reason why the symplectic form δE^* on F maps to the general symplectic form on R_j. The issue is to push forward the symplectic form ω_j of R_j to F by the diffeomorphism φ_j and find a kernel on F for this new symplectic form instead if the canonical one.

We have the following two maps:

$$\varphi_j \colon R_j \to F(\mathbb{1}, 0, 2)$$
$$d\alpha \colon \mathcal{E} \oplus_\rho \mathcal{F} \to \mathcal{R}$$

where $d\alpha$ is the isomorphism (84.61). The new symplectic form on F is defined by

$$\Omega_j(X_0, X_0') = \omega_j\big(d\varphi_j^{-1}X_0, d\varphi_j^{-1}X_0'\big) \tag{84.64}$$

with X_0, $X_0' \in \mathcal{F} \subset \mathcal{E} \oplus_\rho \mathcal{R}$. First, remark that when $X_1 \in \mathcal{E}$ and $X_0 \in \mathcal{F}$, we have

$$\frac{d}{dt}\Big[\alpha\big(e^{tX_1}e^{tX_0}\big)\Big]_{t=0} = d\alpha(X_1 + X_0) \tag{84.65}$$

which is surjective on $d\alpha(\mathcal{E} \oplus_\rho \mathcal{F})$. Second, $d\varphi_j$ is the projection because

$$\begin{aligned} d\varphi_j\left(\frac{d}{dt}\Big[\alpha(e^{tX_1}e^{tX_0})\Big]_{t=0}\right) &= \frac{d}{dt}\Big[\,\mathbf{Ad}(e^{tX_1}e^{tX_0})\Big]_{t=0} \\ &= X_0. \end{aligned}$$

We conclude that

$$d\varphi_j^{-1}(X_0) = d\alpha(X_1 + X_0) \in \mathcal{R}_j$$

where the "$\in \mathcal{R}_j$" is part of the formula: this fixes the choice of $X_1 \in \mathcal{E}$.

Let us compute for example $d\varphi_1^{-1}(A)$. Using the value of $d\alpha(A)$, we find

$$d\varphi_1^{-1}(A) = d\alpha(X_1) + J_1 + J_2$$

where $X_1 \in \mathcal{E}$ must be chosen in such a way that the whole belongs to $\mathcal{R}_1$.

Let us for example compute $d\varphi_1^{-1}(A) = d\alpha(X_1) + J_1 + J_2 \in \mathcal{R}_1$. In order to fix X_1, we we suppose $X_1 = xE_1 + yE_2$, we write

$$d\alpha(X_1) = xJ_1 - xJ_2 + \frac{y}{2}L$$

and we solve

$$d\alpha(X_1) + J_1 + J_2 = aA + bB + cC + dD$$

with respect to x, y, a, b, c, d. This is a linear system of 6 equations for 6 variables. The solution is $b = c = 0$, $x = 1/3$, $a = 4/3$, $d = -4rs/3$ and $y = -8rs/3$. So we have

$$d\varphi_1^{-1}(A) = \frac{4}{3}J_1 + \frac{2}{3}J_2 - \frac{4rs}{3}L = \frac{4}{3}A - \frac{4rs}{3}D. \tag{84.66}$$

84.3 Isospectral deformations of M

In this section, we present a modified version of the oscillatory integral product (83.94) leading to a left invariant associative algebra structure on the space of square integrable functions on R_0. Why is it better that the initial product defined over smooth compactly supported functions? The motivation of considering the square integrable functions is the fact that the spectral triple defined in noncommutative geometry contains the space of square integrable spinors (see the book [953]). The fact to stabilize the space of square integrable functions is then an indispensable step in order to put our results in the framework of spectral geometry.

Theorem 84.16.
Let u and v be smooth compactly supported functions on R_0. Define the following three-point functions:

$$\begin{aligned} S :=& S_V\big(\cosh(a_1-a_2)x_0, \cosh(a_2-a_0)x_1, \cosh(a_0-a_1)x_2\big) \\ & - \bigoplus_{0,1,2} \sinh\big(2(a_0-a_1)\big) z_2; \end{aligned} \tag{84.67}$$

and

$$\begin{aligned} A := \Big[& \cosh\big(2(a_1-a_2)\big)\cosh(2(a_2-a_0))\cosh(2(a_0-a_1)) \\ & \big[\cosh(a_1-a_2)\cosh(a_2-a_0)\cosh(a_0-a_1)\big]^{\dim R_0-2}\Big]^{\frac{1}{2}}. \end{aligned}$$

Then the formula

$$u \star_\theta^{(2)} v := \frac{1}{\theta^{\dim R_0}} \int_{R_0\times R_0} A e^{\frac{2i}{\theta}S} u\otimes v \tag{84.68}$$

extends to $L^2(R_0)$ as a left invariant associative Hilbert algebra structure. In particular, one has the strong closedness property:

$$\int u \star_\theta^{(2)} v = \int uv.$$

Proof. The oscillatory integral product (83.94) may be obtained by intertwining the Weyl product on the Schwartz space $\mathscr{S}$ (in the Darboux global coordinates (83.93)) by the following integral operator [1039]:

$$\tau := F^{-1}\circ(\phi_\theta^{-1})^\star \circ F,$$

F being the partial Fourier transform with respect to the central variable z:

$$F(u)(a,x,\xi) := \int e^{-i\xi z} u(a,x,z)\mathrm{d}z;$$

and ϕ_θ the one parameter family of diffeomorphism(s):

$$\phi_\theta(a,x,\xi) = (a, \frac{1}{\cosh(\frac{\theta}{2}\xi)}x, \frac{1}{\theta}\sinh(\theta\xi)).$$

Set $\mathcal{J} := |(\phi^{-1})^\star \mathrm{Jac}_\phi|^{-\frac{1}{2}}$ and observe that for all $u\in C^\infty\cap L^2$, the function $\mathcal{J}\,(\phi^{-1})^\star u$ belongs to L^2. Indeed, one has

$$\int |\mathcal{J}\,(\phi^{-1})^\star u|^2 = \int |\phi^\star\mathcal{J}|^2\,|\mathrm{Jac}_\phi|\,|u|^2 = \int |u|^2.$$

Therefore, a standard density argument yields the following isometry:

$$T_\theta : L^2(R_0) \longrightarrow L^2(R_0) : u \mapsto F^{-1}\circ m_{\mathcal{J}}\circ(\phi^{-1})^\star\circ F(u),$$

where $m_{\mathcal{J}}$ denotes the multiplication by $\mathcal{J}$. Observing that $T_\theta = F^{-1}\circ m_{\mathcal{J}}\circ F\circ\tau$, one has $\star_\theta^{(2)} = F^{-1}\circ m_{\mathcal{J}}\circ F(\star_\theta)$. A straightforward computation (similar to the one in [961]) then yields the announced formula. □

Let us point out two facts with respect to the above formulas:

(1) The oscillating three-point kernel $A\exp\left(\frac{2i}{\theta}S\right)$ is symmetric under cyclic permutations.

(2) The above oscillating integral formula gives rise to a strongly closed, symmetry invariant, formal star product on the symplectic symmetric space (R_0, ω, s).

(3) Importance of left invariance will be explained in the subsection 84.3.1.

Proposition 84.17.
The space $L^2(R_0)^\infty$ of smooth vectors in $L^2(R_0)$ of the left regular representation closes as a subalgebra of $(L^2(R_0), \star_\theta^{(2)})$.

Proof. First, observe that the space of smooth vectors may be described as the intersection of the spaces $\{V_n\}$ where $V_{n+1} := (V_n)_1$, with $V_0 := L^2(R_0)$ and $(V_n)_1$ is defined as the space of elements a of V_n such that, for all $X \in \mathcal{R}_0$, $X.a$ exists as an element of V_n (we endow it with the projective limit topology).

Let thus $a, b \in V_1$. Then, $(X.a) \star b + a \star (X.b)$ belongs to V_0. Observing that $D \subset V_1$ and approximating a and b by sequences $\{a_n\}$ and $\{b_n\}$ in D, one gets (by continuity of $\star$): $(X.a) \star b + a \star (X.b) = \lim(X.a_n \star b_n + a_n \star (X.b_n)) = \lim X.(a_n \star b_n) = X.(a \star b)$. Hence $a \star b$ belongs to V_1. One then proceeds by induction. □

84.3.1 Complement: importance of left invariance

subsecImpLeftInvarDstar

This complement intends to shortly explain why to express the Dirac operator under the form of left invariant vectors (the $\tilde{t}_i$ in equation (79.584)).

Let R be a Lie group with a simply transitive action on the manifold M, and suppose that the Dirac operator on M reads

$$D|_M = \sigma^i \tilde{X}_i$$

where σ^i are constant and $\tilde{X}_i$ a basis of $\mathcal{R}$ of left invariant vector fields.

Proposition 84.18.
In this setting, if $\star^R$ is a right invariant product on R, the operator D is a derivation.

Proof. Take $u \in \mathrm{Fun}(M)$ and a spinor φ. First $\tilde{X}_i$ acts by derivation on the right invariant product $\star^R$. Indeed proposition 83.10 says that if one defines $u *_\nu v = T_\nu(T_\nu^{-1}u *_M T_\nu^{-1}v)$, then $\tilde{X}$ is a derivation of $*_\nu$ when T_ν is chosen in such a way that $T_\nu\rho_\nu(X)T_\nu^{-1} = \tilde{X}$. Hence have

$$\begin{aligned} D(u \star^R \varphi) &= \sigma^i \tilde{X}_i \cdot (u \star^R \varphi) \\ &= \sigma^i(\tilde{X}_i \cdot u) \star^R \varphi + \sigma^i u \star^R \tilde{X}_i \cdot \varphi \\ &= (Du) \star^R \varphi + u \star^R D\varphi. \end{aligned}$$

□

This result implies in turn that the operator $[D, u]$ is bounded because it is the multiplication operator by Du. Indeed

$$\begin{aligned} [D, u]_{\star^R}\varphi &= D(u \star^R \varphi) - u \star^R D\varphi \\ &= (Du) \star^R \varphi + u \star^R D\varphi - u \star^R D\varphi \\ &= (Du) \star^R \varphi. \end{aligned}$$

84.4 Perspectives

A main achievement of spectral non-commutative geometry is the ability of retrieving the original Riemannian manifold from the data of the spectral triple. Such a result does not exist in the case of *AdS* because the latter is a non-compact *pseudo*-Riemannian manifold. The main lines of such a reconstruction method can however be foreseen in the case of anti de Sitter space.

— Knowing the family of products $\star_\theta^{(2)}$, we know in particular the usual commutative product of functions. That should allow us to find back the manifold AdS_4.

— It is possible to extract the data of the curvature of the manifold from the data of its Dirac operator as the non-differential part of its square. That part will of course appear to be constant and negative (because we know that we were starting from anti de Sitter).

We only quantized an open orbit of AdS_4 because it is a whole physical domain. Quantization of the full space could be very interesting because of a special effect of the noncommutative product: two functions with disjoint supports can have a non vanishing product. What about the physical significance of that property when one multiplies a function supported in the singularity by a function supported in the physical part?

There is another reason to study the quantization of the full space. We will show in section 85.2 that a deformation of the full space by action of the Iwasawa component of $\mathrm{SO}(2,l-1)$ is possible. That quantization has the advantage of deforming the space by the action of the group which is precisely defining the singularity. In other words the *same* group can describe a singularity and a quantization. A work to be done is to try and recover the special causal structure from the data of the quantized manifold. That structure must be in some way contained in the spectral triple.

84.5 Spin structure on the black hole

The group R_1 defines the open orbit with the simple action $\tau(r)[r'u] = [rr'u]$ which fulfills $\pi \circ L_g = \tau(g) \circ \pi$. So we will first study the spin frame and spin bundle over R_1, and then bring the structure over the open orbit $\mathcal{U}$.

84.5.1 Frame bundle over R_1

A frame at $r \in R_1$ is an isometry between $\mathbb{R}^{1,3}$ and T_rR_1. Since dL_r is such an isometry, any isometry between $\mathcal{R}_1$ and T_rR_1 reads $a = dL_r a'$ with $a' \in \mathrm{SO}(\mathcal{R}_1)$. So we consider an isometry $\sigma\colon \mathbb{R}^{1,3} \to \mathcal{R}_1$ and a basis at r reads

$$b(v) = dL_r \circ \sigma a(v) \tag{84.69}$$

EqDefBaseGenebyR1

with $a \in \mathrm{SO}(1,3)$. The basis defined by equation (84.69) will be denoted by $b_a(r)$. Here, we will restrict our bundle to $a \in \mathrm{SO}_0(1,3)$. The fact that b_a is a global section of the frame bundle indicates us that the bundle is trivial by theorem 56.55. The trivialization map is

$$\begin{aligned} \beta\colon B(R_1) &\to R_1 \times \mathrm{SO}_0(1,3) \\ b_a(r) &\mapsto (r,a). \end{aligned} \tag{84.70}$$

84.5.2 Frame bundle over $\mathcal{U}$

The frame bundle over $[ru] \in \mathcal{U}$ is the set of isometries $\mathbb{R}^{1,3} \to T_{[ru]}\mathcal{U}$. If $\tilde{\sigma}$ is a given isometry

$$\tilde{\sigma}\colon \mathbb{R}^{1,3} \to T_{[u]}\mathcal{U}$$

(for example a numeration of the fundamental fields), the fiber over $[ru]$ is given by maps of the form

$$\tilde{b}_a(r)v = d\tau(r)\tilde{\sigma}av. \tag{84.71}$$

This is isomorphic to $R_1 \times \mathrm{SO}_0(1,3)$, in consequence of what we conclude that the frame bundles over $\mathcal{U}$ and R_1 are the same.

Chapter 85

Two notes for further developments

ChapNoteDev

Abstract

This chapter contains two directions that were explored during my thesis and that were not finished for different reasons.

In the first section, we state a result of Unterberger in [1042] which provides a deformation of the complex half-plane, and we show how to translate it as a new noncommutative product on the group $ax + b$, i.e. the Iwasawa subgroup of $\mathrm{SL}(2, \mathbb{R})$. The technique of deformation by group action described in appendix 82.6 then induces a deformed product on the dual of its Lie algebra. We do not study the properties (symmetries, maximal functional space of convergence, symplectic condition to be a true quantization, ...) of this product, but we show that Unterberger's result assures the existence of at least one good functional space. Unfortunately the formula reveals not to be universal; we show the lack of universality on two examples of actions of the group $ax + b$ on AdS_2. The failure is due to divergences of the derivatives of the functions z_i (see equation (82.37)).

This study is motivated by the fact that recent work (not published yet) of P. Bieliavsky provides an universal deformation of the AN of $\mathrm{SL}(2, \mathbb{R})$. We are thus allowed to say that the latter new product is "better" than the one of Unterberger. We do not address the question to know the precise point that makes the lack of universality in Unterberger.

The second section is an application of the extension lemma (lemma 83.23). We show that all the ingredients needed to deform the AN of $\mathrm{SO}(2, n)$ are present. The idea was to deform the AdS black hole using the action of the so-deformed AN. That should provide an alternative way to deform AdS to the one presented in chapter 84, and a quantization of AdS_l using the same group as the group which defines a black hole. That method would use the deformation by group action machinery described in appendix 82.6. The arising question is naturally to know if that quantization is in some sense equivalent to the one given in chapter 84 or not. That question is not answered yet.

85.1 Formula of Unterberger on $\mathrm{SL}(2, \mathbb{R})$

SecEplolUnter

The following results come from [1042] (from page 1219) and provide a deformation [1] of the half-plane

$$D = \{(\xi, \eta) \text{ tel que } \eta > 0\} \subset \mathbb{R}^2.$$

Before to give the precise statement that will be used, we need some definitions. A first product is defined by (we will precise the functional space later):

$$(f \circ g)(\xi, \eta) = \sum_{\alpha,\beta} \frac{(-1)^\alpha}{\alpha!\beta!} (4i\pi)^{-\alpha-\beta} (\partial_q^\alpha \partial_p^\beta \tilde{f})(0,0)(\partial_q^\beta \partial_p^\alpha \tilde{g})(0,0) \tag{85.1}$$

where

$$\tilde{f}(p, q) = f\Big(p + \xi\big(q + \sqrt{1+q^2}\big), \eta\big(q + \sqrt{1+q^2}\big)\Big),$$

1. or, at least, a new noncommutative product.

and the same for $\tilde{g}$. In particular,

$$(f\circ g)(0,1)=4\int f\big(\Psi(q_1,p_1)\big)g\big(\Psi(q_2,p_2)\big)e^{-4i\pi(-q_1p_2+q_2p_1)}\,dq_1\,dp_1\,dq_2\,dp_2.$$

with

$$\Psi(p,q)=\big(p,q+\sqrt{1+q^2}\big)$$

Definition 85.1.
Let r_1, r_2 and n be real numbers with $r_1\geqslant 0$. We denote by $\Sigma^n_{r_1,r_2}$ the space of functions $f\in C^\infty(D)$ such that for all $(j,k)\in\mathbb{N}\times\mathbb{N}$, there exists a $C>0$ such that

$$\left|\left(\frac{\partial}{\partial\xi}\right)^j\left(\eta\frac{\partial}{\partial\eta}\right)^k f(\xi,\eta)\right|\leqslant C\eta^{r_1}(1+\eta)^{r_2}(1+|\xi|)^{n-j}. \tag{85.2}$$

Now, theorem 8.2 in [1042] states

Theorem 85.2.
Let $f\in\Sigma^n_{r_1,r_2}$ and $g\in\Sigma^{n'}_{r_1'nr_2'}$. For each $N\in\mathbb{N}_0$, the function

$$h_N=f\circ g-\sum_{\alpha+\beta\leqslant N-1}\frac{(-1)^\alpha}{\alpha!\beta!}(4i\pi)^{-\alpha-\beta}\sum_{j,k,j',k'}C^{j,k}_{\beta,\alpha}C^{j',k'}_{\alpha,\beta}(e_1^je_2^kf)(e_1^{j'}e_2^{k'}g) \tag{85.3}$$

belongs to the space $\Sigma^{n+n'-N}_{r_1+r_1',r_2+r_2'}$ if constants $C^{j,k}_{\alpha,\beta}$ are defined by the requirement that

$$(\epsilon_2^\beta\epsilon_1^\alpha f)(\xi_0,\eta_0)=\sum_{j,k}C^{j,k}_{\alpha,\beta}(e_1^je_2^kf)(\xi_0,\eta_0)$$

for every smooth function f and $(\xi_0,\eta_0)\in D$ when $j+k\leqslant\alpha+\beta$ and $j\geqslant\alpha$ and $C^{j,k}_{\alpha,\beta}=0$ otherwise. The operators ϵ_i are defined by $\epsilon_1=e_1=\partial_\xi$ and $\epsilon_2=2\big[1+\big(\frac{\eta_0}{\eta}\big)^2\big]^{-1}(\xi_0\partial_\xi+\eta_0\partial_\eta)$. ThoUnterSigmaStable

For our purpose, the point is that there exists a product on $\mathrm{Fun}(D)$ and that theorem 85.2 provides a functional space stabilized by the product. We are now going to translate this result in terms of the Iwasawa subgroup $R=AN$ of $\mathrm{SL}(2,\mathbb{R})$ that is parametrized (see (53.321)) by

$$(a,l)=\begin{pmatrix}e^a & le^a\\ 0 & e^{-a}\end{pmatrix}.$$

The map

$$\begin{aligned}&j\colon R\to R'\\ &(a,l)\mapsto(e^{2a},le^{2a})\end{aligned} \tag{85.4}$$

provides an isomorphism between R and the group

$$R'=\left\{(\alpha,\beta)=\begin{pmatrix}\alpha & \beta\\ 0 & 1\end{pmatrix},\ \alpha>0\right\}.$$

The inverse of j is $j^{-1}(\alpha,\beta)=(\ln\alpha^{1/2},\beta\alpha^{-1})$. The group R' acts on D by

$$(\alpha,\beta)\cdot(\xi,\eta)=(\xi+\beta\alpha^{-1}\eta,\alpha^{-1}\eta) \tag{85.5}$$

which is a freely transitive action. For each choice of "reference point" $(\xi_0,\eta_0)\in D$ we build an identification $i\colon D\to R'$ by the requirement $i(\xi,\eta)\cdot(\xi_0,\eta_0)=(\xi,\eta)$, that is

$$i(\xi,\eta)=\left(\frac{\eta_0}{\eta},\frac{\xi-\xi_0}{\eta}\right). \tag{85.6}$$

Now we can identify D to R by $k\colon D \to R$, $k = j^{-1} \circ i$. For the choice $(\xi_0, \eta_0) = (0, 1)$, we find $k(\xi, \eta) = (k_a(\xi, \eta), k_l(\xi, \eta))$ where

$$k_a(\xi, \eta) = -\frac{1}{2} \ln \eta, \qquad k_l(\xi, \eta) = \xi \tag{85.7}$$

and the function f on R corresponds to the function $\tilde{f} = f \circ k$ on D.

The result of Unterberger is that the function f "can be quantized" if

$$\left| \left(\frac{\partial}{\partial \xi} \right)^j \left(\eta \frac{\partial}{\partial \eta} \right)^k \tilde{f}(\xi, \eta) \right| \leqslant C \eta^{r_1} (1 + \eta)^{r_2} (1 + |\xi|)^{n-j} \tag{85.8}$$

where n, r_1 and r_2 are real numbers and $r_1 \geqslant 0$. We want to see what condition has to be imposed on f in order for $\tilde{f}$ to fulfil this condition. In other words, we want to express the operator

$$A_{ij} = \left(\frac{\partial}{\partial \xi} \right)^j \left(\eta \frac{\partial}{\partial \eta} \right)^k$$

in terms of the coordinates on R. For that we compute $\partial_\xi \tilde{f}$ and $(\eta \partial_\eta) \tilde{f}$ in terms of $\partial_l f$ and $\partial_a f$.

Let us precise that, when we write expressions like $\eta \partial_\eta$, we mean for example

$$(\eta \partial_\eta \tilde{f})(\xi, 2) = 2(\partial_\eta f)(\xi, 2).$$

For $\partial_\xi \tilde{f}$ we have:

$$\begin{aligned}(\partial_\xi \tilde{f})(\xi, \eta) &= (\partial_l f) \circ k(\xi, \eta)(\partial_\xi k_k)(\xi, \eta) \\ &\quad + (\partial_a f) \circ k(\xi, \eta)(\partial_\xi k_a)(\xi, \eta),\end{aligned}$$

using the formula (85.7), we find $(\partial_\xi \tilde{f})(\xi, \eta) = (\partial_l f) \circ k(\xi, \eta)$ and we conclude that

$$\partial_\xi \tilde{f} = (\partial_l f) \circ k. \tag{85.9}$$

For $(\eta \partial_\eta) \tilde{f}$, we find

$$\begin{aligned}(\eta \partial_\eta)(f \circ k)(\xi, \eta) &= \eta \big(\partial_\eta (f \circ k) \big)(\xi, \eta) \\ &= \eta (\partial_a f) \circ k(\xi, \eta)(\partial_\eta k_a)(\xi, \eta) \\ &\quad + \eta (\partial_l f) \circ k(\xi, \eta)(\partial_\eta k_l)(\xi, \eta) \\ &= -\frac{1}{2} (\partial_a f) \circ k(\xi, \eta),\end{aligned}$$

and we conclude that

$$(\eta \partial_\eta)(f \circ k) = -\frac{1}{2} (\partial_a f) \circ k. \tag{85.10}$$

So the operator A_{ij}, expressed on R, reads

$$A_{ij}(f \circ k) = \left(-\frac{1}{2} \right)^j (\partial_a^j \partial_l^i f) \circ k, \tag{85.11}$$

and condition (85.8), with $(\xi, \eta) = k^{-1}(a, l) = (l, e^{-2a})$ reads now

$$\left| \frac{1}{2^k} (\partial_a^k \partial_l^j f)(a, l) \right| \leqslant C e^{-2 r_1 a} (1 + e^{-2a})^{r_2} (1 + |l|)^{n-j} \tag{85.12}$$

with $r_1 \geqslant 0$ and r_2, being any real number. From now on this regularity condition will be referred as the *Unterberger's condition*. That condition characterises a stable functional space for the Unterberger product on R.

We want now to test the deformation of manifold by action of R. A somewhat deceiving result that will be shown is that Unterberger's deformation of R is not an universal deformation in the sense that we will find some action of R on manifold for which the action deformation does not provide a deformation of the manifold.

85.1.1 Action on the dual of its Lie algebra

The action if given by

$$(a,l)\cdot\xi = (y_H + 2y_E l)H^* + y_E e^{-2a}E^*$$

where $\xi = y_H H^* + y_E E^*$ is any point in $\mathcal{R}^*$. The question is to know if the product $(u \star_{\mathcal{R}^*} v)$ makes sense when u and v are compactly supported smooth functions on $\mathcal{R}^*$. In order to address this question, we have to check if for every ξ in $\mathcal{R}^*$, the function

$$(\alpha^\xi u)(a,l) = u\big((a,l)^{-1}\cdot\xi\big) = u\big((y_H - 2y_E e^{-2a}l), y_E e^{2a}\big)$$

fulfils condition (85.12). So we consider

$$f(a,l) = u\big(\underbrace{(y_H - 2y_E e^{-2a}l)}_{z_H(a,l)}, \underbrace{y_E e^{-2a}}_{z_E(a,l)}\big),$$

and we compute

$$\begin{aligned}(\partial_l f)(a,l) &= (\partial_H u)(z_H,z_E)\partial_l(y_H - 2y_E e^{-2a}l)\\ &\quad + (\partial_E u)(z_H,z_E)\partial_l(y_E e^{-2a})\\ &= (\partial_H u)(z_H,z_E)(-2y_E e^{-2a}),\end{aligned}$$

so

$$(\partial_l^j f)(a,l) = (\partial_H^j u)(z_H,z_E)(-2y_E e^{-2a})^j. \tag{85.13}$$

The combination $y_E e^{-2a}$ which goes out is precisely z_E which remains in the derivative of u. But the derivative of u has compact support. Hence, in fact, the coefficient $y_E e^{-2a}$ remains constrained in the domain where the derivative of u does not vanishes. The point is that the coefficient which go out with derivatives is exactly made of z_H and z_E.

So $\mathcal{R}^*$ is as deformable as D. More precisely, a deformation of $\mathcal{R}^*$ by action of R is induced by the deformation of D by Unterberger.

85.1.2 First action on the two dimensional anti de Sitter space

We see AdS_2 as in 78.1.2 and we consider the following action of AN on AdS_2:

$$r\cdot \mathrm{Ad}(g)H = \mathrm{Ad}(gr^{-1})H.$$

It is easy to see what does this action become in terms of the cylinder:

$$\big(e^{y_A H}e^{y_N E}\big)\cdot \mathrm{Ad}\big(e^{x_K T}e^{x_N E}\big) = \mathrm{Ad}\big(e^{x_K T}e^{x_N E - y_N E}e^{y_A H}\big)H$$

where the adjoint action of $e^{y_A H}$ on H is of course trivial. Thus we have

$$(y_A,y_N)\cdot(x_K,x_N) = (x_K, x_N - y_N). \tag{85.14}$$

Notice that only one dimension of AN really acts. This action is thus not a natural one, and has to be seen as an interesting warm up. Using the notations of coordinates (78.2), we consider $x = \phi(\theta,h) \in AdS_2$ and $u \in C_c^\infty(Cyl)$, a compact supported function on AdS_2 and we compute

$$(\alpha^x u)(a,l) = u\big((a,l)^{-1}\cdot x\big) = u\big((-a,-le^{2a})\cdot x\big) = u\big(\theta, h + le^{2a}\big),$$

so that if we pose $f(a,l) = u\big(\theta, h + le^{2a}\big)$, we have

$$(\partial_l f)(a,l) = e^{2a}(\partial_2 u)(\theta, le^{2a}).$$

When one makes $a \to \infty$ and $l \to 0$ in such a way that le^{2a} remains constant, the function $(\partial_l f)$ diverges in an exponential way with respect to a. It contradicts Unterberger's condition (85.12).

85.1.3 Second action on the two dimensional anti de Sitter space

Let us now study the more natural action

$$r \cdot \mathrm{Ad}(g)H = \mathrm{Ad}(rg)H. \tag{85.15}$$

It is in general very difficult to find, for given y_A, y_N, x_K and x_N, the numbers (unique by construction) z_K and z_N such that

$$\mathrm{Ad}(e^{y_A H}e^{y_N E}e^{x_K T}e^{x_N E})H = \mathrm{Ad}(e^{z_K T}e^{z_N E})H.$$

In order to simplify the computations, we use the lemma 82.13 which states that we only have to perform the computation for one (x_K, x_N) in each orbit. We begin by $x_K = x_N = 0$, i.e. the orbit of H itself. First, computations show that

$$\begin{aligned}\mathrm{Ad}(e^{z_K T}e^{z_N E})H =& \big(\cos(2z_K) - \sin(2z_K)z_N\big)H\\ &- 2\big(\cos(2z_K)z_N + \sin(2z_K)\big)E\\ &+ \Big(\big(\cos(2z_K) - 1\big)z_N + \sin(2z_K)\Big)T.\end{aligned}$$

Next,

$$\mathrm{Ad}(e^{aH}e^{lE})H = \begin{pmatrix}1 & -2e^{2a}l\\ 0 & -1\end{pmatrix} = H - 2e^{2a}lE.$$

Comparing with the general form, we find that

$$\mathrm{Ad}(e^{aH}e^{lE})H = \mathrm{Ad}(e^{le^{2a}E})H, \tag{85.16}$$

or $(a,l) \cdot (0,0) = (0, le^{2a})$. What is important in our deformation problem is the function

$$(\alpha^H u)(a,l) = u\big((-a, -le^{2a})) \cdot H\big) = u(0,-l).$$

This function of course satisfies the Unterberger condition when u has a compact support.

The second orbit that we study is the one of $V = \mathrm{Ad}(e^{\pi T/4})H = -2E + T$. One has

$$\mathrm{Ad}(e^{aH}e^{lE})V = -lH + e^{-2a}(e^{4a}l^2 - e^{4a} - 1)E + e^{-2a}T.$$

If we pose $c = \cos(2z_K)$, $s = \sin(2z_K)$ and $b = e^{2a}$, we have to solve the system

$$\begin{cases} c - sz_N = -l & \text{(85.17a)}\\ -2cz_N - 2s = \dfrac{1}{b}(b^2l^2 - b^2 - 1) & \text{(85.17b)}\\ (c-1)z_N + s = \dfrac{1}{b} & \text{(85.17c)}\\ c^2 + s^2 = 1 & \text{(85.17d)}\end{cases}$$

with respect to c, s and z_N. One can check that the following is a solution:

$$c = \frac{b^2l^2 - 2b^2l + b^2 - 1}{b^2l^2 - 2b^2l + b^2 + 1} \tag{85.18a}$$

$$s = \frac{2b(1-l)}{b^2l^2 - 2b^2l + b^2 + 1} \tag{85.18b}$$

$$z_N = \frac{b^2(1-l^2) - 1}{2b}. \tag{85.18c}$$

If we pose $\bar{z}_K(a,l) = z_K(-a, -le^{2a})$ and $\bar{z}_N(a,l) = z_N(-a,-le^{2a})$, we have

$$\bar{z}_K(a,l) = \frac{1}{2}\arcsin\left(\frac{2e^{-2a}(1+le^{2a})}{l^2 + 2le^{-2a} + e^{-4a} + 1}\right) \tag{85.19}$$

$$2\bar{z}_N(a,l) = e^{-2a} - e^{2a}(l^2+1). \tag{85.20}$$

The principle of deformation by action of group leads us to deal with the function

$$f(a,l) = u(\bar{z}_K(a,l), \bar{z}_N(a,l)),$$

which should satisfies Unterberger's condition when u is compactly supported. Notice that z_K is a compact variable, so that u can be non vanishing for all values of z_K without violate the compact support requirement. The derivative of f with respect to a uses the chain rule, and it is apparent the higher order derivatives have to use the Leibniz formula:

$$\frac{\partial f}{\partial a}(a,l) = (\partial_1 u)(\bar{z}_K, \bar{z}_N)\frac{\partial \bar{z}_K}{\partial a}(a,l) + (\partial_2 u)(\bar{z}_K, \bar{z}_N)\frac{\partial \bar{z}_N}{\partial a}(a,l).$$

In order to give an idea of what is going on, here is the first derivative of $\bar{z}_K$ with respect to a:

$$(\partial_a \bar{z}_K)(a,l) = \frac{2e^{2a}}{e^{4a}l^2 + 2e^{2a}l + e^{4a} + 1}.$$

Let us look at the limit $a \to -\infty$ on the line $l = e^{-2a}$. If one performs multiple derivatives of $f(a,l)$ with respect to a, Leibniz rules yields a lot of terms of the form

$$(\partial_1^p \partial_2^q u)\big(\bar{z}_K(a,l), \bar{z}_N(a,l)\big)(\partial_a^i \bar{z}_K)(a,l)^j (\partial_a^k \bar{z}_N)(a,l)^m. \tag{85.21}$$ EqTermGeederrau

On the line $l = e^{-2a}$, the numerator of $(\partial_a^i \bar{z}_K)(a,l)$ is $(e^{4a}+4)^{2i}$ while the numerator is a sum and product of monomials of the form $(e^{4a} + N)$ with $N > 0$. At the limit, this factor in (85.21) goes to a finite number. The factor $(\partial_a^k \bar{z}_N)(a,l)$ is very different because

$$(\partial_a^k \bar{z}_N)(a,l) = (-1)^k 2^{k-1} e^{-2a} - 2^{k-1} e^{2a}(l^2+1).$$

which becomes

$$2^{k-1}e^{-2a}\big((-1)^k - 1\big) - 2^{k-1}e^{2a}$$

on $l = e^{-2a}$. It goes to zero when $a \to -\infty$ and k is even, but is goes to $-\infty$ at the same limit when k is odd. The highest divergence in all the terms of type (85.21) in $(\partial_a^n f)$ is expected for maximal m, so when $i = j = 0$. This is a divergence as

$$x \mapsto e^{2(n-1)x}.$$

Notice that this divergence increases when the order of derivative increases. Hence it contradicts Unterberger's condition which works with parameters r_1 and r_2 who are *constant* with respect to the order of the derivative.

85.2 Deformation of $\mathrm{SO}(2,n)$

SecUnifSOdn

85.2.1 Applying the extension lemma with old Iwasawa

One purpose of this section is to prove that

$$\mathrm{SO}(2,n) \overset{\mathcal{A}\oplus\mathcal{N}}{=} SU(1,1) \oplus SU(1,n-1).$$

For that purpose, we will decompose the Iwasawa algebra of $\mathfrak{so}(2,n)$ as a (symplectic) direct sum and we will compare the root spaces decomposition of each of the two parts with the ones of $\mathfrak{su}(1,1)$ and $\mathfrak{su}(1,n)$. Hence prescription on $\mathfrak{s}_1$ is to be two dimensional and to contains one and only one element of $\mathcal{A}$. This condition will impose a change of variable in the "new" Iwasawa.

The first point is to decompose the Lie algebra $\mathcal{R} := \mathcal{A} \oplus \mathcal{N}$ (from the Iwasawa decomposition of $\mathfrak{so}(2,n)$) into two parts $\mathfrak{s}_1$ and $\mathfrak{s}_2$ such that, as Lie algebras,

$$\mathcal{R} = \mathfrak{s}_1 \oplus_\rho \mathfrak{s}_2$$

where ρ is the adjoint action in $\mathfrak{so}(2, n)$. Recall that

$$\mathcal{A} \oplus \mathcal{N} = \{H_1, H_2, M, N, V_i, W_j\}.$$

From symplectic considerations which will appear further, we want $\mathfrak{s}_1$ and $\mathfrak{s}_2$ to be even dimensional. One can easily remark that

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, N\} \\ \mathfrak{s}_2 &= \{H_2, M, V_i, W_j\} \end{aligned} \tag{85.22}$$

works. What one has to check is that $\mathfrak{s}_1$ and $\mathfrak{s}_2$ are closed under ad: $[s, t] \in \mathfrak{s}_i$ if s, $t \in \mathfrak{s}_i$, and that $\mathfrak{s}_1$ acts on $\mathfrak{s}_2$: $[A, s] \in \mathfrak{s}_2$ if $A \in \mathfrak{s}_1$ and $s \in \mathfrak{s}_2$. This is a rather strong constraint.

Let us now explore systematically the possibilities. In a first time, we will not pay attention to the symplectic part. The question is to explicitly find all the possibilities of $\mathfrak{s}_i$ such that $\mathcal{A} \oplus \mathcal{N}$ is semi-direct product of $\mathfrak{s}_1$ and $\mathfrak{s}_2$.

As notational convention, when one write $E = \{W_i\}$, we mean that *all* the W_i are in the set E. If we want to say that *one particular* W_a is in E, we use the indices $a, b, c \ldots$.

If $H_1 \in \mathfrak{s}_2$, one can only finds H_2 and M in $\mathfrak{s}_1$ because in

$$\begin{aligned} \mathfrak{s}_1 &= \{W_i, N, V_i, \ldots\} \\ \mathfrak{s}_2 &= \{H_1, \ldots\}, \end{aligned} \tag{85.23}$$

the Lie algebra $\mathfrak{s}_1$ doesn't acts on $\mathfrak{s}_2$. So with $H_1 \in \mathfrak{s}_2$, the only possibility is

$$\begin{aligned} \mathfrak{s}_1 &= \{H_2, M\} \\ \mathfrak{s}_2 &= \{H_1, V_i, W_i, N\}. \end{aligned} \tag{85.24}$$

But $\mathfrak{s}_2$ is not closed for ad. First conclusion: $H_1 \in \mathfrak{s}_1$.

Let us consider the case $H_2 \in \mathfrak{s}_2$. In this case, one can only put H_1 and N in $\mathfrak{s}_1$: $H_2 \in \mathfrak{s}_2$ implies

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, N\} \\ \mathfrak{s}_2 &= \{H_2, M, V_i, W_j\}. \end{aligned} \tag{85.25}$$

From now, the question becomes "who can belong to $\mathfrak{s}_1$ in the same time as H_1¿"

The first step is to show that M can't. Let us consider H_1, $M \in \mathfrak{s}_1$. Since $[V_i, W_j] = \delta_{ij} M$, in order for $\mathfrak{s}_2$ to be closed for ad, one has to put some V_a and (or) W_b in $\mathfrak{s}_1$. If $W_a \in \mathfrak{s}_1$, V_a must also belongs to $\mathfrak{s}_1$ because $\mathfrak{s}_1$ must acts on $\mathfrak{s}_2$. For the same reason, $V_a \in \mathfrak{s}_A$ implies $W_a \in \mathfrak{s}_1$. Thus, $M \in \mathfrak{s}_1$ imply at least

$$\{H_1, M, V_i, W_j\} \subset \mathfrak{s}_1.$$

Now, it is also clear that $N \in \mathfrak{s}_2$ is not possible because of the action: $[W_i, N] = -2V_i$. We are left with

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, M, N, V_i, W_j\} \\ \mathfrak{s}_2 &= \{H_2\}, \end{aligned} \tag{85.26}$$

but we want even dimensional spaces. Conclusion: H_1, $M \in \mathfrak{s}_1$ is not possible.

Our second point is to show that $H_1, V_a \in \mathfrak{s}_1$ is also not possible. For, let us consider that one actually has $H_1, V_a \in \mathfrak{s}_1$, ant let us explore the consequences. It is clear that W_a and N can't be in $\mathfrak{s}_2$ in the same time. If $W_a \in \mathfrak{s}_1$, M must also be in $\mathfrak{s}_1$ in order to close under ad. But we just see that it was impossible. On the other hand, if N belongs to $\mathfrak{s}_1$, since $[N, W_a] = 2V_a \in \mathfrak{s}_1$, W_a can't belongs to $\mathfrak{s}_2$. Then we are left with the precedent case: $W_a \in \mathfrak{s}_1$. We conclude that $H_1, V_a \in \mathfrak{s}_1$ is not possible. From now, one sees that the only possibility with H_1, $N \in \mathfrak{s}_1$ is

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, N\} \\ \mathfrak{s}_2 &= \{H_2, M, V_i, W_j\}. \end{aligned} \tag{85.27}$$

The two last cases to explore are H_2 or (and) W_a in $\mathfrak{s}_1$ with H_1. If $W_a \in \mathfrak{s}_1$, then $H_2 \in \mathfrak{s}_1$ because of the action of $\mathfrak{s}_1$ on $\mathfrak{s}_2$ and $[W_a, H_2] = -W_a \in \mathfrak{s}_1$. On the other hand we had yet seen that $V_i \in \mathfrak{s}_2$. So we are left with

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, H_2, W_a, W_b\} \\ \mathfrak{s}_2 &= \{M, N, W_{\neq a,b}, V_i\}, \end{aligned} \tag{85.28}$$

and more generally, one can take in $\mathfrak{s}_1$ any even number of W_i.

One checks that the latest possibility (which is a special case of the previous) works:

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, H_2\} \\ \mathfrak{s}_2 &= \{M, N, W_i, V_i\}. \end{aligned} \tag{85.29}$$

The decomposition of $\mathcal{A} \oplus \mathcal{N}$ as semi-direct product of $\mathfrak{s}_1$ and $\mathfrak{s}_2$ are:

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, H_2, \underbrace{W_a, \dots, W_b}_{\text{even or zero}}\} \\ \mathfrak{s}_2 &= \{M, N, W_{\dots}, V_i\}, \end{aligned} \tag{85.30}$$

and

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, N\} \\ \mathfrak{s}_2 &= \{H_2, M, V_i, W_j\}. \end{aligned} \tag{85.31}$$

The symplectic conditions

Let us now turn our attention to the symplectic matter. There are two symplectic constraint on the choice of the $\mathfrak{s}_i$. The first one is that each one must be a **symplectic Lie algebra**: if Ω_i is the symplectic 2-form on $\mathfrak{s}_i$, then for any x, y, $z \in \mathfrak{s}_i$,

$$\Omega_i([x,y],z) + \Omega_i([y,z],x) + \Omega_i([z,x],y) = 0. \tag{85.32}$$ eq:symple_Lie

The second one is the fact that the action of $\mathfrak{s}_1$ on $\mathfrak{s}_2$ must be symplectic in the sense that $\forall X \in \mathfrak{s}_1$, $\operatorname{ad} X \in \mathfrak{sp}(\Omega_2)$. So one has to check that

$$\Omega_2\big(\operatorname{Ad}_s A, \operatorname{Ad}_s B\big) = \Omega_2(A, B) \tag{85.33}$$

for any $s \in S_1$ and A, $B \in \mathfrak{s}_2$. This fact was crucially used in the proof of proposition 83.22.

We first check that

$$\begin{aligned} \mathfrak{s}_1 &= \{H_1, H_2, W_a, \dots, W_b\} \\ \mathfrak{s}_2 &= \{M, N, W_{\dots}, V_i\} \end{aligned} \tag{85.34}$$

doesn't works. Indeed, one must have

$$\Omega_2\Big(e^{\operatorname{ad} H_1} M, e^{\operatorname{ad} H_1} N\Big) = \Omega_2(M, N),$$

but $[H_1, M] = 0$, $[H_1, N] = 2N$ and $\exp(\operatorname{ad} H_1) = id + [H_1, .] + \dots$, then

$$\Omega_2\Big(e^{\operatorname{ad} H_1} M, e^{\operatorname{ad} H_1} N\Big) = \Omega_2(M, e^2 N),$$

so that $\Omega_2(M, N) = 0$. Note that this is more general: $\Omega_2(M, s) = 0$ for all $s \in \mathfrak{s}_2$ such that $[H_1, s] = \alpha s$ because

$$\Omega_2(M, s) \stackrel{!}{=} \Omega_2(e^{\operatorname{ad} H_1} M, e^{\operatorname{ad} H_1} s) = \Omega_2(M, e^{\alpha} s).$$

This imposes

$$\begin{aligned} \Omega_2(M, V_i) &= 0 \\ \Omega_2(M, N) &= 0 \\ \Omega_2(M, W_i) &= 0. \end{aligned} \tag{85.35}$$

Thus Ω_2 is degenerate. Now, we check that the second works:

$$\begin{aligned}\mathfrak{s}_1 &= \{H_1, N\}\\ \mathfrak{s}_2 &= \{H_2, M, V_i, W_j\}.\end{aligned} \tag{85.36}$$

The condition (85.32) gives

$$\begin{aligned}\Omega_2(V,M) &= 0\\ \Omega_2(W,M) &= 0\\ 2\Omega_2(V,W) + \Omega_2(M,H_2) &= 0.\end{aligned} \tag{85.37}$$

When $s \in \mathfrak{s}_2$ is such that $[H_1, s] = \alpha s$, $\Omega_2(H_2, s) = 0$. Now the matrix of Ω_2 looks like

$$\Omega_2 = \left(\begin{array}{c|c|c|c|c} & H_2 & M & V_i & W_j \\ \hline H_2 & 0 & & 0 & 0 \\ \hline M & & 0 & 0 & 0 \\ \hline V_i & 0 & 0 & 0 & \\ \hline W_j & 0 & 0 & & 0 \end{array}\right) \tag{85.38}$$

We have to check that the last entries keep free.

$$\begin{aligned}\Omega_2(M,H_2) &\overset{!}{=} \Omega_2\Big(e^{\operatorname{ad} H_1} M, e^{\operatorname{ad} H_1} H_2\Big)\\ &\overset{!}{=} \Omega_2\Big(e^{\operatorname{ad} N} M, e^{\operatorname{ad} N} H_2\Big).\end{aligned} \tag{85.39}$$

Since $[H_1, M] = [H_1, H_2] = [N, M] = [N, H_2] = 0$, these two conditions are fulfilled. On the other hand,

$$\begin{aligned}\Omega_2(V_a, W_b) &\overset{!}{=} \Omega_2\Big(e^{\operatorname{ad} H_1} V_a, e^{\operatorname{ad} H_1} W_b\Big)\\ &= \Omega_2(eV_A, e^{-1} W_b)\\ &= \Omega_2(V_a, W_b)\\ &\overset{!}{=} \Omega_2\Big(e^{\operatorname{ad} N} V_a, e^{\operatorname{ad} N} W_b\Big)\\ &= \Omega_2(V_a, W_b + 2V_b).\end{aligned} \tag{85.40}$$

Thus $\Omega_2(V_a, W_b)$ has no constraints and $\Omega_2(V_a, V_b) = 0$. Finally, it is easy to see that

$$\begin{aligned}\Omega_2(W_a, W_b) &\overset{!}{=} \Omega_2\Big(e^{\operatorname{ad} H_1} W_a, e^{\operatorname{ad} H_1} W_b\Big)\\ &= e^{-2}\Omega_2(W_a, W_b),\end{aligned} \tag{85.41}$$

so that $\Omega_2(W_a, W_b) = 0$. Now, one can write the possible form for Ω_2 as

$$\Omega_2 = \left(\begin{array}{c|c|c|c|c|c|c} & H_2 & M & V_1 & V_2 & W_1 & W_2 \\ \hline H_2 & 0 & a & 0 & 0 & 0 & 0 \\ \hline M & -a & 0 & 0 & 0 & 0 & 0 \\ \hline V_1 & 0 & 0 & 0 & 0 & a/2 & a/2 \\ \hline V_2 & 0 & 0 & 0 & 0 & a/2 & a/2 \\ \hline W_1 & 0 & 0 & -a/2 & -a/2 & 0 & 0 \\ \hline W_2 & 0 & 0 & -a/2 & -a/2 & 0 & 0 \end{array}\right) \tag{85.42}$$

85.2.2 Decomposition as split extension

When we try to decompose $\mathcal{A} \oplus \mathcal{N}$ as symplectic direct sum with a strict respect to chosen basis matrices, there are only two possibilities:

$$\begin{aligned}\mathfrak{s}_1 &= \{J_1, J_2, \overbrace{V_a, \dots, V_b}^{\text{even}}\}\\ \mathfrak{s}_2 &= \{L, M, W_i, V_{\dots}\}\end{aligned} \tag{85.43}$$

and

$$\begin{aligned}\mathfrak{s}_1 &= \{J_2, \overbrace{V_a, \dots, V_b}^{\text{odd}}\}\\ \mathfrak{s}_2 &= \{J_1, W_j, M, L, V_{\dots}\}.\end{aligned} \tag{85.44}$$

In the first possibility, $\mathfrak{s}_1$ is not symplectic (because it is abelian); while the second one only works in low dimensional cases: there must not be any V_a. This fact leads us to consider the change of basis (66.126) in $\mathcal{A}$: $H_1 = J_1 - J_2$ and $H_2 = J_1 + J_2$.

If $H_1 \in \mathfrak{s}_2$, then $L, V_i, W_i \in \mathfrak{s}_2$ because $\mathfrak{s}_1$ must act on $\mathfrak{s}_2$. Hence $M \in \mathfrak{s}_2$ and H_2 remains alone in $\mathfrak{s}_1$. That proves that $H_1 \in \mathfrak{s}_1$. If we suppose that $H_2 \in \mathfrak{s}_2$, we find

$$\begin{aligned}\mathfrak{s}_1 &= \{H_1, L\}\\ \mathfrak{s}_2 &= \{H_2, V_i, W_j, M\}.\end{aligned} \tag{85.45}$$

The case $H_1, H_2 \in \mathfrak{s}_1$ leads to

$$\begin{aligned}\mathfrak{s}_1 &= \{H_1, H_2, \overbrace{V_a, \dots V_b}^{\text{even}}\}\\ \mathfrak{s}_2 &= \{M, L, W_i, V_{\text{others}}\}.\end{aligned} \tag{85.46}$$

The symplectic condition excludes the second decomposition. Indeed for each s such that $[H_1, s] = \alpha s$ (i.e. $s = V_i, W_j, L$), we have

$$\Omega_2\big(e^{adH_1}M, e^{\operatorname{ad} H_1}s\big) = e^{\alpha}\Omega_2(M,s) \overset{!}{=} \Omega_2(M,s).$$

Hence $\Omega_2(M,s) = 0$. This proves that the decomposition (85.46) imposes the symplectic form Ω_2 to be degenerate. We are left with decomposition (85.45).

Root space decomposition of $SU(1,n)$ can be found on pages 314–315 of [865]: it has $\dim \mathcal{A} = 1$, $\dim \mathcal{G}_{2f} = 1$ and $\dim \mathcal{G}_f = 2(n-1)$. In $\mathfrak{s}_2$, we have $V_i \in \mathcal{G}_1$, $W_j \in \mathcal{G}_1$, $M \in \mathcal{G}_2$, and when we look at $AdS_l = \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1)$, we have $l-3$ matrices V_i and W_j. Therefore $\mathfrak{s}_2$ is nothing else than the $\mathcal{A} \oplus \mathcal{N}$ of $\mathfrak{su}(1, l-2)$ (recall $l \geqslant 3$). The analysis shows that $\mathfrak{s}_1$ is the $\mathcal{A} \oplus \mathcal{N}$ of $\mathfrak{su}(1,1)$.

85.2.3 Conclusion and perspectives

For our AdS_l black hole, the algebra of the group which defines the singularity is the split extension

$$(\mathcal{A} \oplus \mathcal{N})_{\mathfrak{so}(2,l-1)} = (\mathcal{A} \oplus \mathcal{N})_{\mathfrak{su}(1,1)} \oplus_{\text{ad}} (\mathcal{A} \oplus \mathcal{N})_{\mathfrak{su}(1,l-2)}.$$

A deformation of the corresponding groups is given in [1039]. The extension lemma 83.23 yields an oscillatory integral universal deformation formula for proper actions of the Iwasawa subgroup of $\mathrm{SO}(2, l-1)$. That remark provides an alternative way to deform the black hole to the one presented in section 84.1.

The availability of a quantization of AdS_l by action of AN is an opportunity to embed our black hole toy model in the framework of noncommutative geometry. Indeed, the quantization of AdS_l is the data of the anti de Sitter manifold and the action of the group AN; that is precisely the data which defines the black hole of chapter 79. So we would be able to "see" the causal issue from the data of the deformed spectral triple. Remark that a causal structure (in the physical meaning of the term) is a special property of *pseudo*-Riemannian manifolds for which spectral geometry does not exist yet.

An important remaining problem with that method is the fact that the extension lemma does not assure the existence of a stable functional space for the new product. So there is still a lot of analytic work to be done.

Chapter 86

Gravitation and noncommutative geometry

References about gravitation and noncommutative geometry are [962, 819, 953, 941]. About the spectral action, one can consult [1043], with a prediction for the Higgs mass! I also would like to give the reference [1044] arguing that topos are the right direction to deal with quantum gravity.

86.1 Inner and outer automorphisms

Let M be a paracompact smooth manifold. When $\mathfrak{A}$ is any unital $*$-algebra (as $C^\infty(M)$), we define $\mathrm{Aut}(\mathfrak{A})$ as the set of maps $\alpha\colon \mathfrak{A}\to\mathfrak{A}$ such that

$$\alpha(ab)=\alpha(a)\alpha(b),\qquad \alpha(\mathbb{1})=\mathbb{1}\qquad \alpha(a^*)=\alpha(a)^*. \tag{86.1}$$

Proposition 86.1.
The group $\mathrm{Diff}(M)$ *of diffeomorphism of* M *is diffeomorphic to the group* $\mathrm{Aut}\big(C^\infty(M)\big)$. *The isomorphism is given by* $\varphi\mapsto\alpha_\varphi$ *with*

$$\alpha_\varphi(f)(x)=f\big(\varphi^{-1}(x)\big).$$

Proof. No proof. □

For each $u\in U(\mathfrak{A})$, we define the **inner automorphism** associated with u by

$$\alpha_u(a)=uau^*. \tag{86.2}$$

It satisfies $\alpha_{u^*}\circ\alpha_u=\alpha_u\circ\alpha_{u^*}=\mathrm{Id}$. We denote by $\mathrm{Inn}(\mathfrak{A})$ the subgroup of inner automorphisms:

$$\mathrm{Inn}(\mathfrak{A})=\{\alpha_u \text{ tel que } u\in U(\mathfrak{A})\}.$$

One can show that $\mathrm{Inn}(\mathfrak{A})$ is a normal subgroup of $\mathrm{Aut}(\mathfrak{A})$: for every $u\in U(\mathfrak{A})$ and $\varphi\in\mathrm{Aut}(\mathfrak{A})$, the automorphism $\varphi^{-1}\circ\alpha_u\circ\varphi$ belongs to $\mathrm{Inn}(\mathfrak{A})$. So we can define the group of **outer automorphism** as

$$\mathrm{Out}(\mathfrak{A})=\mathrm{Aut}(\mathfrak{A})/\mathrm{Inn}(\mathfrak{A}),$$

and we have the short exact sequence

$$\mathbb{1}_{\mathrm{Aut}(\mathfrak{A})}\longrightarrow\mathrm{Inn}(\mathfrak{A})\longrightarrow\mathrm{Aut}(\mathfrak{A})\longrightarrow\mathrm{Out}(\mathfrak{A})\longrightarrow\mathbb{1}_{\mathrm{Inn}(\mathfrak{A})}. \tag{86.3}$$

In the case of a commutative algebra, $\mathrm{Inn}(\mathfrak{A})$ reduces to identity, so that $\mathrm{Aut}(\mathfrak{A})=\mathrm{Out}(\mathfrak{A})$.

Proposition 86.2.
In the particular case of $\mathfrak{A}=C^\infty(M)$ *for a smooth manifold* M, *we have*

$$\mathrm{Aut}\big(C^\infty(M)\big)=\mathrm{Out}\big(C^\infty(M)\big)=\mathrm{Diff}(M).$$

Proof. By proposition 72.14, every character of $C^\infty(M)$ is of the form ω_x for a certain $x \in M$. If α is an automorphism of $C^\infty(M)$, then $\alpha^{-1}(\omega_x)$ defined by

$$\alpha^{-1}(\omega_x)f = \omega_x\big(\alpha^{-1}(f)\big)$$

is still a character. Thus, there exists a $y \in M$ such that $\alpha^{-1}(\omega_x) = \omega_y$. That correspondence provides a continuous bijection $\phi\colon M \to M$ defined by

$$\alpha^{-1}(\omega + x) = \omega_{\phi(x)}.$$

Applying a function $f \in C^\infty(M)$ to that equality, we find

$$\alpha^{-1}(f)(x) = f\big(\phi(x)\big),$$

or

$$\alpha(f)(x) = f\big(\phi^{-1}(x)\big). \tag{86.4}$$

Using the chain rule on that expression and the fact that α and f are smooth, we find that ϕ is itself smooth. What we found is that $\alpha \leftrightarrow \phi$ is a group isomorphism between $\mathrm{Aut}\big(C^\infty(M)\big)$ and $\mathrm{Diff}(M)$. □

Since $\mathrm{Out}(\mathfrak{A})$ is the group of diffeomorphism in the commutative case, we want the remaining ($\mathrm{Inn}(\mathfrak{A})$ in the noncommutative case) to be the group of internal automorphisms of the theory, namely the gauge transformations.

Of course, taking a diffeomorphism in general changes the metric, so that we have to use a principle to choose the correct physical metric of the theory. That principle is the minimizing of an action, namely the Einstein-Hilbert action

$$S_{EH} = \int_M r(x)\sqrt{g(x)}dx$$

where r is the scalar curvature up to some constant factor. In the case of a boson theory, the correct action reveals to be the Yang-Mills one

$$S_{YM} = \int F \star F$$

where F is a curvature form.

Let $(\mathfrak{A}, \mathscr{H}, \pi, D, J)$ be a real spectral triple (the representation π of $\mathfrak{A}$ on $\mathscr{H}$ is explicitly written). If α_u is an element of $\mathrm{Inn}(\mathfrak{A})$, we define a new representation of $\mathfrak{A}$ on $\mathscr{H}$ by

$$\pi_u = \pi \circ \alpha_u,$$

and we have the adapted Dirac operator

$$D_u = D + A + \epsilon JAJ^* \tag{86.5}$$

where $A = u[D, u^*]$. That Dirac operator is "adapted" in the sense of the following lemma.

Lemma 86.3.
For every unitary element u of $\mathfrak{A}$, the spectral triples $(\mathfrak{A}, \mathscr{H}, \pi, D, J)$ and $(\mathfrak{A}, \mathscr{H}, \pi_u, D_u, J)$ are equivalent and the unitary operator which provides the equivalence is $U = uJuJ^*$ where

$$D_u = D + u[D, u^*] \pm Ju[D, u^*]J^*$$

with the sign $\pm$ is taken as $-$ if and only if $n = 1 \mod 4$.

Proof. No proof. □

More generally a **inner fluctuation** of the metric a transformation of the form

$$D \mapsto D_A = D + A + JAJ^*$$

where A is any gauge potential. The **spectral action principle** claims that the gravitation coupled with the inner degrees of freedom is given by the action

$$S_B(D, A) = \mathrm{Tr}_{\mathscr{H}}\, \chi\left(\frac{D_A^2}{\Lambda^2}\right) \tag{86.6}$$ EqBozSpectralAct

where Tr is the usual trace on the Hilbert space $\mathscr{H}$, Λ is a mass scale parameter and χ is a function which cuts the eigenvalues of D_A^2 larger than Λ^2. Here the index B refers to "boson".

Proposition 86.4.
The spectral action (86.6) remains unchanged under inner fluctuations

$$A \mapsto A^u = uAu^* + u[D, u^*]$$

for every $u \in \mathrm{Inn}(\mathfrak{A})$.

Proof. No proof. □

That lemma has the consequence that $A + JAJ^* = 0$, so that there are no inner fluctuation of the metric in the case of commutative triple. If $A = \sum_j a_j[D, b_j]$, using the fact that $JJ^* = \mathbb{1}$ and $J[D, b_j]J^* = [D, Jb_jJ^*]$, we find

$$JAJ^* = \sum_j Ja_jJ^*[D, Jb_jJ^*],$$

and with the identification $JaJ^* = a_j^*$, we find $JAJ^* = -A^*$, so that the commutative case does not have inner fluctuations.

¡¡ Avertissement/question à la lectrice !! 86.5
Here we have JAJ^* acting on the right as A^*. Maybe because of something like $Ja = a^*$, so that

$$JAJ^*a = JA(a^*) = J(Aa^*) = aA^*.$$

This makes the action of JAJ^* commute with the one of B which is on the left.

We know that the right representation of $\mathfrak{A}$ on $\mathscr{H}$ is defined by $\xi b = b^0\xi$. Now, for $u \in U(\mathfrak{A})$ we define the **adjoint representation** of u by

$$\mathrm{Ad}(u)\xi = u\xi u^* \tag{86.7}$$

Proposition 86.6.
Let $(\mathfrak{A}, \mathscr{H}, D, J)$ be a real spectral triple. For all gauge potential $A \in \Omega^1_D(\mathfrak{A})$ and for all $u \in U(\mathfrak{A})$, we have

$$\mathrm{Ad}(u)(D + A + \epsilon' JAJ^{-1})\,\mathrm{Ad}(u^*) = D + \gamma_u(A) + \epsilon' J\gamma_u(A)J^{-1} \tag{86.8}$$

where $\gamma_u(A) = u[D, u^*] + uAu^*$.

Proof. No proof. □

Moreover a inned fluctuation of an inner fluctuation is an inner fluctuation by the following proposition.

Proposition 86.7.
Let $(\mathfrak{A}, \mathscr{H}, D)$ be a spectral triple and $D_A = D + A$ for some potential vector A. Then for every $B \in \Omega^1_{D_A}(\mathfrak{A})$ with $B = B^*$ (i.e. for every potential vector with respect to the new Dirac operator D_A) we have

$$D_A + B = D + A' \tag{86.9}$$

with $A' = A + B \in \Omega^1_D(\mathfrak{A})$.

Remark that the inner fluctuation $D \mapsto D+A+\epsilon' JAJ^{-1}$ has to be seen as a two times process. First, we have the fluctuation $D \mapsto D+A$ which is a fluctuation relative to $\mathfrak{A}$ and then we have the fluctuation $D+A \mapsto D+A+\epsilon' JAJ^{-1}$ which is a fluctuation relative to the action of $\mathfrak{A}^0 = J\mathfrak{A}J^*$. The result $D' = D + A + \epsilon' JAJ^{-1}$ is the most general operator such that $JD'J^{-1} = \epsilon D'$.

86.2 Spectral action

One speaks about spectral action and its applications to physics in [1045, 1043, 1046]

The only observables quantities in gravitation are the ones which are invariant under the full diffeomorphism group (that is the principle of gauge invariance). For example, if K is the scalar curvature, and F any function, the quantity

$$\int_M F(K)\sqrt{g}dx$$

is an observable. That quantity is of course almost impossible to observe in real world experiments because it needs a knowledge of curvature on the hole space. That observable is highly non local.

The **spectral action principle** claims that the action to be used at the level of the functional intergral for quantum gravity after Wick rotation is given by

$$\operatorname{Tr} f\left(\frac{D}{\Lambda}\right) \tag{86.10}$$

where D is the Dirac operator, f is an even positive function and Λ a real parameter.

86.3 Gauge theory

In order to build a gauge theory over M for the group G, we begin by building a G-principal bundle over M and we define $\operatorname{Aut}(P)$ as the group of automorphisms of P, i.e. the diffeomorphisms of P which commute with the action of G. We consider G_V, the group of vertical automorphism of P. We have the short exact sequence

$$\mathbb{1} \longrightarrow G_V \longrightarrow \operatorname{Aut}(P) \longrightarrow \operatorname{Diff}(M) \longrightarrow \mathbb{1} \tag{86.11}$$

where the arrow $\alpha\colon \operatorname{Aut}(P) \to \operatorname{Diff}(M)$ is defined by $\varphi \mapsto \alpha_\varphi$,

$$\alpha_\varphi(x) = \pi\big(\varphi(\pi^{-1}x)\big)$$

which is well defined because $\pi(\xi) = \pi(\xi')$ implies $\pi(\varphi\xi) = \pi(\varphi\xi')$. The map φ belongs to the kernel of that arrow when $\alpha_\varphi(x) = x$ for every x. That map φ must satisfies $\varphi(\pi^{-1}x) \in \pi^{-1}x$, so that $\varphi \in G_V$. Now the way to build a gauge theory for the group G in the noncommutative geometry setting is as follows

(1) we search for an algebra $\mathfrak{A}$ such that $\operatorname{Inn}(\mathfrak{A}) \simeq G_V$,

(2) we build a spectral triple over $\mathfrak{A}$,

(3) we compute the spectral action $S_B(D, A)$.

Let M be a Riemannian spin 4-dimensional manifold and $(\mathfrak{A}, \mathscr{H}, D)$ be its spectral triple.

Proposition 86.8.
The spectral action

$$S_G(D, \Lambda) = \operatorname{Tr}_{\mathscr{H}} \chi\left(\frac{D^2}{\Lambda^2}\right))$$

only depends on the spectrum of D.

Proof. No proof. □

We denote by $\mathrm{Spec}(M,D)$ the spectrum of the Dirac operator with eigenvalues repeated as many times as their multiplicity. Two manifolds are said to be **isospectral** when $\mathrm{Spec}(M,D) = \mathrm{Spec}(M',D')$. The proposition says that the spectral action is a spectral invariant. Notice that there exists isospectral manifold that are not isometric, so *one cannot heard the shapes of a drum*. If D and D' are two isospectral Dirac operators corresponding to different metrics g and g', we have

$$S_B(D) = S_B(D'),$$

but

$$S_{EH}(g) \neq S_{EH}(g').$$

Thus the spectral invariance which leads to the spectral action S_B is stronger than the usual diffeomorphism invariance which leads to the Einstein-Hilbert action S_{EH}.

86.4 Scalar curvature

Let $(\mathfrak{A},\mathscr{H},D)$ be a spectral triple of metric dimension 4. We define the **scalar curvature** as the function over $\mathfrak{A}$

$$\mathrm{R}(a) = \fint aD^{-2}. \tag{86.12}$$ EqDefScalCurTriple

That definition is motivated by the following theorem.

Theorem 86.9.
Let $(\mathfrak{A},\mathscr{H},D)$ be the spectral triple of a 4-dimensional spin Riemannian manifold. Then

$$\mathrm{R}(f) = \frac{1}{24\pi^2}\int_M f(x)s(x)\sqrt{g}dx$$

where $s = -R$ is the scalar curvature.

Let us now study how does the curvature (86.12) changes under fluctuations of the form $D \mapsto D + A$. In the expansion of $\mathrm{Tr}\left(f(D/\Lambda)\right)$, the coefficient of Λ^2 is given by $\fint |D|^{-2}$, so that we have to look at the integral $\fint (D+A)^2$.

Proposition 86.10.
If we suppose that

— $\fint AD^{-3} = 0$ *for all* $A \in \Omega^1_D(\mathfrak{A})$,
— $\mathrm{R}(a) = \fint ads^2$ *is a trace on* $\mathfrak{A}$,

then

(1) the term $\fint (D+A)^{-2}$ *is independent of* A,
(2) the functional R *is invariant under the inner fluctuation* $D \mapsto D + A$.

86.5 Boundary conditions and heat kernel expansion

SecGravBoundCC

86.5.1 Example

As example of using the spectral action with heat kernel expansions and formulas (67.17), let us perform the "warm up" that Connes and Chamseddine propose in [965], this is the usual Dirac operator

$$D = \gamma^i(\partial_i + \omega_i) \tag{86.13}$$ EqFormgenDiracD

Let (M,g) be a compact four dimensional Riemannian manifold and V, a vector bundle which carries a Clifford module structure, i.e. a map $\gamma\colon \mathrm{Cl}(M) \to \mathrm{End}(V)$. If $\{e_i\}$ is a local basis of V, we define $\gamma_i = \gamma(e_i)$ where we choose the basis $\{e_i\}$ in such a way that γ_i are constant, cf lemma 68.40. In that case, we pose $\nabla_i = \partial_i + \omega_i$ and we have $D = \gamma^i\nabla_i$ as well as

$$D^2 = \gamma^i\gamma^j\nabla_i\nabla_j.$$

In order to make the link with the general Dirac type operator (62.126), we notice that what we do here is only to put ourself in a case where the element r of $\Gamma(\mathrm{End}\, V)$ reads $\omega_i\gamma^i$ with $\omega_i \in C^\infty(M)$.

The computation of $[\nabla_i, \nabla_j]$ with $\nabla_i = \partial_i + \omega_i$ produces

$$[\nabla_i, \nabla_j] = (\partial_i\omega_j) - (\partial_j\omega_i)$$

where $(\partial_i\omega_j)$ denotes the multiplication operator by the function $\partial_i\omega_j$. Thus we have

$$\begin{aligned} D^2 &= \gamma^i\gamma^j\nabla + i\nabla_j \\ &= (-2g^{ij} - \gamma^j\gamma^i)(\partial_i\omega_j - \partial_j\omega_i + \nabla_j\nabla_i) \\ &= -2g^{ij}\nabla_j\nabla_i - \gamma^j\gamma^i(\partial_i\omega_j - \partial_j\omega_i) - \underbrace{\gamma^j\gamma^i\nabla_j\nabla_i}_{=D^2}. \end{aligned} \tag{86.14}$$

Thus we have

$$D^2 = -g^{ij}\nabla_i\nabla_j - \frac{1}{2}\gamma^i\gamma^j(\partial_i\omega_j - \partial_j\omega_i). \tag{86.15}$$

This is a general form for the square of operator of the form (86.13). The operator we look at now is the one given by a form ω which satisfies the differential equation

$$\gamma^i\gamma^j(\partial_i\omega_j - \partial_j\omega_i) = \frac{R}{2}, \tag{86.16}$$

so that $D^2 = -(g^{ij}\nabla_i\nabla_j - \frac{1}{4}R)$, which immediately gives us $E = -\frac{1}{4}R$ in the formulas (67.17). Using proposition 67.4 on our present Dirac operator (here, $\phi = 0$) we find

$$S = -\frac{1}{2}\Pi_+\big(\gamma_n\gamma_a\nabla_a\chi\big)\Pi_+. \tag{86.17}$$

We also have $\nabla = \nabla'$ and $\Phi = 0$ as well as $\nabla'_a\chi = K_{ab}\chi\gamma^n\gamma^b$ where K_{ab} is the **extrinsic curvature** defined by

$$K_{ab} = \Gamma^n_{ab}.$$

We also define the scalar quantity $K = h^{ab}K_{ab}$ and it is easy to see that $\gamma^a\gamma^bK_{ab} = -K$. Using these formulas,

$$\begin{aligned} S &= -\frac{1}{2}\Pi_+(\gamma_n\gamma^a\nabla'_a\chi)\Pi_+ = -\frac{1}{2}(\gamma_n\gamma^aK_{ab}\chi\gamma^n\gamma^b)\Pi_+ \\ &= \frac{1}{2}\Pi_+(\gamma_n\gamma^aK_{ab}\gamma^b\chi\gamma^n)\Pi_+ = -\frac{1}{2}\Pi_+(K\gamma_n\chi\gamma^n)\Pi_+ \\ &= \frac{1}{2}\Pi_+K\gamma_n\gamma^n\chi = -\frac{1}{2}\Pi_+K\chi\Pi_+, \end{aligned}$$

where we used the commutation relations $\gamma^n\gamma^b = -\gamma^b\gamma^n$, $\chi\gamma^b = \gamma^b\chi$, $\chi\gamma^n = \gamma^n\chi$ and $\gamma_n\gamma^n = -1$. Using now the definition $\Pi_+ = \frac{1}{2}(\mathrm{Id} + \chi)$ and the property $\chi^2 = \mathrm{Id}$, we conclude

$$S = -\frac{1}{2}K\Pi_+. \tag{86.18}$$

Since we are working on a four dimensional space, $\mathrm{Tr}(\mathrm{Id}) = 4$, and $\mathrm{Tr}(S) = -\frac{1}{2}J\,\mathrm{Tr}(\Pi_+) = -K$ because Π_+ is a projector on a space which has half of the dimension of the total space and K is a scalar. For the same reason, $\mathrm{Tr}(E) = -\frac{1}{4}\,\mathrm{Tr}(R) = -\frac{1}{4}R\,\mathrm{Tr}(\mathrm{Id}) = -R$. The formulas (67.17) reduce to

$$a_0(\Delta, \chi) = \frac{1}{4\pi^2}\int_M \sqrt{g}d^4x, \tag{86.19a}$$

$$a_2(\Delta, \chi) = -\frac{1}{48\pi^2}\left(\int_M R\sqrt{g}d^4x + 2\int_{\partial M} K\sqrt{h}d^3x\right). \tag{86.19b}$$

We are interested in the following case.

ThoExpActSpect

Theorem 86.11.
Let $(\mathfrak{A}, \mathscr{H}, D)$ be a spectral triple such that the expansion (67.7) holds. Then the spectral action $S = \mathrm{Tr}\left(f(D/\Lambda)\right)$ can be expanded with respect to Λ in the following way:

$$\mathrm{Tr}\left(f(D/\Lambda)\right) \sim \sum_{\beta\in\Pi} f_\beta \Lambda^\beta \fint |D|^{-\beta} + f(0)\zeta_D(0) + \dots \tag{86.20}$$

where the sum is taken over the dimension spectrum of the triple. The function f_β is given by

$$f_\beta = \int_0^\infty f(v) v^{\beta-1} dv. \tag{86.21}$$

Eqfbetaintdonne

Notice that from the definition of the dimension spectrum, we have $\mathrm{Re}(\beta) \geqslant 0$ and the negatives power involve all the Taylor expansion of f at zero.

Proof. Let us act on a test function $k(u)$ that we expand as

$$k(u) = \int_0^\infty e^{-su} h(s) ds,$$

and we formally write $k(t\Delta) = \int_)^\infty e^{-st\Delta} h(s) ds$, so that from the hypothesis we have

$$\mathrm{Tr}\left(k(t\Delta)\right) \sim \sum_{alpha} t^\alpha \int_0^\infty s^\alpha h(s) ds. \tag{86.22}$$

When $\alpha < 0$, we have the formula

$$s^\alpha = \frac{1}{\Gamma(-\alpha)} \int_0^\infty e^{-sv} v^{-\alpha-1} dv,$$

so that

$$\int_0^\infty s^\alpha h(s) ds = \frac{1}{\Gamma(-\alpha)} \int_0^\infty \int_0^\infty e^{-sv} v^{-\alpha-1} dv\, h(s) ds = \frac{1}{\Gamma(-\alpha)} \int_0^\infty k(v) v^{-\alpha-1} dv.$$

That shows the asymptotic expansion

$$\mathrm{Tr}\left(k(t\Delta)\right) \sim \sum_\alpha a_\alpha t^\alpha \frac{1}{\Gamma(-\alpha)} \int_0^\infty k(v) v^{-\alpha-1} dv. \tag{86.23}$$

EqTrktDeldv

By result (67.9), the term $\alpha = 0$ reduces to

$$\zeta_D(0) \int_0^\infty k(v) v^{-1} dv.$$

¡¡ Avertissement/question au lecteur !! 86.12
Cela devrait être $\zeta_D(0) f(0)$, mais je ne vois pas trop comment ...

Using now the fact that $2a_\alpha/\Gamma(-\alpha) = \mathrm{Res}_{s=-2\alpha} \zeta_D(s)$, expression (86.23) becomes

$$\mathrm{Tr}\left(k(t\Delta)\right) \sim \sum_\alpha t^\alpha \frac{1}{2} \mathrm{Res}_{s=-2\alpha} \zeta_D(s) \int_0^\infty k(v) v^{-\alpha-1} dv. \tag{86.24}$$

EqinterExptrbiftD

Let us study the terms with $\alpha > 0$ in the expansion (86.24). The summation set in the expansion (67.7) can be very big, but equation (86.24) shows that only the α on which we have pole are to be taken into account, so from definition 80.25 of dimension spectrum, the sum over α reduces to a sum over $\alpha \in \Pi$.

If we pose $s' = s + 2\alpha$, we have

$$\mathrm{Res}_{s=-2\alpha} \mathrm{Tr}\left(|D|^{-s}\right) = \mathrm{Res}_{s'=0} \mathrm{Tr}\left(|D|^{2\alpha}|D| - s'\right) = \fint |D|^{2\alpha}$$

If we pose $\beta = -2\alpha$, we find $\fint |D|^{-\beta}$.

□

Let us summarize what we obtained up to here. Let D be an elliptic operator on M and suppose $\partial M = \varnothing$. If we pose $\Delta = D^2$, the Mellin transform provides the following relation between $\mathrm{Tr}(\Delta^{-z})$ (with $z \in \mathbb{C}$) and $\mathrm{Tr}(e^{-t\Delta})$:

$$\mathrm{Tr}(\Delta^{-z})\Gamma(z) = \frac{1}{\sqrt{\pi}} \int_0^\infty \mathrm{Tr}(e^{-t\Delta}) t^{z-1} dt. \tag{86.25}$$

EaIntDevuspi

We know an heat kernel expansion for $\mathrm{Tr}(e^{-t\Delta})$ and we know, from theorem 86.11 (expression (86.21)), the values of a_α when $\mathrm{Tr}(e^{-tD^2}) \sim \sum_\alpha a_\alpha t^\alpha$. So we know what to put into the integral 86.25.

86.6 Boundary

The spectral action principle is given by $S = \mathrm{Tr}\left(f(D/\Lambda)\right)$ with an even function $f\colon \mathbb{R} \to \mathbb{R}$ such that $f(0) = 1$. For practical computations, the simplest is to take $f(x) = e^{-x^2}$ and to male an heat kernel expansion:

$$\mathrm{Tr}\left(f(D/\Lambda)\right) \sim \Lambda^{(n/2)} Vol(M) + \Lambda^{(n/2)-1} \int_M r(x)dx + \Lambda^{(n/2)-2} \dots$$

If the manifold M has boundary, we pose the **elliptic conditions**

$$\mathrm{dom}(D) = \{s \text{ tel que } Ts|_{\partial M} = s|_{\partial M}\} \tag{86.26}$$

where $T = \gamma c(n)$. Here γ is the parity and $c(n)$ is the multiplication by the normal vector to the boundary. In particular $T^2 = \mathrm{Id}$. When such conditions are imposed, one still have an heat kernel expansion under the form

$$\mathrm{Tr}\left(f(D/\Lambda)\right) \sim \Lambda^{(n/2)} Vol(M) \tag{86.27}$$

$$+ \Lambda^{(n-1)/2} \cdot 0 \qquad \text{this term is zero for some reasons} \tag{86.28}$$

$$+ \Lambda^{(n/2)/2} \int_M \text{scalar curvature} \tag{86.29}$$

$$+ \Lambda^{(n-3)/2} \int_{\partial M} \text{mean curvature.} \tag{86.30}$$

One can prove that the two latter terms are divergent on an asymptotically flat space.

Chapter 87

Levy Processes and such

Source: [1047].

87.1 Lévy process

Let G be a semigroup and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. Here $\mathcal{F}$ is a σ-algebra and $\mathbb{P}$ is a probability measure on Ω.

Definition 87.1.
*A **Lévy process** is, for each s and t such that $0 \leqslant s \leqslant t < \infty$, a map $X_{st} \colon (\Omega, \mathcal{F}, \mathbb{P}) \to (G, \mathcal{B})$ where $\mathcal{B}$ is the Borel σ-algebra on G such that*

(1) $\mathbb{P}(X_{ss} = e) = 1$,

(2) $X_{ss} = \lim_{t \searrow s} X_{st}$

(3) $X_{st} X_{tu} = X_{su}$,

(4) if $s_1 \leqslant t_1 \leqslant s_2 \leqslant t_2 \leqslant \ldots \leqslant t_n$, then the random variables $X_{s_1 t_1}$, $X_{s_2 t_2}$, …, $X_{s_n t_n}$ are independent.

(5) $X_{s+h,t+h} = X_{st}$ *in the sense that*

$$\mathbb{P}_{X_{st}} = \mathbb{P}_{X_{s+h,t+h}} \tag{87.1}$$

where $\mathbb{P}_X(A) = \mathbb{P}\big(X^{-1}(A)\big)$.

If G is a group, we add the condition $X_{st} = X_s^{-1} X_t$.

The map $X \colon \Omega \to G$ gives rise to the map

$$\begin{aligned} j_X \colon L^\infty(G, B) &\to L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \\ j_X(f) &= f \circ X \end{aligned} \tag{87.2}$$

Now we can look at the properties of a Lévy process on j instead of X. Firstly, j_X is a $*$-algebra homomorphism such that $j_X(1) = 1$.

87.2 Quantum probability space and stochastic processes

Definition 87.2.
*A **quantum probability space** is a pair (A, Φ) is an unital $*$-algebra A and a state Φ on A. A **quantum random variable** j over a quantum probability space (A, Φ) on a $*$-algebra B is an homomorphism $j \colon B \to A$ of $*$-algebras.*

Definition 87.3.
*A **quantum stochastic process** is a family of quantum random variables $(j_j)_{t \in I}$. Its **marginal distribution** is the set of maps*

$$\begin{aligned} \varphi_j \colon B &\to \mathbb{C} \\ \varphi_j &= \Phi \circ j_t. \end{aligned} \tag{87.3}$$

Definition 87.4.
*An **operator process** if a family of elements $(X_t)_{t\in I}$ of a quantum probability space ($X_t \in A$).*

From an operator process and a choice of $b \in B$, we can define a stochastic process on $\mathbb{C}\langle b, b^*\rangle$ by defining

$$j_t(b) = X_t \tag{87.4}$$

and extend it to $\mathbb{C}\langle b, b^*\rangle$ as $*$-homomorphism.

On the other hand, if $\{j_t \colon B \to A\}_{t\in I}$ is a stochastic process, we define an operator process by choosing $x \in B$ and defining

$$X_t = j_t(x). \tag{87.5}$$

Definition 87.5.
*We say that maps $f_i \colon B \to A$ are **tensor, or boson independent** with respect to Φ if*
(1) $\Phi\big(f_1(b_1)\cdots f_n(b_n)\big) = \Phi\big(f_1(b_1)\big)\ldots\Phi\big(f_n(b_n)\big)$ for every $b_i \in B$,
(2) the images commute: $[f_k(b_k), f_l(b_l)] = 0$ for every $k \neq l$.

If $j_1, j_2 \colon B \to A$ are maps from B to an algebra A, we define the **convolution** by

$$j_1 * j_2 = m_A \circ (j_1 \otimes j_2) \circ \Delta. \tag{87.6}$$

Notice that, if j_1 and j_2 are homomorphisms, it is not guaranteed in general that the convolution $j_1 * j_2$ is an homomorphism. In the case of independent variables, however, it is true: if j_1 and j_2 are independent random variables, then $j_2 * j_2$ is still a random variable because

$$\begin{aligned}(j_1 * j_2)(ab) &= m_A \circ (j_1 \otimes j_2)\big(a_{(1)}b_{(1)} \otimes a_{(2)}b_{(2)}\big)\\ &= j_1(a_{(1)})j_1(b_{(1)})j_2(a_{(2)})j_2(b_{(2)})\\ &= j_1(a_{(1)})j_2(a_{(2)})j_1(b_{(1)})j_2(b_{(2)})\\ &= (j_1 * j_2)(a)(j_1 * j_2)(b).\end{aligned} \tag{87.7}$$

where $a_{(1)}b_{(1)}$ stands for $(ab)_{(1)}$.

87.3 Lévy process

Definition 87.6.
*Let (B, Δ, ϵ) be a bialgebra. A **Lévy process** on (B, Δ) over the quantum probability space (A, Φ) is a quantum stochastic process $(j_{st})_{0\leqslant s\leqslant t}$ which satisfies the following requirements*

Increment property *We have*

$$j_{rs} * j_{st} = j_{rt} \qquad \text{for all } 0 \leqslant r \leqslant s \leqslant t \tag{87.8a}$$

$$j_{tt}(b) = \epsilon(b)1_A \qquad \text{for all } 0 \leqslant t, b \in B \tag{87.8b}$$

Independence of increments *for every $n \in \mathbb{N}$ and $0 \leqslant s_1 \leqslant t_1 \leqslant s_2 \leqslant \ldots \leqslant s_n \leqslant t_n$, the maps $j_{s_1t_1}, \ldots j_{s_nt_n}$ are independent with respect to Φ.*

Stationarity of the increments *the distribution $\varphi_{st} = j_{st} \circ j_{st}$ depends only on the difference $t - s$, in other words we have $\varphi_{s+h,t+h} = \varphi_{st}$ for every $0 \leqslant s \leqslant t$ and $h \geqslant 0$.*

Weak continuity $\lim_{t\searrow s} j_{st}(a) = j_{ss}(a)$ *for every $a \in A$.*

One say that the process $j_{st} \colon B \to (A, \Phi)$ and $k_{st} \colon B \to (A', \Phi')$ are **equivalent** if

$$\Phi\big(j_{s_1t_1}(b_1)\ldots j_{s_nt_n}(b_n)\big) = \Phi'\big(k_{s_1t_1}(b_1)\ldots k_{s_nt_n}(b_n)\big), \tag{87.9}$$

that is if all the expectation values that we could compute are equal.

If (j_{st}) is a Lévy process, we define

$$\varphi_{t-s} = \Phi \circ j_{st} \tag{87.10}$$

for $0 \leqslant s \leqslant t$. This is well defined from the stationarity of increments. We can as well write $\varphi_t = \Phi \circ j_{0t}$.

87.4 Schürmann triple

Proposition 87.7.
Let B be an involutive coalgebra and $(\varphi_t)_{t\geqslant 0}$, a convolution semigroup of linear functionals on B. If we define

$$L = \lim_{t\searrow 0} \frac{1}{t}(\varphi_t - \epsilon), \tag{87.11}$$

then the following are equivalents:

(1) the maps φ_t are states on B;

(2) the map $L\colon B \to \mathbb{C}$ satisfies $L(1_B) = 0$ and is hermitian and conditionally positive.

Lemma 87.8.
If we define $\varphi_t = \varphi_{0t} = \Phi \circ j_{0t}$, we get a convolution semigroup.

Proof. The first condition comes from

$$\varphi_0(b) = \Phi\big(j_{00}(b)\big) = \Phi\big(\epsilon(b)1_A\big) = \epsilon(b)\Phi(1_A) = \epsilon(b) \tag{87.12}$$

since $\Phi(1_A) = 1$ because Φ is a state on A.

Using the fact that $j_{0,s+t} = j_{0s} * j_{s,s+t}$, we have

$$\varphi_{s+t}(a) = \varphi_{0,s+t}(a) = \Phi\big((j_{0s} * j_{s,s+t})(a)\big). \tag{87.13}$$

But,

$$(j_{0s} * j_{s,s+t})(a) = m \circ (j_{0s} \otimes j_{s,s+t})(a_{(1)} \otimes a_{(2)}) = j_{0s}(a_{(1)})j_{s,s+t}(a_{(2)}). \tag{87.14}$$

Thus, since the intervals $0, s$ and $s, s+t$ are independent, the equation (87.13) becomes

$$\begin{aligned}
\varphi_{s+t}(a) &= \Phi\big(j_{0s}(a_{(1)})j_{s,s+t}(a_{(2)})\big) \\
&= \Phi\big(j_{0s}(a_{(1)})\big)\Phi\big(j_{s,s+t}(a_2)\big) \\
&= \varphi_s(a_{(1)})\varphi_t(a_{(2)}) \\
&= (\varphi_s \otimes \varphi_t)\Delta a \\
&= (\varphi_s * \varphi_t)(a).
\end{aligned} \tag{87.15}$$

It remains to be justified that $\Phi\big(j_{s,s+t}(a)\big) = \varphi_t(a)$. This comes from the fact that $\varphi_{s+h,t+h} = \varphi_{st}$, so that $\varphi_{s,s+t} = \varphi_{0+s,s+t} = \varphi_{0,t}$. □

A linear functional ω on a bialgebra B is **conditionally positive** if $\omega(a^*a) \geqslant 0$ for every a such that $\epsilon(a) = 0$. The functional is **Hermitian** if $\omega(a^*) = \overline{\omega(a)}$.

In our case, we know that the functionals φ_t are in fact states. We define

$$L = \lim_{t\searrow 0} \frac{1}{t}(\varphi_t - \epsilon). \tag{87.16}$$

This is conditionally positive because

$$L(a^*a) = \lim_{t\searrow 0} \frac{1}{t}\big(\varphi_t(a^*a) - \epsilon(a^*a)\big) = \lim_{t\searrow 0} \frac{\varphi_t(a^*a)}{t} \geqslant 0 \tag{87.17}$$

because φ_t is a state and thus positive.

Definition 87.9.
*Let B be an unital $*$-algebra with an unital hermitian character $\epsilon\colon B \to \mathbb{C}$. A **Schürmann triple** on (B, ϵ) is a triple (ρ, η, L) with*

(1) $\rho\colon B \to \mathcal{L}(D)$ is a $*$-representation of B on a pre-Hilbert space D.

(2) The map $\eta\colon B \to D$ is linear and

$$\eta(ab) = \rho(a)\eta(b) + \eta(a)\epsilon(b), \tag{87.18}$$

*such a map is called a ρ-ϵ-1-**cocycle**.*

(3) The map $L\colon B \to \mathbb{C}$ is an hermitian linear functional whose $\epsilon - \epsilon - 2-$coboundary is the map

$$(a, b) \mapsto -\langle \eta(a^*), \eta(b)\rangle, \tag{87.19}$$

that means that

$$-\langle \eta(a^*), \eta(b)\rangle = (\partial L)(a, b) = \epsilon(a)L(b) - L(ab) + L(a)\epsilon(b) \tag{87.20}$$

EqCondFiffAssetaL

for every $a, b \in B$.

ItemPropCorSchr

Corollary 87.10.
A Schürmann triple has the following immediate properties.

(1) $\eta(1) = 0$.

(2) $L(1) = 0$.

ItemPropCorSchriii

(3) The condition (87.20) is equivalent to ask

$$L(ab) = -\langle \eta(a^*), \eta(b)\rangle \tag{87.21}$$

EqConsSimplAssetaL

for every $a, b \in K_1$.

Proof. (1) Using the cocycle property and the fact that $\rho(1)$ is the identity,

$$\eta(1 \cdot 1) = \rho(1)\eta(1) + \eta(1)\epsilon(1) = 2\eta(1). \tag{87.22}$$

(2) Now, writing the compatibility relation (87.20) with $a = b = 1$ and taking into account $\eta(1) = 0$ we have

$$0 = \epsilon(1)L(1) - L(1 \cdot 1) + L(1)\epsilon(1) = L(1). \tag{87.23}$$

(3) If a and b belong to K_1, the relation (87.20) reduces to (87.21). Now suppose that the condition (87.21) holds for every a and b in K_1. Taking any $a, b \in B$ we consider $a - \epsilon(a)1$ and $b - \epsilon(b)1$ that are elements of K_1. Taking into account the fact that $L(1) = 0$, we have

$$L(ab) - \epsilon(b)L(a) - \epsilon(a)L(b) = \langle \eta(a^*), \eta(b)\rangle. \tag{87.24}$$

□

Let (π, η, L) be a Schürmann triple on the compact quantum group $\mathcal{A}_q$ (see section 73.2). The map $\eta\colon \mathcal{A}_q \to \mathscr{L}(D)$ satisfies the cocycle condition

$$\eta(ab) = \pi(a)\eta(b) + \eta(a)\epsilon(b). \tag{87.25}$$

PropCocycleDeteretavjnmu

Proposition 87.11.
The cocycle η is determined by its values on the elements v_j^* with $j = 1, \dots, n-1$.

Proof. Since $\epsilon(1) = 1$, we have

$$\eta(1) = \pi(1)\eta(1) + \eta(1)\epsilon(1) = \big(\pi(1) + 1\big)\eta(1). \tag{87.26}$$

Since the representation is nondegenerate we have $\pi(1) = \mathrm{Id}$. The only possibility is thus $\eta(1) = 0$.

If a and b belong to K_1 we have $\eta(ab) = \pi(a)\eta(b)$. Since each element in $\mathcal{A}_q$ is a polynomial in u_{ij} and u_{ij}^*, we only have to fix the value of η on these elements.

First, $\eta(u_{jj}^*) = \eta(v_j^*)$ because $\eta(1) = 0$. Using lemma 73.6, taking the adjoint,

$$q\eta(u_{ll}^* u_{ij}^*) = \eta(u_{ij}^* u_{ll}^*). \tag{87.27}$$

Using the cocycle property and the fact that $\epsilon(u_{ll}^*) = 1$,

$$\big(q\pi(u_{ll}^*) - \mathrm{Id}\,\big)\eta(u_{ij}^*) = \pi(u_{ij}^*)\eta(u_{ll}^*). \tag{87.28}$$

The operator $q\pi(u_{li}^*) - \mathrm{Id}$ is invertible because of lemma 73.14 and $q < 1$. Thus, when $i \neq j$ and $i, j \leqslant n-1$,

$$\eta(u_{ij}^*) = \big(q\pi(u_{li})^* - \mathrm{Id}\,\big)^{-1}\pi(u_{ij}^*)\eta(u_{ll}^*). \tag{87.29}$$

Let us now compute $\eta(v_j)$ for $j = 1, \dots, n-1$. Taking the relation (73.52),

$$\begin{aligned} 0 = \eta(1) &= \sum_{p=1}^{n} \eta(u_{jp}u_{jp}^*) \\ &= \eta(u_{jj}u_{jj}^*) + \sum_{p\neq j} \eta(u_{jp}u_{jp}^*) \\ &= \pi(u_{jj})\eta(u_{jj}^*) + \eta(u_{jj})\underbrace{\epsilon(u_{jj}^*)}_{=1} + \sum_{p\neq j} \eta(u_{jp}u_{jp}^*), \end{aligned} \tag{87.30}$$

so that

$$\eta(v_j) = \eta(u_{jj}) = -\pi(u_{jj})\eta(u_{jj}^*) - \sum_{p} \pi(u_{jj})\eta(u_{jp}^*). \tag{87.31}$$

Eqetauipiu

Thus the vectors $\eta(u_{jj})$ and $\eta(v_j)$ are fixed for $j \leqslant n-1$.

We are going to compute $\eta(u_{ij})$ with $i \neq j$ and $i, j \leqslant n-1$ using the lemma 73.7. We pose $k = \max(i, j)$. Since $\epsilon(u_{kk}) = 1$ and $\epsilon(u_{ij}) = 0$, the relation $\eta(u_{ij}u_{kk}) = q\eta(u_{kk}u_{ij})$ leads to

$$\pi(u_{ij})\eta(u_{kk}) = q\pi(u_{kk})\eta(u_{ij}). \tag{87.32}$$

What we obtain is

$$\eta(u_{ij}) = -\big(\mathrm{Id} - q\pi(u_{kk})\big)^{-1}\pi(u_{ij})\eta(u_{kk}). \tag{87.33}$$

In order to compute the value of $\eta(v_n^*)$, we start from the definition (73.49) of the involution. Taking into account the fact that in our case $D = 1$,

$$u_{nn}^* = D^{nn} = \sum_{\sigma\colon \{1\dots n-1\}\to\{1,\dots,n-1\}} (-1)^{|\sigma|} u_{1\sigma(1)} \dots u_{n-1,\sigma(n-1)}. \tag{87.34}$$

Taking η and using the cocycle property we have

$$\begin{aligned} \eta(v_n^*) &= \eta(u_{nn}^*) \\ &= \sum_{\sigma(n-1)\neq n-1} (-1)^{|\sigma|}\pi\big(u_{1\sigma(1)}\dots u_{n-2,\sigma(n-1)}\big)\eta(u_{n-1,\sigma(n-1)}) \\ &\quad + \sum_{\sigma(n-1)= n-1} (-1)^{|\sigma|}\pi\big(u_{1\sigma(1)}\dots u_{n-2,\sigma(n-2)}\big)\eta(u_{n-1,n-1}) \\ &\quad + \sum_{\sigma(n-1)= n-1} (-1)^{|\sigma|}\eta\big(u_{1\sigma(1)}\dots u_{n-2,\sigma(n-2)}\big)\underbrace{\epsilon(u_{n-1,n-1})}_{=1} \end{aligned} \tag{87.35}$$

Applying many times the cocycle the term $\eta\big(u_{1\sigma(1)}\dots u_{n-2,\sigma(n-2)}\big)$ decomposes into basic elements containing $\eta(u_{i\sigma(i)})$. This fixes $\eta(v_n^*)$.

Taking the equation (87.31) with $j = n$ we have

$$\eta(v_n) = \eta(u_{nn}) = -\pi(u_{nn})\eta(u_{nn}^*) - \sum_{p=1}^{n} \pi(u_{nn})\eta(u_{np}^*). \tag{87.36}$$

That fixes $\eta(v_n)$. Now we consider the previously proved equation

$$\eta(u_{ij}) = \big(q\pi(u_{kk}) - \mathrm{Id}\big)^{-1}\pi(u_{ij})\eta(v_k) \tag{87.37}$$

with $i \neq j$ and $n = \max(i, j)$. Then

$$\eta(u_{ij}) = \big(q\pi(u_{nn}) - \mathrm{Id}\big)^{-1}\pi(u_{ij})\eta(v_n). \tag{87.38}$$

This fixes $\eta(u_{in})$ and $\eta(u_{nj})$.

□

87.5 A representation

PropReprezThetasuqn

Proposition 87.12.
The following is a one dimensional representation of $\mathrm{SU}_q(n)$ *if* $\sum_k \theta_k = 1$*:*

$$\pi(u_{jk}) = e^{i\theta_j}\delta_{jk} \tag{87.39}$$

Proof. First we have

$$\pi(D) = \sum_{\sigma\in S(n)} (-q)^{|\sigma|} e^{i\theta_1}\delta_{1\sigma(1)} \dots e^{i\theta_n}\delta_{n\sigma(n)} = e^{i(\theta_1+\dots+\theta_n)} = 1. \tag{87.40}$$

Equations (73.33) all reduce to $0 = 0$ because of the Kronecker delta's and the condition on the indices. □

Since the condition $D = 1$ imposes the sum of the θ's to be zero, we will write the representation π by

$$\epsilon_{\theta_1\dots\theta_{n-1}}(u_{ij}) = e^{i\theta_j}\delta_{jk}. \tag{87.41}$$

Notice that the counit ϵ is $\epsilon_{0,\dots,0}$. We define

$$\epsilon'_k = \frac{\partial}{\partial\theta_k}\epsilon_{\theta_1\dots\theta_{n-1}}\Big|_{\theta_k=0}. \tag{87.42}$$

As an example, in $\mathrm{SU}_q(3)$ we have

$$\epsilon_{\theta_1\theta_2} = \begin{pmatrix} e^{i\theta_1} & 0 & 0 \\ 0 & e^{i\theta_2} & 0 \\ 0 & 0 & e^{-i(\theta_1+\theta_2)} \end{pmatrix} \tag{87.43}$$

in the sense that $\epsilon_{\theta_1\theta_2}(u_{ik}) = \big(\epsilon_{\theta_1\theta_2}\big)_{ik}$. Thus we have

$$\begin{aligned} \epsilon'_{\theta_1}(u_{ik}) &= \frac{d}{d\theta_1}\Big[\epsilon_{\theta_1\theta_2}(u_{ik})\Big]_{\theta_1=0} \\ &= \frac{d}{d\theta_1}\begin{pmatrix} e^{i\theta_1} & 0 & 0 \\ 0 & e^{i\theta_2} & 0 \\ 0 & 0 & e^{-i(\theta_1+\theta_2)} \end{pmatrix}_{ik}\Big|_{\theta_1=0} \\ &= \begin{pmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{pmatrix}. \end{aligned} \tag{87.44}$$

Higher order derivative are defined the same way:

$$\epsilon''_{kl} = \frac{\partial}{\partial\theta_k}\frac{\partial}{\partial\theta_l}\epsilon_{\theta_1\dots\theta_{n-1}}\Big|_{\substack{\theta_k=0\\ \theta_l=0}}. \tag{87.45}$$

87.5.1 Ideals

Let $\mathcal{A}_q$ be the Hopf C^*-subalgebra of $\mathrm{SU}_q(n)$ generated by $\{1, u_{ij}\}$, and a Schürmann triple (π, η, L) where π is a representation of $\mathcal{A}_q$ on $\mathfrak{B}(\mathscr{H})$. We define the ideal

$$K_1 = \ker(\epsilon) \subset \mathcal{A}_q. \tag{87.46}$$

The ideal K_1 is generated by the elements of the form u_{ij} $(i \neq j)$ and $u_{ii} - 1$. Then we define

$$K_2 = \mathrm{Span}\{ab \text{ tel que } a, b \in K_1\} \tag{87.47}$$

and more generally

$$K_n = \mathrm{Span}\{a_1 \dots a_n \text{ tel que } a_i \in K_1\}. \tag{87.48}$$

We also define

$$K_\infty = \bigcap_{m \geqslant 1} K_m. \tag{87.49}$$

Each of these K_m is a two-sided ideal since

$$ba_1 \dots a_m = (ba_1)a_2 \dots a_m \tag{87.50}$$

belongs to K_m if $a_1 \dots a_m$ belongs to K_m.

We define

$$\begin{aligned} v_j &= u_{jj} - 1 \\ d_j &= \frac{v_j - v_j^*}{2i} \end{aligned} \tag{87.51}$$

Since $\epsilon(1) = 1$, the unit does not belong to K_1, but the elements u_{ij}, u_{ij}^*, v_j and d_j belong to K_1 when $i \neq j$.

Lemuijvkkuij

Lemma 87.13.
If $k = \max(i, j)$ we have

$$u_{ij}v_k - qv_ku_{ij} = (q-1)u_{ij}. \tag{87.52}$$

Proof. This is a computation:

$$\begin{aligned} u_{ij}v_k &= u_{ij}u_{kk} - u_{ij} \\ &= qu_{kk}u_{ij} - u_{ij} && \text{because } k = \max(i, j) \\ &= (qu_{kk} - 1)u_{ij} \\ &= (qu_{kk} - q)u_{ij} + (q-1)u_{ij} \\ &= qv_ku_{ij} + (q-1)u_{ij}. \end{aligned} \tag{87.53}$$

The proof is completed. □

Lesvjvjvuuuujjiiv

Lemma 87.14.
We have

$$v_jv_j^* = u_{jj}u_{jj}^* - v_j - v_j^* - 1. \tag{87.54}$$

Proof. This is a computation:

$$\begin{aligned} v_jv_j^* &= (u_{jj} - 1)(u_{jj}^* - 1) \\ &= u_{jj}u_{jj}^* - u_{jj} - u_{jj}^* + 1 \\ &= u_{jj}u_{jj}j * -v_j - u_{jj}^* + 1 - 1 \\ &= u_{jj}u_{jj}^* - v_j - v_j^* - 1. \end{aligned} \tag{87.55}$$

□

PropuudansKKiii

Proposition 87.15.
We have

ItemuudansKKii
ItemuudansKKiii
ItemuudansKKiv
ItemPropuudansKKiiiavii

(1) $u_{ij} \in K_\infty$ *if* $i \neq j$;
(2) $1 - u_{jj}u_{jj}^* \in K_\infty$;
(3) $u_{ii}u_{jj} - u_{jj}u_{ii} \in K_\infty$;
(4) $u_{ii}u_{jj}^* - u_{jj}^*u_{ii} \in K_\infty$;
(5) $v_j + v_j^* \in K_2$;
(6) $\sum_k v_k \in K_2$;
(7) $\sum_{k=1}^n d_k \in K_2$.

Proof. (1) Since $\epsilon(u_{ij}) = \delta_{ij}$, we have $u_{ij} \in K_1$ when $i \neq j$
Now, using lemma 87.13 we have

$$u_{ij} = (q-1)^{-1}(u_{ij}v_k - qv_ku_{ij}). \tag{87.56}$$

(2) Since $\epsilon(u_{jj}) = 1$ we have $1 - u_{jj}u_{jj}^* \in K_1$. Using formula (73.52), we have

$$\begin{aligned} 1 - u_{jj}u_{jj}^* &= \delta_{jj} - u_{jj}u_{jj}^* \\ &= \sum_k u_{jk}u_{jk}^* - u_{jj}u_{jj}^* \\ &= \sum_{k\neq j} u_{jk}u_{jk}^*. \end{aligned} \tag{87.57}$$

Since each of the terms in that sum belong to K_∞, we have the result.

(3) The combination $u_{ii}u_{jj} - u_{jj}u_{ii}$ belongs to $\ker(\epsilon) = K_1$. Let us suppose $i < j$ (if not, consider change the sign). Using the relation (73.33d), we have

$$u_{ii}u_{jj} - u_{jj}u_{ii} = (q - q^{-1})u_{uj}u_{ji} \in K_\infty. \tag{87.58}$$

(4) Let's begin with $i = j$. We have

$$u_{ii}u_{ii}^* - u_{ii}^*u_{ii} = u_{ii}u_{ii}^* - 1 + 1 - u_{ii}^*u_{ii} \in K_\infty \tag{87.59}$$

where we used the point (2).
If $i \neq j$, we use the relation (73.53a) which says that $u_{ii}u_{jj}^* = u_{jj}^*u_{ii}$, so that the combination we are looking at is zero.

(5) Lemma 87.14 shows that

Eqvvuiiuiivv
$$v_j + v_j^* = u_{jj}u_{jj}^* - 1 - v_jv_j^* \tag{87.60}$$

The fact that $v_j \in K_1$ and item (2) show that the right hand side belong to K_2.

(6) Let us decompose the sum defining the determinant:

$$1 = \sum_{\sigma\in S_n} (-q)^{|\sigma|}u_{1\sigma(1)}\dots u_{n\sigma(n)} = u_{11}\dots u_{nn} + \sum_{\sigma\neq \mathrm{Id}} (-q)^{|\sigma|}u_{1\sigma(1)}\dots u_{n\sigma(n)}. \tag{87.61}$$

The last sum belongs to K_∞ since there is at least one k with $k \neq \sigma(k)$. Thus, replacing u_{jj} by $v_j + 1$, we have

Eqinteunvuvuuvnldots
$$1 = (v_1 + 1)\dots(v_n + 1) + k_\infty. \tag{87.62}$$

The product can be written under the form

$$(v_1 + 1)\dots(v_n + 1) = \sum_{p=1}^n \sum_{i_1<\dots<i_p} v_{i_1}\dots v_{i_p} + 1. \tag{87.63}$$

In the latter sum, a part of the term $p = 1$, each term belongs to K_2, so equation (87.62) reads

$$1 = \sum_{i_1} v_{i_1} + k_2 + k_\infty, \tag{87.64}$$

and we conclude that $\sum_i v_i = k_2 + k_\infty \in K_2$.

(7) We know that $v_1^* + \cdots + v_n^* \in K_2$, so that

$$d_1 + \cdots + d_n = \frac{1}{2i}\big((v_1 + \ldots v_n) - (v_1^* + \cdots + v_n^*)\big) \in K_2. \tag{87.65}$$

□

Proposition 87.16.
We have

$$\mathcal{A}_q/K_\infty \simeq C(S^{n-1}) \tag{87.66}$$

Proof. Since u_{ij} belongs to K_∞ when $i \neq j$, the algebra $\mathcal{A}_q/K_\infty$ is generated by the elements u_{jj} and u_{jj}^* as well as 1. That algebra is commutative from points (3) and (4) of proposition 87.15.

We consider the following map [1]:

$$\begin{aligned} \varphi\colon S^{n-1} &\to \Delta(\mathcal{A}_q/K_\infty) \\ t_1, \ldots t_{n-1} &\mapsto \omega_{t_1,\ldots t_{n-1}} \end{aligned} \tag{87.67}$$

where $\omega_t\colon \mathcal{A}_q/K_\infty \to \mathbb{C}$ $(t \in S^{n-1})$ is defined by

$$\omega_t(u_{kk}) = \begin{cases} e^{it_k} & \text{if } k \neq n \\ e^{-i(t_1+\cdots+t_{n-1})} & \text{if } k = n \end{cases} \tag{87.68}$$

and $\omega_t(1) = 1$. Notice that the value of ω_t on u_{nn} is imposed by the fact that $\omega_t(1 - u_{11} \ldots u_{nn}) = 0$ and the fact that ω_t has to be multiplicative.

The map φ is injective because the value of $\varphi(t_1, \ldots, t_{n-1})$ on the element u_{kk} fixes the values of t_k.

It is also surjective because $1 - u_{jj}u_{jj}^*$ belongs to K_∞; thus each ω in $\Delta(\mathcal{A}_q/K_\infty)$ satisfies $\omega(u_{kk}u_{kk}^*) = 1$ and

$$\omega(u_{kk})\omega(u_{kk})^* = 1, \tag{87.69}$$

so that there exists a unique $t_k \in [0, 2\pi[$ such that $\omega(u_{kk}) = e^{it_k}$.

Now we conclude the proof using the Gelfand theorem 72.15 which states that for every C^*-algebra $\mathfrak{A}$, we have $\mathfrak{A} \simeq C\big(\Delta(\mathfrak{A})\big)$. Indeed, we just proved that

$$S^{n-1} \simeq \Delta(\mathcal{A}_q/K_\infty). \tag{87.70}$$

Thus we have

$$C(S^{n-1}) \simeq C\big(\Delta(\mathcal{A}_q/K_\infty)\big) \simeq \mathcal{A}_q/K_\infty. \tag{87.71}$$

□

87.5.2 Decomposition

Proposition 87.17.
Every element $x \in \mathfrak{A}_q$ decomposes into

$$x = c_0 1 + \sum_{j=1}^{n-1} c_j d_j + \sum_{1 \leqslant i \leqslant j \leqslant n} x_{ij} v_i v_j^* + k_3 \tag{87.72}$$

for some $c_0, c_j, c_{ij} \in \mathbb{C}$ and $k_3 \in K_3$.

1. Here $\Delta(\mathcal{A}_q/K_\infty)$ stands for the structure space of $\mathcal{A}_q/K_\infty$, see definition 72.1.

Proof. First, x is a complex polynomial in the variables u_{ij} and u_{ij}^*. If we replace u_{jj} by $v_j + 1$, we have a polynomial

$$x = p(v_j, v_j^*, u_{ij}, u_{ij}^*) \tag{87.73}$$

with $i \neq j$. We know that $u_{ij} \in K_\infty$ (proposition 87.15), so that, modulo elements of K_3, we can neglect the terms with u_{ij} $(i \neq j)$. It remains a polynomial in the variables v_j and v_j^*. The terms of order m belong to K_m, so that we only have to consider the first two terms. What we have is then a polynomial of degree 2 in v_j and v_j^*:

$$x = c_0 1 + \sum_{j=1}^{n} (\alpha_j v_j + \beta_j v_j^*) + \sum_{i,j=1}^{n} \sum_{\epsilon_1,\epsilon_2 \in \{1,*\}} \gamma_{ij}^{\epsilon_1,\epsilon_2} v_i^{\epsilon_1} v_j^{\epsilon_2} + k_3. \tag{87.74}$$ EqxSumvvdgammaUn

Here the sum over ϵ_i means that we have the terms $v_i v_j$, $v_i v_j^*$, $v_i^* v_j$ and $v_i^* v_j^*$.

The first degree terms are of the form

$$\alpha_j v_j + \beta_j v_j^* = \tilde{c}(\alpha, \beta) \frac{v_j + v_j^*}{2} + c(\alpha, \beta) \frac{v_j - v_j^*}{2i}. \tag{87.75}$$

Since

$$d_n = d_1 + \cdots + d_{n-1} + k_2, \tag{87.76}$$

up to changing the coefficients of the d_k's $(j \leqslant n-1)$ and in the higher order terms, we can reduce the sum to $\sum_{j=1}^{n-1} d_j$.

So let's say that

$$\sum_{j=1}^{n} c_j d_j \in K_1/K_2. \tag{87.77}$$

From element (7) of proposition 87.15, we know that there exists an element $k_2 \in K_2$ such that

$$d_n = k_2 - d_1 - \ldots - d_{n-1}, \tag{87.78}$$

so that in K_1/K_2 we have $d_n = -\sum_{j=1}^{n-1} d_j$ and a basis of K_1/K_2 is

$$B_1 = \{d_1, \ldots, d_{j-1}\}. \tag{87.79}$$

¡¡ Avertissement/question à la lectrice !! 87.18
Ok, this is a generating part for K_1/K_2. Why is it free? In what sense?

The decomposition (87.74) reads now

$$x = c_0 1 + \sum_{j=1}^{n-1} c_j d_j + \sum_{j=1}^{n} \tilde{c}_j (v_j + v_j^*) + \sum_{i,j=1}^{n} \sum_{\epsilon_1,\epsilon_2} \gamma_{ij}^{\epsilon_1,\epsilon_2} v_i^{\epsilon_1} v_j^{\epsilon_2} + k_3. \tag{87.80}$$ EqxSumvvdgammaDeux

Using relation (87.60),

$$v_j + v_j^* = v_j v_j^* + \underbrace{1 - u_{jj} u_{jj}^*}_{\in K_\infty}, \tag{87.81}$$

We can replace $\sum_{j=1}^{n} \tilde{c}_j (c_j + v_j^*)$ by $\sum_{j=1}^{n} \tilde{c}_j v_j v_j^*$.

Let us prove that modulo elements in K_3, the elements

$$v_i v_j, \quad v_i^* v_j, \quad v_i^* v_j^* \tag{87.82}$$

can be written as combinations of $v_i v_j^*$. Since $v_i + v_i^* \in K_2$, we have

$$(v_i + v_i^*) v_j \in K_3, \tag{87.83}$$

so that $v_i v_j = v_i^* v_j + k_3$. Making the same with $(v_i + v_i^*) v_j^*$ and $v_i (v_j + v_j^*)$, we have

$$\sum_{\epsilon_1,\epsilon_2} \gamma_{ij}^{\epsilon_1,\epsilon_2} v_i^{\epsilon_1} v_j^{\epsilon_2} = \sum_{i,j=1}^{n} c_{ij} v_i v_j^* + k_3. \tag{87.84}$$

We can continue the simplification because $v_i v_j^* = v_j^* v_i$ in K_2/K_3. Indeed

$$v_i v_j^* - v_j^* v_i = (u_{ii} - 1)(u_{jj}^* - 1) - (u_{jj}^* - 1)(u_{ii} - 1) = u_{ii}u_{jj}^* - u_{jj}^* u_{ii} \in K_\infty. \tag{87.85}$$

Thus we have for example:

$$v_2 v_1^* = v_1^* v_2 = v_1 v_2^*, \tag{87.86}$$

and (up to redefinition of c_{ij}),

$$\sum_{i,j=1}^{n} c_{ij} v_i v_j^* = \sum_{1 \leqslant i \leqslant j \leqslant n} c_{ij} v_i v_j^* + k_3. \tag{87.87}$$

Since $v_1^* + \cdots + v_n^* \in K_2$, we have

$$v_i v_1^* + \cdots + v_i v_{n-1}^* + v_i v_n^* \in K_3, \tag{87.88}$$

so that the term $v_i v_n^*$ is a combination of the terms $v_i v_j$ with $j < n$. Finally the sum (87.80) reduces to

$$x = c_0 1 + \sum_{j=1}^{n-1} c_j d_j + \sum_{1 \leqslant i \leqslant j \leqslant n-1} c_{ij} v_i v_j^* + k_3. \tag{87.89}$$

□

The left hand side belongs to K_1 while the first term in the right hand side belongs to K_2 (see proposition 87.15).

¡¡ Avertissement/question à la lectrice !! 87.19
Why does it prove that $d_j \in K_1/K_2$?

(1) K_m is not a vector space. So when we write K_i/K_{i+1}, do we mean the set difference *or the class with respect to*

$$x \sim x + k_{i+1} \tag{87.90}$$

or

$$x \sim x k_{i+1} \quad ? \tag{87.91}$$

(2) Why K_1/K_2 could not be empty?

(3) Why $\alpha v_j + \beta v_j^*$ does not *belong to K_2?*

The decomposition leads us to decompose a basis of $\mathcal{A}_q$ into

$$B = \{1\} \cup B_1 \cup B_2 \cup B_3 \tag{87.92}$$

where

$$\begin{aligned} B_1 &= \{d_i \text{ tel que } 1 \leqslant i \leqslant n-1\} \\ B_2 &= \{v_i v_j^* \text{ tel que } 1 \leqslant i \leqslant j \leqslant n-1\} \end{aligned} \tag{87.93}$$

and B_3 is a basis of whatever remains. We have $B_1 \subset K_1$, $B_2 \subset K_2$ and B_3 can be chosen as a subset of K_3.

We have the homomorphisms $\epsilon_\theta \colon \mathrm{SU}_q(n) \to \mathbb{C}$ and we define the derivatives

$$\begin{aligned} \epsilon'_k &= \frac{\partial}{\partial \theta_k} \epsilon_\theta|_{\theta_k = 0} \\ \epsilon''_{kl} &= \frac{\partial}{\partial \theta_l} \frac{\partial}{\partial \theta_k} \epsilon_\theta|_{\theta_k = \theta_l = 0} \end{aligned} \tag{87.94}$$

Proposition 87.20.
The functionals ϵ, ϵ'_k and ϵ''_{kl} with $1 \leqslant k \leqslant j \leqslant n$ separate the points $1, d_1, \ldots, d_{n-1}$, $v_i v_j^*$ with $1 \leqslant i \leqslant j \leqslant n-1$.

Proof. We have to prove that for each x in the enumerated elements, there exist a functional ω in the list such that $\omega(x) \neq 0$ and $\omega(y) \neq 0$ for $y \neq x$ where $x, y \in \{1, d_i, v_i v_j^*\}_{1 \leqslant i \leqslant j \leqslant n-1}$.

Let us begin with $x = 1$. For him, the functional is $\omega = \epsilon$. We have $\epsilon(1) = 1$ and $\epsilon(d_i) = \epsilon(v_i v_j^*) = 0$. Let us pass to $x = d_j$. We prove that $\omega = \epsilon_k'$ with $k \neq j$ works. We have

$$\epsilon_\theta(d_j) = \frac{1}{2i}\big(\epsilon_\theta(u_{jj}) - \epsilon_\theta(u_{jj}^*)\big) = \frac{1}{2i}\big(e^{i\theta_j} - e^{-i\theta_j}\big). \tag{87.95}$$

If $k \neq j$, we have $\epsilon_k'(d_j) = 0$ while, if $j = k$, we have

$$\epsilon_j'(d_j) = \frac{1}{2i}\frac{\partial}{\partial\theta_j}\big(e^{i\theta_j} - e^{-i\theta_j}\big)_{\theta_j=0} = 1. \tag{87.96}$$

In summary,

$$\epsilon'(d_j) = \delta_{kj}. \tag{87.97}$$

In order to see that ϵ_k' separates d_k, we still have to prove that $\epsilon_k'(v_i v_j^*) = 0$ for every i and j.

¡¡ Avertissement/question au lecteur !! 87.21
It seems to me that the functional we are looking at are ϵ and its derivatives. We are not working with arbitrary $\theta_1, \dots, \theta_{n-1}$. For arbitrary θ, we have

$$\epsilon_k'(ab) = \frac{\partial}{\partial\theta_k}\big(\epsilon_\theta(a)\epsilon_\theta(b)\big) = \epsilon_k'(a)\epsilon_\theta(b) + \epsilon_\theta(a)\epsilon_k'(b). \tag{87.98}$$

□

We consider the functionals

$$a \mapsto \epsilon(\theta, a) = \epsilon_\theta(a) \tag{87.99}$$

and the derivatives

$$\begin{aligned} \epsilon_l'(\theta, a) &= \frac{\partial}{\partial\theta_l}\big(\epsilon(\theta, a)\big) \\ \epsilon_{kl}''(\theta, a) &= \frac{\partial}{\partial\theta_k}\frac{\partial}{\partial\theta_l}\big(\epsilon(\theta, a)\big) \end{aligned} \tag{87.100}$$

We have the Leibniz rules

$$\begin{aligned} \epsilon_l'(\theta, ab) &= \frac{\partial}{\partial\theta_l}\big(\epsilon(\theta, a)\epsilon(\theta, b)\big) \\ &= \epsilon_l'(\theta, a)\epsilon(\theta, b) + \epsilon(\theta, a)\epsilon_l'(\theta, b), \end{aligned} \tag{87.101}$$

and

$$\begin{aligned} \epsilon_{kl}''(\theta, ab) = {} & \epsilon_{kl}''(\theta, a)\epsilon(\theta, b) + \epsilon_l'(\theta, a)\epsilon_k'(\theta, b) \\ & + \epsilon_k'(\theta, a)\epsilon_l'(\theta, b) + \epsilon(\theta, a)\epsilon_{kl}''(\theta, b). \end{aligned} \tag{87.102}$$

EqLeibnizepsppkl

We are interested in computing them on the basics elements 1, d_j, $v_i v_j^*$ that are generating $\mathrm{SU}_q(n)$ modulo K_3. Computations show that

$$\begin{aligned} \epsilon(\theta, 1) &= 1 \\ \epsilon(\theta, d_j) &= \frac{1}{2}\big(e^{i\theta_j} - e^{-i\theta_j}\big) & \epsilon(0, d_j) &= 0 \\ \epsilon(\theta, v_j) &= e^{i\theta_j} - 1 & \epsilon(0, v_j) &= 0 \\ \epsilon(\theta, v_i v_j^*) &= (e^{i\theta_i} - 1)(e^{-i\theta_j} - 1) & \epsilon(0, v_i v_j^*) &= 0. \end{aligned} \tag{87.103}$$

Taking the derivative,

$$\begin{aligned} \epsilon_l'(\theta, 1) &= 0 \\ \epsilon_l'(\theta, d_j) &= \frac{1}{2} i\delta_{lj}(e^{i\theta_l} + e^{-i\theta_l}) & \epsilon_l'(0, d_j) &= i\delta_{lj} \\ \epsilon_l'(\theta, v_j) &= i\delta_{kl}e^{i\theta_j} & \epsilon_l'(0, v_j) &= i\delta_{lj} \\ \epsilon_l'(\theta, v_i v_j^*) &= \delta_{ki}e^{i\theta_i}(e^{-i\theta_j} - 1) + (e^{i\theta_i} - 1)\delta_{kj}e^{-i\theta_j} & \epsilon_l'(0, v_i v_j^*) &= 0. \end{aligned} \tag{87.104}$$

¡¡ Avertissement/question à la lectrice !! 87.22
Pourquoi dans $\epsilon_l'(\theta, d_j)$, y'a pas un i qui descend à cause de la dérivée ?

Taking once again the derivative,

$$\begin{aligned}
&\epsilon_{kl}''(\theta, 1) = 0 \\
&\epsilon_{kl}''(\theta, d_j) = \frac{1}{2}\delta_{lj}\delta_{kl}(e^{i\theta_l} - e^{-i\theta_j}) \quad \epsilon_{kl}''(0, d_j) = 0 \\
&\epsilon_{kl}''(\theta, v_j) = \delta_{lj}\delta_{kj}e^{i\theta_j} \qquad\qquad \epsilon_{kl}''(0, v_j) = \delta_{lj}\delta_{kj}
\end{aligned} \tag{87.105}$$

and

$$\epsilon_{kl}''(\theta, v_i v_j^*) = -\delta_{ki}\delta_{li}e^{i\theta_i}(e^{-i\theta_j} - 1) + \delta_{li}\delta_{kj}e^{i\theta_i}e^{-i\theta_j} \tag{87.106a}$$

$$- \delta_{kj}\delta_{lj}e^{-i\theta_j}(e^{i\theta_i} - 1) + \delta_{lj}\delta_{ki}e^{-i\theta_j}e^{i\theta_i} \tag{87.106b}$$

SubEgenspppsuryvs

From these results, we deduce the following proposition.

Proposition 87.23.
The functional family

$$\begin{aligned}
&a \mapsto \epsilon(a) \\
&a \mapsto \epsilon_l'(a) \\
&a \mapsto \epsilon_{kl}''(a)
\end{aligned} \tag{87.107}$$

with $1 \leqslant k \leqslant l \leqslant n-1$ separates the points 1, d_j, $v_j v_j^*$ with $1 \leqslant i \leqslant j \leqslant n-1$.

Proof. For each point $x \in \{1, d_j, v_i v_j^*\}$, we have to find a functional ω in the given family such that $\omega(y) \neq 0$ if and only if $y = x$.

For $x = 1$, the functional ϵ makes the work:

$$\epsilon(1) \neq 0, \quad \epsilon(d_j) = 0, \quad \epsilon(v_i v_j^*) = 0. \tag{87.108}$$

The functional ϵ_l' separates d_l, indeed

$$\epsilon_l'(0, 1) = 0, \quad \epsilon_l'(0, d_j) = \delta_{lj}, \quad \epsilon_l'(0, v_i v_j^*) = 0. \tag{87.109}$$

And finally the functional ϵ_{kl}'' separates $v_k v_l^*$ because

$$\epsilon_k'' l(0, 1) = 0, \quad \epsilon_k'' l(0, d_j) = 0, \quad \epsilon_k'' l(0, v_i v_j^*) = \delta_{li}\delta_{kj} + \delta_{ki}\delta_{kj}. \tag{87.110}$$

The last is nonzero if and only if $k = i$ and $l = j$. □

PropDecompxczepsApp

Proposition 87.24.
The decomposition

$$x = c_0 1 + \sum_{j=1}^{n-1} c_j d_j + \sum_{1 \leqslant i \leqslant j \leqslant n-1} c_{ij} v_i v_j^* + k_3 \tag{87.111}$$

is unique and $c_0 = \epsilon(x)$, $c_j = \epsilon_j'(x)$ and $c_{ij} = \epsilon_{ij}''(x)$.

Proof. Applying ϵ to x we get $\epsilon(x) = c_0$. We also have

$$\epsilon_l'(x) = \sum_{j=1}^{n-1} c_j \underbrace{\epsilon_l'(d_j)}_{\delta_{lj}} = c_l \tag{87.112}$$

and

$$\epsilon_{kl}''(x) = \sum_{i \leqslant i \leqslant j \leqslant n-1} c_{ij}\delta_{lj}\delta_{kj} = c_{kl}. \tag{87.113}$$

Then the element k_3 is unique as the difference

$$k_3 = x - c_0 1 - \sum c_j d_j - \sum c_{ij} v_i v_j^*. \tag{87.114}$$

□

We define the map

$$\begin{aligned} P\colon \mathcal{A}_q &\to K_2 \\ x &\mapsto x - \epsilon(x)1 - \sum_{j=1}^{n-1} \epsilon'_j(x) d_j. \end{aligned} \tag{87.115}$$

This map satisfies because

$$\begin{aligned} \epsilon\big(P(x)\big) &= \epsilon(x) - \epsilon(x) = 0 \\ \epsilon'_k\big(P(x)\big) &= \epsilon'_k(x) - \sum \epsilon'_j(x)\epsilon'_k(d_j) = \epsilon'_k(x) - \epsilon'_k(x) = 0, \end{aligned} \tag{87.116}$$

so that

$$P\big(P(x)\big) = P(x). \tag{87.117}$$

87.6 Gaussian process

We follow [1047].

Proposition 87.25.
If L is conditionally positive and hermitian, then the following conditions are equivalent:

1. $\eta = 0$
2. $L|_{K_2} = 0$
3. *L is a ϵ-derivation, that means*

$$L(ab) = \epsilon(a)L(b) + L(a)\epsilon(b) \tag{87.118}$$

   for every $ab, \in \mathcal{A}$.

4. *The states φ_t are homomorphisms: $\varphi_t(ab) = \varphi_t(a)\varphi_t(b)$ for every $a, b \in \mathcal{A}$ and $t \geqslant 0$.*

A Schürmann triple with such a L is a **drift**.

PropProcessusGaussien

Proposition 87.26.
Let L be an hermitian conditionally positive linear functional on $\mathcal{A}$. The following conditions are then equivalent:

1. $L|_{K_3} = 0$;
2. $\eta|_{K_2} = 0$;
3. $L(b^*b) = 0$ *for every* $b \in K_2$;
4. $\eta(ab) = \epsilon(a)\eta(b) + \eta(a)\epsilon(b)$ *for every* $a, b \in \mathcal{A}$;
5. $\pi(a) = \epsilon(a)1$ *for every* $a \in \mathcal{A}$.

We say that the triple (π, η, L) with a L satisfying the above conditions describes a **Gaussian**.

Proposition 87.27.
A generator of a Gaussian process satisfies

$$L = \sum \alpha_k \epsilon'_k + \sum B_{ij}\epsilon''_{ij} \tag{87.119}$$

where the ϵ'_k and ϵ''_{ij} are the derivatives of the representation defined in proposition 87.12.

We suppose now that π is not Gaussian (that is π is not $\epsilon 1$), but

$$\pi = \pi_1 \oplus \pi_2 \tag{87.120}$$

where π_1 is Gaussian.

¡¡ Avertissement/question à la lectrice !! 87.28
Il est dit que l'espace sur lequel π agit est

$$H_1 = \bigcap_{\alpha \in K_1} \ker \pi(a). \tag{87.121}$$

Cependant par définition de K_1, $\epsilon(a) = 0$ dès que $a \in K_1$, donc $\pi_1(a) = 0$.
Comment ça marche ?

87.6.1 Poisson process

PropDefPoissonnL

Proposition 87.29.
Let $L\colon \mathcal{A} \to \mathbb{C}$ be a linear hermitian conditionally positive functional. Then the following conditions are equivalent:

(1) there exists a state $\varphi\colon \mathcal{A} \to \mathbb{C}$ and a real $\lambda > 0$ such that

$$L(b) = \lambda\big(\varphi(b) - \epsilon(b)\big) \tag{87.122}$$

for every $b \in \mathcal{A}$;

(2) there exists a Schürmann triple (π, η, L) such that the cocycle is trivial, that is there exists a $h \in H$ such that

$$\eta(b) = \big(\pi(b) - \epsilon(b)\big)h. \tag{87.123}$$

A generator L satisfying these properties is a **Poisson generator**. The map η is the **coboundary** of h.

¡¡ Avertissement/question au lecteur !! 87.30
Il faut expliquer cette histoire d'opérateurs $\mathcal{O}$.

87.6.2 Gaussian cocycle

Definition 87.31.
Let π be a representation of $\mathcal{A}$ on an Hilbert space H. We define

$$H_\epsilon = \bigcap_{a \in K_1} \ker \pi(a). \tag{87.124}$$

*The restriction of π on H_ϵ is the **Gaussian part** of π.*

In order to see that this definition makes sense, we have to prove that $\pi(x)f \in H_\epsilon$ whenever $x \in \mathcal{A}$ and $f \in H_\epsilon$. If $a \in K_1$, we have $\pi(a)\pi(x)f = \pi(ax)f = 0$ because K_1 is an ideal.

LempiepsHepsUni

Lemma 87.32.
The representation respects the decomposition $H = H_\epsilon \oplus H_\epsilon^\perp$. Namely,

(1) The representation acts as the counit on its Gaussian part:

$$\pi(a)|_{H_\epsilon} = \epsilon(a)\,\mathrm{Id}\,|_{H_\epsilon}. \tag{87.125}$$

(2) We have $\pi(\mathcal{A}_q)H_\epsilon^\perp \subset H_\epsilon^\perp$.

Proof. (1) Let $f \in H_\epsilon$. We use proposition 87.24 taking into account the fact that $\pi(d_j)f = \pi(v_j^*)f = \pi(k_3)f = 0$. It remains $\pi(x)f = c_0 f = \epsilon(x)f$.

(2) Let $h \in H_\epsilon^\perp$. For every $v \in H_\epsilon$ we have $\langle v, h\rangle = 0$. If $x \in \mathcal{A}_q$, we have

$$\langle v, \pi(x)h\rangle = \langle \pi(x^*)v, h\rangle = 0 \tag{87.126}$$

since $\pi(x^*)v = \epsilon(x^*)v \in H_\epsilon$.

□

Proposition 87.33.
Let $w = \sum_{j=1}^{n} v_j^* v_j$. Then

$$H_\epsilon = \ker \pi(w) = \bigcap_{j=1}^{n} \ker \pi(v_j). \tag{87.127}$$

Proof. Let $f \in H_\epsilon$. Since $\pi(v_j)f = 0$ for every j, taking the sum we have $\pi(w)f = 0$. Thus $H_\epsilon \subset \ker \pi(w)$. Let now $f \in \ker \pi(w)$. We have

$$\begin{aligned} 0 &= \langle f, \pi(w)f \rangle \\ &= \sum_{j=1}^{n} \langle \pi(v_j)f, \pi(v_j)f \rangle \\ &= \sum_{j} \|\pi(v_j)f\|^2. \end{aligned} \tag{87.128}$$

Thus for every j we have $\|\pi(v_j)f\| = 0$ and $f \in \ker \pi(v_j)$. That proves the inclusion

$$\ker \pi(w) \subset \bigcap_{j=1}^{n} \ker \pi(v_j). \tag{87.129}$$

Let $f \in \ker \pi(v_j)$ and $i < j$. We have $\pi(u_{jj})f = f$ and from (73.33b) we know that

$$\pi(u_{ij})f = \pi(u_{ij})\pi(u_{jj})f = q\pi(u_{jj})\pi(u_{ij})f, \tag{87.130}$$

in other words,

$$\big[\operatorname{Id} - q\pi(u_{jj})\big]\pi(u_{ij})f = 0. \tag{87.131}$$

From corollary 73.15, the operator $\operatorname{Id} - q\pi(u_{jj})$ is invertible. Applying the inverse on the equation (87.131) brings $\pi(u_{ij})f = 0$ and then $f \in \bigcap_{i<j} \ker \pi(u_{ij})$. Thus we have

$$\ker \pi(v_j) \subset \bigcap_{i<j} \ker \pi(u_{ij}). \tag{87.132}$$

Combining with (87.129), we have

$$\ker \pi(w) \subset \bigcap_{1 \leqslant i < j \leqslant n} \ker \pi(u_{ij}). \tag{87.133}$$

Let us now prove that

$$\bigcap_{j=1}^{n} \ker \pi(v_j) \subset \bigcap_{1 \leqslant j < i \leqslant n} \ker \pi(u_{ij}). \tag{87.134}$$

For that consider $f \in H$ such that $\pi(v_j)f = 0$ for every $j = 1, \cdots, n$ and apply π to the relation (73.52). For each i we have

$$\operatorname{Id} = \sum_{s=1}^{n} \pi(u_{si}^*)\pi(u_{si}). \tag{87.135}$$

Applying to f and cutting the sum in three parts,

$$f = \sum_{s<i} \pi(u_{si}^*)\pi(u_{si})f + \pi(u_{ii}^*)\pi(u_{ii})f + \sum_{s>i} \pi(u_{si}^*)\pi(u_{si})f. \tag{87.136}$$

The inclusion (87.132) shows that $\pi(u_{si})f = 0$ when $s < i$; in the same time $\pi(u_{ii})f = f$. We are thus left with

$$0 = \sum_{s>i} \pi(u_{si}^*)\pi(u_{si})f \tag{87.137}$$

That implies in particular that

$$0 = \sum_{s>i} \langle f, \pi(u_{si})^*\pi(u_{si})f \rangle = \sum_{s>i} \|\pi(u_{si})f\|, \tag{87.138}$$

and that $\pi(u_{si})f = 0$. What we just proved is that

$$\bigcap_{j=1}^{n} \ker \pi(v_j) \subset \bigcap_{1\leqslant j<i\leqslant n} \ker \pi(u_{ij}). \tag{87.139}$$

Combining with (87.133), we have

$$\bigcap_{j=1}^{n} \ker \pi(v_j) \subset \bigcap_{i\neq j} \ker \pi(u_{ij}). \tag{87.140}$$

Since K_1 is made of u_{ij} $(i \neq j)$ and v_j, we also have

$$\bigcap_{j=1}^{n} \ker \pi(v_j) \subset H_\epsilon \tag{87.141}$$

EqInclkerpivjHesp

and then $\ker \pi(w) \subset H_\epsilon$ because of (87.129). The inclusion (87.141) also shows that the intersection $\bigcap_{j=1}^{n} \ker \pi(v_j)$ is equal to H_ϵ. This concludes the proof. □

Proposition 87.34.
Let $h_1, \dots, h_n \in H$. If (π, η, L) is a Gaussian triple and if η satisfies $\eta(v_j^*) = h_j$ for $j = 1, \dots, n-1$, then η reads

$$\begin{aligned} \eta\colon \mathcal{A}_q &\to H \\ a &\mapsto \sum_{k=1}^{n-1} i\epsilon_k'(a)h_k. \end{aligned} \tag{87.142}$$

EqDefEtaCocyGaussU

Proof. The first point to be shown is that the given η is a Gaussian π-ϵ-1-cocycle. For that we have

$$\begin{aligned} \eta(ab) &= \sum_{k=1}^{n-1} i\Big(\epsilon_k'(a)\epsilon(b) + \epsilon(a)\epsilon_k'(b)\Big)h_k \\ &= \eta(a)\epsilon(b) + \epsilon(a)\eta(b). \end{aligned} \tag{87.143}$$

Since π is part of a Gaussian triple, it satisfies $\epsilon(a) = \pi(a)$. Thus we obtain the cocycle condition $\eta(ab) = \pi(a)\eta(b) + \eta(a)\epsilon(b)$. The fact for a cocycle to be Gaussian is $\eta|_{K_2} = 0$ (proposition 87.26). Here, since $\epsilon_k'(ab) = \epsilon_k'(a)\epsilon(b) + \epsilon(a)\epsilon_k'(b)$, we have

$$\eta(ab) = 0 \tag{87.144}$$

whenever a and b belong to K_1. We proved that the map η of equation (87.142) is a Gaussian cocycle.

Since $v_j = u_{jj} - 1$, we have

$$\begin{aligned} \epsilon_k'(0, v_j^*) &= \frac{\partial}{\partial\theta_k}\big[\epsilon(\theta, v_j^*)\big]_{\theta=0} \\ &= \frac{\partial}{\partial\theta_k}\big[\epsilon(\theta, u_{jj}^*) - \epsilon(\theta, 1)\big]_{\theta=0} \\ &= \frac{\partial}{\partial\theta_k}\big[e^{-i\theta_j}\big] \\ &= -i\delta_{kj}, \end{aligned} \tag{87.145}$$

so that

$$\eta(v_j^*) = \sum_{k=1}^{n-1} i(-i)\delta_{kj}h_k = h_j. \tag{87.146}$$

Since proposition 87.11 states that a cocycle is determined by its values on the elements v_j^* $(j = 1, \dots n-1)$, the formula (87.142) determines the unique cocycle such that $\eta(v_j^*) = h_j$. □

PropGaussCocycle

Proposition 87.35.
Every Gaussian cocycle reads

$$\eta(a) = \sum_{k=1}^{n-1} i\epsilon'_k(a)h_k \tag{87.147}$$

with $h_k = \eta(v_j^*)$.

Proof. For each $x \in \mathcal{A}_q$ we have the decomposition

$$x = \epsilon(x)1 + \sum_{j=1}^{n-1} \epsilon'_j(x)d_j + k_2. \tag{87.148}$$

Applying η and taking into account $\eta(1) = \eta(k_2) = 0$ (the second equality is because η is Gaussian), we have

$$\eta(x) = \sum_{k=1}^{n-1} \epsilon'_k(x)\eta(d_k). \tag{87.149}$$

It remains to be proven that $\eta(d_j) = i\eta(v_j^*)$. We have

$$\begin{aligned} \eta(v_j^*) = \eta(u_{jj}^*) &= \sum_{k}^{n-1} \epsilon'_k(u_{jj}^*)\eta(d_k) \\ &= -\sum_{k=1}^{n-1} i\delta_{jk}\eta(d_k) \\ &= -i\eta(d_j). \end{aligned} \tag{87.150}$$

¡¡ Avertissement/question à la lectrice !! 87.36
Ici, Anna fait un truc plus compliqué.

□

87.6.3 Gaussian generator

Following proposition 87.29, a Gaussian generator has to vanish on K_3 and being hermitian, conditionally positive.

Let us begin to investigate linear functional ψ vanishing on K_3. Since an element of $\mathcal{A}_q$ decomposes into

$$x = \epsilon(x)1 + \sum_{j=1}^{n-1} \epsilon'_j(x)d_j + \sum_{1\leqslant i\leqslant j\leqslant n-1} \epsilon'_{ij}(x)v_iv_j^* + k_3, \tag{87.151}$$

the functional ψ is determined by its values on the elements 1, d_j $(j = 1, \dots, n-1)$ and $v_iv_j^*$ $(1 \leqslant i \leqslant j \leqslant n-1)$. But the functionals ϵ, ϵ'_k and ϵ''_{kl} are separating these points, so

$$\psi = \alpha_0\epsilon + \sum_{k=1}^{n-1} \alpha_k\epsilon'_k + \sum_{k,l}^{n-1} \beta_{kl}\epsilon''_{kl}. \tag{87.152}$$

Since ϵ''_{kl} is symmetric with respect to k, l, we can choose $\beta_{kl} = \beta_l k$.

Propproppsidefposhermconsdpos

Proposition 87.37.
Let ψ be of the form

$$\psi = \alpha_0\epsilon + \sum_{k=1}^{n-1} \alpha_k\epsilon'_k + \sum_{k,l}^{n-1} \beta_{kl}\epsilon''_{kl} \tag{87.153}$$

Eqproppsidefposhermconsdpos

and denote by B the matrix made of the numbers β_{ij}. Then

(1) $\psi(1) = 0$ *if and only if* $\alpha_0 = 0$.

(2) The functional ψ *is hermitian if and only if* $\alpha_i \in \mathbb{R}$ *and* $\beta_{ij} \in \mathbb{R}$.

¡¡ Avertissement/question à la lectrice !! 87.38
Et α_0 *aussi doit être réel pour que* ψ *soit hermitienne ?*

(3) If the matrix B *is positive defined, the functional* ψ *is conditionally positive.*

(4) If ψ *is conditionally positive and hermitian, then* B *is positive defined.*

Proof. (1) Since $\epsilon'_k(1) = \epsilon''_{kl}(1) = 0$ and $\epsilon(1) = 1$, we have $\psi(1) = \alpha_0$.

(2) Let us begin with the elements of B_1. If a belongs to the span of B_1,

$$\psi(a^*) = \sum_{k=1}^{n-1} \alpha_k \epsilon'_k(a^*) = \sum_{k=1}^{n-1} \alpha_k \overline{\epsilon'_k(a)}, \tag{87.154}$$

so $\psi(a^*) = \overline{\psi(a)}$ if and only if $\alpha_k \in \mathbb{R}$.

¡¡ Avertissement/question à la lectrice !! 87.39
Il me semble qu'il faut aussi prendre $\alpha_0 \in \mathbb{R}$.

For elements in B_2, we use the formula (87.106b):

$$\begin{aligned} \psi\big((v_k v_l^*)^*\big) &= \sum_{i,j=1}^{n-1} B_{ij} \epsilon''_{ij}(v_l v_k^*) \\ &= \sum_{i,j} B_{ij}(\delta_{jl}\delta_{ik} + \delta_{jk}\delta_{il}) \\ &= B_{kl} + B_{lk} \\ &= 2B_{kl} \end{aligned} \tag{87.155}$$

because B is symmetric. On the other hand we have similarly

$$\psi(v_k v_l^*) = 2B_{kl}, \tag{87.156}$$

so that B_{kl} has to be real when ψ is Hermitian.

(3) Let $a \in K_1$. Using Leibniz rule we have

$$\begin{aligned} \epsilon(a^*a) &= 0 \\ \epsilon'_i(a^*a) &= 0 \\ \epsilon''_{ij}(a^*a) &= \epsilon'_j(a^*)\epsilon'_i(a) + \epsilon'_i(a^*)\epsilon'_i(a). \end{aligned} \tag{87.157}$$

Thus

$$\psi(a^*a) = 2\sum_{ij} B_{ij}\epsilon'_i(a^*)\epsilon'_j(a) = 2\sum_{ij} B_{ij}\overline{\epsilon'_i(a)}\epsilon'_j(a). \tag{87.158}$$

Let us notice that $\overline{\epsilon'_i(a)}\epsilon'_j(a)$ is a general product $x_i x_j$ with $x \in \mathbb{R}^{n-1}$ because, if $a = \sum_k x_k d_k$,

$$\epsilon_i(a) = \sum_k x_k \epsilon'_i(d_k) = i x_i, \tag{87.159}$$

so that $\overline{\epsilon'_i(a)}\epsilon'_j(a) = (-i)i x_i x_j = x_i x_j$. By definition if B is positive defined, the right hand side of equation (87.158) is positive and ψ is conditionally positive.

(4) If ψ is Hermitian, we already see that B_{ij} are real. Now positivity of

$$\sum_{ij} B_{ij}\overline{\epsilon'_i(a)}\epsilon'_j(a) \tag{87.160}$$

shows that B is positive defined.

¡¡ Avertissement/question au lecteur !! 87.40
Que signifie une matrice «définie positive» ?

□

Corollary 87.41.
If ψ is a Gaussian generator, there exist reals numbers $\alpha_1, \cdots, \alpha_{n-1}$ and a symmetric positive defined matrix B such that

$$\psi = \sum_{k=1}^{n-1} \alpha_k \epsilon'_k + \sum_{i,j=1}^{n-1} B_{ij} \epsilon''_{ij}. \tag{87.161}$$

Proof. By definition, a Gaussian generator is Hermitian, conditionally positive and vanishes on K_3. We already know that it is of the form

$$\psi = \alpha_0 \epsilon + \sum_{k=1}^{n-1} \alpha_k \epsilon'_k + \sum_{k,l}^{n-1} B_{kl} \epsilon''_{kl}. \tag{87.162}$$ EqFormgenpsiGauss

If ψ is Hermitian, $\alpha_i = 0$ and $B_{ij} \in \mathbb{R}$. If ψ is Hermitian and conditionally positive, then B is positive definite.

¡¡ Avertissement/question à la lectrice !! 87.42
Pourquoi $\alpha_0 = 0$? C'est-à-dire : pourquoi $\psi(1) = 0$?

□

Proposition 87.43.
consider the Gaussian cocycle (proposition 87.35)

$$\eta(a) = \sum_{k=1}^{n-1} i\epsilon'_k(a) h_k. \tag{87.163}$$

(1) If the numbers $\langle h_k, h_l \rangle$ are real, then the formula

$$\psi = \frac{1}{2} \sum_{k,l=1}^{n-1} \langle h_k, h_l \rangle \epsilon''_{kl} \tag{87.164}$$

defines a conditionally positive Hermitian functional associated with the cocycle η.

(2) If one of $\langle h_k, h_l \rangle$ is not real, then there are no Hermitian functional associated with η.

Proof. (1) We denote $\alpha_{kl} = \langle h_k, h_l \rangle$. If these numbers are real, we have $\alpha_{kl} = \alpha_{lk}$ since $\langle h_k, h_l \rangle = \overline{\langle h_l, h_k \rangle}$. From corollary 87.10(3), the functional ψ is associated with η if

$$-\langle \eta(a^*), \eta(b) \rangle = \epsilon(a)\psi(b) - \psi(ab) + \psi(a)\epsilon(b). \tag{87.165}$$

Using the Leibniz rule (87.102) (reduced by the fact that a and b belong to K_1), we have

$$\begin{aligned}
\psi(ab) &= \frac{1}{2} \sum_{kl} \alpha_{kl} \epsilon''_{kl}(ab) \\
&= \frac{1}{2} \alpha_{kl} \big(\epsilon'_l(a)\epsilon'_k(b) + \epsilon'_k(a)\epsilon'_l(b) \big) \\
&= \sum_{kl} \langle h_k, h_l \rangle \epsilon'_k(a)\epsilon'_l(b) \\
&= \sum_{kl} \langle \overline{\epsilon'_k(a) h_k}, \epsilon'_l(b) h_l \rangle \\
&= \langle \eta(a^*), \eta(b) \rangle.
\end{aligned} \tag{87.166}$$

The functional ψ is hermitian because it is of form (87.153) with $\alpha_i = 0$ and $\alpha_{kl} \in \mathbb{R}$. Moreover the matrix $B_{kl} = \langle h_k, h_l \rangle$ is symmetric and real and defines a Gaussian cocycle, so that the functional ψ is Hermitian.

(2) The number $\langle h_k, h_l\rangle$ decomposes in real and imaginary parts as

$$\langle h_k, h_l\rangle = \beta_{kl} + id_{kl} \tag{87.167}$$

with

$$\begin{aligned}\beta_{kl} &= \frac{1}{2}\big(\langle h_k, h_l\rangle + \langle h_l, h_k\rangle\big)\\ d_{kl} &= \frac{-i}{2}\big(\langle h_k, h_l\rangle - \langle h_l, h_k\rangle\big)\end{aligned} \tag{87.168}$$

If ψ is Hermitian, for a and b in K_1 we have

$$\begin{aligned}\psi(ab) &= \langle \eta(a^*), \eta(b)\rangle\\ &= \sum_{kl}\langle \epsilon'_k(a^*)h_k, \epsilon'_l(b)h_l\rangle\\ &= \sum_{kl}\overline{\epsilon'_k(a^*)}\epsilon'_l(b)\langle h_k, h_l\rangle\\ &= \sum_{kl}\epsilon'_k(a)\epsilon'_l(b)\langle h_k, h_l\rangle\end{aligned} \tag{87.169}$$

Computing that way $\psi(v_iv_j)$ and $\psi(v_j^*v_i^*)$ and taking into account the formula $\epsilon'_k(v_i) = i\delta_{ki}$, we found that ψ is only Hermitian if

$$\langle h_i, h_j\rangle = \langle h_j, h_i\rangle, \tag{87.170}$$

which means that $\langle h_i, h_j\rangle$ is real.

□

Let π be the representation of $\mathcal{A}_q$ on H and

$$H = H_\epsilon \oplus H_\epsilon^\perp \tag{87.171}$$

be the decomposition of H into Gaussian and non-Gaussian parts. The Schürmann triple is completed by $\psi\colon \mathcal{A}_q \to \mathbb{C}$ and $\eta\colon \mathcal{A}_q \to H$. We consider the corresponding decomposition of η:

$$\eta = \eta^\epsilon \oplus \eta^\perp \tag{87.172}$$

where $\eta^\epsilon = \mathtt{proj}_{H_\epsilon} \circ \eta$ and $\eta^\perp = \mathtt{proj}_{H_\epsilon^\perp} \circ \eta$. Let us denote by h^ϵ the projection of h on H_ϵ. We want to see under which conditions η is a cocycle associated with a generator.

A Gaussian generator reads

$$\eta(a) = \sum_{k=1}^{n-1} i\epsilon'_k(a)h_k \tag{87.173}$$

where $h_i = \eta(v_i^*)$. A condition to have a generator is to have $\langle h_i, h_j\rangle \in \mathbb{R}$. Now, the Gaussian part of the cocycle reads

$$\eta^\epsilon(a) = \sum_k i\epsilon'_k(a)\,\mathtt{proj}_\epsilon\, h_k. \tag{87.174}$$

It will accept a generator if and only if

$$\langle \mathtt{proj}_\epsilon\, h_i, \mathtt{proj}_\epsilon\, h_j\rangle \tag{87.175}$$

is real.

On the other hand, we can check that η^ϵ is a cocycle for the representation π, i.e.

$$\eta^\epsilon(ab) = \pi(a)\eta^\epsilon(b) + \eta^\epsilon(a)\epsilon(b), \tag{87.176}$$

or

$$\mathtt{proj}_\epsilon\, \eta(ab) = \pi(a)\,\mathtt{proj}_\epsilon\, \eta(b) + \mathtt{proj}_\epsilon\, \eta(a)\epsilon(b). \tag{87.177}$$

From lemma 87.32 we have $\pi(a)\eta^\epsilon(b) = \epsilon(a)\eta^\epsilon(b)$. We write the second term under the form

$$\mathtt{proj}_\epsilon\Big(\pi(a)\big(\eta^\epsilon(b) + \eta^\perp(b)\big)\Big) = \epsilon(a)\eta^\epsilon(b) + \mathtt{proj}_\epsilon\, \pi(a)\eta^\perp. \tag{87.178}$$

The last term vanishes by lemma 87.32.

Chapter 88

Complements

88.1 Alternative formalism for the quantum mechanics

app:Wigner

We can a little reformulate the axioms of the quantum mechanics. Since we are in a Hilbert space $\mathscr{H}$ we can speak about orthogonal projections; if $\phi \in \mathcal{R}$, we can consider the projection on the space spanned by ϕ:

$$P_\phi e_k = \frac{\langle \phi, e_k \rangle}{\|\phi\|}\phi$$

where $\{e_i\}$ is a basis of $\mathscr{H}$. It is pretty clear that

$$\mathrm{Tr}(P_\phi P_\psi) = \frac{|\langle \psi, \phi \rangle|^2}{\|\psi\|\|\phi\|}. \tag{88.1}$$

If $\psi \in \mathcal{R}$ and $\phi \in \mathcal{R}'$ are unimodular, then

$$P(\mathcal{R} \to \mathcal{R}') = \mathrm{Tr}(P_\phi P_\psi), \tag{88.2}$$

so we can express the axioms in terms of projections instead of rays. For notational convenience, we put

$$\mathscr{H}_1 = \{\psi \in \mathscr{H} \text{ tel que } \|\psi\| = 1\}. \tag{88.3}$$

We denote by $\mathscr{S}$ the space of the projections into one dimensional subspaces of $\mathscr{H}$ (in other words $\mathscr{S}$ is the space of physical states) and for $P, Q \in \mathscr{S}$, the transition probability is $P \cdot Q = \mathrm{Tr}(PQ)$. Now a **quantum symmetry** is a map $T\colon \mathscr{S} \to \mathscr{S}'$ such that $(TP) \cdot (TQ) = P \cdot Q$.

One can prove the following:

Theorem 88.1.
If $T\colon \mathscr{S} \to \mathscr{S}'$ is a quantum symmetry, then there exists an operator $U\colon \mathscr{H} \to \mathscr{H}'$ such that

(1) $P_{U\phi} = TP_\phi$,

(2) $U(\xi + \eta) = U(\xi) + U(\eta)$,

(3) $\langle U\xi, U\eta \rangle = \kappa(\langle \xi, \eta \rangle)$ item:cond_3

where P_ψ is the projection onto the one dimensional space spanned by ψ and $\kappa\colon \mathbb{C} \to \mathbb{C}$ fulfils $\kappa(\lambda) = \lambda$ or $\kappa(\lambda) = \overline{\lambda}$ and

(1) $U(\lambda\xi) = \kappa(\lambda)\xi$.

tho:pre_Wigner

Here is why this implies Wigner's theorem as given by theorem 69.2. Let us consider some $\varphi_i \in \mathscr{H}$ such that $\|\varphi_i\| = 1$ and P_{φ_i}, the corresponding projections. Let

$$\Delta(P_1, P_2, P_3) = \langle \varphi_1, \varphi_2 \rangle\langle \varphi_2, \varphi_3 \rangle\langle \varphi_3, \varphi_1 \rangle.$$

It is clear that this expression doesn't depend on the choice of φ_i in its ray. We have

$$\begin{aligned}\Delta(TP_1;TP_2,TP_3) &= \Delta(P_{U\varphi_1},P_{U\varphi_2},P_{U\varphi_3})\\ &= \langle U\varphi_1,U\varphi_2\rangle\langle U\varphi_2,U\varphi_3\rangle\langle U\varphi_3,U\varphi_1\rangle\\ &= \kappa(\langle U\varphi_1,U\varphi_2\rangle)\kappa(\langle U\varphi_2,U\varphi_3\rangle)\kappa(\langle U\varphi_3,U\varphi_1\rangle)\\ &= \kappa\big(\Delta(P_1,P_2,P_3)\big).\end{aligned} \tag{88.4}$$

eq:Delta_T

We can see from this that the choice of $\kappa(\lambda)=\lambda$ or $\kappa(\lambda)=\overline{\lambda}$ is determined by the data of T if $\dim\mathscr{H}\geqslant 2$. In the case where $\dim\mathscr{H}=1$, Δ is always equals to 1 and the equality (88.4) don't give any informations. In the case $\dim\mathscr{H}\geqslant 2$, we can choice φ_1 and φ_2 such that $\langle\varphi_1,\varphi_2\rangle$ takes any value $z\in\mathbb{C}$ with $\|z\|\leqslant 1$. Taking $\varphi_3=\varphi_1+\varphi_2$, we find

$$\Delta(P_1,P_2,P_3)=z(1+\overline{z})^2.$$

which is easily non real for a suitable choice of $z\in\mathbb{C}$. Let us suppose that we have an operator U which satisfies the theorem 88.1. If $\kappa(\lambda)=\lambda$, then

$$U(z\psi+z'\phi)=U(z\psi)+U(z'\phi)=zU(\psi)+zU(\phi) \tag{88.5}$$

and

$$\langle U\psi,U\phi\rangle=\kappa(\langle\psi,\phi\rangle)=\langle\psi,\phi\rangle, \tag{88.6}$$

so that U is linear. If $\kappa(\lambda)=\overline{z}$, then

$$U(z\psi)=\overline{z}U\psi \tag{88.7}$$

and

$$\langle U\xi,U\eta\rangle=\kappa(\langle\xi,\eta\rangle)=\overline{\langle\xi,\eta\rangle}. \tag{88.8}$$

88.2 Description of Maxima computations

SecAppMaxima

This appendix is devoted to the description of the Maxima implementation of my computation related to groups acting on the open orbit. I will adopt mixed notation between the full correct syntax of Maxima and a more suitable mathematical notation.

Fist of all I explicitly gave to Maxima the full 6×6 matrices J_1,W,M,L,J_2 and V. The commutator between two matrices is naturally defined by

$$\mathtt{com}(X,Y)=XY-YX.$$

88.2.1 Projections and decompositions

In order to compute projection of a matrix on a basis (or on a sub-space), we introduce a scalar product on the matrices space:

$$\mathtt{pprod}(X,Y)=\sum_{k,l=1}^{J}X_{kl}Y_{kl} \tag{88.9}$$

EqDefMaxpprod

when X and Y are 6×6 matrices. Now projection of a matrix X on Y is given by

$$\mathtt{proj}(X,Y)=\frac{\mathtt{pprod}(X,Y)}{\mathtt{pprod}(Y,Y)}. \tag{88.10}$$

Now the decomposition of the matrix X in the basis J_1,W,M,L,J_2,V is given by a set of six numbers:

$$\mathtt{decomp6}(X)=[\mathtt{proj}(X,J_1),\cdots,\mathtt{proj}(X,V)].$$

This decomposition works in the sense that the following equality is satisfied:

$$X=\mathtt{decomp6}(X)[1]J_1+\cdots+\mathtt{decomp6}[6]V \tag{88.11}$$

EqMaxPropDec

Because the set $\{J_1, W, M, L, J_2, V\}$ is orthogonal for the product (88.9). A function is also defined in order to easily compute the table of commutators:

$$\texttt{table6}(X, Y) = \texttt{decomp6}(\texttt{com}(X, Y)).$$

The function `table6c` returns the same as `table6` plus a check of equality (88.11):

$$\begin{aligned}\texttt{table6c}(X, Y) = [\,&\texttt{table6}(X, Y),\\ &\texttt{table6}(X, Y)[1] * J1 + \texttt{table6}(X, Y)[2] * W\\ &+ \texttt{table6}(X, Y)[3] * M + \texttt{table6}(X, Y)[4] * L\\ &+ \texttt{table6}(X, Y)[5] * J2 + \texttt{table6}(X, Y)[6] * V\\ &- \texttt{com}(X, Y)];\end{aligned} \tag{88.12}$$

it must return six numbers and a vanishing matrix.

The function `combi6` creates a matrix whose coefficients are given under the form of a vector v of length 6:

$$\texttt{combi6}(v) = v[1]J_1 + \cdots + v[6]V.$$

Functions `decomp4`, `table4` and `table4c` are the same as the 6 corresponding ones, but they decompose into the basis $\{A, B, C, D\}$ instead of $\{J_1, W, M, L, J_2, V\}$. Notice however that the matrices A, B, C and D are not necessary orthogonal in the sense of product (88.9). So we cannot compute the projection of X on A as $\texttt{proj}(X, A)$. In order to correct this problem, we remark that A is the only element containing J_1, so the projection of X on A is the coefficient of J_1 in X. So we implement the following:

$$\texttt{decomp4}(X) = [\texttt{proj}(X, J_1), \texttt{proj}(X, W), \texttt{proj}(X, M), \texttt{proj}(X, L)].$$

88.2.2 Symplectic computations

The script *mesmatrices* contains the general symplectic matrix

$$\texttt{omega} = \begin{pmatrix} 0 & -\alpha & -\beta & -\gamma \\ & 0 & -\delta & -\sigma \\ & & 0 & -\epsilon \\ & & & 0 \end{pmatrix}.$$

The command $\texttt{symple}(X, Y)$ returns $\omega(X, Y)$. As we are searching for ω in the basis $\{A, B, C, D\}$, it is implemented as

$$\texttt{symple}(X, Y) = \sum_{k=1}^{4}\sum_{l=1}^{4} \texttt{omega}[k, l]\,\texttt{decomp4}(X)[k]\,\texttt{decomp4}(Y)[l].$$

The command `cycle` computes cyclic sums which appears in the definition of a symplectic algebra (see equation (59.4)):

$$\texttt{cycle}(X, Y, Z) = \texttt{symple}(\texttt{com}(X, Y), Z) + \texttt{symple}(\texttt{com}(Y, Z), X) + \texttt{symple}(\texttt{com}(Z, X), Y). \tag{88.13}$$

When matrices A, B, C and D are given, the symplectic form is fixed by solving the condition $\texttt{cycle}(X, Y, Z) = 0$. This is done by the Maxima function `solve`:

$$\texttt{solve}\Big([\texttt{cycle}(A, B, C) = 0, \texttt{cycle}(A, B, D) = 0, \texttt{cycle}(C, A, D) = 0, \texttt{cycle}(B, D, C) = 0], [\alpha, \beta, \gamma, \delta, \sigma, \epsilon]\Big).$$

88.2.3 $\mathrm{Ad}(g)^*(\delta\xi^*) = E^*$

For the computations with ξ^* (see equation 84.54), we begin to define the function `xistar` which computes $\xi^*(X)$:

$$\texttt{xistar}(X) := \xi_A * \texttt{decomp4}(X)[1] + \xi_B * \texttt{decomp4}(X)[2] + \xi_C * \texttt{decomp4}(X)[3] + \xi_D * \texttt{decomp4}(X)[4],$$

and then we define the function `delxistar` which computes $\delta\xi^*(X,Y) = \xi^*([X,Y])$:

$$\begin{aligned}\texttt{delxistar}(X,Y) := {} & \xi_A * \texttt{decomp4}(\texttt{com}(X,Y))[1] + \xi_B * \texttt{decomp4}(\texttt{com}(X,Y))[2] \\ & + \xi_C * \texttt{decomp4}(\texttt{com}(X,Y))[3] + \xi_D * \texttt{decomp4}(\texttt{com}(X,Y))[4];\end{aligned}$$

In order to accelerate computation of the g such that $\mathrm{Ad}(g)^*\delta\xi^* = E^*$ (see page 4199), we compute by Maxima products as

$$e^{aA} p e^{-aA},$$

using the extensions `linalg-utilities.lisp` and `matrixexp.lisp` for the matrix exponential. We apply ξ^* to the result — this operation does not require Maxima.

88.3 Some of my favourite mistakes

In this appendix, we explain a lot of mistakes and false ideas that one can commit.

err:decomp

Error 88.2.
The reasoning at page 3607 which deduce $X(t) = e$ from $[X(t)] = cst$ is not a particular case of lemma 53.61, nor a re-proof of it. Here, we decompose $X(t) = gk(t)$ with $k(t)$ in the "right space" from arguments which are true in our case but not in general.

err:Intt_Aut

Error 88.3.
Since $\mathrm{ad}(\mathfrak{a})$ is a subalgebra of $\partial(\mathfrak{a})$, it is the Lie algebra of an unique Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$ from theorem 53.7. Then we deduce that $\mathrm{Int}(\mathfrak{a})$ is a Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$. There is a topological problem: $\mathrm{Int}(\mathfrak{a})$ is defined as an analytic subgroup of $\mathrm{GL}(\mathfrak{a})$. There is no evidence that the topology of the subgroup of $\mathrm{Aut}(\mathfrak{a})$ given by 53.7 is the one of $\mathrm{Int}(\mathfrak{a})$.

So theorem 53.7 is not applicable to find relations between $\mathrm{Int}(\mathfrak{a})$ and $\mathrm{Aut}(\mathfrak{a})$.

err:gross

Error 88.4.
An easy (but false) counter example. If G is a Lie group and H, a closed Lie subgroup of G on which we consider the trivial topology (the open subsets of H are H itself and $\varnothing$). Then it is clear that $\iota\colon H \to H'$ is not continuous, so that H is not a topological subgroup. So the second item of theorem 53.113 is false.

In the definition of a *Lie* subgroup, we want H to be a submanifold. But proposition 49.113 says that a submanifold has at least the induced topology. So a Lie subgroup with the trivial topology doesn't exist. See also the example at page 3161.

err:gp_meme_alg

Error 88.5.
We can say that $e^X e^Y$ can be written as e^Z for a certain $Z \in \mathfrak{g} = \mathfrak{h}$ given by the Campbell-Backer-Hausdorff formula. Since the Lie algebras are the same, the CBH formula of G and H are the same. Then we can forget the assumption which says that H is a subgroup of G. Thus two groups which have the same algebra are the same on a neighbourhood of e. This is false because there are no guaranties for X, $Y \in A$ that $[X,Y]$ still belongs to A.

err:f_dege

Error 88.6.
If $\mathfrak{g}^\perp = \mathfrak{g}$, then f is degenerate. But we said in the assumptions that f were nondegenerate. So $\mathfrak{g} = 0$. Be careful: we had shown that *if f is nondegenerate*, then $f = \lambda B$. Here we pose $f = \mathrm{Tr}$ and we have to show that it is nondegenerate in order to prove that $B = \lambda\,\mathrm{Tr}$.

Error 88.7.

ProperrProdInvarDiffeo

Proposition 88.8.
Let $\phi\colon G_1 \to G_2$ be a Lie group diffeomorphism such that

$$\phi \circ L_g = L_{g'} \circ \phi. \tag{88.14}$$

Eqerrphideformdiff

If K is a left invariant kernel on G_1, then $K' = \phi^* K$ is a left invariant kernel on G_2.

Proof. Just a computation using left invariance of K:

$$\begin{aligned} L_g^* K' &= (\phi \circ L_g)^* K \\ &= (L_{g'} \circ \phi)^* K \\ &= \phi^* L_{g'}^* K \\ &= \phi^* K \\ &= K'. \end{aligned}$$

□

The mistake is the fact that equation (88.14) should better be written

$$\phi \circ L_g(h) = L_{g'} \circ \phi(h)$$

where g' depends on h. But the kernel is a *three* points function.

88.4 Frequently asked questions

88.4.1 Are open orbits simply connected?

Yes, they are! See 84.3.

88.4.2 Why not de Sitter?

Troisième partie

Matlab

Chapitre 89

Matlab

(1) Les corrigés sont rédigés pour Octave. De petites différences avec Matlab existent.

(2) Les exercices des séances sont tirés des notes « Introduction au logicile Matlab » qui fut donné à Louvain-la-Neuve sous le nom BIR1200. Les exercices des tests sont dûs à Laurent Claessens et Yannick Voglaire.

(3) Merci à J.J. pour m'avoir signalé que `VerbatimInput` créait des problèmes avec `hyperref`, puis à Tanguy Briançon et Jean-Côme Charpentier pour les avoir résolu.

(4) Merci de me signaler toute erreur ou imprécision. Plus vous vous plaignez, plus les étudiants de l'année prochaine auront un document de qualité :)

89.1 Exercices des séances

89.1.1 Bases et calcul matriciel

Exercise 19 exoSC_serie1-0001

Évaluez

$$\frac{\sqrt{\sqrt{|x|}+1}\big(\sin(\exp x^3)+1\big)}{\arctan(x^2)+\Big(\ln\big(\sqrt{|x|}+1\big)\Big)^{3/2}} \tag{89.1}$$

pour $x=-1.2$.

corrSC_serie1-0001

Correction of the exercise 19

Pour produire ce genre d'expressions complexes, une bonne idée est de la diviser en plusieurs morceaux afin d'éviter de devoir travailler avec trop de parenthèses en même temps.

```
function y = ma_fonction(arg)
  sq = sqrt(abs(arg))+1;
  a = sqrt(sq);
  b = sin( exp(arg.^3)+1 );
  numerateur = a*b; % 1.3371
  c = atan(arg.^2);
  d = (log(sq)).^(3/2);
  denominateur = c+d; % 1.6001
  y = numerateur/denominateur;  % 0.83566
end

reponse = ma_fonction(-1.2)
```

tex/matlab/SC_exo_1–1.m

Exercise 20 exoSC_serie1-0002

Donner une instruction pour construire le vecteur

$$v = (100.5, 90.5, \ldots, 10.5, 0.5). \tag{89.2}$$

corrSC_serie1-0002

Correction of the exercise 20

Une bonne lecture est
http://en.wikibooks.org/wiki/Octave_Programming_Tutorial/Vectors_and_matrices.

```
v = 100.5:-10:0.5
```

tex/matlab/SC_exo_1–2.m

La subtilité de cet exercice est que l'on peut demander un pas négatif.

Exercise 21 exoSC_serie1-0003

Quelques manipulations de matrices.

(1) Construire une matrice $A = (a_{ij})$ de genre 6×6, définie par

$$A = I + u^t u/4 \tag{89.3}$$

où $u = (1, 2, 3, 4, 5, 6)$.

(2) Ajouter 2 à l'élément a_{23} et multiplier par[1] $\ln(2)$ la deuxième colonne de A ; on appellera B la nouvelle matrice ainsi obtenue.

(3) Calculer la matrice inverse de B et vérifier que le produit BB^{-1} donne (approximativement) l'identité.

(4) Résoudre le système $Ax = u^t$ et vérifier que la colonne x obtenue est bien solution.

corrSC_serie1-0003

Correction of the exercise 21

```
u = 1:6         # u = 1:1:6 fait la même chose.
A = eye(6)+(u'*u)/4

B = A

B(2,3) = A(2,3)+2
B(:,2) = log(2)*B(:,2)     # B(:,2) représente la deuxième colone de B

C = inv(B)*B
comparaison = eye(6)-C

erreur_max = max(max(abs(comparaison)))
# max appliqué à une matrice retourne le vecteur
# qui contient le plus grand de chaque colonne.
# Donc il faut appliquer deux fois max pour prendre le max de ces max

solution = A\u'

verification = A*solution - u'
erreur_max = max(abs(verification))
```

tex/matlab/SC_exo_1–3.m

1. Matlab donne-t-il le logarithme en base e ou en base 10 ?

Exercise 22 exoSC_serie1-0004

Quelques exercices sur les matrices.

(1) Donner des instructions (les plus simples possibles) pour produire la matrice A de genre 10×10 ayant la forme suivante :

$$A = \begin{pmatrix} \pi & 0 & 0 & \cdots & -1 \\ 0 & \pi & 0 & \cdots & 0 \\ 0 & 0 & \pi & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & \cdots & 0 & 0 & \pi \end{pmatrix} \tag{89.4}$$

. Remarque : les éléments représentés par des pointillés sont tous nuls sauf sur la diagonale principale de A.

(2) Calculer les trois premiers éléments de la diagonale principale de A^{-1} et A^5.

corrSC_serie1-0004

Correction of the exercise 22

```
A = pi*eye(10)
A(1,10) = -1
A(10,1) = 1

inverse = inv(A)
puissance = A^5

diag(inverse)(1:3)
   %0.28903
   %0.31831
   %0.31831


diag(puissance)(1:3)
    %11.665
   %306.020
   %306.020
```

tex/matlab/SC_exo_1–4.m

Exercise 23 exoSC_serie1-0005

On considère la combustion du propane

$$C_3H_5 + 5O_2 + (\text{air en excès}) \rightarrow 3CO_2 + 4H_20 + (\text{air en excès}) \tag{89.5}$$

en présence d'un excès d'air de 25%, ce qui signifie que l'air fournit est égal à 125% de ce qui est requis pour une combustion complète. On demande de calculer le nombre de moles d'air nécessaires à l'entrée pour 100 moles de gaz sortant (celui-ci étant composé de CO_2, de H_20, de O_2 et de N_2). Pour répondre à cette question, on notera

— P le nombre de moles de propane entrant ;

— A le nombre de moles d'air entrant ;

— C le nombre de moles de CO_2 sortant ;

— W le nombre de moles de H_2O sortant ;

— N le nombre de moles de N_2 sortant ;

— X le nombre de moles O_2 sortant ;

toutes ces quantités sont pour 100 moles de gaz sortant.

(1) Montrer que ces quantités sont liées par les équations suivantes :

$$\begin{cases} 3P = C & (89.6a) \\ 4P = W & (89.6b) \\ 0.21A = C + \dfrac{W}{2} + X & (89.6c) \\ 0/79A = N & (89.6d) \\ 0.21A = (1.25)(5P)\ C + W + N + X = 100 & (89.6e) \end{cases}$$

(on considère que l'air entrant est composé de 21% de O_2 et de 79% de N_2).

(2) Résoudre ce système et déterminer en particulier A.

corrSC_serie1-0005

Correction of the exercise 23

En ce qui concerne l'équation chimique, on a

$$P(C_3H_8) + \cdots O_2 + A(0.21O_2 + 0.79N_2) \to C(CO_2) + W(H_2O) + \frac{N}{0/79}(0.21O_2 + 0.79N_2) \quad (89.7)$$

Le $N/0.79$ sert à faire qu'il y ait N moles de N_2 qui sortent, comme demandé. À partir de là, il faut comprendre que le nombre de moles de O_2 qui sortent est $(0.21) * (N/0.79)$. Cela est X.

Les trois petits points signifie que ce O_2 est « virtuel ». En réalité il est inclus dans l'air, et donc dans le A

- bilan carbone : $3P = C$
- bilan hydrogène : $8P = 2W$
- bilan azote : $0.79A = N$

En ce qui concerne le bilan d'oxygène, il y a deux choses à faire.

- D'abord le $5O_2$ doit venir de l'air, et on sait qu'il en fait $5P$, donc on peut croire que $0.21A = 5P$. Hélas, les choses ne sont pas aussi simple : le A fournit un *excès* d'air. Donc le $0.21A$ n'est en réalité pas $5P$, mais 125% de $5P$. Nous avons donc $0.21A = (1.25) * (5P)$
- Le bilan oxygène (à compter en atomes, et non en molécules O_2 !!) s'écrit $0.21A = C + (W/2) + X$

Enfin, on impose qu'il y ait exactement 100 moles qui sortent, c'est-à-dire $C + W + B + X = 100$.

Le reste du problème est pour Matlab.

```
M=
[3       0    -1     0    0 0
 4       0    0    -1     0 0
(1.25)*5   -0.21 0       0    0 0
 0       0.21   -1     -0.5   0 -1
 0       0    1    1    1 1
 0       0.79   0    0    -1   0
 ]
u = [0 0 0 0 100 0]

reponse = M\u'
   %3.1484
   %93.7031
   % 9.4453
   %12.5937
```

```
    %74.0255
    % 3.9355

air = reponse(2)  % Parce qu'on a mis A dans la 2ième colone de M
 %93.703
```

tex/matlab/SC_exo_1–5.m

89.1.2 Vecteurs à éléments équidistants

Exercise 24 exoSC_serie2-0001

Pour calculer une somme.

(1) Construire le vecteur $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{1000})$.

(2) En déduire le vecteur $(1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \dots, \frac{1}{10^6})$

(3) Utiliser ce dernier vecteur pour calculer $\sum_{k=1}^{1000} \frac{1}{k^2}$.

corrSC_serie2-0001

Correction of the exercise 24

```
longueur = 1000
a = 1:1:longueur
b = 1./a
c = b.^2

sum(c)  %1.6439
```

tex/matlab/SC_exo_2–1.m

Notez que l'on s'est bien garder de coder « en dur » le nombre 1000. Nous avons une fois pour toutes posé `longueur=1000` et puis nous avons utilisé la variable.

Il est fortement conseillé de travailler d'abord avec `longueur=4` pour voir si tout va bien. Au moins jusqu'à 4 vous devriez être capables de faire les calculs à la main et détecter des erreurs si il y en a.

Exercise 25 exoSC_serie2-0002

Calculer

$$\sum_{n=0}^{100} \sin^n(x) \tag{89.8}$$

pour $x = \pi/5$.

corrSC_serie2-0002

Correction of the exercise 25

```
function y = somme(arg)
  taille = 10
  v = 0:taille
  w = sin(arg).^v
    # L'astuce est de faire a.^v pour faire [a^i for i in v]
  y = sum(w)
end

somme(pi/5) %2.4189
```

tex/matlab/SC_exo_2–2.m

Exercise 26 exoSC_serie2-0003

Donner un tableau de valeurs de la fonction $f(x) = \exp(\sin^2(x))$ pour 20 valeurs de x équidistantes de 0 à $\pi/2$.

corrSC_serie2-0003

Correction of the exercise 26

```
debut = 0
fin = pi/2
nombre = 20

v = debut:(fin-debut)/(nombre-1):fin

# Juste pour rire : vérification que c'est bien équidistant :
for i = 2:length(v)      # length(v) donne la longueur du vecteur v
  v(i)-v(i-1)
endfor

function y = f(x)
  y = exp(sin(x).^2)
endfunction

f(v)
%1.0000
%1.0068
%...
%2.6998
%2.7183
```

tex/matlab/SC_exo_2–3.m

Exercise 27 exoSC_serie2-0004

Représenter les graphes des fonctions suivantes :

(1) $f\colon [0, 2\pi] \to \mathbb{R}$, $x \mapsto \sin(x) + \frac{1}{3}\sin(3x) + \frac{1}{5}\sin(5x) + \frac{1}{7}\sin(7x)$,

(2) $f\colon [-3, 3] \to \mathbb{R}$, $x \mapsto \sqrt[3]{x^2|x-2|}$,

(3) $f\colon [10^{-2}, \pi] \to \mathbb{R}$, $x \mapsto \sqrt{x}\sin(1/x)$.

corrSC_serie2-0004

Correction of the exercise 27

```
function y = f1(x)
  y = sin(x)+sin(3*x)/3+sin(5*x)/5+sin(7*x)/7
end

function y = f2(x)
  y = (x.^2.*abs(x-2)).^(1/3)
end

function y = f3(x)
  y = sqrt(x).*sin(1./x)
end

# Créer un vecteur avec les valeurs où on va calculer la fonction
echantillon = 0:0.1:2*pi
```

```
# Créer le vecteur avec les valeurs de la fonction
v = f1(echantillon)
# Créer le graphique
plot(echantillon,v)
# Enregistrer le graphique dans exo24_f1.ps
print -dps exo24_f1.ps

echantillon = -3:0.1:3
v = f2(echantillon)
plot(echantillon,v)
print -dps exo24_f2.ps

echantillon = 10^(-2):0.1:pi
v = f3(echantillon)
plot(echantillon,v)
print -dps exo24_f3.ps

# Notez la différence avec ce zoom sur la partie 0->0.5
# avec un pas de 0.001 au lieu de 0.1

echantillon = 10^(-2):0.001:0.4
v = f3(echantillon)
plot(echantillon,v)
print -dps exo24_f3_zoom.ps

# Il y a une oscillation infinie qu'on devine maintenant
# mais qui était presque invisible sur le graphique avec un pas de 0.1.
```

tex/matlab/SC_exo_2–4.m

Exercise 28 exoSC_serie2-0005

Dans un problème d'écoulement turbulent d'un fluide, l'équation

$$1 = \sqrt{c_f}\big(-0.4 + 1.74\ln(Re\sqrt{c_f})\big) \tag{89.9}$$ EqReyn

relie le coefficient de friction c_f au nombre de Reynolds Re. Pour $Re = 10^4$, représenter la fonction qui, à c_f, associe le second membre de l'équation (89.9), c_f allant de 0 à 0.05. corrSC_serie2-0005

Correction of the exercise 28

```
function y = f(fric)
  Re = 10^4
  y = sqrt(fric).*(0.4 +1.74.*log(Re.*sqrt(fric) ) )
end


debut = 0
fin = 0.05
pas = (fin-debut)/1000
echantillon = debut+pas:pas:fin
  % On ne commence pas à zéro parce que le log n'y a pas de sens.

valeurs = f(echantillon)

```

```
plot(echantillon,valeurs)
print -dps exo2-5.ps
```

tex/matlab/SC_exo_2–5.m

89.1.3 Polynômes et approximations au sens des moindres carrés

Exercise 29 exoSC_serie3-0001

Trouver les racines du polynôme $z^8 + z + 1$ et les représenter dans le plan complexe. corrSC_serie3-0001

Correction of the exercise 29

Un polynôme est donné par un vecteur qui contient les coefficients. Ainsi, le polynôme x^2-2x+3 sera représenté par $p = [1, -2, 3]$. Attention : dans la tête de Matlab, ce p reste un vecteur. Lui demander de faire `plot(p)` ne va pas du tout lui faire tracer le graphe du polynôme.

En ce qui concerne le tracé, il se fait que Matlab place automatiquement les nombres complexes dans le plan complexe. Ainsi, si z est un nombre complexe, `plot(z)` affichera le point du plan qui correspond à z.

```
p = [1,0,0,0,0,0,0,1,1]
racines = roots(p)

plot(racines,'*') % Le * est pour que les points soient des ↩
   petites étoiles.
print -dps exo31.ps
```

tex/matlab/SC_exo_3–1.m

Exercise 30 exoSC_serie3-0002

Le tableau ci-dessous donne la force électromotrice E (en volts) d'une pile, en fonction de la température absolue T en K.

T	290	300	310	320	330
E	1.15053	1.14950	1.14788	1.14656	1.14527

On estime que les valeurs de la force électromotrice peuvent être approchées par les valeurs d'un polynôme du troisième degré

$$E = aT^3 + bT^2 + cT + d. \tag{89.10}$$

(1) Déterminer les constantes a, b, c et d.

(2) Représenter, dans un même diagramme, le graphe du polynôme et les points correspondant aux données.

(3) Estimer la valeur de E pour $T = 316\,\mathrm{K}$.

corrSC_serie3-0002

Correction of the exercise 30

```
% Ici, nous avons des données expérimentales, et puis une courbe ↩
   théorique
% La stratégie sera la suivante :
% 1. Nous mettons les données expérimentales dans les vecteurs x ↩
   et y.
% 2. Nous calculerons la courbe théorique avec un polyfit.
% 3. Pour tracer la courbe théorique, nous allons procéder comme ↩
   d'habitude.
%   Nous commencerons par créer un vecteur d'abcisse X
```

```
%   et puis nous calculerons le vecteur d'ordonnées correspondant ↩
    Y.

x = [290,300,310,320,330]
y = [1.15053,1.14950,1.1478,1.14656,1.14527]

% Pour trouver le polynôme de degré 3 qui passe le mieux par les ↩
    points
% donnés par les vecteurs x et y, il faut utiliser la commande ↩
    polyfit.
p = polyfit(x,y,3)
% 5.1667e-08   -4.8093e-05    1.4770e-02   -3.4821e-01

% Juste pour s'amuser à voir à quoi ressemble le polynôme, ↩
    écrivons-le :
polyout(p,"T")

X = 250:350
Y = polyval(p,X)
% polyval est la commande pour évaluer un polynôme en un point.
% Ici, on l'évalue en tous les points d'abscisse qu'on veut tracer↩
    .

plot(X,Y,':',x,y,'o')
print -dps exo32.ps
% Noter que je trace de 250 à 350 de façon très arbitraire.
% Rien dans les données expérimentales ne montre la croissance de ↩
    la
% fonction entre 250 et 280, ni celle entre 340 et 350.
% D'un point de vue scientifique, la méfiance est de rigueur ↩
    lorsqu'on
% extrapole des données en-dehors du domaine des expériences.

% Pour évaluer la valeur de p au point 316 :
polyval(p,316)  %1.1470
```

tex/matlab/SC_exo_3–2.m

Exercise 31 exoSC_serie3-0003

Le tableau ci-dessous donne les valeurs de l'énergie E consommée par différents animaux dans la course, en liaison avec la masse m de ces animaux.

Animal	souris	écureil	rat	chien (petit)	chien (gros)	mouton	cheval
Masse en gramme m	21	236	384	2.6×10^3	1.8×10^4	3.9×10^4	5.8×10^5
Énergie (cal/g/km)	13	3.7	4.4	1.7	0.92	0.58	0.15

Dans le plan $(\ln(E), \ln(m))$, trouver une droite qui passe approximativement par les points donnés ; quelle est la pente de cette droite ? Représenter, dans un même diagramme, cette droite et les points correspondants aux données. corrSC_serie3-0003

Correction of the exercise 31

Dans cet exercice, la subtilité est d'utiliser les logarithmes des données expérimentales, et non les données elles-mêmes. Nous commençons donc par créer les vecteurs `lnx` et `lny` qui contiennent les logarithmes des données expérimentales.

```
x = [21,236,384,2.6*10^3,1.8*10^4,3.9*10^4,5.8*10^5]
y = [13,3.7,4.4,1.7,0.92,0.58,0.15]

lnx = log(x)
lny = log(y)

% Une droite, est un polynôme de degré 1.
droite = polyfit(lnx,lny,1)
% -0.42109   3.84651

abcisses = [2,15]
ordonnees = polyval(droite,abcisses)

plot(abcisses,ordonnees,':',lnx,lny,'o')
print -dps exo33.ps
```

tex/matlab/SC_exo_3–3.m

Exercise 32 exoSC_serie3-0004

Le tableau ci-dessous donne la conductivité molaire Λ (en $\Omega^{-1}\text{cm}^2/\text{mol}$) de l'acide chlorhydrique, en fonction de la concentration c (en mol/dm^3),

c	0.0005	0.001	0.005	0.01	0.02	0.05
Λ	422.74	421.36	415.80	412.24	407.24	399.09

On considère que la relation entre Λ et c est donnée approximativement par une formule du type

$$\Lambda = a_0 + a_1 c^{1/2}. \qquad \text{EqLLamazau} \quad (89.11)$$

Trouver les coefficients a_0 et a_1 à partir des données et représenter dans un même diagramme les valeurs donnes et le graphe de la fonction définie par (89.11).

Indication : Le second membre de (89.11) peut être vu comme un polynôme du premier degré en $c^{1/2}$. corrSC_serie3-0004

Correction of the exercise 32

```
x = [0.0005,0.001,0.005,0.01,0.02,0.05]
y = [422.74,421.36,415.80,412.00,407.24,399.09]

p = polyfit(sqrt(x),y,1) %-118.00   424.64
polyout(p,"x")

abcisses = 0.0005:0.0001:0.05
ordonnees = p(1)*sqrt(abcisses)+p(2)

plot(abcisses,ordonnees,':',x,y,'o')
print -dps exo34.ps
```

tex/matlab/SC_exo_3–4.m

89.1.4 Intégration numérique et résolution d'équations différentielles

Exercise 33 exoSC_serie4-0001

La chaleur spécifique C_v (en J/K/mol) d'un solide monoatomique varie en fonction de la température absolue T suivant la loi

$$C_V = \frac{9R}{x_m^3}\int_0^{x_m} \frac{e^x x^4}{(e^x-1)^2}dx \qquad (89.12)$$

où $R = 8.314\text{J/K/mol}$ et $x_m = \Theta_D/T$, Θ_D étant la température de Debye, qui dépend du solide considéré. Déterminer la chaleur spécifique pour $T = 300\text{K}$, dans le cas du cuivre, pour lequel $\Theta_D = 313\text{K}$. corrSC_serie4-0001

Correction of the exercise 33

```
function y=f(x)
  y = (exp(x)*x.^4)/(exp(x)-1).^2;
endfunction

% La fonction suivante donne la valeur de l'intégrale
% de la fonction demandée entre 0 et xm.
function y=Integrale(xm)
y = quad('f',0,xm);
endfunction

R = 8.314;
xm = 313/300;

reponse = (9*R/xm^3)*Integrale(xm)  % 23.636
```

tex/matlab/SC_exo_4–1.m

Notez que les fonctions qui sont destinées à être intégrées doivent accepter des entrées vectorielles. D'où le fait qu'il ne faille pas oublier de mettre des points un peu partout.

Exercise 34 exoSC_serie4-0002

La fonction

$$f(x) = e^{-x^2} \int_0^x e^{t^2} dt \tag{89.13}$$

est appelée *intégrale de Dawson*.

(1) Calculer $f(1)$

(2) Déterminer une solution de l'équation $f(x) = 0.5$. (Indication : cette solution est proche de 0.7)

corrSC_serie4-0002

Correction of the exercise 34

Le point à comprendre est que `fzero` ne permet que de résoudre $f(x) = 0$. Or ici nous avons besoin de $f(x) = 0.5$. C'est pour cela que nous définissons la fonction intermédiaire `Dawsonbis` qui vaut $f(x) - 0.5$.

```
function y=Integrande(t)
y = exp(t.^2);
end

% La fonction qui à x fait correspondre l'intégrale de f entre 0 ↩
   et x
function y=Integrale(x)
y=quad('Integrande',0,x);
end

% La fonction Dawson de l'énoncé
function y=Dawson(x)
y=exp(-x.^2)*Integrale(x);
```

```
13 end
14
15 reponseA = Dawson(1)  % 0.53808
16
17 function y=Dawsonbis(x)
18 y=Dawson(x)-0.5;
19 end
20
21 reponseB = fzero('Dawsonbis',0.7) % 0.66607
```

tex/matlab/SC_exo_4–2.m

Exercise 35 exoSC_serie4-0003

On considère le problème

$$\begin{cases} y' = 1 + y^2 & (89.14a) \\ y(0) = 0. & (89.14b) \end{cases}$$

(1) Calculer la solution sur l'intervalle $[0, 1.5]$.

(2) La solution exacte de ce problème est $y(x) = \tan(x)$. Représenter dans un même diagramme la solution exacte et la solution approchée calculée plus haut.

corrSC_serie4-0003

Correction of the exercise 35

Nous suivons le plan suivant

— Nous commençons par définir la fonction qui donne le problème de Cauchy, c'est-à-dire $f(x, y) = 1 + y^2$. Notez qu'il faut bien définir une fonction de deux variables x et y, même si on n'en utilise une seule.

— Ensuite, nous résolvons le système en mettant la solution dans les vecteurs `x` et `y`.

Notez que j'ai l'impression que pour ode45, la syntaxe est notablement différente entre Octave et Matlab.

— La commande `print -dps exo43.ps` sert à enregistrer le graphique dans le fichier `exo43.ps`.

```
1 function retour=Cauchy(x,y)
2 retour = 1+y.^2;
3 end
4
5 [x,y] = ode45(@Cauchy,[0 1.5],0)
6
7 X = 0:0.001:1.5;
8 Y = tan(X);
9
10 plot(x,y,'o',X,Y)
11 print -dps exo43.ps
```

tex/matlab/SC_exo_4–3.m

Exercise 36 exoSC_serie4-0004

On considère le problème

$$\begin{cases} y' = x^2 + y^2/4 & (89.15a) \\ y(0) = 0.2. & (89.15b) \end{cases}$$

(1) Calculer la solution sur l'intervalle $[0, 1.5,]$.

(2) Représenter, dans un même diagramme, le graphe de la solution y et celui de sa dérivée y'.

corrSC_serie4-0004

Correction of the exercise 36

Par définition, si on connait un point (x_0, y_0) du graphe de la solution au problème de Cauchy, la valeur de la dérivée en ce point est donnée par $f'(x_0) = x_0^2 + y_0^2/4$. Nous pouvons donc facilement calculer f' sur les abscisses `X` où Matlab a fournit la solution parce que nous y connaissons le `y` correspondant.

De plus, la formule qui donne $f'(x_0)$ en fonction de x_0 et y_0 est précisément celle qui définit le problème de Cauchy. Pas besoin de la retaper.

```
function retour=Cauchy(x,y)
retour = x.^2+(y.^2)/4;
end

[X,Y] = ode45(@Cauchy,[0 1.5],0.2);

z = Cauchy(X,Y)

plot(X,Y,X,z,'o')
print -dps exo44.ps
```

tex/matlab/SC_exo_4–4.m

89.1.5 Exercices variés

Exercise 37 exoSC_serie5-0001

On considère la suite $\{x_n\}$ définie par

$$\begin{aligned} x_1 &= 0 \\ x_n &= \sqrt{2 + x_{n-1}}. \end{aligned} \tag{89.16}$$

(1) Écrire un programme permettant de calculer x_{12}.

(2) Écrire une fonction permettant de calculer x_n en fonction de n.

corrSC_serie5-0001

Correction of the exercise 37

```
% Cette fonction donne x_{n+1} en fonction de x_n.
function y = recurrence(x)
  y=sqrt(2+x);
end

n = 12
x = 0;

% On va appliquer n fois la récurrence
for i = 1:n
  x = recurrence(x);
endfor

x

% Pour faire une fonction qui calcule le terme n, il suffit de ↩
   faire
```

```
% une fonction qui contient n fois la récurrence.
function y = xn(n)
  x = 0;
  for i = 1:n
    x = recurrence(x);
  endfor
  y = x;
end

% Pour calculer le terme numéro 100 de la suite :
TermeCent = xn(100)
```

tex/matlab/SC_exo_5–1.m

Exercise 38 exoSC_serie5-0002

Écrire une programme permettant de construire à partir des vecteurs donnés

$$\begin{aligned} x &= (x_1, x_2, \dots, x_n) \\ y &= (y_1, y_2, \dots, y_n) \\ z &= (z_1, z_2, \dots, z_n) \end{aligned} \tag{89.17}$$

la matrice

$$\begin{pmatrix} x_1 & z_1 & 0 & \cdots & 0 \\ y_1 & x_2 & z_2 & \ddots & \vdots \\ 0 & y_2 & x_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & z_{n-1} \\ 0 & \cdots & 0 & y_{n-1} & x_n \end{pmatrix}. \tag{89.18}$$

Considérons maintenant la matrice

$$\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}. \tag{89.19}$$

de genre $n \times n$. En utilisant le programme qui précède, écrivez un programme qui permettant de représenter, en fonction de n variant de 1 à 20, le déterminant de la matrice A. corrSC_serie5-0002

Correction of the exercise 38

<++>

Exercise 39 exoSC_serie5-0003

Reprenons notre intégrale de Dawson

$$f(x) = e^{-x^2} \int_0^x e^{t^2} dt. \tag{89.20}$$

(1) Calculer $f(1)$;

(2) Représenter le graphe de f sur l'intervalle $[0, 0.5]$.

corrSC_serie5-0003

Correction of the exercise 39

<++>

Exercise 40 exoSC_serie5-0004

Étant donnés deux vecteurs $x = (x_1, x_2, \dots, x_n)$ et $y = (y_1, y_2, \dots, y_n)$, écrivez un programme donnant le vecteur $z = (z_1, z_2, \dots, z_n)$ dont la composante numéro i est celui des deux nombres x_i, y_i qui est le plus grand en valeur absolue.

corrSC_serie5-0004

Correction of the exercise 40

```
% La fonction suivante retourne le plus grand en valeur absolue de a et b.
function y=maxvalabs(a,b)
  if abs(a)>abs(b)
    y = a;
  else
    y = b;
  end
end

% deux vecteurs de tests.
x = [1,4,8,-4,-4];
y = [-2,3,7,-5,5];

% On crée le vecteur z qui a la m\ême longueur que x.
% Peu importe ce qu'il y a dedans parce qu'on va le redéfinir juste après.
z = 1:length(x)

for i = 1:length(x)
  z(i) = maxvalabs(x(i),y(i));
endfor

z
```

tex/matlab/SC_exo_5–4.m

Exercise 41 exoSC_serie5-0005

Représenter le graphe de la fonction

$$\begin{aligned} f\colon [1,10] &\to \mathbb{R} \\ x &\mapsto \begin{cases} (\ln(x)+2)^2 & \text{si } \ln(x)-x+2 \geqslant 0 \\ x^2-4x & \text{si } \ln(x)-x+2<0. \end{cases} \end{aligned} \tag{89.21}$$

corrSC_serie5-0005

Correction of the exercise 41

Quelques éléments de technique :

— Pour définir la fonction, vu que le test d'égalité ne semble pas exister dans Matlab, il ne faut pas dire que la fonction vaut x^2-4x si $\ln(x)-x+2<0$ (strict) et $\big(\ln(x)+2\big)^2$ sinon.

— Pour tracer, la procédure habituelle serait de faire $Y = f(X)$ après avoir définit un vecteur d'abscisses X. Hélas, f ne s'applique pas bien à un vecteur (à cause du fait que x arrive dans un `if`). Il faut donc faire à la main le passage de composante à composante. C'est à cela que sert la boucle `for`.

```
function y=f(x)
```

```
  if log(x)-x+2 < 0
    y = x.^2-4*x;
  else
    y = (log(x)+2).^2-4*x;
  endif
endfunction

X = 1:0.1:10;

Y = 1:length(X);
for i=1:length(X)
  Y(i)=f(X(i));
endfor

plot(X,Y)
print -dps exo55.ps
```

tex/matlab/SC_exo_5–5.m

Exercise 42 exoSC_serie5-0006

Rechercher la solution sur l'intervalle $[0, 6,]$ du problème

$$\begin{cases} y' = f(x, y) & \text{(89.22a)} \\ y(0) = -1 & \text{(89.22b)} \end{cases}$$

si

$$f(x, y) = \begin{cases} -y^2 + x & \text{si } y \geqslant 0 \\ -y + x^2 & \text{si } y < 0 \end{cases} \tag{89.23}$$

corrSC_serie5-0006

Correction of the exercise 42

<++>

89.2 Anciens tests et examens

89.2.0.1 BIR1200 en 2009

Exercise 43 exoMatlab0001

Représenter le graphe de la fonction

$$\begin{aligned} f\colon [-3, 3] &\to \mathbb{R} \\ x &\mapsto (1 + x)e^{-x^2+3x\cos(x)} - (1 + x^4)^2 \sin(x). \end{aligned} \tag{89.24}$$

corrMatlab0001

Correction of the exercise 43

```
function y=f(x)
y=(1+x).*exp(-x.^2+2*x.*cos(x))-(1+x.^4).^2.*sin(x)
end

X = -3:0.1:3
Y = f(X)
plot(X,Y)


```

```
print -dps exo0001.ps
```

tex/matlab/exo0001.m

Exercise 44 exoMatlab0002

Calculez

$$\sum_{k=1}^{100} \frac{1}{\sin(k)}.$$

Note : Étymologiquement, c'est cette somme qui a donné son nom aux assiettes gyros spéciales avec frites. corrMatlab0002

Correction of the exercise 44

```
u = 1:100
v = 1./sin(u)

reponse = sum(v)     %-60.588
```

tex/matlab/exo0002.m

Exercise 45 exoMatlab0003

Le roi, pour remercier l'inventeur de l'assiette gyros avec frites de sa belle invention, lui a dit qu'il mettrait 1 euro sur la première frite, 1/4 euros sur la seconde, *et cætera* jusqu'à la centième frite. Calculez

$$\sum_{k=1}^{100} \frac{1}{k^2}$$

pour savoir combien tout cela va coûter au trésor royal. corrMatlab0003

Correction of the exercise 45

```
u = 1:100
v = 1./u.^2

reponse = sum(v)     % 1.635
```

tex/matlab/exo0003.m

Exercise 46 exoMatlab0004

Calculez

$$\sum_{k=1}^{100} 2^k.$$

corrMatlab0004

Correction of the exercise 46

```
u = 1:100
v = 2.^u

reponse = sum(v)     % 2.5353e+30
```

tex/matlab/exo0004.m

Exercise 47 exoMatlab0005

Tracer le graphe du polynôme $P(x) = x^7 - 5x^2 + 2$ entre -2 et 2. corrMatlab0005

Correction of the exercise 47

```
p = [1 0 0 0 0 -5 0 2]

X = -2:0.1:2
Y = polyval(p,X)

plot(X,Y)
print -dps exo0005.ps
```

tex/matlab/exo0005.m

Exercise 48 exoMatlab0006

Pour quelle valeur proche de zéro est-ce que $\cos(x) = x$? corrMatlab0006

Correction of the exercise 48

```
function y=f(x)
y=cos(x)-x
endfunction

fzero('f',0)  % 0.73909
```

tex/matlab/exo0006.m

Exercise 49 exoMatlab0007

Pour quelle valeur proche de zéro est-ce que $\ln(x) = \sin(x)$? corrMatlab0007

Correction of the exercise 49

<++>

Exercise 50 exoMatlab0008

Lorsqu'on achète une assiette gyros spéciale avec frites, on observe les prix suivants, en fonction du nombre de frites :

(1) avec 20 frites, 5.2 euros ;

(2) avec 50 frites, 7.5 euros ;

(3) avec 73 frites, 9.6 euros ;

(4) avec 100 frites, 14 euros.

En supposant une progression linéaire du prix en fonction du nombre de frites, estimez combien coûterait une assiette gyros avec 307 frites. corrMatlab0008

Correction of the exercise 50

<++>

Exercise 51 exoMatlab0009

Représentez dans le plan complexe les racines du polynôme

$$z^9 + z^8 + z^7 + \cdots + z^2 + z + 1 \tag{89.25}$$

en les marquant d'un rond.

Représentez ensuite le polynôme pour z allant de 0 à 3. corrMatlab0009

Correction of the exercise 51

<++>

Exercise 52 exoMatlab0010

Thorgal, XIII, Kid Paddle et Gaston se rendent dans un Kebab. Thorgal prend 8 assiettes gyros spéciales, 10 frites et 6 boissons. Il paie 43 euros. XIII mange 2 frites et deux boissons, mais

ne prend pas d'assiette gyros spéciale, et il paie 9 euros. Kid Paddle se contente d'une assiette gyros spéciale et d'une boisson et paie 4.5 euros.

Gaston voudrait prendre une assiette gyros spéciale avec frites. Combien devra-t-il payer ?

Indice : si A, F et B désignent les prix des assiettes, des frites et des boissons, il faut résoudre le système

$$\begin{cases} 8A + 6B + 10F = 43 \\ 2B + 2F = 9 \\ A + B = 4.5 \end{cases}$$

et en déduire la valeur de $A + F$. corrMatlab0010

Correction of the exercise 52

<++>

Exercise 53 exoMatlab0011

Représentez sur le même diagramme les graphes des fonctions

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

et

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

sur l'intervalle $[-2, 2]$. corrMatlab0011

Correction of the exercise 53

<++>

Exercise 54 exoMatlab0012

Donnez une valeur approximative de

$$\int_0^{\pi/2} \sin(7x)\cos(5x)dx \tag{89.27}$$

corrMatlab0012

Correction of the exercise 54

<++>

Exercise 55 exoMatlab0013

Considérez la fonction

$$f(x) = \int_0^x \cos\big(\sin(t)\big)dt. \tag{89.28}$$

(1) Trouvez une valeur de x proche de 0.3 telle que $f(x) = \frac{1}{4}$. Nous nommons x_0 cette valeur.

(2) Tracez le graphe de f sur un intervalle raisonnable autour de x_0, et marquez le point $(x_0, \frac{1}{4})$ par un petit cercle.

corrMatlab0013

Correction of the exercise 55

```
function y=intergrande(t)
y=cos(sin(t));
end

function y=f(x)
y=quad('intergrande',0,x)-1/4;
end

```

```
reponse = fzero('f',0.3)     % 0.25265

X = 0:0.1:5;
Y = f(X);
plot(X,Y,'o')
print -dps exo0013.ps
```

tex/matlab/3nov14h00–q1.m

Exercise 56 exoMatlab0014

Un restaurateur prétend que ses ventes d'assiettes gyros spéciales avec frites ont progressé avec le temps selon la courbe donnée par la fonction

$$g(t) = e^{\cos(t)} \sin(t-1)^2 + 2\sqrt{t^3+7t}. \tag{89.29}$$

Représentez cette courbe sur un graphique pour l'intervalle de temps $t \in [0, 10]$. corrMatlab0014

Correction of the exercise 56

```
function y=f(t)
y=exp(cos(t)).*sin(t-1).^2+2*sqrt(t.^3+7*t)
end

X = 0:0.1:10
Y = f(X)
plot(X,Y)
print -dps exo3novQ2.ps
```

tex/matlab/3nov14h00–q2.m

Exercise 57 exoMatlab0015

Résolvez le problème de Cauchy

$$\begin{cases} y' = -(x^2+y^2) & \text{(89.30a)} \\ y(0) = 1. & \text{(89.30b)} \end{cases}$$

Calculez et tracez la solution sur l'intervalle $[0, 2]$. corrMatlab0015

Correction of the exercise 57

```
function z=f(x,y)
z=-(x.^2+y.^2)
end

[x,y] = ode45(@f,[0,2],1)
plot(x,y)

print -dps exo3novQ3.ps
```

tex/matlab/3nov14h00–q3.m

Exercise 58 exoMatlab0016

En relativité, on démontre que si la longueur d'un objet est l_0, alors un observateur en mouvement à la vitesse v mesurera une longueur donnée par

$$l(v) = l_0\sqrt{1-\frac{v^2}{c^2}} \tag{89.31}$$

où $c = 3 \cdot 10^8$ m/s est une constante physique.

Tracez le graphe de la longueur $l(v)$ observée en fonction de v dans le cas d'un objet de taille $l_0 = 1.3$ m, pour v allant de 0 à $3 \cdot 10^8$ m/s. corrMatlab0016

Correction of the exercise 58

```
function y=l(lz,v)
c=3*10^8
y=lz*sqrt(1-(v.^2/c^2))
end
c=3*10^8
lz = 1.3
X = 0:10000:c
Y = l(lz,X)
plot(X,Y)
print -dps exo3novQ4.ps
```

tex/matlab/3nov14h00–q4.m

Exercise 59 exoMatlab0017

La production mondiale de pétrole, en milliers de barils par jours, de la dernière décennie est donnée par

Année	1997	1998	1999	2000	2001	2002	2003	2004
Production	72231	73588	72377	74916	74847	74478	77031	80326

Année	2005	2006	2007
Production	81255	81659	81533

(1) Donnez les valeurs correspondantes, en milliards de barils par an. Nous vous rappelons que le facteur de conversion est de 365 jours par an, et de 10^6 milliers par milliards. Enregistrez le résultat dans le vecteur `consommation`.

(2) Trouvez les meilleurs constantes a et b telles que le vecteur `consommation` soit approximé par la droite

$$aT + b \tag{89.32}$$

où T est l'année.

(3) Tracez, dans un même diagramme, les données et la droite trouvée, et prolonger la droite jusqu'en 2050. Quelle consommation mondiale serait atteinte selon cette prolongation linéaire ?

Questions bonus (à ne faire que si il vous reste du temps, ne comptent pas pour des points) :

(1) Comparez les résultats obtenus avec la réserve globale de pétrole qui reste sous nos pieds en 2009 : environ 1240 milliards de barils. En particulier, calculez la somme

$$\sum_{T=2009}^{2050} aT + b. \tag{89.33}$$

(2) Où peut-on trouver des assiettes gyros spéciales avec frites à Louvain-la-Neuve ? corrMatlab0017

Correction of the exercise 59

```
annee =  [1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007]
cons =  [72231 73588 72377 74916 74847 74478 77031 80326 81255 ↩
   81659 81533]
```

```
3
4  consommation = cons*365/(10^6)
5  p = polyfit(annee,consommation,1)
6
7  long_terme = 1997:2050
8  theorie = polyval(p,long_terme)
9
10 plot(annee,consommation,'o',long_terme,theorie)
11 print -dps petrole.ps
12
13 % Consomation extrapolée année par année :
14 extracons = polyval(p,2009:2050)
15 % Consomation en 2050 :
16 polyval(p,2050)
17
18 % Somme de notre consommation entre 2009 et 2050 :
19 sum(extracons)
```

tex/matlab/exo0017.m

Les assiettes gyros spéciales avec frites se trouvent uniquement à Louvain la Neuve au Coup de théâtre situé sur la place de l'université, à gauche de l'entrée du bâtiment de dons de sang.

Exercise 60 exoMatlab0018

Représentez le graphe de la fonction

$$f(x) = (1 + x^{e^3}) \sin(x^3) + \cos(3x)$$

sur l'intervale $[-3, 2]$. corrMatlab0018

Correction of the exercise 60

<++>

Exercise 61 exoMatlab0019

Représentez le graphe de la fonction

$$f(x) = 37 \tan(x^2) \sin(x^3) + \cos(3x)$$

sur l'intervale $[-3, 2]$. corrMatlab0019

Correction of the exercise 61

<++>

Exercise 62 exoMatlab0020

(1) Représentez dans un même diagramme les fonctions $\sin(x)$ et $1/x$ entre 0.2 et π.

(2) Trouvez les deux premières solutions positives de l'équation

$$\sin(x) = 1/x$$

après avoir repéré approximativement ces solutions à l'aide du diagramme de la sous-question précédente.

(3) Représentez dans le même diagramme les deux fonctions, et les deux racines marquées d'une croix ou d'un rond. corrMatlab0020

Correction of the exercise 62

<++>

Exercise 63 exoMatlab0021

(1) Représentez dans un même diagramme les fonctions $\sin(x^2)$ et $\exp(x)/4$ entre 0 et 2.
(2) Trouvez les deux premières solutions positives de l'équation

$$\sin(x^2) = \exp(x)/4$$

après avoir repéré approximativement ces solutions à l'aide du diagramme de la sous-question précédente.
(3) Représentez dans le même diagramme les deux fonctions, et les deux racines marquées d'une croix ou d'un rond.

corrMatlab0021

Correction of the exercise 63

```
function y=f1(x)
y=sin(x.^2)
endfunction

function y=f2(x)
y=exp(x)/4
endfunction

function y=f(x)
y=f1(x)-f2(x)
endfunction

X = 0:0.05:2
Y1 = f1(X)
Y2 = f2(X)

t1 = 0.7
t2 = 1.3

r1 = fzero('f',t1)
r2 = fzero('f',t2)

r1    %0.74452
r2    %1.3525

plot(X,Y1,X,Y2,r1,f1(r1),'o',r2,f1(r2),'o')
print -dps exo2-1.ps
```

tex/matlab/exo2–1.m

Exercise 64 exoMatlab0022

(1) Construisez une fonction `assiettegyrosspeciale` d'une variable x qui renvoie la valeur du polynôme

$$3x^8 + 6x^5 - 2x^4 + 6x^3 - x + 1$$

au point x (utilisez le vecteur des coefficients du polynôme).
(2) Construisez une fonction `avecfrites` de deux variables a et b qui renvoie comme résultat l'intégrale de a à b de la fonction définie ci-dessus.
(3) Donnez la valeur de la fonction `avecfrites` lorsque $a = 2$ et $b = 5$.

corrMatlab0022

Correction of the exercise 64

<++>

Exercise 65 exoMatlab0023

(1) Calculez le polynôme p, de degré égal à 10, vérifiant les conditions

$$p(x) = (-1)^x \quad \text{pour} \quad x = 0, 1, \ldots, 10.$$

(2) Représentez ce polynôme, en marquant d'une croix les points du graphe correspondant aux données.

(3) Déterminez les racines du polynôme.

corrMatlab0023

Correction of the exercise 65

<++>

Exercise 66 exoMatlab0024

On considère la fonction $f : [-1, 1] \to \mathbb{R} : x \mapsto 1/(1 + 25x^2)$.

(1) Créez une fonction Matlab pour f.

(2) Construisez le vecteur v contenant 11 points équidistants entre -1 et 1, les extrémités -1 et 1 faisant partie de ces points.

(3) Donnez (sous forme du vecteur de ses coefficients) le polynôme d'approximation de degré 11 pour les données $v, f(v)$.

(4) Représentez dans un même diagramme les graphes de la fonction f et de son polynôme d'approximation.

corrMatlab0024

Correction of the exercise 66

<++>

Exercise 67 exoMatlab0025

Le professeur Gyros a observé, pendant 12 jours, la croissance d'une population de bactéries dans une assiette de frites. Malheureusement il a égaré une partie des résultats et voudrait les reconstituer à partir du tableau partiel suivant (on ne se préoccupera pas des unités dans lesquelles est exprimée la taille de la population).

Jour (j)	1	3	6	7	10	12
Population ($P(j)$)	12	16	30	35	63	90

En supposant que la croissance de population obéisse à peu près à une loi du type $P(j) = Ce^{\lambda j}$, trouvez les coefficients C et λ et utilisez ces estimations pour compléter le tableau ci-dessus.

Indication : L'approximation demandée revient à une approximation par un polynôme du premier degré pour $\ln(P(j))$.

corrMatlab0025

Correction of the exercise 67

<++>

Exercise 68 exoMatlab0026

Définissez une fonction qui, à un entier n, associe la matrice A_n suivante, de genre $(n + 1) \times (n + 1)$:

$$A_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \frac{1}{(n)^3} & 0 & \frac{1}{(2)^4} & 0 & & & \vdots \\ 0 & \frac{1}{(n-1)^3} & 0 & \frac{1}{(3)^4} & & & \vdots \\ 0 & 0 & \frac{1}{(n-2)^3} & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \ddots & \frac{1}{(n-1)^4} & 0 \\ \vdots & & & & \frac{1}{(2)^3} & 0 & \frac{1}{(n)^4} \\ 0 & \cdots & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

Indication : Utilisez la commande `diag(v,k)`*. Si vous n'arrivez pas à créer la fonction, construisez simplement la matrice pour* $n = 7$ *(elle est alors de genre* 8×8*).*

corrMatlab0026

Correction of the exercise 68

<++>

Exercise 69 exoMatlab0027

Définissez une fonction qui, à un entier n, associe la matrice A_n suivante, de genre $(n+1)\times(n+1)$:

$$A_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \frac{1}{(n)^4} & 0 & \frac{1}{(2)^3} & 0 & & & \vdots \\ 0 & \frac{1}{(n-1)^4} & 0 & \frac{1}{(3)^3} & & & \vdots \\ 0 & 0 & \frac{1}{(n-2)^4} & 0 & \ddots & & 0 \\ \vdots & & & \ddots & \ddots & \frac{1}{(n-1)^3} & 0 \\ \vdots & & & & \frac{1}{(2)^4} & 0 & \frac{1}{(n)^3} \\ 0 & \cdots & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

Indication : Utilisez la commande `diag(v,k)`*. Si vous n'arrivez pas à créer la fonction, construisez simplement la matrice pour $n = 7$ (elle est alors de genre 8×8).* corrMatlab0027

Correction of the exercise 69

<++>

Exercise 70 exoMatlab0028

Un restaurateur aimerait connaitre *le nombre de cuivre*, c'est-à-dire l'épaisseur idéale d'une tranche de viande pour ses gyros. Pour ce faire, il a effectué des tests de découpe à différentes épaisseurs et les a fait gouter à un échantillon de 3 personnes représentatives de la société belge. Après avoir obtenu ses données et effectué de long calculs, il est arrivé au fait que le nombre de cuivre devait être égal au déterminant de la matrice

$$A = (v^t.v)^2 + 42(v^t.v) - 2I$$

où le vecteur v contient les épaisseurs allant de 0.7 à 8.76 millimètres en exactement 13 valeurs extrémités comprises, et où I est la matrice identité de genre 13×13.

Bonne âme que vous êtes, vous décidez de l'aider, et vous

(1) construisez en une opération le vecteur v ;

(2) construisez la matrice A en un tournemain ;

(3) donnez fièrement la valeur du nombre de cuivre [2].

corrMatlab0028

Correction of the exercise 70

<++>

Exercise 71 exoMatlab0029

(1) Construisez en une opération le vecteur v dont les 5 composantes sont équidistantes et comprises dans l'intervalle $[\pi, 2\pi]$, π et 2π étant les première et dernière composantes de v.

(2) Construisez une matrice $A = (a_{ij})$ de genre 5×5, définie par

$$A = (v^t.v)^3 + 4I$$

où I est la matrice identité de genre 5×5.

corrMatlab0029

Correction of the exercise 71

<++>

Exercise 72 exoMatlab0030

2. et vous rendez compte qu'il s'est royalement planté dans ses calculs et que vous ne connaitrez jamais la valeur du nombre de cuivre.

(1) Construisez la matrice A donnée par

$$A = v^t.v$$

où $v = (10, 9, 8, 7, 6, 5, 4, 3, 2, 1)$.

(2) Construisez le vecteur u défini par

$$u = w.A^3 + 4v$$

où I est la matrice identité de genre 10×10 et w est le vecteur de même taille que v dont les composantes sont le carré des composantes correspondantes de v.

(3) Remplacez la troisième composante de u par le nombre 7.

corrMatlab0030

Correction of the exercise 72

<++>

Exercise 73 exoMatlab0031

Un échantillon contenant trois substances radioactives voit son activité décroître suivant une loi du type

$$f(t) = ae^{-\lambda_1 t} + be^{-\lambda_2 t} + ce^{-\lambda_3 t}.$$

On suppose que $\lambda_1 = 1.23$, $\lambda_2 = 0.26$, $\lambda_3 = 0.1$, le temps t étant exprimé en jours.

(1) En résolvant un système linéaire de trois équations à trois inconnues, calculez a, b et c sachant que

$$f(0) = 100, f(2) = 62, f(6) = 32.$$

(2) Avec les valeurs de a, b, c obtenues ci-dessus, définissez la fonction f dans un fichier et utilisez cette définition pour représenter le graphe de f sur l'intervalle $[0, 8]$.

(3) Trouvez $\tau \in [0, 6]$ tel que $f(\tau) = 50$.

corrMatlab0031

Correction of the exercise 73

<++>

Exercise 74 exoMatlab0032

Dans le jargon du métier, le système suivant s'appelle « l'équation de l'assiette gyros spéciale avec frites ».

$$\begin{aligned} 3x + 7y &= 3 \\ 6x + 3y + 2z &= -2 \\ 5x + 2z &= 9. \end{aligned}$$

Pour votre culture générale, et pour réussir l'examen, résolvez-le.

corrMatlab0032

Correction of the exercise 74

```
1
2 A = [3,7,0,5;6,3,2,-2;3,0,1,4;0,6,5,-7]
3 u= [3,-2,9,1]
4 x = inv(A)*u'
5
6 %   -0.58003
7 %   -0.61881
8 %    3.48267
9 %    1.81436
```

tex/matlab/exo3novQ6.m

Exercise 75 exoMatlab0033

Résolvez le système linéaire suivant :

$$\begin{aligned} 3x + 7y + 5t &= 3 \\ 6x + 3y + 2z - 2t &= -2 \\ 3x + 1z + 4t &= 9 \\ 6y + 5z - 7t &= 1. \end{aligned}$$

corrMatlab0033

Correction of the exercise 75

<++>

Exercise 76 exoMatlab0034

(1) Construisez une fonction qui prend comme variables deux vecteurs lignes v et w de même longueur, et qui en renvoie le produit scalaire.

(2) Calculez le produit scalaire du vecteur $v = (\sin(1), \sin(2), \ldots, \sin(100))$ par le vecteur $w = (1/1, 1/2, \ldots, 1/100)$.

corrMatlab0034

Correction of the exercise 76

<++>

Exercise 77 exoMatlab0035

(1) Construisez une fonction qui prend comme variables deux vecteurs v et w à trois composantes, et qui en renvoie le produit vectoriel

$$v \times w = (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_3w_2).$$

(2) Construisez une fonction qui à un réel x associe la norme du produit vectoriel des vecteurs $v = (1, 2, x)$ et $w = (4, 5, 6)$.

(3) Donnez la valeur de cette fonction au point $x = 2$.

corrMatlab0035

Correction of the exercise 77

<++>

Exercise 78 exoMatlab0036

(1) Construisez la fonction

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x^2 & \text{si } x < 0, \\ x^3 & \text{si } x \geqslant 0, \end{cases}$$

de manière qu'elle puisse prendre comme argument un vecteur x (utilisez `for` et `if`).

(2) Représentez la fonction f sur l'intervalle $[-3, 3]$.

corrMatlab0036

Correction of the exercise 78

<++>

Exercise 79 exoMatlab0037

Entourez la ou les syntaxes correctes pour

(1) résoudre le système linéaire $Ax = b$:

`Ax=b,` `A*x=b,` `x=b/A,` `b=A\x,` `x=b\A,` `x=A\b,` `x=solve(A,b)`

(2) créer un vecteur v à 10 composantes équidistantes entre a et b ($a, b \in \mathbb{R}$, $a < b$) :

`v=a:10:b,` `v=a:b:10,` `v(10)=a:b,` `v=a:(b-a)/10:b,` `v=a:(b-a)/9:b`

(3) constuire un vecteur contenant les composantes 3 à 5 d'un vecteur v à 7 composantes :

`v(3,4,5)`, `[v(3):v(5)]`, `[v(3) v(4) v(5)]`, `v(3:5)`, `v[3:5]`, `v(3,5)`

corrMatlab0037

Correction of the exercise 79

<++>

Exercise 80 exoMatlab0038

La différence entre le nombre de frites dans une assiette gyros spéciale et le nombre de frites hors d'une telle assiette dans le monde est donnée en fonction du temps (en années du calendrier courant) par le polynôme

$$-7\,815\,222\,954\,x + 11\,815\,611\,x^2 - 5\,954\,x^3 + x^4. \tag{89.34}$$

(1) Trouvez les dates auxquelles il y a eu, dans le monde, autant de frites dans et hors d'une assiette gyros spéciale.

(2) Déduisez-en que l'an 0 fut une année très spéciale.

corrMatlab0038

Correction of the exercise 80

<++>

Exercise 81 exoMatlab0039

Calculez la somme

$$\sum_{k=1}^{1000} \sin(\exp((k-1)^2).$$

corrMatlab0039

Correction of the exercise 81

<++>

Exercise 82 exoMatlab0040

Pour quelle valeur proche de zéro est-ce que $\ln(x) = \sin(x)$?

corrMatlab0040

Correction of the exercise 82

<++>

Exercise 83 exoLCexoMatlab0001

Résolvez le problème de Cauchy

$$\begin{cases} y' = -3(y^2+1) \\ y(\frac{\pi}{4}) = 5 \end{cases}$$

entre $\frac{\pi}{4}$ et $\frac{3\pi}{4}$.

La solution exacte est donnée par $y(x) = 3\cot(x) + 2$. Représentez dans un même diagramme cette solution exacte et l'approximation trouvée.

corrLCexoMatlab0001

Correction of the exercise 83

<++>

Exercise 84 exoLCexoMatlab0002

Si vous aimez les assiettes gyros spéciales avec frites, résolvez le système d'équations suivant :

$$\begin{cases} x + y + z = 0 \\ x = 3y \\ 5z = 8 + x. \end{cases}$$

Si vous ne les aimez pas, résolvez-le également.

Dans les deux cas, donnez le déterminant de la matrice du système et définissez une variable `gyros` qui contient la valeur de y après résolution.

corrLCexoMatlab0002

Correction of the exercise 84

<++>

Exercise 85 exoLCexoMatlab0003

Calculez la profondeur en mètres à laquelle il faut enterrer une assiette gyros spéciale avec frites pour ne plus sentir ses douces effluves. Autrement dit, calculez

$$\sum_{n=1}^{100} \big(\cos(n) + \sin(n)\big).$$

Note purement informative : C'est un nombre négatif. corrLCexoMatlab0003

Correction of the exercise 85

<++>

Exercise 86 exoLCexoMatlab0004

Tracez le graphe de $f(x) = \frac{\sin(x)}{x}$ pour $x \in [-20, 20]$. corrLCexoMatlab0004

Correction of the exercise 86

<++>

Exercise 87 exoLCexoMatlab0005

Calculer la valeur de

$$\int_{-7}^{7} \big(7x^{11} - 5x^3 + \pi \sin(x)\big).$$

corrLCexoMatlab0005

Correction of the exercise 87

<++>

Exercise 88 exoLCexoMatlab0006

Optimiser des situations commerciales revient régulièrement à résoudre des équations différentielles. En cherchant à optimiser la section des frites pour la préparation des assiettes gyros spéciales avec frites, un ingénieur en machines à découper les frites doit résoudre le problème de Cauchy suivant pour $x \in [0, 2]$:

$$\begin{cases} y' = -\sin(x)y \\ y(0) = e. \end{cases}$$

Enregistrez la solution dans les vecteurs `x` et `y`. La solution exacte du problème est donnée par $y(x) = e^{\cos(x)}$. Tracez cette solution exacte et la solution approchée dans un même diagramme. corrLCexoMatlab0006

Correction of the exercise 88

```
function z=cauchy(x,y)
z= -sin(x)*y
end

function y=exact(x)
y=exp(cos(x))
end

e = exp(1)
[x,y]=ode45(@cauchy,[0,2],e)

X = 0:0.1:2
Y = exact(X)
plot(x,y,'o',X,Y)
```

```
print -dps exo3novQ5.ps
```

tex/matlab/exo3novQ5.m

Exercise 89 exoLCexoMatlab0007

En regardant dans les nuages, Chuck Norris en a observé un qui avait la forme d'une assiette gyros spéciale avec frittes, et un autre qui avait la forme du graphe de

$$f(x) = u\cos(x) + v\sin^2(x)$$

où $u = 9.81$ et $v = \sqrt{\pi/2}$.

En ce qui concerne l'assiette gyros, Chuck Norris l'a obtenue dans un restaurant végérarien sans problèmes. En ce qui concerne le graphe de la fonction, il te demande

(1) de tracer le graphe de cette fonction lorsque $x \in [-5, 5]$,

(2) de trouver pour quelle valeur non loin de $x = 1$, est-ce que $f(x) = 6$.

corrLCexoMatlab0007

Correction of the exercise 89

<++>

Exercise 90 exoLCexoMatlab0008

Un jour, Cauchy était en train de manger une assiette gyros spéciale avec frites quand il s'est demandé quelle fonction vérifiait les conditions

$$\begin{cases} y' = 2xy \\ y(0) = 0. \end{cases}$$

Hélas, Cauchy ne disposait pas d'un ordinateur, et ne connaissait donc que la solution exacte $y(x) = e^{x^2} - e$. Trouvez, grâce à Matlab, une solution approchée au problème de Cauchy et tracez-la entre 0 et 5 sur le même diagramme que la solution exacte.

corrLCexoMatlab0008

Correction of the exercise 90

<++>

Exercise 91 exotestMAT1151-G210001

Dessiner la fonction $y(x) = xe^{-x/a}$ entre -2 et 10 pour la valeur du paramètre $a = 3$. Mettre, sur le même graphe, un petit rond sur le maximum de la fonction.

Pour information, la dérivée de y est

$$y'(x) = -\frac{1}{3}xe^{-x/3} + e^{-x/3}.$$

corrtestMAT1151-G210001

Correction of the exercise 91

```
function y=f(x)
a=3
y=x.*exp(-x/a)
endfunction

function y=fp(x)
y=-(1/3)*x.*exp(-x/3)+exp(-x/3)
endfunction

X=-2:0.1:10
Y=f(X)

xmax=fzero('fp',2)
```

```
ymax=f(xmax)
plot(X,Y,xmax,ymax,'o')

print -dps G21-1.ps
```

tex/matlab/G21–1.m

Exercise 92 exotestMAT1151-G210002

Donner

$$\sum_{i=5}^{1000} \frac{1}{\sin(i)}$$

corrtestMAT1151-G210002

Correction of the exercise 92

```
v=5:1000
w=sin(v)
x=1./w
sum(x)
```

tex/matlab/G21–2.m

Exercise 93 exotestMAT1151-G210003

Abusons des extrapolations...J'ai regardé les deux derniers chiffres du code barre de mes 4 derniers achats. Le résultat est :

Numéro de l'article	1	2	3	4
Fin du code barre	82	96	98	90

En utilisant une approximation polynomiale d'ordre 3, deviner les deux derniers chiffres du prochain article que j'achèterai.

corrtestMAT1151-G210003

Correction of the exercise 93

```
x=[1,2,3,4]
y=[82,96,98,90]

P=polyfit(x,y,3)
polyval(P,5)
```

tex/matlab/G21–3.m

Exercise 94 exotestMAT1151-G220001

Écrire la matrice

$$A = \begin{pmatrix} 4 & 0 & 0 & 1 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

avec un minimum de commandes.

corrtestMAT1151-G220001

Correction of the exercise 94

Exercise 95 exotestMAT1151-G220002

Le travail d'une grue qui soulève un bloc de masse m d'une hauteur h à la surface de la Terre est donné par

$$W(h) = \int_R^{R+h} \frac{GMm}{r^2} dr \tag{89.39}$$

où $R = 6.500.000$ mètres est le rayon de la Terre et le produit GM vaut environ $4.144 \cdot 10^{14}$

Tracer le graphique du travail nécessaire pour monter une masse $m = 1$ à la hauteur h en fonction de h entre $h = 0$ et $h = 10.000$.

Questions bonus (ne comptent pas pour des points)

(1) Quelles sont les unités de W dans le SI ?

(2) Est-ce que vous êtes capables d'interpréter le résultat ?

(3) Essayez de tracer le graphique beaucoup plus loin que $h = 10000$.

corrtestMAT1151-G220002

Correction of the exercise 95

```
function y=f(x)
y=4.144*(10^(14))./x.^2
endfunction

function y=W(x)
R=6500000
y=quad(@f,R,R+x)
endfunction

h=10000

X=0:100:10000
Y=1:length(X)
for i=1:length(X)
  Y(i)=W(X(i))
endfor
plot(X,Y)
print -dps G22-2.ps
```

tex/matlab/G22–2.m

Exercise 96 exotestMAT1151-G220003

Calculer

$$\sum_{k=30}^{90} e^{-k/10}.$$

corrtestMAT1151-G220003

Correction of the exercise 96

```
v=30:90
w=-v./10
x=exp(w)
sum(x)
```

tex/matlab/G22–3.m

Exercise 97 exotestMAT1151-G310001

Il y a un proverbe qui dit que celui qui calculera

$$\sum_{j=4}^{250} \frac{1}{\sqrt{j}} \cos(x^j) \tag{89.40}$$

pour $x = 0$ gagnera des points à son test de Matlab.

Vérifier le proverbe, c'est-à-dire donner la valeur de la somme. corrtestMAT1151-G310001

Correction of the exercise 97

```
u = 4:250
v = 1./sqrt(u)
reponse = sum(v)     % 27.910
```

tex/matlab/G31–1.m

Exercise 98 exotestMAT1151-G310002

Chuck Norris, un p'ti blond aux yeux bleus et une étudiante en Matlab sont dans un bateau. Après avoir percuté un iceberg, le pt'it blond se jette à l'eau en jurant que

$$2x + 3y - z - 20 = 0; \tag{89.41}$$

Chuck Norris termine le trajet à la nage jusqu'au Texas d'où il envoie un SMS à l'étudiante disant que

$$y + z = 4. \tag{89.42}$$

Après avoir trouvé un bateau de rechange, l'étudiante découvre dans sa poche un papier sur lequel il est écrit que

$$-2x + y + z = 0. \tag{89.43}$$

Comment va elle s'y prendre pour calculer les nombres x, y et z en sachant qu'elle n'utilise que Matlab ?

(1) Prouver que le système a une solution unique en donnant le déterminant de la matrice correspondante.

(2) Donner les valeurs de x, y et z qui résolvent le système. corrtestMAT1151-G310002

Correction of the exercise 98

```
A = [2,3,-1;0,1,1;-2,1,1]
u = [20,4,0]

deter = det(A)
v = A\u'
```

tex/matlab/G31–2.m

Exercise 99 exotestMAT1151-G310003

Nous considérons la fonction

$$f(x) = x^2 \int_0^x \sqrt{t}\cos(t)dt. \tag{89.44}$$

(1) Donner $f(1)$,

(2) tracer f pour $x \in [0, 3]$, corrtestMAT1151-G310003

Correction of the exercise 99

```
function retour=phi(t)
retour = sqrt(t)*cos(t);
end
```

```
function retour = f(x)
retour = x.^2*quad('phi',0,x);
end

f(1)  #0.53120
X=0:0.1:3
Y=[]
for i = 1:length(X)
  Y(i)=f(X(i))  ;
end

plot(X,Y)
print -dps G31-3.ps
```

tex/matlab/G31–3.m

Exercise 100 exotestMAT1151-G320001

Cauchy est un petit garçon plein de problèmes. L'un d'entre eux est de trouver la fonction $f\colon [0,4] \to \mathbb{R}$ telle que pour chaque $x \in \mathbb{R}$ on ait

$$f'(x) = \frac{1}{2}f(x) + 2e^{x/2}\cos(2x), \tag{89.45}$$

et telle que

$$f(0) = 0. \tag{89.46}$$

Le petit Augustin Cauchy serait très content que tu lui traces une solution approchée (trouvée par Matlab) sur le même graphique que la solution exacte donnée par

$$y(x) = \sin(2x)e^{x/2}. \tag{89.47}$$

corrtestMAT1151-G320001

Correction of the exercise 100

```
function retour=Cauchy(x,y)
retour = y/2 + 2*exp(x/2).*cos(2*x)
end

function retour=solution(x)
retour = sin(2*x).*exp(x/2)
end

[x,y]=ode45(@Cauchy,[0,4],0)

X = 0:0.1:4
Y=solution(X)
plot(x,y,X,Y)
print -dps G32-1.ps
```

tex/matlab/G32–1.m

Exercise 101 exotestMAT1151-G320002

Dora l'exploratrice voudrais explorer la convergence de la série $\sum_{n=1}^{\infty} 2^{\cos(n)}$. Sa carte lui a dit quelque chose en anglais, mais comme Dora ne comprends pas tellement bien cette langue, elle a (comme toujours) besoin de ton aide. Voici ce que tu dois faire :

(1) Donne $\sum_{n=1}^{100} 2^{\cos(n)}$,

(2) Définis la fonction $s(N) = \sum_{n=1}^{N} 2^{\cos(n)}$ et donne les valeurs de $s(k)$ pour tous les k entiers entre 100 et 200.

(3) Prouve que la série ne converge pas.

Note : Si tu fais seulement le point 1, Dora sera très contente et son cousin Diego fera le reste lui-même. En d'autres termes : **les secondes et troisièmes partie sont facultatives et ne comptent pas pour des points**. corrtestMAT1151-G320002

Correction of the exercise 101

```
u=1:100;
v=2.^(cos(2*u));
sum(v)  % 111.89

function retour=s(N)
u=1:N;
v=2.^(cos(2*u));
retour=sum(v);
end

for i=100:200
s(i)
end
% Le dernier est 223.88
```

tex/matlab/G32–2.m

Pour prouver que la série ne converge pas, montrons que la suite des $2^{\cos(n)}$ ne tend pas vers zéro. Pour cela, considérons la suite x_i des développements décimaux de 2π à i décimales, puis prenons $y_i = x_i \cdot 10^i$. Cela est juste la suite 6, 6.2, 6.28, ...

Vu que le cosinus est une fonction croissante entre 6 et 2π, nous avons $\cos(y_i) \geqslant \cos(6)$, et donc $2^{\cos(y_i)} \geqslant 2^{\cos(6)}$. Cela fournit une sous-suite des $2^{\cos(n)}$ qui ne tend pas vers zéro.

Exercise 102 exotestMAT1151-G320003

Blanche Neige a mesuré ses sept nains, mais le Schtroumpf farceur a effacé certains résultats (en plus de lui avoir fait un paquet cadeau qui lui a détruit la coiffure dont elle était super fière). Voici les résultats qui restent :

Numéro du nain	1	2	3	4	5	6	7
Taille	130	120	?	123	114	?	131

(89.48)

Trouver un polynôme de degré 3 qui correspond le mieux possible aux données restantes, et donner une approximation de la taille des nains numéros 3 et 6. corrtestMAT1151-G320003

Correction of the exercise 102

```
num = [1,2,4,5,7]
tailles = [130,120,123,114,131]

P=polyfit(num,tailles,3)
% P =   1.1404e-01  -2.5063e-02  -5.9937e+00   1.3451e+02

polyval(P,3)  % 119.39
polyval(P,6)  % 122.28
```

tex/matlab/G32–3.m

89.3 Conseils généraux

89.3.1 Écriture d'une fonction

Lorsque vous créez une fonction, veillez aux éléments suivants

(1) La première ligne doit être de la forme

```
function retour=nom_de_fonction(x)
```

La valeur de la fonction sera celle de la variable `retour` lorsqu'on arrivera à la fin de la fonction. De plus, cette fonction doit être sauvée dans le fichier `nom_de_fonction.m`.

Noms des fichiers

(1) Tenez vous en à l'alphabet latin et aux chiffres arabes (et non le contraire).

(2) Le nom de fichier ne peut pas *commencer* par des chiffres

(3) Évitez absolument de mettre des caractères spéciaux qui peuvent avoir un sens mathématique : « (», « ' »,« - »,« + »,...

(4) Évitez de mélanger les majuscules et les minuscules.

(5) Ne mettez pas de points dans le nom de vos fichiers (à part `.m`).

89.3.2 Conception des fonctions

Une fonction doit prendre un nombre (ou une matrice) en entrée et sortir un nombre à la fin. Une fonction doit seulement générer des nombres (ou des matrices). Vous ne devez pas mettre de commandes comme `plot` dans une fonction. Même si cela fonctionne de temps en temps, ce n'est pas une bonne idée. En principe, vous devez concevoir vos fonctions de telle façon qu'elles n'affichent rien.

Les commandes qui doivent *afficher* des résultats se mettent dans un script (blank M-file). Créez un tel script par exercice, même si l'exercice se décompose en sous-questions.

89.3.3 Autres

Si un résultat dépend d'un calcul intermédiaire, ne faites *jamais* le calcul dans la fenêtre principale pour en copier-coller le résultat dans votre script. Faites faire le calcul dans votre script, et enregistrez le résultat dans une variable. Ainsi vous gardez toutes les décimales que Matlab avait calculées sans les afficher. Et cela, même si ladite réponse intermédiaire est un nombre entier.

Quatrième partie

Exercices

Chapitre 90

Mathématique générale pour le starter SVT (Besançon)

Contributeurs :

L'équipe SVT Rédaction de la plupart des exercices.

Laurent Claessens LaTeX, corrections et publication

Carlotta Donadello LaTeX, corrections

Pauline Klein détectrice de coquilles.

Nous avons effectué une certaine « classification » des exercices en y ajoutant des petits symboles.

(1) Le symbole ⊙⊙ marque les exercices à faire à tout prix. Il faut les faire tous.

(2) Le symbole ⊖⊖ signifie que l'exercice ne va en principe pas apporter de nouvelles techniques. Ces exercices sont à faire après avoir fait les exercices de type ⊙⊙, si on a encore des doutes.

(3) Le symbole ⊛⊛ indique que l'exercice sera plus difficile et qu'il vaut mieux l'éviter avant d'avoir bien compris les exercices de type ⊙⊙.

(4) Le symbole ⊘⊘ est appliqué aux exercices qui sont des compléments de la matière, mais qui ne sont pas à faire de façon obligatoire.

90.1 Théorie

90.1.1 Suites

Propufulimite

Proposition 90.1.
Soit $(u_n)_{n\in\mathbb{N}}$, une suite définie par récurrence par

$$u_{n+1} = f(u_n) \tag{90.1}$$

où f est une fonction suffisamment gentille[1]*. Si la suite $(u_n)_{n\in\mathbb{N}}$ converge vers un nombre réel u, alors u est une solution de l'équation $u = f(u)$.*

Lorsque u est une solution de $u = f(u)$ on dit que u est un point fixe de f. Attention : la proposition (90.1) ne garantit pas l'existence de la limite, ni que toute solution de $u = f(u)$ soit une limite. D'ailleurs l'équation $u = f(u)$ peut avoir plusieurs solutions sans que la suite n'ai de limites finis. Cela est le cas lorsque la suite diverge vers $\pm\infty$.

Exemple 90.2.
Soit la suite définie par

$$\begin{cases} u_{n+1} = u_n^2 \\ u_0 = 2. \end{cases} \tag{90.2}$$

1. Nous ne rentrons pas dans les détails des hypothèses exactes. Sachez qu'il faut au moins que la fonction soit continue et bien définie.

Les premiers termes sont $u_0 = 2$, $u_1 = 2^2 = 4$, $u_3 = 4^2 = 16$, etc. Cette suite diverge. Pourtant l'équation $u = u^2$ a des solutions : $u = 0$ et $u = 1$.

Si au lieu d'avoir $u_0 = 2$, on avait eu $u_0 = \frac{1}{2}$, alors nous aurions $u_2 = \frac{1}{4}$, $u_3 = \frac{1}{16}$, etc. Cette suite converge vers 0, qui est bien solution de $u = f(u)$. △

Propsuiteborncv

Proposition 90.3.
Toute suite monotone et bornée converge. En particulier, si une suite est décroissante et bornée vers le bas, alors elle converge. De même, si une suite est croissante et bornée vers le haut, alors elle converge.

90.1.2 Techniques pour majorer et minorer

Dans de nombreux exercices sur les suites, une difficulté est de majorer ou minorer des expressions contenant u_n sachant que u_n est dans un certain intervalle. La technique la plus puissante pour ce faire demande une utilisation intensive des dérivées ; nous n'allons pas parler de cela ici, mais sachez que ça existe.

Exemple 90.4.
Trouver des bornes pour la quantité

$$a_n = u_n\big(1 - \ln(u_n)\big) \tag{90.3}$$

EqExpuumulnuR

sachant que $1 < u_n < e$.

Trouvons une borne supérieure pour l'expression (90.3), c'est-à-dire, trouvons M tel que nous soyons certain d'avoir $a_n < M$. Pour ce faire, nous remplaçons tous les u_n par la valeur qui rend l'expression la plus grande possible. Le premier u_n doit être remplacé par e. Le u_n qui se trouve dans le logarithme doit par contre être remplacé par 1 parce qu'il arrive dans terme qui se soustrait ; pour minorer, il faut soustraire la quantité la plus petite possible. Nous pouvons donc dire que

$$a_n < e(1 - \ln(1)) = e. \tag{90.4}$$

De la même façon, si nous voulons minorer a_n, c'est-à-dire trouver un m tel que $m < a_n$, nous devons remplacer les u_n par les valeurs qui rendent a_n le plus petit possible. Le premier u_n doit être remplacé par 1, tandis que le second doit être remplacé par e (pour soustraire le plus possible). Nous trouvons

$$a_n > 1(1 - \ln(e)) = 0. \tag{90.5}$$

Nous avons donc

$$0 < u_n\big(1 - \ln(u_n)\big) < e \tag{90.6}$$

dès que $1 < u_n < e$.

Notez que cela ne sont pas les bornes optimales. Il est possible (en travaillant plus) de prouver que, sous les mêmes hypothèses, $u_n \leqslant 1$. △

90.2 Rappels : exponentielles et logarithmes

Définition 90.5.
Soit $a > 0$ un nombre réel. La fonction exponentielle en base a est la fonction définie par $x \mapsto a^x$.

Le domaine de $x \mapsto a^x$ est $\mathbb{R}$. La fonction exponentielle est croissante si $a > 1$, décroissante si $0 < a < 1$, constante si $a = 1$. Son image est

— $]0, +\infty[$ si $a \neq 1$,

— $\{1\}$ si $a = 1$.

La fonction exponentielle satisfait les propriétés suivantes pour tous x et y dans $\mathbb{R}$:

— $a^0 = 1$;

— $a^{x+y} = a^x a^y$;
— $a^{x-y} = \dfrac{a^x}{a^y}$;
— $a^{xy} = (a^x)^y$.

Si $a \neq 0$ la fonction exponentielle est strictement monotone sur $\mathbb{R}$ et par conséquence elle admet une fonction réciproque. D' où la définition suivante

Définition 90.6.
Soit $a > 0$, $a \neq 1$. La fonction logarithme de base a, $\log_a$, est définie par la relation $x = a^{\log_a x}$.

La fonction logarithme satisfait les propriétés suivantes pour tous x et y dans $\mathbb{R}$:

— $\log_a 1 = 0$;
— $\log_a x + \log_a y = \log_a(xy)$;
— $\log_a x - \log_a y = \log_a\left(\dfrac{x}{y}\right)$;
— $\log_a(x^y) = y \log_a x$.

En outre, la formule suivante permet de «changer de base» :

$$\log_a(x) = \frac{\log_b x}{\log_b a}. \tag{90.7}$$

Cela est particulièrement important parce que nous permet d'établir la relation entre le logarithme en base a et le logarithme népérien (de base e).

90.3 Exponentielles et logarithmes

Exercise 103 exologarithme-0003

Une marque bien connue de céréales pour commencer la journée en pleine forme contient 0.4 g de sel dans chaque portion de 30 g. Cela représente 30% des repères nutritionnels journaliers.

(1) Combien de grammes de sel faudrait-il manger par jour pour suivre ces fameux repères nutritionnels ?
(2) Combien de grammes de sel contient une boîte de 375 g ?

Note : cet exercice n'a rien à voir avec les logarithmes. corrlogarithme-0003

Correction of the exercise 103

(1) $4/2 \simeq 1.3$
(2) 5.

Exercise 104 ⊙⊙ exologarithme-0001

Calculez sans calculatrice

(1) $\log_2 8$;
(2) $\log_{10} 10000$;
(3) $\log_4 \dfrac{1}{16}$;
(4) $\log_9 27$;

corrlogarithme-0001

Correction of the exercise 104

(1) 3 parce que $2^3 = 8$;
(2) 4 parce que $10^4 = 10000$;
(3) -2 en utilisant le fait que $1/16 = 16^{-1}$ et que $\log_4(x^{-1}) = -\log_4(x)$, puis le fait que $16 = 4^2$;

(4) $\frac{3}{2}$ parce que $27 = 3^3 = (\sqrt{9})^3 = 9^{3/2}$.

Exercise 105 exologarithme-0002

Calculer sans utiliser de calculatrices : $e^{-3\ln(8)}$ corrlogarithme-0002

Correction of the exercise 105

Nous utilisons le fait que $a^{xy} = (a^x)^y$. Cela nous permet de mettre le -3 en puissance du tout. Ensuite nous utilisons la formule $e^{\ln(a)} = a$:

$$e^{-3\ln(1/8)} = \left(e^{\ln(1/8)}\right)^{-3} = (1/8)^{-3} = 8^3 = 512. \tag{90.8}$$

90.4 Fonctions et graphes

Exercise 106 ⊙⊙ exoTD1_1

Déterminer l'ensemble de définition de la fonction $x \mapsto f(x)$ dans les cas suivants :

(1) $f(x) = \ln(x+1)$;

(2) $f(x) = \ln(|x|+1)$;

(3) $f(x) = \ln(|x+1|)$;

(4) $f(x) = \sqrt{x^2+x}$;

(5) $f(x) = \sqrt{\cos(x)}$;

(6) $f(x) = \tan(x^2+1)$;

(7) $f(x) = \dfrac{\sqrt{2x+1}-3}{\sqrt{5-x}-1}$;

(8) $f(x) = (1+x)^x$.

Exercise 107 ⊛⊛ exoTD1_2

Dessiner le graphe et signaler la périodicité et les symétries éventuelles de la fonction $x \mapsto f(x)$ dans les cas suivants :

(1) $f(x) = 1 + \cos(2x)$;

(2) $f(x) = \tan\left(x + \dfrac{\pi}{2}\right)$;

(3) $f(x) = \begin{cases} x^2 - x - 1 & \text{si } x \leqslant 1, \\ -x & \text{si } x > 1. \end{cases}$

Exercise 108 ⊘⊘ exoTD1_3

Soient $f(x) = x^2 - 3$ et $g(x) = \ln(1+x)$. Écrivez les expressions explicites et trouvez les domaines des fonctions composées $h_1(x) = f \circ g(x)$ et $h_2(x) = g \circ f(x)$.

Exercise 109 ⊖⊖ exoTD1_4

Soit f la fonction $x \mapsto x^3$. Dessinez les graphes des fonctions

(1) $f(x) - 1$;

(2) $f(x+3)$;

(3) $f(x-1)$;

(4) $-f(x)$;

(5) $f(-x)$;

(6) $|f(x)|$;

(7) $\min\{f(x), 0\}$;

(8) $\max\{1, f(x)\}$.

90.5 Limites du côté de l'infini

Exercise 110 exoSVT-0001

Déterminer lorsqu'elles existent les limites suivantes

(1) $\lim_{x\to+\infty} x^2$;
(2) $\lim_{x\to+\infty} x^2 - x^4$;
(3) $\lim_{x\to+\infty} \frac{x^2-2}{x+1}$;
(4) $\lim_{x\to+\infty} \frac{x+1}{x^3}$;
(5) $\lim_{x\to+\infty} \sin x$;
(6) $\lim_{x\to-\infty} 2^x$;
(7) $\lim_{x\to+\infty} \frac{x}{\sqrt{x^2+1}}$;

corrSVT-0001

Correction of the exercise 110

(1) ∞
(2) $x^2 - x^4 = x^2(1-x^4) \to -\infty$.
(3) ∞
(4) 0
(5) Non déterminé.
(6) 0
(7) Par mise en évidence de x^2 sous la racine,

$$\frac{x}{\sqrt{x^2+1}} = \frac{x}{x\sqrt{1+\frac{1}{x^2}}} = \frac{1}{\sqrt{1+\frac{1}{x^2}}} \to 1. \tag{90.9}$$

La limite est donc 1.

Exercise 111 exoTD3-0003

Déterminer les limites des suites suivantes.

(1) $u_n = \frac{n^2+3n+1}{1+n^2+n^3}$,
(2) $u_n = \frac{|\sin(n)|}{n+1}$,
(3) $u_n = \frac{2^n}{n^5+3n+1}$,
(4) $u_n = (n^4+1)(n^3+2)e^{-n}$,
(5) $u_n = \frac{\ln(n)}{n+1}$,
(6) $u_n = \sqrt{n^2+n+1} - \sqrt{n^2-n+1}$,
(7) $u_n = \sqrt{n+\sqrt{n^2+1}} - \sqrt{n+\sqrt{n^2-1}}$ (pour $n \in \mathbb{N}_0$),
(8) $u_n = \frac{n-\sqrt{n^2+1}}{n-\sqrt{n^2-1}}$ (pour $n \neq 0$).

corrTD3-0003

Correction of the exercise 111

(1) Mettons le plus haut degré en évidence au numérateur et au dénominateur :

$$u_n = \frac{n^2\left(1+\frac{2n}{n^2}+\frac{1}{n^2}\right)}{n^3\left(1+\frac{1}{n^3}+\frac{n^2}{n^3}\right)} = \frac{1+\frac{3}{n}+\frac{1}{n^2}}{n\left(1+\frac{1}{n^3}+\frac{1}{n}\right)}. \tag{90.10}$$

Cette dernière expression montre que le limite est zéro lorsque $n \to \infty$.

(2) C'est le produit d'une suite bornée ($|\sin(n)|$) par une suite qui tend vers zéro ($\frac{1}{n+1}$). Cette suite converge donc vers zéro.

(3) Il s'agit d'une exponentielle en base deux (2^n) divisée par un polynôme ; la limite est donc ∞ parce que l'exponentielle va toujours plus vite.

(4) Il faut se rappeler que $e^{-n} = \frac{1}{e^n}$. La formule donnée est donc un polynôme *divisée* par une exponentielle. Le résultat tend vers zéro.

(5) Ici encore il s'agit d'une limite remarquable. Le logarithme au numérateur tend vers l'infini moins vite que le polynôme au dénominateur. Le résultat est que la limite est zéro.

(6) Il faut multiplier et diviser par le binôme conjugué. Nous avons

$$\begin{aligned} u_n &= \frac{\left(\sqrt{n^2+n+1}-\sqrt{n^2-n+1}\right)\left(\sqrt{n^2+n+1}+\sqrt{n^2-n+1}\right)}{\sqrt{n^2+n+1}+\sqrt{n^2-n+1}} \\ &= \frac{(n^2+n+1)-(n^2-n+1)}{\sqrt{n^2+n+1}+\sqrt{n^2-n+1}} \\ &= \frac{2n}{\sqrt{n^2+n+1}+\sqrt{n^2-n+1}}. \end{aligned} \tag{90.11}$$

Maintenant il faut mettre la plus haute puissance de n en évidence dans le dénominateur. La manipulation à faire est

$$\begin{aligned} \sqrt{n^2+n+1} &= \sqrt{n^2\left(1+\frac{1}{n}+\frac{1}{n^2}\right)} \\ &= |n|\sqrt{1+\frac{1}{n}+\frac{1}{n^2}} \\ &= n\sqrt{1+\frac{1}{n}+\frac{1}{n^2}} \end{aligned} \tag{90.12}$$

où nous avons pu enlever la valeur absolue parce que nous considérons que $n \to \infty$. Ce n est donc certainement positif. Nous avons donc

$$\begin{aligned} u_n &= \frac{2n}{\sqrt{n^2+n+1}+\sqrt{n^2-n+1}} \\ &= \frac{2n}{n\sqrt{1+\frac{1}{n}+\frac{1}{n^2}}+n\sqrt{1-\frac{1}{n}+\frac{1}{n^2}}}. \end{aligned} \tag{90.13}$$

À ce moment nous pouvons simplifier par n. Il reste 2 au numérateur tandis que le dénominateur tend vers 2. La limite est donc 1.

(7) Ici il faut commencer par multiplier et diviser par le binôme conjugué, c'est-à-dire par

$$\sqrt{n+\sqrt{n^2+1}}+\sqrt{n+\sqrt{n^2-1}}. \tag{90.14}$$

Nous trouvons

$$\sqrt{n+\sqrt{n^2+1}}-\sqrt{n+\sqrt{n^2-1}} = \frac{\sqrt{n^2+1}-\sqrt{n^2-1}}{\sqrt{n}\sqrt{1+\sqrt{1+\frac{1}{n}}}+\sqrt{n}\sqrt{1+\sqrt{1-\frac{1}{n}}}}. \tag{90.15}$$

Eqtdtzzztsqg

Maintenant le dénominateur tend clairement vers l'infini tandis que le numérateur est encore une forme indéterminée. Il faut donc le traiter encore une fois avec le coup du binôme conjugué :

$$\sqrt{n^2+1}-\sqrt{n^2-1} = \frac{n^2+1-n^2+1}{\sqrt{n^2+1}+\sqrt{n^2-1}}. \tag{90.16}$$

Par conséquent le numérateur de (90.15) tend vers zéro lorsque $n \to \infty$. Au final, nous avons

$$\lim_{n\to\infty} \sqrt{n+\sqrt{n^2+1}}-\sqrt{n+\sqrt{n^2-1}} = 0 \tag{90.17}$$

(8) Nous multiplions et divisons par le binôme conjugué des du numérateur *et* du dénominateur. Nous trouvons alors

$$\begin{aligned} u_n &= \frac{(n-\sqrt{n^2+1})(n+\sqrt{n^2+1})(n+\sqrt{n^2-1})}{(n-\sqrt{n^2-1})(n+\sqrt{n^2+1})(n+\sqrt{n^2-1})} \\ &= \frac{\big(n^2-(n^2+1)\big)\big(n+\sqrt{n^2-1}\big)}{\big(n+\sqrt{n^2+1}\big)\big(n^2-(n^2-1)\big)} \\ &= \frac{-n-\sqrt{n^2-1}}{n+\sqrt{n^2+1}} \\ &= \frac{-n-n\sqrt{1-\frac{1}{n^2}}}{n+n\sqrt{1+\frac{1}{n^2}}} \\ &= \frac{-1-\sqrt{1-\frac{1}{n^2}}}{1+\sqrt{1+\frac{1}{n^2}}}, \end{aligned} \tag{90.18}$$

et la dernière ligne tend vers 1 lorsque $n \to \infty$.

Remarquez qu'on a effectué la manipulation

$$\sqrt{n^2-1} = \sqrt{n^2\left(1-\frac{1}{n^2}\right)} = n\sqrt{1-\frac{1}{n^2}} \tag{90.19}$$

avant de simplifier par n.

90.6 Limite de suites

Exercise 112 exoTD3-0001

Vrai ou faux ?

(1) Une suite est toujours soit majorée soit minorée.
(2) Toute suite convergente est monotone.
(3) Toute suite convergente est bornée.
(4) Si une suite est monotone et bornée, alors elle converge.
(5) Si une suite converge, alors elle est monotone et bornée.
(6) Une suite positive qui converge vers zéro est décroissante à partir d'un certain rang.
(7) Si la suite $(|x_n|)_{n\in\mathbb{N}}$ converge vers ℓ, la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ ou vers $-\ell$.
(8) Si la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ, la suite (x_{2n+n^2}) converge vers ℓ.
(9) Si la suite $(x_n - y_n)_{n\in\mathbb{N}}$ converge vers 0, les suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ convergent vers la même limite.
(10) Le produit d'une suite qui converge vers 0 et d'une suite quelconque converge vers 0.
(11) Toute suite encadrée par deux suites convergentes est convergente.
(12) La différence de deux suites équivalentes converge vers 0.
(13) Le quotient de deux suites équivalentes non nulle tend vers 1.

corrTD3-0001

Correction of the exercise 112

Il faut souvent penser à la suite $u_n = (-1)^n$ qui est un exemple de suite qui ne converge pas tout en restant bornée et qui prend « presque tout le temps » les mêmes valeurs.

(1) Faux. Par exemple la suite $(-1)^n n$. Les termes pairs tendent vers l'infini et les termes impairs vers moins l'infini. Les premiers termes sont

$$0, -1, 2, -3, 4, -5, \dots. \tag{90.20}$$

ItemTD31b

(2) Faux. La suite $\frac{(-1)^n}{n}$ tend vers zéro, mais oscille entre les positifs et les négatifs. Les premiers termes sont
$$-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{4}, \dots. \tag{90.21}$$

(3) Vrai. À partir d'un certain moment, la suite doit se « stabiliser » autour de la valeur de la limite.

(4) Vrai. C'est la proposition 90.3.

(5) Faux. Même exemple que pour le point (2). Cette suite n'est ni monotone croissante ni monotone décroissante.

(6) Faux. Nous prenons presque le même exemple. Prenons la suite
$$0, 1, 0, \frac{1}{2}, 0, \frac{1}{3}, 0, \frac{1}{4}, 0, \frac{1}{5}, \cdots \tag{90.22}$$
C'est la suite des $\frac{1}{n}$ dans laquelle nous avons inséré des 0 à une place sur deux. Cette suite n'est décroissante à partir d'aucun moment parce qu'elle n'arrête pas de passer de 0 à un nombre strictement positif.

(7) Faux. Si nous prenons la suite $u_n = (-1)^n$, nous avons $(|u_n|) = 1$. Cette dernière suite converge vers 1, mais la suite de départ ne converge pas.

(8) Vrai. Une sous-suite d'une suite convergente converge vers la même limite.

(9) Faux. Prenons les suites $x_n = n$ et $y_n = n$, c'est-à-dire deux fois la même suite. Évidemment, $(x_n - y_n) = 0$ est une suite qui converge vers 0, pourtant aucune des deux suites de départ ne converge.

(10) Faux. La suite $x_n = \frac{1}{n}$ converge vers 0. Mais si nous la multiplions par $y_n = 4n$, le produit est $x_n y_n = 4$ qui converge vers 4.

(11) Faux. À ne pas confondre avec le résultat qui dit qu'une suite encadrée par deux suites qui convergent vers *la même limite* est convergente. Prenons par exemple les suites $x_n = -2$ et $y_n = 2$. Ce sont deux suites convergentes qui encadrent la suite $z_n = (-1)^n$. Cette dernière ne converge pas.

(12) Faux. Les deux suites peuvent être équivalentes et ne pas converger, penser à $u_n = n^2 + n$ et $v_n = n^2$.
Par contre si les deux suites convergent, alors elles convergent vers la même limite. Supposons $(x_n \to \ell)$ et $(y_n \to \ell')$; dans ce cas les règles de calcul de limites s'appliquent et nous avons
$$\lim_{n\to\infty} \frac{u_n}{v_n} = \frac{\ell}{\ell'}. \tag{90.23}$$
Mais si cette limite doit valoir 1, alors nous devons avoir $\ell = \ell'$.

(13) Faux. Si la suite que l'on met au dénominateur s'annule, alors la quotient n'a pas de sens. Attention à la différence entre « suite non nulle » et « suite qui ne s'annule jamais ».

Exercise 113 exoTD3-0002
Modèle malthusien (ou géométrique). Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par
$$\begin{cases} u_{n+1} = au_n & \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.24}$$
où a et $x \geqslant 0$ sont deux nombres réels.

(1) Montrer que $u_n = xa^n$ pour tout $n \in \mathbb{N}_0$.

(2) Soient $a = \frac{1}{2}$ et $x = 1000$, en supposant que n représente un nombre de mois et u_n un nombre d'individus, dans combien de temps la population sera inférieure à 1 (cas d'extinction) ?

(3) Soient $a = 2$ et $x = 2$, en supposant que n représente un nombre de mois et u_n un nombre d'individus, dans combien de temps la population atteindra 1000 individus ?

corrTD3-0002

Correction of the exercise 113

(1) Pour u_0, nous avons $u_0 = x = xa^0$; pour rappel, $a^0 = 1$ pour tout a.
Supposons que la formule soit vraie pour u_k, c'est-à-dire que nous avons $u_k = xa^k$ pour un certain k. Dans ce cas, nous avons

$$u_{k+1} = au_k = axa^k = xa^{k+1} \tag{90.25}$$

parce que $aa^k = a^{k+1}$. Donc la formule proposée est également correcte pour u_{k+1}. Par récurrence, elle est donc correcte pour tous les k.

(2) Nous avons

$$u_n = \frac{1000}{2^n}. \tag{90.26}$$

Il faut donc savoir à partir de quel n nous avons $2^n \geqslant 1000$. La réponse est $n = 10$ ($2^{10} = 1024$).

(3) Dans le cas présent, $u_n = 22^n = 2^{n+1}$. Cette suite dépasse 1000 pour $n + 1 = 10$, c'est-à-dire $n = 9$.

Exercise 114 exoTD3-0004

Modèle arithmétique. Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = u_n + b \quad \forall n \in \mathbb{N}_0 \\ u_0 = x \end{cases} \tag{90.27}$$

où b et $x \geqslant 0$ sont deux nombres réels. Montrer que $u_n = bn + x$ pour tout n.

corrTD3-0004

Correction of the exercise 114

Lorsque $n = 0$, la formule est vraie parce que $u_0x = bn + x$ si $n = 0$.

Supposons maintenant que la formule soit vraie pour un certain k, c'est-à-dire que $u_k = kb + x$. Dans ce cas, le terme suivant de la suite vaut

$$u_{k+1} = u_k + b = bk + x + b = (k + 1)b + x, \tag{90.28}$$

ce qui est bien la formule demandée pour le terme u_{k+1}.

Par récurrence, la formule est vraie pour tous les termes de la suite.

Exercise 115 exoTD3-0005

Modèle malthusien contrôlé. Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = au_n + b, \quad \text{pour tous } n \in \mathbb{N},\ n > 0 \\ u_0 = x \end{cases}$$

où $a \neq 1$, b et $x \geqslant 0$ sont des nombres réels.

(1) Calculer les premiers 4 termes de la suite si $a = 2$ et $x = 10$.

(2) Montrer que $u_n = a^n(x + \frac{b}{a-1}) - \frac{b}{a-1}$ pour tout $n \in \mathbb{N}$.

(3) En supposant que $a = 2$ et $x = 100$ trouver b tel que la population se stabilise, c'est-à-dire telle que la population tende vers une valeur finie ou nulle.

corrTD3-0005

Correction of the exercise 115

Pour les premiers termes nous avons

$$\begin{aligned} u_0 &= 100 \\ u_1 &= 20 + b \\ u_2 &= 2(20 + b) + b = 40 + 3b \\ u_3 &= 2(40 + 3b) + b = 80 + 7b \\ u_4 &= 2(80 + 7b) + b = 160 + 15b. \end{aligned} \tag{90.29}$$

Une façon de voir le concept de récurrence est de considérer que nous avons deux suites. La première est celle de l'énoncé :

$$\begin{cases} u_{n+1} = au_n + b \quad \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.30}$$

et la seconde est

$$v_n = a^n \left(x + \frac{b}{a-1} \right) - \frac{b}{a-1}. \tag{90.31}$$

Ce que nous devons faire est de montrer que $u_n = v_n$ pour tout n.

Pour $n = 0$, cela est vrai parce que $u_n = x$ (par définition) tandis qu'en posant $n = 0$ dans la définition de v_n nous trouvons

$$v_0 = a^0 \left(a + \frac{b}{a-1} \right) - \frac{b}{a-1} = x \tag{90.32}$$

parce que $a^0 = 1$.

Supposons que pour un certain k, nous ayons $u_n = v_n$, et montrons qu'alors $u_{k+1} = v_{k+1}$. Notre supposition nous dit que

$$u_k = v_k = a^k \left(x + \frac{b}{a-1} \right) - \frac{b}{a-1}. \tag{90.33}$$

Maintenant, nous utilisons la formule de définition de u_{k+1} :

$$\begin{aligned} u_{k+1} &= au_k + b \\ &= a \left[a^k \left(x + \frac{b}{a-1} \right) - \frac{b}{a-1} \right] + b \\ &= a^{k+1} \left(x + \frac{b}{a-1} \right) - \frac{ab}{a-1} + b \\ &= a^{k+1} \left(x + \frac{b}{a-1} \right) + \frac{-ab + b(a-1)}{a-1} \\ &= a^{k+1} \left(x + \frac{b}{a-1} \right) + \frac{-b}{a-1} \\ &= v_{k+1}. \end{aligned} \tag{90.34}$$

Donc nous avons bien trouvé que $u_{k+1} = v_{k+1}$ sous l'hypothèse que $u_k = v_k$.

La preuve par récurrence est terminée : nous sommes maintenant sûr que $u_n = v_n$ pour tout n.

Exercise 116 exoTD3-0006

Dynamique des populations : modélisation.

On considère une population de dix mille micro-organismes. Parmi ces micro-organismes, certains portent un gène qui les rend phosphorescents. La population X est celle portant ce gène et Y est le reste de la population.

Chaque jour, on mesure la taille de chaque population et on constate que 20% des micro-organismes de X mutent vers Y alors que 5% des micro-organismes de Y mutent vers X.

Sachant qu'au jour 0, un quart des micro-organismes est phosphorescent et en supposant que le nombre de ces micro-organismes reste constant, quelle est la population de X et de Y au bout de 1 jour, 2 jours, 5 jours et 10 jours ? corrTD3-0006

Correction of the exercise 116

Nous notons N la quantité totale de micro-organismes (c'est-à-dire $N = 10000$), puis X_n la quantité des micro-organismes phosphorescents au jour n, et Y_n la quantité de ceux qui ne le sont pas, nous avons

$$X_0 = \frac{N}{4} \tag{90.35a}$$
$$Y_0 = \frac{3N}{4}, \tag{90.35b}$$

et la loi d'évolution est

$$X_{n+1} = X_n - \frac{20X_n}{100} + \frac{5Y_n}{100} \tag{90.36a}$$
$$Y_{n+1} = Y_n + \frac{20X_n}{100} - \frac{5Y_n}{100}. \tag{90.36b}$$

En simplifiant les fractions,

$$X_{n+1} = \frac{4}{5}X_n + \frac{1}{20}Y_n. \tag{90.37}$$ EqunevolXnTn

Étant donné que la quantité totale de micro-organismes est constante et vaut N, nous avons toujours $X_n + Y_n = N$. Nous pouvons donc remplacer Y_n par $(N - X_n)$ dans l'équation (90.37) pour trouver

$$X_{n+1} = \frac{3}{4}X_n + \frac{N}{20}. \tag{90.38}$$

Et nous sommes maintenant dans un modèle, pour X_n, du même type que celui de l'exercice 115 avec $a = \frac{3}{4}$, $b = \frac{N}{20}$ et $x = X_0 = \frac{N}{4}$. L'évolution de X_n est donc donnée par

$$\begin{aligned} X_n &= \left(\frac{3}{4}\right)^n \left(X_0 + \frac{N/20}{\frac{3}{4} - 1}\right) - \frac{N/20}{\frac{3}{4} - 1} \\ &= \left(\frac{3}{4}\right)^n \left(\frac{N}{4} - 4\frac{N}{20}\right) + 4\frac{N}{20} \\ &= \left(\frac{3}{4}\right)^n \frac{N}{20} + \frac{N}{5} \end{aligned} \tag{90.39}$$ EqrefXnnform

où nous avons utilisé le fait que $X_0 = \frac{N}{4}$. À partir de cette expression, nous pouvons retrouver Y_n simplement en calculant X_n et en faisant $Y_n = N - X_n$.

Pour trouver la population de type X au bout de 1 jour, il suffit de poser $n = 1$ dans l'expression (90.39). Le résultat est $\frac{19}{80}N = 2375$. Au bout de deux jours il en reste $X_2 = 2281.25$, $X_5 = 2118.65$ et $X_10 = 2028.16$.

Exercise 117 exoTD3-0007

Modèle à ressources limitées (compétition). Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = u_n f(u_n) & \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.40}$$

où $x \geqslant$ est un nombre réel et $f\colon [0, \infty[\to [0, a]$ est la fonction définie par

$$f(y) = \begin{cases} a & \text{si } y < 2 \\ \frac{2a}{y} & \text{si } y \geqslant 2. \end{cases} \tag{90.41}$$

Déterminer la limite de la suite $(u_n)_{n\in\mathbb{N}}$ dans les cas suivants.

(1) $x = 3$ et $a = \frac{1}{2}$. Conseil : commencer par montrer que $0 < u_n < 2$ pour tout $n \geqslant 1$.
(2) $x = 1$ et $a = 2$.

corrTD3-0007

Correction of the exercise 117

Ici nous ne pouvons pas utiliser la proposition 90.1 pour trouver des candidats limites parce que la fonction qui lie u_{n+1} à u_n n'est pas continue.

(1) Nous avons $u_0 = 3$ et

$$f(y) = \begin{cases} \frac{1}{2} & \text{si } y < 2 \\ \frac{1}{y} & \text{si } y \geqslant 2. \end{cases} \tag{90.42}$$

Calculons quelques termes.

$$\begin{aligned} u_0 &= 3 \\ u_1 &= f(u_0)u_0 = f(3) \cdot 3 = \frac{1}{3}3 = 1 \\ u_2 &= f(u_1)u_1 = f(1) = \frac{1}{2} \\ u_3 &= f(u_2)u_2 = f(\frac{1}{2})\frac{1}{2} = \frac{1}{4}. \end{aligned} \tag{90.43}$$

Ce que nous remarquons est que, à partir du moment où la suite passe en dessous de 1, le nombre est divisé par deux à chaque pas. Cela est à cause du fait que dès que $u_n < 1$, alors $u_{n+1} = f(u_n)u_n = \frac{1}{2}u_n$.
La suite tend donc vers zéro.

(2) Cette fois nous avons

$$f(y) = \begin{cases} 2 & \text{si } u < 2 \\ \frac{4}{y} & \text{si } y \geqslant 2 \end{cases} \tag{90.44}$$

Calculons quelques termes :

$$\begin{aligned} u_0 &= 1 \\ u_1 &= 2u_0 = 2 \\ u_2 &= f(2)u_2 = \frac{4}{2}2 = 4 \\ u_3 &= f(4) \cdot 4 = 1 \cdot 4 = 4. \end{aligned} \tag{90.45}$$

La suite reste alors bloquée à 4. En effet si $u_n = 4$, alors

$$u_{n+1} = f(u_n)u_n = \frac{4}{u_n} \cdot 4 = \frac{4}{4} \cdot 4 = 4. \tag{90.46}$$

La limite est donc 4.

Notez que dans cet exercice nous avons eu de la chance : rien qu'en calculant les premiers termes, nous avons pu comprendre comment se comporte la suite et déterminer la limite ... ce ne sera pas toujours le cas.

Exercise 118 exoTD3-0008

Modèle à ressources limitées (partage). Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = u_n f(u_n) & \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.47}$$

où $x \geqslant 0$ est un nombre réel et $f\colon [0, \infty[\to [0, a]$ est la fonction définie par

$$f(y) = \begin{cases} a & \text{si } y < 2 \\ 0 & \text{si } y \geqslant 2 \end{cases} \tag{90.48}$$

(1) Déterminer la limite de la suite $(u_n)_{n\in\mathbb{N}}$ dans les cas suivants :
   (1a) $a = \frac{1}{2}$ et $x \geqslant 0$ quelconque,
   (1b) $a = 1$ et $x = 1$,
   (1c) $a = 1$ et $x = 3$,
   (1d) $a = 2$ et $x \geqslant 0$ quelconque.

(2) On considère $x = 1$ avec pour unité le million d'individus et le mois comme unité de temps. Au bout de combien de temps la population sera-t-elle intérieure à 100.000 individus dans le cas où $a = \frac{9}{10}$?
Même question dans le cas $a = \frac{11}{10}$.

corrTD3-0008

Correction of the exercise 118

(1) Commençons par dire quelque chose de valable dans tous les cas : si $u_n = 0$, alors $u_{n+1} = 0$ parce que

$$u_{n+1} = u_n f(u_n) \tag{90.49}$$

qui vaut zéro lorsque $u_n = 0$.
D'autre part, si $u_n \geqslant 2$, alors $u_{n+1} = u_n f(u_n) = 0$. Par conséquent, dès que la suite dépasse 2, elle retombe immédiatement à zéro.

   (1a) Les cas $x = 0$ et $x \geqslant 2$ sont traités par la remarque que nous venons de faire.
   Il reste à voir ce qui se passe si $x \in]0, 2[$. Si $u_n \in]0, 1[$, alors $f(u_n) = \frac{1}{2}$, et donc

$$u_{n+1} = \frac{u_n}{2}. \tag{90.50}$$

   Par conséquent, dès que la suite entre dans l'intervalle $]0, 2[$, la suite y reste (parce que dans ce cas, $\frac{u_n}{2}$ est encore entre 0 et 2) et est décroissante (la moitié à chaque pas). Nous en déduisons que dans ce cas, la suite tend vers 0.
   En conclusion, pour tous les $x \geqslant 0$, la suite va tendre vers 0.

   (1b) Nous avons

$$f(y) = \begin{cases} 1 & \text{si } y < 2 \\ 0 & \text{si } y \geqslant 2 \end{cases} \tag{90.51}$$

   Le calcul des premiers termes est immédiat : $u_0 = 1$, $u_1 = 1 \cdot f(1) = 1$, et puis la suite reste en fait constante égale à 1.

   (1c) Cette fois, nous partons de $u_0 = 3$. Donc $f(u_0) = 0$, et la suite tombe immédiatement à zéro et y reste. Elle converge donc vers zéro.

   (1d) Si $u_0 \geqslant 2$, alors $f(u_0) = 0$ et la suite est immédiatement zéro.
   Si $u_0 < 2$, alors $u_1 = 2u_0 > u_1$. Tant que la suite reste plus petite que 2, la valeur est doublée à chaque pas. À un moment donné, elle va dépasser 2, et à ce moment, elle retombe immédiatement à zéro. La suite tend donc vers zéro.

(2) Dans cette question, nous partons avec $u_0 = 1$, et on demande pour que n nous aurons $u_n \leqslant 0.1$. (cent mille, c'est zéro virgule un million) Afin de comprendre comment les choses se passent, calculons les premiers termes. La fonction f est donnée par

$$f(y) = \begin{cases} \frac{9}{10} & \text{si } y < 2 \\ 0 & \text{si } y \geqslant 2 \end{cases} \tag{90.52}$$

EqLaFnfcandaad

donc

$$\begin{aligned} u_0 &= 1 \\ u_1 &= \frac{9}{10}u_0 = \frac{9}{10} \\ u_2 &= \frac{9}{10}u_1 = \frac{9}{10} \cdot \frac{9}{10} = \left(\frac{9}{10}\right)^2 \\ u_3 &= \frac{9}{10}u_2 = \left(\frac{9}{10}\right)^3, \end{aligned} \tag{90.53}$$

etc. Nous voyons que

$$u_n = \left(\frac{9}{10}\right)^n. \tag{90.54}$$

La question revient à trouver pour quel n nous avons

$$\left(\frac{9}{10}\right)^n = 0.1, \tag{90.55}$$

la réponse est donnée par

$$n = \log_{9/10}(0.1) = 21.8, \tag{90.56}$$

et il faut donc 22 mois pour passer en dessous de la limite des 100 mille.

Si par contre nous prenons $a = 11/10$, alors nous avons, pour les premiers termes la suite

$$u_n = \left(\frac{11}{10}\right)^n, \tag{90.57}$$

et la suite est croissante. Mais dès que la suite dépasse la valeur 2, elle retombe à zéro parce que

$$f(y) = \begin{cases} \frac{11}{10} & \text{si } y < 2 \\ 0 & \text{si } y \geqslant 2 \end{cases} \tag{90.58}$$

En calculant

$$n = \log_{11/10}(2), \tag{90.59}$$

nous trouvons que la suite dépasse 2 en u_8. Par conséquent $u_9 = 0$.

Exercise 119 exoTD3-0009

Modèle de Hassel.

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = \frac{au_n}{(1+u_n)^b} & \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.60}$$

où a, b et $x \geqslant 0$ sont des nombres réels.

(1) On suppose que $a = \frac{1}{2}$ et que b et x sont des nombres positifs quelconques. Montrer que $u_n \geqslant 0$ pour tout $n \in \mathbb{N}$ et que la suite (u_n) est décroissante. En déduire que $\lim_{n\to\infty} u_n = 0$.

(2) On suppose que $a = 2$, $b = 1$ et $x = 2$. Montrer que $\lim_{n\to\infty} u_n = 1$.

(3) On suppose que $a = 2$, $b = 1$ et $x = \frac{1}{2}$. Montrer que $\lim_{n\to\infty} u_n = 1$.

(4) On suppose que $a = 2$, $b = \frac{1}{2}$ et $x = 2$. Montrer que $\lim_{n\to\infty} u_n = 3$.

corrTD3-0009

Correction of the exercise 119

C'est la combinaison des propositions 90.1 et 90.3 qui va nous permettre de nous en sortir.

Lorsque u_{n+1} est donné directement en fonction de u_n, c'est-à-dire lorsque $u_{n+1} = f(u_n)$, les « candidats limites » sont les nombres u tels que

$$u = f(u). \tag{90.61}$$

Dans le cas qui nous intéresse ici, les candidats seront donnés par l'équation

$$u = \frac{au}{(1+u)^b}. \tag{90.62}$$

Equuuuhp

La première solution à cette équation est $u = 0$. Afin de trouver les solutions différentes de zéro, nous pouvons simplifier l'équation (90.62) par u :

$$1 = \frac{a}{(1+u)^b}, \tag{90.63}$$

et donc $(1+u)^b = a$, c'est-à-dire

$$u = a^{1/b} - 1. \tag{90.64}$$

Dans tous les cas nous savons donc que la limite de la suite (si elle existe !) sera donnée par zéro ou par $a^{1/b} - 1$.

(1) Nous allons prouver que la suite est décroissante et bornée vers le bas par zéro. Cela prouvera qu'elle est convergente. Après, il faudra déterminer si la limite est zéro ou bien $a^{1/b} - 1$.
Montrons que $u_n > 0$ pour tout n. Nous faisons cela par récurrence. D'abord l'énoncé dit que u_0 est positif. Ensuite, si u_n est positif, alors

$$u_{n+1} = \frac{1}{2}\frac{u_n}{(1+u_n)^b} \tag{90.65}$$

est également positif.
Pour prouver la décroissance, nous calculons

$$\frac{u_{n+1}}{u_n} = \frac{a}{(1+u_n)^b} = \frac{1}{2}\frac{1}{(1+u_n)^b}. \tag{90.66}$$

Maintenant, le truc est de remarquer que si $u_n > 0$, alors, $1 + u_n > 1$ et donc $(1+u_n)^b > 1$. Cela fait que la fraction est plus petite que 1 et donc que

$$\frac{u_{n+1}}{u_n} < 1. \tag{90.67}$$

Cela signifie que $u_n > u_{n+1}$ et donc que la suite est décroissante.
Nous savons maintenant que la suite est convergente. Les deux candidats à être limite sont 0 et $a^{1/b} - 1 = (\frac{1}{2})^{1/b} - 1 < 0$. Étant donné que nous avons prouvé que la suite reste positive, le second candidat limite n'est pas possible (il est strictement négatif parce que $(\frac{1}{2})^{1/b}$ est toujours négatif). La limite est donc zéro.

(2) Si $a = 2$, $b = 1$ et $u_0 = 2$, nous avons comme candidats les nombres 0 et $2^1 - 1 = 1$. Notre but est donc de montrer que la suite est convergente et d'écarter la possibilité que la limite soit zéro.
La stratégie sera donc de montrer que la suite est toujours plus grande ou égale à 1, et de montrer qu'elle est décroissante. Comme ça elle sera convergente (décroissante et bornée vers le bas), et la limite ne pourra pas être 0.
Supposons que $u_n > 1$. Alors $u_{n+1} = \frac{2u_n}{1+u_n} > 1$. En effet, $2u_n - (1+u_n) = u_n - 1 > 0$, ce qui signifie que $2u_n > (1+u_n)$ et donc que $\frac{2u_n}{1+u_n} > 1$. La suite est donc bornée vers le bas par 1. Elle est également décroissante parce que

$$\frac{u_{n+1}}{u_n} = \frac{2}{1+u_n} < 1 \tag{90.68}$$

si $u_n > 1$.
La suite est donc décroissante, bornée vers le bas par 1 (donc convergente) et ses candidats limites sont 0 et 1. La seule limite possible est donc 1.

(3) La suite est donnée par

$$u_{n+1} = \frac{2u_n}{1+u_n} \tag{90.69}$$

et $u_0 = \frac{1}{2}$. Les candidats limites sont $u = 0$ et $u = a^{1/b} - 1 = 2 - 1 = 1$. Étant donné que la suite part de $\frac{1}{2}$, nous devrions montrer que la suite est croissante et bornée par le haut.
Pour la croissance, supposons que $u \leqslant 1$ et calculons

$$\frac{u_{n+1}}{u_n} = \frac{2}{1+u_n} \geqslant \frac{2}{1+1} = 1. \tag{90.70}$$

Donc tant que la suite est plus petite que 1, elle est croissante.

Pour prouver que la suite est bornée par 1, nous allons prouver que si $\frac{1}{2} \leqslant u_n \leqslant 1$, alors $u_{n+1} \leqslant 1$. En effet demander $u_{n+1} \leqslant 1$ est équivalent à demander

$$2u_n \leqslant 1 + u_n, \tag{90.71}$$

parce que $1 + u_n > 0$. C'est-à-dire $u_n \leqslant 1$. Donc par récurrence tous les termes de la suite sont plus petits ou égaux à 1. Or nous venons de dire que tant que la suite est plus petite que 1, elle est croissante. La suite est donc toujours croissante.

La suite étant croissante et bornée par 1,elle est convergente. Mais comme les deux seuls candidats sont 0 et 1, vu qu'on part de $\frac{1}{2}$, la seule limite possible est 1.

(4) Les candidats limites sont $u = 0$ et $u = 2^{1/(1/2)} - 1 = 2^2 - 1 = 3$. La suite part de 2. Donc si nous pouvions montrer que la suite est croissante et bornée, ce serait gagné.

La suite est donnée par

$$u_{n+1} = \frac{2u_n}{\sqrt{1 + u_n}} \tag{90.72}$$

Nous montrons que tant que $u_n \leqslant 3$, la suite est croissante. Il existe plusieurs possibilités.

(4a) Nous avons la majoration

$$\frac{u_{n+1}}{u_n} = \frac{2}{\sqrt{1 + u_n}} \geqslant \frac{2}{\sqrt{1 + 3}} = 1. \tag{90.73}$$

La suite est donc croissante tant qu'elle est en dessous de 3. Il peut encore arriver que la suite dépasse 3. Nous montrons à présent qu'il n'en est rien. Nous allons montrer que si u_n est plus petit que 3, alors u_{n+1} est encore plus petit que 3.

Pour cela nous considérons la fonction

$$f(x) = \frac{2x}{\sqrt{1 + x}}. \tag{90.74}$$

C'est une fonction croissante pour les x positifs parce que sa dérivée vaut

$$\frac{2}{\sqrt{1 + x}} - \frac{x}{(1 + x)^{3/2}} = \frac{2(x + 1) - x}{(x + 1)^{3/2}} = \frac{x + 2}{(x + 1)^{3/2}} \geqslant 0. \tag{90.75}$$

Donc la plus grande valeur de $f(x)$ pour $x \in [0, 3]$ est $f(3) = 3$. En particulier, la plus grande valeur que pourrait prendre

$$u_{n+1} = \frac{2u_n}{\sqrt{1 + u_n}} = f(u_n) \tag{90.76}$$

lorsque u_n est plus petit que 3 est 3. La suite est donc bornée par 3.

(4b) Une autre façon de prouver le fait que la suite soit bornée par 3 est de procéder par l'absurde en supposant que la suite passe par une valeur plus grande que 3. Tout d'abord nous pouvons perfectionner la majoration (90.73) en écrivant

$$\frac{u_{n+1}}{u_n} = \frac{2}{(1 + u_n)^{1/2}}. \tag{90.77}$$

Si $u_n > 3$, la suite est décroissante. Supposons que $u_n < 3$, $u_{n+1} > 3$ et calculons u_{n+2}. Nous nous attendons à avoir $u_{n+2} < u_{n+1}$ parce que $u_{n+1} > 3$. Calculons :

$$\begin{aligned} n_{n+2} &= \frac{2u_{n+1}}{(1 + u_{n+1})^{1/2}} \\ &= u_{n+1} \underbrace{\frac{2(1 + u_n)^{1/4}}{\big((1 + u_n)^{1/2} + 2u_n\big)^{1/2}}}_{A}. \end{aligned} \tag{90.78}$$

Nous allons montrer que si $u_n < 3$, alors $A > 1$, ce qui signifierais que $u_{n+2} > u_{n+1}$ et contredirait le fait que la suite est croissante tant que $u_n > 3$. Demander $A > 1$ revient à demander

$$2(1+u_n)^{1/4} < \big((1+u_n)^{1/2} + 2u_n\big)^{1/2}. \tag{90.79}$$

En élevant deux fois au carré nous trouvons l'inéquation

$$-4u_n^2 + 9u_n + 9 > 0. \tag{90.80}$$ EqzztzzznqunA

Le polynôme du membre de gauche n'est positif que entre $-3/4$ et 3. Donc si $u_n > 3$, l'inégalité (90.80) n'est jamais satisfaite. Par conséquent $A < 1$ et $u_{n+2} < u_{n+1}$, ce qui contredit le fait que $u_n > 3$.

Exercise 120 exoTD3-0010

Contrôle par introduction d'une population stérile.

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_0 = 1 \\ u_{n+1} = \frac{2u_n}{1+u_n}\frac{u_n}{u_n+s} \quad \forall n \in \mathbb{N}_0 \end{cases} \tag{90.81}$$

où s est un nombre réel positif.

(1) Montrer que, pour $s \geqslant 2$, $u_n \in [0,1]$ pour tout $n \in \mathbb{N}$.

(2) Montrer que, pour $s \geqslant 2$, la suite (u_n) est décroissante.

(3) En déduire que, en choisissant $s \geqslant 2$, on amène la population vers l'extinction. corrTD3-0010

Correction of the exercise 120

(1) On démontre ce premier point par récurrence. D'abord par définition $u_0 = 1$, donc $u_0 \in [0,1]$. Ensuite si $u_n \in [0,1]$ pour touts indices n entre 0 et k alors on a

$$u_{k+1} = \frac{2u_k}{1+u_k}\frac{u_k}{u_k+s} \leqslant \frac{2}{1}\frac{1}{s} \leqslant 1.$$

La première inégalité est obtenue en observant que pour majorer une fraction il faut majorer le numérateur et minorer le dénominateur. Dans notre cas, le numérateur est toujours $\leqslant 2 \cdot 1$ parce que $u_k \leqslant 1$ et le dénominateur est supérieur à s, parce que $u_k \geqslant 0$.

(2) Pour démontrer la décroissance, nous écrivons

$$\frac{u_{n+1}}{u_n} = \frac{2u_n}{(1+u_n)(s+u_n)} \leqslant \frac{2u_n}{(u_n+s)} \tag{90.82}$$

parce que $u_n \in [0,1]$, donc en remplaçant dans le dénominateur $(1+u_n)$ par 1, la fraction diminue. De la même manière, si nous remplaçons dans le dénominateur u_n par zéro et s par 2, nous agrandissons la fraction :

$$\frac{2u_n}{u_n+s} \leqslant \frac{2u_n}{2} = u_n \leqslant 1. \tag{90.83}$$

Par conséquent la suite est décroissante.

Une méthode alternative pour montrer la décroissance est de remarquer que la récurrence définissant la suite peut être écrite sous la forme

$$u_{n+1} = u_n g(u_n) \tag{90.84}$$

avec

$$g(x) = \frac{2x}{(1+x)(x+s)}. \tag{90.85}$$

Cette fonction est une fonction rationnelle. Pour la majorer il faut majorer son numérateur et minorer sont dénominateur, en sachant que $x \in [0,1]$. On obtient alors

$$g(x) \leqslant \frac{2}{s} \leqslant 1. \tag{90.86}$$

Nous avons donc $u_{n+1} = g(u_n)u_n \leqslant u_n \leqslant 1$.

(3) La suite est bornée et décroissante. Elle est donc convergente. Rien n'oblige cependant que la limite soit exactement zéro. Pour le montrer, nous allons chercher les candidats limites, et montrer que seul 0 convient. Il y a plusieurs façons de procéder.

(3a) Pour cela, écrivons la suite sous la forme

$$u_{n+1} = u_n g(u_n) \tag{90.87}$$

avec

$$g(x) = \frac{2x}{(1+x)(x+s)}. \tag{90.88}$$

Le premier candidat limite est $u = 0$. Les autres s'obtiennent en résolvant l'équation $g(x) = 1$. D'abord nous remarquons que $g(0) = 0$. Ensuite, en passant à la dérivée, nous voyons que la fonction g est croissante :

$$g'(x) = \frac{2(x+1)(s+x) - 2x(x+1) - 2x(s+x)}{(x+1)^2(x+s)^2}. \tag{90.89}$$

Le dénominateur est certainement positif. Le numérateur vaut $-2x^2 + 2s$. Étant donné que nous considérons $x \in [0, 1]$ et $s > 2$, nous avons $g'(x) > 0$. Le maximum de la fonction g entre 0 et 1 est alors

$$g(1) = \frac{1}{s+1} < 1. \tag{90.90}$$

La fonction g ne passe donc jamais par 1 entre $x = 0$ et $x = 1$.

Le seul candidat limite est donc zéro.

(3b) La recherche directe des points fixe revient à résoudre l'équation

$$x = \frac{2x^2}{(x+1)(x+s)}. \tag{90.91}$$

La solution $x = 0$ est évidente. Cherchons les autres solutions. Elles reviennent à résoudre

$$x^2 + x(1 + s - 2) + s = 0. \tag{90.92}$$

Les racines sont

$$x = \frac{-(s-1) \pm \sqrt{(s-1)^2 - 4s}}{2}. \tag{90.93}$$

Ce qui se trouve dans la racine carrée est automatiquement plus petit que $(s-1)^2$, par conséquent la racine carrée est plus petite que $(s-1)$ et la combinaison $-(s-1) \pm \sqrt{\cdots}$ est négative. Par conséquent les derniers candidats limites sont négatifs et donc à rejeter parce que nous savons que la suite est positive.

Étant donné que nous savons que la suite est convergente, nous savons qu'elle converge vers zéro.

Exercise 121 [exoTD3-0011]

Les lapins de Fibonacci.

On considère des couples de lapins tels que, chaque mois, chaque couple donne naissance à un nouveau couple qui devient lui-même productif dès l'âge de deux mois. Après n mois, leur nombre u_n est donné par la suite

$$\begin{cases} u_{n+2} = u_{n+1} + u_n \quad \forall n \in \mathbb{N} \\ u_0 = 0 \\ u_1 = 1. \end{cases} \tag{90.94}$$

(1) Montrer que pour tout $n \in \mathbb{N}$,

$$u_n = \frac{1}{\sqrt{5}}\Big[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \Big]. \tag{90.95}$$

(2) Justifier, dans ce cas très idéalisé, le modèle ci-dessus. Expliquer pourquoi a-t-on $u_{n+2} = u_{n+1} + u_n$.

corrTD3-0011

Correction of the exercise 121

(1) Supposons que la formule soit vraie pour tous les termes jusqu'à $k+1$, et prouvons qu'elle est encore vraie pour le terme numéro $k+2$. Pour simplifier la notation, nous notons

$$A = \frac{1+\sqrt{5}}{2} \quad B = \frac{1-\sqrt{5}}{2}. \tag{90.96}$$

En utilisant la définition de u_{k+2} et en supposant que $u_{k+1} = \frac{1}{\sqrt{5}}(A^{k+1} - B^{k+1})$ et $u_k = \frac{1}{\sqrt{5}}(A^k - B^k)$, nous avons

$$\begin{aligned} u_{k+2} &= u_{k+1} + u_k \\ &= \frac{1}{\sqrt{5}}(A^{k+1} + A^k - B^{k+1} - B^k) \\ &= \frac{1}{\sqrt{5}}\big(A^k(A+1) - B^k(B+1)\big). \end{aligned} \tag{90.97}$$

EqFiboUnMOchaAB

Nous voudrions obtenir

$$u_{k+2} = \frac{1}{\sqrt{5}}(A^{k+2} - B^{k+2}). \tag{90.98}$$

En comparant avec l'expression (90.97), nous devons prouver que $A+1 = A^2$ et $B+1 = B^2$. Cela est vrai parce que

$$A + 1 = \frac{1+\sqrt{5}}{2} = \frac{3+\sqrt{5}}{2}, \tag{90.99}$$

mais

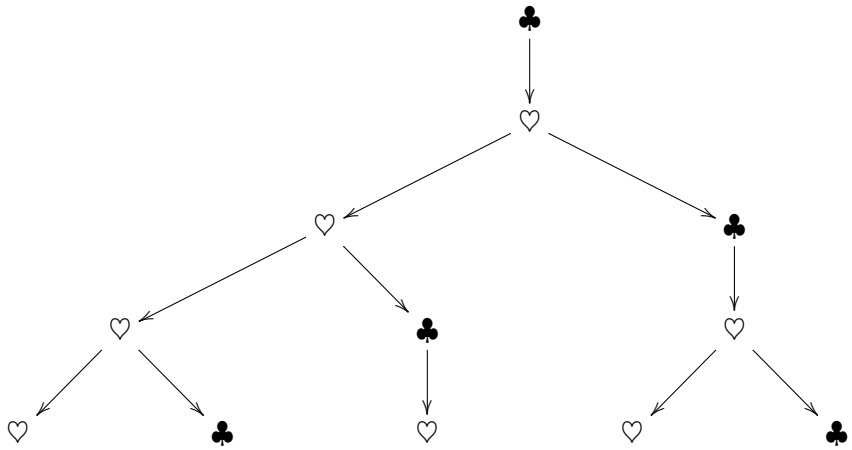
$$A^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{1+5+2\sqrt{5}}{4} = \frac{3+\sqrt{5}}{2}. \tag{90.100}$$

Il en va de même pour B : $B+1 = \frac{3-\sqrt{5}}{2}$, et

$$B^2 = \frac{1+5-2\sqrt{5}}{4} = \frac{6-2\sqrt{5}}{4} = \frac{3-\sqrt{5}}{2}. \tag{90.101}$$

(2) Faisons un petit schéma. Les ♡ désignent les couples en âge de se reproduire tandis que les ♣ désignent les couples qui viennent de naître.

(90.102)



À chaque nouvelle étape, les ♡ restent, mais chaque ♡ donne naissance à un ♣, tandis que les ♣ deviennent des ♡. Donc à l'étape n, le nombre de ♡ est égal au nombre de ♡ de l'étape $n-1$ plus le nombre de ♣ de l'étape $n-1$. Mais le nombre de ♣ de l'étape $n-1$ est égal au nombre de ♡ de l'étape $n-2$. Par conséquent $u_n = u_{n-1} + u_{n-2}$.

Exercise 122 exoTD3-0012

Modèle logistique.

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = 2u_n(1-bu_n) \quad \forall n \in \mathbb{N}_0 \\ u_0 = x, \end{cases} \tag{90.103}$$

où $b > 0$ et $x \geqslant 0$ sont des nombres réels.

(1) Montrer que si $x \in \{0, \frac{1}{b}\}$, alors $\lim_{n\to\infty} u_n = 0$.

(2) Montrer que si $x \in]0, \frac{1}{2b}]$, alors $\lim_{n\to\infty} u_n = \frac{1}{2b}$.

(3) Montrer que si $x \in [\frac{1}{2b}, \frac{1}{b}[$, alors $\lim_{n\to\infty} u_n = \frac{1}{2b}$.

corrTD3-0012

Correction of the exercise 122

Avant de nous lancer dans l'action, nous pouvons déterminer quelles sont les limites possibles. Le « candidats limites » sont les u tels que

$$u = 2u(1-bu). \tag{90.104}$$

La première solution est $u = 0$. Ensuite nous pouvons simplifier l'équation par u, ce qui donne l'équation $1 = 2 - 2bu$, et donc la solution $u = \frac{1}{2b}$. Nous retenons que les limites possibles de la suite (u_n) sont zéro et $\frac{1}{2b}$.

(1) Si $u_0 = 0$, alors la suite reste nulle tout le temps parce que u_n est en facteur dans la définition de u_{n+1}. Si par contre $u_0 = \frac{1}{b}$, alors $1 - bu_0 = 0$, et par conséquent $u_1 = 0$. Dès que $u_1 = 0$, tous les termes suivants de la suite seront nuls.

(2) Nous savons que les limites possibles (si la suite est convergente) sont 0 et $\frac{1}{2b}$. Dans le cas où u_0 est strictement entre 0 et $\frac{1}{2b}$, si nous voulons prouver que la limite existe et vaut $\frac{1}{2b}$.
La stratégie sera de prouver que la suite est croissante et bornée. En effet, si elle est croissante et borné, elle converge. Mais si le premier terme est plus grand que 0 et qu'elle est croissante, elle ne peut pas converger vers 0. Elle devra donc converger vers $\frac{1}{2b}$.

Bornée Montrons que si $u_k < \frac{1}{2b}$, alors $u_{k+1} < \frac{1}{2b}$. Demander

$$2u_k(1-bu_k) < \frac{1}{2b} \tag{90.105}$$

revient à demander

$$-2bu_k^2 + 2u_k - \frac{1}{2b} < 0. \tag{90.106}$$

Étudions donc la fonction $f(x) = -2bx^2 + 2x - \frac{1}{2b}$. C'est une fonction qui s'annule pour les valeurs suivantes de x :

$$\frac{-2 \pm \sqrt{4 - 4(-2b)\left(-\frac{1}{2b}\right)}}{-4b} = \frac{1}{2b}. \tag{90.107}$$

C'est une parabole (fonction du second degré) qui ne s'annule qu'une seule fois, et qui est donc toujours positive ou bien toujours négative. Dans notre cas, elle est toujours négative parce que $f(0) = -\frac{1}{2b} < 0$.
Cela est une très bonne nouvelle parce que nous avons démontré que *quelle que soit la valeur de* u_k, nous avons *toujours* $u_{k+1} < \frac{1}{2b}$. Ce qui nous intéresse pour l'instant est que la suite soit bornée par $\frac{1}{2b}$.

Bornée, alternative Calculons la différence $u_{n+1}-\frac{1}{2b}$ et montrons que cette différence est toujours négative. Nous avons

$$u_{n+1}-\frac{1}{2b}=2u_n(1-bu_n)-\frac{1}{2b} \tag{90.108a}$$
$$=-2bu_n^2+2u_n-\frac{1}{2b} \tag{90.108b}$$
$$=-\frac{1}{2b}\big(4b^2u_n^2-4bu_n+1\big) \tag{90.108c}$$
$$=-\frac{1}{2b}(2bu_n-1)^2, \tag{90.108d}$$

qui est effectivement toujours négatif. Par conséquent $u_{n+1}-\frac{1}{2b}<0$ et $u_{n+1}<\frac{1}{2b}$.

Croissante Calculons

$$\frac{u_{k+1}}{u_k}=2(1-bu_k)>2\left(1-b\frac{1}{2b}\right)=1. \tag{90.109}$$

Pour cette majoration, nous avons utilisé le fait que $1-bu_k>1-b\frac{1}{2b}$ parce que si $u_k<\frac{1}{2b}$, alors $1-u_k>1-\frac{1}{2b}$.

Conclusion Lorsque u_n est entre 0 et $\frac{1}{2b}$, nous avons montré qu'elle est croissante et bornée. Elle est donc convergente. Or les seules bornes possibles sont 0 et $\frac{1}{2b}$. La suite ne peut donc converger que vers $\frac{1}{2b}$.

(3) La bonne surprise est que si $u_k>\frac{1}{2b}$, alors $u_{k+1}<\frac{1}{2b}$, parce que nous avons déjà vu que de toutes façons la fonction $-2bu_n^2+2u_n$ ne retourne que des valeurs plus petites que $\frac{1}{2b}$. Donc si $u_0>\frac{1}{2b}$, il n'en reste pas moins que $u_1<\frac{1}{2b}$, et que par conséquent la suite converge vers $\frac{1}{2b}$ par le point précédent.

Exercise 123 exoTD3-0013

Modèle de Gompertz.

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases}u_{n+1}=u_n-u_n\ln(u_n) & \forall n\in\mathbb{N}_0\\ u_0=x.\end{cases} \tag{90.110}$$

Montrer que si $x\in\,]0,e[$, alors $\lim_{n\to\infty}u_n=1$.

corrTD3-0013

Correction of the exercise 123

Candidats limite Comme d'habitude, nous commençons par trouver les limites possibles en résolvant l'équation

$$u=u-u\ln(u). \tag{90.111}$$

Eqnumulnu

Notez que $u=0$ pourrait être une limite parce que la fonction $u-u\ln(u)$ n'est pas définie en $u=0$, donc la proposition 90.1 ne peut pas conclure. Quoi qu'il en soit, les autres limites possible se trouvent en simplifiant l'équation (90.111) par u. Nous tombons sur l'équation $1=1-\ln(u)$, c'est-à-dire $\ln(u)=0$ et donc $u=1$.

Deux limites possibles : 0 et 1.

Croissance En ce qui concerne la croissance de la suite, nous avons

$$\frac{u_{n+1}}{u_n}=1-\ln(u_n). \tag{90.112}$$

Ici nous voyons une différence entre $u_n\in\,]0,1[$ et $u_n\in\,]1,e[$.

(1) Si $u_n\in\,]0,1[$, alors $1-\ln(u_n)>1$ et la suite est croissante. En d'autres termes, $u_n<1$ implique $u_{n+1}>u_n$.

(2) Si $u_n\in\,]1,e[$, nous avons $\frac{u_{n+1}}{u_n}<1$ et la suite est donc décroissante. En d'autres termes, si $u_n>1$, alors $u_{n+1}<u_n$.

Bornes La fonction qui définit la récurrence est

$$f(x) = x - x\ln(x). \tag{90.113}$$

La dérivée vaut $f'(x) = -\ln(x)$.

(1) Nous commençons par considérer le cas où la suite prend au moins une valeur en dessous de 1.
Si $x < 1$, nous avons $f'(x) > 0$. Cela signifie que la fonction est croissante et que parmi les $x \in]0, 1[$, celui qui rend $f(x)$ la plus grande possible est $x = 1$. Donc si $u_n < 1$, nous avons

$$u_{n+1} = f(u_n) < f(1) = 1 \tag{90.114}$$

La conclusion est que si $u_n < 1$, alors $u_{n+1} < 1$. Dès que la suite passe par une valeur plus basse que 1, elle y reste. Nous avons par ailleurs montré que pour les $u_n < 1$, la suite était croissante. Dès que la suite prend une valeur plus petite que 1, elle est donc bornée et croissante et par conséquent convergente. Les seuls candidats étant 0 et 1, la suite converge vers 1.

(2) Traitons à présent du cas où la suite reste toujours au dessus de 1. Nous avons déjà vu que si $u_n > 1$, alors $u_{n+1} < u_n$.
Si $x > 1$, nous avons $f'(x) < 0$ et f est donc une fonction décroissante. Son minimum *parmi les x plus grands que* 1 est donc atteint pour $x = 1$. C'est-à-dire que

$$f(x) < f(1) = 1 \tag{90.115}$$

pour tout $x > 1$. En ce qui concerne notre suite, si $u_n > 1$, nous avons

$$u_{n+1} = f(u_n) < f(1) = 1. \tag{90.116}$$

Si la suite passe par une valeur supérieure à 1, elle retourne instantanément à une valeur inférieure à 1.

En conclusion, si $x \in]0, 1[$ nous avons $u_0 < 1$ et la suite reste constamment inférieure à 1. Dans ce cas elle converge vers 1.

Si par contre $x \in]1, e[$, nous avons $u_0 \in]1, 0[$ et par conséquent $u_1 < 1$. La suite converge donc également vers 1.

Exercise 124 exoTD3-0014

Modèle avec prédation et prélèvement.

Soit $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1} = a\frac{u_n^2}{u_n^2+b^2} - Eu_n \quad \forall n \in \mathbb{N} \\ u_0 = x, \end{cases} \tag{90.117}$$

où $a > 0$, b et $E > 0$ sont des réels.

(1) Chercher les limites possibles d'une telle suite.

(2) Si $E = 0$ et $a^2 > 4b^2$, en utilisant la question précédente et le fait que la fonction $t \mapsto \frac{at^2}{t^2+b^2}$ est croissante, montrer que pour $0 < x < \frac{a-\sqrt{a^2-4b^2}}{2}$, alors

$$0 < a\frac{u_n}{u_n^2 + b^2} < 1 \tag{90.118}$$

pour tout $n \in \mathbb{N}$.

(3) Si $E = 0$ et $a^2 > 4b^2$, montrer que si $0 < x < \frac{a-\sqrt{a^2-4b^2}}{2}$ alors la suite (u_n) converge vers 0.

(4) Si $a > 2$ et $b = 1$, montrer que si $E > (a-2)/2$, alors la suite est décroissante et si $x > 0$, elle converge vers 0.

(5) Quelle est la morale de cette dernière question ?

corrTD3-0014

Correction of the exercise 124

(1) La suite $(u_n)_{n\in\mathbb{N}}$ est définie par une formule de récurrence du type $u_{n+1} = g(u_n)$, donc si sa limite est un nombre ce nombre sera un point fixe de la fonction qui apparaît dans la définition de la suite.
La fonction qu'on doit considérer est alors

$$g(x) = a\frac{x^2}{x^2+b^2} - Ex.$$

Les points fixes de g satisfont l'équation

$$\ell = a\frac{\ell^2}{\ell^2+b^2} - E\ell,$$

qui est équivalente à

$$(E+1)\ell(\ell^2+b^2) - a\ell^2 = 0,$$

et encore

$$\ell\left((E+1)(\ell^2+b^2) - a\ell\right) = 0.$$

Le produit de deux facteurs est zéro si et seulement si au moins un entre les deux facteurs est nul. Les solutions sont alors 0, ℓ_1 et ℓ_2, où ℓ_1 et ℓ_2 sont les racines du polynôme $(E+1)(\ell^2+b^2) - a\ell$.

$$\ell_1 = \frac{a+\sqrt{a^2-4(E+1)^2b^2}}{2(E+1)}, \qquad \ell_2 = \frac{a-\sqrt{a^2-4(E+1)^2b^2}}{2(E+1)}.$$

Itemtzziqtrois

(2) Supposons maintenant que $E = 0$ et que $a^2 > 4b^2$. On aura alors

$$\ell_1 = \frac{a+\sqrt{a^2-4b^2}}{2}, \qquad \ell_2 = \frac{a-\sqrt{a^2-4b^2}}{2}.$$

Notez que la deuxième de ces hypothèses veut dire simplement que les racines ℓ_1 et ℓ_2 sont bien deux nombres réels distincts. En fait, il est facile de voir que $0 < \ell_2 < \ell_1$.
On a

$$\frac{au_n}{u_n^2+b^2} < 1 \quad \text{si et seulement si } au_n < u_n^2 + b^2,$$

ce qui peut s'écrire aussi comme

$$u_n^2 - au_n + b^2 > 0.$$

Cette dernière condition est satisfaite pour $u_n < \ell_2$ ou $u_n > \ell_1$. Il nous faut alors démontrer que si le premier terme de la suite u_0 est entre $[0, \ell_2]$ touts les termes suivantes sont contenus dans le même intervalle.
Notons d'abord que si $0 < u_0 < \ell_2$ alors $0 < \dfrac{au_0}{u_0^2+b^2} < 1$ et par conséquent $0 < u_1 < u_0$.
Supposons maintenant que les termes u_n soient dans l'intervalle $[0, \ell_2]$ pour tous les indices n entre 0 et k alors on aura

$$0 < u_{k+1} = \frac{au_k}{u_k^2+b^2}u_k < u_k < \ell_2.$$

Méthode alternative : la condition que $t \mapsto \frac{at^2}{t^2+b^2}$ est croissante implique que la suite $(u_n)_{n\in\mathbb{N}}$ est monotone. Autrement dit, cela implique que soit

$$\frac{au_n}{u_n^2+b^2} < 1 \quad \text{pour tout } n,$$

soit

$$\frac{au_n}{u_n^2+b^2} > 1 \quad \text{pour tout } n,$$

Il nous faut simplement vérifier laquelle entre les deux inégalités est vérifiée lorsque $n = 0$. Notez que l'égalité est vérifié si et seulement si $u_n = \ell_1$ ou $u_n = \ell_2$. Dans ce deux cas la suite devient constante.

(3) Les candidats limites de la suite se trouvent en résolvant l'équation

$$x = \frac{ax^2}{x^2+b^2}. \tag{90.119}$$

La première solution est $x = 0$. Pour trouver les autres solutions, nous simplifions l'équation par x pour trouver $ax = x^2 + b^2$ ou encore $x^2 - ax + b^2$ dont les solutions sont

$$\frac{a \pm \sqrt{a^2-4b^2}}{2}. \tag{90.120}$$

Montrons que le seul candidat possible est zéro. D'abord par la question (2) nous savons que

$$\frac{u_{n+1}}{u_n} = \frac{au_n}{u_n^2+b^2} < 1. \tag{90.121}$$

La suite est donc décroissante. D'autre part à part zéro, le plus petit candidat limite est

$$\frac{a-\sqrt{a^2-4b^2}}{2}, \tag{90.122}$$

mais par hypothèse, nous avons

$$0 < x < \frac{a-\sqrt{a^2-4b^2}}{2}, \tag{90.123}$$

ce qui signifie que la suite commence déjà en dessous de ce candidat limite. Une suite décroissante ne peut pas converger vers une limite plus grande que son premier terme. Donc le seul candidat limite est zéro.

(4) Dans les nouvelles conditions $a > 2$, $b = 1$ et $E > (a-2)/2$ nous avons

$$u_{n+1} = a\frac{u_n^2}{u_n^2+1} - Eu_n = \left(a\frac{u_n}{u_n^2+1} - E\right)u_n. \tag{90.124}$$

Décroissante Il nous suffit de prouver que pour tout indice n dans $\mathbb{N}$ la quantité

$$a\frac{u_n}{u_n^2+1} - E$$

est inférieure à 1. L'hypothèse sur E nous permet d'écrire

$$a\frac{u_n}{u_n^2+1} - E < a\frac{u_n}{u_n^2+1} - \frac{(a-2)}{2}.$$

En mettant au même dénominateur, nous avons

$$a\frac{u_n}{u_n^2+1} - E < \frac{2au_n - au_n^2 - a + 2u_n^2 + 2}{2(u_n^2+1)} = -a\frac{(u_n-1)^2}{2(u_n^2+1)} + 1 < 1. \tag{90.125}$$

bornée Nous montrons maintenant que la suite est bornée inférieurement par zéro.

PROBLÈME : avec $E = 1000$, $a = 3$, $u_0 = 1$, c'est un contre-exemple !!

90.7 Étude de fonctions, première partie

Exercise 125 ⊙⊙ | exoTD2-1

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to 1} \dfrac{x^2+x-2}{x^2+2x-3}$;

(2) $\lim_{x\to 0} \dfrac{\sqrt{x+1}-1}{x}$;

(3) $\lim_{x\to 0} \dfrac{\sin x}{x}$;

(4) $\lim_{x\to 0} \dfrac{\ln x}{\sqrt{x}}$;

(5) $\lim_{x\to 0} \dfrac{e^x}{1+x^5}$;

(6) $\lim_{x\to 0} \dfrac{e^{x^2}-1}{\sin^2(x)}$;

(7) $\lim_{x\to 0} \left(\dfrac{1}{\ln(x)} - \dfrac{1}{x-1}\right)$;

(8) $\lim_{x\to 0} \dfrac{1-\cos x}{x^2}$;

(9) $\lim_{x\to 0} \left(1+\dfrac{1}{x^2}\right)^x$.

ExobgEPck

(10) $\lim_{x\to\infty} \left(1+\dfrac{1}{x^2}\right)^x$.

corrTD2-1

Correction of the exercise 125

(1) En utilisant sage,

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x)=(x**2+x-2)/(x**2+2*x-3)
sage: f.limit(x=1)
x |--> 3/4
```

Si nous voulons la calculer à la main, nous commençons par remarquer qu'en remplaçant x par 1 nous tombons sur l'indétermination 0/0. L'idée est de factoriser le numérateur et le dénominateur. Pour cela nous cherchons les racines. Pour le numérateur, les racines sont données par

$$x_\pm = \frac{-1\pm\sqrt{1-4\cdot(-2)}}{2} = \frac{-1\pm 3}{2}, \tag{90.126}$$

c'est-à-dire -2 et 1. Le numérateur se factorise donc en $(x+2)(x-1)$. De la même façon le dénominateur se factorise en $(x-1)(x+3)$. Nous avons donc

$$\frac{x^2+x-2}{x^2+2x-3} = \frac{(x-1)(x+2)}{(x-1)(x+3)} = \frac{x+2}{x+3}. \tag{90.127}$$

Maintenant en remplaçant x par 1 nous trouvons le résultat 3/4.

(2) Nous utilisons la technique du binôme conjugué :

$$\frac{\sqrt{x+1}-1}{x}=\frac{\big(\sqrt{x+1}-1\big)\big(\sqrt{x+1}+1\big)}{x\big(\sqrt{x+1}+1\big)}=\frac{1}{\sqrt{x+1}+1}. \tag{90.128}$$

La limite lorsque $x\to 0$ vaut alors $1/2$.

(3) Ceci est une limite de base qu'il faut connaitre : la limite est 1.

(4) En remplaçant, nous obtenons $\frac{-\infty}{0}$, et donc la limite est $-\infty$ sans indétermination.

(5) Ici il n'y a pas d'indéterminations : en remplaçant on trouve $\frac{1}{1+0}=1$.

(6) Lorsqu'il y a un sinus tout seul, un truc courant est de diviser et multiplier par x. Ici nous avons $\sin^2(x)$, donc nous allons multiplier et diviser par x^2 :

$$\frac{x^2(e^{x^2}-1)}{x^2\sin^2(x)}. \tag{90.129}$$

Lorsqu'on prend la limite pour $x\to 0$ nous savons que $\frac{x^2}{\sin^2(x)}$ tend vers 1. Nous nous retrouvons donc à devoir calculer

$$\lim_{x\to 0}\frac{e^{x^2}-1}{x^2}=\lim_{y\to 0}\frac{e^y-1}{y}, \tag{90.130}$$

où nous avons posé $y=x^2$. La dernière limite est une des limites remarquables. Il est également possible de la faire en utilisant la règle de l'Hospital :

$$\lim_{x\to 0}\frac{e^{x^2}-1}{x^2}=\lim_{x\to 0}\frac{2xe^{x^2}}{2x}=\lim_{x\to 0}e^{x^2}=1. \tag{90.131}$$

(7) Ici il n'y a pas tellement d'indéterminations non plus : lorsque $x\to 0$, le logarithme tend vers $-\infty$, et le premier terme tende vers zéro. En même temps, le second terme tend vers $\frac{1}{0-1}=-1$, donc la différence tend vers 1.

(8) Nous multiplions et divisons par le binôme conjugué $1+\cos(x)$, et nous trouvons

$$\frac{1-\cos(x)}{x^2}=\frac{\big(1-\cos(x)\big)\big(1+\cos(x)\big)}{x^2\big(1+\cos(x)\big)}=\frac{1-\cos^2(x)}{x^2\big(1+\cos(x)\big)}=\frac{\sin^2(x)}{x^2}\frac{1}{1+\cos(x)}. \tag{90.132}$$

La limite du premier facteur est une limite remarquable qui vaut 1. Le second facteur tend vers $\frac{1}{2}$. La limite du tout vaut par conséquent $\frac{1}{2}$.

(9) Nous allons calculer le logarithme de la limite à calculer :

$$\spadesuit=\ln\left[\left(1+\frac{1}{x^2}\right)^x\right]=x\ln\left(1+\frac{1}{x^2}\right). \tag{90.133}$$

EqBtkUpe

Maintenant en posant $t=\frac{1}{x^2}$, nous avons

$$\spadesuit=\frac{1}{\sqrt{t}}\ln(1+t), \tag{90.134}$$

dont nous devons prendre la limite lorsque $t\to\infty$. Cela est une des limites remarquables, qui vaut zéro [2].

Donc la limite que nous cherchons est

$$\lim_{x\to 0}e^{\spadesuit}=e^0=1. \tag{90.135}$$

La limite (90.133) peut également être calculée en utilisant la règle de l'Hospital :

$$x\ln\left(1+\frac{1}{x^2}\right)=\frac{\ln\left(1+\frac{1}{x^2}\right)}{1/x}, \tag{90.136}$$

2. Le logatithme tend plus vite vers ∞ que les puissances.

en prenant la dérivée du numérateur et du dénominateur nous devons calculer

$$\lim_{x\to 0}\frac{\frac{-2}{\left(1+\frac{1}{x^2}\right)x^2}}{-1/x^2}=\lim_{x\to 0}\frac{2}{1+\frac{1}{x^2}}=0. \tag{90.137}$$

(10) Nous refaisons la même manipulation que pour le point (10). Maintenant nous devons calculer

$$\lim_{t\to 0}\frac{\ln(1+t)}{\sqrt{t}}. \tag{90.138}$$

Cette limite peut être déterminée à l'aide de la règle de l'Hospital. Les dérivées du numérateur et du dénominateur sont respectivement $\frac{1}{1+t}$ et $\frac{1}{2\sqrt{t}}$, par conséquent

$$\lim_{t\to 0}\frac{\ln(1+t)}{\sqrt{t}}=\lim_{t\to 0}\frac{2\sqrt{t}}{1+t}=0. \tag{90.139}$$

La limite demandée est donc $e^0=1$.

Exercise 126 ⊙⊙ exoTD2A-2 Il faut connaître les valeurs des limites suivantes :

Sinus x sur x : $\lim_{x\to 0}\frac{\sin x}{x}=1$;

Nombre de Néper : $\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^x=e$.

Un adage fort utile est « l'exponentielle va plus vite et le logarithme va moins vite ». Cela se traduit par les limites suivantes :

(1) $\lim_{x\to 0} x\ln(x)=0$

(2) $\lim_{x\to 0}\sqrt{x}\ln(x)=0$

(3) $\lim_{x\to\infty}\frac{\ln(x)}{x}=0$

(4) $\lim_{x\to\infty}\frac{\ln(x)}{\sqrt{x}}=0$

(5) $\lim_{x\to\infty}\frac{e^x}{x}=\infty$

(6) $\lim_{x\to-\infty} xe^x=0$

En utilisant ces limites (dites *remarquables*, en raison de leur importance) nous pouvons calculer des autres limites importantes :

(1) $\lim_{x\to 0}\frac{1-\cos x}{x^2}=\frac{1}{2}$;

(2) $\lim_{x\to\pm\infty}\left(1+\frac{a}{x}\right)^x=e^a$;

(3) $\lim_{x\to 0}(1+x)^{\frac{1}{x}}=e$;

(4) $\lim_{x\to 0}\frac{\ln(1+x)}{x}=1$;

(5) $\lim_{x\to 0}\frac{e^x-1}{x}=1$.

Exercise 127 ⊖⊖ exoTD2B_1

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to+\infty} x^2$;

(2) $\lim_{x\to+\infty} x^2-x^4$;

(3) $\lim_{x\to+\infty}\frac{x^2-2}{x+1}$;

(4) $\lim_{x\to+\infty}\frac{x+1}{x^3}$;

(5) $\lim_{x\to+\infty}\sin x$;

(6) $\lim_{x\to-\infty} 2^x$;

(7) $\lim_{x\to+\infty} \dfrac{x^4+x-1}{2x^4-x}$;

(8) $\lim_{x\to+\infty} \dfrac{x}{\sqrt{x^2+1}}$.

Exercise 128 ⊙⊙ exoTD2_2

Déterminer dans chaque cas le domaine de définition, la périodicité et/ou les symétries éventuelles et les limites aux extrêmes du domaine. Calculer ensuite la dérivée de la fonction, la où elle est définie.

(1) $f_1(x) = \ln(x^4)$;

(2) $f_2(x) = \ln^4(x)$;

(3) $f_3(x) = \ln(4x)$;

(4) $f_4(x) = (x^2-2x)e^x$;

(5) $f_5(x) = \sqrt{x^2-2}$;

(6) $f_6(x) = \dfrac{\sin x+\cos x}{\sin x-\cos x}$;

(7) $f_7(x) = e^{x^3/2}$;

(8) $f_8(x) = \sqrt{e^{x^3}}$;

(9) $f_9(x) = e^{\sqrt{3+\cos x}}$;

(10) $f_{10}(x) = \dfrac{x^2+1}{(x+1)^3}$;

90.8 Étude de fonctions, suite

Exercise 129 exoTD4-0001

Déterminer dans les cas suivants les intervalles de $\mathbb{R}$ sur lesquels la fonction f est bijective et donner dans chaque cas l'expression de la bijection réciproque.

(1) $f(x) = \frac{1}{1-x}$,

ItemexoTD1ii

(2) $f(x) = x^2-2x+1$.

corrTD4-0001

Correction of the exercise 129

(1) La fonction $f(x) = \frac{1}{1-x}$ est représentée sur la figure 90.1. Elle a deux intervalles de monotone (dé)croissance : $]-\infty,1[$ et $]1,\infty[$. Elle va donc être bijective sur ces deux intervalles.
Sur l'intervalle $]-\infty,1[$, la fonction f prend toutes les valeurs de $]0,\infty[$, tandis que sur l'intervalle $]1,\infty[$, elle prend les valeurs $]-\infty,0[$; en formules,

$$f\big(]-\infty,1[\big) =]0,\infty[\tag{90.140a}$$
$$f\big(]1,\infty[\big) =]-\infty,0[. \tag{90.140b}$$

Trouver l'application inverse revient à trouver pour quel x nous avons $f(y) = y$. L'équation à résoudre est

$$y = \frac{1}{1-x}, \tag{90.141}$$

à résoudre par rapport à x pour un y donné. La solution est

$$x = \frac{y-1}{y}. \tag{90.142}$$

Donc $f^{-1}(y) = \frac{y-1}{y}$.

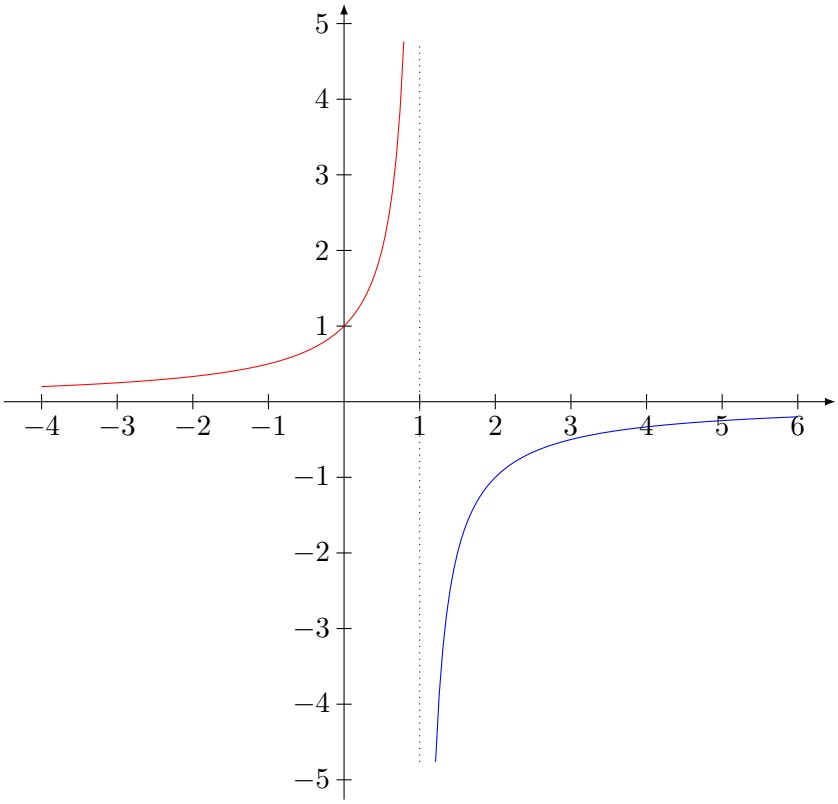


FIGURE 90.1: Le graphe de la fonction $x \mapsto \frac{1}{1-x}$. LabelFigGrapheunsurunmoinsx

Une autre façon de voir cela est d'écrire la définition de la fonction réciproque :

$$f\big(f^{-1}(y)\big) = y, \tag{90.143}$$

et donc

$$\frac{1}{1 - f^{-1}(y)} = y, \tag{90.144}$$

et résoudre cette équation par rapport à $f^{-1}(y)$ en fonction de y.

(2) Cette fois, le graphe est plus difficile à tracer, donc nous allons calculer la dérivée pour savoir sur quel(s) intervalle(s) la fonction est monotone. Nous avons

$$f(x) = x^2 - 2x + 1 \tag{90.145a}$$
$$f'(x) = 2x - 2. \tag{90.145b}$$

La dérivée est donc positive pour $x > 1$ et négative sur $x < 1$. Elle va donc être une bijection sur $]-\infty, 1]$ et sur $[1, \infty[$. Il reste à voir sur quelle est l'image de ces intervalles par f. Étant donné que $f(1) = 0$, nous avons

$$f\big([1, \infty[\big) = [0, \infty[\tag{90.146a}$$
$$f\big(]-\infty, 1]\big) = [0, \infty[. \tag{90.146b}$$

Pour information, nous avons tracé la fonction à la figure 90.2.

Cette fois pour trouver la fonction inverse, il faut résoudre l'équation $f(x) = y$, c'est-à-dire

$$\begin{aligned} x^2 - 2x + 1 &= y \\ x^2 - 2x + (1 - y) &= 0. \end{aligned} \tag{90.147}$$

Cette équation est à résoudre par rapport à x en fonction de y. La solution est

$$x = 1 \pm \sqrt{y}. \tag{90.148}$$

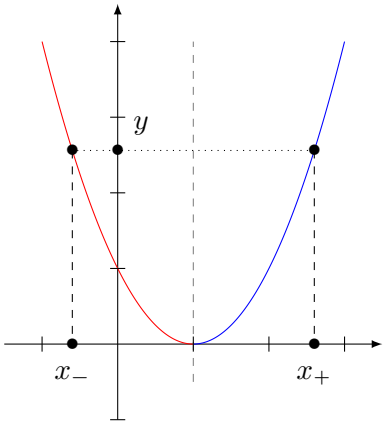


FIGURE 90.2: La fonction de l'exerice 129.(2). Remarquer la symétrie autour du sommet, comme toute fonction du second degré. Pour un y donné, il y a *deux* x sur lesquels fa fonction vaut y. LaTeXPgExoCUd

La question est de savoir si il faut choisir le signe plus ou le signe moins.

Si nous choisissons le signe moins, nous obtenons $x \leqslant 1$, et donc nous tombons dans l'intervalle de bijection de gauche, tandis qu'avec le signe plus, nous avons $x \geqslant 1$, et donc nous sommes dans l'intervalle de droite.

La fonction inverse de

$$\begin{aligned} f\colon [1,\infty[&\to [0,\infty[\\ x &\mapsto x^2-2x+1 \end{aligned} \tag{90.149}$$

est $f^{-1}(y) = 1-\sqrt{y}$, tandis que la fonction inverse de

$$\begin{aligned} f\colon]-\infty,1] &\to [0,\infty[\\ x &\mapsto x^2-2x+1 \end{aligned} \tag{90.150}$$

est $f^{-1}(y) = 1+\sqrt{y}$.

Exercise 130 exoTD4-0002

Les fonctions trigonométriques réciproques.

(1) Montrer que la fonction sinus est une bijection croissante de $[-\frac{\pi}{2},\frac{\pi}{2}]$ dans $[-1,1]$.

(2) On note arcsin (se prononce « arc sinus ») sa bijection réciproque : $\arcsin\colon [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$. Montrer que arcsin est dérivable sur $]-1,1[$ avec

$$\big(\arcsin(x)\big)' = \frac{1}{\sqrt{1-x^2}}. \tag{90.151}$$

Conseil : utiliser la formule $\cos^2(t)+\sin^2(t)=1$ pour tout $t\in\mathbb{R}$.

(3) Montrer que la fonction cosinus est une bijection décroissante de $[0,\pi]$ dans $[-1,1]$. On note arccos (prononcer « arc cosinus ») sa bijection réciproque

$$\arccos\colon [-1,1] \to [0,\pi]. \tag{90.152}$$

Montrer que arccos est dérivable sur $]-1,1[$ avec

$$\big(\arccos(x)\big)' = -\frac{1}{\sqrt{1-x^2}}. \tag{90.153}$$

(4) Montrer que la fonction tan est une bijection croissante de $]-\pi/2,\pi/2[$ dans $\mathbb{R}$. On note arctan (prononcer « arc tangente ») sa bijection réciproque. Montrer que arctan est dérivable sur $\mathbb{R}$ et que

$$\big(\arctan(x)\big)' = \frac{1}{1+x^2}. \tag{90.154}$$

corrTD4-0002

Correction of the exercise 130

(1) La dérivée est cosinus. Or $\cos(x)$ est *strictement* positive $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$. La fonction sinus est donc strictement croissante sur cet intervalle et par conséquent elle y est bijective. Nous avons donc prouvé que

$$\sin\colon]-\frac{\pi}{2}, \frac{\pi}{2}[\to]-1, 1[\tag{90.155}$$

est une bijection. Quid des derniers deux points ? $\cos(-\pi/2) = -1$ et $\cos(\pi/2) = 1$, et ce sont les seuls points sur lesquels les valeurs 1 et -1 sont prises. Nous pouvons donc les ajouter au domaine de bijectivité.

(2) Pour calculer la dérivée d'une fonction inverse, le truc est dériver la définition en utilisant la règle de dérivation des fonctions composées. C'est-à-dire, pour calculer la dérivée de f^{-1}, nous écrivons

$$f\big(f^{-1}(y)\big) = y, \tag{90.156}$$

et nous dérivons par rapport à y, ce qui donne

$$f'\big(f^{-1}(y)\big)(f^{-1})'(y) = 1. \tag{90.157}$$

De là, si nous savons la dérivée de f, nous pouvons trouver la dérivée de f^{-1} en isolant $f^{-1}(y)$.
En ce qui concerne arcsin, nous écrivons

$$\sin\big(\arcsin(x)\big) = x, \tag{90.158}$$

et nous dérivons par rapport à x :

$$\cos\big(\arcsin(x)\big)\arcsin'(x) = 1 \tag{90.159}$$

eqcosasnpu

Oui, mais pour tout $\heartsuit$ entre $-\pi/2$ et $\pi/2$, nous avons $\cos(\heartsuit) = \sqrt{1-\sin^2(\heartsuit)}$, donc [3]

$$\cos\big(\underbrace{\arcsin(x)}_{\heartsuit}\big) = \sqrt{1-\sin^2\big(\underbrace{\arcsin(x)}_{\heartsuit}\big)}. \tag{90.160}$$

En sachant que $\sin^2(\arcsin(x)) = x^2$, nous avons donc

$$\cos\big(\arcsin(x)\big) = \sqrt{1-x^2}. \tag{90.161}$$

En reportant cela dans l'équation (90.159), nous trouvons

$$\sqrt{1-x^2}\arcsin'(x) = 1, \tag{90.162}$$

et par conséquent

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}. \tag{90.163}$$

(3) En ce qui concerne la tangente, nous devons savoir la formule

$$\tan'(x) = \frac{1}{\cos^2(x)}, \tag{90.164}$$

et alors nous avons successivement, en dérivant et en isolant : Eqsubatantancos

$$\tan\big(\arctan(x)\big) = x \tag{90.165a}$$
$$\tan'\big(\arctan(x)\big)\arctan'(x) = 1 \tag{90.165b}$$
$$\arctan'(x) = \cos^2\big(\arctan(x)\big). \tag{90.165c}$$

EqsubatantancosddQZLe

Il faut maintenant trouver ce que vaut $\cos\big(\arctan(x)\big)$.

3. Si $\heartsuit$ n'est pas entre $-\pi/2$ et $\pi/2$, le cosinus peut être négatif, et la formule serait fausse.

Méthode conseillée Nous utilisons la formule $\cos^2(\heartsuit)+\sin^2(\heartsuit)=1$ avec $\heartsuit=\arctan(x)$:

$$\cos^2\big(\arctan(x)\big)+\sin^2\big(\arctan(x)\big)=1. \tag{90.166}$$

Nous divisons par $\cos^2\big(\arctan(x)\big)$ pour faire apparaitre la tangente :

$$\begin{aligned} 1+\frac{\sin^2\big(\arctan(x)\big)}{\cos^2\big(\arctan(x)\big)} &= \frac{1}{\cos^2\big(\arctan(x)\big)}\\ 1+\tan\big(\arctan(x)\big)^2 &= \frac{1}{\cos^2(\arctan(x))}\\ 1+x^2 &= \frac{1}{\cos^2\big(\arctan(x)\big)}. \end{aligned} \tag{90.167}$$

Méthode alternative Afin d'exprimer $\cos\big(\arctan(x)\big)$, il serait bon d'exprimer $\cos(x)$ en termes de $\tan(x)$, de façon que $\cos\big(\arctan(x)\big)$ s'exprime en termes de $\tan\big(\arctan(x)\big)=x$. Pour cela, nous partons de la définition de la tangente :

$$\tan(x)=\frac{\sin(x)}{\cos(x)}=\frac{\sqrt{1-\cos^2(x)}}{\cos(x)}. \tag{90.168}$$

Par conséquent,

$$\tan^2(x)=\frac{1-\cos^2(x)}{\sin(x)}, \tag{90.169}$$

et donc

$$\cos^2(x)\big(\tan^2(x)+1\big)=1, \tag{90.170}$$

et au final,

$$\cos^2(x)=\frac{1}{\tan^2(x)+1}. \tag{90.171}$$

Nous retournons maintenant à notre dernière expression ((3)) :

$$\cos^2\big(\arctan(x)\big)=\frac{1}{\tan^2\big(\arctan(x)\big)+1}=\frac{1}{x^2+1}. \tag{90.172}$$

Quelle que soit la méthode choisie, nous avons

$$\cos^2\big(\arctan(x)\big)=\frac{1}{1+x^2}, \tag{90.173}$$

et nous concluons en reprenant la formule (90.165c) :

$$\arctan'(x)=\frac{1}{1+x^2}. \tag{90.174}$$

Exercise 131 exoTD4-0003

On considère les fonctions f_1 et f_2 définies par

$$\begin{aligned} f_1&: x\mapsto \arctan\frac{x}{\sqrt{1-x^2}}\\ f_2&: x\mapsto \arcsin(x). \end{aligned} \tag{90.175}$$

(1) Déterminer les ensembles de définitions de ces deux fonctions et calculer leurs dérivées.

(2) En déduire une relation entre f_1 et f_2.

corrTD4-0003

Correction of the exercise 131

En ce qui concerne les domaines, pour la première nous devons avoir $x \in]-1,1[$ à cause de la racine et du dénominateur. Étant donné que la fonction tangente peut prendre n'importe quelle valeur, il n'y a pas de conditions sur la fonction arctan.

Pour le domaine de f_2, c'est $x \in [-1,1]$ parce que ce sont les seules valeurs possibles pour la fonction sinus.

En utilisant les règles de dérivation composées et les résultats de l'exercice 130, nous voyons que $f_1'(x) = f_2'(x) = \frac{1}{\sqrt{1-x^2}}$.

Les deux fonctions ont la même dérivée, de plus, elles ont la même valeur en 0 parce que

$$f_1(0) = \arctan(0) = 0 \tag{90.176}$$

tandis que

$$\arcsin(0) = 0 \tag{90.177}$$

parce que c'est en zéro que le sinus vaut zéro. Ce sont deux fonctions qui ont le même domaine (à part les points extrêmes) et qui sont égales en un point. Elles sont donc égales sur leur domaine commun, c'est-à-dire sur $]-1,1[$.

Exercise 132 exoTD4-0004

On considère maintenant la fonction

$$g(x) = \arcsin \frac{x}{\sqrt{1+x^2}}. \tag{90.178}$$

Simplifier l'expression de g.

Conseil : utiliser les résultats de l'exercice 131. corrTD4-0004

Correction of the exercise 132

<+CorrTD4-0004+>

Exercise 133 exoTD4-0005

<+ExoTD4-0005+> corrTD4-0005

Correction of the exercise 133

<+CorrTD4-0005+>

90.9 Intégration

Exercise 134 exoTD5-00001

Calculer les intégrales suivantes

(1) $\int_0^1 x^5 + 3x^2 + 3\,dx$;

(2) $\int_1^2 x^{1/3} + 4x^{1/2}\,dx$;

(3) $\int_1^2 x^{1/3} + 4x^{1/2}\,dx$;

(4) $\int_{-\pi}^0 \cos(x)\,dx$;

(5) $\int_{-2}^1 e^x\,dx$;

(6) $\int_0^1 \frac{1}{\sqrt{x}}\,dx$.

corrTD5-00001

Correction of the exercise 134

Il suffit d'appliquer la formule (14.911) et d'utiliser la linéarité de l'intégrale et le tableau des primitives des fonctions fondamentales.

(1) $\int_0^1 x^5+3x^2+3\,dx = \left[\frac{1}{6}x^6+x^3+3x\right]_0^1 = \frac{1}{6}+4 = \frac{25}{6}$;

(2) $\int_1^2 x^{1/3}+4x^{1/2}\,dx = \left[\frac{3}{4}x^{\frac{4}{3}}+4\frac{2}{3}x^{\frac{3}{2}}\right]_1^2 = \frac{3}{2}2^{\frac{1}{3}}+\frac{16}{3}2^{\frac{1}{2}}-\frac{41}{12}$;

(3) $\int_{-\pi}^0 \cos(x)\,dx = [\sin(x)]_{-\pi}^0 = 0$;

(4) $\int_{-2}^1 e^x\,dx = [e^x]_{-2}^1 = e-e^{-2}$;

(5) $\int_0^1 \frac{1}{\sqrt{x}}\,dx = \int_0^1 x^{-1/2}\,dx = \left[2x^{\frac{1}{2}}\right]_0^1 = 2.$

Exercise 135 exoTD5-00002

Calculer les intégrales suivantes par la méthode du changement de variable

(1) $\int_0^{\pi/3} \sin(3x)\,dx$;

(2) $\int_{-13}^{-11} e^{4x}\,dx$;

(3) $\int_1^2 \frac{1}{x+1}\,dx$;

(4) $\int_{-\pi/2}^0 \sin^2(x)\cos(x)\,dx$. Conseil : le meilleur changement de variable dans ce cas est $y = \sin(x)$;

(5) $\int_1^{\pi} \frac{4x^3}{x^4+5}\,dx$.

corrTD5-00002

Correction of the exercise 135

(1) $I_1 = \int_1^{\pi} \frac{4x^3}{x^4+5}\,dx = \left[\ln(x^4+5)\right]_1^{\pi} = \ln\left(\frac{\pi^4+5}{6}\right)$.

(2) On utilise le changement de variable $u = 4x+1$. On a alors $du = 4dx$, les bornes du domaine d'intégration deviennent $u(0) = 1$ et $u(1) = 5$ et notre intégrale s'écrit de la forme suivante

$$I_2 = \int_0^1 \frac{1}{(4x+1)^4}\,dx = \int_1^5 \frac{1}{4u^4}\,du = \left[-\frac{1}{12u^3}\right]_1^5 = -\frac{1}{12\times 125}+\frac{1}{12} = \frac{31}{375}.$$

(3) Ici on vise à écrire la fonction à intégrer comme $\frac{1}{u^2+1}$, pour avoir $\arctan(u)+C$ comme primitive. Pour ainsi faire, nous devons utiliser le changement de variable $u = \frac{x}{\sqrt{5}}$. On a alors $du = \frac{1}{\sqrt{5}}dx$, les bornes du domaine d'intégration deviennent $u(0) = 0$ et $u(\sqrt{5}) = 1$

$$I_3 = \int_0^{\sqrt{5}} \frac{1}{x^2+5}\,dx = \int_0^1 \frac{\sqrt{5}}{5(u^2+1)}\,du = \frac{1}{\sqrt{5}}[\arctan(u)]_0^1 = \frac{\pi}{4\sqrt{5}}.$$

(2) On utilise le changement de variable $u = \sin(x)$, qui comporte $du = \cos(x)\,dx$, $u(-\pi/2) = -1$ et $u(0) = 0$. Nous avons alors que $I_4 = \int_{-1}^0 u^2\,du = \left[\frac{u^3}{3}\right]_{-1}^0 = \frac{1}{3}$.

(5) Le bon changement de variable nous est donné par l'énoncé. On a $du = e^x dx$, les bornes d'intégration deviennent $u(0) = e^0 = 1$ et $u(1) = e$ et notre intégrale sera

$$I_5 = \int_0^1 \frac{e^x}{e^{2x}+4}\,dx = \int_1^e \frac{1}{u^2+4}\,du.$$

Il faut maintenant travailler comme dans le point précédent de cet exercice. Soit $t = \frac{u}{2}$ alors

$$I_5 = \int_{1/2}^{e/2} \frac{2}{4(t^2+1)}\,dt = \frac{1}{2}\left[\arctan\left(\frac{e}{2}\right) - \arctan\left(\frac{1}{2}\right)\right].$$

(6) Le bon changement de variable nous est donné par l'énoncé. On a $(1+\tan^2(u))du = dx$, les bornes d'intégration deviennent $\arctan(0) = 0$ et $\arctan(1) = \pi/4$ et notre intégrale sera

$$I_6 = \int_0^1 \frac{1}{(1+x^2)^{\frac{3}{2}}}\,dx = \int_0^{\pi/4} \frac{1+\tan^2(u)}{(1+\tan^2(u))^{\frac{3}{2}}}\,du = \int_0^{\pi/4} \frac{1}{(1+\tan^2(u))^{\frac{1}{2}}}\,du.$$

On observe alors que

$$\frac{1}{(1+\tan^2(u))^{\frac{1}{2}}} = \cos(u),$$

et donc

$$I_6 = \int_0^{\pi/4} \cos(u)\,du = \frac{\sqrt{2}}{2}.$$

(3) On utilise le changement de variable $x = a\sin(u)$, qui comporte $dx = a\cos(u)\,du$ et, en utilisant la fonction $u(x) = \arcsin(x/a)$, $u(a) = \pi/2$ et $u(0) = 0$. On obtient

$$I_7 = \int_0^{\pi/2} \sqrt{a^2 - a^2\sin^2(u)}a\cos(u)\,du = a^2\int_0^{\pi/2} |\cos(u)|\cos(u)\,du.$$

Cette dernière intégrale est égale à $a^2\displaystyle\int_0^{\pi/2}\cos^2(u)\,du$, car toutes les valeurs prises par la fonction cosinus lorsque x varie entre 0 et $\pi/2$ sont positives. En intégrant par parties on trouve que

$$\int_0^{\pi/2}\cos^2(u)\,du = \int_0^{\pi/2}\sin^2(u)\,du,$$

d'où on peut écrire

$$\int_0^{\pi/2}\cos^2(u)\,du = \int_0^{\pi/2}\frac{\sin^2(u)+\cos^2(u)}{2}\,du = \frac{\pi}{4}.$$

La velauer de l'intégrale I_7 est $a^2\frac{\pi}{4}$.

(8) On intègre par parties

$$\begin{aligned} I_8 &= \int_{-\pi}^0 e^x\sin(x)\,dx = [e^x\sin(x)]_{-\pi}^0 - \int_{-\pi}^0 e^x\cos(x)\,dx \\ &= 0 - [e^x\cos(x)]_{-\pi}^0 - \int_{-\pi}^0 e^x\sin(x)\,dx = -1 - e^{-\pi} - I_8 \end{aligned}$$

On conclut que $I_8 = \frac{-1-e^{-\pi}}{2}$.

(9) On intègre par parties

$$\begin{aligned} I_9 &= \int_1^2 x^2\ln(2x)\,dx = \left[\frac{x^3}{3}\ln(2x)\right]_1^2 - \int_1^2 \frac{x^2}{3}\,dx \\ &= \left[\frac{8}{3}\ln(4) - \frac{1}{3}\ln(2)\right] - \left[\frac{8}{9} - \frac{1}{9}\right] = \ln\left(\frac{4^{8/3}}{2^{1/3}}\right) - \frac{7}{9} = \ln(32) - \frac{7}{9}. \end{aligned}$$

(10) Il faut réduire la fraction rationnelle en élément simples à intégrer. Pour le faire on commence par regarder le dénominateur et remarquer qu'il peut s'écrire comme $x^3(x^2+1)$. Notre objectif sera alors d'écrire $\frac{x^4+1}{x^5+x^3}$ comme une somme entre deux fractions rationnelles de dénominateur respectif x^3 et x^2+1. Comme le polynôme au numérateur a degré 4 il faut prévoir un polynôme de degré 2 au numérateur de x^3 et un polynôme de degré 1 au numérateur de x^2+1. On a alors

$$\frac{x^4+1}{x^5+x^3} = \frac{Ax^2+Bx+C}{x^3} + \frac{Dx+E}{x^2+1}.$$

Cela va nous donner un système de 5 équations pour les 5 inconnues A, B, C, D, E. On trouve que $A=-1$, $B=0$, $C=1$, $D=2$, $E=0$.

$$I_{10} = \int_1^2 \frac{x^4+1}{x^5+x^3}\,dx = \int_1^2 \frac{-x^2+1}{x^3} + \frac{2x}{x^2+1}\,dx = \int_1^2 \frac{-1}{x} + \frac{1}{x^3} + \frac{2x}{x^2+1}\,dx.$$

Par un calcul désormais immédiat on trouve $I_{10} = \ln(5/4) + 3/8$.

(4) $I_{11} = \displaystyle\int_{-3}^0 \frac{1}{x^2-3x+2}dx$. Il est facile de vérifier que $\frac{1}{x^2-3x+2} = -\frac{1}{x-1} + \frac{1}{x-2}$. En écrivant la fonction à intégrer comme la somme de deux termes nous avons alors

$$I_{11} = -\int_{-3}^0 \frac{1}{x-1}\,dx + \int_{-3}^0 \frac{1}{x-2}\,dx = \left[-\ln(|x-1|) + \ln(|x-2|)\right]_{-3}^0 = \ln\left(\frac{8}{5}\right).$$

(5) Par parties : $I_{12} = \displaystyle\int_1^2 \ln^2(x)dx = \left[x\ln^2(x)\right]_1^2 - \int_1^2 2\ln(x)dx = \left[x\ln^2(x) - 2x\ln(x) + 2x\right]_1^2 = 2\left(\ln^2(2) - 2\ln(2) + 1\right)$.

(13) On va essayer le changement de variable $u = e^x + 1$. On a alors $du = e^x\,dx$ et on pourra écrire $e^x = u - 1$. Les bornes d'intégration deviennent $u(0) = 2$ et $u(1) = e+1$.

$$I_{13} = \int_0^1 \frac{e^x-1}{e^x+1}dx = \int_2^{e+1} \frac{u-2}{u(u-1)}\,du.$$

La fonction de u à intégrer peut s'écrire comme la somme de deux fractions rationnelles plus simples de dénominateur respectif u et $u-1$: pour le faire nous devons trouver A et B dans $\mathbb{R}$ tels que

$$\frac{u-2}{u(u-1)} = \frac{A}{u} + \frac{B}{u-1}$$

On obtient $A=2$ et $B=-1$, donc notre intégrale devient

$$I_{13} = \int_2^{e+1} \frac{2}{u} - \frac{1}{u-1}\,du = \ln\left(\frac{(e+1)^2}{4e}\right).$$

(14) Nous utilisons le changement de variable $u = 1+2e^{-x}$. On a alors $du = -2e^{-x}dx$ et $\frac{1}{1-u}du = dx$.

$$\begin{aligned} I_{14} &= \int_1^2 \frac{e^{-2x}}{(1+2e^{-x})^2}\,dx = \int_{1+2/e}^{1+2/e^2} \left(\frac{u-1}{2}\right)^2 \frac{1}{u^2(1-u)}\,dx = -\int_{1+2/e}^{1+2/e^2} \frac{u-1}{4u^2}\,dx \\ &= \frac{1}{4}\left[\ln(x) + \frac{1}{x}\right]_{1+2/e^2}^{1+2/e} = \frac{1}{4}\left[\ln\left(\frac{e+2}{e^2+2}\right) + \frac{e}{e+2} - \frac{e^2}{e^2+2}\right]. \end{aligned}$$

(15) Le bon changement de variable nous est donné par l'énoncé. On a $dt = 2xdx$, les bornes d'intégration deviennent $t(0) = 1$ et $t(1) = 2$ et notre intégrale sera

$$I_{15} = \int_0^1 \frac{x^3}{x^2+1}dx = \int_1^2 \frac{t-1}{2t}dt = \left[\frac{t}{2} - \frac{1}{2}\ln(t)\right]_1^2 = \frac{1}{2}\left(1 - \ln(2)\right).$$

Exercise 136 exoTD5-00003

Calculer par parties les intégrales suivantes.

(1) $\int_0^\pi x\cos(x)\,dx$;

(2) $\int_1^2 x^2 e^x\,dx$;

(3) $\int_1^2 \ln(x)\,dx$;

(4) $\int_{-\pi}^0 e^x\sin(x)\,dx$;

(5) $\int_0^{\pi/2} \cos^2(x)\,dx$;

(6) $\int_1^2 x^2\ln(2x)\,dx$.

Exercise 137 exoTD5-a-0001

Calculer les intégrales suivantes

(1) $\int_0^1 x^5+3x^2+3\,dx$;

(2) $\int_1^2 x^{1/3}+4x^{1/2}\,dx$;

(3) $\int_{-\pi}^0 \cos(x)\,dx$;

(4) $\int_{-2}^1 e^x\,dx$;

(5) $\int_0^1 \frac{1}{\sqrt{x}}\,dx$.

corrTD5-a-0001

Correction of the exercise 137

(1) Lorsqu'on a une somme à intégrer, on peut prendre l'intégrale terme à terme :

$$\begin{aligned}\int_0^1 x^5+3x^2+3\,dx &= \int_0^1 x^5+3\int_0^1 x^2+3\int_0^1 1dx\\ &= \left[\frac{x^6}{6}\right]_0^1+3\left[\frac{x^3}{3}\right]_0^1+3\,[x]_0^1\\ &= \frac{25}{6}.\end{aligned} \tag{90.179}$$

Nous avons utilisé le fait que $\int x^n = \frac{x^{n+1}}{n+1}$. Par exemple, une primitive de x^5 est $\frac{x^6}{6}$.

(2) Ici nous utilisons la même technique, par exemple $\int x^{1/3}dx = \frac{x^{4/3}}{4/3} = \frac{3}{4}x^{4/3}$. La réponse est

$$-\frac{2^{1/3}}{24}\left(412^{2/3}-128\cdot 2^{1/6}-36\right). \tag{90.180}$$

(3) Une primitive de $\cos(x)$ est $\sin(x)$. Donc nous avons

$$\int_{-\pi}^0 \cos(x)dx = [\sin(x)]_{-\pi}^0 = \sin(0)-\sin(-\pi) = 0. \tag{90.181}$$

(4) Une primitive de e^x est la fonction e^x elle-même. Nous avons donc

$$\int_{-2}^1 e^x dx = [e^x]_{-2}^1 = e-e^{-2}. \tag{90.182}$$

(5) Étant donné que $\frac{1}{\sqrt{x}} = x^{-1/2}$, nous pouvons utiliser la formule usuelle d'intégration de puissance de x :

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \left[\frac{x^{1/2}}{1/2} \right]_0^1 = 2. \tag{90.183}$$

Exercise 138 exoTD5-a-0002

Calculer les intégrales suivantes par la méthode du changement de variable

(1) $\int_0^{\pi/3} \sin(3x)\, dx$;

(2) $\int_{-13}^{-11} e^{4x}\, dx$;

(3) $\int_1^2 \frac{1}{x+1}\, dx$;

(4) $\int_{-\pi/2}^{0} \sin^2(x)\cos(x)\, dx$. Conseil : le meilleur changement de variable dans ce cas est $y = \sin(x)$;

(5) $\int_1^{\pi} \frac{4x^3}{x^4+5}\, dx$.

corrTD5-a-0002

Correction of the exercise 138

(1) $\frac{2}{3}$;

(2) $-\frac{e^{-52}}{4} + \frac{e^{-44}}{4}$;

(3) $\ln(3) - \ln(2)$;

(4) $\frac{1}{3}$;

(5) $\ln(\pi^4 + 5) - \ln(6)$.

Exercise 139 exoTD5-a-0003

Calculer par parties les intégrales suivantes.

(1) $\int_0^{\pi} x\cos(x)\, dx$;

(2) $\int_1^2 x^2 e^x\, dx$;

(3) $\int_1^2 \ln(x)\, dx$;

(4) $\int_{-\pi}^{0} e^x \sin(x)\, dx$;

(5) $\int_0^{\pi/2} \sin^2(x)\, dx$;

(6) $\int_1^2 x^2 \ln(2x)\, dx$.

corrTD5-a-0003

Correction of the exercise 139

(1) -2 ;

(2) $2e^2 - e$;

(3) $2\ln(2) - 1$;

(4) $-\frac{1}{2}e^{-\pi} - \frac{1}{2}$;

(5) Étant donné que c'est une intégrale qui va revenir plus tard, essayons de la faire d'abord avec des bornes générales a et b. Soit donc à calculer

$$\int_a^b \sin^2(x)dx. \tag{90.184}$$

Nous faisons par partie :

$$\begin{aligned} f &= \sin(x) \quad g' = \sin(x) \\ f' &= \cos(x) \quad g = -\cos(x). \end{aligned} \tag{90.185}$$

Nous avons donc

$$\int_a^b \sin^2(x) = [-\sin(x)\cos(x)]_a^b + \int_a^b \cos^2(x)dx. \tag{90.186}$$ EqRelccscintab

Ajoutons la quantité $\int_a^b \sin^2(x)dx$ des deux côtés de cette équation. Dans le membre de gauche nous trouvons $2\int_a^b \sin^2(x)dx$. Dans le membre de droite, nous avons par contre

$$[-\sin(x)\cos(x)]_a^b + \int_a^b \cos^2(x)dx + \int_a^b \sin^2(x)dx. \tag{90.187}$$

La somme des deux intégrales est en réalité $\int_a^b \cos^2(x) + \sin^2(x)dx = \int_a^b 1dx = b - a$.
Au final, nous avons

$$2\int_a^b \sin^2(x)dx = [-\sin(x)\cos(x)]_a^b + (b - a). \tag{90.188}$$

À partir de là, nous pouvons trouver aussi (gratuitement) la valeur de $\int_a^b \cos^2(x)dx$ en substituant dans l'équation (90.186).
Pour être plus schématique, ce que nous avons fait est la chose suivante. Si nous notons $A = \int_a^b \sin^2(x)dx$, $B = \int_a^b \cos^2(x)dx$ et $c = [-\sin(x)\cos(x)]_a^b$, la relation (90.186) nous dit que

$$A = c + B. \tag{90.189}$$

Mais nous savons que

$$A + B = \int_a^b \sin^2(x)dx + \cos^2(x)dx = b - a. \tag{90.190}$$

Nous devons donc simplement résoudre le système

$$\begin{cases} A = c + B & \text{(90.191a)} \\ A + B = b - a. & \text{(90.191b)} \end{cases}$$

Si nous trouvons A et B en termes de c et $(b - a)$, nous avons trouvé toutes les intégrales qui nous intéressent. En substituant la valeur de A donné par la première équation dans la seconde, nous trouvons

$$c + B + B = (b - a), \tag{90.192}$$

c'est-à-dire

$$B = \frac{1}{2}(b - a) - \frac{c}{2}. \tag{90.193}$$

En remettant cette valeur de B dans la première, nous trouvons la valeur de A :

$$A = c + B = c + \frac{1}{2}(b - a) - \frac{c}{2} = \frac{1}{2}(b - a) + \frac{c}{2}. \tag{90.194}$$

Au final, ce que nous avons trouvé est que

$$\int_a^b \sin^2(x)dx = \frac{1}{2}(b - a) + \frac{1}{2}[-\sin(x)\cos(x)]_a^b \tag{90.195a}$$ subEqIntsincdxab

$$\int_a^b \cos^2(x)dx = \frac{1}{2}(b - a) - \frac{1}{2}[-\sin(x)\cos(x)]_a^b. \tag{90.195b}$$ subEqIntcoscdxab

Dans le cas de cet exercice, nous devons trouver $\int_0^{\pi/2}\cos^2(x)dx$, c'est-à-dire poser $a=0$ et $b=\pi/2$ dans l'équation (90.195b) :

$$\int_0^{\pi/2}\cos^2(x)dx=\frac{1}{2}\frac{\pi}{2}-\frac{1}{2}\Big(-\sin(\pi/2)\underbrace{\cos(\pi/2)}_{=0}+\underbrace{\sin(0)}_{=0}\cos(0)\Big)=\frac{\pi}{4}. \tag{90.196}$$

(6) $-\frac{\ln(2)}{3}+\frac{8\ln(4)}{3}-\frac{7}{9}=5\ln(2)-\frac{7}{9}$.

Exercise 140 exoTD5-0001

Calculer les intégrales suivantes.

(1) $\int_0^{2\pi}|\cos(x)|dx$.

(2) $\int_{-1}^{1}x|x|dx$.

(3) $\int_0^{\pi}\cos(2x)dx$.

(4) $\int_{-2}^{-1}\frac{1}{x}dx$.

(5) $\int_0^1\frac{1}{1+4x^2}dx$.

(6) $\int_0^1\frac{x^2-1}{x^2+1}dx$.

(7) $\int_0^1\sqrt{4-x^2}dx$.
Conseil : poser $x=2\cos(t)$, avec $t\in[\frac{\pi}{3},\frac{\pi}{2}]$.

corrTD5-0001

Correction of the exercise 140

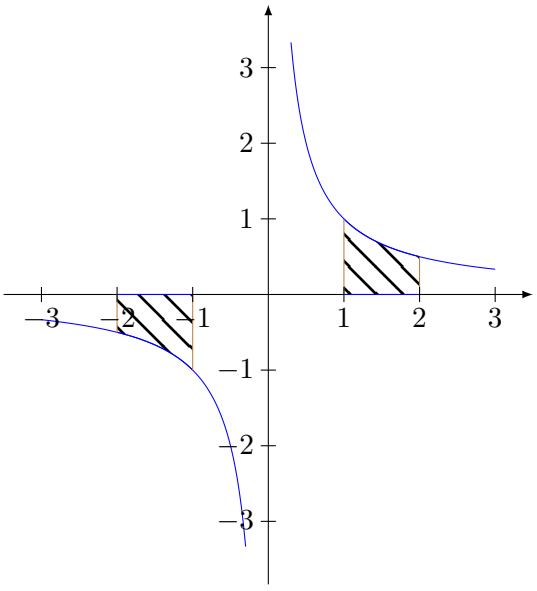


FIGURE 90.3: La surface achurée à gauche est la même que la surface achurée à droite. Seul le signe change lorsqu'on veut calculer l'intégrale. LabelFigUnSurxInt

(1) La technique usuelle pour intégrer une fonction dans laquelle se trouve une valeur absolue est de diviser l'intervalle d'intégration en sous-intervalles sur lesquels nous sommes sûr du

signe. Dans le cas de la fonction cosinus entre 0 et 2π, nous divisons de la façon suivante :

$$[0,2\pi]=\underbrace{[0,\frac{\pi}{2}]}_{\cos(x)\geqslant 0}\cup\underbrace{[\frac{\pi}{2},\frac{3\pi}{2}]}_{\cos(x)\leqslant 0}\cup\underbrace{[\frac{3\pi}{2},2\pi]}_{\cos(x)\geqslant 0}. \tag{90.197}$$

Sur les intervalles où $\cos(x)\geqslant 0$, nous avons $|\cos(x)|=\cos(x)$ et sur les intervalles où $\cos(x)\leqslant 0$, nous avons $|\cos(x)|=-\cos(x)$. L'intégrale se découpe donc de la façon suivante :

$$\begin{aligned}I&=\int_0^{\pi/2}\cos(x)dx+\int_{\pi/2}^{3\pi/2}(-)\cos(x)dx+\int_{3\pi/2}^{2\pi}\cos(x)dx\\&=[\sin(x)]_{x=0}^{x=\pi/2}-[\sin(x)]_{x=\pi/2}^{x=3\pi/2}+[\sin(x)]_{x=3\pi/2}^{x=2\pi}\\&=(1-0)-(-1-1)+(0-(-1))\\&=4.\end{aligned} \tag{90.198}$$

(2) Encore une fois, nous décomposons l'intervalle d'intégration en morceaux de telle manière que le signe de x soit constant sur chacun des morceaux :

$$\begin{aligned}\int_{-1}^{1}x|x|dx&=\int_{-1}^{0}(-x^2)dx+\int_0^1x^2dx\\&=\left[-\frac{x^3}{3}\right]_{-1}^{0}+\left[\frac{x^3}{3}\right]_0^1\\&=0-\left(-\frac{1}{3}\right)+\frac{1}{3}-0\\&=\frac{2}{3}.\end{aligned} \tag{90.199}$$

(3) En posant $y=2x$, nous trouvons $dy=2dx$, et par conséquent $dx=\frac{dy}{2}$. En ce qui concerne les bornes, si $x=0$, alors $y=0$ et si $x=\pi$, alors $y=2\pi$. L'intégrale à calculer devient

$$\int_0^{\pi}\cos(2x)dx=\int_0^{2\pi}\cos(y)\frac{dy}{2}=0. \tag{90.200}$$

(4) L'erreur à ne pas faire est de se souvenir que $\ln(x)$ est une primitive de $\frac{1}{x}$, et de dire

$$\int_{-2}^{-1}\frac{1}{x}dx=[\ln(x)]_{-2}^{-1}=\ln(-1)-\ln(-2). \tag{90.201}$$

À moins que vous sachiez comment définir le logarithme d'un nombre négatif, ce calcul n'a pas de sens.

En regardant la figure 90.3, nous voyons que les aires à gauche sont les mêmes que les aires à droite, sauf que le signe change. Nous pouvons donc écrire

$$\int_{-2}^{-1}\frac{1}{x}dx=-\int_1^2\frac{1}{x}dx=-\Big(\ln(2)-\ln(1)\Big)=\ln(1)-\ln(2). \tag{90.202}$$

Notez que le résultat est négatif comme le dessin l'indique (la fonction est négative entre -2 et -1).

Une autre façon de faire est de retenir l'intégrale

$$\int\frac{1}{x}dx=\ln|x| \tag{90.203}$$

avec les valeurs absolues.

(5) L'intégrale la plus ressemblante que nous connaissons « par cœur » $\int \frac{1}{1+x^2}dx = \arctan(x)$. Le truc est de poser $y = 2x$ de façon que $4x^2$ se transforme en y^2. Nous avons $dx = dy/2$, et en ce qui concerne les bornes, si $x = 0$, $y = 0$ et $y = 2$ lorsque $x1$. Par conséquent,

$$\int_0^1 \frac{dx}{1+4x^2} = \frac{1}{2}\int_0^2 \frac{dy}{1+y^2} = \frac{1}{2}\left[\arctan(y)\right]_0^2 = \frac{1}{2}\arctan(2). \tag{90.204}$$

Notez que $\arctan(0) = 0$.

(6) Lorsque le numérateur et le dénominateur sont des polynômes du même degré, il faut faire apparaitre le dénominateur au numérateur en ajoutant et en enlevant ce qu'il faut. Dans notre cas, nous avons

$$\frac{x^2-1}{x^2+1} = \frac{x^2+1-2}{x^2+1} = \frac{x^2+1}{x^2+1} - \frac{2}{x^2+1} = 1 - 2\frac{1}{x^2+1}. \tag{90.205}$$

Par conséquent

$$\begin{aligned}\int_0^1 \frac{x^2-1}{x^2+1} &= \int_0^1 1dx - 2\int_0^1 \frac{dx}{x^2+1}\\ &= \left[x\right]_0^1 - 2\left[\arctan(x)\right]_0^1\\ &= 1 - 2\arctan(1)\\ &= 1 - \frac{\pi}{2}.\end{aligned} \tag{90.206}$$

Le fait que $\arctan(1) = \frac{\pi}{4}$ vient du fait qu'en $\pi/4$, nous avons $\sin(\pi/4) = \cos(\pi/4)$ et donc $\tan(\pi/4) = 1$.

(7) Le changement de variable $x = 2\cos(t)$ est donné dans le sens inverse de l'habitude, mais ce n'est pas grave. Nous avons $dx = -2\sin(t)dt$. Pour avoir $x = 0$, il faut $t = \pi/2$ et pour avoir $x = 1$, il faut $\cos(t) = \frac{1}{2}$, et donc $t = \pi/3$. L'intégrale à calculer est donc

$$\int_0^1 \sqrt{4-x^2}dx = \int_{\pi/2}^{\pi/3} \sqrt{4-4\cos^2(t)}(-2)\sin(t)dt. \tag{90.207}$$

Pour pouvons simplifier l'expression à intégrer en sortant ce qu'on peut sortir de la racine carrée :

$$\sqrt{4-4\cos^2(t)} = \sqrt{4\big(1-\cos^2(t)\big)} = 2\sqrt{1-\cos^2(t)} = 2\sqrt{\sin^2(t)}. \tag{90.208}$$

Est-ce que nous pouvons dire $\sqrt{\sin^2(t)} = \sin(t)$? Étant donné que nous regardons l'intégrale avec t entre $\pi/2$ et $\pi/3$, le sinus est toujours positif, et par conséquent nous pouvons écrire $\sqrt{\sin^2(t)} = \sin(t)$. Au final, l'intégrale que nous devons faire est

$$-4\int_{\pi/2}^{\pi/3} \sin^2(t)dt. \tag{90.209}$$

Cette intégrale est « connue » parce qu'elle a déjà été calculée : elle est donnée à l'équation (90.195a). En remplaçant a par $\pi/2$ et b par $\pi/3$ nous trouvons

$$\begin{aligned}-4\int_{\pi/2}^{\pi/3} \sin^2(x)dx &= -4\frac{1}{2}\left(\frac{\pi}{3}-\frac{\pi}{2}\right) - 4\frac{1}{2}\left[-\sin(x)\cos(x)\right]_{\pi/2}^{\pi/3}\\ &= -2\left(-\frac{\pi}{6}\right) - 2\Big(-\underbrace{\sin(\pi/3)}_{=\sqrt{3}/2}\underbrace{\cos(\pi/3)}_{=1/2} + \sin(\pi/2)\underbrace{\cos(\pi/2)}_{=0}\Big)\\ &= \frac{\pi}{3} - 2\left(-\frac{\sqrt{3}}{2}\frac{1}{2}\right)\\ &= \frac{\pi}{3} + \frac{\sqrt{3}}{2}.\end{aligned} \tag{90.210}$$

Exercise 141 exoTD5-0002

Calculer l'aire de la surface limitée par les droites d'équations $y = 0$, $x- = 0$, $x = 2$ et la parabole d'équation $y = x^2 - x$. Faire un dessin.

corrTD5-0002

Correction of the exercise 141

<+CorrTD5-0002+>

The result is on figure 90.4.

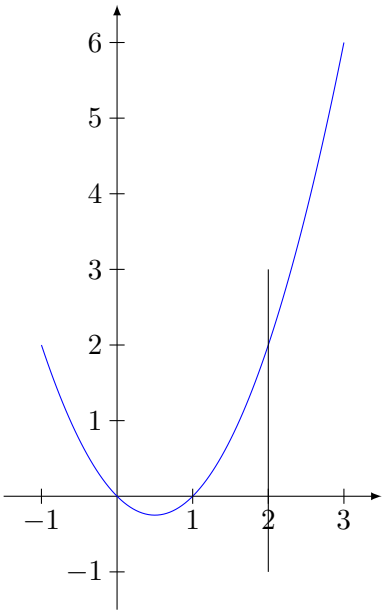


FIGURE 90.4: <+Type your caption here+>

LabelFigAireParabole

Exercise 142 exoTD5-0003

Calculer les intégrales suivantes.

(1) $\int_{-3}^{0} \frac{1}{x^2 - 3x + 2} dx$. Conseil : écrire la fraction sous la forme $\frac{a}{x-1} + \frac{b}{x-2}$.

(2) $\int_{1}^{2} \ln^2(x) dx$.

(3) $\int_{0}^{1} \frac{e^x - 1}{e^x + 1} dx$.

(4) $\int_{1}^{2} \frac{e^{-2x}}{(1 + 2e^{-x})^2} dx$.

(5) $\int_{0}^{1} \frac{x^3}{x^2 + 1} dx$. Conseil : poser $t = x^2 + 1$.

corrTD5-0003

Correction of the exercise 142

<+CorrTD5-0003+>

Exercise 143 exoTD5-0004

Calculer les intégrales suivantes.

(1) $\int_{0}^{1} x e^x dx$

(2) $\int_{0}^{1} x^2 e^x dx$

(3) $\int_{1}^{2} \arctan(x) dx$

(4) $\int_{-1}^{1} x\left(\cos^{10}(x)+\frac{\sin^4(x)-e^{x^2}+1}{\sqrt{|x|}}\right)dx.$

corrTD5-0004

Correction of the exercise 143

(1) 1

(2) $e-2$

(3) $-\frac{\pi}{4}+\frac{\ln(2)}{2}-\frac{\ln(5)}{2}+2\arctan(2)$

(4)

Exercise 144 exoTD5-0005

Calculer les primitives suivantes.

Itemintizcxexsq

(1) $\int te^{-t^2}dt$.

ItemTDczcii

(2) $\int\frac{3t^2+2}{t^3}$.

corrTD5-0005

Correction of the exercise 144

(1) Cette intégrale se fait par changement de variables en posant $u=t^2$. En ce qui concerne la dérivée,

$$\frac{du}{dt}=2t \tag{90.211}$$

et par conséquent $dt=\frac{du}{2t}$. En remplaçant dans l'intégrale demandée,

$$I=\int te^{-u}\frac{du}{2t}. \tag{90.212}$$

Les deux t restants se simplifient et il reste

$$\frac{1}{2}\int e^{-u}du=-\frac{1}{2}e^{-u}+C=-\frac{1}{2}e^{-t^2}+C \tag{90.213}$$

où nous avons fait le changement de variable inverse afin d'exprimer la réponse en termes de la variable originale t.

Notez que pour bien obtenir toutes les primitives, nous devons écrire « $+C$ » dans la réponse. Cela signifie que pour tout choix de constante C, la fonction

$$F(t)=-\frac{1}{2}e^{-t^2}+C \tag{90.214}$$

est une primitive de la fonction

$$f(t)=te^{-t^2}. \tag{90.215}$$

(2) Nous pouvons couper la fraction en deux :

$$\frac{3t^2+2}{t^3}=\frac{2t^2}{t^3}+\frac{2}{t^3}=\frac{3}{t}+2t^{-3}. \tag{90.216}$$

Pour rappel, $t^{-3}=\frac{1}{t^3}$. L'intégrale à calculer devient donc

$$\begin{aligned}\int\frac{3t^2+2}{t^3}dt&=3\int\frac{1}{t}dt+2\int t^{-3}dt\\&=3\ln(t)+2\frac{t^{-2}}{-2}+C\\&=3\ln(t)-\frac{1}{t^2}+C.\end{aligned} \tag{90.217}$$

90.10 Équations différentielles

Exercise 145 exoSVT-0004

Vérifier que $x(t) = \sin(t)$ est une solution de l'équation différentielle

$$x'' = -x. \tag{90.218}$$

EqRNBDxb

Pour quelles valeurs des paramètres A et B la fonction $x(t) = A\sin(t) + B\cos(t)$ est-elle solution de l'équation (90.218) ?

corrSVT-0004

Correction of the exercise 145

<+CorrSVT-0004+>

Exercise 146 exoSVT-0003

Nous considérons l'équation différentielle

$$\frac{dx}{dt} = -x\ln(c/2\pi)\sin(t) + \frac{x}{t}. \tag{90.219}$$

(1) Vérifier que la fonction

$$x(t) = te^{\ln(c/2\pi)\cos(t)} \tag{90.220}$$

avec $c > 0$ est une solution.

(2) Calculer la limite $\lim_{t\to\infty} x(t)$.

(3) Calculer $x(0)$.

(4) Pour quelle valeur(s) de t avons nous $x(t) = c$?

corrSVT-0003

Correction of the exercise 146

(1) En dérivant la fonction proposée par rapport à t, nous trouvons

$$x'(t) = -te^{\ln(c/2\pi)\cos(t)}\ln(c/2\pi)\sin(t) + e^{\ln(c/2\pi)\cos(t)}. \tag{90.221}$$

Nous vérifions maintenant l'équation différentielle

$$x'(t) = -x\ln(c/2\pi)\sin(t) + \frac{x}{t} \tag{90.222}$$

en remplaçant x et x' par leurs valeurs.

$$-te^{\ln(c/2\pi)\cos(t)}\ln(c/2\pi)\sin(t) + e^{\ln(c/2\pi)\cos(t)} \stackrel{?}{=} -te^{\ln(c/2\pi)\cos(t)}\ln(c/2\pi)\sin(t) + \frac{te^{\ln(x)\cos(t)}}{t} \tag{90.223}$$

Un peu de calcul montre que cela est correct.

(2) Lorsque $t \to \infty$, le facteur exponentiel ne tend vers rien de précis (parce que $\cos(t)$ oscille). Il est cependant toujours positif et inférieur à $c/2\pi$. Le facteur t par contre tend vers l'infini. Nous avons par conséquent

$$\lim_{t\to\infty} x(t) = \infty. \tag{90.224}$$

(3) $t = 2\pi$.

Exercise 147 exoTD6A-0002

(1) Vérifier que la fonction $x_1(t) = \sin(t)$ est une solution de l'équation différentielle

$$\frac{dx}{dt} = \sqrt{1 - x^2}, \tag{90.225}$$

arcsin

pour t dans l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

(2) Vérifier que les deux fonctions constantes $x_2(t) = 1$ et $x_3(t) = -1$ sont des solutions de (90.225) pour tout t.
(3) Dessiner les trajectoires des trois solutions.
(4) Est-il possible de prolonger la solution x_1 après $t = \frac{\pi}{2}$? Avant $t = -\frac{\pi}{2}$? Discuter.
(5) Est-il possible de trouver une solution si la donnée de départ est $x(0) = 2$?
(6) Est-il possible de trouver une solution si la donnée de départ est $x(0) = \alpha$, avec α quelconque dans l'intervalle $[-1, 1]$? Ce point est difficile si on ne connaît pas la méthode de séparation de variables.
(7) Dessiner "toutes" les trajectoires trouvées.

corrTD6A-0002

Correction of the exercise 147

(1) Si $x_1(t) = \sin(t)$ nous avons

$$\frac{dx_1}{dt} = \cos(t), \tag{90.226}$$

et

$$\sqrt{1 - x_1^2(t)} = \sqrt{1 - \sin^2(t)} = \sqrt{\cos^2(t)} = \cos(t). \tag{90.227}$$

Attention : en général $\sqrt{\cos^2(t)} = |\cos(t)|$. Ici nous pouvons enlever les valeurs absolues parce que $\cos(t) > 0$ lorsque $t \in]-\pi/2, \pi/2[$.
(2) La vérification est immédiate parce que la dérivée des solutions constante est nulle.
(3) Voir la figure figure 90.5.

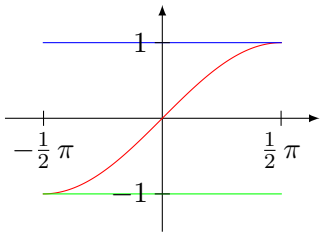


FIGURE 90.5: Trois de solutions à l'équation différentielle de l'exercice 147.

LabelFigSolsEqDiffSin

(4) Non parce que l'équation différentielle $x'(t) = \sqrt{1 - x(t)^2}$ dit que x' doit toujours être positive. Mais si nous prolongeons $x_1(t) = \sin(t)$ en dehors de l'intervalle $]-\pi/2, \pi/2[$, la dérivée devient négative.
Il est cependant possible de la prolonger sur $[-\pi/2, \pi/2]$.
(5) Étant donné que l'équation que doit satisfaire $x(t)$ (pour tout t) contient $\sqrt{1 - x(t)^2}$, le nombre $x(t)$ doit rester entre -1 et 1 pour tout t.
(6) La méthode de séparation des variables consiste à mettre tous les x d'un côté (y compris dx) et tous les t de l'autre, y compris dt. Nous trouvons

$$\frac{dx}{\sqrt{1 - x^2}} = dt. \tag{90.228}$$

Nous prenons l'intégrale des deux côtés en n'oubliant pas la constante d'intégration :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x)=1/sqrt(1-x**2)
sage: f.integrate(x)
x |--> arcsin(x)
```

Notez que Sage ne vous donne constante d'intégration [4] ni domaines [5] (nous y reviendrons). Nous avons :

$$t = \arcsin(x) + C. \tag{90.229}$$

Par conséquent,

$$x(t) = \sin(t + C). \tag{90.230}$$

Ici nous avons changé C en $-C$ pour la commodité d'écriture.
Nous demandons d'avoir $x(0) = \alpha$, c'est-à-dire

$$x(0) = \sin(C) = \alpha, \tag{90.231}$$

ou encore $C = \arcsin(\alpha)$. La solution de l'équation que nous cherchons est donc

$$x(t) = \sin\big(t + \arcsin(\alpha)\big). \tag{90.232}$$

Maintenant nous devons réfléchir sur les domaines. . .Regardons donc encore une fois la condition d'existence $x'(t) \geqslant 0$:

$$x'(t) = \cos(t + C) \geqslant 0 \tag{90.233}$$

implique $t + C \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ et par conséquent

$$t \in [-\frac{\pi}{2} - \arcsin(\alpha), \frac{\pi}{2} - \arcsin(\alpha)]. \tag{90.234}$$

Notons que le domaine dépend de α ! Nous prenons aussi la convention que $\arcsin(\alpha)$ est le nombre z entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ tel que $\sin(z) = \alpha$.

(7) Quelques solutions sont dessinées à la figure 90.6.

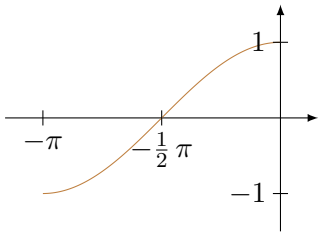


FIGURE 90.6: Des solutions pour l'équation différentielle de l'exercice 147. LabelFigSolsSinpA

Exercise 148 exoTD6A-0003

(1) Vérifier que le système d'équation différentielles systWVHnZx

$$\begin{cases} \dfrac{dx}{dt} = y & \text{(90.235a)} \\ \dfrac{dy}{dt} = -x & \text{(90.235b)} \end{cases}$$

est équivalent à l'équation différentielle de deuxième degré

$$\frac{d^2x}{dt^2} = -x, \tag{90.236}$$ eq2deg

au sens que $x(t) = g(t)$ est une solution de (90.236) si et seulement si $x(t) = g(t)$, $y(t) = g'(t)$ est une solution de ((1)).

(2) Trouver les solutions constantes de ((1)).

4. Vous pouvez voir votre cours de math comme un cours d'anticipation des fautes que vous commettriez si vous vous fiiez trop facilement à votre calculatrice ou votre ordinateur.

5. Il ne peut pas savoir quel est le contexte dans lequel vous êtes en train de calculer !

(3) Vérifier que la fonction $x(t) = A\cos(t) + B\sin(t)$ est une solution de (90.236) pour tous A, B dans $\mathbb{R}$. Nous acceptons que *toutes* les solutions de l'équation ((1)) sont de cette forme.

(4) Dessiner sur le plan x-y les trajectoires des solutions $(x(t), y(t))$.

(5) Trouver toutes les solutions de (90.236) telles que $x(0) = 1$. Trouver sur le plan x-y les trajectoires correspondantes. Comment faut-il faire pour choisir une unique solution de (90.236) ? Discuter.

corrTD6A-0003

Correction of the exercise 148

(1) Soit g une fonction qui satisfait $g''(t) = -g(t)$. Si nous posons

$$x(t) = g(t) \tag{90.237a}$$
$$y(t) = g'(t) \tag{90.237b}$$

alors le couple (x, y) est une solution du système [SubesqoSZWaq]

$$\begin{cases} x'(t) = y & \text{(90.238a)} \\ y'(t) = -x. & \text{(90.238b)} \end{cases}$$

En effet d'une part $x'(t) = g'(t) = y(t)$, et d'autre part $y'(t) = dg'/dt = g''(t) = -g(t) = x(t)$. Dans le sens inverse si le couple de *fonctions*[6] (x, y) satisfait au système (90.238), alors en posant $g(t) = x(t)$ nous obtenons une fonction g qui satisfait $g'' = -g$.

(2) Si x est une fonction constante, $x' = 0$ et la première équation du système donne $y = 0$. À ce moment, la seconde équation du système donne $x(t) = 0$. La seule solution constante est $(0, 0)$.

(3) En dérivant nous obtenons successivement :

$$x(t) = A\cos(t) + B\sin(t) \tag{90.239a}$$
$$x'(t) = -A\sin(t) + B\cos(t) \tag{90.239b}$$
$$x''(t) = -A\cos(t) - B\sin(t). \tag{90.239c}$$

Nous avons donc $x'' = -x$ et la fonction proposée est donc une solution de l'équation.

(4) La trajectoire à dessiner est la courbe d'équation paramétrique

$$\begin{cases} x(t) = A\cos(t) + B\sin(t) & \text{(90.240a)} \\ y(t) = -A\sin(t) + B\cos(t) & \text{(90.240b)} \end{cases}$$

Le calcul montre que pour tout t nous avons

$$x(t)^2 + y(t)^2 = A^2 + B^2, \tag{90.241}$$

ce qui montre que nous avons un cercle de rayon $\sqrt{A^2 + B^2}$.

Les trajectoires possibles sont donc de deux natures : d'abord il y a la trajectoire constante $(x, y) = (0, 0)$ et ensuite il y a des cercles de tous les rayons.

Encore une fois remarquons que les constantes A et B ne jouent pas un rôle mineur. En effet la solution avec $A = 2$, $B = 3$ est :

$$\begin{cases} x_1(t) = 2\cos(t) + 3\sin(t) & \text{(90.242a)} \\ y_1(t) = -2\sin(t) + 3\cos(t). & \text{(90.242b)} \end{cases}$$

C'est un cercle de rayon $\sqrt{13}$. La solution avec $A = 3$, $B = 2$ est

$$\begin{cases} x_2(t) = 3\cos(t) + 2\sin(t) & \text{(90.243a)} \\ y_2(t) = -3\sin(t) + 2\cos(t). & \text{(90.243b)} \end{cases}$$

6. Nous n'insisterons jamais assez sur le fait que x et y sont ici des fonctions.

C'est le même cercle ! Il y a par contre une différence de point de départ :

$$\begin{pmatrix} x_1(0) \\ y_1(0) \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \tag{90.244}$$

tandis que

$$\begin{pmatrix} x_2(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}. \tag{90.245}$$

Ces deux solutions parcourent la même orbite, mais ne partent pas du même points !

Fixer les constantes A et B permet de fixer à la fois le rayon de l'orbite et son point de départ.

(5) Demander $x(0) = 1$ impose $A = 1$, mais B reste libre. Nous avons mis quelques trajectoires sur la figure 90.7.

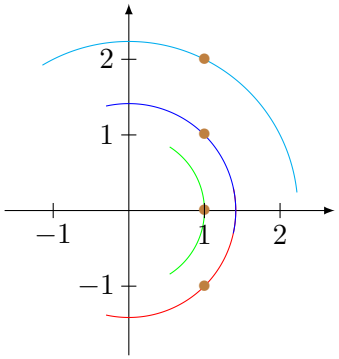


FIGURE 90.7: Les trajectoires sont tracées pour t entre 0 et 1. Les points correspondent à $t = 0$. La courbe rouge correspond à $B = -1$. Nous voyons qu'elle est sur la même trajectoire que $B = 1$, mais elle ne part pas du même point. LabelFigTrajs

Exercise 149 exoTD6b-0001

Pour quelles valeur(s) des constantes a et b la fonction $f(x) = x^4e^{2x}$ est-elle une solution de l'équation différentielle

$$y' - ay = bx^3e^{2x} \, ? \tag{90.246}$$

corrTD6b-0001

Correction of the exercise 149

En injectant la forme de y proposée dans l'équation, le membre de gauche devient

$$y' - ay = 4x^3e^{2x} + 2x^4e^{2x} - ax^4e^{2x}, \tag{90.247}$$

tandis que le membre de droite devient bx^3e^{2t}.

Nous devons donc voir pour quelles valeurs de a et b ces deux grandeurs sont égales pour tout x. En simplifiant par e^{2x} (qui ne s'annule jamais),

$$4x^3 + 2x^4 - ax^4 = bx^3e^{2x}, \tag{90.248}$$

et donc $a = 2$, $b = 4$.

Exercise 150 exoTD6b-0002

Déterminez l'ensemble des fonctions qui satisfont les équations différentielles suivantes.

(1) $y' = (t^3 + t)e^{-y}$
(2) $y' = 1 + y^2$
(3) $y' = \frac{\cos(t)}{1+e^y}$
(4) $y' = y^2$
(5) $y' = y^{\frac{1}{3}}$

corrTD6b-0002

Correction of the exercise 150

La technique de base est toujours la même : écrire y' sous la forme dy/dt et puis rassembler d'un côté les termes en y et de l'autre ceux en t.

(1) Nous avons successivement

$$\frac{dy}{dt} = t^3 + t, \tag{90.249a}$$

$$e^y dy = (t^3 + t)dt \tag{90.249b}$$

$$e^y = \frac{t^4}{4} + \frac{t^2}{2} + C. \tag{90.249c}$$

La dernière ligne est simplement le fait d'avoir intégré des deux côtés. Nous isolons y en passant au logarithme :

$$y(t) = \ln\left(\frac{t^4}{4} + \frac{t^2}{2} + C\right). \tag{90.250}$$

Une méthode alternative à résoudre soi-même est de demander à un ordinateur de le faire :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: t=var('t')
sage: y=function('y',t)
sage: DE=diff(y,t)-(t**3+t)*exp(-y)
sage: desolve(DE,[y,t])
e^y(t) == 1/4*t^4 + 1/2*t^2 + c + 1/4
```

Notez qu'il ne termine pas complètement le travail parce qu'il ne fait pas le logarithme pour isoler y. Par contre, il faut avouer que le gros est fait. Notez aussi qu'un peu naïvement, il écrit $c + \frac{1}{4}$ alors que c est une constante arbitraire ; on peut simplement écrire c.

(2) Nous avons à intégrer les deux membres de

$$\frac{dy}{1 + y^2} = dt, \tag{90.251}$$

ce qui donne $\arctan(y) = t + C$. La solution est

$$y(t) = \tan(t + C). \tag{90.252}$$

Tant qu'aucun contexte n'est donné, nous n'avons pas grand chose à dire sur le domaine (et pourtant il y aurait des choses à dire !).

(3) Nous devons intégrer

$$(1 + e^y)dy = \cos(t)dt. \tag{90.253}$$

La solution est

$$y + e^y = \sin(t) + C. \tag{90.254}$$

Il n'y a pas de moyen simples de résoudre cette équation pour isoler y. Voici donc un exemple d'équation différentielle que nous ne pouvons pas résoudre explicitement.

(4) À intégrer :

$$\frac{dy}{y^2} = dt, \tag{90.255}$$

solution :

$$y = -\frac{1}{t + C}. \tag{90.256}$$

(5) Dans le même ordre d'idée que le précédent :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: t=var('t')
sage: y=function('y',t)
sage: DE=diff(y,t)-y**(1/3)
sage: desolve(DE,[y,t])
3/2*y(t)^(2/3) == c + t
```

Sage nous dit que

$$\frac{3}{2}y^{2/3} = C + t \tag{90.257}$$

La solution recherchée est donc

$$y(t) = \left(\frac{2}{3}(C+t)\right)^{-3/2}. \tag{90.258}$$

Exercise 151 exoTD6b-0003

Trouver les fonctions qui satisfont les équations et les conditions suivantes :

(1) $y' = (t^3 + t)e^{-y}$, $y(1) = 1$
(2) $y' = 1 + y^2$, $y(0) = 0$
(3) $y' = 1 + y^2$, $y(0) = 1$
(4) $y' = \frac{\cos(t)}{1+e^y}$, $y(\frac{\pi}{2}) = 3$
(5) $y' = y^2$, $y(1) = 2$
(6) $y' = y^2$, $y(1) = 0$
(7) $y' = y^{\frac{1}{3}}$, $y(0) = -1$
(8) $y' = y^{\frac{1}{3}}$, $y(0) = 0$

corrTD6b-0003

Correction of the exercise 151

Ce sont les mêmes équations que pour l'exercice 150. Nous repartirons donc de la solution, et nous fixerons la constante.

(1) En plaçant $t = 1$ dans la solution nous avons

$$y(1) = \ln\left(\frac{3}{4} + C\right), \tag{90.259}$$

et par conséquent, pour avoir $y(1) = 1$, nous demandons

$$C = e - \frac{3}{4}. \tag{90.260}$$

(2) Nous avons $y(0) = \tan(C)$. Demander $y(0) = 0$ demande $\tan(C) = 0$ c'est-à-dire $C = k\pi$. Les solutions sont donc

$$y_k(t) = \tan(t + k\pi). \tag{90.261}$$

(3) C'est la même équation que la précédente. La condition sur C s'écrit $\sin(C) = \cos(C)$, c'est-à-dire $C = \frac{\pi}{4} + k\pi$. Les solutions sont donc

$$y_k(t) = \tan(t + k\pi + \frac{\pi}{4}). \tag{90.262}$$

(4) Pour cette équation nous n'étions pas parvenu à écrire une solution explicite, mais nous pouvons tout de même fixer la constante :

$$y(\pi/2) + e^{y(\pi/2)} = \sin(\pi/2) + C, \tag{90.263}$$

mais $y(\pi/2) = 3$ et $\sin(\pi/2) = 1$, donc

$$2 + e^3 = C. \tag{90.264}$$

L'équation qui donne y (et que nous ne savons pas résoudre) est alors

$$y + e^y = \sin(t) + 3 + e^3. \tag{90.265}$$

(5) La solution en $t = 1$ est

$$y(1) = -\frac{1}{1+C}. \tag{90.266}$$

Pour obtenir $y(1) = 2$ nous avons besoin de $C = -3/2$. La solution est donc

$$y(t) = -\frac{1}{t - \frac{3}{2}}. \tag{90.267}$$

(6) Cette fois nous devons avoir

$$-\frac{1}{1+C} = 0, \tag{90.268}$$

ce qui est impossible. Voici donc un exemple d'équation différentielle avec condition initiale impossible.

(7) L'équation à résoudre pour C est

$$y(0) = \left(\frac{2}{3}C\right)^{-2/3} = -1. \tag{90.269}$$

Cela est impossible. Notez que

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: solve(x**(-2/3)==-1,x)
[x == I, x == -I]
```

Sage trouve des solutions parmi les nombres complexes. Si vous voulez utiliser un ordinateur pour travailler, il faut pouvoir interpréter ses réponses !

(8) Ici nous devons résoudre

$$\left(\frac{2}{3}C\right)^{-2/3} = 0, \tag{90.270}$$

ce qui donne $C = 0$ et par conséquent

$$y(t) = \left(\frac{2t}{3}\right)^{-2/3}. \tag{90.271}$$

Exercise 152 ⊙⊙ | exoTD6b-0004

Résoudre les problèmes suivants

(1) $y' - 2ty = t \qquad y(1) = e^{-\frac{1}{2}}$

(2) $y' + y\tan(t) = \sin(2t) \qquad y(0) = 6$

(3) $y' + y\cos(t) = 0 \qquad y(0) = 10$

(4) ⊛⊛ $t^3y' + (2 - 3t^2)y = t^3$

(5) $y' - y = t$

corrTD6b-0004

Correction of the exercise 152

(1) L'équation homogène correspondante est

$$y'_H - 2ty_H = 0. \tag{90.272}$$

Esseghoni

En remplaçant y'_H par $\frac{dy_H}{dt}$ et en passant tous les t à droite et tous les y à gauche,

$$\frac{dy_H}{y} = 2tdt. \tag{90.273}$$

En intégrant des deux côtés (c'est-à-dire en prenant les primitives et en n'oubliant pas la constante d'intégration), nous trouvons $\ln(y_H) = t^2 + C$. Nous avons donc

$$y_H(t) = e^{t^2}e^C. \tag{90.274}$$

Afin de simplifier les notations, nous allons écrire

$$y_H = Ke^{t^2} \tag{90.275}$$

avec $K = e^C$. Pour chaque $K \in \mathbb{R}$, la fonction $y_H(t) = Ke^{t^2}$ vérifie l'équation homogène (90.272).

Pour résoudre l'équation complète (non homogène), nous utilisons la *méthode de variations des constantes*, c'est-à-dire que nous cherchons les solutions sous la forme $y(t) = K(t)e^{t^2}$. Nous remplaçons cette expression pour y dans l'équation de départ :

$$K'e^{t^2} + 2te^{t^2} - 2tKe^{t^2} = t \tag{90.276}$$

où nous avons utilisé le fait que $y' = K'e^{t^2} + 2Kte^{t^2}$. Après simplifications, nous trouvons

$$K' = te^{-t^2}. \tag{90.277}$$

La fonction K est donc une primitive de la fonction te^{-t^2}, c'est-à-dire

$$K = \int te^{-t^2}dt. \tag{90.278}$$

Cette intégrale est l'intégrale de l'exercice 144.(1). La réponse est

$$K(t) = -\frac{1}{2}e^{-t^2} + C. \tag{90.279}$$

La solution de l'équation différentielle proposée est donc

$$y(t) = K(t)e^{t^2} = \big(-\frac{1}{2}e^{-t^2} + C\big)e^{t^2}, \tag{90.280}$$

ou encore, après simplifications,

$$y(t) = -\frac{1}{2} + Ce^{t^2}. \tag{90.281}$$

Nous pouvons vérifier que cela est bien la solution de l'équation proposée en remplaçant dans l'équation de départ. D'abord $y' = 2tCe^{t^2}$, donc

$$y' - 2ty = 2tCe^{t^2} - 2t\big(-\frac{1}{2} + Ce^{t^2}\big) = 2tCe^{t^2} + t - 2tCe^{t^2} = t, \tag{90.282}$$

ce qu'il fallait. Nous avons donc bien, pour chaque C, une solution de l'équation différentielle $y' - 2ty = t$.

Maintenant, nous demandons en plus que $y(1) = e^{-1/2}$. Le point à comprendre est que nous avons une solution à l'équation différentielle *sans contrainte* pour chaque C. Nous pouvons fixer C de telle manière à avoir une solution qui satisfait la contrainte :

$$y(1) = -\frac{1}{2} + Ce^1 = -\frac{1}{2} + eC. \tag{90.283}$$

Si nous voulons que cela soit égal à $e^{-1/2}$, nous devons résoudre l'équation

$$-\frac{1}{2} + eC = e^{-1/2}. \tag{90.284}$$

La solution est

$$C = e^{-1}\big(e^{-1/2} + \frac{1}{2}\big). \tag{90.285}$$

(2)
```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage:  t=var('t')
sage: y=function('y',t)
sage: DE=diff(y,t)+y*tan(t)-sin(2*t)
sage: desolve(DE,[y,t])
(c - 2*cos(t))/sec(t)
```

La solution est

$$y(t) = \big(-2\cos(t) + C\big)\cos(t). \tag{90.286}$$

Pour obtenir la valeur de la constante C nous utilisons la condition initiale :

$$y(0) = (-2 + C) = 6 \tag{90.287}$$

par conséquent $C = 8$.

(3) Ici nous avons seulement une équation homogène. Il ne faudra donc pas utiliser la méthode de la variation des constantes. D'abord nous écrivons l'équation différentielle sous la forme

$$\frac{dy}{dt} = -y\cos(t), \tag{90.288}$$

ensuite nous remettons tous les y à gauche et tous les t à droite :

$$\frac{dy}{y} = -\cos(t)dt. \tag{90.289}$$

En intégrant des deux côtés, nous avons

$$\ln(y) = -\sin(t) + C \tag{90.290}$$

et donc

$$y = Ke^{-\sin(t)} \tag{90.291}$$

Eqsbzqiiikesint

où nous avons posé $K = e^C$.

Pour vérifier le résultat, nous commençons par calculer la dérivée :

$$y' = -K\cos(t)e^{-\sin(t)}, \tag{90.292}$$

et ensuite nous remettons dans l'équation :

$$y' + y\cos(t) = -K\cos(t)e^{-\sin(t)} + Ke^{-\sin(t)}\cos(t) = 0, \tag{90.293}$$

comme il se doit.

Nous devons maintenant trouver pour quelle valeur de la constante K nous avons $y(0) = 10$. Pour cela nous posons $t = 0$ dans l'équation (90.291) :

$$y(0) = Ke^{-\sin(0)} = K. \tag{90.294}$$

Il faut donc poser $K = 10$ pour obtenir $y(0) = 10$. La réponse est donc

$$y(t) = 10e^{-\sin(t)}. \tag{90.295}$$

(4) L'équation homogène à résoudre est

$$t^3y'_H + (2 - 3t^2)y_H = 0. \tag{90.296}$$

En séparant les variables,

$$\frac{dy_H}{y_H} = -\frac{2 - 3t^2}{t^3}. \tag{90.297}$$

La primitive du membre de droite est presque l'exercice 144.(2). En intégrant (90.297), nous trouvons

$$\ln(y_H) = 3\ln(t) + t^{-2} + C, \tag{90.298}$$

donc

$$y_H(t) = Ke^{3\ln(t)+t^{-2}} = K\left(e^{\ln(t)}\right)^3 e^{t^{-2}} = Kt^3e^{1/t^2}. \tag{90.299}$$

Cela est la solution générale de l'équation homogène associée à notre problème.

En ce qui concerne le problème non homogène (c'est-à-dire avec le second membre), la technique de variation des constantes nous indique de chercher une fonction $K(t)$ pour laquelle la solution serait sous la forme

$$y(t) = K(t)t^3e^{t^{-2}}. \tag{90.300}$$

Afin de trouver la fonction K, nous injectons (90.300) dans l'équation différentielle à résoudre : nous calculons $t^3y' + (2 - 3t^2)y$ et nous imposons que le résultat soit t^3. Pour la dérivée de y nous avons

$$\begin{aligned} y' &= K't^3e^{t^{-2}} + K\left(t^3e^{t^{-2}}\right)' \\ &= K't^3e^{t^{-2}} + K\left(3t^2e^{t^{-2}} + t^3(-2)t^{-3}e^{t^{-2}}\right) \\ &= e^{t^{-2}}\big[K't^3 + K(3t^2 - 2)\big]. \end{aligned} \tag{90.301}$$

Maintenant nous pouvons calculer l'équation

$$\begin{aligned} t^3y' + (2 - 3t^2)y &= t^3e^{t^{-2}}\big[K't^3 + K(3t^2 - 2)\big] + (2 - 3t^2)Kt^3e^{t^{-2}} \\ &= t^3e^{t^{-2}}t^3K'. \end{aligned} \tag{90.302}$$

La fonction K est donc contrainte par la relation $t^3e^{t^{-2}}t^3K' = t^3$, c'est-à-dire

$$K' = e^{-t^{-2}}t^{-3}. \tag{90.303}$$

La fonction K est donc donné par l'intégrale

$$K = \int e^{-t^{-2}}t^{-3}dt \tag{90.304}$$

qui s'obtient en posant $u = t^{-2}$, $dt = \frac{du}{-2t^{-3}}$. Le résultat est que

$$K(t) = \frac{1}{2}\left(e^{-t^{-2}} + C\right) \tag{90.305}$$

Maintenant la solution générale de notre équation différentielle s'écrit

$$y(t) = K(t)t^3e^{t^{-2}} = \frac{t^3 + Ct^3e^{t^{-2}}}{2}. \tag{90.306}$$

Remarquons que cette solution n'est pas valable en $t = 0$.

Exercise 153 exoSVT-0005

Nous vous suggérons de regarder la vidéo
http://www.youtube.com/watch?v=XYjUUGQhehs
qui explique comment s'obtient l'équation différentielle du circuit électrique RC.

Le résultat est que si on impose une tension V aux bornes du circuit comprenant une résistance R et un condensateur de capacité C, l'équation qui régit la tension u_C aux bornes du condensateur est

$$V = RC\frac{du_C}{dt} + u_C. \tag{90.307}$$ EqnmKqqq

La charge sur le condensateur est donnée par $q(t) = Cu_C(t)$.

Déterminer la fonction $u_C(t)$ en résolvant l'équation différentielle (90.307). En déduire la fonction $q(t)$ en supposant que $q(0) = 0$. Calculer $\lim_{t\to\infty} q(t)$. corrSVT-0005

Correction of the exercise 153

Histoire d'avoir des notations qui ressemblent à ce qu'on a d'habitude, nous écrivons y pour u_C et nous mettons l'équation différentielle sous la forme

$$y' = -\frac{y}{RC} + \frac{V}{RC}. \tag{90.308}$$ EqHUbHWZ

Cela est une équation différentielle « presque » à variables séparée. Nous commençons par résoudre l'équation homogène associée

$$y' = -\frac{y}{RC}. \tag{90.309}$$

Nous avons successivement

$$y' = -\frac{y}{RC} \tag{90.310a}$$
$$\frac{dy}{y} = -\frac{1}{RC}dt \tag{90.310b}$$
$$\ln(y) = -\frac{t}{RC} + A \tag{90.310c}$$
$$y(t) = Ke^{-t/RC}. \tag{90.310d}$$

Afin de trouver la solution de l'équation complète nous utilisons la méthode de variation des constantes. Nous posons

$$y(t) = K(t)e^{-t/RC} \tag{90.311}$$

et nous récrivons l'équation (90.308) :

$$K'e^{-t/RC} - \frac{1}{RC}Ke^{-t/RC} = -\frac{1}{RC}Ke^{-t/RC} + \frac{V}{RC} \tag{90.312a}$$
$$K' = \frac{V}{RC}e^{t/RC} \tag{90.312b}$$
$$K = Ve^{t/RC} + B \tag{90.312c}$$

où nous avons noté B la constante d'intégration. La solution complète s'écrit donc

$$y(t) = \big(Ve^{t/RC} + B\big)e^{-t/RC} = V + Be^{-t/RC} \tag{90.313}$$

où B est une constante arbitraire. La fonction $q(t)$ est obtenue en multipliant par C la fonction $u_C(t)$:

$$q(t) = VC + BCe^{-t/RC}. \tag{90.314}$$

Afin d'avoir $q(0) = 0$ nous devons fixer $B = V$, c'est-à-dire

$$q(t) = VC(1 + e^{-t/RC}). \tag{90.315}$$

Cela est la quantité de charges électriques sur le condensateur après un temps t. Nous avons

$$\lim_{t\to\infty} q(t) = VC, \tag{90.316}$$

qui correspond à la formule connue de la charge d'un condensateur lorsque le système est à l'équilibre.

Exercise 154 exoTD6-0001

On considère pour $t \in \mathbb{R}$, l'équation différentielle

$$y'(t) + \sin(t)y(t) = \exp(\cos(t)). \tag{90.317}$$ EqEqdiffexoysvec

(1) Déterminer la primitive F de $t \mapsto \sin(t)$ qui s'annule en $\pi/2$.

(2) Montrer que y est solution de (90.317) si et seulement si $\Big(y(t)\exp\big(F(t)\big)\Big)' = 1$.
La fonction $t \mapsto \exp\big(F(t)\big)$ est appelée *facteur intégrant* associé à (90.317).

(3) En déduire toutes les solutions de (90.317).

(4) Déterminer la solution de (90.317) telle que $y(0) = 0$.

corrTD6-0001

Correction of the exercise 154

(1) Les primitives de la fonction $t \mapsto \sin(t)$ sont les fonctions $F(t) = -\cos(t) + C$ où C est n'importe quelle constante. Il y a une primitive par choix de constante. Nous voulons sélectionner celle qui vérifie $F(\pi/2) = 0$:

$$F(\pi/2) = -\cos(\pi/2) + C = C. \tag{90.318}$$

Afin que cela vaille zéro, nous avons besoin de $C = 0$. La fonction F dont nous parlons dans cet exercice est donc

$$F(t) = -\cos(t). \tag{90.319}$$

(2) Explicitons la condition $(ye^F)' = 1$ en utilisant la règle de dérivation de produit :

$$(ye^F) = y'e^F + yF'e^F = e^F\big(y' + y\sin(t)\big). \tag{90.320}$$

En imposant que cela soit 1, nous trouvons la condition

$$e^{F(t)}\big(y'(t) + y(t)\sin(t)\big) = 1, \tag{90.321}$$

en passant $e^{F(t)}$ de l'autre côté, et en remplaçant $F(t)$ par sa valeur $-\cos(t)$, nous trouvons

$$y'(t) + y(t)\sin(t) = e^{\cos(t)}, \tag{90.322}$$

c'est-à-dire l'équation de départ.

(3) Étant donné que $(ye^F)' = 1$, il existe une constante C telle que $ye^F = t + C$ (c'est la primitive de 1). Nous avons donc comme solution générale :

$$y(t) = e^{\cos(t)}(t + C). \tag{90.323}$$ Eqszzuvecttc

(4) Parmi toutes les solutions données par l'équation (90.323), nous devons sélectionner celle telle que $y(0) = 0$. En posant $t = 0$ dans la formule générale de $y(t)$, nous trouvons que

$$y(0) = 0. \tag{90.324}$$

Pour que cela soit zéro, il faut sélectionner $C = 0$. La solution de l'équation de départ qui s'annule en zéro est donc

$$y(t) = te^{\cos(t)}. \tag{90.325}$$

Exercise 155 exoTD6-0002
Résoudre les équations différentielles suivantes.

(1) $y'(x) = y(x) + x$ pour $x \in \mathbb{R}$.
(2) $x(x+1)y'(x) + (1+x)y(x) = 1$ pour $x \in]-1, 0[$.
(3) $xy'(x) - y(x) = x^2 \arctan(x)$ pour $x \in]0, \infty[$.
(4) $xy'(x) - (x+2)y(x) + x(x+1) = 0$ pour $x \in]-\infty, 0[$.

corrTD6-0002

Correction of the exercise 155
<+CorrTD6-0002+>

Exercise 156 exoTD6-0003
Étude de la suite (u_n) définie par $u_n = \int_0^{\pi/4} \big(\tan(x)\big)^n dx$.

corrTD6-0003

Correction of the exercise 156
<+CorrTD6-0003+>

Exercise 157 exoTD6-0004
Soit m un entier naturel. Calculer l'intégrale $I_m = \int_0^{\pi/2} \big(\sin(t)\big)^m dt$.

corrTD6-0004

Correction of the exercise 157
<+CorrTD6-0004+>

Exercise 158 exoTD6A-0001
Soient α et k deux nombres réels positifs fixés. Vérifier que la fonction

$$x(t) = ke^{\ln\left(\frac{C}{k}\right)e^{-\alpha t}} \tag{90.326}$$

est une solution de l'équation différentielle

$$\frac{dx}{dt} = \alpha x \ln\left(\frac{k}{x}\right), \tag{90.327}$$

pour tout choix de $C > 0$, $C \in \mathbb{R}$.

Calculez la limite de $x(t)$ lorsque t tend vers $+\infty$ et vérifiez que la constante C est la valeur de la fonction $x(t)$ à l'instant $t = 0$.

corrTD6A-0001

Correction of the exercise 158

90.11 Révisions

Exercise 159 ⊖⊖ exorevisions-0001
Calculer

(1) $\log_6(216)$
(2) $\log_3 \frac{1}{9}$
(3) $\log_{16} 4$.

corrrevisions-0001

Correction of the exercise 159
<+Corrrevisions-0001+>

Exercise 160 ⊖⊖ exorevisions-0002
Donner l'ensemble de définition des fonctions suivantes.

(1) $f(x) = \ln(|x|)$
(2) $f(x) = \cos(\frac{1}{x})$
(3) $f(x) = \sqrt{x^2 - 1}$

corrrevisions-0002

Correction of the exercise 160

<+Corrrevisions-0002+>

Exercise 161 ⊖⊖ exorevisions-0003

Tracer les fonctions suivantes :

(1) $f(x) = \ln(|x|)$

(2) $f(x) = \cos(x+1)$

(3) $f(x) = \cos(x) + 1$

corrrevisions-0003

Correction of the exercise 161

<+Corrrevisions-0003+>

Exercise 162 ⊖⊖ exorevisions-0004

Calculer les limites

(1) $\lim_{x\to 2} \dfrac{x^2+x-6}{4(x^2-4)}$

(2) $\lim_{x\to\infty} \ln(x+1)$

(3) $\lim_{x\to 0} \dfrac{\sin(x)}{x}$

(4) $\lim_{x\to\infty} x^3 - 2x^2$.

corrrevisions-0004

Correction of the exercise 162

<+Corrrevisions-0004+>

90.12 Anciennes interrogations

Pour rappel, toutes les interrogations, devoirs surveillés et examens se font avec uniquement du papier et de quoi écrire (pas de notes, pas de calculatrices ou autre équipements). Les réponses doivent être justifiées un minimum.

90.12.1 Septembre 2010

Exercise 163 exointerro-0002

(1) Dessiner le graphe de la fonction $x \mapsto f(x)$ dans les cas suivants :

(1a) $f(x) = ex$;

(1b) $f(x) = \begin{cases} x^2+1 & \text{si } x \leqslant 0, \\ (x+1)^2 & \text{si } x > 0. \end{cases}$;

(1c) $f(x) = |x|$;

(1d) $f(x) = \log_2(2^{x-1})$.

(2) Donner un exemple de fonction paire et un exemple de fonction impaire.

corrinterro-0002

Correction of the exercise 163

Les graphes sont tracés sur la figure 90.8.

(1) (1a) La constante e est juste un nombre (qui vaut approximativement 2.718). La fonction est donc une simple droite de coefficient angulaire e.

(1b) Remarquez que à droite, elle est plus pentue qu'à gauche.

(1c) La fonction valeur absolue est une fonction de base.

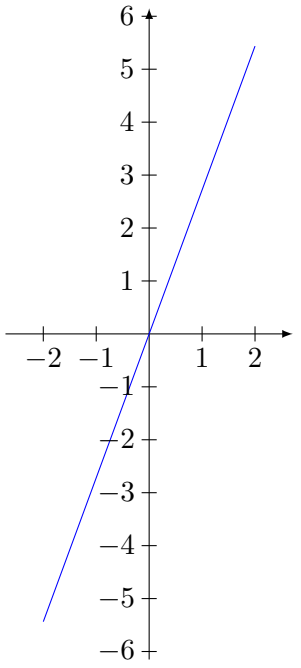
(1d) Nous avons toujours $\log_a(a^b) = b$. Donc ici il s'agit seulement de la fonction $y = x - 1$ qui est une droite.

(2) De nombreux exemples sont possibles. Citons les plus courantes.

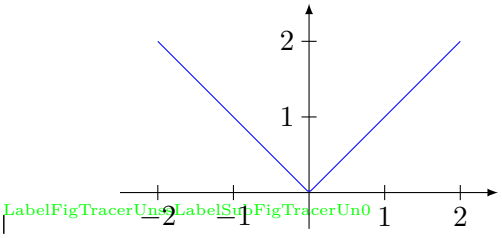
Fonctions paires $\cos(x)$, x^2, x^4, toutes les fonctions constantes.

Fonctions impaires $\sin(x)$, x^3, x, x^5, la fonction constante zéro.

Notez que la fonction constante $y = 0$ est une fonction paire et impaire en même temps. La fonction $x^3 + 1$ n'est pas une fonction impaire.

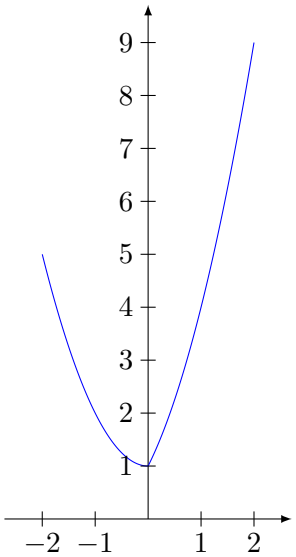


(a) La fonction $y = ex$

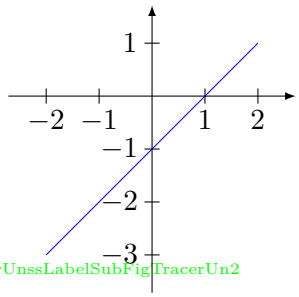


LabelFigTracerUnssLabelSubFigTracerUn0 LabelFigTracerUnssLabelSu

(b) La fonction $y = |x|$



(c) La fonction $y = x^2 + 1$ à gauche et $(x + 1)^2$ à droite.



LabelFigTracerUnssLabelSubFigTracerUn2 LabelFigTracerUnssLabelSubFigT

(d) La fonction $y = \log_2(2^{x-1})$

FIGURE 90.8: Les fonctions de la question 163.

LabelFigTracerUn

Exercise 164 exointerro-0003

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to 3} \dfrac{x^2-9}{x^2-2x-3}$;

(2) $\lim_{x\to+\infty} \dfrac{x^2-2x}{2x-3}$;

(3) $\lim_{x\to 0} \dfrac{1-\cos x}{\sin^2 x}$;

(4) $\lim_{x\to+\infty} \left(1+\dfrac{3}{x}\right)^x$. Conseil : essayer la substitution $y = x/3$;

(5) $\lim_{x\to 0} \dfrac{x}{\sqrt{x+4}-2}$.

corrinterro-0003

Correction of the exercise 164

(1) En remplaçant, nous trouvons $\frac{0}{0}$, donc il y a une indétermination à lever. Pour cela, nous factorisons le numérateur et le dénominateur. Le numérateur est un produit remarquable : $a^2-b^2=(a+b)(a-b)$, et donc $x^2-9=(x+3)(x-3)$. En ce qui concerne le dénominateur, nous avons x^2-2x-3 pour

$$x=\frac{2\pm\sqrt{4-4\cdot 1\cdot(-3)}}{2}=\frac{2\pm\sqrt{16}}{2}=\frac{2\pm 4}{2}, \tag{90.328}$$

et donc $x=3$ ou $x=-1$.
Par conséquent le dénominateur s'écrit

$$x^2-2x-3=(x-3)(x+1), \tag{90.329}$$

et $(x-3)$ se simplifie dans la fraction. Par conséquent nous devons calculer la limite

$$\lim_{x\to 3}\frac{x+3}{x+1}=\frac{6}{2}=3. \tag{90.330}$$

(2) Nous mettons x^2 en évidence au numérateur et x au dénominateur. Cela donne

$$\frac{x^2-2x}{2x-3}=\frac{x^2\left(1-\frac{2}{x}\right)}{x\left(2-\frac{3}{x}\right)}. \tag{90.331}$$

Lorsque nous prenons la limite pour $x\to\infty$, les fractions $\frac{2}{x}$ et $\frac{3}{x}$ tendent vers zéro, tandis qu'après simplification par x, il reste

$$\lim_{x\to\infty}\frac{x}{2}=\infty. \tag{90.332}$$

(3) En remplaçant, nous trouvons $\frac{0}{0}$, ce qui est une indétermination à lever. Ici le truc est de multiplier et diviser par le binôme conjugué $(1+\cos(x))$:

$$\begin{aligned}\frac{1-\cos(x)}{\sin^2(x)}&=\frac{\big(1-\cos(x)\big)\big(1+\cos(x)\big)}{\sin^2(x)\big(1+\cos(x)\big)}\\&=\frac{1-\cos^2(x)}{\sin^2(x)\big(1+\cos(x)\big)}\\&=\frac{1}{1+\cos(x)}\end{aligned} \tag{90.333}$$

où nous avons utilisé le fait que $1-\cos^2(x)=\sin^2(x)$, et simplifié par $\sin^2(x)$. Maintenant nous pouvons conclure le calcul :

$$\lim_{x\to 0}\frac{1}{1+\cos(x)}=\frac{1}{2}. \tag{90.334}$$

(4) Lorsque nous posons $y = \frac{x}{3}$, nous remplaçons $\frac{3}{x}$ par $\frac{1}{y}$ et x par $3y$. D'autre part, lorsque x tend vers l'infini, la nouvelle variable y tend aussi vers l'infini, donc nous avons

$$\lim_{y\to\infty}\left(1+\frac{1}{y}\right)^{3y}=\left[\left(1+\frac{1}{y}\right)^{y}\right]^{3}. \tag{90.335}$$

La limite de ce qu'il y a dans le crochet est une limite connue : elle vaut e. Donc nous avons

$$\lim_{y\to\infty}\left[\left(1+\frac{1}{y}\right)^{y}\right]^{3}=e^{3} \tag{90.336}$$

(5) Ici encore le truc est de multiplier et diviser par la binôme conjugué de $\sqrt{x+4}-2$. Nous avons

$$\frac{x}{\sqrt{x+4}-2}=\frac{x\big(\sqrt{x+4}+2\big)}{\big(\sqrt{x+4}-2\big)\big(\sqrt{x+4}+2\big)}=\frac{x\big(\sqrt{x+4}+2\big)}{x+4-4}. \tag{90.337}$$

Après simplification par x, nous restons avec le calcul

$$\lim_{x\to 0}\sqrt{x+4}+2=4. \tag{90.338}$$

Exercise 165 exointerro-0004

Calculer :

(1) $\log_9 3$;

(2) $\log_{10} 100$;

(3) $\log_2 \frac{1}{4}$.

corrinterro-0004

Correction of the exercise 165

Pour rappel la définition du logarithme est $\log_a(x) = y$ lorsque $a^y = x$. Encore pour rappel, nous avons $\sqrt{a} = a^{1/2}$, et $\frac{1}{a} = a^{-1}$.

(1) Nous aurons $\log_9(3) = y$ si $9^y = 3$. Quelle valeur de y ? Nous savons que $\sqrt{9} = 3$, et donc $9^{1/2} = 3$, c'est-à-dire que $\log_9(3) = \frac{1}{2}$.

(2) Nous aurons $\log_{10}(100) = y$ si $10^y = 100$. Or $10^2 = 100$, donc $\log_{10}(100) = 2$.

(3) Ici, nous devons nous rendre compte que $\frac{1}{4} = 4^{-1}$. Donc nous restons avec la question : pour que y est-ce que $2^y = 4^{-1}$? La réponse est $y = -2$, donc $\log_2 \frac{1}{4} = -2$.

Exercise 166 exointerro-0005

Déterminer l'ensemble de définition de la fonction $x \mapsto f(x)$ dans les cas suivants :

(1) $f(x) = \ln(x-2)$;

(2) $f(x) = \frac{1}{\pi x}$;

(3) $f(x) = \frac{1}{\ln(x)}$;

(4) $f(x) = \sqrt{x+1}$.

corrinterro-0005

Correction of the exercise 166

(1) Ce qui se trouve dans le logarithme doit être positif (strictement), donc ici $x - 2 > 0$, et par conséquent $x > 2$. L'ensemble de définition est donc $]2, \infty[$.

(2) Le dénominateur doit être non nul, donc $\pi x \neq 0$, c'est-à-dire $x \neq 0$. Le domaine de définition est donc $\mathbb{R}_0$.

(3) Ici nous avons deux éléments à problèmes. D'abord nous avons un dénominateur qui demande $\ln(x) \neq 0$, et ensuite nous avons un logarithme qui demande $x > 0$. Le domaine de la fonction sont les x qui satisfont aux deux conditions en même temps. Nous avons $\ln(x) = 0$ si $x = 1$, donc le domaine sera tous les x positifs différents de 1 : $]0, \infty[\backslash\{1\}$.

(4) Nous avons une racine, donc ce qui se trouve dedans doit être positif (ou nul). Nous posons donc $x+1 \geqslant 0$, et par conséquent $x \geqslant -1$. Le domaine de définition est par conséquent le domaine est $[1, \infty[$.

Exercise 167 exointerro-0007

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to-\infty} x^3$;

(2) $\lim_{x\to+\infty} \frac{x-2}{2x-3}$;

(3) $\lim_{x\to 0} \frac{x-2}{2x-3}$;

(4) $\lim_{x\to 2} \frac{x^2-4}{(x-2)(x+1)}$;

(5) $\lim_{x\to+\infty} \frac{\sin(x)}{2}$.

corrinterro-0007

Correction of the exercise 167

(1) Il n'y a pas d'indétermination : lorsque x tend vers l'infini, x^3 tend vers $-\infty$.

(2) Mise en évidence de x au numérateur et au dénominateur et simplification par x :

$$\frac{x-2}{2x-3} = \frac{x\left(1-\frac{2}{x}\right)}{x\left(2-\frac{3}{x}\right) = \frac{1-\frac{2}{x}}{2-\frac{3}{x}}}. \tag{90.339}$$

Lorsque nous faisons la limite $x \to \infty$, les fractions $\frac{2}{x}$ et $-\frac{3}{x}$ tendent vers zéro et il ne reste que

$$\lim_{x\to\infty} \frac{1-\frac{2}{x}}{2-\frac{3}{x}} = \frac{1}{2}. \tag{90.340}$$

(3) Ici par contre, il n'y a pas d'indéterminations : si on remplace x par zéro, nous obtenons immédiatement $\frac{2}{3}$.

(4) En remplaçant, nous avons $\frac{0}{0}$, et par conséquent une indétermination à lever. Nous pouvons factoriser le numérateur en $x^2-4=(x+2)(x-2)$ et simplifier la fraction par $x-2$:

$$\frac{x^2-4}{(x-2)(x+1)} = \frac{(x+2)(x-2)}{(x-2)(x+1)} = \frac{x+2}{x+1}, \tag{90.341}$$

et

$$\lim_{x\to 2} \frac{x+2}{x+1} = \frac{4}{3}. \tag{90.342}$$

(5) La fonction $\frac{\sin(x)}{2}$ est une fonction périodique et oscille donc tout le temps. Elle n'a pas de limite.

Exercise 168 exointerro-0008

Soient $f(x)=x^2-2x$ et $g(x)=1/(1+x)$. Écrivez les expressions explicites et trouvez les domaines des fonctions composées $h_1(x)=f\circ g(x)$ et $h_2(x)=g\circ f(x)$.

corrinterro-0008

Correction of the exercise 168

(1) Nous avons

$$(f\circ g)(x) = f\big(g(x)\big) = g(x)^2 - 2g(x) = \frac{1}{(1+x)^2} - \frac{2}{1+x}. \tag{90.343}$$

Le domaine de définition est donné par $x \neq 1$ à cause du dénominateur.

(2) Nous avons

$$(g\circ f)(x) = g\big(f(x)\big) = \frac{1}{1+f(x)} = \frac{1}{x^2-2x+1}. \tag{90.344}$$

Le domaine de définition est donné par $\mathbb{R}$ moins les x tels que $x^2-2x+1=0$. Étant donné que $x^2-2x+1=(x-1)^2$, le domaine est $x \neq 1$.

90.12.2 DS octobre 2010, un

Exercise 169 exoDS2010-1-0001

Soient $f(x) = \sqrt{x-1}$ et $g(x) = \dfrac{1}{\ln(x)}$.

(1) Trouvez les domaines de définition de f et de g.

(2) Écrivez les expressions explicites et trouvez les domaines de définition des fonctions composées $h_1(x) = f \circ g(x)$ et $h_2(x) = g \circ f(x)$.

corrDS2010-1-0001

Correction of the exercise 169

(1) La racine $\sqrt{x-1}$ demande $x - 1 \geqslant 0$, donc pour la fonction f, le domaine est $x \in [1, +\infty[$.
En ce qui concerne la fonction g, il faut penser à deux choses. La première est le dénominateur : $\ln(x) \neq 0$. La seconde est le logarithme : $x > 0$.
La première condition demande $x \neq 1$. Donc en combinant les deux, nous avons tous les nombres strictement plus grands que zéro, sauf 1 : $x \in]0, \infty[\backslash\{1\}$.

(2) (2a) D'abord nous avons

$$(f \circ g)(x) = f\big(g(x)\big) = \sqrt{g(x) - 1} = \sqrt{\frac{1}{\ln(x)} - 1}. \tag{90.345}$$

Ensuite nous regardons les conditions pour le domaine. Il y en a trois parce qu'il y a une racine carrée, un dénominateur et un logarithme.

(2a.i) Pour la racine carrée, $\frac{1}{\ln(x)} - 1 \geqslant 0$, ce qui donne l'inéquation

$$\frac{1}{\ln(x)} \geqslant 1. \tag{90.346}$$

Attention : le logarithme peut être négatif. Il n'est donc pas exact de faire directement passer le $\ln(x)$ de l'autre côté :

$$1 \geqslant \ln(x). \tag{90.347}$$

Ce passage n'est vrai que si $\ln(x) > 0$. Mais si $\ln(x) < 0$, nous n'avons certainement pas $\frac{1}{\ln(x)} \geqslant 1$. Par conséquent, les conditions sont $1 \geqslant \ln(x)$ **et** $\ln(x) > 0$. Cela donne $x > 1$ et $x \leqslant e$.

(2a.i) Pour le dénominateur, $\ln(x) \neq 0$ implique $x \neq 1$.

(2a.i) Pour le logarithme, nous devons avoir $x > 0$.

Maintenant nous devons regarder l'intersection de ces trois conditions. Nous devons avoir $x > 1$, $x > 0$, $x \leqslant e$ et $x \neq 1$. Cela fait donc

$$x \in]1, e]. \tag{90.348}$$

(2b) Nous avons

$$(g \circ f)(x) = g\big(f(x)\big) = \frac{1}{\ln\big(f(x)\big)} = \frac{1}{\ln\big(\sqrt{x-1}\big)}. \tag{90.349}$$

Encore une fois, nous avons trois conditions : un dénominateur, un logarithme et une racine carrée.

(2b.ii) Pour le dénominateur, il faut $\ln\big(\sqrt{x-1}\big) \neq 0$, ce qui donne $\sqrt{x-1} \neq 1$, et donc $x - 1 \neq 1$, et finalement $x \neq 2$.

(2b.ii) Pour la racine, nous demandons $x - 1 \geqslant 0$, et par conséquent $x \geqslant 1$.

(2b.ii) Pour le logarithme, nous avons besoin de $\sqrt{x-1} > 0$, c'est-à-dire $x > 1$. Pour rappel, la valeur d'une racine carrée n'est jamais négative, mais elle est nulle sur l'argument est zéro.

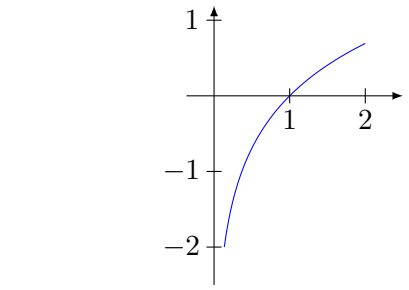
En remettant tout ensemble, nous devons avoir simultanément $x > 1$, $x \neq 2$ et $x \geqslant 1$, c'est-à-dire

$$x \in]1, \infty[\backslash\{2\}. \tag{90.350}$$

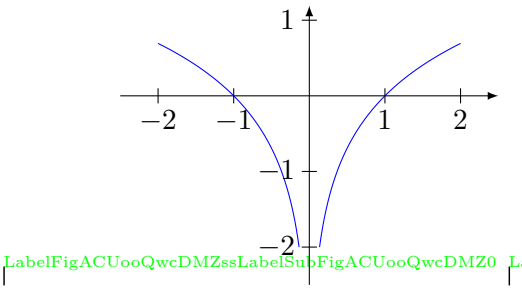
Exercise 170 exoDS2010-1-0002

Indiquez le graphe correspondant à chacune des fonctions suivantes (figure 90.9). Justifiez vos réponses

(1) $\ln(x)$;
(2) $\ln(x^2)$;
(3) $\ln(|x|)$;
(4) $|\ln(x)|$;
(5) $\ln(x+1)$;
(6) $\ln(x)+1$;
(7) $\sqrt{\ln(x)}$.

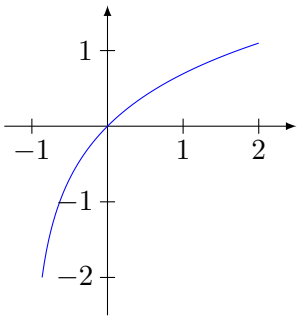


(a) logarithme

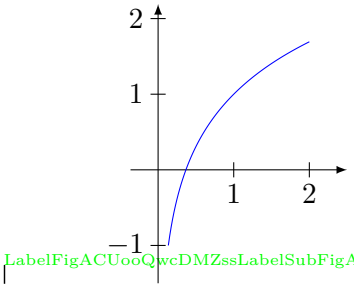


LabelFigACUooQwcDMZssLabelSubFigACUooQwcDMZ0 LabelFigACUooQwcDMZssLabelSubFigACU

(b) lnvalabsolue

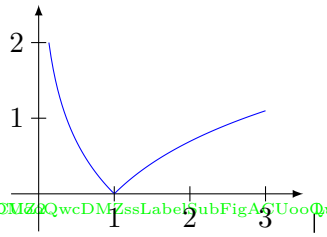


(c) lnxplusun

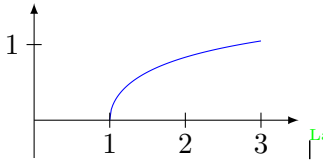


LabelFigACUooQwcDMZssLabelSubFigACUooQwcDMZ2 LabelFigACUooQwcDMZssLabelSubFigACUooQwcDMZ

(d) valabsolueln

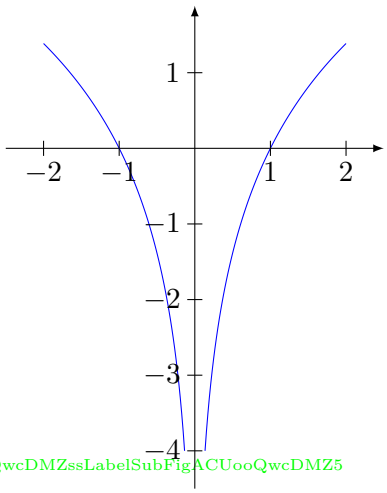


(e) unplusln



LabelFigACUooQwcDMZssLabelSubFigACUooQwcDMZ5 LabelFigACUooQwcDMZssLabelSubFig

(f) sqrtln



(g) lnsqrt

FIGURE 90.9: Les graphes à considérer de la question 170. LabelFigACUooQwcDMZ corrDS2010-1-0002

Correction of the exercise 170

Chaque graphique de la figure 90.9 a sa petite particularité qui lui permet de se distinguer des autres.

(1) Le graphique 90.9(a) est le graphique bien reconnaissable de $\ln(x)$: en fait le domaine de cette fonction est $]0,+\infty[$ et $\ln(1)=0$.

(2) Le graphique 90.9(b) est le même que le logarithme, mais symétrique par rapport à l'axe vertical. C'est donc $\ln(|x|)$.

(3) Le graphique 90.9(c) est le même que $\ln(x)$, mais décalé de 1 vers la gauche (il tend vers $-\infty$ en -1 au lieu de zéro). C'est donc le $\ln(x+1)$.

(4) Le graphique 90.9(d) est le logarithme décalé de 1 vers le haut, c'est donc $\ln(x)+1$.

(5) Le graphique 90.9(e) est $|\ln(x)|$ parce que c'est le même que le logarithme pour $x\geqslant 1$, et il est le symétrique par rapport à l'axe horizontal de $\ln(x)$ là où $\ln(x)$ est négatif.

(6) Le graphique 90.9(f) est $\sqrt{\ln(x)}$. La façon de le voir est de remarquer que le graphe commence en 1. Il n'existe pas avant $x=1$, ce qui signifie qu'il n'existe pas là où $\ln(x)$ est négatif.

(7) Le graphique 90.9(g) est $\ln(x^2)$ parce qu'il est symétrique, mais plus pentu que $\ln(|x|)$.

Notez que $\ln(x^2)$ et $\ln(|x|)$ se ressemblent parce que tant le carré que la valeur absolue on pour effet de « rendre positif les nombres négatifs ». Cependant comme x^2 avance plus vite que x, c'est lui qui est le plus penché.

Exercise 171 exoDS2010-1-0003

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to 4}\frac{(x-4)^2}{x^2-16}$;

(2) $\lim_{x\to-\infty}\frac{x^{36}-2x^{25}-12}{2x^{27}-3x^{12}+1200}$;

(3) $\lim_{x\to 0}\frac{\sin(2x)}{x}$;

(4) $\lim_{x\to+\infty}\frac{\sin(x)}{x}$;

(5) $\lim_{x\to 0}\frac{\sqrt{x}}{\sqrt{4-x}-\sqrt{4+x}}$;

(6) $\lim_{x\to+\infty}\frac{\ln(x)}{x^{1/1234}}$.

corrDS2010-1-0003

Correction of the exercise 171

Pour les limites des fonctions rationnelles (c'est-à-dire, des fractions de polynômes), il y a deux techniques à ne pas confondre. La première est la factorisation et simplification ; la seconde est la mise en évidence du plus haut degré au numérateur et au dénominateur.

La mise en évidence se fait lorsqu'on a des limites en $\pm\infty$. La factorisation se fait pour les limites en des nombres.

(1) Ici il faut factoriser le dénominateur et simplifier.

$$\lim_{x\to 4}\frac{(x-4)^2}{x^2-16}=\lim_{x\to 4}\frac{(x-4)^2}{(x-4)(x+4)}=\lim_{x\to 4}\frac{(x-4)}{(x+4)}=\frac{0}{8}=0. \tag{90.351}$$

(2) Ici, il faut mettre en évidence la plus haute puissance de x au numérateur et au dénominateur.

$$\begin{aligned}\lim_{x\to-\infty}\frac{x^{36}-2x^{25}-12}{2x^{27}-3x^{12}+1200}&=\lim_{x\to-\infty}\frac{x^{36}\left(1-\frac{2x^{25}}{x^{36}}-\frac{12}{x^{36}}\right)}{x^{27}\left(2-\frac{3x^{25}}{x^{27}}-\frac{12}{x^{27}}\right)}\\&=\frac{x^9\left(1-\frac{2}{x^{11}}-\frac{12}{x^{36}}\right)}{2-\frac{3}{x^{15}}+\frac{1200}{x^{27}}}\\&=\lim_{x\to-\infty}\frac{x^9}{2}\\&=-\infty.\end{aligned} \tag{90.352}$$

(3) La limite que l'on connaît est celle $\lim_{x\to 0}\frac{\sin(2x)}{2x}=1$. Ici le problème est que nous avons juste x au dénominateur. L'astuce est de multiplier et diviser par 2 :

$$\lim_{x\to 0}\frac{\sin(2x)}{x}=\lim_{x\to 0}\frac{2\sin(2x)}{2x}=2. \tag{90.353}$$

(4) La fonction $\sin(x)$ est bornée, et nous la divisions par quelque chose qui tend vers l'infini. La limite est donc 0. Plus formellement on peut dire

$$0\leqslant\left|\frac{\sin(x)}{x}\right|\leqslant\frac{1}{|x|}, \tag{90.354}$$

tandis que $\frac{1}{x}\to 0$ si $x\to\infty$.

(5) Ici l'astuce est de multiplier et diviser par le « binôme conjugué » du dénominateur, à savoir

$$\sqrt{4-x}+\sqrt{4+x}. \tag{90.355}$$

Cela fait apparaitre au dénominateur le produit

$$(\sqrt{4-x}-\sqrt{4+x})(\sqrt{4-x}+\sqrt{4+x})=(4-x)-(4+x) \tag{90.356}$$

par le produit remarquable $(\spadesuit-\clubsuit)(\spadesuit+\clubsuit)=\spadesuit^2-\clubsuit^2$. La limite à calculer est donc

$$\lim_{x\to 0}\frac{\sqrt{x}\left(\sqrt{4-x}+\sqrt{4+x}\right)}{(4-x)-(4+x)}=\lim_{x\to 0}\frac{4\sqrt{x}}{-2x}=\lim_{x\to 0}\frac{-2}{\sqrt{x}}=-\infty. \tag{90.357}$$

(6) Ici nous utilisons le fait que « le logarithme est le moins fort ». Donc le logarithme du numérateur tend moins vite vers l'infini que la puissance de x qui est au dénominateur. Nous avons donc

$$\lim_{x\to\infty}\frac{\ln(x)}{x^{1/1234}}=0. \tag{90.358}$$

Exercise 172 exoDS2010-1-0004
Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par

$$\begin{cases} u_{n+1}=u_n+2^n,\\ u_0=0. \end{cases} \tag{90.359}$$

(1) Démontrer que $u_n=2^n-1$.

(2) Calculer, si elle existe, la limite de la suite $(u_n)_n$.

corrDS2010-1-0004

Correction of the exercise 172

(1) Il faut prouver par récurrence. D'abord la formule proposée est vraie pour $n=0$ parce que $2^0-1=0=u_0$ (pour rappel, $a^0=1$ pour tout $a\neq 0$). Ensuite, il faut vérifier que la formule est correcte pour $k+1$ en supposant qu'elle soit correcte pour k. En posant $n=k+1$ dans la formule à démontrer, nous avons

$$u_{k+1}=2^k+u_k=2^k+2^k-1=2^{k+1}-1, \tag{90.360}$$

ce qui est bien la formule avec $n=k+1$.

(2) La limite est donc $\lim_{n\to\infty}u_n=\infty$ parce que $2^n\to\infty$; c'est une exponentielle.

Exercise 173 exoDS2010-1-0005
Considérons la suite définie par

$$u_n=(-1)^n+\frac{1}{n}, \tag{90.361}$$

pour $n\geqslant 1$, $n\in\mathbb{N}$.

(1) Calculer la valeur des termes u_1, u_2, u_3 et u_4.

(2) Démontrer que la suite $(u_n)_{n\in\mathbb{N}*}$ est bornée.

(3) Est-ce que la suite $(u_n)_{n\in\mathbb{N}*}$ est monotone ?

(4) On admet que la suite $(|u_n|)_{n\in\mathbb{N}*}$ est convergente. Calculer sa limite.

(5) Calculer la valeur des premiers 3 termes des sous-suites $(u_{2n})_{n\in\mathbb{N}*}$ (sous-suite des indices pairs) et $(u_{2n-1})_{n\in\mathbb{N}*}$ (sous-suite des indices impairs), pour $n \geqslant 1$.

(6) Est-ce que la suite $(u_n)_{n\in\mathbb{N}*}$ est convergente ? Et les suites $(u_{2n})_{n\in\mathbb{N}*}$ et $(u_{2n-1})_{n\in\mathbb{N}*}$?

corrDS2010-1-0005

Correction of the exercise 173

(1) Pour rappel,

$$(-1)^n = \begin{cases} 1 & \text{si } n \text{ est pair} \\ -1 & \text{si } n \text{ est impair.} \end{cases} \tag{90.362}$$

Les premiers termes sont donc

$$\begin{aligned} u_1 &= (-1) + 1 = 0 \\ u_2 &= (-1)^2 + \frac{1}{2} = \frac{3}{2} \\ u_3 &= (-1)^3 + \frac{1}{3} = -\frac{2}{3} \\ u_4 &= (-1)^4 + \frac{1}{4} = \frac{5}{4}. \end{aligned} \tag{90.363}$$

(2) Comme $1/n$ est surement compris entre 0 et 1, la suite $(u_n)_n$ est majorée par 2 et minorée par -1.
Méthode alternative : nous avons

$$\left|(-1)^n + \frac{1}{n}\right| \leqslant |(-1)^n| + \left|\frac{1}{n}\right| \leqslant 1 + \frac{1}{n}. \tag{90.364}$$

Or $\frac{1}{n}$ est une suite qui tend vers zéro, ce qui fait que $1 + \frac{1}{n}$ est bornée.

(3) La suite est croissante entre u_1 et u_2, et décroissante entre u_2 et u_3. Plus généralement, les termes pairs seront positifs tandis que les termes impairs seront négatifs ; elle n'arrête donc pas de monter et descendre au dessus de zéro et en dessous de zéro.

(4) La valeur absolue de u_n est

$$\left|(-1)^n + \frac{1}{n}\right| = 1 + \frac{(-1)^n}{n}. \tag{90.365}$$

La suite $\frac{(-1)^n}{n}$ tend vers 0, comme on a vu au cours. La limite de $(|u_n|)_n$ est donc 1.
Méthode alternative : si la suite $(|u_n|)$ converge, elle converge vers la même limite que n'importe quelle de ses sous-suites. La sous-suite des termes pairs est facile :

$$|u_{2n}| = \left|(-1)^{2n} + \frac{1}{n}\right| = 1 + \frac{1}{n}, \tag{90.366}$$

et cela tend vers 1.

(5)

$$\begin{aligned} u_2 &= \frac{3}{2}, \quad u_1 = 0, \\ u_4 &= \frac{5}{4}, \quad u_3 = -\frac{2}{3}, \\ u_6 &= \frac{7}{6}, \quad u_5 = -\frac{4}{5}. \end{aligned} \tag{90.367}$$

(6) La suite des termes pairs est (rappel : pour tout n, le nombre $2n$ est pair et donc $(-1)^{2n} = 1$)

$$u_{2n} = (-1)^{2n} + \frac{1}{2n} = 1 + \frac{1}{2n}. \tag{90.368}$$

Cette suite tend vers 1 parce que $\frac{1}{2n}$ tend vers zéro.
La suite des termes impairs par contre vaut

$$u_{2n+1} = (-1)^{2n+1} + \frac{1}{2n+1} = -1 + \frac{1}{2n+1}, \tag{90.369}$$

et cette suite tend vers -1.
Nous avons donc trouvé deux sous-suites de (u_n) qui tendent vers des limites différentes. La suite (u_n) elle-même ne converge donc pas.

90.12.3 DS octobre 2010, deux

Exercise 174 exoDS2010bis-0001
Soient $f(x) = \dfrac{1}{1-x^2}$ et $g(x) = \ln(x)$.

(1) Trouvez les domaines de définition de f et de g.
(2) Écrivez les expressions explicites et trouvez les domaines de définition des fonctions composées $h_1(x) = f \circ g(x)$ et $h_2(x) = g \circ f(x)$.

corrDS2010bis-0001

Correction of the exercise 174

(1) Pour la fonction f, nous avons un dénominateur qui ne peut pas s'annuler : $1 - x^2 \neq 0$, c'est-à-dire $x^2 \neq 1$. Le domaine est donc

$$\text{Domaine}(f) = \mathbb{R}\backslash\{-1, 1\}. \tag{90.370}$$

Pour rappel, $x^2 = 1$ lorsque $x = 1$ ou $x = -1$.
En ce qui concerne la fonction g, la seule condition est que $x > 0$ à cause du logarithme.

(2) (2a) Pour $f \circ g$, nous avons

$$(f \circ g)(x) = f\big(g(x)\big) = \frac{1}{1 - g(x)^2} = \frac{1}{1 - \ln(x)^2}. \tag{90.371}$$

Dans cette expression nous voyons une fraction et un logarithme. Il y aura donc deux conditions.

(2a.i) Pour le dénominateur, $1 - \ln(x)^2 \neq 0$. Cela demande $\ln(x)^2 \neq 1$, et donc les deux conditions

$$\begin{aligned} \ln(x) &\neq 1 \\ \ln(x) &\neq -1. \end{aligned} \tag{90.372}$$

La première demande $x \neq e$, tandis que la seconde demande $x \neq -e$.

(2a.i) Pour le logarithme, la condition est $x > 0$.

En résumé, nous devons avoir $x > 0$ en même temps que $x \neq e$ et $x \neq -e$. Le domaine de $f \circ g$ est donc

$$\text{Domaine}\big(f \circ g\big) =]0, \infty[\backslash\{e\}. \tag{90.373}$$

Notez que la condition $x \neq -e$ est redondante par rapport à la condition $x > 0$.

(2b) Pour $g \circ f$, nous avons

$$(g \circ f)(x) = g\big(f(x)\big) = \ln\big(f(x)\big) = \ln\left(\frac{1}{1-x^2}\right). \tag{90.374}$$

Ici nous voyons une fraction et un logarithme. Il y aura donc deux conditions.

(2b.ii) Pour la fraction, le dénominateur ne peut pas s'annuler : $1-x^2\neq 0$. Cela demande $x\neq 1$ et $x\neq -1$.

(2b.ii) Pour le logarithme, nous devons avoir $\frac{1}{1-x^2}>0$, c'est-à-dire $1-x^2>0$.

Le domaine est donc

$$\text{Domaine}(g\circ f)=]-1,1[. \tag{90.375}$$

Exercise 175 exoDS2010bis-0002

Indiquez le graphe correspondant à chacune des fonctions suivantes (figure 90.10). Justifiez vos réponses

(1) $\cos(x)$;
(2) $\cos(x+\pi/2)$;
(3) $\cos(e^x)$;
(4) $\cos(x)+1$;
(5) $\cos(4x)$;
(6) $|\cos(x)|$;
(7) $\sqrt{\cos(x)}$.

corrDS2010bis-0002

Correction of the exercise 175

Chaque fonction a sa particularité qu'il faut reconnaitre.

(1) La fonction $\cos(x)$ elle-même est celle qui vaut 1 en 0, qui s'annule en $\frac{\pi}{2}$ et qui oscille. C'est donc le graphe 90.10(a).

(2) La fonction $\cos(x+\frac{\pi}{2})$ est la même que la fonction cosinus, mais décalée de $\frac{\pi}{2}$ vers la gauche. C'est le graphe 90.10(d).

(3) La fonction $\cos(e^x)$ est une fonction qui oscille de plus en plus vite parce que ce qui se trouve dans le cosinus (c'est-à-dire e^x) monte de plus en plus vite. Le graphe qui correspond est 90.10(e).

(4) Le graphe de la fonction $\cos(x)+1$ est le même que celui de $\cos(x)$, mais décalé de 1 vers le haut. C'est le graphe 90.10(c).

(5) La fonction $\cos(4x)$ oscille quatre fois plus vite que le cosinus (parce que $4x$ avance 4 fois plus vite que x). C'est donc le graphe 90.10(g).

(6) La fonction $|\cos(x)|$ est la même que $\cos(x)$ sauf que partout où $\cos(x)$ est négatif, il devient positif. C'est le graphe 90.10(b)

(7) La fonction $\sqrt{\cos(x)}$ est reconnaissable au fait qu'elle n'est pas définie là où le cosinus est négatif. Il y a donc des « trous » dans son domaine. C'est le graphe 90.10(f)

Exercise 176 exoDS2010bis-0003

Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim_{x\to 3}\frac{(x-3)(x+15)}{x^2-9}$;
(2) $\lim_{x\to-\infty}\frac{x^{50}-3200x^5+1}{5x^{17}-3x+4}$;
(3) $\lim_{x\to 0}\frac{\sin(x)}{x}$;
(4) $\lim_{x\to+\infty} x\sin(\frac{1}{x})$. Conseil : essayer la substitution $y=1/x$;
(5) $\lim_{x\to 0}\frac{\sqrt{1-x^2}-\sqrt{1+x^2}}{5x}$;
(6) $\lim_{x\to+\infty} x^{-3}\ln(x)$.

corrDS2010bis-0003

Correction of the exercise 176

Pour les limites des fonctions rationnelles (c'est-à-dire : des fractions de polynômes), il y a deux techniques à ne pas confondre. La première est la factorisation et simplification ; la seconde est la mise en évidence du plus haut degré au numérateur et au dénominateur.

La mise en évidence se fait lorsqu'on a des limites en $\pm\infty$. La factorisation se fait pour les limites en des nombres.

(1) Ici il faut factoriser le dénominateur et simplifier.

$$\lim_{x\to 3}\frac{(x-3)(x+15)}{x^2-9}=\lim_{x\to 3}\frac{(x-3)(x+15)}{(x-3)(x+3)}=\lim_{x\to 3}\frac{x+15}{x+3}=\frac{18}{6}=3. \tag{90.376}$$

(2) Ici, il faut mettre en évidence la plus haute puissance de x au numérateur et au dénominateur.

$$\begin{aligned}\lim_{x\to-\infty}\frac{x^{50}-3200x^5+1}{5x^{17}-3x+4}&=\\&=\lim_{x\to-\infty}\frac{x^{50}\left(1-\frac{3200}{x^{45}}+\frac{1}{x^{50}}\right)}{x^{17}\left(5-\frac{3}{x^{16}}+\frac{4}{x^{17}}\right)}=\\&=\lim_{x\to-\infty}\frac{x^{33}\left(1-\frac{3200}{x^{45}}+\frac{1}{x^{50}}\right)}{\left(5-\frac{3}{x^{16}}+\frac{4}{x^{17}}\right)}=\\&=-\infty.\end{aligned} \tag{90.377}$$

(3) On connaît la limite remarquable $\lim_{x\to 0}\frac{\sin(x)}{x}=1$.

(4) Après le changement de variable on trouve la même limite que au point précédent.

$$\lim_{x\to+\infty}x\sin(\frac{1}{x})=\lim_{y\to 0}\frac{\sin(y)}{y}=1. \tag{90.378}$$

(5) Ici l'astuce est de multiplier et diviser par le « binôme conjugué » du numérateur, à savoir

$$\sqrt{1-x^2}+\sqrt{1+x^2}. \tag{90.379}$$

par le produit remarquable $(\spadesuit-\clubsuit)(\spadesuit+\clubsuit)=\spadesuit^2-\clubsuit^2$, cela fait apparaitre au numérateur le produit

$$(\sqrt{1-x^2}-\sqrt{1+x^2})(\sqrt{1-x^2}+\sqrt{1+x^2})=1-x^2-1-x^2=-2x^2. \tag{90.380}$$

La limite à calculer est donc

$$\begin{aligned}\lim_{x\to 0}\frac{\sqrt{1-x^2}-\sqrt{1+x^2}}{5x}&=\lim_{x\to 0}\frac{-2x^2}{5x(\sqrt{1-x^2}+\sqrt{1+x^2})}\\&=\lim_{x\to 0}\frac{-2x}{5(\sqrt{1-x^2}+\sqrt{1+x^2})}\\&=\frac{0}{10}\\&=0.\end{aligned} \tag{90.381}$$

(6) La fonction dont on veut calculer la limite peut s'écrire comme

$$f(x)=x^{-3}\ln(x)=\frac{\ln(x)}{x^3}. \tag{90.382}$$

Nous utilisons le fait que « la croissance du logarithme est lente ». Donc le logarithme du numérateur tend moins vite vers l'infini que la puissance de x qui est au dénominateur. Nous avons donc

$$\lim_{x\to\infty}\frac{\ln(x)}{x^3}=0. \tag{90.383}$$

Exercise 177 [exoDS2010bis-0004]

Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par

$$\begin{cases}u_{n+1}=u_n+n+1,\\u_0=0.\end{cases} \tag{90.384}$$

(1) Démontrer que $u_n = \frac{n(n+1)}{2}$.

(2) Calculer, si elle existe, la limite de la suite $(u_n)_n$.

corrDS2010bis-0004

Correction of the exercise 177

(1) Il faut prouver par récurrence. D'abord la formule proposée est vraie pour $n = 0$ parce que $0(0+1)/2 = 0$. Ensuite, il faut vérifier que la formule est correcte pour $k+1$ en supposant qu'elle soit correcte pour tous indices n entre 0 et k. En posant $n = k+1$ dans la formule à démontrer, nous avons

$$u_{k+1} = u_k + k + 1 = \frac{k(k+1)}{2} + k + 1 = \frac{(k+1)(k+2)}{2}, \tag{90.385}$$

ce qui est bien la formule avec $n = k+1$.

(2) La limite est donc $\lim_{n\to\infty} u_n = \infty$ parce que u_n est la somme de tous les nombres entiers positifs entre 0 et n, donc c'est une suite monotone croissante qui n'est pas bornée.

Exercise 178 exoDS2010bis-0005

Considérons la suite définie par

$$u_n = \frac{(-1)^n n}{n+1}, \tag{90.386}$$

pour $n \geqslant 1$, $n \in \mathbb{N}$.

(1) Calculer la valeur des termes u_1, u_2, u_3 et u_4.

(2) Démontrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ est bornée.

(3) Est-ce que la suite $(u_n)_{n\in\mathbb{N}^*}$ est monotone ?

(4) On admet que la suite $(|u_n|)_{n\in\mathbb{N}^*}$ est convergente. Calculer sa limite.

(5) Calculer la valeur des premiers 3 termes des sous-suites $(u_{2n})_{n\in\mathbb{N}^*}$ (sous-suite des indices pairs) et $(u_{2n-1})_{n\in\mathbb{N}^*}$ (sous-suite des indices impairs), pour $n \geqslant 1$.

(6) Est-ce que la suite $(u_n)_{n\in\mathbb{N}^*}$ est convergente ? Et les suites $(u_{2n})_{n\in\mathbb{N}^*}$ et $(u_{2n-1})_{n\in\mathbb{N}^*}$?

corrDS2010bis-0005

Correction of the exercise 178

(1) Pour rappel,

$$(-1)^n = \begin{cases} 1 & \text{si } n \text{ est pair} \\ -1 & \text{si } n \text{ est impair.} \end{cases} \tag{90.387}$$

Les premiers termes sont donc

$$\begin{aligned} u_1 &= \frac{(-1)^1 1}{2} = -\frac{1}{2} \\ u_2 &= \frac{(-1)^2 2}{3} = \frac{2}{3} \\ u_3 &= \frac{(-1)^3 3}{4} = -\frac{3}{4} \\ u_4 &= \frac{(-1)^4 4}{5} = \frac{4}{5}. \end{aligned} \tag{90.388}$$

(2) Nous avons

$$\left|\frac{(-1)^n n}{n+1}\right| = \frac{n}{n+1}. \tag{90.389}$$

Or $\frac{n}{n+1}$ est inférieur à 1 pour tout n entier positif. Ce qui fait que $\frac{(-1)^n n}{n+1}$ est bornée entre -1 et 1.

(3) La suite est croissante entre u_1 et u_2, et décroissante entre u_2 et u_3. Plus généralement, les termes pairs seront positifs tandis que les termes impairs seront négatifs ; elle n'arrête donc pas de monter et descendre au dessus de zéro et en dessous de zéro.

(4) La suite $(|u_n|)$ converge vers 1. Pour le voir on peut calculer la limite de la fonction $f(x) = \frac{x}{x+1}$ lorsque x tend vers $+\infty$.

(5)
$$\begin{aligned} u_2 &= \frac{2}{3}, & u_1 &= -\frac{1}{2}, \\ u_4 &= \frac{4}{5}, & u_3 &= -\frac{3}{4}, \\ u_6 &= \frac{6}{7}, & u_5 &= -\frac{5}{6}. \end{aligned} \tag{90.390}$$

(6) La suite des termes pairs est (rappel : pour tout n, le nombre $2n$ est pair et donc $(-1)^{2n} = 1$)
$$u_{2n} = \frac{2n}{2n+1}. \tag{90.391}$$
Cette suite tend vers 1.
La suite des termes impairs par contre vaut
$$u_{2n+1} = -\frac{2n+1}{2n+2}, \tag{90.392}$$
et cette suite tend vers -1.
Nous avons donc trouvé deux sous-suites de (u_n) qui tendent vers des limites différentes. La suite (u_n) elle-même ne converge donc pas.

Exercise 179 exoDS_2010_1
Calculer :

(1) $\log_8 2$;

(2) $\log_6 1$;

(3) $\log_3 \frac{1}{9}$.

Exercise 180 exoDS_2010_2bis
Soient $f(x) = \dfrac{1}{1-x^2}$ et $g(x) = \ln(x)$.

(1) Trouvez les domaines de définition de f et de g.

(2) Écrivez les expressions explicites et trouvez les domaines de définition des fonctions composées $h_1(x) = f \circ g(x)$ et $h_2(x) = g \circ f(x)$.

Exercise 181 exoDS_2010_4bis
Déterminer, lorsqu'elles existent, les limites suivantes :

(1) $\lim\limits_{x\to 3} \dfrac{(x-3)(x+15)}{x^2-9}$;

(2) $\lim\limits_{x\to -\infty} \dfrac{x^{50} - 3200x^5 + 1}{5x^{17} - 3x + 4}$;

(3) $\lim\limits_{x\to 0} \dfrac{\sin(x)}{x}$;

(4) $\lim\limits_{x\to +\infty} x \sin(\dfrac{1}{x})$. Conseil : essayer la substitution $y = 1/x$;

(5) $\lim\limits_{x\to 0} \dfrac{\sqrt{1-x^2} - \sqrt{1+x^2}}{5x}$;

(6) $\lim\limits_{x\to +\infty} x^{-3} \ln(x)$.

Exercise 182 exoDS_2010_6bis
Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par
$$\begin{cases} u_{n+1} = u_n + n + 1, \\ u_0 = 0. \end{cases} \tag{90.393}$$

(1) Démontrer que $u_n = \dfrac{n(n+1)}{2}$.

(2) Calculer, si elle existe, la limite de la suite $(u_n)_n$.

Exercise 183 exoDS_2010_5bis

Considérons la suite définie par

$$u_n = \frac{(-1)^n n}{n+1}, \tag{90.394}$$

pour $n \geqslant 1$, $n \in \mathbb{N}$.

(1) Calculer la valeur des termes u_1, u_2, u_3 et u_4.

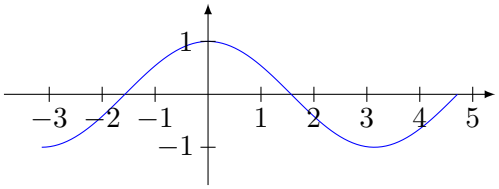
(2) Démontrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ est bornée.

(3) Est-ce que la suite $(u_n)_{n\in\mathbb{N}^*}$ est monotone ?

(4) On admet que la suite $(|u_n|)_{n\in\mathbb{N}^*}$ est convergente. Calculer sa limite.

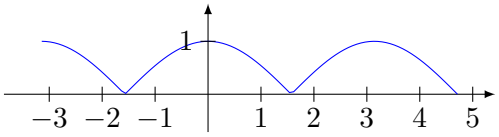
(5) Calculer la valeur des premiers 3 termes des sous-suites $(u_{2n})_{n\in\mathbb{N}^*}$ (sous-suite des indices pairs) et $(u_{2n-1})_{n\in\mathbb{N}^*}$ (sous-suite des indices impairs), pour $n \geqslant 1$.

(6) Est-ce que la suite $(u_n)_{n\in\mathbb{N}^*}$ est convergente ? Et les suites $(u_{2n})_{n\in\mathbb{N}^*}$ et $(u_{2n-1})_{n\in\mathbb{N}^*}$?



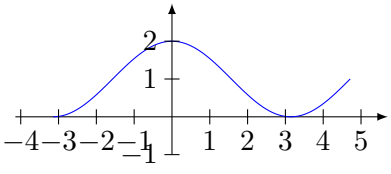
LabelFigExerciceGraphesbissscosinus

(a)



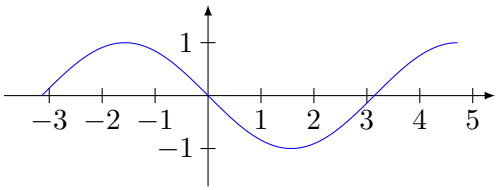
LabelFigExerciceGraphesbisssvalabsoluecosinus

(b)



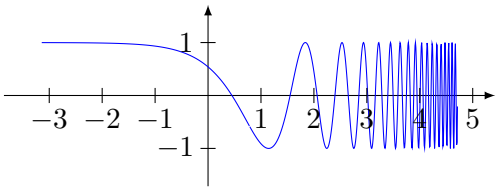
LabelFigExerciceGraphesbisscosxplusun

(c)



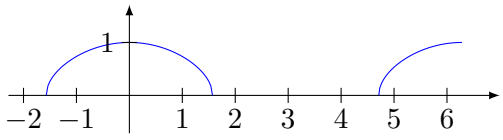
LabelFigExerciceGraphesbissssinus

(d)



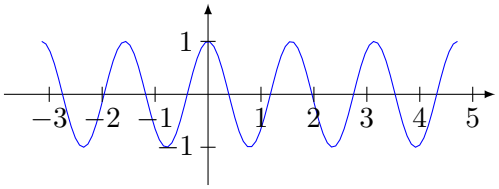
LabelFigExerciceGraphesbisscosex

(e)



LabelFigExerciceGraphesbissssqrtcos

(f)



LabelFigExerciceGraphesbisscosquattrox

(g)

FIGURE 90.10: Les graphes à considérer de la question 175.LabelFigExerciceGraphesbis

90.12.4 Épreuve complémentaire décembre 2010

90.12.4.1 A traiter

Exercise 184 exoExamenDecembre2010-0001

(1) Tracer le graphe de la fonction $f(x) = x$ pour x entre 0 et 5.

(2) Calculer l'intégrale $\int_0^5 x\,dx$, qui correspond à l'aire de la région entre l'axe des x et le graphe de f.

(3) Tracer le graphe de la fonction *partie entière* $x \mapsto [x]$, définie par

$$[x] = \text{le plus grand nombre entier qui est plus petit de } x.$$

Exemples : $[4.67] = 4$, $[2] = 2$, $[0.34] = 0$.

(4) Calculer l'intégrale $\int_0^5 [x]\,dx$. Conseil : écrire cette intégrale comme la somme de 5 intégrales $\int_0^1 \ldots + \cdots + \int_4^5 \ldots$.

(5) Tracer le graphe de la fonction *mantisse*, $m(x) = x - [x]$, pour x entre 0 et 5.

(6) Calculer l'intégrale $\int_0^5 x - [x]\,dx$.

corrExamenDecembre2010-0001

Correction of the exercise 184

(1) Le graphique est un segment de droite entre les points $(0, 0)$ et $(5, 5)$.

(2)

$$\int_0^5 x dx = \left[\frac{x^2}{2}\right]_0^5 = \frac{25}{2} - 0 = \frac{25}{2}. \tag{90.395}$$

(3) La fonction partie entière est tracée sur la figure 90.11.

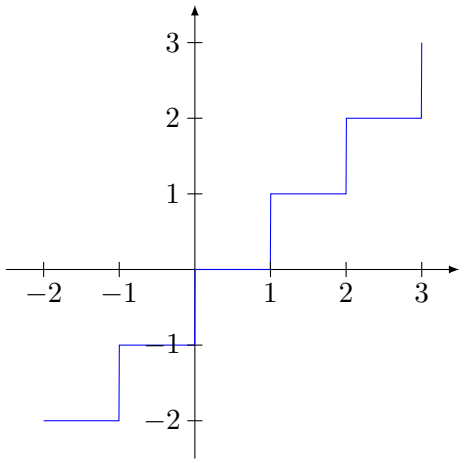


FIGURE 90.11: La fonction partie entière. LabelFigPartieEntiere

(4) En découpant l'intervalle d'intégration, nous avons le calcul suivant :

$$\begin{aligned}\int_0^5 [x]dx &= \int_1^1 0dx + \int_1^2 1dx + \int_2^3 2dx + \int_3^4 3dx + \int_4^5 4dx \\ &= 1 + 2 + 3 + 4 = 10.\end{aligned} \tag{90.396}$$

Notez que $[x] = 2$ pour x entre 2 et 3. C'est pour cela que nous avons

$$\int_2^3 [x]dx = \int_2^3 2dx. \tag{90.397}$$

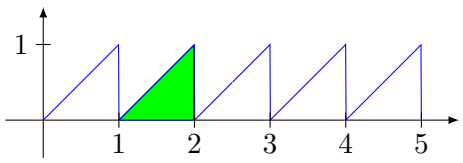


FIGURE 90.12: La fonction mantisse. Remarquez que du point de vue de la surface, ce sont des petits triangles. LabelFigMantisse

(5) La fonction mantisse est tracée sur la figure 90.12.

(6) Nous pouvons utiliser les résultats précédents :

$$\int_0^5 x-[x]dx=\int_0^5 xdx-\int_0^5[x]dx=\frac{25}{2}-10=\frac{5}{2}. \tag{90.398}$$

Une autre façon est de se souvenir que cette intégrale représente la surface sous la courbe (voir figure 90.12). Cette surface est formée de cinq triangles dont les surfaces valent $\frac{1}{2}$; pour rappel, la surface d'un triangle est donnée par $\frac{\text{base}\times\text{hauteur}}{2}$. Donc la surface totale est de $\frac{5}{2}$.

Exercise 185 exoECdecembre2010-0001

Modèle malthusien contrôlé. Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_{n+1}=au_n+b, & \text{pour tous } n\in\mathbb{N},\ n>0\\ u_0=x \end{cases}$$

où $a\neq 1$, b et $x\geqslant 0$ sont des nombres réels.

(1) Calculer les premiers 4 termes de la suite si $a=2$ et $x=10$.

(2) Montrer que $u_n=a^n(x+\frac{b}{a-1})-\frac{b}{a-1}$ pour tout $n\in\mathbb{N}$.

corrECdecembre2010-0001

Correction of the exercise 185

Voir la correction de l'exercice 115.

Exercise 186 exoECdecembre2010-0002

Calculer les dérivées des fonctions suivantes

(1) $f(x)=\tan(x)$. Rappel : $\tan(x)=\sin(x)/\cos(x)$;

(2) $g(x)=5\sin(e^x)$;

(3) $h(x)=\ln(x)\cos(x^2)$.

corrECdecembre2010-0002

Correction of the exercise 186

<+CorrECdecembre2010-0002+>

Exercise 187 exoECdecembre2010-0003

Trouver la solution générale de l'équation

$$y'=yx^{1/2}. \tag{90.399}$$

corrECdecembre2010-0003

Correction of the exercise 187

<+CorrECdecembre2010-0003+>

Exercise 188 exoECdecembre2010-0004

<+ExoECdecembre2010-0004+>

corrECdecembre2010-0004

Correction of the exercise 188

<+CorrECdecembre2010-0004+>

90.12.5 Examen décembre 2010

Exercise 189 exoExamen-0001

(1) Tracer le graphe de la fonction $f(x) = x$ pour x entre 0 et 5.

(2) Calculer l'intégrale $\int_0^5 x\,dx$, qui correspond à l'aire de la région entre l'axe des x et le graphe de f.

(3) Tracer le graphe de la fonction *partie entière* $x \mapsto [x]$, définie par

$$[x] = \text{le plus grand nombre entier qui est plus petit de } x.$$

Exemples : $[4.67] = 4$, $[2] = 2$, $[0.34] = 0$.

(4) Calculer l'intégrale $\int_0^5 [x]\,dx$. Conseil : écrire cette intégrale comme la somme de 5 intégrales $\int_0^1 \ldots + \cdots + \int_4^5 \ldots$.

(5) Tracer le graphe de la fonction *mantisse*, $m(x) = x - [x]$, pour x entre 0 et 5.

(6) Calculer l'intégrale $\int_0^5 x - [x]\,dx$.

Exercise 190 exoExamenDecembre2010-0002

Calculer les limites suivantes :

(1) $\lim_{x\to+\infty} \frac{x^3+4x+13}{x+x^5+2}$;

(2) $\lim_{x\to0} \frac{x^3+4x+13}{\sqrt{x}+x^5+2}$;

(3) $\lim_{x\to0} \frac{\sin^2(x)}{x}$;

(4) $\lim_{x\to5} \frac{x^2-25}{x-5}$;

(5) $\lim_{x\to+\infty} \frac{e^x}{x^5+1}$;

(6) $\lim_{x\to-\infty} \frac{e^x}{x^5+1}$.

corrExamenDecembre2010-0002

Correction of the exercise 190

(1) Nous mettons en évidence la plus grande puissance de x au numérateur et au dénominateur. Attention : celle du dénominateur est 5 :

$$\frac{x^3+4x+13}{x+x^5+2} = \frac{x^3\left(1+\frac{4}{x^2}+\frac{12}{x^3}\right)}{x^5\left(\frac{1}{x^4}+1+\frac{2}{x^5}\right)}. \tag{90.400}$$

À ce moment, nous simplifions par x^3, et nous nous souvenons qu'à la limite $x \to \infty$, les termes du type $\frac{4}{x^2}$ tendent vers 0. Il reste donc $\lim_{x\to\infty} \frac{1}{x^2} = 0$.
La réponse est donc zéro.

(2) Ici, en remplaçant simplement x par zéro dans l'expression, nous ne tombons sur aucune indétermination : $\frac{13}{2}$.

(3) La limite très connue est $\lim_{x\to0} \frac{\sin(x)}{x} = 1$. Ici nous écrivons

$$\frac{\sin^2(x)}{x} = \sin(x)\frac{\sin(x)}{x}. \tag{90.401}$$

Nous avons alors

$$\lim_{x\to0} \sin(x)\frac{\sin(x)}{x} = \lim_{x\to0}\sin(x)\lim_{x\to0}\frac{\sin(x)}{x} \tag{90.402}$$

(4) En remplaçant x par 5, nous tombons sur l'indétermination $\frac{0}{0}$. Nous parvenons à la lever en factorisant le numérateur et en simplifiant :

$$\frac{x^2-25}{x-5}=\frac{(x-5)(x+5)}{x-5}=x+5. \tag{90.403}$$

Maintenant, la limite vaut $\lim_{x\to 5} x+5=10$.

(5) Nous avons vu que, pour les limites en l'infini, l'exponentielle « avance plus vite que tout polynôme ». La limite est donc ∞.

(6) En remplaçant nous trouvons $\frac{0}{-\infty}$. Cela n'est pas une indétermination : ça vaut 0.

Exercise 191 exoExamenDecembre2010-0003

On considère les trois suites suivantes, définies pour $n\geqslant 1$,

$$u_n=\frac{1}{n},\quad v_n=\frac{1}{n+1},\quad w_n=u_nv_n.$$

(1) Écrire les premiers 4 termes de chaque suite.

(2) Montrer que la suite $(u_n)_n$ est bornée et décroissante.

(3) Les trois suites convergent. Trouver la limite de chaque suite.

(4) Monter que $w_n=\int_n^{n+1}\frac{1}{x^2}\,dx$, pour tout $n\geqslant 1$.

(5) Esquisser le graphe de la fonction $1/x^2$ pour x dans $[1,4]$. Donner une intérpretation graphique des intégrales w_1, w_2, w_3.

(6) On définit la suite $(s_n)_n$ par $s_n=w_1+w_2+\cdots+w_n$. Démontrer, par récurrence que $s_n=\int_1^{n+1}\frac{1}{x^2}\,dx$.

(7) Calculer la limite de $(s_n)_n$. Conseil : calculer l'intégrale du point précédent.

corrExamenDecembre2010-0003

Correction of the exercise 191

<+CorrExamenDecembre2010-0003+>

Exercise 192 exoExamenDecembre2010-0004

Calculer les intégrales suivantes

(1) $\int_0^1 x^7+x^{1/3}\,dx$;

(2) $\int_{\pi/2}^{2\pi}\cos(x)\,dx$;

(3) $\int_0^1\frac{4}{4x+1}\,dx$;

(4) $\int_0^{\pi/2}x^2\sin(x)\,dx$. Conseil : intégrer par parties ;

(5) $\int_0^1 e^x\sqrt{1+e^x}\,dx$. Conseil : essayer le changement de variable $y=1+e^x$.

corrExamenDecembre2010-0004

Correction of the exercise 192

<+CorrExamenDecembre2010-0004+>

Exercise 193 exoExamenDecembre2010-0005

On considère l'équation différentielle linéaire

$$y'=-xy+x. \tag{90.404}$$ nonhomog

(1) On commence par considérer l'équation homogène associée

$$y'=-xy. \tag{90.405}$$ homog

(1a) Dire si l'équation (90.405) a une solution constante. Si oui, la trouver.

(1b) Trouver la solution générale de l'équation (90.405).

(2) Trouver une solution particulière $P(x)$ de (90.404) de la forme $P(x) = ax + b$, avec a et b dans $\mathbb{R}$.

(3) Vérifiez que la fonction $1 - 3e^{-x^2/2}$ est une solution du problème suivant

$$\begin{cases} y' = -xy + x, \\ y\big(\sqrt{\ln(3)}\big) = 1 - \sqrt{3}. \end{cases} \tag{90.406}$$

corrExamenDecembre2010-0005

Correction of the exercise 193

(1) (1a) Dire que $y(x)$ est une constante, c'est dire que $y(x) = C$ pour un certain $C \in \mathbb{R}$. Dans ce cas $y'(x) = 0$ et l'équation devient $0 = -xC$ pour tout x. Le seul choix de constante C qui convient est $C = 0$.

Par conséquent, la fonction $y(x) = 0$ est une solution constante de l'équation $y' = xy$.

(1b) Pour trouver la solution générale, nous écrivons $y' = dy/dx$ et nous mettons tous les y à gauche et tous les x à droite :

$$\begin{aligned} \frac{dy}{dx} &= -xy \\ \frac{dy}{y} &= -xdx \\ \ln(y) &= -\frac{x^2}{2} + C. \end{aligned} \tag{90.407}$$

Pour obtenir la dernière ligne, nous avons intégré des deux côtés de l'équation. Nous prenons maintenant l'exponentielle des deux membres afin d'isoler le y :

$$y(x) = e^{-\frac{x^2}{2}+C} = e^{-x^2/2}e^C = Ke^{-x^2/2} \tag{90.408}$$

où nous avons posé $K = e^C$.

(2) Si $y(x) = ax + b$, alors $y'(x) = a$. En mettant cela dans l'équation (90.404), nous trouvons

$$a = -x(ax + b) + x, \tag{90.409}$$

ou encore

$$-ax^2 + (1 - b)x - a = 0. \tag{90.410}$$ EqPouraetbÃlzuzc

Attention : cette équation n'est pas une équation pour trouver x, mais une équation pour a et b qui doit être vérifiée pour tout x. Lorsque $x = 0$, cette équation devient $a = 0$. Nous devons donc avoir $a = 0$. L'équation (90.410) devient alors

$$(1 - b)x = 0 \tag{90.411}$$

qui doit encore être vraie pour tout x. Avec $x = 1$, nous avons la condition $b - 1 = 0$ et donc $b = 1$. Les valeurs de a et b telles que $P(x) = ax + b$ soit solution sont donc $a = 0$ et $b = 1$. C'est donc $P(x) = 1$.

(3) Nous avons

$$y(x) = 1 - 3e^{-x^2/2}y'(x) = -3(-x)e^{-x^2/2} = 3xe^{-x^2/2}. \tag{90.412a}$$

En remplaçant ces valeurs de y et y' cela dans l'équation $y' = -xy + x$, nous avons

$$3xe^{-x^2/2} = -x\big(1 - 3e^{-x^2/2}\big) + x. \tag{90.413}$$

En distribuant le x dans le membre de droite, nous trouvons que cette équation est bien vérifiée pour tout x.
En remplaçant x par $\sqrt{\ln(3)}$ dans la formule de $y(x)$, nous avons

$$\begin{aligned} y\big(\sqrt{\ln(3)}\big) &= 1 - 3e^{-\left(\sqrt{\ln(2)}\right)^2/2} \\ &= 1 - 3e^{-\ln(3)/2} \\ &= 1 - 3\sqrt{e^{-\ln(3)}} \end{aligned} \tag{90.414}$$

où nous avons utilisé le fait que $e^{a/2} = \sqrt{e^a}$. Maintenant en appliquant la formule $e^{-a} = \frac{1}{e^a}$ avec $a = \ln(3)$, nous trouvons

$$y\big(\sqrt{\ln(3)}\big) = 1 - 3\sqrt{\frac{1}{3}} = 1 - \sqrt{\frac{9}{3}} = 1 - \sqrt{3}. \tag{90.415}$$

90.12.6 DS décembre 2011

Exercise 194 exoExosenvrac-0001 **(5 points)**

Compléter le tableau suivant en considérant $\alpha \neq -1$, $\alpha \neq -2$.

Primitive $\int f(x)\,dx$	Function $f(x)$	Dérivée $f'(x)$
$\sin(x) + C$	...	...
...	$\sin(3x)$	...
...	...	$-\frac{1}{x^2}$
...	$\ln(x)$	...
...	...	$(\alpha+1)x^\alpha$
$\frac{e^{5x}}{5} + C$	...	...
...	$\tan(x) = \frac{\sin(x)}{\cos(x)}$	...
...	$\arcsin(x)$	...

(90.416)

corrExosenvrac-0001

Correction of the exercise 194

Primitive $\int f(x)\,dx$	Fonction $f(x)$	Dérivée $f'(x)$
$\sin(x)+C$	$\cos(x)$	$-\sin(x)$
$-\frac{1}{3}\cos(3x)+C$	$\sin(3x)$	$3\cos(3x)$
$\ln(x)+Cx+D$	$\frac{1}{x}+C$	$-\frac{1}{x^2}$
$x\ln(x)-x+C$	$\ln(x)$	$\frac{1}{x}$
$\frac{x^{\alpha+2}}{\alpha+2}+Cx+D$	$x^{\alpha+1}+C$	$(\alpha+1)x^{\alpha}$
$\frac{e^{5x}}{5}+C$	e^{5x}	$5e^{5x}$
$-\ln\left(\lvert\cos(x)\rvert\right)+C$	$\tan(x)=\frac{\sin(x)}{\cos(x)}$	$\tan^2(x)+1=\frac{1}{\cos^2(x)}$
$x\arcsin(x)+\sqrt{1-x^2}+C$	$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$

(90.417)

Exercise 195 expExosenyrac-0015 **(6 points)**

Déterminer dans chaque cas le domaine de définition, la périodicité et/ou les symétries éventuelles et les limites aux extrêmes du domaine. Calculer ensuite la dérivée de la fonction, la où elle est définie.

(1) $f_1(x)=x^3-1$;

(1a) Domaine de définition : ...

(1b) La fonction est périodique ? Si oui, trouver sa période.

(1c) La fonction est paire ? Impaire ? Expliquer.

(1d) Limites aux extrêmes du domaine : ...

(1e) Dérivée : ...

(2) $f_2(x)=e^{\cos(x)}$;

(2a) Domaine de définition : ...

(2b) La fonction est périodique ? Si oui, trouver sa période.

(2c) La fonction est paire ? Impaire ? Expliquer.

(2d) Limites aux extrêmes du domaine : ...

(2e) Dérivée : ...

(3) $f_3(x) = \frac{x}{x-2}$;

(3a) Domaine de définition : ...

(3b) La fonction est périodique ? Si oui, trouver sa période.

(3c) La fonction est paire ? Impaire ? Expliquer.

(3d) Limites aux extrêmes du domaine : ...

(3e) Dérivée : ...

corrExosenvrac-0015

Correction of the exercise 195

(1) $f(x) = x^3 - 1$.

(1a) Le domaine est $\mathbb{R}$.

(1b) La fonction n'est pas périodique.

(1c) La fonction n'est ni paire ni impaire :

$$f(-x) = (-x)^3 - 1 = -x^3 - 1, \tag{90.418}$$

qui n'est ni $f(x)$ ni $-f(x)$.

(1d) Les bords du domaine sont $\pm\infty$. Les limites sont

$$\lim_{x\to\infty} x^3 - 1 = \infty \tag{90.419a}$$

$$\lim_{x\to-\infty} x^3 - 1 = -\infty. \tag{90.419b}$$

(1e) La dérivée est $f'(x) = 3x^2$.

(2) $f(x) = e^{\cos(x)}$

(2a) Le domaine est $\mathbb{R}$ entier.

(2b) La fonction est périodique de période 2π.

(2c) Elle est paire parce que

$$f(-x) = e^{\cos(-x)} = e^{\cos(x)} = f(x). \tag{90.420}$$

Ici ce qui joue est $\cos(-x) = \cos(x)$.

(2d) Les bords du domaine sont $x = \pm\infty$. La fonction cosinus oscillant sans fin entre 1 et -1, la fonction oscille entre e^1 et e^{-1}. Les limites n'existent donc pas.

(2e) La dérivée :

$$f'(x) = -\sin(x)e^{\cos(x)}. \tag{90.421}$$

(3) $f_3(x) = \frac{x}{x-2}$.
   (3a) Le domaine est $\mathbb{R}\backslash\{2\}$.
   (3b) La fonction n'est pas périodique.
   (3c) Ni paire ni impaire :
   $$f(-x) = \frac{-x}{-x-2} = \frac{x}{x+2}, \tag{90.422}$$
   qui est ni $f(x)$ ni $f(-x)$.
   (3d) Les bords du domaine sont $x = \pm\infty$ et $x = 2$. Nous avons les limites
   $$\lim_{x\to-\infty} f_3(x) = 1 \tag{90.423a}$$
   $$\lim_{x\to\infty} f_3(x) = 1 \tag{90.423b}$$
   $$\lim_{x\to 2} f_3(x) = \text{n'existe pas.} \tag{90.423c}$$
   La troisième limite n'existe pas parce que la limite à gauche et à droite ne coïncident pas.
   (3e) La dérivée est
   $$f_3'(x) = \frac{1}{x-2} - \frac{x}{(x-2)^2} = \frac{-2}{(x-2)^2}. \tag{90.424}$$

Exercise 196 expExosenvrac-0015A **(2 points)**

Trouver les intervalles sur lesquels les fonctions $f_1(x) = x^3 + 1$ et $f_2 = \frac{x}{x-2}$ sont strictement monotones et déterminer leurs fonctions réciproques. corrExosenvrac-0015A

Correction of the exercise 196

(1) La fonction f_1 est partout strictement croissante parce que sa dérivée, $f_1'(x) = 3x^2$, est partout positive. Notez que $f'(0) = 0$, mais comme cela n'arrive qu'à un seul point et la dérivée ne change pas de signe à droite et à gauche de $x = 0$, la fonction reste strictement croissante.
   La fonction réciproque se trouve en résolvant par rapport à x l'équation
   $$y = x^3 + 1. \tag{90.425}$$
   La solution est
   $$x = \sqrt[3]{y-1}. \tag{90.426}$$

(2) La fonction f_2 demande immédiatement $x \neq 2$. En ce qui concerne la dérivée,
   $$f_2'(x) = \frac{1}{x-1} - \frac{x}{(x-2)^2} = \frac{-2}{(x-2)^2}. \tag{90.427}$$
   La fonction est toujours décroissante (sauf évidemment en $x = 2$ où la fonction n'existe pas). En ce qui concerne la fonction réciproque, il faut résoudre
   $$y = \frac{x}{x-2} \tag{90.428}$$
   pour obtenir y en fonction de x. Nous avons successivement
   $$y = \frac{x}{x-2} \tag{90.429a}$$
   $$y(x-2) = x \tag{90.429b}$$
   $$yx - 2y - x = 0 \tag{90.429c}$$
   $$x(y-1) = 2y \tag{90.429d}$$
   $$x = \frac{2y}{y-1} \tag{90.429e}$$
   La fonction réciproque est donc
   $$f_2^{-1}(y) = \frac{2y}{y-1}. \tag{90.430}$$

Exercise 197 exoExosenvrac-0006 **(3 points)**

Résoudre par séparation des variables l'équation différentielle

$$x' = tx^2. \tag{90.431}$$ eqexo6

Une fois obtenue la solution générale de l'équation (90.431) trouver son unique solution qui satisfait la condition $x(2) = 1$. corrExosenvrac-0006

Correction of the exercise 197

L'équation $x' = tx^2$ peut se résoudre par séparation de variables parce que le membre de droite de l'équation est le produit d'une fonction de t, à l'occurrence l'identité, et une fonction qui ne dépend explicitement que de x : x^2.

La seule solution constante de l'équation est $x(t) = 0$.

Si x n'est pas nulle on peut écrire l'équation sous la forme

$$\frac{1}{x^2}x' = t, \tag{90.432}$$

et intégrer par rapport à t les deux côtés. Suite au changement de variable classique $x = x(t)$ dans l'intégrale à gauche, on obtient

$$\int \frac{1}{x^2}\,dx = \int t\,dt, \tag{90.433}$$

et donc

$$-\frac{1}{x} = \frac{t^2}{2} + C, \qquad C \in \mathbb{R}. \tag{90.434}$$

La solution générale de l'équation est alors $x(t) = -\dfrac{1}{\frac{t^2}{2} + C}$.

La solution du problème de Cauchy

$$\begin{cases} x' = tx^2, \\ x(2) = 1, \end{cases} \tag{90.435}$$

n'est pas la solution constante. Il faut alors supposer qu'elle soit de la forme $x(t) = -\dfrac{1}{\frac{t^2}{2} + C}$ et calculer $x(2) = -\dfrac{1}{2 + C}$. La condition $-\dfrac{1}{2 + C} = 1$ nous permet de trouver la valeur de la constante C qui détermine l'unique solution du problème de Cauchy. Ici : $C = -3$. **Exercise 198** exoExosenvrac-0009 **(4 points)**

Calculer les intégrales suivantes.

(1) $\displaystyle\int_0^1 \cos\left(x\frac{\pi}{2}\right) + x^4\,dx$;

(2) $\displaystyle\int_0^1 xe^{x^2}\,dx$, par changement de variable ;

(3) $\displaystyle\int_0^1 x^2e^x\,dx$, par parties.

(4) Utiliser le fait que $\displaystyle\int_{\pi/2}^{\pi} \sin^2(x)\,dx = \frac{\pi}{4}$ pour calculer la valeur de $\displaystyle\int_{\pi/2}^{\pi} \cos^2(x)\,dx$. corrExosenvrac-0009

Correction of the exercise 198

(1) $\displaystyle\int_0^1 \cos\left(x\frac{\pi}{2}\right) + x^4\,dx = \left[\frac{2}{\pi}\sin\left(x\frac{\pi}{2}\right) + \frac{x^5}{5}\right]_0^1 = \frac{2}{\pi} + \frac{1}{5}$;

(2) Le changement de variable à utiliser est $y = x^2$. On obtient $\displaystyle\int_0^1 xe^{x^2}\,dx = \int_0^1 \frac{e^y}{2}\,dy = \frac{e-1}{2}$;

(3) $\int_0^1 x^2e^x\,dx = \left[x^2e^x\right]_0^1 - \int_0^1 2xe^x\,dx = \left[x^2e^x\right]_0^1 - \left[2xe^x\right]_0^1 + 2\int_0^1 e^x\,dx = e - 2e + 2e - 2 = e - 2.$

(4) La valeur de $\int_{\pi/2}^{\pi} \cos^2(x)\,dx$ est $\frac{\pi}{4}$. Il y a plusieurs méthodes pour obtenir ce résultat mais celle qui nous parait la plus rapide est de se souvenir du fait que $\cos^2(x) + \sin^2(x) = 1$ et intégrer les deux côtés de cette égalité par rapport à x. On obtient alors $\int_{\pi/2}^{\pi} \cos^2(x)\,dx + \frac{\pi}{4} = \frac{\pi}{2}$.

90.12.7 Épreuve complémentaire, 15 décembre 2010

Exercise 199 exoEC-0003

Calculer les dérivées des fonctions suivantes

(1) $f(x) = \tan(x)$. Rappel : $\tan(x) = \sin(x)/\cos(x)$;

(2) $g(x) = 5\sin(e^x)$;

(3) $h(x) = \ln(x)\cos(x^2)$.

Exercise 200 EC-0004

Trouver la solution générale de l'équation

$$y' = yx^{1/2}.$$

Chapitre 91

Géométrie analytique (Besançon)

91.1 Espaces vectoriels normés

91.1.1 Normes

Exercise 201 ⊙⊙ | exoEspVectoNorme0001

On considère l'espace vectoriel $\mathbb{R}^n$ muni de la norme euclidienne.

(1) Montrer que la norme euclidienne est **convexe** c'est-à-dire que quels que soient $x, y \in \mathbb{R}^n$ et $t \in [0,1]$ on a
$$\|tx + (1-t)y\| \leqslant t\|x\| + (1-t)\|y\|. \tag{91.1}$$

(2) Montrer que la norme est lipschitzienne de constante 1, c'est-à-dire quels que soient $x, y \in \mathbb{R}^n$ on a
$$| \|x\| - \|y\| | \leqslant \|x - y\|. \tag{91.2}$$

(3) Montrer l'**identité du parallélogramme** : quels que soient $x, y \in E$ on a
$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{91.3}$$

corrEspVectoNorme0001

Correction of the exercise 201

(1) D'abord, en utilisant l'inégalité triangulaire,
$$\|tx + (1-t)y\| \leqslant \|tx\| + \big\|(1-t)y\big\|. \tag{91.4}$$
Ensuite, étant donné que $t \in [0,1]$, nous avons $t > 0$ et $1 - t > 0$, donc nous pouvons les sortir de la norme sans mettre de valeur absolue (voir le point (3) de la définition 7.158 et le fait que nous sachions, par ailleurs, que la norme euclidienne est une norme).
Notez que ce point est valable pour toute norme : nous n'avons utilisé que des propriétés de définition des normes.

(2) Ce point est valable pour toute les normes en vertu de la proposition 7.159.

(3) Utilisons le fait que la norme euclidienne découle du produit scalaire [1] : $\|x\|^2 = x \cdot x$. Nous avons donc
$$\begin{aligned}\|x-y\|^2 + \|x+y\|^2 &= (x-y)\cdot(x-y) + (x+y)\cdot(x+y)\\ &= x\cdot x - x\cdot y - y\cdot x + y\cdot y\\ &\quad + x\cdot x + x\cdot y + y\cdot x + y\cdot y\\ &= 2x\cdot x + 2y\cdot y\\ &= 2\|x\|^2 + 2\|y\|^2.\end{aligned} \tag{91.5}$$

1. Nous mentionnons ce fait autour de la définition 10.10.

Exercise 202 exoGeomAnal-0040
Soit $\|.\|_p$ la norme p sur $\mathbb{R}^n$. Si A est une matrice, montrer que

$$\|A\|_p = \sup_{\|x\|_p=1} \|Ax\|. \tag{91.6}$$

Voir la définition 10.55. corrGeomAnal-0040

Correction of the exercise 202
Nous allons montrer que les ensembles sur lesquels ont prend le supremum sont en réalité les mêmes :

$$\underbrace{\left\{ \frac{\|Ax\|_p}{\|x\|_p \text{ tel que } x \neq 0} \right\}}_{A} = \underbrace{\{\|Ax\|_p \text{ tel que } \|x\|_p = 1\}}_{B}. \tag{91.7}$$

Attention : ce sont des sous-ensembles de réels ; pas de sous-ensembles de $\mathbb{M}(\mathbb{R})$ ou des sous-ensembles de $\mathbb{R}^n$.

Pour la première inclusion, prenons un élément de A, et prouvons qu'il est dans B. C'est-à-dire que nous prenons $x \in \mathbb{R}^n$ et nous considérons le nombre $\|Ax\|_p/\|x\|_p$. Le vecteur $y = x/\|x\|$ est un vecteur de norme 1, donc la norme de Ay est un élément de B, mais

$$\|Ay\|_p = \frac{\|Ax\|_p}{\|x\|_p}. \tag{91.8}$$

Nous avons donc $A \subset B$.

L'inclusion $B \subset A$ est immédiate.

Exercise 203 ⊛⊛ exoGeomAnal-0041
On va montrer que pour $p = \infty$, la norme matricielle p vérifie pour tout $A = (a_{i,j}) \in \mathbb{M}_n(\mathbb{R})$:

$$\|A\|_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}| \tag{91.9}$$ zzqnGAAP

(1) Soit $x \in \mathbb{R}^n$ tel que $||x||_\infty = 1$. On pose $b = Ax$ (on a donc $||Ax||_\infty = \max_{1 \leqslant i \leqslant n} |b_i|$).
   (1a) Montrer que $|b_i| \leqslant \sum_{j=1}^{n} |a_{ij}|$ pour tout $i \in \{1, .., n\}$.
   (1b) En déduire que

$$||Ax||_\infty \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

   puis que

$$||A||_\infty \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

(2) On va montrer l'inégalité inverse. Pour cela soit i^*, l'indice correspondant à l'égalité suivante :

$$\sum_{j=1}^{n} |a_{i^*j}| = \max_{i \in \{1,\dots,n\}} \sum_{j=1}^{n} |a_{ij}|. \tag{91.10}$$

   L'indice i^* est donc le numéro de la ligne qui réalise le maximum dont on parle dans le membre de droite. Soit par ailleurs x^* le vecteur de norme $||x||_\infty = 1$ défini par :

$$x_j^* = \begin{cases} 1 & \text{si } a_{i^*j} \geqslant 0 \\ -1 & \text{si } a_{i^*j} < 0 \end{cases}$$

   (2a) Monter que $a_{i^*j} x_j^* = |a_{i^*j}|$, sans sommation sur j (ni sur i ni sur i^*).
   (2b) Montrer que

$$||Ax^*||_\infty \geqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

(2c) En déduire le résultat (91.9).

corrGeomAnal-0041

Correction of the exercise 203

(1) Par définition de l'action d'une matrice sur un vecteur, nous avons

$$b_i = \sum_k a_{ik}x_k. \tag{91.11}$$

En prenant la valeur absolue, nous trouvons

$$|b_i| \leqslant \sum_k |a_{ik}||x_k| \leqslant \sum_k |a_{ik}| \tag{91.12}$$

parce que $|x_k| \leqslant 1$.
Pour chaque $x \in \mathbb{R}^n$ tel que $\|x\|_\infty = 1$ nous avons

$$\|Ax\|_\infty = \max_{1\leqslant i\leqslant n} |b_i| \leqslant \max_{1\leqslant i\leqslant n} \sum_k |a_{ik}|. \tag{91.13}$$

Donc en ce qui concerne la norme de A, nous avons

$$\|A\|_\infty = \sup_{\|x\|_\infty=1} \|Ax\|_\infty, \tag{91.14}$$

mais tous les éléments de l'ensemble sur lequel le supremum est pris sont plus petit que $\max \sum_k |a_{ik}|$.

(2) Nous passons à présent à l'inégalité inverse.

(2a) Le vecteur x^* est fait exprès pour mettre un signe moins lorsque a_{i^*j} est négatif et un signe plus là où a_{i^*j} est positif. L'effet est donc de rendre $a_{i^*j}x_j^*$ toujours positif.

(2b) Par définitions,

$$\|Ax^*\|_\infty = \max_{i\in\{1,\dots,n\}} |(Ax^*)_i| = \max_i \left|\sum_k a_{ik}x_k^*\right|. \tag{91.15}$$

Nous pouvons majorer en remplaçant i par i^* en effet, le membre de droite prend un maximum sur tous les i possibles alors que i^* est un i possible. Nous avons alors

$$\|Ax^*\|_\infty \geqslant \left|\sum_k a_{i^*k}x_k^*\right| = \left|\sum_k |a_{i^*k}|\right| = \sum_k |a_{i^*k}|. \tag{91.16}$$

Maintenant la définition de i^* nous permet de dire

$$\|Ax^*\|_\infty \geqslant \max_{i\in\{1,\dots,n\}} \sum_k |a_{ik}|. \tag{91.17}$$

Par conséquent

$$\|A\|_\infty \geqslant \max_{i\in\{1,\dots,n\}} \sum_k |a_{ik}|. \tag{91.18}$$

Ayant obtenu les inégalités dans les deux sens, les quantités sont égales.

Exercise 204 ⊛⊛ exoGeomAnal-0042

Montrer que la norme matricielle $\|.\|_1$ satisfait la formule

$$||A||_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}| \tag{91.19}$$

pour tout $A \in \mathbb{M}_n(\mathbb{R})$.

corrGeomAnal-0042

Correction of the exercise 204

Nous suivons essentiellement les mêmes étapes et astuces que pour l'exercice 203. Soit $x \in \mathbb{R}^n$ avec $\|x\|_1 = 1$ et $b = Ax$. Alors nous avons

$$\|Ax\|_1 = \|b\|_1 = \sum_{i=1}^{n} |b_i| \tag{91.20a}$$
$$= \sum_i \left| \sum_{j=1}^{n} a_{ij}x_j \right| \tag{91.20b}$$
$$\leqslant \sum_i \sum_j |a_{ij}x_j| \tag{91.20c}$$
$$= \sum_j |x_j| \sum_i |a_{ij}| \tag{91.20d}$$

À ce moment nous ne pouvons pas remplacer $\sum_j |x_j|$ par $\|x\|_1 = 1$ parce que la somme sur j se poursuit dans le reste de la formule ! Nous devons donc faire une étape supplémentaire :

$$\|Ax\|_1 \leqslant \sum_j |x_j| \sum_i |a_{ij}| \tag{91.21a}$$
$$\leqslant \sum_j |x_j| \left(\max_{j\in\{1,\dots,n\}} \sum_i |a_{ij}| \right) \tag{91.21b}$$
$$= \max_j \sum_i |a_{ij}|. \tag{91.21c}$$

Pour l'inégalité inverse, nous considérons j^*, le j qui réalise le maximum $\max_j \sum_i |a_{ij}|$, c'est-à-dire

$$\max_j \sum_i |a_{ij}| = \sum_i |a_{ij*}|, \tag{91.22}$$

et nous considérons le vecteur x^* donné par

$$x_j^* = \begin{cases} 1 & \text{si } a_{ij*} \geqslant 0 \\ -1 & \text{si } a_{ij*} < 0. \end{cases} \tag{91.23}$$

Dans ce cas nous avons

$$\|Ax^*\|_1 = \sum_i \left| \sum_j a_{ij}x_j^* \right| \tag{91.24a}$$
$$= \sum_i \sum_j |a_{ij}| \tag{91.24b}$$
$$= \sum_j \sum_i |a_{ij}| \tag{91.24c}$$
$$\geqslant \max_j \sum_i |a_{ij}| \tag{91.24d}$$

La dernière majoration est le fait qu'une somme d'éléments positifs est plus grande que le plus grands de ses éléments. Nous avons donc prouvé que

$$\|A\|_1 = \sup_{\|x\|_1=1} \|Ax\|_1 \geqslant \max_j \sum_i |a_{ij}|. \tag{91.25}$$

Exercise 205 exoGeomAnal-0043

Trouver les normes matricielles 1,2 et ∞ de la matrice :

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix}$$

corrGeomAnal-0043

Correction of the exercise 205

Les normes 1 et ∞ sont relativement simples :

$$\|A\|_1 = \max\{5,5,5\} = 5 \tag{91.26}$$

et

$$\|A\|_\infty = \max\{4,7,4\}. \tag{91.27}$$

Pour la norme $\|A\|_2$, nous devons calculer le rayon spectral de A^tA. La première chose à faire est de calculer les valeurs propres. Ce sont des calculs à faire à l'ordinateur.

```
   ----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                         |
| Type notebook() for the GUI, and license() for information.          |
----------------------------------------------------------------------
sage: A=matrix(3,3,[2,1,1,2,3,2,1,1,2])
sage: M=A.transpose()*A
sage: M
[ 9  9  8]
[ 9 11  9]
[ 8  9  9]
sage: M.characteristic_polynomial()
x^3 - 29*x^2 + 53*x - 25
sage: solve(  x^3 - 29*x^2 + 53*x - 25,x)
[x == -3*sqrt(19) + 14, x == 3*sqrt(19) + 14, x == 1]
```

Les valeurs propres de cette matrice $M = A^tA$ sont

$$\{\lambda_i\} = \{1, 14 + 3\sqrt{19}, 14 - 3\sqrt{19}\}. \tag{91.28}$$

La plus grande valeur propre est donc $14 + 3\sqrt{19}$, et la norme 2 de la matrice est

$$\|A\|_2 = \sqrt{\rho(A^tA)} = \sqrt{14 + 3\sqrt{19}} \simeq 5.203 \tag{91.29}$$

Exercise 206 exoGeomAnal-0044

Soit $\|.\|$ une norme algébrique sur $\mathbb{M}_n(\mathbb{R})$. Montrer que pour toute matrice $A \in \mathbb{M}_n(\mathbb{R})$, on a

$$\rho(A) \leqslant \|A\|. \tag{91.30}$$

Indices :

(1) Montrer que si une matrice B vérifie $\|B\| < 1$ alors $\mathbb{1} - B$ est inversible. Pour cela, considérez la série $\sum_{k=0}^{\infty} B^k$, montrez qu'elle converge vers un inverse de B. Par convention, $B^0 = \mathbb{1}$.

(2) Si λ est valeur propre de A, que peut-on dire de $A - \lambda\mathbb{1}$?

corrGeomAnal-0044

Correction of the exercise 206

Nous considérons l'espace vectoriel normé

$$\big(\mathbb{M}_n(\mathbb{R}), \|.\|\big). \tag{91.31}$$

Nous acceptons que cet espace est **complet**, c'est-à-dire que toute suite de Cauchy dans cet espace converge vers un élément de cet espace.

(1) Soit la suite des sommes partielles dans $\mathbb{M}_n(\mathbb{R})$

$$x_n = \sum_{k=0}^{n} B^k. \tag{91.32}$$

Nous allons prouver que cette suite est de Cauchy : $x_n - x_m = \sum_{k=n}^{m} B^k$, nous avons

$$\|x_n - x_m\| \leqslant \sum_{k=n}^{m} \|B^k\| \leqslant \sum_{k=n}^{m} \|B\|^k. \tag{91.33}$$

Étant donné que $\|B\|$ est un réel plus petit que 1, nous sommes dans le cas d'une série de puissance réelles et nous savons que pour tout $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que $n, m \geqslant N$ implique

$$\sum_{k=n}^{m} \|B\|^k < \epsilon. \tag{91.34}$$

Nous savons donc que la suite (x_n) est de Cauchy, et nous acceptions que cela prouve qu'elle converge vers un élément de $\mathbb{M}_n(\mathbb{R})$.
Prouvons à présent que $\sum_{k=0}^{\infty} B^k$ est l'inverse de $\mathbb{1} - B$. Pour cela, remarquons que

$$\begin{aligned}(\mathbb{1} - B)\sum_{k=0}^{N} B^k &= \sum_{k=0}^{N} B^k - \sum_{k=1}^{N+1} B^k && (91.35a)\\ &= B^0 - B^{N+1} && (91.35b)\\ &= \mathbb{1} - B^{N+1}. && (91.35c)\end{aligned}$$

Par conséquent

$$\begin{aligned}(\mathbb{1} - B)\sum_{k=0}^{\infty} B^k &= \lim_{N\to\infty} (\mathbb{1} - B)\sum_{k=0}^{N} B^k\\ &= \lim_{N\to\infty} \mathbb{1} - B^{N+1}.\end{aligned} \tag{91.36}$$

Mais nous avons $\lim_{N\to\infty} B^{N+1} = 0$ parce que $\|B^{N+1}\| \leqslant \|B\|^{N+1} \to 0$.

(2) Si λ est une valeur propre de A, alors la matrice $A - \lambda\mathbb{1}$ n'est pas inversible. Évidement si $\lambda = 0$ nous avons $\lambda < \|A\|$. Nous supposons que $\lambda \neq 0$. La matrice

$$\mathbb{1} - \frac{A}{\lambda} \tag{91.37}$$

n'est pas non plus inversible, ce qui implique que

$$\left\| \frac{A}{\lambda} \right\| \geqslant 1, \tag{91.38}$$

et par conséquent $\|A\| \geqslant \lambda$.

91.1.2 Topologie

Exercise 207 ⊙⊙ exoEspVectoNorme0004
Soit $(x_n)_{n\in\mathbb{N}}$ une suite convergente dans $\mathbb{R}^N$.

(1) Montrer que $(x_n)_{n\in\mathbb{N}}$ est bornée.

(2) Montrer que la suite de réels $(\|x_n\|)_{n\in\mathbb{N}}$ est convergente. Examiner la réciproque.
corrEspVectoNorme0004

Correction of the exercise 207

(1) Si la suite $(x_n)_{n\in\mathbb{N}}$ n'est pas bornée alors pour tout $M \in \mathbb{N}$ il existe n_M tel que $\|x_{n_M}\| > M$. La sous-suite $(x_{n_M})_{M\in\mathbb{N}}$ ne converge pas dans $\mathbb{R}$, donc $(x_n)_{n\in\mathbb{N}}$ n'est pas convergente. Cela veut dire que être bornée est une condition nécessaire pour la convergence.

(2) Soit $L \in \mathbb{R}^N$ la limite de la suite $(x_n)_{n\in\mathbb{N}}$. Alors pour chaque $\varepsilon > 0$ fixé il existe un $\bar{n}$ tel que si $n \geqslant \bar{n}$

$$\|x_n - L\| \leqslant \varepsilon. \tag{91.39}$$

La suite $(\|x_n\|)_{n\in\mathbb{N}}$ est convergente parce que pour tout $n \geqslant \bar{n}$ on a

$$|\|x_n\| - \|L\|| \leqslant \|x_n - L\| \leqslant \varepsilon. \tag{91.40}$$

On a utilisé la proposition 7.159.

Notez que la suite $\|x_n\|$ peut être convergente sans que la suite (x_n) soit elle-même convergente. Par exemple, la suite $x_n = (-1)^n$ n'est pas convergente, tandis que la suite des normes $\|x_n\| = 1$ est constante, et donc convergente.

Exercise 208 ⊙⊙ exoEspVectoNorme0005

On définit dans $\mathbb{R}^2$ la suite $(u_n) = (x_n, y_n)$ par

$$\begin{cases} x_{n+1} = x_n^2 + y_n^2 & \text{(91.41a)} \\ y_{n+1} = 2x_n y_n & \text{(91.41b)} \end{cases}$$

avec comme valeur initiale $(x_0, y_0) = (1/2, 1/2)$. Montrer que la suite (u_n) est convergente et calculer sa limite. corrEspVectoNorme0005

Correction of the exercise 208

Nous commençons par calculer le premier terme de la suite,

$$(x_1, y_1) = \left(\frac{1}{4} + \frac{1}{4}, 2\frac{1}{4}\right) = \left(\frac{1}{2}, \frac{1}{2}\right) = (x_0, y_0). \tag{91.42}$$

La suite (x_n, y_n) est donc la suite constante. Sa limite est $(1/2, 1/2)$.

Exercise 209 ⊙⊙ exoEspVectoNorme0006

Dans $\mathbb{R}^2$ on considère l'ensemble $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0,\ y > 0 \text{ et } xy = 1\}$. Montrer que A est un fermé de $\mathbb{R}^2$. corrEspVectoNorme0006

Correction of the exercise 209

Première méthode Soit $(x_n, y_n)_{n\in\mathbb{N}}$ une suite dans A convergente dans $\mathbb{R}^2$ et appellons $(\bar{x}, \bar{y})$ sa limite. Comme le produit est une application bilinéaire continue on a $\lim_n x_n y_n = \bar{x}\bar{y}$, donc $\bar{x}\bar{y} = 1$. En outre, comme toutes les normes de $\mathbb{R}^2$ sont équivalentes, nous pouvons écrire que la convergence de $(x_n, y_n)_{n\in\mathbb{N}}$ signifie que pour tout $\varepsilon > 0$ fixé il existe un $\bar{n}$ tel que si $n \geqslant \bar{n}$ alors

$$\|(x_n, y_n) - (\bar{x}, \bar{y})\|_\infty \leqslant \varepsilon, \tag{91.43}$$

cela implique la convergence des suites $(x_n)_n$ et $(y_n)_n$ vers $\bar{x}$ et $\bar{y}$ respectivement, parce que pour tout $n \geqslant \bar{n}$

$$|x_n - \bar{x}| \leqslant \max\{|x_n - \bar{x}|, |y_n - \bar{y}|\} = \|(x_n, y_n) - (\bar{x}, \bar{y})\|_\infty \leqslant \varepsilon, \tag{91.44}$$

le même calcul étant valable pour $\|y_n - \bar{y}\|$. Le point $(\bar{x}, \bar{y})$ est donc dans dans la fermeture de la region $x > 0$, $y > 0$, qui est $[0, \infty[\times[0, \infty[$. De plus la condition $\bar{x}\bar{y} = 1$ nous dit que ni $\bar{x}$ ni $\bar{y}$ est égal à zéro, donc $\bar{x} > 0$ et $\bar{y} > 0$. Bref, le point $(\bar{x}, \bar{y})$ est dans A. Nous disons alors que A est fermé parce qu'il contient tous ses points d'accumulation.

Deuxième méthode L'ensemble A est la branche d'hyperbole $y = 1/x$ située dans la partie $x > 0$ du plan. La méthode usuelle pour prouver qu'un ensemble est fermé est de prouver que son complémentaire est ouvert, et la méthode usuelle pour prouver qu'un ensemble est ouvert est de montrer qu'autour de chaque point de l'ensemble, il existe une boule contenue dans l'ensemble.

Les points du complémentaire de A sont les points (a, b) de $\mathbb{R}^2$ qui satisfont une des trois propriétés suivantes

$$a \leqslant 0 \tag{91.45a}$$
$$b \leqslant 0 \tag{91.45b}$$
$$ab \neq 1. \tag{91.45c}$$

subExynec106

Si $ab \neq 1$, il existe une boule autour de (a, b) dans laquelle tout (x, y) satisfait $xy \neq 1$. En effet, nous pouvons considérer la fonction $f(x, y) = xy - 1$. Si cette fonction est strictement positive en un point, alors elle reste strictement positive dans une boule (proposition 12.494). Le problème est déjà réglé pour tous les points (x, y) qui satisfont à la propriété (91.45c).

Si un point (a, b) est hors de A parce que $a \leqslant 0$, alors soit $a = 0$ (dans ce cas, $ab \neq 1$ et nous retombons dans le cas précédent), soit $a < 0$. Dans ce dernier cas, il existe une boule centrée en (a, b) qui reste dans le demi-plan $x < 0$.

Le cas $b < 0$ se traite de la même manière. Nous avons donc montré qu'autour de chaque point hors de A, il existe une boule qui n'intersecte pas A. Cela signifie que le complémentaire de A est ouvert, et donc que A est fermé.

Exercise 210 ⊛⊛ exoEspVectoNorme0007

Soit A une partie non vide de $\mathbb{R}^N$ et soit $x \in \mathbb{R}^N$.

(1) Montrer qu'il existe $\inf_{y\in A} \|x - y\|$. On pose $d(x, A) = \inf_{y\in A} \|x - y\|$.

(2) Montrer que $d(x, A) = 0$ si et seulement si $x \in \overline{A}$.

La fonction d ainsi définie donne la distance entre un point et un ensemble. Cet exercice justifie la définition (7.231).

corrEspVectoNorme0007

Correction of the exercise 210

(1) La proposition 1.461 nous indique que l'ensemble $\{\|x - y\| \text{ tel que } y \in A\}$ aura un infimum si et seulement si il est borné par le bas. Cela est certainement le cas parce que c'est un ensemble de nombres positifs, qui accepte donc zéro pour minorant.

(2) Si x n'est pas dans $\bar{A}$, il existe une boule de rayon δ autour de x qui n'intersecte pas A. Or, par définition,
$$B(x, \delta) = \{y \in \mathbb{R}^N \text{ tels que } \|y - x\| < \delta\}, \tag{91.46}$$
donc tous les points de A sont à distance au moins égale à δ de x (les points à distance moins de δ seraient dans la boule). Donc δ est un minorant de l'ensemble $\{\|x - y\| \text{ tels que } y \in A\}$, et l'infimum de cet ensemble ne peut pas avoir zéro comme infimum.

Montrons à présent que les points x de $\bar{A}$ sont tels que $d(x, A) = 0$. Si $x \in \bar{A}$, nous avons une suite (x_n) dans A qui converge vers x. Cette convergence signifie que $\lim \|x - x_n\| = 0$ (ne pas confondre $\lim x_n = x$ qui est une limite dans $\mathbb{R}^N$ et $\lim \|x_n - x\| = 0$ qui est une limite dans $\mathbb{R}$).

Pour tout ε, il existe donc un x_n dans A tel que $\|x_n - x\| \leqslant \varepsilon$, et donc $\inf_{y\in A} \|x - y\| \leqslant \varepsilon$. Cela prouve que cet infimum est zéro (parce que nous savons par ailleurs qu'il est positif ou nul).

Exercise 211 ⊛⊛ exoEspVectoNorme0009

On identifie $M_n(\mathbb{R})$ avec $\mathbb{R}^{n^2}$; plus précisément, on identifie une matrice

$$A = (a_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant n} \tag{91.47}$$

avec le vecteur $x = (x_1, x_2, \ldots, x_{n^2}) \in \mathbb{R}^{n^2}$, où $a_{i,j} = x_{(n-1)i+j}$. Soit $GL_n(\mathbb{R})$ l'ensemble des matrices $n \times n$ inversibles. Montrer que $GL_n(\mathbb{R})$ est un ouvert de $\mathbb{R}^{n^2}$. corrEspVectoNorme0009

Correction of the exercise 211

L'identification entre les vecteurs et les matrices consiste simplement à « déplier » la matrice pour en faire un vecteur. Par exemple, en dimension deux,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \in \mathbb{R}^4. \tag{91.48}$$

En dimension 3,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix} \in \mathbb{R}^9. \tag{91.49}$$

Une matrice est inversible si et seulement si son déterminant est non nul. Or le déterminant est un polynôme en les composantes de la matrice. En dimension deux, nous avons

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc, \tag{91.50}$$

mais en écriture « dépliée », nous pouvons aussi bien écrire

$$\det \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = ad - bc. \tag{91.51}$$

En dimension 3, le déterminant est donc un polynôme des 9 variables qui apparaissent dans le vecteur « déplié ». En général, dans $\mathbb{R}^{n^2}$, nous considérons donc le polynôme $\det\colon \mathbb{R}^{n^2} \to \mathbb{R}$ qui à un vecteur $X \in \mathbb{R}^{n^2}$ fait correspondre le déterminant de la matrice obtenue en « repliant » le vecteur X.

Donc dans $\mathbb{R}^{n^2}$, l'ensemble des matrices inversibles est donné par l'ensemble des vecteurs sur lesquels le polynôme det ne s'annule pas, c'est-à-dire

$$\{X \in \mathbb{R}^{n^2} \text{ tels que } \det(X) \neq 0\}. \tag{91.52}$$

Mais le déterminant est un polynôme, et donc une fonction continue. Cet ensemble est par conséquence ouvert par le corolaire 12.495.

Exercise 212 exoOutilsMath-0002

Prouver que pour tout vecteur X et tout scalaire λ, nous avons

$$\|\lambda X\| = |\lambda| \|X\|. \tag{91.53}$$

corrOutilsMath-0002

Correction of the exercise 212

<+CorrOutilsMath-0002+>

Exercise 213 exoOutilsMath-0003

Prouver que $\sin(30^\circ) = \frac{1}{2}$ et $\cos(30^\circ) = \frac{\sqrt{3}}{2}$ en vous inspirant de la figure 91.1. corrOutilsMath-0003

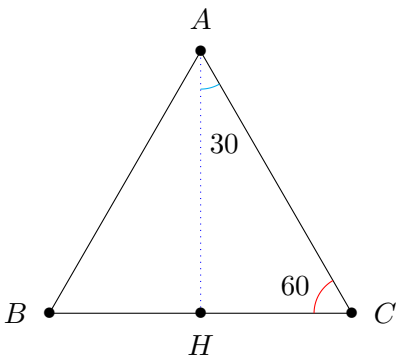


FIGURE 91.1: Un triangle équilatéral de côté 1. LabelFigTriangleRectangle

Correction of the exercise 213

Le segment AC a une longueur 1, et le segment HC une longueur $\frac{1}{2}$ parce que nous sommes dans un triangle équilatéral. Si nous nommons x la longueur de la hauteur AH, le théorème de Pythagore dit que

$$\left(\frac{1}{2}\right)^2 + x^2 = 1^2, \tag{91.54}$$

donc $x^2 = 1 - \frac{1}{4} = \frac{3}{4}$, d'où le fait que

$$\sin(60^\circ) = \frac{\sqrt{3}}{2}. \tag{91.55}$$

Afin de trouver le cosinus de 30° nous utilisons la relation

$$\sin^2(30) + \cos^2(30) = 1, \tag{91.56}$$

donc $\cos^2(30^\circ) = \frac{1}{4}$, et le résultat en découle.

Exercise 214 exoOutilsMath-0005

Donner les coordonnées cylindriques des points suivants :

(1) $(x, y, z) = (6, 6, 8)$.

corrOutilsMath-0005

Correction of the exercise 214

(1) $(r, \theta, z) = (6\sqrt{2}, \pi/4, 8)$

Exercise 215 exoOutilsMath-0006

Dessiner la trajectoire d'un mobile dont la position en fonction du temps est donnée en coordonnées cylindriques par

$$\varphi(t) = (R, t, t). \tag{91.57}$$

corrOutilsMath-0006

Correction of the exercise 215

<+CorrOutilsMath-0006+>

91.2 Courbes et surfaces

Exercise 216 exoCourbesSurfaces0001

Les courbes qui ont les mêmes graphes ne sont pas toujours équivalentes. Quelques exemples.

(1) On considère les courbes planes $([0, 1], f)$ et $([-1, 1], g)$ définies par

$$f(t) = (t, t) \quad \text{et} \quad g(t) = (t^2, t^2). \tag{91.58}$$

Montrer que les deux courbes ont le même graphe (image), mais ne sont pas équivalentes.

(2) Même question avec les courbes $([0,2\pi],f_1)$ et $([0,6\pi],g_1)$ définies par $f_1(t)=(\cos t,\sin t)$ et $\vec{g_1}(t)=(\cos t,\sin t)$.

(3) Montrer que les courbes $]-\pi,\pi[,f)$ et $(\mathbb{R},g)$ définies par

$$f(t)=(\cos t,\sin t)\quad\text{et}\quad g(t)=\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)\tag{91.59}$$

sont équivalentes.

Indice : parmi les formules de trigonométries, nous rappelons

$$\frac{1-\tan^2\left(\frac{x}{2}\right)}{1+\tan^2\left(\frac{x}{2}\right)}=\cos(x)\tag{91.60a}$$

$$\sin(2x)=2\sin(x)\cos(x).\tag{91.60b}$$

corrCourbesSurfaces0001

Correction of the exercise 216

(1) Nous devons prouver que $\text{Graph}(f)=\text{Graph}(g)$. Si $(t,t)\in\text{Graph}(f)$, alors $(t,t)=g(\sqrt{t})$ où $\sqrt{t}$ existe parce que $t\geqslant 0$. Nous avons donc $\text{Graph}(f)\subset\text{Graph}(g)$. Inversement, si $(t^2,t^2)\in\text{Graph}(g)$, alors $(t^2,t^2)=f(t^2)$ où $t^2\in[0,1]$ lorsque $t\in[-1,-1]$.
Les graphes sont donc identiques. Si ils étaient équivalents (définition 21.35), alors il y aurait une application inversible $\theta\colon[-1,1]\to[0,1]$ telle que $g(t)=f\big(\theta(t)\big)$, c'est-à-dire

$$(t^2,t^2)=\big(\theta(t),\theta(t)\big).\tag{91.61}$$

Cela impose $\theta(t)=t^2$, mais cette fonction n'est pas inversible sur $[-1,1]$.

(2) Les deux graphes sont évidement les mêmes parce que les valeurs de t entre 2π et 6π ne créent aucun nouveau points (périodicité de sinus et cosinus). En ce qui concerne l'équivalence des deux courbes, nous devons avoir une application $\theta\colon[0,6\pi]\to[0,2\pi]$ telle que

$$\big(\cos(t),\sin(t)\big)=\big(\cos\theta(t),\sin\theta(t)\big),\tag{91.62}$$

ce qui impose $\theta(t)=t+2k\pi$ pour un certain k entier. Aucune de ces fonctions n'est une bijection entre $[0,2\pi]$ et $[0,6\pi]$.

(3) Il nous faut une fonction $\theta\colon]-\pi,\pi[\to\mathbb{R}$ telle que $g\big(\theta(t)\big)=f(t)$, c'est-à-dire

$$\left(\frac{1-\theta^2(t)}{1+\theta^2(t)},\frac{2\theta(t)}{1+\theta^2(t)}\right)=\big(\cos(t),\sin(t)\big).\tag{91.63}$$

En résolvant pour la première composante, nous trouvons

$$\theta(t)=\pm\sqrt{\frac{1-\cos(t)}{1+\cos(t)}}\tag{91.64}$$

Eqthetafromunzzzucs

où le signe est à discuter en fonction du signe de $\sin(t)$. En utilisant l'indice, nous savons que cela se récrit avantageusement sous la forme

$$\theta(t)=\tan\left(\frac{t}{2}\right).\tag{91.65}$$

Sous cette forme, il est immédiatement apparent que c'est un difféomorphisme. Il reste à vérifier que

$$\sin(t)=\frac{2\theta(t)}{1+\theta^2(t)}.\tag{91.66}$$

En utilisant θ sous la forme (91.64), cela se vérifie très facilement.

Exercise 217 ⊙⊙ exoCourbesSurfaces0002

Trouver le domaine de définition des fonctions $x(t)$ et $y(t)$, et tracer les courbes suivantes :

(1) $x(t) = t + \dfrac{1}{t}$, $y(t) = t + \frac{1}{2t^2}$ CourbSDi CourbSDii

(2) $x(t) = \cos^2(t)$, $y(t) = \cos^3(t)\sin(t)$ CourbSDvii

(3) $x(t) = \sin(2t)$, $y(t) = \sin(3t)$ corrCourbesSurfaces0002

Correction of the exercise 217

Pour esquisser le graphe d'une courbe paramétrique il faut d'abord étudier

(1) les ensembles de définition des fonctions $x(t)$ et $y(t)$,

(2) la périodicité, la parité de ces fonctions,

(3) le sens de variation de x et de y,

(4) les points particuliers correspondant à des valeurs remarquables de t, les tangentes en ces points (s'il y en a),

(5) les éventuels points d'inflexion et/ou de rebroussement,

(6) les branches infinies (s'il y en a) et l'existence des asymptotes.

Parfois, si la courbe est compliquée, il faut aussi calculer en quelques points la tangente et la normale. La normale est la perpendiculaire à la tangente, et comme expliqué dans la section 21.11, la concavité est tournée dans le sens de la normale dirigée dans le même sens que la dérivée seconde. Itemzzdexoi

(1) Les fonctions $x(t)$ et $y(t)$ sont définies sur le même domaine $\mathbb{R}\backslash\{0\}$. Cela veut dire que les équations décrivent l'union de deux courbes. La fonction $x(t)$ est une fonction impaire et $x(t) = x(1/t)$. La fonction $y(t)$ n'a pas de symétries évidentes. L'étude des dérivées

$$\begin{aligned} x'(t) &= 1 - \frac{1}{t^2}, \\ y'(t) &= 1 - \frac{1}{t^3}, \end{aligned} \tag{91.67}$$

nous donne plusieurs indices. La fonction x est croissante pour $|t| > 1$ et décroissante si $|t| < 1$. La fonction y est croissante pour $t < 0$ et pour $t > 1$, sinon elle est décroissante. Le point $t = 1$ est le minimum global pour les deux fonctions. Cela veut dire que la courbe de support $]0, +\infty[$ n'est jamais ni à gauche de $x = 2$ ni au dessus de $y = 3/2$. Le point $t = -1$ est un point de maximum global pour la fonction x (la courbe passe par le point $(-2, -1/2)$). La courbe de support $]-\infty, 0[$ n'est jamais à droite de $x = -2$.

Les limites aux extrêmes du domaine sont les suivantes

$$\begin{aligned} &\lim_{t\to 0^\pm} x(t) = \pm\infty, \quad \lim_{t\to\pm\infty} x(t) = \pm\infty, \\ &\lim_{t\to 0^\pm} y(t) = +\infty, \quad \lim_{t\to\pm\infty} y(t) = \pm\infty. \end{aligned} \tag{91.68}$$

Les courbes ne sont donc pas bornées. On peut étudier leur comportement asymptotique. Si t est très grand les termes $1/t$ et $1/(2t^2)$ sont très faibles. La courbe ressemble fort à la droite $x = y$ lorsque t est grand. La même chose arrive quand t est négatif et $|t|$ est très grand.

Au contraire, si t et positif et proche de zéro, les termes $1/t$ et $1/(2t^2)$ deviennent dominants. Le comportement de la courbe est alors très proche du comportement de la fonction $y = x^2/2$ pour $x > 0$. Si t est négatif et de valeur absolue petite alors la courbe ressemble encore beaucoup à $y = x^2/2$, pour $x < 0$. Bien que $y(t)$ ne soit pas une fonction paire elle devient « presque » paire là où son terme dominant est $1/(2t^2)$, qui est pair.

Le résultat est sur la figure 91.2.

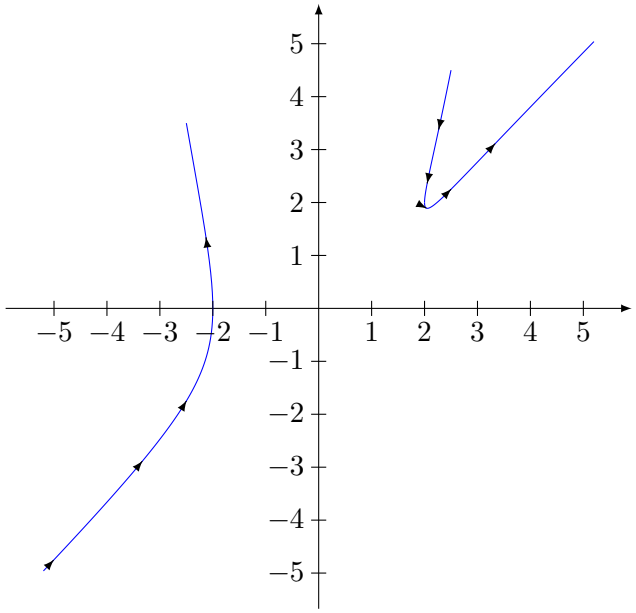


FIGURE 91.2: La courbe de l'exercice 217(1) LabelFigCbCartTui

Une difficulté de l'exercice est de déterminer, pour les t entre zéro et un, quelle branche est au dessus de l'autre ? Comment justifier le sens des flèches dans la figure 91.3 ?
Si $\ell > 2$, il y a deux valeurs de $t \in]0, \infty[$ pour lesquelles $x(t) = \ell$; elles sont données par

$$t_1 = \frac{\ell + \sqrt{\ell^2 - 4}}{2} > 1, \quad t_2 = \frac{\ell - \sqrt{\ell^2 - 4}}{2} < 1. \tag{91.69}$$

Nous pouvons calculer :

$$y(t_1) - y(t_2) = \sqrt{\ell^2 - 4} + \frac{-2\ell\sqrt{\ell^2 - 4}}{(\ell^2 - 2)^2 - \ell^2(\ell^2 - 4)} = \sqrt{\ell^2 - 4} - \frac{\ell\sqrt{\ell^2 - 4}}{2}. \tag{91.70}$$

Étant donné que $\ell > 2$, le second terme est plus grand que le premier et la différence est négative. Par conséquent,

$$y(t_1) < y(t_2), \tag{91.71}$$

et donc la branche des $t > 1$ est en dessous de la branche des $y < 1$, ce qui justifie le sens des flèches. Itemzzdexoii

(2) Le domaine des fonctions $x(t)$ et $y(t)$ est $\mathbb{R}$. Les deux fonctions sont périodique de période π. Pour le voir il suffit de remarquer que $2y(t) = x(t)\sin(2t)$. Étant donnée la périodicité, nous considérons seulement la portion du domaine $[0, \pi]$. On peut déjà calculer $x(0) = 1$ et $y(0) = 0$. La fonction x est toujours positive. Cela veut dire que la courbe est toute à droite de l'axe des y. La fonction y est positive pour t dans $[0, \pi/2]$ et négative pour t dans $[\pi/2, \pi]$.
Les dérivées sont

$$\begin{aligned} x'(t) &= -2\cos(t)\sin(t), \\ y'(t) &= -3\cos^2(t)\sin^2(t) + \cos^4(t). \end{aligned} \tag{91.72}$$

On remarque que la droite tangente à la courbe est verticale pour $t = 0$ et $t = \pi$.
La valeur $t = \pi/2$ est critique pour les deux fonctions, $x(\pi/2) = y(\pi/2) = 0$. La fonction x est décroissante pour $t \in]0, \pi/2[$ et croissante pour $t \in]\pi/2, \pi[$. La valeur $t = \pi/2$ n'est pas une valeur de maximum ou de minimum pour y.
Trouvons les autres points critiques de la fonction y, c'est-à-dire résolvons $y'(t) = 0$. Les points correspondants à $\cos(t) = 0$ ayant déjà été discutés, nous pouvons simplifier par $\cos^2(t)$ et résoudre

$$-3\sin^2(t) + \cos^2(t) = 0. \tag{91.73}$$

Il y a au moins deux façons de procéder. La première consiste à remplacer $\cos^2(t)$ par $1-\sin^2(t)$ et puis résoudre par rapport à $\sin(t)$. Dans ce cas nous tombons sur[2] $\sin(t)=\pm\frac{1}{2}$ et par conséquent $t_1=\pi/6$ et $t_2=5\pi/6$. La seconde méthode consiste à encore diviser par $\cos^2(t)$ et tomber sur $\tan^2(t)=\frac{1}{3}$. Nous trouvons évidement les mêmes solutions.

Un calcul montre que y est croissante sur $[0,t_1]$ et $[t_2,\pi/2]$, sinon elle est décroissante.

Lorsque t s'approche de $\pi/2$ la valeur $\sin(t)$ s'approche de 1. Cela veut dire que $y(t)$ devient presque $\cos^3(t)$ et la courbe ressemble à $y=x^{3/2}$ si $\cos(t)$ est positif et à $y=-x^{3/2}$ si $\cos(t)$ est négatif.

Le résultat est à la figure 91.3.

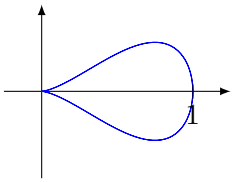


FIGURE 91.3: La courbe de l'exercice (2).

LabelFigCbCartTuii

Itemzzdexoiii

(3) La période de $x(t)$ est π et la période de $y(t)$ est $2\pi/3$. La période de la courbe qu'on va tracer sera alors 2π qui est le plus petit multiple commun et non nul de π et de $2\pi/3$.

Afin de simplifier les calculs, nous pouvons remarquer que

$$x(t+\pi)=x(t) \tag{91.74a}$$
$$y(t+\pi)=-y(t). \tag{91.74b}$$

La courbe a donc une symétrie par rapport à l'axe horizontal, et il est suffisant d'étudier la courbe entre 0 et π. Le reste se déduisant par symétrie. Voir la figure 91.4.

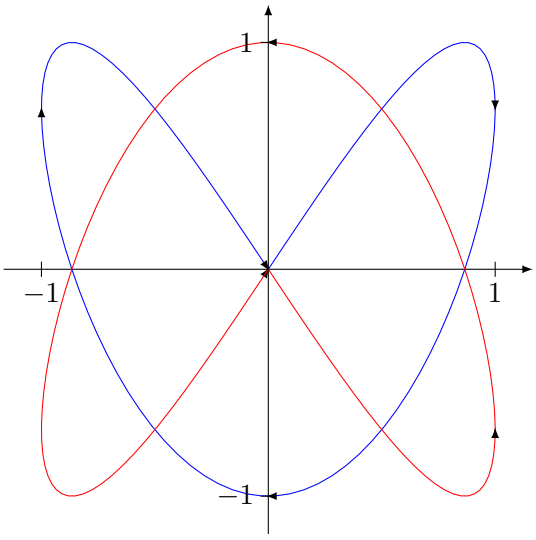


FIGURE 91.4: Le graphique de la courbe de l'exercice 217(3). La courbe en bleu représente les valeurs du paramètre entre 0 et π tandis qu'en rouge nous avons tracé le paramètre entre π et 2π.

LabelFigExoParamCb

Dans cette correction nous allons cependant faire « comme si » nous n'avions pas remarqué la symétrie.

Une simple, mais longe, étude des signes de x et y nous dit que la courbe $\gamma(t)=(x(t),y(t))$ sera dans le premier quadrant pour t contenu dans l'union des intervalles

$$I_1=\left[0,\frac{\pi}{3}\right],\qquad I_6=\left[\frac{4\pi}{3},\frac{3\pi}{2}\right].$$

2. Nous nous sommes restreints à $t\in[0,\pi]$.

De mme $\gamma(t)$ est dans le deuxième quadrant si t est dans

$$I_4 = \left[\frac{2\pi}{3}, \pi\right], \qquad I_8 = \left[\frac{5\pi}{3}, 2\pi\right],$$

$\gamma(t)$ est dans le troisième quadrant si t est dans

$$I_3 = \left[\frac{\pi}{2}, \frac{2\pi}{3}\right], \qquad I_7 = \left[\frac{3\pi}{2}, \frac{5\pi}{3}\right],$$

et enfin $\gamma(t)$ est dans le quatrième quadrant si t est dans

$$I_2 = \left[\frac{\pi}{3}, \frac{\pi}{2}\right], \qquad I_5 = \left[\pi, \frac{4\pi}{3}\right].$$

Il faut donc trouver tous les 9 points d'intersection de γ avec les axes

$$\begin{array}{lll} \gamma(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \gamma\left(\frac{\pi}{3}\right) = \begin{pmatrix} \sin\left(\frac{2\pi}{3}\right) \\ 0 \end{pmatrix} & \gamma\left(\frac{\pi}{2}\right) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ \gamma\left(\frac{2\pi}{3}\right) = \begin{pmatrix} \sin\left(\frac{4\pi}{3}\right) \\ 0 \end{pmatrix} & \gamma(\pi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \gamma\left(\frac{4\pi}{3}\right) = \begin{pmatrix} \sin\left(\frac{8\pi}{3}\right) \\ 0 \end{pmatrix} \\ \gamma\left(\frac{3\pi}{2}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \gamma\left(\frac{5\pi}{3}\right) = \begin{pmatrix} \sin\left(\frac{10\pi}{3}\right) \\ 0 \end{pmatrix} & \gamma(2\pi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{array} \tag{91.75}$$

L'étude des dérivées premières

$$\begin{cases} x'(t) = 2\cos(2t), \\ y'(t) = 3\cos(3t), \end{cases} \tag{91.76}$$

nous dit que la fonction x a quatre points critiques dans $[0, 2\pi]$

$$\tfrac{\pi}{4}(\max) \quad \tfrac{3\pi}{4}(\min) \quad \tfrac{5\pi}{4}(\max) \quad \tfrac{7\pi}{4}(\min), \tag{91.77}$$

et que la fonction y en a six

$$\begin{array}{lll} \frac{\pi}{6}(\max) & \frac{\pi}{2}(\min) & \frac{5\pi}{6}(\max) \\ \frac{7\pi}{6}(\min) & \frac{3\pi}{2}(\max) & \frac{11\pi}{6}(\min). \end{array} \tag{91.78}$$

Les points critiques de x correspondent aux points de γ où la droite tangente est verticale. De même les points critiques de y correspondent aux points de γ où la droite tangente est horizontale. Pour bien tracer la courbe il faut calculer le points de γ pour tous ces dix valeurs de t (omis).

Il faut aussi trouver les droites tangentes à γ aux points d'intersection avec les axes. Ces information sont suffisantes pour esquisser γ.

Le résultat est sur la figure 91.5.

Exercise 218 ⊛⊛ exoCourbesSurfaces0003

On fait rouler sans glissement un cercle de rayon 1 sur l'axe (Ox). Déterminer et tracer la courbe décrite par un point du cercle (cette courbe est appelée **cycloïde**). corrCourbesSurfaces0003

Correction of the exercise 218

Si nous supposons une vitesse angulaire ω, la roue avance horizontalement d'une longueur ωt en un temps t. Si le centre du cercle se trouve en $(0, 1)$ au départ, cela signifie que le centre de la roue décrit le mouvement $C(t) = (\omega t, 1)$. Un point du cercle décrit un cercle autour de ce centre. Si nous suivons le point P situé en $(0, 1)$ au départ, nous avons l'équation $P(t) = C(t) + \big(\cos(\omega t), \sin(\omega t)\big)$. En posant $\omega = 1$, nous avons

$$\begin{cases} x(t) = t + \cos(t) & \text{(91.79a)} \\ y(t) = 1 + \sin(t). & \text{(91.79b)} \end{cases}$$

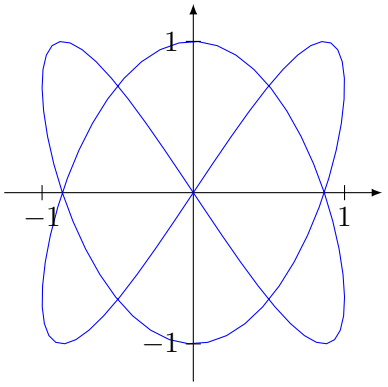
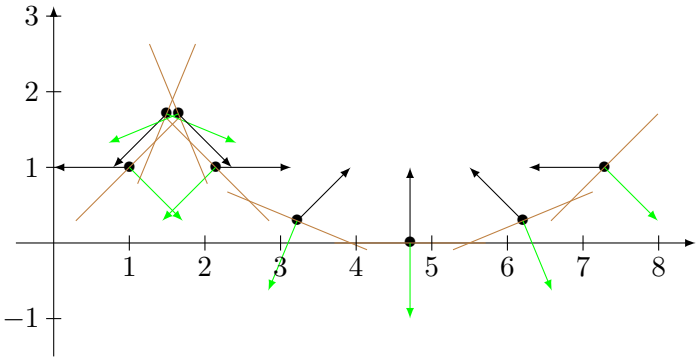
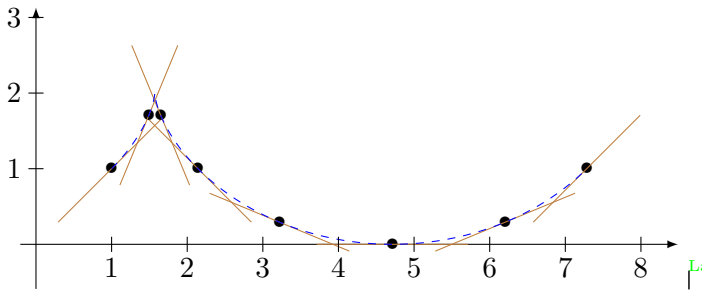


FIGURE 91.5: Le dessin de l'exercice (3). LabelFigCbCartTuiii



LabelFigSubfiguresCDUTraceCycloidessSS

(a) Quelques points de repères



LabelFigSubfiguresCDUTraceCycloidessSS2T

(b) La courbe

FIGURE 91.6: La cycloïde de l'exercice 218. LabelFigSubfiguresCDUTraceCycloide

Il est facile de trouver quelques points. Les dérivées premières et secondes donnent des indices :

$$\begin{cases} x'(t) = 1 - \sin(t) & (91.80a) \\ y'(t) = \cos(t). & (91.80b) \end{cases}$$

et

$$\begin{cases} x''(t) = -\cos(t) & (91.81a) \\ y''(t) = -\sin(t). & (91.81b) \end{cases}$$

Nous voyons que la fonction $x(t)$ est toujours croissante. La courbe que nous cherchons n'aura donc pas d'auto-intersections. La dérivée seconde donne la concavité ; ses valeurs remarquables sont évidement tous les multiples de $\pi/2$.

Nous avons mis sur la figure 91.6(a) la tangente, la dérivée seconde et la normale intérieure en divers points. La courbe est tracée à la figure 91.6(b).

Exercise 219 exoCourbesSurfaces0005

Étudier (comme dans l'exercice 217) les courbes définies en coordonnées polaires (r, θ) par :

(1) $\odot\odot^{\text{CSCi}} r(\theta) = 1 + \cos\theta$ (la **cardioïde**)

(2) $\odot\odot^{\text{CSCii}} r(\theta) = \sin(2\theta)$

(3) $\circledast\circledast^{\text{CSCiii}} r(\theta) = \dfrac{\sin(\theta)}{\theta}$.

(4) $\circledast\circledast^{\text{CSCiv}} r(\theta) = \dfrac{\theta - 1}{\theta + 1}$

(5) $\circledast\circledast^{\text{CSCv}} r(\theta) = \cos(\theta) - \cos(2\theta)$

(6) $\odot\odot^{\text{CSCvi}} r(\theta) = \dfrac{\cos(\theta)}{1 + \sin(\theta)}$

corrCourbesSurfaces0005

Correction of the exercise 219

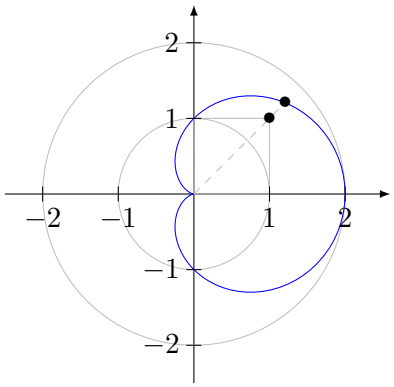


FIGURE 91.7: La cardioïde de l'exercice 219.(1). LabelFigCardioideexo

(1) La courbe est périodique de période 2π. La fonction $r(\theta)$ est symétrique au sens que $r(\theta) = r(2\pi - \theta)$. Cela est une conséquence du fait que la fonction $\cos(t)$ est paire. Donc nous trouvons une courbe symétrique par rapport à l'axe Ox. Nous avons ensuite $r(\pi) = 0$, ce qui nous pousse à étudier la courbe seulement sur l'intervalle $[0, \pi]$. Nous avons $r(0) = 2$ et $r(\pi/2) = 1$; cela nous fait déjà deux points. Comme la dérivée $r'(\theta) = -\sin(\theta)$ est négative sur $[0, \pi]$, la distance entre la courbe et l'origine est monotone décroissante. Il est intéressant de remarquer que $r(\frac{\pi}{4}) = 1 + \frac{1}{\sqrt{2}} > \sqrt{2}$, ce qui fait que la courbe passe *au dessus* du carré qui passerait par $(1, 1)$, voir figure 91.7.

(2) Une première chose à faire est de trouver quelles sont les valeurs acceptables de θ. Vu que r doit être positif, nous devons avoir $\theta \in [0, \pi/2]$ or $\theta \in [3\pi/2, 2\pi]$. La fonction $r(\theta)$ est périodique de période π, donc la courbe correspondant au deuxième intervalle n'est que une rotation de π de la première courbe. Pour cette raison, nous considérons ici seulement le cas $\theta \in [0, \pi/2]$. La fonction r est monotone croissante sur $\theta \in [0, \pi/4]$ et décroissante sinon. Nous avons $r(0) = 0$, $r(\frac{\pi}{4}) = 1$ et $r(\frac{\pi}{2}) = 0$. La courbe est dessinée sur la figure 91.8

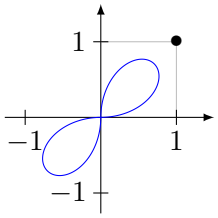


FIGURE 91.8: La courbe de l'exercice 219.(2). LabelFigCSCii

Notez qu'en $\theta = \pi/4$ nous avons $r = 1$, et que par conséquent la courbe passe au-dessous du point $(1, 1)$.

(3) La courbe ici n'est pas périodique. Pour décrire de façon efficace une courbe en coordonnées polaires il faut permettre au paramètre θ de prendre ses valeurs dans tout $\mathbb{R}$ et pas seulement dans l'intervalle $[0, 2\pi]$, par contre le rayon r ne pourra jamais être négatif. Le domaine de définition de la fonction r est donc l'union de tous les intervalles sur lesquels la fonction sin est positive moins le point $\theta = 0$

$$\text{Domaine}_r = \{[2k\pi, (2k+1)\pi] \,:\, k \in \mathbb{Z}\}\backslash\{0\}. \tag{91.82}$$

Étant donné que $r(-\theta) = r(\theta)$, la courbe est symétrique par rapport à l'axe Ox. Nous pouvons donc nous concentrer sur les valeurs positives de θ. Notre courbe se divisera donc en une infinité de branches. La première sera $]0, \pi]$, la seconde sera $[2\pi, 3\pi]$, etc.

Le « point de départ » de la la première branche est $\theta = 0$. Certes, $r(0)$ n'est pas défini, mais la limite existe :

$$\lim_{\theta \to 0} r(\theta) = \lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1. \tag{91.83}$$

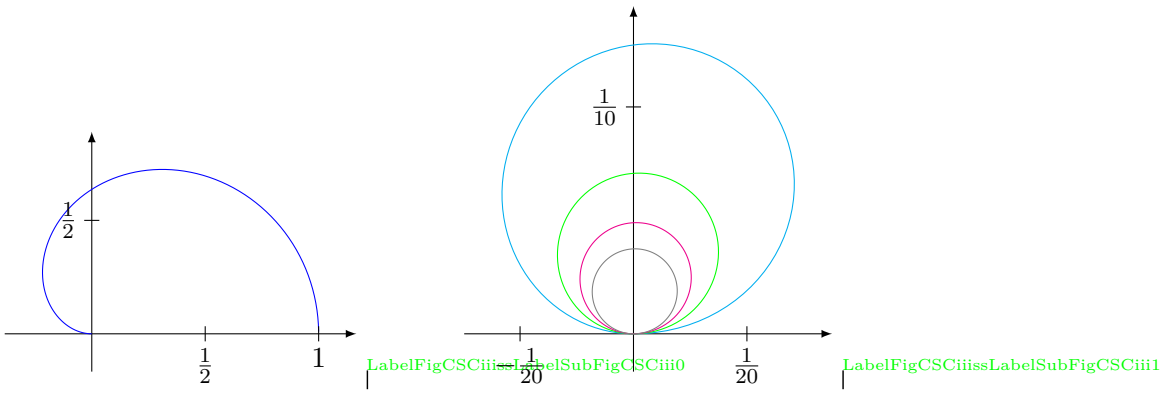
La courbe part donc du point $(1, 0)$.

En ce qui concerne la dérivée de r, nous avons

$$r'(\theta) = \frac{\cos(\theta)}{\theta} - \frac{\sin(\theta)}{\theta^2}. \tag{91.84}$$

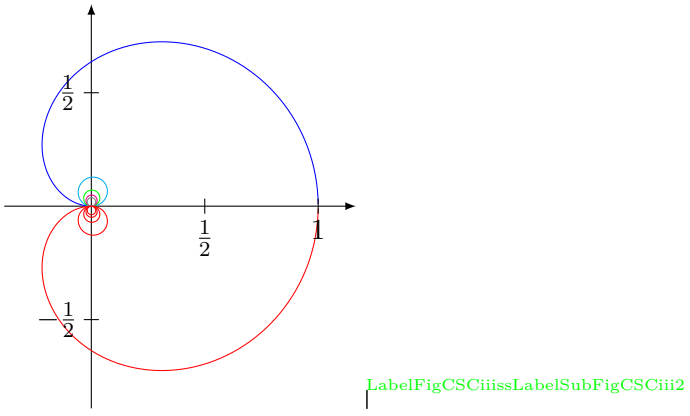
Cette fonction n'est jamais nulle dans l'intervalle $]0, \pi]$, par conséquent on sait que r sera une fonction monotone (décroissante) sur la première branche de la courbe. Calculons quelques valeurs. D'abord, $\lim_{\theta \to 0} r(\theta) = 1$, $r(\frac{\pi}{2}) = \frac{2}{\pi} < 1$ et $r(\pi) = 0$. Nous avons aussi $r(\frac{\pi}{4}) = \frac{2\sqrt{2}}{\pi} < 1$. La fonction $r'(\theta)$ est a un zéro dans chaque intervalle de la forme $[2k\pi, (2k+1)\pi]$ pour $k > 0$, en fait $r(n\pi) = 0$ pour tout $n > 0$ dans $\mathbb{N}$. Trouver ces points critiques est un peu plus difficile, mais on comprend sans difficulté est que la valeur du maximum locale de r est un nombre de plus en plus petit. Ces branches sont dont des courbes de longueur décroissante.

Le résultat est sur la figure 91.9.



(a) La première branche pour les angles positifs.

(b) Quelques autres branches pour les angles positifs.



(c) Le dessin au complet

FIGURE 91.9: La courbe de l'exercice 219.(3). Les deux premières branches sont dessinées. Notez que la seconde est beaucoup plus petite que la première, et qu'elle part de $(0, 0)$. LabelFigCSCiii

See also the subfigure 91.9(a) See also the subfigure 91.9(b) See also the subfigure 91.9(c)

(4) La première chose à faire est d'étudier le signe de la fonction $r(\theta)$ et de considérer le domaine où elle est positive. Un tableau de signe montre que r est positive lorsque $\theta \in]-\infty,-1[\cup [1,\infty[$. Commençons par le second morceau.
Pour $\theta = 1$ nous avons $r(\theta) = 0$. La courbe commence donc à l'origine et part directement selon un angle d'un peu moins de 90 degrés (1 radian est moins de l'angle droit). Ensuite, le rayon n'arrête pas d'augmenter, et la limite $\lim_{r\to\infty} r(\theta) = 1$. Nous avons donc une spirale qui devient de plus en plus dense lorsqu'elle s'approche du rayon 1.
En ce qui concerne l'autre partie, nous avons $\lim_{r\to-1^-} r(\theta) = \infty$, donc la courbe part vers l'infini en tendant vers la droite d'angle -1. D'autre par la limite $\lim_{r\to-\infty} r(\theta) = 1$. La courbe est donc une spirale qui vient s'accumuler autour du cercle de rayon 1 (mais depuis l'extérieur, cette fois).
Cette double spirale est dessinée à la figure 91.10.

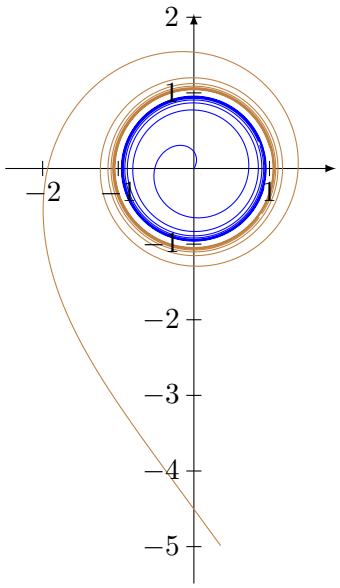
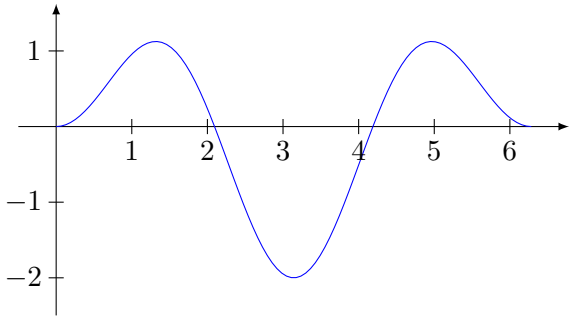
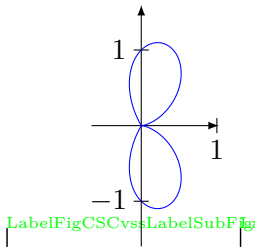


FIGURE 91.10: La courbe de l'exercice 219.(4).

LabelFigCSCiv



(a) Le graphique de $r(\theta)$



(b) La courbe polaire correspondante

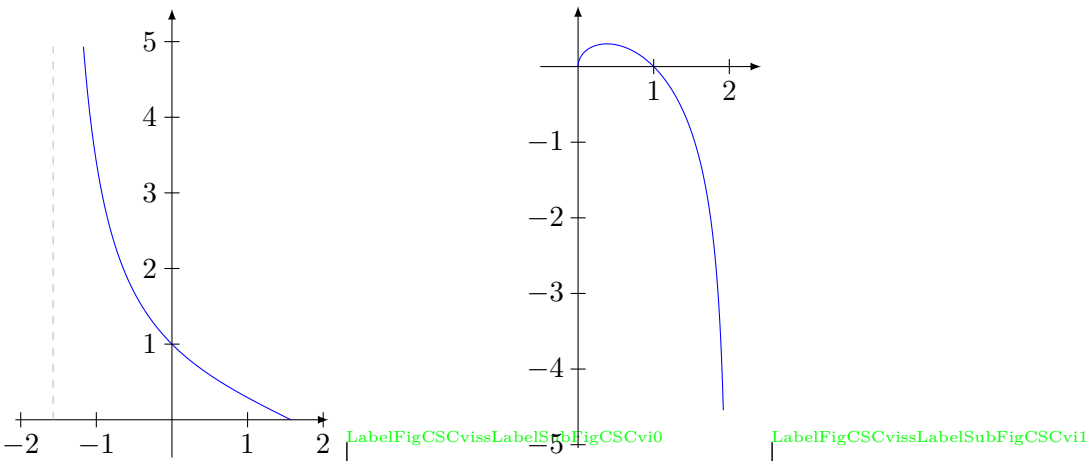
FIGURE 91.11: La courbe de l'exercice 219.(5).

LabelFigCSCv

(5) La chose la plus compliquée dans cet exercice est en réalité de tracer la courbe $r(\theta)$ dans le plan (r,θ). Elle est tracée à la figure 91.11(a). Nous trouvons les points d'annulation de $r(\theta)$ en remplaçant $\cos(2\theta)$ par $\sin^2(\theta) - \cos^2(\theta) = 1 - 2\cos^2(\theta)$. Ainsi nous avons l'équation

$$r(\theta) = -2\cos^2(\theta) + \cos(\theta) + 1. \tag{91.85}$$

En posant $y = \cos(\theta)$, cette fonction s'annule pour $y = 1$ ou $y = -\frac{1}{2}$, ce qui correspond aux valeurs $\theta = 0$, $\theta = 2\pi/3$ et $\theta = 4\pi/3$.



LabelFigCSCvissLabelSubFigCSCvi0 LabelFigCSCvissLabelSubFigCSCvi1

(a) Le graphique de $r(\theta)$ (b) La courbe polaire correspondante

FIGURE 91.12: La courbe de l'exercice 219.(5). LabelFigCSCvi

La courbe demandée est sur la figure 91.11(b).

(6) Étant donné que le dénominateur est toujours positif (ou nul), le domaine de variation acceptable de θ est la domaine sur lequel $\cos(\theta)$ est positif. Nous allons donc étudier la courbe pour θ : $-\frac{\pi}{2} \to \frac{\pi}{2}$. Lorsque θ tend vers $-\frac{\pi}{2}$ par des valeurs supérieures, $r(\theta)$ tend vers l'infini tandis que lorsque θ vaut zéro, le rayon $r(0) = 1$.

La fonction $r(\theta)$ est tracée sur la figure 91.12(a), et la courbe paramétrique correspondante est sur la figure 91.12(b).

Exercise 220 exoCourbesSurfaces0006

Quelques longueurs de courbes.

(1) Déterminer la longueur de la courbe $y = \sqrt{x}(1 - \frac{x}{3})$ pour $0 \leqslant x \leqslant 3$.

(2) Calculer la longueur de la néphroïde représentée paramétriquement par

$$\begin{cases} x(t) = 3\cos t - \cos(3t) & (91.86a) \\ y(t) = 3\sin t - \sin(3t) & (91.86b) \end{cases}$$

où $t \in [0, 2\pi]$.

corrCourbesSurfaces0006

Correction of the exercise 220

(1) Comme expliqué à la page 1969, nous voyons le graphe de la fonction $y = f(x)$ comme le chemin $\gamma(t) = \big(t, f(t)\big)$. Nous avons donc

$$\begin{aligned} \gamma(t) &= \big(t, \sqrt{t}(1 - \frac{t}{3})\big) \\ \gamma'(t) &= \big(1, \frac{1-t}{2\sqrt{t}}\big), \end{aligned} \tag{91.87}$$

et par conséquent l'élément de longueur à intégrer est

$$\|\gamma'(t)\| = \sqrt{1 + \frac{(1-t)^2}{4t}} = \frac{t+1}{2\sqrt{t}}, \tag{91.88}$$

et la longueur de l'arc considéré vaut

$$\int_0^3 \|\gamma'(t)\| = \Big[\frac{1}{3}t^{3/2} + t^{1/2}\Big]_{t=0}^{3} = \frac{1}{3}3^{3/2} + \sqrt{3} = 2\sqrt{3}. \tag{91.89}$$

(2) En calculant les dérivées, nous trouvons

$$\|\gamma'(t)\|^2 = 18\big(1 - \sin(t)\sin(3t) - \cos(t)\cos(3t)\big) \tag{91.90}$$

que l'on peut simplifier en utilisant des formules de trigonométrie

$$\|\gamma'(t)\|^2 = 36\sin^2(t). \tag{91.91}$$

Nous allons devoir intégrer la racine carrée de cela entre 0 et 2π. Il ne faut pas dire que $\sqrt{\sin^2(t)} = \sin(t)$ et ensuite calculer

$$\sqrt{36}\int_0^{2\pi} \sin(t)dt, \tag{91.92}$$

parce que $\sin(t)$ est négatif entre π et 2π. Ce qu'il faut faire est $\sqrt{\sin^2(t)} = |\sin(t)|$ et profiter du fait que $|\sin(t)|$ est symétrique par rapport à π :

$$l = 2\sqrt{36}\int_0^{\pi} \sin(t)dt = 4\sqrt{36} = 24. \tag{91.93}$$

Exercise 221 exoCourbesSurfaces0007

On considère une courbe plane $([\theta_1, \theta_2], \vec{\gamma})$ définie, en coordonnées polaires, par

$$\vec{\gamma}(\theta) = \big(r(\theta)\cos\theta, r(\theta)\sin\theta\big), \tag{91.94}$$

où r est une fonction C^1 sur $[\theta_1, \theta_2]$.

Quelques applications de la formule.

(1) Est-ce que $\vec{\gamma}$ est rectifiable ?

(2) Montrer que la longueur de cet arc est donné par

$$\ell(\vec{\gamma}) = \int_{\theta_1}^{\theta_2} \sqrt{r^2(\theta) + (r')^2(\theta)}\, d\theta. \tag{91.95}$$ EqFormDemExoLpola

(3) Calculer la longueur de la courbe d'équation (en coordonnées polaires) $r(\theta) = \cos^2\theta$, $\theta \in [0, \pi]$.

(4) Calculer la longueur de la cardioïde, définie en coordonnées polaires par $r(\theta) = 1 - \cos\theta$, où $\theta \in [-\pi, \pi]$.

corrCourbesSurfaces0007

Correction of the exercise 221

(1) Le théorème 21.9 nous renseigne qu'un arc paramétré compact C^1 est automatiquement rectifiable. Dans notre cas, la fonction γ est effectivement de classe C^1 parce que ses composantes sont des produits de $r(\theta)$ et d'une fonction trigonométrique.

(2) En suivant la formule (21.29) du même théorème 21.9, nous avons $l(\gamma) = \int_{\theta_1}^{\theta_2} \|\gamma'(t)\|dt$. Il reste simplement à calculer la norme :

$$\gamma'(t) = \big(r'(t)\cos(t) - r(t)\sin(t), r'(t)\sin(t) + r(t)\cos(t)\big). \tag{91.96}$$

Nous trouvons alors

$$\|\gamma'(t)\| = \sqrt{r(t)^2 + r'(t)^2}, \tag{91.97}$$

et par conséquent la formule annoncée.

(3) En utilisant la formule, nous devons calculer l'intégrale

$$\begin{aligned} l &= \int_0^\pi \sqrt{\cos^2(\theta) + 4\cos^2(\theta)\sin^2(\theta)}d\theta \\ &= 2\int_0^{\pi/2} \cos(\theta)\sqrt{1 + 4\sin^2(\theta)}d\theta \\ &= 2\int_0^1 \sqrt{1 + 3u^2}du \\ &= 2\Big[\frac{1}{2}\sqrt{(4x^2 + 1)}x + \frac{1}{4}\operatorname{arcsinh}(2x)\Big]_0^1 \end{aligned} \tag{91.98}$$

où nous avons utilisé le changement de variable $u = \sin(\theta)$. La dernière intégrale se calcule en utilisant la formule (18.89). Le résultat est

$$l = \sqrt{5} + \frac{\operatorname{arcsinh}(2)}{2}. \tag{91.99}$$

(4) En utilisant la même formule (91.95), nous tombons sur l'intégrale

$$l = \sqrt{2}\int_{-\pi}^{\pi} \sqrt{1 - \cos(\theta)}d\theta. \tag{91.100}$$

Nous utilisons la formule $1 - \cos(\theta) = 2\sin^2\frac{\theta}{2}$ et nous effectuons le changement de variables $u = \theta/2$:

$$l = 4\int_{-\pi/2}^{\pi/2} |\sin(u)|du. \tag{91.101}$$

Attention à ne pas oublier la valeur absolue : sans elle l'intégrale est nulle, ce qui serait absurde pour une longueur. Au final nous avons $l = 8$.

Exercise 222 exoCourbesSurfaces0008

On considère le cercle $\mathcal{C}_R$ de rayon R centré à l'origine qui a l'équation paramétrique $f(t) = (R\cos t, R\sin t)$, où $t \in [0, 2\pi]$.

(1) Déterminer l'abscisse curviligne du point $f(t)$ par rapport au point $A = (R, 0)$. Calculer la longueur de $\mathcal{C}_R$.

(2) Déterminer le vecteur tangent unitaire $\tau(t)$, le vecteur unitaire normal $\nu(t)$ en $f(t)$, la courbure $c(t)$ et le rayon de courbure $R(t)$ de $\mathcal{C}_R$ en $f(t)$.

(3) Le vecteur $\nu(t)$ pointe vers l'intérieur ou vers l'extérieur du disque de rayon R ? Que se passe-t-il lorsqu'on remplace f par le paramétrage équivalente

$$g(t) = \big(R\cos(-t), R\sin(-t)\big), \tag{91.102}$$

où $t \in [0, 2\pi]$?

corrCourbesSurfaces0008

Correction of the exercise 222

(1) Nous reprenons les indications des équations (21.132) et (21.133) pour construire les coordonnées normales. Vu que $\gamma(0) = A$, nous pouvons faire commencer l'intégrale qui définit ϕ en zéro :

$$\phi(t) = \int_0^t \|f'(u)\|du = \int_0^t Rdu = tR. \tag{91.103}$$

Le point $f(t)$ a donc pour abscisse curviligne le nombre Rt. Nous pouvons aller plus loin en écrivant complètement le paramétrage. D'abord, $\phi^{-1}(s) = \frac{s}{R}$, par conséquent

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s) = \gamma(\frac{s}{R}) = \Big(R\cos(\frac{s}{R}), R(\sin\frac{s}{R})\Big). \tag{91.104}$$

Ensuite il faut savoir la longueur de la courbe pour savoir sur quel intervalle varie le paramètre s. Cette longueur est donné par $l(f)$ qui se calcule par

$$l(f) = \int_0^{2\pi} \|f'(t)\|dt = R\int_0^{\pi} 1dt = 2\pi R. \tag{91.105}$$

À partir de maintenant, nous passons de façon simple des coordonnées f (c'est-à-dire les radians) aux coordonnées normales par l'égalité $f(t) = \gamma_N(Rt)$. Cela est une égalité dans $\mathbb{R}^2$.

(2) En ce qui concerne le vecteur tangent, c'est le corolaire 21.51 (et même en fait le 21.52) qui nous indique la marche à suivre. Notez que lorsqu'on demande quelque chose au point $f(t)$, nous pouvons faire nos calculs avec les coordonnées normales en considérant le point $\gamma_N(Rt)$ (qui est le même). Quoi qu'il en soit, nous allons utiliser les coordonnées f et calculer

$$\tau(t) = \frac{f'(t)}{\|f'(t)\|} = \big(-\sin(t), \cos(t)\big). \tag{91.106}$$

Pour le reste, c'est la définition 21.59 qui joue. Nous devons donc travailler en coordonnées normales. Nous posons donc $s = Rt$ et nous calculons tous les éléments au point $\gamma_N(s)$. Nous avons

$$\gamma'_N(s) = \Big(-\sin(\frac{s}{R}), \cos(\frac{s}{R})\Big) \tag{91.107a}$$

$$\gamma''_N(s) = \frac{1}{R}\Big(-\cos(\frac{s}{R}), -\sin(\frac{s}{R})\Big). \tag{91.107b}$$

Donc

$$\nu(s) = -\Big(\cos(\frac{s}{R}), \sin(\frac{s}{R})\Big). \tag{91.108}$$

En terme des coordonnées $f(t)$, nous avons $\nu(t) = -\big(\cos(t), \sin(t)\big)$.
La courbure du cercle en $\gamma_N(s)$ vaut $c(s) = \|\gamma''_N(s)\| = \frac{1}{R}$, et son rayon de courbure vaut alors R ...ce n'est peut-être pas pour rien que cela se nomme « rayon de courbure ».

(3) Le vecteur $\nu(s)$ pointe vers l'intérieur du cercle. Il est même exactement l'opposé du rayon partant du centre.

Si nous prenons le paramétrage g, il ne se passe rien pour la direction de la courbure. En effet, le vecteur tangent change de signe, mais lorsque nous prenons la dérivée seconde, le signe moins sort deux fois et par conséquent le vecteur normal ν continue à pointer vers l'intérieur.

Exercise 223 exoGeomAnal-0019

(1) Donner la formule qui permet de calculer la longueur de la courbe γ paramétrée pas la fonction $r : [3, 5] \to \mathbb{R}^3$, $r : t \mapsto (x(t), y(t), z(t))$.

(2) Soit $\mathcal{D}$ le cube de sommets $A = (0, 0, 0)$, $B = (0, 1, 0)$, $C = (1, 1, 0)$, $D = (1, 0, 0)$, $E = (0, 0, 1)$, $F = (0, 1, 1)$, $G = (1, 1, 1)$ et $H = (1, 0, 1)$.

(2a) Dessiner $\mathcal{D}$ dans l'espace.

(2b) Paramétrer les trois courbes suivantes d'extrêmes A et G :

(2b.ii) γ_1 est l'union des deux morceaux de droite entre A et B et entre B et G ;

(2b.ii) γ_2 est l'union des trois morceaux de droite entre A et B, entre B et F et entre F et G ;

(2b.ii) γ_3 est l'union des deux morceaux de droite entre A et $P = (0, 1, 1/2)$ et entre P et G.

(2c) Calculer la longueur de la courbe qui vous parait la plus courte (motiver votre choix). Vous pouvez donner la valeur de l'intégrale sans calculer des intégrales si vous n'avez pas pu trouver le bon paramétrage au point précédent, mais il faudra la motiver à l'aide d'un dessin pour qu'elle soit prise en compte.

corrGeomAnal-0019

Correction of the exercise 223

<+CorrGeomAnal-0019+>

Exercise 224 exoGeomAnal-0038

Trouver le paramétrage normal du cercle de rayon R centré dans l'origine.

Un paramétrage de cette courbe est $(R\cos\theta, R\sin\theta)$.

corrGeomAnal-0038

Correction of the exercise 224

<+GeomAnal-0038+>

91.3 Limite et continuité

Exercise 225 ⊙⊙ exoLimiteContinue0001

On considère la fonction $f(x,y) = \frac{xy}{x^2+y^2}$ pour $(x,y) \neq (0,0)$. Montrer que f est constante sur les droites passant par $(0,0)$. La fonction f a-t-elle une limite en $(0,0)$?

corrLimiteContinue0001

Correction of the exercise 225

En coordonnées paramétriques, une droite passant par l'origine est de la forme $\gamma(t) = (at, bt)$ pour des constantes a et b. Sur cette droite, la fonction vaut

$$f\big(\gamma(t)\big) = f(at, bt) = \frac{abt^2}{a^2t^2 + b^2t^2} = \frac{ab}{a^2 + b^2}. \tag{91.109}$$

Nous voyons donc que le long de la droite de paramètres a, b, la fonction est constante et vaut $ab/(a^2 + b^2)$.

Note : pour obtenir ce résultat sans utiliser d'équations paramétrique, on pouvait poser $y = px + m$, et traiter le cas particulier de la droite verticale à part.

Le corolaire 12.225 montre alors que la fonction n'a pas de limite pour $(x,y) \to (0,0)$ parce que nous avons trouvé toute une série de chemins le long desquels les limites sont différentes. Par exemple pour $a = 0$ et $b = 1$, nous avons $\lim_{t\to 0} f(\gamma(t)) = 0$ et avec $a = b = 1$, nous avons $\lim_{t\to 0}(f \circ \gamma)(t) = \frac{1}{2}$.

Exercise 226 ⊖⊖ exoLimiteContinue0002

Calculer les limites suivantes, ou indiquer si elles n'existent pas.

(1) $\lim_{(x,y)\to(0,0)} \frac{x}{x^2+y^2}$;

(2) $\lim_{(x,y)\to(0,0)} \frac{(x+2y)^3}{x^2+y^2}$.

corrLimiteContinue0002

Correction of the exercise 226

(1) Si nous considérons le chemin $\gamma(t) = (t,t)$, nous avons $(f \circ \gamma)(t) = \frac{t}{2t^2} = \frac{1}{2t}$, et la limite $t \to 0$ n'existe pas. En vertu du corolaire 12.226, la limite de f en $(0,0)$ n'existe pas.

Ceci est un cas typique d'un fait qui arrive souvent : lorsqu'on a une fraction de polynômes dont le dénominateur est de plus haut degré que le numérateur, alors il n'y a pas de limites (parce que le dénominateur tend plus vite vers zéro que le numérateur). Attention toutefois à ne pas en faire une règle générale.

(2) Nous voyons que le degré du numérateur est plus haut que celui du dénominateur. Cela nous incite à croire que la limite va exister et valoir 0. En effet, en passant aux coordonnées polaires,

$$f(r\cos\theta, r\sin\theta) = r(\cos\theta + 2\sin\theta)^3 < r3^3 = 27r. \tag{91.110}$$

La limite de cela lorsque $r \to 0$ est vaut 0. La limite de la fonction est donc zéro.

Notez qu'il est important de borner $f(r\cos\theta, r\sin\theta)$ par un nombre qui ne dépend pas de θ avant de tester la limite $r \to 0$. Ici, nous avons borné $\cos\theta + 2\sin\theta$ par 3, et $27r$ est le s_r de la proposition 18.242.

Exercise 227 ⊙⊙ exoLimiteContinue0003
Soit

$$\begin{aligned} f\colon \mathbb{R}^2\backslash\{(0,0)\} &\to \mathbb{R} \\ (x,y) &\mapsto \frac{x^2y^2}{x^2y^2+(x-y)^2}. \end{aligned} \tag{91.111}$$

Montrer que

$$\lim_{y\to 0}\lim_{x\to 0} f(x,y) = \lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0, \tag{91.112}$$

mais que $\lim_{(x,y)\to(0,0)} f(x,y)$ n'existe pas. corrLimiteContinue0003

Correction of the exercise 227
Il faut d'abord bien comprendre ce que signifie la notation

$$\lim_{y\to 0}\lim_{x\to 0} f(x,y). \tag{91.113}$$

Lorsque nous écrivons $\lim_{x\to 0} f(x,y)$, nous considérons $f(x,y)$ comme fonction de la seule variable x ; la variable y est alors traitée comme un paramètre. Pour considérer y comme un paramètre, il y a deux cas à traiter : $y=0$ et $y\neq 0$. Dans le cas $y=0$, la fonction f se réduit à la fonction constante 0. Nous devons donc considérer la fonction

$$g(x) = 0 \tag{91.114}$$

dont le domaine est $x\neq 0$. Évidement, la limite de g lorsque $x\to 0$ est zéro. Dans le ces $y\neq 0$ nous considérons la fonction

$$g(x) = \frac{x^2y^2}{x^2y^2+(x-y)^2}, \tag{91.115}$$

dont le domaine est $\mathbb{R}$. Nous trouvons que $\lim_{x\to 0} g(x) = 0$.

Tout cela pour dire que $\lim_{x\to 0} f(x,y) = 0$. Par conséquent, $\lim_{y\to 0}\lim_{x\to 0} f(x,y) = 0$.

De la même façon, nous avons $\lim_{y\to 0} f(x,y) = 0$.

Étudions maintenant la limite simultanée $\lim_{(x,y)\to(0,0)} f(x,y)$. Une bonne astuce pour regarder plusieurs chemins d'un coup est de considérer le chemin $\gamma(t) = (t,kt)$. Cela teste toute les droites (sauf une, laquelle ?) en un seul calcul. Nous avons

$$f(t,kt) = \frac{k^2t^4}{k^2t^4+t^2(1-k)^2} = \frac{k^2t^2}{k^2t^2+(1-k)^2}. \tag{91.116}$$

La limite de cela lorsque $t\to 0$ vaut $\frac{0}{(1-k)^2} = 0$ pour tout k sauf pour $k=1$. Le chemin qui correspond à $k=1$ apparaît donc spécial. Regardons le de plus près. Nous avons

$$f(t,t) = \frac{t^4}{t^4} = 1, \tag{91.117}$$

et donc la limite le long de ce chemin est 1. Nous avons donc trouvé deux chemins de limites différentes, et nous concluons que $\lim_{(x,y)\to(0,0)} f(x,y)$ n'existe pas.

Exercise 228 ⊙⊙ exoLimiteContinue0004
Soit

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{R} \\ f(x,y) &= \begin{cases} (x+y)\sin\frac{1}{x}\sin\frac{1}{y} & \text{si } xy\neq 0 \\ 0 & \text{sinon} \end{cases} \end{aligned} \tag{91.118}$$

Montrer que les deux limites itérées

$$\lim_{y\to 0}\lim_{x\to 0} f(x,y) \tag{91.119a}$$

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y) \tag{91.119b}$$

n'existent pas, mais que $\lim_{(x,y)\to(0,0)} f(x,y)$ existe et est égale à 0. corrLimiteContinue0004

Correction of the exercise 228

Pour calculer $\lim_{x\to 0} f(x,y)$, nous coupons la fonction en deux termes :

$$f(x,y)=x\sin\frac{1}{x}\sin\frac{1}{y}+y\sin\frac{1}{x}\sin\frac{1}{y}. \tag{91.120}$$ EqLC4fffssusxusy

Ne pas oublier que y est une simple constante différente de zéro ; en particulier $\sin\frac{1}{y}$ et $y\sin\frac{1}{y}$ sont deux simples constantes.

Le premier terme du membre de droite de (91.120) a une limite pour $x\to 0$. En effet,

$$0\leqslant|x\sin\frac{1}{x}|<|x| \tag{91.121}$$

parce que $\sin\frac{1}{x}<1$. Par conséquent, $\lim_{x\to 0} x\sin\frac{1}{x}=0$. Le second terme de (91.120) n'a par contre pas de limite lorsque $x\to 0$ parce que $\sin\frac{1}{x}$ oscille sans fin entre -1 et 1.

Nous en déduisons que $\lim_{x\to 0} f(x,y)$ n'existe pas.

En ce qui concerne la limite simultanée,

$$0\leqslant|f(x,y)|=|x+y||\sin\frac{1}{x}||\sin\frac{1}{y}|\leqslant|x+y|\to 0. \tag{91.122}$$

La fonction $|f(x,y)|$ est donc coincée entre 0 et la fonction $|x+y|$. Évidement, chacune de ces deux fonctions tend vers zéro lorsque $(x,y)\to(0,0)$.

Notez que pour montrer que les limites itérées n'existaient pas, ce qui a joué est que nous avions deux termes dont un avait une limite et l'autre pas. Si nous avions deux termes dont aucun n'avait une limite, nous n'aurions pas pu conclure.

Exercise 229 ⊙⊙ exoLimiteContinue0005

Étudier l'existence des limites suivantes :

ex_limcont_0005_uno

(1) $\lim_{(x,y)\to(0,0)}\frac{x^2y}{x+y}$;

(2) $\lim_{(x,y,z)\to(0,0,0)}\frac{xyz+z^3}{2x^3+yz^2}$;

(3) $\lim_{(x,y)\to(0,0)}\frac{|x|+|y|}{x^2+y^2}$;

(4) $\lim_{(x,y)\to(0,0)}\frac{x^4y}{x^2-y^2}$;

(5) $\lim_{(x,y,z)\to(0,0,0)}\frac{xy+yz}{x^2+2y^2+3z^2}$.

corrLimiteContinue0005

Correction of the exercise 229

(1) Nous essayons les chemins (t,kt). Ce que nous trouvons est

$$f(t,kt)=\frac{kt^2}{1+k}\to 0 \tag{91.123}$$ EqLC5ftkt

si $k\neq -1$. Mais si $k=-1$, nous ne sommes pas dans le domaine de f. Le calcul (91.123) ne nous a donc pas apporté grand chose. Il nous a cependant appris que lorsque $k\to -1$, nous avons un zéro qui arrive au dénominateur. Nous essayons donc un chemin « plus dynamique », où au lieu de considérer un k fixé, nous prenons $k(t)$ de telle façon que le dénominateur tende vers zéro plus vite que le numérateur, c'est-à-dire par exemple $1-k=t^3$. Cela donne le calcul

$$f\big(t,(t^3-1)t\big)=\frac{t^3-1}{t}. \tag{91.124}$$

Et nous voyons que la fonction f n'admet pas de limites le long de ce chemin parce que $\lim_{t\to 0}(t^3-1)/t$ n'existe pas.

(2) Nous essayons les chemins (t, at, bt). Nous trouvons

$$f(t, at, bt) = \frac{ab + b^3}{2 + ab^2}, \tag{91.125}$$

si $2 + ab^2 \neq 0$. On voit que les limites faites sur des chemins différents sont différentes. Par exemple, si $a = b = 1$ on a $\lim f(t, t, t) = 2/3$ et si $b = 0$ on a $f(t, at, 0) = 0$. La fonction f n'a donc pas une limite en l'origine.

(3) Nous passons aux coordonnées polaires

$$\lim_{(x,y)\to(0,0)} \frac{|x| + |y|}{x^2 + y^2} = \lim_{r\to 0} \frac{r(|\cos\theta| + |\sin\theta|)}{r^2} \leqslant \lim_{r\to 0} \frac{2}{r}. \tag{91.126}$$

Donc si nous considérions $+\infty$ comme une limite valide alors la limite de f en l'origine serait $+\infty$. En fait, dans la proposition 8.257 à propos de limite, il est explicitement dit que la limite doit être un élément de $\mathbb{R}$ et cela n'est pas le cas de $+\infty$. Pas conséquent nous devons conclure que la limite de f n'existe pas (ou mieux : que elle n'existe pas dans $\mathbb{R}$).

(4) Ce point peut être résolu de la même manière que le point (1). La fonction f est définie dehors des lignes $x = y$ et $x = -y$. Nous essayons d'abord la limite sur les droites (t, kt), pour $k \neq \pm 1$, et nous trouvons que si une limite existe elle sera zéro. Ensuite nous trouvons une courbe passant par l'origine sur laquelle le dénominateur décroît plus vite que le numérateur, par exemple $(t, t(1 + t^4))$ et voyons que la limite de f n'existe pas.

(5) Ce point peut être fait encore une fois par la méthode des chemins. Cependant, nous l'utilisons comme une excuse pour présenter les coordonnées ellipsoïdales. L'équation décrivant une ellipsoïde E centrée en l'origine et dont les axes sont parallèle aux axes coordonnés prend la forme

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \tag{91.127}$$ Ellissoide

où a, b et c sont les longueurs des demi-axes. Vous pouvez vérifier que tout point de la forme

$$\begin{aligned} x &= a\cos\theta\sin\phi \\ y &= b\sin\theta\sin\phi \\ z &= c\cos\phi, \end{aligned} \tag{91.128}$$

avec θ dans $[0, 2\pi[$ et ϕ dans $[0, \pi]$, est dans E. En fait E coïncide avec l'ensemble des points de cette forme. Multiplions maintenant les deux côtés de (91.127) par un scalaire $\rho > 0$. Cela revient à dilater (si $\rho > 1$) ou contracter (si $\rho < 1$) E tout en gardant ses proportions. La réunion de toutes les contractions et les dilatations possibles de E est l'espace $\mathbb{R}^3$. Avec cette justification intuitive, nous introduisons les coordonnées suivantes, dont les coordonnées sphériques sont un cas particulier (pour quels a, b et c retrouvons-nous les coordonnées sphériques ?)

$$\begin{array}{rcl} \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} & \to & \mathbb{R}^3 \\ (\rho, \theta, \phi) & \mapsto & (x = a\rho\cos\theta\sin\phi, y = b\rho\sin\theta\sin\phi, z = c\rho\cos\phi). \end{array} \tag{91.129}$$

ce changement de variables est un difféomorphisme de la bande $\mathbb{R}^+ \times [0, 2\pi[\times[0, \pi]$ dans $\mathbb{R}^3\backslash\{(0, 0)\}$.

En ce qui concerne notre limite, si nous passons aux coordonnées ellipsoïdales avec $a = 1$, $b = 1/\sqrt{2}$ et $c = 1/\sqrt{3}$ nous trouvons

$$\lim_{(x,y,z)\to(0,0,0)} \frac{xy + yz}{x^2 + 2y^2 + 3z^2} = \lim_{\rho\to 0} \frac{1}{\sqrt{2}}\cos\theta\sin\theta\sin^2\phi + \frac{1}{\sqrt{6}}\sin\theta\sin\phi\cos\phi. \tag{91.130}$$

La quantité à droite ne dépend même pas de ρ et prend des valeurs différentes sur les différents droites par l'origine, la limite n'existe donc pas.

Exercise 230 ⊖⊖ exoLimiteContinue0006

Étudier les limites suivantes :

(1) $\lim_{(x,y)\to(1,0)} \frac{\ln(x+e^y)}{\sqrt{x^2+y^2}}$;

(2) $\lim_{(x,y)\to(0,0)} \frac{x^3y}{x^4+y^4}$;

(3) $\odot\odot \lim_{(x,y)\to(0,0)} \frac{(x^2+y^2)^2}{x^2-y^2}$;

(4) $\lim_{(x,y)\to(0,0)} \frac{1-\cos xy}{y^2}$;

(5) $\lim_{(x,y)\to(0,0)} \frac{\sin x}{\cos y - \cosh x}$.

corrLimiteContinue0006

Correction of the exercise 230

(1) La fonction à considérer est le rapport de deux fonctions continues $f_1 = \ln(x+e^y)$ et $f_2 = \sqrt{x^2+y^2}$. Les deux fonctions sont bornées en $(1,0)$, et f_2 ne s'annule pas au voisinage de $(1,0)$. La limite donc existe et est le rapport entre les deux limites : $\ln 2$.

(2) Passez aux coordonnés polaires. La limite n'existe pas.

(3) En passant aux coordonnés polaires, nous avons

$$\frac{(x^2+y^2)^2}{x^2-y^2} = \frac{r^4}{r^2\big(\cos^2(\theta)-\sin^2(\theta)\big)} = \frac{r^2}{\cos^2(\theta)-\sin^2(\theta)} = \frac{r^2}{\cos(2\theta)}. \tag{91.131}$$

Certes, *pour chaque* θ, cette fraction tend vers zéro lorsque r tend vers zéro. Mais faire la limite avec θ constant revient à faire la limite le long d'une droite. Cependant, si en même temps que faire $r \to 0$, nous prenons un angle qui tend vers 45 degrés, nous faisons tendre $\cos(2\theta)$ vers zéro. Cela reviendrait à suivre une spirale.

Nous voyons que la fonction n'a pas de limite en suivant le chemin $r(t) = t$ et $\theta(t)$ tel que $\cos\big(2\theta(t)\big) = t^3$. Est-ce qu'un tel chemin existe ? Oui parce que si nous prenons $\theta(t) = \frac{\arccos(t^3)}{2}$, nous avons bien $\cos\big(2\theta(t)\big) = t^3$. Ce chemin est tracé à la figure 91.13.

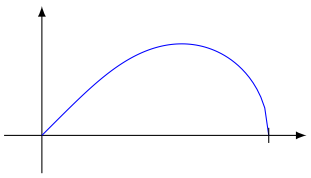


FIGURE 91.13: Un chemin possible le long duquel calculer une limite. $r(t) = t$, $\theta(t) = \arccos(t^3)/2$. LabelFigSpiraleLimite

Pour rappel, nous sommes intéressés par $t \to 0$, donc le fait que $\arccos(t^3)$ n'existe pas pour $|t| > 1$ ne nous embête pas.

(4) Dans ce point et le suivant il faut utiliser les développements asymptotiques (de Taylor) pour le calcul des limites.

La fonction $\cos(xy)$ admet le développement suivant :

$$\cos xy = 1 - (xy)^2 + o((xy)^3).$$

Nous écrivons alors la limite sous la forme

$$\lim_{(x,y)\to(0,0)} \frac{1-\cos(xy)}{y^2} = \lim_{(x,y)\to(0,0)} \frac{(xy)^2}{y^2} = \lim_{(x,y)\to(0,0)} x^2 = 0. \tag{91.132}$$

(5) Il existe des fonctions $a \in o(x^2)$, $b \in o(y^4)$ et $c \in o(x^4)$ telles que

$$\sin(x) = x + a(x) \tag{91.133a}$$
$$\cos(y) = 1 - y^2 + b(y) \tag{91.133b}$$
$$\cosh(x) = 1 + x^2 + c(x). \tag{91.133c}$$

Pourquoi développer le numérateur à l'ordre 1 et le dénominateur à l'ordre 2 ? Parce que le dénominateur est une différence. Nous développons jusqu'à l'ordre qu'il faut pour avoir

quelque chose de non nul. En l'occurrence dans $\cos(y)-\cosh(x)$, à l'ordre zéro nous avons $1-1=0$ tandis qu'à l'ordre 2 nous avons $-(x^2+y^2)$ qui est non nul. Nous écrivons donc la fonction donnée sous la forme

$$f(x,y)=-\frac{x+a(x)}{x^2+y^2+b(y)+c(x)}. \tag{91.134}$$

À ce moment, nous comprenons que la limite ne va pas exister parce que le degré du dénominateur est plus grand que celui du numérateur. Pour le montrer rigoureusement, prenons le chemin $\gamma(t)=(t,t)$:

$$f(t,t)=-\frac{t+a(t)}{2t^2+b(t)+c(t)}=-\frac{1+\frac{a(t)}{t}}{2t+\frac{b(t)}{t}+\frac{c(t)}{t}}. \tag{91.135}$$

Lorsque nous faisons la limite $t\to 0$, nous avons, par construction,

$$\lim_{t\to 0}\frac{a(t)}{t}=\lim_{t\to 0}\frac{b(t)}{t}=\lim_{t\to 0}\frac{c(t)}{t}=0, \tag{91.136a}$$

et par conséquent la limite $\lim_{t\to 0} f(t,t)=\frac{1}{0}$ n'existe pas.

Exercise 231 ⊖⊖ exoLimiteContinue0007

Soit

$$\begin{aligned} f&\colon \mathbb{R}^2\to\mathbb{R}\\ f(x,y)&=\begin{cases}\frac{x\ln(1+y)}{e^{xy}-1} & \text{si } x\neq 0, y\neq 0\\ 1 & \text{si } x=0 \text{ ou } y=0\end{cases}\end{aligned} \tag{91.137}$$

Calculer $\lim_{(x,y)\to(0,0)} f(x,y)$.

corrLimiteContinue0007

Correction of the exercise 231

Encore une fois on peut calculer la limite à l'aide des développements asymptotiques.

$$\lim_{(x,y)\to(0,0)}\frac{x\ln(1+y)}{e^{xy}-1}=\lim_{(x,y)\to(0,0)}\frac{xy}{1+xy-1}=1. \tag{91.138}$$

Exercise 232 ⊙⊙ exoLimiteContinue0008

Soit

$$\begin{aligned} f&\colon \mathbb{R}^2\to\mathbb{R}\\ f(x,y)&=\begin{cases}\frac{x^2y}{x^4+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0)\end{cases}\end{aligned} \tag{91.139}$$

(1) Montrer que $\lim_{x\to 0}(\lim_{y\to 0} f(x,y))=\lim_{y\to 0}(\lim_{x\to 0} f(x,y))=0$.

(2) Montrer que pour tout $(v_1,v_2)\in\mathbb{R}^2$ on a $\lim_{t\to 0} f(tv_1,tv_2)=0$.

(3) Montrer que f n'est pas continue en $(0,0)$.

corrLimiteContinue0008

Correction of the exercise 232

(1) Comme pour les exercices 227 et 228

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{x^2y}{x^4+y^2}\right)=\lim_{x\to 0}\frac{0}{x^4}=0, \tag{91.140}$$

de même pour l'autre limite.

(2) Nous calculons la limite de la fonction sur chaque droite passant par l'origine. Soit (v_1, v_2) un vecteur fixé, alors

$$\lim_{t\to 0} f(tv_1, tv_2) = \lim_{t\to 0} \frac{v_1^2 v_2 t^3}{t^2(v_1^4 t^2 + v_2^2)} = \lim_{t\to 0} t \underbrace{\frac{v_1^2 v_2}{v_1^4 t^2 + v_2^2}}_{A}, \tag{91.141}$$

cette limite existe et sa valeur est zéro, parce que A est une fonction bornée de t.

(3) Il est suffisant de remarquer que la limite pour $t \to 0$ de f restreinte à la parabole (t, t^2) n'est pas zéro. En fait

$$\lim_{t\to 0} f(t, t^2) = \lim_{t\to 0} \frac{t^4}{t^4 + t^4} = \frac{1}{2}. \tag{91.142}$$

Exercise 233 ⊖⊖ exoLimiteContinue0009

Étudier la continuité des fonctions suivantes :

(1) $f(x,y) = \begin{cases} \frac{x^4 y}{x^4+y^6} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases}$

(2) $f(x,y) = \begin{cases} \frac{xy^4}{x^4+y^6} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases}$

(3) $f(x,y) = \begin{cases} y^2 \sin\frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon.} \end{cases}$

(4) $f(x,y) = \begin{cases} xe^{\arctan\frac{y}{x}} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$

corrLimiteContinue0009

Correction of the exercise 233

Pour rappel, une fonction est continue en (a,b) si $\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$. Donc si, comme c'est souvent le cas dans cet exercice, $f(0,0) = 0$, nous devons vérifier si la limite de $f(x,y)$ lorsque (x,y) tend vers $(0,0)$ est égale à zéro ou non. Si oui, la fonction est continue, si non, la fonction n'est pas continue.

(1) Cette fonction est continue en l'origine. Pour le voir nous passons aux coordonnées polaires

$$\begin{aligned} \lim_{(x,y)\to(0,0)} \frac{x^4 y}{x^4 + y^6} &= \lim_{r\to 0} \frac{r^5 \cos^4\theta \sin\theta}{r^4(\cos^4\theta + r^2 \sin^6\theta} = \\ &= \lim_{r\to 0} \frac{r^5 \cos^4\theta \sin\theta}{r^4(\cos^4\theta + r^2 \sin^6\theta} = \\ &= \lim_{r\to 0} r \underbrace{\frac{\sin\theta}{1 + r^2 \tan^4\theta \sin^2\theta}}_{A}. \end{aligned} \tag{91.143}$$

Pour faire le dernier passage nous avons supposé $\cos\theta \neq 0$, cela ne pose pas de problèmes parce que nous pouvons considérer l'axe vertical séparément et il est facile de voir que $\lim_{y\to 0} f(0,y) = 0$. La fonction A est borné dans l'intervalle $[-1,1]$, donc la limite de f existe et vaut zéro.

(2) Cette fonction n'est pas continue en l'origine. Il suffit de considérer la limite le long de la courbe $(t^{3/2}, t)$

$$\lim_{t\to 0} \frac{t^{4+3/2}}{2t^6} \tag{91.144}$$

n'est clairement pas zéro.

(3) Cette fonction doit être testée en tous les points de la forme $(a,b) = (a,0)$ parce que c'est en ces points que l'on a un « problème », à savoir un zéro dans un dénominateur.

La fonction f est le produit de y^2 et d'une fonction bornée. On a alors

$$0 \leqslant |y^2| \left|\sin\frac{x}{y}\right| \leqslant |y|^2. \tag{91.145}$$

Quand y tend vers zéro, $|f|$ est coincée entre deux fonctions qui tendent vers zéro, elle ; ne peut pas s'échapper et donc doit tendre vers zéro. Ce théorème est dit «des deux gendarmes» («dei due carabinieri» en italien) ou «de l'étau».

(4) Passons aux coordonnées polaires : $f(r\cos\theta, r\sin\theta) = r\cos\theta e^{\theta}$. La fonction f est le produit de x et d'une fonction qui prends ses valeurs entre $-e^{2\pi}$ et $e^{2\pi}$ et est donc continue partout.

Exercise 234 ⊛⊛ exoLimiteContinue0010

On définit la fonction f sur $\mathbb{R}^2\backslash\{(x,x) \text{ tel que } x \in \mathbb{R}\}$ par

$$f(x,y) = \frac{\sin x - \sin y}{x-y}. \tag{91.146}$$

Peut-on prolonger f en une fonction continue sur $\mathbb{R}^2$? corrLimiteContinue0010

Correction of the exercise 234

Le but de l'exercice est de déterminer la limite $\lim_{(x,y)\to(a,a)} f(x,y)$ pour tout a. Le théorème des accroissements finis en une dimension nous enseigne que pour chaque x et y, il existe un nombre $\xi(x,y)$ tel que

$$\sin(y) = \sin(x) + \cos\big(\xi(x,y)\big)(y-x). \tag{91.147}$$

Cela est le théorème 12.192. Donc la fonction f peut être écrite sous la forme

$$f(x,y) = \frac{-\cos\big(\xi(x,y)\big)(y-x)}{x-y} = \cos\big(\xi(x,y)\big). \tag{91.148}$$

Ceci est une composée de fonctions. Il nous faut trouver la limite $\lim_{(x,y)\to(a,a)} \xi(x,y)$. Nous savons que pour tout x et y, le nombre $\xi(x,y)$ est strictement compris entre x et y, donc

$$\big|\xi(x,y) - a\big| \leqslant \max\big\{|x-a|, |y-a|\big\}. \tag{91.149}$$

Par conséquent, $\lim_{(x,y)\to(a,a)} \xi(x,y) = a$, et

$$\lim_{(x,y)\to(a,a)} f(x,y) = \cos(a). \tag{91.150}$$

Nous pouvons donc prolonger f par continuité en posant

$$f(x,y) = \begin{cases} \dfrac{\sin(x)-\sin(y)}{x-y} & \text{si } x \neq y \\ \cos(x) & \text{si } x = y \end{cases} \tag{91.151}$$

Exercise 235 ⊙⊙ exoLimiteContinue0011

Étudier la continuité des applications suivantes :

(1) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} \dfrac{\sin(xy)}{\sqrt{x^2+y^2}} & \text{si} \quad (x,y) \neq (0,0), \\ 0 & \text{si} \quad (x,y) = (0,0). \end{cases}$

(2) $g : \mathbb{R}^3 \to \mathbb{R}$, $g(x,y,z) = \begin{cases} \dfrac{4xyz}{x^2+y^2+z^2} & \text{si} \quad (x,y,z) \neq (0,0,0), \\ 0 & \text{si} \quad (x,y,z) = (0,0,0)\ . \end{cases}$

corrLimiteContinue0011

Correction of the exercise 235

(1) Lorsqu'on a le sinus d'une fonction qui tend vers zéro, il est souvent bon de multiplier et diviser par la fonction afin de faire apparaitre la limite connue $\lim_{x\to 0}\frac{\sin(x)}{x}=1$. Cela permet de simplifier les expressions. Ici, on a

$$\sin(xy)=\frac{xy\sin(xy)}{xy}. \tag{91.152}$$

Le calcul de la limite s'effectue donc de la façon suivante :

$$\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{\sqrt{x^2+y^2}}=\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{xy}\frac{xy}{\sqrt{x^2+y^2}}=\lim_{(x,y)\to(0,0)}\frac{xy}{\sqrt{x^2+y^2}}=0. \tag{91.153}$$

Le fait que la limite finale soit zéro peut être vu en passant aux coordonnées polaires. La fonction f est donc continue à l'origine.

(2) La fonction est continue, pour le voir il suffit de passer aux coordonnées sphériques.

Exercise 236 exoExamDec2011-0002

Soient $f\colon\mathbb{R}\to\mathbb{R}$ une fonction continue et $a\in\mathbb{R}$ tels que $f(a)>0$.

(1) Montrer que l'ensemble $\{x\in\mathbb{R}\text{ tel que }f(x)>0\}$ est ouvert.

(2) Soit A un ouvert contenant a. Montrer que

$$\int_A f(x)dx>0 \tag{91.154}$$

corrExamDec2011-0002

Correction of the exercise 236

<+CorrExamDec2011-0002+>

Exercise 237 exoGeomAnal-0035

Soit $f:\mathbb{R}^2\to\mathbb{R}$ la fonction

$$f(x,y)=\begin{cases}\frac{y\sin^2(2x)}{x} & \text{si } x\neq 0\\ 0 & \text{sinon.}\end{cases} \tag{91.155}$$

(1) Montrer que f est une fonction continue sur $\mathbb{R}^2$.

(2) Calculer les dérivées partielles de f au point $(\pi/6,4)$.

(3) Calculer les dérivées partielles de f au point $(0,2)$.

(4) Les dérivées partielles sont-elles continues au point $(0,2)$?

(5) La fonction f est-elle différentiable au point $(0,2)$?

corrGeomAnal-0035

Correction of the exercise 237

<+GeomAnal-0035+>

91.4 Calcul différentiel

91.4.1 Dérivées partielles

Exercise 238 exoCalculDifferentiel0001

Étudier la continuité de la fonction $f:\mathbb{R}^2\to\mathbb{R}$, l'existence et la continuité des dérivées partielles premières dans les cas suivants :

(1) $\ominus\ominus f(x,y)=\begin{cases}(x^2+y^2)\sin\left(\frac{1}{x^2+y^2}\right) & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases}$

(2) $\odot\odot f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon.} \end{cases}$

(3) $\ominus\ominus f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon.} \end{cases}$

(4) $\odot\odot f(x,y) = \begin{cases} \frac{x\sin y - y\sin x}{x^2+y^2} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon.} \end{cases}$

(5) $\ominus\ominus f(x,y) = \begin{cases} \frac{\sin(xy)}{|x|+|y|} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon.} \end{cases}$

(6) $\ominus\ominus f(x,y) = \begin{cases} e^{1/(x^2+y^2-1)} & \text{si } x^2+y^2<1 \\ 0 & \text{sinon.} \end{cases}$

(7) $\odot\odot f(x,y) = \begin{cases} y^2\sin\frac{x}{y} & \text{si } y\neq 0 \\ 0 & \text{sinon.} \end{cases}$

(8) $\odot\odot f(x,y) = \begin{cases} (x^2+y^2)^x & \text{si } (x,y)\neq(0,0) \\ 1 & \text{sinon.} \end{cases}$

(9) $\ominus\ominus f(x,y) = \sin(|xy|)$

(10) $\ominus\ominus f(x,y) = \begin{cases} 1-x^2-y^2 & \text{si } x^2+y^2\leqslant 1 \\ 0 & \text{sinon.} \end{cases}$

(11) $\odot\odot f(x,y) = \begin{cases} x^2 & \text{si } |x|>y \\ y^2 & \text{sinon.} \end{cases}$

corrCalculDifferentiel0001

Correction of the exercise 238

(1) L'apparition de la combinaison x^2+y^2 est un signal impérieux pour passer aux coordonnées polaires :

$$f(\rho\cos\theta,\rho\sin\theta)=\rho^2\sin\frac{1}{\rho^2}. \tag{91.156}$$

Cela est une fonction qui ne dépend pas de θ et donc le fait que $\lim_{\rho\to 0} f=0$ implique que $\lim_{(x,y)\to(0,0)} f(x,y)=0$. La fonction est donc continue parce que $f(0,0)=0$.
En ce qui concerne les dérivées partielles,

$$\frac{\partial f}{\partial x}(x,y)=2x\sin\left(\frac{1}{x^2+y^2}\right)-2x\cos\left(\frac{1}{x^2+y^2}\right)\frac{1}{x^2+y^2}. \tag{91.157}$$

EqCDundsdfx

En regardant par exemple la fonction (91.157) le long du chemin $\gamma(t)=(t,t)$, nous voyons que la limite $\lim_{(x,y)\to(0,0)}\partial_x f(x,y)$ n'existe pas. La dérivée partielle n'est donc pas continue.

(2) En ce qui concerne la continuité, le passage en polaire est conseillé parce que nous avons une fraction de polynômes :

$$f(r\cos\theta,r\sin\theta)=r(\cos^3\theta-\sin^3\theta). \tag{91.158}$$

Avec les notations de la proposition 18.242, nous avons donc

$$E_r=\{\rho(\cos^3\theta-\sin^3\theta) \text{ tels que } \rho\in[0,r],\theta\in[0,2\pi]\}. \tag{91.159}$$

Vu que la fonction $\cos^3\theta-\sin^3\theta$ est une fonction bornée de θ par exemple par 2, nous avons

$$s_r<2r. \tag{91.160}$$

Par conséquent $\lim_{r\to 0} s_r = 0$, ce qui prouve que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. La fonction est donc continue.

En ce qui concerne la dérivabilité,

$$\partial_x f(x,y) = \frac{x^4 + 3x^2y^2 + 2xy^3}{(x^2+y^2)^2}. \tag{91.161}$$

Cette fonction n'a pas de limite lorsque $(x,y) \to (0,0)$ parce que par exemple pour les chemins $\gamma(t) = (t, kt)$ nous avons

$$\partial_x f(t,kt) = \frac{2k^3 + 3k^2 + 1}{k^4 + 2k^2 + 1}, \tag{91.162}$$

qui dépend clairement de k et donc du chemin choisit. La dérivée par rapport à x n'est donc pas continue en $(0,0)$. Cependant, cette dérivée partielle existe :

$$\partial_x f(0,0) = \lim_{x\to 0} \frac{\frac{x^3}{x^2} - 0}{x} = 1. \tag{91.163}$$ EqcddzuijDifcv

De la même façon, $\partial_y f(0,0) = -1$.

(3) La fonction f est continue en $(0,0)$ parce que en polaires elle s'écrit

$$f(\rho,\theta) = \rho\cos\theta|\sin\theta|. \tag{91.164}$$

Cette fonction est toujours en norme plus petite que ρ, et par conséquent nous avons $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. En ce qui concerne la dérivée, nous avons

$$\partial_x f(x,y) = \frac{|y|}{\sqrt{x^2+y^2}}\left(1 - \frac{x^2}{x^2+y^2}\right). \tag{91.165}$$

Sur la droite (t,kt) nous avons

$$\begin{aligned}\frac{|t||k|}{|t|\sqrt{1+k^2}}\left(1 - \frac{t^2}{t^2(1+k^2)}\right) &= \frac{|k|}{\sqrt{1+k^2}}\left(1 - \frac{1}{1+k^2}\right)\\ &= \frac{|k|}{\sqrt{1+k^2}}\left(\frac{1+k^2-1}{1+k^2}\right)\\ &= \frac{|k|k^2}{(1+k^2)^{3/2}}.\end{aligned} \tag{91.166}$$

Cela ne dépend pas de t. La limite n'existe donc pas. La dérivée selon x n'est donc pas continue en $(0,0)$.

La dérivée dans la direction de x au point $(0,0)$ existe et vaut

$$\partial_x f(0,0) = \lim_{x\to 0} \frac{\frac{0}{\sqrt{x}} - 0}{x} = 0. \tag{91.167}$$ EqCCDZUpxf

En ce qui concerne la dérivée par rapport à y aux point autres que $(0,0)$, nous avons

$$\partial_y f(x,y) = \frac{|y|}{\sqrt{x^2+y^2}} - \frac{|y|x^2}{(x^2+y^2)^{3/2}}. \tag{91.168}$$

La dérivée $\partial_y f(0,0)$ se calcule en utilisant le même type de calcul que celui donné dans l'équation (91.167).

(4) Cette fonction qui mélange des fonctions trigonométriques et des polynômes est la candidate parfaite pour utiliser un développement. La théorie du polynôme de Taylor nous enseigne qu'il existe une fonction $a \in o(x^5)$ (c'est-à-dire $\lim_{x\to 0} \frac{a(x)}{x^5} = 0$) telle que

$$\begin{aligned}\sin(x) &= x - \frac{x^3}{3!} + a(x)\\ \sin(y) &= y - \frac{y^3}{3!} + a(y).\end{aligned} \tag{91.169}$$

En substituant dans la définition de f, nous avons

$$\begin{aligned}f(x,y) &= \frac{xy - \frac{xy^3}{3} + xa(y) - yx + \frac{yx^3}{3} - ya(x)}{x^2+y^2}\\ &= \frac{xy}{6}\frac{x^2-y^2}{x^2+y^2} + \frac{xa(y) - ya(x)}{x^2+y^2}.\end{aligned} \tag{91.170}$$

La limite de la première partie est zéro, comme expliqué dans l'exemple 18.244.

Pour la seconde partie, nous avons

$$0 \leqslant \left|\frac{xa(y)}{x^2+y^2}\right| \leqslant \left|\frac{xa(y)}{y^2}\right|. \tag{91.171}$$

Cela peut être vu comme le produit de la fonction x par la fonction $a(y)/y^2$, dont les limites valent zéro. Nous avons donc montré que

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0, \tag{91.172}$$

et par conséquent que la fonction est continue en $(0,0)$.

Passons aux dérivées partielles. En ce qui concerne la dérivée $\partial_x f$, nous avons

$$\frac{\partial f}{\partial x}(x,y) = \frac{\sin(y) - y\cos(x)}{x^2+y^2} + \frac{2x\big(x\sin(y) - y\sin(x)\big)}{(x^2+y^2)^2}. \tag{91.173}$$

Encore une fois nous remplaçons les fonctions trigonométriques par leur développements. Pour rappel, en ce qui concerne le cosinus, il existe une fonction b telle que

$$\cos(x) = 1 - \frac{x^2}{2} + b(x) \tag{91.174}$$

avec $\lim_{x\to 0} \frac{b(x)}{x^4} = 0$. En remplaçant les différentes expressions pour $\cos(x)$, $\sin(x)$ et $\sin(y)$, nous trouvons

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) = &\frac{y}{x^2+y^2}\left(\frac{x^2}{2} - \frac{y^2}{6}\right) + \frac{a(y) - yb(x)}{x^2+y^2}\\ &+ 2x\frac{xy}{6}\frac{y^2-x^2}{(x^2+y^2)^2} + 2x\frac{xa(y) - ya(x)}{(x^2+y^2)}.\end{aligned} \tag{91.175}$$

Nous pouvons montrer terme à terme que le tout tend vers 0 lorsque $(x,y) \to (0,0)$.

En passant aux polaires, le premier terme devient $\frac{1}{6}r\sin(\theta)\big(3\cos^2(\theta) - \sin^2(\theta)\big)$. Cela est r multiplié par une fonction bornée de θ. La limite est donc 0. L'autre terme polynomial se traite de la même manière. En ce qui concerne le terme

$$\frac{a(y) - yb(x)}{x^2+y^2}, \tag{91.176}$$

nous avons

$$0 \leqslant \left|\frac{yb(x)}{x^2+y^2}\right| < \left|\frac{yb(x)}{x^2}\right|. \tag{91.177}$$

La limite de ce qui se trouve à droite est zéro par la propriété de $b(x)$. Au final nous avons prouvé que

$$\lim_{(x,y)\to(0,0)} \partial_x f(x,y) = 0. \tag{91.178}$$

Attention : cela ne prouve pas que $\partial_x f$ soit une fonction continue. Il faut encore vérifier si $\partial_x f(0,0)$ est nul ou non. Cela se fait en utilisant directement la définition :

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t}. \tag{91.179}$$

Étant donné que $f(t,0) - f(0,0) = 0$ pour tout t, la limite est nulle, ce qui prouve que $\partial_x f(0,0) = 0$ et par conséquent que $\partial_x f$ est continue en $(0,0)$.

Vu que dans la fonction x et y arrivent de façon symétrique, la même chose sera vrai pour $\partial_y f$. Les deux dérivées partielles étant continues en $(0,0)$, la fonction est donc différentiable en $(0,0)$ par la proposition 12.295.

(5) Nous avons

$$\left|\frac{\sin(xy)}{|x|+|y|}\right| \leqslant \left|\frac{\sin(xy)}{|x|}\right| = |y| \left|\frac{\sin(xy)}{xy}\right| \tag{91.180}$$

où nous avons multiplié et divisé par $|y|$ pour faire apparaitre la combinaison $\sin(xy)/xy$ dont nous savons que la limite est 1. Nous avons donc $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

En ce qui concerne les dérivées partielles, nous rappelons que la dérivée de $|x|$ vaut $\operatorname{sgn}(x)$, c'est-à-dire 1 si x est positif et -1 si x est négatif. Nous avons

$$\partial_x f(x,y) = \frac{y\cos(xy)}{|x|+|y|} + \frac{\operatorname{sgn}(x)\sin(xy)}{(|x|+|y|)^2}. \tag{91.181}$$ EqCD15diffc

Nous utilisons encore une fois les développements. La majorité des termes tendent vers zéro ; par exemple,

$$\frac{y\cos(xy)}{|x|+|y|} = \frac{1}{|x|+|y|}\left(y - \frac{y^3x^2}{2} + b(xy)\right). \tag{91.182}$$

En passant aux polaires, le terme central disparaît, et le terme en $b(xy)$ disparaît parce que, proche de $(0,0)$, nous avons évidement $|xy| < |x|$ et par conséquent $|b(xy)| < b(x)$. En faisant le même genre de jeux avec l'autre terme de (91.181), nous devons calculer

$$\lim_{(x,y)\to(0,0)} \frac{y}{|x|+|y|} + \operatorname{sgn}(x)\frac{xy}{(|x|+|y|)^2}. \tag{91.183}$$

Si on passe en coordonnées polaires, on voit que cette limite dépend de θ, et donc n'existe pas. La dérivée partielle $\partial_x f$ n'est en particulier pas continue en $(0,0)$. Le nombre $\partial_x f(0,0)$ existe pourtant :

$$\partial_x f(0,0) = \lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x} = \lim_{x\to 0} \frac{\frac{\sin(0)}{|x|+0} - 0}{0} = 0. \tag{91.184}$$

(6) Pour cette fonction la zone à problème n'est pas le point $(0,0)$. Les points à étudier sont les points sur le cercle $x^2 + y^2 = 1$.

Heureusement, l'exercice est assez vite fait parce que la fonction donnée est à symétrie radiale. En coordonnées polaires, un point du cercle est de la forme $(1,\theta)$.

Un bon exemple de chemin, en polaires, est $\big((1-t),\theta\big)$. Note : le chemin choisit doit aller vers le point à tester lorsque $t \to 0$. Ici, le chemin $\big((1-t),\theta\big)$ tend vers le point $(1,\theta)$ qui est bien un point sur le cercle, c'est-à-dire un point à tester. Dans les coordonnées polaires, nous avons $x^2 + y^2 = r$, et le long de notre chemin nous avons $r = 1 - t$. Nous avons donc

$$f(r,\theta) = \begin{cases} e^{1/(r^2-1)} & \text{si } r < 1 \\ 0 & \text{si } r \geqslant 1. \end{cases} \tag{91.185}$$

Nous avons

$$\lim_{r\to 1^-} e^{1/(r^2-1)} = 0, \tag{91.186}$$

eqSYoHJK

par conséquent la fonction est continue. Insistons sur le fait que la limite (91.186) est calculée avec $r < 1$. Si nous avions calculé $r \to 1$, la limite n'aurait pas existé.

En ce qui concerne la dérivée par rapport à x, nous devons considérer trois zones bien distinctes du plan. La première est constituée des points intérieurs au cercle $r = 1$, la seconde est le cercle lui-même et la troisième est l'extérieur du cercle.

Commençons par considérer un point (x, y) du cercle $r = 1$ et calculons

$$\frac{\partial f}{\partial x}(x, y) = \lim_{\epsilon\to 0} \frac{f(x + \epsilon, y) - f(x, y)}{\epsilon}. \tag{91.187}$$

Nous supposons $x > 0$ (le même calcul sera vrai avec $x < 0$). Alors $f(x + \epsilon, y) = 0$ tant que $\epsilon > 0$; nous devons donc calculer

$$\frac{\partial f}{\partial x}(x, y) = \lim_{\epsilon\to 0^-} e^{1/((x+\epsilon)^2+y^2-1)} = 0. \tag{91.188}$$

La dérivée partielle par rapport à x vaut donc zéro sur le pourtour du cercle.

Si (x, y) est à l'extérieur du cercle, nous avons bien évidemment $\partial_x f(x, y) = 0$.

Pour (x, y) à l'intérieur du cercle nous utilisons les règles de dérivation pour obtenir

$$\frac{\partial f}{\partial x}(x, y) = 2x \frac{e^{1/(x^2+y^2-1)}}{(x^2 + y^2 - 1)}. \tag{91.189}$$

La continuité de $\partial_x f$ consiste à voir si cette expression tend vers zéro lorsque (x, y) tend vers un point du cercle.

En passant aux polaires, la limite en suivant la droite du rayon du cercle revient à étudier la limite

$$\lim_{r\to 1^-} \frac{e^{1/(r^2-1)}}{r^2 - 1}. \tag{91.190}$$

En passant au logarithme,

$$\ln\left(\frac{e^{1/(r^2-1)}}{r^2 - 1}\right) = \frac{1}{r^2 - 1} - \ln(r^2 - 1). \tag{91.191}$$

Une façon de voir cette limite est de considérer le chemin $r(\epsilon) = \sqrt{1 - \epsilon}$ (avec $\epsilon > 0$) et de regarder la limite

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} - \ln(\epsilon). \tag{91.192}$$

En mettant au même dénominateur et en appliquant la règle de l'Hospital,

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} - \ln(\epsilon) = \lim_{\epsilon\to 0^+} \frac{1 - \epsilon \ln \epsilon}{\epsilon} = \lim_{\epsilon\to 0} -\ln(\epsilon) - 1 = +\infty. \tag{91.193}$$

Nous avons donc vu que en calculant la limite $(x, y) \to (1, \theta_0)$ [3] le long du le chemin

$$\gamma(t) = \big(r(t), \theta(t)\big) = (\sqrt{1 - t}, \theta_0), \tag{91.194}$$

nous avons

$$\lim_{t\to 0} \frac{\partial f}{\partial x}\big(\gamma(t)\big) = +\infty. \tag{91.195}$$

La dérivée partielle dans la direction de x n'est donc pas continue.

3. Ici (x, y) est le point qui bouge, donné en cartésiennes et $(1, \theta_0)$ avec θ_0 fixé est un point du cercle donné en polaires.

(7) La fonction $y^2 \sin\left(\frac{x}{y}\right)$ a des problèmes sur tous les points du type $(a,0)$. Pour voir la continuité, nous devons donc calculer

$$\lim_{(x,y)\to(a,0)} y^2 \sin\left(\frac{x}{y}\right). \tag{91.196}$$

Nous traitons cela de façon classique :

$$0 \leqslant |f(x,y)| \leqslant |y^2| \to 0. \tag{91.197}$$

Par la règle de l'étau nous avons la continuité de la fonction.
Il s'agit maintenant d'étudier la continuité des dérivées partielles. La dérivée partielle par rapport à x vaut

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} y\cos\frac{x}{y} & \text{si } y\neq 0 \\ 0 & \text{sinon.} \end{cases} \tag{91.198}$$

Pour la seconde ligne, nous avons utilisé le calcul direct

$$\frac{\partial f}{\partial x}(a,0) = \lim_{t\to 0} \frac{f(a+t,0)-f(a,0)}{t} = 0. \tag{91.199}$$

Par le même raisonnement que précédemment, $\lim_{(x,y)\to(a,0)} y\cos\frac{x}{y} = 0$, donc la dérivée $\partial_x f$ est continue aux points $(a,0)$.
Pour la dérivée partielle par rapport à y nous avons

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} 2y\sin\left(\frac{x}{y}\right) - x\cos\left(\frac{x}{y}\right) & \text{si } y\neq 0 \\ 0 & \text{sinon.} \end{cases} \tag{91.200}$$

Encore une fois, la seconde ligne provient d'un calcul utilisant directement la définition :

$$\frac{\partial f}{\partial y}(a,0) = \lim_{t\to 0} \frac{f(a,t)-f(a,0)}{t} = \lim_{t\to 0} t\sin\left(\frac{a}{t}\right) = 0. \tag{91.201}$$

Afin de voir la continuité de $\partial_y f$, il faut simplement calculer la limite $\lim_{(x,y)\to(a,0)} \frac{\partial f}{\partial y}(x,y)$ et voir si c'est égal à 0 :

$$\lim_{(x,y)\to(a,0)} 2y\sin\frac{x}{y} - x\cos\frac{x}{y}. \tag{91.202}$$

EqcCDusellcy

Le premier terme tend vers 0, tandis que le second est un produit d'une fonction qui tend vers a (la fonction x) par la fonction $\cos\frac{x}{y}$ qui n'a pas de limite. Si $a\neq 0$, la limite (91.202) n'existe pas. La fonction $\partial_y f$ est donc continue en $(0,0)$.
En résumé, la domaine de continuité de $\partial_y f$ est $\mathbb{R}^2\backslash\{y=0\}\cup\{(0,0)\}$.

(8) Lorsque nous devons calculer une limite d'une fonction exposant une autre fonction, le passage au logarithme est toujours une bonne idée :

$$(x^2+y^2)^x = e^{x\ln(x^2+y^2)}. \tag{91.203}$$

En passant aux polaires, il est vite vu que

$$\lim_{(x,y)\to(0,0)} x\ln(x^2+y^2) = 0, \tag{91.204}$$

et donc $\lim_{(x,y)\to(0,0)} f(x,y) = e^0 = 1$. La fonction est donc continue en $(0,0)$.
En ce qui concerne la dérivée par rapport à x par contre, elle n'existe même pas en $(0,0)$ parce que

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = \lim_{t\to 0}\frac{(t^2)^t-1}{t} = \lim_{t\to 0}\frac{e^{t\ln(t^2)}-1}{t}. \tag{91.205}$$

Cette limite se fait en utilisant la règle de l'Hospital. La dérivée du numérateur vaut

$$\big(\ln(t^2)+2\big)e^{t\ln(t^2)}, \tag{91.206}$$

dont la limite n'existe pas lorsque $t\to 0$.

(9) En tant que composée de fonctions continues, la fonction $f(x,y)=\sin(|xy|)$ est continue. Pour la dérivée, nous avons

$$\partial_x f(x,y)=\begin{cases} y\cos(xy) & \text{si } xy>0\\ -y\cos(xy) & \text{si } xy<0\\ ? & \text{dans les autres cas.}\end{cases} \tag{91.207}$$

Nous n'avons pas besoin de savoir ce qu'il se passe dans les autres cas (ce sont les points sur les axes). En effet, si nous essayons de voir la limite vers le point $(0,1)$ en suivant le chemin $(t,1)$, la limite n'existe pas parce que nous avons 1 lorsqu'on considère $t>0$ et -1 avec $t<0$.

Remarque : la fonction f peut être vue comme la composée de $(x,y)\mapsto xy$, de $t\mapsto|t|$ et de $u\mapsto\sin(u)$. La fonction valeur absolue n'étant pas dérivable en 0, on pouvait s'attendre à ce que la fonction f ne soit pas dérivable aux points où $xy=0$.

(10) Dans cet exercice, les points à étudier sont ceux du cercle $x^2+y^2=1$. La continuité se règle en coordonnées polaires parce que la fonction vaut simplement $f(\rho,\theta)=1-\rho^2$ dont la limite vaut bien 0 lorsque $\rho\to1$.

En ce qui concerne la dérivée par rapport à x nous avons

$$\partial_x f(x,y)=\begin{cases} -2x & \text{si } x^2+y^2<1\\ 0 & \text{si } x^2+y^2>1\\ ? & \text{dans les autres cas.}\end{cases} \tag{91.208}$$

Cette fonction n'est évidement pas continue sur les points du cercle tels que $x\neq0$.

(11) Les points de raccordement entre les deux zones sont les points $(-a,a)$ et (a,a) avec $a\geqslant0$. Évidement, $x^2=y^2$ en tous ces points, donc la fonction est continue. Pour les dérivées,

$$\partial_x f(x,y)=\begin{cases} 2x & \text{si } |x|>y\\ 0 & \text{sinon.}\end{cases} \tag{91.209}$$

Parmi les points du type $(-a,a)$ et (a,a), seul au point $(0,0)$ cette fonction est continue.

Exercise 239 exoCalculDifferentiel0002

Déterminer l'ensemble de définition puis calculer les dérivées partielles des fonctions définies par

(1) $f(x,y)=\arctan(xy)$
(2) $g(x,y)=\arctan\left(\frac{x}{y}\right)$
(3) $h(x,y,z)=e^{\frac{x}{y}}+e^{\frac{z}{y}}$
(4) $\varphi(x,y)=x^2\sin(y)$
(5) $\psi(x,y)=\sqrt{1-x^2-y^2}$
(6) $\theta(x,y)=\ln(x+y)$

Étudier la différentiabilité de ces fonctions. corrCalculDifferentiel0002

Correction of the exercise 239

En ce qui concerne les ensembles de définition, les règles les plus courantes sont

(1) $\ln(x)$ demande $x>0$ (stricte) ;
(2) $1/x$ demande $x\neq0$;
(3) $\sqrt{x}$ demande $x\geqslant0$ (non stricte).

C'est parti ...

(1) Il n'y a pas de conditions d'existence. Les dérivées partielles sont

$$\frac{\partial f}{\partial x} = \frac{y}{1+(xy)^2} \tag{91.210a}$$
$$\frac{\partial f}{\partial y} = \frac{x}{1+(xy)^2}. \tag{91.210b}$$

(2) L'arc tangente est bien définie pour tous les réels. La seule condition est donc $x \neq 0$. Les dérivées partielles sont

$$\frac{\partial f}{\partial x} = \frac{-y}{x^2+y^2} \tag{91.211a}$$
$$\frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}. \tag{91.211b}$$

Pour rappel, $(\arctan(x))' = \frac{1}{1+x^2}$.

(3) Nous avons les fractions x/y et z/y. Toutes deux demandent $y \neq 0$. C'est la seule condition. Le domaine est donc tout $\mathbb{R}^2$ moins la droite $y = 0$. Les dérivées partielles sont

$$\frac{\partial f}{\partial x} = \frac{1}{y}e^{x/y}, \tag{91.212a}$$
$$\frac{\partial f}{\partial y} = -\frac{x}{y^2}e^{x/y} - \frac{z}{y^2}e^{z/y}, \tag{91.212b}$$
$$\frac{\partial f}{\partial z} = \frac{1}{y}e^{z/y}. \tag{91.212c}$$

(4) Le domaine est $\mathbb{R}^2$. Les dérivées sont

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x\sin(y)\\ \frac{\partial f}{\partial y} &= x^2\cos(y).\end{aligned} \tag{91.213}$$

Cette fonction est différentiable en tant que produit de fonctions différentiables.

(5) Condition : $1-x^2-y^2 \geqslant 0$, c'est-à-dire $x^2+y^2 \leqslant 1$, c'est-à-dire la boule centrée en $(0,0)$ et de rayon 1, y compris le bord. Les dérivées sont

$$\frac{\partial f}{\partial x} = \frac{-x}{\sqrt{1-x^2-y^2}} \tag{91.214a}$$
$$\frac{\partial f}{\partial y} = \frac{-y}{\sqrt{1-x^2-y^2}}. \tag{91.214b}$$

Il faut vérifier la différentiabilité sur le bord du domaine, c'est-à-dire sur les points de la forme (a,b) avec $a^2+b^2=1$. Étant donné que les dérivées partielles ne sont pas continues en ces points, la fonction n'est pas différentiable sur le bord.

(6) Condition : $x+y>0$. C'est le demi-plan strictement au dessus de la droite $y=-x$. $\partial_x f(x,y) = 1/(x+y)$ et $\partial_y f(x,y) = 1/(x+y)$.
Le domaine de la fonction étant ouvert (ne contient pas la droite), ça n'a pas de sens de se demander si la fonction est différentiable sur les points de la droite $x=-y$. En chaque point du domaine, il existe un voisinage sur lequel les dérivées partielles sont continues, donc la fonction est différentiables (proposition 12.295).

91.4.2 Différentielles

Exercise 240 ⊙⊙ exoCalculDifferentiel0003
Soit L une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$.

(1) Montrer que L est continue.

(2) Montrer que L admet des dérivées partielles premières que l'on explicitera.

(3) Montrer que L est différentiable et expliciter sa différentielle.

corrCalculDifferentiel0003

Correction of the exercise 240

La première chose à faire est de savoir la forme générale d'une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$. Tout vecteur $x \in \mathbb{R}^n$ s'écrit sous la forme $x = \sum_{i=1}^n x_i e_i$ où $\{e_i\}$ est la base canonique de $\mathbb{R}^n$. En utilisant la linéarité de L,

$$L(x) = \sum_i L(x_i e_i) = \sum_i x_i L(e_i) = \sum_i a_i x_i \tag{91.215}$$

où nous avons nommé $a_i = L(e_i)$. Maintenant L est continue parce que c'est une somme de fonctions continues (dont les projections $x \mapsto x_i$). En ce qui concerne les dérivées partielles, nous calculons en utilisant $L(x) = \sum_i a_i x_i$ et en utilisant la linéarité de la dérivée :

$$\begin{aligned} \frac{\partial L}{\partial x_k}(x) &= \sum_i a_i \frac{\partial x_i}{\partial x_k} \\ &= \sum_i a_i \delta_{ik} \\ &= a_k \end{aligned} \tag{91.216}$$

parce que la dérivée de x_i par rapport à x_k vaut 1 si $i = k$ et vaut zéro si $i \neq k$.

Pour montrer qu'elle est différentiable, nous allons procéder en deux étapes. D'abord nous allons trouver un candidat pour la différentielle, et ensuite nous allons voir que ce candidat fonctionne. Le candidat est donné par les formules du lemme 12.255 :

$$dL(x).u = \sum_i \frac{\partial L}{\partial x_i}(x) u_i = \sum_i a_i u_i = L(u). \tag{91.217}$$

Cela nous dit que *si L est différentiable*, alors sa différentielle est L elle-même. Testons donc cela dans le critère (10.682) :

$$\lim_{h \to 0} \frac{\|L(a+h) - L(a) - T(h)\|}{\|h\|} \tag{91.218}$$

où T est notre candidat différentielle, c'est-à-dire $T(h) = L(h)$. En utilisant la linéarité de L, nous voyons que $L(a+h) = L(a) + L(h)$ et donc le numérateur est identiquement nul. Cela prouve que L est différentiable et que la différentielle est bien celle que nous avons devinée.

Exercise 241 exoCalculDifferentiel0004

On considère l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{x^3 y}{x^4 + y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \tag{91.219}$$

(1) Étudier la continuité de f au point $(x,y) = (0,0)$ (indication : remarquer que $2x^2 y \leqslant x^4 + y^2$).

(2) Montrer que f admet des dérivées partielles premières au point $(0,0)$ mais n'est pas différentiable en $(0,0)$.

corrCalculDifferentiel0004

Correction of the exercise 241

(1) La petite formule $2x^2 y \leqslant x^4 + y^2$ est un cas particulier de la formule $a^2 + b^2 \geqslant 2ab$ qui vient simplement du fait que $a^2 - 2ab + b^2 = (a-b)^2 \geqslant 0$. Tester la continuité de f en $(0,0)$ revient à calculer sa limite lorsque $(x,y) \to (0,0)$ et puis de regarder si la limite est égale à la valeur

de la fonction en $(0,0)$. Ici, $f(0,0)=0$, donc nous devons voir si $\lim_{(x,y)\to(0,0)} f(x,y)=0$ ou non. Étant donné que

$$\frac{2x^2y}{x^4+y^2}\leqslant 1, \tag{91.220}$$

nous avons toujours

$$0\leqslant |f(x,y)|\leqslant \left|\frac{x^3y}{x^4+y^2}\right|\leqslant \left|\frac{x}{2}\right|\to 0. \tag{91.221}$$

Le nombre $|f(x,y)|$ est par conséquent coincé entre 0 et une fonction qui tend vers zéro. La règle de l'étau conclut que la limite existe au vaut zéro.

(2) Les dérivées partielles en $(0,0)$ se calculent par la définition :

$$\partial_x f(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=0, \tag{91.222}$$

et

$$\partial_y f(0,0)=\lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}=0. \tag{91.223}$$

Donc les deux dérivées partielles existent et sont nulles. Le « candidat » différentiel est donc $T(u)=0$. Nous testons le critère de différentiabilité (10.682) avec $T=0$:

$$\begin{aligned}\lim_{(h_1,h_2)\to(0,0)}\frac{|f\big((0,0)+(h_1,h_2)\big)-f(0,0)-T(h_1,h_2)|}{\|(h_1,h_2)\|_{\mathbb{R}^2}}\\ =\lim_{(h_1,h_2)\to(0,0)}\left|\frac{h_1^3h_2}{(h_1^4+h_2^2)}\frac{1}{\sqrt{h_1^2+h_2^2}}\right|.\end{aligned} \tag{91.224}$$

En suivant le chemin (t,t), nous trouvons la limite zéro. Par contre en suivant le chemin (t,t^2), nous trouvons la limite $\frac{1}{2}$. La limite n'existe donc pas et la fonction n'est pas différentiable en $(0,0)$.

Exercise 242 exoCalculDifferentiel0021

Étudier la différentiabilité des fonctions suivantes. Elles proviennent de l'exercice 238.

(1) $$f(x,y)=\begin{cases}(x^2+y^2)\sin\left(\frac{1}{x^2+y^2}\right) & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases}$$

(2) $$f(x,y)=\begin{cases}\frac{x^3-y^3}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases}$$

corrCalculDifferentiel0021

Correction of the exercise 242

(1) Nous avons vu autour de l'équation (91.163) que $\partial_x f(0,0)=1$ et $\partial_y f(0,0)=-1$. Par conséquent la différentielle de f en $(0,0)$, si elle existe, vaut $T(x,y)=0$. Afin de voir si cela est bien la différentielle, nous calculons

$$\begin{aligned}\lim_{(x,y)\to(0,0)}\frac{\frac{x^3-y^3}{x^2-y^2}-(x-y)}{\sqrt{x^2+y^2}}&=\lim_{(x,y)\to(0,0)}\frac{-xy^2+yx^2}{(x^2+y^2)^{3/2}}\\ &=\lim_{r\to 0} r^3\frac{\cos\theta\sin\theta(\cos\theta-\sin\theta)}{r^3}.\end{aligned} \tag{91.225}$$

Cette dernière limite étant dépendante de θ, nous en déduisons que la limite qui définit la différentielle n'existe pas. La fonction n'est donc pas différentiable en $(0,0)$.

91.4.3 Dérivée d'applications composées

Exercise 243 exoCalculDifferentiel0006

On considère l'application $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$\varphi(x, y, z) = (x^2 - y^2, y^2 - z^2, z^2 - x^2) \tag{91.226}$$

et une application $f : \mathbb{R}^3 \to \mathbb{R}$ de classe C^1. Écrire, en termes des dérivées partielles de f,

(1) la matrice jacobienne de l'application $g = f \circ \varphi : \mathbb{R}^3 \to \mathbb{R}$;

(2) le nombre $\frac{\partial g}{\partial x}(t, t, t) + \frac{\partial g}{\partial y}(t, t, t) + \frac{\partial g}{\partial z}(t, t, t)$ pour tout réel t.

corrCalculDifferentiel0006

Correction of the exercise 243

Nommons u, v et w les variables de f. La matrice jacobienne de la fonction $g\colon \mathbb{R}^3 \to \mathbb{R}$ est la matrice colonne

$$\begin{pmatrix} \frac{\partial g}{\partial x}(x, y, z) \\ \frac{\partial g}{\partial y}(x, y, z) \\ \frac{\partial g}{\partial z}(x, y, z) \end{pmatrix}. \tag{91.227}$$

Les dérivées de g se calculent en termes de celles de f en utilisant la règle des fonctions composées :

$$\begin{aligned} \frac{\partial g}{\partial x}(x, y, z) &= \frac{\partial f}{\partial u}\big(\varphi(x, y, z)\big) \underbrace{\frac{\partial \varphi_u}{\partial x}(x, y, z)}_{=2x} \\ &\quad + \frac{\partial f}{\partial v}\big(\varphi(x, y, z)\big) \underbrace{\frac{\partial \varphi_v}{\partial x}(x, y, z)}_{=0} \\ &\quad + \frac{\partial f}{\partial w}\big(\varphi(x, y, z)\big) \underbrace{\frac{\partial \varphi_w}{\partial x}(x, y, z)}_{=-2x} \\ &= 2x\big(\partial_u f(\varphi) + \partial_w f(\varphi)\big) \end{aligned} \tag{91.228}$$

où nous avons noté, sur la dernière ligne, simplement φ au lieu de $\varphi(x, y, z)$ pour alléger la notation. De la même façon pour les autres nous trouvons

$$\begin{aligned} \partial_y g(x, y, z) &= 2y\big(-\partial_u f(\varphi) + \partial_v f(\varphi)\big) \\ \partial_z g(x, y, z) &= 2z\big(-\partial_v f(\varphi) + \partial_w f(\varphi)\big). \end{aligned} \tag{91.229}$$

Nous calculons maintenant la somme de ces trois expressions en posant $x = y = z = t$. Les termes partent deux à deux et il ne reste que zéro. **Exercise 244** exoCalculDifferentiel0007

Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe C^2. On pose $g(x, y) = f(x^2 - y^2, 2xy)$. Calculer Δg en fonction de Δf, où $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ est l'opérateur **laplacien**.

corrCalculDifferentiel0007

Correction of the exercise 244

Il s'agit d'un exercice de dérivation de fonction composée ; une des choses importantes est de bien noter à quel point sont calculées les dérivées partielles : certaines sont calculées en (x, y) et d'autre en $(x^2 - y^2, 2xy)$.

Écrivons

$$g(x, y) = (f \circ \varphi)(x, y) \tag{91.230}$$

avec $\varphi(x, y) = (x^2 - y^2, 2xy)$. Étant donné que nous allons en avoir beaucoup besoin, calculons directement les dérivées de φ :

$$\begin{aligned} \frac{\partial \varphi_1}{\partial x} &= 2x \qquad \frac{\partial \varphi_2}{\partial x} = 2y \\ \frac{\partial \varphi_1}{\partial y} &= -2y \qquad \frac{\partial \varphi_2}{\partial y} = 2x. \end{aligned} \tag{91.231}$$

En utilisant la formule donné au théorème 12.301, nous avons

$$\begin{aligned}\frac{\partial g}{\partial x}(x,y) &= \frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big)\frac{\partial \varphi_1}{\partial x}(x,y) + \frac{\partial f}{\partial x_2}\big(\varphi(x,y)\big)\frac{\partial \varphi_2}{\partial x}(x,y)\\ &= 2x\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big) + 2y\frac{\partial f}{\partial x_2}\big(\varphi(x,y)\big).\end{aligned} \tag{91.232}$$

Afin de trouver $\frac{\partial^2 g}{\partial x^2}(x,y)$, il s'agit maintenant de calculer la dérivée de $\frac{\partial g}{\partial x}(x,y)$ par rapport à x. Pour ce faire, nous voyons l'expression

$$\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big) \tag{91.233}$$

comme la fonction composée de $\frac{\partial f}{\partial x_1}$ et de φ. Par souci de simplification des notations, nous allons adopter les notations suivantes :

$$\begin{aligned}\partial_1 f &= \frac{\partial f}{\partial x_1} & \partial_2 f &= \frac{\partial f}{\partial x_2}\\ \partial_{11}^2 f &= \frac{\partial^2 f}{\partial x_1^2} & \partial_{22}^2 f &= \frac{\partial^2 f}{\partial x_2^2}\\ \partial_{12}^2 f &= \frac{\partial^2 f}{\partial x_1 \partial x_2} & \partial_{21}^2 f &= \frac{\partial^2 f}{\partial x_2 \partial x_1}.\end{aligned} \tag{91.234}$$

Étant donné que la fonction f est C^2, nous avons $\partial_{12}^2 f = \partial_{21}^2 f$ par le théorème 12.343. En utilisant en même temps la formule de dérivation de produit,

$$\begin{aligned}\frac{\partial^2 g}{\partial x}(x,y) &= 2x\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big)\right)\\ &\quad + 2\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big)\\ &\quad + 2y\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x_2}\big(\varphi(x,y)\big)\right)\\ &= 2x\left(\frac{\partial^2 f}{\partial x_1 \partial x_1}\big(\varphi(x,y)\big)\frac{\partial \varphi_1}{\partial x}(x,y) + \frac{\partial^2 f}{\partial x_2 \partial x_1}\big(\varphi(x,y)\big)\frac{\partial \varphi_2}{\partial x}(x,y)\right)\\ &\quad + 2\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big)\\ &\quad + 2y\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\big(\varphi(x,y)\big)\frac{\partial \varphi_1}{\partial x}(x,y) + \frac{\partial^2 f}{\partial x_2 \partial x_2}\big(\varphi(x,y)\big)\frac{\partial \varphi_2}{\partial x}(x,y)\right)\\ &= 4x^2\partial_{11}^2 f\big(\varphi(x,y)\big) + 8xy\partial_{12}^2 f\big(\varphi(x,y)\big) + 2\partial_1 f\big(\varphi(x,y)\big) + 4y^2\partial_{22}^2 f\big(\varphi(x,y)\big).\end{aligned} \tag{91.235}$$

En ce qui concerne le calcul de $\frac{\partial^2 g}{\partial y^2}(x,y)$, nous avons le même genre de jeu :

$$\frac{\partial g}{\partial y}(x,y) = -2y\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big) + 2x\frac{\partial f}{\partial x_2}\big(\varphi(x,y)\big). \tag{91.236}$$

La dérivée de cela par rapport à y donne

$$\begin{aligned}\frac{\partial^2 g}{\partial y^2}(x,y) &= -2\partial_1 f\big(\varphi(x,y)\big) - 2y\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x_1}\big(\varphi(x,y)\big)\right) + 2x\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x_2}\big(\varphi(x,y)\big)\right)\\ &= -2\partial_1 f\big(\varphi(x,y)\big) + 4y^2\partial_{11}^2 f\big(\varphi(x,y)\big) - 8xy\partial_{12}^2 f\big(\varphi(x,y)\big) + 4x^2\partial_{22}^2 f\big(\varphi(x,y)\big).\end{aligned} \tag{91.237}$$

En effectuant la somme, nous trouvons

$$\Delta g(x,y) = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 4(x^2+y^2)\Delta f(x^2-y^2, 2xy). \tag{91.238}$$

Notez que nous n'avons pas $\Delta g = 4(x^2+y^2)\Delta f$.

Exercise 245 exoCalculDifferentiel0008

Trouver toutes les fonctions f de $\mathbb{R}^2$ dans $\mathbb{R}$ telles que

(1) ⊙⊙$\frac{\partial f}{\partial x}(x,y) = -\sin x \sin y$ et $\frac{\partial f}{\partial y}(x,y) = \cos x \cos y$.

(2) ⊖⊖$\frac{\partial f}{\partial x}(x,y) = 2xy + y^3$ et $\frac{\partial f}{\partial y}(x,y) = x^2 + 3y^2x$. Comparer $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ et $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$.

(3) ⊙⊙Même question avec $\frac{\partial f}{\partial x}(x,y) = -y$ et $\frac{\partial f}{\partial y}(x,y) = x$.

(4) ⊖⊖Même question avec $\frac{\partial f}{\partial x}(x,y) = y + \cos x$ et $\frac{\partial f}{\partial y}(x,y) = x^2 + y^2$.

corrCalculDifferentiel0008

Correction of the exercise 245

Notez que dans tous les points de cet exercice, les dérivées données sont des fonctions C^∞. Donc les fonctions qu'on cherche sont également C^∞ ; elles vérifient en particulier le théorème 12.343. Une façon de s'assurer qu'un énoncé est impossible est de vérifier que $\partial^2_{xy} f \neq \partial^2_{yx} f$. Si les dérivées partielles croisées ne sont pas égales, alors il est impossible de trouver une fonction qui satisfait les contraintes.

(1) Le fait de vérifier $\partial_x f(x,y) = -\sin(x)\sin(y)$ implique que

$$f(x,y) = \int -\sin(x)\sin(y)dx = \sin(y)\cos(x) + c_1(y). \tag{91.239}$$

La subtilité à comprendre est que la constante d'intégration est une fonction qui ne dépend pas de x, mais qui peut dépendre de y ! Ce qui nous avons donc obtenu est qu'il existe une fonction $c_1(y)$ telle que

$$f(x,y) = \sin(y)\cos(x) + c_1(y). \tag{91.240}$$ CD81fc1

Pour voir ce que nous dit la seconde condition, nous pouvons calculer $\partial_y f$ en utilisant la forme (91.240) :

$$\frac{\partial f}{\partial y}(x,y) = \cos(y)\cos(x) + c_1'(y). \tag{91.241}$$

Si nous voulons que cela soit toujours égal à $\cos(y)\cos(x)$, il faut $c_1'(y) = 0$ et donc que c_1 soit une constante. Toutes les fonctions répondant à la question sont donc de la forme

$$f(x,y) = \sin(y)\cos(x) + C \tag{91.242}$$

pour une constante C.

(2) La condition sur $\partial_x f$ dit que

$$f(x,y) = x^2y + xy^3 + c(y) \tag{91.243}$$

pour une certaine fonction c. La condition sur $\partial_y f$ nous dit alors que $c'(y) = 0$ et donc que c est constante.

En dérivant $2xy + y^3$ par rapport à y, nous trouvons

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = 2x + 3y^2, \tag{91.244}$$

tandis qu'en dérivant $x^2 + 3xy^2$ par rapport à x nous trouvons

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y} = 2x + 3y^2, \tag{91.245}$$

qui est la même chose.

Nous avons donc, dans ce cas, $\partial^2_{yx} f = \partial^2_{xy} f$ par le thé.

(3) La condition $\partial_x f = -y$ impose $f(x,y) = -yx + c_1(y)$, tandis que la condition $\partial_y f(x,y) = x$ impose $f(x,y) = yx + c_2(x)$. Les deux sont incompatibles. Il n'y a donc pas de solutions à cet exercice.

À noter que $\partial^2_{yx} f(x,y) = -1$ tandis que $\partial^2_{xy} f(x,y) = 1$. Les deux dérivées secondes sont dont différentes.

(4) Encore une fois les deux conditions sont incompatibles parce que nous trouvons

$$\begin{aligned} f(x,y) &= xy + \sin(x) + c_1(y) \\ f(x,y) &= x^2y + \frac{y^3}{3} + c_2(x). \end{aligned} \tag{91.246}$$

Il est certes possibles d'arranger les fonctions $c_1(y)$ et $c_2(x)$ pour obtenir les termes $\sin(x)$ et $\frac{y^3}{3}$, mais il n'est cependant pas possible de les arranger pour obtenir les termes xy et x^2y. Ici encore, les dérivées secondes ne sont pas égales : $\partial^2_{yx} f = 1$ tandis que $\partial_{xy} f = 2x$.

Exercise 246 ⊛⊛ | exoCalculDifferentiel0009

Trouver toutes les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$, C^1 sur $\mathbb{R}^2$ telles que

$$\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial y}(x,y) = 0 \tag{91.247}$$

pour tout $(x,y) \in \mathbb{R}^2$. On pourra effectuer le changement de variables $x = u + v$, $y = u - v$. corCalculDifferentiel0009

Correction of the exercise 246

Nous allons montrer deux méthodes pour résoudre cet exercice.

Première méthode Commençons par écrire le changement de variable de façon très explicite et dans les deux sens :

$$x(u,v) = u + v \tag{91.248a}$$
$$y(u,v) = u - v \tag{91.248b}$$

et

$$u(x,y) = \frac{1}{2}(x+y) \tag{91.249a}$$
$$v(x,y) = \frac{1}{2}(x-y) \tag{91.249b}$$

Nous introduisons la fonction $\tilde{f}$ qui est la fonction f « dans les coordonnées (u,v) » :

$$\tilde{f}(u,v) = f\big(x(u,v), y(u,v)\big). \tag{91.250}$$

La fonction f s'exprime en termes de $\tilde{f}$ de la façon suivante :

$$f(x,y) = \tilde{f}\big(u(x,y), v(x,y)\big). \tag{91.251}$$

Nous pouvons donc écrire les dérivées partielles de f en termes de celles de $\tilde{f}$ en utilisant la règle de dérivation des fonctions composées :

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{\partial \tilde{f}}{\partial u}\big(u(x,y),v(x,y)\big)\frac{\partial u}{\partial x}(x,y) + \frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big)\frac{\partial v}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) &= \frac{\partial \tilde{f}}{\partial u}\big(u(x,y),v(x,y)\big)\frac{\partial u}{\partial y}(x,y) + \frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big)\frac{\partial v}{\partial y}(x,y). \end{aligned} \tag{91.252}$$

Les dérivées de u et v par rapport à x et y sont connues. Nous avons donc

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{1}{2}\frac{\partial \tilde{f}}{\partial u}\big(u(x,y),v(x,y)\big) + \frac{1}{2}\frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big) \\ \frac{\partial f}{\partial y}(x,y) &= \frac{1}{2}\frac{\partial \tilde{f}}{\partial u}\big(u(x,y),v(x,y)\big) - \frac{1}{2}\frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big). \end{aligned} \tag{91.253}$$

En substituant ces valeurs dans l'équation pour f, nous trouvons la condition suivante pour $\tilde{f}$:

$$\frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big) = 0, \tag{91.254}$$

c'est-à-dire que $\tilde{f}$ ne dépend pas de v. Il existe donc une fonction ψ telle que $\tilde{f}(u,v)=\psi(u)$. Cela donne, sur f, la forme suivante :

$$f(x,y)=\tilde{f}\big(u(x,y),v(x,y)\big)=\psi\big(u(x,y)\big)=\psi\big((x+y)/2\big). \tag{91.255}$$

La conclusion est que la fonction f ne peut être qu'une fonction de $x+y$, ou encore que les variables x et y ne peuvent apparaitre que sous la combinaison $x+y$.

Par exemple les fonctions suivantes sont bonnes :

$$f(x,y)=x+y \tag{91.256a}$$
$$f(x,y)=\frac{\cos(x+y)}{(x+y)^2}. \tag{91.256b}$$

Mais la fonction $f(x,y)=x^2+y^2$ n'est pas bonne.

Exercise 247 exoCalculDifferentiel0010

Trouver toutes les fonctions $f:\mathbb{R}^2\to\mathbb{R}$, C^1 sur $\mathbb{R}^2$ telles que

$$\frac{\partial f}{\partial x}(x,y)+2x\frac{\partial f}{\partial y}(x,y)=0 \tag{91.257}$$ eCDuzEqares

pour tout $(x,y)\in\mathbb{R}^2$. On pourra effectuer le changement de variables $x=u$, $y=v+u^2$. corrCalculDifferentiel0010

Correction of the exercise 247

Nous effectuons le changement de variable suivant :

$$\begin{aligned} x(u,v)&=u & u(x,y)&=x\\ y(u,v)&=v+u^2 & v(x,y)&=y-x^2. \end{aligned} \tag{91.258}$$

Nous allons avoir besoin des dérivées de ce changement de variable :

$$\begin{aligned} \frac{\partial u}{\partial x}&=1 & \frac{\partial v}{\partial x}&=-2x\\ \frac{\partial u}{\partial y}&=0 & \frac{\partial v}{\partial y}&=1. \end{aligned} \tag{91.259}$$

Maintenant nous considérons la fonction $\tilde{f}$ comme étant la fonction f « vue dans les variables u et v », c'est-à-dire

$$\tilde{f}(u,v)=f\big(x(u,v),y(u,v)\big). \tag{91.260}$$

Nous pouvons aussi voir f comme donnée en termes de $\tilde{f}$ par la formule inverse :

$$f(x,y)=\tilde{f}\big(u(x,y),v(x,y)\big). \tag{91.261}$$

Nous calculons les dérivées de f en termes de celles de $\tilde{f}$ en utilisant la formule de dérivation de fonctions composées :

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y)&=\frac{\partial\tilde{f}}{\partial u}(u,v)\underbrace{\frac{\partial u}{\partial x}(x,y)}_{=1}+\frac{\partial\tilde{f}}{\partial v}(u,v)\underbrace{\frac{\partial v}{\partial x}(x,y)}_{=-2x}\\ &=\partial_u\tilde{f}(u,v)-2x\partial_v\tilde{f}(u,v). \end{aligned} \tag{91.262}$$

De la même façon,

$$\begin{aligned} \frac{\partial f}{\partial y}(x,y)&=\frac{\partial\tilde{f}}{\partial u}(u,v)\frac{\partial u}{\partial y}(x,y)+\frac{\partial\tilde{f}}{\partial v}(u,v)\frac{\partial v}{\partial y}(x,y)\\ &=\partial_v\tilde{f}(u,v). \end{aligned} \tag{91.263}$$

Il est sous-entendu que lorsque nous écrivons $\partial_v \tilde{f}(u,v)$, nous entendons

$$\frac{\partial \tilde{f}}{\partial v}\big(u(x,y),v(x,y)\big), \tag{91.264}$$

en tant que fonction de x et y.

L'équation (91.257) devient donc $\partial_u \tilde{f} - 2x\partial_v \tilde{f} + 2x\partial_v \tilde{f} = 0$, c'est-à-dire simplement

$$\frac{\partial \tilde{f}}{\partial u}(u,v) = 0. \tag{91.265}$$

La résolution de cette équation est qu'il doit exister une fonction $\psi\colon \mathbb{R} \to \mathbb{R}$ telle que

$$\tilde{f}(u,v) = c + \psi(v). \tag{91.266}$$

La fonction f donnée en termes des variables x et y devient

$$f(x,y) = \tilde{f}\big(u(x,y),v(x,y)\big) = c + \psi\big(v(x,y)\big) = c + \psi(y - x^2). \tag{91.267}$$

Notez que l'on peut écrire plus simplement $f(x,y) = \psi(y-x^2)$ parce que la fonction ψ étant arbitraire, on peut la redéfinir pour inclure la constante.

Exercise 248 exoCalculDifferentiel0011

Trouver toutes les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1 sur $\mathbb{R}^2\backslash\{(0,0)\}$, continues sur $\mathbb{R}^2$ telles que

$$x\frac{\partial f}{\partial x}(x,y) + y\frac{\partial f}{\partial y}(x,y) = \sqrt{x^2+y^2} \tag{91.268}$$ EqeCDuuegares

pour tout $(x,y) \in \mathbb{R}^2\backslash\{(0,0)\}$. On pourra effectuer le changement de variables $x = r\cos\theta$, $y = r\sin\theta$ avec $r > 0$ et $\theta \in \mathbb{R}$. corrCalculDifferentiel0011

Correction of the exercise 248

Passons aux coordonnées polaires, c'est-à-dire posons

$$\begin{aligned} x &= r\cos(\theta) \quad & r &= (x^2+y^2)^{1/2} \\ y &= r\sin(\theta) \quad & \theta &= \arctan\left(\frac{y}{x}\right). \end{aligned} \tag{91.269}$$

Nous considérons la fonction $\tilde{f}$ définie par

$$\tilde{f}(r,\theta) = f\big(x(r,\theta),y(r,\theta)\big), \tag{91.270}$$ EqeCDchemuudevarffuxy

et ensuite nous voyons $\tilde{f}$ comme une fonction composée. Nous avons

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial \tilde{f}}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial \tilde{f}}{\partial \theta}\frac{\partial \theta}{\partial x} \\ &= \partial_r \tilde{f}\frac{x}{\sqrt{x^2+y^2}} + \partial_\theta \tilde{f}\frac{1}{1+\left(\frac{y}{x}\right)^2}\left(-\frac{y}{x^2}\right) \\ &= \partial_r \tilde{f}\frac{x}{\sqrt{x^2+y^2}} - \partial_\theta \tilde{f}\frac{y}{x^2+y^2}. \end{aligned} \tag{91.271}$$

En ce qui concerne la dérivée par rapport à y, nous faisons

$$\frac{\partial f}{\partial y} = \partial_r \tilde{f}\frac{y}{\sqrt{x^2+y^2}} + \partial_\theta \frac{x}{x^2+y^2}. \tag{91.272}$$

Nous pouvons maintenant chercher à résoudre l'équation (91.268). Nous l'écrivons sous la forme

$$1 = \frac{x}{\sqrt{x^2+y^2}}\frac{\partial f}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\frac{\partial f}{\partial y}. \tag{91.273}$$

Nous avons donc

$$1 = \left(\frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2} \right) \partial_r \tilde{f} + \left(\frac{-yx+xy}{(x^2+y^2)^{3/2}} \right) \partial_\theta \tilde{f} \tag{91.274}$$

L'équation se réduit donc à

$$\frac{\partial \tilde{f}}{\partial r} = 1, \tag{91.275}$$

EqeCDuueqsimp

et par conséquent

$$\tilde{f}(r,\theta) = r + c(\theta). \tag{91.276}$$

Pour obtenir cela, nous avons pris l'intégrale de (91.275) par rapport à r et nous avons considéré que la constante d'intégration pouvait dépendre de θ. Nous trouvons maintenant la forme générale de f en utilisant la définition (91.270) « à l'envers » :

$$f(x,y) = \tilde{f}\big(r(x,y),\theta(x,y)\big) = r(x,y) + c\big(\theta(x,y)\big) = \sqrt{x^2+y^2} + c\big(\arctan\left(\frac{y}{x}\right)\big). \tag{91.277}$$

Au final, ce que nous avons prouvé, c'est que pour toute fonction $c\colon \mathbb{R} \to \mathbb{R}$, la fonction

$$f(x,y) = \sqrt{x^2+y^2} + c\big(\arctan\left(\frac{y}{x}\right)\big) \tag{91.278}$$

est une solution de l'équation proposée. De plus, toutes les solutions s'écrivent de cette manière pour une certaine fonction c.

Étant donné que nous demandons des fonctions C^1, nous demandons que la fonction c soit C^1.

Exercise 249 ⊙⊙ exoCalculDifferentiel0012

Trouver toutes les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$, de classe C^2 sur $\mathbb{R}^2$ qui vérifient

(1) $\frac{\partial^2 f}{\partial x^2} = 0$

(2) $\frac{\partial^2 f}{\partial x \partial y} = 0$

(3) ⊖⊖ $\frac{\partial^2 f}{\partial x^2} = \cos(x+y)$

Pour chaque exercice, donner quelques exemples explicites. corrCalculDifferentiel0012

Correction of the exercise 249

(1) Cette équation décrit le mouvement d'une particule dans le plan dont nous savons qu'elle ne subit aucune accélération dans la direction x et dont nous n'avons aucune information sur le mouvement dans la direction y.

L'équation

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 0 \tag{91.279}$$

signifie que la fonction $\frac{\partial f}{\partial x}(x,y)$ ne dépend pas de x. Il existe donc une fonction ψ telle que

$$\frac{\partial f}{\partial x}(x,y) = \psi(y). \tag{91.280}$$

En intégrant les deux membres par rapport à x nous trouvons

$$f(x,y) = x\psi(y) + \varphi(y). \tag{91.281}$$

Dans cette formule, φ est la constante d'intégration. Vu que nous avons effectué une intégrale par rapport à x, cette constante d'intégration ne peut pas dépendre de x. Elle peut cependant dépendre de y.

Note : la fonction $\varphi(y)$ est n'importe quelle fonction de y, et peut en particulier avoir une partie constante. Les fonctions suivantes sont bonnes :

$$\begin{aligned} &f(x,y) = xy + 1 \\ &f(x,y) = \cos(y) \quad \text{c'est à dire } \psi(y) = 0 \text{ et } \varphi(y) = \cos(y) \end{aligned} \tag{91.282}$$

(2) Nous avons $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)=0$, et donc $\partial_y f=\psi(y)$ pour une certaine fonction. En intégrant par rapport à y nous trouvons

$$f(x,y)=\int\psi(y)dy+\varphi(x)+C \tag{91.283}$$

où $\varphi(x)+C$ est la constante d'intégration. Étant donné que $\int\psi(y)dy$ peut être n'importe quelle fonction (nous ne sommes intéressés qu'aux fonctions C^2), nous avons la forme générale

$$f(x,y)=\psi(y)+\varphi(x). \tag{91.284}$$

Les fonctions qui répondent à la question sont donc des fonctions qui ont une partie en x et une partie en y complètement séparées. En autres exemples, les fonctions suivantes fonctionnent :

$$\begin{aligned}f(x,y)&=\cos(y)+\sin(x)-4\\ g(x,y)&=\frac{x+y}{xy}.\end{aligned} \tag{91.285}$$

La seconde peut paraitre étonnante, mais ce n'est rien d'autre que $\frac{1}{x}+\frac{1}{y}$.

(3) Nous avons

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)=\cos(x+y), \tag{91.286}$$

et par conséquent,

$$\frac{\partial f}{\partial x}=\int\cos(x+y)dx=\sin(x+y)+\psi(y) \tag{91.287}$$

où ψ est la constante d'intégration par rapport à x. En intégrant encore,

$$f(x,y)=-\cos(x+y)+x\psi(y)+\varphi(y). \tag{91.288}$$

Exercise 250 ⊖⊖ exoCalculDifferentiel0013

Résoudre l'équation

$$2\frac{\partial^2 f}{\partial x^2}(x,y)-5\frac{\partial^2 f}{\partial x\partial y}(x,y)+2\frac{\partial^2 f}{\partial y^2}(x,y)=0 \tag{91.289}$$

avec $f:\mathbb{R}^2\to\mathbb{R}$ de classe C^2 en utilisant le changement de variables $u=x+2y$, $v=x+\frac{1}{2}y$. corrCalculDifferentiel0013

Correction of the exercise 250

Nous effectuons le changement de variables

$$\begin{aligned}u&=x+2y\\ v&=x+\frac{1}{2}y,\end{aligned} \tag{91.290}$$

avec les dérivées partielles

$$\begin{aligned}\frac{\partial u}{\partial x}&=1 & \frac{\partial u}{\partial y}&=2\\ \frac{\partial v}{\partial x}&=1 & \frac{\partial v}{\partial y}&=\frac{1}{2}.\end{aligned} \tag{91.291}$$

Nous posons

$$f(x,y)=\tilde{f}\big(u(x,y),v(x,y)\big), \tag{91.292}$$

et puis nous écrivons l'équation demandée en termes de $\tilde{f}$. Nous avons

$$\frac{\partial f}{\partial x}=\partial_u\tilde{f}+\partial_v\tilde{f}, \tag{91.293}$$

EqeCDzzutEqderuncx

donc

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \partial^2_{uu}\tilde{f}\frac{\partial u}{\partial x} + \partial^2_{vu}\tilde{f}\frac{\partial v}{\partial x} + \partial^2_{uv}\tilde{f}\frac{\partial u}{\partial x} + \partial^2_{vv}\tilde{f}\frac{\partial v}{\partial x}\\ &= \partial^2_{uu}\tilde{f} + 2\partial^2_{uv}\tilde{f} + \partial^2_{vv}\tilde{f}.\end{aligned} \tag{91.294}$$

Pour calculer $\partial^2_{xy}f$, nous pouvons calculer $\partial^2_{yx}f$ et donc dériver par rapport à y l'équation (91.293) :

$$\begin{aligned}\frac{\partial^2 f}{\partial x\partial y} &= \frac{\partial}{\partial y}\big(\partial_u\tilde{f} + \partial_v\tilde{f}\big)\\ &= 2\partial^2_{uu}\tilde{f} + \frac{5}{2}\partial^2_{uv}\tilde{f} + \frac{1}{2}\partial^2_{vv}\tilde{f}\end{aligned} \tag{91.295}$$

où nous avons utilisé le fait que $\partial^2_{uv}\tilde{f} = \partial^2_{vu}\tilde{f}$. En ce qui concerne les dérivées par rapport à y,

$$\frac{\partial f}{\partial y} = 2\partial_u\tilde{f} + \frac{1}{2}\partial_v\tilde{f} \tag{91.296}$$

et

$$\begin{aligned}\frac{\partial^2 f}{\partial y^2} &= 2\partial^2_{uu}\tilde{f}\frac{\partial u}{\partial y} + 2\partial_{vu}\tilde{f}\frac{\partial v}{\partial y} + \frac{1}{2}\partial^2_{uv}\tilde{f}\frac{\partial u}{\partial y} + \frac{1}{2}\partial^2_{vv}\tilde{f}\frac{\partial v}{\partial y}\\ &= 4\partial^2_{uu}\tilde{f} + 2\partial^2_{uv}\tilde{f} + \frac{1}{4}\partial^2_{vv}\tilde{f}.\end{aligned} \tag{91.297}$$

En écrivant l'équation à résoudre avec ces expressions, nous trouvons $-\frac{9}{12}\partial^2_{uv}\tilde{f} = 0$, c'est-à-dire

$$\frac{\partial}{\partial u}\left(\frac{\partial\tilde{f}}{\partial v}\right) = 0. \tag{91.298}$$

Étant donné que la dérivée de $\partial_v\tilde{f}$ par rapport à u est nulle, nous savons qu'il existe une fonction ψ de v uniquement telle que

$$\frac{\partial\tilde{f}}{\partial v} = \psi(v). \tag{91.299}$$

En intégrant par rapport à v, et en se souvenant que la constante d'intégration peut être une fonction de u, nous trouvons qu'il existe une fonction φ de u telle que

$$\tilde{f}(u,v) = \int\psi(v)dv + \varphi(u). \tag{91.300}$$

La fonction ψ étant arbitraire, sa primitive est arbitraire. Quitte à redéfinir ψ, nous écrivons

$$\tilde{f}(u,v) = \psi(v) + \varphi(u). \tag{91.301}$$

Nous exprimons maintenant la réponse en termes de x et y :

$$f(x,y) = \tilde{f}\big(u(x,y),v(x,y)\big) = \psi(x+\frac{1}{2}y) + \varphi(x+2y). \tag{91.302}$$

Exercise 251 exoCalculDifferentiel0015

Soient $f,g:\mathbb{R}^N\to\mathbb{R}$ deux applications différentiables en $x_0\in\mathbb{R}^N$.

(1) Montrer que les applications $f+g$ et fg sont différentiables en x_0 et que

$$D(f+g)(x_0) = Df(x_0) + Dg(x_0) \tag{91.303a}$$
$$D(fg)(x_0) = f(x_0)Dg(x_0) + g(x_0)Df(x_0). \tag{91.303b}$$

(2) En déduire que pour tout $n\in\mathbb{N}_0$ nous avons

$$Df^n(x_0) = nf^{n-1}(x_0)Df(x_0). \tag{91.304}$$

corrCalculDifferentiel0015

Correction of the exercise 251

Les deux premières égalités sont respectivement la proposition 12.269 et le lemme 12.270. En ce qui concerne la différentielle de la puissance n de f, nous procédons par récurrence. Si la formule est valable pour n, alors

$$\begin{aligned} F(f^{n+1})(x_0) &= D(ff^n)(x_0) \\ &= f(x_0)Df^n(x_0) + f^n(x_0)Df(x_0) \\ &= f(x_0)nf^{n-1}(x_0)Df(x_0) + f^n(x_0)Df(x_0) \\ &= nf^n(x_0)Df(x_0) + f^n(x_0)Df(x_0) \\ &= (n+1)f^n(x_0)Df(x_0). \end{aligned} \tag{91.305}$$

Exercise 252 exoCalculDifferentiel0016

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une application différentiable sur $\mathbb{R}^N$ telle que pour tout $x \in \mathbb{R}^N$,

$$df(x) = 0. \tag{91.306}$$

Montrer qu'il existe une constante $C \in \mathbb{R}$ telle que $f(x) = C$. corrCalculDifferentiel0016

Correction of the exercise 252

Cet exercice est une partie de la proposition 12.315.

Exercise 253 ⊛⊛ !⊛⊛ exoCalculDifferentiel0017

Soit $f : \mathbb{R}^N \to \mathbb{R}$ une application deux fois différentiable sur $\mathbb{R}^N$ telle que

$$d^2f(x) = 0 \tag{91.307}$$

pour tout $x \in \mathbb{R}^N$. Montrer qu'il existe $a \in \mathbb{R}^N$ et $b \in \mathbb{R}$ tels que

$$f(x) = a \cdot x + b. \tag{91.308}$$

corrCalculDifferentiel0017

Correction of the exercise 253

Il faut d'abord bien comprendre ce que signifie $d^2(f)$ parce que déjà df est une application linéaire. En ce qui concerne les espaces, nous avons

$$\begin{aligned} f &: \mathbb{R}^N \to \mathbb{R} \\ df &: \mathbb{R}^N \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}) \\ d(df) &: \mathbb{R}^n \to \mathcal{L}\big(\mathbb{R}^N, \mathcal{L}(\mathbb{R}^N, \mathbb{R})\big). \end{aligned} \tag{91.309}$$

En général, nous avons toujours, si $f : A \to B$, alors $df : A \to \mathcal{L}(A, B)$. Nous avons juste appliqué ce principe à f puis à df lui-même.

La proposition 12.315 dit que lorsque df est nulle, f est constante. En extrapolant un peu ce résultat, nous acceptons que si $d(df)$ est nulle, alors df est constante en tant que application de $\mathbb{R}^N$ dans $\mathcal{L}(\mathbb{R}^N, \mathbb{R})$. Il existe donc $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ tel que

$$T = df(x) \tag{91.310}$$

CD17dfxT

pour tout x. Si nous considérons la base canonique $\{e_i\}$, nous pouvons expliciter (91.310) en l'appliquant le vecteurs e_k :

$$T(e_k) = df(x).e_k = \sum_i \frac{\partial f}{\partial x_i}(x)\delta_{ik} = \frac{\partial f}{\partial x_k}(x) \tag{91.311}$$

parce que $(e_k)_i = \delta_{ik}$. Nous notons $a_k = T(e_k)$; c'est un réel bien défini. Maintenant nous avons l'équation

$$\frac{\partial f}{\partial x_k}(x) = a_k \tag{91.312}$$

pour tout $x \in \mathbb{R}^N$. Si nous l'intégrons par rapport à x_k, nous trouvons

$$f(x) = x_k a_k + \varphi(x_1, \ldots, \hat{x}_k, \ldots, x_N) \tag{91.313}$$

où φ est la « constante » d'intégration. Elle dépend de tous les x_i sauf de x_k (c'est le sens du chapeau que nous avons mis sur x_k). L'équation (91.313) est valable pour chaque k. Nous avons donc en réalité N équations. En dimension deux, nous aurions les équations

$$f(x_1, x_2) = x_1 a_1 + \varphi_1(x_2) \tag{91.314a}$$
$$f(x_1, x_2) = x_2 a_2 + \varphi_2(x_1). \tag{91.314b}$$

La seule façon de choisir le fonction φ_1 et φ_2 de telle façon à avoir les deux égalités en même temps est de prendre $\varphi_1(x_2) = a_2 x_2 + C$ et $\varphi_2(x_1) = a_1 x_1 + C$. La même chose se passe en dimension plus grande : la seule façon de choisir les fonctions φ_k de façon à satisfaire toutes les équations (91.313) en même temps est de prendre

$$\varphi_k(x_1, \ldots, \hat{x}_k, \cdots, x_N) = a_1 x_1 + \cdots + a_{k-1} x_{k-1} + a_{k+1} x_{k+1} + \cdots + a_N x_N + C, \tag{91.315}$$

et donc nous devons avoir,

$$f(x) = a_1 x_1 + \cdots + a_N x_N + C, \tag{91.316}$$

c'est-à-dire exactement $f(x) = a \cdot x + C$ où C est une constante.

Exercise 254

Soient $f : \mathbb{R}^2 \to \mathbb{R}$ et $g : \mathbb{R}^2 \to \mathbb{R}$ et $h : \mathbb{R} \to \mathbb{R}$ définies par

$$\begin{aligned} g(x, y) &= f(y, x) \\ h(x) &= f(x, x) \end{aligned} \tag{91.317}$$

pour tout x et y dans $\mathbb{R}$. Calculer les dérivées partielles premières de g et h en fonction de celles de f, en supposant f de classe C^1 sur $\mathbb{R}^2$.

Correction of the exercise 254

Le plus simple est de calculer à partir des définitions. Nous avons

$$\begin{aligned} \frac{\partial g}{\partial x}(x, y) &= \lim_{t \to 0} \frac{g(x + t, y) - g(x, y)}{t} \\ &= \lim_{t \to 0} \frac{f(y, x + t) - f(y, x)}{t} \\ &= \frac{\partial f}{\partial y}(y, x); \end{aligned} \tag{91.318}$$

ensuite

$$\begin{aligned} \frac{\partial g}{\partial y}(x, y) &= \lim_{t \to 0} \frac{g(x, y + t) - g(x, y)}{t} \\ &= \lim_{t \to 0} \frac{f(y + t, x) - f(y, x)}{t} \\ &= \frac{\partial f}{\partial x}(y, x). \end{aligned} \tag{91.319}$$

Pour h, le plus simple est de considérer h comme une fonction composée :

$$h(x) = (f \circ \varphi)(x, y) \tag{91.320}$$

où $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2$ est donnée par $\varphi(x, y) = (x, x)$. En utilisant la formule de dérivation des fonctions composées,

$$\begin{aligned} h'(x) &= \partial_1 f\big(\varphi(x, y)\big) \underbrace{\frac{\partial \varphi_1}{\partial x}(x, y)}_{=1} + \partial_2 f\big(\varphi(x, y)\big) \underbrace{\frac{\partial \varphi_2}{\partial x}(x, y)}_{=1} \\ &= \frac{\partial f}{\partial x}(x, x) + \frac{\partial f}{\partial y}(x, x). \end{aligned} \tag{91.321}$$

Exercise 255 exoCalculDifferentiel0020

Soient $f : \mathbb{R} \to \mathbb{R}$ de classe C^2 et $g : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$g(x,y) = \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{si } x \neq y \\ f'(x) & \text{si } x = y \end{cases} \tag{91.322}$$

Montrer que g est de classe C^1 sur $\mathbb{R}^2$. corrCalculDifferentiel0020

Correction of the exercise 255

Nous devons montrer que g est continue, et puis que ses dérivées partielles existent et sont également continues.

Commençons par prouver que g est continue. Nous devons prouver que pour tout $h \in \mathbb{R}$,

$$\lim_{(x,y)\to(h,h)} \frac{f(x)-f(y)}{x-y} = f'(h). \tag{91.323}$$ EqeCDzvgtlly

Étant donné que f est de classe C^2, nous pouvons effectuer un développement en suivant la formule (12.1327).

$$f(x) = f(h) + f'(h)(x-h) + \alpha(x)(x-h) \tag{91.324a}$$ EqeCDzzszfx

$$f(y) = f(h) + f'(h)(y-h) + \alpha(y)(y-h) \tag{91.324b}$$

où la fonction α a la propriété que $\lim_{x\to h} \alpha(x) = 0$. En substituant ces expressions dans la limite à calculer (91.323), nous trouvons

$$\lim_{(x,y)\to(h,h)} \frac{f(x)-f(y)}{x-y} = f'(h) + \lim_{(x,y)\to(h,h)} \big(\alpha(x)-\alpha(y)\big) = f'(h), \tag{91.325}$$

ce qu'il fallait. La fonction g est donc continue. Il s'agit maintenant de prouver la continuité des dérivées partielles aux points (h,h) (la fonction f étant C^2, le quotient qui définit g est certainement C^1 aux points autres que (h,h)).

En vertu du théorème 12.294, pour prouver que g est C^1, il suffit de prouver que les dérivées partielles existent et sont continues. Nous devons donc vérifier que $\lim_{(x,y)\to(h,h)} \partial_x f(x,y) = \partial_x f(h,h)$ (et idem pour y). D'abord, si (x,y) n'est pas de la forme (h,h), nous avons

$$\partial_x g(x,y) = \frac{f'(x)(x-y) - \big(f(x)-f(y)\big)}{(x-y)^2}. \tag{91.326}$$ Eqzzdzpxgxyun

Ensuite,

$$\partial_x f(h,h) = \lim_{t\to 0} \frac{g(h+t,h)-g(h,h)}{t} = \lim_{t\to 0} \frac{\frac{f(h+t)-f(h)}{(h+t)-h} - f'(h)}{t}. \tag{91.327}$$

Pour effectuer cette limite, il ne suffit pas d'écrire $f(h+t)$ en substituant x par $h+t$ dans (91.324a). En effet, simplifications, il reste

$$\lim_{t\to 0} \frac{\frac{f'(h)t+\alpha(h+t)t}{t} - f'(h)}{t} = \lim_{t\to 0} \frac{\alpha(h+t)}{t}. \tag{91.328}$$

Cette dernière limite est indéterminée ; il faut donc développer jusqu'à un ordre supérieur :

$$f(h+t) = f(h) + tf'(h) + \frac{t^2}{2} f''(h) + \beta(t)t^2 \tag{91.329}$$

En effectuant le calcul, nous trouvons que

$$\partial_x f(h,h) = \frac{1}{2} f''(h). \tag{91.330}$$

Nous calculons par ailleurs $\lim_{(x,y)\to(h,h)} \partial_x g(x,y)$ en substituant le développement (voir (12.1328))

$$f(y) = f(x) + f'(x)(y-x) + f''(x)\frac{(y-x)^2}{2} + \beta(y-x)(y-x)^2 \tag{91.331}$$

dans l'équation (91.326) :

$$\begin{aligned}\lim_{(x,y)\to(h,h)} \partial_x g(x,y)&\\ &= \lim_{(x,y)\to(h,h)} \frac{f'(x)(x-y) - \Big(- f'(x)(y-x) - f''(x)\frac{(y-x)^2}{2} - \alpha(y-x)(y-x)^2\Big)}{(x-y)^2}\\ &= \lim_{(x,y)\to(h,h)} \frac{f''(x)}{2} - \alpha(y-x)\\ &= \frac{f''(h)}{2}.\end{aligned} \tag{91.332}$$

Étant donné que $\lim_{(x,y)\to(h,h)} \partial_x g(x,y) = \partial_x g(h,h)$, nous avons la continuité de la dérivée première dans la direction de x. La dérivée dans la direction de y se traite de la même façon, et nous obtenons la continuité.

La fonction g est donc C^1 par le théorème 12.294 parce que ses dérivées partielles sont continues.

91.5 Intégrales multiples

Exercise 256 ⊙⊙ exoIntegralesMultiples0001

Soit $D = [0,1]^2$. Calculer $\iint_D \frac{dx\,dy}{(x+y+1)^2}$.

corrIntegralesMultiples0001

Correction of the exercise 256

Nous décomposons l'intégrale double en deux intégrales simples en utilisant le théorème de Fubini (voir les théorèmes 20.177 et 20.26.8) :

$$I = \iint_{[0,1]\times[0,1]} \frac{1}{(x+y+1)^2} dxdy = \int_{[0,1]} \left[\int_{[0,1]} \frac{1}{(x+y+1)^2} dx \right] dy. \tag{91.333}$$

La première intégrale à faire est l'intégrale $\int_{[0,1]} \frac{1}{(x+y+1)^2} dx$. Cela est une intégrale par rapport à x dans laquelle nous devons considérer y comme constante. Cela est donc une intégrale de la forme $\int_0^1 \frac{1}{(x+a)^2}$ avec $a = y+1$. Ce type d'intégrale s'effectue en posant $t = x + a$.

Nous posons donc $t = x + y + 1$, et nous calculons :

$$\int_0^1 \frac{1}{(x+y+1)^2} dx = \int_{y+1}^{y+2} \frac{1}{t^2} dt = \left[-\frac{1}{t} \right]_{y+1}^{y+2} = \frac{1}{y+1} - \frac{1}{y+2}. \tag{91.334}$$

Maintenant nous pouvons poursuivre :

$$I = \int_0^1 \left(\frac{1}{y+1} - \frac{1}{y+2} \right) dy = 2\ln(2) - \ln(3). \tag{91.335}$$

Pour obtenir cela, nous avons intégré séparément les deux termes, en utilisant les changements de variables $t_1 = y+2$ et $t_2 = y+1$. Ne pas oublier que $\int \frac{dx}{x} = \ln(x)$.

Exercise 257 exoIntegralesMultiples0002

Calculer les intégrales suivantes sur les domaines donnés.

(1) ⊙⊙ $I_1 = \iint_D (x+y)e^{-x}e^{-y}dxdy$ où $D = \big\{(x,y) \in \mathbb{R}^2 \text{ tels que } x,y \geqslant 0, x+y \leqslant 1\big\}$.

(2) ⊙⊙ $I_2 = \iint_D (x^2+y^2)dxdy$ où $D = \{(x,y) \in \mathbb{R}^2$ tels que $x^2+y^2 < x, x^2+y^2 > y\}$. ItemexoIntegralesMultiples0002ii

(3) ⊙⊙ $I_3 = \iint_D \frac{xy}{1+x^2+y^2}dxdy$ où $D = \{(x,y) \in [0,1]^2$ tels que $x^2+y^2 \geqslant 1\}$. ItemexoIntegralesMultiples0002iii

(4) ⊖⊖ $I_4 = \iint_D \frac{1}{x+xy+y+1}dxdy$ où $D = [0, \frac{\pi}{2}] \times [0, \frac{1}{2}]$.

(5) ⊙⊙ $I_5 = \iiint_D z\,dxdydz$ où $D = \{(x,y,z) \in (\mathbb{R}^+)^3$ tels que $y^2+z \leqslant 1, x^2+z \leqslant 1\}$.

(6) ⊖⊖ $I_5 = \iint_D xy\,dxdy$ où $D = \left\{(x,y) \in \mathbb{R}^2 \text{ tels que } x,y>0, \frac{x^2}{a^2}+\frac{y^2}{b^2} \leqslant 1\right\}$ avec $a,b>0$.

corrIntegralesMultiples0002

Correction of the exercise 257

La véritable difficulté de l'exercice est de décomposer le domaine correctement. Après, ce sont des intégrales à effectuer.

(1) Dans le domaine, x peut prendre n'importe quelle valeurs entre 0 et 1. Mais une fois que x est fixé, la variable y ne peut varier qu'entre 0 et $1-x$. Nous décomposons donc l'intégrale de la façon suivante :

$$I_1 = \int_0^1 \left[\int_0^{1-x} (x+y)e^{-x}e^{-y}dy\right]dx = \int_0^1 \left((x+1)e^{-x} - 2e^{-1}\right)dx = 2-5e^{-1}. \tag{91.336}$$

(2) La combinaison x^2+y^2 nous incite à passer en polaires. Le domaine est donné par les inéquations

$$\begin{cases} r^2 < r\cos(\theta) & \text{(91.337a)} \\ r^2 > r\sin(\theta). & \text{(91.337b)} \end{cases}$$

Étant donné que r est toujours positif, on peut simplifier ces inéquations par r sans toucher au sens des inégalités. Le domaine sera donc $\sin(\theta) < r$ et $r < \cos(\theta)$. Notons que quand $\sin(\theta) < 0$, alors le domaine de variation de r est $[0, \cos(\theta)[$ (et non $]\sin(\theta), \cos(\theta)[$). Il faudra donc intégrer séparément la partie du domaine avec $\sin(\theta) < 0$ et celle avec $\sin(\theta) > 0$.
Si $\sin(\theta) < 0$ la condition $\cos(\theta) > r \geqslant 0$ nous dit que θ varie entre $-\pi/2$ et 0.
Si $\sin(\theta) > 0$ la condition $\cos(\theta) > r > \sin(\theta)$ nous dit que θ varie entre 0 et $\pi/4$.
L'intégrale à effectuer est donc

$$I_2 = \underbrace{\int_{-\pi/2}^{0}\int_0^{\cos\theta} r^3\,drd\theta}_{=\frac{3\pi}{64}} + \underbrace{\int_0^{\pi/4}\int_{\sin\theta}^{\cos\theta} r^3\,drd\theta}_{=\frac{1}{8}} = \frac{3\pi}{64}+\frac{1}{8}. \tag{91.338}$$

Le dessin du domaine est lui aussi très intéressant à étudier. L'équation $x^2+y^2 = x$ est un cercle parce que, en reformant le carré, nous avons successivement

$$\begin{aligned} x^2 - x + y^2 &= 0 \\ \left(x-\frac{1}{2}\right)^2 - \frac{1}{4} + y^2 &= 0 \\ \left(x-\frac{1}{2}\right)^2 + y^2 &= \left(\frac{1}{2}\right)^2, \end{aligned} \tag{91.339}$$

ce qui est l'équation du cercle de rayon 1/2 et de centre $(\frac{1}{2}, 0)$. Par conséquent l'inéquation $x^2+y^2 < x$ correspond à l'intérieur de ce cercle.
De la même façon, l'inéquation $x^2+y^2 > y$ correspond à l'extérieur du cercle de rayon $\frac{1}{2}$ et de centre $(0, \frac{1}{2})$. Le domaine est dessiné sur la figure 91.14. On y voit que pour la partie en dessous de $y = 0$, il n'y a pas de contraintes sur le rayon (à part de rester dans le cercle).

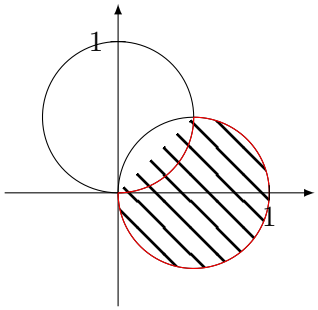
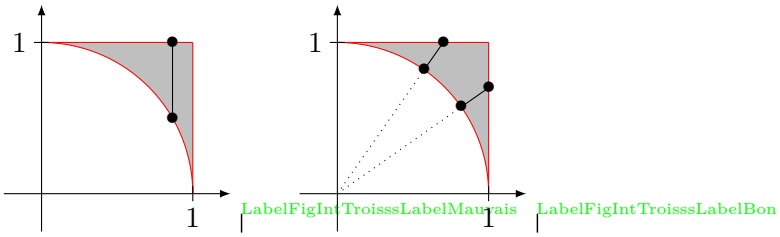


FIGURE 91.14: Domaine d'intégration pour l'exercice 257.(2). LabelFigDeuxCercles



(a) Intégrer en cartésiennes. (b) Intégrer en polaires.

FIGURE 91.15: Domaine d'intégration pour l'exercice 257.(3). LabelFigIntTrois

(3) Nous devons intégrer sur la partie du carré de côté 1 qui ne se trouve pas dans le cercle. Nous pourrions, pour chaque x entre 0 et 1, pour y entre $\sqrt{1-x^2}$ et 1, comme indiqué sur la figure 91.15(a), mais nous pouvons aussi intégrer en polaires comme indiqué sur la figure 91.15(b).

Dans le cas des polaires, on intègre θ de 0 à $\pi/2$, et puis pour r, le début de l'intégration est en $r = 1$ tandis que la fin de l'intégration est donnée par l'intersection entre la droite d'angle θ et soit le côté vertical soit le côté horizontal du carré selon que θ soit plus petit ou plus grand que $\pi/4$.

Pour $\theta < \pi/4$, si ρ est la valeur de r à laquelle la droite d'angle θ intersecte le côté vertical, et $y = \rho \sin\theta$ la hauteur à laquelle ça se passe, nous avons $\rho^2 = 1 + y^2$, c'est-à-dire $\rho = 1/\cos(\theta)$.

Pour les angles entre $\pi/4$ et $\pi/2$, ce que nous obtenons est $\rho = 1/\sin\theta$.

L'intégrale à calculer en polaire est donc

$$I_3 = \int_0^{\pi/4} \int_1^{1/\cos\theta} f\big(r\cos\theta, r\sin\theta\big) r\, dr d\theta + \int_{\pi/4}^{\pi/2} \int_1^{1/\sin\theta} f\big(r\cos\theta, r\sin\theta\big) r\, dr d\theta. \tag{91.340}$$

Bonne chance.

En cartésiennes, les choses se révèlent plus facile. En effet nous devons faire

$$I_3 = \int_0^1 \int_{\sqrt{1-x^2}}^1 \frac{xy}{1+x^2+y^2}\, dy dx. \tag{91.341}$$

Il faut bien comprendre que nous commençons par intégrer par rapport à y en traitant x comment une « vulgaire » constante que nous pouvons même sortir de l'intégrale :

$$\begin{aligned} I_3 &= \int_0^1 x \left[\int_{\sqrt{1-x^2}}^1 \frac{y}{(1+x^2)+y^2} dy \right] dx \\ &= \int_0^1 x \left[\frac{1}{2} \ln(x^2+y^2+1) \right]_{\sqrt{1-x^2}}^1 dx \\ &= \frac{1}{2} \int_0^1 \Big(x \ln(x^2+2) - x\ln(2) \Big) dx. \end{aligned} \tag{91.342}$$

Nous avons utilisé la primitive classique $\int y/(a+y^2) = \ln(y^2+a)/2$ lorsque a est positif, ce qui est le cas de $1+x^2$. À partir de là, nous sommes sur une intégrale usuelle, très similaire à (18.79). Le résultat est :

$$I_3 = \frac{3}{4}\ln\left(\frac{3}{2}\right) - \frac{1}{4}. \tag{91.343}$$

Une des morales de cet exercice est qu'il y a des situations dans lesquelles la façon de paramétrer le domaine change radicalement la difficulté du calcul effectif des intégrales. Ici c'est nettement plus simple en cartésiennes.

(4) Attention à l'ordre des intégrales. Ceci est faux :

$$I_4 = \int_0^{\pi/2}\int_0^{1/2} \frac{1}{x+xy+y+1}dxdy. \tag{91.344}$$

En effet dans cette écriture c'est x qu'on intègre de 0 à $\frac{1}{2}$, or dans le domaine la variable x est dans $[0, \frac{\pi}{2}]$. Ce qui est correct est

$$I_4 = \int_0^{1/2}\int_0^{\pi/2} \frac{1}{x+xy+y+1}dxdy. \tag{91.345}$$

La première chose à faire est de factoriser le dénominateur : $x+xy+y+1 = (x+1)(y+1)$. Ensuite, nous pouvons diviser l'intégrale en deux intégrales indépendantes en sortant $\frac{1}{y+1}$ de l'intégrale sur x :

$$I_4 = \int_0^{1/2}\frac{1}{y+1}dy\int_0^{\pi/2}\frac{1}{x+1}dx = \ln\left(\frac{\pi}{2}+1\right)\ln\left(\frac{3}{2}\right). \tag{91.346}$$

(5) Une façon de paramétrer le domaine est le suivant :

$$\begin{aligned} z&\colon 0\to 1\\ y&\colon 0\to\sqrt{1-z}\\ x&\colon 0\to\sqrt{1-z}. \end{aligned} \tag{91.347}$$

Nous laissons donc aller z de 0 à 1, et pour chacune des valeurs de z, nous contraignons x et y en fonction de z par les équations $y^2+z\leqslant 1$ et $x^2+z\leqslant 1$. L'intégrale est donc, en se rappelant que $\int z\,dx = xz$,

$$\begin{aligned} I_4 &= \int_0^1\int_0^{\sqrt{1-z}}\int_0^{\sqrt{1-z}} z\,dxdydz\\ &= \int_0^1\int_0^{\sqrt{1-z}}[xz]_{x=0}^{x=\sqrt{1-z}}dydz\\ &= \int_0^1\int_0^{\sqrt{1-z}} z\sqrt{1-z}\,dydz\\ &= \int_0^1[yz\sqrt{1-z}]_0^{\sqrt{1-z}}\,dz\\ &= \int_0^1 z(1-z)\,dz\\ &= \frac{1}{6}. \end{aligned} \tag{91.348}$$

(6) Nous avons envie d'utiliser une version un peu modifiée des coordonnées polaires : $x = ar\cos(\theta)$, $y = br\sin(\theta)$, de telle façon à épouser les formes du domaine. Pour intégrer, il faut comprendre et utiliser le théorème 14.281. Ici,

$$\phi(r,\theta) = \big(ar\cos(\theta), br\sin(\theta)\big), \tag{91.349}$$

et l'intégrale devient

$$\int_D f(x,y)dxdy = \int_{\phi^{-1}(D)} f\big(\phi(r,\theta)\big)|J_\phi(r,\theta)|drd\theta. \tag{91.350}$$

Le jacobien est donné par

$$|J_\phi(r,\theta)| = \det\begin{pmatrix} \frac{\partial\phi_1}{\partial r} & \frac{\partial\phi_1}{\partial\theta} \\ \frac{\partial\phi_2}{\partial r} & \frac{\partial\phi_2}{\partial\theta} \end{pmatrix} = \begin{pmatrix} a\cos(\theta) & -ar\sin(\theta) \\ b\sin(\theta) & br\cos(\theta) \end{pmatrix} = abr. \tag{91.351}$$

Notez que dans le cas particulier $a = b = 1$, nous retrouvons le jacobien usuel r des coordonnées polaires. Il faut trouver $\phi^{-1}(D)$. Pour parcourir le quart d'ellipse proposé, il faut $r\colon 0 \to 1$ et $\theta\colon 0 \to \pi/2$. Nous devons calculer

$$\begin{aligned} I_5 &= \int_{[0,1]\times[0,\frac{\pi}{2}]} (ar\cos\theta)(br\sin\theta)abr\, drd\theta \\ &= a^2b^2\int_0^{\pi/2}\int_0^1 r^3\cos\theta\sin\theta\, drd\theta \\ &= \frac{a^2b^2}{8} \end{aligned} \tag{91.352}$$

où nous avons utilisé l'intégrale (18.87).

Exercise 258 exoIntegralesMultiples0003

Soit D le disque de centre $(0,1)$ et de rayon 1 du plan. Calculer

$$\iint_D (x^2+y^2)dxdy. \tag{91.353}$$

corrIntegralesMultiples0003

Correction of the exercise 258

L'équation qui décrit le cercle D en coordonnées polaires est $r = 2\sin(\theta)$. Dans l'intégrale il faut donc faire varier r entre 0 et $2\sin(\theta)$ et θ entre 0 et π. L'intégrale à calculer devient

$$\begin{aligned} \int_0^\pi\int_0^{2\sin(\theta)} 2r^2\sin(\theta)\, drd\theta &= \int_0^\pi \Big[\frac{2r^3}{3}\sin(\theta)\Big]_{r=0}^{r=2\sin(\theta)} d\theta \\ &= \int_0^\pi \Big(\frac{16\sin^4(\theta)}{3}\Big)d\theta \\ &= 2\pi. \end{aligned} \tag{91.354}$$

Exercise 259 exoIntegralesMultiples0004

Soit $D = \{x^2+y^2 \leqslant 1, 0 \leqslant z \leqslant 1-(x^2+y^2)\} \subset \mathbb{R}^3$. Calculer le volume de D. corrIntegralesMultiples0004

Correction of the exercise 259

En coordonnées cylindriques, les bornes d'intégration sont données par

$$\begin{cases} r \in [0,1] & \text{(91.355a)} \\ z \in [0, 1-r^2] & \text{(91.355b)} \\ \theta \in [0, 2\pi]. & \text{(91.355c)} \end{cases}$$

La fonction à intégrer est 1 (pour le volume) qu'il faut multiplier par le jacobien des coordonnées cylindriques (r) :

$$\int_0^1\int_0^{1-r^2}\int_0^{2\pi} r\, d\theta dz dr = -2\pi\int_0^1 r(r^2-1)\, dr = \frac{\pi}{2}. \tag{91.356}$$

Exercise 260 exoIntegralesMultiples0005

Soit $D = \{x \geqslant 0, y \geqslant 0, z \geqslant 0, x + y + z \leqslant 1\}$. Calculer :

$$\iiint_D \frac{dx\,dy\,dz}{(1+x+y+z)^3}. \tag{91.357}$$

corrIntegralesMultiples0005

Correction of the exercise 260

Le domaine d'intégration est donné par les bornes

$$\begin{cases} x \in [0,1] & \text{(91.358a)} \\ y \in [0, 1-x] & \text{(91.358b)} \\ z \in [0, 1-x-y]. & \text{(91.358c)} \end{cases}$$

L'intégrale à calculer devient

$$I = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} \frac{1}{(1+x+y+z)^3} dz dy dx. \tag{91.359}$$

L'intégrale selon z se fait avec la formule $\int \frac{1}{(a+t)^3} dt = -\frac{1}{2}\frac{1}{(a+t)^2}$ avec $a = 1+x+y$. Il vient

$$I = \int_0^1 \int_0^{1-x} \left[-\frac{1}{2} \frac{1}{(1+x+y+z)^2} \right]_{z=0}^{z=1-x-y} dy dx = \frac{\ln(2)}{2} - \frac{5}{16}. \tag{91.360}$$

Exercise 261 exoIntegralesMultiples0006

Soit $D = \{(x,y); x \geqslant 0, y \geqslant 0, x^2 + y^2 \leqslant 1\}$. Calculer $\iint_D (4 - x^2 - y^2)\, dxdy$.

corrIntegralesMultiples0006

Correction of the exercise 261

Le domaine est le quart de disque de rayon 1 et de centre $(0,0)$. Il est donc conseillé de passer aux coordonnées polaires. En comptant le jacobien r, l'intégrale à effectuer est

$$I = \int_0^{\pi/2} \int_0^1 (4 - r^2) r\, dr d\theta = \frac{7\pi}{8}. \tag{91.361}$$

Exercise 262 exoIntegralesMultiples0007

Identifier les ensembles suivants et calculer leur aire s'ils sont dans $\mathbb{R}^2$, leur volume s'ils sont dans $\mathbb{R}^3$.

(1) $D = \left\{(x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}$ avec $a, b > 0$;

(2) $D = \left\{(x,y,z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1\right\}$ avec $a, b, c > 0$; qu'obtient-on dans le cas particulier où D est la boule unité de $\mathbb{R}^3$?

(3) $D = \left\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 \leqslant R\,,\ 0 \leqslant z \leqslant h\right\}$ avec $R, h > 0$;

(4) $D = \left\{(x,y,z) \in \mathbb{R}^3 \mid x \geqslant 0\,,\ y \geqslant 0\,,\ z \geqslant 0\,,\ x + y + z \leqslant 1\right\}$;

(5) $D = \left\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 \leqslant z^2/h^2\,,\ 0 \leqslant z \leqslant h\right\}$ avec $h > 0$.

corrIntegralesMultiples0007

Correction of the exercise 262

(1) Ceci est une ellipse de demi-grands axes a et b. Les coordonnées qui s'invitent d'elles mêmes sont

$$\begin{cases} x = ar\cos(\theta) & \text{(91.362a)} \\ y = br\sin(\theta), & \text{(91.362b)} \end{cases}$$

dont le jacobien est abr.

$$A = \int_0^1 \int_0^{2\pi} abr\, d\theta dr = ab\pi. \tag{91.363}$$

Notez que si $a = b = R$, c'est le cercle de rayon R et nous avons la formule habituelle πR^2 pour l'aire.

(2) Ici c'est la même chose, mais avec trois variables. On prend les coordonnées sphériques modifiées

$$\begin{cases} x = a\rho\cos(\theta)\sin(\phi) & (91.364a)\\ y = b\rho\sin(\theta)\sin(\phi) & (91.364b)\\ z = c\rho\cos(\phi), & (91.364c)\end{cases}$$

dont le jacobien vaut $abc\rho^2\sin(\phi)$. L'intégrale à calculer est donc

$$V = abc\int_0^1\int_{-\pi}^{\pi}\int_0^{\pi}\rho^2\sin(\phi)\,d\phi d\theta d\rho = \frac{4\pi}{3}abc. \tag{91.365}$$

Si $a = b = c = R$, nous avons la sphère de rayon R, et nous retrouvons la formule classique du volume de la sphère $V = \frac{4}{3}\pi R^3$.

(3) L'ensemble proposé est un cylindre plein de rayon R et de hauteur h. Les coordonnées cylindriques sont là pour ça :

$$V = \int_0^R\int_0^h\int_0^{2\pi} r d\theta dz dr = \pi R^2 h. \tag{91.366}$$

(4) Il s'agit du volume contenu en dessous du plan $x + y + z - 1 = 0$ dans le premier octant. En fait, il s'agit de la même pyramide qu'on a déjà rencontrée dans l'exercice 260. L'intégrale à faire est

$$V = \int_0^1\int_0^{1-x}\int_0^{1-x-y} dzdydx = \frac{1}{6}. \tag{91.367}$$

(5) Ceci est un cône posé sur sa pointe et de hauteur h. En effet, passons aux coordonnées cylindriques, nous avons $r^2 < z^2/h^2$, et donc $r < z/h$ parce que r est toujours positif. À chaque hauteur z, nous avons donc un disque de rayon z/h. Pour calculer le volume, c'est l'intégrale

$$V = \int_0^h\int_0^{z/h}\int_0^{2\pi} r\,d\theta dr dz = \frac{\pi h}{3}. \tag{91.368}$$

Exercise 263 exoIntegralesMultiples0008

Quel est le volume délimité par deux cylindres de révolution d'axes (Ox) et (Oy) et de même rayon $R > 0$?

corrIntegralesMultiples0008

Correction of the exercise 263

Les équations du domaine sont

$$\begin{cases} x^2 + z^2 \leqslant R^2 & (91.369a)\\ y^2 + z^2 \leqslant R^2. & (91.369b)\end{cases}$$

Vu que z est la variable qui arrive dans les deux équations, le plus simple est de faire varier z d'abord et de faire varier x et y en fonction :

$$\begin{aligned} z &\in [-R, R]\\ x &\in [-\sqrt{R^2 - z^2}, \sqrt{R^2 - z^2}]\\ y &\in [-\sqrt{R^2 - z^2}, \sqrt{R^2 - z^2}]. \end{aligned} \tag{91.370}$$

L'intégrale à calculer devient

$$\begin{aligned} V &= \int_{-R}^{R}\int_{-\sqrt{R^2-z^2}}^{\sqrt{R^2-z^2}}\int_{-\sqrt{R^2-z^2}}^{\sqrt{R^2-z^2}} 1\,dydxdz\\ &= 2\int_{-R}^{R}\int_{-\sqrt{R^2-z^2}}^{\sqrt{R^2-z^2}}\sqrt{R^2 - z^2}\,dxdz\\ &= 4\int_{-R}^{R}(R^2 - z^2)\,dz\\ &= \frac{16}{3}R^3. \end{aligned} \tag{91.371}$$

Exercise 264 exoIntegralesMultiples0009

Soit $D = \{(x^2 + y^2)^2 \leqslant xy\}$. Calculer la surface du domaine, ainsi que l'intégrale

$$\iint_D \sqrt{xy}\, dx\, dy. \tag{91.372}$$

corrIntegralesMultiples0009

Correction of the exercise 264

En coordonnées polaires, le domaine d'intégration est $r^2 < \sin(\theta)\cos(\theta)$. L'angle θ ne peut donc prendre que les valeurs telles que $\sin(\theta)\cos(\theta) \geqslant 0$, c'est-à-dire de 0 à $\pi/2$ et puis de π à $3\pi/2$.

En ce qui concerne la surface, nous avons

$$S = 2\int_0^{\pi/2}\int_0^{\sqrt{\sin(\theta)\cos(\theta)}} r\, dr d\theta = 2\int_0^{\pi/2} \frac{1}{2}\sin(\theta)\cos(\theta) d\theta = \frac{1}{2}. \tag{91.373}$$

En ce qui concerne l'intégrale, la fonction que nous devons intégrer étant la même en (x, y) qu'en $(-x, -y)$, nous pouvons simplement intégrer une des deux parties, et multiplier par deux :

$$I = 2\int_0^{\pi/2}\int_0^{\sqrt{\sin(\theta)\cos(\theta)}} r^2\sqrt{\cos(\theta)\sin(\theta)}\, dr d\theta. \tag{91.374}$$

Ne pas oublier qu'il y a un r qui provient du jacobien des coordonnées polaires. Nous avons

$$\begin{aligned} I &= 2\int_0^{\pi} \sqrt{\sin(\theta)\cos(\theta)} \left[\frac{r^3}{3}\right]_0^{\sqrt{\cos(\theta)\sin(\theta)}} \\ &= \frac{2}{3}\int_0^{\pi} \cos^2(\theta)\sin^2(\theta) d\theta. \end{aligned} \tag{91.375}$$

Cette dernière intégrale se fait comme indiqué par la formule (18.90). Nous trouvons

$$I = \frac{2}{3}\left[\frac{\theta}{8} - \frac{4\theta}{32}\right]_0^{\pi} = \frac{\pi}{12}. \tag{91.376}$$

Exercise 265 exoIntegralesMultiples0010

Soient des nombres a, b, c et d tels que $0 < a \leqslant b$ et $0 < c \leqslant d$. Calculer l'aire de $D = \{ax^2 \leqslant y \leqslant bx^2, \frac{c}{x} \leqslant y \leqslant \frac{d}{x}\}$. Indication : poser $u = \frac{y}{x^2}$ et $v = xy$.

corrIntegralesMultiples0010

Correction of the exercise 265

L'intégrale à calculer est

$$\int_D 1 dA.$$

Comme vous pouvez voire dans la figure, la region D n'est pas une region du premier ou du deuxième type. Cependant D jouit d'une certaine régularité parce que son bord est donné par deux morceaux d'hyperbole et deux morceaux de parabole. En fait, la méthode la plus simple pour décrire D est d'introduire des nouvelles variables adaptées à sa geométrie. Nous choisissons $u = y/x^2$ et $v = xy$. On peut alors écrire $D = \{(u, v) : u \in [a, b],\ v \in [c, d]\}$.

La transformation inverse, qui nous donne x et y en fonction de u et v, est $x = (v/u)^{1/3}$ et $y = (uv^2)^{1/3}$. Il faut bien comprendre que ces deux transformations sont définies et de classe $\mathcal{C}^1$ seulement hors de l'origine, d'où l'importance des hypothèses $0 < a \leqslant b$ et $0 < c \leqslant d$. Le jacobien à considérer est

$$J = \det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \frac{1}{3u}. \tag{91.377}$$

Notre intégrale est donc

$$\int_a^b\int_c^d \frac{1}{3u} dv du = \frac{d-c}{3}\ln\left(\frac{b}{a}\right).$$

Exercise 266 exoIntegralesMultiples0011

En utilisant un changement de variables, calculer l'intégrale de f sur D avec

(1) $D = \{(x,y) \in \mathbb{R}^2 \mid \pi^2 < x^2+y^2 \leqslant 4\pi^2\}$ et $f(x,y) = \sin\sqrt{x^2+y^2}$;

(2) $D = \left\{(x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}$ avec $a,b>0$ et $f(x,y) = x^2+y^2$;

(3) $D = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 \leqslant 1\,,\ 0 \leqslant z \leqslant h\}$ avec $h>0$ et $f(x,y,z) = z$;

(4) $D = \{(x,y,z) \in \mathbb{R}^3 \mid 1 \leqslant x^2+y^2+z^2 \leqslant 4\}$ et $f(x,y,z) = (x^2+y^2+z^2)^\alpha$.

(5) $D = \{(x,y) \in \mathbb{R}^2 \mid 0 < x^2 \leqslant y \leqslant 2x^2\,,\ 1/x \leqslant y \leqslant 2/x\}$ et $f(x,y) = x+y$ (changement de variable $u = y/x^2$, $v = xy$) ;

(6) $D = \{(x,y,z) \in \mathbb{R}^3 \mid x \geqslant 0\,,\ y \geqslant 0\,,\ z \geqslant 0\,,\ x^2+y^2+z^2 \leqslant 1\}$ et $f(x,y,z) = xyz$.

corrIntegralesMultiples0011

Correction of the exercise 266

(1) Nous passons aux coordonnées polaires $x = r\cos(\theta)$ et $y = r\sin(\theta)$. Le domaine d'intégration est alors $D = \{(r,\theta) \,:\, r \in [\pi, 2\pi],\ \theta \in [0, 2\pi]\}$. Il s'agit d'une couronne circulaire.

$$\begin{aligned}\int_0^{2\pi}\int_\pi^{2\pi} \sin(r)\, r\, dr d\theta &= \int_0^{2\pi}\left([-r\cos(r)]_\pi^{2\pi} + \int_\pi^{2\pi}\cos(r)\right) d\theta = \\ &= \int_0^{2\pi} -3\pi + [\sin(r)]_\pi^{2\pi}\, d\theta = -6\pi^2.\end{aligned} \tag{91.378}$$

(2) Nous passons aux coordonnées elliptiques $x = ar\cos(\theta)$ et $y = br\sin(\theta)$. La description du domaine est alors très simple $D = [0,1]\times[0,2\pi]$. La fonction à intégrer devient $a^2r^2\cos^2(\theta) + b^2r^2\sin^2(\theta)$. Le jacobien du changement de variable est le détermiant

$$\left|\begin{matrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{matrix}\right| = \left|\begin{matrix}a\cos(\theta) & -ar\sin(\theta)\\ b\sin(\theta) & br\cos(\theta)\end{matrix}\right| = abr. \tag{91.379}$$

On a

$$\begin{aligned}&\int_0^{2\pi}\int_0^1 \left(a^2r^2\cos^2(\theta) + b^2r^2\sin^2(\theta)\right) abr\, dr d\theta = \\ &= ab\frac{1}{4}\int_0^{2\pi}\left(a^2\cos^2(\theta) + b^2\sin^2(\theta)\right) d\theta = \\ &= ab\frac{a^2}{4}\int_0^{2\pi}\cos^2(\theta)d\theta + \frac{b^2}{4}\int_0^{2\pi}\sin^2(\theta)d\theta \\ &= ab\frac{(a^2+b^2)\pi}{4}.\end{aligned} \tag{91.380}$$

(3) Il faut passer aux coordonnées cylindriques $x = r\cos(\theta)$, $y = r\sin(\theta)$ et $z = z$. Le domaine d'intégration est $D = \{(r,\theta,z) \,:\, r \in [0,1], \theta \in [0,2\pi], z \in [0,h]\}$. Le jacobien de cette transformation est r comme pour les coordonnées polaires

$$\begin{aligned}&\int_0^{2\pi}\int_0^1\int_0^h z\, r\, dz dr d\theta = \\ &= \int_0^{2\pi}\int_0^1 \frac{rh^2}{2}\, dr d\theta = \\ &= \frac{\pi h^2}{2}.\end{aligned} \tag{91.381}$$

(4) Le domaine d'intégration est le volume entre deux sphères centrées dans l'origine. Leur rayons respectifs sont 1 et 2. Il est donc naturel de travailler ici en coordonnées sphériques. Le jacobien est $\rho^2\sin(\phi)$ et la fonction à intégrer est $\rho^{2\alpha}$. Remarque : toute puissance de ρ est intégrable dans l'intervalle $[1,2]$. Cependant il faut écrire séparément les cas $\alpha = -3/2$

et $\alpha \neq -3/2$. Si $\alpha \neq -3/2$ alors

$$\begin{aligned}&\int_0^{\pi}\int_0^{2\pi}\int_1^2 \rho^{2(\alpha+1)}\sin(\phi)\,d\rho d\theta d\phi = \\ &= 2\pi\left[\frac{\rho^{2(\alpha+1)+1}}{2(\alpha+1)+1}\right]_1^2\int_0^{\pi}\sin(\phi)d\phi = \\ &= \frac{4\pi}{2(\alpha+1)+1}\left(2^{2(\alpha+1)+1}-1\right).\end{aligned} \tag{91.382}$$

Si $\alpha = -3/2$ alors

$$\int_0^{\pi}\int_0^{2\pi}\int_1^2 \rho^{-1}\sin(\phi)\,d\rho d\theta d\phi = 4\pi\ln(2). \tag{91.383}$$

(5) Voir l'exercice 265.

(6) Le domaine d'intégration est l'huitième de la sphère unitaire contenu dans le premier octant. On peut utiliser des coordonnées sphériques ou cylindriques. Ce corrigé est en coordonnées cylindriques, mais vous pouvez essayer les sphériques pour voir si vous obtenez le même résultat. On décrit le domaine d'intégration de la façon suivante :

$$\begin{aligned}&z \in [0,1]\\ &r \in [0,\sqrt{1-z^2}]\\ &\theta \in [0,\pi/2].\end{aligned} \tag{91.384}$$

Le jacobien est simplement r et l'intégrale devient

$$\begin{aligned}&\int_0^1\int_0^{\pi/2}\int_0^{\sqrt{1-z^2}} r^3 z\cos(\theta)\sin(\theta)\,dr d\theta dz = \\ &= \frac{1}{4}\int_0^1\int_0^{\pi/2}(1-z^2)^2 z\cos(\theta)\sin(\theta)\,d\theta dz = \\ &= \frac{1}{4}\int_0^1(1-z^2)^2 z\,dz\int_0^{\pi/2}\cos(\theta)\sin(\theta)\,d\theta\end{aligned} \tag{91.385}$$

Avec les changements de variables $\tilde{z} = 1-z^2$ et $t = \sin(\theta)$ on obtient enfin

$$\frac{1}{4}\left[\frac{1}{6}\tilde{z}^3\right]_0^1\left[\frac{t^2}{2}\right]_0^1 = \frac{1}{48}. \tag{91.386}$$

Exercise 267 exoIntegralesMultiples0012

Soit $R > 0$, $D_R = \{x^2+y^2 \leqslant R^2, x > 0, y > 0\}$ et $K_R = [0,R]^2$. Montrer que :

$$\iint_{D_R} e^{-(x^2+y^2)}\,dx\,dy \leqslant \iint_{K_R} e^{-(x^2+y^2)}\,dx\,dy \leqslant \iint_{D_{2R}} e^{-(x^2+y^2)}\,dx\,dy. \tag{91.387}$$ gaussian

En déduire l'existence et la valeur de

$$\lim_{R\to+\infty}\int_0^R e^{-t^2}\,dt. \tag{91.388}$$

corrIntegralesMultiples0012

Correction of the exercise 267

La première inégalité est vérifiée parce que la fonction $f(\xi) = e^{\xi}$ est positive pour tout ξ dans $\mathbb{R}$ et D_R a une aire inférieure à l'aire de K_R. Les deux intégrales peuvent être considérées comme le volumes des régions solides entre le plan x-y et le graphe de $e^{-x^2+y^2}$ au dessus de D_R et de K_R respectivement

$$\iint_{D_R}\int_0^{e^{-(x^2+y^2)}} 1\,dz\,dx\,dy \leqslant \iint_{K_R}\int_0^{e^{-(x^2+y^2)}} 1\,dz\,dx\,dy. \tag{91.389}$$

Les deux solides sont égaux au dessus de D_R mais l'un de deux continue à l'extérieur de D_R : son volume sera alors plus importante.

La deuxième inégalité se démontre de la même façon. Il faut juste remarquer que l'exponentielle est une fonction positive et que l'aire de K_R est inférieure à l'aire de D_{2R}.

Nous voulons maintenant résoudre la deuxième partie de l'exercice. Il faut d'abord comprendre le lien entre le trois intégrales dans l'inégalité et l'intégrale

$$\int_0^R e^{-t^2}\,dt.$$

En fait, ce dernier est la racine carrée de l'intégrale au milieu

$$\iint_{K_R} e^{-(x^2+y^2)}\,dx\,dy = \int_0^R e^{-x^2}\,dx \int_0^R e^{-y^2}\,dy = \left(\int_0^R e^{-t^2}\,dt\right)^2. \tag{91.390}$$

Les intégrales sur D_R et D_{2R} sont simples à calculer en coordonnées polaires

$$\iint_{D_R} e^{-(x^2+y^2)}\,dx\,dy = \int_0^{2\pi}\int_0^R e^{-r^2} r\,dr\,d\theta = \pi[e^{-r^2}]_0^R. \tag{91.391}$$

Comme les inégalités (91.387) sont valides pour tout R la limite qu'on veut calculer est coincée entre deux limites identiques

$$\lim_{R\to+\infty}\left(\pi(e^{-R^2}-1)\right)^{1/2} \leqslant \lim_{R\to+\infty}\int_0^R e^{-t^2}\,dt \leqslant \lim_{R\to+\infty}\left(\pi(e^{-4R^2}-1)\right)^{1/2}. \tag{91.392}$$

Nous obtenons alors

$$\lim_{R\to+\infty}\int_0^R e^{-t^2}\,dt = \sqrt{\pi}. \tag{91.393}$$

Exercise 268 exoIntegralesMultiples0013

Soient $a, R > 0$. Dans le plan (yOz), soit D le disque de centre $(0, a, 0)$ et de rayon a. En tournant autour de l'axe (Oz), le disque D engendre un domaine T (appelé un **tore** plein). Calculer le volume de T (c'est-à-dire l'intégrale triple $\iiint_T dx\,dy\,dz$). corrIntegralesMultiples0013

Correction of the exercise 268

On peut décrire ce solide en coordonnées sphériques. Considérons d'abord la section de T contenue dans le plan y-z, pour $y > 0$. Il s'agit du cercle D de rayon a centré au point $(0, a, 0)$. Les points de D correspondent aux points (ρ, θ, ϕ) tels que

$$\begin{array}{l}\theta = \frac{\pi}{2},\\ \phi \in [0, \pi]\\ \rho \in [0, 2a\sin(\phi)]\,.\end{array} \tag{91.394}$$

On obtient tous les autres points dans T en faisant tourner D autour de l'axe vertical : cela correspond à faire varier θ dans l'intervalle $[0, 2\pi]$.

L'intégrale à calculer est alors

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{2a\sin(\phi)} \rho^2\sin(\phi)\,d\rho\,d\phi\,d\theta = \frac{16a^3\pi}{3}\int_0^{\pi}\sin^4(\phi)\,d\phi = 2a^3\pi^2. \tag{91.395}$$

Exercise 269 exoGeomAnal-0050

Soit C_1 le disque de rayon 2 centré en $(-1, 0)$ et C_2 le disque de rayon 2 centré en $(1, 0)$. Calculer l'aire de l'intersection entre C_1 et C_2.

Indice : faire un dessin et exploiter les symétries. corrGeomAnal-0050

Correction of the exercise 269

L'intégrale à calculer est

$$\int_0^1 \sqrt{4-(x+1)^2}. \tag{91.396}$$

Nous pouvons faire le changement de variables $u = x+1$ puis $u = 2\cos\theta$. Ce faisant nous devons calculer

$$4\int_0^{\pi/3} \sin^2(\theta)d\theta = \frac{2\pi}{3} - \frac{\sqrt{3}}{2}. \tag{91.397}$$

Exercise 270 exoIntegralesMultiples0014

Calculer les intégrales suivantes à l'aide de la formule d'intégration par parties

$$\int_0^{+\infty} e^{-y}\sin(3y)dy \quad \text{et} \quad \int_0^{+\infty} e^{-y}\cos(3y)dy. \tag{91.398}$$

En utilisant la périodicité des fonctions trigonométriques et un changement de variable, en déduire les valeurs des intégrales

$$\int_{2\pi}^{+\infty} e^{-y}\sin(3y)dy \quad \text{et} \quad \int_{2\pi}^{+\infty} e^{-y}\cos(3y)dy. \tag{91.399}$$

corrIntegralesMultiples0014

Correction of the exercise 270

<+CorrIntegralesMultiples0014+>

Exercise 271 exoIntegralesMultiples0015

On note par $(I_m)_{m\in\mathbb{N}}$ les intégrales $I_m = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^m e^{-\frac{x^2}{2}}dx$, $m \in \mathbb{N}$.

(1) Que peut-on dire des valeurs I_m avec m impair ?

(2) Exprimer I_{2m+2} en fonction de I_{2m}.

(3) En déduire une expression pour I_{2m} en fonction de m.

corrIntegralesMultiples0015

Correction of the exercise 271

<+CorrIntegralesMultiples0015+>

Exercise 272 exoIntegralesMultiples0016

Soit C le cercle de rayon 1 et de centre $(1,0)$. Calculer l'aire de

$$C \cap \{(x,y) \in \mathbb{R}^2 \text{ tel que } y > 1/2\}. \tag{91.400}$$

corrIntegralesMultiples0016

Correction of the exercise 272

<+CorrIntegralesMultiples0016+>

Exercise 273 exoIntegralesMultiples0017

<+ExoIntegralesMultiples0017+>

corrIntegralesMultiples0017

Correction of the exercise 273

<+CorrIntegralesMultiples0017+>

Exercise 274 exoIntegralesMultiples0018

<+ExoIntegralesMultiples0018+>

corrIntegralesMultiples0018

Correction of the exercise 274

<+CorrIntegralesMultiples0018+>

Exercise 275 exoIntegralesMultiples0019

<+ExoIntegralesMultiples0019+>

corrIntegralesMultiples0019

Correction of the exercise 275

<+CorrIntegralesMultiples0019+>

Exercise 276 exoIntegralesMultiples0020

<+ExoIntegralesMultiples0020+> corrIntegralesMultiples0020

Correction of the exercise 276

<+CorrIntegralesMultiples0020+>

Exercise 277 exoIntegralesMultiples0021

<+ExoIntegralesMultiples0021+> corrIntegralesMultiples0021

Correction of the exercise 277

<+CorrIntegralesMultiples0021+>

Exercise 278 exoIntegralesMultiples0022

<+ExoIntegralesMultiples0022+> corrIntegralesMultiples0022

Correction of the exercise 278

<+CorrIntegralesMultiples0022+>

Exercise 279 exoIntegralesMultiples0023

<+ExoIntegralesMultiples0023+> corrIntegralesMultiples0023

Correction of the exercise 279

<+CorrIntegralesMultiples0023+>

Exercise 280 exoIntegralesMultiples0024

<+ExoIntegralesMultiples0024+> corrIntegralesMultiples0024

Correction of the exercise 280

<+CorrIntegralesMultiples0024+>

Exercise 281 exoIntegralesMultiples0025

<+ExoIntegralesMultiples0025+> corrIntegralesMultiples0025

Correction of the exercise 281

<+CorrIntegralesMultiples0025+>

Exercise 282 exoIntegralesMultiples0026

<+ExoIntegralesMultiples0026+> corrIntegralesMultiples0026

Correction of the exercise 282

<+CorrIntegralesMultiples0026+>

Exercise 283 exoIntegralesMultiples0027

<+ExoIntegralesMultiples0027+> corrIntegralesMultiples0027

Correction of the exercise 283

<+CorrIntegralesMultiples0027+>

Exercise 284 exoIntegralesMultiples0028

<+ExoIntegralesMultiples0028+> corrIntegralesMultiples0028

Correction of the exercise 284

<+CorrIntegralesMultiples0028+>

Exercise 285 exoIntegralesMultiples0029

<+ExoIntegralesMultiples0029+> corrIntegralesMultiples0029

Correction of the exercise 285

<+CorrIntegralesMultiples0029+>

Exercise 286 exoIntegralesMultiples0030

<+ExoIntegralesMultiples0030+> corrIntegralesMultiples0030

Correction of the exercise 286

<+CorrIntegralesMultiples0030+>

91.6 Autres exercices

Cette section contient entre autres des exercices donnés à des devoirs, interrogations et DS.

Exercise 287 exodevoir1-0004

(1) Trouver les domaines de définition des fonctions suivantes :

(1a) $f_1(x,y) = \ln(x^2+y^2-2)$;

(1b) $f_2(x,y) = \dfrac{x^2-y^2}{x-y}$;

(1c) $f_3(x,y,z) = \dfrac{1}{\sqrt{z^2-x^2-y^2}}$.

(2) Dessiner les graphes et quelques courbes de niveau des fonctions suivantes :

(2a) $g_1(x,y) = x+y$;

(2b) $g_2(x,y) = 1-x^2-y^2$;

(2c) $g_3(x,y) = x-y^2$.

corrdevoir1-0004

Correction of the exercise 287

(1) (1a) La fonction f_1 est une fonction composée. Soit $g(x,y) = x^2+y^2-2$ et $h(\xi) = \ln(\xi)$, alors $f_1 = h \circ g(x,y)$. La fonction g est définie partout sur $\mathbb{R}^2$, et le domaine du logarithme est $\{\xi \in \mathbb{R} : \xi > 0\}$. Le domaine de f_1 est alors l'ensemble des points (x,y) tels que $g(x,y) > 0$, c'est-à-dire le complémentaire du disque fermé de rayon $\sqrt{2}$ centré dans l'origine.

(1b) La fonction f_2 est le rapport entre deux polynômes. Les fonctions polynomiales sont définies partout, mais le rapport est bien défini seulement où le dénominateur est non nul. Le domaine de f_2 est donc l'ensemble $\mathbb{R}^2 \backslash x = y$.

(1c) La racine carrée est une fonction définie sur l'ensemble des nombres positifs et en zéro. Ici, comme la racine est au dénominateur, la valeur zéro doit être exclue. Le domaine de définition de f_3 est alors donné par la relation $\{(x,y,z) : z^2-x^2-y^2 > 0\}$. Il s'agit de l'intérieur d'un cône.

(2) (2a) $g_1(x,y) = x+y$, 91.19(a), 91.16. ;

(2b) $g_2(x,y) = 1-x^2-y^2$, 91.19(b), 91.17. ;

(2c) $g_3(x,y) = x-y^2$, 91.19(c), 91.18.

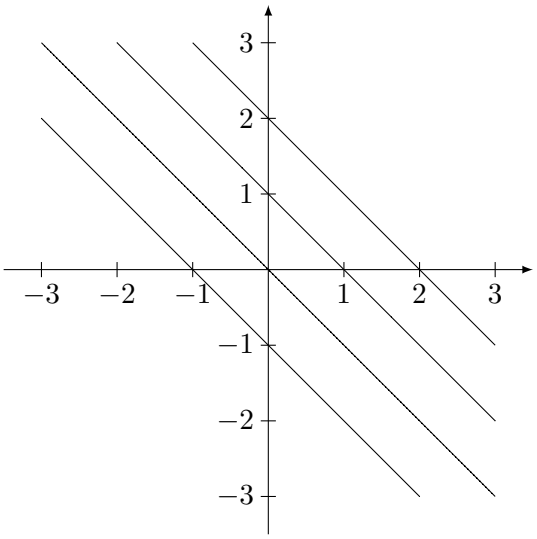


FIGURE 91.16: LabelFigMCQueGF

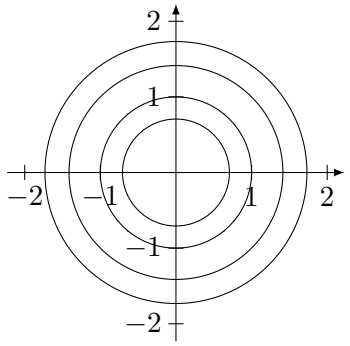


FIGURE 91.17: LabelFigPHTVjfk

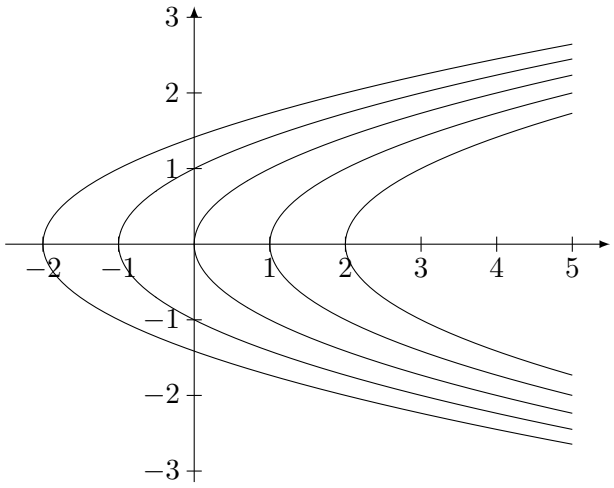


FIGURE 91.18: LabelFigIWuPxFc

Exercise 288 exodevoir1-0005

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction

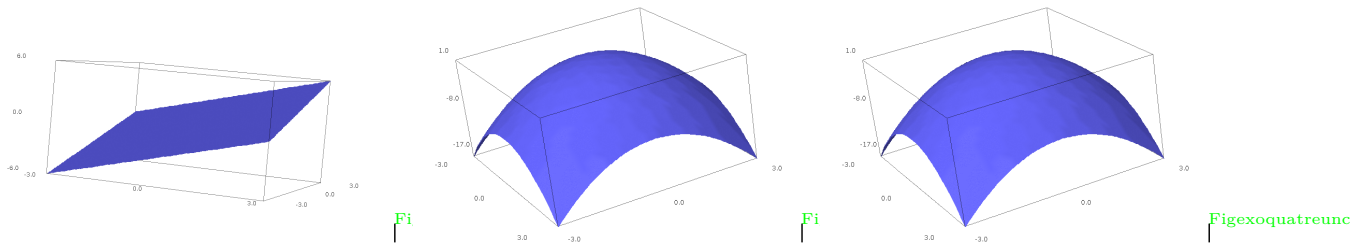
$$f(x,y) = \begin{cases} (x^3+y)\sin\left(\frac{x}{x^2+y^2}\right), & \text{si } (x,y) \neq (0,0) \\ 0, & \textit{sinon.} \end{cases} \tag{91.401}$$

Est-elle une fonction continue ? corrdevoir1-0005

Correction of the exercise 288

La fonction f est continue. Il suffit d'observer que la fonction sinus est bornée et écrire

$$-|x^3+y| \leqslant \left|(x^3+y)\sin\left(\frac{x}{x^2+y^2}\right)\right| \leqslant |x^3+y|. \tag{91.402}$$



Fi Fi Figexoquatreunc

(a) (b) (c)

FIGURE 91.19: Des dessins pour l'exercice 287

La limite de $|x^3 + y|$ lorsque $(x, y) \to (0, 0)$ est zéro (pour le voir on peut passer en coordonnées polaires) et donc, par la règle de l'étau, la limite de $f(x, y)$ en l'origine est zéro.

Exercise 289 exodevoir1-0006

Soit $f : \mathbb{R}^2\backslash\{(0, 0)\} \to \mathbb{R}$ la fonction

$$f(x, y) = \frac{x^3 y}{x^6 + 3y^2}. \tag{91.403}$$

(1) Montrer que $\lim_{x\to 0} \lim_{y\to 0} f(x, y) = \lim_{y\to 0} \lim_{x\to 0} f(x, y) = 0$.

(2) Montrer que $\lim_{t\to 0} f(t, at) = 0$, pour toute constante $a \in \mathbb{R}$.

(3) Montrer que la fonction f n'admet pas de limites en $(0, 0)$.

corrdevoir1-0006

Correction of the exercise 289

(1) La limite double $\lim_{y\to 0} \lim_{x\to 0} f(x, y)$ est nulle parce que

$$\lim_{x\to 0} f(x, y) = \lim_{x\to 0} \frac{x^3 y}{x^6 + 3y^2} = 0, \tag{91.404}$$

et $\lim_{y\to 0} 0 = 0$. De même $\lim_{x\to 0} \lim_{y\to 0} f(x, y)$ est zéro parce que

$$\lim_{y\to 0} f(x, y) = \lim_{y\to 0} \frac{x^3 y}{x^6 + 3y^2} = 0, \tag{91.405}$$

et $\lim_{x\to 0} 0 = 0$.

(2)

$$\lim_{t\to 0} f(t, at) = \lim_{t\to 0} \frac{at^4}{t^6 + 3a^2t^2}, \tag{91.406}$$

le terme dominant au dénominateur lorsque $t \to 0$ est $3a^2t^2$, qui tend vers zéro moins vite que t^4. La valeur de la limite est donc zéro.

(3) Pour montrer cela on calcule la limite le long de la courbe $t \mapsto (t, t^3)$:

$$\lim_{t\to 0} f(t, t^3) = 1/4 \neq 0. \tag{91.407}$$

L'existence de la limite serait en contradiction avec le théorème d'unicité de la limite. Par conséquent la limite n'existe pas.

Exercise 290 exodevoir2-0002

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $f(x, y) = x \cos(y)$. Calculer

(1) les dérivées partielles de f ;

(2) la dérivée de f suivant le vecteur $\vec{v} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ au point $(1, \pi)$.

Écrire l'équation du plan tangent au graphe de f au point $(1, \pi, -1)$.

[Bonus :] Trouver le vecteur normal à ce plan. corrdevoir2-0002

Correction of the exercise 290

(1) Les dérivées partielles de f sont :

$$\partial_x f(x, y) = \cos(y), \qquad \partial_y f(x, y) = -x \sin(y).$$

(2) $\nabla f(x, y) \cdot v = \frac{\cos(y)}{\sqrt{2}} + \frac{x \sin(y)}{\sqrt{2}}$. La valeur de cette fonction au point $(1, \pi)$ est $-\frac{1}{\sqrt{2}}$.

(3) Comme on a vu à la section 4.4 du poly le plan tangent au graphe de f au point $(1,\pi)$ est le graphe de la fonction $T_{(1,\pi)}$

$$T_{(1,\pi)}(x,y) = f(1,\pi) + \nabla f(1,\pi) \cdot (x-1, y-\pi) = -1 - (x-1) = -x.$$

Cela veut dire que le plan tangent est l'ensemble $\mathcal{T} \subset \mathbb{R}^3$,

$$\mathcal{T} = \{(x,y,z) \in \mathbb{R}^3 \,|\, z = -x\},$$

L'équation $x + z = 0$ est l'équation du plan $\mathcal{T}$.

Le vecteur normal au plan $\mathcal{T}$ un vecteur de norme 1 qui est orthogonal à tous les vecteurs contenus dans $\mathcal{T}$. Dans cet exemple le vecteur normal est $\vec{n} := \frac{1}{\sqrt{2}}(1,0,1)$, parce que l'équation du plan nous dit que si le point $P := (x,y,z)$ appartient à $\mathcal{T}$ alors $P \cdot \vec{n} = 0$. En général, le vecteur normal est le vecteur qui a pour composantes les coefficients de x, y et z dans l'équation du plan. Pour comprendre comment ça arrive il faut un peu réfléchir sur comment on définit un plan.

Soit $P = (p_1,p_2,p_3)$ un point du plan et $\vec{n} = (n_1,n_2,n_3)$ le vecteur normal au plan. Si $Q = (x,y,z)$ est un autre point du plan alors le vecteur $\vec{PQ}$ est dans le plan. On a alors $(Q-P) \cdot \vec{n} = 0$, c'est-à-dire $n_1x + n_2y + n_3z = n_1p_1 + n_2p_2 + n_3p_3$. Le terme de droite est un nombre réel et il nous donne la distance entre le plan et l'origine. À gauche on voit que les coefficients de x, y et z sont bien les composantes de $\vec{n}$. Voici un exercice supplémentaire (facultatif!!) sur les plans.

Exercice : C'est connu depuis l'antiquité que pour trouver un plan il nous suffit une des combinaisons suivantes :

— trois points du plan ;

— un point du plan et une droite contenue dans le plan qui ne passe pas par le point ;

— deux droites parallèles dans le plan ;

— deux droites qui ont un point en commun dans le plan.

On vient de dire que pour trouver un plan il nous faut un point du plan P et un vecteur normal au plan $\vec{n}$. Démontrer que chacune des combinaisons proposées est équivalente à la couple «point du plan + vecteur normal». Pour la preuve on peut utiliser la proposition B.2 du poly.

Exercise 291 exodevoir2-0003

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable en l'origine. On sait que $f(x,x) = 2x$, et $f(x,-x) = 0$, pour tout $x \in \mathbb{R}$. Trouver $\nabla f(0,0)$. corrdevoir2-0003

Correction of the exercise 291

Soient g_1 et g_2 deux fonctions de $\mathbb{R} \to \mathbb{R}^2$, $g_1(x) = (x,x)$ et $g_2(x) = (x,-x)$.

On a alors que $f(x,x) = f \circ g_1(x)$ et $f(x,-x) = f \circ g_2(x)$. On peut calculer les dérivées de ces deux fonctions scalaires

$$\frac{d(f\circ g_1)}{dx}(0) = 2, \qquad \frac{d(f\circ g_2)}{dx}(0) = 0.$$

D'autre part la définition de dérivée d'une fonction composée nous dit que

$$\frac{d(f\circ g_1)}{dx}(0) = \nabla f(0,0) \cdot \frac{dg_1}{dx}(0),$$

et

$$\frac{d(f\circ g_2)}{dx}(0) = \nabla f(0,0) \cdot \frac{dg_2}{dx}(0).$$

On obtient alors le système de deux équations

$$\begin{cases} \partial_x f(0,0) + \partial_y f(0,0) = 2 \\ \partial_x f(0,0) - \partial_y f(0,0) = 0. \end{cases} \tag{91.408}$$

La seule solutions de ce système est $\partial_x f(0,0) = \partial_y f(0,0) = 1$.

Cet exercice est très difficile si on ne connaît pas la méthode de résolution. Cependant, la méthode n'est pas trop compliquée une fois qu'on a vu la solution et peut s'avérer utile en plusieurs situations.

Exercise 292 exodevoir2-0004

Soient $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(s,t) = (\cos s, \sin t, s+t)$, et $g : \mathbb{R}^3 \to \mathbb{R}^2$, $g(x,y,z) = (x^2+y^2, z^2x)$.

Trouver la matrice jacobienne $F = g \circ f$ au point $(\pi, 3)$. corrdevoir2-0004

Correction of the exercise 292

Il y a deux façons de résoudre cet exercice. La première consiste à trouver d'abord une formule explicite pour F, l'utiliser pour calculer le gradient $\nabla F(x,y)$ et enfin évaluer cette fonction au point $(\pi, 3)$. La deuxième consiste à appliquer ce qu'on a étudié à la section 4.3.3 du poly sur la différentielle des fonctions composées. Il nous faut alors calculer les matrices $\nabla g(f(\pi,3))$ et $\nabla f(\pi,3)$ et les multiplier entre elles **dans cet ordre**.

Le résultat est

$$\nabla F(\pi,3) = \begin{pmatrix} 0 & 2\cos(3)\sin(3) \\ -2(\pi+3) & -2(\pi+3) \end{pmatrix}. \tag{91.409}$$

Exercise 293 exodevoir2-0005

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \begin{cases} \frac{xy}{|x|+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

(1) Calculer les dérivées partielles de f en $(0,0)$.

(2) Montrer que f n'est pas différentiable. corrdevoir2-0005

Correction of the exercise 293

(1) Les dérivées partielles de f en l'origine sont

$$\partial_x f(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h(|h|+0)} = 0, \tag{91.410}$$

et

$$\partial_y f(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h(0+h^2)} = 0. \tag{91.411}$$

(2) Si f est différentiable en l'origine alors sa fonction différentielle est nulle. Cela veut dire que f est différentiable en l'origine si et seulement si

$$\lim_{(x,y)\to(0,0)} \frac{|f(x,y)|}{\sqrt{x^2+y^2}} = 0.$$

En fait, il est facile de voir que cette limite n'existe pas. Pour le démontrer nous allons utiliser la méthode des chemins. Soit λ un nombre réel.

$$\lim_{t\to 0} \frac{|f(t,\lambda t)|}{|t|\sqrt{1+\lambda}} = \lim_{t\to 0} \frac{|\lambda t^2|}{||t| + \lambda^2 t^2||t|\sqrt{1+\lambda}}. \tag{91.412}$$

Le terme dominant dans la limite est $\frac{|\lambda t^2|}{t^2\sqrt{1+\lambda}} = \frac{|\lambda|}{\sqrt{1+\lambda}}$. Comme cette quantité dépends de λ la limite n'existe pas et la fonction f n'est pas différentiable en l'origine.

Exercise 294 exodevoir2-0006

Trouver, si elle existe, une fonction $g : \mathbb{R}^3 \to \mathbb{R}$ telle que

$$\begin{aligned} \partial_x g(x,y,z) &= 3x^2 y\sin(yz) + \frac{3y}{x}, \\ \partial_y g(x,y,z) &= x^3\sin(yz) + x^3 yz\cos(yz) + 3\ln(x), \\ \partial_z g(x,y,z) &= x^3y^2\cos(yz) + 6z. \end{aligned} \tag{91.413}$$

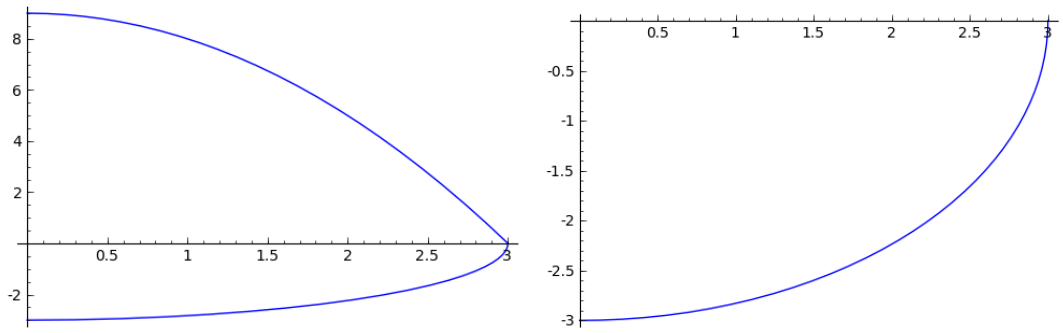


FIGURE 91.20: Les projections sur le plans z-x et x-y de la région décrite dans l'exercice 6.1 (a) exodevoir2

corrdevoir2-0006

Correction of the exercise 294

On commence par intégrer par rapport à la variable x la dérivée partielle de g par rapport à x. Cette procédure nous permet de trouver la fonction f à moins d'une fonction K_1 qui ne dépend pas de x et joue le rôle de «constante d'intégration». Nous avons

$$f(x,y,z) = \int \partial_x g(x,y,z)\,dx = x^3 y\sin(yz) + 3y\ln(x) + K_1(y,z). \tag{91.414}$$ parrapportax

De même, en intégrant par rapport à y et z les dérivées partielles correspondantes on obtient

$$f(x,y,z) = \int \partial_y g(x,y,z)\,dy = x^3 y\sin(yz) + 3y\ln(x) + K_2(x,z), \tag{91.415}$$ parrapportay

$$f(x,y,z) = \int \partial_z g(x,y,z)\,dz = x^3 y\sin(yz) + 3z^2 + K_3(x,y). \tag{91.416}$$ parrapportaz

Pour compléter notre exercice il nous faut maintenant trouver des fonctions K_1, K_2 et K_3 telles que les équations (91.414), (91.415) et (91.416) soient vérifiées simultanément.

On a alors $K_1 = K_2 = 3z^2$ et $K_3 = 3y\ln(x)$.

Exercise 295 exodevoir2-0007

(1) Traduire les équations suivantes du système de coordonnées donné (cartésiennes, cylindriques, sphériques) vers chacun des deux autres. Esquisser les volumes ou les surfaces correspondants.

   (a) (Cartésiennes) $-\sqrt{9-x^2-y^2} \leqslant z \leqslant 9-x^2-y^2$, $x>0$, $y<0$;

   (b) (Sphériques) $\cos\theta\sin\phi + \cos\phi = 0$.

(2) Traduire les équations suivantes, données en coordonnées cylindriques, en coordonnées cartésiennes. Dessiner le volume correspondant.

$$-1 \leqslant z \leqslant 1, \qquad \sqrt{1-z^2} \leqslant r \leqslant 1.$$

corrdevoir2-0007

Correction of the exercise 295

(1) (a) La région à décrire ressemble à un quart d'œuf. En coordonnées cylindriques $\theta \in [-\pi/2, 0]$ à cause des inégalités $x>0$ et $y<0$ et $z \in [-\sqrt{9-r^2}, 9-r^2]$. L'intersection entre la boule de rayon 3 centrée en l'origine et le paraboloïde elliptique $z = 9-r^2$ est le disque de rayon 3, centre $(0,0,0)$ sur le plan x-y. Dans la figure, il faut faire attention parce que les échelles des deux axes ne sont pas les mêmes.

Les équations en coordonnées sphériques sont $\theta \in [-\pi/2, 0]$, $\rho\cos(\phi) \leqslant 9 - \rho^2\sin^2(\phi)$, pour ϕ dans l'intervalle $[0, \pi/2]$ et $\rho \leqslant 3$ pour ϕ dans l'intervalle $[\pi/2, \pi]$.

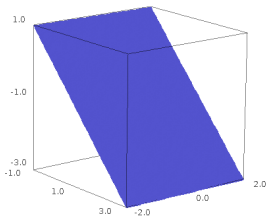


FIGURE 91.21: La surface décrite dans l'exercice 6.1.(b) exo6devoir2un

(b) Il suffit de multiplier l'équation par ρ pour comprendre que il s'agit de l'équation du plan $x+z=0$. En coordonnées cylindriques on peut l'écrire comme $r\cos(\theta)+z=0$.

(2) Les inéquations en coordonnées cylindriques sont

$$\begin{cases} -1 \leqslant z \leqslant 1, \\ \sqrt{1-z^2} \leqslant \sqrt{x^2+y^2}, \\ \sqrt{x^2+y^2} \leqslant 1. \end{cases} \tag{91.417}$$

On peut enlever les racines carrées dans la deuxième et la troisième. On voit alors que la région à décrire est bornée entre la boule de centre $(0,0,0)$ et rayon 1 et le cylindre de base $x^2+y^2 \leqslant 1$, symétrique par rapport à l'axe des z centré en l'origine et de hauteur totale 2.

Exercise 296 exodevoir2-0008

Soit S la partie de $\mathbb{R}^2$ bornée par les courbes $xy=1$, $xy=2$, $y=3$ et $y=4$. Évaluer l'intégrale

$$\int\int_S \sin(xy)\,dA,$$

à l'aide d'un changement de variables. corrdevoir2-0008

Correction of the exercise 296

On introduit les nouvelles variables $s=xy$ et $t=y$. Le domaine d'intégration est un rectangle dans le plan s-t : $S=[1,2]\times[3,4]$. La fonction à intégrer devient une fonction de la seule variable s : $\sin(xy)=\sin(s)$. Pour compléter l'intégrale il faut encore calculer le déterminant de la matrice Jacobienne de la transformation $(s,t)\mapsto(x,y)$. On écrit $x=s/t$ et $y=t$, donc la matrice Jacobienne est

$$\begin{pmatrix} \frac{1}{t} & -\frac{s}{t^2} \\ 0 & 1 \end{pmatrix}, \tag{91.418}$$

et son déterminant est $1/t$.

On a alors

$$\int_1^2\int_3^4 \frac{\sin(s)}{t}\,dt\,ds = \int_1^2 \sin(s)\,ds \int_3^4 \frac{1}{t}dt = (-\cos(2)+\cos(1))\,(\ln(4)-\ln(3))\,. \tag{91.419}$$

Exercise 297 exodevoir2-0009

Calculer les intégrales suivantes :

(1) le volume de la région solide bornée inférieurement par la sphère $x^2+y^2+z^2=4$ et supérieurement par $x^2+y^2+z^2=4z$. Conseil : utiliser les coordonnées sphériques.

(2) $\int\int\int_E x\,dV$, où E est la région bornée par les plans $x+y-z=0$, $z=2$, $x=0$, $y=0$.

corrdevoir2-0009

Correction of the exercise 297

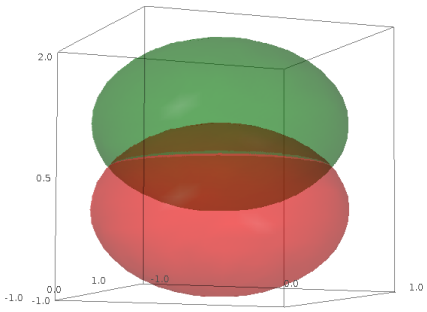


FIGURE 91.22: Les deux sphères

exo8devoir2

(1) On voit sur le dessin que l'intersection entre les deux sphères $x^2+y^2+z^2=4$ et $x^2+y^2+(z-2)^2=4$ est un cercle. Pour trouver ses équations on passe d'abord en coordonnées sphériques : $x^2+y^2+z^2=4$ devient $\rho=2$ et $x^2+y^2+(z-2)^2=4$ devient $\rho=4\cos(\phi)$. L'intersection est donnée par le sicles de rayon 2 et de $\phi=\arccos(1/2)=\pi/3$. La région dont on veut calculer le volume est alors donnée par $\phi\in[0,\pi/3]$, $\rho\in[2,4\cos(\phi)]$ et $\theta\in[0,2\pi]$. Le volume est donné par la valeur de l'intégrale suivante

$$\begin{aligned}\int_0^{2\pi}\int_0^{\pi/3}\int_2^{4\cos(\phi)} 1\,\rho^2\sin(\phi)\,d\rho\,d\phi\,d\theta &= 2\pi\int_0^{\pi/3}\frac{(4\cos(\phi))^3-2^3}{3}\sin(\phi)\,d\phi = \\ &= \frac{2\pi}{3}4^2\left(-\cos^4(\pi/3)+\cos^4(0)\right)+\frac{2\pi}{3}2^3\left(\cos(\pi/3)-\cos(0)\right) = \\ &= \frac{2\pi}{3}(-1+16-4)=\frac{22\pi}{3}.\end{aligned} \tag{91.420}$$

(2) Le domaine d'intégration E est une pyramide renversée. Le rôle joué par les trois variables est symétrique. Les côtés de la pyramide sont des morceaux de droite entre les sommets $(0,0,0)$, $(0,0,2)$ $(0,2,2)$ et $(2,0,2)$. La pyramide est dont contenue dans le premier octant.
On peut choisir librement une variable de départ : ici je choisis z. On voit tout de suite que z est dans l'intervalle $[0,2]$. Ensuite, je considère z fixée et je dois décrire le triangle qui est la section de la pyramide à la hauteur z. Les sommet de cet triangle sont donnés par $(0,0,z)$, $(0,z,z)$ et $(z,0,z)$. Encore une fois on peut choisir la variable qu'on on veut traiter en première. Je choisis y, elle parcourt l'intervalle $[0,z]$. Soient maintenant fixées z et y : la variable x parcourt un petit morceau de droite entre 0 et $z-y$. Maintenant nous écrivons l'intégrale : comme l'intervalle d'intégration pour x dépend de z et de y il faut intégrer d'abord par rapport à x. Ensuite, comme z apparaît dans la borne de y il faut intégrer par rapport à y. La variable dont les bornes sont simplement des nombres réels est toujours la dernière.

$$\int_0^2\int_0^z\int_0^{z-y} x\,dx\,dy\,dz=\frac{1}{2}\int_0^2\int_0^z(z-y)^2\,dy\,dz=\frac{1}{6}\int_0^2 z^3\,dz=\frac{1}{24}16=\frac{2}{3}. \tag{91.421}$$

Exercise 298 exoCourbesSurfaces0004

Tracer la courbe paramétrique définie par :

$$\begin{cases} x(t)=t^2+2t & \text{(91.422a)}\\ y(t)=\dfrac{1+2t}{t^2}. & \text{(91.422b)}\end{cases}$$

corrCourbesSurfaces0004

Correction of the exercise 298

La fonction x est définie pour toute valeur de t, mais y n'est pas définie en $t=0$. Notre courbe sera alors l'union de deux branches, la première définie pour $t>0$ et l'autre pour $t<0$. On

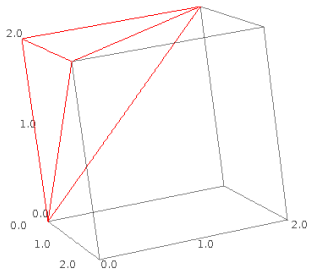


FIGURE 91.23: Le domaine E exo8devoir2deuxieme

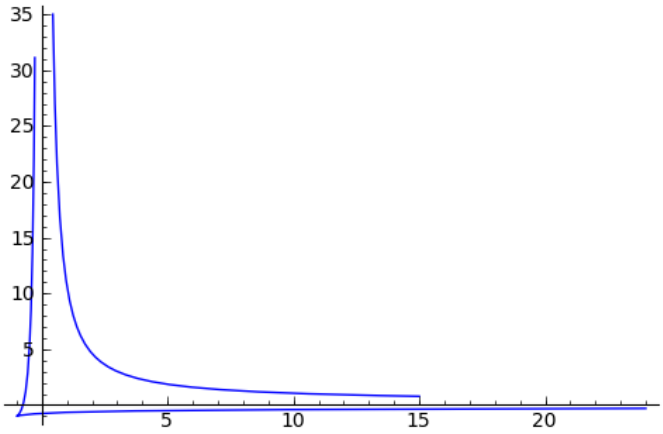


FIGURE 91.24: La courbe de l'exercice 298 figdevoir3exo1

remarque tout de suite que les limites de x et y lorsque $t \to 0^{\pm}$ sont, respectivement, $0^{\pm}$ et $+\infty$ (la courbe ressemble à $(2t, 1/t^2)$). Par contre, lorsque $t \to \pm\infty$ on a $x \to +\infty$ et $y \to 0^{\pm}$ (la courbe ressemble à $(t^2, 2/t)$).

Ni x ni y ont des symétries remarquables.

On cherche tout de suite les intersections avec les axes : $x(t) = 0$ pour $t = 0$ et $t = -2$. Comme $t = 0$ ne fait pas partie du domaine la seule intersection avec l'axe verticale est $(0, y(-2) = -3/4)$. Le seul zéro de y est $t = -1/2$, l'intersection avec l'axe horizontale correspondante est $(x(-1/2) = -3/4, 0)$.

Les dérivées de x et de y sont $\dot{x}(t) = 2t + 2$ et $\dot{y}(t) = -2/t^3 - 2/t^2$. Le point $t = -1$ est donc un point critique pour les deux fonctions. Plus précisément, il s'agit d'un minimum pour x et y : $x(-1) = -1$, $y(-1) = -1$.

Il faut bien comprendre qu'il s'agit d'un minimum pour la branche qui correspond à $t < 0$ mais que pour $t > 0$ y pourrait bien prendre des valeurs encore plus petits. Pourriez-vous dire pourquoi cela ne peut pas arriver à x ? On peut voir un exemple de ce comportement dans le deuxième exercice de ce devoir. Nous pouvons observer que pour $t > 0$ x est strictement croissante et y strictement décroissante.

Les deux branches de cette courbe sont tracées dans la figure 91.24.

Exercise 299 [exodevoir3-0002

Trouver le domaine de définition des fonctions $x(t)$ et $y(t)$, et tracer la courbe

$$x(t) = t^2 + \frac{2}{t}, \qquad y(t) = t + \frac{1}{t}.$$

corrdevoir3-0002

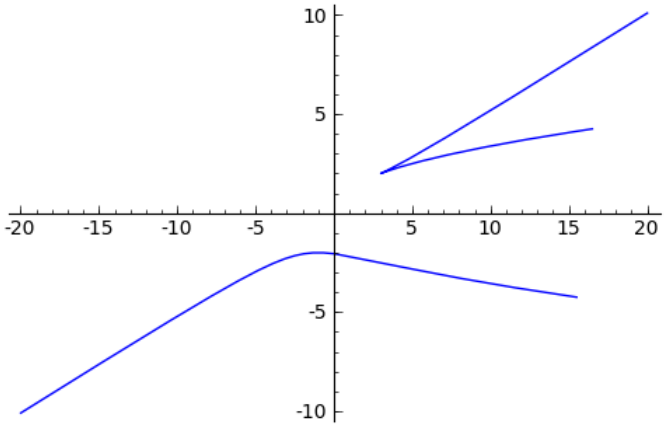


FIGURE 91.25: La courbe de l'exercice 299 figdevoir3exo2

Correction of the exercise 299

Le domaine de définition de x et y est $\mathbb{R}\backslash\{0\}$. On aura donc une courbe en deux morceaux. Le comportement de $(x(t), y(t))$ au voisinage de $t = 0^{\pm}$ est proche au comportement de la droite $(2z, z)$ pour $z \to \pm\infty$ et que lorsque $t \to \pm\infty$ la courbe rassemble à la parabole (t^2, t). On aura donc l'impression que les deux branches soient symétriques par rapport à l'axe horizontale lorsque $|t|$ est grand.

On remarque que la fonction y est impaire $y(-t) = -y(t)$. La fonction x n'as pas de symétries évidentes.

On cherche les intersections avec les axes : $x(t) = 0$ pour $t = -2^{1/3}$, mais $y(t) = 0$ n'a pas de solutions. La seule intersection de la courbe $(x(t), y(t))$ avec les axes est le point $(0, y(-2^{1/3}))$.

On voit que $\dot{x}(t) = 2t - 2/t^2$ est nulle au point $t = 1$, positive si $t > 1$ et négative sinon. La valeur $x(1) = 3$ est donc la plus petite valeur prise par x dans dans la branche de $t > 0$. Dans l'autre branche x est une fonction décroissante de t.

Pour ce qui concerne la dérivée $\dot{y}(t) = 1 - 1/t^2$, est nulle aux points $t = \pm 1$, positive si $|t| > 1$ et négative sinon. La dérivée seconde de y est $+2/t^3$, cela veut dire que 1 est un point de minimum de y et -1 est un point de maximum.

Les deux branches de cette courbe sont tracées dans la figure 91.25.

Exercise 300 exoCourbesSurfaces0009

Considérer l'ellipse d'équation paramétrique $\gamma(t) = (a\cos t, b\sin t)$, où $a,\ b > 0$ et $t \in [0, 2\pi]$. Calculer l'abscisse curviligne par rapport au point $A = (a, 0)$, puis par rapport au point $B = (0, b)$. corrCourbesSurfaces0009

Correction of the exercise 300

Nous cherchons le point $\sigma(s)$ tel que la longueur d'arc entre $\sigma(s)$ et $A = (a, 0)$ soit égale à s. Si X et Y sont des points de l'ellipse, nous allons noter

$$d(X, Y) \tag{91.423}$$

la longueur d'arc entre X et Y, parcourue dans le sens trigonométrique.

La fonction qui joue un rôle fondamental dans cet exercice est la fonction

$$\phi(t) = \int_0^t \|\gamma'(u)\| du. \tag{91.424}$$

Il se fait que, dans le cas de l'ellipse, cette fonction est compliquée à calculer [4]. Nous allons donc simplement l'appeler « ϕ ». Par définition et parce que $\gamma(0) = A$,

$$d(A, \gamma(t)) = \phi(t). \tag{91.425}$$

4. Elle est même impossible à exprimer à l'aide des fonctions simples.

Par conséquent

$$d\Big(A,\gamma\big(\phi^{-1}(s)\big)\Big)=\phi\big(\phi(s)\big)=s. \tag{91.426}$$

Le paramétrage à choisir est donc

$$\sigma_A=\gamma\circ\phi^{-1}. \tag{91.427}$$

Cela est un résultat déjà trouvé dans le cours.

En ce qui concerne le point B, nous avons $t\in[0,2\pi]$ et

$$\gamma_B(t)=\begin{pmatrix}a\cos(t+\frac{\pi}{2})\\ b\sin(t+\frac{\pi}{2})\end{pmatrix}. \tag{91.428}$$

Cela est le paramétrage de l'ellipse dont le « premier » point est B. Le lien entre γ_B et γ est

$$\gamma_B(t)=\gamma(t+\frac{\pi}{2}). \tag{91.429}$$

La réponse peut donc être donné en considérant la fonction

$$\phi_B(t)=\int_0^t\|\gamma_B'(u)\|du, \tag{91.430}$$

et ensuite

$$\sigma_B(s)=\gamma_B\circ\phi_B^{-1} \tag{91.431}$$

EqReponseSigmaB

Pour le sport, nous allons exprimer cela en termes de γ et ϕ. En suivant le même raisonnement que précédemment, et en effectuant le changement de variable $v=u+\pi/2$,

$$\begin{aligned}d\big(B,\gamma_B(t)\big)&=\int_0^t\|\gamma_B'(u)\|du\\ &=\int_0^t\|\gamma_A'\big(u+\frac{\pi}{2}\big)\|du\\ &=\int_{\pi/2}^{t+\pi/2}\|\gamma'(v)\|dv\\ &=\phi\big(t+\frac{\pi}{2}\big)-\phi\big(\frac{\pi}{2}\big).\end{aligned} \tag{91.432}$$

À partir de là, nous pouvons calculer ϕ_B^{-1} en termes de l'inverse de ϕ de la façon suivante. Nous avons

$$t=\phi_B^{-1}(s) \tag{91.433}$$

si et seulement si (en prenant ϕ_B des deux côtés) :

$$s=\phi(t+\frac{\pi}{2})-\phi(\frac{\pi}{2}). \tag{91.434}$$

D'où nous déduisons successivement :

$$\begin{aligned}s+\phi(\frac{\pi}{2})&=\phi(t+\frac{\pi}{2})\\ \phi^{-1}\Big(s+\phi(\frac{\pi}{2})\Big)&=t+\frac{\pi}{2}\\ t&=\phi^{-1}\Big(s+\phi(\frac{\pi}{2})\Big)-\frac{\pi}{2}.\end{aligned} \tag{91.435}$$

Nous avons donc prouvé que

$$\phi_B^{-1}(s)=\phi^{-1}\Big(s+\phi(\frac{\pi}{2})\Big)-\frac{\pi}{2}. \tag{91.436}$$

Le sens de la réponse (91.431) est que le point

$$\gamma_B\big(\phi_B^{-1}(s)\big) \tag{91.437}$$

est à la distance s du point B. Autrement dit :

$$d\Big(B, \gamma_B\big(\phi_B^{-1}(s)\big)\Big) = s. \tag{91.438}$$

Cela peut s'exprimer maintenant en termes de ϕ et de γ de la façon suivante :

$$\begin{aligned}\gamma_B\big(\phi_B^{-1}(s)\big) &= \gamma\big(\phi_B^{-1}(s) + \frac{\pi}{2}\big)\\ &= \gamma\Big(\phi^{-1}\big(s + \phi(\frac{\pi}{2})\big)\Big)\end{aligned} \tag{91.439}$$

Exercise 301 exodevoir3-0004

Calculer la longueur de chacune des courbes suivantes.

(1) $\overrightarrow{r}(t) = (2\sin(t), -2\cos(t), 2t)$ pour t dans l'intervalle $[0, 2\pi]$;

(2) $\overrightarrow{r}(t) = (e^{-t}\cos(t), e^{-t}\sin(t), e^{-t})$ pour t dans l'intervalle $[0, 2]$;

(3) $\overrightarrow{r}(t) = (t^2, 4t + 3, -4\ln(t))$ pour t dans l'intervalle $[1, 2]$.

corrdevoir3-0004

Correction of the exercise 301

Soit L la courbe de $\mathbb{R}^n$ paramétrée par $t \to \mathbf{r}(t) = (r_1(t), \dots, r_n(t))$, pour t dans l'intervalle I. Pour calculer la longueur de L il nous faut utiliser la formule

$$\text{Longueur de } L = \int_I |\overrightarrow{r'}(t)|\, dt = \int_I \sqrt{\left(\frac{dr_1(t)}{dt}\right)^2 + \cdots + \left(\frac{dr_n(t)}{dt}\right)^2}\, dt. \tag{91.440}$$

On obtient alors :

(1) $\overrightarrow{r'}(t) = (2\cos(t), 2\sin(t), 2)$, et t dans $[0, 2\pi]$,

$$\text{Longueur de la première courbe } = \int_0^{2\pi} \sqrt{4\cos^2(t) + 4\sin^2(t) + 4}\, dt = 4\sqrt{2}\pi. \tag{91.441}$$

(2) $\overrightarrow{r'}(t) = (-e^{-t}\cos(t) - e^{-t}\sin(t), -e^{-t}\sin(t) + e^{-t}\cos(t), -e^{-t})$ et t dans $[0, 2]$,

$$\begin{aligned}&\text{Longueur de la deuxième courbe } =\\ &\int_0^2 \sqrt{e^{-2t}(\cos(t) + \sin(t))^2 + e^{-2t}(\cos(t) - \sin(t))^2 + e^{-2t}}dt =\\ &\int_0^2 \sqrt{3e^{-2t}}\, dt = \sqrt{3}(1 - e^2).\end{aligned} \tag{91.442}$$

(3) $\overrightarrow{r'}(t) = (2t, 4, -4/t)$ pour t dans l'intervalle $[1, 2]$.

$$\begin{aligned}&\text{Longueur de la troisième courbe } = \int_1^2 \sqrt{4t^2 + 16 + \frac{16}{t^2}}\, dt =\\ &\int_1^2 2t + \frac{4}{t}\, dt = 3 + 4\ln(2).\end{aligned} \tag{91.443}$$

Exercise 302 exocontrolecontinu0001

Pour chacun des ensembles suivants, trouver son intérieur, son adhérence, sa frontière et l'ensemble de ses points d'accumulation. Préciser si ces ensembles sont ouverts, fermés, bornés. Faire un dessin de chaque situation.

(1) $A_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 < y \leqslant \sqrt{1 - x^2},\ x \in [-2, 0.5[\}$;

(2) $A_2 = \{(x^2, y) \in \mathbb{R}^2 \mid x \in \mathbb{N}, y \in \mathbb{R}\}$.

corrcontrolecontinu0001

Correction of the exercise 302

(1) D'abord remarquons que la fonction $\sqrt{1-x^2}$ n'est définie que pour $|x| \leqslant 1$. Nous devons donc restreindre le domaine de $x \in [-1, \frac{1}{2}[$. Afin de trouver les points d'intersection entre x^2 et $\sqrt{1-x^2}$, nous commençons par résoudre l'équation au carré :

$$x^4 = 1 - x^2 \tag{91.444}$$

en posant $u = x^2$, c'est-à-dire l'équation $u^2 + u - 1 = 0$. Nous ne retenons que les solutions avec $u \geqslant 0$ et ce que nous trouvons est $x = \pm x_0$ avec

$$x_0 = \frac{\sqrt{\sqrt{5}-1}}{2}\sqrt{2}. \tag{91.445}$$

Étant donné que $\sqrt{5}$ est un peu plus grand que 2, ces solutions sont en valeur absolues plus petites que 1. L'intersection $x = -x_0$ est donc certainement importante : c'est à partir de là que x^2 devient plus petit que $\sqrt{1-x^2}$. Étant donné que $x_0 > 1/2$, l'autre intersection n'est pas importante.

Un dessin de la situation est à la figure 91.26.

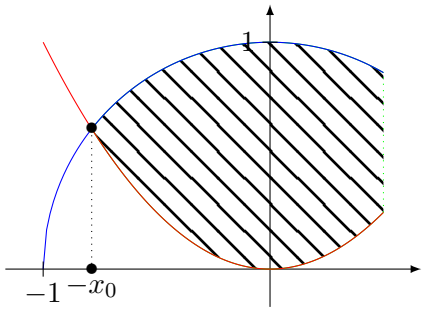


FIGURE 91.26: Pour l'exercice 302 LabelFigexamssepti

Le domaine n'est pas ouvert parce que les points sur le bord du dessus ($y = \sqrt{1-x^2}$) sont dans le domaine ; le domaine n'est pas non plus fermé parce que les points sur le bord droit $x = 1/2$ ne sont pas dans le domaine.

Le domaine est borné.

L'intérieur est donné par les équations $x^2 < y < \sqrt{1-x^2}$ avec $x \in]-1, \frac{1}{2}[$. Notez que le point d'intersection à gauche n'est pas dans l'intérieur. L'adhérence est donnée par les équations $x^2 \leqslant y \leqslant \sqrt{1-x^2}$ avec $x \in [-1, \frac{1}{2}]$.

La frontière est constituée du graphe des deux fonctions x^2 et $\sqrt{1-x^2}$ entre $x = -1$ et $x = \frac{1}{2}$ ainsi que du segment $x = \frac{1}{2}$ avec y entre $\frac{1}{4}$ et $\frac{3}{4}$.

(2) Ceci est un exercice très similaire à l'exemple 9.79. Pour chaque $x \in \mathbb{N}$ nous avons toute la droite verticale correspondante, voir la figure 91.27.

L'ensemble est non borné, il est fermé (et non ouvert). Sa fermeture est lui-même, et son intérieur est vide. Sa frontière est lui-même.

Exercise 303 exocontrolecontinu0002

(1) Trouver le domaine de définition de la fonction

$$f(x,y) = \frac{\sqrt{y-3}}{\ln(x+y)}. \tag{91.446}$$

(2) Tracer, si elles existent, les courbes de niveau de la fonction $f(x,y) = x^2 + (y-3)^2$ aux niveaux -1, 0 et 1.

corrcontrolecontinu0002

Correction of the exercise 303

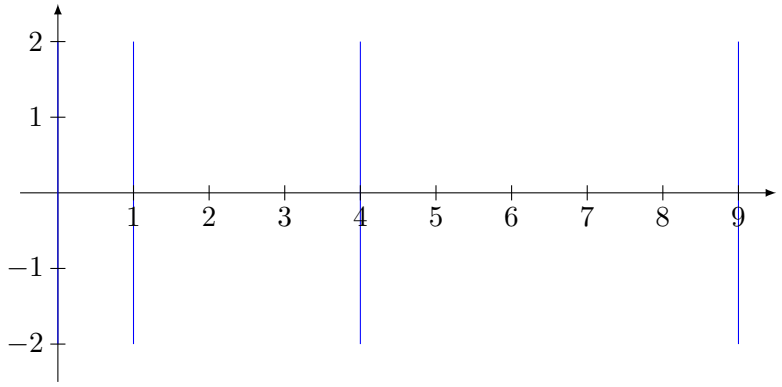


FIGURE 91.27: Pour l'exerice 302. LabelFigexamsseptii

(1) Le domaine de définition de la fonction

$$f(x,y)=\frac{\sqrt{y-3}}{\ln(x+y)}. \tag{91.447}$$

est la partie de $\mathbb{R}$ déterminée par le conditions

$$y\geqslant 3,\qquad ,x+y>0,\quad \ln(x+y)\neq 0.$$

cette dernière condition peut s'écrire aussi $x+y\neq 1$.

(2) La courbe de niveau -1 n'existe pas, celle de niveau 0 est donnée par le point $(0,3)$, la courbe de niveau 1 est un cercle de rayon 1 centrée au point $(0,3)$.

Exercise 304 exocontrolecontinu0003

Soit $f:\mathbb{R}^2\to\mathbb{R}$ la fonction

$$f(x,y)=\begin{cases}\frac{y\sin^2(2x)}{x}, & \text{si } x\neq 0\\ 0, & sinon.\end{cases} \tag{91.448}$$

(1) Montrer que f est une fonction continue sur $\mathbb{R}^2$.
(2) Calculer les dérivées partielles de f au point $(\pi/6,4)$.
(3) Calculer les dérivées partielles de f au point $(0,2)$.
(4) Les dérivées partielles sont-elles continues au point $(0,2)$?
(5) La fonction f est-elle différentiable au point $(0,2)$?

corrcontrolecontinu0003

Correction of the exercise 304

(1) Les points où la continuité n'est pas assurée sont les points de la forme $(0,a)$ avec $a\in\mathbb{R}$. Nous devons donc calculer, pour tout a la limite

$$\lim_{(x,y)\to(0,a)}\frac{y\sin^2(2x)}{x} \tag{91.449}$$

Nous pouvons calculer séparément les deux limites

$$\lim_{(x,y)\to(0,a)} y=0 \tag{91.450a}$$

$$\lim_{(x,y)\to(0,a)}\frac{\sin^2(2x)}{x}=0. \tag{91.450b}$$

Étant donné que les deux limites existent, la limite du produit est égale au produit des limites (proposition 12.6). Notre limite vaut donc zéro et la fonction est continue (parce que la limite vers $(0,a)$) est égale à la valeur en $(0,a)$.

(2) Le calcul montre que SubeqsDerpartGAzztc

$$\frac{\partial f}{\partial x} = \frac{4y\sin(2x)\cos(2x)}{x} - y\frac{\sin^2(2x)}{x^2} \tag{91.451a}$$
$$\frac{\partial f}{\partial y} = \frac{\sin^2(2x)}{x}. \tag{91.451b}$$

En posant $x = \frac{\pi}{6}$, $y = 4$ nous trouvons

$$\frac{\partial f}{\partial x}(\frac{\pi}{6}, 4) = \frac{24\sqrt{3}}{\pi} - \frac{108}{\pi^2} \tag{91.452a}$$
$$\frac{\partial f}{\partial y}(\frac{\pi}{6}, 4) = \frac{9}{2\pi}. \tag{91.452b}$$

(3) Le point $(0, 2)$ n'étant pas dans le domaine sur lequel la fonction est « facile », nous devons utiliser la définition de la dérivée directionnelle :

$$\frac{\partial f}{\partial x}(0,2) = \lim_{t\to 0}\frac{f(t,2) - f(0,2)}{t} \tag{91.453a}$$
$$= \lim_{t\to 0}\frac{2\sin^2(2t)}{t^2} \tag{91.453b}$$
$$= 2\lim_{t\to 0}\frac{2\sin(2t)}{2t}\frac{2\sin(2t)}{2t} \tag{91.453c}$$
$$= 8. \tag{91.453d}$$

Dans ce calcul, nous avons multiplié le numérateur et le dénominateur par 4. La dérivée selon y est plus simple parce que $f(0, 2 + t) = f(0, 2) = 0$, ce qui donne

$$\frac{\partial f}{\partial y}(0,2) = \lim_{t\to 0}\frac{f(0,2+t) - f(0,2)}{t} = 0. \tag{91.454}$$ EqAGzztczdttozfrac

(4) Nous devons vérifier si en prenant la limite $(x, y) \to (0, 2)$ des fonctions (91.451) sont bien les nombres $\partial_x f(0, 2)$ et $\partial_y f(0, 2)$. La première limite à calculer est

$$\lim_{(x,y)\to(0,2)} \frac{4y\sin(2x)\cos(2x)}{x} - y\frac{\sin^2(2x)}{x^2}. \tag{91.455}$$ EqlimxyAGzztcqymyfrac

En décomposant et en sachant que $\lim_{x\to 0}\sin(2x)/x = 2$, nous trouvons que la limite (91.455) vaut zéro, ce qui ne correspond pas à la valeur de $\partial_x f(0, 2)$ donnée par (91.454). La dérivée partielle par rapport à x n'est donc pas continue en $(0, 2)$. En ce qui concerne la dérivée par rapport à y, nous avons

$$\lim_{(x,y)\to(0,2)} \frac{\sin^2(2x)}{x} = 0, \tag{91.456}$$

cette dérivée directionnelle est donc continue.

Exercise 305 exocontrolecontinu0004

Tracer le domaine d'intégration de l'intégrale

$$\int_0^2 \int_0^{\sqrt{2x-x^2}} \sqrt{x^2+y^2}\,dy\,dx,$$

et calculer sa valeur (à l'aide des coordonées polaires). corrcontrolecontinu0004

Correction of the exercise 305

Le domaine d'intégration est donné par une moitié du cerce de centre $(1, 0)$ et de rayon 1, plus précisement la moitié contenue dans le premier quadrant.

L'équation de ce cercle en coordonnées polaires est $r = 2\cos(\theta)$. Le fait qu'on ne considère que la partie contenue dans le premier quadrant se traduit par $\theta \in [0, \pi/2]$.

Notre intégrale est alors

$$\int_0^{\pi/2}\int_0^{2\cos(\theta)} r^2\,dr\,d\theta = \frac{8}{3}\int_0^{\pi/2}\cos^3(\theta)\,d\theta.$$

La valeur de cette intégrale est 16/9. Une méthode pour le voir est la suivante (on utilise plusieurs fois l'intégration par parties)

$$\begin{aligned}\int_0^{\pi/2}\cos^3(\theta)\,d\theta &= \sin(\theta)\cos^2(\theta)\big|_0^{\pi/2} + 2\int_0^{\pi/2}\sin^2(\theta)\cos(\theta)\,d\theta \\ &= 2\sin^3(\theta)\big|_0^{\pi/2} - 4\int_0^{\pi/2}\sin^2(\theta)\cos(\theta)\,d\theta.\end{aligned} \tag{91.457}$$

Cela nous dit que

$$2\sin^3(\theta)\big|_0^{\pi/2} = 6\int_0^{\pi/2}\sin^2(\theta)\cos(\theta)\,d\theta = 3\int_0^{\pi/2}\cos^3(\theta)\,d\theta \tag{91.458}$$

c'est-à-dire

$$\int_0^{\pi/2}\cos^3(\theta)\,d\theta = \frac{2}{3}.$$

Exercise 306 exocontrolecontinu0005

Calculer la longueur de la courbe

$$\begin{array}{rccc}\gamma: & [0,4\pi] & \to & \mathbb{R}^3 \\ & t & \mapsto & \left(t\cos(t), t\sin(t), \frac{2\sqrt{2}}{3}t^{3/2}\right).\end{array} \tag{91.459}$$

corrcontrolecontinu0005

Correction of the exercise 306

La longueur de la courbe γ est donnée par la formule

$$\text{Longueur} = \int_0^{4\pi}\|\gamma'(t)\|\,dt. \tag{91.460}$$

La norme de $\gamma'(t)$ est

$$\sqrt{(\cos(t)-t\sin(t))^2+(\sin(t)+t\cos(t))^2+2t} = \sqrt{1+t^2+2t} = 1+t.$$

La longueur de γ est alors

$$\int_0^{4\pi} 1+t\,dt = 4\pi+8\pi^2.$$

Exercise 307 exocontrolecontinu0006

Soit f la fonction $f(x,y)=x^2-3x-5y^3$. Calculer la dérivée de f suivant le vecteur

$$\frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix}, \tag{91.461}$$

au point $(2,3)$.

corrcontrolecontinu0006

Correction of the exercise 307

La fonction f est différentiable, parce que c'est un polynôme. Le gradient de f est $\nabla f(x,y) = (2x-3, -15y^2)$, qui au point $(2,3)$ vaut $(1,-135)$. La dérivée de f suivant le vecteur donné au point $(2,3)$ est

$$\frac{1}{\sqrt{5}}\nabla f(2,3)\cdot\begin{pmatrix}1\\2\end{pmatrix} = \frac{1}{\sqrt{5}}(1-270) = -\frac{269}{5}. \tag{91.462}$$

Exercise 308 exocontrolecontinu0007

Soient $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y)=x\ln(yx^2)$, et $g:\mathbb{R}\to\mathbb{R}^2$, $g(t)=(t,e^t)$.

Trouver la matrice jacobienne $F = g \circ f$ au point $(\pi/2, 1)$. corrcontrolecontinu0007

Correction of the exercise 308

La méthode la plus rapide pour compléter l'exercice est de calculer d'abord les matrices jacobiennes de f et g, ou pour mieux dire, le gradient de f et le jacobien de g.

$$\nabla f(x,y) = \begin{pmatrix} \ln(yx^2) + 2 & \frac{x}{y} \end{pmatrix}, \tag{91.463}$$

$$\frac{d}{dt} g(t) = \begin{pmatrix} 1 \\ e^t \end{pmatrix}. \tag{91.464}$$

La fonction $F = g \circ f$ est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Sa matrice jacobienne doit alors être une matrice 2×2. On la trouve par la formule

$$\begin{aligned} &\nabla F(\pi/2, 1) = \nabla g(f(\pi/2,1)) \cdot \nabla f(\pi/2, 1) = \\ &\begin{pmatrix} 1 \\ \left(\frac{\pi}{2}\right)^\pi \end{pmatrix} \cdot \begin{pmatrix} \ln(\pi^2/4) + 2 & \frac{\pi}{2} \end{pmatrix} = \\ &\begin{pmatrix} \ln(\pi^2/4) + 2 & \frac{\pi}{2} \\ \left(\frac{\pi}{2}\right)^\pi (\ln(\pi^2/4) + 2) & \left(\frac{\pi}{2}\right)^{\pi+1} \end{pmatrix} \end{aligned} \tag{91.465}$$

Exercise 309 exocontrolecontinu0008

Trouver, si elle existe, une fonction $g : \mathbb{R}^3 \to \mathbb{R}$ telle que

$$\begin{aligned} &\partial_x g(x,y,z) = yz + y\cos(xy), \\ &\partial_y g(x,y,z) = xz + x\cos(xy), \\ &\partial_z g(x,y,z) = xy, \\ &g(0,1,2) = 5. \end{aligned} \tag{91.466}$$

corrcontrolecontinu0008

Correction of the exercise 309

Une telle fonction existe et est unique. Pour la trouver il faut d'abord calculer les integrales partielles

$$\begin{aligned} g(x,y,z) &= \int \partial_x g(x,y,z)\, dx = xyz + \sin(xy) + f_1(y,z), \\ g(x,y,z) &= \int \partial_y g(x,y,z)\, dy = xyz + \sin(xy) + f_2(x,z), \\ g(x,y,z) &= \int \partial_z g(x,y,z)\, dz = xyz + f_3(x,y). \end{aligned} \tag{91.467}$$

On obtient alors $g(x,y,z) = xyz + \sin(xy) + f_1(y,z) = xyz + \sin(xy) + f_2(x,z) = xyz + f_3(x,y)$. Cela nous dit que $g(x,y,z) = xyz + \sin(xy) + K$, où K est une constante. Pour trouver la valeur de K nous utilisons la dernière condition. Comme

$$g(0,1,2) = K,$$

c'est évident que K doit être 5. **Exercise 310** exocontrolecontinu0009

Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$

$$f(x,y) = \begin{cases} \frac{(x^3+y^2)}{\sqrt{x^2+y^2}} \sin\left(\frac{x}{x^2+y^2}\right), & \text{si } (x,y) \neq (0,0), \\ 0, \quad \text{sinon.} \end{cases} \tag{91.468}$$

(1) Calculer les dérivées partielles de f au point $(0,0)$.

(2) Établir si la fonction f est différentiable au point $(0,0)$.

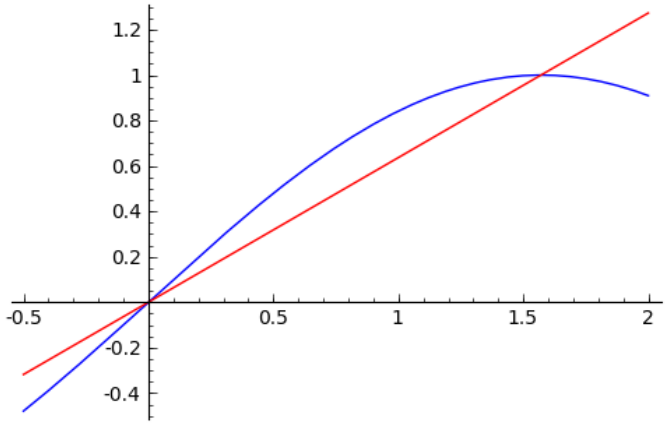


FIGURE 91.28: La région E

Exercise 311 exocontrolecontinu0010

Soit E la partie de $\mathbb{R}^2$ delimitée par les courbes $y = \sin(x)$ et $y = \frac{2}{\pi}x$ dans le premier quadrant. Dessiner E et calculer l'intégrale

$$\int_E x + 2y \, dS.$$

corrcontrolecontinu0010

Correction of the exercise 311

Voici un dessin de la région E. La courbe $y = \sin(x)$ est dessinée en bleu et la courbe $y = \frac{2}{\pi}x$ en rouge. Le points d'intersection des deux courbes dans le premier quadrant sont $(0, 0)$ et $(\pi/2, 1)$.

$$\begin{aligned}\int_E x + 2y \, dS &= \int_0^{\pi/2} \int_{2x/\pi}^{\sin(x)} x + 2y \, dy \, dx \\ &= \int_0^{\pi/2} x\sin(x) - \frac{2}{\pi}x^2 + \sin^2(x) - \frac{4x^2}{\pi^2} \, dx \\ &= \int_0^{\pi/2} x\sin(x) - \frac{(2\pi+4)}{\pi^2}x^2 + \sin^2(x) \, dx.\end{aligned} \tag{91.469}$$

Cette intégrale peut être calculée en considérant un terme à la fois. D'abord on obtient

$$\int_0^{\pi/2} x\sin(x) \, dx = -x\cos(x)|_0^{\pi/2} + \int_0^{\pi/2} \cos(x) \, dx = 1;$$

et

$$-\int_0^{\pi/2} \frac{(2\pi+4)}{\pi^2}x^2 \, dx = -\frac{(\pi^2+2\pi)}{12}.$$

Enfin, on observe

$$\int_0^{\pi/2} \sin^2(x) \, dx = -\cos(x)\sin(x)|_0^{\pi/2} + \int_0^{\pi/2} \cos^2(x) \, dx,$$

ce qui nous permet d'écrire

$$\int_0^{\pi/2} \sin^2(x) + \cos^2(x) \, dx = 2\int_0^{\pi/2} \sin^2(x) \, dx,$$

et donc

$$\int_0^{\pi/2} \sin^2(x) \, dx = \frac{\pi}{4}.$$

Exercise 312 exocontrolecontinu0011

(1) Donner la définition de norme.

(2) La fonction d de $\mathbb{R}^2 \times \mathbb{R}^2$ vers $\mathbb{R}$, $d(x,y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$ est la distance induite par la norme supremum dans $\mathbb{R}^2$. Vérifier que d est invariante par translations, c'est-à-dire que

$$d(x - z, y - z) = d(x, y),$$

pour tout choice de x, y et z dans $\mathbb{R}^2$.

corrcontrolecontinu0011

Correction of the exercise 312

(1) Pour la définition de norme, voir le cours.

(2) pour tout choice de x, y et z dans $\mathbb{R}^2$ on a

$$\begin{aligned} d(x - z, y - z) &= \max\{|(x_1 - z_1) - (y_1 - z_1)|, |(x_2 - z_2) - (y_2 - z_2)|\} \\ &= \max\{|(x_1 - y_1) + (z_1 - z_1)|, |x_2 - y_2| + (z_2 - z_2)\} \\ &= \max\{|x_1 - y_1|, |x_2 - y_2|\} = d(x, y). \end{aligned} \tag{91.470}$$

Exercise 313 exocontrolecontinu0012

Trouver le paramétrage normal du cercle de rayon R centré dans l'origine.
Un paramétrage de cette courbe est $(R\cos\theta, R\sin\theta)$.

corrcontrolecontinu0012

Correction of the exercise 313

Le paramétrage « usuelle » du cercle est

$$\sigma(t) = (R\cos(t), R\sin(t)). \tag{91.471}$$

Les formules qui donnent le paramétrage normal sont les formules (21.132) et (21.133). Nous devons donc d'abord calculer

$$\phi(t) = \int_0^{2\pi R} \|\sigma'(u)\| du = Rt. \tag{91.472}$$

Ensuite, $\gamma_N(s) = (\sigma \circ \phi^{-1})(s)$ où

$$\phi^{-1}(s) = \frac{s}{R}. \tag{91.473}$$

Donc le paramétrage normal est

$$\gamma_N(s) = \sigma\left(\frac{s}{R}\right) = \begin{pmatrix} R\cos\left(\frac{s}{R}\right) \\ R\sin\left(\frac{s}{R}\right). \end{pmatrix} \tag{91.474}$$

Exercise 314 exocontrolecontinu0013

Calculer le volume de la partie E de $\mathbb{R}^3$ dans le premier octant entre les deux sphères $x^2 + y^2 + z^2 = 1$ et $x^2 + y^2 + z^2 = 4$ et au dessus du cône $z = \sqrt{\frac{x^2 + y^2}{2}}$. Conseil : passer aux coordonées sphèriques.

Bonus : Dessiner E.

corrcontrolecontinu0013

Correction of the exercise 314

Il nous faut calculer $\int_E 1\, dV$. Le volume E peut être décrit en coordonnées sphèriques de la façon suivante. D'abord E est contenu dans le premier octant, c'est-à-dire que les valeurs de θ à considérer sont entre 0 et $\pi/2$, et que z ne peut que être positif. De plus, E est contenu entre les deux sphères centrées dans l'origine de rayons 1 et 2. Cela nous dit que $1 < r < 2$ pour tous le points de E. Enfin on traduit léquation du cône en coordonées sphèriques et on obtient

$$2\cos^2(\phi) = \sin^2(\phi),$$

soit $\tan(\phi) = \sqrt{2}$. À chaque point de E correspond un valeur de ϕ compris entre 0 et $\arctan(\sqrt{2})$.

L'intégrale à calculer est alors

$$\int_0^{\arctan(\sqrt{2})} \int_0^{\pi/2} \int_1^2 r^2 \sin\phi \, dr \, d\theta \, d\phi = \frac{7\pi}{6}(\cos(\arctan(\sqrt{2})) - 1).$$

Exercise 315 exoGeomAnal-0015

Pour chacun des ensembles suivants, trouver son intérieur, son adhérence, sa frontière et l'ensemble de ses points d'accumulation. Préciser si ces ensembles sont ouverts, fermés, bornés. Faire un dessin de chaque situation.

(1) $A_1 = \{(x, y) \in \mathbb{R}^2 \mid x - 1 \leqslant y \leqslant \sin(x),\ x \in]0, \pi/2]\}$;

(2) $A_2 = \{(\cos(\pi x), y) \in \mathbb{R}^2 \mid x \in \mathbb{N}, y \in [0, +\infty[\}$.

corrGeomAnal-0015

Correction of the exercise 315

<+CorrGeomAnal-0015+>

Exercise 316 exoGeomAnal-0016

(1) Trouver le domaine de définition de la fonction

$$f(x, y) = \frac{y^3 + 2011}{\ln(x/50)\sqrt{x - 34}}. \tag{91.475}$$

(2) Tracer, si elles existent, les courbes de niveau de la fonction $f(x, y) = (x+1)^2 + y$ aux niveaux -1, 0 et 1.

corrGeomAnal-0016

Correction of the exercise 316

<+CorrGeomAnal-0016+>

Exercise 317 exoGeomAnal-0018

(1) Montrer que la limite

$$\lim_{(x,y)\to(0,0)} \frac{x - y^2}{x + y^2}, \tag{91.476}$$

n'existe pas.

(2) Calculer, si elle existe, la valeur de la limite

$$\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x + y}. \tag{91.477}$$

corrGeomAnal-0018

Correction of the exercise 317

<+CorrGeomAnal-0018+>

Exercise 318 exoGeomAnal-0021

Soit f la fonction $f(x, y, z) = \sin(x)\ln(5y) + 120ze^x$. Trouver le domaine de f et calculer sont gradient au point $P = (\pi, 0.2, 0.05)$. Calculer la dérivée de f au point P suivant le vecteur

$$\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{91.478}$$

corrGeomAnal-0021

Correction of the exercise 318

<+CorrGeomAnal-0021+>

Exercise 319 exoGeomAnal-0023

Calculer l'intégrale de la fonction $f(x, y) = y$ sur la région de $\mathbb{R}^2$ bornée par la parabole $y^2 = 2x + 6$ et la droite $y = x - 1$. Commencer par faire un dessin du domaine d'intégration.

corrGeomAnal-0023

Correction of the exercise 319

<+CorrGeomAnal-0023+>

Exercise 320 exoGeomAnal-0027

Calculer la longueur de la courbe $\rho(\theta) = 1 + \cos(\theta)$ (cardioïde).
Rappel : $\cos(\theta) = \cos^2(\theta/2) - \sin^2(\theta/2)$ pour tout t.

corrGeomAnal-0027

Correction of the exercise 320

<+CorrGeomAnal-0027+>

Exercise 321 exoGeomAnal-0017

(1) Expliquer pourquoi il n'existe pas une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\partial_x f(x,y) = 3x + y, \qquad \partial_y f(x,y) = 4\sin(xy). \tag{91.479}$$

(2) Soit $h : \mathbb{R}^3 \to \mathbb{R}$ une fonction dont les dérivées partielles existent partout dans $\mathbb{R}^3$. Nous savons que $\partial_z h$ est continue sur $\mathbb{R}^3\backslash\{(0,1,0)\}$ et que $\partial_x h$ et $\partial_y h$ sont continue partout dans $\mathbb{R}^3$. Qu'est-ce que pouvons nous dire de la différentiabilité de h ?

corrGeomAnal-0017

Correction of the exercise 321

(1) Étant donné que $\partial_x f(x,y) = 3x + y$, nous savons que $f(x,y)$ doit être de la forme

$$f(x,y) = \frac{3}{2}x^2 + yx + a(y) \tag{91.480}$$ EqGAzzusui

où $a(y)$ est une fonction seulement de y. Aucune fonction de ce type n'a $4sin(x,y)$ comme dérivée par rapport à y. En effet, la dérivée de (91.480) par rapport à y est

$$3x + y + a'(y), \tag{91.481}$$ EqGAzzuuii

et il n'y a aucun choix de fonction $a(y)$ tel que (91.481) se réduise à $4\sin(xy)$.

(2) Vu que toutes les dérivées partielles de h sont continues sur $\mathbb{R}^3\backslash\{0,1,0\}$, la fonction h y est différentiable (proposition 12.295)

Exercise 322 exoGeomAnal-0020

(1) Calculer la valeur de l'intégrale

$$\mathcal{I} = \int_0^2 \int_0^y e^{x+5y}(x-y)\,dx\,dy.$$

(2) On veut récrire $\mathcal{I}$ dans le nouvelles variables $s = x + 5y$ et $t = x - y$.

(2a) Trouver le domaine d'intégration dans le nouvelles variables. Faire un dessin ;

(2b) Trouver le Jacobien du changement de variables ;

(2c) Écrire la fonction à intégrer par rapport aux nouvelles variables.

corrGeomAnal-0020

Correction of the exercise 322

<+CorrGeomAnal-0020+>

Exercise 323 exoGeomAnal-0022

Trouver, si elle existe, une fonction $g : \mathbb{R}^3 \to \mathbb{R}$ telle que

$$\begin{aligned}
\partial_x g(x,y,z) &= \frac{3\sin^2(z)}{x} + 2y(xy - z),\\
\partial_y g(x,y,z) &= 2x(xy - z),\\
\partial_z g(x,y,z) &= 3\sin(2z)\ln(x) - 2(xy - z),\\
g(0,1,\pi) &= 5.
\end{aligned} \tag{91.482}$$

corrGeomAnal-0022

Correction of the exercise 323

<+CorrGeomAnal-0022+>

Exercise 324 exoGeomAnal-0024

Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$

$$f(x,y) = \begin{cases} \frac{(x^3+y^2)}{\sqrt{x^2+y^2}} \sin\left(\frac{x}{x^2+y^2}\right), & \text{si } (x,y) \neq (0,0), \\ 0, \qquad \text{sinon.} \end{cases} \tag{91.483}$$

(1) Calculer les dérivées partielles de f au point $(0,0)$.

(2) Établir si la fonction f est différentiable au point $(0,0)$.

corrGeomAnal-0024

Correction of the exercise 324

<+CorrGeomAnal-0024+>

Exercise 325 exoGeomAnal-0025

Soit E la partie de $\mathbb{R}^3$ délimitée par le paraboloïde $z = x^2 + y^2 - 3$ et le cône $z = \frac{1}{2}\sqrt{x^2+y^2}$.

(1) Dessiner E ;

(2) le décrire en coordonnées sphériques **et** cylindriques ;

(3) calculer son volume.

corrGeomAnal-0025

Correction of the exercise 325

<+CorrGeomAnal-0025+>

Exercise 326 exoGeomAnal-0026

(1) Soit A une partie de $\mathbb{R}^2$.

Donner la définition de l'adhérence, de l'intérieur et de la frontière de A.

(2) Est-ce que l'adhérence d'un ensemble corresponds toujours à l'union de la frontière et de l'intérieur d'un ensemble ? Si non, donnez un exemple.

corrGeomAnal-0026

Correction of the exercise 326

<+CorrGeomAnal-0026+>

Exercise 327 exoDS2011-0001

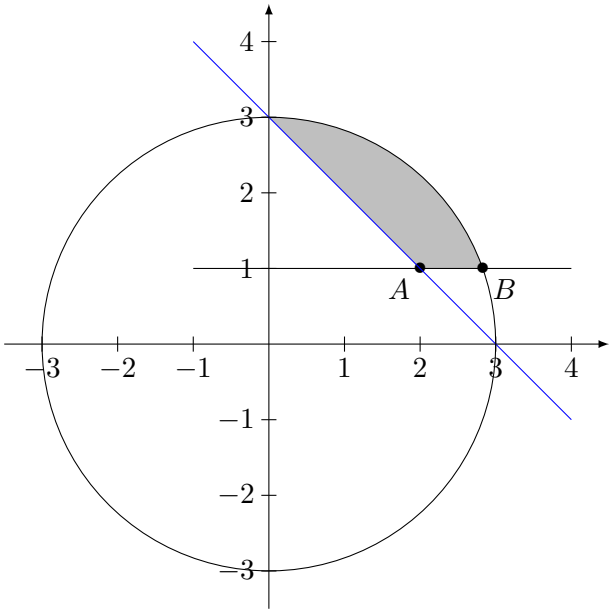
Pour chacun des ensembles suivants, trouver son intérieur, son adhérence, sa frontière. Préciser si ces ensembles sont ouverts, fermés, bornés. Faire un dessin de chaque situation.

(1) $A_1 = \{(x,y) \in \mathbb{R}^2 \mid 3 - y < x \leqslant \sqrt{9-y^2},\ y > 1\}$;

(2) $A_3 = \{(n, 2n+1) \in \mathbb{R}^2 \mid n \in \mathbb{N}\}$.

corrDS2011-0001

Correction of the exercise 327



(1)

Les deux points d'intersection A et B s'obtiennent en posant $y = 1$ dans les équations $x = 3 - y$ et $x = \sqrt{9 - y^2}$. Les résultats sont $A = (2, 1)$ et $B = (2\sqrt{2}, 1)$

L'intérieur et l'adhérence s'obtiennent en mettant respectivement les inégalités strictes et non strictes dans la définition. L'ensemble n'est ni ouvert, ni fermé. Il est borné. Sa frontière s'écrit comme l'union de trois pièces :

$$\begin{aligned}\partial A_1 = &\{(x, 1) \text{ tel que } x \in [2, 2\sqrt{2}]\}\\ &\cup \{x = 3 - y \text{ tel que } x \in [0, 2]\}\\ &\cup \{x = \sqrt{9 - y^2} \text{ tel que } x \in [0, 2\sqrt{2}]\}.\end{aligned} \tag{91.484}$$

(2) L'ensemble A_2 est une suite de points dans $\mathbb{R}^2$. Son intérieur est vide, son adhérence est lui-même, il n'est pas ouvert, il est fermé, et il est non borné.

<++>

Exercise 328 exoDS2011-0002

Calculer la longueur de la courbe paramétrée $([0, 1], \gamma)$, où γ est la fonction

$$\begin{array}{rccc}\gamma : & \mathbb{R} & \to & \mathbb{R}^3\\ & t & \mapsto & \begin{pmatrix}\frac{2}{3}t^3\\ 4\\ \frac{t^2}{2} - \frac{t^4}{4}\end{pmatrix}.\end{array} \tag{91.485}$$

corrDS2011-0002

Correction of the exercise 328

Nous avons la dérivée du chemin :

$$\gamma'(t) = \begin{pmatrix}2t^2\\ 0\\ t - t^3\end{pmatrix}, \tag{91.486}$$

puis la norme

$$\|\gamma'(t)\| = \sqrt{t^2 + t^6 + 2t^4} = t + t^3. \tag{91.487}$$

Attention, ici pour enlever la racine carrée dans l'égalité

$$\sqrt{(t + t^3)^2} = t + t^3, \tag{91.488}$$

Eqsarttipte

il faut bien justifier qu'avec $t \in [0, 1]$, $t + t^3 > 0$. L'égalité (91.488) n'est pas vraie pour toute valeur de t.

En ce qui concerne la longueur,

$$l(\gamma) = \int_0^1 (t + t^3)dt = \frac{3}{4}. \tag{91.489}$$

Exercise 329 exoDS2011-0003

Quelques calculs sur les normes et les fonctions à valeurs vectorielles.

(1) Soit $v = (2, 1, 6)$. Calculer les normes $\|v\|_1$, $\|v\|_2$, $\|v\|_\infty$.

(2) Trouver un vecteur $w_1 \in \mathbb{R}^2$ de norme 1 par rapport à la norme $\| \cdot \|_\infty$ qui a la même direction du vecteur $w = (2, 3)$. Dessiner les vecteurs w, w_1 et la boule unitaire de $\| \cdot \|_\infty$.

(3) Soit $F : \mathbb{R} \to \mathbb{R}^3$ la fonction de composantes $f_1(x) = x$, $f_2(x) = \dfrac{1}{\ln(x + 1)}$, $f_3(x) = \sqrt{x^2 - 4}$. Trouver le domaine de définition de F et ses limites aux extrêmes du domaine, si elles existent. corrDS2011-0003

Correction of the exercise 329

(1) $9, \sqrt{41}, 6$.

(2) Un vecteur qui a la même direction que $(2, 3)$ est un vecteur de la forme $\lambda(2, 3)$ pour un certain λ. Nous fixons λ afin que la norme soit 1, c'est-à-dire

$$\max\{|2\lambda|, |3\lambda|\} = 1 \tag{91.490}$$

Nous trouvons $\lambda = 1/3$ et par conséquent $w_1 = (\frac{2}{3}, 1)$.

(3) Nous avons les restrictions suivantes sur le domaine :

$$\begin{cases} \ln(x + 1) \neq 0 & \text{(91.491a)} \\ x + 1 > 0 & \text{(91.491b)} \\ x^2 - 4 \geqslant 0. & \text{(91.491c)} \end{cases}$$

L'intersection des trois conditions est $x \geqslant 2$. Cela est donc le domaine de notre fonction. Les limites du domaine sont donc $x = 2$ et $x = \infty$. Nous avons

$$\lim_{x \to 2} \begin{pmatrix} x \\ 1/\ln(x + 1) \\ \sqrt{x^2 - 4} \end{pmatrix} = \begin{pmatrix} 2 \\ 1/\ln(3) \\ 9 \end{pmatrix}. \tag{91.492}$$

La limite en l'infini n'existe pas parce que dans la définition de limite, nous demandons que le vecteur limite ait une norme finie. Autant pour les fonctions $\mathbb{R} \to \mathbb{R}$, nous avons donné un sens à une expression du type $\lim_{x\to\infty} f(x) = \infty$, dans le cas de fonctions $\mathbb{R} \to \mathbb{R}^3$, nous n'avons pas donné de sens à des expressions du type

$$\lim_{x \to 2} \begin{pmatrix} x \\ 1/\ln(x + 1) \\ \sqrt{x^2 - 4} \end{pmatrix} = \begin{pmatrix} \infty \\ 0 \\ \infty \end{pmatrix}. \tag{91.493}$$ EqLimInfnsDS

La réponse (91.493) est donc fausse.

Exercise 330 exoDS2011-0004

Considérons courbe paramétrée $(]0, \pi/2], \gamma_{pol}(\theta))$, où la fonction γ_{pol} est la fonction de $]0, \pi/2]$ dans $\mathbb{R}^2$ $\gamma_{pol}(\theta) = (r(\theta) = 2\cos(\theta), \theta)$ (coordonnées polaires).

(1) Trouver un paramétrage du même arc géométrique de la forme $\gamma_{cart}(x) = (x, y(x))$ (coordonnées cartésiennes). Conseil : essayez d'écrire d'abord l'équation cartésienne de la courbe.

(2) La courbe est une partie d'un cercle. Trouver son centre, son rayon et dessiner la courbe dans le plan cartésien x-y.

(3) Déterminer la longueur de la courbe. Vous pouvez utiliser quelconque méthode qui vous vient à l'esprit quitte à la justifier.

Bonus : Dessiner la courbe dans le demi-plan ρ-θ.

corrDS2011-0004

Correction of the exercise 330

En termes du paramètre θ, la courbe s'écrit

$$\gamma(\theta) = \begin{pmatrix} r(\theta)\cos(\theta) \\ r(\theta)\sin(\theta) \end{pmatrix} = \begin{pmatrix} 2\cos^2(\theta) \\ 2\cos(\theta)\sin(\theta) \end{pmatrix}. \tag{91.494}$$

Étant donné que sur $\theta \in [0, \pi/2]$ la fonction $2\cos^2(\theta)$ est injective, il est possible d'utiliser xx comme paramètre. Nous avons $x = 2\cos^2(\theta)$, c'est-à-dire

$$\cos(\theta) = \sqrt{\frac{x}{2}} \tag{91.495a}$$

$$\sin(\theta) = \sqrt{1 - \frac{x}{2}}. \tag{91.495b}$$

Pour la seconde ligne, nous avons utilisé la formule $\sin(\theta) = \sqrt{1 - \cos^2(\theta)}$. Une fois de plus, cette formule se justifie par le fait que $\theta \in [0, \pi/2]$ et que $\sin(\theta)$ y est toujours positive. Nous avons au final

$$\gamma(x) = \begin{pmatrix} x \\ 2\sqrt{\frac{x}{2}\left(1 - \frac{x}{2}\right).} \end{pmatrix} \tag{91.496}$$

avec $x \in [0, 2]$. Note : préciser l'intervalle de x n'est pas facultatif.

De nombreux étudiants ont cru malin d'écrire

$$\theta = \arccos\sqrt{\frac{x}{2}} \tag{91.497}$$

et

$$\gamma(x) = \begin{pmatrix} x \\ 2\sqrt{\frac{x}{2}}\sin\left(\arccos\sqrt{\frac{x}{2}}\right) \end{pmatrix}. \tag{91.498}$$

EqCartfgxygammaqqn

Quelques remarques à propos de cette idée :

(1) Ce n'était pas la réponse attendue parce qu'elle est triviale. Si vous avez une fonction en termes du paramètre t donnée par

$$\gamma(t) = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}, \tag{91.499}$$

vous pouvez toujours poser $t = f^{-1}(t)$ et puis répondre

$$\gamma(x) = \begin{pmatrix} x \\ g\big(f^{-1}(x)\big) \end{pmatrix}. \tag{91.500}$$

La réponse en arccos n'était pas attendue, mais elle reste néanmoins correcte. Soyons généreux. Pour un prochain (ou examen), essayez de ne pas recommencer. De notre côté, nous essayerons de formuler la question de façon à vous empêcher de répondre ainsi ;)

(2) Lorsque vous utiliser la fonction inverse d'une fonction non injective (par exemple cosinus), vous *devez* préciser les domaines ! Dans le cas présent, vous devez considérer la fonction

$$\arccos\colon [0, 1] \to [0, \frac{\pi}{2}], \tag{91.501}$$

et justifier que lorsque $x \in [0, 2]$, la fonction $\sqrt{\frac{x}{2}}$ reste entre 0 et 1.

Pour comprendre de quel cercle il s'agit, nous pouvons partir de l'équation (91.498) et calculer r^2 en sommant les carrés des deux composantes. Le résultat est que

$$x^2 + y^2 = x^2 + 4\frac{x}{2}\left(1 - \frac{x}{2}\right), \tag{91.502}$$

et l'équation de notre cercle est $x^2 + y^2 = 2x$. Pour trouver le centre et le rayon, le mieux est de compléter le carré :

$$x^2 - 2x = (x-1)^2 - 1 \tag{91.503}$$

et par conséquent

$$(x-1)^2 + y^2 = 1. \tag{91.504}$$

Cela est l'équation d'un cercle de centre $(1, 0)$ et de rayon 1. Étant donné que dans le contexte de cet exercice nous avons $y \geqslant 0$ [5], nous n'avons en réalité que la moitié supérieure du cercle. Sa longueur vaut donc π.

Il y avait aussi moyen de calculer la longueur en utilisant une des équations paramétriques (polaires, ou cartésiennes) et en considérant la formule correspondante.

Exercise 331 exoGeomAnal-0045

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction $f(x, y) = \dfrac{x^2y^5}{4x^4 + y^{10}}$.

(1) Montrer que $\lim_{t\to 0} f(t, at) = 0$, pour tout $a \in \mathbb{R}$.

(2) Montrer que f n'admet pas de limites en $(0, 0)$.

corrGeomAnal-0045

Correction of the exercise 331

Nous avons

$$f(t, at) = \frac{t^7a^5}{4t^4 + a^{10}t^{10}} = \frac{t^3a^5}{4 + a^{10}t^6} \to 0. \tag{91.505}$$

Il y a plusieurs façons de trouver des chemins sur lesquels la limite ne vaut pas zéro. Une façon (pas la plus simple) est de chercher un chemin de la forme $\gamma(t) = (t, a(t)t)$ et résoudre l'équation

$$\frac{t^3a^5}{4 + a^{10}t^6} = \alpha \tag{91.506}$$

pour trouver $a(t)$. La résolution est un peu longuette, et en posant $y = a^5$ nous trouvons

$$y = \frac{t^5 \pm \sqrt{t^6(1 - 16\alpha^2)}}{2\alpha t^6}. \tag{91.507}$$

Si α est assez petit cela existe et définit bien $a(t)$. Le long du chemin ainsi construit nous aurons $\lim_{t\to 0} f(t, a(t)t) = \alpha$.

Sinon avec un peu de flair on peut essayer $(t, t^{2/5})$ pour qui la limite est 1/5.

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2*y**5/(4*x**4+y**10)
sage: t=var('t')
sage: f(t,t**(5/2))
t^(29/2)/(t^25 + 4*t^4)
sage: f(t,t**(5/2)).limit(t=0)
0
sage: f(t,t**(2/5)).limit(t=0)
1/5
```

5. Pouvez-vous justifier cela en lisant l'énoncé ?

Exercise 332 exoGeomAnal-0046

Soit $f : \mathbb{R}^3 \to \mathbb{R}$ la fonction $f(x, y, z) = x^2 + \sin(yz) + ye^x$.

(1) Trouver les dérivées partielles de f.

(2) Calculer le gradient de f au point $(1, \pi, 1/2)$.

(3) Trouver la dérivée de f suivant le vecteur $\vec{v} = (1/2, 1/2, 1/3)$ au point $(1, \pi, 1/2)$.

(4) Justifier la différentiabilité de f sur $\mathbb{R}^3$.

corrGeomAnal-0046

Correction of the exercise 332

Les dérivées partielles sont

$$\frac{\partial f}{\partial x} = 2x + ye^x \tag{91.508a}$$
$$\frac{\partial f}{\partial y} = z\cos(yz) + e^x \tag{91.508b}$$
$$\frac{\partial f}{\partial z} = y\cos(yz). \tag{91.508c}$$

Le gradient consiste à mettre les dérivées partielles en colonne :

$$\nabla f(1, \pi, \frac{1}{2}) = \begin{pmatrix} \pi e + 2 \\ e \\ 0 \end{pmatrix}. \tag{91.509}$$

De nombreux étudiants se sont lancés dans le calcul de la dérivée directionnelle avec la définition en termes de limites. La réponse est évidemment correcte, mais ce n'est pas le plus simple.

Il ne faut utiliser les définitions des limites, dérivées ou différentielles qu'aux points où il y a des problèmes : bord de domaine, dérivée partielles non continues, des infinis, ...

Dans les cas où la fonction est C^∞ (polynômes, sinus, cosinus et composées), le plus simple est d'utiliser les dérivées partielles et d'utiliser les formules du lemme 12.255. Dans notre cas, nous écrivons la dérivée directionnelle en tant que produit scalaire avec le gradient :

$$\frac{\partial f}{\partial v} = \nabla f \cdot v = (\pi e + 2)\frac{1}{2} + e\frac{1}{2} + 0\frac{1}{3} = \frac{1}{2}(\pi e + 2) + \frac{e}{2}. \tag{91.510}$$

Pour justifier la différentiabilité, il y a plusieurs façons. Dans l'ordre de simplicité :

(1) Il s'agit d'une composée de fonctions différentiables (sans dénominateurs qui s'annulent).

(2) Les dérivées partielles sont continues partout, et en particulier au voisinage de n'importe quel point. La proposition 12.295 conclut alors que la fonction est différentiable en chaque point.

(3) Le candidat différentielle est

$$T(u) = (2x + ye^x)u_1 + (z\cos(yz) + e^x)u_2 + (y\cos(yz))u_3. \tag{91.511}$$

En substituant cela dans la définition, et en calculant la limite, nous obtenons zéro.

C'est la seconde méthode qui est « conseillée ». En effet la première est trop simpliste et risque de provoquer des erreurs de temps en temps. La troisième est trop sophistiquée et amène des calculs dans lesquels il est facile de se tromper. Il est cependant important de comprendre les trois méthodes parce qu'il y a des exercices dans lesquels il n'y a pas de choix.

Exercise 333 exoGeomAnal-0047

(1) Trouver, si elle existe, une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ qui satisfait les conditions suivantes :

$$\begin{aligned} \partial_x f(x, y) &= \cos(x) + y^2, \\ \partial_y f(x, y) &= y\sin(x) + 2y, \\ f(\pi, 2) &= 4. \end{aligned} \tag{91.512}$$

(2) Trouver le gradient au point $(0,0)$ d'une fonction $g : \mathbb{R}^2 \to \mathbb{R}$ telle que $g(x,2x) = 2x$, et $g(x,-x) = -2x$.

corrGeomAnal-0047

Correction of the exercise 333

(1) Il n'y a pas de fonctions répondant aux conditions parce qu'elle devrait satisfaire

$$\frac{\partial^2 f}{\partial y \partial y} = 2y \tag{91.513}$$

et

$$\frac{\partial^2 f}{\partial x \partial y} = 0. \tag{91.514}$$

Ces deux fonctions n'étant pas égales, il n'y a pas de fonctions f.

Remarque 91.1.
Beaucoup d'étudiants se sont lancés avec plus ou moins de succès dans l'intégration des deux conditions. On obtient alors

$$f(x,y) = -\sin(x) + xy^2 + \alpha(y) \tag{91.515a}$$
$$f(x,y) = y\cos(x) + 2xy + \beta(x). \tag{91.515b}$$

Il faut alors justifier qu'il n'existe pas de fonctions α et β qui rendent égales ces deux expressions.
Dans ce cas ci, cela n'était pas la méthode la plus simple.

(2) Ici le « truc » est de calculer les dérivées partielles de g dans les directions $u = (1,2)$ et $v = (1,-1)$. D'abord $g(0,0) = 0$, ensuite

$$\frac{\partial g}{\partial u}(0,0) = \lim_{t\to 0} \frac{g(t,2t) - g(0,0)}{t} = \lim_{t\to 0} \frac{2t}{t} = 2 \tag{91.516}$$

et

$$\frac{\partial g}{\partial v} = \lim_{t\to 0} \frac{g(t,-t)}{t} = -2. \tag{91.517}$$

Les dérivées directionnelles étant les combinaisons linéaires des dérivées partielles nous obtenons le système d'équations

$$\left\{ \frac{\partial f}{\partial x}(0,0) + 2\frac{\partial f}{\partial y}(0,0) = 2\frac{\partial f}{\partial x}(0,0) - \frac{\partial f}{\partial y}(0,0) = -2 \right. \tag{91.518a}$$

pour les *nombres* $\partial_x f(0,0)$ et $\partial_y f(0,0)$. Le résultat est

$$\frac{\partial f}{\partial x} = -\frac{2}{3} \quad \frac{\partial f}{\partial y} = \frac{4}{3}. \tag{91.519}$$

Nous avons donc

$$\nabla f(0,0) = -\frac{2}{3}\begin{pmatrix}1\\0\end{pmatrix} + \frac{4}{3}\begin{pmatrix}0\\1\end{pmatrix} = \frac{2}{3}\begin{pmatrix}-1\\3\end{pmatrix}. \tag{91.520}$$

Exercise 334 exoGeomAnal-0048
Calculer les intégrales suivantes.

(1) $\int\int_D (x+y)^3\, dA$, où D est la partie bornée de $\mathbb{R}^2$ délimitée par les droites : $y = 0$, $x = 2$, $x = y$.

(2) L'aire à l'intérieur de la cardioïde $r(\theta) = 1 + \cos(\theta)$.

(3) Le volume du cylindre rectangle de hauteur 3 qui a comme base inférieure le cercle $(x+1)^2 + y^2 = 1$ dans le plan $z = 0$. Passer aux coordonnées cylindriques. La réponse $Volume = 3\pi$ ne vaut pas de points si on ne calcule pas l'intégrale multiple en coordonnées cylindriques.

corrGeomAnal-0048

Correction of the exercise 334

(1) La variable x va de 0 à 2 tandis que pour chaque x, la variable y va de 0 à x. Nous devons donc effectuer l'intégrale

$$\int_0^2 \int_0^x (x+y)^3 dy\, dx. \tag{91.521}$$

Cela est un peu de calcul et la réponse est 24 :

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=(x+y)**3
sage: f.integrate(y,0,x).integrate(x,0,2)
(x, y) |--> 24
```

(2) Les bornes d'intégration sont $\theta\colon 0 \to 2\pi$ et $r\colon 0 \to 1+\cos(\theta)$. L'intégrale en coordonnées polaires est donc

$$\int_0^{2\pi} d\theta \int_0^{1+\cos(\theta)} r\, dr\, d\theta. \tag{91.522}$$

Ne pas oublier le jacobien qui vaut r en coordonnées polaires. Une fois encore c'est un peu de calcul et la réponse est $\frac{3}{2}\pi$:

```
sage: f(r,t)=r
sage: f.integrate(r,0,1+cos(t)).integrate(t,0,2*pi)
(r, t) |--> 3/2*pi
```

(3) L'équation $(x+1)^2+y^2=1$ décrit un cercle centré en $(-1,0)$. Le volume étant invariant par translation nous allons considérer le cylindre de même hauteur, mais dont la base est centrée en $(0,0)$. En coordonnées cylindriques,

$$\begin{cases} x = r\cos(\theta) & \text{(91.523a)} \\ y = r\sin(\theta) & \text{(91.523b)} \\ z = z, & \text{(91.523c)} \end{cases}$$

les équations qui définissent le cylindre sont $r \leqslant 1$, $0 \leqslant z \leqslant 3$. Nous avons donc l'intégrale

$$V = \int_0^1 dr \int_0^3 dz \int_0^{2\pi} d\theta\, r. \tag{91.524}$$

Le résultat est celui attendu :

```
sage: f(r,theta,z)=r
sage: f.integrate(theta,0,2*pi).integrate(z,0,3).integrate(r,0,1)
(r, theta, z) |--> 3*pi
```

Exercise 335 exoGeomAnal-0049

Démontrer la proposition suivante.

Soit $U \subset \mathbb{R}^m$ un ouvert connexe par arcs. Soit $f : U \to \mathbb{R}^n$. Les conditions suivantes sont équivalentes :

(1) f est constante ;

(2) les dérivées partielles $\partial_1 f, \dots, \partial_m f$ existent et sont nulles sur U ;

(3) f est différentiable et $df(a) = 0$ pour tout $a \in U$.

corrGeomAnal-0049

Correction of the exercise 335

Il s'agit de la proposition 12.315 dont la clef est le théorème des accroissements finis 12.314.

91.7 Exercices pour aller plus loin

Exercise 336 exoGeomAnal-0001

Montrer que la définition d'arc géométrique compact a un sens, c'est-à-dire que si (I, γ) est un arc compact, alors tous les arcs équivalents sont également compacts. corrGeomAnal-0001

Correction of the exercise 336

<+CorrGeomAnal-0001+>

Exercise 337 exoGeomAnal-0002

Durant la démonstration du théorème 21.39, nous avons supposé que θ était strictement croissante. Cela avait permis de rentrer $\theta'(s)$ dans la norme. Comment doit on adapter la preuve si on suppose que θ est strictement décroissante ? corrGeomAnal-0002

Correction of the exercise 337

<+CorrGeomAnal-0002+>

Exercise 338 exoGeomAnal-0003

Nous avons montré durant l'exemple 21.44 que les angles exprimés en radians forment des coordonnées normales pour le cercle. Quid des degrés ? Refaire l'exemple en supposant un paramétrage $f(t) = \big(\cos(t), \sin(t)\big)$ où t représente maintenant l'angle en degré : $t \in [0, 360]$. Avons nous encore un paramétrage normal ? corrGeomAnal-0003

Correction of the exercise 338

<+CorrGeomAnal-0003+>

Exercise 339 exoGeomAnal-0004

Pourquoi l'équation (21.90) est ce qu'on obtient en posant $\theta'(t) = 1$ dans l'équation (21.86) ? corrGeomAnal-0004

Correction of the exercise 339

<+CorrGeomAnal-0004+>

Exercise 340 exoGeomAnal-0005

Prouver que l'application $\|.\|_\infty$ sur $\mathbb{R}^3$ donnée par

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant 3} |x_i| \tag{91.525}$$

est bien une norme. Dessiner la boule unité qui lui correspond. corrGeomAnal-0005

Correction of the exercise 340

Nous avons quatre propriétés à vérifier. Dans la suite v est un élément de $\mathbb{R}^3$.

(1) L'application $\|.\|_\infty$ sur $\mathbb{R}^3$ prends ses valeurs en $\mathbb{R}^+$:
   - la valeur absolue est une fonction non négative : pour tout r dans $\mathbb{R}$ on a $|r| = r$ si r est positif et $|r| = -r$ si r est négatif. Il est clair que $|r| = 0$ si et seulement si $r = 0$;
   - le maximum parmi trois nombres $\geqslant 0$ ne peut que être $\geqslant 0$.

(2) $\|v\|_\infty = 0$ si et seulement si $v = (0, 0, 0)^t$.
   - $\|(0, 0, 0)^t\|_\infty = \max\{0, 0, 0\} = 0$;
   - soit v un vecteur non nul tel que $\|v\|_\infty = 0$. Si v est non nul au moins une parmi les composantes de v est non nulle. Sans perdre en généralité on dira qu'il s'agit de v_1. Alors $|v_1| > 0$. La définition de maximum nous dit que $|v_1| \leqslant \|v\|_\infty$. On a obtenu la suite d'inégalités $0 \geqslant v_1 > 0$. Absurde.

(3) Pour tout λ in $\mathbb{R}$ on a $\|\lambda v\|_\infty = |\lambda|\|v\|_\infty$.
— soit $v = (v_1, v_2, v_3)^t$, alors $\lambda v = (\lambda v_1, \lambda v_2, \lambda v_3)^t$. Nous savons que $|\lambda v_i| = |\lambda||v_i|$ et que la multiplication fois un nombre positif ne change pas le sens des inégalités : si a et b sont dans $\mathbb{R}$ et $a < b$, alors $|\lambda|a < |\lambda|b$. On obtient alors

$$\|\lambda v\|_\infty = \max\{|\lambda||v_1|, |\lambda||v_2|, |\lambda||v_3|\} = |\lambda| \max\{|v_1|, |v_2|, |v_3|\} = |\lambda|\|v\|_\infty.$$

(4) Inégalité triangulaire : $\|v + w\|_\infty \leqslant \|v\|_\infty + \|w\|_\infty$.
— soient v et w deux vecteurs dans $\mathbb{R}^3$. La somme $v + w$ est un vecteur de composantes $(v_1 + w_1, v_2 + w_2, v_3 + w_3)$. On a alors

$$\|v + w\|_\infty = \max_{1\leqslant i\leqslant 3}\{|v_i + w_i|\}.$$

Pour fixer les idées nous pouvons supposer que le maximum soit $|v_1 + w_1|$. Une propriété de la valeur absolue nous dit que $|v_1 + w_1| \leqslant |v_1| + |w_1|$. Pour conclure il nous suffit de remarquer que $|v_1| \leqslant \max\{|v_1|, |v_2|, |v_3|\} = \|v\|_\infty$ et, de même, $|w_1| \leqslant \|w\|_\infty$.

La boule unité est un cube centré dans l'origine et de côté 2.

Exercise 341 exoGeomAnal-0006

Montrer que la frontière d'une partie A d'un espace vectoriel normé peut s'exprimer de la façon suivante :

$$\partial A = \complement\big(\operatorname{Int}(A) \cup \operatorname{Int}(\complement A)\big). \tag{91.526}$$

corrGeomAnal-0006

Correction of the exercise 341

Soit A une partie d'un espace vectoriel normé. Un point P est dans la frontière de A si pour tout $r > 0$ dans $\mathbb{R}^+$ la boule $B(P, r)$ de centre P et rayon r a intersection non nulle avec A et avec son complémentaire $\complement A$. La boule $B(P, r)$ contient des points de A et des points de $\complement A$

Cela veut dire qu'il n'existe pas un seul $\bar{r}$ dans $\mathbb{R}^+_{>0}$ tel que $B(P, \bar{r}) \subset \operatorname{Int} A$, parce que l'intersection entre $\operatorname{Int}(A)$ et $\complement A$ est nulle. Le point P est donc dans le complémentaire de $\operatorname{Int}(A)$. De même on voit que P est dans le complémentaire de $\operatorname{Int}(\complement A)$. Le premier lemme dans l'annexe A nous permet de conclure

$$\complement\left(\operatorname{Int}(\complement A)\right) \cap \complement\left(\operatorname{Int}(A)\right) = \complement\left(\operatorname{Int}(\complement A) \cup \operatorname{Int}(A)\right).$$

Exercise 342 exoGeomAnal-0007

Les ensembles suivants sont-ils des pavés ? Écrire les pavés comme produits d'intervalles.

(1) Le graphe de $f(x) = 2x$ dans $\mathbb{R}^2$.
(2) $\{(3, y) \text{ tel que } y \in \mathbb{R}\} \subset \mathbb{R}^2$.
(3) $\{(3, y) \text{ tel que } y \in]1, 7]\} \subset \mathbb{R}^2$.
(4) $\{(x, y, 0) \text{ tel que } x \in \mathbb{R},\ y \in [0, 1]\} \subset \mathbb{R}^3$.

corrGeomAnal-0007

Correction of the exercise 342

<+CorrGeomAnal-0007+>

Exercise 343 exoGeomAnal-0008

Montrer que $\bar{A}$ est le plus petit fermé qui contient A.

corrGeomAnal-0008

Correction of the exercise 343

Soit B un fermé tel que $A \subset B$. Nous voulons prouver que $\bar{A} \subset B$. Soit $a \in \bar{A}$; par le corolaire 7.254, il existe une suite (x_n) dans A telle que $a = \lim a_n$. Mais $a_n \in A \subset B$, donc la suite (a_n) est une suite dans B, et sa limite appartient alors à $\bar{B} = B$. Nous avons donc montré que $a \in B$, et par conséquent que $\bar{A} \subset B$, ce qu'il fallait démontrer.

Exercise 344 exoGeomAnal-0009

Montrer que si A est une partie d'un espace vectoriel normé V et si B est une partie d'un espace vectoriel normé W, alors $\overline{A \times B} = \bar{A} \times \bar{B}$.

corrGeomAnal-0009

Correction of the exercise 344

Nous commençons par prouver que $\overline{A\times B}\subset\bar{A}\times\bar{B}$. Un élément (a,b) dans $\overline{A\times B}$ peut toujours être vu comme la limite d'une suite dans $A\times B$. Soit donc (a_n,b_n) une suite telle que $\lim(a_n,b_n)=(a,b)$ avec $a_n\in A$ et $b_n\in B$. Par le lemme 10.78, nous avons une convergence « composante par composante » : $a_n\to a$ et $b_n\to b$. Mais la suite (a_n) est contenue dans A, donc sa limite, a, est dans $\bar{A}$. De la même manière, $b\in\bar{B}$. Par conséquent $(a,b)\in\bar{A}\times\bar{B}$.

Il faut maintenant prouver l'inclusion inverse. Soit $(a,b)\in\bar{A}\times\bar{B}$. Nous avons des suites (a_n) et (b_n) dans A et B respectivement qui convergent vers a et b. En utilisant à nouveau le lemme 10.78 (mais cette fois dans le sens inverse), nous avons

$$\lim\underbrace{(a_n,b_n)}_{\in A\times B}=(a,b)\in\overline{A\times B}.\tag{91.527}$$

La dernière appartenance est simplement le fait qu'une suite convergente contenue dans $A\times B$ converge dans $\overline{A\times B}$.

Exercise 345 exoGeomAnal-0010

On considère l'ensemble $A=\{\frac{1}{n}$ tel que $n\in\mathbb{N}_0\}$. Déterminer les points d'accumulation de A dans $\mathbb{R}$. Montrer que A n'est ni fermé ni ouvert. corrGeomAnal-0010

Correction of the exercise 345

<+CorrGeomAnal-0010+>

Exercise 346 exoGeomAnal-0011

Soient A et B deux ensembles de $\mathbb{R}^N$. Montrer que :

(1) $\overline{A\cup B}=\overline{A}\cup\overline{B}$.

(2) $\overline{A\cap B}\subset\overline{A}\cap\overline{B}$ et donner un exemple où cette inclusion est stricte.

(3) L'intérieur de $A\cap B$ est égal à l'intersection des intérieurs de A et B. Que peut-on affirmer de l'intérieur de $A\cup B$?

(4) L'intérieur de $\complement A$ est $\complement\overline{A}$.

corrGeomAnal-0011

Correction of the exercise 346

Démontrer que la fonction $f:\mathbb{R}^2\to\mathbb{R}$

$$f(x,y)=\begin{cases}(x^2+y^2)\sin\left(\frac{1}{x^2+y^2}\right) & \text{si } (x,y)\neq(0,0),\\ 0, & \text{si } (x,y)=(0,0)\end{cases}\tag{91.528}$$

est différentiable à tout point de $\mathbb{R}^2$ (et par conséquence est continue partout), mais que ses dérivées partielles ne sont pas définies en $(0,0)$.

Cet exercice est assez difficile. Sur $\mathbb{R}^2\backslash(0,0)$ on peut utiliser le la proposition 12.295. Pour prouver la différentiabilité en l'origine il faut utiliser la définition.

Exercise 347 exoGeomAnal-0012

Démontrer la règle de différentiation des produits en utilisant la règle de la chaine et le lemme 12.274. Indice : trouver une forme bilinéaire B telle que $(fg)(x)=B\big(f(x),g(x)\big)$, et ensuite trouver une fonction $s\colon\mathbb{R}^m\to\mathbb{R}^n\times\mathbb{R}^n$ telle que $fg=B\circ s$. corrGeomAnal-0012

Correction of the exercise 347

<+CorrGeomAnal-0012+>

Exercise 348 exoGeomAnal-0013

Nous avons donné une démonstration du théorème de Weierstrass 7.146 (toute fonction sur un compact de $\mathbb{R}^m$ est bornée et atteint ses bornes). Pouvons nous prouver le cas plus général où on remplace $\mathbb{R}^m$ par un espace vectoriel normé en nous contentant de recopier mot à mot en remplaçant $\mathbb{R}^m$ par V ? Quelles sont les étapes qui demandent un surplus de justification ?

Cet exercice est difficile. Effectuer lesdits « surplus de justification » l'est encore plus parce qu'il faut vérifier quels sont les lemmes et propositions antérieures qui tiennent en remplaçant $\mathbb{R}^m$

par V dans les énoncés et les démonstration. Sachez pour votre culture générale que quasiment tout ce que nous disons sur les fonctions continues de $\mathbb{R}^m$ dans $\mathbb{R}$ est valable à peu près mot à mot pour des fonctions continues d'un espace vectoriel normé quelconque dans $\mathbb{R}$ (et même souvent vers un autre espace vectoriel normé quelconque). corrGeomAnal-0013

Correction of the exercise 348

<+CorrGeomAnal-0013+>

Exercise 349 exoGeomAnal-0014

Donner un exemple de sous-ensemble A de $\mathbb{R}^3$ tel que $\bar{A}$ n'est pas égal à l'ensemble des points d'accumulation de A. corrGeomAnal-0014

Correction of the exercise 349

<+CorrGeomAnal-0014+>

Exercise 350 exoCalculDifferentiel0005

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction C^1.

(1) Montrer que si f admet un minimum global au point (x_0, y_0) (c'est-à-dire que pour tout $(x, y) \in \mathbb{R}^2$, nous avons $f(x, y) \geqslant f(x_0, y_0)$), alors
$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0. \tag{91.529}$$
Pour ce faire, considérer les fonctions d'une variable réelle $f_1(t) = f(x_0 + t, y_0)$ et $f_2(t) = f(x_0, y_0 + t)$.

(2) Montrer que $f(x, y) = x^2 + 2xy + y^2 + x$ n'admet pas de minimum.

(3) Montrer que $f(x, y) = x^2 - y^2$ n'admet pas de minimum.

(4) Montrer que si $f(x, y) = 2x^2 + 2xy + y^2 + x + 3y$ admet un minimum alors ce minimum est atteint au point $(x_0, y_0) = (1, -5/2)$. Montrer que f admet effectivement un minimum en ce point.
Indice : remarquer que pour tous réels h et k, $f(1 + h, -5/2 + k) - f(1, -5/2) \geqslant 0$. corrCalculDifferentiel0005

Correction of the exercise 350

<+CorrCalculDifferentiel0005+>

Exercise 351 exoCalculDifferentiel0014

Résoudre l'équation
$$2xy\frac{\partial f}{\partial x}(x, y) + (1 + y^2)\frac{\partial f}{\partial y}(x, y) = 0 \tag{91.530}$$
où $f\colon \{(x, y) \in \mathbb{R}^2 \text{ tel que } x > 0\} \to \mathbb{R}$ est une fonction de classe C^1. On pourra effectuer le changement de variables $x = (u^2 + v^2)/2$, $y = u/v$ avec $v > 0$. corrCalculDifferentiel0014

Correction of the exercise 351

<+CorrCalculDifferentiel0014+>

Exercise 352 exoCalculDifferentiel0018

Une application $f : U \to \mathbb{R}$ de classe C^2 sur un ouvert U de $\mathbb{R}^3$ est dite **harmonique** si et seulement si $\Delta f = 0$, où
$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{91.531}$$
est le laplacien de f.

(1) Montrer que si f est harmonique sur $\mathbb{R}^2$ et de classe C^3 alors $\frac{\partial f}{\partial x}$ et $y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}$ sont harmoniques. Note : deviner quelle est la formule du laplacien sur $\mathbb{R}^2$.

(2) Vérifier que l'application $f\colon \mathbb{R}^3 \to \mathbb{R}$ définie par
$$f(x, y, z) = \arctan\frac{y}{x} + \arctan\frac{z}{y} + \arctan\frac{x}{z} \tag{91.532}$$
est harmonique sur $\mathbb{R}^3$.

corrCalculDifferentiel0018

Correction of the exercise 352

<+CorrCalculDifferentiel0018+>

Chapitre 92

Exercices d'analyse numérique (Louvain-la-Neuve)

Exercise 353 exoSerieUn0002

Donner un exemple de problème stable de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^0(\mathbb{R})\backslash C^1(\mathbb{R})$. Justifier. corrSerieUn0002

Correction of the exercise 353

La fonction

$$x(d) = \begin{cases} 0 & \text{si } x \geqslant 0 \\ x & \text{si } x > 0 \end{cases} \tag{92.1}$$

est continue, mais pas C^1 (non dérivable en $x = 0$). La dérivée est partout bornée par 1, et donc le problème est stable.

Un autre exemple très classique serait de prendre $x(d) = |d|$. Dans ce cas, on peut prendre n'importe que η et $K = 1$. Le calcul est que

$$|x(d) - x(d_0)| < K|d - d_0| \tag{92.2a}$$
$$\big||d| - |d_0|\big| < |d - d_0|. \tag{92.2b}$$

Cette dernière inéquation est correcte, comme on peut le voir en mettant au carré les deux membres.

Exercise 354 exoSerieUn0003

Donner un exemple de problème instable de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^0(\mathbb{R})$. Justifier corrSerieUn0003

Correction of the exercise 354

Un exemple assez classique de fonction dont la dérivée n'est pas bornée sans pour autant que la fonction aie un comportement immoral [1] est $x \mapsto \sqrt{x}$. Afin d'avoir une fonction définie sur $\mathbb{R}$ tout entier, nous regardons la fonction

$$x(d) = \sqrt{|d|}. \tag{92.3}$$

Si nous considérons maintenant $d_0 = 0$ et n'importe quel η, nous avons

$$\frac{|x(d) - x(d_0)|}{|d - d_0|} = \frac{\sqrt{d}}{d} = \frac{1}{\sqrt{d}}. \tag{92.4}$$

Il n'est pas possible de trouver un K qui majore ce rapport. Le problème est donc mal conditionné.

Attention : dans ce calcul nous avons supposé $d > 0$. Pensez à adapter au cas $d < 0$.

Exercise 355 exoSerieUn0001

Démontrer que tout problème de la forme $x = x(d)$ avec $d \in \mathbb{R}$ et $x \in C^1(\mathbb{R})$ est stable. corrSerieUn0001

1. Penser à $x \mapsto x \sin(1/x)$.

Correction of the exercise 355

Il faut démontrer qu'une fonction C^1 sur $\mathbb{R}$ vérifie automatiquement la condition 34.26(2) de la définition de la stabilité. Pour cela, remarquons qu'une fonction C^1 possède une dérivée continue, et donc bornée sur tout compact [2]

Prenons $\eta > 0$ et $d_0 \in \mathbb{R}$ et puis un d tel que $|d - d_0| < \eta$. Par le théorème des bornes atteintes, la fonction x' est bornée sur l'intervalle $[d_0 - \eta, d_0 + \eta]$. Appelons K un majorant de x' sur cet intervalle. La fonction

$$f(d) = x(d_0) + K|d - d_0| \tag{92.5}$$

majore $x(d)$, et donc on a

$$\big|x(d) - x(d_0)\big| \leqslant K|d - d_0|. \tag{92.6}$$

Attention : vérifier si ce raisonnement est correct avec $d_0 > d$, et adapter au besoin.

Exercise 356 exoSerieUn0004

Montrer que l'évaluation de $x(d) = \cos(d)$ est bien conditionnée près de $d = 0$. corrSerieUn0004

Correction of the exercise 356

Proche de zéro, la dérivé de $\cos(d)$ est petite. Nous allons donc prendre d_0 proche de 0 et un ϵ et choisir η suffisamment petit pour que

$$\frac{|\cos(d) - \cos(d_0)|}{|d - d_0|} < \epsilon \tag{92.7}$$

tant que $|d - d_0| < \epsilon$. Dans ce cas, nous avons que $K^{\eta}_{abs}(d_0) < \epsilon$, et donc que

$$K^{\eta}_{rel}(d) = \epsilon \frac{d}{\cos(d)}. \tag{92.8}$$

Lorsque d est proche de zéro, la fraction reste proche de zéro.

Ce que nous concluons est que $K^{\eta}_{rel}(d)$ peut être rendu aussi petit que l'on veut (prendre ϵ petit) en choisissant d dans un petit voisinage de 0 et η suffisamment petit pour ne pas déborder du voisinage.

Exercise 357 exoSerieUn0005

Exprimer le conditionnement relatif du problème composé $x = f(d)$ avec $f = f_1 \circ f_2$ en termes des conditionnements relatifs des problèmes composants $x_i = f_i(d_i) \quad i = 1, 2$. corrSerieUn0005

Correction of the exercise 357

Le conditionnement absolu est donné par la norme de la dérivée. En ce qui concerne le conditionnement relatif, nous avons

$$\begin{aligned} K_{f_1 \circ f_2}(d) &= |(f_1 \circ f_2)'(d)| \frac{|d|}{|(f_1 \circ f_2)(d)|} \\ &= f_1'\big(f_2(d)\big) g'(d) \frac{|d|}{|(f_1 \circ f_2)(d)|} \end{aligned} \tag{92.9}$$ EqSUZCKfg

Mais le conditionnement de f_1 au point $f_2(d)$ est donné par

$$K_{f_1}\big(f_2(d)\big) = f_1'\big(f_2(d)\big) \frac{|f_2(d)|}{|f_1\big(f_2(d)\big)|}, \tag{92.10}$$

nous pouvons faire apparaitre cette expression dans (92.9) en multipliant et divisant par $f_2(d)$. Ainsi nous avons

$$K_{f_1 \circ f_2}(d) = K_{f_1}\big(f_2(d)\big) f_2'(d) \frac{|d|}{f_2(d)} = K_{f_1}\big(f_2(d)\big) K_{f_2}(d). \tag{92.11}$$

2. Un compact est un ensemble fermé et borné, typiquement un intervalle du type $[a, b]$.

Cependant, pour cette formule nous avons besoin que les fonctions f_1 et f_2 soient telles que l'approximation du conditionnement par la dérivée fonctionne.

Exercise 358 exoSerieUn0006

Que peut on dire du conditionnement du problème consistant a retrancher un nombre d de 1 ? corrSerieUn0006

Correction of the exercise 358

Le conditionnement absolu est simple avec $x(d) = 1 - d$:

$$\sup_{|d-d_0|<\eta} \frac{|x(d) - x(d_0)|}{|d - d_0|} = 1. \tag{92.12}$$

Le conditionnement relatif est donc

$$K_{rel}^{\eta}(d) = \frac{|d|}{|d-1|}. \tag{92.13}$$

C'est donc autour de $d = 0$ que le problème est bien conditionné.

Exercise 359 exoSerieUn0007

Calculer le conditionnement relatif pour

(1) $x - a^d = 0$ avec $a > 0$.

(2) $d - x + 1 = 0$.

corrSerieUn0007

Correction of the exercise 359

(1) Nous supposons que a est une constante et nous devons calculer

$$K = \sup_{|d-d_0|<\eta} \frac{|a^d - a^{d_0}|}{|d - d_0|}. \tag{92.14}$$

Si $a > 1$, l'expression dans le supremum est croissante en d[3] et donc le supremum est atteint lorsque $d = d_0 + \eta$ et le conditionnement absolu vaut

$$K_{abs}^{\eta} = \frac{|a^{d_0}(a^{\eta} - 1)|}{\eta}. \tag{92.15}$$

Dans ce cas, le conditionnement relatif vaut

$$K_{rel}(d) = \frac{|d(a^{\eta} - 1)|}{\eta}. \tag{92.16}$$

Nous pouvons obtenir une majoration de K_{rel} en utilisant le théorème des accroissements finis. En effet, il existe un $c \in [d, d_0]$ tel que

$$a^d = a^{d_0} + f'(c)(d - d_0) \tag{92.17}$$

où $f(x) = a^x$ et $f'(c) = \ln(a)a^c$. Donc

$$K_{abs} = \sup_{|d-d_0|<\eta} \ln(a)a^c. \tag{92.18}$$

Lorsque $a > 1$, la fonction $\ln(a)a^c$ est croissante en c et donc nous avons

$$K_{abs}(d_0) \leqslant \ln(a)a^{d_0+\eta} \tag{92.19}$$

et

$$K_{rel}(d) \leqslant \ln(a)da^{\eta}. \tag{92.20}$$

3. Je ne vois pas pourquoi, mais en la traçant, ça a l'air d'être le cas.

Il y a encore une troisième façon de travailler, en utilisant le corolaire 34.33. Nous savons alors que $K_{\text{abs}}(d) \simeq |x'(d)|$, c'est-à-dire

$$K_{\text{abs}}(d)^\eta \simeq |\ln(a)a^d| \tag{92.21}$$

lorsque η est petit. Le conditionnement relatif vaut alors

$$K_{\text{rel}} \simeq d|\ln(a)|. \tag{92.22}$$

Donc plus a est proche de 1, mieux le problème est conditionné, et plus d est petit, mieux c'est aussi.

(2) Lorsque $x(d) = d + 1$, nous avons

$$K_{abs}^\eta(d_0) = \sup_{|d-d_0|<\eta} \frac{|x(d) - x(d_0)|}{|d - d_0|} = 1, \tag{92.23}$$

et donc

$$K_{rel}^\eta = \frac{|d|}{|d+1|}. \tag{92.24}$$

Exercise 360 exoSerieUn0008

Étudier le conditionnement de la formule par radicaux donnant la plus grande racine de $P(x) = x^2 + 2px - q$. corrSerieUn0008

Correction of the exercise 360

Nous voulons utiliser le corolaire 34.33. Pour cela, nous devons d'abord montrer que le problème est stable. De la même manière que, avant, nous prouvions la stabilité en montrant que la dérivée de x est bornée, ici nous allons argumenter que la norme du gradient est bornée pour dire que le problème est stable [4].

Supposons que $q > -p^2$. La plus grande des deux racines de $x^2 + 2px - q = 0$ est donnée par

$$x(p, q) = -p + \sqrt{p^2 + q}. \tag{92.25}$$

Nous avons donc

$$\frac{\partial x}{\partial p} = -1 + \frac{p}{\sqrt{p^2+q}} = \frac{-x(p,q)}{\sqrt{p^2+q}} \tag{92.26}$$

et

$$\frac{\partial x}{\partial q} = \frac{1}{2\sqrt{p^2+q}}, \tag{92.27}$$

et la norme du gradient est donc

$$\|\nabla x(p_0, q_0)\| = \frac{\sqrt{4x^2(p_0,q_0)+1}}{2\sqrt{p_0^2+q_0}}. \tag{92.28}$$

Cette quantité reste bornée au voisinage de tout point vérifiant $q > -p^2$. Nous pouvons donc utiliser le corolaire 34.33. En fait, en calculant la norme de ∇x, nous avons fait coup double : d'une part nous avons validé les hypothèses du corolaire, et d'autre part, nous en avons calculé la conclusion.

Pour η suffisamment petit nous avons donc

$$K_{\text{abs}}\big((p_0, q_0), \eta\big) \sim \|\nabla x(p_0, q_0)\| = \frac{\sqrt{4x^2(p_0,q_0)+1}}{2\sqrt{p_0^2+q_0}}. \tag{92.29}$$

4. Nous n'allons pas entrer dans les détails de la preuve que la stabilité est impliquée par la borne sur le gradient, mais c'est un bon exercice.

Le conditionnement relatif vaut quant à lui

$$K_{\text{rel}}\big((p_0,q_0),\eta\big) = K_{\text{abs}}\big((p_0,q_0),\eta\big)\frac{\sqrt{p_0^2+q_0^2}}{|x(p_0,q_0)|} \sim \frac{\sqrt{4x^2(p_0,q_0)+1}}{|x(p_0,q_0)|}\frac{\sqrt{p_0^2+q_0^2}}{2\sqrt{p_0^2+q_0}}. \tag{92.30}$$

On en déduit que le problème est mal conditionné quand q_0 est proche de $-p_0^2$, c.-à-d. quand les deux racines sont proches l'une de l'autre, ou quand $x(p_0,q_0)=0$ (pourquoi ? ?).

Exercise 361 exoSerieUn0009

Que peut-on dire du conditionnement de $x=(d-1)^2$? corrSerieUn0009

Correction of the exercise 361

Nous devons d'abord calculer le supremum de

$$\begin{aligned}\frac{|(d-1)^2-(d_0-1)^2|}{|d-d_0|} &= \frac{|(d+d_0-2)(d-d_0)|}{|d-d_0|}\\ &= |d+d_0-2|.\end{aligned} \tag{92.31}$$

Si $d-2>0$, alors le supremum est atteint en $d=d_0+\eta$ et alors on a

$$K_{\text{abs}} = 2d_0+\eta-2 \tag{92.32}$$

et

$$K_{\text{rel}} = \frac{d_0(2d_0+\eta-2)}{(d_0-1)^2}. \tag{92.33}$$

Si on a $d-2<0$, alors le supremum est atteint en $d=d_0-\eta$ et alors

$$K_{\text{abs}} = |2d_0-\eta-2|. \tag{92.34}$$

Notez que dans ce cas ci, nous n'avons pas de garanties sur le signe de ce qui se trouve dans la valeur absolue.

Exercise 362 exoSerieTrois0002

On considère le problème de Cauchy suivant ($a\in\mathbb{R}$) :

$$\left\{\begin{array}{rcll} x'(t) &=& x_0e^{at}(a\cos t-\sin t) & t>0\\ x(0) &=& x_0. & \end{array}\right. \tag{92.35}$$

Étudier le conditionnement de ce problème en fonction de la donnée initiale. corrSerieTrois0002

Correction of the exercise 362

Nous allons noter f au lieu de x la fonction recherchée parce que sinon il y a des problèmes de notation à n'en plus finir parce que la donnée est notée x_0. Nous nous voyons mal écrire des expressions comme $x(x)(t)$. Au lieu de ça, nous écrivons $f_{x_0}(t)$ la solution au temps t du problème avec x_0 comme paramètre.

La solution exacte du problème est donnée par

$$f_{x_0}(t) = x_0e^{at}\cos(t). \tag{92.36}$$

Pour trouver le conditionnement absolu, il faut étudier le rapport

$$\frac{\|f_x-f_{x_0}\|}{|x-x_0|} \tag{92.37}$$

où la norme du numérateur est la norme sur l'espace de fonctions considéré (par exemple la norme supremum). Nous avons que

$$f_x(t)-f_{x_0}(t) = (x-x_0)e^{at}\cos(t), \tag{92.38}$$

et donc

$$\|f_x - f_{x_0}\| = |x - x_0| \cdot \|t \mapsto e^{at}\cos(t)\| \tag{92.39}$$

parce que les normes sont $\mathbb{R}$-linéaire nonobstant une valeur absolue. Nous trouvons donc

$$K_{\text{abs}}(x_0) = \|t \mapsto e^{at}\cos(t)\|. \tag{92.40}$$

Notons que cela est infini pour la norme supremum si $a > 0$ et si on considère $\mathbb{R}$ comme domaine de variation de t. Quoi qu'il en soit, pour le conditionnement relatif,

$$\|f_{x_0}\| = |x_0| \cdot \|t \mapsto e^{at}\cos(t)\|, \tag{92.41}$$

et donc

$$K_{\text{rel}}(x_0) = \|t \mapsto e^{at}\cos(t)\| \cdot \frac{|x_0|}{\|f_{x_0}\|} = 1. \tag{92.42}$$

Au passage, il y a une simplification par l'infini si $a > 0$.

Exercise 363 exoSerieDeux0001

Démontrer que la norme opérateur $|\ |_{\text{op}}$ définit bien une norme sur l'espace $\text{Lin}(\mathbb{R}^2, \mathbb{R}^2)$ des applications linéaires de $\mathbb{R}^2$ vers $\mathbb{R}^2$. corrSerieDeux0001

Correction of the exercise 363

Nous devons vérifier les trois conditions de la définition 7.158.

(1) Si $\sup_{|x|=1}\{|\alpha(x)|\} = 0$, c'est que $\alpha(x) = 0$ pour tout x de norme 1, et nous en déduisons que $\alpha(x) = 0$ pour tout x.

(2) Nous avons

$$\begin{aligned} \|\lambda\alpha\| &= \sup\{|(\lambda\alpha)(x)|\} \\ &= \sup\{|\lambda| \cdot |\alpha(x)|\} \\ &= |\lambda| \sup\{|\alpha(x)|\} \\ &= |\lambda| \cdot \|\alpha\|. \end{aligned} \tag{92.43}$$

(3) Nous avons

$$\begin{aligned} \|\alpha + \beta\| &= \sup_{|x|=1}\{|\alpha(x) + \beta(x)|\} \\ &\leqslant \sup\{|\alpha(x)| + |\beta(x)|\} \\ &\leqslant \sup\{|\alpha(x)|\} + \sup\{|\beta(x)|\} \\ &= \|\alpha\| + \|\beta\|. \end{aligned} \tag{92.44}$$

Dans ces calculs, notez qu'on a utilisé le fait que la valeur absolue est une norme sur $\mathbb{R}$.

Exercise 364 exoSerieDeux0002

Calculer $|\alpha|_{\text{op}}$ lorsque α est ItemExoDeuxDeuxa

(1) une homothétie de rapport $\lambda \neq 0$ dans $\mathbb{R}^n$,

(2) une rotation d'angle θ centrée en l'origine du plan $\mathbb{R}^2$,

(3) une rotation d'angle θ autour de l'axe supporté par $(1, 1, 1)$ dans $\mathbb{R}^3$. corrSerieDeux0002

Correction of the exercise 364

(1) Si $\alpha(x) = \lambda x$, nous devons faire le calcul

$$\sup_{|x|=1}\{|\lambda x|\} = |\lambda|. \tag{92.45}$$

(2) Si A est une rotation, par définition la norme de Ax est la même que celle de x, donc

$$\sup_{|x|=1}\{\|Ax\|\} = 1. \tag{92.46}$$

Il est à noter que la norme $\|.\|$ dans le supremum est la norme sur $\mathbb{R}^2$.

(3) Si nous prenons une base orthonormée dans laquelle un des vecteurs est $(1,1,1)$, nous retombons sur le cas précédent. Souvenez vous qu'on peut passer d'une base orthonormée à une autre avec des isométries. Le changement de base n'affecte donc pas la norme dont on prend le supremum.

Exercise 365 exoSerieDeux0003

Calculer les différentielles $f_{\star p}$ (comme applications linéaires) des fonctions f suivantes aux points p donnés.

(1) $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^3$; $p = 2$.

(2) $f : \mathbb{R}^3 \to \mathbb{R} : x \mapsto \langle \alpha(x), (1,2,0) \rangle$ avec $p = (a,b,c)$, où α est la rotation autour de l'axe $(1,1,1)$.

(3) $f : \mathbb{R}^n \to \mathbb{R} : x \mapsto \langle \alpha(x), w_0 \rangle$ avec $p \in \mathbb{R}^n$, où $\alpha \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ et où $w_0 \in \mathbb{R}^n$ est un vecteur fixé.

(4) $f \colon \mathbb{R}^2 \to \mathbb{R}^3$, $f(x,y) = (x, y, x^2 + y^2)$ au point $p = (1,1)$.

corrSerieDeux0003

Correction of the exercise 365

(1) En dimension 1, la différentielle est l'application linéaire de coefficient directeur donné par la dérivée. Dans le cas de $f(x) = x^3$, nous avons que

$$df_2(x) = f'(2)x = 12x. \tag{92.47}$$

(2) L'application proposée est une application linéaire. Elle est donc égale à sa différentielle. Nous allons le voir explicitement dans l'exercice (3). ItemExoDeuxTroisc

(3) Nous supposons que l'application α s'écrive $\alpha(x) = \sum_{ij} \alpha_{ij} x_j e_i$, et nous supposons que $w_0 = \sum_k w_k e_k$. En omettant les sommes sous-entendues, nous pouvons écrire la fonction f sous la forme

$$\begin{aligned} f(x) &= \langle \alpha(x), w_0 \rangle \\ &= \langle \alpha_{ij} x_j e_i, w_k e_k \rangle \\ &= \sum_{ijk} \alpha_{ij} x_j w_k \underbrace{\langle e_i, e_k \rangle}_{=\delta_{ik}} \\ &= \sum_{ij} \alpha_{ij} x_j w_i \end{aligned} \tag{92.48}$$

Les dérivées partielles s'écrivent

$$\begin{aligned} \frac{\partial f}{\partial x_k} &= \sum_{ij} w_i \alpha_{ij} \underbrace{\frac{\partial x_j}{\partial x_k}}_{=\delta_{jk}} \\ &= \sum_i w_i \alpha_{ik}. \end{aligned} \tag{92.49}$$

Notez que ces coefficients sont constants. Cela est le fait que la différentielle d'une application linéaire est linéaire. Vérifions maintenant qu'elle est même égale à elle même. Appliquons df au vecteur $v = v_k e_k$:

$$\begin{aligned} df(v) &= \sum_k \frac{\partial f}{\partial x_k} v_k \\ &= \sum_{ki} w_i \underbrace{\alpha_{ik} v_k}_{=\alpha(v)_i} \\ &= \sum_i w_i \alpha(v)_i \\ &= \langle w, \alpha(v) \rangle. \end{aligned} \tag{92.50}$$

Nous retombons donc bien sur la fonction f de départ.

(4) La matrice de la différentielle est donnée par

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(1,1) & \frac{\partial f_1}{\partial y}(1,1) \\ \frac{\partial f_2}{\partial x}(1,1) & \frac{\partial f_2}{\partial y}(1,1) \\ \frac{\partial f_3}{\partial x}(1,1) & \frac{\partial f_3}{\partial y}(1,1) \end{pmatrix} \tag{92.51}$$

où $f_1(x,y) = x$, $f_2(x,y) = y$ et $f_3(x,y) = x^2 + y^2$. En calculant les dérivées partielles nous trouvons la matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{pmatrix}, \tag{92.52}$$

et donc son action sur un vecteur est donnée par

$$df_{(1,1)} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ 2v_x + 2v_y \end{pmatrix}. \tag{92.53}$$

EqActionDiffdeuxtrois

Exercise 366 exoSerieDeux0004

Calculer les normes opérateurs des différentielles de l'exercice 365. corrSerieDeux0004

Correction of the exercise 366

(1) La différentielle $df(x) = 12x$ est une homothétie de rapport 12 et sa norme est donc 12 en vertu de ce que nous avons dit à l'exercice 364(1).

(2) Il est préférable de lire d'abord la résolution du point (3). Au vu de ce qui y est dit, il faut calculer $\|\alpha^t(w)\|$ lorsque α est la rotation et $w = (1,2,0)$. Si α est une rotation, c'est une matrice orthogonale et sa transposée est égale à son inverse. En particulier, α^t est encore un opérateur de rotation autour de l'axe $(1,1,1)$. Cette rotation ne change évidement [5] pas la norme de w, et donc la norme de la différentielle est égale à

$$\|(1,2,0)\| = \sqrt{5}. \tag{92.54}$$

ItemCDeuxquatre

(3) Oublions un instant l'application α et voyons la norme de l'application $x \mapsto \langle x, w \rangle$. Nous devons calculer

$$\sup_{|x|=1} \{|\langle x, w \rangle|\}. \tag{92.55}$$

Le produit scalaire est maximum lorsque x est parallèle à w. Dans ce cas, elle vaut $|w|$ parce que $\|x\| = 1$.

Que se passe-t-il lorsque l'on met maintenant l'application α ? Il n'est évidement plus du tout garanti que $\alpha(x)$ sera parallèle à w lorsque x est parallèle à w. Nous pouvons par contre écrire ceci :

$$\langle \alpha(x), w \rangle = \langle x, \alpha^t(w) \rangle \tag{92.56}$$

en utilisant la transposée. Nous tombons donc dans le même cas que si α n'était pas là, mais avec le vecteur $\alpha^t(w)$ au lieu de w. La norme recherchée est donc

$$|\alpha^t(w)|. \tag{92.57}$$

(4) Nous devons calculer le supremum de la norme du vecteur (dans $\mathbb{R}^3$) donné par l'équation (92.53) lorsque $\|(v_x, v_y)\| = 1$, c'est-à-dire

$$\|df_{(1,1)}\| = \sup_{x^2+y^2=1} \{x^2 + y^2 + 4(x+y)^2\}. \tag{92.58}$$

5. Évidemment ?

Pour maximiser la fonction $x^2+y^2+4(x+y)^2$ sur le cercle, nous passons en coordonnées polaires : $x=\cos(\alpha)$ et $y=\sin(\alpha)$. Nous trouvons

$$5+8\sin(\alpha)\cos(\alpha)=5+4\sin(2\alpha). \tag{92.59}$$

Cette fonction prend son maximum lorsque $\sin(2\alpha)=1$ et donc la norme vaut

$$\|df_{(1,1)}\|=9. \tag{92.60}$$

Exercise 367 exoSerieDeux0005

Considérons le problème consistant a trouver la plus grande racine ξ du polynôme $P(x)=x^2+ax+b$ où $a,b\leqslant 1$. Estimer le conditionnement relatif $K_{\text{rel}}(a,b)$. Interpréter le résultat. corrSerieDeux0005

Correction of the exercise 367

La plus grande racine est donnée, en fonction de $(a,b)\in\mathbb{R}^2$, par

$$\xi(a,b)=\frac{-a+\sqrt{a^2-4b}}{2}. \tag{92.61}$$

Le gradient de cette fonction se calcule facilement :

$$\begin{aligned}(\partial_a\xi)(a,b)&=-\frac{1}{2}\frac{\xi(a,b)}{\sqrt{a^2-4b}}\\(\partial_b\xi)(a,b)&=-\frac{1}{\sqrt{a^2-4b}}.\end{aligned} \tag{92.62}$$

La norme du gradient vaut

$$\|\nabla\xi(a,b)\|=\frac{\sqrt{\xi^2+4}}{2\sqrt{a^2-4b}}\simeq K^{\eta}_{\text{abs}}(a,b) \tag{92.63}$$

si η est assez petit. Nous trouvons ensuite le conditionnement relatif

$$K_{\text{rel}}\simeq\frac{\sqrt{\xi^2+4}\sqrt{a^2+b^2}}{\xi\sqrt{a^2-4b}}. \tag{92.64}$$

Pour que le conditionnement soit bon, il faut que $(a,b)\to(0,0)$. Il faut étudier cette limite.

Exercise 368 exoSerieDeux0006

Appliquer la méthode de Newton au problème précédant, étudier le comportement des différentes suites obtenues

Aide : raisonner sur un dessin avant de vous lancer dans un démonstration rigoureuse. corrSerieDeux0006

Correction of the exercise 368

La formule d'itération de Newton est construite pour que x_{n+1} soit l'abscisse d'intersection entre la tangente à f en x_n et l'axe des x. Faisons une petit dessin à la figure 92.1 avec $a<0$ et $b>0$ pour se fixer les idées.

Supposons d'abord que $x_n<r_0$. Alors le point est sur la partie décroissante de P, et la dérivée est de plus en plus petite (en valeur absolue) jusqu'au sommet $S=(r_0+r_1)/2$. Cela fait que $x_{n+1}>x_n$. La suite des x_i ainsi obtenue est donc croissante.

D'un autre côté, nous avons $P(x_n)>0$ et $P'(x_n)<0$, mais pour tout $x\in[x_n,r_0]$, nous avons $P'(x)<P(x_n)$, donc $x_{n+1}<r_0$. Nous avons donc $x_{n+1}\in[x_n,r_0,]$ de telle sorte que la suite des x_i soit bornée et croissante et donc convergente.

Supposons maintenant que $x_n\in]r_0,S,[$. Dans ce cas, nous avons $x_{n+1}<r_0$ et donc on repasse tout de suite au cas précédent.

Regardez ce qu'il se passe quand $x_n=S$ et puis essayez de refaire tous les raisonnements pour les cas où x_n se trouve à droite de C.

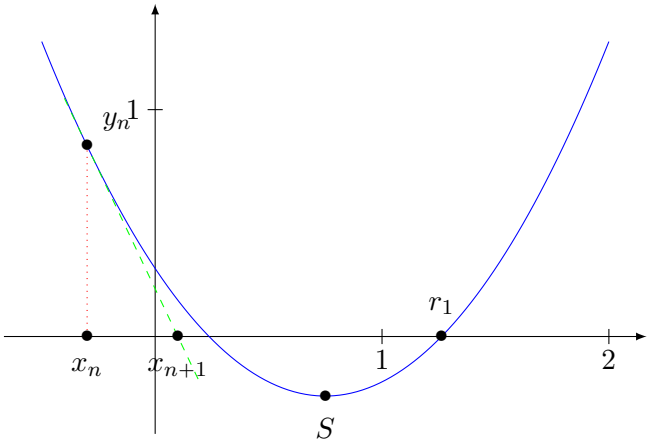


FIGURE 92.1: La méthode de Newton pour une parabole LabelFigMethodeNewton

Exercise 369 exoSerieTrois0001

Soit $x = x(d)$ un problème C^1 en dimension un. Soit $\{x_n = x_n(d_n)\}$ un algorithme fortement consistant. Démontrer que sous ces hypothèses, si l'algorithme est stable alors il est nécessairement convergent. corrSerieTrois0001

Correction of the exercise 369

L'étude de la convergence passe par l'étude de la quantité

$$|x(d) - x_n(d_n)| \tag{92.65}$$

qui devrait être petite quand d_n est proche de d. En ajoutant et retranchant $x_n(d)$

$$\begin{aligned} |x(d) - x_n(d_n)| &= \big|x(d) + x_n(d) - x_n(d) - x_n(d_n)\big| \\ &\leqslant \big|x(d) - x_n(d)\big| + \big|x_n(d) - x_n(d_n)\big| \end{aligned} \tag{92.66}$$ Eqtroiszerounbaspt

Nous allons étudier les deux termes séparément.

D'abord, F_n est un problème stable. Cela implique que $F_n(x_n, d_n) = 0$ a une unique solution $x(d)$. Or nous avons que

$$\begin{aligned} F_n(x, d_n) &= 0 \\ F_n(x_n, d_n) &= 0. \end{aligned} \tag{92.67}$$

La première ligne est la forte consistance et la seconde est la définition de x_n comme solution de $F_n(x_n, d_n)$. L'unicité de la solution implique donc que $x = x_n$ en tant que fonctions. Le premier terme de la somme (92.66) est donc nul.

Pour le second terme, nous savons que la fonction x_n est continue au point d_n parce que x_n est solution du problème stable F_n. Donc il existe un δ qui fait que $\big|x_n(d) - x_n(d_n)\big| \leqslant \epsilon$ lorsque $|d - d_n| \leqslant \delta$.

Exercise 370 exoSerieTrois0004

Vérifier la convergence (au sens de la définition 34.70) de la méthode de Newton dans le cas de la recherche de la plus grande racine de $x^2 + ax + b = 0$ $(a, b \geqslant 1)$.

Lire la section 34.6 peut d'avérer utile pour comprendre. Comprendre la résolution de l'exercice 368 est aussi un atout. corrSerieTrois0004

Correction of the exercise 370

En vertu de la remarque 34.71, nous devons simplement considérer la suite numérique $\{x_n\}$ et en prouver la convergence vers la solution cherchée.

Nous avons déjà montré à l'exercice 368 que la suite de ces x_n était de toute façon convergente (parce que bornée et croissante ou bien bornée et décroissante d'après le côté). Cette limite est obligatoirement $x(a_0, b_0)$ parce que ce dernier est l'unique solution à l'équation

$$\ell = \ell - \frac{f(\ell)}{f'(\ell)}. \tag{92.68}$$ eqellsollimTrtr

ATTENTION : remarquez que ce dernier point n'est pas tout à fait exact. Il y a deux solutions à l'équation (92.68). En regardant cependant attentivement la façon dont la suite des x_n évolue dans l'exercice 368, tu dois pouvoir te convaincre que celle dont nous parlons ici est la correcte, si nous commençons la récurrence du bon côté, c'est-à-dire du côté de la plus grande racine.

Exercise 371 exoSerieTrois0003

Prouver la stabilité et estimer le conditionnement relatif asymptotique correspondant au problème de Newton de l'exercice 371.

Relire la section 34.6 devient de plus en plus utile. corrSerieTrois0003

Correction of the exercise 371

La stabilité de l'algorithme revient à la stabilité de tous les problèmes intermédiaires. Le problème numéro n est le problème qui consiste à trouver x_{n+1} en termes de la donnée de la fonction (a et b). Dans notre cas, l'équation (34.118) s'écrit

$$F_n(x_{n+1}, x_n, (a,b)) = x_{n+1} - x_n + \frac{x_n^2 + ax_n + b}{2x_n + a}. \tag{92.69}$$ EqProbNDeNewtonTT

La solution (unique) de ce problème est

$$x_{n+1}(a,b) = x_n - \frac{x_n^2 + ax_n + b}{2x_n + a} \tag{92.70}$$ EqSolNewtonTT

Le problème (92.69) a une solution unique donnée par (92.70). Cette solution est une fonction de a et b après substitution de x_n par sa valeur en termes de x_{n-1} puis x_{n-2}, etc. En tant que composée de fonctions C^1 (chacune des fonctions $x_i(x_{i-1}, a, b)$), la fonction $x_{n+1}(a,b)$ est C^1, de telle sorte que le résultat de l'exercice 353 montre que le problème est stable. Notons que cela n'est pas vrai si x_n est sur le sommet de la parabole, mais nous avons vu dans l'exercice 368 que si x_0 n'était pas sur le sommet, alors les x_n n'y seraient pas non plus.

Le conditionnement absolu peut être approximé par la norme du gradient de la fonction x_{n+1} :

$$\begin{aligned} \frac{\partial F_n}{\partial a} &= \frac{x_n^2 - b}{(2x_n + a)^2} \\ \frac{\partial F_n}{\partial b} &= \frac{1}{2x_n + a}. \end{aligned} \tag{92.71}$$

En prenant la norme nous trouvons

$$K_{\text{abs}} \simeq \frac{\sqrt{(x_n^2 - b)^2 + (2x_n + a)^2}}{(2x_n + a)^2}, \tag{92.72}$$

et donc nous avons le conditionnement relatif

$$\begin{aligned} K_{\text{rel}} &= K_{\text{abs}} \cdot \frac{\sqrt{a^2 + b^2}}{|x_{n+1}(a,b,x_n)|} \\ &= \frac{\sqrt{(x_n^2 - b)^2 + (2x_n + a)^2}\sqrt{a^2 + b^2}}{|2x_n + a| \cdot |x_{n+1}(a,b,x_n)|} \end{aligned} \tag{92.73}$$ EqKReltroitroti

où $x_{n+1}(a,b,x_n)$ est donné par la formule (92.70).

Si l'algorithme converge[6], alors les x_n convergent vers la solution exacte. Or nous savons la solution exacte du problème depuis le jardin d'enfance ; nous avons donc :

$$\lim x_n = \frac{1}{2}\big(-a + \sqrt{a^2 - 4b}\big). \tag{92.74}$$ EqSilimExisteTroistrois

Le conditionnement relatif asymptotique est obtenu en substituant (92.74) dans l'expression (92.73) parce que quand une suite a une limite, sa limite supérieure est égale à sa limite.

6. Voir exercice 370

Exercise 372 exoSerieQuatre0001

Si $b = 10$, $n = 3$, $m = 6$ et $s = 0$ que représente le nombre représenté en virgule fixe $[0012, 200000]$? Même question si $b = 3$. corrSerieQuatre0001

Correction of the exercise 372

Le $s = 0$ indique qu'on a affaire à des nombres positifs. Dans le cas de la base 10 ($b = 10$), le nombre $[0012, 200000]$ représente

$$0 \cdot 10^3 + 0 \cdot 10^2 + 1 \cdot 10^1 + 2 \cdot 10^0 + 2 \cdot 10^{-1} + 0 \cdot 10^{-2} + \cdots + 0 \cdot 10^{-6} = 12.2. \tag{92.75}$$

C'est le 12.2 usuel comme on apprend à l'écrire au jardin d'enfants.

En base 3, c'est le même jeu, sauf qu'on remplace les 10 par des 3. Nous avons donc

$$1 \cdot 3^1 + 2 \cdot 3^0 + 2 \cdot 3^{-1} = 3 + 2 + \frac{2}{3} = \frac{17}{3}. \tag{92.76}$$

Exercise 373 exoSerieQuatre0002

En virgule flottante, que signifient les nombres suivants ?

(1) $\{[34], e = 2, b = 10, s = 0\}$

(2) $\{[10001], e = 6, b = 2, s = 0\}$ corrSerieQuatre0002

Correction of the exercise 373

Ici, comme dans l'exercice 372, le $s = 0$ signifie qu'on parle de nombre positifs. La formule utilisée pour les nombres à virgule flottante est :

$$x = (-1)^s b^e \sum_{j=1}^{t} a_j b^{-j}. \tag{92.77}$$ EqRepreFlotNOrm

Cela pourrait ne pas être la même que celle donnée dans la définition 34.6.

(1) Nous écrivons le nombre en suivant la formule (92.77) :

$$10^2 \cdot \left(3 \cdot 10^{-1} + 4 \cdot 10^{-2}\right) = 10^2 \cdot 0.34 = 34. \tag{92.78}$$

(2) Le premier 1 est le a_1, et le denier est a_5. La formule (92.77) devient donc

$$2^6 \cdot (1 \cdot 2^{-1} + 1 \cdot 2^{-5}) = 2^6\left(\frac{1}{2} + \frac{1}{2^5}\right) = 2^5 + 2 = 34. \tag{92.79}$$

Exercise 374 exoSerieQuatre0003

Représenter le réel $234,5$ en virgule flottante avec $b = 2$. corrSerieQuatre0003

Correction of the exercise 374

La première chose à faire pour ce genre d'exercice est d'écrire le nombre en somme de puissances de deux :

$$235.5 = 2^7 + 2^6 + 2^5 + 2^3 + 2 + 2^{-1}, \tag{92.80}$$ EqQTdecun

ensuite on met la plus grande puissance de 2 en évidence [7]

Nous écrivons maintenant la suite des 0 et 1 en fonction des coefficients des puissances de 2 : 111010101. Les zéros correspondent aux puissances 4, 2 et 0 qui ne sont pas présentes dans la décomposition (92.80). Au final, le nombre s'écrit

$$\{[111010101], 2, 7, 0\}. \tag{92.81}$$

Le 7 est la puissance mise en évidence, le 2 est la base et le 0 indique qu'on a un nombre positif.

7. Si on avait précisé un L et que la plus grande puissance dépassait L, l'exercice serait impossible.

Exercise 375 exoSerieQuatre0004

Combien de nombres réels peut-on représenter en virgule flottante avec $b = 4$, $L = -2$, $U = 2$, $t = 2$? corrSerieQuatre0004

Correction of the exercise 375

Calculons la quantité de nombres réels positifs représentables, et multiplions la par deux. Le nombre zéro ne pose pas de problèmes parce qu'il n'est de toutes façons pas représentable.

Pour un e donné, les nombres représentables sont les sommes possibles

$$\sum_{j=1}^{t} a_j b^{-j}, \tag{92.82}$$

c'est-à-dire toutes les combinaisons valides des a_j. Le coefficient a_1 peut prendre les valeurs entre 1 et $b-1$, tandis que les autres vont de 0 à $b-1$. Ici, on a $b = 4$ et $t = 2$, c'est-à-dire les combinaisons 10, 11, 12, 13, 20, 21, 22, 23, 30, 31, 32, 33 y a donc 12 possibilités.

En ce qui concerne e, il y a les possibilités de $L = -2$ à $U = 2$, c'est-à-dire 5 possibilités. En tout nous pouvons donc écrire $12 \times 5 = 60$ nombres positifs, et donc 120 nombres en tout.

Exercise 376 exoSerieQuatre0005

Une machine représente les nombres en virgule flottante en base $b = 10$ et avec $t = 4$ chiffres significatifs. On considère le problème $x = F(d) = d - 1$ avec $d = 1,00098$. L'erreur relative après codage par $\hat{d} = \text{fl}(d) = 0,1001.10^1$ est approximativement $\epsilon_d \simeq \frac{1}{50000}$. Or la résolution machine, livre une solution approchée $\hat{x} = 1.10^{-3}$ dont l'erreur relative s'élève à

$$\epsilon_x = \left| \frac{F(\hat{d}) - x}{x} \right| \simeq \frac{1}{50}. \tag{92.83}$$

Expliquer ce phénomène de propagation en termes du conditionnement de F. corrSerieQuatre0005

Correction of the exercise 376

Certes, le nombre d rentre très bien dans la machine, mais nous allons voir que le nombre $1-d$, lui, rendre nettement plus mal. En effet, une des choses importantes vues au cours théorique est que

$$F(\hat{d}) \sim F(d)(1 + K_{\text{rel}}(d)\rho_d) \tag{92.84}$$

où K_{rel} est le conditionnement relatif du problème F tandis que ρ_d est l'écart relatif entre d et $\hat{d}$: $\rho_d = (\hat{d} - d)/d$. Dans notre cas, nous avons

$$\rho_d = \frac{1.001 - 1.00098}{1.00098} = 0.00001998\ldots \sim \frac{1}{50000}. \tag{92.85}$$

Par contre, l'erreur relative sur $F(d)$ est donnée par le nombre

$$\frac{F(\hat{d}) - F(d)}{F - d} = K_{\text{rel}}(d)\rho_d. \tag{92.86}$$ EqEcartRelSurF

Nous devons évaluer $K_{\text{rel}}(d)$. D'abord nous calculons le conditionnement absolu

$$K_{\text{abs}}(d) \simeq |F'(d)| = |1 - d|' = 1. \tag{92.87}$$

Ensuite le relatif est donné par

$$K_{\text{rel}}(d) = \frac{|d|}{|1 - d|} = \frac{1.00098}{1 - 1.00098} \sim 1021. \tag{92.88}$$

L'erreur relative sur $F(d)$ est donc environ 1000 fois plus grande que celle sur d.

Exercise 377 exoSerieQuatre0006

Pour approximer $\ln(1,02)$ par une série de Taylor tronquée, combien de termes faut-il sommer pour que l'erreur de troncature soit inférieure a 10^{-5} ? corrSerieQuatre0006

Correction of the exercice 377

Ceci est un exercice classique qui devrait être tout à fait du type de ceux fait aux exercices d'analyse. Faites moi signe si il y a un problème.

Exercise 378 exoSerieQuatre0007

Supposons que $\hat{x} = 0,937$ ait trois décimales exactes par rapport a x. Majorer l'erreur relative. Si $f(x) = \sqrt{1-x}$, majorer l'erreur relative commise en remplaçant $f(x)$ par $f(\hat{x})$. corrSerieQuatre0007

Correction of the exercice 378

Nous ne connaissons pas x, mais nous savons que son écriture décimale commence par 0.937, et donc est dans $[0.937, 0.938[$. L'erreur relative est donnée par

$$\rho(x) = \left| \frac{x - \hat{x}}{x} \right| = \frac{x - \hat{x}}{x} \tag{92.89}$$

dont il faut trouver une majoration. Il s'agit donc de trouver pour quelle valeur de x cette expression est la plus grande possible. Nous avons

$$\rho'(x) = \frac{\hat{x}}{x^2} \tag{92.90}$$

qui est toujours positive. Donc la fonction ρ est une fonction positive de x et la majoration se fait en prenant la plus grande valeur possible pour x. Au final nous trouvons

$$\rho_x \leqslant \frac{0.938 - 0.937}{0.938}. \tag{92.91}$$

Nous notons α ce nombre.

Maintenant nous voulons étudier la différence entre $f(x)$ et $f(\hat{x})$ pour la fonction $f(t) = \sqrt{1-t}$. Pour ce faire, nous reprenons la formule (92.86) en utilisant f. Un premier pas est de majorer ρ_x par la constante α :

$$\rho_{f(\hat{x})} \simeq K_{\text{rel}}(x)\rho(x) \leqslant K_{\text{rel}}(x)\alpha \tag{92.92}$$ EqQSrholega

Nous devons maintenant évaluer $K_{\text{rel}}(x)$. D'abord

$$K_{\text{abs}}(x) \simeq |f'(x)| = \frac{1}{2\sqrt{1-x}}, \tag{92.93}$$

et ensuite

$$K_{\text{rel}}(x) = \frac{1}{2}\left| \frac{x}{x-1} \right| = \frac{x}{2(1-x)}. \tag{92.94}$$

Hélas ce n'est pas terminé parce que nous ne savons pas la valeur de x. Afin de terminer nos majoration de l'équation (92.92), nous devons majorer $K_{\text{rel}}(x)$ en fonction de x. Encore une fois, cela se fait en utilisant la dérivée. Le résultat est que $K_{\text{rel}}(x)$ est maximum pour $x = 0.938$.

Donc la majoration que nous pouvons faire est

$$\rho_{f(\hat{x})} \leqslant \frac{0.938 - 0.937}{0.938} \cdot \frac{0.938}{2(0.938 - 1)} = \frac{0.001}{0.124} \simeq 0.008. \tag{92.95}$$

Exercise 379 exoSerieQuatre0008

Pour les évaluations suivantes quels sont les problèmes mal conditionnés ?

(1) $\cos(x)$ pour $x \simeq \frac{\pi}{2}$

(2) $\sqrt{x+1} - \sqrt{x}$ pour x grand

(3) $\sqrt[6]{1+x} - 1$ pour $x \simeq 0$.

corrSerieQuatre0008

Correction of the exercice 379

Le code suivant (à copier-coller dans un fenêtre Sage) résout l'exercice. Il utilise le module `outilsMAT1151` disponible sur gitorious.

```
# -*- coding: utf8 -*-
import outilsMAT1151

var('x')
fa(x) = cos(x)
fb(x) = sqrt(x+1)-sqrt(x)
fc(x) = (1+x)**(1/6)-1
print outilsMAT1151.ConditionnementRelatif(fa,pi/2)
print outilsMAT1151.ConditionnementRelatif(fb,oo)
print outilsMAT1151.ConditionnementRelatif(fc,0)
```

Les réponses retournées sont

```
+Infinity
1/2
1
```

Donc le premier est mal conditionné, le second est plutôt bien conditionné et le troisième est un peu moins bien.

Exercise 380 exoSerieCinq0001

On considère la suite de nombres

$$I_n := \int_0^1 \frac{x^n}{x+10} dx \tag{92.96}$$

(1) Donner une formule de récurrence sur les $\{I_n\}$ (aide : quel faut il faire quand le numérateur a un degré plus haut que le dénominateur ?)

(2) Étudier la propagation des erreurs pour l'évaluation machine de I_N pour N grand.

corrSerieCinq0001

Correction of the exercise 380

(1) Nous pouvons calculer I_0 directement en effectuant l'intégrale :

$$I_0 = \int_0^1 \frac{1}{x+10} dx = \ln(11) - \ln(10). \tag{92.97}$$

Nous devons maintenant chercher à exprimer I_n en fonction de I_{n-1}.

Par division euclidienne, on obtient

$$\frac{x^n}{x+10} = x^{n-1} - 10\frac{x^{n-1}}{x+10}.$$

Dès lors,

$$I_n = \int_0^1 x^{n-1}\, dx - 10I_{n-1} = \left[\frac{x^n}{n}\right]_0^1 - 10I_{n-1} = \frac{1}{n} - 10I_{n-1}.$$

(2) Le conditionnement relatif κ_n de l'application $I_{n-1} \mapsto I_n = \frac{1}{n} - 10I_{n-1}$ vaut environ

$$\kappa_n \sim 10\frac{I_{n-1}}{I_n}.$$

L'erreur de propagation à l'étape n est donnée par

$$\kappa_1\kappa_2 \dots \kappa_n\rho_0 \sim 10^n \frac{I_0}{I_1}\frac{I_1}{I_2} \cdots \frac{I_{n-1}}{I_n}\rho_0 = 10^n \frac{I_0}{I_n}\rho_0, \tag{92.98}$$

EqCUkkkrho

où ρ_0 est l'erreur relative sur I_0. En effet, afin de lancer l'algorithme, nous devons mettre la valeurs de I_0 « à la main », et cette erreur se propage à cause des conditionnements des différents problèmes de l'algorithme.

Comment se comporte l'équation (92.98) lorsque $n \to \infty$? À cause du 10^n, la seule chance que la propagation de l'erreur ne tende pas vers l'infini est que $I_n \to \infty$. Nous allons voir que ce n'est pas le cas.

Sur l'intervalle $[0,1]$, nous avons $x^n \leqslant x^{n-1}$, donc $I_n \leqslant I_{n-1}$; la suite des I_n est donc une suite décroissante. Étant bornée vers le bas par 0, nous savons que c'est une suite convergente.

Donc la propagation d'erreur tend bien vers l'infini lorsque n est grand, ce qui montre que l'algorithme par récurrence est très mauvais pour évaluer l'intégrale pour des grands n.

Exercise 381 exoSerieCinq0002

Soit A une matrice carrée de taille N. Démontrer que

$$\det A = \sum_{\sigma \in \mathrm{Sym}(N)} \epsilon(\sigma) A_{1\sigma(1)} ... A_{N\sigma(N)}. \tag{92.99}$$

corrSerieCinq0002

Correction of the exercise 381

Nous pouvons décomposer le groupe $\mathrm{Sym}(N+1)$ en $N+1$ classes de la façon suivante :

$$\begin{aligned} P_1 &= \begin{pmatrix} 1 & 2 & 3 & \dots & N+1 \\ 1 & \cdot & \cdot & \dots & \cdot \end{pmatrix} \\ P_2 &= \begin{pmatrix} 1 & 2 & 3 & \dots & N+1 \\ 2 & \cdot & \cdot & \dots & \cdot \end{pmatrix} \\ &\vdots \\ P_{N+1} &= \begin{pmatrix} 1 & 2 & 3 & \dots & N+1 \\ N+1 & \cdot & \cdot & \dots & \cdot \end{pmatrix}. \end{aligned} \tag{92.100}$$

La première classe contient toutes les permutations qui envoient 1 sur 2, le second toutes les permutations qui envoient 1 sur 2, etc. Nous avons « évidement » l'union disjointe

$$\mathrm{Sym}(N+1) = \bigcup_{i=1}^{N+1} P_i. \tag{92.101}$$

Pour une matrice A d'ordre N, nous posons

$$F(A) = \sum_{\sigma \in \mathrm{Sym}(N)} \epsilon(\sigma) A_{1\sigma(1)} \dots A_{N\sigma(N)}. \tag{92.102}$$

Le but est de prouver que la fonction F est le déterminant.

Pour $N = 2$, nous avons

$$\sum_{\sigma \in \mathrm{Sym}(2)} \epsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} = A_{11}A_{22} - A_{12}A_{21} = \det(A),$$

puisque $\mathrm{Sym}(2) = \{\mathrm{Id}, (1,2)\}$ avec $\epsilon(\mathrm{Id}) = 1$ et $\epsilon(1,2) = -1$.

Si maintenant A est d'ordre $N+1$, nous calculons $F(A)$ en décomposant la somme sur $\mathrm{Sym}(N+1)$ en sommes sur les classes P_i :

$$\begin{aligned} F(A) = &\sum_{\sigma \in P_1} \epsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{N+1,\sigma(N+1)} \\ &+ \sum_{\sigma \in P_2} \epsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{N+1,\sigma(N+1)} \\ &\vdots \\ &+ \sum_{\sigma \in P_{N+1}} \epsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{N+1,\sigma(N+1)} \end{aligned} \tag{92.103}$$

EqCDsommeSymFA

Commençons par étudier la première ligne. Par définition de P_1, nous avons $\sigma(1) = 1$ pour tous les σ de la somme. Par conséquent, le facteur A_{11} peut se mettre en évidence. La première ligne vaut donc

$$A_{11} \sum_{\sigma \in P_1} \epsilon(\sigma) A_{2\sigma(2)} \dots A_{N+1,\sigma(N+1)} = A_{11} F(B) \tag{92.104}$$

EqCDPremierParq

où

$$B = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & A_{22} & \dots & A_{2,N+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & A_{N+1,2} & \dots & A_{N+1,N+1} \end{pmatrix}. \tag{92.105}$$

En effet, parmi les éléments B_{1k}, seul B_{11} n'est pas nul et vaut 1. Par conséquent, dans le cas de la matrice B, la somme sur tout $\mathrm{Sym}(N+1)$ se réduit à la somme sur P_1.

Afin d'utiliser l'hypothèse de récurrence, il suffit de remarquer que pour toute matrice M, nous avons

$$F(M) = F \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix}. \tag{92.106}$$

Pour ce remarquer, considérons la matrice

$$M = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & B_{11} & B_{12} & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & B_{N,1} & B_{N,2} & B_{NN} \end{pmatrix}, \tag{92.107}$$

et étudions la somme

$$F(M) = \sum_{\sigma \in P_1} \epsilon(\sigma) M_{2\sigma(2)} \dots M_{N+1,\sigma(N+1)}. \tag{92.108}$$

Nous avons une bijection

$$\begin{aligned} \varphi \colon \mathrm{Sym}(N) &\to P_1 \\ \varphi(\eta)(n) &= \eta(n-1) + 1. \end{aligned} \tag{92.109}$$

Cette bijection vérifie $\epsilon\big(\varphi(\eta)\big) = \epsilon(\eta)$. En utilisant cette bijection, nous avons

$$\begin{aligned} F(M) &= \sum_{\eta \in \mathrm{Sym}(N)} \epsilon\big(\varphi(\eta)\big) M_{2,\varphi(\eta)(2)} \dots M_{N+1,\varphi(\eta)(N+1)} \\ &= \sum_{\eta \in \mathrm{Sym}(N)} \epsilon(\eta) M_{2,\eta(1)+1} \dots M_{N+1,\eta(N)+1} \\ &= \sum_{\eta \in \mathrm{Sym}(N)} \epsilon(\eta) B_{1,\eta(1)} \dots B_{N,\eta(N)} \\ &= F(B). \end{aligned} \tag{92.110}$$

La formule (92.104) est donc bien A_{11} fois le déterminant du mineur correspondant, c'est-à-dire le premier terme du développement du déterminant de A en suivant la première ligne.

Nous regardons maintenant le second terme de la somme (92.103). Dans cette somme nous avons toujours $\sigma(1) = 2$ et donc nous pouvons factoriser A_{12}, et nous devons étudier

$$A_{12} \sum_{\sigma \in P_2} \epsilon(\sigma) A_{2\sigma(2)} \dots A_{N+1}\sigma(N+1) = A_{12} F(B) \tag{92.111}$$

où

$$B = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ A_{21} & 0 & A_{23} & \cdots & A_{2,N+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N+1,1} & 0 & A_{N+1,3} & \dots & A_{N+1,N+1} \end{pmatrix}. \tag{92.112}$$

Cette fois pour utiliser la relation de récurrence, nous devons prouver que permuter deux colonnes revient à changer le signe de F. De cette façon nous pouvons nous ramener au cas précédent, plus un signe, comme il se doit lorsqu'on calcule un déterminant.

Considérons η_{kl} la permutation de k et l qui laisse invariant tous les autres nombres. Nous considérons la matrice N dont les éléments sont

$$N_{ij} = M_{i,\eta_{kl}(j)}, \tag{92.113}$$

c'est-à-dire la matrice M dont nous avons permuté les colonnes k et l. L'application $\sigma \mapsto \sigma \circ \eta$ est une bijection à l'intérieur de $\mathrm{Sym}(N)$ et de plus $\epsilon\Big(\sigma \circ \eta\Big) = -\epsilon(\sigma)$. Nous avons donc

$$\begin{aligned} F(N) &= \sum_{\sigma\in\mathrm{Sym}(n)} \epsilon(\sigma) N_{1\sigma(1)} \dots N_{n,\sigma(n)} \\ &= \sum_{\sigma} \epsilon(\sigma) M_{1,(\sigma\circ\eta)(1)} \dots M_{n,(\sigma\circ\eta)(n)} \\ &= \sum_{\sigma} \epsilon(\sigma\circ\eta) M_{1\sigma(1)} \dots M_{n\sigma(n)} \\ &= -F(M). \end{aligned} \tag{92.114}$$

Pour obtenir l'avant-dernière ligne, nous avons changé la sommation de σ vers $\sigma \circ \eta$.

Exercise 382 exoSerieCinq0003

En utilisant la méthode de Gauss, calculer la matrice inverse de

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix}. \tag{92.115}$$

corrSerieCinq0003

Correction of the exercise 382

Il s'agit d'utiliser la méthode de la matrice compagnon. Faites moi signe si cela pose un problème.

Exercise 383 exoSerieCinq0004

Soit V un vectoriel réel de dimension finie et $V \times V \to \mathbb{R} : (u, v) \mapsto \langle u, v\rangle$ un produit scalaire euclidien sur V. Exprimer la norme opérateur d'une application linéaire $\alpha : V \to V$ telle que $\langle \alpha u, v\rangle = \langle u, \alpha v\rangle$ en termes de son spectre.

corrSerieCinq0004

Correction of the exercise 383

Le fait que l'application α vérifie $\langle \alpha u, v\rangle = \langle u, \alpha v\rangle$ implique que la matrice de α est une matrice symétrique (pourquoi ?). Elle possède par conséquent une base orthonormale de vecteurs propres. Notons f_i cette base :

$$\alpha f_i = \lambda_i f_i. \tag{92.116}$$

Certains de λ_i peuvent être nuls, ce n'est pas un problème.

La norme de α est donnée par

$$\|\alpha\| = \sup_{\|x\|=1} \{|\alpha(x)|\}. \tag{92.117}$$

Si nous décomposons x dans la base choisie des vecteurs propres de α_i, nous avons $x = x_i f_i$ (somme sous-entendue sur les i) avec la contrainte

$$\sum_i x_i^2 = 1. \tag{92.118}$$

Nous avons

$$\alpha x = x_i \alpha f_i = x_i \lambda_i f_i, \tag{92.119}$$

et comme la base des f_i est orthonormale, nous avons

$$|\alpha(x)|^2 = \sum_i x_i^2 \lambda_i^2. \tag{92.120}$$

Nous devons donc trouver pour quel x sur la sphère unité la somme $\sum_i x_i^2\lambda_i^2$ est la plus grande. Les coefficients λ_i étant fixés, il est intuitivement clair que le maximum est atteint lorsque les composantes de x sont toutes nulles sauf celle qui correspond à la valeur propre la plus grande. Par conséquent, la norme de l'opérateur α serait égal à la valeur absolue de sa plus grande valeur propre.

Prouvons cela.

D'abord le supremum est bien un maximum parce que nous faisons un supremum sur une sphère ($\sum x_i^2 = 1$) et que la sphère est compacte[8].

Pour trouver un extrémum sous contrainte, nous utilisons la méthode des multiplicateurs de Lagrange. Dans le cas qui nous occupe, le lagrangien est

$$L(x_i, \sigma) = \sum_i x_i^2\lambda_i^2 + \sigma(x_1^2 + \cdots + x_n^2 - 1). \tag{92.121}$$

Le système d'équations à résoudre est donné par $\nabla L = 0$:

$$\begin{cases} \dfrac{\partial L}{\partial x_i} = 2x_i(\lambda^2 + \sigma) & \text{(92.122a)} \\ \dfrac{\partial L}{\partial \sigma} = x_1^2 + \cdots + x_n^2 - 1. & \text{(92.122b)} \end{cases}$$

Si il existe au moins deux λ_i différents, alors nous ne pouvons fixer σ que pour annuler $2\lambda_i - \sigma$ pour un seul des λ_i. Les x_i de toutes les autres valeurs propres doivent alors être zéro.

Les extrémums sont donc donnés par les vecteurs x dont toutes les composantes sont nulles sauf celles qui correspondent à une valeur propre donnée. Par conséquent, la norme de α est donnée par

$$\|\alpha\| = \max\{|\lambda_1|, \ldots, |\lambda_n|\}, \tag{92.123}$$

et est donc bien égal à la plus grande valeur propre.

Exercise 384 exoSerieCinq0006

Démontrer qu'une matrice symétrique définie positive peut s'écrire sous la forme T^tT où T est triangulaire.

corrSerieCinq0006

Correction of the exercise 384

Notons d'abord que la matrice T^tT est toujours symétrique parce que $(T^tT)^t = T^tT$.

Prouvons maintenant par récurrence que la matrice A peut être écrite sous la forme T^tT où T est triangulaire. Supposons que A puisse être écrite sous la forme

$$A = \begin{pmatrix} & & x_1 \\ & T_0^tT_0 & \vdots \\ & & x_n \\ x_1 & \ldots x_n & x_{n+1} \end{pmatrix} \tag{92.124}$$

où T_0 est une matrice triangulaire. Nous allons chercher la matrice triangulaire T telle que $T^tT = A$ sous la forme

$$T = \begin{pmatrix} * & * & * & t_1 \\ 0 & * & * & \vdots \\ 0 & 0 & a & t_n \\ 0 & 0 & 0 & t_{n+1} \end{pmatrix} \tag{92.125}$$

où les $*$ (ainsi que le a) forment la matrice T_0.

8. Cela n'est pas vrai en dimension infinie.

En égalisant $A = T^tT$, nous trouvons d'abord

$$t_{n+1}^2 = x_{n+1}. \tag{92.126}$$ EqCCtnpuxnpu

Mais nous avons vu que $x_{n+1} > 0$ en tant qu'élément diagonal de A. Par conséquent la relation (92.126) déterme bien t_{n+1}. Nous avons même le choix du signe. Quel luxe !

Ensuite nous trouvons $x_n = t_n(a + t_{n+1})$, ce qui fournit t_n sous la forme

$$t_n = \frac{x_n}{a + t_{n+1}}. \tag{92.127}$$

Nous avons dit que nous avions le choix du signe pour t_{n+1}. Nous le choisissons positif. Comment prouver que le dénominateur n'est pas nul ? En fait a est l'élément en bas à droite de la matrice T_0 qui est la matrice qui réalise $T_0^tT = A_0$, et donc nous pouvons supposer que a est strictement positif exactement comme nous venons de prendre t_{n+1} strictement positif. En effet, a n'est rien d'autre que le « t_{n+1} » du pas de récurrence précédent.

Exercise 385 exoSerieCinq0005

Démontrer que pour une matrice réelle A symétrique définie positive, la méthode de triangulation de Gauss ne nécessite pas de permutations de lignes. (aide : ne pas chercher à utiliser le résultat de l'exercice précédent) corrSerieCinq0005

Correction of the exercise 385

Lorsqu'on applique la méthode de Gauss, ne pas vouloir faire de permutations de lignes revient à demander qu'à chaque étape, l'élément en haut à gauche puisse servir de pivot, c'est-à-dire qu'il soit non nul.

La matrice étant symétrique, elle est diagonalisable par une matrice orthogonale et toutes ses valeurs propres sont strictement positives parce qu'on suppose que A est définie positive. Par conséquent nous avons

$$A = B^tDB \tag{92.128}$$

où B est orthogonale et D est la matrice diagonale qui contient les valeurs propres sur sa diagonale. Nous en déduisons que les éléments diagonaux de A sont positifs :

$$A_{ii} = B_{ik}^t \underbrace{D_{kl}}_{\delta_{kl}\lambda_l} B_{li} = B_{il}^t \lambda_l B_{li} = \sum_l (B_{il})^2 \lambda_l > 0. \tag{92.129}$$

En ce qui concerne le premier pas de la méthode de Gauss, c'est donc facile : nous venons de prouver que les éléments diagonaux sont non nuls, en particulier A_{11} est non nul et peut servir de pivot.

En ce qui concerne la seconde étape, l'élément A_{22} a changé, il est devenu

$$A_{22} - A_{12}\frac{A_{21}}{A_{11}} \tag{92.130}$$ EqCCnouvelbg

qu'il convient de prouver être non nul. Pour cela nous allons utiliser une récurrence pour prouver que, après chaque pas de Gauss, la matrice « cofacteur » qui reste est encore symétrique et sans valeurs propres nulles.

En effet, si nous nommons B la matrice obtenue après une étape de Gauss, les éléments qui ont changés ont changé de la façon suivante :

$$B_{ij} = A_{ij} - A_{1j}\frac{A_{1i}}{A_{11}} = A_{ij} - \frac{A_{1j}A_{1i}}{A_{11}}. \tag{92.131}$$

En utilisant le fait que A est symétrique, nous avons alors que

$$B_{ji} = A_{ji} - \frac{A_{1i}A_{1j}}{A_{11}} = B_{ij}. \tag{92.132}$$

Donc la matrice qui reste est encore symétrique. Attention : *toute* la matrice n'est pas symétrique :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & -2 & -3 \end{pmatrix}. \tag{92.133}$$

La matrice symétrique dont nous parlons est la petite $\begin{pmatrix} 1 & -2 \\ -2 & -3 \end{pmatrix}$.

Étant donné que le méthode de Gauss ne change pas le rang d'une matrice (mais elle change ses valeurs propres !), la nouvelle matrice ne peut pas avoir de valeurs propres nulles, et donc rentre dans les hypothèse qui font en sorte que tous ses éléments diagonaux sont non nuls.

Par récurrence, nous avons prouvé qu'à chaque étape de l'algorithme de Gausse, tous les éléments diagonaux sont non nuls pourvu que la matrice de départ était symétrique et définie positive.

Remarque 92.1.
Nous pouvons prouver que le second pas de Gauss fonctionne bien de façon directe en utilisant le résultat de l'exercice 384 de la façon suivante. Exprimons l'élément (92.130) en termes de la matrice triangulaire T. Nous avons toujours

$$A_{ij} = (T^tT)_{ij} = T^t_{ik}T_{kj} = \sum_k T_{ki}T_{kj}. \tag{92.134}$$

Notez que, la matrice T étant triangulaire, la somme sur les k dans l'équation (92.134) est souvent assez courte. Par exemple parmi les T_{1k}, seul T_{11} est non nul, et $\sum_k T_{k2} = T_{12} + T_{22}$.

Le nombre (92.130) devient alors, avec les sommes sous-entendues

$$\begin{aligned} T_{k2}T_{k2} - T_{i1}T_{i2}\frac{T_{j2}T_{j1}}{T_{l1}T_{l1}} &= T_{k2}T_{k2} - \frac{T_{11}T_{21} \cdot T_{12}T_{11}}{T_{11}T_{11}} \\ &= T_{12}T_{12} + T_{22}T_{22} - T_{12}T_{12} \\ &= T_{22}T_{22} \neq 0. \end{aligned} \tag{92.135}$$

La diagonale de T ne peut pas contenir d'éléments nuls, sinon son déterminant serait nul, alors que $0 \neq \det A = (\det T)^2$. Cela n'est hélas pas suffisant pour conclure l'exercice parce que nous manquons encore de vision sur ce qu'il se passe au pas suivant (que devient l'élément A_{33} ?).

Les exercices qui suivent proviennent d'examens d'années précédentes.

Exercise 386 exoexamens-0000

On définit les problèmes suivants :

$$F_1(x_1, d_1) = x_1 - \frac{1}{4}\sin^2(d_1) = 0, \tag{92.136}$$

et

$$F_2(x_2, d_2) = x_2^2 - x_2 + d_2 \tag{92.137}$$

pour $x_2 \geqslant \frac{1}{4}$ et $d \in [0, \frac{1}{4}]$. Nous considérons alors le problème composite

$$F(x, d) = F_2(x, x_1(d)) = 0 \tag{92.138}$$

où $x_1(d)$ est solution du premier problème.

(1) Le problème composite (92.138) est-il stable ?

(2) Exprimer le conditionnement relatif du problème composite du problème composite en fonction de ceux des deux problèmes.

(3) Pour quelles valeurs de d le problème (92.138) est-il bien conditionné ?

correxamens-0000

Correction of the exercise 386

Pour être un problème stable, il faut d'abord avoir une solution unique. Explicitons le problème composite :

$$F(x,d) = x^2 - x + \frac{1}{4}\sin^2(d). \tag{92.139}$$

Les solutions de $F(x,d) = 0$ sont données par

$$x(d) = \frac{1 \pm \sqrt{1-\sin^2(d)}}{2}. \tag{92.140}$$

Attention à ne pas oublier la valeur absolue :

$$\sqrt{1-\sin^2(d)} = \sqrt{\cos^2(d)} = |\cos(d)| \neq \cos(d), \tag{92.141}$$

donc

$$x(d) = \frac{1 \pm |\cos(d)|}{2} \tag{92.142}$$

Étant donné que nous cherchons uniquement les solutions $x \geqslant \frac{1}{2}$, nous ne gardons que la solution avec un plus. Au final l'unique solution au problème composite est

$$x(d) = \frac{1 + |\cos(d)|}{2}. \tag{92.143}$$

À cause de la valeur absolue, cette fonction n'est pas C^1, mais sa dérivée reste bornée, donc elle vérifie la seconde condition de la stabilité. Vérifiez que l'exercice 355 tient encore si nous remplaçons « C^1 » par « dérivée bornée ».

À cause de la formule de dérivation des fonctions composées, le conditionnement relatif du problème composite est le produit des conditionnements relatifs, voir exercice 357. Plus précisément :

$$K_{f_1 \circ f_2}(d) = K_1\big(f_2(d)\big)K_2(d), \tag{92.144}$$

et non $K_1(d)K_2(d)$.

Calculons les deux conditionnements relatifs.

$$\begin{aligned} K^{(1)}_{\text{abs}}(d) &= \frac{1}{2}|\sin(d)\cos(d)| \\ K^{(1)}_{\text{rel}}(d) &= 2\frac{d\cos(d)}{\sin(d)}. \end{aligned} \tag{92.145}$$

et, étant donnée que

$$x_2(d) = \frac{1+\sqrt{1-4d^2}}{2}, \tag{92.146}$$

EqComposeK0exam

nous avons

$$\begin{aligned} K^{(2)}_{\text{abs}}(d) &= \frac{2d}{\sqrt{1-4d^2}} \\ K^{(2)}_{\text{rel}}(d) &= \frac{4d}{\sqrt{1-4d^2}\big(1+\sqrt{1-4d^2}\big)}. \end{aligned} \tag{92.147}$$

En utilisant la formule (92.146), nous avons

$$K_{\text{rel}}(d) = \frac{2d\sin(d)}{1+|\cos(d)|}, \tag{92.148}$$

et le problème est bien conditionné lorsque d est petit ou bien lorsque $\sin(d)$ est petit.

Pour trouver le conditionnement relatif du problème composite, nous pouvions aussi directement partir de la solution

$$x(d) = \frac{1 + |\cos(d)|}{2}. \tag{92.149}$$

Pour la dérivée, attention que $|\cos(x)|' \neq |\sin(x)|$, mais grâce au fait que la formule du conditionnement absolue elle-même aie des valeurs absolues, nous avons quand même

$$K_{\text{abs}}(d) = \frac{|\sin(d)|}{2}, \tag{92.150}$$

et donc

$$K_{\text{rel}}(d) = \frac{|\sin(d)|}{2} \cdot \frac{d}{\frac{1+|\cos(d)|}{2}} = \frac{d\sin(d)}{1+|\cos(d)|}. \tag{92.151}$$

Exercise 387 exoexamens-0001

Pour tout entier n positif, on définit la quantité

$$I_n = \int_0^1 \frac{1}{(1+x^2)^n} dx. \tag{92.152}$$

(1) Donner une formule de récurrence pour les I_n. (aide : exprimer le numérateur comme $1 = 1 + x^2 - x^2$, décomposer en somme de deux intégrales et utiliser une intégration par partie sur l'un des deux termes).

(2) Étudier la propagation des erreurs pour l'évaluation machine de I_n pour n grand.

correxamens-0001

Correction of the exercise 387

En utilisant l'aide, nous avons

$$I_n = \int_0^1 \frac{1}{(1+x^2)^n} dx = \int_0^1 \frac{1}{(1+x^2)^{n-1}} dx - \int_0^1 \frac{x^2}{(1+x^2)^n} dx. \tag{92.153}$$

Dans le premier terme, nous reconnaissons I_{n-1} ; c'est donc bien parti pour la récurrence. Le second terme est plus compliqué à traiter. L'aide propose de faire une intégrale par partie. Il y a beaucoup de façons de diviser $\frac{x^2}{(1+x^2)^n}$ en produit de deux facteurs. Presque toutes se heurtent à une primitive très difficile à calculer. Il se fait que la bonne façon de couper la fonction en deux est de poser

$$\begin{aligned} u &= x \\ v' &= \frac{x}{(1+x^2)^n}. \end{aligned} \tag{92.154}$$

La raison est que dans v', nous voyons la dérivée du dénominateur apparaitre au dénominateur. En intégrant nous trouvons

$$v = \frac{1}{2(1-n)(1+x^2)^{n-1}}. \tag{92.155}$$

L'intégration par partie donne

$$\begin{aligned} \int_0^1 \frac{x^2}{(1+x^2)^n} &= \left[\frac{x}{2(1-n)(1+x^2)^{n-1}} \right]_0^1 - \int_0^1 \frac{1}{2(1-n)(1+x^2)^{n-1}} dx \\ &= \frac{1}{2(1-n)2^{n-1}} - \frac{1}{2(1-n)} I_{n-1}. \end{aligned} \tag{92.156}$$

La formule de récurrence devient donc

$$I_n = I_{n-1} - \frac{1}{2^n(1-n)} + \frac{1}{2(1-n)} I_{n-1}, \tag{92.157}$$

c'est-à-dire

$$I_n = \frac{3-2n}{2(1-n)} I_{n-1} - \frac{1}{2^n(1-n)}. \tag{92.158}$$

En ce qui concerne les conditionnements, le conditionnement du pas numéro n est donné par la dérivée de I_n par rapport à I_{n-1} :

$$K_{\text{abs}}^{(n)} = |\frac{3-2n}{2(1-n)}|, \tag{92.159}$$

et donc

$$K_{\text{rel}}^{(n)} = |\frac{3-2n}{2(1-n)}\frac{I_{n-1}}{I_n}|. \tag{92.160}$$

En prenant le produit des conditionnements pour tous les pas, nous avons

$$K_{\text{rel}} = |\prod_{k=1}^{n}\left(\frac{3-k}{2(1-k)}\right)\frac{I_0}{I_n}| \tag{92.161}$$

EqExpKrePart

où il n'est pas compliqué de voir que $I_0 = 1$ en partant directement de la définition. Il est facile de voir que le produit est plus petit que 1 ; hélas il est également facile de voir que I_n est également plus petit que 1 directement par la définition (la fonction qu'on intègre est plus petite que 1).

Nous ne pouvons donc rien conclure de l'expression (92.161). Cela est différent de ce qu'il se passait dans l'exercice 380, dans lequel la formule analogue avait permis de conclure.

Pour conclure, il faut donc un peu repartir en arrière, et écrire le problème du pas numéro n en détail :

$$I_n(d) = \frac{3-2n}{2(1-n)}d - \frac{1}{2^n(1-n)} \tag{92.162}$$

en gardant à l'esprit que n est grand et que $d \leqslant 1$. Nous trouvons

$$K_{\text{abs}}^{(n)}(d) = \left|\frac{3-2n}{2(1-n)}\right|, \tag{92.163}$$

et, après mise au même dénominateur et simplifications,

$$\begin{aligned} K_{\text{rel}}^{(n)}(d) &= \frac{3-2n}{2(1-n)} \cdot \frac{d}{\frac{(3-2n)d}{2(1-n)} - \frac{1}{2^n(1-n)}} \\ &= \left|\frac{(3-2n)d}{(3-2n)d-1}\right|. \end{aligned} \tag{92.164}$$

N'oublions pas que n est grand, donc en réalité $3-2n$ est négatif et cette fraction est plus petite que 1. Nous avons donc prouvé que pour tout n (plus grand que 3), le conditionnement relatif est plus petit que 1. Notons aussi que pour aucun n ce conditionnement n'est nul.

Chapitre 93

Pour des ingénieurs (Louvain-la-Neuve)

93.1 Déterminants et systèmes d'équations

Exercise 388 exoINGE1121La0007

En n'utilisant que les propriétés des déterminants, démontrer l'égalité

$$\det\begin{pmatrix}1&a&a^2\\1&b&b^2\\1&c&c^2\end{pmatrix}. \tag{93.1}$$

Ce déterminant est le déterminant de Vandermonde. corrINGE1121La0007

Correction of the exercise 388

Pour calculer le déterminant, il faut échelonner la matrice de façon usuelle et faire quelques mises en évidence en utilisant des produits remarquables :

$$\begin{aligned}\det\begin{pmatrix}1&a&a^2\\1&b&b^2\\1&c&c^2\end{pmatrix}&=\det\begin{pmatrix}1&a&a\\0&b-a&a^2-b^2\\0&c-a&c^2-a^2\end{pmatrix}\\&=(b-a)(c-a)\det\begin{pmatrix}1&a&a^2\\0&1&-(a+b)\\0&1&-(c+b)\end{pmatrix}\\&=(b-a)(c-a)\det\begin{pmatrix}1&a&a^2\\0&1&-(a+b)\\0&0&-c+a\end{pmatrix}\\&=(b-a)(c-a)(a-c)\det\begin{pmatrix}1&a&a^2\\0&1&-(a+b)\\0&0&1\end{pmatrix}\end{aligned} \tag{93.2}$$

Le déterminant de la dernière matrice est 1 parce que le déterminant d'une matrice triangulaire supérieure est égale au produit de ses éléments diagonaux.

Exercise 389 exoLineraire0029

Calculer les déterminants des matrices suivantes

$$\begin{aligned}a&=\begin{pmatrix}4&5\\3&7\end{pmatrix}\qquad b=\begin{pmatrix}0&5\\7&2\end{pmatrix}\\c&=\begin{pmatrix}5&3&2\\1&2&3\\6&7&9\end{pmatrix}\quad d=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\end{aligned} \tag{93.3}$$

corrLineraire0029

Correction of the exercise 389

<+Lineaire0029+>

Exercise 390 exoLineraire0030

Montrer, sans effectuer, que

$$\det\begin{pmatrix} bc & a^2 & a^2 \\ b^2 & ac & b^2 \\ c^2 & c^2 & ab \end{pmatrix} = \det\begin{pmatrix} bc & ab & ca \\ ab & ca & bc \\ ca & bc & ab \end{pmatrix}$$
$$\det\begin{pmatrix} a_1+b_1 & a_2+b_2 & a_3+b_3 \\ b_1+c_1 & b_2+c_2 & b_3+c_3 \\ c_1+a_1 & c_2+a_2 & c_3+a_3 \end{pmatrix} = 2\det\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}. \tag{93.4}$$

corrLineraire0030

Correction of the exercise 390

<+Lineaire0030+>

Exercise 391 exoLineraire0031

Calculer le plus simplement possible

$$\det\begin{pmatrix} a & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & a & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & a & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \ldots & a & 1 \\ 1 & 1 & & 1 & \ldots & 1 & a \end{pmatrix} \tag{93.5}$$

corrLineraire0031

Correction of the exercise 391

<+Lineaire0031+>

Exercise 392 exoINGE1121La0006

Calculer le déterminant des matrices suivantes.

(1) $C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 3 & 2 & 1 \\ 0 & 5 & 4 & 3 & 2 & 1 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix}$.

corrINGE1121La0006

Correction of the exercise 392

(1) La matrice C est triangulaire *inférieure*, donc son déterminant n'est pas le produit des éléments diagonaux. Il y a une subtilité sur les signes. Si nous développons le déterminant selon la première ligne, seul le 1 reste, mais il vient avec un signe :

$$\det\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 3 & 2 & 1 \\ 0 & 5 & 4 & 3 & 2 & 1 \\ 6 & 5 & 4 & 3 & 21 & \end{pmatrix} = (-1)\det\begin{pmatrix} 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 4 & 3 & 2 \\ 0 & 5 & 4 & 3 & 2 \\ 6 & 5 & 4 & 3 & 2 \end{pmatrix} \tag{93.6}$$

Nous développons ce déterminant en suivant sa première ligne, et il ne reste que le 2, qui vient avec un signe +. Notez que ce 2, dans la matrice originale était sur une case qui aurait eut un signe moins.

En continuant de la sorte, le déterminant à calculer est

$$(-1)\cdot 2\cdot(-3)\cdot 4\cdot(-5)\cdot 6 = -720. \tag{93.7}$$

Exercise 393 exoLineraire0001

Exercice 5, page 85. Résoudre les systèmes

(1) $\begin{cases} x+2y=4 \\ 2x+y=5 \end{cases}$

(2) $\begin{cases} x_1+2x_2+3x_3=0 \\ 2x_1+x_2+3x_3=0 \\ 3x_1+2x_2+x_3=0 \end{cases}$

(3) $\begin{cases} x+y+z-3t=2 \\ x-2y+2z+15t=-3 \\ x+y-z-9t=0 \end{cases}$

(4) $\begin{cases} x+4y+z=12 \\ x+y-z=0 \\ 2x+y=4 \\ x+z=4 \end{cases}$

(5) $\begin{cases} 2x+y-3z-t=1 \\ x-4y+3z+4t=-4 \end{cases}$

corrLineraire0001

Correction of the exercise 393

(1) $x=2$, $y=1$.

(2) Le système se met sous forme de matrice et puis se manipule de la façon suivante.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 2 & 1 & 3 & 0 \\ 3 & 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -4 & -8 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right). \tag{93.8}$$

En pratique, il n'est pas nécessaire de continuer, parce que maintenant la solution se lit facilement. En effet, la dernière ligne dit $z=0$. La seconde ligne dit $y+z=0$, (donc $y=0$) et la première ligne dit $x+2y+3z=0$, donc $x=0$.

(3) Ce système a 3 équations pour 4 inconnues. Nous ne nous attendons donc pas à pouvoir le résoudre complètement. La stratégie va donc être de tout exprimer en termes de la dernière variable. Nous n'allons donc pas essayer de mettre de zéros dans la dernière ligne.

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -3 & 2 \\ 1 & -2 & 2 & 15 & -3 \\ 1 & 1 & -1 & -9 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & -3 & 2 \\ 1 & -2 & 2 & 15 & -3 \\ 0 & 0 & -2 & -6 & -2 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & -3 & 2 \\ 0 & -3 & 1 & 18 & -5 \\ 0 & 0 & -2 & -6 & -2 \end{array}\right) \tag{93.9}$$

À partir d'ici, nous pouvons exprimer toutes les variables en fonction de t. La dernière ligne dit que $-2z-6t=-2$, ce qui donne

$$z=1-3t. \tag{93.10}$$

La seconde ligne dit que $-3y+z+18t=-5$. En injectant dedans la valeur $z=1-3t$, nous trouvons

$$y=2+5t, \tag{93.11}$$

Et enfin, la première ligne dit que $x+y+z-3t=2$. En injectant les valeurs déjà trouvées de y et z en fonction de t, nous trouvons

$$x=t-1. \tag{93.12}$$

Notons qu'il y a quand même moyen de se simplifier un peu la vie en ajoutant encore plus de zéros dans la matrice (tout en en cassant aucun !). En effet, on peut utiliser la troisième ligne pour mettre des zéros sur la colonne des z des premières et deuxièmes lignes, et puis utiliser la deuxième ligne pour mettre un zéro sur la colonne des y dans la première ligne :

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & -6 & 1 \\ 0 & -3 & 0 & 15 & -6 \\ 0 & 0 & 1 & 3 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & 5 & -2 \\ 0 & 0 & 1 & 3 & 1 \end{array}\right). \tag{93.13}$$

Sur cette dernière matrice, les solutions se lisent encore plus facilement.

(4) Cette fois, il y a plus d'équations que d'inconnues. Nous nous attendons donc à ce qu'il n'y ait soit pas de solutions, soit qu'il y ait une ligne « en trop ». Nous pouvons commencer par mettre des zéros sur la colonne du z de la première et seconde ligne en utilisant la troisième.

$$\left(\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 1 & 1 & -1 & 0 \\ 2 & 1 & 0 & 4 \\ 1 & 0 & 1 & 4 \end{array}\right) \sim \left(\begin{array}{ccc|c} 0 & 4 & 0 & 8 \\ 2 & 1 & 0 & 4 \\ 2 & 1 & 0 & 4 \\ 1 & 9 & 1 & 4 \end{array}\right). \tag{93.14}$$

La deuxième et la troisième ligne sont identique. Nous pouvons donc simplement barrer une des deux et continuer comme si nous avions que trois équations. Il est toujours bien de simplifier la première ligne par 4. Cette première ligne ne contient qu'un seul coefficient non nul. Elle donne donc tout de suite $y = 2$. À partir de là, le système est simple à résoudre. Nous avons

$$\begin{cases} y = 2 \\ 2x + y = 4 \\ x + z = 4. \end{cases} \tag{93.15}$$

Sachant que $y = 2$, la seconde équation donne $x = 1$, et sachant que $x = 1$, la troisième donne $1 + z = 4$, c'est-à-dire $z = 3$.

(5) Il y a deux équations pour 4 inconnues, donc on va pouvoir laisser deux variables non résolues. Exprimons x et z en termes de y et t (tout autre choix est bon). La somme des deux équations donne tout de suite $3x + 3y + 3t = -3$, et donc

$$x = -1 - y - t. \tag{93.16}$$

En remettant cela dans la première équation, nous avons $-2 - 2y - 2t + y - 3z - t = 1$, ce qui donne

$$z = \frac{3 + y + 3t}{3}. \tag{93.17}$$

Exercise 394 exoINGE1121La0016

(INGE1121, 1.1) Résoudre les systèmes d'équations suivants :

(1) $\begin{cases} x_1 - 2x_2 + 3x_3 - 2x_4 = 0 \\ 3x_1 - 7x_2 - 2x_3 + 4x_4 = 0 \\ 4x_1 + 3x_2 + 5x_3 + 2x_4 = 0 \end{cases}$

(2) $\begin{cases} 2x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 2x_2 - x_3 + 3x_4 = 4 \\ 3x_1 + 3x_2 + 3x_3 - 3x_4 = 9 \end{cases}$

(3) $\begin{cases} x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ 3x_1 - x_2 - 4x_3 = 0 \end{cases}$

(4) $\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ x_1 + 4x_2 + 7x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 0 \end{cases}$

(5) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 - x_3 + x_4 = -4 \\ x_1 - x_2 + x_3 + x_4 = 2 \end{cases}$

(6) $\begin{cases} x_1 + 3x_2 + 3x_3 = 1 \\ x_1 + 3x_2 + 4x_3 = 0 \\ x_1 + 4x_2 + 3x_3 = 3 \end{cases}$

(7) $\begin{cases} x_1 - 3x_2 + 2x_3 = -6 \\ -3x_1 + 3x_2 - x_3 = 17 \\ 2x_1 - x_2 = 3 \end{cases}$

(8) $\begin{cases} x_1 - 2x_2 + 3x_3 - 2x_4 = 0 \\ 3x_1 - 7x_2 - 2x_3 + 4x_4 = 0 \\ 4x_1 + 3x_2 + 5x_3 + 2x_4 = 0 \end{cases}$

(9) $\begin{cases} 2x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 2x_2 - x_3 + 3x_4 = 4 \\ 3x_1 + 3x_2 + 3x_3 - 3x_4 = 9 \end{cases}$

(10) $\begin{cases} x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ 3x_1 - x_2 - 4x_3 = 0 \end{cases}$

(11) $\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 + 5x_2 + 2x_3 = 0 \\ x_1 + 4x_2 + 7x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 0 \end{cases}$

(12) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 - x_3 + x_4 = -4 \\ x_1 - x_2 + x_3 + x_4 = 2 \end{cases}$

(13) $\begin{cases} x_1+3x_2+3x_3=1 \\ x_1+3x_2+4x_3=0 \\ x_1+4x_2+3x_3=3 \end{cases}$ (14) $\begin{cases} x_1-3x_2+2x_3=-6 \\ -3x_1+3x_2-x_3=17 \\ 2x_1-x_2=3 \end{cases}$

corrINGE1121La0016

Correction of the exercise 394

Nous résolvons les systèmes en utilisant Sage avec le script suivant.

```
# -*- coding: utf8 -*-
"""
Ce script Sage résout un certain nombre
de systèmes d'équations linéaires du cours INGE1121
"""

import outilsINGE

def exercise_1_1_bcdefhi():
   # Exercice 1.1.b (INGE1121)
   A=matrix([ [1,-2,3,-2,0],[3,-7,-2,4,0],[4,3,5,2,0] ])
   v=vector((0,0,0,0,0))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.c (INGE1121)
   A=matrix([ [2,1,-2,3],[3,2,-1,3],[3,3,3,-3] ])
   v=vector((0,4,9))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.d (INGE1121)
   A=matrix([ [1,2,-3],[2,5,2],[3,-1,-4] ])
   v=vector((0,0,0))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.e (INGE1121)
   A=matrix([ [1,2,-1],[2,5,2],[1,4,7],[1,3,3] ])
   v=vector((0,0,0,0))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.f (INGE1121)
   A=matrix([ [1,1,1,1],[1,1,1,-1],[1,1,-1,1],[1,-1,1,1] ])
   v=vector((0,4,-4,2))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.h (INGE1121)
   A=matrix([ [1,3,3],[1,3,4],[1,4,3] ])
   v=vector((1,0,3))
   print outilsINGE.SolveLinearSystem(A,v)
   # Exercice 1.1.i (INGE1121)
   A=matrix([ [1,-3,2],[-3,3,-1],[2,-1,0] ])
   v=vector((-6,17,3))
   print outilsINGE.SolveLinearSystem(A,v)
```

Le résultat est le suivant :

```
The given matrix corresponds to the system
x1 - 2*x2 + 3*x3 - 2*x4 == 0
3*x1 - 7*x2 - 2*x3 + 4*x4 == 0
4*x1 + 3*x2 + 5*x3 + 2*x4 == 0
And the solutions are
[
[x1 == -23/16*r19, x2 == -5/16*r19, x3 == 15/16*r19, x4 == r19, x5 == r18]
```

```
]
The given matrix corresponds to the system
2*x1 + x2 - 2*x3 + 3*x4 == 0
3*x1 + 2*x2 - x3 + 3*x4 == 4
3*x1 + 3*x2 + 3*x3 - 3*x4 == 9
And the solutions are
[
[x1 == 3*r20 - 7, x2 == -4*r20 + 11, x3 == r20, x4 == 1]
]
The given matrix corresponds to the system
x1 + 2*x2 - 3*x3 == 0
2*x1 + 5*x2 + 2*x3 == 0
3*x1 - x2 - 4*x3 == 0
And the solutions are
[
[x1 == 0, x2 == 0, x3 == 0]
]
The given matrix corresponds to the system
x1 + 2*x2 - x3 == 0
2*x1 + 5*x2 + 2*x3 == 0
x1 + 4*x2 + 7*x3 == 0
x1 + 3*x2 + 3*x3 == 0
And the solutions are
[
[x1 == 9*r21, x2 == -4*r21, x3 == r21]
]
The given matrix corresponds to the system
x1 + x2 + x3 + x4 == 0
x1 + x2 + x3 - x4 == 4
x1 + x2 - x3 + x4 == -4
x1 - x2 + x3 + x4 == 2
And the solutions are
[
[x1 == 1, x2 == -1, x3 == 2, x4 == -2]
]
The given matrix corresponds to the system
x1 + 3*x2 + 3*x3 == 1
x1 + 3*x2 + 4*x3 == 0
x1 + 4*x2 + 3*x3 == 3
And the solutions are
[
[x1 == -2, x2 == 2, x3 == -1]
]
The given matrix corresponds to the system
x1 - 3*x2 + 2*x3 == -6
-3*x1 + 3*x2 - x3 == 17
2*x1 - x2 == 3
And the solutions are
[
[x1 == 37, x2 == 71, x3 == 85]
]
```

Exercise 395 exoLineraire0002

Exercice 7, page 86. Discuter les systèmes suivants

(1) $\begin{cases} mx+y=1 \\ x+my=1 \end{cases}$

(2) $\begin{cases} \lambda x+y+m=1 \\ x+\lambda y+z=1 \\ x+y+\lambda z=1 \end{cases}$

corrLineraire0002

Correction of the exercise 395

<+Lineraire0002+>

Exercise 396 exoINGE1121La0010

Pour chacun des systèmes suivants $A \cdot X = B$,

(1) Résoudre le système par échelonnement,

(2) Calculer A^{-1},

(3) Vérifier votre réponse en calculant $A^{-1}B$. Qu'êtes-vous censé obtenir ?

Les énoncés sont

(1)

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 3 & 2 & 2 \\ 5 & 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 10 \\ 1 \\ 4 \end{pmatrix} \tag{93.18}$$

corrINGE1121La0010

Correction of the exercise 396

Nous utilisons Sage pour fournir la réponse. Le code suivant résout le système et donne l'inverse de la matrice :

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_1_3():
   A=matrix([[2,1,-2],[3,2,2],[5,4,3]])
   v=vector((10,1,4))
   print outilsINGE.SolveLinearSystem(A,v)
   print "Matrice inverse :"
   print A.inverse()
```

La sortie est ici :

```
The given matrix corresponds to the system
2*x1 + x2 - 2*x3 == 10
3*x1 + 2*x2 + 2*x3 == 1
5*x1 + 4*x2 + 3*x3 == 4
And the solutions are
[
[x1 == 1, x2 == 2, x3 == -3]
]
Matrice inverse :
[  2/7  11/7  -6/7]
[ -1/7 -16/7  10/7]
[ -2/7   3/7  -1/7]
```

Exercise 397 exoINGE1121La0009

Discuter et résoudre le système

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 1 & -1 \\ 3 & 1 & r & 3 \\ -2 & 6 & 4 & s \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}. \tag{93.19}$$

corrINGE1121La0009

Correction of the exercise 397

Nous échelonons la matrice de façon usuelle. Tellement usuelle que nous en confions les calculs à Sage.

```
# -*- coding: utf8 -*-

def exercise_1_5():
    var('r,s,y1,y2,y3,y4')
    A=matrix([[1,2,3,4,y1],[2,-1,1,-1,y2],[3,1,r,3,y3],[-2,6,4,s,y4]])
    A[3]=A[3]+A[1]
    A[1]=A[1]-2*A[0]
    A[2]=A[2]-3*A[0]
    print A
    A[2]=A[2]-A[1]
    A[3]=A[3]+A[1]
    print A
```

D'un point de vue technique, remarquez que la numérotation des lignes et colonnes commence à zéro, et non à un ! C'est comme ça dans beaucoup de langages de programmation.

La sortie est

```
[         1          2          3          4         y1]
[         0         -5         -5         -9 -2*y1 + y2]
[         0         -5      r - 9         -9 -3*y1 + y3]
[         0          5          5      s - 1    y2 + y4]
[                1                2                3                4                y1]
[                0               -5               -5               -9       -2*y1 + y2]
[                0                0            r - 4                0    -y1 - y2 + y3]
[                0                0                0            s - 10 -2*y1 + 2*y2 + y4]
```

À partir d'ici, nous pouvons discuter un petit peu. D'abord, la dernière ligne donne deux cas

(1) $s = 10$. Dans ce cas, il y a deux sous-cas

  (1a) $-2y_1 + 2y_2 + y_4 = 0$. Ici, on peut simplement barrer la dernière ligne, et continuer. Il y aura une infinité de solutions.

  En continuant, nous tombons sur la troisième ligne. Il y a une discussion à faire selon que $r = 4$ ou non.

  (1b) $-2y_1 + 2y_2 + y_4 \neq 0$. Ici, la dernière ligne dit qu'il n'y a pas de solutions.

La troisième ligne donne une discussion similaire selon que $r = 4$ ou non en fonction de $-y_1 - y_2 + y_3$. Les cas avec des solutions sont

(1) $s = 10$, $r = 4$, $-2y_1 + 2y_2 + y_4 = 0$, $-y_1 - y_2 + y_3 = 0$. Dans ce cas, le système devient

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & -5 & -5 & -9 & -2y_1 + y_2 \end{array}\right). \tag{93.20}$$

La solution de ce système comporte deux paramètres et est

$$\begin{aligned} x_1 &= -\frac{2}{5}\alpha - \beta + y_1 + \frac{2}{5}y_2 \\ x_2 &= -\frac{9}{5}\alpha - \beta + \frac{1}{5}y_2 \\ x_3 &= \beta \\ x_4 &= \alpha. \end{aligned} \tag{93.21}$$ EqSolsUn15

Notez que les y ne sont pas des paramètres. Ce sont des données du problème.

(2) $s = 10$, $r \neq 4$, $-2y_1 + 2y_2 + y_4 = 0$, $-y_1 - y_2 + y_3 \neq 0$. Dans ce cas, seule la dernière ligne doit être barrée dans le système. Il reste un système de trois équations.

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & -5 & -5 & -9 & -2y_1 + y_2 \\ 0 & 0 & r-4 & 0 & -y_1 - y_2 + y_3 \end{array}\right). \tag{93.22}$$

Les solutions sont les mêmes que celles (93.21), mais x_3 n'est plus un paramètre :

$$\beta = x_3 = \frac{-y_1 - y_2 + y_3}{r-4}. \tag{93.23}$$

(3) $s \neq 10$, $r = 4$, $-2y_1 + 2y_2 + y_4 =\neq$, $-y_1 - y_2 + y_3 = 0$. Ce cas-ci est similaire au précédent.

(4) $s \neq 10$, $r \neq 4$, $-2y_1 + 2y_2 + y_4 \neq 0$, $-y_1 - y_2 + y_3 \neq 0$. Ici, c'est un système complet de 4 équations à 4 inconnues à résoudre.

Un problème de Bachet (XVIIème siècle) Exercise 398 exoLineraire0003

Trois hommes ont chacun certaine somme d'écus. Le premier donne des siens aux deux autres autant qu'ils en ont chacun ; en après le second en donne aux deux autres autant qu'ils en ont chacun ; finalement le troisième en donne aux deux autres autant qu'ils en ont chacun : cela fait, chacun se trouve 8 écus. On demande combien chacun en avait du commencement.

Extrait des *Problèmes plaisants et délectables qui se font par les nombres*, publié en 1612. corrLineraire0003

Correction of the exercise 398

Disons qu'au début, ils ont respectivement x, y et z écus. Lorsque le premier donne, il en donne y au second et z au troisième. Leurs avoirs sont donc

$$\begin{aligned} x - y - z \\ 2y \\ 2z. \end{aligned} \tag{93.24}$$

À ce moment, le second donne $x - y - z$ au premier et $2z$ au troisième. Ils ont donc

$$\begin{aligned} 2(x - y - z) \\ 2y - (x - y - z) - 2z \\ 4z. \end{aligned} \tag{93.25}$$

En simplifiant :

$$\begin{aligned} 2(x - y - z) \\ 3y - x - z \\ 4z. \end{aligned} \tag{93.26}$$

Maintenant, le troisième donne $2(x - y - z)$ au premier et $3y - x - z$ au second. Ils se retrouvent donc avec

$$\begin{aligned} 4(x - y - z) \\ 2(3y - x - z) \\ -x - y + 7z. \end{aligned} \tag{93.27}$$

Petite vérification : la somme des trois vaut bien $x+y+z$. Il faut maintenant juste maintenant résoudre le système donné en égalant ces trois quantités à 8. La solution est donnée par $x=13$, $y=7$ et $z=4$.

Un problème stupide (XXième siècle) **Exercise 399** exoLineraire0004

Un père a 25 ans de plus que son fils. Dans 7 ans, il aura 5 fois l'âge de son fils. Que fait le père ?

corrLineraire0004

Correction of the exercise 399

<+Lineraire0004+>

93.2 Opérations sur les matrices

Exercise 400 exoLineraire0005

On donne les matrices

$$A=\begin{pmatrix}1&2\\3&4\\5&6\end{pmatrix},\quad B=\begin{pmatrix}-3&-2\\1&-5\\4&3\end{pmatrix}. \tag{93.28}$$

Déterminer la matrice D telle que $A+B-D=0$.

corrLineraire0005

Correction of the exercise 400

Si $A+B-D=0$, alors $D=A+B$. Petit calcul :

$$D=\begin{pmatrix}1&2\\3&4\\5&6\end{pmatrix}+\begin{pmatrix}-3&-2\\1&-5\\4&3\end{pmatrix}=\begin{pmatrix}-2&0\\4&-1\\9&9\end{pmatrix}. \tag{93.29}$$

Exercise 401 exoLineraire0006

Si $A=\begin{pmatrix}1&-1\\2&-1\end{pmatrix}$ et $B=\begin{pmatrix}1&1\\4&-1\end{pmatrix}$, montrer que $(A+B)^2=A^2+B^2$.

corrLineraire0006

Correction of the exercise 401

En général, nous avons

$$(A+B)^2=(A+B)(A+B)=A^2+AB+BA+B^2. \tag{93.30}$$

Il n'est pas vrai de dire $AB=BA$, parce que ce n'est pas toujours le cas avec des matrices. Ici, nous voulons que $(A+B)^2=A^2+B^2$, c'est-à-dire $AB+BA=0$. Calculons les produits AB et BA avec les matrices proposées :

$$AB=\begin{pmatrix}1&-1\\2&-1\end{pmatrix}\begin{pmatrix}1&1\\4&-1\end{pmatrix}=\begin{pmatrix}-3&2\\-2&3\end{pmatrix}, \tag{93.31}$$

tandis que

$$BA=\begin{pmatrix}3&-2\\2&-3\end{pmatrix}. \tag{93.32}$$

Nous avons donc bien $AB+BA=0$ dans ce cas ci.

Exercise 402 exoLineraire0007

Montrer que les matrices

$$A=\begin{pmatrix}0&1&-1\\4&-3&4\\3&-3&4\end{pmatrix}\quad B=\begin{pmatrix}-1&0&a\\0&-1&b\\0&0&1\end{pmatrix}\quad, \tag{93.33}$$

avec a, $b\in\mathbb{R}$ sont involutives, c'est-à-dire telles que $X^2=\mathbb{1}$.

corrLineraire0007

Correction of the exercise 402

Il s'agit de calculer et de voir que

$$AA = BB = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{93.34}$$

Exercise 403 exoLineraire0008

Soit

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}. \tag{93.35}$$

Trouver toutes les matrices X telles que $AX = A$. corrLineraire0008

Correction of the exercise 403

Nous cherchons la matrice X sous la forme

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{93.36}$$

Le produit AX vaut

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a + c & 2b + d \\ 2a + c & 2b + d \end{pmatrix}. \tag{93.37}$$

Les équations pour a, b, c et d sont donc

$$\begin{cases} 2a + c = 2 \\ 2b + d = 1. \end{cases} \tag{93.38}$$

La matrice X générale est donc de la forme

$$X = \begin{pmatrix} a & b \\ 2 - 2a & 1 - 2b \end{pmatrix}. \tag{93.39}$$

Exercise 404 exoLineraire0009

Démontrer que si les matrices A et B sont telles que $AB = A$ et $BA = B$, alors $A^2 = A$ et $B^2 = B$. corrLineraire0009

Correction of the exercise 404

Par définition, $A^2 = AA$. En remplaçant le premier A par AB, nous avons

$$A^2 = ABA. \tag{93.40}$$

Maintenant, nous remplaçons BA par B, ce qui donne

$$A^2 = AB, \tag{93.41}$$

et nous savons par hypothèse que $AB = A$. Le même genre de petit jeu donne $B^2 = B$.

Exercise 405 exoLineraire0010

Montrer que

$$\begin{pmatrix} \lambda & \lambda \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^n \\ 0 & \lambda^n \end{pmatrix} \tag{93.42}$$

pour tout $\lambda \in \mathbb{R}$ et tout $n \in \mathbb{N}_0$. corrLineraire0010

Correction of the exercise 405

Cet exercice se fait par récurrence. Lorsque $n = 1$, c'est vrai. Supposons que ce soit vrai pour n, et montrons que c'est vrai pour $n + 1$. Nous avons

$$\begin{pmatrix} \lambda & \lambda \\ 0 & \lambda \end{pmatrix}^{n+1} = \begin{pmatrix} \lambda & \lambda \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} \lambda & \lambda \\ 0 & \lambda \end{pmatrix}^n. \tag{93.43}$$

En remplaçant la dernière matrice par l'hypothèse de récurrence, et en effectuant le produit, on trouve ce qu'il faut.

Exercise 406 exoLineraire0011

À quelles conditions sur les matrices A et B peut-on écrire

(1) $(A-B)^2 = A^2 + B^2$

(2) $(A+B)(A-B) = A^2 - B^2$.

corrLineraire0011

Correction of the exercise 406

(1) $(A-B)^2 = A^2 - AB - BA + b^2$, cela vaut $A^2 + B^2$ lorsque $AB = -BA$.

(2) $(A+B)(A-B) = A^2 - AB + BA - B^2$, cela vaut $A^2 - B^2$ lorsque $AB = BA$.

Exercise 407 exoLineraire0012

Trouve toutes les matrices B qui commutent avec la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{93.44}$$

corrLineraire0012

Correction of the exercise 407

Si nous posons

$$X = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix}, \tag{93.45}$$

nous trouvons

$$\begin{aligned} AX &= \begin{pmatrix} e & f & g & h \\ i & j & k & l \\ m & n & o & p \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ XA &= \begin{pmatrix} 0 & a & b & c \\ 0 & e & f & g \\ 0 & i & j & k \\ 0 & m & n & o \end{pmatrix}. \end{aligned} \tag{93.46}$$

En égalisant les deux, nous trouvons

$$j = e = i = m = n = o = 0, \tag{93.47}$$

ainsi que $f = a$, $b = g$, $h = c$, $j = e$, $k = f$, $l = g$, $n = i$, $o = j$, $p = k = f$. Étant donné que $e = i = 0$, nous avons directement $j = n = 0$. Remarquez que certaines variables arrivent dans plusieurs équations. Après avoir bien posé toutes les égalités, le résultat est

$$\begin{pmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{pmatrix}. \tag{93.48}$$

Exercise 408 exoINGE1121La0008

Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Montrer par récurrence que $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

corrINGE1121La0008

Correction of the exercise 408

Lorsque $n = 1$, la proposition est évidente. Supposons maintenant que la proposition soit vraie pour un certain entier k, et prouvons qu'alors c'est vrai pour $k+1$.

Nous devons calculer

$$A^{k+1} = A^k A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{93.49}$$

EqRecur18INGE

Pour écrire cela, nous avons utilisé l'hypothèse de récurrence

$$A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}. \tag{93.50}$$

En effectuant le produit matriciel dans (93.49), nous trouvons

$$A^{k+1} = \begin{pmatrix} 1 & k+1 \\ 0 & 1 \end{pmatrix}, \tag{93.51}$$

ce qui est bien ce que nous voulions.

93.3 Espaces vectoriels

Exercise 409 exoLineraire0013

Exercice 1, page 79. Quels sont parmi les ensembles suivants qui ont une structure d'espace vectoriel réel ?

(1) L'ensemble des polynômes à coefficients réels en une indéterminée, de degré $\leqslant 3$,
(2) L'ensemble des polynômes à coefficients réels en une indéterminée, de degré 3.
(3) L'ensemble des triplets (a, b, c) de nombres réels tels que $a + 2b - c = 0$.
(4) L'ensemble des triplets (a, b, c) de nombres réels tels que $a^2 + b^2 + c^2 = 1$.
(5) L'ensemble des triplets (a, b, c) de nombres réels tels que $a - b = 1$.
(6) L'ensemble des triplets (a, b, c) de nombres réels tels que $a + b = 0$ et $a + c = 0$.
(7) L'ensemble des couples de nombres entiers.
(8) L'ensemble des couples (a, b) de réels tels que $a^2 + b^2 \neq 0$.

corrLineraire0013

Correction of the exercise 409

(1) oui.
(2) Non, et la différence avec le précédent est que la somme de deux polynômes de degré trois peut n'être que de degré deux : $(x^3 + x^2) - (x^3 + x) = x^2 - x$ par exemple.
(3) Oui, c'est un plan. Si (a, b, c) et (a', b', c') vérifient la condition, regardons que $(a + a', b + b', c + c')$ vérifie la condition :

$$(a + a') + 2(b + b') - (c + c') = 0. \tag{93.52}$$

(4) Non, par exemple $(1, 0, 0)$ est dans l'ensemble, mais pas $(2, 0, 0)$ qui est le double.
(5) Non, c'est un plan afin. Par exemple $(2, 1) + (5, 4) = (7, 5)$, mais $(7 - 5) \neq 1$, alors que $(2, 1)$ et $(5, 4)$ vérifient la condition.
(6) Oui, c'est l'intersection de deux plans. Si (a, b, c) et (a', b', c') vérifient les deux conditions, c'est facile de voir que $(a + a', b + b', c + c')$ vérifie les deux conditions.
(7) Non, si $(n, m) \in \mathbb{N}^2$, il suffit de le multiplier par $\sqrt{2}$ et ce n'est plus dans $\mathbb{N}^2$.
(8) Non, par exemple $(a, b) - (a, b) = (0, 0)$ et évidement, $(0, 0)$ n'est pas dans l'ensemble.

Exercise 410 exoLineraire0014

Exercice 2, page 19. Dans $\mathbb{R}^3$, calculer la somme vectorielle $5(-3\sqrt{2}, \frac{1}{10}, 2) - (2, -1, \frac{3}{5}) + 7(\sqrt{2}, 0, 2)$.

corrLineraire0014

Correction of the exercise 410

On commence par calculer les produits à part :

$$\begin{aligned} 5(-3\sqrt{2},\frac{1}{10},2)&=(-14\sqrt{2},\frac{2}{2},10)\\ 7(\sqrt{2},0,2)&=(7\sqrt{2},0,14), \end{aligned} \tag{93.53}$$

et puis on fait la somme

$$(-14\sqrt{2},\frac{2}{2},10)-(2,-1,\frac{3}{5})+=(7\sqrt{2},0,14)=(-8\sqrt{2}-2,\frac{3}{2},\frac{117}{5}). \tag{93.54}$$

Exercise 411 exoLineraire0018

Exercice 3, page 79. Le vecteur $(1,2,3)$ de $\mathbb{R}^3$ est-il combinaison linéaire des vecteurs $(1,1,1)$, $(3,5,2)$ et $(1,3,0)$? Même question pour le vecteur $(0,-2,1)$. corrLineraire0018

Correction of the exercise 411

En général, il faut trouver les coefficients λ_1, λ_2 et λ_3 tels que

$$\lambda_1\begin{pmatrix}1\\1\\1\end{pmatrix}+\lambda_2\begin{pmatrix}3\\5\\2\end{pmatrix}+\lambda_3\begin{pmatrix}1\\3\\0\end{pmatrix}=\begin{pmatrix}a\\b\\c\end{pmatrix}. \tag{93.55}$$

Cela revient à résoudre le système

$$\begin{cases}\lambda_1+3\lambda_2+\lambda_3=a\\ \lambda_1+5\lambda_2+3\lambda_3=b\\ \lambda_1+2\lambda_2=c\end{cases} \tag{93.56}$$

Nous résolvons cela en termes de matrices :

$$\left(\begin{array}{ccc|c}1&3&1&a\\1&5&3&b\\1&2&0&c\end{array}\right)\sim\left(\begin{array}{ccc|c}0&1&1&a-c\\0&2&2&b-a\\1&2&0&c\end{array}\right) \tag{93.57}$$

Au niveau des coefficients, la seconde ligne est juste le double de la première, donc pour avoir une solution, il faut absolument que le terme indépendant de la seconde soit également la moitié : $b-a=2(a-c)$. Nous pouvons continuer et tomber sur la matrice suivante :

$$\left(\begin{array}{ccc|c}0&0&0&3a-b-2c\\0&2&2&b-a\\1&2&0&c\end{array}\right). \tag{93.58}$$

Lorsqu'on regarde le vecteur

$$\begin{pmatrix}a\\b\\c\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}, \tag{93.59}$$

on a la condition $3a-b-2c=3-2-6\neq 0$, donc ce vecteur ne peut pas être écrit comme combinaison des trois vecteurs donnés. Avec l'autre, par contre, ça fonctionne mieux :

$$\begin{pmatrix}a\\b\\c\end{pmatrix}=\begin{pmatrix}0\\-2\\1\end{pmatrix}, \tag{93.60}$$

donne le système

$$\begin{cases}\lambda_1+2\lambda_2=1\\ 2\lambda_2+2\lambda_3=-2.\end{cases} \tag{93.61}$$

Comme ce sont deux équations avec 3 inconnues, il n'y a pas d'espoir de trouver une seule solution. Nous résolvons donc par rapport à λ_1 et λ_3 en laissant λ_2 comme paramètre. Nous trouvons $\lambda_1 = 1 - 2\lambda_2$ et $\lambda_3 = -1 - \lambda_2$.

Pour chaque λ_2, cela donne un triple $(\lambda_1, \lambda_2, \lambda_3)$ tel que

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}. \tag{93.62}$$

Prenez par exemple, $\lambda_2 = 0$, on trouve $\lambda_1 = 1$ et $\lambda_3 = -1$.

Exercise 412 exoLineraire0015

Exercice 5, page 80. Trouver une condition nécessaire et suffisante sur les nombres réels a, b, et c pour que les vecteurs $(0, 1, -1)$, $(5, -2, 1)$ et (a, b, c) forme une partie liée de $\mathbb{R}^3$. corrLineraire0015

Correction of the exercise 412

Une condition très intéressante pour savoir si trois vecteurs forment une base et de les mettre dans une matrice et de voir le déterminant. Si il est non nul, c'est une base ; si il est nul, ce n'est pas une base. Ici, on a

$$\begin{vmatrix} 0 & 5 & a \\ 1 & -2 & b \\ -1 & 1 & c \end{vmatrix} = -5(c + b) + a(1 - 2). \tag{93.63}$$

La condition pour que les trois soient non libres est que le déterminant soit nul, c'est-à-dire

$$a + 5(b + c) = 0. \tag{93.64}$$

Exercise 413 exoLineraire0016

Exercice 6, page 80. Les parties suivantes de $\mathbb{R}^3$ sont-elles libres ?

(1) $\{(1, 2, 1), (1, 1, 1)\}$,
(2) $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$,
(3) $\{(0, 0, 0), (3, 2, 1)\}$,
(4) $\{(2, 1, -3), (1, 4, 0), (4, 9, -3)\}$,
(5) $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}$,
(6) $\{(1, 0, 1), (0, -8, 0), (1, 7, -2)\}$,
(7) $\{(-1, 0, 1)\}$

corrLineraire0016

Correction of the exercise 413

(1) Oui parce que deux vecteurs dans $\mathbb{R}^3$, il devraient être multiples l'un de l'autre pour ne pas être libres.
(2) Oui, base canonique
(3) Oui : tout les vecteurs sont libres avec $(0, 0, 0)$.
(4) On utilise la critère du déterminant (voir exercice 412). Ici nous avons

$$\begin{vmatrix} 2 & 1 & 4 \\ 1 & 4 & 9 \\ -3 & 0 & -3 \end{vmatrix} = -3(1 \cdot 9 - 4 \cdot 4) = 0. \tag{93.65}$$

Donc le système n'est pas libre.
(5) Plus de 3 vecteur dans $\mathbb{R}^3$ ne peuvent pas être libres.
(6) Le déterminant

$$\begin{vmatrix} 1 & 0 & 1 \\ 0 & -5 & 7 \\ 1 & 0 & -2 \end{vmatrix} \tag{93.66}$$

est non nul, donc la partie est libre.

(7) Un seul vecteur, c'est toujours libre.

Exercise 414 exoLineraire0017

Exercice 7, page 80. Dans l'espace vectoriel E_0, déterminer l'ensemble des points qui sont combinaisons linéaire des points des parties suivantes de E_0 :

(1) Une sphère de centre 0.

(2) Une droite ne passant pas par o.

(3) Un point distinct de o.

(4) Un plan passant par o.

(5) Un plan ne passant pas par o.

(6) La réunion de deux droites, à discuter suivant les positions relatives de ces deux droites.

corrLineraire0017

Correction of the exercise 414

(1) Tous parce que, en particulier, la base canonique est contenue dans la sphère

(2) Non parce qu'une droite est donné seulement par deux vecteurs. Tous les points de D sont de la forme

$$a + \lambda(b - a) \tag{93.67}$$

si $a, b \in D$. En faisant des combinaisons de tels vecteurs, on ne peut pas faire d'autre choses que des combinaisons des vecteurs a et $(b - a)$.

(3) Avec un seul point, on ne génère qu'une seule droite

(4) On ne génère que le plan (qui est, lui-même, un espace vectoriel).

(5) oui

Exercise 415 exoLineraire0019

Exercice 8, page 80. Parmi les sous-ensembles suivants, lesquels sont des parties génératrices de $\mathbb{R}^3$? Lesquels sont des bases de $\mathbb{R}^3$?

(1) $\{(1,2,1),(1,1,1)\}$,

(2) $\{(1,0,0),(0,1,0),(0,0,1)\}$,

(3) $\{(0,0,0),(3,2,1)\}$,

(4) $\{(2,1,-3),(1,4,0),(4,9,-3)\}$,

(5) $\{(1,0,0),(0,1,0),(0,0,1),(0,1,1),(1,0,1),(0,0,1)\}$,

(6) $\{(1,0,1),(0,-5,0),(1,7,-2)\}$,

(7) $\{(-1,0,1)\}$.

corrLineraire0019

Correction of the exercise 415

<+Lineraire0019+>

Exercise 416 exoLineraire0020

Exercice 9, page 81. Quelles sont les composantes du vecteur $(1,2) \in \mathbb{R}^2$ dans les bases suivantes ?

(1) $\{(1,0),(0,1)\}$,

(2) $\{(0,1),(1,0)\}$,

(3) $\{(1,2),(1,1)\}$,

(4) $\{(1,1),(2,1)\}$,

(5) $\{(-1,3),(4,0)\}$

corrLineraire0020

Correction of the exercise 416

(1) C'est la base canonique, donc le vecteur a pour composantes $(1,2)$.

(2) Il faut résoudre
$$\lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{93.68}$$
Facile, la réponse est que les coordonnées $(\lambda_1, \lambda_2) = (2, 1)$.

(3) À résoudre, le système
$$\begin{cases} a + b = 1 \\ 2a + b = 2 \end{cases}. \tag{93.69}$$
La solution est $a = 1$ et $b = 0$, donc les coordonnées sont $(1, 0)$.

(4) À résoudre :
$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \tag{93.70}$$
la solution est $(\lambda_1, \lambda_2) = (1,)1$.

(5) Le système est
$$\begin{cases} -a + 4b = 1 \\ 3a = 2 \end{cases}, \tag{93.71}$$
donc $(2/3, -5/12)$.

Exercise 417 exoLineraire0021

Exercice 10, page 81. Quel est le vecteur de $\mathbb{R}^3$ qui a pour composante $(1, 2, -1)$ dans la base $e_1 = (1, 1, 1)$, $e_2 = (0, 1, 2)$, $e_3 = (0, -1, 0)$ de $\mathbb{R}^3$? corrLineraire0021

Correction of the exercise 417

Le vecteur qui a pour composante $(\lambda_1, \lambda_2, \lambda_3)$ dans la base e_1, e_2, e_3 est le vecteur $\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3$. Ici il s'agit de
$$1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + (-1) \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}. \tag{93.72}$$

Exercise 418 exoLineraire0022

Exercice 11, page 81. Donner une base de $\mathbb{R}^2$ dans laquelle les composantes du vecteur $(1, 2)$ sont $(0, 1)$. corrLineraire0022

Correction of the exercise 418

On demande une base dans lequel le second vecteur de base est $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Presque n'importe quoi est bon comme premier vecteur (juste pas un multiple de celui qu'on a déjà choisit). Par exemple
$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{93.73}$$

Exercise 419 exoLineraire0023

Décrire une base de $\mathbb{R}^3$ dans laquelle les composantes des vecteurs $(7, -5, 2)$ et $(\sqrt{3}, 29, -8)$ sont aussi simples que possible. corrLineraire0023

Correction of the exercise 419

Dans la base
$$\begin{pmatrix} 7 \\ -5 \\ 2 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \sqrt{3} \\ 29 \\ -8 \end{pmatrix}, \tag{93.74}$$
les deux vecteurs donnés ont évidement les coordonnées $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Exercise 420 exoLineraire0024

Exercice 13, page 81. Dans $\mathbb{R}^2$, on donne les bases
$$E = \begin{cases} e_1 = (1, 0) \\ e_2 = (0, 1) \end{cases} \quad F = \begin{cases} f_1 = (1, 1) \\ f_2 = (1, 2) \end{cases} \quad G = \begin{cases} g_1 = (1, 2) \\ g_2 = (2, 1). \end{cases} \tag{93.75}$$

(1) Quelles sont les composantes du vecteur (a,b) dans les bases E, F et G ?
(2) Quelles sont les composantes du vecteur $7e_1-2e_2$ dans les bases F et G ?
(3) Quelles sont les composantes du vecteur f_1+3f_2 dans les bases E et F ?
(4) Quelles sont les composantes du vecteur g_1 dans les bases E et F ?
corrLineraire0024

Correction of the exercise 420

(1) La base E étant la base canonique, le vecteur $\begin{pmatrix}a\\b\end{pmatrix}$ a pour coordonnées (a,b). Pour la base F, il faut résoudre le système

$$\begin{pmatrix}a\\b\end{pmatrix}=\lambda_1\begin{pmatrix}1\\1\end{pmatrix}+\lambda_2\begin{pmatrix}1\\2\end{pmatrix}. \tag{93.76}$$

La résolution donne donc $(2a-b,b-a)$. Pour trouver les coordonnées de ce même vecteur dans la base G, il faut résoudre

$$\begin{cases}\lambda_1+2\lambda_2=a\\ 2\lambda_1+\lambda_2=b.\end{cases} \tag{93.77}$$

Cela se résous facilement par substitution, et le résultat est $\lambda_1=\frac{2b-a}{3}$ et $\lambda_2=\frac{2a-b}{3}$.

(2) Nous avons d'abord $7e_1-2e_2=\begin{pmatrix}7\\-2\end{pmatrix}$. Ensuite, nous devons résoudre

$$\begin{pmatrix}7\\-2\end{pmatrix}=\lambda_1\begin{pmatrix}1\\2\end{pmatrix}+\lambda_2\begin{pmatrix}2\\1\end{pmatrix}. \tag{93.78}$$

Cela donne $(-11/3,16/3)$.

Exercise 421 exoLineraire0025

Dans l'espace vectoriel des polynômes à coefficients réels de degré $\leqslant 2$, trouver une base du sous-espace engendré par les polynômes $5x^2+3x+2$, $-5x^2+9x+8$, $-10x^2+18x+12$ et x^2+9x+6.
corrLineraire0025

Correction of the exercise 421

<+Lineraire0025+>

Exercise 422 exoLineraire0026

Exercice 15, page 82. Les fonctions e^x, $\sin(x)$ et $\cos(x)$ sont-elles une partie libre de l'espace vectoriel des fonctions continues de $\mathbb{R}$ vers $\mathbb{R}$? Même question pour les fonctions $\sin^2(x)$, $\cos^2(x)$, $\sin(2x)$ et $\cos(2x)$.
corrLineraire0026

Correction of the exercise 422

Il faut essayer de trouver des constantes $a,b,c\in\mathbb{R}$ telles que la fonction

$$be^x+b\sin(x)+c\cos(x) \tag{93.79}$$

soit nulle, c'est-à-dire telle que cette somme soit zéro pour tout x. Cela n'est pas possible (et donc la partie est libre). En effet, comme $e^x\to\infty$, de toutes façons, il faut $a=0$. Maintenant, en $x=0$, $\sin(x)=0$ et $\cos(x)=1$, nous devons donc avoir $c=0$. De là, $b=0$.

Exercise 423 exoLineraire0027

Dans l'espace vectoriel F des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on considère le sous-ensemble L des fonctions f de la forme

$$f(x)=A(x)\cos(x)+B(x)\sin(x) \tag{93.80}$$

où $A(x)$ et $B(x)$ sont deux polynômes en x de degré $\leqslant 1$.

(1) Montrer que L est un sous-espace de F,
(2) Montrer que les fonctions $f_1(x)=\cos(x)$, $f_2(x)=\sin(x)$, $f_3(x)=x\cos(x)$, $f_4(x)=x\sin(x)$ est une base de L.
corrLineraire0027

Correction of the exercise 423

<+Lineraire0027+>

93.3.1 Orthogonalité

Exercise 424 exoLineraire0028

Exercice 17, page 82. On donne les vecteurs $x = (\frac{1}{3}, \frac{2}{3}, \frac{-2}{3})$ et $y = (\frac{2}{3}, \frac{1}{3}, \frac{2}{3})$ de $\mathbb{R}^3$.

(1) Montrer qu'ils sont orthogonaux,

(2) Construire un vecteur z perpendiculaire à x et y.

(3) Donner une base orthonormée de $\mathbb{R}^3$ qui contient le vecteur $v = \frac{1}{\sqrt{6}}(1, 2, 1)$.

corrLineraire0028

Correction of the exercise 424

(1) Pour voir qu'ils sont orthogonaux : produit scalaire.

$$x \cdot y = \begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix} \cdot \begin{pmatrix} 2/3 \\ 1/3 \\ 2/3 \end{pmatrix} = \frac{1}{9} + \frac{2}{9} - \frac{4}{9} = 0, \tag{93.81}$$

donc ils sont orthogonaux.

(2) Le produit vectoriel de x et y donne un vecteur perpendiculaire à la fois à x et à y :

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \end{vmatrix} = 2e_1 - 2e_2 - e_3. \tag{93.82}$$

Exercise 425 exoINGE1121La0001

Calculer une base orthonormée du sous-espace de $\mathbb{R}^4$ défini par

$$V = \{(\alpha, \alpha + \beta, 2\alpha + \beta - 2\gamma, \beta + 2\gamma) \text{ tel que } \alpha, \beta, \gamma \in \mathbb{R}\}. \tag{93.83}$$

corrINGE1121La0001

Correction of the exercise 425

Étant donné que l'espace V est donné avec 3 paramètre, il peut être au plus de dimension 3. Nous trouvons 3 vecteurs de cet espace en posant successivement $\alpha = 1$, $\beta = 1$ et $\gamma = 1$ en laissant à chaque fois les deux autres paramètres à zéro. Les trois vecteurs trouvés sont :

$$\begin{aligned} v_1 &= (1, 1, 2, 0) \\ v_2 &= (0, 1, 1, 1) \\ v_3 &= (0, 0, -2, 2) \end{aligned} \tag{93.84}$$

Afin de vérifier que ces trois vecteurs sont linéairement indépendants, nous les mettons dans une matrice et nous vérifions que le rang de la matrice est bien 3 :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & -2 \\ 0 & 1 & 2 \end{pmatrix}. \tag{93.85}$$

Le rang de cette matrice est bien 3 parce que par exemple le déterminant de la sous-matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & -2 \end{pmatrix} \tag{93.86}$$

est non nul.

Les vecteurs v_i trouvé forment donc une base de V. Nous trouvons une base orthogonale en appliquant la méthode de Gram-Schmidt. Le résultat est

$$w_1 = (1, 1, 2, 0) w_2 = (-1/2, 1/2, 0, 1) w_1 = (2/3, 2/3, -2/3, 2), \tag{93.87}$$

dont les normes sont respectivement $\sqrt{6}$, $\sqrt{3/2}$ et $4/\sqrt{3}$. Une base orthonormée est donnée par les vecteurs w_i divisés par leurs normes.

Exercise 426 exoINGE1121La0002

Calculer une base orthonormée du sous-espace de $\mathbb{R}^4$ défini par les équations

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 0 \\ x_1 - 2x_2 + 2x_3 &= 0. \end{aligned} \tag{93.88}$$

corrINGE1121La0002

Correction of the exercise 426

La première chose à faire est de trouver les vecteurs de l'espace V. Pour ce faire, nous résolvons le système donné par la matrice

$$\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -2 & 2 & 0 \end{pmatrix}. \tag{93.89}$$

Le rang de cette matrice étant 2, nous savons que l'espace des solutions sera de dimension donnée par

$$\dim V = \dim(\mathbb{R}^4) - 2 = 2. \tag{93.90}$$

Il est assez facile de trouver deux vecteurs linéairement indépendants dans V, par exemple en posant $x_3 = 1$ et $x_4 = 0$ et puis $x_3 = 0$ et $x_4 = 1$. Les vecteurs ainsi trouvés sont

$$\begin{aligned} v_1 &= (-\frac{4}{3}, \frac{1}{3}, 1, 0) \\ v_2 &= (\frac{2}{3}, \frac{1}{3}, 0, 1). \end{aligned} \tag{93.91}$$

Ces deux vecteurs sont linéairement indépendants parce que la matrice qu'ils forment est de rang deux.

La méthode de Gram-Schmidt appliqué à ces deux vecteurs fournit

$$\begin{aligned} w_1 &= (-\frac{4}{3}, \frac{1}{3}, 1, 0) \\ w_2 &= (\frac{4}{13}, \frac{11}{26}, \frac{7}{26}, 1), \end{aligned} \tag{93.92}$$

et leurs normes sont $\sqrt{26}/3$ et $\sqrt{35/26}$.

Exercise 427 exoINGE1121La0003

Soit E le sous-espace de $\mathbb{R}^3$ engendré par

$$\{(1,1,0), (1,2,-2), (-2,-3,2), (-3,-4,2)\}. \tag{93.93}$$

Donner une base orthonormée de E. Compléter le cas échéant cette base en une base orthonormée de $\mathbb{R}^3$. corrINGE1121La0003

Correction of the exercise 427

Ici, la tentation est d'appliquer la méthode de Gram-Schmidt. Cela n'est pas directement possible parce qu'avant d'appliquer Gram-Schmidt, il nous faut une partie libre. La matrice formée par les vecteurs donnés est de rang deux :

$$\begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -3 & -4 \\ 0 & -2 & 2 & 2 \end{pmatrix}. \tag{93.94}$$

En substituant les lignes $L_2 \to L_2 - L_1$ et $L_3 \to L_3/2$, nous trouvons que la deuxième et la troisième ligne deviennent les mêmes, de telle sorte qu'il reste

$$\begin{pmatrix} 1 & 1 & -2 & -3 \\ 0 & 1 & -1 & -1 \end{pmatrix}, \tag{93.95}$$

dont le rang est deux.

Nous savons donc maintenant que $\dim(E) = 2$. Prenons donc deux vecteurs linéairement indépendants dans E et appliquons leur Gram-Schmidt. Un choix possible est

$$\begin{aligned} v_1 &= (1, 1, 0) \\ v_2 &= (1, 2, -2) \end{aligned} \tag{93.96}$$

dont le Gram-Schmidt est

$$\begin{aligned} w_1 &= (1, 1, 0) \\ w_2 &= (-\frac{1}{2}, \frac{1}{2}, -2) \end{aligned} \tag{93.97}$$

Exercise 428 exoINGE1121La0004

En utilisant les propriétés de l'inverse et de la transposée d'un produit, démonter que

(1) si A et B sont orthogonales, alors AB est orthogonale,

(2) si A est orthogonale et si $B = AT$ avec T inversible, alors TB^{-1} est orthogonale.

corrINGE1121La0004

Correction of the exercise 428

Dans les deux cas, il faut utiliser le fait que $(AB)^{-1} = B^{-1}A^{-1}$ et $(AB)^t = B^tA^t$.

(1) Calculons le produit $(AB)^t(AB)$:

$$(AB)^t(AB) = B^t \underbrace{A^tA}_{1} B = B^tB = 1 \tag{93.98}$$

parce que $AA^t = BB^t = 1$ du fait que A et B soient orthogonales.

(2) La matrice TB^{-1} est $T(AT)^{-1} = TT^{-1}A^{-1} = A^{-1}$. Mais on sait que si A est orthogonale, alors A^{-1} est également orthogonale.

Exercise 429 exoINGE1121La0005

Soit $W = \{(x_1, x_2, x_3, x_4) \text{ tel que } x_1 + 2x_2 + x_3 + 3x_4 = 0\}$.

(1) Montrer que W est un sous-espace de $\mathbb{R}^4$. Déterminer sa dimension et une base.

(2) Construire le sous-espace vectoriel $W^\perp$ orthogonal à W. Déterminer sa dimension et une base.

corrINGE1121La0005

Correction of the exercise 429

Afin de vérifier que W est un espace vectoriel, il faut faire trois choses

(1) Prouver que $(0, 0, 0, 0) \in W$. La vérification est immédiate : $0 + 2 \cdot 0 + 0 + 3 \cdot 0 = 0$.

(2) Si x et y sont dans W, il faut que $x + y$ soient dans W. Testons donc l'équation de définition de W sur le vecteur

$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4). \tag{93.99}$$

Il faut vérifier que la combinaison suivante est nulle :

$$(x_1 + y_1) + 2(x_2 + y_2) + (x_3 + y_3) + 3(x_4 + y_4) \tag{93.100}$$

En regroupant les termes en x et en y, on obtient

$$\underbrace{x_1 + 2x_2 + x_3 + 3x_4}_{x\in W \Rightarrow =0} + \underbrace{y_1 + 2y_2 + y_3 + 3y_4}_{y\in W \Rightarrow =0}. \tag{93.101}$$

(3) Si $x \in W$, il faut que $\lambda x \in W$ pour tout $\lambda \in \mathbb{R}$. La démonstration est similaire à la précédente. Il suffit d'écrire la condition pour le vecteur λx et de mettre λ en évidence pour obtenir zéro.

Les points du plan sont les points $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ qui vérifient l'équation

$$x_1 + 2x_2 + x_3 + 3x_4 = 0. \tag{93.102}$$

La matrice de cette équation est la toute bête matrice

$$\begin{pmatrix} 1 & 2 & 1 & 3 \end{pmatrix}, \tag{93.103}$$

dont le rang est un. La dimension de l'espace de solutions d'un système de rang un dans $\mathbb{R}^4$ est 3. L'espace W est donc de dimension 3.

Supposons avoir trouvé une base orthogonale w_1, w_2, w_3 de W. Les vecteurs de $W^\perp$ perpendiculaires à w_1, w_2 et w_3. Ils sont donc linéairement indépendants de ceux de W (par la proposition 10.16). Si $w'_1, \cdots, w'_l$ sont une base orthonormale de $W^\perp$, nous avons alors $3 + l$ vecteurs linéairement indépendants dans $\mathbb{R}^4$. Nous en déduisons que $l = 1$. Donc $\dim(W^\perp) = 1$.

Trouvons une base de W. Par définition les vecteurs de W s'écrivent sous la forme

$$(-2x_2 - x_3 - 3x_4, x_2, x_3, x_4) \tag{93.104}$$

où x_2, x_3 et x_4 sont des paramètres. Pour trouver trois vecteurs de base de cet espace, posons successivement $x_2 = 1$, $x_3 = 1$ et $x_3 = 1$. Les vecteurs trouvés sont

$$\begin{aligned} v_1 &= (-2, 1, 0, 0) \\ v_2 &= (-1, 0, 1, 0) \\ v_3 &= (-3, 0, 0, 1). \end{aligned} \tag{93.105}$$

EqLesTroisvi

Pour vérification que ces trois vecteurs sont bien linéairement indépendants, nous pouvons les mettre dans une matrice

$$\begin{pmatrix} -2 & -1 & -3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{93.106}$$

et puis vérifier que le rang de cette matrice est bien 3.

Nous savons que $W^\perp$ est de dimension 1. En trouver une base revient à trouver un vecteur perpendiculaire aux trois vecteurs v_i donnés en (93.105). Nous pouvons bien entendu chercher ce vecteur sous la forme $x = (x_1, x_2, x_3, x_4)$ et résoudre le système d'équation donné par

$$\begin{cases} \langle x, v_1 \rangle = 0 & \text{(93.107a)} \\ \langle x, v_2 \rangle = 0 & \text{(93.107b)} \\ \langle x, v_3 \rangle = 0 & \text{(93.107c)} \end{cases}$$

Heureusement, il y a un « truc » pour trouver. Il suffit de prendre l'équation de W et de prendre comme vecteur x le vecteur dont les coordonnées sont les coefficients dans l'équation de W. Ici, l'équation de W est

$$x_1 + 2x_2 + x_3 + 3x_4 = 0. \tag{93.108}$$

Je prétends donc que le vecteur $(1, 2, 1, 3)$ est perpendiculaire aux trois. La vérification est aisée.

93.4 Valeurs propres et vecteurs propres

Exercise 430 exoLineraire0032

Exercice 1, page 88. Déterminer les valeurs propres et vecteurs propres des matrices suivantes.

(1) $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$

(2) $B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(3) $E = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 0 \\ 2 & 0 & 10 \end{pmatrix}$.

Sont-elles diagonalisables ?

corrLineraire0032

Correction of the exercise 430

(1) Nous avons

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \tag{93.109}$$

et donc

$$A - \lambda \mathbb{1} = \begin{pmatrix} 1-\lambda & 2 \\ 0 & 4-\lambda \end{pmatrix}, \tag{93.110}$$

dont le déterminant est $(1-\lambda)(3-\lambda)$. Les seules valeurs propres possibles sont donc

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 3. \end{aligned} \tag{93.111}$$

Il faut maintenant regarder, pour chacune de ces valeurs si il y a vraiment des vecteurs qui correspondent. Pour $\lambda_1 = 1$, nous devons résoudre

$$\begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{93.112}$$

La réponse est que tous les vecteurs de la forme $\begin{pmatrix} x \\ 0 \end{pmatrix}$ sont des vecteurs propres de la matrice A pour la valeur propre 1. Pour λ_2, nous faisons le calcul

$$(A - 3\mathbb{1}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{93.113}$$

dont les solutions sont données par $x = y$, c'est-à-dire par tous les vecteurs multiples de $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Vérifions, juste pour le fun que la matrice A multiplie bien le vecteur $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ par 3. En effet,

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2+4 \\ 0+6 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix}. \tag{93.114}$$

Bien joué !

(2) Passons à la matrice

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{93.115}$$

Le déterminant de $B - \lambda\mathbb{1}$ est donné par

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 1). \tag{93.116}$$

Donc les valeurs propres sont données par $\lambda = \pm 1$.
Commençons par $\lambda = 1$. Nous devons résoudre

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{93.117}$$

Donc $-x+y=0$ et $x-y=0$. La seule contrainte est $x=y$. Il y a deux vecteurs linéairement indépendants qui vérifient cette contrainte :

$$\begin{pmatrix}1\\1\\1\end{pmatrix}, \text{et} \quad \begin{pmatrix}1\\1\\0\end{pmatrix}. \tag{93.118}$$

Pour $\lambda=-1$, on doit résoudre

$$\begin{pmatrix}1&1&0\\1&1&0\\0&0&2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}. \tag{93.119}$$

La résolution donne $x+y=0$ et $z=0$. L'espace des vecteurs qui vérifient cette condition est engendré par le vecteur $\begin{pmatrix}1\\-1\\0\end{pmatrix}$.

Les matrices A et B sont diagonalisables parce qu'elles possèdent une base de vecteurs propres.

(3) Passons à la matrice

$$E=\begin{pmatrix}1&3&5\\0&1&0\\2&0&10\end{pmatrix}. \tag{93.120}$$

Nous avons

$$\det(E-\lambda\mathbb{1})=(1-\lambda)^2(10-\lambda)-10(1-\lambda)=-\lambda(\lambda-1)(\lambda-11). \tag{93.121}$$

Il faudra donc passer en revue les trois valeurs possibles $\lambda=0$, $\lambda=1$ et $\lambda=11$. Pour $\lambda=0$, nous avons à résoudre

$$\begin{pmatrix}1&3&5\\0&1&0\\2&0&10\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{93.122}$$

ce qui donne $x+5z=0$ et $y=0$, donc on a l'espace engendré par le vecteur $\begin{pmatrix}1\\0\\-1/5\end{pmatrix}$. Pour la valeur propre $\lambda=1$, on trouve le vecteur $\begin{pmatrix}-9/2\\-5/3\\1\end{pmatrix}$, et enfin pour $\lambda=11$, il faut résoudre le système

$$\begin{cases}x+3y+5z=11x\\y=11y\\2x+10z=11z\end{cases}, \tag{93.123}$$

ce qui donne immédiatement $y=0$ et puis $2x-z=0$. L'espace propre correspondant est engendré par le vecteur $\begin{pmatrix}1/2\\0\\1\end{pmatrix}$. Vérification :

$$\begin{pmatrix}1&3&5\\0&1&0\\2&0&10\end{pmatrix}\begin{pmatrix}1/2\\0\\1\end{pmatrix}=\begin{pmatrix}11/2\\0\\11\end{pmatrix}. \tag{93.124}$$

Exercise 431 exoLineraire0033

Exercice 5, page 89. Si λ_1, λ_2,..., λ_n sont les valeurs propres et v_1, v_2,..., v_n sont des vecteurs propres correspondants d'une matrice $n\times n$ A , quelles sont les valeurs propres et vecteurs propres des matrices

(1) αA pour $\alpha \in \mathbb{R}$,

(2) A^k pour $k \in \mathbb{N}_0$,

(3) A^{-1} si elle existe.

corrLineraire0033

Correction of the exercise 431

Si v est une vecteur propre de valeur propre λ, c'est que l'équation suivante est vérifiée :

$$Av = \lambda v. \tag{93.125}$$

(1) Calculons $(\alpha A)v = \alpha(Av) = \alpha\lambda v$, donc v est vecteur propre de αA pour la valeur $\alpha\lambda$.

(2) Nous avons
$$A^2v = AAv = A\lambda v = \lambda Av = \lambda\lambda v = \lambda^2 v. \tag{93.126}$$
Le vecteur v est donc vecteur propre de A^2, de valeur propre λ.

(3) Essayons de voir si nous avons une égalité $A^{-1}v = \mu v$ pour un certain μ. En appliquant A des deux côtés, nous avons $v = \mu Av = \mu\lambda v$, donc nous avons $\mu = \frac{1}{\lambda}$.

Noter que $\lambda = 0$ n'est pas possible parce qu'on a supposé que A est inversible.

Exercise 432 exoLineraire0034

Exercice 6, page 89. Si y est un vecteur propre de $B = R^{-1}AR$ correspondant à la valeur propre λ, démontrer que $x = Ry$ est un vecteur propre de A.

corrLineraire0034

Correction of the exercise 432

Nous devons calculer $A(Ry)$, en sachant que $R^{-1}ARy = \lambda y$. En appliquant R des deux côtés, nous avons

$$A(Ry) = \lambda(Ry), \tag{93.127}$$

ce qui prouve que le vecteur Ry est vecteur propre de A pour la valeur propre λ.

Exercise 433 exoLineraire0035

Exercice 9, page 90. On donne[1] la matrice

$$M = \begin{pmatrix} a & 0 & c \\ 2 & a & 3(b-1) \\ -3b & c & a \end{pmatrix}. \tag{93.128}$$

(1) Déterminer les valeurs c pour que $\det M = 0$ quand $a = 1$.

(2) Pour ces valeurs de c, déterminer les valeurs propres et les vecteurs propres de M. Dans quel(s) cas existe-t-il une base formée de vecteurs propres ?

corrLineraire0035

Correction of the exercise 433

En posant $a = 1$, le déterminant de la matrice est

$$\begin{vmatrix} 1 & 0 & c \\ 2 & 1 & 3(b-1) \\ -3b & c & 1 \end{vmatrix} = 2c^2 + 3c + 1, \tag{93.129}$$

qui s'annule lorsque $c = -1/2$ ou bien lorsque $c = -1$.

Lorsque $c = -1$, les valeurs propres sont données par l'équation

$$\begin{vmatrix} 1-\lambda & 0 & -1 \\ 2 & 1-\lambda & 3(b-1) \\ -3b & -1 & 1-\lambda \end{vmatrix} = -(\lambda-3)\lambda^2 = 0. \tag{93.130}$$

1. si si, on vous la donne, vous pouvez la garder.

Les valeurs propres sont $\lambda_1 = 0$ et $\lambda - 2 = 3$. Noter que valeur λ_1 est une racine double. Afin de trouver les vecteurs propres pour la valeur $\lambda_1 =$, nous résolvons maintenant l'équation $Mv = 0$:

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 2 & 1 & 3(b-1) & \\ -3b & -1 & 1 & \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 2-3b & 0 & 3(b-1)+1 & \\ 1-3b & -1 & 0 & 0 \end{array}\right). \tag{93.131}$$

Nous voyons que la deuxième ligne est $(2 - 3b)$ fois la première, donc elle saute. Il reste $z = x$ et $y = (1 - 3b)x$, et donc l'espace engendré par le vecteur

$$\begin{pmatrix} 1 \\ 1 \\ 1-3b \end{pmatrix}. \tag{93.132}$$

Nous passons maintenant à $\lambda_2 = 3$, c'est-à-dire à résoudre l'équation $Mv = 3v$. Sous forme matricielle nous avons

$$\left(\begin{array}{ccc|c} -2 & 0 & -1 & 0 \\ 2 & -2 & 3(b-1) & 0 \\ -3b & -1 & -2 & 0 \end{array}\right), \tag{93.133}$$

dont la solution fournit le vecteur

$$\begin{pmatrix} -1/2 \\ (3b-4)/2 \\ 1 \end{pmatrix}. \tag{93.134}$$

Le fait qu'il n'y ait que deux vecteurs propres linéairement indépendants fait qu'il n'y a pas de bases de vecteurs propres, et donc la matrice n'est pas diagonalisable.

Exercise 434 exoLineraire0036

Exercice 12, page 91. Les matrices complexes

$$A = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix} \tag{93.135}$$

sont-elles diagonalisables ? Si oui, donner pour chacune d'entre elles une base (du vectoriel complexe $\mathbb{C}^2$) dans laquelle elle est diagonale. corrLineraire0036

Correction of the exercise 434

Nous avons

$$\det \begin{pmatrix} 1-\lambda & 0 \\ 0 & i-\lambda \end{pmatrix} = (1-\lambda)(i-\lambda), \tag{93.136}$$

et donc les deux valeurs propres sont 1 et i. Pour $\lambda = 1$, on résou

$$\begin{pmatrix} 0 & 0 \\ 0 & i-1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{93.137}$$

donc x est libre et $(i - 1)y = 0$, c'est-à-dire que nous avons le vecteur

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{93.138}$$

Pour la valeur propre $\lambda = i$, par contre, nous résolvons

$$\begin{pmatrix} 1-i & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{93.139}$$

donc y est libre et $(1 - u)x = 0$, c'est-à-dire $x = 0$. Le vecteur propre associé est donc

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{93.140}$$

Il y a deux vecteurs linéairement indépendants, donc c'est une base et la matrice est diagonalisable.

Faisons maintenant la matrice

$$B = \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix}. \tag{93.141}$$

Ses valeurs propres sont encore $\lambda_1 = 1$ et $\lambda_2 = i$. Pour $\lambda_1 = 1$, il faut résoudre $Av = v$, c'est-à-dire

$$\begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}, \tag{93.142}$$

et la solution est $x + y = x$ et $iy = y$. Cela donne $y = 0$ et x libre, c'est-à-dire qu'une base de cet espace propre est $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Pour $\lambda_2 = i$, on doit résoudre $Av = iv$, c'est-à-dire $x + y = ix$ et $iy = it$. Pas de contraintes sur y, mais $x = y/(1 - i)$. Nous avons donc le vecteur propre

$$\begin{pmatrix} 1/(1-i) \\ 1 \end{pmatrix} = \begin{pmatrix} i+1 \\ 2 \end{pmatrix}. \tag{93.143}$$

Exercise 435 exoLineraire0037

Exercice 14, page 91. Soit

$$\begin{aligned} A_1 &= \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix} \\ A_2 &= \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}. \end{aligned} \tag{93.144}$$

Trouver une matrice orthogonale P qui transforme A en une matrice diagonale. corrLineraire0037

Correction of the exercise 435

Comme ce sont des matrices symétriques, il existe une base dans laquelle elles sont diagonales. Et de plus, cette base est elle-même orthonormale, donc le changement de base est donné par une matrice orthogonale.

Voyons les vecteurs propres de la matrice

$$A = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix}, \tag{93.145}$$

Nous avons $\det(1-\lambda\mathbb{1}) = -(\lambda-6)(\lambda-3)(\lambda-2)$. Nous devons résoudre successivement les systèmes $Av = 6v$, $Av = 3v$ et $Av = 2v$. Les vecteurs trouvés sont respectivement

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{93.146}$$

En calculant les produits scalaires, il est vite vu que ces trois vecteurs sont deux à deux orthogonaux. La matrice qui fait passer de la base canonique à la base de ces trois vecteurs s'obtient en mettant simplement ces trois vecteurs en colonnes :

$$\begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}. \tag{93.147}$$

Hélas, cette matrice n'est pas orthogonale parce que les trois vecteurs obtenus ne sont pas de norme 1. Il faut donc un peu les modifier. La norme du premier est $\sqrt{1^2 + 4 = 1} = \sqrt{6}$, donc en

divisant le vecteur par $\sqrt{6}$, nous avons un vecteur de norme 1. La norme du second est $\sqrt{3}$ et celle du troisième est $\sqrt{2}$, donc la matrice orthogonale recherchée est

$$\begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \end{pmatrix}. \tag{93.148}$$

Exercise 436 exoLineraire0038

Quelques questions d'interprétation géométrique des matrices.

(1) Écrire la matrice A de la symétrie par rapport au plan $x = -z$ de $\mathbb{R}^3$.
(2) Écrire la matrice B de la symétrie par rapport au plan $x = -y$ de $\mathbb{R}^3$.
(3) Faire le calcul de $C = BA$ et décrire géométriquement la transformation représentée par la matrice C.
(4) Donner les valeurs propres réelles et les vecteurs propres correspondants de C. corrLineraire0038

Correction of the exercise 436

(1) La symétrie par rapport au plan $x = -z$ doit laisser invariante les vecteurs qui sont dans le plan, c'est-à-dire les vecteurs tels que $x = -z$:

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{et} \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{93.149}$$

En même temps, la symétrie doit changer le signe du vecteur $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ qui est perpendiculaire au plan donné. Afin d'écrire la matrice de l'application, il faut trouver l'image des vecteurs de bases e_1, e_2 et e_3. Pour cela, il faut décomposer ces vecteurs de la base canonique dans la base donnée par

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \text{et} \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \tag{93.150}$$

Nous trouvons

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tag{93.151}$$

donc, en nous rappelant quel vecteur reste inchangé par A et quel doit changer de signe,

$$Ae_1 = A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2}A\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \tag{93.152}$$

Vu que nous avons dit que e_2 ne changeait pas, nous avons tout de suite

$$Ae_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{93.153}$$

Et enfin,

$$Ae_3 = A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2}A\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2}A\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \tag{93.154}$$

Nous pouvons maintenant écrire la matrice de la symétrie en écrivant simplement les vecteurs Ae_1, Ae_2 et Ae_3 en colonne :

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{93.155}$$

(2) La symétrie par rapport au plan $x = -y$ laisse inchangés les vecteurs

$$f_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{et} f_2 \quad = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \tag{93.156}$$

tandis qu'elle change le signe de

$$f_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \tag{93.157}$$

Pour écrire la matrice, il faut d'abord calculer Ae_1, Ae_2 et Ae_3. Pour ce faire, nous décomposons e_1, e_2 et e_3 dans la base $\{f_1, f_2, f_3\}$:

$$\begin{aligned} e_1 &= \frac{1}{2}(f_3 + f_3) \\ e_2 &= \frac{1}{2}(f_3 - f_2) \\ e_3 &= f_1, \end{aligned} \tag{93.158}$$

donc

$$\begin{aligned} Ae_1 &= \frac{1}{2}(f_2 - f_3) = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \\ Ae_2 &= \frac{1}{2}(-f_3 - f_2) = \begin{pmatrix} -1 \\ \\ 0 \end{pmatrix} \\ Ae_3 &= f_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{93.159}$$

La matrice de la symétrie se note alors

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{93.160}$$

(3) La matrice produit est

$$AB = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{93.161}$$

Tout vecteur perpendiculaire à un des deux plans doit être retourné par cette matrice, et tout vecteur dans les deux plans en même temps doit être laissé inchangé. Il s'agit donc de la symétrie par rapport à la droite intersection des deux plans. Calculer les vecteurs propres et valeurs propres pour s'en convaincre.

Exercise 437 exoLineraire0039

Exercice 16, page 92. On considère la matrice

$$M = \begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix} \tag{93.162}$$

avec $a, b \in \mathbb{R}$.

(1) Calculer les valeurs propres et les vecteurs propres de M

(2) Pour quelle(s) valeur(s) de a et de b la matrice M possède-t-elle deux directions propres orthogonales ?

corrLineraire0039

Correction of the exercise 437

<+Lineaire0039+>

Exercise 438 exoLineraire0040

Exercice 17, page 93. Dans $\mathbb{R}^2$, on considère les transformations suivantes :

(1) La symétrie par rapport à la droite $y = 2x$,

(2) la rotation d'angle $\alpha \neq 0$ autour de l'origine,

(3) la symétrie par rapport à l'origine.

Calculer les valeurs et les vecteurs propres de ces transformations et, si possible, donner pour chacune d'elles une base de $\mathbb{R}^2$ dans laquelle sa matrice est diagonale. corrLineraire0040

Correction of the exercise 438

(1) Le vecteur qui est sur cet axe ne bouge pas, tandis que celui qui est perpendiculaire change de signe. Ces deux vecteurs sont donc des vecteurs propres, de valeurs propres 1 et -1 respectivement.

(2) Dans une rotation, aucune direction n'est conservée, sauf cas exceptionnels comme $\alpha = 2\pi$ (où rien ne bouge) ou bien $\alpha = \pi$ qui renverse tout.

(3) Tous les vecteurs sont renversés. En particulier $Ae_1 = -e_1$ et $Ae_2 = -e_2$. Donc la matrice est diagonale dans la base canonique, avec -1 comme valeur propre double.

Exercise 439 exoLineraire0041

Exercice 18, page 93. Sans calculer, déterminer les valeurs et vecteurs propres des transformations suivantes de $\mathbb{R}^3$.

(1) La symétrie orthogonale par rapport au plan passant par les points $(0,0,0)$, $(0,1,1)$ et $(1,0,1)$.

(2) Le symétrie orthogonale par rapport à la droite d'équation $x = -\frac{y}{3} = -\frac{z}{2}$.

corrLineraire0041

Correction of the exercise 439

(1) Les deux vecteurs directeurs du plan sont conservés, et le vecteur perpendiculaire est retourné. Les conservés sont ceux qui sont vecteurs propres de valeur propre 1, et ce sont

$$\begin{pmatrix}0\\1\\1\end{pmatrix} \quad \text{et} \quad \begin{pmatrix}1\\0\\1\end{pmatrix}. \tag{93.163}$$

Pour trouver un vecteur perpendiculaire au plan, il faut un vecteur perpendiculaire aux deux vecteurs directeurs du plan. Pour ce faire, on calcule le produit vectoriel. C'est

$$\begin{pmatrix}1\\1\\-1\end{pmatrix}. \tag{93.164}$$

Ce dernier est un vecteur propre de valeur propre -1.

(2) Les vecteurs sur la droite sont conservés. À multiple près, il y en a un seul et c'est

$$\begin{pmatrix}1\\-3\\-2\end{pmatrix}. \tag{93.165}$$

Celui-là est de valeur propre 1. Tous ceux qui lui sont perpendiculaires (le plan perpendiculaire à cette droite) sont de valeur propre -1.

Exercise 440 exoLineraire0042

Exercice 19, page 93. On considère la transformation linéaire A de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$\begin{pmatrix} 1 & 1 & -2 \\ 2 & 1 & -2 \\ 3 & 1 & -4 \end{pmatrix}. \tag{93.166}$$

(1) Chercher les valeurs propres et les vecteurs propres de A.

(2) Cette matrice est-elle diagonalisable ? Si oui, donner la matrice d'un changement de base qui permet de la diagonaliser.

(3) On donne le vecteur $v_1 = (1, 1, 2)$. Calculer $v_2 = Av_1$ et $v_3 = A^2v_1$.

(4) Montrer que les vecteurs v_1, v_2 et v_3 forment une base de $\mathbb{R}^3$.

(5) Donner les composantes du vecteur propre de A de norme $2\sqrt{6}$ correspondant à la plus grande valeur propre de A dans la base $\{v_1, v_2, v_3\}$.

corrLineraire0042

Correction of the exercise 440

Je n'ai pas corrigé cet exercice. Mais si tu as une correction à proposer, n'hésite surtout pas à me l'envoyer.

Exercise 441 exoINGE1121La0020

Trouver les valeurs propres et les vecteurs propres de la matrice

$$\begin{pmatrix} 2 & 0 & \alpha \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{93.167}$$

où α est un paramètre réel. Discuter selon les valeurs de α.

corrINGE1121La0020

Correction of the exercise 441

Pour les valeurs propres, nous calculons le polynôme caractéristique :

$$P_A(\lambda) = \det \begin{pmatrix} 2-\lambda & 0 & \alpha \\ 0 & -1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = -(2-\lambda)^2(1+\lambda). \tag{93.168}$$

Cela fournit les valeurs propres $\lambda_1 = 2$ (de multiplicité 2) et $\lambda_2 = -1$.

Commençons par trouver les vecteurs propres correspondants à la valeur propre $\lambda_2 = -1 =$. Pourquoi nous commençons par celle-là ? Parce qu'il n'y a pas de discussion sur α pour lui. La matrice du système est

$$\begin{pmatrix} 3 & 0 & \alpha \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}. \tag{93.169}$$

La troisième ligne dit tout de suite que $z = 0$m et donc la première ligne donne $x = 0$ (sans discussions). Ensuite, y est sans contraintes. L'espace propre de la valeur propre -1 est donc toujours engendré par

$$v_{-1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{93.170}$$

En ce qui concerne l'espace propre pour la valeur $\lambda_1 = 2$, nous avons le système de matrice

$$\begin{pmatrix} 0 & 0 & \alpha \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{93.171}$$

La seconde ligne dit que $y = 0$. La première dit que $\alpha z = 0$. Il y a donc deux possibilités :

(1) $\alpha \neq 0$. Alors $z = 0$ est obligatoire et l'espace propre est engendré par

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{93.172}$$

Dans ce cas, même si la valeur propre était de multiplicité deux, il n'y a que une seule dimension de vecteurs propres.

(2) Si $\alpha = 0$, alors on ne va pas discuter longtemps : la matrice est déjà diagonale depuis le début. On lit tout de suite dessus que les vecteurs propres sont les trois vecteurs de la base canonique.

Exercise 442 exoexoMatrices-0001

Si la matrice A est d'ordre 3 et de rang 2, que pouvons nous dire des valeurs propres de A et de leur multiplicité.

correxoMatrices-0001

Correction of the exercise 442

D'abord le fait que le rang soit plus petit que l'ordre implique que le déterminant est nul, et que nous avons une valeur propre nulle. La multiplicité de cette valeur propre ne peut être que 1. Si elle était de multiplicité deux, le rang n'aurait été que de un.

En ce qui concerne les autres valeurs propres, on ne peut rien dire. En effet n'importe quelle matrice de la forme

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \tag{93.173}$$

avec a et b non nuls est dans les hypothèses de l'exercice.

93.5 Triangularisation

Exercise 443 exoINGE1121La0021

Pour chacune des matrices A suivantes, expliquer si il existe ou non une matrice inversible B telle que la matrice $B^{-1}AB$ soit diagonale.

Si une telle matrice existe, pouvez-vous la choisir orthogonale ?

(1) $A = \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}$,

(2) $B = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$,

(3) $C = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$,

(4) $D = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$.

corrINGE1121La0021

Correction of the exercise 443

(1) La matrice A est symétrique. Il existe donc une matrice orthogonale qui la diagonalise.

(2)

(3)

(4) Le polynôme caractéristique est donné par

$$\det \begin{pmatrix} 1-\lambda & 2 \\ 1 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) - 2. \tag{93.174}$$

Les solutions sont $\lambda_1 = 0$ et $\lambda_2 = 3$.

Pour trouver le vecteur propre correspondant à la valeur zéro, on résous le système homogène de la matrice

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}. \tag{93.175}$$

La réponse est donnée par le vecteur $v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ et tous ses multiples.

Pour trouver les vecteurs de la valeur 3, nous résolvons le système de

$$\begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}, \tag{93.176}$$

et la réponse est $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Les deux vecteurs v_1 et v_2 n'étant pas orthogonaux, il n'est pas possible de trouver une matrice orthogonale qui diagonalise la matrice de départ.

Exercise 444 exoINGE1121La0011

Déterminer les valeurs propres des matrices suivantes :

(1) $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$,

(2) $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,

(3) $R = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$,

(4) $D = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Pour chacune des valeurs propres, rechercher l'ensemble des vecteurs propres correspondants et donner la dimension de l'espace vectoriel engendré par ces vecteurs propres. corrINGE1121La0011

Correction of the exercise 444

<+INGE1121La0011+>

93.6 Formes quadratiques

Exercise 445 exoINGE1121La0012

(INGE1121, 6.1) Réduire les formes quadratiques suivantes en somme pondérée de carrés par la méthode des valeurs propres et des vecteurs propres. Donner le genre des formes quadratiques.

(1) $p(X) = 2x_1x_3 + 2x_3x_4$.

(2) $p(X) = 2x_1^2 - x_2^2 + 2x_3^2 + 4x_1x_2 + 8x_1x_3 + 4x_2x_3$

(3) $p(X) = x_1x_2 + x_1x_3 + x_2x_3$

(4) $p(X) = 2x_1^2 - 2x_1x_2 + 2x_2^2 + 2x_3^2$

(5) $p(X) = x_1^2 + x_1x_2 + 2x_2^2$

corrINGE1121La0012

Correction of the exercise 445

(1) Il s'agit d'abord de remplir la matrice A correspondante à la forme quadratique. Le terme en x_1x_3 correspond aux cases $(1,3)$ et $(3,1)$ de la matrice. Comme le coefficient est 2, nous

mettons 1 dans chaque case. De la même manière, le terme $2x_3x_4$ fait venir un 1 dans les cases $(3,4)$ et $(4,3)$. La matrice est donc

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{93.177}$$

Cette façon de construire la matrice fournit toujours une matrice symétrique. Cette dernière aura donc une base de vecteurs propres. Si nous avions mit simplement 2 dans les cases $(1,3)$ et $(2,4)$, nous n'aurions pas eut la symétrie. Cela est l'intérêt de mettre la moitié dans chacune des deux cases.

Nous pouvons maintenant nous lancer dans la recherche des valeurs propres de cette matrice :

$$\begin{aligned} \det(A - \lambda\mathbb{1}) &= \det \begin{pmatrix} -\lambda & 0 & 1 & 0 \\ 0 & -\lambda & 0 & 0 \\ 1 & 0 & -\lambda & 1 \\ 0 & 0 & 1 & -\lambda \end{pmatrix} \\ &= -\lambda \det \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix} \\ &= -\lambda^2 \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} - \lambda \det \begin{pmatrix} 1 & 0 \\ 1 & -\lambda \end{pmatrix} \\ &= -\lambda^2(\lambda^2 - 2). \end{aligned} \tag{93.178}$$

Les valeurs propres sont donc 0 (de multiplicité deux) et $\pm\sqrt{2}$. Nous savons donc qu'il existe des nouvelles variables y_i telles que la forme quadratique s'écrive

$$\sqrt{2}(y_1^2 - y_2^2). \tag{93.179}$$

exoLa0012quddiafy

Afin d'exprimer ces nouvelles coordonnées en fonction des x_i, nous devons trouver les vecteurs propres de la matrice.

Pour la valeur propre $\sqrt{2}$, nous devons résoudre le système donné par la matrice

$$\begin{pmatrix} -\sqrt{2} & 0 & 1 & 0 \\ 0 & -\sqrt{2} & 0 & 0 \\ 1 & 0 & -\sqrt{2} & 1 \\ 0 & 0 & 1 & -\sqrt{2} \end{pmatrix} \tag{93.180}$$

La deuxième ligne dit tout de suite que $x_2 = 0$. En prenant $x_4 = \mu$ comme paramètre, nous trouvons ensuite

$$x_3 = \sqrt{2}\mu, \tag{93.181}$$

et $-\sqrt{2}x_1 + x_3 = 0$, ce qui donne $\sqrt{2}(\mu - x_1) = 0$ et donc

$$x_1 = \mu. \tag{93.182}$$

Une base de l'espace propre pour la valeur propre $\sqrt{2}$ est alors donné par

$$\begin{pmatrix} 1 \\ 0 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{93.183}$$

De la même manière, nous trouvons les espaces propres pour les valeurs $-\sqrt{2}$ et 0. Les réponses sont

$$\begin{aligned}\sqrt{2} \rightsquigarrow \begin{pmatrix} 1 \\ 0 \\ \sqrt{2} \\ 1 \end{pmatrix}, \qquad -\sqrt{2} \rightsquigarrow \begin{pmatrix} 1 \\ 0 \\ -\sqrt{2} \\ 1 \end{pmatrix} \\ 0 \rightsquigarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.\end{aligned} \tag{93.184}$$

Notez que tous ces vecteurs sont deux à deux orthogonaux (était-ce prévisible ?).

Une erreur à ne pas faire est de croire que la matrice orthogonale qui diagonalise A est la matrice obtenue en mettant ces quatre vecteurs en colonnes. En effet, ces vecteurs ne sont pas normés. Pour obtenir la bonne matrice, il faut les d'abord diviser chacun de ces vecteurs par leur norme. Ce que nous obtenons est

$$B = \begin{pmatrix} 1/2 & 1/2 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 & 0 \\ 1/2 & 1/2 & -1/\sqrt{2} & 0 \end{pmatrix}. \tag{93.185}$$

La forme quadratique a la forme (93.179) dans les variables Y données par $X = BY$, c'est-à-dire

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 & 0 \\ 1/2 & 1/2 & -1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}. \tag{93.186}$$

Tu paries que si on substitue les valeurs

$$\begin{aligned} x_1 &= \frac{1}{2}(y_1 + y_2) + \frac{y_3}{\sqrt{2}} \\ x_2 &= y_4 \\ x_3 &= \frac{1}{\sqrt{2}}(y_2 - y_1) \\ x_4 &= \frac{1}{2}(y_1 + y_2) - \frac{y_3}{\sqrt{2}} \end{aligned} \tag{93.187}$$

dans l'expression $2x_1x_3 + 2x_3x_4$, on obtient bien $\sqrt{2}(y_2^2 - y_1^2)$?

En ce qui concerne le *genre* de la forme quadratique, il est indéterminé parce qu'il y a une valeur propre strictement négative et une strictement positive.

(2) La matrice est

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}. \tag{93.188}$$

Ses valeurs et vecteurs propres sont

$$\begin{aligned} 1 &\to v_1 = (1, 1, 0) \\ -1/2 &\to v_2 = (1, 0, 1) \\ -1/2 &\to v_3 = (1, -1, -1). \end{aligned} \tag{93.189}$$

La matrice suivante est formée des vecteurs d'une base orthonormale de vecteurs propres :

$$B = \begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ \frac{1}{3}\sqrt{\frac{3}{2}} & -\frac{1}{3}\sqrt{\frac{3}{2}} & \frac{2}{3}\sqrt{\frac{3}{2}} \\ \frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} & -\frac{1}{3}\sqrt{3} \end{pmatrix} \tag{93.190}$$

Étant donné les valeurs propres, la matrice n'a aucun genre particulier.

Exercise 446 exoINGE1121La0013

(INGE1121, 6.2)Donner le genre de la forme quadratique associée à la matrice

$$A = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}. \tag{93.191}$$

corrINGE1121La0013

Correction of the exercise 446

Nous pourrions évidement calculer les valeurs propres de la matrice et en regarder les signes. Hélas, la matrice étant de taille 3, le polynôme caractéristique sera de degré trois, et il n'est pas certain que nous serons capables de le résoudre.

Heureusement, en ce qui concerne la recherche du genre d'une application quadratique, les mineurs principaux font autant l'affaire que les valeurs propres. Ici les mineurs principaux sont

$$m_1 = \det \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}, \quad m_2 = \det \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}, \quad m_3 = \det(3). \tag{93.192}$$

Les calculs vite faits bien faits donnent $m_1 = 20$, $m_2 = 10$ et $m_3 = 3$, de telle sorte qu'ils sont tous strictement positifs. Le genre est donc *défini positif*.

Exercise 447 exoINGE1121La0017

(INGE 1121, 6.3) Soit p la forme quadratique associée à la matrice

$$A = \begin{pmatrix} 2 & 0 & -2 & 0 \\ 0 & 5 & 0 & -1 \\ -2 & 0 & 2 & 0 \\ 0 & -1 & 0 & 5 \end{pmatrix}. \tag{93.193}$$

(1) Calculer le range de A.
(2) Sachant que $(1, 1, -1, 1)$ est vecteur propre de A, diagonaliser A au moyen d'une matrice orthogonale.
(3) Écrire la forme quadratique X^tAX comme somme pondérée de carrés en fonction des x_i.
(4) Déterminer le genre de la forme quadratique $X \mapsto X^tAX$.
(5) Écrire p comme somme pondérée de carrés en fonction de x_1, x_2, x_3 et x_4.

corrINGE1121La0017

Correction of the exercise 447

Pour trouver le rang, faisons un tout petit peu d'échelonnement :

$$\begin{aligned} L_3 &\to L_2 + L_1 \\ L_4 &\to 5L_4 + L_2 \end{aligned} \tag{93.194}$$

donne

$$\begin{pmatrix} 2 & 0 & -2 & 0 \\ 0 & 5 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 24 \end{pmatrix}. \tag{93.195}$$

Après avoir barré la troisième ligne, nous trouvons facilement une sous-matrice de taille 3×3 qui a un déterminant non nul. Par conséquent, la rang de la matrice A est 3.

Le polynôme caractéristique est de degré 4, mais nous savons que 0 est une valeur propre, donc nous n'aurons en réalité que du degré 3 à résoudre. Mais comme nous savons un vecteur propre (donné dans l'énoncé), nous allons savoir une seconde valeur propre et le degré du polynôme caractéristique sera seulement deux.

Pour savoir la valeur propre du vecteur donné, il suffit de lui appliquer la matrice, on trouve

$$\begin{pmatrix} 2 & 0 & -2 & 0 \\ 0 & 5 & 0 & -1 \\ -2 & 0 & 2 & 0 \\ 0 & -1 & 0 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ -4 \\ 4 \end{pmatrix}, \tag{93.196}$$

et donc 4 est une valeur propre de vecteur propre $(1, 1, -1, 1)$.

Après quelques laborieux calculs, le polynôme caractéristique se factorise en

$$P_A(\lambda) = (\lambda - 4)(\lambda - 6)\lambda(\lambda - 4), \tag{93.197}$$

donc les valeurs propres sont $\{4, 6, 0\}$ et 4 est double.

Pour les vecteurs propres, nous résolvons les systèmes d'équations correspondants et les réponses sont

$$\begin{aligned} 6 &\to (0, 1, 0, -1) \\ 0 &\to (1, 0, 1, 0) \\ 4 &\to (1, 0, -1, 0) \\ 4 &\to (0, 1, 0, 1) \end{aligned} \tag{93.198}$$

EqVectosDSProps

Notez que l'espace propre pour la valeur 4 étant de dimension 2, il y a de nombreuses autres possibilités : toute les combinaisons linéaires des deux vecteurs proposés ici [2] sont corrects. Pour les autres valeurs, seuls les multiples sont corrects.

La méthode de Gram-Schmidt (ne pas oublier de normaliser après) appliquée aux vecteurs (93.198) fournit la base orthonormée

$$\begin{aligned} (0, \sqrt{2}22, 0, -\sqrt{2}/2) \\ (\sqrt{2}/2, 0, \sqrt{2}/2, 0) \\ (\sqrt{2}/2, 0, -\sqrt{2}/2, 0) \\ (0, \sqrt{2}/2, 0, \sqrt{2}/2) \end{aligned} \tag{93.199}$$

La matrice B qui diagonalise A est la matrice obtenue en mettant ces quatre vecteurs en colonne, et nous avons

$$B = \begin{pmatrix} 0 & \sqrt{2}/2 & \sqrt{2}/2 & 0 \\ \sqrt{2}/2 & 0 & 0 & \sqrt{2}/2 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ -\sqrt{2}/2 & 0 & 0 & \sqrt{2}/2 \end{pmatrix} \tag{93.200}$$

et puis

$$B^T AB = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}. \tag{93.201}$$

Les coordonnées qui font que la forme quadratiques est belle sont données par $X = BY$, et donc par

$$\begin{aligned} x1 &= \sqrt{2}y_2/2 + \sqrt{2}y_3/2 \\ x2 &= \sqrt{2}y_1/2 + \sqrt{2}y_4/2 \\ x3 &= \sqrt{2}y_2/2 - \sqrt{2}y_3/2 \\ x4 &= -\sqrt{2}y_1 1/2 + \sqrt{2}y_4/2 \end{aligned} \tag{93.202}$$

Dans ces coordonnées, la forme quadratique devient simplement une somme des carrés des y_i avec les valeurs propres comme coefficients :

$$p(X) = 6y_1^2 + 4y_3^2 + 4y_4^2. \tag{93.203}$$

2. Ce sont ceux que me donne mon ordinateur. J'insiste : soyez capable de résoudre les exercices à l'ordinateur ; vous n'allez pas trainer toute votre vie à faire des calculs à la main.

Notez que cela ne dépend pas de y_2 ; cela correspond au fait qu'une des valeurs propres est nulle.

Le genre de la forme quadratique est semi-définie positive parce que toutes ses valeurs propres sont positives ou nulles.

Exercise 448 exoINGE1121La0014

(INGE1121, 6.4) Soit la forme quadratique

$$p(X) = 2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3 + x_2x_4 - "x_3x_4. \tag{93.204}$$

(1) Calculer le rang et le déterminant de la matrice symétrique associée A :

(2) Qu'en déduire concernant les valeurs propres de A ?

(3) -1 est une valeur propre de A. Calculer sa multiplicité dans le polynôme caractéristique de A.

(4) Construire une matrice orthogonale B telle que B^tAB soit diagonale.

corrINGE1121La0014

Correction of the exercise 448

La matrice correspondante est

$$A = \begin{pmatrix} 0 & 1 & -1/2 & 1/2 \\ 1 & 0 & -1/2 & 1/2 \\ -1/2 & -1/2 & 0 & -1 \\ 1/2 & 1/2 & -1 & 0 \end{pmatrix}. \tag{93.205}$$

Afin de trouver le rang de cette matrice, nous échelonnons un petit peu. Il y a plusieurs façons de procéder. On peut par exemple faire $L_2 \to L_2 - 2L_4$ et $L_3 \to L_3 + L_4$:

$$A = \begin{pmatrix} 0 & 1 & -1/2 & 1/2 \\ 0 & -1 & 3/2 & 1/2 \\ 0 & 0 & -1 & -1 \\ 1/2 & 1/2 & -1 & 0 \end{pmatrix}, \tag{93.206}$$

et ensuite $L_2 \to L_2 + L_1$:

$$A = \begin{pmatrix} 0 & 1 & -1/2 & 1/2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \\ 1/2 & 1/2 & -1 & 0 \end{pmatrix}. \tag{93.207}$$

À partir de là, nous voyons que la deuxième et la troisième ligne sont égales au signe près, donc le déterminant de la matrice est nul et le rang ne peut pas être 4. Pour prouver que le rang est 3, il suffit de prendre à peu près n'importe quelle sous-matrice de taille 3 et de calculer le déterminant. Par exemple la matrice obtenue en enlevant la troisième ligne et la troisième colonne :

$$\det \begin{pmatrix} 0 & 1 & 1/2 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix} = 1/2 \neq 0. \tag{93.208}$$

Le rang de la matrice est donc 3.

Le fait que le déterminant soit nul montre qu'il y aura une valeur propre nulle. Mais le fait que le rang soit 3 dit également que la valeur propre nulle sera seulement de multiplicité un. Il y aura donc 3 valeurs propres non nulles.

Pour les valeurs propres, il faut calculer et résoudre le polynôme caractéristique

$$P_A(\lambda) = \det \begin{pmatrix} -\lambda & 1 & -1/2 & 1/2 \\ 1 & -\lambda & -1/2 & 1/2 \\ -1/2 & -1/2 & -\lambda & -1 \\ 1/2 & 1/2 & -1 & -\lambda \end{pmatrix}. \tag{93.209}$$

Les solutions avec les multiplicités sont $\{-1,-1,0,2\}$, et les vecteurs propres correspondants sont

$$\begin{aligned} 2 &\to (1,1,-1,1)\\ 0 &\to (1,1,1,-1)\\ -1 &\to (1,-1,0,0)\\ -1 &\to (0,0,1,1) \end{aligned} \tag{93.210}$$

La forme quadratique n'a pas de genre particulier. Notre ami Gram-Schmidt, suivit d'une renormalisation,nous fournit la base orthonormale suivante :

$$\begin{aligned} &(1/2,1/2,-1/2,1/2)\\ &(1/2,1/2,1/2,-1/2)\\ &(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0,0)\\ &(0,0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}) \end{aligned} \tag{93.211}$$

En mettant ces vecteurs en colonne, on obtient la matrice de changement de variables

$$\begin{aligned} x_1 &= \frac{\sqrt{2}}{2}y_3+\frac{1}{2}y_1+\frac{1}{2}y_2\\ x_2 &= -\frac{\sqrt{2}}{2}y_3+\frac{1}{2}y_1+\frac{1}{2}y_2\\ x_3 &= \frac{1}{\sqrt{2}}y_4-\frac{1}{2}y_1+\frac{1}{2}y_2\\ x_4 &= \frac{\sqrt{2}}{2}y_4+\frac{1}{2}y_1-\frac{1}{2}y_2 \end{aligned} \tag{93.212}$$

En mettant ces valeurs dans $p(X)$, nous trouvons bien

$$p\big(X(Y)\big)=2y_1^2-y_3^2-y_4^2, \tag{93.213}$$

qui est bien la forme dont les valeurs propres de A sont les coefficients.

Cette forme n'est pas d'un genre particulier [3].

Exercise 449 exoINGE1121La0018

(INGE1121, 6.5) Sachant que $(-1,0,1,0)$ est un vecteur propre de la matrice

$$A=\begin{pmatrix} 2 & 1 & -1 & 1\\ 1 & 0 & 1 & 1\\ -1 & 1 & 2 & 1\\ 1 & 1 & 1 & 0 \end{pmatrix} \tag{93.214}$$

(1) Diagonaliser A au moyen d'une matrice orthogonale

(2) Écrire la forme quadratique X^tAX sous forme d'une somme pondérée de carrés.

corrINGE1121La0018

Correction of the exercise 449

Calculons Av afin de savoir la valeur propre associée au vecteur donné :

$$\begin{pmatrix} 2 & 1 & -1 & 1\\ 1 & 0 & 1 & 1\\ -1 & 1 & 2 & 1\\ 1 & 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1\\ 0\\ 1\\ 0 \end{pmatrix}=\begin{pmatrix} -3\\ 0\\ 3\\ 0 \end{pmatrix}. \tag{93.215}$$

La valeur propre est donc 3. Nous savons donc que $(\lambda-3)$ pourra être factorisé dans le polynôme caractéristique.

Pour le reste de l'exercice c'est standard et c'est résolu de la façon suivante :

3. À part d'être le genre qui peut arriver à l'examen.

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_6_5():
   A=matrix(QQ,4,4,[2,1,-1,1,1,0,1,1,-1,1,2,1,1,1,1,0])
   x=outilsINGE.QuadraticForm(A)
   print x
```

qui retourne

```
Hi guy; I'm the quadratic form associated with the matix
[ 2  1 -1  1]
[ 1  0  1  1]
[-1  1  2  1]
[ 1  1  1  0]
My eigenvalues and eigenvectors are :
3 -> (1, 0, -1, 0)
3 -> (0, 1, 2, 1)
-1 -> (1, 0, 1, -2)
-1 -> (0, 1, 0, -1)
I've the following orthonormal basis of eigenvectors :
(1/2*sqrt(2), 0, -1/2*sqrt(2), 0)
(1/2, 1/2, 1/2, 1/2)
(1/6*sqrt(6), 0, 1/6*sqrt(6), -1/3*sqrt(6))
(-1/2*sqrt(1/3), 3/2*sqrt(1/3), -1/2*sqrt(1/3), -1/2*sqrt(1/3))
A matrix B such that B^tAB is diagonal is
[   1/2*sqrt(2)            1/2    1/6*sqrt(6) -1/2*sqrt(1/3)]
[             0            1/2             0  3/2*sqrt(1/3)]
[  -1/2*sqrt(2)            1/2    1/6*sqrt(6) -1/2*sqrt(1/3)]
[             0            1/2   -1/3*sqrt(6) -1/2*sqrt(1/3)]
I'm quite pretty in the following variables ...
x1 = -1/2*sqrt(1/3)*y4 + 1/2*sqrt(2)*y1 + 1/6*sqrt(6)*y3 + 1/2*y2
x2 = 3/2*sqrt(1/3)*y4 + 1/2*y2
x3 = -1/2*sqrt(1/3)*y4 - 1/2*sqrt(2)*y1 + 1/6*sqrt(6)*y3 + 1/2*y2
x4 = -1/2*sqrt(1/3)*y4 - 1/3*sqrt(6)*y3 + 1/2*y2
Look at me when I wear my cool variables
3*y1^2 + 3*y2^2 - y3^2 - y4^2
```

Un point délicat peut être la résolution du polynôme caractéristique. Nous trouvons

$$(1+\lambda)\lambda^3 - 5\lambda^2 + 3\lambda + 9 = 0, \tag{93.216}$$ EqDHpolycar

mais nous savons que $\lambda = 3$ est une solution, donc nous pouvons diviser (Euclide) le membre de gauche de (93.216) par $\lambda - 3$ et trouver

$$(1+\lambda)(\lambda-3)(\lambda^2 - 2\lambda - 3) = 0 \tag{93.217}$$

dont les solutions sont 3 et -1, chacune étant de multiplicité deux.

Conseil technique : on peut utiliser Gram-Schmidt séparément sur les deux vecteurs de valeur propre -1 et sur les deux de valeur propre 3. En effet, ces deux « paquets » de vecteurs propres sont orthogonaux.

Exercise 450 exoINGE1121La0015

(INGE1121, 6.6)Donner le genre de la forme quadratique

$$p(X) = 9x_1^2 + 17x_2^2 + 10x_3^2 - 6x_1x_2 + 12\sqrt{2}x_1x_3 + 4\sqrt{2}x_2x_3. \tag{93.218}$$

corrINGE1121La0015

Correction of the exercise 450

Une façon économique de procéder est de calculer les mineurs principaux de la matrice associée. Nous avons

$$m_1 = \det \begin{pmatrix} 9 & -3 & 6\sqrt{2} \\ -3 & 17 & 2\sqrt{2} \\ 6\sqrt{2} & 2\sqrt{2} & 10 \end{pmatrix} = \det \begin{pmatrix} 9 & -3 & 6\sqrt{2} \\ 0 & 48 & 12\sqrt{2} \\ 0 & 2 & \sqrt{2}/2 \end{pmatrix} = 0, \tag{93.219}$$

puis

$$m_2 = \det \begin{pmatrix} 9 & -3 \\ -3 & 17 \end{pmatrix} = 9 \cdot 17 - 9 = 144, \tag{93.220}$$

et enfin

$$m_3 = \det(9) = 9. \tag{93.221}$$

Deux strictement positifs et un nul, cela fait *semi définie positive*.

Exercise 451 exoINGE1121La0019

(INGE1121, 6.7) Déterminer le genre des formes quadratiques associées aux matrices suivantes en utilisant au maximum les propositions relatives aux mineurs, mineurs principaux primaires, ...

(1)
$$A = \begin{pmatrix} -6 & 1 & -2 \\ 1 & -3 & 1 \\ -2 & 1 & -1 \end{pmatrix} \tag{93.222}$$

(2)
$$A = \begin{pmatrix} 5 & 2 & 10 \\ 2 & -7 & 3 \\ 10 & 3 & 0 \end{pmatrix}. \tag{93.223}$$

corrINGE1121La0019

Correction of the exercise 451

(1) Nous y allons avec la méthode des mineurs principaux primaires : $\det(A) = -3$,

$$\det \begin{pmatrix} -6 & 1 \\ 1 & -3 \end{pmatrix} = 17 \tag{93.224}$$

et $\det(-6) = 6$.

Nous sommes dans le cas où les mineurs principaux primaires d'ordre impairs sont plus petits que zéros tandis que ceux d'ordre pair sont plus grands que zéro. La forme quadratique correspondante est dont définie négative.

(2) Dans ce cas, les mineurs principaux primaires sont 775, -39 et 5. La forme quadratique est de genre indéfinie.

93.7 Interrogation de mars 2010

Exercise 452 exoMars20100001

Donner un exemple de matrice de rang 2 dont le déterminant est nul.

corrMars20100001

Correction of the exercise 452

Si on veut un *rang* égal à deux, il faut que la matrice *contienne* une sous-matrice 2×2 de déterminant non nul (le plus simple est la matrice identité). Le plus simple pour obtenir un *déterminant* nul est d'avoir une ligne de zéros.

On peut donc considérer cette matrice :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{93.225}$$

Il y a énormément d'autres exemples possibles. En tout cas, il faut une matrice carré (pour que le concept de déterminant ait un sens) de taille au moins égale à trois.

Exercise 453 exoMars20100002

Résoudre le système suivant, en discutant selon les valeurs du paramètre $k \in \mathbb{R}$.

$$\begin{cases} 2x - y + 4z = 1 & \text{(93.226a)} \\ 2y - 6z = 8 & \text{(93.226b)} \\ -4x + 4y + (k-11)z = 6 & \text{(93.226c)} \end{cases}$$

corrMars20100002

Correction of the exercise 453

La matrice « augmentée » du système est

$$\left(\begin{array}{ccc|c} 2 & -1 & 4 & 1 \\ 0 & 2 & -6 & 8 \\ -4 & 4 & (k-11) & 6 \end{array}\right). \tag{93.227}$$

Diviser la seconde ligne par deux est toujours une bonne idée pour y voir plus clair. Ensuite on commence à échelonner en faisant $L_3 \to L_3 + 2L_1$:

$$\left(\begin{array}{ccc|c} 2 & -1 & 4 & 1 \\ 0 & 1 & -3 & 4 \\ 0 & 2 & k-3 & 8 \end{array}\right). \tag{93.228}$$

Ensuite on fait $L_3 \to L_3 - 2L_2$ pour mettre un dernier bon zéro où il faut :

$$\left(\begin{array}{ccc|c} 2 & -1 & 4 & 1 \\ 0 & 1 & -3 & 4 \\ 0 & 0 & k+3 & 0 \end{array}\right). \tag{93.229}$$

Sur cette matrice, on voit tout de suite que la discussion portera sur $k = -3$ ou $k \neq -3$. D'ailleurs, en calculant le déterminant de la matrice initiale, on aurait vu qu'il était un multiple de $k + 3$, donc le déterminant est nul si et seulement si $k = -3$.

Commençons par $k = -3$.

Dans ce cas, la dernière ligne devient $0 = 0$ et on peut la barrer. Il reste deux équations pour trois inconnues. Posons $z = \lambda$, c'est-à-dire que nous prenons z comme paramètre. Alors la seconde ligne devient $y - 3\lambda = 4$, c'est-à-dire $y = 4 + 3\lambda$. La première ligne devient alors $2x - 4 - 3\lambda + 4\lambda = 1$, c'est-à-dire $x = (5 - \lambda)/2$. Les solutions sont donc

$$\begin{pmatrix} \frac{5-\lambda}{2} \\ 4 + 3\lambda \\ \lambda \end{pmatrix}. \tag{93.230}$$

Pour chaque $\lambda \in \mathbb{R}$, il y a une solution. Il y a donc une infinité de solutions qui correspondent au cas de déterminant nul.

Prenons maintenant $k \neq -3$.

Dans ce cas, la dernière ligne devient $(k+3)z = 0$, c'est-à-dire $z = 0$. Les autres lignes donnent alors $y = 4$ et $x = 5/2$. L'unique solution correspondant aux cas de déterminant non nul est

$$\begin{pmatrix} 5/2 \\ 4 \\ 0 \end{pmatrix}. \tag{93.231}$$

Exercise 454 exoMars20100003

Soit A une matrice carré d'ordre 5. Montrer que si la somme des éléments de chaque ligne de A est le même réel α, alors α est une valeur propre de A.

Donner le vecteur propre de A associé à la valeur propre α.

corrMars20100003

Correction of the exercise 454

Si on n'a aucune idée de la solution, le mieux est d'essayer sur un petit exemple pour se donner une idée. Prenons par exemple cette petite matrice dont la somme des éléments de chaque ligne est 4 :

$$\begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}. \tag{93.232}$$

Cherchons les vecteurs propres pour la valeur propre 4. Le système à résoudre est

$$\begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 4 \begin{pmatrix} x \\ y \end{pmatrix}, \tag{93.233}$$

dont les solutions sont $x = y$, c'est-à-dire les multiples de $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Qu'est-ce qu'on parie que dans le cas de l'exercice la réponse sera que le vecteur propre est

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} ? \tag{93.234}$$

Regardons donc le produit

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{22} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{12} + a_{13} + a_{14} + a_{15} \\ a_{21} + a_{22} + a_{23} + a_{24} + a_{25} \\ a_{31} + a_{32} + a_{33} + a_{34} + a_{35} \\ a_{41} + a_{42} + a_{43} + a_{44} + a_{45} \\ a_{51} + a_{52} + a_{53} + a_{54} + a_{55} \end{pmatrix}. \tag{93.235}$$

Par hypothèse, le vecteur du membre de droite vaut

$$\begin{pmatrix} \alpha \\ \alpha \\ \alpha \\ \alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \tag{93.236}$$

donc on a bien un vecteur propre de valeur propre α.

Exercise 455 exoMars20100004

Soit E le sous-espace de $\mathbb{R}^3$ engendré par $\{(1,3,2),(0,2,-1),(3,7,7),(2,2,6)\}$.

(1) Donner une base orthonormée de E.

(2) Déterminer la dimension de E.

corrMars20100004

Correction of the exercise 455

Vu qu'on est dans $\mathbb{R}^3$, l'espace E est au maximum de dimension 3. Il est directement évident que les deux premiers sont linéairement indépendants. Donc $\dim(E) \geqslant 2$. La question est de savoir si il est de dimension 2 ou 3.

Nommons v_1, v_2, v_3 et v_4 les quatre vecteurs proposés.

Il y a de nombreuses façons de procéder. Une façon est de considérer la matrice formée par les 4 vecteurs proposés et de voir, en l'échelonnant, quelle est la matrice carré de taille maximum qu'on peut trouver dedans. En partant de

$$\begin{pmatrix} 1 & 0 & 3 & 2 \\ 3 & 2 & 7 & 2 \\ 2 & -1 & 7 & 6 \end{pmatrix}, \tag{93.237}$$

on peut arriver à

$$\begin{pmatrix} 1 & 0 & 3 & 2 \\ 1 & 0 & 3 & 2 \\ 2 & -1 & 7 & 6 \end{pmatrix}. \tag{93.238}$$

Dans cette matrice, c'est tout de suite visible qu'il n'y a aucune matrice de taille 3 de déterminant non nul parce que toute la première ligne est égale à la seconde.

Une autre façon de procéder (plus piétonne) est de vérifier si les vecteurs v_3 et/ou v_4 sont des combinaisons de v_1 et v_2. Si ils sont de telles combinaisons, alors les deux premiers génèrent tout E. Si en revanche v_3 n'est pas une combinaison de v_1 et v_2, alors ce sont ces trois vecteurs v_1, v_2 et v_3 qui généreront tout E.

Vérifier si trois vecteurs dans $\mathbb{R}^3$ sont linéairement indépendants revient à calculer le déterminant de la matrice formée par ces trois vecteurs. Dans notre cas, on doit vérifier les deux déterminants

$$\det \begin{pmatrix} 1 & 0 & 3 \\ 3 & 2 & 7 \\ 2 & -1 & 7 \end{pmatrix} \tag{93.239}$$

et

$$\det \begin{pmatrix} 1 & 0 & 2 \\ 3 & 2 & 2 \\ 2 & -1 & 6 \end{pmatrix}. \tag{93.240}$$

Le calcul montre que ces deux déterminants sont nuls, donc chacun des deux vecteurs v_3 et v_4 n'est combinaison linéaire de v_1 et v_2.

Une base de E est donc donnée par $\{v_1, v_2\}$. La dimension est donc 2.

Pour trouver une base orthogonale, on utilise Gram-Schmidt. Nous posons $f_1 = v_1$, et puis

$$f_2 = v_2 - \frac{\langle v_2, f_1 \rangle}{\langle f_1, f_1 \rangle} f_1 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} - \frac{4}{14} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} -2 \\ 8 \\ -11 \end{pmatrix}. \tag{93.241}$$

Une base orthogonale de E est donc donnée par

$$f_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad \text{et} \quad f_2 = \begin{pmatrix} -2 \\ 8 \\ -11 \end{pmatrix}. \tag{93.242}$$

Vérification : $\langle f_1, f_2 \rangle = -2 + 24 - 22 = 0$. Notez que j'ai « oublié » le $\frac{1}{7}$ dans f_2 parce que de toutes façons, il n'est pas encore normé.

Une base orthonormale se trouve en divisant par la norme :

$$g_1 = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad \text{et} \quad g_2 = \frac{1}{3\sqrt{21}} \begin{pmatrix} -2 \\ 8 \\ -11 \end{pmatrix}. \tag{93.243}$$

Chapitre 94

Pour les géographes (Bruxelles)

Exercise 456 exoJanvier001

Esquisser dans le système d'axes le graphe de la fonction indiquée :

(1) x^2+1
(2) x^3
(3) $\tan(x)$
(4) $\log_2(x)$
(5) $|\ln(x)|$
(6) $\ln|x|$
(7) $\sin(x+1)$
(8) $\sin(2\pi x)$
(9) e^{-x}.

corrJanvier001

Correction of the exercise 456

Exercise 457 exoJanvier002

Calculer $(1+\sqrt{3}i)^{39}$ et en donner l'expression la plus simple possible.

corrJanvier002

Correction of the exercise 457

Commençons par mettre le nombre proposé sous forme trigonométrique $\rho e^{i\theta}$. La norme de $1+\sqrt{3}$ est $\sqrt{1+(\sqrt{3})^2}=2$. Afin de trouver l'angle θ, nous savons que

$$1+\sqrt{3}i=2(\frac{1}{2}+\frac{\sqrt{3}}{2}i), \tag{94.1}$$

qui doit être égal à $2e^{i\theta}$. L'angle θ est donc l'angle qu'il faut pour que $\cos(\theta)=\frac{1}{2}$ et $\sin(\theta)=\frac{\sqrt{3}}{2}$, c'est-à-dire $\theta=\pi/3$.

Nous devons donc calculer $\left(2e^{i\pi/3}\right)^{39}=2^{39}e^{39i\pi/3}=-2^{39}$, parce que $\frac{39\pi}{3}=12\pi$, et que l'angle 12π est le même que l'angle π et que $e^{i\pi}=-1$.

Exercise 458 exoJanvier003

Soient $n\in\mathbb{N}_0$ et a_0, a_1,..., $a_n\in\mathbb{R}$ ($a_n\neq 0$). Vérifier que si $z\in\mathbb{C}$ est solution de l'équation $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$, alors $\bar{z}$ l'est également.

corrJanvier003

Correction of the exercise 458

On sait que z est solution, c'est-à-dire

$$a_nz^n+\cdots+a_1z+a_0=0$$

donc en prenant la conjugué des deux membres de l'équation, on obtient

$$\overline{a_nz^n+\cdots+a_1z+a_0}=\bar{0}$$

et, comme $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$ et $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ pour tout $z_1, z_2 \in \mathbb{C}$, on en déduit

$$\bar{a}_n \bar{z}^n + \cdots + \bar{a}_1 \bar{z} + \bar{a}_0 = \bar{0}$$

c'est-à-dire, comme $\bar{a} = a$ pour tout $a \in \mathbb{R}$,

$$a_n \bar{z}^n + \cdots + a_1 \bar{z} + a_0 = 0$$

ce qui montre que $\bar{z}$ est solution de l'équation.

Exercise 459 [exoJanvier004]

(1) Énoncer un critère de convergence (au choix) pour les séries et en donner le nom.

(2) La série $\sum_{k=2}^{\infty} \frac{1}{(\ln(k))^k}$ converge-t-elle ? Justifier brièvement.

[corrJanvier004]

Correction of the exercise 459

(1) Nous pouvons par exemple donner le **critère de la racine**. Soit $\sum_k a_k$ une série dont les termes sont positifs. Supposons que la limite

$$\lim_{k\to\infty} \sqrt[k]{a_k}$$

existe et vaut α, alors si $\alpha < 1$, la série converge, et si $\alpha > 1$ la série diverge.

(2) Cette série converge par application du critère de la racine ; en effet, dans ce cas ci $a_k = \frac{1}{\ln(k)^k}$ et donc

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \lim_{k\to\infty} \frac{1}{\ln(k)} = 0 < 1.$$

Exercise 460 [exoJanvier005]

Montrer par l'absurde que $\sqrt{2} \notin \mathbb{Q}$.

[corrJanvier005]

Correction of the exercise 460

Supposons $\sqrt{2}$ soit un rationnel. Alors il existe p, q des entiers positifs tels que

$$\sqrt{2} = \frac{p}{q}$$

soit une fraction irréductible.

On en déduit que $2q^2 = p^2$ donc p^2 est pair, et donc p est pair. En posant $p = 2r$, on obtient alors $2q^2 = 4r^2$, et par conséquent $q^2 = 2r^2$ est pair également, et de même pour q. On en déduit que la fraction $\frac{p}{q}$ n'était pas irréductible puisque p et q sont tous les deux divisibles par 2. Contradiction.

Exercise 461 [exoJanvier006]

Donner l'aire d'un hexagone régulier de côté a.

[corrJanvier006]

Correction of the exercise 461

Un hexagone régulier de côté a est la réunion de 6 triangles équilatéraux de côté a. Chacun de ces triangles a une base a et une hauteur $\sqrt{a^2 - \frac{a^2}{4}} = \frac{\sqrt{3}a}{2}$ (par Pythagore), donc l'aire de l'hexagone est :

$$3a\frac{\sqrt{3}a}{2} = \frac{3\sqrt{3}}{2}a^2$$

Exercise 462 [exoJanvier007]

Charles-Édouard souhaite repeindre sa voiture. Pour ce faire, il dispose de pots de peinture verte, jaune, bleue, rouge et noire en suffisance. Il souhaite cependant repeindre séparément le coffre, le capot, le pare-choc, le toit et chacune des quatre portières. De combien de façons différentes peut-il le faire sachant qu'il souhaite qu'au moins une des portières soit verte et que le pare-choc ne peut pas être bleu ? Justifier brièvement.

[corrJanvier007]

Correction of the exercise 462

Charly a cinq couleurs a disposition, et a huit éléments à repeindre.

En ce qui concerne les portières, sans tenir compte de la restriction, il a 5^4 possibilités différentes. Or il faut qu'au moins une des portières soit verte, ce qui exclut les 4^4 cas où cette couleur n'est pas utilisée. Soit $5^4 - 4^4$ possibilités.

Pour le pare-chocs, il a 4 possibilités, et pour chacun des trois autres éléments, 5 possibilités.

Au total, il y a donc $4 \times 5^3 \times (5^4 - 4^4)$ possibilités pour barioler la voiture de Charles-Edouard.

Exercise 463 exoJanvier008

Soit f la fonction définie par $f(x) = \frac{1}{2-x}$ pour tout $x \in \mathbb{R}\backslash\{2\}$. Montrer par récurrence que

$$f^n(x) = \frac{n-(n-1)x}{(n+1)-nx} \tag{94.2}$$

pour tout $n \in \mathbb{N}_0$.

Rappel : $f^n(x) = \underbrace{f \circ \ldots \circ f}_{n}(x)$.

corrJanvier008

Correction of the exercise 463

Pour $n = 1$, c'est évident. Supposons pour n fixé

$$f^n(x) = \frac{n-(n-1)x}{(n+1)-nx}$$

et calculons :

$$\begin{aligned} f^{n+1}(x) &= f(f^n(x)) \\ &= f\left(\frac{n-(n-1)x}{(n+1)-nx}\right) \\ &= \frac{1}{2-\frac{n-(n-1)x}{(n+1)-nx}} \\ &= \frac{(n+1)-nx}{2((n+1)-nx)-(n-(n-1)x)} \\ &= \frac{(n+1)-((n+1)-1)x}{((n+1)+1)-(n+1)x} \end{aligned}$$

ce qui est bien la formule pour $f^{n+1}(x)$.

Exercise 464 exoJanvier009

Soit $f\colon \mathbb{R} \to \mathbb{R}$, une fonction dérivable.

(1) Donner la définition de $f'(x)$.

(2) En utilisant la définition de $f'(x)$, en déduire que si f et g sont deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ dérivables, alors

$$(fg)' = f'g + fg'. \tag{94.3}$$

corrJanvier009

Correction of the exercise 464

(1) La dérivée de la fonction f est donnée par

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

(2) Soit $x \in \mathbb{R}$, alors

$$\begin{aligned} (fg)'(x) &= \lim_{h\to 0} \frac{(fg)(x+h)-(fg)(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h} \\ &= \lim_{h\to 0} \frac{(f(x+h)-f(x))g(x+h)+f(x)(g(x+h)-g(x))}{h} \\ &= f'(x)g(x)+f(x)g'(x) \end{aligned}$$

en ayant utilisé la règle « la limite du produit est le produit des limites lorsque celles-ci existent ».

Exercise 465 exoJanvier010
Énoncer le théorème des accroissements finis, et l'illustrer par un dessin. corrJanvier010

Correction of the exercise 465
Soit $f : [a, b] \to \mathbb{R}$ une fonction continue et dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{94.4}$$

Exercise 466 exoJanvier011
Énoncer la règle de l'Hospital. corrJanvier011

Correction of the exercise 466
Si f et g sont deux fonctions dérivables en a et telles que $f(a) = g(a) = 0$, et si de plus la limite

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{94.5}$$

existe. Alors

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{94.6}$$

Exercise 467 exoJanvier012
À une soirée avec n personnes, il y a au moins deux personnes ayant le même nombre d'amis présents à la soirée. Vrai ou faux ? Justifier brièvement. corrJanvier012

Correction of the exercise 467
Soit $f : P \to \{0, 1, \dots, n-1\}$ la fonction qui a chaque personne (les éléments de P) associe son nombre d'amis présents à la soirée.

Supposons d'abord que tout le monde ait au moins un ami. Alors f est à valeurs dans $\{1, \dots, n-1\}$ qui contient $n-1$ éléments. Or il y a n personnes dans P, donc f n'est pas injective (principe des tiroirs). Ceci montre qu'au moins deux personnes ont le même nombre d'amis dans ce cas.

Si par contre un des convives n'a aucun ami présent à la soirée, alors soit il n'est pas seul dans ce cas (auquel cas il y a bien deux personnes qui ont le même nombre d'amis : zéro), soit il est seul dans ce cas, et les $n-1$ personnes restantes ont toutes au moins un ami, ce qui nous ramène au premier cas.

Exercise 468 exoJanvier013
Pour quelles valeurs du paramètre a, les courbes d'équations $y = e^{2x} + 1$ et $y = 2ae^x$ ont-elles des points d'intersection ? Pour quelle(s) valeur(s) de a n'ont-elles qu'un seul point d'intersection ? Vérifier qu'alors, elles possèdent la même tangente en ce point. corrJanvier013

Correction of the exercise 468
Les intersections se trouvent en résolvant $e^{2x} + 1 = 2ae^x$ ce qui se fait en posant par exemple $t = e^x$, et donc il faut résoudre (pour $t > 0$) l'équation

$$t^2 - 2at + 1 = 0$$

dont les solutions existent si $|a| \geqslant 1$ (car le discriminant est $4(a^2 - 1)$) et sont alors données par

$$t = a \pm \sqrt{a^2 - 1}.$$

Comme $\sqrt{a^2 - 1} < |a|$ dès que $|a| > 1$, on en déduit que $t < 0$ si $a \leqslant -1$, et $t > 0$ si $a \geqslant 1$, ce qui montre qu'il n'y a d'intersection que pour $a \geqslant 1$.

De plus, le calcul du discriminant et la discussion précédente montrent que l'intersection est unique si $a = 1$, et cette intersection a lieu en $x = \ln(1) = 0$.

Il reste à voir que la dérivée de $e^{2x}+1$ (en $x=0$) est égale à la dérivée de $2ae^x$ (en $x=0$). Or la première dérivée est $2e^{2x}$, et la deuxième est $2ae^x$, ce qui donne 2 dans les deux cas pour $a=1$ et $x=0$.

Exercise 469 exoJanvier014

On considère un fil (de fer, par exemple) de longueur 10 cm que l'on coupe en morceaux. On utilise le premier morceau pour former un carré. Avec le second, on construit un cercle. Où faut il couper le fil pour que la somme des aires de ces deux figures géométriques soit minimales ? corrJanvier014

Correction of the exercise 469

L'aire d'un carré de périmètre a est $\frac{a^2}{16}$; l'aire d'un cercle de circonférence C est $\frac{C^2}{4\pi}$. On cherche à maximiser

$$f(a,C)=\frac{a^2}{16}+\frac{C^2}{4\pi}$$

sous la contrainte $a+C=10$.

On définit

$$g(a)=16\pi f(a,10-a)=\pi a^2+4(10-a)^2$$

qu'on cherche à maximiser. Dès lors on résout $g'(a)=0$ et on trouve $a=\frac{40}{\pi+4}$, d'où $b=\frac{10\pi}{\pi+4}$. Il faut donc couper à a centimètres du bord (ou à b centimètres).

Chapitre 95

Pour les pharmaciens (Bruxelles)

page 33, 2e quadri, séance 3

ex 23 D'après la page 16bis, une *fonction sinusoïdale* est une fonction de la forme $f(x) = a\sin(\omega x + \varphi)$. Pour avoir $f(0) = 0$, on peut par exemple choisir $\varphi = 0$. Par ailleurs, le maximum de la fonction sinus est 1, donc la maximum de $f(x)$ est a ; choisissons donc $a = 2$. Reste à déterminer ω pour qu'un maximum soit atteint en $x = 1$. Or on sait qu'un maximum de la fonction sinus est atteint en $\frac{\pi}{2}$ par exemple ; choisissons donc $\omega = \frac{\pi}{2}$, de sorte que si $x = 1$, on obtient $2\sin(\frac{\pi}{2}) = 2$ qui est bien le maximum de f.

La fonction choisie est donc f définie par $f(x) = 2\sin(\frac{\pi}{2}x)$.

ex 25 Notons f la fonction définie par

$$f(x) = \frac{x2 + px + q2}{x}$$

et calculons

$$f'(x) = \frac{x2 - q2}{x2}$$

de sorte que $f'(x) = 0$ si et seulement si $x = \pm q$ (avec $q \neq 0$, pour que la dérivée ait un sens). On remarque également, par le signe de la dérivée, que l'extrémum obtenu en $-|q|$ est un maximum, l'autre étant un minimum.

Dès lors nous aurons deux cas possibles : si on impose $f(-q) = 0$ et $f(q) = 4$, on obtient $q = 1$ et $p = 2$, et si on impose $f(-q) = 4$ et $f(q) = 0$, on obtient $q = -1$ et $p = 2$.

Dans les deux cas, les extrémas sont $(-1, 0)$ (maximum) et $(1, 4)$ (minimum).

ex 29 Si le carton fait a de largeur et b de hauteur, alors la surface imprimable –tenant compte des marges– est $(a - 3)(b - 2)$ et est fixée à 54. On veut donc minimiser la surface de carton, donnée par ab, sachant que $a = 3 + \frac{54}{b-2}$.

Définissons $f(b) = 3b + \frac{54b}{b-2}$, et trouvons-en le minimum. Sa dérivée est

$$f'(b) = 3 + \frac{54(b-2) - 54b}{(b-2)2} = \frac{3(b-2)2 - 108}{(b-2)2}$$

et s'annule pour $b = -4$ (à rejeter, n'a pas de sens pour une longueur) ou $b = 8$. D'après le signe de la dérivée, $b = 8$ fournit bien un minimum.

En conclusion, $b = 8$ et $a = 12$.

ex 30 Soient a et b ces nombres. On sait $a, b \geqslant 0$ et $a + b = 12$, donc $b = 12 - a$.

(1) On veut minimiser $a2 + b2 = a2 + (12 - a)2 = f(a)$. La dérivée $f'(a) = 2a - 2(12 - a)$ s'annule pour $a = 6$. La solution est donc $a = b = 6$.

(2) On veut maximiser $ab2$ (ou $ba2$, mais il suffit d'échanger les nombres pour retomber sur le premier cas). On définit $f(a) = a(12-a)2$, et la dérivée

$$f'(a) = (12-a)2 - 2a(12-a)$$

s'annule lorsque $a = 12$ (mais alors $b = 0$, à rejeter, ceci n'est pas un maximum) ou lorsque $a = 4$; les solutions sont donc $(4, 8)$ et le symétrique $(8, 4)$.

(3) On veut maximiser $ba3$ (même remarque que ci-dessus), donc on définit $f(a) = (12-a)a3$ dont la dérivée est

$$f'(a) = -a3 + 3(12-a)a2$$

et s'annule pour $a = 0$ (pas un maximum) ou $9 = a$; donc les solutions sont $(9, 3)$ et $(3, 9)$.

ex 32 On cherche (x, y) tel que $y2 = 4ax$ et minimisant la distance

$$d = \|(x,y) - (2a,a)\| = \sqrt{(x-2a)2 + (y-a)2} = \sqrt{\left(\frac{y2}{4a} - 2a\right)2 + (y-a)2}$$

Remarquons que minimiser d revient à minimiser $d2$, donc posons

$$f_a(y) = \left(\frac{y2}{4a} - 2a\right)2 + (y-a)2$$

et calculons la dérivée

$$f'_a(y) = 2\left(\frac{y2}{4a} - 2a\right)\frac{y}{a} + 2(y-a) = \frac{y3}{4a2} - 2a$$

qui s'annule lorsque $y = 2a$. Donc la solution est $(x, y) = (a, 2a)$

ex 33 Soit x la distance "sur le rivage" par rapport au premier bateau où sera débarqué le passager (x est entre 0 et 5). Alors il s'agit de minimiser

$$d(x) = \sqrt{9 + x2} + \sqrt{(5-x)2 + 81}$$

donc on calcule la dérivée

$$d'(x) = \frac{x}{\sqrt{9+x2}} - \frac{5-x}{\sqrt{(5-x)2+81}} = \frac{x\sqrt{(5-x)2+81} + (x-5)\sqrt{9+x2}}{\sqrt{9+x2}\sqrt{(5-x)2+81}}$$

qui s'annule lorsque $3|x| = |x-5|$ càd lorsque $3x = 5-x$ (car $x \in [0,5]$), donc $x = \frac{5}{4}$.

Le trajet minimal du bateau est donc $d(\frac{5}{4}) = 13$.

Une autre de manière de voir le problème est de considérer le principe de réflexion : le trajet minimal est alors donné par $\sqrt{52 + 122} = 13$.

95.1 Outils math

Exercise 470 exoOutilsMath-0001

Montrer que si X et Y sont orthogonaux, alors

$$\|X+Y\|^2 = \|X\|^2 + \|Y\|^2. \tag{95.1}$$

Montrer sur un dessin pourquoi cette équation n'est rien d'autre que le théorème de Pythagore. corrOutilsMath-0001

Correction of the exercise 470

<+CorrOutilsMath-0001+>

Exercise 471 exoOutilsMath-0004

Vérifier graphiquement et par calcul que $\tan(\theta+\pi)=\tan(\theta)$. corrOutilsMath-0004

Correction of the exercise 471

<+CorrOutilsMath-0004+>

Exercise 472 exoOutilsMath-0007

Donner les coordonnées sphériques des points

(1) $(x,y,z)=(3,-3,3\sqrt{2})$
(2) $(x,y,z)=\left(\frac{1}{4},\frac{\sqrt{3}}{4},\frac{\sqrt{3}}{2}\right)$
(3) $(x,y,z)=(0,0,-5)$

corrOutilsMath-0007

Correction of the exercise 472

Détaillons le premier exercice. Le changement de variable pour les coordonnées sphériques est donné par les équations (18.805) :

$$x=\rho\sin(\theta)\cos(\varphi)=3 \tag{95.2a}$$
$$y=\rho\sin(\theta)\sin(\varphi)=-3 \tag{95.2b}$$
$$z=\rho\cos(\theta)=3\sqrt{2}. \tag{95.2c}$$

La première chose à calculer est la coordonnées radiale ρ. Elle est donnée par

$$\rho=\sqrt{x^2+y^2+z^2}=6. \tag{95.3}$$

Maintenant nous pouvons trouver θ en écrivant l'équation pour z :

$$6\cos(\theta)=3\sqrt{2}. \tag{95.4}$$

Nous trouvons $\cos(\theta)=\sqrt{2}/2$. De là, θ vaut soit $\frac{\pi}{4}$ soit $-\frac{\pi}{4}$. Nous devons choisir $\theta=\frac{\pi}{4}$ parce que par définition des coordonnées polaires, nous avons toujours $\theta\in[0,\pi]$.

Écrivons maintenant l'équation pour y :

$$5\frac{\sqrt{2}}{2}\sin(\varphi)=-3, \tag{95.5}$$

donc $\sin(\varphi)=-\frac{1}{\sqrt{2}}$. L'angle φ vaut alors soir $-\frac{\pi}{4}$ soit $\frac{5\pi}{4}$. Pour choisir, nous regardons l'équation pour x qui nous indique que $\cos(\varphi)$ doit être positif. Par conséquent nous choisissons $\varphi=-\frac{\pi}{4}$. Par convention, nous prenons toutefois l'angle φ entre 0 et 2π, donc nous ne prenons pas $\varphi=-\frac{\pi}{4}$ mais l'angle équivalent $\varphi=\frac{7\pi}{4}$.

Les coordonnées (r,θ,ϕ) sont :

(1) $(\rho,\theta,\varphi)=(6,\frac{\pi}{4},\frac{7\pi}{4})$;
(2) $\left(2,\frac{\pi}{2},-\frac{\pi}{3}\right)$;
(3) $\left(1,\frac{\pi}{6},\frac{\pi}{3}\right)$;
(4) $(5,\pi,0)$.

Exercise 473 exoOutilsMath-0008

Trouver les coordonnées cartésiennes des points suivants :

(1) $(r,\theta)=(6,\frac{3\pi}{4})$;
(2) $(r,\theta)=(1,-\frac{\pi}{6})$;
(3) $(r,\theta)=(\sqrt{2},-\pi)$.
(4) $(\rho,\theta,\varphi)=(3,\frac{\pi}{4},\frac{\pi}{6})$,
(5) $(r,\theta,z)=(8,\frac{2\pi}{3},-3)$

corrOutilsMath-0008

Correction of the exercise 473

(1) $(x, y) = (-3\sqrt{2}, 3\sqrt{2})$
(2) $(x, y) = (\sqrt{3}/2, -1/2)$
(3) $(x, y) = (-\sqrt{2}, 0)$
(4) $(x, y, z) = ((3\sqrt{2}\sqrt{3}/4, 3\sqrt{2}/4, 3\sqrt{2}/3))$
(5) $(x, y, z) = (-4, 4\sqrt{3}, -3)$.

Exercise 474 exoOutilsMath-0009

Dessiner la trajectoire d'un mobile dont les équations du mouvement en coordonnées sphériques sont

$$P(t) = (1, \frac{\pi}{2}, t). \tag{95.6}$$

corrOutilsMath-0009

Correction of the exercise 474

<+CorrOutilsMath-0009+>

Exercise 475 exoOutilsMath-0010

Donner l'équation cartésienne d'une sphère centrée à l'origine et de rayon R. Quelle est son équation en coordonnées sphériques ?

corrOutilsMath-0010

Correction of the exercise 475

En coordonnées cartésiennes, l'équation de la sphère centrée en $(0, 0, 0)$ est $x^2 + y^2 + z^2 = R^2$. En coordonnées sphériques, la sphère prend la forme plus simple

$$\rho = R. \tag{95.7}$$

Exercise 476 exoOutilsMath-0011

Donner l'équation cartésienne d'un cylindre infini de rayon R et d'axe $\overrightarrow{OZ}$. Quelle est son équation en coordonnées cylindriques ?

corrOutilsMath-0011

Correction of the exercise 476

Le cylindre est composé de cercles horizontaux de rayon R. Son équation cartésienne est donc $x^2 + y^2 = R^2$. Il n'y a pas de contraintes sur z.

En coordonnées cylindriques, l'équation devient

$$r = R. \tag{95.8}$$

Exercise 477 exoOutilsMath-0012

Donner les coordonnées polaires du point dont les coordonnées cartésiennes sont $(x, y) = (\sqrt{3}, 1)$.

corrOutilsMath-0012

Correction of the exercise 477

Nous avons représenté la situation sur la figure 95.1. Tout d'abord nous trouvons la longueur du segment qui joint l'origine au point $P = (\sqrt{3}, 1)$. Nous avons $l = \sqrt{(\sqrt{3})^2 + 1^2} = 2$.

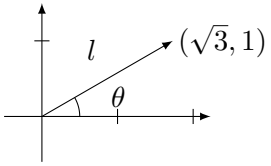


FIGURE 95.1: Il s'agit de trouver l et θ en sachant que $x = \sqrt{3}$ et $y = 1$. LabelFigExoPolaire

Nous savons donc déjà que la coordonnée radiale sera $r = 2$. Afin de trouver l'angle, nous utilisons par exemple le fait que le sinus de θ est la hauteur divisée par l'hypoténuse :

$$\sin(\theta) = \frac{1}{2}. \tag{95.9}$$

Cela nous permet de savoir que $\frac{\pi}{6}$ ou bien $\frac{\pi}{6} + \pi$. Nous choisissons $\frac{\pi}{6}$ parce que l'angle $\frac{\pi}{6} + \pi$ donnerait des coordonnée négatives.

Exercise 478 exoOutilsMath-0013

Donner les coordonnées cartésiennes du point dont les coordonnées polaires sont $(r, \theta) = (4, \frac{\pi}{4})$. corrOutilsMath-0013

Correction of the exercise 478

Nous utilisons le changement de variable (18.751) :

$$\begin{cases} x = r\cos\frac{\pi}{4} = 2\sqrt{2} & \text{(95.10a)} \\ y = r\sin\frac{\pi}{4} = 2\sqrt{2} & \text{(95.10b)} \end{cases}$$

Les coordonnées cartésiennes du point sont donc $(2\sqrt{2}, 2\sqrt{2})$.

Exercise 479 exoOutilsMath-0014

Donner l'équation cartésienne du cercle de centre $(-1, 2)$ et de rayon 1. corrOutilsMath-0014

Correction of the exercise 479

Nommons S le cercle de centre $(-1, 2)$ et de rayon 1. Nous avons $(x, y) \in S$ si et seulement si la distance entre (x, y) et $(-1, 2)$ vaut exactement 1 :

$$f\big((x, y), (-1, 2)\big) = 1. \tag{95.11}$$

L'équation du cercle S est donc

$$(-1 - x)^2 + (2 - y)^2 = 1^2, \tag{95.12}$$

c'est-à-dire

$$(x + 1)^2 + (y - 2)^2 = 1. \tag{95.13}$$

Exercise 480 exoOutilsMath-0015

Donner l'équation polaire du cercle de centre $(0, 0)$ et de rayon 3. corrOutilsMath-0015

Correction of the exercise 480

Un point appartient au cercle si et seulement si sa distance à l'origine vaut 3. L'équation du cercle est donc

$$r = 3 \tag{95.14}$$

et pas de contraintes sur la coordonnée angulaire θ.

Morale : les coordonnée polaires sont très pratiques pour travailler sur des cercles centrés en l'origine.

Exercise 481 exoOutilsMath-0016

(1) Donner un vecteur orthogonal à $(1, 2, 3)$. Merci de ne pas répondre « le vecteur nul ».
(2) Quel est l'équation du lieu géométrique des vecteurs orthogonaux à $(1, 1, 1)$? corrOutilsMath-0016

Correction of the exercise 481

(1) Il s'agit de trouver n'importe quel vecteur dont le produit scalaire avec $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ est nul. Par exemple le vecteur

$$v = \begin{pmatrix} 1 \\ -1/2 \\ 0 \end{pmatrix} \tag{95.15}$$

fait l'affaire.

(2) Un vecteur (x, y, z) est perpendiculaire à $(1, 1, 1)$ si le produit scalaire est nul. Dans ce cas, l'équation est

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = x + y + z = 0. \tag{95.16}$$

Cela est un plan. C'est le plan perpendiculaire au vecteur $(1, 1, 1)$.

Exercise 482 exoOutilsMath-0017

Donner la projection du vecteur $\frac{1}{\sqrt{3}}(1, 1, 1)$ sur la droite qui contient le vecteur $\frac{2}{\sqrt{2}}(1, 0, 1)$. exoOutilsMath-0017

Correction of the exercise 482

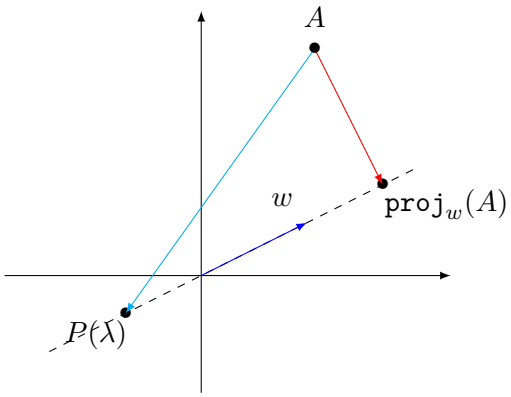


FIGURE 95.2: Pour l'exercice 482. À partir du point A, il faut trouver quel vecteur tombe perpendiculairement à la droite contenant w. LabelFigExoProjection

Étant donné que la projection est une application linéaire, nous commençons par déterminer le vecteur $\mathtt{proj}_{(1,0,1)}(1, 1, 1)$; ensuite il suffira de multiplier la réponse par $1/\sqrt{3}$.

La projection dont nous parlons sera certainement un multiple de $(1, 0, 1)$. Tout l'exercice se réduit à savoir quel multiple. Soit

$$P(\lambda) = \lambda \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \tag{95.17}$$

Il faut trouver λ de telle manière que le vecteur qui joint $(1, 1, 1)$ à $P(\lambda)$ soit perpendiculaire à $(1, 0, 1)$, voir figure 95.2.

Si nous nommons $v(\lambda)$ le vecteur qui va de $(1, 1, 1)$ à $P(\lambda)$, nous avons

$$v(\lambda) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} \lambda \\ 0 \\ \lambda \end{pmatrix} = \begin{pmatrix} 1 - \lambda \\ 1 \\ 1 - \lambda \end{pmatrix} \tag{95.18}$$

Nous devons donc résoudre l'équation

$$v(\lambda) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 0, \tag{95.19}$$

c'est-à-dire $(1 - \lambda) + (1 - \lambda) = 0$, et par conséquent $\lambda = 1$. La projection de $(1, 1, 1)$ est finalement

$$\mathtt{proj}_{(1,0,1)}(1, 1, 1) = P(1) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{95.20}$$

et

$$\mathtt{proj}_{(1,0,1)} \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}} P(1) = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{95.21}$$

Exercise 483 exoOutilsMath-0018
Donner l'équation polaire de la droite qui passe par les points $(1,0)$ et $(0,1)$. corrOutilsMath-0018

Correction of the exercise 483
Il s'agit de la droite d'équation cartésiennes $x+y=1$. L'équation en coordonnées polaires est obtenue en remplaçant x par $r\cos(\theta)$ et y par $r\sin(\theta)$:

$$r\cos(\theta)+r\sin(\theta)=1, \tag{95.22}$$

ou encore, en factorisant,

$$r\big(\cos(\theta)+\sin(\theta)\big)=1. \tag{95.23}$$

Cela est l'équation, mais maintenant il faut préciser le domaine de variation des variables. En effet si nous prenons par exemple $\theta=\pi$, nous voyons $-r=1$. Mais le rayon doit toujours être positif. Tous les angles ne sont donc pas acceptables. D'ailleurs sur un dessin nous voyons tout de suite que la droite ne passe pas par tous les angles.

Les angles acceptables sont ceux tels que $\cos(\theta)+\sin(\theta)>0$. L'équation

$$\cos(\theta)=-\sin(\theta) \tag{95.24}$$

a pour solutions les angles $\theta_1=\frac{3\pi}{4}$ et $\theta_2=-\frac{\pi}{4}$. Les angles acceptables sont donc ceux entre θ_1 et θ_2. L'équation complète de la droite est alors

$$r\big(\cos(\theta)+\sin(\theta)\big)=1, \tag{95.25}$$

avec $\theta\in\left]-\frac{\pi}{4},\frac{3\pi}{4}\right[$.

Exercise 484 exoOutilsMath-0019
Donner l'équation polaire de la courbe d'équation $y=\frac{1}{x}$. corrOutilsMath-0019

Correction of the exercise 484
En remplaçant x par $r\cos(\theta)$ et y par $r\sin(\theta)$, nous trouvons immédiatement l'équation

$$r^2=\frac{1}{\sin(\theta)\cos(\theta)}. \tag{95.26}$$

Il faut trouver le domaine de variation de θ. Étant donné que r^2 doit être positif, seuls les θ tels que $\sin(\theta)\cos(\theta)>0$ sont éligibles. Nous avons donc le domaine de variation

$$\theta\in\left]0,\frac{\pi}{2}\right[\cup\left]\pi,\frac{3\pi}{2}\right[. \tag{95.27}$$

En confirmation, le graphique de la figure 95.3 montre bien que la courbe n'est présente que dans les premiers et troisièmes quadrants.

Exercise 485 exoOutilsMath-0020
Soient X et Y, deux vecteurs dans $\mathbb{R}^3$. Si $P=\texttt{proj}_X Y$, que vaut $(Y-P)\cdot X$? corrOutilsMath-0020

Correction of the exercise 485
Par construction, la projection d'un point P sur la droite de X se fait le long de la droite perpendiculaire à X qui passe par P. Le vecteur $Y-P$ sera donc perpendiculaire à X et le produit scalaire sera nul. Faire un dessin sur lequel on visualise P, X et $X-P$.

Nous pouvons obtenir le résultat par calcul. Nous savons que

$$\|\texttt{proj}_X Y\|=\frac{X\cdot Y}{\|X\|}. \tag{95.28}$$

Donc le vecteur $\texttt{proj}_X Y$ est le vecteur de norme $X\cdot Y/\|X\|$ parallèle à X. Par conséquent

$$\texttt{proj}_X Y=\frac{X\cdot Y}{\|X\|}\frac{X}{\|X\|}. \tag{95.29}$$

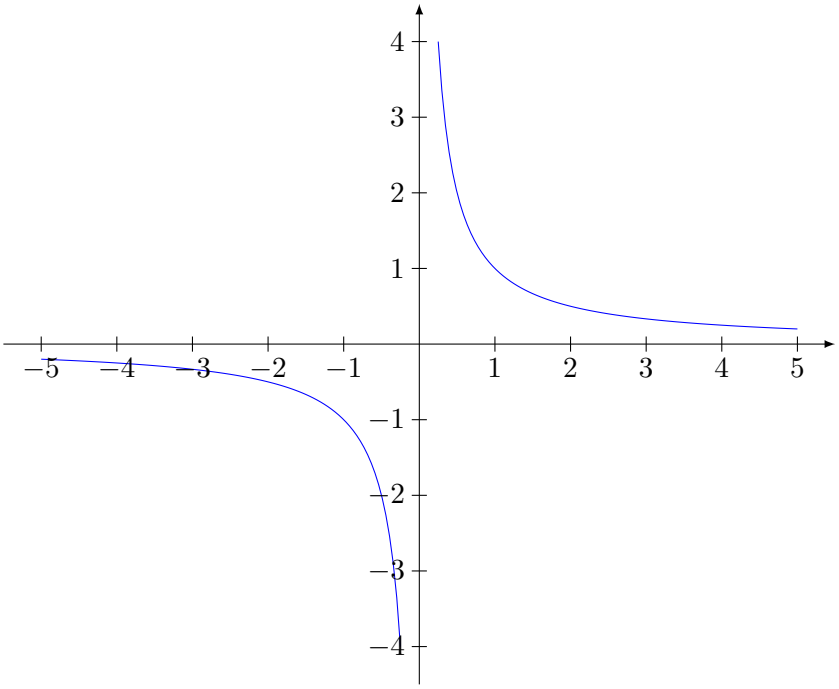


FIGURE 95.3: La fonction $y = 1/x$. LabelFigExoUnSurxPolaire

Donc

$$\begin{aligned}(Y - \mathtt{proj}_X Y) \cdot X = Y \cdot X - \frac{X \cdot Y}{\|X\|^2} X \cdot X \\ = Y \cdot X - Y \cdot X \\ = 0\end{aligned} \tag{95.30}$$

Exercise 486 exoOutilsMath-0021

Donner les vecteurs de norme 1 parallèles à $(-1, 3, 6)$. corrOutilsMath-0021

Correction of the exercise 486

Il s'agit de trouver les multiples de $v = (-1, 3, 6)$ qui ont une norme 1. La norme de v est

$$\|(-1, 3, 6)\| = \sqrt{1 + 9 + 36} = \sqrt{47}. \tag{95.31}$$

Les multiples de norme 1 seront donc

$$\pm \frac{1}{\sqrt{47}} \begin{pmatrix} -1 \\ 3 \\ 6 \end{pmatrix}. \tag{95.32}$$

Exercise 487 exoOutilsMath-0022

Calculer les dérivées partielles des fonctions suivantes :

(1) $f(x, y) = \cos(xy)$
(2) $f(x, y) = \sqrt{R^2 - x^2 - y^2}$.

corrOutilsMath-0022

Correction of the exercise 487

(1) $\frac{\partial f}{\partial x} = -y \sin(xy)$, $\frac{\partial f}{\partial y} = -x \sin(y)$.
(2)

$$\begin{aligned}\frac{\partial \sqrt{R^2 - x^2 - y^2}}{\partial x} = \frac{-x}{\sqrt{R^2 - x^2 - y^2}} \\ \frac{\partial \sqrt{R^2 - x^2 - y^2}}{\partial y} = \frac{-y}{\sqrt{R^2 - x^2 - y^2}}\end{aligned} \tag{95.33}$$

```
sage: f(x,y)=cos(x*y)
sage: f.diff(x)
(x, y) |--> -y*sin(x*y)
sage: f.diff(y)
(x, y) |--> -x*sin(x*y)
sage: var('R')
R
sage: f(x,y)=sqrt(R^2-x^2-y^2)
sage: f.diff(x)
(x, y) |--> -x/sqrt(R^2 - x^2 - y^2)
sage: f.diff(y)
(x, y) |--> -y/sqrt(R^2 - x^2 - y^2)
```

Notez l'utilisation de `f.diff(x)` pour la dérivée partielle dans la direction x et `f.diff(y)` pour celle dans la direction y.

Notez aussi qu'il a fallu utiliser `var('R')` pour déclarer que `R` sera une variable.

Exercise 488 exoOutilsMath-0023

Donner les points suivants en coordonnées polaires :

(1) $(x,y)=(1,1)$;
(2) $(x,y)=(2\sqrt{3},-2)$;
(3) $(x,y)=(-1,1)$.

corrOutilsMath-0023

Correction of the exercise 488

(1) $(r,\theta)=(\sqrt{2},\pi/4)$;
(2) $(r,\theta)=(4,-\pi/6)$;
(3) $(r,\theta)=(\sqrt{2},3\pi/4)$.

Exercise 489 exoOutilsMath-0024

Donner en coordonnées cylindriques l'équation du cône de sommet $(0,0,1)$ et de rayon à la base R.

corrOutilsMath-0024

Correction of the exercise 489

En coordonnées cylindriques, il est souvent pratique de penser en termes de « tranches » horizontales. Ici à la hauteur h nous avons un cercle de rayon $R(1-h)$. L'équation de ce cercle est

$$\begin{cases} r=(1-h)R & (95.34a)\\ z=h. & (95.34b)\end{cases}$$

Cela est pour une tranche. L'équation du cône est alors

$$r=(1-z)R \tag{95.35}$$

avec $z\in[0,1]$.

Exercise 490 exoOutilsMath-0025

Donner en coordonnées sphériques l'équation de l'équateur de la sphère de rayon R centrée en $(0,0,0)$.

corrOutilsMath-0025

Correction of the exercise 490

L'équateur est le cercle horizontal de rayon R. Son équation est alors

$$\begin{cases} r=R & (95.36a)\\ \theta=\dfrac{\pi}{2}. & (95.36b)\end{cases}$$

Imposer la valeur de θ à $\frac{\pi}{2}$ est ce qu'il faut pour dire que nous voulons être dans le plan horizontal $z = 0$.

Exercise 491 exoOutilsMath-0026

Soit S la sphère de rayon $\frac{1}{\sqrt{3}}$ centrée en $(0, 0, 0)$. Donner en coordonnées sphériques l'équation de la partie de S dont la hauteur est plus grande que $\frac{1}{2}$. corrOutilsMath-0026

Correction of the exercise 491

Pour que nous soyons sur la sphère de rayon $1/\sqrt{3}$, il faut certainement imposer $\rho = \frac{1}{\sqrt{3}}$. Afin d'être à une hauteur plus haute que $\frac{1}{2}$, il faut imposer que $z > \frac{1}{2}$, c'est-à-dire

$$\rho\cos(\theta) > \frac{1}{2}, \tag{95.37}$$

ou encore : $\theta < \frac{\pi}{6}$. L'équation recherchée est

$$\begin{cases} r = \dfrac{1}{\sqrt{3}} & \text{(95.38a)} \\ \theta \in [0, \dfrac{\pi}{6}]. & \text{(95.38b)} \end{cases}$$

Exercise 492 exoOutilsMath-0027

Quelle est l'équation de la tangente de la fonction $f(x) = 1/x$ au point $a = 2$? corrOutilsMath-0027

Correction of the exercise 492

Pour trouver le coefficient directeur de la tangente, il faut étudier le rapport

$$\begin{aligned} \frac{f(x) - f(2)}{x - 2} &= \frac{\frac{1}{x} - \frac{1}{2}}{x - 2} \\ &= \frac{2 - x}{2x(x - 2)} \\ &= -\frac{1}{2x}. \end{aligned} \tag{95.39}$$

En prenant la limite $x \to 2$ nous trouvons $-\frac{1}{4}$ comme coefficient directeur.

Il s'agit donc de trouver la droite de coefficient directeur $\frac{-1}{4}$ et qui passe par le point $\big(2, f(2)\big) = (2, \frac{1}{2})$. Si nous écrivons la droite sous la forme $y = ax+b$, la première chose est que $a = -\frac{1}{4}$. Ensuite, il faut que l'équation

$$y = -\frac{1}{4}x + b \tag{95.40}$$

soir vérifiée pour le point $(x, y) = (2, \frac{1}{2})$. Nous trouvons l'équation pour b :

$$-\frac{1}{4} \cdot 2 + b = \frac{1}{2}, \tag{95.41}$$

donc $b = 1$ et l'équation de la tangente recherchée est

$$y = -\frac{1}{4}x + 1. \tag{95.42}$$

Exercise 493 exoOutilsMath-0028

Donner une approximation de $\ln(1.0001)$ sachant que $\ln(1) = 0$. corrOutilsMath-0028

Correction of the exercise 493

Nous utilisons l'approximation (12.395) avec les paramètres suivants :

$$\begin{aligned} x = 1 \qquad & \Delta, x = 0.0001 \\ f(x) = \ln(x), \quad & f'(x) = \frac{1}{x}. \end{aligned} \tag{95.43}$$

Nous avons donc

$$\ln(1.0001) \simeq \ln(1) + 0.0001 = 0.0001. \tag{95.44}$$

Confirmation par Sage :

```
----------------------------------------------------------------------
| Sage Version 4.5.3, Release Date: 2010-09-04                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: numerical_approx(ln(1.0001))
0.000099995000333297 3
```

Exercise 494 exoOutilsMath-0029

Donner une approximation de $\sqrt{4.01}$.

corrOutilsMath-0029

Correction of the exercise 494

Il s'agit d'utiliser la formule $f(x+\Delta x)\simeq f(x)+f'(x)\Delta x$ avec

$$\begin{aligned} x&=4, & \Delta x&=0.01\\ f(x)&=\sqrt{x}, & =f'(x)&=\frac{1}{2\sqrt{x}}. \end{aligned} \tag{95.45}$$

Nous avons alors

$$\sqrt{4.01}\simeq\sqrt{4}+0.01\cdot\frac{1}{2\sqrt{4}}=2+\frac{1}{400}=2.0025. \tag{95.46}$$

Confirmation par Sage :

```
----------------------------------------------------------------------
| Sage Version 4.5.3, Release Date: 2010-09-04                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: numerical_approx(sqrt(4.01))
2.00249843945008
```

Exercise 495 exoOutilsMath-0030

Calculer les différentielles des fonctions suivantes :

(1) $f(x,y)=x^2y+x^3y$ au point $(1,2)$;
(2) $f(x,y)=\sin(xy)+\cos(x^2+y^4)$ au point $(0,0)$;
(3) $f(x,y,z)=x^2yz+y^2z^2+e^{xyz}$ au point $(1,0,1)$.

Vous devez dans chacun de cas donner la valeur de $df_{(a,b)}\begin{pmatrix}u_1\\u_2\end{pmatrix}$ (à adapter pour les cas de $f\colon\mathbb{R}^3\to\mathbb{R}$).

corrOutilsMath-0030

Correction of the exercise 495

(1) D'abord les dérivées partielles :

$$\frac{\partial f}{\partial x}(x,y)=2xy+2x^2y \tag{95.47a}$$
$$\frac{\partial f}{\partial y}(x,y)=x^2+x^3. \tag{95.47b}$$

Donc

$$\begin{aligned} df_{(a,b)}\begin{pmatrix}u_1\\u_2\end{pmatrix}&=\frac{\partial f}{\partial x}(a,b)u_1+\frac{\partial f}{\partial y}(a,b)u_2\\ &=(2ab+3a^2b)u_1+(a^2+a^3)u_2. \end{aligned} \tag{95.48}$$

(2) Les dérivées partielles :

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y)&=-2x\sin(x^2+y^4)+y\cos(xy)\\ \frac{\partial f}{\partial y}(x,y)&=-4y^3\sin(x^2+y^4)+x\cos(xy) \end{aligned} \tag{95.49}$$

Nous avons de toutes façons que les deux sont nulles lorsque $x = y = 0$, dont la différentielle est nulle :

$$df_{(0,0)} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0. \tag{95.50}$$

(3) Les dérivées partielles sont

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xyz + yze^{xyz} \\ \frac{\partial f}{\partial y} &= x^2z + 2yz^2 + xze^{xyz} \\ \frac{\partial f}{\partial z} &= x^2y + 2y^2z + xye^{xyz}. \end{aligned} \tag{95.51}$$

Si on pose $x = 1$, $y = 0$ et $z = 1$, nous trouvons que seule $\partial_y f$ est non nulle et vaut deux. Donc

$$df_{(1,0,1)} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 2u_2. \tag{95.52}$$

Exercise 496 exoOutilsMath-0031

Dans quelle direction (u_1, u_2) la pente de la fonction

$$f(x, y) = x^2 + 2xy \tag{95.53}$$

est-elle maximale au point $(-1, 1)$?

corrOutilsMath-0031

Correction of the exercise 496

<+CorrOutilsMath-0031+>

Exercise 497 exoOutilsMath-0032

Montrer que la force de gravitation dérive d'un potentiel. En d'autres termes, montrer que le champ de vecteur

$$F(r) = k\frac{r}{\|r\|^3} \tag{95.54}$$

est un champ de gradients.

Attention : ici le symbole « r » représente le *vecteur* position. Ce n'est pas le r des coordonnées polaires !

corrOutilsMath-0032

Correction of the exercise 497

Nous devons trouver une fonction $V(x, y, z)$ telle que $\nabla V(r) = k\frac{r}{\|r\|^3}$, c'est-à-dire telle que

$$\nabla V \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{k}{(x^2 + y^2 + z^2)^{3/2}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{95.55}$$

Cette équation revient aux trois équations suivantes :

$$\begin{aligned} \frac{\partial V}{\partial x}(x, y, z) &= \frac{kx}{(x^2 + y^2 + z^2)^{3/2}} \\ \frac{\partial V}{\partial y}(x, y, z) &= \frac{ky}{(x^2 + y^2 + z^2)^{3/2}} \\ \frac{\partial V}{\partial z}(x, y, z) &= \frac{kz}{(x^2 + y^2 + z^2)^{3/2}}. \end{aligned} \tag{95.56}$$

Il est facile de voir que la fonction

$$V(x, y, z) = \frac{-k}{\sqrt{x^2 + y^2 + z^2}} \tag{95.57}$$

fait l'affaire. Nous pouvons aussi l'écrire

$$V(r) = -k\frac{1}{\|r\|}. \tag{95.58}$$

Exercise 498 exoOutilsMath-0033

Soit F un champ de vecteurs et σ_1 un chemin. Montrer que si σ_2 est le chemin opposé, alors la circulation de F le long de σ_1 a le signe opposé à celle le long de σ_2, c'est-à-dire

$$\int_{\sigma_1} F = -\int_{\sigma_2} F. \tag{95.59}$$

corrOutilsMath-0033

Correction of the exercise 498

<+CorrOutilsMath-0033+>

Exercise 499 exoOutilsMath-0034

Prouver que $\nabla \cdot (\nabla \times F) = 0$.

corrOutilsMath-0034

Correction of the exercise 499

<+CorrOutilsMath-0034+>

Exercise 500 exoOutilsMath-0035

Soit f, une fonction et si F est un champ de vecteurs ? Parmi les combinaisons suivantes, lesquelles sont des fonctions, des champs de vecteurs, ou n'existent pas ?

(1) $\nabla(\nabla f)$
(2) $\nabla(\nabla \cdot f)$
(3) $\nabla(\nabla \times f)$
(4) $\nabla \cdot (\nabla f)$
(5) $\nabla \cdot (\nabla \cdot f)$
(6) $\nabla \cdot (\nabla \times f)$
(7) $\nabla \times (\nabla f)$
(8) $\nabla \times (\nabla \cdot f)$
(9) $\nabla \times (\nabla \times f)$
(10) $\nabla(\nabla F)$
(11) $\nabla(\nabla \cdot F)$
(12) $\nabla(\nabla \times F)$
(13) $\nabla \cdot (\nabla F)$
(14) $\nabla \cdot (\nabla \cdot F)$
(15) $\nabla \cdot (\nabla \times F)$
(16) $\nabla \times (\nabla F)$
(17) $\nabla \times (\nabla \cdot F)$
(18) $\nabla \times (\nabla \times F)$

corrOutilsMath-0035

Correction of the exercise 500

Étant donné que f est une fonction, $\nabla \cdot f$ et $\nabla \times f$ n'ont pas de sens. Donc seul ∇f a un sens et est un champ de vecteurs. Par conséquent, $\nabla \times (\nabla f)$ et $\nabla \cdot (\nabla f)$ ont un sens tandis que $\nabla(\nabla f)$ n'a pas de sens.

Étant donné que F est un champ de vecteurs, ∇F n'a pas de sens tandis que $\nabla \times F$ et $\nabla \cdot F$ ont un sens. Le premier est un nouveau champ de vecteurs tandis que le second est une fonction. En ce qui concerne $\nabla \times F$,

(1) $\nabla(\nabla \times F)$, non ;
(2) $\nabla \cdot (\nabla \times F)$, oui ;
(3) $\nabla \times (\nabla \times F)$, oui.

En ce qui concerne $\nabla \times F$ qui est une fonction,

(1) $\nabla(\nabla \cdot F)$, oui ;
(2) $\nabla \cdot (\nabla \cdot F)$, non ;
(3) $\nabla \times (\nabla \cdot F)$, non.

Exercise 501 exoOutilsMath-0036

Est-ce qu'il existe un champ de vecteur A tel que $F = \nabla \times A$ si

$$F(x, y, z) = \begin{pmatrix} x^2y \\ z \\ xyz \end{pmatrix}? \tag{95.60}$$

corrOutilsMath-0036

Correction of the exercise 501

Si $F = \nabla \times A$, alors $\nabla \cdot F = 0$. Vérifions cela :

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_y}{\partial y} = 2xy + 0 + xy \neq 0. \tag{95.61}$$

Étant donné que la divergence de F n'est pas nulle, il n'existe pas de champ A tel que $F = \nabla \times A$.

Exercise 502 exoOutilsMath-0037

Montrer qu'il n'existe pas de fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$ telle que

$$\begin{pmatrix} \cos(x) \\ \sin(x) \\ xy \end{pmatrix} = \nabla f. \tag{95.62}$$

corrOutilsMath-0037

Correction of the exercise 502

Si la fonction f existait, nous aurions $\nabla \times F = \nabla \times (\nabla f) = 0$. Vérifions donc que $\nabla \times F \neq 0$:

$$\begin{aligned} \nabla \times \begin{pmatrix} \cos(x) \\ \sin(x) \\ xy \end{pmatrix} &= \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ \cos(x) & \sin(x) & xy \end{vmatrix} \\ &= (x-0)e_x - (y-0)e_y + \big(\cos(x) - 0\big)e_z \\ &= \begin{pmatrix} x \\ -y \\ \cos(x) \end{pmatrix} \neq 0. \end{aligned} \tag{95.63}$$

Exercise 503 exoOutilsMath-0038

Donner en coordonnées cylindriques l'équation de la droite verticale passant par le point $(x, y, z) = (1, 1, 0)$.

Par « verticale », nous voulons dire perpendiculaire au plan XY.

corrOutilsMath-0038

Correction of the exercise 503

Le point donné a comme coordonnée cylindriques

$$(\sqrt{2}, \frac{\pi}{4}, 0). \tag{95.64}$$

La droite verticale qui passe par ce point est contrainte par le fait de ne pas pouvoir changer r ni θ ; seule la variable z est sans contraintes. L'équation est donc en deux morceaux :

$$\begin{cases} r = \sqrt{2} & \text{(95.65a)} \\ \theta = \dfrac{\pi}{4}. & \text{(95.65b)} \end{cases}$$

Exercise 504 exoOutilsMath-0039

Soit la fonction $f(x) = x\cos(2x)$. Donner l'équation de la tangente au graphe de f en $x = \frac{\pi}{2}$.

corrOutilsMath-0039

Correction of the exercise 504

Le coefficient directeur de la tangente est donné par la valeur de la dérivée en $\pi/2$.

```
sage: f(x)=x*cos(2*x)
sage: (f.diff(x))(pi/2)
-1
```

Si vous voulez voir le détail :

```
sage: f.diff(x)
x |--> -2*x*sin(2*x) + cos(2*x)
```

Notre tangente sera donc de la forme

$$y = -x + b \tag{95.66}$$

et il faut encore trouver la constante b. Pour cela nous savons que la tangente doit passer par le point $(\frac{\pi}{2}, f(\frac{\pi}{2}))$. Donc

$$f\left(\frac{\pi}{2}\right) = -\frac{\pi}{2} + b. \tag{95.67}$$

Étant donné que $f(\pi/2) = -\pi/2$, nous avons $b = 0$ et au final

$$y = -x. \tag{95.68}$$

Exercise 505 exoOutilsMath-0040

Une personne au haut d'une colline de hauteur $h_0 > 0$ lance verticalement une masse avec une vitesse (verticale) $v_0 > 0$. La hauteur de la masse en fonction du temps est donnée par

$$h(t) = h_0 - \frac{gt^2}{2} + v_0 t.$$

(1) Pour quelle valeur de t la hauteur est-elle maximale ?

(2) Quelle est la hauteur maximale atteinte ?

Les réponses peuvent évidemment dépendre de h_0, g et v_0. corrOutilsMath-0040

Correction of the exercise 505

La dérivée de la hauteur en fonction du temps est

$$h'(t) = -gt + v_0, \tag{95.69}$$

cette dérivée s'annule en $t = \frac{v_0}{g}$ qui correspond à un maximum parce que v_0 et g sont positifs (faire un tableau de signe de la dérivée).

La hauteur maximale atteinte est donnée par

$$h\left(\frac{v_0}{g}\right) = h_0 + \frac{v_0^2}{2g}. \tag{95.70}$$

```
sage: var('h0,g,v0,t')
(h0, g, v0, t)
sage: h(t)=h0-g*t**2/2+v0*t
sage: solve(h.diff(t)==0,t)
[t == v0/g]
sage: h(v0/g)
1/2*v0^2/g + h0
```

Notez que la fonction `solve` retourne une *liste* de solutions. Ici la liste se réduit à un seul élément. Si on ne veut pas recopier à la main la solution, on peut faire comme ceci :

```
sage: h( solve(h.diff(t)==0,t)[0].rhs()  )
1/2*v0^2/g + h0
```

Ici nous avons écrit `[0]` pour indiquer que nous voulions la première solution (la numérotation des listes commence à zéro, et non à un !), et ensuite nous avons utilisé `.rhs()` pour demander le membre de droite de la solution.

Exercise 506 exoOutilsMath-0041

Calculer les dérivées directionnelles de la fonction $f(x,y)=x^2-y^2$ dans les directions suivantes :

(1) $u=(0,1)$
(2) $u=(1,0)$
(3) $u=(1,1)/\sqrt{2}$

Pour quels points la dérivée dans la direction $(1,1)$ s'annule ?

corrOutilsMath-0041

Correction of the exercise 506

(1) Commençons par $u=(0,1)$.

$$\begin{aligned}\frac{\partial f}{\partial u}(x,y)&=\lim_{t\to 0}\frac{f(x+0t,y+t)-f(x,y)}{t}\\&=\lim_{t\to 0}\frac{x^2-(y+t)^2-(x^2-y^2)}{t}\\&=\lim_{t\to 0}\frac{-2ty-t^2}{t}\\&=-2y.\end{aligned}\tag{95.71}$$

(2) De la même manière, nous trouvons $\partial_u f(x,y)=2y$ lorsque $u=(0,1)$.
(3) Prenons maintenant $u=(1,1)/\sqrt{2}$. Nous avons

$$\begin{aligned}\frac{\partial f}{\partial u}(x,y)&=\lim_{t\to 0}\frac{f\big(x+\frac{t}{\sqrt{2}},y+\frac{t}{\sqrt{2}}\big)-f(x,y)}{t}\\&=\lim_{t\to 0}\frac{\big(x^2+2\frac{xt}{\sqrt{2}}+\frac{t}{2}\big)-\big(y^2+\frac{2yt}{\sqrt{2}}+\frac{t^2}{2}\big)-x^2+y^2}{t}\\&=\sqrt{2}(x-y).\end{aligned}\tag{95.72}$$

Remarque 95.1.
Cet exemple nous illustre deux principes plus généraux :

(1) La dérivée directionnelle dans la direction $(1,0)$ est exactement la dérivée partielle ∂_x, et la dérivée directionnelle dans la direction $(0,1)$ est exactement la dérivée partielle ∂_y.
(2) La dérivée directionnelle dans une direction (u_1,u_2) est la moyenne pondérée des dérivées partielles.

Exercise 507 exoOutilsMath-0042

Calculer la dérivée directionnelle de $f(x,y)=x^2+y^2$ dans la direction $u=\begin{pmatrix}0\\-1\end{pmatrix}$. Comparer le résultat avec $\partial_y f$.

corrOutilsMath-0042

Correction of the exercise 507

le calcul est habituel :

$$\lim_{t\to 0}\frac{f(x,y-t)-f(x,y)}{t}=-2y.\tag{95.73}$$

Cela est $-\partial_y f(x,y)$.

Encore une fois nous avons trouvé la dérivée directionnelle comme « pondération » par u_1 et u_2 des dérivées partielles. Ici $u_1=0$ et $u_2=-1$.

Exercise 508 exoOutilsMath-0043

Soit la fonction $f(x,y)=xy\cos(x+y)$. Calculer les dérivées partielles ainsi que

$$\frac{\partial f}{\partial u}(\frac{\pi}{5},\frac{2\pi}{15})\tag{95.74}$$

où $u = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$.

corrOutilsMath-0043

Correction of the exercise 508

Les dérivées partielles sont données par

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= y\cos(x+y) - xy\sin(x+y)\\ \frac{\partial f}{\partial y}(x,y) &= x\cos(x+y) - xy\sin(x+y).\end{aligned} \tag{95.75}$$

Étant donne que ce sont des fonctions continues, nous pouvons utiliser la formule de la dérivée directionnelle en termes des dérivées partielles :

$$\frac{\partial f}{\partial u}(a,b) = u_1\frac{\partial f}{\partial x}(a,b) + u_2\frac{\partial f}{\partial y}(a,b). \tag{95.76}$$

Ici $(a,b) = \left(\frac{\pi}{5}, \frac{2\pi}{15}\right)$. Nous avons alors

$$\begin{aligned}\frac{\partial f}{\partial x}(\frac{\pi}{5},\frac{2\pi}{15}) &= \frac{2\pi}{15}\cos(\frac{\pi}{3}) - \frac{2\pi^2}{75}\sin(\frac{\pi}{3}) = \frac{\pi}{15} - \frac{\sqrt{3}\pi^2}{75}\\ \frac{\partial f}{\partial y}(\frac{\pi}{5},\frac{2\pi}{15}) &= \frac{\pi}{10} - \frac{\sqrt{3}\pi^2}{75}.\end{aligned} \tag{95.77}$$

Donc

$$\frac{\partial f}{\partial u}(\frac{\pi}{5},\frac{2\pi}{15}) = \frac{1}{\sqrt{2}}\left[\frac{\pi}{15} - \frac{\sqrt{3}\pi^2}{75} - \frac{\pi}{10} + \frac{\sqrt{3}\pi^2}{75}\right]. \tag{95.78}$$

En calculant, le résultat est $-\frac{\pi}{30\sqrt{2}}$.

Exercise 509 exoOutilsMath-0044

Soit $f(x,y) = x^2e^{\cos(y)}$, et u, le vecteur de direction $\pi/6$. Calculer

$$\frac{\partial f}{\partial u}(1,2) \tag{95.79}$$

corrOutilsMath-0044

Correction of the exercise 509

Les dérivées partielles sont données sont

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2xe^{\cos(y)}\\ \frac{\partial f}{\partial y} &= -x^2\sin(y)e^{\cos(y)}.\end{aligned} \tag{95.80}$$

En ce qui concerne la dérivée directionnelle, il faut d'abord comprendre que la « direction » d'angle $\pi/6$ est le vecteur

$$u = \begin{pmatrix}\cos\frac{\pi}{6}\\ \sin\frac{\pi}{6}\end{pmatrix} \tag{95.81}$$

et par conséquent,

$$\begin{aligned}\frac{\partial f}{\partial u}(x,y) &= u_1\frac{\partial f}{\partial x} + u_2\frac{\partial f}{\partial y}\\ &= \cos\left(\frac{\pi}{6}\right)2xe^{\cos(y)} - \sin\left(\frac{\pi}{6}\right)x^2\sin(y)e^{\cos(y)}.\end{aligned} \tag{95.82}$$

En remplaçant $x = 1$ et $y = 2$, nous trouvons

$$\frac{\partial f}{\partial u}(1,2) = \sqrt{3}e^{\cos(2)} - \frac{1}{2}\sin(2)e^{\cos(2)}. \tag{95.83}$$

Exercise 510 exoOutilsMath-0045

Donner un vecteur simultanément perpendiculaire aux vecteurs $u = (-1, 0, 0)$ et $v = (0, 1, 1)$. corrOutilsMath-0045

Correction of the exercise 510

<+CorrOutilsMath-0045+>

Exercise 511 exoOutilsMath-0046

Donner l'équation du plan tangent à la fonction $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ aux points situés à la verticale de

(1) $(a, b) = (0, 0)$;
(2) $(a, b) = (\frac{R}{2}, \frac{R}{3})$
(3) $(a, b) = (0, R)$. Ici le calcul ne fonctionne pas. Faites les calculs et un dessin pour comprendre pourquoi.

: corrOutilsMath-0046

Correction of the exercise 511

<+CorrOutilsMath-0046+>

Exercise 512 exoOutilsMath-0047

Soit $v = (1, 2, 3)$.

(1) Donner un vecteur non nul perpendiculaire à v ;
(2) Donner un vecteur de norme 1 perpendiculaire à v.

corrOutilsMath-0047

Correction of the exercise 512

Le vecteur (x, y, z) sera perpendiculaire à $(1, 3, 3)$ si (et seulement si) $x + 2y + z = 0$. On peut par exemple prendre

$$w = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}. \tag{95.84}$$

En ce qui concerne un vecteur de norme 1, il suffit de prendre $w/\|w\|$, c'est-à-dire

$$\frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}. \tag{95.85}$$

Il y a évidemment de nombreuses autres possibilités.

Exercise 513 exoOutilsMath-0048

Donner la tangente au graphe de la fonction

$$f(x) = \cos(x) \sin\left(x + \frac{\pi}{4}\right) \tag{95.86}$$

au point d'abscisse $x = \frac{\pi}{2}$. corrOutilsMath-0048

Correction of the exercise 513

La première chose à faire est de calculer la dérivée de f :

$$f'(x) = -\sin(x) \sin\left(x + \frac{\pi}{4}\right) + \cos(x) \cos\left(x + \frac{\pi}{4}\right). \tag{95.87}$$

Le coefficient directeur de la droite que nous cherchons est donc donné par

$$f'\left(\frac{\pi}{2}\right) = -\sin\frac{\pi}{2} \sin\frac{3\pi}{4} + \cos\frac{\pi}{2} \cos\frac{3\pi}{4} = -\frac{\sqrt{2}}{2}. \tag{95.88}$$

L'équation de la droite est donc de la forme

$$y = -\frac{\sqrt{2}}{2}x + b \tag{95.89}$$

où la constante b est encore à fixer. Nous fixons cette constante en imposant

$$y\left(\frac{\pi}{2}\right) = f\left(\frac{\pi}{2}\right). \tag{95.90}$$

Étant donné que $f(\pi/2) = 0$, nous devons poser

$$b = \frac{\sqrt{2}\pi}{4} \tag{95.91}$$

et par conséquent

$$y = -\frac{\sqrt{2}}{2}x + \frac{\sqrt{2}\pi}{4}. \tag{95.92}$$

Exercise 514 exoOutilsMath-0049

Donner en coordonnées cylindriques l'équation du cercle horizontal centré en $(0, 0, 1)$ et de rayon $\sqrt{2}$.

corrOutilsMath-0049

Correction of the exercise 514

Étant donné que le cercle demandé est horizontal de hauteur 1, il est certain qu'il faut imposer $z = 1$. À part cela, il s'agit simplement des équations en coordonnées polaires du cercle de rayon $\sqrt{2}$, c'est-à-dire $r = \sqrt{2}$. Au final les équations demandées sont

$$\begin{cases} z = 1 & \text{(95.93a)} \\ r = \sqrt{2}. & \text{(95.93b)} \end{cases}$$

Exercise 515 exoOutilsMath-0050

Au lancer du poids, si on lance le poids avec une vitesse initiale $v_0 > 0$ et un angle α, alors on atteint une distance

$$d(\alpha) = \frac{2v_0^2}{g}\cos(\alpha)\sin(\alpha). \tag{95.94}$$

(1) Avec quel angle faut-il lancer le poids afin d'obtenir le meilleur résultat possible ?

(2) À quelle distance arrive le poids dans ces conditions ?

Pour rappel, g est une constante positive. Pour des raisons physiques évidentes (faire un dessin de la situation), la réponse doit être un angle $\alpha \in [0, \pi/1]$. Les autres solutions éventuelles doivent être rejetées.

corrOutilsMath-0050

Correction of the exercise 515

L'exercice revient à maximiser la fonction d donnée. Pour ce faire, on dérive :

$$d'(\alpha) = \frac{2v_0^2}{g}\big(\cos^2(\alpha) - \sin^2(\alpha)\big). \tag{95.95}$$

Note : ne pas croire que la parenthèse fait 1. Nous avons $d'(\alpha) = 0$ si

$$\cos(\alpha) = \pm\sin(\alpha). \tag{95.96}$$

Si nous ne considérons que les angles $\alpha \in [0, \frac{\pi}{2}]$, la seule solution est $\alpha = \frac{\pi}{4}$. Afin d'être certain que cet angle donne un maximum (et non un minimum ou un point d'inflexion), il faut faire le tableau de signe de d' pour α entre 0 et $\frac{\pi}{2}$.

Étant donné que pour $\alpha = 0$, nous avons $\cos(\alpha) > \sin(\alpha)$ et que pour $\alpha = \pi/2$ nous avons $\sin(\alpha) > \cos(\alpha)$, nous voyons que $\alpha = \frac{\pi}{4}$ est bien un maximum de d.

La distance atteinte est donc donnée par

$$d\big(\frac{\pi}{4}\big) = \frac{2v_0^2}{g}\cos\frac{\pi}{4}\sin\frac{\pi}{4} = \frac{v_0^2}{g}. \tag{95.97}$$

Exercise 516 exoOutilsMath-0051

Quel est le travail nécessaire pour soulever une masse m à la hauteur h en suivant le chemin $\sigma(t) = \big(\cos(t), \sin(t), t\big)$, $t \in [0, h]$?

corrOutilsMath-0051

Correction of the exercise 516

Le champ de force est donné par

$$F(x, y, z) = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix}, \tag{95.98}$$

et le chemin suivit est donné par

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}. \tag{95.99}$$

La première chose à faire est de calculer la vitesse du chemin (c'est-à-dire sa dérivée) :

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix}. \tag{95.100}$$

Ensuite, la quantité à intégrer est le produit scalaire $F\big(\sigma(t)\big) \cdot \sigma'(t)$:

$$\begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix} = -mg. \tag{95.101}$$

Par conséquent le travail est donné par

$$W = \int_0^h -mg dt = -mgh. \tag{95.102}$$

Exercise 517 exoOutilsMath-0052

Quel est le travail d'un satellite qui fait un demi-tour de la Terre en suivant une orbite de rayon R autour de l'équateur ?

corrOutilsMath-0052

Correction of the exercise 517

Le champ de gravitation est donné par

$$F(x, y, z) = \frac{-1}{(x^2 + y^2 + z^2)^{3/2}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{95.103}$$

et le chemin est

$$\sigma(t) = \begin{pmatrix} R\cos(t) \\ R\sin(t) \\ 0 \end{pmatrix} \tag{95.104}$$

où R est le rayon de l'orbite.

Nous avons

$$\sigma'(t) = \begin{pmatrix} -R\sin(t) \\ R\cos(t) \\ 0 \end{pmatrix} \tag{95.105}$$

et

$$\begin{aligned} F\big(\sigma(t)\big) &= \frac{-1}{\big(R^2\cos^2(t)+R^2\sin^2(t)\big)^{3/2}}\begin{pmatrix} R\cos(t)\\ R\sin(t)\\ 0\end{pmatrix}\\ &= \frac{R}{R^3}\begin{pmatrix}\cos(t)\\ \sin(t)\\ 0\end{pmatrix}\\ &= \frac{1}{R^2}\begin{pmatrix}\cos(t)\\ \sin(t)\\ 0\end{pmatrix}. \end{aligned} \tag{95.106}$$

Le produit scalaire est facile à calculer :

$$F\big(\sigma(t)\big)\cdot\sigma'(t)=0. \tag{95.107}$$

Donc le travail est nul, quel que soit le morceau de trajet parcouru. C'est pour cela que les satellites peuvent tourner indéfiniment sans avoir besoin de propulsion. Si ce n'était pas le cas, la Terre serait déjà tombée sur le Soleil depuis longtemps.

Exercise 518 | exoOutilsMath-0053

Calculer le rotationnel de

$$F(x,y,z)=\begin{pmatrix}\cos(xy)\\ \sin(yz)\\ e^{xz}\end{pmatrix}. \tag{95.108}$$

corrOutilsMath-0053

Correction of the exercise 518

Il s'agit simplement d'utiliser la formule

$$\begin{aligned} \nabla\times F &= \begin{vmatrix} e_x & e_y & e_z\\ \partial_x & \partial_y & \partial_z\\ \cos(xy) & \sin(yz) & e^{xz}\end{vmatrix}\\ &= e_x\big(0-y\cos(yz)\big)-e_y\big(ze^{xz}-0\big)+e_z\big(0+x\sin(xy)\big)\\ &= -y\cos(yz)e_x - ze^{xz}e_y + x\sin(xy)e_z\\ &= \begin{pmatrix}-y\cos(yz)\\ -ze^{xz}\\ x\sin(xy)\end{pmatrix}. \end{aligned} \tag{95.109}$$

Exercise 519 | exoOutilsMath-0054

Calculer les dérivées partielles premières des fonctions

$$\begin{aligned} f(x,y)&=x^2y^2\ln(x^2+y^2)\\ g(x,y)&=xy\sin\frac{1}{x^2+y^2}. \end{aligned} \tag{95.110}$$

corrOutilsMath-0054

Correction of the exercise 519

Ceci est un pur calcul comme un ordinateur peut le faire :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2*y**2*log(x**2+y**2)
sage: f.diff(x)
(x, y) |--> 2*x^3*y^2/(x^2 + y^2) + 2*x*y^2*log(x^2 + y^2)
```

```
sage: f.diff(y)
(x, y) |--> 2*x^2*y^3/(x^2 + y^2) + 2*x^2*y*log(x^2 + y^2)

sage: g(x,y)=x*y*sin(1/(x**2+y**2))
sage: g.diff(x)
(x, y) |--> -2*x^2*y*cos(1/(x^2 + y^2))/(x^2 + y^2)^2 + y*sin(1/(x^2 + y^2))
sage: g.diff(y)
(x, y) |--> -2*x*y^2*cos(1/(x^2 + y^2))/(x^2 + y^2)^2 + x*sin(1/(x^2 + y^2))
```

Exercise 520 exoOutilsMath-0055

Quelle est la différentielle de la fonction

$$f(x,y,z) = e^x \cos(yz) + xyz \tag{95.111}$$

au point $(0, \pi/2, 1)$?

Calculer la dérivée directionnelle de f en ce point dans la direction du vecteur $(1,1,1)$. corOutilsMath-0055

Correction of the exercise 520

La première chose à faire est de calculer les dérivées partielles et de les évaluer au point donné :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=exp(x)*cos(y*z)+x*y*z
sage: f.diff(x)
(x, y, z) |--> y*z + e^x*cos(y*z)
sage: f.diff(y)
(x, y, z) |--> -z*e^x*sin(y*z) + x*z
sage: f.diff(z)
(x, y, z) |--> -y*e^x*sin(y*z) + x*y
sage: f.diff(x)(0,pi/2,1)
1/2*pi
sage: f.diff(y)(0,pi/2,1)
-1
sage: f.diff(z)(0,pi/2,1)
-1/2*pi
```

Nous avons donc

$$\begin{aligned}
\frac{\partial f}{\partial x}(0,\frac{\pi}{2},1) &= \frac{\pi}{2} \\
\frac{\partial f}{\partial y}(0,\frac{\pi}{2},1) &= -1 \\
\frac{\partial f}{\partial z}(0,\frac{\pi}{2},1) &= -\frac{\pi}{2}
\end{aligned} \tag{95.112}$$

Notez que lorsque Sage écrit `1/2*pi` il veut dire $\frac{1}{2}\pi$ et non $\frac{1}{2\pi}$.

Maintenant nous utilisons les super formules qui lient les dérivées partielles à la différentielle et aux dérivées directionnelles :

$$\begin{aligned}
\frac{\partial f}{\partial u}(a,b,c) &= \nabla f(a,b,c) \cdot u \\
&= u_1 \frac{\partial f}{\partial x}(a,b,c) + u_2 \frac{\partial f}{\partial y}(a,b,c) + u_3 \frac{\partial f}{\partial z}(a,b,c) \\
&= df_{(a,b,c)}(u).
\end{aligned} \tag{95.113}$$

Quelques remarques sur ces formules :

(1) La première égalité tient uniquement pour $\|u\| = 1$. Les trois autres expressions sont par contre égales pour n'importe quel vecteur.
(2) En termes de calculs, c'est évidemment l'expression du milieu qu'on utilise.
(3) Elles sont importantes à retenir.

En ce qui concerne la différentielle nous avons

$$df_{(0,\frac{\pi}{2},1)}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix} = \frac{\pi}{2}u_1 - u_2 - \frac{\pi}{2}u_3. \tag{95.114}$$ Eqormzzccdfabc

En ce qui concerne la dérivée directionnelle dans la direction du vecteur $(1,1,1)$, il faut prendre $u = (1,1,1)/\sqrt{3}$ parce que nous voulons un vecteur de norme 1. Donc nous utilisons (95.114) avec $u_1 = u_2 = u_2 = 1/\sqrt{3}$:

$$\frac{\partial f}{\partial u}(0,\frac{\pi}{2},1) = df_{(0,\frac{\pi}{2},1)}\begin{pmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix} = \frac{-1}{\sqrt{3}}. \tag{95.115}$$

Exercise 521 exoOutilsMath-0056

Donner une approximation de la fonction

$$f(x,y,z) = xe^x + ye^y + ze^z \tag{95.116}$$

au point $(1.0001; 1.999; 2.98)$.

Vous pouvez garder e, e^2 et e^3 dans la réponse. corrOutilsMath-0056

Correction of the exercise 521

Nous utilisons la formule d'approximation (12.670) avec trois variables au lieu de deux. Nous avons l'approximation

$$\begin{aligned} f(1.0001, 1.999, 2.98) = f(1,2,3) &+ 0.0001\frac{\partial f}{\partial x}(1,2,3) \\ &- 0.0001\frac{\partial f}{\partial y}(1,2,3) \\ &- 0.02\frac{\partial f}{\partial z}(1,2,3). \end{aligned} \tag{95.117}$$

Les dérivées partielles se calculent facilement

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: a(x)=x*exp(x)
sage: f(x,y,z)=a(x)+a(y)+a(z)
sage: f
(x, y, z) |--> x*e^x + y*e^y + z*e^z
sage: f(1,2,3)
e + 2*e^2 + 3*e^3
sage: f.diff(x)(1,2,3)
2*e
sage: f.diff(y)(1,2,3)
3*e^2
sage: f.diff(z)(1,2,3)
4*e^3
```

Notez la petite astuce de définir `a(x)=x*exp(x)` pour ne pas devoir taper trois fois la même expression. Nous avons donc comme approximation :

$$e+2e^2+3e^3+\frac{2e}{10000}-\frac{3e^2}{10000}-\frac{2}{100}4e^3=\frac{5001}{5000}e+\frac{19997}{10000}e^2+\frac{77}{25}e^3. \tag{95.118}$$

Exercise 522 exoOutilsMath-0057

Donner l'équation du plan tangent au graphe de la fonction

$$f(x,y)=x^2-y^2 \tag{95.119}$$

au point $(2,1,3)$. Trouver un vecteur orthogonal à ce plan. corrOutilsMath-0057

Correction of the exercise 522

Les dérivées partielles de f au point $(2,1)$ sont calculables comme d'habitude :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2-y**2
sage: f(2,1)
3
sage: f.diff(x)(2,1)
4
sage: f.diff(y)(2,1)
-2
```

Nous utilisons la formule du plan tangent

$$\begin{aligned}T_{(2,1)}(x,y)&=f(2,1)+\frac{\partial f}{\partial x}(2,1)(x-2)+\frac{\partial f}{\partial y}(2,1)(y-1)\\&=3+4(x-2)+3(y-1)\\&=4x+3y-12.\end{aligned} \tag{95.120}$$

Pour trouver une vecteur perpendiculaire à ce plan, il suffit de trouver deux vecteurs parallèles à ce plan et puis de trouver une vecteur qui est simultanément perpendiculaire à ces deux vecteurs.

Le plan passant par l'origine et étant parallèle à notre plan tangent est le plan d'équation

$$z=4x+3y. \tag{95.121}$$

Il possède par exemple les deux vecteurs suivants :

$$v_1=\begin{pmatrix}1\\0\\4\end{pmatrix}\quad v_2=\begin{pmatrix}0\\1\\3\end{pmatrix}. \tag{95.122}$$

Un vecteur (x,y,z) simultanément perpendiculaire à v_1 et à v_2 doit vérifier

$$\begin{cases}x+4z=0 & \text{(95.123a)}\\ y+3z=0. & \text{(95.123b)}\end{cases}$$

Vu qu'il y a moins d'équations que d'inconnues, on peut par exemple choisir $x=1$ et puis les équations fixent $z=-1/4$ et $y=3/4$. Par conséquent un vecteur perpendiculaire au plan est

$$\begin{pmatrix}1\\-1/4\\3/4\end{pmatrix}. \tag{95.124}$$

Exercise 523 exoOutilsMath-0058

Pour chacun des champs de vecteurs suivants,

(1) calculer le rotationnel

(2) donner un potentiel scalaire.

$$\begin{aligned} G(x,y) &= \begin{pmatrix} 3x^2y \\ x^3+y^3 \end{pmatrix}, \\ H(x,y,z) &= \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \\ F(x,y) &= \begin{pmatrix} y\cos(xy)e^{xy} + \sin(xy)ye^{xy} + 1 \\ x\cos(xy)e^{xy} + \sin(xy)xe^{xy} \end{pmatrix} \end{aligned} \tag{95.125}$$

Les champs F et G sont vus dans $\mathbb{R}^3$ avec zéro comme troisième composante. corrOutilsMath-0058

Correction of the exercise 523

Tous les rotationnels sont nuls (sinon on ne demanderait pas un potentiel scalaire). En ce qui concerne les potentiels,

(1) $V_G = x^3y + y^4/4$,

(2) $V_H = x^2 + y^2 + z^2$;

(3) $V_F = \sin(xy) + e^{xy} + x$.

Exercise 524 exoOutilsMath-0059

Montrer que le champ de vecteurs

$$F(x,y) = \begin{pmatrix} x^2\sin(xy) \\ y^2 + x\sin(y^2) \\ 0 \end{pmatrix} \tag{95.126}$$

n'est pas un champ de gradient (c'est-à-dire qu'il ne dérive pas d'un potentiel). corrOutilsMath-0059

Correction of the exercise 524

Si F était le gradient de la fonction f, alors le rotationnel de F serait nul : $\nabla \times (\nabla f) = 0$. Calculons le rotationnel de F :

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ x^2\sin(xy) & y^2 + x\sin(y^2) & 0 \end{vmatrix} = \big(\sin(y^2) - x^3\cos(xy)\big)e_z. \tag{95.127}$$

Cela n'est donc pas nul et nous en déduisons qu'il n'existe pas de fonctions f telles que $F = \nabla f$.

Exercise 525 exoOutilsMath-0060

Calculer la circulation du champ de vecteurs $F(x,y) = (x^2, xy)$ le long du carré de sommets $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$ parcouru dans le sens direct (le sens trigonométrique). corrOutilsMath-0060

Correction of the exercise 525

La difficulté de cet exercice est de paramétrer le chemin. Cela se fait en quatre morceaux :

$$\begin{aligned} \sigma_1(t) &= (t,0) \\ \sigma_2(t) &= (1,t) \\ \sigma_3(t) &= (1-t,0) \\ \sigma_4(t) &= (0,1-t). \end{aligned} \tag{95.128}$$

L'intégrale se fait donc en quatre parties. La première est

$$\begin{aligned} I_1 &= \int_0^1 F(t,0)\cdot\begin{pmatrix}1\\0\end{pmatrix}dt \\ &= \int_0^1 \begin{pmatrix}t^2\\0\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}dt \\ &= \int_0^1 t^2dt \\ &= \frac{1}{3}. \end{aligned} \tag{95.129}$$

La seconde :

$$\begin{aligned} I_2 &= \int_0^1 F(1,t)\cdot\begin{pmatrix}0\\1\end{pmatrix}dt \\ &= \int_0^1 tdt \\ &= \frac{1}{2}. \end{aligned} \tag{95.130}$$

La troisième :

$$I_3 = \int_0^1 F(1-t,0)\cdot\begin{pmatrix}-1\\0\end{pmatrix} = -\int_0^1 (1-t)^2dt = -\frac{1}{3}. \tag{95.131}$$

Et la quatrième :

$$I_4 = \int_0^1 F(0,1-t)\cdot\begin{pmatrix}0\\-1\end{pmatrix}dt = 0. \tag{95.132}$$

Au final, l'intégrale cherchée vaut

$$\frac{1}{3}+\frac{1}{2}-\frac{1}{3}=\frac{1}{2}. \tag{95.133}$$

Exercise 526 exoOutilsMath-0061

Calculer la longueur de la courbe paramétrée

$$\begin{aligned} \sigma\colon [1,2] &\to \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix}t\\ \ln(t)\\ 2\sqrt{2t}\end{pmatrix}. \end{aligned} \tag{95.134}$$

corrOutilsMath-0061

Correction of the exercise 526

Il faut intégrer la norme du vecteur tangent. Ici c'est

$$\sigma'(t) = \begin{pmatrix}1\\ 1/t\\ \sqrt{2}/\sqrt{t}\end{pmatrix}, \tag{95.135}$$

et donc

$$\|\sigma'(t)\| = \sqrt{\frac{t^2+1+2t}{t^2}} = \frac{t+1}{t}. \tag{95.136}$$

Nous avons alors

$$l(\sigma) = \int_1^2 \frac{t+1}{t}dt = 1+\ln(2). \tag{95.137}$$

Exercise 527 exoOutilsMath-0062

Soit

$$V(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}. \tag{95.138}$$

(1) Déterminer $F = \nabla V$.
(2) Calculer la circulation de F le long du segment de droite joignant les points $(1, 2, 3)$ et $(3, 4, 5)$.
(3) Si σ est un chemin général joignant ces deux mêmes points, que vaut $\int_\sigma F$?

corrOutilsMath-0062

Correction of the exercise 527

Le gradient de V est le champ de vecteurs

$$F(x, y, z) = \frac{-1}{(x^2 + y^2 + z^2)^{3/2}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{95.139}$$

Il n'est toutefois pas utile de le savoir pour calculer la circulation de F étant donné que nous connaissons le potentiel.

Quel que soit le chemin σ reliant le point a au point b, nous aurons $\int_\sigma F = V(b) - V(a)$. Ici nous avons donc

$$\int_\sigma F = V(3, 4, 5) - V(1, 2, 3) = \frac{1}{\sqrt{50}} - \frac{1}{\sqrt{14}}. \tag{95.140}$$

Exercise 528 exoOutilsMath-0063

Calculer la circulation de

$$F(x, y, z) = \begin{pmatrix} \cos(z) \\ e^x \\ e^y \end{pmatrix} \tag{95.141}$$

le long du chemin $\sigma : [0, 2] \to \begin{pmatrix} 1 \\ t \\ e^t \end{pmatrix}$.

corrOutilsMath-0063

Correction of the exercise 528

Nous avons

$$F\big(\sigma(t)\big) = F(1, t, e^t) = \begin{pmatrix} \cos(e^t) \\ e^1 \\ e^t \end{pmatrix}, \tag{95.142}$$

et

$$\sigma'(t) = \begin{pmatrix} 0 \\ 1 \\ e^t \end{pmatrix}, \tag{95.143}$$

et le produit scalaire est

$$F\big(\sigma(t)\big) \cdot \sigma'(t) = e + e^{2t}. \tag{95.144}$$

Par conséquent la circulation que nous cherchons vaut

$$I = \int_0^2 (e + e^{2t}) dt = 2e + \left[\frac{1}{2} e^{2t}\right]_0^2 = 2e + \frac{1}{2}(e^4 - 1). \tag{95.145}$$

Exercise 529 exoOutilsMath-0064

Calculer la circulation du champ

$$F(x, y, z) = \begin{pmatrix} x^2 \\ xy \\ 1 \end{pmatrix} \tag{95.146}$$

le long de $\sigma(t) = (t, t^2, 1)$ pour $t \in [0, 1]$.

corrOutilsMath-0064

Correction of the exercise 529

Nous avons les trois choses usuelles à calculer :

$$\begin{aligned} F\big(\sigma(t)\big) = \begin{pmatrix} t^2 \\ t^3 \\ 1 \end{pmatrix}, \sigma'(t) &= \begin{pmatrix} 1 \\ 2t \\ 0 \end{pmatrix} \\ F\big(\sigma(t)\big) \cdot \sigma'(t) &= t^2 + 2t^4. \end{aligned} \tag{95.147}$$

L'intégrale à calculer est donc

$$\int_0^1 (t^2 + 2t^4) = \frac{11}{15}. \tag{95.148}$$

Exercise 530 exoOutilsMath-0065

Soient $F(x, y, z) = -ye_x + xe_y$ et

$$G(x, y, z) = \frac{F(x, y, z)}{x^2 + y^2}. \tag{95.149}$$

Calculer le rotationnel et la divergence des champs F et G. corrOutilsMath-0065

Correction of the exercise 530

Le champ de vecteurs F est

$$F(x, y, z) = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}, \tag{95.150}$$

sa divergence est nulle, et son rotationnel est

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ -y & x & 0 \end{vmatrix} = 2e_z. \tag{95.151}$$

En ce qui concerne G, la divergence est également nulle et

$$\nabla \times G = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ -\frac{u}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \end{vmatrix} = 0 \tag{95.152}$$

Exercise 531 exoOutilsMath-0066

Soit $f\colon \mathbb{R}^3 \to \mathbb{R}$ une fonction de classe C^1, et un champ de vecteurs $F\colon \mathbb{R}^3 \to \mathbb{R}^3$. Vérifier que

$$\nabla \cdot (fF) = f\nabla \cdot F + F \cdot \nabla f. \tag{95.153}$$

corrOutilsMath-0066

Correction of the exercise 531

<+CorrOutilsMath-0066+>

Exercise 532 exoOutilsMath-0067

Soit le champ de vecteurs

$$r(x, y, z) = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{95.154}$$

Calculer $\nabla(\|r\|)$ et $\nabla \cdot (\|r\|r)$. corrOutilsMath-0067

Correction of the exercise 532

Nous avons $\|r\| = \sqrt{x^2 + y^2 + z^2}$. Par conséquent,

$$\nabla\big(\|r\|\big) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2+z^2}} \\ \frac{y}{\sqrt{x^2+y^2+z^2}} \\ \frac{z}{\sqrt{x^2+y^2+z^2}} \end{pmatrix} = \frac{1}{\|r\|} r \tag{95.155}$$

D'autre part,

$$\|r\|r = \begin{pmatrix} \sqrt{x^2+y^2+z^2}x \\ \sqrt{x^2+y^2+z^2}y \\ \sqrt{x^2+y^2+z^2}z \end{pmatrix}, \tag{95.156}$$

et les calculs montrent que

$$\begin{aligned} \nabla \cdot (\|r\|r) &= \frac{2x^2+y^2+z^2}{\sqrt{x^2+y^2+z^2}} + \frac{x^2+2y^2+z^2}{\sqrt{x^2+y^2+z^2}} + \frac{x^2+y^2+2z^2}{\sqrt{x^2+y^2+z^2}} \\ &= \frac{4(x^2+y^2+z^2)}{\sqrt{x^2+y^2+z^2}} \\ &= 4\sqrt{x^2+y^2+z^2} \\ &= 4\|r\|. \end{aligned} \tag{95.157}$$

Exercise 533 exoOutilsMath-0068

Soit

$$F(x,y) = \begin{pmatrix} 2x^2y+y \\ x+2xy^2 \end{pmatrix} e^{x^2+y^2}. \tag{95.158}$$

(1) Montrer que $\nabla \times F = 0$.

(2) Montrer que F dérive d'un potentiel scalaire.

(3) Calculer la circulation de F le long du demi-cercle unité centré en $(0,0)$ parcouru de $(1,0)$ à $(-1,0)$ dans le sens trigonométrique.

corrOutilsMath-0068

Correction of the exercise 533

(1) Ce point est juste un calcul. Tout se simplifie.

(2) Nous devons avoir en même temps

$$\frac{\partial V}{\partial x}(x,y,z) = y(2x^2+1)e^{x^2+y^2} \tag{95.159}$$

et

$$\frac{\partial V}{\partial y}(x,y,z) = x(1+2y^2)e^{x^2+y^2}. \tag{95.160}$$

Cherchons un potentiel sous la forme

$$V(x,y) = f(x,y)e^{x^2+y^2}. \tag{95.161}$$

Nous avons

$$\frac{\partial V}{\partial x} = \frac{\partial f}{\partial x}e^{x^2+y^2} + 2xfe^{x^2+y^2}. \tag{95.162}$$

Nous devons donc trouver f telle que

$$\frac{\partial f}{\partial x} + 2xf = y(2x^2+1). \tag{95.163}$$

La fonction $f(x,y) = xy$ fonctionne.
Il suffit donc de vérifier que le potentiel

$$V(x,y,z) = xye^{x^2+y^2} \tag{95.164}$$

est bien tel que $\nabla V = F$.

(3) Nous avons $V(1,0) = V(-1,0) = 0$, par conséquent, quel que soit le chemin les reliant, nous avons toujours que la circulation le long d'un chemin joignant $(1,0)$ à $(-1,0)$ est nulle.

Exercise 534 exoOutilsMath-0069

Trouver un vecteur simultanément orthogonal à $\begin{pmatrix}1\\1\\0\end{pmatrix}$ et à $\begin{pmatrix}0\\1\\1\end{pmatrix}$.

corrOutilsMath-0069

Correction of the exercise 534

<+0069+>

Exercise 535 exoOutilsMath-0070

Trouver l'aire du triangle dont les sommets sont les points $(0,0)$, $(2,1)$ et $(-1,1)$.

corrOutilsMath-0070

Correction of the exercise 535

L'aire du triangle est le moitié de celle du parallélogramme construit sur les deux vecteurs $v=(-1,1)$ et $v=(2,1)$. Cette dernière aire vaut

$$\det\begin{pmatrix}2&1\\-1&1\end{pmatrix}=2-(-1)=3. \tag{95.165}$$

Donc l'aire du triangle vaut $\frac{3}{2}$.

Exercise 536 exoOutilsMath-0071

Déterminer l'équation du plan passant par le point $\begin{pmatrix}1\\4\\7\end{pmatrix}$ et orthogonal au vecteur $\begin{pmatrix}2\\1\\5\end{pmatrix}$.

corrOutilsMath-0071

Correction of the exercise 536

Trouvons d'abord le plan passant par $(0,0,0)$ qui est perpendiculaire au vecteur donné. Ce la est $z=ax+by$. Nous cherchons donc a et b tels que

$$\begin{pmatrix}x\\y\\ax+by\end{pmatrix}\cdot\begin{pmatrix}2\\1\\5\end{pmatrix}=0 \tag{95.166}$$

pour tout x et y. L'équation est

$$(2+5a)x+(1+5b)y=0. \tag{95.167}$$

Nous devons donc avoir $a=-2/5$ et $b=-1/5$.

Maintenant nous cherchons le plan de la forme

$$z=-\frac{2}{5}x-\frac{1}{5}y+c \tag{95.168}$$

qui passe par le point $(1,4,7)$. Il faut donc

$$7=-\frac{2}{5}1-\frac{1}{5}4+c, \tag{95.169}$$

ce qui donne $c=\frac{41}{5}$.

Exercise 537 exoOutilsMath-0072

Déterminer l'équation du plan passant par le point $\begin{pmatrix}1\\2\\3\end{pmatrix}$ et parallèle aux vecteurs $\begin{pmatrix}2\\6\\1\end{pmatrix}$ et $\begin{pmatrix}1\\4\\7\end{pmatrix}$.

corrOutilsMath-0072

Correction of the exercise 537

<+CorrOutilsMath-0072+>

Exercise 538 exoOutilsMath-0073

Calculer le volume du parallélépipède formé par les vecteurs e_x, $3e_y-e_z$ et $4e_x+2e_y-e_z$.

corrOutilsMath-0073

Correction of the exercise 538

Le volume est donné par la valeur absolue du déterminant construit sur les trois vecteurs :

$$\det\begin{pmatrix}1&0&0\\0&3&-1\\4&2&-1\end{pmatrix}=\det\begin{pmatrix}3&-1\\2&-1\end{pmatrix}=-3-(-2)=-1. \tag{95.170}$$

Le volume est donc $V=1$.

Exercise 539 exoOutilsMath-0074

Trouver l'équation du plan passant par les points

$$a=\begin{pmatrix}1\\2\\0\end{pmatrix}\quad b=\begin{pmatrix}0\\1\\-2\end{pmatrix}\quad c=\begin{pmatrix}4\\0\\1\end{pmatrix}. \tag{95.171}$$

corrOutilsMath-0074

Correction of the exercise 539

Nous commençons par trouver le plan passant par $(0,0,0)$ parallèle aux vecteurs $b-a$ et $c-a$. Les vecteurs

$$\begin{pmatrix}1\\1\\2\end{pmatrix},\quad\begin{pmatrix}3\\-2\\1\end{pmatrix} \tag{95.172}$$

doivent appartenir à $z=ax+by$. Le système est

$$\begin{cases}2=a+b & \text{(95.173a)}\\ 1=3a-2b, & \text{(95.173b)}\end{cases}$$

et la solution est le plan $z=x+y$.

Maintenant nous trouvons le plan $z=x+y+c$ qui passe par $(4,0,1)$. La solution est $c=-3$. Le plan cherché est donc

$$z=x+y-3. \tag{95.174}$$

Pour vérification, si $f(x,y)=x+y-3$, nous avons bien $f(1,2)=0$, $f(0,1)=-2$ et $f(4,0)=1$.

Exercise 540 exoOutilsMath-0075

En mécanique classique, le **moment** M d'une force F par rapport à un point O est le produit de la forme $\|F\|$ par la distance d entre O et la ligne qui porte F, voir figure 95.4.

Le vecteur moment $\overline{M}$ est le vecteur de taille M et orthogonal au plan défini par O et par F.

Vérifier que $\overline{M}=R\times F$ où R est le vecteur qui lie O à l'origine de F, voir figure

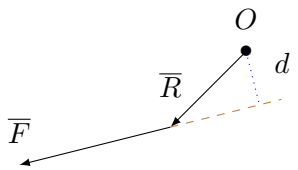


FIGURE 95.4: Moment de force.

LabelFigMomentForce

corrOutilsMath-0075

Correction of the exercise 540

En ce qui concerne la direction, c'est correct : c'est perpendiculaire au plan qui contient les vecteurs R et f. En ce qui concerne la norme, il faut vérifier que

$$\|F\times R\|=\|F\|d. \tag{95.175}$$

En regardant le dessin, on voit immédiatement que $\|F\|d$ est bien la surface du parallélogramme construit sur F et R.

Remarque 95.2.
En réalité pour la direction il reste une ambigüité parce que nous n'avons pas fixé le sens. Le moment pourrait autant être $R \times F$ que $F \times R$.

Exercise 541 exoOutilsMath-0076

Soient deux milieux d'indice de réfraction n_1 et n_2 et séparés par une surface plane perpendiculaire à un vecteur unitaire $\overline{N}$. D'après la loi de Snell, on a

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{n_2}{n_1} \tag{95.176}$$

où θ_1 et θ_2 sont les angles d'incidence (par rapport à $\overline{N}$) d'un rayon lumineux et son angle de réfraction.

Vérifier que

$$n_1(\overline{N} \times a) = n_2(\overline{N} \times b) \tag{95.177}$$

où a et b désignent deux vecteurs unitaires le long des rayons d'incidence et de réfraction.

Voir la situation sur la figure 95.5.

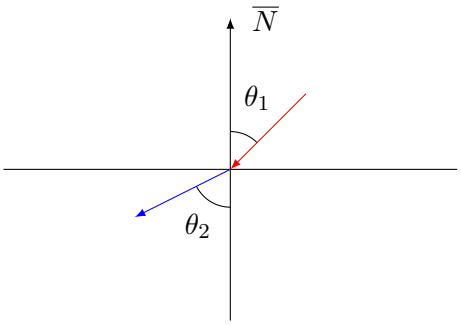


FIGURE 95.5: Angles de réfraction et de réflexion. LabelFigRefraction

corrOutilsMath-0076

Correction of the exercise 541

Les vecteurs $N \times a$ et $N \times b$ sont dans le même sens. Il faut voir leur normes. En utilisant la formule de la proposition 18.55, la norme de $N \times a$ est $\sin(\theta_1)$ parce que N et a sont de norme 1. Nous avons donc

$$\begin{cases} \|N \times a\| = \sin(\theta_1) & \text{(95.178a)} \\ \|N \times b\| = \sin(\theta_2) & \text{(95.178b)} \end{cases}$$

La loi de Snell montre alors que $\|N \times a\| n_1 = \|N \times b\| n_2$.

Exercise 542 exoOutilsMath-0077

Soit $g(x, y, z) = (x + y)\sqrt{x^2 + y^2 + z^2}$. Considérons

$$\tilde{g}(\rho, \theta, \varphi) = (g \circ f)(\rho, \theta, \varphi) \tag{95.179}$$

où f est le changement de coordonnées sphériques :

$$f \begin{pmatrix} \rho \\ \theta \\ \varphi \end{pmatrix} = \begin{pmatrix} \rho \sin(\theta) \cos(\varphi) \\ \rho \sin(\theta) \sin(\varphi) \\ \rho \cos(\theta) \end{pmatrix}. \tag{95.180}$$

Déterminer les fonctions

$$\frac{\partial \tilde{g}}{\partial \rho}, \quad \frac{\partial \tilde{g}}{\partial \theta}, \quad \frac{\partial \tilde{g}}{\partial \varphi}. \tag{95.181}$$

corrOutilsMath-0077

Correction of the exercise 542

La fonction $\tilde{g}$ est donnée par

$$\tilde{g}(\rho,\theta,\varphi)=\rho^2\sin\theta\big(\cos\varphi+\sin\varphi\big). \tag{95.182}$$

Les dérivées partielles sont

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial\rho}&=2\rho\sin(\theta)\big(\cos(\varphi)+\sin(\varphi)\big)\\ \frac{\partial\tilde{g}}{\partial\theta}&=\rho^2\cos(\theta)\big(\cos(\varphi)+\sin(\varphi)\big)\\ \frac{\partial\tilde{g}}{\partial\varphi}&=\rho^2\sin(\theta)\big(\cos(\varphi)-\sin(\varphi)\big)\end{aligned} \tag{95.183}$$

Exercise 543 exoOutilsMath-0078

Prouver que les coordonnées cylindriques forment un système curviligne orthogonal. exoOutilsMath-0078

Correction of the exercise 543

Les coordonnées cylindriques sont données par le changement de variables

$$M(r,\theta,z)=\begin{pmatrix}r\cos(\theta)\\ r\sin(\theta)\\ z\end{pmatrix}. \tag{95.184}$$

Les dérivées de M par rapport aux coordonnées cylindriques sont

$$\begin{aligned}\frac{\partial M}{\partial r}&=\frac{\partial x}{\partial r}e_x+\frac{\partial y}{\partial r}e_y+\frac{\partial z}{\partial r}e_z\\ &=\cos(\theta)e_x+\sin(\theta)e_y\\ \frac{\partial M}{\partial\theta}&=\frac{\partial x}{\partial\theta}e_x+\frac{\partial\theta}{\partial r}e_y+\frac{\partial z}{\partial\theta}e_z\\ &=-r\sin(\theta)e_x+r\cos(\theta)e_y\\ \frac{\partial M}{\partial z}&=\frac{\partial x}{\partial z}e_x+\frac{\partial y}{\partial z}e_y+\frac{\partial z}{\partial z}e_z\\ &=e_z.\end{aligned} \tag{95.185}$$

Les vecteurs de la base locale des coordonnées cylindriques sont donc donnés en normalisant ces trois vecteurs :

$$\begin{aligned}e_r&=\cos(\theta)e_x+\sin(\theta)e_y\\ e_\theta&=-\sin(\theta)e_x+\cos(\theta)e_y\\ e_z&=e_z\end{aligned} \tag{95.186}$$

Le fait que cette base locale soit orthogonale signifie que ces trois derniers vecteurs sont orthogonaux. Vérification :

$$\begin{aligned}e_r\cdot e_\theta&=\begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix}\cdot\begin{pmatrix}-\sin\theta\\ \cos\theta\\ 0\end{pmatrix}=0\\ e_r\cdot e_z&=\begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix}\cdot\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}=0\\ e_\theta\cdot e_z&=\begin{pmatrix}-\sin\theta\\ \cos\theta\\ 0\end{pmatrix}\cdot\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}=0.\end{aligned} \tag{95.187}$$

Exercise 544 exoOutilsMath-0079

Un fil électrique rectiligne suivant l'axe z est parcouru d'un courant électrique Ie_z où I est une constante. Soit $p\in\mathbb{R}^3$. Le valeur du champ magnétique en p est donnée par

$$B(p)=\frac{I\times d}{\|d\|^2} \tag{95.188}$$

où d est le vecteur qui relie le point p au fil.

(1) Faire un dessin de la situation ; indiquer le vecteur d.
(2) Calculer le la divergence et le rotationnel de B.

corrOutilsMath-0079

Correction of the exercise 544

Le dessin est sur la figure 95.6.

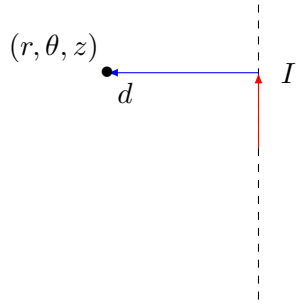


FIGURE 95.6: La façon naturelle de décrire la situation sont les coordonnées cylindriques. LaberFigExMagnetique

Écrire la situation en coordonnées cylindriques est relativement facile. D'abord le vecteur courant est Ie_z. Le vecteur d est un multiple de e_r parce que ce dernier est exactement fait pour pointer vers l'axe vertical. En ce qui concerne le multiple à fixer, la longueur de d est r (par définition de la coordonnées cylindrique r) et sa direction est du fil vers le point, c'est-à-dire dans la même direction que e_r. Nous avons donc

$$d = re_r. \tag{95.189}$$

Nous avons par conséquent

$$B(r,\theta,z) = \frac{re_r \times e_z}{r^2} = \frac{e_\theta}{r}. \tag{95.190}$$

N'oubliez pas de vérifier pourquoi $e_r \times e_z = e_\theta$. Calculer la divergence et le rotationnel de ce champ se fait en utilisant les formules (24.123) et (24.136) en posant $B_r = 0$, $B_\theta = -\frac{1}{r}$ et $B_z = 0$. Nous trouvons

$$\begin{aligned} \nabla \cdot B &= 0 \\ \nabla \times B &= 0. \end{aligned} \tag{95.191}$$

Remarque 95.3.
La divergence d'un champ magnétique est *toujours* nulle. Le fait que le rotationnel soit nul par contre est juste une propriété du cas particulier que nous regardons ici.

Exercise 545 exoOutilsMath-0080

Soit le champ de vecteurs donné en coordonnées polaires suivant :

$$F(r,\theta) = r\sin(\theta)e_r + 2e_\theta. \tag{95.192}$$

Exprimer ce champ en coordonnées cartésiennes.

corrOutilsMath-0080

Correction of the exercise 545

La difficulté est d'exprimer les vecteurs de la base locale $\{e_r, e_\theta\}$ en termes des vecteurs de la base cartésienne $\{e_x, e_y\}$. Nous savons que les vecteurs de la base locale des coordonnées polaires sont donnés par

$$\begin{cases} e_r = \cos(\theta)e_x + \sin(\theta)e_y & \text{(95.193a)} \\ e_\theta = -\sin(\theta)e_x + \cos(\theta)e_y. & \text{(95.193b)} \end{cases}$$

En utilisant les formules $x = r\cos(\theta)$ et $y = r\sin(\theta)$, nous avons immédiatement

$$\cos(\theta) = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}} \tag{95.194}$$

et

$$\sin(\theta) = \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}}. \tag{95.195}$$

Par conséquent

$$\begin{cases} e_r = \frac{1}{\sqrt{x^2+y^2}}(xe_x + ye_y) & \text{(95.196a)} \\ e_\theta = \frac{1}{\sqrt{x^2+y^2}}(-ye_x + xe_y). & \text{(95.196b)} \end{cases}$$

En remplaçant cela dans la formule de F,

$$\begin{aligned} F(x,y) &= ye_r + 2e_\theta \\ &= \frac{y}{\sqrt{x^2+y^2}}(xe_x + ye_y) + 2\frac{1}{\sqrt{x^2+y^2}}(-ye_x + xe_y) \\ &= \frac{1}{\sqrt{x^2+y^2}}\big((yx-2y)e_x + (y^2+2x)e_y\big). \end{aligned} \tag{95.197}$$

Exercise 546 exoOutilsMath-0081

Soit la fonction

$$f(x,y,z) = \frac{z}{x^2+y^2}. \tag{95.198}$$

(1) Écrire cette fonction en coordonnées cylindriques.

(2) Calculer le gradient de f.

(3) Si G est le gradient de f, que vaut $\nabla \times G$?

corrOutilsMath-0081

Correction of the exercise 546

Le passage en cylindriques se fait en remplaçant x^2+y^2 par r :

$$\tilde{f}(r,\theta,z) = \frac{z}{r}. \tag{95.199}$$

Le gradient est alors donné par

$$\nabla\tilde{f} = -\frac{z}{r}e_r + \frac{1}{r}e_z. \tag{95.200}$$

Le rotationnel de ce gradient est nul parce que le rotationnel d'un gradient est toujours nul.

Exercise 547 exoOutilsMath-0082

Trouver les fonctions **harmoniques** radiales. C'est-à-dire les fonctions f ne dépendant que de ρ et telles que $\Delta f = 0$.

corrOutilsMath-0082

Correction of the exercise 547

Étant donné que f ne dépend que de ρ, la formule (24.130) du laplacien se réduit à

$$\Delta f = \frac{1}{\rho^2\sin\theta}\frac{\partial}{\partial\rho}\left(\rho^2\sin\theta\frac{\partial f}{\partial\rho}\right). \tag{95.201}$$

Demander que cela soit nul revient à poser l'équation

$$2\frac{\partial f}{\partial\rho} + \rho\frac{\partial^2 f}{\partial\rho^2} = 0. \tag{95.202}$$

Histoire de retomber sur des notations habituelles, nous cherchons une fonction $f(x)$ telle que $2f'(x) + xf''(x) = 0$. Nous posons d'abord $g(x) = f'(x)$. L'équation pour g est

$$2g(x) + xg'(x) = 0, \tag{95.203}$$

c'est-à-dire

$$\frac{g'}{f} = -\frac{2}{x}. \tag{95.204}$$

En intégrant des deux côtés,

$$\ln\big(g(x)\big) = -2\ln(x) + C. \tag{95.205}$$

Les fonctions g qui satisfont l'équation sont donc

$$g(x) = Kx^{-2}. \tag{95.206}$$

Étant donné que $g = f'$, une intégration donne

$$f(x) = -\frac{K}{x} + C, \tag{95.207}$$

et donc les fonctions harmoniques radiales sont les fonctions de la forme

$$f(\rho) = -\frac{K}{\rho} + C. \tag{95.208}$$

Exercise 548 exoOutilsMath-0083

Un vent se dirige vers le sud, mais la force de Coriolis le détourne un peu vers l'est. En termes de coordonnées sphériques sur la Terre (de rayon R), le vent est donné par

$$u(\rho, \theta, \varphi) = \sin(\varphi)e_\theta + e_\varphi. \tag{95.209}$$

(1) Calculer $\nabla \cdot u$ et $\nabla \times u$.

(2) Est-ce qu'il existe une fonction $p(\rho, \theta, \varphi)$ telle que $u = \nabla p$?

corrOutilsMath-0083

Correction of the exercise 548

Nous avons $F_\rho = 0$, $F_\theta = \sin(\varphi)$ et $F_\varphi = 1$. Par conséquent,

$$\begin{aligned}\nabla \cdot u &= \frac{1}{\rho^2 \sin(\theta)}\left(\frac{\partial}{\partial\theta}(\rho \sin\theta \sin\varphi) + \frac{\partial}{\partial\varphi}(\rho)\right) \\ &= \frac{\sin\varphi\cos\theta}{R\sin\theta}.\end{aligned} \tag{95.210}$$

Notez qu'on a le droit de remplacer ρ par R *après* avoir fait les dérivées.

Pour le rotationnel,

$$\nabla \times u = \frac{1}{R\sin\theta}(\cos\theta - \cos\varphi)e_\rho + \frac{1}{R}e_\theta + \frac{\sin\varphi}{R}e_\varphi. \tag{95.211}$$

Étant donné que le rotationnel de u n'est pas nul, il n'existe pas de fonction dont le gradient est u.

Exercise 549 exoOutilsMath-0084

Soit le champ de vecteurs

$$F(x, y, z) = \frac{1}{x^2 + y^2}\begin{pmatrix} -y \\ x \\ 4 \end{pmatrix}. \tag{95.212}$$

(1) Écrire F en coordonnées cylindriques.

(2) Donner la divergence de F.

corrOutilsMath-0084

Correction of the exercise 549

En coordonnées cartésiennes le champ donné est

$$F(x, y, z) = \frac{1}{r^2}(-ye_x + xe_y + 4e_z). \tag{95.213}$$

Étant donné que nous avons vu que $e_\theta = \frac{1}{r}(-ye_x + xe_y)$, nous avons

$$F(r,\theta,z) = \frac{1}{r^2}(re_\theta + 4e_z). \tag{95.214}$$

Nous avons donc $F_r = 0$, $F_\theta = 1/r$ et $F_z = 4/r^2$. En utilisant la formule nous trouvons

$$\nabla \cdot F = 0. \tag{95.215}$$

Cela est logique étant donné que ce champ de vecteurs décrit un mouvement qui est la superposition d'un écoulement à vitesse constante le long de l'axe z et d'une rotation autour de cet axe. Les écoulements à vitesse constante et les rotations sont des mouvements de fluides incompressibles.

Exercise 550 exoOutilsMath-0085

(1) Exprimer, en coordonnées polaires dans le plan, l'équation d'une demi-droite partant de l'origine.

(2) Exprimer, en coordonnées sphériques de l'espace $\mathbb{R}^3$, les équations d'une demi-droite de $\mathbb{R}^3$ passant par l'origine, puis l'équation d'une droite passant par l'origine.

(3) Équations d'une droite passant par l'origine en coordonnées cartésiennes ?

corrOutilsMath-0085

Correction of the exercise 550

<+CorrOutilsMath-0085+>

Exercise 551 exoOutilsMath-0086

On pose, pour tout $(x,y) \in \mathbb{R}^2$, $f(x,y) = \sin(xy) + \cos(x^2+y^4)$.

(1) Écrire les deux dérivées partielles $\frac{\partial f}{\partial x}(x,y)$ et $\frac{\partial f}{\partial y}(x,y)$ au point $(x,y) \in \mathbb{R}^2$.

(2) Écrire le gradient $\nabla f_{(0,0)}$ de la fonction f au point $(0,0)$.

(3) Écrire l'équation (cartésienne) du plan tangent au graphe de f au point $(0,0,1)$.

corrOutilsMath-0086

Correction of the exercise 551

<+CorrOutilsMath-0086+>

Exercise 552 exoOutilsMath-0087

On pose, pour tout $(x,y,z) \in \mathbb{R}^3$, $f(x,y,z) = xy + e^{xy}\sin(yz)$. Calculer la dérivée directionnelle de f au point $(0,0,0)$ dans la direction du vecteur $U_0 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, puis du vecteur $U_1 = (0,0,1)$.

corrOutilsMath-0087

Correction of the exercise 552

<+CorrOutilsMath-0087+>

Exercise 553 exoOutilsMath-0088

Soit la fonction $f(x,y) = x\sin\big(y\ln(2x)\big)$.

(1) Donner l'équation du plan tangent au graphe de f au point $(e,0)$.

(2) Donner trois points distincts situés dans ce plan.

Note : vous remarquerez avec étonnement que la dérivée de la fonction $x \mapsto \ln(ax)$ ne dépend pas de a !

corrOutilsMath-0088

Correction of the exercise 553

Les dérivées partielles de f se calculent :

$$\begin{aligned} \frac{\partial f}{\partial x} &= y\cos\big(y\ln(2x)\big) + \sin\big(y\ln(2x)\big) \\ \frac{\partial f}{\partial y} &= x\ln(2x)\cos\big(y\ln(2x)\big). \end{aligned} \tag{95.216}$$

Au point $(e,0)$ nous avons

$$\begin{aligned}\frac{\partial f}{\partial x}(e,0)&=0\\ \frac{\partial f}{\partial y}(e,0)&=e\ln(2e).\end{aligned}\tag{95.217}$$

Le plan a donc pour équation

$$T_{(e,0)}(x,y)=ye\ln(2e)\tag{95.218}$$

parce que $f(e,0)=0$.

Pour trouver trois points dans ce plan, il suffit d'écrire $\big(x,y,T_{(e,0)}(x,y)\big)$ pour n'importe quel choix de x et y. Par exemple $(0,0,0)$, $(0,1,e\ln(2e))$ et $(1,0,0)$.

Exercise 554 exoOutilsMath-0089

Calculer $\frac{\partial f}{\partial u}(a,b,c)$ où $f(x,y,z)=xy+\sin(z)$ et u est la direction donnée par le vecteur $(1,1,2)$. Ici a et b sont deux réels quelconques. corrOutilsMath-0089

Correction of the exercise 554

Les dérivées partielles de f sont $\frac{\partial f}{\partial x}=y$, $\frac{\partial f}{\partial y}=x$ et $\frac{\partial f}{\partial z}=\cos(z)$.

Attention : lorsqu'on cherche une dérivée directionnelle, il faut s'assurer que le vecteur « direction » soit de norme 1. Ici nous devons donc considérer $u=(1,1,2)/\sqrt{6}$. Le calcul est donc

$$\frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}b\\a\\\cos(c)\end{pmatrix}=\frac{1}{\sqrt{6}}\big(b+a+2\cos(c)\big).\tag{95.219}$$

Exercise 555 exoOutilsMath-0090

Soit

$$f(x,y,z)=\frac{z}{x^2+y^2}.\tag{95.220}$$

(1) Calculer $\nabla f(a,b,c)\cdot v$ si $v=(1,2,-3)$.

(2) Donner $df_{(a,b,c)}(v)$.

(3) Calculer la circulation du champ de vecteurs ∇f le long du cercle $\sigma(t)=\big(2\cos(t),5,2\sin(t)\big)$ avec $t\in[0,2\pi]$. corrOutilsMath-0090

Correction of the exercise 555

Les dérivées partielles sont données par

$$\frac{\partial f}{\partial x}=\frac{-2xz}{(x^2+y^2)^2},\quad\frac{\partial f}{\partial y}=\frac{-2yz}{(x^2+y^2)^2},\quad\frac{\partial f}{\partial z}=\frac{1}{x^2+y^2}.\tag{95.221}$$

Nous avons donc

$$\nabla f(a,b,c)=\begin{pmatrix}\frac{-2ac}{(a^2+b^2)^2}\\\frac{-2bc}{(a^2+b^2)^2}\\\frac{1}{a^2+b^2}\end{pmatrix}\tag{95.222}$$

et le produit scalaire avec $(1,2,-3)$ est alors

$$\nabla f(a,b,c)\cdot\begin{pmatrix}1\\2\\-3\end{pmatrix}=\frac{-2ac}{(a^2+b^2)^2}+2\frac{-2bc}{(a^2+b^2)^2}-3\frac{1}{a^2+b^2}.\tag{95.223}$$

La valeur de $df_{(a,b,c)}(v)$ est exactement la même.

La circulation de ∇f le long d'un chemin $\sigma\colon[a,b]\to\mathbb{R}^3$ est donnée par $f\big(\sigma(b)\big)-f\big(\sigma(a)\big)$. Ici nous avons donc

$$\int_\sigma\nabla f=f\big(\sigma(2\pi)\big)-f\big(\sigma(0)\big)=0.\tag{95.224}$$

La circulation d'un champ de vecteur conservatif le long d'un chemin fermé est toujours nulle.

Exercise 556 exoOutilsMath-0091

Soit le champ de vecteurs

$$F(x,y,z)=\begin{pmatrix}\sin(yz)\\ xz\cos(yz)\\ xy\cos(yz)\end{pmatrix}. \tag{95.225}$$

(1) Calculer la circulation de ce champ de vecteurs le long du segment de droite qui joint le point $(0,0,0)$ et le point $(4,1,0)$.

(2) Donner le champ de vecteurs $\nabla\times F$.

(3) Calculer le nombre $(\nabla\cdot F)\big(\frac{1}{\sqrt{3}},\pi,\frac{1}{4}\big)$.

corrOutilsMath-0091

Correction of the exercise 556

Ici, il était possible de remarquer que le champ de vecteur proposé dérivait du potentiel $V(x,y,z)=x\sin(yz)$. En remarquant cela, la réponse à la première question est $V(4,1,0)-V(0,0,0)=0-0=0$. Le rotationnel est également automatiquement nul.

Si on ne remarque pas que le champ dérive d'un potentiel, il faut un peu plus travailler.

(1) Le chemin est $\sigma(t)=(4t,t,0)$, et la dérivée vaut $\sigma'(t)=(4,1,0)$. Nous avons aussi

$$F\big(\sigma(t)\big)=\begin{pmatrix}0\\ 0\\ 4t^2\end{pmatrix}. \tag{95.226}$$

Nous avons donc

$$F\big(\sigma(t)\big)\cdot\sigma'(t)=0 \tag{95.227}$$

et l'intégrale est donc nulle.

(2) Le rotationnel est donné par le déterminant

$$\nabla\times F=\begin{vmatrix}e_x & e_y & e_z\\ \partial_x & \partial_y & \partial_z\\ \sin(yz) & xz\cos(yz) & xy\cos(yz)\end{vmatrix}. \tag{95.228}$$

En calculant les dérivées, on remarque que tout s'annule.

(3) En ce qui concerne la divergence,

$$\nabla\cdot F=-xz^2\sin(yz)-xy^2\sin(yz), \tag{95.229}$$

et par conséquent

$$(\nabla\cdot F)\big(\frac{1}{\sqrt{3}},\pi,\frac{1}{4}\big)=-\frac{1}{2}\sqrt{\frac{2}{3}}\left(\frac{1}{16}+\pi^2\right). \tag{95.230}$$

Note. Pour calculer la divergence à l'aide de Sage, on peut faire comme ceci :

```
sage: var('x,y,z')
(x, y, z)
sage: fx=sin(y*z)
sage: fy=x*z*cos(y*z)
sage: fz=x*y*cos(y*z)
sage: expression=symbolic_expression(fx.diff(x)+fy.diff(y)+fz.diff(z))
sage: divergence=expression.function(x,y,z)
sage: divergence(1/sqrt(3),pi,1/4)
-1/6*pi^2*sqrt(2)*sqrt(3) - 1/96*sqrt(2)*sqrt(3)
```

Exercise 557 exoOutilsMath-0092

Intégrer la fonction $f(x,y) = \frac{1}{x^2+y^2}$ sur la couronne de rayon interne R_1 et de rayon externe R_2 corrOutilsMath-0092

Correction of the exercise 557

En coordonnées polaires, le domaine proposé est donné par $\theta\colon 0 \to 2\pi$ et $r\colon R_1 \to R_2$. La fonction, elle, est donnée par $f(r,\theta) = \frac{1}{r^2}$. Donc l'intégrale est

$$I = \int_{R_1}^{R_2}\int_0^{2\pi} \frac{1}{r^2} r\, d\theta dr = 2\pi \int_{R_1}^{R_2} \frac{dr}{r} = 2\pi\big(\ln(R_2) - \ln(R_1)\big). \tag{95.231}$$

Exercise 558 exoOutilsMath-0093

Soit T le domaine triangulaire déterminé par les points $(1,1)$, $(3,1)$ et $(1,5)$.

(1) Écrire ce domaine sous forme de domaine « vertical ».

(2) Écrire ce domaine sous forme de domaine « horizontal ».

(3) En calculer la surface en utilisant une intégrale double.

(4) En calculer la surface en calculant un déterminant. corrOutilsMath-0093

Correction of the exercise 558

<+CorrOutilsMath-0093+>

Exercise 559 exoOutilsMath-0094

Calculer

$$I = \int_B e^{(x^2+y^2+z^2)^{3/2}} dxdydz \tag{95.232}$$

où B est la boule unité de $\mathbb{R}^3$ centrée à l'origine. corrOutilsMath-0094

Correction of the exercise 559

Tant la forme de la fonction que celle du domaine d'intégration demandent de passer aux coordonnées sphériques. Cédant à la pression, nous passons aux sphériques. Le domaine est donné par

$$\begin{aligned} \rho&\colon 0 \to 1 \\ \theta&\colon 0 \to \pi \\ \varphi&\colon 0 \to 2\pi. \end{aligned} \tag{95.233}$$

La fonction quant à elle vaut

$$f(\rho,\theta,\varphi) = e^{\rho^3}. \tag{95.234}$$

Nous devons donc effectuer l'intégrale

$$I = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin(\theta) d\theta \int_0^1 e^{\rho^3} r^2 d\rho = \frac{3\pi}{3}\int_0^1 e^{\rho^3} 3\rho^2 d\rho. \tag{95.235}$$

Pour effectuer la dernière intégrale, nous faisons le changement de variable $u = \rho^3$, $du = 3\rho^2 d\rho$:

$$I = \frac{4\pi}{3}\int_0^1 e^u du = \frac{4\pi}{3}(e-1). \tag{95.236}$$

Exercise 560 exoOutilsMath-0095

On considère la fonction

$$f\colon (x,y,z) \in \mathbb{R}^2 \to x^2yz + 2y^2\sin(xyz). \tag{95.237}$$

(1) Calculer les dérivées partielles premières de f.

(2) Quelle est la différentielle de f au point $(1,1,\pi)$?

(3) Donner une approximation de $f(1+10^{-2}, 1-10^{-3}, \pi+10^{-4})$.

(4) Soit $\overrightarrow{F}(x,y,z)=\nabla f(x,y,z)$ le gradient de f au point (x,y,z). Que vaut la circulation de $\overrightarrow{F}$ le long d'une courbe d'extrémités $(1,2,\pi)$ et $(1,1,\frac{\pi}{2})$.

corrOutilsMath-0095

Correction of the exercise 560

Les dérivées partielles de f sont :

$$\begin{aligned}\frac{\partial f}{\partial x}&=2y^3z\cos(xyz)+2xyz\\ \frac{\partial f}{\partial y}&=2xy^2z\cos(xyz)+x^2z+4y\sin(xyz)\\ \frac{\partial f}{\partial z}&=2xy^3\cos(xyz)+x^2y.\end{aligned}\tag{95.238}$$

Nous avons donc

$$\frac{\partial f}{\partial x}(1,1,\pi)=0,\quad \frac{\partial f}{\partial y}(1,1,\pi)=-\pi,\quad \frac{\partial f}{\partial z}(1,1,\pi)=-1.\tag{95.239}$$

La différentielle est donc donnée par

$$df_{(1,1,\pi)}(u)=-\pi u_2-u_3.\tag{95.240}$$

L'approximation est

$$\begin{aligned}f(1+10^{-2},1-10^{-3},\pi+10^{-4})&\simeq f(1,1,\pi)-10^{-3}(-\pi)+10^{-4}(-1)\\ &=\pi(1+10^{-3})-10^{-4}.\end{aligned}\tag{95.241}$$

En ce qui qui concerne la circulation de ∇f, nous avons un potentiel (f) est donc pas de problèmes :

$$\int_\sigma\nabla f=f(1,1,\frac{\pi}{2})-f(1,2,\pi)=\frac{\pi}{2}+2-\left(-\frac{3\pi}{2}+2\right)=-\frac{3\pi}{2}+2.\tag{95.242}$$

Les calculs peuvent être faits avec Sage :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y,z)=x**2*y*z+2*y**2*sin(x*y*z)
sage: f.diff(x)
(x, y, z) |--> 2*y^3*z*cos(x*y*z) + 2*x*y*z
sage: f.diff(y)
(x, y, z) |--> 2*x*y^2*z*cos(x*y*z) + x^2*z + 4*y*sin(x*y*z)
sage: f.diff(z)
(x, y, z) |--> 2*x*y^3*cos(x*y*z) + x^2*y
sage: f.diff(x)(1,1,pi)
0
sage: f.diff(x)(x=1,y=1,z=pi)
0
sage: f.diff(x)
sage: f.diff(y)(x=1,y=1,z=pi)
-pi
sage: f.diff(z)(x=1,y=1,z=pi)
-1
sage: f(1,1,pi)
pi
```

```
sage: f(1,1,pi/2)
1/2*pi + 2
sage: f(1,2,pi)
2*pi
sage: f(1,1,pi/2)-f(1,2,pi)
-3/2*pi + 2
```

Notez les deux possibilités pour calculer $\partial_x f(a,b,c)$. On peut écrire `f.diff(x)(a,b,c)` ou bien `f.diff(x)(x=a,y=b,z=c)`.

Exercise 561 exoOutilsMath-0096

On considère le champ de vecteurs de classe C^1 sur $\mathbb{R}^2\backslash\{(0,0)\}$

$$F(x,y) = \left(\frac{u}{x^2+y^2}, \frac{-x}{x^2+y^2}\right). \tag{95.243}$$

(1) Calculer le rotationnel de F.

(2) Soit $\sigma\colon t \in [0,2\pi] \to \big(\cos(t),\sin(t)\big)$ le cercle unité parcouru une fois dans le sens direct. Calculer la circulation de F le long de σ.

(3) Le champ F peut-il dériver d'un potentiel scalaire sur $\mathbb{R}^2\backslash\{(0,0)\}$?

corrOutilsMath-0096

Correction of the exercise 561

(1) Le rotationnel s'écrit

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ \frac{y}{x^2+y^2} & \frac{-x}{x^2+y^2} & 0 \end{vmatrix} = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) e_z. \tag{95.244}$$

En ce qui concerne les dérivées partielles, nous avons

$$\begin{aligned} \frac{\partial F_y}{\partial x} &= \frac{x^2-y^2}{(x^2+y^2)^2} \\ \frac{\partial F_x}{\partial y} &= \frac{x^2-y^2}{(x^2+y^2)^2}, \end{aligned} \tag{95.245}$$

et par conséquent $\nabla \times F = 0$.

(2) Nous avons

$$F\big(\sigma(t)\big) = F\big(\cos(t),\sin(t),0\big) = \begin{pmatrix} \sin(t) \\ -\cos(t) \\ 0 \end{pmatrix}, \tag{95.246}$$

et

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 0 \end{pmatrix}. \tag{95.247}$$

Il faut donc intégrer

$$\int_\sigma F = \int_0^{2\pi} (-1)dt = -2\pi. \tag{95.248}$$

(3) Le fait que l'intégrale de F le long d'un chemin fermé ne soit pas nulle implique que le champ ne dérive pas d'un potentiel. En effet, dès qu'un champ dérive d'un potentiel, son intégrale sur un chemin fermé vaut zéro.

Exercise 562 exoOutilsMath-0097

(1) Montrer que le champ de vecteurs sur $\mathbb{R}^2$

$$F(x,y) = \begin{pmatrix} 6xy + 4x^3y^4 \\ 3x^2 + 4x^4y^3 \end{pmatrix} \tag{95.249}$$

dérive d'un potentiel scalaire. Donner un tel potentiel.

(2) Quel est le rotationnel de F ?

(3) Montrer que le champ de vecteurs sur $\mathbb{R}^2$

$$F(x,y) = \begin{pmatrix} y\sin(xy) \\ -x\sin(xy) \end{pmatrix} \tag{95.250}$$

n'est pas un champ de gradients, c'est-à-dire qu'il ne dérive pas d'un potentiel scalaire.

corrOutilsMath-0097

Correction of the exercise 562

(1) Le potentiel $V(x,y,z)$ doit vérifier les deux équations

$$\begin{aligned} \frac{\partial V}{\partial x} &= 6xy + 4x^3y^4 \\ \frac{\partial V}{\partial y} &= 3x^2 + 4x^4y^3. \end{aligned} \tag{95.251}$$

Il est facile de voir que la fonction $V(x,y,z) = 3x^2y + x^4y^4$ satisfait à ces deux équations en même temps.

(2) Étant donné que le champ dérive d'un potentiel, son rotationnel est nul :

$$\nabla \times F = \nabla \times (\nabla V) = 0. \tag{95.252}$$

(3) Le rotationnel du champ proposé est donné par

$$\nabla \times F = \begin{pmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_y \\ y\sin(xy) & -x\sin(xy) & 0 \end{pmatrix} = -2xy\cos(xy) + 2\sin(xy). \tag{95.253}$$

(4) Le fait que le rotationnel de F soit nul implique que F ne peut pas être un champ de gradients.

Exercise 563 exoOutilsMath-0098

Soit la courbe paramétrée de $\mathbb{R}^2$

$$\begin{aligned} \sigma\colon [0,\pi/2] &\to \mathbb{R}^2 \\ \sigma(t) &= \begin{pmatrix} \cos^3(t) \\ \sin^3(t) \end{pmatrix}. \end{aligned} \tag{95.254}$$

(1) Montrer que $\|\sigma'(t)\| = 3\sin(t)\cos(t)$

(2) Calculer la longueur de cette courbe.

corrOutilsMath-0098

Correction of the exercise 563

Nous avons d'abord

$$\sigma'(t) = \begin{pmatrix} -3\cos^2(t)\sin(t) \\ 3\sin^2(t)\cos(t) \end{pmatrix}. \tag{95.255}$$

Par conséquent,

$$\|\sigma'(t)\|^2 = 9\cos^2(t)\sin^2(t)\big(\cos^2(t) + \sin^2(t)\big) = 9\cos^2(t)\sin^2(t). \tag{95.256}$$

En ce qui concerne la longueur,

$$l(\sigma) = \int_0^{\pi/2} 3\sin(t)\cos(t)dt. \tag{95.257}$$

Cette intégrale se règle avec le changement de variables $u = \sin(t)$, $du = \cos(t)dt$. Lorsque $t = 0$, $u = 0$ et lorsque $t = \pi/2$, $u = 1$. Par conséquent

$$l(\sigma) = 3\int_0^1 udu = \frac{3}{2}. \tag{95.258}$$

Une autre façon de résoudre l'intégrale est d'écrire $\sin(t)\cos(t) = \frac{\sin(2t)}{2}$. Nous nous retrouvons alors avec

$$l(\sigma) = \frac{3}{2}\int_0^{\pi/2} \sin(2t)dt. \tag{95.259}$$

Une primitive de $\sin(2t)$ est facile à trouver : c'est $-\cos(2t)/2$.

Une troisième façon de calculer

$$I = \int_0^{\pi/2} \cos(t)\sin(t)dt \tag{95.260}$$

est de faire par parties :

$$\begin{aligned} u &= \cos(t) \quad u' = -\sin(t) \\ v' &= \sin(t) \quad v = -\cos(t). \end{aligned} \tag{95.261}$$

De la sorte,

$$I = \left[-\cos^2(t)\right]_0^{\pi/2} - \int_0^{\pi/2} \sin(t)\cos(t)dt, \tag{95.262}$$

c'est-à-dire $I = -1 - I$, et par conséquent $I = -1/2$.

Exercise 564 exoOutilsMath-0099

Nous considérons les points du plan $A = (1, 1)$, $B = (2, 3)$ et $C = (5, 2)$.

(1) Calculer les composantes des vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$.

(2) Calculer l'aire du parallélogramme déterminé par ces deux vecteurs.

corrOutilsMath-0099

Correction of the exercise 564

Les coordonnées des vecteurs sont $\overrightarrow{AB} = (2, 3) - (1, 1) = (1, 2)$ et $\overrightarrow{AC} = (5, 2) - (1, 1) = (4, 1)$. L'aire du parallélogramme déterminé par deux vecteurs est la valeur absolue du déterminant formé par les deux vecteurs :

$$\begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} = 1 - 8 = -7. \tag{95.263}$$

La surface demandée est donc 7.

Exercise 565 exoOutilsMath-0100

On pose pour tout $x \in]-1, \infty[$, $f(x) = x\ln(x)$. Écrire l'équation (cartésienne) de la droite tangente au graphe de la fonction f au point $(x_0, y_0) = (1, \ln 2)$.

corrOutilsMath-0100

Correction of the exercise 565

<+CorrOutilsMath-0100+>

Exercise 566 exoOutilsMath-0101

Calculer l'intégrale de la fonction $f(x, y) = xy^2$ sur les domaines suivants :

(1) $D = \{(x, y) \in \mathbb{R}^2 \text{ tel que } 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$;

(2) $D = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x \geqslant 0, y \geqslant 0, x + y \leqslant 1\}$;

(3) $D = \{(x, y) \in \mathbb{R}^2 \text{ tel que } 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant x^2\}$.

corrOutilsMath-0101

Correction of the exercise 566

(1) Les bornes sont

$$\begin{cases} x\colon 0 \to 1 & (95.264a)\\ y\colon 0 \to 1 & (95.264b)\end{cases}$$

et l'intégrale est

$$\int_0^1\int_0^1 xy^2dxdy = \int_0^1 y^2\left[\frac{x^2}{2}\right]_0^1 dy = \int_0^1 \frac{y^2}{2} = \frac{1}{6}. \tag{95.265}$$

(2) Cette fois les bornes sont

$$\begin{cases} x\colon 0 \to 1 & (95.266a)\\ y\colon 0 \to 1-x, & (95.266b)\end{cases}$$

et l'intégrale à calculer est

$$\int_0^1\int_0^{1-x} xy^2\,dy\,dx. \tag{95.267}$$

L'intégrale sur y est

$$\int_0^{1-x} xy^2dy = x\left[\frac{y^3}{3}\right]_0^{1-x} = \frac{x}{3}(1-x)^3. \tag{95.268}$$

L'intégrale sur x restante à faire est alors

$$\int_0^1 \frac{x}{3}(1-x)^3 = \frac{1}{60}. \tag{95.269}$$

(3) Les bornes sont

$$\begin{cases} x\colon 0 \to 1 & (95.270a)\\ y\colon 0 \to x^2, & (95.270b)\end{cases}$$

et l'intégrale est

$$\int_0^1\int_0^{x^2} xy^2\,dydx = \frac{1}{3}\int_0^1 x^7dx = \frac{1}{24}. \tag{95.271}$$

Vérification :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x*y**2
sage: f.integrate(y,0,x**2).integrate(x,0,1)
(x, y) |--> 1/24
```

Exercise 567 exoOutilsMath-0102

Calculer l'intégrale de f sur le domaine D dans les cas suivants :

(1) $f(x,y) = x$, $D = \{(x,y) \in \mathbb{R}^2$ tel que $y \geqslant 0, x-y \geqslant 0, x+2y-4 \leqslant 0\}$;
(2) $f(x,y) = x+y$, $D = \{(x,y) \in \mathbb{R}^2$ tel que $0 \leqslant x \leqslant 1, x^2 \leqslant y \leqslant x\}$;
(3) $f(x,y) = \cos(xy)$, $D = \{(x,y) \in \mathbb{R}^2$ tel que $1 \leqslant x \leqslant 2, 0 \leqslant xy \leqslant \frac{\pi}{2}\}$;
(4) $f(x,y) = xy$, $D = \{(x,y) \in \mathbb{R}^2$ tel que $x \geqslant 0, y \geqslant 0, xy+x+y \leqslant 1\}$;
(5) $f(x,y) = 1/(x+y)^3$, $D = \{(x,y) \in \mathbb{R}^2$ tel que $1 \leqslant x \leqslant 3, y \geqslant 2, x+y \leqslant 5\}$.

corrOutilsMath-0102

Correction of the exercise 567

(1) Pour cet exercice, vous devriez faire un petit dessin. Le domaine d'intégration est le triangle de sommets $(0,0)$, $(\frac{4}{3},\frac{4}{3})$, $(4,0)$. Nous le décomposons en deux parties. Celle de gauche est

$$\begin{cases} x\colon 0\to\frac{4}{3} & (95.272a)\\ y\colon 0\to x. & (95.272b)\end{cases}$$

et celle de droite est

$$\begin{cases} x\colon \frac{4}{3}\to 4 & (95.273a)\\ y\colon 0\to 2-\frac{x}{2}. & (95.273b)\end{cases}$$

L'intégrale à calculer est donc

$$\int_0^{4/3}\int_0^x x\,dydx+\int_{4/3}^4\int_0^{2-x/2} x\,dydx=\frac{64}{81}+\frac{320}{81}=\frac{128}{27}. \tag{95.274}$$

```
sage: f(x,y)=x
sage: f.integrate(y,0,x).integrate(x,0,4/3)
(x, y) |--> 64/81
sage: f.integrate(y,0,2-x/2).integrate(x,4/3,4)
(x, y) |--> 320/81
sage: 64/81+320/81
128/27
```

Exercise 568 exoOutilsMath-0103

Soit le domaine

$$D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } -1\leqslant x\leqslant 1, x^2\leqslant y\leqslant 4-x^3\}. \tag{95.275}$$

Calculer l'aire de D.

corrOutilsMath-0103

Correction of the exercise 568

<+CorrOutilsMath-0103+>

Exercise 569 exoOutilsMath-0104

On pose $D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1, x^2+y^2\geqslant 1\}$. Calculer

$$\iint_D \frac{xy}{1+x^2+y^2}dxdy. \tag{95.276}$$

corrOutilsMath-0104

Correction of the exercise 569

<+CorrOutilsMath-0104+>

Exercise 570 exoOutilsMath-0105

Calculer l'intégrale double

$$\iint_D \frac{1}{1+x^2+y^2}dxdy \tag{95.277}$$

où $D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1, x^2+y^2\leqslant 1\}$.

corrOutilsMath-0105

Correction of the exercise 570

<+CorrOutilsMath-0105+>

Exercise 571 exoOutilsMath-0106

Soit $D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } x^2+y^2-2x\leqslant 0\}$.

(1) Montrer que D est un disque. Préciser le centre et le rayon.
(2) Calculer $\iint_D \sqrt{x^2+y^2}dxdy$.

corrOutilsMath-0106

Correction of the exercise 571

<+CorrOutilsMath-0106+>

Exercise 572 exoOutilsMath-0107

Calculer l'intégrale double $\iint_D f(x,y)dxdy$ dans les cas suivants :

(1) $D = \{(x,y) \in \mathbb{R}^2 \text{ tel que } x \geqslant 0, y \geqslant 0, 1 \leqslant x^2+y^2 \leqslant 4\}$ et

$$f(x,y) = \frac{xy}{x^2+y^2}. \tag{95.278}$$

(2) $D = \{(x,y) \in \mathbb{R}^2 \text{ tel que } x \geqslant 0, 1 \leqslant x^2+y^2 \leqslant 2y\}$ et

$$f(x,y) = \sqrt{x^2+y^2}. \tag{95.279}$$

corrOutilsMath-0107

Correction of the exercise 572

<+CorrOutilsMath-0107+>

Exercise 573 exoOutilsMath-0108

Soit la fonction, donnée en coordonnées polaires, $f(r,\theta) = \theta$. Intégrer cette fonction sur le cercle de rayon R.

corrOutilsMath-0108

Correction of the exercise 573

Le chemin sur lequel on intègre est $\sigma(t) = (R\cos(t), R\sin(t))$. L'astuce à comprendre pour cet exercice est que

$$f\big(\sigma(t)\big) = t, \tag{95.280}$$

étant donné que t est l'angle correspondant au point $\big(R\cos(t), R\sin(t)\big)$. L'intégrale à calculer est donc

$$\int_\sigma f d\sigma = \int_0^{2\pi} t\|\sigma'(t)\|dt = R\int_0^{2\pi} t\,dt = 2R\pi^2. \tag{95.281}$$

Exercise 574 exoOutilsMath-0109

Soit le champ de vecteurs

$$F(x,y) = \begin{pmatrix} y\cos(x) \\ xy \end{pmatrix}. \tag{95.282}$$

Intégrer la fonction $(x,y) \mapsto \nabla \cdot F(x,y,z)$ sur le carré de sommets $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$.

corrOutilsMath-0109

Correction of the exercise 574

La fonction à intégrer est $f(x,y) = \nabla \cdot F(x,y) = -y\sin(x) + x$, et le domaine d'intégration est $[0,1]\times[0,1]$, c'est-à-dire

$$I = \int_0^1 dx \int_0^1 \big(-y\sin(x) + x\big)dy = \int_0^1 \big(x - \frac{\sin(x)}{2}\big)dx = \frac{\cos(1)}{2}. \tag{95.283}$$

Exercise 575 exoOutilsMath-0110

Calculer le flux du champ de vecteurs

$$F(x,y,z) = e_x \tag{95.284}$$

au travers du cylindre de rayon R et de hauteur h autour de l'axe z.

Même question si le cylindre est autour de l'axe x.

Remarque : ces cylindres sont considérés *avec* leur « couvercles ».

corrOutilsMath-0110

Correction of the exercise 575

Un paramétrage du cylindre autour de l'axe z est

$$\phi(\theta, z) = \begin{pmatrix} R\cos\theta \\ R\sin\theta \\ z \end{pmatrix}. \tag{95.285}$$

Les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \\ 0 \end{pmatrix}, T_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{95.286}$$

Le vecteur normal est donc

$$T_\theta \times T_z = R\cos(\theta)e_x + R\sin(\theta)e_y. \tag{95.287}$$

C'est un vecteur dirigé vers l'extérieur.

Le champ de vecteurs considéré est constant : $F(\theta, z) = e_x$. Nous avons donc

$$F(\theta, z) \cdot (T_\theta \times T_z) = R\cos(\theta) \tag{95.288}$$

et le flux vaut

$$\Phi = \int_0^{2\pi} d\theta \int_0^h R\cos(\theta) dz = 0. \tag{95.289}$$

En ce qui concerne les couvercles haut au bas, ils sont paramétrés par

$$\phi_1(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ h \end{pmatrix}, \phi_2(r, \theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ 0 \end{pmatrix}. \tag{95.290}$$

Les vecteurs normaux correspondants sont dans la direction de e_z, de façon que le produit scalaire avec $F(r, \theta)$ soit nul. Le flux total est donc nul.

Regardons maintenant le cylindre le long de l'axe x. Un paramétrage est

$$\phi(\theta, x) = \begin{pmatrix} x \\ R\cos(\theta) \\ R\sin(\theta) \end{pmatrix}, \tag{95.291}$$

et le vecteurs tangents sont

$$T_\theta = \begin{pmatrix} 0 \\ -R\sin\theta \\ R\cos\theta \end{pmatrix}, T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{95.292}$$

Le vecteur normal est alors donné par

$$T_\theta \times T_x = R\cos(\theta)e_y + R\sin(\theta)e_z. \tag{95.293}$$

Nous avons par conséquent $F(\theta, x) \cdot (T_\theta \times T_x) = 0$. Pas de flux par le côté du cylindre.

Regardons les « couvercles ». Le premier est donné par le paramétrage

$$\phi_1(r, \theta) = \begin{pmatrix} 0 \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{95.294}$$

Le vecteur normal serait $T_r \times T_\theta = re_x$, et le flux

$$\Phi = \int_0^{2\pi} d\theta \int_0^R r\, dr = \pi R^2. \tag{95.295}$$

Le second couvercle est donné par

$$\phi_2(r,\theta) = \begin{pmatrix} h \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{95.296}$$

Le vecteur normal est encore re_x, et le flux est à nouveau πR^2.

Le flux total serait donc $2\pi R^2$.

Cela n'est pas possible parce que tous les vecteurs qui « rentrent » d'un côté doivent « sortir » de l'autre côté. L'erreur est le le premier vecteur normal est un vecteur qui pointe vers l'intérieur du cylindre, tandis que le second pointe vers l'extérieur. Si nous choisissons, par convention, de prendre uniquement les vecteurs extérieurs, il faut changer le vecteur normal du premier couvercle en $-re_x$. Le premier flux vaudra donc

$$-\pi R^2, \tag{95.297}$$

de telle sorte que le flux total sera nul.

Exercise 576 exoOutilsMath-0111

Soit le champ de vecteurs

$$F(x,y,z) = \frac{1}{|x|}e_x. \tag{95.298}$$

Calculer son flux au travers du cube de côté 2 dont le point le plus proche de l'origine est (1, 1, 1). corrOutilsMath-0111

Correction of the exercise 576

Seules les deux faces $x = 1$ et $x = 3$ sont perpendiculaires au champ de vecteurs donné. Les quatre autres faces sont parallèles et donc de flux nul.

La première face prend le paramétrage

$$\phi_1(y,z) = \begin{pmatrix} 1 \\ y \\ z \end{pmatrix} \tag{95.299}$$

avec $(u,z) \in [1,3]\times[1,3]$. Le vecteur normal est e_x, mais comme nous voulons suivre la convention de prendre des vecteurs normaux *extérieurs*, nous prenons $-e_x$. Nous avons aussi $F\big(\phi_1(y,z)\big) = e_x$, donc

$$\Phi_1 = \int_1^3 dy \int_1^3 e_x \cdot (-e_x)dz = -4. \tag{95.300}$$

La seconde face est donnée par

$$\phi_2(y,z) = \begin{pmatrix} 3 \\ y \\ z \end{pmatrix}. \tag{95.301}$$

Le vecteur normal est e_x, et il pointe vers l'extérieur. Nous avons $F\big(\phi_2(y,z)\big) = \frac{1}{3}e_x$, et donc le second flux est

$$Phi_2 = \int_1^3 dy \int_1^3 \frac{1}{3}e_x \cdot (e_x) = \frac{4}{3}. \tag{95.302}$$

Le flux total au travers du cube est donné par la somme :

$$\Phi = -4 + \frac{4}{3} = -\frac{8}{3}. \tag{95.303}$$

Exercise 577 exoOutilsMath-0112

Soit $\alpha \in [0, 2\pi]$ et $a \in \mathbb{R}^+$, et le champ de vecteurs

$$F(x,y,z) = a\cos(\alpha)e_x + a\sin(\alpha)e_y. \tag{95.304}$$

Calculer le flux de F au travers du carré centré à l'origine, dans le plan $x = 0$ et de côté $2c$. corrOutilsMath-0112

Correction of the exercise 577

Le carré a pour paramétrage

$$\phi(y,z)=\begin{pmatrix}0\\y\\z\end{pmatrix} \tag{95.305}$$

avec $y,z\colon -c\to c$. Le vecteur normal est e_x. Nous avons donc comme produit scalaire :

$$F\big(\phi(y,z)\big)\cdot e_x=a\cos(\alpha), \tag{95.306}$$

et par conséquent un flux

$$\Phi=\int_{-c}^{c}dy\int_{-c}^{c}dz a\cos(\alpha)=4c^2a\cos(\alpha). \tag{95.307}$$

Exercise 578 exoOutilsMath-0113

Intégrer la fonction $f(x,y,z)=|z|$ sur l'ellipsoïde (pleine)

$$\frac{x^2}{4}+\frac{y^2}{16}+\frac{z^2}{36}\leqslant 9. \tag{95.308}$$

Le paramétrage « recommandée » est

$$\phi(\rho,\theta,\varphi)=\begin{pmatrix}2\rho\sin(\theta)\cos(\varphi)\\4\rho\sin(\theta)\sin(\varphi)\\6\rho\cos(\theta)\end{pmatrix}. \tag{95.309}$$

Ensuite, refaire tous les calculs qu'on a fait pour les coordonnées sphériques. corrOutilsMath-0113

Correction of the exercise 578

Dans le paramétrage recommandée,

$$\phi(\rho,\theta,\varphi)=\begin{pmatrix}2\rho\sin(\theta)\cos(\varphi)\\4\rho\sin(\theta)\sin(\varphi)\\6\rho\cos(\theta)\end{pmatrix}, \tag{95.310}$$

les paramètres ont comme domaines :

$$\begin{aligned}\rho&\colon 0\to 3\\ \theta&\colon 0\to\pi\\ \varphi&\colon 0\to 2\pi.\end{aligned} \tag{95.311}$$

Les vecteurs tangents sont :

$$\begin{aligned}T_\rho&=\begin{pmatrix}2\sin\theta\cos\varphi\\4\sin\theta\sin\varphi\\6\cos\theta\end{pmatrix},\\ T_\theta&=\begin{pmatrix}2\rho\cos\theta\cos\varphi\\4\rho\cos\theta\sin\varphi\\-6\rho\sin\theta\end{pmatrix},\\ T_\varphi&=\begin{pmatrix}-2\rho\sin\theta\sin\varphi\\4\rho\sin\theta\cos\varphi\\0\end{pmatrix}.\end{aligned} \tag{95.312}$$

Ici il est question d'intégrer sur un volume. L'élément de volume se crée avec le produit mixte. Un calcul (le faire[1] !!) utilisant la formule (10.99) donne

$$T_\rho\cdot(T_\theta\times T_\varphi)=2\cdot 4\cdot 6\cdot\rho^2\sin(\theta)=48\rho^2\sin(\theta). \tag{95.313}$$

1. Oui, vraiment, fais le.

Ça, c'est pour le paramétrage de notre volume.

La fonction à intégrer est

$$f\big(\phi(\rho,\theta,\varphi)\big) = |6\rho\cos(\theta)|. \tag{95.314}$$

Étant donné que ρ est toujours positif, nous pouvons retirer la valeur absolue de ρ. Par contre sur l'intervalle d'intégration, $\cos(\theta)$ change de signe. Il faudra donc y penser. Quoi qu'il en soit, l'intégrale à effectuer est

$$\begin{aligned}\int_V f &= \int_0^3 d\rho \int_0^\pi d\theta \int_0^{2\pi} d\varphi\, 48 \cdot 6 \cdot \rho|\cos(\theta)| \cdot \rho^2\sin(\theta) \\ &= 6 \cdot 48 \cdot 2\pi \int_0^3 \rho^3 d\rho \int_0^\pi \sin(\theta)|\cos(\theta)|.\end{aligned} \tag{95.315}$$

L'intégrale sur ρ vaut $\frac{81}{4}$. Celle sur θ se coupe en deux :

$$\int_0^{\pi/2} \sin(\theta)\cos(\theta) - \int_{\pi/2}^{\pi} \sin(\theta)\cos(\theta) = 1. \tag{95.316}$$

Au final, l'intégrale vaut

$$\frac{6 \cdot 2 \cdot 48 \cdot 81\pi}{4} = 11664\pi. \tag{95.317}$$

Exercise 579 exoOutilsMath-0114

Donner un paramétrage $\phi\colon [0,2\pi]\times\mathbb{R} \to \mathbb{R}^3$ du cylindre (creux) infini de rayon 2 autour de l'axe x. En déduire l'équation du plan tangent à ce cylindre au point $(-1,1,-\sqrt{3})$. corrOutilsMath-0114

Correction of the exercise 579

Un paramétrage du cylindre situé autour de l'axe x est donné par une adaptation des coordonnées cylindriques :

$$\phi(x,\theta) = \begin{pmatrix} x \\ 2\cos\theta \\ 2\sin\theta \end{pmatrix}. \tag{95.318}$$

Pour trouver le plan tangent, la meilleur façon est souvent de trouver deux vecteurs tangents, et d'en déduire un vecteur normal en prenant le produit vectoriel. Les vecteurs tangents à le paramétrage sont

$$T_\theta = \begin{pmatrix} 0 \\ -2\sin\theta \\ 2\cos\theta \end{pmatrix}, \quad T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{95.319}$$

Afin de savoir quels sont les vecteurs tangents *au point demandé*, il faut trouver à quels x et θ correspond le point $(-1,1,-\sqrt{3})$. Manifestement, $x=-1$. En ce qui concerne θ, nous avons

$$\begin{aligned}2\cos\theta &= 1 \\ 2\sin\theta &= -\sqrt{3}.\end{aligned} \tag{95.320}$$

Cela fait $\theta = -\pi/3$. Les vecteurs tangents (et donc les générateurs du plan tangent) sont alors

$$T_\theta(-1,-\frac{\pi}{3}) = \begin{pmatrix} 0 \\ \sqrt{3} \\ 1 \end{pmatrix}, \quad T_x(-1,-\frac{\pi}{3}) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{95.321}$$

Un vecteur normal au cylindre au point $\phi(-1,-\pi/3)$ est donné par

$$\begin{vmatrix} e_x & e_y & e_z \\ 0 & \sqrt{3} & 1 \\ 1 & 0 & 0 \end{vmatrix} = e_y - \sqrt{3}e_z = \begin{pmatrix} 0 \\ 1 \\ -\sqrt{3} \end{pmatrix}. \tag{95.322}$$

L'équation du plan tangent est donc de la forme

$$y - \sqrt{3}z + d = 0 \tag{95.323}$$

pour un certain d à fixer. Nous le fixons en imposant que ce plan passe par le point $(-1, 1, -\sqrt{3})$. Nous demandons donc que

$$1 - \sqrt{3}(-\sqrt{3}) + d = 0, \tag{95.324}$$

ce qui fait $d = -4$. Le plan est donc

$$y - \sqrt{3}z - 4 = 0. \tag{95.325}$$

Exercise 580 exoOutilsMath-0115

Soit la région $D \subset \mathbb{R}^2$ paramétrée par

$$\begin{aligned} \phi\colon [0,\pi] \times [1,2] &\to \mathbb{R}^2 \\ (x,h) &\mapsto \begin{pmatrix} x \\ \sin(x) + h \end{pmatrix}. \end{aligned} \tag{95.326}$$

Dessiner le domaine et calculer

$$\int_D f(x,y)dS \tag{95.327}$$

où $f(x,y) = y\cos(x)$.

corrOutilsMath-0115

Correction of the exercise 580

En ce qui concerne le domaine, pour chaque x, la coordonnée y varie de $\sin(x)+1$ à $\sin(x)+2$. Le domaine est donc celui de la figure 95.7.

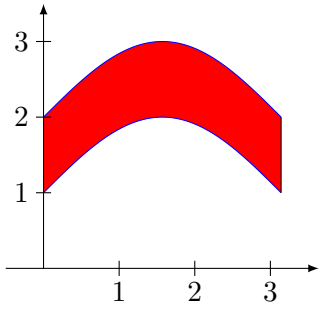


FIGURE 95.7: Le domaine d'intégration

LabelFigExSinLarge

La formule d'intégrale à utiliser est

$$\int_D f\, dS = \int_0^{\pi}\int_1^2 \tilde{f}(x,h)\|T_x \times T_h\|\, dh\, dx \tag{95.328}$$

où $T_x = (1, \cos(x))$ et $T_h = (0,1)$. Nous avons donc $T_x \times T_h = e_z$ et par conséquent $\|T_x \times T_h\| = 1$.

En ce qui concerne la fonction, nous devons exprimer $f(x,y) = y\cos(x)$ en termes de x et y. Le changement de variables $(x,y) \to (x,h)$ est donné par

$$\begin{cases} x = x & \text{(95.329a)} \\ y = \sin(x) + h, & \text{(95.329b)} \end{cases}$$

par conséquent

$$\tilde{f}(x,h) = \big(\sin(x) + h\big)\cos(x). \tag{95.330}$$

Nous restons donc avec l'intégrale

$$\int_D f\, dS = \int_0^{\pi}\int_1^2 \big(\sin(x) + h\big)\cos(x)\, dh\, dx \tag{95.331}$$

à calculer. Cette intégrale est nulle parce que la fonction de x à intégrer est antisymétrique par rapport à $\pi/2$. Si on ne voit pas cela, il y a moyen de calculer l'intégrale en suivant les mêmes pas que ceux décrits dans la correction de l'exercice 563.

Exercise 581 exoOutilsMath-0116

Soit le volume $V \subset \mathbb{R}^3$ donné par le paramétrage

$$\phi(r,\theta,z) = \begin{pmatrix} \frac{R}{h} zr\cos(\theta) \\ \frac{R}{h} zr\sin(\theta) \\ z \end{pmatrix} \tag{95.332}$$

où R et h sont des constantes positives et les domaines de variations des variables sont $r \in [0,1]$, $\theta \in [0,2\pi]$, $z \in [0,h]$.

Décrire le domaine et en calculer le volume. corrOutilsMath-0116

Correction of the exercise 581

Des trois variables r, θ et z, la plus facile à comprendre est z. Cette variables correspond au z usuel. Le problème revient donc à comprendre quelle forme géométrique nous avons à chaque hauteur z. Pour chaque z et r, nous avons un cercle de rayon

$$\frac{R}{h} zr \tag{95.333}$$

situé à la hauteur z. Lorsque r varie de 0 à 1, ce rayon varie de 0 à $\frac{R}{h}z$. Notre volume V consiste donc en un disque plein de rayon Rz/h à chaque hauteur. Lorsque $z = 0$, ce rayon est 0 et lorsque $z = h$, ce rayon devient R. Nous avons donc un cône de hauteur h posé sur la pointe.

Le volume est donné par

$$V = \int_0^1 dr \int_0^{2\pi} d\theta \int_0^h dz |T_z \cdot (T_r \times T_\theta)| \tag{95.334}$$

où

$$\begin{aligned} T_r &= \frac{\partial\phi}{\partial r} = \begin{pmatrix} \frac{R}{h} z\cos(\theta) \\ \frac{R}{h} z\sin(\theta) \\ 0 \end{pmatrix} \\ T_\theta &= \frac{\partial\phi}{\partial\theta} = \begin{pmatrix} -\frac{R}{h} zr\sin(\theta) \\ \frac{R}{h} zr\cos(\theta) \\ 0 \end{pmatrix} \\ T_z &= \frac{\partial\phi}{\partial z} = \begin{pmatrix} \frac{R}{h} r\cos(\theta) \\ \frac{R}{h} r\sin(\theta) \\ 1 \end{pmatrix}. \end{aligned} \tag{95.335}$$

Le calcul donne

$$\begin{aligned} T_z \cdot (T_r \times T_\theta) &= \begin{vmatrix} \frac{R}{h} z\cos\theta & \frac{R}{h} z\sin\theta & 0 \\ -\frac{R}{h} zr\sin\theta & \frac{R}{h} zr\cos\theta & 0 \\ \frac{R}{h} r\cos\theta & \frac{R}{h} r\sin\theta & 1 \end{vmatrix} \\ &= \frac{R^3}{h^3} \begin{vmatrix} z\cos\theta & z\sin\theta & 0 \\ -zr\sin\theta & zr\cos\theta & 0 \\ r\cos\theta & r\sin\theta & h/R \end{vmatrix} \\ &= \frac{R^2}{h^2}\big(z^2 r\cos^2(\theta) + z^2 r\sin^2(\theta)\big) \\ &= \frac{R^2}{h^2} z^2 r. \end{aligned} \tag{95.336}$$

L'intégrale qui donne le volume est au final :

$$V = \int_0^1 \int_0^{2\pi} \int_0^h \frac{R^2}{h^2} z^2 r\, dz\, d\theta\, dr = \frac{\pi R^2 h}{3}. \tag{95.337}$$

Exercise 582 exoOutilsMath-0117

Intégrer la fonction $f(x,y,z)=|z|$ sur l'ellipsoïde (creuse)

$$\frac{x^2}{4}+\frac{y^2}{16}+\frac{z^2}{36}=9. \tag{95.338}$$

À la différence de l'exercice 578, c'est ici une intégrale de surface. corrOutilsMath-0117

Correction of the exercise 582

Nous utilisons le paramétrage avec $\rho=3$, c'est-à-dire

$$\phi(\theta,\varphi)=\begin{pmatrix}6\sin(\theta)\cos(\varphi)\\12\sin(\theta)\sin(\varphi)\\18\cos(\theta)\end{pmatrix} \tag{95.339}$$

avec $\theta\in[0,\pi]$ et $\varphi\in[0,2\pi]$. La fonction à intégrer est $f\big(\phi(\theta,\varphi)\big)=18|\cos(\theta)|$. Nous écrivons donc

$$\int_\phi f=18\int_0^\pi\int_0^{2\pi}|\cos(\theta)||J_\phi|d\varphi d\theta. \tag{95.340}$$

Pour déterminer J_ϕ nous calculons le produit vectoriel

$$T_\theta\times T_\varphi=\begin{vmatrix}e_x & e_y & e_z\\6\cos(\theta)\cos(\varphi) & 12\cos(\theta)\sin(\varphi) & -18\sin(\theta)\\-6\sin(\theta)\sin(\varphi) & 12\sin(\theta)\cos(\varphi) & 0\end{vmatrix}=216\sin(\theta)\begin{pmatrix}\sin(\theta)\cos(\varphi)\\\sin(\theta)\sin(\varphi)\\\cos(\theta)\end{pmatrix}. \tag{95.341}$$

Le jacobien est la norme de cela :

$$\begin{aligned}\|T_\theta\times T_\varphi\|&=216\sin(\theta)\sqrt{\sin^2(\theta)\cos^2(\varphi)+\sin^2(\varphi)\sin^2(\theta)+\cos^2(\theta)} &(95.342a)\\&=216|\sin(\theta)| &(95.342b)\\&=216\sin(\theta). &(95.342c)\end{aligned}$$

Notons que nous avons enlevé la valeur absolue parce que $\theta\in[0,2\pi]$. L'intégrale à calculer est

$$I=18\cdot 216\int_0^\pi\int_0^{2\pi}|\cos(\theta)|\sin(\theta)d\varphi d\theta. \tag{95.343}$$

Nous n'enlevons pas la valeur absolue sur le cosinus parce que $\cos(\theta)$ n'est pas de signe constant sur $\theta\in[0,\pi]$. Nous devons donc couper l'intégrale en deux parties avant d'intégrer. Le gros de l'intégrale se calcule comme suit :

```
sage: f(x)=sin(x)*cos(x)
sage: f.integrate(x)
x |--> -1/2*cos(x)^2
sage: f.integrate(x,0,pi/2)-f.integrate(x,pi/2,pi)
x |--> 1
```

Si vous vous demandez comment on trouve une primitive de $\sin(\theta)\cos(\theta)$, pensez à faire par partie. En collant les bouts,

$$I=3888\cdot 2\pi=7776\pi. \tag{95.344}$$

Exercise 583 exoOutilsMath-0118

(1) Calculer l'intégrale double

$$\iint_\Omega(x^2y+\sqrt{x}y^2)\,dxdy \tag{95.345}$$

où $\Omega=[0,1]\times[1,2]$.

(2) Soit Ω la région du plan délimitée par les droites d'équations $y = 0$, $x = 0$, $x = 1$ et la courbe d'équation $y = x^2$. Calculer l'intégrale double

$$\iint_{\Omega} (x^2 + x)\sqrt{y}\, dxdy. \tag{95.346}$$

corrOutilsMath-0118

Correction of the exercise 583

(1) Lorsqu'on écrit $\Omega = [0, 1] \times [1, 2]$, cela signifie que x va de 0 à 1 et que y va de 1 à 2. L'intégrale à calculer est donc

$$\int_0^1 dx \int_1^2 dy(x^2y + \sqrt{x}y^2) = \frac{37}{18}. \tag{95.347}$$

(2) L'intervalle de x est

$$x\colon 0 \to 1. \tag{95.348}$$

Pour chaque x, la variable y est limitée par x^2, donc

$$y\colon 0 \to x^2. \tag{95.349}$$

D'où l'intégrale

$$\begin{aligned}\int_0^1 dx \int_0^{x^2} dy(x^2 + x)\sqrt{y} &= \int_0^1 dx(x^2 + x)\left[\frac{2}{3}y^{3/2}\right]_{y=0}^{y=x^2} \\ &= \frac{2}{3}\int_0^1 (x^2 + x)x^3 \\ &= \frac{11}{45}.\end{aligned} \tag{95.350}$$

Attention : il ne peut pas rester de x ou de y dans la réponse finale ! Ici l'ordre d'intégration est important.

Exercise 584 exoOutilsMath-0119

On considère la transformation

$$T\colon \begin{pmatrix} u \\ v \end{pmatrix} \in [1, 2] \times [1, 3] \to \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u^2 - v^2 \\ 2uv \end{pmatrix}. \tag{95.351}$$

Soit

$$\mathcal{R} = \left\{ T\begin{pmatrix} u \\ v \end{pmatrix};\ \begin{pmatrix} u \\ v \end{pmatrix} \in [1, 2] \times [1, 3] \right\} \tag{95.352}$$

l'image de T. On admettra que $T\colon [1, 2] \times [1, 3] \to \mathcal{R}$ est bijective.

(1) Calculer le déterminant de la matrice jacobienne de T.

(2) Calculer l'aire de $\mathcal{R}$ en utilisant un changement de variable (un paramétrage) que l'on précisera.

corrOutilsMath-0119

Correction of the exercise 584

(1) Le déterminant de la matrice jacobienne est donné par

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2u & -2v \\ 2v & 2u \end{vmatrix} = 4(u^2 + v^2). \tag{95.353}$$

(2) Nous paramétrons la surface par $\phi(u,v) = T(u,v)$. L'aire est donnée par

$$\int_1^2 du \int_1^3 dv \|T_u \times T_v\|. \tag{95.354}$$

La norme $\|T_u \times T_v\|$ est la jacobien que nous venons de calculer. D'où l'intégrale

$$Aire = \int_1^2 du \int_1^3 dv(4u^2 + 4v^2) = \frac{160}{3}. \tag{95.355}$$

Exercise 585 exoOutilsMath-0120

Soit

$$\Omega = \{(x,y); x \geqslant 0, y \geqslant 0, 1 \leqslant x^2 + y^2 \leqslant 4\}. \tag{95.356}$$

Dessiner Ω puis calculer

$$I = \iint_\Omega \frac{xy^2 + 1}{\sqrt{x^2 + y^2}} dxdy \tag{95.357}$$

en utilisant les coordonnées polaires. corrOutilsMath-0120

Correction of the exercise 585

La surface est représentée à la figure 95.8. C'est la surface contenue entre les cercles de rayon 1 et 2, dans le cadrant où x et y sont positifs.

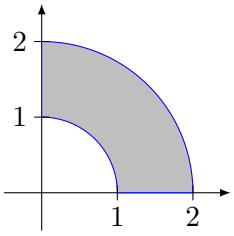


FIGURE 95.8: La surface de l'exercice 585. LabelFigCouroneExam

En coordonnées polaires cette surface est paramétrée par

$$\begin{aligned} r &: 1 \to 2 \\ \theta &: 0 \to \frac{\pi}{2}. \end{aligned} \tag{95.358}$$

En comptant le jacobien r des coordonnées polaires, l'intégrale à calculer est

$$I = \int_1^2 dr \int_0^{\pi/2} d\theta \frac{r^3 \cos(\theta)\sin^2(\theta) + 1}{r} r. \tag{95.359}$$

La partie de l'intégrale qui peut poser problème est celle sur θ :

$$\int_0^{\pi/2} \cos(\theta)\sin^2(\theta) d\theta. \tag{95.360}$$

On pose $u = \sin(\theta)$, et cela devient

$$\int_0^1 u^2 du = \frac{1}{3}. \tag{95.361}$$

Au final, le résultat est

$$I = \frac{\pi}{2} + \frac{5}{4}. \tag{95.362}$$

Exercise 586 exoOutilsMath-0121

Soit V la région de l'espace $\mathbb{R}^3$ déterminée par les conditions $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ et $0 \leqslant z \leqslant xy$.

(1) Calculer le volume de V.
(2) Calculer l'intégrale triple $\iiint_V (x+y)\,dxdydz$.

corrOutilsMath-0121

Correction of the exercise 586

Le volume que l'on considère est paramétré par

$$\begin{cases} x\colon 0 \to 1 & (95.363a)\\ y\colon 0 \to 1 & (95.363b)\\ z\colon 0 \to xy. & (95.363c)\end{cases}$$

(1) Le volume est alors donné par

$$V = \int_0^1 dx \int_0^1 dy \int_0^{xy} dz 1 = \int_0^1 dx \int_0^1 xy\,dy = \frac{1}{4}. \tag{95.364}$$

(2) L'intégrale à calculer est

$$\begin{aligned}\int_0^1 dx \int_0^1 dy \int_0^{xy} (x+y) &= \int_0^1 dx \int_0^1 dy (x+y)[z]_0^{xy}\\ &= \int_0^1 dx \int_0^1 xy(x+y)dy\\ &= \frac{1}{3}.\end{aligned} \tag{95.365}$$

Exercise 587 exoOutilsMath-0122

Soient $a>0$, $b>0$, $c>0$ et soit l'ellipsoïde

$$E = \left\{(x,y,z); \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \leqslant 1\right\}. \tag{95.366}$$

(1) On considère l'application

$$T\colon \begin{pmatrix}u\\v\\w\end{pmatrix} \in \mathbb{R}^3 \to \begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}au\\bv\\cw\end{pmatrix} \in \mathbb{R}^3. \tag{95.367}$$

Monter que lorsque (u,v,w) décrit la boule unité de $\mathbb{R}^3$, alors $T(u,v,w)$ décrit l'ellipsoïde E.

(2) Calculer le déterminant de la matrice jacobienne de T. Calculer le volume de l'ellipsoïde E.

corrOutilsMath-0122

Correction of the exercise 587

(1) Prouvons que T est une bijection entre la boule unité et E. Si $(x,y,z) \in E$, alors l'unique élément de $\mathbb{R}^3$ dont l'image par T est (x,y,z) est le vecteur $\begin{pmatrix}x/a\\y/b\\z/c\end{pmatrix}$. En effet,

$$T\begin{pmatrix}x/a\\y/b\\z/c\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}. \tag{95.368}$$

De plus si $(x,y,z) \in E$, on a

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \leqslant 1, \tag{95.369}$$

c'est-à-dire que le point $\begin{pmatrix}x/a\\y/b\\z/c\end{pmatrix}$ est dans la boule unité.

(2) La matrice jacobienne de T est la matrice

$$J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}. \tag{95.370}$$

Son déterminant vaut abc.

Afin de trouver le volume de E, nous pouvons utiliser les sphériques modifiées de l'exercice 578.

Une façon plus rapide est d'utiliser la formule du changement de variable du théorème 18.251. Par rapport aux notations de ce théorème, nous avons $f(x, y, z) = 1$. Nous avons donc

$$V = \int_E 1\, dxdudz = \int_{T(Boule)} 1\, dxdydz = \int_{Boule} |J_T(u, v, w)| dudvdw. \tag{95.371}$$

Ici nous venons de voir que $J_T(u, v, w)$ était la constante abc qui peut sortir de l'intégrale. Donc

$$V = abc \int_{Boule} 1\, dudvdw = abc\frac{4\pi}{3}. \tag{95.372}$$

La dernière intégrale était simplement le volume de la sphère.

Exercise 588 exoOutilsMath-0123

Soit $D = \{(u, v) \in \mathbb{R}^2; u^2 + v^2 \leqslant 1\}$ le disque unité centré à l'origine. Vérifier que la surface paramétrée

$$\Phi\colon (u, v) \in D \to \begin{pmatrix} u - v \\ u + v \\ uv \end{pmatrix} \in \mathbb{R}^3 \tag{95.373}$$

est régulière. Calculer l'aire de cette surface. corrOutilsMath-0123

Correction of the exercise 588

Au vu de la définition 20.58, il faut vérifier que les vecteurs T_u et T_v sont non nuls et non colinéaires. Les vecteurs tangents à le paramétrage sont vite calculés :

$$T_u = \begin{pmatrix} 1 \\ 1 \\ v \end{pmatrix} \quad T_v = \begin{pmatrix} -1 \\ 1 \\ u \end{pmatrix}. \tag{95.374}$$

Ils ne sont évidemment jamais nuls et ne sont jamais colinéaires parce qu'il n'existe pas de multiples de T_u qui soit égal à T_v.

La surface S sont nous devons calculer l'aire est l'image par Φ du disque unité. Le bon paramétrage de la surface est donc $\Phi\colon D \to S$. Nous devons donc calculer

$$Aire = \int_D \|T_u \times T_v\|\, dudv. \tag{95.375}$$

Un petit calcul montre que

$$T_u \times T_v = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 1 & v \\ -1 & 1 & u \end{vmatrix} = (u - v)e_x - (u + v)e_y + 2e_z. \tag{95.376}$$

Nous devons donc calculer l'intégrale de la fonction

$$\|T_u \times T_v\| = \sqrt{2u^2 + 2v^2 + 4} \tag{95.377}$$

sur le disque $D \equiv u^2 + v^2 \leqslant 1$. Nous faisons cela en polaires : nous posons $u = r\cos(\theta)$, $v = r\sin(\theta)$, et nous n'oublions pas le jacobien des polaires :

$$Aire = \int_0^{2\pi} d\theta \int_0^1 r\sqrt{2r^2 + 4} dr. \tag{95.378}$$

En posant $t = 2r^2 + 4$, nous avons

$$\begin{aligned} Aire &= 2\pi \int_4^6 \sqrt{t}\frac{dt}{4} \\ &= \frac{\pi}{2}\int_4^6 \sqrt{t}dt \\ &= \frac{\pi}{3}(6\sqrt{6} - 8). \end{aligned} \tag{95.379}$$

Exercise 589 exoOutilsMath-0124
On considère la fonction

$$f\colon (x, y, z) \in \mathbb{R}^3 \to x^2yz + 2y^2\sin(xyz). \tag{95.380}$$

(1) Calculer les dérivées partielles premières de f.
(2) Quelle est la différentielle de f au point $(1, 1, \pi)$?
(3) Donner une approximation de $f(1 + 10^{-2}, 1 - 10^{-3}, \pi + 10^{-4})$.

corrOutilsMath-0124

Correction of the exercise 589

(1) En oubliant pas la règle de Leibniz, les dérivées partielles sont données par Subeqsderpartzutz

$$\frac{\partial f}{\partial x} = 2y^3z\cos(xyz) + 2xyz \tag{95.381a}$$
$$\frac{\partial f}{\partial y} = 2xy^2z\cos(xyz) + x^2z + 4y\sin(xyz) \tag{95.381b}$$
$$\frac{\partial f}{\partial z} = 2xy^3\cos(xyz) + x^2y. \tag{95.381c}$$

(2) Étant donné que ces dérivées partielles sont continues, la différentielle de f au point $(1, 1, \pi)$ est l'application

$$df_{(1,1,\pi)}(u_1, u_2, u_4) = \frac{\partial f}{\partial x}(1, 1, \pi)u_1 + \frac{\partial f}{\partial y}(1, 1, \pi)u_2 + \frac{\partial f}{\partial z}(1, 1, \pi)u_3. \tag{95.382}$$

Ici en remplaçant $x = 1$, $y = 1$ et $z = \pi$ dans les équations (95.381) nous trouvons

$$\frac{\partial f}{\partial x}(1, 1, \pi) = 0 \tag{95.383a}$$
$$\frac{\partial f}{\partial y}(1, 1, \pi) = -\pi \tag{95.383b}$$
$$\frac{\partial f}{\partial z}(1, 1, \pi) = -1 \tag{95.383c}$$
$$\tag{95.383d}$$

(3) En ce qui concerne l'approximation nous avons

$$f(1 + 10^{-2}, 1 - 10^{-3}, \pi + 10^{-4} \simeq f(1, 1, \pi) + 10^{-3}\pi - 10^{-4} \tag{95.384}$$

avec $f(1, 1, \pi) = \pi$.

Exercise 590 exoOutilsMath-0125
On considère la fonction

$$f\colon (x, y) \in \mathbb{R}^2 \to (x - iy)^2 \tag{95.385}$$

où $i = \sqrt{-1}$.

(1) Vérifier que les parties réelles et imaginaires de $f(x,y)$ sont données par

$$\begin{aligned} Re\, f(x,y) &= x^2 - y^2 \\ Im\, f(x,y) &= -2xy. \end{aligned} \tag{95.386}$$

(2) Montrer que le champ

$$F(x,y) = \begin{pmatrix} Re\, f(x,y) \\ Im\, f(x,y) \end{pmatrix} \tag{95.387}$$

est à divergence nul et rotationnel nul.

corrOutilsMath-0125

Correction of the exercise 590

Nous avons

$$f(x,y) = x^2 - 2ixy + (iy)^2 = x^2 - y^2 - 2ixy. \tag{95.388}$$

La partie réelle est donc bien $x^2 - y^2$ et la partie imaginaire est $2xy$. Le champ de vecteurs que nous regardons est donc

$$F(x,y) = \begin{pmatrix} x^2 - y^2 \\ -2xy \end{pmatrix}. \tag{95.389}$$

La divergence est

$$\nabla \cdot F = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} = 2x - 2x = 0. \tag{95.390}$$

Le rotationnel est

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ x^2 - y^2 & -2xy & 0 \end{vmatrix} = (-2y + 2y)e_z = 0. \tag{95.391}$$

Exercise 591 exoOutilsMath-0126

On considère la courbe paramétrée (cycloïde)

$$\sigma\colon t \in [0, 2\pi] \to \sigma(t) = \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix}. \tag{95.392}$$

(1) Déterminer le vecteur vitesse $\sigma'(t)$.

(2) Calculer la norme euclidienne de $\sigma'(t)$.

(3) Calculer la longueur de la courbe paramétrée. (on pourra utiliser l'identité $1 - \cos(2t) = 2\sin^2(t/2)$).

corrOutilsMath-0126

Correction of the exercise 591

Le vecteur vitesse est la dérivée composante par composante :

$$\sigma'(t) = \begin{pmatrix} 1 - \cos(t) \\ \sin(t) \end{pmatrix}. \tag{95.393}$$

Nous avons

$$\|\sigma'(t)\|^2 = 1 + \cos^2(t) - 2\cos(t) + \sin^2(t) = 2(1 - \cos(t)) = 4\sin^2\left(\frac{t}{2}\right), \tag{95.394}$$

et par conséquent

$$\|\sigma'(t)\| = 2\sin\left(\frac{t}{2}\right). \tag{95.395}$$

La longueur est l'intégrale de cela :

$$\begin{aligned} l(\sigma) &= \int_0^{2\pi} 2\sin\left(\frac{t}{2}\right) dt \\ &= 4\int_0^{2\pi} \sin\left(\frac{t}{2}\right)\frac{dt}{2} \\ &= 4\int_0^{\pi} \sin(u)du \\ &= -4[\cos(u)]_{u=0}^{u=\pi} \\ &= 8 \end{aligned} \tag{95.396}$$

où nous avons posé $u = t/2$.

Exercise 592 exoOutilsMath-0127

Soit F le champ de vecteurs

$$F\colon (x,y,z) \in \mathbb{R}^3 \to \begin{pmatrix} yz+3x^2 \\ xz+3y^2 \\ xy+3z^2 \end{pmatrix} \in \mathbb{R}^3. \tag{95.397}$$

(1) Montrer que F dérive d'un potentiel scalaire que l'on déterminera.

(2) En déduire la circulation du champ F le long d'un chemin joignant les points $(0,1,2)$ et $(1,0,2)$.

corrOutilsMath-0127

Correction of the exercise 592

On vérifie facilement que la fonction suivante est un potentiel :

$$f(x,y,z) = xyz + 3(x^2+y^2+z^2). \tag{95.398}$$

La circulation d'un champ de vecteur dérivant d'un potentiel le long d'un chemin est égale à la différence de potentiel aux extrémités. Ici nous avons

$$W = f(0,1,2) - f(1,0,2) = 15 - 15 = 0. \tag{95.399}$$

Exercise 593 exoOutilsMath-0128

(1) Vérifier que $y\ln(y) - y$ est une primitive de $\ln(y)$.

(2) Calculer l'intégrale double

$$\iint_D x\ln(y)\,dxdy \tag{95.400}$$

où

$$D = \{(x,y) \text{ tel que } x \in [0,1], e^x \leqslant y \leqslant e^{2x}\}. \tag{95.401}$$

ImtemzuhtE

(3) Soient $0 < a < b$ donnés. Dessiner le *quart de couronne* de la partie $\{(x,y)$ tel que $x \geqslant 0, y \geqslant 0\}$ du plan et situé entre les cercles de rayon a et b.

(4) Calculer, en passant aux coordonnées polaires, l'intégrale double

$$\iint_E \ln(x^2+y^2)\,dxdy \tag{95.402}$$

où E est l'ensemble dessiné au point (3).

corrOutilsMath-0128

Correction of the exercise 593

(1) La vérification se fait en dérivant.

(2) Le calcul de l'intégrale est

$$\begin{aligned} I &= \int_0^1 dx \int_{e^x}^{e^{2x}} x\ln(y)dy \\ &= \int_0^1 (2x^2e^{2x} - xe^{2x} - x^2e^x + xe^x). \end{aligned} \tag{95.403}$$

Calculons séparément les quatre primitives à savoir pour cette intégrale. Elles se font toutes par partie selon le modèle $A = \int xe^x dx$. On pose $u = x$, $dv = e^x$, et nous avons

$$A = \int xe^x dx = xe^x - \int e^x = xe^x - e^x. \tag{95.404}$$

Ensuite,

$$\int x^2e^x = x^2e^x - 2\int xe^x = x^2e^x - 2xe^x + 2e^x. \tag{95.405}$$

De la même façon,

$$\int xe^{2x}dx = e^{2x}\left(\frac{x}{2} - \frac{1}{4}\right) \tag{95.406}$$

et

$$\int x^2e^{2x}dx = \frac{x^2}{2}e^{2x} - \int 2x\frac{1}{2}e^{2x} = e^{2x}\left(\frac{x^2}{2} - \frac{x}{2} + \frac{1}{4}\right). \tag{95.407}$$

En remettant les bouts ensemble, nous trouvons

$$I = \frac{e^2}{4} - e + \frac{9}{4}. \tag{95.408}$$

(3) Le dessin est à la figure 95.9.

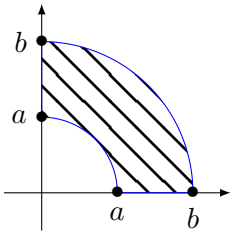


FIGURE 95.9: Le dessin de l'exerice 593. LabelFigratrap

(4) L'intégrale à calculer (en n'oubliant pas le jacobien) est

$$\int_a^b dr \int_0^{2\pi} d\theta r\ln(r^2) = 2\pi\int_a^b \ln(r^2)rdr \tag{95.409a}$$

$$= 2\pi\frac{1}{2}\int_{a^2}^{b^2}\ln(u)du \tag{95.409b}$$

$$= \pi\left[u\ln(u) - u\right]_{a^2}^{b^2}. \tag{95.409c}$$

Exercise 594 exoOutilsMath-0129

(1) Calculer le jacobien de la transformation

$$\begin{pmatrix} r \\ \theta \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ z \end{pmatrix}. \tag{95.410}$$

(2) En déduire la valeur de l'intégrale triple

$$\int_D (x^2+y^2+z^2)dxdydz \tag{95.411}$$

où D est le cylindre

$$D = \{(x,y,z); x^2+y^2 \leqslant 1, 0 \leqslant z \leqslant 1\}. \tag{95.412}$$

corrOutilsMath-0129

Correction of the exercise 594

La transformation proposée est celle des coordonnés cylindriques. Pour le jacobien, le calcul est dans le cours, et le résultat est $J = r$. En ce qui concerne le calcul de l'intégrale,

$$I = \int_0^1 dz \int_0^{2\pi} d\theta \int_0^1 dr(r^2+z^3)r \tag{95.413a}$$

$$= 2\pi \int_0^2 dz \left[\frac{r^4}{4} + z^3\frac{r^2}{2}\right] \tag{95.413b}$$

$$= 2\pi \left[\frac{z}{4} + \frac{z^4}{>8}\right]_0^1 \tag{95.413c}$$

$$= \frac{3\pi}{4}. \tag{95.413d}$$

Exercise 595 exoOutilsMath-0130

Est-ce qu'il existe une fonction f telle que

$$\nabla f = \begin{pmatrix} \cos(y) \\ -x\sin(y) \end{pmatrix} \tag{95.414}$$

corrOutilsMath-0130

Correction of the exercise 595

<+130+>

Exercise 596 exoOutilsMath-0131

Soit la fonction $f(x,y) = x^2\sin(y) + e^{\cos(x)}$.

(1) Calculer $df_a(u)$ avec $a = (\pi, -\pi/4)$ et $u = (1,5)$.

(2) Calculer $\nabla \cdot (\nabla f)(a)$.

corrOutilsMath-0131

Correction of the exercise 596

Les dérivées partielles sont classiques. En ce qui concerne $\nabla \cdot (\nabla f)(a)$, il faut comprendre qu'il y a implicitement des parenthèses de la façon suivante :

$$(\nabla \cdot (\nabla f))(a). \tag{95.415}$$

Cela donne

$$\nabla \cdot \nabla f = \nabla \cdot \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}, \tag{95.416}$$

c'est-à-dire que nous devons calculer la somme des dérivées secondes.

Voici comment on peut faire les calculs avec Sage.

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: f(x,y)=x**2*sin(y)+exp(cos(x))
sage: f.diff(x)
(x, y) |--> 2*x*sin(y) - e^cos(x)*sin(x)
```

```
sage: f.diff(y)
(x, y) |--> x^2*cos(y)

sage: f.diff(x)(pi,-pi/4)
-pi*sqrt(2)
sage: f.diff(y)(pi,-pi/4)
1/2*pi^2*sqrt(2)

sage: f.diff(x)(pi,-pi/4)+5*f.diff(y)(pi,-pi/4)
5/2*pi^2*sqrt(2) - pi*sqrt(2)

sage: D=f.diff(x,2)+f.diff(y,2)
sage: D
(x, y) |--> -x^2*sin(y) + e^cos(x)*sin(x)^2 - e^cos(x)*cos(x) + 2*sin(y)
sage: D(pi,-pi/4)
1/2*pi^2*sqrt(2) - sqrt(2) + e^(-1)
```

Vous noterez la notation `f.diff(x,2)` pour calculer la dérivée seconde de f par rapport à x.

Les résultats sont

$$df_a(u) = \frac{5}{2}\pi^2\sqrt{2} - \pi\sqrt{2}, \tag{95.417}$$

puis

$$\nabla \cdot (\nabla f) = -x^2\sin(y) + e^{\cos(x)}\sin^2(x) - e^{\cos(x)}\cos(x) + 2\sin(y). \tag{95.418}$$

En calculant cela au point a nous trouvons

$$\nabla \cdot (\nabla f)(a) = \frac{1}{2}\pi^2\sqrt{2} + e^{-1}. \tag{95.419}$$

Exercise 597 exoOutilsMath-0132

Soit en coordonnées polaires $F(r,\theta) = \frac{1}{r^2}e_r$. Calculer en coordonnées polaires l'intégrale $\int_C F$ où C est le cercle de rayon R centré à l'origine.

Calculer $\nabla \times F$ et $\nabla \cdot F$. corrOutilsMath-0132

Correction of the exercise 597

Ce champ de vecteurs est donné en coordonnées polaires. Pour l'intégrale, si γ est le chemin qui parcourt le cercle,

$$\int_\gamma F = \int_{0J^{2\pi}} F\big(\gamma(t)\big) \cdot \gamma'(t)dt = 0 \tag{95.420}$$

parce que le champ de vecteurs est toujours dans la direction e_r (c'est-à-dire radial) alors que γ' est toujours dans la direction e_θ qui lui est perpendiculaire.

En ce qui concerne le rotationnel et la divergence, les résultats sont

$$\nabla \times F = 0 \tag{95.421a}$$

$$\nabla \cdot F = \frac{1}{r}\left(\frac{\partial}{\partial r}\frac{1}{r}\right) = -\frac{1}{r^3}. \tag{95.421b}$$

Exercise 598 exoOutilsMath-0133

Soient F, un champ de vecteur, x, u des éléments de $\mathbb{R}^3$ et f, une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$. Déterminer si les choses suivantes sont des vecteurs, des fonctions, des champs de vecteurs, des nombres ou autre chose :

(1) $\nabla \times F(x)$

(2) $f\nabla \cdot F$

(3) $\frac{\partial f}{\partial u}$.

corrOutilsMath-0133

Correction of the exercise 598

(1) Il faut comprendre le parenthèsage implicite $(\nabla \times F)(x)$ parce que $\nabla \times (F(x))$ n'a pas de sens. $\nabla \times F$ est un champ de vecteurs ; donc $\nabla \times F(x)$ est un vecteur.

(2) $\nabla \times F$ est une fonction ; le produit de deux fonctions est une fonction. Le tout est une fonction.

(3) $\frac{\partial f}{\partial u}$ est une fonction.

Exercise 599 exoOutilsMath-0134

Soit

$$F(x,y) = \begin{pmatrix} y + \cos(y) \\ -x\sin(y) + x \\ 0 \end{pmatrix}. \tag{95.422}$$

(1) Calculer $\nabla \times F$

(2) Soit $\sigma(t) = \big(t, \sin(t^2), 1\big)$ avec $t \in [0, \sqrt{\pi}]$. Calculer $\int_\sigma F$.

corrOutilsMath-0134

Correction of the exercise 599

Le calcul du rotationnel donne immédiatement $\nabla \times F = 0$. En ce qui concerne l'intégrale, le plus simple est de chercher un potentiel (qui existe parce que le rotationnel est nul). Nous trouvons le potentiel

$$f(x,y,z) = xy + x\cos(y). \tag{95.423}$$

Nous avons alors

$$\int_\sigma F = f(\sqrt{\pi}, 0, 1) - f(0,0,1) = \sqrt{\pi}. \tag{95.424}$$

Exercise 600 exoOutilsMath-0135

Soit en coordonnées cylindriques la fonction $f(r,\theta,z) = rz^2$. Calculer l'intégrale de f sur le volume délimité par le cylindre de rayon R et de hauteur h.

corrOutilsMath-0135

Correction of the exercise 600

Les bornes du cylindre (plein) en coordonnées cylindriques sont :

$$\begin{cases} r : 0 \to R & \text{(95.425a)} \\ z : 0 \to h & \text{(95.425b)} \\ \theta : 0 \to 2\pi. & \text{(95.425c)} \end{cases}$$

L'intégrale à calculer est donc, en n'oubliant pas le jacobien,

$$\int_0^R dr \int_0^h dz \int_0^{2\pi} d\theta r^2 z^2 = \frac{2}{9}\pi R^3 h^3 \tag{95.426}$$

Exercise 601 exoOutilsMath-0136

Soit la transformation $T : \mathbb{R}^3 \to \mathbb{R} \times \mathbb{R} \times]-\pi/2, \pi/2[$ donnée par

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} xy \\ x + y \\ \arctan(z) \end{pmatrix} \tag{95.427}$$

Donner la matrice jacobienne de cette transformation et déterminer si T est une bijection ou non.

corrOutilsMath-0136

Correction of the exercise 601

La matrice jacobienne se calcule facilement :

$$\begin{pmatrix} y & x & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \frac{1}{1+z^2} \end{pmatrix}. \tag{95.428}$$

L'application T n'est pas une bijection parce que ce n'est pas injectif. Par exemple $T(1,2,0) = T(2,1,0)$. Plus généralement vu que x et y ont des rôles symétriques dans T, nous avons toujours $T(x,y,z) = T(y,x,z)$.

Beaucoup d'étudiants ont pris l'initiative de calculer le déterminant de la matrice jacobienne, bien que ce n'était pas demandé. Le voici :

$$J_T = \det \begin{pmatrix} y & x & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \frac{1}{1+z^2} \end{pmatrix} = \frac{y-x}{1+z^2}. \tag{95.429}$$

Bien entendu, ceux qui se sont vautrés dans le calcul ont perdu des points : si vous prenez l'initiative de me montrer que vous n'êtes pas capable de calculer un déterminant, j'apprécie l'initiative à sa juste valeur.

Exercise 602 exoOutilsMath-0137

Calculer la matrice jacobienne de

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} xe^{yz} \\ ye^{zx} \\ ze^{xy} \end{pmatrix}. \end{aligned} \tag{95.430}$$

corrOutilsMath-0137

Correction of the exercise 602

<+CorrOutilsMath-0137+>

Exercise 603 exoOutilsMath-0138

On note $\|X\|$ la norme euclidienne du vecteur $X = (x,y,z) \in \mathbb{R}^3$. Calculer le gradient ∇f_{X_0} où $X_0 = (x_0, y_0, z_0) \neq (0,0,0)$ de l'application

$$\begin{aligned} f\colon \mathbb{R}^3 \backslash \{(0,0,0)\} &\to \mathbb{R} \\ X &\mapsto f(X) = \frac{2\sqrt{2}}{\|X\|}. \end{aligned} \tag{95.431}$$

corrOutilsMath-0138

Correction of the exercise 603

<+CorrOutilsMath-0138+>

Exercise 604 exoOutilsMath-0139

Soit le plan $x+y+z=1$. Intégrer la fonction $f(x,y,z) = xy^2$ sur la portion de ce plan donnée par $|x| < 2$ et $-x < y < x$.

corrOutilsMath-0139

Correction of the exercise 604

La surface est paramétrée par

$$\phi(x,y) = \begin{pmatrix} x \\ y \\ 1-x-y \end{pmatrix}, \tag{95.432}$$

les bornes étant données par

$$\begin{cases} x\colon 0 \to 2 & \text{(95.433a)} \\ y\colon -x \to x. & \text{(95.433b)} \end{cases}$$

Pourquoi x part de zéro au lieu de -2 ? Parce que pour $x < 0$, il n'y a pas de y vérifiant la condition $-x < y < x$. L'intégrale à calculer est celle de la fonction f sur la surface paramétrée par ϕ, c'est-à-dire

$$\int_\phi f = \int_0^2 dx \int_{-x}^{x} dy\, xy^2 \| \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \|. \tag{95.434}$$

L'élément de surface est calculé par

$$dS = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{95.435}$$

La norme vaut $\sqrt{3}$ et l'intégrale recherchée est

$$\int_\phi f = \sqrt{3}\int_0^2 \int_{-x}^{x} xy^2 dydx = \sqrt{3}\frac{64}{15}. \tag{95.436}$$

Exercise 605 exoOutilsMath-0140
Calculer l'intégrale de $f(x,y) = y\sin(x)$ sur le rectangle $[2,3]\times[0,\pi]$. corrOutilsMath-0140

Correction of the exercise 605
L'intégrale à calculer est

$$\int_2^3 \int_0^\pi y\sin(x)dydx = \frac{\pi^2}{2}\big(\cos(2)-\cos(3)\big). \tag{95.437}$$

Exercise 606 exoOutilsMath-0141
Calculer $\nabla \times F$ avec

$$F(x,y,z) = \begin{pmatrix} x+y \\ \sin(z) \\ e^{yz} \end{pmatrix}. \tag{95.438}$$

Est-ce que F dérive d'un potentiel ? corrOutilsMath-0141

Correction of the exercise 606
Le rotationnel vaut

$$\nabla \times F = \begin{vmatrix} e_x & e_y & e_z \\ \partial_x & \partial_y & \partial_z \\ x+y & \sin(z) & e^{yz} \end{vmatrix} = \begin{pmatrix} ze^{yz}-\cos(z) \\ 0 \\ -1 \end{pmatrix}. \tag{95.439}$$

Ce rotationnel étant non nul, le champ ne dérive pas d'un potentiel.

Exercise 607 exoOutilsMath-0142
Soit le domaine D défini par

$$0 \leqslant x+y \leqslant 5 \tag{95.440a}$$
$$-2 \leqslant x-y \leqslant 3. \tag{95.440b}$$

Calculer $\int_D (x^2-y^2)dxdy$.

Note : x^2-y^2 est un produit remarquable. L'auteur de ces lignes vous recommande chaudement d'y penser et d'effectuer un changement de variables (ou de choisir un paramétrage) sensé. corrOutilsMath-0142

Correction of the exercise 607
Le changement de variable qui s'invite est

$$\begin{cases} u = x+y & \text{(95.441a)} \\ v = x-y. & \text{(95.441b)} \end{cases}$$

Ce qui nous intéresse est surtout d'avoir x et y en fonction de u et v, c'est-à-dire

$$\begin{cases} x = \dfrac{u+v}{2} & \text{(95.442a)} \\ y = \dfrac{u-v}{2} & \text{(95.442b)} \end{cases}$$

Nous utilisons donc le paramétrage

$$\phi(u,v)=\begin{pmatrix}\frac{u+v}{2}\\ \frac{u-v}{2}\end{pmatrix} \tag{95.443}$$

et les bornes $u\colon 0\to 5$, $v\colon -2\to 3$. Le jacobien en est donné par la norme du vecteur

$$\frac{\partial\phi}{\partial u}\times\frac{\partial\phi}{\partial v}=\begin{vmatrix}e_x & e_y & e_z\\ \frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & -\frac{1}{2} & 0\end{vmatrix}=\begin{pmatrix}0\\0\\-1/2\end{pmatrix}. \tag{95.444}$$

Nous avons aussi $f\big(\phi(u,v)\big)=uv$ par le produit remarquable $x^2-y^2=(x+y)(x-y)=uv$. Bref, l'intégrale à calculer est

$$\int_\phi f=\int_0^5 du\int_{-2}^3 dv\frac{1}{2}(uv)=\frac{125}{8}. \tag{95.445}$$

Exercise 608 exoOutilsMath-0143 Épreuve complémentaire mai 2012.

On considère le champ de vecteurs

$$F(x,y)=\begin{pmatrix}2xy+y^2\\ 2xy+x^2\end{pmatrix}. \tag{95.446}$$

(1) Calculer directement (en utilisant seulement la définition) la circulation de F le long du chemin

$$\sigma\colon f\in[0,1]\to\begin{pmatrix}t\\ t^2\end{pmatrix}. \tag{95.447}$$

(2) Montrer que F dérive d'un potentiel scalaire que l'on précisera.

(3) Retrouver alors la valeur de la circulation de F le long du chemin σ. corrOutilsMath-0143

Correction of the exercise 608

Nous avons d'abord

$$\sigma'(t)=\begin{pmatrix}1\\2t\end{pmatrix}, \tag{95.448}$$

et ensuite

$$F\big(\sigma(t)\big)=F(t,t^2)=\begin{pmatrix}2t^3+t^4\\ 2t^3+t^2\end{pmatrix}. \tag{95.449}$$

Du coup l'intégrale demandée vaut

$$\int_0^1 F\big(\sigma(t)\big)\cdot\sigma'(t)dt=\int_0^1\begin{pmatrix}2t^3+t^4\\ 2t^3+t^2\end{pmatrix}\cdot\begin{pmatrix}1\\2t\end{pmatrix}dt=\int_0^1(2t^3+t^4)+2t(2t^3+t^2)dt=2. \tag{95.450}$$

En ce qui concerne le potentiel, nous devons trouver une fonction $f(x,y)$ telle que

$$\begin{cases}\frac{\partial f}{\partial x}=2xy+y^2 & (95.451a)\\ \frac{\partial f}{\partial y}=2xy+x^2. & (95.451b)\end{cases}$$

En intégrant la première il est vite vu que la fonction

$$f(x,y)=x^2y+xy^2 \tag{95.452}$$

fait l'affaire. L'intégrale vaut alors

$$\int_\sigma f=f\big(\sigma(1)\big)-f\big(\sigma(0)\big)=f(1,1)-f(0,0)=2-0=2. \tag{95.453}$$

Cela confirme le résultat obtenu plus haut.

Exercise 609 exoOutilsMath-0144 Épreuve complémentaire mai 2012.

(1) Calculer l'intégrale double

$$\iint_{\Omega}(x^3y^2+x^2y^3)dxdy \tag{95.454}$$

où $\Omega=[0,1]\times[0,2]$.

(2) Soit Δ le triangle délimité par les droites d'équations respectives $x=0$, $y=0$ et $y=-x+1$. Représenter Δ puis calculer l'intégrale double

$$\iint_{\Delta}x^3ydxdy. \tag{95.455}$$

corrOutilsMath-0144

Correction of the exercise 609

La première intégrale est très simple. Il faut intégrer

$$\int_0^2\left(\int_0^1(x^3y^2+x^2y^3)dx\right)dy. \tag{95.456}$$

Attention au fait que les bornes $[0,2]$ sont bien pour l'intégrale vis-à-vis de y, et non de x.

```
sage: f(x,y)=x**3*y**2+x**2*y**3
sage: f.integrate(x,0,1).integrate(y,0,2)
(x, y) |--> 2
```

La réponse est donc 2.

Pour la seconde, le domaine est le triangle de sommets $(0,1)$, $(0,0)$ et $(1,0)$. Les bornes sont donc

$$\begin{cases} x\colon 0\to 1 & \text{(95.457a)}\\ y\colon 0\to 1-x. & \text{(95.457b)}\end{cases}$$

L'intégrale à calculer est alors

$$\int_0^1 dx\left(\int_0^{1-x}x^3y\,dy\right). \tag{95.458}$$

Notez l'importance de l'ordre des bornes. Nous avons

```
sage: f(x,y)=x**3*y
sage: f.integrate(y,0,1-x).integrate(x,0,1)
(x, y) |--> 1/120
```

La réponse est donc 1/120.

Exercise 610 exoOutilsMath-0145

Épreuve complémentaire de mai 2012.

Soit

$$D=\{(x,y)\in\mathbb{R}^2 \text{ tel que } x\geqslant 0, y\geqslant 0, 1\leqslant x^2+y^2\leqslant 4\}. \tag{95.459}$$

(1) Représenter D.

(2) Calculer l'intégrale double

$$\iint_{D}\cos(x^2+y^2)dxdy \tag{95.460}$$

en utilisant les coordonnées polaires.

corrOutilsMath-0145

Correction of the exercise 610

Le domaine est la couronne de rayon interne 1 et de rayon externe 2 prise dans le quart de plan. Les bornes polaires sont donc

$$\begin{cases} r\colon 1 \to 2 & \text{(95.461a)} \\ \theta\colon 0 \to \dfrac{\pi}{2}, & \text{(95.461b)} \end{cases}$$

et en n'oubliant pas le jacobien r, l'intégrale à calculer est

$$\int_0^{\pi/2} d\theta \int_1^2 r\cos(r^2)dr. \tag{95.462}$$

L'intégrale sur θ revient à multiplier par $\frac{\pi}{2}$. En ce qui concerne celle sur r, un bon changement de variable est $u = r^2$ avec quoi il reste

$$\frac{\pi}{2}\int_1^4 \cos(u)\frac{du}{2} = \frac{\pi}{4}\big(\sin(4) - \sin(1)\big). \tag{95.463}$$

Exercise 611 exoOutilsMath-0146

Examen mai 2012.

On considère la courbe paramétrée (la cycloïde)

$$\sigma\colon t \in [0, 2\pi] \to \sigma(t) = \begin{pmatrix} t - \sin t \\ 1 - \cos t \end{pmatrix}. \tag{95.464}$$

(1) Déterminer le vecteur vitesse $\sigma'(t)$

(2) Calculer la norme euclidienne de $\sigma'(t)$

(3) Calculer la longueur de la courbe paramétrée. (on pourra utiliser l'identité $1 - \cos t = 2\sin^2(\frac{t}{2})$)

corrOutilsMath-0146

Correction of the exercise 611

Le vecteur vitesse est le vecteur dérivé :

$$\sigma'(t) = \begin{pmatrix} 1 - \cos t \\ \sin(t) \end{pmatrix}. \tag{95.465}$$

La norme est

$$\begin{aligned} \|\sigma'(t)\| &= \sqrt{(1-\cos t)^2 + \sin^2(t)} && \text{(95.466a)} \\ &= \sqrt{2 - 2\cos t} && \text{(95.466b)} \\ &= \sqrt{2}\sqrt{1 - \cos(t)} && \text{(95.466c)} \\ &= 2\sqrt{\sin^2\left(\frac{t}{2}\right)} && \text{(95.466d)} \\ &= 2\sin\left(\frac{t}{2}\right) && \text{(95.466e)} \end{aligned}$$

où nous avons utilisé la formule rappelée dans l'énoncé. Nous notons que $\sin(t/2)$ est positif lorsque $t \in [0, 2\pi]$; c'est pour cela que nous avons pu enlever la racine carrée sans mettre de valeur absolue dans la dernière ligne. En ce qui concerne la longueur nous avons alors

$$l(\sigma) = \int_0^{2\pi} 2\sin(t/2)dt = 8. \tag{95.467}$$

Exercise 612 exoOutilsMath-0147

Examen mai 2012.

On considère les trois points du plan $A = (1, 1)$, $B = (2, 3)$ et $C = (5, 2)$.

(1) Calculer les composantes des vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$.
(2) Calculer l'aire du parallélogramme déterminé par ces deux vecteurs.

corrOutilsMath-0147

Correction of the exercise 612

Les vecteurs demandés sont

$$\overrightarrow{AB} = B - A = (2,3) - (1,1) = (1,2) \tag{95.468a}$$
$$\overrightarrow{AC} = (5,2) - (1,1) = (4,1). \tag{95.468b}$$

L'aire du parallélogramme est donnée par la norme du produit vectoriel :

$$(1,2) \times (4,1) = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 2 & 0 \\ 4 & 1 & 0 \end{vmatrix} = -7e_z. \tag{95.469}$$

L'aire demandée vaut donc 7.

95.2 Autres retrouvés

Exercise 613 exoDerive-0003

Montrer que la tangente à un cercle est perpendiculaire au rayon.

corrDerive-0003

Correction of the exercise 613

Nous avons déjà vu dans la correction de l'exercice 616 que pour un cercle

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \tag{95.470}$$

Un point général du cercle a pour abscisse $x = R\cos(\theta)$. En remplaçant nous trouvons le coefficient directeur suivant pour la tangente :

$$y'\big(R\cos(\theta)\big) = -\frac{1}{\tan(\theta)}. \tag{95.471}$$

Par conséquent une droite perpendiculaire à la tangente aurait comme coefficient directeur le nombre $\tan(\theta)$. Or cela est bien le coefficient directeur du rayon qui joint le point $(0,0)$ au point $\big(R\cos(\theta), R\sin(\theta)\big)$.

Exercise 614 exoDerive-0001

Donner la dérivée des fonctions suivantes :

(1) $x^2 - x^4$;
(2) $\cos(2x)$
(3) $\frac{x^2-2}{x+1}$;
(4) $\frac{x+1}{x^3}$;
(5) $\sin x$;
(6) $\tan(x)$;
(7) $x^2 + \frac{1}{x}$;
(8) 2^x ;
(9) $\frac{x}{\sqrt{x^2+1}}$;
(10) $\frac{1}{x^2-4}$;

(11) $\dfrac{1}{|x^2-4|}$.

corrDerive-0001

Correction of the exercise 614

Le programme suivant par Sage résout l'exercice :

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

var('x')
f=x**2-x**4
print f.diff(x)
f=cos(2*x)
print f.diff(x)
f=(x**2-2)/(x+1)
print f.diff(x)
f=(x+1)/(x**3)
print f.diff(x)
f=tan(x)
print f.diff(x)
f=x**2+1/x
print f.diff(x)
f=2**x
print f.diff(x)
f=x/sqrt(x**2+1)
print f.diff(x)
f=sin(x)
print f.diff(x)
f=1/(x**2-4)
print f.diff(x)
f=1/abs(x**2-4)
print f.diff(x)
```

tex/sage/corrDerive_0001.sage

Le résultat est :

```
-4*x^3 + 2*x
-2*sin(2*x)
2*x/(x + 1) - (x^2 - 2)/(x + 1)^2
-3*(x + 1)/x^4 + 1/x^3
tan(x)^2 + 1
2*x - 1/x^2
2^x*log(2)
-x^2/(x^2 + 1)^(3/2) + 1/sqrt(x^2 + 1)
cos(x)
-2*x/(x^2 - 4)^2
-2*(x^2 - 4)*x/abs(x^2 - 4)^3
```

Notez

(1) la syntaxe `a**b` pour signifier a^b ;

(2) la syntaxe `abs(x)` pour la valeur absolue ;

(3) la fonction `log` dans Sage[2] est le logarithme *népérien*, c'est-à-dire en base e. Le logarithme que vous utilisez pour définir le PH d'une solution acide, lui, est en base 10 !

Donnons quelques détails sur certains.

(1) Pour calculer la dérivée de la tangente, on écrit

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \tag{95.472}$$

et ensuite on utilise la formule du quotient :

$$\tan'(x) = \frac{\cos(x)\cos(x) - \sin(x)\big(-\sin(x)\big)}{\cos^2(x)} = \frac{1}{\cos^2(x)}. \tag{95.473}$$

Notez que cela est bien égal à $\tan^2(x)+1$. Parfois la même fonction peut s'écrire de plusieurs façons différentes.

(2) La fonction 2^x n'est pas à confondre avec x^2. Le truc pour trouver la dérivée de 2^x est de se souvenir que le logarithme est la fonction inverse de l'exponentielle. Donc

$$2^x = e^{\ln(2^x)} = e^{x\ln(2)}. \tag{95.474}$$

Nous avons utilisé le fait que $\ln(a^b) = b\ln(a)$. Par conséquent,

$$(2^x)' = \ln(2)e^{x\ln(2)} = \ln(2)2^x. \tag{95.475}$$

Cette astuce de faire $f(x) = e^{\ln(f(x))}$ est très classique pour calculer la dérivée de fonctions dans lesquelles le x apparaît dans une puissance.

(3) La fonction

$$f(x) = \frac{1}{|x^2-4|} \tag{95.476}$$

requiert une attention particulière à cause de la valeur absolue. La fonction valeur absolue est définie par

$$|(|\,x) = \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{si } x < 0. \end{cases} \tag{95.477}$$

Donc nous avons

$$\big|'\big|\,(x) = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x < 0. \end{cases} \tag{95.478}$$

EqDerivDeValAbs

Notez que cette fonction n'a pas de dérivée en zéro parce que son graphe fait un angle ! La formule (95.478) peut être écrite sous la forme

$$\big|'\big|\,(x) = \frac{x}{|x|}. \tag{95.479}$$

C'est cette dernière formule qui est manifestement utilisée par Sage. La dérivée de la fonction, après simplification, est donc

$$f'(x) = \begin{cases} \frac{-2x}{(x^2+4)^2} & \text{si } -2 < x < 2 \\ \frac{2x}{(x^2+4)^2} & \text{sinon.} \end{cases} \tag{95.480}$$

La fonction n'est pas dérivable aux point $x = 2$ et $x = -2$.

2. Comme dans tous les systèmes sérieux.

Exercise 615 exoDerive-0006

Pour quelle(s) valeur(s) du paramètre a la fonction $f(x) = x + \frac{a}{x}$ est-elle croissante autour du point $x = \frac{1}{10}$?

corrDerive-0006

Correction of the exercise 615

La croissance de la fonction revient à la positivité de la dérivée. La dérivée de f est donnée par

$$f'(x) = 1 - \frac{a}{x^2}, \tag{95.481}$$

et donc

$$f'\left(\frac{1}{10}\right) = 1 - 100a. \tag{95.482}$$

Nous devons donc résoudre l'inéquation

$$1 - 100a > 0 \tag{95.483}$$

par rapport à a. La solution est

$$a < \frac{1}{100} \tag{95.484}$$

Comment faire avec Sage ?

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,a')
(x, a)
sage: f=x+a/x
sage: derive=f.diff(x)
sage: derive(1/10)
/home/moky/Sage/local/lib/python2.6/site-packages/IPython/iplib.py:2073:
DeprecationWarning: Substitution using function-call syntax and unnamed
arguments is deprecated and will be removed from a future release
of Sage; you can use named arguments instead,like EXPR(x=..., y=...)
exec code_obj in self.user_global_ns, self.user_ns -1/10/x^2 + 1
sage: derive(x=1/10)
-100*a + 1
sage: solve(derive(x=1/10)>0,a)
[[a < (1/100)]]
```

Notes :

(1) Il faut dire que a va être une variable. C'est pour cela qu'on a commencé par `var('x,a')`. Étant donné que Sage comprend tout seul que `x` est une variable, on aurait pu ne pas déclarer `x`.

(2) Lorsqu'on veut dériver, il faut dire par rapport à quelle variable on veut dériver. Exemple :

```
sage: f
x + a/x
sage: f.diff(a)
1/x
```

C'est pour cela qu'il faut taper `f.diff(x)`.

(3) Le long message d'erreur `DeprecationWarning` est dû à une faute de ma part. En effet la fonction `derive` est une fonction de deux variables : `x` et `a`. C'est pour cela que je dois préciser que je veux l'évaluer avec `x=1/10` et en laissant `a` comme variable. Exemples :

```
sage: derive
-a/x^2 + 1
sage: derive(a=5)
-5/x^2 + 1
sage: derive(a=x)
-1/x + 1
```

(4) Enfin, vous noterez la fonction `solve` qui sert à résoudre des équations et inéquations.

Exercise 616 exoDerive-0004
Quelle est l'équation de la tangente au cercle de rayon R au point d'angle $\pi/4$? corrDerive-0004

Correction of the exercise 616

La partie du cercle dans le premier quadrant est donnée par la fonction

$$y(x) = \sqrt{R^2 - x^2} \tag{95.485}$$

dont la dérivée vaut

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \tag{95.486}$$ EqzqDerv

Le point du cercle qui correspond à l'angle $\frac{\pi}{4}$ est le point qui correspond à $x = R\cos(\pi/4) = \frac{R}{\sqrt{2}}$. En mettant cette valeur de x dans l'expression (95.486) nous trouvons

$$y'\big(\frac{R}{\sqrt{2}}\big) = -1. \tag{95.487}$$

Notez que le R s'est simplifié.

La tangente que nous cherchons est donc la droite de coefficient directeur -1 qui passe par $(x, y) = \big(R\frac{1}{\sqrt{2}}, R\frac{1}{\sqrt{2}}\big)$. Si $y = ax + b$, nous avons immédiatement $a = -1$ Ensuite b se trouve par l'équation

$$\frac{R}{\sqrt{2}} = -\frac{R}{\sqrt{2}} + b, \tag{95.488}$$

d'où nous tirons $b = R\sqrt{2}$. En définitive,

$$y = -x + R\sqrt{2}. \tag{95.489}$$

Exercise 617 exoDerive-0008
Épreuve complémentaire mai 2012

On considère la fonction

$$f\colon (x, y, z) \in \mathbb{R}^3 \to x^2yz + 2y^2\sin(xyz). \tag{95.490}$$

(1) Calculer les dérivées partielles premières de f.

(2) Soit $F(x, y, z) = \nabla f(x, y, z)$ (le gradient de f au point x, y, z). Que vaut le rotationnel de F ?

(3) Que vaut la circulation du champ de vecteurs F le long d'une courbe reliant les points $(1, 2, \pi)$ et $(1, 1, \frac{\pi}{2})$.

corrDerive-0008

Correction of the exercise 617

Les dérivées partielles premières sont un calcul usuel :

```
sage: f(x,y,z)=x**2*y*z+2*y**2*sin(x*y*z)
sage: f.diff(x)
(x, y, z) |--> 2*y^3*z*cos(x*y*z) + 2*x*y*z
sage: f.diff(y)
(x, y, z) |--> 2*x*y^2*z*cos(x*y*z) + x^2*z + 4*y*sin(x*y*z)
sage: f.diff(z)
(x, y, z) |--> 2*x*y^3*cos(x*y*z) + x^2*y
```

C'est-à-dire

$$\frac{\partial f}{\partial x} = 2xyz + 2y^3 z\cos(xyz) \tag{95.491a}$$
$$\frac{\partial f}{\partial y} = x^2 z + 4y\sin(xyz) + 2xy^2 z\cos(xyz) \tag{95.491b}$$
$$\frac{\partial f}{\partial z} = 2xy^3\cos(xyz) + x^2 y. \tag{95.491c}$$

Le rotationnel de $F = \nabla f$ est nul parce c'est la dérivée d'un potentiel (même pas besoin de faire des calculs).

En ce qui concerne la circulation, nous utilisons le fait que nous connaissons le potentiel :

$$\int_\sigma \nabla f = f(1,1,\frac{\pi}{2}) - f(1,2,\pi) = -\frac{3\pi}{2} + 2. \tag{95.492}$$

Exercise 618 exoDerive-0005

Donner une approximation de $\cos(46^\circ)$. corrDerive-0005

Correction of the exercise 618

L'angle proche de 46° dont nous connaissons le cosinus est 45°. En passant aux radians,

$$\cos(46^\circ) = \cos\left(\frac{\pi}{4} + \frac{\pi}{180}\right). \tag{95.493}$$

Nous pouvons donc donner une approximation de $\cos(46^\circ)$ en partant de la fonction cosinus en $\frac{\pi}{4}$ et de sa dérivée :

$$\begin{aligned}\cos\left(\frac{\pi}{4} + \frac{\pi}{180}\right) &\simeq \cos\left(\frac{\pi}{4}\right) - \frac{\pi}{180}\sin\left(\frac{\pi}{4}\right)\\ &= \frac{\sqrt{2}}{2} - \frac{\pi}{180}\frac{\sqrt{2}}{2}\\ &= \frac{1}{\sqrt{2}}\left(1 - \frac{\pi}{180}\right).\end{aligned} \tag{95.494}$$

Notez toutefois que la technique a ses limites. Nous avons certes réussi à exprimer le cosinus en faisant disparaitre les fonctions trigonométriques ; mais il n'en reste pas mois que pour avoir une réelle approximation numérique, il faut encore être capable de calculer $\sqrt{2}$ et π.

Pour vérification :

```
----------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                       |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: x=23*pi/90
sage: exacte=cos(x)
sage: approximation=(1/sqrt(2))*(1-pi/180)
sage: numerical_approx(exacte-approximation)
-0.000107069232665902
```

La valeur exacte est donc un peu plus petite que la valeur calculée. Cela s'explique par le fait que, si on regarde la courbe de cosinus (figure 95.10), on voit que la pente s'accélère [3], ce qui a pour effet que la tangente a tendance à surévaluer la fonction.

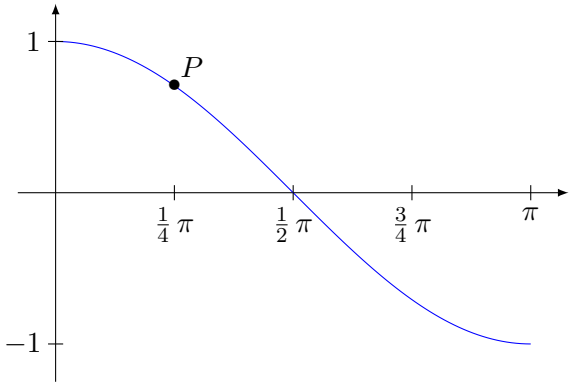


FIGURE 95.10: La fonction cosinus autour de $\frac{\pi}{2}$. LabelFigFnCosApprox

Exercise 619 exoDerive-0009

Soit

$$f(x,y) = x^2 \exp(3y \sin x). \tag{95.495}$$

(1) Déterminer la différentielle de f au point $(\frac{\pi}{4}, 0)$.

(2) Quelle est la dérivée directionnelle de f au point $(\frac{\pi}{4}, 0)$ dans la direction

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}. \end{pmatrix} \tag{95.496}$$

corrDerive-0009

Correction of the exercise 619

Pour les calculs, pas de complications :

```
sage: f(x,y)=x**2*exp(3*y*sin(x))
sage: f.diff(x)
(x, y) |--> 3*x^2*y*e^(3*y*sin(x))*cos(x) + 2*x*e^(3*y*sin(x))
sage: f.diff(y)
(x, y) |--> 3*x^2*e^(3*y*sin(x))*sin(x)
sage: f.diff(x)(x=pi/4,y=0)
1/2*pi
sage: f.diff(y)(x=pi/4,y=0)
3/32*pi^2*sqrt(2)
```

C'est-à-dire que

$$\frac{\partial f}{\partial x}(\frac{\pi}{4}, 0) = \frac{\pi}{2} \tag{95.497a}$$

$$\frac{\partial f}{\partial y}(\frac{\pi}{4}, 0) = \frac{3}{32}\pi^2\sqrt{2}. \tag{95.497b}$$

La différentielle est donc

$$df_{(\pi/4,0)} = \frac{\pi}{2}dx + \frac{3}{32}\pi^2\sqrt{2}dy. \tag{95.498}$$

3. En effet, la dérivée de la dérivée est $-\cos(x)$ qui à cet endroit est négatif, donc la dérivée, c'est-à-dire la pente de la tangente, devient de plus en plus négative.

En ce qui concerne la dérivée directionnelle, nous faisons

$$\frac{\partial f}{\partial u}(\frac{\pi}{4},0)=df_{(\pi/4,0)}(u)=\frac{\pi}{2}\frac{1}{\sqrt{2}}-\frac{3}{32}\pi^2\sqrt{2}\frac{1}{\sqrt{2}}=\frac{\pi}{2}\frac{1}{\sqrt{2}}-\frac{3}{32}\pi^2. \tag{95.499}$$

Exercise 620 exoDerive-0007

Soit $f\colon\mathbb{R}\to\mathbb{R}$ une fonction continue. Soient F et G deux primitives de f. Montrer que le nombre $F(b)-F(a)=G(b)-G(a)$. En d'autres termes, le nombre $F(b)-F(a)$ ne dépend pas du choix de la primitive de f.

corrDerive-0007

Correction of the exercise 620

<+CorrDerive-0007+>

Exercise 621 exoDerive-0002

Dériver les fonctions suivantes.

(1) $\sin\big(\ln(x)\big)$

(2) $\dfrac{\sin x}{x}$;

(3) e^{x^2}

(4) $\cos(x)^{\sin(x)}$

corrDerive-0002

Correction of the exercise 621

Le programme suivant par Sage résout l'exercice :

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

var('x')
f=sin(ln(x))
print f.diff(x)
f=sin(x)/x
print f.diff(x)
f=exp(x**2)
print f.diff(x)
f=cos(x)**(sin(x))
print f.diff(x)
```

tex/sage/corrDerive_0002.sage

Le résultat est :

```
cos(log(x))/x
cos(x)/x - sin(x)/x^2
2*x*e^(x^2)
(log(cos(x))*cos(x) - sin(x)^2/cos(x))*cos(x)^sin(x)
```

Chapitre 96

Exercices de calcul différentiel et intégral 1

96.1 Prérequis

Cette section contient quelques exercices du type de ce qui est plus ou moins censé être connu à l'entrée de l'université dans diverses sections scientifiques [1].

Exercise 622 exoINGE1114-0006

Déterminer l'ensemble des x qui satisfont

(1) $|x| \leqslant 2$
(2) $\left|5 - \frac{1}{x}\right| < 1$
(3) $\frac{1}{|x+2|} > \frac{1}{|x-1|}$
(4) $|x^2 - 2| \leqslant 1$
(5) $|3 - 2x| < 1$.

corrINGE1114-0006

Correction of the exercise 622

Une technique de base dans ces exercices est de savoir que

$$|\text{truc}| < \text{machin} \tag{96.1}$$

implique la double inégalité

$$-\text{machin} < \text{truc} < \text{machin}, \tag{96.2}$$

Eqmactrucmach

lorsque machin > 0.

(1) $x \in [-2, 2]$.
(2) Ici, nous appliquons la règle (96.2) :

$$-1 < 5 - \frac{1}{x} < 1. \tag{96.3}$$

Nous résolvons les deux inégalités $-1 < 5 - \frac{1}{x}$ et $5 - \frac{1}{x} < 1$ séparément. Les solutions du problème seront l'*intersection* des solutions de chacune des deux inégalités.

La première donne $-5 < -1/x$, et donc $6 > 1/x$. Cette inégalité est certainement vraie pour les x négatifs. Pour les $x > 0$, elle est vraie quand $x > 1/6$. Les solutions de cette première inégalité sont donc $]-\infty, 0[\cup]\frac{1}{6}, \infty[$.

La seconde inégalité donne $\frac{1}{x} > 4$. Cette inégalité n'est valable pour aucun négatif. Pour les positifs, les solutions sont $x \in]0, \frac{1}{4}[$.

La solution de l'exercice est donc donnée par

$$\left(]-\infty, 0[\cup]\frac{1}{6}, \infty[\right) \cap]0, \frac{1}{4}[. \tag{96.4}$$

C'est-à-dire $x \in]\frac{1}{6}, \frac{1}{4}[$.

1. Ils proviennent surtout d'un cours pour ingénieur de Louvain-la-Neuve.

(3) Étant donné que les valeurs absolues sont positives, nous pouvons inverser les fractions (en changeant le sens de l'inégalité) et nous avons

$$|x+2| < |x-1|. \tag{96.5}$$

Nous décomposons cette inégalité en quatre possibilités selon que $x+2$ et $x-1$ soient positifs ou négatifs :

(3a) $x+2 \geqslant 0$ et $x-1 \geqslant 0$, conditions remplie lorsque $x \geqslant 1$

(3b) $x+2 \geqslant 0$ et $x-1 \leqslant 0$, conditions remplie lorsque $x \in [-2,1]$

(3c) $x+2 \leqslant 0$ et $x-1 \geqslant 0$, conditions jamais remplie.

(3d) $x+2 \leqslant 0$ et $x-1 \leqslant 0$, conditions remplie lorsque $x \leqslant -2$

Résolvons l'inéquation dans les quatre cas. Le premier cas donne $|x+2| = x+2$ et $|x-1| = x-1$, et donc

$$x+2 < x-1 \tag{96.6}$$

Cette inéquation n'est jamais satisfaite et ne fournit donc aucune solutions.

Le second cas donne $|x+2| = x+2$ et $|x-1| = -x+1$, et donc

$$x+2 < -x+1, \tag{96.7}$$

ce qui donne $x < -1/2$. Ces solutions ne sont des vraies solutions de l'exercice que lorsque $x \in [-2,1]$. Ce second cas fournit donc les solutions $x \in [-2, -\frac{1}{2}[$.

Le troisième cas n'est pas à traiter.

Le quatrième cas donne

$$-x-2 < -x+1, \tag{96.8}$$

qui est toujours satisfaite. Les solutions fournies par ce cas sont donc $x \in]-\infty, -2]$.

Au final, les solutions de l'exercice sont

$$x \in]-\infty, -\frac{1}{2}[\backslash \{-2\}. \tag{96.9}$$

Pourquoi retirer -2 ? Simplement parce que -2 est exclu dès le départ par les conditions d'existence.

(4)

(5) Nous traduisons le problème en

$$-1 \leqslant x^2 - 2 \leqslant 1. \tag{96.10}$$

La première inégalité est $x^2 - 1 \geqslant 0$. (un petit tableau de signe pour ce binôme ?). Les solutions sont $x \in]-\infty, -1] \cup [1, \infty[$.

Les solutions de la seconde sont $x \in [-\sqrt{3}, \sqrt{3}]$.

L'intersection des solutions des deux inégalités est

$$x \in [-\sqrt{3}, -1] \cup [1, \sqrt{3}]. \tag{96.11}$$

(6)

(7) Ici encore, on coupe en deux inégalités

$$-1 < 3 - 2x < 1. \tag{96.12}$$

Les solutions sont respectivement $2 > x$ et $x > 1$ et donc $x \in]1, 2[$.

Exercise 623 [exoINGE1114-0008]

Trouver les supremum, infimum, maximum et minimum, si ils existent dans chacun des ensembles suivants.

(1) $\{\frac{1}{n}$ tel que $n \in \mathbb{N}^*\}$.

(2) $\{0\} \cup \{\frac{1}{n}$ tel que $n \in \mathbb{N}^*\}$.
(3)
(4)
(5) $\{x$ tel que $0 \leqslant x \leqslant \sqrt{2}, x \in \mathbb{R}\backslash\mathbb{Q}\}$.
(6) $\{\frac{1}{n} + (-1)^n$ tel que $n \in \mathbb{N}^*\}$
(7)
(8)
(9)

corrINGE1114-0008

Correction of the exercise 623

(1) $\sup E = 1$ et $1 \in E$, donc $\max E = 1$. Par contre, $\inf E = 0$ et $0 \notin E$, donc il n'y a pas de minimum.
(2) Ici, 0 est dans l'ensemble, donc la différence avec le précédent est que $\min E = 0$.
(3)
(4)
(5) Pour rappel, $0 \in \mathbb{Q}$ et $\sqrt{2} \notin \mathbb{Q}$. Il n'y a donc pas de minimum (parce que 0 n'est pas dans l'ensemble), mais $\inf E = 0$ et $\max E = \sup E = \sqrt{2}$.
(6) Dans la suite des $\frac{1}{n} + (-1)^n$, un terme sur deux s'approche de 1 et l'autre terme sur deux s'approche de -1 (dessiner sur un graphe les cinq ou six premiers termes pour s'en persuader). Nous avons donc $\inf E = -1$ et $\max E = \sup E = \frac{1}{2} + 1 = \frac{3}{2}$. Il n'y a pas de minimum.

Exercise 624 exoINGE1114-0009

Montrer que si a et b sont des nombres réels, alors

$$\sup\{a, b\} = \frac{1}{2}\big(a + b + |a - b|\big) \tag{96.13a}$$
$$\inf\{a, b\} = \frac{1}{2}\big(a + b - |a - b|\big). \tag{96.13b}$$

SubEqsupexoSerUnab

corrINGE1114-0009

Correction of the exercise 624

Prouvons l'égalité (96.13b). Il n'y a que deux possibilités. Soit $a > b$, soit $a < b$. Si $a > b$, alors $|a - b| > 0$ et nous pouvons supprimer les valeurs absolues. Ce que nous avons est alors

$$\frac{1}{2}(a + b + a - b) = a, \tag{96.14}$$

et effectivement, quand $a > b$, nous avons $\sup\{a, b\} = a$. Par contre, si nous supposons que $a < b$, nous savons que $|a - b| = -(a - b)$ et donc

$$\frac{1}{2}(a + b + |a - b|) = \frac{1}{2}(a + b - a + b) = b. \tag{96.15}$$

Exercise 625 exoSerieUn0010

Soit A, un ensemble non vide tel que $\sup A = \inf A$. Que pouvez-vous dire de A ? corrSerieUn0010

Correction of the exercise 625

L'ensemble A est alors un singleton. En effet, si A possédait deux éléments différents$x$ et y, il y aurait forcément un des deux qui est plus grand que l'autre. Supposons que $x > y$. Dans ce cas,

$$\inf A \leqslant y < x \leqslant \sup A, \tag{96.16}$$

ce qui serait contraire à l'hypothèse.

Exercise 626 exoSerieUn0011

Soit A, un ensemble non vide minoré avec $\inf A > 0$. Montrer que

$$\sup\{\frac{1}{x} \text{ tel que } x \in A\} = \frac{1}{\inf A}. \tag{96.17}$$

corrSerieUn0011

Correction of the exercise 626

Nommons $B = \{\frac{1}{x} \text{ tel que } x \in A\}$. Nous allons procéder en plusieurs étapes. D'abord, nous montrons que $1/\inf A$ est un majorant de B, et ensuite, nous montrerons que $1/\inf A$ est plus petit que tous les majorants.

Pour tout x dans a, nous avons $\inf A \leqslant x$, et donc

$$\frac{1}{\inf A} \geqslant \frac{1}{x}. \tag{96.18}$$

Étant donné que $\inf A > 0$, nous pouvons inverser les deux membres en retournant le sens de l'inégalité :

$$\frac{1}{\inf A} \geqslant y \tag{96.19}$$

pour tout $y \in B$, ce qui signifie que $1/\inf A$ est un majorant de B.

Soit maintenant m, un majorant de B. Nous voulons prouver que $m \geqslant 1/\inf A$, ou encore que $1/m \leqslant \inf A$. Étant donné que m est un majorant de B, nous avons que pour tout $x \in A$,

$$\begin{aligned} m &\geqslant \frac{1}{x} \\ \frac{1}{m} &\leqslant x \\ \frac{1}{m} &\leqslant \inf A. \end{aligned} \tag{96.20}$$

Exercise 627 exoSerieUn0012

Montrer que si A et B sont des ensembles bornés tels que $A \subseteq B$, alors

$$\inf B \leqslant \inf A \leqslant \sup A \leqslant \sup B. \tag{96.21}$$

corrSerieUn0012

Correction of the exercise 627

Remarquons pour commencer que inf et inf B existent parce que A et B sont bornés. Prouvons que $\inf B \leqslant \inf A$. Par définition, $\inf B \leqslant x$ pour tout $x \in B$. Étant donné que $A \subseteq B$, nous avons en particulier que $\inf B \leqslant x$ pour tout $x \in A$. Cela prouve que $\inf B$ est un minorant de A, et donc que $\inf B \leqslant \inf A$.

Exercise 628 exoSerieUn0016

Montrer que

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2. \tag{96.22}$$

corrSerieUn0016

Correction of the exercise 628

Vérifions avec $n = 2$:

$$1 + 3 = 4 = 2^2, \tag{96.23}$$

c'est bon.

Supposons maintenant avoir prouvé la thèse pour un certain n. Nous allons montrer qu'alors la thèse est encore vraie pour $n + 1$. En écrivant la formule pour $n + 1$, nous avons

$$\underbrace{1 + 3 + 5 + \cdots + (2n - 1)}_{=n^2} + 2(n + 1) - 1 = n^2 + 2n + 2 - 1 = (n + 1)^2, \tag{96.24}$$

ce qu'il fallait. Pour information, le fait que les termes entre accolades valent n^2 est l'hypothèse de récurrence.

Exercise 629 exoSerieUn0017

Montrer que

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{n(n+1)}=\frac{n}{n+1}. \tag{96.25}$$

corrSerieUn0017

Correction of the exercise 629

Supposons avoir réussi pour n, et regardons ce que devient la formule proposée avec $n+1$ en utilisant l'hypothèse de récurrence :

$$\underbrace{\frac{1}{1\cdot 2}+\cdots+\frac{1}{n(n+1)}}_{n/(n+1)}+\frac{1}{(n+1)(n+2)}. \tag{96.26}$$

En mettant au même dénominateur, nous trouvons

$$\frac{n(n+2)+1}{(n+1)(n+2)}=\frac{n^2+2n+1}{(n+1)(n+2)}=\frac{n+1}{n+2}, \tag{96.27}$$

comme il le fallait.

Exercise 630 exoSerieUn0018

Montrer par récurrence que pour tout $n\geqslant 1$ et pour tout $x\geqslant 0$, on a $(1+x)^n\geqslant 1+nx$. corrSerieUn0018

Correction of the exercise 630

Pour $n=1$, on a $1+x\geqslant 1+n$, ce qui est manifestement vrai. Pour $n+1$, nous avons

$$\begin{aligned}(1+x)^{n+1}=(1+x)(1+x)^n&\geqslant(1+x)(1+nx)\\&=1+nx+x+\underbrace{nx^2}_{\geqslant 0}\\&\geqslant 1+x(n+1),\end{aligned} \tag{96.28}$$

ce qui est bien la thèse avec $n+1$.

Exercise 631 exoINGE11140023

Trouver les réels x qui satisfont l'équation

$$e^{2x}-2e^{-2x}=1. \tag{96.29}$$

corrINGE11140023

Correction of the exercise 631

Étant donné que l'inconnue x apparaît seulement dans la combinaison e^{2x}, il est naturel de poser $y=e^{2x}$. Nous avons alors $e^{-2x}=\frac{1}{y}$ et l'équation à résoudre devient

$$y-\frac{2}{y}-1=0 \tag{96.30}$$

où nous savons que $y\neq 0$ parce qu'une exponentielle ne s'annule jamais. En multipliant par y, nous trouvons une équation du second degré $y^2-y-2=0$, dont les solutions sont

$$y=\frac{1\pm 3}{2}, \tag{96.31}$$

c'est-à-dire $y=2$ ou $y=-1$. La seconde possibilité doit être rejetée parce qu'elle ne correspond à aucune solution en x : $e^{2x}=-1$ est impossible. La seconde solution, $y=2$, donne

$$\begin{aligned}e^{2x}&=2\\2x&=\ln(2)\\x&=\frac{\ln(2)}{2}.\end{aligned} \tag{96.32}$$

La solution de l'exercice est donc $x=\ln(2)/2$. Cette solution peut aussi être exprimée sous la forme $\ln(\sqrt{2})$.

Remarque 96.1.
Une erreur classique est de prendre le logarithme des deux membres de l'équation et d'écrire

$$\ln(e^{2x}) - 2\ln(e^{-2x}) = \ln(1). \tag{96.33}$$

La fonction logarithme n'est pas une fonction linéaire. Il n'est donc pas vrai que $\ln(a+2b) = \ln(a) + 2\ln(b)$.

Parmi les variantes de cette erreur, notons l'oubli de prendre le logarithme du second membre : laisser le 1 inchangé est une circonstance aggravante. Si on prend le logarithme du membre de gauche, il faut aussi prendre celui du membre de droite. Il est également faux de croire que $\ln(1) = 1$ ou que $\ln(1) = e$. Que vaut le logarithme de 1 ?

Exercise 632 exoINGE11140024

Résoudre les équations et inéquations suivantes :

(1) $2^x = 8$
(2) $2^{x-1} - 2^{x-3} = 2^{3-x} - 2^{1-x}$
(3) $e^x + 5e^{-x} = 6$
(4) $2^x > 4^{x+1}$
(5) $\left(\frac{9}{4}\right)^x - \frac{8}{27} = 0$
(6) $2\log_2(x) + \log_x(2) = 3$
(7) $\log_4(1-x) < 0$
(8) $\left(\frac{1}{3}\right)^{3x} < \left(\frac{1}{3}\right)^{2x-9}$.

corrINGE11140024

Correction of the exercise 632

(1) $x = 3$.
(2) On peut poser $y = 2^{x-1}$ et alors $2^{x-3} = y/4$, $2^{3-x} = 4/y$ et $2^{1-x} = 1/y$. Ainsi, on a l'équation suivante pour y :

$$y - \frac{y}{4} = \frac{4}{y} - \frac{1}{y}. \tag{96.34}$$

Étant donné que $y = 0$ n'est pas possible, on peut multiplier tout par y et obtenir une équation du second degré en y :

$$\frac{3y^2}{4} - 3 = 0, \tag{96.35}$$

ou encore $y^2 - 4 = 0$, ce qui donne les solutions $y = 2$ et $y = -2$. La solution $y = 2$ donne $2^{x-1} = 2$, et donc $x = 2$. La solution $y = -2$ donne $2^{x-1} = -2$ qui est impossible. La solution de l'exercice est donc $x = 2$.

Noter qu'on peut également poser $y = 2^x$ ou bien $y = 2^{x-3}$. Ces substitutions mènent à des calculs un peu différents, mais c'est tout aussi bien.

(3) En posant $y = e^x$, nous tombons sur l'équation

$$y^2 - 6y + 5 = 0, \tag{96.36}$$

dont les solutions sont $y = 1$ et $y = 5$. Les solutions en x sont donc $x = \ln(1) = 0$ et $x = \ln(5)$.

(4)
(5)
(6) Le problème de cet exercice est le logarithme en base x. Notez déjà les conditions d'existence : x ne peut pas être négatif, ni nul, ni égal à un. Ensuite, nous utilisons la formule de changement de base des logarithmes pour transformer $\log_x$ en $\log_2$:

$$\log_x(2) = \frac{\log_2(2)}{\log_2(x)}. \tag{96.37}$$

La fraction se simplifie du fait que $\log_2(2) = 1$. L'équation que nous devons regarder devient donc :

$$2\log_2(x) + \frac{1}{\log_2(x)} = 3. \tag{96.38}$$

La bonne idée est de poser $y = \log_2(x)$ et de trouver l'équation

$$2y^2 - 3y + 1 = 0, \tag{96.39}$$

dont les solutions sont $y = 1$ et $y = 1/2$. La solution $\log_2(x) = 1$ donne $x = 2$ et la solution $\log_2(x) = \frac{1}{2}$ donne $x = \sqrt{2}$.

(7) Avant de commencer, notez les conditions d'existence $1 - x > 0$ qui demande $x < 1$. Nous savons que $\log_4(A) < 0$ lorsque $A < 1$. Cela est valable pour toute expression A, tant que la base du logarithme est plus grande que 1. Donc les solutions de l'équation sont données par $x > 0$. En tenant compte des conditions d'existence

$$x \in]0, 1[. \tag{96.40}$$

(8) La fonction

$$x \mapsto \left(\frac{1}{3}\right)^x \tag{96.41}$$

étant décroissante, l'équation est vraie quand

$$3x > 2x - 9, \tag{96.42}$$

c'est-à-dire pour $x > -9$.

Exercise 633 exoINGE11140025

Résoudre les équations et inéquations suivantes :

(1) $\cos(x) = 1$
(2) $2\sin\frac{x}{2} = \frac{1}{2}$
(3) $\cos(\frac{\pi}{2} - x) = \frac{\sqrt{3}}{2}$
(4) $\sin(x) = -\cos(x)$
(5) $-3\cos^2(3x) + \frac{5}{2}\cos(3x) - 1 = 0$
(6) $\tan(2x) = 3\tan(x)$
(7) $\sin(x - \frac{\pi}{3}) > \sin(x)$
(8) $\sin(2x) = 0$
(9) $\cos(40x) = 4$
(10) $\sin(x) = \cos(x)$
(11) $\sin^2(x) + \cos^2(x) = \frac{1}{2}$
(12) $|\cos(2x)| < \frac{1}{2}$
(13) $\arcsin(x) = -\frac{\pi}{4}$

corrINGE11140025

Correction of the exercise 633

(1) Dans le cercle trigonométrique, nous avons $\cos(x) = 1$ lorsque $x = 0$. Les solutions sont donc

$$x \in \{\ldots, -4\pi, -2\pi, 0, 2\pi, 4\pi, \ldots\} = \{2k\pi\}_{k\in\mathbb{Z}}. \tag{96.43}$$

(2)
(3)
(4)

(5) Comme à chaque fois que l'inconnue se trouve dans une fonction spéciale (cosinus, logarithme, exponentielle), un petit changement de variable est tout indiqué. Dans ce cas ci nous posons $y = \cos(3x)$, et nous trouvons l'équation

$$-y^2 + \frac{5}{2}y - 1 = 0, \tag{96.44}$$

dont les solutions sont $y = 2$ et $y = 1/2$.
La solution $y = 2$ correspond à $\cos(3x) = 2$, qui n'a aucune solution. La solution $y = 1/2$ par contre correspond à $\cos(3x) = \frac{1}{2}$. Cela donne les solutions suivantes pour x :

$$x = \frac{\pi}{6} + \frac{2k\pi}{3} \tag{96.45a}$$

$$x = \frac{\pi}{2} + \frac{2k\pi}{3} \tag{96.45b}$$

(6)

(7)

(8) Nous savons que $\sin(2x) = 0$ lorsque $2x = k\pi$, c'est-à-dire

$$x = \frac{k\pi}{2}. \tag{96.46}$$

Ici, comme partout nous sous-entendons $k \in \mathbb{Z}$.

(9) La fonction cosinus ne passe *jamais* par la valeur 4.

(10) Sur le cercle trigonométrique, nous avons $\sin(x) = \cos(x)$ lorsque $x = \frac{\pi}{4}$, mais aussi lorsque $x = \frac{5\pi}{4}$. Les solutions sont donc

$$x = \frac{\pi}{4} + k\frac{\pi}{2}. \tag{96.47}$$

(11) La somme $\cos^2(x) + \sin^2(x)$ fait toujours 1, il n'y a donc pas de solutions.

(12) Nous savons que $\cos(y) = \frac{1}{2}$ lorsque $y = \pi/3$ et $-\pi/3$. Pour toutes les valeurs intermédiaires, $\cos(y) < \frac{1}{2}$. Étant donné que nous regardons la valeur absolue, les valeurs de y entre $2\pi/3$ et $4\pi/3$ sont également à prendre. Au final,

$$\begin{aligned} 2x \in]-\frac{\pi}{3}, \frac{\pi}{3}[+ 2k\pi \cup]\frac{2\pi}{3}, \frac{4\pi}{3}[+ 2k\pi \\ =]-\frac{\pi}{3}, \frac{\pi}{3}[+ k\pi. \end{aligned} \tag{96.48}$$

(13) Prenons le sinus des deux membres de $\arcsin(x) = -\frac{\pi}{4}$. Le premier membre devient x, tandis que le second membre devient $\sin(-\pi/4) = -1/\sqrt{2}$. La solution est donc

$$x = -\frac{1}{\sqrt{2}}. \tag{96.49}$$

Exercise 634 exoINGE11140027

Construire les fonctions inverses des fonctions suivantes sur leur domaine de définition.

(1) $f(x) = \sqrt[5]{x+1}$

(2) $g(x) = \frac{3x-1}{x+4}$.

corrINGE11140027

Correction of the exercise 634

Pour rappel, l'inverse de la fonction f est donnée par

$$f^{-1}(y) = \{x \text{ tel que } f(x) = y\}. \tag{96.50}$$

Cela est une fonction pour les valeurs de y où l'ensemble se réduit à un singleton.

(1) L'équation à résoudre est

$$\sqrt[5]{x+1} = y. \tag{96.51}$$

Donc nous avons $f^{-1}(y) = y^5 - 1$. Notez que cette fonction est définie partout, mais la fonction de départ f ne prend que des valeurs positives. Par conséquent, la fonction f^{-1} ne doit être vue que entre 0 et ∞.

(2) Cette fonction a une asymptote verticale en $x = -4$ et prend toutes les valeurs réelles sauf 3. Son inverse va donc *à priori* être définie sur $\mathbb{R}\backslash\{3\}$. Résolvons l'équation qui donne $g^{-1}(y)$:

$$y = \frac{3x-1}{x+4} \tag{96.52}$$

par rapport à x. La réponse est que

$$g^{-1}(y) = \frac{1+4y}{3-y}. \tag{96.53}$$

Notez la condition d'existence $y \neq 3$. La fonction de départ ne prenant jamais la valeur 3, cette condition d'existence n'en est pas vraiment une parce que f^{-1} n'était déjà définie que sur $\mathbb{R}\backslash\{3\}$.

96.2 Limites et continuité

Exercise 635 exoINGE11140028

Calculer les limites suivantes ou prouver leur non-existence.

(1)

$$\lim_{x\to\infty} \frac{\sin(x)\cos(x)}{x} \tag{96.54}$$

(2)

$$\lim_{x\to\infty} \frac{x+\sin(x)}{x-\sin(x)} \tag{96.55}$$

(3)

$$\lim_{x\to\infty} \sin(x)\frac{x^2+2x+3}{x}. \tag{96.56}$$

Comment résoudre cet exercice avec Sage ? Les lignes de code à rentrer pour résoudre la première limite sont dans l'ordre ItemSageGenda

(1) `var('x')` ItemSageGendb

(2) `f(x)=sin(x)*cos(x)/x` ItemSageGendc

(3) `limit(f(x),x=oo)`

La ligne (1) déclare que la lettre `x` désignera une variable. La ligne (2) définit la fonction f. Notez que la multiplication est notée par la petite étoile `*`. Pour la ligne (3), notez que l'infini est écrit par deux petits « o ».

La fonction $f(x) = x^2+2x+3$ s'encode avec `f(x)=x**2+2*x+3`. Le symbole pour faire l'exposant est la double étoile `**`. corrINGE11140028

Correction of the exercise 635

(1) Cette fonction est une très bonne candidate à l'utilisation de l'étau. En effet, nous avons une fonction trigonométrique $(\sin(x)\cos(x))$ multipliée par une fonction qu'on connait bien $(1/x)$. Pour tout $x > 0$, nous avons les inégalités

$$-\frac{1}{x} \leqslant \frac{\sin(x)\cos(x)}{x} \leqslant \frac{1}{x} \tag{96.57}$$

Les deux fonctions $\pm\frac{1}{x}$ tendent vers zéro lorsque x va vers l'infini ; par conséquent la fonction au milieu tend vers zéro. Note que l'on a utilisé le fait que $-1 \leqslant \sin(x)\cos(x) \leqslant 1$.

ItemB40028ex

(2) Nous commençons par faire les manipulations suivantes :

$$\frac{x+\sin(x)}{x-\sin(x)}=\frac{x\left(1+\frac{\sin(x)}{x}\right)}{x\left(1-\frac{\sin(x)}{x}\right)}=\frac{1+\frac{\sin(x)}{x}}{1-\frac{\sin(x)}{x}}. \tag{96.58}$$

En utilisant la règle de l'étau, nous savons que $\lim_{x\to\infty}\frac{\sin(x)}{x}=0$. Par conséquent nous avons

$$\lim_{x\to\infty}\frac{x+\sin(x)}{x-\sin(x)}=1. \tag{96.59}$$

Remarque que, à côté de fonctions qui tendent vers l'infini, le sinus peut être considéré comme une constante.

Exercise 636 exoINGE11140029

Une fonction périodique admet-elle une limite lorsque x tend vers l'infini ? corrINGE11140029

Correction of the exercise 636

Non, sauf si elle est constante.

Exercise 637 exoINGE11140030

Calculer les limites suivantes en utilisant les règles de calcul.

(1) $\lim_{x\to\pm\infty}\sqrt{x^2+1}-\sqrt{x^2-1}$

(2) $\lim_{x\to0}\frac{\sqrt{x+4}-2}{x}$.

corrINGE11140030

Correction of the exercise 637

(1) Ici, il faut multiplier et diviser par la binôme conjugué et d'appliquer le produit remarquable pour faire disparaitre les racines du numérateur.

$$\begin{aligned}\sqrt{x^2+1}-\sqrt{x^2-1}&=\frac{\left(\sqrt{x^2+1}-\sqrt{x^2-1}\right)\left(\sqrt{x^2+1}+\sqrt{x^2-1}\right)}{\left(\sqrt{x^2+1}+\sqrt{x^2-1}\right)}\\&=\frac{x^2+1-(x^2-1)}{\left(\sqrt{x^2+1}+\sqrt{x^2-1}\right)}\\&=\frac{2}{\left(\sqrt{x^2+1}+\sqrt{x^2-1}\right)}.\end{aligned} \tag{96.60}$$

À partir de cette dernière expression, il est clair que

$$\lim_{x\to\pm\infty}=0. \tag{96.61}$$

(2) Ici, l'astuce est de multiplier et diviser par le binôme conjugué de ce qui se trouve au numérateur :

$$\frac{\sqrt{x+4}-2}{x}=\frac{\left(\sqrt{x+4}-2\right)\left(\sqrt{x+4}+2\right)}{x\left(\sqrt{x+4}+2\right)}. \tag{96.62}$$

L'avantage de cette façon de faire est qu'on peut maintenant utiliser le produit remarquable $(A+B)(A-B)=A^2-B^2$ au numérateur pour faire disparaitre la racine. Après simplifications, nous nous retrouvons avec

$$\frac{1}{\sqrt{x+4}+2}, \tag{96.63}$$

dont la limite pour $x\to0$ est $\frac{1}{4}$.

Exercise 638 exoINGE11140031

Calculer les limites suivantes en utilisant les règles de calcul.

ItemexoINGE_11140031xsinusx

(1) $\lim_{x\to 0} x\sin(\frac{1}{x})$.

(2) $\lim_{x\to\infty}\left(\frac{x}{x+1}\right)^{x+2}$.

ItemexoINGE_11140031

(3) $\lim_{x\to 0}\frac{\sin(\alpha x)}{\sin(\beta x)}$.
Pour effectuer cet exercice avec Sage, il faut taper les lignes suivantes :

(3a) `var('x,a,b')`

(3b) `f(x)=sin(a*x)/sin(b*x)`

(3c) `limit(f(x),x=0)`

Noter qu'il faut déclarer les variables `x`, `a` et `b`.

(4) $\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^{x}$.

(5) $\lim_{x\to\pm\infty}\frac{\sqrt{x^2+1}-x}{x-2}$
Les lignes pour résoudre par Sage sont :

(5a) `f(x)=(sqrt(x**2+1))/(x-2)`

(5b) `limit(f(x),x=oo)`

(5c) `limit(f(x),x=-oo)`

Noter la commande pour la racine carré : `sqrt`. Étant donné que cette fonction diverge en $x = 2$, si tu veux la tracer, il faut procéder en deux fois :

(5a) `plot(f,(-100,1.9))`

(5b) `plot(f,(2.1,100))`

La première ligne trace de -100 à 1.9 et la seconde de 2.1 à 100. Ces graphiques vous permettent déjà de voir les limites. Attention : ils ne sont pas des *preuves* ! Mais ils sont de sérieux indices qui peuvent vous inspirer dans vos calculs.

(6) $\lim_{x\to 2}\frac{\sqrt{x+2}-2}{\sqrt{x+7}-3}$.

(7) $\lim_{h\to 0}\frac{(x+h)^3-x^2}{h}$.

corrINGE11140031

Correction of the exercise 638

(1) Cet exercice, comme beaucoup qui contiennent un cosinus ou un sinus en facteur, se traite avec la règle de l'étau. En effet, étant donné que $\sin(\frac{1}{x})$ est coincé entre -1 et 1, nous avons toujours les inéquations

$$-x \leqslant x\sin(\frac{1}{x}) \leqslant x. \tag{96.64}$$

Or, nous savons que $-x$ et x tendent tous les deux vers zéro lorsque x tend vers zéro. Nous concluons que ce qui se trouve au milieu tend également vers zéro lorsque $x \to 0$.

(2) Ce genre d'exercices avec un x dans l'exposant d'une fraction doit faire penser à la limite qui définit le nombre e. Commençons par inverser la fraction en mettant un signe dans l'exposant, et séparons en deux facteurs :

$$\left(\frac{x}{x+1}\right)^{x+2} = \left(\frac{x+1}{x}\right)^{-x-2} = \left(\frac{x+1}{x}\right)^{-x}\left(\frac{x+1}{x}\right)^{-2} \tag{96.65}$$

Le second facteur tend vers 1 lorsque x tend vers l'infini. Pour le second,

$$\left(\frac{x+1}{x}\right)^{-x} = \left(1+\frac{1}{x}\right)^{-x} = \left(\left(1+\frac{1}{x}\right)^{x}\right)^{-1}. \tag{96.66}$$

La limite du tout fait donc e^{-1}.

(3) La subtilité de cet exercice est de faire apparaitre la limite $\lim_{x\to 0} \sin(x)/x = 1$ que l'on connaît. Pour ce faire, on commence par multiplier et diviser par x :

$$\frac{\sin(\alpha x)}{\sin(\beta x)} = \frac{x}{x}\frac{\sin(\alpha x)}{\sin(\beta x)} \tag{96.67}$$

À ce niveau, nous avons les combinaisons $\sin(\alpha x)/x$ et $x/\sin(\beta x)$ qui apparaissent. Ce n'est pas encore tout à fait ce que l'on veut. Nous multiplions et divisons par $\alpha\beta$ pour mettre tout comme il faut :

$$\frac{\sin(\alpha x)}{\sin(\beta x)} = \frac{x}{x}\frac{\alpha\beta}{\alpha\beta}\frac{\sin(\alpha x)}{\sin(\beta x)} \tag{96.68}$$

La limite de cette expression se décompose en trois morceaux. D'abord la constante α/β, et ensuite les deux limites

$$\lim_{x\to 0} \frac{\beta x}{\sin(\beta x)} \to 1 \tag{96.69a}$$

$$\lim_{x\to 0} \frac{\sin(\alpha x)}{\alpha x} \to 1. \tag{96.69b}$$

La réponse est donc que

$$\lim_{x\to 0} \frac{\sin(\alpha x)}{\sin(\beta x)} = \frac{\alpha}{\beta}. \tag{96.70}$$

(4) Pour cet exercice et les autres avec les exponentielles, voir la sections 97.3.

(5) Pour cet exercice, on peut mettre le plus haut degré de x en évidence en sortant le x^2 de la racine. Ce dernier sort sous la forme de la valeur absolue : $\sqrt{x^2} = |x|$. De plus, lors de la simplification par x, la valeur absolue se transforme en $\pm x$ suivant que x est positif ou négatif. Nous avons

$$\frac{\sqrt{x^2+1}-x}{x-2} = \frac{|x|\sqrt{1+\frac{1}{x}}-x}{x(1-\frac{2}{x})} = \frac{\pm\sqrt{1+\frac{1}{x^2}}-1}{1-\frac{2}{x}}. \tag{96.71}$$

Lorsqu'on fait la limite $x \to \pm\infty$, le dénominateur tend de toutes façons vers un. Le numérateur tend, lui, vers zéro si le signe est positif et vers -2 si le signe est négatif. Donc

$$\lim_{x\to\infty} = 0 \tag{96.72a}$$

$$\lim_{x\to-\infty} = -2 \tag{96.72b}$$

(6) Il s'agit d'utiliser l'astuce du binôme conjugué tant pour le numérateur que pour le dénominateur. Il faut donc multiplier et diviser par $\sqrt{x+2j}+2$ et par $\sqrt{x+7}+3$ en même temps. En utilisant les produits remarquables, et en simplifiant par $(x-2)$, on trouve

$$\frac{\sqrt{x+7}+3}{\sqrt{x+2}+2}, \tag{96.73}$$

dont la limite vaut

$$\lim_{x\to 2} = \frac{3}{2}. \tag{96.74}$$

(7) Si on développe le $(x+h)^3$, nous trouvons au numérateur

$$x^3 + 3hx^2 + 3xh^2 + h^3 - x^3. \tag{96.75}$$

Les x^3 se simplifient et il ne reste plus que des termes qui contiennent h. En écrivant la fraction complète, les h se simplifient avec le dénominateur et nous tombons sur

$$\lim_{h\to 0} 2x^2 + x^2 = 3x^2 \tag{96.76}$$

qui ne dépend plus de h.

Exercise 639 exoINGE11140032

Étudier le domaine de définition et la continuité des fonctions définies ci-dessous.

(1) $f(x) = \begin{cases} \frac{\sin(x)}{x} & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}$

(2) $f(x) = \begin{cases} x\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$

(3) $f(x) = \begin{cases} \frac{\sqrt{x+1}}{\sqrt{x}} & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}$

(4) $f(x) = \begin{cases} \frac{x+1}{x}\sin(\frac{x}{x+1}) & \text{si } x \neq 0 \text{ et } x \neq -1 \\ 0 & \text{si } x = 0 \text{ ou } x = -1. \end{cases}$

(5) $f(x) = (1 + |x|)e^{|x|}$.

(6) $f(x) = \ln(1 + |x|)$.

corrINGE11140032

Correction of the exercise 639

Pour la continuité, ces exercices se réduisent à calculer la limite des fonctions en le point à problème, et vérifier si cette limite est égale ou non à la valeur de la fonction en ce point, en vertu du théorème qui lie la limite et la continuité.

(1) Cette fonction est définie partout, et vu que $\sin(x)/x \to 1$, elle est continue partout.

(2) Elle est définie partout, et elle est continue parce que

$$\lim_{x\to 0} x\sin(\frac{1}{x}) = 0. \tag{96.77}$$

Cette limite a été calculée à l'exercice 638(1).

(3) La fonction n'est définie que pour $x \geqslant 0$. De plus, Nous avons que

$$\lim_{x\to 0} \frac{\sqrt{x+1}}{x} = \infty, \tag{96.78}$$

donc cette fonction n'est pas continue en 0.

(4) La fonction est définie sur tout $\mathbb{R}$. Il faut maintenant regarder les deux points de discontinuité possibles. Le premier revient à calculer la limite

$$\lim_{x\to 0} \frac{x+1}{x}\sin\left(\frac{x}{x+1}\right), \tag{96.79}$$

qui est un cas du type $\frac{\sin(A)}{A}$ avec A qui tend vers zéro. Nous savons que cette limite vaut 1. Nous avons donc que $\lim_{x\to 0} f(x) = 1$, mais $f(0) = 0$. La fonction n'est donc pas continue en 0.

Pour le second point, nous devons calculer

$$\lim_{x\to -1} \frac{x+1}{x}\sin\left(\frac{x}{x+1}\right). \tag{96.80}$$

Pour cette limite, nous utilisons la règle de l'étau. Le sinus reste borné entre -1 et 1, tandis que le facteur $\frac{x+1}{x}$ tend vers zéro. Nous avons donc $\lim_{x\to -1} f(x) = 0$ et donc la fonction est continue en -1.

(5) Cette fonction est définie et continue sur tout $\mathbb{R}$ parce qu'elle est une composée de fonctions continues : la valeur absolue et l'exponentielle.

(6) Cette fonction est définie tant que $1 + |x| > 0$, ce qui est toujours le cas. Elle est également continue partout.

96.3 Suites numériques

Exercise 640 exoINGE11140033

Calculer la limite des suites suivantes.

(1) $s_n = \frac{1}{n}$
(2) $s_n = \frac{2n-1}{3n+2}$
(3) $s_n = \frac{(-1)^n}{n}$

corrINGE11140033

Correction of the exercise 640

(1) Cette suite est la suite type qui tend vers zéro. En effet, prenons n'importe quel $\epsilon > 0$, et trouvons à partir de quel moment dans la suite, $s_n < \epsilon$. Nous cherchons dons les n pour lesquels $\frac{1}{n} < \epsilon$. Cette inéquation est facile à résoudre : la réponse est

$$n > \frac{1}{\epsilon}. \tag{96.81}$$

Prenons par exemple $\epsilon = 0.1$. Effectivement, dès que $n > 10$, nous avons $s_n = \frac{1}{n} < 0.1$.

(2) Cette suite ressemble à une fonction, donc nous allons appliquer à peu près les mêmes recettes. D'abord nous mettons n en évidence au numérateur et au dénominateur et nous simplifions :

$$s_n = \frac{2}{3 - \frac{2}{n}} - \frac{1/n}{3 - \frac{2}{n}}. \tag{96.82}$$

À ce moment nous appliquons le théorème qui dit que $\lim(a_n - b_n) = \lim(a_n) - \lim(b_n)$ (quand les limites existent). Ici,

$$a_n = \frac{2}{3 - \frac{2}{n}} \quad b_n = \frac{1/n}{3 - \frac{2}{n}}. \tag{96.83}$$

En décomposant encore une fois chacun des deux termes en utilisant le théorème qui dit que $\lim(a_n/b_n) = \frac{\lim a_n}{\lim b_n}$, et en utilisant le fait que $\frac{1}{n} \to 0$, nous avons que le limite du tout vaut $\frac{2}{3}$, comme nous nous y attendions depuis le début.

(3) Nous pouvons utiliser la technique de l'étau en écrivant que pour tout n,

$$-\frac{1}{n} \leqslant \frac{(-1)^n}{n} \leqslant \frac{1}{n}. \tag{96.84}$$

Les deux suites extrêmes tendent vers 0, donc celle du milieu tend vers zéro.

Exercise 641 exoINGE11140034

Prouver que la suite $s_n = \frac{2n+7}{3n+2}$ est croissante et bornée, donc qu'elle admet une limite. Calculer cette limite.

corrINGE11140034

Correction of the exercise 641

Afin de prouver la croissance, nous calculons $s_{n+1} - s_n$, et nous prouvons que cette différence est toujours positive. En mettant au même dénominateur et en simplifiant, nous trouvons

$$s_{n+1} - s_n = \frac{2(n+1)-7}{3(n+1)+2} - \frac{2n-7}{3n+2} = \frac{25}{9n^2+21n+1}, \tag{96.85}$$

qui est toujours positif quand n est positif. Note qu'ici, on ne considère que les $n > 0$. Nous travaillons avec des suites, pas avec des fonctions ! Nous pouvons prouver que cette suite est majorée de la façon suivante :

$$s_n = \frac{2n-7}{3n+2} < \frac{2n}{3n+2} < \frac{2n}{3n} = \frac{2}{3}. \tag{96.86}$$

Les inégalités sont vraies parce que on ajoute 7 au numérateur et on enlève 2 au dénominateur. La suite est donc majorée par $\frac{2}{3}$.

Maintenant que nous savons que la suite est majorée et croissante, nous savons qu'elle est convergente. Le calcul de la limite se fait de façon usuelle, et le résultat est $\frac{2}{3}$.

Exercise 642 exoINGE11140035

La suite $s = \{s_n\}_{n\in\mathbb{N}\setminus\{0\}}$ définie par $s_1 = 1$ et $s_n = \sqrt{3s_{n-1}}$ pour tout $n > 1$ est croissante et bornée. Le prouver. Quelle en est la limite ?

corrINGE11140035

Correction of the exercise 642

Pour prouver que la suite est croissante, nous calculons $\frac{s_n}{s_{n-1}}$, et nous prouvons que cette fraction est plus grande que 1 pour tout n. Nous avons

$$\frac{s_n}{s_{n-1}} = \frac{\sqrt{3s_{n-1}}}{s_{n-1}} = \frac{\sqrt{3}}{\sqrt{s_{n-1}}}. \tag{96.87}$$

Pour que cela soit plus grand que 1, il faut d'abord s'assurer que $s_{n-1} < 3$. Cela se fait par récurrence. En effet, si $s_k < 3$, alors

$$s_{k+1} = \sqrt{3s_k} < \sqrt{3 \cdot 3} = 3. \tag{96.88}$$

Nous avons donc bien que tous les éléments de la suite sont plus petits que 3, et donc que la suite est croissante et bornée.

La borne se trouve en utilisant l'astuce suivante. Nous savons que pour tout n,

$$s_{n+1} = \sqrt{3s_n}. \tag{96.89}$$ EqLimPresNsurtrois

Si nous notons $l = \lim s_n$, et si nous prenons la limite des deux côtés de (96.89), nous trouvons

$$l = \sqrt{3l}, \tag{96.90}$$

dont la solution est $l = 3$.

Exercise 643 exoINGE11140036

Les suites suivantes sont-elles bornées, croissantes, décroissantes, convergentes ? Dans le cas où elles sont convergentes, calculer leur limite.

(1) $s_n = 2 - \frac{n-1}{10}$
(2) $s_n = (-1)^n n$
(3) $s_n = \frac{(-1)^{n-1}}{n+1}$.

corrINGE11140036

Correction of the exercise 643

(1) Étant donner que multiplier une suite par un nombre constant ou l'additionner à un nombre constant ne change par le fait qu'elle soit bornée ou non, la suite $s_n = 2 - \frac{n-1}{10}$ sera bornée si et seulement si $\frac{n-1}{10}$ est bornée, ce qui arrivera si et seulement si $n - 1$ est bornée. Or on sait bien que la suite $n - 1$ n'est pas bornée.

Comme la suite n'est pas bornée, elle ne peut pas être convergente. Étudions la (dé)croissance en calculant la différence entre deux termes consécutifs :

$$s_{n+1} - s_n = 2 - \frac{(n+1)-1}{10} - \left(2 - \frac{n-1}{10}\right) = -\frac{1}{10}. \tag{96.91}$$

La différence étant négative, la suite est décroissante.

(2) La convergence de cette suite se règle par l'étau :

$$-\frac{1}{n+1} < \frac{(-1)^{n-1}}{n+1} < \frac{1}{n+1}. \tag{96.92}$$

La suite proposée converge donc vers zéro.

(3) Cette suite n'est ni croissante ni décroissante (parce qu'elle est positive et négative un élément sur deux). Elle ne converge pas non plus parce qu'elle n'est pas bornée.

Exercise 644 exoINGE11140037

Étudier la convergence des suites suivantes en écrivant une expression formelle pour le terme général.

(1) $\sqrt{2}$, $\sqrt{2\sqrt{2}}$, $\sqrt{2\sqrt{2\sqrt{2}}}$, ...

(2) 0.2, 0.23, 0.233, 0.2333,...

corrINGE11140037

Correction of the exercise 644

(1) Cette suite a pour formule de récurrence $s_1 = \sqrt{2}$ et $s_{n+1} = \sqrt{2s_n}$. Sa convergence se traite exactement comme ce qui a été fait pour l'exercice 642.

(2) La formule qui donne la suite est $s_1 = 0.2$ et $s_{n+1} = s_n + \frac{3}{10^n}$.

96.4 Limites

Exercise 645 exoGeneral0010

Calculer les limites suivantes ($\frac{0}{0}$, $\frac{\infty}{\infty}$, $\infty - \infty$,$0 \cdot \infty$).

(1) $\lim_{x\to 2} \frac{x^2+x-6}{x^2-4}$.

(2) $\lim_{x\to \frac{\pi}{4}} \frac{1-\tan(x)}{\cos(2x)}$.

(3) $\lim_{x\to \frac{\pi}{2}} \frac{\sec(x)}{\tan(x)}$.

(4) $\lim_{x\to 0} \frac{x^4-2x^3}{2x-\sin(2x)}$.

(5) $\lim_{x\to \pm\infty} \frac{x}{\sqrt{1+x^2}}$.

(6) $\lim_{x\to 0} \frac{a^x-b^x}{x}$.

(7) $\lim_{x\to 0} \frac{e^x-e^{-x}}{\sin(x)}$.

(8) $\lim_{x\to 1} \left(\frac{x}{x-1} - \frac{1}{\ln(x)}\right)$.

(9) $\lim_{x\to 0} \frac{e^x-e^{\sin(x)}}{x-\sin(x)}$.

(10) $\lim_{x\to 0} \frac{x+\sin(2x)}{x-\sin(2x)}$.

(11) $\lim_{x\to 0} \left(\frac{1}{\sin(x)} - \frac{\cos(x)}{\sin(x)}\right)$.

(12) $\lim_{x\to 0} \frac{\sqrt{x+9}-3}{x}$.

(13) $\lim_{x\to 0} \frac{\tan(x)}{x}$.

(14) $\lim_{x\to >0} x^m \ln(x)$.

(15) $\lim_{x\to <\frac{\pi}{2}} \big(\tan(3x) - \tan(x)\big)$.

(16) $\lim_{x\to\infty} x\sin(a/x)$.

(17) $\lim_{x\to \frac{\pi}{2}} \frac{\ln\big(\sin(x)\big)}{(\pi-2x)^2}$.

Ici, a et b sont des réels positifs, et m est un entier positif.

corrGeneral0010

Correction of the exercise 645

(1) Le numérateur et le dénominateur s'annulent en $x = 2$, donc nous pouvons certainement simplifier la fraction par $(x - 2)$. En effet, une simple factorisation du second degré nous montre que $x^2+x-6 = (x-2)(x+3)$, tandis que le dénominateur est un produit remarquable : $x^2 - 4 = (x + 2)(x - 2)$. Il reste donc

$$\lim_{x\to 2} \frac{(x-2)(x+3)}{(x-2)(x+2)} = \lim_{x\to 2} \frac{x+3}{x+2} = \frac{5}{4}. \tag{96.93}$$

(2) En utilisant la règle de l'Hospital, nous trouvons

$$\lim_{x\to \frac{\pi}{4}} \frac{1-\tan(x)}{\cos(2x)} = \lim_{x\to \frac{\pi}{4}} \frac{-1/\cos^2(x)}{-2\sin(2x)} \lim_{x\to \frac{\pi}{4}} \frac{1}{2\sin(2x)\cos^2(x)} = 1. \tag{96.94}$$

(3)

(4)

(5) Pour rappel : $(a^x)' = a^x \ln(x)$. En sachant cela, la règle de l'Hospital montre que

$$\lim_{x\to 0} \frac{a^x - b^x}{\sin(x)} = \lim_{x\to 0} \frac{a^x \ln(a) - b^x \ln(b)}{\cos(x)} = \ln(a) - \ln(b). \tag{96.95}$$

(6)

(7)

(8)

(9)

(10) Règle de l'Hospital immédiate :

$$\lim_{x\to 0} \frac{1-\cos(x)}{\sin(x)} = \lim_{x\to 0} \frac{\sin(x)}{\cos(x)} = 0. \tag{96.96}$$

(11)

(12)

(13) On a un produit de type $0 \times \infty$. Afin d'utiliser la règle de l'Hospital, nous la écrivons $x^m \ln(x) = \ln(x)/(x^{-m})$, et nous trouvons

$$\lim_{x\to 0} \frac{\ln(x)}{x^{-m}} = \lim_{x\to 0} \frac{1/x}{-mx^{-m-1}} = \lim_{x\to 0} \frac{1}{-mx^{-m}} = 0. \tag{96.97}$$

(14)

(15)

(16)

(17)

(18)

Exercise 646 exoGeneral0011

Calculer les limites suivantes (1^∞, ∞^0, 0^0).

(1) $\lim_{x\to 0} \cos(x)^{\operatorname{cotg}(x)}$

(2) $\lim_{x\to 1} (2-x)^{1/(x-1)}$

(3) $\lim_{x\to_> 1} \left(\frac{1}{x}\right)^{\tan(x)}$

(4) $\lim_{x\to_> 0} (2\cos(x) - e^x)^{\ln(x)}$

(5) $\lim_{x\to_> 0} x^{\sin(x)}$

(6) $\lim_{x\to_> 0} \big(\sin(x)\big)^{\sin(x)}$

(7) $\lim_{x\to 0} (e^x + x)^{1/x}$

(8) $\lim_{x\to\infty} (x^2 + x + 1)^{1/x}$

(9) $\lim_{x\to_> 0} x^{2x}$

(10) $\lim_{x\to\infty} \left(\frac{2}{x} + 1\right)^x$

(11) $\lim_{x\to 1} \ln(x)^{(x-1)}$

(12) $\lim_{x\to_< \frac{\pi}{2}} (\tan x)^{\cos(x)}$

(13) $\lim_{x\to_> 0} \operatorname{cotg}(x)^{\frac{1}{\ln(x)}}$

(14) $\lim_{x\to\infty} \big(\cos(\frac{a}{x}) + \beta \sin(\frac{a}{x})\big)^x$.

corrGeneral0011

Correction of the exercise 646

En règle générale, pour calculer des limites de fonctions qui se présentent sous la forme $f(x)^{g(x)}$, il faut faire la manipulation suivante :

$$f(x)^{g(x)} = e^{\ln\left(f(x)^{g(x)}\right)} = e^{g(x)\ln\left(f(x)\right)}. \tag{96.98}$$ EqManipExpLnlimite

À partir de là, nous devons calculer la limite de $g(x)\ln\big(f(x)\big)$, et puis prendre l'exponentielle en justifiant par le faire que l'exponentielle est une application continue (et donc commute avec la limite).

(1)

(2)

(3)

(4)

(5) En vertu de la manipulation (96.98), nous devons calculer

$$\lim_{x\to 0}\sin(x)\ln(x)=\lim_{x\to 0}\frac{\ln(x)}{1/\sin(x)}=\lim_{x\to 0}\frac{1/x}{-\cos(x)/\sin(x)^2}=-\lim_{x\to 0}\frac{1}{x}\frac{\sin(x)^2}{\cos(x)}. \tag{96.99}$$

EqPassageE

Nous éliminons un des sinus et le $\frac{1}{x}$ en utilisant la limite $\lim_{x\to 0}\frac{\sin(x)}{x}=1$. La limite (96.99) vaut en définitive zéro, de telle sorte que

$$\lim_{x\to_> 0}x^{\sin(x)}=e^0=1. \tag{96.100}$$

(6)

(7)

(8)

(9)

(10) Nous avons

$$\lim_{x\to\infty}\left(\frac{2}{x}+1\right)^x=\lim_{x\to\infty}\exp\left(x\ln(\frac{2}{x}+1)\right). \tag{96.101}$$

La limite à calculer est donc la suivante qui se traite par l'Hospital :

$$\lim_{x\to\infty}x\ln\left(\frac{2}{x}+1\right)=\lim_{x\to\infty}\frac{(-2/x^2)/\left(\frac{2}{x}+1\right)}{-1/x^2}=2. \tag{96.102}$$

La limite recherchée est donc e^2.

(11) Après avoir fait le coup de l'exponentielle, nous devons calculer la limite

$$\lim_{x\to 1}(x-1)\ln\big(\ln(x)\big)=\lim_{x\to 1}=\lim_{x\to}\frac{\ln\big(\ln(x)\big)}{\frac{1}{x-1}}=-\lim_{x\to}\frac{(x-1)^2}{x\ln(x)}=\lim_{x\to 1}\frac{2(x-1)}{1/x}=0. \tag{96.103}$$

La réponse attendue est donc $e^0=1$.

(12)

(13)

(14)

(15)

Exercise 647 exo0013

Vrai ou faux (démontrez ou donnez un contre-exemple) :

ItemAExo0013

(1) Si u_n est telle que $u_{2n}\to 0$ et $u_{2n+1}\to 0$, alors $u_n\to 0$.

(2) Si u_n est telle que $u_{2n}\to 0$ et $u_{3n}\to 0$ et $u_{4n}\to 0$, et ..., alors $u_n\to 0$.

(3) $\exists$ une suite qui prend une infinité de fois les valeurs 1,2.

(4) $\exists$ une suite qui prend une infinité de fois les valeurs 1,2,3.

ItemEnumiE13

(5) $\exists$ une suite qui prend une infinité de fois toutes les valeurs entières.

(6) Toute suite périodique et convergente est constante.

(7) Si $\|x_k\|\to\|x\|$, alors $x_k\to x$.

(8) Toute suite monotone qui a une sous-suite bornée converge.

(9) Il existe une suite telle que $\lim_{n\to\infty}u_n=\infty$ mais qui n'est jamais monotone.

corr0013

Correction of the exercise 647

(1) Tout nombre entier peut être écrit sous la forme $2n$ ou $2n+1$. Soit $\epsilon > 0$, et K tel que $k > K$ implique que x_{2k} et x_{2k+1} soient plus petit que ϵ. Dans ce cas, à partir de $2K$, la suite des x_k est plus petite que ϵ.

(2) Si p_i est la suite des nombres premiers, alors une suite qui a les propriétés imposées peut très bien avoir $x_{p_i} = 1$.

(3) Oui : la suite $1, 2, 1, 2, 1, 2, \ldots$

(4) idem.

(5) La suite $-1, 0, 1, -2, -1, 0, 1, 2, -3, -2, -1, 0, 1, 2, 3, \ldots$ qui consiste à énumérer de $-n$ à n puis de $-(n+1)$ à $n+1$ et ainsi de suite.

(6) Soient $a > b$, deux valeurs atteintes une infinité de fois par la suite. Soit $r \in \mathbb{R}$. Prouvons que la suite ne peut pas converger vers r. Disons que $r \neq a$, dans ce cas, si $|r - a| = \sigma$, pour tout $K > 0$, il existe un $k > K$ tel que $x_k = a$ et donc tel que $|x_k - r| \leqslant \epsilon$ pour $\epsilon = \sigma/2$ par exemple. Même raisonnement si $r \neq b$.

Donc une suite périodique qui converge ne peut pas prendre deux valeurs distinctes (ici : a et b). Nous en déduisons qu'elle doit être constante.

(7) faux, comme le montre le contre-exemple de la suite constante $x_k = -1$. Nous avons $\|x_k\| \to |1|$ et pourtant x_k ne converge pas vers 1.

(8) La sous-suite bornée est monotone et bornée, donc elle converge. Soit a sa limite. Supposons pour fixer les idées que la suite est monotone croissante. La suite x_k ne peut pas prendre de valeurs plus grande que a. En effet si elle prend la valeur $a + \delta$, elle ne peut plus, ensuite, prendre de valeurs intermédiaires entre a et $a + \delta$ parce qu'elle est monotone croissante. Par conséquent, il n'est pas possible d'avoir une sous-suite convergente parce que cette dernière devrait avoir des termes tels que $|a - x_k| < \delta/2$.

La suite elle-même est donc bornée par a et converge donc.

(9) Pour trouver un exemple, il suffit de trouver une suite qui fait deux pas en avant et un pas en arrière. Par exemple $u_n = n + (-1)^n$. Nous avons

$$u_{n+1} - u_n = \begin{cases} -1 & \text{si } n\text{est pair} \\ 3 & \text{si } n\text{est impair,} \end{cases} \tag{96.104}$$

donc cette suite n'est jamais monotone.

Exercise 648 exo0017

Déterminez le supremum (sauf pour l'exercice (7)), l'infimum (sauf pour l'exercice (7)), la limite, les limites supérieure et inférieure (s'ils existent) de chacune des suites ci-dessous.

(1) $(\frac{1}{2}, \frac{-2}{3}, \frac{3}{4}, \frac{-4}{5}, \frac{5}{6}, \ldots)$

(2) $(0, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots)$

(3) $1 + \frac{(-1)^k}{k}$

(4) $\frac{i^k}{k}$

(5) $\cos(\frac{\pi}{4} + k\frac{\pi}{2})$

(6) $(1 + \frac{25}{k^2})$

ItemGExo0017

(7) $\frac{\sin(k)}{k}$

(8) $\lfloor \frac{15}{7+k} \rfloor k^2 + (1 - \lfloor \frac{15}{7+k} \rfloor)\frac{1}{k}$

où $\lfloor r \rfloor$ désigne le plus grand entier inférieur à r. Ainsi $\lfloor 2, 5 \rfloor = 2$ et $\lfloor -2, 5 \rfloor =$ -3. corr0017

Correction of the exercise 648

Pour les questions où la suite est définie en termes de k, nous supposons qu'elle commencent à $k = 1$ (et non $k = 0$).

(1) Cette suite contient une sous-suite qui tend vers 1 et une qui tend vers -1. Tous les termes sont, par ailleurs, plus petit que 1 en norme.

$$\sup = 1, \quad \inf = -1, \quad \limsup = 1, \quad \liminf = -1, \quad \lim = NAN \tag{96.105}$$

(2) Étant donné que le nombre 1 est présent dans toutes les queues de suites (et que c'est le plus petit), nous avons que la limite inférieure est 1.

$$\sup = NAN, \quad \inf = 0, \quad \limsup = NAN, \quad \liminf = 1, \quad \lim = NAN \tag{96.106}$$

(3) Cette suite oscille autour de 1, avec des sauts de plus en plus petits. Le plus grand terme de la suite est celui avec $k = 2$, le plus petit est avec $k = 1$.

$$\sup = 1 + \frac{1}{2}, \quad \inf = 0, \quad \limsup = 1, \quad \liminf = 1, \quad \lim = 1 \tag{96.107}$$

(4) Étant donné que nous n'avons pas considéré de relations d'ordre sur $\mathbb{C}$, les concepts d'infimum, supremum et de limites supérieures et inférieures n'ont pas de sens. Le concept de limite, lui, par contre a encore un sens. Comme la norme de i^k est constamment 1, la limite est zéro.

$$\sup = NAN, \quad \inf = NAN, \quad \limsup = NAN, \quad \liminf = NAN, \quad \lim = 0 \tag{96.108}$$

(5) Cette suite n'est rien d'autre qu'une alternance de $\cos(\pi/4)$ et $-\cos(\pi/4)$.

$$\sup = \cos\frac{\pi}{4}, \quad \inf = -\cos\frac{\pi}{4}, \quad \limsup = \cos\frac{\pi}{4}, \quad \liminf = -\cos\frac{\pi}{4}, \quad \lim = NAN \tag{96.109}$$

(6) Cette suite étant strictement décroissante, il n'y a pas de limite supérieure et nous avons

$$\sup = 1 + 25, \quad \inf = 1, \quad \limsup = NAN, \quad \liminf = 1, \quad \lim = 1 \tag{96.110}$$

(7) Pour calculer la limite supérieure, remarquons que nous avons toujours

$$0 < \sup\{x_k \text{ tel que } k \geqslant l\} \leqslant \frac{1}{l}. \tag{96.111}$$

Pour cette raison, la limite supérieure est nulle. La limite inférieure est également nulle, pour la même raison. Une autre manière de le voir est de voir que

$$\left|\frac{\sin(k)}{k}\right| < \frac{1}{k}, \tag{96.112}$$

et que donc la limite de la suite est zéro. Par la proposition de la page 44 du cours (point (ii)), la limite est, dans ce cas, égale à la limite supérieure et à la limite inférieure.

(8) Il faut remarquer qu'à partir de $k = 9$, la valeur de $\lfloor \frac{15}{7+k} \rfloor$ est nulle. Pour les grands k, la suite n'est donc rien d'autre que $k \mapsto 1/k$ qui tend vers zéro. Cela donne que zéro est la limite inférieure, supérieure et la limite tout court.

L'infimum de la suite est donc zéro et le supremum est atteint en $k = 8$ et vaut 64. Ce terme est le dernier pour lequel la valeur entière vaut 1.

Exercise 649 exo0016

Soient (a_n) une suite convergente dans $\mathbb{R}$ et $\sigma : \mathbb{N} \to \mathbb{N}$ une fonction injective. Posons $b_n = a_{\sigma(n)}$. Prouvez que (b_n) converge et que $\lim b_n = \lim a_n$. corr0016

Correction of the exercise 649

La première chose à remarquer est que la suite $x_n = \sigma(n)$ tend vers l'infini. Maintenant, si nous désignons par a la limite de la suite (a_n), et si nous prenons $\epsilon > 0$, nous avons un N tel que $n > N$ implique $|a_n - a| \leqslant \epsilon$. Prenons K tel que $k > K$ implique $\sigma(k) > N$. Dans ce cas, $|b_k - a| = |a_{\sigma(k)} - a| \leqslant \epsilon$, ce qui prouve que (b_n) converge vers la même limite que (a_n).

Other resolution

Remarque 96.2.
Pour résoudre cet exercice, il faut bien comprendre la définition de la limite : celle-ci dit qu'à ϵ fixé, il n'y a qu'un nombre fini de termes de la suite qui sont « ϵ-loin » de la limite. Toute la question de l'exercice est de savoir combien de ces termes « problématiques » on met dans b_n. L'injectivité de σ assure que chaque terme « problématique » ne peut apparaitre qu'une fois au maximum, et donc qu'ils sont en nombre fini.

Notons a la limite de la suite (a_n), et montrons que $(b_n \stackrel{\text{def}}{=} a_{\sigma(n)})$ converge vers a. Par définition de la convergence de (a_n), on sait que

$$\forall \epsilon > 0 \, \exists K' \text{ tel que } (k' \geqslant K' \Rightarrow |a_{k'} - a| < \epsilon)$$

Pour $\epsilon > 0$ fixé, notons

$$E = \big\{ i \,\big|\, \sigma(i) < K' \big\} \quad \text{et} \quad K = 1 + \max E$$

qui existe bel et bien, car l'ensemble E est fini d'après l'injectivité de σ. Si $k \geqslant K$, alors $k \notin E$ et donc $\sigma(k) \geqslant K'$ et

$$|b_k - a| = \big|a_{\sigma(k)} - a\big| < \epsilon$$

ce qui prouve la convergence de (b_k) vers a.

Exercise 650 exo0024
Déterminez si les limites suivantes existent et dans l'affirmative calculez les.

(1) $\lim_{x\to 0} x \sin(\frac{1}{x})$
(2) $\lim_{x\to 0} \frac{\sin(\sin(x))}{x}$
(3) $\lim_{x\to+\infty} (\ln(x))^{\frac{1}{1-\ln(x)}}$

corr0024

Correction of the exercise 650

(1) Nous avons

$$-|x| \leqslant x \sin(\frac{1}{x}) \leqslant |x|, \tag{96.113}$$

et donc $\lim_{x\to 0} x \sin(\frac{1}{x}) = 0$ par la règle de l'étau.

(2) $\lim_{x\to 0} \frac{\sin(\sin(x))}{x} = 1$ par l'Hospital.

(3) Nous avons

$$\ln(x)^{\frac{1}{1-\ln(x)}} = e^{\ln(\ln(x)^{1/(1-\ln(x))})} = e^{\frac{\ln\left(\ln(x)\right)}{1-\ln(x)}}. \tag{96.114}$$

Par la règle de l'Hospital, nous trouvons que

$$\lim_{x\to\infty} \frac{\ln\left(\ln(x)\right)}{1-\ln(x)} = -\lim_{x\to\infty} \frac{1}{\ln(x)} = 0. \tag{96.115}$$

Nous avons donc

$$\lim_{x\to\infty} e^{\frac{\ln(\ln(x))}{1-\ln(x)}} = e^0 = 1. \tag{96.116}$$

96.5 Dérivées et optimisation

Exercise 651 exoGeneral0012
Déterminer les extréma des fonctions suivantes :

(1) $y = 2\sin(x) + \cos(2x)$
(2) $y = x^2 + 250/x$
(3) $y = e^x/x$.

corrGeneral0012

Correction of the exercise 651

La stratégie est toujours de dériver la fonction, et puis de chercher les zéros de la dérivée.

(1) $y(x) = 2\sin(x) + \cos(2x)$. Nous avons, en utilisant une petite formule de trigonométrie

$$y'(x) = 2\cos(x) - 2\sin(2x) = 2\cos(x) - 2\sin(x)\cos(x) \tag{96.117}$$

Cette dérivée s'annule si $\cos(x) = 0$, ou bien si $\sin(x) = \frac{1}{2}$. Les valeurs, entre π et 2π, qui le font sont

$$\frac{\pi}{2}, \frac{3\pi}{2}, \frac{\pi}{6}, \frac{2\pi}{3}. \tag{96.118}$$

Les extrémums globaux sont $x = 3\pi/2$ et $x = \pi/6$ où la fonction vaut respectivement -3 et $3/2$.

(2) $y(x) = x^2 + \frac{250}{x}$, la dérivée est

$$y'(x) = 2x - \frac{250}{x}. \tag{96.119}$$

Pour résoudre $y' = 0$, nous commençons par multiplier l'équation par x^2 pour trouver $2x^3 - 250 = 0$, dont la seule solution réelle est $x = 5$.

(3) $y(x) = e^x/x$, la dérivée est

$$y'(x) = \frac{(x-1)e^x}{x^2}, \tag{96.120}$$

qui s'annule pour $x = 1$. Rappelez vous que e^x ne s'annule jamais.

Exercise 652 exoGeneral0013

Que vaut la dérivée de la fonction $y = 2\ln^2\big((2x)^2\big)$?

corrGeneral0013

Correction of the exercise 652

Exercice de dérivation de fonctions composées, en n'oubliant pas de se simplifier la vie avec la formule $\ln(a^2) = 2\ln(a)$. La réponse est

$$y'(x) = \frac{8\ln(4x^2)}{x}. \tag{96.121}$$

Cette fonction peut également être dérivée par Sage avec les commandes suivantes :

(1) `var('x')`

(2) `f(x)=2*log( (2*x)**2 )**2`

(3) `f.diff(x)`

Sage donne la réponse sous forme d'une fonction de x.

Exercise 653 exoGeneral0014

La fonction x^x avec $x > 0$ possède un minimum. Donner la position de ce minimum.

corrGeneral0014

Correction of the exercise 653

Afin de dériver x^x, il faut écrire

$$x^x = e^{\ln(x^x)} = e^{x\ln(x)}, \tag{96.122}$$

dont la dérivée vaut

$$\big(\ln(x) + 1\big)e^{x\ln(x)}. \tag{96.123}$$

Cela s'annule lorsque $\ln(x) = -1$, c'est-à-dire en $x = 1/e$.

Exercise 654 exoGeneral0015

Dans le plan euclidien, on considère le point $c = (3, 5)$. Par ce point, on trace une droite D qui coupe l'axe Ox en un point a d'abscisse positive et l'axe Oy en un point b d'ordonnée positive.

Comment doit être la droite D pour que le triangle Oab ait une aire minimum ? (interrogation de janvier 1973) corrGeneral0015

Correction of the exercise 654

La droite D passe par $(3,5)$, par $(0,b)$ et par $(a,0)$. Étant donné que deux points fixent la droite, la connaissance de b fixe a. Nous commençons par chercher la surface du triangle qui passe par $(0,b)$ et par $(3,5)$.

La droite est

$$y = \frac{5-b}{3}x + b, \tag{96.124}$$

et elle coupe l'axe Ox en $x = -\frac{3b}{5-b}$. Le triangle a donc une base de longueur $-3b/(5-b)$ et de hauteur b, donc la surface est

$$S(b) = \frac{-3b^2}{5-b}, \tag{96.125}$$

qu'on va essayer de maximiser. La dérivée est

$$S'(b) = \frac{3b^2 - 30b}{b^2 - 10b + 25}, \tag{96.126}$$

et s'annule pour $b = 10$.

Exercise 655 exoGeneral0016

Un bateau mouille l'ancre à 3 km du rivage. En face d'un point situé 4 km plus loin le long de la côte, un autre bateau est ancré à 9 km du rivage. Un canot du premier bateau doit conduire un passager au rivage et aller ensuite rejoindre l'autre bateau. Quel est le trajet minimum du canot ? corrGeneral0016

Correction of the exercise 655

La situation est représentée sur la figure 96.1.

Cet exercice peut être résolu de façon simple en remarquant que le canot doit aller de A vers I puis de I vers B. Ce trajet est le même que celui qui consiste à aller de A vers I puis de I vers B', si B' est le symétrique de B par rapport à la côte. Ce qu'il faut faire est donc simplement fixer x pour que le trajet AIB' soit une droite.

La réponse est $x = 1$.

Sans cette astuce, la distance à parcourir, en fonction de x s'exprime avec Pythagore :

$$d(x) = \sqrt{x^2+9} + \sqrt{(4-x)^2+81}, \tag{96.127}$$

dont la dérivée est

$$d'(x) = \frac{x}{\sqrt{x^2+9}} - \frac{4-x}{\sqrt{(4-x)^2+81}}. \tag{96.128}$$

Après mise au même dénominateur, nous voyons que cela s'annule en $x = 1$, comme précédemment déduit.

96.6 Primitives et intégration

Exercise 656 exoGeneral0017

Calculer les primitives suivantes (exercice 1, page 40)

(1) $\int (x + \sqrt{x})dx$

(2) $\int \left(\frac{3}{\sqrt{x}} - \frac{x\sqrt{x}}{4}\right) dx$

(3) $\int \frac{dx}{\sqrt[4]{x}}$

(4) $\int \frac{1}{3x-7}dx$

(5) $\int \tan(2x)dx$

(6) $\int \frac{1}{\cos^2(7x)}dx$

(7) $\int \operatorname{cotg}(x/3)dx$

(8) $\int e^{-\sin(x)}\cos(x)dx$

(9) $\int \tan(x)\sec^2(x)dx$

(10) $\int \frac{x}{\sqrt{2x^2+3}}dx$

(11) $\int \frac{\arctan(x)}{1+x^2}dx$

(12) $\int \frac{\cos(x)}{16+\sin^2(x)}dx$

(13) $\int \frac{1}{x\ln(x)}dx$

(14) $\int \frac{1}{\sqrt{9-x^2}}dx$

(15) $\int \frac{1}{4-9x^2}dx$

(16) $\int \frac{1}{\sqrt{x^2+9}}dx$

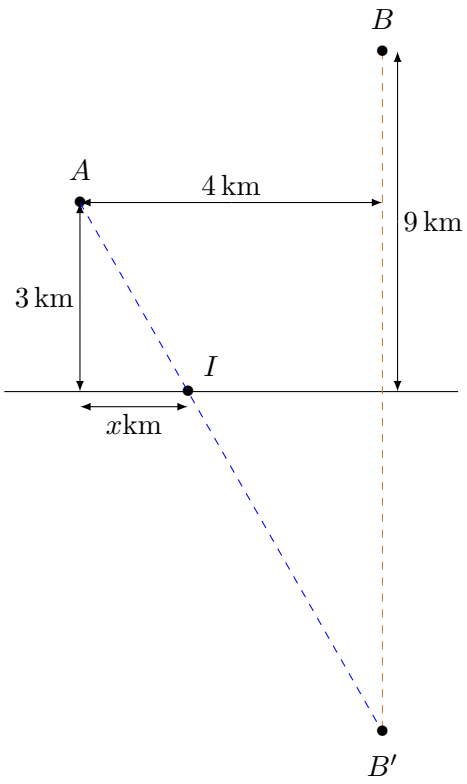


FIGURE 96.1: Petit dessin pour l'exercice 649.

LabelFigBateau

(17) $\int \frac{1}{x\sqrt{1-\ln^2(x)}}dx$
(18) $\int \frac{\sin(2x)}{\sqrt{1+\cos^2(x)}}dx$
(19) $\int \frac{e^x}{\sqrt{1-e^{2x}}}dx$
(20) $\int \frac{\arccos^2(x)}{\sqrt{1-x^2}}dx$
(21) $\int 2^{x^2}xdx$
(22) $\int \frac{x}{\sqrt{1-x^4}}dx$
(23) $\int \frac{e^x}{1+e^{2x}}dx$
(24) $\int \frac{1}{a^2x^2-b^2}dx$
(25) $\int \frac{1}{\cos^2(x)\sqrt{\tan(x)-1}}dx$
(26) $\int \frac{\cos(\ln(x))}{x}dx$
(27) $\int \frac{e^{x/2}}{\sqrt{e^x-1}}$
(28) $\int \sqrt{\frac{e^x}{1-e^x}}dx$
(29) $\int \frac{5x^4+3}{\cos^2(x^5+3x)}dx$
(30) $\int \frac{\sqrt{1+\ln(x)}}{x}dx$
(31) $\int \frac{1-\arctan(x)}{1+x^2}dx$
(32) $\int \operatorname{cotg}(x)\ln\big(\sin(x)\big)dx$
(33) $\int x^2(1-x^3)^2dx$

corrGeneral0017

Correction of the exercise 656

Il y a deux types d'intégrales remarquables. Le premier est

$$\int \frac{f'}{f} = \ln(f) \tag{96.129}$$

pour toute fonction f.

Le second est

$$\int ff' = \frac{f^2}{2} \tag{96.130}$$

pour toute fonction f. Cette formule se prouve simplement faisant le changement de variable $u = f$, donc $du = f'dx$. Avec ce changement de variable, nous trouvons $\int ff'dx = \int udu = \frac{u^2}{2}$.

Afin de se ramener à une de ces deux, il faut soit faire un changement de variable, soit essayer de voir si une partie de la fonction à intégrer ne ressemble pas à une autre partie.

(1) Utiliser la formule $\int x^n = \frac{x^{n+1}}{n+1}$, et trouver

$$\int (x+\sqrt{x})dx = \frac{x^2}{2} + \frac{2}{3}x^{3/2}. \tag{96.131}$$

(2)

(3)

(4) Faire le changement de variable $u = 3x - 7$, $du = 2dx$, ce qui amène

$$\int \frac{dx}{3x-7} = \frac{1}{3}\int \frac{du}{u} = \frac{\ln(u)}{3} = \frac{1}{3}\ln(3x-7). \tag{96.132}$$

ItemIntCinq

(5) Après avoir effectué le changement de variable $u = 2x$, il nous reste à faire

$$I = \int \tan(2x) = \frac{1}{2}\int \frac{\sin(u)}{\cos(u)}du. \tag{96.133}$$

La subtilité est de voir que le numérateur $\sin(u)$ est la dérivée (au signe près) du dénominateur $\cos(u)$, donc nous sommes dans le cas de l'intégrale remarquable f'/f avec $f(x) = \cos(x)$:

$$I = -\frac{1}{2}\ln\big(\cos(u)\big) = -\frac{1}{2}\ln\big(\cos(2x)\big). \tag{96.134}$$

(6) Nous savons que $\big(\tan(x)\big)' = 1/\cos^2(x)$, donc nous posons $u = 7x$, $du = du/7$ et nous trouvons

$$I = \int \frac{dx}{\cos^2(7x)} = \frac{1}{7}\int \frac{du}{\cos^2(u)} = \frac{1}{7}\tan(7x). \tag{96.135}$$

(7) Même idée que (5).

(8) Ici, la bonne idée est de poser $u(x) = -\sin(x)$, de telle manière à avoir $dx = -du/\cos(x)$. L'intégrale devient

$$\int e^{-\sin(x)}\cos(x)dx = \int e^u(-du) = -e^{-\sin(x)}. \tag{96.136}$$

(9) Nous savons que $\big(\tan(x)\big)' = \sec^2(x)$, donc nous sommes dans le cas remarquable $\int ff'$ avec $f = \tan(x)$. La réponse est donc

$$I = \frac{\tan(x)^2}{2}. \tag{96.137}$$

(10) Le truc à voir, c'est que $\big(\sqrt{2x^2+3}\big)' = \frac{2x}{\sqrt{2x^2+3}}$, donc

$$I = \frac{1}{2}\int \frac{2x}{\sqrt{2x^2+3}} = \frac{1}{2}\int \big(\sqrt{2x^2+3}\big)' = \frac{\sqrt{2x^2+3}}{2}. \tag{96.138}$$

(11) La formule à repérer est $\big(\arctan(x)\big)' = 1/(1+x^2)$, donc nous sommes encore dans le cas remarquable $\int ff'$ avec $f(x) = \arctan(x)$. La réponse est donc tout de suite

$$I = \frac{f(x)^2}{2} = \frac{1}{2}\arctan(x)^2. \tag{96.139}$$

(12) Lorsque nous avons des combinaisons simples de fonctions trigonométriques, effectuer un changement de variable où u est une des fonctions. Ici, le changement $u = \sin(x)$ fait l'affaire :

$$\int \frac{\cos(x)}{1+\sin^2(x)} = \int \frac{du}{16+u^2} = \frac{1}{4}\arctan\left(\frac{u}{4}\right) = \frac{1}{4}\arctan\left(\frac{\sin(x)}{4}\right). \tag{96.140}$$

(13) Cette intégrale est un cas remarquable $\int(f'/f)$ avec $f(x) = \ln(x)$.

ItemdixseptUnQuatre

(14)
```
sage: var('x')
sage: f(x)=1/sqrt(9-x**2)
sage: f.integrate(x)
x |--> arcsin(1/3*x)
```

(15)
```
sage: f(x)=1/(4-9*x**2)
sage: f.integrate(x)
x |--> -1/12*log(3*x - 2) + 1/12*log(3*x + 2)
```

(16)
```
sage: f(x)=1/sqrt(x**2+9)
sage: f.integrate(x)
x |--> arcsinh(1/3*x)
```
Noter ici le sinus hyperbolique dû au signe différent par rapport au point (14). Item001717

(17) Ici, c'est le changement de variable $u = \ln(x)$ qui fonctionne parce qu'il amène gratuitement $du = dx/x$, c'est-à-dire le x du dénominateur :

$$\int \frac{dx}{x\sqrt{1-\ln^2(x)}} = \int \frac{du}{\sqrt{1-u^2}} = \arcsin(u) = \arcsin\big(\ln(x)\big). \tag{96.141}$$

(18)
```
sage: f(x)=sin(2*x)/sqrt(1+cos(x)**2)
sage: f.integrate(x)
x |--> -2*sqrt(cos(x)^2 + 1)
```

(19) Ici, afin de faire la même chose que dans le numéro (17), on peut être tenté de poser $u = e^{2x}$ parce que c'est ce qui arrive dans la racine. Hélas, ça ne fonctionne pas (essayez pour voir). Le bon changement est $u = e^x$, et $du = e^x dx$, de telle manière à obtenir

$$\int \frac{e^x dx}{\sqrt{1-e^{2x}}} = \int \frac{du}{\sqrt{1-u^2}} = \arcsin(e^x). \tag{96.142}$$

(20)
```
sage: f(x)=acos(x)**2/sqrt(1-x**2)
sage: f.integrate(x)
x |--> -1/3*arccos(x)^3
```

(21) Le dénominateur ressemble à la formule connue de l'intégrale $\int \frac{1}{\sqrt{1-x^2}}$. Hélas, nous avons x^4 au lieu de x^2. Du coup, nous allons essayer un changement de variable qui fait que $u^2 = x^4$, comme ça nous aurons $\sqrt{1-u^2}$ au dénominateur, comme dans la formule connue. Le changement est $u = x^2$, $du = 2xdx$. Il vient :

$$\int \frac{xdx}{\sqrt{1-x^4}} = \frac{1}{2}\int \frac{du}{\sqrt{1-u^2}}, \tag{96.143}$$

que nous savons faire.

(22)

(23) Ici, il s'agit de trouver un changement de variable qui amène sur $u^2 - 1$, pour cela, nous commençons par récrire le dénominateur de la façon suivante :

$$a^2x^2 - b^2 = b^2(\frac{a^2}{b^2}x^2 - 1), \tag{96.144}$$

et puis nous faisons le changement de variable $u = ax/b$, qui donne

$$I = \frac{1}{b^2}\int \frac{(b/a)du}{u^2-1} = -\frac{1}{ab}\frac{1}{2}\ln\left|\frac{1+u}{1-u}\right|. \tag{96.145}$$

(24) Ici, il faut voir que $1/\cos^2(x)$ est la dérivée de $\tan(x)$, donc le changement de variable $u = \tan(x)$ amène l'intégrale

$$\int \frac{du}{\sqrt{u-1}} \tag{96.146}$$

(25) Le x au dénominateur et la présence d'un logarithme indiquent le changement de variable $u = \ln(x)$, qui amène à

$$\int \frac{\cos\big(\ln(x)\big)dx}{x} = \int \cos(u)du = \sin(u) = \sin\big(\ln(x)\big). \tag{96.147}$$

(26)

(27) En écrivant l'intégrale sous la forme

$$\int \frac{e^{x/2}dx}{\sqrt{1-e^x}}, \tag{96.148}$$

le changement de variable $u = e^{x/2}$ se propose tout seul. Nous avons $dx = 2e^{-x/2}du$ et il reste à intégrer

$$\int \frac{2du}{\sqrt{1-u^2}}. \tag{96.149}$$

(28)

(29) Encore une fois, la présence d'un x au dénominateur et d'un logarithme suggère de poser $u = \ln(x)$. Il reste

$$I = \int \frac{\sqrt{1+\ln(x)}}{x}dx = \int \sqrt{1+u}du, \tag{96.150}$$

cette dernière intégrale se règle par le changement de variable $v = 1 + u$, et nous trouvons

$$I = \frac{v^{3/2}}{3/2} = \frac{2}{3}(1+u)^{3/2} = \frac{2}{3}\big(1+\ln(x)\big)^{3/2}. \tag{96.151}$$

(30)

(31) Chose à remarquer : $\Big(\ln\big(\sin(x)\big)\Big)' = \operatorname{cotg}(x)$, donc on est dans le cas $\int(f'f)$ avec $f = \ln\big(\sin(x)\big)$, et donc

$$I = \frac{f(x)^2}{2} = \frac{1}{2}\ln^2\big(\sin(x)\big). \tag{96.152}$$

(32) Le changement de variable $u = 1 - x^3$ amène

$$\frac{1}{3}\int u^2 du = \frac{1}{3}\frac{u^3}{3} = \frac{1}{9}(1-x^3)^3. \tag{96.153}$$

Exercise 657 exoGeneral0018

Quelques fractions rationnelles à intégrer (exercice 2.1, page 41, numéros 1, 2 et 10).

(1) $\int \frac{dx}{3x^2-2x+4}$

(2) $\int \frac{1}{x^2+3x+1}dx$

(3) $\int \frac{3x+1}{1-4x^2}dx$

(4) $\int \frac{7x+1}{6x^2+x-1}dx$

corrGeneral0018

Correction of the exercise 657

Le secret pour intégrer une fraction avec un second degré au dénominateur est de faire apparaitre une combinaison de la forme $(\alpha x^2 + A) + B$ au dénominateur, à la place du binôme donné.

(1) Nous cherchons A et B pour que $3x^2 - 2x + 4 = (\sqrt{3}x + A)^2 + B$. Pour ce faire, nous développons le carré (produit remarquable) et nous égalisons les termes de degré égaux :

$$(\sqrt{3}x + A)^2 = 3x^2 + 2A\sqrt{3}x + A^2, \tag{96.154}$$

doit être égalé à $3x^2 - 2x + 4$. Nous en déduisons que $A = -1/\sqrt{3}$ et $B = 11/3$. L'intégrale à calculer se récrit donc

$$\int \frac{dx}{3x^2-2x+4} = \int \frac{dx}{\big(\sqrt{3}-\frac{1}{\sqrt{3}}\big)^2 + \frac{11}{3}}. \tag{96.155}$$

Nous posons maintenant $u = \sqrt{3}x - 1/\sqrt{3}$ et nous avons

$$\frac{1}{\sqrt{3}} \int \frac{du}{u^2 + \left(\sqrt{\frac{11}{3}}\right)^2}, \tag{96.156}$$

qui se résous par la formule usuelle.

(2)

(3) Cette intégrale se coupe en deux :

$$I = \int \frac{3x}{1 - 4x^2} + \int \frac{1}{1 - 4x^2}. \tag{96.157}$$

Le deuxième morceau est usuel : il se traite avec le changement de variable $v = 2x$. Pour le premier morceau, on a un x au numérateur, et un dénominateur dont la dérivée contient juste un terme en x. Nous posons donc $u = 1 - 4x^2$, $du = -8xdx$. Nous avons alors

$$I = \int \frac{3x\left(-\frac{du}{8x}\right)}{u} + \int \frac{dv/2}{1 - v^2} = -\frac{3}{8}\ln(u) + \frac{1}{4}\ln\left|\frac{1+v}{1-v}\right|. \tag{96.158}$$

(4) Encore une fois, nous écrivons le dénominateur sous la forme $(\alpha x + A)^2 + B$. Nous trouvons que

$$6x^2 + x - 1 = \left(\sqrt{6}x + \frac{1}{2\sqrt{6}}\right)^2 - \frac{25}{24}. \tag{96.159}$$

Le changement de variable $u = \sqrt{6}x + \frac{1}{2\sqrt{6}}$ amène $du = \sqrt{6}dx$ et

$$x = \frac{u - \frac{1}{2\sqrt{6}}}{\sqrt{6}} \tag{96.160}$$

donc

$$I = \int \frac{\frac{7u - \frac{7}{2\sqrt{6}}}{\sqrt{6}} \frac{1}{\sqrt{6}du}}{u^2 - \left(\sqrt{\frac{25}{24}}\right)^2} \tag{96.161}$$

Cette intégrale se coupe en deux. Une de la forme $\int u/(u^2 - A^2)$, et une de la forme $\int 1/(u^2 - A^2)$. Ces deux ont déjà été traitées.

Exercise 658 exoGeneral0019

Quelques intégrations par partie (exercice 3, page 42).

(1) $\int xe^x dx$

(2) $\int x^2 \ln(x) dx$

(3) $\int \arcsin(x) dx$

(4) $\int e^x \sin(x) dx$

corrGeneral0019

Correction of the exercise 658

(1) Ceci est une intégrale par parie tout ce qu'il y a de plus traditionnel. Nous posons

$$\begin{aligned} u = x \quad & dv = e^x dx \\ du = dx \quad & v = e^x, \end{aligned} \tag{96.162}$$

nous trouvons alors

$$I = xe^x - \int e^x dx = (x-1)e^x. \tag{96.163}$$

(2) Étant donné qu'intégrer $\ln(x)$ ne nous ragoute pas trop, mais que le dériver serait du meilleur effet en société, nous essayons par partie avec les choix suivants :

$$\begin{aligned} u &= \ln(x) & dv &= x^2 \\ du &= \frac{1}{x} & v &= \frac{x^3}{3}. \end{aligned} \tag{96.164}$$

Nous obtenons

$$I = \frac{x^3 \ln(x)}{3} - \int \frac{x^3}{3} = \frac{x^3 \ln(x)}{3} - \frac{x^3}{9}. \tag{96.165}$$

(3) Pour celui-ci, il se passe que l'on sait dériver $\arcsin(x)$, donc nous allons faire $\int \arcsin(x)dx = \int 1 \cdot \arcsin(x)dx$, et poser

$$\begin{aligned} u &= \arcsin(x) & dv &= dx \\ du &= \frac{1}{\sqrt{1-x^2}} & v &= x, \end{aligned} \tag{96.166}$$

ce qui amène à

$$I = x \arcsin(x) - \int \frac{xdx}{\sqrt{1-x^2}} = x \arcsin(x) - \sqrt{1-x^2}. \tag{96.167}$$

(4) Celle-ci est un exemple typique de ce qu'il se passe quand on a deux fonctions « cycliques », c'est-à-dire deux fonctions dont les dérivées bouclent. Posons

$$\begin{aligned} u &= e^x & dv &= \sin(x)dx \\ du &= e^x dx & v &= -\cos(x), \end{aligned} \tag{96.168}$$

ce qui amène

$$I = -e^x \cos(x) = \int \cos(x)e^x, \tag{96.169}$$

dont l'intégrale n'a pas l'air plus sympathique que celle de départ. Si nous refaisons par partie,

$$\begin{aligned} u &= e^x & dv &= \cos(x)dx \\ du &= e^x dx & v &= \sin(x), \end{aligned} \tag{96.170}$$

nous tombons sur

$$I = e^x\big(\sin(x) - \cos(x)\big) - \int \sin(x)e^x. \tag{96.171}$$

L'intégrale qui reste est exactement celle que nous devons calculer. Donc

$$I = e^x\big(\sin(x) - \cos(x)\big) - I, \tag{96.172}$$

ou encore

$$I = \frac{1}{2}e^x\big(\sin(x) - \cos(x)\big). \tag{96.173}$$

Exercise 659 exoGeneral0020

Exercice 5, page 47. Soit $f(x)$, une fonction définie et dérivable sur $[a, b]$. Déterminer

(1) $\int_a^b f'(x)dx$,

(2) la dérivée au point $t = b$ de la fonction

$$F(t) = \int_a^t f'(x)dx, \tag{96.174}$$

(3) la dérivée au point $t = b$ de la fonction

$$F(t) = \int_a^t f(x)dx. \tag{96.175}$$

corrGeneral0020

Correction of the exercise 659

(1) La fonction f est évidement une primitive de f', donc

$$\int_a^b f'(x)dx = f(b) - f(a). \tag{96.176}$$

(2) Par construction, $F(t)$ est une primitive de f', donc sa dérivée en $t = b$ vaut $f'(b)$.

(3) Cette fois, F est une primitive de f, donc sa dérivée en b vaut $f(b)$.

Exercise 660 exoGeneral0021

Exercice 6, page 48. Soit $F(x) = \int_{-\infty}^{x} e^{-t^2}dt$. Parmi les affirmations suivantes, lesquelles sont vraies ? Justifier.

(1) F est croissante
(2) F est décroissante
(3) F n'est ni croissante ni décroissante
(4) F ne s'annule jamais
(5) F est impaire
(6) F est paire
(7) F admet un maximum en $x = 0$
(8) F admet un minimum en $x = 0$
(9) F admet un point d'inflexion en $x = 0$.

Classer par ordre croissant les nombres suivants :

$$0, \quad 1, \quad \int_2^3 e^{-x^2}dx, \quad \int_0^1 e^{-x^2}dx, \quad \int_{-3}^{-2} e^{-x^2}dx. \tag{96.177}$$

corrGeneral0021

Correction of the exercise 660

(1) VRAI : sa dérivée est positive. On peut aussi justifier en disant que c'est la surface sous une courbe toujours positive, donc plus on prend un domaine large, plus la surface est grande.
(2) FAUX
(3) FAUX
(4) VRAI parce que la fonction est strictement croissante avec $\lim_{x\to\infty} f(x) = 0$.
(5) FAUX : une fonction croissante n'est ni impaire, ni paire
(6) FAUX
(7) FAUX : fonction croissante
(8) FAUX
(9) Calculons la dérivée seconde. D'abord, $F'(t) = e^{-t^2}$, ensuite $F''(t) = -2te^{-t^2}$. La dérivée seconde change effectivement de signe en $t = 0$, donc il y a un point d'inflexion.

Pour classer les nombres dans l'ordre, remarquons que la fonction est une surface sous une fonction positive, donc les nombres proposés sont toujours positifs (donc zéro est le plus petit). D'autre part, la fonction est toujours plus petite que 1, donc les intégrales sur des domaines de taille 1 sont toujours plus petites que 1. Reste à classer les autres nombres.

Parce que la fonction est paire, la surface entre -3 et -2 est la même qu'entre 2 et 3. De plus, la fonction est décroissante lorsqu'on s'éloigne de 0, donc la surface proche de zéro est plus grande que les deux autres.

Exercise 661 exoGeneral0022

Exercice 1, page 49. Calculer l'aire de la surface comprise entre les deux paraboles $y = 6x - x^2$ et $y = x^2 - 2x$.

corrGeneral0022

Correction of the exercise 661

Les courbes s'intersectent en $x = 0$ et $x = 4$. La surface entre les deux est

$$S = \int_0^4 y_1(x) - \int_0^4 y_2(x) = \int_0^4 (8x - 2x^2) = \left[\frac{8x^2}{2} - 2\frac{x^3}{3}\right]_0^4 = 64 - \frac{128}{3} = \frac{64}{3}. \tag{96.178}$$

Note : nous avons calculé $\int(y_1 - y_2)$, mais en réalité, c'est un coup de bol. Si la fonction y_2 avait été au dessus de y_1, nous aurions dû calculer $\int(y_2 - y_1)$. Heureusement, il y a un truc : il suffit de voir si le résultat final est positif ou négatif. Si nous avions eut un résultat négatif, alors c'est qu'il fallait calculer l'autre possibilité (ça ne change qu'un signe, donc il ne faut même pas refaire le calcul).

Exercise 662 exoGeneral0023

Exercice 2, page 49. Calculer l'aire de la surface comprise entre les courbes $xy = 12$ et $x + y = 8$. corrGeneral0023

Correction of the exercise 662

Les courbes s'intersectent en $x = 6$ et $x = 2$. L'intégrale à calculer est

$$\int_2^6 \frac{12}{x} - (8 - x) = \left[12\ln(x) + \frac{x^2}{2} - 8x\right]_2^6 = 12\ln(3) - 16. \tag{96.179}$$

Étant donné que c'est négatif, en fait nous savons qu'il fallait calculer le contraire : la surface est $16 - 12\ln(3)$.

Exercise 663 exoGeneral0024

Exercice 2, page 49. Calculer l'aire de la surface comprise entre les deux courbes $x^2 + y^2 = 4$ et $x^2 + y^2 = 4x$. corrGeneral0024

Correction of the exercise 663

Trouver les intersections entre les deux courbes est aisé : il suffit de résoudre le petit système

$$\begin{cases} x^2 + y^2 = 4 & \text{(96.180a)} \\ x^2 + y^2 = 4x, & \text{(96.180b)} \end{cases}$$

dont on déduit immédiatement que $4 = 4x$, et donc que $x = 1$. Ensuite, $y = \pm\sqrt{3}$. Les deux courbes s'intersectent donc en $(1, \pm\sqrt{3})$. Nous allons calculer la surface au dessus de l'axe, et puis multiplier par deux. Avant que la courbe $x^2 + y^2 = 4x$ ne quitte le cercle, c'est-à-dire entre 0 et 1, la surface est donné par cette courbe. Après, entre 1 et 2, elle est donné par le cercle lui-même.

Nous décomposons la surface à calculer en plusieurs intégrales à effectuer :

$$S_1 = \int_0^1 \sqrt{4x - x^2}dx, \quad S_2 = \int_1^2 \sqrt{4 - x^2}dx, \tag{96.181}$$

et la surface totale à faire vaut

$$S = 2(S_1 + S_2). \tag{96.182}$$

L'intégrale S_2 se règle en posant $u = x/2$, ce qui amène à $\sqrt{4 - 4u^2} = 2\sqrt{1 - u^2}$. L'intégrale de cette fonction se calcule en posant $u = \cos(v)$, et en utilisant le fait que $\sqrt{1 - \cos^2(v)} = \sin(v)$. Ce que nous avons est donc

$$\int \sqrt{1 - u^2}du = \int \sqrt{1 - \cos^2(v)}\big(-\sin(v)\big)dv = -\int \sin^2(v). \tag{96.183}$$

La dernière intégrale s'effectue en utilisant la formule de trigonométrie $2\sin^2\frac{x}{2} = 1 - \cos(x)$.

L'intégrale S_1 est plus problématique. Il faut faire apparaitre quelque chose de la forme $u^2 + B$. Cherchons donc A et B tels que

$$4x - x^2 = -(x + A)^2 + B. \tag{96.184}$$

La solution est $A = -2$ et $B = 4$, et la première intégrale devient

$$S_1 = \int_0^1 \sqrt{-(x-2)^2 - 4} = \int_{-2}^{-1} \sqrt{-u^2 + 4} = S_2 \tag{96.185}$$

en posant le changement de variable $u = x - 2$.

Note : en remarquant que la seconde courbe est en réalité aussi un cercle de même rayon, centré en $(0, 2)$, on peut considérablement simplifier le calcul parce que les surfaces rouges et cyan sont en réalité les mêmes. Cela se voit dans le calcul : au final la difficulté est levée en faisant un changement de variable $u = x - 2$ (ce qui est une translation), et on retombe sur la même intégrale que la simple.

Exercise 664 exoGeneral0025

Exercice 13, page 49. Calculer le volume du solide engendré par la rotation autour de Ox de la surface comprise entre la courbe $y = \sqrt{x}e^x$, l'axe Ox et la droite $x = 1$. corrGeneral0025

Correction of the exercise 664

La fonction est $y = \sqrt{x}e^x$, qui n'existe pas en dessous de $x = 0$, qui vaut 0 en $x = 0$ et qui est croissante. En utilisant la formule générale du solide de révolution,

$$V = \pi \int_0^1 xe^{2x} dx, \tag{96.186}$$

qui est à intégrer par partie. Nous trouvons

$$V = \pi \left[\frac{(2x-1)e^{2x}}{4} \right]_0^1 = \pi \left(\frac{e^2}{4} + \frac{1}{4} \right). \tag{96.187}$$

Exercise 665 exoGeneral0026

Exercice 14, page 50. Quel est le volume du solide engendré par la rotation autour de Ox de la surface située dans le premier quadrant entre le système d'axe et la courbe $y = 2^{-x}$. Pour quelle section perpendiculaire à l'axe Ox ce volume sera-t-il partagé en deux parties égales ? corrGeneral0026

Correction of the exercise 665

Comme d'habitude, nous écrivons $2^{-x} = e^{-x\ln(2)}$ avant même de commencer à réfléchir. L'intégrale est assez facile :

$$V = \pi \int_0^\infty e^{-2\ln(2)x} = \left[-\pi \left(\frac{e^{-2\ln(2)x}}{2\ln(2)} \right) \right]_0^\infty = \frac{\pi}{a}. \tag{96.188}$$

Par contre, si nous n'intégrons que de 0 à d, nous trouvons

$$-\pi \left[\frac{e^{-2\ln(2)x}}{2\ln(2)} \right]_0^d = -\pi \frac{e^{-2\ln(2)d}}{a} + \frac{\pi}{2\ln(2)}. \tag{96.189}$$

Pour que l'un soit la moitié de l'autre, il faut

$$1 - e^{-2\ln(2)} = \frac{1}{2}, \tag{96.190}$$

c'est-à-dire $d = \frac{1}{2}$.

Exercise 666 exoGeneral0027

Exercice 17, page 50. La parabole d'équation $y = x^2/3$ partage le disque d'équation $x^2 + y^2 \leqslant 4$ en deux régions. Calculer le volume de révolution engendré par la plus petite de ces régions

(1) autour de l'axe Ox,

(2) autour de l'axe Oy.

corrGeneral0027

Correction of the exercise 666

La parabole « sort » du disque aux points $(\pm\sqrt{3}, 1)$. La partie plus petite est celle qui est au dessus de l'axe Ox. En faisant tourner autour de Ox, nous calculons

$$V = 2\pi \int_0^{\sqrt{3}} \left((4 - x^2) - \frac{x^2}{9} \right) = \frac{28\sqrt{3}}{5}. \tag{96.191}$$

Pour calculer la rotation dans l'autre sens, il faut regarder sa feuille dans l'autre sens, c'est-à-dire inverser x et y dans les formules. De $y = \frac{x^2}{3}$, nous tirons la nouvelle fonction $y = \sqrt{3x}$, et nous intégrons ça entre 1 et 0 :

$$V = \pi \int_2^1 (4 - x^2) + \pi \int_1^0 3x = \frac{19\pi}{6}. \tag{96.192}$$

96.6.1 Longueur d'un arc de courbe

Exercise 667 exoInter0012

Exercice 23, page 51. Calculer la longueur des arcs de courbe suivants.

(1) $y = \ln(1 - x^2)$ pour $0 \leqslant x \leqslant \frac{1}{2}$

(2) $y = x^{3/2}$ pour $0 \leqslant x \leqslant 5$

(3) $y = \ln(x)$ pour $1 \leqslant x \leqslant 2\sqrt{2}$.

corrInter0012

Correction of the exercise 667

À tous les coups, la formule à utiliser est la formule (20.476), la seule vraie difficulté est de calculer l'intégrale.

(1) Ici, $y'(x) = -2x/(1 - x^2)$, et nous devons calculer l'intégrale

$$\begin{aligned} l &= \int_0^{1/2} \sqrt{1 + \frac{4x^2}{(1 - x^2)^2}} \\ &= \int_0^{1/2} \sqrt{\frac{(x^2 + 1)^2}{(1 - x_2)^2}} \\ &= \int_0^{1/2} \frac{x^2 + 1}{1 - x^2}. \end{aligned} \tag{96.193}$$

La division euclidienne de $x^2 + 1$ par $-x_2 + 1$ donne que

$$\frac{x^2 + 1}{1 - x^2} = -1 + \frac{2}{1 - x^2}, \tag{96.194}$$

de telle sorte que la longueur demandée vaut

$$\begin{aligned} l &= \int_0^{1/2} -1 + \int_0^{1/2} \frac{2}{1 - x^2} \\ &= [\ln(1 + x) + 1 - x - \ln(1 - x)]_0^{1/2} \\ &= \ln(3) - \frac{1}{2}. \end{aligned} \tag{96.195}$$

(2) La longueur d'arc de $y = x^{3/2}$ avec $0 \leqslant x \leqslant 5$. L'intégrale à calculer est

$$l = \int_0^5 \sqrt{1 + \left(\frac{3}{2} x^{1/2} \right)^2} dx = \int_0^5 \sqrt{1 + \frac{9x}{4}} dx. \tag{96.196}$$

Cette intégrale s'effectue en faisant le changement de variables $t = 1 + \frac{9x}{4}$, $dx = \frac{4}{9}dx$, ce qui amène à

$$\int \sqrt{t}\frac{4}{9}dt = \frac{4}{9}\frac{t^{3/2}}{3/2} = \frac{8}{27}\left(1 + \frac{9x}{4}\right)^{3/2}, \tag{96.197}$$

et donc

$$l = \frac{8}{27}\left[\left(1 + \frac{9x}{4}\right)^{3/2}\right]_0^5 = \frac{335}{27}. \tag{96.198}$$

Exercise 668 exoInter0013

Exercice 27, page 51. Déterminer a pour que la longueur d'arc de la cycloïde

$$\begin{aligned} x &= a(t - \sin t) \\ y &= a(1 - \cos t) \end{aligned} \tag{96.199}$$

vaille 8. Astuce : penser aux petites formules de trigonométrie pour vous simplifier la vie. corrInter0013

Correction of the exercise 668

La difficulté est de savoir entre quelle et quelle borne du paramètre t se situe un « arc » de cycloïde. Manifestement, la fonction $x(t)$ n'est pas périodique, et même croissante. La fonction $y(t)$, par contre, est périodique de période 2π. Notez que la vitesse de la courbe

$$v = (\dot{x}(t), \dot{y}(t)) = a(1 - \cos(t), \sin(t)) \tag{96.200}$$

a pour norme

$$\|v\| = 2a^2\big(1 - \cos(t)\big), \tag{96.201}$$

s'annule précisément tous les 2π. Nous prenons donc cela comme un bras de cycloïde.

En utilisant la formule de trigonométrie $1 - \cos(t) = 2\sin^2\frac{x}{2}$, et la formule de la longueur d'arc (20.478), nous avons la longueur donnée par

$$l = |a|\sqrt{2}\int_0^{2\pi}\sqrt{1 - \cos(t)} = 2|a|\int_0^{\pi}|\sin(u)|2du = 4|a|\int_0^{\pi}\sin(u)du \tag{96.202}$$

où nous avons utilisé le changement de variable $u = \frac{x}{2}$, ainsi que supprimé la valeur absolue parce que la fonction sinus est toujours positive entre 0 et π. La longueur de l'arc est donc

$$l = 4a\left[-\cos(u)\right]_0^{\pi} = 8a, \tag{96.203}$$

et donc $a = 1$ répond à la question.

96.6.2 Aire d'une surface de révolution

Exercise 669 exoInter0015

Exercice 36, page 52. Trouver l'aire du cône engendré par la rotation du segment de droite $y = 2x$ autour de Ox pour $0 \leqslant x \leqslant 2$. corrInter0015

Correction of the exercise 669

La surface demandée est donnée par

$$S = 2\pi\int_0^2\sqrt{1+4}\,2xdx = 2\pi 2\sqrt{5}\left[\frac{x^2}{2}\right]_0^2 = 8\pi\sqrt{5}. \tag{96.204}$$

Exercise 670 exoInter0014

Trouver l'aire des surfaces engendrées par la rotation des arcs de courbes suivants

(1) $4y = x^3$ pour $0 \leqslant x \leqslant 1$, autour de Ox

(2) $y^2 + 4x = 2\ln(y)$ pour $1 \leqslant y \leqslant 3$ autour de Oy.

corrInter0014

Correction of the exercise 670

(1) La courbe $4y = x^3$ s'écrit tout aussi bien $y = \frac{x^3}{3}$, et la surface de révolution correspondante est

$$S = 2\pi \int_0^1 \sqrt{1 + \left(\frac{3x^2}{4}\right)^2} \frac{x^3}{4} dx \tag{96.205}$$

Le changement de variable miracle qu'il faut voir est $u = 1 + \frac{9x^4}{16}$, parce que $dx = \frac{16}{27}\frac{1}{x^3}du$, ce qui fait que les x^3 se simplifient et l'intégrale à calculer devient

$$2\pi \int \sqrt{u} du = \frac{8\pi}{27} \frac{u^{3/2}}{3/2} = \frac{\pi}{3 \cdot 27} \left(1 + \frac{9x^4}{16}\right)^{3/2}. \tag{96.206}$$

Nous avons donc

$$S = \frac{4\pi}{27} \left[\left(1 + \frac{9x^4}{16}\right)\right]_0^1 = \frac{61\pi}{432}. \tag{96.207}$$

(2) La courbe donnée est

$$x = \frac{2\ln(y) - y^2}{4}. \tag{96.208}$$

Il est important d'écrire $x = x(y)$, et non $y = y(x)$ comme d'habitude, parce que nous allons maintenant faire une rotation le long de l'axe Oy au lieu de Ox. Cela fait qu'il faut changer $x \leftrightarrow y$ dans toutes les formules utilisées d'habitude.

$$\begin{aligned} I(y) &= \int \sqrt{1 + \left(\frac{\frac{2}{y} - 2y}{4}\right)^2} \frac{2\ln(y) - y^2}{4} du \\ &= \int \frac{y^2 + 1}{2y} \left(\frac{2\ln(y) - y^2}{4}\right) dy \\ &= \int \frac{y\ln(y)}{4} dy + \int \frac{\ln(y)}{4y} dy - \frac{y^3}{8} - \frac{y}{8}. \end{aligned} \tag{96.209}$$

La seconde intégrale est une intégrale remarquable : c'est $\int ff'$ avec $f = \ln$, donc l'intégrale vaut $\frac{\ln(y)^2}{2}$. La seule difficulté qui reste est de calculer

$$J = \int y\ln(y) dy. \tag{96.210}$$

Pour cela, nous faisons par partie :

$$\begin{aligned} u &= \ln(y) \quad dv = y \\ du &= \frac{1}{y} dy \qquad v = \frac{y^2}{2}. \end{aligned} \tag{96.211}$$

Alors, nous avons

$$J = \frac{y^2\ln(y)}{2} - \frac{1}{2}\int y dy = \frac{y^2\ln(y)}{2} - \frac{y^2}{4}. \tag{96.212}$$

En remettant les bouts ensemble,

$$I(y) = \frac{y^2\ln(y)}{8} - \frac{y^2}{8} + \frac{\ln(y)^2}{8} - \frac{y^4}{32}, \tag{96.213}$$

et la surface demandée vaut

$$S = 2\pi\big(I(3) - I(1)\big) = \frac{\pi}{4}\big(\ln(3)^2 + 9\ln(3) - 28\big). \tag{96.214}$$

Exercise 671 exoInter0016

Soit $y = 1/x$ avec $x \geqslant 1$. Calculer l'aire sous la courbe et le volume de révolution ; en déduire que pour peindre la surface plane (infinie), il suffit de remplir le volume (fini) ...et qu'on peut même récupérer une grande partie de la peinture après. Où est l'erreur ? corrInter0016

Correction of the exercise 671

La surface générée par la révolution de la courbe $y = 1/x$ autour de l'axe Ox entre 1 et l'infini est donnée par

$$S = 2\pi \int_0^\infty \sqrt{1 + \left(-\frac{1}{x^2}\right)^2} \frac{1}{x} dx = 2\pi \int_0^\infty \frac{\sqrt{x^4 + 1}}{x^3} dx. \tag{96.215}$$

Lorsque x est grand, $x^4 + 1$ est à peu près comme x^4, et donc le numérateur est à peu près comme x^2. Donc, lorsque x est grand, la fonction sous l'intégrale se comporte comme $1/x$, dont l'intégrale diverge.

Nous concluons que la surface est infinie.

En ce qui concerne le volume, le calcule est simple :

$$V = \pi \int_0^\infty \frac{1}{x^2} dx = \pi \left[-\frac{1}{x}\right]_1^\infty = \pi, \tag{96.216}$$

donc il y a un volume fini, entouré par une surface infinie.

96.7 Équations différentielles

96.7.1 Équations à variables séparées

Exercise 672 exoEquaDiff0001

Résoudre les équations suivantes pour $y(x)$.

(1) $y'\sqrt{1 - x^2} + xy = 0$,

(2) $(1 - y^2)dy - ydx = 0$.

corrEquaDiff0001

Correction of the exercise 672

(1) En mettant tous les y à gauche et tous les x à droite, et en utilisant le fait que $y' = dy/dx$, nous trouvons

$$\begin{aligned} \frac{y'}{y} &= -\frac{x}{\sqrt{1 - x^2}} \\ \frac{\frac{dy}{dx}}{y} &= -\frac{x}{\sqrt{1 - x^2}} \\ \frac{dy}{y} &= -\frac{x}{\sqrt{1 - x^2}} dx \end{aligned} \tag{96.217}$$

En prenant l'intégrale des deux côtés, nous avons

$$\ln(y) = \sqrt{1 - x^2} + C, \tag{96.218}$$ EqDiff001presol

donc la solution s'écrit

$$y(x) = e^{\sqrt{1-x^2}+C} = Ke^{\sqrt{1-x^2}}, \tag{96.219}$$

si $K = e^C$.

Remarque : étant donné que $K = e^C$, on pourrait croire que la solution n'est valable que pour K positif. En réalité il n'en est rien parce que l'équation (96.218) devrait contenir une valeur absolue de y.

(2) L'équation $(1-y^2)dy - ydx$ se récrit sous la forme

$$\frac{1-y^2}{y}dy = dx, \tag{96.220}$$

qui peut être intégrée des deux côtés :

$$x = \int \frac{1-y^2}{y}dy = \ln(y) - \frac{y^2}{2}, \tag{96.221}$$

donc la solution est $2\ln(y) - y^2 = 2x + C$.
Note : nous n'avons pas la solution sous la forme $y = y(x)$, mais sous la forme $x = x(y)$...la vie ne peut pas être parfaite !

96.7.2 Équations homogènes

Exercise 673 exoEquaDiff0002
Résoudre les équations suivantes.

(1) $4ydy + xdy = 0$.

(2) $(2x+3y)dx + (x-y)dy = 0$,

corrEquaDiff0002

Correction of the exercise 673

(1) L'équation est

$$y' = \frac{-4y}{x}. \tag{96.222}$$

En posant $y = xz$ et $y' = z + xz'$, nous trouvons

$$\begin{aligned} z + xz' &= \frac{-4xz}{x} = -4z \\ xz' &= -5z \\ \frac{z'}{z} &= -\frac{5}{x}, \end{aligned} \tag{96.223}$$

ce qui fait $\ln(z) = -5\ln(x) + C$, que l'on remet dans la variable y :

$$y = Kx^{-4}. \tag{96.224}$$

(2) En divisant par dx, l'équation devient $(2x+3y) + (y-x)y' = 0$, que l'on remet sous la forme

$$y' = \frac{3y+2x}{x-y}. \tag{96.225}$$

Comme indiqué dans la méthode générale pour ce genre d'équations, il faut poser $u = y/x$, c'est-à-dire $y = ux$ et $y' = u + xu'$, ce qui donne

$$\begin{aligned} u + xu' &= \frac{3ux+2x}{x-ux} = \frac{3u+2}{1-u} \\ xu' &= \frac{3u+2}{1-u} - u \\ x\frac{du}{dx} &= \frac{u^2+2u+2}{1-u}. \end{aligned} \tag{96.226}$$

Ici, la subtilité est de remettre tous les u d'un côté et tous les x de l'autre, y compris les du et dx. Ce que l'on obtient est

$$\frac{1-u}{u^2+2u+2}du = \frac{dx}{x}, \tag{96.227}$$

qui peut être intégré [2] des deux côtés :

$$\ln(x) + K = 2\arctan(u+1) - \frac{1}{2}\ln(u^2+2u+2). \tag{96.228}$$

Nous remettons maintenant les y au lieu des u et, en utilisant les propriétés des logarithmes,

$$K = \ln(y^2+2xy+2x^2) - 4\arctan\left(\frac{y+x}{2}\right). \tag{96.229}$$

Il est à remarquer qu'il n'est pas possible d'isoler y dans cette expression. Nous ne pouvons donc pas donner la solution sous le forme explicite $y = y(x)$. Cela arrive souvent dans le cadre des équations différentielles.

96.7.3 Équations linéaires

Exercise 674 exoEquaDiff0003

Résoudre les équations suivantes.

(1) $y' + 2xy = 4x$

(2) $(x-2)y' = y + 2(x-2)^2$

(3) $y' + y\,\mathrm{cotg}(x) = 5e^{\cos(x)}$, avec la condition initiale $x_0 = \frac{\pi}{2}$, $y_0 = -4$.

corrEquaDiff0003

Correction of the exercise 674

(1) Nous commençons par résoudre l'équation homogène

$$\frac{y'_H}{y_H} = \frac{1}{x-2} \tag{96.230}$$

dont la solution générale est

$$y_H(x) = C(x-2). \tag{96.231}$$

Nous trouvons la solution générale de cette éqiation non homogène en appliquand la méthode de variation des constantes. Nous trouvons l'équation différentielle $C' = 2(x-2)$ pour la fonction $C(x)$. La solution est $C(x) = x^2 - 4x + K$ et donc nous avons

$$y(x) = (x^2 - 4x + K)(x-2). \tag{96.232}$$

(2) L'équation homogène, $y'_H + y_H\,\mathrm{cotg}(x) = 0$, se résous en sachant qu'une primitive de $\mathrm{cotg}(x)$ est $\ln\big(\sin(x)\big)$. La solution de l'homogène est donc

$$y_H(x) = \frac{K}{\sin(x)}. \tag{96.233}$$

La méthode de variation des constantes demande donc de substituer

$$\begin{aligned} y(x) &= \frac{K(x)}{\sin(x)} \\ y'(x) &= \frac{K'(x)}{\sin(x)} - \frac{K(x)\cos(x)}{\sin^2(x)} \end{aligned} \tag{96.234}$$

dans l'équation non homogène. Encore une fois, les termes en K non dérivés se simplifient et nous restons avec

$$K'(x) = 5\sin(x)e^{\cos(x)}, \tag{96.235}$$

donc la solution est $K(x) = -5e^{\cos(x)}$. La solution au problème posé est donc

$$y(x) = \frac{-5e^{\cos(x)}}{\sin(x)}. \tag{96.236}$$

2. Voir le rappel 18.2.2.

96.7.4 Problèmes divers

Exercise 675 exoEquaDiff0004

Soit l'équation différentielle

$$\frac{dy}{dx} + yf(x) = \sin(x). \tag{96.237}$$

(1) Déterminer $f(x)$ de telle manière que cette équation accepte $y = \cos(x)$ comme solution particulière.

(2) Quelle est la solution générale de l'équation différentielle obtenue en remplaçant $f(x)$ par la fonction obtenue ?

(3) Quelle est la solution particulière qui admet un extrémum en un point d'abscisse $\frac{\pi}{4}$?

corrEquaDiff0004

Correction of the exercise 675

(1) Nous substituons $y(x) = \cos(x)$ et $y'(x) = -\sin(x)$ dans l'équation, et nous trouvons

$$-\sin(x) + \cos(x)f(x) = \sin(x), \tag{96.238}$$

d'où nous déduisons $f(x) = 2\tan(x)$.

(2) Nous avons donc l'équation différentielle

$$y' + 2y\tan(x) = \sin(x) \tag{96.239}$$

à résoudre. C'est une équation linéaire, donc nous commençons par l'homogène associée : $y'_H + 2y_H\tan(x) = 0$ qui se transforme en

$$\frac{y'_H}{y_H} = -2\tan(x), \tag{96.240}$$

et dont la solution est

$$y_H(x) = Ke^{-2\ln(1/\cos(x))} = K\cos^2(x). \tag{96.241}$$

Nous trouvons la solution générale de l'équation non homogène par la méthode de *variations des constantes*. Nous posons dons

$$y(x) = K(x)\cos^2(x), y'(x) \quad = K'\cos^2(x) - 2K\cos(x)\sin(x) \tag{96.242}$$

et nous injectons cette fonction dans l'équation de départ. Après simplifications, il ne reste qu'une équation pour K :

$$K'(x) = \frac{\sin(x)}{\cos^2(x)}, \tag{96.243}$$

dont la solution est $K(x) = \frac{1}{\cos(x)} + C$, ce qui donne la solution générale de l'équation de départ :

$$y(x) = \left(\frac{1}{\cos(x)} + C\right)\cos^2(x) = \cos(x)\big(1 + C\cos(x)\big). \tag{96.244}$$

(3) Un tout petit peu de calcul montre que

$$y'(\pi/4) = -\frac{1}{\sqrt{2}} - C, \tag{96.245}$$

qui s'annule lorsque $C = -\frac{1}{\sqrt{2}}$. La solution demandée est donc

$$y(x) = \cos(x)\big(1 + \frac{\cos(x)}{\sqrt{2}}\big). \tag{96.246}$$

Exercise 676 exoEquaDiff0005

Déterminer, sans résoudre l'équation, le lieu des minimums et maximums des solutions de

$$y' = xy - 1. \tag{96.247}$$

Montrer que toutes les courbes solutions coupent les axes sous un angle constant. corrEquaDiff0005

Correction of the exercise 676

Les minimums et maximums de $y(x)$ sont là où $y'(x) = 0$, donc dans la partie du plan où $xy = 1$. L'angle sous lequel $y(x)$ coupe l'axe Oy est donné par la dérivée en 0, c'est-à-dire par $y'(0)$. Par l'équation, nous voyons que $y'(0) = -1$.

L'angle sous lequel la solution $y(x)$ coupe l'axe Ox est donné par $y'(x_0)$ où x_0 est tel que $y(x_0) = 0$. Nous avons $y'(x_0) = x_0 y(x_0) - 1 = -1$.

Exercise 677 exoEquaDiff0006

Dans quelle région du plan les fonctions solutions de l'équation différentielle

$$y' = \sin(x^2 + y^2) \tag{96.248}$$

sont-elles croissantes, décroissantes ? corrEquaDiff0006

Correction of the exercise 677

Les solutions passant par le point (x, y) seront croissantes en ce point si $y'(x) > 0$, c'est-à-dire si $\sin(x^2 + y^2) > 0$.

Exercise 678 exoEquaDiff0007

Montrer que la substitution $y = t/x$ réduit l'équation différentielle

$$y(1 - xy)dx - x(1 + xy)dy = 0 \tag{96.249}$$

à une équation à variables séparables. Résoudre cette équation. corrEquaDiff0007

Correction of the exercise 678

Le changement de fonction inconnue $y(x) = \frac{t(x)}{x}$ induit le changement suivant pour la dérivée :

$$y'(x) = \frac{1}{x}\left(t' - \frac{t}{x}\right), \tag{96.250}$$

que nous substituons dans l'équation proposée[3]. Nous trouvons

$$\begin{aligned} \frac{t}{x}(1-t) - x(1+t)\frac{1}{x}\left(t' - \frac{t}{x}\right) &= 0 \\ \frac{t - t^2}{x} - \frac{(1+t)(xt' - t)}{x} &= 0 \\ -xt'(1+t) + 2t &= 0 \\ \frac{(1+t)}{2t}t' &= \frac{1}{x}. \end{aligned} \tag{96.251}$$

où nous avons fait une simplification par x. La solution que nous allons obtenir n'est donc pas censée être valable en $x = 0$. Ici, et dans tout le chapitre sur les équations différentielles, nous ne discutons pas le domaine de validité des solutions.

En passant le dx de l'autre côté, et en intégrant des deux côtés,

$$\int \frac{(1+t)}{2t}dt = \int \frac{dx}{x}. \tag{96.252}$$

Nous avons donc $\ln(x) + C = \frac{\ln(t)}{2} + \frac{t}{2}$. Ici, nous allons faire quelque chose de subtil. La constante C peut valoir n'importe quoi, donc il n'y a pas de mal à changer de constante et écrire $C = \ln(K)$ où

3. Le verbe « proposer » est évidement un euphémisme.

K est une nouvelle constante. Alors, dans le membre de gauche nous avons $\ln(x)+\ln(K) = \ln(Kx)$. Cette manipulation a fait « rentrer » la constante dans le logarithme.

Nous faisons de même avec le dénominateur 2 du membre de droite : nous le faisons passer à gauche, et nous le rentrons dans le logarithme sous forme d'un carré, nous avons donc

$$\ln(Kx^2) = \ln(t) + t, \tag{96.253}$$

donc, en prenant l'exponentielle des deux côtés,

$$\begin{aligned} Kx^2 &= e^{\ln(t)}e^t = xye^{xy} \\ Kx &= ye^{xy}. \end{aligned} \tag{96.254}$$

Encore une fois, nous ne pouvons pas écrire la solution sous forme explicite $y = y(x)$.

Exercise 679 exoEquaDiff0008

Chercher la solution de l'équation

$$y' + \frac{y}{1+x} = x \tag{96.255}$$

telle que $y'(0) = 1$. corrEquaDiff0008

Correction of the exercise 679

L'équation donnée étant une équation linéaire nous résolvons d'abord l'équation homogène associée :

$$y'_H + \frac{y_H}{1+x} = 0. \tag{96.256}$$

En remettant les y_H d'un côté et les x de l'autre et en intégrant, nous trouvons la solution générale de l'homogène sous la forme

$$y_H(x) = \frac{C}{1+x}. \tag{96.257}$$

La méthode de variation des constantes demande d'introduire

$$\begin{aligned} y(x) &= \frac{C(x)}{1+x} \\ y'(x) &= \frac{C'(x)}{1+x} - \frac{C(x)}{(1+x)^2} \end{aligned} \tag{96.258}$$

dans l'équation de départ. Les termes en C non dérivés se simplifient, et nous restons avec

$$C'(x) = x(1+x), \tag{96.259}$$

que nous intégrons immédiatement : $C(x) = \frac{x^2}{2} + \frac{x^3}{3} + K$, et donc

$$y(x) = \frac{3x^2 + 2x^3 + K}{6(1+x)}. \tag{96.260}$$ EqDel008SScAu

Nous pouvons résoudre les conditions initiales sans dériver (96.260). En effet, l'équation de départ dit que $y'(0) + \frac{y(0)}{1+0} = 0$. En demandant $y'(0) = 1$, cette équation dit que $y(0) = K/6$. Nous posons maintenant $x = 0$ dans (96.260) et nous demandons que le résultat soit -1. Ce que nous trouvons est $k = -6$, et donc la solution

$$y(x) = \frac{3x^2 + 2x^3 - 6}{6(1+x)}. \tag{96.261}$$

Exercise 680 exoEquaDiff0009

En tout point d'une courbe d'équation $y = y(x)$, le carré de la pente de la tangente est égal au produit de l'abscisse et de l'ordonnée du point. Trouver l'équation de cette courbe, sachant qu'elle passe par le point $(1, 1)$. corrEquaDiff0009

Correction of the exercise 680

L'équation différentielle à résoudre est $y'(x)^2 = xy$. Cela se résous en prenant la racine carrée des deux côtés, et en ramenant les y d'un côté et les x de l'autre :

$$\frac{dy}{\sqrt{y}} = \pm\sqrt{x}dx. \tag{96.262}$$

Cette équation peut être intégrée des deux côtés. Cela donne

$$y(x) = \left(\frac{x^{3/2}}{3} + C\right)^2. \tag{96.263}$$

En particulier, $y(1) = \left(\frac{1}{3} + C\right)^2$, donc pour que $y(1) = 1$, il faut $C = \frac{2}{3}$ ou bien $C = -\frac{4}{3}$.

96.7.5 Équations différentielles du second ordre

Exercise 681 exoEquaDiff0010

Exercice II.1, page 59. Déterminer la solution générale de chacune des équations différentielles suivantes.

(1) $y'' - 2y' - 3y = 0$
(2) $y'' + 4y' + 4y = 0$
(3) $y'' - 4y' + 5y = 0$
(4) $y'' - 4y' + 3y = 0$
(5) $2y'' - y' = 0$
(6) $y'' + 4y = 0$
(7) $y'' + y' + y = 0$
(8) $y'' + 2y' + y = 0$
(9) $y'' - 7y' + 10y = 0$
(10) $4y'' + 12y' + 9y = 0$

corrEquaDiff0010

Correction of the exercise 681

(1) Le polynôme caractéristique est $x^2 - 2x - 3 = 0$, et ses solutions sont

$$\begin{aligned}\lambda_1 &= 3\\ \lambda_2 &= -1,\end{aligned} \tag{96.264}$$

donc les solutions de l'équation sont données par

$$y(x) = Ae^{3x} + Be^{-x} \tag{96.265}$$

avec A, $B \in \mathbb{R}$.

(2) Le polynôme caractéristique est $x^2 + 4x + 4 = 0$, et sa racine est

$$\lambda_1 = \lambda_2 = -2. \tag{96.266}$$

C'est une racine double. Dans le cas d'une racine double, les solutions « de base » de l'équation différentielle sont $y_1(x) = e^{\lambda x}$ et $y_2(x) = xe^{\lambda x}$. Dans notre cas,

$$y(x) = Ae^{-2x} + Bxe^{-2x}. \tag{96.267}$$

(3) Cette fois, les racines du polynôme caractéristique $\lambda^2 - 4\lambda + 5 = 0$ sont complexes :

$$\frac{4 \pm \sqrt{16 - 20}}{2} = 2 \pm i. \tag{96.268}$$

Les solutions de base de l'équation différentielle sont donc $y_1(x) = e^{(2+i)x}$ et $y_2(x) = e^{(2-i)x}$. Ce sont des solutions complexes, et il existe une astuce pour en extirper les solutions réelles. D'abord, nous séparons les parties réelles et imaginaires dans les exponentielles :

$$e^{(2+i)x} = e^{2x}e^{ix}, \tag{96.269}$$

et ensuite, nous utilisons la formule d'Euler $e^{i\omega x} = \cos(\omega x) + i\sin(\omega x)$:

$$\begin{aligned} y_1(x) &= e^{2x}\big(\cos(x) + i\sin(x)\big) \\ y_2(x) &= e^{2x}\big(\cos(x) - i\sin(x)\big). \end{aligned} \tag{96.270}$$

Nous savons que toutes les combinaisons de y_1 et y_2 sont des solutions de l'équation différentielle. Les combinaisons qui nous intéressent sont les combinaisons qui ne contiennent plus de nombres complexes. Il n'est pas très difficile de les trouver :

$$y_1(x) + y_2(x) = 2e^{2x}\cos(x), \tag{96.271}$$

et

$$i\big(y_1(x) - y_2(x)\big) = -2e^{2x}\sin(x). \tag{96.272}$$

Les solutions réelles de l'équation sont donc

$$y(x) = Ae^{2x}\cos(x) + Be^{2x}\sin(x). \tag{96.273}$$

(4)

$$y(x) = Ae^{3x} + Be^{x} \tag{96.274}$$

(5) Les racines du polynôme caractéristique sont $\lambda_1 = 1/2$ et $\lambda_2 = 0$. Pour rappel, $e^0 = 1$, donc la seconde racine ne donne pas lieu à une exponentielle dans la solution :

$$y(x) = Ae^{x/2} + B \tag{96.275}$$

(6) Les solutions complexes sont

$$\begin{aligned} y_1(x) &= e^{-2i} \\ y_2(x) &= e^{2i} \end{aligned} \tag{96.276}$$

Nous en extirpons les solutions réelles en en prenant la somme et la différence multipliée par i. Nous avons donc la solution réelle générale

$$y(x) = A\cos(2x) + B\sin(2x). \tag{96.277}$$

(7) Cet exemple est justement traité dans une introduction à Maxima [4]. Les solutions complexes sont données par

$$\begin{aligned} y_1(x) &= \exp\big(\frac{-1+\sqrt{3}i}{2}x\big) \\ y_2(x) &= \exp\big(\frac{-1-\sqrt{3}i}{2}x\big) \end{aligned} \tag{96.278}$$

Les solutions réelles sont

$$\begin{aligned} y_1(x) &= e^{-x/2}\cos(\frac{\sqrt{3}}{2}x) \\ y_2(x) &= e^{-x/2}\sin(\frac{\sqrt{3}}{2}x) \end{aligned} \tag{96.279}$$

(8) Le polynôme caractéristique admet la racine double $\lambda = -1$. La solution générale est donc

$$y(x) = Ae^{-x} + Bxe^{-x}. \tag{96.280}$$

(9) Nous avons $\lambda_1 = 5$ et $\lambda_2 = 2$, donc

$$y(x) = Ae^{5x} + Be^{2x}. \tag{96.281}$$

(10) Racine double $\lambda = -3/2$, donc

$$y(x) = Ae^{-3x/2} + Bxe^{-3x/2}. \tag{96.282}$$

4. http://michel.gosse.free.fr/documentation/introduction/node6.html

Exercise 682 exoEquaDiff0011

Exercice II.2, page 59. Pour chacune des équations différentielles suivantes, déterminer la solution particulière répondant aux conditions initiales indiquées.

(1) $y'' + 10y' + 16y = 0$, CI : $x_0 = 0$, $y_0 = 1$, $y'_0 = -12$
(2) $y'' - 6y' + 10y = 0$, CI : $x_0 = 0$, $y_0 = 1$, $y'_0 = 4$
(3) $y'' + 9y = 0$, CI : $x_0 = \frac{\pi}{3}$, $y_0 = 0$, $y'_0 = 1$
(4) $y'' - 6y' + 9y = 0$, CI : $x_0 = \frac{1}{3}$, $y_0 = 0$, $y'_0 = e$.

corrEquaDiff0011

Correction of the exercise 682

(1) Le polynôme caractéristique est

$$\lambda^2 + 10\lambda + 16 = 0, \tag{96.283}$$

dont les racines sont $\lambda_1 = -8$ et $\lambda_2 = -2$. La solution générale s'écrit donc

$$y(x) = ae^{-8x} + Be^{-2t}. \tag{96.284}$$

Il faut maintenant fixer les paramètres A et B pour que $y(0) = 0$ et $y'(0) = 12$. Pour ce faire, nous calculons $y(x)$:

$$y'(x) = -8Ae^{-8x} - 2Be^{-2x}. \tag{96.285}$$

Nous avons donc les contraintes suivantes sur A et B : $y(0) = A + B = 1$ et $y'(0) = -8A - 2B = 12$. Cela est un petit système de deux équations à deux inconnues que nous résolvons facilement :

$$\begin{aligned} A &= -\frac{7}{3} \\ B &= \frac{10}{3} \end{aligned} \tag{96.286}$$

La solution répondant aux conditions posées est donc

$$\frac{56}{3}e^{-8x} - \frac{20}{3}e^{-2x}. \tag{96.287}$$

(2) Les solutions du polynôme caractéristique sont $\lambda_1 = 3 + i$ et $\lambda_2 = 3 - i$. La première chose à faire est de trouver les solutions réelles qui en découlent. Elles sont

$$y(x) = Ae^{3x}\sin(x) + Be^{3x}\sin(x). \tag{96.288}$$

Il faut maintenant fixer A et B pour que $y(0) = 1$ et $y'(0) = 4$. Le système d'équation à résoudre est

$$\begin{cases} B = 1 & \text{(96.289a)} \\ 3B + A = 4. & \text{(96.289b)} \end{cases}$$

La réponse est $A = 1$ et $B = 1$.

(3) L'équation caractéristique est $\lambda^2 + 9 = 0$, et les racines sont $\lambda = \pm 3i$. Les solutions de l'équation différentielle sont donc les fonctions de la forme

$$y(x) = A\cos(3x) + B\sin(3x). \tag{96.290}$$

Étant donné que $y(\frac{\pi}{3}) = 0$, nous trouvons $-A = 0$ et la condition $y'(\frac{\pi}{3}) = 1$ donne $-3B = 0$, donc en réalité $A = B = 0$ et la solution à l'équation répondant aux conditions posées est la fonction identiquement nulle :

$$y(x) = 0. \tag{96.291}$$

Exercise 683 exoEquaDiff0012

Exercice II.3, page 59. Déterminer la solution générale de chacune des équations différentielles suivantes :

(1) $y'' - 5y' + 4y = 3 - 2x$
(2) $y'' - 6y' + 13y = 39$
(3) $y'' - 2y' + y = 6e^x$
(4) $y'' + 4y = \cos(2x)$

corrEquaDiff0012

Correction of the exercise 683

(1) La solution à l'équation sans second membre est $y_H(x) = Ae^x + Be^{4x}$. Nous devons maintenant trouver une solution particulière à l'équation avec un second membre. Comme le second membre est un polynôme, nous essayons un polynôme, soit $y_P(x) = ax + b$. Nous calculons

$$\begin{aligned} y(x) &= ax + b \\ y'(x) &= a \\ y''(x) &= 0, \end{aligned} \tag{96.292}$$

et nous remettons dans l'équation de départ :

$$0 - 5a + 4(ax + b) = 3 - 2x. \tag{96.293}$$

Cela est une équation pour a et b dont la solution est $a = -1/2$ et $b = 1/8$. La solution particulière que nous avons trouvée est

$$y_P(x) = -\frac{x}{2} + \frac{1}{8}, \tag{96.294}$$

d'où nous déduisons que la solution générale à l'équation donnée est

$$y(x) = Ae^x + Be^{4x} - \frac{x}{2} + \frac{1}{8}. \tag{96.295}$$

(2) Le polynôme caractéristique de l'équation sans second membre est $\lambda^2 - 6\lambda + 13 = 0$, et ses racines sont $3 \pm 2i$. La solution *réelle* générale de l'équation sans second membre est

$$y_H(x) = e^{3x}\big(A\cos(2x) + B\sin(2x)\big). \tag{96.296}$$

Nous devons encore trouver une solution particulière de l'équation homogène. Étant donné que le second membre est une constante, nous essayons une constante. Il est facile de voir que la fonction constante $y_P(x) = 3$ fonctionne.

(3) La solution générale de l'équation sans second membre est

$$y_H(x) = Ae^x + Bxe^x \tag{96.297}$$

Le second membre étant une exponentielle, nous sommes incités à essayer ae^x comme solution particulière. Hélas, cela ne fonctionne pas parce que cela est une solution de l'équation sans second membre. Alors nous devons trouver quelque chose de plus complexe. L'essai que nous faisons est

$$y_P(x) = (ax^2 + bx + c)e^x. \tag{96.298}$$

Étant donné que e^x et xe^x sont des solutions de l'équation sans second membre, il est couru d'avance que les termes $(bx+c)e^x$ ne vont pas jouer. Nous pouvons donc déjà poser $b = c = 0$. Nous avons donc

$$\begin{aligned} y_P(x) &= ax^2e^x \\ y'_P(x) &= 2axe^x + ax^2e^x \\ y'_P(x) &= (2ax + 2a + ax^2 + 2ax)e^x. \end{aligned} \tag{96.299}$$

En remettant tout ça dans l'équation de départ, nous trouvons une équation pour a que nous résolvons. La réponse est que une solution particulière est donnée par $y_P(x) = 3x^2e^x$.

(4) L'équation sans second membre est $y''+4y=0$, dont l'équation caractéristique est $\lambda^2+4=0$. Les solutions sont $\pm 2i$, et donc les solutions réelles sont

$$y_H(x) = A\cos(2x) + B\sin(2x). \tag{96.300}$$

Il s'agit maintenant de trouver une solution particulière. Comme le second membre est $\cos(2x)$, nous voudrions essayer

$$y_P(x) = a\cos(2x) + b\sin(2x), \tag{96.301}$$

mais cela ne va pas fonctionner parce que $\sin(2x)$ et $\cos(2x)$ sont déjà des solutions de l'équation sans second membre. Nous essayons alors

$$y_P(x) = ax\cos(2x) + bx\sin(2x). \tag{96.302}$$

Afin de fixer les constantes a et b, nous injectons cette fonction dans l'équation. Pour cela nous commençons par en calculer les dérivées :

$$y_P'(x) = \sin(2x)(-2ax+b) + \cos(2x)(2bx+a), \tag{96.303}$$

et

$$y_P''(x) = \sin(2x)(-4bx+-4a) + \cos(2x)(-4ax+4b). \tag{96.304}$$

L'équation devient donc

$$y_P''(x) + 4y_P'(x) = -4(a\sin(2x) - b\cos(2x)) = \cos(2x). \tag{96.305}$$

Il faut donc choisir $a=0$ et $b=1/4$. La solution particulière que nous venons de construire est donc

$$y_P(x) = \frac{x}{4}\cos(2x), \tag{96.306}$$

et la solution générale à l'équation qui nous intéresse est

$$y(x) = A\cos(2x) + B\sin(2x) + \frac{x}{4}\cos(2x). \tag{96.307}$$

Exercise 684 exoEquaDiff0013

Exercice 6, page 60. Déterminer la courbe solution de chacune des équations différentielles suivantes, répondant aux conditions imposées.

(1) $y''-2y'=2x$. On demande que la tangente au point $(0,2)$ soit parallèle à l'axe des abscisses.
(2) $y''-6y'+9y=x^2e^{3x}$.
   (2a) la tangente au point $(0,1)$ a un coefficient directeur nul,
   (2b) la tangente au point $(1,0)$ a une pente égale à $e^3/4$.

corrEquaDiff0013

Correction of the exercise 684

(1) La solution générale de l'équation différentielle sans second membre est

$$y_H(x) = A + Be^{2x}. \tag{96.308}$$

Pour trouver une solution particulière de l'équation avec second membre, nous essayons $y_P(x) = ax^2+bx$. En injectant cet essai dans l'équation différentielle, nous trouvons une équation pour a et b dont la solution donne

$$y_P(x) = -\frac{x^2}{2} - \frac{x}{2}. \tag{96.309}$$

La solution générale de l'équation différentielle avec second membre est donc

$$y(x) = A + Be^{2x} - \frac{x^2}{2} - \frac{x}{2}. \tag{96.310}$$

La première condition à poser sur A et B est $y(0) = B + A = 2$. Pour la seconde, nous commençons par calculer $y'(x)$:

$$y'(x) = 2e^{2x}B - t - \frac{1}{2}. \tag{96.311}$$

La second condition à imposer est donc

$$y'(0) = 2B - \frac{1}{2} = 0. \tag{96.312}$$

En résolvant ce petit système algébrique, nous trouvons les valeurs de A et B qui conviennent : $A = 3/4$ et $B = 1/4$.

(2) La solution de l'équation différentielle sans second membre est $y_H(x) = Ae^{3x} + Bxe^{3x}$. Pour trouver une solution particulière, il est certain que $(ax + b)e^{3x}$ ne va pas fonctionner parce que c'est solution de l'équation sans second membre. Il faut donc prendre un degré plus haut. On essaye donc $y_P(x) = ax^4 + bx^3$. En remettant dans l'équation de départ, nous trouvons les valeurs de a et b :

$$y_P(x) = \frac{x^4}{12}. \tag{96.313}$$

La solution générale à l'équation donnée est donc

$$y(x) = Ae^{3x} + Bxe^{3x} + \frac{x^4}{12}e^{3x}. \tag{96.314}$$

Nous devons maintenant trouver les bonnes constantes A et B dans les situations demandées.

(2a) Nous voulons une tangente au point $(0, 1)$, donc en particulier il faut que la courbe passe par là. Donc nous posons la première contrainte : $y(0) = 1$. Ensuite, nous demandons que le coefficient directeur de la tangente en ce point soit nul, c'est-à-dire que $y'(0) = 0$. Cela pose la seconde contrainte :

$$y'(0) = 0 = \frac{1}{12}(12B), \tag{96.315}$$

donc $B = 1$. La solution demandée est donc

$$y(x) = e^{3x} + xe^{3x} + \frac{x^4}{12}e^{3x}. \tag{96.316}$$

(2b) Cette fois, il faut résoudre les contraintes

$$\begin{aligned} y(1) &= 0 \\ y'(1) &= \frac{e^3}{4}. \end{aligned} \tag{96.317}$$

La réponse est que $A = 0$ et $B = -1/12$, et la solution demandée à l'équation différentielle est

$$\frac{xe^{3x}}{12}(x^3 - 1). \tag{96.318}$$

Exercise 685 exoEquaDiff0015

Quelle est la courbe passant par l'origine et admettant en chaque point une pente égale au triple de l'abscisse du point ? corrEquaDiff0015

Correction of the exercise 685

<+EquaDiff0015+>

Exercise 686 exoEquaDiff0014

Déterminer les fonctions nulles à l'origine et telles que leur dérivée première soit égale à l'opposé de la dérivée seconde. corrEquaDiff0014

Correction of the exercise 686

<+EquaDiff0014+>

Exercise 687 exoEquaDiff0016

En tout point d'une courbe, l'angle formé par la tangente en ce point et l'axe des abscisses est complémentaire de l'angle formé par la droite joignant le point à l'origine et l'axe des abscisses. Quelles sont ces courbes ?

corrEquaDiff0016

Correction of the exercise 687

Considérons un point $\big(x, y(x)\big)$ sur la courbe. Nous nommons β l'angle que fait la droite joignant ce point à l'origine, par définition,

$$\tan(\beta) = \frac{y(x)}{x}. \tag{96.319}$$

Si α est l'angle que fait la tangente avec l'axe Ox, alors

$$\tan(\alpha) = y'(x). \tag{96.320}$$

La condition imposée est $\alpha + \beta = 90$. Un peu de trigonométrie montre que

$$\tan(\beta) = \tan(90 - \alpha) = \frac{1}{\tan(\alpha)}, \tag{96.321}$$

et donc

$$y'(x) = \frac{x}{f(x)} \tag{96.322}$$

est l'équation différentielle à résoudre. Nous la remettons sous la forme

$$\frac{dy}{dx} = \frac{x}{y}, \tag{96.323}$$

ou encore : $ydy = xdx$, d'où la solution

$$y^2 = x^2 + C. \tag{96.324}$$

96.8 Fonctions de deux variables réelles

96.8.1 Tracer

Exercise 688 exoFoncDeuxVar0001

Représenter les courbes de niveaux des surfaces suivantes correspondant aux cotes $z = 0$, $z = 1$, $z = -1$, $z = h$.

(1) $z^2 = x^2 + y^2$
(2) $x + y + z = 5$
(3) $z = xy$
(4) $z = x^2 + y^2$
(5) $4x^2 + y^2 = 4$.

corrFoncDeuxVar0001

Correction of the exercise 688

<+FoncDeuxVar0001+>

96.8.2 Limites à deux variables

Exercise 689 exoFoncDeuxVar0010

Calculer les limites suivantes :

(1) $\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}$
(2) $\lim_{(x,y)\to(0,0)} \frac{(xy)^2}{(x+y)^2+(x-y)^2}$

(3) $\lim_{(x,y)\to(0,0)} \frac{xy^3}{x^2+y^2}$

(4) $\lim_{(x,y)\to(0,0)} \frac{x\sin(y)}{\sqrt{x^2+y^2}}$

corrFoncDeuxVar0010

Correction of the exercise 689

(1) Ici la méthode des chemins pour est particulièrement éclairante. Regardons d'abord la fonction sur la droite $x = y$. Nous avons

$$f(x,y) = \frac{x-x}{2x} = 0. \tag{96.325}$$

Donc la fonction est nulle sur toute la ligne.

Si nous regardons maintenant la ligne verticale $x = 0$, nous avons

$$f(0,y) = \frac{-y}{y} = -1, \tag{96.326}$$

donc la fonction vaut -1 sur toute la ligne verticale.

(2)

(3) Regardons la technique des coordonnées polaires. Nous remplaçons x par $r\cos(\theta)$ et y par $r\sin(\theta)$:

$$f(r,\theta) = \frac{r^4\cos(\theta)\sin^3(\theta)}{r^2} = r^2\cos(\theta)\sin^3(\theta). \tag{96.327}$$

Cette fonction tend vers zéro quand $r \to 0$. Nous avons donc

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0. \tag{96.328}$$

Pour cet exercice nous pouvons aussi utiliser la règle de l'étau en écrivant d'abord

$$0 \leqslant |f(x,y)| \leqslant \frac{|x||y^3|}{|x^2+y^2|}. \tag{96.329}$$

Mais on a $|x| \leqslant \sqrt{x^2+y^2}$, $|y| \leqslant \sqrt{x^2+y^2}$ et $|x^2+y^2| = \left(\sqrt{x^2+y^2}\right)^2$, donc

$$0 \leqslant |f(x,y)| \leqslant \frac{\sqrt{x^2+y^2}\left(\sqrt{x^2+y^2}\right)^3}{\left(\sqrt{x^2+y^2}\right)^2} = \left(\sqrt{x^2+y^2}\right)^2 \to 0. \tag{96.330}$$

(4) En passant aux polaires, nous avons

$$f(r,\theta) = \frac{r\cos\theta\sin\big(r\sin\theta\big)}{r} = \cos(\theta)\sin\big(r\sin\theta\big). \tag{96.331}$$

La limite de cette dernière fonction lorsque $r \to 0$ vaut zéro.

Une autre façon de procéder consiste à multiplier et diviser par y de telle façon à faire apparaitre $\sin(y)/y$ dont nous connaissons la limite :

$$f(x,y) = \frac{\sin(y)}{y} \cdot \frac{xy}{\sqrt{x^2+y^2}}. \tag{96.332}$$

La limite du premier facteur est 1, tandis que le second peut être traité de façon classique en prenant la valeur absolue et en majorant $|x|$ par $\sqrt{x^2+y^2}$.

Exercise 690 exoFoncDeuxVar0011

(INGE1121, 8.2) L'énoncé suivant est-il vrai ou faux ? Justifier.

Si la limite en (a,b) de $f(x,y)$ existe sur une infinité de chemins différents passant par (a,b) et vaut chaque fois le même nombre ℓ, alors $\lim_{(x,y)\to(a,b)} f(x,y)$ existe et vaut ℓ. corrFoncDeuxVar0011

Correction of the exercise 690

L'énoncé est faux. Pensez à une fonction dont « tous les chemins sauf un » ont la même limite. Par exemple

$$f(x,y)=\begin{cases}1 & \text{si } y\neq 0\\ 0 & \text{si } y=0\end{cases} \tag{96.333}$$

Cette fonction vaut 1 partout sauf sur une droite. Il y a donc une infinité de chemins suivant lesquels on a une limite en $(0,0)$ qui vaut 1 (tous les chemins qui ne passent pas par $y=0$). Mais on a un chemin pour lequel la limite est zéro (le chemin horizontal $y=0$).

Donc cette fonction n'a pas de limite en $(0,0)$.

Essayez de trouver d'autres exemples. Trouvez un exemple de fonction pour lequel on a une infinité de chemins pour lesquels la limite est 1 et en même temps une infinité de chemins pour lesquels la limite est $\sqrt{2}$.

Exercise 691 exoFoncDeuxVar0012

(INGE1121 8.3) Étudier la continuité des fonctions suivantes :

(1)
$$f(x,y)=\begin{cases}\frac{2xy}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0)\end{cases} \tag{96.334}$$

(2)
$$f(x,y)=\begin{cases}\frac{y^2-x^2}{y^4+x^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0)\end{cases} \tag{96.335}$$

(3)
$$f(x,y)=\begin{cases}y\cos(\frac{1}{y}) & \text{si } y\neq 0\\ 0 & \text{si } y=0\end{cases} \tag{96.336}$$

corrFoncDeuxVar0012

Correction of the exercise 691

L'énoncé « étudier la continuité de » revient à demander « vérifier si la limite est égale à la valeur ». Dans le cadre de cet exercice, nous devons simplement calculer les limites des fonctions et voir si ces limites sont égales à la valeur donnée.

(1) Regardons la ligne $x=y$, c'est-à-dire le chemin $(x,y)=(t,t)$. La limite de f le long de ce chemin vaut
$$\lim_{t\to 0}\frac{2t^2}{t^2+t^2}=1, \tag{96.337}$$
tandis que le long du chemin $x=0$ (c'est-à-dire $(0,t)$) nous avons
$$\lim_{t\to 0}\frac{0}{t^2}=0. \tag{96.338}$$
Les deux limites n'étant pas égales, la limite de f pour $(x,y)\to(0,0)$ n'existe pas et *a forciori* la fonction n'est pas continue.

(2) Elle n'est pas continue parce que la limite n'existe même pas. Prenez par exemple le chemin vertical $x=0$. Le long de ce chemin, la fonction vaut $f(0,y)=\frac{1}{y^2}$ qui n'a certainement pas de limites pour $y\to 0$.

(3) En valeur absolue, $\cos(1/y)$ est borné par 1, donc

$$0 \leqslant |f(x,y)| \leqslant |y| \tag{96.339}$$

et donc elle tend vers zéro lorsque $(x,y) \to (0,0)$. Cette fonction est donc continue.

Exercise 692 exoFoncDeuxVar0013

(INGE 1121, 8.4) Dessiner les courbes de niveau de la fonction

$$f(x,y) = \frac{x^2}{y} \tag{96.340}$$

et calculer $\lim_{(x,y)\to(0,0)} f(x,y)$.

corrFoncDeuxVar0013

Correction of the exercise 692

Les courbes de niveau sont données par

$$y = \frac{x^2}{C}, \tag{96.341}$$

sauf la courbe de hauteur zéro qui est l'axe $x = 0$. Les points $y = 0$ ne font pas partie du domaine.

Étant donné que toutes les courbes de niveau s'intersectent en $(0,0)$, la limite n'existe pas.

Exercise 693 exoFoncDeuxVar0014

(INGE 1121, 8.5) Calculer

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x)}{xy}. \tag{96.342}$$

corrFoncDeuxVar0014

Correction of the exercise 693

La fonction proposée est le produit de $\sin(x)/x$ qui reste bornée (et ne tend pas vers zéro [5]) au abords de $(0,0)$ par la fonction $\frac{1}{y}$ qui n'est pas bornée, et qui n'a pas de limite en $(0,0)$.

La limite de f n'existe donc pas.

Exercise 694 exoFoncDeuxVar0015

(INGE 1121, 8.6) Calculer

$$\lim_{(x,y)\to(0,0)} \sin\left(\sqrt{x^2+y^2}\right) \ln\left(\sqrt{x^2+y^2}\right). \tag{96.343}$$

corrFoncDeuxVar0015

Correction of the exercise 694

Lorsqu'un a un sinus de quelque chose qui tend vers zéro et qu'on ne sait pas très bien quoi faire, une bonne idée est toujours de multiplier et diviser par ce qui est dans le sinus. De cette façon, le sinus disparaît. Ici, nous avons

$$f(x,y) = \frac{\sin\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}} \cdot \sqrt{x^2+y^2} \cdot \ln\sqrt{x^2+y^2}. \tag{96.344}$$

Lors de la limite, le premier morceau est la limite bien connue qui fait 1. Il reste à trouver la limite de la fonction

$$\sqrt{x^2+y^2}\ln\sqrt{x^2+y^2}. \tag{96.345}$$

Cette fonction est la composée $f(x,y) = (g \circ r)(x,y)$ où

$$\begin{aligned} r(x,y) &= \sqrt{x^2+y^2} \\ g(t) &= t\ln(t). \end{aligned} \tag{96.346}$$

5. Vers quoi elle tend ?

Évidement, la limite de r lorsque on tend vers $(0,0)$ est zéro. Voyons la limite de $g(t)$ lorsque $t \to 0$. Cela se règle en utilisant la règle de l'Hopital :

$$\lim_{t\to 0} t\ln(t) = \lim_{t\to 0} \frac{\ln(t)}{1/t} = \lim_{t\to 0} \frac{1/r}{-1/r^2} = \lim_{t\to 0} -t = 0. \tag{96.347}$$

Donc

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0. \tag{96.348}$$

Exercise 695 exoFoncDeuxVar0016

(INGE 1121 8.7) Soit la fonction

$$f(x,y) = \frac{x^2+y^2}{x+y}. \tag{96.349}$$

(1) Donner le domaine de définition de f,

(2) montrer que les courbes de niveau sont des cercles et esquisser les graphes des courbes de niveau de hauteur -2, -1, 1 et 2,

(3) calculer la limite de f lorsque (x,y) tend vers $(0,0)$.

corrFoncDeuxVar0016

Correction of the exercise 695

Le domaine de définition de f est manifestement l'ensemble de (x,y) tels que $x \neq -y$. L'équation de la courbe de niveau de hauteur C est

$$\frac{x^2+y^2}{x+y} = C, \tag{96.350}$$

c'est-à-dire

$$x^2 - Cx + y^2 - Cy = 0. \tag{96.351}$$

Afin de mettre cela sous une forme qui ressemble plus à un cercle, nous reformons les carrés parfait à partir de x^2+Cx et y^2+Cy. Pour cela, nous cherchons a et b tels que

$$x^2 - Cx = (x+a)^2 + b. \tag{96.352}$$ EqCourbeCxSeize

Égalisant les termes de degrés égaux en x, nous trouvons $a = -C/2$ et $b = -C^2/4$. L'équation de la courbe de niveau (96.352) se met donc sous la forme

$$\big(x - \frac{C}{2}\big)^2 - \frac{C^2}{4} + \big(y - \frac{C}{2}\big)^2 - \frac{C^2}{4} = 0 \tag{96.353}$$

Cette dernière équation se récrit avantageusement sous la forme

$$\left(x - \frac{C}{2}\right)^2 + \left(y - \frac{C}{2}\right)^2 = \frac{C^2}{2}, \tag{96.354}$$

ce qui est le cercle de rayon $\frac{|C|}{\sqrt{2}}$ et de centre $(\frac{C}{2}, \frac{C}{2})$. Notez la valeur absolue dans le rayon.

$$\sqrt{\left(\frac{C}{2}\right)^2 + \left(\frac{C}{2}\right)^2} = \frac{|C|}{\sqrt{2}} \tag{96.355}$$

de l'origine. Il est donc hors de question que la limite existe quand $(x,y) \to (0,0)$.

Exercise 696 exoFoncDeuxVar0018

(INGE1121, 8.8)Calculer les limites suivantes ou monter qu'elles n'existent pas.

(1)
$$\lim_{(x,y)\to(0,0)} \frac{2xy^2}{x^4+2y^4}.$$

(2)
$$\lim_{(x,y)\to(2,3)} \frac{x+y}{2-y}$$

(3)
$$\lim_{(x,y,z)\to(0,0,0)} \frac{2xy+yz}{x^2+y^2+z^2}.$$

(4)
$$\lim_{(x,y)\to(\frac{\pi}{2},2)} \frac{xy+1}{2+\cos(x)}.$$

(5)
$$\lim_{(x,y,z)\to(0,0,0)} \frac{2x^2+y^2-z^2}{x^2+y^2+z^2}$$

(6)
$$\lim_{(x,y)\to(1,-1)} \frac{y^2+x}{(x-1)(y+2)}$$

(7)
$$\lim_{(x,y,z)\to(0,0,0)} \frac{x^2+y^3+2z^3}{xyz^2}.$$

(8)
$$\lim_{(x,y)\to(0,0)} \frac{x^4-y^4}{x^2+y^2}.$$

(9)
$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+2y^2}.$$

(10)
$$\lim_{(x,y)\to(0,0)} \frac{x^3-2x^2y+3y^2x-y^3}{x^2+y^2}.$$

(11)
$$\lim_{(x,y)\to(0,0)} \frac{x^2y^3}{y^5-x^5}.$$

(12)
$$\lim_{(x,y,z)\to(0,0,0)} \frac{x^2-z^2}{x^3-z^3}.$$

(13)
$$\lim_{(x,y)\to(1,-1)} e^{-xy^2}.$$

(14)
$$\lim_{(x,y)\to(0,0)} \ln\sqrt{1-x^2-y^2}$$

(15)
$$\lim_{(x,y)\to(0,0)} \frac{x^2-2xy+2y^2}{x^2+2y^2}.$$

(16)
$$\lim_{(x,y)\to(2,-1)} \ln\frac{1+x+3y}{3y^2-x}$$

(17)
$$\lim_{(x,y)\to(0,0)} \frac{e^{xy}\sin^2(xy)}{xy}.$$

(18)
$$\lim_{(x,y)\to(-2,1)} \frac{x^2-4x+4}{xy+2y-x-2}$$

(19)
$$\lim_{(x,y)\to(0,0)} \exp\left(\frac{-1}{x^2+y^2}\right)$$

corrFoncDeuxVar0018

Correction of the exercise 696

(1) Le degré du dénominateur est plus grand que celui du numérateur. Nous nous attendons donc à avoir une limite qui n'existe pas. Prenons le chemin (t,t) par exemple :

$$f(t,t)=\frac{1}{t}, \tag{96.356}$$

et dont la limite le long de ce chemin n'existe pas.

(2) Le point $(2,3)$ est à l'intérieur du domaine de continuité de la fonction, donc sa limite est égale à la valeur de la fonction au point donné, c'est-à-dire -5. En pratique, il s'agit juste de remplacer x et y par les valeurs 2 et 3.

(3) Un bon chemin à tenter est toujours (t,kt,kt). Nous avons

$$f(t,kt,kt)=\frac{k^2+2k}{1+2k^2}, \tag{96.357}$$

et donc la limite dépend de k. Deux k différents correspondent donc à deux chemins le long desquels la limite est différente.

La limite n'existe donc pas.

(4) Le domaine de continuité de cette fonction est $\mathbb{R}^2$ tout entier parce que c'est une fraction de fonctions continues dont le dénominateur ne s'annule pas. Il n'y a donc aucune indéterminations et la limite s'obtient en remplaçant par les valeurs : $(\pi + 1)/2$.

(5) C'est une limite de polynômes dont le numérateur et le dénominateurs ont le même degré. Utiliser le chemin (t, kt, kt) est en général une bonne idée : le résultat a de bonnes chances de dépendre de k et donc la limite de ne pas exister. Ici nous avons

$$f(t, kt, kt) = \frac{2}{1 + 2k^2}. \tag{96.358}$$

Cette expression dépend effectivement de k et donc deux k différents correspondent à deux chemins le long desquels la limite de f est différente.

(6) Le domaine de continuité de cette fonction est tous les points de $\mathbb{R}^2$ sauf les points avec $x = 1$ ou $y = -2$. Étant donné que la limite que l'on demande n'est pas vers un de ces points, nous la calculons en remplaçant les valeurs :

$$\lim_{(x,y)\to(-1,1)} \frac{y^2 + x}{(x - 1)(y + 2)} = \frac{0}{-6} = 0. \tag{96.359}$$

J'insiste : 0/6 n'est pas une indétermination.

(7) Si nous prenons le chemin (t, kt, kt), nous trouvons

$$\lim_{t\to 0} f(t, kt, kt) = \lim_{t\to 0} \frac{1 + 3k^3t}{k^3t^2} = \infty. \tag{96.360}$$

La limite de la fonction n'existant pas le long de ces chemins, la fonction n'a pas de limite en $(0, 0, 0)$.

(8) Grâce au produit remarquable $x^4 - y^4 = (x^2 - y^2)(x^2 + y^2)$, nous pouvons simplifier la fraction et obtenir $f(x, y) = x^2 - y^2$, dont la limite vaut zéro.

(9) Le degré du numérateur (3) est plus grand que celui du dénominateur, donc nous partons avec l'idée que la limite va exister et sera zéro. Prouvons cela. D'abord

$$0 \leqslant |f(x, y)| = \frac{|x||y|^2}{x^2 + 2y^2}. \tag{96.361}$$

Ensuite nous utilisons les majorations quasi classiques (mais un peu adaptées au $2y^2$) $|x| \leqslant \sqrt{x^2 + y^2}$ et $|y| \leqslant \sqrt{x^2 + 2y^2}$. Donc

$$0 \leqslant \frac{\sqrt{x^2 + 2y^2}\big(\sqrt{x^2 + 2y^2}\big)^2}{\big(\sqrt{x^2 + y^2}\big)^2} = \sqrt{x^2 + 2y^2}, \tag{96.362}$$

mais la limite de l'expression à droite vaut manifestement zéro. Donc la limite de la fonction est zéro.

(10) Prouvons que la limite est nulle (parce que le degré du numérateur est plus grand que celui du dénominateur). Nous majorons la valeur absolue du numérateur par

$$|x^3 - 2x^2y + 3y^2x - y^3| \leqslant |x|^3 + 2x^2|y| + 3y^2|x| + |y|^3. \tag{96.363}$$

Ensuite nous utilisons la majoration $|x| \leqslant \sqrt{x^2 + y^2}$ (idem pour y) et après simplification nous trouvons

$$0 \leqslant |f(x, y)| \leqslant \frac{9}{2}\sqrt{x^2 + y^2}, \tag{96.364}$$

dont la limite vaut zéro.

(11) ATTENTION : il ne suffit pas de dire que sur le chemin (t,t) la limite n'existe pas. En effet, dans la technique des chemins, on demande que les chemins soient compris dans le domaine de définition de la fonction. Or aucun point de la forme (t,t) n'est dans le domaine de définition de f.

Pour cette fonction, les droites (t,kt) fonctionnent très bien (penser à rejeter $k=1$). En effet,

$$f(t,kt)=\frac{k^3t^5}{t^5(k^5-1)}=\frac{k^3}{k^5-1}, \tag{96.365}$$

qui dépend de k. Nous avons donc tout plein de chemins qui ont des limites différentes.

(12) En utilisant une division euclidienne, nous factorisons le numérateur en

$$x^3-y^3=(x-z)(x^2+xz+z^2), \tag{96.366}$$

tandis que le numérateur se factorise en $(x+z)(x-z)$. Après simplification nous trouvons

$$f(x,y)=\frac{x+z}{x^2+xz+z^2}, \tag{96.367}$$

dont le domaine de continuité contient maintenant le point $(1,1,1)$. La limite se trouve donc en remplaçant x, y et z par leurs valeurs. La réponse est $2/3$.

(13) Nous voyons cette fonction comme composée $f=h\circ g$ avec

$$\begin{aligned} g(x,y)&=-xy^2\\ h(t)&=e^t. \end{aligned} \tag{96.368}$$

Ce sont deux fonctions qui sont partout continues, donc la fonction f est partout continue. De ce fait, la limite se calcule simplement en remplaçant les valeurs. Nous avons e^{-1}.

(14) Le domaine d'existence de la fonction est donné par $x^2+y^2<1$. Le point proposé est bien à l'intérieur de ce domaine, donc il suffit de remplacer par les valeurs, on obtient donc

$$\ln\sqrt{1}=\ln(1)=0. \tag{96.369}$$

(15) Le chemin $(t,0)$ donne $\frac{t^2}{t^2}\to 1$. Mais le chemin (t,t) donne $\frac{t^2}{3t^2}\to\frac{1}{3}$. La limite de la fonction n'existe donc pas.

(16) C'est la composée de $g(x,y)=(1+x+3y)/(3y^2-x)$ et de $h(t)=\ln(t)$. Nous avons que $\lim_{(x,y)\to(2,-1)}g(x,y)=0$, et que la limite du logarithme pour $t\to 0$ est moins l'infini. La limite n'existe donc pas.

(17) Cette fonction est la composée $f(x,y)=h\big(g(x,y)\big)$ où $g(x,y)=xy$ et

$$h(t)=\frac{e^t\sin^2(t)}{t}. \tag{96.370}$$

Il faut donc calculer la limite

$$\lim_{t\to 0}\frac{e^t\sin^2(t)}{t}. \tag{96.371}$$

Pour ce faire, nous faisons la manipulation suivante :

$$\frac{e^t\sin^2(t)}{t}=\frac{e^t\sin^2(t)}{t}\frac{t}{t}=te^t\left(\frac{\sin(t)}{t}\right)^2. \tag{96.372}$$

Lorsque nous faisons la limite $t\to 0$, la grande parenthèse tend vers 1, tandis que $te^t\to 0$.

Nous avons donc que la limite de f vaut zéro.

Une autre façon de résoudre cet exercice est de multiplier et diviser par xy, de telle façon à faire apparaitre le terme

$$\left(\frac{\sin(xy)}{xy}\right)^2 \tag{96.373}$$

dont la limite est 1.

(18) En remplaçant les valeurs de x et y nous trouvons $16/0$, ce qui n'existe pas. Afin de prouver que la limite n'existe effectivement pas, prenons un chemin. Le chemin $(t-2,1+t)$ fait l'affaire :

$$f(t-2,1+t)=\frac{(t-2)^2-4(t-2)+4}{(t-2)(1+t)+2(1+t)-(t-2)-2}=\frac{t^2-8t+16}{t^2}, \tag{96.374}$$

dont la limite $t\to 0$ est ∞. La limite de cet exercice n'existe donc pas.

(19) Nous pouvons voir la fonction comme composée $f=h\circ g$ avec

$$g(x,y)=\frac{1}{x^2+y^2}h(t) \quad =e^{-t}. \tag{96.375}$$

Dans ce cas, la limite de g vaut ∞ et ensuite, $e^{-\infty}=0$. Cela fournit la bonne réponse mais l théorème 12.488 sur lequel se base la méthode des composées ne permet pas de passer l'infini de g à h aussi facilement. La solution pour être plus rigoureux est de poser

$$g(x,y)=x^2+y^2h(t) \quad =e^{-1/t}. \tag{96.376}$$

D'ailleurs, lorsque x et y n'arrivent que dans la combinaison x^2+y^2, c'est toujours une bonne idée de faire une composée de cette façon.

Maintenant, la limite de g est zéro et $\lim_{x\to 0}e^{-1/t}=0$.

96.8.3 Dérivées partielles, différentielles totales

Exercice 697 exoFoncDeuxVar0002

Calculer les dérivées partielles $\partial_x f$, $\partial_y f$, $\partial_x^2 f$, $\partial_{xy}^2 f$, $\partial_{yx}^2 f$ et $\partial_y^2 f$ des fonctions suivantes.

(1) $2x^3+3x^2y-2y^2$

(2) $\ln(xy^2)$

(3) $\tan(x/y)$

(4) $\frac{xy^2}{x+y}$

corrFoncDeuxVar0002

Correction of the exercise 697

Le script Sage suivant (`exoDV002.sage`) résout l'exercice :

```
# -*- coding: utf8 -*-

def LesCalculs(f):
   print "Pour la fonction %s"%str(f)
   print "d_x",f.diff(x).simplify_full()
   print "d_y",f.diff(y).simplify_full()
   print "d^2_x",f.diff(x).diff(x).simplify_full()
   print "d_xd_y",f.diff(x).diff(y).simplify_full()
   print "d_yd_x",f.diff(y).diff(x).simplify_full()
   print "d^2_y",f.diff(y).diff(y).simplify_full()
   print ""

def exercise_DV002():
   var('x,y')
   fa(x,y)=2*x**3+3*x**2*y-2*y**2
   fb(x,y)=ln(x*y**2)
   fc(x,y)=tan(x/y)
   fd(x,y)=x*y**2/(x+y)
   LesCalculs(fa)
   LesCalculs(fb)
   LesCalculs(fc)
   LesCalculs(fd)
```

La sortie est :

```
Pour la fonction (x, y) |--> 2*x^3 + 3*x^2*y - 2*y^2
d_x (x, y) |--> 6*x^2 + 6*x*y
d_y (x, y) |--> 3*x^2 - 4*y
d^2_x (x, y) |--> 12*x + 6*y
d_xd_y (x, y) |--> 6*x
d_yd_x (x, y) |--> 6*x
d^2_y (x, y) |--> -4

Pour la fonction (x, y) |--> log(x*y^2)
d_x (x, y) |--> 1/x
d_y (x, y) |--> 2/y
d^2_x (x, y) |--> -1/x^2
d_xd_y (x, y) |--> 0
d_yd_x (x, y) |--> 0
d^2_y (x, y) |--> -2/y^2

Pour la fonction (x, y) |--> tan(x/y)
d_x (x, y) |--> 1/(y*cos(x/y)^2)
d_y (x, y) |--> -x/(y^2*cos(x/y)^2)
d^2_x (x, y) |--> 2*sin(x/y)/(y^2*cos(x/y)^3)
d_xd_y (x, y) |--> -(2*x*sin(x/y) + y*cos(x/y))/(y^3*cos(x/y)^3)
d_yd_x (x, y) |--> -(2*x*sin(x/y) + y*cos(x/y))/(y^3*cos(x/y)^3)
d^2_y (x, y) |--> 2*(x^2*sin(x/y) + x*y*cos(x/y))/(y^4*cos(x/y)^3)

Pour la fonction (x, y) |--> x*y^2/(x + y)
d_x (x, y) |--> y^3/(x^2 + 2*x*y + y^2)
d_y (x, y) |--> (2*x^2*y + x*y^2)/(x^2 + 2*x*y + y^2)
d^2_x (x, y) |--> -2*y^3/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d_xd_y (x, y) |--> (3*x*y^2 + y^3)/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d_yd_x (x, y) |--> (3*x*y^2 + y^3)/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
d^2_y (x, y) |--> 2*x^3/(x^3 + 3*x^2*y + 3*x*y^2 + y^3)
```

Exercise 698 exoFoncDeuxVar0003

Si $z = x^2y - 3y$, calculer Δz et dz. corrFoncDeuxVar0003

Correction of the exercise 698

<+FoncDeuxVar0003+>

96.8.4 Différentiabilité, accroissements finis

Exercise 699 exoFoncDeuxVar0019

(INGE1121, 9.1) Étudier la différentiabilité en $(0,0)$ des fonctions suivantes

(1)

$$f(x,y) = \begin{cases} \frac{8x^3y^3}{4x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \tag{96.377}$$

(2)

$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \tag{96.378}$$

(3)

$$f(x,y) = \begin{cases} x \ln \sqrt{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \tag{96.379}$$

corrFoncDeuxVar0019

Correction of the exercise 699

Pour prouver la différentiabilité, nous allons prouver que les dérivées partielles $\partial_x f$ et $\partial_y f$ sont continues. La fonction sera alors différentiable par le théorème 12.490.

(1) Calculons

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{24x^2y^3(4x^2+y^2)-64x^4y^3}{(4x^2+y^3)^2}\\ &= \frac{8x^2y^3(4x^2+3y^2)}{(4x^2+y^2)}.\end{aligned} \tag{96.380}$$

Cette fonction est continue partout en dehors du point $(0,0)$ (parce que c'est une fraction de polynômes). Au point $(0,0)$, elle est également continue parce qu'elle a une limite qui vaut zéro.

Attention : ce qui fait la continuité de la *dérivée* n'est pas le fait que la limite de $\partial_x f$ soit zéro, mais simplement le fait que cette limite existe. Le fait qu'ici la dérivée vaille numériquement la même chose que la fonction elle-même est une coïncidence.

En ce qui concerne l'autre dérivée partielle nous avons la même chose.

Les dérivées partielles étant continues, la fonction est différentiable.

(2) Pas différentiable.

(3) La dérivée partielle dans la direction x est

$$\frac{\partial f}{\partial x} = \ln\sqrt{x^2+y^2} + \frac{2x^2}{\sqrt{x^2+y^2}}. \tag{96.381}$$

La limite du premier terme a l'air problématique, et le second terme tendant simplement vers zéro, il n'y a aucune chance qu'il compense. Pour en être sûr, prenons le chemin (t,t) et calculons

$$\lim_{t\to 0} \ln(|t|) + \frac{2}{\sqrt{2}}t. \tag{96.382}$$

Cette limite n'existe pas.

La fonction proposée n'est pas différentiable en $(0,0)$. Pas besoin de calculer la dérivée partielle par rapport à y.

Exercise 700 exoMaximisation-0001

Soit la fonction

$$f(x,y) = \begin{cases} \frac{x^2y^a}{(x^2+y^2)^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{sinon.}\end{cases} \tag{96.383}$$

Prouver que f admet des dérivées partielles seulement si $a>0$. corrMaximisation-0001

Correction of the exercise 700

En ce qui concerne $\partial_y f$, il n'y a aucun problème parce que

$$\frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = 0 \tag{96.384}$$

parce que $f(0,t)=f(0,0)=0$ pour tout $t\neq 0$.

La dérivée par rapport à x est plus douteuse parce que $f(t,0)$ existe (et vaut 0) lorsque $a>0$ du fait que $y^a=0$ se trouve au numérateur, mais $f(t,0)$ n'existe pas lorsque $a<0$ parce que $y^a=0$ se trouve au dénominateur. La limite ne peut donc pas être calculée et la dérivée partielle $\partial_x f(0,0)$ n'existe pas.

Exercise 701 exoFoncDeuxVar0026

(INGE1121, 9.2) Démontrer que les fonctions suivantes sont différentiables sur leur domaine de définition.

(1)
$$f(x,y)=\frac{xy}{y^2-x^2} \tag{96.385}$$
(2)
$$f(x,y,z)=\frac{xyz}{x-y}. \tag{96.386}$$

corrFoncDeuxVar0026

Correction of the exercise 701

(1)

(2) Cette fonction n'existe pas sur la droite $x=y$. Ailleurs elle est un quotient de polynômes et est donc différentiable.

Exercise 702 exoFoncDeuxVar0021

(INGE1121 9.3) Soit la fonction f de $\mathbb{R}^2$ dans $\mathbb{R}^3$ définie par
$$f(x,y)=\big(y\sin(x),e^xy^2,x^2y\big). \tag{96.387}$$

(1) Donner la matrice jacobienne de f.

(2) Calculer $f(0,1)$ et utiliser la jacobienne pour approcher $f(0.01,1.02)$.

corrFoncDeuxVar0021

Correction of the exercise 702

(IGNE 1121, 9.3) La fonction f proposée est formée des trois fonctions
$$\begin{aligned}f_1(x,y)&=y\sin(x)\\ f_2(x,y)&=e^xy^2\\ f_3(x,y)&=x^2y.\end{aligned} \tag{96.388}$$

Il s'agit maintenant de calculer toutes les dérivées partielles au point $(0,1)$ et puis de les mettre en matrice. Nous avons les dérivées partielles
$$\begin{aligned}\frac{\partial f_1}{\partial x}&=y\cos(x) & \frac{\partial f_1}{\partial y}&=\sin(x)\\ \frac{\partial f_2}{\partial x}&=e^xy^2 & \frac{\partial f_2}{\partial y}&=2ye^x\\ \frac{\partial f_3}{\partial x}&=2xy & \frac{\partial f_3}{\partial y}&=x^2.\end{aligned} \tag{96.389}$$

En évaluant tout ça au point $(0,1)$ la matrice jacobienne devient
$$J_f(0,1)=\begin{pmatrix}1&0\\1&2\\0&0\end{pmatrix}. \tag{96.390}$$

Lorsqu'on calcule $f(x+a,x+b)$ où l'on suppose que a et b sont petits, on a la notation compacte
$$f(x+a,y+b)=f(x,y)+J(x,y)\cdot\begin{pmatrix}a\\b\end{pmatrix}. \tag{96.391}$$

Ici nous avons évidement $x=0$, $y=1$, $a=0.01$ et $b=0.02$ et donc
$$f(0+0.01,1+0.02)\simeq f(0,1)+(0.01,0.05,0)=(0.01,1.05,0). \tag{96.392}$$

Exercise 703 exoFoncDeuxVar0022

(INGE1121, 9.4) Donner la jacobienne des fonctions suivantes

(1)
$$f(x,y)=xy\sin(xy) \tag{96.393}$$
(2)
$$f(x,y)=4x-5y \tag{96.394}$$
(3)
$$f(x,y)=\frac{x}{2x+5x} \tag{96.395}$$
(4)
$$f(x,y,z)=\frac{xy}{z} \tag{96.396}$$

corrFoncDeuxVar0022

Correction of the exercise 703

Pour les fonctions à valeurs réelles, la jacobienne n'est autre que le vecteur gradient, c'est-à-dire les différentes dérivées partielles mises les unes à côté des autres.

(1) Ici nous avons
$$\begin{aligned}\frac{\partial f}{\partial x}&=y\sin(xy)+xy^2\cos(xy)\\ \frac{\partial f}{\partial y}&=x^2y\cos(xy)+x\sin(xy),\end{aligned} \tag{96.397}$$
et la matrice jacobienne est la matrice ligne
$$\begin{pmatrix}y\sin(xy)+xy^2\cos(xy) & x^2y\cos(xy)+x\sin(xy).\end{pmatrix} \tag{96.398}$$

(2)

(3)

Exercise 704 exoFoncDeuxVar0023

(INGE1121, 9.5) Pour la fonction
$$g(u,v,w)=uv-(\cos u)v+u^2w, \tag{96.399}$$
donner une approximation de $g(r,1+s,t)$ pour r, s et t petits.

corrFoncDeuxVar0023

Correction of the exercise 704

Nous partons du point $(0,1,0)$ et nous faisons un pas (r,s,t). La nouvelle valeur de la fonction sera approximée par
$$g(0,1,0)+r\frac{\partial g}{\partial u}(0,1,0)+s\frac{\partial g}{\partial v}(0,1,0)+w\frac{\partial g}{\partial w}(0,1,0). \tag{96.400}$$
Nous avons
$$\begin{aligned}g(0,1,0)&=-1\\ \frac{\partial g}{\partial u}&=v+\sin(u)v+2uv=1\\ \frac{\partial g}{\partial v}&=u-\cos(u)=-1\\ \frac{\partial g}{\partial w}&=u^2=0.\end{aligned} \tag{96.401}$$
Au final,
$$g(r,1+s,t)=-1+r-s. \tag{96.402}$$

Exercise 705 exoDerrivePartielle-0000

(INGE1121, 9.6) Pour chacune des fonctions suivantes, déterminer la dérivée directionnelle au point P dans la direction du vecteur u.

(1) $f(x,y)=2x+y$, $P-(3,1)$, $v=(2,-3)$.

(2)

(3)

(4) $f(x,y) = x\ln(y)$, $P = (2,2)$, $(1,-1)$.

(5) $f(x,y,z) = \sin(xyz)$, $P = (\pi,1,1)$, $v = (2,1,2)$.

corrDerrivePartielle-0000

Correction of the exercise 705

(1) Les dérivées partielles sont faciles :

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2\\ \frac{\partial f}{\partial y} &= 1.\end{aligned} \tag{96.403}$$

Pour dériver dans une direction, il faut considérer le vecteur de norme 1 dans la direction donnée. Ici $\|v\| = \sqrt{4+9} = \sqrt{13}$, et la dérivée demandée vaut

$$\frac{\partial f}{\partial v} = \frac{2}{\sqrt{13}}\frac{\partial f}{\partial x} + \frac{-3}{\sqrt{13}}\frac{\partial f}{\partial y} = \frac{1}{\sqrt{13}}. \tag{96.404}$$

(2)

(3)

(4) Les dérivées partielles sont

$$\begin{aligned}\frac{\partial f}{\partial x} &= \ln(y)\\ \frac{\partial f}{\partial y} &= \frac{x}{y}.\end{aligned} \tag{96.405}$$

En y substituant le point $x = 2$, $y = 2$, nous avons $\partial_x f(P) = \ln(2)$ et $\partial_y f(P) = 1$.
Le vecteur normalisé dans la direction v est $(1/\sqrt{2}, -1/\sqrt{2})$. Donc nous avons

$$\frac{\partial f}{\partial v}(P) = \frac{1}{\sqrt{2}}\ln(2) - \frac{1}{\sqrt{2}}. \tag{96.406}$$

(5) Les dérivées partielles sont

$$\begin{aligned}\frac{\partial f}{\partial x} &= yz\cos(xyz) & \frac{\partial f}{\partial x}(P) &= -1\\ \frac{\partial f}{\partial y} &= xz\cos(xyz) & \frac{\partial f}{\partial y}(P) &= -\pi\\ \frac{\partial f}{\partial z} &= xy\cos(xyz) & \frac{\partial f}{\partial z}(P) &= -\pi\end{aligned} \tag{96.407}$$

Le vecteur normé dans la direction donnée est $v = (2/3, 1/3, 2/3)$, et la dérivée directionnelle est

$$\frac{\partial f}{\partial v}(P) = -\frac{2}{3} - \pi. \tag{96.408}$$

Exercise 706 exoDerrivePartielle-0001

(INGE1121, 9.7) Soit f une fonction de classe C^1 vérifiant $f(x,y) = f(y,x)$ pour tout $x, y \in \mathbb{R}$. Montrer au moyen d'un exemple que l'une des égalités suivantes est fausse, et démontrer l'autre :

$$\frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial x}(b,a) \tag{96.409a}$$ subEqDPZUa

$$\frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial y}(b,a) \tag{96.409b}$$ subEqDPZUb

(96.409c)

corrDerrivePartielle-0001

Correction of the exercise 706

La fonction $f(x,y)=xy$ est un exemple qui fonctionne bien parce que

$$\frac{\partial f}{\partial x}(x,y)=y. \tag{96.410}$$

Nous avons alors $\partial_x f(a,b)=b$ et $\partial_x f(b,a)=a$ qui ne sont pas égaux. Cela prouve que la relation (96.409a) n'est pas correcte.

Prouvons maintenant que la relation (96.409b) est vraie. En écrivant la définition de la dérivée partielle, nous avons

$$\begin{aligned}\frac{\partial f}{\partial x}&=\lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}\\&=\lim_{h\to 0}\frac{f(b,a+h)-f(b,a)}{h}\\&=\frac{\partial f}{\partial y}(b,a)\end{aligned} \tag{96.411}$$

où nous avons utilisé le fait que $f(a+h,b)=f(b,a+h)$.

Exercise 707 exoFoncDeuxVar0025

(INGE1121, 9.8) La production d'une société est donnée par $Q(K,L)=60K^{1/2}L^{1/3}$, où K représente le capital investi (une unité correspondant à 1000 euros) et où L désigne le travail mesuré en heures de travail. On suppose que $K=900.000$ et $L=1000$. Estimer l'effet de l'addition de 1000 euros au capital.

corrFoncDeuxVar0025

Correction of the exercise 707

Nous voulons voir ce que ça donne sur f d'ajouter 1 à x pour la fonction

$$f(x,y)=60x^{1/2}y^{1/3} \tag{96.412}$$

lorsque $x=900.000$ et $y=1000$. Nous utilisons la formule des accroissements finis avec les dérivées partielles :

$$f(900.000+1,1000)=(900.000,1000)+1\cdot\frac{\partial f}{\partial x}(900.000,1000). \tag{96.413}$$ EqDVVCpref

La dérivée partielle vaut

$$\frac{\partial f}{\partial x}=\frac{60}{2\sqrt{x}}y^{1/3}. \tag{96.414}$$

En remplaçant avec le nombres donnés, la formule (96.413) donne

$$\begin{aligned}f(x,y)&=60\cdot 300\sqrt{10}\cdot 10=18.000\sqrt{10}\\(\partial_x f)(900.000,1000)&=\frac{60\cdot 10}{2\cdot 300\sqrt{10}}=\frac{1}{\sqrt{10}}.\end{aligned} \tag{96.415}$$

Donc en ajoutant 1000 euros au capital, on n'augmente que de $1/\sqrt{10}$ la productivité.

96.8.5 Plan tangent

Exercise 708 exoFoncDeuxVar0027

(INGE1121, 9.10) Chercher les équations du plan tangent et de la droite normale au point P aux graphes de la fonction f dans les cas suivants

(1) $f(x,y)=4x^2+9y^2$, avec $P=(-2,-1,25)$
(2) $f(x,y)=2e^{-x}\cos(y)$ avec $P=(0,\frac{\pi}{3},1)$
(3) $f(x,y)=\ln\frac{x+y}{2y}$ avec $P=(0,2,\ln(1/2))$.

corrFoncDeuxVar0027

Correction of the exercise 708

(1) Calculer les dérivées partielles au point $(x, y) = (-2, -1)$:

$$\begin{aligned}\frac{\partial f}{\partial x} &= 8x = -16\\ \frac{\partial f}{\partial y} &= 18y = -18.\end{aligned} \tag{96.416}$$

Le plan tangent est donné par

$$\begin{aligned}T_{(-2,-1)}(x, y) &= 25 + -16(x + 2) - 18(y + 1)\\ &= -16x - 18y - 25.\end{aligned} \tag{96.417}$$

En guise de petite vérification, noter que $T_{(-2,-1)}(-2, -1) = 25 = f(-2, -1)$.
Des vecteurs directeurs du plan peuvent être trouvés en prenant deux vecteurs dans le plan translaté à l'origine :

$$U(x, y) = -16x - 18y. \tag{96.418}$$

Deux points dans le plan U sont par exemple

$$\begin{aligned}P_1 &= (1, 0, -16)\\ P_2 &= (0, 1, -18).\end{aligned} \tag{96.419}$$

Pour trouver un vecteur perpendiculaire au plan U nous prenons le produit vectoriel de P_1 et P_2 :

$$v = P_1 \times P_2 = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -16 \\ 0 & 1 & -18 \end{vmatrix} = \begin{pmatrix} 16 \\ 18 \\ 1 \end{pmatrix}. \tag{96.420}$$

Toute droite normale au plan est parallèle à ce vecteur.
Afin de trouver la droite normale au plan T passant par le point $(-2, -1, 25)$ nous faisons une translation :

$$r(\lambda) = \begin{pmatrix} -2 \\ -1 \\ 25 \end{pmatrix} + \lambda \begin{pmatrix} 16 \\ 18 \\ 1 \end{pmatrix} = \begin{pmatrix} 16\lambda - 2 \\ 18\lambda - 1 \\ \lambda + 25 \end{pmatrix}. \tag{96.421}$$

(2)

(3) Nous commençons par calculer les dérivées partielles qui serviront de « coefficients angulaires » pour le plan tangent. En l'occurrence les dérivées partielles sont

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{1}{x + y}\\ \frac{\partial f}{\partial y} &= \frac{-x}{xy + y^2}.\end{aligned} \tag{96.422}$$

Nous évaluons ces dérivées partielles en $(x, y) = (0, 2)$:

$$\begin{aligned}(\partial_x f)(0, 2) &= \frac{1}{2}\\ (\partial_y f)(0, 1) &= 0.\end{aligned} \tag{96.423}$$

L'équation du plan tangent est alors

$$T_P(x, y) = f(P) + x \cdot \frac{\partial f}{\partial x}(P) + (y - 2)\frac{\partial f}{\partial y}(P). \tag{96.424}$$

Le second terme est nul et il reste

$$T_{(0,2)}(x, y) = \ln\left(\frac{1}{2}\right) + \frac{x}{2}. \tag{96.425}$$

Pour trouver la normale, nous allons trouver deux vecteurs dans le plan et en prendre le produit vectoriel. Attention cependant : il faut réellement prendre des vecteurs dans le plan, ce qui est délicat pour ce plan parce qu'il ne passe pas par l'origine. Pour éviter tout problèmes, nous commençons par prendre le plan

$$U(x,y)=\frac{x}{2}. \tag{96.426}$$

Deux vecteurs de ce plan sont par exemple

$$\begin{aligned} v&=(1,0,\frac{1}{2})\\ w&=(0,1,0). \end{aligned} \tag{96.427}$$

Le vecteur perpendiculaire est alors donné par le produit vectoriel

$$u\times w=\begin{vmatrix} e_x & e_y & e_z\\ 1 & 0 & 1/2\\ 0 & 1 & 0 \end{vmatrix}=\begin{pmatrix} -1/2\\ 0\\ 1 \end{pmatrix}. \tag{96.428}$$

La droite que nous recherchons est celle qui est parallèle à ce vecteur et qui passe par le point $P=(0,2,1)$. L'équation paramétrique est donnée par

$$r(\lambda)=\begin{pmatrix} 0\\ 2\\ 1 \end{pmatrix}+\lambda\begin{pmatrix} -1/2\\ 0\\ 1 \end{pmatrix}=\begin{pmatrix} -\lambda/2\\ 2\\ 1+\lambda \end{pmatrix}. \tag{96.429}$$

Exercise 709 exoDerrivePartielle-0002

(INGE1121, 9.11) En quels points de la surface $z=4x^2+9y^2$ la normale est-elle parallèle à la droite passant par les deux points $P=(-2,4,3)$ et $Q=(5,-1,2)$? corrDerrivePartielle-0002

Correction of the exercise 709

Le vecteur directeur qu'on nous donne est

$$v=P-Q=(-7,5,1). \tag{96.430}$$

Nous devons trouver un point du graphe où le vecteur normal est parallèle à v. Pour ce faire, nous calculons le vecteur normal au point $\big(x,y,z(x,y)\big)$ pour un (x,y) quelconque, et puis nous fixerons x et y.

Les dérivées partielles de z par rapport à x et y sont

$$\begin{aligned} \frac{\partial z}{\partial x}&=8x\\ \frac{\partial z}{\partial y}&=18y. \end{aligned} \tag{96.431}$$

Au point (x,y) cela donne les deux vecteurs tangents

$$\begin{aligned} t_1&=(1,0,8x)\\ t_2&=(0,1,8y). \end{aligned} \tag{96.432}$$

Le vecteur normal est donné par

$$n=t_1\times t_2=\begin{vmatrix} e_x & e_y & e_z\\ 1 & 0 & 8x\\ 0 & 1 & 18y \end{vmatrix}=\begin{pmatrix} -8x\\ -18y\\ 1 \end{pmatrix}. \tag{96.433}$$

Afin que n soit parallèle à v (c'est-à-dire soit un multiple) nous cherchons à résoudre l'équation

$$\begin{pmatrix} -7\\ 5\\ 1 \end{pmatrix}=\lambda\begin{pmatrix} -8x\\ -18y\\ 1 \end{pmatrix}. \tag{96.434}$$

Il faut résoudre cela par rapport à λ, x et y. Il n'y a cependant que x et y qui nous intéressent.

La dernière ligne montre que $\lambda = 1$ ensuite les deux autres lignes fournissent les équations

$$\begin{cases} -7 = -8x & \text{(96.435a)} \\ 5 = -18y, & \text{(96.435b)} \end{cases}$$

et donc le point recherché est $(\frac{7}{8}, -\frac{5}{18})$.

96.8.6 Dérivées de fonctions composées

Exercise 710 exoDerrivePartielle-0003

(INGE1121, 9.12) Calculer $\frac{\partial w}{\partial x}$ et $\frac{\partial w}{\partial y}$ dans les cas suivants :

(1) $w = u\cos(v)$, $u = x^2 + y^2$, $v = xy$.

(2) $w = 2uv + 3v^2$, $u = x\sin(y)$, $v = y\sin(x)$.

corrDerrivePartielle-0003

Correction of the exercise 710

(1) La formule de dérivation des fonctions composées, dans ce cas, devient

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial u}\big(u(x,y), v(x,y)\big)\frac{\partial u}{\partial x}(x,y) + \frac{\partial w}{\partial v}\big(u(x,y), v(x,y)\big)\frac{\partial v}{\partial x}(x,y). \tag{96.436}$$

Les calculs donnent

$$2x\cos(v) - u\sin(v) = 2x\cos(x,y) - (x^2 + y^2)\sin(xy). \tag{96.437}$$

En ce qui concerne la dérivée dans le sens de y nous avons

$$\frac{\partial w}{\partial y}(x,y) = -(x^2 + y^2)x\sin(xy) + 2y\cos(xy) \tag{96.438}$$

(2) Ici c'est le même jeu, les réponses sont

$$\begin{aligned} \frac{\partial w}{\partial x} &= 2xy\sin(y)\cos(x) + 6y^2\sin(x)\cos(x) + 2y\sin(x)\sin(y) \\ \frac{\partial w}{\partial y} &= 2xy\sin(x)\cos(y) + 2x\sin(x)\sin(y) + 6y\sin^2(x) \end{aligned} \tag{96.439}$$

Exercise 711 exoFoncDeuxVar0017

(INGE1121, 9.13) Calculer $\frac{\partial w}{\partial r}$ et $\frac{\partial w}{\partial s}$ dans les cas suivants :

(1) $w = u^2 + uv$, $u = r^2$, $v = 2r + s$,

(2) $w = e^{xy}$, $y = r + s$, $x = rs$.

corrFoncDeuxVar0017

Correction of the exercise 711

(1)

(2) En ce qui concerne les dérivées de w nous avons

$$\begin{aligned} \frac{\partial w}{\partial x} &= ye^{xy} \\ \frac{\partial w}{\partial y} &= xe^{xy} \end{aligned} \tag{96.440}$$

En utilisant la formule de dérivée de fonctions composées, nous trouvons

$$\begin{aligned} \frac{\partial w}{\partial r} &= (2rs + s^2)e^{r^2s + rs^2} \\ \frac{\partial w}{\partial s} &= (r^2 + 2rs)e^{r^2s + rs^2} \end{aligned} \tag{96.441}$$

Exercise 712 exoDerrivePartielle-0004
(INGE1121, 9.15) Calculer $\frac{\partial r}{\partial u}$, $\frac{\partial r}{\partial v}$ et $\frac{\partial r}{\partial t}$ lorsque $r = x\ln(y)$, $x = 3u + vt$ et $y = uvt$. corrDerrivePartielle-0004

Correction of the exercise 712
La formule de dérivation des fonctions composées s'écrit

$$\begin{aligned}\frac{\partial r}{\partial u} &= \frac{\partial r}{\partial x}\big(x(u,v,t),y(u,v,t)\big)\frac{\partial x}{\partial u}(u,v,t)\\ &\quad+\frac{\partial r}{\partial y}\big(x(u,v,t),y(u,v,t)\big)\frac{\partial y}{\partial u}(u,v,t),\end{aligned} \tag{96.442}$$

et le même genre de formules pour $\frac{\partial r}{\partial v}$ et $\frac{\partial r}{\partial t}$.
Les résultats sont

$$\begin{aligned}\frac{\partial r}{\partial u} &= 3\ln(uvt) + 3 + \frac{vt}{u},\\ \frac{\partial r}{\partial v} &= t\ln(uvt) + \frac{3u}{v} + t,\\ \frac{\partial r}{\partial t} &= v\ln(uvt) + \frac{3u}{t} + v\end{aligned} \tag{96.443}$$

Exercise 713 exoDerrivePartielle-0005
(INGE1121, 9.16) Calculer dw/dt lorsque $w = \ln(u+v)$, $u = e^{-t}$, $v = t^3 - t^2$. corrDerrivePartielle-0005

Correction of the exercise 713
Nous utilisons la formule de dérivation des fonctions composées (ici c'est un cas classique qui n'a même pas plusieurs variables). En utilisant la formule $\big(\ln(f)\big)' = f'/f$, nous avons

$$\big(\ln(u+v)\big)' = \frac{u'+v'}{u+v} = \frac{-e^{-t}+3t^2+2t}{t^3+t^2+e^{-t}}. \tag{96.444}$$

Exercise 714 exoFoncDeuxVar0030
(INGE1121, 9.17) Calculer dz/dt si $z = 3x^2 + xy + 2$ avec $x = 3t + \sin(t)$ et $y = t^3 + t$. corrFoncDeuxVar0030

Correction of the exercise 714
Il s'agit d'une dérivation de fonctions composée. Par exemple la dérivée de x^2 par rapport à t est $2xx'$. Nous avons alors

$$\frac{d}{dt}(3x^2+xy) = 6xx' + x'y + xy' = 6x(3+\cos(t)) + \big(3+\cos(t)\big)(t^3+t) + x(3t^2+t). \tag{96.445}$$

Exercise 715 exoFoncDeuxVar0024
(INGE1121, 9.18)Soit $F\colon \mathbb{R}^2 \to \mathbb{R}$ définie par $F(\rho,\theta)$éf$\big(x(\rho,\theta),y(\rho,\theta)\big)$ avec

$$\begin{aligned}f(u,v) &= \ln(u-v)\\ x(\rho,\theta) &= \rho\cos(\theta)\\ y(\rho,\theta) &= \rho\sin(\theta).\end{aligned} \tag{96.446}$$

(1) Calculer $\frac{\partial F}{\partial \rho}$ et $\frac{\partial F}{\partial \theta}$ à la fois par la formule de dérivation des fonctions composées, et en explicitant $F(\rho,\theta)$.
(2) Calculer $\frac{\partial F}{\partial \rho}(1,0)$.

corrFoncDeuxVar0024

Correction of the exercise 715
Si nous notons u et v les variables de f, nous écrivons la formule générale

$$\frac{\partial F}{\partial \rho} = \frac{\partial f}{\partial u}\big(x(\rho,\theta),y(\rho,\theta)\big)\frac{\partial x}{\partial \rho}(\rho,\theta) + = frac\partial f\partial v\big(x(\rho,\theta),y(\rho,\theta)\big)\frac{\partial y}{\partial \rho}(\rho,\theta) \tag{96.447}$$

Nous avons

$$\begin{aligned}\frac{\partial f}{\partial u}(u,v)&=\frac{1}{u-v} & \frac{\partial f}{\partial v}(u,v)&=\frac{-1}{u-v}\\ \frac{\partial x}{\partial \rho}&=\cos(\theta) & \frac{\partial y}{\partial \rho}&=\sin(\theta),\end{aligned}\tag{96.448}$$

et donc

$$\begin{aligned}\frac{\partial F}{\partial \rho}(\rho,\theta)&=\frac{1}{x-y}\cos(\theta)+\frac{-1}{x-y}\sin(\theta)\\ &=\frac{\cos(\theta)}{\rho(\cos\theta-\sin\theta)}-\frac{\sin(\theta)}{\rho(\cos\theta-\sin\theta)}\\ &=\frac{1}{\rho}.\end{aligned}\tag{96.449}$$

EqVQdFdtho

Avec le même genre de calculs ;

$$\frac{\partial F}{\partial \theta}=\frac{\sin(\theta)+\cos(\theta)}{\sin(\theta)-\cos(\theta)}.\tag{96.450}$$

En substituant $\rho=1$ et $\theta=$ dans l'équation (96.449), nous avons

$$\frac{\partial F}{\partial \rho}(0,1)=\frac{1}{1}=1.\tag{96.451}$$

Exercise 716 exoFoncDeuxVar0020

(INGE1121, 9.19) Calculer les dérivées partielles en $(0,0)$ de la fonction $F(x,y)=f\big(g(x,y),y\big)$ sachant que

$$\begin{array}{llll} g(0,0)=1 & \frac{\partial g}{\partial x}(0,0)=-1 & \frac{\partial g}{\partial y}(0,0)=3\\ f(0,0)=2 & \frac{\partial f}{\partial x}(0,0)=-1 & \frac{\partial f}{\partial y}(0,0)=7\\ f(2,0)=3 & \frac{\partial f}{\partial x}(1,0)=-2 & \frac{\partial f}{\partial y}(1,0)=8\\ f(0,2)=4 & \frac{\partial f}{\partial x}(0,1)=-4 & \frac{\partial f}{\partial y}(0,1)=9.\end{array}\tag{96.452}$$

corrFoncDeuxVar0020

Correction of the exercise 716

En notant u et v les variables de f nous avons

$$\frac{\partial F}{\partial x}(0,0)=\frac{\partial f}{\partial u}\big(g(0,0),0\big)\frac{\partial g}{\partial x}(0,0)+\frac{\partial f}{\partial v}\big(g(0,0),0\big)\frac{\partial y}{\partial x}.\tag{96.453}$$

Le second terme disparaît parce que $\partial_x y=0$. En remplaçant chaque morceau par sa valeur, nous trouvons

$$\frac{\partial f}{\partial u}(1,0)\frac{\partial g}{\partial x}(0,0)=-2\cdot(-1)=2.\tag{96.454}$$

Ici, il faut remarquer que les notations de l'énoncé sont faites pour induire en erreur : lorsqu'il est écrit $\frac{\partial f}{\partial x}$, on veut dire « la dérivée de f par rapport à sa première variable ». Ce n'est donc pas le même x que les autres.

En ce qui concerne la dérivée de F par rapport à y, il faut faire le même raisonnement en tenant compte du fait que $\partial_y y=1$. Le résultat est

$$\frac{\partial F}{\partial y}(0,0)=2.\tag{96.455}$$

Exercise 717 exoDerrivePartielle-0006

(INGE1121, 9.20) Si $f(x,y)$ et $y=g(x)$ sont des fonctions de classe C^2, déterminer $w'(x)=\frac{dw}{dx}$ et $w''(x)=\frac{d}{dx}\left(\frac{dw}{dx}\right)(x)$ sachant que w est la fonction définie par $w(x)=f\big(x,g(x)\big)$.

corrDerrivePartielle-0006

Correction of the exercise 717

La première dérivée est standard :

$$w'(x) = \frac{\partial f}{\partial u}\big(x, g(x)\big) + \frac{\partial f}{\partial v}\Big(x, g(x)\Big)g'(x) \tag{96.456}$$

où nous avons implicitement utilisé les faits que $\frac{\partial x}{\partial x} = 1$ et $\frac{\partial g}{\partial x} = g'(x)$ parce que g est une authentique fonction de une seule variable.

Pour calculer $w''(x)$, nous devons calculer la dérivée de tout cela. Allons y petit bout par petit bout. D'abord

$$\left[\frac{\partial f}{\partial u}\Big(x, g(x)\Big)\right]' = \frac{\partial^2 f}{\partial u^2}\Big(x, g(x)\Big) + \frac{\partial^2 f}{\partial v \partial u}\Big(x, g(x)\Big)g'(x). \tag{96.457}$$

Ensuite, pour le second terme nous utilisons la règle de Leibniz (càd $(uv)' = u'v + v'u$) :

$$\begin{aligned}\left[\frac{\partial f}{\partial v}\Big(x, g(x)\Big)g'(x)\right]' = \left[\frac{\partial f}{\partial v}\Big(x, g(x)\Big)\right]' g'(x)\\ + \frac{\partial f}{\partial v}\Big(x, g(x)\Big)g''(x).\end{aligned} \tag{96.458}$$

Enfin,

$$\begin{aligned}\left[\frac{\partial f}{\partial v}\Big(x, g(x)\Big)\right]' &= \frac{\partial}{\partial u}\left(\frac{\partial f}{\partial v}\right)\Big(x, g(x)\Big)\\ &\quad + \frac{\partial}{\partial v}\left(\frac{\partial f}{\partial v}\right)\Big(x, g(x)\Big)g'(x)\\ &= \frac{\partial^2 f}{\partial u \partial v}\Big(x, g(x)\Big) + \frac{\partial^2 f}{\partial v^2}\Big(x, g(x)\Big)g'(x).\end{aligned} \tag{96.459}$$

La réponse est obtenue en recollant tous les morceaux.

96.8.7 Dérivées de fonctions implicites

Exercise 718 exoFoncDeuxVar0004

Dériver les fonctions implicites suivantes

(1) $y^3 + 2y^2 + xy - 4 = 0$

corrFoncDeuxVar0004

Correction of the exercise 718

<+FoncDeuxVar0004+>

Exercise 719 exoFoncDeuxVar0005

Les surfaces $x^2y^2 + 2x + z^3 = 16$ et $3x^2 + y^2 - 2z = 9$ ont le point $(2, 1, 2)$ situé dans leur courbe d'intersection. Quelles sont les équations des plans tangents respectifs aux deux surfaces en ce point ? Quel est l'angle formé par les deux normales ?

corrFoncDeuxVar0005

Correction of the exercise 719

<+FoncDeuxVar0005+>

Exercise 720 exoFoncDeuxVar0006

Soit la surface d'équation $(x + 1)(y + 1) = z + 1$. On demande

(1) l'équation du plan tangent à l'origine

(2) de décrire l'intersection du plan tangent et de la surface,

(3) la nature de cette surface.

corrFoncDeuxVar0006

Correction of the exercise 720

<+FoncDeuxVars0006+>

Exercise 721 exoFoncDeuxVar0007

Écrire les équations de la droite tangente à la courbe

$$\begin{cases} 3x^2y + y^2z = 2 & \text{(96.460a)} \\ 2xz - x^2y = 3 & \text{(96.460b)} \end{cases}$$

au point $(1, -1, 1)$. corrFoncDeuxVar0007

Correction of the exercise 721

<+FoncDeuxVar0007+>

96.8.8 Extrema

Exercise 722 exoFoncDeuxVar0008

Rechercher les points de la surface $z^2 = x^2 + y^2 + xy + x + y + 1$ les plus rapprochés de l'origine. corrFoncDeuxVar0008

Correction of the exercise 722

<+FoncDeuxVar0008+>

Exercise 723 exoFoncDeuxVar0009

Partager le nombre positif a en trois parties telles que leur produit soit maximum. corrFoncDeuxVar0009

Correction of the exercise 723

<+FoncDeuxVar0009+>

Exercise 724 exoFoncDeuxVar0029

(INGE1121, 10.1) Rechercher les extrémums des fonctions suivantes

(1) $f(x, y) = 2 - \sqrt{x^2 + y^2}$

(2) $f(x, y) = x^3 + 3xy^2 - 15x - 12y$

(3) $f(x, y) = \frac{x^3}{3} + \frac{4y^3}{3} - x^2 - 3x - 4y - 3$

corrFoncDeuxVar0029

Correction of the exercise 724

Les corrigés sont créés par le script Sage `exo101.sage`

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_1_A():
   var('x,y')
   f(x,y)=2-sqrt(x**2+y**2)
   print outilsINGE.Extrema(f)
def exercise_10_1_B():
   var('x,y')
   f(x,y)=x**3+3*x*y**2-15*x-12*y
   print outilsINGE.Extrema(f)
def exercise_10_1_C():
   var('x,y')
   f(x,y)=x**3/3+4*y**3/3-x**2-3*x-4*y-3
   print outilsINGE.Extrema(f)
```

Des réponses :

(1)
```
The function :
(x, y) |--> -sqrt(x^2 + y^2) + 2
Derivative x and y :
(x, y) |--> -x/sqrt(x^2 + y^2)
(x, y) |--> -y/sqrt(x^2 + y^2)
Hessian matrix :
[-sqrt(x^2 + y^2)*y^2/(x^4 + 2*x^2*y^2 + y^4)                          x*y/(x^2 + y^2)^(3/2)
[                        x*y/(x^2 + y^2)^(3/2) -sqrt(x^2 + y^2)*x^2/(x^4 + 2*x^2*y^2 + y^4
Critical points :
(0, 0)
At (0,0), the Hessian is
   power::eval(): division by zero
```

Ici nous voyons que Sage a du mal à calculer la matrice hessienne en $(0,0)$. En effet, nous tombons sur une division par zéro. Pour résoudre l'exercice, il faut se rendre compte que la fonction $(x,y) \mapsto \sqrt{x^2+y^2}$ est toujours positive et est nulle seulement au point $(0,0)$. Donc f est toujours plus petite ou égale à deux tandis que $f(0,0) = 2$. Le point est donc un maximum global.

(2)
```
The function :
(x, y) |--> x^3 + 3*x*y^2 - 15*x - 12*y
Derivative x and y :
(x, y) |--> 3*x^2 + 3*y^2 - 15
(x, y) |--> 6*x*y - 12
Hessian matrix :
[6*x 6*y]
[6*y 6*x]
Critical points :
(2, 1)
(1, 2)
(-1, -2)
(-2, -1)
At (2,1), the Hessian is
   (12, 6)
   (6, 12)
   Primary principal minors are [108, 12]
   Hessian positive defined
   local minimum
At (1,2), the Hessian is
   (6, 12)
   (12, 6)
   Primary principal minors are [-108, 6]
   Undefinite Hessian
   «selle» point
At (-1,-2), the Hessian is
   (-6, -12)
   (-12, -6)
   Primary principal minors are [-108, -6]
   Undefinite Hessian
   «selle» point
At (-2,-1), the Hessian is
   (-12, -6)
   (-6, -12)
   Primary principal minors are [108, -12]
   Hessian negative defined
```

```
   local maximum
```

Petite note sur la façon dont on trouve les points critiques. Le système est

$$\begin{cases} 3x^2 + 3y^2 - 15 = 0 & \text{(96.461a)} \\ 6xy - 12 = 0. & \text{(96.461b)} \end{cases}$$

De la seconde équation, nous isolons x : $x = 2/y$. Précisons qu'il n'y a pas de solutions avec $y = 0$. En remettant le x trouvé en fonction de y dans la première équation nous trouvons

$$\frac{12}{y^2} + 3y^2 - 15 = 0, \tag{96.462}$$

ce qui revient à l'équation bicarrée

$$3y^4 - 15y^2 + 12 = 0. \tag{96.463}$$

En posant $u = y^2$ nous avons $u = 4$ ou $u = 1$ (résolution d'une équation du second degré pour u). Nous avons alors les quatre possibilités $y = \pm 2$ et $y = \pm 1$. Pour chacune de ces possibilités, la formule $x = 2/y$ fournit le x correspondant.

(3)
```
The function :
(x, y) |--> 1/3*x^3 - x^2 + 4/3*y^3 - 3*x - 4*y - 3
Derivative x and y :
(x, y) |--> x^2 - 2*x - 3
(x, y) |--> 4*y^2 - 4
Hessian matrix :
[2*x - 2       0]
[      0     8*y]
Critical points :
(3, 1)
(-1, 1)
(3, -1)
(-1, -1)
At (3,1), the Hessian is
   (4, 0)
   (0, 8)
   Primary principal minors are [32, 4]
   Hessian positive defined
   local minimum
At (-1,1), the Hessian is
   (-4, 0)
   (0, 8)
   Primary principal minors are [-32, -4]
   Undefinite Hessian
   «selle» point
At (3,-1), the Hessian is
   (4, 0)
   (0, -8)
   Primary principal minors are [-32, 4]
   Undefinite Hessian
   «selle» point
At (-1,-1), the Hessian is
   (-4, 0)
   (0, -8)
   Primary principal minors are [32, -4]
   Hessian negative defined
   local maximum
```

Exercise 725 exoFoncDeuxVar0028

Déterminer les valeurs extrêmes et les points de selle des fonctions suivantes.

(1) $f(x,y)=x^2+4x+y^2-2y$.
(2) $f(x,y)=e^{x^2+xy}$.
(3) $f(x,y)=e^x\sin(y)$.

corrFoncDeuxVar0028

Correction of the exercise 725

Certains corrigés de cet exercice ont étés réalisés par Sage. Le script utilisé est `exo103.sage`

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_3_A():
   var('x,y')
   f(x,y)=x**2+4*x+y**2-2*y
   print outilsINGE.Extrema(f)

def exercise_10_3_H():
   var('x,y')
   f(x,y)=exp(x**2+x*y)
   print outilsINGE.Extrema(f)

def exercise_10_3_Q():
   var('x,y')
   f(x,y)=exp(x)*sin(y)
   print outilsINGE.Extrema(f)
```

Des réponses :

(1)
```
The function :
(x, y) |--> x^2 + y^2 + 4*x - 2*y
Derivative x and y :
(x, y) |--> 2*x + 4
(x, y) |--> 2*y - 2
Hessian matrix :
[2 0]
[0 2]
Critical points :
(-2, 1)
At (-2,1), the Hessian is
   (2, 0)
   (0, 2)
   Primary principal minors are [4, 2]
   Hessian positive defined
   local minimum
```

(2)
(3)
(4)
(5)
(6)
(7)

(8)
```
The function :
(x, y) |--> e^(x^2 + x*y)
Derivative x and y :
(x, y) |--> (2*x*e^(x^2) + y*e^(x^2))*e^(x*y)
(x, y) |--> x*e^(x^2 + x*y)
Hessian matrix :
[(4*x*y*e^(x^2) + y^2*e^(x^2) + 2*(2*x^2 + 1)*e^(x^2))*e^(x*y)                    (x*y*e
[                  (x*y*e^(x^2) + (2*x^2 + 1)*e^(x^2))*e^(x*y)
Critical points :
(0, 0)
At (0,0), the Hessian is
   (2, 1)
   (1, 0)
   Primary principal minors are [-1, 2]
   Undefinite Hessian
   «selle» point
```

(9)

(10)

(11)

(12)

(13)

(14)

(15)

(16)

(17)
```
The function :
(x, y) |--> e^x*sin(y)
Derivative x and y :
(x, y) |--> e^x*sin(y)
(x, y) |--> e^x*cos(y)
Hessian matrix :
[ e^x*sin(y)  e^x*cos(y)]
[ e^x*cos(y) -e^x*sin(y)]
Critical points :
   I'm not able to solve these equations.
   I'm not able to solve these equations.
At (   ,I), the Hessian is
   I'm not able to solve these equations.
At (   ,I), the Hessian is
   I'm not able to solve these equations.
```

Ici, Sage n'est pas capable de résoudre les équations qui annulent le jacobien. Les équations à résoudre sont pourtant faciles :

$$\begin{cases} e^x \cos(y) = 0 & (96.464a) \\ e^x \sin(y) = 0 & (96.464b) \end{cases}$$

Étant donné que l'exponentielle ne s'annule jamais, il faudrait avoir en même temps $\cos(y) = 0$ et $\sin(y) = 0$, ce qui est impossible. La fonction n'a donc aucun extrémum local.

(18)

(19)

(20)

Exercise 726 exoDerrivePartielle-0007

Considérons la fonction

$$f(x,y)=xy^2e^{-(x^2+y^2)/4}. \tag{96.465}$$

(1) Montrer qu'il y a une infinité de points critiques.

(2) Déterminer leur nature.

corrDerrivePartielle-0007

Correction of the exercise 726

Voici la fonction Sage qui fournit les informations :

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_4():
  var('x,y')
  f(x,y)=x*y**2*exp(-(x**2+y**2)/4)
  print outilsINGE.Extrema(f)
```

tex/sage/exo104.sage

La sortie est

```
The function :
(x, y) |--> x*y^2*e^(-1/4*x^2 - 1/4*y^2)
Derivative x and y :
(x, y) |--> -1/2*(x^2 - 2)*y^2*e^(-1/4*x^2 - 1/4*y^2)
(x, y) |--> -1/2*(x*y^3 - 4*x*y)*e^(-1/4*x^2 - 1/4*y^2)
Hessian matrix :
[                1/4*(x^3 - 6*x)*y^2*e^(-1/4*x^2 - 1/4*y^2) 1/4*((x^2 - 2)*y^3 - 4*(x^2 - 2)*y
[1/4*((x^2 - 2)*y^3 - 4*(x^2 - 2)*y)*e^(-1/4*x^2 - 1/4*y^2)         1/4*(x*y^4 - 10*x*y^2 + 8*x
Critical points :
(r17, 0)
(-sqrt(2), -2)
(sqrt(2), -2)
(-sqrt(2), 2)
(sqrt(2), 2)
At (r17,0), the Hessian is
   (0, 0)
   (0, 2*r17*e^(-1/4*r17^2))
   Primary principal minors are [0, 0]
   Hessian positive semidéfinite
   I don't conclude
   Hessian negative semidefinite
   I don't conclude
At (-sqrt(2),-2), the Hessian is
   (4*sqrt(2)*e^(-3/2), 0)
   (0, 4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), 4*sqrt(2)*e^(-3/2)]
   Hessian positive defined
   local minimum
At (sqrt(2),-2), the Hessian is
   (-4*sqrt(2)*e^(-3/2), 0)
   (0, -4*sqrt(2)*e^(-3/2))
```

```
   Primary principal minors are [32*e^(-3), -4*sqrt(2)*e^(-3/2)]
   Hessian negative defined
   local maximum
At (-sqrt(2),2), the Hessian is
   (4*sqrt(2)*e^(-3/2), 0)
   (0, 4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), 4*sqrt(2)*e^(-3/2)]
   Hessian positive defined
   local minimum
At (sqrt(2),2), the Hessian is
   (-4*sqrt(2)*e^(-3/2), 0)
   (0, -4*sqrt(2)*e^(-3/2))
   Primary principal minors are [32*e^(-3), -4*sqrt(2)*e^(-3/2)]
   Hessian negative defined
   local maximum
```

Notez la présence de `r1` comme paramètres dans les solutions. Tous les points avec $y = 0$ sont des points critiques. Cependant, Sage[6] ne parvient pas à conclure la nature de ces points $(x, 0)$.

Notons que le nombre $f(x, y)$ a toujours le signe de x parce que y^2 et l'exponentielle sont positives. Toujours ? En tout cas lorsque $x \neq 0$. Prenons un point $(a, 0)$ avec $a > 0$. Dans un voisinage de ce point, nous avons $f(x, y) > 0$ parce que si $a > 0$, alors $x > 0$ dans un voisinage de a. Le point $(a, 0)$ est un minimum local parce que $0 = f(a, 0) \leqslant f(x, y)$ pour tout (x, y) dans un voisinage de $(a, 0)$.

De la même façon, les points $(a, 0)$ avec $a < 0$ sont des maximums locaux parce que dans un voisinage, la fonction est négative.

Le point $(0, 0)$ n'est ni maximum ni minimum local. C'est un point de selle.

Exercise 727 exoMaximisation-0002

Soit la fonction

$$g(x, y) = (x^4 + 1)(2 + \cos(y)). \tag{96.466}$$

Prouver que le point $(0, 0)$ est un point de selle. corrMaximisation-0002

Correction of the exercise 727

Le calcul de la matrice hessienne donne

$$H(x, y) = \begin{pmatrix} 12x^2(\cos(y) + 2) & -4x^3 \sin(y) \\ -4x^3 \sin(y) & -(x^4 + 1)\cos(y) \end{pmatrix}. \tag{96.467}$$

En $(0, 0)$, cette matrice est semi-définie positive, et nous ne pouvons rien conclure. Il faut donc trouver autre chose.

Afin de prouver qu'un point est un point de selle, la méthode la plus directe est de trouver, dans chaque voisinage, un point où la fonction est plus grande et un point où la fonction est plus petite.

Dans le cas qui nous occupe, $g(0, 0) = 3$. Est-ce qu'on peut trouver un point (x, y) très proche de $(0, 0)$ tel que $g(x, y) > 3$? Et un autre tel que $g(x, y) < 3$? Supposons que ϵ est petit. Alors

$$g(\epsilon, 0) = 3(\epsilon^4 + 1) > 3, \tag{96.468}$$

tandis que

$$g(0, \epsilon) = 2 + \cos(\epsilon) < 3. \tag{96.469}$$

Exercise 728 exoDerrivePartielle-0008

(INGE1121, 10.5) Partager le nombre 1000 en trois nombres dont le produit soit maximum. corrDerivePartielle-0008

6. ou, plus précisément, le programme que j'ai écrit avec Sage.

Correction of the exercise 728

Juste pour rester général, nous allons noter $1000 = A$. Ainsi nous verrons si le résultat dépend du nombre à partager. Si x, y et z sont les nombres, nous avons

$$z = A - x - y \tag{96.470}$$

et nous devons maximiser $f(x, y) = xyz = xy(A - x - y)$.

Cela se fait avec les méthodes habituelles et la fonction

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_5():
   var('x,y,A')
   assume(A>0)
   f(x,y)=x*y*(A-x-y)
   print outilsINGE.Extrema(f)
```

Notez la présence de la ligne

```
assume(A>0)
```

qui dit à Sage que A est un nombre positif (ici c'est 1000).

La sortie est :

```
The function :
(x, y) |--> (A - x - y)*x*y
Derivative x and y :
(x, y) |--> A*y - 2*x*y - y^2
(x, y) |--> A*x - x^2 - 2*x*y
Hessian matrix :
[         -2*y A - 2*x - 2*y]
[A - 2*x - 2*y          -2*x]
Critical points :
(0, 0)
(0, A)
(A, 0)
(1/3*A, 1/3*A)
At (0,0), the Hessian is
   (0, A)
   (A, 0)
   Primary principal minors are [-A^2, 0]
   Hessian negative semidefinite
   I don't conclude
At (0,A), the Hessian is
   (-2*A, -A)
   (-A, 0)
   Primary principal minors are [-A^2, -2*A]
   Undefinite Hessian
   «selle» point
At (A,0), the Hessian is
   (0, -A)
   (-A, -2*A)
   Primary principal minors are [-A^2, 0]
```

```
   Hessian negative semidefinite
   I don't conclude
At (1/3*A,1/3*A), the Hessian is
   (-2/3*A, -1/3*A)
   (-1/3*A, -2/3*A)
   Primary principal minors are [1/3*A^2, -2/3*A]
   Hessian negative defined
   local maximum
```

Il faut faire un peu le ménage dans toutes les solutions trouvées. Les points avec x ou y qui sont nuls sont à rejeter parce qu'ils ne vont certainement pas maximiser le produit xyz. Reste la solution $x = A/3$, $y = A/3$, qui est un maximum local.

La solution du problème est donc $(x, y, z) = (\frac{A}{3}, \frac{A}{3}, \frac{A}{3})$. Il faut couper en trois parts égales, et cela ne dépend pas du nombre choisi.

Exercise 729 exoDerrivePartielle-0009

(INGE1121, 10.6) Quelles sont les dimensions d'une boîte parallélépipédique rectangulaire de volume fixé V dont la surface soit la plus petite possible. corrDerrivePartielle-0009

Correction of the exercise 729

Si x, y et z sont les dimensions de la boîte, la contrainte est $xyz = V$. La surface est donnée par

$$S(x, y, z) = 2xy + 2xz + 2yz, \tag{96.471}$$

et en posant $z = V/xy$, nous devons minimiser la fonction

$$f(x, y) = xy + \frac{V}{y} + \frac{V}{x}. \tag{96.472}$$

La maximisation de cette fonction se fait comme d'habitude. Le code Sage est

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_6():
   var('x,y,V')
   assume(V>0)
   f(x,y)=x*y+V/y+V/x
   print outilsINGE.Extrema(f)
```

La sortie est :

```
The function :
(x, y) |--> x*y + V/x + V/y
Derivative x and y :
(x, y) |--> (x^2*y - V)/x^2
(x, y) |--> (x*y^2 - V)/y^2
Hessian matrix :
[2*V/x^3       1]
[      1 2*V/y^3]
Critical points :
(-2*V^(1/3)/(I*sqrt(3) + 1), 1/2*(I*sqrt(3) - 1)*V^(1/3))
(2*V^(1/3)/(I*sqrt(3) - 1), 1/2*(-I*sqrt(3) - 1)*V^(1/3))
(V^(1/3), V^(1/3))
At (-2*V^(1/3)/(I*sqrt(3) + 1),1/2*(I*sqrt(3) - 1)*V^(1/3)), the Hessian is
   (-1/4*(I*sqrt(3) + 1)^3, 1)
```

```
   (1, 16/(I*sqrt(3) - 1)^3)
   Primary principal minors are [-4*(I*sqrt(3) + 1)^3/(I*sqrt(3) - 1)^3 - 1, -1/4*(I*sqrt(3) +
   Hessian positive defined
   local minimum
   Hessian negative defined
   local maximum
At (2*V^(1/3)/(I*sqrt(3) - 1),1/2*(-I*sqrt(3) - 1)*V^(1/3)), the Hessian is
   (1/4*(I*sqrt(3) - 1)^3, 1)
   (1, 16/(-I*sqrt(3) - 1)^3)
   Primary principal minors are [4*(I*sqrt(3) - 1)^3/(-I*sqrt(3) - 1)^3 - 1, 1/4*(I*sqrt(3) -
   Hessian positive defined
   local minimum
   Hessian negative defined
   local maximum
At (V^(1/3),V^(1/3)), the Hessian is
   (2, 1)
   (1, 2)
   Primary principal minors are [3, 2]
   Hessian positive defined
   local minimum
```

Notez la présence de `I` qui est la façon dont Sage note le nombre imaginaire $\sqrt{-1}$. Les solutions complexes étant à rejeter, nous ne devons garder que la dernière : $x = y = z = \sqrt[3]{V}$

Exercise 730 exoDerrivePartielle-0010

(INGE1121, 10.9) Notons $f(x, y, z)$ le carré de la distance du point $(0, 0, 2)$ à un point de la surface $z = xy$. Déterminer les extrémums locaux. corrDerrivePartielle-0010

Correction of the exercise 730

En utilisant le théorème de Pythagore, la fonction f s'écrit

$$f(x, y, z) = x^2 + y^2 + (z - 2)^2, \tag{96.473}$$

mais comme nous ne nous intéressons qu'aux points tels que $z = xy$, nous regardons la fonction

$$f(x, y) = x^2 + y^2 + (xy - 2)^2. \tag{96.474}$$

Nous devons étudier les extrémums de cette fonction.

Le code Sage qui fait les calculs est :

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_9():
   var('x,y')
   f(x,y)=x**2+y**2+(x*y-2)**2
   print outilsINGE.Extrema(f)
```

La sortie :

```
The function :
(x, y) |--> (x*y - 2)^2 + x^2 + y^2
Derivative x and y :
(x, y) |--> 2*x*y^2 + 2*x - 4*y
(x, y) |--> 2*(x^2 + 1)*y - 4*x
```

```
Hessian matrix :
[2*y^2 + 2 4*x*y - 4]
[4*x*y - 4 2*x^2 + 2]
Critical points :
(I*sqrt(3), -I*sqrt(3))
(-I*sqrt(3), I*sqrt(3))
(-1, -1)
(1, 1)
(0, 0)
At (I*sqrt(3),-I*sqrt(3)), the Hessian is
   (-4, 8)
   (8, -4)
   Primary principal minors are [-48, -4]
   Undefinite Hessian
   «selle» point
At (-I*sqrt(3),I*sqrt(3)), the Hessian is
   (-4, 8)
   (8, -4)
   Primary principal minors are [-48, -4]
   Undefinite Hessian
   «selle» point
At (-1,-1), the Hessian is
   (4, 0)
   (0, 4)
   Primary principal minors are [16, 4]
   Hessian positive defined
   local minimum
At (1,1), the Hessian is
   (4, 0)
   (0, 4)
   Primary principal minors are [16, 4]
   Hessian positive defined
   local minimum
At (0,0), the Hessian is
   (2, -4)
   (-4, 2)
   Primary principal minors are [-12, 2]
   Undefinite Hessian
   «selle» point
```

Il faut rejeter les solutions qui contiennent le nombre imaginaire `I`. Pour le reste, la solution est là.

Le point délicat est de résoudre les équations

$$\begin{cases} (y^2+1)x - 2y = 0 & \text{(96.475a)} \\ (x^2+1)y - 2x = 0. & \text{(96.475b)} \end{cases}$$

Nous isolons x dans la première équation :

$$x = \frac{2y}{(y^2+1)}, \tag{96.476}$$

et nous le remettons dans la seconde :

$$\left(\frac{4y^2}{(y^2+1)^2} + 1\right) y - \frac{4y}{y^2+1} = 0. \tag{96.477}$$

La première solution qu'on voit est $y = 0$ qui implique $x = 0$. Le point $(0, 0)$ est donc un point critique.

Si $y \neq 0$ nous pouvons simplifier par y et mettre au même dénominateur $(y^2 + 1)^2$:

$$4y^2 + (y^2 + 1)^2 - 4(y^2 + 1) = 0. \tag{96.478}$$

En simplifiant nous trouvons une équation bicarrée : $y^4 + 2y^2 - 3 = 0$. Nous posons $u = y^2$ et nous trouvons l'équation $u^2 + 2u - 3 = 0$. En ce qui concerne u nous avons les solutions $u = 1$ et $u = -3$. La solution $u = -3$ est à rejeter parce que $u = y^2$ (c'est sans doute ici qu'arrivent les solutions complexes trouvées par Sage). Reste donc $y = \pm 1$.

Exercise 731 exoMaximisation-0000

(INGE1121, 10.11) Soit $f(x, y) = x^3 - 3xy^2$.

(1) Montrer que le seul point critique est $(0, 0)$ et déterminer la nature de la matrice hessienne en ce point.

(2) Déterminer la nature du point critique en considérant le comportement de la fonction sur les droites passant par l'origine.

corrMaximisation-0000

Correction of the exercise 731

Cette fois, les calculs ne sont pas compliqués. Le code Sage est :

```
# -*- coding: utf8 -*-

import outilsINGE

def exercise_10_11():
   var('x,y')
   f(x,y)=x**3-3*x*y**2
   print outilsINGE.Extrema(f)
```

La sortie est :

```
The function :
(x, y) |--> x^3 - 3*x*y^2
Derivative x and y :
(x, y) |--> 3*x^2 - 3*y^2
(x, y) |--> -6*x*y
Hessian matrix :
[ 6*x -6*y]
[-6*y -6*x]
Critical points :
(0, 0)
At (0,0), the Hessian is
   (0, 0)
   (0, 0)
   Primary principal minors are [0, 0]
   Hessian positive semidéfinite
   I don't conclude
   Hessian negative semidefinite
   I don't conclude
```

Il n'est pas capable de conclure, mais les informations sont intéressantes quand même et répondent à la première question.

D'abord, notons que $f(0, 0) = 0$.

Si nous prenons la droite [7] (t, kt) et calculons f dessus :

$$f(t, kt) = t^3 - 3kt^3 = t^3(1 - 3k). \tag{96.479}$$

Cela change de signe quand t change de signe. Dans tout voisinage de $(0, 0)$, il y a donc des points avec $f > 0$ et des points avec $f < 0$. Nous concluons que le point $(0, 0)$ est un point de selle.

Lorsque nous demandons d'étudier une fonction, nous demandons les éléments suivants : domaine de définition, croissance, extrémums, points d'inflexion, asymptote et dessiner le graphe.

Exercise 732 exo1

Calculer

(1) $f'(2)$ si $f(x) = \ln(x^2 + 1)$
(2) $g'(e)$ si $g(x) = x^x$
(3) $h'(\frac{\pi}{6})$ si $h(x) = xe^{\sin(x)}$
(4) $\ell'(\frac{\pi}{3})$ si $\ell(x) = \frac{x}{5+\cos(x)}$.

corr1

Correction of the exercise 732

(1) $f(x) = \ln(x^2 + 1)$, utiliser la formule $(\ln(u))' = \frac{u'}{u}$,

$$f'(x) = \frac{2x}{x^2 + 1}, \tag{96.480}$$

$f'(2) = \frac{4}{5}$.

(2) $g(x) = x^x$, passer au logarithme :

$$x^x = e^{\ln(x^x)} = e^{x \ln(x)}. \tag{96.481}$$

À partir de là, nous utilisons la formule $(e^u)' = u'e^u$ avec $u(x) = x \ln(x)$. Ce que nous trouvons est

$$g'(x) = x^x(\ln(x) + 1), \tag{96.482}$$

$g'(e) = 2e^e$.

(3) $h(x) = xe^{\sin(x)}$, cela est un produit de x par $e^{\sin(x)}$. Chacun des deux facteurs est relativement facile à dériver. Ce que nous trouvons à la fin est

$$h'(x) = x \cos(x)e^{\sin(x)} + e^{\sin(x)}, \tag{96.483}$$

$h'(\frac{\pi}{6}) = \frac{\pi\sqrt{3}}{12}\sqrt{e} + \sqrt{e}$.

(4) $\ell(x) = \frac{x}{5+\cos(x)}$, c'est une fraction sans surprises :

$$\ell'(x) = \frac{x \sin(x)}{\big(\cos(x) + 5\big)^2} + \frac{1}{\cos(x) + 5}, \tag{96.484}$$

$\ell'(\frac{\pi}{3}) = \frac{2\pi\sqrt{3}}{363} + \frac{2}{11}$.

Exercise 733 exo2

Une population de bactéries croît selon la loi $p(t) = a2^{bt}$. Au temps $t = 2$, on mesure $p(2) = 640$, et au temps $t = 3$, on mesure $p(3) = 5120$.

(1) Calculer a et b.
(2) En combien de temps la population est-elle doublée ?

7. Notez que ce chemin ne donne pas toutes les droites lorsque k varie : il manque la droite verticale $(0, t)$.

corr2

Correction of the exercise 733

Nous avons $p(t) = a2^{bt}$, et nous mesurons

$$p(2) = a2^{2b} = 640 \quad \text{et} \quad p(3) = a2^{3b} = 5120. \tag{96.485}$$

Ce qu'il nous faut, c'est trouver les nombres a et b qui satisfont les équations

$$\begin{cases} a2^{2b} = 640 & (96.486a) \\ a2^{3b} = 5120. & (96.486b) \end{cases}$$

Afin de résoudre ce système, la subtilité est de prendre le rapport des deux équations pour éliminer le a :

$$\frac{p(2)}{p(3)} = \frac{a2^{2b}}{a2^{3b}} = \frac{640}{5120} = \frac{1}{8}. \tag{96.487}$$

Cela fait $2^{2b-3b} = 1/8$, ou encore $2^{-b} = 1/8$, ce qui revient à $2^b = 8$. Nous trouvons $b = 3$. L'équation $a2^{2b} = 64$ devient alors $a2^6 = 640$, et donc $a = 10$. L'équation pour $p(t)$ est maintenant

$$p(t) = 10 \cdot 2^{3t}, \tag{96.488}$$

et en particulier,

$$p(t + T) = 2^{3T} p(t), \tag{96.489}$$

et donc on résous l'équation $2^{3T} = 2$ pour trouver le temps de doublement de population : $T = \frac{1}{3}$.

Exercise 734 exo3

On désire construire une boîte cylindrique, avec couvercle, de volume V donné. Quelles sont les dimensions de la boîte la plus économique (c'est-à-dire celle dont la surface totale est minimum) ?

corr3

Correction of the exercise 734

Les paramètres de la boîte sont le rayon R et la hauteur h qui sont contraints par $V = \pi R^2 h$, et on veut minimiser

$$S(R, h) = 2\pi Rh + 2\pi R^2. \tag{96.490}$$

Cela ressemble à une fonction à deux variables (R et h), mais ce n'en est pas une. En effet, dès que R est choisi, h est automatiquement choisit à cause de la contrainte $V = \pi R^2 h$. Nous pouvons donc remplacer h par $V/\pi R^2$ dans la fonction $S(R, h)$, et nous trouvons la surface en fonction de R tout seul :

$$S(R) = \frac{2V}{R} + 2\pi R^2. \tag{96.491}$$

Pour minimiser, nous cherchons l'annulation de la dérivée

$$S'(R) = -\frac{2V}{R^2} + 4\pi R. \tag{96.492}$$

Nous trouvons $S'(R) = 0$ pour

$$R = \left(\frac{V}{2\pi}\right)^{1/3}, \tag{96.493}$$

qui est la seule solution réelle. Nous croyons que cela est un minimum et non un maximum parce que $S(R) \to \infty$ pour $R \to \pm\infty$. Nous retrouvons la hauteur que l'on cherche en remettant cette valeur de R dans l'équation qui lie R et h :

$$h = \frac{V}{\pi R^2} = 2^{-1/3} \left(\frac{V}{\pi}\right)^{4/3}. \tag{96.494}$$

Exercise 735 exo4

Étudier la fonction définie par $f(x) = \ln(x^2 + \frac{5}{9})$.

corr4

Correction of the exercise 735

La fonction $\ln(x^2+\frac{5}{9})$ est définie tant que $x^2+\frac{5}{9}>0$, c'est-à-dire toujours. La dérivée est

$$f'(x)=\frac{2x}{x^2+\frac{5}{9}},\tag{96.495}$$

qui a le signe de x. Donc

(1) $f(x)$ est croissante pour $x>0$,

(2) $f(x)$ est décroissante pour $x<0$.

Nous calculons la dérivée seconde :

$$f''(x)=-\frac{162x^2-9}{81x^4+9x^2+25}.\tag{96.496}$$

Le dénominateur est toujours positif, donc cette dérivée a le signe de $9-162x^2$. Il y a un point d'inflexion en

$$x=\pm\frac{1}{3\sqrt{2}}.\tag{96.497}$$

Il y a un extrémum en $f'(x)=0$, c'est-à-dire en $x=0$, et c'est un minimum (regarder la croissance pour le voir).

Il n'y a aucune asymptote parce que $f(x)\to\infty$ pour $x\to\pm\infty$, et

$$\lim_{x\to\infty}\frac{f(x)}{x}=0,\tag{96.498}$$

voir page 65 du cours.

Exercise 736 exo5

Étudier la fonction $f(x)=e^{-x^2}$. corr5

Correction of the exercise 736

La fonction est $g(x)=e^{-x^2}$, dont le domaine est $\mathbb{R}$. La dérivée est

$$g'(x)=-2xe^{-x^2}.\tag{96.499}$$

Étant donné que l'exponentielle est toujours positive, cela a le signe opposé à x, donc

(1) Croissante pour $x<0$,

(2) décroissante pour $x>0$,

et maximum pour $x=0$.

La dérivée seconde est

$$g''(x)=(4x^2-2)e^{-x^2},\tag{96.500}$$

qui s'annule quand $x=\pm\sqrt{\frac{1}{2}}$, ce sont les points d'inflexion. Il y a asymptote horizontale $y=0$ en $\pm\infty$.

Exercise 737 exoTP40001

On fait tourner autour de l'axe Ox la région du plan comprise entre la courbe $y=e^x\sqrt{2x+1}$ et les deux axes. Que vaut le volume du solide obtenu ? corrTP40001

Correction of the exercise 737

D'abord, la fonction ne commence à exister qu'à partir de $x=-1/2$, donc il faut calculer le volume entre $-1/2$ et 0. En utilisant la formule $V=\pi\int_a^b f(x)^2dx$, nous devons calculer

$$V=\int_{-1/2}^{0}e^{2x}(2x+1)dx.\tag{96.501}$$

Cela se fait par partie. Nous posons

$$\begin{aligned} u &= 2x+1 \qquad dv = e^{2x} \\ du &= 2 \qquad\quad v = \frac{1}{2}e^{2x}. \end{aligned} \tag{96.502}$$

Donc l'intégrale à calculer est

$$I = \int e^{2x}(2x+1) = \frac{1}{2}e^{2x}(2x+1) - \int e^{2x} = xe^{2x}, \tag{96.503}$$

et le volume est

$$V = \pi[xe^{2x}]_{-1/2}^{0} = \frac{\pi}{2e}. \tag{96.504}$$

Exercise 738 exoTP40002

Déterminer la solution de l'équation différentielle

$$y' + 2y\tan(x) = \cos(x) \tag{96.505}$$

qui s'annule en $x = \pi$.

corrTP40002

Correction of the exercise 738

L'équation homogène est

$$y'_H + 2y_H\tan(x) = 0, \tag{96.506}$$

qui donne

$$\frac{y'_H}{y_H} = -2\tan(x). \tag{96.507}$$

En intégrant des deux côtés,

$$y_H(x) = K\cos^2(x). \tag{96.508}$$

Afin de trouver la solution générale de l'équation non homogène, nous posons

$$y(x) = K(x)\cos^2(x), \tag{96.509}$$

dont la dérivée est

$$y' = K'\cos^2(x) + -2K\sin(x)\cos(x). \tag{96.510}$$

En remettant le tout dans l'équation de départ, nous trouvons l'équation suivante pour K :

$$K'(x)\frac{1}{\cos(x)}. \tag{96.511}$$

Cette intégrale se règle en posant $u = \sin(x)$, donc $\cos(x) = \sqrt{1-u^2}$, donc

$$\int \frac{1}{\sqrt{1-u^2}}\frac{du}{\cos(x)} = \int \frac{du}{\sqrt{1-u^2}\sqrt{1-u^2}}. \tag{96.512}$$

Nous avons donc

$$K(x) = \frac{1}{2}\ln\left(\frac{\sin(x)+1}{\sin(x)-1}\right) + C. \tag{96.513}$$

Au final, la solution générale de l'équation différentielle est

$$y(x) = \frac{\cos^2(x)}{2}\ln\left|\frac{\sin(x)+1}{\sin(x)-1}\right| + C\cos^2(x). \tag{96.514}$$

I s'agit maintenant de trouver la solution qui vaut zéro lorsque $x = \pi$. Pour cela, nous calculons $y(\pi)$. Ce sont les valeurs absolues du logarithme qui nous sauvent la vie, et on trouve qu'avec $C = 0$, le tout vaut zéro.

Exercise 739 exoTP40003

Calculer la longueur d'arc de la courbe d'équations

$$\begin{cases} x = e^t \cos(t) \\ y = e^t \sin(t), \end{cases} \tag{96.515}$$

$0 \leqslant t \leqslant 4\pi$. Dessiner cet arc de courbe. corrTP40003

Correction of the exercise 739

Nous avons les fonctions

$$\begin{aligned} x(t) &= e^t \cos(t) \\ y(t) &= e^t \sin(t), \end{aligned} \tag{96.516}$$

et donc les dérivées

$$\begin{aligned} x'(t) &= e^t \cos(t) - e^t \sin(t) \\ y'(t) &= e^t \sin(t) + e^t \cos(t). \end{aligned} \tag{96.517}$$

Maintenant nous allons utiliser la formule $L = \int_{t_0}^{t_1} \sqrt{x'(t)^2 + y'(t)^2} dt$. En développant les carrés et en regroupant, nous trouvons

$$L = \sqrt{2} \int_0^{4\pi} e^t = \sqrt{2}(e^{4\pi} - 1). \tag{96.518}$$

Reste le délicat point de dessiner la fonction.

Exercise 740 exoTP40004

On donne la surface d'équation $z = 100 - \sqrt{x^2 + y^2}$. Calculer l'aire de la portion de cette surface située à l'intérieur du cylindre d'équation $x^2 + y^2 = 4x$. corrTP40004

Correction of the exercise 740

<+TP40004+>

Exercise 741 exoTP40005

Une équipe de biologistes et de géographes part en expédition vers une île lointaine. Leur navire devra parcourir d kilomètres. Parmi les coûts du voyage, il y a ceux du carburant et ceux du personnel. Le coût par heure en carburant est directement proportionnel au carré de la vitesse, de la forme kv^2. Quant au coût pas heure en personnel, il est évidemment indépendant de la vitesse, soit p le cout par heure en personnel.

Calculer, en fonction des constantes k et p, la vitesse v en km/h à laquelle il faut naviguer pour minimiser le coût total du voyage. corrTP40005

Correction of the exercise 741

Le coût total de l'expédition est le coût du personnel plus celui du carburant. Afin d'exprimer le tout en fonction de v, il faut se rendre compte que la durée de l'expédition est la distance à parcourir divisée par la vitesse : d/v. À partir de là, la prix du carburant est

$$kv^2 \left(\frac{d}{v} \right) = kdv, \tag{96.519}$$

et le prix du personnel est

$$p \left(\frac{d}{v} \right), \tag{96.520}$$

ce qui donne le coût total en fonction de la vitesse sous la forme

$$C(v) = kdv + \frac{pd}{v}. \tag{96.521}$$

Pour en trouver le maximum, il faut trouver la valeur de v où la dérivée de cette fonction s'annule. La dérivée vaut

$$C'(v) = kd - \frac{pd}{v^2}, \tag{96.522}$$

et cela vaut zéro pour

$$v = \pm\sqrt{\frac{p}{k}}. \tag{96.523}$$

La solution négative est évidement à rejeter. La vitesse optimale est donc

$$v = \sqrt{p/k}. \tag{96.524}$$

Il est intéressant de noter que cette vitesse ne dépend pas de la distance à parcourir.

Exercise 742 exoTP50001

Espace vectoriel de polynômes, bases et matrices.

(1) Montrer que les vecteurs

$$B = \{1, 1-x, 1+x-x^2, 1+x+x^2-x^3\} \tag{96.525}$$

forment une base de l'espace vectoriel des polynômes de degré $\leqslant 3$ en x.

(2) Calculer les composantes du polynôme $1+x+x^2+x^3$ dans la base B.

(3) Chercher la matrice de l'application linéaire φ définie par

$$\varphi\big(p(x)\big) = p'(x) \tag{96.526}$$

pour tout polynôme $p(x)$ dans cette base B.

corrTP50001

Correction of the exercise 742

(1) Montrons que la partie est génératrice, c'est-à-dire que tout polynôme de la forme

$$ax^3 + bx^2 + cx + d \tag{96.527}$$ Eqpolyatrouver

peut être écrit sous la forme

$$\begin{aligned} &\lambda_1(1)\\ &+\lambda_2(1-x)\\ &+\lambda_3(1+x-x^2)\\ &+\lambda_4(1+x+x^2-x^3). \end{aligned} \tag{96.528}$$ EqPOlycofLL

Pour ce faire, nous égalisons les coefficients des termes en chaque puissance de x entre le polynôme (96.527) et (96.528). Nous trouvons

$$\begin{aligned} x^3 &\rightsquigarrow \quad a = -\lambda_4\\ x^2 &\rightsquigarrow \quad b = \lambda_4 - \lambda_3\\ x &\rightsquigarrow \quad c = \lambda_4 + \lambda_3 - \lambda_2\\ \text{terme indépendant} &\rightsquigarrow \quad d = \lambda_4 + \lambda_3 + \lambda_3 + \lambda_1, \end{aligned} \tag{96.529}$$

d'où nous déduisons que

$$\begin{aligned} \lambda_4 &= -a\\ \lambda_3 &= -b - a\\ \lambda_2 &= -c - b - 2a\\ \lambda_3 &= d + c + 2b + 4a. \end{aligned} \tag{96.530}$$ EqSolspolyTPc

Cela montre que chaque vecteur de l'espace considéré s'écrit de façon unique comme combinaison des vecteurs de la partie B. Ceci est exactement dire que B est une base.

(2) Trouver les coefficients de $1+x+x^2+x^3$ est facile : il suffit de reprendre la solution (96.530) avec $a=b=c=d=1$. Cela fait $\lambda_4 = -1$, $\lambda_3 = -2$, $\lambda_2 = -4$, $\lambda_1 = 8$.

(3) C'est ici que les choses compliquées commencent. Il s'agit d'abord de voir les polynômes de base comme des vecteurs. Cela se fait par l'identification suivante :

$$1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad 1-x=\begin{pmatrix}0\\1\\1\\1\end{pmatrix},\quad 1+x-x^2=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad 1+x+x^2-x^3=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{96.531}$$

Maintenant nous devons voir comment la dérivation transforme ces vecteurs les uns en combinaisons des autres. En effet, dans la matrice de φ, la première colonne sera les composantes de l'image du premier vecteur, et ainsi de suite pour les autres colonnes. Nous avons

$$\begin{aligned}
\varphi\begin{pmatrix}1\\0\\0\\0\end{pmatrix}&=(1)'=0=\begin{pmatrix}0\\0\\0\\0\end{pmatrix},\\
\varphi\begin{pmatrix}0\\1\\0\\0\end{pmatrix}&=(1-x)'=-1=\begin{pmatrix}-1\\0\\0\\0\end{pmatrix},\\
\varphi\begin{pmatrix}0\\0\\1\\0\end{pmatrix}&=(1+x-x^2)'=-2x+1=2(1-x)-1\\
&=2\begin{pmatrix}0\\1\\0\\0\end{pmatrix}-\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=\begin{pmatrix}-1\\2\\0\\0\end{pmatrix},\\
\varphi\begin{pmatrix}0\\0\\0\\1\end{pmatrix}&=(1+x+x^2-x^3)'=3(1+x-x^2)+(1-x)-3\\
&=3\begin{pmatrix}0\\0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\end{pmatrix}-3\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=\begin{pmatrix}-3\\1\\3\\0\end{pmatrix}.
\end{aligned} \tag{96.532}$$

La matrice de φ consiste à mettre les images de φ en colonnes, c'est-à-dire

$$\begin{pmatrix}0&-1&-1&-3\\0&0&2&1\\0&0&0&3\\0&0&0&0\end{pmatrix}. \tag{96.533}$$

Exercise 743 exoTP50002

Trouver toutes les solutions du système

$$\begin{cases}x+2y+z-t=3\\2x-y+t=2\\5y+2z-3t=4\end{cases} \tag{96.534}$$

en utilisant la méthode de Gauss. corrTP50002

Correction of the exercise 743

Étant donné que nous avons seulement 3 équations pour 4 inconnues, nous savons que nous allons pouvoir résoudre en laissant (au moins) une variable dans la solution. Étant donné qu'il

n'y a pas de zéros dans la colonne du t, je propose de choisir celle-là : c'est la colonne la moins appétissante, donc autant la choisir. Au départ, le système s'écrit sous la forme

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 3 \\ 2 & -1 & 0 & 1 & 2 \\ 0 & 5 & 2 & -3 & 4 \end{array}\right). \tag{96.535}$$

Nous allons essayer de ne pas casser les zéros déjà présents. La première chose à faire est d'annuler la première case de la première ligne en utilisant la seconde : $L_1 \to L_1 - L_2/2$. Nous trouvons

$$\left(\begin{array}{cccc|c} 0 & 5/2 & 1 & -3/2 & 2 \\ 2 & -1 & 0 & 1 & 2 \\ 0 & 5 & 2 & -3 & 4 \end{array}\right). \tag{96.536}$$

Ensuite, nous pouvons annuler la case du z dans la première ligne en utilisant la troisième : $L_1 \to L_1 - L_3/2$. Nous avons alors

$$\left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 1 & 2 \\ 0 & 5 & 2 & -3 & 4 \end{array}\right). \tag{96.537}$$

La première ligne est maintenant complètement nulle, on peut la barrer. Il ne reste donc que deux équations pour quatre inconnues, et nous pouvons donc choisir une seconde variable par rapport à laquelle nous ne voulons pas résoudre le système (une variable qui va être reléguée au rang de simple paramètre dans la solution). Étant donné que la colonne des y est sans zéros, nous choisissons celle-là. Le système est donc

$$\begin{cases} 2x - y + t = 2 & \text{(96.538a)} \\ 5y + 2z - 3t = 4. & \text{(96.538b)} \end{cases}$$

Comme nous avons décidé de laisser y et t, ce système est résolu. La première équation donne directement

$$x = \frac{1}{2}(y - t + 2) \tag{96.539}$$

et la seconde donne

$$z = \frac{1}{2}(3t - 5y + 4). \tag{96.540}$$

Exercise 744 exoTP50003

Donner toutes les valeurs propres et vecteurs propres des applications linéaires $E \to E$ suivantes.

(1) Une rotation de 180 degrés autour de

$$\begin{cases} x = y \\ y = z \end{cases} \tag{96.541}$$

(2) Une rotation de 90 degrés autour de l'axe Ox.

(3) La symétrie par rapport à $x + y + z = 0$.

(4) La symétrie par rapport à l'axe Oz composé avec l'homothétie de rapport 5.

corrTP50003

Correction of the exercise 744

Une bonne idée est de lire le point 8.18.1 avant de se lancer dans cet exercice.

(1) Nous savons que le vecteur

$$f_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{96.542}$$

est sur la droite, et donc ne va pas bouger. Ce vecteur est donc un vecteur propre de valeur propre 1. Les vecteurs perpendiculaires à la droite vont changer de signe. Il n'est pas très difficile d'en trouver deux linéairement indépendants. Un premier qui est perpendiculaire à f_1 est par exemple

$$f_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{96.543}$$

Pour vérifier qu'il est perpendiculaire, penser au produit scalaire. Pour en trouver f_3 qui est à la fois perpendiculaire à la droite et à f_2, prenons le produit vectoriel :

$$f_3 = f_1 \times f_2 = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{vmatrix} = -e_1 + 2e_2 - e3 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}. \tag{96.544}$$

Par définition, si A est la rotation demandée, nous avons

$$\begin{aligned} Af_1 &= f_1 \\ Af_2 &= -f_2 \\ Af_3 &= -f_3 \end{aligned} \tag{96.545}$$

Ce sont donc tous les trois des vecteurs propres, de valeurs propres respectivement 1, -1 et -1.

Au cas où vous voudriez savoir la matrice de la rotation, il faut d'abord exprimer e_1, e_2 et e_3 en termes de f_1, f_2 et f_3 :

$$\begin{aligned} e_1 &= \frac{1}{3}f_1 + \frac{1}{2}f_2 - \frac{1}{6}f_3 \\ e_2 &= \frac{1}{3}f_1 + \frac{1}{3}f_3 \\ e_3 &= \frac{1}{3}f_1 - \frac{1}{2}f_2 - \frac{1}{6}f_3 \end{aligned} \tag{96.546}$$

Maintenant, trouver Ae_1, Ae_2 et Ae_3 revient à changer quelques signes dans ces équation, et nous trouvons

$$A = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}. \tag{96.547}$$

Vous pouvez même vous amuser à trouver les valeurs propres et vecteurs propres de cette matrice (vous les savez déjà !).

(2) La rotation de 90 degrés autour de Ox laisse e_x invariant, puis fait $e_y \to e_z$ et $e_z \to -e_y$ (ne pas oublier ce signe, faire un dessin !). Nous avons donc

$$\begin{aligned} A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\ A\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\ A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \end{aligned} \tag{96.548}$$

et donc la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{96.549}$$

C'est certain que e_1 est vecteur propre de valeur propre 1. Pour les autres, voyons l'équation caractéristique :

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2+1) = 0. \tag{96.550}$$

Seul $\lambda = 1$ est solution, et comme cette racine est une racine simple, il n'y a que un seul vecteur propre correspondant. Nous concluons que e_1 (et ses multiples) est l'unique vecteur propre. Nous pouvons cependant le vérifier en résolvant l'équation

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{96.551}$$

Le système correspondant est

$$\begin{cases} x = x \\ -z = y \\ y = z. \end{cases} \tag{96.552}$$

Les deux dernières équations imposent $y = z = 0$, tandis que la première laisse x libre. Nous avons donc bien que les multiples de e_1.

(3) Trouvons d'abord deux vecteurs linéairement indépendants dans le plan $x+y+z = 0$. Prenons par exemple

$$f_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \tag{96.553}$$

et

$$f_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{96.554}$$

Noter qu'ils ne sont pas perpendiculaires, mais ce n'est pas important. L'important est que ces deux vecteurs soient linéairement indépendants et qu'ils restent inchangés par la symétrie considérée. Ils sont donc vecteurs propres de valeur propre 1. Pour trouver un vecteur perpendiculaire au plan (et qui sera donc vecteur propre de valeur propre -1), nous prenons le produit vectoriel :

$$f_3 = f_1 \times f_2 = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{vmatrix} = e_1 + e_2 + e_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{96.555}$$

Cela est un troisième vecteur propre. Sa valeur propre est -1.

(4) La symétrie autour de l'axe Oz fait en sorte que e_z ne change pas, et que les deux autres changent de signe. Ensuite, la dilatation de facteur 5 multiplie tout par 5. Nous avons donc au final

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &\to \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} &\to \begin{pmatrix} O \\ -5 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &\to \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{96.556}$$

Ce sont les trois vecteurs propres, de valeurs propres 5, -5 et -5.

Exercise 745 exoTP50004

On donne la matrice

$$M = \begin{pmatrix} a & 0 & b \\ b & a & 0 \\ b & 0 & a \end{pmatrix} \tag{96.557}$$

avec $a, b \in \mathbb{R}$.

(1) La matrice M est-elle inversible ? Discuter suivant les valeurs de a et b si nécessaire.

(2) Déterminer les valeurs propres et les vecteurs propres de M.

(3) Pour quelles valeurs de a et de b la matrice M est-elle diagonalisable ?

corrTP50004

Correction of the exercise 745

Une matrice est inversible si et seulement si son déterminant est non nul. Ici, le déterminant vaut

$$\det(M) = a(a^2 - b^2). \tag{96.558}$$

La matrice M sera donc non inversible si $a = 0$ ou bien si $a = \pm b$. Dans tous les autres cas, elle est inversible.

Pour les vecteurs et valeurs propres, nous regardons l'équation caractéristique

$$\begin{vmatrix} a-\lambda & 0 & b \\ b & a-\lambda & 0 \\ b & 0 & a-\lambda \end{vmatrix} = (a-\lambda)\big((a-\lambda)^2 - b^2\big) = 0. \tag{96.559}$$

La première valeur propre évidente est $\lambda_1 = a$. Les deux autres sont les solutions de $(a-\lambda)^2 - b^2$, qui est une équation usuelle du second degré. Les solutions sont $a + b$ et $a - b$. Au final :

$$\begin{aligned} \lambda_2 &= a \\ \lambda_2 &= a + b \\ \lambda_3 &= a - b. \end{aligned} \tag{96.560}$$

La proposition 14.8 de la page 202 du cours dit que si une matrice a trois valeurs propres distinctes, alors elle est diagonalisable. Ici, c'est le cas tant que $b \neq 0$. Si $b = 0$, alors les choses sont moins claires. Lorsque $a = 0$, l'équation caractéristique devient

$$(a-\lambda)^3 = 0, \tag{96.561}$$

dont la solution $\lambda = a$ est une racine triple. Il y a donc de fortes chances qu'il y ait trois vecteurs propres correspondants. En effet, dans le cas $b = 0$, la matrice M se réduit à

$$M_{b=0} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}. \tag{96.562}$$

Là, c'est bien clair que les trois vecteurs de base e_1, e_2 et e_3 sont vecteurs propres de valeur propre a (sauf si $a = 0$). En résumé, la matrice M est diagonalisable tout le temps sauf si $a = 0$ et $b = 0$ en même temps.

96.8.9 Quelques fautes usuelles

Pour l'exercice 737, les fautes les plus souvent commises sont

(1) $f' = e^{2x}$ implique $f = \frac{1}{2}e^x$. Cela n'est pas vrai. La dérivée de e^{2x} est $2e^{2x}$. Le 2 reste dans l'exponentielle.

(2) Lorsqu'on intègre par partie, il faut aussi mettre les bornes pour le morceau qui n'est pas dans la nouvelle intégrale :

$$\int_a^b fg' = [fg]_a^b - \int_a^b f'g. \tag{96.563}$$

Pour l'exercice 738, les fautes les plus souvent commises sont

(1) Lorsqu'on a trouvé la solution générale $y_k(x)$ qui dépend du paramètre k (ou C), il faut encore trouver la valeur du paramètre k telle que $y_k(\pi) = 0$.

Pour l'exercice 739, les fautes les plus souvent commises sont

(1) Ne pas oublier que $e^0 = 1$.

96.9 Intégrales de surface, Stokes et Green

Exercice 6

(*a*) **La suite** $[k \to \frac{1}{k}]$ **est convergente.**

Nous allons montrer que cette suite converge vers 0. Il faut donc prouver la chose suivante :

$$\forall \epsilon > 0 \quad \exists K_\epsilon \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_\epsilon, \quad |x_k - x| < \epsilon \tag{96.564}$$

Remarque : On pourrait également montrer que cette suite est *de Cauchy* pour prouver qu'elle est convergente sans devoir déterminer sa limite.

Pour prouver que (96.564) s'applique bien à la suite des $\frac{1}{k}$ il nous faut montrer que

$$\forall \epsilon > 0 \quad \exists K_\epsilon \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_\epsilon, \quad \frac{1}{k} < \epsilon \tag{96.565}$$

Ceci est une conséquence immédiate de l'exercice précédent. On peut également le montrer de la manière suivante : à ϵ positif donné, si nous arrivons à déterminer l'indice K_ϵ de (96.565) tel que $\forall k \geqslant K_\epsilon, \quad \frac{1}{k} < \epsilon$, il est clair que la suite satisfait à la définition. Or, $\frac{1}{k} < \epsilon \leftrightarrow \frac{1}{\epsilon} < k$. Donc si nous prenons $K_\epsilon := \lceil 1/\epsilon \rceil + 1$, on a bien que $\forall k \geqslant K_\epsilon$, $\frac{1}{k} < \epsilon$, ce qui est ce qu'il fallait démontrer.

(*b*) **La suite** $(1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \ldots)$ **est convergente.**

On remarque que cette suite tend vers zéro. (Il suffit de voir que le numérateur est borné et que le dénominateur tend vers l'infini). Si on l'écrit sous la forme standard, on obtient :

$$x_1 = 1, x_k = \frac{(-1)^k}{k} \quad \forall k \geqslant 2$$

Donc, ce que nous voulons voir est que $x_k \longrightarrow_{k\to\infty} 0$, i.e. :

$$\forall \epsilon > 0 \quad \exists K_\epsilon \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_\epsilon, \quad |\frac{(-1)^k}{k}| < \epsilon \tag{96.566}$$

Étant donné que $|(-1)^k| = 1 \, \forall k$, il est clair que l'équation (96.566) est la même que l'équation (96.565), et donc que l'on peut affirmer que pour tout $\epsilon > 0$, il suffit de prendre $K \geqslant \frac{1}{\epsilon}$ et la condition est satisfaite.

Exercice 7

Ici il est demandé de prouver de nouvelles règles de calcul en repartant de la définition de la convergence vers l'infini :

$$x_k \longrightarrow \infty \quad \text{si} \quad \forall M > 0 \quad \exists K_M \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M, x_k \geqslant M \tag{96.567}$$

(a) $\lim(x_k + y_k) = +\infty$**.**

On veut voir la chose suivante :

$$\forall M>0 \quad \exists K_M \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M, x_k + y_k \geqslant M \tag{96.568}$$

Soit $M>0$. Comme x_k et y_k convergent à l'infini, on sait que

$$\begin{cases} \exists K_M^x \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M^x, x_k \geqslant \frac{M}{2} \\ \exists K_M^y \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M^y, y_k \geqslant \frac{M}{2}, \end{cases}$$

et donc il suffit de prendre $K_M = \max(K_M^x, K_M^y)$ dans (96.568) pour s'assurer que la définition est satisfaite.

(b) $\lim(x_k y_k) = +\infty$**.**

On veut voir la chose suivante :

$$\forall M>0 \quad \exists K_M \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M, x_k y_k \geqslant M \tag{96.569}$$

Soit $M>0$. Comme x_k et y_k convergent à l'infini, on sait que

$$\begin{cases} \exists K_M^x \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M^x, x_k \geqslant \sqrt{M} \\ \exists K_M^y \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M^y, y_k \geqslant \sqrt{M}, \end{cases}$$

et donc il suffit de prendre $K_M = \max(K_M^x, K_M^y)$ dans (96.569) pour s'assurer que la définition est satisfaite.

(d) Soit z_k **une suite tendant vers un réel** a **strictement positif. Prouvez que** $\lim(x_k z_k) = +\infty$**.**

Le but de l'exercice est toujours le même, c'est-à-dire de prouver que

$$\forall M>0 \quad \exists K_M \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_M, \ x_k z_k \geqslant M \tag{96.570}$$

Soit $M>0$. On sait que :

$$\begin{cases} \forall \tilde{M}>0 \ \exists K_{\tilde{M}}^x \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_{\tilde{M}}^x, \ x_k \geqslant \tilde{M} \\ \forall \epsilon>0 \ \exists K_\epsilon^z \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_\epsilon^z, \ |z_k - a| < \epsilon, \end{cases} \tag{96.571}$$

Prenons un ϵ tel que $a-\epsilon>0$. Par la deuxième partie de (96.571) on voit qu'il existe un indice K_ϵ^z tel que $\forall k \geqslant K_\epsilon^z, \ z_k > a-\epsilon > 0$.
Prenons un $\tilde{M}$ tel que $M = \tilde{M}(a-\epsilon)$. Par la première partie de (96.571) on voit qu'il existe un indice $K_{\tilde{M}}^x$ tel que $\forall k \geqslant K_{\tilde{M}}^x, \ x_k \geqslant \tilde{M}$.
et donc il suffit de prendre $K_M = \max(K_{\tilde{M}}^x, K_\epsilon^z)$ dans (96.570) pour avoir que

$$\forall k \geqslant K_M, \ x_k z_k \geqslant \tilde{M}(a-\epsilon) = M.$$

Exercice 8

Une suite x_k est bornée si $\exists N>0$ tel que $\forall k, |x_k| < N$.
On veut voir que $\frac{x_k}{y_k} \longrightarrow 0$, i.e.

$$\forall \epsilon>0 \quad \exists K_\epsilon \in \mathbb{N} \quad \text{tq} \quad \forall k \geqslant K_\epsilon, \ |\frac{x_k}{y_k}| < \epsilon \tag{96.572}$$

Soit $\epsilon>0$. Comme la suite x_k est bornée, on a $|\frac{x_k}{y_k}| < \frac{N}{|y_k|} \ \forall k$. On utilise maintenant le fait que $y_k \longrightarrow \infty$. Prenons $M = \frac{N}{\epsilon}$. On peut écrire que $\exists K_M$ tel que $\forall k \geqslant K_M, \ y_k \geqslant M = \frac{N}{\epsilon}$, et donc si dans (96.572) on prend $K_\epsilon = K_M$ on a :

$$\forall k \geqslant K_\epsilon, \ |\frac{x_k}{y_k}| < \frac{N}{|y_k|} < \frac{N}{N/\epsilon} = \epsilon.$$

Exercice 9

Pour cet exercice, on peut utiliser les règles de calcul. Il faut faire attention que ces règles ne s'appliquent que si toutes les limites existent !

(a) $x_k = \frac{k+2}{k}\cos(k\pi)$

On voit que cette suite va dans deux directions différentes, $+1$ et -1 à cause du facteur $\cos(k\pi) = (-1)^k$. Elle ne converge donc pas. Pour le prouver, on peut prendre deux suites partielles de la suite x_k qui convergent vers des limites différentes.
Choisissons

$$\left\{ \begin{array}{lll} y_k & = x_{2k} & = \frac{(2k)+2}{2k}(-1)^{2k} \\ z_k & = x_{2k+1} & = \frac{(2k+1)+2}{2k+1}(-1)^{2k+1} \end{array} \right.$$

Comme $x_k = \frac{k+1}{k} = 1 + \frac{1}{k}$ et que $\frac{1}{k} \to 0$, nous pouvons appliquer les règles de calcul et en déduire que $x_k \to 1$. On fait la même chose pour y_k.

(c) $x_k = \frac{k^3+k+1}{5k^3+2}$

Nous avons que

$$\forall k, \quad x_k = \left(\frac{k^3}{k^3}\right)\frac{1 + \frac{1}{k} + \frac{1}{k^3}}{5 + \frac{2}{k^3}} = \frac{1 + \frac{1}{k} + \frac{1}{k^3}}{5 + \frac{2}{k^3}}$$

Comme

$$\forall k \geqslant 1 \ \frac{1}{k^3} \leqslant \frac{1}{k^2} \leqslant \frac{1}{k}$$

et comme $\frac{1}{k} \to 0$, nous pouvons appliquer la règle de l'étau pour voir que

$$\frac{1}{k^3} \to 0 \quad \text{et} \quad \frac{1}{k^2} \to 0.$$

En appliquant les règles de calcul à la suite x_k transformée, on voit donc que $x_k \to \frac{1}{5}$.

(d) $x_k = \frac{k+(-1)^k}{k-(-1)^k}$

On peut le voir par exemple par la règle de l'étau :

$$\forall k \geqslant 0, \quad \frac{k-1}{k+1} \leqslant \frac{k+(-1)^k}{k-(-1)^k} \leqslant \frac{k+1}{k-1}.$$

Or, comme les deux suites qui bornent la suite x_k convergent toutes les deux vers 1, il est clair que x_k converge aussi vers 1.

(d) $x_k = x_{k-1}^2 + 1$, $x_1 = 1$

Suite définie par récurrence. Ses premiers éléments sont

$$1,\ 2,\ 5,\ 26,\ 677,\ \ldots$$

Toute limite admissible réelle finie l de cette suite doit satisfaire à

$$l = l^2 + 1$$

ce qui implique qu'elle ne peut avoir de limite réelle finie. En regardant ses premiers éléments, on remarque immédiatement qu'elle semble converger à l'infini. Nous allons le prouver en utilisant la définition.

Soit $M > 0$. On a

$$x_k \geqslant k \, \forall k.$$

En effet (par récurrence sur k) : il est clair que $x_1 \geqslant 1$. Supposons que $x_k \geqslant k$. Ceci implique t-il que $x_{k+1} \geqslant k+1$? Par définition des x_k, $x_{k+1} = x_k^2 + 1$. Par l'hypothèse de récurrence, on a donc $x_{k+1} \geqslant (k)^2 + 1 \geqslant k+1$ ce qui prouve le résultat. Comme la suite $y_k = k$ converge à l'infini, il en est de même pour la suite x_k.

96.10 Continuité de fonctions réelles

Travaux Personnels

BAC2 en sciences mathématiques et physiques

Exercice 1. Calculer les limites suivantes

a) $\lim_{n\to\infty} \left(1 + \frac{2}{n-4}\right)^n$

b) $\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{\sqrt{n}}$

c) $\lim_{x\to\infty} \left(1 + \frac{\alpha}{x}\right)^x$

d) $\lim_{x\to 0} \frac{\log(1+\alpha x)}{x}$

e) $\lim_{x\to\infty} \frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_m x^m}$ où $a_j, b_j \in \mathbb{C}$ et $n, m \geqslant 0$

f) $\lim_{x\to 0} \frac{\sqrt{1-\cos x}}{x}$

Exercice 2. Prouver que

a) $\lim_{x\to\infty} x^{\frac{1}{x}} = \lim_{x\to 0^+} x^x = 1$

b) $\lim_{x\to\infty} \frac{x^{\ln x}}{\mathrm{e}^x} = 0$ càd e^x croit plus vite que $x^{\ln x}$

Exercice 3. Prouver que

$$\cosh 2x \;=\; \cosh^2 x + \sinh^2 x, \qquad \sinh 2x \;=\; 2\sinh x \cosh x$$

Exercice 4. Prouver que

a) $1 + \cos z + \cos 2z + \cdots + \cos nz = \cos \frac{nz}{2} \cdot \frac{\sin(n+1)z/2}{\sin z/2}$

b) $1 + \sin z + \sin 2z + \cdots + \sin nz = \sin \frac{nz}{2} \cdot \frac{\sin(n+1)z/2}{\sin z/2}$

Aide : $\sum_{k=0}^{n} \mathrm{e}^{\mathrm{i}kz} = \frac{1-\mathrm{e}^{\mathrm{i}(n+1)z}}{1-\mathrm{e}^{\mathrm{i}z}} = \mathrm{e}^{\mathrm{i}nz/2} \cdot \frac{\mathrm{e}^{\mathrm{i}(n+1)z/2} - \mathrm{e}^{-\mathrm{i}(n+1)z/2}}{\mathrm{e}^{\mathrm{i}z/2} - \mathrm{e}^{-\mathrm{i}z/2}}$

Rappelons qu'une fonction $f\colon \mathbb{C} \supset D \to \mathbb{C}$ est **uniformément continue** si pour tout $\epsilon > 0$ il existe un $\delta > 0$ tel que

$$|x-y| < \delta \implies |f(x) - f(y)| < \epsilon \quad \text{pour tout } x, y \in D.$$

Prouver que la fonction $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ est continue, mais n'est pas uniformément continue.

96.11 Intégrales, longueur de courbes, EDO's linéaires

Exercice 2. Soient $n, m \in \mathbb{N} \cup \{0\}$. Calculer

$$\int_0^1 x^n(1-x)^m\,dx \quad \text{et} \quad \int_{-1}^1 (1+x)^n(1-x)^m\,dx$$

Solution : Posons $I_{n,m} := \int_0^1 x^n(1-x)^m\,dx$. Intégration par partie donne la formule récursive

$$I_{n,m} \;=\; \frac{m}{n+1} I_{n+1,m-1}.$$

Avec $I_{n+m,0} = \frac{1}{n+m+1}$ nous obtenons

$$I_{n,m} \;=\; \frac{n!\,m!}{(n+m+1)!}$$

La substitution $x := 2t - 1$ fournit

$$\int_{-1}^1 (1+x)^n(1-x)^m\,dx \;=\; 2^{n+m+1}\int_0^1 t^n(1-t)^m\,dt \;=\; 2^{n+m+1}\,\cdot\, I_{n,m}.$$

Exercice 3. Soient $a, b > 0$. Calculer

$$\int_0^{\pi/2} \frac{d\varphi}{a^2\sin^2\varphi + b^2\cos^2\varphi}$$

Solution :

$$= \int_0^{\pi/2} \frac{1/\cos^2\varphi}{a^2\tan^2\varphi + b^2} d\varphi \;=\; \int_0^\infty \frac{dt}{a^2t^2+b^2} \;=\; \frac{\pi}{2ab}$$

Exercice 4. Calculer la longueur de l'arc de la parabole $y = x^2,\ x \in [0, b]$.

Solution :

$$s \;=\; \int_0^b \sqrt{1+4x^2}\,dx \;=\; \frac{b}{2}\sqrt{1+4b^2} + \frac{1}{4}\ln\left(2b + \sqrt{1+4b^2}\right)$$

Exercice 5. La **parabole de Neil** ν est la courbe définie par $\nu(t) = (t^2, t^3)$, pour $t \in \mathbb{R}$.

a) Esquisser la parabole de Neil.

b) Quelle est la signification du paramètre t ?

Solution : $t = \tan\alpha$

c) Calculer la longueur de l'arc $\{\nu(t) \mid t \in [0, \tau]\}$.

Solution :

$$s \;=\; \int_0^\tau \sqrt{4t^2+9t^4}\,d\tau \;=\; \frac{8}{27}\left(\left(1+\frac{9}{4}\tau^2\right)^{3/2} - 1\right)$$

Exercice 6. La **hélice** γ de pas $2\pi h$ est la courbe dans $\mathbb{R}^3$ définie par

$$\gamma(t) \;=\; (r\cos t, r\sin t, ht)\,.$$

a) Esquisser la hélice.

b) Expliquer le mot "pas".

c) Calculer la longueur de l'arc sur la hélice si on fait un tour.

Solution : $\int_0^{2\pi} \sqrt{r^2+h^2}\,dt \;=\; 2\pi\sqrt{r^2+h^2}$

Exercice 7. Calculer un système fondamental réel pour

a) $y^{(4)} - y = 0$,

b) $y^{(4)} + 4y'' + 4y = 0$,

c) $y^{(4)} - 2y^{(3)} + 5y'' = 0$.

Solution :

a) $\mathrm{e}^x, \mathrm{e}^{-x}, \cos x, \sin x$

b) $\cos\sqrt{2}x, x\cos\sqrt{2}x, \sin\sqrt{2}x, x\sin\sqrt{2}x$

c) $1, x, \mathrm{e}^x\cos 2x, \mathrm{e}^x\sin 2x$

Exercice 8. Déterminer une solution particulière de l'équation $y'' + y = q$ pour

a) $q = x^3$,

b) $q = \sinh x$,

c) $q = 1/\sin x$.

Solution :

a) $x^3 - 6x$

b) $\frac{1}{2}\sinh x$

c) $\sin x \cdot \ln|\sin x| - x\cos x$

Exercice 9. L'équation différentielle $m\ddot{y} = mg - k\dot{y}$ décrit la chute d'un corps soumit à la gravitation si la friction est proportionnelle à la vitesse ("un homme tombant de l'avion").

Calculer la solution avec $y(0) = 0, \dot{y}(0) = 0$. Calculer la "vitesse finale" $v_\infty = \lim_{t\to\infty}\dot{y}(t)$.

Solution :

L'équation homogène $\ddot{y} + k/m \cdot y = 0$ possède les solutions $c_1 + c_2\mathrm{e}^{-k/m\,\cdot\,t}$, où $c_1, c_2 \in \mathbb{R}$.

L'équation inhomogène $\ddot{y} + k/m \cdot y = g$ possède comme solution particulière une fonction lineaire, càd $y_p = (mg/k)t)$. En tenant compte des conditions initiales nous obtenons

$$y(t) = \frac{mg}{k}\left(t - \frac{m}{k}(1 - \mathrm{e}^{-k/m\,\cdot\,t})\right).$$

En particulier, $v_\infty = mg/k$.

Exercice 10. Regardons l'ensemble des solutions de l'équation différentielle $P(\mathrm{D})y = 0$.

Montrer l'équivalence entre les propositions suivantes :

(1) Pour toute solution y on a $\lim_{t\to\infty} y(t) = 0$

(2) Pour toute racine z du polynôme caractéristique on a $\mathrm{Re}\, z < 0$.

Dans ce cas, l'équation différentielle est appelé **asymptotiquement stable**.

Solution : On a $\lim_{t\to\infty} y(t) = 0$ pour toute solution y ssi c'est vrai pour tout élément d'un système fondamental. On a $\lim_{t\to\infty} t^k\mathrm{e}^{\mathfrak{gl}t} = 0$ ssi $\mathrm{Re}\,\mathfrak{gl} < 0$, d'où l'affirmation suit.

96.12 Calcul de limites

Exercice 11. Déterminez si les limites suivantes existent et dans l'affirmative calculez les en utilisant, s'il y a lieu, la règle de l'Hospital ou la règle de l'étau.

(1) $\lim_{x\to+\infty} \frac{x+1}{x^2+2}$

(2) $\lim_{x\to+\infty} \frac{\sin(x)}{x}$

(3) $\lim_{x\to0} \frac{\sin(x)}{x}$

(4) $\lim_{x\to+\infty} \frac{x^n}{e^x}$

(5) $\lim_{x\to+\infty}(1+\frac{a}{x})^x$

(6) $\lim_{x\to0}(\frac{1}{\sin(x)}-\frac{1}{x})$

(7) $\lim_{x\to+\infty} \cos(2\pi x)$

(8) $\lim_{x\to+\infty} \frac{1}{\sin(x)+2}(x)+\ln(x)\cos(x)$

(9) $\lim_{x\to+\infty} \frac{\ln(x)(\sin(x)+2)}{x}$

(10) $\lim_{x\to+\infty} x^{\frac{1}{x}}$

Exercice 12. Déterminez si les limites suivantes existent et dans l'affirmative calculez-les.

(1) $\lim_{x\to0} x\sin(\frac{1}{x})$

(2) $\lim_{x\to0} \frac{\sin(\sin(x))}{x}$

(3) $\lim_{x\to+\infty}(\ln(x))^{\frac{1}{1-\ln(x)}}$

Exercice 13. Calculez les limites suivantes :

(1) $\lim_{x\to+\infty} \frac{\ln(x)}{x^a}$

(2) $\lim_{x\to+\infty} \frac{\ln(x)^a}{x^b}$

(3) $\lim_{x\to+\infty} a^x$

(4) $\lim_{x\to+\infty} a^{\frac{1}{x}}$

où a et b sont des réels positifs.

Exercice 14. Déterminez, pour chacune des suites suivantes, si elle converge et dans l'affirmative calculez sa limite.

(1) $k\to\cos(2\pi k)$

(2) $k\to\cos(\frac{\pi}{3}k)$

(3) $k\to k(a^{\frac{1}{k}}-1)$

où a est une réel.

Exercice 15. Calculez les limites suivantes si elles existent.

(1) $\lim_{x\to+\infty}\cos x$

(2) $\lim_{x\to\pm\infty}\sqrt{2x^4+3}-x^2$

Exercice 16. Déterminez si la limite de chacune des suites suivantes existe et dans l'affirmative calculez la.

(1) $\lim_{k\to+\infty}(\frac{ak+1}{k})^k$

(2) $\lim_{k\to+\infty} \frac{1}{\sin(\frac{\pi}{6}k)+1}(k)+\ln(k)\cos(\frac{\pi}{5}k)$

(3) $\lim_{k\to+\infty} \frac{\ln(k)(\sin(\frac{\pi}{3}k)+1)}{k}$

(4) $\lim_{k\to+\infty} \sqrt[3k]{k}(1+\frac{1}{3k})^{3k}$

où a est un réel.

96.13 Dérivabilité

Exercice 17. Déterminez l'ensemble des points où les fonctions suivantes sont continues et celui où elles sont dérivables. Prouvez soigneusement vos résultats.

(1) $x \to x]$

(2) $x \to |x|$

(3) $x \to \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(4) $x \to x^2$

Exercice 18. Étudiez la dérivabilité et la continuité de la dérivée de chacune des fonctions suivantes :

(1) $x \to \begin{cases} 0 & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

(2) $x \to \begin{cases} \sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(3) $x \to \begin{cases} x\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(4) $x \to \begin{cases} x^2\sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

Le but de cet exercice est aussi d'exhiber des exemples illustrant les différents types de comportements possibles, relativement à la continuité et la dérivabilité, d'une fonction en un point.

Exercice 19. Étudiez la dérivabilité et la continuité de la dérivée de chacune des fonctions suivantes :

(1) $x \to \begin{cases} \frac{2x+a}{1+e^{\frac{1}{x}}} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(2) $x \to \begin{cases} \frac{\sin(x)}{x} & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

(3) $x \to \begin{cases} e^{\frac{-1}{x}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$

(4) $[-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R} : x \to \begin{cases} (\frac{\sin(2x)}{x})^{x+1} & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

où a et b sont des réels.

Exercice 20. Considérons la fonction

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto f(x) = \begin{cases} x & \text{si } x \text{ est rationnel} \\ 0 & \text{si } x \text{ est irrationnel} \end{cases}$$

Vérifiez que f est continue en 0 mais n'est ni dérivable à gauche ni dérivable à droite en 0.

Exercice 21.

(1) Soit (X, d) un espace métrique et $f\colon (X, d) \to \mathbb{R}$ une application continue. Montrer que l'ensemble

$$\{x \mid f(x) = 0\}$$

est fermé.

(2) Soit $f\colon \mathbb{R} \to \mathbb{R}$ une application continue. Montrer que l'ensemble

$$\{x \in \mathbb{R} \mid f(x) = x\}$$

des points fixes de f est fermé.

Exercice 22. Soit A un sous-ensemble de l'espace métrique (X, d). Montrer que la fonction

$$\mathrm{dist}_A \colon X \to \mathbb{R}, \quad x \mapsto \inf_{a \in A} d(a, x)$$

est continue.

Exercice 23. Soient (X, d_X), (Y, d_Y) deux espaces métriques. Une application $f \colon X \to Y$ est **Lipschitzienne** si il il existe une constante $L \geqslant 0$ telle que

$$d_Y\big(f(x), f(x')\big) \leqslant L\, d_X(x, x') \quad \text{pour tout } x, x' \in X.$$

Dans ce cas, on dit que f est L**-Lipschitzienne**.

(1) Montrer qu'une application Lipschitzienne est continue.

(2) Montrer qu'une application $f \colon \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$ est Lipschitzienne. Quelle est la plus petite constante L qui convienne ?

(3) Montrer que les fonctions $z \mapsto |z|$, $z \mapsto \overline{z}$, $z \mapsto \mathrm{Re}\, z$ et $z \mapsto \mathrm{Im}\, z$ de $\mathbb{C}$ dans $\mathbb{R}$ sont Lipschitziennes. Quelle sont les plus petites constantes L qui conviennent ?

(4) Montrer que la fonction $\mathrm{dist}_A \colon X \to \mathbb{R}$ de l'Exercice 13 est Lipschitzienne.

96.14 Intégration

Exercice 24. Soient $n, m \in \mathbb{N} \cup \{0\}$. Calculer

$$\int_0^1 x^n (1-x)^m \, dx \quad \text{et} \quad \int_{-1}^1 (1+x)^n (1-x)^m \, dx$$

Exercice 25. Soient $a, b > 0$. Calculer

$$\int_0^{\pi/2} \frac{d\varphi}{a^2 \sin^2 \varphi + b^2 \cos^2 \varphi}$$

Exercice 26. Calculer la longueur de l'arc de la parabole $y = x^2$, $x \in [0, b]$.

Exercice 27. La **parabole de Neil** ν est la courbe définie par $\nu(t) = (t^2, t^3)$, $t \in \mathbb{R}^n$.

(1) Esquisser la parabole de Neil.

(2) Quelle est la signification du paramètre t ?

(3) Calculer la longueur de l'arc $\{\nu(t) \mid t \in [0, \tau]\}$.

Exercice 28. Une **hélice** γ de pas $2\pi h$ est une courbe dans $(\mathbb{R}^n)^3$ définie par

$$\gamma(t) \;=\; (r\cos t, r \sin t, ht)\,.$$

(1) Esquisser γ et expliquer le mot "pas".

(2) Calculer la longueur de l'arc sur la hélice si on fait un tour.

Exercice 29. Calculez la longueur des arcs de courbe suivants :

(1) $y = \ln(1 - x^2)$ $\qquad 0 \leqslant x \leqslant \frac{1}{2}$

(2) $y = x^{3/2}$ $\qquad 0 \leqslant x \leqslant 5$

(3) $y = 1 - \ln(\cos x)$ $\qquad 0 \leqslant x \leqslant \frac{\pi}{4}$

(4) l'arc de cubique déterminé par $y = x^3 + x^2 + x + 1$ avec $0 \leqslant x \leqslant 1$.

96.15 Quelques corrections

Ce qui suit sont des corrections d'exercices donnée sur les feuilles distribuées au début de l'année.

31.

(1) $df_{(1,1)}$ et $dg_{(\sqrt{2},\frac{\pi}{4})}$

$$\begin{array}{l}\frac{\partial f}{\partial x}(x,y) = \frac{1}{y}\ln(\frac{x}{y})e^{\frac{x}{y}} + \frac{1}{x}e^{\frac{x}{y}}\\ \frac{\partial f}{\partial x}(1,1) = e\\ \frac{\partial f}{\partial y}(x,y) = -\frac{x}{y^2}\ln(\frac{x}{y})e^{\frac{x}{y}} - xe^{\frac{x}{y}}\\ \frac{\partial f}{\partial x}(1,1) = e\end{array}$$

Par les règles de calcul, f est différentiable en $(1,1)$. La différentielle $df_{(1,1)}$ est donc représentée dans les bases canoniques de $\mathbb{R}^2$ et $\mathbb{R}$ par la matrice jacobienne (ici gradient) :

$$df_{(1,1)} = (e \quad -e)$$

$$\begin{array}{lcllcl}\frac{\partial g_1}{\partial r}(r,\theta) & = & \cos(\theta) & \frac{\partial g_1}{\partial \theta}(r,\theta) & = & -r\sin(\theta)\\ \frac{\partial g_1}{\partial r}(\sqrt{2},\frac{\pi}{4}) & = & \frac{\sqrt{2}}{2} & \frac{\partial g_1}{\partial \theta}(\sqrt{2},\frac{\pi}{4}) & = & -1\\ \frac{\partial g_2}{\partial r}(r,\theta) & = & \sin(\theta) & \frac{\partial g_2}{\partial \theta}(r,\theta) & = & r\cos(\theta)\\ \frac{\partial g_2}{\partial r}(\sqrt{2},\frac{\pi}{4}) & = & -\frac{\sqrt{2}}{2} & \frac{\partial g_2}{\partial \theta}(\sqrt{2},\frac{\pi}{4}) & = & 1\end{array}$$

La fonction g est également différentiable en $(\sqrt{2},\frac{\pi}{4})$ et sa matrice Jacobienne est :

$$dg_{(\sqrt{2},\frac{\pi}{4})} = \begin{pmatrix}\frac{\sqrt{2}}{2} & -1\\ \frac{\sqrt{2}}{2} & 1\end{pmatrix}$$

(2) $\tilde{f} = e^{\cos(\theta)}\ln(\cos(\theta))$.

(3) On voit d'abord que $g(\sqrt{2},\frac{\pi}{4}) = (1,1)$. Donc

$$\begin{array}{rcl}d\tilde{f}_{(\sqrt{2},\frac{\pi}{4})} & = & df_{g(\sqrt{2},\frac{\pi}{4})} \circ dg_{(\sqrt{2},\frac{\pi}{4})}\\ & = & df_{(1,1)} \circ dg_{(\sqrt{2},\frac{\pi}{4})}\end{array}$$

et la matrice jacobienne de la différentielle de la composée est donc :

$$d\tilde{f}_{(\sqrt{2},\frac{\pi}{4})} = (e \quad -e)\begin{pmatrix}\frac{\sqrt{2}}{2} & -1\\ \frac{\sqrt{2}}{2} & 1\end{pmatrix} = (0 \quad -2e)$$

32.

(1) $\frac{\partial g}{\partial u} = e^v\frac{\partial f}{\partial x}(\star,\star) + 2uv\frac{\partial f}{\partial y}(\star,\star)$

(2) $\frac{\partial g}{\partial v} = ue^v\frac{\partial f}{\partial x}(\star,\star) + (1+u^2)\frac{\partial f}{\partial y}(\star,\star)$

où $(\star,\star) = (ue^v, v(1+u^2))$.

33.

Soit $g : \mathbb{R}^2 \to \mathbb{R} : (x,y) \to f(x^2-y^2)$. Dérivées partielles de :

$$(x,y) \to y(\partial_x g)(x,y) + x(\partial_y g)(x,y)?$$

Nommons cette fonction h.

(1) $\partial_x g(x,y) = 2xf'(x^2-y^2)$

(2) $\partial_y g(x,y) = -2yf'(x^2-y^2)$

et donc $h(x,y) = 0\,\forall (x,y) \in \mathbb{R}^2$.

34.

$h(t) = f(t, g(t^2))$.

(1) $h'(t) = \frac{\partial f}{\partial x}(\star,\star) + \frac{\partial f}{\partial y}(\star,\star)2tg'(t^2)$

(2) $h''(t) = \frac{\partial^2 f}{\partial x}(\star,\star) + 4tg'(t^2)\frac{\partial^2 f}{\partial x \partial y}(\star,\star) + 4t^2(g'(t^2))^2\frac{\partial^2 f}{\partial y^2}(\star,\star)$
$+[2g'(t^2) + 4t^2 g''(t^2)]\frac{\partial f}{\partial y}(\star,\star)$

où $(\star,\star) = (t, g(t^2))$.

35.

$$h : \mathbb{R}^2 \to \mathbb{R} : (u,v) \to f(g(ue^v), g(v)(1+u^2))^{g(u+v)}$$

Comme toujours il vaut mieux faire ce genre d'exercices prudemment. Renommons donc les diverses composantes de cette fonction.

Soit $l(u,v) = f(g(ue^v), g(v)(1+u^2))$. On a alors :

(1) $\frac{\partial l}{\partial u}(u,v) = \frac{\partial f}{\partial x}(\star,\star)g'(ue^v)e^v + \frac{\partial f}{\partial y}(\star,\star)g(v)2u$

(2) $\frac{\partial l}{\partial v}(u,v) = \frac{\partial f}{\partial x}(\star,\star)g'(ue^v)ue^v + \frac{\partial f}{\partial y}(\star,\star)g'(v)(1+u^2)$

où $(\star,\star) = (g(ue^v), g(v)(1+u^2))$.

Alors $h(u,v) = l(u,v)^{g(u+v)} = e^{g(u+v)\ln(l(u,v))}$ qui est facile à dériver :

(1) $\frac{\partial h}{\partial u} = [g'(u+v)\ln(l(u,v)) + \frac{g(u+v)}{l(u,v)}\frac{\partial l}{\partial u}(u,v)]l(u,v)^{g(u+v)}$

(2) $\frac{\partial h}{\partial v} = [g'(u+v)\ln(l(u,v)) + \frac{g(u+v)}{l(u,v)}\frac{\partial l}{\partial v}(u,v)]l(u,v)^{g(u+v)}$

26.

(1) $(x,y) \to 3x^2 + x^3y + x$.
Combinaison linéaire de fonctions continues et différentiables sur $\mathbb{R}^2$ (Exercice : prouver rigoureusement que les polynômes sont bien des fonctions continues et différentiables sur $\mathbb{R}^2$).

(2) $(x,y) \to \begin{cases} e & \text{si } xy \neq 0 \\ e^{x+y} & \text{sinon} \end{cases}$ N.B. : Il est toujours utile de se représenter le domaine de chacune des fonctions.
La première remarque est que cette fonction est clairement continue et différentiable en tout point hors de $\{xy = 0\}$ (fonction constante). Sur $\{xy = 0\}$?

(2a) Continuité :
Prenons un point dans $\{xy = 0\}$, par exemple le point $(a,0)$ (Remarquez que le cas $(0,b)$ est réglé par symétrie). Pour voir si la fonction est continue en ce point il faut voir si

$$\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0) = e^a.$$

Si on prend deux manières différentes d'aller vers $(a,0)$ ($y = 0$ puis $x = a$) on voit que si $a \neq 1$ la fonction ne peut pas être continue. Et en $(1,0)$? Si on $(x,y) \to (1,0)$ avec d'abord $y = 0$ puis $y \neq 0$ on aura regardé toutes les manières de tendre vers $(1,0)$. Or dans les deux cas les limites valent $e = f(1,0)$, ce qui prouve que la fonction est continue en $(1,0)$ (et $(0,1)$ par symétrie).

(2b) Différentiabilité :
Comme la fonction est discontinue en tout point $(a,0)$ et $(0,b)$ avec $a \neq 1$ et $b \neq 1$ elle est aussi non différentiable en chacun de ces points. Il reste donc les points $(1,0)$ et $(0,1)$. Comme toujours, nous regardons d'abord les dérivées directionnelles en $(1,0)$:

$$\frac{\partial f}{\partial u}(1,0) = \lim_{t\to 0} \frac{f(1+tu_1, tu_2) - e}{t}$$

Il y a deux possibilités : $u_2 = 0$ et donc $u = (\pm 1.0)$ ou$u_2 \neq 0$ (pourquoi ne regarde-t-on que ces deux cas ?).

(2b.ii) si $u \neq (\pm 1, 0)$.
$\frac{\partial f}{\partial u}(1,0) = \lim_{t\to 0} \frac{e-e}{t} = 0$.

(2b.ii) si $u = (\pm 1, 0)$, i.e. si $u = (1,0) = e_1$
$\frac{\partial f}{\partial u}(1,0) = \frac{\partial f}{\partial x}(1,0) = \lim_{t\to 0} \frac{f(1+t,0)-e}{t} = \lim_{t\to 0} \frac{e^{1+t}-e}{t} =^H 0$.

Conclusion :
Si f était différentiable en $(1,0)$, on aurait que sa différentielle prendrait la forme suivante :

$$\begin{aligned} df_{(1,0)}u &= \tfrac{\partial f}{\partial x}(1,0)u_1 + \tfrac{\partial f}{\partial y}(1,0)u_2 \\ &= eu_1 \quad \forall u \in \mathbb{R}^2 \end{aligned}$$

Sa différentielle satisferait également à :

$$df_{(1,0)}u = \frac{\partial f}{\partial u}(1,0) = 0 \quad \forall u \neq (\pm 1, 0) \in \mathbb{R}^2$$

Les deux propriétés étant contradictoires, la fonction f ne peut être différentiable en $(1,0)$ (ni en $(0,1)$ par symétrie).

(3) $\to \begin{cases} \frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases}$
Continue et différentiable sur $\mathbb{R} - \{y = 0\}$. Sur l'axe $y = 0$ elle n'est pas continue.

(4) $\to \begin{cases} x + ay & \text{si } x > 0 \\ x & \text{sinon} \end{cases}$ Si $a = 0$ fonction continue et différentiable sur $\mathbb{R}^2$. Si $a \neq 0$, fonction continue et différentiable partout en dehors de l'axe $x = 0$. Sur cet axe, elle est discontinue en tout point sauf en $(0,0)$ où elle est continue. Mais elle n'est pas différentiable en $(0,0)$ car toutes ses dérivées directionnelles n'y sont pas définies.

(5) $\to \begin{cases} \frac{xy^5}{x^6+y^6} & \text{si } x \neq y \\ 0 & \text{sinon} \end{cases}$ Fonction continue et différentiable partout en dehors de la droite $x = y$.
La fonction est discontinue en chacun des points de cette droite.

30.

(1) $(u,v) \to u^3 + 12u^2v - 5v^3$

(1a) $\frac{\partial f}{\partial u} = 3u^2 + 24uv$

(1b) $\frac{\partial f}{\partial v} = 12u^2 - 15v^2$

(2) $(u,v) \to f(u^2)\ln(v)$

(2a) $\frac{\partial f}{\partial u} = 2uf'(u^2)\ln(v)$

(2b) $\frac{\partial f}{\partial v} = \frac{f(u^2)}{v}$

(3) $(x,y) \to \tan(x + y^2)$

(3a) $\frac{\partial f}{\partial x} = \frac{1}{cos^2(x+y^2)}$

(3b) $\frac{\partial f}{\partial v} = \frac{2y}{cos^2(x+y^2)}$

(4) $(r, \theta) \to r^\theta$

(4a) $\frac{\partial f}{\partial r} = \theta r^{\theta-1}$

(4b) $\frac{\partial f}{\partial \theta} < ++ >= \ln(r) r^\theta$

(5) $(x, y) \to (x + 3)e^x$

(5a) $\frac{\partial f}{\partial x} = e^x(x + 4)$

(5b) $\frac{\partial f}{\partial y} = 0$

(6) $(u, v) \to \ln(f(uv))$

(6a) $\frac{\partial f}{\partial u} = \frac{vf'(uv)}{f(uv)}$

(6b) $\frac{\partial f}{\partial v} = \frac{uf'(uv)}{f(uv)}$

32.

(1) $\frac{\partial g}{\partial u} = e^v \frac{\partial f}{\partial x}(\star,\star) + 2uv\frac{\partial f}{\partial y}(\star,\star)$

(2) $\frac{\partial g}{\partial v} = ue^v \frac{\partial f}{\partial x}(\star,\star) + (1+u^2)\frac{\partial f}{\partial y}(\star,\star)$

où $\frac{\partial f}{\partial x}(\star,\star) = \frac{\partial f}{\partial x}(ue^v, v(1+u^2))$ et $\frac{\partial f}{\partial y}(\star,\star) = \frac{\partial f}{\partial y}(ue^v, v(1+u^2))$.

34. $h(t) = f(t, g(t^2))$.

(1) $h'(t) = \frac{\partial f}{\partial x}(\star,\star) + \frac{\partial f}{\partial y}(\star,\star)2tg'(t^2)$

(2) $h''(t) = \begin{array}{l} \frac{\partial^2 f}{\partial x^2}(\star,\star) + 4tg'(t^2)\frac{\partial^2 f}{\partial x \partial y}(\star,\star) + 4t^2(g'(t^2))^2\frac{\partial^2 f}{\partial y^2}(\star,\star) \\ +[2g'(t^2) + 4t^2 g''(t^2)]\frac{\partial f}{\partial y}(\star,\star) \end{array}$

où $(\star,\star) = (t, g(t^2))$.

96.16 Intégration

96.16.1 Série A

Exercice 11

(1) $\int \frac{x^3+3x+1}{x}dx = \frac{x^3}{3} + 3x + \ln(x)$

(2) $\int x^2 dx = \frac{x^3}{3}$

(3) $\int 3(x^2+1)^2 dx = \int 3x^4 + 6x^2 + 3dx = \frac{3}{5}x^5 + 2x^3 + 3x$

(4) $\int (3x^2-6x)^3(x-1)dx = \frac{1}{12}(3x^2-6x)^4$

Exercice 12

(1) $\int \sin^2(x^2+1)\cos(x^2+1)xdx = \frac{1}{6}\sin(x^2+1)^3$

(2) $\int \tan(x)dx = -\ln|\cos(x)|$

(3) $\int \frac{1}{(2+\sqrt{x})\sqrt{x}}dx = 2\ln(2+\sqrt{x})$

(4) $\int \frac{\ln(x)}{x(1-\ln^2(x)}dx = \frac{1}{2}\ln\left|1-\ln^2(x)\right|$

Travaux perso 2 ———————

1. Soient deux réels x et y vérifiant $0 < x < y$. On veut montrer que pour tout naturel $k \geqslant 2$, on a

$$0 < \sqrt[k]{y} - \sqrt[k]{x} < \sqrt[k]{y-x}.$$

La première inégalité vient de l'inégalité $x < y$ élevée à la puissance $\frac{1}{k}$.

On peut ré-écrire la deuxième, sachant que $x > 0$, en divisant par $\sqrt[k]{x}$ pour obtenir

$$\sqrt[k]{\frac{y}{x}} - 1 - \sqrt[k]{\frac{y}{x} - 1} < 0 \quad \text{avec } \frac{x}{y} > 1 \tag{96.573}$$

ce qui s'écrit encore $f(t) < 0$ en posant $f(t) \overset{\text{def}}{=} \sqrt[k]{t} - \sqrt[k]{t-1} - 1$. On peut alors étudier la fonction f. Étant donné que $f(1) = 0$, il suffirait que f soit strictement décroissante sur $]1;\infty[$ pour qu'on ait l'inégalité voulue, à savoir $f(t) < 0$ dès que $t > 1$.

Pour le montrer, on voit que

$$f'(t) = \frac{1}{k}\left(t^{\frac{1-k}{k}} - (t-1)^{\frac{1-k}{k}}\right)$$

d'où on tire les équivalences suivantes

$$f'(t) < 0 \tag{96.574}$$
$$\iff \quad t^{\frac{1-k}{k}} < (t-1)^{\frac{1-k}{k}} \tag{96.575}$$
$$\iff \quad t^{1-k} < (t-1)^{1-k} \tag{96.576}$$
$$\iff \quad t > t-1 \tag{96.577}$$
$$\iff \quad 0 > -1 \tag{96.578}$$

où la dernière inégalité est manifestement vraie, ce qui prouve la première inégalité et achève l'exercice.

2.

Exercice 1

(1) Par exemple, $B(x,r)$ avec $x \in \mathbb{R}^n$ et $r > 0$.

(2) On utilise la densité de $\mathbb{Q}$ dans $\mathbb{R}$ pour voir que $B(q,r)$ ($q \in \mathbb{Q}^n$ et $r > 0$) est également une base.

On observe ensuite que seuls les r « petits » sont utiles, donc on se restreint aux boules de la forme $B(q,1/n)$ ($q \in \mathbb{Q}^n$ et $n \in \mathbb{N}_0$). Cet ensemble de boules est une base dénombrable de la topologie usuelle sur $\mathbb{R}^n$.

96.17 Équations différentielles du premier ordre

96.17.1 Exercices

Exercice 104. En remplaçant y dans l'équation par $f(t) = t^4 e^{2t}$, l'équation devient, après simplifications

$$\left((b + 4\,a + 4)\,t^4 + (8\,a + 16)\,t^3\right) = 0$$

qui doit être vraie pour toute valeur de t. Un polynôme est nul si et seulement chacun des coefficients est nul, donc l'équation se ramène au système

$$\begin{cases} b + 4a + 4 = 0 \\ 8a + 16 = 0 \end{cases}$$

dont l'unique est solution est $(a,b) = (-2,4)$.

Exercice 105.

(1) $y(t) = \ln(\frac{t^4}{4} + \frac{t^2}{2} + K)$

(2) $y(t) = \tan(t + K)$

(3) L'intégration directe donne la relation

$$y(t) + e^{y(t)} = \sin(t) + K \tag{96.579}$$

eqyplusexpy

et il faut encore justifier l'éventuelle existence et/ou unicité d'un tel $y(t)$. Pour ce faire, nous aurons besoin du lemme suivant.

Lemme 96.3.
La fonction $f : \mathbb{R} \to \mathbb{R} : z \mapsto z + e^z$ est une bijection.

Démonstration. La dérivée de f est strictement positive, donc f est strictement croissante, donc f est injective.
Par ailleurs, étant donné que

$$\lim_{z\to-\infty} f(z) = -\infty \quad \text{et} \quad \lim_{z\to+\infty} f(z) = +\infty,$$

le théorème des valeurs intermédiaires (qui affirme que l'image d'une fonction continue est un intervalle) dit que l'image de f contient n'importe quel intervalle arbitrairement grand, donc contient $\mathbb{R}$ entier. Ceci prouve la surjectivité de f. □

Ce lemme montre qu'il existe une (unique) fonction $g : \mathbb{R} \to \mathbb{R}$ qui est réciproque de f. On en déduit, en récrivant l'équation (96.579) sous la forme

$$f(y(t)) = \sin(t) + K$$

et en lui appliquant g, que $y(t) = g(\sin(t) + K)$ existe et est univoquement définie.

(4) L'équation $y' = y^2$ pourrait poser un problème pour trouver des solutions y pour lesquelles il existe t tel que $y(t) = 0$, car on ne peut alors pas diviser par y^2.

On commence par remarquer que $y(t) = 0$ (fonction identiquement nulle) est une solution de l'équation.

Si y ne s'annule pas sur un certain intervalle fixé, l'équation s'y écrit

$$\frac{y'}{y^2} = 1$$

dont les solutions sont de la forme $y(t) = \frac{-1}{t+K}$ (où K est une constante). Une telle solution ne peut pas tendre vers 0, dès lors une solution définie sur un intervalle est soit identiquement nulle, soit ne s'annule pas du tout.

Si on ne s'intéresse qu'à des fonctions définies sur des intervalles, il n'y a donc que ces solutions : soit $y(t) = 0$, soit $y(t)$ est de la forme $\frac{-1}{t+K}$ pour une certaine constante K.

La solution générale sur un domaine quelconque s'obtient en prenant l'union sur des intervalles disjoints de solutions du type précédent.

(5) L'équation $y' = y^{\frac{1}{3}}$ pose le même problème que l'équation précédente, mais la solution est différente.

On remarque à nouveau que la fonction nulle est solution, et on s'intéresse aux autres solutions.

Si y est une solution qui ne s'annule pas sur un certain intervalle fixé, elle satisfait

$$\frac{y'}{y^{1/3}} = 1$$

et donc est de la forme $y(t) = \pm(\frac{2x}{3} + K)^{2/3}$. Une telle solution est définie pour $x \geqslant {}^{3K}/_{2}$ et tend vers 0 lorsque $x \to {}^{3K}/_{2}$, et donc peut se recoller avec une solution nulle « sur le bord gauche de l'intervalle » (sur le bord droit, la solution tend vers $\pm\infty$).

La solution générale sur un intervalle est donc une solution de la forme

$$y(t) = \begin{cases} 0 & \text{si } x \leqslant {}^{3K}/_{2} \\ \pm(\frac{2x}{3} + K)^{2/3} & \text{si } x > {}^{3K}/_{2} \end{cases}$$

pour un certaine constante K, ou alors $y(t)$ est identiquement nulle.

(6) L'équation est équivalente à

$$\frac{yy'}{y^2+1} = -\sin(t)$$

ce qu'on intègre pour obtenir

$$\frac{1}{2}\ln(y^2+1) = -\cos(t) + K$$

c'est-à-dire $y^2 = -1 + Ke^{-2\cos(t)}$ pour une certaine constante (positive) K. Selon la valeur de K, ces solutions sont définies ou non sur tout $\mathbb{R}$:

(6a) Si $K > e^2$, le membre de droite est strictement positif pour tout t et on peut en prendre la racine (solution sur $\mathbb{R}$)

(6b) Si $K < e^2$, le membre de droite est négatif pour certaines valeurs de t et on ne peut pas en prendre la racine.

(6c) Si $K = e^2$, le choix d'une racine du membre de droite (positive ou négative) donne une fonction qui n'est pas dérivable aux points où elle est nulle (car la racine n'est pas dérivable en ces points). Par contre, si on change de choix de signe pour la racine à chaque fois que le membre de droite s'annule, la fonction obtenue est dérivable.

Ces trois cas seraient plus clairs sur une illustration.

Exercice 106. Il s'agit ici de reprendre les solutions générales de l'exercice ci-dessus en sélectionnant la ou les solutions qui satisfont au problème de Cauchy. Dans les cas agréables cela revient simplement à déterminer la constante. Dans les autres cas, il faut vérifier l'existence et l'unicité de la solution.

Exercice 116. Dans chacun de ces exercices, il s'agit d'intégrer la fonction donnée sur un domaine précisé.

(1) Le domaine d'intégration est un rectangle, le choix des bornes est donc simple :

$$\begin{aligned}\int_0^2 \left(\int_0^1 (4-x^2-y^2)dx\right) dy &= \int_0^2 \left[4x - \frac{x^3}{3} - y^2 x\right]_{x=0}^{x=1} dy \\ &= \int_0^2 4 - \frac{1}{3} - y^2 dy \\ &= \left[\left(4-\frac{1}{3}\right)y - \frac{y^3}{3}\right]_0^2 \\ &= 8 - \frac{10}{3}\end{aligned}$$

(2)

96.18 Théorème de la fonction implicite

96.18.1 Exercices

Exercice 129.

(1) Une telle fonction Z doit vérifier $F(1,1,Z(1,1)) = 0$, donc en particulier, en notant $z_0 = Z(1,1)$, il faut $z_0 + \ln(z_0) = 1$. Une solution évidente est $z_0 = 1$.
Montrons que cette solution est unique : la fonction auxiliaire $g : \mathbb{R}_0^+ \to \mathbb{R} : z \mapsto z + \ln(z)$ est strictement croissante (sa dérivée est strictement positive sur son domaine) et en particulier injective. L'équation en z_0 se réécrit sous la forme $g(z_0) = 1$, et l'injectivité nous assure l'unicité de la solution.
Au point $(x_0, y_0, z_0) = (1,1,1)$ on peut appliquer le théorème de la fonction implicite puisque $\frac{\partial F}{\partial z}(1,1,1) = 1 + 1 = 2 \neq 0$, et donc on a un voisinage U de $(1,1)$ et une unique fonction $Z : U \to \mathbb{R}$ satisfaisant à la condition énoncée.

(2) Notons $Z = Z(x,y)$ pour la simplicité. Pour tout (x,y) dans U, nous avons donc $Z + \ln(Z) - xy = 0$. En particulier on peut dériver cette identité par rapport à x et à y, d'où

$$\begin{aligned}\frac{\partial Z}{\partial x} + \frac{1}{Z}\frac{\partial Z}{\partial x} - y = 0 \\ \frac{\partial Z}{\partial y} + \frac{1}{Z}\frac{\partial Z}{\partial y} - x = 0\end{aligned}$$

pour tout $(x,y) \in U$, et on en tire (on note ∂_x au lieu de $\frac{\partial}{\partial x}$)

$$\partial_x Z = \frac{y}{1+\frac{1}{Z}} = \frac{yZ}{1+Z} \quad \text{et} \quad \partial_y Z = \frac{x}{1+\frac{1}{Z}} = \frac{xZ}{1+Z}$$

On peut enfin dériver l'une ou l'autre de ces égalités pour obtenir les dérivées secondes, en remplaçant ensuite les occurrences de $\partial_x Z$ et $\partial_y Z$ par leur expression ci-dessus, par exemple

$$\begin{aligned}\partial^2_{xy} Z &= \frac{(x\partial_x Z + Z)(1+Z) - xZ\partial_x Z}{(1+Z)^2} \\ &= \frac{(xyZ + Z(1+Z)^2)}{(1+Z)^3}\end{aligned}$$

où ∂^2_{xy} désigne la dérivée partielle seconde par rapport à y puis par rapport à x.

Exercice 130. Si $x = 1$, l'équation nous donne $1 = y$, donc la fonction doit vérifier $Y(1) = 1$. Montrons d'abord qu'une telle fonction existe : on considère

$$F : \mathbb{R}_0^+ \to \mathbb{R}_0^+ : (x, y) \mapsto x^y - y^x$$

et on vérifie que $\partial_y F(1, 1) = -1 \neq 0$. Le théorème de la fonction implicite s'applique, et fournit effectivement un voisinage U de 1 et une unique fonction $Y : U \subset \mathbb{R} \to \mathbb{R}$ vérifiant $Y(1) = 1$ et vérifiant l'équation $x^{Y(x)} - Y(x)^x$.

Pour calculer la dérivée, on peut ré-écrire l'équation (en notant $Y = Y(x)$ pour simplifier la notation) sous la forme

$$Y \ln(x) = x \ln(Y)$$

et en dérivant par rapport à x on obtient

$$Y' \ln(x) + Y \frac{1}{x} = \ln(Y) + x \frac{Y'}{Y}$$

d'où on tire en réarrangeant les termes

$$Y'(x) = \frac{Y}{x} \frac{x \ln(Y) - Y}{Y \ln(x) - x} = \frac{\ln(Y) - \frac{Y}{x}}{\ln(x) - \frac{x}{Y}}$$

Exercice 131. L'équation implicite pour $(x, y) = (0, 0)$ devient $ze^z = 0$ dont l'unique solution est $z = 0$, donc une telle fonction Z doit vérifier $Z(0, 0) = 0$. Pour vérifier l'existence et l'unicité de la fonction Z, on considère la fonction

$$F : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \mapsto ze^z - x - y.$$

On calcule $\partial_z F(0, 0, 0) = 1 \neq 0$, de sorte que le théorème de la fonction implicite s'applique et fournit une unique fonction Z telle que demandée.

Pour écrire le polynôme de Taylor il suffit de calculer les dérivées de Z, ce qu'on fait en utilisant la relation $Ze^Z = x + y$:

$$\partial_x Z(1 + Z)e^Z = 1 \Rightarrow \partial_x Z = \frac{e^{-Z}}{1 + Z}$$
$$\partial_y Z(1 + Z)e^Z = 1 \Rightarrow \partial_y Z = \frac{e^{-Z}}{1 + Z}$$

où on note comme toujours $Z = Z(x, y)$ pour simplifier l'écriture. On calcule également les dérivées secondes :

$$\partial^2_{xx} Z = \partial^2_{yx} Z = \partial^2_{yy} Z = -\frac{Z + 2}{(1 + Z)^3} e^{-2Z}$$

et donc le polynôme de Taylor à l'ordre 2 s'écrit

$$Z(x, y) = x + y - x^2 - 2xy - y^2 + o(\|(x, y)\|^2)$$

Exercice 132. Pour $x = 3/4$, l'équation $F(3/4, y, z) = 0$ devient le système

$$\begin{cases} \frac{9}{16} + y^2 + z^2 - 1 = 0 \\ \frac{9}{16} + y^2 - \frac{3}{4} = 0 \end{cases}$$

qui a exactement les quatre solutions $(y, z) = (\pm\frac{\sqrt{3}}{4}, \pm\frac{1}{2})$.

Le jacobien partiel de F par rapport à (y, z) est donné par

$$\begin{vmatrix} 2y & 2z \\ 2y & 0 \end{vmatrix} = -4yz \qquad (*)$$

exo132-jacobienpartiel

et donc est non-nul en chacun des quatre points $(\pm\frac{\sqrt{3}}{4}, \pm\frac{1}{2})$, ce qui permet d'appliquer le théorème de la fonction implicite. Ceci prouve l'existence des quatre fonctions φ demandées, correspondant chacune à un des points ci-dessus.

Pour calculer les dérivées, on sait que ces fonctions vérifient

$$\begin{cases} x^2 + Y(x)^2 + Z(x)^2 = 1 \\ x^2 + Y(x)^2 = x \end{cases}$$

pour tout x dans un voisinage de $3/4$. En particulier on peut dériver ces deux équations pour obtenir (on note $Y = Y(x)$ et $Z = Z(x)$ pour alléger la notation) :

$$\begin{cases} 2x + 2YY' + 2ZZ' = 0 \\ 2x + 2YY' = 1 \end{cases}$$

d'où on tire

$$\begin{cases} Z' = \frac{-1}{2Z} \\ Y' = \frac{1-2x}{2Y} \end{cases}$$

où on remarque que la division par Y et par Z est bien définie si x est assez proche de $3/4$ puisque le jacobien partiel (*) est non nul.

Remarque 96.4.
La formule donnant la dérivée ne dépend pas explicitement du point autour duquel on fait le calcul, mais dépend bien sûr encore de la valeur de Z en ce point.

Exercice 133. Étant donnée la relation, on vérifie que la fonction Y définie implicitement par

$$e^{yx} - 1 = x^2 + y$$

doit satisfaire à $Y(0) = 0$. Une première tentative montre que la limite demandée est donc du type indéterminé « $\frac{0}{0}$ ». Afin d'appliquer la *règle de L'hospital* on veut d'abord vérifier que la fonction Y est bien dérivable autour de 0.

Pour ce faire, on considère

$$F : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto e^{xy} - 1 - x^2 - y$$

et on a $\partial_y F(0, 0) = -1 \neq 0$, donc le théorème de la fonction implicite s'applique et assure l'existence d'une fonction Y de classe C^∞ autour de 0 vérifiant

$$e^{xY} - 1 = x^2 + Y$$

où on note $Y = Y(x)$ pour alléger la notation. En particulier sa dérivée doit satisfaire à l'équation

$$e^{xY}(Y + xY') = 2x + Y'$$

et donc, pour $x = 0$ on obtient

$$Y(0) = Y'(0)$$

ce qui montre que la dérivée s'annule en 0.

La limite devient donc

$$\lim_{x\to 0} \frac{Y(x)}{\cos(x) - 1} = \lim_{x\to 0} \frac{Y'(x)}{-\sin(x)}$$

et est à nouveau du type indéterminé « $\frac{0}{0}$ ».

Dérivons à nouveau l'équation satisfaite par Y' pour obtenir

$$e^{xY}(Y + xY')^2 + e^{xY}(Y' + Y' + xY'') = 2 + Y''$$

ce qui, pour $x = 0$, donne

$$Y(0)^2 + 2Y'(0) = 2 + Y''(0)$$

et donc $Y''(0) = -2$ puisque $Y(0) = Y'(0) = 0$.

La limite devient donc

$$\lim_{x\to 0} \frac{Y(x)}{\cos(x) - 1} = \lim_{x\to 0} \frac{Y''(x)}{-\cos(x)} = 2$$

et la réponse attendue est donc 2.

Exercice 134. Au voisinage du point $(0, 1)$ la courbe peut s'écrire sous la forme $y = Y(x)$ par application du théorème de la fonction implicite (le vérifier !). La tangente à la courbe est alors donnée par l'équation

$$y - 1 = Y'(0)(x - 0)$$

ce qui implique de calculer $Y'(0)$.

Par définition $Y(x)$ vérifie $Y(x)^2 + \sin(xY(x)) = 1$ et donc sa dérivée vérifie

$$2Y(x)Y'(x) + \cos(xY(x))(Y(x) + xY'(x)) = 0$$

et donc en $x = 0$, on obtient

$$2Y(0)Y'(0) + Y(0) = 0$$

ce qui donne $Y'(0) = -1/2$ puisque sachant que $Y(0) = 1$.

Exercice 135. Nous savons que pour une surface donnée sous forme implicite $F(x, y, z) = k$, le vecteur gradient ∇F en un point de la surface est normal à cette surface.

Dans le cadre de cet exercice, on est donc ramené à chercher les points sur la surface dont le gradient est un multiple de $(1, -2, 2)$ (qui est le vecteur normal au plan donné).

Le gradient au point (x, y, z) est donné par $(8x, 32y, 16z)$, et on veut qu'il existe $\lambda \neq 0$ tel que $(8x, 32y, 16z) = (\lambda, -2\lambda, 2\lambda)$. Cette condition combinée à la condition d'appartenance à la surface fournit une équation pour λ qu'il suffit de résoudre.

La fin de l'exercice dépend de la manière de compléter l'énoncé, puisqu'il faut expliciter l'équation des plans tangents.

Exercice 136.

(1) Étant donné que $M = F^{-1}(0, 0)$ et que $\{(0, 0)\}$ est un ensemble fermé (ne contient qu'un seul point !), on en déduit que M est l'image réciproque par une fonction continue d'un fermé, donc M est un ensemble fermé (vérifier !). Par ailleurs, M est complètement contenue dans la sphère de rayon 1, donc est bornée. Ces deux propriétés fournissent la compacité.

L'ensemble M est donné sous forme implicite par l'annulation de deux fonctions, dont les gradients sont $(1, 1, 1)$ et $(2x, 2y, 2z)$. Ces deux vecteurs linéairement dépendant si et seulement si $x = y = z$, ce qui n'est pas possible pour un point de M. On en déduit que les gradients sont indépendants sur M, et donc que M est une variété C^1 de dimension $3 - 2 = 1$.

En fait, c'est l'intersection d'une sphère et d'un plan, c'est donc un cercle.

(2) Pour $y = 0$, on observe qu'il faut que $X(0) + Z(0) = 0$ et $X(0)^2 + Z(0)^2 = 1$. On en déduit que $(X(0), Z(0))$ vaut soit $\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$, soit $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. Pour chacun de ces points le théorème de la fonction implicite s'applique et fournit donc deux paires de fonctions (X_1, Z_1) (avec $X_1(0) = \frac{\sqrt{2}}{2}$) et (X_2, Z_2) (avec $X_2(0) = -\frac{\sqrt{2}}{2}$).

(3) La meilleure approximation polynomiale de degré 1 est donnée par le polynôme de Mc Laurin d'ordre 1, donc on calcule la dérivée première des fonctions X_1 et X_2.

En dérivant les équations qui définissent X_1 et Z_1 (et X_2 et Z_2 également, ce sont les mêmes !), on obtient les relations

$$\begin{cases} X_1' + 1 + Z_1' = 0 \\ 2X_1X_1' + 2y + 2Z_1Z_1' = 0 \end{cases}$$

ce qui, pour $y = 0$, fournit $X_1'(0) = Z_1'(0) = -1/2$. Et en mettant un indice $_2$ partout, on obtient la même chose $X_2'(0) = Z_2'(0) = -1/2$.
Le polynôme de Mc Laurin pour X_1 et X_2 s'écrit donc

$$X_1(y) = \frac{\sqrt{2}}{2} - \frac{1}{2}y + o(|y|)$$
$$X_2(y) = -\frac{\sqrt{2}}{2} - \frac{1}{2}y + o(|y|)$$

Exercice 137.

(1) Considérons l'application $F : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 : (t, v) \mapsto v - \varphi_t(v)$. On sait que $F(0, v_0) = 0$ par définition de v_0. On sait également que pour $t = 0$, la différentielle de l'application partielle $\mathbb{R}^2 \to \mathbb{R}^2 : v \mapsto F(v)$ vaut $\mathrm{Id} - d\varphi_0$ et est donc inversible par hypothèse sur le spectre. On en déduit que le théorème de la fonction implicite s'applique, et qu'il existe un voisinage $]-\epsilon, \epsilon[$ de $t = 0$, un voisinage U de $v = v_0$ et une unique application

$$V :]-\epsilon, \epsilon[\to U$$

telle que $F(t, V(t)) = 0$ pour tout $t \in]-\epsilon, \epsilon[$. C'est-à-dire $V(t) = \varphi_t(V(t))$ pour $t \in]-\epsilon, \epsilon[$, et donc $V(t)$ est un point fixe de φ_t.

(2) Les applications $\varphi_t : (x, y) \mapsto (x + t, y)$ n'ont pas de point fixe, sauf pour $t = 0$. A titre d'information, on dit que c'est une action (la variable t « agit » puisqu'elle translate vers la droite) *libre* (sans point fixe autre que $t = 0$) de la droite $\mathbb{R}$ sur le plan $\mathbb{R}^2$.

96.19 Intégrales curvilignes

96.19.1 Exercices

Exercice 144′

(1) Un paramétrage est donné, il reste à intégrer

$$\int_0^{2\pi} \left\|\gamma'(t)\right\| dt$$

où $\gamma'(t) = (a(1 - \cos(t)), a\sin(t))$, c'est-à-dire

$$\int_0^{2\pi} a\sqrt{2 - 2\cos(t)}dt = 8a$$

où on a utilisé l'égalité trigonométrique

$$1 - \cos(t) = 2\sin^2\left(\frac{t}{2}\right).$$

(2) On sait qu'on peut paramétriser cet astroïde par

$$\begin{cases} \sqrt[3]{x} = \sqrt[3]{a}\cos(t) \\ \sqrt[3]{y} = \sqrt[3]{a}\sin(t) \end{cases}$$

ce qui donne le chemin

$$\gamma(t) = (a\cos^3(t), a\sin^3(t)) \Rightarrow \gamma'(t) = (-3a\cos^2(t)\sin(t), 3a\sin^2(t)\cos(t))$$

et l'intégrale devient, grace aux relations trigonométriques,

$$\int_0^{2\pi} 3a\sqrt{\sin^2(t)\cos^2(t)}dt = \frac{3a}{2}\int_0^{2\pi} |\sin(2t)|\, dt$$

où il faut encore faire attention au signe. Par périodicité et par symétrie, on se ramène à

$$6a\int_0^{\frac{\pi}{2}} \sin(2t)dt = 6a$$

(3) Le chemin peut être paramétrisé de la manière suivante

$$\gamma(t) = \begin{cases} (t+1,0) & -1 \leqslant t \leqslant 0 \\ (1-t,t) & 0 \leqslant t \leqslant 1 \\ (0,2-t) & 1 \leqslant t \leqslant 2 \end{cases}$$

c'est un chemin C^1 par morceaux. On a

$$\gamma'(t) = \begin{cases} (1,0) & -1 < t < 0 \\ (-1,1) & 0 < t < 1 \\ (0,-1) & 1 < t < 2 \end{cases} \quad \text{et donc} \quad \|\gamma'(t)\| = \begin{cases} 1 & -1 < t < 0 \\ \sqrt{2} & 0 < t < 1 \\ 1 & 1 < t < 2 \end{cases}$$

et l'intégrale devient donc

$$\int_{-1}^{0} (t+1)dt + \int_{0}^{1} \sqrt{2}dt + \int_{1}^{2} (2-t)\underset{\cdot}{t} = 1 + \sqrt{2}$$

Exercice 143 = 145′.

(1) Il faut calculer

$$\int_0^{\frac{1}{2}} \sqrt{1 + \frac{4x^2}{(1-x^2)^2}} dx$$

(2) Calculer

$$\int_0^5 \sqrt{1 + \frac{9}{4}x} dx$$

(3) Calculer

$$\int_0^{\frac{\pi}{4}} \sqrt{1 + \tan^2(x)}\, dx = \int_0^{\frac{\pi}{4}} \frac{1}{\cos(x)} dx$$

Exercice 146′

(1) On paramétrise ce cercle de la façon usuelle mais en prenant attention au sens

$$\gamma(t) = (\cos(-t), \sin(-t)) = (\cos(t), -\sin(t)) \qquad t \in [0, 2\pi]$$

d'où on tire $\gamma'(t) = (-\sin(t), -\cos(t))$ et l'intégrale devient

$$\int_0^{2\pi} (-\sin^3(t) - \cos^3(t))dt = 0.$$

L'intégrale est nulle parce que ces deux fonctions ($\sin^3$ et $\cos^3$) possèdent un centre de symétrie, et qu'on les intègre sur une période complète.

(2) Le chemin se paramétrise par

$$\gamma(t) = (\cos(t), \sin(t)) \Rightarrow \gamma'(t) = (-\sin(t), \cos(t))$$

et donc l'intégrale devient

$$\int_0^{2\pi} \big\langle G(\gamma(t)), \gamma'(t) \big\rangle\, dt = \int_0^{2\pi} \big(-\sin^3(t) + \cos^3(t)\big)\, dt = 0$$

Cette intégrale est nulle pour les même raisons que ci-dessus.

(3) Le paramétrage est donnée, et on a

$$\gamma'(t) = (-a\sin(t), a\cos(t), b)$$

donc l'intégrale devient

$$\int_0^{2\pi} \Big[\big(a\sin(t) - bt\big)(-a\sin(t)) + \big(bt - a\sin(t)\big)a\cos(t) + ab(\cos(t) - \sin(t))\Big]dt \tag{96.580}$$

et vaut $-\pi a(a+2b)$ après calculs.

(4) Sur le chemin donné, l'intégrale vaut

$$\int_\gamma -xdx - ydy - zdz = \int_\gamma -\frac{1}{2}df$$

où $f(x,y,z) = x^2+y^2+z^2$. Cette intégrale est donc nulle puisque le chemin γ est fermé (c'est un cercle).

On peut même aller plus loin : sur la sphère $f \equiv 1$ donc $df = 0$.

Exercice 144 = 147′. Ayant $y = x^3+x^2+x+1$ on a $dy = (3x^2+2x+1)dx$ donc il faut calculer

$$\int_0^1 (x(3x^2+2x+1) + x^3 + x^2 + x + 1)dx$$

Exercice 145 = 148′. Il faut calculer

(1)
$$\int_0^2 (x^2 + \frac{x^2}{2})dx$$

(2)
$$\int_0^2 (\frac{x^3}{2} + \frac{x^3}{2})dx$$

(3)
$$\int_0^1 (16y^4 + 4y^4)dy$$

(4) La ligne brisée est un peu inutile...

Ces quatres résultats sont identiques, ce qui laisse penser qu'en réalité la 1-forme $\omega = 2xydx + x^2dy$ est une forme exacte. En effet, on vérifie aisément que $\omega = d(x^2y)$.

Exercice 146 = 149′. On peut paramétriser l'hélice par les équations

$$\begin{cases} x = \cos(2\pi t) \\ y = \sin(2\pi t) \\ z = t \end{cases} \qquad t \in [0;1]$$

et l'intégrale devient

$$\int_0^1 \left(-2\pi\sin^2(2\pi t) + 2\pi\cos^2(2\pi t) + t^2\right)dt = \frac{1}{3}$$

On pouvait aussi remarquer que

$$ydx + xdy + z^2dz = df \quad \text{où } f(x,y,z) = xy + \frac{1}{3}z^3$$

et donc l'intégrale vaut bien $f(1,0,1) - f(1,0,0) = 1/3$.

96.20 Intégrales de surface

Exercice 147.

(1) La sphère est paramétrisée en coordonnées sphériques (avec un rayon r constant) par

$$\begin{cases} x = r\cos(u)\sin(v) \\ y = r\sin(u)\sin(v) \\ z = r\cos(v) \end{cases} \qquad \text{avec} \quad \begin{cases} u \in [0; 2\pi[\\ v \in [0; \pi[\end{cases}$$

c'est-à-dire le paramétrage est

$$F : [0; 2\pi[\times [0; \pi[\to \mathbb{R}^3 : (u, v) \mapsto (r\cos(u)\sin(v), r\sin(u)\sin(v), r\cos(v))$$

et on calcule que l'élément de surface vaut

$$\left\| \frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v} \right\| dudv = r^2 \sin(v) dudv$$

Dès lors, la surface de la sphère vaut

$$\int_0^{\pi} \int_0^{2\pi} r^2 \sin(v) dudv = 2r^2\pi \Big[-\cos(v) \Big]_0^{\pi} = 4\pi r^2$$

(2) En coordonnées cylindriques (ρ, θ, z), le cylindre plein de rayon r tel qu'il est décrit a pour équation

$$\rho \leqslant r\cos(\theta)$$

et la sphère s'écrit $\rho^2 + z^2 = r^2$. On peut donc paramétriser la surface demandée via

$$\begin{aligned} F_1 &\equiv x = \rho\cos\theta \\ F_2 &\equiv y = \rho\sin\theta \\ F_3 &\equiv z = \pm\sqrt{r^2 - \rho^2} \end{aligned}$$

où le signe $\pm$ indique qu'il y a deux morceaux de surface.
On peut calculer le produit vectoriel

$$\frac{\partial F}{\partial \theta} \wedge \frac{\partial F}{\partial \rho} = \begin{vmatrix} \overrightarrow{e_x} & \overrightarrow{e_y} & \overrightarrow{e_z} \\ -\rho\sin\theta & \rho\cos\theta & 0 \\ \cos\theta & \sin\theta & \pm\frac{-\rho}{r^2-\rho^2} \end{vmatrix} = (\pm\frac{-\rho^2\cos\theta}{r^2-\rho^2}, \pm\frac{-\rho^2\sin\theta}{r^2-\rho^2}, -\rho)$$

dont la norme donne l'élément de surface

$$d\sigma = \frac{r\rho}{\sqrt{r^2-\rho^2}} d\rho d\theta$$

ce qui permet de calculer l'intégrale sur l'un des morceaux :

$$\begin{aligned} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{r\cos\theta} \frac{r\rho}{\sqrt{r^2-\rho^2}} d\rho d\theta &= r \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \Big[-\sqrt{r^2-\rho^2} \Big]_0^{r\cos\theta} \theta \\ &= r \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (-r\,|\sin\theta| + r) \\ &= d\theta = \pi r^2 - 2r^2 = r^2(\pi - 2) \end{aligned}$$

(3) Le morceau de cône se accepte le paramétrage suivant en coordonnées cylindriques :

$$\rho = |z| \quad 0 < z < b$$

et puisque $z > 0$, on a $z = \rho$. On peut donc expliciter le changement de coordonnées :

$$\begin{cases} F_1 \equiv x = \rho\cos\theta \\ F_2 \equiv y = \rho\cos\theta \\ F_3 \equiv \rho = \rho \end{cases} \qquad \text{avec } 0 < \rho < b$$

l'élément de surface peut alors se calculer comme suit

$$\frac{\partial F}{\partial \rho} \wedge \frac{\partial F}{\partial \theta} = \begin{vmatrix} \overrightarrow{e_x} & \overrightarrow{e_y} & \overrightarrow{e_z} \\ \cos\theta & \sin\theta & 1 \\ -\rho\sin\theta & \rho\cos\theta & 0 \end{vmatrix} = (-\rho\cos\theta, -\rho\sin\theta, \rho),$$

donc $d\sigma = |\rho|\sqrt{2} = \rho\sqrt{2}$, et l'intégrale devient

$$\int_0^b \int_0^{2\pi} \rho^2\sqrt{2} d\theta\rho = \frac{2\sqrt{2}\pi}{3} b^3$$

(4) La sphère se paramétrise en coordonnées sphériques par

$$\begin{cases} x = \cos u \sin v \\ y = \sin u \sin v \\ z = \cos v \end{cases} \qquad \text{avec } \begin{cases} u \in [0, 2\pi[\\ v \in [0, \pi[\end{cases}$$

dont l'élément de volume a déjà été calculé et vaut $\sin(v)du$ẹ.
L'équation du cône devient

$$\cos^2 v = \sin^2 v$$

c'est-à-dire $\sin v = |\cos v|$ (car $\sin v \geqslant 0$), donc les choix possibles sont

$$v = \frac{\pi}{4} \qquad \text{ou} \qquad v = \frac{3\pi}{4}$$

Étant donné qu'on veut être *dans* le cône, il faut

$$v \in \left[0, \frac{\pi}{4}\right] \cup \left[\frac{3\pi}{4}, \pi\right[$$

et donc l'intégrale devient

$$\int_0^{2\pi} \left(\int_0^{\frac{\pi}{4}} \sin(v)dv + \int_{\frac{3\pi}{4}}^{\pi} \sin(v)dv \right) du = \int_0^{2\pi} (2 - \sqrt{2})du = 2\pi(2 - \sqrt{2}).$$

(5) En coupant le cylindre le long de la droite

$$\begin{cases} x = 0 \\ y = 1 \end{cases}$$

on remarque que la surface limitée par l'hélice

$$\begin{cases} x = \sin t \\ y = \cos t \\ z = \frac{t}{2\pi} \end{cases}$$

est un triangle dont la base est la circonférence du cylindre et la hauteur est le pas de l'hélice, c'est-à-dire son aire vaut $\frac{1}{2}(2\pi \cdot 1) = \pi$.

Exercice 150.

(1) Si $P(x,y) = 2(x^2+y^2)$ et $Q(x,y) = (x+y)^2$, l'intégrale demandée rentre dans les conditions du théorème de Green, puisque le domaine est le périmètre γ d'un triangle plein T, dont le bord admet clairement en chaque point un vecteur normal extérieur. Dès lors

$$\begin{aligned}\int_\gamma 2(x^2+y^2)dx + (x+y)^2dy &= \iint_T (2x-2y)dxdy \\ &= \int_1^2 \int_1^y 2(x-y)dxdy \\ &= \int_1^2 \Big[x^2 - 2xy\Big]_1^y dy \\ &= \int_1^2 (-y^2 - 1 + 2y)dy = -\frac{1}{3}\end{aligned}$$

(2) Le théorème de Green s'applique, et il s'agit donc d'intégrer

$$-\iint_S 4xydxdy$$

où S est le disque de rayon R. Par passage en coordonnées polaires, on trouve

$$-\int_0^{2\pi}\int_0^R 4\rho^3 \sin\theta\cos\theta d\rho d\theta = -R^4\frac{1}{2}\Big[\sin^2(\theta)\Big]_0^{2\pi} = 0$$

ce qu'on aurait pu deviner en utilisant les symétries du problème.

(3) D'après le théorème de Green, on a

$$\int_\gamma dx + xdy = \int_0^1 \int_{y^2}^{\sqrt{y}} dxdy = \frac{1}{3}$$

(4) Tout le monde sait que l'ellipse E a pour aire πab. On peut aussi le calculer en utilisant le théorème de Green :

$$\iint_E dxdy = \int_\gamma (dx + xdy)$$

mais c'est plus compliqué.

(5) Par le théorème de Green, cette intégrale vaut

$$-\iint_D x^2+y^2 = -\int_0^{2\pi}\int_0^1 \rho^3 d\rho d\theta = -\frac{\pi}{2}$$

où D est le disque unité. Ne pas oublier le signe qui vient du fait que l'on demande de tourner dans le sens horloger, alors que $\int_{\partial D}\omega$ n'est égal à $\int_\gamma \omega$ que lorsque γ parcours ∂D dans le sens trigonométrique, c'est-à-dire dans l'autre sens.

Exercice 151.

(1) Si on applique le théorème de Stokes dans le plan $z=0$, cela revient en fait à appliquer le théorème de Green, en effet le rotationnel vaut $(0,0,2x-2y)$ et le vecteur normal vaut $(0,0,1)$. L'intégrale devient donc une intégrale sur le disque D unité dans le plan $z=0$

$$\iint_D \nabla\times G \cdot dS = \iint_D (2x-2y)dxdy$$

et elle vaut zéro par les symétries (ou par calcul).

Si on applique le théorème de Stokes à la demi-sphère unité, on se rappelle que le rotationnel vaut

$$\nabla \times G = (0, 0, 2x - 2y)$$

et que pour calculer le flux de $\nabla \times G$ au travers de la demi-sphère S, on peut utiliser le théorème de la divergence. En effet, celui-ci établit que

$$\iint_D \nabla \times G \cdot dS + \iint_S \nabla \times G \cdot dS = \iiint_V \nabla \cdot \nabla \times G$$

où V est la demi-sphère unité pleine. Or $\nabla \cdot \nabla \times G = 0$. On en déduit que l'intégrale est nulle. D'après le calcul ci-dessous, nous avons donc

$$\iint_S \nabla \times \ \cdot dS = 0 - 0 = 0.$$

(2) Soit $\gamma(t) = (\cos(t), \sin(t), 0)$. L'intégrale demandée est

$$\int_0^{2\pi} (\cos^3(t)\sin(t) + \sin^3(t)\cos(t))dt = 0$$

ou, par la formule de stokes :

$$\iint_D \langle (0, 2, 0), (0, 0, 1) \rangle \, dxdy = 0$$

où $(0, 0, 1)$ représente le vecteur normal au disque unité D dans le plan $z = 0$.

(3) Utilisons le théorème de Stokes. Le rotationnel de $G = (y+z, z+x, x+y)$ vaut $\nabla \times G = (0, 0, 0)$. Voilà qui est réglé.

(4) Le rotationnel de $G = (y, z, x)$ vaut $(-1, -1, -1)$. Le vecteur normal au disque $x^2 + z^2 = a^2$ dans le plan $y = 0$ (ordre des paramètres : (x, z))vaut $(1, 0, 0) \wedge (0, 0, 1) = (0, -1, 0)$. On peut donc calculer l'intégrale de

$$\iint_D dxdy = \pi a^2$$

puisque $\iint_D dxdy$ est l'aire du disque de rayon a. On en conclut que

$$\int_\gamma ydx + zdy + xdz = \pi a^2$$

où γ est le cercle donné, orienté par le vecteur $(0, 0, 1)$ au point $(1, 0, 0)$.

Chapitre 97

Exercices de calcul différentiel et intégral 2

97.1 Supremum, maximum

Exercise 1 exo0001

Déterminez, s'ils existent, les supremum, infimum, maximum et minimum dans $\mathbb{R}$ des sous-ensembles de $\mathbb{R}$ ci-dessous.

(1) $]10, 36]$
(2) $[10, 36[$
(3) $]5, +\infty[$
(4) $]8, 9[$
(5) $\{\frac{1}{2}, \frac{-2}{3}, \frac{3}{4}, \frac{-4}{5}, \frac{5}{6}, ...\}$
(6) $\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, ...\}$ itemexo1g
(7) $\{\frac{1}{k} \mid k \in \mathbb{N}_0\}$
(8) $\{\frac{(-1)^k k}{k+1} \mid k \in e\mathbb{N}_0\}$
(9) $\{\sin(\frac{\pi k}{100}) \mid k \in \mathbb{Z}\}$
(10) $\{\sin(\frac{\pi k}{3}) \mid k \in \mathbb{Z}\}$ itemexo1h itemexo1k
(11) $\{x^2 \mid x \in]-1, \frac{1}{2}[\}$
(12) $\text{Image}([x \to e^x]) := \{e^x \mid x \in e\mathbb{R}\}$
(13) $\text{Image}(e^x cos(x)) := \{e^x cos(x) \mid x \in \mathbb{R}\}$
(14) $\text{Image}(\varphi(x)) := \{\varphi(x) \mid x \in [-1, 1]\}$
(15) $\text{Image}(|\varphi(x)|) := \{|\varphi(x)| \mid x \in [-1, 1]\}$

où

$$\varphi = \begin{cases} \sqrt{1-x^2} & x \leqslant 0 \\ x - 2 & x > 0 \end{cases}$$

corr0001

Correction of the exercise 1

Par convention, si le supremum n'existe pas, nous disons qu'il vaut l'infini. Cela est logique pour la raison suivante : étant donné que toute partie majorée de $\mathbb{R}$ a un supremum, le fait de ne pas en avoir signifie que la partie considérée n'est pas majorée. Dans ce cas, il est logique de dire que son supremum est infini. (bien que cela ne soit pas vrai au sens strict de la définition)

num	sup	inf	max	min
a	36	10	36	NAN
b	36	10	NAN	10
c	∞	5	NAN	NAN
d	9	8	NAN	NAN
e	1	-1	NAN	NAN
f	$1/2$	0	$1/2$	NAN
g	1	0	1	NAN
h	1	-1	NAN	NAN
i	1	-1	1	-1
j	$\sin \pi/3$	$\sin 4\pi/3$	$\sin \pi/3$	$\sin 4\pi/3$
k	1	0	NAN	0
l	∞	0	NAN	NAN
m	∞	$-\infty$	NAN	NAN
n	1	-2	1	NAN
o	2	0	NAN	0

Exercise 2 exo0002

Reprenez les exercices (7), (8) et (11), et justifiez soigneusement vos réponses. corr0002

Correction of the exercise 2

Pour le (7), la suite est décroissante, donc le maximum est atteint (et est le premier terme), et l'infimum sera la limite si elle existe. Cette limite sera le minimum si elle est atteinte. Ici, la limite est zéro, et n'est évidement pas atteinte.

Pour le (8), la suite est alternée. Il faut donc bien se garder de prendre le plus petit en norme comme étant le minimum. La limite des termes positifs est 1, et celle des termes négatifs est -1. Ce sont les candidats supremum et infimum. Étant donné que ces deux sous-suites sont croissantes en norme, ils sont effectivement supremum et infimum. Comme ils ne sont pas atteints, il n'y a pas de maximum, ni de minimum. En effet, quand il y a un supremum, un maximum ne peut que lui être égal.

La fonction présente dans l'exercice (11) est paire et toujours positive ou nulle. Le minimum (et donc infimum) est donc zéro : il est atteint et aucune valeurs négatives n'est atteinte. La fonction $x \mapsto x^2$ est une fonction croissante de $|x|$, donc il faut chercher le supremum sur le point du domaine le plus éloigné de zéro, c'est-à-dire les points arbitrairement proches de -1. Prouvons que 1 est le supremum de l'ensemble considéré. En effet, pour tout ϵ, le nombre $1+\epsilon$ n'est pas dans l'ensemble parce qu'il serait l'image de $\pm\sqrt{1+\epsilon}$ qui n'est pas dans $]-1, \frac{1}{2}[$.

Exercise 3 exo0003

Soient A et B des sous-ensembles bornées de $\mathbb{R}$

(1) Prouver que si $A \subset B$ alors $\sup(A) \leqslant \sup(B)$. Que peut on dire pour inf ?

(2) Est il vrai que si $A \subset B$ et $A \neq B$ alors $\sup(A) < \sup(B)$?

(3) Établir des relations liant $\sup(A)$ et $\sup(B)$ à $\sup(A \cup B)$ d'une part, et à $\sup(A \cap B)$ d'autre part (si $A \cap B \neq \varnothing$).

corr0003

Correction of the exercise 3 itemCorr3a

(1) Étant donné que $A \subset B$, nous avons que, si $s_B = \sup(B)$, alors $\forall x \in B$, $x \leqslant s_B$, et en particulier,

$$\forall x \in A,\ x \leqslant s_B. \tag{97.1}$$ EqForallsAsBgeq

Si maintenant, $s_A = \sup(A)$ et $s_A > s_B$, alors il existe un $\epsilon > 0$ tel que $s_B = s_A - \epsilon$, et donc $\exists y \in A$ tel que $y > s_B$, ce qui contredirait (97.1).

De la même manière, si $A \subset B$, alors $\inf(A) \geqslant \inf(B)$.

(2) Non. Par exemple $A = \{2\}$ et $B = \{1, 2\}$.

(3) Nous avons

$$\sup(A \cup B) = \max\{\sup(A), \sup(B)\}. \tag{97.2}$$

En effet, appelons m ce maximum. Pour tout $x \in A \cup B$, nous avons $x \leqslant m$, et d'autre part, si $\epsilon > 0$, alors $m - \epsilon$ est soit plus petit que s_A, soit plus petit que s_B (soit les deux). Si $m < s_A$, alors il existe un $x \in A \subset A \cup B$ tel que $x > m - \epsilon$, et, de la la même manière, si $m < s_B$, alors il existe un $x \in B \subset A \cup B$ tel que $x > m - \epsilon$.

Nous n'avons par contre pas de rapport direct entre $\sup(A)$, $\sup(B)$ et $\sup(A \cap B)$, comme le montre l'exemple $A = \{1, 5\}$ et $B = \{1, 7\}$. Par contre, nous avons que

$$\sup(A \cap B) \leqslant \min\{\sup(A), \sup(B)\}. \tag{97.3}$$

En effet, en utilisant les résultats du point (1), et en tenant compte du fait que que $A \cap B \subset A$, nous avons $\sup(A) \geqslant \sup(A \cap B)$ et $\sup B \geqslant \sup(A \cap B)$.

Exercise 4 exo00035
Si $A + B = \{a + b \text{ tel que } a \in A, b \in B\}$, prouver que

$$\sup(A + B) = \sup(A) + \sup(B). \tag{97.4}$$

corr00035

Correction of the exercise 4

Une propriété fondamentale du supremum est la suivante : pour tout $\epsilon > 0$,

$$\begin{aligned} &\exists a \in A \text{ tel que } a \geqslant \sup(A) - \epsilon/2 \\ &\exists b \in B \text{ tel que } b \geqslant \sup(B) - \epsilon/2. \end{aligned} \tag{97.5}$$

Pour tout ϵ, nous avons donc, pour a et b bien choisis dans A et B,

$$\sup(A) + \sup(B) - \epsilon \leqslant a + b \leqslant \sup(A + B). \tag{97.6}$$

Par conséquent [1], nous avons

$$\sup(A) + \sup(B) \leqslant \sup(A + B). \tag{97.7}$$ EqAplusBsupAsupBIneqPrem

Afin d'obtenir l'inégalité dans l'autre sens, regardons ce que l'on peut dire de $\sup(A+B)-\epsilon$. Nous écrivons la propriété fondamentale du supremum :

$$\exists x \in A + B \text{ tel que } \sup(A + B) - \epsilon < x, \tag{97.8}$$

en d'autres termes (puisque $x \in A + B$), il existe $a \in A$ et $b \in B$ tels que

$$\sup(A + B) - \epsilon \leqslant a + b \leqslant \sup(A) + \sup(B). \tag{97.9}$$ EqAplusBsupAsupB

La dernière équation est juste le fait que $a \leqslant \sup(A)$ et $b \leqslant \sup(B)$. Les inégalités (97.9) étant vraies pour tout ϵ, nous avons

$$\sup(A + B) \leqslant \sup(A) + \sup(B). \tag{97.10}$$

Ceci combiné à l'inégalité (97.7) donne l'égalité attendue.

Exercise 5 exo0004

Soient deux fonctions bornées f et g de E vers $\mathbb{R}$.

$$f, g : E \to \mathbb{R}$$

Prouvez que

$$\sup\{f(x) + g(x) \mid x \in E\} \leqslant \sup\{f(x) \mid x \in E\} + \sup\{g(x) \mid x \in E\}$$

1. Il est important que vous méditiez sur le fait que si $x + \epsilon \geqslant y$ pour tout $\epsilon > 0$, alors $x \geqslant y$.

Montrer par un exemple que l'égalité n'a pas toujours lieu. corr0004

Correction of the exercise 5

Nous avons

$$\begin{aligned}\sup\{f(x) \text{ tel que } x\in E\}+\sup\{g(y) \text{ tel que } y\in E\}&\\ &=\sup\{f(x)+g(y) \text{ tel que } (x,y)\in E\times E\}\\ &\geqslant \sup\{f(x)+g(y) \text{ tel que } (x,y)\in E\times E, x=y\}\\ &=\sup\{f(x)+g(x) \text{ tel que } x\in E\}.\end{aligned} \tag{97.11}$$

La première inégalité semble un peu mystérieuse. Elle n'est en réalité qu'une application de l'exercice 4 avec

$$\begin{aligned}A&=\{f(x) \text{ tel que } x\in E\}\\ B&=\{g(y) \text{ tel que } y\in E\}.\end{aligned} \tag{97.12}$$

Intuitivement, l'égalité n'aura lieu que quand f et g prennent leur supremum au même point. Prenons par exemple $E=[0,1]$ et puis $f(x)=x$ et $g(x)=1-x$. Le supremum de f et g sont tout deux 1, tandis que le supremum de la somme est la constante 1. Donc $\sup\{f+g\}=1$ tandis que $\sup\{f\}+\sup\{g\}=2$.

Preuve alternative...due à un étudiant en séance d'exercices.

Prenons $M=\sup\{f(x) \text{ tel que } x\in E\}$. Par définition du supremum, $M\geqslant f(x)$ pour tout $x\in E$. De la même manière, nous prenons $\tilde{M}=\sup\{g(y) \text{ tel que } y\in E\}$. Pour tout $x\in E$, nous avons

$$f(x)+g(x)\leqslant M+\tilde{M}, \tag{97.13}$$

c'est-à-dire que $M+\tilde{M}$ est un majorant de l'ensemble $\{f(x)+g(x) \text{ tel que } x\in E\}$. Étant donné que le supremum est le minimum de l'ensemble des majorants, nous en déduisons que

$$\sup\{f(x)+g(x) \text{ tel que } x\in E\}\leqslant M+\tilde{M}. \tag{97.14}$$

Exercise 6 exo0005

Soit E un sous-ensemble de $\mathbb{R}$. Prouver que si

$$x=\lim x_k$$

où $x_k\in E$ pour tout k, et si $x\in \text{Maj}(E)$ avec [2]

$$\text{Maj}(E)=\{y \text{ tel que } y\geqslant e,\ \forall e\in E\}, \tag{97.15}$$

alors $x=\sup(E)$. corr0005

Correction of the exercise 6

Il faut prouver que $\forall\epsilon>0$, $\exists y\in E$ tel que $y>x-\epsilon$. Étant donné que $x=\lim x_k$, il existe un x_k tel que $|x-x_k|<\epsilon$, mais vu que $x\in M(E)$, nous avons $x_k<x$ et donc $x-x_k<\epsilon$, ou encore : $x-\epsilon<x_k$.

97.2 Suites

Exercise 7 exo0006

Donner quand c'est possible un exemple de suite

(1) convergente

(2) périodique (non constante)

2. Vous reconaitrez l'ensemble des majorants de E.

(3) ayant une limite infinie
(4) n'ayant pas de limite
(5) à valeurs entières convergeant vers π
(6) croissante et bornée
(7) bornée mais non convergente
(8) bornée, monotone décroissante
(9) bornée, monotone croissante mais non convergente
(10) croissante et n'ayant pas de limite
(11) décroissante divergente ayant une sous-suite convergente

corr0006

Correction of the exercise 7

(1) $x_k = 1$ pour tout k,
(2) $x_k = (-1)^k$
(3) $x_k = k$,
(4) $x_k = (-1)^k$,
(5) Impossible parce que pour $\epsilon = 0.01$, il n'y a pas d'entiers n tels que $|n - \pi| \leqslant \epsilon$,
(6) $x_k = 1 - \frac{1}{k}$,
(7) $x_k = (-1)^k$,
(8) $x_k = \frac{1}{k}$,
(9) impossible par le théorème de la page 43,
(10) impossible, en effet, par le point précédent, il faudrait une suite non bornée. Donc pour tout M, il existe un K tel que $x_K > M$. Mais, étant donné que la suite est croissante, pour tout $k > K$, $x_k \geqslant x_K > M$. Cela prouve que la suite a l'infini comme limite.
(11) Impossible parce que décroissant et divergent implique non bornée. Donc $\forall M < 0$, il existe un K tel que $k > K$ implique $x_k < M$. Toute sous-suite « converge » donc également vers $-\infty$.

Exercise 8 exo0007

Démontrez que les suites suivantes sont convergentes :

(1) $k \mapsto \frac{1}{k}$
(2) $(1, \frac{1}{2}, \frac{-1}{3}, \frac{1}{4}, \frac{-1}{5}, \frac{1}{6}, \ldots)$
(3) $x_k = \frac{(i)^k}{k}$
(4) $x_k = \frac{1}{k^2}$
(5) $k \mapsto \frac{1}{k+3}$

corr0007

Correction of the exercise 8

(1) La suite $k \mapsto \frac{1}{k}$ est monotone décroissante, bornée vers le bas par zéro, elle est donc convergente. Il n'est pas compliqué de prouver qu'elle converge vers zéro.
(2) Le candidat limite est 0, et en effet,

$$\left\| \frac{(-1)^k}{k} - 0 \right\| = \left| \frac{(-1)^k}{k} \right| = \frac{1}{k} \to 0, \tag{97.16}$$

donc le critère de convergence s'applique.
(3) Étant donné que $|i^n| = 1$ dans $\mathbb{C}$, le raisonnement du point précédent s'applique.
(4) $\frac{1}{k^2}$ est une sous-suite de $\frac{1}{k}$ qui converge,

(5) $k \mapsto \frac{1}{k+3}$ est également une sous-suite (décalée de 3).

Exercise 9 exo0008

Une suite réelle $\{a_n\}$ satisfait la propriété $\forall N,\ \exists n_1, n_2 > N$ tel que $|a_{n_1}| < 5$ et $|a_{n_2}| > 5$. Construire une suite convergente et une suite divergente ayant cette propriété. corr0008

Correction of the exercise 9

Une suite convergente vers 5 est faisable. Par exemple

$$k \mapsto 5 + \frac{(-1)^k}{k}. \tag{97.17}$$

Pour une suite divergente, simplement prendre $4, 6, 4, 6, \ldots$

Exercise 10 exo0009

On considère deux suites croissantes, $\{x_n\}$ et $\{y_n\}$. Démontrez ou infirmez à l'aide d'un contre-exemple,

(1) la suite $\{x_n + y_n\}$ est croissante.

(2) la suite $\{x_n.y_n\}$ est croissante.

corr0009

Correction of the exercise 10

(1) vrai.

(2) faux. Le secret pour construire un contre-exemple est de faire intervenir la suite $-1/k$ qui est croissante tout en restant négative. Avec cela, nous construisons les suites $x_k = -\frac{1}{k}$ et $y_k = x^3$. Le produit vaut $(xy)_k = -x^2$ qui est décroissante.

Exercise 11 exo0010

Soient x_k et y_k deux suites tendant vers $+\infty$ et z_k une suite tendant vers un réel a strictement positif ($a > 0$). Démontrez que :

(1) $\lim(x_k + y_k) = +\infty$

(2) $\lim(x_k y_k) = +\infty$

(3) $\lim(x_k + z_k) = +\infty$

(4) $\lim(x_k z_k) = +\infty$

(5) $\lim(-x_k) = -\infty$

corr0010

Correction of the exercise 11

(1)

(2) Soient $M > 1$ et K, L tels que $k > K$ et $l > L$ impliquent $x_k > M$ et $y_k > M$. Dans ce cas, $m > \max\{K, L\}$ implique $(xy)_m > M^2 > M$.

(3) Soient $M > 0$, K tel que $k > K$ implique $x_k > M + 1$ et L tel que $l > L$ implique $|z_l - a| < \frac{1}{2}$. Alors $m > \max\{K, L\}$ implique $(x + z)_k > M + 1 \pm \frac{1}{2} > M + \frac{1}{2} > M$.

(4) Soient $M > 0$ et K tels que $k > K$ implique $x_k > \frac{M}{a} + 1$, et $z_k > a - \epsilon$. Un tel k peut être trouvé pour tout choix de ϵ. Nous choisissons ϵ de façon à avoir $a - \epsilon > 0$, et nous minorons

$$x_k z_k > \left(\frac{M}{a} + 1\right)(a - \epsilon) = M + a - \frac{\epsilon M}{a} - \epsilon. \tag{97.18}$$

Si nous prenons ϵ assez petit pour que $a - \frac{\epsilon M}{a} - \epsilon$, nous trouvons $x_k z_k > M$.

(5) Soit $M < 0$, il faut trouver un K tel que $-x_k < M$ dès que $k > K$. Évidement, le K tel que $x_k > -M$ fonctionne.

Exercise 12 exo0011

Soit x_k une suite bornée et y_k une suite tendant vers $+\infty$. Démontrez que la suite $[k \to \frac{x_k}{y_k}]$ tend vers 0. corr0011

Correction of the exercise 12

Nous prouvons que $|x_k/y_k| \to 0$. La suite x_k étant bornée, nous pouvons dire que

$$\left|\frac{x_k}{y_k}\right| < \frac{M}{y_k} \tag{97.19}$$

où nous considérons (quitte à passer à une queue de suite) que y_k est toujours positif. Notez que le M est prit indépendant de k. À partir d'ici, l'exercice revient à montrer que si $x_k \to \infty$, alors $(1/x_k) \to 0$.

Exercise 13 exo0012

Les suites suivantes convergent elles et si oui vers quel nombre ? Démontrez soigneusement toutes vos affirmations.

(1) $x_k = \frac{k+2}{k}\cos(k\pi)$
(2) $k \mapsto \frac{\frac{1}{k^3}+\frac{1}{k}+1}{\frac{5}{k^3}+2}$
(3) $k \mapsto \frac{k^3+k+1}{5k^3+2}$
(4) $\frac{k+(-1)^k}{k-(-1)^k}$
(5) $x_k = i^k$
(6) $k \mapsto k^2+2$
(7) $x_k = x_{k-1}^2 + 1, \qquad x_1 = 1$

corr0012

Correction of the exercise 13

(1) Nous savons que $(k+2)/k \to 1$. Prenons donc K tel que $k > K$ implique $(k+2)/k > \frac{1}{2}$. D'autre part, $\cos(k\pi) = (-1)^{k+1}$, donc à tout moment de la suite, il y a un élément plus petit que $-1/2$ et un autre plus grand que $1/2$. Il n'y a donc pas de convergence parce qu'une telle suite ne peut pas être de Cauchy.

(2) On utilise la proposition de la page 39 du cours qui dit que la limite d'un quotient est le quotient des limites. Or, la suite $\frac{1}{k^3} + \frac{1}{k} + 1$ tend vers 1 et la suite $\frac{5}{k^3} + 2$ tend vers 2. La suite proposée tend donc vers $1/2$.

(3) Le secret est de mettre en évidence le terme de plus haut degré, et de simplifier :

$$x_k = \frac{k^3(1 + \frac{1}{k^2} + \frac{1}{k^3})}{5 + \frac{2}{k^3}}, \tag{97.20}$$

ensuite on se souvient que la limite du quotient est le quotient des limites (quand elles existent, ce qui est le cas après simplification par k^3). Nous obtenons donc convergence vers $1/5$.

(4) En mettant k en évidence et en simplifiant, nous tombons sur la suite

$$x_k = \frac{1 + \frac{(-1)^k}{k}}{1 - \frac{(-1)^k}{k}}, \tag{97.21}$$

où les limites du numérateur et du dénominateur existent indépendamment. La limite est donc le quotient des limites, c'est-à-dire 1.

(5) Cette suite est dans $\mathbb{C}$. La sous-suite i^{4k} est constante (et vaut 1), tandis que la sous-suite i^{4k+1} est constante et vaut i. Cette suite n'est donc pas de Cauchy.

(6) Cette suite diverge (tend vers l'infini).

(7) Nous avons $x_2 = 2$, et ensuite $x_k > 2^{(2^k)}$. Cela se prouve par récurrence, en effet si $x_k > 2^{(2^k)}$,

$$x_{k+1} = x_k^2 + 1 > \left(2^{(2^k)}\right)^2 + 1 > 2^{(2^{k+1})}. \tag{97.22}$$

Étant donné que la suite $y_k = 2^{(2^k)}$ tend vers l'infini, la suite des x_k (qui est toujours plus grande) tend également vers l'infini.

Other resolution

Considérons la suite $x_{k+1} = x_k^2 + 1$. Si cette suite a une limite x, alors la suite $x \mapsto x_k^2$ a pour limite limite x^2. Pour chaque k, nous avons $x_{k+1} - x_k^2 = 1$, de sorte que la suite $k \mapsto x_{k+1} - x_k^2$ est constante et a 1 pour limite. Mais la différence de deux suites convergentes a pour limite la différence des limites, de telle sorte qu'en passant à la limite, la suite $x \mapsto x_{k+1} - x_k^2$ a pour limite $x - x^2$.

En d'autres termes, si nous supposons que $x_k \to x$, alors nous passons à la limite dans la relation de récurrence

$$x_{k+1} = x_k^2 + 1, \tag{97.23}$$

et nous trouvons l'équation

$$x = x^2 + 1 \tag{97.24}$$

que doit satisfaire le candidat limite x. Mais il est vite vu que cette équation n'a pas de solution réelle. La suite des x_k ne peut donc pas converger.

Nous referons abondamment usage de cette technique pour résoudre l'exercice 19.

Exercise 14 exo0015

Construire une suite $\{x_n\}$ telle que pour tout $a \in \mathbb{R}^n$, il existe une sous-suite $\{x_{n_k}\}$ qui converge vers a. corr0015

Correction of the exercise 14

Étant donné qu'il n'y a qu'une quantité dénombrable de rationnels, une simple adaptation de la question (5) de l'exercice 647 donne une suite qui contient une infinité de fois chaque rationnel. Toute suite de rationnels est une sous-suite de cette suite, et en particulier si r est un réel quelconque, une suite de rationnels qui converge vers r est une sous-suite de la suite considérée.

Other resolution

En réalité, toute suite qui énumère tous les rationnels convient. Il ne faut pas spécialement que tous les rationnels arrivent chacun une infinité de fois. En effet, soient $r \in \mathbb{R}$, et $B = B(r, \epsilon)$, une boule de rayon ϵ autour de r. Si x_k est une suite qui énumère tous les rationnels, pour tout N, il existe une infinité de $x_n \in B$ $(b > N)$, parce que cette boule contient une infinité de rationnels, qui ne peuvent donc pas être tous énumérés avant le Nième élément de la suite x_k.

Exercise 15 exo0018

Reprenez les réponses a, b, c, e et g de l'exercice précédent et justifiez les soigneusement. corr0018

Correction of the exercise 15

Il faut *toujours* justifier soigneusement. Cet exercice n'est donc pas à proprement parler un exercice. Si vous avez quelque chose à faire ici, c'est que vous avez mal fait quelque chose plus haut ...pas bien :p

Exercise 16 exo0019

Soient x_k et y_k deux suites bornées de nombres réels.

(1) Prouvez que $\limsup(x_k + y_k) \leqslant \limsup(x_k) + \limsup(y_k)$.

(2) Donnez un exemple dans lequel l'inégalité précédente est stricte.

(3) Prouvez que si l'une des deux suites est convergente, il y a égalité.

(4) Donnez un exemple pour lequel $\limsup(x_k y_k) > \limsup(x_k)\limsup(y_k)$.

(5) Trouvez une condition suffisante pour que

$$\limsup(x_k y_k) \leqslant \limsup(x_k)\limsup(y_k). \tag{97.25}$$

EqLimSupxyLeqLSxLSy

(6) Donnez un exemple où cette dernière inégalité est stricte.

(7) Prouvez que si

$$x_k, y_k \geqslant 0 \qquad \forall k,$$

alors dès que l'une des deux suites converge, on a l'égalité.

corr0019

Correction of the exercise 16

(1) Formellement, une suite $(x_k)_{k\in\mathbb{N}}$ est une fonction $f : \mathbb{N} \to \mathbb{R} : k \mapsto x_k$, et le résultat de l'exercice 5 s'applique. On en déduit que pour tout indice j, nous avons

$$\sup\{x_k + y_k \text{ tel que } k \geqslant j\} \leqslant \sup\{x_k \text{ tel que } k \geqslant j\} + \sup\{y_k \text{ tel que } k \geqslant j\}.$$

Ici, pour chaque j, nous avons pris $E_j = \{k \text{ tel que } k \geqslant j\}$ en guise de E de l'exercice 5. Il suffit maintenant de passer à la limite $(j \to \infty)$ pour obtenir le résultat.

(2) Prendre $x_k = (-1)^k$ et $y_k = -x_k$. L'inégalité devient $0 < 2$.

(3) Si l'une des suites, disons x_k, converge, alors

$$\limsup(x_k) = \lim(x_k) = -\lim(-x_k) = -\limsup(-x_k)$$

et on en déduit que

$$\begin{aligned}\limsup(y_k) &= \limsup(y_k + x_k + (-x_k)) \\ &\leqslant \limsup(y_k + x_k) + \limsup(-x_k) \\ &= \limsup(y_k + x_k) - \limsup(x_k)\end{aligned}$$

ce qui fournit l'inégalité inverse de celle obtenue au premier point, d'où l'égalité.

(4) Prendre $x_k = (-1, 0, -1, 0, \ldots)$ et $y_k = x_k$. L'inégalité devient $1 > 0$.

(5) Ce qui a fait fonctionner l'exemple du point précédent, c'est qu'un produit de nombres négatifs a donné un nombre positif, ce qui a permis, dans le produit, de passer au dessus de zéro, tandis que dans chacune des suites séparément, la limite supérieure était zéro. Ce stratagème ne fonctionne pas sans nombres négatifs.

Nous allons voir que la condition x_k, y_k répond à la question posée. Soit $l \in \mathbb{N}$ fixé. Pour tout $k \geqslant l$, nous savons

$$0 \leqslant x_k \leqslant \sup\{x_i \text{ tel que } i \geqslant l\} \text{ et } 0 \leqslant y_k \leqslant \sup\{y_i \text{ tel que } i \geqslant l\}$$

et on en déduit

$$0 \leqslant x_k y_k \leqslant \sup\{x_i \text{ tel que } i \geqslant l\}\sup\{y_i \text{ tel que } i \geqslant l\}$$

et donc

$$\sup\{x_k y_k \text{ tel que } k \geqslant l\} \leqslant \sup\{x_i \text{ tel que } i \geqslant l\}\sup\{y_i \text{ tel que } i \geqslant l\}$$

d'où le résultat attendu en prenant la limite pour $l \to +\infty$.

(6) Prendre $x_k = (1, 0, 1, 0, \ldots)$ et $y_k = (0, 1, 0, 1, 0, \ldots)$, c'est-à-dire la même suite décalée d'un indice pour que le produit soit la suite nulle. L'inégalité devient $0 < 1$.

(7) Supposons que l'une des suites converge, par exemple $(x_k) \to r$. Discutons deux cas : si $r \neq 0$, on a

$$\limsup(x_k) = \lim(x_k) = \frac{1}{\lim(\frac{1}{x_k})} = \frac{1}{\limsup(\frac{1}{x_k})} \tag{97.26}$$

et on en déduit

$$\begin{aligned}\limsup(y_k) &= \limsup(y_k \cdot x_k/x_k)) \\ &\leqslant \limsup(y_k x_k) \limsup\left(\frac{1}{x_k}\right) \\ &= \frac{\limsup(y_k x_k)}{\limsup(x_k)}\end{aligned} \tag{97.27}$$

ce qui fournit l'inégalité inverse de (97.25), d'où l'égalité. Lorsque $r = 0$, l'inégalité déjà obtenue est en fait

$$\limsup(x_k y_k) \leqslant 0 \tag{97.28}$$

mais d'après la condition sur x_k et y_k, on sait aussi que

$$0 \leqslant \limsup(x_k y_k) \tag{97.29}$$

et on en déduit l'égalité.

Exercise 17 exo0020

On considère la suite de nombres réels $k \mapsto x_k = (1 + \frac{1}{k})^k$ Item0020a

(1) Montrez que cette suite est convergente (pour ce faire prouvez qu'elle est monotone croissante et bornée). Item0020beborne

(2) Prouvez que sa limite $e \overset{def}{=} \lim x_k$ est supérieure à 2 et inférieure à 3.

(3) Déterminez

(3a) $\lim(1 + \frac{1}{k})^{k+1}$

(3b) $\lim(1 + \frac{1}{k})^{\frac{k}{2}}$

(3c) $\lim(1 + \frac{1}{2k})^k$

(3d) $\lim(1 - \frac{1}{k})^k$

en termes de e. corr0020

Correction of the exercise 17

Nous considérons la suite $x_k = \left(1 + \frac{1}{k}\right)^k$

(1) Montrons que (x_k) est croissante. Pour cela, utilisons la formule du binôme de Newton :

$$\begin{aligned}\left(1 + \frac{1}{k}\right)^k &= \sum_{i=0}^{k} \binom{k}{i} 1^{(1-i)} \left(\frac{1}{k}\right)^i \\ &= \sum_{i=0}^{k} \frac{k \cdot (k-1) \cdot \ldots \cdot 2 \cdot 1}{i! \cdot (k-i) \cdot (k-i-1) \cdot \ldots \cdot 2 \cdot 1} \frac{1}{k^i} \\ &= \sum_{i=0}^{k} \frac{k \cdot (k-1) \cdot \ldots \cdot (k-i+1)}{i!} \frac{1}{k^i} \\ &= \sum_{i=0}^{k} \frac{1}{i!} \cdot \frac{k}{k} \cdot \frac{k-1}{k} \cdot \frac{k-2}{k} \cdot \ldots \cdot \frac{k-(i-1)}{k} \\ &= \sum_{i=0}^{k} \frac{1}{i!} \cdot 1 \cdot \left(1 - \frac{1}{k}\right) \cdot \left(1 - \frac{2}{k}\right) \cdot \ldots \cdot \left(1 - \frac{i-1}{k}\right)\end{aligned}$$

Afin de soulager la notation, écrivons

$$A_k(i) = \frac{1}{i!} \cdot 1 \cdot \left(1 - \frac{1}{k}\right) \cdot \left(1 - \frac{2}{k}\right) \cdot \ldots \cdot \left(1 - \frac{i-1}{k}\right), \tag{97.30}$$

de façon à avoir $x_k = \sum_{i=0}^{k} A_k(i)$. Manifestement, tant que $i < k+1$ (ce qui est toujours le cas dans les sommes considérées), $A_k(i) \geqslant 0$ et $A_{k+1}(i) - A_k(i) \geqslant 0$ pour tout k et tout i. Donc

$$x_{k+1} - x_k = \sum_{i=0}^{k} (A_{k+1}(i) - A_k(i)) + A_{k+1}(k) \geqslant 0. \tag{97.31}$$

Cela prouve que la suite est croissante.

Par ailleurs, on observe que la suite est majorée par 3 : on peut majorer les facteurs du type $1 - \frac{\cdot}{k}$ par 1, et on en déduit que

$$\begin{aligned} x_k &\leqslant \sum_{i=0}^{k} \frac{1}{i!} = \sum_{i=0}^{k} \frac{1}{1 \cdot 2 \cdot 3 \cdot \ldots \cdot i} \\ &\leqslant \sum_{i=0}^{k} \frac{1}{1 \cdot \underbrace{2 \cdot 2 \cdot \ldots \cdot 2}_{i-1\text{fois}}} \\ &= 1 + \sum_{i=1}^{k} \frac{1}{2^{i-1}} < 1 + 2 = 3 \end{aligned} \tag{97.32}$$

où on utilise le fait que les sommes partielles de la série géométrique de raison $\frac{1}{2}$ sont plus petites que la somme de cette série, c'est-à-dire 2.

(2) Le premier point prouve déjà que la limite, notée e est inférieure à 3. De plus, la suite étant croissante on a forcément

$$e \geqslant x_1 = 2$$

ce qui est le résultat annoncé.

(3) Dans l'ordre : e, $\sqrt{e}$, $\sqrt{e}$, $1/e$.

(3a) Nous pouvons écrire $\left(1 + \frac{1}{k}\right)^{k+1} = \left(1 + \frac{1}{k}\right) \cdot \left(1 + \frac{1}{k}\right)^{k}$, et puis utiliser le produit de deux suites convergentes.

(3b) Il faut utiliser le fait que si x_k est convergente, $\lim \sqrt{x_k} = \sqrt{\lim x_k}$. Nous allons prouver cela « à la main » mais sachez que c'est une conséquence de la continuité de la fonction $t \mapsto \sqrt{t}$. Si la suite (x_k) converge vers $r \neq 0$, alors nous pouvons faire la manipulation suivante en supposant que k est assez grand pour que $|x_k - r| \leqslant \epsilon$:

$$\begin{aligned} |\sqrt{x_k} - \sqrt{r}| &= \left| \frac{(\sqrt{x_k} - \sqrt{r})(\sqrt{x_k} + \sqrt{r})}{\sqrt{x_k} + \sqrt{r}} \right| \\ &= \left| \frac{x_k - r}{\sqrt{x_k} + \sqrt{r}} \right| \\ &\leqslant \frac{|x_k - r|}{\sqrt{r}} \\ &\leqslant \frac{\epsilon}{\sqrt{r}} \end{aligned} \tag{97.33}$$

EqInegssqrtLimCQu

où nous avons tenu compte du fait que $\sqrt{x_k} + \sqrt{r} \geqslant \sqrt{r}$. Pour tout $\epsilon > 0$, il existe un K tel que $k > K$ implique les inégalités (97.33). Si ϵ' est donné, il suffit de prendre $\epsilon < \sqrt{r}\epsilon'$ pour obtenir

$$|\sqrt{x_k} - \sqrt{r}| \leqslant \epsilon'. \tag{97.34}$$

(3c) Si nous comparons les suites

$$x_k = \left(1 + \frac{1}{k}\right)^k \quad \text{et} \quad y_k = \left(1 + \frac{1}{2k}\right)^k, \tag{97.35}$$

nous voyons que $x_{2k} = \left(1 + \frac{1}{2k}\right)^{2k}$, et donc $y_k = \sqrt{x_{2k}}$. Or, il est évident que la limite de la suite (x_{2k}) est la même que la limite de la suite (x_k).

(3d) Il faut essayer de transformer $1-\frac{1}{k}$ en $1+\frac{1}{k}$. Pour cela, on voit que

$$y_k = \left(1-\frac{1}{k}\right)^k = \left(\frac{k-1}{k}\right)^k = \left(\frac{k}{k-1}\right)^{-k}. \tag{97.36}$$

Étant donné que $k > k-1$, nous exprimons $\frac{k}{k-1}$ sous la forme $1+x$, et nous trouvons

$$y_k = \left(1+\frac{1}{k-1}\right)^{-k} = \left(1+\frac{1}{k-1}\right)^{-k+1}\left(1+\frac{1}{k-1}\right)^{-1} \to 1/e \tag{97.37}$$

par la règle du produit de deux suites convergentes, et le fait que $\lim(1/x_k) = 1/\lim(x_k)$.

97.2.1 Suites définies par récurrence

Exercise 18 exo0021

Soient a et b deux nombres réels fixés. Définissons la suite (x_n) comme suit : $x_1 = a$, $x_2 = b$ et, pour $n > 2$, x_n est la moyenne des deux termes précédents, c'est-à-dire, $x_n = \frac{1}{2}(x_{n-2} + x_{n-1})$.

(1) Faites un dessin (Aide : supposez $a < b$ et pensez aux points milieux).

(2) Prouvez (par induction) que

$$x_{n+1} - x_n = \left(-\frac{1}{2}\right)^{n-1}(b-a)$$

(3) Montrez que

$$x_{n+1} - x_1 = \left[\sum_{k=1}^{n}\left(-\frac{1}{2}\right)^{k-1}\right](b-a)$$

(4) Prouvez que $x_n \to \frac{(a+2b)}{3}$.

corr0021

jour peut-

Correction of the exercise 18

(1)

(2) Nous avons

$$x_{n+2} - x_n = \frac{x_n - x_{n+1}}{2} - x_{n+1} = -\frac{1}{2}(x_{n+1} - x_n). \tag{97.38}$$

(3) Nous avons

$$x_{n+1} - x_1 = x_{n+1} - x_n + x_n - x_{n-1} + x_{n-1} - \ldots + x_3 - x_2 + x_2 - x_1 \tag{97.39}$$

que l'on regroupe deux à deux pour obtenir le résultat.

(4) Si nous posons $q = -1/2$, nous avons, en utilisant le résultat précédent, que

$$x_{n+1} = \left(\sum_{k=0}^{n-1} q^k\right)(b-a) + a \tag{97.40}$$

où nous avons utilisé le fait que $\sum_{k=1}^{n} q^{k-1} = \sum_{k=0}^{n-1} q^k$. Maintenant, il faut savoir prendre la limite, c'est-à-dire sommer cette série géométrique. Cela se fait en lisant wikipedia : la somme d'une série géométrique de raison q est $1/(1-q)$. De là, la conclusion vient aisément.

as de des-

Correction alternative

(1)

(2) Nous voulons montrer pour tout $n \geqslant 1$ la proposition

$$P(n) \equiv x_{n+1} - x_n = \left(-\frac{1}{2}\right)^{n-1}(b-a).$$

La proposition $P(1)$ est vraie, on trouve $b-a = b-a$. Supposons $P(i)$ vraie pour $i \geqslant 1$ fixé. Nous avons alors successivement

$$\begin{aligned} x_{i+2} - x_{i+1} &= \frac{1}{2}(x_i + x_{i+1}) - x_{i+1} \\ &= -\frac{1}{2}(x_{i+1} - x_i) \\ &= -\frac{1}{2}\left(-\frac{1}{2}\right)^{i-1}(b-a) \quad \text{car } P(i+1) \text{ est supposée vraie} \\ &= \left(-\frac{1}{2}\right)^{i}(b-a) \end{aligned}$$

ce qui est exactement la relation $P(i+1)$. Par le principe de récurrence, $P(n)$ est vraie pour tout naturel $n \geqslant 1$, ce qu'on voulait démontrer.

(3) On veut une égalité sur $x_{n+1} - x_1$, cest-à-dire la différence entre deux termes éloignés dans la suite. L'astuce usuelle dans ce cas est de faire apparaitre les différences des termes successifs :

$$\begin{aligned} x_{n+1} - x_1 &= x_{n+1} + (-x_n + x_n) + (-x_{n-1} + x_{n-1}) + \dots \\ &\quad + (-x_3 + x_3) + (-x_2 + x_2) - x_1 \\ &= (x_{n+1} - x_n) + (x_n - x_{n-1}) + (x_{n-1} + \dots \\ &\quad - x_3) + (x_3 - x_2) + (x_2 - x_1) \\ &= \sum_{i=1}^{n}(x_{i+1} - x_i) \\ &= \sum_{i=1}^{n}\left[\left(-\frac{1}{2}\right)^{i-1}(b-a)\right] \\ &= \sum_{i=1}^{n}\left[\left(-\frac{1}{2}\right)^{i-1}\right](b-a) \end{aligned}$$

ce qui prouve l'égalité souhaitée.

(4) On se rappelle de la formule de la *série géométrique*

$$\sum_{k=0}^{i} q^k = \frac{1-q^{i+1}}{1-q}$$

ce qui, appliquée au cas qui nous intéresse, donne (avec $q = {}^{-1}\!/_{2}$)

$$\begin{aligned} x_{n+1} &= x_1 + \sum_{i=1}^{n}\left[\left(-\frac{1}{2}\right)^{i-1}\right](b-a) \qquad \text{(regardez bien les indices !)} \\ &= x_1 + \frac{1-(-\frac{1}{2})^n}{1-(-\frac{1}{2})}(b-a) \\ &= x_1 + \frac{1-(-\frac{1}{2})^n}{\frac{3}{2}}(b-a) \end{aligned}$$

et, étant donné que $\left|-\frac{1}{2}\right| < 1$, en passant à la limite on a

$$\lim_{n\to\infty}(-\frac{1}{2})^n = 0$$

ce qui montre que

$$\lim_{n\to\infty}(x_{n+1}) = x_1 + \frac{1-0}{\frac{3}{2}}(b-a) = a + \frac{2}{3}(b-a) = \frac{2}{3}b + \frac{1}{3}a$$

qui est le résultat attendu.

Exercise 19 exo0022

Démontrez que la suite définie par la récurrence suivante est convergente et calculez sa limite

$$x_k = \frac{1}{1+x_{k-1}}, \quad x_1 = 1. \tag{97.41}$$

Pour ce faire, ItemAExo0022

(1) Trouvez un candidat limite en passant à la limite dans la relation de récurrence,
(2) déterminez quels termes de la suite sont plus grands ou plus petit que le candidat,
(3) montrez que la suite des termes plus grands et la suite des termes plus petits convergent toutes deux vers la même limite,
(4) déduisez en que le candidat trouvé en (1) est bien la limite des x_k.

corr0022

Correction of the exercise 19

On obtient des candidats limite en appliquant les règles de calculs sur la relation de récurrence $x_k(1+x_{k-1}) = 1$. En effet, en supposant que la suite (x_k) converge vers un réel x, alors ce réel doit satisfaire $x(1+x) = 1$ puisque les suites $k \mapsto x_k$ et $k \mapsto x_{k-1}$ convergent toutes les deux vers x. Ceci donne les deux *candidats limite* solutions de cette équation :

$$x_- = \frac{-1-\sqrt{5}}{2} \quad \text{et} \quad x_+ = \frac{-1+\sqrt{5}}{2}. \tag{97.42}$$

Par récurrence, on prouve aisément (le faire !) que tous les termes de la suite sont strictement positifs. Cela exclu la possibilité x_- qui est strictement négatif[3]. Nous savons donc que *si la limite de la suite existe*, alors cette limite est x_+.

Vérifions d'abord si la suite ne serait pas bornée vers le haut ou vers le bas par x_+. Supposons que $x_k < x_+$. Note que au moins x_1 satisfait cette condition, donc nous ne travaillons pas dans le vide. Nous avons alors $1 + x_k \geqslant 1 + x_+$, et donc

$$x_{k+1} = \frac{1}{1+x_k} \leqslant \frac{1}{1+x_+}. \tag{97.43}$$

Mais x_+ est justement solution de l'équation $1/(1+x) = x$, donc $1/(1+x_+) = x_+$ et nous avons

$$x_{k+1} \leqslant x_+ \tag{97.44}$$

dès que $x_k \geqslant x_+$. Inversement, nous trouvons que

$$x_{k+1} \geqslant x_+ \tag{97.45}$$

dès que $x_k \geqslant x_+$. Donc la suite oscille en réalité autour de x_+. Les termes impairs seront tous plus petits que x_+ et les termes pairs seront tous plus grand que x_+. Les suites (x_{2k}) et (x_{2k+1}) sont donc à regarder séparément. Calculons pour voir si ces suites sont croissantes ou décroissantes :

$$x_{k+2} = \frac{1}{1+\frac{1}{1+x_k}} = \frac{x_k+1}{x_k+2}, \tag{97.46}$$ EqNombreOrxkPlusDeux

et donc

$$x_{k+2} - x_k = -\frac{x_k^2 + x_k - 1}{x_k + 2}. \tag{97.47}$$

3. Une suite de réels strictement positifs peut avoir une limite nulle, mais pas strictement négative. . .méditez là dessus.

Le dénominateur est toujours positif, tandis que le signe du numérateur dépend de x_k, et la valeur de changement de signe n'est autre que x_+. Pas mal hein !

Nous avons que

$$\begin{aligned} x_k > x_+ &\Rightarrow x_{k+2} < x_k \\ x_k < x_+ &\Rightarrow x_{k+2} > x_k \end{aligned} \tag{97.48}$$

Donc la suite des (x_{2k}) est plus grande que x_+ et décroissante, tandis que la suite des x_{2k+1} est plus petite que x_+ et décroissante. Or nous savons qu'une suite bornée et monotone est convergente. Donc les deux suites sont convergentes.

D'après l'équation (97.46), nous avons $x_{k+2}(x_k+2) = x_k + 1$, et donc la seule limite possible vérifie $x(x+2) = x+1$, dont la seule solution acceptable est, encore une fois, x_+.

Nous sommes maintenant dans le cas d'une suite x_k donc les deux sous-suites x_{2k} et x_{2k+1} sont convergentes et convergent vers la même limite, c'est-à-dire le cas de la question (1) de l'exercice 647 : la suite des x_k converge vers x_+.

97.3 Calcul de limites

SecCalcLimFHtQNu

Exercise 20 exo0023

Déterminez si les limites suivantes existent et dans l'affirmative calculez les en utilisant, si il y a lieu, la règle de l'Hospital ou la règle de l'étau.

(1) $\lim_{x\to+\infty} \frac{x+1}{x^2+2}$
(2) $\lim_{x\to+\infty} \frac{\sin(x)}{x}$
(3) $\lim_{x\to 0} \frac{\sin(x)}{x}$
(4) $\lim_{x\to+\infty} \frac{x^n}{e^x}$
(5) $\lim_{x\to+\infty}(1+\frac{a}{x})^x$
(6) $\lim_{x\to 0}(\frac{1}{\sin(x)} - \frac{1}{x})$
(7) $\lim_{x\to+\infty} \cos(2\pi x)$
(8) $\lim_{x\to+\infty} \frac{x}{\sin(x)+2} + \ln(x)\cos(x)$
(9) $\lim_{x\to+\infty} \frac{\ln(x)(\sin(x)+2)}{x}$
(10) $\lim_{x\to+\infty} x^{\frac{1}{x}}$

Item0023h

(Sont présentés ici différents types de problèmes auxquels on peut être confronté lors du calcul de limites de fonctions. Cet exercice est un exercice de drill : n'essayez pas de justifier *à fond* chaque étape du calcul.) corr0023

Correction of the exercise 20

Nous allons faire un usage intensif (et sans justifications trop poussées) des choses dites autour de la règle de l'Hospital.

(1) Nous commençons par mettre en évidence le plus haut degré de x au numérateur et au dénominateur :

$$\frac{x+1}{x^2+2} = \frac{x\left(1+\frac{1}{x}\right)}{x^2\left(1+\frac{2}{x^2}\right)} = \frac{1+\frac{1}{x}}{x\left(1+\frac{2}{x^2}\right)}. \tag{97.49}$$

Fixons $\epsilon > 0$, et considérons X_1 tel que $x > X_1$ implique $1 + \frac{1}{x} < 1 + \epsilon$. Nous pouvons aussi choisir X_2 tel que $x > X_2$ implique $x\left(1+\frac{2}{x^2}\right)$. En prenant le maximum de X_1 et X_2, nous trouvons

$$\frac{1+\frac{1}{x}}{x\left(1+\frac{2}{x^2}\right)} < \frac{1+\epsilon}{M}. \tag{97.50}$$

En prenant M arbitrairement grand, nous pouvons rendre cette fraction arbitrairement petite. La limite cherchée est donc zéro.

Other resolution

En utilisant la règle de l'Hospital, nous trouvons tout de suite

$$\lim_{x\to\infty} \frac{x+1}{x^2+2} = \lim_{x\to\infty} \frac{1}{2x} = 0. \tag{97.51}$$

(2) Le candidat limite n'est pas compliqué à deviner : c'est zéro parce que $\sin(x)$ reste borné tandis que le x au dénominateur vient l'écraser. Nous avons très vite une majoration

$$\left|\frac{\sin(x)}{x}\right| < \frac{1}{x} \to 0. \tag{97.52}$$

La limite est donc zéro. Notez que nous ne pouvons pas applique la règle de l'Hospital parce que la limite $\lim_{x\to\infty}\sin(x)$ n'existe pas.

(3) Ici par contre, la règle de l'Hospital fonctionne parce que $\lim_{x\to 0}\sin(x)$ existe, et vaut zéro. Nous avons donc

$$\lim_{x\to 0}\frac{\sin(x)}{x} = \lim_{x\to 0}\frac{\cos(x)}{1} = 1. \tag{97.53}$$

(4) Si n est un entier positif, nous trouvons

$$\lim_{x\to\infty}\frac{x^n}{e^x} = \lim_{x\to\infty}\frac{nx^{n-1}}{e^x} = \ldots = \lim_{x\to\infty}\frac{n!}{e^x} = 0, \tag{97.54}$$

en appliquant n fois la règle de l'Hospital. Dans le cas où n est un entier négatif, nous tombons sur un cas $0\cdot 0 = 0$.

(5) Le changement de variable $u = x/a$ donne

$$\left(1+\frac{a}{x}\right)^x = \left(1+\frac{1}{u}\right)^{au} = \left[\left(1+\frac{1}{u}\right)^u\right]^a. \tag{97.55}$$

En utilisant le fait que $\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t = e$ (voir la remarque à la page 141), nous trouvons alors

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a. \tag{97.56}$$

(6) En mettant au même dénominateur, nous trouvons un cas $\frac{0}{0}$ qui peut se traiter en utilisant deux fois la règle de l'Hospital :

$$\begin{aligned}\lim_{x\to\infty}\left(\frac{1}{\sin(x)}-\frac{1}{x}\right) &= \lim_{x\to 0}\left(\frac{x-\sin(x)}{x\sin(x)}\right)\\ &= \lim_{x\to 0}\frac{1-\cos(x)}{x\cos(x)+\sin(x)}\\ &= \lim_{x\to 0}\frac{\sin(x)}{-x\sin(x)+\cos(x)+\cos(x)}\\ &= \frac{0}{2} = 0.\end{aligned} \tag{97.57}$$

(7) La limite n'existe pas parce que $\forall X > 0, \exists x_0, x_1 > X$ tels que $\cos(2\pi x_0) = 1$ et $\cos(2\pi x_1) = 0$.

(8) Nous savons que $\sin(x)+2 \in [1,3]$, donc

$$\frac{x}{\sin(x)+2} + \ln(x)\cos(x) > \frac{x}{3} + \ln(x)\cos(x) > \frac{x}{3} - \ln(x), \tag{97.58}$$

qui est un cas $\infty - \infty$. Étudions la fonction

$$f(x) = \frac{x}{3} - \ln(x). \tag{97.59}$$

Eq0023frac

La dérivée de f vaut $f'(x) = \frac{1}{3} - \frac{1}{x} > \frac{1}{4}$. La fonction $f(x)$ majore donc une droite de coefficient directeur $\frac{1}{4}$ et tend donc vers l'infini.

(9) Nous majorons d'abord $|\sin(x)+2|$ par 3, et puis nous appliquons la règle de l'Hospital :

$$\lim_{x\to\infty}\left|\frac{\ln(x)\big(\sin(x)+2\big)}{x}\right| < \lim_{x\to\infty} 3\frac{\ln(x)}{x} = \lim_{x\to\infty}\frac{1/x}{1} = 0. \tag{97.60}$$

(10) Pour ce dernier, nous utilisons le passage par $y = e^{\ln(y)}$:

$$x^{1/x} = e^{\ln(x^{1/x})} = e^{\frac{1}{x}\ln(x)}. \tag{97.61}$$

Mais nous savons déjà que $\lim_{x\to\infty} \frac{1}{x}\ln(x) = 0$, donc la limite cherchée est $e^0 = 1$.

Other resolution

(1)

(2) $\lim_{x\to\infty} \frac{\sin(x)}{x}$
Par la règle de l'étau : $\frac{-1}{x} \leqslant \frac{sin(x)}{x} \leqslant \frac{1}{x}$. Comme $\frac{1}{x}$ tend vers 0 quand x tend vers ∞, on a $\lim_{x\to\infty} \frac{\sin(x)}{x} = 0$

(3) $\lim_{x\to 0} \frac{\sin(x)}{x}$
Un cas $\frac{0}{0}$. On peut donc appliquer la règle de l'Hospital :

$$\lim_{x\to 0} \frac{\sin(x)}{x} =^H \lim_{x\to 0} \frac{\cos(x)}{1} = \cos(0) = 1$$

où dans la dernière limite on utilise le fait que la fonction $\cos(x)$ est continue en 0.

(4) $\lim_{x\to\infty} \frac{x^n}{e^x}$
Un cas $\frac{\infty}{\infty}$. On peut donc appliquer la règle de l'Hospital :

$$\begin{aligned}\lim_{x\to\infty} \frac{x^n}{e^x} &=^H \lim_{x\to\infty} \frac{nx^{n-1}}{e^x}\\ &=^H \lim_{x\to\infty} \frac{n(n-1)x^{n-2}}{e^x}\\ &=^H \dots\\ &=^H \lim_{x\to\infty} \frac{n!}{e^x} = 0\end{aligned} \tag{97.62}$$

où dans la dernière limite on utilise le fait que la fonction $e^x \to \infty$ quand $x \to \infty$, et $n!$ reste borné.

(5) $\lim_{x\to\infty}(1 + \frac{a}{x})^x$
Ici, on transforme : $(1 + \frac{a}{x})^x = e^{x\ln(1+\frac{a}{x})}$. Comme l'exponentielle est continue, on peut passer à la limite dans l'exponentielle, et l'exercice devient de calculer :

$$\lim_{x\to\infty} x\ln(1 + \frac{a}{x}) = \lim_{x\to\infty} \frac{\ln(1 + \frac{a}{x})}{\frac{1}{x}}$$

qui est un cas $\frac{0}{0}$. On a donc, par l'Hospital :

$$\lim_{x\to\infty} \frac{\ln(1 + \frac{a}{x})}{\frac{1}{x}} =^H \lim_{x\to\infty} \frac{1/(1 + \frac{a}{x})\frac{-a}{x^2}}{\frac{-1}{x^2}} = a\lim_{x\to\infty} \frac{1}{(1 + \frac{a}{x})} = a$$

Cet argument est valable quel que soit a, et le résultat final est donc

$$\lim_{x\to\infty}(1 + \frac{a}{x})^x = e^a$$

(6)

(7) $\lim_{x\to\infty} \cos(2\pi x)$
On voit que cette fonction ne peut converger en l'infini. Pour le prouver, on peut par exemple prendre deux manières différentes d'aller vers l'infini qui donneront deux limites différentes de la fonction. Si on prend $x_k = k$, et $y_k = \frac{2k+1}{2}$, les deux suites tendent vers l'infini, mais $\lim_{k\to\infty} \cos(2\pi x_k) = 1$ alors que $\lim_{k\to\infty} \cos(2\pi y_k) = -1$, ce qui prouve que la fonction ne converge pas quand x tend vers l'infini.

(8)

(9) $\lim_{x\to\infty} \frac{\ln(x)(\sin(x)+2)}{x}$
Par la règle de l'étau :

$$\frac{\ln(x)}{x} \leqslant \frac{\ln(x)(\sin(x)+2)}{x} \leqslant 3\frac{\ln(x)}{x} \ \forall x$$

et, $\lim_{x\to\infty} \frac{\ln(x)}{x} =^H \lim_{x\to\infty} \frac{1/x}{1} = 0$.

Exercise 21 exo0025

Calculez les limites suivantes :

(1) $\lim_{x\to+\infty} \frac{\ln(x)}{x^a}$

(2) $\lim_{x\to+\infty} \frac{\ln(x)^a}{x^b}$

(3) $\lim_{x\to+\infty} a^x$

(4) $\lim_{x\to+\infty} a^{\frac{1}{x}}$

où a et b sont des réels positifs. (Le but de ces exercices est, aussi, d'acquérir un peu de culture. On y compare la croissance du logarithme et des polynômes) corr0025

Correction of the exercise 21

(1) $\lim_{x\to+\infty} \frac{\ln(x)}{x^a}$
Si $a > 0$ alors, par l'Hospital, nous trouvons $\lim_{x\to\infty} \frac{1}{ax^a}$, dont la limite est nulle. Si $a \leqslant 0$, elle vaut tout aussi clairement l'infini.

(2) $\lim_{x\to+\infty} \frac{\ln(x)^a}{x^b}$
Dans le cas où a ou b est nul, le résultat ne fait aucun doute. Si a et b n'ont pas le même signe, le numérateur et le dénominateur vont dans le même sens, et il n'y a aucun problèmes non plus. Il ne reste donc qu'à étudier le cas où a et b sont strictement positifs (le cas négatif se traite en passant à l'inverse). Pour tout x nous avons alors les inégalités

$$\frac{\ln(x)^{\lfloor a\rfloor}}{x^b} \leqslant \frac{\ln(x)^a}{x^b} \leqslant \frac{\ln(x)^{\lceil a\rceil}}{x^b}. \tag{97.63}$$ Eq0025EtauGrand

où $\lfloor a\rfloor$ représente le plafond du réel a, et $\lceil a\rceil$, son plancher. On calcule la limite du terme le plus à gauche en appliquant $\lfloor a\rfloor$ fois la règle de l'Hospital. Ce qu'il faut calculer devient :

$$\lim_{x\to+\infty} \frac{a(a-1)(a-2)\dots(a-(\lfloor a\rfloor-1))\ln(x)^{a-\lfloor a\rfloor}}{b^{\lfloor a\rfloor}x^b}.$$

On peut maintenant utiliser l'étau, en sachant que pour x suffisamment grand

$$0 \leqslant \frac{\ln(x)^a}{x^b} \leqslant \frac{\ln(x)}{x^b}$$

et nous avons vu à l'exercice précédent que cette dernière tend vers 0. Le terme le plus à droite de (97.63) se traite de la même façon, et sa limite est également nulle. La règle de l'étau conclut que

$$\lim_{x\to+\infty} \frac{\ln(x)^a}{x^b} = 0 \tag{97.64}$$

lorsque a et b sont des réels strictement positifs.

(3) $\lim_{x\to+\infty} a^x$
Il y a trois cas possibles (on suppose a positif). Quels autres cas y aurait-il eu si on avait admis a négatif ?

(3a) $a < 1$. Nous avons $x^x < \epsilon$ dès que $x > \log_a(\epsilon) = \frac{\ln(\epsilon)}{\ln(a)}$. La dernière expression est plus grande que zéro dès que $\epsilon < 1$ et $a < 1$.

(3b) $a = 1$. $\lim_{x\to+\infty} a^x = 1$
Trivial

(3c) $a > 1$. La dérivée de $x \mapsto a^x$ est $a^x \ln(a)$. Étant donné que $a > 1$, nous avons $\ln(a) > 0$ et $a^x > 1$ (dès que x est assez grand). La dérivée de a^x majore donc $\ln(a)$, et il existe une constante c telle que la fonction $x \mapsto a^x + c$ soit plus grande que $x \to \ln(a)x$. Donc $\lim_{x\to\infty} a^x = \infty$ lorsque $a > 1$.

Other resolution
Pour le cas $a > 1$. Par l'étau : $a^x = (1 + b)^x \geqslant 1 + nx$ pour $x \geqslant n$.

Exercise 22 exo0026
Déterminez, pour chacune des suites suivantes, si elle converge et dans l'affirmative calculez sa limite.

(1) $k \mapsto \cos(2\pi k)$
(2) $k \mapsto \cos(\frac{\pi}{3}k)$
(3) $k \mapsto k(a^{\frac{1}{k}} - 1)$

où a est une réel. corr0026

Correction of the exercise 22

(1) La suite $k \mapsto \cos(2k\pi)$ n'est en réalité autre que la suite constante $x_k = 1$. Elle converge donc vers 1.

(2) Parmi les sous-suites de la suite $k \mapsto \cos(\frac{\pi}{3}\pi)$, se trouvent les sous-suites constantes $x_k = \cos(\frac{\pi}{3} + 2k\pi)$ et $y_k = \cos(\frac{4\pi}{3} + 2k\pi) < 0$. La suite ne peut donc pas converger.

(3) Nous étudions la limite de la fonction

$$f(x) = x\big(a^{1/x} - 1\big) \tag{97.65}$$

lorsque $x \to \infty$. Nous tombons sur une indétermination $\infty \times 0$ qui se lève en utilisant la règle de l'Hospital. Attention : la règle de l'Hospital ne peut pas être utilisée telle quelle : il faut utiliser $fg = f/(1/g)$ afin de retrouver un cas de type $\frac{0}{0}$. Le calcul est donc le suivant :

$$\begin{aligned}\lim_{x\to\infty} x\big(a^{1/x} - 1\big) &= \lim_{x\to\infty} \frac{(a^{1/x} - 1)}{1/x}\\ &= \lim_{x\to\infty} \frac{\frac{-a^{1/x}\ln(a)}{x^2}}{-1/x^2}\\ &= \lim_{x\to\infty} a^{1/x}\ln(a)\\ &= \ln(a).\end{aligned} \tag{97.66}$$

Exercise 23 exo0027
Déterminez si la limite de chacune des suites suivantes existe et dans l'affirmative calculez la.

(1) $\lim_{k\to+\infty}(\frac{ak+1}{k})^k$
(2) $\lim_{k\to+\infty} \frac{k}{\sin(\frac{\pi}{5}k)+1} + \ln(k)\cos(\frac{\pi}{5}k)$
(3) $\lim_{k\to+\infty} \frac{\ln(k)(\sin(\frac{\pi}{3}k)+1)}{k}$
(4) $\lim_{k\to+\infty} \sqrt[3k]{k}(1 + \frac{1}{3k})^{3k}$

où a est un réel. corr0027

Correction of the exercise 23

(1) $\lim_{k\to\infty}\left(\frac{ak+1}{k}\right)^k$. Cette limite se calcule en essayant de se ramener à la limite qui définit l'exponentielle. Pour ce faire, nous calculons

$$\lim_{k\to\infty}\left(\frac{ak+1}{k}\right)^k=\lim_{k\to\infty}\left(a\big(1+\frac{1/a}{k}\big)\right)^k=e^{1/a}\lim_{k\to\infty}a^k. \tag{97.67}$$

Si $|a|<1$, alors $a^k\to 0$ et la limite est zéro. Si $a<-1$, l'alternance de signe empêche de trouver une limite. Dans ce cas, nous ne pouvons que déterminer une limite supérieure et une limite inférieure. Si $a>1$, nous avons $a^k\to\infty$ et donc la limite recherchée est l'infini.
Ici nous avons implicitement utilisé les résultats de l'exercice 11.

(2) $\lim_{k\to+\infty}\frac{k}{\sin(\frac{\pi}{5}k)+1}+\ln(k)\cos(\frac{\pi}{5}k)$. Étant donné que $\sin(\pi k/5)$ n'est jamais égal à -1, le premier terme est toujours bien défini, et toujours positif. En réalité, ce premier terme est même toujours plus grand que $\frac{k}{2}$. Le limite que nous cherchons est donc plus grande que celle de $\frac{k}{2}+\ln(k)\cos\left(\frac{\pi k}{5}\right)$. Cette dernière limite vaut l'infini par la règle de l'étau :

$$\frac{k}{2}-\ln(k)\leqslant\frac{k}{2}+\ln(k)\cos\left(\frac{k\pi}{5}\right)\leqslant\frac{k}{2}+\ln(k). \tag{97.68}$$

L'expression la plus à droite tend clairement vers l'infini, tandis que celle la plus à gauche tend également vers l'infini en vertu du travail fait en dessous de l'équation (97.59) dans la correction de la question (8) l'exercice 20.

(3) $\lim_{k\to+\infty}\frac{\ln(k)\left(\sin(\frac{\pi}{3}k)+1\right)}{k}$. Si nous majorons la quantité $\sin(x)+1$ par 2, et nous trouvons

$$\lim_{k\to\infty}\frac{\ln(k)\left(\sin(k\pi/3)+1\right)}{k}\leqslant\lim_{k\to\infty}\frac{2\ln(k)}{k}=0. \tag{97.69}$$

(4) $\lim_{k\to+\infty}\sqrt[3k]{k}(1+\frac{1}{3k})^{3k}$. D'abord, la suite $\left(1+\frac{1}{3k}\right)^{3k}$ est une sous-suite de $\left(1+\frac{1}{l}\right)^l$ qui converge vers e. Ensuite, nous avons $\sqrt[3k]{k}=\sqrt[3]{\sqrt[k]{k}}$, dont la limite est 1 en vertu de la troisième suite particulière de la page 47 du cours.
Au final, la limite est e.

97.3.1 Limites à deux variables

Exercise 24 exo0028

Déterminez si les limites suivantes existent et dans l'affirmative calculez les.

(1) $\lim_{(x,y)\to(0,0)}\frac{x^2y}{x^4+y^6+1}$
(2) $\lim_{(x,y)\to(0,0)}\frac{xy^5}{x^6+y^6}$
(3) $\lim_{(x,y)\to(0,0)}\frac{x^2-xy+3y^3}{\sqrt{x^2+y^2}}$
(4) $\lim_{(x,y)\to(0,0)}\frac{\ln(3x^4+x^2+y^2)}{x^2+y^2}$ Item0028e
(5) $\lim_{(x,y)\to(0,1)}\frac{\sin(3x^3+x^2+y^2-2y+1)}{x^2+y^2-2y+1}$
(6) $\lim_{(x,y)\to(0,0)}\frac{x-y}{\ln(x^2+y^2+1)}$
(7) $\lim_{(x,y)\to(1,1)}\frac{x-y}{\ln(x^2+y^2-1)}$
(8) $\lim_{(x,y)\to(0,0)}\frac{xy^2}{x^2+y^4}$

corr0028

Correction of the exercise 24

(1) $\lim_{(x,y)\to(0,0)}\dfrac{x^2y}{x^4+y^6+1}$

Cette fonction est le quotient de deux polynômes. Le polynôme au dénominateur est non nul en $(0,0)$, donc la fonction est continue en $(0,0)$, et

$$\lim_{(x,y)\to(0,0)}f(x,y)=f(0,0)=0. \tag{97.70}$$

Nous avons utilisé le fait qu'une fonction est continue en un point si et seulement si elle vaut sa limite en ce point.

(2) $\lim_{(x,y)\to(0,0)} \frac{x^2-xy+y^3}{\sqrt{x^2+y^2}}$

On va appliquer la règle de l'étau trois fois, aux trois fonctions :

$$\frac{x^2}{\sqrt{x^2+y^2}}, \frac{|xy|}{\sqrt{x^2+y^2}}, \frac{|y^3|}{\sqrt{x^2+y^2}}$$

$$\begin{aligned}
\frac{x^2}{\sqrt{x^2+y^2}} &\leqslant \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2} \longrightarrow_{(x,y)\to(0,0)} 0\\
\frac{|xy|}{\sqrt{x^2+y^2}} &\leqslant \frac{|y|\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}} = |y| \longrightarrow_{(x,y)\to(0,0)} 0\\
\frac{|y|^3}{\sqrt{x^2+y^2}} &\leqslant \frac{|y|(x^2+y^2)}{\sqrt{x^2+y^2}} = |y|\sqrt{x^2+y^2} \longrightarrow_{(x,y)\to(0,0)} 0
\end{aligned}$$

ce qui prouve que $\lim_{(x,y)\to(0,0)} \frac{x^2-xy+y^3}{\sqrt{x^2+y^2}} = 0$

Other resolution

Considérons une boule de rayon ϵ autour de $(0,0)$. Tout point à l'intérieur de cette boule est de la forme $(r\cos\theta, r\sin\theta)$ avec $r\in[0,\epsilon]$ et $\theta\in[0,2\pi]$. Sur un tel point, la fonction vaut

$$\begin{aligned}
\left|\frac{r^2\cos^2\theta - r^2\cos\theta\sin\theta + 3r^3\sin^3\theta}{r}\right| &= r\left|\cos^2\theta - \cos\theta\sin\theta + 3r\sin^3\theta\right|\\
&\leqslant 5r \leqslant 5\epsilon.
\end{aligned} \tag{97.71}$$

La fonction est donc bornée par 5ϵ dans la boule de rayon ϵ, et converge donc vers zéro.

(3) $\lim_{(x,y)\to(0,0)} \frac{x-y}{\ln(x^2+y^2+1)}$

Nous allons appliquer la méthode décrite au point 12.19. L'ensemble des valeurs que f atteint dans une boule de rayon δ autour de $(0,0)$ se calcule en utilisant les coordonnées polaires. En effet, tout point de cette boule s'écrit sous la forme $(x,y)=(r\cos\theta, r\sin\theta)$ pour $r\in[0,\delta]$ et $\theta\in[0,2\pi]$. Nous avons donc

$$E_\delta = \Big\{\frac{r(\cos\theta-\sin\theta)}{\ln(r^2+1)}\Big\}_{\substack{r\in]0,\delta]\\ \theta\in[0,2\pi]}}. \tag{97.72}$$

Remarquez que nous avons volontairement écrit $r\in]0,\delta]$ en excluant le 0 de l'intervalle. De toutes façons, f n'existe pas en $(0,0)$. Notons tout de suite que pour tout δ, la valeur 0 est dans E_δ, en prenant $\sin\theta=\cos\theta$.

Voyons ce que nous pouvons trouver d'autre dans cet ensemble. Il est assez naturel, pour des questions de facilité, d'essayer avec $\sin\theta=0$ et $\cos\theta=1$. Pour tout δ, l'ensemble E_δ contient tout les nombres

$$\frac{r}{\ln(r^2+1)} \tag{97.73}$$

avec $r\in]0,\delta]$. Il n'est pas très difficile, en utilisant la règle de l'Hopital de voir que

$$\lim_{r\to0}\frac{r}{\ln(r^2+1)} = \infty. \tag{97.74}$$

L'ensemble E_δ contient donc des valeurs arbitrairement élevées en même temps que 0. Il ne peut donc pas y avoir convergence.

Other resolution

Bien que $\frac{x}{\ln(x)} \longrightarrow_{x\to0^+} 1$, on va voir que la limite n'existe pas ici. Prenons deux manières de tendre vers $(0,0)$ qui donneront deux limites pour f différentes. Ceci suffira pour prouver que la fonction n'admet pas de limite en $(0,0)$.

$$\begin{array}{lll} x=0,\ y\to 0^+ & \lim_{(x,y)\to(0,0)} \frac{x-y}{\ln(x^2+y^2+1)} & = -\lim_{y\to 0^+} \frac{y}{\ln(1+y^2)} \\ & & =^H -\lim \frac{1+y^2}{2y} = -\infty \\ x\to 0^+,\ y=0 & \lim_{(x,y)\to(0,0)} \frac{x-y}{\ln(x^2+y^2+1)} & = \lim_{x\to 0^+} \frac{x}{\ln(1+x^2)} \\ & & =^H \lim \frac{1+x^2}{2x} = +\infty \end{array}$$

(4) $\lim_{(x,y)\to(1,1)} \frac{x-y}{\ln(x^2+y^2-1)}$

Passer en coordonnées polaires demande de poser $x = 1 = r\cos\theta$ et $y = 1 + r\sin\theta$. Nous trouvons donc l'expression suivante pour $f(r,\theta)$ (où le centre des coordonnées polaires est $(1,1)$, et non $(0,0)$) :

$$\frac{r(\cos\theta-\sin\theta)}{\ln\big(r^2+2r(\cos\theta+\sin\theta)+1\big)}. \tag{97.75}$$

Avec $\theta = \pi/4$, nous trouvons que cela vaut toujours 0, ce qui prouve que dans tout voisinage de $(1,1)$, la fonction $f(x,y)$ prend la valeur zéro. En prenant $\cos\theta = 1$ et $\sin\theta = 0$, nous trouvons par contre

$$f(r) = \frac{r}{\ln(r^2+2r+1)} = \frac{r}{2\ln(r+1)}, \tag{97.76}$$

donc la limite pour $r \to 0$ vaut 1. Cela prouve que, en choisissant un voisinage de $(1,1)$ assez petit, la fonction f prend des valeurs arbitrairement proches de 1. Cela rend impossible la condition (12.522) pour quelque l que ce soit.

Other resolution

Prenons deux manières de tendre vers $(1,1)$ qui donneront deux limites pour f différentes. Ceci suffira pour prouver que la fonction n'admet pas de limite en $(1,1)$.

$$\begin{array}{lll} x=1,\ y\to 1 & \lim_{(x,y)\to(1,1)} \frac{x-y}{\ln(x^2+y^2-1)} & = \lim_{y\to 1} \frac{1-y}{\ln(y^2)} \\ & & =^H \lim \frac{-1}{-2y/y^2} = \frac{1}{2} \\ x\to 1,\ y=1 & \lim_{(x,y)\to(1,1)} \frac{x-y}{\ln(x^2+y^2-1)} & = \lim_{x\to 1} \frac{x}{\ln(x^2)} \\ & & =^H \lim \frac{1}{-2x/x^2} = -\frac{1}{2} \end{array}$$

(5) $\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4}$

Prenons deux manières de tendre vers $(0,0)$ qui donneront deux limites pour f différentes. Ceci suffira pour prouver que la fonction n'admet pas de limite en $(0,0)$.

$$\begin{array}{lll} x=y^2 & \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4} & = \lim_{y\to 0} \frac{y^4}{2y^4} = \frac{1}{2} \\ x=0,\ y\to 0 & \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4} & = \lim_{y\to 0} \frac{0}{y^4} = 0 \end{array}$$

Exercise 25 exo0029

Démontrer que les limites suivantes n'existent pas :

(1) $\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^3+y^3}$

(2) $\lim_{(x,y)\to(1,1)} \tan(\pi(\frac{xy}{2}))$

(3) $\lim_{(x,y)\to(1,0)} \frac{1}{2x+y^2-2}$

(4) $\lim_{(x,y)\to(1,0)} \cos\left(\frac{1}{|x-1|y}\right)$

corr0029

Correction of the exercise 25

(1) Nous avons $f(0,y)=0$, tandis que $f(x,x)=1/2$, donc il n'y a pas convergence.

(2) La tangente de $\pi/2$ n'est pas définie, mais on sait que la limite à gauche vaut $+\infty$, tandis que la limite à droite vaut $-\infty$. Nous allons jouer là-dessus pour prouver que la limite proposée n'existe pas. Ce qu'il faut faire, c'est trouver deux chemins $\gamma_i\colon [0,1]\to\mathbb{R}^2$ tels que $\gamma_1(1)=\gamma_2(t)=(1,1)$, mais tels que $\lim_{t\to1}(f\circ\gamma_1)(t)\neq\lim_{t\to1}(f\circ\gamma_2)(t)$.
Les chemins
$$\begin{aligned}\gamma_1(t)&=(1,t)\\ \gamma_2(t)&=(1,2-t)\end{aligned}\tag{97.77}$$
font l'affaire. Toute la subtilité était de mettre $y\to1$ avec dans un cas $y>1$ et dans l'autre cas, $t<1$.

(3) Ici encore, on peut trouver un chemin qui fait tendre vers $+\infty$ et un autre qui fait tendre vers $-\infty$. Cela est d'ailleurs souvent le cas lorsqu'on est en présence d'un cas $1/0$: on peut souvent trouver un chemin qui fait $1/+0$ et un chemin qui fait $1/-0$.
Ici, nous prenons $\gamma_1(t)=(t,0)$, donc
$$f\big(\gamma_1(t)\big)=\frac{1}{2t-2}\to-\infty,\tag{97.78}$$
tandis qu'avec $\gamma_2(t)=(2-t,0)$, nous trouvons
$$(f\circ\gamma_2)(t)=\frac{1}{2(2-t)-2}\to\infty.\tag{97.79}$$

(4) Cette fois, le x ne peut rien pour faire basculer le signe du dénominateur. Nous prenons donc un chemin où $y\to0$ par les négatifs, et un chemin avec $y\to0$ par les positifs, par exemple $\gamma_1(t)=(t,1-t)$ et $\gamma_2(t)=(t,-1+t)$.

Exercise 26 exo0030

Calculer les limites suivantes si elles existent (un changement de variables peut vous simplifier la vie) :

(1) $\lim_{(x,y)\to(1,3)}\frac{\sin((x-1)^4+(x-1)^2+y^2-6y+9)}{(x-1)^2+(y-3)^2}$

(2) $\lim_{(x,y)\to(1,1)}\frac{(x-y)(x+y)}{x^3+x^2y-y^2x+y^3+2}$

(3) $\lim_{(x,y)\to(0,0)}\frac{x^2+y^2}{2x^2+y^2+x^3}$

(4) $\lim_{(x,y)\to(0,0)}\tan\left[\frac{\pi}{2}\frac{2x^2+2y^2}{x^2+y^2+x^3}\right]$

corr0030

Correction of the exercise 26

(1) Le changement de variable est assez visible : $u=x-1$ et $v=y-3$. Dans les nouvelles variables, la limite à trouver est
$$\lim_{(u,v)\to(0,0)}\frac{-\sin(u^4+u^2+v^2)}{u^2+v^2}.\tag{97.80}$$
Cet exercice est maintenant similaire à l'exercice 24(5). Afin de faire apparaitre $\sin(x)/x$, nous multiplions et divisons l'expression par $u^4+u^2+v^2$, et nous arrivons sur
$$\lim_{(u,v)\to(0,0)}\frac{u^4+u^2+v^2}{u^2+v^2}=\lim\frac{u^4}{u^2+v^2}+\underbrace{\lim\frac{u^2+v^2}{u^2+v^2}}_{=1}.\tag{97.81}$$
La première limite se calcule comme d'habitude, par passage aux coordonnées polaires, et elle vaut 0. Au final, nous nous retrouvons avec $1\cdot(1+0)=1$.

(2) Ici, le dénominateur vaut 4 en $(x, y) = (1, 1)$, donc la fonction est continue dans un voisinage de ce point. Pas de problèmes, la limite est zéro.

(3) Un passage en polaire et une simplification par r^2 amène l'expression

$$\frac{1}{\cos^2(\theta) + 1 + r\cos^3(\theta)}. \tag{97.82}$$

Remarquons que quelle que soit la valeur de r, avec $\theta = \pi/2$, nous trouvons que cela vaut 1, tandis qu'avec $\theta = 0$, cette expression est plus petite que $1/2$. Donc, dans n'importe quelle boule centrée en $(0, 0)$, la fonction prend au moins une fois la valeur 1 et une fois une valeur plus petite que $1/2$. Cela prouve que la limite n'existe pas.

(4) Regardons ce qu'il se passe dans la tangente. En passant aux coordonnées polaires, nous trouvons que

$$\lim_{(x,y)\to(0,0)} \frac{2(x^2 + y^2)}{x^2 + y^2 + x^3} = \lim_{r\to 0} \frac{2}{1 + r\cos^3(\theta)} = 2. \tag{97.83}$$

Étant donné que celle limite existe (et vaut 2), la limite demandée vaut $\tan(\pi) = 0$.

Exercise 27 exoLimSup0001

Supposons que a_i $(i = 1, 2, 3, 4)$ sont des fonctions C^∞ de $\mathbb{R}^2$ dans $\mathbb{R}$. Pour les fonctions suivantes, représentez dans le plan leur domaine de définition(s) et écrivez les limites qu'il faudra étudier pour déterminer leur continuité sur $\mathbb{R}^2$:

(1)
$$f(x, y) = \begin{cases} a_1 & \text{si } x > 0 \\ a_2 & \text{si } x \leqslant 0 \end{cases} \tag{97.84}$$

(2)
$$f(x, y) = \begin{cases} a_1 & \text{si } x > 0 \\ a_2 & \text{sinon} \end{cases} \tag{97.85}$$

(3)
$$f(x, y) = \begin{cases} a_1 & \text{si } x = y \\ a_2 & \text{sinon} \end{cases} \tag{97.86}$$

(4)
$$f(x, y) = \begin{cases} a_1 & \text{si } x > e^y \\ a_2 & \text{sinon} \end{cases} \tag{97.87}$$

(5)
$$f(x, y) = \begin{cases} a_1 & \text{si } x > 0, y > 0 \\ a_2 & \text{si } x \geqslant 0, y \leqslant 0 \\ a_3 & \text{si } x < 0, y < 0 \\ a_4 & \text{sinon} \end{cases} \tag{97.88}$$

(6)
$$f(x, y) = \begin{cases} a_1 & \text{si } xy > 0 \\ a_2 & \text{si } xy < 0 \\ a_3 & \text{sinon} \end{cases} \tag{97.89}$$

97.4 Limite et continuité

Exercise 28 exocontinueSupl1

Donner quand c'est possible un exemple de fonction

(1) continue sur $[0, 1]$

(2) continue sur $]0, 1[$ et non continue sur $[0, 1]$

(3) continue sur $[0, 1]$ telle que $f|_{]0,1]} = \frac{1}{x}$

(4) continue nulle part sur $[0, 1]$

corrcontinueSupl1

Correction of the exercise 28

(1) La fonction constante $f(x) = 0$

(2) La fonction

$$f(x) = \begin{cases} 1 & \text{si } x = 0 \text{ ou si} x = 1 \\ 0 & \text{sinon.} \end{cases} \tag{97.90}$$

(3) Impossible : quelle que soit la valeur réelle mise en zéro, il y aura toujours un point où $f(x)$ sera plus grande que cette valeur dans tout voisinage de $x = 0$.

(4) La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{si } x \notin \mathbb{Q} \end{cases} \tag{97.91}$$ EqSupl1Dirich

Celle-là, c'est un classique, ne l'oubliez pas !

Exercise 29 exocontinueSupl2

Donner une définition précise et illustrer par un exemple.

(1) f n'est pas continue en a.

(2) g est un prolongement continu de f.

(3) $\lim_{x\to a} f(x)$ n'existe pas.

corrcontinueSupl2

Correction of the exercise 29

Pour les définitions précises, voir cours. Voici des exemples.

(1) N'importe quelle fonction qui fait un saut en la valeur a, ou qui y tend vers l'infini. Par exemple $f(x) = \frac{1}{x-a}$.

(2) La fonction $f(x) = \frac{x^2-1}{x-1}$ n'existe pas en $x = 1$. La fonction $g(x) = x + 1$, toutefois, est égale à f partout et est continue en $x = 1$.

(3) La fonction (97.91) fait évidement l'affaire. Moins tordu, on peut prendre n'importe quelle fonction non continue en a.

Exercise 30 exo0031

A partir de la définition, prouver la continuité des fonctions suivantes :

$$\cos x, \sin x, e^x, \ln x$$

Indice : utiliser : $1 - nx \leqslant (1 - x)^n \leqslant 1 - \frac{nx}{1+(n-1)x}$ si $0 < x < 1$.

corr0031

Correction of the exercise 30

Pas de corrections pour cet exercice. Ça fait appel à des séries et tout ça ...

La marge est tro
pondre à ces que
une chose à cre

Au delà de l'exercice de calcul qu'il représente, cet énoncé donne l'occasion de réfléchir au sens ~~de la vie~~ de ce qu'on fait quand on démontre quelque chose. Le point crucial de cet exercice est « à partir des définitions ». Mais quelle est la définition de $\cos(x)$? Quelle est la définition de « un angle de x radians » ? Quelle est la définition de la longueur d'un arc de courbe ou de l'aire d'une surface définie par une telle courbe ?

Exercise 31 exo0032

Déterminez l'ensemble des points où les fonctions suivantes sont continues et celui où elles sont dérivables. Prouvez soigneusement vos résultats.

(1) $x \to x$

(2) $x \to |x|$

(3) $x \to \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(4) $x \to x^2$

corr0032

Correction of the exercise 31

(1) La fonction $\mathbb{R} \to \mathbb{R} : x \mapsto x$ est partout dérivable et partout continue. Cela a été démontré à titre d'exemple dans la partie rappels.

(2) La fonction $x \mapsto |x|$ est continue sur $\mathbb{R}$ et dérivable sur $\mathbb{R}\backslash\{0\}$. Commençons par montrer la non-dérivabilité en 0.

Analysons la limite suivante

$$\lim_{x\to 0} \frac{|x|-|0|}{x-0} = \begin{cases} \lim_{\substack{x\to 0\\ x>0}} \frac{|x|}{x} = \lim_{\substack{x\to 0\\ x>0}} \frac{x}{x} = \lim_{\substack{x\to 0\\ x>0}} 1 = 1 \\ \lim_{\substack{x\to 0\\ x<0}} \frac{|x|}{x} = \lim_{\substack{x\to 0\\ x<0}} \frac{-x}{x} = \lim_{\substack{x\to 0\\ x<0}} -1 = -1 \end{cases} \tag{97.92}$$

On observe que deux restrictions de cette limite fournissent deux valeurs différentes, ce qui montre que la limite n'existe pas ; la fonction n'est donc pas dérivable en 0.

Montrons la continuité en 0, c'est-à-dire

$$\lim_{x\to 0} |x| = |0| = 0$$

ce qui, par définition, équivaut à

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in \mathbb{R}\, |x-0| < \delta \Rightarrow ||x|-|0|| = |x| < \epsilon$$

qui est vérifiée pour $\delta \stackrel{\text{def}}{=} \epsilon$.

Reste à voir la dérivabilité en $x = a \neq 0$. Pour fixer les idées, prenons $a < 0$ et choisissons une boule B autour de a, complètement contenue dans $\mathbb{R}^-$ (c'est possible car $a < 0$). La notion de limite étant locale, nous avons :

$$\lim_{x\to a} \frac{|x|-|a|}{x-a} = \lim_{\substack{x\to a\\ x\in B}} \frac{-x-(-a)}{x-a} = -1$$

où on a utilisé le fait que $a, x \in \mathbb{R}^-$ pour avoir $|x| = -x$ et $|a| = -a$. Dès lors la dérivée en $a < 0$ existe et vaut -1 ; un argument similaire montre que la dérivée en $a > 0$ existe et vaut 1.

(3) Voyons que la fonction, notons-la f, n'est pas continue (donc pas dérivable) en 0. En effet, la limite

$$\lim_{\substack{x\to 0\\ x>0}} f(x) = \lim_{\substack{x\to 0\\ x>0}} \frac{1}{x} = +\infty$$

n'est pas égale à $f(0) = 0$.

Il n'est pas nécessaire de tester la dérivabilité en zéro : si elle n'est pas continue, elle n'est pas dérivable. Pour les curieux, calcul suivant prouve néanmoins directement que la fonction n'est pas dérivable en zéro :

$$\lim_{\epsilon\to 0} \frac{f(\epsilon)-f(0)}{\epsilon} = \lim_{\epsilon\to 0} \frac{\epsilon \sin(\frac{1}{\epsilon})-0}{\epsilon} = \lim_{\epsilon\to 0} \sin(\frac{1}{\epsilon}), \tag{97.93}$$

qui n'est pas définie.

Par contre, f est dérivable en tout point $a \in \mathbb{R}\backslash\{0\}$, car dans une boule assez petite autour de a, $f(x) = \frac{1}{x}$ est dérivable. Or la dérivabilité est une notion locale, donc f est dérivable en $a \neq 0$, et donc également continue.

(4) Montrons que $x \mapsto x^2$ est dérivable sur $\mathbb{R}^2$. En effet si $a \in \mathbb{R}$, alors

$$\lim_{\substack{x\to a\\ x\neq a}} \frac{x^2-a^2}{x-a} = \lim_{\substack{x\to a\\ x\neq a}} (x+a) = 2a$$

donc la dérivée existe en tout point $a \in \mathbb{R}$ et vaut $2a$.

Remarquons que cela colle avec la formule usuelle :

$$(x^2)' = 2x$$

Exercise 32 exo0033

Quelques questions sur les polynômes.

(1) Prouver que les seules fonctions polynomiales (à coefficients réels) $p[x]$ telles que, pour tout $x>0$, $p[x]/x^2 \in [2,3]$ sont de la forme $p[x]=Ax^2$ pour A une constante réelle.

(2) Donner un exemple de fonction continue $f:[0,\infty)\to\mathbb{R}^n$ qui n'est pas un polynôme mais telle que, $\forall x>0$,

$$2 \leqslant \frac{f(x)}{x^2} \leqslant 3.$$

corr0033

Correction of the exercise 32

(1) Notons

$$p(x)=a_0+a_1x+a_2x^2+a_3x^3+\cdots+a_nx^n$$

avec $a_n\neq 0$, de sorte que (en appliquant les règles de calculs)

$$\lim_{\substack{x\to 0\\ x>0}} \frac{p(x)}{x^2} = \begin{cases} \pm\infty & \text{si } a_0 \text{ ou } a_1\text{est non-nul} \\ a_2 & \text{si } a_0=a_1=0 \end{cases}$$

Dans le premier cas, le signe $\pm$ dépend des signes de a_0 et a_1. De la même manière (en mettant le terme de plus haut degré en évidence pour pouvoir calculer la limite)

$$\lim_{\substack{x\to +\infty\\ x>0}} \frac{p(x)}{x^2} = \begin{cases} +\infty & \text{si } n>2\text{et}a_n>0 \\ -\infty & \text{si } n>2 \text{ et } a_n<0 \\ a_2 & \text{si } n=2 \end{cases}$$

Ceci montre que la fonction $p(x)\div x$ prendra des valeurs arbitrairement grande (en valeur absolue) en s'approchant de 0 et de $+\infty$, sauf si $a_0=a_1=0$ et $n=2$, c'est-à-dire si $p(x)=a_2x^2$.

(2) Par exemple $x^{2\frac{\sin(x)+5}{2}}$.

Exercise 33 exo0034

Soit f une fonction continue de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Supposons que (i) $f(1)=1$ et que (ii) $f(x+y)=f(x)+f(y)$ pour tout x et y réel.

(1) Prouver que $f(x)=x$ pour tout rationnel $x>0$.

(2) Prouver que $f(0)=0$ et que $f(-x)=-f(x)$ pour tout x dans $\mathbb{R}^n$.

(3) Prouver que $f(x)=x$ pour tout x réel.

corr0034

Correction of the exercise 33

(1) Si $\frac{p}{q}$ est un quotient d'entiers strictement positifs, alors

$$\begin{aligned} qf(\frac{p}{q}) &= \underbrace{f(\frac{p}{q})+\cdots+f(\frac{p}{q})}_{q \text{ fois}} \\ &= f(\underbrace{\frac{p}{q}+\ldots+\frac{p}{q}}_{q \text{ fois}}) \\ &= f(q\frac{p}{q}) = f(p) = f(\underbrace{1+\ldots+1}_{p \text{ fois}}) \\ &= \underbrace{f(1)+\cdots+f(1)}_{p \text{ fois}} = p\,f(1) = p \end{aligned}$$

donc $f(\frac{p}{q})=\frac{p}{q}$.

(2) On trouve $f(0) = f(0+0) = f(0) + f(0) = 2f(0)$ donc $f(0) = 0$. Par ailleurs, pour $x \in \mathbb{R}$, on a

$$0 = f(0) = f(x + (-x)) = f(x) + f(-x).$$

(3) On a montré que si $r < 0$ est un rationnel, alors $-r > 0$ donc $f(-r) = -r$ d'après le premier point, or $f(r) = -f(-r)$ d'après le point précédent, donc $f(r) = r$. Par ailleurs, si x est un réel quelconque, il existe une suite $(r_1, r_2, \ldots)$ de rationnels qui tend vers x. Dès lors

$$x = \lim_{i\to+\infty} r_i = \lim_{i\to+\infty} f(r_i) = f\left(\lim_{i\to+\infty} r_i\right) = f(x),$$

parce que la continuité de f permet d'inverser la limite et f. Ceci prouve le résultat demandé.

Exercise 34 [exo0036]

Étudiez la continuité, la dérivabilité et la continuité de la dérivée de chacune des fonctions suivantes :

(1) $x \mapsto \begin{cases} \frac{2x+a}{1+e^{\frac{1}{x}}} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

(2) $x \mapsto \begin{cases} \frac{\sin(x)}{x} & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

(3) $x \mapsto \begin{cases} e^{\frac{-1}{x}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$

(4) $[-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R} : x \to \begin{cases} (\frac{\sin(2x)}{x})^{x+1} & \text{si } x \neq 0 \\ 1 & \text{sinon} \end{cases}$

où a et b sont des réels. [corr0036]

Correction of the exercise 34

Comme pour l'exercice précédent, toutes ces fonctions sont clairement continues, dérivables et à dérivée continue sur $\mathbb{R}_0$ (attention cependant à la dernière fonction !). La question intéressante est donc de savoir ce qu'il se passe en 0.

(1) Notons f cette fonction. On observe que

$$\lim_{x\to 0} f(x) = \begin{cases} \lim_{\substack{x\to 0 \\ x>0}} \frac{2x+a}{1+e^{\frac{1}{x}}} = 0 \\ \lim_{\substack{x\to 0 \\ x<0}} \frac{2x+a}{1+e^{\frac{1}{x}}} = a \end{cases}$$

en utilisant les règles de calculs. La différence entre les deux résultats provient du fait que $\lim_{x\to 0} e^{1/x}$ vaut soit ∞ soit 0 suivant que $x \to 0$ en venant par les positifs ou par les négatifs. Donc f est continue en 0 si et seulement si le paramètre $a = 0$.

Regardons la dérivabilité de f en 0 lorsque $a = 0$. On obtient :

$$\lim_{x\to 0} \frac{f(x)}{x} = \begin{cases} \lim_{\substack{x\to 0 \\ x>0}} \frac{2}{1+e^{\frac{1}{x}}} = 0 \\ \lim_{\substack{x\to 0 \\ x<0}} \frac{2}{1+e^{\frac{1}{x}}} = 2 \end{cases}$$

et donc la fonction n'est pas dérivable en 0.

(2) Notons f cette fonction. On sait déjà que

$$\lim_{\substack{x\to 0 \\ x\neq 0}} f(x) = \lim_{\substack{x\to 0 \\ x\neq 0}} \frac{\sin x}{x} = 1 = f(0)$$

ce qui prouve que f est continue en 0.

En fait f est dérivable en 0 puisque, en utilisant deux fois la règle de l'Hopital,

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{f(\epsilon)-f(0)}{\epsilon} &= \lim_{\epsilon\to 0}\frac{\sin(\epsilon)-\epsilon}{\epsilon^2}\\ &= \lim_{\epsilon\to 0}\frac{\cos(\epsilon)-1}{\epsilon}\\ &= \lim_{\epsilon\to 0}\frac{-\sin(\epsilon)}{1}\\ &= 0,\end{aligned} \tag{97.94}$$

c'est-à-dire $f'(0)=0$.

Par ailleurs si $x\neq 0$, les formules usuelles de dérivation donnent

$$f'(x)=\frac{x\cos(x)-\sin(x)}{x^2}$$

et la continuité de la dérivée revient alors à étudier la continuité de

$$f'(x)=\begin{cases}\frac{x\cos(x)-\sin(x)}{x^2} & \text{si } x\neq 0\\ 0 & \text{si } x=0\end{cases}$$

via les méthodes usuelles. Calculons donc grâce à l'Hospital :

$$\lim_{\substack{x\to 0\\ x\neq 0}}\frac{x\cos(x)-\sin(x)}{x^2}=\lim_{\substack{x\to 0\\ x\neq 0}}\frac{\cos(x)-x\sin(x)-\cos(x)}{2x}=\lim_{\substack{x\to 0\\ x\neq 0}}\frac{-\sin(x)}{2}=0$$

ce qui prouve que f' est bien continue en 0.

En fait, la fonction $\frac{\sin(x)}{x}$ est même infiniment dérivable en 0, et *analytique*.

(3) Notons f cette fonction. Montrons que f est dérivable en 0. En effet, en appliquant encore la règle de l'Hospital (notez l'astuce de calcul pour éviter de tourner en rond) :

$$\lim_{\substack{x\to 0\\ x\neq 0}}\frac{f(x)}{x}=\begin{cases}\lim_{\substack{x\to 0\\ x>0}}\frac{e^{\frac{-1}{x}}}{x}=\lim_{\substack{x\to 0\\ x>0}}\frac{\frac{1}{x}}{e^{\frac{1}{x}}}=\lim_{\substack{x\to 0\\ x>0}}\frac{1}{e^{\frac{1}{x}}}=0\\ \lim_{\substack{x\to 0\\ x<0}}0=0\end{cases}$$

ce qui prouve bien la dérivabilité en 0.

Par ailleurs, on trouve évidemment avec les règles de calculs :

$$f'(x)=\begin{cases}\frac{1}{x^2}e^{\frac{-1}{x}} & \text{si } x>0\\ 0 & \text{si } x\leqslant 0\end{cases}$$

et, toujours grâce à l'Hospital (avec la même astuce de calcul), on voit que cette fonction est également continue en 0.

En fait, cette fonction est indéfiniment dérivable, et toutes ses dérivées en 0 sont nulles malgré que la fonction elle-même soit non nulle pour $x>0$; en d'autres termes c'est une fonction qui est très plate autour de 0, mais pas constante.

(4) Au lieu de calculer $\lim f(x)$, nous calculons $\lim e^{\ln f(x)}$ (qui est égal, bien entendu) :

$$\begin{aligned}\lim_{\substack{x\to 0\\ x\neq 0}}\left(\frac{\sin(2x)}{x}\right)^{(x+1)} &= \lim_{\substack{x\to 0\\ x\neq 0}}\exp(x+1)\ln\left(\frac{\sin(2x)}{x}\right)\\ &= \exp\lim_{\substack{x\to 0\\ x\neq 0}}(x+1)\ln\left(\frac{\sin(2x)}{x}\right)\\ &= \exp\big(1\cdot\ln(2)\big)=2\end{aligned}$$

où on a utilisé la continuité des fonctions exp et ln, les règles de calculs[4] et la règle de l'Hospital. Comme $2\neq f(0)=1$, la fonction n'est pas continue en 0.

4. Es-tu capable de justifier le fait que $\lim_{x\to 0}\frac{\sin(2x)}{x}=2$?

Exercise 35 exoreserve0001

Prouver que la fonction $f\colon \mathbb{R} \to \mathbb{R}$

$$f(x) = \frac{\sinh(x)}{x} \tag{97.95}$$

est de classe C^∞. corrreserve0001

Correction of the exercise 35

Le seul problème se trouve éventuellement en $x = 0$. En utilisant la règle de l'Hospital, nous voyons très vite que f y est continue. En ce qui concerne f', nous avons

$$f'(x) = \frac{x\cosh(x) - \sinh(x)}{x^2}. \tag{97.96}$$

Encore une fois, la règle de l'Hospital nous dit que cette fonction est continue en $x = 0$.

Pour passer aux dérivées d'ordre supérieur, nous remarquons qu'elles peuvent toujours s'écrire sous la forme

$$f^{(n)}(x) = \frac{P(x)\sinh(x) + Q(x)\cosh(x)}{x^{n+1}}. \tag{97.97}$$ Eq1907fnchshQ

Cela se voit par récurrence en utilisant la règle de Leibniz pour la dérivation de produits. La limite de (97.97) lorsque $x \to 0$ se calcule en faisant $n + 1$ fois la règle de l'Hospital. À ce moment, le dénominateur est devenu 1 et le numérateur est toujours une combinaison de polynômes, de sinus et cosinus hyperboliques.

Exercise 36 exo0037

Montrez que la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x \mapsto f(x) &= \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{si } x \neq \mathbb{Q} \end{cases} \end{aligned} \tag{97.98}$$

est discontinue sur $\mathbb{R}$ et que la fonction $f\colon \mathbb{Q} \to \mathbb{R}$, $x \mapsto 1$ est continue sur $\mathbb{Q}$. corr0037

Correction of the exercise 36

Pour un même réel a, on peut trouver une suite $(q_i)_{i\in\mathbb{N}}$ de rationnels et une suite $(r_i)_{i\in\mathbb{N}}$ d'irrationnels qui tendent toutes deux vers a. Mais alors $f(q_i) = 1$ et $f(r_i) = 0$ pour tout i. Donc $f(q_i) \to 1$ et $f(r_i) \to 0$, ce qui montre que la limite

$$\lim_{x\to a} f(x)$$

n'existe pas.

Par contre, restreinte à $\mathbb{Q}$ comme proposé, la fonction devient constante... Il devient alors clair que, pour tout $a \in \mathbb{Q}$,

$$\forall \epsilon, \exists \delta : \forall x \in \mathbb{Q} : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$$

puisque $|f(x) - f(a)| = 0$ (donc la condition est satisfaite pour n'importe quel δ).

Exercise 37 exo0038

Définissons $f\colon [0,1] \to [0,1]$ comme suit. Si x est irrationnel, alors $f(x) = 0$. Si $x \neq 0$ est rationnel, posons $x = \frac{m}{n}$ avec $pgcd(m,n) = 1$, et $f(x) = \frac{1}{n}$. Finalement, $f(0) = 0$.

Prouvez que f est continue en tout irrationnel et en 0, et discontinue en tout rationnel non nul. corr0038

Correction of the exercise 37

Remarquons d'abord que si x est irrationnel ou nul, et que (q_i) est une suite de rationnels qui tend vers x, alors $f_i \stackrel{\text{def}}{=} f(q_i)$ tend vers 0.

En effet, par l'absurde, supposons qu'il existe $N \in \mathbb{N}_0$ et une sous-suite (f_{i_k}) telle que pour tout k, $f_{i_k} > \frac{1}{N}$. Mais alors $N!q_{i_k}$ est toujours un entier (car le dénominateur de q_{i_k} est inférieur à N, donc divise $N!$) et

$$\lim_{k\to\infty} N!q_{i_k} = N!x \Rightarrow x = \frac{1}{N!}\lim_{k\to\infty} N!q_{i_k}$$

ce qui montre que x est rationnel, car une suite d'entiers convergente est toujours constante à partir d'un certain rang (et possède une limite entière). Ceci est une contradiction avec « x irrationnel ». Par ailleurs, si $x = 0$, alors $q_{i_k} = 0$ à partir d'un certain k, mais alors $f_{i_k} = f(q_{i_k}) = 0$ est une contradiction avec $f_{i_k} > \frac{1}{N}$. D'où la thèse.

Il est maintenant clair que si a est irrationnel ou nul, la limite

$$\lim_{x\to a} f(x) = \begin{cases} \lim_{\substack{x\to a\\ x\in\mathbb{Q}}} f(x) = 0 \\ \lim_{\substack{x\to a\\ x\in\mathbb{R}\setminus\mathbb{Q}}} f(x) = \lim_{\substack{x\to a\\ x\in\mathbb{R}\setminus\mathbb{Q}}} 0 = 0 \end{cases}$$

est nulle, donc f est continue en a. Par ailleurs, si a est rationnel

$$\lim_{x\to a} f(x) \text{ n'existe pas car } \quad \lim_{\substack{x\to a\\ x\in\mathbb{R}\setminus\mathbb{Q}}} f(x) = 0 \neq f(a)$$

donc f n'est pas continue en a.

Exercise 38 exo0039

Donnez un exemple d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ partout discontinue et d'une fonction non constante et partout continue $g\colon \mathbb{R} \to \mathbb{R}$ telles que $g \circ f$ est partout continue (Aide : considérez $g(x) = |x|$). corr0039

Correction of the exercise 38

Posons $g(x) = |x|$ et

$$f(x) = \begin{cases} -1 & \text{si } x \in \mathbb{Q} \\ 1 & \text{sinon.} \end{cases}$$

alors $g(f(x)) = 1$ pour tout x, donc $g \circ f$ est continue.

Exercise 39 exo0040

Considérons la fonction

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x \mapsto f(x) &= \begin{cases} x & \text{si } x \text{ est rationnel} \\ 0 & \text{si } x\text{est irationnel} \end{cases} \end{aligned} \tag{97.99}$$

Vérifiez que f est continue en 0 mais n'est ni dérivable à gauche ni dérivable à droite en 0. corr0040

Correction of the exercise 39

On a clairement pour tout $x \in \mathbb{R}$,

$$0 \leqslant |f(x)| \leqslant |x|$$

or le membre de droite tend vers 0 quand $x \to 0$, donc la règle de l'étau s'applique et $\lim_{x\to 0} f(x) = 0 = f(0)$.

Par contre

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{f(x)}{x} = \begin{cases} \lim_{\substack{x\to 0\\ x<0\\ x\in\mathbb{Q}}} \frac{f(x)}{x} = 1 \\ \lim_{\substack{x\to 0\\ x<0\\ x\in\mathbb{R}\setminus\mathbb{Q}}} \frac{f(x)}{x} = 0 \end{cases}$$

ce qui montre que la limite à gauche n'existe pas. Idem pour la limite à droite.

Exercise 40 exocontinueSup0003

Étudier la continuité des fonctions suivantes :

(1) $f(x,y)=\begin{cases} xy & \text{si } x>0 \\ x^2y^2 & \text{si } x\leqslant 0 \end{cases}$

(2) $f(x,y)=\begin{cases} \lfloor xy \rfloor & \text{si } x>0, y>0 \\ xy & \text{sinon} \end{cases}$

(3) $f(x,y)=\begin{cases} e^{x-y^2} & \text{si } x=y \\ e & \text{sinon} \end{cases}$

(4) $f(x,y)=\begin{cases} 1 & \text{si } x>e^y \\ 1-(x+y) & \text{sinon} \end{cases}$

(5) $f(x,y)=\begin{cases} xy & \text{si } x>0, y>0 \\ (x-1) & \text{si } x\geqslant 0, y\leqslant 0 \\ x(y-1) & \text{si } x<0, y<0 \\ (x-1)(y-1) & \text{sinon} \end{cases}$

Exercise 41 exocontinueSup0004

Prouver que les fonctions sinus et cosinus sont continues et dérivables. corrcontinueSup0004

Correction of the exercise 41

(1) Il suffit de prouver que ces fonctions sont continues en 0.
en effet : Supposons qu'on ait que $\lim_{x\to 0}\sin x=0$, $\lim_{x\to 0}\cos x=1$. Regardons la limite en a réel quelconque

$$\lim_{x\to a}\sin x=\lim_{t\to 0}[\sin(a+t)]=\sin a\lim_{t\to 0}[\cos t]+\cos a\lim_{t\to 0}[\sin t]=\sin a$$

en utilisant les formules d'addition des arcs. Le même raisonnement s'applique à $\cos x$.

(2) $\lim_{t\to 0}\sin t=0$?
Affirmation $|\sin x|\leqslant|x|$, $\forall x\in\mathbb{R}^n$.
en effet Si $|x|\geqslant 1$, c'est évident.
Pour les réels $|x|\leqslant 1$, considérons juste les valeurs de x positives (c'est suffisant), et regardons ce que ça donne dans le cercle trigonométrique.

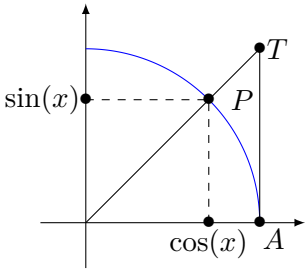


FIGURE 97.1: Dessin qui permet de calculer quelques limites. LabelFigDessinLim

L'inégalité suivante est claire sur la figure 97.1 :

$$\text{Aire triangle } OAP\leqslant\text{Aire portion de disque } OAP$$

Or, l'aire du triangle OAP est en fait $\frac{1}{2}\text{base}\times\text{hauteur}=\frac{1}{2}*1*\sin x$. Et celle de la portion de disque OAP est en fait $\frac{1}{2}r^2\theta=\frac{1}{2}*1*x$. Ceci implique immédiatement que

$$\frac{1}{2}\sin x\leqslant\frac{1}{2}x$$

ce qui nous donne le résultat.

(3) $\lim_{t\to 0}\cos t = 1$?

On sait que $1 \geqslant \cos^2 x = 1 - \sin^2 x \geqslant 1 - x^2$ par le point (ii). Donc,

$$\sqrt{1-x^2} \leqslant \cos x \leqslant 1$$

ce qui par la règle de l'étau nous donne le résultat.

(4) Les fonctions cos et sin sont dérivables partout.

Note : Il aurait été beaucoup plus efficace de prouver directement qu'elles sont dérivables et d'en déduire qu'elles sont continues.
Regardons en $a \in \mathbb{R}^n$:

$$\lim_{x\to a}\frac{\cos x - \cos a}{x-a} = \lim_{x\to a}\frac{-2\sin(\frac{x+a}{2})\sin(\frac{x-a}{2})}{x-a} = -\sin a\left[\lim_{x\to a}\frac{\sin(\frac{x-a}{2})}{\frac{x-a}{2}}\right]$$

$$\lim_{x\to a}\frac{\sin x - \sin a}{x-a} = \lim_{x\to a}\frac{2\cos(\frac{x+a}{2})\sin(\frac{x-a}{2})}{x-a} = \cos a\left[\lim_{x\to a}\frac{\sin(\frac{x-a}{2})}{\frac{x-a}{2}}\right]$$

Ceci nous montre que, à condition que nous arrivions à prouver que $\lim_{t\to 0}\frac{\sin t}{t} = 1$, nous aurons montré que

$$(\cos(x))'|_{x=a} = -\sin(a),$$
$$(\sin(x))'|_{x=a} = \cos(a),$$

ce qui prouve que cos et sin sont dérivables en tout point, et même qu'elles sont dérivables une infinité de fois et que chacune de ces dérivées est elle-même continue.
Il reste donc à prouver que $\lim_{t\to 0}\frac{\sin t}{t} = 1$. Prenons une fois de plus $x > 0$. Si nous reprenons le graphe de la figure 97.1, nous pouvons écrire les inégalités

$$\text{Aire } \Delta OAP \leqslant \text{Aire portion de disque } OAP \leqslant \text{Aire } \Delta OAT$$

Sachant que la longueur du segment AT est en fait égale à la tangente de l'angle qui le sous-tend, nous avons donc les inégalités :

$$\tfrac{1}{2}\sin x \leqslant \tfrac{1}{2}x \leqslant \tfrac{1}{2}\tan x$$

$$\Rightarrow 1 < \frac{x}{\sin x} < \frac{1}{\cos x}$$

$$\Rightarrow 1 > \frac{\sin x}{x} > \cos x$$

et cette dernière égalité permet d'utiliser l'étau et d'obtenir notre résultat.

Exercise 42 exocontinueSup0005
Prouver que la fonction $x \mapsto e^x$ est continue. corrcontinueSup0005

Correction of the exercise 42

(1) Il suffit de prouver qu'elle est continue en 0.

(2) Affirmation $\forall x \in [0,1], \forall n$,

$$1 - nx \leqslant (1-x)^n \leqslant 1 - \frac{nx}{1+(n-1)x}$$

en effet :

(2a) $1 - nx \leqslant (1-x)^n$, $\forall x \in [0,1], \forall n$, par induction directe sur n.

(2b) $(1-x)^n \leqslant 1-\frac{nx}{1+(n-1)x}$, $\forall x \in [0,1], \forall n$.
Réécrivons la autrement :

$$(1-x)^n \leqslant \tfrac{1-x}{1+(n-1)x} \iff (1+(n-1)x)(1-x)^n \leqslant 1-x$$

Et, à nouveau par induction directe sur n.

(3) $\lim_{x\to 0} e^x = 1$?
Nous allons nous restreindre, sans perte de généralité (pourquoi ?), à regarder les $x \in [0,1]$. Il suffit pour cela de montrer que $\lim_{x\to 0} e^{-x} = 1$ (pourquoi ?). Nous savons que, par définition,

$$e^{-x} = \lim_{n\to\infty} (1-\frac{x}{n})^n.$$

En utilisant l'affirmation au point (ii), $\forall n,\ \forall x \in [0,1]$

$$1-x \ \leqslant\ (1-\frac{x}{n})^n \ \leqslant\ 1-\frac{x}{1+(1-\frac{1}{n})x}.$$

et donc en passant à la limite,

$$1-x \leqslant e^{-x} \leqslant 1-\frac{x}{1+x}$$

ce qui par la règle de l'étau, suffit à prouver le résultat.

(4) e^x est dérivable une infinité de fois et chacune de ses dérivées est continue et égale à e^x.
Ceci se prouve facilement en utilisant les mêmes inégalités qu'au point précédent.

97.5 Dérivées partielles et différentiabilité

Exercise 43 exo0042

Déterminez l'ensemble des points où les fonctions sont continues et celui où elles sont différentiables.

(1) $(x,y) \mapsto 3x^2 + x^3y + x$

(2) $(x,y) \mapsto \begin{cases} e & \text{si } xy \neq 0 \\ e^{x+y} & \text{sinon} \end{cases}$

(3) $(x,y) \mapsto \begin{cases} \frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases}$

(4) $(x,y) \mapsto \begin{cases} \frac{x-y}{\ln(x^2+y^2+1)} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases}$

(5) $(x,y) \mapsto \begin{cases} x+ay & \text{si } x > 0 \\ x & \text{sinon} \end{cases}$

(6) $(x,y) \mapsto \begin{cases} \frac{xy^5}{x^6+y^6} & \text{si } x \neq y \\ 0 & \text{sinon} \end{cases}$

corr0042

Correction of the exercise 43

(1) C'est une fonction polynomiale à deux variables, elle est différentiable (et donc continue) en tout point de $\mathbb{R}^2$ par les règles de calcul. Vous devriez essayer de prouver rigoureusement que tout polynôme est différentiable.

(2) Sur l'ouvert des points (a,b) vérifiant $ab \neq 0$, la fonction est constante, donc différentiable et continue.
Il faut maintenant vérifier la continuité et la différentiabilité en les points (a,b) vérifiant $ab = 0$, c'est-à-dire que a ou b est nul. Traitons le cas $a = 0$ (le cas $b = 0$ étant identique par symétrie).
Si $a = 0$ et $b \neq 1$, alors

$$f(a,b) = f(0,b) = e^b \quad \text{et} \quad \lim_{\substack{(x,y)\to(0,b) \\ y=b}} f(x,y) = e$$

or $f(a,b) \neq e$ donc la fonction n'est pas continue.

Si $a = 0$ et $b = 1$, la fonction est continue. En effet, nous avons l'inégalité

$$0 \leqslant |f(x,y) - e| = \begin{cases} |e^{x+y} - e| \\ 0 \end{cases} \leqslant \left|e^{x+y} - e\right|$$

et le membre de droite tend vers 0 quand $(x,y) \to (0,1)$ par continuité de e^{x+y}. Dès lors $f(x,y)$ tend vers $e = f(0,1)$ par l'étau.

En ce qui concerne la différentiabilité, il suffit de voir ce qui se passe en $(0,1)$ et on voit qu'elle n'est pas différentiable en calculant les dérivées directionnelles.

(3) Autour de tout point (a,b) avec $b \neq 0$, on peut choisir une boule sur laquelle la fonction est le quotient de deux fonctions différentiables dont le dénominateur ne s'annule pas, donc est différentiable.

Pour les points du type $(a,0)$, la fonction n'est pas continue car la limite restreinte

$$\lim_{\substack{y\to 0 \\ x=a+y}} \frac{a+y}{y}$$

n'existe pas si $a \neq 0$, et vaut $1 \neq 0 = f(a,0)$ si $a = 0$.

(4) Sur $\mathbb{R}^2\backslash\{(0,0)\}$, la fonction est quotient de fonctions différentiables dont le dénominateur ne s'annule pas, donc est différentiable.

Pour voir que la fonction n'est pas continue en $(0,0)$, nous calculons grâce à la règle de l'Hospital :

$$\lim_{\substack{x\to 0 \\ y=0 \\ x<0}} \frac{x}{\ln(x^2+1)} = \lim_{\substack{x\to 0 \\ y=0 \\ x<0}} \frac{x^2+1}{2x} = -\infty \neq 0$$

ce qui prouve que la limite

$$\lim_{(x,y)\to(0,0)} \frac{x-y}{\ln(x^2+y^2+1)}$$

ne vaut pas 0.

(5) Clairement si le paramètre a est nul, la fonction (notons-la f) est différentiable (donc continue) partout. Supposons donc $a \neq 0$ dans la suite.

Par ailleurs, si on considère le point (b,c) avec $b > 0$, il existe une boule autour de (b,c) sur laquelle la fonction s'écrit $f(x,y) = x + ay$, et donc est différentiable.

De la même manière autour de (b,c) avec $b < 0$, il y a une petite boule sur laquelle la fonction s'écrit $f(x,y) = x$, et donc est différentiable.

Considérons un point de la forme $(0,c)$. Alors nous avons

$$\lim_{(x,y)\to(0,c)} f(x,y) \begin{cases} \lim_{\substack{(x,y)\to(0,c) \\ x\leqslant 0}} x = 0 \\ \lim_{\substack{(x,y)\to(0,c) \\ x>0}} x + ay = ac \end{cases}$$

donc la limite du membre de gauche existe si et seulement si $ac = 0$, c'est-à-dire si et seulement si $c = 0$.

Reste à étudier la différentiabilité en $(0,0)$. Voyons la dérivée directionnelle suivante :

$$\begin{aligned} \frac{\partial f}{\partial(1,1)}(0,0) &= \lim_{\substack{h\to 0 \\ h\neq 0}} \frac{f(0+h,0+h) - f(0,0)}{h} \\ &= \begin{cases} \lim_{\substack{h\to 0 \\ h>0}} \frac{f(h,h)}{h} = \lim_{\substack{h\to 0 \\ h>0}} \frac{h+ah}{h} = 1+a \\ \lim_{\substack{h\to 0 \\ h<0}} \frac{f(h,h)}{h} = \lim_{\substack{h\to 0 \\ h<0}} \frac{h}{h} = 1 \end{cases} \end{aligned} \tag{97.100}$$

Cette dérivée directionnelle n'existe pas, donc la fonction n'est pas différentiable en $(0,0)$.

(6) Notons f cette fonction. En les points (a,b) avec $a\neq b$, la fonction est le quotient de fonctions différentiables (dont le dénominateur ne s'annule pas) donc est différentiable.
Regardons la continuité en un point de la forme (a,a). Pour s'approcher de (a,a) sans croiser la droite $x=y$, on choisit de s'approcher via la perpendiculaire $x+y=2a$, et on calcule

$$\lim_{\substack{y\to a\\ x=2a-y\\ y\neq a}} f(x,y)=\lim_{y\to a}\frac{(2a-y)y^5}{(2a-y)^6+y^6}=\begin{cases}\frac{1}{2} & \text{si } a\neq 0\\ -\frac{1}{2} & \text{si } a=0\end{cases}$$

qui dans aucun cas n'est égale à $f(a,a)=0$, donc la fonction n'est continue en aucun de ces points.

Other resolution

(1) Cette fonction est différentiable et continue partout. En effet, la candidate différentielle est donnée par la formule

$$df_a(u)=\frac{\partial f}{\partial x}(a)u_x+\frac{\partial f}{\partial y}(a)u_y, \tag{97.101}$$

que l'on peut calculer et injecter dans la définition (10.682), afin de vérifier si ça fonctionne. Item0042b

(2) Nous testons la continuité en étudiant les limites. Lorsque $xy\neq 0$, alors cette condition est également satisfaite dans un voisinage de (x,y). Dans toute cette partie, la fonction est donc continue et différentiable (parce que constante). Nous étudions donc la continuité sur l'axe Y, c'est-à-dire

$$\lim_{(x,y)\to(0,a)} f(x,y). \tag{97.102}$$

Le long du chemin $\gamma_1(t)=(0,(1-t)a)$, la fonction vaut e^{ta}, et donc la limite est e^a. Le long du chemin $\gamma_2(t)=(1-t,a)$, la fonction est constante et vaut e. La limite ne peut donc exister que si $e=e^a$, c'est-à-dire $a=1$.
Première conclusion : sur les axes, la fonction est au mieux continue en $(0,1)$ et[5] $(1,0)$. Il n'est pas très compliqué de voir que, en fait, f est continue en ces points.
Pour la différentiabilité, elle est assurée dans la zone $xy\neq 0$. Sur les axes, elle ne peut avoir lieu que en $(1,0)$ ou $(0,1)$ parce qu'ailleurs, elle n'est pas continue. Cherchons la différentielle de f en $(1,0)$ en partant de sa dérivée directionnelle.

$$\frac{\partial f}{\partial u}(1,0)=\lim_{t\to 0}\frac{f(1+tu_1,tu_2)-e}{t}. \tag{97.103}$$

La valeur de $f(1+tu_1,tu_2)$ dépend de si u_2 est nul ou non. Ce que nous trouvons (il faut utiliser la règle de l'Hopital une fois) est

$$\frac{\partial f}{\partial u}(0,1)=\begin{cases}0 & \text{si } u_2=0\\ e(u_1+u_2) & \text{si } u_2\neq 0.\end{cases} \tag{97.104}$$ Eq0042Diffu

En posant $u=(1,0)$ et puis $u=(0,1)$, nous en déduisons en particulier les dérivées partielles

$$\begin{aligned}\frac{\partial f}{\partial x}(1,0)&=0\\ \frac{\partial f}{\partial y}(1,0)&=e.\end{aligned} \tag{97.105}$$

Par conséquent, *si* $df_{(1,0)}$ *existe*, nous avons

$$\begin{aligned}df_{(1,0)}&=edx_1\\ df_{(1,0)}(u)&=eu_1.\end{aligned} \tag{97.106}$$ Eq0042CandidatUn

5. Parce qu'il est évident que tout ce que nous faisons sur l'axe Y est valable sur l'axe X.

Par ailleurs, la formule (97.104) nous donne directement une autre expression pour la différentielle, si elle existe :

$$df_a(u) = \frac{\partial f}{\partial u}(0,1) = \begin{cases} 0 & \text{si } u_2 = 0 \\ e(u_1 + u_2) & \text{si } u_2 \neq 0. \end{cases} \quad \text{Eq0042CandidatDeux} \quad (97.107)$$

Si les expressions (97.106) et (97.107) coïncident, nous ne pouvons rien dire. Mais si elles ne sont pas les mêmes, alors nous pouvons conclure que la différentielle de f en $(0,1)$ n'existe pas.

Mais ces deux expressions ne sont manifestement pas égales sur $u = (a, 0)$ quand $a \neq 0$: la première vaut ea, tandis que la seconde vaut 0.

(3) La fonction est évidement continue et différentiable lorsque $y \neq 0$. La question est de savoir ce qu'il se passe sur les axes. L'exercice (2) était du même genre, mais ce qui sauvait deux points sur les axes, c'était le fait que la fonction était égale sur ces points à ce qu'elle vaut en dehors des axes.

Nous étudions donc ce qu'il se passe sur l'axe $y = 0$. La limite

$$\lim_{(x,y)\to(a,0)} f(x,y) \quad (97.108)$$

n'existe pas quand $a \neq 0$. En effet, si on s'approche de $(a, 0)$ en suivant un chemin horizontal, on trouve zéro, tandis qu'en suivant un chemin vertical, on trouve l'infini. De plus, même en $(0, 0)$, la limite n'existe pas, comme en témoignent les chemins $\gamma_1(t) = (t, t)$ et $\gamma_2(t) = (2t, t)$.

Cette fonction n'étant continue nulle part sur l'axe $y = 0$, il ne sert à rien de tester sa différentiabilité.

(4)

(5) Nous supposons, bien entendu que $a \neq 0$, sinon c'est de la triche. Sur les points $x \neq 0$, c'est continu et différentiable. Vérifions ce qu'il se passe en $(0, y_0)$. Voyons avec un chemin qui vient de la gauche, et un qui vient de la droite (les deux limites suivantes sont prises avec $t > 0$) :

$$\lim_{t\to 0} f\big(-t, y_0\big) = 0, \quad (97.109)$$

mais

$$\lim_{t\to 0} f\big(t, y_0\big) = ay_0 \quad (97.110)$$

Donc la fonction peut être continue en $y_0 = 0$, mais nulle part ailleurs. Il n'est pas très compliqué de vérifier que, effectivement, f est continue au point $(0, 0)$. Nous ne devons donc étudier la différentiabilité qu'en $(0, 0)$.

Si la différentielle de f en $(0, 0)$ existait, alors $\frac{\partial f}{\partial u}(0, 0)$ existerait aussi. Calculons cette dérivée directionnelle par sa définition :

$$\frac{\partial f}{\partial u}(0,0) = \lim_{t\to 0} \frac{f(tu_1, tu_2)}{t}. \quad \text{EqLimdsduf0042} \quad (97.111)$$

Si $u_1 \neq 0$, prendre la limite avec $t > 0$ donne $u_1 + au_2$, tandis que la limite avec $t < 0$ est égale à u_1. Par conséquent, la limite (97.111) n'existe pas, et la fonction n'est pas différentiable en $(0, 0)$.

(6)

Other resolution

(1) $(x, y) \to 3x^2 + x^3y + x$.

Combinaison linéaire de fonctions continues et différentiables sur $\mathbb{R}^2$ (Exercice : prouver rigoureusement que les polynômes sont bien des fonctions continues et différentiables sur $\mathbb{R}^2$).

(2)

$$(x,y) \mapsto \begin{cases} e & \text{si } xy \neq 0 \\ e^{x+y} & \text{sinon} \end{cases} \tag{97.112}$$

N.B. : Il est toujours utile de se représenter le domaine de chacune des fonctions.

La première remarque est que cette fonction est clairement continue et différentiable en tout point hors de $\{xy = 0\}$ (fonction constante). Sur $\{xy = 0\}$?

(2a) Continuité :

Prenons un point dans $\{xy = 0\}$, par exemple le point $(a, 0)$ (Remarquez que le cas $(0, b)$ est réglé par symétrie). Pour voir si la fonction est continue en ce point il faut voir si

$$\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0) = e^a.$$

Si on prend deux manières différentes d'aller vers $(a, 0)$ ($y = 0$ puis $x = a$) on voit que si $a \neq 1$ la fonction ne peut pas être continue. Et en $(1, 0)$? Si on $(x, y) \to (1, 0)$ avec d'abord $y = 0$ puis $y \neq 0$ on aura regardé toutes les manières de tendre vers $(1, 0)$. Or dans les deux cas les limites valent $e = f(1, 0)$, ce qui prouve que la fonction est continue en $(1, 0)$ (et $(0, 1)$ par symétrie).

(2b) Différentiabilité :

Comme la fonction est discontinue en tout point de la forme $(a, 0)$ avec $a \neq 1$ et $(0, b)$ avec $b \neq 1$ elle est aussi non différentiable en chacun de ces points. Il reste donc les points $(1, 0)$ et $(0, 1)$. Comme toujours, nous regardons d'abord les dérivées directionnelles en $(1, 0)$:

$$\frac{\partial f}{\partial uj}(1,0) \;=\; \lim_{t\to 0} \frac{f(1+tu_1, tu_2) - e}{t}$$

Il y a deux possibilités : $u_2 = 0$ et donc $u = (\pm 1.0)$ ou$u_2 \neq 0$ (pourquoi ne regarde-t-on que ces deux cas ?).

(2b.ii) si $u \neq (\pm 1, 0)$.
$\frac{\partial f}{\partial u}(1,0) \;=\; \lim_{t\to 0} \frac{e-e}{t} \;=\; 0.$

(2b.ii) si $u = (\pm 1, 0)$, i.e. si $u = (1, 0) = e_1$
$\frac{\partial f}{\partial u}(1,0) = \frac{\partial f}{\partial x}(1,0) = \lim_{t\to 0} \frac{f(1+t,0)-e}{t} = \lim_{t\to 0} \frac{e^{1+t}-e}{t} =^H 0.$

Conclusion :

Si f était différentiable en $(1, 0)$, on aurait que sa différentielle prendrait la forme suivante :

$$\begin{aligned} df_{(1,0)}u &= \frac{\partial f}{\partial x}(1,0)u_1 + \frac{\partial f}{\partial y}(1,0)u_2 \\ &= eu_1 \;\; \forall u \in \mathbb{R}^2 \end{aligned}$$

Sa différentielle satisferait également à :

$$df_{(1,0)}u = \frac{\partial f}{\partial u}(1,0) = 0 \;\; \forall u \neq (\pm 1, 0) \in \mathbb{R}^2$$

Les deux propriétés étant contradictoires, la fonction f ne peut être différentiable en $(1, 0)$ (ni en $(0, 1)$ par symétrie).

(3)

$$(x,y) \mapsto \begin{cases} \frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases} \tag{97.113}$$

Continue et différentiable sur $\mathbb{R} - \{y = 0\}$. Sur l'axe $y = 0$ elle n'est pas continue.

(4)

$$(x,y) \mapsto \begin{cases} x + ay & \text{si } x > 0 \\ x & \text{sinon} \end{cases} \tag{97.114}$$

Si $a = 0$ fonction continue et différentiable sur $\mathbb{R}^2$. Si $a \neq 0$, fonction continue et différentiable partout en dehors de l'axe $x = 0$. Sur cet axe, elle est discontinue en tout point sauf en $(0,0)$ où elle est continue. Mais elle n'est pas différentiable en $(0,0)$ car toutes ses dérivées directionnelles n'y sont pas définies.

(5)
$$(x,y) \mapsto \begin{cases} \frac{xy^5}{x^6+y^6} & \text{si } x \neq y \\ 0 & \text{sinon} \end{cases} \tag{97.115}$$

Fonction continue et différentiable partout en dehors de la droite $x = y$. La fonction est discontinue en chacun des points de cette droite.

Exercise 44 exo0043

Supposons que a_1, a_2, a_3 et a_4 soient des fonctions C^∞ de $\mathbb{R}^2$ dans $\mathbb{R}$. Pour les fonctions suivantes, représentez dans le plan leur domaine de définition(s) et écrivez les limites qu'il faudra étudier pour déterminer leur continuité sur $\mathbb{R}^2$:

(1)
$$f(x,y) = \begin{cases} a_1 & \text{si } x > 0 \\ a_2 & \text{si } x \leqslant 0 \end{cases} \tag{97.116}$$

(2)
$$f(x,y) = \begin{cases} a_1 & \text{si } x > 0, y > 0 \\ a_2 & \text{sinon} \end{cases} \tag{97.117}$$

(3)
$$f(x,y) = \begin{cases} a_1 & \text{si } x = y \\ a_2 & \text{sinon} \end{cases} \tag{97.118}$$

(4)
$$f(x,y) = \begin{cases} a_1 & \text{si } x > e^y \\ a_2 & \text{sinon} \end{cases} \tag{97.119}$$

(5)
$$f(x,y) = \begin{cases} a_1 & \text{si } x > 0, y > 0 \\ a_2 & \text{si } x \geqslant 0, \leqslant 0 \\ a_3 & \text{si } x < 0, y < 0 \\ a_4 & \text{sinon} \end{cases} \tag{97.120}$$

(6)
$$f(x,y) = \begin{cases} a_1 & \text{si } xy > 0 \\ a_2 & \text{si } xy < 0 \\ a_3 & \text{sinon} \end{cases} \tag{97.121}$$

Note : ici et dans ces exercices, lorsque nous écrivons « a_1 », nous sous-entendons bien entendu « $a_1(x,y)$ ». Le but de cet exercice est de permettre aux étudiants de cerner rapidement quelles sont les zones à problèmes. corr0043

Correction of the exercise 44

Il suffit de dessiner les « zones à problèmes », qui sont les frontières entre les zones (c'est-à-dire les points qui ne sont pas dans l'intérieur d'une des zones). Cela est dessiné à la figure 97.2.

Exercise 45 exo0044

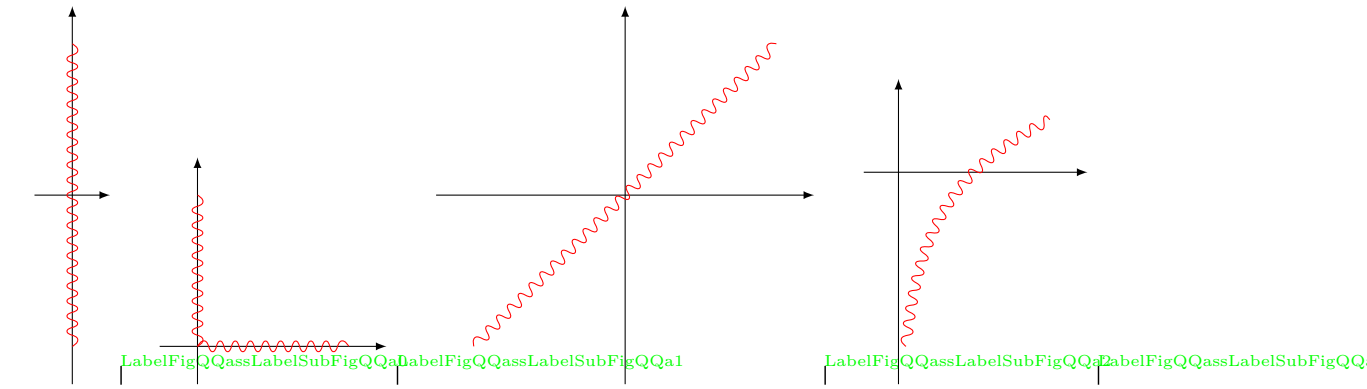
Étudier la continuité et la différentiabilité des fonctions suivantes :

(1)
$$f(x,y) = \begin{cases} xy & \text{si } x > 0 \\ x^2y^2 & \text{si } x \leqslant 0 \end{cases} \tag{97.122}$$

Item0044b

(2)
$$f(x,y) = \begin{cases} \sin(xy) & \text{si } x > 0 \\ xy & \text{sinon} \end{cases} \tag{97.123}$$

(3)
$$f(x,y) = \begin{cases} e^{x-y^2} & \text{si } x = y \\ e & \text{sinon} \end{cases} \tag{97.124}$$



(a) Exercice numéro 1 (b) Exercice numéro 2 (c) Exercice numéro 3 (d) Exercice numéro 4

(e) Exercice numéro 5 (f) Exercice numéro 6

FIGURE 97.2: Les zones à problèmes pour l'exercice 44.

LabelFigQQa

(4)

$$f(x,y) = \begin{cases} 1 & \text{si } x > e^y \\ 1-(x+y) & \text{sinon} \end{cases} \tag{97.125}$$

(5)

$$f(x,y) = \begin{cases} xy & \text{si } x>0, y>0 \\ (x-1)y & \text{si } x \geqslant 0, y \leqslant 0 \\ x(y-1) & \text{si } x<0, y<0 \\ (x-1)(y-1) & \text{sinon} \end{cases} \tag{97.126}$$

corr0044

Correction of the exercise 45

(1) La fonction est différentiable et continue en tout point hors de la droite $x=0$ car tant xy que x^2y^2 sont des fonctions polynomiales. Sur la droite $x=0$ la fonction est continue partout mais n'est différentiable qu'en $(0,0)$.

Étudions plus précisément la continuité en $(0,a)$, et pour cela regardons ce qu'il se passe dans la boule $B\big((0,a),r\big)$. D'une part, si $x>0$, alors

$$|xy| < |x| \cdot |a+r| < r|a+r| \to 0 \tag{97.127}$$

quand $r \to 0$. Et d'autre part, si $x<0$, alors

$$|x^2y^2| < |x^2| \cdot |a+r|^2 < r^2|a+r|^2 \to 0. \tag{97.128}$$

La fonction est donc continue en $(0,a)$. Faut voir maintenant si elle y est différentiable. Afin de tester la différentiabilité en $(0,a)$, nous calculons $\partial_x f(0,a)$:

$$\frac{\partial f}{\partial x}(0,a) = \lim_{t\to 0} \frac{f(t,a)-f(0,a)}{t}. \tag{97.129}$$

Si $t>0$, alors cette limite vaut $\lim_{t\to 0}\frac{ta}{t}=a$. Si par contre $t<0$, alors la limite vaut $\lim_{t\to 0}\frac{t^2a^2}{t}=0$. La dérivée partielle n'existe donc que si $a=0$. Donc f n'est différentiable en aucun point de la forme $(0,a)$ avec $a\neq 0$.

Il est assez facile de calculer que

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0, \tag{97.130}$$

donc le candidat différentielle en $(0,0)$ est $T_{(0,0)}(u)=0$. Voyons si cette forme linéaire satisfait la définition :

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y)-f(0,0)-T_{(0,0)}(x-0)}{\|(x,y)\|} = \lim_{(x,y)\to(0,0)} \frac{f(x,y)}{\sqrt{x^2+y^2}}. \tag{97.131}$$

Nous étudions cette limite en coordonnées polaires. À gauche, $f(r,\theta)=r^4\cos^2(\theta)\sin^2(\theta)$, et à droite, $f(r,\theta)=r^2\cos(\theta)\sin(\theta)$. Dans les deux cas, $\lim_{r\to 0} f(r,\theta)/r=0$.

La fonction f est donc différentiable en $(0,0)$, et sa différentielle est l'application linéaire identiquement nulle.

(2) La fonction de cet exercice est représentée sur la figure 97.3.

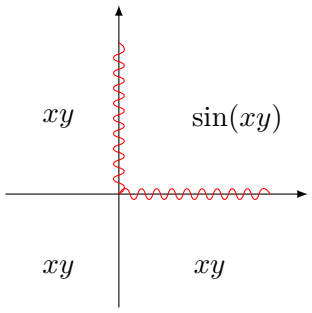


FIGURE 97.3: La fonction de l'exercice 45(2). En rouge, les zones où la fonction n'est pas trivialement différentiable. LabelFigQCb

Sur les axes, la fonction est nulle, mais $\sin(0)=0$, donc c'est une fonction continue partout. Calculons les dérivées partielles en $(a,0)$. La dérivée partielle par rapport à x ne pose pas de problèmes et vaut 0 :

$$\frac{\partial f}{\partial x}(a,0)=0. \tag{97.132}$$

Celle par rapport à y demande de traiter les cas $t<0$ et $t>0$ séparément :

$$\frac{\partial f}{\partial y}(a,0) = \lim_{t\to 0}\frac{f(a,t)}{t} = \begin{cases} \lim_{t\to 0}\frac{\sin(ta)}{t} = a\lim_{t\to 0}\frac{\sin(ta)}{ta} = a & \text{si } t>0 \\ \lim_{t\to 0}\frac{ta}{t}=a & \text{si } t<0. \end{cases} \tag{97.133}$$

Donc la candidate différentielle en $(a,0)$ est

$$T(u_1,u_2)=au_2, \tag{97.134}$$

ou encore $T=ady$. Il faut maintenant vérifier dans la définition. D'abord, $T\big((x,y)-(a,0)\big)=T(x-a,y)=ay$, et $f(a,0)=0$ donc nous vérifions si la limite suivante est zéro ou non :

$$\lim_{(x,y)\to(a,0)} \frac{f(x,y)-ay}{\sqrt{(x-a)^2+y^2}}. \tag{97.135}$$ Eqd0044

Il faut calculer différemment suivant que $y > 0$ ou $y < 0$.
Si $y > 0$, alors nous utilisons le fait que

$$\sin(xy) - ay = ay\left(\frac{\sin(ay)}{ay} - 1\right) \tag{97.136}$$

La limite (97.135) devient donc

$$\lim \frac{ay}{\sqrt{(x-a)^2 + y^2}}\left(\frac{\sin(xy)}{ay} - 1\right). \tag{97.137}$$

En utilisant les coordonnées polaires autour de $(a, 0)$, nous voyons que

$$\frac{ay}{\sqrt{(x-a)^2 + y^2}} = \frac{a}{2}\sin(\theta). \tag{97.138}$$

Pour calculer ce qui se trouve dans la parenthèse de (97.137), nous faisons

$$\frac{\sin(xy)}{ay} = \frac{x\sin(xy)}{axy} \to \frac{x}{a}. \tag{97.139}$$

En recollant les bouts,

$$\lim \underbrace{\frac{ay}{\sqrt{(x-a)^2 + y^2}}}_{=\frac{a}{2}\sin(\theta)}\Big(\underbrace{\frac{\sin(xy)}{ay}}_{\to\frac{x}{a}} - 1\Big). \tag{97.140}$$

Le tout tend donc vers zéro. Ceci conclut le calcul de (97.135) lorsque $y > 0$.
Faisons maintenant le calcul de (97.135) pour $t < 0$. En utilisant les coordonnées polaires $x = a + r\cos(\theta)$ et $y = r\sin(\theta)$, nous trouvons

$$\lim_{(x,y)\to(0,0)} \frac{xy - ay}{\sqrt{(x-a)^2 + y^2}} = \lim_{r\to 0} \frac{r^2\cos(\theta)\sin(\theta)}{r} = 0. \tag{97.141}$$

Donc la fonction considérée est différentiable en $(a, 0)$.
C'est un bon exercice d'écrire la différentielle, et de refaire tous les calculs sur le point $(0, a)$.

(3) En un point (a, b) avec $a \neq b$, la fonction est constante donc différentiable et continue.
En un point de type (a, a), on a :

$$\lim_{(x,y)\to(a,a)} f(x,y) = \begin{cases} \lim_{\substack{(x,y)\to(a,a)\\ x\neq y}} f(x,y) = e \\ \lim_{\substack{(x,y)\to(a,a)\\ x=y}} f(x,y) = e^{a-a^2} \end{cases}$$

Donc la fonction f est continue en (a, a) si et seulement si $a = a^2 + 1$, c'est-à-dire jamais.

(4) La « zone à problèmes » est formée de points du type (e^b, b). En dehors, la fonction est différentiable.
Le calcul de limite montre que la fonction est continue en (e^b, b) si et seulement si $e^b + b = 0$; cette équation possède une unique solution, notons-la b_0. Clairement $b_0 < 0$.
Pour vérifier la différentiabilité en (e^{b_0}, b_0), calculons

$$\begin{aligned} \frac{\partial f}{\partial(1,0)}(e^{b_0}, b_0) &= \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(e^{b_0}+h, b_0) - f(e^{b_0}, b_0)}{h} \\ &= \begin{cases} \lim_{\substack{h\to 0\\ h>0}} \frac{1-1}{h} = 0 \\ \lim_{\substack{h\to 0\\ h<0}} \frac{1-(e^{b_0}+h+b_0)-1}{h} = -1 \end{cases} \end{aligned} \tag{97.142}$$

où on a bien sûr utilisé le fait que $e^{b_0} + b_0 = 0$. Étant donné que cette limite, qui donne la dérivée partielle, n'existe pas la fonction n'est pas différentiable.

(5) La « zone à problèmes » est formée par la réunion des deux axes. Considérons le point (a,b) sur les axes.

Point de vue continuité : si $b=0$, alors la fonction est continue pour $a>0$, non-continue si $a<0$. Lorsque $a=0$, la fonction est continue si $b=1$, non-continue sinon.

Elle n'est nulle part différentiable sur les axes.

Exercise 46 exo0045

La figure 97.4 représente le domaine d'une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$, et sur chacune des parties, elle est définie différemment.

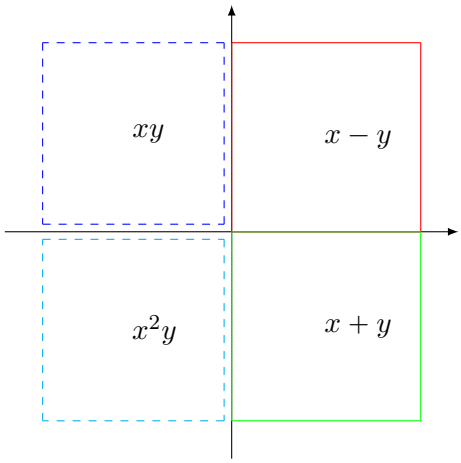


FIGURE 97.4: La fonction de l'exercice 46. LabelFigExoXLVL

Donner l'expression de cette fonction sous la forme

$$f(x,y)=\begin{cases}\dots & \text{si } \dots\\ \dots & \text{si } \dots\\ \dots & \text{si } \dots\\ \dots & \text{si } \dots\end{cases} \tag{97.143}$$

Ensuite, étudier la continuité et la différentiabilité de cette fonction sur $\mathbb{R}^2$. corr0045

Correction of the exercise 46

L'expression de f est ici

$$f(x,y)=\begin{cases}xy & \text{si } x<0, y>0\\ x-y & \text{si } x\geqslant 0, y\geqslant 0\\ x^2y & \text{si } x>0, y<0\\ x+y & \text{sinon.}\end{cases}$$

On note que les deux axes forment une zone à problèmes. La zone hors des axes est un ouvert sur lequel f est différentiable car composée de polynômes. Analysons chacun des points de la forme (a,b) dans la zone à problèmes (c'est-à-dire si $ab=0$).

Si $a=0$ et $b>0$ Un tel point $(0,b)$ est sur l'axe verticale, dans la moitié supérieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y>0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} f(x,y)=\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x\geqslant 0}} x-y=0-b=-b$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y>0\\ x<0}} xy = 0b = 0$$

qui sont différentes puisque b est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b>0$.

Si $a=0$ et $b<0$ Un tel point $(0,b)$ est sur l'axe verticale, dans la moitié inférieure. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $y<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x\geqslant 0}} x^2y = 0^2b = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} f(x,y) = \lim_{\substack{(x,y)\to(0,b)\\ y<0\\ x<0}} x+y = 0+b = b$$

qui sont différentes puisque b est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(0,b)$ avec $b<0$.

Si $a>0$ et $b=0$ Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié droite. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x>0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y\geqslant 0}} x-y = a-0 = a$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x>0\\ y<0}} x^2y = a^20 = 0$$

qui sont différentes puisque a est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a>0$.

Si $a<0$ et $b=0$ Un tel point $(a,0)$ est sur l'axe horizontal, dans la moitié gauche. Pour calculer la limite de f en ce point, on peut restreindre notre étude au demi-plan ouvert $x<0$, ce qui revient à comparer la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y>0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y>0}} xy = a0 = 0$$

avec la limite

$$\lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y\leqslant 0}} f(x,y) = \lim_{\substack{(x,y)\to(a,0)\\ x<0\\ y\leqslant 0}} x+y = a+0 = a$$

qui sont différentes puisque a est supposé non-nul.

Conclusion : f n'est pas continue en un point du type $(a,0)$ avec $a<0$.

Si $a = 0$ et $b = 0$ Le cas du point $(0,0)$ est particulier, puisque il est adhérent aux quatre composantes du domaine où la fonction est définie différemment. Pour étudier la continuité, il faut donc étudier quatre limites. Ces limites ont déjà été étudiées ci-dessus et valent toutes 0, ce qui prouve la continuité de f en $(0,0)$.

En ce qui concerne la différentiabilité, on sait qu'il est nécessaire que toutes les dérivées directionnelles existent. Calculons la dérivée dans la direction $(0,1)$ (au point $(0,0)$) :

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f((0,0)+t(0,1))-f(0,0)}{t} = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(0,t)}{t} = \dots$$

qu'on sépare en deux cas, car $f(0,t)$ possède une formule différente si $t<0$ ou si $t\geqslant 0$:

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(0,t)}{t} = \begin{cases} \lim_{\substack{t\to 0\\ t<0}} \frac{f(0,t)}{t} = \lim_{\substack{t\to 0\\ t<0}} \frac{0+t}{t} = 1 \\ \lim_{\substack{t\to 0\\ t\geqslant 0}} \frac{f(0,t)}{t} = \lim_{\substack{t\to 0\\ t\geqslant 0}} \frac{0-t}{t} = -1 \end{cases}$$

ce qui prouve que la limite n'existe pas, donc que la dérivée directionnelle n'existe pas, et finalement que la fonction n'est pas différentiable.

Conclusion : La fonction donnée est continue hors des axes et au point $(0,0)$, mais discontinue partout ailleurs sur les axes. Elle est différentiable hors des axes, mais ne l'est pas sur les axes.

Exercise 47 [exo0046]

Prouvez que toutes les dérivées directionnelles en $(0,0)$ de la fonction suivante existent mais que cette même fonction n'est pas différentiable en $(0,0)$ car elle n'est pas continue en $(0,0)$.

$$(x,y)\mapsto \begin{cases} \frac{xy^2}{x^2+y^4} & \text{si } (x,y)\neq(0,0) \\ 0 & \text{sinon} \end{cases} \tag{97.144}$$

[corr0046]

Correction of the exercise 47

Pour prouver que cette fonction n'est pas continue en $(0,0)$, il suffit de prouver que sa limite n'est pas égale à sa valeur ...ou tout simplement que la limite n'existe pas, c'est bon aussi. Si je passe en polaire,

$$f(r,\theta) = \frac{r\cos(\theta)\sin(\theta)^2}{\cos(\theta)^2 + r^2\sin(\theta)^4}. \tag{97.145}$$

Si nous prenons la limite $r\to 0$ avec θ constant, cette limite vaut zéro. Mais nous pouvons également prendre une limite sur un chemin plus subtil ...un chemin en polaire. Prenons n'importe que $r(t)\to 0$, et choisissons $\theta(t)$ tel que $\cos\big(\theta(t)\big) = r(t)$. Étant donné que $r(t)<1$, un tel θ existe toujours. Une conséquence est que $\sin\big(\theta(t)\big)\to 1$. Nous avons (nous ne recopions pas toujours toutes les dépendances en t) :

$$\begin{aligned} \lim_{t\to 0} f\big(r(t),\theta(t)\big) &= \lim_{t\to 0} \frac{r(t)\cos\big(\theta(t)\big)\sin\big(\theta(t)\big)^2}{\cos\big(\theta(t)\big)^2 + r(t)^2\sin\big(\theta(t)\big)^4} \\ &= \lim_{t\to 0} \frac{\cos(\theta)^2\sin(\theta)^2}{\cos(\theta)^2+\cos(\theta)^2\sin(\theta)^4} \\ &= \lim_{t\to 0} \frac{\sin(\theta)^2}{1+\sin(\theta)^4} \\ &= \frac{1}{2}. \end{aligned} \tag{97.146}$$

Donc, la limite en $(0,0)$ de f n'existe pas, ce qui entraîne que la fonction n'est pas continue et qu'elle n'est donc pas différentiable.

Une autre façon, plus simple si on y pense, de déduire la non continuité de f en $(0,0)$ est de regarder la limite le long du chemin $\gamma(t) = (t^2,t)$. Pourquoi ce chemin ? Parce que le dénominateur

est égal à x^2+y^4. En utilisant ce chemin, nous avons alors les même degrés dans les deux termes du dénominateur. Nous calculons

$$\lim_{t\to 0}(f\circ\gamma)(t)=\lim_{t\to 0}\frac{t^2\cdot t^2}{t^4+t^4}=\frac{1}{2}. \tag{97.147}$$

Ce résultat avait déjà été déduit dans l'exemple 12.266 du rappel théorique.

Other resolution

Soit (u,v) un vecteur fixé. Calculons la dérivées directionnelles de la fonction, au point $(0,0)$, dans cette direction :

$$\begin{aligned}\frac{\partial f}{\partial(u,v)}(0,0)&=\lim_{\substack{t\to 0\\ t\neq 0}}\frac{f(tu,tv)-f(0,0)}{t}\\ &=\lim_{\substack{t\to 0\\ t\neq 0}}\frac{tut^2v^2}{t(t^2u^2+t^4v^4)}\\ &=\lim_{\substack{t\to 0\\ t\neq 0}}\frac{uv^2}{u^2+t^2v^4}\\ &=\begin{cases}\frac{uv^2}{u^2}=\frac{v^2}{u} & \text{si } u\neq 0\\ 0 & \text{sinon}\end{cases}\end{aligned}$$

donc, quel que soit le vecteur (u,v), la dérivée directionnelle existe. Mais la fonction n'est pas continue car :

$$\lim_{\substack{(x,y)\to(0,0)\\ x=y^2\\ y\neq 0}}f(x,y)=\lim_{\substack{(x,y)\to(0,0)\\ x=y^2\\ x\neq 0}}\frac{y^2y^2}{y^4+y^4}=\lim_{\substack{(x,y)\to(0,0)\\ x=y^2\\ x\neq 0}}\frac{1}{2}=\frac{1}{2}$$

et $\frac{1}{2}\neq f(0,0)=0$.

Exercise 48 exo0047

Prouvez que la fonction suivante f est continue en $(0,0)$, que toutes les dérivées directionnelles en $(0,0)$ de cette fonction existent mais qu'elle n'est pas différentiable en $(0,0)$ car l'application qui envoie un vecteur u sur la dérivée de la fonction f dans la direction u en $(0,0)$ n'est pas linéaire.

$$f:(x,y)\to\begin{cases}x & xy>0\\ y & xy\leqslant 0\end{cases}$$

corr0047

Correction of the exercise 48

La continuité en $(0,0)$ est évidente parce que $f(r,\theta)$ est majorée par r dans la boule $B(0,r)$. Calculons la dérivée directionnelle de f dans la direction u au point $(0,0)$. Ce que nous trouvons est

$$\frac{\partial f}{\partial u}(0,0)=\lim_{t\to 0}\frac{f(tu_1,tu_2)}{t}=\begin{cases}u_1 & \text{si } u_1u_2>0\\ u_2 & \text{si } u_1u_2\leqslant 0\end{cases} \tag{97.148}$$

Voyons si cette application est linéaire par rapport à u (une application définie par morceau, c'est toujours un peu mal parti pour être linéaire). Afin d'alléger la notation, nous notons A cette application.

Nous avons

$$A\big((1,2)+(-3,4)\big)=A(-2,6)=6, \tag{97.149}$$

tandis que

$$A(1,2)+A(-3,4)=5\neq 6. \tag{97.150}$$

L'application A n'est donc pas linéaire, ce qui fait que f n'est pas différentiable en $(0,0)$.

Exercise 49 exo0048

Prouvez que la fonction suivante f est continue en $(0,0)$, que toutes les dérivées directionnelles en $(0,0)$ de cette fonction existent, que l'application qui envoie un vecteur u sur la dérivée de la fonction f dans la direction u en $(0,0)$ est linéaire mais que cette même fonction n'est pas différentiable en $(0,0)$.

$$(x,y) \to \begin{cases} \frac{x^2y\sqrt{x^2+y^2}}{x^4+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases}$$

(aide : pour vérifier que f est continue en $(0,0)$ prouvez que $\frac{x^2y}{x^4+y^2}$ est borné ; pour prouver que $\frac{x^2y}{x^4+y^2}$ est borné considérez les courbes de niveaux de f) corr0048

Correction of the exercise 49

Exercise 50 exo0049

Calculez les dérivées partielles premières par rapport à la première variable et à la seconde variable des fonctions suivantes

(1) $(u,v) \mapsto u^3 + 12u^2v - 5v^3$
(2) $(u,v) \mapsto f(u^2)\ln(v)$
(3) $(x,y) \mapsto tg(x+y^2)$
(4) $(r,\theta) \mapsto r^\theta$
(5) $(x,y) \mapsto (x+3)e^x$
(6) $(u,v) \mapsto \ln(f(uv))$

où f est une fonction de classe C^1 de $\mathbb{R}$ dans $\mathbb{R}$. corr0049

Correction of the exercise 50

(1) $(u,v) \mapsto u^3 + 12u^2v - 5v^3$.

$$\begin{aligned} \frac{\partial f}{\partial u} &= 3u^2 + 24uv \\ \frac{\partial f}{\partial v} &= 12u^2 - 15v^2 \end{aligned} \tag{97.151}$$

(2) $(u,v) \mapsto f(u^2)\ln(v)$.

$$\begin{aligned} \frac{\partial f}{\partial u} &= 2uf'(u^2)\ln(v) \\ \frac{\partial f}{\partial v} &= \frac{f(u^2)}{v} \end{aligned} \tag{97.152}$$

(3) $(x,y) \mapsto \tan(x+y^2)$.

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{cos^2(x+y^2)} \\ \frac{\partial f}{\partial v} &= \frac{2y}{cos^2(x+y^2)} \end{aligned} \tag{97.153}$$

(4) $(r,\theta) \mapsto r^\theta$

$$\begin{aligned} \frac{\partial f}{\partial r} &= \theta r^{\theta-1} \\ \frac{\partial f}{\partial \theta} &= \ln(r)r^\theta \end{aligned} \tag{97.154}$$

(5) $(x,y) \mapsto (x+3)e^x$

$$\begin{aligned} \frac{\partial f}{\partial x} &= e^x(x+4) \\ \frac{\partial f}{\partial y} &= 0 \end{aligned} \tag{97.155}$$

(6) $(u,v)\mapsto \ln(f(uv))$.

$$\begin{aligned}\frac{\partial f}{\partial u}&=\frac{vf'(uv)}{f(uv)}\\ \frac{\partial f}{\partial v}&=\frac{uf'(uv)}{f(uv)}\end{aligned}\tag{97.156}$$

Exercise 51 exo0050

Soient

$$\begin{aligned}f\colon \mathbb{R}_0^+\times\mathbb{R}_0^+&\to\mathbb{R}\\ (x,y)&\mapsto e^{x/y}\ln(\frac{x}{y})\end{aligned}\tag{97.157}$$

et

$$\begin{aligned}g\colon]0,\infty[\times]0,\frac{\pi}{2}[&\to\mathbb{R}^2\\ (r,\theta)&\mapsto(r\cos(\theta),r\sin(\theta)).\end{aligned}\tag{97.158}$$

(1) Calculez $df(1,1)$ et $dg(\sqrt{2},\frac{\pi}{4})$

(2) Calculez $\tilde{f}\stackrel{def}{=}f\circ g$

(3) Calculez $d\tilde{f}(\sqrt{2},\frac{\pi}{4})$ à partir de df et dg.

corr0050

Correction of the exercise 51

La différentielle $df_{(1,1)}$ existe parce que dans un voisinage de ce point, la fonction est parfaitement C^∞. Pour la calculer, nous calculons les dérivées partielles. D'une part

$$\frac{\partial f}{\partial x}(1,1)=\left[\frac{1}{y}e^{x/y}\ln(\frac{x}{y})+e^{x/y}\frac{1}{x/y}\frac{1}{y}\right]_{(1,1)}=e\ln(1)+e=e.\tag{97.159}$$

De la même manière,

$$\frac{\partial f}{\partial y}(1,1)=-e,\tag{97.160}$$

ce qui nous donne

$$df_{(1,1)}=e(dx-dy),\tag{97.161}$$

c'est-à-dire

$$df_{(1,1)}(u)=e(u_1-u_2).\tag{97.162}$$

Juste pour le plaisir, nous allons calculer les dérivées partielles de g en utilisant la définition, mais vous devriez remarquer qu'il est tout à fait suffisant d'appliquer les règles de calcul composante par composantes, sans vous laisser effrayer par le fait que les variables s'appellent r et θ au lieu des habituelles x et y.

$$\begin{aligned}\frac{\partial g}{\partial r}(r,\theta)&=\lim_{t\to0}\frac{g(r+t,\theta)-g(r,\theta)}{t}\\ &=\lim_{t\to0}\frac{\big((r+t)\cos(\theta),(r+t)\sin(\theta)\big)-\big(r\cos(\theta),r\sin(\theta)\big)}{t}\\ &=\lim_{t\to0}\frac{\big(t\cos(\theta),0\big)}{t}\\ &=\big(\cos(\theta),\sin(\theta)\big).\end{aligned}\tag{97.163}$$

En utilisant la même technique, ou les règles usuelles de calcul,

$$\frac{\partial g}{\partial\theta}=\big(-r\sin(\theta),r\cos(\theta)\big).\tag{97.164}$$

Magie ! Remarque que si tu fais le produit scalaire entre $\partial_\theta g$ et $g(r,\theta)$, tu obtiens zéro. En effet, g est la fonction qui transforme des coordonnées polaires en cartésiennes. Mais un déplacement

infinitésimal de θ (c'est-à-dire $\partial_\theta g$) correspond à un vecteur tangent au cercle de rayon r. Cette tangente est perpendiculaire au rayon.

Nous avons donc, au point demandé :

$$\begin{aligned}\frac{\partial g}{\partial r}(\sqrt{2},\frac{\pi}{4}) &= \left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),\\ \frac{\partial g}{\partial \theta}(\sqrt{2},\frac{\pi}{4}) &= (-1,1).\end{aligned} \tag{97.165}$$

EqDiffPArtg0050

En remplaçant $g(r,\theta)$ par sa valeur dans f, nous trouvons

$$\tilde{f}(r,\theta) = e^{1/\tan(\theta)}\ln\left(\frac{1}{\tan(\theta)}\right). \tag{97.166}$$

Remarquons que cette fonction ne dépend que de l'angle. Afin de calculer maintenant la différentielle de la composée, analysons la formule donnée dans le cours :

$$d(f\circ g)(a) = df\big(g(a)\big)\circ dg(a). \tag{97.167}$$

En notant cela avec nos notations plus habituelles,

$$d(f\circ g)_a(u) = df_{g(a)}\big(dg_a(u)\big). \tag{97.168}$$

Dans notre cas, nous avons que $g(\sqrt{2},\frac{\pi}{4}) = (1,1)$, c'est-à-dire précisément le point où nous avons déjà calculé la différentielle de f. Cool, nous n'avons pas perdu notre temps. Écrite dans le langage des formes, les équations (97.165) signifient ceci :

$$dg_{(\sqrt{2},\frac{\pi}{4})} = \begin{pmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}dr + \begin{pmatrix}-1\\ 1\end{pmatrix}d\theta. \tag{97.169}$$

En déballant un peu tout,

$$\begin{aligned}d(f\circ g)_{(\sqrt{2},\frac{\pi}{4})}(u_r,u_\theta) &= df_{g(\sqrt{2},\frac{\pi}{4})}\left(u_r\begin{pmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}+u_\theta\begin{pmatrix}-1\\ 1\end{pmatrix}\right)\\ &= (edx-edy)\begin{pmatrix}\frac{u_r}{\sqrt{2}}-u_\theta\\ \frac{u_r}{\sqrt{2}}+u_\theta\end{pmatrix}\\ &= e\left(\frac{u_r}{\sqrt{2}}-u_\theta\right)-e\left(\frac{u_r}{\sqrt{2}}+u_\theta\right)\\ &= -2eu_\theta.\end{aligned} \tag{97.170}$$

Exercise 52 exo0051

Calculez les différentielles en $(0,0)$ des fonctions suivantes :

(1) $f_1(x,y) = \begin{cases}\dfrac{\sin(\pi(15x^3+x^2+(y-2)^2))}{\pi(x^2+(y-2)^2)} & \text{quand le dénominateur est non nul}\\ 1 & \text{ailleurs}\end{cases}$

(2) $f_2(x,y) = (x+1)^{x+y}$

(3) $g_1 : (\mathbb{R}^n)^2 \longrightarrow (\mathbb{R}^n)^2 : (u,v) \longrightarrow (\ln(\sin^2(u)+1), uv)$

(4) $g_2 : \mathbb{R} \longrightarrow \mathbb{R}^3 : (u,v) \longrightarrow (\ln(\tan^2(u)+1), uv-\pi, \cos(\cos(uv)))$

(5) $h = f_1\circ g_1$

corr0051

Correction of the exercise 52

(1) Le dénominateur ne s'annule qu'au point $(0,2)$. Étudions ce qu'il se passe ailleurs. Les dérivées partielles par rapport à x et y ne sont pas compliquées à calculer (quoiqu'un peu long) :

$$\frac{\partial f_1}{\partial x}(0,0) = 0, \tag{97.171}$$

et

$$\frac{\partial f_1}{\partial y}(0,0) = -1, \tag{97.172}$$

donc si la différentielle existe (et elle existe),

$$df_{1(0,0)} = -dy. \tag{97.173}$$

(2) $f_2(x,y) = (x+1)^{x+y}$. Nous avons

$$\begin{aligned}\frac{\partial f_2}{\partial x}(x,y) &= (x+1)^{x+y}\left(\frac{x+y}{x+1} + \ln(x+1)\right),\\ \frac{\partial f_2}{\partial y}(x,y) &= (x+1)^{x+1}\ln(x+1),\end{aligned} \tag{97.174}$$

donc

$$df_{2(0,0)} = 0. \tag{97.175}$$

Exercise 53 exo0052

Donner les équations des plans tangents en $(0,0)$ aux graphes des fonctions f_1, f_2 et h de l'exercice 52. corr0052

Correction of the exercise 53

Exercise 54 exo0053

Soit f une fonction C^1 de $\mathbb{R}^2$ vers $\mathbb{R}$. Donnez une expression des dérivées partielles de la fonction g, définie ci-dessous, en fonction des dérivées partielles de la fonction f.

$$g : \mathbb{R}^2 \to \mathbb{R} : (u,v) \to f(ue^v, v(1+u^2))$$

corr0053

Correction of the exercise 54

Exercise 55 exo0054

Soit f une fonction C^1 de $\mathbb{R}$ vers $\mathbb{R}$. Donnez une expression de la fonction suivante en fonction de la dérivée de f.

$$(x,y) \to y(\partial_x g)(x,y) + x(\partial_y g)(x,y)$$

si

$$g : \mathbb{R}^2 \to \mathbb{R} : (x,y) \to f(x^2 - y^2)$$

corr0054

Correction of the exercise 55

Exercise 56 exo0055

Soient $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe C^2 et soit $g : \mathbb{R} \to \mathbb{R}$ une autre fonction de classe C^2. On définit la fonction $h : \mathbb{R} \to \mathbb{R}$ par la formule suivante

$$h(t) \stackrel{def}{=} f(t, g(t^2))$$

Calculez la dérivée première et la dérivée seconde de la fonction h en fonction des dérivées partielles de la fonction f et des dérivées de la fonction g. corr0055

Correction of the exercise 56

Exercise 57 exo0056

Soit f une fonction C^1 de $\mathbb{R}^2$ vers $\mathbb{R}$ et g une fonction de classe C^1 de $\mathbb{R}$ dans $\mathbb{R}$. Donnez une expression des dérivées partielles de la fonction h, définie ci-dessous, en fonction des dérivées partielles de la fonction f et des dérivées de la fonction g.

$$h : \mathbb{R}^2 \to \mathbb{R} : (u, v) \to f(g(ue^v), g(v)(1 + u^2))^{g(u+v)}$$

corr0056

Correction of the exercise 57

Exercise 58 exo0057

Soient $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}$, $h : \mathbb{R}^3 \to \mathbb{R}^2 = (h_1, h_2)$. Calculer les gradients (ou matrices Jacobiennes) des fonctions suivantes en termes des dérivées (partielles) de f, g, h.

(1) $w_1(x, y) = g(x + y, 1 - f(x^y))$.

(2) $w_2(t) = f(g(1 + t, 1 - t^t))$.

(3) $w_3(x, y) = h(g(x, y), f(x + \ln(xy)), y)$.

(4) $w_4(x, y, z) = (h_1(x, y, z), h_2(x, y, z), f(x + y + z))$.

corr0057

Correction of the exercise 58

Exercise 59 exo0058

Soit

$$f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \to 4x^2 + y^2$$

(1) Calculez les dérivées partielles $(\frac{\partial}{\partial x} f)(a, b)$ et $(\frac{\partial}{\partial y} f)(a, b)$ de la fonction f par rapport à sa première et par rapport à sa seconde variable au point (a, b) Donnez en une interprétation géométrique

(2) Calculez l'équation du plan tangent au graphe de f au point $(a, b, f(a, b))$.

(3) Calculez la différentielle $(df)(a, b)$ de f au point (a, b). Quel lien y a-t-il entre ce plan tangent et le graphe de $df(a, b)$?

(4) Calculez la dérivée directionnelle de f au point $(1, 2)$ dans la direction $(\frac{1}{2}, \frac{\sqrt{3}}{2})$.

(5) Calculez le gradient ∇f de f au point (a, b). Donnez en une interprétation géométrique.

(6) Dessinez quelques ensembles de niveau de la fonction f. Calculez $\nabla f(1, 1)$ et écrire l'équation de la tangente à la courbe

$$\{(x, y) \mid 4x^2 + y^2 = 5\}$$

au point $(1, 1)$.

corr0058

Correction of the exercise 59

Les dérivées partielles sont

$$(\partial_x f)(a, b) = 8a, \quad \text{et} \quad (\partial_y f)(a, b) = 2b. \tag{97.176}$$

Nous voyons tout de suite que la fonction monte donc plus vite en suivant la direction x (coefficient 8) qu'en suivant la direction y (coefficient 2).

Le plan tangent à f en $\big(a, b, f(a, b)\big)$ est donné par

$$\begin{aligned} T_{(a,b)}(x, y) &= f(a, b) + \frac{\partial f}{\partial x}(a, b) \cdot \big((x, y) - (a, b)\big)_x + \frac{\partial f}{\partial y}(a, b) \cdot \big((x, y) - (a, b)\big)_y \\ &= 4a^2 + b^2 + 8a(x - a) + 2b(y - b) \\ &= 8ax + 2by - 4a^2 - b^2. \end{aligned} \tag{97.177}$$

Petite vérification : $T_{(a,b)}(a, b) = f(a, b)$.

La différentielle $df_{(a,b)}$ existe parce que f est polynomiale. En utilisant les dérivées partielles,

$$df_{(a,b)}(u_1,u_2)=8au_1+2bu_2. \tag{97.178}$$

En termes de matrices,

$$df_{(a,b)}=\begin{pmatrix}8a\\2b\end{pmatrix}. \tag{97.179}$$

La différentielle est l'application linéaire dont la matrice est formée des « coefficients angulaires » du plan tangent.

On demande de calculer $\frac{\partial f}{\partial u}(a,b)$ avec $u=(\frac{1}{2},\frac{\sqrt{3}}{2})$ et $(a,b)=(1,2)$. Étant donné que f est différentiable, nous avons

$$\begin{aligned}\frac{\partial f}{\partial u}(a,b)=df_{(a,b)}(u)&=\frac{\partial f}{\partial x}(a,b)u_1+\frac{\partial f}{\partial y}(a,b)u_2\\&=8au_1+2bu_2\\&=4+2\sqrt{3}.\end{aligned} \tag{97.180}$$

Le gradient est encore une variation sur le même thème :

$$\nabla f(a,b)=\begin{pmatrix}\frac{\partial f}{\partial x}(a,b)&\frac{\partial f}{\partial y}(a,b)\end{pmatrix}=\begin{pmatrix}8a&2b\end{pmatrix}, \tag{97.181}$$

c'est le vecteur qu'il faut suivre en partant de (a,b) pour voir la fonction f monter le plus vite possible. Nous avons, juste en remplaçant, que

$$\nabla f(1,1)=(8,2). \tag{97.182}$$

Afin de comprendre le lien entre la tangente à une courbe de niveau et le plan tangent, il faut lire la page 94 du cours, ou alors les notes idéologiques de la sous-section 12.25.3.

Le vecteur tangent à la courbe de niveau $4x^2+y^2=5$ en $(1,1)$ est un vecteur qui est perpendiculaire à $\nabla f(1,1)=(8,2)$. Le vecteur $1,-4$ par exemple fait l'affaire. Nous cherchons donc la droite qui passe par $(1,1)$ et dont le vecteur directeur est $(1,-4)$. L'équation de cette droite est

$$y=-4x+5. \tag{97.183}$$

Exercise 60 exo0059

Considérons la fonction :

$$f(x,y)=\begin{cases}\dfrac{\tan(\pi(15x^5+x^2+(y-2)^2))}{\pi(x^2+(y-2)^2)} & \text{quand le dénominateur est non nul}\\ 1 & \text{ailleurs}\end{cases}$$

(1) Cette fonction est-elle différentiable en $(0,0)$? Si oui, calculez $\nabla f(0,0)$ et $df_{(0,0)}(2,3)$.

(2) Écrivez l'équation du plan tangent au graphe de f en $(0,0,f(0,0))$.

(3) Soit $g:\mathbb{R}^2\longrightarrow\mathbb{R}^2:(u,v)\longrightarrow(\ln(\sin^2(u)),uv-\pi)$. Calculez $g(\frac{\pi}{2},2)$. Calculez $d(f\circ g)_{(\frac{\pi}{2},2)}$ à partir de df et dg.

(4) Étudiez la continuité et la différentiabilité de f.

corr0059

Correction of the exercise 60

Exercise 61 exo0060

Soit $F\colon\mathbb{R}^2\to\mathbb{R}^2$, $(x,t)\mapsto F(x,t)$ une fonction de classe C^2. On dit que F satisfait l'équation des cordes vibrantes (ou équation des ondes), si

$$\frac{\partial^2F}{\partial t^2}=c^2\frac{\partial^2F}{\partial x^2} \tag{97.184}$$

EqONde0060

où c représente la vitesse de propagation de l'onde.

(1) Prouvez que si f et g sont deux solutions de l'équation et a et b deux réels alors $af + bg$ est encore une solution de l'équation. (l'ensemble des solutions de l'équation est un sous-espace vectoriel de l'espace vectoriel des fonctions C^2 de $\mathbb{R}^2$ vers $\mathbb{R}$)

(2) Vérifiez que si ϕ et ψ sont deux fonctions de classe C^2 de $\mathbb{R}$ dans $\mathbb{R}$ alors

$$F : \mathbb{R}^2 \to \mathbb{R} : (x,t) \to \phi(x+ct) + \psi(x-ct)$$

est une solution de l'équation.

(3) Soit le changement de variable

$$\begin{cases} \xi = x + ct \\ \eta = x - ct \end{cases}$$

Écrivez l'équation dans ces nouvelles variables et prouvez que les solutions exhibées au point précédent sont les seules possibles.

(4) Si vous êtes étudiant en physique, vous devriez être capable d'expliquer pourquoi le paramètre c de l'équation d'onde est la vitesse de l'onde. Si vous êtes étudiant en mathématique, cela ne vous dispense moralement pas de vous poser la question. Remarquez l'analogie entre l'équation d'onde (97.184) et l'équation

$$\Box\psi(\bar{x},t) = 0 \tag{97.185}$$

où $\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta$.

corr0060

Correction of the exercise 61

Le fait que l'ensemble des solutions soit un espace vectoriel découle de la linéarité de l'opération de dérivation.

Calculons maintenant les dérivées par rapport à x et à t $F(x,t) = \phi(x+ct) + \psi(x-ct)$. Cela est un exercice de dérivation partielle usuelle :

$$\begin{aligned} \frac{\partial F}{\partial x}(x,t) &= \phi'(x+ct) + \psi(x-ct), \\ \frac{\partial^2 F}{\partial x^2}(x,t) &= \phi''(x+ct) + \psi''(x-ct). \end{aligned} \tag{97.186}$$

De la même manière, mais en tenant compte du signe et du coefficient c, nous trouvons

$$\begin{aligned} \frac{\partial F}{\partial t}(x,t) &= c\phi'(x+ct) - c\psi(x-ct), \\ \frac{\partial^2 F}{\partial t^2}(x,t) &= c^2\phi''(x+ct) + c^2\psi''(x-ct). \end{aligned} \tag{97.187}$$

En comparant, nous trouvons directement que $\partial_t^2 F = c^2\partial_x^2 F$.

Posons

$$G(\xi,\eta) = F\big(\frac{1}{2}(\xi+\eta), \frac{1}{2c}(\xi-\eta)\big), \tag{97.188}$$

tandis que, pour simplifier la notation, nous notons $A = \big(\frac{1}{2}(\xi+\eta), \frac{1}{2c}(\xi-\eta)\big)$. Toujours exercice de dérivation partielle :

$$\frac{\partial G}{\partial \xi}(\xi,\eta) = \frac{\partial F}{\partial x}(A) \cdot \frac{1}{2} + \frac{\partial F}{\partial t}(A) \cdot \frac{1}{2c}, \tag{97.189}$$

et puis

$$\frac{\partial^2 G}{\partial\eta\partial\xi}(\xi,\eta) = \frac{1}{4}\frac{\partial^2 F}{\partial x^2}(A) - \frac{1}{4c^2}\frac{\partial^2 F}{\partial t^2}(A). \tag{97.190}$$

L'équation d'onde est donc en fait exactement

$$\frac{\partial^2 G}{\partial\eta\partial\xi} = 0. \tag{97.191}$$

EqOndeNouvVariable

En particulier, $\partial_\xi G$ est une constante par rapport à η, nous écrivons donc

$$\frac{\partial G}{\partial \xi} = a(\xi). \tag{97.192}$$

Étant donné que G est C^2, nous savons que a est C^1 ; par conséquent, cette fonction admet une primitive C^2 que nous notons ϕ. L'équation d'onde dans les nouvelles variables (97.191) a pour solution

$$G(\xi, \eta) = \phi(\xi) + C \tag{97.193}$$

où C est une constante *par rapport à* ξ, c'est-à-dire que C peut dépendre de η. Au final,

$$G(\xi, \eta) = \phi(\xi) + \psi(\eta), \tag{97.194}$$

ce qu'il fallait démontrer.

Exercise 62 exo0061

Soit $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \to f(x, y)$ une fonction de classe C^2. On définit le **laplacien** de f par la formule suivante :

$$\Delta f \stackrel{def}{=} \frac{\partial^2}{\partial x^2} f + \frac{\partial^2}{\partial y^2} f$$

(1) Calculez le laplacien en coordonnées polaires. Autrement dit, posez

$$\tilde{f}(r, \theta) \stackrel{def}{=} f(r\cos(\theta), r\sin(\theta)) \tag{97.195}$$

et prouvez que :

$$(\Delta f)(r\cos(\theta), r\sin(\theta)) = (\frac{\partial^2}{\partial r^2}\tilde{f})(r, \theta) + \frac{1}{r}(\frac{\partial}{\partial r}\tilde{f})(r, \theta) + \frac{1}{r^2}(\frac{\partial^2}{\partial \theta^2}\tilde{f})(r, \theta)$$

(2) Vérifiez que la fonction suivante $F : (x, y) \to \ln(\sqrt{x^2 + y^2})$ est une solution de l'équation aux dérivée partielles $\Delta f = 0$

corr0061

Correction of the exercise 62

97.6 Séries et séries de puissances

Exercise 63 exo0062

Donner quand c'est possible un exemple de série ne contenant pas une infinité de termes nuls (sinon c'est facile)

(1) absolument convergente,

(2) dont la somme vaut deux,

(3) de somme infinie mais dont le terme principal tend vers 0,

(4) divergente (et pas de somme infinie),

(5) qui converge simplement mais pas absolument,

(6) convergente mais dont le terme principal ne tend pas vers 0.

corr0062

Correction of the exercise 63

(1) La série

$$\sum_{i=1}^{\infty} \frac{1}{n^2}$$

converge absolument.

(2) On connaît la somme de la série géométrique de raison q, on en déduit que

$$\sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{1-\frac{1}{2}} = 2$$

comme demandé.

(3) La série harmonique

$$\sum_{i=0}^{\infty} \frac{1}{i}$$

est l'exemple standard de cette situation.

(4) Un exemple simple est

$$\sum_{i=0}^{\infty} (-1)^i$$

dont la suite des sommes partielles est $(1, 0, 1, 0, 1, \ldots)$.

(5) La série harmonique alternée

$$\sum_{i=0}^{\infty} (-1)^i \frac{1}{i}$$

est un bel exemple.

(6) Ceci est impossible : si la suite des sommes partielles converge, elle doit être de Cauchy, et en particulier la différence entre deux de ses termes successifs, c'est-à-dire le terme général, doit pouvoir être rendue aussi petite que voulu (voir la proposition 10.100).

Exercise 64 exo0063

Supposer que $\sum_{n=1}^{\infty} a_n$ est une série convergente à termes positifs. Montrer qu'il existe une série $\sum_{n=1}^{\infty} \tilde{a}_n$ telle que

$$\tilde{a}_n > 0, \quad \lim_{n\to\infty} \frac{\tilde{a}_n}{a_n} = 0, \quad \text{et} \quad \sum_{n=1}^{\infty} \tilde{a}_n < \infty. \tag{97.196}$$

(Dans l'ensemble des séries convergentes à termes positifs, il n'y a pas de plus lent) corr0063

Correction of the exercise 64

La suite $\tilde{a}_n = a_n/n$ fait l'affaire.

Exercise 65 exo0064

Montrez que la série convergente

$$\frac{1}{2} - \frac{1}{2} + \frac{1}{3} - \frac{1}{3} + \frac{1}{4} - \frac{1}{4} + \cdots \tag{97.197}$$

admet un réarrangement divergent

$$\frac{1}{2} - \frac{1}{2} + \frac{1}{3} + \frac{1}{4} - \frac{1}{3} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} - \frac{1}{4} + \cdots$$

corr0064

Correction of the exercise 65

Si h_k désigne le $k$ième terme de la série harmonique, la somme proposée n'est autre que $\sum_k (h_k - h_k)$. Nous pouvons le réarranger en espaçant les $-h_k$ de telle manière à laisser la série monter de 1 entre deux arrivée d'un terme négatif. Plus précisément,

$$h_1, \ldots, h_{S_1}, -h_1, h_{S_1+1}, \ldots, h_{S_2}, -h_2, h_{S_2+1}, \ldots, h_{S_3}, \ldots \tag{97.198}$$ EqSerieArran0064

où les indices S_k sont définis de telle sorte à avoir

$$(\sum_{i=S_k}^{S_{k+1}} h_i) - h_k > 1. \tag{97.199}$$

De cette façon, la série des sommes partielles de (97.198) est divergente.

Exercise 66 exo0065

Soient $\sum a_n$ une série convergente dans $\mathbb{R}$ et $\sigma\colon \mathbb{N} \to \mathbb{N}$ une fonction injective. Posons $b_n = a_{\sigma(n)}$. $\sum b_n$ converge telle toujours dans $\mathbb{R}$? Si non, quand a-t-on cette convergence ? corr0065

Correction of the exercise 66

Prenons $\sum_k \frac{(-1)^k}{k}$. Cette somme converge (critère des séries alternées). Considérons maintenant l'application injective $\sigma(k) = 2k$. La série $\sum_k \frac{(-1)^{\sigma(k)}}{\sigma(k)}$, par contre, diverge (le montrer !).

Exercise 67 exo0066

Déterminez si les séries ci-dessous sont absolument convergentes, convergentes ou divergentes.

(1) $\sum_{k=1}^{\infty} \frac{1}{3^k}$
(2) $\sum_{k=2}^{\infty} \frac{1}{(\ln(k))^k}$
(3) $\sum_{k=1}^{\infty} \frac{k!}{k^k}$
(4) $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$
(5) $\sum_{k=1}^{\infty} \frac{1}{k^2-\cos(k)}$
(6) $\sum_{k=1}^{\infty} k^3 e^{-3k}$
(7) $\sum_{k=1}^{\infty} \frac{(-1)^k \ln(k)}{k}$
(8) $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^a}$
(9) $\sum_{k=1}^{\infty} \frac{\cos(k)+i\sin(k)}{k^2}$
(10) $\sum_{k=1}^{\infty} \frac{\cos(\frac{k\pi}{2})+i\sin(\frac{k\pi}{2})}{k^a}$

où a est un nombre réel. corr0066

Correction of the exercise 67

(1) $\sum_k (1/3^k)$. C'est une série de puissances avec $q = 1/3$. Étant donné que $|q| < 1$, nous avons convergence absolue.

(2) $\sum_k (1/\ln(k)^k)$. Cette somme est toute désignée pour faire fonctionner le critère de la racine. Nous avons

$$\sqrt[k]{\frac{1}{\ln(k)^k}} = \frac{1}{\ln(k)}, \tag{97.200}$$

et bien entendu, la limite supérieure de cette suite est zéro. Donc, convergence.

Other resolution

Nous pouvons aussi dire que

$$\frac{1}{\ln(k)^k} < \frac{1}{\ln(3)^k}, \tag{97.201}$$

et maintenant nous pouvons comparer avec une série de puissance de raison $1/\ln(3)$.

Remarquez que la majoration $1/\ln(k)^k < 1/\ln(2)^k$ n'est pas suffisante parce que $\ln(2) < 1$. En effet, le logarithme est croissant, et $\ln(e) = 1$, alors que $e > 2$ en vertu de l'exercice 17(2).

(3) $\sum_k (k!/k^k)$. Une fois n'est pas coutume, nous avons une suite qui croît plus vite que la factorielle. Appliquons le critère du quotient :

$$\frac{a_{k+1}}{a_k} = \frac{(k+)!}{(k+1)^{k+1}} \frac{k^k}{k!} = \frac{(k+1)k^k}{(k+1)^{k+1}} = \frac{k^k}{(k+1)^k} = \left(\frac{k}{k+1}\right)^k. \tag{97.202}$$

Le truc est maintenant de voir que cette suite n'est pas loin d'être la suite qui définit le nombre e, pour rappel

$$\left(1 + \frac{1}{k}\right)^k \to e. \tag{97.203}$$

Nous avons

$$\left(\frac{k}{k+1}\right)^k = \left(\frac{k+1}{k}\right)^{-k} = \left[\left(\frac{k+1}{k}\right)^k\right]^{-1} \to 1/e < 1. \tag{97.204}$$

Donc la série converge absolument.

(4) $\sum_k (1/k(k+1))$. Chaque terme de cette série est plus petit que $\frac{1}{k^2}$. Or la série des $1/k^2$ converge, donc le critère de comparaison donne la convergence.

(5) $\sum_k (1/(k^2 - \cos(k)))$. D'abord, remarquons que tous les termes de cette série sont positif. Ensuite, Nous avons

$$\frac{1}{|k^2 - \cos(k)|} < \frac{1}{|k^2 - 1|} < \frac{1}{(k-1)^2}. \tag{97.205}$$

La dernière inégalité est due au fait que $k^2 - 1 = (k+1)(k-1)$.

(6) $\sum_k (k^3 e^{-3k})$. Nous faisons le quotient :

$$\frac{(k+1)^4 e^{-3(k+1)}}{k^3 e^{-3k}} = e^{-3} \left(\frac{k+1}{k} \right)^3 \to e^{-3} < 1, \tag{97.206}$$

donc la série converge.

(7) $\sum_k \big((-1)^k \ln(k)\big)/k$. C'est une série alternée construite sur une série dont le terme général tend vers zéro. Il y a donc convergence. Il faut voir maintenant si la convergence est absolue ou non. Cela est vite réglé par comparaison :

$$\sum_{k=1}^{\infty} \frac{\ln(k)}{k} > \sum_k \frac{1}{k}, \tag{97.207}$$

qui diverge. Il y a donc convergence simple mais pas absolue.

(8) $\sum_k \big((-1)^k / k^a\big)$. Lorsque $a < 0$, la série diverge. Cela ne veut pas dire, cependant, qu'elle tend vers $\pm\infty$. Lorsque $a > 0$, il y a convergence simple par le critère des séries alternées. Il faut encore étudier dans quel cas il y a convergence absolue. Pour ce faire, nous regardons la série non alternée

$$\sum_{k=1}^{\infty} \frac{1}{k^a}, \tag{97.208}$$

qui converge si et seulement si $a > 1$. En résumé, nous avons

(8a) Diverge si $a < 0$,

(8b) Converge simplement mais pas absolument si $0 \leqslant a \leqslant 1$,

(8c) converge absolument quand $a > 1$.

(9) $\sum_k (\cos(k) + i \sin(k))/k^2$. La norme du numérateur vaut tout le temps 1, donc la série converge absolument, et donc simplement.

(10) La série converge absolument quand $a > 1$. Si $a \leqslant 0$, alors la série diverge parce que le terme général ne tend pas vers zéro. Lorsque $0 < a \leqslant 1$, il n'y a pas de convergence absolue parce que nous tombons sur la série de Riemann (10.349). Par contre, le critère d'Abel assure la convergence simple sur cet intervalle. En résumé :

(10a) si $a \leqslant 0$, alors la série diverge,

(10b) si $0 < a \leqslant 1$, il y a convergence simple,

(10c) si $a > 1$, alors il y a convergence absolue.

Exercise 68 exo0067

Déterminez et représentez dans le plan de Gauss le domaine de convergence des séries suivantes :

(1) $\sum_{k=1}^{\infty} z^k$

(2) $\sum_{k=1}^{\infty} \frac{(z+1)^k}{k}$

(3) $\sum_{k=1}^{\infty} \frac{k!(z+i)^k}{k^k}$

(4) $\sum_{k=1}^{\infty} \frac{z^k}{k!}$

(5) $\sum_{k=1}^{\infty} (k+1)^a (z-5+i)^k$

(6) $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}\sqrt{k+3}(z+1-i)^k}{(k+2)^2}$

(7) $\sum_{k=1}^{\infty} \frac{(-z+2)^k}{2k+\ln(k)}$

corr0067

Correction of the exercise 68

(1) La série
$$\sum_{k=0}^{\infty} z^k$$
est une série de puissance de coefficients $c_k = 1$ et de centre $z_0 = 0$. Pour trouver le rayon de convergence, nous utilisons la formule (15.67)
$$\limsup \sqrt[k]{|c_k|} = \limsup \sqrt[k]{1} = 1.$$
Le rayon de convergence est donc donné par $R = 1/1 = 1$.
Si z vérifie $|z| = 1$, alors $\left|z^k\right| = 1$ et
$$\lim_{k\to\infty} z^k \neq 0$$
ce qui prouve que la série ne converge pas sur le bord du disque de convergence.
Remarquons que cette série est en réalité la série géométrique de raison z, dont on connaît la somme explicitement depuis la proposition 10.135.

(2) La série
$$\sum_{k=1}^{\infty} \frac{(z+1)^k}{k}$$
est une série de puissance de coefficients $c_k = \frac{1}{k}$ et centrée en $z_0 = -1$. Nous calculons
$$\alpha = \limsup \sqrt[k]{\frac{1}{k}} = \limsup \frac{1}{\sqrt[k]{k}} = 1, \tag{97.209}$$
donc le rayon de convergence est $R = 1/\alpha = 1$. Le théorème 15.25 dit tout ce qu'on veut savoir dans tout le plan complexe [6], sauf *sur* le cercle de rayon 1 centré en -1.
Étudions la convergence absolue sur le bord du disque, c'est-à-dire pour les points z tels que $|z+1| = 1$. Le terme général de la série devient
$$\frac{|z+1|^k}{k} = \frac{1}{k}, \tag{97.210}$$
qui n'est autre que le terme général de la série harmonique. Donc la série des modules (qui donne la convergence absolue) est la série harmonique. On en déduit que la série ne converge pas absolument sur le bord du disque de convergence.
Pour appliquer Abel, toujours pour z tel que $|z+1| = 1$, majorons les sommes partielles suivantes
$$\begin{aligned}
\left|\sum_{k=1}^{n} (z+1)^k\right| &= \left|-1 + \sum_{k=0}^{n} (z+1)^k\right| \\
&= \left|-1 + \frac{1-(z+1)^{n+1}}{1-(z+1)}\right| \\
&\leqslant 1 + \left|\frac{1-(z+1)^{n+1}}{-z}\right| \\
&\leqslant 1 + \frac{1+\left|(z+1)^{n+1}\right|}{|z|} \\
&= 1 + \frac{1+|(z+1)|^{n+1}}{|z|} \\
&= 1 + \frac{2}{|z|}
\end{aligned}$$

6. N'hésitez pas à faire un petit dessin de la solution, afin de vous assurer que vous voyez bien qui sont z, z_0, R et le disque de convergence.

ce qui fournit une majoration (indépendante de n) pour tout z vérifiant $z \neq 0$ et $|z+1| = 1$. On peut alors appliquer le critère de Abel, puisque la suite $\frac{1}{k}$ est clairement décroissante et tend vers 0.

Le point $z = 0$ reste à analyser. Si on récrit la série de l'énoncé avec cette valeur de z, on retrouve la série harmonique (qui est divergente).

Conclusion : Si $|z+1| > 1$ ou si $z = 0$, la série diverge ; si $|z+1| = 1$ et $z \neq 0$, la série converge simplement ; si $|z+1| < 1$, la série converge absolument (et simplement).

(3) La série

$$\sum_{k=1}^{\infty} \frac{k!(z+i)^k}{k^k} \tag{97.211}$$

eqseriepuissances-exo-c

est une série de puissances de coefficients $c_k = \frac{k!}{k^k}$ et de centre $z_0 = -i$. La formule

$$\alpha = \limsup \sqrt[k]{\frac{k!}{k^k}} \tag{97.212}$$

ne nous ragoute pas trop. Nous utilisons donc plutôt la formule alternative (15.68) :

$$\lim_{k\to\infty} \frac{(k+1)!}{(k+1)^{k+1}} \frac{k^k}{k!} = \lim_{k\to\infty} \frac{k^k}{(k+1)^k} = \lim_{k\to\infty} \frac{1}{\left(1+\frac{1}{k}\right)^k} = \frac{1}{e} \tag{97.213}$$

qui montre que le rayon de convergence est $R = e$.

Pour z tel que $|z+1| = e$, on observe que le module du terme général est donné par

$$a_k = \frac{k!e^k}{k^k} \tag{97.214}$$

Nous voulons donc vérifier si la série $\sum_k \frac{k!}{k^k} e^k$ converge. Pour ce faire, nous calculons

$$\frac{a_{k+1}}{a_k} = \frac{(k+1)ek^k}{(k+1)^{k+1}} = e\frac{k^k}{(k+1)^k}. \tag{97.215}$$

EqCalculPuiss0067ee

La suite $\big(k/(k+1)\big)^k$ tend vers 1, donc la suite des a_{k+1}/a_k tend vers 1. Nous pouvons cependant dire plus. En vertu de l'exercice 17(1), la suite $x_k = \left(\frac{k+1}{k}\right)^k$ qui définit e est monotone croissante, donc la suite (97.215) est une suite monotone décroissante qui tend vers 1. Chacun de ses termes est donc plus grand que 1.

Le fait que ce rapport soit plus grand que 1 montre que $a_{k+1} \geqslant a_k$, c'est-à-dire que la suite a_k est croissante. En particulier, le terme général de la série (97.211) ne peut pas tendre vers 0 sur le bord du disque de convergence.

Notez que nous avons bien prouvé que la série ne converge pas sur le bord, et non seulement qu'elle ne converge pas absolument.

(4) La série

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}$$

est de la forme ci-dessus, avec $c_k = 1/k!$ et $z_0 = 0$. Calculons la limite

$$\lim_{k\to\infty} \frac{c_{k+1}}{c_k} = \lim_{k\to\infty} \frac{k!}{(k+1)!} = \lim_{k\to\infty} \frac{1}{k+1} = 0$$

qui montre que la série possède un rayon de convergence infini, c'est-à-dire qu'elle converge absolument quel que soit $z \in \mathbb{C}$.

(5) La série dépendant des paramètres $a \in \mathbb{R}$ et $z \in \mathbb{C}$

$$\sum_{k=1}^{\infty} (k+1)^a (z-5+i)^k$$

est une série de puissance (en z) de coefficients $c_k = (k+1)^a$ et centrée en $z_0 = 5-i$. Le calcul de la limite

$$\lim_{k\to\infty} \frac{(k+2)^a}{(k+1)^a} = 1^a = 1$$

montre que le rayon de convergence vaut $R = 1$, indépendamment de a.

Si $|z-5+i| = 1$, alors la série des modules, qui permet d'établir la convergence absolue, devient

$$\sum_{k=1}^{\infty} (k+1)^a$$

et est manifestement équivalente à la série de Riemann $\sum_{k=1}^{\infty} k^a$. La série de l'énoncé est donc absolument convergente si et seulement si $a < -1$.

Par ailleurs, le module du terme général étant $(k+1)^a$ (toujours pour z sur le bord du disque), il tend vers 0 si et seulement si $a < 0$. On en déduit que la série diverge dès que $a \geqslant 0$.

Pour $-1 \leqslant a < 0$, le « critère de Riemann » a échoué, et nous savons qu'elle ne converge pas absolument. Nous essayons donc d'utiliser le critère d'Abel pour savoir si la série convergerait simplement. Étant donnée la condition sur a, le facteur $(k+1)^a$ forme une suite décroissant vers 0. On peut donc essayer de majorer les sommes partielles suivante :

$$\begin{aligned}
\left| \sum_{k=1}^{n} (z-5+i)^k \right| &= \left| -1 + \frac{1-(z-5+i)^{n+1}}{1-(z-5+i)} \right| \\
&\leqslant 1 + \left| \frac{1-(z-5+i)^{n+1}}{1-(z-5+i)} \right| \\
&\leqslant 1 + \frac{1+\left|(z-5+i)^{n+1}\right|}{|1-(z-5+i)|} \\
&= 1 + \frac{2}{|1-(z-5+i)|}
\end{aligned}$$

ce qui est valable dès que $1-(z-5+i) \neq 0$. On en déduit que la série converge simplement sur le disque $|z-5+i| = 1$ dès que $-1 \leqslant a < 0$ et $z \neq 6-i$.

Si $z = 6-i$, la série initiale devient la série

$$\sum_{k=1}^{\infty} (k+1)^a$$

et elle ne converge que si $a < -1$ (qui est un cas déjà étudié) et diverge sinon.

Conclusion : Indépendamment de a, le rayon de convergence vaut 1, et la série converge absolument si $|z-5+i| < 1$; diverge si $|z-5+i| > 1$. En ce qui concerne les points z vérifiant $|z-5+i| = 1$, la série y converge absolument si et seulement si $a < -1$; converge simplement si $a < 0$ et $z \neq 6-i$; diverge si $a \geqslant 0$ ou si $(0 > a \geqslant -1$ et $z = 6-i)$.

(6) Pour le rayon de convergence, nous calculons la limite

$$\alpha = \lim_{k\to\infty} \left| \frac{(-1)^{k+2}\sqrt{k+4}}{(k+3)^2} \cdot \frac{(k+2)^2}{(-1)^{k+1}\sqrt{k+3}} \right| = \lim_{k\to\infty} \left(\frac{k+2}{k+3} \right)^2 \sqrt{\frac{k+4}{k+3}} = 1 \tag{97.216}$$

Le rayon de convergence est donc 1 sans discussions.

Sur le bord, $|z+1-i|=1$, la convergence absolue est étudiée par la série

$$\sum_k \frac{\sqrt{k+1}}{(k+2)^2}, \tag{97.217}$$

qui se compare à la série $\sum_k k^{-3/2}$, qui converge. Il y a donc convergence absolue sur le bord du disque de convergence.

(7) Ce qui est tout à fait boulversifiant dans la série

$$\sum_{k=1}^{\infty} \frac{(-z+2)^k}{2k+\ln(k)},$$

c'est le signe devant le z. Qu'à cela ne tienne, nous faisons simplement la manipulation suivante pour mettre ce signe en évidence :

$$(-z+2)^k = \big(-(z-2)\big)^k = (-1)^k(z-2). \tag{97.218}$$

Nous sommes donc avec une série de puissances de coefficients $c_k = \frac{(-1)^k}{2k+\ln(k)}$, et centrée en $z_0=2$. Le calcul de la limite

$$\lim_{k\to\infty}\left|\frac{c_{k+1}}{c_k}\right| = \lim_{k\to\infty}\frac{2k+\ln(k)}{2(k+1)+\ln(k+1)} = 1$$

montre que le rayon de convergence est $R=1$.

Pour un point z vérifiant $|z-2|=1$, la série des modules (permettant d'analyser la convergence absolue) devient

$$\sum_{k=1}^{\infty}\frac{1}{2k+\ln(k)}$$

qui est équivalente à la série harmonique, et donc diverge. Il n'y a donc pas convergence absolue sur le bord du disque. Rien n'est perdu cependant, il est toujours possible que nous ayons une convergence simple.

Pour analyser la convergence simple, toujours sur le bord du disque, utilisons le critère d'Abel : la suite $\frac{1}{2k+\ln(k)}$ est clairement décroissante et tend vers 0. On écrit :

$$\begin{aligned}\left|\sum_{k=1}^{n}(-z+2)^k\right| &= \left|-1+\sum_{k=0}^{n}(-z+2)^k\right| \\ &\leqslant 1+\left|\sum_{k=1}^{n}(-z+2)^k\right| \\ &= 1+\frac{|1-(-z+2)^{n+1}|}{|-1+z-2|} \\ &\leqslant 1+\frac{1+|z-2|^{n+1}}{|z-1|} \\ &= 1+\frac{1}{|z-1|}.\end{aligned} \tag{97.219}$$

Cela nous fournit une borne pour les sommes partielles, tant que $z\neq 1$. Nous avons donc prouvé la convergence (non absolue) sur le disque de convergence moins le point $z=1$. Ce dernier point reste à étudier.

Pour analyser le point $z=1$, reprenons l'énoncé initial : on retrouve la série

$$\sum_{k=1}^{\infty}\frac{1}{2k+\ln(k)}$$

qui est équivalente à la série harmonique et donc diverge.

Conclusion : La série converge absolument si z vérifie $|-z+2|<1$; converge simplement si $|-z+2|=1$ et $z\neq 1$; diverge si $|-z+2|>1$ ou si $z=1$.

97.7 Exercices de topologie

Si A_n est une suite d'ensemble, le symbole

$$\bigcap_{n=1}^{\infty} A_n \tag{97.220}$$

désigne l'ensemble des éléments qui sont dans A_n pour tout $n \in \mathbb{N}$. Remarquez que l'infini *n'est pas* un élément de $\mathbb{N}$! L'intersection se fait donc de $n = 1$ à l'infini ; l'infini non compris.

Prenons comme exemple le cas du point (6) de l'exercice 70. Étant donné que $A_n =]-\frac{1}{n}, \frac{1}{n}[$, on pourrait croire que $A_\infty =]0, 0[= \varnothing$, et que par conséquent, l'intersection $\bigcap_{n=1}^{\infty}$ est vide.

97.7.1 Exercices ultra basiques

Exercise 69 exo0071

Donner si possible un exemple d'espace

(1) Ouvert dans $\mathbb{R}^2$.

(2) Fermé dans $\mathbb{R}^3$.

(3) Fermé et ouvert.

(4) Compact ne contenant que 8 points.

(5) Fermé non borné.

(6) Fermé, ouvert et borné.

corr0071

Correction of the exercise 69

(1) Une boule ouverte $\{(x, y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 < 1\}$

(2) Un cercle $\{(cos(t), sin(t), 1) \text{ tel que } t \in \mathbb{R}\}$

(3) Pour un exemple non trivial, il faut être dans un espace topologique non connexe. Par exemple dans l'hyperboloïde à deux nappes d'équation $x^2 + y^2 = z^2 - 1$ vu comme espace topologique, la nappe $z > 0$ est ouverte et fermée. L'hyperboloïde lui-même est également un ouvert fermé de lui-même (c'est vrai pour n'importe quel espace topologique).

Plus généralement, dans un espace topologique non connexe, les composantes connexes sont ouvertes et fermées.

(4) Dans $\mathbb{R}^n$, n'importe quel ensemble de 8 points est compact (fermé et borné). Nous pouvons aussi montrer que n'importe quel espace topologique qui possède un nombre fini de points est compact.

En effet, soit l'ensemble $A = \{a_1, \dots a_n\}$, et prenons un recouvrement de cet ensemble par des ouverts $\mathcal{O}_\alpha$. Comme c'est un recouvrement, nous avons $A \subseteq \bigcup_\alpha \mathcal{O}_\alpha$. Soit α_1 tel que $a_1 \in \mathcal{O}_{\alpha_1}$ (ceci existe parce que a_1 est dans un des $\mathcal{O}_\alpha$). Plus généralement, nous prenons i tel que a_i soit dans $\mathcal{O}_{\alpha_i}$ pour $i = 1, \dots n$.

Maintenant, $A \subseteq \bigcup_{i=1}^{n} \mathcal{O}_{\alpha_i}$, ce qui fait que nos $\mathcal{O}_{\alpha_i}$ forment un sous-recouvrement fini de A.

(5) Une hélice $\{(cos(t), sin(t), t) \text{ tel que } t \in \mathbb{R}\}$ dans $\mathbb{R}^3$. Une simple droite dans $\mathbb{R}^2$ est bon aussi.

(6) Pour obtenir un exemple non trivial (pas $\varnothing$ dans un espace métrique, ni un espace topologique métrique borné vu comme sous-espace de lui-même), il faut se placer dans un espace non connexe. Par exemple dans l'espace topologique formé de l'une union disjointe de deux sphères $S^2 \sqcup S^2$, chaque sphère est ouverte, fermée et bornée.

Exercise 70 exo0072

Déterminez l'intérieur, l'adhérence et la frontière des sous-ensembles de $\mathbb{R}$ suivants (Précisez si l'ensemble est ouvert, fermé, borné, compact).

(1) $]-\sqrt{3}, \sqrt{3}[\cup [10, 11]$

Item72b

(2) $[2,3[\backslash\{e\}$

(3) $\mathbb{Z}$

(4) $\mathbb{Q}$

(5) $\{\frac{1}{i} \mid i \in \mathbb{Z}_0\}$

ItemF0072

(6) $\bigcap_{n=1}^{\infty}]-\frac{1}{n}, \frac{1}{n}[$

corr0072

Correction of the exercise 70

Quelques remarques. N'oubliez pas que l'adhérence d'un ensemble non vide n'est *jamais* vide : l'ensemble lui-même est toujours dans son adhérence.

(1)

(2)

(3) L'ensemble $\mathbb{Z}$ n'a pas d'intérieur. En effet, prenons $n \in \mathbb{Z}$ et considérons la boule $B(n,e)$. Cette boule contient évidement des éléments qui ne sont pas dans $\mathbb{Z}$. Nous avons donc prouvé qu'aucune boule centrée en n n'est entièrement comprise dans $\mathbb{Z}$. Cela prouve que $\mathbb{Z}$ n'as pas d'intérieur.

L'adhérence de $\mathbb{Z}$ est réduite à $\mathbb{Z}$ lui-même parce que si $x \notin \mathbb{Z}$, il existe un r tel que $B(x,r) \cap \mathbb{Z} = \varnothing$, ce qui prouve que x n'est pas adhérent à $\mathbb{Z}$.

(4) L'ensemble $\mathbb{Q}$ n'as pas d'intérieur pour la même raison que $\mathbb{Z}$: il existe un élément hors de $\mathbb{Q}$ aussi proche que l'on veut de tout rationnel. L'adhérence de $\mathbb{Q}$, par contre est tout $\mathbb{R}$. En effet, si $x \in \mathbb{R}$ et si $r > 0$, la boule $B(x,r)$ intersecte toujours $\mathbb{Q}$.

(5)

(6) Notez que

$$\bigcap_{n=1}^{\infty}]-\frac{1}{n}, \frac{1}{n}[= \{0\}. \tag{97.221}$$

Les réponses complètes sont dans le tableau suivant. Parmi les propriétés d'être ouvert, fermé, borné, compact ou connexe par arcs (c.p.a), les ensembles cités jouissent des propriétés mentionnées mais pas des autres.

(a) **Intérieur** $]-\sqrt{3},\sqrt{3}[\cup]10,11[$
Adhérence $[-\sqrt{3},\sqrt{3}] \cup [10,11]$
Frontière $\{-\sqrt{3},\sqrt{3},10,11\}$
Propriétés borné

(b) **Intérieur** $]2;3[\backslash\{e\}$
Adhérence $[2;3]$
Frontière $\{2,e,3\}$
Propriétés borné

(c) **Intérieur** $\varnothing$
Adhérence $\mathbb{Z}$
Frontière $\mathbb{Z}$
Propriétés fermé

(d) **Intérieur** $\mathbb{R}$
Adhérence $\mathbb{R}$
Frontière $\varnothing$
Propriétés ouvert, fermé, connexe par arcs.

(e) **Intérieur** $\varnothing$

Adhérence $\{1/i \text{ tel que } i \in \mathbb{Z}_0\} \cup \{0\}$

Frontière $\{1/i \text{ tel que } i \in \mathbb{Z}_0\} \cup \{0\}$

Propriétés

(f) **Intérieur** $\varnothing$

Adhérence $\{0\}$

Frontière $\{0\}$

Propriétés fermé,b borné, compact, connexe par arcs.

Exercise 71 exo0073

Soit $(E, ||.||)$ un espace vectoriel normé, et x_n une suite convergente dans E dont x est la limite. Montrer que $\{x_n\} \bigcup \{x\}$ est compact. corr0073

Correction of the exercise 71

Prenons un recouvrement par des ouverts. Parmi tous les ouverts, il y en a au moins un qui contient x, disons A_1. Cet ouvert contient une boule de rayon r. Prenons maintenant $K \in \mathbb{N}$ tel que $n > K$ implique $|x_n - x| \leqslant r$. Tous les éléments de la suite à partir de K (y compris la limite) sont donc contenus dans A_1.

Il ne reste plus que K éléments à mettre dans un nombre finis d'ouverts. Ça c'est facile : on prend un ouvert par point.

Si cet exercice t'a plu, alors l'exercice 86 va sûrement te donner tout plein de plaisir.

Exercise 72 exo0074

Considérons intervalles

$$A_n = \left[\frac{-3n+2}{n}, \frac{2n^2-n}{n^2}\right].$$

Que pouvez-vous dire de $\bigcup_{n=1}^{\infty} A_n$? Et de $\bigcap_{n=1}^{\infty} A_n$? corr0074

Correction of the exercise 72

On observe que

$$A_n = \left[-3 + \frac{2}{n}, 2 - \frac{1}{n}\right]$$

où $-3 + \frac{2}{n}$ est une suite décroissante, et $2 - \frac{1}{n}$ est croissante. On a donc une suite d'intervalles emboîtés

$$A_1 \subset A_2 \subset A_3 \ldots$$

donc la réunion vaut $]-3, 2[$ et l'intersection vaut $A_1 = [-1, 1]$. On remarque que chaque A_n est fermé, mais la réunion est ouverte.

Exercise 73 exo0075

Démontrez les lois de *de Morgan* : si $A_n \subset \Omega \; \forall n$, alors

$$\left(\bigcup_n A_n\right)^c = \bigcap_n A_n^c, \quad \left(\bigcap_n A_n\right)^c = \bigcup_n A_n^c$$

où A^c est le complémentaire de A dans Ω. corr0075

Correction of the exercise 73

Montrons la première relation : soit $x \in \Omega$. Alors nous avons les équivalences suivantes :

$$x \in \left(\bigcup_n A_n\right)^c \iff \neg\left(x \in \bigcup_n A_n\right) \iff \neg\,(\exists n \text{ tel que } x \in A_n)$$
$$\iff \forall n, \neg(x \in A_n) \iff \forall n, x \in (A_n)^c \iff x \in \bigcap_n (A_n)^c$$

où $\neg$ désigne la négation logique (et n est élément de $\mathbb{N}$).

La deuxième relation se traite de façon similaire. On peut également la déduire de la première relation de la manière suivante : pour montrer que $(\bigcap_n A_n)^c = \bigcup_n (A_n{}^c)$, il suffit de poser $B_n = A_n^c$, de prendre le complémentaire dans les deux membres et d'utiliser le fait que $(X^c)^c = X$ pour se ramener à la première relation.

Exercise 74 exo0076

Représentez dans le plan les sous-ensembles de $\mathbb{R}^2$ suivants et déterminez leur intérieur, leur adhérence et leur frontière (précisez si l'ensemble est ouvert, fermé, borné, compact ou [7] connexe par arcs).

Item76a

(1) $\{(x,y) \mid 3x+4y=2\}$ Item76b

(2) $\{(x,y) \mid 1-xy>0\}$ Item76c

(3) $\{(x,y) \mid |x|=1, |y|\neq 1\}$ Item76d

(4) $\{(x,y) \mid (x-3)^2+(y+2)^2 \leqslant 4\}$ Item76e

(5) $\{(x,y) \mid x^2-2x+y^2+4y-5 \leqslant 0\}$ Item76f

(6) $\{(x,y) \mid x^2+y^2 \leqslant 3 \text{ et } -1 \leqslant y < 1\}$ Item76g

(7) $\{(\cos(t),\sin(t)) \mid t \in \mathbb{R}\}$ Item76h

(8) $\{(v\cos(u), v\sin(u)) \mid u \in \mathbb{R} \text{ et } v \in]-2,1]\}$ Item76i

(9) $\{(x_0,y_0)(1-t)+t(x_1,y_1) \mid t \in [0,1]\}$ Item76j

(10) $\{(t,\frac{t}{k}) \mid t \in [0,1] \text{ et } k \in \mathbb{N}_0\}$ Item76k

(11) $\{(t,\frac{t}{k}) \mid t \in]0,1] \text{ et } k \in \mathbb{N}_0\}$ Item76l

(12) $\bigcup_{n\geqslant 1}\{\frac{1}{n}\} \times [0,1]$ corr0076

Correction of the exercise 74

Quelques notes

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

(10) Le point $(0,0)$ est dans l'ensemble, et donc dans l'adhérence. L'adhérence peut donc être notée indifféremment $E \cup [0,1]\times\{0\}$ ou $E \cup]0,1]\times\{0\}$.

(11) Dans ce cas ci, par contre, le point $(0,0)$ n'est pas dans l'ensemble, donc l'adhérence doit être écrite $E \cup [0,1]\times\{0\}$.

(12)

Parmi les propriétés d'être ouvert, fermé, borné, compact ou connexe par arcs (c.p.a), les ensembles cités (toujours noté E) jouissent des propriétés mentionnées mais pas des autres.

sdf

7. ce « ou » n'est évidement pas exclusif.

Ex.	Intérieur	Adhérence	Frontière	Propriétés
(1)	$\varnothing$	E	E	fermé, c.p.a
(2)	E	$\{1-xy \geqslant 0\}$	$\{1-xy=0\}$	ouvert, c.p.a
(3)	$\varnothing$	$\{\lvert x\rvert = 1\}$	$\{\lvert x\rvert = 1\}$	Néant
(4)	$\{(x-3)^2+(y+2)^2<4\}$	E	$\{(x-3)^2+(y+2)^2=4\}$	fermé, borné, compact, c.p.a
(5)	$\{(x-1)^2+(y+2)^2<10\}$	E	$\{(x-1)^2+(y+2)^2=10\}$	fermé, borné, compact, c.p.a
(6)	$\{x^2+y^2<3 \text{ et } \lvert y\rvert<1\}$	$\{x^2+y^2\leqslant 3 \text{ et } \lvert y\rvert\leqslant 1\}$	$\{x^2+y^2\leqslant 3 \text{ et } \lvert y\rvert=1\}$ $\cup\{x^2+y^2=3 \text{ et } \lvert y\rvert\geqslant 1\}$	borné, c.p.a
sdfds (7)	$\varnothing$	$E=\{x^2+y^2=1\}$	E	fermé, borné, compact, c.p.a
(8)	$E=\{x^2+y^2<4\}$	$\{x^2+y^2\geqslant 4\}$	$\{x^2+y^2=4\}$	ouvert, borné, c.p.a
(9)	$\varnothing$	E	E	fermé, borné, compact, c.p.a C'est le segment joignant (x_0,y_0) à (x_1,y_1)
(10)	$\varnothing$	$E\cup[0,1]\times\{0\}$	$E\cup]0,1]\times\{0\}$	borné, c.p.a
(11)	$\varnothing$	$E\cup[0,1]\times\{0\}$	$E\cup[0,1]\times\{0\}$	borné
(12)	$\varnothing$	$E\cup\{0\}\times[0,1]$	$E\cup\{0\}\times[0,1]$	borné

Pour l'ensemble (5), remarquez que $x^2-2x+y^2+yy-5=(x-1)^2+(y+2)^2-10$; c'est le coup classique de reformer les carrés.

Parmi les propriétés d'être ouvert, fermé, borné ou compact, les ensembles considérés jouissent des propriétés citées et pas des autres.

(1) Fermé (droite).

(2) Fermé (hyperbole).

(3) Aucune propriété (union disjointe de deux demi-droites et de deux segments ouverts).

(4) Fermé, borné, compact (disque).

(5) Idem.

(6) Borné (morceau de disque)

(7) Fermé, borné, compact (cercle).

(8) Ouvert, borné (disque ouvert de rayon 2).

(9) Fermé, borné, compact (segment de droite ou point).

(10) Borné (bouquet infini de segments)

(11) Idem (idem sans le point 0).

(12) Borné (union disjointe d'une infinité de segments).

Exercise 75 exo0077

Représentez vous dans l'espace les sous-ensembles de $\mathbb{R}^3$ suivants et déterminez leur intérieur, leur adhérence et leur frontière (précisez si l'ensemble est ouvert, fermé, borné, compact).

(1) $\{(x,y,z)\mid (x-3)^2+(y)^2+z^2\leqslant 9\}$

(2) $\{(x,y,z)\mid x^2+y^2\leqslant 7\}$ Item0077c

(3) $\{(x,y,z)\mid x^2\leqslant 3\}$

(4) $\{(x,y,z)\mid x<y\}$

(5) $\{(\cos(t),\sin(t),t)\mid t\in\mathbb{R}\}$ corr0077

Correction of the exercise 75

Même esprit que l'exercice précédent.

(1) Fermé, borné, compact (sphère).

(2) Fermé (cylindre plein).

(3) Fermé (pavé infini).

(4) Ouvert (demi espace).

(5) Fermé (hélice).

Exercise 76 exo0078

Justifiez soigneusement l'exercice (2) de l'exercice 70. corr0078

Correction of the exercise 76

Soit $A=[2,3[\backslash\{e\}$ et vérifions que $\operatorname{Int}A=]2,3[\backslash\{e\}$:

(1) Vérifions d'abord que 2 n'est pas un point intérieur à A. En effet, si on considère la boule $B(2,\epsilon)$ autour de 2, le point $2-\epsilon/2$ est bien dans cette boule mais $2-\epsilon/2<2$ (car $\epsilon>0$), donc $2-\epsilon/2\notin A$.

(2) De plus, $]2,3[\backslash\{e\}$ est la réunion des ouverts $]2,e[$ et $]e,3[$, c'est donc un ouvert, qui est contenu dans A.

On en déduit que $]2,3[\backslash\{e\}$ est le plus grand ouvert contenu dans A, c'est donc bien l'intérieur de A.

Similairement, pour l'adhérence, on veut montrer que $\operatorname{Adh}A=[2,3]$:

(1) Notons d'abord que pour n assez grand, les quantités $e + 1/n$ et $3 - 1/n$ sont dans A. Les suites correspondantes, $x_n = e + 1/n$ et $y_n = 3 - 1/n$, tendent vers e et 3 respectivement. On en déduit que e et 3 doivent être dans l'adhérence de A.

(2) D'autre part, l'ensemble $[2, 3]$ étant fermé, c'est forcément le plus petit fermé contenant A.

Le fait que A n'est pas ouvert ni fermé résulte des définitions car A n'est ni égal à son adhérence, ni égal à son intérieur. Le fait que A n'est pas connexe par arcs peut se justifier comme suit : A est inclus dans la réunion $]-\infty, e[\cup]e, +\infty[$. Ces deux ouverts $]-\infty, e[$ et $]e, +\infty[$ sont disjoints, et aucun d'eux ne contient entièrement A. Donc A n'est pas connexe, et en particulier n'est pas connexe par arcs.

Exercise 77 exo0079

Justifier soigneusement l'exercice (3) de l'exercice 75. corr0079

Correction of the exercise 77

Soit $E = \{(x, y, z) \text{ tel que } x^2 \leqslant 3\}$. Prouvons que E est fermé, car si (p_k) est une suite de points $(p_k = (x_k, y_k, z_k))$ de E (c'est-à-dire $(x_k)^2 \leqslant 3$) convergente dans $\mathbb{R}^3$ vers une limite $p = (x, y, z)$, alors en particulier la suite x_k tend vers x. Comme $(x_k)^2 \leqslant 3$, on peut passer à la limite dans l'inégalité et donc $x^2 \leqslant 3$, ce qui prouve que $p = (x, y, z)$ est dans E, et donc $\operatorname{Adh} E \subset E$. Comme par ailleurs $E \subset \operatorname{Adh} E$, on a bien l'égalité.

Remarque 97.1.
De manière plus générale, si $f : \mathbb{R}^n \to \mathbb{R}^m$ est continue, et si $F \subset \mathbb{R}^m$ est un fermé de $\mathbb{R}^m$, alors l'ensemble

$$f^{-1}(F) \stackrel{\text{def}}{=} \{p \in \mathbb{R}^n \text{ tel que } f(p) \in F\}$$

est un ensemble fermé dans $\mathbb{R}^n$. Par ailleurs, si on suppose F ouvert au lieu de fermé, alors $f^{-1}(F)$ est ouvert.

Dans ce cas ci, $f : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \mapsto x^2$ est continue, et $F = [3, +\infty[$ est fermé. L'ensemble E est bien l'ensemble des points $p \in \mathbb{R}^3$ tels que $f(p) \in F$.

Prouvons que $\operatorname{Int} E = \{(x, y, z) \text{ tel que } x^2 < 3\}$. Tout d'abord, par la remarque ci-dessus, on observe que $\{(x, y, z) \text{ tel que } x^2 < 3\}$ est ouvert (image réciproque de l'ouvert $]3, +\infty]$) et donc est inclus dans l'intérieur de E. Par ailleurs, les points (x, y, z) vérifiant $x^2 = 3$ ne sont pas dans l'intérieur : si $B((3, y, z), \epsilon)$ désigne la boule de centre $(3, y, z)$ et de rayon ϵ, le point $(3 + \epsilon/2, y, z)$ appartient à cette boule mais n'est pas dans E ; donc aucune boule centrée en $(3, y, z)$ ne peut être contenue entièrement dans E. Ce qui prouve le résultat.

Prouvons que E est connexe par arcs. Si (x, y, z) et (a, b, c) sont des points de E, alors le chemin

$$\gamma : [0, 1] \to \mathbb{R}^3 : t \mapsto t(x, y, z) + (1 - t)(a, b, c)$$

est continu (il paramétrise le segment de droite joignant (a, b, c) à (x, y, z)). La condition pour être dans E s'écrit $|x| \leqslant 3$ et $|a| \leqslant 3$. Par l'inégalité triangulaire, on sait que

$$|tx + (1 - t)a| \leqslant |t|\,|x| + |1 - t|\,|a| \leqslant t\sqrt{3} + (1 - t)\sqrt{3} = \sqrt{3}$$

où on a utilisé le fait que $t \geqslant 0$ et $1 - t \geqslant 0$ (car $t \in [0, 1]$). Ceci prouve que $\gamma(t) \in E$ pour tout $t \in [0, 1]$, et on a donc bien un chemin continu reliant (x, y, z) et (a, b, c) contenu dans E.

Exercise 78 exo0080

Prouver que $\mathbb{R}_0$ n'est pas un sous-espace connexe de $\mathbb{R}$. corr0080

Correction of the exercise 78

L'ensemble $\mathbb{R}_0$ n'est pas un sous-espace connexe de $\mathbb{R}$ parce que $\mathbb{R}_0 = \mathbb{R}_0^- \cup \mathbb{R}_0^+$ est la réunion de deux ouverts non vides.

97.7.2 Exercices simplement basiques

Les exercices qui suivent ne seront en principe pas vus aux séances (faute de temps, plus que faute d'envie), mais ils sont certainement très intéressants à regarder pour celles et ceux qui désirent en savoir un peu plus sur la topologie.

Exercise 79 exo0081

Deux métriques à peine bizarres...

(1) Pour tout ensemble non vide X (par exemple l'ensemble des habitants de Mars, l'ensemble des étoiles fixes dans la voie lactée, ou l'ensemble des bouteilles de bière à Bruxelles) nous définissons la « métrique discrète » par

$$d(x,y) := \begin{cases} 0 & \text{si } x = y, \\ 1 & \text{si } x \neq y. \end{cases}$$

Expliquer pourquoi le nom est approprié. Montrer qu'il s'agit bien d'une métrique. Décrire $B_r(x)$ pour tout $r > 0$ et pour tout $x \in X$.

(2) Dans $X = \mathbb{R}^2$ nous regardons la « métrique bureaucratique » (ou la « métrique SNCF ») définie par

$$d(x,y) := \begin{cases} |x-y| & \text{si } x, y \text{ sont linéairement dépendants,} \\ |x| + |y| & \text{si } x, y \text{ sont linéairement indépendants.} \end{cases}$$

Expliquer pourquoi le nom est approprié. Montrer qu'il s'agit bien d'une métrique.

corr0081

Correction of the exercise 79

(1) La topologie associée est la topologie discrète (les singletons sont ouverts, car une boule ouverte de rayon 1 (ou $\frac{1}{2}$) ne contient qu'un seul point.

(2) C'est la distance usuelle si x et y sont sur une même droite (même ligne de chemin de fer), mais dans les autres cas il faut repasser par le centre.

Exercise 80 exo0082

Soit (X,d) un espace métrique. Prouver l'inégalité quadrilatère

$$|d(a,b) - d(u,v)| \leqslant d(a,u) + d(b,v) \tag{97.222}$$

pour tout $a, b, u, v \in X$.

corr0082

Correction of the exercise 80

Pour tous a, b, u, v on a, par l'inégalité triangulaire

$$d(a,b) \leqslant d(a,u) + d(u,v) + d(v,b) \text{ et } d(u,v) \leqslant d(u,a) + d(a,b) + d(b,v)$$

d'où on tire

$$d(a,b) - d(u,v) \leqslant d(a,u) + d(v,b) \text{ et } d(u,v) - d(a,b) \leqslant d(u,a) + d(b,v)$$

ce qui montre l'inégalité annoncée.

Exercise 81 exo0083

Montrer que pour tout $x \in \mathbb{R}^n$ et tout $r > 0$ la boule

$$B(x,r) := \{y \in \mathbb{R}^n \mid |x-y| < r\}$$

est ouverte.

corr0083

Correction of the exercise 81

Afin de prouver que la boule est ouverte, nous allons utiliser le théorème 7.8 : nous prenons un point $p \in B(x,r)$, et nous allons montrer qu'il existe une boule autour de p qui est contenue dans $B(x,r)$.

Étant donné que $p \in B(x,r)$, nous avons $d(p,x) < r$. Prouvons que la boule $B\big(p, r-d(p,x)\big)$ est contenue dans $B(x,r)$. Pour cela, nous prenons $p' \in B\big(p, r-d(p,x)\big)$, et nous essayons de prouver que $p' \in B(x,r)$. En effet, en utilisant l'inégalité triangulaire,

$$d(x,p') \leqslant d(x,p) + d(p,p') \leqslant d(x,p) + r - d(p,x) = r. \tag{97.223}$$

Exercise 82 exo0084

Soient $f\colon (X,d_X) \to (Y,d_Y)$ et $g\colon (Y,d_Y) \to (Z,d_Z)$ des applications continues entre espaces métriques. Montrer que la composition $g \circ f\colon (X,d_X) \to (Z,d_Z)$ est continue. corr0084

Correction of the exercise 82

Soit $\epsilon > 0$, on sait qu'il existe η tel que

$$d(f(x),f(y)) < \eta \Rightarrow d(g(f(x)),g(f(y))) < \epsilon$$

par la continuité de g en $f(x)$. Or on sait qu'il existe δ tel que $d(f(x),f(y)) < \eta$ est vrai dès que $d(x,y) < \delta$ par continuité de f en x. Donc si $d(x,y) < \delta$, on a bien $d(g(f(x)),g(f(y))) < \epsilon$.

Exercise 83 exo0085

Soit $f\colon E \to \mathbb{R}$, une fonction continue sur l'espace métrique E.

(1) Montrer que l'ensemble des x tels que $f(x) = 0$ est fermé.

(2) Montrer que l'ensemble des points fixes[8] de f est fermé.

corr0085

Correction of the exercise 83

(1) Il suffit de regarder le complémentaire, c'est-à-dire l'ensemble S des x vérifiant $f(x) \neq 0$. Vérifions que pour tout $x \in S$, on peut placer une boule ouverte centrée en x complètement incluse à S. Soit $\epsilon = |f(x)|$. Alors il existe $\delta > 0$ tel que si $d(x,y) < \delta$, on a $d(f(x),f(y)) = |f(x) - f(y)| < \epsilon$. La boule ouverte $\{y \in X$ tel que $d(x,y) < \delta$ est alors une boule incluse dans S : l'image d'un point y ne peut pas être nulle, sinon l'inégalité $|f(x) - f(y)| = \epsilon < \epsilon$ n'est pas satisfaite.

(2) On utilise le premier point sur la fonction g définie par $g(x) = f(x) - x$.

Exercise 84 exo0086

Soit E, un espace métrique et $A \subset E$. La **distance** entre A et $x \in E$ est définie par

$$d(A,x) = \inf_{a \in A} d(a,x). \tag{97.224}$$

Prouver que la $x \mapsto d(A,x)$ définit une fonction continue (pourquoi est-ce que cette fonction est bien définie ?). corr0086

Correction of the exercise 84

D'abord, on voit que cette fonction est bien définie, puisque la distance est toujours minorée par 0 donc l'infimum existe toujours.

Pour la continuité, prenons $x \in X$ et $\epsilon > 0$. On veut trouver $\delta > 0$ tel que

$$d(x,y) < \delta \Rightarrow \left| \inf_{a \in A} d(a,x) - \inf_{a \in A} d(a,y) \right| < \epsilon$$

8. C'est-à-dire les points x tels que $f(x) = x$.

Pour tout $a \in A$, on a successivement

$$\begin{aligned} d(x,a) \leqslant d(a,y) + d(x,y) &\Rightarrow \inf_{a\in A} d(x,a) \leqslant \inf_{a\in A} \big(d(a,y) + d(x,y)\big) \\ &\Rightarrow \inf_{a\in A} d(x,a) \leqslant \inf_{a\in A} d(a,y) + d(x,y) \\ &\Rightarrow \inf_{a\in A} d(x,a) - \inf_{a\in A} d(a,y) \leqslant d(x,y) \end{aligned}$$

et de même en échangeant les rôles de x et y, ce qui conduit à

$$\inf_{a\in A} d(y,a) - \inf_{a\in A} d(a,x) \leqslant d(x,y)$$

donc on en déduit la majoration

$$\left|\inf_{a\in A} d(y,a) - \inf_{a\in A} d(a,x)\right| \leqslant d(x,y) \tag{97.225}$$

egex13

dès lors le choix $\delta = \epsilon$ conduit à l'inégalité désirée.

Exercise 85 exo0087

À propos de fonctions Lipschitziennes.

(1) Montrer qu'une application Lipschitzienne est continue.

(2) Montrer qu'une application $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$ est Lipschitzienne. Quelle est la plus petite constante L qui convienne ?

(3) Montrer que les fonctions $z \mapsto |z|$, $z \mapsto \overline{z}$, $z \mapsto \operatorname{Re} z$ et $z \mapsto \operatorname{Im} z$ sont Lipschitziennes. Quelle sont les plus petites constantes L qui conviennent ?

(4) Montrer que la fonction $d(A, \cdot\,)\colon X \to \mathbb{R}$ de l'exercice 84 est Lipschitzienne.

corr0087

Correction of the exercise 85

(1) Soit $\epsilon > 0$, on choisit $\delta = \frac{\epsilon}{L}$, on a ce qu'il faut.

(2) On note que $|ax + b - (ay + b)| = |a|\,|x - y|$. La plus petite constante est donc $L = |a|$.

(3) (3a) On note que $||z| - |z'|| \leqslant |z - z'|$ donc $L = 1$ convient. Ce choix est minimal (prendre $z = 2$ et $z' = 1$).

(3b) On note que $\left|\overline{z} - \overline{z'}\right| = \left|z - \overline{z'}\right| = |z - z'|$ donc $L = 1$ convient et est minimal.

(3c) On note que $|\Re z - \Re z'| = |\Re(z - z')| \leqslant |z - z'|$. Dès lors $L = 1$ convient et est minimal (prendre $z = 2$ et $z' = 1$).

(3d) Idem (prendre $z = 2i$ et $z' = i$).

(3e) Cela découle directement de l'inégalité (97.225) utilisée pour montrer la continuité.

Exercise 86 exo0088

Nous considérons l'ensemble $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. Un sous-ensemble de $\bar{\mathbb{R}}$ qui ne contient pas ∞ est dit ouvert si et seulement si il est ouvert pour $\mathbb{R}$, et un sous-ensemble contenant ∞ sera ouvert si son complémentaire est compact dans $\mathbb{R}$.

(1) Prouver que ces ouverts définissent bien une topologie sur $\bar{\mathbb{R}}$.

(2) Prouver que $\bar{\mathbb{R}}$ est un espace compact.

Rendez vous compte que $\mathbb{R}$ est non compact, mais qu'on l'a rendu compact en *ajoutant* un point en plus. L'espace topologique ainsi défini est le **compactifié** d'Alexandroff de l'ensemble des réels $\mathbb{R}$.

corr0088

Correction of the exercise 86

Exercise 87 exo0089

Montrer que l'ensemble $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ n'est ni ouvert ni fermé.

corr0089

Correction of the exercise 87

L'ensemble proposé n'est évidement pas ouvert, parce que le point 1 ne satisfait pas au théorème 7.8.

Pour qu'il soit fermé, il faudrait que le complémentaire soit ouvert. Hélas, le complémentaire n'est pas ouvert parce que 0 (qui n'est pas dans l'ensemble) ne satisfait pas au théorème : dans tout ouvert autour de 0, il y a des points de la forme $1/n$.

97.8 Fonctions d'une variable réelle (suite)

Exercise 88 exo0090

Déterminer l'image des fonctions suivantes et prouvez soigneusement vos réponses.

(1) $f : \mathbb{R} \to \mathbb{R} : x \to \sin(x)$

(2) $f : [0, 3] \to \mathbb{R} : x \to xe^x$

(3) $f :]\pi, +\infty[\to \mathbb{R} : y \to y \sin(y)$

corr0090

Correction of the exercise 88

(1) $[-1, 1]$,

(2) L'image est $[0, 3e^3]$ par ce que la fonction est strictement croissante (vérifier la dérivée).

(3) L'image est $\mathbb{R}$ parce que si $M \in \mathbb{R}$, il existe $x_0 > M$ tel que $\sin(y_0) = 1$, et pour ce y_0, nous avons $f(y_0) = y_0 > M$, donc f prend des valeurs arbitrairement élevées. De la même façon, en prenant y_0 tel que $\sin(y_0) = -1$, nous trouvons que f prend des valeurs arbitrairement basses. La continuité et le théorème des valeurs intermédiaires font le reste.

Exercise 89 exo0091

Prouvez que si $f :]a, b[\to \mathbb{R}$ est une fonction dérivable sur $]a, b[$ et que $f' :]a, b[\to \mathbb{R} : x \to f'(x)$ est bornée sur $]a, b[$ alors f est uniformément continue sur $]a, b[$.

Qu'est-ce qu'il se passe si nous considérons l'intervalle $]a, \infty[$ au lieu de $]a, b[$?

corr0091

Correction of the exercise 89

Le théorème des accroissements finis nous dit qu'entre x et y, il existe un c tel que

$$f(x) - f(y) = f'(c)(x - y) < M(x - y) \tag{97.226}$$

si M est une majoration de f' sur $]a, b[$.

Si maintenant on me donne un ϵ, il me suffit de prendre $\delta < \frac{\epsilon}{M}$, et j'ai

$$|f(x) - f(y)| < \epsilon \tag{97.227}$$

pour tout x et y tels que $|x - y| < \delta$.

Exercise 90 exo0092

Prouvez que si une fonction f définie sur sous-ensemble de $\mathbb{R}$ est uniformément continue sur I_1 et sur I_2, deux intervalles qui se chevauchent alors elle est uniformément continue sur $I_1 \cup I_2$. corr0092

Correction of the exercise 90

Soit $\epsilon > 0$. Il faut trouver un δ tel que $|f(x) - f(y)| < \epsilon$ dès que $x, y \in I_1 \cup I_2$ et $|x - y| < \delta$. Nous savons qu'il existe un tel δ_1 pour I_1 et un δ_2 pour I_2.

Dans ce genre d'exercice, il est d'usage de choisi $\delta = \min\{\delta_1, \delta_2\}$, afin que δ fonctionne à la fois pour I_1 et pour I_2. Cependant, cela n'est pas suffisant ici parce que nous n'avons aucune garantie sur ce qu'il se passerait sir $x \in I_1 \backslash I_2$ et $y \in I_2 \backslash I_2$ (c'est-à-dire si aucun des deux n'est dans l'intersection de I_1 et I_2).

La solution, pour empêcher cette situation est de prendre, en plus, δ plus petit que la moitié[9] de la taille de l'intersection entre I_1 et I_2. De cette manière, deux éléments x et y tels que $|x-y| \leqslant \delta$ sont automatiquement tous les deux dans I_1 ou tous les deux dans I_2.

Exercise 91 exo0093

Cet exercice sera utile dans la suite.

(1) Soient $a, b \in \mathbb{R}, a < b$ et $f : [a, b[\to \mathbb{R}$ une fonction continue. Prouver que f est uniformément continue sur $[a, b[$ si et seulement si f peut être prolongée par continuité au point b.

(2) Montrer que la fonction $f(x) = \frac{1}{x}$ n'est pas uniformément continue sur $]0, \infty[$ mais qu'elle l'est sur $[a, \infty[$ pour tout $a > 0$.

(Aide : pour $\Rightarrow$, prouver que si $(b_n) \subset [a, b[$ et $\lim b_n = b$ alors la limite $\lim f(b_n)$ existe dans $\mathbb{R}$.) corr0093

Correction of the exercise 91

(1) Si la fonction f peut être prolongée par continuité sur $[a, b,]$, alors la nouvelle fonction est une fonction continue sur un compact, et est donc uniformément continue. La fonction f est alors la restriction à $[a, b[$ d'une fonction uniformément continue, ce qui fait que f est elle-même uniformément continue.

Exercise 92 exo0094

Déterminer si les fonctions suivantes sont uniformément continues

(1) $[0, b] \to \mathbb{R} : x \mapsto x^2$
(2) $]0, b[\to \mathbb{R} : x \mapsto x^2$
(3) $]0, 1] \to \mathbb{R} : x \mapsto x \sin(\frac{1}{x})$
(4) $[0, +\infty[\to \mathbb{R} : x \mapsto x^2$
(5) $[a, 1] \to \mathbb{R} : x \mapsto \sin(\frac{1}{x})$
(6) $]0, 1] \to \mathbb{R} : x \mapsto \sin(\frac{1}{x})$
(7) $[0, \infty[\to \mathbb{R} : x \mapsto \sin(x^2)$
(8) $[0, 2] \to \mathbb{R} : x \mapsto \sqrt{x}$
(9) $[1, +\infty[\to \mathbb{R} : x \mapsto \sqrt{x}$
(10) $[0, +\infty[\to \mathbb{R} : x \mapsto \sqrt{x}$
(11) $]0, 1[\to \mathbb{R} : x \mapsto x^x$
(12) $]0, 1] \to \mathbb{R} : x \mapsto \ln(x)$
(13) $[1, \infty[\to \mathbb{R} : x \mapsto \ln(x)$

où $b \in \mathbb{R}^+$ et où $a \in]0, 1[$. corr0094

Correction of the exercise 92

(1) La fonction est continue sur un compact, donc uniformément continue.

(2) Cette fonction est une restriction d'une fonction uniformément continue.

(3) Nous considérons la fonction

$$\begin{aligned} f\colon]0,1] &\to \mathbb{R} \\ x &\mapsto x \sin(\frac{1}{x}). \end{aligned} \tag{97.228}$$

Étant donné que $\lim_{x\to 0} f(x) = 0$ (le prouver), la fonction

$$\begin{aligned} g\colon [0,1] &\to \mathbb{R} \\ x &\mapsto \begin{cases} f(x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} \end{aligned} \tag{97.229}$$

est continue sur $[0, 1]$ et y est donc uniformément continue. Maintenant, la fonction f est la restriction de g au domaine $]0, 1]$, et y est donc uniformément continue.

(4) Vue sur l'intervalle $[a, 1]$, la fonction $x \mapsto \sin(\frac{1}{x})$ est continue sur un compact, et donc uniformément continue.o

9. En réalité, prendre exactement la taille de l'intersection est ne pose problème qu'au cas où les deux intervalles sont ouverts et que un point est à l'infimum de l'intersection et l'autre au supremum.

(5) Pour tout $\delta > 0$, il existe un choix de x et y avec $|x-y| \leqslant \delta$, mais avec $\frac{1}{x} = \frac{\pi}{2} + 2k\pi$ et $\frac{1}{y} = \frac{3\pi}{2} + 2k'\pi$. Pour un tel choix, nous avons

$$\left|\sin(\frac{1}{x}) - \sin(\frac{1}{y})\right| = 2, \tag{97.230}$$

donc nous n'avons pas uniforme convergence.

(6) Cette fonction n'est pas uniformément continue parce que si ϵ est donné, nous allons montrer qu'aucun choix de δ ne convient. Il suffit en effet de prendre x assez grand pour que $|x^2 - (x+\delta)^2| > 2\pi$, et alors, dans la boule $B(x,\delta)$, il y a un endroit où la fonction $\sin(x^2)$ prend la valeur 1 et un endroit où elle prend la valeur -1.

(7) La fonction est continue sur un compact.

(8) Nous savons que la dérivée de $\sqrt{x}$ est bornée sur $[1,\infty[$, donc la fonction y est uniformément continue en vertu de l'exercice 89.

(9) La fonction $\sqrt{x}$ étant uniformément continue sur $[0,2]$ et sur $[1,\infty[$, elle est uniformément continue sur $[0,\infty[$ par le principe des intervalles chevauchant, exercice 90.

(10) En utilisant l'astuce (habituelle) $x^x = e^{\ln(x^x)} = e^{x\ln(x)}$, nous pouvons montrer que

$$\lim_{x\to 0^+} x^x = 1. \tag{97.231}$$

Nous pouvons donc prolonger par continuité la fonction x^x sur l'intervalle compact $[0,1]$ en disant qu'elle vaut 1 en $x=0$. Cette prolongation est uniformément continue, et donc la fonction de départ est uniformément continue.

(11) Étant donné que $\lim_{x\to 0} \ln(x) = -\infty$, pour tout δ, il existe x et y tels que $|x-y| < \delta$ et avec $|\ln(x) - \ln(y)| > 1$.

(12) Cette fonction est uniformément continue parce que sa dérivée est bornée.

Exercise 93 exo0095

Soit P une fonction polynomiale de degré n de $\mathbb{R}$ dans $\mathbb{R}$. Supposons que P ait exactement n zéros distincts.

(1) Démontrez que la fonction dérivée P' possède exactement $n-1$ zéros distincts si $n \geqslant 1$.

(2) Quel est le nombre de zéros distincts de $P^{(k)}$, où $k \in \mathbb{N}$?

(3) Prouver que P est uniformément continue si et seulement si $deg(P) \leqslant 1$.

corr0095

Correction of the exercise 93

Par le théorème de Rolle, il y a un zéro de la dérivée entre deux zéros de la fonction. Comme la fonction possède n zéros, la dérivée en possède *au moins* $n-1$. Nous savons par ailleurs que la dérivée est un polynôme de degré $n-1$ et possède donc *au plus* $n-1$ zéros. La combinaison des deux donne la conclusion.

Par récurrence, la dérivée $k$ième a $n-k$ zéros distincts.

Si P est de degré 0 ou 1, alors elle est uniformément continue parce que linéaire (ou affine). Si le degré est plus grand ou égal à 2, il faut encore montrer que le polynôme n'est pas uniformément continue.

Il y a au plus $n + (n-1) + \cdots + 1$ point distincts où P ou une de ses dérivées s'annule. Au delà de ces points, P et ses dérivées ne changent plus de signe, et sont donc toutes positives ou toutes négatives. En particulier, P' est une fonction qui tend soit vers l'infini soit vers moins l'infini (c'est ici que nous utilisons le fait que le degré est plus grand ou égal à 2). Soit x dans cette zone, et notons m, le minimum de P' sur $B(x,\delta)$, alors

$$|P(x) - P(x+\delta)| > m\delta, \tag{97.232}$$

et il suffit de prendre x assez grand pour que $m > \epsilon/\delta$ pour contredire l'uniforme continuité.

Exercise 94 exo0096

Soit f une fonction continue de $[-1,1]$ dans $\mathbb{R}$, dérivable sur $]-1,1[$ et telle que $f(-1)=f(1)=0$. En utilisant la fonction auxiliaire $g(x)=e^{\lambda x}f(x)$ montrez que pour tout $\lambda\in\mathbb{R}$, il existe $a\in]-1,1[$ tel que $f'(a)+\lambda f(a)=0$. corr0096

Correction of the exercise 94

Nous cherchons deux points x_1 et x_2 tels que

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}=-\lambda f(a). \tag{97.233}$$

Étant donné que $g(-1)=g(1)=0$, le théorème des accroissements finis dit qu'il y a un point a dans $]-1,1[$ tel que $g'(a)=0$. Mais

$$g'(x)=\lambda e^{\lambda x}f(x)+e^{\lambda x}f'(x), \tag{97.234}$$

donc la condition $g'(a)=0$ dit que $f'(a)=-\lambda f(a)$.

Exercise 95 exo0097

Soit f une fonction inversible définie sur un sous-ensemble connexe par arc I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Démontrez que si f est continue sa réciproque l'est aussi.

Indication : commencer par prouver que f est monotone croissante ou décroissante. corr0097

Correction of the exercise 95

Prouvons que f est nécessairement monotone, c'est-à-dire que $f(y)-f(x)$ a un signe constant tant que $y>x$. L'application $(x,y)\mapsto f(y)-f(x)$ est une application continue de $I\times I$ vers $\mathbb{R}$ qui ne passe par zéro que lorsque $x=y$ (parce que f est inversible), donc elle ne peut pas changer de signe dans la partie $y>x$.

Maintenant, le théorème de la bijection 12.51 prouve l'exercice.

Exercise 96 exo0098

Montrez que dans le théorème précédent l'hypothèse de connexité de l'ensemble I est capitale. Trouvez une fonction f définie sur un ensemble non connexe par arc dont la réciproque existe mais n'est pas continue. corr0098

Correction of the exercise 96

La fonction

$$f(x)=\begin{cases}x & \text{si } x<1\\ 3-x & \text{si } 1<x\leqslant 2\end{cases} \tag{97.235}$$

Elle vérifie $f^{-1}(1)=2$, mais $\lim_{y\to 1^-}f^{-1}(y)=1$, ce qui prouve que f^{-1} n'est pas continue.

Exercise 97 exo0099

Soit f une fonction définie sur tout $\mathbb{R}$ qui satisfait les trois propriétés suivantes :

(1) f est dérivable et $f'=f$

(2) $f(0)=1$

(3) $f(x+y)=f(x)f(y)\qquad\forall x,y\in\mathbb{R}$

Démontrez que $Im(f)=\mathbb{R}_0^+$. corr0099

Correction of the exercise 97

D'abord, remarquons que $f'(0)=1$, donc $f'>0$ sur un voisinage de 0. Donc, si ϵ est dans ce voisinage, $f(\epsilon)>1$. Disons $f(\epsilon)=1+\delta$. En utilisant la propriété d'additivité, nous avons

$$f(\epsilon+\epsilon)=f(\epsilon)^2=(1+\delta)^2>1, \tag{97.236}$$

et plus généralement,

$$f(n\epsilon)=(1+\delta)^n. \tag{97.237}$$

Évidement, $(1+\delta)^n$ tend vers l'infini lorsque n tend vers l'infini. Nous en déduisons que $[1,\infty[\subset$ Image(f).

Reste à voir que les valeurs entre 0 et 1 sont dans l'image et que les valeurs négatives n'y sont pas. Afin de voir que les valeurs entre 1 et 0 sont dans l'image de f, nous prenons le même argument que précédemment, mais en utilisant le fait que $f(-\epsilon)=1-\delta$ et que $(1-\delta)^n\to 0$ lorsque $n\to\infty$.

Montrons que $f(x)\leqslant 0$ n'est pas possible. Bien entendu, $f(x)=0$ n'est pas possible parce que nous aurions, pour tout y l'identité $f(x+y)=f(x)f(y)=0$, et donc $f=0$.

Si $f(x_0)<0$, alors il existe $x_1>x$ tel que $f(x_1)=0$ (parce que nous avons déjà montré que $f(n\epsilon)\to\infty$ quand $n\to\infty$). Prenons

$$x_1=\inf\{x>x_0 \text{ tel que } f(x)=0\}. \tag{97.238}$$

Étant donné que f est continue sur $[x_0,x_1]$, il existe $c\in\,]x_0,x_1[$ tel que $f'(c)>0$. Mais, par définition de x_1, nous avons $f(c)<0$, cela contredit la propriété $f'=f$.

Exercise 98 exo0100

Montrer que toute fonction Lipschitzienne sur $\mathbb{R}$ est uniformément continue. La réciproque est-elle vraie ?

97.9 Développements de Taylor et Maclaurin

Exercise 99 exoDevel0001

Afin de se familiariser avec la notion de fonction de type $o(x^n)$.

(1) Donner un exemple de fonction satisfaisant :
  (1a) $f(x)\in o(|x^3|)$,
  (1b) $f(x)\in o(|x-1|)$,
  (1c) $f(x)\in o(\sin x)$.

(2) Vrai ou faux ? (Justifier ou donner un contre-exemple) :
  (2a) Si $f(x)\in o(x)$ alors $xf(x)\in o(x^2)$.
  (2b) Si $f(x)\in o(x)$ alors $\frac{f(x)}{x}\in o(x)$.
  (2c) Si $f(x)\in o(x^3)$ alors $f(x)\in o(x^2)$, $f(x)\in o(x)$, et $f(x)\in o(1)$.

corrDevel0001

Correction of the exercise 99

(1) Pour la première vague, nous avons
  (1a) $f(x)=x^4$ fait l'affaire parce que $\frac{x^4}{x^3}\to 0$
  (1b) Sous-entendu, nous demandons une fonction dans $o\big(|x-1|\big)$ pour $x\to 1$. La fonction $f(x)=(x-1)^2$ fonctionne.
  (1c) La fonction $f(x)=x\sin(x)$.

(2) Pour la deuxième vague,
  (2a) Nous avons
$$\lim_{x\to 0}\frac{xf(x)}{x^2}=\lim_{x\to 0}\frac{f(x)}{x}=0, \tag{97.239}$$
  donc VRAI.
  (2b) Nous avons $\lim_{x\to 0} f(x)/x^2\neq 0$, mais $x^2\in o(x)$, donc FAUX.
  (2c) En utilisant l'hypothèse $\lim_{x\to 0} f(x)/x^3$, nous trouvons
$$\begin{aligned}\lim_{x\to 0}\frac{f(x)}{x^2}&=\lim_{x\to 0}x\frac{f(x)}{x}=0\\ \lim_{x\to 0}\frac{f(x)}{x}&=\lim_{x\to 0}x^2\frac{f(x)}{x^3}=0\\ \lim_{x\to 0}\frac{f(x)}{1}&=\lim_{x\to 0}x^3\frac{f(x)}{x^3}=0.\end{aligned} \tag{97.240}$$

Exercise 100 exoDevel0002
Soit $f(x) = e^x$.

(1) Déterminez les développements de Taylor d'ordre 1 et 2 de la fonction f autour de 0.
(2) Esquissez le graphe de f et de ces deux développements de Taylor.
(3) Calculez $e^{0,1}$ avec une erreur inférieure à $0,001$

corrDevel0002

Correction of the exercise 100
Nous devons écrire

$$P(x-a) = f(a) + f'(a)(x-a) + \frac{f''(x)}{2!}(x-a)^2 \tag{97.241}$$

avec $a = 0$ et $f(x) = e^x$. Étant donné que les dérivées de tout ordre de l'exponentielle sont e^x, nous avons toujours $f^{(k)}(0) = 1$, et donc

$$P(x) = 1 + x + \frac{x^2}{2}. \tag{97.242}$$

Les graphes sont sur la figure 97.5.

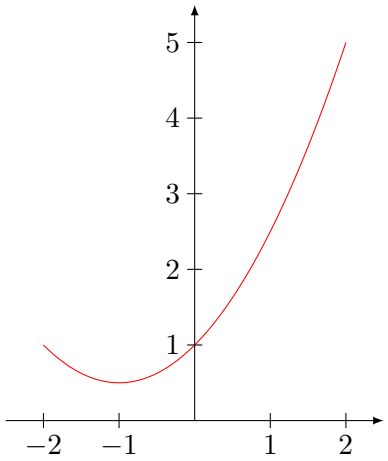


FIGURE 97.5: En bleu : la fonction $x \mapsto e^x$, en cyan l'approximation d'ordre 0, en vert son développement d'ordre 1 et en rouge celui d'ordre 2. LabelFigLaurin

Afin d'évaluer l'erreur commise entre $e^{0,1}$ et $P(0,1)$, nous calculons le reste donné par la proposition 12.472. Il existe un c entre 0 et $0,1$ tel que

$$R_2(x) = \frac{e^c}{3} \cdot 0,001, \tag{97.243}$$

où c est entre 0 et $0,1$. La valeur de e^c est bornée par $e < 3$, et donc

$$R_2(0,1) < 0,001, \tag{97.244}$$

et l'estimation

$$e^{0,1} = 1 + x + \frac{x^2}{2} = 1 + 0,1 + \frac{0,01}{2} = 1,105 \tag{97.245}$$

est acceptable. En effet, la valeur « exacte » est $1,10517$.

Exercise 101 exoDevel0003
En utilisant un développement de Taylor de la fonction $[x \to \ln(x)]$ autour de 1

(1) Calculez $\ln(1,2)$ avec une erreur inférieure à $0,0001$.
(2) Estimez l'erreur commise dans le calcul de $\ln(1,2)$ si on remplace $\ln(x)$ par son développement de Taylor d'ordre 384 autour de 1.

corrDevel0003

Correction of the exercise 101

Le développement du logarithme autour de 1 est donné par

$$P(x-1) = x - 1 - \frac{(x-1)^2}{2!} + \frac{(x-1)^3}{3!}, \tag{97.246}$$

EqDevLn

et est supposé approximer la fonction $\ln(x)$ autour de $x = 1$. En posant $x = 1$ dans (97.246), nous trouvons 0, ce qui est la valeur exacte du logarithme en $x = 1$. Notons toutefois qu'en remplaçant avec $x = 0$, nous trouvons un nombre qui est, en norme, borné par e, ce qui est une nettement plus mauvaise approximation de $\ln(0) = -\infty$.

L'erreur commise en remplaçant $\ln(x)$ par le développement à l'ordre 384 est donné par

$$R_{384}(1,2) = \frac{\ln^{(384)}(c)}{(385)!} \cdot (0,2)^{385} = \frac{384!}{c^{384}} \frac{1}{385!} \cdot (0,2)^{385} \tag{97.247}$$

pour un certain $c \in]1, 1.2[$. Cette valeur est majorée par

$$\frac{(0,2)^{385}}{385}. \tag{97.248}$$

Exercise 102 exoDevel0004

Donnez le développement de Taylor d'un polynôme autour de $a \in \mathbb{R}$ à l'ordre k.

Exercise 103 exoDevel0009

D'autres exercices pour comprendre les fonctions du type o.

(1) Prouver que si $f \in o(x^k)$, alors $f \circ \varphi \in o\big(\varphi(x)^k\big)$ lorsque $\lim_{x\to 0} \varphi(x) = 0$.
(2) Si $f \in o(x^k)$, montrer que $f(x)^n \in o(x^{kn})$.
(3) Montrer que $f \in o(x^n)$ et si $g \in o(x^k)$, alors $f \circ g \in o(x^{kn})$.

corrDevel0009

Correction of the exercise 103

(1)
(2) Cela est juste un calcul :

$$\lim_{x\to 0} \frac{f(x)^n}{x^{kn}} = \lim_{x\to 0} \left(\frac{f(x)}{x^k} \right)^n = 0. \tag{97.249}$$

(3) Nous avons

$$\lim_{x\to 0} \frac{f\big(g(x)\big)}{x^{kn}} = \lim_{x\to 0} \frac{g(x)^n}{x^{kn}} \frac{f\big(g(x)\big)}{g(x)^n}. \tag{97.250}$$

EqLimFcirtcGkn

Le premier point montre que $\lim_{x\to 0} f\big(g(x)\big)/g(x)^n = 0$, et le second point montre que $g(x)^n \in o(x^{kn})$, donc la limite (97.250) est nulle.

Exercise 104 exoDevel0005

Déterminez les séries de Mc Laurin (séries de Taylor autour de 0) des fonctions suivantes

(1) $x \mapsto \sin(x)$
(2) $y \mapsto \cos(y)$
(3) $t \mapsto e^t$
(4) $x \mapsto \ln(1+x)$
(5) $t \mapsto \sqrt{1+t}$
(6) $y \mapsto \frac{1}{1+y}$

corrDevel0005

Correction of the exercise 104

(1)
(2)
(3)

(4)

(5) $\sqrt{1+t} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + o(x^3)$

Exercise 105 exoDevel0006

Déterminez les développements de Mc Laurin jusqu'à l'ordre 3 des fonctions suivantes (utilisez les résultats de l'exercice 104).

(1) $t \mapsto 3t^4 + t^3 + t^2 + t$

(2) $x \mapsto \frac{\cos(x)-1}{x}$

(3) $x \mapsto \frac{\cos(2x)}{\sin(x)+1}$

(4) $y \mapsto \frac{y^2+y}{\cos(y)}$

(5) $x \mapsto \sqrt{\cos(2x)}$

corrDevel0006

Correction of the exercise 105

Ici, le point crucial qui joue est l'unicité du polynôme dont il est question dans la proposition 12.471. Nous utiliserons aussi toutes les propriétés de combinaisons des o.

(1) Le principe d'unicité dit qu'il existe un seul polynôme $P(t)$ d'ordre 3 tel que

$$3t^4 + t^3 + t^2 + t = P(t) + o(t^3), \tag{97.251}$$ EqCondPOlupolyTay

mais $3t^4 \in o(t^3)$, donc en fait le polynôme $P(t) = t^3 + t^2 + t =$ vérifie la condition (97.251). C'est donc lui, le développement de Taylor d'ordre 3 de $3t^4 + t^3 + t^2 + t$.

(2) Nous reprenons le développement de $\cos(x) - 1$ de l'exercice 104. Nous avons

$$\cos(x) - 1 = -\frac{x^2}{2} + \frac{x^4}{4!} + o(x^4), \tag{97.252}$$

donc

$$\frac{\cos(x) - 1}{x} = -\frac{x^3}{2} + \frac{x^3}{4!} + \frac{o(x^4)}{x} \tag{97.253}$$

Étant donné que $o(x^4)/x = o(x^3)$, nous avons que

$$\frac{\cos(x) - 1}{x} = -\frac{x^3}{2} + \frac{x^3}{4!} + o(x^3). \tag{97.254}$$

Par unicité,

$$-\frac{x^3}{2} + \frac{x^3}{4!} \tag{97.255}$$

est le développement de Taylor de la fonction proposée.

(3)

(4)

(5) Étant donné les solutions de l'exercice 105, nous savons que

$$\begin{aligned}\sqrt{1 + \big(\cos(x) - 1\big)} &= \sqrt{1 + \big(-2x^2 + \frac{2x^4}{3} + o(x^4)\big)} \\ &= 1 + \frac{A}{2} - \frac{A^2}{8} + \frac{A^3}{16} + o(A^3),\end{aligned} \tag{97.256}$$

si nous appelons A la quantité $-2x^2 + \frac{2x^4}{3} + o(x^4)$. Nous ne cherchons que l'ordre 3, donc nous voulons débusquer tous les termes en $o(x^3)$, et les rassembler sous le seul symbole « $+o(x^3)$ ». En réalité, $A = -2x^2 + o(x^3)$, donc $A^2 \in o(x^3)$, et $o(A^3) \in o(x^3)$, il ne reste que

$$\sqrt{1 + \big(\cos(x) - 1\big)} = 1 - x^2 + o(x^3). \tag{97.257}$$

Note : si nous avions cherché les termes jusqu'à l'ordre 4, alors il aurait fallu par exemple garder le terme en x^4 dans A^2, et au final,

$$\sqrt{1 + \big(\cos(x) - 1\big)} = 1 - x^2 - \frac{x^4}{4} + o(x^4). \tag{97.258}$$

Exercise 106 exoDevel0007

Déterminez les développements de Taylor jusqu'à l'ordre 2 des fonctions suivantes.

(1) $(x,y)\mapsto(1+x+y)^2$ autour de $(0,0)$
(2) $(x,y)\mapsto(1+x+y)^2$ autour de $(-1,0)$
(3) $(u,v)\mapsto(u+v)e^u$ autour de $(0,1)$
(4) $(u,v)\mapsto\cos(uv)$ autour de $(0,0)$
(5) $(u,v)\mapsto\sin(vu^2)v^3$ autour de $(0,0)$

Exercise 107 exoDevel0008

Supposons que $f\in o(t^{23})$ et $g(t)\in o(t^3)$. Montrer que

(1) $\frac{f(x)}{x^5}\in o(x^{18})$
(2) $x^3g(x)\in o(x^6)$
(3) $f(t)g(t)\in o(t^{26})$
(4) $f(t^2)\in o(t^{46})$
(5) $f(g(t))\in o(t^{69})$
(6) $o(t^9)\subset o(t^5)$
(7) $(u+v)f(u)\in o(\|(u,v)\|^{24})$
(8) $f(uv)\in o(\|(u,v)\|^{46})$

corrDevel0008

Correction of the exercise 107

(1)

(2) Considérons la fonction (continue) F donnée par

$$F(t)=\begin{cases}\frac{f(t)}{t^k} & \text{si } t\neq 0\\ 0 & \text{si } t=0.\end{cases}\tag{97.259}$$

La fonction que nous étudions est $(F\circ\varphi)(x)$. Étant donné la continuité de F, nous avons

$$\lim_{x\to0}(F\circ\varphi)(x)=F\big(\lim_{x\to0}\varphi(x)\big)=F(0)=0.\tag{97.260}$$

Exercise 108 exoreserve0002

Calculer

$$\lim_{(x,y)\to(t,t)}\frac{\sin(x)-\sin(y)}{x-y}\tag{97.261}$$

pour chaque $t\in\mathbb{R}$. Peut-on prolonger f par continuité sur la droite $x=y$? Conseil : utiliser la formule des accroissements finis pour la différence $\sin(x)-\sin(y)$. corrreserve0002

Correction of the exercise 108

Nous nous mettons dans une boule B centrée en (t,t) de rayon suffisamment petit pour que tous les x et y soient dans la même période de la fonction sinus. Pour chaque (x,y), il existe un $\xi\in[x,y]$ tel que

$$\sin(x)-\sin(y)=\cos(\xi)(x-y).\tag{97.262}$$

Un tel ξ existant pour chaque couple (x,y) nous pouvons considérer une fonction ξ définie sur B telle que $\sin(x)-\sin(y)=\cos\big(\xi(x,y)\big)(x-y)$. Certes, il existe plusieurs telles fonctions (en particulier n'importe quelle valeur convient en $\xi(t,t)$), mais nous en considérons une. Nous avons donc

$$\frac{\sin(x)-\sin(y)}{x-y}=\cos\big(\xi(x,y)\big).\tag{97.263}$$

Étant donné que $\xi(x,y)\in[x,y]$ lorsque $x\neq y$, nous avons la limite

$$\lim_{(x,y)\to(t,t)}\xi(x,y)=t.\tag{97.264}$$

En effet, soient $\epsilon < 0$ et δ tels que pour tout $(x,y) \in B\big((t,t),\delta\big)$, nous ayons $|x-y| < \epsilon$. Nous avons alors

$$\big|\xi(x,y) - x\big| < |y-x| < \epsilon. \tag{97.265}$$

Notez que nous avons établi une limite pour ξ en ayant aucune idée de sa continuité ! Finalement nous avons

$$\lim_{(x,y)\to(t,t)} f(x,y) = \lim_{(x,y)\to(t,t)} \cos\big(\xi(x,y)\big) = \cos\big(\lim_{(x,y)\to(t,t)} \xi(x,y)\big) = \cos(t). \tag{97.266}$$

La fonction

$$f(x,y) = \frac{\sin(x) - \sin(y)}{x-y} \tag{97.267}$$

peut donc être continument prolongée par $\cos(x)$ sur la ligne $x = y$.

97.10 Optimisation sans contraintes

Exercise 109 exoOptimSS0001

Déterminez les extrémums locaux et globaux des fonctions suivantes

(1) $\mathbb{R} \to \mathbb{R} : x \mapsto e^x \cos(x)$
(2) $\mathbb{R}_0^+ \to \mathbb{R} : x \mapsto \frac{\ln(x)}{x}$
(3) $\mathbb{R} \to \mathbb{R} : t \mapsto \sin^3(t)$
(4) $\mathbb{R} \to \mathbb{R} : t \mapsto (1-\cos(t))^2$
(5) $]-3,3[\to \mathbb{R} : t \mapsto t^3 - 3t + 2$
(6) $[-3,3] \to \mathbb{R} : t \mapsto t^3 - 3t + 2$
(7) $[-2,5[\to \mathbb{R} : y \mapsto 3x^5 - 10x^3 - 45x + 7$
(8) $\mathbb{R} \to \mathbb{R} : x \mapsto \frac{|x+1|}{(2+2x+x^2)}$
(9) $\mathbb{R} \to \mathbb{R} : x \mapsto \sqrt[3]{x^2}$

corrOptimSS0001

Correction of the exercise 109

(1) La fonction $e^x \cos(x)$ n'a pas d'extrémum globaux, mais la dérivée

$$f'(x) = e^x\big(\cos(x) - \sin(x)\big) \tag{97.268}$$

s'annule en $a_i = \frac{\pi}{4} + 2k\pi$ et $b_i = \frac{5\pi}{4} + 2k\pi$. Étant donné que $f''(x) = -2e^x \sin(x)$, les premiers points a_i sont des maximums locaux, et les points b_i sont des minimums locaux.

(2) La dérivée de f est $f'(x) = \frac{1-\ln(x)}{x^2}$ et s'annule en $x = e$, qui et un point de maximum global.

(3)

(4)

(5) La dérivée de f s'annule en $t = \pm 1$, où la dérivée seconde vaut respectivement 6 et -6, ce qui fait que 1 est minimum local, et -1 est maximum local. Ils ne sont cependant pas extrémum globaux.

(6) Idem que le point précédent, mais -3 est minimum global et 3 est maximum global.

(7)

(8) Le dénominateur ne s'annule jamais, donc c'est une fonction qui s'étudie de façon usuelle pour $x > 1$ et pour $x < -1$ séparément (à cause de la valeur absolue). Il faut cependant étudier le point $x = -1$ de façon séparée, parce que la fonction n'y est pas dérivable.

Nous avons $f(-1) = 0$, alors que la fonction est toujours strictement positive ailleurs. Nous concluons que $x = -1$ est minimum global.

Exercise 110 exoOptimSS0002

Soit $f : [a,b] \to \mathbb{R}$. Démontrez que si $f'(b) > 0$ alors b est un maximum local de la fonction f.

corrOptimSS0002

Correction of the exercise 110

La définition de la dérivée sur le bord du domaine de définition (voir page 74) est

$$f'(b) = \lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \frac{f(b-\epsilon)-f(b)}{\epsilon}. \tag{97.269}$$

L'hypothèse dit que ce nombre est strictement positif, donc $\exists \epsilon_0$ tel que $\epsilon < \epsilon_0$ implique

$$\frac{f(b-\epsilon)-f(b)}{\epsilon} > 0, \tag{97.270}$$

et donc $f(b-\epsilon)-f(b)>0$, ce qui fait que f est un maximum.

Exercise 111 exoOptimSS0003

Déterminez les extrémums locaux et globaux des fonctions suivantes

(1) $\mathbb{R}^2 \to \mathbb{R} : (x,y) \mapsto x^3 - y^3 + 3x^2 + 3y^2 - 9x$
(2) $\mathbb{R}^2 \to \mathbb{R} : (x,y) \mapsto x^3 + y^3 + 3xy + 43$
(3) $\{(x,y) \in \mathbb{R}^2 \mid x>0, y>0, x+y<1\} \mapsto \mathbb{R} : (x,y) \mapsto x^3 - 3y^2 + x - 2y - 1$
(4) $\{(x,y) \in \mathbb{R}^2 \mid x\geqslant 0, y\geqslant 0, x+y\leqslant 1\} \mapsto \mathbb{R} : (x,y) \mapsto x^3 - 3y^2 + x - 2y - 1$

corrOptimSS0003

Correction of the exercise 111

(1) Notons tout de suite que la fonction f n'adment pas d'extrémum globaux. En effet nous avons la limite $\lim_{x\to\pm\infty} f(x,0) = \pm\infty$. La différentielle de f est donnée par

$$df(x,y) = (3x^2 + 6x - 9, -3y^2 + 6y), \tag{97.271}$$

qui s'annule en $(-3,0)$, $(-3,2)$, $(1,0)$, $(1,2)$. Ce sont donc ces seuls points qui sont susceptibles d'être des extrémums locaux. La matrice des dérivées secondes est

$$d^2f(x,y) = \begin{pmatrix} 6x+6 & 0 \\ 0 & -6y+6 \end{pmatrix}. \tag{97.272}$$

Les valeurs propres sont données par l'équation

$$\det\left(d^2f(x,y) - \lambda\mathbb{1}\right) = \begin{vmatrix} 6x+6-\lambda & 0 \\ 0 & -6y+6-\lambda \end{vmatrix}, \tag{97.273}$$

dont les solutions sont $\lambda_1(x,y) = 6x+6$ et $\lambda_2(x,y) = -6y+6$.

Il suffit maintenant de calculer λ_1 et λ_2 pour les différents points critiques. Nous avons

$$\begin{cases} \lambda_1(-3,0) = -12 & \text{(97.274a)} \\ \lambda_2(-3,0) = 6, & \text{(97.274b)} \end{cases}$$

$$\begin{cases} \lambda_1(-3,2) = -12 & \text{(97.275a)} \\ \lambda_2(-3,2) = -6, & \text{(97.275b)} \end{cases}$$

$$\begin{cases} \lambda_1(1,0) = 12 & \text{(97.276a)} \\ \lambda_2(1,0) = 6, & \text{(97.276b)} \end{cases}$$

$$\begin{cases} \lambda_1(1,2) = 12 & \text{(97.277a)} \\ \lambda_2(1,2) = -6, & \text{(97.277b)} \end{cases}$$

Le point $(1,0)$ est donc minimum local parce que d^2f y est définie positive, et le point $(-3,2)$ est maximum local parce que d^2f y est définie négative.

(2) La différentielle est
$$df = (3x^2 + 3y; 3y^2 + 3x), \tag{97.278}$$
donc les points critiques sont donnés par le système
$$\begin{cases} x^2 + y = 0 & (97.279a) \\ y^2 + x = 0, & (97.279b) \end{cases}$$
dont les solutions réelles sont $(0,0)$ et $(-1,-1)$. La matrice des dérivées secondes est donnée par
$$d^2 f(x,y) = \begin{pmatrix} 6x & 3 \\ 3 & 6y \end{pmatrix}, \tag{97.280}$$
dont nous devons chercher les valeurs propres. L'équation $\det(A - \lambda \mathbb{1}) = 0$ est
$$(6x - \lambda)(6y - \lambda) - 9 = 0, \tag{97.281}$$
et les solutions sont
$$\lambda_\pm(x,y) = 3(x+y) \pm 3\sqrt{y^2 - 2xy + y^2 + 1}. \tag{97.282}$$
Étant donné que $\lambda_\pm(0,0) = \pm 3$, ce point n'est ni un maximum ni un minimum local. L'autre point critique est par contre un maximum local strict parce que $\lambda_+(-1,-1) = -3$ et $\lambda_-(-1,-1) = -9$, ce qui fait que $d^2 f(-1,-1)$ est définie négative.

(3) La différentielle vaut
$$df = (3x^2 + 1, -6y - 2), \tag{97.283}$$
qui ne s'annule jamais. Il n'y a donc aucun extrémum local. Notons aussi que pour tout $(x,y) \in \mathbb{R}^2$, nous avons, pour tout $\epsilon > 0$
$$\begin{aligned} f(x+\epsilon, y) &> f(x,y) \\ f(x-\epsilon, y) &< f(x,y), \end{aligned} \tag{97.284}$$
le domaine étant ouvert, il n'y a donc pas d'extrémum global.

(4) Cette fois, le domaine est compact, et il existe certainement un maximum et un minimum global, qu'il faut aller chercher sur les bords.

Exercise 112 exoOptimSS0004
Soit $\{(x_i, y_i) \in \mathbb{R}^2 \mid 1 \leqslant i \leqslant n\}$ un ensemble fini de points de $\mathbb{R}^2$.

(1) Déterminez la droite D pour laquelle la somme des carrés des distances entre les points et leurs projection parallèlement à l'axe y sur la droite D est minimale.

(2) Appliquez le point un aux données suivantes $(x_1, y_1) = (1, \frac{1}{2})$, $(x_2, y_2) = (3, 1)$, $(x_2, y_2) = (5, 2)$.

corrOptimSS0004

Correction of the exercise 112

Une droite verticale n'est solution que si tous les points donnés sont alignés verticalement. Nous cherchons donc une droite oblique sous la forme $y = ax + b$, où a et b sont les paramètres inconnus que nous voulons optimiser.

Le point (x_i, y_i) se projette verticalement sur la droite sur le point $(x_i, ax_i + b)$, et la distance est donc $y_i - ax_i - b$. La fonction à minimiser est donc
$$f(a,b) = \sum_i (y_i - ax_i - b)^2. \tag{97.285}$$
Les différentielles sont
$$\begin{aligned} \partial_a f(a,b) &= 2\sum_i (y_i - ax_i - b)(-x_i) \\ \partial_b f(a,b) &= 2\sum_i (y_i - ax_i - b)(-1) \end{aligned} \tag{97.286}$$

Les points critiques sont donnés par la solution du système d'équation

$$\begin{cases} (\sum_i x_i^2)a + (\sum_i x_i)b = \sum_i y_i x_i & (97.287a) \\ (\sum_i x_i)a + nb = \sum_i y_i. & (97.287b) \end{cases}$$

Exercise 113 exoOptimSS0005

Calculez le volume de la plus spacieuse boîte de forme parallélépipède rectangle dont trois côtés sont sur les axes, dont un coin est l'origine du système d'axes orthonormés $0xyz$ et dont le coin opposé (x, y, z) est dans le morceau de plan $\{(x, y, z) \in (\mathbb{R}_0^+)^3 \mid \frac{x}{2} + \frac{y}{3} + \frac{z}{4} = 1\}$. corrOptimSS0005

Correction of the exercise 113

La fonction à minimiser est $V(x, y, z) = xyz$, qui se réduit à un problème à deux variables en utilisant la contrainte,

$$V(x, y) = 4xy - 2x^2y - \frac{4}{3}xy^2, \tag{97.288}$$

qui est une fonction usuelle de $\mathbb{R}^2$ vers $\mathbb{R}$ à maximiser. Les équations pour les points critiques sont

$$\begin{aligned} \partial_x V &= 4y - 4xy - \frac{4}{3}y^2 = 0 \\ \partial_y V &= 4x - 2x^2 - \frac{8}{3}xy = 0. \end{aligned} \tag{97.289}$$

Évidement, nous refusons toute solution avec x ou y nuls, donc nous pouvons simplifier la première par y et la seconde par x, ce qui donne le système

$$\begin{cases} 2 - 2x - \frac{2}{3}y = 0 & (97.290a) \\ 2 - x - \frac{4}{3}y = 0. & (97.290b) \end{cases}$$

La solution est $(\frac{2}{3}, 1)$. Étant donné que nous regardons V sur un compact, et que le maximum n'est certainement pas sur un des bords (V s'y annule), le maximum global (qui existe par compacité) doit être à l'intérieur, et donc sur un maximum local. Il ne peut donc être que $(\frac{2}{2}, 1)$.

Nous pouvons le vérifier directement :

$$d^2V(x, y) = \begin{pmatrix} -4y & 4 - 4x - \frac{8}{3}y \\ 4 - 4x - \frac{8}{3}y & -\frac{8}{3}x, \end{pmatrix} \tag{97.291}$$

donc

$$d^2V(\frac{3}{3}, 1) = 4\begin{pmatrix} -1 & -1/9 \\ -1/9 & -2/3 \end{pmatrix}, \tag{97.292}$$

et ses valeurs propres sont

$$\lambda_\pm = \frac{-15 \pm \sqrt{13}}{18}, \tag{97.293}$$

qui sont, effectivement, toutes deux négatives, ce qui prouve un maximum global.

Exercise 114 exoOptimSS0006

Déterminer les extrémums (locaux et globaux) et points de selle éventuels de la fonction

$$f(x, y) = (x - 1)y(y - x) \tag{97.294}$$

dans le domaine **fermé** $D \subset \mathbb{R}^2$ limité par les droites $y = x$, $y = 0$ et $x = 1$. corrOptimSS0006

Correction of the exercise 114

D'abord, remarquer que $f \equiv 0$ sur le bord du domaine. La fonction est également toujours positive à l'intérieur du domaine. Nous avons donc minimum global sur les bords, et nous ne recherchons que des minimums locaux à l'intérieur. Les équations pour les points critiques sont

$$\begin{cases} \partial_x f(x, y) = y(y - 2x + 1) = 0 & (97.295a) \\ \partial_y f(x, y) = (x - 1)(2y - x) = 0, & (97.295b) \end{cases}$$

dont les solutions sont les points $(0,0)$, $(\frac{2}{3},\frac{1}{3})$, $(1,0)$, $(1,1)$. Parmi eux, seul $(\frac{2}{3},\frac{1}{3})$ n'est pas sur le bord, donc c'est lui le maximum local.

97.11 Équations différentielles

97.11.1 Équations différentielles du premier ordre

Exercise 115 exoEqsDiff0001

Pour quelles valeurs des constantes a et b la fonction $t \mapsto t^4 e^{2t}$ est-elle une solution de l'équation différentielle

$$y'' + 2ay' + by = 12t^2 e^{2t}$$

corrEqsDiff0001

Correction of the exercise 115

En calculant les dérivées successives de la solution imposée, nous trouvons

$$\begin{aligned} y(t) &= t^4 e^{2t} \\ y'(t) &= (2t^4 + 4t^3)e^{2t} \\ y''(t) &= (4t^4 + 16t^3 + 12t^2)e^{2t}. \end{aligned} \tag{97.296}$$

En injectant le tout dans l'équation, nous trouvons l'équation

$$(4 + 4a + b)t^4 + (16 + 8a)t^3 = 0. \tag{97.297}$$

Cette équation doit être vraie pour tout t, donc nous en déduisons $a = -2$ et $b = 4$.

Exercise 116 exoEqsDiff0002

Déterminez l'ensemble des fonctions qui satisfont les équations différentielles suivantes.

ItemaEqsDiff00002

(1) $y' = (t^3 + t)e^{-y}$
(2) $y' = 1 + y^2$
(3) $y' = \frac{\cos(t)}{1+e^y}$
(4) $y' = y^2$ zero
(5) $y' = y^{\frac{1}{3}}$ unique

ItemfEqsDiff00002

(6) $yy' + (1 + y^2)\sin(t) = 0$ crasse

corrEqsDiff0002

Correction of the exercise 116

(1) Nous manipulons un tout petit peu l'équation en utilisant $y' = dy/dt$:

$$\begin{aligned} e^y y' &= (t^3 + t) \\ e^y dy &= (t^3 + t)dt, \end{aligned} \tag{97.298}$$

que nous intégrons des deux côtés pour trouver $e^y = \frac{t^4}{4} + \frac{t^2}{2} + C$. De là, nous déduisons la solution générale :

$$y(t) = \ln\left(\frac{t^4}{4} + \frac{t^2}{2} + C\right). \tag{97.299}$$

(2) $y' = 1 + y^2$. Nous avons $u(t) = 1$ et $f(y) = 1 + y^2$ qui ne s'annule jamais, donc $I = J = \mathbb{R}$. Petit calcul :

$$G(y) = \int \frac{dy}{1 + y^2} = \arctan(y). \tag{97.300}$$

La solution est donc implicitement donnée par $\arctan\big(y(t)\big) = t + C$, et donc explicitement par

$$y(t) = \tan(t + C). \tag{97.301}$$

Notons que, dans ce cas ci, nous sommes parvenu à isoler $y(t)$ dans l'équation $G\big(y(t)\big) = U(t) + C$. Cela n'est pas toujours possible, comme nous le verrons dans d'autres exercices.

(3) Nous avons $u(t) = \cos(t)$ et $f(y) = \frac{1}{1+e^y}$, et donc $U(t) = \sin(t)$ et $G(t) = y + e^y$. La solution se présente sous forme implicite

$$y + e^y = \sin(t) + C. \tag{97.302}$$

Il n'est pas possible de résoudre cette équation pour isoler y, donc nous devons nous contenter de cette forme. Toutefois, nous pouvons nous demander sur quel domaine cette formule définit correctement la fonction $y(t)$. Il faudrait trouver les domaines I et J telles que la fonction $z \mapsto z + e^z$ soit bijective entre I et J.

Démontrons que la fonction $f\colon z \to z + e^z$ est bijective sur $\mathbb{R}$. D'abord, elle est sur surjective parce que $\lim_{z\to\pm\infty} f(z) = \pm\infty$. Ensuite, elle est injective parce que sa dérivée est toujours strictement positive. Si nous notons g son inverse, alors

$$y(t) = g\big(\sin(t) + C\big) \tag{97.303}$$

est la solution.

(4) $y' = y^2$. Nous avons $u(t) = 1$, donc $I = \mathbb{R}$ et $f(y) = y^2$, donc deux possibilités de domaines connexes où f ne s'annule pas : $J_1 = \mathbb{R}_0^-$ et $J_2 = \mathbb{R}_0^+$. Nous trouvons immédiatement que

$$G(y) = -\frac{1}{y} \quad U(t) \quad = t, \tag{97.304}$$

donc

$$y(t) = \frac{-1}{t + C} \tag{97.305}$$

EqSol105d

sont les solutions qui ne s'annulent pas.

Dès que $y(t_0) \neq 0$, nous avons $y(t) = -1/(t + C)$ sur un voisinage de t_0. Cependant, cette fonction ne s'annule jamais. Donc une fonction qui ne s'annule pas en un point ne s'annule jamais. La seule solution qui s'annule en un point est la solution identiquement nulle $y(t) = 0$ pour tout t.

La solution (97.305) est valable sur les intervalles $I_1 =]-\infty, -C[$ et $I_2 =]-C, \infty, [$. Une constante différente peut être choisie sur ces deux domaines, vu que nous n'avons de toutes façon pas la continuité en $t = -C$.

La solution générale consiste à découper $\mathbb{R}$ en une suite d'intervalles I_k et de poser

$$y(t) = \begin{cases} -\frac{1}{t+C_k} \\ \text{ou bien} \\ 0 \end{cases} \tag{97.306}$$

sur l'intervalle I_k. La seule contrainte sur le choix du découpage et des constantes C_k est que $-C_k \notin I_k$. Il y a donc *énormément* de solutions, si on n'impose pas d'être continue sur un intervalle maximum.

(5) $y' = y^{1/3}$. Nous avons $y'/y^{1/3} = 1$, et donc, si $y(t_0) \neq 0$, alors sur un voisinage de t_0, la solution est donnée par $\frac{3}{2}y^{2/3} = t + C$,

$$y(t) = \left(\frac{2t}{3} + C\right)^{3/2}. \tag{97.307}$$

L'étude du domaine de cette solution est intéressante. Ce domaine est à priori l'intervalle $A =]-\frac{3C}{2}, \infty[$. Pourquoi ne pas mettre le point $-3C/2$ dans le domaine ? Parce que $y = 0$ en ce point, et nous avons dès le départ écarté les solutions avec $y = 0$.

Cependant, même si y' n'est pas défini en $-3C/2$, il n'en reste pas moins que nous pouvons raccorder $y(t)$ avec la solution $y(t) \equiv 0$ en ce point :

$$y(t) = \begin{cases} 0 & \text{si } t \leqslant -3C/2 \\ \big(\frac{2t}{3} + C\big)^{3/2} & \text{si } t > -3C/2 \end{cases} \tag{97.308}$$

Cela est une solution continue de l'équation différentielle donnée, dont la dérivée n'existe pas en un seul point.

(6) Nous remettons l'équation sous la forme

$$y' = -\frac{1+y^2}{y}\sin(t). \tag{97.309}$$

Nous avons $f(y) = -\frac{1+y^2}{y}$ qui ne s'annule pas, mais qui n'est pas continue en $y = 0$. Donc une fonction $y\colon \mathbb{R} \to \mathbb{R}_0^+$ ou $y\colon \mathbb{R} \to \mathbb{R}_0^-$ est solution de l'équation proposée si et seulement si

$$\frac{1}{2}\ln(y^2+1) = \cos(t) + C, \tag{97.310}$$

c'est-à-dire

$$y(t) = \pm\sqrt{Ke^{2\cos(t)} - 1}. \tag{97.311}$$ EqSolGeneBacExpCis

Cela fait une double infinité de solutions : pour chaque $K \in \mathbb{R}$ et pour chaque choix de $\pm$, nous avons une solution.

Ici encore, le domaine vaut le coup d'œil. Notons tout de suite que lorsque $K < 0$, le domaine est vide. L'expression $e^{2\cos(t)}$ varie entre e^{-2} et e^2, donc

(6a) Si $K < e^{-2}$, alors le domaine est vide.

(6b) Si $K > e^2$, alors le domaine est tout $\mathbb{R}$.

(6c) Si $K \in]e^{-2}, e^2[$, alors la solution n'existe pas pour tous les t, et le domaine est une suite d'intervalles.

(6d) Si $k = e^2$, alors $y(t_0) = 0$ lorsque $\cos(t_0) = -1$, mais la solution continue à exister après t_0, parce que l'expression sous la racine redevient strictement positive. Cependant, la dérivée de la fonction

$$y(t) = \sqrt{Ke^{2\cos(t)} - 1} \tag{97.312}$$

en le point t_0 n'existe pas : elle vaut -1 à gauche et 1 à droite (voir le calcul de la sous-section 18.241). Il y a cependant moyen d'écrire une solution dont la dérivée existe en faisant

$$y(t) = \begin{cases} \sqrt{Ke^{2\cos(t)} - 1} & \text{si } t \leqslant t_0 \\ -\sqrt{Ke^{2\cos(t)} - 1} & \text{si } t > t_0 \end{cases} \tag{97.313}$$

C'est-à-dire, en combinant plusieurs choix de signes $\pm$ à différents points de l'intervalle.

Exercise 117 exoEqsDiff0003

Trouvez la ou les solutions maximales des problèmes de Cauchy suivants, à chaque fois décrivez la famille de fonctions et les intervalles de $\mathbb{R}$ sur lesquels elles sont définies (en général une fonction et un intervalle par problème de Cauchy). Démontrez soigneusement vos résultats.

(1) $y' = (t^3 + t)e^{-y}$, $y(1) = 1$
(2) $y' = 1 + y^2$, $y(0) = 0$
(3) $y' = 1 + y^2$, $y(0) = 1$
(4) $y' = \frac{\cos(t)}{1+e^y}$, $y(\frac{\pi}{2}) = 3$
(5) $y' = y^2$, $y(1) = 2$
(6) $y' = y^2$, $y(1) = 0$
(7) $y' = y^{\frac{1}{3}}$, $y(0) = -1$
(8) $y' = y^{\frac{1}{3}}$, $y(0) = 0$
(9) $yy' + (1+y^2)\sin(t) = 0$, $y(0) = 1$

corrEqsDiff0003

Correction of the exercise 117

(1) Nous avons vu dans l'exercice 116(1) que si il existe une solution, elle doit s'écrire sous la forme

$$y(t) = \ln\left(\frac{t^4}{4} + \frac{t^2}{2} + C\right) \tag{97.314}$$

où le domaine est déterminé par la valeur de la constante par l'inégalité

$$\frac{t^4}{4} + \frac{t^2}{2} + C > 0. \tag{97.315}$$

Nous demandons $y(1) = 1$, c'est-à-dire

$$y(1) = \ln\left(\frac{1}{4} + \frac{1}{2} + C\right) = 1, \tag{97.316}$$

ce qui permet de fixer la constante $C = e - \frac{3}{4}$.

(2) Nous avons déjà trouvé que $y(t) = \tan(t + C)$. La condition de Cauchy dit que $y(0) = \tan(C) = 0$, donc $C \in \{0, \pi\}$. Il y a donc les deux solutions

$$\begin{aligned} y_1(t) &= \tan(t) \\ y_2(t) &= \tan(t + \pi). \end{aligned} \tag{97.317}$$

(3) La même avec $y(0) = 1$. La condition $1 = \tan(C)$ demande $C \in \{\frac{\pi}{4}, \frac{5\pi}{4}\}$, et donne deux solutions

$$\begin{aligned} y_1(t) &= \tan(t + \frac{\pi}{4}) \\ y_2(t) &= \tan(t + \frac{5\pi}{4}), \end{aligned} \tag{97.318}$$

dont les domaines de définitions respectifs sont

$$\begin{aligned} I_1' &=]-\pi - \frac{\pi}{4}, \pi - \frac{\pi}{4}[\\ I_2' &=]-\pi - \frac{5\pi}{4}, \pi - \frac{5\pi}{4}[. \end{aligned} \tag{97.319}$$

Une bonne question à se poser est de comprendre pourquoi cela ne viole pas le théorème d'unicité de la solution.

(4) Nous savons déjà que la solution est donnée par la formule implicite

$$y + e^y = \sin(t) + C. \tag{97.320}$$

La condition de Cauchy se traduit en $3 + e^3 = \sin(\frac{\pi}{2}) + C$, ce qui donne $C = 2 + e^3$. La fonction $z \mapsto z + e^z$ étant bijective sur tout $\mathbb{R}$, cette formule définit bien $y(t)$ pour tout t.

(5) Nous avons deux solutions mutuellement exclusives : $y(t) = -1/(t+C)$ et $y \equiv 0$. La condition de Cauchy $y(1) = 2$ demande d'utiliser la première. Le calcul détermine que $C = -3/2$.
Cette solution n'existe pas en $t = \frac{3}{2}$. Donc on peut mettre n'importe quelle solution sur $]-\infty, -\frac{3}{2}[$, et puis la solution $y(t) = \frac{-1}{t-\frac{3}{2}}$ sur $]\frac{3}{2}, \infty[$.

(6) Seule la solution identiquement nulle satisfait à la condition.

(7) La condition de Cauchy amène la condition $y(0) = C^{3/2} = 1$, donc $C = 1$. La solution continue maximale est donc

$$y(t) = \begin{cases} 0 & \text{si } x \leqslant \frac{3}{2} \\ \left(\frac{2x}{3} + 1\right)^{3/2} & \text{si } x > \frac{3}{2} \end{cases} \tag{97.321}$$

(8) La solution $y \equiv 0$ fait l'affaire, mais il y en a une autre : celle avec $C = 0$ qui est

$$y(t) = \begin{cases} \left(\frac{2x}{3}\right)^{3/2} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases} \tag{97.322}$$

La seconde ligne de la définition est importante parce qu'on donne une condition de Cauchy en $t = 0$, donc nous voulons des solutions qui soient au moins dérivables en zéro, sinon le problème n'a pas de sens. Le fait que la fonction y ainsi définie soit continue est évident. Il faut vérifier qu'elle est aussi dérivable en $t = 0$. Pour cela, il suffit de voir que

$$\lim_{x \to 0^+} \frac{y(0) - y(x)}{x} = 0, \tag{97.323}$$

ce qui n'est pas très compliqué.

(9) En vertu de la solution (97.311) déjà trouvée , nous devons résoudre l'équation $\pm\sqrt{Ke^2-1}=1$ par rapport à K, donc nous devons faire le choix du signe positif et prendre $K=2e^{-2}$, qui est plus petit que e^2, mais plus grand que e^{-2}, donc la solution existe, mais n'est pas partout définie.

Exercise 118 exoEqsDiff0004

Résolvez les problèmes de Cauchy suivants

(1) $y'-2ty=t \qquad y(1)=e^{-\frac{1}{2}}$

(2) $y'+y\tan(t)=\sin(2t) \qquad y(0)=6$

(3) $y'+y\cot(x)=5e^{\cos(x)} \qquad y(\frac{\pi}{2})=-4$

(4) $y'+y\cos(t)=0 \qquad y(0)=10$

(5) $x^3y'+(2-3x^2)y=x^3$

(6) $y'-2y\cot(2x)=1-2x\cot(2x)-2\csc(2x)$

corrEqsDiff0004

Correction of the exercise 118

(1) $y'-2ty=t$ est une équation linéaire non homogène, donc nous résolvons d'abord l'équation homogène $y'_H=2ty_H$, qui amène

$$\ln(y_H)=t^2+C, \tag{97.324}$$

c'est-à-dire

$$y_H(t)=Ce^{t^2}. \tag{97.325}$$

La méthode de variation des constantes nous dit de chercher la solution du système non homogène sous la forme $y(t)=C(t)e^{t^2}$. En remplaçant dans l'équation, nous trouvons $C'(t)=te^{-t^2}$, c'est-à-dire

$$C(t)=-\frac{1}{2}e^{-t^2}+K, \tag{97.326}$$

et donc

$$y(t)=(-\frac{1}{2}e^{-t^2}+K)e^{t^2}=-\frac{1}{2}+Ke^{t^2}. \tag{97.327}$$

Le problème de Cauchy dit de résoudre

$$y(1)=-\frac{1}{2}+Ke=e^{-1/2}, \tag{97.328}$$

donc de prendre $K=e^{-3/2}+\frac{1}{2e}$.

(2) Nous commençons par résoudre l'équation homogène $y'_H+y_H\tan(t)=0$. Cela donne

$$\ln(y_H)=-\ln\left(\frac{1}{\cos(t)}\right)+C. \tag{97.329}$$

Ici, nous effectuons un changement de nom pour la constante : $C\to\ln(C)$, et nous utilisons la formule de somme des logarithmes. Nous obtenons $\ln(y_H)=\ln(C\cos(y))$, et donc

$$y_H(t)=C\cos(t) \tag{97.330}$$

comme solution générale de l'équation homogène. En ce qui concerne la solution de l'équation non homogène, nous posons

$$\begin{aligned} y(t)&=K(t)\cos(t)\\ y'(t)&=K'\cos(t)-K\sin(t), \end{aligned} \tag{97.331}$$

que nous remettons dans l'équation de départ :

$$K'\cos(t)-K\sin(t)+\frac{K\cos(t)\sin(t)}{\cos(t)}=\sin(2t). \tag{97.332}$$

Les deux termes en K non dérivé se simplifient et il reste

$$K'(t) = \frac{\sin(2t)}{\cos(t)}. \tag{97.333}$$

En utilisant la formule $\sin(2t) = 2\sin(t)\cos(t)$, cette équation s'intègre facilement. La réponse est :

$$K(t) = -2\cos(t) + C, \tag{97.334}$$

et la solution du problème non homogène est

$$y(t) = K(t)\cos(t) = \big(C - 2\cos(t)\big)\cos(t). \tag{97.335}$$

La résolution du problème de Cauchy est $y(0) = -2 + C = 6$, donc $C = 8$.

(3) $y' + y\operatorname{cotg}(x) = 5e^{\cos(x)}$. L'équation homogène associée est $y'_H + y_H \operatorname{cotg}(x) = 0$, dont la solution est donnée par

$$\ln(y_H) = -\ln\big(C\sin(x)\big), \tag{97.336}$$

et donc par

$$y_H(t) = \frac{C}{\sin(t)} \tag{97.337}$$

où nous avons renommé $C \to C^{-1}$. La méthode de variations des constantes demande de substituer ceci dans l'équation :

$$\begin{aligned} y(t) &= \frac{K(t)}{\sin(t)} \\ y'(t) &= \frac{K'}{\sin(t)} - K\frac{\cos(t)}{\sin^2(t)}. \end{aligned} \tag{97.338}$$

Nous trouvons

$$\frac{K'}{\sin(t)} - \underbrace{K\frac{\cos(t)}{\sin^2(t)} + \frac{K}{\sin(t)}\operatorname{cotg}(t)}_{=0} = 5e^{\cos(t)} \tag{97.339}$$

où, comme toujours, les termes en K se simplifient, et il reste $K'(t) = 5\sin(t)e^{\cos(t)}$, d'où $K(t)$ se trouve. La solution finale est alors

$$y(t) = \frac{-5e^{\cos(t)+C}}{\sin(t)}. \tag{97.340}$$

La condition de Cauchy demande $y(\frac{\pi}{2}) = \frac{-5+C}{1} = -4$, et donc $C = -1$. La solution est donc

$$y(t) = \frac{-5e^{\cos(t)} - 1}{\sin(t)}. \tag{97.341}$$

(4)

(5) $x^3y' + (2 - 3x^2)y = x^3$. Nous allons commencer par diviser l'équation par x^3. Nous devrons discuter à la fin ce qu'il se passe en $x = 0$. L'équation homogène est $y'_H/y_H = (3x^2 - 2)/x^3$, et la solution est

$$y_H(t) = Kx^3e^{-x^2}. \tag{97.342}$$

Nous appliquons la méthode de variation des constantes, et nous mettons

$$\begin{aligned} y(x) &= K(x)x^3e^{1/x^2} \\ y'(x) &= K'x^3e^{1/x^2} + K(3x^2 - 2)e^{1/x^2} \end{aligned} \tag{97.343}$$

dans l'équation de départ. Après simplification nous trouvons

$$\begin{aligned} K'(x) &= x^{-3}e^{-1/x^2} \\ K(x) &= \frac{1}{2}e^{-1/x^2} + C \end{aligned} \tag{97.344}$$

La solution à l'équation non homogène est donc

$$y(x) = \frac{x^3}{2} + x^3 e^{1/x^3}. \tag{97.345}$$

Cette solution n'est pas définie en $x = 0$, et, pire, $\lim_{x\to 0} y(x) \neq 0$. Donc la seule solution possible qui passe par $x = 0$ est la solution identiquement nulle.

(6) L'équation homogène est $y'_H/y_H = 2\operatorname{cotg}(2x)$, dont la solution est

$$y_H(x) = K\sin(2x). \tag{97.346}$$

Exercise 119 exoEqsDiff0005

Soit $p(t)$ le nombre d'individus d'une population à l'instant t. Un modèle de population classique et très simple est celui régi par $p' = Kp$ où $K \in \mathbb{R}$ est le taux de croissance de la population (taux de natalité moins le taux de mortalité). La résolution de cette équation différentielle montre que la croissance de la population est exponentielle, ce qui est généralement satisfaisant tant qu'il n'y a pas de problèmes de surpopulation (territoire et ressource illimités). Pour tenir compte de tels problèmes, on suppose plutôt que p est régie par l'équation différentielle $p' = Kp(M - p)$ où $M \in \mathbb{R}_0^+$ est le seuil de surpopulation.

(1) Sachant que $p(0) = p_0$ déterminez $p(t)$.
(2) Déterminez le comportement de $p(t)$ lorsque t tend vers $+\infty$.
(3) Dessiner le champ de pentes correspondant à cette équation et en déduire l'allure des solutions sans utiliser la résolution de l'équation différentielle. Remarquez que $p = M$ est une solution stable, alors que $p = 0$ est instable.

97.11.2 Équations différentielles du second ordre

Exercise 120 exoEqsDiff0006

Trouvez toutes les solutions des équations différentielles suivantes et des problèmes de Cauchy lorsqu'ils sont spécifiés.

(1) $y'' - 2y' = 0$
$y(0) = 0, y'(0) = 1$
(2) $3y'' + 4y' + y = 0$
(3) $y'' - 6y' + 9y = 0$
$y(1) = 0, y'(1) = 1$
(4) $y'' - 4y' + 4y = 0$
(5) $y'' + y = 0$
$y(3) = 0, y'(3) = 1$
(6) $y'' - 2y' + 3y = 0$
(7) $y^{4/} - 2y''' - 3y'' + 8y' - 4y = 0$

corrEqsDiff0006

Correction of the exercise 120

(1) Le polynôme caractéristique est $x^2 - 2x = 0$, donc $x = 0$ et $x = 2$. La solution générale est donc

$$y(t) = A + Be^{2t}. \tag{97.347}$$

Les dérivées se calculent aisément :

$$\begin{aligned} y'(t) &= 2Be^{2t} \\ y''(t) &= 4Be^{2t}. \end{aligned} \tag{97.348}$$

Les données de Cauchy demandent de résoudre le système

$$\begin{cases} y(0) = A + B = 0 & \text{(97.349a)} \\ y'(0) = 2B = 1, & \text{(97.349b)} \end{cases}$$

donc $A = -\frac{1}{2}$ et $B = \frac{1}{2}$.

(2) Polynôme caractéristique : $3x^2 + 4x + 1 = 0$, solutions : $x = -\frac{1}{3}$ et $x = -1$, donc l'équation différentielle est résolue par
$$y(t) = Ae^{-x/3} + Be^{-x}. \tag{97.350}$$

(3) Cette fois, le polynôme caractéristique a une racine double $x = 3$. Le système fondamental de solutions est alors donné par
$$\begin{aligned} y_1(t) &= e^{-3t} \\ y_2(t) &= te^{-3t}, \end{aligned} \tag{97.351}$$
par la proposition 2 de la page 331. La donnée de Cauchy est résolue avec
$$y(t) = Ay_1(t) + By_2(t) \tag{97.352}$$
et $A = \frac{e^3}{6}$ et $B = -\frac{e^3}{6}$.

(4) Le polynôme caractéristique est $x^2 - 4x + 4 = 0$, qui accepte une racine double en $x = 2$, donc la solution générale de l'équation est
$$y(t) = Ae^{2t} + Bte^{2t}. \tag{97.353}$$

(5) Le polynôme caractéristique est $x^2 + x = 0$, et ses solutions sont $x = 0$ et $x = -1$. La solution générale est donc
$$y(t) = A + Be^{-t}. \tag{97.354}$$
La donnée de Cauchy se résous par le système
$$\begin{cases} y(3) = A + Be^{-3} = 0 & \text{(97.355a)} \\ y'(3) = -Be^{-3} = 1, & \text{(97.355b)} \end{cases}$$
qui a pour solution $A = 1$ et $B = -e^3$.

(6) Le polynôme caractéristique est donné par $x^2 - 2x + 3 = 0$, dont les deux solutions complexes conjuguées sont
$$\begin{aligned} x_1 &= 1 + \sqrt{2}i \\ x_2 &= 1 - \sqrt{2}i. \end{aligned} \tag{97.356}$$
Un ensemble fondamental de solutions complexes est donné par
$$\begin{aligned} y_1(t) &= e^{(1+\sqrt{2}i)t} \\ y_2(t) &= e^{(1-\sqrt{2}i)t} \end{aligned} \tag{97.357}$$
Les combinaisons linéaires $\frac{1}{2}(y_1 + y_2)$ et $\frac{1}{2i}(y_1 - y_2)$ donnent les solutions réelles :
$$y(t) = Ae^t \cos(\sqrt{2}t) + Be^t \sin(\sqrt{t}). \tag{97.358}$$
En effet, $e^{(a+bi)t} = e^{at}e^{it} = e^{at}\big(\cos(bt) + i\sin(bt)\big)$.

(7) Le polynôme caractéristique se factorise en
$$(x - 2)(x - 1)^2(x + 2), \tag{97.359}$$
donc la solution générale est
$$y(t) = Ae^{2t} + Be^{-t} + Cte^{-t} + De^{-2t}. \tag{97.360}$$

Exercise 121 exoEqsDiff0007

Trouvez toutes les solutions des équations différentielles suivantes et des problèmes de Cauchy lorsqu'ils sont spécifiés. (Pour chacun des membres de droite des équations suivantes il existe un sous-espace de fonctions de dimension finie stable par dérivation.)

(1) $y'' - 2y' = e^t \sin(t)$
$y(0) = 0, y'(0) = 1$

(2) $y'' + y = -2\sin(t) + 4t\cos(t)$

(3) $y'' - y = e^t + 2$
$y(3) = 0, y'(3) = 1$

(4) $y'' - 2y' + 3y = t^3 + \sin(t)$

(5) $y'' - 4y' + 4y = t^3 e^{2t} + te^{2t}$
$y(0) = 0, y'(0) = 1$

(6) $y''' - y'' - 4y' + 4y = 2t^2 - 4t - 1 + 2t^2 e^{2t} + 5te^{2t} + e^{2t}$

corrEqsDiff0007

Correction of the exercise 121

(1) $y'' - 2y' = e^t \sin(t)$. Un système fondamental pour l'équation homogène est donné par

$$\begin{cases} y_1 = 1 \\ y_2 = e^{2t}. \end{cases} \tag{97.361}$$

Étant donné que le second membre contient e^t, il est naturel de mettre e^t dans un essai de solution particulière. Ensuite, la présence de $\sin(t)$ nous incite à mettre $A\cos(t) + B\sin(t)$. Donc nous essayons

$$y_P(t) = e^t\big(A\cos(t) + B\sin(t)\big). \tag{97.362}$$

Un petit calcul de dérivation avec cette fonction montre que

$$\begin{aligned} y_P'' - 2y_P' - e^t\sin(t) \\ = e^t\big(\sin(t)(-2B - 2(A-B) - 1) + \cos(t)(2A - 2(B+A))\big), \end{aligned} \tag{97.363}$$

que nous devons égaler à zéro pour trouver A et B. Le système à résoudre est

$$\begin{cases} -2B - 2(A-B) - 1 = 0 \\ 2A - 2(B+A) = 0, \end{cases} \tag{97.364}$$

et la solution est $A = -\frac{1}{2}$, $B = 0$. Nous avons donc pour solution particulière

$$y_P(t) = -\frac{1}{2}e^t\sin(t). \tag{97.365}$$

(2) $y'' + y = -2\sin(t) + 4t\cos(t)$. L'équation caractéristique est $x^2 + 1 = 0$, et donc $x = \pm i$. Un système fondamental de solutions réelles est donné par

$$\begin{cases} y_1 = \cos(t) \\ y_2 = \sin(t). \end{cases} \tag{97.366}$$

Nous notons P l'opérateur $f \mapsto P(x)$ défini par

$$P(f)(x) = f''(x) + f(x) + 2\sin(x) - 4x\cos(x). \tag{97.367}$$

La fonction $y_P(x)$ est solution particulière de l'équation si $P(y_P) = 0$. Le second membre contenant du sinus et du cosinus, nous allons essayer une solution qui sera une combinaison de sinus et cosinus. Au niveau des coefficients, nous allons mettre des polynômes. Nous essayons

$$y_P(x) = (ax^3 + bx^2 + cx + d)\cos(x) + (a'x^3 + b'x^2 + c'x + d)\sin(x). \tag{97.368}$$

Les termes d et d' correspondent à $\cos(x)$ et $\sin(x)$, qui sont solutions de l'homogène. Ils ne vont donc pas contribuer à obtenir le second membre ; nous pouvons donc directement les oublier et essayer

$$y_P(x) = (ax^3 + bx^2 + cx)\cos(x) + (a'x^3 + b'x^2 + c'x)\sin(x). \tag{97.369}$$

Le calcul montre que

$$\begin{aligned} P\big(y_P(x)\big) = \big(-6ax^2 + (6a' - 4b)x - 2c + 2b' + 2\big)\sin(x) \\ + \big(6a'x^2 + (4b' + 6a - 4)x + 2c' + 2b\big)\cos(x). \end{aligned} \tag{97.370}$$

Afin que cela soit zéro, il faut $a = a' = 0$, $b = 0$, $b' = 1$, $c' = 0$ et $b' = 1$. Une solution particulière de l'équation non homogène est donc donnée par

$$y_P(x) = 2x\cos(x) + x^2\sin(x). \tag{97.371}$$

(3) $y'' - y = e^t + 2$. Un système fondamental pour l'homogène est

$$\begin{cases} y_1 = e^x \\ y_2 = e^{-x}. \end{cases} \tag{97.372}$$

Étant donné que e^x est déjà une solution de l'homogène, nous ne mettons pas e^x dans un essai de solution particulière, mais nous essayons plutôt axe^x, donc nous essayons

$$y_P(x) = axe^x - 2. \tag{97.373}$$

Le -2 est là pour obtenir le 2 du membre de droite. En remettant dans l'équation, nous trouvons $(2a-1)e^x$, qui s'annule pour $a = 1/2$, donc

$$y_P(x) = \frac{x}{2}e^x - 2, \tag{97.374}$$

et la solution générale de l'équation recherchée est

$$y(x) = Ae^x + Be^x + \frac{x}{2}e^x - 2. \tag{97.375}$$

En ce qui concerne le problème de Cauchy, nous devons calculer la dérivée :

$$y'(3) = -Be^{-3} + Ae^3 + \frac{3}{2}e^3 + \frac{1}{2}e^3, \tag{97.376}$$

et nous voyons que les conditions imposent

$$A = -\frac{e^{-3}(7e^3 - 6)}{4}, \quad B = \frac{e^6 + 2e^3}{4}. \tag{97.377}$$

(4) $y'' - 2y' + 3y = t^3 + \sin(x)$. Un système fondamental de solutions réelles est

$$\begin{cases} y_1(x) = e^x \cos(\sqrt{2}x) \\ y_2(x) = e^x \sin(\sqrt{2}x) \end{cases} \tag{97.378}$$

Nous allons chercher une solution particulière en deux parties. Une pour obtenir x^3 et une pour obtenir $\sin(x)$. Afin d'obtenir x^3, nous essayons un polynôme :

$$\begin{aligned} y_{P1} &= ax^3 + bx^2 + cx + d \\ y'_{P1} &= 3ax^2 + 2bx + c \\ y''_{P1} &= 6ax + 2b \end{aligned} \tag{97.379}$$

Afin que $y''_{P1} - 2y'_{P1} + 3y_{P1} = x^3$, nous devons choisir

$$a = \frac{1}{3}, \quad b = \frac{2}{3}, \quad c = \frac{2}{9}, \quad d = -\frac{8}{27}. \tag{97.380}$$

Nous cherchons le second morceau de telle façon à avoir $y''_{P2} - 2y'_{P2} + 3y_{P2} = \sin(x)$. Pour cela, nous regardons $y_{P2}(x) = A\cos(x) + B\sin(x)$ et il est vite vu que $A = B = \frac{1}{4}$. La solution particulière ainsi construite est

$$y_P(x) = \frac{x^3}{3} + \frac{2x^2}{3} + \frac{2x}{9} - \frac{8}{27} + \frac{1}{4}\cos(x) + \frac{1}{4}\sin(x). \tag{97.381}$$

Exercise 122 exoEqsDiff0008

Trouvez toutes les solutions des équations différentielles suivantes et des problèmes de Cauchy lorsqu'ils sont spécifiés. (A l'inverse de l'exercice précédent, ici, les membres de droite des équations ne sont pas inclus dans des sous-espaces de fonctions de dimension finie stables par dérivation.)

(1) $\begin{cases} y''-6y'+9y=\frac{e^{3x}}{x^2} \\ y(0)=0, y'(0)=1 \end{cases}$
(2) $y''-y=e^{-t}\sin(e^{-t})+\cos(e^{-t})$
(3) $y''+y=\csc(t)$
(4) $\begin{cases} y'''+y=e^{x/2} \\ y(2)=0, y'(2)=4, y''(2)=8 \end{cases}$

corrEqsDiff0008

Correction of the exercise 122

(1) $y''-6y'+9y=\frac{e^{3x}}{x^2}$. Le polynôme caractéristique de l'homogène a une racine double $x=3$, donc un système fondamental est donné par

$$\begin{aligned} y_1(x)&=e^{3x}\\ y_2(x)&=xe^{3x}. \end{aligned} \tag{97.382}$$

Le Wronskien est

$$W=\det\begin{pmatrix} e^{3x}x & xe^{3x} \\ 3e^{3x} & e^{3x}+3xe^{3x}\end{pmatrix}=e^{6x}, \tag{97.383}$$

et les deux autres qu'il faut calculer sont

$$\begin{aligned} W_1&=\det\begin{pmatrix} 0 & xe^{3x} \\ \frac{e^{3x}}{x^2} & e^{3x}(1+3x)\end{pmatrix}=-\frac{e^{6x}}{x},\\ W_2&=\det\begin{pmatrix} e^{3x} & 0 \\ 3e^{3x} & \frac{e^{3x}}{x^2}\end{pmatrix}=\frac{e^{6x}}{x^2}. \end{aligned} \tag{97.384}$$

En utilisant la méthode usuelle, nous trouvons

$$\begin{aligned} c_1'&=\frac{W_1}{W}=-\frac{1}{x}\\ c_2'&=\frac{W_2}{W}=\frac{1}{x^2}, \end{aligned} \tag{97.385}$$

ce qui donne $c_1(x)=-\ln(x)$ et $c_2(x)=-1/x$. Une solution particulière est donc donnée par

$$y_P(x)=-\ln(x)e^{3x}-e^{3x}. \tag{97.386}$$

Notons que le deuxième morceau est même une solution de l'équation homogène, nous pouvons donc l'oublier dans la solution particulière. En fin de compte, la solution générale est

$$y(x)=Ae^{3x}+Bxe^{3x}-\ln(x)e^{3x}. \tag{97.387}$$

Il n'y a pas de solutions au problème de Cauchy parce que $y(0)$ n'est pas définie. Remarquez que le $v(x)$ est e^{3x}/x^2, donc 0 est en dehors du domaine de continuité de $v(x)$, et donc les théorèmes ne s'appliquent pas. Il n'est donc pas choquant qu'il n'y ait pas de solutions.

(2) $y''-t=e^{-t}\sin(e^{-t})+\cos(e^{-t})$. L'équation homogène a pour système fondamental $y_1(x)=e^x$ et $y_2(x)=e^{-x}$. Les déterminants à calculer sont

$$\begin{aligned} W&=\det\begin{pmatrix} e^x & e^{-x} \\ e^x & -e^{-x}\end{pmatrix}=-1-1=-2,\\ W_1&=\det\begin{pmatrix} 0 & e^{-x} \\ e^{-x}\sin(e^{-x})+\cos(e^{-x}) & -e^{-x}\end{pmatrix}=-e^{-2x}\sin(e^{-x})+e^{-x}\cos(e^{-x}),\\ W_2&=\det\begin{pmatrix} e^x & 0 \\ e^x & e^{-x}\sin(e^{-x})+\cos(e^{-x})\end{pmatrix}=\sin(e^{-x})+e^x\cos(e^{-x}). \end{aligned} \tag{97.388}$$

De là, nous déduisons

$$\begin{aligned} c_1'&=\frac{e^{-2x}}{2}\big(\sin(e^{-x})+e^x\cos(e^{-x})\big)\\ c_2'&=-\frac{1}{2}\big(\sin(e^{-x})+e^x\cos(e^{-x})\big). \end{aligned} \tag{97.389}$$

La seconde s'intègre assez facilement et donne $c_2(x)=e^x\cos(e^{-x})/2$, tandis que pour trouver c_1, il faut changer de variable et poser $u=e^{-x}$ pour trouver

$$c_1=-\frac{1}{2}\int\big(u\sin(u)+\cos(u)\big)=-\frac{1}{2}\int\big(u\sin(u)\big)'=-\frac{1}{2}e^{-x}\sin(e^{-x}). \tag{97.390}$$

(3) $y'' + y = 1/\sin(x)$. Le polynôme caractéristique admet les deux solutions complexes $\pm i$, et donc nous avons le système fondamental réel

$$\begin{aligned} y_1 &= \cos(x) \\ y_2 &= \sin(x). \end{aligned} \tag{97.391}$$

Les calculs sont assez simples :

$$\begin{aligned} W &= \det \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} = 1 \\ W_1 &= \det \begin{pmatrix} 0 & \sin(x) \\ \frac{1}{\sin(x)} & \cos(x) \end{pmatrix} = -1 \\ W_2 &= \det \begin{pmatrix} \cos(x) & 0 \\ -\sin(x) & \frac{1}{\sin(x)} \end{pmatrix} = \tan(x). \end{aligned} \tag{97.392}$$

Après une séance d'intégration, nous trouvons la solution de l'équation différentielle sous la forme

$$y(x) = -x\cos(x) + \ln\big(\sin(x)\big)\sin(x) + A\cos(x) + B\sin(x). \tag{97.393}$$

(4) $y''' + y = e^{x/2}$. Le polynôme caractéristique admet deux solutions complexes conjuguées :

$$x = \frac{1 \pm \sqrt{2}i}{2}, \tag{97.394}$$

en plus de la solution réelle $x = -1$. Un système fondamental de solutions réelle est donné par

$$\begin{aligned} y_1 &= e^{-x} \\ y_2 &= e^{x/2}\cos(\omega x) \\ y_2 &= e^{x/2}\sin(\omega x) \end{aligned} \tag{97.395}$$

où nous avons posé $\omega = \sqrt{3}/2$. Le Wronskien est

$$\begin{aligned} W = & \\ & \begin{vmatrix} e^{-x} & e^{x/2}\cos(\omega x) & e^{x/2}\sin(\omega x) \\ -e^{-x} & -\frac{1}{2}\big[\sqrt{3}e^{x/2}\sin(\omega x) - e^{x/2}\cos(\omega x)\big] & -\frac{1}{2}\big[\sqrt{3}e^{x/2}\sin(\omega x) + e^{x/2}\cos(\omega x)\big] \\ e^{-x} & -\frac{1}{2}\big[\sqrt{3}e^{x/2}\sin(\omega x) + e^{x/2}\cos(\omega x)\big] & -\frac{1}{2}\big[\sqrt{3}e^{x/2}\sin(\omega x) - e^{x/2}\cos(\omega x)\big] \end{vmatrix} \\ & = 3\omega. \end{aligned} \tag{97.396}$$

Les autres calculs ne sont pas plus faciles et donnent

$$\begin{aligned} W_1 &= \omega e^{3x/2} \\ W_2 &= -\frac{1}{2}\big(3\sin(\omega x) + \sqrt{3}\cos(\omega x)\big) \\ W_3 &= -\frac{1}{2}\big(3\sin(\omega x) + \sqrt{3}\cos(\omega x)\big). \end{aligned} \tag{97.397}$$

Nous trouvons ensuite facilement que

$$c_1(x) = \frac{2}{9}e^{3x/2}. \tag{97.398}$$

Nous trouvons aussi

$$\begin{aligned} c_2'(x) &= -\frac{3\sin(\omega x) + \sqrt{3}\cos(\omega x)}{3\sqrt{3}} \\ c_3'(x) &= -\frac{\sqrt{3}\sin(\omega x) - 3\cos(\omega x)}{3\sqrt{3}} \\ c_2(x) &= -\frac{1}{3\sqrt{3}}\big(2\sin(\omega x) - 2\sqrt{3}\cos(\omega x)\big) \\ c_3(x) &= -\frac{1}{3\sqrt{3}}\big(2\sqrt{3}\sin(\omega x) + 2\cos(\omega x)\big). \end{aligned} \tag{97.399}$$

Exercise 123 exoEqsDiff0009

La solution générale d'une équation différentielle du second ordre est typiquement une famille à deux paramètres de fonctions. Le problème de « Cauchy » consiste à chercher dans cette famille une fonction dont la valeurs en un point et la valeur de la dérivée en ce même point sont spécifiées. Le problème des « valeur aux limites » consiste à chercher dans cette même famille une fonction dont les valeurs en deux points distincts sont spécifiées. Étudiez l'existence et l'unicité des solutions de ce problème dans le cadre des équations différentielles linéaires à coefficients constants. En particulier donnez un exemple de tel problème qui ne possède pas de solutions.

97.11.3 Équations différentielles : modélisation

Exercise 124 exoEqsDiff0010

Lorsqu'on étire un ressort, la force de rappel est proportionnelle à l'allongement du ressort. Nous allons décrire le mouvement d'une masse m fixée à l'extrémité d'un ressort qu'on étire d'une longueur y_0 et qu'on lâche. On suppose qu'il n'y a aucun effet d'amortissement. On rappelle que l'on a le théorème fondamental de la dynamique $F = m\gamma$, liant l'accélération γ et la somme des forces F appliquées à la masse m.

(1) Les forces en présence sont : le poids (mg) et la force de rappel du ressort $-ky$. On suppose que la position de l'extrémité du ressort est y relativement à la position d'équilibre du ressort avec la masse en son extrémité. Établir que l'on a l'équation différentielle $my'' = -ky$, et la résoudre pour obtenir la position $y(t)$ de la masse à l'instant t. Quel type de mouvement obtient on ?

(2) On suppose maintenant que la masse est plongée dans un liquide (amortisseur) et que la force de friction est proportionnelle à la vitesse et vaut $-cy'$. Écrire l'équation régissant le mouvement de la masse, puis résoudre en distinguant 3 cas suivant la valeur de $c^2 - 4km$. Quel type de mouvement obtient on dans chacun des cas ?

Exercise 125 exoEqsDiff0011

Régulateur de Foucault.

Un point P de masse m est accroché à un fil sans masse enroulé autour d'un cylindre homogène de rayon R, de masse M et d'axe horizontal fixe. La chute du point P entraîne la rotation du cylindre. Ce cylindre, muni d'ailettes est soumis à la résistance de l'air que l'on représentera par un couple de frottement $-f\theta'$, où $\omega = \theta'$ représente la vitesse de rotation du cylindre. Le système étant abandonné sans vitesse initiale, on veut déterminer la fonction ω. On note z la longueur parcourue par le point P.

La mise en équation fait apparaitre les relations suivantes :

(1) - en appliquant le théorème du moment cinétique

$$J\theta'' = -f\theta' + RT,$$

où $J = MR^2/2$ est le moment d'inertie du cylindre et T est une force.

(2) - en appliquant le principe fondamental de la dynamique pour le point P

$$mz'' = mg - T$$

(1) Quelle est la longueur Δz parcourue par M lorsque le cylindre tourne d'un angle $\Delta\theta$? En déduire une relation entre z'' et θ''.

(2) Trouver une équation différentielle vérifiée par ω.

(3) Résoudre cette équation. Que peut-on conclure pour le mouvement de la poulie lorsque t devient grand ?

Exercise 126 exoEqsDiff0012

On cherche à modéliser l'évolution dans le temps de la quantité d'une substance (par exemple la pénicilline) contenue dans le sang lors d'une injection en continu. La substance est injectée à raison de I grammes par minute. Par ailleurs la substance est éliminée du sang à une vitesse proportionnelle à la quantité de substance présente au temps t.

(1) Justifier que la quantité de substance y présente dans le sang au temps t satisfait l'équation différentielle
$$y' = -ky + I,$$
pour une certaine constante $k > 0$.

(2) Chercher une expression de y en fonction de t, sachant qu'au temps $t = 0$ (début de l'injection), on a $y = 0$. Calculer la limite de y quand t tend vers l'infini. Interprétation.

Exercise 127 exoEqsDiff0013

L'équation différentielle $m\ddot{y} = mg - k\dot{y}$ décrit la chute d'un corps soumit à la gravitation si la friction est proportionnelle à la vitesse.

(1) Calculer la solution avec $y(0) = 0, \dot{y}(0) = 0$.

(2) Calculer la « vitesse finale » $v_\infty = \lim_{t\to\infty} \dot{y}(t)$.

Exercise 128 exoEqsDiff0014

Calculer un système fondamental réel pour

(1) $y^{(4)} - y = 0$

(2) $y^{(4)} + 4y'' + 4y = 0$

(3) $y^{(4)} - 2y^{(3)} + 5y'' = 0$

Exercise 129 exoEqsDiff0015

Déterminer une solution particulière de l'équation $y'' + y = q$ pour

(1) $q = x^3$

(2) $q = \sinh x$

(3) $q = 1/\sin x$

Exercise 130 exoEqsDiff0016

Regardons l'ensemble des solutions de l'équation différentielle $P(\mathrm{D})y = 0$. Montrer que cette équation différentielle est **asymptotiquement stable**, c'est-à-dire que pour toute solution y, on a $\lim_{t\to\infty} y(t) = 0$ si et seulement si pour toute racine z du polynôme caractéristique on a $\mathrm{Re}\, z < 0$.

97.12 Intégrales multiples

Exercise 131 exoIntMult0001

Intégrez les fonctions suivantes sur les domaines spécifiés

(1) $4 - x^2 - y^2$ sur $\{(x, y) \in \mathbb{R}^2 \mid 0 < x < 1, 0 < y < 2\}$

(2) $x^2 + y^2$ sur $\{(x, y) \in \mathbb{R}^2 \mid 0 < x < y < 2\}$

(3) e^{x+y} sur $\{(x, y) \in \mathbb{R}^2 \mid 1 < sup\{|x|, |y|\} < 2\}$

(4) xyz sur $\{(x, y, z) \in \mathbb{R}^3 \mid 0 < x, 0 < y, 0 < z, x + y + z < 1\}$

(5) $\sqrt{1 - x^2 - y^2}$ sur la région du plan délimitée par les courbes $y = x$ et $x^2 + y^2 = y$. corrIntMult0001

Correction of the exercise 131

(1) Comme expliqué dans les rappels (page 1900), nous utilisons Fubini :
$$\begin{aligned}\int_{[0,2]\times[0,1]} f(x, y) &= \int_0^2 \int_0^1 (4 - x^2 - y^2)\, dxdy \\ &= \int_0^2 \left[4x - \frac{x^3}{3} - y^2 x\right]_0^1 dy \\ &= \int_0^2 \left(4 - \frac{1}{3} - y^2\right) dy \\ &= \frac{14}{3}.\end{aligned} \tag{97.400}$$

(2) Cet exercice a été fait à la page 1901. L'intégrale à calculer est

$$\int_0^2 \left(\int_0^y (x^2+y^2)dx \right) dy = \frac{16}{3}. \tag{97.401}$$

(3) Un dessin du domaine d'intégration est donné à la figure 97.6. Une stratégie consiste à intégrer la fonction sur tout le grand carré, et ensuite soustraire l'intégrale sur le petit. Pour cela, nous utilisons la sous-additivité de l'intégrale.

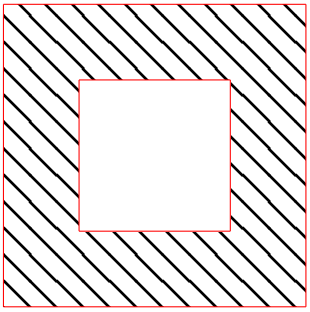


FIGURE 97.6: Un domaine d'intégration. LabelFigIntDeuxCarres

L'intégrale sur le grand carré vaut

$$\int_{-2}^2 \left(\int_{-2}^2 e^{x+y}dx \right) dy = \int_{-2}^2 e^y [e^x]_{-2}^2 dy = (e^2-e^{-2})^2, \tag{97.402}$$

tandis que l'intégrale sur le petit carré central vaut

$$\int_{-1}^1 \left(\int_{-1}^1 e^x e^y dx \right) dy = \left([e^x]_{-1}^1 \right)^2 = \left(e-e^{-1} \right)^2. \tag{97.403}$$

Au final,

$$\int_E f = (e^2-e^{-2})^2 - (e-e^{-1})^2. \tag{97.404}$$

(4) Nous commençons par intégrer verticalement. Le domaine va de $z=0$ à $z=1$. Pour chacun de ces z, la variable y peut varier de 0 à $1-z$, et pour chacun de ces (z,y) fixés, la variable x peut varier de 0 à $1-z-y$, donc l'intégrale à calculer est

$$\begin{aligned}
I &= \int_0^1 dz \Big(\int_0^{1-z} dy \int_0^{1-z-y} dx\,(xyz) \Big) \\
&= \int_0^1 dz \int_0^{1-z} \left[\frac{x^2}{2} yz \right]_0^{1-z-y} \\
&= \frac{1}{2} \int_0^1 dz \int_0^{1-z} yz(z^2+2yz-2z+y^2-2y+1)dy \\
&= \frac{1}{2} \int_0^1 dz \int_0^{1-z} [y(z^3-2z^2+z)+y^2(2z^2-2z)+y^3 z]dy \\
&= \frac{1}{2} \int_0^1 z(1-z)^4 \left(\frac{1}{2}+\frac{2}{3}+\frac{1}{4} \right) dz \\
&= \frac{17}{24} \int_0^1 z(1-z)^4 dz.
\end{aligned} \tag{97.405}$$

EqLongCalzxyzInt

Cette dernière intégrale s'effectue en changeant de variable en posant $u=1-z$. Nous tombons[10] sur

$$I = \frac{17}{24} \int_0^1 (u^4-u^5)du = \frac{1}{720}. \tag{97.406}$$

10. Je crois qu'il y a une faute de coefficient dans le calcul (97.405) ; en tout cas la réponse 1/720 est correcte, parce que c'est ce que dit maxima.

```
integrate( integrate(integrate(x*y*z,x,0,1-z-y),y,0,1-z) ,z,0,1 ).
```

(5) Cet exercice est fait comme exemple à la page 1778, l'intégrale à calculer est donnée à l'équation (18.853). Le résultat est

$$I = \frac{3\pi - 5\sqrt{2}}{36}. \tag{97.407}$$

Exercise 132 exoIntMult0002

Intégrez les fonctions suivantes sur les domaines spécifiés. Un changement de variables peut grandement vous simplifier la vie.

(1) $y - x$ $\quad \{(x, y) \in \mathbb{R}^2 \mid -3 < y - x < 1, 1 < y + x < 5\}$
(2) $e^{a(x^2+y^2)}$ $\quad \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < R^2\}$
(3) z^2 $\quad \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 < 1\}$
(4) $\sqrt{x^2 + y^2 + z^2}$ $\quad \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 < 1\}$
(5) z $\quad \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 < z^2, 0 < z < 1\}$

où a est un réel et R un réel positif. corrIntMult0002

Correction of the exercise 132

(1) Nous utilisons le changement de variable

$$\begin{cases} x = (u + v)/2 \\ y = (u - v)/2 \end{cases} \tag{97.408}$$

dont le jacobien est donné par

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} = -\frac{1}{2}. \tag{97.409}$$

Nous devons donc calculer

$$\int_E f(x, y) = \int_E (y - x) = \int_{g^{-1}(E)} -v \cdot \frac{1}{2}. \tag{97.410}$$

Ne pas oublier que le jacobien doit être pris en valeur absolue. Le domaine d'intégration est donné par $-1 < v < 3$ et $1 < u < 5$. Donc l'intégrale à calculer est

$$I = -\int_{-1}^{3} dv \int_{1}^{5} du \frac{v}{2} = -8. \tag{97.411}$$

(2) En passant aux coordonnées polaires, nous devons simplement calculer l'intégrale

$$\int_0^{2\pi} d\theta \int_0^R e^{aR} \cdot r\, dr = 2\pi \left[\frac{e^{ar^2}}{2a} \right]_0^R = \frac{\pi}{a}(e^{aR^2} - 1). \tag{97.412}$$

(3) Nous passons en coordonnées cylindriques, dans lesquelles le domaine d'intégration s'écrit $r^2 + z^2 < 1$, donc nous devons calculer

$$\int_0^{2\pi} d\theta \int_0^1 dr \int_{-\sqrt{1-r^2}}^{\sqrt{1-r^2}} z^2 r\, dr = \frac{4\pi}{15}. \tag{97.413}$$

Il est également possible de faire cet exercice en coordonnées sphériques.

(4) La fonction $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ et le domaine $x^2 + y^2 + z^2 < 1$ se prêtent parfaitement bien aux coordonnées sphériques. Il faut donc intégrer

$$I = \int_0^{2\pi} d\theta \int_0^{\pi} d\varphi \int_0^1 r \cdot r^2 \sin\varphi\, dr = \pi. \tag{97.414}$$

(5) Nous calculons en coordonnées cylindriques. Le domaine est $r^2 < z^2$ et $0 < z < 1$, donc

$$I = \int_0^1 dz \int_0^{2\pi} d\theta \int_0^z zr\, dr = \frac{\pi}{4}. \tag{97.415}$$

Calculer le volume ou la surface d'un domaine revient à intégrer la fonction constante 1 sur le domaine. Si nous effectuons un changement de variables, le jacobien intervient toutefois.

Exercise 133 exoIntMult0003

Calculez l'aire de la région du plan délimitée par les droites $\theta = 0$, $\theta = \frac{\pi}{4}$ et l'arc de circonférence $\rho = \cos\theta$ intercepté par ces deux droites. corrIntMult0003

Correction of the exercise 133

Nous intégrons en polaires. L'angle θ va de 0 à $\pi/4$ ($x = y$), et pour chacun de ces angles, le rayon va de 0 à $\cos(\theta)$. L'intégrale est donc

$$I = \int_0^{\pi/4} d\theta \int_0^{\cos(\theta)} r\, dr = \frac{\pi + 2}{16} \tag{97.416}$$

où nous avons utilisé la primitive $\int \cos(\theta)^2 d\theta = \frac{\sin(2\theta)+2\theta}{4}$.

Exercise 134 exoIntMult0004

Calculez l'aire délimitée par la courbe $\rho^2 = a^2 \cos 2\theta$. corrIntMult0004

Correction of the exercise 134

$\rho^2 = a^2 \cos(2\theta)$. La surface délimitée par cette courbe contient deux parties : une à gauche de l'axe vertical et une à droite. Nous allons calculer l'aire de la partie à droite. Afin d'obtenir l'aire totale, il faudra multiplier par deux.

Pour chaque θ, le rayon du domaine va de 0 à $a\sqrt{\cos(2\theta)}$. Étant donné que r est toujours positif, les angles ne vont que de $-\pi/4$ à $\pi/4$, donc

$$I = \int_{-\pi/4}^{\pi/4} d\theta \int_0^{a\sqrt{\cos(2\theta)}} r\, dr = \frac{a^2}{2}. \tag{97.417}$$

L'aire totale vaut donc a^2.

Exercise 135 exoIntMult0005

Calculez les volumes des domaines suivants

(1) Ellipsoïde. $\{(x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\}$

(2) Tranche d'ellipsoïde. $\{(x, y, z) \in \mathbb{R}^3 \mid \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{4} < 1, 0 < z < 1\}$

(3) Intersection de deux cylindres. $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 < 1, y^2 + z^2 < 1\}$ corrIntMult0005

Correction of the exercise 135

(1) Le domaine

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \tag{97.418}$$

se prête bien au changement de variables

$$\begin{cases} x = au \\ y = bv \\ z = cs \end{cases}, \tag{97.419}$$

dont le jacobien est

$$J = \det \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} = abc. \tag{97.420}$$

Nous devons donc intégrer

$$\int_{u^2+v^2+s^2<1} abc\, du\, dv\, ds = \frac{4\pi abc}{3}. \tag{97.421}$$

(2) Nous refaisons le même changement de variable :

$$\begin{cases} x = 3u \\ y = 3v \\ z = 2s \end{cases}, \tag{97.422}$$

dont le jacobien vaut 12. Le domaine devient $u^2 + v^2 + s^2 < 1$, limité par $0 < s < \frac{1}{2}$. Le volume cherché est donc 12 fois le volume du morceau de sphère contenu entre les hauteurs 0 et $\frac{1}{2}$.
Il faut donc intégrer

$$\int_0^{1/2} S(z)dz \tag{97.423}$$

où $S(z) = \pi r(z)^2$ est la surface de la section de la sphère à la hauteur z. Le rayon $r(z)$ est donné par Pythagore : $z^2 + r(z)^2 = 1$. Nous avons donc

$$V = 12\int_0^{1/2} \pi(1 - z^2)dz = \frac{11\pi}{2}. \tag{97.424}$$

(3) Intersection de deux cylindres. Ici, si nous passons en coordonnées cylindriques, nous pouvons simplifier l'expression d'un des deux cylindres, mais nous compliquerions l'expression de l'autre. Nous restons donc en cartésiennes. Les contraintes $x^2 + y^2 < 1$ et $y^2 + z^2 < 1$ font que le plus simple est de laisser y varier de -1 à 1 et de faire z et x s'adapter :

$$V = \int_{-1}^{1} \Big(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \big(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx\big)dz\Big)dy. \tag{97.425}$$

L'avantage de cet ordre d'intégration est que les intégrales sur x et sur z donnent le même résultat : $2\sqrt{1-y^2}$. Il reste donc à calculer

$$\int_{-1}^{1} 4(1 - y^2)dy = \frac{16}{3}. \tag{97.426}$$

Exercise 136 exoIntMult0006
Calculez le volume du solide limité par $z^2 = x^2 + y^2$; $x^2 + y^2 = 2x$ et $z = 0$. corrIntMult0006

Correction of the exercise 136

Pour votre culture générale, et parce que c'est important dans d'autres exercices, sachez que l'équation $x^2 + y^2 = 2x$ décrit un cercle. En coordonnées cylindriques, le domaine devient

$$\begin{cases} z^2 = r^2 \\ r^2 = 2r\cos(\theta) \\ z = 0. \end{cases} \tag{97.427}$$

Dès que $\cos(\theta) < 0$, le r a un problème, donc le domaine d'intégration en θ se limite à $]-\frac{\pi}{2}, \frac{\pi}{2}[$, et l'intégrale à calculer est

$$\int_{-\pi/2}^{\pi/2} d\theta \int_0^{2\cos(\theta)} dr \int_0^r rdz = \frac{32}{9}. \tag{97.428}$$

Exercise 137 exoIntMult0007
Calculez le volume commun à la sphère $x^2 + y^2 + z^2 = 4$ et au cylindre $x^2 + y^2 = 1$. corrIntMult0007

Correction of the exercise 137

Ici nous passons encore en cylindriques [11] et le domaine d'intégration devient

$$\begin{cases} r^2 + z^2 = 4 \\ r^2 = 1, \end{cases} \tag{97.429}$$

11. Il ne faut pas sauter sur les sphériques dès qu'on voit une sphère !

donc

$$V = \int_0^{2\pi} d\theta \int_0^1 dr \int_{-\sqrt{4-r^2}}^{\sqrt{4-r^2}} r dz = 2\left(\frac{8}{3} - \sqrt{3}\right) \cdot 2\pi. \tag{97.430}$$

Exercise 138 exoIntMult0008

Calculez le volume délimité par les paraboloïdes $z = x^2 + y^2$ et $z = \frac{1}{2}(x^2 + y^2) + 2$. corrIntMult0008

Correction of the exercise 138

Nommons les paraboloïdes $P_1 \equiv z = x^2 + y^2$ et $P_2 \equiv \frac{1}{2}(x^2 + y^2) + 2$, et passons en coordonnées cylindriques :

$$\begin{aligned} P_1 &\equiv z = r^2 \\ P_2 &\equiv z = \frac{1}{2}r^2 + 2. \end{aligned} \tag{97.431}$$

La paraboloïde P_2 est plus haute que la paraboloïde P_1 jusqu'à $r = 2$ (solution de l'équation $r^2 = \frac{1}{2}r^2 + 2$). Nous allons donc calculer le volume contenu entre les deux paraboloïdes en faisant la différence

$$\int_{r\leqslant 2} P_2 - \int_{r\leqslant 2} P_1, \tag{97.432}$$

c'est-à-dire

$$\int_0^{2\pi} d\theta \int_0^2 dr \int_0^{(r^2/2)+2} dz \cdot r - \int_0^{2\pi} d\theta \int_0^2 dr \int_0^{r^2} dz \cdot r = 12\pi - 8\pi = 4\pi. \tag{97.433}$$

Exercise 139 exoIntMult0009

Calculez les intégrales suivantes. Représenter le domaine et effectuer un changement de variable peut être utile.

(1)

$$\int_0^1 \left(\int_0^{\sqrt{1-x^2}} e^{\sqrt{x^2+y^2}} dy\right) dx \tag{97.434}$$

(2)

$$\int_0^4 \left(\int_{\sqrt{y}}^2 y e^{x^5} dx\right) dy \tag{97.435}$$

corrIntMult0009

Correction of the exercise 139

(1) Il s'agit d'intégrer la fonction $(x, y) \mapsto e^{\sqrt{x^2+y^2}}$ sur un quart de disque de rayon 1. Le passage en polaire s'impose donc. L'intégrale à calculer est

$$I = \int_0^{\pi/2} d\theta \int_0^1 r e^r dr. \tag{97.436}$$

L'intégrale sur r se fait par partie en posant $u = r$ et $dv = e^r dr$. Nous avons alors

$$I = \int_0^{\pi/2} [re^r - e^r]_0^1 d\theta = \frac{\pi}{2}. \tag{97.437}$$

(2) Cette intégrale devient subitement plus facile si on intègre d'abord par rapport à y et puis par rapport à x :

$$\int_0^4 \left(\int_{\sqrt{y}}^2 y e^{x^5} dx\right) dy = \int_0^2 dx \int_0^{x^2} y e^{x^5} dy = \frac{e^{32} - 1}{10}. \tag{97.438}$$

Exercise 140 exoIntMult0011

Soit $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ une fonction continue et positive. Calculer le volume de la région limitée par la surface de révolution obtenue par rotation de la courbe $\{(x, y, z) \in R^3 \mid y = f(x), z = 0\}$ autour de l'axe Ox et par les plans $\{(x, y, z) \mid x = a\}$ et $\{(x, y, z) \mid x = b\}$. corrIntMult0011

Correction of the exercise 140

Le volume d'un domaine est l'intégrale de la fonction constante 1 sur le domaine. Nous passons aux coordonnées cylindriques « autour de x »

$$\begin{cases} x = x \\ y = r\cos(\theta) \\ z = r\sin(\theta) \end{cases}, \tag{97.439}$$

dont le jacobien vaut r. Nous avons donc à calculer :

$$V = \int_a^b dx \int_0^{2\pi} d\theta \int_0^{f(x)} r dr = 2\pi \int_a^b \frac{1}{2} f(x)^2 = \pi \int_a^b f^2, \tag{97.440}$$

formule que certains étudiants ont probablement déjà vue.

Exercise 141 exoIntMult0012

Intégrer la fonction $(x, y, z) \mapsto x+y$ sur le domaine délimité par les surfaces suivantes $\{(x, y, z) \mid y = 1 - x^2\}$, $\{(x, y, z) \mid z = 0\}$, $\{(x, y, z) \mid z = 2\}$ et $\{(x, y, z) \mid y = 1 - x\}$. corrIntMult0012

Correction of the exercise 141

Ce qui est déjà certain, c'est que l'intégrale sur z va de 0 à 2. Les solutions de l'équation $1 - x^2 = 1 - x$, étant $x = 0$ et $x = 1$, l'intégrale sur x va de 0 à 1. Pour chaque x, la variable t prend les valeurs de $1 - x$ à $1 - x^2$. L'intégrale à calculer est donc

$$I = \int_0^2 dz \int_0^1 dx \int_{1-x}^{1-x^2} (x + y)dy = \int_0^2 = 2\int_0^1 \left[xy + \frac{y^2}{2}\right]_{1-x}^{1-x^2} dx = \frac{11}{30}. \tag{97.441}$$

Exercise 142 exoIntMult0013

Intégrer les fonctions 1 et $(x, y, z) \to y^2$ sur le domaine délimité par les surfaces suivantes $\{(x, y, z) \mid x^2 - \frac{y^4}{4} + z^2 = 1\}$, $\{(x, y, z) \mid y = 0\}$ et $\{(x, y, z) \mid y = 2\}$. corrIntMult0013

Correction of the exercise 142

Nous passons à des coordonnées presque cylindriques :

$$\begin{cases} x = r\cos(\theta) \\ y = \sqrt{2}u \\ z = r\sin(\theta) \end{cases}, \tag{97.442}$$

dont le jacobien est $\sqrt{2}r$. Les bornes du domaine deviennent $0 < u < \sqrt{2}$ et $r^2 - u^4 = 1$, donc les intégrales à calculer sont

$$\int_0^{\sqrt{2}} du \int_0^{2\pi} d\theta \int_0^{\sqrt{1+u^4}} \sqrt{2} r f(r, \theta, u) dr. \tag{97.443}$$

(1) L'intégrale de la fonction 1 est très simple et donne $\frac{18\pi}{5}$.

(2) La fonction $y^2 = 2u^2$ demande le calcul suivant :

$$\int_0^{\sqrt{2}} du \int_0^{2\pi} d\theta \int_0^{\sqrt{1+u^4}} 2u^2\sqrt{2} r dr = \frac{152\pi}{21}. \tag{97.444}$$

97.13 Théorème de la fonction implicite

Exercise 143 exoImplicite0001

Soit $F : \mathbb{R}^2 \times \mathbb{R}_0^+ \to \mathbb{R} : (x, y, z) \mapsto z + \ln(z) - xy$

(1) Prouver qu'il existe une fonction $Z(x, y)$ dans un voisinage de $(1, 1)$ telle que

$$F(x, y, Z(x, y)) = 0 \tag{97.445}$$

Prouver l'unicité d'une telle fonction.

(2) Calculer $(\partial_x Z)(x, y)$, $(\partial_y Z)(x, y)$ et $(\partial^2_{xy} Z)(x, y)$ en fonction de x, y, $Z(x, y)$.

corrImplicite0001

Correction of the exercise 143

Nous considérons la fonction

$$\begin{aligned} F\colon \mathbb{R}^2 \times \mathbb{R}_0^+ &\to \mathbb{R} \\ (x, y, z) &\mapsto z + \ln(z) - xy, \end{aligned} \tag{97.446}$$

et nous nous intéressons à la fonction $Z(x, y)$ définie au voisinage de $(x, y) = (1, 1)$ par la condition

$$F\big(x, y, Z(x, y)\big) = 0. \tag{97.447}$$

Nous pouvons facilement trouver $z_0 = Z(1, 1)$, étant donné que sa définition est que

$$z_0 + \ln(z_0) - 1 = 0. \tag{97.448}$$

Facile de voir que $z_0 = 1$ remplit la condition. Dans les notations du théorème de la fonction implicite, nous avons donc $\tilde{x} = (1, 1)$ et $\tilde{y} = z_0$. Le déterminant à vérifier pour le théorème est une matrice 1×1 parce qu'il n'y a que une seule variable z que l'on veut exprimer en termes des autres (x et y) :

$$\frac{\partial F}{\partial z}(1, 1, z_0) = 1 + \frac{1}{z_0}. \tag{97.449}$$

Cela n'est pas nul parce que $z_0 = 1$. Il existe donc un voisinage $\mathcal{U}$ de $(1, 1)$ dans $\mathbb{R}^2$ et un voisinage V de z_0 dans $\mathbb{R}$ tels que $\forall (x, y) \in \mathcal{U}$, il existe un et un seul $Z(x, y)$ solution de $F\big((x, y), Z(x, y)\big) = 0$.

Pour calculer les dérivées de Z, il faut bien se rendre compte que, pour chaque x et y, nous avons l'égalité

$$Z(x, y) + \ln\big(Z(x, y)\big) - xy = 0. \tag{97.450}$$

Nous pouvons dériver cette égalité partiellement par rapport à x pour obtenir [12]

$$(\partial_x Z)(x, y) + \frac{(\partial_x Z)(x, y)}{Z(x, y)} - y = 0. \tag{97.451}$$ EqPartialXZexoI

Nous pouvons aisément isoler $(\partial_x Z)(x, y)$ dans cette équation :

$$\frac{\partial Z}{\partial x} = \frac{yZ}{Z + 1}. \tag{97.452}$$

Notez que la fonction « inconnue » apparaît dans l'expression de la dérivée. C'est la vie, et c'est pourquoi ce chapitre parle de fonctions *implicite*.

De la même manière, nous trouvons

$$\frac{\partial Z}{\partial y} = \frac{xZ}{Z + 1}. \tag{97.453}$$

12. Nous n'allons plus, dans l'avenir, toujours écrire explicitement les dépendances en x et y

Pour la dérivée seconde $\partial_{xy}Z$, d'abord nous savons que Z est une fonction C^2 [13], donc l'ordre des dérivées n'a pas d'importance. Dérivons donc par rapport à x l'expression de $\partial_y Z$:

$$\frac{\partial^2 Z}{\partial x\partial y} = \frac{Z}{Z+1} + x\left(\frac{\partial_x Z}{Z+1} - Z\frac{\partial_x Z}{(Z+1)^2}\right). \tag{97.454}$$

Nous devons encore substituer $\partial_x Z$ par sa valeur (97.451) pour trouver la réponse finale

$$\frac{\partial^2 Z}{\partial x\partial y}(x,y) = \frac{Z}{z+1} + x\left(\frac{yZ}{(Z+1)^2} - \frac{yZ^2}{(Z+1)^2}\right). \tag{97.455}$$

Exercise 144 exoImplicite0002

Calculer Y' où $Y : D \subset \mathbb{R} \to \mathbb{R}$ est définie implicitement dans un voisinage de 1 par l'équation $y^x = x^y$. corrImplicite0002

Correction of the exercise 144

À faire.

Exercise 145 exoImplicite0003

Écrire le développement de Taylor d'ordre 2 en $(0,0)$ de la fonction $Z : D \subset \mathbb{R}^2 \to \mathbb{R}$ définie implicitement par l'équation $ze^z = x + y$. Prouver l'unicité de Z. corrImplicite0003

Correction of the exercise 145

Nous avons $f\big((x,y),z\big) = ze^z - x - y$. Évidemment, $z(0,0) = 0$, et pour que le théorème de la fonction implicite fonctionne, il faut que

$$\frac{\partial F}{\partial z}\big((0,0),0\big) = (ze^z + e^z)_{x=y=z=0} \neq 0. \tag{97.456}$$

Cette condition est vérifiée. Le développement de $z(x,y)$ à l'ordre 2 autour de $(0,0)$ est donné par

$$z(0,0) + x\partial_x z(0,0) + y\partial_y z(0,0) + \frac{x^2}{2}\partial_x^2 z(0,0) + \frac{y^2}{2}\partial_y^2 z(0,0) + xy\partial_{xy}^2 z(0,0). \tag{97.457}$$

Tout l'exercice se réduit donc à calculer les dérivées partielles de $z(x,y)$ par rapport à x et à y jusqu'à l'ordre 2, et de les évaluer en $(0,0)$. La relation de définition de $z(x,y)$ est

$$z(x,y)e^{z(x,y)} - x - y = 0. \tag{97.458}$$ EqDefZxyImplI

Nous en extrayons la valeur de $z(0,0) = 0$. Pour le reste, nous dérivons la relation $F\big(x,y,z(x,y)\big) = 0$, et nous évaluons en $(0,0)$. Par exemple pour $\partial_x z(0,0)$, nous commençons par écrire

$$(\partial_1 F)(0,0,z(0,0)) + (\partial_3 F)(0,0,z(0,0,0))\frac{\partial z}{\partial x}(0,0) = 0. \tag{97.459}$$

Dans cette équation, $\partial_1 F$ dénote la dérivée de F par rapport à sa première variable. Celle-là est toujours égale à -1 parce que $F(x,y,z) = ze^z - x - y$. D'autre part,

$$(\partial_3 F)(x,y,z) = ze^z + e^z, \tag{97.460}$$

donc, en isolant $\partial_x z(0,0)$, nous avons

$$\frac{\partial z}{\partial x}(0,0) = \frac{1}{e^{z(0,0)}(z(0,0)+1)} = 1. \tag{97.461}$$

Le même genre de raisonnements amène les autres dérivées. Pour la dérivée par rapport à y, nous écrivons

$$\frac{\partial F}{\partial x_1}\underbrace{\frac{\partial x_1}{\partial y}}_{=0} + \frac{\partial F}{\partial x_2}\underbrace{\frac{\partial x_2}{\partial y}}_{=-1} + \underbrace{\frac{\partial F}{\partial x_3}}_{=ze^z+e^z}\frac{\partial z}{\partial y} = 0. \tag{97.462}$$

13. parce que F est C^2, voir l'énoncé du théorème.

Les résultats sont

$$\begin{aligned}\frac{\partial z}{\partial x}(0,0)&=1\\ \frac{\partial z}{\partial y}(0,0)&=1\\ \frac{\partial^2 z}{\partial x^2}(0,0)&=-2\end{aligned}\tag{97.463}$$

Exercise 146 exoImplicite0004

Soit $F:\mathbb{R}^3\to\mathbb{R}^2:(x,y,z)\to(x^2+y^2+z^2-1,x^2+y^2-x)$

(1) Prouver qu'il existe 4 fonctions $\varphi=x\to(Y(x),Z(x))$ définie dans un voisinage de $\frac{3}{4}$ telles que $F(x,Y(x),Z(x))=0$.

(2) Calculer $Y'(x)$, $Z'(x)$ et $Y''(x)$.

corrImplicite0004

Correction of the exercise 146

Cet exercice est particulièrement facile parce que l'équation

$$F(x,y,z)=(x^2+y^2+z^2-1,x^2+y^2-x)=(0,0)\tag{97.464}$$

peut être résolue explicitement pour $Y(x)$ et $Z(x)$. En effet, la deuxième composante dit que

$$x^2+Y(x)^2-x=0,\tag{97.465}$$

donc

$$Y^{(\pm)}(x)=\pm\sqrt{x-x^2}.\tag{97.466}$$

Pour chacune de ces deux solutions possibles, l'équation pour la première composante devient

$$Z(x)^2=1-x,\tag{97.467}$$

de telle façon à avoir

$$Z^{(\pm)}(x)=\pm\sqrt{1-x}.\tag{97.468}$$

Il y a donc bien 4 possibilités.

Exercise 147 exoImplicite0005

Calculez

$$\lim_{x\to0}\frac{y(x)}{\cos(x)-1}$$

où y est la fonction définie implicitement au voisinage de 0 par la relation $e^{yx}-1=x^2+y$. corrImplicite0005

Correction of the exercise 147

La relation de définition de $y(x)$ est

$$F\big(x,y(x)\big)=e^{xy}-1-x^2-y=0.\tag{97.469}$$

Nous trouvons que $y(0)=0$, ce qui fournit une indétermination dans la limite à calculer. Afin d'utiliser la règle de l'Hospital, nous devons trouver la dérivée de y au point $x=0$.

Pour cela, nous savons que

$$\frac{\partial F\big(x,y(x)\big)}{\partial x}=(\partial_1F)\big(x,y(x)\big)+(\partial_2F)\big(x,y(x)\big)y'(x),\tag{97.470}$$

donc

$$y'(x)=-\frac{(\partial_1F)(x,y)}{\partial_2F(x,y)}.\tag{97.471}$$

Nous calculons par ailleurs facilement que

$$\begin{aligned}\frac{\partial F}{\partial x_1}(x,y) &= ye^{xy} - 2x\\ \frac{\partial F}{\partial x_2}(x,y) &= xe^{xy} - 1,\end{aligned} \tag{97.472}$$

donc

$$y'(x) = -\frac{ye^{xy} - 2x}{xe^{xy} - 1}. \tag{97.473}$$

Nous devons donc calculer

$$\lim_{x\to 0} \frac{y'(x)}{-\sin(x)}. \tag{97.474}$$

Hélas, nous tombons encore une fois sur une indétermination parce que

$$y'(0) = \frac{y(0)e^{0y(0)} - 2\cdot 0}{0\cdot e^{0\cdot y(0)} - 1}. \tag{97.475}$$

Nous devons donc utiliser encore une fois la règle de l'Hospital. C'est encore un peu de calcul et la réponse est

$$y''(0) = 2, \tag{97.476}$$

ce qui fait qu'au final,

$$\lim_{x\to 0} \frac{y(x)}{\cos(x) - 1} = \lim_{x\to 0} \frac{y''(x)}{-\cos(x)} = -2. \tag{97.477}$$

Exercise 148 exoImplicite0006

Écrire l'équation de la droite tangent à la courbe

$$y^2 + \sin(xy) - 1 = 0$$

au point $(0,1)$. Donner les coordonnées d'un point où la tangente à cette courbe est horizontale. corrImplicite0006

Correction of the exercise 148

La définition de la fonction $y(x)$ est que

$$y^2 + \sin(xy) - 1 = 0 \tag{97.478}$$ EqNNDerImpliSix

pour tout x.

Nous vérifions que $y(0) = 1$ résous bien cette équation et que le théorème de la fonction implicite s'applique dans un voisinage de $x = 0$ et $y = 1$.

Nous dérivons cette équation par rapport à x (qui apparaît dans le y) :

$$y'\big(2y + x\cos(xy)\big) = -y\cos(xy). \tag{97.479}$$ EqDerrSixImpli

Pour savoir le coefficient directeur de la tangente au point $(0,1)$, nous devons poser $x = 0$ et $y = 1$ dans cette équation. En tenant compte du fait que $\cos(0) = 1$, nous trouvons

$$y'(0) = \frac{1}{2}. \tag{97.480}$$

La droite recherchée est celle qui passe par le point $(0,1)$ et qui a $-1/2$ comme coefficient directeur [14].

Trouver une tangente horizontale est un petit peu plus subtil. Si nous posons $y'(x) = 0$ dans la relation (97.479), nous trouvons l'équation suivante pour x :

$$y(x)\cos\big(xy(x)\big) = 0. \tag{97.481}$$ EqRechercheTgHorSix

14. Si vous ne savez pas comment en déduire une équation paramétrique ou cartésienne de la droite, demandez à madame Aude.

La première chose que nous voulons faire est de trouver un x tel que $y(x) = 0$. Hélas, la relation de définition (97.478) donne $0 = -1$ lorsqu'on pose $y = 0$. Cela montre que la fonction y ainsi définie ne passe jamais par zéro.

La seconde chose à faire pour annuler le membre de gauche de (97.481) est d'essayer de trouver x tel que $xy(x) = \frac{\pi}{2}$. Cela demanderais $y = \pi/2x$. Encore une fois, cet essai échoue si on remplace dans l'équation (97.478).

Le troisième essai est de chercher un x tel que $xy(x) = -\pi/2$. Cette fois, en remplaçant dans (97.478), ça fonctionne. Nous trouvons

$$\frac{\pi^2}{4x^2} = 2, \tag{97.482}$$

ce qui fait $x = \pm\frac{\pi}{2\sqrt{2}}$.

Nous ne sommes, cependant, pas sorti de l'auberge pour autant. En effet, nous avons prouvé que la fonction $y(x)$ existait dans un voisinage de $x = 0$ et $y = 1$. Rien ne prouve que ce voisinage va jusqu'à $\pm\pi/2\sqrt{2}$. Nous devons donc rappliquer le théorème de la fonction implicite pour prouver que la relation (97.478) définit bien une fonction $y(x)$ dans un voisinage de $(\tilde{x}, \tilde{y}) = (\pi/2\sqrt{2}, -\sqrt{2})$. Il est d'abord vrai que $F(\tilde{x}, \tilde{y}) = 0$, et ensuite,

$$\frac{\partial F}{\partial y}(\tilde{x}, \tilde{y}) = 2\tilde{y} + \tilde{x}\cos(\tilde{x}\tilde{y}) = -2\sqrt{2} \neq 0, \tag{97.483}$$

donc la fonction y est bien définie au voisinage du point que nous avons sélectionné.

Exercise 149 exoImplicite0007

Quelles sont les équations des plans tangents à la surface

$$4x^2 + 16y^2 + 8z^2 = 1,$$

parallèles au plan $x - 2y + 2z + 7 = 0$? corrImplicite0007

Correction of the exercise 149

La relation $4x^2 + 16y^2 + 8z^2 = 1$ définit bien $z(x, y)$ pourvu que $z \neq 0$, en effet

$$\frac{\partial F}{\partial z} = 16z. \tag{97.484}$$

En procédant comme d'habitude, nous pouvons trouver les dérivées partielles $\partial_x z$ et $\partial_y z$. Nous pouvons même utiliser la formule explicite

$$z(x, y) = \pm\sqrt{\frac{1 - 4x^2 - 16y^2}{8}}. \tag{97.485}$$

Ce faisant, nous avons oublié une série de points : ceux où $z = 0$. Pour traiter ces points, nous savons que si $y \neq 0$, alors nous avons l'expression

$$y(x, z) = \sqrt{\frac{1 - 4x^2 - 8z^2}{16}} \tag{97.486}$$

qui permet de trouver les plans tangents. Il reste enfin les points avec $y = z = 0$ donnés par $x = \pm 1/4$ qui doivent être traités séparément.

Exercise 150 exoImplicite0008

Soit la fonction

$$F : \mathbb{R}^3 \to \mathbb{R}^2 : (x, y, z) \mapsto (x + y + z,\ x^2 + y^2 + z^2 - 1). \tag{97.487}$$

On définit $M \subset \mathbb{R}^3$ comme l'ensemble

$$M = \{(x, y, z) \in \mathbb{R}^3 \text{ tels que } F(x, y, z) = (0, 0)\}. \tag{97.488}$$

(1) Montrer que M est compact. Montrer que M est une variété C^1 dans $\mathbb{R}^3$. Quelle est sa dimension ?

(2) Prouver qu'il existe des fonctions $\phi : \mathbb{R} \to \mathbb{R}^2 : y \mapsto (X(y), Z(y))$ définies dans un voisinage V de $y = 0$ telles que $F(X(y), y, Z(y)) = (0, 0)$ $\forall y \in V$.

(3) Donner une approximation des fonctions $y \to X(y)$ autour de $y = 0$ par un polynôme du premier degré en une indéterminée.

corrImplicite0008

Correction of the exercise 150

Un point (x, y, z) est dans M lorsque

$$\begin{cases} x + y + z = 0 \\ x^2 + y^2 + z^2 = 1. \end{cases} \tag{97.489}$$

La seconde équation est celle d'une sphère de rayon 1, donc M est inclus dans la sphère et est donc bornée. D'autre part, tant la première équation (un plan dans $\mathbb{R}^3$) que la seconde définissent des ensembles fermés. En tant qu'intersection de fermés, l'ensemble M est fermé. Maintenant que nous savons que M est ferme et borné, nous savons que M est compact.

Maintenant nous posons

$$\begin{aligned} F\colon \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x, y, z) &\mapsto (x + y + z, x^2 + y^2 + z^2). \end{aligned} \tag{97.490}$$

Les relations de définition pour $X(y)$ et $Z(y)$ sont

$$\begin{cases} X(y) + y + Z(y) = 0 \\ X(y)^2 + y^2 + z(y)^2 - 1 = 0. \end{cases} \tag{97.491}$$

EqHuitEqsqDefXZ

Étant donné que

$$\begin{vmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial z} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2x & 2z \end{vmatrix} = 2(z - x). \tag{97.492}$$

Si $y = 0$, les équations deviennent

$$\begin{cases} x + z = 0 \\ x^2 + y^2 = 1, \end{cases} \tag{97.493}$$

de façon que $z - x \neq 0$. Donc le condition $F\big(X(y), y, Z(y)\big) = 0$ définit bien les fonctions X et Z au voisinage de $y = 0$ et de $(x, z) = (0, 0)$.

L'approximation de $X(y)$ cherchée est évidemment Taylor, c'est-à-dire

$$X(0) + y(\partial_y X)(0). \tag{97.494}$$

Pour trouver $X(0)$, nous devons résoudre le système

$$\begin{cases} X(0) + Z(0) = 0 \\ X(0)^2 + Z(0)^2 = 1, \end{cases} \tag{97.495}$$

donc les solutions sont $X(0) = \pm 1/2$ et $Z(0) = -X(0)$. Nous avons donc le choix de travailler autour de deux points différents :

$$(x, y, z) = (\frac{1}{2}, 0, -\frac{1}{2}) \tag{97.496}$$

EqPossXYzeroUn

ou bien

$$(x, y, z) = (-\frac{1}{2}, 0, \frac{1}{2}). \tag{97.497}$$

EqPossXYzerodeux

Nous trouvons $X'(0)$ en dérivant les deux équations de définition (97.491) par rapport à y et en résolvant le système par rapport à $X'(0)$. Le système à résoudre est

$$\begin{cases} X'(y) + 1 + Z'(y) = 0 \\ 2X(y)X'(y) + 2y + 2Z(y)Z'(y) = 0. \end{cases} \tag{97.498}$$

La résolution donne

$$X'(y) = \frac{-2y + 2Z(y)}{2X(y) - 2Z(y)}. \tag{97.499}$$

Étant donné que nous connaissons les valeurs de $X(0)$ et $Z(0)$ données par (97.496), nous trouvons

$$X'(0) = -\frac{1}{2}. \tag{97.500}$$

Dans ce cas, l'approximation est

$$X(y) \sim \frac{1}{2} - \frac{1}{2}y. \tag{97.501}$$

Si nous avions choisit de travailler avec la possibilité (97.497), alors nous aurions obtenu

$$X(y) \sim -\frac{1}{2} - \frac{1}{2}y. \tag{97.502}$$

Exercise 151 exoImplicite0009

Soit $\Phi : \mathbb{R}\times\mathbb{R}^2 \to \mathbb{R}^2 : (t, v) \to \Phi(t, v) =: \varphi_t(v)$ une famille à un paramètre de difféomorphismes de $\mathbb{R}^2$ ($\varphi_t : \mathbb{R}^2 \to \mathbb{R}^2 : v \to \varphi_t(v)$ est un difféomorphisme de $\mathbb{R}^2$ pour $t \in \mathbb{R}$).

(1) Montrer que si φ_0 possède un point fixe v_0 et que le spectre de $(d\varphi_0)(v_0)$ ne contient pas 1 (le réel 1 n'est pas une valeur propre de l'opérateur $(d\varphi_0)(v_0)$ de R^2) alors il existe $\varepsilon > 0$ tel que pour $t \in]-\varepsilon, \varepsilon[$ le difféomorphisme φ_t possède aussi un point fixe.

(2) Montrer en exhibant un exemple que l'hypothèse sur le spectre de $(d\varphi_0)(v_0)$ est essentielle. corrImplicite0009

Correction of the exercise 151

À faire.

97.14 Variétés et extrémums liés

Exercise 152 exoVariete0001

Les ensembles suivants sont ils des variétés ? Si oui déterminer leur dimension et leur espace tangent (en un point arbitraire)

(1) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$
(2) $\{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 < 1\}$
(3) $\{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 \leqslant 1\}$
(4) $\{(x, y) \in \mathbb{R}^2 \mid |x| + |y| = 1\}$
(5) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$
(6) $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$
(7) $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2\}$
(8) $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$
(9) $\{(x, y, z) \in \mathbb{R}^3 \mid ((x - 1)^2 + y^2)((x + 1)^2 + y^2) = z\}$
(10) $\{(x, y) \in \mathbb{R}^2 \mid ((x - 1)^2 + y^2)((x + 1)^2 + y^2) = 1\}$ corrVariete0001

Correction of the exercise 152

(1) Si nous prenons un point (x, y) différent du pôle sud S, alors nous avons une carte autour du point donnée par

$$\begin{aligned} F_1 :]-\frac{\pi}{2}, \frac{3\pi}{2}[&\to \mathbb{R}^2 \\ \theta &\mapsto \big(\cos(\theta), \sin(\theta)\big). \end{aligned} \tag{97.503}$$

La différentielle de cette application est

$$dF_1(\theta) = \big(-\sin(\theta), \cos(\theta)\big), \tag{97.504}$$

qui est toujours de rang maximum (c'est-à-dire 1) parce qu'il n'arrive jamais que $dF_1(\theta) = (0,0)$.

Autour du point S, nous pouvons prendre la carte $F_2 :]-\pi, 0, [\to \mathbb{R}^2$ définie de la même manière.

Les cartes F_1 et F_2 forment un atlas pour le cercle (de nombreux autres choix sont évidement possibles). Notons toutefois qu'il n'est pas possible de trouver une seule carte qui paramétrise tout le cercle parce que le cercle est fermé alors que les cartes doivent partir d'ensembles ouverts.

(2)

(3)

(4) Il est apparent que cela ne va pas être une variété à cause des coins. Étudions donc ce qu'il s'y passe en regardant la figure 97.7

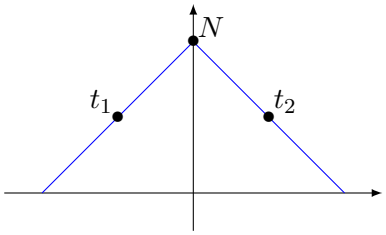


FIGURE 97.7: Ceci n'est pas une variété. Même pas le dessin d'une variété. LabelFigCoinPasVar

Supposons avoir une carte

$$t \mapsto F\big(x(t), y(t)\big). \tag{97.505}$$

Il y a un seul t_0 tel que $F(t_0) = N$. Si t_1 est tel que $f(t_1)$ est à gauche de N, nous avons

$$\frac{x(t_1)}{y(t_1)} = -1, \tag{97.506}$$

et si t_2 est tel que $F(t_2)$ est à droite de N, alors

$$\frac{x(t_2)}{y(t_2)} = 1. \tag{97.507}$$

Il n'y a donc que (au maximum) un seul point (t_0) où le rapport $x(t)/y(t) \neq \pm 1$. Il n'est donc pas possible d'avoir $x(t)$ et $y(t)$ continues parce que $y(t)$ ne s'annulant pas, si $x(t)$ et $y(t)$ étaient continues, alors le rapport serait continu.

(5)

(6) L'ensemble $x^2 + y^2 = 1$ dans $\mathbb{R}^3$ est un cylindre dont l'équation en coordonnées cylindrique est $r = 1$. Pour avoir un atlas, il faut donc simplement prendre le même atlas que le cercle de rayon 1, et de le « multiplier » par $\mathbb{R}$.

(7) L'ensemble donné par l'équation $x^2 + y^2 = z^2$ est un cône. Nous avons donc des doutes quant au fait que ce serait une variété, à cause du sommet. Nous allons montrer que ce n'est pas une variété parce que nous pouvons trouver trois vecteurs tangents linéairement indépendants en $(0,0,0)$, alors qu'une variété de dimension deux ne peut pas en avoir plus que deux (proposition 7.342). Nous considérons les trois chemins

$$\begin{aligned} \gamma_1(t) &= (-t, 0, t) \\ \gamma_2(t) &= (t, 0, t) \\ \gamma_3(t) &= (0, t, t). \end{aligned} \tag{97.508}$$

Il n'est pas très compliqué de prouver que ces trois chemins sont contenus dans l'ensemble. Les vecteurs tangents correspondants sont

$$\begin{aligned}\gamma_1'(0) &= (-1,0,1)\\ \gamma_2'(0) &= (1,0,1)\\ \gamma_2'(0) &= (0,1,1).\end{aligned} \tag{97.509}$$

Ces trois vecteurs sont linéairement indépendants et génèrent tout $\mathbb{R}^3$.

(8) Cet ensemble est la sphère. Pour créer un atlas, il ne faut pas oublier de mettre deux cartes par angles. Il est facile de trouver une carte, centrée en le pôle nord, qui ne contient ni le pôle nord ni le pôle sud en donnant un angle et une distance. Ensuite, il suffit de prendre une seconde carte centrée au pôle ouest et qui ne contient ni le pôle est ni le pôle ouest (mais bien les deux autres pôles).

(9) Nous allons utiliser la caractérisation de la proposition 20.7 avec la fonction

$$G(x,y,z) = \big((x-1)^2+y^2\big)\big((x+1)^2+y^2\big) - z = 0. \tag{97.510}$$

Il est certain que $G=0$ est exactement l'ensemble considéré. Il faut encore vérifier que la fonction G satisfait aux hypothèses de la proposition, c'est-à-dire que la différentielle dG est de rang maximum. Ici, le rang maximum est 1. Nous avons

$$dG(x,y,z) = \begin{pmatrix} 4x(y^2+x^2-1)\\ 4y(y^2+x^2+1)\\ -1 \end{pmatrix}. \tag{97.511}$$

Grâce au -1, c'est toujours de rang 1. Donc oui, l'ensemble considéré est une variété.

(10) Cette fois, le -1 n'apparaît pas, donc nous ne pouvons pas utiliser la même fonction pour prouver que $G=0$ est une variété.
Autour de $x=0$, nous pouvons résoudre par rapport à y :

$$y = \pm\sqrt{\sqrt{4x^2+1}-x^2-1}. \tag{97.512}$$

Ces deux courbes sont représentées à la figure 97.8.

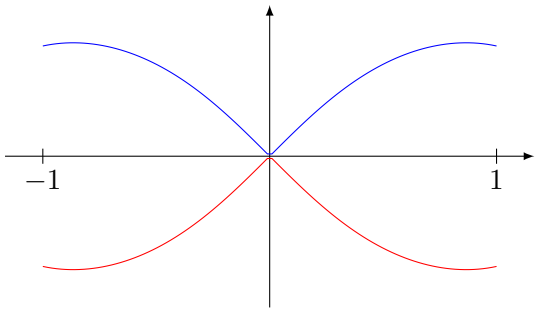


FIGURE 97.8: Ce à quoi ça ressemble. En bleu la partie sans le signe, et en rouge avec le signe moins. LabelFigExoVarj

En prenant un chemin bleu à gauche et rouge à droite, puis un autre chemin rouge à gauche et bleu à droite, on a deux chemins qui ont des vecteurs tangents linéairement indépendants. En n'importe quel autre point, il n'y a qu'un seul vecteur tangent. Donc l'ensemble proposé n'est pas une variété.

Exercise 153 exoVariete0002

Trouver les extrémums de $f : \mathbb{R}^3 \to \mathbb{R} : (x,y,z) \mapsto x^2y$ relativement à la sphère $\{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2=1\}$.

corrVariete0002

Correction of the exercise 153

Nous avons $f(x,y,z)=x^2y$ et $G(x,y,z)=x^2+y^2+z^2-1$. Le système à résoudre est

$$\begin{cases} 2xy+2\lambda x=0\\ x^2+2\lambda y=0\\ 2\lambda z=0\\ x^2+y^2+z^2-1=0. \end{cases} \tag{97.513}$$

Il y a directement deux cas à séparer : $\lambda=0$ et $\lambda\neq 0$.

Commençons par $\lambda=0$, alors $x=0$ et $y^2+z^2=1$. Cela fait tout un cercle de points candidats. Sur ce cercle, la fonction vaut zéro, et en dehors du cercle, le signe de la fonction est le signe de y. Donc les points du cercle avec $y>0$ sont des minimums locaux (parce que dans un voisinage, la fonction est positive), et les points du cercle avec $y<0$ sont des minimums locaux. Les deux points avec $y=0$ ne sont ni l'un ni l'autre.

Analysons maintenant ce qu'il se passe si $\lambda\neq 0$. Dans ce cas, $z=0$ et le système devient

$$\begin{cases} 2x(y+\lambda)=0\\ x^2+2\lambda y=0\\ x^2+y^2-1=0. \end{cases} \tag{97.514}$$

Si $x=0$, alors $y=\pm1$ et $\lambda=0$, ce qui fait rejeter la possibilité $x=0$. Donc $x\neq 0$ et $y=-\lambda$. Il reste les équations

$$\begin{cases} x^2-2\lambda^2=0\\ x^2+\lambda^2-1=0. \end{cases} \tag{97.515}$$

Il y a quatre solutions : $x=\pm\sqrt{2/3}$ et $y=\pm\sqrt{1/3}$. Ces points sont maximums et minimums locaux.

Exercise 154 exoVariete0003

Trouver les extrémums de la fonction $f:\mathbb{R}^3\to\mathbb{R}:(x,y,z)\mapsto xy^2z^3$ relativement à la portion de plan $S:=\{(x,y,z)\in\mathbb{R}^3\mid x+y+z=12,x,y,z>0\}$.

Aide : reconsidérer le même problème relativement à sa fermeture

$$\overline{S}:=\{(x,y,z)\in\mathbb{R}^3\mid x+y+z=12,x,y,z\geqslant 0\},$$

qui est compacte mais plus une variété, pour prouver que le point que vous obtenez par la méthode de Lagrange est bien un maximum. corrVariete0003

Correction of the exercise 154

Nous avons

$$L(x,y,z,\lambda)=xy^2z^3+\lambda(x+y+z-12). \tag{97.516}$$

Les équations à résoudre sont

$$\begin{cases} y^2z^3+\lambda=0\\ 2xyz^3+\lambda=0\\ 3xy^2z^2+\lambda=0\\ x+y+z-12=0, \end{cases} \tag{97.517}$$

avec la contrainte que x, y et z doivent être plus grand ou égaux à zéro. La solution est $x=2$, $y=4$ et $z=6$. Ce point est un candidat, et il faut savoir si il est réellement minimum ou maximum, local ou global.

Si nous relaxons la condition $x,y,z>0$, et que nous demandons seulement $x,y,z\geqslant 0$, alors le domaine est compact et la fonction a certainement un minimum et un maximum local. Sur le bord du nouveau domaine (c'est-à-dire là où une des coordonnées est nulle), la fonction vaut zéro, tandis qu'elle est strictement positive ailleurs. Le minimum global est donc atteint sur le bord. Il existe alors un maximum global à l'intérieur du domaine. Étant donné que $(2,4,6)$ est le seul candidat, l'existence d'un maximum prouve que ce point est maximum global.

Exercise 155 exoVariete0004

Trouver les extrémums de la fonction $d : \mathbb{R}^6 \to \mathbb{R} : (v, v') \mapsto \|v - v'\|^2$, c'est-à-dire

$$(x, y, z, x', y', z') \mapsto (x - x')^2 + (y - y')^2 + (z - z')^2 \tag{97.518}$$

relativement au plan $S := \{(x, y, z, x', y', z') \in R^6 \mid x = y = z, x' = 1, y' = 0\}$.

Aide : reconsidérer le même problème relativement à l'ensemble $S \cap C$ où

$$C := \{(x, y, z, x', y', z') \in R^6 \mid (x - x')^2 + (y - y')^2 + (z - z')^2 < 10\} \tag{97.519}$$

qui est borné et à sa fermeture $\overline{S \cap C}$ qui est compact mais n'est plus une variété, pour prouver que le point que vous obtenez par la méthode de Lagrange est bien un minimum. corrVariete0004

Correction of the exercise 155

La portion d'espace est donnée par les fonctions

$$\begin{aligned} G_1 &= x - y \\ G_2 &= x - z \\ G_3 &= x' - 1 \\ G_4 &= y'. \end{aligned} \tag{97.520}$$

Les gradients sont

$$\begin{aligned} \nabla G_1 &= (1, -1, 0, 0, 0, 0) \\ \nabla G_2 &= (1, 0, -1, 0, 0, 0) \\ \nabla G_3 &= (0, 0, 0, 1, 0, 0) \\ \nabla G_4 &= (0, 0, 0, 0, 1, 0), \end{aligned} \tag{97.521}$$

qui sont linéairement indépendants. Les candidats seront donc uniquement les points tels que $\nabla L = 0$ où

$$\begin{aligned} &L(x, y, z, x', y', z', \lambda_1, \lambda_2, \lambda_3) \\ &\quad = (x - x')^2 + (y - y')^2 + (z - z')^2 + \lambda_1(x - y) + \lambda_2(x - y) + \lambda_3(x' - 1) + \lambda_4 y'. \end{aligned} \tag{97.522}$$

La résolution du système d'équations est donnée par l'unique point

$$(x, y, z, x', y', z') = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 1, 0, \frac{1}{2}), \tag{97.523}$$

et la fonction vaut $\frac{1}{2}$ en ce point.

Si au lieu de regarder la fonction sur tout le domaine S, nous ne la regardons seulement sur la fermeture de l'intersection avec l'ensemble

$$C = \{(x, y, z, x', y', z') \text{ tel que } (x - x')^2 + (y - y')^2 + (z - z')^2 \leqslant 10\}, \tag{97.524}$$

alors il y a un minimum et un maximum. La fonction vaut 10 sur le bord de $\overline{S \cap C}$, tandis qu'elle est plus petite que 10 à l'intérieur. Le minimum n'est donc pas sur le bord. Il y a un seule candidat à être minimum à l'intérieur, c'est le point trouvé.

Ce point est donc minimum global de f relativement à l'ensemble $\overline{S \cap C}$, et même relativement à $S \cap C$. Le fait que ce soit un minimum local par rapport à S découle maintenant du lemme 17.75.

Exercise 156 exoVariete0005

Trouver les extrémums de la fonction

$$\begin{aligned} f\colon \mathbb{R}^3 &\to \mathbb{R} \\ (x, y, z) &\mapsto x^2 + y^2 + z^2 \end{aligned} \tag{97.525}$$

relativement à la surface $S := \{z^2 = x^2 + y^2 + xy + x + y + z + 1\}$. corrVariete0005

Correction of the exercise 156

Attention : exercice difficile [15] !

Le lagrangien du problème est

$$L(x,y,z,\lambda) = x^2 + y^2 + z^2 + \lambda(x^2 + y^2 + xy + x + y + z - 1 - z^2), \tag{97.526}$$

et les équations à résoudre sont

$$\begin{cases} 2x + 2\lambda x + \lambda y + \lambda = 0 \\ 2y + 2\lambda y + \lambda x + \lambda = 0 \\ 2z + \lambda - 2\lambda z = 0 \\ x^2 + y^2 + xy + x + y + z - 1 - z^2 = 0. \end{cases} \tag{97.527}$$

Il est possible de résoudre les trois premières équations pour x, y et z :

$$x = -\frac{\lambda}{2\lambda + 2}, \quad y = -\frac{\lambda}{3\lambda + 2}, \quad z = \frac{\lambda}{2\lambda - 2}. \tag{97.528}$$

En remettant dans la dernière équation, cela donne une équation épouvantable pour λ :

$$33\lambda^4 - 22\lambda^3 - 44\lambda^2 - 8\lambda + 16 = 0. \tag{97.529}$$

En réfléchissant beaucoup, on peut prouver que les solutions sont sont comprises entre 1/2 et 3/2. Cela fait que $|x| < \frac{3}{7}$, $|y| < \frac{3}{7}$ et $|z| < \frac{3}{2}$, et donc que

$$f(x,y,z) < \frac{513}{196} \simeq 2.61. \tag{97.530}$$

97.15 Intégrales curvilignes

Exercise 157 exoVariete0006

Calculez la longueur des arcs de courbe suivants :

(1) $y = \ln(1 - x^2)$ pour $0 \leqslant x \leqslant \frac{1}{2}$

(2) $y = x^{3/2}$ pour $0 \leqslant x \leqslant 5$

(3) $y = 1 - \ln(\cos x)$ pour $0 \leqslant x \leqslant \frac{\pi}{4}$

corrVariete0006

Correction of the exercise 157

(1) Il s'agit d'appliquer la formule (21.62). Il faut donc intégrer

$$L = \int_0^{1/2} \sqrt{1 + \left(\frac{2x}{1 - x^2}\right)^2} = \int_0^{1/2} \frac{x^2 + 1}{x^1 - 1}. \tag{97.531}$$

La division euclidienne de $x^2 + 1$ par $x^2 - 1$ donne 1 comme quotient et 2 comme reste, c'est-à-dire que $x^2 + 1 = (x^2 - 1) + 2$. Cela suggère de découper la fraction comme

$$\frac{x^2 - 1}{x^2 - 1} + \frac{2}{x^2 - 1} = 1 + \frac{2}{x^2 - 1}. \tag{97.532}$$

Ces deux fractions sont dans le formulaire. Après calcul, la réponse est $\ln(\frac{3}{4}) - \frac{11}{6}$.

(2)

$$L = \int_0^5 \sqrt{1 + \frac{9}{4}x} = \frac{335}{27}. \tag{97.533}$$

Notez que pour effectuer l'intégrale, le changement de variable $u = 1 + 9x/4$ est conseillé.

15. Si vous trouvez une méthode plus simple, merci de me le faire savoir

(3) L'intégrale sur laquelle on tombe est $\int \frac{1}{\cos(x)}dx$. Il faut savoir[16] que $1/sin(x) = \text{cosec}(x)$. Cette intégrale est donc dans le formulaire.

Exercise 158 exoVariete0007

Calculez $\int_\gamma xdy+ydx$ où γ est l'arc de cubique déterminé par $y = x^3+x^2+x+1$ avec $0 \leqslant x \leqslant 1$. corrVariete0007

Correction of the exercise 158

Comme d'habitude quand on donne un chemin par une équation $y = y(x)$, le chemin est paramétré par $t \mapsto (t, y(t))$. Ici nous avons

$$\gamma(t) = (t, t^3 + t^2 + t + 1) \tag{97.534}$$

avec $0 \leqslant t \leqslant 1$. Le vecteur tangent à ce chemin est donné par

$$\gamma'(t) = (1, 3t^2 + 2t + 1), \tag{97.535}$$

et il faut calculer

$$\int_0^5 \omega_{\gamma(t)}\big(\gamma'(t)\big) = \int_0^5 tdy \begin{pmatrix} 1 \\ 3t^2 + 2t + 1 \end{pmatrix} + (t^3 + t^2 + t + 1)dx \begin{pmatrix} 1 \\ 3t^2 + 2t + 1 \end{pmatrix}. \tag{97.536}$$

Maintenant, il faut se rendre compte que $dy \begin{pmatrix} 1 \\ 3t^2 + 2t + 1 \end{pmatrix} = 1$ et que $dx \begin{pmatrix} 1 \\ 3t^2 + 2t + 1 \end{pmatrix}$. L'intégrale à calculer est finalement

$$\int_0^5 \big[t(3t^2 + 2t + 1) + (t^3 + t^2 + t + 1)1\big] = 780. \tag{97.537}$$

Exercise 159 exoVariete0008

Calculez $\int_\gamma 2xydx+x^2dy$ le long des différents chemins suivants joignant l'origine $(0, 0)$ au point $a = (2, 1)$

(1) la droite $y = \frac{x}{2}$
(2) la parabole $y = \frac{x^2}{4}$
(3) la parabole $y^2 = \frac{x}{2}$
(4) la ligne brisée Oba où $b = (0, 1)$

Donnez une interprétation de ces résultats. Pour ce faire, sachez qu'en physique, il y a un adage qui dit que *le travail d'une force conservative ne dépend pas du chemin suivit*. En mathématique, il y a un théorème qui dit la même chose. Citer ce théorème et prouver que la forme $\omega = 2xydx + x^2dy$ vérifie les hypothèses de ce théorème. corrVariete0008

Correction of the exercise 159

(1) Le chemin proposé est $\gamma(t) = (t, t/2)$, et l'intégrale est

$$\int_0^2 t^2dx \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} + t^2dy \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \int_0^2 \left(t^2 + \frac{t^2}{4}\right) = 4. \tag{97.538}$$

(2) Le nouveau chemin est

$$\gamma(t) = (t, \frac{t^2}{4}), \tag{97.539}$$

et le chemin dérivé est $\gamma'(t) = (1, t/2)$. L'intégrale est

$$\int_0^2 \omega_{\gamma(t)} \begin{pmatrix} 1 \\ t/2 \end{pmatrix} = \int_0^2 (\frac{t^3}{2} + \frac{t^3}{2}) = 4 \tag{97.540}$$

16. oui oui, c'est une notation qui s'utilise de temps en temps.

(3) Afin de ne pas s'ennuyer avec des racines carrés, il est plus facile de paramétrer le chemin en regardant x comme fonction de y, c'est-à-dire

$$\gamma(t) = (2t^2, t), \tag{97.541}$$

avec t parcourant l'intervalle $[0, 1]$. Nous trouvons

$$\omega_{\gamma(t)} = 4t^3 dx + 4t^4 dy. \tag{97.542}$$

L'intégrale est alors

$$\int_0^1 \omega_{\gamma(t)} \begin{pmatrix} 4t \\ 1 \end{pmatrix} = \int_0^1 16t^4 + 4t^4 = 4. \tag{97.543}$$

(4) Le premier chemin est vertical et est $\gamma_1(t) = (0, t)$, tandis que le second est horizontal à la hauteur 1 et est $\gamma_2(t) = (t, 1)$. Pour le premier, $t \in [0, 1]$, tandis que pour le second, $t \in [0, 2]$. Nous avons $\omega_{\gamma_1(t)} = 0$, donc le premier chemin ne compte pas. Il reste à intégrer

$$\omega_{\gamma_2(t)} = 2tdx + t^2 dy, \tag{97.544}$$

donc

$$\int_0^2 (2tdx + t^2 dy) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 4. \tag{97.545}$$

Toutes les intégrales valent 4. La raison est que la forme ω est exacte,et que l'intégrale d'une forme différentielle exacte entre deux points ne dépend pas du chemin suivit. Afin de prouver que ω est exacte, il faut trouver une fonction $f(x, y)$ pour laquelle $\omega = df$. Il faut donc que

$$\frac{\partial f}{\partial x} = 2xy \quad \text{et} \quad \frac{\partial f}{\partial y} = x^2. \tag{97.546}$$

La première équation montre, par simple intégration par rapport à x, que $f(x, y) = x^2 y + c(y)$. La fonction $c(y)$ est la constante d'intégration ; elle est constante par rapport à x, mais peut dépendre de y. La seconde équation fixe cette constante en fonction de y. Nous avons que $c(y)$ doit être une constante. Donc

$$\omega = d(x^2 y + c) \tag{97.547}$$

pour n'importe quel $c \in \mathbb{R}$. Un simple calcul montre que $f(2, 1) - f(0, 0) = 4$.

Exercise 160 exoVariete0009

Calculez $\int_\gamma ydx + xdy + z^2 dz$ où γ est une hélice de pas 1 tracée sur le cylindre $x^2 + y^2 = 1$. corrVariete0009

Correction of the exercise 160

Pour l'hélice, il faut quelque chose qui tourne dans le plan xy, et qui monte dans la direction z. Le paramétrage

$$\gamma(t) = \big(\cos(t), \sin(t), t/2\pi\big) \tag{97.548}$$

est parfaite. Nous avons

$$\omega_{\gamma(t)} = \sin(t)dx + \cos(t)dt + \frac{t^2}{4\pi^2} dz, \tag{97.549}$$

et donc

$$\int_\gamma \omega = \int_0^{2\pi} \big(-\sin(t)^2 + \cos(t)^2 + \frac{t^2}{8\pi^3}\big) = \frac{1}{3}. \tag{97.550}$$

Notons que la forme ω est exacte. En effet, il faut une fonction f telle que

$$\begin{aligned} \frac{\partial f}{\partial x} &= y, \\ \frac{\partial f}{\partial y} &= x, \\ \frac{\partial f}{\partial z} &= z^2. \end{aligned} \tag{97.551}$$

La dernière nous dit que $f(x,y,z) = \frac{z^3}{3} + c(x,y)$. La seconde équation dit donc que $c(x,y) = xy + h(y)$, et enfin la première équation montre que $h = 0$. En résumé, $\omega = df$ avec

$$f(x,y,z) = xy + \frac{z^3}{3}. \tag{97.552}$$

Maintenant, on a

$$\int_\gamma \omega = (f \circ \gamma)(2\pi) - (f \circ \gamma)(0) = f(1,0,1) - f(1,0,0) = \frac{1}{3}. \tag{97.553}$$

Exercise 161 exoVariete0010

Calculer les intégrales suivantes

(1) $\int_\gamma 1ds$ où γ est l'arc de cycloïde donné par

$$\gamma(t) = (a\,(t - \sin(t)), a\,(1 - \cos(t))) \qquad 0 \leqslant t \leqslant 2\pi.$$

(2) $\int_\gamma 1ds$ où γ est l'astroïde d'équation $x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$ et où a est un réel positif.

(3)

$$\int_\gamma (x+y)ds \tag{97.554}$$

où γ est le contour du triangle ABO de sommets

$$A = (1,0) \quad B = (0,1) \quad O = (0,0).$$

corrVariete0010

Correction of the exercise 161

Il faut diviser le chemin en trois morceaux. Attention à l'ordre : nous intégrons sur le chemin parcourant le triangle dans le sens ABC et non dans le sens inverse ACB.

Le chemin σ_1 joint A à B : $\sigma_1(t) = (1-t)A + tB = \begin{pmatrix} 1-t \\ t \end{pmatrix}$. L'intégrale à effectuer est donc

$$\int_{\sigma_1} (x+y)d\sigma_1 = \int_0^1 (1-t) + t\,dt = 1. \tag{97.555}$$

Le chemin qui joint B à C est $\sigma_2(t) = (0, 1-t)$ et l'intégrale le long de ce chemin vaut

$$\int_{\sigma_2} (x+y)d\sigma_2 = \int_0^1 (1-t)dt = \frac{1}{2}. \tag{97.556}$$

La troisième intégrale est sur $\sigma_3(t) = (t,0)$:

$$\int_{\sigma_3} (x+y)d\sigma_3 = \int_0^1 t\,dt = \frac{1}{2}. \tag{97.557}$$

L'intégrale de $x+y$ le long de tout le triangle vaut donc $1 + \frac{1}{2} + \frac{1}{2} = 2$.

Exercise 162 exoVariete0011

Calculer les intégrales suivantes :

(1)

$$\int_\gamma y^2dx + x^2dy \tag{97.558}$$

où γ est le cercle de rayon 1 centré en l'origine et parcouru dans le sens horlogique.

(2)

$$\int_\gamma G \tag{97.559}$$

où G est le champs de vecteurs qui vaut (y^2, x^2) en (x,y) et où γ est le cercle de rayon 1 centré en l'origine et parcouru dans le sens trigonométrique.

(3)
$$\int_\gamma (y-z)dx + (z-y)dy + (x-y)dz \tag{97.560}$$
où γ est l'arc d'hélice
$$\{(a\cos(t), a\sin(t), bt) \text{ tel que } 0 \leqslant t \leqslant 2\pi\}$$
et où a et b sont des réels positifs.

(4)
$$\int_\gamma (y+z)dx + (z+x)dy + (x+y)dz \tag{97.561}$$
où γ est le cercle à l'intersection de la sphère unité de $\mathbb{R}^3$ et du plan d'équation $x+y+z=0$, parcouru dans le sens indiqué par le vecteur $(1,1,-2)$ au point $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0)$.

corrVariete0011

Correction of the exercise 162

(1)

(2)

(3)

(4) Cet exercice est intéressant parce qu'il est difficile de trouver une bonne expression pour le chemin. Heureusement, cela n'est pas nécessaire. En effet, dans le plan, $P \equiv x+y+z=0$, nous avons
$$\omega|_P(x,y,z) = -xdx - ydy - zdz. \tag{97.562}$$
Si maintenant nous considérons la forme $\omega' = -xdx - ydy - zdz$, il est facile de voir qu'elle est exacte et que $\omega' = df$ avec
$$f(x,y,z) = -\frac{x^2}{2} - \frac{y^2}{2} - \frac{z^2}{2}. \tag{97.563}$$
Étant donné que le chemin γ considéré est contenu dans le plan P, nous avons
$$\int_\gamma \omega = \int_\gamma \omega' = 0 \tag{97.564}$$
parce que l'intégrale d'une forme exacte sur un chemin fermé est nulle.

97.16 Intégrales de surface, Stokes et Green

Exercise 163 exoVariete0012

Quelques calculs d'aires.

(1) Calculez la surface de la sphère de rayon R centrée à l'origine $(0,0,0)$.

(2) Calculez $\iint_{S\cap C} 1 d\sigma$ où S est la sphère de rayon R centrée à l'origine $(0,0,0)$ et C le cylindre circulaire de diamètre $((0,0,0),(R,0,0))$ parallèle à $0z$.

(3) Calculez $\iint_S \sqrt{x^2+y^2} d\sigma$ où S est la surface latérale du cône $S := \{(x,y,z) \mid x^2+y^2-z^2 = 0, 0 < z < b\}$ et où b est un réel positif.

(4) Calculez l'aire de la calotte sphérique découpée par le cône $z^2 = x^2+y^2$ dans la sphère unité.

(5) Calculez l'aire de la portion de surface cylindrique $x^2+y^2=1$ limitée par l'hélice de pas un, la droite $x=0, y=1$ et la surface $z=0$.

corrVariete0012

Correction of the exercise 163

(1) La sphère est une variété de dimension 2 dans $\mathbb{R}^3$. Nous avons donc besoin de cartes pour intégrer dessus. Les coordonnées sphériques sont une carte qui recouvrent toute la sphère sauf un bout de mesure nulle (voir page 486). Nous allons donc considérer

$$\begin{aligned} F\colon \mathbb{R}^+\backslash\{0\}\times]0,2\pi[\times]0,\pi[&\to \mathbb{R}^3 \\ (r,\theta,\varphi) &\mapsto \begin{pmatrix} R\cos(\theta)\sin(\varphi) \\ R\sin(\theta)\sin(\varphi) \\ R\cos\varphi \end{pmatrix}. \end{aligned} \tag{97.565}$$

Les vecteurs tangents à ce paramétrage sont

$$\frac{\partial F}{\partial\theta} = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\cos\theta\sin\varphi \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial\varphi} = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\sin\theta\cos\varphi \\ -R\sin\varphi \end{pmatrix}, \tag{97.566}$$

et l'élément de surface est la norme de

$$\begin{aligned} \partial_\theta F\times\partial_\varphi F &= \begin{vmatrix} e_1 & e_2 & e_3 \\ -R\sin\theta\sin\varphi & R\cos\theta\sin\varphi & 0 \\ R\cos\theta\cos\varphi & R\sin\theta\cos\varphi & -R\sin\varphi \end{vmatrix} \\ &= (-R^2\cos\theta\sin^2\varphi)e_1 - (R^2\sin\theta\sin^2\varphi)e_2 + (-R^2\sin\varphi\cos\varphi)e_3. \end{aligned} \tag{97.567}$$

La norme du tout vaut

$$d\sigma(\theta,\varphi) = \|\partial_\theta F\times\partial_\varphi F\| = R^2\sin\varphi. \tag{97.568}$$

Donc l'intégrale à calculer est

$$\int_0^{2\pi}\int_0^{\pi} R^2\sin\varphi\, d\varphi d\theta = 2\pi R^2[-\cos\varphi]_0^{\pi} = 4\pi R^2. \tag{97.569}$$

(2) Un point (x,y) est sur le cercle lorsque $(x-\frac{R}{2})^2+y^2=\frac{R^2}{4}$. En passant aux coordonnées polaires, l'intérieur du cercle est

$$r^2 - Rr\cos\theta < 0, \tag{97.570}$$

que l'on peut simplifier par r parce que r est toujours strictement positif. Le cylindre (plein) dont on parle est donc donné par

$$r < R\cos\theta \tag{97.571}$$

en coordonnées cylindriques. En coordonnées cylindriques, la sphère s'écrit

$$r^2+z^2=R^2. \tag{97.572}$$

Ce sur quoi nous intégrons est le morceau de sphère au dessus (et en dessous) du cercle. Il est donc naturel d'utiliser ce dernier pour paramétriser la surface sur laquelle on veut intégrer. La carte est donc

$$F(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ \sqrt{R^2-r^2} \end{pmatrix} \tag{97.573}$$

EqParmsCylSphere

à prendre sur l'ouvert $r < R\cos\theta$. Un peu de calcul montre que

$$\|\partial_r F\times\partial_\theta F\| = \frac{rR}{\sqrt{R^2-r^2}}. \tag{97.574}$$

Il faut intégrer cela avec $\theta\in]-\pi/2,\pi/2[$ et $r\in]0,R\cos\theta[$:

$$I = \int_{-\pi/2}^{\pi/2}\int_0^{R\cos\theta}\frac{rR}{\sqrt{R^2-r^2}}drd\theta = R^2(\pi-2). \tag{97.575}$$

Ce résultat doit encore être multiplié par deux pour tenir compte de la partie de surface en dessous. Le résultat est donc

$$2R^2(\pi-2). \tag{97.576}$$

(3) Une carte pour le cône (à part le sommet) est donnée par le cercle sur lequel il se projette :

$$F(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r \end{pmatrix}, \tag{97.577}$$

avec $r\colon 0 \to b$ et $\theta\colon 0 \to 2\pi$. Nous avons

$$\frac{\partial F}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial \theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}, \tag{97.578}$$

et

$$d\sigma(r,\theta) = r\sqrt{2}. \tag{97.579}$$

Nous devons donc calculer l'intégrale

$$\int_0^b \int_0^{2\pi} r d\sigma(r,\theta) = \int_0^b \int_0^{2\pi} r^2\sqrt{2} = \pi\sqrt{2}\frac{b^3}{2}. \tag{97.580}$$

(4) Encore une fois, le cône va vers le haut et vers le bas. Il y aura donc lieu de multiplier le résultat par deux. Cherchons à quelle hauteur le cône coupe la sphère. Lorsque $y = 0$, nous avons les équations $z = x$ et $x^2 + z^2 = 1$, ce qui donne la hauteur $z = 1/\sqrt{2}$. Nous devons donc intégrer sur le morceau de sphère qui flotte au dessus du cercle de rayon $1/\sqrt{2}$. Nous reprenons le paramétrage en coordonnées cylindrique donnée par (97.573), dont nous connaissons déjà l'élément de surface, et nous devons simplement calculer

$$\int_0^{1/\sqrt{2}} \int_0^{2\pi} \frac{r}{\sqrt{1-r^2}} d\theta dr = 2\pi(1 - \frac{1}{\sqrt{2}}). \tag{97.581}$$

Après multiplication par deux, nous avons la réponse

$$S = 2\pi(2 - \sqrt{2}). \tag{97.582}$$

(5) Une bonne carte pour le cylindre est donnée par

$$F(\theta, z) = \begin{pmatrix} \cos\theta \\ \sin\theta \\ z \end{pmatrix} \tag{97.583}$$

Nous voyons avec un tout petit peu de calcul que $d\sigma(r,\theta) = 1$, de façon que la surface demandée soit

$$\int_0^{2\pi} \int_0^{\theta/2\pi} 1 dz d\theta = \pi. \tag{97.584}$$

Cela est exactement la surface du triangle de hauteur 1 et de base donnée par la circonférence de la base du cylindre, un peu comme si le cylindre n'était pas vraiment courbé, ce qui est confirmé par le fait que l'élément de surface est la constante 1.

Exercise 164 exoVariete0013

Calculer les flux des champs de vecteurs suivants à travers les surfaces suivantes :

(1) $G = (x^2, y^2, z^2)$ où S est la surface extérieure du cube $0 \leqslant x, y, z \leqslant a$.

(2) $G = (x+y, y+z, z+x)$ où S est la surface totale du cylindre $x^2 + y^2 = 1$ limité par les plans $z = 0$ et $z = 1$.

(3) $G = (y^2, x^2, z^2)$ où S est la surface du paraboloïde $x^2 + y^2 = z$ limité par le plan $z = 1$. Indice : la réponse n'est pas $2\pi/3$.

corrVariete0013

Correction of the exercise 164

(1) Soient C le cube et B le cube plein dont C est le bord. Par le théorème de la divergence,

$$\iint_C G \cdot dS = \iiint_B \nabla \cdot G$$

or la divergence de G se calcule aisément puisque

$$\nabla \cdot G = 2(x+y+z)$$

et donc l'intégrale recherchée est

$$\int_0^a \int_0^a \int_0^a 2(x+y+z)dxdydz = 3a^4.$$

(2) La surface totale S borde un cylindre plein C. Le théorème de la divergence s'applique, et l'intégrale recherchée est

$$\iint_S G \cdot dS = \iiint_C \nabla \cdot G = \iiint_C 3$$

ce qui revient à calculer le triple du volume du cylindre. Ce volume vaut π, et la réponse attendue est donc 3π.

(3) Ici la surface S du paraboloïde ne délimite pas un volume fermé. Néanmoins, il semble plus facile de calculer le flux au travers du « couvercle » C d'équations $z = 1, x^2 + y^2 \leqslant 1$ que le flux au travers de S. On va donc appliquer le théorème de la divergence sous la forme :

$$\iint_S G \cdot dS + \iint_C G \cdot dS = \iint_{S \cup C} G \cdot dS = \iiint_V \nabla \cdot G$$

où V est le volume délimité par la surface S et son couvercle C.

La divergence de G vaut $2z$, et donc en passant en coordonnées cylindriques on trouve

$$\iiint_V \nabla \cdot G = \int_0^{2\pi} \int_0^1 \int_0^{\sqrt{z}} 2z\rho d\rho dz d\theta = \frac{2\pi}{3}.$$

Sur le couvercle, $G = (y^2, x^2, 1)$ et le couvercle est paramétrisé par

$$\begin{cases} x = \rho\cos(t) \\ y = \rho\sin(t) \\ z = 1 \end{cases}$$

avec pour vecteur normal unitaire extérieur $(0,0,1)$. Donc le flux au travers de C vaut

$$\iint_C \big\langle (y^2, x^2, 1), (0,0,1) \big\rangle = \iint_C 1 = \int_0^{2\pi} \int_0^1 \rho d\rho d\theta = \pi$$

ce qui n'est pas insensé puisqu'on a intégré la fonction 1 sur un disque, donc on retrouve l'aire de ce disque.

En conclusion, l'intégrale recherchée au départ vaut

$$\iint_S G \cdot dS = \iiint_V \nabla \cdot G - \iint_C G \cdot dS = \frac{2\pi}{3} - \pi = -\frac{\pi}{3}$$

Exercise 165 exoVariete0014

Soit $D \subset \mathbb{R}^3$ un ouvert borné dont le bord ∂D est une variété de dimension deux. Soit ν un champ de vecteurs unitaire normal à ∂D en tout points. Soit $G : \partial D \to \mathbb{R}^3$ un champ de vecteurs constant et soit $c : \partial D \to \mathbb{R}$ la fonction définie sur ∂D qui associe à v le cosinus de l'angle entre $G(v)$ et $\nu(v)$. Calculez $\int_{\partial D} c(v)d\sigma$. corrVariete0014

Correction of the exercise 165

Soit D ouvert borné dans l'espace, dont le bord est une variété de dimension 2 ayant un champ de vecteurs unitaire normal ν. Soit G un champ de vecteurs constant et $c : \partial D \to \mathbb{R}$ la fonction définie sur ∂D qui associe au point $v \in \partial D$ le cosinus de l'angle entre $G(v)$ et $\nu(v)$.

Par définition du produit scalaire, on sait que

$$c(v) = \frac{\langle G(v), \nu(v)\rangle}{\|G(v)\| \, \|\nu(v)\|} = \frac{\langle G(v), \nu(v)\rangle}{\|G(v)\|}$$

où $\|G(v)\|$ est une constante (puisque le champ G est constant). Dès lors nous avons

$$\iint_{\partial D} c(v) = \frac{1}{\|G\|} \iint_{\partial D} \langle G(v), \nu(v)\rangle$$

où la dernière intégrale est, par définition, le flux de G au travers de ∂D. Comme G est constant, sa divergence vaut 0, et l'intégrale est donc nulle par le théorème de la divergence.

Exercise 166 exoVariete0015

Calculer le flux du champ de vecteurs $G(x, y, z) = (y,\ x\ , yz)$ à travers le cornet de glace limité par l'ellipsoïde

$$\frac{x^2}{16} + \frac{y^2}{9} + \frac{z^2}{16} = 1 \tag{97.585}$$

lorsque $y \geqslant 0$ et par le cône d'équation $x^2 + z^2 = (y + 4)^2$ lorsque $y < 0$. Faire un dessin peut vous aider. corrVariete0015

Correction of the exercise 166

Le cornet de glace S délimite un volume fermé V constitué d'un demi-ellipsoïde E et d'un bout de cône C. Notons que les deux bouts se recollent bien en $y = 0$ parce que sur ce plan, l'ellipsoïde et le cône ont tous les deux pour équation $x^2 + y^2 = 16$.

Le théorème de la divergence s'applique, et l'intégrale recherchée vaut

$$\iint_S G \cdot dS = \iiint_E y + \iiint_C y$$

puisque $\nabla \cdot G = y$.

Pour calculer l'intégrale sur le demi-ellipsoïde, prenons des coordonnées sphériques un peu modifiées :

$$\begin{cases} x = 4r\cos\phi\sin\theta \\ y = 3r\cos\theta \\ z = 4r\sin\phi\sin\theta \end{cases}$$

dont le jacobien vaut $48r^2\sin\theta$. L'intégrale vaut [17]

$$\iiint_E y = \int_0^{2\pi}\int_0^{\frac{\pi}{2}}\int_0^1 144r^3\sin\theta\cos\theta dr d\theta d\phi = 144\frac{1}{4}\frac{1}{2}2\pi = 36\pi.$$

Afin de calculer l'intégrale sur le cône, nous utilisons les coordonnées cylindriques :

$$\begin{cases} x = \rho\cos t \\ y = \tilde{y} - 4 \\ z = \rho\sin t \end{cases}$$

17. Pouvez-vous justifier le fait que le rayon va de 0 à 1 ?

dont le jacobien vaut ρ. L'intégrale vaut

$$\iiint_C y = \int_0^{2\pi}\int_0^4\int_0^{\tilde{y}} \tilde{y} - 4 d\rho d\tilde{y} dt = \frac{-64\pi}{3}$$

L'intégrale recherchée au départ vaut donc $\frac{44\pi}{3}$.

Exercise 167 exoVariete0016

Calculer.

(1) $\oint_\gamma 2(x^2+y^2)dx + (x+y)^2 dy$ où γ est le périmètre du triangle dont les sommets sont les points $a = (1,1)$, $b = (2,2)$ et $c = (1,2)$.

(2) $\oint_\gamma xy^2 dx - x^2 y dy$ où γ est la circonférence $x^2 + y^2 = R^2$.

(3) $\oint_\gamma dx + x dy$ où γ est le cycle formé par $y = x^2$ et $y^2 = x$.

(4) l'aire de l'ellipse $x = a\cos\theta$, $y = b\sin\theta$.

(5) $\oint_\gamma -x^2 y dx + xy^2 dy$ où γ est le cercle de rayon 1 centré en l'origine $(0,0)$ et parcouru dans le sens horlogique.

corrVariete0016

Correction of the exercise 167

(1) Lorsqu'on ne précise rien, c'est qu'on demande que le chemin est dans le « bon » sens. Il n'y a donc pas de problèmes de sens, on peut appliquer le théorème de Green les yeux fermés. Tout ce que nous avons à faire est d'intégrer la fonction

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \tag{97.586}$$

où $P(x,y) = 2(x^2+y^2)$ et $Q(x,y) = (x+y)^2$ sur le triangle. Cela se fait ainsi :

$$2\int_1^2 dy \int_1^y dx (x-y) = -\frac{1}{3}. \tag{97.587}$$

(2) Juste pour rappel, le bon sens est le sens trigonométrique. Si D est le disque, l'intégrale à faire est

$$-4\int_D xy\, dx\, dy = \int_0^{2\pi}\int_0^R r^2 \cos\theta \sin\theta dr d\theta = 0. \tag{97.588}$$

(3) La forme est $\omega = dx + xdy$, donc $P = 1$ et $Q = x$, ce qui donne que la fonction à intégrer sur la surface est 1. L'intégrale est donc

$$\int_0^2 \int_{x^2}^{\sqrt{x}} dy dx = \frac{1}{3}. \tag{97.589}$$

(4) Nous avons repéré au point précédent que la forme $\omega = dx + xdy$ donnait 1 comme fonction en prenant le passage à Green. Ici, nous devons intégrer la fonction 1 sur l'ellipse. Nous allons donc intégrer la forme $\omega = dx + xdy$ sur le contour. Le chemin (dans le bon sens !) qui paramètre l'ellipse est donné par

$$\gamma(t) = \begin{pmatrix} a\cos(t) \\ b\sin(t) \end{pmatrix} \tag{97.590}$$

où $t\colon 0 \to 2\pi$. La dérivée du chemin (son vecteur tangent) est donné par

$$\gamma'(t) = \begin{pmatrix} -a\sin(t) \\ b\cos(t) \end{pmatrix}. \tag{97.591}$$

Nous devons donc intégrer

$$\begin{aligned} \int_0^{2\pi} (dx + a\cos(t)dy) \begin{pmatrix} -a\sin(t) \\ b\cos(t) \end{pmatrix} &= \int_0^{2\pi} \big(-a\sin(t) + ab\cos(t)\big) dt \\ &= ab\int_0^{2\pi} \cos^2(t) dt \\ &= \pi ab. \end{aligned} \tag{97.592}$$

(5) Cette fois, le sens de parcours est prescrit, et il est le sens inverse de celui où Green fonctionne. Nous devons donc ajouter un signe :

$$\int_\gamma -x^2ydx + xy^2dy = -\int_D (x^2+y^2)dxdy = -\int_0^1\int_0^{2\pi} r^3 d\theta dr = -\frac{\pi}{2}. \tag{97.593}$$

Exercise 168 exoVariete0017

Calculer.

(1) $\oint_\gamma y^2dx + z^2dy + x^2dz$ où γ est le périmètre du triangle de sommets $(1,0,0)$, $(0,1,0)$, $(0,0,1)$,

(2) $\oint_\gamma y^2dx + x^2dy + z^2dz$ où γ est la circonférence $x^2+y^2=1, z=0$. On prendra comme surfaces successivement le disque dans le plan $z=0$ et la sphère unité,

(3) la circulation du champ de vecteurs $G(-x^3, y^3, 2x+z^2)$ le long du cercle

$$\begin{cases} z=0 \\ x^2+y^2=1 \end{cases}$$

directement et ensuite par la formule de Stokes.

(4)

$$\oint_\gamma (y+z)dx + (z+x)dy + (x+y)dz$$

où γ est le cercle $\gamma := \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2=1, x+y+z=0\}$ parcouru dans le sens indiqué par le vecteur $(1,1,-2)$ au point $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$,

(5)

$$\oint_\gamma ydx + zdy + xdz \tag{97.594}$$

où γ est l'intersection de $x^2+y^2+z^2=a^2$ et $y=0$.

corrVariete0017

Correction of the exercise 168

(1) Le chemin γ est dans le cadre d'application du théorème de Stokes puisqu'il est le bord d'un triangle T, dès lors

$$\int_\gamma y^2dx + z^2dy + x^2dz = \iint_T \nabla \times G \cdot dS$$

où $G = (y^2, z^2, x^2)$. Le rotationnel vaut

$$\nabla \times G = \begin{vmatrix} \vec{e_x} & \vec{e_y} & \vec{e_z} \\ \partial_x & \partial_y & \partial_z \\ y^2 & z^2 & x^2 \end{vmatrix} = (-2z, -2x, -2y)$$

Paramétrons le triangle via

$$F(u,v) = \begin{pmatrix} 1-u-v \\ u \\ v \end{pmatrix} \tag{97.595}$$

avec $u \in]0,1[$ et $v \in]0, 1-u[$. Le dessin de la figure 97.9 montre que le paramétrage choisie a la bonne orientation vis-à-vis de la convention prise par le théorème de Green.

Le vecteur normal à la surface, de ce paramétrage, vaut

$$\frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v} = (1,1,1)$$

et donc l'intégrale devient

$$\int_0^1 \int_0^{1-u} -2dvdu = -1$$

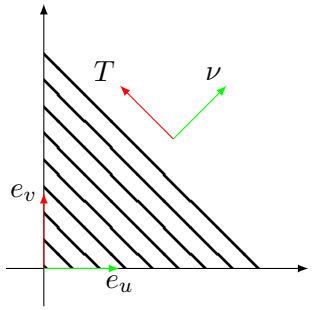


FIGURE 97.9: Question d'orientation. La base (ν, T) a la même orientation que la base $(1_u, 1_v)$. Cela fait que la carte choisie est de bonne orientation. LabelFigTriangleUV

(2) Le rotationnel de G vaut

$$\nabla \times G = 2 \begin{pmatrix} 0 \\ 0 \\ x - y \end{pmatrix}. \tag{97.596}$$

Le paramétrage du cercle, en coordonnées cylindriques, est

$$F(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix}. \tag{97.597}$$

Nous devons donc intégrer

$$\int_0^1 \int_0^{2\pi} \langle \nabla \times G, N(r, \theta) \rangle d\theta dr. \tag{97.598}$$ EqEtapeUn153

Il est vite vu que

$$N(r, \theta) = \frac{\partial F}{\partial r} \times \frac{\partial F}{\partial \theta} = \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}, \tag{97.599}$$

donc le produit scalaire dans l'intégrale (97.598) donne à intégrer $2r^2(\cos\theta - \sin\theta)$. La symétrie de la fonction $\cos\theta - \sin\theta$ fait que son intégrale sur $[0, 2\pi]$ est nulle.

Si nous voulons utiliser Stokes en considérant la demi-sphère au lieu du cercle, c'est plus compliqué. Ce que nous devons calculer est le flux du champ de vecteur $\nabla \times G$ au travers de la demi-sphère. Nous venons de voir que ce flux au travers du « couvercle » est nulle. Donc nous pouvons appliquer le théorème de la divergence et dire

$$\int_{\text{demi-sphere}} \nabla \times G + \int_{\text{couvercle}} \nabla \times G = \int_{\text{volume}} \nabla \cdot (\nabla \times G) = 0 \tag{97.600}$$

parce que la divergence d'un rotationnel est toujours nulle.

(3) Nous devons calculer

$$\int_\gamma \langle G, \dot{\gamma} \rangle \tag{97.601}$$

où

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 0 \end{pmatrix} \quad \gamma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 0 \end{pmatrix} \tag{97.602}$$

et $t\colon 0 \to 2\pi$. Le produit scalaire entre γ' et G vaut $x^3\sin(t) + y^3\cos(t)$, donc nous devons calculer

$$\int_0^{2\pi} \left[x^3 \sin(t) + y^3 \cos(t) \right] dt = 0. \tag{97.603}$$

Si nous voulons faire cela en utilisant Stokes, nous remarquons que

$$\nabla \times G = -2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{97.604}$$

tandis que le vecteur normal est toujours $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Le produit scalaire entre les deux étant toujours nul, nous retrouvons le résultat comme quoi l'intégrale demandée est nulle.

(4) Le rotationnel de ce champ de vecteurs est toujours nul.

(5) L'équation $x^2 + z^2 = a^2$, $y = 0$ est un cercle dans le plan (x, z). Nous utilisons Stokes :

$$\int_{\text{cercle}} \langle G, T \rangle = \int_{\text{disque}} \langle \nabla \times G, N \rangle d\sigma_F \tag{97.605}$$

avec le paramétrage

$$F(x, z) = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} \tag{97.606}$$

Le vecteur normal est donné par

$$N = \frac{\partial F}{\partial x} \times \frac{\partial F}{\partial z} = - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{97.607}$$

N'oubliez pas de faire des petits dessins en 3D pour montrer que le bon ordre des variables est bien (x, z). Le rotationnel du champ de vecteur est donné par

$$\nabla \times G = \nabla \times \begin{pmatrix} y \\ z \\ x \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{97.608}$$

Le produit scalaire de $\nabla \times G$ avec N est la constante 1. La réponse est donc la surface du cercle de rayon a : πa^2.

Exercise 169 exoVariete0018

Soit γ la courbe d'équations $\frac{x^2}{2} + \frac{y^2}{2} + z^2 = 1$, $x^2 + y^2 = y$, $z \geqslant 0$ où γ est parcourue dans le sens indiqué par le vecteur $(1, 0, 0)$ au point $(0, 0, 1)$. Calculez

$$\int_\gamma y^3 dx + (xy + 3xy^2) dy + z^4 dz$$

(1) directement

(2) en appliquant la formule de Stokes

corrVariete0018

Correction of the exercise 169

<+Variete0018+>

Exercise 170 exoVariete0019

Même question pour

$$\int_\gamma 2z dx + x dy + y dz$$

où γ est la courbe d'équations $x^2 + y^2 = z$, $z = y$ et est parcourue dans le sens indiqué par le vecteur $(1, 0, 0)$ au point $(0, 0, 0)$.

corrVariete0019

Correction of the exercise 170

<+Variete0019+>

Exercise 171 exoVariete0020

Considérons sur $\mathbb{R}^3$ le volume $V : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$

$$V(u,v,w) = \frac{1}{3!}\sum_{\sigma\in S_3} \epsilon(\sigma) u_{\sigma(1)} v_{\sigma(1)} w_{\sigma(1)}$$

où $\epsilon(\sigma)$ est la signature de la permutation σ. Soient $\mathbb{R} > 0$ et $0 < a < R$. Soit

$$M_a = \left\{ (x_1,x_2,x_3) \in \mathbb{R}^3 \mid \sum_{i=1}^{3} x_i^2 = R^2 \text{ et } x_3 > -a \right\}$$

(1) Démontrez que M_a est une variété dans $(\mathbb{R}^n)^3$.

(2) Démontrez que, si l'on définit pour deux vecteurs tangents à M_a au point $x \in M_a$:

$$\omega_x(A,B) = V(x,A,B)\ \forall A,B \in T_x(M_a)$$

on obtient une 2-forme différentielle ω sur M_a.

(3) Calculer $\int_{M_a} \omega$.

(4) Que vaut $\lim_{a\to R} \int_{M_a} \omega$?

(5) Retrouvez ce résultat par le théorème de Stokes.

corrVariete0020

Correction of the exercise 171

<+Variete0020+>

97.17 Autres

Exercise 172 exoTP20090001

Un ensemble $\Omega \subset \mathbb{R}^n$ est dit **étoilé par rapport à** $x_0 \in \Omega$ si pour tout $y \in \Omega$, le segment

$$[x_0,y] \stackrel{def}{=} \{(1-t)x_0 + ty \text{ tel que } t \in [0,1]\}$$

est inclus dans Ω. Un ensemble Ω étoilé par rapport à un de ses points est dit **étoilé**.

(1) Soit $f\colon A \subset \mathbb{R} \to \mathbb{R}$ une fonction dérivable sur l'ensemble connexe A. Montrer que si $f'(a) = 0$ pour tout $a \in A$, alors f est constante sur A.

(2) Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ où Ω est étoilé. Montrer que si $f \in C^1(\Omega,\mathbb{R})$ est telle que $df_a = 0$ pour tout $a \in \Omega$, alors f est constante.

corrTP20090001

Correction of the exercise 172

(1) Un connexe dans $\mathbb{R}$ est un intervalle. Prouvons d'abord que si f est dérivable sur $]a,b[$ avec $f'(x) = 0$ pour tout $x \in]a,b[$, alors f est constante sur $]a,b[$. En effet, si $x,y \in]a,b[$, alors le théorème des accroissements finis (page 184) montre qu'il existe un $c \in]x,y[$ tel que

$$f(y) - f(x) = f'(c)(y-x). \tag{97.609}$$

Étant donné que $f'(c) = 0$, nous en déduisons que $f(x) = f(y)$.

La propriété est donc prouvée dans le cas où A est un intervalle ouvert. Si l'ensemble A est un intervalle de la forme $[a,b]$, alors on peut utiliser la continuité pour prouver que si f est continues en a et b et constante sur $]a,b[$, alors elle est constante sur $[a,b]$. Pour s'en convaincre, supposons que $f(x) = y_0$ sur $]a,b[$ et que $f(a) = y_1$. Dans ce cas, étant donné que f est continue en a, il existe un voisinage de a dans lequel f prend ses valeurs uniquement dans $B(y_1,\epsilon)$. Si nous choisissons ϵ de telle manière à exclure y_0 de la boule (ce qui est possible si $y_1 \neq y_0$), alors nous contredisons la continuité de f en a parce que tout voisinage de a contient un point de $]a,b[$ où f vaut y_0.

(2) Considérons a_0, un point par rapport auquel Ω est étoilé. Soit $a \in \Omega$. Considérons le segment de droite $\gamma\colon [0,1]1 \to \Omega$ tel que $\gamma(0) = a_0$ et $\gamma(1) = a$. Nous considérons maintenant la fonction

$$\begin{aligned} g\colon [0,1] &\to \mathbb{R} \\ t &\mapsto (f \circ \gamma)(t). \end{aligned} \tag{97.610}$$

Cette fonction sert à ramener le problème sur Ω à un problème sur une droite. Nous avons que

$$g'(t) = \frac{dg}{dt}(t) = \frac{d}{dt}\big[f\big(\gamma(t)\big)\big] = \sum_i \frac{\partial f}{\partial x_i}\big(\gamma(t)\big) \cdot \frac{\partial \gamma_i}{\partial t}(t). \tag{97.611}$$

Par hypothèse, $df = 0$, ce qui veut dire que $\frac{\partial f}{\partial x_i}(p) = 0$ pour tout $p \in \Omega$, et en particulier

$$\frac{\partial f}{\partial x_i}\big(\gamma(t)\big) = 0 \tag{97.612}$$

pour tout $t \in [0,1]$. Donc nous avons prouvé que $g'(t) = 0$, et donc que g est une constante sur $[0,1]$. Pour tout $a \in \Omega$ nous avons donc que $f(a) = f(a_0)$. Cela finit de prouver que f est une constante.

Exercise 173 exoTP20090002

Soit $T : A \subset \mathbb{R}^n \to A$ une application. Supposons qu'il existe un entier $p > 1$ tel que $T^p \stackrel{\text{def}}{=} \underbrace{T \circ T \circ \ldots \circ T}_{p \text{ fois}}$ soit une application contractante.

(1) Si A est fermé, prouver à l'aide du théorème de Banach que T possède un unique point fixe.

(2) Montrer qu'il existe un ensemble $A \subset \mathbb{R}^2$ et une application $T : A \to A$ qui n'est pas une contraction, mais telle que T^2 est une contraction.

Remarque 97.2.
Cet énoncé est utilisé dans la démonstration du théorème de Cauchy-Lipschitz 17.39.

corrTP20090002

Correction of the exercise 173

Étant donné que T^p est contractante, elle possède un unique point fixe. Si x_0 est un point fixe de T, alors il est point fixe de T^p, en effet, pour tout $k > 1$, nous avons

$$T^k(x_0) = T^{k-1}T(x_0) = T^{k-1}(x_0). \tag{97.613}$$

Donc le point fixe de T^p est l'unique candidat point fixe de T. Montrons que si a est le point fixe de T^p, alors a est aussi un point fixe de T. Par définition,

$$T^p(a) = a. \tag{97.614}$$

En appliquant T des deux côtés et en tenant compte du fait que $T \circ T^k = T^k \circ T$, nous avons

$$T^p\big(T(a)\big) = T(a), \tag{97.615}$$

ce qui prouve que $T(a)$ est lui-même un point fixe de T^p. Mais le point fixe de T^p étant unique, cela implique que $T(a) = a$, et donc que a est un point fixe de T.

En ce qui concerne les exemples, une multitude d'exemples sont possibles. Le plus facile est de prendre une application $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ qui envoie tous les points de $\mathbb{R}^2$ sur $(0,0)$, sauf quelques-uns qui vont vers un autre point, en choisissant cet autre point de telle manière que T l'envoie sur $(0,0)$. De cette manière T n'est pas une contraction parce qu'elle n'est pas continue, et T^2 est une contraction parce que $T^2 = 0$.

Exercise 174 exoTP20090003

Considérons un tore T dans $\mathbb{R}^3$, c'est-à-dire la surface de révolution engendrée par la rotation d'un cercle de rayon r autour d'un axe coplanaire et disjoint de ce cercle, à distance R du centre du cercle.

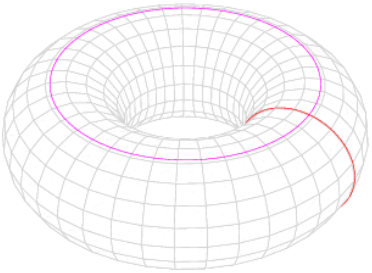


FIGURE 97.10: Les cercles du tore, figure de wikipedia. FigToreWiki

(1) Donner un paramétrage du tore en fonction de deux angles φ et θ bien choisis (la figure 97.10 devrait vous aider).
(2) Donner une carte autour d'un point (x_0, y_0, z_0) sur le tore et vérifier les conditions de la définition de variété sous forme paramétrique.
(3) Écrire l'équation du plan tangent au tore au point (x_0, y_0, z_0).
(4) Calculer le volume intérieur au tore.

corrTP20090003

Correction of the exercise 174

Équations paramétriques Soit C le cercle d'équation $(x-R)^2+z^2=r^2$, avec $r<R$. Il s'écrit sous forme paramétrique (avec $\varphi\in\mathbb{R}$) :

$$\begin{cases} x = R + r\cos\varphi \\ y = 0 \\ z = r\sin\varphi \end{cases}$$

Considérons la rotation de $\mathbb{R}^3$ d'angle θ autour de l'axe $0z$ (qui est bien disjoint et coplanaire à C) : c'est une transformation linéaire dont la matrice est

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Le tore T peut alors s'obtenir par rotation des points du cercle C sous l'action de R_θ ($\theta\in\mathbb{R}$) :

$$R_\theta \begin{pmatrix} R+r\cos\varphi \\ 0 \\ r\sin\varphi \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R+r\cos\varphi \\ 0 \\ r\sin\varphi \end{pmatrix} = \begin{pmatrix} (R+r\cos\varphi)\cos\theta \\ (R+r\cos\varphi)\sin\theta \\ r\sin\varphi \end{pmatrix}$$

d'où les équations paramétriques

$$T \equiv \begin{cases} x = (R+r\cos\varphi)\cos\theta \\ y = (R+r\cos\varphi)\sin\theta \\ z = r\sin\varphi \end{cases} \quad \text{où } \varphi,\theta\in\mathbb{R} \qquad (97.616)$$

torus_parameq

Ces équations ne fournissent évidemment pas une bijection entre l'ensemble des paramètres φ,θ et les points du tore. C'est un revêtement.

Mentionnons au passage l'existence d'une équation sous forme implicite : le système ci-dessus est équivalent à

$$\begin{cases} \sqrt{x^2+y^2} - R = r\cos\varphi \\ z = r\sin\varphi \end{cases}$$

et donc équivalent à $(\sqrt{x^2+y^2}-R)^2+z^2=r^2$. Le tore a donc pour équation, en éliminant la racine,

$$(x^2+y^2+z^2+R^2-r^2)^2 = 4R^2(x^2+y^2).$$

Cartes et atlas Il y a bien entendu une infinité d'atlas faisant du tore une variété, et un seul suffit à notre bonheur. Nous allons néanmoins donner plusieurs possibilités, pour montrer plusieurs approches du problème.

Version 1 Étant donné le point $p_0 = (x_0, y_0, z_0)$, nous allons définir une carte (U, F_{p_0}), où U sera un ouvert (relativement au tore) bien choisi autour de p_0, et F_{p_0} une restriction de le paramétrage ci-dessus.

L'idée est que la formule qui donne le paramétrage du tore, ci-dessus, est un bon candidat "carte", mais trois problèmes s'enchainent alors :

D'abord, ce n'est pas une bijection. Pour pallier à ce problème, nous allons restreindre l'espace des paramètres φ, θ jusqu'à ce qu'on puisse "inverser" la formule.

Deuxième problème, c'est qu'il est difficile d'inverser explicitement des formules contenant des sinus et cosinus. Nous allons donc appliquer le théorème de la fonction réciproque pour nous "simplifier la vie" (tout est question de point de vue).

Troisième problème : le théorème de la fonction réciproque ne fonctionne que d'un ouvert de $\mathbb{R}^n$ dans un ouvert de $\mathbb{R}^n$, avec le même "n". Le tore "vit" dans $\mathbb{R}^3$ mais n'en est pas un ouvert, nous allons donc l'épaissir (en laissant varier le petit rayon r) pour obtenir des "coordonnées toriques" valables dans un voisinage du tore. Ces coordonnées toriques seront inversibles, et nous permettrons de nous en sortir.

Soit F l'application "coordonnées toriques" définie par

$$\begin{aligned} F\colon \mathbb{R}_0^+ \times \mathbb{R} \times \mathbb{R} &\to \mathbb{R}^3 \\ (\rho, \varphi, \theta) &\mapsto \begin{pmatrix} (R + r\cos\varphi)\cos\theta \\ (R + \rho\cos\varphi)\sin\theta \\ \rho\sin\varphi \end{pmatrix} \end{aligned} \tag{97.617}$$

(de sorte que le tore T est obtenu en fixant $\rho = r$ et en laissant varier φ, θ). Le déterminant de la matrice jacobienne de F, donné par $J = -\rho(R + \rho\cos\varphi)$, est non nul sur le tore T puisqu'alors $\rho = r > 0$ et $R - \rho\cos\varphi \geqslant R - \rho = R - r > 0$.

Soit $p_0 = (x_0, y_0, z_0)$ un point du tore, et soit (φ_0, θ_0) un couple de réels tels que $F(r, \varphi_0, \theta_0) = p_0$. Alors la différentielle de F en (r, φ_0, θ_0) est un isomorphisme linéaire (car le déterminant de la jacobienne est non-nul), et le théorème de la fonction réciproque nous fournit alors un ouvert $\tilde{W} \subset \mathbb{R}_0^+ \times \mathbb{R} \times \mathbb{R}$ autour de (r, φ_0, θ_0) et un ouvert $\tilde{U} \subset \mathbb{R}^3$ autour de p_0 tels que $F_{|\tilde{W}} : \tilde{W} \to \tilde{U}$ est un difféomorphisme. Notons $G : \tilde{U} \to \tilde{W}$ sa réciproque.

Soit F_{p_0} la restriction de $F_{|\tilde{W}}$ à $\rho = r$, c'est-à-dire

$$F_{p_0} : W \subset \mathbb{R}^2 \to \mathbb{R}^3 : (\varphi, \theta) \mapsto F(r, \varphi, \theta)$$

où $W = \{(\varphi, \theta) \in \mathbb{R}^2 \text{ tel que } (r, \varphi, \theta) \in \tilde{W}\}$, et soit $U = F(W) = \tilde{U} \cap T$. Alors (U, F_{p_0}) est une carte autour de p_0 :

(1) F_{p_0} est continue, car c'est une restriction de F (qui est continue).
(2) F_{p_0} est une injection de W dans U, car c'est une restriction de $F_{|\tilde{W}}$ (qui est un difféomorphisme, donc en particulier injective).
(3) F_{p_0} est une surjection, par définition de U.
(4) Par les deux derniers points, F_{p_0} est une bijection ; de plus sa réciproque est continue car c'est la restriction de G à U (et G est différentiable donc en particulier continue).
(5) F_{p_0} est de classe C^1 (comme application d'un ouvert de $\mathbb{R}^2$ dans $\mathbb{R}^3$) car est composée d'applications différentiables (sinus, cosinus, produits d'applications différentiables,...).
(6) Le rang de la différentielle de F_{p_0} est maximal, car sa matrice jacobienne (en (φ_0, θ_0)) est la matrice jacobienne de F (en (r, φ_0, θ_0)) dont on a enlevé la première colonne (correspondant aux dérivées par rapport à ρ). Or cette dernière était de rang trois, il reste donc une matrice de rang deux.

Ceci prouve qu'on a bien une carte autour de p_0, et comme on l'a fait en un point arbitraire du tore, on a bien une variété de dimension 2 dans $\mathbb{R}^3$.

Version 2 L'atlas ci-dessus est composé de cartes données par le théorème de la fonction implicite. On peut bien sûr choisir nous même la carte qu'on veut autour d'un point p_0 fixé sur le tore.

Notons f le paramétrage du tore :

$$\begin{aligned} f\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R}^3 \\ (\varphi,\theta) &\mapsto \begin{pmatrix} (R+r\cos\varphi)\cos\theta \\ (R+r\cos\varphi)\sin\theta \\ r\sin\varphi \end{pmatrix} \end{aligned} \tag{97.618}$$

et donnons nous (φ_0,θ_0) tels que $p_0 = f(\varphi_0,\theta_0)$.

On définit

$$W =]\varphi_0-\pi;\varphi_0+\pi[\times]\theta_0-\pi;\theta_0+\pi[$$

et on pose $U = f(W)$.

Montrons que $f_W : W \to U$ est une bijection : c'est évidemment une surjection par définition de U, et il reste à montrer qu'elle est injective. Supposons avoir (u,v) et (x,y) dans W tels que $f(u,v) = f(x,y)$. C'est-à-dire :

$$\begin{aligned} (R+r\cos u)\cos v &= (R+r\cos x)\cos y \\ (R+r\cos u)\sin v &= (R+r\cos x)\sin y \\ r\sin u &= r\sin x \end{aligned} \tag{97.619}$$

En sommant les carrés des deux premières équations on obtient

$$(R+r\cos u)^2 = (R+r\cos x)^2$$

ce qui montre que $\cos u = \cos x$ (car $r < R \Rightarrow R + r\cos x > 0$). D'autre part, on sait que $\sin u = \sin x$. Deux angles dans un intervalle de longueur 2π n'ont même sinus et même cosinus que s'ils sont égaux, ce qui prouve que $u = x$. Avec cette information, les deux premières équations donnent alors v et y ont également même sinus et même cosinus, donc $v = y$. On en déduit $(u,v) = (x,y)$, et l'injectivité.

Les autres points à montrer pour avoir une carte (continûment différentiable, inverse continue, de rang maximal) s'obtiennent de la même manière que dans la première version.

Version 2 bis Plutôt que de construire une carte autour d'un point donné du tore, on peut définir un atlas de quelques cartes bien choisies recouvrant la totalité du tore. C'est par exemple le cas si on définit les ouverts (f désigne le paramétrage, comme ci-dessus)

$$\begin{aligned} W_1 &=]0;2\pi[\times]0;2\pi[\qquad U_1 = f(W_1) \\ W_2 &=]-\pi;\pi[\times]-\pi;\pi[\qquad U_2 = f(W_2) \\ W_3 &=]-\frac{\pi}{2};\frac{3\pi}{2}[\times]-\frac{\pi}{2};\frac{3\pi}{2}[\qquad U_3 = f(W_3) \end{aligned}$$

Dans cette situation, U_1, U_2 et U_3 recouvrent bien le tore. L'application de carte est dans chaque cas la restriction de f à W_i, c'est-à-dire $f_i : W_i \to U_i$. Pour vérifier que chaque (U_i, f_i) définit une carte, on peut procéder comme ci-dessus.

Version 3 Une autre manière de construire des cartes se base sur l'idée qu'un tore dont on enlève un cercle bien choisi est un cylindre. Et un cylindre est homéomorphe à un anneau : si on se donne le cylindre

$$C \equiv \{(x,y,z) \text{ tel que } x^2+y^2 = 1, 1 < z < 2\}$$

et l'anneau

$$A \equiv \{(x,y) \text{ tel que } 1 < x^2+y^2 < 4\}$$

on a l'homéomorphisme :

$$C \to A : (x,y,z) \mapsto (xz,yz)$$

Il n'est alors pas très dur de construire des cartes basées sur ce principe.

Version 4 [Idée de Patrick Weber] Toujours sur l'idée que construire une carte revient à enlever des points au tore et obtenir une surface qu'on sait déplier, on peut se baser sur le paramétrage

$$\begin{aligned} f\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R}^3 \\ (\varphi,\theta) &\mapsto \begin{pmatrix} (R+r\cos\varphi)\cos\theta \\ (R+r\cos\varphi)\sin\theta \\ r\sin\varphi \end{pmatrix} \end{aligned} \tag{97.620}$$

et enlever le cercle $\varphi=\theta$. Posons

$$\begin{aligned} W_1 &= \{(\varphi,\theta) \text{ tel que } 0<\varphi<2\pi \text{et} \varphi<\theta<\varphi+2\pi\} \\ W_2 &= \{(\varphi,\theta) \text{ tel que } 0<\varphi<2\pi \text{et} \varphi-\pi<\theta<\varphi+\pi\} \end{aligned}$$

et $U_i=f(W_i)$. Alors pour $i=1,2$, f est un homéomorphisme entre W_i et U_i, et on obtient un atlas à deux cartes.

Plan tangent

Version 1 Ayant l'équation du tore sous forme implicite, on peut l'utiliser pour déterminer le plan tangent au tore en un point $p_0=(x_0,y_0,z_0)$: il est donné par l'équation

$$\begin{aligned} &4(x_02+y_02+z_02-R2-r2)x_0(x-x_0) \\ &+4(x_02+y_02+z_02-R2-r2)y_0(y-y_0) \\ &+4(x_02+y_02+z_02+R2-r2)z_0(z-z_0)=0. \end{aligned} \tag{97.621}$$

Ceci n'est valable que si la forme implicite est "la bonne" (il faut que le gradient soit non-nul en chaque point du tore), ce qui est vrai mais est à démontrer.

Version 2 On sait que le plan tangent sera le plan perpendiculaire au "petit rayon" du tore. Ce dernier est le vecteur qui relie (en notant encore F les "coordonnées toriques" définites ci-dessus) le point $F(0,\varphi_0,\theta_0)$ et le point $F(r,\varphi_0,\theta_0)$. Ce vecteur a pour composantes

$$F(r,\varphi_0,\theta_0)-F(0,\varphi_0,\theta_0)=\big(r\cos\varphi\cos\theta, r\cos\varphi\sin\theta, r\sin\varphi\big)$$

et donc le plan tangent a pour équation

$$\cos\varphi\cos\theta(x-(R+r\cos\varphi)\sin\theta)+\cos\varphi\sin\theta(y-(R+r\cos\varphi)\cos\theta)+\sin\varphi(z-r\sin\varphi)=0$$

Volume du tore

Version 1 Utilisons les "coordonnées toriques" définies ci-dessus pour écrire le changement de coordonnées

$$\begin{cases} x=(R+\rho\cos\varphi)\cos\theta \\ y=(R+\rho\cos\varphi)\sin\theta \\ z=\rho\sin\varphi \end{cases}$$

pour $0<\rho<r$, $0<\varphi<2\pi$ et $0<\theta<2\pi$. Ceci paramétrise l'intérieur du tore (à un ensemble de mesure nulle près), et la valeur absolue du jacobien de cette transformation est $\rho(R+\rho\cos\varphi)$. Le volume du tore est donc l'intégrale

$$\begin{aligned} &\int_0^{2\pi}\int_0^{2\pi}\int_0^r \rho(R+\rho\cos\varphi)d\rho d\varphi d\theta \\ =&\int_0^{2\pi}\int_0^{2\pi}\left(\frac{1}{2}Rr^2+\frac{1}{3}r^3\cos\varphi\right)d\varphi d\theta \\ =&\int_0^{2\pi}\pi Rr^2 d\theta \\ =&2\pi^2Rr^2 \end{aligned}$$

avec la remarque que c'est exactement le volume d'un cylindre de hauteur $2\pi R$ et de rayon r, correspondant au tore qu'on aurait tranché comme un saucisson pour le redresser.

Version 2 Une autre approche consiste à coup le tore "horizontalement". La mesure d'une tranche horizontale est l'aire d'un anneau dont le grand rayon est $R+\sqrt{r^2-z^2}$ et le rayon intérieur est $R-\sqrt{r^2-z^2}$ annulaires, et à intégrer pour z entre $-r$ et r. C'est-à-dire

$$\begin{aligned}\int_{-r}^{r}\Big(\pi(R+\sqrt{r^2-z^2})^2-\pi(R-\sqrt{r^2-z^2})^2\Big)\,dz &= \pi\int_{-r}^{r}4R\sqrt{r^2-z^2}dz\\ &= 4\pi R\int_{-r}^{r}\sqrt{r^2-z^2}dz\end{aligned}\tag{97.622}$$

où l'intégrale qui reste à calculer est exactement l'aire d'un demi-disque de rayon r, c'est-à-dire $\frac{1}{2}\pi r^2$. On en déduit que le volume du tore est $2\pi^2r^2R$.

Bonnes vacances !

97.18 Suites de fonctions

Exercise 175 exo111
On considère les fonctions

$$f_n(x)=(1-x^4)^n\tag{97.623}$$

pour $n\in\mathbb{N}$.

(1) La suite $\{f_n\}$ converge-t-elle sur $[-1,1]$? Si oui, vers quelle fonction ?
(2) La suite $\{f_n\}$ converge-t-elle uniformément sur $[-1,1]$?
(3) La suite $\{f_n\}$ converge-t-elle uniformément sur tout compact de $[-1,1]$?

corr111

Correction of the exercise 175

(1) La suite de fonctions proposée converge (ponctuellement) vers la fonction

$$f(x)=\begin{cases}0 & \text{si } x\neq 0\\ 1 & \text{si } x=0\end{cases}\tag{97.624}$$

(2) En vertu du théorème 12.357, si la convergence était uniforme, la fonction limite devrait être continue, ce qui n'est pas le cas. La convergence n'est donc pas uniforme sur $[-1,1]$.
(3) Soit K, un compact dans $]0,1]$, et appelons a son minimum. Soit $\epsilon>0$, et choisissons N tel que $f_n(a)<\epsilon$ dès que $n>N$. Étant donné que $f(x)=0$ sur K et que toutes les fonctions en jeu sont positives, nous avons :

$$\|f_n(x)-f(x)\|=f_n(x)\leqslant f_n(a)<\epsilon,\tag{97.625}$$

ce qui prouve que la suite de fonctions est uniformément convergente sur tout compact de $]0,1]$.

Exercise 176 exo112
On considère les fonctions

$$f_n(x)=\begin{cases}0 & \text{si } \frac{1}{n}\leqslant x\leqslant 1\\ 1 & \text{pour } x=0\\ 1-nx & \text{si } 0<x<\frac{1}{n}\end{cases}\tag{97.626}$$

(1) La suite $\{f_n\}$ converge-t-elle sur $[0,1]$? Si oui, vers quelle fonction ?
(2) La suite $\{f_n\}$ converge-t-elle uniformément sur $[0,1]$? Et sur $]0,1]$?

(3) La suite $\{f_n\}$ converge-t-elle uniformément sur tout compact de $]0,1]$?

corr112

Correction of the exercise 176

(1) La suite converge sur $[0,1]$ vers la fonction

$$f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0. \end{cases} \tag{97.627}$$

(2) La convergence n'est pas uniforme parce que la limite n'est pas continue. Sur $]0,1]$, la convergence n'est pas uniforme non plus parce que $\forall n$, il existe $x \in]0,1]$ tel que $f_n(x) = \frac{1}{2}$ et $f(x) = 0$, à cause du théorème des valeurs intermédiaires.

(3) Soit a, le minimum du compact K. On a que $|f_n(x)| \leqslant |f_n(a)|$ pour tout $x \in K$. Autrement dit,

$$\|f_n\|_\infty \leqslant |f_n(a)|. \tag{97.628}$$

Or, la suite numérique $|f_n(a)|$ tends vers zéro lorsque $n \to \infty$, donc nous avons convergence uniforme sur le compact K.

Exercise 177 exo113

Soit

$$f_n(x) = \begin{cases} 1 & \text{pour } n < x < n+1 \\ 0 & \text{sinon.} \end{cases} \tag{97.629}$$

(1) La suite $\{f_n\}$ converge-t-elle sur $\mathbb{R}$?

(2) La suite $\{f_n\}$ converge-t-elle uniformément sur $\mathbb{R}$?

(3) La suite $\{f_n\}$ converge-t-elle uniformément sur tout compact de $\mathbb{R}$?

corr113

Correction of the exercise 177

(1) La suite converge vers $f(x) = 0$ sur $\mathbb{R}$.

(2) Pas de convergence uniforme sur $\mathbb{R}$ parce que pour tout n, il existe un $x \in \mathbb{R}$ tel que $f_n(x) = 1$, par exemple $x = n + \frac{1}{2}$.

(3) Soit K, un compact de $\mathbb{R}$ et M, un majorant entier de K. $\forall n > M$, on a $f_n(x) = 0$, donc la convergence uniforme sur tout compact est triviale (pour chaque compact, il existe un moment où la suite ne bouge même plus).

Exercise 178 exo114

Soit

$$f_n(x) = \begin{cases} 0 & \text{pour } |x| > 2 \\ x+2 & \text{pour } -2 \leqslant x \leqslant -1 \\ 1 & \text{pour } -1 < x \leqslant -1/n \\ -nx & \text{pour } -1/n < x < 1/n \\ -1 & \text{pour } 1/n \leqslant x < 1 \\ x-2 & \text{pour } 1 \leqslant x < 2 \end{cases} \tag{97.630}$$

Montrer que la suite $\{f_n\}$ converge sur $\mathbb{R}$, qu'elle ne converge pas uniformément sur $\mathbb{R}$ et qu'elle ne converge pas uniformément sur tout compact de $\mathbb{R}$.

corr114

Correction of the exercise 178

(1)

(2) Pour cause de discontinuité de la limite f en 0, nous n'avons pas de convergence uniforme, ni sur $\mathbb{R}$, ni sur tout compact de $\mathbb{R}$.

Exercise 179 exo115

Montrer que la suite

$$f_n(x) = \frac{\sin(nx)}{n} \tag{97.631}$$

converge uniformément sur $[0, 2\pi]$. corr115

Correction of the exercise 179

Nous avons que

$$|f_n(x)| = \frac{|\sin(nx)|}{n} \leqslant \frac{1}{n}. \tag{97.632}$$

Si $\epsilon > 0$ et $N = 1/\epsilon$, nous avons que $\forall n > N$,

$$\frac{|\sin(nx)|}{n} \leqslant \frac{1}{n} < \frac{1}{n} < \epsilon, \tag{97.633}$$

ce qui prouve l'uniforme convergence de la suite vers $f = 0$.

Exercise 180 exo116

Étudier la convergence sur $[0, 1]$ de la suite $f_n(x) = nxe^{-nx^2}$. Calculer

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx \quad \text{et} \quad \int_0^1 \lim_{n\to\infty} f_n(x)dx. \tag{97.634}$$

corr116

Correction of the exercise 180

La première chose à faire est d'un peu regarder à quoi ressemble la fonction, et en particulier, calculer sa dérivée qui va nous permetre de trouver le maximum de la fonction :

$$f_n'(x) = ne^{-nx^2}(1 - 2nx^2). \tag{97.635}$$

D'autre part, pour tout $x \in [0, 1]$, on a évidement $\lim_{n\to\infty} f_n(x) = 0$, donc $f_n \to 0$. Pour l'uniforme convergence, nous calculons

$$\sup_{x\in[0,1]} |f_n(x) - f(x)| = \sup_{x\in[0,1]} |f_n(x)| = \max_{x\in[0,1]} f_n(x) = e^{-1/2}\sqrt{\frac{n}{2}}. \tag{97.636}$$

Dans ce calcul, le suprémum de la valeur absolue est devenu le maximum sans valeur absolue parce que f_n est positive et continue, or une fonction continue sur un compact atteint ses bornes. En conclusion, il n'y a pas convergence uniforme sur $[0, 1]$.

Une primitive de f_n est donnée par

$$F_n(x) = -\frac{e^{-nx^2}}{2}, \tag{97.637}$$

et on vérifie que

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \frac{1}{2} \quad \text{mais} \quad \int_0^1 \lim_{n\to\infty} f_n(x)dx = 0. \tag{97.638}$$

Le fait qu'il n'y ait pas égalité nous permet de vérifier immédiatement que la suite ne converge pas uniformément, en vertu du théorème 14.205. En particulier, nous notons que

$$\int_0^1 f_n(x)dx = \frac{1}{2}(1 - e^{-n}). \tag{97.639}$$ Eqintfexpmoin

Exercise 181 exo117

Trouver un exemple de suite de fonctions $\{f_n(x)\}$ qui converge sur $[0, 1]$, ne converge pas uniformément, et telle que

(1) $\lim_{n\to\infty} \int_0^1 f_n(x)dx = \infty$

(2) $\lim_{n\to\infty} \int_0^1 f_n(x)dx = A \neq \int_0^1 \lim_{n\to\infty} f_n(x)dx = \infty$

(3) $\lim_{n\to\infty} \int_0^1 f_n(x)dx = A = \int_0^1 \lim_{n\to\infty} f_n(x)dx = \infty$

corr117

Correction of the exercise 181

(1) La fonction $f_n(x) = nxe^{-nx^2}$ de l'exercice 180 fait presque ce qu'il faut, mais ce serait bien que la formule (97.639) diverge un peu plus vis-à-vis de n. Essayons donc $f_n(x) = n^2xe^{-nx^2}$. Par construction, nous avons

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \infty, \tag{97.640}$$

et, étant donné que pour tout x nous avons $\lim_{n\to\infty} f(x)=0$, la permutation de la limite avec l'intégrale donne

$$\int_0^1 \lim_n f_n = 0. \tag{97.641}$$

Le simple fait que les deux intégrales ne coïncident pas prouve que la suite des f_n ne converge pas uniformément sur $[0, 1]$.

(2) La suite de fonctions de l'exercice 180 répond à la question.

(3) Faisant les calculs, nous voyons que pour la suite de fonctions de l'exercice 176, nous avons

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 \lim_{n\to\infty} f_n(x)dx = 0. \tag{97.642}$$

Il est intéressant de remarquer ce, *dans ce cas*, on peut permuter la limite et l'intégrale sans changer la valeur de l'intégrale.

Exercise 182 exo118

Étudier la convergence sur $]0, \pi[$ de la suite de fonctions

$$f_n(x) = \frac{1 - \cos(nx)}{nx^2}. \tag{97.643}$$

corr118

Correction of the exercise 182

Pour tout $x \in]0, \pi[$, nous avons

$$|f_n(x)| = \frac{|1 - \cos(nx)|}{nx^2} \leqslant \frac{2}{nx^2}, \tag{97.644}$$

donc nous avons que $f_n \to 0$. En ce qui concerne l'uniforme convergence, nous avons, en utilisant le résultat $\lim_{x\to 0} \sin(x)/x = 1$:

$$\lim_{x\to 0} \frac{1 - \cos(nx)}{nx^2} = \lim_{x\to 0} \frac{n \sin(nx)}{2nx} = \frac{n}{2} \lim_{x\to 0} \frac{\sin(nx)}{x} = \frac{n}{2}, \tag{97.645}$$

donc $\sup_{x\in]0,\pi[} \left|\frac{1-\cos(nx)}{nx^2}\right| \geqslant \frac{n}{2}$. Cela empêche la convergence uniforme vers $f(x) = 0$: pour l'avoir, il faudrait au moins que ce supprémum tende vers zéro quand n tend vers l'infini.

Cependant, lorsque $x \geqslant a > 0$, nous avons $f_n(x) \leqslant \frac{2}{na^2}$, ce qui donne la convergence uniforme sur tout compact.

Exercise 183 exo1182

On considère les fonctions $f_n(x) = n^2x(1 + n^2x^2)^{-1}$, $n \in \mathbb{N}$.

(1) Montrer que la suite $\{f_n\}$ converge pour tout $x \in \mathbb{R}$ et déterminer la fonction $f = \lim_{n\to\infty} f_n$. Item1182ii

(2) la suite $\{f_n\}$ converge-t-elle uniformément sur $\mathbb{R}$? Item1182iii

(3) la suite $\{f_n\}$ converge-t-elle uniformément sur $]0, \infty[$? Item1182iv

(4) la suite $\{f_n\}$ converge-t-elle uniformément sur tout compact de $]0, \infty[$?

Justifier les réponses aux point (2), (3), (4). corr1182

Correction of the exercise 183

(1) Lorsque $x \neq 0$, nous avons que

$$\lim_{n\to\infty} \frac{n^2 x}{1+n^2x^2} = \frac{1}{x}, \tag{97.646}$$

et $f_n(0) = 0$ pour tout n. Par conséquent, la limite de la suite des fonctions f_n est

$$f(x) = \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases} \tag{97.647}$$

(2) Le fait que la limite soit discontinue fait qu'il n'y a pas convergence uniforme sur $\mathbb{R}$ (plus généralement, il n'y a pas convergence uniforme sur aucun ouvert contenant zéro).

(3) Nous avons que $(f_n - f)(x) = -\frac{1}{x(n^2x^2+1)} \to -\infty$ pour $x \to 0$, donc il n'y a pas de convergence uniforme sur $]0,\infty[$.

(4) Étant donné que sur $]0,\infty[$, la fonction $|f_n - f|$ est décroissante, si a est le minimum du compact K, alors

$$\sup_{x\in K} |f_n - f| = \frac{1}{a(1+n^2a^2)}. \tag{97.648}$$

Pour cette raison, $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$, et on a une convergence uniforme sur tout compact de $]0,\infty[$.

Exercise 184 exo119

Étudier la convergence sur $[0,1]$ de la suite de fonctions

$$f_n(x) = \begin{cases} 0 & \text{pour } x \in [0, 1 - \frac{1}{n+1}] \\ \frac{1}{x^n} - 1 & \text{pour } x \in]1 - \frac{1}{n+1}, 1] \end{cases} \tag{97.649}$$

corr119

Correction of the exercise 184

Remarquons d'abord que toutes les fonctions f_n sont discontinues en $1 - \frac{1}{n+1}$, et que pour tout n, nous avons $f_n(1) = 0$. D'autre part, le domaine $[0, 1 - \frac{1}{n+1}]$ sur lequel f_n est nulle est croissant (avec n) et tend vers $[0,1]$ entier. Donc, pontuellement, nous avons la convergence $f_n \to 0$.

La convergence uniforme sur $[0,1]$ demande d'avoir

$$\lim_{n\to\infty} \|f_n - f\|_\infty = 0. \tag{97.650}$$

Le calcul de cette quantité est aisé :

$$\lim_{n\to\infty} \|f_n - f\|_\infty = \lim_{n\to\infty} \sup_{x\in[0,1]} |f_n| = \lim_{n\to\infty} \left(\frac{1}{\left(1 - \frac{1}{n+1}\right)^n} - 1 \right), \tag{97.651}$$

où la limite se calcule en remarquant que $1 - \frac{1}{n+1} = \frac{n}{n+1} = \left(1 + \frac{1}{n}\right)^{-1}$, et donc nous nous retrouvons avec la limite

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n - 1 = e - 1 \neq 0. \tag{97.652}$$

Il n'y a donc pas convergence uniforme sur $[0,1]$.

La suite converge par contre uniformément sur tout compact de $[0,1[$. En effet, soit b le maximum du compact $K \subset [0,1[$. Dès que $1 - \frac{1}{n+1} > b$, nous avons $f_n|_K = 0$.

97.19 Séries de fonctions

Exercise 185 exo_I-1-10

Sur $[0,1]$, on considère les fonctions

$$f_n(x) = \begin{cases} \frac{1}{n} & \text{pour } \frac{1}{2^n} < x \leqslant \frac{1}{2^{n-1}} \\ 0 & \text{ailleurs.} \end{cases} \tag{97.653}$$

La série $\sum_{n=1}^{\infty} f_n(x)$ converge-t-elle

(1) ponctuellement ?

(2) uniformément ?

(3) normalement ?

corr_I-1-10

Correction of the exercise 185

(1) D'abord, $f_n(0) = 0$ pour tout n, donc $f(0) = 0$. Maintenant, pour $p \in \mathbb{N}$ et $x \in]\frac{1}{2^p}, \frac{1}{2^{p-1}}]$, si $n \geqslant p$, alors

$$\sum_{k=1}^{n} f_k(x) = \frac{1}{p}, \tag{97.654}$$

donc $\lim_{n\to\infty} \left| \sum_{k=1}^{n} f_k(x) - \frac{1}{p} \right| = 0$, et nous avons donc convergence ponctuelle vers la fonction

$$f(x) = \begin{cases} \frac{1}{p} & \text{si } \frac{1}{2^p} < x \leqslant \frac{1}{2^{p-1}} \\ 0 & \text{en } x = 0. \end{cases} \tag{97.655}$$

(2) Pour la convergence uniforme, remarquer que

$$\left| f(x) - \sum_{k=1}^{n} f_k(x) \right| < \frac{1}{n+1}, \tag{97.656}$$

donc $\forall \epsilon > 0$, $\exists N$ tel que $n > N$ implique

$$\| f - \sum_{k=1}^{n} f_k \|_\infty < \epsilon, \tag{97.657}$$

en l'occurrence, le N qui fonctionne est donné par $\frac{1}{N+1} < \epsilon$. Nous avons donc convergence uniforme.

(3) Étant donné que $\|f_k\|_\infty = 1/k$, la série numérique $\sum_{k=1}^{\infty} \|f_k\| = \sum_{k=1}^{\infty} \frac{1}{k}$ diverge et nous n'avons pas de convergence normale.

Nous avons donc un exemple du fait que la convergence uniforme n'implique pas la convergence normale.

Exercise 186 exo_I-1-11

Montrer que la **fonction de Riemann**

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} \tag{97.658}$$

est C^∞ pour $x > 1$.

Exercise 187 exo_I-1-12

Étudier la convergence de la série $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^x}$. Est-ce que cette somme définit une fonction continue ?

corr_I-1-12

Correction of the exercise 187

En $x = 0$, nous avons la série $\sum_n (-1)^n$ qui ne converge pas. En $x < 0$, le terme général de la série ne converge pas vers zéro, donc il y a divergence de la série. Travaillons donc avec $x > 0$.

Prenons un compact K dans $]0,\infty[$, nommons α son minimum. Pour tout $x\in K$, nous avons

$$\frac{1}{n^x}\leqslant\frac{1}{n^\alpha}. \tag{97.659}$$

Sur le compact, la suite de fonctions $b_n(x)=\frac{1}{n^x}$ converge uniformément vers zéro, de telle sorte que le critère d'Abel (théorème 12.367) conclue à la convergence uniforme. Nous avons donc la convergence uniforme sur tout compact dans $]0,\infty[$. Le théorème 12.370 dit alors que

$$f(x)=\sum_{n=1}^{\infty}\frac{(-1)^n}{n^x} \tag{97.660}$$

est une fonction continue sur $]0,\infty[$.

Exercise 188 exo_I-1-13

Sur $[\frac{1}{2},1]$, étudier la convergence de la série

$$f(x)=\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{x-1}{x}\right)^n. \tag{97.661}$$

Étudier la convergence de la série dérivée ; en déduire que

$$\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}=\ln(2). \tag{97.662}$$

Exercise 189 exo_I-1-14

Pour $x\in[1,3]$, soit

$$f(x)=\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{x-3}{2x+1}\right)^n. \tag{97.663}$$

Montrer que

$$\frac{df}{dx}=\frac{7}{(2x+1)(x+4)} \tag{97.664}$$

en justifiant les différentes étapes du calcul. corr_I-1-14

Correction of the exercise 189

Par le critère des séries de puissances (poser $y=(x-3)/(2x+1)$), nous avons la convergence uniforme de $f(x)$ pour $\left|\frac{x-3}{2x+1}\right|<1$, donc pour $x\in[1,3]$. La série des dérivées se somme en utilisant la même technique que pour l'exercice 188 :

$$\begin{aligned}
g(x)&=\sum_{n=1}^{\infty}\left(\frac{x-3}{2x+1}\right)^{n-1}\frac{2x+1-2(x-3)}{(2x+1)^2}\\
&=\frac{7}{(2x+1)^2}\sum_{n=1}^{\infty}\left(\frac{x-3}{2x+1}\right)^{n-1}\\
&=\frac{7}{(2x+1)^2}\sum_{n=0}^{\infty}\left(\frac{x-3}{2x+1}\right)^{n}\\
&=\frac{7}{(2x+1)^2}\frac{1}{1-\frac{x-3}{2x+1}}\\
&=\frac{7}{(x+4)(2x+1)}.
\end{aligned} \tag{97.665}$$

Étant donné que la série $\sum_{n=0}^{\infty}\left(\frac{x-3}{2x+1}\right)^n$ converge uniformément sur $[1,3]$, sur cet intervalle, $g(x)$ est effectivement la dérivée de $f(x)$.

Exercise 190 exo_I-1-15

Soit $f\colon[a,b]\to[a,b]$, une fonction appartenant à $C^1\big([a,b]\big)$, et soit $\{f_n\}$, la suite des fonctions itérées définies par

$$f_0(x)=x,\quad f_n(x)=f\big(f_{n-1}(x)\big). \tag{97.666}$$

(1) Montrer que si $\sup_{x\in[a,b]}|f'(x)|<1$, alors la série

$$\sum_{k=1}^{\infty}\big(f_k(x)-f_{k-1}(x)\big) \tag{97.667}$$

converge uniformément.

(2) En déduire que la suite des fonctions f_n converge uniformément vers une fonction constante C, que $f(C)=C$ et que C est unique.

(Indication : utiliser la formule des accroissements finis). corr_I-1-15

Correction of the exercise 190

Afin de majorer la différence $|f_k(x)-f_{k-1}(x)|=|f\big(f_{k-1}(x)\big)-f\big(f_{k-2}(x)\big)|$, nous utilisons la formule des accroissements finis sur f : pour tout x_1, $x_2\in[a,b]$, il existe un $x^*\in[x_1,x_2]$ tel que $f(x_2)-f(x_1)=f'(x^*)(x_2-x_1)$. Nous obtenons successivement

$$\begin{aligned}
|f_k(x)-f_{k-1}(x)|&=\Big|f\big(f_{k-1}(x)\big)-f\big(f_{k-2}(x)\big)\Big|\\
&=\Big|f'(x^*)\big(f_{k-2}(x)-f_{k-1}(x)\big)\Big|\\
&\leqslant\lambda\big|f_{k-2}(x)-f_{k-1}(x)\big|\\
&\leqslant\lambda^2\big|f_{k-3}(x)-f_{k-2}(x)\big|\\
&\;\;\vdots\\
&\leqslant\lambda^{k-1}\big|f_{k-k}(x)-f_{k-(k-1)}(x)\big|.
\end{aligned}\tag{97.668}$$

où $\lambda=\sup_{[a,b]}|f'(x)|$. Nous avons donc la majoration

$$|f_k(x)-f_{k-1}(x)|\leqslant\lambda^{k-1}|f(x)-x|. \tag{97.669}$$

Si M est le maximum de $|f(x)-x|$ sur le compact $[a,b]$, alors

$$|f_k(x)-f_{k-1}(x)|\leqslant M\lambda^{k-1}, \tag{97.670}$$

mais si $\lambda<1$, la série $\sum_k\lambda^k$ converge. Le critère de Weierstrass (théorème 12.372) montre alors que la série

$$g(x)=\sum_{k=1}^{\infty}\big(f_k(x)-f_{k-1}(x)\big) \tag{97.671}$$ EqSeriegfkmoinsfkmoinsun

converge uniformément sur $[a,b]$. La suite des sommes partielles de la série (97.671) est donnée par f_1-f_0, $-f_0+f_2$, $-f_0+f_3$,...c'est-à-dire que

$$g(x)=-x+\lim_{n\to\infty}f_n(x). \tag{97.672}$$

On peut exprimer ce résultat autrement en disant que $g(x)$ est la limite de ses sommes partielles qui sont données par $g_n(x)=-x+f_n(x)$. Nous avons démontré que le membre de gauche a une limite uniforme, donc le membre de droite possède également une limite uniforme. Ceci prouve que la suite de fonctions $\{f_n(x)\}$ converge uniformément vers une certaine fonction $C(x)$.

Afin de prouver que la fonction $C(x)$ est constante, nous majorons la différence de f_n prise en deux points différents :

$$\begin{aligned}
\big|f_n(x_1)-f_n(x_2)\big|&=\Big|f\big(f_{n-1}(x_1)\big)-f\big(f_{n-1}(x_2)\big)\Big|\\
&=|f'(x_1^*)||f_{n-1}(x_1)-f_{n-1}(x_2)|\\
&\;\;\vdots\\
&=|f'(x_1^*)|\dots|f'(x_n^*)||x_1-x_2|
\end{aligned}\tag{97.673}$$

où les points n_i^* sont des points donnés par le théorème des accroissements finis. La norme $|x_1-x_2|$ est évidement majorée par $|b-a|$, tandis que les nombres $|f'(x_j^*)|$ sont tous majorés par $\lambda<1$. Au final, nous avons

$$\big|f_n(x_1)-f_n(x_2)\big|\leqslant\lambda^n|b-a|. \tag{97.674}$$

En prenant n assez grand, cela peut être rendu aussi petit que souhaité. La fonction $C(x)$ est donc constante.

L'unicité de C en tant que point fixe de f est facile : en effet, si F est un point fixe de f, alors $f_1(F)=f_2(F)=\ldots$ Donc le point F est un point fixe de la limite des f_n. Donc seule la limite vers laquelle f_n tend peut être un point fixe de f.

Prouvons maintenant que C est effectivement un point fixe de f. Par définition, nous avons $f_n(C)=f\big(f_{n-1}(C)\big)$. Soit $\epsilon>0$. En vertu de la continuité des fonctions f_n, il existe un N tel que $n>N$ implique

$$\begin{aligned}f_n(C)&=C+\epsilon_1\\ f_{n-1}(C)&=C+\epsilon_2\end{aligned} \tag{97.675}$$

avec $|\epsilon_1|$ et $|\epsilon_2|$ plus petits que ϵ. Pour un tel n, nous avons alors $C+\epsilon_1=f(C+\epsilon_2)$. Par la continuité de f, il existe un ϵ_3 avec $|\epsilon_3|<\epsilon$ tel que $f(C+\epsilon_2)=f(C)+\epsilon_3$ (choisir n suffisamment grand pour que ϵ_2 soit suffisamment petit. Nous avons donc

$$f(C)=C+\epsilon_1+\epsilon_3 \tag{97.676}$$

où ϵ_1 et ϵ_3 peuvent être arbitrairement petits.

Exercise 191 exo_I-1-16

Théorème de Borel Soient $\{C_n\}$, une suite quelconque de nombres réels et f une fonction C^∞ de $\mathbb{R}$ dans $[0,1]$, égale à 1 dans $[-\frac{1}{2},\frac{1}{2}]$ et à 0 dans le complémentaire de $[-1,1]$. On considère la série

$$u(x)=\sum_{n=0}^{\infty}f\left(\frac{x}{t_n}\right)\frac{x^n}{n!}C_n. \tag{97.677}$$

(1) Montrer que par un choix convenable de la suite des nombres $t_n>0$, la série converge pour tout $x\in\mathbb{R}$ et que pour tout entier $m\geqslant 1$, la série des dérivées $m$ièmes

$$\frac{d^m}{dx^m}\left[f\left(\frac{x}{t_n}\right)\frac{x^n}{n!}\right]C_n \tag{97.678}$$

est uniformément convergente sur $\mathbb{R}$.

(2) En déduire que la fonction u est C^∞ sur $\mathbb{R}$ et que

$$\frac{d^m}{dx^m}u\bigg|_{x=0}=C_m \tag{97.679}$$

pour tout $m\in\mathbb{N}$.

corr_I-1-16

Correction of the exercise 191

Cet exercice est énoncé (de façon un peu différente) et corrigé sur le site bibmath.net [18]. La présente correction s'en inspire largement.

Nous posons

$$u_n(x)=\frac{C_n}{n!}x^nf(\frac{x}{t_n}). \tag{97.680}$$

Prouvons que la série des u_n converge normalement (ce qui est plus fort que uniformément). Pour cela, nous majorons $|u_n(x)|$ pour tout x. Si $|x/t_n|\geqslant 1$, il n'y a pas de problèmes : $u_n(x)=0$. Sinon, nous avons

$$|u_n(x)|\leqslant\frac{|C_n|}{n!}\|f\|_\infty|t_n|^n \tag{97.681}$$ EqMajorationBorel16

18. L'énoncé : http://www.bibmath.net/exercices/bde/analyse/distrib/testeno.pdf, et la correction : http://www.bibmath.net/exercices/bde/analyse/distrib/testcor.pdf

où la norme supremum $\|f\|_\infty$ existe parce que f est à support compact. Nous choisissons les nombres t_n de telle manière à avoir en même temps $t_n \leqslant 1$ et $|C_n|t_n \leqslant 1$. Pour cela, nous posons

$$t_n = \begin{cases} 1 & \text{si } |C_n| \leqslant 1 \\ 1/|C_n| & \text{si } |C_n| \geqslant 1 \end{cases} \tag{97.682}$$

Nous avons donc

$$|u_n(x)| \leqslant \|f\|_\infty \frac{1}{n!} \tag{97.683}$$

qui est le terme général d'une série convergente. La série $u = \sum_n u_n$ étant normalement convergente, elle est uniformément convergente. Afin de prouver que nous pouvons la dériver terme à termes, nous prouvons que pour tout m, la série $\sum_n u_n^{(m)}$ converge normalement. D'abord, remarquons que $\|f^{(m)}\|_\infty$ existe parce que $f^{(m)}$ est à support compact.

En utilisant la règle de Leibniz, nous trouvons la série des dérivées $m$ièmes :

$$u_n^{(m)}(x) = \sum_{l=0}^{m} \binom{l}{m} \frac{x^{n-l}}{(n-l)!} \frac{C_n}{t_n^{m-l}} f^{(m-l)}(\frac{x}{t_n}). \tag{97.684}$$

Afin de majorer cela de la même façon que nous avons trouvé la majoration (97.681), nous écrivons

$$\frac{C_n x^{n-l}}{t_n^{m-l}} = \frac{C_n}{t_n} \frac{x^{n-l}}{t_n^{n-l}} t_n^{n-m-1} \tag{97.685}$$

EqPreMajBorel

dont chacun des trois facteurs est plus petit ou égal à 1 lorsque $n > m$. Comprenons bien que nous nous fixons un degré de dérivation m, et puis que nous voulons sommer sur n la série $\sum_n u^{(k)}(x)$. Les m premier termes valent quelque chose qui ne nous intéresse pas (dans les séries, nous pouvons toujours modifier les premiers termes sans changer la convergence), tandis que les termes suivants se majorent grâce au fait que (97.685) est plus petit ou égal à 1 pour ces termes :

$$|u_n^{(m)}(x)| \leqslant \sum_{l=0}^{m} \binom{l}{m} \|f^{(k-l)}\|_\infty \frac{1}{(n-l)!}. \tag{97.686}$$

Il suffit de poser $M_1 = \max\{\|f^{(m-l)}\|_\infty\}_{l=0,\dots,m}$ et $M_2 = \max\{\binom{l}{m}\}_{l=0,\dots,m}$ pour trouver la borne

$$|u_n^{(m)}(x)| \leqslant k M_1 M_2 \frac{1}{(n-m)!}, \tag{97.687}$$

dont la série converge. Ce que nous avons donc prouvé est que

$$\|u_n^{(m)}\|_\infty \leqslant M_n^{(m)} \tag{97.688}$$

où la somme $\sum_n M_n^{(m)}$ converge. La série $\sum_n u_n^{(m)}$ converge donc normalement et uniformément (voir remarque 1 de la page I.17).

Nous pouvons donc dériver terme à terme. Il reste à déterminer la valeur de la $m$ième dérivée en $x = 0$. Fixons un ordre m de dérivation et prouvons que pour tout $\epsilon > 0$, nous avons $|u^{(m)}(0) - C_m| \leqslant \epsilon$. Soit donc $\epsilon > 0$ et prenons un entier $N > m$ tel que pour tout x,

$$\left| \sum_{n \geqslant N} u_n^{(m)}(x) \right| \leqslant \epsilon. \tag{97.689}$$

Maintenant, nous considérons un voisinage de zéro dans lequel

$$u_n^{(m)}(x) = C_n \frac{x^{n-m}}{(n-m)!} \tag{97.690}$$

pour tout $n < N$. À ce moment, nous avons pour tout x dans ce voisinage

$$\left| u^{(m)}(x) - \sum_{n<N} \frac{x^{n-m}}{(n-m)!} C_n \right| = \left| \sum_{n \geqslant N} u_n^{(m)}(x) \right| \leqslant \epsilon. \tag{97.691}$$

En particulier, en $x = 0$ nous avons

$$|u^{(m)}(0) - C_m| \leqslant \epsilon. \tag{97.692}$$

Nous en déduisons que $u^{(m)}(0) = C_m$.

Note Dans le corrigé que j'ai suivit, il est dit que sur un voisinage de zéro, $u_n(x) = \frac{C_n}{n!}x^n$. Il me semble que cela n'est pas toujours vrai parce que si $C_n \to \infty$, alors $t_n \to 0$ et tout voisinage de zéro contiendra un x tel que pour un certain n, nous ayons $|x/t_n| \in [\frac{1}{2}, 1]$ sur lequel f ne vaut pas 1. C'est pour cela que je fais appel à un N et un ϵ afin de prouver que la différence entre $u_n(x) = \frac{C_n}{n!}x^n$ et la réalité n'est pas si grande que ça.

Après avoir fait le théorème de Borel, essayons de voir ce qu'il dit. Le théorème de la page I.29 nous donne facilement, sous forme de séries de puissances, une fonction C^∞ sur un intervalle dont les dérivées en zéro sont données à l'avance. L'exercice que nous venons de faire nous permet de trouver une fonction C^∞ sur tout $\mathbb{R}$ dont les dérivées sont données à l'avance.

97.20 Existence d'intégrales

En vertu des différents théorèmes, l'étude de la convergence et de l'existence d'une intégrale d'une fonction sur $[a, \infty[$ se fait dans l'ordre suivant :

— si l'intégrale de f existe, alors elle converge,

— si la fonction est positive et que son intégrale n'existe pas, alors elle ne converge pas,

— si le signe de la fonction est variable et que l'intégrale n'existe pas, alors elle peut converger ou non, selon les cas.

Exercise 192 exo_I-2-1

Utiliser les critères des fonctions tests pour étudier l'existence des intégrales suivantes : ExoPointPremierIdeuxUn

(1) $\int_0^\infty (1 + 2x^2)^{-1/2}\, dx$,

(2) $\int_0^\infty P(x)e^{-x^2}dx$ où $P(x)$ est un polynôme,

(3) $\int_2^\infty x^\alpha e^{-x} \ln^\beta x\, dx$, en fonction de α et β, ItemDExoI21

(4) $\int_1^2 \frac{\sqrt{x}}{\ln x}dx$,

(5) $\int_0^1 \frac{dx}{e^x-1}$,

(6) $\int_2^\infty \frac{\ln(x)}{(1+x^3)^{1/p}}dx$, en fonction de p,

(7) $\int_2^\infty \frac{dx}{x^2-1}$.

corr_I-2-1

Correction of the exercise 192

Nous donnons systématiquement deux preuves. La première utilise les fonctions test, tandis que la seconde utilise une majoration (ou minoration) explicite qui donne le résultat. Notez qu'à chaque fois, les deux méthodes reviennent au même.

(1) Intuitivement, le résultat est clair parce que, pour des grands x, la fonction $f(x) = (1 + 2x^2)^{-1/2}$ se comporte comme $1/x$ (compter les degrés de x), alors que $1/x$ ne s'intègre pas vers l'infini. La réponse est donc que l'intégrale ne converge pas, et il faudra la comparer à une fonction de type $1/x$.

En effet, nous utilisons le corolaire 20.94 avec $\alpha = 1$:

$$L = \lim_{x\to\infty} \frac{x}{\sqrt{1 + 2x^2}} = \frac{1}{\sqrt{2}} \neq 0. \tag{97.693}$$

L'intégrale n'existe donc pas.

Justification alternative Nous cherchons directement une fonction de type a/x qui majore $f(x)$ pour les grands x. Posons $g(x) = a/x$. Nous avons

$$f(x) - g(x) = \frac{x - a\sqrt{1+2x^2}}{x\sqrt{1+2x^2}}, \tag{97.694}$$

dont le dénominateur est toujours positif ou nul dans l'ensemble considéré. Dès que $a < 1/\sqrt{2}$, il existe un x_0 tel que $\forall x > x_0$, nous avons $f(x) > g(x)$. Pour prouver cela, remarquez que

$$\lim_{x\to\infty} \left(x - a\sqrt{1+2x^2}\right) = \lim_{x\to\infty} x(1 - \sqrt{2}a) = \pm\infty \tag{97.695}$$

où le $\pm$ est fixé par la valeur de a. Nous utilisons le théorème 20.196 pour conclure à la non existence.

(2) Nous avons $\lim_{x\to\infty} x^2 P(x)e^{-x^2} = 0$, d'où nous concluons à l'existence de l'intégrale.
Justification alternative Sans perte de généralité, nous pouvons supposer que $P(x)$ est positif. En effet, pour x assez grand, un polynôme ne change plus de signe, et nous pouvons éventuellement considérer $-P(x)$ au lieu de $P(x)$. Soit Q, un polynôme de degré deux plus haut que P, que nous prenons positif. Une propriété de l'exponentielle est que $\exists x_0 > 0$ tel que $x > x_0$ implique $e^{-x^2} > Q(x)$ et $Q(x) > P(x)$, et donc

$$\frac{P(x)}{Q(x)} > P(x)e^{-x^2} \tag{97.696}$$

pour $x > x_0$. Mais, la fraction rationnelle P/Q est intégrable sur $[x_0, \infty[$ parce qu'elle peut être majoré par une fonction de la forme a/x^2, comme dans l'exercice (1). Le théorème 20.196 conclut à l'intégrabilité de f sur $[0, \infty[$.

(3) Pour tout α et β, nous avons

$$\lim_{x\to\infty} x^\alpha e^{-x} \ln(x)^\beta = 0, \tag{97.697}$$

d'où l'existence de l'intégrale s'ensuit.
Justification alternative La fonction $f(x) = x^\alpha e^{-x} \ln^\beta(x)$ est positive sur l'ensemble considéré. Étant donné que $\ln(x) < x$ (pour x assez grand), lorsque $\beta > 0$, nous pouvons majorer $f(x)$ par $e^{-x}x^{\alpha+\beta}$. Une fois de plus, l'exponentielle fait son travail : la fonction $g(x) = e^{-x}x^{\alpha+\beta}$ est intégrable quel que soient α et β.
Si $\beta < 0$, alors $\ln(x)^\beta$ est même majorée par 1, et c'est la conclusion est encore plus simple. ItemDCorrI21

(4) Le changement de variable $t = 1/\ln(x)$ mène à étudier l'existence de

$$\int_{1/\ln(2)}^{\infty} t^{-1}e^{-1/2t}dt. \tag{97.698}$$

Étant donné que $\lim_{t\to\infty} t^1 t^{-1} e^{-1/2t} = 1$, cette intégrale n'existe pas.
Justification alternative La fonction $f(x) = \sqrt{x}/\ln(x)$ est bornée sur tout intervalle de la forme $[1+\epsilon, 2]$; nous ne sommes donc intéressés que par l'intégrale sur $[1, 1+\epsilon]$. Sur cet intervalle, nous avons $\ln(x) < x - 1$, et donc

$$\frac{\sqrt{x}}{\ln(x)} > \frac{\sqrt{x}}{x-1} > \frac{1}{2}\frac{1}{x-1} \tag{97.699}$$

où $\frac{1}{2}$ est une minoration de $\sqrt{x}$ entre 1 et $1+\epsilon$. L'intégrale de $1/(x-1)$ ne converge pas en $x = 1$.
Notez qu'un changement de variable $t = \ln(x)$ fait tout aussi bien le travail : nous tombons sur

$$\int_0^{\ln(2)} \frac{1}{t}e^{3t/2}dt \tag{97.700}$$

dans laquelle $e^{3t/2}$ peut être minorée par 1, alors que l'intégrale de $1/x$ en $x = 0$ n'existe pas.

(5) Le changement de variable $u = 1/(e^x - 1)$ donne

$$\int_{e-1}^{\infty} \frac{1}{1+u} du, \tag{97.701}$$

mais $\lim_{u\to\infty} u/(1+u) = 1$, donc cette intégrale n'existe pas.

Justification alternative Lorsque x est proche de zéro et si a est suffisamment négatif, nous avons $e^x - 1 < ax$. Donc, nous avons la majoration

$$\frac{1}{e^x - 1} > -\frac{1}{ax}, \tag{97.702}$$

alors que la fonction $1/x$ n'est pas intégrable.

(6) Prenons α quelconque et calculons

$$L_p(\alpha) = \lim_{x\to\infty} x^\alpha \frac{\ln(x)}{(1+x^3)^{1/p}} = \lim_{x\to\infty} x^{\alpha-\frac{3}{p}} \ln(x) = \begin{cases} \infty & \text{si } \alpha - 3/p \geqslant 0 \\ 0 & \text{si } \alpha - 3/p < 0 \end{cases} \tag{97.703}$$

L'intégrale existera si $\alpha > 1$ et si $L \neq \infty$, c'est-à-dire si $\alpha < 3/p$. L'intégrale existe donc si on a un α tel que $1 < \alpha < \frac{3}{p}$. Au contraire, l'intégrale n'existera pas si il existe un α tel que $\frac{3}{p} \leqslant 1 \leqslant \alpha \leqslant 1$. En d'autres termes, l'intégrale existe si $\frac{3}{p} > 1$, et n'existe pas si $\frac{3}{p} \leqslant 1$.
En définitive, l'intégrale existera si et seulement si $p \in]0, 3[$.

Justification alternative Prouvons que l'intégrale

$$\int_2^{\infty} \frac{\ln(x)}{x^{1+\epsilon}} \tag{97.704}$$

converge pour tout $\epsilon > 0$. Pour tout $\alpha > 0$, nous avons

$$\lim_{x\to\infty} \frac{\ln(x)}{x^\alpha} = \lim_{x\to\infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0, \tag{97.705}$$

donc la fonction $x \mapsto \ln(x)/x^\alpha$ est bornée pour tout $\alpha > 0$. Maintenant,

$$\frac{\ln(x)}{x^{1+\epsilon}} = \frac{\ln(x)}{x^{\epsilon/2}} \cdot \frac{1}{x^{1+\epsilon/2}} \tag{97.706}$$

où le premier terme peut être majoré par une certaine constante M. Par conséquent,

$$\frac{\ln(x)}{x^{1+\epsilon}} < M \frac{1}{x^{1+\epsilon/2}}, \tag{97.707}$$

dont le membre de droite a une intégrale qui converge sur $[2, \infty[$ pour tout $\epsilon > 0$. Nous avons donc existence de l'intégrale lorsque $3/p > 1$.

(7) Prenons $\alpha = \frac{3}{2} > 1$. Nous avons

$$\lim_{x\to\infty} \frac{x^{\frac{3}{2}}}{x^2 - 1} = 0, \tag{97.708}$$

donc l'intégrale existe.

Justification alternative Il est évident que pour tout M, l'intégrale $\int_2^M f(x)dx$ ne pose pas de problèmes. Nous sommes donc en train de chercher une fonction g qui majore f pour les grands x. Afin d'avoir la bonne puissance au dénominateur, il faut chercher $g(x) = a/x^2$. On vérifie que $a > 1$ fait l'affaire. En effet, la quantité

$$\frac{a}{x^2} - \frac{1}{x^2 - 1} = \frac{x^2(a-1) - a}{x^2(x-1)(x+1)}. \tag{97.709}$$

est positive pour tout x assez grand dès que $a > 1$.

Exercise 193 exo_I-2-2
Soit

$$I = \int_1^\infty \frac{\sin(x)}{x^\beta} dx. \tag{97.710}$$

Montrer que

(1) I existe si $\beta > 1$,
(2) I ne converge pas si $\beta \leqslant 0$ (utiliser le critère de Cauchy pour montrer que la suite $I_n = \int_1^{n\pi} n^{-\beta} \sin(x)dx$ ne converge pas),
(3) I converge mais n'existe pas pour $0 < \beta \leqslant 1$ (s'inspirer de la démonstration du cours p.I.60).

corr_I-2-2

Correction of the exercise 193

(1) Si $\beta > 1$, nous devons étudier l'intégrale $\int_1^\infty \left|\frac{\sin(x)}{x^\beta}\right| dx$. La fonction $g(x) = 2/x^\beta$ majore l'intégrante alors que l'intégrale $\int_1^\infty g(x)dx$ converge.
(2) Si $\beta < 0$, calculons la différence entre deux termes successifs de la suite

$$I_n = \int_1^{n\pi} x^{-\beta} \sin(x)dx. \tag{97.711}$$

EqInACalcDiff

Selon la définition (20.352) de la convergence d'une intégrale et le fait qu'une suite numérique ne peut converger que si elle est de Cauchy, pour que I converge, la différence $I_{n+1} - I_n$ doit tendre vers zéro lorsque n tend vers l'infini. (voir le théorème de la page 46 de première année : dans $\mathbb{R}$ ou $\mathbb{R}^n$, une suite est convergente si et seulement si elle est de Cauchy) Supposons pour simplifier que n est pair (donc $\sin(x)$ est positive)

$$I_n - I_{n+1} = \int_{n\pi}^{(n+1)\pi} \frac{\sin(x)}{x^\beta} dx > (n\pi)^{-\beta} \int_{n\pi}^{(n+1)\pi} \sin(x)dx = 2(n\pi)^{-\beta}. \tag{97.712}$$

Cela ne converge pas vers zéro quand $n \to \infty$. Donc l'intégrale I ne converge pas quand $\beta \leqslant 0$. Elle n'existe donc pas non plus.

Pour résumer, nous avons toutes les situations possibles :

— intégrale qui existe (et donc converge),
— intégrale qui n'existe pas et qui converge [19]
— intégrale qui ne converge pas (et qui n'existe donc pas).

Exercise 194 exo_I-2-3
Étudier l'existence de

$$\int_a^b \frac{P(x)}{Q(x)} dx \tag{97.713}$$

où $P(x)$ et $Q(x)$ sont des polynômes sans zéros communs, $Q(a) = 0$, $Q(x) \neq 0$ pour $x \in]a, b]$.

corr_I-2-3

Correction of the exercise 194

L'intégrante est bornée sur l'intervalle $[a + \epsilon, b]$, donc nous nous bornons [20] à étudier l'intégrale

$$\int_a^{a+\epsilon} \left|\frac{P(x)}{Q(x)}\right| dx. \tag{97.714}$$

Étant donné que $Q(a) = 0$, il existe un entier $q \geqslant 1$ et un polynôme $S(x)$ ne s'annulant pas en a tels que

$$Q(x) = (x - a)^q S(x). \tag{97.715}$$

19. La terminologie « existe » provient du cas de dimension plus que un, et n'est effectivement pas très heureuse ici.
20. Il n'y a pas que les intégrales qui sont bornées.

Par continuité, si ϵ est assez petit, le polynôme S ne s'annule pas sur $[a, a+\epsilon]$. Si M est une minoration de $|P(x)/S(x)|$ sur l'intervalle considéré, nous avons

$$\int_a^{a+\epsilon} \left|\frac{P(x)}{Q(x)}\right| dx = \int_a^{a+\epsilon} \frac{|P(x)|}{(x-a)^q|S(x)|} dx \geqslant M \int_a^{a+\epsilon} \frac{1}{(x-a)^q} dx. \tag{97.716}$$

Étant donné que, par construction, $q \geqslant 1$, cette dernière intégrale n'existe jamais.

Exercise 195 exo_I-2-4

Lorsqu'un point matériel pesant est assujetti à décrire une courbe polie d'équations

$$s \mapsto \big(x(s), y(s), z(s)\big) \tag{97.717}$$

où s est la longueur d'arc, le mouvement est décrit par (expression du temps de parcours) :

$$t - t_0 = \int_{s_0}^{s} \frac{ds}{\sqrt{2g(h - z(s))}} \tag{97.718}$$

où g est l'accélération de la pesanteur et h une constante. Étudier l'existence de l'intégrale

$$\int_{s_0}^{s_1} \frac{ds}{\sqrt{2g(h - z(s))}} \tag{97.719}$$

lorsque $z(s_1) = h$, en supposant que $z(s)$ est de classe C^2. corr_I-2-4

Correction of the exercise 195

Nous pouvons supposer que $z(s) \neq h$ en dehors de $s = s_1$, et nous faisons le changement de variable $\epsilon = s_1 - s$. Nous obtenons donc l'intégrale

$$\int_{s_1 - s_0}^{0} \frac{-d\epsilon}{\sqrt{2g(h - z(s_1 - \epsilon))}} \tag{97.720}$$

dans laquelle nous remplaçons (en vertu du théorème 1, page 206 du cours de première)

$$z(s_1 - \epsilon) = z(s_1) + a\epsilon + b\epsilon^2 + o(\epsilon^2) \tag{97.721}$$

où $a = z'(s_1)$ et $b = z''(s_1)^2/2$. Lorsque $a \neq 0$ et que ϵ est petit,

$$\frac{1}{\sqrt{2g(a\epsilon + b\epsilon^2 + o(\epsilon^2))}} < \frac{M}{\sqrt{\epsilon}} \tag{97.722}$$

pour une certaine constante M. Nous déduisons donc que l'intégrale existe lorsque la dérivée première de $z(s)$ en s_1 n'est pas nulle. Pour le même genre de raisons, l'intégrale n'existe pas quand $f'(s_1) = 0$, ce qui correspond à une situation où la courbe $z(s)$ « s'arrête » sur la valeur h.

97.21 Fonctions définies par des intégrales

Exercise 196 exo_I-3-1

Soient f une fonction continue et

$$F(x) = \int_{a-x}^{a+x} f(t)dt. \tag{97.723}$$

Calculer $\frac{dF}{dx}$. corr_I-3-1

Correction of the exercise 196

Nous vérifions facilement que toutes les hypothèses de la proposition 14.321 sont satisfaites avec $\partial f/\partial x = 0$, $\varphi(x) = a - x$ et $\psi(x) = a + x$. La formule donne alors

$$\frac{dF}{dx} = f(a+x) + f(a-x). \tag{97.724}$$

Une autre possibilité est de séparer l'intégrale en deux parties :

$$F(x) = \int_{a-x}^{a} f(t)dt + \int_{a}^{a+x} f(t)dt \tag{97.725}$$

et d'appliquer le théorème fondamental de l'analyse :

$$\begin{aligned}\frac{dF}{dx} &= \frac{d}{d(a+x)}\int_a^{a+x} f(t)dt + \frac{d}{d(a-x)}\left(-\int_a^{a-x} f(t)dt\right)(-1)\\ &= f(a+x) + f(a-x).\end{aligned} \tag{97.726}$$

Le dernier signe moins vient du changement de variable $x \mapsto a - x$ fait pour calculer la dérivée.

Exercise 197 exo_I-3-2

Soit f, une fonction continue sur $[a, b]$. Montrer que la solution $y(x)$ de l'équation différentielle

$$\frac{d^n y}{dx^n} = f(x) \tag{97.727}$$

telle que $y(a) = y'(a) = \ldots = y^{\underline{n-1}}(a) = 0$, peut s'écrire

$$y(x) = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)dt. \tag{97.728}$$

EqSolProposEqDiffEx

corr_I-3-2

Correction of the exercise 197

Lorsque $n = 1$, la solution proposée (97.728) devient

$$y(x) = \int_a^x f(t)dt, \tag{97.729}$$

donc $y'(x) = f(x)$ et $y(a) = 0$. Il faut maintenant étudier le cas où $n > 1$. En utilisant la formule (14.1012), nous trouvons

$$\begin{aligned}\frac{dy}{dx} &= \int_a^x \frac{\partial}{\partial x}\left(\frac{(x-t)^{n-1}}{(n-1)!}f(t)\right)dt + \underbrace{\left[\frac{(x-t)^{n-1}}{(n-1)!}f(t)dt\right]_{t=x}}_{=0}\cdot 1\\ &= \int_a^x \frac{(x-t)^{n-2}}{(n-2)!}f(t)dt.\end{aligned} \tag{97.730}$$

Si $n = 2$, nous avons donc $\frac{dy}{dx} = \int_a^x f(t)dt$, ce qui fait $y''(x) = f(x)$ et $y(a) = y'(a) = 0$. Par récurrence,

$$\frac{d^p y}{dy^p} = \int_a^x \frac{(x-t)^{n-(p+1)}}{\big(n-(p+1)\big)!}f(t)dt, \tag{97.731}$$

et en particulier,

$$\frac{d^{n-1}y}{dx^{n-1}}\int_a^x f(t)dt, \tag{97.732}$$

ce qui donne immédiatement $y^{\underline{n}}(x) = f(x)$ et $y(a) = y'(a) = \ldots = y^{\underline{n-1}}(a) = 0$.

97.22 Convergence, continuité et dérivation sous le signe intégral

Exercise 198 exo_I-3-4

Étudier la convergence uniforme des intégrales suivantes :

(1) $\int_1^\infty \frac{\sin(x)}{x^2}dx$,

(2) $\int_0^\infty xe^{-xt}dt$ pour $x \in [0, 1]$,

(3) $\int_0^1 \big(\ln(xt)\big)^{1/3}dt$ pour $x \in [1,3]$,

(4) $\int_0^\infty \frac{\sin(xt)}{1+t^2}dt$,

(5) $\int_0^\infty \frac{t\cos(xt)}{1+t^2}dt$,

(6) $\int_0^\infty \frac{\sin(xt)}{t}dt$.

corr_I-3-4

Correction of the exercise 198

Nous allons utiliser le critère de Weierstrass, théorème 17.12.

(1) La fonction $t \mapsto \frac{10}{t^2}$ majore $|\sin(xt)/t^2|$, or l'intégrale $\int_0^\infty \frac{10}{t^2}dt$ existe.

(2) Le théorème 17.17 affirme que si l'intégrale était uniformément convergente, alors la fonction

$$I(x) = \int_0^\infty xe^{-xt}dt \tag{97.733}$$

serait continue. Nous allons vérifier cela. D'abord, si $x = 0$, nous avons évidement $I(0) = 0$. Si $x \neq 0$, alors

$$I(x) = \lim_{T\to\infty} \int_0^T xe^{-xt}dt = \lim_{T\to\infty} \big(-e^{-xT} + 1\big) = 1. \tag{97.734}$$

Cette fonction I n'est donc pas continue en 0 et il ne peut donc pas y avoir de convergence uniforme sur l'intervalle $[0,1]$.

Ce résultat peut également être vu directement sur la définition. Pour avoir uniforme convergence, il faudrait que $\forall \epsilon > 0$ $\exists T_0$ (indépendant de x) tel que $T \geqslant T_0$ implique

$$\left|\int_T^\infty xe^{-xt}dt\right| < \epsilon \tag{97.735}$$

pour tout x. Or, nous savons que

$$\int_T^\infty xe^{-xt}dt = \begin{cases} e^{-xT} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0, \end{cases} \tag{97.736}$$

donc $\sup_{x\in[0,1]} \int_0 1^\infty xe^{-xt}dt = 1$, et un T_0 convenable ne peut pas être trouvé.

Nous avons toutefois convergence uniforme sur tout compact de $]0,1]$ parce que si x_0 est le minimum du compact, alors $T_0 = -\ln(\epsilon)/x_0$ fonctionne.

(3) Nous pouvons majorer la norme de $\ln(xt)$ de façon indépendante de x de la façon suivante :

$$|\ln(xt)| = |\ln(x) + \ln(t)| \leqslant |\ln(x)| + |\ln(t)| \leqslant \ln(3) - \ln(t) = \ln(3/t) \tag{97.737}$$

où nous avons utilisé le fait que $|\ln(t)| = -\ln(t)$ dans l'intervalle considéré. Maintenant, pour $t \in [0,1]$, nous avons $\ln(3/t) < 3/t$, et au final, nous avons

$$|\ln(xt)|^{1/3} < \frac{3^{1/3}}{t^{1/3}}, \tag{97.738}$$

tandis que

$$\int_0^1 \frac{3^{1/3}}{t^{1/3}} \tag{97.739}$$

existe.

(4) La norme $|\sin(xt)/(1+t^2)|$ se majore par $1/(1+t^2)$ dont l'intégrale existe.

(5) L'intégrale n'existe pas en $x = 0$ parce que l'intégrale de $t/(1 + t^2)$ ne converge pas. Notez la différence avec l'exercice précédent où l'absence d'un t au numérateur faisait converger l'intégrale. En ce qui concerne le cas $x \neq 0$, nous utilisons le critère d'Abel avec $\varphi(x, t) = \cos(xt)$ et $\psi(x, t) = 1/(1 + t^2)$.
Nous avons

$$|\int_0^T \cos(xt)dt| = \frac{1}{|x|}\sin(xT) \leqslant \frac{1}{|x|}. \tag{97.740}$$

Si x est restreint à un compact de ne contenant pas 0, alors cette quantité peut être bornée uniformément en x par un certain nombre M. D'autre part, nous avons que $\lim_{t\to\infty} \psi(x, t) = 0$ de façon uniforme en x (parce que x n'apparaît même pas dans la fonction).
Le critère d'Abel conclut à l'uniforme convergence sur tout compact ne contenant pas 0.

(6) Si nous posons

$$F(x) = \int_0^\infty \frac{\sin(xt)}{t}dt, \tag{97.741}$$

nous trouvons $F(0) = 0$ et, via le changement de variable $u = xt$: $F(x) = F(1)$ quand $x > 0$, et $F(x) = F(-1)$ pour $x < 0$. Donc, si nous voulons que l'intégrale converge uniformément, il faut que $F(1) = F(-1) = 0$. Nous verrons cependant que

$$\int_0^\infty \frac{\sin(t)}{t}dt = \frac{\pi}{2} \tag{97.742}$$

EqIntSinSurt

dans lemme 2, page III.21 (attention : preuve difficile). L'intégrale ne convergera donc uniformément sur aucun intervalle contenant zéro.
Sans utiliser le résultat (97.742), nous pouvons utiliser le critère d'Abel avec

$$\varphi(x, t) = \sin(xt), \quad \psi(x, t) = \frac{1}{t}. \tag{97.743}$$

Nous obtenons facilement que

$$|\int_0^T \varphi(x, t)dt| \leqslant \frac{1}{|x|}M \tag{97.744}$$

si M est choisi assez grand. Cette quantité peut être majorée de façon uniforme par rapport à x quand x est restreint à un compact ne contenant pas zéro. Le critère d'Abel fournit donc la convergence uniforme sur tout compact ne contenant pas zéro.

Exercise 199 exo_I-3-5
Sur l'intervalle $]0, 1[$ la fonction

$$F(x) = \int_0^1 \frac{e^t}{\sin(xt)^{1/3}}\, dt \tag{97.745}$$

est-elle continue ? Est-elle C^1 ? Justifier. corr_I-3-5

Correction of the exercise 199
Nous considérons la fonction

$$f(x, t) = \frac{e^t}{\sin(xt)^{1/3}}. \tag{97.746}$$

La stratégie que nous allons suivre est la suivante :

(1) montrer que $\int_0^1 f(x, t)dt$ existe $\forall x \in]0, 1[$,
(2) on montre que $F(x)$ est continue par le coup du compact pour $x \in]0, 1[$,
(3) pour étudier la dérivabilité, on pose

$$G(x) = \int_0^1 \frac{\partial f}{\partial x}(x, t)dt, \tag{97.747}$$

dont on prouve d'abord l'existence et ensuite l'uniforme convergence.

D'abord, montrons que pour tout $x \in]0,1]$, l'intégrale de la fonction $f_x : t \mapsto f(x,t)$ par rapport à t existe. Pour cela nous allons utiliser le corolaire 20.95. En nous souvenant que la page 191 du cours de première dit que $\lim_{x\to 0} \sin(x)/x = 1$, nous avons

$$\lim_{t\to 0} t^{1/3} f(x,t) = \lim_{t\to 0} \left(\frac{xt}{\sin(xt)} \right)^{\frac{1}{3}} \frac{e^t}{x^{1/3}} = \frac{1}{x^{1/3}} < \infty. \tag{97.748}$$

Nous appliquons le corolaire 20.95 chaque fonction $f_x(t) = f(x,t)$, dont nous concluons que pour tout $x \neq 0$,

$$\int_0^1 f_x(t)dt \tag{97.749}$$

existe. Ceci montre que la fonction $F(x)$ est bien définie. Nous étudions maintenant sa continuité. Nous faisons le coup du compact. Soit donc un compact de $]0,1]$ dont le minimum est $x_0 > 0$, nous avons la majoration

$$f(x,t) \leqslant \frac{e^t}{\sin(x_0 t)^{1/3}}, \tag{97.750}$$

dont l'intégrale existe : ce n'est autre que $F(x_0)$ dont nous venons de prouver l'existence. Par le critère de Weierstrass, l'intégrale $\int_0^1 f(x,t)\,dt$ est uniformément convergente sur le compact considéré. Par le coup du compact et le théorème 17.17, la fonction F est alors continue sur l'ensemble $]0,1]$.

Nous pouvons maintenant étudier la dérivabilité sous le signe intégral de F. Pour cela, considérons

$$G(x) = \int_0^1 \frac{\partial f}{\partial x}(x,t)dt = -\int_0^1 \frac{te^t}{3} \frac{\cos(xt)}{\sin(xt)^{4/3}}. \tag{97.751}$$ EqIntPartialAvUtCorr35

Nous allons prouver que cette intégrale existe et qu'elle converge uniformément en x sur tout compact de $]0,1$. Ainsi, elle sera la dérivée de F et continue, ce qui prouve que F est C^1.

Pour prouver l'existence, nous procédons comme pour F :

$$\lim_{t\to 0} t^{1/3} \left| \frac{\partial f}{\partial x} \right| = \frac{e^t}{3} \frac{\cos(xt)}{x^{4/3}} \left(\frac{xt}{\sin(xt)} \right)^{\frac{4}{3}} = \frac{1}{3x^{4/3}}. \tag{97.752}$$

Nous en déduisons que l'intégrale (97.751) existe pour tout $x \neq 0$.

Maintenant, nous voudrions bien borner $\frac{cos(xt)}{\sin(xt)^{4/3}}$ de façon indépendante de x pour $x \in K \subset]0,1[$. Nous serions donc content que cette fonction soit soit croissante soit décroissante sur $]0,1[$. En réalité, elle est décroissante parce que si nous posons

$$l(y) = \frac{\cos(y)}{\sin(y)^{4/3}}, \tag{97.753}$$

alors

$$l'(y) = -\frac{\sin(y)\cos(y)^{4/3} + \frac{4}{3}\sin(y)^{1/3}\cos(y)}{\sin(y)^{8/3}}, \tag{97.754}$$

qui est négatif tant que $y \in [0, \frac{\pi}{2}]$. Nous pouvons donc majorer la dérivée de f comme ceci :

$$\left| \frac{\partial f}{\partial x} \right| (x,t) \leqslant \frac{te^t}{3} \frac{\cos(x_0 t)}{\sin(x_0 t)^{4/3}}. \tag{97.755}$$

L'intégrale du membre de droite existe : c'est $G(x_0)$ dont nous venons de prouver l'existence. Donc $G(x)$ converge uniformément sur tout compact de $]0,1]$. Cela prouve deux choses. Premièrement, $G(x)$ est la dérivée de $F(x)$, et secondement G est continue. Ces deux points sont exactement le fait que F est C^1 sur $]0,1]$.

Exercise 200 exo_1-3-6 Sachant que pour $x > 0$,

$$\int_0^\infty \frac{x}{x^2+t^2}dt = \frac{\pi}{2}, \tag{97.756}$$ EqPredicI36

montrer que pour $x > 0$,

$$\int_0^\infty \frac{dt}{(x^2+t^2)^2} = \frac{\pi}{4x^3}. \tag{97.757}$$

corr_I-3-6

Correction of the exercise 200

L'intégrale peut être calculée explicitement pour chaque x :

$$F(x) = \int_0^\infty \frac{x}{x^2+t^2}dt = \lim_{T\to\infty}\int_0^T \frac{x}{x^2+t^2}dt = \lim_{T\to\infty}\left[\arctan\frac{t}{x}\right]_0^T. \tag{97.758}$$

D'où nous déduisons

$$F(x) = \begin{cases} \frac{\pi}{2} & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -\frac{\pi}{2} & \text{si } x < 0 \end{cases} \tag{97.759}$$

Voyons maintenant comment se comporte la dérivée de cette fonction. Pour cela, nous regardons l'intégrale de la dérivée, et nous allons prouver sa convergence uniforme sur tout compact. Nous avons $\partial_x\left(\frac{x}{x^2+t^2}\right) = (t^2-x^2)/(x^2+t^2)^2$, donc

$$\left|\frac{\partial}{\partial x}\left(\frac{x}{x^2+t^2}\right)\right| \leqslant \frac{t^2+x^2}{(x^2+t^2)^2} \leqslant \frac{1}{x^2+t^2} \tag{97.760}$$

et l'intégrale de cette dernière expression converge uniformément sur tout compacts de $\mathbb{R}\backslash\{0\}$. Mais sur chacun de ces compacts, $F(x)$ est constante, c'est-à-dire que pour tout $x \neq 0$,

$$G(x) = \int_0^\infty \frac{\partial}{\partial x}\left(\frac{x}{x^2+t^2}\right)dt = \frac{dF}{dx} = 0. \tag{97.761}$$

Par ailleurs, $G(x)$ peut être écrite en termes de $F(x)$ et de l'intégrale que l'on cherche :

$$0 = G(x) = \int_0^\infty \frac{x^2+t^2-2x^2}{(x^2+t^2)^2}dt = \frac{1}{x}F(x) - 2x^2\int_0^\infty \frac{dt}{(x^2+t^2)^2}. \tag{97.762}$$

Donc pour $x \neq 0$,

$$\int_0^\infty \frac{dt}{(x^2+t^2)^2} = \frac{F(x)}{2x^3} = \begin{cases} \frac{\pi}{4x^3} & \text{si } x > 0 \\ -\frac{\pi}{4x^3} & \text{si } x < 0. \end{cases} \tag{97.763}$$

Exercise 201 exo_I-3-7

Sachant que pour $m \geqslant 0$,

$$J(m) = \int_0^\infty \frac{\cos(mx)}{1+x^2}dx = \frac{\pi}{2}e^{-m}, \tag{97.764}$$

calculer pour $m \geqslant 0$,

$$I(m) = \int_0^\infty \frac{x\sin(mx)}{1+x^2}dx, \tag{97.765}$$

en justifiant les différentes étapes du calcul.

corr_I-3-7

Correction of the exercise 201

Le développement de cet exercices est semblable à celui de l'exercice 200. D'abord, nous étudions la convergence de l'intégrale de la dérivée

$$\frac{\partial}{\partial m}\left(\frac{\cos(mx)}{1+x^2}\right) = -\frac{x\sin(mx)}{1+x^2}. \tag{97.766}$$

Pour cela, nous utilisons le critère d'Abel avec

$$\begin{aligned}\varphi(m,x) &= -\sin(mx)\\ \psi(m,x) &= \frac{x}{1+x^2}.\end{aligned} \tag{97.767}$$

Une primitive de $\sin(mx)$ est donnée par $-\frac{\cos(mx)}{m}$, donc

$$\int_0^X \sin(mx)dx = \frac{1}{m}\big(1-\cos(mX)\big) \tag{97.768}$$ EqPrimInt1378

qui ne peut pas être bornée indépendamment de m. Le critère d'Abel ne peut donc être appliqué que dans un compact. Nous faisons donc à nouveau le coup du compact, et nous appliquons Abel sur chaque compact (en la variable m, pas x ! !) de $]0,\infty[$. Le cas $m=0$ devra donc être traité à part. Si m est borné vers le bas par m_0, alors nous pouvons borner (97.768) par

$$\Big|\int_0^X \sin(mx)dx\Big| \leqslant \frac{2}{m_0}. \tag{97.769}$$

Par ailleurs, les fonctions $m\mapsto\psi(m,x)$ sont constantes (et valent $x/(1+x^2)$) et tendent uniformément vers zéro quand x tend vers l'infini. Le critère d'Abel s'applique donc sur le compact dont le minimum est m_0 et nous trouvons que l'intégrale

$$\int_0^\infty \frac{x\sin(mx)}{1+x^2}dx \tag{97.770}$$

est uniformément convergente sur tout compact. Nous en concluons que $I(m)=-\frac{\partial J}{\partial m}(m)$ est continue. Et nous avons

$$I(m)=\begin{cases}\frac{\pi}{2}e^{-m} & \text{si } x>0\\ 0 & \text{si } x=0.\end{cases} \tag{97.771}$$

Exercise 202 exo_I-3-8

La fonction

$$x\mapsto\int_\pi^\infty \frac{e^{-xt}\sin(t)}{(t-\pi)^{\frac{1}{2}}}dt \tag{97.772}$$

est-elle continue sur $]0,1[$?

corr_I-3-8

Correction of the exercise 202

Grâce à l'exponentielle décroissante, la fonction tend vers zéro à l'infini assez vite pour que l'intégrale converge de ce côté. En ce qui concerne l'éventuel problème en π, nous calculons

$$\lim_{t\to\pi}\frac{e^{-xt}\sin(t)}{\sqrt{t-\pi}} = e^{-\pi x}\lim_{t\to\pi}\frac{\sin(t)}{\sqrt{t-\pi}} = \lim_{t\to\pi}2\sqrt{t-\pi}\cos(t)=0, \tag{97.773}$$

il n'y a donc pas de singularité en π, contrairement à ce que l'on aurait pu craindre à première vue.

La fonction $f(x,t)$ sous le signe intégral est une fonction continue et bornée (comme toute fonction continue sur un compact) sur le compact $x\in[0,1]$ et $t\in[\pi,\pi+1]$. Nous la majorons par une constante sans importance. L'important est que pour l'intervalle non compact $t\in[\pi+1,\infty[$, nous pouvons faire la majoration

$$\left|\frac{e^{-xt}\sin(t)}{(t-\pi)^{\frac{1}{2}}}\right| \leqslant e^{-xt}. \tag{97.774}$$

Sur $x>0$, nous avons $\int_\pi^\infty e^{-xt}dt=\frac{e^{-\pi x}}{x}$, qui est un nombre bien défini tant que $x>0$. Le théorème 17.12 (critère de Weierstrass) conclut que $F(x)$ converge uniformément sur tout compact de $]0,1]$. Elle y est donc continue.

Exercise 203 exo_I-3-9

La fonction

$$F(x) = \int_0^\infty \frac{\sin(xt) + \cos(xt)}{1 + t^2} dt \tag{97.775}$$

est-elle continue sur $\mathbb{R}$? Est-elle dérivable sur $]0, \infty[$?

corr_I-3-9

Correction of the exercise 203

Utilisons le critère d'Abel sur tout compact. Lorsque x est restreint à parcourir un compact, nous pouvons borner

$$\int_0^T \big[\sin(xt) + \cos(xt)\big] dt \tag{97.776}$$

de façon indépendante de T et x. La fonction $\psi(x, t) = 1/(1 + t^2)$ vérifie ce qu'il faut pour Abel. La convergence uniforme sur tout compact implique la continuité de F sur $\mathbb{R}$.

Pour la dérivabilité, nous regardons l'intégrale de la dérivée :

$$G(x) = -\int_0^\infty t \frac{\sin(xt) - \cos(xt)}{t^2 + 1} dt. \tag{97.777}$$

Sur chaque compact de $]0, \infty[$, le critère d'Abel donne la convergence uniforme. Donc F est dérivable sur cet ensemble, et sa dérivée y vaut G.

Exercise 204 exo_I-3-10

Soit la **fonction d'Euler**

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt. \tag{97.778}$$

(1) démontrer que $\Gamma(x) \in C^\infty\big(]0, \infty[\big)$,

(2) prouver la relation

$$\Gamma(x + 1) = x\Gamma(x), \tag{97.779}$$ EqGammaFacto

en déduire que $\Gamma(n) = (n - 1)!$ pour tout entier $n \geqslant 1$.

(3) montrer que $\Gamma(\frac{1}{2}) = 2\int_0^\infty e^{-u^2} du$,

(4) montrer que

$$\big(\Gamma(\frac{1}{2})\big)^2 = \lim_{R\to\infty} \iint_D e^{-(u^2+v^2)} du\, dv \tag{97.780}$$

où D est un disque de rayon R centré à l'origine. En déduire que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(5) Calculer

$$\int_0^\infty e^{-t}\sqrt{t} dt. \tag{97.781}$$

(6) montrer que

$$\int_0^1 x(\ln x)^{-1/3} dx = \frac{-2^{1/3}}{2}\Gamma(\frac{2}{3}). \tag{97.782}$$

corr_I-3-10

Correction of the exercise 204

Certains résultats de cet exercice sont aussi à voir dans le cours, à la page I.78.

(1) Afin d'étudier la $n$ième dérivée de $\Gamma(x)$, nous divisons l'intégrale en deux parties :

$$I_n(x) = \int_0^1 \frac{\partial^n}{\partial x^n}\left(e^{-t}t^{x-1}\right) dt = \int_0^1 e^{-t}t^{x-1}(\ln(t))^n dt, \tag{97.783}$$

et

$$J_n(x) = \int_0^\infty \frac{\partial^n}{\partial x^n}\left(e^{-t}t^{x-1}\right) dt = \int_0^\infty e^{-t}t^{x-1}(\ln(t))^n dt, \tag{97.784}$$

avec $n \geqslant 0$. Afin d'étudier la convergence de I et de J, nous posons

$$L = \lim_{\substack{t\to 0\\ t>0}} |e^{-t}t^{x-1}(\ln t)^n|t^\alpha = \lim_{\substack{t\to 0\\ t>0}} t^{\alpha+x-1}|\ln t|^n = (-1)^n \lim_{\substack{t\to 0\\ t>0}} (\ln t)^n t^{\alpha+x-1}, \tag{97.785}$$

et nous utilisons le critère des fonctions tests, corolaire 20.94. Dès que $\alpha \leqslant 1-x$, nous avons $L = \infty$. Donc si $x \leqslant 0$, nous prenons $\alpha = 1$ et nous concluons à ce que $I_n(x)$ n'existe pas. Si, par contre, $\alpha + x - 1 > 0$, alors $L = 0$. Donc dès que $x > 0$, nous pouvons trouver un $\alpha < 1$ tel que $\alpha + x - 1 > 0$, et donc $I_n(x)$ existe quand $x > 0$.

En conclusion, $I_n(x)$ existe si et seulement si $x > 0$. Maintenant que nous savons l'existence de $I_n(x)$, nous étudions sa convergence quand $x > 0$.

Sur un compact dont le minimum est ϵ, nous avons

$$|e^{-t}t^{x-1}(\ln t)^n| < e^{-t}t^{\epsilon-1}|\ln t|^n, \tag{97.786}$$

mais l'intégrale du membre de droite entre 0 et 1 existe (c'est $I_n(\epsilon)$), donc $I_n(x)$ converge uniformément sur tout compact de $]0,\infty[$. En conséquence de quoi, nous avons

(1a) $I_n(x)$ est continue,

(1b) $I_n(x) = \frac{d^n}{dx^n}(I_0(x))$

pour tout $n \geqslant 1$. Cela prouve que $I_0(x)$ est C^∞ sur $]0,\infty[$.

Nous passons maintenant à l'étude de $J_n(x)$. En utilisant l'astuce (15.353), nous avons

$$\lim_{t\to\infty} |e^{-t}t^{x-1}(\ln t)^n|t^\alpha = 0, \tag{97.787}$$

de telle sorte que $J_n(x)$ existe. Son type de convergence est étudiée sur un compact en x dont le maximum est A. Si $t \geqslant 1$, nous avons

$$e^{-t}t^{x-1}(\ln t)^n \leqslant e^{-t}t^{A-1}(\ln t)^n. \tag{97.788}$$

(2) Le calcul de $\Gamma(x+1)$ revient au le calcul de

$$\int_0^\infty e^{-t}t^x, \tag{97.789}$$

qui peut s'intégrer par partie en posant $u = t^x$ et $dv = e^{-t}dt$. Le résultat est

$$\Gamma(x+1) = 0 + x\int_0^\infty e^{-t}t^{x-1}dt = x\Gamma(x). \tag{97.790}$$

(3) Le calcul de $\Gamma(\frac{1}{2})$ est un simple changement de variable dans l'intégrale. En posant $u = t^{1/2}$, nous trouvons

$$\Gamma(\frac{1}{2}) = \int_0^\infty e^{-t}t^{-\frac{1}{2}}dt = 2\int_0^\infty e^{-u^2}du. \tag{97.791}$$

(4) Nous avons

$$\Gamma(\frac{1}{2})^2 = \Gamma(\frac{1}{2})\Gamma(\frac{1}{2}) = 4\int_0^\infty e^{-u^2}du\int_0^\infty e^{-v^2}dv. \tag{97.792}$$

La fonction $(u,v) \mapsto e^{-(u^2+v^2)}$ étant intégrable sur $[0,\infty[\times[0,\infty[$, le théorème de Fubini 14.309, et en particulier la formule (14.993) s'appliquent et nous avons

$$I = \frac{1}{4}\Gamma(\frac{1}{2})^2 = \iint\limits_{\substack{u>0\\ v>0}} e^{-(u^2+v^2)}dudv. \tag{97.793}$$

Cette intégrale se calcule en effectuant le changement de variable polaire : $u = r\cos\theta$, $v = r\sin\theta$:

$$I = \iint\limits_{\substack{r>0\\ 0\leqslant\theta\leqslant\frac{\pi}{2}}} e^{-r^2}rdrd\theta = \int_0^{\pi/2}\left(\int_0^\infty e^{-r^2}rdr\right)d\theta = \frac{\pi}{2}\int_0^\infty \frac{1}{2}e^{-u}du = \frac{\pi}{4} \tag{97.794}$$

où nous avons effectué un changement de variable $u = r^2$ pour effectuer la dernière intégrale. La formule $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ est maintenant prouvée.

(5) Il suffit de remarquer que l'intégrale que nous devons calculer n'est autre que $\Gamma(\frac{3}{2})$. En utilisant la formule (97.779), nous trouvons

$$\int_0^\infty e^{-t}\sqrt{t}dt = \Gamma(\frac{1}{2}+1) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{\sqrt{\pi}}{2}. \tag{97.795}$$

(6) Un premier changement de variable $u = \ln x$ donne

$$I = \int_0^1 x(\ln x)^{-1/3}dx = \int_{-\infty}^0 e^u u^{-1/3}e^u du = -\int_0^\infty e^{2u}u^{-1/3}. \tag{97.796}$$

À partir de là, le changement de variable $t = -2u$ fourni la solution.

Exercise 205 exo_I-3-11

Soit la **fonction** β **d'Euler**

$$\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt. \tag{97.797}$$

(1) montrer que $\beta(x,y)$ existe si et seulement si $x > 0$ et $y > 0$

(2) montrer que la fonction β est symétrique : $\beta(x,y) = \beta(y,x)$

(3) montrer que

$$\beta(x,y) = \int_0^\infty \frac{u^{x-1}}{(1+u)^{x+y}}du \tag{97.798}$$

et en déduire la valeur de

$$\int_0^\infty \frac{u^2}{(1+u)^5}du. \tag{97.799}$$

corr_I-3-11

Correction of the exercise 205

(1) Nous allons étudier l'existence de l'intégrale proposée en utilisant le corolaire 20.95. Pour cela, remarquons d'abord que t^{x-1} peut avoir un problème en $t = 0$, tandis que $(1-t)^{y-1}$ peut avoir un problème en $t = 1$. Pour cette raison, nous étudions séparément l'intégrale sur $[0, \frac{1}{2}]$ et sur $[\frac{1}{2}, 1]$.

Pour étudier l'existence de

$$I = \int_0^{\frac{1}{2}} t^{x-1}(1-t)^{y-1}, \tag{97.800}$$

nous calculons

$$\lim_{t\to 0} t^{\alpha+x-1}(1-t)^{y-1} = \lim_{t\to 0} t^{\alpha+x-1} = L. \tag{97.801}$$

Pour avoir $L < \infty$, nous avons besoin de $\alpha < 1 - x$. Si $x > 0$, il suffit de prendre $\alpha < 1$, et nous avons $L = 0$, ce qui prouve l'existence de l'intégrale. Mais si $x \leqslant 0$, en prenant α entre 1 et $1 - x$, nous avons $L = \infty$ et $\alpha > 1$, donc pas d'existence de l'intégrale. Pour la même raison, nous trouvons que l'intégrale n'existe que si $y > 0$. En conclusion, $\beta(x,y)$ existe si et seulement si $x > 0$ et $y > 0$.

Notez que ce résultat pouvait être deviné très simplement en comptant les degrés : si $x \leqslant 0$, autour de 0, le facteur $(1-t)^{y-1}$ reste borne (pour toute valeur de y), tandis que t^{x-1} croit plus vite que $\frac{1}{t}$, ce qui donne la divergence. Dès que $x > 0$, la fonction t^{x-1} ne croit plus aussi vite que $\frac{1}{t}$ en zéro, et il y a convergence. Idem pour y.

(2) Le changement de variable $1 - t = u$ dans l'intégrale

$$\beta(y,x) = \int_0^1 t^{y-1}(1-t)^{x-1}dt = -\int_1^0 u^{x-1}(1-u)^{y-1}du = \beta(x,y). \tag{97.802}$$

(3) Pour commencer, le changement de variable $t = u/(1+1)$ et $dt = du/(1+u)^2$ fait le travail :

$$\begin{aligned}\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt &= \int_0^\infty \left(\frac{u}{1+u}\right)^{x-1}\left(\frac{1}{1+u}\right)^{y-1}\frac{du}{(1+u)^2}\\ &= \int_0^\infty \frac{u^{x-1}}{(1+u)^{x+y}}du.\end{aligned} \tag{97.803}$$

L'intégrale que nous devons calculer maintenant n'est autre que $\beta(3,2)$. En utilisant la formule

$$\Gamma(x)\Gamma(y) = \beta(x,y)\Gamma(x+y) \tag{97.804}$$

de la page *I*.80 nous trouvons $\beta(3,2) = \Gamma(3)\Gamma(2)/\Gamma(5)$. Nous utilisons maintenant le lien entre Γ et la factorielle :

$$\beta(3,2) = \frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} = \frac{2!1!}{4!} = \frac{1}{12}. \tag{97.805}$$

97.23 Quelques propriétés des espaces fonctionnels

Exercise 206 exo_I-4-1

Étudier la convergence au sens $L^2(M)$ sur $\mathbb{R}$ de la suite de fonctions de l'exercice 178 :

$$f_n(x) = \begin{cases} 0 & \text{pour } |x| > 2\\ x+2 & \text{pour } -2 \leqslant x \leqslant -1\\ 1 & \text{pour } -1 < x \leqslant -1/n\\ -nx & \text{pour } -1/n < x < 1/n\\ -1 & \text{pour } 1/n \leqslant x < 1\\ x-2 & \text{pour } 1 \leqslant x < 2 \end{cases} \tag{97.806}$$

corr_I-4-1

Correction of the exercise 206

Il est aisé de voir que

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}|f_n - f|^2dx = 0; \tag{97.807}$$

pour cela, tracez la fonction, et regardez-en l'aire Nous disons que f_n converge vers f en **moyenne quadratique**. Nous avons

$$f_n(x) - f(x) = \begin{cases} 0 & \text{pour } |x| > \frac{1}{n}\\ -nx-1 & \text{pour } -\frac{1}{n} \leqslant x < 0\\ 1-nx & \text{pour } 0 < x \leqslant \frac{1}{n} \end{cases} \tag{97.808}$$

Exercise 207 exo_I-4-2

Donner un exemple de suite de fonctions qui converge uniformément sur $\mathbb{R}$ et qui ne converge pas sur $\mathbb{R}$ en moyenne quadratique. corr_I-4-2

Correction of the exercise 207

Il suffit de trouver une suite de fonctions qui s'écrase sur zéro tout en gardant une surface constante. Par exemple

$$f_n(x) = \begin{cases} 0 & \text{si } |x| > n^2\\ \frac{1}{n} & \text{sinon} \end{cases} \tag{97.809}$$

La convergence uniforme vers zéro est évidente parce que $\|f_n\|_\infty = \frac{1}{n}$. Cette suite ne converge toutefois pas en moyenne quadratique parce que

$$\|f_n\|_{L^2} = \int_{-\infty}^{+\infty} f_n^2(x)dx = \int_{-n^2}^{n^2}\frac{1}{n^2}dx = 2. \tag{97.810}$$

Il est intéressant de noter que ce contre-exemple ne tient pas si on demande d'étudier la convergence sur $[a,b]$ au lieu de $\mathbb{R}$. En réalité, la convergence uniforme sur $[a,b]$ implique la convergence L^2 sur $[a,b]$ parce que la convergence uniforme $f_n \to f$ dit que dès que $n \geqslant N$, $|f_n(x) - f(x)| < \epsilon$, et donc

$$\|f_n - f\|_{L^2} = \int_a^b |f_n(x) - f(x)|^2 dx \leqslant \epsilon(b-a). \tag{97.811}$$

Exercise 208 exo_I-4-3

Pour une suite de fonctions $f_n : \mathbb{R} \to \mathbb{R}$, donner avec justification les liens entre

(1) la convergence uniforme

(2) la convergence uniforme sur tout compact

(3) la convergence au sens L^2.

corr_I-4-3

Correction of the exercise 208

La convergence uniforme implique l'uniforme sur tout compact, mais l'inverse n'est pas vrai comme le montre l'exemple

$$f_n(x) = \begin{cases} 1 & \text{si } n < x < n+1 \\ 0 & \text{sinon.} \end{cases} \tag{97.812}$$

La convergence uniforme n'implique pas la convergence en moyenne quadratique, comme nous l'avons vu à l'exercice 207.

Afin de voir que la convergence L^2 n'implique pas la convergence uniforme sur tout compact, reprenons l'exemple de l'exercice 178.

La différence entre la limite et les fonctions f_n est composée des deux triangles de base $\frac{1}{n}$ et de hauteur 1. La surface en dessous de $(f_n - f)^2$ n'est donc autre que

$$2\left(\frac{1}{n}\frac{1}{2}\right)^2 = \frac{1}{2n^2} \to 0. \tag{97.813}$$

Il y a donc convergence au sens L^2, mais même pas uniforme sur tout compact, et *a fortiori* pas uniforme.

97.24 Équations différentielles

97.24.1 Équations différentielles résolubles

Exercise 209 exo_II-1-01

Déterminer la solution générale de l'équation

$$y' = te^y. \tag{97.814}$$

Déterminer la solution telle que $y(0) = 0$.

corr_II-1-01

Correction of the exercise 209

Ceci est une équation à variables séparées, c'est-à-dire du type $y' = u(t)f(y)$. Ici, nous avons $u(t) = t$ et $f(y) = e^y$, et la primitive de $1/f$ est donnée par

$$G(y) = \int e^{-y} = -e^{-y}. \tag{97.815}$$

La fonction y est une solution de l'équation que nous regardons si et seulement si il existe une constante $c \in \mathbb{R}$ telle que

$$-e^{-y} = \frac{t^2}{2} + C. \tag{97.816}$$

Eq101ExC

Étant donné qu'une exponentielle est toujours positive, une solution y_C ne peut exister que lorsque $|t| < \sqrt{2C}$. En renommant $C \to C/2$ et en prenant le logarithme de la relation (97.816), nous trouvons

$$y_C(t) = -\ln\left(\frac{C-t^2}{2}\right), \tag{97.817}$$

qui est valable sur l'intervalle $|t| < \sqrt{C}$. Afin de trouver celle qui vérifie $y_C(0) = 0$, nous devons résoudre

$$y_c(0) = -\ln\left(\frac{C}{2}\right) = 0 \tag{97.818}$$

par rapport à C, dont la solution est évidement $C = 2$. La solution recherchée est donc

$$\begin{aligned} y_2 \colon]-\sqrt{2},\sqrt{2}[&\to \mathbb{R} \\ t &\mapsto -\ln\left(\frac{2-t^2}{2}\right). \end{aligned} \tag{97.819}$$

Exercise 210 exo_II-1-02

Déterminer la solution générale de

$$y' = 3y^{2/3}, \tag{97.820}$$

déterminer la solution telle que $y(0) = 27$. Étudier le cas où $y(0) = 0$. corr_II-1-02

Correction of the exercise 210

Ceci est encore une équation à variables séparées, dont la solution se recherche en suivant la même méthode que pour l'exercice 209. En suivant la même notation, nous avons $u(t) = 1$ et $f(y) = 3y^{2/3}$, ce qui donne $U(t) = t$ et $G(y) = y^{1/3}$. Il faut faire ici une remarque : dans la proposition 1 de la page 311, nous demandons que $f(\eta) \neq 0$ pour tout $\eta \in J$ où J est l'ensemble d'arrivée de la fonction y. Les solutions que nous allons trouver devront donc être restreintes au domaine $y(t) \neq 0$. Nous y reviendrons.

La solution générale est donc fournie par l'équation $y^{1/3} = t + C$, c'est-à-dire

$$y(t) = (t+C)^3. \tag{97.821}$$ EqGen102

La solution telle que $y(0) = 27$ est donc

$$\begin{aligned} y \colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto (t+3)^3. \end{aligned} \tag{97.822}$$

Toute solution y tel que $y(\tau) \neq 0$ pour un certain τ a la forme (97.821) sur $\mathbb{R}\backslash\{-C\}$. Cette solution s'étend immédiatement à $\mathbb{R}$ tout entier. La seule solution que le théorème nous a donc fait oublier est la solution identiquement nulle.

À part la solution identiquement nulle, la condition $y(0) = 0$ accepte la solution $y(t) = t^3$, qui est celle avec $C = 0$.

Exercise 211 exo_II-1-03

Résoudre

$$y' = (t^3 + t)e^{-y} \tag{97.823}$$

et déterminer la solution telle que $y(1) = 1$. corr_II-1-03

Correction of the exercise 211

En suivant le notations usuelles, nous avons $u(t) = t^3 + t$ et $f(y) = e^{-y}$. La solution générale est donnée par

$$G\big(y(t)\big) = U(t) + C \tag{97.824}$$

U est une primitive de u et G est une primitive de $1/f$. En l'occurrence,

$$y(t) = \ln\left(\frac{t^4}{4} + \frac{t^2}{2} + C\right). \tag{97.825}$$

Lorsque $C = e - \frac{3}{4}$, nous avons $y(1) = 1$, et la solution est

$$y(t) = \ln\left(\frac{t^4 + 2t^2 + 4e - 3}{4}\right), \tag{97.826}$$

qui existe sur tout $\mathbb{R}$.

Exercise 212 exo_II-1-04

Résoudre

$$y' + y\,\mathrm{cotg}(t) = -\frac{t}{\sin t} \tag{97.827}$$

avec $0 < t < \pi$, et déterminer la solution telle que $y(\pi/2) = -\pi^2/8$. corr_II-1-04

Correction of the exercise 212

À résoudre :

$$y' + \mathrm{cotg}(t) = -\frac{t}{\sin(t)}. \tag{97.828}$$

Ceci est une équation linéaire. Nous commençons par résoudre l'équation homogène $y' + y\,\mathrm{cotg}(t) = 0$, dont la solution est

$$y_H(t) = \frac{K}{\sin(t)}. \tag{97.829}$$

Afin de résoudre l'équation non homogène, nous utilisons la méthode de variations des constantes : nous disons maintenant que K est une fonction de t, et nous remettons le tout dans l'équation de départ pour trouver $K(t)$. Nous regardons donc

$$y(t) = \frac{K(t)}{\sin(t)}. \tag{97.830}$$

Ce que nous trouvons est

$$\frac{K'}{\sin(t)} - K\frac{\cos(t)}{\sin^2(t)} + \frac{K\,\mathrm{cotg}(t)}{\sin(t)} = -\frac{t}{\sin(t)}, \tag{97.831}$$

c'est-à-dire, après simplifications, $K' = -t$. Nous avons donc pour solution finale :

$$y(t) = \frac{1}{\sin(t)}\left(\frac{C - t^2}{2}\right). \tag{97.832}$$

La solution telle que $y(\pi/2) = -\pi^2/8$ est donnée par $C = 0$.

Exercise 213 exo_II-1-05

Résoudre les équations différentielles suivantes :

(1) $\sin(2t)y' + y^2 = 1$,
(2) $y' + \frac{y}{t+1} = \sin(t)$,
(3) $y' - \frac{a}{t}y = e^t t^a$, où a est une constante,
(4) $yy' + (1 + y^2)\sin(t) = 0$,
(5) $y'' + 2y' + y = e^{-t}\ln|t|$.

corr_II-1-05

Correction of the exercise 213

(1) $\sin(2t)y' + y^2 = 1$. Ceci est une équation à variables séparée qui s'écrit sous forme plus traditionnelle

$$y' = \frac{1}{\sin(t)}(1 - y^2). \tag{97.833}$$

En suivant les notations du cours de première, nous avons $u(t) = 1/\sin(t)$ et $f(y) = 1 - y^2$. Si $1 - y^2 = 0$, nous trouvons les deux solutions constantes $y = \pm 1$, sinon nous pouvons continuer la méthode et trouver

$$\begin{aligned} U(t) &= \frac{1}{2}\ln(\tan(t)) \\ G(t) &= \int \frac{dy}{1-y^2} = -\frac{1}{2}\ln\left(\frac{y+1}{y-1}\right). \end{aligned} \tag{97.834}$$

L'équation implicite donnant y est donc

$$-\frac{1}{2}\ln\left(\frac{y+1}{y-1}\right) = \frac{1}{2}\ln(\tan t) + C. \tag{97.835}$$

(2) $y' + y/(t+1) = \sin(t)$. C'est une équation linéaire, dont on cherche d'abord la solution générale de l'homogène associée. Cette solution est

$$y_H(t) = \frac{K}{t+1}. \tag{97.836}$$

Nous appliquons maintenant le technique de la variation des constantes pour trouver la solution générale de l'équation proposée. Maintenant $K = K(t)$, et nous avons

$$y'(t) = \frac{K'(t)}{t+1} - \frac{K(t)}{(t+1)^2} \tag{97.837}$$

que nous remettons dans l'équation de départ pour trouver une équation différentielle pour $K(t)$. Ce que nous trouvons est

$$K'(t) = \sin(t)(t+1), \tag{97.838}$$

dont la solution se ramène au calcul d'une primitive. Le résultat est

$$K(t) = \sin(t) - \cos(t)(t+1) + C, \tag{97.839}$$

ce qui donne la réponse finale :

$$y(t) = \frac{\sin(t) + C}{t+1} - \cos(t). \tag{97.840}$$

(3) $y' - ay/t = e^t t^a$. Ceci est une équation différentielle linéaire qui peut aussi s'écrire $y't^{-a} - at^{a-1}y = e^t$, ce qui met l'équation de départ sous la forme

$$(t^{-a}y)' = e^t, \tag{97.841}$$

ce qui donne tout de suite $t^{-a}y = e^t + C$, et donc la solution

$$y(t) = (e^t + C)t^a. \tag{97.842}$$

(4) $yy' + (1+y^2)\sin(t) = 0$. Cela est une équation à variables séparée, mais elle peut être simplifiée en remarquant que $yy' = (y^2)'/2$. Nous avons alors

$$\frac{1}{2}\frac{(y^2)'}{1+y^2} = -\sin(t), \tag{97.843}$$

que nous récrivons

$$\frac{(1+y^2)'}{1+y^2} = -2\sin(t). \tag{97.844}$$

Cela entraîne que la solution est donnée par l'équation

$$1 + y^2 = Ce^{2\cos(t)}. \tag{97.845}$$

(5) $y'' + 2y' + y = e^{-t}\ln(t)$. C'est une équation linéaire à coefficients constants, dont l'équation homogène est

$$y'' + 2y' + y = 0. \tag{97.846}$$

Son polynôme caractéristique est $r^2 + 2r + 1 = 0$, dont l'unique solution est $r = -1$, de multiplicité 2. En vertu de la théorie générale la solution générale à l'équation homogène est

$$y_H = (At + B)e^{-t} \tag{97.847}$$

pour certaines constantes A et B.
Maintenant, nous utilisons la méthode de la variation des constantes pour trouver la solution générale de l'équation non homogène. Nous posons donc

$$y(t) = \big(A(t)t + B(t)\big)e^{-t}. \tag{97.848}$$

Toujours en vertu de ce qui a été vu en première, les fonctions A et B ne sont pas complètement indépendantes, mais peuvent être choisies de façon à vérifier la condition (32.77)

$$A't + B' = 0. \tag{97.849}$$ EqII105ConditionVC

Cette condition va considérablement simplifier le calcul qui suit. D'abord, nous avons

$$y' = (A't + A + B' - At - B)e^{-t}, \tag{97.850}$$

dans lequel nous utilisons la condition (97.849) pour trouver

$$y' = (A - At - B)e^{-t}. \tag{97.851}$$

En dérivant encore une fois, il vient

$$y'' = (A' + A't - 2A - B' + At + B)e^{-t}, \tag{97.852}$$

dans laquelle il n'y a pas de dérivées secondes de A et B, grâce à l'utilisation de (97.849). Nous pouvons maintenant écrire l'équation $y'' + 2y' + y = e^{-t}\ln|t|$ qui est maintenant une équation différentielle pour A et B. Étant donné que (97.849) tient toujours, nous avons en réalité le système

$$\begin{cases} A' - A't - B' = \ln|t| & \text{(97.853a)} \\ A't + B' = 0. & \text{(97.853b)} \end{cases}$$

En introduisant la seconde équation dans la première, nous trouvons $A' = \ln|t|$, dont nous déduisons

$$A(t) = t\big(\ln|t| - 1\big) + A_0. \tag{97.854}$$

La seconde équation nous dit par ailleurs que $B' = -A't = -t\ln|t|$, dont l'intégration donne

$$B(t) = -\frac{t^2}{4}\big(2\ln|t| - 1\big) + B_0. \tag{97.855}$$

Exercise 214 exo_II-1-06
On considère l'équation différentielle

$$y'' + y = \cos(\omega t) \tag{97.856}$$

où ω est une constante positive.

(1) Déterminer sa solution générale lorsque $\omega \neq 1$ et exprimer cette solution en fonction des conditions initiales $y_0 = y(0)$ et $y_0' = y'(0)$. ItemII106a

(2) Déterminer sa solution générale pour $\omega = 1$ et l'exprimer en fonction de y_0 et y_0'. ItemII106b

(3) Montrer que lorsque ω tend vers 1, la solution obtenue en (1) tend vers la solution obtenue en (2).

corr_II-1-06

Correction of the exercise 214

L'équation homogène est $y'' + y = 0$, dont la solution est $y_H = A\cos(t) + B\sin(t)$. Afin de trouver une solution particulière de l'équation non homogène, nous essayons

$$y = \alpha\cos(\omega t), \tag{97.857}$$

ce qui mène à l'équation $-\omega^2\alpha + \alpha = 1$, donc

$$\alpha = \frac{1}{1-\omega^2}. \tag{97.858}$$

Si $\omega \neq 1$ (on a supposé que $\omega \geqslant 0$), les solutions sont donc

$$y(t) = A\cos(t) + B\sin(t) + \frac{1}{1-\omega^2}\cos(\omega t). \tag{97.859}$$

Étant donné que $y(0) = A + \frac{1}{1-\omega^2}$ et $y'(0) = B$, nous trouvons

$$y_A = \left(y_0 - \frac{1}{1-\omega^2}\right)\cos(t) + y_0'\sin(t) + \frac{1}{1-\omega^2}\cos(\omega t). \tag{97.860}$$

Par contre, si $\omega = 1$, cette solution ne fonctionne pas, et nous cherchons une solution particulière de l'équation non homogène sous la forme $y(t) = \alpha t\cos(t) + \beta t\sin(t)$. En dérivant deux fois et en remplaçant dans l'équation départ, nous trouvons

$$\sin(t)(-2\alpha - \beta t + \beta t) + \cos(t)(-\alpha t + 2\beta + \alpha t) = \cos(t), \tag{97.861}$$

donc $\alpha = 0$ et $\beta = \frac{1}{2}$. Dans le cas $\omega = 1$, la solution est donc

$$y = A\cos(t) + B\sin(t) + \frac{t}{2}\sin(t). \tag{97.862}$$

Dans ce cas-ci, nous avons $y(0) = A$ et $y'(0) = B$, donc

$$y_B = y_0\cos(t) + y_0'\sin(t) + \frac{t}{2}\sin(t). \tag{97.863}$$

Prouver la convergence de $\lim_{\omega\to 0} y_A = y_B$, nous calculons $y_A - y_B$:

$$y_A - y_B = \frac{\cos(\omega t) - \cos(t)}{1-\omega^2} - \frac{t}{2}\sin(t). \tag{97.864}$$

Nous calculons donc (avec la règle de l'Hospital par rapport à ω)

$$\lim_{\omega\to 1}\frac{\cos(\omega t) - \cos(t)}{1-\omega^2} - \frac{t}{2}\sin(t) = \lim_{\omega\to 1}\frac{-t\sin(\omega t)}{2} - \frac{t}{2}\sin(t) = 0. \tag{97.865}$$

97.24.2 Équation de Bernoulli

Exercise 215 exoreserve0004

Résoudre les équations suivantes :

(1) $ty' + y = ty^2$,

(2) $y' = \frac{y}{t}\big(y\ln|t| - 1\big)$,

(3) $y - \cos(t)y' = \cos(t)\big(1 - \sin(t)\big)y^2$.

ItemCII107

corrreserve0004

Correction of the exercise 215

(1) $ty' + y = ty^2$.
Évidement, $y = 0$ est solution. Si $y(t_0) \neq 0$, nous pouvons chercher une solution non triviale dans un voisinage de t_0. En vertu de la méthode générale exposée en 32.10.2, nous posons $z = 1/y$, ce qui donne $y' = -y^2 z'$. Après aménagements (diviser par y^2), nous trouvons

$$-tz' + z = t, \tag{97.866}$$

EqII107Lin

qui est une équation linéaire. La solution générale de l'équation homogène associée est

$$z_H(t) = At. \tag{97.867}$$

Nous utilisons maintenant la méthode de variation des constantes pour trouver la solution générale de (97.866). Nous posons donc $z(t) = A(t)t$ que nous remplaçons dans l'équation de départ $ty' + y = ty^2$. Nous tombons sur l'équation

$$A' = -\frac{1}{t} \tag{97.868}$$

dont la solution générale est $A(t) = \ln\left(\frac{C}{t}\right)$. En définitive, la solution à notre problème est

$$z(t) = t \ln\left(\frac{C}{t}\right), \tag{97.869}$$

et donc

$$y(t) = \frac{1}{t \ln\left(\frac{C}{t}\right)}. \tag{97.870}$$

N'oublions pas de mentionner que $y = 0$ est aussi solution (qui correspond à $C \to 0$).

(2) $y' = -\frac{1}{t}y + \frac{\ln|t|}{t}y^2$.
Nous posons $z = 1/y$, et nous récrivons l'équation sous la forme

$$z' - \frac{1}{t}z = -\frac{\ln|t|}{t} \tag{97.871}$$

qui est une équation linéaire. L'équation linéaire est

$$z'_H = \frac{z_H}{t} \tag{97.872}$$

et la solution est $z_h = At$. Nous utilisons encore la méthode de la variation des constantes en posant $z(t) = A(t)t$. L'équation différentielle à laquelle doit satisfaire $A(t)$ est alors

$$A' = -\frac{\ln(t)}{t^2}. \tag{97.873}$$

Trouver une primitive de $\frac{\ln|t|}{t^2}$ n'est pas trop aisé (ça se fait par partie), mais la solution est

$$A(t) = \frac{1}{t}\big(\ln(t) + 1\big) + Ct, \tag{97.874}$$

donc

$$z(t) = \ln(t) + 1 + Ct. \tag{97.875}$$

(3) $y - \cos(t)y' = \cos(t)\big(1 - \sin(t)\big)y^2$.
Encore une fois, $y = 0$ est solution. En posant $z = 1/y$, nous trouvons l'équation

$$z + \cos(t)z' = \cos(t)\big(1 - \sin(t)\big) \tag{97.876}$$

EqII107EqpourZ

à laquelle z doit satisfaire. L'équation homogène est

$$z'_H = -\frac{z_H}{\cos(t)}. \tag{97.877}$$

Nous résolvons cette équation en utilisant la méthode des équations à variable séparées de la section 32.4. Nous posons donc

$$\begin{aligned} u(t) &= \frac{1}{\cos(t)},\\ f(z) &= -z,\\ U(t) &= \ln\left[\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)\right] \quad \text{(voir formulaire)},\\ G(z) &= \ln\left(\frac{1}{z}\right). \end{aligned} \tag{97.878}$$

La solution z_H est donnée par l'équation

$$\ln\left(\frac{1}{z}\right) = \ln\left[K\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)\right], \tag{97.879}$$

c'est-à-dire

$$z_H(t) = \frac{K}{\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)}. \tag{97.880}$$

Nous appliquons maintenant la méthode de variation des constantes sur cette solution afin de trouver la solution générale de l'équation (97.876). En utilisant la règle de Leibniz, $z' = K'z_H + Kz'_H$, nous trouvons

$$\frac{K}{\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)} + \cos(t)\left(\frac{K'}{\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)} - \frac{K}{2\sin^2\left(\frac{\pi}{4}+\frac{t}{2}\right)}\right) = \cos(t)\big(1-\sin(t)\big). \tag{97.881}$$

Malgré leurs apparences, les deux termes en K se simplifient. En effet, en vertu de l'équation $z'_H = \frac{-z_H}{\cos(t)}$, nous avons

$$\frac{-K}{2\sin^2\left(\frac{\pi}{4}+\frac{t}{2}\right)} = \frac{-K}{\cos(t)\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)}. \tag{97.882}$$

Le travail de voir quel est le lien entre $\sin^2\left(\frac{\pi}{4}+\frac{t}{2}\right)$, $\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)$ et $\cos(t)$ est en réalité fait dans votre formulaire au moment où vous l'avez utilisé pour intégrer u pour obtenir le $U(t)$ de (97.878).

Après cette simplification durement méritée, nous trouvons l'équation suivante pour $K(t)$:

$$\frac{K'}{\tan\left(\frac{\pi}{4}+\frac{t}{2}\right)} = 1-\sin(t). \tag{97.883}$$

Résoudre cela revient à trouver la primitive de

$$\big(1-\sin(t)\big)\tan\left(\frac{\pi}{4}+\frac{t}{2}\right), \tag{97.884}$$

ce qui est relativement compliqué. La réponse est

$$\begin{aligned} K(t) = \ln\left(\sin\left(\frac{2x+\pi}{4}\right)+1\right) + \ln\left(\sin\left(\frac{2x+\pi}{4}\right)-1\right)\\ + 2\ln\sec\left(\frac{2x+\pi}{4}\right) + 2\sin^2\left(\frac{2x+\pi}{4}\right) \end{aligned} \tag{97.885}$$

Nous pouvons un peu simplifier en utilisant le fait que $\ln(a+b)+\ln(a-b) = \ln(a^2-b^2)$:

$$K(t) = \ln\left(-\cos^2\left(\frac{2x+\pi}{4}\right)\right) + 2\ln\sec\left(\frac{2x+\pi}{4}\right) + 2\sin^2\left(\frac{2x+\pi}{4}\right). \tag{97.886}$$

Il me semble toutefois qu'il faudrait prendre des valeurs absolues pour les logarithmes.

Exercise 216 exo_II-1-08

Lorsqu'on étudie le mouvement d'un point matériel *pesant* soumis à une résistance de l'air fonction de la vitesse, on peut montrer que l'équation différentielles de l'hodographe des vitesses est

$$\frac{1}{r}\frac{dr}{d\theta} = \frac{\sin\theta + R(r)}{\cos\theta} \tag{97.887}$$

où r, θ sont les coordonnées polaires et R est la fonction caractérisant le type de résistance.

Résoudre l'équation différentielle pour $R(r) = kr^\alpha$. Quelle est la nature de l'hodographe des vitesses lorsque $R(r) = kr$?

corr_II-1-08

Correction of the exercise 216

Afin de mieux suivre les notations de la théorie (Bernoulli, page 2499) nous allons écrire β au lieu de α. Nous pouvons directement régler son compte au cas $\beta = 0$. En effet, nous trouvons

$$\frac{dr}{r} = \left(\tan(\theta) + \frac{k}{\cos(\theta)}\right) d\theta, \tag{97.888}$$

d'où nous tirons $r(\theta)$ moyennant une simple primitive.

Si $\beta \neq 0$, nous trouvons l'équation

$$r' = r\tan(\theta) + \frac{k}{\cos(\theta)} r^{\beta+1} \tag{97.889}$$

qui est de la forme (32.337) avec

$$\begin{aligned} a(\theta) &= \tan(\theta) \\ b(\theta) &= k/\cos(\theta) \\ \alpha &= \beta + 1. \end{aligned} \tag{97.890}$$

En suivant la méthode générale, poser $z = r^{-\beta}$ fournit l'équation linéaire

$$z' = -\beta\tan(\theta)z + \frac{k}{\cos(\theta)}. \tag{97.891}$$ EqII108LinPourz

L'équation homogène associée, $z'_H + \beta\tan(\theta)z_H = 0$ (qui est à variable séparées), a pour solution

$$z_H = K\cos^\beta(\theta). \tag{97.892}$$

Nous utilisons maintenant la méthode de variations des constantes, c'est-à-dire que nous posons $z(\theta) = K(\theta)z_H(\theta)$. En remettant dans l'équation 97.891, et en effectuant la simplification qui se présente, nous trouvons l'équation suivante pour K :

$$K' = -\frac{\beta k}{\cos^{\beta+1}(\theta)}. \tag{97.893}$$

Nous avons donc

$$K(\theta) = -\beta k \int_0^\theta \cos^{-(\beta+1)}(t)dt, \tag{97.894}$$

qui n'est pas une intégrale facile à calculer.

Lorsque $\beta = 1$, les choses sont plus simples parce que nous savons que

$$\int \frac{dx}{\cos^2(x)} = \tan(x). \tag{97.895}$$

Donc $K(\theta) = -\beta k\tan(\theta) + K$, et en refaisant le changement de variable vers r, nous trouvons la solution

$$r(\theta) = \frac{1}{K\cos(\theta) - k\sin(\theta)}. \tag{97.896}$$

Ceci est une équation polaire pour une courbe que nous devons identifier. Nous la récrivons sous la forme

$$1 = Kr\cos(\theta) - kr\sin(\theta), \tag{97.897}$$

et nous identifions les coordonnés cartésiennes $x = r\cos(\theta)$ et $y = r\sin(\theta)$, donc

$$1 = Kx - ky, \tag{97.898}$$

qui est l'équation d'une droite.

97.24.3 Équations de Ricatti

Exercise 217 exo_II-1-09

Résoudre les équations suivantes :

(1) $(y - \sin(t))' - y\big(y - \sin(t)\big)$,

(2) $y' = \frac{y^2 - t^2 y - 2t}{1 - t^3}$.

Pour la seconde, cherchez des solutions particulières sous la forme $y = at^2 + bt + d$. corr_II-1-09

Correction of the exercise 217

Nous suivons la méthode expliquée en 32.10.3. La première chose à faire est de voir si nous pouvons deviner des solutions particulières.

(1) Ici, nous voyons tout de suite que $y_1(t) = \sin(t)$ est une solution. Nous posons donc

$$y(t) = \sin(t) + \frac{1}{u}, \tag{97.899}$$

et nous trouvons l'équation différentielle

$$u' + \sin(t)u = 1. \tag{97.900}$$ EqII109DiffPouru

Note : en suivant les notations du rappel théorique, nous avons

$$\begin{aligned} a(t) &= 1, \\ b(t) &= -\sin(t), \\ c(t) &= \cos(t). \end{aligned} \tag{97.901}$$

L'équation (97.900) est du type de (32.79). Dans ces notations, nous cherchons donc K qui vérifie l'équation $K'(t) = g(t)/u_H(t)$, c'est-à-dire

$$K'(t) = -\frac{1}{e^{\cos(t)}}. \tag{97.902}$$

Cette primitive n'est absolument pas simple à calculer. La solution à notre équation différentielle est donc

$$z(t) = -e^{\cos(t)} \int_0^t e^{-\cos(u)} du. \tag{97.903}$$

Pour la petite histoire, l'intégrateur en ligne de Wolfram ne trouve pas de forme pour cette intégrale.

97.24.4 Équations homogènes

Exercise 218 exo_II-1-10

Résoudre :

(1) $y' = \frac{y}{y-t}$

(2) $y' = \frac{y}{t - 2(ty)^{1/2}}$.

corr_II-1-10

Correction of the exercise 218

(1) $y' = y/(y-t)$.
Nous posons, conformément à (32.335) $y = tz$, donc $y' = z + tz'$, et nous trouvons

$$tz' = \frac{z - z^2 + z}{z-1} = \frac{z(z-2)}{1-z}. \tag{97.904}$$

Si $z(z-2) = 0$, alors nous avons les solutions particulières $z(t) = 0$ et $z(t) = 2$ qui correspondent aux solutions $y(t) = 0$ et $y(t) = 2t$. Si $z(z-2) \neq 0$, alors nous pouvons faire la manipulation suivante :

$$\frac{z'(1-z)}{z(z-2)} = \frac{1}{t}. \tag{97.905}$$

Remarquez que le numérateur est à peu près la dérivée du numérateur. En posant $u = z(z-2)$, nous récrivons l'équation sous la forme plus simple

$$\frac{-u'/2}{u} = \frac{1}{t}. \tag{97.906}$$

En écrivant $u' = du/dt$ et en faisant passer le dt de l'autre côté, nous avons

$$\frac{du}{u} = -2\frac{dt}{t} \tag{97.907}$$

que nous intégrons des deux côtés :

$$\ln(u) = \ln(Kt^{-2}), \tag{97.908}$$

ce qui amène au final

$$y(t) = t\big(1 \pm \sqrt{1 + Kt^{-2}}\big). \tag{97.909}$$

(2) $y' = y/(t - 2(ty)^{1/2})$.
Nous posons $y = tz$, et après quelques manipulations algébriques nous remettons tout sous la forme

$$\frac{z'(1-2z^{1/2})}{2z^{3/2}} = \frac{1}{t}. \tag{97.910}$$ EqII110EqPourz

Nous utilisons l'astuce de la section 32.2, en commençant par écrire

$$\frac{1-2z^{1/2}}{2z^{3/2}} = \frac{dt}{t}, \tag{97.911}$$

dont nous intégrons les deux membres :

$$-\ln(z) - \frac{1}{\sqrt{z}} = \ln|t| + C, \tag{97.912}$$

où nous avons mit toutes les constantes dans C. Ensuite, nous remettons l'ancienne variable :

$$\ln|t| = -\ln\left|\frac{y}{t}\right| - \frac{1}{\sqrt{y/t}} + C, \tag{97.913}$$

ou encore

$$\ln|y| + \sqrt{\frac{t}{y}} + C = 0. \tag{97.914}$$

Cela est une équation implicite pour $y(t)$.

97.24.5 Équations différentielles exactes. Facteurs intégrants

Exercice 219 exo_II-1-11

Résoudre

(1) $y'(t+y^2)+(y-t^2)=0$,

(2) $\big(y+ty+\sin(y)\big)dt+\big(t+\cos(y)\big)dy=0$,

(3) $y'=-\frac{3t^2y+8ty^2}{t^3+8t^2y+12y^2}$.

corr_II-1-11

Correction of the exercise 219

Nous suivons les méthodes et notations du point 32.10.4.

(1) $y'(t+y^2)+(y-t^2)=0$.

Ici, nous avons

$$\begin{aligned} P(t,y)&=y-t^2\\ Q(t,y)&=t+y^2. \end{aligned} \tag{97.915}$$

Nous vérifions que l'équation est bien exacte : $\partial_y P=\partial_t Q=1$. Nous pouvons donc nous lancer à la recherche d'une fonction $f(t,y)$ telle que $df=Pdt+Qdy$. Nous pouvons la deviner : en intégrant P par rapport à t, nous avons $f(t,y)=yt-\frac{t^3}{3}+C$ où C est une constante *par rapport à* t. Donc C peut dépendre de y. Il n'est pas très difficile de fixer $C(y)$ pour que $\partial_y f=Q$. Nous trouvons

$$f(t,y)=\frac{1}{3}(y^3-t^3)+ty. \tag{97.916}$$ EqII111FonctionIntf

Un simple calcul montre que cette fonction fonctionne.

Ce résultat peut également être calculé en intégrant la forme $Pdt+Qdy$ le long d'un chemin qui relie $(0,0)$ à (t,y). En tant que chemin, nous choisissons la composée de

$$\begin{aligned} \gamma_1(u)&=(0,uy)\\ \gamma_2(u)&=(ut,y). \end{aligned} \tag{97.917}$$

Nous trouvons, en appliquant la définition d'une intégrale de forme différentielle [21] sur un chemin, nous trouvons

$$\begin{aligned} f(t,y)&=\int_{\gamma_1\circ\gamma_2}(Pdt+Qdy)\\ &=\int_0^1(Pdt+Qdy)\big(\gamma_1(u)\big)\big(\gamma_1'(u)\big)du\\ &\quad+\int_0^1(Pdt+Qdy)\big(\gamma_2(u)\big)\big(\gamma_2'(u)\big)du\\ &=\int_0^1\big((P\circ\gamma_1)(u)dt+(Q\circ\gamma_1)(u)dy\big)\begin{pmatrix}0\\y\end{pmatrix}\\ &\quad+\int_0^1\big((P\circ\gamma_2)(u)dt+(Q\circ\gamma_2)(u)dy\big)\begin{pmatrix}t\\0\end{pmatrix}\\ &=\int_0^1 yQ(0,uy)du+\int_0^1 tP(ut,y)du\\ &=\frac{y^3}{3}+ty-\frac{t^3}{3}. \end{aligned} \tag{97.918}$$

Voilà qui confirme la formule (97.916). Nous avons donc maintenant $Q=\frac{\partial f}{\partial y}$ et $P=\frac{\partial f}{\partial t}$. La solution de l'équation différentielle est donc donnée par l'équation implicite

$$\frac{1}{3}(y^3-t^3)+ty=C. \tag{97.919}$$

21. Juste pour être sûr que vous ayez compris : vous savez la différence entre une forme et une forme *différentielle* ?

(2) $\big(y+ty+\sin(y)\big)dt+\big(t+\cos(y)\big)dy=0$.
Nous avons

$$\begin{aligned}P(t,y)&=y+ty+\sin(y)\\Q(t,y)&=t+\cos(y).\end{aligned}\tag{97.920}$$

Un rapide calcul montre que cette équation n'est pas exacte. Nous faisons donc le coup du facteur intégrant (32.351) pour la rendre exacte. L'équation pour $M(t,y)$ est

$$M\big(t+\cos(y)\big)=\big(t+\cos(y)\big)\frac{\partial M}{\partial t}-\big(y+ty+\sin(y)\big)\frac{\partial M}{\partial y}.\tag{97.921}$$

Une très mauvaise idée serait d'essayer de la résoudre en général : tout ce dont nous avons besoin est d'une solution à cette équation. Étant donné que le coefficient devant M est égal à celui devant $\partial_t M$, ce serait bien que le terme en $\partial_y M$ soit nul. Cherchons donc une solution qui ne soit pas fonction de y, mais seulement de t. Dans ce cas, nous trouvons à résoudre

$$M\big(t+\cos(y)\big)=\big(t+\cos(y)\big)\frac{\partial M}{\partial t},\tag{97.922}$$

c'est-à-dire $M=\partial_t M$. Une solution est $M(t,y)=e^t$.
Nous regardons donc maintenant l'équation différentielle

$$e^t\big(y+ty+\sin(y)\big)dt+e^t\big(t+\cos(y)\big)dy=0,\tag{97.923}$$

qui est exacte (vous pouvez le vérifier, si vous ne faites pas confiance en le calcul que nous venons de faire). La fonction qui intègre cette forme est

$$f(t,y)=e^t\big(\sin(y)+yt\big).\tag{97.924}$$

La solution implicite au problème est donc

$$e^t\big(\sin(y)+yt\big)=C.\tag{97.925}$$

Exercise 220 exo_II-1-12
Soit l'équation différentielle

$$P(t,y)+Q(t,y)y'=0.\tag{97.926}$$

(1) Montrer que si les fonctions P et Q sont homogènes de degré n, il existe un facteur intégrant M homogène de degré $m=-(1+n)$. Indication : utiliser la formule d'Euler

$$t\frac{\partial P}{\partial t}+t\frac{\partial P}{\partial y}=nP.\tag{97.927}$$

(2) Montrer que si $P=yp(ty)$ et $Q=yq(ty)$, il existe un facteur intégrant $M=M(ty)$.

(3) Vérifier que dans chacun des cas précédents, il existe une autre façon de résoudre l'équation différentielle.

Juste pour le plaisir, prouvons la formule d'Euler. Le fait que P et Q sont homogènes de degré n signifie que

$$\begin{aligned}P(\lambda t,\lambda y)&=\lambda^n P(t,y)\\Q(\lambda t,\lambda y)&=\lambda^n Q(t,y).\end{aligned}\tag{97.928}$$

Nous considérons $f(\lambda,t,y)=P(\lambda t,\lambda y)$, et nous calculons $(df/d\lambda)(1,t,y)$.

$$\frac{df}{d\lambda}(1,t,y)=\frac{\partial P}{\partial t}(t,y)\frac{d(\lambda t)}{d\lambda}(1)+\frac{\partial P}{\partial y}(t,y)\frac{d(\lambda y)}{d\lambda}=\frac{\partial P}{\partial t}(t,y)t+y\frac{\partial}{\partial y}(t,y).\tag{97.929}$$

D'autre part, par hypothèse, $f(\lambda,t,y)=\lambda^n P(t,y)$, donc

$$\frac{df}{d\lambda}(1,t,y)=\big(n\lambda^{n-1}P(t,y)\big)_{\lambda=1}=nP(t,y).\tag{97.930}$$

Cela conclut la démonstration de la formule d'Euler. corr_II-1-12

Correction of the exercise 220

Récrivons l'équation du facteur intégrant (32.351) :

$$M(\partial_y P - \partial_t Q) = Q(\partial_t M) - P(\partial_y M). \tag{97.931}$$ EqFactIntGenII112

(1) Nous allons prouver que l'équation du facteur intégrant accepte une solution lorsque M est homogène de degré $-(n+1)$, c'est-à-dire quand M satisfait l'équation

$$t\partial_t M + y\partial_y M = -(n+1)M. \tag{97.932}$$

Nous remplaçons dans (97.931) toute les dérivées par rapport à y par des dérivées par rapport à t en utilisant les formules d'Euler :

$$\begin{aligned} \partial_y P &\to \frac{1}{y}(nP - t\partial_t P) \\ \partial_y M &\to \frac{1}{y}\big(-(n+1)M - t\partial_t M\big). \end{aligned} \tag{97.933}$$

En remettant les termes ensembles, nous trouvons

$$\frac{MP}{y} + \frac{t}{y}\partial_t(MP) + \partial_t(MQ). \tag{97.934}$$

Les deux premiers termes du membre de gauche peuvent être mit en une seule dérivée par rapport à t en remarquant que

$$\frac{MP}{y} + \frac{t}{y}\partial_t(MP) = \partial_t\left(\frac{tMP}{y}\right). \tag{97.935}$$

Ce que nous trouvons alors est $\partial_t\left(\frac{tMP}{y} + MQ\right) = 0$, et donc

$$M(Q + tP/y) = C. \tag{97.936}$$

Attention : la constante est une constante *par rapport à* t, elle peut dépendre de y. Ce que nous avons est

$$M = \frac{C(y)}{tP + Qy} \tag{97.937}$$

où $C(y)$ est une fonction de y. Cela est la forme que doit avoir M pour satisfaire à l'équation du facteur intégrant. Pour que M soit, de plus homogène de degré $-(m+1)$, il faut en plus que

$$M(\lambda t, \lambda y) = \lambda^{-m-1}M(t, y), \tag{97.938}$$

ce qui va imposer des restrictions sur $C(y)$. Le but de l'exercice est de prouver que cette restriction peut être satisfaire par un choix convenable de $C(y)$. En utilisant l'homogénéité de P et Q, nous trouvons

$$M(\lambda t, \lambda y) = \frac{M(\lambda y)}{\lambda^{n+1}(tP + yQ)}. \tag{97.939}$$

Donc, nous en déduisons que pour tout y et pour tout λ, $C(y) = C(\lambda y)$. Une fonction constante satisfait à cette exigence. Donc nous avons un facteur intégrant sous la forme

$$M(t, y) = \frac{1}{tP + yQ}. \tag{97.940}$$

(2) Nous reprenons l'équation $P+Qy'=0$, et nous posons $v=ty$. L'équation du facteur intégrant devient
$$M\Big(\partial_y\big(yp(ty)\big)-\partial_t\big(tq(ty)\big)\Big)=Q(\partial_t M)-P(\partial_y M). \tag{97.941}$$
Nous utilisons la règle de Leibniz et la règle de la dérivation de fonctions composées pour effectuer les dérivations. Le membre de gauche devient
$$M\big(p-q+v\partial_v(p-q)\big), \tag{97.942}$$
tandis que pour calculer le membre de droite, nous nous souvenons de l'hypothèse $M=M(ty)=M(v)$, et nous pouvons donc écrire $\partial_y M=(\partial_v M)t$, ce qui mène à l'équation
$$M\big(p-q+v\partial_v(p-q)\big)=-v\frac{\partial M}{\partial v}(p-q). \tag{97.943}$$
Si $f=p-q$, nous avons donc l'équation
$$\frac{M}{\partial_v M}=\frac{\partial_v f}{f}, \tag{97.944}$$
dont les solutions sont $M=k/f$.

(3) Dans le premier cas, si P et Q sont homogènes de même degré, alors l'équation $P+qy'=0$ se récrit $y'=-P/Q$, qui est homogène de degré zéro, et nous pouvons suivre la méthode donnée au point 32.10.1.
Dans le second cas, nous avons l'équation
$$yp(v)dt+tq(v)dy=0, \tag{97.945}$$
et l'hypothèse nous permet de dire que $ydt=dv-tdy$, ce qui met l'équation sous la forme
$$(dv-tdy)p+tqdy=0, \tag{97.946}$$
dans laquelle il n'y a plus de dt. Cette équation peut être mise sous la forme
$$\frac{dy}{y}=\frac{p}{v(p-q)}\frac{dv}{v}, \tag{97.947}$$
qui se résout par quadrature.

97.24.6 L'équation $y'=f\left(\frac{at+by+c}{a't+b'y+c'}\right)$

Exercise 221 exo_II-1-13
Résoudre
(1) $y'=\frac{t+2y+1}{2t-3}$, ItemII113a
(2) $y'=\frac{t+2y+1}{2t+4y+3}$, ItemII113b
(3) $y'=(2t+3y+5)^2$. corr_II-1-13

Correction of the exercise 221

(1) $y'=(t+2y+1)/(2t-3)$. Le déterminant
$$\begin{vmatrix}1&2\\2&0\end{vmatrix}=-4\neq 0. \tag{97.948}$$
Nous devons donc résoudre le système
$$\begin{cases}\bar{t}+2\bar{y}+1 &=0 & (97.949a)\\ 2\bar{t}-3 &=0. & (97.949b)\end{cases}$$

La solution est $\bar{t} = 3/2$ et $\bar{y} = -5/4$. Nous posons donc $\tau = t - \frac{3}{2}$, et $u(\tau) = y(t) + \frac{5}{4}$. Essayons d'effectuer ce changement de variable de façon la plus explicite qui soit. Nous avons

$$\frac{du}{d\tau}(\tau) = \frac{d}{d\tau}\big((y \circ t)(\tau)\big) = y'\big(t(\tau)\big) \cdot \frac{dt}{d\tau} = y'\big(t(\tau)\big) \tag{97.950}$$

où y' est la dérivée de y. Nous pouvons écrire l'équation différentielle qui nous intéresse au point $t(\tau)$:

$$u'(\tau) = y'\big(t(\tau)\big) = \frac{t(\tau) + 2y\big(t(\tau)\big) + 1}{2t(\tau) - 3} = \frac{\tau + 2u(\tau)}{2\tau}. \tag{97.951}$$

EqII113AvecTauHom

C'est cette équation que nous allons résoudre pour $u(\tau)$. C'est une équation homogène, donc nous posons $z(\tau) = u(\tau)/\tau$, et nous avons $u'(\tau) = \tau^2 z'(\tau) + z(\tau)$, et l'équation 97.951 devient

$$z'(\tau) = \frac{1}{2\tau}, \tag{97.952}$$

ce qui donne $z = \frac{1}{2}\ln(k\tau)$, et donc

$$u(\tau) = \frac{1}{2}\tau \ln(k\tau). \tag{97.953}$$

Maintenant nous pouvons en tirer $y(t)$ en utilisant le fait que

$$u(t - \frac{3}{2}) = y(t) + \frac{5}{4}, \tag{97.954}$$

parce que $y(t) + \frac{5}{4} = u(\tau)$.

(2) $y' = \frac{t+2y+1}{2t+4y+3}$. Ici, le déterminant est nul, et nous écrivons l'équation en faisant apparaitre le numérateur au dénominateur (opération possible parce que déterminant est nul) :

$$y'(t) = \frac{2 + 2t + 1}{2(t + 2t + 1) + 1}, \tag{97.955}$$

et nous effectuons le changement de variable $u(t) = t + 2y(t) + 1$ qui donne immédiatement $y' = (u' - 1)/2$. L'équation se met alors sous la forme

$$u' = \frac{4u + 1}{2u + 1}, \tag{97.956}$$

qui est une équation à variables séparées. Pour la résoudre, il faut trouver une primitive de $(2u + 1)/(4u + 1)$, et la réponse est

$$t + \frac{1}{2}\left(\frac{1}{4}\ln(4u + 1) + u\right) = C. \tag{97.957}$$

C'est une formule implicite pour $u(t)$, et donc pour $y(t)$.

(3) $y' = (2t + 3y + 5)^2$. En posant simplement $u = 2t + 3y + 5$, nous trouvons

$$\frac{u'}{3} - \frac{2}{3} = u^2, \tag{97.958}$$

qui est une équation à variables séparées pour laquelle il faut une primitive de $1/(3u^2 + 2)$ par rapport à u. La réponse est

$$\frac{1}{\sqrt{6}} \arctan\left(\frac{3u}{\sqrt{6}}\right) = t + C. \tag{97.959}$$

97.24.7 Équation d'Euler

Exercise 222 exo_II-1-14

Résoudre les équations différentielles

ItemII114a

(1) $y'' + \frac{3y'}{t+1} + \frac{y}{(t+1)^2} = 0$,

ItemII114b

(2) $y'' + \frac{3y'}{t+1} + \frac{y}{(t+1)^2} = 1$,

(3) $y'' = \frac{a+b-1}{t}y' - \frac{ab}{t^2}y$

où a et b sont des constantes.

corr_II-1-14

Correction of the exercise 222

(1) $y'' + \frac{3y'}{t+1} + \frac{y}{(t+1)^2} = 0$. Après multiplication par $(t+1)^2$, nous avons l'équation sous forme usuelle :

$$y''(t+1)^2 + 3y'(t+1) + y = 0 \tag{97.960}$$

EqEulerSSFormeII114

Nous posons $\tau = \ln(t+1)$. Ici, afin d'éviter des confusions entre les dérivations par rapport à t et celles par rapport à τ, il est bon de changer également de nom pour la fonction inconnue :

$$z(\tau) = y\big(t(\tau)\big), \tag{97.961}$$

ou bien

$$z\big(\tau(t)\big) = y(t). \tag{97.962}$$

Maintenant nous calculons $y'(t)$ et $y''(t)$. Pour la première dérivée, nous avons

$$\begin{aligned} y'(t) &= \frac{d}{dt}\big(y(t)\big) \\ &= \frac{d}{dt}\big(z\big(\tau(t)\big)\big) \\ &= z'\big(\tau(t)\big)\frac{d\tau}{dt}(t) \\ &= z'\big(\tau(t)\big)\frac{1}{t+1}, \end{aligned} \tag{97.963}$$

et pour la seconde, nous utilisons Leibniz :

$$\begin{aligned} y''(t) &= \frac{d}{dt}\left(\frac{z'\big(\tau(t)\big)}{t+1}\right) \\ &= \frac{1}{(t+1)^2}\Big[z''\big(\tau(t)\big) - z'\big(\tau(t)\big)\Big]. \end{aligned} \tag{97.964}$$

En remplaçant cela et $y(t) = z\big(\tau(t)\big)$, l'équation de départ devient alors

$$z'' + 2z' + z = 0. \tag{97.965}$$

Pour résoudre cette équation, nous calculons les racines du polynôme caractéristique $r^2 + 2r + 1 = 0$, qui n'a qu'une racine double : $r = -1$. Le système fondamental de solutions est donc $z_1(\tau) = e^{-\tau}$ et $z_2(\tau) = \tau e^{-\tau}$. Nous trouvons donc

$$z(\tau) = Ae^{-\tau} + B\tau e^{-\tau}. \tag{97.966}$$

Nous pouvons maintenant effectuer le changement de variable inverse en utilisant $y(t) = z\big(\tau(t)\big)$:

$$y(t) = (A + B\tau(t))e^{-\tau(t)}, \tag{97.967}$$

et donc $y(t) = \big(B\ln(t+1) + A\big)(t+1)^{-1}$.

(2) $y'' + \frac{3y'}{t+1} + \frac{y}{(t+1)^2} = 1$. Cette équation ne diffère de la précédente que par le second membre. Nous repartons donc de l'équation (97.960) dans laquelle il faut modifier le second membre :

$$y''(t+1)^2 + 3y'(t+1) + y = (t+1)^2. \tag{97.968}$$

Nous connaissons déjà la solution générale de l'équation homogène associée. Il reste à trouver une particulière. Avec la variable τ, le second membre est $e^{2\tau}$. L'équation dont il nous faut une solution particulière est

$$z'' + 2y' + y = e^{2\tau}. \tag{97.969}$$

On essaye $z_P(\tau) = ae^{2\tau}$. C'est très vite vu que $a = \frac{1}{9}$ est ce qu'il faut.

(3) $y'' = \frac{a+b-1}{t}y' - \frac{ab}{t^2}y$. Nous posons $\tau = \ln(t)$, et nous tombons sur l'équation

$$z'' - (a+b)z' + abz = 0. \tag{97.970}$$

L'équation caractéristique a comme solutions $r_1 = a$ et $r_2 = b$. Si $a \neq b$, alors ce sont deux solutions distinctes.

(3a) Si $a \neq b$. Dans ce cas, nous avons $y(t) = z\big(\tau(t)\big) = Ae^{a\tau(t)} + Be^{b\tau(t)}$, et donc

$$y(t) = At^a + Bt^b. \tag{97.971}$$

(3b) Si $a = b$, alors $z(\tau) = Ae^{a\tau} + Be^{b\tau}$, et donc

$$y(t) = \big(A + B\ln(t)\big)t^a. \tag{97.972}$$

97.24.8 Équation dont on peut réduire l'ordre

Exercise 223 exo_II-1-15

On considère l'équation différentielle

$$y'' = \frac{-k^2}{y^2} \tag{97.973}$$

où k est une constante positive.

(1) Déterminer à une quadrature près, la solution générale de cette équation.

(2) Résoudre pour cette équation les problèmes de Cauchy suivants :

(2a) $y(0) = 1$, $y'(0) = \sqrt{2}k$,

(2b) $y(0) = 1$, $y'(0) = 2k$,

(2c) $y(0) = 1$, $y'(0) = k$,

(2d) $y(0) = 1$, $y'(0) = 0$.

corr_II-1-15

Correction of the exercise 223

Nous posons $z = y'$, c'est-à-dire

$$y'(t) = z\big(y(t)\big). \tag{97.974}$$

La dérivée seconde est donnée par

$$y''(t) = z'\big(y(t)\big)y'(t) = z'\big(y(t)\big)z\big(y(t)\big), \tag{97.975}$$

qui peut encore être écrit sous la forme ingénieuse

$$y''(t) = \frac{1}{2}\frac{d}{dy}\Big[z^2(y)\Big]_{y(t)}. \tag{97.976}$$

Si nous notons u la variable de z, nous avons

$$z^2(y) = 2\frac{k^2}{u} + C, \tag{97.977}$$

et donc

$$z(u) = \pm\sqrt{2\frac{k^2}{u} + C}. \tag{97.978}$$

Donc, pour refaire le passage vers les anciennes variables

$$z\big(y(t)\big) = \pm\left(\frac{2k^2}{y(t)} + C\right)^{1/2} = y'(t). \tag{97.979}$$

L'équation (très implicite) pour y prend alors la forme (voir l'astuce de la page 2457)

$$t - t_0 = \pm\int_{y(t_0)}^{y}\left(\frac{2k^2}{\xi} + C\right)^{-1/2} d\xi. \tag{97.980}$$

Le jeu est maintenant de trouver le C, ainsi que le signe $\pm$ à choisir et effectuer l'intégrale dans les différents cas proposés. Calculons donc cette intégrale :

$$I = \int\sqrt{\frac{\xi}{C\xi + k^2}}d\xi. \tag{97.981}$$

Commençons par le changement de variable

$$u = \sqrt{\xi/(C\xi + k^2)}, \quad d\xi = \frac{du}{u}\frac{k^2}{\sqrt{2(C\xi + k^2)^2}} = \frac{2}{k^2}u(C\xi + k^2)^2du. \tag{97.982}$$

L'intégrale à calculer devient alors

$$I = 2k^2\int\frac{u^2}{(1 - Cu^2)^2}du, \tag{97.983}$$

qui se traite avec le changement de variable

$$v = \frac{1}{1 - Cu^2}, \quad du = \frac{(1 - Cu^2)^2}{2uC}dv. \tag{97.984}$$

Ce sur quoi nous tombons est

$$I = \frac{k^2}{C}\int udv = \frac{k^2}{C}\left(\frac{u}{1 - Cu^2} - \int\frac{du}{1 - Cu^2}\right). \tag{97.985}$$

Cette intégrale dépend de la valeur de C.

Nous cherchons la valeur de C en reprenant la formule (97.979) pour $y'(t)$. Nous trouvons

$$y'(0) = \sqrt{2}k = \pm\left(\frac{2k^2}{y(0)} + C\right)^{1/2}. \tag{97.986}$$

Étant donné que $y(0) = 1$, nous en déduisons $C = 0$. Le signe à choisir est $+$ parce que $k > 0$. Les conditions initiales données étant en zéro, nous avons $t_0 = 0$, et nous nous retrouvons à calculer

$$t = \int_1^y\sqrt{\frac{\xi}{2k^2}}d\xi = \frac{2}{3}\frac{1}{\sqrt{2}k}\big(y^{3/2} - 1\big). \tag{97.987}$$

De là, sortir $y(t)$ est une opération algébrique simple :

$$y(t) = \left[\frac{3\sqrt{2}k}{2}\Big(t + \frac{2}{3}\frac{1}{\sqrt{2}k}\Big)\right]^{2/3} = \left(\frac{3k}{\sqrt{2}}t + 1\right)^{2/3}. \tag{97.988}$$

Cherchons maintenant les valeurs de C pour les autres données de Cauchy.

(2) $y(0) = 1$, $y'(0) = 2k$. L'équation (97.979) nous donne C :

$$2k = \pm \left(\frac{2k^2}{1+C} \right)^{1/2}, \tag{97.989}$$

donc $C = 2k^2$. Il faut prendre le signe + parce que k est positif.

(3) $2k^2 + C = k^2$, donc $C = -k^2$,

(4) $2k^2 + C = 0$, donc $C = -2k^2$.

Dans ces trois cas, la formule générale (97.980) devient

$$t = [I]_0^y \tag{97.990}$$

EqtIyII115

parce que $t_0 = 0$ dans les trois cas.

Passons à la résolution du second problème de Cauchy. L'intégrale (97.985) peut être calculée :

$$I = \frac{k^2}{C} \left(\frac{u}{1-k^2u^2} - \int \frac{du}{1-k^2u^2} \right) = \frac{u}{1-k^2u^2} - \frac{1}{2k} \ln \left| \frac{ku+1}{ku-1} \right|. \tag{97.991}$$

Il faut maintenant faire les changements de variables inverse :

$$u = \frac{1}{k} \sqrt{\frac{\xi}{\xi+1}}. \tag{97.992}$$

EqufracII115cici

L'équation (97.990) donne

$$\left[\frac{k^2}{C} \left(\frac{u}{1-k^2u^2} - \int \frac{du}{1-k^2u^2} \right) = \frac{u}{1-k^2u^2} - \frac{1}{2k} \ln \left| \frac{ku+1}{ku-1} \right| \right]_0^y \tag{97.993}$$

dans laquelle il faut remettre (97.992). Après quelques calculs, nous trouvons

$$-\frac{\sqrt{2}}{k} - \frac{1}{2k} \ln \left| \frac{\sqrt{2}+1}{\sqrt{2}-1} \right|. \tag{97.994}$$

Passons au second problème de Cauchy. Cette fois,

$$I = \frac{-u}{1+k^2u^2} + \frac{1}{k} \arctan(ku), \tag{97.995}$$

et

$$u = \frac{1}{k} \sqrt{\frac{\xi}{1-\xi}}. \tag{97.996}$$

Encore une fois, il faut utiliser la formule $t = [I]_0^t$, et puis remplacer.

Faisons le troisième problème de Cauchy. Cette fois, $C = -2k^2$. Étant donné que $y'(0) = 0$ et $y(0) = 1$, l'équation de départ devient

$$y''(0) = -k^2, \tag{97.997}$$

qui est négative. Mais comme $y'(0) = 0$, la dérivée $y'(t)$ est négative, en tout cas pour les petits t. Par conséquent, $0 < y(t) < 1$ sur un voisinage de 0 (parce que $y(0) = 1$), donc la fraction dans

$$y'(t) = \pm \left(\frac{2k^2 - 2k^2y(t)}{y(t)} \right) \tag{97.998}$$

est positive, et il faut choisir le signe $-$ afin que $y'(t)$ soit négatif. Nous devons donc, cette fois, utiliser la formule

$$t = -[I]_0^{y(t)}, \tag{97.999}$$

et tout ce qui s'ensuit.

Exercise 224 exo_II-1-16

Résoudre

(1) $y'' = (y')^2 + y^2$,
(2) $ayy'' + b(y')^2 + cy^2 = 0$

où a et b sont des constantes réelles. corr_II-1-16

Correction of the exercise 224

(1) $y'' = y'^2 + y^2$. Étant donné que la variable t n'intervient pas dans l'équation, poser

$$z\big(y(t)\big) = y'(t) \tag{97.1000}$$

permet de diminuer l'ordre de l'équation. Le calcul de y' et y'' en termes de z est usuel, et il faut encore utiliser l'astuce

$$y''(t) = \frac{1}{2}\frac{d}{dy}\Big(z^2(y)\Big)_{y=y(t)}. \tag{97.1001}$$

L'équation à résoudre devient

$$\frac{1}{2}\frac{d}{dy}\Big(z^2(y)\Big)_{y=y(t)} = z\big(y(t)\big)^2 + y(t)^2, \tag{97.1002}$$

où maintenant nous voyons $y(t)$ comme variable indépendante, ce qui fait que la fonction $z(u)$ vérifie l'équation

$$\frac{1}{2}\frac{d}{du}(z^2) = z^2 + u^2. \tag{97.1003}$$

Pour résoudre cela, nous posons $v(z) = z^2$, et alors v doit vérifier $\frac{v'}{2} = v + u^2$, dont la solution est

$$v(u) = Ce^{2u} - (u^2 + u + \frac{1}{2}). \tag{97.1004}$$

Nous trouvons donc l'équation suivante pour y' :

$$y'(t) = \big(Ce^{2y} - y^2 - y - \frac{1}{2}\big), \tag{97.1005}$$

que nous pouvons résoudre à quadrature près :

$$t - t_0 = \int_{y_0}^{y} \frac{dy}{\big(Ce^{2y} - y^2 - y - \frac{1}{2}\big)}. \tag{97.1006}$$

(2) $ayy'' + by'^2 + cy^2$. Dans ce cas-ci, poser $z\big(y(t)\big) = y'(t)$ n'est pas suffisant ; nous posons plutôt

$$z\big(y(t)\big) = y'(t)^2. \tag{97.1007}$$

La dérivation par rapport à t de cette équation donne

$$z'\big(y(t)\big)y'(t) = 2y'(t)y''(t) \tag{97.1008}$$ EqDerrtChII116

où z' désigne la dérivée de z par rapport à sa variable (c'est-à-dire y, et non t). Simplifions [22] par $y'(t)$, et remplaçons dans l'équation de départ :

$$ay(t)\frac{1}{2}z'\big(y(t)\big) + bz\big(y(t)\big) + cy(t)^2 = 0. \tag{97.1009}$$

Si nous voyons maintenant $y(t)$ comme variable de la fonction z, la fonction $z(y)$ vérifie l'équation différentielle

$$auz' + 2bz + 2cu^2 = 0, \tag{97.1010}$$

qui est une équation linéaire. L'équation homogène a pour solution

$$z_H(u) = Au^{-2b/a}. \tag{97.1011}$$

22. Je ne sais pas très bien ce qu'il se passe si y' vaut zéro en un point.

La méthode de variation des constantes va fournir la solution générale à l'équation non homogène. L'équation différentielle que nous trouvons pour $A(u)$ est

$$A'(u) = -\frac{2c}{a}u^{1+\frac{2b}{a}}. \tag{97.1012}$$

Avant d'écrire $A(u)$, il faut distinguer deux cas. En effet, u^α s'intègre différemment suivant que $\alpha = -1$ ou non. Les deux cas à distinguer sont

(2a) $1 + 2b/a = -1$, c'est-à-dire $a = -b$,

(2b) $1 + 2b/a \neq -1$, c'est-à-dire $a \neq -b$.

Dans le premier cas, le plus simple est en fait de remonter à l'équation de départ et d'y mettre $a = -b$. Ce que nous trouvons est

$$\frac{yy'' - y'^2}{y^2} = -\frac{c}{a}. \tag{97.1013}$$

Si on remarque que le membre de gauche est $\big(\ln|y|\big)''$, alors on trouve

$$\ln|y| = -\frac{c}{2a}t^2 + At + B, \tag{97.1014}$$

ou encore

$$y(t) = \alpha \exp\Big(-\frac{c}{2a}t^2 + At\Big) \tag{97.1015}$$

où α est une constante positive parce qu'elle provient de e^B.

Pour le second cas ($a \neq -b$), nous continuons en résolvant (97.1012) :

$$A(u) = -\frac{c}{a+b}u^{2(1+b/a)} + L, \tag{97.1016}$$

et $z(u) = A(u)u^{-2b/a}$. En substituant l'expression (97.1016) dans $z(u)$,

$$z(u) = -\frac{c}{a+b}u^2 + Lu^{-2b/a}. \tag{97.1017}$$

Maintenant nous nous souvenons que $z\big(y(t)\big) = y'(t)^2$, donc nous trouvons une nouvelle équation pour $y(t)$:

$$y'(t)^2 = -\frac{c}{a+b}y(t)^2 + Ly(y)^{-2b/a}. \tag{97.1018}$$

Cette équation est du type

$$f(y, y') = g(y)y'^2 + h(y), \tag{97.1019}$$

avec $g(y) = 1$, $f(y, y') = 0$ et $h(y) = -\frac{c}{a+b}y^2 + Ly^{-2b/a}$. Il y a une méthode pour elle à la page II.79, mais je propose d'arrêter ici.

97.24.9 Quelques exemples d'équations résolubles

Exercise 225 exo_II-1-17

Résoudre (éventuellement à quadrature près)

(1) $ty' - y + y\ln(y) = t\ln(t)$,

(2) $3y^2y' - ay^3 = 1 + t$ avec $a \neq 0$,

(3) $(y\cos(t))' = ky^3$,

(4) $2tyy' - y^2 - t^2 = 0$,

(5) $y' = ty(t + y) - t(1 + t)$,

(6) $t^2y' + y^2 - yt - t^2 = 0$,

(7) $(7y^2 - 4t - 3ty - 3t^2)y' = y(2 + 3t + y)$,

(8) $y'' = y'(2 + y'\tan(y))$,

(9) $y' = \frac{-ty}{t^2+y^2}$.

corr_II-1-17

Correction of the exercise 225

(1) Si nous récrivons l'équation sous la forme

$$y' = \frac{l\ln(t) + y - t\ln(y)}{t}, \tag{97.1020}$$

nous voyons que l'équation est homogène. Si nous posons $v = yt$, nous trouvons

$$\ln(t) = \int \frac{dv}{\ln(v)}, \tag{97.1021}$$

qui est résolue à quadrature près.

(2) C'est une équation de Bernoulli, sous la forme

$$3y' - ay = (1+t)y^{-2}. \tag{97.1022}$$

En suivant les notations de l'équation (32.337),

$$\begin{aligned} a(t) &= a \\ b(t) &= \frac{1+t}{3} \\ \alpha &= -2. \end{aligned} \tag{97.1023}$$

Nous posons $z = y^3$, et nous avons l'équation différentielle

$$z' - 3au = \frac{1+t}{3}, \tag{97.1024}$$

qui est linéaire. La solution pour y est donnée par

$$y^3 = Ce^{at} - \frac{1}{a^2}(at + a + 1). \tag{97.1025}$$

(3) Après avoir utilisé une règle de Leibniz pour dériver $y\cos(t)$, nous mettons l'équation sous la forme

$$y' = \frac{ky^3}{\cos(t)} + \frac{y\sin(t)}{\cos(t)}, \tag{97.1026}$$

qui est une équation de Bernoulli avec

$$\begin{aligned} a(t) &= \tan(t) \\ b(t) &= \frac{k}{\cos(t)} \\ \alpha &= 3. \end{aligned} \tag{97.1027}$$

La solution est

$$z = \frac{1}{y^2} = C\cos(t)^2 - \frac{k}{2}\cos(t)^2 \ln\left|\frac{1+\sin(t)}{1-\sin(t)}\right| - k\sin(t). \tag{97.1028}$$

(4) $2tyy' - y^2 - t^2 = 0$. En posant $z = y^2$, on trouve $z' = 2yy'$, et donc

$$tz' - z = t^2, \tag{97.1029}$$

qui est une équation linéaire. Cette équation peut soit se résoudre comme une équation linéaire, soit en écrivant

$$\frac{tz' - z}{t^2} = 1, \tag{97.1030}$$

et en remarquant que le membre de gauche n'est autre que $(z/t)'$. Donc $\frac{z}{t} = t + C$, et

$$z = y^2 = t^2 + Ct. \tag{97.1031}$$

(5) $y' = ty(t+y) - t(1+t)$. C'est une équation de Ricatti avec

$$\begin{aligned} a(t) &= t \\ b(t) &= t^2 \\ c(t) &= -t(1+t), \end{aligned} \tag{97.1032}$$

dont on devine la solution particulière $y(t) = 1$. En suivant le schéma général, nous posons $y(t) = 1 + \frac{1}{u(t)}$, et nous trouvons l'équation linéaire

$$u' = -(2t + t^2)u - t \tag{97.1033}$$

pour u. L'équation homogène a pour solution

$$u_H = K \exp\big(-t^2 + \frac{t^3}{3}\big). \tag{97.1034}$$

La méthode de variation des constantes donne K sous la forme

$$K = -\int \exp\big(t^2 + \frac{t^3}{3}\big). \tag{97.1035}$$

(6) $t^2y' + y^2 - yt - t^2 = 0$. Cette équation est Ricatta (avec $y = t$ comme solution particulière), mais elle est également une équation homogène. En posant $y = tv$, nous trouvons l'équation

$$tv' + v^2 = 1, \tag{97.1036}$$

qui est une équation à variables séparées.

(7) $(7y^2 - 4t - 3ty - 3t^2)y' = y(2 + 3t + y)$. C'est une équation presque exacte, dont $M = y$ est un facteur intégrant. Solution :

$$y^2\big(\frac{7}{4}y^2 - 2t - ty - \frac{3}{2}t^2\big) = C. \tag{97.1037}$$

(8) $y'' = y'(2 + y'\tan(y))$. La variable t n'entre n'intervient pas explicitement dans l'équation, donc nous posons $z\big(y(t)\big) = y'(t)$. Par calcul de dérivation,

$$y''(t) = \frac{dz}{dy}\big(y(t)\big) \cdot \underbrace{y'(t)}_{=z}, \tag{97.1038}$$

et l'équation de départ s'écrit

$$\frac{dz}{dy}\big(y(t)\big) \cdot z\big(y(t)\big) = z\big(y(t)\big) \cdot \Big(2 + z\big(y(t)\big) + \tan\big(y(t)\big)\Big). \tag{97.1039}$$

L'équation pour $z(u)$ est donc

$$z'z = z\big(2 + z\tan(u)\big). \tag{97.1040}$$

Une simplification par z (ce qui revient à perdre les solutions constantes pour y) mène à l'équation $z' = 2 + z\tan(u)$, dont les solutions sont données par

$$z = 2\tan(y) + \frac{2K}{\cos(y)}, \tag{97.1041}$$

qu'il faut maintenant identifier y'. L'équation que nous trouvons alors pour y est

$$y' = \frac{2\big(\sin(y) + K\big)}{\cos(y)}, \tag{97.1042}$$

que nous pouvons récrire sous la forme

$$\big(\sin(y) + K\big)' = 2\big(\sin(y) + K\big), \tag{97.1043}$$

et dont les solutions sont donc données par

$$\sin(y) + K = Le^{2t}. \tag{97.1044}$$

(9) $y' = -ty/(t^2+y^2)$. C'est une équation homogène, donc on pose $v(t) = y(t)/t$ et on trouve l'équation

$$tv' + v = -\frac{v}{1+v^2} \tag{97.1045}$$

que l'on peut remettre sous la forme

$$\frac{1}{t}\frac{dt}{dv} = \frac{1+v^2}{v(2+v^2)}, \tag{97.1046}$$

ou encore

$$\frac{dt}{t} = \frac{1+v^2}{v(2+v^2)}dv. \tag{97.1047}$$

Afin d'intégrer le second membre, nous décomposons mettons la fraction en fractions simples que nous cherchons sous la forme

$$\frac{1+v^2}{v(2+v^2)} = \frac{A}{v} + \frac{B}{2+v^2}. \tag{97.1048}$$

Un calcul montre que

$$\frac{1+v^2}{v(2+v^2)} = \frac{1}{2v} + \frac{v}{2(2+v^2)}. \tag{97.1049}$$

L'intégrale est maintenant aisée, et nous trouvons

$$\ln\left|\frac{1}{t}\right| = \frac{1}{2}\ln|v| + \frac{1}{4}\ln|2+v^2| + C. \tag{97.1050}$$

En manipulant les logarithmes, nous trouvons

$$\ln\left|\left(\frac{y^2}{t^2}+2\right) + \left(\frac{y}{t}\right)^2\right| = \ln|Ct^{-4}|. \tag{97.1051}$$

Étant donné que le logarithme est injectif sur $\mathbb{R}$, nous pouvons les simplifier. Le résultat est

$$y^2(y^2+2t^2) = C. \tag{97.1052}$$

97.24.10 L'équation différentielle linéaire du second ordre

Exercise 226 exo_II-1-20

On considère l'équation différentielle

$$t(2-3t)(t-1)y'' - 2(2t-1)y' - 2(1-3t)y = 0. \tag{97.1053}$$

(1) Montrer qu'il existe un seul réel a tel que $y = y^a$ soit solution de l'équation différentielle.

(2) Montrer que le changement de fonction inconnue $y = t^a x$ permet de déterminer la solution générale de l'équation.

corr_II-1-20

Correction of the exercise 226

Exercise 227 exo_II-1-21

On considère l'équation différentielle linéaire du second ordre

$$A(t)y'' + B(t)y' + C(t)y = 0. \tag{97.1054}$$

(1) Montrer que si l'on possède une solution $y_1(t)$ de cette équation, on peut en déduire la solution générale par « double quadrature ». S'inspirer de l'exercice 226.

(2) Montrer que l'équation différentielle peut se ramener à la forme

$$\big(p(t)y'\big)' + q(t)y = 0 \tag{97.1055}$$

et que si $y_1(t)$ en est une solution, sa solution générale est :

$$y(t) = y_1(t)\Big(C_1 + C_2\int_{t_0}^{t}\frac{d\xi}{p(\xi)y_1^2(\xi)}\Big). \tag{97.1056}$$

(3) Montrer que si l'on connaît une solution y_1 de l'équation

$$\big(p(t)y'\big)' + q(t)y = 0, \tag{97.1057}$$

on peut déterminer la solution générale de

$$\big(p(t)y'\big)' + q(t)y = r(t). \tag{97.1058}$$

(4) Résoudre l'équation différentielle

$$t(2-3t)(t-1)y'' - 2(2t-1)y' - 2(1-3t)y = 1. \tag{97.1059}$$

corr_II-1-21

Correction of the exercise 227

Exercise 228 exo_II-1-22

Monter que le changement de fonction inconnue

$$z(t) = \exp\big(-\int a(t)y(t)dt\big) \tag{97.1060}$$

ramène l'équation de Ricatti

$$y' = a(t)y^2 + b(t)y + c(t) \tag{97.1061}$$

à une équation différentielle linéaire du second ordre en la fonction $z(t)$. Quelle propriété de l'équation de Ricatti peut-on alors déduire de l'exercice 227 ?

corr_II-1-22

Correction of the exercise 228

Exercise 229 exo_II-1-23

Résoudre

$$y'' = (1+t^2)y. \tag{97.1062}$$

corr_II-1-23

Correction of the exercise 229

Exercise 230 exo_II-1-24

Résoudre par développement en série autour de $t = 0$ les équations suivantes :

(1) $(1-t^2)y'' - 4ty' - 2y = 0$,

(2) $y'' + 2ty' + 2y = 0$,

(3) $(1-\frac{t^2}{2})y'' + 4ty' - 10y = 0$.

corr_II-1-24

Correction of the exercise 230

97.24.11 Équation différentielle implicite du premier ordre

Exercise 231 exo_II-1-18

Considérons la famille de droites données sous forme paramétrique (z est le paramètre)

$$\begin{cases} x = z + 4 & \text{(97.1063a)} \\ y = 6z + C & \text{(97.1063b)} \end{cases}$$

Laquelle de ces droites passe par le point $(9, 32)$?

Résoudre les équations différentielles

(1) $t = \frac{a+by'^2}{y'}$

(2) $y = \frac{a+by'^2}{y'}$

où a et b sont des constantes réelles. Discuter, pour chacune des équations, le nombre de solutions au problème de Cauchy $y(t_0) = y_0$; comparer en particulier les cas $a = b = 1$ et $a = b = -1$. corr_II-1-18

Correction of the exercise 231

La solution de ce type d'équation est donnée sous forme paramétrique aux pages II.82 et II.84 du cours.

(1) Nous posons $z(t) = y'(t)$, ce qui implique

$$t(z) = \frac{a + bz^2}{z} = f(z). \tag{97.1064}$$

Cette équation doit être vue comme t donné en termes du paramètre z. Le jeu est maintenant de trouver y en fonction du même paramètre z. Ainsi nous aurons la courbe $y(t)$ sous forme paramétrique. Afin d'utiliser la formule du cours, nous calculons

$$f'(z) = \frac{bz^2 - a}{z^2}, \tag{97.1065}$$

et nous posons

$$\begin{aligned} y &= y_0 + \int_{z_0}^{z} \xi f'(\xi) d\xi + C \\ &= y_0 + \int_{z_0}^{z} (b\xi + \frac{a}{\xi}) d\xi + C \\ &= y_0 + b\left(\frac{z^2}{2} - \frac{z_0^2}{2}\right) - a \ln|z/z_0| + C. \end{aligned} \tag{97.1066}$$

Dans cette expression pour $y(z)$, nous pouvons rassembler toutes les constantes arbitraires (y_0, C et z_0) en une seule :

$$y(z) = \frac{b}{2}z^2 - a\ln(z) + C. \tag{97.1067}$$

Pour le problème de Cauchy $y(t_0) = y_0$, il faut d'abord voir pour quelle valeur du paramètre z nous avons $t(z) = t_0$. Cela se fait en résolvant l'équation

$$t_0 = \frac{a + bz_0^2}{z_0} \tag{97.1068}$$

par rapport à z_0. Les solutions sont données par

$$z_0 = \frac{t_0 \pm \sqrt{t_0^2 - 4ab}}{2b}. \tag{97.1069}$$

Noter que si $b = 0$, alors l'équation différentielle de départ est évidente [23], nous supposons donc que $b \neq 0$.

Si $a = b = 1$, nous avons les possibilités

23. Prouvez-vous qu'elle est évidente en écrianve toutes les solutions ! !

(1a) Si $|t_0| > 2$, deux solutions,

(1b) Si $|t_0| < 2$, pas de solutions,

(1c) Si $|t_0| = 2$, une seule solutions.

Si par contre $a = -1 = b$, alors nous avons

$$z_0 = \frac{t_0 \pm \sqrt{t_0^2 + 4}}{-2}, \tag{97.1070}$$

et donc toujours deux solutions.

Dans tous les cas, la solution se trouve sous forme paramétrique

$$\begin{cases} t(z) = \dfrac{a + bz^2}{z}, & \text{(97.1071a)} \\ y(z) = \dfrac{b}{2}z^2 - a\ln(z) + C. & \text{(97.1071b)} \end{cases}$$

Une fois que z_0 est trouvé en fonction de t_0, il convient de résoudre $y(z_0) = y_0$ pour fixer la constante C.

(2) En posant $z = y'$, nous écrivons

$$y = f(z) = \frac{a + bz^2}{z}, \tag{97.1072}$$

qui doit, encore une fois, être vue comme une équation paramétrique. Pour appliquer la formule du cours, nous calculons

$$f'(z) = \frac{bz^2 - a}{z^2}, \tag{97.1073}$$

et nous posons

$$\begin{aligned} t &= t_0 + \int_{z_0}^{z} \frac{b\xi^2 - a}{\xi^3} d\xi + C \\ &= t_0 + b\ln(|z/z_0|) + 2a\left(\frac{1}{z^2} - \frac{1}{z_0^2}\right) + C \end{aligned} \tag{97.1074}$$

Encore une fois, il est important de regrouper toutes les constantes arbitraires (z_0, t_0 et C) en une seule :

$$t(z) = b\ln(z) + \frac{2a}{z^2} + C. \tag{97.1075}$$

Pour résoudre le problème de Cauchy $y(t_0) = y_0$, nous commençons par trouver pour quelle valeur z_0 du paramètre z nous avons $t(z_0) = t_0$. L'équation à résoudre est

$$y_0 = \frac{a + bz_0^2}{z_0}, \tag{97.1076}$$

et les solutions sont

$$z_0 = \frac{y_0 \pm \sqrt{y_0^2 - 4ab}}{2b}. \tag{97.1077}$$

Si $a = b = 1$, nous avons les possibilités

(2a) Si $|y_0| < 2$, alors il y a deux solutions,

(2b) Si $|y_0| > 2$, alors il n'y a pas de solutions,

(2c) Si $|y_0| = 2$, alors il y a une seule solution.

Si $a = -b = 1$, par contre, il y a toujours deux solutions.

Other resolution

(1) Nous commençons par dilater les données $y = az$ et $y = \sqrt{|ab|}u$, de telle façon à mettre l'équation sous la forme

$$u = \frac{1 + \epsilon z'^2}{z'} \tag{97.1078}$$

où $\epsilon = \pm 1$, selon les valeurs de a et b. La solution de l'équation (97.1078) est donné sous forme paramétrique à la page II.82 du cours :

$$\begin{cases} u(\lambda) = \dfrac{1 + \epsilon\lambda^2}{\lambda} & \text{(97.1079a)} \\ z(\lambda) = C + \displaystyle\int_{\lambda_0}^{\lambda} \xi f'(\xi) d\xi & \text{(97.1079b)} \end{cases}$$

où $f(\xi) = \frac{1+\epsilon\xi^2}{\xi}$. L'intégrale n'est pas très compliquée à effectuer :

$$z(\lambda) = C + \left[\frac{\epsilon\xi^2}{2} - \ln(\xi) \right]_{\lambda_0}^{\lambda} = C + \frac{\epsilon}{2}(\lambda^2 - \lambda_0^2) + \ln\left(\frac{\lambda_0}{\lambda}\right). \tag{97.1080}$$

Le C peut être redéfini pour englober toutes les termes contenant λ_0 (qui est une constante arbitraire). Nous avons donc

$$\begin{cases} u(\lambda) = \dfrac{1 + \epsilon\lambda^2}{\lambda} & \text{(97.1081a)} \\ z(\lambda) = C + \dfrac{\epsilon\lambda^2}{2} - \ln(\lambda). & \text{(97.1081b)} \end{cases}$$

Résolvons le problème de Cauchy $z(u_0) = z_0$. Pour cela, cherchons la valeur λ_0 du paramètre pour lequel $u(\lambda) = u_0$. Il vient l'équation

$$\epsilon\lambda_0^2 - u_0\lambda + 1 = 0, \tag{97.1082}$$

dont la solution est

$$\lambda_0 = \frac{u_0 \pm \sqrt{u_0^2 - 4\epsilon}}{2\epsilon}. \tag{97.1083}$$

Il suffit maintenant de remplacer cette valeur dans l'équation (97.1081b) de $z(\lambda)$. Il faut distinguer les cas $\epsilon = \pm 1$.

Étudions d'abord le cas $\epsilon = -1$. Dans ce cas, $\sqrt{u_0^2 - 4\epsilon}$ existe toujours (c'est-à-dire pour tout u_0). Nous trouvons donc deux valeurs pour la constante C dans (97.1081b), données par

$$C^{(i)} = z_0 - \frac{\epsilon(\lambda_0^{(i)})^2}{2} - \ln(\lambda_0^{(i)}) \tag{97.1084}$$

où

$$\lambda_0^{(1)} = \frac{-u_0 - \sqrt{u_0^2 + 4}}{2}, \qquad \lambda_0^{(2)} = \frac{-u_0 + \sqrt{u_0^2 + 4}}{2} \tag{97.1085}$$

Dans le cas où $\epsilon = 1$, la racine n'existe pas toujours.

(1a) Si $-2 < u_0 < 2$, il n'y a aucune solutions.

(1b) Si $u_0 = \pm 2$, il y a une seule solution.

(1c) Si $u_0 > 2$ ou $u_0 < -2$, il y a deux solutions.

Toutes ces solutions s'obtiennent par la même méthode que plus haut.

Cette multiplicité de solutions peut être vue en récrivant l'équation de départ (97.1078) sous la forme

$$z' = \frac{u \pm \sqrt{u^2 + 4\epsilon}}{2\epsilon}, \tag{97.1086}$$

qui sont réellement deux équations différentielles différentes qui peuvent être résolues moyennant le calcul de l'intégrale $\int \sqrt{u^2 + 4\epsilon} du$.

(2) $y = (a + by'^2)/y'$. Comme précédemment, nous écrivons l'équation sous la forme $y = (1 + \epsilon y'^2)/y'$ avec $\epsilon = \pm 1$. Cette fois, la forme paramétrique de la solution est donné à la page II.84 :

$$\begin{cases} y(\lambda) = \dfrac{1 + \epsilon\lambda^2}{\lambda} & \text{(97.1087a)} \\ t(\lambda) = t_0 + \displaystyle\int_{\lambda_0}^{\lambda} \frac{f'(\xi)}{\xi} d\xi & \text{(97.1087b)} \end{cases}$$

où $f(\xi) = (1 + \epsilon\xi^2)/\xi$. L'intégrale n'est pas très difficile à effectuer, et nous avons

$$\begin{cases} y(\lambda) = \dfrac{1 + \epsilon\lambda^2}{\lambda} & \text{(97.1088a)} \\ t(\lambda) = C + \epsilon \ln(\lambda) + \dfrac{1}{2\lambda^2}. & \text{(97.1088b)} \end{cases}$$

Encore une fois, il y avait moyen de résoudre cette équation par une autre voie : $yy' = 1 + \epsilon y'^2$ peut être résolue algébriquement en fonction de y' pour donner

$$y' = \frac{y \pm \sqrt{y^2 + 4\epsilon}}{2\epsilon}, \tag{97.1089}$$

qui sont deux équations à variables séparées.

Exercise 232 exo_II-1-19

Résoudre l'équation différentielle

$$y = (1 + y'^2)^{1/2}. \tag{97.1090}$$

Représenter graphiquement les solutions. Donner les solutions correspondant aux conditions initiales suivantes :

(1) $y(0) = \cosh(1)$,

(2) $y(0) = 1$,

(3) $y(0) = 0$.

corr_II-1-19

Correction of the exercise 232

$y = (1 + y'^2)^{1/2}$. Avant de passer en paramétrique, il faut vérifier que y' n'est pas une constante. En effet, si tel était le cas, alors l'équation $y(\lambda) = (1 + \lambda^2)^{1/2}$ ne serait valable que pour une seule valeur du paramètre ($\lambda = C$), ce qui est un peu pas trop bien pour une équation paramétrique.

Quoi qu'il en soit, la seule solution avec $y' = C$ est $y' = 0$ et donc $y = 1$.

Les autres solutions sont obtenues de façon paramétrique :

$$\begin{cases} y(\lambda) = (1 + \lambda^2)^{1/3} & \text{(97.1091a)} \\ t(\lambda) = t_0 + \displaystyle\int_{\lambda_0}^{\lambda} \frac{f'(\xi)}{\xi} d\xi & \text{(97.1091b)} \end{cases}$$

où $f(\xi) = (1 + \xi^2)^{1/2}$. Nous devons donc intégrer $1/\sqrt{1 + \xi^2}$. Le formulaire donne

$$\int \frac{dv}{\sqrt{v^2 \pm a^2}} = \ln|v + \sqrt{v^2 \pm a^2}| + C. \tag{97.1092}$$

Lemme 97.3.
Nous avons la formule

$$\operatorname{arcsinh}(a) = \ln(a + \sqrt{a^2 + 1}) \tag{97.1093}$$

pour l'inverse du sinus hyperbolique.

Démonstration. Nous avons $x = \operatorname{arcsinh}(a)$, lorsque $e^x - e^{-x} = 2a$. Posons $y = e^x$, et résolvons

$$y + \frac{1}{y} = 2a \tag{97.1094}$$

par rapport à y. Il y a deux solutions : $y = a \pm \sqrt{a^2 + 1}$. La solution $a - \sqrt{a^2 + 1}$ est à rejeter parce que $y = e^x > 0$. Donc il ne reste que

$$x = \ln(a + \sqrt{a^2 + 1}). \tag{97.1095}$$

□

Par conséquent,

$$t(\lambda) = t_0 + \operatorname{arcsinh}(\lambda) - \operatorname{arcsinh}(\lambda_0). \tag{97.1096}$$

Cette relation peut être inversée pour connaitre λ en fonction de t :

$$\lambda(t) = \sinh(t - C) \tag{97.1097}$$

où nous avons, comme toujours, regroupé toutes les constantes dans une seule constante C. De cette façon, nous pouvons donner une forme explicite pour $y(t)$:

$$y(t) = y\big(\lambda(t)\big) = \big(1 + \lambda(t)^2\big)^{1/2} = \cosh(t - C). \tag{97.1098}$$

Nous avons donc comme solution générale de l'équation différentielle : $y(t) = \cosh(t - C)$.

Other resolution

Nous pouvons résoudre $y = (1 + y'^2)^{1/2}$ algébriquement par rapport à $y'(t)$:

$$y'(t) = \pm\sqrt{y^2 - 1}, \tag{97.1099}$$

ce qui mène à

$$\frac{dy}{\sqrt{y^2 - 1}} = \pm dt, \tag{97.1100}$$

et donc $y(t) = \pm\cosh(t - C)$. Le choix de $\pm$ se fixe en remontant à l'équation de départ : $y = (1 + y'^2)^{1/2}$ demande $y > 0$, et donc le choix de $y = \cosh(t + C)$.

Nous pouvons à présent résoudre quelques problèmes de Cauchy.

(1) $y(0) = \cosh(1)$. Nous avons $\cosh(C) = \cosh(1)$, donc $C = \pm 1$ (je rappelle que le cosinus hyperbolique est une fonction paire), et donc deux solutions :

$$y(t) = \cosh(t \pm 1). \tag{97.1101}$$

(2) $y(0) = 1$. Nous trouvons facilement $C = 0$ et donc $y = \cosh(t)$. Ne pas oublier la solution $y = 1$ que nous avons mentionné plus haut.

(3) $y(0) = 0$, pas de solutions.

Exercise 233 exo_II-1-28

Résoudre l'équation de Clairaut

$$y = ty' + \frac{1}{y'^2} \tag{97.1102}$$

et les problèmes de Cauchy suivants :

(1) $y(1) = 2$,

(2) $y(2) = 3$,

(3) $y(0) = 1$,

(4) $y(0) = -1$,

(5) $y(3) = -2$.

corr_II-1-28

Correction of the exercise 233

La théorie se trouve à la page II.85 du cours. En suivant ces notations, nous avons ici $f(y') = 1/y'^2$, dont l'ensemble de définition est $\mathbb{R}_0$. Donc pour tout $c \in \mathbb{R}_0$, la droite

$$y_c(t) = ct + \frac{1}{c^2} \tag{97.1103}$$

EqII128lignes

est solution. Nous trouvons les constantes c qui correspondent au problème de Cauchy $y(t_0) = y_0$ en résolvant l'équation

$$c^2 y_0 - c^3 t_0 - 1 = 0. \tag{97.1104}$$

EqTrouverFDroitesII128

En vertu de ce qui est écrit au bas de la page II.86, l'enveloppe est donnée paramétriquement par

$$\begin{cases} t(\lambda) = -f'(\lambda) & \text{(97.1105a)} \\ y(\lambda) = -f'(\lambda)\lambda + f(\lambda) & \text{(97.1105b)} \end{cases}$$

où $f(\lambda) = 1/\lambda^2$. En remplaçant,

$$\begin{cases} t(\lambda) = \dfrac{2}{\lambda^3} & \text{(97.1106a)} \\ y(\lambda) = \dfrac{3}{\lambda^2}. & \text{(97.1106b)} \end{cases}$$

La relation $t(\lambda)$ peut être inversée : $\lambda = \sqrt[3]{2/t}$, et donc

$$y(t) = 3\left(\frac{2}{t}\right)^{-2/3}. \tag{97.1107}$$

Cette dernière équation est plus belle sous la forme

$$y_E^3(t) = \frac{27}{4}t^2. \tag{97.1108}$$

Étudions maintenant les problèmes de Cauchy proposés.

(1) $y(1) = 2$. Étant donné que le point $(1, 2)$ n'est pas sur y_E, regardons si il ne serait pas sur une des droites, c'est-à-dire : résolvons $y_c(1) = 2$ par rapport à c. Nous trouvons les solutions

$$c_1 = 1 \quad c_2 \quad = \frac{-\sqrt{5}+1}{2} c_3 \quad = \frac{\sqrt{5}+1}{2}. \tag{97.1109}$$

Les trois solutions correspondantes s'obtiennent en mettant c_i dans l'équation des droites (97.1103).

(2) $y(2) = 3$. Le point $(2, 3)$ est sur l'enveloppe, et l'équation (97.1104) à résoudre pour trouver les droites est

$$3c^2 - 2c^3 - 1 = 0, \tag{97.1110}$$

les solution sont alors $c = -\frac{1}{2}$ et $c = 1$. Les droites solutions sont donc

$$\begin{cases} y_1(t) = \dfrac{1}{2}t + 4 & \text{(97.1111a)} \\ y_2(t) = t + 1. & \text{(97.1111b)} \end{cases}$$

(3) $y(0) = 1$. Le point $(0, 1)$ n'est pas sur l'enveloppe, et les droites sont données par $c = \pm 1$.

(4) $y(0) = -1$. Ce n'est pas sur l'enveloppe, et il n'y a pas de solutions à l'équation (97.1104). Pas du tout de solutions dans ce cas-ci.

(5) $y(3) = -2$. Cela n'est pas sur l'enveloppe, et une seule droite $c = -1$, c'est-à-dire $y = -t + 1$.

Exercise 234 exo_II-1-25

Résoudre l'équation de Lagrange

$$y = y' + ty^2. \tag{97.1112}$$

corr_II-1-25

Correction of the exercise 234

Exercise 235 exo_II-1-26

Appliquer le principe d'intégration par différentiation (cf. 5.f, page II.89) pour résoudre l'équation

$$ty'^2 + (t - y)y' - y = 0. \tag{97.1113}$$

corr_II-1-26

Correction of the exercise 235

Exercise 236 exo_II-1-27

Dans un plan rapporté à des coordonnées cartésiennes Oxy, déterminer les courbes telles qu'en tout point de la courbe, le segment de sa tangente en ce point, compris entre les axes de coordonnées, ait une longueur constante.

corr_II-1-27

Correction of the exercise 236

97.25 Systèmes différentiels linéaires à coefficients constants

Exercise 237 exo_II-2-01

Résoudre les systèmes différentiels suivants :

(1)

$$\begin{cases} x' + x + y = 0 & (97.1114a) \\ y' - 2x - y = 0. & (97.1114b) \end{cases}$$

(2)

$$\begin{cases} x' + x + y = 0 & (97.1115a) \\ y' - x + 3y = 0. & (97.1115b) \end{cases}$$

(3)

$$\begin{cases} x' = 2x + y + z & (97.1116a) \\ y' = x + 2y + z & (97.1116b) \\ z' = x + y + 2z & (97.1116c) \end{cases}$$

(4)

$$\begin{cases} x_1' = x_3 & (97.1117a) \\ x_2' = 2x_1 + x_4 & (97.1117b) \\ x_3' = -x_1 + 2x_4 & (97.1117c) \\ x_4' = 0. & (97.1117d) \end{cases}$$

(5)

$$\begin{cases} x' = \dfrac{1}{2}(x - y + z) & (97.1118a) \\ t' = \dfrac{1}{2}(y - z) & (97.1118b) \\ z' = x & (97.1118c) \end{cases}$$

(6)

$$\begin{cases} x' = x - y + e^t & \text{(97.1119a)} \\ y' = x + y + e^{2t}. & \text{(97.1119b)} \end{cases}$$

corr_II-2-01

Correction of the exercise 237

(1) Le système s'écrit sous la forme

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{97.1120}$$

Les valeurs propres de la matrice sont données par l'annulation du déterminant

$$\begin{vmatrix} -1-\lambda & -1 \\ 2 & 1-\lambda \end{vmatrix} = 0. \tag{97.1121}$$

Les solutions sont $\lambda = \pm i$. Le vecteur propre pour la valeur propre i est donné par

$$v_1 = \begin{pmatrix} 1 \\ -i-1 \end{pmatrix} \tag{97.1122}$$

Le vecteur propre pour la valeur $-i$ est le complexe conjugué[24] de v_1 :

$$v_2 = v_1^* = \begin{pmatrix} 1 \\ i-1 \end{pmatrix}. \tag{97.1123}$$

Les deux valeurs propres sont donc dans le cas (3) de la recette de la page 2465. La solution du système s'écrit donc comme une combinaison linéaire

$$\begin{pmatrix} x \\ y \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ -i-1 \end{pmatrix} e^{it} + C_2 \begin{pmatrix} 1 \\ i-1 \end{pmatrix} e^{-it} \tag{97.1124}$$

où $C_1, C_2 \in \mathbb{C}$.
Cette solution est réelle si et seulement si $C_1 = \bar{C}_2$. Si $C_1 = a + bi$ et $C_2 = a - bi$, alors les solutions prennent la forme

$$\begin{cases} x(t) = a\cos(t) + b\sin(t) & \text{(97.1125a)} \\ y(t) = -(a+b)\cos(t) + (a-b)\sin(t). & \text{(97.1125b)} \end{cases}$$

Other resolution

Il y a moyen de résoudre cet exercice en calculant explicitement l'exponentielle de

$$A = \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix}. \tag{97.1126}$$

D'abord, nous considérons M_1, la matrice qui agit comme l'identité sur le vecteur propre de λ_1 et qui donne zéro sur le vecteur propre de λ_2, et M_2, le contraire[25]. Dans ce cas, pour toute fonction analytique (c'est-à-dire pour toute fonction qui se définit par un développement en série de puissances), nous avons

$$f(A) = f(\lambda_1)M_1 + f(\lambda_2)M_2. \tag{97.1127}$$

EqfAII201

24. Parce que si $Av = \lambda v$, alors $A^*v* = \lambda^* v^*$, tandis que $A^* = A$ et que les deux valeurs propres sont complexes conjuguées.

25. Notez déjà que cela n'est possible que parce que nous avons une base de vecteurs propres.

En faisant $f(x) = x^0$, nous avons $f(A) = \mathbb{1}$ et $f(\lambda_1) = f(\lambda_2) = 1$, donc

$$\mathbb{1} = M_1 + M_2. \tag{97.1128}$$

En prenant d'autre part $f(x) = x - i$, nous avons

$$A - i\mathbb{1} = -2iM_2. \tag{97.1129}$$

En combinant (97.1128) et (97.1129), nous trouvons

$$\begin{aligned} M_1 &= \frac{1}{2}(\mathbb{1} - iA) \\ M_2 &= \frac{1}{2}(\mathbb{1} + iA). \end{aligned} \tag{97.1130}$$

Donc, en utilisant la formule (97.1127) avec $f = \exp$, nous avons

$$e^{tA} = e^{it}\frac{1}{2}(\mathbb{1} - i1) + e^{-it}\frac{1}{2}(\mathbb{1} - iA). \tag{97.1131}$$

Maintenant, la solution du système est donnée par

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(t) - \sin(t) & -\sin(t) \\ 2\sin(t) & \cos(t) + \sin(t) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \tag{97.1132}$$

Tu sais comment vérifier que (97.1131) est bien l'exponentielle de tA ? En utilisant la proposition 1 de la page II.40. Dérivons donc l'équation (97.1131) par rapport à t, et posons $t = 0$.

$$\frac{d}{dt}\big(e^{it}\frac{1}{2}(\mathbb{1} - iA) + e^{-it}\frac{1}{2}(\mathbb{1} + iA)\big)_{t=0} = \frac{1}{2}(i\mathbb{1} + A - i\mathbb{1} + A) = A. \tag{97.1133}$$

(2) Dans ce système,

$$A = \begin{pmatrix} -1 & -1 \\ 1 & -3 \end{pmatrix}, \tag{97.1134}$$

et le polynôme caractéristique est

$$p(\lambda) = \lambda^2 + 4\lambda + 4 = 0. \tag{97.1135}$$

Cela a une solution unique $\lambda = -2$ de multiplicité deux. Il n'y a hélas que un seul vecteur propre (à multiple près) :

$$v = \begin{pmatrix} a \\ a \end{pmatrix}. \tag{97.1136}$$

Nous sommes donc dans le cas où il y a un seul vecteur pour une valeur propre multiple, c'est-à-dire le point (5) de la page 2466. Nous cherchons donc des solutions sous la forme

$$\begin{pmatrix} at + c \\ at + d \end{pmatrix} e^{-2t} \tag{97.1137}$$

où l'on a, conformément à la recette, prit le vecteur $\begin{pmatrix} a \\ a \end{pmatrix}$ comme coefficient de $e^{\lambda t}$. Nous injections cette solution dans le système pour tenter de fixer a, c et d. Après quelques calculs et après simplification par e^{-2t}, nous tombons sur le système

$$\begin{cases} a - c + d = 0 & \text{(97.1138a)} \\ a - c + d = 0, & \text{(97.1138b)} \end{cases}$$

donc $a = c - d$. La solution du système est donc

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (c-d)t + c \\ (c-d)t + d \end{pmatrix} e^{-2t}. \tag{97.1139}$$

Cette solution peut être rendue un peu plus jolie en changeant le nom des constantes :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} at+b \\ at+(b-a) \end{pmatrix} e^{-2t}. \tag{97.1140}$$

Cette solution est réelle lorsque $a, b \in \mathbb{R}$.

Other resolution

Nous pouvons, ici aussi, résoudre le système en calculant l'exponentielle de la matrice. Le problème est cependant plus compliqué parce que nous avons un seul vecteur propre. Il n'est donc pas possible de définir les matrices M_1 et M_2 de l'équation (97.1127).

C'est pour parer à cette carence de vecteurs propres qu'on a inventé la **forme de Jordan**. Prenons une matrice B, et mettons la sous la forme $B = \lambda\mathbb{1} + N$, c'est-à-dire

$$B = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}, \tag{97.1141}$$

et

$$N = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \end{pmatrix}. \tag{97.1142}$$

Il est vite vu qu'en calculant les puissances de N, la diagonale de 1 se décale. Dans le cas de matrices 2×2, nous avons $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, et $N^2 = 0$. Donc, $B^2 = (\lambda\mathbb{1} + N)^2 = \lambda^2\mathbb{1} + 2\lambda N$. En calculant la puissance suivante de B, nous trouvons

$$B^3 = B^2(\lambda\mathbb{1} + N) = \lambda^2\mathbb{1} + 3\lambda^2 N, \tag{97.1143}$$

et aucun terme en N^2 pacque $N^2 = 0$. Par conséquent, lorsque B est sous forme de Jordan,

$$f(B) = f(\lambda)\mathbb{1} + f'(\lambda)N. \tag{97.1144}$$ EqExpBJordanII201

Dans le problème qui nous occupe, A n'est pas sous forme de Jordan. Au lieu d'avoir (97.1144), nous avons

$$f(A) = f(\lambda)M_1 + f'(\lambda)M_2 \tag{97.1145}$$

où M_1 et M_2 sont des matrices à déterminer. Nous trouvons facilement que $M_1 = \mathbb{1}$ et $M_2 = A + 2\mathbb{1}$.

(3) C'est un système 3×3, avec pour matrice

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \tag{97.1146}$$

L'équation caractéristique est

$$-\lambda^3 + 6\lambda^2 - 9\lambda + 4 = (\lambda - 1)(-\lambda^2 + 5\lambda - 4) = 0. \tag{97.1147}$$

Les solutions sont $\lambda_1 = 4$ et $\lambda_2 = 1$. La seconde est de multiplicité deux. La valeur propre 4 a pour vecteur propre $\begin{pmatrix} a \\ a \\ a \end{pmatrix}$, et la valeur double 1 a deux vecteurs propres : $\begin{pmatrix} b \\ 0 \\ -b \end{pmatrix}$ et $\begin{pmatrix} 0 \\ c \\ -c \end{pmatrix}$.

Nous sommes donc dans un cas où toutes les valeurs propres ont leur bon nombre de vecteurs propres, et la solution est simplement

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ a \\ a \end{pmatrix} e^{4t} + \begin{pmatrix} b \\ 0 \\ b \end{pmatrix} e^{t} + \begin{pmatrix} 0 \\ c \\ -c \end{pmatrix} e^{t}. \tag{97.1148}$$

Other resolution

Il y a une autre façon de résoudre ce système. Remarquons en effet que $x'+y'+z' = 4(x+y+z)$, donc

$$x + y + z = Ke^{4t}. \tag{97.1149}$$

Par conséquent, nous avons

$$\begin{aligned} x' &= 2x + y + (-x - y + Ke^{4t}) \\ y' &= c + 2y + (-x - y + Ke^{4t}). \end{aligned} \tag{97.1150}$$

Nous en déduisons que

$$\begin{aligned} x' &= x + Ke^{4t} \\ y' &= y + Ke^{4t}, \end{aligned} \tag{97.1151}$$

et donc

$$\begin{aligned} x &= Ce^{t} + \frac{K}{3}e^{4t} \\ y &= De^{t} + \frac{K}{3}e^{4t}. \end{aligned} \tag{97.1152}$$

(4) La matrice est

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{97.1153}$$

Vu qu'il y a une ligne de zéros, le déterminant de cette matrice est nul, et il y aura certainement $\lambda = 0$ comme valeur propre. Le polynôme caractéristique est en effet $\lambda^4 + \lambda^2$, et les valeurs propres sont $\lambda_1 = i$, $\lambda_2 = -i$ et $\lambda_3 = \lambda_4 = 0$. Les vecteurs propres sont

$$v_1 = \begin{pmatrix} 1 \\ -2i \\ i \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2i \\ -i \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ a \\ 0 \\ 0 \end{pmatrix}. \tag{97.1154}$$

Il n'y a que un seul vecteur propre pour la valeur propre zéro. Les deux valeurs propres « comme il faut » fournissent les solutions

$$\bar{x} = C_1 \begin{pmatrix} 1 \\ -2i \\ i \\ 0 \end{pmatrix} e^{it} + C_2 \begin{pmatrix} 1 \\ 2i \\ -i \\ 0 \end{pmatrix} e^{-it}. \tag{97.1155}$$

Les solutions correspondantes à la troisième valeur propre sont à chercher sous la forme $\bar{x} = (\bar{a}t + \bar{b})e^{i\lambda_3 t}$ où $\bar{a} = \begin{pmatrix} 0 \\ a \\ 0 \\ 0 \end{pmatrix}$ et $\lambda_3 = 0$. Nous cherchons donc $\bar{x}$ sous la forme

$$\bar{x} = \begin{pmatrix} 0 + b_1 \\ a_2 t + b_2 \\ 0 + b_3 \\ 0 + b_4 \end{pmatrix}. \tag{97.1156}$$

Le système pour les constantes a_2 et b_i est

$$\begin{cases} 0 = b_3 & \text{(97.1157a)} \\ a_2 = 2b_1 + b_4 & \text{(97.1157b)} \\ 0 = -b_1 + 2b_4 & \text{(97.1157c)} \\ 0 = 0. & \text{(97.1157d)} \end{cases}$$

Nous en déduisons donc la solution générale du système proposé :

$$\bar{x} = C_1 \begin{pmatrix} 1 \\ -2i \\ i \\ 0 \end{pmatrix} e^{it} + C_2 \begin{pmatrix} 1 \\ 2i \\ -i \\ 0 \end{pmatrix} e^{-it} + \begin{pmatrix} 2a/5 \\ at + b_2 \\ 0 \\ a/5 \end{pmatrix} \tag{97.1158}$$

avec C_1, C_2, a et b_2 comme constantes.

(5) La matrice est

$$A = \begin{pmatrix} 1/2 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 \\ 1 & 0 & 0 \end{pmatrix}, \tag{97.1159}$$

dont le polynôme caractéristique est $\det(A - \lambda\mathbb{1}) = -\lambda^3 + \lambda^2 + \frac{\lambda}{4}$. Les solutions sont $\lambda_1 = 0$, $\lambda_2 = \frac{1+\sqrt{2}}{2}$ et $\lambda_3 = \frac{1-\sqrt{2}}{2}$; les vecteurs propres correspondants sont

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -\frac{\sqrt{2}}{2} - 1 \\ 1 \\ -\sqrt{2} \end{pmatrix}, \quad v_3 = \begin{pmatrix} \frac{\sqrt{2}}{2} - 1 \\ 1 \\ \sqrt{2} \end{pmatrix}. \tag{97.1160}$$

La solution du système s'écrit donc sous la forme

$$\bar{x} = A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + B \begin{pmatrix} -\frac{\sqrt{2}}{2} - 1 \\ 1 \\ -\sqrt{2} \end{pmatrix} e^{\lambda_2 t} + C \begin{pmatrix} \frac{\sqrt{2}}{2} - 1 \\ 1 \\ \sqrt{2} \end{pmatrix} e^{\lambda_3 t}. \tag{97.1161}$$

(6) Pour un système non homogène, il y a la proposition 1 de la page II.35 qui dit qu'il faut trouver la solution générale du système homogène associé, et lui ajouter une solution particulière du système non homogène. Dans le cas présent, le système homogène a pour solutions

$$\bar{x}_H = C_1 \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{(1+i)t} + C_2 \begin{pmatrix} 1 \\ i \end{pmatrix} e^{1-i}t. \tag{97.1162}$$

Pour le système non homogène, nous essayons une solution de la forme

$$\bar{x}_P = \begin{pmatrix} \alpha e^t + \beta e^{2t} \\ \gamma e^t + \delta e^{2t} \end{pmatrix}, \tag{97.1163}$$

qu'on injecte dans l'équation

$$\bar{x}'_P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \bar{x}_P + \begin{pmatrix} e^t \\ e^{2t} \end{pmatrix}. \tag{97.1164}$$

Après un certain nombre de calculs, nous trouvons

$$\begin{aligned} \alpha &= 0 \\ \beta &= -1/2 \\ \gamma &= 1 \\ \delta &= 1/2, \end{aligned} \tag{97.1165}$$

et donc la solution particulière est

$$\bar{x}_P = \begin{pmatrix} -\frac{1}{2}e^{2t} \\ e^t + \frac{1}{2}e^{2t}. \end{pmatrix}. \tag{97.1166}$$

Exercise 238 exo_II-2-02

Trouver, par la méthode de variation des constantes, les solutions réelles du système

$$\begin{cases} x' = ay - 2a & (97.1167a)\\ y' = -ax + a. & (97.1167b) \end{cases}$$

Montrer qu'il existe une façon plus facile d'obtenir la solution. corr_II-2-02

Correction of the exercise 238

La matrice du système homogène est $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$, et les valeurs propres sont $\lambda = \pm ia$. Les vecteurs propres correspondants sont

$$v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{97.1168}$$

Le système homogène a donc comme solutions

$$x_H = C_1 \begin{pmatrix} 1 \\ i \end{pmatrix} e^{iat} + C_2 \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{-iat}. \tag{97.1169}$$

Les solutions réelles sont

$$x_H = \begin{pmatrix} A \\ B \end{pmatrix} \cos(at) + \begin{pmatrix} B \\ -A \end{pmatrix} \sin(at). \tag{97.1170}$$

Le principe de la variation des constantes consiste à poser $A = A(t)$ et $B = B(t)$. La dérivation de x ne pose pas trop de problèmes :

$$x' = \begin{pmatrix} A' + aB \\ B' - aA \end{pmatrix} \cos(at) + \begin{pmatrix} B' - aA \\ -A' - aB \end{pmatrix} \sin(at), \tag{97.1171}$$

que l'on remet dans le système de départ. Ce que nous obtenons est

$$\begin{cases} A'\cos(at) + B'\sin(at) = -2a & (97.1172a)\\ B'\cos(at) - A'\sin(at) = a. & (97.1172b) \end{cases}$$

Afin de trouver les fonctions inconnues A et B, nous commençons par résoudre ce système algébriquement par rapport à A' et B'. Nous obtenons

$$\begin{cases} A'(t) = -a\sin(at) - 2a\cos(at) & (97.1173a)\\ B'(t) = -2a\sin(at) + a\cos(at). & (97.1173b) \end{cases}$$

Ces deux équations sont maintenant solubles par simple quadrature.

$$\begin{cases} A(t) = \cos(at) - 2\sin(at) + C_1 & (97.1174a)\\ B(t) = 2\cos(at) + \sin(at) + C_2. & (97.1174b) \end{cases}$$

La solution générale du système est donc

$$\begin{aligned} x &= \begin{pmatrix} \cos(at) - 2\sin(at) + C_1 \\ 2\cos(at) + \sin(at) + C_2 \end{pmatrix} \cos(at) + \begin{pmatrix} 2\cos(at) + \sin(at) + C_2 \\ -\cos(at) + 2\sin(at) - C_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 + C_2\sin(at) + C_1\cos(at) \\ 2 + C_2\cos'(at) - C_1\sin(at) \end{pmatrix}. \end{aligned} \tag{97.1175}$$

Ceci est pour la variation des constantes. Il y a cependant une façon plus facile de résoudre le système.

D'abord, le système se récrit

$$\begin{cases} x' = a(y-2) & (97.1176a)\\ y' = -a(x-1), & (97.1176b) \end{cases}$$

ce qui nous incite à faire le changement de variable $u = x - 1$ et $v = y - 2$. Le système devient

$$\begin{cases} u' = av & \text{(97.1177a)} \\ v' = -au. & \text{(97.1177b)} \end{cases}$$

En dérivant la première équation, $u'' = -a^2 u$, ce qui donne tout ce suite

$$u = C_1 \cos(at) + C_2 \sin(at). \tag{97.1178}$$

Ensuite v se trouve en faisant $v = u'/a$.

Exercise 239 exo_II-2-03 corr_II-2-03

Correction of the exercise 239

Exercise 240 exo_II-2-04 corr_II-2-04

Correction of the exercise 240

Exercise 241 exo_II-2-05 corr_II-2-05

Correction of the exercise 241

Exercise 242 exo_II-2-06 corr_II-2-06

Correction of the exercise 242

Exercise 243 exo_II-2-07 corr_II-2-07

Correction of the exercise 243

Chapitre 98

Exercices pour analyse CTU

98.1 TD 1 : Fonctions monotones, bijections

Exercise 244 exoanalyseCTU-0020

(1) Donner le tableau de variations de la fonction exponentielle.

(2) Expliciter les renseignements fournis par ce tableau de variations.

(3) Donner l'allure de la courbe représentative de la fonction exponentielle.

corranalyseCTU-0020

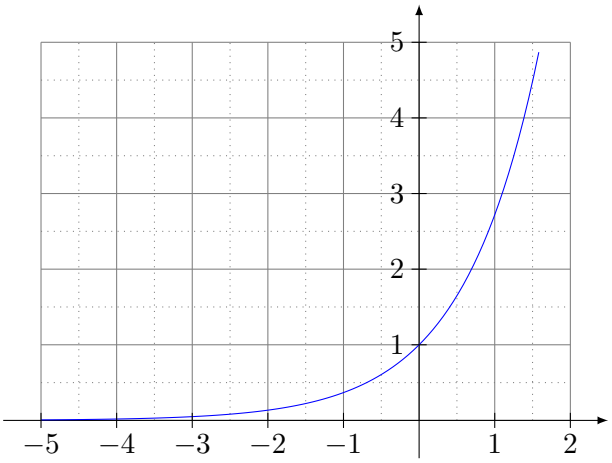
Correction of the exercise 244

Le tableau de variations de la fonction exponentielle :

x	$-\infty$		$+\infty$
			$+\infty$
e^x		$\nearrow$	
	0		

Le tableau indique les limites $\lim_{x\to-\infty} e^x = 0$, $\lim_{x\to\infty} e^x = \infty$. De plus la flèche indique que la fonction est strictement monotone croissante.

Le graphe de la fonction exponentielle (qui doit être absolument connu) est comme ceci :



Exercise 245 exostarterST-0002

Soit f la fonction définie sur $\mathbb{R}$ par : $f(x) = \begin{cases} x+1 & \text{si} \quad x < -1 \\ 2x & \text{si} \quad -1 \leqslant x \leqslant 1 \\ 4-x & \text{si} \quad x > 1 \end{cases}$

(1) La fonction f est-elle monotone sur $\mathbb{R}$?

(2) Tracer la représentation graphique de f.

(3) La fonction f est-elle continue sur $\mathbb{R}$?

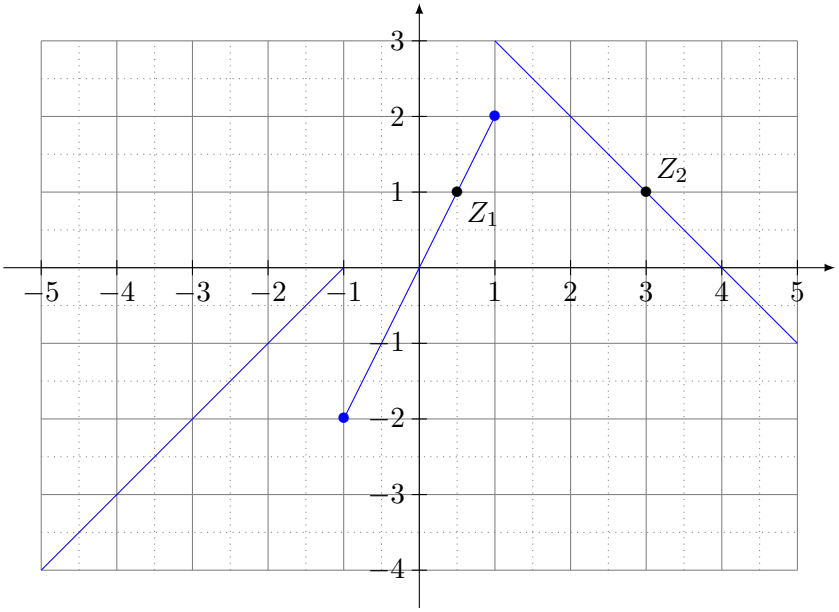
(4) Déterminer les antécédents éventuels de 1 par la fonction f.

corrstarterST-0002

Correction of the exercise 245

La fonction n'est pas monotone parce qu'elle est tantôt croissante (les parties $x+1$ et $2x$), tantôt décroissante (la partie $4-x$).

Le graphe de la fonction f est en trois parties :



Les points bleus représentent des points sur le graphe de la fonction. Le graphe se poursuit à l'infini tant à gauche qu'à droite.

Il est visible qu'elle n'est pas continue en $x=-1$ et $x=1$. La justification précise est que

$$\lim_{x\to -1} f(x) \tag{98.1}$$

n'existe pas parce que $\lim_{x\to -1^-} f(x) = \lim_{x\to -1} x+1 = 0$ alors que $\lim_{x\to -1^+} f(x) = \lim_{x\to -1} 2x = -2$.

La fonction f admet deux antécédents de -1 : $x=\frac{1}{2}$ et $x=3$. Ils sont visibles sur le graphe aux points Z_1 et Z_2. Formellement, il s'agit de résoudre l'équation

$$f(x) = 1. \tag{98.2}$$

Exercise 246 exoanalyseCTU-0022

On pourra répondre en proposant une représentation graphique possible des fonctions.

(1) On sait que la fonction f_1 est strictement croissante sur chacun des intervalles $[a,b]$ et $]b,c]$. Peut-on affirmer qu'elle est strictement croissante sur l'intervalle $[a,c]$?

(2) Déterminer une fonction f_2 et un intervalle I tels que la fonction f_2 ne soit pas bijective de I dans $f_2(I)$.

(3) Déterminer une fonction f_3 et un intervalle J tels que la fonction f_3 soit bijective de J dans $f_3(J)$, mais non monotone.

corranalyseCTU-0022

Correction of the exercise 246

(1) Non : la fonction peut « faire un saut » et redescendre juste en b. Par exemple la fonction

$$f(x) = \begin{cases} x & \text{si } -1 \leqslant x \leqslant 5 \\ x-8 & \text{si } x > 5 \end{cases} \tag{98.3}$$

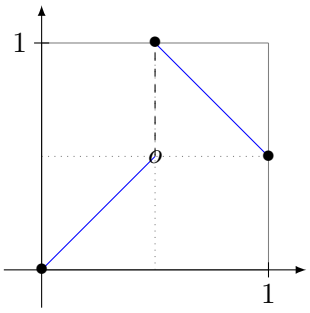
EqKZMooQxBmEN

est croissante sur $[-1,5]$ et sur $]5,10]$ mais $f(4)=4$ alors que $f(6)=-2$.

Il est recommandé de faire un dessin de la fonction (98.3).

(2) Une fonction f est toujours surjective vers l'intervalle $f(I)$. Nous devons donc chercher une fonction qui ne soit pas injective. La fonction sinus en est le parfait exemple. Nous pouvons donc proposer $f_2(x) = \sin(x)$ et $I = \mathbb{R}$.

(3) Voici le graphique d'une fonction bijective de $[0, 1]$ vers lui-même :



Le point blanc indique un point qui n'est pas sur le graphe. Ici $f(1/2) = 1$ et non $f(1/2) = 1/2$. Soyez capable d'en donner une expression analytique.

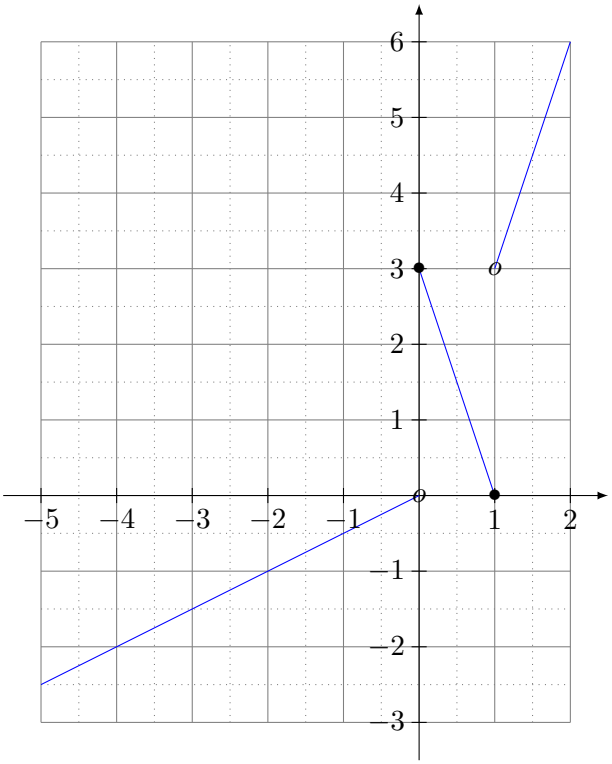
Exercise 247 exoautoanalyseCTU-1 Soit f la fonction définie sur $\mathbb{R}$ par : $f(x) = \begin{cases} \frac{x}{2} & \text{si} \quad x < 0 \\ 3(1-x) & \text{si} \quad 0 \leqslant x \leqslant 1 \\ 3x & \text{si} \quad x > 1 \end{cases}$

(1) Tracer la représentation graphique de f et donner son tableau de variations.

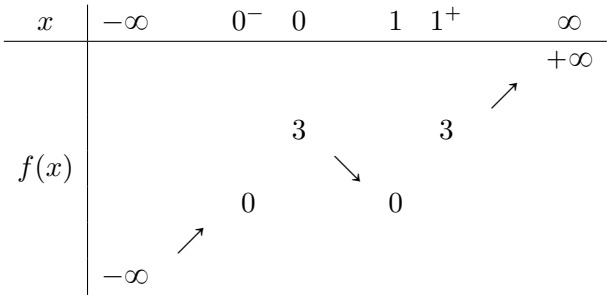
(2) On note y un nombre réel, déterminer les antécédents éventuels de y par la fonction f.

corrautoanalyseCTU-1

Correction of the exercise 247



Le tableau de variations est ceci :



les colonnes 0^- et 1^+ ne correspondent pas à des valeurs effectivement atteintes par la fonction, mais seulement les limites.

En ce qui concerne les antécédents, nous avons de la chance : la fonction est injective. Aucun nombre ne possède plusieurs antécédents. Il est donc possible d'écrire la fonction réciproque f^{-1} de la façon suivante :

$$f^{-1}(y) = \begin{cases} 2y & \text{si } y \in]-\infty, 0[\\ 1 - \frac{y}{3} & \text{si } y \in [0, 3] \\ +\frac{y}{3} & \text{si } y \in]3, \infty[. \end{cases} \tag{98.4}$$

Exercise 248 exoautoanalyseCTU-2

Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = x^3 + x + 1$.

(1) Étudier les variations de f.

(2) Justifier que f est une bijection de $\mathbb{R}$ dans un intervalle que l'on précisera.

(3) Soit g la bijection réciproque de f. Préciser l'ensemble de définition de g et calculer $g'(3)$.

corrautoanalyseCTU-2

Correction of the exercise 248

(1) La fonction dérivée de f est $f'(x) = 3x^2 + 1$. Il s'agit d'une fonction positive définie sur $\mathbb{R}$. La fonction f est donc strictement monotone croissante.

(2) Par le point précédent nous savons que f est strictement monotone sur son ensemble de définition, qui est $\mathbb{R}$ tout entier. La fonction f est donc une bijection de $\mathbb{R}$ dans son image. Pour trouver l'image de f nous nous rappelons que $\lim_{x\to\pm\infty} f(x) = \pm\infty$ et que f est continue (toute fonction polynomiale est continue). Le théorème des valeurs intermédiaires 9.93 nous dit alors que l'image de f est $\mathbb{R}$ tout entier.

(3) Il n'est pas aisé d'écrire g, la bijection réciproque de f, sous une forme analytique explicite. Par conséquent nous sommes obligés à répondre aux questions de ce point en utilisant les informations que nous avons sur f. D'abord, l'ensemble de définition de g est l'image de f, qui est $\mathbb{R}$. Ensuite, la dérivée de g en $x = 3$ peut être calculée à partir de la formule (12.462)

$$g'(3) = \frac{1}{f'(g(3))} = \frac{1}{3(g(3))^2 + 1}.$$

La valeur de $g(3)$ est la solution de $f(x) = 3$, c'est-à-dire $x^3 + x + 1 = 3$, on peut écrire $x(x^2 + 1) = 2$ et remarquer que $x = 1$ est une solution de cette équation. C'est d'ailleurs la seule solution, car f est une bijection. On a donc $g'(3) = 1/4$.

Exercise 249 exodevoir1ajout

Soit f la fonction définie sur $\mathbb{R}^*$ par $f(x) = -x^3 + \frac{1}{x} + 2$.

(1) Étudier les variations de f.

(2) Justifier que f est une bijection de $\mathbb{R}^{*,+}$ dans un intervalle que l'on précisera.

(3) Soit g la bijection réciproque de f. Préciser l'ensemble de définition de g et calculer $g'(2)$.

98.2 TD 2 : Dérivation, fonction composées

Exercise 250 exostarterST-0006

Soient $f(x) = \sqrt{x-1}$ et $g(x) = \dfrac{1}{\ln(x)}$.

(1) Trouvez les domaines de définition de f et de g.

(2) Écrivez les expressions explicites et trouvez les domaines de définition des fonctions composées $h_1(x) = f \circ g(x)$ et $h_2(x) = g \circ f(x)$.

(3) Trouvez l'expression des fonctions dérivées de h_1 et h_2, h_1' et h_2'.

corrstarterST-0006

Correction of the exercise 250

(1) On a $\text{Domaine}_f = \{x \in \mathbb{R} \text{ tels que } x - 1 \geqslant 0\} = [1, +\infty[$ et $\text{Domaine}_g = \{x \in \mathbb{R} \text{ tels que } x > 0, \text{ et } \ln(x) \neq 0\} =]0, 1[\cup]1, +\infty[$.

(2) Les expressions explicites de h_1 et h_2 sont

$$h_1(x) = f \circ g(x) = \sqrt{\frac{1}{\ln(x)} - 1},$$

et

$$h_2(x) = g \circ f(x) = \frac{1}{\ln\left(\sqrt{x-1}\right)}.$$

Leurs ensembles de définition sont

$$\begin{aligned}\text{Domaine}_{h_1} &= \left\{ x \in \mathbb{R} \text{ tels que } \ln(x) \neq 0 \text{ et } \frac{1}{\ln(x)} - 1 \geqslant 0 \right\} \\ &= \{x \in \mathbb{R} \text{ tels que } 0 < \ln(x) \leqslant 1\} =]1, e].\end{aligned}$$

et

$$\begin{aligned}\text{Domaine}_{h_2} &= \left\{ x \in \mathbb{R} \text{ tels que } \sqrt{x-1} > 0, \text{ et } \ln(\sqrt{x-1}) \neq 0 \right\} \\ &= \{x \in \mathbb{R} \text{ tels que } x > 1 \text{ et } x - 1 \neq 1\} =]1, 2[\cup]2, +\infty[.\end{aligned}$$

(3) Pour trouver l'expression des fonctions dérivées de h_1 et h_2, il faut utiliser les formules de la proposition 12.172.

$$h_1'(x) = -\frac{1}{2x \ln^2(x) \sqrt{\frac{1}{\ln(x)} - 1}},$$

$$h_2'(x) = -\frac{1}{2(x-1) \ln^2(\sqrt{x-1})}.$$

Exercise 251 exoDS2012-1-0001

La figure 98.1 répresente le graphe de la fonction $f(x) = \sin(x)$, pour $x \in [-\pi, \pi]$.

(1) Est-ce que la fonction $f_1(x) = \sin(x^2)$ est paire ? Impaire ? Périodique ? Si elle est périodique trouver sa période. Justifier au mieux vos réponses.

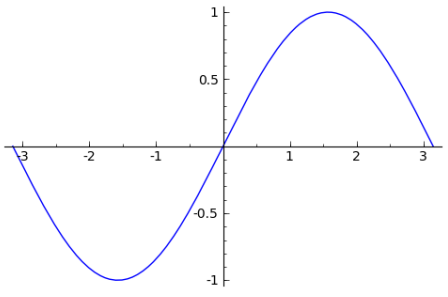
(2) Esquisser le graphe des fonctions suivantes

- $f_1(x) = \sin(2x)$;
- $f_2(x) = 2\sin(x)$;
- $f_3(x) = \sin^2(x)$;
- $f_4(x) = \sin(x-1)$;
- $f_5(x) = \sin(x) - 1$.

corrDS2012-1-0001

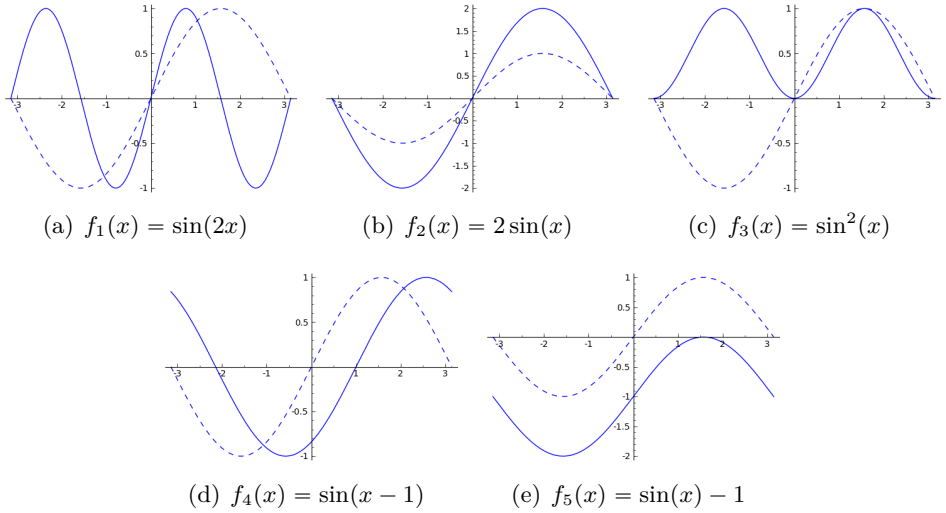
Correction of the exercise 251

FIGURE 98.1: Le graphe de la fonction sinus



graphesin

(2)



(a) $f_1(x) = \sin(2x)$ (b) $f_2(x) = 2\sin(x)$ (c) $f_3(x) = \sin^2(x)$

(d) $f_4(x) = \sin(x-1)$ (e) $f_5(x) = \sin(x) - 1$

FIGURE 98.2: Les graphes des fonctions $f_1, \ldots, f_5$

(1) La fonction $\sin(x^2)$ est paire, parce que $(-x)^2 = x^2$ et donc $\sin((-x)^2) = \sin(x^2)$.
Une fonction paire non nulle ne peut pas être impaire.
La fonction $\sin(x^2)$ n'est pas périodique : si elle l'était il y aurait $T > 0$ tel que $\sin((x+T)^2) = \sin(x^2)$ pour tout x dans $\mathbb{R}$. Cela veut dire que, pour tout x, $(x+T)^2 - x^2$ est un multiple entier de 2π, parce que la fonction sinus est périodique de période 2π. Mais alors $T = -x + \sqrt{2k\pi + x^2}$, ce qui est impossible parce que la période d'une fonction périodique ne peut pas dépendre de x.

Exercise 252 exoDS2012-2-0001

La figure 98.3 répresente le graphe de la fonction $f(x) = \cos(x)$, pour $x \in [-\pi, \pi]$.

(1) Est-ce que la fonction $f(x) = \cos(x) + x^2$ est paire ? Impaire ? Périodique ? Si elle est périodique, trouver sa période. Justifier au mieux vos réponses.

(2) Esquissez le graphe des fonctions suivantes pour $x \in [-\pi, \pi]$.

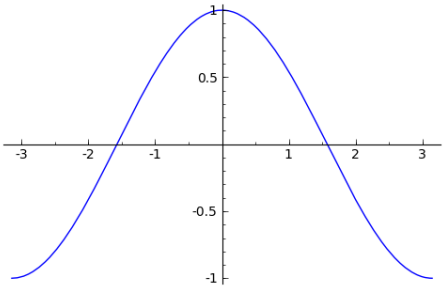
- $f_1(x) = \cos(2x)$;
- $f_2(x) = \frac{\cos(x)}{4}$;
- $f_3(x) = \cos(x) - 2$;
- $f_4(x) = |\cos(x)|$.

corrDS2012-2-0001

Correction of the exercise 252

(1) La fonction $\cos(x) + x^2$ est paire car elle est la somme de deux fonctions paires. La seule fonction qui est paire et impaire en même temps est la fonction nulle, donc $\cos(x) + x^2$ n'est pas impaire.
Cette fonction n'est pas périodique non plus : si elle l'était il y aurait $T > 0$ tel que $\cos(x+T) + (x+T)^2 = \cos(x) + x^2$ pour tout x dans $\mathbb{R}$. Cela voudrait dire que $\cos(T) + T^2 = 1$ (on

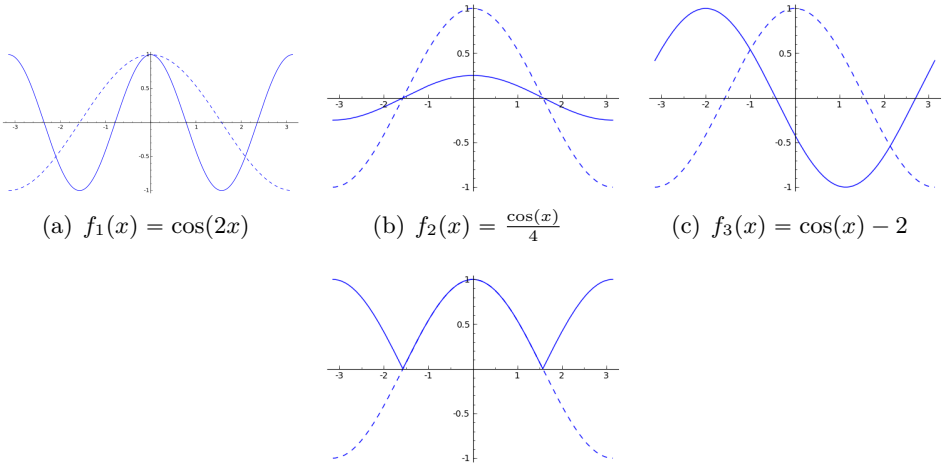
FIGURE 98.3: Le graphe de la fonction cosinus



graphecos

prend $x = 0$), $\cos(2T) + 4T^2 = \cos(T) + T^2$ (on prend $x = T$), et plus en générale, pour tout k entier positif $\cos(kT) + k^2T^2 = 1$. Cela n'est pas possible car $1 - \cos(kT)$ est borné par 2 alors que k^2T^2 peut devenir aussi grand qu'on veut quitte à prendre k assez grand.

(2)



(a) $f_1(x) = \cos(2x)$ (b) $f_2(x) = \frac{\cos(x)}{4}$ (c) $f_3(x) = \cos(x) - 2$

(d) $f_4(x) = |\cos(x)|$

FIGURE 98.4: Les graphes des fonctions $f_1, \ldots, f_4$

Exercise 253 exoanalyseCTU-0024

Soit f la fonction définie sur $\mathbb{R}$ par : $f(x) = \frac{2}{3}x^3 - \frac{1}{2}x^2 - 6x + \frac{58}{6}$.

(1) Montrer que l'équation $f(x) = 0$ admet une unique solution réelle.

(2) Donner un encadrement de cette solution par deux entiers consécutifs.

corranalyseCTU-0024

Correction of the exercise 253

(1) Pour répondre à cette question on commence par calculer la fonction dérivée de f, qui est $f'(x) = 2x^2 - x - 6$. L'étude du signe de f' ne présente aucune difficulté : c'est une fonction négative entre $x = -3/2$ et $x = 2$ et positive ailleurs. Du coup, la fonction f est strictement monotone sur les trois intervalles $I_1 =]-\infty, -3/2[$, $I_2 =]-3/2, 2[$ et $I_3 =]2, +\infty[$, croissante sur I_1 et I_3 et décroissante sur I_2. Il est facile de vérifier que $\lim_{x\to\pm\infty} f(x) = \pm\infty$, $f(-3/2) = 15,291\overline{6}$ et $f(2) = 1$. Puisque f est une fonction continue, le théorème des valeurs intermédiaires nous dit que la valeur 0 sera atteinte une seule fois, pour un x qui est dans l'intervalle I_1.

(2) Il faut calculer $f(-3)$ et $f(-4)$ et remarquer que $f(-3)$ est positive et $f(-4)$ négative. La valeur de x pour laquelle $f(x) = 0$ est donc comprise entre -4 et -3.

Exercise 254 exoautoanalyseCTU-52

Soit f la fonction définie par $f(x) = \dfrac{\ln(1+x) - x}{x^2}$.

(1) Déterminer l'ensemble de définition de f .

(2) Montrer que f peut-être prolongée par continuité en 0.

corrautoanalyseCTU-52

Correction of the exercise 254

(1) L'ensemble de définition de f est $\text{Domaine}_f = \{x \in \mathbb{R} \text{ tels que } x \neq 0 \text{ et } 1+x > 0\} =]-1,0[\cup]0,+\infty[$.

(2) Il faut calculer la limite de f lorsque $x \to 0$. Pour le faire nous avons besoin du théorème de de L'Hôpital

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{\ln(1+x)-x}{x^2} = \lim_{x\to 0} \frac{\frac{1}{1+x}-1}{2x} = \lim_{x\to 0} \frac{-x}{2x(1+x)} = -\frac{1}{2}.$$

La fonction qui prolonge f en $x=0$ est

$$\tilde{f}(x) = \begin{cases} f(x), & \text{si } x \neq 0; \\ -\frac{1}{2}, & \text{si } x = 0. \end{cases}$$

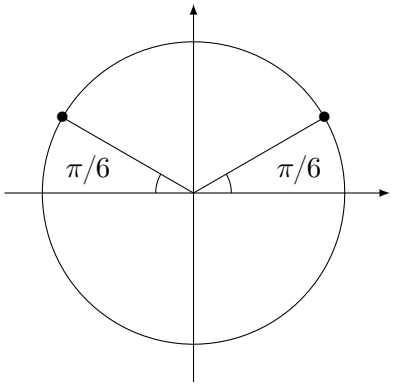
98.3 TD 3 : Rappels de trigonométrie

Exercise 255 exoautoanalyseCTU-3

Résoudre dans $\mathbb{R}$ les équations : **a/** $\sin(x) = \frac{1}{2}$ **b/** $\cos(x) = -\frac{1}{2}$.

corrautoanalyseCTU-3

Correction of the exercise 255



Les valeurs usuelles et les graphes sont autour de l'équation 1649.

Pour résoudre $\sin(x) = \frac{1}{2}$, la première chose est de se souvenir que $\sin(\pi/6) = \frac{1}{2}$. Il faut ensuite trouver les autres angles du cercle trigonométrique donnant le même sinus. Un dessin peut vraiment aider.

Nous avons donc les solutions $x = \pi/6$ et $x = 5\pi/6$ dans le cercle trigonométrique (c'est-à-dire entre 0 et 2π). L'ensemble des solutions dans $\mathbb{R}$ s'obtient en ajoutant les solutions obtenues par la périodicité de la fonction sinus :

$$x \in \left\{ \frac{\pi}{6} + 2k\pi \text{ tel que } k \in \mathbb{Z} \right\} \cup \left\{ \frac{5\pi}{6} + 2k\pi \text{ tel que } k \in \mathbb{Z} \right\}. \tag{98.5}$$

En ce qui concerne l'équation $\cos(x) = -\frac{1}{2}$, nous voyons dans les tables de valeurs usuelles que $\cos(\pi/3) = 1/2$. Donc $\cos(\frac{2\pi}{3}) = \cos(\frac{4\pi}{3}) = -\frac{1}{2}$. Les solutions entre 0 et 2π sont donc $x = 2\pi/3$ et $x = 4\pi/3$. Les solutions dans $\mathbb{R}$ sont au final

$$x \in \left\{ \frac{2\pi}{3} + 2k\pi \text{ tel que } k \in \mathbb{Z} \right\} \cup \left\{ \frac{4\pi}{3} + 2k\pi \text{ tel que } k \in \mathbb{Z} \right\}. \tag{98.6}$$

Exercise 256 exostarterST-0009

(1) (1a) Exprimer $\cos(2x)$ et $\sin(2x)$ en fonction de $\cos(x)$ et $\sin(x)$.

(1b) Exprimer $\tan(2x)$ en fonction de $\tan(x)$.

(1c) Déterminer les valeurs exactes de $\cos\left(\frac{\pi}{8}\right)$ et $\sin\left(\frac{\pi}{8}\right)$.

(2) (2a) Déterminer les valeurs exactes de $\cos\left(\frac{\pi}{12}\right)$, $\sin\left(\frac{7\pi}{12}\right)$ et $\tan\left(\frac{5\pi}{12}\right)$.

(2b) Déterminer un réel A et un réel φ tels que :

(2b.ii) $\frac{1}{2}\cos(x) - \frac{\sqrt{3}}{2}\sin(x) = \cos(x+\varphi)$;

(2b.ii) $\cos(x) + \sin(x) = A\cos(x+\varphi)$.

corrstarterST-0009

Correction of the exercise 256

(1) (1a) C'est un cas particulier d'utilisation des formules, qui devraient être bien connues, pour les valeurs de cosinus et de sinus d'une somme ou d'une différence d'angles.

À retenir 98.1

$$\cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta); \tag{98.7}$$

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta). \tag{98.8}$$

Nous obtenons

À retenir 98.2

$$\cos(2x) = \cos^2(x) - \sin^2(x); \tag{98.9}$$

$$\sin(2x) = 2\sin(x)\cos(x). \tag{98.10}$$

(1b) On sait que $\tan(x) = \frac{\sin(x)}{\cos(x)}$, par conséquent, les formules du point précédent nous donnent

$$\tan(2x) = \frac{\sin(2x)}{\cos(2x)} = \frac{2\sin(x)\cos(x)}{\cos^2(x) - \sin^2(x)} = \frac{2\cos^2(x)\tan(x)}{\cos^2(x)\left(1-\tan^2(x)\right)} = \frac{2\tan(x)}{1-\tan^2(x)}.$$

(1c) L'angle $\frac{\pi}{8}$ est la moitié de l'angle $\frac{\pi}{4}$, dont on connaît les valeurs de sinus et cosinus : $\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$. Pour trouver $\cos\left(\frac{\pi}{8}\right)$ et $\sin\left(\frac{\pi}{8}\right)$ nous écrivons alors un système de deux équations

$$\begin{cases} \frac{\sqrt{2}}{2} = \cos^2\left(\frac{\pi}{8}\right) - \sin^2\left(\frac{\pi}{8}\right), \\ \frac{\sqrt{2}}{2} = 2\cos\left(\frac{\pi}{8}\right)\sin\left(\frac{\pi}{8}\right). \end{cases}$$

Ensuite on utilise la relation $\cos^2(x) + \sin^2(x) = 1$, pour tout $x \in \mathbb{R}$, ce qui nous donne, dans la première équation $\cos\left(\frac{\pi}{8}\right) = \left(\frac{1}{2} + \frac{\sqrt{2}}{4}\right)^{1/2}$ et ensuite $\sin\left(\frac{\pi}{8}\right) = \left(\frac{1}{2} - \frac{\sqrt{2}}{4}\right)^{1/2}$.
Il est possible (mais demande un peu plus de travail) de ne pas utiliser du tout la relation $\cos^2(x) + \sin^2(x) = 1$ car le système nous donne déjà suffisamment d'information.

(2) (2a) Ici aussi, il suffit d'écrire les angles $\frac{\pi}{12}$, $\frac{7\pi}{12}$ et $\frac{5\pi}{12}$ comme sommes ou différences d'angles dont on connaît les valeurs de cos, sin et tan. Les formules vues dans les premiers points de cet exercice nous permettrons de conclure. Une possibilité est la suivante $\frac{\pi}{12} = \frac{\pi}{3} - \frac{\pi}{4}$, $\frac{5\pi}{12} = \frac{\pi}{4} + \frac{\pi}{6}$, $\frac{7\pi}{12} = \frac{\pi}{4} + \frac{\pi}{3}$.

(2b) (2b.ii) Dans la première équation il faut juste chercher un angle φ tel que $\cos(\varphi) = \frac{1}{2}$ et $\sin(\varphi) = \frac{\sqrt{3}}{2}$. La réponse est donc $\varphi = \frac{\pi}{6}$

(2b.ii) Dans cette deuxième équation il faut que $\cos(\varphi) = -\sin(\varphi) = \frac{1}{A}$. On trouve alors $A = \sqrt{2}$ et $\varphi = -\frac{\pi}{4}$.

98.4 TD 4 : Fonction trigonométriques réciproques

Exercise 257 exostarterST-0010-0011

(1) Résoudre les équations : **a/** $\arcsin(x) = \frac{\pi}{4}$ **b/** $\arcsin(x) = \frac{3\pi}{4}$

(2) $\sin\frac{3\pi}{4} = \frac{\sqrt{2}}{2}$. Que vaut $\arcsin(\frac{\sqrt{2}}{2})$? Peut-on comparer $\arcsin(\sin(x))$ et x ?

(3) Résoudre les équations : **a/** $\arccos(x) = \frac{\pi}{4}$ **b/** $\arccos(x) = \frac{3\pi}{4}$

(4) $\cos\frac{5\pi}{4} = -\frac{\sqrt{2}}{2}$. Que vaut $\arccos(-\frac{\sqrt{2}}{2})$? Peut-on comparer $\arccos(\cos(x))$ et x ?

corrstarterST-0010-0011

Correction of the exercise 257

(1) La fonction arcsin est une bijection définie sur l'intervalle $[-1, 1]$ qui prend ses valeurs dans l'intervalle $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$. Il existe donc une seule solution de la première équation, qui est $x = \sqrt{2}/2$. La deuxième équation, par contre, n'a pas de solution.

(2) On peut toujours comparer $\arcsin(\sin(x))$ et x, mais la relation $\arcsin(\sin(x)) = x$ sera satisfaite uniquement si x appartient à l'intervalle $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$. Dans les autres cas, on saura seulement que soit $x = \arcsin(\sin(x)) + 2K\pi$ ou $x = -\arcsin(\sin(x)) + (2K+1)\pi$, avec $K \in \mathbb{Z}$.

(3) La fonction arccos est une bijection définie sur l'intervalle $[-1, 1]$ qui prend ses valeurs dans l'intervalle $[0, \pi]$. Il existe donc une seule solution de la première équation, qui est $x = \sqrt{2}/2$ et une seule solution de la deuxième, qui est $x = -\sqrt{2}/2$.

(4) On peut toujours comparer $\arccos(\cos(x))$ et x, mais la relation $\arccos(\cos(x)) = x$ sera satisfaite uniquement si x appartient à l'intervalle $[0, \pi]$. Dans les autres cas, on saura seulement que soit $x = \arccos(\cos(x)) + 2K\pi$ ou $x = -\arccos(\cos(x)) + 2K\pi$, pour un $K \in \mathbb{Z}$ à déterminer.

Exercise 258 exostarterST-0012

(1) La fonction arcsin est-elle paire ou impaire ?

(2) La fonction arccos est-elle paire ou impaire ? Exprimer $\arccos(-x)$ en fonction de $\arccos(x)$.

(3) Montrer que pour tout réel x de [-1 ; 1] : $\arccos(x) + \arcsin(x) = \frac{\pi}{2}$

corrstarterST-0012

Correction of the exercise 258

(1) La fonction arcsin est impaire. Pour le voir il suffit d'utiliser la relation $\arcsin(\sin(x)) = x$, qui est satisfaite pour tout x dans l'intervalle $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$. On a donc que pour tout y dans le

domaine de arcsin, il existe un x dans $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$ tel que $y = \sin(x)$. On a alors $\arcsin(-y) = \arcsin(-\sin(x)) = \arcsin(\sin(-x)) = -x = -\arcsin(y)$, ce qui revient à dire que arcsin est impaire.

(2) La fonction arccos prend se valeurs sur un intervalle qui n'est pas symétrique par rapport à 0, par conséquent, cos n'est ni paire ni impaire. On peut tout de même essayer d'exprimer $\arccos(-y)$ en fonction de $\arccos(y)$, en se disant que si y est dans le domaine de arccos alors il existe un x entre 0 et π tel que $y = \cos(x)$. Alors $\arccos(-y) = \arccos(-\cos(x)) = \arccos(\cos(\pi - x)) = \pi - \arccos(y)$. Dans cette dernière suite d'égalités on a utilisé la formule pour le cosinus d'une différence d'angles.

(3) Pour tous les x dans l'intervalle $[-1, 1]$ on a, par définition, $\cos(\arccos(x)) = x$ et

$$\cos\left(-\arcsin(x) + \frac{\pi}{2}\right) = \cos(\arcsin(x))\cos\left(\frac{\pi}{2}\right) + \sin(\arcsin(x))\sin\left(\frac{\pi}{2}\right) \tag{98.11a}$$
$$= \sin(\arcsin(x)) \tag{98.11b}$$
$$= x. \tag{98.11c}$$

Donc $\arccos(x) = -\arcsin(x) + \frac{\pi}{2}$.

Exercise 259 exomazhe-0006

On considère les fonctions f_1 et f_2 définies par

$$\begin{aligned} f_1 &: x \mapsto \arctan\left(\frac{x}{\sqrt{1-x^2}}\right) \\ f_2 &: x \mapsto \arcsin(x). \end{aligned} \tag{98.12}$$

(1) Déterminer les ensembles de définitions de ces deux fonctions et calculer leurs dérivées.

(2) En déduire une relation entre f_1 et f_2.

corrmazhe-0006

Correction of the exercise 259

(1) L'ensemble de définition de f_2 a été donné dans le cours, il s'agit de l'intervalle $[-1, 1]$. Pour la fonction f_1 on commence par dire que la racine carrée est définie si $1 - x^2 \geqslant 0$ et que à son tour la fraction rationnelle est définie uniquement si $\sqrt{1-x^2} \neq 0$. La fonction arctan par contre est définie sur $\mathbb{R}$ tout entier. En somme nous avons que $\text{Domaine}_{f_2} =]-1, 1[$.
La dérivée de f_2 a été donné dans le cours, il s'agit de $x \mapsto \frac{1}{\sqrt{1-x^2}}$. Pour calculer la dérivée de f_1 il faut appliquer à plusieurs reprises la formule de dérivation des fonctions composées

$$\begin{aligned} f_1'(x) &= \frac{1}{1 + \left(\frac{x}{\sqrt{1-x^2}}\right)^2} \cdot \frac{\sqrt{1-x^2} + \frac{x^2}{\sqrt{1-x^2}}}{1-x^2} = \\ &= \frac{1-x^2}{(1-x^2)+x^2} \cdot \frac{\frac{(1-x^2)+x^2}{\sqrt{1-x^2}}}{1-x^2} = \frac{1}{\sqrt{1-x^2}}. \end{aligned}$$

Les deux fonctions ont donc la même dérivée. Il faut aussi remarquer que les ensembles de définitions de dérivées sont égaux.

(2) On peut en conclure que la fonction différence entre f_1 et f_2, définie sur la partie commune du domaine $]-1, 1[$, est une constante (c'est-à-dire, une fonction qui a dérivée nulle). Laquelle ? Il suffit de calculer $f_1(x) - f_2(x)$ pour une valeur de x : si on prend par exemple $x = 0$ on a $f_1(0) - f_2(0) = 0$. Cela nous dit que f_1 et f_2 sont égales sur l'intervalle $]-1, 1[$. On ne peut pas les comparer aux points $x = \pm 1$, car f_1 n'est pas définie à ces points.

Exercise 260 exomazhe-0007

On considère maintenant la fonction

$$g(x) = \arcsin \frac{x}{\sqrt{1+x^2}}. \tag{98.13}$$

Simplifier l'expression de g.

Conseil : utiliser les résultats de l'exercice 259. corrmazhe-0007

Correction of the exercise 260

L'ensemble de définition de g est $\mathbb{R}$ tout entier. Cela arrive parce que $1+x^2>0$ pour tout x, $\sqrt{1+x^2}>1$ pour tout x et

$$\left|\frac{x}{\sqrt{1+x^2}}\right| \leqslant 1, \qquad \text{pour tout } x\in\mathbb{R}.$$

Cette dernière inégalité est facile à démontrer si on utilise le fait que $|x|=\sqrt{(x)^2}$ pour tout $x\in\mathbb{R}$.

La dérivée de g est

$$\begin{aligned} g'(x) &= \frac{1}{\sqrt{1-\left(\frac{x}{\sqrt{1+x^2}}\right)^2}} \cdot \frac{\sqrt{1+x^2}-\frac{x^2}{\sqrt{1+x^2}}}{1+x^2} = \\ &= \frac{\sqrt{1+x^2}}{\sqrt{(1+x^2)-x^2}} \cdot \frac{\frac{(1+x^2)-x^2}{\sqrt{1+x^2}}}{1+x^2} = \frac{1}{1+x^2}. \end{aligned}$$

Les fonctions g et arctan ont donc la même dérivée et le même ensemble de définition. On peut en conclure que la fonction différence entre g et arctan est une constante. On trouve $g(0)-\arctan(0)=0$. Les deux fonctions sont donc égales.

Exercise 261 exoautoanalyseCTU-10

On considère la fonction f définie par $f(x)=\tan(2\arctan(x))$.

(1) Déterminer l'ensemble de définition de f.

(2) Montrer que pour tout x de l'ensemble de définition, $f(x)=\dfrac{2x}{1-x^2}$.

corrautoanalyseCTU-10

Correction of the exercise 261

(1) On sait que $\arctan(x)\in\left]-\frac{\pi}{2},\frac{\pi}{2}\right[$ et par conséquent $2\arctan(x)\in\left]-\pi,\pi\right[$. Or, la fonction tan n'est pas définie en $\pm\frac{\pi}{2}$, donc il faut éviter les valeurs de x pour lesquels $2\arctan(x)=\pm\frac{\pi}{2}$, c'est-à-dire qu'il faut imposer la condition $x\neq\pm1$. Le domaine de la fonction f est donc $\mathbb{R}\backslash\{\pm1\}$.

(2) Pour plus de lisibilité on appelle $g(x)$ la fonction $\frac{2x}{1-x^2}$.

La façon classique de résoudre ce type d'exercice comporte deux pas :

— remarquer que $g(0)=0$ et $f(0)=0$ c'est-à-dire que les graphes des fonctions f et g se croisent en correspondance à $x=0$;

— calculer les dérivées de f et de g, montrer qu'elles sont égales et conclure que la fonction différence $f-g$ est constante (et donc il s'agit de la fonction nulle, par le point précédent).

Dans cet exercice en particulier, cette méthode n'est pas très performante, parce que l'expression de la dérivée de f est compliquée.

Nous allons donc utiliser la définition de la fonction tangente, les formules pour calculer sinus et cosinus de $2x$ et la définition de arctan.

$$\begin{aligned} f(x) &= \tan(2\arctan(x)) = \frac{\sin(2\arctan(x))}{\cos(2\arctan(x))} \\ &= \frac{2\sin(\arctan(x))\cos(\arctan(x))}{\cos^2(\arctan(x)) - \sin^2(\arctan(x))} \\ &= \frac{2\frac{\sin(\arctan(x))}{\cos(\arctan(x))}}{1 - \frac{\sin^2(\arctan(x))}{\cos^2(\arctan(x))}} = \frac{2\tan(\arctan(x))}{1 - \tan^2(\arctan(x))} \\ &= \frac{2x}{1-x^2} = g(x), \end{aligned}$$

pour tout x dans l'ensemble de définition de f.

98.5 TD 5 : Trigonométrie hyperbolique

Exercise 262 exostarterST-0015

Remarque 98.3.
Les points 4, 5 et 6 de cet exercice sont donnés en devoir.

On appelle *sinus hyperbolique* (sinh) et *cosinus hyperbolique* (cosh) les fonctions définies sur $\mathbb{R}$ par

$$\sinh(x) = \frac{e^x - e^{-x}}{2}; \qquad \cosh(x) = \frac{e^x + e^{-x}}{2}. \tag{98.14}$$ defcoshetsinh

(1) Trouver les domaines de définition de sinh et cosh, étudier leur parité.

(2) Ces fonctions sont continues et dérivables sur tout leur domaine. Trouver $\sinh'(x)$ et $\cosh'(x)$.

(3) Montrer que pour tout $x \in \mathbb{R}$ on a $\cosh^2(x) - \sinh^2(x) = 1$. Que peut-on en déduire sur l'image de cosh ?

(4) Démontrer les formules suivantes :

(4a) $\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y)$;

(4b) $\cosh(x+y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y)$.

(5) Donner des expressions de $\cosh(2x)$ et $\sinh(2x)$ en fonction de $\cosh(x)$ et $\sinh(x)$.

(6) Simplifier l'expression $f(x) = \cosh\left(\ln(x + \sqrt{x^2-1})\right)$ en utilisant la définition (98.14).

corrstarterST-0015

Correction of the exercise 262

(1) Les fonctions sinh et cosh sont définies sur $\mathbb{R}$ tout entier. La fonction sinh est impaire, car

$$\sinh(-x) = \frac{e^{-x} - e^{-(-x)}}{2} = \frac{e^{-x} - e^x}{2} = -\sinh(x).$$

En procédant de la même manière on trouve que cosh est paire.

(2)

$$\sinh'(x) = \frac{e^x - (-e^{-x})}{2} = \frac{e^x + e^{-x}}{2} = \cosh(x),$$

$$\cosh'(x) = \frac{e^x + (-e^{-x})}{2} = \frac{e^x - e^{-x}}{2} = \sinh(x).$$

(3) On calcule directement en utilisant les formules (98.14). Pour tout $x \in \mathbb{R}$

$$\cosh^2(x) - \sinh^2(x) = \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \tag{98.15a}$$
$$= \frac{1}{4}\left[(e^{2x} + e^{-2x} + 2) - (e^{2x} + e^{-2x} - 2)\right] \tag{98.15b}$$
$$= 1. \tag{98.15c}$$

par la définition on savait déjà que la fonction cosh est strictement positive, cette nouvelle relation nous dit que $\cosh(x) > 1$ pour tout $x \in \mathbb{R}$.

(4) Nous allons simplement utiliser les définitions de sinh et cosh.

$$\sinh(x+y) = \frac{e^{x+y} - e^{-(x+y)}}{2};$$
$$\sinh(x)\cosh(y) + \cosh(x)\sinh(y) = \frac{e^x - e^{-x}}{2}\frac{e^y + e^{-y}}{2} \frac{e^x + e^{-x}}{2}\frac{e^y - e^{-y}}{2}$$
$$= \frac{e^{x+y} + e^{x-y} - e^{-x+y} - e^{-(x+y)}}{4} + \frac{e^{x+y} - e^{x-y} + e^{-x+y} - e^{-(x+y)}}{4}$$
$$= \frac{e^{x+y} - e^{-(x+y)}}{2}.$$

L'autre égalité peut se montrer de façon analogue.

(5) Par le point précédent, en prenant $y = x$: $\cosh(2x) = \cosh^2(x) + \sinh^2(x)$ et $\sinh(2x) = 2\cosh(x)\sinh(x)$.

(6)
$$f(x) = \cosh\left(\ln(x + \sqrt{x^2-1})\right)$$
$$= \frac{e^{\ln(x+\sqrt{x^2-1})} + e^{-\ln(x+\sqrt{x^2-1})}}{2} = \frac{1}{2}\left(x + \sqrt{x^2-1} + \frac{1}{x+\sqrt{x^2-1}}\right) = x$$

98.6 TD 6 : Calcul intégral, introduction

Exercise 263 exomazhe-0008

Calculer l'aire de la surface limitée par les droites d'équations $y = 0$, $x = 0$, $x = 2$ et la parabole d'équation $y = x^2 - x + 1$. Faire un dessin. corrmazhe-0008

Correction of the exercise 263

La fonction $f(x) = x^2 - x + 1$ ne prends que de valeurs positives quand x est dans l'intervalle $[0, 2]$. En effet, la dérivée $f'(x) = 2x - 1$ ne vaut zéro que en $x = 1/2$, est négative sur $[0, 1/2[$ et positive sur $]1/2, 2]$. Cela veut dire que $f(1/2) = 1/4 - 1/2 + 1 = 3/4$ est le minimum de f sur $[0, 2]$.

La valeur de l'aire à calculer est donc l'intégrale de f sur l'intervalle $[0, 2]$

$$\int_0^2 x^2 - x + 1\, dx = \left[\frac{1}{3}x^3 - \frac{1}{2}x^2 + x\right]_0^2 = \frac{8}{3} - \frac{4}{2} + 2 = \frac{8}{3}.$$

Exercise 264 exostarterST-0017

(1) Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = x + 2$.
  (1a) Déterminer la primitive F_1 de f qui s'annule en 2.
  (1b) Tracer les représentations graphiques de f et de F_1.

(2) (2a) Déterminer l'ensemble des primitives de la fonction tan sur l'intervalle $\left]-\frac{\pi}{2}\,;\,\frac{\pi}{2}\right[$.
  (2b) Déterminer la primitive sur l'intervalle $\left]-\frac{\pi}{2}\,;\,\frac{\pi}{2}\right[$ de la fonction tan qui s'annule en $\frac{\pi}{3}$.

corrstarterST-0017

Correction of the exercise 264

(1) L'ensemble des primitives de f est

$$\int x+2\,dx = \frac{1}{2}x^2+2x+C, \qquad C\in\mathbb{R}.$$

Pour répondre à la question il faut trouver une valeur de C telle que si $F_1(x)=\frac{1}{2}x^2+2x+C$ alors $F_1(2)=0$, c'est-à-dire que C est la solution de $6+C=0$. On a alors $C=-6$ et $F_1(x)=\frac{1}{2}x^2+2x-6$.

> **À retenir 98.4**
> Dans cette deuxième partie de l'exercice nous sommes obligés à préciser dans quel intervalle sera définie la primitive que nous intéresse. Cela est du au fait que l'ensemble de définition de la fonction tan consiste en une réunion d'intervalles disjoints. **Cela ne veut pas dire qu'on calcul l'intégrale de** tan **sur l'intervalle** $\left]-\frac{\pi}{2}\,;\,\frac{\pi}{2}\right[$ **!!**

(2) L'ensemble des primitives que nous intéresse est donc

$$\int \tan(x)\,dx = \int \frac{\sin(x)}{\cos(x)}\,dx = -\ln\left(|\cos(x)|\right)+C = \ln\left(\frac{1}{\cos(x)}\right)+C,$$

pour $C\in\mathbb{R}$, et x dans l'intervalle $\left]-\frac{\pi}{2}\,;\,\frac{\pi}{2}\right[$.
Le choix de l'intervalle nous à permis d'omettre la valeur absolue, car la fonction cos est toujours positive sur $\left]-\frac{\pi}{2}\,;\,\frac{\pi}{2}\right[$.
Pour déterminer la primitive qui s'annule en $\frac{\pi}{3}$ il faut résoudre l'équation $\ln\left(\frac{1}{\cos(\pi/3)}\right)+C=0$. On trouve $C=\ln(2)$ et donc la primitive cherchée est $F(x)=\ln\left(\frac{1}{\cos(x)}\right)+\ln(2)=\ln\left(\frac{2}{\cos(x)}\right)$.

Exercise 265 exoanalyseCTU-0002

La proposition 15.93 nous a montré que le logarithme était une primitive de la fonction inverse

$$\begin{aligned} i\colon]0,\infty[&\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned} \tag{98.16}$$

Trouver une primitive de la fonction

$$\begin{aligned} i\colon \mathbb{R}\backslash\{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x}. \end{aligned} \tag{98.17}$$

En d'autres termes, nous demandons une primitive de la fonction inverse sur les négatifs. exoanalyseCTU-0002

Correction of the exercise 265

Le fait que $x\mapsto\ln(x)$ soit une primitive de $\frac{1}{x}$ se traduit par le fait que

$$\ln(x)=\int_1^x \frac{1}{t}dt. \tag{98.18}$$

La primitive de la fonction $x\mapsto 1\frac{1}{x}$ pour les x négatifs sera la fonction de x donnée par

$$\int_{-1}^{x}\frac{1}{t}dt=\int_1^{-x}\frac{1}{-u}(-du)=\int_1^{-x}\frac{1}{u}du=\ln(-x). \tag{98.19}$$

Dans ce calcul x est une constante négative.

Donc pour les x négatifs, la primitive choisie de $\frac{1}{x}$ est la fonction $x \mapsto \ln(-x)$.

Pour toutes les valeurs de x nous avons que la primitive de $\frac{1}{x}$ est $\ln(|x|)$.

Exercise 266 exomazhe-0010

Calculer les intégrales suivantes

(1) $\int_0^1 x^5 + 3x^2 + 3\,dx$;

(2) $\int_1^2 x^{1/3} + 4x^{1/2}\,dx$;

(3) $\int_{-\pi}^0 \cos(x)\,dx$;

(4) $\int_{-2}^1 e^x\,dx$;

(5) $\int_0^1 \frac{1}{\sqrt{x}}\,dx$.

corrmazhe-0010

Correction of the exercise 266

Il suffit d'appliquer la formule (14.911) et d'utiliser la linéarité de l'intégrale et le tableau des primitives des fonctions fondamentales.

(1) $\int_0^1 x^5 + 3x^2 + 3\,dx = \left[\frac{1}{6}x^6 + x^3 + 3x\right]_0^1 = \frac{1}{6} + 4 = \frac{25}{6}$;

(2) $\int_1^2 x^{1/3} + 4x^{1/2}\,dx = \left[\frac{3}{4}x^{\frac{4}{3}} + 4\frac{2}{3}x^{\frac{3}{2}}\right]_1^2 = \frac{3}{2}2^{\frac{1}{3}} + \frac{16}{3}2^{\frac{1}{2}} - \frac{41}{12}$;

(3) $\int_{-\pi}^0 \cos(x)\,dx = [\sin(x)]_{-\pi}^0 = 0$;

(4) $\int_{-2}^1 e^x\,dx = [e^x]_{-2}^1 = e - e^{-2}$;

(5) $\int_0^1 \frac{1}{\sqrt{x}}\,dx = \int_0^1 x^{-1/2}\,dx = \left[2x^{\frac{1}{2}}\right]_0^1 = 2.$

Exercise 267 exodevoir2ajout

Calculer l'aire de la region du plan comprise entre les graphes de $f(x) = x^2 - 2x$ et $g(x) = -x^2 + x$.

98.7 TD 7 : Recherche de primitives

Exercise 268 exostarterST-0018

Déterminer les ensembles de primitives suivants :

(1) $F_1(x) = \int x(x^2+3)\,\mathrm{d}x$

(2) $F_2(x) = \int x\sqrt{1+x^2}\,\mathrm{d}x$

(3) $F_3(x) = \int \frac{x^2}{1+x^3}\,\mathrm{d}x$

(4) $F_4(x) = \int xe^x\,\mathrm{d}x$

(5) $F_5(x) = \int \arcsin(x)\,\mathrm{d}x$

(6) $F_6(x) = \int x^2e^x\,\mathrm{d}x$

(7) $F_7(x) = \int \ln(x)\,\mathrm{d}x$

(8) $F_8(x) = \int \arctan(x)\,\mathrm{d}x$

(9) $F_9(x) = \int x\sin(x)\,\mathrm{d}x$

(10) $F_{10}(x) = \int \sin^2(x)\,\mathrm{d}x$

(11) $F_{11}(x) = \int \frac{\ln(x)}{x}\,\mathrm{d}x$

(12) $F_{12}(x) = \int \frac{x}{1+x^4}\,\mathrm{d}x$

(13) $F_{13}(x) = \int \frac{dx}{x^2+4x+5}$

(14) $F_{14}(x) = \int \frac{8x\,dx}{x^2+4x+5}$

(15) $F_{15}(x) = \int \frac{dx}{x\sqrt{1-\ln(x)}}$ (poser $t = 1-\ln(x)$).

corrstarterST-0018

Correction of the exercise 268

(1) Par parties : $F_1(x) = \int x(x^2+3)\,dx = x\left(\frac{x^3}{3} + 3x\right) - \int \frac{x^3}{3} + 3x\,dx = \frac{x^4}{4} + \frac{3x^2}{2} + C$. À l'aide du changement de variable $u = x^2 + 3$: on a $du = 2xdx$ et donc $F_1(x) = \int x(x^2+3)\,dx =$

$\frac{1}{2}\int u\,du = \frac{u^2}{4} + C = \frac{1}{4}\left(x^4 + 6x^2 + 9\right) + C = \frac{x^4}{4} + \frac{3x^2}{2} + C$. Bien entendu, la constante C a absorbé le $\frac{9}{4}$ dans le dernier passage.

(2) À l'aide du changement de variable $u = x^2 + 1$: on a $du = 2xdx$ et donc $F_2(x) = \int x\sqrt{1+x^2}\,dx = \frac{1}{2}\int \sqrt{u}\,du = \frac{1}{3}u^{3/2} + C = \frac{1}{3}(x^2+1)^{3/2} + C$.

(3) À l'aide du changement de variable $u = x^3 + 1$: on a $du = 3x^2dx$ et donc $F_3(x) = \int \frac{x^2}{1+x^3}\,dx = \frac{1}{3}\int \frac{1}{u}\,du = \frac{1}{3}\ln(|u|) + C = \frac{1}{3}\ln(|x^3+1|) + C$.

(4) Par parties : $F_4(x) = \int xe^x\,dx = xe^x - \int e^x\,dx = (x-1)e^x + C$.

(5) Il faudra utiliser ici l'intégration par parties (comme dans 20.137) et ensuite un changement de variable. On commence par calculer $F_5(x) = \int 1 \times \arcsin(x)\,dx = x\arcsin(x) - \int \frac{x}{\sqrt{1-x^2}}\,dx$. Ensuite on pose $u = 1 - x^2$ et on obtient $F_5(x) = x\arcsin(x) + \frac{1}{2}\int \frac{1}{\sqrt{u}}\,du = x\arcsin(x) + \sqrt{1-x^2} + C$.

(6) Il faudra utiliser ici l'intégration par parties deux fois de suite. $F_6(x) = \int x^2e^x\,dx = x^2e^x - 2\int xe^x\,dx = x^2e^x - 2\left(xe^x - \int e^x\,dx\right) = e^x\left(x^2 - 2x + 2\right) + C$.

(7) Cette primitive a été calculée dans l'exemple 20.137.

(8) Par parties : $F_8(x) = \int \arctan(x)\,\mathrm{d}x = x\arctan(x) - \int \frac{x}{x^2+1}\,dx = x\arctan(x) - \frac{1}{2}\ln(x^2+1) + C$.

(9) Par parties : $F_9(x) = \int x\sin(x)\,dx = -x\cos(x) + \int \cos(x)\,dx = -x\cos(x) + \sin(x) + C$.

(10) Le calcul se fait, comme dans l'exemple 20.144, en s'appuyant sur des formules de trigonométrie.
Première possibilité : $F_{10}(x) = \int \sin^2(x)\,dx = \int \frac{1-\cos(2x)}{2}\,dx = \frac{1}{2}x - \frac{\sin(2x)}{4} + C$.
Deuxième possibilité (avec une intégration par parties d'abord) : on écrit

$$F_{10}(x) = \int \sin^2(x)\,dx = -\cos(x)\sin(x) + \int \cos^2(x)\,dx \tag{98.20}$$

et ensuite on remarque que en intégrant une deuxième fois par parties on n'avance pas du tout. On utilise alors la formule $1 = \cos^2(x) + \sin^2(x)$ et on a $F_{10}(x) = -\cos(x)\sin(x) + \int 1 - \sin^2(x)\,dx$. On a alors que $\int \sin^2(x)\,dx = \frac{1}{2}\left(-\cos(x)\sin(x) + x\right) + C$. Le résultat est le même qu'on a obtenu ci-dessus, parce que $\sin(2x) = 2\cos(x)\sin(x)$.

(11) Par parties on a

$$F_{11}(x) = \int \frac{\ln(x)}{x}\,dx = \ln^2(x) - \int \frac{\ln(x)}{x}\,dx,$$

d'où on déduit que $F_{11}(x) = \frac{1}{2}\ln^2(x) + C$.

(12) À l'aide du changement de variable $u = x^2$: on a $du = 2xdx$ et donc $F_{12}(x) = \int \frac{x}{1+x^4}\,dx = \frac{1}{2}\int \frac{1}{1+u^2}\,du = \frac{1}{2}\arctan(u) + C = \frac{1}{2}\arctan(x^2) + C$.

(13) On observe que $F_{13}(x) = \int \frac{1}{x^2+4x+5}\,dx = \int \frac{1}{(x+2)^2+1}\,dx$. Le changement de variable $t = x + 2$ nous donne $F_{13}(x) = \int \frac{1}{t^2+1}\,dt = \arctan(t) + C = \arctan(x+2) + C$.

(14) Pour calculer cette primitive on voudrait écrire convenablement le polynôme au dénominateur de $\dfrac{8x}{x^2+4x+5}$. On peut d'abord essayer d'en chercher les racines, mais une fois vu qu'elle ne sont pas réelles nous devons abandonner l'espoir de décomposer la fonction rationnelle. Ici nous pouvons observer que le numérateur est de degré 1 et le dénominateur de degré 2. Il faut donc chercher à trouver une fonction à intégrer (par changement de variable) de la forme u'/u. La dérivée de x^2+4x+5 et $2x+4$ donc $F_{14}(x) = 4\displaystyle\int \frac{2x+4}{x^2+4x+5} - \frac{4}{x^2+4x+5}\,dx = 4\ln\left(|x^2+4x+5|\right) - 16\int \frac{1}{x^2+4x+5}\,dx$. L'intégrale qui reste n'est pas encore banale. L'astuce standard dans ce cas consiste à écrire x^2+4x+5 dans la forme $1+$(un carré) et ensuite utiliser un changement de variable pour se ramener à la dérivée de la fonction arctan. En somme $\displaystyle\int \frac{1}{x^2+4x+5}\,dx = \int \frac{1}{1+(x+2)^2}\,dx$ et nous pouvons le calculer facilement à l'aide du changement de variable $u = x+2$: $\displaystyle\int \frac{1}{x^2+4x+5}\,dx = \int \frac{1}{1+u^2}\,du = \arctan(u)+C = \arctan(x+2)+C$. Finalement, $F_{14}(x) = 4\ln\left(|x^2+4x+5|\right) - 16\arctan(x+2)+C$.

(15) Le changement de variable conseillé nous donne $F_{15}(x) = \displaystyle\int \frac{1}{e^{1-t}\sqrt{t}}(-e^{1-t})dt = -\int \frac{1}{\sqrt{t}}\,dt = -2\sqrt{t}+C = -2\sqrt{1-\ln(x)}+C$.

98.8 TD 8 : Intégrale définie

Exercise 269 exomazhe-0011

Calculer les intégrales suivantes

(1) $I_1 = \displaystyle\int_1^{\pi} \frac{4x^3}{x^4+5}\,dx$;

(2) $I_2 = \displaystyle\int_0^{1} \frac{1}{(4x+1)^4}\,dx$;

(3) $I_3 = \displaystyle\int_0^{\sqrt{5}} \frac{1}{x^2+5}\,dx$;

(4) $I_4 = \displaystyle\int_{-\pi/2}^{0} \sin^2(x)\cos(x)\,dx$. Prendre $u = \sin(x)$;

(5) $I_5 = \displaystyle\int_0^{1} \frac{e^x}{e^{2x}+4}\,dx$. Prendre $u = e^x$;

(6) $I_6 = \displaystyle\int_0^{1} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,dx$. Prendre $x = \tan u$;

(7) $I_7 = \displaystyle\int_0^{a} \sqrt{a^2-x^2}\,dx$, $a \in [0\,;+\infty[$. Prendre $x = a\sin(u)$.

(8) $I_8 = \displaystyle\int_{-\pi}^{0} e^x \sin(x)\,dx$;

(9) $I_9 = \displaystyle\int_1^{2} x^2\ln(2x)\,dx$.

(10) $I_{10} = \displaystyle\int_1^{2} \frac{x^4+1}{x^5+x^3}\,dx$.

(11) $I_{11} = \displaystyle\int_{-3}^{0} \frac{1}{x^2-3x+2}dx$. Conseil : écrire la fraction sous la forme $\frac{a}{x-1}+\frac{b}{x-2}$.

(12) $I_{12} = \displaystyle\int_1^{2} \ln^2(x)dx$.

(13) $I_{13} = \displaystyle\int_0^{1} \frac{e^x-1}{e^x+1}dx$.

(14) $I_{14} = \displaystyle\int_1^{2} \frac{e^{-2x}}{(1+2e^{-x})^2}dx$.

(15) $I_{15} = \displaystyle\int_0^{1} \frac{x^3}{x^2+1}dx$. Conseil : poser $t = x^2+1$.

corrmazhe-0011

Correction of the exercise 269

(1) $I_1 = \displaystyle\int_1^{\pi} \frac{4x^3}{x^4+5}\,dx = \Big[\ln(x^4+5)\Big]_1^{\pi} = \ln\left(\frac{\pi^4+5}{6}\right)$.

(2) On utilise le changement de variable $u = 4x+1$. On a alors $du = 4dx$, les bornes du domaine

d'intégration deviennent $u(0) = 1$ et $u(1) = 5$ et notre intégrale s'écrit de la forme suivante

$$I_2 = \int_0^1 \frac{1}{(4x+1)^4}\,dx = \int_1^5 \frac{1}{4u^4}\,du = \left[-\frac{1}{12u^3}\right]_1^5 = -\frac{1}{12\times 125} + \frac{1}{12} = \frac{31}{375}.$$

(3) Ici on vise à écrire la fonction à intégrer comme $\frac{1}{u^2+1}$, pour avoir $\arctan(u) + C$ comme primitive. Pour ainsi faire, nous devons utiliser le changement de variable $u = \frac{x}{\sqrt{5}}$. On a alors $du = \frac{1}{\sqrt{5}}dx$, les bornes du domaine d'intégration deviennent $u(0) = 0$ et $u(\sqrt{5}) = 1$

$$I_3 = \int_0^{\sqrt{5}} \frac{1}{x^2+5}\,dx = \int_0^1 \frac{\sqrt{5}}{5(u^2+1)}\,du = \frac{1}{\sqrt{5}}[\arctan(u)]_0^1 = \frac{\pi}{4\sqrt{5}}.$$

(2) On utilise le changement de variable $u = \sin(x)$, qui comporte $du = \cos(x)\,dx$, $u(-\pi/2) = -1$ et $u(0) = 0$. Nous avons alors que $I_4 = \displaystyle\int_{-1}^0 u^2\,du = \left[\frac{u^3}{3}\right]_{-1}^0 = \frac{1}{3}$.

(5) Le bon changement de variable nous est donné par l'énoncé. On a $du = e^x dx$, les bornes d'intégration deviennent $u(0) = e^0 = 1$ et $u(1) = e$ et notre intégrale sera

$$I_5 = \int_0^1 \frac{e^x}{e^{2x}+4}\,dx = \int_1^e \frac{1}{u^2+4}\,du.$$

Il faut maintenant travailler comme dans le point précédent de cet exercice. Soit $t = \frac{u}{2}$ alors

$$I_5 = \int_{1/2}^{e/2} \frac{2}{4(t^2+1)}\,dt = \frac{1}{2}\left[\arctan\left(\frac{e}{2}\right) - \arctan\left(\frac{1}{2}\right)\right].$$

(6) Le bon changement de variable nous est donné par l'énoncé. On a $(1+\tan^2(u))du = dx$, les bornes d'intégration deviennent $\arctan(0) = 0$ et $\arctan(1) = \pi/4$ et notre intégrale sera

$$I_6 = \int_0^1 \frac{1}{(1+x^2)^{\frac{3}{2}}}\,dx = \int_0^{\pi/4} \frac{1+\tan^2(u)}{(1+\tan^2(u))^{\frac{3}{2}}}\,du = \int_0^{\pi/4} \frac{1}{(1+\tan^2(u))^{\frac{1}{2}}}\,du.$$

On observe alors que

$$\frac{1}{(1+\tan^2(u))^{\frac{1}{2}}} = \cos(u),$$

et donc

$$I_6 = \int_0^{\pi/4} \cos(u)\,du = \frac{\sqrt{2}}{2}.$$

(3) On utilise le changement de variable $x = a\sin(u)$, qui comporte $dx = a\cos(u)\,du$ et, en utilisant la fonction $u(x) = \arcsin(x/a)$, $u(a) = \pi/2$ et $u(0) = 0$. On obtient

$$I_7 = \int_0^{\pi/2} \sqrt{a^2 - a^2\sin^2(u)}a\cos(u)\,du = a^2\int_0^{\pi/2} |\cos(u)|\cos(u)\,du.$$

Cette dernière intégrale est égale à $a^2\displaystyle\int_0^{\pi/2}\cos^2(u)\,du$, car toutes les valeurs prises par la fonction cosinus lorsque x varie entre 0 et $\pi/2$ sont positives. En intégrant par parties on trouve que

$$\int_0^{\pi/2}\cos^2(u)\,du = \int_0^{\pi/2}\sin^2(u)\,du,$$

d'où on peut écrire

$$\int_0^{\pi/2}\cos^2(u)\,du = \int_0^{\pi/2}\frac{\sin^2(u)+\cos^2(u)}{2}\,du = \frac{\pi}{4}.$$

La velauer de l'intégrale I_7 est $a^2\frac{\pi}{4}$.

(8) On intègre par parties

$$I_8 = \int_{-\pi}^{0} e^x \sin(x)\,dx = [e^x \sin(x)]_{-\pi}^{0} - \int_{-\pi}^{0} e^x \cos(x)\,dx$$
$$= 0 - [e^x \cos(x)]_{-\pi}^{0} - \int_{-\pi}^{0} e^x \sin(x)\,dx = -1 - e^{-\pi} - I_8$$

On conclut que $I_8 = \frac{-1-e^{-\pi}}{2}$.

(9) On intègre par parties

$$I_9 = \int_1^2 x^2 \ln(2x)\,dx = \left[\frac{x^3}{3}\ln(2x)\right]_1^2 - \int_1^2 \frac{x^2}{3}\,dx$$
$$= \left[\frac{8}{3}\ln(4) - \frac{1}{3}\ln(2)\right] - \left[\frac{8}{9} - \frac{1}{9}\right] = \ln\left(\frac{4^{8/3}}{2^{1/3}}\right) - \frac{7}{9} = \ln(32) - \frac{7}{9}.$$

(10) Il faut réduire la fraction rationnelle en élément simples à intégrer. Pour le faire on commence par regarder le dénominateur et remarquer qu'il peut s'écrire comme $x^3(x^2+1)$. Notre objectif sera alors d'écrire $\frac{x^4+1}{x^5+x^3}$ comme une somme entre deux fractions rationnelles de dénominateur respectif x^3 et x^2+1. Comme le polynôme au numérateur a degré 4 il faut prévoir un polynôme de degré 2 au numérateur de x^3 et un polynôme de degré 1 au numérateur de x^2+1. On a alors

$$\frac{x^4+1}{x^5+x^3} = \frac{Ax^2+Bx+C}{x^3} + \frac{Dx+E}{x^2+1}.$$

Cela va nous donner un système de 5 équations pour les 5 inconnues A, B, C, D, E. On trouve que $A=-1$, $B=0$, $C=1$, $D=2$, $E=0$.

$$I_{10} = \int_1^2 \frac{x^4+1}{x^5+x^3}\,dx = \int_1^2 \frac{-x^2+1}{x^3} + \frac{2x}{x^2+1}\,dx = \int_1^2 \frac{-1}{x} + \frac{1}{x^3} + \frac{2x}{x^2+1}\,dx.$$

Par un calcul désormais immédiat on trouve $I_{10} = \ln(5/4) + 3/8$.

(4) $I_{11} = \int_{-3}^{0} \frac{1}{x^2-3x+2}dx$. Il est facile de vérifier que $\frac{1}{x^2-3x+2} = -\frac{1}{x-1} + \frac{1}{x-2}$. En écrivant la fonction à intégrer comme la somme de deux termes nous avons alors

$$I_{11} = -\int_{-3}^{0} \frac{1}{x-1}\,dx + \int_{-3}^{0} \frac{1}{x-2}\,dx = [-\ln(|x-1|) + \ln(|x-2|)]_{-3}^{0} = \ln\left(\frac{8}{5}\right).$$

(5) Par parties : $I_{12} = \int_1^2 \ln^2(x)dx = \left[x\ln^2(x)\right]_1^2 - \int_1^2 2\ln(x)dx = \left[x\ln^2(x) - 2x\ln(x) + 2x\right]_1^2 = 2\left(\ln^2(2) - 2\ln(2) + 1\right)$.

(13) On va essayer le changement de variable $u = e^x + 1$. On a alors $du = e^x\,dx$ et on pourra écrire $e^x = u - 1$. Les bornes d'intégration deviennent $u(0) = 2$ et $u(1) = e+1$.

$$I_{13} = \int_0^1 \frac{e^x-1}{e^x+1}dx = \int_2^{e+1} \frac{u-2}{u(u-1)}\,du.$$

La fonction de u à intégrer peut s'écrire comme la somme de deux fractions rationnelles plus simples de dénominateur respectif u et $u-1$: pour le faire nous devons trouver A et B dans $\mathbb{R}$ tels que

$$\frac{u-2}{u(u-1)} = \frac{A}{u} + \frac{B}{u-1}$$

On obtient $A=2$ et $B=-1$, donc notre intégrale devient

$$I_{13} = \int_2^{e+1} \frac{2}{u} - \frac{1}{u-1}\,du = \ln\left(\frac{(e+1)^2}{4e}\right).$$

(14) Nous utilisons le changement de variable $u = 1+2e^{-x}$. On a alors $du = -2e^{-x}dx$ et $\frac{1}{1-u}du = dx$.

$$I_{14} = \int_1^2 \frac{e^{-2x}}{(1+2e^{-x})^2}\,dx = \int_{1+2/e}^{1+2/e^2} \left(\frac{u-1}{2}\right)^2 \frac{1}{u^2(1-u)}\,dx = -\int_{1+2/e}^{1+2/e^2} \frac{u-1}{4u^2}\,dx$$
$$= \frac{1}{4}\left[\ln(x) + \frac{1}{x}\right]_{1+2/e^2}^{1+2/e} = \frac{1}{4}\left[\ln\left(\frac{e+2}{e^2+2}\right) + \frac{e}{e+2} - \frac{e^2}{e^2+2}\right].$$

(15) Le bon changement de variable nous est donné par l'énoncé. On a $dt = 2xdx$, les bornes d'intégration deviennent $t(0) = 1$ et $t(1) = 2$ et notre intégrale sera

$$I_{15} = \int_0^1 \frac{x^3}{x^2+1}dx = \int_1^2 \frac{t-1}{2t}dt = \left[\frac{t}{2} - \frac{1}{2}\ln(t)\right]_1^2 = \frac{1}{2}\left(1 - \ln(2)\right).$$

98.9 TD 9 : Équations différentielles : généralités

Exercise 270 exostarterST-0019

Vérifier que $y(x) = -\sqrt{x^2+1}$ est l'unique solution sur $\mathbb{R}$ de l'équation

$$2yy' - 2x = 0, \tag{98.21}$$

qui satisfait la condition initiale $y(0) = -1$. corrstarterST-0019

Correction of the exercise 270

Il nous faut montrer d'abord que la fonction y satisfait l'équation et ensuite que elle vérifie la condition initiale.

On calcule à part la fonction dérivée y'

$$y'(x) = -\left(\sqrt{x^2+1}\right)' = -\left((x^2+1)^{1/2}\right)' = -\frac{1}{2}(x^2+1)^{-1/2}2x = -\frac{x}{\sqrt{x^2+1}},$$

et on substitue y' et y dans le membre de gauche de l'équation par leurs expressions analytiques. On obtient

$$2\left(-\sqrt{x^2+1}\right)\left(-\frac{x}{\sqrt{x^2+1}}\right) - 2x = 2x - 2x = 0,$$

ce qui veut dire que y satisfait l'équation différentielle.

Ensuite nous devons calculer $y(0)$ et montrer qu'il vaut -1 : on a $y(0) = -\sqrt{0^2+1} = -\sqrt{1} = -1$.

Exercise 271 exostarterST-0020

(1) Déterminer une fonction polynomiale qui soit solution sur $\mathbb{R}$ de l'équation différentielle

$$y' - 4y = 4x^3 - 15x^2 + 2x - 3. \tag{98.22}$$

(2) On considère l'équation différentielle

$$y' - 2y = \sin(x). \tag{98.23}$$

Trouver une solution sur $\mathbb{R}$ de cette équation sous la forme $a\cos(x) + b\sin(x)$, où a et b sont des constantes réelles. corrstarterST-0020

Correction of the exercise 271

(1) La solution cherchée est une fonction polynomiale de degré au plus 3. Il s'agit donc d'une fonction de la forme $y(x) = ax^3+bx^2+cx+d$ avec a, b, c, d réels à déterminer. Nous calculons la dérivée de cette fonction, $y'(x) = 3ax^2 + 2bx + c$, et ensuite nous faisons la substitution dans l'équation

$$\big(3ax^2 + 2bx + c\big) - 4\left(ax^3 + bx^2 + cx + d\right) = 4x^3 - 15x^2 + 2x - 3.$$

Cette équation ne peut être vérifiée que si les coefficients des termes du même degré sont égaux. Pour trouver les bonnes valeurs de a, b, c, d dans $\mathbb{R}$ il suffit donc d'écrire un système comme le suivant

$$\begin{cases} -4a = 4, & \text{ici on impose que les coefficients des termes de degré 3 soient égaux;} \\ 3a - 4b = -15, & \text{ici on impose que les coefficients des termes de degré 2 soient égaux;} \\ 2b - 4c = 2, & \text{ici on impose que les coefficients des termes de degré 1 soient égaux;} \\ c - 4d = -3, & \text{ici on impose que les coefficients des termes de degré 0 soient égaux.} \end{cases}$$

On a alors $a = -1$, $b = 3$, $c = 1$ et $d = 1$. La solution polynomiale de l'équation différentielle est $y(x) = -x^3 + 3x^2 + x + 1$.

(2) Ici aussi il faut procéder par substitution. La dérivée de $y(x) = a\cos(x) + b\sin(x)$ est $y'(x) = -a\sin(x) + b\cos(x)$, donc on obtient

$$-a\sin(x) + b\cos(x) - 2\left(a\cos(x) + b\sin(x)\right) = \sin(x)$$

Cette équation doit être satisfaite pour tout valeur de x. En particulier, si x vaut 0 alors tous les termes avec $\sin(x)$ sont nuls et si x vaut $\pi/2$ alors tous les termes avec $\cos(x)$ sont nuls. Cela permet d'écrire un système de deux équations pour déterminer les deux inconnues a et b :

$$\begin{cases} b - 2a = 0 & \text{ici on prend } x = 0 \\ -a - 2b = 1 & \text{ici on prend } x = \pi/2 \end{cases}$$

Exercise 272 exostarterST-0021

On considère, sur $\mathbb{R}$, l'équation différentielle

$$y' = y + x. \tag{98.24}$$ nonhom

(1) Soient y_1 et y_2 deux solutions de l'équation (98.24). Montrer que la fonction différence $y_h = y_1 - y_2$ est une solution de

$$y' = y. \tag{98.25}$$

(2) Trouver une solution de (98.24) de la forme $ax + b$, où a et b sont des nombres réels.

(3) Comment trouve-t-on la solution générale de (98.24) ?

corrstarterST-0021

Correction of the exercise 272

(1) Soient y_1 et y_2 deux solutions de l'équation (98.24). La fonction différence $y_h = y_1 - y_2$ est dérivable et sa dérivée est donné par $y_h' = y_1' - y_2'$, ce qui implique que

$$y_h' = y_1' - y_2' = (y_1 + x) - (y_2 + x) = y_1 - y_2 = y_h.$$

(2) On remplace y dans l'équation (98.24) par une fonction de la forme $y_p = ax + b$, où a et b sont des nombres réels, et on obtient des conditions sur a et b

$$y_p' = y_p + x \;\Rightarrow\; a = (a+1)x + b, \qquad \text{pour tout } x \in \mathbb{R}.$$

On a alors que $a = b$ et $b = -1$, donc $y_p(x) = -x - 1$.

(3) La solution générale de l'équation (98.24), $\mathcal{Y}$, est donnée par la somme de y_p et de la solution générale de l'équation $y' = y$. Cette dernière équation a comme solution générale $\mathcal{Y}_h = \{Ce^x \,:\, C \in \mathbb{R}\}$, donc on a $\mathcal{Y} = \{Ce^x - x - 1 \,:\, C \in \mathbb{R}\}$.

98.10 TD 10 : Équations différentielles : résolution

Exercise 273 exoautoanalyseCTU-32

Équations différentielles à variables séparables

(1) Déterminer la solution f de l'équation différentielle (E_1) : $(1+y)y' = 4x^3$, qui vérifie la condition initiale $f(5) = 14$.

(2) Déterminer la solution g de l'équation différentielle (E_2) : $y' = xy^2$ qui vérifie la condition initiale $g(1) = -\dfrac{1}{2}$.

(3) On considère l'équation différentielle (E_3) : $y^2y' = x^2$.

(3a) Déterminer la forme générale des solutions.

(3b) Déterminer les solutions φ_1, φ_2 et φ_3 qui vérifient respectivement : $\varphi_1(0) = 0$, $\varphi_2(0) = 1$ et $\varphi_3(0) = -1$.

corrautoanalyseCTU-32

Correction of the exercise 273

(1) On commence par trouver la solution générale de (E_1). Comme l'équation est à variables séparables nous allons intégrer de deux côtés par rapport à la variable x.

$$\int (1+y(x))y'(x)\,dx = \int 4x^3\,dx.$$

Nous faisons le changement de variable $y = y(x)$ dans le membre de gauche pour obtenir une expression facile à intégrer

$$\int (1+y)\,dy = \int 4x^3\,dx.$$

et finalement nous avons

$$y + \frac{y^2}{2} = x^4 + C.$$

Il n'est pas possible de trouver une forme explicite pour y sans prendre en compte la condition initiale, car la fonction $y \mapsto \frac{y^2}{2} + y$ n'admet pas de réciproque. Il vaut mieux, dans ce cas, dire que la solution générale de l'équation est l'ensemble de toutes les fonctions y qui satisfont $y + \frac{y^2}{2} = x^4 + C$ pour un quelque C dans $\mathbb{R}$.
Pour déterminer la solution particulière f nous devons remplacer x par 5 et y par 14 dans $y + \frac{y^2}{2} = x^4 + C$. Cela nous permet de fixer une valeur de C : $14 + \frac{(14)^2}{2} = (5)^4 + C$ implique $C = -513$.
La forme explicite de f est donnée par la formule suivante :

$$f(x) = \frac{-2 + 2\sqrt{1 + 2(x^4 - 513)}}{2},$$

où le fait de connaître la valeur de f en $x = 5$ nous a permis de choisir entre les deux racines

$$r_1(x) = \frac{-2 + 2\sqrt{1 + 2(x^4 - 513)}}{2} \tag{98.26a}$$

$$r_2(x) = \frac{-2 - 2\sqrt{1 + 2(x^4 - 513)}}{2}. \tag{98.26b}$$

(2) On commence par traiter le cas où $y = 0$: on aurait alors que $y' = 0$ et la seule solution possible de l'équation différentielle serait la fonction qui vaut toujours zéro. Cette fonction ne satisfait pas la condition initiale et par conséquent nous pouvons supposer que $g(x) \neq 0$ pour tout x. Cela nous permet de diviser par y^2 les deux membres de l'équation différentielle et de séparer ainsi les variables

$$\frac{y'}{y^2} = x.$$

En intégrant des deux côtés nous obtenons la solution générale $\frac{1}{y} = -\left(\frac{x^2}{2} + C\right)$ avec $C \in \mathbb{R}$, qu'on peut expliciter par $y = \frac{-2}{x^2+C}$. La valeur de C qui correspond à g est facile à trouver : $-\frac{1}{2} = \frac{-2}{1+C}$ implique $C = 3$, c'est-à-dire $g(x) = \frac{-2}{x^2+3}$.

(3) (2a) En intégrant des deux côtés de l'équation nous obtenons $\frac{y^3}{3} = \frac{x^3}{3} + C$, donc la forme explicite de la la solution générale de cette équation est $y = \left(x^3 + C\right)^{1/3}$.

(2b) La solution particulière qui satisfait la condition $\phi_1(0) = 0$ est $\phi_1(x) = x$ car la constante C doit être nulle et $\left(x^3\right)^{1/3} = x$. La solution particulière qui satisfait la condition $\phi_2(0) = 1$ est $\phi_2(x) = \left(x^3 + 1\right)^{1/3}$, celle qui satisfait la condition $\phi_3(0) = -1$ est $\phi_3(x) = \left(x^3 - 1\right)^{1/3}$.

Exercise 274 exoautoanalyseCTU-33

Équations différentielles linéaires du premier ordre sans second membre

(1) Résoudre sur $\mathbb{R}$ l'équation différentielle $y' - xy = 0$ et déterminer la solution φ de cette équation qui vérifie $\varphi(1) = 2$.

(2) Résoudre sur $]0\,;+\infty[$ l'équation différentielle $xy' - y = 0$ et déterminer la solution φ de cette équation qui vérifie $\varphi(1) = 2$.

(3) Résoudre sur $]0\,;+\infty[$ l'équation différentielle $y' - \left(\frac{2}{x^2}\right) y = 0$.

corrautoanalyseCTU-33

Correction of the exercise 274

Dans cet exercice nous allons utiliser la formule (32.396). Si vous oubliez la formule vous pouvez vous en sortir sans problèmes en utilisant la méthode de résolution pour les équations à variables séparables, dont les équations linéaires du premier ordre sans second membre sont un cas particulier.

(1) Dans ce cas, $a(x) = 1$ et $b(x) = -x$ donc la formule nous donne

$$\mathcal{Y} = \left\{ Ke^{\int x\,dx} = Ke^{x^2/2}, \qquad K \in \mathbb{R},\ x \in \mathbb{R} \right\}.$$

Pour déterminer la solution φ nous utilisons la condition donnée dans l'énoncé, ce qui nous donne $K = 2e^{-1/2}$.

(2) Ici $a(x) = x$ et $b(x) = -1$ donc la formule nous donne

$$\mathcal{Y} = \left\{ Ke^{\int \frac{1}{x}\,dx} = Ke^{\ln(x)} = Kx, \qquad K \in \mathbb{R},\ x \in]0\,;+\infty[\right\}.$$

Pour déterminer la solution φ nous utilisons la condition donnée dans l'énoncé, ce qui nous donne $K = 2$. Il faut observer que dans la formule nous avons pu intégrer $1/x$ sans crainte et ensuite omettre la valeur absolue parce que l'énoncé nous dit que l'intervalle sur lequel nos solutions sont définies est $]0\,;+\infty[$.

(3) Par la formule de résolution présentée dans le cours la solution générale de cette équation est

$$\mathcal{Y} = \left\{ Ke^{\int \frac{2}{x^2}\,dx} = Ke^{-2/x}, \qquad K \in \mathbb{R},\ x \in]0\,;+\infty[\right\}.$$

Exercise 275 exoautoanalyseCTU-34bis

(1) On cherche à résoudre sur $I =]0\,;+\infty[$ l'équation différentielle $(E) : y' - \frac{3}{x}y = x$.

(1a) Résoudre l'équation homogène (H) associée à (E) sur I.

(1b) Utiliser la méthode de variation de la constante pour trouver une solution particulière de (E).

(1c) En déduire l'ensemble des solutions de (E) sur l'intervalle I.

(1d) Déterminer la solution de (E) qui prend la valeur 2 en 1.

(2) Résoudre sur $]0\,;+\infty[$ l'équation différentielle $xy' + 2y = \dfrac{x}{x^2+1}$.

corrautoanalyseCTU-34bis

Correction of the exercise 275

La différence entre les deux points de cet exercice est que dans l'énoncé du premier on a détaillé chaque passage de la résolution, alors que l'énoncé du deuxième est plus synthétique. En pratique, comme on verra, on demande essentiellement la même chose.

(1) (1a) L'équation homogène (H) associée à (E) sur I est

$$y' - \frac{3}{x}y = 0.$$

Sa solution générale est donnée par la formule (32.396)

$$\mathcal{Y}_h = \left\{ Ke^{\int \frac{3}{x}\,dx} = Kx^3 \ : \ K \in \mathbb{R},\ x \in I \right\}.$$

(1b) La méthode de variation de la constante consiste à remplacer la constante K dans $\mathcal{Y}_h$ par une fonction $x \mapsto K(x)$ à déterminer et ensuite injecter la fonction "candidate" solution $y_p = K(x)x^3$ dans (E) à la place de l'inconnue y. Nous avons donc

$$K'(x)x^3 + 3K(x)x^2 - \frac{3}{x}(K(x)x^3) = x,$$

c'est-à-dire

$$K'(x)x^3 = x, \ \text{ ou encore } K'(x) = \frac{1}{x^2}.$$

L'ensemble des primitives de $1/x^2$ est $\mathcal{P} = -1/x + C$, avec $C \in \mathbb{R}$ donc une solution particulière y_p est $-x^2$ (il suffit de prendre $C = 0$).

(1c) L'ensemble des solutions de (E) sur l'intervalle I est la somme entre la solution générale $\mathcal{Y}_h$ et y_p.

$$\mathcal{Y} = \left\{ -x^2 + Kx^3 \ : \ C \in \mathbb{R},\ x \in I \right\}. \qquad (98.27)$$

solgenexo34bis1

Remarque 98.5.
On peut aussi écrire $\mathcal{Y}$ comme le produit entre l'ensemble des primitives de K' et la fonction x^3, on écrirait alors $\mathcal{Y} = \left\{ \left(-\dfrac{1}{x} + C \right) x^3 \ : \ C \in \mathbb{R},\ x \in I \right\}$.

(1d) On doit simplement trouver la bonne valeur de K dans (98.27). Si $x = 1$ on a $y = -1^2 + K \cdot 1^3 = -1 + K$ Donc $y(1) = 2$ si et seulement si $K = 3$.

(2) Résoudre sur $]0\,;+\infty[$ l'équation différentielle $xy' + 2y = \frac{x}{x^2+1}$. Ici aussi nous devons trouver d'abord la solution générale de l'équation homogène associé $xy' + 2y = 0$, qui est, par la formule (32.396)

$$\mathcal{Y}_h = \left\{ Ke^{-\int \frac{2}{x}\,dx} = \frac{K}{x^2} \ : \ K \in \mathbb{R},\ x \in I \right\}.$$

Ensuite nous appliquons la méthode de variation de la constante pour déterminer les solutions de l'équation de départ. Il faut injecter dans l'équation la "candidate" solution $\dfrac{K(x)}{x^2}$, ce qui donne

$$x\left(\frac{K'(x)}{x^2} - 2\frac{K(x)}{x^3} \right) + 2\frac{K}{x^2} = \frac{x}{x^2+1},$$

$$K'(x) = \frac{x^2}{x^2+1},$$

Pour trouver les primitives de K' nous pouvons utiliser le fait que $\dfrac{x^2}{x^2+1} = 1 - \dfrac{1}{x^2+1}$, et obtenir $K(x) = x - \arctan(x) + C$ pour $C \in \mathbb{R}$. La solution générale de l'équation est donc

$$\mathcal{Y} = \left\{ \frac{x - \arctan(x) + C}{x^2} + \ : \ C \in \mathbb{R},\ x \in I \right\}.$$

Exercise 276 exoautoanalyseCTU-36

Équations différentielles linéaires du second ordre sans second membre

(1) (1a) Déterminer la solution générale de l'équation différentielle (H_1) : $y'' - 5y' + 6y = 0$.

(1b) Déterminer la solution particulière φ_1 qui vérifie $\varphi_1(0) = -2$ et $\varphi_1'(0) = -2$.

(2) (2a) Déterminer la solution générale de l'équation différentielle (H_2) : $y'' + 2y' + 2y = 0$.

(2b) Déterminer la solution particulière φ_2 qui vérifie $\varphi_2(0) = -2$ et $\varphi_2'(0) = -2$.

(3) (3a) Déterminer la solution générale de l'équation différentielle (H_3) : $y'' - 4y' + 4y = 0$.

(3b) Déterminer la solution particulière φ_3 qui vérifie $\varphi_3(0) = -2$ et $\varphi_3'(0) = -2$.

corrautoanalyseCTU-36

Correction of the exercise 276

(1) (1a) Le polynôme caractéristique de l'équation (H_1) est $r^2 - 5r + 6$, dont les racines sont 2 et 3. Nous sommes donc dans le cas où les deux racines sont réelles et distinctes. La solution générale de cette équation est donc, en appliquant la formule (32.407)

$$\mathcal{Y}_h = \left\{C_1 e^{2x} + C_2 e^{3x} \ : \ C_1,\, C_2 \in \mathbb{R}\right\}.$$

(1b) L'ensemble des deux conditions nous permet d'écrire un système de des équations pour déterminer les deux inconnues C_1 et C_2.

$$\begin{cases} C_1 + C_2 = -2 & \text{qui correspond à la condition } \varphi_1(0) = -2, \\ 2C_1 + 3C_2 = -2 & \text{qui correspond à la condition } \varphi_1'(0) = -2. \end{cases}$$

On a donc $C_1 = -4$ et $C_2 = 2$.

La solution particulière φ_1 est donc $\varphi_1(x) = -4e^{2x} + 2e^{3x}$.

(2) (2a) Le polynôme caractéristique de l'équation (H_2) est $r^2 + 2r + 2$, qui admet deux racines complexes conjuguées : $-1 + i$ et $-1 - i$. La solution générale réelle de l'équation est déterminée à partir de la formule (32.409)

$$\mathcal{Y}_h = \left\{e^{-x}\left(C_1 \cos(x) + C_2 \sin(x)\right) \ : \ C_1,\, C_2 \in \mathbb{R}\right\}. \qquad (98.28)$$

exoautoanal36part2

(2b) Comme on a fait pour l'équation (H_1), il faut utiliser les deux conditions pour écrire un système de deux équations dans les inconnues C_1 et C_2.

$$\begin{cases} C_1 = -2 & \text{qui correspond à la condition } \varphi_2(0) = -2, \\ -C_1 + C_2 = -2 & \text{qui correspond à la condition } \varphi_2'(0) = -2. \end{cases}$$

On a donc $C_1 = -2$ et $C_2 = -4$.

La solution particulière φ_2 est donc $\varphi_2(x) = -2e^{-x}\left(\cos(x) + \sin(x)\right)$.

(3) (3a) Le polynôme caractéristique de l'équation (H_3) est $r^2 - 4r + 4$, qui a une racine double $r = 2$. La solution générale réelle de l'équation est déterminée à partir de la formule (32.410)

$$\mathcal{Y}_h = \left\{(C_1 + C_2 x)e^{2x} \ : \ C_1,\, C_2 \in \mathbb{R},\, x \in I\right\}.$$

(3b) En travaillant comme dans les deux cas précédents nous obtenons $C_1 = -2$ et $C_2 = 2$.

La solution particulière φ_3 est donc $\varphi_3(x) = 2(-1 + x)e^{2x}$.

Exercise 277 exoautoanalyseCTU-37

Équations différentielles linéaires du second ordre avec second membre

Résoudre sur $\mathbb{R}$ chacune des équations différentielles suivantes, puis donner la solution avec la condition initiale $y(0) = 0$ et $y'(0) = 0$.

(1) $y'' + 2y' + 2y = 6e^{-x}$

(2) (2a) $y'' - 2y' + 5y = \cos x$

(2b) $y'' - 2y' + 5y = x$

(2c) $y'' - 2y' + 5y = x + \cos x$

corrautoanalyseCTU-37

Correction of the exercise 277

(1) On connaît la solution générale de l'équation homogène associé à $y'' + 2y' + 2y = 6e^{-x}$ pour l'avoir trouvée dans l'exercice 276. Ici notre préoccupation principale sera donc de trouver une solution particulière de l'équation non homogène. Comme le terme de droite de l'équation est une exponentielle nous allons chercher une solution de la forme $y_p(x) = Ce^{-x}$. On obtient

$$Ce^{-x} - 2Ce^{-x} + 2Ce^{-x} = 6e^{-x},$$

ce qui implique que $C = 6$. La solution particulière y_p est donc $y_p = 6e^{-x}$. La solution générale de l'équation en examen est alors la somme entre la solution générale de l'équation homogène associée, (98.28), et y_p :

$$\mathcal{Y} = \left\{ e^{-x} \left(C_1 \cos(2x) + C_2 \sin(2x) + 6 \right) \,:\, C_1,\, C_2 \in \mathbb{R} \right\}.$$

les conditions initiales fixées nous permettent d'écrire le système de deux équations en deux inconnues C_1 et C_2

$$\begin{cases} C_1 + 6 = 0, & \text{qui correspond à la condition } y(0) = 0, \\ -(C_1 + 6) + 2C_2 = 0 & \text{qui correspond à la condition } y'(0) = 0. \end{cases}$$

On trouve alors $C_1 = -6$ et $C_2 = 0$, et $y(x) = 6e^{-x}\left(1 - \cos(2x)\right)$.

(2) (2a) Le polynôme caractéristique de l'équation homogène $y'' - 2y' + 5y = 0$ est $r^2 - 2r + 5$, dont les racines sont les nombres complexes conjugués $1 + i2$ et $1 - i2$. La solution générale de l'équation homogène est donc $\mathcal{Y}_h = \left\{ e^x \left(C_1 \cos(2x) + C_2 \sin(2x) \right) \,:\, C_1,\, C_2 \in \mathbb{R} \right\}$.
On cherche une solution particulière y_p de l'équation non homogène $y'' - 2y' + 5y = \cos(x)$. Comme le membre de droite est la fonction cosinus, $y_p(x)$ sera de la forme $A\cos(x) + B\sin(x)$. En injectant y_p dans l'équation nous obtenons

$$-\left(A\cos(x) + B\sin(x)\right) - 2\left(B\cos(x) - A\sin(x)\right) + 5\left(A\cos(x) + B\sin(x)\right) = \cos(x),$$

qui correspond au système de deux équations pour les deux inconnues A et B

$$\begin{cases} -A - 2B + 5A = 1, & \text{ces sont les coefficients de cosinus,} \\ -B + 2A + 5B = 0, & \text{ces sont les coefficients de sinus.} \end{cases}$$

On a alors $A = 1/5$ et $B = -1/10$, et $y_p = \dfrac{1}{5}\cos(x) - \dfrac{1}{10}\sin(x)$.
La solution générale de l'équation $y'' - 2y' + 5y = \cos(x)$ est la somme de $\mathcal{Y}_h$ et y_p, c'est-à-dire

$$\mathcal{Y} = \left\{ e^x \left(C_1 \cos(2x) + C_2 \sin(2x) \right) + \frac{1}{5}\cos(x) - \frac{1}{10}\sin(x) \,:\, C_1,\, C_2 \in \mathbb{R} \right\}.$$

Les conditions initiales fixées nous donnent le système suivant pour trouver les valeurs de C_1 et C_2 correspondants à la solution particulière demandée :

$$\begin{cases} C_1 + \frac{1}{5} = 0, & \text{qui correspond à la condition } y(0) = 0, \\ C_1 + 2C_2 - \frac{1}{10} = 0 & \text{qui correspond à la condition } y'(0) = 0. \end{cases}$$

Donc $C_1 = -1/5$ et $C_2 = 3/20$ et $y(x) = e^x\left(-\frac{1}{5}\cos(2x) + \frac{3}{20}\sin(2x)\right) + \frac{1}{5}\cos(x) - \frac{1}{10}\sin(x)$.

(b) On a établi plus tôt dans l'exercice que la solution générale de l'équation homogène est donc $\mathcal{Y}_h = \{e^x (C_1 \cos(2x) + C_2 \sin(2x)) : C_1, C_2 \in \mathbb{R}\}$.
On cherche une solution particulière y_p de l'équation non homogène $y'' - 2y' + 5y = x$. Comme le membre de droite est un polynône de degré un, $y_p(x)$ sera de la forme $Ax+B$. En injectant y_p dans l'équation nous obtenons

$$-2A + 5(Ax + B) = x,$$

qui correspond au système de deux équations pour les deux inconnues A et B

$$5A = 1, \text{ et, } -2A + 5B = 0,$$

On a alors $A = 1/5$ et $B = 2/25$, et $y_p = \frac{1}{5}x + \frac{2}{25}$.
La solution générale de l'équation $y''-2y'+5y = x$ est la somme de $\mathcal{Y}_h$ et y_p, c'est-à-dire

$$\mathcal{Y} = \left\{ e^x (C_1 \cos(2x) + C_2 \sin(2x)) + \frac{1}{5}x + \frac{2}{25} \ : C_1, C_2 \in \mathbb{R} \right\}.$$

Les conditions initiales fixées nous donnent le système suivant pour trouver les valeurs de C_1 et C_2 correspondants à la solution particulière demandée :

$$\begin{cases} C_1 + \frac{2}{25} = 0, & \text{qui correspond à la condition } y(0) = 0, \\ C_1 + 2C_2 + \frac{1}{5} = 0 & \text{qui correspond à la condition } y'(0) = 0. \end{cases}$$

Donc $C_1 = -2/25$ et $C_2 = -3/50$ et $y(x) = e^x \left(-\frac{2}{25}\cos(2x) - \frac{3}{50}\sin(2x)\right) + \frac{1}{5}x + \frac{2}{25}$.

(c) Pour obtenir une solution particulière de cette équation il suffit de sommer les solutions particulières trouvées pour les équations des points (Z.a) et (2.b) de cet exercice (on exploite ici le fait que l'équation soit linéaire). La solution générale de cette équation est donc

$$\mathcal{Y} = \left\{ e^x (C_1 \cos(2x) + C_2 \sin(2x)) + \frac{1}{5}\cos(x) - \frac{1}{10}\sin(x) + \frac{1}{5}x + \frac{2}{25} \ : C_1, C_2 \in \mathbb{R} \right\}.$$

Les conditions initiales fixées nous donnent le système suivant pour trouver les valeurs de C_1 et C_2 correspondants à la solution particulière demandée :

$$\begin{cases} C_1 + \frac{7}{25} = 0, & \text{qui correspond à la condition } y(0) = 0, \\ C_1 + 2C_2 + \frac{1}{10} = 0 & \text{qui correspond à la condition } y'(0) = 0. \end{cases}$$

Donc $C_1 = -7/25$ et $C_2 = 9/100$ et

$$y(x) = e^x \left(-\frac{7}{25}\cos(2x) + \frac{9}{100}\sin(2x)\right) + \frac{1}{5}\cos(x) - \frac{1}{10}\sin(x) + \frac{1}{5}x + \frac{2}{25}.$$

98.11 TD 11 : Développements limités

Exercise 278 [exoautoanalyseCTU-40]

(1) (1a) Déterminer un développement limité de e^{3x} à l'ordre 2 en 0.

(1b) En déduire : $\lim_{x\to 0} \dfrac{e^{3x} - 1}{x}$.

(2) (2a) Déterminer un développement limité de $\ln(1 + x)$ à l'ordre 4 en 0.

(2b) Déterminer : $\lim_{x\to 0} \dfrac{\ln(1 + x)}{x}$.

[corrautoanalyseCTU-40]

Correction of the exercise 278

(1) (1a) Nous allons utiliser la formule de Taylor-Young (12.1366). Soit $f(x) = e^{3x}$, alors le développement cherché est

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + x^2\alpha(x) = 1 + 3x + \frac{9}{2}x^2 + x^2\alpha(x).$$

Remarque 98.6.
Un développement similaire a été fait dans le cours dans l'exemple 15.121 en utilisant la règle de développement d'une fonction composée.

(1b)

$$\lim_{x\to 0}\frac{e^{3x}-1}{x} = \lim_{x\to 0}\frac{3x + \frac{9}{2}x^2 + x^2\alpha(x)}{x} = 3.$$

(2) (2a) Ce développement limité est dans le tableau des développements limités à connaître. Pour le calculer on peut utiliser encore une fois la formule de Taylor-Young (12.1366). Soit $f(x) = \ln(1+x)$, on a

$$\begin{aligned} f(x) =& f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + x^4\alpha(x) = \\ &= 0 + x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + x^4\alpha(x). \end{aligned}$$

(2b)

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = \lim_{x\to 0}\frac{x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + x^4\alpha(x)}{x} = 1.$$

Exercise 279 exoautoanalyseCTU-42

(1) Déterminer le développement limité à l'ordre 5 en 0 de la fonction f définie par $f(x) = \cos(x) + \sin(x)$.

(2) (2a) Déterminer, par la règle de développement des produits de fonctions, le développement limité à l'ordre 3 en 0 de la fonction g définie par $g(x) = \cos(x)\sin(x)$.

(2b) En déduire $g'(0)$, $g''(0)$ et $g^{(3)}(0)$. (Comparer le développement obtenu au point précédent avec la formule de Taylor-Young).

corrautoanalyseCTU-42

Correction of the exercise 279

(1) Il suffit de sommer terme à terme les développements connus de cosinus et de sinus. On obtient

$$f(x) = 1 + x - \frac{x^2}{2} - \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + x^5\alpha(x).$$

(2) (2a) On fait le produits des développements limités à l'ordre 3 des fonctions sinus et cosinus

$$g(x) = \left(1 - \frac{x^2}{2}\right)\left(x - \frac{x^3}{3!}\right) + x^3\alpha(x) = x - \frac{x^3}{2} - \frac{x^3}{3!} + x^3\alpha(x) = x - \frac{2x^3}{3} + x^3\alpha(x).$$

(2b) Par comparaison avec la formule de Taylor-Young nous avons que $g'(0) = 1$, $g''(0) = 0$ et $g^{(3)}(0) = -4$.

Exercise 280 exoautoanalyseCTU-43

(1) Déterminer le développement limité de : $x \mapsto e^x$ à l'ordre 4 en 1.

(2) Déterminer le développement limité de : $x \mapsto \sin x$ à l'ordre 4 en $\frac{\pi}{2}$.

(3) Déterminer le développement limité de : $x \mapsto \dfrac{\ln(1+x)}{\sqrt{1+x}}$ à l'ordre 3 en 0.

corrautoanalyseCTU-43

Correction of the exercise 280

(1) Nous pouvons soit utiliser la formule générale de Taylor-Young, (12.1366), soit se ramener à un développement limité au voisinage de zéro par une translation, ce qui veut dire, développer au voisinage de zéro la fonction $g(x) = e^{1+x}$. Les deux méthodes sont absolument équivalentes.

Nous allons essayer la première, car il y a des exemples de la deuxième méthode dans le chapitre sur les développements limités. Remarquez que $(e^x)^{(m)} = e^x$ pour tout $m \in \mathbb{N}^*$, donc on pourra mettre à facteur e, qui est la valeur de la dérivée de l'exponentielle calculée en $x = 1$. Il faut d'abord appliquer la formule et ensuite faire tous les calculs nécessaires à simplifier au maximum l'expression obtenue

$$\begin{aligned}
e^x =& e\left(1 + (x-1) + \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3!} + \frac{(x-1)^4}{4!}\right) + (x-1)^4\alpha(x-1) \\
&= e\left(1 + (x-1) + \frac{x^2+2x+1}{2} + \frac{x^3-3x^2+3x-1}{6} + \frac{x^4-4x^3+6x^2-4x+1}{24}\right) \\
&\qquad + (x-1)^4\alpha(x-1) \\
&= e\left(\frac{12-4+1}{24} + \frac{3x-x}{6} + \frac{x^2}{4} + \frac{x^4}{24}\right) + (x-1)^4\alpha(x-1) \\
&= e\left(\frac{3}{8} + \frac{x}{3} + \frac{x^2}{4} + \frac{x^4}{24}\right) + (x-1)^4\alpha(x-1).
\end{aligned}$$

(2) Le développement à trouver est égale au développement autour de zéro de la fonction $x \mapsto \sin(x + \pi/2)$. On sait que $\sin(x + \pi/2) = \cos(x)\sin(\pi/2) + \cos(\pi/2)\sin(x) = \cos(x)$, et le développement de la fonction cosinus est connu (voir le tableau des développements dans le cours).

(3) Il faudra utiliser ici la règle pour trouver le développement d'un rapport entre fonctions. Nous pouvons commencer par calculer les développements à l'ordre 3 en 0 des fonctions $\ln(1+x)$ et $\sqrt{1+x}$. Ces deux développements sont dans la liste des développements à connïtre, mais on les rappelle ici pour plus de lisibilité :

$$\begin{aligned}
\ln(x+1) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots; \\
\sqrt{1+x} &= (1+x)^{1/2} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + \ldots.
\end{aligned}$$

Nous allons donc procéder à une division pour déterminer le développement du rapport $\dfrac{\ln(1+x)}{\sqrt{1+x}}$.

$$\begin{array}{rcccccc|l}
 & x & - & \frac{x^2}{2} & + & \frac{x^3}{3} & - & \frac{x^4}{4} & \\
-(& x & + & \frac{x^2}{2} & - & \frac{x^3}{8} & + & \frac{x^4}{16} &) \\
\hline
 & & - & x^2 & + & \frac{11x^3}{24} & - & \frac{5x^4}{16} & \\
 & -(& - & x^2 & - & \frac{x^3}{2} & + & \frac{x^4}{8} &) \\
\hline
 & & & & & \frac{23x^3}{24} & - & \frac{7x^4}{16} & \\
 & & & & -(& \frac{23x^3}{24} & + & \frac{x^4}{48} &) \\
\hline
 & & & & & & - & \frac{11x^4}{24} &
\end{array}
\qquad
\begin{array}{|l}
1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} \\
\hline
x - x^2 + \frac{23x^3}{24}
\end{array}$$

Le développement cherché est donc $\dfrac{\ln(1+x)}{\sqrt{1+x}} = x - x^2 + \frac{23x^3}{24} + x^3\alpha(x)$.

Exercise 281 exoautoanalyseCTU-44

On considère la fonction f définie par $f(x) = \dfrac{\cos(x)}{1-x}$.

(1) Donner l'ensemble de définition de f.

(2) Déterminer le développement limité à l'ordre 3 en 0 de la fonction f.

(3) En déduire les valeurs de $f'(0)$, $f''(0)$ et $f^{(3)}(0)$.

(4) On note (C) la courbe représentative de f. Déterminer l'équation de la tangente T à (C) au point d'abscisse 0.

corrautoanalyseCTU-44

Correction of the exercise 281

On considère la fonction f définie par $f(x) = \dfrac{\cos(x)}{1-x}$.

(1) Domaine$_f = \mathbb{R}\backslash\{1\}$.

(2) Le développement de la fonction cosinus est connu

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} + \dots.$$

Pour trouver le développement au voisinage de 0 de la fonction f nous utilisons la règle pour calculer le développement d'un rapport

$$\begin{array}{r|l}
1 - \frac{x^2}{2} + \frac{x^4}{24} & 1-x \\
-\left(1 - x\right) & \hline
\quad + x - \frac{x^2}{2} & 1 + x + \frac{x^2}{2} + \frac{x^3}{2} \\
\quad -\left(+ x - x^2\right) & \\
\hline
\qquad + \frac{x^2}{2} + \frac{x^4}{24} & \\
\qquad -\left(\frac{x^2}{2} - \frac{x^3}{2}\right) & \\
\hline
\qquad\quad + \frac{x^3}{2} + \frac{x^4}{24} & \\
\qquad\quad -\left(+ \frac{x^3}{2} - \frac{x^4}{2}\right) & \\
\hline
\qquad\qquad + \frac{13x^4}{24} &
\end{array}$$

Le développement limité à l'ordre 3 en 0 de la fonction f est $1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{2} + x^3\alpha(x)$.

Remarque 98.7.
Il était possible aussi de calculer ce développement comme le produit entre le développement de cos et de $\dfrac{1}{1-x}$ au voisinage de 0.

(3) Par la formule de Taylor-Young on a $f'(0) = 1$, $f''(0) = 1$ et $f^{(3)}(0) = 3$.

(4) L'équation de la tangente T à (C) au point d'abscisse 0 est $y = 1 + x$. L'équation de la tangente est toujours l'approximation d'ordre 1 de la fonction f au voisinage de 0.

Exercise 282 exoautoanalyseCTU-51

Calculer les limites suivantes à l'aide des développements limités

(1) $\lim\limits_{x\to 0}\left(\dfrac{\ln(1+x) - \sin(x)}{x}\right)$;

(2) $\lim\limits_{x\to 0}\left(\dfrac{(1+x)^{\frac{1}{x}} - e}{x}\right)$;

(3) $\lim\limits_{x\to \frac{\pi}{3}}\left(\dfrac{\sin(3x)}{1 - 2\cos(x)}\right)$.

corrautoanalyseCTU-51

Correction of the exercise 282

(1) On calcule d'abord

$$\ln(x+1)-\sin(x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}-x+\frac{x^3}{6}+x^4\alpha(x)=-\frac{x^2}{2}+x^2\alpha(x),$$

ce qui nous dit que au voisinage de zéro la fonction $\dfrac{\ln(1+x)-\sin x}{x}$ a le même comportement que $-x/2$. La valeur de la limite est donc zéro.

(2) L'expression $(x+1)^{1/x}$ est **par définition dans ce cours** équivalente à $e^{\frac{1}{x}\ln(x+1)}$, ce qui veut dire que son développement limité sera obtenu par la règle de développement d'une fonction composée.
On calcule d'abord

$$\frac{1}{x}\ln(x+1)=\frac{1}{x}\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\dots\right)=1-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\dots$$

On a alors que $e^{\frac{1}{x}\ln(x+1)}\approx e^{1-\frac{x}{2}+\frac{x^2}{3}-\frac{x^3}{4}+\dots}\approx e^{1-\frac{x}{2}}$ lorsque x est proche ce 0.

Remarque 98.8.
Le développement de l'exponentielle autour de 1 à été calculé dans l'exercice 280, il donc est possible d'utiliser le résultat pour terminer le calcul de cette limite, mais pour des raisons pédagogiques nous allons continuer le calcul.

$$e^{1-\frac{x}{2}}=e+\frac{ex}{2}+\dots$$

Nous avons alors que

$$\lim_{x\to0}\left(\frac{(1+x)^{\frac{1}{x}}-e}{x}\right)=\lim_{x\to0}\left(\frac{\frac{ex}{2}}{x}\right)=\frac{e}{2}.$$

(3) Le développement limité de $\sin(3x)$ lorsque x est dans un voisinage de $\pi/3$ est le développement limité au voisinage de 0 de la fonction $g(x)=\sin(3x-\pi)$

$$g(x)=g(0)+g'(0)x+\dots=\sin(-\pi)+3\cos(-\pi)\left(x-\frac{\pi}{3}\right)+\dots=-3\left(x-\frac{\pi}{3}\right)+\dots.$$

Le développement de cos au voisinage de $\pi/3$ a été calculé dans l'exemple 18.31, nous avons donc

$$\cos(x)=\frac{1}{2}-\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{3}\right)+\dots.$$

La limite a calculer devient alors

$$\lim_{x\to\frac{\pi}{3}}\left(\frac{\sin(3x)}{1-2\cos(x)}\right)=\lim_{x\to\frac{\pi}{3}}\frac{-3\left(x-\frac{\pi}{3}\right)}{\sqrt{3}\left(x-\frac{\pi}{3}\right)}=-\sqrt{3}.$$

98.12 Interrogations des années précédentes

Exercise 283 exoanalyseCTU-0010
Soit f la fonction définie sur $\mathbb{R}$ par $f(x)=x^3+5x-10$.

(1) Que vaut $f(2)$?

(2) Étudier les variations de f.

(3) Justifier que f est une bijection de $\mathbb{R}$ dans un intervalle que l'on précisera.

(4) Soit g la bijection réciproque de f. Préciser l'ensemble de définition de g et calculer $g'(8)$.
corranalyseCTU-0010

Correction of the exercise 283

<+CorranalyseCTU-0010+>

Exercise 284 exoanalyseCTU-0011

(1) (**Question de cours**) Donner le domaine de définition, la dérivée et l'aspect de la représentation graphique de la fonction arcsin.

(2) Soit f_1 la fonction définie par $f_1(x) = \cos(\arcsin(x))$.

(2a) Déterminer l'ensemble de définition de f_1.

(2b) Calculer $f_1'(x)$ pour $-1 < x < 1$. Montrer que $f_1'(x) = -\dfrac{x}{\sqrt{1-x^2}}$.

(3) Soit f_2 la fonction définie par $f_2(x) = \sqrt{1-x^2}$.

(3a) Déterminer l'ensemble de définition de f_2.

(3b) Calculer $f_2'(x)$ pour $-1 < x < 1$. Montrer que $f_1'(x) = f_2'(x)$.

(4) Donner une expression simplifiée de $f_1(x)$ valable pour $x \in]-1, 1[$.
corranalyseCTU-0011

Correction of the exercise 284

<+CorranalyseCTU-0011+>

Exercise 285 exoanalyseCTU-0012

(1) (**Question de cours**) Donner le domaine de définition, la dérivée et l'aspect de la représentation graphique de la fonction arcsin.

(2) On considère la fonction f définie par $f(x) = \arcsin(2x^2 - 1)$.

(2a) Déterminer l'ensemble de définition de f.

(2b) Calculer $f'(x)$ pour $0 < x < 1$. Montrer que $f'(x) = \dfrac{2}{\sqrt{1-x^2}}$.

(3) Montrer que pour tout x de l'intervalle $[0\ ;\ 1[$, $f(x) = 2\arcsin x - \dfrac{\pi}{2}$.
corranalyseCTU-0012

Correction of the exercise 285

<+CorranalyseCTU-0012+>

Exercise 286 exoanalyseCTU-0013

(1) Compléter le tableau suivant.

Primitive $\int f(x)\,dx$	Fonction $f(x)$	Dérivée $f'(x)$
...	e^x	...
...	$\frac{1}{x}$	...
...	$\cos(x)$	...
...	x^α	...

(98.29)

(2) Calculer les intégrales suivantes

(2a) $\displaystyle\int_0^1 x^3 + x^{1/3}\,dx$;

(2b) $\displaystyle\int_\pi^{3\pi/2} 5\sin(x)\,dx$;

(2c) $\displaystyle\int_1^2 e^x + \frac{1}{x}\,dx$.

(3) Calculer l'intégrale suivante par la méthode du changement de variable

$$\int_1^2 \frac{3x^2}{x^3+17}\,dx.$$

corranalyseCTU-0013

Correction of the exercise 286

<+CorranalyseCTU-0013+>

Exercise 287 exoanalyseCTU-0014

(1) Compléter le tableau suivant.

Primitive $\int f(x)\,dx$	Function $f(x)$	Dérivée $f'(x)$
...	x^3	...
...	$x^{1/5}$	...
...	$\cos(x)$	...
...	$\frac{1}{1+x^2}$	...

(98.30)

(2) Calculer les intégrales suivantes

(2a) $\displaystyle\int_2^3 \frac{1}{(x-1)(x+2)}\,dx$;

(2b) $\displaystyle\int_0^{1/2} \frac{1}{\sqrt{1-x^2}}\,dx$;

(2c) $\displaystyle\int_0^{\pi/2} e^{\cos(x)}\sin(x)\,dx$. Conseil : utiliser un changement de variable.

(3) Calculer, par parties, les primitives suivantes

(3a) $\displaystyle\int \ln(x)\,dx$;

(3b) $\displaystyle\int x\sin(x)\,dx$.

corranalyseCTU-0014

Correction of the exercise 287

<+CorranalyseCTU-0014+>

Exercise 288 exoanalyseCTU-0015

Dans tout le problème on se place dans l'intervalle $I =]0, +\infty[$.

On considère l'équation différentielle suivante

$$xy' + y = x^2. \tag{98.31}$$

nonhomC

(1) De quel type est l'équation (98.31) ?

(2) Vérifier que $y(x) = \frac{1}{3}x^2 + \frac{1}{x}$ est l'unique solution de (98.31) qui satisfait la condition initiale $y(1) = \frac{4}{3}$.

(3) Vérifier que $y(x) = \frac{1}{3}x^2$ est l'unique solution de (98.31) qui satisfait la condition initiale $y(1) = \frac{1}{3}$.

(4) Quel rapport y a-t-il entre l'équation (98.31) et l'équation différentielle

$$xy' + y = 0 \quad ? \tag{98.32}$$

homC

(5) Trouver l'unique solution de (98.32) qui satisfait la condition initiale $y(1) = 1$.

corranalyseCTU-0015

Correction of the exercise 288

<+CorranalyseCTU-0015+>

Exercise 289 exoanalyseCTU-0016

On considère l'équation différentielle suivante

$$y' + \sin(x)y = 0, \qquad x \in \mathbb{R}. \tag{98.33}$$

nonhomD

(1) De quel type est l'équation (98.33) ?

(2) Déterminer la solution générale de (98.33).

(3) Quel rapport y a-t-il entre l'équation (98.33) et l'équation différentielle

$$y' + \sin(x)y = \frac{e^{\cos(x)}}{1+x^2} \quad ? \tag{98.34}$$

homD

(4) Déterminer la solution générale de (98.34) en utilisant la méthode de variation de la constante.

corranalyseCTU-0016

Correction of the exercise 289

<+CorranalyseCTU-0016+>

Exercise 290 exoanalyseCTU-0017

(1) On considère l'équation différentielle suivante

$$2y'' + 2y' + 5y = 0, \qquad x \in \mathbb{R}. \tag{98.35}$$

nonhomA

(1a) De quel type est l'équation (98.35) ?

(1b) Déterminer la solution générale de (98.35).

(1c) Déterminer la solution de (98.35) qui vérifie la condition initiale $y(0) = 2$, $y'(0) = 2$.

(2) On considère maintenant l'équation différentielle :

$$2y'' + 2y' + 5y = 25x - 5, \qquad x \in \mathbb{R}. \tag{98.36}$$

homA

(2a) Quel est le type de cette équation différentielle (98.36) ? Quel lien y a-t-il entre l'équation (98.36) et l'équation (98.35) ?

(2b) Déterminer une solution particulière f de (98.36) sous la forme $f(x) = \alpha x + \beta$.

(2c) Déterminer la solution générale de (98.36).

corranalyseCTU-0017

Correction of the exercise 290

<+CorranalyseCTU-0017+>

Exercise 291 exoanalyseCTU-0018

(1) On considère l'équation différentielle suivante

$$xy' + y = 0, \qquad x \in]0, +\infty[. \tag{98.37}$$

nonhomB

(1a) De quel type est l'équation (98.37) ?

(1b) Déterminer la solution générale de (98.37).

(2) On considère maintenant l'équation différentielle :

$$xy' + y = \cos(x), \qquad x \in]0, +\infty[. \tag{98.38}$$

homR

(2a) Quel rapport y a-t-il entre l'équation (98.38) et l'équation (98.37) ?

(2b) Déterminer la solution générale de (98.38) en utilisant la méthode de variation de la constante.

corranalyseCTU-0018

Correction of the exercise 291

<+CorranalyseCTU-0018+>

Exercise 292 exoanalyseCTU-0019

(1) **Question de cours :** Écrire la formule de Taylor-Young pour le développement limité à l'ordre 3 au voisinage de 0 d'une fonction g qu'on suppose continue et au moins 3 fois dérivable.

(2) Donner le développement limité à l'ordre 3 de $x \mapsto \cos(x)$ et de $x \mapsto \sin(x)$.

(3) En déduire développement limité à l'ordre 3 de $f(x) = \cos(x)\sin(x)$ au voisinage de 0.

corranalyseCTU-0019

Correction of the exercise 292

<+CorranalyseCTU-0019+>

98.13 Autres exercices

Exercise 293 exomazhe-0012

Calculer par parties les intégrales suivantes.

(1) $\displaystyle\int_0^{\pi} x\cos(x)\,dx$;

(2) $\displaystyle\int_1^2 x^2 e^x\,dx$;

(3) $\displaystyle\int_1^2 \ln(x)\,dx$;

(4) $\displaystyle\int_{-\pi}^0 e^x \sin(x)\,dx$;

(5) $\displaystyle\int_0^{\pi/2} \cos^2(x)\,dx$;

(6) $\displaystyle\int_1^2 x^2 \ln(2x)\,dx$.

corrmazhe-0012

Correction of the exercise 293

<+Corrmazhe-0012+>

Exercise 294 exomazhe-0013

Remarque 98.9.
Cet exercice est donné en devoir.

Calculer les intégrales suivantes.

(1) $\displaystyle\int_{-3}^0 \frac{1}{x^2-3x+2}dx$. Conseil : écrire la fraction sous la forme $\frac{a}{x-1} + \frac{b}{x-2}$.

(2) $\displaystyle\int_1^2 \ln^2(x)dx$.

(3) $\displaystyle\int_0^1 \frac{e^x-1}{e^x+1}dx$.

(4) $\displaystyle\int_1^2 \frac{e^{-2x}}{(1+2e^{-x})^2}dx$.

(5) $\displaystyle\int_0^1 \frac{x^3}{x^2+1}dx$. Conseil : poser $t = x^2+1$.

corrmazhe-0013

Correction of the exercise 294

<+Corrmazhe-0013+>

Exercise 295 exomazhe-0014

Calculer les intégrales suivantes.

(1) $\displaystyle\int_0^1 xe^x dx$

(2) $\displaystyle\int_0^1 x^2e^x dx$

(3) $\displaystyle\int_1^2 \arctan(x)dx$

(4) $\displaystyle\int_{-1}^1 x\left(\cos^{10}(x)+\frac{\sin^4(x)-e^{x^2}+1}{\sqrt{|x|}}\right)dx.$

corrmazhe-0014

Correction of the exercise 295

<+Corrmazhe-0014+>

Exercise 296 exomazhe-0015

Déterminer dans chaque cas le domaine de définition, la périodicité et/ou les symétries éventuelles et les limites aux extrêmes du domaine. Calculer ensuite la dérivée de la fonction, la où elle est définie.

(1) $f_1(x)=x^3-1$;

(2) $f_2(x)=e^{\cos(x)}$;

(3) $f_3(x)=\dfrac{x}{x-2}$.

corrmazhe-0015

Correction of the exercise 296

<+Corrmazhe-0015+>

Exercise 297 exomazhe-0016

Montrer que la fonction tan est une bijection croissante de $]-\pi/2,\pi/2[$ dans $\mathbb{R}$. On note arctan (prononcer « arc tangente ») sa bijection réciproque. Montrer que arctan est dérivable sur $\mathbb{R}$ et que

$$\big(\arctan(x)\big)'=\frac{1}{1+x^2}. \tag{98.39}$$

corrmazhe-0016

Correction of the exercise 297

<+Corrmazhe-0016+>

Exercise 298 exoautoanalyseCTU-31

Notion d'équation différentielle

(1) Déterminer une fonction polynôme qui soit solution sur $\mathbb{R}$ de l'équation différentielle : $y'-4y=4x^3-15x^2+2x-3$.

(2) Déterminer une équation différentielle du premier ordre telle que la fonction φ définie sur $]0\,;+\infty[$ par $\varphi(x)=x\ln(x)$ soit une solution.

(3) On considère l'équation différentielle : $y'-2y=\sin x$. Trouver une solution sur $\mathbb{R}$ de cette équation sous la forme $a\cos x+b\sin x$, où a et b sont des constantes réelles.

corrautoanalyseCTU-31

Correction of the exercise 298

<+CorrautoanalyseCTU-31+>

Exercise 299 exoautoanalyseCTU-34

Équations différentielles linéaires du premier ordre avec second membre

(1) On cherche à résoudre sur $\mathbb{R}$ l'équation différentielle (E) : $y'+x^2y=x^2$.

(1a) Résoudre l'équation homogène (H) associée à (E).

(1b) Trouver une solution évidente de (E).

(1c) En déduire l'ensemble des solutions de (E).

(1d) Déterminer la solution de (E) qui prend la valeur 4 en 0.

(2) Résoudre sur $]0\,;+\infty[$ l'équation différentielle $2xy'+y=\dfrac{1}{1+x}$.

corrautoanalyseCTU-34

Correction of the exercise 299

<+CorrautoanalyseCTU-34+>

Exercise 300 exoautoanalyseCTU-47

Calculer les primitives suivantes :

$$\int \frac{\mathrm{d}x}{(2x+1)^3} \qquad \int \frac{\mathrm{d}u}{\sqrt{1+u}} \qquad \int \frac{t}{1+t^2}\,\mathrm{d}t \qquad \int \frac{\mathrm{d}t}{1+t^2} \qquad \int \frac{x}{(1+x^2)^2}\,\mathrm{d}x \qquad \int \frac{t+1}{4+t^2}\,\mathrm{d}t$$

$$\int \arctan 2x\,\mathrm{d}x \qquad \int x\ln(x+1)\,\mathrm{d}x \qquad \int xe^{2x}\,\mathrm{d}x \qquad \int \frac{1}{x\sqrt{1-\ln x}}\,\mathrm{d}x \quad (\text{poser } t = 1-\ln x)$$

corrautoanalyseCTU-47

Correction of the exercise 300

<+CorrautoanalyseCTU-47+>

Exercise 301 exoautoanalyseCTU-50

(1) Déterminer le développement limité de : $x \mapsto e^x$ à l'ordre 4 en 1.

(2) Déterminer le développement limité de : $x \mapsto \sin x$ à l'ordre 4 en $\dfrac{\pi}{2}$.

(3) Déterminer, de deux façons différentes, le développement limité de : $x \mapsto \dfrac{\ln(1+x)}{1-x}$ à l'ordre 4 en 0.

(4) Déterminer le développement limité de : $x \mapsto \dfrac{\ln(1+x)}{\sqrt{1+x}}$ à l'ordre 3 en 0.

corrautoanalyseCTU-50

Correction of the exercise 301

<+CorrautoanalyseCTU-50+>

Exercise 302 exoautoanalyseCTU-16

Calculer les intégrales suivantes :

$$J_1 = \int_0^1 (4x+2)(x^2+x+1)\,\mathrm{d}x \qquad J_2 = \int_0^1 xe^{x^2}\,\mathrm{d}x \qquad J_3 = \int_0^{\frac{\pi}{4}} \sin 3x\,\mathrm{d}x \qquad J_4 = \int_0^{\frac{\pi}{2}} \frac{\cos x}{2+\sin x}\,\mathrm{d}x$$

corrautoanalyseCTU-16

Correction of the exercise 302

<+CorrautoanalyseCTU-16+>

Exercise 303 exoautoanalyseCTU-17

(1) Déterminer l'ensemble des primitives de la fonction tan sur l'intervalle $\left]-\dfrac{\pi}{2}\,;\,\dfrac{\pi}{2}\right[$.

(2) Déterminer la primitive sur l'intervalle $\left]-\dfrac{\pi}{2}\,;\,\dfrac{\pi}{2}\right[$ de la fonction tan qui s'annule en $\dfrac{\pi}{3}$.

corrautoanalyseCTU-17

Correction of the exercise 303

<+CorrautoanalyseCTU-17+>

Exercise 304 exoautoanalyseCTU-18

Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = |2x-1|$.

(1) Justifier que la fonction f admet des primitives sur $\mathbb{R}$.

(2) Déterminer la primitive F_1 de f qui s'annule en 2.

(3) Tracer les représentations graphiques de f et de F_1.

corrautoanalyseCTU-18

Correction of the exercise 304

<+CorrautoanalyseCTU-18+>

Exercise 305 exoautoanalyseCTU-20

Déterminer les ensembles de primitives suivants :

$$F_1(x)=\int x(x^2+3)\,\mathrm{d}x \qquad F_2(x)=\int x\sqrt{1+x^2}\,\mathrm{d}x \qquad F_3(x)=\int \frac{x^2}{1+x^3}\,\mathrm{d}x$$
corrautoanalyseCTU-20

Correction of the exercise 305

<+CorrautoanalyseCTU-20+>

Exercise 306 exoautoanalyseCTU-21

Déterminer les ensembles de primitives suivants :

$$G_1(x)=\int \ln(x)\,\mathrm{d}x \qquad G_2(x)=\int \arctan(x)\,\mathrm{d}x \qquad G_3(x)=\int x\sin(x)\,\mathrm{d}x$$
$$G_4(x)=\int xe^x\,\mathrm{d}x \qquad G_5(x)=\int \arcsin(x)\,\mathrm{d}x \qquad G_6(x)=\int x^2e^x\,\mathrm{d}x$$
corrautoanalyseCTU-21

Correction of the exercise 306

<+CorrautoanalyseCTU-21+>

Exercise 307 exoautoanalyseCTU-28

Déterminer les ensembles de primitives suivants :

$$F_1(x)=\int \sin^2 x\,\mathrm{d}x \qquad F_2(x)=\int (x^2+x+1)e^x,\mathrm{d}x \qquad F_3(x)=\int \frac{\ln x}{x}\,\mathrm{d}x \qquad F_4(x)=\int \frac{\ln x}{x^3}\,\mathrm{d}x$$
$$F_5(x)=\int x\sin(x)\,\mathrm{d}x \quad F_6(x)=\int \frac{x}{1+x^4}\,\mathrm{d}x F_7(x)=\int \frac{1}{x^2+4x+5}\,\mathrm{d}x F_8(x)=\int \frac{8x}{x^2+4x+5}\,\mathrm{d}x$$
corrautoanalyseCTU-28

Correction of the exercise 307

<+CorrautoanalyseCTU-28+>

Exercise 308 exoautoanalyseCTU-4

(1) (1a) Rappeler les formules de trigonométrie que vous connaissez.

(1b) Déterminer les valeurs exactes de $\cos\dfrac{\pi}{12}$, $\sin\dfrac{7\pi}{12}$ et $\tan\dfrac{5\pi}{12}$.

(1c) Déterminer un réel A et un réel φ tels que :

a/ $\cos x-\sqrt{3}\sin x=A\cos(x+\varphi)$ **b/** $\cos x+\sin x=A\cos(x+\varphi)$

(2) (2a) Exprimer $\cos(2x)$ et $\sin(2x)$ en fonction de $\cos(x)$ et $\sin(x)$.

(2b) Exprimer $\tan(2x)$ en fonction de $\tan(x)$.

(2c) Déterminer les valeurs exactes de $\cos\dfrac{\pi}{8}$ et $\sin\dfrac{\pi}{8}$.

corrautoanalyseCTU-4

Correction of the exercise 308

<+CorrautoanalyseCTU-4+>

Exercise 309 exoautoanalyseCTU-7

(1) La fonction arcsin est-elle paire ou impaire ?

(2) La fonction arccos est-elle paire ou impaire ? Exprimer $\arccos(-x)$ en fonction de $\arccos(x)$.

(3) Montrer que pour tout réel x de [-1 ; 1] : $\arccos(x)+\arcsin(x)=\dfrac{\pi}{2}$

corrautoanalyseCTU-7

Correction of the exercise 309

<+CorrautoanalyseCTU-7+>

Exercise 310 exoautoanalyseCTU-8

Simplifier les écritures suivantes en précisant à chaque fois le domaine de validité de la formule.

a/ $\sin(\arcsin x)$ **b/** $\cos(\arccos x)$ **c/** $\sin(\arccos x)$

corrautoanalyseCTU-8

Correction of the exercise 310

<+CorrautoanalyseCTU-8+>

Exercise 311 exoautoanalyseCTU-9

On considère les fonctions f et g définies respectivement par $f(x)=\arccos(\cos x)$ et $g(x)=\arcsin(\sin x)$.

(1) Donner l'ensemble de définition des fonctions f et g.

(2) Représenter graphiquement les fonctions f et g.

corrautoanalyseCTU-9

Correction of the exercise 311

<+CorrautoanalyseCTU-9+>

Exercise 312 exoautoanalyseCTU-13

(1) Démontrer les formules suivantes :

(1a) $\cosh^2 x - \sinh^2 x = 1$

(1b) $\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y)$

(1c) $\cosh(x+y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y)$

(2) Donner des expressions de $\cosh(2x)$ et $\sinh(2x)$ en fonction de $\cosh(x)$ et $\sinh(x)$.

corrautoanalyseCTU-13

Correction of the exercise 312

<+CorrautoanalyseCTU-13+>

98.14 Exercices en réserve

Exercise 313 exosession1-0001

corrsession1-0001

Correction of the exercise 313

Exercise 314 exomazhe-0017

Soit f la fonction $f(x) = \dfrac{e^x - 2}{e^x + 1}$.

(1) Préciser l'ensemble de définition, les variations et les limites aux bords du domaine de f.

(2) Montrer que pour tout x dans l'ensemble de définition de f on a

$$f'(x) = -\frac{1}{3}\left(f^2(x) + f(x) - 2\right). \tag{98.40}$$

EQooCHQSooPtlbYZ

(3) Montrer que f réalise une bijection de son domaine vers un intervalle que l'on précisera.

(4) Soit g la bijection réciproque de f. Quel est l'intervalle de définition de g ?

(5) Rappeler la formule qui donne la dérivée de la fonction réciproque.

(6) Utiliser l'équation (98.40) pour montrer que la dérivée de g est $g'(y) = -\frac{3}{y^2+y-2}$.

(7) Déterminer l'expression explicite de g.

corrmazhe-0017

Correction of the exercise 314

(1) L'ensemble de définition de f est $\mathbb{R}$, les limites aux bords du domaine sont

$$\lim_{x\to-\infty} f(x) = -2, \qquad \lim_{x\to+\infty} f(x) = 1.$$

La dérivée de f est

$$f'(x) = \frac{e^x(e^x+1) - e^x(e^x-2)}{(e^x+1)^2} = \frac{3e^x}{(e^x+1)^2},$$

qui est une fonction strictement positive. La fonction f est continue et f' est strictement positive, f est donc strictement croissante entre les valeurs -2 et 1.

(2) Nous pouvons calculer explicitement

$$-\frac{1}{3}\big(f^2(x)+f(x)-2\big)=-\frac{1}{3}\frac{(e^x-2)^2+(e^x-2)(e^x+1)-2(e^x+1)^2}{(e^x+1)^2}$$
$$=-\frac{1}{3}\frac{-9e^x}{(e^x+1)^2}=\frac{3e^x}{(e^x+1)^2}.$$

Cette expression est identique à l'expression de f' trouvée au point précédent.

(3) La fonction f est définie sur $\mathbb{R}$, continue et strictement croissante entre les valeurs -2 et 1. Elle réalise donc une bijection de $\mathbb{R}$ vers $]-2,1[$.

(4) L'intervalle de définition de g est $]-2,1[$.

(5) La formule qui donne la dérivée de la fonction réciproque est

$$g'(y)=\frac{1}{f'(g(y))}.$$

(6) Nous avons

$$g'(y)=\frac{1}{f'(g(y))}=\frac{-3}{(f^2(g(y))+f(g(y))-2)}=\frac{-3}{y^2+y-2}.$$

(7) Pour trouver l'expression explicite de g nous trouvons d'abord l'ensemble des primitives de g'

$$\int g'(y)\,dy=\int\frac{-3}{y^2+y-2}\,dy=\int\frac{-3}{(y+2)(y-1)}\,dy$$

On a

$$\frac{-3}{(y+2)(y-1)}=\frac{1}{(y+2)}-\frac{1}{(y-1)},$$

et donc

$$\int g'(y)\,dy=\int\frac{1}{(y+2)}\,dy-\int\frac{1}{(y-1)}\,dy=\ln(|y+2|)-\ln(|y-1|)+C.$$

Il faut trouver maintenant la valeur de la constante C qui correspond à g (car g est une primitive particulière de g'). Il suffit de calculer la valeur de f à un point, par exemple $f(0)=-1/2$. Ensuite forcement $g(-1/2)=0$, donc

$$\ln(|-1/2+2|)-\ln(|-1/2-1|)+C=0$$

et on trouve $C=\ln(1)=0$. La fonction g est $g(y)=\ln(|y+2|)-\ln(|y-1|)$.

Exercise 315 exoanalyseCTU-0101

(1) Montrer que pour tout $y\in[-1,1]$ nous avons

$$\cos(\arcsin(y))=\sin(\arccos(y)).\tag{98.41}$$

(2) Soit la fonction f définie par

$$f(y)=\arctan(y)+\arctan\left(\frac{1}{y}\right).\tag{98.42}$$

(2a) Trouver l'ensemble de définition de f.

(2b) Montrer que la dérivée de f est nulle pour tout y dans le domaine de f.

(2c) En déduire que pour tout $y<0$,

$$\arctan(y)+\arctan\left(\frac{1}{y}\right)=-\frac{\pi}{2}.\tag{98.43}$$

(2d) Est-il vrai que la fonction f est constante ? Expliquer votre réponse.

corranalyseCTU-0101

Correction of the exercise 315

(1) Pour tout $y \in [-1,1]$ nous avons

$$[\cos(\arcsin(y))]^2 = 1 - [\sin(\arcsin(y))]^2 = 1 - y^2,$$

et analoguement

$$[\sin(\arccos(y))]^2 = 1 - [\cos(\arccos(y))]^2 = 1 - y^2.$$

Les deux expressions sont donc équivalentes.

(2) (2a) L'ensemble de définition arctan est $\mathbb{R}$, mais l'ensemble de définition $1/x$ est $\mathbb{R}\backslash\{0\}$. Par consequent l'ensemble de définition de f est $\mathbb{R}\backslash\{0\}$.

(2b) Nous pouvons calculer la dérivée de f, on a

$$f'(y) = \frac{1}{1+y^2} + \frac{1}{1+\left(\frac{1}{y}\right)^2}\frac{-1}{y^2} = \frac{1}{1+y^2} - \frac{1}{1+y^2} = 0,$$

pour tout y dans le domaine de f.

(2c) On peut calculer $f(-1)$, en sachant que sa valeur est la valeur de f pour tout $y < 0$,

$$\arctan(-1) + \arctan(\frac{1}{-1}) = 2\left(-\frac{\pi}{4}\right) = -\frac{\pi}{2}. \tag{98.44}$$

(2d) La fonction f n'est pas constante sur son domaine, car elle prend des valeurs différents dans les intervalles $]0,+\infty[$ et $]-\infty,0[$. Pour $y > 0$ on a en effet $f(y) = f(1) = \pi/2$.

Exercise 316 exoanalyseCTU-0102

(1) Montrer que la fonction sinh admet une fonction réciproque.

(2) Montrer que $f(y) = \ln(y + \sqrt{y^2+1})$ est la fonction réciproque de sinh.

(3) Calculer $f'(y)$ d'abord en utilisant son expression analytique et ensuite en utilisant la formule de dérivation d'une fonction composée et les propriétés des fonctions sinh et cosh vues dans l'exercice 19

corranalyseCTU-0102

Correction of the exercise 316

(1) La fonction sinh est dérivable (donc continue) et sa dérivée est cosh, qui est une fonction strictement positive. Donc sinh admet une fonction réciproque par le théorème de la bijection.

(2) Pour montrer que $f(y) = \ln(y+\sqrt{y^2+1})$ est la fonction réciproque de sinh nous considérons les fonctions composées $y \mapsto \sinh(f(y))$ et $x \mapsto f(\sinh(x))$. La fonction f est la réciproque de sinh si et seulement si ces deux fonctions sont respectivement l'indentité de y et de x.

$$\sinh(f(y)) = \frac{1}{2}\left(y + \sqrt{y^2+1} - \frac{1}{y+\sqrt{y^2+1}}\right) = \frac{2y^2 + 2y\sqrt{y^2+1}}{2(y+\sqrt{y^2+1})} = y;$$

$$f(\sinh(x)) = \ln\left(\frac{e^x - e^{-x}}{2} + \sqrt{\left(\frac{e^x - e^{-x}}{2}\right)^2 + 1}\right)$$

$$= \ln\left(\frac{e^x - e^{-x}}{2} + \sqrt{\left(\frac{e^x + e^{-x}}{2}\right)^2}\right) = \ln\left(\frac{2e^x}{2}\right) = x.$$

(3) Calculons $f'(y)$ en utilisant son expression analytique

$$f'(y) = \frac{1}{y+\sqrt{y^2+1}}\left(1+\frac{y}{\sqrt{y^2+1}}\right) = \frac{y+\sqrt{y^2+1}}{(y+\sqrt{y^2+1})\sqrt{y^2+1}} = \frac{1}{\sqrt{y^2+1}}.$$

La formule de dérivation d'une fonction composée nous dit que

$$f'(y) = \frac{1}{\cosh(f(y))}$$

ensuite on sait que f est la fonction réciproque de sinh et les propriétés des fonctions sinh et cosh vues dans l'exercice 19 nous disent que $\cosh^2(x) - \sinh^2(x) = 1$. On a donc

$$\cosh(f(y)) = \sqrt{1+\sinh^2(f(y))} = \sqrt{1+y^2}.$$

Exercise 317 exoanalyseCTU-0103

(1) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction paire. Déterminer si les fonctions suivantes sont paires, impaires ou aucun des deux. Justifier chaque affirmation par une petite démonstration ou un contre-exemple.

$$f_1(x) = f(-x), \tag{98.45a}$$
$$f_2(x) = f(x) - 1, \tag{98.45b}$$
$$f_3(x) = \begin{cases} f(x) & \text{si } x > 0, \\ f(-x) & \text{si } x \leqslant 0, \end{cases} \tag{98.45c}$$
$$f_5(x) = \sqrt{(f(x))^2}, \tag{98.45d}$$
$$f_6(x) = |f(x)| + f(x), \tag{98.45e}$$
$$f_7(x) = xf(x), \tag{98.45f}$$
$$f_8(x) = g(f(x)), \text{ où } g \text{ est une fonction impaire.} \tag{98.45g}$$

(2) Refaire la première partie de cet exercice en supposant maintenant que f est impaire.

corranalyseCTU-0103

Correction of the exercise 317

(1) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction paire.

(1a) $f_1(x) = f(-x)$ est paire parce que $f_1(-x) = f(-(-x)) = f(-x) = f_1(x)$.

(1b) $f_2(x) = f(x) - 1$ est paire parce que $f_2(-x) = f(-x) - 1 = f(x) - 1 = f_2(x)$.

(1c) $f_3(x) = \begin{cases} f(x) & \text{si } x > 0, \\ f(-x) & \text{si } x \leqslant 0, \end{cases}$ est paire parce que $f_3(x) = f(|x|)$, donc $f_3(-x) = f(|-x|) = f(|x|) = f_3(x)$

(1d) $f_5(x) = \sqrt{(f(x))^2}$ est paire parce que $f_5(x) = |f(x)|$, donc $f_5(-x) = |f(-x)| = |f(x)| = f_5(x)$.

(1e) $f_6(x) = |f(x)| + f(x)$ est paire comme somme de deux fonctions paires.

(1f) $f_7(x) = xf(x)$ est impaire comme produit d'une fonction paire et une impaire : $f_7(-x) = -xf(-x) = -xf(x) = -f_7(x)$.

(1g) $f_8(x) = g(f(x))$ avec g impaire. Le tout est paire parce que $f_8(-x) = g(f(-x)) = g(f(x)) = f_8(x)$.

(2) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction impaire.

(2a) $f_1(x) = f(-x)$ est impaire et $f_1(-x) = f(-(-x)) = -f(-x) = -f_1(x)$

(2b) $f_2(x) = f(x) - 1$ n'est ni paire ni impaire et $f_2(-x) = f(-x) - 1 = -f(x) - 1$. Elle est paire si et seulement si $f(x) = 0$ pour tout x.

(2c) $f_3(x) = \begin{cases} f(x) & \text{si } x > 0, \\ f(-x) & \text{si } x \leqslant 0, \end{cases}$ est paire parce que $f_3(x) = f(|x|)$. Donc $f_3(-x) = f(|-x|) = f(|x|) = f_3(x)$

(2d) $f_5(x) = \sqrt{(f(x))^2}$ est paire parce que $f_5(x) = |f(x)|$, donc $f_5(-x) = |f(-x)| = |-f(x)| = f_5(x)$

(2e) $f_6(x) = |f(x)| + f(x)$ n'est ni paire ni impaire parce que $f_6(-x) = |f(x)| - f(x) \neq f_6(x)$ et $-f_6(x)$. f_6 est paire si et seulement si $f(x) = 0$ pour tout x.

(2f) $f_7(x) = xf(x)$ est paire parce qu'elle est le produit de deux fonctions impaires : $f_7(-x) = -xf(-x) = xf(x) = f_7(x)$.

(2g) $f_8(x) = g(f(x))$ où g est impaire. Le tout est impair parce que $f_8(-x) = g(-f(x)) = -g(f(x)) = -f_8(x)$

Exercise 318 exosession1-0002

Calculer les intégrales et les primitives suivantes

(1) $\displaystyle\int x^{-4} + x^{1/5}\,dx$;

(2) $\displaystyle\int \ln(x)\,dx$;

(3) $\displaystyle\int x\sin^2(x)\,dx$;

(4) $\displaystyle\int_{-\pi/4}^{\pi/4} \tan(x)\,dx$;

(5) $\displaystyle\int_0^1 \frac{1}{4e^x + e^{-x}}\,dx$. Utiliser le changement de variable $t = e^x$;

(6) $\displaystyle\int_0^{1/2} \sqrt{\frac{\arcsin(x)}{1-x^2}}\,dx$. Utiliser le changement de variable $t = \arcsin(x)$.

corsession1-0002

Correction of the exercise 318

(1) $\displaystyle\int x^{-4} + x^{1/5}\,dx = -\frac{x^{-3}}{3} + \frac{5}{6}x^{6/5} + C$;

(2) $\displaystyle\int \ln(x)\,dx = x\ln(x) - \int x \times \frac{1}{x}\,dx = x\ln(x) - x + C$;

(3)
$$\begin{aligned}\int x\sin^2(x)\,dx &= \int x\frac{1-\cos(2x)}{2}\,dx \\ &= \frac{x^2}{4} - \int \frac{x}{2}\cos(2x)\,dx \\ &= \frac{x^2}{4} - \frac{x}{4}\sin(2x) + \int \frac{\sin(2x)}{4}\,dx \\ &= \frac{x^2}{4} - \frac{x}{4}\sin(2x) - \frac{\cos(2x)}{8} + C\,;\end{aligned}$$

(4) $\displaystyle\int_{-\pi/4}^{\pi/4} \tan(x)\,dx = 0$ car tan est une fonction impaire et $[-\pi/4, \pi/4]$ est symétrique par rapport à l'origine. Si on veut vérifier ce résultat il suffit d'écrire $\tan(x)$ comme $\frac{\sin(x)}{\cos(x)}$ et utiliser la formule pour les primitives des fonctions du type U'/U. On aura alors que les primitives de tan sont $-\ln(|\cos(x)|) + C$.

(5) On veut utiliser le changement de variable $t = e^x$. On a donc $dt = e^x dx$, ce qu'on peut écrire aussi comme $\frac{1}{t}dt = dx$.

$$\begin{aligned}\int_0^1 \frac{1}{4e^x + e^{-x}}\,dx &= \int_{e^0=1}^{e^1=e} \frac{1}{4t + \frac{1}{t}}\frac{1}{t}\,dt \\ &= \int_1^e \frac{1}{4t^2+1}\,dt = \frac{1}{2}\left[\arctan(2t)\right]_{t=1}^{t=e} \\ &= \frac{1}{2}\left(\arctan(2e) - \arctan(2)\right).\end{aligned}$$

(6) On veut utiliser le changement de variable $t = \arcsin(x)$, donc $dt = \frac{1}{\sqrt{1-x^2}}dx$. On peut écrire l'intégrale comme il suit

$$\int_0^{1/2} \sqrt{\frac{\arcsin(x)}{1-x^2}}\,dx = \int_0^{1/2} \sqrt{\arcsin(x)}\frac{1}{\sqrt{1-x^2}}\,dx$$
$$= \int_{\arcsin(0)=0}^{\arcsin(1/2)=\pi/6} \sqrt{t}\,dt = \left[\frac{2}{3}t^{3/2}\right]_0^{\pi/6} = \frac{2}{3}\left(\frac{\pi}{6}\right)^{3/2}.$$

Exercise 319 exosession1-0003

corrsession1-0003

Correction of the exercise 319

<+Corrsession1-0003+>

Exercise 320 exosession1-0004

corrsession1-0004

Correction of the exercise 320

<+Corrsession1-0004+>

Exercise 321 exosession1-0005

Déterminer la solution générale $\mathcal{Y}$ de l'équation différentielle

$$9y'' + y = 4x + 1,$$

ainsi que la solution particulière y_p qui vérifie $y(0) = 2$, $y'(0) = 1$.

corrsession1-0005

Correction of the exercise 321

Nous trouvons facilement par substitution que une solution particulière de l'équation différentielle est $\bar{y}(x) = 4x + 1$.

Pour trouver la solution générale de l'équation homogène associée, $\mathcal{Y}_h$, nous écrivons d'abord le polynôme caractéristique de l'équation, qui est

$$9r^2 + 1,$$

et nous cherchons les racines de $9r^2 + 1 = 0$ dans $\mathbb{C}$. On a $r_{1,2} = \pm\frac{i}{3}$, donc la solution générale est

$$\mathcal{Y}_h = \left\{A\cos\left(\frac{x}{3}\right) + B\sin\left(\frac{x}{3}\right),\ A,\ B \in \mathbb{R}\right\}.$$

La solution générale $\mathcal{Y}$ de l'équation différentielle d'origine est alors

$$\mathcal{Y} = \left\{A\cos\left(\frac{x}{3}\right) + B\sin\left(\frac{x}{3}\right) + \bar{y}(x),\ A,\ B \in \mathbb{R}\right\}.$$

La solution particulière y_p qui vérifie $y(0) = 2$, $y'(0) = 1$ correspond aux valeurs des coefficients $A = 1$ et $B = -9$, en fait

$$A\cos(0) + B\sin(0) + \bar{y}(0) = A + 1$$
$$\text{donc la condition } y(0) = 2 \text{ devient } A + 1 = 2$$
$$-\frac{A}{3}\sin(0) + \frac{B}{3}\cos(0) + \bar{y}'(0) = \frac{B}{3} + 4$$
$$\text{donc la condition } y'(0) = 1 \text{ devient } \frac{B}{3} + 4 = 1.$$

L'unique solution du système

$$\begin{cases} A + 1 = 2, \\ \frac{B}{3} + 4 = 1, \end{cases}$$

est $A = 1$, $B = -9$.

Exercise 322 exosession1-0007

Soit f la solution du problème de Cauchy

$$\begin{cases} y' = -2x + 2xy^2, \\ y(0) = \frac{1}{2}. \end{cases}$$

(1) Pourriez-vous écrire une équation différentielle du deuxième ordre dont f est la solution ?
(2) Déterminer un développement limité de f à l'ordre 2 en 0 en utilisant l'équation différentielle et la formule de Taylor-Young. *Remarque : il n'est pas nécessaire de résoudre l'équation différentielle pour répondre à cette question.*

corrsession1-0007

Correction of the exercise 322

(1) Une équation différentielle du deuxième ordre dont f est la solution est obtenue en dérivant les deux côtés de l'équation par rapport à x

$$y'' = -2 + 2y^2 + 4xyy'.$$

(2) Par la formule de Taylor-Young le développement limité de f à l'ordre 2 en 0 a la forme

$$f(x) = f(0) + f'(0)x + \frac{x^2}{2}f''(0) + x^2\varepsilon(x),$$

oú ε est une fonction dont la limite pour x qui tends vers 0 est nulle.
En utilisant le calcul fait au point précèdent et la condition initiale donnée nous avons

$$f(x) = \frac{1}{2} + 0 + \left(-2 + \frac{1}{2}\right)\frac{x^2}{2} + x^2\varepsilon(x) = \frac{1}{2} - \frac{3}{4}x^2 + x^2\varepsilon(x).$$

Exercise 323 exosession1-0006

Trouver l'unique solution du problème de Cauchy

$$\begin{cases} y' = -2x + 2xy^2, \\ y(0) = \frac{1}{2}. \end{cases}$$

corrsession1-0006

Correction of the exercise 323

L'équation n'est pas linéaire mais à variables séparables. Si $y(x)$ est égale à ± 1 pour un $x\ in\mathbb{R}$ alors $y' = 0$ et la solution est constante. Ce cas ne nous intéresse pas car ces solutions ne satisfont pas la condition initiale imposée. Supposons donc que y soit différent de ± 1. Nous pouvons procéder comme il suit : d'abord

$$y' = 2x(-1 + y^2) \Rightarrow \frac{y'}{(-1 + y^2)} = 2x,$$

ensuite, nous cherchons les primitives (par rapport à x) des deux côtés de l'équation

$$\begin{aligned} &\int \frac{y'}{(-1 + y^2)}\,dx = \int 2x\,dx, \\ &\int \frac{1}{(y-1)(y+1)}\,dy = x^2 + C, \qquad \text{pour } C \in \mathbb{R} \\ &\frac{1}{2}\int \frac{1}{(y-1)} - \frac{1}{(y+1)}\,dy = x^2 + C, \\ &\ln\left(\sqrt{\frac{|y-1|}{|y+1|}}\right) = x^2 + C, \\ &\frac{y-1}{y+1} = Ke^{2x^2}, \qquad \text{pour } K \in \mathbb{R}\backslash\{0\}. \end{aligned}$$

À partir d'ici on peut trouver une expression analytique explicite pour la solution générale de l'équation différentielle

$$\mathcal{Y} = \left\{ y(x) = \frac{1 + Ke^{2x^2}}{1 - Ke^{2x^2}}, \qquad K \in \mathbb{R} \right\} \cup \{y(x) = -1\}.$$

La valeur de K qui correspond à la condition initiale est trouvée par substitution et vaut $K = -1/3$.

Exercise 324 exosession1-0008

(1) Calculer à l'aide de la formule de Taylor-Young, le développement limité à l'ordre 7 de la fonction $f(x) = \cos(2x)$ au voisinage de zéro.

(2) Calculer par une méthode de votre choix le développement limité à l'ordre 7 des fonctions $g(x) = \cos^2(x)$ et $h(x) = \sin^2(x)$ au voisinage de zéro.

(3) Calculer, à l'aide des développements limités, la valeur de la limite

$$\lim x \to 0 \frac{e^x - 1 - x}{\sin^2(x)}.$$

corrsession1-0008

Correction of the exercise 324

(1) La fonction $f(x) = \cos(2x)$ est paire, donc son développement limité au voisinage de zéro ne contient que les termes d'ordre paire

$$f(x) = f(0) + \frac{x^2}{2}f''(0) + \frac{x^4}{4!}f^{(4)}(0) + \frac{x^6}{6!}f^{(6)}(0) + x^7\varepsilon(x).$$

Les dérivées de f sont faciles à calculer : $f^{(n)}(x) = 2^n \cos^{(n)}(2x)$. Par conséquent

$$f(x) = 1 - 2x^2 + \frac{2x^4}{3} - \frac{4x^6}{45} + x^7\varepsilon(x).$$

(2) La fonction $g(x) = \cos^2(x)$ est égale à $\frac{1+\cos(2x)}{2}$, donc nous avons

$$g(x) = \frac{1 + \cos(2x)}{2} = 1 - x^2 + \frac{x^4}{3} - \frac{2x^6}{45} + x^7\varepsilon(x).$$

Analogue ment, $h(x) = \sin^2(x)$ est égale à $1 - \cos(x)$, donc

$$h(x) = 1 - \cos(x) = x^2 - \frac{x^4}{3} + \frac{2x^6}{45} + x^7\varepsilon(x).$$

(3)

$$\lim x \to 0 \frac{e^x - 1 - x}{\sin^2(x)} = \lim x \to 0 \frac{\frac{x^2}{2}}{x^2} = \frac{1}{2}.$$

Exercise 325 exosession2-0001

Soit f la fonction définie par $f(x) = \arccos(\sqrt{1 - x^2})$

(1) Déterminer l'ensemble de définition de f, $\mathcal{D}_f$, et étudier la parité de f.

(2) Calculer $f(0)$, $f(1/2)$, $f(1)$.

(3) Calculer $f'(x)$ pour $0 < x < 1$; on montrera que dans ce cas, $f'(x) = \dfrac{1}{\sqrt{1 - x^2}}$.

(4) En déduire une expression simplifiée de $f(x)$ pour $x \in]0, 1[$. Cette expression est-elle encore valable pour $x = 0$ et pour $x = 1$?

Exercise 326 exosession2-0002

Calculer les intégrales et les primitives suivantes

(1) $\displaystyle\int x(x^{-1/2} + x^{1/5})\, dx$;

(2) $\displaystyle\int \frac{\ln(x)}{x} dx$;

(3) $\displaystyle\int \arcsin(x)\, dx$;

(4) $\displaystyle\int_{\pi/2}^{\pi} \frac{\cos(x) + 1}{\sin(x) + x}\, dx$;

(5) $\displaystyle\int_1^2 \frac{\cos(4+\ln(x))}{x}\,dx$. Utiliser le changement de variable $t = 4 + \ln(x)$;

(6) $\displaystyle\int_0^1 \frac{x^3}{1+x^2}\,dx$. Utiliser le changement de variable $t = x^2 + 1$.

Exercise 327 exosession2-0003

On considère l'équation différentielle linéaire

$$y' + 2y = \frac{2}{3}x^3 - x^2 + 1. \tag{98.46}$$ eq:exo3

(1) Déterminer une fonction polynomiale qui est solution de l'équation différentielle (98.46).

(2) Déterminer la solution générale $\mathcal{Y}$ de l'équation (98.46).

Exercise 328 exosession2-0006

Soit f la solution du problème de Cauchy

$$\begin{cases} xy' - (1+x)y = 0, & \text{pour tout } x \in (0,+\infty), \\ y(1) = \frac{1}{2}. \end{cases}$$

(1) Montrer que f est aussi une solution de l'équation de deuxième ordre

$$xy'' - (2+x)y = 0, \qquad \text{pour tout } x \in (0,+\infty),$$

(2) Soit g la fonction définie par $g(x) = f(x+1)$. Déterminer un développement limité de g à l'ordre 2 en 0 en utilisant l'équation différentielle et la formule de Taylor-Young. *Remarque : il n'est pas nécessaire de résoudre l'équation différentielle pour répondre à cette question.*

Exercise 329 exosession2-0005

Déterminer la solution générale $\mathcal{Y}$ de l'équation différentielle

$$y'' + 2y' + 17y = 4x + 1,$$

ainsi que la solution particulière y_p qui vérifie $y(0) = 2$, $y'(0) = 1$.

Exercise 330 exosession2-0004

Trouver l'unique solution du problème de Cauchy

$$\begin{cases} xy' - (1+x)y = 0, & \text{pour tout } x \in (0,+\infty), \\ y(1) = \frac{1}{2}. \end{cases}$$

Exercise 331 sujetjanvier2016ODE1

(1) Résoudre sur l'intervalle $]-1\,;\,1[$ l'équation différentielle

$$(x-1)y' + y = \frac{1}{\sqrt{1-x^2}}. \tag{98.47}$$ equa1

(2) Déterminer l'unique solution φ de (98.47) telle que $\varphi(0) = 1/\sqrt{2}$.

Exercise 332 sujetjanvier2016ODE2

(1) Montrer que la fonction $f(x) = \dfrac{3e^x}{x^2}$ sur l'intervalle $]0\,;\,+\infty[$ est une solution de l'équation différentielle

$$x^2y'' + 4xy' - (x^2-2)y = 0. \tag{98.48}$$ equa2

(2) On veut résoudre l'équation différentielle (98.48). Pour ce faire il faut effectuer le changement de variable $z = x^2y$.

Exercise 333 sujetjanvier2016ODE3

Nous considérons l'équation différentielle

$$y'' - 3y' - 10y = 2x. \tag{98.49}$$ equa3

(1) Déterminer la solution générale $\mathcal{Y}_H$ de l'équation homogène associée à (98.49).
(2) Trouver une solution particulière, y_P, de la forme $ax + b$, de l'équation (98.49).
(3) En déduire la solution générale de l'équation (98.49).
(4) Existe-t-il une solution y de (98.49) pour laquelle nous avons $\lim_{x\to\infty} y(x) = 0$?
(5) Déterminer l'unique solution de (98.49) qui satisfait les conditions $y(0) = 10$ et $y'(0) = 1$.

Exercise 334 sujetjanvier2016DL

(1) Rappeler le développement limité à l'ordre 6 au voisinage de 0 de la fonction cosinus.
(2) Calculer le développement limité à l'ordre 6 au voisinage de 0 de la fonction

$$f(x) = \frac{1 + ax^2}{1 + bx^2},$$

où a et b sont des réels (leur valeurs exacte n'a pas d'importance pour l'instant).
(3) Déterminer les valeurs de a et b pour lesquelles

$$\lim_{x\to 0} \frac{\cos(x) - f(x)}{x^5} = 0.$$

Exercise 335 exoanalyseCTU-0023

(1) Donner un exemple de fonction bijective de I dans J pour :
(1a) $I = [0, 1]$, $J = [1, 2]$
(1b) $I =]0, 1[$, $J = \mathbb{R}$
(1c) $I = \mathbb{R}^+$, $J = [0, 1[$.
(2) Donner un exemple d'intervalles I et J pour que f soit bijective de I dans J lorsque :
(2a) $f(x) = x^2$
(2b) $f(x) = \ln(x^2)$

corranalyseCTU-0023

Correction of the exercise 335

<+CorranalyseCTU-0023+>

Exercise 336 exoautoanalyseCTU-11

On considère la fonction f définie par $f(x) = \arcsin(2x^2 - 1)$.
(1) Déterminer l'ensemble de définition de f.
(2) Montrer que la fonction f est dérivable sur les intervalles $]-1\ ;\ 0[$ et $]0\ ;\ 1[$, et déterminer l'expression de $f'(x)$.
(3) Montrer que pour tout x de l'intervalle $[0\ ;\ 1[$, $f(x) = 2\arcsin x - \dfrac{\pi}{2}$.
(4) Tracer la courbe représentative de f.

corrautoanalyseCTU-11

Correction of the exercise 336

<+CorrautoanalyseCTU-11+>

Exercise 337 exoautoanalyseCTU-29

Pour tout n de $\mathbb{N}$, on note : $I_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n t \, \mathrm{d}t$.
(1) Calcculer I_0, I_1, I_2.
(2) Établir une relation de récurrence entre I_n et I_{n+2}.
(3) En déduire I_{2p} et I_{2p+1} pour tout $p \in \mathbb{N}$.

corrautoanalyseCTU-29

Correction of the exercise 337

<+CorrautoanalyseCTU-29+>

Exercise 338 exoautoanalyseCTU-30

On définit pour tout $n \in \mathbb{N}$ la fonction f_n sur $\mathbb{R}^+$ par $f_n(x) = x^n e^{-x}$.

Soit $a > 0$, on pose : $I_n(a) = \int_0^a f_n(x)\,\mathrm{d}x$

(1) Déterminer $I_0(a)$.

(2) Montrer que $I_0(a)$ admet une limite lorsque $a \to +\infty$.

On notera cette limite $\int_0^{+\infty} f_0(x)\,\mathrm{d}x$.

(3) Déterminer une relation de récurrence entre $I_n(a)$ et $I_{n+1}(a)$.

(4) Montrer que pour tout n, $I_n(a)$ admet une limite lorsque a tend vers $+\infty$.

On notera cette limite $\int_0^{+\infty} f_n(x)\,\mathrm{d}x$.

(5) Démontrer que : $\int_0^{+\infty} x^n e^{-x}\,\mathrm{d}x = n\,!$

corrautoanalyseCTU-30

Correction of the exercise 338

<+CorrautoanalyseCTU-30+>

Exercise 339 exoautoanalyseCTU-46

Soit la fonction f de la variable réelle x définie par $f(x) = \arctan(x) - \arctan(3x)$.

(1) Déterminer l'ensemble de définition, les variations et les limites aux bords du domaine de la fonction f.

(2) En utilisant le point précédent, montrer que f est une bijection de $I = \left]-\infty, -\frac{1}{\sqrt{3}}\right]$ vers un intervalle que l'on précisera.

(3) Démontrer que pour tout réel x, $f(x) = -\arctan\left(\frac{2x}{1+3x^2}\right)$.

(4) Trouver une expression de la fonction réciproque de f sur I à partir de l'expression donné dans le point précédent.

(5) La fonction f réalise aussi une bijection de $\left]-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right[$ vers $\left]-\frac{\pi}{6}, \frac{\pi}{6}\right[$. Est-il possible d'utiliser entre ces deux intervalles la fonction réciproque trouvée au point précédent ? Motiver votre réponse.

corrautoanalyseCTU-46

Correction of the exercise 339

<+CorrautoanalyseCTU-46+>

Exercise 340 exoanalyseCTU-0008

Déterminer le prolongement par continuité de la fonction

$$f(x) = \frac{\sin(x)}{x} \tag{98.50}$$

en $x = 0$.

Indice : suivre l'exemple 18.240.

corranalyseCTU-0008

Correction of the exercise 340

<+CorranalyseCTU-0008+>

Exercise 341 exoanalyseCTU-0009

Nous considérons les fonctions suivante.

(1) $f_1(x) = \ln(x)$;
(2) $f_2(x) = \ln(x^2)$;
(3) $f_3(x) = \ln(|x|)$;
(4) $f_4(x) = |\ln(x)|$;
(5) $f_5(x) = \ln(x+1)$;
(6) $f_6(x) = \ln(x) + 1$;
(7) $f_7(x) = \sqrt{\ln(x)}$.

(1) Déterminez leur domaines de définition.
(2) Indiquez le graphe correspondant à chacune des fonctions (figure 98.5). Justifiez vos réponses

corranalyseCTU-0009

Correction of the exercise 341

<+CorranalyseCTU-0009+>

Exercise 342 exostarterST-0014modif

Soit $f(x) = \arccos\left(\dfrac{2\sqrt{x}}{1+x}\right)$. Déterminer l'ensemble de définition, le domaine de dérivabilité, et l'expression de $f'(x)$.

Exercise 343 exoautoanalyseCTU-14

Simplifier les expressions suivantes :

(1) $f(x) = \cosh\Big(\ln(x + \sqrt{x^2-1})\Big)$
(2) $f(x) = \sinh\Big(\ln(x + \sqrt{x^2-1})\Big)$

corrautoanalyseCTU-14

Correction of the exercise 343

<+CorrautoanalyseCTU-14+>

Exercise 344 exoanalyseCTU-0005

Pour s'exercer à manipuler les fonctions trigonométrique inverses

(1) Montrer que pour tout $y \in \mathbb{R}$ nous avons

$$\cos\big(\arctan(y)\big) = \frac{1}{\sqrt{1+y^2}} \tag{98.51}$$

et

$$\sin\big(\arctan(y)\big) = \frac{y}{\sqrt{1+y^2}}. \tag{98.52}$$

(2) Montrer que pour tout $y > 0$,

$$\arctan(y) + \arctan(\frac{1}{y}) = \frac{\pi}{2}. \tag{98.53}$$

(3) En déduire que pour tout $y < 0$,

$$\arctan(y) + \arctan(\frac{1}{y}) = -\frac{\pi}{2}. \tag{98.54}$$

corranalyseCTU-0005

Correction of the exercise 344

<+CorranalyseCTU-0005+>

Exercise 345 exoautoanalyseCTU-48

Calculer les intégrales suivantes :

$$I_1 = \int_1^4 \frac{1}{\sqrt{x}}\,\mathrm{d}x \qquad I_2 = \int_1^4 x\ln x\,\mathrm{d}x \qquad I_3 = \int_0^\pi x\sin 3x\,\mathrm{d}x$$

$$I_4 = \int_0^{\frac{\sqrt{3}-1}{3}} \frac{1}{9x^2+6x+2}\,\mathrm{d}x \quad (\text{poser } t = 3x+1) \qquad I_5 = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{1-\cos x}{1+\cos x}\sin x\,\mathrm{d}x \quad (\text{poser } t = \cos x)$$

corrautoanalyseCTU-48

Correction of the exercise 345

<+CorrautoanalyseCTU-48+>

Exercise 346 exoautoanalyseCTU-45

On considère la fonction f de la variable réelle x, définie par $f(x) = \arcsin\left(\dfrac{2x}{1+x^2}\right) - 2\arctan(x)$.

(1) Déterminer l'ensemble de définition de f.

(2) Étudier la parité de f.

(3) Calculer $f(0)$, $f(1)$, $f(\sqrt{3})$ et $\lim\limits_{x\to+\infty} f(x)$.

(4) Justifier que f est dérivable sur chacun des intervalles $]-\infty\,;\,-1[$, $]-1\,;\,1[$ et $]1\,;\,+\infty[$, puis calculer l'expression de $f'(x)$.

(5) En déduire l'expression de $f(x)$.

(6) Établir le tableau de variations de f.

(7) Tracer la courbe représentative de f.

corrautoanalyseCTU-45

Correction of the exercise 346

<+CorrautoanalyseCTU-45+>

Exercise 347 exoautoanalyseCTU-15

Calculer les intégrales suivantes :

$$I_1 = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \tan^2 x\,\mathrm{d}x \qquad I_2 = \int_0^1 (x^3 - 7x + 15)\,\mathrm{d}x \qquad I_3 = \int_{-2}^{-1} \frac{\mathrm{d}t}{t^4} \qquad I_4 = \int_2^8 \sqrt{x}\,\mathrm{d}x$$

$$I_5 = \int_0^1 \frac{\mathrm{d}t}{1+t^2} \qquad I_6 = \int_{-8}^{-2} \frac{\mathrm{d}t}{t} \qquad I_7 = \int_{\frac{1}{2}}^{\frac{\sqrt{3}}{2}} \frac{\mathrm{d}x}{\sqrt{1-x^2}}$$

corrautoanalyseCTU-15

Correction of the exercise 347

<+CorrautoanalyseCTU-15+>

Exercise 348 exoautoanalyseCTU-53

(1) Calculer le développement limité en 0 et à l'ordre 2 des fonctions $x \mapsto \sin(2x)$ et $x \mapsto \ln(1-2x)$.

(2) En déduire le développement limité en 0 et à l'ordre 2 de f définie par $f(x) = \ln(1-2x) - \sin(2x)$.

(3) En déduire les valeurs $f'(0)$ et $f''(0)$.

(4) On note (C_f) la courbe représentative de f.

(4a) Déterminer l'équation de la tangente à (C_f) au point d'abscisse 0.

(4b) Préciser la position relative de (C_f) par rapport à sa tangente.

corrautoanalyseCTU-53

Correction of the exercise 348

<+CorrautoanalyseCTU-53+>

Exercise 349 exoautoanalyseCTU-38

Résoudre sur $\mathbb{R}$ les équations différentielles suivantes :

(1) $y'' + 2y' + y = xe^x$

(2) $y'' + y' - 2y = xe^x$

corrautoanalyseCTU-38

Correction of the exercise 349

<+CorrautoanalyseCTU-38+>

Exercise 350 exoautoanalyseCTU-35

(1) Résoudre sur l'intervalle $]-1\,;\,1[$ l'équation différentielle (E_1) : $(x-1)y' + y = \dfrac{1}{\sqrt{1-x^2}}$.

(2) Résoudre sur l'intervalle $]0\,;\,+\infty[$ l'équation différentielle (E_2) : $xy' - 2y = x^4$.

(3) (3a) Résoudre l'équation différentielle (E_3) : $yy' + 2x = 0$.

(3b) Déterminer la solution φ de (E_3) telle que $\varphi(0) = -2$.

corrautoanalyseCTU-35

Correction of the exercise 350

<+CorrautoanalyseCTU-35+>

Exercise 351 exoautoanalyseCTU-39

(1) (1a) Résoudre sur $\mathbb{R}$ l'équation : $(1 + x^2)y'' + 2xy' = 0$

(1b) Déterminer la solution φ de cette équation telle que $\varphi(1) = 0$ et $\varphi'(0) = 2$.

(2) Résoudre sur l'intervalle $]0\,;\,+\infty[$ l'équation différentielle : $x^2y'' + 4xy' - (x^2 - 2)y = 0$

On pourra poser $z = x^2y$.

corrautoanalyseCTU-39

Correction of the exercise 351

<+CorrautoanalyseCTU-39+>

Exercise 352 exoautoanalyseCTU-41

La fonction f est la solution de l'équation différentielle $y' + 2xy = 2xy^3$ qui prend la valeur $\dfrac{1}{2}$ en 0.

Déterminer un développement limité de f à l'ordre 2 en 0.

corrautoanalyseCTU-41

Correction of the exercise 352

<+CorrautoanalyseCTU-41+>

Exercise 353 exoautoanalyseCTU-49

(1) Résoudre sur $]-1\,;\,+\infty[$ l'équation différentielle (E_1) : $(x + 1)y' - (2x + 3)y = e^{2x}$

(2) Résoudre sur $]-1\,;\,+1[$ l'équation différentielle (E_2) : $y' - y\cos(x) = \dfrac{e^{\sin x}}{\sqrt{1 - x^2}}$

(3) (3a) Résoudre sur $\mathbb{R}$ l'équation différentielle (E_3) : $y'' - 6y' + 13y = 0$.

(3b) Déterminer la solution φ de (E_3) qui vérifie la condition initiale $\varphi(0) = 1$ et $\varphi'(0) = 1$.

(4) Résoudre sur $\mathbb{R}$ l'équation différentielle (E_4) : $y'' - 4y' + 13y = 10\cos(2x) + 25\sin(2x)$.

corrautoanalyseCTU-49

Correction of the exercise 353

<+CorrautoanalyseCTU-49+>

Exercise 354 exoautoanalyseCTU-27

Calculer les intégrales suivantes :

$$I_1 = \int_{-2}^{2} \frac{e^t}{e^t + 3}\,\mathrm{d}t \quad I_2 = \int_0^{\frac{\pi}{2}} \cos^n(x)\sin(x)\,\mathrm{d}x \quad (n \in \mathbb{N}) \quad I_3 = \int_0^1 \frac{1}{\sqrt{4 - x^2}}\,\mathrm{d}x \quad I_4 = \int_{-2}^{2} \sinh x\,\mathrm{d}x$$

$$I_5 = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin t}{1 + \cos^2 t}\,\mathrm{d}t \qquad I_6 = \int_2^4 \frac{1}{x^2 - 1}\,\mathrm{d}x \qquad I_7 = \int_0^2 \frac{t + 1}{t^2 + 4}\,\mathrm{d}t \qquad I_8 = \int_0^{\frac{\pi}{4}} x\cos(2x)\,\mathrm{d}x$$

corrautoanalyseCTU-27

Correction of the exercise 354

<+CorrautoanalyseCTU-27+>

Exercise 355 exomazhe-0009

Completer le tableau suivant.

Primitive $\int f(x)\,dx$	Function $f(x)$	Dérivée $f'(x)$
$\sin(x)+C$	...	...
...	$\sin(3x)$	...
...	...	$-\frac{1}{x^2}$
...	$\ln(x)$	...
...	...	$(\alpha+1)x^\alpha$
$\frac{e^{5x}}{5}+C$	...	...
...	$\tan(x)=\frac{\sin(x)}{\cos(x)}$	...
...	$\arcsin(x)$	...

(98.55)

corrmazhe-0009

Correction of the exercise 355

<+Corrmazhe-0009+>

Exercise 356 exoautoanalyseCTU-19

Calculer les intégrales et les primitives suivantes

(1) $K_1=\displaystyle\int_{\frac{\pi}{4}}^{2\pi}\frac{1}{x^{1/4}}+\cos(3x)\,\mathrm{d}x$;

(2) $K_2=\displaystyle\int x^3(\ln(x)-1)\,\mathrm{d}x$;

(3) $K_3=\displaystyle\int\frac{x^2+15}{x(x+1)(x+5)}\,\mathrm{d}x$;

(4) $K_4=\displaystyle\int\frac{x}{\sqrt{x^2-1}}\arctan\sqrt{x^2-1}\,\mathrm{d}x$;

(5) $K_5=\displaystyle\int_{\frac{\pi}{4}}^{\pi}1+\cos(x)+\cos^2(x)\,\mathrm{d}x$;

(6) $K_6=\displaystyle\int_{\frac{\pi}{2}}^{\pi}\sin(x)+\sin^3(x)\,\mathrm{d}x$.

corrautoanalyseCTU-19

Correction of the exercise 356

<+CorrautoanalyseCTU-19+>

Exercise 357 exoautoanalyseCTU-22

Calculer les intégrales suivantes :

$J_1=\displaystyle\int_2^3\frac{\mathrm{d}x}{x+5}$ $\qquad$ $J_2=\displaystyle\int_2^3\frac{\mathrm{d}x}{3x+5}$ $\qquad$ $J_3=\displaystyle\int_2^3\frac{\mathrm{d}x}{(2x+5)^2}$

$$J_4 = \int_0^1 \frac{\mathrm{d}x}{(4x+1)^4} \qquad J_5 = \int_0^2 \frac{x}{x^2+5}\,\mathrm{d}x \qquad J_6 = \int_0^{\sqrt{5}} \frac{\mathrm{d}x}{x^2+5}$$

corrautoanalyseCTU-22

Correction of the exercise 357

<+CorrautoanalyseCTU-22+>

Exercise 358 exoautoanalyseCTU-23

Calculer les intégrales suivantes en utilisant le changement de variable proposé :

$$I_1 = \int_0^1 \frac{e^x}{e^{2x}+4}\,\mathrm{d}x \quad (\text{prendre } t = e^x) \qquad I_2 = \int_1^2 x\sqrt{1+x^2}\,\mathrm{d}x \quad (\text{prendre } t = x^2)$$

$$I_3 = \int_0^1 \frac{\mathrm{d}x}{(1+x^2)^{\frac{3}{2}}} \quad (\text{prendre } x = \tan t) \qquad I_4 = \int_0^a \sqrt{a^2-x^2}\,\mathrm{d}x \quad (a \in [0\,;+\infty[, \text{ poser } x = a\sin t)$$

corrautoanalyseCTU-23

Correction of the exercise 358

<+CorrautoanalyseCTU-23+>

Exercise 359 exostarterST-0007

Indiquez le graphe correspondant à chacune des fonctions suivantes (figure 98.6). Justifiez vos réponses

(1) $\cos(x)$;
(2) $\cos(x+\pi/2)$;
(3) $\cos(e^x)$;
(4) $\cos(x)+1$;
(5) $\cos(4x)$;
(6) $|\cos(x)|$;
(7) $\sqrt{\cos(x)}$.

corrstarterST-0007

Correction of the exercise 359

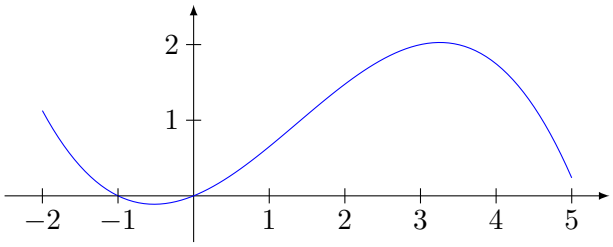
<+CorrstarterST-0007+>

98.15 Autres

Exercise 360 exoanalyseCTU-0001

Quelques questions à propos d'intervalles et de croissance.

(1) Donner un intervalle ouvert contenant le nombre -4.
(2) Donner un exemple de fonction décroissante mais non strictement décroissante.
(3) Supposons que la fonction f dont le graphe est donné ci-dessous soit dérivable. Dresser le tableau de signe de la dérivée de f.



corranalyseCTU-0001

Correction of the exercise 360

<+CorranalyseCTU-0001+>

Exercise 361 exoanalyseCTU-0003

Reprendre le dessin de la figure 18.3 et y dessiner également la fonction sinus de façon à mettre en évidence la symétrie.

corranalyseCTU-0003

Correction of the exercise 361

<+CorranalyseCTU-0003+>

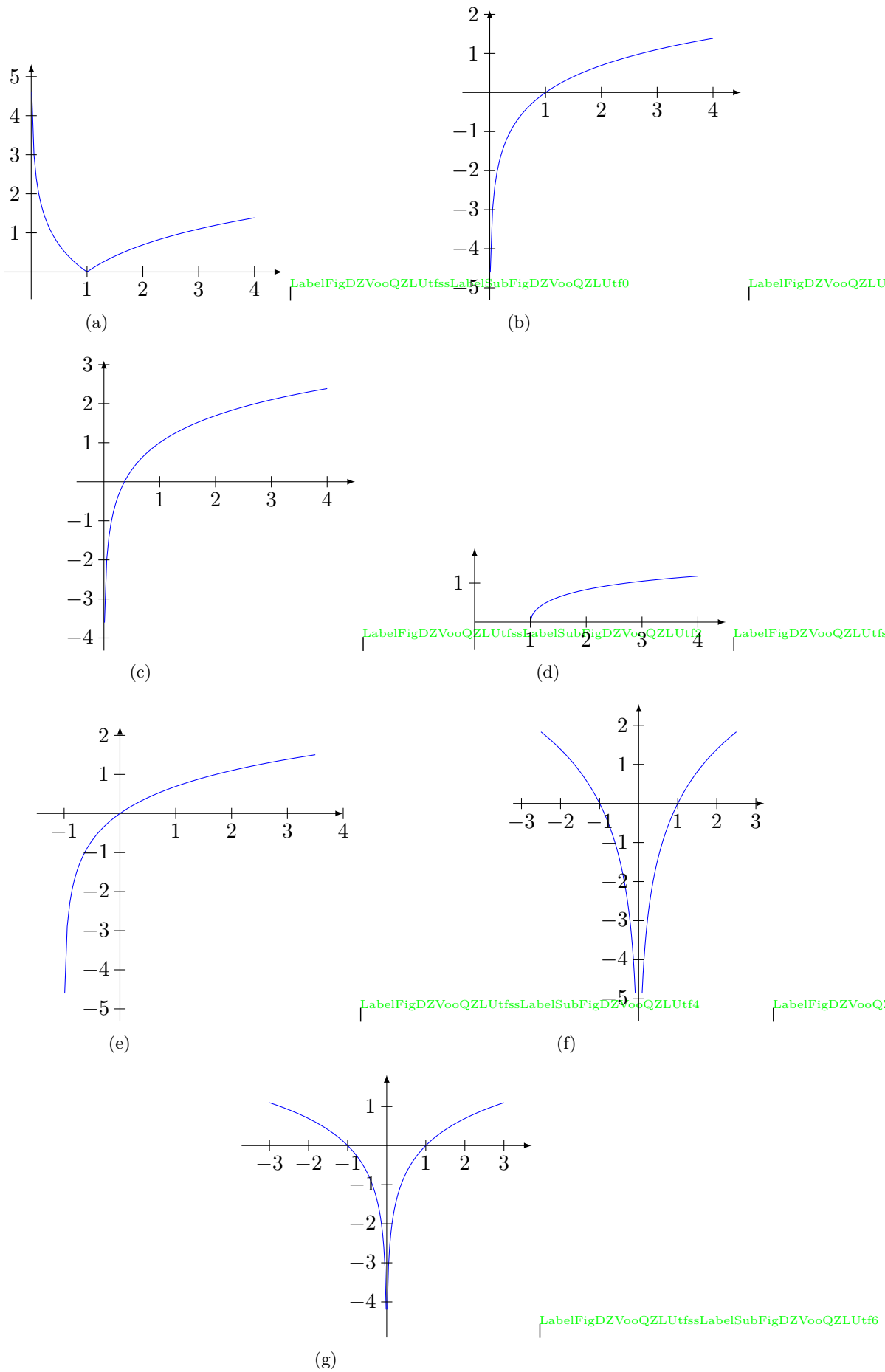
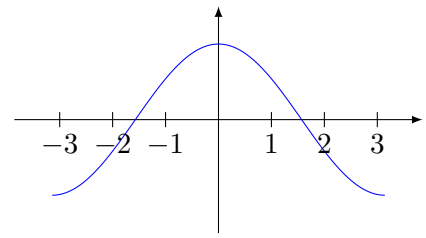
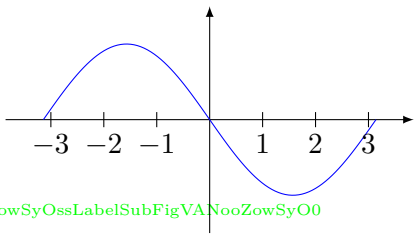


FIGURE 98.5: Les graphes à considérer

LabelFigDZVooQZLUtf

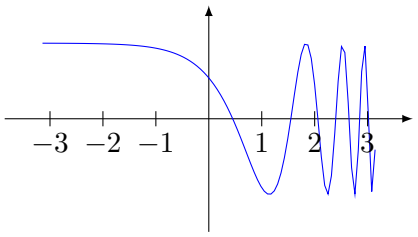


(a)

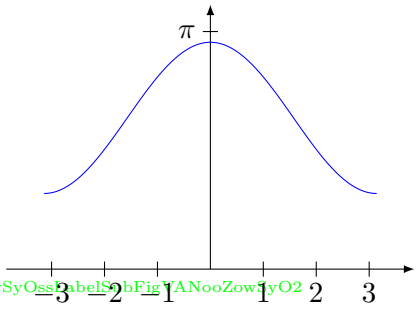


LabelFigVANooZowSyOssLabelSubFigVANooZowSyO0 LabelFigVANooZowSyC

(b)

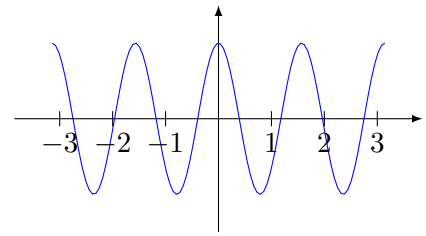


(c)

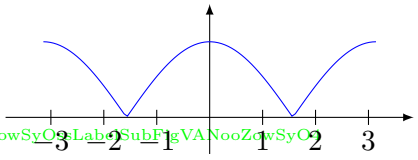


LabelFigVANooZowSyOssLabelSubFigVANooZowSyO2 LabelFigVANooZowSy

(d)

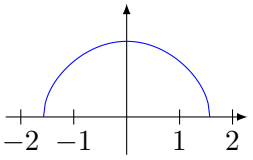


(e)



LabelFigVANooZowSyOssLabelSubFigVANooZowSyO4 LabelFigVANooZowSyC

(f)



LabelFigVANooZowSyOssLabelSubFigVANooZowSyO6

(g)

FIGURE 98.6: Les graphes à considérer de la question 359. LabelFigVANooZowSyO

Chapitre 99

GNU Free Documentation License

Version 1.3, 3 November 2008
Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc.

http://fsf.org/

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom : to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation : a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals ; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical

connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "**publisher**" means any person or entity that distributes copies of the Document to the public.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties : any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts : Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version :

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another ; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of

it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/ .

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM : How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page :

Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation ; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled “GNU Free Documentation License”.

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the “with . . . Texts.” line with this :

with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Bibliographie

[1] Laurent Claessens. Mes inventions personnelles, lire avec prudence et merci de me dire si c'est correct. Attention : je ne prétends pas avoir découvert ce résultat ; un résultat qui a cette référence signifie seulement que j'ai ré-inventé une partie de l'énoncé ou de la preuve. Je serais étonné qu'il y ait ici des résultats vraiment présents nulle part ailleurs.

[2] Mathologer. Numberphile v. Math : the truth about 1+2+3+...=-1/12 13 jan. 2018. URL https://www.youtube.com/watch?v=YuIIjLr6vUA.

[3] Numberphile. Does -1/12 Protect Us From Infinity ? - Numberphile 16 février 2024. URL https://www.youtube.com/watch?v=beakj767uG4.

[4] Wikipédia. Multi-indice — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Multi-indice&oldid=155192272.

[5] jperon. Tri de l'index tenant compte des accents octobre 2022. URL https://github.com/LaurentClaessens/mazhe/issues/162.

[6] A. Casamayou and N Cohen and G. Connan and T. Dumont and L. Fousse and F. Maltey and M. Meulien and M. Mezzarobba and C. Pernet and M. N. Thiéry and Zimmermann P.. Calcul mathématique avec Sage URL https://www.sagemath.org/sagebook/french.html.

[7] Wikipédia. Axiome du choix — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Axiome_du_choix&oldid=116151439.

[8] Wikipédia. Relation binaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Relation_binaire&oldid=183967099.

[9] Wikipédia. Élément maximal — Wikipédia, l'encyclopédie libre URL https://fr.wikipedia.org/w/index.php?title=Élément_maximal&oldid=180618994.

[10] Wikipédia. Axiome du choix — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Axiome_du_choix&oldid=178683927.

[11] O. Teytaud and C. Antonini and J.-B. Bardet and J.-F. Quint and M. De Crisenoy and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau. Les mathématiques pour l'Agrégation URL http://les.mathematiques.free.fr/pdf/agreg.zip.

[12] Wikipédia. Lemme de Zorn — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Lemme_de_Zorn&oldid=164572297.

[13] Patrice Goyer. Communication personnelle

[14] Wikipédia. Algèbre des parties d'un ensemble — Wikipédia, l'encyclopédie libre 19 juillet 2023. URL https://fr.wikipedia.org/w/index.php?title=Alg%C3%A8bre_des_parties_d%27un_ensemble&oldid=205030352.

[15] ChromoMaxwell. Re :

Exo

Associativité de la différence symétrique URL https://forums.futura-sciences.com/mathematiques-superieur/101396-exo-associativite-de-difference-symetrique.html.

[16] Patrice Tauvel. Cours d'algèbre

[17] Wikipédia. Morphisme d'anneaux — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Morphisme_d%27anneaux&oldid=129892066.

[18] Christine Laurent-Thiébaut. Axiomatique des nombres URL http://ljk.imag.fr/membres/Bernard.Ycart/mel/ax/ax.pdf.

[19] Laurent Schwartz. Analyse I

[20] Wikipédia. Loi de composition interne — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Loi_de_composition_interne&oldid=174035538.

[21] Wikipédia. Équipotence — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quipotence&oldid=157942537.

[22] Jérôme Von Buhren. Cardinal d'un ensemble et théorème de Cantor-Bernstein URL http://vonbuhren.free.fr/Prepa/Problemes/Cardinal_ensemble_theoreme_Cantor_Bernstein.pdf.

[23] Wikipédia. Ensemble infini — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Ensemble_infini&oldid=182905419.

[24] . Infinite set — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Infinite_set&oldid=1018913868.

[25] Wikipédia. Axiome du choix dénombrable — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Axiome_du_choix_d%C3%A9nombrable&oldid=165350212.

[26] Ilie Grigorescu. Lecture 4 - finite and infinite sets URL http://www.math.miami.edu/~igrigore/teaching/mth533/lecture4_mth433.pdf.

[27] Wikipédia. Ensemble dénombrable — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Ensemble_d%C3%A9nombrable&oldid=150759981.

[28] Alexander Bogomolny. Cantor-Bernstein-Schroeder theorem URL https://www.cut-the-knot.org/WhatIs/Infinity/Bernstein.shtml.

[29] Wikipédia. Théorème de comparabilité cardinale — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_comparabilit%C3%A9_cardinale&oldid=158933871.

[30] Ramez Labib-Sami. Théorie de la cardinalité -2. Ensembles infini non dénombrables. URL http://www.logique.jussieu.fr/~sami/Ensembles_L3/5_Cardinaux_non_denombrables.pdf.

[31] Wikipédia. Théorème de Cantor — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Cantor&oldid=166939268.

[32] Pedro Tamaroff. Prove that if A is an infinite set then Ax2 is equipotent to A URL https://math.stackexchange.com/questions/1041731/prove-that-if-a-is-an-infinite-set-then-a-times-2-is-equipotent-to-a.

[33] cdrcprds. Ensembles équipotents : ajouter ou soustraire des cardinalités URL https://github.com/LaurentClaessens/mazhe/issues/146#issuecomment-588513588.

[34] MikeMathMan. Cardinality of the Cartesian Product of Two Equinumerous Infinite Sets URL https://math.stackexchange.com/questions/1383755/cardinality-of-the-cartesian-product-of-two-equinumerous-infinite-sets.

[35] Wikipédia. Groupe ordonné 22 aout 2022. URL https://fr.wikipedia.org/wiki/Groupe_ordonn%C3%A9.

[36] Peter Hendrikus Kropholler. GROUPS RINGS AND FIELDS URL http://www.maths.gla.ac.uk/~phk/ulearning.pdf.

[37] Wikipédia. Sous-groupe caractéristique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Sous-groupe_caract%C3%A9ristique&oldid=178218738.

[38] Wikipédia. Primalité dans un anneau — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Primalit%C3%A9_dans_un_anneau&oldid=145407116.

[39] Wikipédia. Primalité dans un anneau — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Primalit%C3%A9_dans_un_anneau&oldid=127919489.

[40] Wikipédia. Diviseur de zéro — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Diviseur_de_z%C3%A9ro&oldid=143858271.

[41] Pierre-Emmanuel Chaput. ANNEAUX EUCLIDIENS, PRINCIPAUX, FACTORIELS URL http://chaput.perso.math.cnrs.fr/documents/ch6-anneau-eucl-princ-fact.pdf.

[42] Fabio Lucchini. gcd(a S)= a gcd(S) (integral domain or ufd) 28 juillet 2023. URL https://math.stackexchange.com/q/4743750.

[43] Patrice Goyer and Jean-Claude Cid. Communication personnelle 28 septembre 2023.

[44] Wikipédia. Corps commutatif — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_commutatif&oldid=152289565.

[45] Patrick Polo. Idéaux premiers et maximaux, localisation, théorème des zéros de Hilbert URL https://webusers.imj-prg.fr/~patrick.polo/M1Galois/ATG05ch3.pdf.

[46] wikipédia. Théorème de Krull — Wikipédia, l'encyclopédie libre 2022. URL https://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Krull&oldid=192263296.

[47] Wikiversité. Arithmétique/Divisibilité et congruences dans Z — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Arithm%C3%A9tique/Divisibilit%C3%A9_et_congruences_dans_Z&oldid=650617. En ligne ; accédé le 25-mai-2017

[48] Wikipédia. Lemme d'Euclide — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Lemme_d%27Euclide&oldid=179832418.

[49] Michel Cretin. Anneaux principaux et factoriels URL http://math.univ-lyon1.fr/~cretin/OralAlgebre/annprincfact.pdf.

[50] Wikipédia. Idéal premier — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Id%C3%A9al_premier&oldid=125964642.

[51] Mongi Amorri. Théorème de Bezout, théorème de Gauss URL http://lux.lyceefrancais-brasilia.net/documents/cours/ts/bezout_gauss.pdf.

[52] Wikipédia. Anneau euclidien — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Anneau_euclidien&oldid=126145344.

[53] Wikiversité. Anneau (mathématiques)/Exercices/Exercices — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Anneau_(math%C3%A9matiques)/Exercices/Exercices&oldid=676468. En ligne ; accédé le 26-octobre-2017

[54] Anonyme. Théorie des groupes URL http://ekladata.com/CKoLuTDRT9S2sx5BPX9DkCDAcdo.pdf.

[55] Wikipédia. Groupe symétrique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Groupe_sym%C3%A9trique&oldid=146656747.

[56] Jean Delcourt. Groupes-permutations URL http://delcourt.u-cergy.fr/StrucAlg/chap_5.pdf.

[57] . Décomposition d'une permutation en un produit de cycles à supports disjoints URL http://math.mad.free.fr/wordpress/wp-content/uploads/cycle.pdf.

[58] Wikiversité. Théorie des groupes/Groupes symétriques finis — Wikiversité 2021. URL https://fr.wikiversity.org/wiki/Théorie_des_groupes/Groupes_symétriques_finis.

[59] François Combes. Algèbre et géométrie

[60] Wikipédia. Permutation — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Permutation&oldid=89805226.

[61] Wikipédia. Taquin — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Taquin&oldid=190177267.

[62] numberphile. Why is this 15-Puzzle Impossible ? URL https://www.youtube.com/watch?v=YI1WqYKHi78.

[63] Sonner. How can I prove that interior product obeys a graded Leibniz rule ? 17 may 2016. URL https://math.stackexchange.com/q/1789506.

[64] Jean-Louis Rouget. Compléments d'algèbre URL https://www.maths-france.fr/MathSpe/Cours/01-structures.pdf.

[65] Wikipédia. Module sur un anneau — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Module_sur_un_anneau&oldid=141582982.

[66] Patrick Polo. Algèbre et théorie de Galois URL https://webusers.imj-prg.fr/~patrick.polo/M1Galois/ATG07chI-III.pdf.

[67] David Harari. Algèbre 1 - Anneaux et modules URL https://www.math.u-psud.fr/~harari/exposes/rings.pdf.

[68] Wikipédia. Algèbre sur un corps — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Alg%C3%A8bre_sur_un_corps&oldid=98992040.

[69] Wikipédia. Caractéristique d'un anneau — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Caract%C3%A9ristique_d%27un_anneau&oldid=136316021.

[70] Wikipédia. Polynôme formel — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Polyn%C3%B4me_formel&oldid=185737458.

[71] Patrick Polo. Groupes et polynômes symétriques, résolution d'équations URL https://webusers.imj-prg.fr/~patrick.polo/M1Galois/ATG07chIXcorr.pdf.

[72] Wikipédia. Corps des fractions — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_des_fractions&oldid=130674526.

[73] Wikipédia. Corps ordonné — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_ordonn%C3%A9&oldid=125185412.

[74] . Nth root algorithm — Wikipedia, The Free Encyclopedia

[75] Wikipédia. Méthode de Héron — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=M%C3%A9thode_de_H%C3%A9ron&oldid=181923161.

[76] Frédéric Paulin. Topologie, analyse et calcul différentiel URL http://www.math-question-center.com/publications-pdf/cours_d'analyse+topologie_et_calcul_differntiel_Frederic_Paulin.pdf.

[77] Wikipédia. Construction des nombres réels — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Construction_des_nombres_r%C3%A9els&oldid=145402164.

[78] Wikipédia. Nombre positif — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Nombre_positif&oldid=146043202.

[79] . Least-upper-bound property — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Least-upper-bound_property&oldid=893001075.

[80] Wikipédia. Valeur absolue — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Valeur_absolue&oldid=181957023.

[81] Wikipédia. Corps valué — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_valu%C3%A9&oldid=172545490.

[82] Wikipédia. Partie entière et partie fractionnaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Partie_enti%C3%A8re_et_partie_fractionnaire&oldid=181404258.

[83] Foys. Re : Ressources pour la topologie générale URL http://www.les-mathematiques.net/phorum/read.php?11,2093686,2194966#msg-2194966.

[84] Jean Luc W. groupe-quotient et création de nouveaux sous-groupes URL https://fr.wikipedia.org/wiki/Discussion:Groupe_quotient.

[85] Pierre Lissy. Théorème de Jordan-Hölder URL http://www.ljll.math.upmc.fr/~lissy/Agreg/developpements/JordanHolder.pdf.

[86] Wikipedia contributors. Solvable group — Wikipedia, The Free Encyclopedia 2022. URL https://en.wikipedia.org/w/index.php?title=Solvable_group&oldid=1104386698.

[87] Wikipédia. Groupe résoluble — Wikipédia, l'encyclopédie libre 2022. URL https://fr.wikipedia.org/w/index.php?title=Groupe_r%C3%A9soluble&oldid=194829259.

[88] Muriel Galley. Groupes résolubles URL http://matthieu.gendulphe.com/Galley.pdf.

[89] Wikipédia. Action de groupe (mathématiques) — Wikipédia, l'encyclopédie libre 2022. URL https://fr.wikipedia.org/w/index.php?title=Action_de_groupe_(math%C3%A9matiques)&oldid=193789140.

[90] Fabrice Castel. Groupes finis URL http://agreg-maths.univ-rennes1.fr/documentation/docs/Groupes_finis.pdf.

[91] Sylvain Gervais and François Nicoleau. Feuille 1 : Notions sur les groupes URL https://www.math.sciences.univ-nantes.fr/~agreg/DocumentAgreg/L3-GroupeGeometrie.pdf.

[92] . Groupe cyclique URL https://moodle.utc.fr/pluginfile.php/49237/mod_resource/content/0/ChapII_groupes/MT10_ChapII6.pdf.

[93] cdrcprds. Questions (9), (15) et (21) d'algèbre URL https://github.com/LaurentClaessens/mazhe/issues/60.

[94] Gilles Costantini. pgcd et ppcm dans $\mathbb{Z}$, théorème de Bezout, applications URL http://gilles.costantini.pagesperso-orange.fr/agreg_fichiers/bezout.pdf.

[95] Congrugences dans $\mathbb{Z}$, anneaux $\mathbb{Z}/n\mathbb{Z}$ URL http://math.univ-lyon1.fr/capes/IMG/pdf/new.congruence.pdf.

[96] Wikipédia. Lemme d'Euclide — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Lemme_d%27Euclide&oldid=112046247.

[97] Wikipédia. Théorème fondamental de l'arithmétique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_fondamental_de_l%27arithm%C3%A9tique&oldid=110036012.

[98] Wikipédia. Théorème de Cauchy (groupes) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Cauchy_(groupes)&oldid=140167231.

[99] . 101. Groupes monogènes, groupes cycliques. Exemples URL http://unemainlavelautre.net/concours_examens/agregation_interne_2019_maths_lecons_et_exercices/agregation_interne_oral_101_groupes_monogenes_groupes_cycliques_exemples.pdf.

[100] Alexei Pantchichkine. Algèbre2 URL http://www-fourier.ujf-grenoble.fr/~panchish/05ensl.pdf.

[101] Niels Borne. Fiche numéro 2 : morphismes, sous-groupe distingué, quotient URL http://math.univ-lille1.fr/~borne/Enseignement/TD2_M308.pdf.

[102] Wikipédia. Lemme d'Euclide 22 juillet 2023. URL https://fr.wikipedia.org/wiki/Lemme_d%27Euclide#Unicit%C3%A9_de_la_forme_irr%C3%A9ductible_d'une_fraction.

[103] Georges Skandalis. Algèbre générale et algèbre linéaire URL http://www.math.univ-paris-diderot.fr/_media/formations/prepa/agreginterne/polycopiealgebre.pdf.

[104] Arnaud Girand. Développements pour l'agrégation externe de mathématiques URL http://perso.univ-rennes1.fr/arnaud.girand/pdf/dvp_agreg/dvp.pdf.

[105] Wikipédia. Formule du binôme de Newton — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Formule_du_bin%C3%B4me_de_Newton&oldid=144161875.

[106] Wikipédia. Identité remarquable — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Identit%C3%A9_remarquable&oldid=178609166.

[107] G. B. Folland. Higher-Order Derivatives and Taylor's Formula in Several Variables URL https://sites.math.washington.edu/~folland/Math425/taylor2.pdf.

[108] Serge Mehl. Preuves diverses de résultats énoncés au fil des pages URL http://serge.mehl.free.fr/base/preuves_diverses.html.

[109] Chritian Blanchet. Propriétés arithmétiques des anneaux commutatifs intègres URL https://webusers.imj-prg.fr/~christian.blanchet/enseignement/2009-10/ch2_anneaux_integres.pdf.

[110] Anonyme. sans titre URL http://www.ac-grenoble.fr/champo/IMG/arithmetique.pdf.

[111] Didier Piau and Bernard Ycart. Lemme de Gauss et décomposition en facteurs premiers URL https://ljk.imag.fr/membres/Bernard.Ycart/mel/ar/node6.html.

[112] Wikipédia. Idéal premier — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Id%C3%A9al_premier&oldid=145977206.

[113] Danny-Jack Mercier. Anneaux factoriels URL http://megamaths.perso.neuf.fr/cours/ari/cann0001.pdf.

[114] Sébastien Pellerin. Développements d'algèbre pour l'oral de l'agrégation URL http://pellerin.xyz/doc/agreg/algebre.pdf.

[115] Wikipédia. Triplet pythagoricien — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Triplet_pythagoricien&oldid=114628048.

[116] Wikipédia. Primalité dans un anneau — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Primalit%C3%A9_dans_un_anneau&oldid=127919489.

[117] Wikipédia. Polynôme primitif 26 juillet 2023. URL https://fr.wikipedia.org/w/index.php?title=Polyn%C3%B4me_primitif&oldid=181739006.

[118] Pierre-Emmanuel Chaput. ANNEAUX DE POLYNÔMES URL http://chaput.perso.math.cnrs.fr/documents/ch7-anneau-de-polynomes.pdf.

[119] Wikipédia. Factorisation des polynômes 30 juillet 2023. URL https://fr.wikipedia.org/wiki/Factorisation_des_polyn%C3%B4mes.

[120] Wikipédia. Racine de l'unité — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Racine_de_l%27unit%C3%A9&oldid=178477716.

[121] Olivier Debarre. Anneaux et corps, préparation à l'agrégation externe URL https://www.math.ens.fr/~debarre/anneauxcorps.pdf.

[122] Wikipédia. Lemme de Gauss (polynômes) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Lemme_de_Gauss_(polyn%C3%B4mes)&oldid=112056131.

[123] Alain Vigne. Communication personnelle

[124] Keith Conrad. Cyclicity of $(Z/(p))^\times$ URL https://kconrad.math.uconn.edu/blurbs/grouptheory/cyclicmodp.pdf.

[125] Wikipédia. Espace vectoriel — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_vectoriel&oldid=143370323.

[126] Espaces vectoriels de dimension infinie URL http://www.les-mathematiques.net/b/e/e/node8.php.

[127] . Tensor product — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Tensor_product&oldid=842778935.

[128] Marie-Pierre Lebaux. Rappels sur les applications linéaires URL https://perso.univ-rennes1.fr/marie-pierre.lebaud/agint/ecrit/algebre-lineaire/applications-lineaires/V-appli-lin.pdf.

[129] Jean-Louis Rouget. Dimension d'un espace vectoriel URL https://www.maths-france.fr/MathSup/Cours/22-dimensions.pdf.

[130] Daniel Ferrand. Étendre le corps URL http://webusers.imj-prg.fr/~daniel.ferrand/ExtCorps.pdf.

[131] Ycart Bernard and Luc Rozoy. Déterminants URL https://ljk.imag.fr/membres/Bernard.Ycart/mel/de/de.pdf.

[132] G. Donald Allen. Matrices and linear algebra URL http://www.math.tamu.edu/~dallen/m640_03c/lectures/chapter2.pdf.

[133] Wikipédia. Mineur (algèbre linéaire) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Mineur_(alg%C3%A8bre_lin%C3%A9aire)&oldid=133960288.

[134] Khalid Koufany. Formes linéaires et dualité URL http://www.iecl.univ-lorraine.fr/~Khalid.Koufany/Alg-Bil/Chap1-Formes-Lineaires.pdf.

[135] . Direct sum of modules — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Direct_sum_of_modules&oldid=835636392.

[136] Laurent Claessens. existence of a real structure on a complex vector space URL https://math.stackexchange.com/q/4034093.

[137] Wikipédia. Somme directe — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Somme_directe&oldid=182206467.

[138] John Douglas Moore. Othogonal complement URL http://web.math.ucsb.edu/~moore/orthogonalcomplements.pdf.

[139] . Real structure — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Real_structure&oldid=990936192.

[140] Wikipédia. Théorème de Cauchy (groupes) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Cauchy_(groupes)&oldid=122684298.

[141] Wikiversité. Groupe (mathématiques)/Exercices/Premiers résultats sur les groupes simples — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Groupe_(math%C3%A9matiques)/Exercices/Premiers_r%C3%A9sultats_sur_les_groupes_simples&oldid=597195. En ligne ; accédé le 31-mai-2017

[142] Wikiversité. Groupe (mathématiques)/Automorphismes d'un groupe cyclique — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Groupe_(math%C3%A9matiques)/Automorphismes_d%27un_groupe_cyclique&oldid=595608. En ligne ; accédé le 3-juin-2017

[143] Structure des groupes d'ordre pq URL http://agregmaths.free.fr/doc/docs_nicolas/developpementAlgebre/grouped'ordrepq.pdf.

[144] Wikipédia. Inverse modulaire — Wikipédia, l'encyclopédie libre 2023. URL https://fr.wikipedia.org/w/index.php?title=Inverse_modulaire&oldid=205601911.

[145] Wikiversité. Exercice : Nombres premiers et fonctions arithmétiques 2022. URL https://fr.wikiversity.org/wiki/Introduction_%C3%A0_la_th%C3%A9orie_des_nombres/Exercices/Nombres_premiers_et_fonctions_arithm%C3%A9tiques#Exercice_1-6.

[146] minutephysics. How quantum computing breaks encryption, Shor's algorithm explained URL https://www.youtube.com/watch?v=lvTqbM5Dq4Q.

[147] Wikipédia. Théorème d'Euler (arithmétique) — Wikipédia, l'encyclopédie libre 5 juin 2023. URL https://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_d%27Euler_(arithm%C3%A9tique)&oldid=204927084.

[148] wikiversité. Introduction à la théorie des nombres : Nombres premiers et fonctions arithmétiques 22 juillet 2023. URL https://fr.wikiversity.org/wiki/Introduction_%C3%A0_la_th%C3%A9orie_des_nombres/Nombres_premiers_et_fonctions_arithm%C3%A9tiques.

[149] Olivier Debarre. TD4 : produit semi-direct URL http://www.math.ens.fr/~debarre/TDC4.pdf.

[150] Christophe Mourougane. Théorie des groupes et géométrie URL http://perso.univ-rennes1.fr/christophe.mourougane/enseignements/2009-10/THGG/poly.groupes.pdf.

[151] Michel Emsalem and Pierre Dèbes. Théorème de Sylow URL http://exo7.emath.fr/ficpdf/fic00023.pdf.

[152] N. G. J. Pagnon. Le groupe symétrique et le groupe alterné URL http://amatheux.com/IMG/pdf/groupe_symetrique.pdf.

[153] Automorphisme de R et continuité URL http://forums.futura-sciences.com/mathematiques-superieur/261451-automorphisme-de-r-continuite.html.

[154] Poirot. Automorphismes et sous-corps premier Avril 2021. URL https://les-mathematiques.net/vanilla/index.php?p=discussion/comment/2213126#Comment_2213126.

[155] Wikipédia. Nombre premier de Sophie Germain — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Nombre_premier_de_Sophie_Germain&oldid=162746200.

[156] Développement Théorème de Sophie Germain. URL https://www.agreg-maths.fr/uploads/versions/930/Dev_SophieGermain_BN.pdf.

[157] Antoine Barrier. Théorème de Sophie Germain URL https://www.agreg-maths.fr/uploads/versions/2235/abarrier_dev_thm_Sophie_Germain.pdf.

[158] Florian Lemonnier. Théorème de Sophie Germain URL https://flemonnier.pagesperso-orange.fr/agregation/developpements/Sophie_Germain.pdf.

[159] Wikipédia. Monoïde — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Mono%C3%AFde&oldid=148198215.

[160] Emmanuel Vieillard Baron. Le lemme chinois en termes de congruences URL https://www.les-mathematiques.net/b/a/d/node10.php.

[161] Sylvain Duchet. Congrugence dans $\mathbb{Z}$, Anneaux $\mathbb{Z}/n\mathbb{Z}$, applications URL http://epsilon.2000.free.fr/Csup/congruences.pdf.

[162] Quentin Guyot. Communication personnelle octobre 2022.

[163] Wikipédia. Anneau factoriel 30 juillet 2023. URL https://fr.wikipedia.org/wiki/Anneau_factoriel#Anneaux_des_polyn%C3%B4mes.

[164] Patrick Polo. Extension de corps, caractéristique, corps de rupture, corps de décomposition, clôtures algébriques URL https://webusers.imj-prg.fr/~patrick.polo/M1Galois/ATG07chVII.pdf.

[165] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Le théorème de Wedderburn URL http://www.les-mathematiques.net/d/a/w/node5.php.

[166] cdrcprds. Lemme 5.43 sur les pgcd de polynômes URL https://github.com/LaurentClaessens/mazhe/issues/52#issuecomment-333251728.

[167] Geneviève Bretenoux and Marie-Thérèse Hogbé and Jacques Queyrut and Dominique Labarsouque and Bernadtte Munos and Catherine Pannier. Théorème de Gauss et polynômes irréductibles 2014. URL https://uel.unisciel.fr/mathematiques/polynomes1/polynomes1_ch02/co/apprendre_ch2_3_05.html.

[168] Wikipédia. Extension de corps — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Extension_de_corps&oldid=133373097.

[169] Wikipédia. Extension de corps — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Extension_de_corps&oldid=133373097.

[170] David Harari. Algèbre 1 - notions de théorie des corps URL https://www.math.u-psud.fr/~harari/exposes/corps.pdf.

[171] Mortajine Abdellatif. Les extensions de corps URL http://www.iecn.u-nancy.fr/~mortajin/chap3-L3-S5.pdf.

[172] Wikipédia. Extension algébrique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Extension_alg%C3%A9brique&oldid=148162879.

[173] Olivier Debarre. Algèbre 2 URL https://www.math.ens.fr/~debarre/Algebre2.pdf.

[174] Wikipédia. Clôture algébrique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Cl%C3%B4ture_alg%C3%A9brique&oldid=159475673.

[175] Herscovich Estanislao. Existence et « unicité » d'une clôture algébrique URL https://www-fourier.ujf-grenoble.fr/~eherscov/MAT4111/ThmSteinitz.pdf.

[176] jeanas. Constante de Weiner 19 septembre 2024. URL https://linuxfr.org/nodes/136626/comments/1969399.

[177] Wikipédia. Corps de décomposition — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_de_d%C3%A9composition&oldid=139881008.

[178] Vincent Beck. Extensions séparables 5 octobre 2010. URL https://www.idpoisson.fr/beck/wp-content/uploads/sites/16/2018/06/cours-ext-sep-poly.pdf.

[179] Patrick Polo. Extensions normales, séparables, galoisiennes. Corps fini. URL http://www.math.jussieu.fr/~polo/M1/ATG07chVIII.pdf.

[180] mathématiques and jobherzt . polynôme irréductible Février 2021. URL https://les-mathematiques.net/vanilla/index.php?p=/discussion/653420/polynome-irreductible.

[181] Olivier Dodane. Le théorème de zéros de Hilbert URL http://www.math.ens.fr/~debarre/nullstellensatz.pdf.

[182] Patrick Polo. Polynômes symétriques et résolutions d'équations URL http://www.math.jussieu.fr/~polo/M1/ATGch9.pdf.

[183] Vincent Beck and Jérôme Malick and Gabriel Peyré. Sur C tout est connexe ! URL http://objagr.gforge.inria.fr/documents/files/connexite-polynome.pdf.

[184] El Hage. Équation générale de degré n URL http://les.mathematiques.free.fr/pdf/gal10.pdf.

[185] Florian Morel Chevillet. Résolubilité par radicaux des équations algébriques URL http://matthieu.gendulphe.com/MorelChevillet.pdf.

[186] Wikipédia. Espace topologique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_topologique&oldid=164610481.

[187] Wikipédia. Base (topologie) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Base_(topologie)&oldid=177248194.

[188] Wikipédia. Base (topologie) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Base_(topologie)&oldid=140151583.

[189] Wikipédia. Voisinage (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Voisinage_(math%C3%A9matiques)&oldid=171863485.

[190] Erwan Aubry. URL https://math.unice.fr/~eaubry/Enseignement/L3/rappelstopo.pdf.

[191] Wikipédia. Prébase — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Pr%C3%A9base&oldid=162334814.

[192] Wikipédia. Topologie induite — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Topologie_induite&oldid=172447185.

[193] Wikiversité. Topologie/Sous-espace topologique URL https://fr.wikibooks.org/wiki/Topologie/Sous-espace_topologique.

[194] . Filtration (mathematics) — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Filtration_(mathematics).

[195] Dirk Hundertmark. Lecture notes functional analysis URL https://www.math.kit.edu/iana1/lehre/funcana2012w/media/fa-lecturenotes.pdf.

[196] Wikiversité. General Topology/Bases 5 décembre 2023. URL https://en.wikibooks.org/w/index.php?title=General_Topology/Bases&oldid=4319527.

[197] Wikipédia. Topologie quotient 2022. URL https://fr.wikipedia.org/wiki/Topologie_quotient.

[198] Jonny Evans. 3.02 Quotient topology : continuous maps URL http://www.homepages.ucl.ac.uk/~ucahjde/tg/html/quot02.html.

[199] Wikipédia. Espace séparé — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_s%C3%A9par%C3%A9&oldid=97013072.

[200] Wikipédia. Équivalent — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quivalent&oldid=149344743.

[201] Wikiversité. Espaces vectoriels normés : Connexité 5 aout 2023. URL https://fr.wikiversity.org/wiki/Espaces_vectoriels_norm%C3%A9s/Connexit%C3%A9.

[202] Wikipédia. Connexité (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Connexit%C3%A9_(math%C3%A9matiques)&oldid=164885091.

[203] Wikipédia. Connexité URL https://fr.wikiversity.org/wiki/Topologie_g%C3%A9n%C3%A9rale/Connexit%C3%A9.

[204] François De Marçay. Théorème de Jordan URL https://www.imo.universite-paris-saclay.fr/~joel.merker/Enseignement/Analyse-Complexe/jordan-courbe.pdf.

[205] Vincent Le Guiec and Sacha Quayle. Le théorème de Jordan 29 avril 2021. URL https://minerve.ens-rennes.fr/images/Le_Théorème_de_Jordan_S.Quayle_V.Le_Gruiec..pdf.

[206] Emmanuel Vieillard Baron. Sous-espaces compacts URL http://www.les-mathematiques.net/a/t/c/node5.php.

[207] Louis Thanhoffer de Volcsey. Are intersections of compact sets compact ? URL https://www.quora.com/Are-intersections-of-compact-sets-compact?share=1.

[208] Wikipédia. Compactifié d'Alexandrov — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Compactifi%C3%A9_d%27Alexandrov&oldid=146811167.

[209] . Finite intersection property — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Finite_intersection_property&oldid=1076876990.

[210] . Limit of a function — Wikipedia, The Free Encyclopedia

[211] Académie française. ambiguïté URL https://www.dictionnaire-academie.fr/article/A9A1402.

[212] Wikipédia. Connexité (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Connexit%C3%A9_(math%C3%A9matiques)&oldid=123744973.

[213] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Connexité URL http://www.les-mathematiques.net/a/a/b/node23.php.

[214] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Espaces métriques compacts URL http://www.les-mathematiques.net/a/a/b/node22.php.

[215] Jean-Marc Decauwert. Convexité URL https://membres-ljk.imag.fr/Bernard.Ycart/mel/ga/node7.html.

[216] Guillaume Carlier. Analyse fonctionnelle URL https://www.ceremade.dauphine.fr/~carlier/poly2010.pdf.

[217] Éric Brunelle. Norme matricielle URL http://www.dms.umontreal.ca/~math1600/6Supplement/Normematricielle-1.pdf.

[218] Aurelien Monteillet. Corrigés d'exercices pour le TD 3 URL http://aurelien.monteillet.com/teaching/TopologieL3/corrige-topo3.pdf.

[219] Raz Kupferman. Topological vector spaces URL http://www.ma.huji.ac.il/~razk/iWeb/My_Site/Teaching_files/TVS.pdf.

[220] Wikipédia. Sous-espace supplémentaire — Wikipédia, l'encyclopédie libre 2022. URL https://fr.wikipedia.org/w/index.php?title=Sous-espace_suppl%C3%A9mentaire&oldid=193382370.

[221] Anne. Supplémentaire topologique (Banach ?) 21 aout 2022. URL https://fr.wikipedia.org/wiki/Discussion:Sous-espace_suppl%C3%A9mentaire#Suppl%C3%A9mentaire_topologique_(Banach_?).

[222] Kr Dpk and Gerd . Topological direct sum of vector spaces 20 aout 2022. URL https://math.stackexchange.com/q/4515661.

[223] proofwiki. Limit of Composite Function 2020. URL https://proofwiki.org/wiki/Limit_of_Composite_Function.

[224] Wikipédia. Anneau topologique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Anneau_topologique&oldid=142162744.

[225] Maria Infusino. Topological vector spaces Summer Semester 2017. URL http://www.math.uni-konstanz.de/~infusino/TVS-SS17/Note2017(July29).pdf.

[226] user408858. Neighborhood of 0 is absorbent in topological vector spaces 2023-06-19. URL https://math.stackexchange.com/q/4721001.

[227] Gabriel Nagy. Topological Vector Spaces I : Basic Theory URL https://www.math.ksu.edu/~nagy/func-an-2007-2008/top-vs-1.pdf.

[228] EDUARD A. NIGSCH. LOCALLY CONVEX SPACES 2017. URL https://www.asc.tuwien.ac.at/~enigsch/lehre/lcs.pdf.

[229] Emmanuel Vieillard-Baron. Application continue URL http://les.mathematiques.free.fr/a/t/g/node10.html.

[230] Nitrogen. If X is compact and $T : XoY$ continuous and bijective, show that T is homeomorphism. URL https://math.stackexchange.com/q/1785556.

[231] David Wilkins. Normed vector spaces and functional analysis URL http://www.maths.tcd.ie/~dwilkins/Courses/212/212PtIII.pdf.

[232] Wikipédia . Espace séquentiel 2022. URL https://fr.wikipedia.org/wiki/Espace_séquentiel.

[233] Huiqiang Jiang. Functional analysis URL http://www.math.pitt.edu/~hqjiang/2303/functional.pdf.

[234] Nicolas Bourbaki. General topology 2 chapters 5 - 10 URL https://books.google.be/books?id=bQwhdmL6IjUC. Oui je sais c'est une honte de citer Bourbaki en anglais en pointant vers une version partielle disponible sur Googlebooks. Mais c'est surtout une honte que ce livre ne soit pas disponible gratuitement en version électronique.

[235] Wikipédia. Algèbre de Banach — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Alg%C3%A8bre_de_Banach&oldid=181944654.

[236] Ron Freiwald. Chapter IV Completeness and Compactness URL http://www.math.wustl.edu/~freiwald/ch4.pdf.

[237] Richard Timoney. Dual Spaces and the Hahn-Banach Theorem URL https://www.maths.tcd.ie/~richardt/321/321-ch3.pdf.

[238] Wikipédia. Suite de Cauchy — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Suite_de_Cauchy&oldid=98046451.

[239] Norbert. First theorem in Topological vector spaces. URL https://math.stackexchange.com/q/155244.

[240] Wikipédia. Continuité — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Continuit%C3%A9&oldid=96548980.

[241] Wikipédia. Théorème des fermés emboités — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_des_ferm%C3%A9s_emboit%C3%A9s&oldid=85987870.

[242] Étienne Matheron. Topologie, analyse fonctionnelle URL https://matheron.perso.math.cnrs.fr/enseignement_fichiers/topanafonc.pdf.

[243] François Golse. Distributions, analyse de Fourier, équations aux dérivées partielles URL http://www.cmls.polytechnique.fr/perso/golse/MAT431-10/POLY431.pdf.

[244] Foys. Déduire la connexité par arcs URL http://www.les-mathematiques.net/phorum/read.php?14,1477834,1487858#msg-1487858.

[245] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Le théorème de Tykhonov URL http://www.les-mathematiques.net/a/a/b/node20.php.

[246] . Equicontinuity — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Equicontinuity&oldid=1078023606.

[247] Wikipédia. Équicontinuité — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quicontinuit%C3%A9&oldid=137910699.

[248] Wikipédia. Continuité uniforme — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Continuit%C3%A9_uniforme&oldid=138238518.

[249] Zied Ammari. Analyse fonctionnelle : Pré-requis URL http://perso.univ-rennes1.fr/zied.ammari/other/pdf/chapitre1.pdf.

[250] Jean Saint Raymond. Semi-normes URL http://www.math.jussieu.fr/~raymond/preprints/seminormes.pdf.

[251] Wikipédia. Hahn–Banach theorem 3 juin 2023. URL https://en.wikipedia.org/wiki/Hahn%E2%80%93Banach_theorem.

[252] Gabriel Nagy. Convexity "Warm-up" II :The Minkowski Functional Fall 07 - Spring 08. URL https://www.math.ksu.edu/~nagy/func-an-2007-2008/conv-w-2.pdf.

[253] Gariel Nagy. Locally Convex Vector Spaces III : The Metric Point of View 2007-2008. URL https://www.math.ksu.edu/~nagy/func-an-2007-2008/lcvs-3.pdf.

[254] geetha290krm. seminorm continuous at zero 2023-12-30. URL https://math.stackexchange.com/q/4835942.

[255] psl2Z. Topology from seminorms dec 03, 2023. URL https://math.stackexchange.com/q/4819595.

[256] Stéphane Mischler. Semi-norme et introduction aux evtlcs URL https://www.ceremade.dauphine.fr/~mischler/Enseignements/AFAENS/Chap1evtlcs.pdf.

[257] Wikipédia. Semi-norme — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Semi-norme&oldid=89981857.

[258] Wikipédia. Norme (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Norme_(math%C3%A9matiques)&oldid=186698313.

[259] Jean Saint Raymond. Semi-normes URL https://webusers.imj-prg.fr/~jean.saint-raymond/preprints/seminormes.pdf.

[260] Wikipédia. Espace vectoriel normé 29-mars-2022. URL http://fr.wikipedia.org/w/index.php?title=Espace_vectoriel_norm%C3%A9&oldid=192365488.

[261] Wikipédia. Théorème de Baire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Baire&oldid=95077621.

[262] Zuoqin Wang. Lecture 16 : multilinear algebra URL http://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds/Notes/Lec16.pdf.

[263] Sha Vuklia. Why are multilinear maps continuous ? 2023-05-18. URL https://math.stackexchange.com/questions/1490776/why-are-multilinear-maps-continuous.

[264] Gilles Dubois. Applications multilinéaires continues 11 september 2024. URL http://gilles.dubois12.free.fr/?page_id=935.

[265] Wikipédia. Forme antisymétrique 3 avril 2023. URL https://fr.wikipedia.org/wiki/Forme_antisym%C3%A9trique.

[266] Emmanuel Vieillard Baron. Déterminant d'une matrice, d'une application linéaire URL http://www.les-mathematiques.net/b/e/d/node5.php.

[267] mdp. Determinant of linear transformation 7 oct 2014. URL https://math.stackexchange.com/q/962435.

[268] Robert Rolland. Le théorème de Müntz-Szász URL http://megamaths.perso.neuf.fr/rr/fichexo_201.pdf.

[269] Wikipédia. Matrice de Sylvester — Wikipédia, l'encyclopédie libre URL https://fr.wikipedia.org/w/index.php?title=Matrice_de_Sylvester&oldid=191741692.

[270] Wikipédia. Résultant — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=R%C3%A9sultant&oldid=90109864.

[271] Emmanuel Pedon. Cours de géométrie affine et euclidienne pour la licence de Mathématiques URL http://pedon.perso.math.cnrs.fr/fichiers/enseignement/CoursGeoLicence.pdf.

[272] user642796. Accumulation points of uncountable sets URL https://math.stackexchange.com/q/310122.

[273] afk. transposée et valeurs propres Juillet 2012. URL https://les-mathematiques.net/vanilla/index.php?p=/discussion/comment/765054/#Comment_765054.

[274] Stéphane Lamy. DEUX DÉVELOPPEMENTS POUR LA LEÇONS 122 URL https://www.math.univ-toulouse.fr/~slamy/teaching/agreg/lecon_122.pdf.

[275] Matthieu Romagny. Endomorphismes cycliques URL https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/endom_cycliques.pdf.

[276] Wikipédia. Polynôme caractéristique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Polyn%C3%B4me_caract%C3%A9ristique&oldid=188987150.

[277] Jean-Etienne Rombaldi. Polynômes d'endomorphismes en dimension finie. Applications URL http://www-fourier.ujf-grenoble.fr/~rombaldi/AgregInterne/Oral1/110.pdf.

[278] Wikipédia. Application bilinéaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Application_bilin%C3%A9aire&oldid=144184234.

[279] Wikipédia. Forme quadratique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Forme_quadratique&oldid=172697192.

[280] Michel Coste. Formes bilinéaires symétriques, formes quadratiques URL https://perso.univ-rennes1.fr/michel.coste/Bil.pdf.

[281] Maria Carrisoza. Algèbre IV : algèbre géométrique Printemps 2012. URL http://licence-math.univ-lyon1.fr/lib/exe/fetch.php?media=p12:algiv:chapitre1.pdf.

[282] Stéphane Lamy. Algèbre bilinéaire URL https://www.math.univ-toulouse.fr/~slamy/teaching/agreg/cours_bilineaire_dec_2019.pdf.

[283] Emmanuel Vieillard Baron. Formes quadratiques réelles URL http://www.les-mathematiques.net/b/c/h/node12.php.

[284] mathphysics. Isotropie 5 aout 2023. URL http://mathphysics.fr/Notes/Isotropie.php.

[285] Transformation laissant invariante une forme quadratique URL http://www.ilemaths.net/forum-sujet-500814.html.

[286] Qmechanic. Interval preserving transformations are linear in special relativity URL https://physics.stackexchange.com/questions/12664/interval-preserving-transformations-are-linear-in-special-relativity.

[287] Wikipédia. Loi d'inertie de Sylvester — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Loi_d%27inertie_de_Sylvester&oldid=156712630.

[288] Wikipédia. Hermitien — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Hermitien&oldid=124476255.

[289] Gilles Dubois. Dualité URL http://gilles.dubois10.free.fr/algebre_lineaire/dualite.html.

[290] Marc Sage. Dualité en dimension finie URL http://www.normalesup.org/~sage/Enseignement/Cours/DualDimFinie.pdf.

[291] Donald Allen. Unitary matrices URL https://www.math.tamu.edu/~dallen/m640_03c/lectures/chapter4.pdf.

[292] La l'Irlandais. De la réduction des endomorphismes URL http://tanopah.jo.free.fr/epilogues/reduction.pdf.

[293] Wikipédia. Réduction de Gauss — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=R%C3%A9duction_de_Gauss&oldid=155723411.

[294] Marie-Pierre Lebaud. Formes quadratiques URL https://perso.univ-rennes1.fr/marie-pierre.lebaud/agint/ecrit/algebre-lineaire/formes-quadratiques/V-formes-quadratiques.pdf.

[295] Matthieu Romagny. Endomorphismes semi-simples URL https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/endom_semi_simples.pdf.

[296] Robert Rollan. Produit tensorile d'espaces vectoriels URL http://robert.rolland.acrypta.com/telechargements/algebre/tensor.pdf.

[297] Marc Sage. Algèbre multilinéaire URL http://www.normalesup.org/~sage/Cours/ProdTens.pdf.

[298] Ministère nationale. Rapport de jury de concours – agrégation de mathématiques, concours externe URL http://agreg.org/Rapports/rapport2011provisoire.pdf.

[299] Hervé Carrieu and Maurice Fadel and Etienne Fieux and Patrice Lassère and Frédéric Rodriguez. Autour des matrices de Frobenius ou compagnon URL http://www.math.univ-toulouse.fr/~lassere/pdf/vfcomp.pdf.

[300] Grégory Vial. Autour du théorème des invariants de similitude URL http://w3.bretagne.ens-cachan.fr/math/people/gregory.vial/files/cplts/ivs.pdf.

[301] Arnaud Moncet. Invariants de similitude URL http://blogperso.univ-rennes1.fr/arnaud.moncet/public/IVS.pdf.

[302] G. Donald Allen. Lectures on linear algebras URL http://www.math.tamu.edu/~dallen/m640_03c/readings.htm.

[303] Commutant d'un endomorphisme URL http://minerve.bretagne.ens-cachan.fr/images/Commutant.pdf.

[304] Bernard Alken. Dans un commentaire de « Dimension du commutant d'une matrice » URL http://www.mathoman.com/index.php/1538-dimension-du-commutant-d-une-matrice.

[305] Wikipédia. Forme linéaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Forme_lin%C3%A9aire&oldid=154647982.

[306] Wikipédia. Théorème de Burnside (problème de 1902) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Burnside_(probl%C3%A8me_de_1902)&oldid=98510166.

[307] Wikipédia. Théorème de Borel-Lebesgue — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Borel-Lebesgue&oldid=96560796.

[308] Pete Clar. The Instructor's Guide to Real Induction 5 aout 2012. URL https://arxiv.org/abs/1208.0973.

[309] Lavrov Misha. Every path homotopic to a Piecewise linear path (Hatcher 1.1.4) 9 juin 2018. URL https://math.stackexchange.com/q/2813173.

[310] Wikipédia. Théorème des suites adjacentes — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_des_suites_adjacentes&oldid=148318442.

[311] Mathieu Mansuy. Autour des séries alternées URL http://www.mathieu-mansuy.fr/pdf/ECS2-complément1.pdf.

[312] LucaMac. A topological space X has the Intermediate Value Property iff it is Connected 2018-08-27. URL https://math.stackexchange.com/q/2896105.

[313] M. Quéffelec and A. Bodin. Topologie générale URL http://exo7.emath.fr/ficpdf/fic00037.pdf.

[314] Ambigraphe. Demande de référence URL https://fr.wikipedia.org/wiki/Discussion:Ouvert_(topologie).

[315] Wikipédia. Connexité par arcs — Wikipédia, l'encyclopédie libre 3 février 2023. URL https://fr.wikipedia.org/w/index.php?title=Connexit%C3%A9_par_arcs&oldid=198763722.

[316] Marc Lelarge. Le théorème de Perron-Frobenius 15 octobre 2014. URL https://www.di.ens.fr/~lelarge/ani/cours/cours1.pdf.

[317] Wikipédia. Équation du second degré — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quation_du_second_degr%C3%A9&oldid=179274356.

[318] Christian Squarcini. Dégénérescence URL http://christian-squarcini.pagesperso-orange.fr/AgregInterne/Algebrelineaire/2_2.pdf.

[319] Jean-Marc Decauwert. Orthogonalité, bases orthonormées URL https://ljk.imag.fr/membres/Bernard.Ycart/mel/ge/node5.html.

[320] Daniel Li. Espaces de Hilbert URL http://www.editions-ellipses.fr/product_info.php?products_id=9387.

[321] user127.0.0.1. Parallelogram law in normed vector space without an inner product. URL https://math.stackexchange.com/questions/641077/parallelogram-law-in-normed-vector-space-without-an-inner-product.

[322] Frank Jones. Chapter 7 : cross product URL http://www.owlnet.rice.edu/~fjones/chap7.pdf.

[323] Maxime Pouvreau. Pseudo-réduction simultanée URL http://minerve.bretagne.ens-cachan.fr/images/Pseudored.pdf.

[324] Jean-François Burnal. Normes Lp URL http://jf.burnol.free.fr/agregnormeslp.pdf.

[325] Wikipédia. Norme équivalente — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Norme_%C3%A9quivalente&oldid=192429578.

[326] William F. Trench. Introduction to real analysis URL http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF.

[327] user8128. All norms of $\mathbb{R}^n$ are equivalent 2018. URL https://math.stackexchange.com/q/2890032.

[328] Wikipédia. Norme d'opérateur — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Norme_d%27op%C3%A9rateur&oldid=125963366.

[329] Wikipedia. Spectral radius — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Spectral_radius&oldid=758072091.

[330] Christian Squarcini. Applications linéaires URL http://christian-squarcini.pagesperso-orange.fr/AgregInterne/Topo/4_3.pdf.

[331] Gilles Costantini. Espaces vectoriels de dimension infinie, normes usuelles, équivalence de normes URL http://gilles.costantini.pagesperso-orange.fr/agreg_fichiers/evn.pdf.

[332] Gilles Dubois. séries commutativement convergentes URL http://gilles.dubois10.free.fr/analyse_reelle/seriescomconv.html.

[333] Wikipédia. Théorème de réarrangement de Riemann 2024. URL https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_de_r%C3%A9arrangement_de_Riemann.

[334] Florian Lemonnier. Théorème de réarrangement de Riemann URL https://perso.eleves.ens-rennes.fr/people/florian.lemonnier/agregation/developpements/Riemann.pdf.

[335] H. Jerome Keisler. Elementaty calculus, an infinitesimal approach URL http://www.math.wisc.edu/~keisler/keislercalc-810.pdf.

[336] Wikipédia. Suite arithmético-géométrique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Suite_arithm%C3%A9tico-g%C3%A9om%C3%A9trique&oldid=88168841.

[337] Prime.mover. Sum of Sequence of Products of Consecutive Reciprocals URL https://proofwiki.org/wiki/Sum_of_Sequence_of_Products_of_Consecutive_Reciprocals.

[338] Daniel Daigle. Cardinalité URL http://mysite.science.uottawa.ca/asavag2/mat2762/cardinal.pdf.

[339] Alan D. Sokal. A really simple elementary proof of the uniform boundedness theorem URL https://arxiv.org/pdf/1005.1585.pdf.

[340] Wikipédia. Théorème de Banach-Steinhaus — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Banach-Steinhaus&oldid=92377069.

[341] . Strong operator topology — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Strong_operator_topology&oldid=849989429.

[342] Sylvain. Théorème de l'application ouverte URL https://agreg-maths.fr/uploads/versions/59/thm_app_ouverte_scourte.pdf.

[343] Wikipédia. Théorème de Banach-Schauder — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Banach-Schauder&oldid=179020717.

[344] Nicolas. Théorèmes de l'application ouverte et du graphe fermé URL http://agregmaths.free.fr/doc/docs_nicolas/developpementAnalyse/Theoremedel'applicationouverteetdugrapheferme.pdf.

[345] . Tensor product — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Tensor_product&oldid=846586202.

[346] Martin Brandenburg. unicity of tensor product decomposition October 15, 2023. URL https://math.stackexchange.com/q/4787198.

[347] Arjeh Cohen. Tensor product URL https://www.win.tue.nl/~amc/ow/lba/lba3.pdf.

[348] Mike Hill. Lecture 16 : tensor products URL https://www.math.ucla.edu/~mikehill/Teaching/Math5651/Lecture16.pdf.

[349] Mike Hill. LECTURE 16 : TENSOR PRODUCTS URL https://www.math.ucla.edu/~mikehill/Teaching/Math5651/Lecture16.pdf.

[350] Martin Brandenburg. every bilinear map is a tensor product if the dimensions matche 2023-oct-21. URL https://math.stackexchange.com/q/4791265.

[351] Frédéric Hélein. Calcul différentiel extérieur et géométrie 10 novembre 2016. URL https://webusers.imj-prg.fr/~frederic.helein/caldiffext.pdf.

[352] Davide Giraudo. Every linear mapping on a finite dimensional space is continuous URL https://math.stackexchange.com/q/113008.

[353] Livio Flaminio. Éléments de géométrie différentielle URL http://www-gat.univ-lille1.fr/~flaminio/M403/2009-2010/cours_geo_top.pdf.

[354] Espaces métriques et espaces normés URL http://www.les-mathematiques.net/a/a/b/node4.php.

[355] Jerome E. Marsden and Tudor S. Ratiu. Manifolds, tensor analysis and applications URL http://www.springer.com/gp/book/9780387967905.

[356] F. Laudenbach. Calcul différentiel et intégral URL http://books.google.fr/books?id=Ws9A7ZoRJNEC.

[357] Pierre Lairez. Théorème de dérivation d'une limite Communication personnelle

[358] Marie-Claure David and Frédéric Haglund and Daniel Perrin. Géométrie affine URL http://webens.math.u-psud.fr/~mcld/GAEL/Gael/bar.pdf.

[359] Marie-Claude David and Frédéric Haglund and Daniel Perrin. Géometrie affine. Document de travail pour la préparation au CAPES. Deuxième partit : barycentre URL http://omphale.math.u-psud.fr/~mcld/GAEL/Gael/bar.pdf.

[360] Marie-Claude David and Frédéric Haglund and Daniel Perrin. Géometrie affine. Document de travail pour la préparation au CAPES. Troisième partie : convexité URL http://omphale.math.u-psud.fr/~mcld/GAEL/Gael/conv.pdf.

[361] Jean-Marc Decauwert. Géométrie affine URL http://ljk.imag.fr/membres/Bernard.Ycart/mel/ga/ga.pdf.

[362] Wikiversité. Barycentre/Théorème de l'associativité du barycentre — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Barycentre/Th%C3%A9or%C3%A8me_de_l%27associativit%C3%A9_du_barycentre&oldid=582092. En ligne ; accédé le 16-janvier-2017

[363] Wikipédia. Espace topologique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_topologique&oldid=152161195.

[364] Daniel Perrin. À propos de la définition de la limite d'une fonction en un point URL https://www.math.u-psud.fr/~perrin/CAPES/analyse/fonctions/definitiondelimite.pdf.

[365] monde. Annonce Frido 2019 URL https://linuxfr.org/news/le-frido-livre-collaboratif-de-mathematique-de-niveau-agregation-et-un-peu-plus.

[366] monde. Annonce Frido 2018 URL https://linuxfr.org/news/le-frido-2018-livre-libre-de-mathematique-pour-l-agregation.

[367] Usual limit from filter URL https://math.stackexchange.com/questions/3364039/usual-limit-from-filter.

[368] Foys. Limites épointées URL http://www.les-mathematiques.net/phorum/read.php?18,2195094,2195240.

[369] Donatien Alphonse François Sade. L'histoire de Juliette URL https://fr.wikisource.org/wiki/L%E2%80%99histoire_de_Juliette/cinqui%C3%A8me_partie.

[370] Wikipédia. Intégrale de Kurzweil-Henstock — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Int%C3%A9grale_de_Kurzweil-Henstock&oldid=178545707.

[371] Foys. Re : Limites épointées URL http://www.les-mathematiques.net/phorum/read.php?18,2195094,2255758#msg-2255758.

[372] Wikipédia. Droite réelle achevée — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Droite_r%C3%A9elle_achev%C3%A9e&oldid=144800302.

[373] Ud779. Limits of functions and left hand right hand limit URL https://math.stackexchange.com/questions/1418673/limits-of-functions-and-left-hand-right-hand-limit.

[374] . Leçon nr 63 URL https://www.capes-de-maths.com/lecons/lecon63.pdf.

[375] Wikiversité. Fonctions d'une variable réelle : Continuité 4 aout 2023. URL https://fr.wikiversity.org/wiki/Fonctions_d%27une_variable_r%C3%A9elle/Continuit%C3%A9#Th%C3%A9or%C3%A8mes_sur_les_fonctions_continues.

[376] Wikipédia. Fonction Cauchy-continue — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Fonction_Cauchy-continue&oldid=157229246.

[377] Wikiversité. Topologie générale/Complétude — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Topologie_g%C3%A9n%C3%A9rale/Compl%C3%A9tude&oldid=753513. En ligne ; accédé le 24-juillet-2019

[378] Provatiscus. Uniforme continuité utilisee sans justification page 656 URL https://github.com/LaurentClaessens/mazhe/issues/124.

[379] Adrien Fontaine. Développement : théorème de Sarkiowski URL http://perso.eleves.bretagne.ens-cachan.fr/~afontain/dvptthmdesarkowski.pdf.

[380] Pierre Monmarché. Développements URL http://laurent.claessens-donadello.eu/pdf/1-total.pdf.

[381] Wikipédia. Théorème de Heine — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Heine&oldid=145940281.

[382] Callus. URL http://math.stackexchange.com/questions/1829332/prove-that-if-v-is-finite-dimensional-then-v-is-even-dimensional.

[383] J. L. Littlewoor. Every polynomyal has a root URL http://www.math.univ-toulouse.fr/~bauval/Littlewood-95-8.pdf.

[384] Wikipédia. Théorème fondamental de l'algèbre — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_fondamental_de_l%27alg%C3%A8bre&oldid=139327371.

[385] Wikipédia. Trigonalisation — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Trigonalisation&oldid=109085753.

[386] Henry C. King. Unitary diagonalization of matrices URL http://www-users.math.umd.edu/~hck/Normal.pdf.

[387] Wikipedia. Spectral theorem — Wikipedia, The Free Encyclopedia URL http://en.wikipedia.org/w/index.php?title=Spectral_theorem&oldid=575488135.

[388] theorem for normal triangular matrices URL http://planetmath.org/theoremfornormaltriangularmatrices.

[389] Michel Granget. Normes matricielles, conditionnement URL http://math.univ-angers.fr/~granger/ananum/Chapitre_II.pdf.

[390] David Michel. Théorème de Lie-Kolchin URL https://perso.eleves.ens-rennes.fr/people/David.Michel/agreg/TheoremeLieKolchin.pdf.

[391] David Delaunay. Nilpotence URL http://mp.cpgedupuydelome.fr/pdf/Réductiondesendomorphismes-Nilpotence.pdf.

[392] Pierre-Emmanuel Jabin. Analyse numérique, correction du TD 8 URL http://math.unice.fr/~jabin/CTD3-8.pdf.

[393] Gallouët !Thierry. Norme et conditionnement d'une matrice URL https://www.i2m.univ-amu.fr/~gallouet/licence.d/anum.d/anum-tg2.pdf.

[394] martini. Why is the operator 2-norm of a diagonal matrix its largest value ? URL https://math.stackexchange.com/q/413375.

[395] Wikipédia. Théorème de Thalès — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Thal%C3%A8s&oldid=188328664.

[396] Wikipédia. Angle inscrit dans un demi-cercle — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Angle_inscrit_dans_un_demi-cercle&oldid=187962800.

[397] URL https://www.math.hmc.edu/calculus/tutorials/quotient_rule/proof.pdf.

[398] Wikiversité. Fonctions d'une variable réelle/Dérivabilité — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Fonctions_d%27une_variable_r%C3%A9elle/D%C3%A9rivabilit%C3%A9&oldid=718842. En ligne ; accédé le 24-mai-2018

[399] Wikiversité. Fonction dérivée/Dérivée d'un quotient — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Fonction_d%C3%A9riv%C3%A9e/D%C3%A9riv%C3%A9e_d%27un_quotient&oldid=675129. En ligne ; accédé le 24-mai-2018

[400] Wikipédia. Opérations sur les dérivées — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Op%C3%A9rations_sur_les_d%C3%A9riv%C3%A9es&oldid=99099598.

[401] Andrew D. Hwang. Open Unit Ball diffeomorphic to the Open Unit Cube URL https://math.stackexchange.com/q/1241724.

[402] Wikiversité. Fonctions d'une variable réelle/Dérivabilité — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Fonctions_d%27une_variable_r%C3%A9elle/D%C3%A9rivabilit%C3%A9&oldid=819258. En ligne ; accédé le 14-novembre-2020

[403] Wikiversité. Fonctions d'une variable réelle/Dérivabilité — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Fonctions_d%27une_variable_r%C3%A9elle/D%C3%A9rivabilit%C3%A9&oldid=753518. En ligne ; accédé le 19-juillet-2019

[404] Wikipédia. Théorème de Rolle — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Rolle&oldid=132483699.

[405] Guillaume Connan. Dérivation URL http://gconnan.free.fr/les%20pdf/Deriv.pdf.

[406] Keith Konrad. The remainder in Taylor series URL http://www.math.uconn.edu/~kconrad/blurbs/analysis/TaylorRemainder.pdf.

[407] Wikipédia. Règle de L'Hôpital — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=R%C3%A8gle_de_L%27H%C3%B4pital&oldid=131491717.

[408] . Squeeze theorem — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Squeeze_theorem&oldid=903141752.

[409] Kévin Santugino. Comprendre les dérivées partielles et leurs notations URL https://www.math.u-bordeaux.fr/~ksantugi/downloads/PolyDeriveesPartielles.pdf.

[410] Wikipédia. Application multilinéaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Application_multilin%C3%A9aire&oldid=161449382.

[411] Wikipédia. Convergence uniforme — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Convergence_uniforme&oldid=95493463.

[412] Suites et séries de fonctions URL http://blog.psi945.fr/public/maths-psi/cours-psi-suites-series-fonctions.pdf.

[413] F. Poupaud. Analyse fonctionnelle pour la licence URL http://math.unice.fr/~rascle/pdffiles/coursanapp/ana-fonc.pdf.

[414] Nicole Bopp. Un complément à la leçon sur l'équation fonctionnelle de la fonction exponentielle URL http://irma.math.unistra.fr/~bopp/CAPES/cours/equation-felle-exp.pdf.

[415] Michael Gechele. Théorème de Stone-Weierstrass URL http://michael.gechele.perso.neuf.fr/Agregation/Theoreme_Stone_Weierstrass.pdf.

[416] Primitives et intégrales URL http://math.univ-lyon1.fr/capes/IMG/pdf/new.primitive.pdf.

[417] E.F. Whittlesey. ANALYTICFUNCTIONS IN BANACH SPACES URL https://www.ams.org/journals/proc/1965-016-05/S0002-9939-1965-0184092-2/S0002-9939-1965-0184092-2.pdf.

[418] Jaap Eldering. Inversion is Smooth URL https://math.stackexchange.com/q/2145515.

[419] G. B. Folland. Taylor's formula URL https://sites.math.washington.edu/~folland/Math425/taylor.pdf.

[420] Wikipedia. Développement de Taylor URL http://fr.wikipedia.org/wiki/Développement_de_Taylor.

[421] Wikiversité. Calcul différentiel/Théorèmes utiles — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Calcul_diff%C3%A9rentiel/Th%C3%A9or%C3%A8mes_utiles&oldid=657239. En ligne ; accédé le 10-juillet-2017

[422] Jens Wagemaker. Let $A \in SO(3)$ show that 1 is always an eigenvalue. (Proof check and alternative proof ?) URL https://math.stackexchange.com/q/4183457.

[423] Yoann Gelineau. Générateurs du groupe linéaire URL http://math.univ-lyon1.fr/~gelineau/devagreg/Generateurs_Groupe_Lineaire.pdf.

[424] Sandrine Caruso. Générateurs de $\mathrm{GL}_n(K)$ et $\mathrm{SL}_n(K)$ URL http://sandrine.toonywood.org/pageperso/agreg/geneSL.pdf.

[425] Wikipédia. Théorème de Cayley-Hamilton — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Cayley-Hamilton&oldid=104470625.

[426] David Monniaux. URL http://david.monniaux.free.fr/dotclear/index.php/post/2014/06/09/Mémoire-à-effaçage-rapide. David Monniaux esquisse la preuve dans un commentaire à cette note

[427] Pierre Lissy. Décomposition polaire URL http://www.ljll.math.upmc.fr/~lissy/Agreg/developpements/DecPol.pdf.

[428] Wikipédia. Décomposition polaire — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=D%C3%A9composition_polaire&oldid=90240040.

[429] Jean-Etienne Rombaldi. Endomorphismes symétriques d'un espace vectoriel. Applications URL http://www-fourier.ujf-grenoble.fr/~rombaldi/AgregInterne/Oral1/120.pdf.

[430] Brice Loustau. Développements d'algèbre pour l'agrégation URL http://myismail.net/docs/divers/agreg/Sortie/LessonsAlgLoustau.pdf.

[431] Michel Coste. Sous-groupes compacts du groupe linéaire URL http://agreg-maths.univ-rennes1.fr/documentation/docs/ssgrpecompact.pdf.

[432] Daniel Li. Notions fondamentales de la théorie des probabilités URL http://labomathlens.free.fr/Liens/ProbaM1/PROBA01.pdf.

[433] Daniel Li. Tribus et mesures URL http://www.editions-ellipses.fr/product_info.php?products_id=9387.

[434] Daniel Saada. Tribu de Borel et tribu de Baire d'un espace topologique URL http://www.daniel-saada.eu/fichiers/16-Tribus_de_Baire.pdf.

[435] Wikipédia. Tribu-trace — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Tribu-trace&oldid=155238311.

[436] Daniel Li. Construction de la mesure de Lebesgue URL http://www.editions-ellipses.fr/product_info.php?products_id=9387.

[437] Wikipédia. Mesure (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Mesure_(math%C3%A9matiques)&oldid=145203435.

[438] . Field of sets — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Field_of_sets&oldid=1055198868.

[439] François de Marçay. Théorie abstraite de l'intégration et théorème de Radon-Nikodym URL https://www.imo.universite-paris-saclay.fr/~joel.merker/Enseignement/Integration/abstraite-integration.pdf.

[440] Mark. Two « simultaneously » sigma-finite measure 23 juillet 2023. URL https://math.stackexchange.com/q/4741254.

[441] ProofWiki. Outer Measure of Limit of Increasing Sequence of Sets — ProofWiki, URL http://www.proofwiki.org/w/index.php?title=Outer_Measure_of_Limit_of_Increasing_Sequence_of_Sets&oldid=85952.

[442] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Quelques résultats d'unicité URL http://www.les-mathematiques.net/a/i/d/node1.php#T42.

[443] James Norris. Uniqueness of measures URL http://www.statslab.cam.ac.uk/~james/Lectures/pm3.pdf.

[444] Wikipédia. Tribu borélienne 24 aout 2024. URL https://fr.wikipedia.org/wiki/Tribu_bor%C3%A9lienne.

[445] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Ensembles négligeables et complétion de tribus URL http://www.les-mathematiques.net/a/i/c/node1.php.

[446] Wikipédia. Complétion d'une mesure — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Compl%C3%A9tion_d%27une_mesure&oldid=89819650.

[447] B. Mauray. Intégration jusqu'au théorème de Lebesgue URL https://webusers.imj-prg.fr/~bernard.maurey/agreg/Textes/Lebesgue.pdf.

[448] Gerald Teschl. Topics in real and functional analysis URL http://www.mat.univie.ac.at/~gerald/ftp/book-fa/fa.pdf.

[449] Emmanuel Vieillard Baron. Les fonctions mesurables URL http://www.les-mathematiques.net/a/i/b/node1.php.

[450] Wikipédia. Fonction étagée — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Fonction_%C3%A9tag%C3%A9e&oldid=104836104.

[451] John K. Hunter. Measure theory URL https://www.math.ucdavis.edu/~hunter/measure_theory/measure_notes.pdf.

[452] Daniel Li. Intégration sur un espace produit URL http://www.editions-ellipses.fr/product_info.php?products_id=9387.

[453] Chat GPT. Chat GPT avec un bon prompt, même pas compliqué la plupart des fois. URL https://chatgpt.com.

[454] Rémi Peyre and Frédéric Simon. Travaux dirigés de probabilités URL http://www.normalesup.org/~rpeyre/pro/enseignement/td09pbas-e.pdf.

[455] Wikipédia. Ensemble de Vitali — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Ensemble_de_Vitali&oldid=100925109.

[456] c. Re : mesure non nulle et existence d'un ouvert URL http://www.les-mathematiques.net/phorum/read.php?4,608327,608430#msg-608430.

[457] Daniel Choï. Mathématiques pour la mécanique URL http://www.meca.unicaen.fr/Enseignement/Document/CoursChoi/mathmeca.pdf.

[458] Wikipédia. Lebesgue integration 7 janvier 2024. URL https://en.wikipedia.org/wiki/Lebesgue_integration.

[459] Gilles Dubois. URL http://dubois.gilles.pagesperso-orange.fr/analyse_reelle/intlimites.html.

[460] Prove the Countable additivity of Lebesgue Integral. URL https://math.stackexchange.com/questions/668958/prove-the-countable-additivity-of-lebesgue-integral.

[461] Daniel Li. Construction de l'intégrale de Lebesgue URL http://www.editions-ellipses.fr/product_info.php?products_id=9387.

[462] David Madore. Sur la rédaction des maths et la recherche de l'inambiguïé URL http://www.madore.org/~david/weblog/d.2018-10-24.2562.html.

[463] Yao Jianfeng. Statistique et logiciels URL http://perso.univ-rennes1.fr/jian-feng.yao/pedago/statlog/.

[464] user1770201. Showing that the Countable Sum of Measures is Mutually Singular to another Measure 18 décembre 2012. URL https://math.stackexchange.com/q/261618.

[465] Laurent Claessens. Why the sum is -oo ? 23 juillet 2023. URL https://en.wikipedia.org/wiki/Talk:Hahn_decomposition_theorem.

[466] Bruce Driver. Analysis tools with applications URL http://www.math.ucsd.edu/~bdriver/240-01-02/Lecture_Notes/anal.pdf.

[467] B. Mauray. Préparation à l'agrégation, analyse URL http://www.math.jussieu.fr/~maurey/agreg/index.html.

[468] Thierry Gallouët and Raphaèle Herbin. Mesure, intégration, probabilités URL http://www.cmi.univ-mrs.fr/~gallouet/licence.d/mes-int-pro.pdf.

[469] Daniel Li. Changement de variables dans les intégrales sur un ouvert de $\mathbb{R}^N$ URL http://li.perso.math.cnrs.fr/textes/Integration/change-v.pdf.

[470] Bruno Demange. Théorème de changement de variables URL https://www-fourier.ujf-grenoble.fr/~demange/integration/2012/integration-chap5.pdf.

[471] Wikipédia. Intégrale impropre — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Int%C3%A9grale_impropre&oldid=103902526.

[472] Rong-Qing Jia . Chapter 3. Absolutely Continuous Functions URL https://sites.ualberta.ca/~rjia/Math418/Notes/chap3.pdf.

[473] Rong-Qing Jia . Chapter 2. Functions of Bounded Variation URL https://sites.ualberta.ca/~rjia/Math418/Notes/chap2.pdf.

[474] Wikipédia. Bounded variation 21 juillet 2024. URL https://en.wikipedia.org/wiki/Bounded_variation.

[475] Jean Abdou Samra. Un ouvert connexe par arcs est connexe par arcs C1 5 novembre 2022. Communication personnelle

[476] Romain Boillaud. Séries de fonctions URL http://www.rblld.fr/cours/chap6.pdf.

[477] Wikipédia. Rayon de convergence — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Rayon_de_convergence&oldid=161945407.

[478] Gérard Eguether. Ex-séries entières URL http://iecl.univ-lorraine.fr/~Gerard.Eguether/zARTICLE/EX.pdf.

[479] G. Costantini. RÈGLES DE D'ALEMBERT ET DE CAUCHY POUR LES SÉRIES À TERMES RÉELS POSITIFS

[480] Règles de d'Alembert et de Cauchy URL https://uel.unisciel.fr/mathematiques/serie/serie_ch01/co/apprendre_10.html.

[481] . Cauchy–Hadamard theorem — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Cauchy%E2%80%93Hadamard_theorem&oldid=1060324209.

[482] Wikiversité. Série numérique/Produit de Cauchy — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=S%C3%A9rie_num%C3%A9rique/Produit_de_Cauchy&oldid=788554. En ligne ; accédé le 13-mars-2021

[483] Wikipédia. Produit de Cauchy — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Produit_de_Cauchy&oldid=178673280.

[484] Georges Skandalis. Analyse, résumés et exercices URL https://www.math.univ-paris-diderot.fr/_media/formations/prepa/agreginterne/polycopieanalyse.pdf.

[485] Rahul Krishna. Approximations by Taylor polynomials URL http://math.columbia.edu/~krishna/08092011.pdf.

[486] Vogel Thomas. Analytic functions URL http://www.math.tamu.edu/~tvogel/410/sect74a.pdf.

[487] Jimmie Lawson. Chapter 4 partitions of unity and smooth functions URL https://www.math.lsu.edu/~lawson/Chapter4.pdf.

[488] J. Parizet. Algèbre engendrée par une matrice carrée URL https://clong.pagesperso-orange.fr/semana/documents/parizet/algA.pdf.

[489] N. P. Landsman. Lectures notes on C^*-algabras, Hilbert C^*-modules, and quantum mechanics math-ph/9807030v1

[490] Dan Klain. The matrix exponential (with exercices) URL http://faculty.uml.edu/dklain/exponential.pdf.

[491] numberphile . A proof that e is irrational URL https://www.youtube.com/watch?v=xOXsDfMMTjs&pp=ugMICgJmchABGAE%3D.

[492] ENS Cachan. Théorème de stabilité de Lyapunov URL http://minerve.bretagne.ens-cachan.fr/images/Lyapunov.pdf.

[493] Rached Mneimé. Réduction des endomorphismes

[494] Jaap Eldering. Inversion is Smooth October 6 2022. URL https://math.stackexchange.com/q/2145515.

[495] Matthieu Romagny. Représentations linéaires des groupes finis URL http://www.math.jussieu.fr/~romagny/agreg/theme/rlgf_paysage.pdf.

[496] Sébastien Pellerin. Développements d'analyse pour l'oral de l'agrégation URL http://pellerin.xyz/doc/agreg/analyse.pdf.

[497] Kevin Quirin. Théorème de Müntz URL http://kevin.quirin.free.fr/Trucs/agreg/Dvt_Muntz.pdf.

[498] Pierre Lissy. Déterminant de Cauchy et application au théorème de Müntz URL http://www.ljll.math.upmc.fr/~lissy/Agreg/developpements/CauchyMuntz.pdf.

[499] Benjamin Dadoun. Théorème de Müntz URL http://benjamin.dadoun.free.fr/muntz.pdf.

[500] Jean-Louis Rouget. Intégrales dépendant d'une paramètre URL http://www.maths-france.fr/MathSpe/Cours/09-integrales-a-parametres.pdf.

[501] Wikipédia. Lemme de Hadamard — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Lemme_de_Hadamard&oldid=90659396.

[502] Jean-Pierre Demailly. Analyse numérique et équations différentielles

[503] bibmath. Point fixe Février 2023. URL https://www.bibmath.net/dico/index.php?action=affiche&quoi=./p/pointfixe.html.

[504] Clémence Minazzo and Kelsey Rider. Théorèmes du point fixe et applications aux équations différentielles URL http://math.unice.fr/~eaubry/Enseignement/M1/memoire.pdf.

[505] Ivan Nourdin. Leçons d'analyse, probabilités algèbre et géométrie

[506] Sandrine Caruso. Théorème de Cauchy-Lipschitz URL http://boumbo.toonywood.org/sandrine/pageperso/agreg/cauchy-lipschitz.pdf.

[507] Franck Boyer. Équations différentielles ordinaires URL http://www.latp.univ-mrs.fr/~fboyer/Enseignement/Agreg/cours_EDO_Agreg_FBoyer.pdf.

[508] Théo Pierron. Théorème de Cauchy-Lipschitz global URL http://perso.eleves.ens-rennes.fr/~tpier758/agreg/dvpt/maths/cauchy_lipschitz.pdf.

[509] Thierry Audibert. Le problème de Cauchy, résultats fondamentaux URL http://www.univenligne.fr/Documents/PolycopiesAgreg/EDO1.pdf.

[510] Thomas Budzinski. Théorème de Cauchy-Lipschitz URL https://www.eleves.ens.fr/home/budzinsk/polys/Nonolympique/2014_cauchylipschitz.pdf.

[511] Wikipédia. Théorème d'inversion locale — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_d%27inversion_locale&oldid=100975433.

[512] Antoine Chambert-Loir. Introduction aux groupes et algèbres de Lie URL http://www.math.u-psud.fr/~chambert/enseignement/2004-05/lie/lie.pdf.

[513] E. P. Ban. Lie groups URL http://www.math.uu.nl/people/ban/lecnot.html.

[514] Wikipédia. Extremum — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Extremum&oldid=177462314.

[515] Joachim Stubbe. URL http://mathaa.epfl.ch/prob/enseignement/analyse2/series/AnalysisE.pdf.

[516] Laurent Guillopé. Optimisation sous contrainte URL https://www.math.sciences.univ-nantes.fr/~guillope/l3-osc/osc.pdf.

[517] Bruno Galerne. Optimisation, algorithmique URL http://www.math-info.univ-paris5.fr/~bgalerne/m1_opti_algo/poly_opti_algo.pdf.

[518] Wikipédia. Fonction convexe — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Fonction_convexe&oldid=157832047.

[519] Wikiversité. Fonctions convexes/Définition et premières propriétés — Wikiversité, URL http://fr.wikiversity.org/w/index.php?title=Fonctions_convexes/D%C3%A9finition_et_premi%C3%A8res_propri%C3%A9t%C3%A9s&oldid=381775. En ligne ; accédé le 9-février-2014

[520] Wikiversité. Fonctions convexes/Fonctions convexes dérivables — Wikiversité, URL http://fr.wikiversity.org/w/index.php?title=Fonctions_convexes/Fonctions_convexes_d%C3%A9rivables&oldid=379779. En ligne ; accédé le 9-février-2014

[521] Stephen Boyd and Lieven Vandenberghe. Convex optimization URL http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

[522] Document 33 : fonctions convexes URL http://math.univ-lyon1.fr/capes/IMG/pdf/new.convexe.pdf.

[523] Wikipédia. Inégalité arithmético-géométrique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=In%C3%A9galit%C3%A9_arithm%C3%A9tico-g%C3%A9om%C3%A9trique&oldid=99682278.

[524] Wikipédia. Inégalité de Kantorovitch — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=In%C3%A9galit%C3%A9_de_Kantorovitch&oldid=102417007.

[525] John K. Hunter. Chapter 13 : Metric, normed and topological spaces URL https://www.math.ucdavis.edu/~hunter/intro_analysis_pdf/ch13.pdf.

[526] Jean-François Burnol. Normes Lp URL http://jf.burnol.free.fr/agregnormeslp.pdf.

[527] Georges Comte. Lemme de Morse URL http://gc83.perso.sfr.fr/Agregation/LemmedeMorse.pdf.

[528] Wikipédia. Calcul du volume de l'hypersphère — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Calcul_du_volume_de_l%27hypersph%C3%A8re&oldid=147651804.

[529] Florian Bouguet. Ellipsoïde de John-Lowner URL http://florian.bouguet.free.fr/doc/developpements/john-loewner.pdf.

[530] Martin Henk. Löewnr-John ellipsoids URL https://www.math.uni-bielefeld.de/documenta/vol-ismp/24_henk-martin.pdf.

[531] Shuzhong Zhang. Supplement : the Löwner-John ellipsoids URL http://www.isye.umn.edu/courses/ie8534/pdf/Loewner-John_ellipsoid.pdf.

[532] B. Mauray. Analyse fonctionnelle et théorie spectrale (version longue) URL http://www.math.jussieu.fr/~maurey/ts012/poly/lths.pdf.

[533] Michel Rascle. Analyse fonctionnelle pour la licence URL http://math.unice.fr/~rascle/pdffiles/coursanapp/ana-fonc.pdf.

[534] Mourad Besbes. Chapitre 5 : espaces métriques complets, espaces de Banach URL http://besbes.mourad.free.fr/Enseignement/Topologie/Chapitre5.pdf.

[535] C. Antonini and JF. Quint and P. Borgnat and J. Bérard and E. Lebeau and E. Souche and A. Chateau and O. Teytaud. Complété d'un espace métrique URL http://www.les-mathematiques.net/a/a/b/node26.php.

[536] Liou. Completion of a metric space URL https://www.math.ncku.edu.tw/~fjmliou/advcal/Completion.pdf.

[537] Wikipédia. Espace complet — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_complet&oldid=91125353.

[538] B. Maurey. Holomorphie URL http://www.math.jussieu.fr/~maurey/agreg/Textes/Holomorphe045.pdf.

[539] Samuel and Nicolas and Ivik . autour de 2000-2002. Provient de notes de TP données à l'Université Libre de Bruxelles. Je ne suis pas tout à fait certain des auteurs

[540] Romain Boillaud. Séries entières URL http://www.rblld.fr/cours/chap8.pdf.

[541] Philippe Picart. Définition de cos et sin par les séries entières URL http://trucsmaths.free.fr/telech/exp_cos_sin.pdf.

[542] Wikipédia. Pi — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Pi&oldid=133197907.

[543] Alfred Gray. Modern differential geometry of curves and surfaces with mathematica URL http://webmath2.unito.it/paginepersonali/sergio.console/CurveSuperfici/.

[544] Wikiversité. Trigonométrie/Cosinus et sinus dans le cercle trigonométrique — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Trigonom%C3%A9trie/Cosinus_et_sinus_dans_le_cercle_trigonom%C3%A9trique&oldid=801599. En ligne ; accédé le 28-décembre-2020

[545] David Cimasoni. Cours de géométrie I, semestre de printemps URL http://www.unige.ch/math/folks/cimasoni/GeometrieI.pdf.

[546] Sylvain. Générateurs de Isom(E) URL https://agreg-maths.fr/developpements/147.

[547] Christophe Mourougane. Théorie des groupes et géométrie URL http://perso.univ-rennes1.fr/christophe.mourougane/enseignements/2008-9/THGG/poly.groupes.pdf.

[548] Daniel Perrin. Isométries du plan URL https://www.math.u-psud.fr/~perrin/CAPES/geometrie/isometries11-12.pdf.

[549] Éric Jourdain. Angles URL https://perso.univ-rennes1.fr/eric.jourdain/GEEU/Cours/Angles.pdf.

[550] Wikipédia. Angle — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Angle&oldid=130370952.

[551] B. Guesmi. Théorème de l'angle inscrit URL http://s999cdfd874dd6835.jimcontent.com/download/version/1427398317/module/5509391568/name/theoremedel'arccapable.pdf1.pdf.

[552] wikipédia. Polygone 13 juillet 2022. URL https://fr.wikipedia.org/wiki/Polygone.

[553] Wikipedia contributors. Convex polygon — Wikipedia, The Free Encyclopedia 2022. URL https://en.wikipedia.org/wiki/Convex_polygon.

[554] Groupes diédraux URL http://theoriedesgroupes.perso.sfr.fr/cours/Diedraux.pdf.

[555] Pierre Lissy. URL http://www.ljll.math.upmc.fr/~lissy/Agreg/developpements/Died.pdf.

[556] Pierre Renfer. Roulettes et colliers URL http://les.mathematiques.free.fr/pdf/collier.pdf.

[557] Nicole Bopp. Un complément à la leçon sur les isométries d'un polygone régulier. URL http://irma.math.unistra.fr/~bopp/CAPES/cours/groupe-fini-isometrie.pdf.

[558] Théo Pierron. Groupes de pavage du plan URL http://perso.eleves.ens-rennes.fr/~tpier758/agreg/dvpt/useless/gpe_pavage.pdf.

[559] Frédéric Hélein. Géométrie euclidienne URL https://webusers.imj-prg.fr/~frederic.helein/cours/euclide.pdf.

[560] Laurent Dietrich. Les sous-groupes finis de SO(3) URL http://www.math.cmu.edu/~ldietric/doc/ssgrso3.pdf.

[561] Sous groupes finis de SO(3,R) URL http://minerve.bretagne.ens-cachan.fr/images/Sous-groupes_finis_SO3.pdf.

[562] Pierre Bernard. Entiers algébriques et racines de l'unité URL http://allken-bernard.org/pierre/weblog/?p=2061.

[563] Marc Sage. Sons, fréquences, harmoniques, tons : le compromis du piano URL http://www.normalesup.org/~sage/Musique/Harmonie.pdf.

[564] Décomposition en éléments simples. Cas complexe. URL https://cahier-de-prepa.fr/pc-carnot/download?id=2.

[565] Wikipédia. Chiffrement RSA — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Chiffrement_RSA&oldid=146774926.

[566] URL https://www.torproject.org/about/torusers.html.en.

[567] URL https://www.torproject.org/projects/torbrowser.html.en.

[568] Wikipédia. Polynôme cyclotomique — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Polyn%C3%B4me_cyclotomique&oldid=93751495.

[569] Adam Parusinski. Théorème de Dirichlet (version faible) URL http://math.unice.fr/~parus/AGREG/Dev/dirichlet.pdf.

[570] Michel Merle. corps finis URL http://www-math.unice.fr/~merle/Complements/corps_finis.pdf.

[571] Gabriel Peyré. Corps finis URL https://objectifagregation.github.io/.

[572] François Rodier. Corps finis URL http://iml.univ-mrs.fr/~rodier/Cours/RappelCorpsfinis.pdf.

[573] Stephane Vento. La loi de réciprocité quadratique URL http://www.proba.jussieu.fr/pageperso/nourdin/LeSiteDeLAgregatif/vento2.pdf.

[574] . Chevalley–Warning theorem — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Chevalley%E2%80%93Warning_theorem&oldid=792119700.

[575] Christian Squarcini. Anneaux factoriels URL http://christian-squarcini.pagesperso-orange.fr/AgregInterne/Anneauxcorps/3_2.pdf.

[576] Wikipédia. Fonction de Möbius — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Fonction_de_M%C3%B6bius&oldid=91236321.

[577] Wikipédia. Corps fini — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Corps_fini&oldid=92489466.

[578] Wikipédia. Nombre constructible — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Nombre_constructible&oldid=103130793.

[579] Tracés usuels « à la règle et au compas » URL http://dokeos3.u-cergy.fr/dokeos20/courses/EEMEM2CRPE/scorm/2012_2013SCORMEEMEUE3EC2M2P1_SCORM/page_04.htm.

[580] Wikipédia. Théorème de Wantzel — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Wantzel&oldid=103220511.

[581] Wikipédia. Nombre de Fermat — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Nombre_de_Fermat&oldid=103080861.

[582] Ismaël Bailleul. Sous-variétés de $\mathbb{R}^n$, théorème d'inversion locale et conséquences URL https://agreg-maths.univ-rennes1.fr/documentation/docs/CoursSousVarietes.pdf.

[583] Geoffrey Deperle. Théorème du point fixe de Brouwer URL https://agreg-maths.fr/uploads/versions/4665/Th%C3%A9or%C3%A8me%20du%20point%20fixe%20de%20Brouwer.pdf.

[584] William K Allard. The Brouwer fixed point theorem URL http://www.math.duke.edu/~wka/math204/fixed.pdf.

[585] Kenneth Kuttler. Topics in Analysis URL http://www.math.byu.edu/~klkuttle/SobolevSpacesB.pdf.

[586] Hervé Le Dret. Notes de cours M2 Équations aux dérivées partielles élliptiques URL https://www.ljll.math.upmc.fr/~ledret/M2Elliptique/chapitre4.pdf.

[587] Bachir Bekka. Théorème du point fixe de Markov-Kakutani : applications à l'existence de mesures de Haar et aux chaînes de Markov URL http://agreg-maths.univ-rennes1.fr/documentation/docs/PointFixe.pdf.

[588] Guillaume CARLIER. Analyse complexe 2012-2013. URL https://www.ceremade.dauphine.fr/~carlier/analysecomplexe.pdf.

[589] Wikipédia. Force conservative — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Force_conservative&oldid=168767525.

[590] Wikipédia. Travail d'une force — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Travail_d%27une_force&oldid=178096219.

[591] Olivier Casterá. Différentielles totales exactes URL http://o.castera.free.fr/pdf/Differentielle_totale_exacte.pdf.

[592] Wikipédia. Théorème de Jordan — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Jordan&oldid=124990679.

[593] Intégration de fractions rationnelles URL http://blascheck.franck.free.fr/IMG/pdf/IntegFracRat.pdf.

[594] Ophélie Rouby. Théorème de Rothstein-Trager URL http://perso.eleves.bretagne.ens-cachan.fr/~oroub842/dvlpts/Rothstein_Trager.pdf.

[595] Frédéric Chyzak. Intégration symbolique des fractions rationnelles URL http://www.enseignement.polytechnique.fr/informatique/profs/Frederic.Chyzak/CF-2004/notes-IntRat.pdf.

[596] Jerry Kazdan. Taylor's Theorem - Integral Remainder URL https://www2.math.upenn.edu/~kazdan/361F15/Notes/Taylor-integral.pdf.

[597] Sylvain Golénia. Approximation de ln(2) URL http://mathaoutils.blogspot.com/2018/11/approximation-de-ln2.html.

[598] Louis Fauchier-Magnan. Introduction à la théorie de Morse URL http://cqfd.epfl.ch/webdav/site/cqfd/shared/projets/igat/IntroductionàlathéoriedeMorse-LouisFauchier-Magnan.pdf.

[599] Reduction Formula for Integral of Power of Sine URL https://proofwiki.org/wiki/Reduction_Formula_for_Integral_of_Power_of_Sine.

[600] Michel Stainer. Compléments sur les séries numériques URL http://michel.stainer.pagesperso-orange.fr/PSIx/Cours/Ch03-ComplSeriesNum.pdf.

[601] Wallis's Product URL https://proofwiki.org/wiki/Wallis%27s_Product.

[602] Jacek Cichoń. Stirling approximation formula URL http://cs.pwr.edu.pl/cichon/Math/StirlingApp.pdf.

[603] Wikipédia. Formule de Stirling — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Formule_de_Stirling&oldid=93224272.

[604] Jean-Louis Rouget. Formule de Stirling URL https://www.maths-france.fr/MathSpe/GrandsClassiquesDeConcours/SeriesNumeriques/FormuleDeStirling.pdf.

[605] Stirling's Formula URL https://proofwiki.org/wiki/Stirling%27s_Formula.

[606] Keith Conrad. Differentiating under the integral sign URL https://kconrad.math.uconn.edu/blurbs/analysis/diffunderint.pdf.

[607] Christophe Caignaert. Courbes URL http://c.caignaert.free.fr/chapitre15/node1.html.

[608] gb. Cercle circonscrit à une courbe URL http://www.les-mathematiques.net/phorum/read.php?8,750556,750586.

[609] Mohammad Ghomi. Lecture note on differential geometry URL http://people.math.gatech.edu/~ghomi/LectureNotes/LectureNotes5U.pdf.

[610] Wikipedia. Convex curve — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Convex_curve&oldid=685739052.

[611] Robert Rolland. Géométrie projective URL http://robert.rolland.acrypta.com/telechargements/geometrie/pr0.pdf.

[612] Daniel Bertrand. Géométrie projective URL http://www.math.jussieu.fr/~bertrand/Enseignement/epdf/Algeo09_chap2.pdf.

[613] Nicolas Jacon. Géométrie projective URL http://njacon.perso.math.cnrs.fr/jacon_Geometrieproj.pdf.

[614] François Labourie. Géométrie affine et projective URL http://math.unice.fr/~labourie/preprints/pdf/geomproj.pdf.

[615] Matthieu Romagny. Droites projectives, homographies URL https://perso.univ-rennes1.fr/matthieu.romagny/agreg/droite_projective_Halberstadt.pdf.

[616] Arnaud Bodin. L'inversion URL http://math.univ-lille1.fr/~bodin/geometrie/ch_inversion.pdf.

[617] Laurent Dietrich. Le groupe circulaire URL http://www.math.cmu.edu/~ldietric/doc/groupecirc.pdf.

[618] Florian Bouget. Le groupe circulaire URL http://florian.bouguet.free.fr/doc/developpements/Groupe_circulaire.pdf.

[619] 182 : utilisation des nombres complexes en géométrie. Homographies URL http://minerve.bretagne.ens-cachan.fr/images/182_2015-2016.pdf.

[620] Sébastien Pellerin. Action du groupe modulaire sur le demi-plan de Poincaréé URL http://dynamaths.free.fr/docs/lecons/developpement_algebre_6.pdf.

[621] Alexandre Alessandri. Action du groupe modulaire sur le demi-plan de Poincaré URL http://minerve.bretagne.ens-cachan.fr/images/Dvt_psl2.pdf.

[622] Gabriel Nagy. Operator theory in Hilbert spaces URL https://www.math.ksu.edu/~nagy/real-an/2-07-op-th.pdf.

[623] Aline Kurtzmann. Espérance conditionnelle URL http://www.iecn.u-nancy.fr/~kurtzman/cours/agreg/esperance-conditionnelle.pdf.

[624] Wikipédia. Théorème de projection sur un convexe fermé — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_projection_sur_un_convexe_ferm%C3%A9&oldid=90246160.

[625] Michel Merle. Espaces de Hilbert URL http://www-math.unice.fr/~merle/Algebre_et_geometrie/hilbert.pdf.

[626] Douglas N. Arnold. Functional analysis URL http://www.math.umn.edu/~arnold/502.s97/functional.pdf.

[627] Wikipédia. Base de Hilbert URL https://fr.wikipedia.org/wiki/Base_de_Hilbert?oldid=128533376.

[628] Christiane Schomblond. Théorie quantique des champs, QED, QCD URL http://homepages.ulb.ac.be/~cschomb/intfonc+QED+QCD.pdf.

[629] Remsirems. Re : Réponses à quelques questions URL http://linuxfr.org/nodes/110155/comments/1675813.

[630] Jonathan Bain. Kochen-Specker and Measurement URL http://faculty.poly.edu/~jbain/philqm/lectureslides/08.KS&Measurement.pdf.

[631] Michel Kern. Introduction à la méthode des éléments finie URL http://mms2.ensmp.fr/ef_paris/formulation/polycop/f_coursEF.pdf.

[632] Antoine Bensalah. Communication électronique privée

[633] Wikipédia. Série de Laurent — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=S%C3%A9rie_de_Laurent&oldid=181017723.

[634] Pascal Dingoyan. URL http://www.math.jussieu.fr/~dingoyan/lescours/lm367-pageweb/lm367-chapitre5.pdf.

[635] Noach Dana-Picard. The logarithm of a complex number URL http://ndp.jct.ac.il/tutorials/complex/node26.html.

[636] Guillaume Carlier. Analyse complexe URL https://www.ceremade.dauphine.fr/~carlier/analysecomplexe.pdf.

[637] Wikiversité. Fonctions d'une variable complexe/Le logarithme complexe — Wikiversité, URL https://fr.wikiversity.org/w/index.php?title=Fonctions_d%27une_variable_complexe/Le_logarithme_complexe&oldid=617412. En ligne ; accédé le 16-décembre-2016

[638] psykokwak. Propriétés de l'indice et application au calcul d'intégrale URL https://agreg-maths.fr/uploads/versions/1961/DEV%20indice.pdf.

[639] Wikipédia. Anneau principal — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Anneau_principal&oldid=141994549.

[640] Corentin. Re : formule de Cauchy et dérivation URL http://www.les-mathematiques.net/phorum/read.php?4,385272,385289.

[641] Philippe Charpentier. Analyse complexe URL http://www.math.u-bordeaux1.fr/~pcharpen/enseignement/fichiers-master1/Analyse_Complexe.pdf.

[642] Wikipédia. Pôle (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=P%C3%B4le_(math%C3%A9matiques)&oldid=155205828.

[643] Wikipédia. Singularité isolée — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Singularit%C3%A9_isol%C3%A9e&oldid=155006512.

[644] Christian Kkein. Séries de Laurent, résidus URL http://math.u-bourgogne.fr/IMB/klein/Complement_dAnalyse_files/chap607.pdf.

[645] Marie. Partition d'un entier en parts fixées URL https://www.agreg-maths.fr/uploads/versions/812/Partitiond'unentierenpartsfixées.pdf.

[646] Quadriviuum Tremens. Une introduction à la Théorie Analytique des Nombres - Hors-Série URL https://www.youtube.com/watch?v=P7_OWtw6C-Y.

[647] Wikipédia. Intégrale de Fresnel — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Int%C3%A9grale_de_Fresnel&oldid=132170602.

[648] David Sirajuddin. Fresnel integrals URL http://itcanbeshown.com/integrals/FresnelIntegrals/fresnel_integrals.pdf.

[649] Jean-Marie Lion. Fonctions holomorphes d'une variable URL http://perso.univ-rennes1.fr/jean-marie.lion/coursholo.pdf.

[650] Oliver Knill. A short introduction to several complex variables URL https://people.math.harvard.edu/~knill/teaching/severalcomplex_1996/severalcomplex.pdf.

[651] Klaus Fritzche and Hans Grauert. From holomorphic functions to complex manifolds URL https://books.google.com/books?id=jSeRz36zXIMC&pg=PA33.

[652] Laurent Claessens. local inverse of analytic function URL https://math.stackexchange.com/q/4345396.

[653] Franck Boyer. Analyse fonctionnelle URL https://www.math.univ-toulouse.fr/~fboyer/_media/enseignements/m1af/poly_af_m1.pdf.

[654] Nikko and Bruno and Steff. l'inverse d'une application linéaire continue et bijective est-elle continue ? URL http://www.les-mathematiques.net/phorum/read.php?2,60410,60446.

[655] Wikipédia. Théorème d'Ascoli — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_d%27Ascoli&oldid=92958118.

[656] Keith Conrad. L^p spaces for $0 < p < 1$ URL https://www.math.uconn.edu/~kconrad/blurbs/analysis/lpspace.pdf.

[657] B. Maurey. URL http://www.math.jussieu.fr/~maurey/IntegProba/Cours/CoursIP101-14.pdf.

[658] Bosses glissantes URL http://www.bibmath.net/dico/index.php?action=affiche&quoi=./b/bosseglissante.html.

[659] Wikipédia. Inégalité de Young — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=In%C3%A9galit%C3%A9_de_Young&oldid=181063084.

[660] t.b.. On the equality case of the Hölder and Minkowski inequalities URL https://math.stackexchange.com/q/87940.

[661] Danarmk. URL http://linuxfr.org/nodes/110155/comments/1675589.

[662] B. Mauray. Convolution, inégalités, approximations et régularisation URL https://webusers.imj-prg.fr/~bernard.maurey/agreg/ag045/Convolution045.pdf.

[663] Schilling. Minkowsky's inequality for integrals URL http://math.ucsd.edu/~lni/math240/suppl-1.pdf.

[664] . Minkowski inequality — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Minkowski_inequality&oldid=1015403274.

[665] Daniel Fischer. When does equality hold in the Minkowski's inequality $\|f+g\|_p \leqslant \|f\|_p+\|g\|_p$? URL https://math.stackexchange.com/q/581305.

[666] D. Sarason. Math 202B Solutions URL https://math.berkeley.edu/~sarason/Class_Webpages/solutions_202B_assign5.pdf.

[667] Charles Suquet. Intégration, analyse de Fourier,Probabilités URL http://math.univ-lille1.fr/~suquet/ens/IFP/Cours/cours04/CoursIFP04.html.

[668] Wikipédia. Théorème de Riesz-Fischer — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Riesz-Fischer&oldid=90074115.

[669] Elias M. Stein and Rami Shakarchi. Real analysis URL http://home.ustc.edu.cn/~liweiyu/documents/(Princeton%20Lectures%20in%20Analysis)%20Elias%20M.%20Stein,%20Rami%20Shakarchi-Real%20analysis_%20measure%20theory,%20integration,%20and%20Hilbert%20spaces.%20Vol.3.-Princeton%20University%20Press%20(2005).pdf.

[670] JanTinbergen1991. Lp -space is a Hilbert space if and only if p=2 URL https://math.stackexchange.com/questions/3017814/lp-space-is-a-hilbert-space-if-and-only-if-p-2.

[671] Thierry Gallouët. Les espaces Lp URL https://www.i2m.univ-amu.fr/perso/thierry.gallouet/licence.d/int/part4.pdf.

[672] Jon A. Wellner. Some Basic Large Sample Theory URL https://sites.stat.washington.edu/jaw/COURSES/580s/581/LECTNOTES/ch2e-jaw.pdf.

[673] Wikipédia. Uniform integrability 17 juin 2024. URL https://en.wikipedia.org/wiki/Uniform_integrability.

[674] Mittens. Is ... linear for ...? 2024-06-28. URL https://math.stackexchange.com/q/4173760.

[675] Mittens. Prove that ... (Vitali) June 28 2024. URL https://math.stackexchange.com/q/4938755.

[676] Jean-Marc Bouclet. Convolution URL https://www.math.univ-toulouse.fr/~bouclet/Notes-de-cours-exo-exam/Agregation/convolution.pdf.

[677] convolution, associativity of URL https://planetmath.org/convolutionassociativityof.

[678] Gilles Leborgne. Introduction à la théorie des distributions URL http://www.isima.fr/~leborgne/IsimathDistributions/distrib.pdf.

[679] Daniel Raban. Math 245C Lecture 16 Notes May 6, 2019. URL https://pillowmath.github.io/Math%20245C/Lec16.pdf.

[680] Philippe Spindel. Éléments de géométrie différentielle pour la mécanique analytique et le gravitation URL https://portail.umons.ac.be/EN2/universite/facultes/fs/services/institut_physique/mecanique_et_gravitation/Documents/Textespédagogiques/geodiff.pdf.

[681] Hart Smith. Lecture 2 : convolution URL https://sites.math.washington.edu/~hart/m526/Lecture2.pdf.

[682] Léonard Gallardo. Chapitre 3 : les séries de Fourier URL http://www.lmpt.univ-tours.fr/~gallardo/EspHilbert3.pdf.

[683] Charles Suquet. Espaces de Hilbert et séries de Fourier URL http://math.univ-lille1.fr/~suquet/ens/IFP/Cours/cours04/Chap7ifp04.pdf.

[684] user17762. Showing the inequality ... 9 mai 2012. URL https://math.stackexchange.com/q/143188.

[685] Olof Hanner. On the uniform convexity of L^p and l^p URL https://projecteuclid.org/download/pdf_1/euclid.afm.1485893271.pdf.

[686] Assaf Naor. Proof of the uniform convexity lemma URL https://web.math.princeton.edu/~naor/homepagefiles/inequality.pdf.

[687] David Giraudo. URL https://math.stackexchange.com/questions/95307/showing-left-fracab2-rightp-left-fraca-b2-rightp-leq-frac12/95312#95312.

[688] Kai-Seng Chou and Tianwen Luo. Chapter 4 : The Lebesgue spaces URL https://www.math.cuhk.edu.hk/course_builder/1415/math5011/MATH5011_Chapter_4.2014.pdf.

[689] gerw. How to prove Clarkson's inequality ? URL https://math.stackexchange.com/questions/1607683/how-to-prove-clarksons-inequality.

[690] Ryszard Szwarc. Clarkson inequality for complex numbers May 24, 2022. URL https://math.stackexchange.com/q/4457490.

[691] R. Fortet. Remarques sur les espaces uniformément convexes URL http://www.numdam.org/article/BSMF_1941__69__23_0.pdf.

[692] M .A Usman and F. A. Hammed and M. O Olayiwola. Uniformly convex spaces URL http://www.sciencepub.net/researcher/0101/10_0506_UNIFORMLY_CONVEX.pdf.

[693] Jean-François Burnol. Le théorème de dualité des espaces L^p URL http://jf.burnol.free.fr/0506L312annexeDualiteLp.pdf.

[694] Brian Bockelman. Functional analysis notes URL http://www.math.unl.edu/~s-bbockel1/928/Hahn_Banach_Theorem.html.

[695] gerw. Show that the norm of a complex functional is equal to the norm of its real part 2017-03-16. URL https://math.stackexchange.com/q/2189895.

[696] Haïm Brézis. Analyse fonctionnelle Théorie et applications

[697] Alan Sokal. The Hahn-Banach theorem and duality of Banach spaces URL https://www.ucl.ac.uk/~ucahad0/3103_handout_6.pdf.

[698] Téofil Adamski. Le théorème de prolongement de Tietze URL https://perso.eleves.ens-rennes.fr/people/teofil.adamski/agreg/developpements/Tietze.pdf.

[699] Terence Tao. Brouwer's fixed point and invariance of domain theorems, and Hilbert's fifth problem 13 June 2011. URL https://terrytao.wordpress.com/2011/06/13/brouwers-fixed-point-and-invariance-of-domain-theorems-and-hilberts-fifth-problem/.

[700] Wikipédia. Espace Lp — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Espace_Lp&oldid=92618226.

[701] Didier Smets. Base d'analyse fonctionnelle URL https://www.ljll.math.upmc.fr/smets/MM005/MM005_Chapitre_6.pdf.

[702] Wikipédia. Théorème de Jordan — Wikipédia, l'encyclopédie libre 18 mars 2023. URL https://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Jordan&oldid=198580114.

[703] Sylvie Benzoni. Topologie faible URL http://math.univ-lyon1.fr/~benzoni/Agreg/Topologie.pdf.

[704] Sylvie Benzoni. Distributions tempérées URL http://math.univ-lyon1.fr/~benzoni/Agreg/Distributions.pdf.

[705] MoebiusCorzer Giraudo. Show smooth functions of compact support are dense in the Schwartz space URL http://math.stackexchange.com/questions/229751/show-smooth-functions-of-compact-support-are-dense-in-the-schwartz-space.

[706] Richard Gomez. Séries de Fourier URL http://megamaths.perso.neuf.fr/rg/fourier.pdf.

[707] David Delaunay. Séries de Fourier URL http://mp.cpgedupuydelome.fr/pdf/SériesdeFourier.pdf.

[708] B. Maurey. Notes de cours : analyse hilbertienne et de Fourier URL http://www.math.jussieu.fr/~maurey/FourHilb/PolyFH.pdf.

[709] Richard Zekri. Analyse de Fourier URL http://lumimath.univ-mrs.fr/infoetudiant/ANALYSEDEFOURIERMASTER2009.pdf.

[710] Michel Crouzeix. Compléments d'analyse URL https://perso.univ-rennes1.fr/michel.crouzeix/publis/anep.pdf.

[711] Bachir Bekka. Le théorème d'inversion de Fourier du point de vue des distributions URL http://agreg-maths.univ-rennes1.fr/documentation/docs/Theoreme-Inversion-Fourier.pdf.

[712] Erdős László. Fourier transform URL https://www.mathematik.uni-muenchen.de/~lerdos/WS08/QM/four08.pdf.

[713] Isabelle Gallagher. Chapitre B. Distributions URL https://www.math.ens.fr/~gallagher/ChapitreB2017-2018.pdf.

[714] Wikipédia. General Leibniz rule 4-jan-2024. URL https://en.wikipedia.org/wiki/General_Leibniz_rule.

[715] Calvin Khor. multiindex partial derivative higher order chain rule 2024-jan-03. URL https://math.stackexchange.com/q/4837970.

[716] danarmk. Topologie sur D(Omega) URL https://github.com/LaurentClaessens/mazhe/issues/87#issuecomment-1731779089.

[717] Daniel Li. Notions sur les distributions URL http://labomathlens.free.fr/Liens/AF/distrib.pdf.

[718] Claude Zuily. Éléments de distributions et d'équations aux dérivées partielles.

[719] Sabin Lessard. Cours de processus stochastiques URL http://www.dms.umontreal.ca/~lessards/ProcessusStochastiquesLessard2014.pdf.

[720] Peter Young. Singular Fourier transforms and the Integral Representation of the Dirac Delta Function

[721] Daniel Choï. Méthode des éléments-finis par l'exemple URL http://meca.unicaen.fr/~choi/pdf/cours-mef.pdf.

[722] Michael Reiter and Arthur Schuste. Fourier Transform and Sobolev Spaces URL http://www.mat.univie.ac.at/~stein/lehre/SoSem08/sobolev_fourier.pdf.

[723] Norbert Heuer. LTCC module computation methods URL http://www.ltcc.ac.uk/courses/AppliedComputationalMethods/ACM2007Chap1.pdf.

[724] Arkhnor. Densité des fonctions Cinfty à support compact URL http://www.ilemaths.net/sujet-densite-des-fonctions-c-infty-a-support-compact-587937.html.

[725] Gilles Leborgne. Complément : espaces de Sobolev fractionnaires URL http://ws.isima.fr/~leborgne/IsimathDistributions/sobolev_frac.pdf.

[726] Richard B. Melrose. Lecture notes for 18.155, Fall 2002 URL https://ocw.mit.edu/courses/mathematics/18-155-differential-analysis-fall-2004/lecture-notes/section10.pdf.

[727] Julien Vovelle. Équations différentielles, cours no 2 URL http://math.univ-lyon1.fr/~vovelle/2Cours.pdf.

[728] A. Munnier. Théorie des équations différentielles ordinaires URL http://www.iecn.u-nancy.fr/~munnier/files/cours_edo.pdf.

[729] Olivier Debarre. Équations différentielles URL http://www.math.ens.fr/~debarre/EquaDiff.pdf.

[730] Sylvie Benzoni-Gavage. Propagation d'ondes URL http://math.univ-lyon1.fr/~benzoni/Ondes.pdf.

[731] Sonia Fliss. L'équation de transport à coefficients variables URL http://perso.ensta-paristech.fr/~fliss/Sonia_Fliss_web_page/Enseignement_files/Amphi3_2014.pdf.

[732] Sylvie Benzoni-Gavage. Équations différentielles ordinaires URL http://math.univ-lyon1.fr/~benzoni/EDO-M1.pdf.

[733] Wikipédia. Flot (mathématiques) — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Flot_(math%C3%A9matiques)&oldid=134130419.

[734] Alexander Grigor'yan. Ordinary differential equation URL https://www.math.uni-bielefeld.de/~grigor/odelec2009.pdf.

[735] Christopher P. Grant. Theory of ordinary differential equation URL http://www.math.pitt.edu/~bard/bardware/classes/2920/Grant_4july2007.pdf.

[736] John B.. Picard-Lindelöf theorem with parameter : what regularity ? URL http://math.stackexchange.com/questions/2126539/picard-lindelöf-theorem-with-parameter-what-regularity.

[737] Kevin Quirin. Développements URL http://kevin.quirin.free.fr/Trucs/agreg/developpements.pdf.

[738] Wikipédia. Équations de Lotka-Volterra — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quations_de_Lotka-Volterra&oldid=113756354.

[739] Wikipédia. Équation de Riccati — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89quation_de_Riccati&oldid=162760831.

[740] Predrag Cvitanović. Partial differential equations URL http://www.cns.gatech.edu/~predrag/courses/PHYS-6124-12/StGoChap6.pdf.

[741] Ralph Chill. Introduction aux équations aux dérivées partielles URL http://www.math.univ-metz.fr/~chill/edp.pdf.

[742] Gianluca Bontempi. Un cours d'analyse numérique et de matlab <https ://-web.archive.org/web/20090203224115/

[743] Wikipedia. Two's complement — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Two%27s_complement&oldid=706335033.

[744] Python Software Foundation. Floating Point Arithmetic : Issues and Limitations URL https://docs.python.org/3/tutorial/floatingpoint.html.

[745] Wikipédia. IEEE 754 — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=IEEE_754&oldid=120234746.

[746] Comprendre les nombres à virgules flottantes URL http://blog.netinfluence.ch/2009/09/24/comprendre-les-nombres-a-virgule-flottante/.

[747] Harald Schmidt. IEEE754 Converter URL http://www.h-schmidt.net/FloatConverter/IEEE754.html.

[748] David Goldberg. What Every Computer Scientist Should Know About Floating-Point Arithmetic URL http://docs.oracle.com/cd/E19957-01/806-3568/ncg_goldberg.html#11655.

[749] Emmanuel Frénod. Calcul scientifique URL http://web.univ-ubs.fr/lmam/frenod/IMG/DocEtudiant/MTH1504/calculscientifique.pdf.

[750] Franck Boyer. Agrégation externe de mathématiques. Analyse numérique URL http://www.math.univ-toulouse.fr/~fboyer/_media/enseignements/agreg/cours_an_agreg_fboyer_2014.pdf.

[751] Antoine Chambert-Loir. Autour de la méthode de Newton URL http://perso.univ-rennes1.fr/antoine.chambert-loir/2005-06/agreg/newton.pdf.

[752] Wikipédia. Vitesse de convergence — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Vitesse_de_convergence&oldid=110252913.

[753] Martin Pinto Campos and Laurence Halpern. Approximations polynomiales de fonctions et et des intégrales URL http://documents.lamacs.fr/cours/macs1/cours-macs1-pinto1.pdf.

[754] Ming Yang. Matrix decomposition URL http://users.eecs.northwestern.edu/~mya671/files/Matrix_YM_.pdf.

[755] Benjamin Ambrosio. Chapitre 3 : Méthodes directes de résolution du système $Ax = b$ URL http://lmah.univ-lehavre.fr/~ambrosio/CoursMN/Chapitre3.pdf.

[756] Emmanuel Frénod. Méthodes directes de résolution des systèmes linéaires URL http://web.univ-ubs.fr/lmam/frenod/IMG/DocEtudiant/MTH1504/methodes-directes.pdf.

[757] Christian Lécot. Analyse numérique matricielle URL https://www.lama.univ-savoie.fr/~lecot/data/M1_ANM.pdf.

[758] Wikipédia. Matrice à diagonale dominante — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Matrice_%C3%A0_diagonale_dominante&oldid=132268176.

[759] Habib Joulak. Méthodes des différences finies en élasticité URL https://ori-nuxeo.univ-lille1.fr/nuxeo/site/esupversions/0e05f3d2-fe12-4d67-b74f-f622869539f2.

[760] Wikipédia. Variable aléatoire à densité 5 décembre 2022. URL https://fr.wikipedia.org/wiki/Variable_al%C3%A9atoire_%C3%A0_densit%C3%A9.

[761] Loi d'une variable aléatoire 5 aout 2022. URL https://www.lesmath.com/2022/08/loi-d-une-variable-aleatoire.html.

[762] Philippe Briand. Probabilités de base URL http://www.lama.univ-savoie.fr/~briand/proba/g12_cours.pdf.

[763] Vincent Bansaye. Variables aléatoires, espérance, indépendance URL http://www.cmap.polytechnique.fr/~bansaye/CoursTD2.pdf.

[764] Frédéric Plachet. Modèles de durée Aout 2021. URL https://www.planchet.net/EXT/ISFA/fp-isfa.nsf/0/1430AD6748CE3AFFC1256F130067B88E/$FILE/Seance2.pdf.

[765] Mark McClure. Differentiable function has measurable derivative? Jan 15 2014. URL https://math.stackexchange.com/q/639690.

[766] Wikipédia. Joint probability distribution 24 aout 2024. URL https://en.wikipedia.org/wiki/Joint_probability_distribution.

[767] Maverick Meerkat and Did. Expected value as integral of survival function URL https://math.stackexchange.com/questions/642021/expected-value-as-integral-of-survival-function.

[768] Gabriel Lepetit. Processus de Poisson URL https://www-fourier.ujf-grenoble.fr/~lepetitg/agregation/Processus_de_Poisson.pdf.

[769] Alfio Marazzi. URL http://www.iumsp.ch/Unites/us/Alfio/polybiostat/ch06.pdf.

[770] . Variables aléatoires URL http://www.cpgedupuydelome.fr/IMG/pdf/10_-_variables_aleatoires_cours_complet.pdf.

[771] JMoravitz. Is this probability statement correct? 28 oct 2022. URL https://math.stackexchange.com/q/4563280.

[772] Pierre-Loïc Méliot. Notion de conditionnement URL https://www.math.u-psud.fr/~meliot/agreg/condition.pdf.

[773] Cédric Boutillier. URL https://github.com/LaurentClaessens/mazhe/issues/16#issue-180279120.

[774] Automaths. Paradoxe des deux enfants – Episode 2 ! URL https://automaths.blog/2018/08/25/paradoxe-des-deux-enfants-episode-2/.

[775] Wikipédia. Inégalité de Jensen — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=In%C3%A9galit%C3%A9_de_Jensen&oldid=90235295.

[776] Gérard Letac. Calcul des probabilités, Deug 2ième année URL http://les.mathematiques.free.fr/pdf/proba.zip.

[777] Chafaï Djalil. Probability space 2019. URL https://djalil.chafai.net/blog/2019/11/11/probability-space/.

[778] Daniel Saada. Note sur le théorème de transfert pour les variables aléatoires réelles URL http://www.les-mathematiques.net/phorum/file.php?12,file=17435,filename=th_du_tranfert.pdf.

[779] Wikipédia. Variable aléatoire à densité 8 mai 2022. URL https://fr.wikipedia.org/w/index.php?title=Variable_al%C3%A9atoire_%C3%A0_densit%C3%A9&oldid=193504954.

[780] Christophe Sabot. Différents types de convergence de suites de v.a. URL http://math.univ-lyon1.fr/~sabot/Convergence.pdf.

[781] Antonin Macé and Thomas Pradeau. Exposé de maîtrise. Estimations de densités de probabilité URL http://www.eleves.ens.fr/home/amace/travaux/exposemaitrise.pdf.

[782] Wikipedia. Continuous mapping theorem — Wikipedia, The Free Encyclopedia URL http://en.wikipedia.org/w/index.php?title=Continuous_mapping_theorem&oldid=545523937.

[783] Wikipédia. Inégalité de Markov — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=In%C3%A9galit%C3%A9_de_Markov&oldid=100208179.

[784] Math1000. Sum of exponential random variables ? Jul 5, 2020. URL https://math.stackexchange.com/q/146653.

[785] Charles Suquet. Simulation URL http://math.univ-lille1.fr/~suquet/ens/Agr/simul06.pdf.

[786] Yoann Gelineau. Vecteurs gaussiens URL http://math.univ-lyon1.fr/~gelineau/files/vecteurs_gaussiens.pdf.

[787] Scilab. rand URL https://help.scilab.org/docs/5.4.0/fr_FR/rand.html.

[788] Wikipédia. Nombre normal — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Nombre_normal&oldid=89848922.

[789] Wikipédia. Théorème de Glivenko-Cantelli — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Th%C3%A9or%C3%A8me_de_Glivenko-Cantelli&oldid=90841113.

[790] Vincent Rivoirard and Gilles Stoltz. Estimation de densité de probabilités URL http://www.math.ens.fr/statenaction/PDF/Densite-Principal.pdf.

[791] Nils Berglund. Chaînes de Markov URL http://www.univ-orleans.fr/mapmo/membres/berglund/markov.pdf.

[792] Arnaud Guyader. Espérance conditionnelle et chaines de Markov URL https://perso.univ-rennes2.fr/system/files/users/jegou_n/EsperanceConditionnelle.pdf.

[793] Pierre Brémaud. Basic Markov Chains 9 décembre 2015. URL https://www.di.ens.fr/~busic/cours/M2amis/markov.pdf.

[794] David Madore. Réflexions sans intérêt sur les files d'attente 14 sept 2019. URL http://www.madore.org/~david/weblog/d.2021-09-14.2696.files-attente.html#d.2021-09-14.2696.

[795] Charles Suquet. Martingales 2009-2010. URL https://math.univ-lille1.fr/~suquet/Polys/Mrt.pdf.

[796] Massimiliano Gubinelli. Comportement asymptotique des martingales URL http://www.ceremade.dauphine.fr/~mgubi/e1112/pd3.pdf.

[797] Jean Lacroix and Pierre Priouret. Probabilités approfondies URL http://www.proba.jussieu.fr/cours/processus.pdf.

[798] Wikipedia. Optional stopping theorem — Wikipedia, The Free Encyclopedia URL http://en.wikipedia.org/w/index.php?title=Optional_stopping_theorem&oldid=547547174.

[799] Ari Davidovsky. Martingales, Stopping Times, and the Optional Stopping Theorem URL https://math.uchicago.edu/~may/REU2019/REUPapers/Davidovsky.pdf.

[800] Nizar Touzi. Martingales en temps discret et chaine de Markov URL http://www.cmap.polytechnique.fr/~touzi/MAP432-Poly.pdf.

[801] Wikipédia. Stopped process 6 juillet 2024. URL https://en.wikipedia.org/wiki/Stopped_process.

[802] Nasser Saheb. Interarrivée URL https://dept-info.labri.fr/ENSEIGNEMENT/formation-a-distance/cibermiage-processus/c1.7/ch02/seq07/sequence.htm.

[803] Yoann Gelineau. Processus de Poisson URL http://math.univ-lyon1.fr/~gelineau/devagreg/Processus_Poisson.pdf.

[804] Benoît Cadre. Le processus de Poisson 2008-2009. URL https://perso.univ-rennes1.fr/jean-christophe.breton/agreg/AGREG/COURS/Agreg_Poisson.pdf.

[805] Wikipédia. Étoile de Kleene — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=%C3%89toile_de_Kleene&oldid=149113546.

[806] Wikipédia. Onde gravitationnelle — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Onde_gravitationnelle&oldid=185499357.

[807] Wikipédia. Expérience de Michelson et Morley — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Exp%C3%A9rience_de_Michelson_et_Morley&oldid=176995513.

[808] Zach Weinersmith. URL http://www.smbc-comics.com/comic/2008-02-21.

[809] Gro-Tsen. URL https://bsky.app/profile/gro-tsen.bsky.social/post/3kxzv27uwno2s.

[810] Umberto P.. Why Pythagorean theorem is all about 2 ? 2024-09-10. URL https://math.stackexchange.com/q/4969768.

[811] Numberphile. 5040 and other Anti-Prime Numbers - Numberphile URL https://www.youtube.com/watch?v=2JM2oImb9Qg.

[812] Denis Serre. Matrices

[813] S. Helgason. Differential geometry, Lie groups, and symmetric spaces

[814] Michael Barr and Charles Wells. Toposes, Triples and Theories <http ://www.case.edu/artsci/math/wells/pub/ttt.html>.

[815] Bob Coecke. Introducing categories to the practicing physicist 0808.1032v1

[816] Nicolai Reshetikhin. Topological Quantum Field Theory and Related Algebraic Structures URL http://staff.science.uva.nl/~nresheti/masterclass-08.html.

[817] . Locally convex topological vector space — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Locally_convex_topological_vector_space&oldid=1068316371.

[818] Wikipedia. Pole (complex analysis) — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Pole_(complex_analysis)&oldid=702671984.

[819] G.Landi. An Introduction to Noncommutative spaces and their Geometry hep.th/9701078

[820] Wikipédia. Partition de l'unité — Wikipédia, l'encyclopédie libre URL http://fr.wikipedia.org/w/index.php?title=Partition_de_l%27unit%C3%A9&oldid=122365411.

[821] Marius Crainic. chapter 5 : partition of unity URL https://webspace.science.uu.nl/~crain101/topologie11/chapter5.pdf.

[822] R. Wallach. Construction of Haar measure URL https://mathweb.ucsd.edu/~nwallach/haarmeasure.pdf.

[823] Laurent Claessens. Use of Zorn lemma in a proof of Kakutani theorem URL https://math.stackexchange.com/q/4402080.

[824] Dragan Miličić. Notes on representations of compact groups URL https://www.math.utah.edu/~milicic/Math_6260/compact.pdf.

[825] asteroid (17536). modular function 2013. URL https://planetmath.org/ModularFunction.

[826] Wikibooks. Measure Theory/Riesz' representation theorem — Wikibooks, The Free Textbook Project URL https://en.wikibooks.org/w/index.php?title=Measure_Theory/Riesz%27_representation_theorem&oldid=4042731.

[827] S. Kumaresan. Riesz-Markov Representation Theorem URL https://4dspace.mtts.org.in/expository-article-download.php?ai=160.

[828] Desvl. The Riesz-Markov-Kakutani Representation Theorem URL https://desvl.xyz/2020/09/19/The-Riesz-Markov-Kakutani-Representation-Theorem/.

[829] Robin Graham and Josh Swanson and Debbie Matthews. (Compact) Lie Groups and Representation Theory Lecture Notes 2015. URL https://courses.washington.edu/mnotes/liegroups/notes.pdf.

[830] Wang Zuoquin. Smooth functions URL http://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds/Notes/Lec03.pdf.

[831] Jean Gallier. Notes on differential geometry and Lie Groups URL https://www.cis.upenn.edu/~cis610/diffgeom2.pdf.

[832] Zuoquin Wang. 16F URL https://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds.

[833] José Carlos Santos. What is the topology on a Lie algebra $\mathfrak{g}$? URL https://math.stackexchange.com/q/3169462.

[834] Zuoqin Wang. 16F-Manifols URL http://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds/.

[835] Math Curator Zanachan. Manifolds with boundary (part 4)- Smooth Invariance of Domains URL https://www.youtube.com/watch?v=cqO6bLLhxqc.

[836] Zuoquin Wang. Differential manifold URL http://staff.ustc.edu.cn/~wangzuoq/Courses/18F-Manifolds/.

[837] Andrea Rincon. The Frobenius theorem URL https://www.mathi.uni-heidelberg.de/~lee/Andrea_Rincon.pdf.

[838] Shoshichi Kobayashi and Katsumi Nomizu. Foundations of differential geometry. Vol. I Reprint of the 1963 original, A Wiley-Interscience Publication

[839] . Immersion (mathematics) — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Immersion_(mathematics)&oldid=864823918.

[840] . Embedding — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Embedding&oldid=876329076.

[841] Patrick Girardet. Immersions and Embeddings URL https://mathweb.ucsd.edu/~eizadi/250A-2019/Patrick-Girardet.pdf.

[842] Johachim Wehler. Lie groupe URL https://www.mathematik.uni-muenchen.de/~wehler/LieAlgebrasLieGroupsScript.pdf.

[843] Michael Kunzinger. 1.1. submanifolds URL https://www.mat.univie.ac.at/~mike/teaching/ws1415/submf.pdf.

[844] Marc Henneaux. Groupes et représentations : groupe des rotations à 3 dimensions, groupe de Lorentz et groupe de Poincaré URL http://www.solvayinstitutes.be/pdf/Marc/SO3SO3_1_18112014.pdf.

[845] wikipedia . Lie algebra representation 2022. URL https://en.wikipedia.org/wiki/Lie_algebra_representation.

[846] . Lie bracket of vector fields — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Lie_bracket_of_vector_fields&oldid=859293850.

[847] Wikipedia. Semi-simple operator — Wikipedia, The Free Encyclopedia URL http://en.wikipedia.org/w/index.php?title=Semi-simple_operator&oldid=544555496.

[848] Hans Samelson. Notes on Lie algebras URL http://www.math.cornell.edu/~hatcher/Other/Samelson-LieAlg.pdf.

[849] Arthur A. Sagle and Ralph E. Walde. Introduction to Lie groups and Lie algebras Pure and Applied Mathematics, Vol. 51

[850] F. Wisser. Classification of complex and real semi-simple Lie algebras URL http://www.mat.univie.ac.at/~cap/files/wisser.pdf.

[851] C. Schomblond. Électrodynamique classique URL http://homepages.ulb.ac.be/~cschomb/notes.html.

[852] David Sénéchal. Mécanique quantique URL http://www.physique.usherbrooke.ca/senechal/doc/PHY731.pdf.

[853] Dragan Miličić. Lectures on Lie groups URL http://www.math.utah.edu/~milicic/Eprints/lie.pdf.

[854] Jürgen Berndt. Lie group actions on manifolds URL http://www.mth.kcl.ac.uk/~berndt/sophia.pdf.

[855] G. Hochschild. The structure of Lie groups

[856] Jean-Pierre Serre. Algèbres de Lie semi-simples complexes

[857] J. F. Cornwell. Group theory in physics, volume 2

[858] Shlomo Sternberg. Lie algebras URL http://www.math.harvard.edu/~shlomo/index.html.

[859] Anthony W. Knapp. Representation theory of semi-simple groups An overview based on examples

[860] Thomas R. Covert. The Xartan matrix of a root system URL http://ocw.mit.edu/courses/mathematics/18-06ci-linear-algebra-communications-intensive-spring-2004/projects/final_project_5.pdf.

[861] Brian G. Wybourne. Classical groups for physicists Very complete discussion about computation of root and weight space ; many examples on $\mathfrak{su}(3)$.

[862] Robert N Cahn. Semi-simple Lie algebras and their representations http ://phyweb.lbl.gov/ rncahn/www/liealgebras/texall.pdf

[863] James E. Humphreys. Reflexion groups and Coxeter groups URL http://books.google.com/books?id=ODfjmOeNLMUC&hl=en.

[864] Nicolas Boulanger and Sophie Buyl and Francis Dolan. Semi-simple Lie algebras and representations URL http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveI/Modave2005.pdf.

[865] Anthony W. Knapp. Lie groups beyond an introduction

[866] James E. Humphreys. Introduction to Lie algebras and representation theory Second printing, revised

[867] Philipp Fahr Dierk. Enveloping algebras of finite dimensional nilpotent Lie algebras URL http://www.math.uni-bielefeld.de/~philfahr/enveloping/enveloping.html.

[868] D. P. Želobenko. Compact Lie groups and their representations Translated from the Russian by Israel Program for Scientific Translations, Translations of Mathematical Monographs, Vol. 40

[869] New groups from old URL http://www.oup.co.uk/pdf/0-19-853548-1.pdf.

[870] Zuoqin Wang. URL https://staff.ustc.edu.cn/~wangzuoq/Courses/13F-Lie/Lie.html.

[871] . Lie group — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Lie_group&oldid=876239769.

[872] J. J. Duistermaat and J. A. C Kolk. Lie groups

[873] Lie groups, Lie algebras and the Exponential Map URL http://www.cis.upenn.edu/~cis610/cis61005sl8.pdf.

[874] Sigurdur Helgason. Introduction to Lie groups URL https://ocw.mit.edu/courses/mathematics/18-755-introduction-to-lie-groups-fall-2004/lecture-notes/chap2_topic1to7.pdf.

[875] Zuogin Wang. URL http://staff.ustc.edu.cn/~wangzuoq/Courses/13F-Lie/Notes/Lec11.pdf.

[876] Wang Zuoqin. LECTURE 5 : LIE GROUPS AND THEIR LIE ALGEBRAS URL https://staff.ustc.edu.cn/~wangzuoq/Courses/13F-Lie/Notes/Lec%2005.pdf.

[877] . Closed-subgroup theorem — Wikipedia, The Free Encyclopedia URL https://en.wikipedia.org/w/index.php?title=Closed-subgroup_theorem&oldid=876758988.

[878] orangeskid. Is a smooth Lie group an analytic Lie group? URL https://math.stackexchange.com/questions/4231641/is-a-smooth-lie-group-an-analytic-lie-group.

[879] joshphysics. Integrating the generator of the infinitesimal special conformal transformation URL http://physics.stackexchange.com/questions/121920/integrating-the-generator-of-the-infinitesimal-special-conformal-transformation.

[880] Gye-Seon Lee. Review of Differential Geometry URL https://www.mathi.uni-heidelberg.de/~lee/GonencSS16.pdf.

[881] Christian Kassel. Quantum groups

[882] Jacques Faraut. Groupes et algèbres de Lie URL http://les-mathematiques.u-strasbg.fr/phorum/download.php/2,5426/GroupesdeLie.ps](http://les-mathematiques.u-strasbg.fr/phorum/download.php/2,5426/GroupesdeLie.ps.

[883] Paul ES Wormer. Angular momentum theory URL http://www.theochem.ru.nl/~pwormer/angmom.pdf.

[884] Vicror Piercey. Verma modules URL http://math.arizona.edu/~vpiercey/VermaModules.pdf.

[885] Singh Anupam. Representation theory of Lie algebras URL http://www.isibang.ac.in/~statmath/conferences/gt/lie_algebra.pdf. Workshop on Group Theory 18 Dec 2006 - 5 Jan 2007, Bangalore

[886] Nathan Jacobson. Basic algebra. II

[887] Thomas Timmermann. An invitation to quantum groups and duality URL http://dx.doi.org/10.4171/043. From Hopf algebras to multiplicative unitaries and beyond

[888] Roland Vergnioux. Haar integrals on finite and compact quantum groups URL http://www.math.unicaen.fr/~vergnioux/.

[889] Leonid I. Korogodski and Yan S. Soibelman. Algebras of functions on quantum groups : part I

[890] T. Tjin. An introduction to quantized Lie groups and algebras hep.th/9111043

[891] Anthony Giaquinto and James J. Zhang. Bialgebra actions, twists, and universal deformation formulas hep-th/9411140

[892] Pierre Bieliavsky and Philippe Bonneau and Francesco D'Andrea and Victor Gayral and Yoshiaki Maeda and Yannick Voglaire. Drinfel'd twists and multiplicative unitaries for Kahlerian Lie groups Not published yet

[893] Rolf Farnsteiner. Lie algebras with coalgebra splitting URL http://www.mathematik.uni-bielefeld.de/~rolf/Lie-Co-Split.pdf.

[894] S. Helgason. Differential geometry and symmetric spaces

[895] Xavier Bekaert and Nicolas Boulanger. The unitary representations of the Poincaré group in any spacetime dimension URL http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf. Lecture notes for the second edition of the Modave summer school.

[896] F. A. Dolan. Character Formulae and Partition Functions in Higher Dimensional Conformal Field Theory hep.th/0508031v2

[897] David Renard. Groupes et représentations URL http://www.cmls.polytechnique.fr/perso/renard/Groupes.pdf.

[898] Frédéric Paulin. Introduction aux groupes de Lie pour la physique URL https://www.math.u-psud.fr/~paulin/notescours/cours_centrale.pdf.

[899] D. B. Westra. SU(2) and SO(3) URL https://www.mat.univie.ac.at/~westra/so3su2.pdf.

[900] S. Sternberg. Group theory and physics

[901] Equations alignment using amsmath Package URL https://www.resurchify.com/latex_tutorial/latex_alignment.php.

[902] W. A. Stein and . Sage Mathematics Software (Version 4.3.3) URL http://www.sagemath.org.

[903] S. Ferrara and C. Frønsdal. Conformal fields in higher dimensions hep.th/0006009

[904] J. Engquist and P. Sundell and L. Tamassia. Singleton strings hep.th/0701081v1

[905] J. Engquist and P. Sundell and L. Tamassia. On singleton composites in non-compact WZW models hep.th/0701051

[906] L. Brinka and R. R. Metsaevb and M. A. Vasilievd. How massless are massless fields in AdS_d hep.th/0005136

[907] J. A. Teschner. Weak quasitriangular Quasi-Hopf algebra structure of minimal models hep.th/9512096

[908] Wikipedia. Vector bundle — Wikipedia, The Free Encyclopedia 17 octobre 2023. URL https://en.wikipedia.org/w/index.php?title=Vector_bundle&oldid=1171634676.

[909] Allen Hatcher. Vector bundles and K-theory URL http://www.math.cornell.edu/~hatcher/VBKT/VBpage.html.

[910] Wang Zuoquin. Differential forms URL http://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds/Notes/Lec17.pdf.

[911] Amitai Yuval. Exterior derivative well-defined 2017-sept-30. URL https://math.stackexchange.com/q/2451863.

[912] Zuoqin Wang. LECTURE 22 : THE EXTERIOR DERIVATIVE URL http://staff.ustc.edu.cn/~wangzuoq/Courses/21F-Manifolds/Notes/Lec22.pdf.

[913] Kevin Guo. exterior derivative in local coordinates 2022-08-08. URL https://math.stackexchange.com/q/4508419.

[914] rmdmc89. Formula for the exterior derivative 2016-10-18. URL https://math.stackexchange.com/q/1973454.

[915] Gregory L. Naber. Topology, geometry, and gauge fields Interactions

[916] J. Dieudonné. Éléments d'analyse. Tome IV. Chapitres XVIII à XX Nouveau tirage, Cahiers Scientifiques, Fasc. 34

[917] Mathew A. Johnson. The Hodge decomposition theorem URL http://www.math.uiuc.edu/~mjohns51/Presentations/Hodge_Theorem.pdf.

[918] Sylvain Cappell and Dennis DeTurck and Herman Gluck and Edward Y. Miller. Cohomology of harmonic forms on Riemannian manifolds with boundary URL http://www.math.upenn.edu/~deturck/papers/har-coho-7.pdf.

[919] Vladimir G. Ivancevic and Tijana T. Ivancevic. Undergraduate lecture notes in De Rham-Hodge theory 0807.4991

[920] Wang Zuoquin. Orientation and integration URL http://staff.ustc.edu.cn/~wangzuoq/Courses/16F-Manifolds/Notes/Lec18.pdf.

[921] Kelvin Lois. Every connected orientable smooth manifold has exactly two orientations, Lee Proposition 15.9 2019. URL https://math.stackexchange.com/q/3153700.

[922] Frédéric Paulin. Géométrie différentielle élémentaire 2006-2007. URL https://www.imo.universite-paris-saclay.fr/~frederic.paulin/notescours/cours_geodiff.pdf.

[923] Laurent Claessens. commute two sums when defining integral of differential manifold URL https://math.stackexchange.com/q/4917916.

[924] Ben Steffan. Diffeomorphism between half plane and plane URL https://math.stackexchange.com/q/4912999.

[925] Stephen Montgomery-Smith. Upper Half Plane ; not diffeomorphic to the whole plane URL https://math.stackexchange.com/q/572785.

[926] Math Curator Zanachan. Manifolds with Boundary (part 3)- Diffeomorphisms preserve boundary and interior URL https://www.youtube.com/watch?v=ABnMezxfw6M.

[927] anonyme. Manifolds with boundary URL https://math.ucr.edu/~res/math260s10/manwithbdy.pdf.

[928] jws. integral of a vector field in $\mathbb{R}^n$ URL http://math.stackexchange.com/questions/1764223/integral-of-a-vector-field-in-mathbbrn.

[929] James Mracek. Integration over manifolds URL http://www.math.toronto.edu/mat1300/integration.12.pdf.

[930] Eckhard Meinrenken. Group action on manifold URL http://www.math.toronto.edu/mein/teaching/action.pdf.

[931] J. E. Marsden. <http ://www.cds.caltech.edu/ marsden>.

[932] C. Schomblond. Théorie hamiltonienne à contraintes URL http://homepages.ulb.ac.be/~cschomb/notes.html.

[933] Alan Weinstein. Lagrangian mechanics and groupoids URL http://math.berkeley.edu/~alanw/. Appeared in Mechanics Day, Fields Institute Proc. v. 7, AMS, 1995.

[934] P. Dazord A. Coste and A. Weinstein. Groupoïdes symplectiques (in french) URL http://math.berkeley.edu/~alanw/. Appeared in Publ. Dept. Math. Univ. Claude-Bernard Lyon I, 1987.

[935] Antony Wassermann. Operators on Hilbert space URL http://www.dpmms.cam.ac.uk/~ajw/.

[936] Jean-Yves Girard. Introduction aux algèbres d'opérateurs I : Des espaces de Hilbert aux algèbres stellaires URL http://recherche.ircam.fr/equipes/repmus/mamux/AlgOpGirard.pdf.

[937] A. D. Andrew and W. L. Green. Spectral theory of operators on Hilbert spaces URL http://people.math.gatech.edu/~green/Spring2002/Ma7334/spectral-theory.pdf.

[938] Olaf Post. Periodic manifolds, spectral gaps, and eigenvalues in gaps URL http://www.math.hu-berlin.de/~post/papers/promo-e.pdf.

[939] Paul Garrett. Operators on Hilbert spaces URL http://www.math.umn.edu/~garrett/m/fun/Notes/04a_ops_hsp.pdf.

[940] Michael F. Whittaker. Fredholm modules and spectral triples in noncommutative geometry URL http://michaelwhittaker.ca/STandFredholmNCG.pdf.

[941] Joseph C. Várilly. An introduction to noncommutative geometry physics/9709045

[942] Vaughan FR Jones. von Neumann algebras URL http://math.berkeley.edu/~reb/courses/209/VonNeumann2.pdf.

[943] François Trèves. Topological vector spaces, distributions and kernels

[944] J. Dieudonné. Éléments d'analyse. Tome III : Chapitres XVI et XVII

[945] A. A. Kirillov. Elements of the theory of representations Translated from the Russian by Edwin Hewitt, Grundlehren der Mathematischen Wissenschaften, Band 220

[946] J. Dieudonné. Éléments d'analyse. Tome II : Chapitres XII à XV

[947] André Fortin and André Garon. Les éléments finis : de la théorie à la pratique URL http://www.giref.ulaval.ca/~afortin/cours_elements_finis/documents/notes_elements_finis.pdf.

[948] Gerd Grubb. Boundary operators on Sobolev spaces URL http://www.math.ku.dk/~grubb/notes/epd3.pdf.

[949] V. Maslov. Méthodes opératorielles Translated from the Russian by Djilali Embarek. An English translation by V. Golo, N. Kulman and G. Voropaeva exists under the title Operational methods.

[950] Michael E. Taylor. Pseudodifferential operators

[951] J. Dieudonné. Éléments d'analyse. Tome VII. Chapitre XXIII. Première partie Cahiers Scientifiques, Fasc. XL

[952] M. A. Rieffel. Deformation quantization for actions of $\mathbb{R}^d$

[953] Alain Connes. Noncommutative geometry URL http://www.alainconnes.org/downloads.html.

[954] Vincent Acary. EDP elliptiques. Laplace et Poisson URL https://bipop.inrialpes.fr/people/acary/Teaching/Chapitre4.pdf.

[955] Frédéric Paulin. Topologie algébrique élémentaire URL http://www.dma.ens.fr/edition/NotesCours/cours_ENS_topalg.pdf.

[956] Ottmar Loos. Symmetric spaces. I : General theory

[957] S. Kobayashi and K. Nomizu. Foundation of differential geometry

[958] Martin Kerin and David Wraith. Homogeneous metrics on spheres URL http://www.maths.tcd.ie/pub/ims/bull51/M5102.pdf.

[959] Jacques Dixmier. Les C^*-algèbres et leurs représentations Reprint of the second (1969) edition

[960] P. Bieliavsky. Semisimple symplectic symmetric spaces

[961] P. Bieliavsky. Strict quantization of solvable symmetric spaces math.QA/0010004

[962] Alain Connes and Matilde Marcolli. Noncommutative Geometry, Quantum Fields and Motives URL http://www.alainconnes.org/downloads.html.

[963] D. V. Vassilevich. Heat kernel expansion ; user's manual hep.th/0306138

[964] Thomas P. Branson and Peter B. Gilkey. Residues of the eta function for an operator of Dirac type with local boundary conditions URL http://www.math.uiowa.edu/ftp/branson/paper27.ps.

[965] Ali H. Chamseddine and Alain Connes. Quantum gravity boundary terms from spectral action 0705.1786

[966] P. Bieliavsky. Équation de champs dans l'espace de Minkowski et son compactifié

[967] H. Blaine Lawson and Marie-Louise Michelsohn. Spin geometry

[968] Joseph C. Várilly and Pawel Witkowski. Dirac Operators and Spectral Geometry URL http://toknotes.mimuw.edu.pl/sem3/files/Varilly_dosg.pdf.

[969] Blake Mellor. Spin$^{\mathbb{C}}$-manifolds URL http://math.berkeley.edu/~alanw/ps/mellor.ps.

[970] Thomas P. Branson and Peter B. Gilkey. Residues of the eta function for an operator of Dirac type URL http://www.math.uiowa.edu/ftp/branson/paper24.ps.

[971] W. Fulton and J. Harris. Representation theory

[972] C. Chevalley. The Alegraic theory of spinors A deep discussion about Clifford algebras.

[973] Nick Brönn. Dirac operators A good explanation of the links between Clifford algebra and the problem of "square root" for the Laplace operator. It is done for the Euclidean metric and in Riemannian spaces. I found it on the internet but I don't remember where.

[974] S.Weinberg. The quantum theory of fields

[975] M. E. Peskin and D. V.Schroeder. An introduction to quantum field theory Title says what it is.

[976] F.Schwabl. Advanced quantum mechanics

[977] G.Svetlichny. Preparation for gauge theory Almost all you need —and wish— to know about differential geometry in order to understand the gauge theories of the mathematical physics : Lie groups, fiber bundles, connections, gauge transformation, spin bundle and so on, math-ph/9902027

[978] Sami Virtanen. Quantum mechanics 3 URL http://www.fyslab.hut.fi/kurssit/Tfy-44.135/.

[979] L.Landau and E.Lifchitz. Physique Théorique Théorie des champs

[980] A. Collinucci and A. Wijns. Topology of fiber bundles and instantons URL http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf.

[981] Paul Ginsparg. Applied conformal field theory URL http://arxiv.org/abs/hep-th/9108028v1.

[982] A. N. Schellekens. Conformal field theory URL http://www.nikhef.nl/~t58/CFT.pdf.

[983] joshphysics. Integrating the generator of the infinitesimal special conformal transformation URL http://physics.stackexchange.com/questions/121920/integrating-the-generator-of-the-infinitesimal-special-conformal-transformation.

[984] Stéphane Detournay. Aspects of two-dimensional conformal field theory URL http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveI/CoursPropre.pdf.

[985] Pierre Harpe and Vaughan Jones. An introduction to C^*-algebra URL http://www.unige.ch/math/biblio/preprint/cstar/liste.html.

[986] Jacques Dixmier. Traces sur les C-algebres URL http://www.numdam.org/article/AIF_1963__13_1_219_0.pdf.

[987] Klaus Thomsen. No precise title <www.imf.au.dk/kurser/analyse2/E04/thomaslassen.pdf>

[988] Claire Anantharaman-Delaroche. On tensor products of group C^*-algebras and related topics URL http://www.univ-orleans.fr/mapmo/membres/anantharaman/publications/Lausanne07.pdf.

[989] S. L. Woroniwicz. Compact quantum groups URL http://fuw.edu.pl/~slworono/PDF-y/CQG3.pdf.

[990] William Gordon Ritter. Introduction to quantum group theory math.QA/0201080

[991] H. Tjerk Koelink. On $*$-representations of the Hopf $*$-algebra associated with the quantum group $\mathrm{U}_q(n)$ URL http://www.numdam.org/article/CM_1991__77_2_199_0.pdf.

[992] Johan Kustermans. Locally compact quantum groups URL http://staff.science.uva.nl/~riveen/seminar2008_files/Kustermans-LNM1185-2005.pdf.

[993] Maxime Rey. Algèbres de Hopf combinatoires URL http://join.inria.fr/attached-documents/Stages-2005-rapports/rapport-2005-rey.pdf.

[994] Ludwik Dąbrowski and Giovanni Landi and Andrzej Sitarz and Walter Suijlekom and Joseph C. Várilly. The Dirac operator on $\mathrm{SU}_q(2)$ math.QA/0411609v2

[995] Kazimierz Brągiel. The twisted SU(3) group. Irreducible $*$-representations of the C^*-algebra $C(\mathrm{S}_\mu\mathrm{U}(3))$ URL http://dx.doi.org/10.1007/BF00420012.

[996] K. R. Goodearl. Commutation relations for arbitrary quantum minors URL http://www.math.ucsb.edu/~goodearl/comm.rel.qminors.pdf.

[997] W. Pusz and S. L. Woronowicz. Twisted second quantization URL http://dx.doi.org/10.1016/0034-4877(89)90006-2.

[998] Vyjayanthi Chari and Andrew Pressley. A guide to quantum groups

[999] Serge Levendorskii and Yan Soibelman. Algebras of functions on compact quantum groups, Schubert cells and quantum tori URL http://projecteuclid.org/getRecord?id=euclid.cmp/1104203139.

[1000] Salvatore Triolo. A note on semifinite von Neumann algebra URL http://math.unipa.it/~salvo/Publication_files/bollettino.pdf.

[1001] V. Jones and V. S. Sunder. Introduction to subfactors

[1002] Jacques Dixmier. C^*-algebras Translated from the French by Francis Jellett, North-Holland Mathematical Library, Vol. 15

[1003] Fabio Cipriani. Noncommutative Dirichlet forms URL http://stubber.math-inf.uni-greifswald.de/algebra/qpt/cipriani-10jan2008.pdf.

[1004] Johan Kustermans and Stefaan Vaes. Locally compact quantum groups URL http://archive.numdam.org/article/ASENS_2000_4_33_6_837_0.pdf.

[1005] William Arveson. The canonical anticommutation relations URL http://web.math.berkeley.edu/~arveson/Dvi/car.pdf.

[1006] Fabio Cipriani. Dirichlet forms and Markovian semigroups on standard forms of von Neumann algebras URL http://dx.doi.org/10.1006/jfan.1996.3063.

[1007] Fabio Cipriani and Jean-Luc Sauvageot. Derivations as square roots of Dirichlet forms URL http://dx.doi.org/10.1016/S0022-1236(03)00085-5.

[1008] A. Ashtekar and J. Lewandowski. Background Independent Quantum Gravity : A Status Report gr-qc/0404018

[1009] J. Baez. Degenerate Solutions of General Relativity from Topological Field Theory gr-qc/9702051

[1010] J. Baez. An Introduction to Spin Foam Models of Quantum Gravity and BF Theory gr-qc/9905087v1

[1011] Cédric Bény. The Ashtekar Hamiltonian for general relativity URL http://ampc013.uwaterloo.ca/~cbeny/ashtekar.pdf.

[1012] Pierre Bieliavsky and Marianne Rooman and Philippe Spindel. Regular Poisson structures on massive non-rotating BTZ black holes hep-th/0206189

[1013] Pierre Bieliavsky and Stephane Detournay and Michel Herquet and Marianne Rooman and Philippe Spindel. Global geometry of the $2+1$ rotating black-hole hep.th/0306293v1

[1014] Máximo Bañados and Claudio Teitelboim and Jorge Zanelli. The lack hole in three-dimensional space-time hep-th/9204099v3

[1015] Máximo Bañados and Marc Henneaux and Claudio Teitelboim and Jorge Zanelli. Geometry of the $2+1$ black hole gr-qc/9302012v1

[1016] P. Bieliavsky and S. Detournay and P. Spindel and M. Rooman. Star products on extended massive non-rotating BTZ black holes hep-th/0403257

[1017] Pierre Bieliavsky and Stephane Detournay and Philippe Spindel and Marianne Rooman. Noncommutative locally anti-de Sitter black holes math.QA/0507157

[1018] Clément Hyvrier. Étude géométrique de certaines variétés localement anti-de Sitter : les trous noirs « BTZ » massifs et sans moment angulaire Mémoire de DEA, Université libre de Bruxelles, faculté des sciences, département de mathématiques

[1019] Laurent Claessens and Stephane Detournay. Solvable symmetric black hole in anti-de Sitter spaces math.DG/0510442

[1020] Pierre Bieliavsky and Yannick Voglaire and Laurent Claessens and Daniel Sternheimer. Quantized Anti de Sitter spaces and non-formal deformation quantizations of symplectic symmetric spaces 0705.4179v1

[1021] Laurent Claessens. Locally anti de Sitter spaces and deformation quantization. Ph.D. Thesis URL http://hdl.handle.net/2078.1/5354.

[1022] J. Faraut and G. Ólafsson. Causal semisimple symmetric spaces, the geometry and harmonic analysis URL http://www.math.lsu.edu/~olafsson/pdf_files/hobook.pdf.

[1023] J. Figueroa-O'Farrill and O. Madden and S. F. Ross and J. Simón. Quotients of $\mathrm{AdS}_{p+1}\times S^q$: causally well-behaved spaces and black holes hep-th/0402094

[1024] M. Banados and A. Gomberoff and C. Martinez. Anti-de Sitter space and black holes hep-th/9805087v1

[1025] O. Madden and S. F. Ross. Quotients of anti-de Sitter space hep-th/0401205

[1026] M. Bañados. Constant curvature black holes gr-qc/9703040

[1027] S. Åminneborg and I. Bengtsson and S. Holst and P. Peldán. Making anti-de Sitter black holes gr-qc/9604005

[1028] S. Holst and P. Peldán. Black holes and causal structure in anti-de Sitter isometric space-times gr-qc/9705067

[1029] P. Bieliavsky and M. Massar. Strict deformation quantization for actions of a class of symplectic Lie groups Noncommutative geometry and string theory (Yokohama, 2001) math/0011144

[1030] Pierre Bieliavsky and Stephane Detournay and Marianne Rooman and Philippe Spindel. BTZ black holes, WZW models and noncommutative geometry hep-th/0511080

[1031] Stephane Detournay. Deformations of anti-de Sitter black holes PhD thesis hep-th/0611031

[1032] Laurent Claessens. The horizon of the BTZ black hole 0912.2245

[1033] A. Franc. Structures de spin et opérateurs de Dirac sur les espaces riemanniens symmétriques compacts simplement connexes

[1034] Moshé Flato. Deformation view of physical theories

[1035] J. M. Gracia-Bondia and J. C. Várilly. Algebras of distributions suitable for phase-space quantum mechanics. I

[1036] P. Bieliavsky and S. Gutt and M. Bordemann and S. Waldmann. Traces for star products on the dual of a Lie algebra math.QA/0202126

[1037] Michael Reed and Barry Simon. Methods of modern mathematical physics. II. Fourier analysis, self-adjointness

[1038] Pierre Bieliavsky and Philippe Bonneau and Yoshiaki Maeda. Universal deformation formulae for three-dimensional solvable Lie groups math.QA/0308189

[1039] P. Bieliavsky and M. Massar. Oscillatory integral formulae for left-invariant star products on a class of Lie groups

[1040] P. Bieliavsky. Symplectic Symmetric Spaces math.DG/0703358

[1041] Maxima 5.10.0 URL http://maxima.sourceforge.net. A libre software for the manipulation of symbolic and numerical expressions.

[1042] André Unterberger. The calculus of pseudodifferential operators of Fuchs type

[1043] Ali H. Chamseddine and Alain Connes. The spectral action principle hep.th/9606001

[1044] Chris J. Isham. Topos methods in the fundations of physics 1004.3564v1

[1045] Ali H. Chamseddine and Alain Connes. Why the standard model 0706.3688

[1046] Ali H. Chamseddine and Alain Connes. A Universal Action Formula hep.th/9606056

[1047] Uwe Franz. The theory of quantum Lévy processes math/0407488v1